Foundations and Trends® in Communications and Information Theory

A Toolbox for Refined Information-Theoretic Analyses with Applications

Suggested Citation: Neri Merhav and Nir Weinberger (2025), "A Toolbox for Refined Information-Theoretic Analyses with Applications", Foundations and Trends® in Communications and Information Theory: Vol. 22, No. 1, pp 1–184. DOI: 10.1561/0100000142.

Neri Merhav
Technion – Israel Institute of Technology
merhav@ee.technion.ac.il

Nir Weinberger
Technion – Israel Institute of Technology
nirwein@technion.ac.il

now
the essence of knowledge
Boston — Delft

Contents

A Toolbox for Refined Information-Theoretic Analyses with Applications

Neri Merhav and Nir Weinberger

Technion – Israel Institute of Technology, Israel;
merhav@ee.technion.ac.il, nirwein@technion.ac.il

ABSTRACT

This monograph offers a toolbox of mathematical techniques that have been effective and widely applicable in information-theoretic analyses. The first tool is a generalization of the method of types to Gaussian settings, and then to general exponential families. The second tool is Laplace and saddle-point integration, which allow to refine the results of the method of types, and can obtain various precise asymptotic results. The third is the type class enumeration method, a principled method to evaluate the exact random-coding exponent of coded systems, which results in the best known exponent in various problems. The fourth is a subset of tools aimed at evaluating the expectation of non-linear functions of random variables, either via integral representations, by a refinement of Jensen's inequality via change-of-measure, by complementing Jensen's inequality with a reversed inequality, or by a class of generalized Jensen's inequalities that are applicable for functions beyond convex/concave. Various examples of all these tools are provided throughout the monograph.

Neri Merhav and Nir Weinberger (2025), "A Toolbox for Refined Information-Theoretic Analyses with Applications", Foundations and Trends® in Communications and Information Theory: Vol. 22, No. 1, pp 1–184. DOI: 10.1561/0100000142.

1

Introduction

This monograph is concerned with a set of analytical tools for information-theoretic analyses. The use of analytical methods to address challenging combinatorial problems is a classical method in mathematics, and includes various widely used techniques such as Stirling's approximation, Chernoff's bound, transform methods (with interchanging summation or integration order), among others. Analytical techniques also formed the basis of the inception of information-theory by Shannon [182]: On the face of it, and even at a deeper look, efficient coding for noisy channels is a formidable combinatorial problem, in a high dimensional space. Shannon addressed that challenge using analytical techniques:

1. The asymptotic equipartition property, and the estimation of volumes in high dimensional spaces, which allows to evaluate the size of high-probability sets. In the proof of the *noisy channel coding theorem* for discrete memoryless channels (DMCs), this allows to show that when an n-dimensional codeword is transmitted, the set of likely outputs has size roughly given by $e^{nH(Y|X)}$, where $H(Y|X)$ is the conditional entropy of the channel output Y conditioned on the input X, and the total set of likely outputs has roughly size of $e^{nH(Y)}$ (where $H(Y)$ is the entropy of Y).

2. The random-coding argument, which establishes the existence of optimal codes by evaluating the ensemble-average of randomly chosen code, and forms the basis for achievability (direct) results.

3. Convexity of information-measures, which is used to establish data-processing theorems, and consequently forms the basis for impossibility (converse) results.

Combining these ideas directly led, among other results, to the analytical formula for the capacity of DMCs, given by $C = \max_{P_X} I(X;Y)$ (where $I(X;Y) = H(Y) - H(Y|X)$ is the mutual information). Since Shannon's work, these ideas have been continuously extended and refined in numerous ways.

The goal of this monograph is to follow this path and propose a set of advanced analytical tools that are affirmed to be efficient and widely applicable for information-theoretic problems, allowing to obtain accurate and refined performance measure characterizations. Sections 2 and 3 to follow address the problem of estimating volumes in high dimensions, first, via a generalized method of types and, second, via the more advanced saddle-point method; Section 4 describes the *type class enumeration method* (TCEM), a tight analysis method of the performance of random-coding ensembles, and Section 5 considers various aspects of convexity and Jensen's inequality, mostly related to the computation of the expected values of non-linear functions. We next describe each one of these with more detail.

In Section 2, we describe a generalization of the method of types [38], [41], which was originally developed for finite alphabets, to Gaussian distributions, which are distributions over a continuous alphabet, and more generally, to distributions from exponential families. We introduce the notion of a typical set with respect to (WRT) a given parametric family of probability distributions. Such typical sets are defined in a way that the probability of each vector in the set is roughly the same for all possible distributions in the defined parametric family. This generalizes both the notion of weak typicality (a family consisting of a single distribution), and the notion of strong typicality for finite alphabets (the family is the set of all possible PMFs). Moreover, it allows to define, *e.g.*, typical sets for the Gaussian distribution. A key property of typical sets is their *volume*, because if an event of interest can be represented as the union over typical sets, then its probability can be accurately determined on the exponential scale using the volumes of these sets, and the probability of a single representative element from each of these sets. We thus develop a general method to evaluate the volumes of typical sets, and demonstrate its use on memoryless Gaussian sources, on Gaussian sources conditioned on other vectors, and on Gaussian sources with memory. We then generalize this method to distributions from an exponential family.

While the method of types is a general and widely applicable approach that leads to useful exponential bounds, there are settings which require more delicate analysis, and thus, more advanced tools. In Section 3, we begin by describing the Laplace method of integration, and exemplify its use in the problems of universal coding and extreme-value statistics. We then discuss the closely-related saddle-point method of integration in the complex plane, and show how it allows to accurately evaluate the size of type classes, volumes of hyper-spheres, and large-deviations probabilities, not only in the exact exponential rate, but also with the exact pre-exponential factor. We show that this method is applicable beyond parametric models. We further demonstrate its use for the evaluation of the number of lattice points in an L_1 ball, and the evaluation of the volume of an intersection of a hyper-sphere and hyperplane, refining the analysis of Section 2.

In Section 4, we consider coded settings and ensembles of random codes. We introduce the TCEM, which is a principled method for deriving the error exponent of random codes. We first describe the standard techniques commonly used to derive bounds on the error exponent, such as Jensen's inequality and its implications, and various types of union bounds. While these methods indeed turned out to be effective in the error-exponent analysis of basic settings, such as point-to-point channels and standard decoding rules, there is no general guarantee that they are accurate in more advanced scenarios. Indeed, we survey various settings in which these methods are sub-optimal, and do not provide the exact random-coding error exponent. As an alternative, we show that ensemble-average error probabilities (and other related performance measures) may be expressed via *type class* *enumerators* (TCEs), and specifically, via their (non-integer) moments and tail probabilities. We demonstrate this both on basic settings as well as more involved ones. We explore the probabilistic and statistical properties of TCEs, and then discuss a number of settings in multi-user information theory, in distributed compression and in hypothesis testing, for generalized decoding rules such as those allowing erasures and list outputs, and for the analysis of the typical random code. We outline how the TCEM is used in each of these settings, and how it allows to obtain, among other things, exact error-exponents for optimal decoding rules. In Appendix B we show that the exponents obtained by the TCEM can also be computed effectively.

In Section 5, we address the problem of evaluating the expectation of a non-linear function $f(\cdot)$ of a random variable (RV) X. In many cases, this function is either convex or concave, and so a natural course of action is to bound it using Jensen's inequality. However, there is no guarantee that the resulting bound is tight enough for the intended application. We present two general and useful strategies that can be employed in such cases. The first one is based on finding an *integral representation* of the function. Then, we interchange the expectation and integral order, and obtain an alternative expression for $\mathbb{E}\{f(X)\}$. The technique is useful if computing the inner expectation is simpler than the original expectation, or if it can be evaluated more accurately. After evaluating the inner expectation, the expectation $\mathbb{E}\{f(X)\}$ of interest can be computed by solving a one-dimensional integral. For example, when $f(t) = \ln(t)$, this allows to replace the evaluation of the expected logarithm with the evaluation of its moment-generating function (MGF). This is especially appealing since if $X = \sum_{i=1}^{n} X_i$ is the sum of n independent and identically distributed (IID) RVs, then its MGF is the n-th power of the MGF of just one of them. In accordance, this transforms the original expectation, which is an integral in $\mathbb{R}^n$, to a one-dimensional integral. We focus on the logarithmic function $f(t) = \ln(t)$ (and its integer powers), as well as the power function $f(t) = t^\rho$ for some $\rho > 0$ (even non-integer), and exemplify the use of this technique in a multitude of problems such as differential entropy for generalized multivariate Cauchy densities, ergodic capacity of the Rayleigh single-input multiple-output (SIMO) channel, and moments of guesswork.

The second strategy exploits convexity or concavity properties, but goes beyond the standard Jensen's inequality. This strategy may come in various flavors. First, a change of measure can be performed before using Jensen's inequality, and then the alternative measure can be optimized over a given class to improve the bound. As a notable example, when $f(t) = \ln(t)$, this reproduces the Donsker–Varadhan variational characterization of the Kullback–Leibler (KL) divergence. Second, one may use Jensen's inequality, but accompany it with an inequality in the opposite direction, *i.e.*, a reverse Jensen's inequality (RJI), in order to evaluate its tightness. We provide a few techniques, all of which rely on a general form of such a RJI. Third, the "supporting-line" approach used to prove Jensen's inequality may be generalized to cases in which the function whose expected value is sought of is not convex/concave, but takes a more complicated form, such as the composition or a multiplication of a different function with a convex/concave function. A generalized version of Jensen's inequality can still be derived, by properly optimizing the supporting line. We exemplify the use of this technique in various problems involving the evaluation of data compression performance and channel capacity.

In summary, we present a diverse toolbox of analytical techniques, indispensable to every information-theorist aiming to obtain tight and accurate results. We mention in passing other analytical techniques widely used in information theory, such as central-limit theorems extensively used in non-vanishing error regimes [198], concentration of measure bounds [169], statistical-physics methods such as the cavity and the replica method [151], and various methods described in the recent book [56]. These complement the tools outlined in this monograph.

This monograph was invited and written following a plenary talk by the first author, at the 2023 IEEE International Symposium on Information Theory (ISIT 2023), Taipei, Taiwan, June 25-30, 2023. It should be pointed out that some of the proposed techniques (like in Sections 2, 4, and many parts of Section 5) are original, while others are not new (like in Section 3).

2

Extension of the Method of Types to Continuous Alphabets

2.1 Introduction

In their renowned 1981 book (with its second edition [41]), Csiszár and Körner introduced the groundbreaking concept of the *method of types*. This method has since emerged as a cornerstone within classical Shannon theory, offering a remarkably powerful and versatile mathematical analytical tool-set. Its primary application is in providing a set of tools that is predominantly used to establish achievability theorems, and occasionally also converse results. Additionally, this method's utility extends to the evaluation of error probability exponential decay rates (referred to as *error exponents*) and the exponential growth rates of subsets of sequences as functions of the blocklength (or the dimension).

The method of types is a fundamental combinatorial approach, originally crafted for memoryless sources and channels with finite alphabets. In essence, for a given finite alphabet $\mathcal{X}$, this method involves partitioning the space of all $|\mathcal{X}|^n$ sequences of length n from $\mathcal{X}$ into distinct equivalence classes termed *type classes*. Each type class encompasses sequences that share the same empirical distribution, characterized by a specific array of relative frequencies pertaining to the $|\mathcal{X}|$ alphabet letters. An alternative perspective on type classes is that within each such class, any sequence can be derived through permutations of other sequences. The strength of the method of types emanates from a concept of elegant simplicity: Despite the exponential growth of the size of each type class with n (its exponential rate being determined by the entropy of the corresponding empirical distribution), the number of distinct type classes experiences only a polynomial growth with n. This interplay of growth dynamics yields a crucial outcome: The likelihood of any event expressed as a union of type classes is dominated by the exponential behavior driven by the most probable type class contained within the event. Similarly, when dealing with the size of a set defined as a union of type classes, this size experiences an exponential dominance dictated by the largest type class within that set.

Csiszár provides a comprehensive exploration of the method of types in [38], encompassing foundational principles, as well as numerous applications. These applications span a wide spectrum, including the derivation of error exponents for source coding, channel coding, source-channel coding, hypothesis testing, the type covering lemma, the packing lemma, the capacity evaluation for arbitrarily varying channels, rate-distortion coding, as well as multi-terminal source and channel coding theorems. In addition, Csiszár also undertakes in [38] a meticulous survey of several notable extensions to the method of types. Foremost among these are second-order and higher-order types, with recognition attributed to prior work by Billingsley [18], Boza [22], Whittle [221], Davisson, Longo, and Sgarro [47], as well as Natarajan [155]. Furthermore, the exploration extends to finite-state types, proposed by Weinberger, Merhav, and Feder [214]. Csiszár's comprehensive survey [38] ends with a section addressing continuous alphabets. This section's outset acknowledges that extensions of the type concept to continuous alphabets remain largely uncharted. It proceeds to navigate this challenge by adopting a discretization strategy through fine quantization (see Tridenski and Somekh-Baruch [204] for a recent application of this approach). Nonetheless, this approach reveals vulnerabilities when grappling with probability density functions (PDFs) that are supported by the entirety of the real line or half of it. In such cases, achieving arbitrarily high resolution quantization, a requisite of the traditional method of types, becomes unattainable. While acknowledging that coarsely quantizing the tails of distributions generally entails minimal impact due to their low probability, certain technical intricacies arise, particularly concerning the uniformity of convergence across a class of distributions. This concern becomes particularly salient when confronted with the need to interchange limit operations, such as the limits as n grows large and the quantization resolution increases concurrently. Furthermore, the cost associated with achieving high-resolution quantization manifests as an escalated computational workload in the calculation of the desired exponential rate. This is due to the fact that the number of free parameters to optimize is equal to the number of quantization levels minus one.

Within this section, our central proposition emerges: The extension of type classes and the critical components of the method of types to continuous alphabets is not only viable but also remarkably intuitive. This assertion is particularly pertinent when considering PDFs originating from the broader exponential family [115], [146], and especially when dealing with the Gaussian PDF, as expounded in, *e.g.*, [9], [88], [90], [116], [124], [138], [196]. Notably, our approach circumvents the need for the discrete approximations proposed in [38].

Our methodology revolves around the partitioning of a suitably chosen high probability set in the space of n-sequences into equivalence classes, referred to as type classes, in analogy to their finite-alphabet counterparts. This construction retains two pivotal attributes, analogous to their roles in the customary finite-alphabet context:

1. It is possible to devise a computable expression that characterizes the exponential growth rate of the size or volume of each type class as a function of n. This expression, which is always a certain form of entropy or differential entropy, remains amenable to calculation independently of n and aligns with the concept of *single-letter expression* in the jargon of information theorists.

2. The number of distinct type classes relevant to the problem at hand exhibits sub-exponential growth WRT n. This assures that the quantity of distinct types relevant to our problem expands in a manner manageable for analysis.

By "computable expression" in the first point, we refer to an expression whose computational complexity remains fixed as n varies. In relation to the second point, when we mention types "relevant to the problem at hand," we imply scenarios where the aggregate number of distinct type classes might conceivably be boundless, yet the vast majority beyond a sub-exponential subset hold minimal importance and can be disregarded, given their inconsequential collective impact on the quantity of interest. This might be due to their associated probabilities being negligibly small.

In the upcoming sections, we embark on a concise exploration of the fundamental concepts that underlie the extension of the method of types to encompass continuous alphabets. Our journey begins with the Gaussian scenario before encompassing the broader domain of exponential families. Throughout these discussions, we will interweave illustrative examples to provide practical context for the concepts being elucidated.

2.2 Various Definitions of Type Classes

2.2.1 Type Classes and the Method of Types

As mentioned earlier, in the memoryless, finite-alphabet case, we define a type class as the set of all sequences that share the same empirical distribution. More precisely, given a sequence $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, with $x_i \in \mathcal{X}$, $i = 1, 2, \ldots, n$, $\mathcal{X}$ being a finite alphabet of finite size $|\mathcal{X}|$, the empirical distribution, $\hat{P}_{\boldsymbol{x}}$, associated with $\boldsymbol{x}$ is the vector $\{\hat{P}_{\boldsymbol{x}}(x),\ x \in \mathcal{X}\}$, where $\hat{P}_{\boldsymbol{x}}(x) = n_{\boldsymbol{x}}(x)/n$, and $n_{\boldsymbol{x}}(x)$ being the number of occurrences of the letter $x \in \mathcal{X}$ in $\boldsymbol{x}$. Thus, the *type* of $\boldsymbol{x} \in \mathcal{X}^n$ is defined by

$$\mathcal{T}_n(\boldsymbol{x}) \triangleq \left\{\boldsymbol{x}' \in \mathcal{X}^n \colon \hat{P}_{\boldsymbol{x}'} = \hat{P}_{\boldsymbol{x}}\right\}. \tag{2.1}$$

An alternative, equivalent, definition of $\mathcal{T}_n(\boldsymbol{x})$ is the set of all $\boldsymbol{x}' \in \mathcal{X}^n$ that can be obtained as permutations of $\boldsymbol{x}$. Since $\mathcal{T}_n(\boldsymbol{x})$ corresponds to a particular empirical probability distribution, say, $\hat{P}$, it would be sometimes convenient to denote it by $\mathcal{T}_n(\hat{P})$. Similar notation applies to type classes of pairs of n-vectors, $(\boldsymbol{x}, \boldsymbol{y})$ (and triples, and so on), where in the alternative notation, $\hat{P}$ is understood to be the joint empirical distribution.

The definition (2.1) lies at the heart of the *method of types* [38], [41, Chapter 2], [36, Section 11.1], which we next succinctly describe. To this end, let $\doteq$ denote equality on the exponential scale, *i.e.*, two positive sequences $\{a_n\}$ and $\{b_n\}$ satisfy that $a_n \doteq b_n$ if $\lim_{n\to\infty} \frac{1}{n} \ln \frac{a_n}{b_n} = 0$. The method of types can be summarized by the following properties:

1. The number of possible types: For sequences of length n, a type is equivalent to an empirical distribution of the form $(\frac{n_1}{n}, \ldots, \frac{n_{|\mathcal{X}|}}{n})$, where $n_j \in \{0, 1, \ldots, n\}$ and $\sum_{j=1}^{|\mathcal{X}|} n_j = n$. The "stars-and-bars" method shows that the number of possible types is given by (see [41, Exercise 2.1])
$$\binom{n + |\mathcal{X}| - 1}{|\mathcal{X}| - 1} \leq (n+1)^{|\mathcal{X}|}. \tag{2.2}$$
This shows that the number of types increases polynomially with n, as long as the alphabet is finite, *i.e.*, $|\mathcal{X}| < \infty$.

2. The probability of a sequence: If $\boldsymbol{x} \in \mathcal{T}_n(Q)$ for some Q then
$$P(\boldsymbol{x}) = \prod_{i=1}^{n} P(x_i) \tag{2.3}$$
$$= \prod_{x\in\mathcal{X}} [P(x)]^{nQ(x)} \tag{2.4}$$
$$= \exp\left[n \sum_{x\in\mathcal{X}} Q(x) \ln P(x)\right] \tag{2.5}$$
$$= \exp\left[n \sum_{x\in\mathcal{X}} Q(x) \ln Q(x) + n \sum_{x\in\mathcal{X}} Q(x) \ln \frac{P(x)}{Q(x)}\right] \tag{2.6}$$
$$= \exp\left[-n(H(Q) + D(Q||P))\right], \tag{2.7}$$
where
$$H(Q) \triangleq \sum_{x\in\mathcal{X}} Q(x) \ln \frac{1}{Q(x)} \tag{2.8}$$
is the entropy, and
$$D(Q||P) \triangleq \sum_{x\in\mathcal{X}} Q(x) \ln \frac{Q(x)}{P(x)} \tag{2.9}$$
is the KL divergence.[1]

[1] Assuming the conventions, based on continuity arguments, $0 \ln \frac{0}{0} = 0$, $0 \ln \frac{0}{p} = 0$ and $q \ln \frac{q}{0} = \infty$ for $p, q \in (0, 1)$ [36, Section 2.3].

3. The size of a type class: It can be shown [41, Lemma 2.3] that the size of a type class is

$$\frac{e^{nH(Q)}}{\binom{n+|\mathcal{X}|-1}{|\mathcal{X}|-1}} \leq |\mathcal{T}_n(Q)| \leq e^{nH(Q)}. \tag{2.10}$$

Thus, as long as $|\mathcal{X}| < \infty$, it holds that

$$|\mathcal{T}_n(Q)| \doteq e^{nH(Q)}. \tag{2.11}$$

Establishing this property is the main technical challenge in the derivation of the method of types.

4. The probability of observing a type: The probability of observing a n-length sequence of a type Q from a memoryless source with distribution P satisfies

$$\Pr\left[\boldsymbol{X} \in \mathcal{T}_n(Q)\right] \doteq e^{-nD(Q||P)}. \tag{2.12}$$

This is a direct consequence of (2.7) and (2.11).

The method of types can be easily extended to both joint and conditional type classes. For joint types over finite alphabets $\mathcal{X}$ and $\mathcal{Y}$, we consider a pair of a sequences $(\boldsymbol{x}, \boldsymbol{y}) \in (\mathcal{X} \times \mathcal{Y})^n$ over $\mathcal{X} \times \mathcal{Y}$, with $x_i \in \mathcal{X}$ and $y_i \in \mathcal{Y}$, $i = 1, 2, \ldots, n$. The empirical distribution, $\hat{P}_{\boldsymbol{x},\boldsymbol{y}}$, associated with $(\boldsymbol{x}, \boldsymbol{y})$ is the vector $\{\hat{P}_{\boldsymbol{x},\boldsymbol{y}}(x),\ x \in \mathcal{X}, y \in \mathcal{Y}\}$, where $\hat{P}_{\boldsymbol{x},\boldsymbol{y}}(x) = n_{\boldsymbol{x},\boldsymbol{y}}(x, y)/n$, $n_{\boldsymbol{x},\boldsymbol{y}}(x, y)$ being the number of occurrences of the pair of letters $(x, y) \in \mathcal{X} \times \mathcal{Y}$ in $(\boldsymbol{x}, \boldsymbol{y})$, and the type class of $(\boldsymbol{x}, \boldsymbol{y}) \in (\mathcal{X} \times \mathcal{Y})^n$ is

$$\mathcal{T}_n(\boldsymbol{x}, \boldsymbol{y}) \triangleq \left\{(\boldsymbol{x}', \boldsymbol{y}') \in (\mathcal{X} \times \mathcal{Y})^n \colon \hat{P}_{\boldsymbol{x}',\boldsymbol{y}'} = \hat{P}_{\boldsymbol{x},\boldsymbol{y}}\right\}. \tag{2.13}$$

The properties of the method of types for this case are obtained by replacing $|\mathcal{X}|$ with $|\mathcal{X}||\mathcal{Y}|$ and understanding types Q as joint types, Q_{XY}. Thus, $|\mathcal{T}_n(Q)| \doteq e^{nH(Q_{XY})}$ where $H(Q_{XY})$ is now a joint entropy. A pair of sequences also leads to a *conditional type class*, where the conditional type class of $\boldsymbol{y} \in \mathcal{Y}^n$ given a *fixed* $\boldsymbol{x} \in \mathcal{X}^n$ is defined as

$$\mathcal{T}_n(\boldsymbol{y}|\boldsymbol{x}) \triangleq \left\{\boldsymbol{y}' \in \mathcal{Y}^n \colon \hat{P}_{\boldsymbol{x},\boldsymbol{y}'} = \hat{P}_{\boldsymbol{x},\boldsymbol{y}}\right\}. \tag{2.14}$$

Note that while a joint type class is a subset of $(\mathcal{X} \times \mathcal{Y})^n$, a conditional type is a subset of $\mathcal{Y}^n$. The size of a conditional type class is exponentially on the same order as $e^{nH(Q_{Y|X}|Q_X)}$, that is, determined by the conditional entropy $H(\tilde{Y}|\tilde{X})$ for $(\tilde{X},\tilde{Y}) \sim Q_{XY}$. These definitions and properties can be directly generalized to multiple length-n sequences, as long the number of sequences and all the alphabets are finite and fixed, independent of n.

With a careful handling of a few technical aspects, the method of types is also extended to Markov chains over finite alphabets. Considering, for simplicity, a first-order Markov chain, the empirical Markov distribution is essentially $\{\hat{P}_{\boldsymbol{x}}(x,x'),\ (x,x') \in \mathcal{X}^2\}$, where $\hat{P}_{\boldsymbol{x}}(x,x') = n_{\boldsymbol{x}}(x,x')/n$, and $n_{\boldsymbol{x}}(x,x')$ is the number of occurrences of the letter $(x,x') \in \mathcal{X}^2$ as *consecutive* symbols in $\boldsymbol{x}$. Thus, ignoring edge effects, this is merely a joint type of $\overline{\boldsymbol{x}} = (x_1,\ldots,x_{n-1})$ and $\overline{\boldsymbol{x}}' = (x_2,x_3,\ldots,x_n)$, the latter being a shifted version of the former. The technical details are related to the fact that the set of joint types is not arbitrary. Specifically, upon adopting the (convenient) cyclic convention that x_1 follows x_n, the number of incoming transitions to each state must be equal to the number of outgoing transitions from that state, thus, the joint empirical distribution $\hat{P}_{\boldsymbol{x}}(x,x')$ must have identical marginals. Various variants thus have been developed [18], [22], [47], [155], [221] (see [38, Section VII.A] for a survey), but overall, the properties of the method of types remain essentially the same: The number of possible types is polynomial, and the other properties follow by replacing entropy $H(\tilde{X})$ with $\tilde{X} \sim Q$ with first-order conditional entropy $H(\tilde{X}_2|\tilde{X}_1)$, where $(\tilde{X}_1,\tilde{X}_2) \sim Q_{X_1X_2}$, and the KL divergence by the conditional KL divergence

$$D(Q_{X_2|X_1}||P_{X_2|X_1}|Q_X) \triangleq \sum_{x_1\in\mathcal{X},x_2\in\mathcal{X}} Q_X(x_1)Q_{X_2|X_1}(x_2|x_1)\ln\frac{Q_{X_2|X_1}(x_2|x_1)}{P_{X_2|X_1}(x_2|x_1)}. \quad (2.15)$$

Our goal in the rest of this section would be to generalize this concept beyond finite alphabets, including both memoryless sources and sources with memory. The main challenges are how to handle the fact that infinite alphabets potentially yield infinitely many types, and how to asses the size of these generalized type classes. We will show that the number of types can be essentially bounded by a finite number, which grows sub-exponentially with n, and that the sizes (or volumes) of various interesting definitions of generalized type classes scale as e^{nH} for a suitable entropy term.

2.2.2 Types for General Alphabets

Our first step is to develop generalizations of the definition of the basic type class. Clearly, the two previously provided definitions hold specifically for finite-alphabet memoryless systems. However, when considering the more general scenario, a broader definition becomes necessary. The essential requirement for formulating a comprehensive method of types is that sequences falling within the same type class exhibit matching probabilities, at least in the exponential scale. In cases where the data may be governed by a single probability distribution (or PDF, in continuous scenarios) denoted as P, the definition is as follows:

$$\mathcal{T}_n(P) \triangleq \left\{ \boldsymbol{x} \in \mathcal{X}^n \colon -\frac{\ln P(\boldsymbol{x})}{n} = H \right\}. \tag{2.16}$$

Here, H represents a constant, which for a discrete alphabet, typically signifies the entropy rate of distribution P, and for a continuous alphabet, signifies the differential entropy rate. This definition underscores the fundamental property that all sequences within a given type class share a consistent probabilistic behavior. It encapsulates the notion that their probabilities, when viewed through the lens of logarithmic scaling, converge to a common value, thereby enabling a more inclusive approach in diverse scenarios. In certain instances, intricate technical nuances necessitate the incorporation of a certain tolerance factor, denoted as $\epsilon > 0$. This becomes particularly pertinent in continuous scenarios, as we will soon delve into.[2] This leads us to introduce the notion of an *ϵ-inflated type class*, represented as follows:

$$\mathcal{T}_{n,\epsilon}(P) \triangleq \left\{ \boldsymbol{x} \in \mathcal{X}^n \colon \left| -\frac{\ln P(\boldsymbol{x})}{n} - H \right| \leq \epsilon \right\}. \tag{2.17}$$

Equations (2.16) and (2.17) define the notion of *weak typicality* [36, Section 3.3]. However, there are instances where we require this property of almost equal log-probabilities (or log-densities) not solely for one

[2] In the next section, when we explore the saddle-point method, we will see how to circumvent the need for this tolerance factor.

specific source P, but concurrently for all sources within a given class. The most prominent case is that of the class of all memoryless sources over a finite alphabet $\mathcal{X}$. If we let $\mathcal{P}(\mathcal{X})$ be the set of all PMFs over $\mathcal{X}$, then we define the type class of $\boldsymbol{x}$ WRT the class $\mathcal{P}(\mathcal{X})$ as

$$\mathcal{T}_n(\boldsymbol{x}) \triangleq \{\boldsymbol{x}' \in \mathcal{X}^n \colon P(\boldsymbol{x}') = P(\boldsymbol{x}),\ \forall P \in \mathcal{P}(\mathcal{X})\}\,. \tag{2.18}$$

The definition in (2.18) aligns with the concept of a type class in (2.1), as if $\boldsymbol{x} \in \mathcal{T}_n(\boldsymbol{x})$ then any permutation of its letters also belongs to $\mathcal{T}_n(\boldsymbol{x})$.

Now, consider a parametric family of sources, $\{P_\theta \colon \theta \in \Theta\}$, where θ is the parameter and Θ is the parameter space. We define the type class of $\boldsymbol{x}$ WRT the class $\{P_\theta \colon \theta \in \Theta\}$ (see also [150]) as

$$\mathcal{T}_n(\boldsymbol{x}) \triangleq \{\boldsymbol{x}' \in \mathcal{X}^n \colon P_\theta(\boldsymbol{x}') = P_\theta(\boldsymbol{x}),\ \forall \theta \in \Theta\}\,. \tag{2.19}$$

Indeed, when the set $\{P_\theta \colon \theta \in \Theta\}$ encompasses all memoryless sources with a given finite alphabet $\mathcal{X}$ of size $|\mathcal{X}|$, the parameter vector θ can be construed as the vector comprising $|\mathcal{X}| - 1$ letter probabilities, with the $|\mathcal{X}|$-th probability completing their sum to unity. This alignment of definitions corresponds precisely to the definition in (2.18). The rationale underlying this correspondence stems from the fact that the probability of a sequence $\boldsymbol{x}$ under any memoryless source depends on $\boldsymbol{x}$ solely via the empirical distribution $\hat{P}$ associated with $\boldsymbol{x}$. As a result, any two sequences sharing the same empirical distribution must invariably possess identical probabilities across all memoryless sources indexed by distinct θ values. In essence, this expansive definition of a type class seamlessly envelops the well-established definition applicable to memoryless sources, effectively encompassing it as a specific case. More generally, the ϵ-inflated type class of $\boldsymbol{x}$ is defined as

$$\mathcal{T}_{n,\epsilon}(\boldsymbol{x}) \triangleq \left\{\boldsymbol{x}' \in \mathcal{X}^n \colon \left|\frac{\ln P_\theta(\boldsymbol{x}')}{n} - \frac{\ln P_\theta(\boldsymbol{x})}{n}\right| \le \epsilon,\ \forall \theta \in \Theta\right\} \tag{2.20}$$

$$= \bigcap_{\theta\in\Theta} \left\{\boldsymbol{x}' \in \mathcal{X}^n \colon \left|\frac{\ln P_\theta(\boldsymbol{x}')}{n} - \frac{\ln P_\theta(\boldsymbol{x})}{n}\right| \le \epsilon\right\}. \tag{2.21}$$

The definitions in (2.18)-(2.21) correspond to the notion of *strong typicality* [41, Chapter 2]. It is evident that broadening the scope of reference sources, achieved by expanding the parametric family, causes the type classes to contract. Indeed, then the requirement $\frac{1}{n}|\ln P_\theta(\boldsymbol{x}') - \ln P_\theta(\boldsymbol{x})| \le \epsilon$ is imposed over a larger classes of sources, the intersection in (2.21) is taken over more sets, and hence this intersection becomes smaller. Conversely, focusing on a subset of $\{P_\theta\colon \theta \in \Theta\}$ results in the expansion of type classes. At the far end of this spectrum, when the subclass of sources becomes a singleton, we are back to weak typicality.

As a pertinent example that illustrates this, consider the class of memoryless, zero-mean Gaussian sources parameterized by the variance, denoted as $\theta = \sigma^2$. The PDF for this class is expressed as follows:

$$P_{\sigma^2}(\boldsymbol{x}) = \frac{\exp\left\{-\sum_{i=1}^n x_i^2/(2\sigma^2)\right\}}{(2\pi\sigma^2)^{n/2}}. \tag{2.22}$$

Since $P_{\sigma^2}(\boldsymbol{x})$ depends on $\boldsymbol{x}$ only via $\sum_{i=1}^n x_i^2$, it is clear that all sequences $\{\boldsymbol{x}\}$ with a given norm (*i.e.*, all sequences pertaining to points on the surface of a given Euclidean hyper-sphere centered at the origin) have the same PDF, $P_{\sigma^2}(\boldsymbol{x})$. Thus, a natural definition of type classes WRT the class of zero-mean, memoryless Gaussian sources parameterized by their variance, also known as *Gaussian types* (or "power types"), are surfaces of n-dimensional hyper-spheres centered at the origin. Expanding this parametric class by introducing a mean parameter μ leads us to consider $\theta = (\sigma^2, \mu)$ ($\sigma^2 > 0$, $\mu \in \mathbb{R}$). Consequently, the PDF becomes:

$$P_{\sigma^2,\mu}(\boldsymbol{x}) = \frac{\exp\left\{-\sum_{i=1}^n (x_i - \mu)^2/(2\sigma^2)\right\}}{(2\pi\sigma^2)^{n/2}}. \tag{2.23}$$

In this context, $P_{\sigma^2,\mu}(\boldsymbol{x})$ depends on $\boldsymbol{x}$ exclusively through $\sum_{i=1}^n x_i^2$ and $\sum_{i=1}^n x_i$. Accordingly, the definition of a type class involves the intersection of a hyper-sphere surface defined by a particular radius and a hyper-plane defined by a specific value of $\sum_{i=1}^n x_i$. This type class is notably smaller compared to the type class relative to $\{P_{\sigma^2}(\boldsymbol{x}),\ \sigma^2 > 0\}$, which was solely defined by the hyper-sphere surface without any additional intersection with a hyper-plane.

More generally, consider a parametric class of memoryless sources that form an *exponential family* (see, *e.g.*, Lehmann [110, Section 1.4]). This means that the single-letter marginal is of the form,

$$P_\theta(x) = \frac{\exp\left\{\sum_{j=1}^k \theta_j \phi_j(x)\right\}}{Z(\theta)}, \tag{2.24}$$

where $\theta = (\theta_1, \ldots, \theta_k)$ is the parameter vector, $\phi_i : \mathcal{X} \to \mathbb{R}$ are given functions, and $Z(\theta)$ is a normalization constant, given by

$$Z(\theta) \triangleq \sum_{x \in \mathcal{X}} \exp\left\{\sum_{j=1}^k \theta_j \phi_j(x)\right\}, \tag{2.25}$$

in the discrete case, or

$$Z(\theta) \triangleq \int_{\mathcal{X}} \exp\left\{\sum_{j=1}^k \theta_j \phi_j(x)\right\} \mathrm{d}x, \tag{2.26}$$

in the continuous case. Moving on to n-sequences, by considering the product form,

$$P_\theta(\boldsymbol{x}) = \prod_{i=1}^n P_\theta(x_i) = \frac{\exp\left\{\sum_{j=1}^k \theta_j \sum_{i=1}^n \phi_j(x_i)\right\}}{[Z(\theta)]^n}, \tag{2.27}$$

type classes are defined by a given combination of values of the statistics, $\sum_{i=1}^n \phi_j(x_i)$, for $j = 1, \ldots, k$. The class of memoryless Gaussian sources parameterized by σ^2 only, is an exponential family with $k = 1$, a transformed parameter, $\theta = \theta_1 = -\frac{1}{2\sigma^2}$, $\phi_1(x) = x^2$ and accordingly, $Z(\theta) = \sqrt{2\pi\sigma^2} = \sqrt{-\pi/\theta}$. The broader class, parameterized by (σ^2, μ), is also an exponential family with $k = 2$, $\theta_1 = -\frac{1}{2\sigma^2}$, $\theta_2 = \frac{\mu}{\sigma^2}$, $\phi_1(x) = x^2$, $\phi_2(x) = x$, and

$$Z(\theta) = \sqrt{2\pi\sigma^2} \exp\left\{\frac{\mu^2}{2\sigma^2}\right\} = \sqrt{-\frac{\pi}{\theta_1}} \exp\left\{-\frac{\theta_2^2}{4\theta_1}\right\}. \tag{2.28}$$

The class of memoryless sources from the alphabet $\mathcal{X} = \{1, 2, \ldots, |\mathcal{X}|\}$ is yet another example of an exponential family with $k = |\mathcal{X}| - 1$, a parameter transformation,

$$\theta_j = \ln\left(\frac{p_j}{1 - \sum_{l=1}^{|\mathcal{X}|-1} p_l}\right), \tag{2.29}$$

for $j = 1, 2, \ldots, |\mathcal{X}| - 1$,

$$\phi_j(x) = \begin{cases} 1, & x = j \\ 0, & x \neq j \end{cases} \tag{2.30}$$

and

$$Z(\theta) = \frac{1}{1 - \sum_{j=1}^{|\mathcal{X}|-1} p_j}. \tag{2.31}$$

In summary, we observe that exponential families are general enough to include at least two important special cases of memoryless sources: Finite-alphabet memoryless sources and Gaussian memoryless sources, but of course they include many more [110, Section 1.4].

The method of types for exponential families is useful for assessing the exponential order of certain sums or integrals (depending on whether the alphabet is discrete or continuous) of functions that depend on $\boldsymbol{x}$ only via the set of statistics $\{\phi_j, j = 1, 2, \ldots, k\}$, *i.e.*,

$$\sum_{\boldsymbol{x}} f\left(\sum_{i=1}^{n} \phi_1(x_i), \sum_{i=1}^{n} \phi_2(x_i), \ldots, \sum_{i=1}^{n} \phi_k(x_i)\right) \tag{2.32}$$

in the discrete alphabet case, or

$$\int_{\mathbb{R}^n} f\left(\sum_{i=1}^{n} \phi_1(x_i), \sum_{i=1}^{n} \phi_2(x_i), \ldots, \sum_{i=1}^{n} \phi_k(x_i)\right) \mathrm{d}\boldsymbol{x} \tag{2.33}$$

in the continuous alphabet case. Most notably, the method is useful when f is an exponential function of $\{\phi_j,\ j = 1, 2, \ldots, k\}$, for example, an exponential function of a linear combination of $\{\phi_j\}$, possibly multiplied by an indicator function for the event that the vector $\{\phi_j,\ j = 1, \ldots, k\}$ lies in a certain region in $\mathbb{R}^k$.

2.2.3 Markov and Conditional Type Classes

The exponential family also lends itself to handle sources with certain structures of memory, most notably, Markov sources, where

$$P_\theta(\boldsymbol{x}) = \frac{\exp\left\{\sum_{j=1}^{k} \theta_j \sum_{i=0}^{n-1} \phi_j(x_i, x_{i+1})\right\}}{Z_n(\theta)}, \tag{2.34}$$

and so, type classes are defined according to a given combination of values of the statistics $\sum_{i=0}^{n-1} \phi_j(x_i, x_{i+1})$, as an extension of finite-alphabet Markov types [18], [22], [47], [155], [221]. In addition, a parallel extension of conditional type classes to the continuous alphabet case can also be defined, either WRT an exponential family of conditional distributions (in the discrete case) or conditional PDFs (in the continuous case). In the memoryless case, we define the single-letter conditional probability function pertaining to an exponential family, as

$$P_\theta(x|y) = \frac{\exp\left\{\sum_{j=1}^k \theta_j \phi_j(x,y)\right\}}{Z(y,\theta)}, \tag{2.35}$$

with $Z(y,\theta)$ being a normalization constant such that $P_\theta(x|y)$ sums or integrates (over x) to unity. Here, the conditional type class of $\boldsymbol{x}$ given $\boldsymbol{y}$ is the set of all $\{\boldsymbol{x}'\}$ such that for the given $\boldsymbol{y}$, $\sum_{i=1}^n \phi_j(x_i', y_i) = \sum_{i=1}^n \phi_j(x_i, y_i)$, for all $j = 1, 2, \ldots, k$. For example, a Gaussian conditional type class is defined WRT the class

$$P_{\sigma^2,a}(x|y) = \frac{\exp\{-(x-ay)^2/(2\sigma^2)\}}{\sqrt{2\pi\sigma^2}}, \tag{2.36}$$

which is a conditional exponential family with $k = 2$, $\theta_1 = -\frac{1}{2\sigma^2}$, $\theta_2 = \frac{a}{\sigma^2}$, $\phi_1(x,y) = x^2$, $\phi_2(x,y) = xy$, and

$$Z(y,\theta) = \sqrt{2\pi\sigma^2} \exp\left\{\frac{a^2y^2}{2\sigma^2}\right\}. \tag{2.37}$$

In this case, the conditional type class is defined by prescribed values of $\sum_{i=1}^n x_i^2$ and $\sum_{i=1}^n x_i y_i$.

In the sequel, we will demonstrate the usefulness of the concepts of type classes and conditional type classes in several applications.

2.3 Simple Gaussian Types

As delineated in Section 2.2, in the finite alphabet case, the conventional method of types hinges on an explicit formulation for the exponential growth rate of the size of a given type class (as a function of the sequence length n). Similarly, when dealing with the continuous Gaussian scenario, a prerequisite is obtaining a specific, well-defined, expression for the volume of the associated type class. Let us consider the simplest scenario–that of typicality WRT zero-mean, Gaussian, IID sources, characterized by their variance. In this context, the corresponding type class, as detailed in Section 2.2, is defined by a hyper-spherical surface, denoted, with a slight abuse of notation, as

$$\mathcal{T}_n(s) \triangleq \left\{ \boldsymbol{x} \in \mathbb{R}^n \colon \frac{1}{n} \sum_{i=1}^{n} x_i^2 = s \right\}. \tag{2.38}$$

Strictly speaking, the volume of $\mathcal{T}_n(s)$ is zero, if viewed as an object in the space $\mathbb{R}^n$, because its real dimension is $n-1$, as it is the surface area of a hyper-sphere of radius $\sqrt{ns}$. The surface area of an n-dimensional hyper-sphere of radius R is given by $2\pi^{n/2}R^{n-1}/\Gamma(n/2)$, where $\Gamma(\cdot)$ is the Gamma function, defined as

$$\Gamma(u) \triangleq \int_0^\infty t^{u-1}e^{-t}\,\mathrm{d}t, \tag{2.39}$$

whose value for $u = n/2$, (n being a positive integer) is given by

$$\Gamma\left(\frac{n}{2}\right) = \begin{cases} (\frac{n}{2}-1)!, & n \text{ is even} \\ 2^{-(n-1)/2}\cdot\sqrt{\pi}\times 1\times 3\times\ldots(n-2), & n \text{ is odd} \end{cases}. \tag{2.40}$$

Thus, the surface area of an n-dimensional hyper-sphere of radius $\sqrt{ns}$ is the volume of $\mathcal{T}_n(s)$ in $n-1$ dimensions:

$$\mathrm{Vol}\{\mathcal{T}_n(s)\} = \frac{2\pi^{n/2}(\sqrt{ns})^{n-1}}{\Gamma(n/2)} \tag{2.41}$$

$$\overset{(*)}{\sim} \frac{2\pi^{n/2}(ns)^{(n-1)/2}}{\sqrt{4\pi/n}(n/2e)^{n/2}} \tag{2.42}$$

$$= \frac{(2\pi es)^{n/2}}{\sqrt{\pi s}} \tag{2.43}$$

$$= \frac{e^{nh(X)}}{\sqrt{\pi s}}, \tag{2.44}$$

where the notation $a_n \sim b_n$, for two positive sequences, $\{a_n\}$ and $\{b_n\}$, means that $a_n/b_n \to 1$ as $n \to \infty$, $(*)$ follows from Stirling's approximation

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n, \tag{2.45}$$

and X is a Gaussian RV $X \sim N(0,s)$, whose differential entropy is $h(X) = \frac{1}{2}\ln(2\pi es)$. Hence, $\mathrm{Vol}\{\mathcal{T}_n(s)\}$ is of the exponential order of

$e^{nh(X)}$ in parallel to the fact that in the finite-alphabet case, the size of a type class is exponentially e^{nH}, where H is the empirical entropy associated with the type (see Section 2.2.1 and [38, Section VII.A]). This result is not a coincidence and we will encounter it repeatedly in the sequel.

Now, if our purpose is to integrate over $\mathbb{R}^n$ a certain function that depends on $\boldsymbol{x}$ only via $\sum_{i=1}^n x_i^2$, we can proceed as follows:

$$\int_{\mathbb{R}^n} f\left(\sum_{i=1}^n x_i^2\right) \mathrm{d}\boldsymbol{x} \overset{(a)}{=} \int_0^\infty \mathrm{d}R \int_{\{\boldsymbol{x}:\ \sum_i x_i^2 = R^2\}} f\left(\sum_{i=1}^n x_i^2\right) \mathrm{d}\boldsymbol{x} \tag{2.46}$$

$$\overset{(b)}{=} \int_0^\infty \mathrm{d}(\sqrt{ns}) \cdot \mathrm{Vol}\left\{\mathcal{T}_n(s)\right\} f(ns) \tag{2.47}$$

$$\overset{(c)}{\sim} \frac{\sqrt{n}}{2} \int_0^\infty \frac{\mathrm{d}s}{\sqrt{s}} \cdot \frac{(2\pi e s)^{n/2}}{\sqrt{\pi s}} f(ns) \tag{2.48}$$

$$= \frac{1}{2}\sqrt{\frac{n}{\pi}} \cdot (2\pi e)^{n/2} \int_0^\infty \mathrm{d}s \cdot s^{n/2-1} f(ns), \tag{2.49}$$

where (a) follows by expressing the integral in two stages, thus combining the contributions of all hyper-sphere surfaces, where the inner integration WRT $\boldsymbol{x}$ in the first line is over the hyper-sphere surface, whose dimension is $n-1$, (b) follows by a change of the outer integration variable $R = \sqrt{ns}$, and (c) follows by substituting (2.44) for $\mathrm{Vol}\{\mathcal{T}_n(s)\}$. We have thus simplified an n-dimensional integral to a one-dimensional integral.

Example 2.1. Consider the calculation of the probability of the event $\sum_{i=1}^n X_i^2 \geq nA$, where $\{X_i\}$ are IID, zero-mean, Gaussian RVs with variance σ^2 and $A > \sigma^2$. In this case, using (2.49),

$$\Pr\left\{\sum_{i=1}^n X_i^2 \geq nA\right\}$$

$$= \int_{\mathbb{R}^n} (2\pi\sigma^2)^{-n/2} \exp\left\{-\sum_{i=1}^n x_i^2/(2\sigma^2)\right\} \cdot \mathbb{1}\left\{\sum_{i=1}^n x_i^2 \geq nA\right\} \mathrm{d}\boldsymbol{x} \tag{2.50}$$

$$\sim \frac{1}{2}\sqrt{\frac{n}{\pi}}(2\pi e)^{n/2} \int_A^\infty s^{n/2-1} (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{sn}{2\sigma^2}\right\} \mathrm{d}s. \tag{2.51}$$

For $A > \sigma^2$, this integral is dominated by the value of the integrand at $s = A$, and therefore, the above expression is of the exponential order of

$$\exp\left\{-\frac{n}{2}\left[\frac{A}{\sigma^2} - \ln\left(\frac{A}{\sigma^2}\right) - 1\right]\right\}. \tag{2.52}$$

At this juncture, one might inquire about the necessity of employing the method of types for the aforementioned example, as well as for several other instances elaborated upon in the subsequent sections of this section. After all, the exponential rate of the aforementioned probability can be readily derived through a straightforward application of Chernoff's bound, renowned for its exponential accuracy [49]. However, the rationale for employing the method of types, not just in this simple instance, but also in the forthcoming sections, is threefold:

- *General applicability.* While the chosen example was intentionally simple, serving as an illustrative vehicle for the underlying technique, the method of types possesses a generality and adaptability that extends to more intricate scenarios. Consider, for instance, an event that encompasses a vector of diverse empirical statistics, confined within a specific spatial region. Such intricate events are beyond the capabilities of Chernoff's bound.

- *Broad utility.* The capacity to gauge the volume of a type class holds significance beyond the mere evaluation of probabilities linked to rare events. Its utility extends to deriving universal hypothesis testing strategies and universal decoders in instances where the source and/or channel characteristics are unknown. For a comprehensive understanding, refer to works such as [115], [116], [124], [138], all of which underscore its importance. This aspect will be elaborated upon in Section 2.8.

- *Enhanced precision.* Through the utilization of the Laplace method of one-dimensional integration, we will come to realize in Section 3 that we can attain not only the accurate exponential order found in the last integral (akin to Chernoff's bound or general large-deviations bounds), but also an asymptotically precise pre-exponential factor.

It is imperative to bear these considerations in mind as we delve into subsequent sections.

2.4 More Refined Gaussian Types

Let us proceed to the next phase. Consider a scenario where the function we need to integrate depends on $\boldsymbol{x}$, not solely through $\sum_{i=1}^n x_i^2$, but also through $\sum_{i=1}^n x_i$. For instance, this arises when calculating the probability of an event like $\{\sum_{i=1}^n (X_i - \mu)^2 \geq nA\}$. In this situation, we must engage with more refined type classes, defined by specific values of both $\sum_{i=1}^n x_i^2$ and $\sum_{i=1}^n x_i$. In simpler terms, our type class now takes the form of an intersection between a hyper-sphere surface and a hyper-plane. Unlike the previous case where we dealt with a simple hyper-sphere, here, an apparent closed-form formula for the volume of this $(n-2)$-dimensional construct, defined by $\sum_{i=1}^n x_i^2 = ns$ and $\sum_{i=1}^n x_i = n\mu$ for given constants $s > 0$ and $\mu \in \mathbb{R}$ (with $s > \mu^2$), is not readily available. At this point, an exact solution to this challenge remains elusive. Nevertheless, we can furnish an approximation that can be continually honed as n becomes increasingly large. This approximation suffices to derive the precise exponential scale of the desired expression, thereby serving our immediate purpose. Subsequently, we will acquaint ourselves with more advanced techniques that, on occasion, enable a significantly more accurate assessment.

Let $\epsilon > 0$ be arbitrarily small and consider the ϵ-inflated version of the type class described above

$$\mathcal{T}_n(s, \mu, \epsilon) \triangleq \left\{ \boldsymbol{x} \colon \left| \frac{1}{n} \sum_{i=1}^n x_i^2 - s \right| \leq \epsilon, \ \left| \frac{1}{n} \sum_{i=1}^n x_i - \mu \right| \leq \epsilon \right\}. \tag{2.53}$$

Further consider an auxiliary PDF of n IID Gaussian RVs of mean μ and variance $s - \mu^2$, that is,

$$g(\boldsymbol{x}) = \frac{\exp\left\{ -\frac{1}{2(s-\mu^2)} \sum_{i=1}^n (x_i - \mu)^2 \right\}}{[2\pi(s-\mu^2)]^{n/2}}. \tag{2.54}$$

Then,

$$1 \overset{(a)}{\geq} \int_{\mathcal{T}_n(s,\mu,\epsilon)} g(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \tag{2.55}$$

$$= \int_{\mathcal{T}_n(s,\mu,\epsilon)} \frac{\exp\left\{ -\frac{1}{2(s-\mu^2)} \sum_{i=1}^n (x_i - \mu)^2 \right\}}{[2\pi(s-\mu^2)]^{n/2}} \, \mathrm{d}\boldsymbol{x} \tag{2.56}$$

$$= \int_{\mathcal{T}_n(s,\mu,\epsilon)} \frac{\exp\left\{-\frac{1}{2(s-\mu^2)}\left[\sum_{i=1}^n x_i^2 - 2\mu\sum_{i=1}^n x_i + n\mu^2\right]\right\}}{\left[2\pi(s-\mu^2)\right]^{n/2}} \, \mathrm{d}\boldsymbol{x} \quad (2.57)$$

$$\overset{(b)}{\geq} \int_{\mathcal{T}_n(s,\mu,\epsilon)} \frac{\exp\left\{-\frac{\left[n(s+\epsilon)-2\mu\cdot n(\mu-\epsilon\cdot\mathrm{sgn}(\mu))+n\mu^2\right]}{2(s-\mu^2)}\right\}}{\left[2\pi(s-\mu^2)\right]^{n/2}} \mathrm{d}\boldsymbol{x} \quad (2.58)$$

$$= \mathrm{Vol}\left\{\mathcal{T}_n(s,\mu,\epsilon)\right\} \cdot \frac{\exp\left\{-\frac{n\left[s+\epsilon-\mu^2+2\epsilon|\mu|\right]}{2(s-\mu^2)}\right\}}{\left[2\pi(s-\mu^2)\right]^{n/2}} \quad (2.59)$$

$$= \mathrm{Vol}\left\{\mathcal{T}_n(s,\mu,\epsilon)\right\} \cdot \frac{\exp\left\{-\frac{n}{2} - \frac{n\epsilon(2|\mu|+1)}{2(s-\mu^2)}\right\}}{\left[2\pi(s-\mu^2)\right]^{n/2}} \quad (2.60)$$

$$= \mathrm{Vol}\left\{\mathcal{T}_n(s,\mu,\epsilon)\right\} \cdot \frac{\exp\left\{-\frac{n\epsilon(2|\mu|+1)}{2(s-\mu^2)}\right\}}{\left[2\pi e(s-\mu^2)\right]^{n/2}}, \quad (2.61)$$

where (a) follows since the probability of $\mathcal{T}_n(s,\mu,\epsilon)$ under $g(\boldsymbol{x})$ must be less than 1, (b) follows since within $\mathcal{T}_n(s,\mu,\epsilon)$ it holds that $\sum_{i=1}^n x_i^2 \leq n(s+\epsilon)$ and $|\sum_{i=1}^n x_i - \mu| \leq n\epsilon$, which implies that

$$2\mu\sum_{i=1}^n x_i \geq 2\mu\left[n\mu - n\epsilon\cdot\mathrm{sign}(\mu)\right] = 2n\mu^2 - 2n\epsilon|\mu|. \quad (2.62)$$

Consequently,

$$\mathrm{Vol}\left\{\mathcal{T}_n(s,\mu,\epsilon)\right\} \leq \left[2\pi e(s-\mu^2)\right]^{n/2} \cdot \exp\left\{\frac{n\epsilon(2|\mu|+1)}{2(s-\mu^2)}\right\}. \quad (2.63)$$

To establish a lower bound, consider the application of the weak law of large numbers (WLLN), which asserts that as n approaches infinity, the probability of the complement of $\mathcal{T}_n(s,\mu,\epsilon)$ under the PDF g diminishes to zero, for any fixed $\epsilon > 0$. Notably, we can even allow ϵ to tend towards zero, albeit at a pace that remains gentle relative to the growth of n. This probability can be readily bounded from above by employing either Chebyshev's or Chernoff's bound. Denote the resultant upper bound for this probability as δ_n. This leads us to the following expression:

$$1-\delta_n$$

$$\leq \int_{\mathcal{T}_n(s,\mu,\epsilon)} g(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \quad (2.64)$$

$$= \int_{\mathcal{T}_n(s,\mu,\epsilon)} \frac{\exp\left\{-\frac{1}{2(s-\mu^2)}\left[\sum_{i=1}^n x_i^2 - 2\mu\sum_{i=1}^n x_i + n\mu^2\right]\right\}}{\left[2\pi(s-\mu^2)\right]^{n/2}}\,\mathrm{d}\boldsymbol{x} \quad (2.65)$$

$$\leq \int_{\mathcal{T}_n(s,\mu,\epsilon)} \frac{\exp\left\{-\frac{\left[n(s-\epsilon)-2\mu\cdot n(\mu+\epsilon\cdot\mathrm{sgn}(\mu))+n\mu^2\right]}{2(s-\mu^2)}\right\}}{\left[2\pi(s-\mu^2)\right]^{n/2}}\,\mathrm{d}\boldsymbol{x} \quad (2.66)$$

$$= \mathrm{Vol}\left\{\mathcal{T}_n(s,\mu,\epsilon)\right\}\cdot\frac{\exp\left\{-\frac{n}{2(s-\mu^2)}\left[s-\epsilon-\mu^2-2\epsilon|\mu|\right]\right\}}{\left[2\pi(s-\mu^2)\right]^{n/2}} \quad (2.67)$$

$$= \mathrm{Vol}\left\{\mathcal{T}_n(s,\mu,\epsilon)\right\}\cdot\frac{\exp\left\{\frac{n\epsilon(2|\mu|+1)}{2(s-\mu^2)}\right\}}{\left[2\pi e(s-\mu^2)\right]^{n/2}}, \quad (2.68)$$

where the steps are justified similarly to the justification of the steps in (2.61). Consequently,

$$\mathrm{Vol}\left\{\mathcal{T}_n(s,\mu,\epsilon)\right\} \geq (1-\delta_n)\cdot\left[2\pi e(s-\mu^2)\right]^{n/2}\cdot\exp\left\{-\frac{n\epsilon(2|\mu|+1)}{2(s-\mu^2)}\right\}. \quad (2.69)$$

As we allow ϵ to approach infinitesimally small values, we discern that the volume of $\mathcal{T}_n(s,\mu,\epsilon)$ essentially aligns with the exponential order given by:

$$\left[2\pi e(s-\mu^2)\right]^{n/2} = (2\pi e s)^{n/2}\cdot\left(1-\frac{\mu^2}{s}\right)^{n/2}. \quad (2.70)$$

The first factor, $(2\pi es)^{n/2}$, corresponds to e^{nh}, as we previously deduced in Section 2.3. Concurrently, the subsequent factor, $(1-\mu^2/s)^{n/2}$, embodies the volume reduction attributed to the intersection with the (ϵ-inflated) hyper-plane, namely $n(\mu-\epsilon) \leq \sum_{i=1}^n x_i \leq n(\mu+\epsilon)$. Consequently, it becomes evident that there is no sacrifice in terms of the exponential order when the hyper-sphere intersects with the hyper-plane that encompasses the origin ($\mu = 0$). Stated differently, the majority of volume is captured by elements within $\mathcal{T}_n(s,\mu,\epsilon)$ that exhibit a property where the sum of their coordinates is relatively modest (in absolute value).

As evident, the underpinning of the volume's upper and lower bound derivation is straightforward, yet this same concept retains its relevance in more intricate scenarios. When faced with an ϵ-enlarged type class, characterized by linear and quadratic criteria on $\boldsymbol{x}$, we construct an auxiliary Gaussian PDF, denoted as $g(\cdot)$, which exhibits two key attributes:

1. The likelihood of the type class under $g(\cdot)$ converges towards unity as n approaches infinity.

2. The value of the PDF of all sequences situated within the type class are virtually the same, differing only exponentially by a factor that scales with ϵ. This value of the PDF is denoted as g_0.

The volume of the type class then aligns with the exponential order of $1/g_0$. In the prior calculation, g_0 equates to $[2\pi e(s-\mu^2)]^{-n/2}$, leading to an exponential volume of $1/g_0 = [2\pi e(s-\mu^2)]^{n/2}$. Given that our objective is to pinpoint the correct exponential order rather than striving for precise evaluation at this stage, the demand in the first item mentioned earlier can actually be considerably relaxed. It is even permissible for the type class probability to approach zero, as long as the rate of decay remains sub-exponential in n.

Example 2.2. Consider the calculation of the probability of the event $\{\sum_{i=1}^n (X_i - A)^2 \geq nB\}$ when $\{X_i\}$ are IID, zero-mean Gaussian RVs with variance σ^2. To this end, we cover the set $\mathcal{E} \triangleq \{\boldsymbol{x}\colon \sum_{i=1}^n (x_i - A)^2 \geq nB\}$ by a finite number of ϵ-inflated types classes, $\{\mathcal{T}_n(s_i, \mu_j, \epsilon)\}$, where $s_i = i\epsilon$ and $\mu_j = j\epsilon$, where i and j are odd integers. Since

$$\sum_{i=1}^n (x_i - A)^2 = \sum_{i=1}^n x_i^2 - 2A\sum_{i=1}^n x_i + nA^2, \tag{2.71}$$

these are all the type classes with the property $s_i - 2A\mu_j + A^2 > B$. To avoid the necessity of dealing with contributions of infinitely many such type classes, we proceed as follows. Let us partition the set $\mathcal{E}$ into two disjoint subsets, $\mathcal{E}_1 \triangleq \{\boldsymbol{x}\colon nB \leq \sum_{i=1}^n (x_i - A)^2 < nC\}$ and $\mathcal{E}_2 \triangleq \{\boldsymbol{x}\colon \sum_{i=1}^n (x_i - A)^2 \geq nC\}$, for some $C > B$ arbitrarily large. The idea is that $\mathcal{E}_1$ contains finitely many types, whereas the contribution of $\mathcal{E}_2$ can be upper bounded by simple (crude) bound, which for large enough C, would yield an exponential decay faster than that of $\mathcal{E}_1$, and so, the contribution of $\mathcal{E}_2$ can be neglected altogether. We will thus show that $\Pr\{\mathcal{E}\} \doteq \Pr\{\mathcal{E}_1\}$. Specifically, first observe that

$$\sum_{i=1}^{n}(x_i - A)^2 = \sum_{i=1}^{n} x_i^2 - 2A\sum_{i=1}^{n} x_i + nA^2 \tag{2.72}$$

$$\leq \sum_{i=1}^{n} x_i^2 + 2|A| \cdot \left|\sum_{i=1}^{n} x_i\right| + nA^2 \tag{2.73}$$

$$\overset{(*)}{\leq} \sum_{i=1}^{n} x_i^2 + 2|A| \cdot \sqrt{n\sum_{i=1}^{n} x_i^2} + nA^2 \tag{2.74}$$

$$= n \cdot \left(\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2} + |A|\right)^2, \tag{2.75}$$

where $(*)$ follows from the Schwarz–Cauchy inequality. Therefore,

$$\Pr\{\mathcal{E}_2\}$$

$$= \Pr\left\{\sum_{i=1}^{n}(X_i - A)^2 \geq nC\right\} \tag{2.76}$$

$$\leq \Pr\left\{n \cdot \left(\sqrt{\frac{1}{n}\sum_{i=1}^{n} X_i^2} + |A|\right)^2 \geq nC\right\} \tag{2.77}$$

$$= \Pr\left\{\sum_{i=1}^{n} X_i^2 \geq n(\sqrt{C} - |A|)^2\right\} \tag{2.78}$$

$$\overset{(*)}{\leq} \exp\left\{-\frac{n}{2}\left[\frac{(\sqrt{C} - |A|)^2}{\sigma^2} - \ln\left(\frac{(\sqrt{C} - |A|)^2}{\sigma^2}\right) - 1\right]\right\}, \tag{2.79}$$

where $(*)$ follows from Chernoff's bound [229, Prop. 13.1.3, p. 374]. By selecting large enough C, it becomes apparent that $\Pr\{\mathcal{E}_2\}$ must decay with an arbitrarily fast exponential rate. In particular, it can be made faster (and hence negligible) compared to the contribution of $\mathcal{E}_1$. It is therefore enough to confine attention to $\mathcal{E}_1$. Now, within $\mathcal{E}_1$, there are finitely many type classes $\{\mathcal{T}_n(s_i, \mu_j, \epsilon)\}$, as i cannot exceed $C/(2\epsilon)$ and $|j|$ cannot exceed $\sqrt{C}/(2\epsilon)$ (because in the definition of $\mathcal{T}_n(s, \mu, \epsilon)$, $|\mu|$ cannot exceed $\sqrt{s}$, or else $\mathcal{T}_n(s, \mu, \epsilon)$ would be empty for small ϵ). It follows then that the total number of ϵ-inflated type classes is less than $C^{3/2}/(2\epsilon^2)$. Therefore, $\Pr\{\mathcal{E}_1\} \doteq \Pr\{\mathcal{E}\}$ is determined by the probability of the dominant type class within $\mathcal{E}_1$. Each such type class with $\mathcal{E}_1$ contributes the product $\text{Vol}\{\mathcal{T}_n(s_i, \mu_j, \epsilon) \doteq [2\pi e(s_i - \mu_j^2)]^{n/2}$ times the PDF within that type class, $g(\boldsymbol{x}) \doteq (2\pi\sigma^2)^{-n/2} e^{-ns_i^2/(2\sigma^2}$, which is given by

$$\begin{aligned}
&\left[2\pi e(s_i - \mu_j^2)\right]^{n/2} \cdot (2\pi\sigma^2)^{-n/2} e^{-ns_i^2/(2\sigma^2)} \\
&= \exp\left\{-\frac{1}{2}\left[\frac{s_i}{\sigma^2} - \ln\left(\frac{s_i - \mu_j^2}{\sigma^2}\right) - 1\right]\right\}. \qquad (2.80)
\end{aligned}$$

Since there are finitely many type classes within $\mathcal{E}_1$, the probability of $\mathcal{E}$ (or, equivalently, of $\mathcal{E}_1$) is then dominated by the maximum of this expression over all type classes within $\mathcal{E}_1$, which in the limit of small $\epsilon > 0$ becomes

$$\begin{aligned}
&\exp\left\{-\frac{1}{2} \inf_{\{(s,\mu):\, s - 2A\mu + A^2 \geq B\}} \left[\frac{s}{\sigma^2} - \ln\left(\frac{s - \mu^2}{\sigma^2}\right) - 1\right]\right\} \\
&= \exp\left\{-\frac{1}{2} \inf_{\{(s,\mu):\, s - 2A\mu + A^2 \geq B\}} \left[\frac{s - \mu^2}{\sigma^2} - \ln\left(\frac{s - \mu^2}{\sigma^2}\right) - 1 + \frac{\mu^2}{\sigma^2}\right]\right\}. \qquad (2.81)
\end{aligned}$$

Note that the exponential rate of the probability of $\mathcal{E}$ is strictly positive as long as the set $\{(s, \mu) \colon s - 2A\mu + A^2 \geq B\}$ does not include the pair $(s, \mu) = (\sigma^2, 0)$, which amounts to the condition $\sigma^2 < B - A^2$. In addition, the objective function of this minimization can be interpreted as the KL divergence between two Gaussian PDFs, $\mathcal{N}(\mu, s - \mu^2)$ and $\mathcal{N}(0, \sigma^2)$, in analogy to the form of exponential rates of probabilities of rare events that are computed using the traditional method of types, where the KL divergence between two finite-alphabet distributions is minimized subject to a constraint (or constraints) corresponding to the event in question (see also the fourth property of the method of types in Section 2.2.1, and the calculation near the end of Section 2.3). This also agrees with basic foundations in large-deviations theory [49].

2.5 Conditional Gaussian Types

In analogy to the finite-alphabet case, the notion of conditional types exists also in the Gaussian case. Given a sequence $\boldsymbol{y} = (y_1, y_2, \ldots, y_n) \in$

$\mathbb{R}^n$, a conditional Gaussian type class is defined as the set of $\{\boldsymbol{x}\}$ with given values of $\frac{1}{n}\sum_{i=1}^n x_i^2$ and $\frac{1}{n}\sum_{i=1}^n x_i y_i$. In the ϵ-inflated version, this amounts to

$$\mathcal{T}_n(s,c,\epsilon|\boldsymbol{y}) = \left\{\boldsymbol{x}\colon \left|\frac{1}{n}\sum_{i=1}^n x_i^2 - s\right| \leq \epsilon,\ \left|\frac{1}{n}\sum_{i=1}^n x_i y_i - c\right| \leq \epsilon\right\}, \quad (2.82)$$

where $s \geq c^2/P_y$ and $P_y \triangleq \frac{1}{n}\sum_{i=1}^n y_i^2$, due to the Schwarz–Cauchy inequality. In fact, this is an extension of the refined Gaussian types considered in Section 2.4, where $y_i = 1$ for all i. To estimate the volume of this conditional type class, consider the Gaussian channel,

$$g(\boldsymbol{x}|\boldsymbol{y}) = \frac{\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \alpha y_i)^2\right\}}{(2\pi\sigma^2)^{n/2}}, \quad (2.83)$$

and let us select the parameters of this channel to be

$$\alpha = \frac{c}{P_y}, \quad (2.84)$$

and

$$\sigma^2 = s - \frac{c^2}{P_y}, \quad (2.85)$$

for reasons that will become apparent shortly. It is easy to check that the channel $g(\boldsymbol{x}|\boldsymbol{y})$ has the two desired properties: It assigns a high probability and an approximately uniform distribution within $\mathcal{T}_n(s,c,\epsilon|\boldsymbol{y})$, which is of the exponential order of

$$g_0 = \left[2\pi e(s - c^2/P_y)\right]^{-n/2} = e^{-nh(X|Y)}, \quad (2.86)$$

where $h(X|Y)$ is the conditional entropy of a Gaussian zero-mean, RV X, with variance s, given a jointly Gaussian, zero-mean, RV Y with variance P_y and $\mathbb{E}\{XY\} = c$. The expression $s - c^2/P_y$ is then the conditional variance of X given Y, which is also the minimum mean squared error (MMSE) in estimating X based on Y.

Example 2.3. Consider a simplified version of the problem of universal decoding of [116] for the additive white Gaussian noise (AWGN) channel,

$$Y_i = \alpha X_i + Z_i, \qquad i = 1, 2, \ldots, n, \quad (2.87)$$

where $\{Z_i\}$ are IID, zero-mean Gaussian RVs with variance σ^2, α is an unknown fixed parameter, $\{X_i\}$ are the channel inputs, and $\{Y_i\}$ are the channel outputs. Consider now a random codebook for channel coding, where $M = e^{nR}$ codewords of length n are selected independently at random where each codeword is drawn under a PDF, $q(\boldsymbol{x})$, which is uniform across the surface of a hyper-sphere of radius $\sqrt{nP}$. In [116], it is shown that in the limit of small $\epsilon > 0$,

$$\begin{aligned}&\mathrm{Vol}\left\{\mathcal{T}_n\left(P, \frac{1}{n}\sum_{i=1}^n x_i y_i, \epsilon\right)\right\}\\ &\doteq \exp\left\{\frac{1}{2}\ln\left[2\pi e\left(P - \frac{\left[\frac{1}{n}\sum_{i=1}^n x_i y_i\right]^2}{\frac{1}{n}\sum_{i=1}^n y_i^2}\right)\right]\right\} \qquad (2.88)\end{aligned}$$

can serve as a universal decoding metric (independent of the unknown α), which achieves the same random coding exponent as that of the maximum-likelihood (ML) decoder, that is cognizant of α. This is equivalent to a decoder that maximizes $|\sum_{i=1}^n x_i y_i|$ among all codewords. This is the Gaussian analogue to the well known universal minimum entropy decoder, or, equivalently, the maximum mutual information (MMI) decoder [41] for DMCs, which achieves the random coding error exponent for ensembles of fixed composition codes. In [116], the problem is more general in the sense that an interference signal may also be present, and so, more interesting decoders are derived (see also [88], [90] for further developments).

The notion of a conditional Gaussian type can be easily extended to account for conditioning on more than one vector $\boldsymbol{y}$. Let $\boldsymbol{y}^1, \boldsymbol{y}^2, \ldots \boldsymbol{y}^k$ be k given vectors in $\mathbb{R}^n$, where k is fixed, independently of n. Consider the conditional type defined by

$$\begin{aligned}&\mathcal{T}_n\left(s, c_1, \ldots, c_k, \epsilon|\boldsymbol{y}^1, \ldots, \boldsymbol{y}^k\right) \triangleq\\ &\left\{\boldsymbol{x}\colon \left|\frac{1}{n}\sum_{i=1}^n x_i^2 - s\right| \le \epsilon,\ \left|\frac{1}{n}\sum_{i=1}^n x_i y_i^j - c_j\right| \le \epsilon,\ \forall\ j = 1, \ldots, k\right\}. \qquad (2.89)\end{aligned}$$

Here, we can use a conditional PDF of the form

$$g(\boldsymbol{x}|\boldsymbol{y}^1,\ldots,\boldsymbol{y}^k) = \frac{\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n\left(x_i - \sum_{j=1}^k \alpha_i y_i^j\right)^2\right\}}{(2\pi\sigma^2)^{n/2}}, \tag{2.90}$$

and tune the parameters $(\sigma^2, \alpha_1, \ldots, \alpha_k)$ such that the conditional type class $\mathcal{T}_n(s, c_1, \ldots, c_k, \epsilon|\boldsymbol{y}^1, \ldots, \boldsymbol{y}^k)$ would have high probability for large n. The resulting volume would then be of the exponential order of $\exp\{nh(X|Y_1,\ldots,Y_k)\}$ where

$$h(X|Y_1,\ldots,Y_k) = \frac{1}{2}\ln\left(2\pi e \cdot \text{MMSE}\{X|Y_1,\ldots,Y_k\}\right), \tag{2.91}$$

and $\text{MMSE}\{X|Y_1,\ldots,Y_k\}$ is the MMSE of estimating X based on $Y_1,\ldots,Y_k$ where $(X, Y_1,\ldots,Y_k)$ is a zero-mean Gaussian vector with $\mathbb{E}\{X^2\} = s$, $\mathbb{E}\{XY_j\} = c_j$ and a given covariance matrix of $(Y_1,\ldots,Y_k)$ with $\mathbb{E}\{Y_m Y_l\} = \frac{1}{n}\sum_{i=1}^n y_i^m y_i^l$. It is not necessary to find the coefficients of the optimal (linear) estimator of X based on $(Y_1,\ldots,Y_k)$ in order to calculate $\text{MMSE}\{X|Y_1,\ldots,Y_k\}$, as it is possible to calculate the latter directly from the covariance matrix of $(X, Y_1,\ldots,Y_k)$. This is based on the following information-theoretic consideration, similar to [36, Section 12.6]: Let $\Lambda(Y_1,\ldots,Y_k)$ and $\Lambda(X, Y_1,\ldots,Y_k)$ denote the covariance matrices of $(Y_1,\ldots,Y_k)$ and $(X, Y_1,\ldots,Y_k)$, respectively. These matrices must both be positive definite, otherwise, the problem is singular (*i.e.*, there is a redundant constraint, such as $Y_3 = Y_2 + Y_1$) . Now, on the one hand,

$$\begin{aligned} &h(X|Y_1,\ldots,Y_k) \\ &= h(X, Y_1,\ldots,Y_k) - h(Y_1,\ldots,Y_k) &(2.92)\\ &= \frac{1}{2}\ln\left[(2\pi e)^{k+1}|\Lambda(X,Y_1,\ldots,Y_k)|\right] \\ &\quad - \frac{1}{2}\ln\left[(2\pi e)^k|\Lambda(Y_1,\ldots,Y_k)|\right] &(2.93)\\ &= \frac{1}{2}\ln\left[2\pi e\cdot\frac{|\Lambda(X,Y_1,\ldots,Y_k)|}{|\Lambda(Y_1,\ldots,Y_k)|}\right], &(2.94)\end{aligned}$$

and on the other hand, denoting by $(\alpha_1^*,\ldots,\alpha_k^*)$ the coefficients of the optimal (linear) estimator, we have

$$h(X|Y_1,\ldots,Y_k) = h\left(X - \sum_{i=1}^k \alpha_i^* Y_i \middle| Y_1,\ldots,Y_k\right) \tag{2.95}$$

$$\overset{(*)}{=} h\left(X - \sum_{i=1}^{k} \alpha_i^* Y_i\right) \tag{2.96}$$

$$= \frac{1}{2} \ln\left(2\pi e \cdot \text{MMSE}\left\{X|Y_1, \dots, Y_k\right\}\right), \tag{2.97}$$

where $(*)$ follows from the orthogonality principle, which in the Gaussian case implies independence between $X - \sum_{i=1}^{k} \alpha_i^* Y_i$ and $(Y_1, \dots, Y_k)$. By equating the two expressions of $h(X|Y_1, \dots, Y_k)$, we have

$$\text{MMSE}\left\{X|Y_1, \dots, Y_k\right\} = \frac{|\Lambda(X, Y_1, \dots, Y_k)|}{|\Lambda(Y_1, \dots, Y_k)|}. \tag{2.98}$$

Thus, the volume can be calculated directly, without recourse of finding first the optimal coefficients.

2.6 Gauss–Markov Types

So far, we have dealt with Gaussian types defined by empirical second order statistics that correspond to memoryless Gaussian sources, namely, the empirical mean and the empirical second moment. As we mentioned before, in the finite-alphabet case, the method of types has been extended to Markov-types, namely, types defined by counts of transitions between consecutive letters along a sequence, that it, the number of time indices $\{i\}$ along an n-sequence $\boldsymbol{x}$ such that $x_{i-1} = a$ and $x_i = b$, where $a, b \in \mathcal{X}$ [38], [47], [155]. But what would be the corresponding Markov extension of Gaussian types?

The simplest definition of a first-order Gauss–Markov type class is defined as the set of all $\boldsymbol{x} \in \mathbb{R}^n$ with prescribed values of empirical variance, $\frac{1}{n}\sum_{i=1}^{n} x_i^2$ and the empirical first autocorrelation, $\frac{1}{n}\sum_{i=1}^{n} x_i x_{i-1}$ (for a given x_0). The ϵ-inflated version would then be naturally defined as

$$\mathcal{T}_n(s_0, s_1, \epsilon) \triangleq \left\{\boldsymbol{x}\colon \left|\frac{1}{n}\sum_{i=1}^{n} x_i^2 - s_0\right| \le \epsilon, \ \left|\frac{1}{n}\sum_{i=1}^{n} x_i x_{i-1} - s_1\right| \le \epsilon\right\}, \tag{2.99}$$

where $|s_1| \le s_0$. What is the volume of $\mathcal{T}_n(s_0, s_1, \epsilon)$?

The basic idea is the same as before: We seek a Gaussian PDF, which assigns to $\mathcal{T}_n(s_0, s_1, \epsilon)$ a high probability, and at the same time, it is approximately uniform (in the exponential sense) across $\mathcal{T}_n(s_0, s_1, \epsilon)$. Given s_0 and s_1, let

$$\sigma^2 = s_0 - \frac{s_1^2}{s_0} \tag{2.100}$$

and

$$\rho = \frac{s_1}{s_0}, \tag{2.101}$$

and consider the first-order Gauss–Markov process (also known as first-order autoregressive (AR) process),

$$X_i = \rho X_{i-1} + Z_i, \quad i = 1, 2, \ldots, n, \quad X_0 = x_0, \tag{2.102}$$

where $\{Z_i\}$ are IID, zero-mean, Gaussian RVs with variance σ^2, and x_0 is a fixed initial condition. The joint PDF of a given sample $\boldsymbol{x}$ from this process, conditioned on x_0 is given by

$$\begin{aligned} g(\boldsymbol{x}|x_0) &= \frac{\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \rho x_{i-1})^2\right\}}{(2\pi\sigma^2)^{n/2}} \qquad (2.103) \\ &= \frac{\exp\left\{-\frac{1}{2\sigma^2}\left[(1+\rho^2)\sum_{i=1}^n x_i^2 - 2\rho\sum_{i=1}^n x_i x_{i-1} + \rho^2(x_0^2 - x_n^2)\right]\right\}}{(2\pi\sigma^2)^{n/2}}. \qquad (2.104) \end{aligned}$$

It is apparent that $g(\boldsymbol{x}|x_0)$ depends on $\boldsymbol{x}$ only via $\sum_{i=1}^n x_i^2$, $\sum_{i=1}^n x_i x_{i-1}$, and $\Delta \triangleq x_0^2 - x_n^2$. Assuming that the value of Δ is fixed (*i.e.*, independent of n), it is readily seen that within $\mathcal{T}_n(s_0, s_1, \epsilon)$, the PDF $g(\boldsymbol{x}|x_0)$, with the choices (2.100) and (2.101), is essentially [neglecting ϵ and Δ, which do not affect the exponential rate of (2.104)], $g_0 = [2\pi e(s_0 - s_1^2/s_0)]^{-n/2}$. Also, by the ergodicity of the process, $\mathcal{T}_n(s_0, s_1, \epsilon)$ has high probability for large n and fixed $\epsilon > 0$, and so, both conditions are satisfied. The volume is, therefore, of the exponential of order of

$$\begin{aligned} \frac{1}{g_0} &= \left[2\pi e\left(s_0 - \frac{s_1^2}{s_0}\right)\right]^{n/2} \qquad (2.105) \\ &= \exp\left\{\frac{n}{2}\ln\left[2\pi e\left(s_1 - \frac{s_1^2}{s_0}\right)\right]\right\} \qquad (2.106) \end{aligned}$$

$$= \exp\left\{\frac{n}{2}\ln(2\pi e\sigma^2)\right\} \tag{2.107}$$

$$= e^{nh(X_2|X_1)}, \tag{2.108}$$

where $h(X_2|X_1)$ is the conditional differential entropy of X_2 given X_1, in analogy to the parallel result for finite-alphabet Markov types, where the size of a type class is exponentially $e^{nH(X_2|X_1)}$, where $H(X_2|X_1)$ is the conditional entropy associated with the corresponding Markov process (see Section 2.2.1).

The intuitive explanation for this expression of the volume is as follows: Consider the linear transformation that maps a realization $\boldsymbol{z} = (z_1, \ldots, z_n)$ of the random vector $\boldsymbol{Z} = (Z_1, \ldots, Z_n)$ into $\boldsymbol{x} = (x_1, \ldots, x_n)$, which is a realization of $\boldsymbol{X} = (X_1, \ldots, X_n)$. This transformation, which is given by $x_t = \sum_{i=0}^{t} \rho^i z_{t-i}$, can be represented by an $n \times n$ triangular transformation matrix, *i.e.*,

$$\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ \rho & 1 & 0 & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \rho^{n-1} & \rho^{n-2} & \ldots & \rho & 1 \end{pmatrix} \cdot \begin{pmatrix} z_1 \\ z_2 \\ \cdots \\ z_n \end{pmatrix}. \tag{2.109}$$

Now, consider the ϵ-inflated surface of the hyper-sphere of radius $\sqrt{n\sigma^2}$ of $\boldsymbol{z}$-sequences, which form the Gaussian type of the driving noise process, $\{Z_i\}$. The volume of this type class is exponentially $[2\pi e\sigma^2]^{n/2} = [2\pi e(s_0 - s_1^2/s_0)]^{n/2}$. But these typical $\boldsymbol{z}$-sequences are mapped into corresponding $\boldsymbol{x}$-sequences, by the above triangular transformation matrix whose diagonal terms are all equal to 1, and hence its Jacobian is also equal to 1. In other words, the transformation from $\boldsymbol{z}$ to $\boldsymbol{x}$ preserves volumes, and so, the volume of the ϵ-inflated surface of a hyper-sphere of typical $\boldsymbol{z}$-sequences is transformed by the above matrix into a hyper-ellipsoid of typical $\boldsymbol{x}$-sequences of exactly the same volume.

Example 2.4. Consider the calculation of the exponential decay rate of

$$\Pr\left\{\sum_{t=1}^{n} X_t X_{t-1} \geq \rho \sum_{t=1}^{n} X_t^2\right\}, \tag{2.110}$$

for some $\rho > 0$, where $\{X_t\}$ are IID zero-mean, Gaussian RVs with variance σ^2. Since the volume of the type class is of the exponential order of $[2\pi e(s_0 - s_1^2/s_0)]^{n/2} = [2\pi e s_0(1 - s_1^2/s_0^2)]^{n/2}$, and the PDF within a type class is of the exponential scale of $(2\pi\sigma^2)^{-n/2}\exp\{-ns_0/(2\sigma^2)\}$, the exponent is given by

$$\inf_{s_0\geq 0,\ s_1/s_0\geq\rho}\left[\frac{s_0}{2\sigma^2} + \frac{1}{2}\ln(2\pi\sigma^2) - \frac{1}{2}\ln(2\pi e s_0) - \frac{1}{2}\ln\left(1 - \frac{s_1^2}{s_0^2}\right)\right]$$
$$= \inf_{s_0\geq 0}\left[\frac{s_0}{2\sigma^2} + \frac{1}{2}\ln(2\pi\sigma^2) - \frac{1}{2}\ln(2\pi e s_0)\right] - \frac{1}{2}\ln(1-\rho^2) \tag{2.111}$$
$$= -\frac{1}{2}\ln(1-\rho^2). \tag{2.112}$$

More generally, consider a k-th order Gauss–Markov type, defined by

$$\mathcal{T}_n(s_0, s_1, \ldots, s_k, \epsilon) \triangleq \left\{\boldsymbol{x}\colon \left|\frac{1}{n}\sum_{i=1}^{n} x_i x_{i-j} - s_j\right| \leq \epsilon,\ \forall j = 0, 1, \ldots, k\right\}, \tag{2.113}$$

for given $(s_0, s_1, \ldots, s_k)$ and some $(x_0, x_{-1}, \ldots, x_{-(k-1)})$. It is assumed that the $(k+1)\times(k+1)$ matrix S whose (i,j)-th entry $(i,j \in \{0,1,\ldots,k\})$ is $s_{|i-j|}$ is a positive definite matrix. Here, we find a matching k-th order AR process,

$$X_t = \sum_{i=1}^{k} a_i X_{t-i} + Z_t, \qquad t = 1, 2, \ldots, \tag{2.114}$$

where $\{Z_t\}$ is again Gaussian white noise with variance σ^2, such that $\mathbb{E}\{X_t X_{t-i}\} = s_i$ for all $i = 0, 1, \ldots, k$. Given $(s_0, s_1, \ldots, s_k)$, the corresponding parameter vector, $(\sigma^2, a_1, \ldots, a_k)$ of the AR process is obtained by solving the Yule–Walker equations [162, Eqs. (12-41a), (12-41b)],

$$\sum_{i=1}^{k} a_i s_{|i-j|} = s_j, \qquad j = 1, 2, \ldots, k, \tag{2.115}$$

and

$$\sigma^2 = s_0 - \sum_{i=1}^{k} a_i s_i. \tag{2.116}$$

The corresponding PDF g, which is given by

$$g(\boldsymbol{x}) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{t=1}^{n}\left(x_t - \sum_{i=1}^{k} a_i x_{t-i}\right)^2\right\}, \tag{2.117}$$

has the two desired properties of exponential uniformity within the Gauss–Markov type $\mathcal{T}_n(s_0, s_1, \ldots, s_k, \epsilon)$ and assigning high probability to $\mathcal{T}_n(s_0, s_1, \ldots, s_k, \epsilon)$. Here too, $g_0 = (2\pi e\sigma^2)^{-n/2}$ which implies that the volume of the type class is essentially

$$\frac{1}{g_0} = (2\pi e\sigma^2)^{n/2}. \tag{2.118}$$

The intuition is the same as before: The mapping from $\boldsymbol{x}$ to $\boldsymbol{z}$ is by a triangular matrix whose diagonal entries are all equal to 1, and so is its Jacobian. Therefore, it preserves volumes, and so is the inverse transformation, which maps the hyper-sphere surface of volume $(2\pi e\sigma^2)^{n/2}$ in the $\boldsymbol{z}$-domain into the typical hyper-ellipsoid of the same volume in the $\boldsymbol{x}$-domain. Now, let $S(e^{i\omega}) = \sigma^2/|1 - \sum_{j=1}^{k} a_j e^{-j\omega i}|^2$ is the spectrum of $\{X_t\}$, where $i \triangleq \sqrt{-1}$. Since σ^2 is the variance of the innovation process, the differential entropy rate of $\{X_t\}$ is given by

$$\begin{aligned}
h &= \lim_{n\to\infty} \frac{h(\boldsymbol{X})}{n} && (2.119)\\
&= \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln[2\pi e S(e^{i\omega})]\, \mathrm{d}\omega && (2.120)\\
&= \frac{1}{2}\ln(2\pi e) + \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln S(e^{i\omega})\, \mathrm{d}\omega && (2.121)\\
&= \frac{1}{2}\ln(2\pi e) + \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln\left[\frac{\sigma^2}{\left|1 - \sum_{j=1}^{k} a_j e^{-j\omega i}\right|^2}\right] \mathrm{d}\omega && (2.122)\\
&= \frac{1}{2}\ln(2\pi e) + \frac{1}{2}\ln\sigma^2 - \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln\left|1 - \sum_{j=1}^{k} a_j e^{-j\omega i}\right| \mathrm{d}\omega && (2.123)\\
&= \frac{1}{2}\ln(2\pi e\sigma^2), && (2.124)
\end{aligned}$$

where in the last equality we have used the Kolmogorov–Szegö relation [162, p. 491] between the spectrum and the innovation variance, while noting that $\int_{-\pi}^{\pi} \ln[1 - \sum_{j=1}^{k} a_j e^{-j\omega i}]\, \mathrm{d}\omega = 0$ since all zeroes of the function $1 - \sum_{j=1}^{k} a_j z^{-j}$ must be within the unit circle. This implies that the volume continues to be of the exponential order of e^{nh}. Similarly as before, σ^2 can be found directly from the covariance matrix of $(s_0, s_1, \ldots, s_k)$, as the ratio between the determinants of the covariance matrix of order $(k+1) \times (k+1)$, and the covariance matrix of order $k \times k$.

Is it possible to calculate the volume of a type class that is defined by prescribed values of both the empirical autocorrelation and the correlation with a given $\boldsymbol{y}$? This turns out to be considerably harder (see the discussion in [116]) and it requires more advanced tools that will be provided in the next section.

2.7 Types Classes Pertaining to Exponential Families

So far, we have considered various kinds of Gaussian types, which are defined WRT given values of first and second order empirical statistics, like the empirical mean, the empirical second moment, the empirical correlation and autocorrelation, and so on. We now move on to extend the scope to deal with types associated with empirical moments or arbitrary functions. As described in Section 2.1, consider the type class of all sequences $\{\boldsymbol{x}\}$ that share the same combination of values of statistics $\frac{1}{n}\sum_{i=1}^{n} \phi_j(x_i)$, $j = 1, 2, \ldots, k$. More formally, consider the ϵ-inflated type class

$$\mathcal{T}_n(\boldsymbol{q}, \epsilon) \triangleq \left\{\boldsymbol{x}\colon \left|\frac{1}{n}\sum_{i=1}^{n} \phi_j(x_i) - q_j\right| \leq \epsilon,\ \forall\, 1 \leq j \leq k\right\}, \tag{2.125}$$

where $\boldsymbol{q} = (q_1, \ldots, q_k)$. What is the volume of $\mathcal{T}_n(\boldsymbol{q}, \epsilon)$? Using the same general idea as before, we seek a PDF of $\boldsymbol{x}$ which would assign to all members of $\mathcal{T}_n(\boldsymbol{q}, \epsilon)$ approximately the same PDF (in the exponential scale), and at the same time, the probability of $\mathcal{T}_n(\boldsymbol{q}, \epsilon)$ would be large for large n. As discussed in Section 2.1, consider the PDF

$$P_\theta(\boldsymbol{x}) = \frac{\exp\left\{\sum_{j=1}^{k} \theta_j \sum_{i=1}^{n} \phi_j(x_i)\right\}}{[Z(\theta)]^n}, \tag{2.126}$$

where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$ and

$$Z(\boldsymbol{\theta}) = \int_{\mathcal{X}} \exp\left\{\sum_{j=1}^{k} \theta_j \phi_j(x)\right\} \mathrm{d}x, \tag{2.127}$$

assuming that $\mathcal{X}$ is a continuous alphabet, and where it is understood that in the discrete case, the integration over $\mathcal{X}$ is replaced by summation. Clearly, $P_{\boldsymbol{\theta}}(\boldsymbol{x})$ assigns exponentially the same PDF to all members of $\mathcal{T}_n(\boldsymbol{q}, \epsilon)$, but $\mathcal{T}_n(\boldsymbol{q}, \epsilon)$ has high probability only if $\boldsymbol{\theta}$ is tuned accordingly for the given vector, $\boldsymbol{q}$. If we can select $\boldsymbol{\theta}$ such that

$$\mathbb{E}\left\{\phi_j(X)\right\} \equiv \frac{\partial \ln Z(\boldsymbol{\theta})}{\partial \theta_j} = q_j, \tag{2.128}$$

simultaneously for all $1 \leq j \leq k$, then by the WLLN, $\mathcal{T}_n(\boldsymbol{q}, \epsilon)$ would have high probability. Let us assume then, that $\boldsymbol{q}$ is such that there exists a parameter vector $\boldsymbol{\theta}$ that solves the set of k simultaneous equations (2.128), which can be presented in the vector form as

$$\nabla \ln Z(\boldsymbol{\theta}) = \boldsymbol{q}. \tag{2.129}$$

Let $\boldsymbol{\theta} = G(\boldsymbol{q})$ denote solution to this vector equation. In other words, $G(\boldsymbol{q})$ is the inverse mapping of $F(\boldsymbol{\theta}) \triangleq \nabla \ln Z(\boldsymbol{\theta})$, provided that it exists. The PDF of every $\boldsymbol{x} \in \mathcal{T}_n(\boldsymbol{q}, \epsilon)$ is exponentially

$$\frac{\exp\left\{n \sum_{j=1}^{k} \theta_j q_j\right\}}{[Z(\boldsymbol{\theta})]^n} = \frac{\exp\left\{n q^T G(\boldsymbol{q})\right\}}{[Z(G(\boldsymbol{q}))]^n}, \tag{2.130}$$

and so, the volume of $\mathcal{T}_n(\boldsymbol{q}, \epsilon)$ is of the exponential order of the reciprocal

$$\exp\left\{n\left[\ln Z(G(\boldsymbol{q})) - \boldsymbol{q}^T G(\boldsymbol{q})\right]\right\}. \tag{2.131}$$

Note that the (differential) entropy associated with $P_{\boldsymbol{\theta}}$ is given by

$$h[\boldsymbol{q}] = \mathbb{E}\left\{\ln \frac{1}{P_{\boldsymbol{\theta}}(X)}\right\} = \ln Z(\boldsymbol{\theta}) - \boldsymbol{q}^T \boldsymbol{\theta} = \ln Z(G(\boldsymbol{q})) - \boldsymbol{q}^T G(\boldsymbol{q}), \tag{2.132}$$

and so, once again, the volume is of the exponential order of $e^{nh[\boldsymbol{q}]}$.

It is interesting to relate the asymptotic evaluation of the log-volume of a type class to the *principle of maximum entropy* (see, *e.g.*, [36, Chapter 12] and references therein). We argue that $h[\boldsymbol{q}]$ is the largest possible differential entropy of any RV, X, that satisfies the moment constraints, $\mathbb{E}\{\phi_j(X)\} = q_j$, $j = 1, 2, \ldots, k$. To see why this is true, consider the following chain of equalities:

$$\sup_{\{X:\, \mathbb{E}\{\phi_j(X)\}=q_j,\ 1\le j\le k\}} h(X)$$

$$= \sup_X \inf_{\boldsymbol{\theta}} \left[h(X) + \sum_{j=1}^k \theta_j \left(\mathbb{E}\{\phi_j(X)\} - q_j\right) \right] \tag{2.133}$$

$$= \sup_f \inf_{\boldsymbol{\theta}} \int_{-\infty}^{\infty} \mathrm{d}x f(x) \left[\ln \frac{1}{f(x)} + \sum_{j=1}^k \theta_j \left(\phi_j(x) - q_j\right) \right] \tag{2.134}$$

$$= \sup_f \inf_{\boldsymbol{\theta}} \int_{-\infty}^{\infty} \mathrm{d}x f(x) \left[\ln \frac{\exp\left\{\sum_{j=1}^k \theta_j \phi_j(x)\right\}}{f(x)} - \sum_{j=1}^k \theta_j q_j \right] \tag{2.135}$$

$$= \sup_f \inf_{\boldsymbol{\theta}} \int_{-\infty}^{\infty} \mathrm{d}x f(x) \left[\ln \frac{P_{\boldsymbol{\theta}}(x) \cdot Z(\boldsymbol{\theta})}{f(x)} - \boldsymbol{q}^T \boldsymbol{\theta} \right] \tag{2.136}$$

$$\stackrel{(a)}{=} \inf_{\boldsymbol{\theta}} \sup_f \left\{ -D(f \| P_{\boldsymbol{\theta}}) + \ln Z(\boldsymbol{\theta}) - \boldsymbol{q}^T \boldsymbol{\theta} \right\} \tag{2.137}$$

$$= \inf_{\theta} \left\{ \ln Z(\boldsymbol{\theta}) - \boldsymbol{q}^T \boldsymbol{\theta} \right\} \tag{2.138}$$

$$\stackrel{(b)}{=} \ln Z(G(\boldsymbol{q})) - \boldsymbol{q}^T G(\boldsymbol{q}) \tag{2.139}$$

$$= h[\boldsymbol{q}], \tag{2.140}$$

where (a) follows from the minimax theorem and fact that the objective is convex in $\boldsymbol{\theta}$ and concave in f, and (b) follows from the fact that the minimizing $\boldsymbol{\theta}$ is $\boldsymbol{\theta}^* = G(\boldsymbol{q})$, which is obtained by equating to zero the gradient of the convex function $\ln Z(\boldsymbol{\theta}) - \boldsymbol{q}^T \boldsymbol{\theta}$. As can be seen, the maximizing f is exactly $P_{\boldsymbol{\theta}}$ with $\boldsymbol{\theta} = G(\boldsymbol{q})$.

Example 2.5. The volume of the "Laplacian type class," where $k = 1$ and $\phi_1(x) = |x|$ is exponentially

$$(2eq)^n = \exp(nh[q]), \tag{2.141}$$

where

$$h[q] = \ln(2eq) \tag{2.142}$$

is the differential entropy of a Laplacian RV with $\mathbb{E}\{|X|\} = q$. More generally, the "generalized Gaussian type class" is defined for $k = 1$ and $\phi_1(x) = |x|^m$ (for arbitrary $m > 0$), where the volume exponent is given by the differential entropy of the generalized Gaussian RV with $\mathbb{E}\{|X|^m\} = q$, which is given by

$$h[q] = \frac{1}{m} \ln\left(\frac{meq}{2c_m}\right), \tag{2.143}$$

where

$$c_m = \left[\frac{m}{2^{1+1/m}\Gamma(1/m)}\right]^m. \tag{2.144}$$

The method of types for exponential families is flexible enough to evaluate exponential rates of moments and probabilities of events defined WRT statistics that are different from the sufficient statistics of underlying PDF. Consider the following example.

Example 2.6. Suppose that $X_1, X_2, \ldots, X_n$ are IID, zero-mean, Gaussian RVs with variance σ^2 and we wish to assess the probability that $\sum_{i=1}^n |X_i| \geq nA$, where $A \geq \sqrt{\frac{2}{\pi}}\sigma$. In such a case, we may define type classes as above with $k = 2$, $\phi_1(x) = |x|$ and $\phi_2(x) = x^2$, where ϕ_1 is needed to support the statistics of the event in question, and ϕ_2 is for the underlying Gaussian PDF. Then, letting $\boldsymbol{q} = [q_1, q_2]$, each type class, $\mathcal{T}_n(q_1, q_2, \epsilon)$, contributes a probability of the exponential order of

$$e^{nh[q_1,q_2]} \cdot (2\pi\sigma^2)^{-n/2} e^{-nq_2/(2\sigma^2)}, \tag{2.145}$$

and so, the dominant type class contributes an exponential order of

$$\inf_{(q_1,q_2):\ q_1 \geq A,\ q_2 \geq q_1^2} \left\{\frac{q_2}{2\sigma^2} - h[q_1, q_2]\right\} + \frac{1}{2}\ln(2\pi\sigma^2), \tag{2.146}$$

where the constraint $q_2 \geq q_1^2$ follows from the inequality $\frac{1}{n}\sum_{i=1}^n x_i^2 \geq (\frac{1}{n}\sum_{i=1}^n |x_i|)^2$.

Finally, we point out that extension to conditional types and Markov types can be carried out conceptually straightforwardly following the same ideas described above in the context of Gaussian types. In both cases, the main tool is corresponding the exponential family, which is defined in (2.34) for Markov types and in (2.35) for conditional types.

2.8 Further Applications

The Gaussian method of types has found application in various contexts and levels of generality across prior research. In this section, we provide a brief overview of these contexts along with some of the outcomes achieved.

In [116], the challenge of universal decoding for memoryless Gaussian channels with unknown deterministic interference was tackled, and the method of Gaussian types played a central role in the analysis. As highlighted in [116, Eq. (5)], the universal decoding metric for the Gaussian channel hinges on the volume of the conditional Gaussian type class of a channel input vector $\boldsymbol{x}$, given a channel output vector $\boldsymbol{y}$. The effectiveness of this decoding metric is contingent on having an explicit formula for the exponential rate of this volume.

The extension from the memoryless case to Gaussian channels with intersymbol interference remained an open question after [116], as estimating the corresponding volume was non-trivial. The gap was eventually bridged in [88] and [90] using more advanced methodologies to be discussed later. A similar connection between universal decoding metrics and volumes of conditional type classes was observed in a broader context of universal decoding for arbitrary channels concerning a specific class of decoding metrics [124]. Additional insights can be found in [147, Section 4].

The method of Gaussian types has also played a pivotal role in deducing random coding exponents for typical random codes in distinctive scenarios. For instance, in the context of the colored Gaussian channel [138] and the dirty-paper channel [196], this method was crucial. Both studies relied on the concept of conditional type classes and their associated volumes, and the presence of explicit expressions played a vital role in achieving exponentially tight results.

In [9, Section IV.A], the method of Gaussian types found application in addressing the problem of optimal guessing subject to a fidelity constraint for memoryless Gaussian sources. This corresponds to a parallel result for finite-alphabet memoryless sources, for which the conventional method of types is employed. By leveraging outcomes pertaining to the exponential order of the volume of both simple Gaussian types and conditional Gaussian types, the optimal achievable guessing exponent was deduced. The crux of this derivation revolves around creating a continuous version of the type-covering lemma. This lemma establishes the capability to encompass a Euclidean hyper-sphere with a radius of $\sqrt{n\sigma^2}$ using exponentially $\exp\{\frac{n}{2}\ln\frac{\sigma^2}{D}\}$ Euclidean hyper-spheres, each with a radius of $\sqrt{nD}$, where $D < \sigma^2$. This type-covering result was reaffirmed and expanded to support successive refinement coding theorems in [232], also employing Gaussian types. Interestingly, it seems that the authors of [232] were unaware of the initial version of this result in [9]. Gaussian types were also harnessed by Kelly and Wagner in [102] concerning the reliability of source coding with side-information (the Wyner–Ziv problem [225]). Moreover, Scarlett [173] and Scarlett and Tan [178] employed Gaussian types (termed "power types") for second-order asymptotic analyses in their respective works. Similar methods were explored in [94] within the domain of compression for similarity queries. Additional related references include [198] and [210]. Furthermore, an analogous type-covering lemma for Laplacian type classes was established in [230] (also covered in [231]).

The method of types extended to general exponential families found application in [115] within the domain of model order estimation. Just as mentioned previously, in this context as well, the existence of an expression for the volume of a type class played a pivotal role in deducing the model order selection criterion. Additionally, in [146], the method of types was employed for exponential families within the context of a continuous-alphabet extension of widely recognized lower bounds for mismatched capacity, utilizing random coding analysis. This showcases the versatility of the method across diverse problem domains.

3

The Laplace Method of Integration and the Saddle-Point Method

3.1 Introduction

The Laplace method of integration (see, *e.g.*, [23, Chapter 4], [120, Section 4.2]) is a powerful technique for approximating definite integrals of the form:

$$\int_a^b g(x) e^{nf(x)} \, \mathrm{d}x, \tag{3.1}$$

where the parameter n is large ($n \gg 1$), and the functions f and g exhibit sufficient regularity WRT the real variable x. Importantly, these functions are assumed to remain independent of n. More generally, x may designate a d-dimensional vector, where d is independent of the large parameter n, whereas the integration occurs over $\mathbb{R}^d$ or a subset thereof.

The significance of this method is twofold. Firstly, it offers intrinsic utility by itself, providing accurate asymptotic approximations for integrals. However, its greater importance lies in its role as the foundation for the saddle-point method, an extension that applies the principles of the Laplace method to the integration of complex functions along contours within the complex plane. The saddle-point method finds broad applications across diverse disciplines, including physics, probability, statistics, and engineering. Notably, this section emphasizes that the method holds promise in information theory as well. In many instances, the saddle-point method can serve as a viable alternative to the extended method of types discussed in Section 2. This advantage becomes particularly apparent when it comes to circumventing the need for ϵ-inflation of type classes, a strategy employed in Section 2. The Laplace method and the saddle-point method offer a distinct advantage by not only yielding the accurate exponential rate, as demonstrated in Section 2, but also by providing the correct pre-exponential term. Remarkably, this method furnishes approximations that exhibit asymptotic precision. Specifically, as the large parameter n grows without bound, the ratio between the approximation and the actual value converges to unity, signifying an increasingly faithful representation of the underlying quantity.

It is important to note that the content presented in this section exhibits some overlap with the material found in [120, Sections 4.2 and 4.3] and in [23, Chapters 4 and 5]. As a result, several intricate technical aspects related to the Laplace method and the saddle-point method are either succinctly addressed or occasionally omitted (though appropriately cross-referenced to [23], [120]). These intricate technicalities are associated with the assessment of the approximation error terms pertaining to these methods and regularity conditions. Instead, the focus here lies on considering these methods in the context of their capacity to stand as valid alternatives to the generalized method of types, as described in Section 2. This pertains to both its discrete and continuous alphabet variations. For readers seeking a more comprehensive treatment with meticulous attention to detail and rigor, we recommend delving into the pertinent chapters of [23] and [120].

3.2 The Laplace Method of Integration

Commencing with the Laplace method, we turn our attention to an illustrative example tied to the domain of universal source coding (as expounded in references such as [46] and [36, Section 13.2]). This example serves as a compelling application that underscores the significance of the Laplace method within information theory.

Example 3.1 (Universal coding)**.** Consider a family of binary memoryless (Bernoulli) sources defined over the alphabet $\{0,1\}$, parameterized by $\theta \in [0,1]$, which represents the probability of emitting a $'1'$. The probability mass function of this source is given by:

$$P_\theta(\boldsymbol{x}) = (1-\theta)^{n-n_1}\theta^{n_1}, \tag{3.2}$$

where $\boldsymbol{x} \in \{0,1\}^n$, and $n_1 \leq n$ is the count of occurrences of $'1'$ in $\boldsymbol{x}$. When dealing with an unknown θ, a universal code is devised using the Shannon code,[1] designed for the weighted mixture of these sources [46]:

$$P(\boldsymbol{x}) = \int_0^1 \mathrm{d}\theta \cdot w(\theta) P_\theta(\boldsymbol{x}) = \int_0^1 \mathrm{d}\theta \cdot w(\theta) e^{nf(\theta)}, \tag{3.3}$$

where $w(\cdot)$ is a positive function that integrates to unity across the interval $[0,1]$, and

$$f(\theta) = \ln(1-\theta) + p \ln\left(\frac{\theta}{1-\theta}\right), \tag{3.4}$$

with $p \triangleq \frac{n_1}{n}$. This necessitates the computation of an integral involving an exponential function of n (in this case, across the interval $[0,1]$) to evaluate the performance of this universal code. An asymptotically exact evaluation of such an integral is crucial in the quest of characterizing, not only the main term of the achievable compression ratio, but also the redundancy terms (see Example 3.2 below).

Consider first an integral of the form:

$$F_n \triangleq \int_{-\infty}^{+\infty} e^{nf(x)}\, \mathrm{d}x, \tag{3.5}$$

where the function $f(\cdot)$ is independent of n. It will be assumed that the function f satisfies the following assumptions:

1. f is real and continuous.
2. f has a unique global maximum at $x = x_0$: $f(x) < f(x_0)\ \ \forall x \neq x_0$, and $\exists b > 0,\ c > 0$ such that $|x - x_0| \geq c$ implies $f(x) \leq f(x_0) - b$.

[1] The Shannon code for lossless source coding for a distribution $P(\boldsymbol{x})$ is a variable length code whose length function is given by $\lceil -\log_2 P(\boldsymbol{x}) \rceil$ bits [36, Section 5.9].

3. The integral defining F_n converges for all large enough n. Without essential loss of generality, let this sufficiently large n be $n = 1$, *i.e.*, $\int_{-\infty}^{+\infty} e^{f(x)}\, \mathrm{d}x < \infty$.

4. The derivative $f'(x)$ exists at a certain open neighborhood of $x = x_0$, and $f''(x_0) < 0$. Thus, $f'(x_0) = 0$.

These assumptions pave the way to approximate $f(x)$, at the vicinity of $x = x_0$, by a second-order Taylor series expansion,

$$f(x) \approx f(x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 = f(x_0) - \frac{|f''(x_0)|}{2}(x - x_0)^2, \quad (3.6)$$

which renders F_n as being dominated by the constant $e^{nf(x_0)}$, multiplied by a Gaussian integral, namely, the integral of $\exp\{-\frac{n}{2}|f''(x_0)|(x-x_0)^2\}$, whereas the combined contribution of all the range away from x_0 is negligibly small for large n. Accordingly, as shown in [23, Chapter 4] and [120, Section 4.2], we arrive at the Laplace method approximation, given by

$$\int_{-\infty}^{+\infty} e^{nf(x)}\, \mathrm{d}x \sim e^{nf(x_0)} \cdot \sqrt{\frac{2\pi}{n|f''(x_0)|}}, \quad (3.7)$$

where the pre-exponential factor $\sqrt{\frac{2\pi}{n|f''(x_0)|}}$ is found by integrating a Gaussian PDF of variance $1/[n|f''(x_0)|]$ to 1, *i.e.*,

$$\sqrt{\frac{n|f''(x_0)|}{2\pi}} \int_{-\infty}^{+\infty} \exp\left[-\frac{n|f''(x_0)|}{2}(x - x_0)^2\right] \mathrm{d}x = 1. \quad (3.8)$$

This approximation continues to apply if F_n is defined as an integral over any finite or half-infinite interval that contains the maximizer $x = x_0$ as an internal point. On the other hand, if the maximizer x_0 falls at one of the endpoints of the integration range, and $f'(x_0)$ does not vanish, the Gaussian integral approximation ceases to apply, and the local behavior around the maximum would be approximated by an exponential $\exp\{-n|f'(x_0)|(x - x_0)\}$ instead, which gives a different pre-exponential factor, yet the exponential factor $e^{nf(x_0)}$ would continue to be present. Specifically, consider, for example, the integral

$$F_n = \int_{x_0}^{\infty} e^{nf(x)}\, \mathrm{d}x \quad (3.9)$$

where the maximum of $f(x)$ across the range $[x_0, \infty)$ is attained at $x = x_0$, and where $f'(x_0) < 0$. Then, F_n is approximated by a first-order Taylor series expansion, according to

$$F_n \sim \int_{x_0}^{\infty} \exp\left\{n\left[f(x_0) + f'(x_0)(x - x_0)\right]\right\} \mathrm{d}x \tag{3.10}$$

$$= \int_{x_0}^{\infty} \exp\left\{n\left[f(x_0) - \left|f'(x_0)\right|(x - x_0)\right]\right\} \mathrm{d}x \tag{3.11}$$

$$= \frac{e^{nf(x_0)}}{n|f'(x_0)|}. \tag{3.12}$$

Returning to the case when x_0 is an internal point of the integration range and both $g(x_0) > 0$ and $f'(x_0) = 0$, a further extension is the following:

$$\int_{-\infty}^{+\infty} g(x)e^{nf(x)}\, \mathrm{d}x \sim g(x_0)e^{nf(x_0)} \cdot \sqrt{\frac{2\pi}{n|f''(x_0)|}}, \tag{3.13}$$

where g is a function that does not depend on n.

The Laplace method also has an extension to the case where the integration variable x represents a d-dimensional vector, where d is a positive integer that does not grow with n. The integration now takes place over $\mathbb{R}^d$ or a subset thereof, with x_0 positioned as an internal point within the integration region. In this case, the Gaussian approximation becomes multi-dimensional too, and so we must replace $|f''(x_0)|$ in both (3.7) and (3.13) with the absolute value of the determinant of the Hessian matrix $\nabla^2 f(x_0)$ of f evaluated at $x = x_0$. Additionally, the factor n that multiplies $|f''(x_0)|$ should be substituted with n^d. This adjustment arises from a corresponding approximation involving a multi-dimensional Gaussian integral. Overall, the approximation takes the form

$$\int_{\mathbb{R}^d} g(x)e^{nf(x)}\, \mathrm{d}x \sim \left(\frac{2\pi}{n|\nabla^2 f(x_0)|}\right)^{d/2} g(x_0)e^{nf(x_0)}, \tag{3.14}$$

where $|\nabla^2 f(x_0)|$ is the determinant of the Hessian of f at $x = x_0$. If the global maximum of f is achieved by more than one point, and the number of maximizers is finite or countable, then the contributions from all of these maximizers should be aggregated or summed together. See [120, Section 4.6] for additional details.

3.3 Examples of the Laplace Method

We next demonstrate the use of Laplace method in a few examples.

Example 3.2 (Universal coding revisited). Applying the Laplace integral approximation to Example 3.1, we have

$$P(\boldsymbol{x}) = \int_0^1 w(\theta) \exp\left\{n\left[\ln(1-\theta) + p\ln\left(\frac{\theta}{1-\theta}\right)\right]\right\} \mathrm{d}\theta \tag{3.15}$$

$$\sim w(p)e^{-nH(p)}\sqrt{\frac{2\pi p(1-p)}{n}}, \tag{3.16}$$

where we recall that $p = \frac{n_1}{n}$ and $H(p) \triangleq -p\ln p - (1-p)\ln(1-p)$ is thus the empirical entropy of $\boldsymbol{x}$. So, the compression ratio corresponding to the Shannon code WRT the mixture is

$$\frac{-\ln P(\boldsymbol{x})}{n} = H(p) + \frac{\ln n}{2n} - \frac{\ln\left[w(p)\sqrt{2\pi p(1-p)}\right]}{n} + o\left(\frac{1}{n}\right). \tag{3.17}$$

The principal component of the normalized redundancy can be expressed as $\frac{\ln n}{2n}$, a well-established result (for more details, refer to [108]). Similarly, when considering a mixture encompassing all sources with an alphabet size of r, this entails integration over $r-1$ letter probabilities, resulting in a dominant redundancy term of $\frac{(r-1)\ln n}{2n}$.

Example 3.3 (Extreme Value Statistics). Consider a set of non-negative, IID RVs $\{X_i\}_{i=1}^n$, each characterized by the PDF $p(x)$. Our goal is to evaluate the expectation of the minimum value among these variables, $\mathbb{E}\{\min_{1\le i\le n} X_i\}$. Let us explore the following sequence of equalities to facilitate this assessment. Denoting the cumulative distribution function of each X_i by $F(x)$, we have

$$\mathbb{E}\left\{\min_{1\le i\le n} X_i\right\} \stackrel{(a)}{=} \int_0^\infty \Pr\left\{\min_{1\le i\le n} X_i \ge x\right\} \mathrm{d}x \tag{3.18}$$

$$= \int_0^\infty \Pr\left[\bigcap_{i=1}^n \{X_i \ge x\}\right] \mathrm{d}x \tag{3.19}$$

$$\stackrel{(b)}{=} \int_0^\infty [1-F(x)]^n \,\mathrm{d}x \tag{3.20}$$

$$= \int_0^\infty \exp\{n\ln[1-F(x)]\} \,\mathrm{d}x, \tag{3.21}$$

where (a) follows from the integral identity $\mathbb{E}\{X\} = \int_0^\infty \Pr\{X \geq t\}\ \mathrm{d}t$, which holds for any non-negative RV,[2] and (b) due to the IID assumption. Hence, we may use the Laplace method with $f(x) = \ln[1 - F(x)]$. Let us first assume that $p(0) > 0$. Then, the maximum of $f(x)$ is obtained at the edge-point of the integration domain, $x_0 = 0$ and $f'(0) = -p(0) < 0$. Therefore, the approximation is not by a Gaussian integral, but a simple exponential,

$$\int_0^\infty \exp\{-n|f'(0)|x\}\ \mathrm{d}x = \frac{1}{n|f'(0)|}, \tag{3.22}$$

which yields

$$\mathbb{E}\left\{\min_{1\leq i\leq n} X_i\right\} \sim \frac{1}{np(0)}. \tag{3.23}$$

However, if $p(0) = 0$ while $p'(0) > 0$, the Laplace approximation is executed through a Gaussian integral over half of the real line. In such a scenario, the outcome is as follows:

$$\mathbb{E}\left\{\min_{1\leq i\leq n} X_i\right\} \sim \frac{1}{2}\sqrt{\frac{2\pi}{np'(0)}}. \tag{3.24}$$

The last example in this section supports the Stirling approximation.

Example 3.4 (The Stirling formula)**.** Beginning from the identity

$$\int_0^\infty \mathrm{d}x \cdot e^{-sx} \tag{3.25}$$

and differentiating both sides n times WRT s, the left-hand side becomes $(-1)^n \int_0^\infty x^n e^{-sx}\ \mathrm{d}x$, and the right-hand side (RHS) gives $(-1)^n n!/s^{n+1}$, which together yield the identity

$$n! = s^{n+1} \int_0^\infty x^n e^{-sx}\ \mathrm{d}x, \tag{3.26}$$

holding true for every $s > 0$. On substituting $s = n$, we get

$$n! = n^{n+1} \int_0^\infty x^n e^{-nx}\ \mathrm{d}x = n^{n+1} \int_0^\infty e^{n(\ln x - x)}\ \mathrm{d}x. \tag{3.27}$$

Assessing this integral using the Laplace method, we have $f(x) = \ln x - x$, which is maximized at $x_0 = 1$, with $f(x_0) = f''(x_0) = -1$. Thus,

$$n! \sim n^{n+1} e^{-n\cdot 1}\sqrt{\frac{2\pi}{n\cdot 1}} = \left(\frac{n}{e}\right)^n \sqrt{2\pi n}, \tag{3.28}$$

which is the well-known Stirling formula for approximating $n!$.

[2]We will delve into this and other integral identities in Section 5.

3.4 The Saddle-Point Method

We now broaden our focus to encompass integrals along paths within the complex plane, a concept that arises more frequently than one might anticipate. As previously mentioned, the extension of the Laplace integration technique to complex functions is referred to as the saddle-point method or the steepest descent method, with explanations for these names becoming apparent in the forthcoming presentation. Specifically, our current interest lies in evaluating an integral represented as follows:

$$F_n = \int_{\mathcal{P}} g(z)e^{nf(z)} \, \mathrm{d}z. \tag{3.29}$$

In this context, the variable z takes on complex values and $\mathcal{P}$ designates a certain path within the complex plane, originating from a point A and concluding at a point B. Our initial focus will be on the case $g(z) \equiv 1$, and we make the assumption that $\mathcal{P}$ exclusively lies within a region where the function f is analytic (see, *e.g.*, [159]).

At first glance, the reader might question the relevance of complex integrals when dealing with quantities that are inherently real — such as probabilities, expectations, volumes of high-dimensional objects, and more. The answer lies in the fact that even if these quantities are real, there are instances where expressing a certain term in a calculation as an inverse Fourier transform or an inverse Laplace transform, or inverse Z-transform, becomes useful and beneficial. These inverse transforms are represented through complex integrals. To illustrate, consider the following straightforward example: Computing the volume of an n-dimensional hyper-sphere with radius R. This task can be approached by interpreting the volume as the integral of $U(R^2 - \sum_{i=1}^{n} x_i^2)$ over $\mathbb{R}^n$, where $U(t)$ signifies the Heaviside unit step function. Next, we express $U(t)$ as the inverse Laplace transform of $1/s$, subsequently we interchange the integration order, and finally, we apply the saddle-point method to evaluate the complex integration. As we proceed, we will delve into the detailed execution of this concept.

The first observation of significance is that the integral's value depends solely upon the endpoints, A and B, regardless of the of the particular path $\mathcal{P}$. To illustrate, let us consider an alternative path denoted as $\mathcal{P}'$, connecting points A and B, while ensuring that the function f remains free of singularities within the enclosed region formed by $\mathcal{P} \cup \mathcal{P}'$. Under these conditions, the integral of $e^{nf(z)}$ across the closed path encompassing both $\mathcal{P}$ and $\mathcal{P}'$ — traversing from A to B via $\mathcal{P}$ and then returning from B to A through $\mathcal{P}'$ — vanishes, indicating that the integrals along $\mathcal{P}$ and $\mathcal{P}'$ between A and B hold identical values. In essence, this imparts us with the liberty to select our preferred integration path, so long as we exercise caution to avoid traversing too closely to the opposing side of any potential singularity point. This consideration gains significance as we proceed with our upcoming analyses.

Another fundamental key property of analytic complex functions is the *maximum-modulus theorem*. This theorem essentially states that the magnitude of an analytic function lacks any maxima. Although a comprehensive proof of this theorem is beyond our scope, its essence can be captured as follows: Consider an analytic function expressed as:

$$f(z) = u(z) + jv(z) = u(x, y) + jv(x, y), \tag{3.30}$$

where u and v are real-valued functions. When f is analytic, the Cauchy–Riemann conditions [86, Section 4.3] must hold for the partial derivatives of u and v:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}; \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}. \tag{3.31}$$

Taking the second-order partial derivative of u, we arrive at:

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 v}{\partial x \partial y} = \frac{\partial^2 v}{\partial y \partial x} = -\frac{\partial^2 u}{\partial y^2}, \tag{3.32}$$

where the first and third equalities stem from the Cauchy–Riemann conditions. Alternatively, we can write:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \tag{3.33}$$

which is recognized as the Laplace equation. Consequently, any point where $\frac{\partial u}{\partial x} = \frac{\partial u}{\partial y} = 0$ cannot be a local maximum or minimum of u. If it were a local maximum along the x-direction, then $\frac{\partial^2 u}{\partial x^2} < 0$, implying that $\frac{\partial^2 u}{\partial y^2}$ must be positive, making it a local minimum along the y-direction, and vice versa. Put simply, points where partial derivatives of u are zero are, in fact, saddle points. This line of reasoning applies to the modulus of the integrand $e^{nf(z)}$ due to:

$$\left|\exp\{nf(z)\}\right| = \exp\left[n\text{Re}\{f(z)\}\right] = e^{nu(z)}. \tag{3.34}$$

Furthermore, if $f'(z) = 0$ at some $z = z_0$, then $u'(z_0) = 0$ as well, establishing that z_0 is a saddle point of $|e^{nf(z)}|$. Thus, points where f exhibits zero derivatives are saddle points.

Armed with this foundational understanding, let us return to our integral F_n in (3.29). Given the flexibility to select the path $\mathcal{P}$, suppose we can identify a trajectory that crosses a saddle point z_0 (hence the name of the method), and where the maximum value of $|e^{nf(z)}|$ along $\mathcal{P}$ is achieved at $z = z_0$. In this scenario, much like in the Laplace method, we anticipate that the integral's dominant contribution would stem from $e^{nf(z_0)}$. Naturally, this path would be suitable only if it intersects the saddle point z_0 along a direction WRT which z_0 represents a local maximum of $|e^{nf(z)}|$ or equivalently, of $u(z)$. Moreover, for the application of our prior Laplace method findings, we aim to configure $\mathcal{P}$ so that any point z in proximity to z_0, where the Taylor expansion reads (due to the fact that $f'(z_0) = 0$):

$$f(z) = f(z_0) + \frac{1}{2}f''(z_0)(z - z_0)^2 + o\left(|z - z_0|^2\right), \tag{3.35}$$

where the second term, $\frac{1}{2}f''(z_0)(z-z_0)^2$, is exclusively real and negative, and where the exact form of the error term $o(|z-z_0|^2)$ can be determined from f. More precisely, it can be shown by the mean value theorem that for every $\epsilon > 0$ there exists $\delta > 0$ such that $|z - z_0| \leq \delta$ implies

$$\left|f(z) - f(z_0) - \frac{f''(z_0)}{2} \cdot (z - z_0)^2\right| \leq \epsilon|z - z_0|^2, \tag{3.36}$$

and the correspondence between δ and ϵ depends on the particular smoothness properties of the function f. Consequently, it assumes a local behavior akin to a negative parabola, mirroring the behavior observed in the Laplace method. This implication manifests itself in the phase of $\frac{1}{2}f''(z_0)(z - z_0)^2$, given by:

$$\arg\left\{f''(z_0)\right\} + 2\arg(z - z_0) = \pi, \tag{3.37}$$

or equivalently:

$$\arg(z - z_0) = \frac{\pi - \arg\{f''(z_0)\}}{2} \triangleq \theta. \tag{3.38}$$

In essence, $\mathcal{P}$ should traverse z_0 along the direction θ. This orientation is called the *axis* of z_0 and can be demonstrated to be the direction of steepest descent from the summit at z_0 — hence the name *steepest-descent method*. Notably, it is worth mentioning that in the $\theta - \pi/2$ direction, which stands perpendicular to the axis, $\arg[f''(z_0)(z-z_0)^2] = \pi - \pi = 0$. Consequently, $f''(z_0)(z - z_0)^2$ emerges as real and positive in this direction, akin to a positive parabolic pattern. This indicates that along this direction, z_0 constitutes a local minimum.

Visually speaking, our strategy involves the selection of a path $\mathcal{P}$ connecting A to B, constructed as three distinct segments (as depicted in Figure 3.1): $A \to A'$ and $B' \to B$ form the arbitrary initial and final sections of the integral path. The middle part, connecting A' to B' and localized near z_0, consists of a straight line aligned with the axis of z_0.

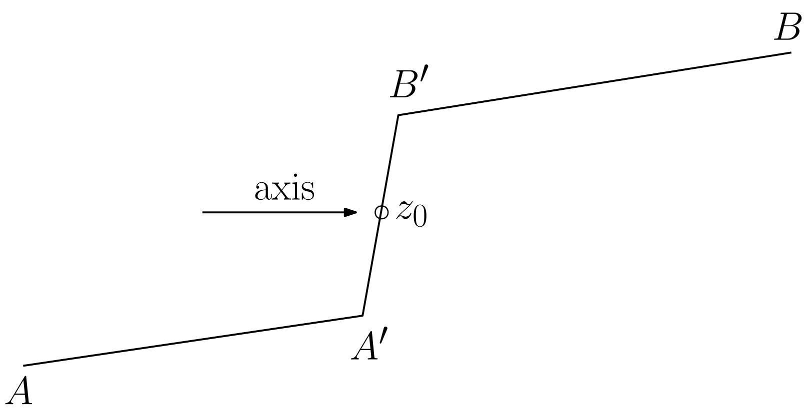

Figure 3.1: A path $\mathcal{P}$ from A to B, passing via z_0 along the axis.

Accordingly, let us decompose F_n into its three parts:

$$F_n = \int_A^{A'} e^{nf(z)}\, dz + \int_{A'}^{B'} e^{nf(z)}\, dz + \int_{B'}^{B} e^{nf(z)}\, dz. \tag{3.39}$$

As for the first and the third terms,

$$\left| \int_A^{A'} e^{nf(z)}\, dz + \int_{B'}^{B} e^{nf(z)}\, dz \right|$$

$$\leq \int_{A}^{A'} |e^{nf(z)}|\ \mathrm{d}z + \int_{B'}^{B} |e^{nf(z)}|\ \mathrm{d}z \tag{3.40}$$

$$= \int_{A}^{A'} e^{n\mathrm{Re}\{f(z)\}}\ \mathrm{d}z + \int_{B'}^{B} e^{n\mathrm{Re}\{f(z)\}}\ \mathrm{d}z, \tag{3.41}$$

whose contribution is negligible compared to $e^{nf(z_0)}$, exactly like the tails in the Laplace method. As for the middle integral,

$$\int_{A'}^{B'} e^{nf(z)}\ \mathrm{d}z \sim e^{nf(z_0)} \int_{A'}^{B'} \exp\left\{\frac{nf''(z_0)(z-z_0)^2}{2}\right\} \mathrm{d}z. \tag{3.42}$$

By transitioning from the complex integration variable z to the real variable x, ranging from $-\delta$ to $+\delta$, with $z = z_0 + xe^{j\theta}$ (following the axis direction), we end up with exactly the Gaussian integral encountered in the Laplace method, resulting in:

$$\int_{A'}^{B'} \exp\{nf''(z_0)(z-z_0)^2/2\}\ \mathrm{d}z = e^{j\theta}\sqrt{\frac{2\pi}{n|f''(z_0)|}} \tag{3.43}$$

where the factor $e^{j\theta}$ is due to the change of variable ($\mathrm{d}z = e^{j\theta}\mathrm{d}x$). Thus,

$$F_n \sim e^{j\theta} \cdot e^{nf(z_0)}\sqrt{\frac{2\pi}{n|f''(z_0)|}}, \tag{3.44}$$

and somewhat more generally,

$$\int_{\mathcal{P}} g(z)e^{nf(z)}\ \mathrm{d}z \sim e^{j\theta}g(z_0)e^{nf(z_0)}\sqrt{\frac{2\pi}{n|f''(z_0)|}}. \tag{3.45}$$

3.5 Examples of the Saddle-Point Method

We next demonstrate the use of the saddle-point method in a few examples.

Example 3.5 (The size of a type class of binary sequences). To count the number of binary sequences of length n with exactly k 1's and $(n-k)$ 0's, we use the notation m_k. Let us examine the complex function

$$M(z) = (1+z^{-1})^n = \sum_{k=0}^{n} m_k z^{-k}. \tag{3.46}$$

The second equality expresses the fact that $M(z)$ can be viewed as the Z-transform of the sequence $\{m_k\}_{k=0}^n$, and so, m_k is given by the inverse Z-transform of $M(z)$:

$$m_k = \frac{1}{2\pi j}\oint_{\mathcal{P}}(1+z^{-1})^n z^{k-1}\mathrm{d}z \tag{3.47}$$

$$= \frac{1}{2\pi j}\oint_{\mathcal{P}}\frac{1}{z}\exp\left\{n\left[\ln(1+z^{-1})+q\ln z\right]\right\}\mathrm{d}z, \tag{3.48}$$

where $q = k/n$ and $\oint_{\mathcal{P}}$ denotes integration along an arbitrary counter-clockwise closed path $\mathcal{P}$ that surrounds the origin. Here, $g(z) = 1/z$ and

$$f(z) = \ln(1+z^{-1}) + q\ln z = \ln(1+z) - (1-q)\ln z, \tag{3.49}$$

whose saddle point is $z_0 = \frac{1-q}{q}$. If we choose $\mathcal{P}$ to be the circle $|z| = \frac{1-q}{q}$, it intersects the point z_0, situated on the real line, in a vertical manner. Remarkably, this alignment corresponds to the axis of z_0. A straightforward calculation yields

$$f''(z_0) = \frac{q^3}{1-q} \tag{3.50}$$

which gives

$$m_k \sim \frac{e^{j\pi/2}}{z_0}\cdot e^{nf(z_0)}\cdot\frac{1}{2\pi j}\cdot\sqrt{\frac{2\pi}{n|f'(z_0)|}} \tag{3.51}$$

$$= \frac{e^{j\pi/2}}{(1-q)/q}\cdot e^{nH(q)}\cdot\frac{1}{2\pi j}\cdot\sqrt{\frac{2\pi(1-q)}{nq^3}} \tag{3.52}$$

$$= \frac{e^{nH(q)}}{\sqrt{2\pi nq(1-q)}}, \tag{3.53}$$

where, as before, $H(q) \triangleq -q\ln q - (1-q)\ln(1-q)$ is the binary entropy function.

Our next example addresses continuous alphabets.

Example 3.6 (Surface area of a hyper-sphere). This example is closely connected to the concept of simple Gaussian-type classes, as discussed in Section 2. While there exists an exact closed-form expression for the surface area of an n-dimensional Euclidean hyper-sphere [see (2.44)], we explore this example to illustrate the asymptotic accuracy of the saddle-point method. Our starting point is the representation of the surface area of an n-dimensional Euclidean hyper-sphere with radius r as follows:

$$S_n(r) = 2r \int_{\mathbb{R}^n} \delta\left(r^2 - \sum_{i=1}^n x_i^2\right) \mathrm{d}\boldsymbol{x}, \tag{3.54}$$

where $\delta(\cdot)$ designates the Dirac delta function. To see why this true, observe that $S_n(r)$ integrates to

$$\begin{aligned}
V_n(R) &= \int_0^R S_n(r)\ \mathrm{d}r && (3.55)\\
&= \int_0^R 2r \int_{\mathbb{R}^n} \delta\left(r^2 - \sum_{i=1}^n x_i^2\right) \mathrm{d}\boldsymbol{x}\mathrm{d}r && (3.56)\\
&= \int_{\mathbb{R}^n} \left[\int_0^R 2r\delta\left(r^2 - \sum_{i=1}^n x_i^2\right) \mathrm{d}r\right] \mathrm{d}\boldsymbol{x} && (3.57)\\
&= \int_{\mathbb{R}^n} \left[\int_0^{R^2} \delta\left(r^2 - \sum_{i=1}^n x_i^2\right) \mathrm{d}(r^2)\right] \mathrm{d}\boldsymbol{x} && (3.58)\\
&= \int_{\mathbb{R}^n} U\left(R^2 - \sum_{i=1}^n x_i^2\right) \mathrm{d}\boldsymbol{x} && (3.59)\\
&= \mathrm{Vol}\left\{\boldsymbol{x}\colon \sum_{i=1}^n x_i^2 \leq R^2\right\}, && (3.60)
\end{aligned}$$

where $U(\cdot)$ is the unit step function. Thus, the integral of $S_n(r)$ across the interval $[0, R]$ yields the volume of a hyper-sphere of radius R, and so, $S_n(r)$ is the surface area of a hyper-sphere of radius r. We next represent the Dirac delta function as the inverse Fourier transform of the unit function, *i.e.*,

$$\delta(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega t}\ \mathrm{d}\omega, \tag{3.61}$$

and so, referring to Section 2, the surface area of sphere of radius $\sqrt{ns}$ is given as follows. Let $\vartheta > 0$ be some positive real, to be chosen shortly. Then,

$$
\begin{aligned}
& S_n(\sqrt{ns}) \\
&= 2\sqrt{ns}\int_{\mathbb{R}^n} \mathrm{d}\boldsymbol{x}\cdot\delta\left(ns-\sum_{i=1}^n x_i^2\right) && (3.62)\\
&\stackrel{(a)}{=} 2\sqrt{ns}e^{n\vartheta s}\int_{\mathbb{R}^n} \mathrm{d}\boldsymbol{x}\cdot\exp\left\{-\vartheta\sum_{i=1}^n x_i^2\right\}\cdot\delta\left(ns-\sum_{i=1}^n x_i^2\right) && (3.63)\\
&= 2\sqrt{ns}e^{n\vartheta s}\int_{\mathbb{R}^n} \mathrm{d}\boldsymbol{x}\cdot\exp\left\{-\vartheta\sum_{i=1}^n x_i^2\right\}\times \\
&\qquad \int_{-\infty}^{+\infty}\frac{\mathrm{d}\omega}{2\pi}\cdot\exp\left\{j\omega\left(ns-\sum_{i=1}^n x_i^2\right)\right\} && (3.64)\\
&= \sqrt{ns}e^{n\vartheta s}\int_{-\infty}^{+\infty}\frac{\mathrm{d}\omega}{\pi}\cdot e^{j\omega ns}\int_{\mathbb{R}^n}\mathrm{d}\boldsymbol{x}\cdot\exp\left\{-(\vartheta+j\omega)\sum_{i=1}^n x_i^2\right\} && (3.65)\\
&= \sqrt{ns}e^{n\vartheta s}\int_{-\infty}^{+\infty}\frac{\mathrm{d}\omega}{\pi}\cdot e^{j\omega ns}\left[\int_{\mathbb{R}}\mathrm{d}x\cdot e^{-(\vartheta+j\omega)x^2}\right]^n && (3.66)\\
&\stackrel{(b)}{=} \sqrt{ns}e^{n\vartheta s}\int_{-\infty}^{+\infty}\frac{\mathrm{d}\omega}{\pi}\cdot e^{j\omega ns}\left(\frac{\pi}{\vartheta+j\omega}\right)^{n/2} && (3.67)\\
&= \sqrt{ns}\pi^{n/2-1}\int_{-\infty}^{+\infty}\mathrm{d}\omega\cdot\exp\left\{n\left[(\vartheta+j\omega)s-\frac{1}{2}\ln(\vartheta+j\omega)\right]\right\} && (3.68)\\
&= \sqrt{ns}\cdot\pi^{n/2-1}\cdot\frac{1}{j}\cdot\int_{\vartheta-j\infty}^{\vartheta+j\infty}\mathrm{d}z\cdot\exp\left\{n\left[zs-\frac{1}{2}\ln z\right]\right\}, && (3.69)
\end{aligned}
$$

where in (a) we have multiplied the expression by $e^{n\vartheta s}$ outside the integral and by $e^{-\vartheta\sum_i x^2}$ inside the integral, but $e^{-\vartheta\sum_i x^2} = e^{-n\vartheta s}$ wherever the delta function of the integrand does not vanish, and so, these two multiplications cancel each other. This step is crucial for the subsequent steps. In (b) we have applied complex Gaussian integration. In this case, we have

$$f(z) = zs - \frac{1}{2}\ln z, \tag{3.70}$$

and the integration is along an arbitrary vertical straight line $\mathrm{Re}\{z\} = \vartheta$. We select this straight line to cross the saddle-point, that is, $\vartheta = z_0 = \frac{1}{2s}$, where

$$f(z_0) = \frac{1}{2}\ln(2es) \tag{3.71}$$

and

$$f''(z_0) = 2s^2. \tag{3.72}$$

Once again, the axis is vertical, and so,

$$S_n(\sqrt{ns}) \sim \sqrt{ns} \cdot \pi^{n/2-1} \cdot \frac{1}{j} \cdot e^{j\pi/2} \cdot \exp\left\{\frac{n}{2}\ln(2es)\right\} \cdot \sqrt{\frac{2\pi}{2s^2 n}} \quad (3.73)$$

$$= \frac{(2\pi es)^{n/2}}{\sqrt{\pi s}}, \quad (3.74)$$

which agrees with (2.44) from Section 2. Note that the representation of $\delta\left(ns - \sum_{i=1}^n x_i^2\right)$ as an inverse Fourier transform converted the integrand into an exponential function of $(ns - \sum_{i=1}^n x_i^2)$, which is a product form and hence can be represented as a product of identical integrals, which is actually one-dimensional integral raised to the power of n.

In the above derivation, when we shifted the vertical integration path from the imaginary axis, $\{z\colon \text{Re}\{z\} = 0\}$, to the parallel vertical line $\{z\colon \text{Re}\{z\} = \vartheta\}$, we have actually replaced the inverse Fourier transform by the inverse Laplace transform. By the same token, we can handle the volume of the n-dimensional hyper-sphere as

$$V_n(ns) = \int_{\mathbb{R}^n} U\left(ns - \sum_{i=1}^n x_i^2\right) \mathrm{d}\boldsymbol{x} \quad (3.75)$$

with the representation of the unit step function as the inverse Laplace transform of $1/z$, which amounts to substituting

$$U\left(ns - \sum_{i=1}^n x_i^2\right) = \frac{1}{2\pi j}\int_{\text{Re}\{z\}=\vartheta} \frac{\mathrm{d}z}{z} \cdot \exp\left\{z\left(ns - \sum_{i=1}^n x_i^2\right)\right\}, \quad (3.76)$$

and interchanging the order of the integration. The saddle-point approximation of this expression is very similar to the above, and is simply obtained by multiplying (3.74) by $1/z_0 = 2s$. We next demonstrate how this is done in the context of assessing a probability of a large-deviations event.

Example 3.7 (Large deviations). This example delves into a topic that was extensively studied by Bahadur and Rao [13]. Here, we offer a partial exposition to illustrate the application of the saddle-point method. Consider a set of IID RVs $X_1, X_2, \ldots, X_n$, all of which are independent copies of a real RV X with a PDF $p(x)$. Additionally, let A be a constant greater than the expected value of X. We aim to evaluate the probability of a large-deviations event, namely, $\{\sum_{i=1}^n X_i \geq nA\}$, utilizing the saddle-point method. Introducing θ as an arbitrary positive real number, we have from (3.76):

$$\begin{aligned}
&\Pr\left\{\sum_{i=1}^n X_i \geq nA\right\} \\
&= \int_{\mathbb{R}^n} U\left(\sum_{i=1}^n x_i - nA\right) \prod_{i=1}^n p(x_i)\ \mathrm{d}\boldsymbol{x} && (3.77) \\
&= \int_{\mathbb{R}^n} \frac{1}{2\pi j} \int_{\mathrm{Re}\{z\}=\theta} \frac{\mathrm{d}z}{z} \cdot \exp\left\{z\left(\sum_{i=1}^n x_i - nA\right)\right\} \cdot \prod_{i=1}^n p(x_i)\ \mathrm{d}\boldsymbol{x} && (3.78) \\
&= \frac{1}{2\pi j} \int_{\mathrm{Re}\{z\}=\theta} \frac{e^{-znA}}{z} \cdot \mathrm{d}z \int_{\mathbb{R}^n} \prod_{i=1}^n \left[p(x_i)e^{zx_i}\right]\ \mathrm{d}\boldsymbol{x} && (3.79) \\
&= \frac{1}{2\pi j} \int_{\mathrm{Re}\{z\}=\theta} \frac{e^{-znA}}{z} \cdot \mathrm{d}z \left[\int_{\mathbb{R}} p(x)e^{zx}\ \mathrm{d}x\right]^n && (3.80) \\
&= \frac{1}{2\pi j} \int_{\mathrm{Re}\{z\}=\theta} \frac{\mathrm{d}z}{z} \cdot \exp\left\{n\left[\ln\left(\int_{\mathbb{R}} p(x)e^{zx}\ \mathrm{d}x\right) - zA\right]\right\}, && (3.81)
\end{aligned}$$

and we can apply[3] the saddle-point method with $g(z) = 1/z$ and

$$f(z) = \ln\left(\int_{\mathbb{R}} p(x)e^{zx}\ \mathrm{d}x\right) - zA. \tag{3.82}$$

Consider the function f confined to the reals, namely, $f(s)$, where $s \in \mathbb{R}$. Since $f(s)$ is a convex function, it can be shown that its derivative vanishes uniquely at some finite real $s = s_\star > 0$, provided that $A < x_{\max} \triangleq \sup_{\{x:\ p(x)>0\}} x$. Then, $z = s_\star$ is a saddle-point of f.

At this point, we have to distinguish between two cases — non-lattice and lattice RVs.

Non-lattice RVs: Let us first assume that p is such that $z = s_\star$ is the only saddle-point of f in the entire complex plane (shortly, we

[3] There is a non-trivial issue concerning the non-analyticity of the logarithmic function, whose argument, $\int_{\mathbb{R}} p(x)e^{zx}\ \mathrm{d}x$, may surround the origin infinitely many times while z exhausts the vertical line $\mathrm{Re}\{z\} = \theta$, because the origin is a singular point of the logarithmic function. This requires to pass among different branches of the logarithmic function along the journey from $\theta - j\infty$ to $\theta + j\infty$. This issue is discussed in detail in [143].

also address situations where this is not the case). In this case, a simple application of the saddle-point method suggests to select $\theta = s_\star$, where the axis is vertical, and so,

$$\Pr\left\{\sum_{i=1}^{n} X_i \geq nA\right\}$$
$$\sim \frac{1}{s_\star} \cdot \frac{e^{j\pi/2}}{2\pi j} \cdot \exp\left\{n\left[\ln\left(\int_{\mathbb{R}} p(x)e^{s_\star x}\ \mathrm{d}x\right) - s_\star A\right]\right\} \cdot \sqrt{\frac{2\pi}{nV(s_\star)}} \tag{3.83}$$
$$= \frac{\exp\left\{n\left[\ln\left(\int_{\mathbb{R}} p(x)e^{s_\star x}\ \mathrm{d}x\right) - s_\star A\right]\right\}}{s_\star\sqrt{2\pi nV(s_\star)}}, \tag{3.84}$$

where $V(s) = f''(s) = \mathrm{Var}_s\{X\}$, with the latter being defined as the variance of X WRT the PDF that is proportional to $p(x)e^{sx}$, *i.e.*, the tilted PDF. It is worth highlighting the intriguing similarity between the exponential term

$$\exp\left\{n\left[\ln\left(\int_{\mathbb{R}} p(x)e^{s_\star x}\ \mathrm{d}x\right) - s_\star A\right]\right\}, \tag{3.85}$$

and Chernoff's bound, as $s_\star$ minimizes $f(s)$ over the real numbers. At the same time, $z = s_\star$ is determined as the saddle-point that dominates the integration along the vertical line defined by $\mathrm{Re}\{z\} = s_\star$. This observation aligns with the modulus theorem: Given that $z = s_\star$ minimizes $|e^{nf(z)}| = e^{nf(s)}$ horizontally along the real line, it maximizes $|e^{nf(z)}|$ along the vertical direction of the integration path. While the exponential behavior of the saddle-point approximation mirrors that of Chernoff's bound, known for its exponential tightness [49], it is noteworthy that the former provides a more refined characterization, including the correct pre-exponential factor, which is given by $1/[s_\star\sqrt{2\pi nV(s_\star)}]$. In Appendix A we provide an alternative justification of the tightness of Chernoff's bound, which is based on the method of types and the minimax theorem (though, without the correct pre-exponential factor).

Lattice RVs: As previously mentioned, in the earlier derivation, we made the assumption that $z = s_\star$ represents the sole saddle-point of the function f across the entire complex plane. However, this assumption does not hold universally. Let us consider a scenario in which X is a lattice RV, implying that X can only assume values that are integer multiples of a constant $\Delta > 0$, that is,

$$p(x) = \sum_{i=-\infty}^{\infty} \alpha_i \delta(x - i\Delta), \tag{3.86}$$

where $\delta(\cdot)$ is the Dirac delta function and $\{\alpha_i\}$ are non-negative reals which sum up to unity. Consider the vertical line of integration, $z = s_\star + j\omega$, $-\infty < \omega < \infty$. In this scenario, it becomes evident that if $s_\star$ is a saddle-point of $e^{nf(z)}$, then so are the points $s_\star + j\Omega k$, where k ranges over all integers ($k = 0, \pm 1, \pm 2, ...$), and Ω is defined as $\Omega = 2\pi/\Delta$. This is due to the periodic nature of $|e^{nf(z)}|$, which is equivalent to $e^{n\mathrm{Re}\{f(z)\}}$, along the vertical direction with a period of Ω. Indeed,

$$\begin{aligned}
&\mathrm{Re}\left\{f(s_\star + jk\Omega)\right\} \\
&= \mathrm{Re}\left\{\ln\left[\int_{\mathbb{R}} p(x) e^{(s_\star + jk\Omega)x}\, \mathrm{d}x\right] - (s_\star + jk\Omega)A\right\} && (3.87)\\
&= \mathrm{Re}\left\{\ln\left[\sum_{i=-\infty}^{\infty} \alpha_i e^{(s_\star + jk\Omega)i\Delta}\right]\right\} - s_\star A && (3.88)\\
&= \mathrm{Re}\left\{\ln\left[\sum_{i=-\infty}^{\infty} \alpha_i e^{s_\star i\Delta} e^{jki\Omega\Delta}\right]\right\} - s_\star A && (3.89)\\
&= \mathrm{Re}\left\{\ln\left[\sum_{i=-\infty}^{\infty} \alpha_i e^{s_\star i\Delta} e^{j2\pi ik}\right]\right\} - s_\star A && (3.90)\\
&= \mathrm{Re}\left\{\ln\left[\sum_{i=-\infty}^{\infty} \alpha_i e^{s_\star i\Delta}\right]\right\} - s_\star A && (3.91)\\
&= \mathrm{Re}\{f(s_\star)\}. && (3.92)
\end{aligned}$$

In such a situation, during the integration along the line $\mathrm{Re}\{z\} = s_\star$, the contributions from all saddle-points, $s_\star + jk\Omega$ for $k = 0, \pm 1, \pm 2, \ldots$, carry equal significance, collectively dominating the exponential rate of the integral. This has a notable impact on the pre-exponential factor, which now needs to be adjusted to reflect this collective contribution. Therefore, the modified pre-exponential factor is given by:

$$\frac{1}{\sqrt{2\pi n V(s_\star)}} \cdot \sum_{k=-\infty}^{\infty} \frac{e^{-jk\Omega A n}}{s_\star + jk\Omega}$$

$$= \sqrt{\frac{2\pi}{nV(s_\star)}} \cdot \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-j\omega nA} \cdot \frac{1}{s_\star + j\omega} \cdot \left[\sum_{k=-\infty}^{\infty} \delta(\omega - k\Omega)\right] \mathrm{d}\omega \tag{3.93}$$

$$\stackrel{(*)}{=} \sqrt{\frac{2\pi}{nV(s_\star)}} \cdot \left\{ \left[e^{-s_\star t} U(t)\right] \star \left[\frac{1}{\Omega} \sum_{k=-\infty}^{\infty} \delta\left(t - \frac{2\pi k}{\Omega}\right)\right] \right\} \Bigg|_{t=-nA} \tag{3.94}$$

$$= \frac{1}{\Omega} \sqrt{\frac{2\pi}{nV(s_\star)}} \sum_{k=-\infty}^{\infty} e^{-s_\star(-nA - 2\pi k/\Omega)} U\left(-nA - \frac{2\pi k}{\Omega}\right) \tag{3.95}$$

$$= \frac{1}{\Omega} \sqrt{\frac{2\pi}{nV(s_\star)}} \cdot \exp\left\{-s_\star \left[(-nA) \bmod \left(\frac{2\pi}{\Omega}\right)\right]\right\} \cdot \sum_{k=0}^{\infty} e^{-s_\star \cdot 2\pi k/\Omega} \tag{3.96}$$

$$= \sqrt{\frac{2\pi}{nV(s_\star)}} \cdot \frac{\exp\left\{-s_\star \left[(-nA) \bmod \left(\frac{2\pi}{\Omega}\right)\right]\right\}}{\Omega(1 - e^{-2\pi s_\star/\Omega})} \tag{3.97}$$

$$= \sqrt{\frac{2\pi}{nV(s_\star)}} \cdot \frac{\Delta \exp\{-s_\star[(-nA) \bmod \Delta]\}}{2\pi(1 - e^{-s_\star \Delta})} \tag{3.98}$$

$$= \sqrt{\frac{1}{2\pi nV(s_\star)}} \cdot \frac{\Delta \exp\{-s_\star[(-nA) \bmod \Delta]\}}{1 - e^{-s_\star \Delta}}, \tag{3.99}$$

where in $(*)$ we have used the fact that the inverse Fourier transform of the product of two frequency-domain functions is equal to the convolution between the individual inverse Fourier transforms. The oscillatory factor in the numerator, $\exp\{-s_\star[(-nA) \bmod \Delta]\}$, illustrates the granularity inherent in the probability quanta related to the lattice-like nature of the involved RVs (also discussed in [143]). It is worth noting that the non-lattice scenario can be considered as a specific case of the lattice scenario, where $\Delta \to 0$.

Our final example pertains to the enumeration of codewords within a hyper-cubical lattice subject to an L_1 power constraint. The motivation here is to evaluate the coding rate of a hyper-cubical lattice code (defined below). In a nutshell, when the hyper-cubes are exceptionally small, this count approximates the ratio between the volume of the L_1 hypersphere defining the power constraint and the volume of the hyper-cube. However, the saddle-point method provides a more precise estimation.

Example 3.8 (Number of codewords of a power-limited lattice code). Let us examine a hyper-cubical lattice code, where the codewords take the form of $(k_1\Delta, k_2\Delta, \ldots, k_n\Delta)$, with $\Delta > 0$ given, $\{k_i\}$ being integers, and adhering to the L_1 power constraint $\Delta \sum_{i=1}^n |k_i| \leq nQ$. What is the number M of lattice codewords that can be found? We can establish the following sequence of equalities:

$$M = \sum_{k_1=-\infty}^{\infty} \cdots \sum_{k_n=-\infty}^{\infty} U\left[nQ - \Delta \sum_{i=1}^{n} |k_i|\right] \tag{3.100}$$

$$= \sum_{k_1=-\infty}^{\infty} \cdots \sum_{k_n=-\infty}^{\infty} \frac{1}{2\pi j} \int_{\mathrm{Re}\{z\}=\theta} \frac{\mathrm{d}z}{z} \exp\left\{z\left[nQ - \Delta \sum_{i=1}^{n} |k_i|\right]\right\} \tag{3.101}$$

$$= \frac{1}{2\pi j} \int_{\mathrm{Re}\{z\}=\theta} \mathrm{d}z \frac{e^{nQz}}{z} \sum_{k_1=-\infty}^{\infty} \cdots \sum_{k_n=-\infty}^{\infty} \exp\left\{-\Delta z \sum_{i=1}^{n} |k_i|\right\} \tag{3.102}$$

$$= \frac{1}{2\pi j} \int_{\mathrm{Re}\{z\}=\theta} \mathrm{d}z \frac{e^{nQz}}{z} \left[\sum_{k=-\infty}^{\infty} \exp\{-\Delta z|k|\}\right]^n \tag{3.103}$$

$$= \frac{1}{2\pi j} \int_{\mathrm{Re}\{z\}=\theta} \mathrm{d}z \frac{e^{nQz}}{z} \left[\frac{e^{\Delta z}+1}{e^{\Delta z}-1}\right]^n \tag{3.104}$$

$$= \frac{1}{2\pi j} \int_{\mathrm{Re}\{z\}=\theta} \frac{\mathrm{d}z}{z} \exp\left\{n\left[Qz - \ln\tanh\left(\frac{\Delta z}{2}\right)\right]\right\}. \tag{3.105}$$

Thus, the saddle-point method can be applied with $g(z) = 1/z$ and

$$f(z) = Qz - \ln\tanh\left(\frac{\Delta z}{2}\right) \tag{3.106}$$

$$= Qz - \ln\sinh\left(\frac{\Delta z}{2}\right) + \ln\cosh\left(\frac{\Delta z}{2}\right). \tag{3.107}$$

The derivative of f vanishes at

$$z = s_\star = \frac{1}{\Delta} \ln\left(\frac{\Delta}{Q} + \sqrt{\frac{\Delta^2}{Q^2}+1}\right), \tag{3.108}$$

but similarly as in Example 3.7, here too, $\mathrm{Re}\{f(z)\}$ is periodic in the vertical direction with period $\Omega = 2\pi/\Delta$, and so, there are infinitely many saddle-points $\{s_\star + jk\Omega,\ k = 0, \pm 1, \pm 2, \ldots\}$, and M is exponentially $e^{nf(s_\star)}$ with the same pre-exponential factor as in the lattice case of Example 3.7, except that $(-nA)$ mod Δ is replaced by (nQ) mod Δ and $V(s_\star)$ is replaced by $|f''(s_\star)|$. Therefore, the coding rate (in nats per channel use) is of the form,

$$R = \frac{\ln M}{n} = f(s_\star) - \frac{\ln n}{2n} + o\left(\frac{\ln n}{n}\right), \tag{3.109}$$

with

$$f(s_\star) = \frac{Q}{\Delta}\ln\left(\frac{\Delta}{Q} + \sqrt{\frac{\Delta^2}{Q^2}+1}\right) + \ln\left(\frac{\Delta}{Q} + \sqrt{\frac{\Delta^2}{Q^2}+1}+1\right) - \ln\left(\frac{\Delta}{Q} + \sqrt{\frac{\Delta^2}{Q^2}+1}-1\right). \tag{3.110}$$

It is easy to verify that when $\Delta/Q \ll 1$, the exponential factor, $e^{nf(s_\star)}$ is approximately $\frac{(2eQ)^n}{\Delta^n}$, which is exponentially the ratio between volume of the L_1-hyper-sphere of 'radius' nQ and the volume of the hyper-cube, Δ^n. We skip the details of calculating $f''(s_\star)$ for the pre-exponent.

In conclusion, we note that a similar calculation for the more traditional L_2 power constraint involves dealing with the infinite summation $\sum_k e^{-z\Delta^2 k^2}$ (instead of $\sum_k e^{-\Delta z|k|}$ as in our previous analysis). Although this expression lacks an apparent closed-form representation, the same fundamental behavior persists: The rate remains primarily determined by the log-volume ratio, subtracting $\frac{\ln n}{2n}$, with some negligible terms.

3.6 Discussion — Extension to the Multivariate Case

In Example 3.6, we witnessed the powerful capability of the saddle-point method in assessing type class measures without the need for the ϵ-inflation technique employed in Section 2. When confronted with the task of integrating over $\boldsymbol{x}$ a function of the form $f(\sum_{i=1}^n x_i^2)$, we can conveniently rewrite this as an equivalent integral over $f(r)S_n(r)$ WRT r. This transformation effectively replaces the n-dimensional integration with a one-dimensional integration, which, in certain cases, can be well-approximated using either the Laplace method or the saddle-point method.

In Section 2, we explored more intricate type classes defined as intersections between hyper-sphere surfaces and hyper-planes, such as $\sum_{i=1}^{n} x_i = nc$. Evaluating the Lebesgue measure of such objects involves integrating a product of delta functions, specifically $\delta(ns - \sum_i x_i^2) \cdot \delta(nc - \sum_i x_i)$. To compute this measure, we represent each delta function as an inverse Laplace transform separately, each with its own complex integration variable, *i.e.*,

$$\int_{\mathbb{R}^n} \delta\left(ns - \sum_{i=1}^{n} x_i^2\right) \cdot \delta\left(nc - \sum_{i=1}^{n} x_i\right) \, \mathrm{d}\boldsymbol{x}$$

$$= \frac{1}{(2\pi j)^2} \int_{\mathbb{R}^n} \int_{\theta - j\infty}^{\theta + j\infty} \int_{\nu - j\infty}^{\nu + j\infty} \mathrm{d}z_1 \mathrm{d}z_2 \cdot e^{z_1\left(ns - \sum_{i=1}^{n} x_i^2\right) + z_2\left(nc - \sum_{i=1}^{n} x_i\right)} \, \mathrm{d}\boldsymbol{x} \tag{3.111}$$

$$= \frac{1}{(2\pi j)^2} \int_{\theta - j\infty}^{\theta + j\infty} \int_{\nu - j\infty}^{\nu + j\infty} \mathrm{d}z_1 \mathrm{d}z_2 \cdot \int_{\mathbb{R}^n} e^{z_1\left(ns - \sum_{i=1}^{n} x_i^2\right) + z_2\left(nc - \sum_{i=1}^{n} x_i\right)} \, \mathrm{d}\boldsymbol{x} \tag{3.112}$$

$$= \frac{1}{(2\pi j)^2} \int_{\theta - j\infty}^{\theta + j\infty} \int_{\nu - j\infty}^{\nu + j\infty} \mathrm{d}z_1 \mathrm{d}z_2 \cdot e^{n(zs + z'c)} \left[\int_{\mathbb{R}} e^{-(z_1 x^2 + z_2 x)} \mathrm{d}x\right]^n \tag{3.113}$$

$$= \frac{1}{(2\pi j)^2} \int_{\theta - j\infty}^{\theta + j\infty} \int_{\nu - j\infty}^{\nu + j\infty} \mathrm{d}z_1 \mathrm{d}z_2 \cdot e^{n(z_1 s + z_2 c)} \left[\exp\left\{\frac{z_2^2}{4z_1^2}\right\} \sqrt{\frac{\pi}{z_1}}\right]^n \tag{3.114}$$

$$= \frac{\pi^{n/2}}{(2\pi j)^2} \int_{\theta - j\infty}^{\theta + j\infty} \int_{\nu - j\infty}^{\nu + j\infty} \mathrm{d}z_1 \mathrm{d}z_2 \cdot \exp\left\{n\left[z_1 s + z_2 c + \frac{z_2^2}{4z_1^2} - \frac{\ln z_1}{2}\right]\right\}, \tag{3.115}$$

where θ and ν are arbitrary positive reals. In cases like this, an extension of the saddle-point method to the multivariate setting is required. As outlined in [158], the extension of the saddle-point method to integration over more than one complex variable is analogous to the previously mentioned extension of the Laplace integration method to the d-dimensional case. Namely, it is based on an approximation by an integral of a d-dimensional Gaussian integral with an inverse covariance matrix given by the Hessian of f at the saddle-point z_0. Conceptually, it can also be thought of as a succession of d univariate integration operations of one coordinate at a time. In [158, Theorem 2.1], an explicit theorem is provided to this end, where the result is the same as in (3.14), except that x_0 should be replaced by z_0 (under proper conditions).

Building on these insights, if we encounter the need to integrate a function of the form $f(\sum_{i=1}^n x_i^2, \sum_{i=1}^n x_i)$, we can transform it into a two-dimensional integration of f multiplied by the Lebesgue measure of the corresponding type class, following a similar procedure to what was just described. These considerations are applicable to types defined by any fixed number of constraints, including those related to conditional types (*e.g.*, constraints involving $\sum_{i=1}^n x_i y_i$) and constraints associated with Gauss–Markov types (such as constraints specifying values of $\sum_{i=1}^n x_i x_{i-\ell}$ for $\ell = 1, 2, \ldots, k$). Notably, the saddle-point method allows for the combination of constraints, even those involving $\sum_{i=1}^n x_i y_i$ and $\sum_{i=1}^n x_i x_{i-\ell}$. This capability resolved an outstanding challenge posed in [116] and was successfully addressed in [90], particularly in the context of the Gaussian intersymbol interference channel, thanks to the versatility of the saddle-point method.

Extending this generality further, instead of linear and quadratic constraints, situations may arise with constraints involving combinations of empirical means of arbitrary functions, denoted as $\sum_{i=1}^n \phi_j(x_i)$ for $j = 1, 2, \ldots, k$. The associated saddle-point integration in these cases will involve exponential functions of linear combinations of these statistics. It is important to note that the coefficients of these linear combinations can be complex in general. In essence, this entails working with exponential families characterized by complex parameters.

3.7 Further Applications

The saddle-point method has found extensive applications in various disciplines, including probability theory, mathematical statistics, and physics, with notable usage in statistical physics. While less common in the information theory community, there have been exceptions in the last two decades.

In Example 3.1, we demonstrated how the Laplace integration method can be effectively employed to approximate Bayesian mixtures of memoryless sources, particularly relevant to universal source coding [46], [108]. Schwartz also utilized this approximation to derive a model order estimator from a Bayesian perspective within a sequence of nested parametric families [181].

Several researchers have applied the Laplace and saddle-point methods to obtain more refined bounds on the error probability of channel coding and decoding, including characterizations of the pre-exponential factor, in addition to the exponential one. Notable contributors to this area include Atluğ and Wagner [6], Font-Segura, Vázquez-Vilar, Martinez, and Guillén i Fàbregas [67], Honda [87], Martinez and Guillén i Fàbregas [114], [113], and Scarlett, Martinez and Guillén i Fàbregas [174]. These methods have also been applied to derive sharper bounds on the probability of error in binary hypothesis testing [211].

Furthermore, the saddle-point and Laplace methods have been applied to finite blocklength analysis and higher-order asymptotics of achievable coding rates. Researchers like Anade, Gorce, Mary, and Perlaza [7], Erseghe [62], Moulin [153], Polyanskiy [166], Tan and Tomamichel [199], Yavas, Kostina, and Wigger [228], and Lancho, Östman, Durisi, Koch and Vázquez-Vilar [109] have contributed to this area.

In the work by Huleihel, Salamatian, Merhav, and Médard [90], the saddle-point approximation was applied to assess the log-volume of a conditional Gaussian type class related to the Gaussian intersymbol interference channel, with implications for mismatched universal decoding. This addressed an open problem from [116].

In [143], the saddle-point approximation was used to refine the evaluation of the probability that a randomly selected codeword would fall within a sphere of a specified radius from a given source vector, based on a given distortion measure. The precise pre-exponential factor allowed for the characterization of redundancy rates. In [122], the method was applied to lossless data compression in the context of the set partitioning problem.

Lastly, in [151, Section 4.7], Mézard and Montanari establish a valuable link between the saddle-point method, Sanov's theorem, and the method of types, providing further insights into the connections between these powerful techniques.

4

The Type Class Enumeration Method

4.1 Introduction

In Section 2, we considered probabilistic properties of a single random vector, or a finite collection of vectors, and developed a generalized version of *the method of types* [38], [41]. In this section, we ascend one hierarchical level and consider analysis of *coding* problems, and specifically the problem of evaluating the *error exponent* in coded systems. Such problems involve an *exponential* number of random vectors, and so, the method we propose in this section will require additional analytical tools.

Starting from Shannon [182], the common method of proving achievability results in information theory is via *random-coding* analysis, in which the error probability is averaged over an ensemble of randomly selected codebooks. While the random-coding argument was originally invoked to find the *capacity* C of noisy channels [36], it was broadly adapted to other settings as well. In this section, we will focus on error exponent analysis [60], [65, Chapters 7-9], [71, Chapter 5], [40], [42], which is a refined performance measure of coded systems. The error exponent refers to the largest exponential decay rate of the error probability of a sequence of codes at increasing blocklength n, for a given rate R below the capacity C. Since the error probability of the optimal codebook can be *upper bounded* by the average of the error probability over an ensemble of random codebooks, the error exponent can be *lower bounded* by the *random-coding error exponent* — the exponential decay rate of the ensemble-average error probability.

Moreover, the random-coding error exponent is interesting as *a paradigm on its own right*, since it is by now well-established that random codes, or *random-like* codes (*e.g.*, turbo codes [17] and low-density parity-check (LDPC) codes [69]; see [171]) are highly efficient [31]. In fact, in some applications, the codebook is routinely redrawn at random, for example, in order to preserve the security of the transmitted information. So, when a communication system uses such a random code, it is the random-coding error probability (or exponent) that is a relevant measure to the long-term performance of the system, rather than just serving as a lower bound to the best achievable exponent.

The analysis of the random-coding error exponent has lead to the proposal and usage of a large number of analytical bounding methods. We next outline several of them, in order to contrast them later on with our type of techniques.

First, the error probability of the optimal ML decoder can be upper bounded by the error probability of simpler, sub-optimal decoder. For example, the error probability of the typicality decoder [36, Chapter 7] decays to zero at all rates below capacity, just as the ML decoder. So, analyzing the typicality decoder can be used to prove lower bounds on the capacity. However, this decoder has poor performance in terms of the error exponent.

Second, as popularized by Gallager [70], Jelinek [99] and Forney [68], the use of convexity properties and *Jensen-style* inequalities. These include, for example, the inequality $\mathbb{E}[Z^\rho] \le (\mathbb{E}[Z])^\rho$ for a non-negative RV and $0 \le \rho \le 1$, or the power distribution inequality

$$\Big(\sum_j a_j\Big)^\rho \le \sum_j a_j^\rho \tag{4.1}$$

(see [212, Appendix 3A] for a comprehensive list of such inequalities).

Third, the use of *Chernoff-style* bounds, in which an indicator of an error event, based on likelihoods, is replaced by their ratio. For example, a pairwise error event of an ML decoder over the channel W from $\boldsymbol{x}$ to $\boldsymbol{y}$ is upper bounded as

$$\mathbb{1}\left\{W(\boldsymbol{y}|\boldsymbol{x}_j) \geq W(\boldsymbol{y}|\boldsymbol{x}_i)\right\} \leq \left[\frac{W(\boldsymbol{y}|\boldsymbol{x}_j)}{W(\boldsymbol{y}|\boldsymbol{x}_i)}\right]^{\lambda} \tag{4.2}$$

for any $\lambda \geq 0$.

Fourth, *refined union bounds*, in which the simple union bound over events $\{\mathcal{A}_j\}$ is replaced by a quantity lower than the sum of probabilities of each event. These bounds include, *a truncated union bound*

$$\Pr\left[\bigcup_j \mathcal{A}_j\right] \leq \min\left[1, \sum_j \Pr[\mathcal{A}_j]\right], \tag{4.3}$$

a union bound with a power parameter $0 \leq \rho \leq 1$ (also known as *Gallager's union bound* [71, p. 136])

$$\Pr\left[\bigcup_j \mathcal{A}_j\right] \leq \left(\sum_j \Pr[\mathcal{A}_j]\right)^{\rho}, \tag{4.4}$$

or *a union bound with intersection* of an event $\mathcal{G}$

$$\Pr\left[\bigcup_j \mathcal{A}_j\right] \leq \sum_j \Pr\left[\mathcal{A}_j \cap \mathcal{G}\right] + \Pr\left[\mathcal{G}^c\right], \tag{4.5}$$

where $\mathcal{G}^c$ is the complement of $\mathcal{G}$. As an illustrative example, such a union bound can be used to bound the probability of an error event in channel coding, since this event is a union of the events that one of the alternative codewords is decoded. The above bounding methods then lead to tractable, computable, bounds on the random-coding exponent, and other quantities of interest.

Nonetheless, in typical channel coding problems, codebooks with a positive coding rate R have an exponential number of codewords e^{nR}, and so, the analysis of the error probability involves evaluation of the probability of a union of an *exponential* number of events. In some cases, it can be shown that a bound obtained via these methods is actually tight. For example, in the simple case of a point-to-point DMC, Gallager has shown that its random-coding error exponent, obtained using (4.4), is tight, by *lower* bounding the error probability [72]. However, there is no *general* claim that these bounding methods lead to the *exact* random-coding error exponent, that is, that the final result is the true exponential decay rate of the expected error probability over the random ensemble of codebooks. In fact, in various scenarios they are *strictly loose*.

In this section, we introduce the *type class enumeration method* (TCEM) for the analysis of random codes, which is an original viable alternative or complement to the aforementioned techniques. It is a principled method, whose main virtue is that it preserves exponential tightness along all steps of the derivation of the exponent. It is therefore guaranteed to obtain the *exact* exponent. The TCEM achieves that by refraining from using the various bounding techniques mentioned above, and thus avoids the need to optimize over various parameters (which cannot always be done in a closed-form), and leads to explicit expressions. More often than not, it does so in a "single-pass", *i.e.*, without separately lower and upper bounding the random-coding error exponent. Consequently, ensemble-tight random-coding exponents can be obtained in a multitude of coding problems. Moreover, as mentioned, and as we shall survey, in coding problems that go beyond basic ones, the error exponents obtained by the TCEM are oftentimes strictly larger than those achieved using the above bounding techniques.

For this section, we recall the usual notation convention for an equality or an inequality in the exponential scale: For two positive sequences $\{a_n\}$ and $\{b_n\}$, the notation $a_n \doteq b_n$ means that $\lim_{n\to\infty} \frac{1}{n} \ln \frac{a_n}{b_n} = 0$, and $a_n \stackrel{.}{\leq} b_n$ means that $\lim_{n\to\infty} \frac{1}{n} \ln \frac{a_n}{b_n} \leq 0$. Accordingly, $a_n \doteq 1$ means that a_n is sub-exponential, and $a_n \doteq e^{-n\infty}$ means that a_n decays at a super-exponential rate (*e.g.*, double-exponentially).

The main idea of the TCEM is that each codeword can be categorized according to a *joint type* (empirical distribution) with an additional length-n vector, and that and the union bound is exponentially tight for a union of the polynomial number of events. Indeed, for k_n events $\{\mathcal{E}_i\}_{i=1}^{k_n}$

$$\max_{1\leq m\leq k_n} \Pr\left[\mathcal{E}_m\right] \leq \Pr\left[\bigcup_{m=1}^{k_n} \mathcal{E}_m\right] \leq k_n \cdot \max_{1\leq m\leq k_n} \Pr\left[\mathcal{E}_m\right], \tag{4.6}$$

and so if $k_n \doteq 1$ then

$$\Pr\left[\bigcup_{m=1}^{k_n} \mathcal{E}_m\right] \doteq \max_{1\le m\le k_n} \Pr\left[\mathcal{E}_m\right]. \tag{4.7}$$

Therefore, the analysis of a coding problem can be based on a *type class enumerator* (TCE), which counts the number of randomly selected codewords in a suitably defined type class. For illustration, one may recall that for binary symmetric channels (BSCs), the *distance spectrum of a codebook*, namely, the number of pairs of codewords at each of the $n+1$ possible Hamming distances, plays an important role in determining its error probability (*e.g.*, [112, Chapter 2]). Indeed, a specific form of TCEs for BSCs was used by [14] to analyze various random-coding exponents. The TCEM can be thought of as a considerable generalization of this fundamental idea.

In the TCEM, the codebook is drawn at random, and consequently, the TCEs are RVs. The random-coding error exponent thus depends on their probabilistic and statistical properties, such as moments or tail bounds. Each TCE is typically a binomial RV $N \sim \text{Binomial}(e^{nA}, e^{-nB})$ (or a close variant of such variables), defined by e^{nA} independent trials for belonging to a type class, each with success probability e^{-nB}. It exhibits an interesting *phase transition* at $A = B$: If the number of trials dominates the success rate, $A > B$, then the TCE is tightly concentrated around its exponentially large expected value $e^{n(A-B)}$ (double-exponential concentration). We refer to these as *typically populated* types (as coined in [197]). Otherwise, if $B > A$, then the TCE is typically zero, and the probability that it is strictly positive is exponentially less than $e^{-n(B-A)}$. We refer to these as *typically empty* types. The transition between these regimes is sharp, and is rooted in a statistical-mechanical perspective on random coding. This perspective is based on an analogy to Derrida's random energy model (REM) [50]–[52], [151, Chapters 5 and 6], which is a spin glass model with high degree of disorder, and which is well known in the literature of statistical physics of magnetic materials. The phase transition in the REM is analogous to the one exhibited for the TCEs, and we refer the reader to [119] and [120, Chapter 6] for a thorough exposition.

The TCEM hence involves the following steps: (1) Expressing the error probability (or other quantity of interest) using suitably defined TCEs. (2) Evaluating the necessary probabilistic and statistical properties of the TCEs (moments or tail probabilities). (3) Plugging in these properties in the expression for the error probability, and evaluating the resulting expression. (4) Developing an efficient procedure to compute the exponent. This last step is equally important, since in some cases, the resulting expression for the exponent may appear involved or challenging to compute. We show in Appendix B how efficient methods can be developed.

For the sake of simplicity of the exposition, we focus in this section on DMCs, for which the standard method of types [38], [41] is applicable. However, given the generalized method of types described in Section 2, these ideas can be extended to other channels, including Gaussian channels (which have continuous alphabets) and channels with memory, without requiring a substantial modification, *e.g.*, [138], [196].

The outline of this section is as follows. For methodological reasons, our first step will invoke the TCEM for problems in which error exponents are already well-established, namely, error exponents for DMCs (random-coding [40], [42], expurgated [70, Section V], [98]) and the correct decoding exponent for rates above capacity [10], [57]. This will exemplify the technique of the TCEM in a familiar setting, and serve as a basis for the rest of the section. We will then derive the basic statistical and probabilistic properties of TCEs, to wit, tail probabilities and moments. Later, we will demonstrate the TCEM in more advanced settings, namely: (1) The error exponent of superposition coding in a broadcast asymmetric DMC for the optimal bin-index decoder. (2) The random-binning error exponent of distributed compression [189]. (3) The random-coding error exponents of generalized decoders, such as Forney's erasure/list decoder [68] and a generalized version of the likelihood decoder [227]. (4) The error exponent of the typical random code [14].

In the last subsection, we will survey the wide applicability of the TCEM, and its ability to provide exact random-coding exponents in a multitude of information-theoretic problems: The problem could be a channel coding or a source coding problem; the problem could involve a single user and point-to-point channels, or multiple users operating in a distributed manner over a network [59]; the code could have a fixed length, be a convolutional/trellis code [100], [209], [212], or have variable encoding length (with feedback) [26]; the decoder could be the optimal ML decoder, the universal MMI decoder [77], a mismatched decoder [179], an erasure/list decoder [68] that is allowed to output an erasure or more than a single codeword, a list decoder that outputs a list of possible codewords [61], [223], a bin-index decoder, which is the optimal ML decoder in which the codeword is only known to belong to a bin; a likelihood decoder which randomly decodes a message based on a posterior probability distribution [227]; a joint detector-decoder that is required to make a decision in addition to decoding the message [213]; and more. Moreover, beyond the random-coding error exponent, other exponents can also be derived using the TCEM, *e.g.*, the error exponent of the typical random code [14] and large-deviations from this typical code [197], [205].

4.2 Basic Coding Problems

To obtain a quick glance on the underlying ideas, we first consider the basic problems of the random-coding and expurgated exponents for a DMC, and then the correct decoding exponent (for rates above capacity). Along the way, we will introduce several useful techniques, such as the summation–maximization equivalence, tail integration, and, later on, exponential tightness of the union bound for pairwise independent events. For the sake of convenience, we begin with a short background of classic error exponents for DMCs.

4.2.1 A Short Background: Error Exponents of DMCs

Consider a DMC W with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$, and a codebook $\mathcal{C}_n = \{\boldsymbol{x}_m\}$ whose codewords $\boldsymbol{x}_m \in \mathcal{X}^n$ have blocklength n, and it has rate R, that is $|\mathcal{C}_n| = e^{nR}$.[1] One of the most important problems of error exponent analysis [41, Chapter 10], [71, Chapter 5],

[1]Throughout, we will ignore integer constraints on large quantities such as e^{nR} (which should be $\lceil e^{nR} \rceil$), since these do not affect any of the analyses or the results.

is to find the maximum achievable error exponent achieved at any rate R, also known as the *reliability function,* $E^*(R)$. This establishes the existence of a sequence of codes $\{\mathcal{C}_n^*\}$ of rate R, whose error probability decays with the maximal exponent[2]

$$E^*(R) = \sup_{\{\mathcal{C}_n\}} \limsup_{n\to\infty} -\frac{1}{n} \ln P_\mathsf{e}(\mathcal{C}_n^*), \tag{4.8}$$

where $P_\mathsf{e}(\mathcal{C}_n)$ is the error probability of the codebook $\mathcal{C}_n$ (for a given, implicit, decoding rule). As expected, pointing out a particular sequence of codes achieving the reliability function is a formidable problem. The random-coding argument shows that $E^*(R)$ is lower bounded by the exponent achieved by random codes. Specifically, we consider a random ensemble in which each codeword $\boldsymbol{X}_m \in \mathcal{X}^n$ is chosen randomly, independent of all other codewords, and in identical way: In the IID ensemble, each symbol of the codeword is drawn independently from some distribution P_X, and in the fixed-composition ensemble, each codeword is chosen uniformly at random from a type class $\mathcal{T}_n(P_X)$. While both ensembles can be analyzed using the TCEM, we will focus on the latter since it is more common when invoking the method of types, and since it typically leads to larger random-coding exponents. The average error probability for a random codebook $\mathfrak{C}_n$ chosen from the ensemble will be denoted by $\overline{P}_\mathsf{e} \triangleq \mathbb{E}[P_\mathsf{e}(\mathfrak{C}_n)]$. For a given ensemble, the random-coding error exponent at rate R is then given by

$$E_{\text{rc}}(R) \triangleq \lim_{n\to\infty} -\frac{1}{n} \ln \mathbb{E}\left[P_\mathsf{e}(\mathfrak{C}_n)\right], \tag{4.9}$$

whenever the limit exists, for which it holds that $E^*(R) \geq E_{\text{rc}}(R)$.

The random-coding error exponent was studied by two different schools. First, an approach lead by Gallager [71, Chapter 5], which is based on analytical techniques such as refined union bounds, and later, by Csiszár, Körner and Marton [40]–[42], who developed and used the method of types [38] to this problem. Since the TCEM is based on the method of types, we will next describe the latter [41, Chapter 10]. For a DMC W, and a fixed-composition input P_X, this random-coding error

[2]It is unclear if the following limit exists [41, Exercise 10.7], and so we take the conservative definition of limit-superior.

exponent takes the form $E_{\rm rc}(R) = \max_{P_X} E_{\rm rc}(R, P_X)$, where, with a slight abuse of notation,

$$E_{\rm rc}(R, P_X) \triangleq \min_{Q_{Y|X}} \left\{ D(Q_{Y|X}||W|P_X) + \Big[I(P_X \times Q_{Y|X}) - R \Big]_+ \right\}. \quad (4.10)$$

It was also shown that this exponent can be achieved using the MMI decoder, and does not require the optimal ML decoder. In parallel, it was proved that the *sphere packing bound* [19], [65], [82], [183] $E_{\rm sp}(R) \triangleq \max_{P_X} E_{\rm sp}(R, P_X)$, where

$$E_{\rm sp}(R, P_X) \triangleq \min_{Q_{Y|X}:\, I(P_X \times Q_{Y|X}) \leq R} D(Q_{Y|X}||W|P_X), \quad (4.11)$$

is an upper bound on the reliability function $E^*(R) \leq E_{\rm sp}(R)$. Remarkably, there exists a critical rate $R_{\rm cr}$ such that for any $R \geq R_{\rm cr}$ it holds that

$$E^*(R) = E_{\rm rc}(R) = E_{\rm sp}(R), \quad (4.12)$$

and so at high rates, the reliability function is exactly known, and random-coding is optimal. At low rates, $R < R_{\rm cr}$, however, the ensemble-average error probability may be highly affected by codes with large error probability. This has lead to the idea of *expurgating* the ensemble from these codes, and to the development of the expurgated exponent.[3] The expurgated exponent $E_{\rm ex}(R)$ is a lower bound on the reliability function $E^*(R) \geq E_{\rm ex}(R)$, and improves on the random-coding error exponent at low rates. Let the Bhattacharyya distance between $\boldsymbol{x}, \tilde{\boldsymbol{x}} \in \mathcal{X}^n$ be defined as

$$d_{\rm B}(\boldsymbol{x}, \tilde{\boldsymbol{x}}) \triangleq -\ln \sum_{\boldsymbol{y} \in \mathcal{Y}^n} \sqrt{W(\boldsymbol{y}|\boldsymbol{x}) \cdot W(\boldsymbol{y}|\tilde{\boldsymbol{x}})}. \quad (4.13)$$

Since it only depends on the joint type, $Q_{X\tilde{X}} = \hat{Q}_{\boldsymbol{x},\tilde{\boldsymbol{x}}}$, we also denote, with a slight abuse of notation, $d_{\rm B}(Q_{X\tilde{X}})$ as the Bhattacharyya between some $(\boldsymbol{x}, \tilde{\boldsymbol{x}}) \in \mathcal{T}_n(Q_{X\tilde{X}})$. The expurgated exponent [38], [40]–[42] is given by $E_{\rm ex}(R) = \max_{P_X} E_{\rm ex}(R, P_X)$ where

[3] The development of the expurgated error exponent also involves expurgating bad codewords from codebooks.

$$E_{\text{ex}}(R, P_X) \triangleq \min_{Q_{X\tilde{X}} :\, Q_X = Q_{\tilde{X}} = P_X,\, I(Q_{X\tilde{X}}) \leq R} [d_{\text{B}}(Q_{X\tilde{X}}) + I(Q_{X\tilde{X}})] - R. \quad (4.14)$$

Also remarkably, Shannon, Gallager and Berlekamp [184] showed that the expurgated exponent is tight at zero rate $E^*(0) = E_{\text{ex}}(0)$, and also used this result to derive an improved upper bound at intermediate rates, known as the *straight line bound*.

4.2.2 Random Coding Error Exponent of a DMC via Type Class Enumeration

We next show how to derive the random-coding error exponent $E_{\text{rc}}(R)$ via the TCEM. As usual, we fix the transmitted codeword $\boldsymbol{X}_1 = \boldsymbol{x}$ and the output vector $\boldsymbol{y}$, and then write the probability that one of the $e^{nR} - 1$ (random) competing codewords in $\mathfrak{C}_n \backslash \{\boldsymbol{X}_1\}$ is decoded instead of $\boldsymbol{X}_1$. This amounts to

$$\overline{P}_{\mathsf{e}} = \sum_{\boldsymbol{x} \in \mathcal{X}^n} \sum_{\boldsymbol{y} \in \mathcal{Y}^n} \Pr[\boldsymbol{X}_1 = \boldsymbol{x}] \, W(\boldsymbol{y}|\boldsymbol{x}) \times \Pr\left[\bigcup_{m=2}^{e^{nR}} \{\boldsymbol{X}_m \text{ has higher score than } \boldsymbol{X}_1 = \boldsymbol{x}\}\right]. \quad (4.15)$$

The next reasonable step is to further bound the inner probability by a union bound, and as said, while a naive union bound fails, the clipped union bound (4.3) or Gallager's union bound (4.4) both lead to the exact random-coding error exponent in this basic setting. However, the TCEM proceeds differently.

Let us denote by Q_{XY} a generic joint type of $(\boldsymbol{x}, \boldsymbol{y})$, where $Q_X = P_X$ matches the type of the fixed-composition ensemble, and where for brevity, henceforth, we will often make this implicit. We further consider the class of α-decoders, which decide using a score function $\alpha(Q_{XY})$ that depends only on the joint type of the output vector $\boldsymbol{y}$ and the candidate codeword. Specifically, if $\hat{Q}_{\boldsymbol{x},\boldsymbol{y}}$ is the joint type of $(\boldsymbol{x}, \boldsymbol{y})$ then the decoded codeword is $X(\hat{j})$ where

$$\hat{j}(\boldsymbol{y}) = \arg\max_{1 \leq j \leq e^{nR}} \alpha(\hat{Q}_{\boldsymbol{x}_j, \boldsymbol{y}}), \quad (4.16)$$

where ties are arbitrarily broken. Let the expected log-likelihood of a joint type Q_{XY} be

$$f(Q_{XY}) \triangleq \mathbb{E}_Q \left[\ln W(Y|X)\right]. \tag{4.17}$$

It can be easily noted that the MMI decoder, $\alpha(Q_{XY}) = I(Q_{XY})$, and the ML decoder, $\alpha(Q_{XY}) = f(Q_{XY})$, are both α-decoders. We now introduce a suitable TCE.

Definition 4.1 (TCE for random-coding exponent)**.** For a codebook $\mathcal{C}_n = \{\boldsymbol{x}_m\}$, an output vector $\boldsymbol{y}$, and a joint type Q_{XY} such that $\hat{Q}_{\boldsymbol{y}} = Q_Y$, let

$$N_{\boldsymbol{y}}(Q_{XY}, \mathcal{C}_n) \triangleq \left|\{m > 1 \colon (\boldsymbol{x}_m, \boldsymbol{y}) \in \mathcal{T}_n(Q_{XY})\}\right|. \tag{4.18}$$

The TCE $N_{\boldsymbol{y}}(Q_{XY}, \mathcal{C}_n)$ counts the number of incorrect codewords in $\mathcal{C}_n$ whose joint type with $\boldsymbol{y}$ is Q_{XY}. By the method of types (the fourth property in Section (2.2.1)), when $\boldsymbol{X}_m \sim \text{Uniform}[\mathcal{T}_n(Q_X)]$ then

$$\Pr\left[(\boldsymbol{X}_m, \boldsymbol{y}) \in \mathcal{T}_n(Q_{XY})\right] = k_n \cdot e^{-nI(Q_{XY})} \tag{4.19}$$

for some $k_n \doteq 1$. So, for a random codebook $\mathfrak{C}_n = \{\boldsymbol{X}_m\}$,

$$\begin{aligned} N_{\boldsymbol{y}}(Q_{XY}, \mathfrak{C}_n) &= \sum_{m=2}^{e^{nR}} \mathbb{1}\left\{(\boldsymbol{X}_m, \boldsymbol{y}) \in \mathcal{T}_n(Q_{XY})\right\} \\ &\sim \text{Binomial}\left(e^{nR} - 1, k_n \cdot e^{-nI(Q_{XY})}\right). \end{aligned} \tag{4.20}$$

More generally, since any $\boldsymbol{X}_m$ has a unique joint type with $\boldsymbol{y}$, then viewed as a collection of TCEs, it holds that

$$\{N_{\boldsymbol{y}}(Q_{XY})\}_{Q_{XY}} \sim \text{Multinomial}\left(e^{nR}, \{p(Q_{XY})\}_{Q_{XY}}\right) \tag{4.21}$$

where $p(Q_{XY}) \doteq e^{-nI(Q_{XY})}$.

Since the probability distribution of $N_{\boldsymbol{y}}(Q_{XY}, \mathfrak{C}_n)$ only depends on $\boldsymbol{y}$ through its type, for brevity, we will omit both $\mathfrak{C}_n$ and $\boldsymbol{y}$ from the notation of TCEs (with a slight abuse of notation). We then have

$$\overline{P}_{\mathsf{e}} = \sum_{Q_{XY}} \Pr\left[(\boldsymbol{X}_1, \boldsymbol{y}) \in \mathcal{T}_n(Q_{XY})\right] \times$$

$$\Pr\left[\bigcup_{\tilde{Q}_{XY}\colon Q_Y=\tilde{Q}_Y,\, \alpha(\tilde{Q}_{XY})\geq\alpha(Q_{XY})} \mathbb{1}\left\{N(\tilde{Q}_{XY})\geq 1\right\}\right]. \quad (4.22)$$

The substantial difference between this bound and (4.15), is that its inner probability is a union over a polynomial number of types, rather than an exponential number of codewords. For such a union of polynomial number of events, even the regular union bound is exponentially tight, as shown in (4.6), and therefore

$$\overline{P}_{\mathsf{e}} \doteq \max_{Q_{XY}} \max_{\tilde{Q}_{XY}} \Pr\left[(\boldsymbol{X}_1,\boldsymbol{y})\in\mathcal{T}_n(Q_{XY})\right]\cdot\Pr\left[N(\tilde{Q}_{XY})\geq 1\right] \quad (4.23)$$

$$\overset{(*)}{\doteq} \max_{Q_{XY}} \max_{\tilde{Q}_{XY}} \exp\left[-n\cdot D(Q_{XY}||P_X\times W)\right]\cdot\Pr\left[N(\tilde{Q}_{XY})\geq 1\right], \quad (4.24)$$

where the inner maximization is over the set

$$\left\{\tilde{Q}_{XY}\colon Q_Y=\tilde{Q}_Y,\ \alpha(\tilde{Q}_{XY})\geq\alpha(Q_{XY})\right\}, \quad (4.25)$$

and in $(*)$ we have used the method of types [(2.12) in Section 2.2.1]. In the next section, we will derive various properties of TCEs, and specifically, tight tail bounds on $\Pr[N(\tilde{Q}_{XY})\geq 1]$. After inserting the tail bound of Theorem 4.1 back to (4.24) we obtain

$$\overline{P}_{\mathsf{e}} \doteq \exp\left[-n\cdot E_{\mathrm{rc},\alpha}(R)\right], \quad (4.26)$$

where

$$E_{\mathrm{rc},\alpha}(R,P_X) \triangleq \min_{Q_{Y|X},\tilde{Q}_{Y|X}} D(Q_{Y|X}||W|P_X)+\left[I(P_X\times\tilde{Q}_{Y|X})-R\right]_+, \quad (4.27)$$

for which the minimization is over the set

$$\Big\{Q_{Y|X},\ \tilde{Q}_{Y|X}\colon (P_X\times Q_{Y|X})_Y=(P_X\times\tilde{Q}_{Y|X})_Y,\ \alpha(P_X\times\tilde{Q}_{Y|X})\geq\alpha(P_X\times Q_{Y|X})\Big\}. \quad (4.28)$$

This recovers a similar bound from [40] obtained in a different way. For example, if $\alpha(Q_{XY})$ is the MMI rule, the input the minimization is over $\{I(P_X \times \tilde{Q}_{Y|X}) \geq I(P_X \times Q_{Y|X})\}$. We recover the *random-coding error exponent* (4.10). If $\alpha(Q_{XY}) = f(Q_{XY}) \triangleq \mathbb{E}_Q[\ln W(Y|X)]$ is the ML rule, then we achieve the same exponent. Indeed, since the ML is the optimal decoder in terms of error probability, its error probability can only be lower. On the other hand, $\tilde{Q}_{XY} = Q_{XY}$ belongs to the set of inner minimization, and so the exponent cannot be larger (see [40, Proof of Lemma 4] for a direct proof, which does not utilize the optimality of the ML rule).

At this point, we pause for a brief comment on a delicate technical issue. The minimization in (4.27) should, in principle, be taken over the set of possible types for sequences of length n, and then the limit $n \to \infty$ should be taken. As a general rule, in almost all other derivations based on the TCEM, we obtain an expression of the form

$$\max_{Q \in \mathcal{Q}_n} \exp\left[-ng(Q) + n\epsilon_n\right], \tag{4.29}$$

where $\mathcal{Q}_n$ is a set that is typically the intersection of some feasible set $\mathcal{Q}$ and the set of possible types of length n, $g(Q)$ is some function, and $\epsilon_n = o(n)$ does *not* depend on Q. Then, the exponent $E(Q)$ is obtained by evaluating

$$E(Q) = \lim_{n\to\infty} -\frac{1}{n} \ln \max_{Q \in \mathcal{Q}_n} e^{-ng(Q)+n\epsilon_n} = \lim_{n\to\infty} \min_{Q \in \mathcal{Q}_n} g(Q). \tag{4.30}$$

However, typically, under mild assumptions on the probabilistic model of the problem, it holds that $g(Q)$ is *uniformly continuous* in $\mathcal{Q}$, and so the limit results in a minimization over the feasible set $\mathcal{Q}$, that is,

$$E(Q) = \min_{Q \in \mathcal{Q}} g(Q) \tag{4.31}$$

That being said, there are cases that uniform convergence is not simple to establish, see [215], [217], [218] for two such cases. Moreover, in [215], in which the TCEM was used to derive random-coding error exponents for a distributed hypothesis testing problem, such an issue was an obstacle for proving that the obtained exponent is indeed exact.

For general decoding scores, the random-coding error exponent $E_{\text{rc},\alpha}(R, P_X)$ is lower than the standard random-coding error exponent, and on the face of it, is difficult to compute. Indeed, the clipping operation is the result of the phase transition of the TCE at $R = I(\tilde{Q}_{XY})$. This leads to an exponent expression whose feasible set may be partitioned into two subsets, one with the additional constraint $I(\tilde{Q}_{XY}) \leq R$, and the other one with the additional constraint $I(\tilde{Q}_{XY}) > R$. Then, the objective function of (4.27) may be separately minimized over each of these subsets, and the value of $E_{\text{rc},\alpha}(R, P_X)$ is given by the minimum of the two minimal values. In this method, the first optimization problem is typically simpler. Indeed, if we re-parameterize $\tilde{Q}_{XY} = P_X \cdot \tilde{Q}_{Y|X}$ for a fixed input distribution P_X, then $I(P_X \cdot \tilde{Q}_{Y|X})$ is a convex function of $\tilde{Q}_{Y|X}$, and the sub-level set $\{I(\tilde{Q}_{XY}) \leq R\}$ is a convex set. When the constraint set $\{\alpha(P_X \times \tilde{Q}_{Y|X}) \geq \alpha(P_X \times Q_{Y|X})\}$ is also convex (which occurs when $\alpha(Q_{XY})$ is linear in Q_{XY}, as for the ML decoder), the resulting optimization problem has a convex feasible set. Since the objective function of (4.27) is jointly convex in $(Q_{Y|X}, \tilde{Q}_{Y|X})$, this results in a convex optimization problem, which can be efficiently solved [21]. However, the second optimization problem is problematic since its constraint set $\{I(\tilde{Q}_{XY}) > R\}$ is a super-level set that is not a convex set. Nonetheless, in Appendix B we show a method to efficiently compute the exponent, which only requires solving convex optimization problems (assuming $\alpha(Q)$ is linear). As we have seen, the phase transition ($I > R$ or $I < R$) holds generally for TCEs, and so this issue occurs for almost any exponent derived by this method, which may take much more complicated form. Nonetheless, methods similar to the one described in Appendix B can usually be developed to efficiently compute the exponent, even for such complicated scenarios, see *e.g.*, [215, Section VI], [64, Section V], and [216, Appendix A].

4.2.3 Expurgated Exponent of a DMC via Type Class Enumeration

We next move on to shortly discuss the expurgated exponent. Assuming that the ML decoder is used, the pairwise error probability for two codewords $\boldsymbol{x}$ and $\tilde{\boldsymbol{x}}$ is upper bounded by the *Bhattacharyya bound* (*e.g.*, [41, Problem 10.20]) as

$$P_{\mathsf{e}}(\boldsymbol{x}, \tilde{\boldsymbol{x}}) \triangleq \Pr\left[W(\boldsymbol{Y}|\tilde{\boldsymbol{x}}) \geq W(\boldsymbol{Y}|\boldsymbol{x})\right] \leq \exp\left[-n \cdot d_{\mathrm{B}}(\boldsymbol{x}, \tilde{\boldsymbol{x}})\right]. \quad (4.32)$$

Thus, for a given code $\mathcal{C}_n$,

$$P_{\mathsf{e}}(\mathcal{C}_n) \overset{(a)}{\leq} \frac{1}{e^{nR}} \sum_{m=1}^{e^{nR}} \sum_{\tilde{m}=1}^{e^{nR}} \mathbb{1}\{\tilde{m} \neq m\} \cdot P_{\mathsf{e}}(\boldsymbol{x}_m, \boldsymbol{x}_{\tilde{m}}) \tag{4.33}$$

$$\overset{(b)}{\leq} \frac{1}{e^{nR}} \sum_{m=1}^{e^{nR}} \sum_{\tilde{m}=1}^{e^{nR}} \mathbb{1}\{\tilde{m} \neq m\} \cdot \exp\left[-n \cdot d_{\mathrm{B}}(\boldsymbol{x}_m, \boldsymbol{x}_{\tilde{m}})\right], \tag{4.34}$$

where (a) follows from the regular union bound and (b) follows from (4.32). We next introduce a suitable TCE for the expurgated exponent.

Definition 4.2 (TCE for expurgated exponent)**.** For a joint type $Q_{X\tilde{X}}$, a codebook $\mathcal{C}_n = \{\boldsymbol{x}_m\}$, and a codeword index $m = 1, \ldots, e^{nR}$, let

$$\overline{N}_m(Q_{X\tilde{X}}, \mathfrak{C}_n) \triangleq \left|\{\tilde{m} \colon \tilde{m} \neq m,\ (\boldsymbol{x}_m, \boldsymbol{x}_{\tilde{m}}) \in \mathcal{T}_n(Q_{X\tilde{X}})\}\right|, \tag{4.35}$$

count the number of codewords in the codebook $\mathcal{C}_n$ which have a joint type $Q_{X\tilde{X}}$ with $\boldsymbol{x}_m$. By the method of types [(2.12), Section (2.2.1)], when $\boldsymbol{X}_{\tilde{m}} \sim \text{Uniform}[\mathcal{T}_n(Q_X)]$ then

$$\Pr\left[(\boldsymbol{X}_{\tilde{m}}, \boldsymbol{x}_m) \in \mathcal{T}_n(Q_{X\tilde{X}})\right] = k_n \cdot e^{-nI(Q_{X\tilde{X}})} \tag{4.36}$$

for some $k_n \doteq 1$. So, for a random codebook $\mathfrak{C}_n = \{\boldsymbol{X}_m\}$ it holds that

$$\overline{N}_m(Q_{X\tilde{X}}, \mathfrak{C}_n) \triangleq \sum_{\tilde{m}=1}^{e^{nR}} \mathbb{1}\{\tilde{m} \neq m\} \cdot \mathbb{1}\left\{(\boldsymbol{X}_{\tilde{m}}, \boldsymbol{X}_m) \in \mathcal{T}_n(Q_{X\tilde{X}})\right\}$$

$$\sim \text{Binomial}\left(e^{nR} - 1, k_n \cdot e^{-nI(Q_{X\tilde{X}})}\right). \tag{4.37}$$

It should be noted that $\{\overline{N}_m(Q_{X\tilde{X}}, \mathfrak{C}_n)\}_{m=1}^{e^{nR}}$ is a collection of an exponential number of *dependent* RVs.

As for the TCE for random-coding, we will omit $\mathfrak{C}_n$ from the notation of TCEs (with a slight abuse of notation). Evidently, the upper bound in (4.34) can be expressed using the TCEs as

$$P_{\mathsf{e}}(\mathcal{C}) \leq \frac{1}{e^{nR}} \sum_{m=1}^{e^{nR}} \sum_{Q_{X\tilde{X}}} \overline{N}_m(Q_{X\tilde{X}}) \cdot \exp\left[-n \cdot d_{\mathrm{B}}(Q_{X\tilde{X}})\right]. \tag{4.38}$$

In Appendix C, we show how the bound (4.38) and the properties of $\{\overline{N}_m(Q_{X\tilde{X}})\}_{m=1}^{e^{nR}}$ derived in the next section can be used to derive the classic expurgated exponent (4.14).

4.2.4 The Correct Decoding Exponent of a DMC

One of the first demonstrations of the usefulness of the TCEM was for the *correct* decoding exponent of a DMC at rates above capacity [119]. Following Arimoto [10], the correct decoding error probability of the ML decoder begins with the identity

$$P_{\mathsf{c}}(\mathcal{C}_n) = \frac{1}{e^{nR}} \sum_{\boldsymbol{y} \in \mathcal{Y}^n} \max_m W(\boldsymbol{y}|\boldsymbol{X}_m) \tag{4.39}$$

$$= \lim_{\beta \to \infty} \frac{1}{e^{nR}} \sum_{\boldsymbol{y} \in \mathcal{Y}^n} \left[\sum_m W^\beta(\boldsymbol{y}|\boldsymbol{X}_m) \right]^{1/\beta}. \tag{4.40}$$

Let us consider the TCE

$$N_{\boldsymbol{y}}(Q_{XY}) \triangleq |\{m \geq 1 \colon (\boldsymbol{x}_m, \boldsymbol{y}) \in \mathcal{T}_n(Q_{XY})\}|, \tag{4.41}$$

which is only slightly different from the random-coding TCE of Definition 4.1, and so, we abuse the notation and denote them similarly. Assuming an ensemble of random codebooks, we next evaluate the ensemble-average of the correct decoding probability. Recalling that $f(Q_{XY}) \triangleq \mathbb{E}_Q[\ln W(Y|X)]$, we fix a finite β and $\boldsymbol{y}$, and write the ensemble average using TCEs as

$$\mathbb{E}\left\{ \left[\sum_m W^\beta(\boldsymbol{y}|\boldsymbol{X}_m) \right]^{1/\beta} \right\}$$

$$= \mathbb{E}\left\{ \left[\sum_{Q_{XY}} N(Q_{XY}) \cdot e^{n\beta f(Q_{XY})} \right]^{1/\beta} \right\} \tag{4.42}$$

$$\overset{(*)}{\doteq} \mathbb{E}\left\{ \left[\max_{Q_{XY}} N(Q_{XY}) \cdot e^{n\beta f(Q_{XY})} \right]^{1/\beta} \right\} \tag{4.43}$$

$$= \mathbb{E}\left\{ \left[\max_{Q_{XY}} N^{1/\beta}(Q_{XY}) \cdot e^{n f(Q_{XY})} \right] \right\} \tag{4.44}$$

$$\overset{(*)}{\doteq} \mathbb{E}\left[\sum_{Q_{XY}} N^{1/\beta}(Q_{XY}) \cdot e^{n f(Q_{XY})} \right] \tag{4.45}$$

$$= \sum_{Q_{XY}} \mathbb{E}\left[N^{1/\beta}(Q_{XY}) \right] \cdot e^{n f(Q_{XY})}, \tag{4.46}$$

where $(*)$ both follow from the fact that the number of types is polynomial in n to interchange a summation with a maximum in both directions. We refer to this as the *summation–maximization equivalence*, which is frequently used to manipulate probabilities to a form that allows for a direct substitution of TCE moments. As can be seen, (4.46) requires evaluating the fractional $1/\beta$ moment of the TCE $N(Q_{XY})$. Next, since $|\mathcal{T}_n(Q_Y)| \doteq e^{nH(Q_Y)}$, we obtain

$$\overline{P}_{\mathsf{c}} \triangleq \mathbb{E}\left[P_{\mathsf{c}}(\mathfrak{C}_n)\right] \tag{4.47}$$

$$\stackrel{(a)}{=} \mathbb{E}\left[\lim_{\beta\to\infty} \frac{1}{e^{nR}} \sum_{\boldsymbol{y}\in\mathcal{Y}^n} \left[\sum_m W^\beta(\boldsymbol{y}|\boldsymbol{X}_m)\right]^{1/\beta}\right] \tag{4.48}$$

$$\stackrel{(b)}{=} \lim_{\beta\to\infty} \mathbb{E}\left[\frac{1}{e^{nR}} \sum_{\boldsymbol{y}\in\mathcal{Y}^n} \left[\sum_m W^\beta(\boldsymbol{y}|\boldsymbol{X}_m)\right]^{1/\beta}\right] \tag{4.49}$$

$$\stackrel{(c)}{\doteq} \lim_{\beta\to\infty} \sum_{Q_{XY}} \mathbb{E}\left[N^{1/\beta}(Q_{XY})\right] \cdot e^{n[f(Q_{XY})+H(Q_Y)-R]}, \tag{4.50}$$

where (a) follows from Arimoto's identity (4.40); (b) is an interchange of the order of order of the limit and the expectation operator. It is justified by first noting that $[\sum_m W^\beta(\boldsymbol{y}|\boldsymbol{X}_m)]^{1/\beta}$ is the L_β-norm of the vector $(W^\beta(\boldsymbol{y}|\boldsymbol{X}_1), W^\beta(\boldsymbol{y}|\boldsymbol{X}_2), \ldots, W^\beta(\boldsymbol{y}|\boldsymbol{X}_{e^{nR}}))$, and

$$\frac{1}{e^{nR}} \sum_{\boldsymbol{y}\in\mathcal{Y}^n} \left[\sum_m W^\beta(\boldsymbol{y}|\boldsymbol{X}_m)\right]^{1/\beta} \tag{4.51}$$

is a monotonic non-negative and non-increasing function of β, and thus has a finite limit when $\beta \to \infty$ for any $\{\boldsymbol{X}_m\}$. Also, this function is bounded by its value at $\beta = 1$, which is integrable

$$\mathbb{E}\left[\frac{1}{e^{nR}} \sum_{\boldsymbol{y}\in\mathcal{Y}^n} \sum_m W(\boldsymbol{y}|\boldsymbol{X}_m)\right] = |\mathcal{Y}|^n < \infty. \tag{4.52}$$

Then, the order interchange is justified by the dominated convergence theorem. Finally, (c) follows from (4.46).[4]

[4]Strictly speaking, if the number of joint types is k_n, which is polynomial in n, then the polynomial factor hidden by the notation of exponential equality in (4.46) is between $\frac{1}{k_n}$ and $k_n^{1/\beta}$, which belongs to $(\frac{1}{k_n}, k_n)$, an interval that is independent of β.

In the next section, we will evaluate the moments $\mathbb{E}[N^{1/\beta}(Q_{XY})]$ (concretely, using Theorem 4.2 and the standard method of types). Using this result, we obtain[5]

$$\overline{P}_{\mathsf{c}} \doteq \exp\left[-n \cdot E_{\mathrm{c}}(R, P_X)\right] \tag{4.53}$$

where

$$E_{\mathrm{c}}(R, P_X) \triangleq \min\left\{E_{-}(R, P_X), E_{+}(R, P_X)\right\}, \tag{4.54}$$

with

$$E_{-}(R, P_X) \triangleq \min_{Q_{XY}\,:\, I(Q_{XY})>R} \left[I(Q_{XY}) - f(Q_{XY}) - H(Q_Y)\right], \tag{4.55}$$

as well as

$$\begin{aligned} &E_{+}(R, P_X) \\ &\triangleq \lim_{\beta\to\infty} \min_{Q_{XY}\,:\, I(Q_{XY})\le R} \left[\frac{1}{\beta} I(Q_{XY}) - \frac{1}{\beta} R + R - f(Q_{XY}) - H(Q_Y)\right] \end{aligned} \tag{4.56}$$

$$= \min_{Q_{XY}\,:\, I(Q_{XY})\le R} \left[R - f(Q_{XY}) - H(Q_Y)\right]. \tag{4.57}$$

Therefore,

$$\begin{aligned} E_{\mathrm{c}}(R, P_X) &= \min_{Q_{XY}} \left[\max\{R, I(Q_{XY})\} - f(Q_{XY}) - H(Q_Y)\right] && (4.58)\\ &= \min_{Q_{XY}} \left\{D(Q_{XY}\|P_X \times W) + \left[R - I(Q_{XY})\right]_+\right\}, && (4.59) \end{aligned}$$

where the second equality uses the identity

$$\begin{aligned} &D(Q_{XY}\|P_X \times W) - I(Q_{XY}) + H(Q_Y) \\ &= \sum_{x\in\mathcal{X}, y\in\mathcal{Y}} Q_{XY}(x,y) \ln \frac{Q_{XY}(x,y)}{P_X(x)W(y|x)} \\ &\quad - \sum_{x\in\mathcal{X}, y\in\mathcal{Y}} Q_{XY}(x,y) \ln \frac{Q_{XY}(x,y)}{P_X(x)Q(y)} \end{aligned}$$

[5] Again, the rigorous argument requires exchange of limits between β and n. Specifically, after substituting the moments $\mathbb{E}[N^{1/\beta}(Q_{XY})]$ we get that $E_{\mathrm{c}}(R, P_X) = \lim_{n\to\infty} \lim_{\beta\to\infty} \Gamma(n, \beta)$ for some function $\Gamma(n, \beta)$. It can be shown that $\lim_{n\to\infty} \Gamma(n, \beta)$ converges uniformly in β (we omit the details), and so the Moore-Osgood theorem [201, Section 3.1] assures that the order of limits may be interchanged.

$$+ \sum_{x\in\mathcal{X},y\in\mathcal{Y}} Q_{XY}(x,y) \ln \frac{1}{Q_Y(y)} \tag{4.60}$$

$$= \sum_{x\in\mathcal{X},y\in\mathcal{Y}} Q_{XY}(x,y) \ln \frac{1}{W(y|x)} \tag{4.61}$$

$$= -f(Q_{XY}). \tag{4.62}$$

We also remark that in all the above expressions, $Q_X = P_X$ is implicitly assumed.

This bound recovers the Körner–Dueck exponent [57], which is known to be optimal (after minimizing over the input distribution P_X). In [120, Chapter 6], this example was studied in detail for a BSC, and compared with Arimoto's approach in [10]. Arimoto started as in (4.40), but continued by upper bounding these moments using Jensen's inequality, which interchanges between the expectation operator and the $1/\beta$-power [similarly to (4.40)]. As was shown in [119, Section 3], for a BSC with crossover probability p, the correct-decoding exponent is

$$\overline{P}_{\mathsf{c}} \doteq \exp\left[-n \cdot D(\delta_{\mathrm{GV}}(R)||p)\right] \tag{4.63}$$

$$= \exp\left[-n \cdot \left(\delta_{\mathrm{GV}}(R) \ln \frac{1}{p} + (1-\delta_{\mathrm{GV}}(R)) \ln \frac{1}{1-p}\right) - H(\delta_{\mathrm{GV}}(R))\right], \tag{4.64}$$

where $\delta_{\mathrm{GV}}(R)$ is the (smaller) solution to $\ln 2 - H(\delta) = R$ (recall that $H(q) \triangleq -q\ln q - (1-q)\ln(1-q)$ is the binary entropy function). If one now follows (4.50), and takes $\beta \to \infty$, then after using Jensen's inequality, the following bound is obtained:

$$\overline{P}_{\mathsf{c}} \overset{\cdot}{\leq} \exp\left[-n \cdot \left(\min\left\{\ln \frac{1}{p}, \ln \frac{1}{1-p}\right\}\right) - H(\delta_{\mathrm{GV}}(R))\right]. \tag{4.65}$$

Evidently, the exponent in (4.65) is *strictly* smaller than the exact exponent in (4.64), and so using a Jensen-based derivation and taking $\beta \to \infty$ leads to a loose bound on the exponent. Arimoto was able to recover his bound from the sub-optimality of Jensen's inequality by replacing the limit $\beta \to \infty$ with a maximization over β, and obtained an exponent that matches that of Körner–Dueck [57] (after also minimizing over the input distribution P_X). However, there is no general guarantee that such optimization over β can lead to an exact exponent, and the additional optimization over β may be intractable. Indeed, in more complicated settings, such as the ones discussed in Section 4.4 to follow, the optimization over parameters (such as β) for derivations that are based on Jensen inequality cannot be performed analytically, and even if so, they lead to strictly sub-optimal bounds.

4.3 Probabilistic and Statistical Properties of Type Class Enumerators

In the previous section, we showed how to analyze basic coded systems via the TCEM. In this section, we turn to analyze the probabilistic and statistical properties of TCEs. Motivated by the discussion up until now, we let n be the blocklength, and let $A, B > 0$ be two constants. We will be interested in RVs of the form $N \sim \text{Binomial}(k'_n \cdot e^{nA}, k''_n \cdot e^{-nB})$, where $k'_n \cdot e^{nA}$ is the number of trials, and $k''_n \cdot e^{-nB}$ is the probability of a successful trial, and where $k'_n \doteq k''_n \doteq 1$. Specifically, we will be interested in tail probabilities and moments of these RVs. As we shall see, the asymptotic logarithm normalized by n of these probabilities/moments will be a continuous function of (A, B) in some open set. Thus, in that open set, the polynomial pre-factors k'_n and k''_n do not affect the asymptotic result, as it can be sandwiched by the corresponding results for $N \sim \text{Binomial}(e^{n(A\pm\epsilon)}, e^{-n(B\pm\epsilon)})$, with any arbitrarily small $\epsilon > 0$. As we shall next see, such binomial RVs experience a phase transition at $B = A$, and therefore we will separate the analysis to the cases of $B < A$ and $B > A$. As a side note, for $A = B$, the distribution of such a binomial RV tends to that of a Poisson RV.

We begin by studying the tail probabilities of a generic TCE N.

Theorem 4.1. *Assume that* $N \sim \text{Binomial}(e^{nA}, e^{-nB})$ *and* $\lambda \in \mathbb{R}$. *Then, the upper tail is*

$$\lim_{n\to\infty} -\frac{1}{n} \ln \Pr\left[N > e^{n\lambda}\right] = \begin{cases} [B-A]_+\,, & [A-B]_+ \geq \lambda \\ \infty, & \textit{elsewhere} \end{cases}, \tag{4.66}$$

and the lower tail is

$$\lim_{n\to\infty} -\frac{1}{n}\ln \Pr\left[N < e^{n\lambda}\right] = \begin{cases} 0, & A-B<\lambda \\ \infty, & A-B>\lambda \end{cases}. \tag{4.67}$$

The proof of Theorem 4.1, as well as all other theorems in this section, is deferred to Appendix D.

We continue with the moments of the TCE N. While the first moment (expected value) of N is trivially given by $\mathbb{E}[N] = e^{n(A-B)}$, error exponent analysis typically requires to evaluate general moments, which are possibly fractional.

Theorem 4.2. *Assume that* $N \sim \text{Binomial}(e^{nA}, e^{-nB})$*. Then, for any* $s > 0$

$$\mathbb{E}\left[N^s\right] \doteq \begin{cases} e^{n(A-B)s}, & A>B \\ e^{-n(B-A)}, & A<B \end{cases}. \tag{4.68}$$

Importantly, for $A < B$ the moment is asymptotically independent of s.

In various advanced settings, the analysis of the TCEM also requires to evaluate probabilistic and statistical properties of a *pair* of dependent TCEs, or even a family $\{N_j\}_{j=1}^{k_n}$ of sub-exponential number of TCEs ($k_n \doteq 1$), which are possibly *dependent*. For example, let (U_1, U_2) be a pair of dependent Bernoulli RVs so that $\Pr[U_j = 1] = e^{-nB_j}$ for $j = 1, 2$. Typically, U_j are indicators for disjoint events, *e.g.*, U_1 is the event in which a random vector belongs to some type class, and U_2 belongs to a different type class. Thus, only one at most of the U_j is 1. Now, assume that we draw e^{nA} IID RVs from the distribution of (U_1, U_2), and let N_j denote the corresponding number of successes for U_j, for $j = 1, 2$. While strictly speaking the TCEs are dependent RVs, we next show they are asymptotically independent in the regime we consider. Indeed, let us condition on the event that $N_1 = e^{n\nu}$ for some $\nu \in [0, A)$. Then, $N_2 | N_1 = e^{n\nu} \sim \text{Binomial}(e^{nA} - e^{n\nu}, \Pr[U_1 = 1|U_2 = 0])$. Evidently, the number of trials is $e^{nA} - e^{n\nu} \sim e^{nA}$ and the success probability is

$$\Pr\left[U_1 = 1 | U_2 = 0\right] = \frac{\Pr\left[U_1 = 1, U_2 = 0\right]}{\Pr\left[U_2 = 0\right]} \tag{4.69}$$

$$= \frac{\Pr\left[U_1 = 1\right]}{1 - e^{-nB_2}} \tag{4.70}$$

$$\sim e^{-nB_1}. \tag{4.71}$$

So, up to factors that tend to 1, the parameters of the conditional binomial N_1 are exactly as those of the unconditional binomial. It is easy to generalize the above argument to a sub-exponential number of TCEs, that is, $\{N_j\}_{j=1}^{k_n}$ where $k_n \doteq 1$, thus showing that they are asymptotically independent. Indeed, if we consider the full set of TCEs, *i.e.*, $\{N(Q)\}$ of all possible types, for which it must hold that $\sum N(Q) = e^{nA}$, then each trial is successful for exactly one of the types. Thus, the joint distribution of $\{N(Q)\} \sim \text{Multinomial}(e^{nA}, \{p_Q\}_Q)$, where $p_Q \doteq e^{-nB(Q)}$. It is well-known that for a large number of trials, the multinomial distribution tends to an *independent* Poisson distribution [152, Theorem 5.6].

In the context of superposition codebooks to be discussed later on in Section 4.4, the joint distribution of TCEs were analyzed in the arXiv version of [215, Appendix D]. As we will also see in that setting, it is required to analyze the probability of an intersection of upper tail events of TCEs $\{N_j\}_{j=1}^{k_n}$, when $k_n \doteq 1$ and where $N_j \sim \text{Binomial}(e^{nA_j}, e^{-nB_j})$. This result is also obtained from the joint distribution of the TCEs, and is addressed by the following theorem.

Theorem 4.3. *Assume that* $N_j \sim \text{Binomial}(e^{nA_j}, e^{-nB_j})$ *for* $j = 1, \ldots, k_n$ *and* $k_n \doteq 1$. *Assume that* $\lambda \in \mathbb{R}$, *and* $\lambda \neq A_j - B_j$ *for all* $j = 1, \ldots, k_n$. *Then,*

$$\Pr\left[\bigcap_{j=1}^{k_n} \left\{N_j \leq e^{n\lambda}\right\}\right] \doteq \mathbb{1}\left\{\min_{1\leq j\leq k_n} \{B_j - A_j + [\lambda]_+\} > 0\right\}. \tag{4.72}$$

Evidently, depending on the parameters $\{(A_j, B_j)\}_{j=1}^{k_n}$, this probability exhibit a sharp transition from being zero on the exponential scale, that is, decaying super-exponentially to being asymptotically 1. In the former case, this is due to a single tail event whose probability decays super-exponentially, and in the latter case, this is due to the exponential decay of the probability of each of the events $\{N_j > e^{n\lambda}\}_{j=1}^{k_n}$ (see Appendix D for a formal proof).

4.4 Advanced Coding Problems

In this section, we demonstrate how the TCEM can be used in various advanced coding problems. We show how the error probability in these problems can be expressed via suitably defined TCEs, which share similar properties to the TCE analyzed in the previous section. For brevity, we will not state here the resulting exponents – they can be found in the cited references – and typically provide the exact exponent for the random ensemble of interest. Moreover, oftentimes the exponents of the TCEM are also the best possible known, and are strictly better than exponent bounds obtained via classic bounding techniques.

4.4.1 Superposition Coding

We begin with the asymmetric broadcast channel (or a broadcast channel with a degraded message set) [16], [33], [73], [106], [107], which is a prototypical example for a multiuser channel [59]. This setting introduces new aspects, for which the TCEM is especially useful in deriving exact random-coding error exponents. We focus on a simple version of this setting, in which a single transmitter wishes to communicate different messages to two receivers with different channels, and so possibly different point-to-point capacities. The first channel is referred to as the *strong user* channel, and the second as the *weak user* channel, for reasons to be made clear shortly. We denote the strong user (resp. weak user) channel by W_y (resp. W_z), which is from the input alphabet $\mathcal{X}$ to the output alphabet $\mathcal{Y}$ (resp. $\mathcal{Z}$).

Rather than drawing a regular random code for this channel, Cover [33] and then Bergman [16], proposed to use *superposition coding*, or a *hierarchical codebook*. In this coding method, the rate is split as $R = R_z + R_y$, and the message is thus determined by two indices. In the random coding regime, the codebook is constructed as follows. An auxiliary alphabet $\mathcal{U}$ is chosen along with a joint input type P_{UX}. Then, e^{nR_z} *cloud centers* $\tilde{\mathfrak{C}}_n = \{\boldsymbol{U}_i\}_{i=1}^{e^{nR_z}}$ are drawn from the fixed-composition ensemble of input type P_U. For each cloud center, a sub-codebook $\mathfrak{C}_n(i) = \{\boldsymbol{X}_{i,j}\}_{j=1}^{e^{nR_y}}$ of *satellite codewords* is chosen uniformly at random from the conditional type class $\mathcal{T}_n(P_{X|U}|\boldsymbol{U}_i)$. Alternatively, this sub-codebook is referred to as a *bin* [127]. The random codebook is then $\mathfrak{C}_n = \bigcup_{i=1}^{e^{nR_z}} \mathfrak{C}_n(i) = \{\boldsymbol{X}_{i,j}\}$ which has size $e^{n(R_y+R_z)}$ and thus is capable of sending messages at a total rate of $R = R_y + R_z$. The weak user is only intended to decode the sub-codebook, that is, to decide which sub-codebook $\{\mathcal{C}_n(i)\}$ contains the transmitted message, and thus achieve a rate of R_z (the rate of the common message, indexed by i). The strong user decodes the codeword and achieves the total rate R (the common and the private message, indexed by (i, j)).

Let us focus on the weak user. For a given hierarchical codebook $\mathcal{C}_n$, the ML decoder, which minimizes the error probability of the weak user, uses the likelihood

$$W_z\left(\boldsymbol{Z}|\mathcal{C}_n(i)\right) \triangleq \Pr\left[\boldsymbol{Z} = \boldsymbol{z}|i\right] \tag{4.73}$$

$$= \frac{1}{e^{nR_y}} \sum_{j=1}^{e^{nR_y}} W_z(\boldsymbol{z}|\boldsymbol{x}_{i,j}) \tag{4.74}$$

$$= \frac{1}{e^{nR_y}} \sum_{j=1}^{e^{nR_y}} e^{n\cdot\alpha_z(\hat{Q}_{\boldsymbol{z},\boldsymbol{x}_{i,j}})}, \tag{4.75}$$

with the choice $\alpha_z(Q) = f_z(Q) \triangleq \mathbb{E}_Q[\ln W_z(\boldsymbol{Z}|\boldsymbol{X})]$, that is, using the true channel likelihood. One can also replace this choice with a channel-independent one, *e.g.*, that of the MMI rule $\alpha_z(Q) = I(Q_{XZ})$. In any case, the score of this decoder for a single message i is comprised of a sum over an *exponential* number of e^{nR_y} satellite codewords. The complex structure of this decoding rule compared to the standard ML decoding rule substantially complicates the analysis of the error exponent. Indeed, for the point-to-point channel considered before, an error event from $\boldsymbol{x}$ to $\tilde{\boldsymbol{x}}$ given output $\boldsymbol{y}$ occurs for the ML decoder whenever $W(\boldsymbol{y}|\tilde{\boldsymbol{x}}) \geq W(\boldsymbol{y}|\boldsymbol{x})$. This event can be expressed using the corresponding joint types as $f(\hat{Q}_{\tilde{\boldsymbol{x}}\boldsymbol{y}}) \geq f(\hat{Q}_{\boldsymbol{x}\boldsymbol{y}})$, and directly leads to a simple constraint $f(\tilde{Q}_{XY}) \geq f(Q_{XY})$ in (4.24) (when $\alpha(\cdot) = f(\cdot)$). However, $W_z(\cdot|\mathcal{C}_n(i))$ is *not* a memoryless channel, and so the event of making an error from i to $\tilde{i}$, that is, $W_z(\boldsymbol{Z}|\mathcal{C}_n(\tilde{i})) \geq W_z(\boldsymbol{Z}|\mathcal{C}_n(i))$ cannot be expressed as a simple relation between types as before. A naive use of a union bound or Jensen-type inequalities to analyze this sum of exponential number of terms typically fails in providing the exact exponent. The TCEM ameliorates this by partitioning the summation over the e^{nR_y} private codewords of the strong user according to their joint type $\hat{Q}_{\boldsymbol{z},\boldsymbol{x}_{i,j}}$, thus transforming the sum over an exponential number of likelihoods to a sum over a polynomial number of average likelihoods. To show this, we next evaluate the ensemble-average error probability of the weak user. Nonetheless, we will do this in a slightly different way compared to standard channel coding, in order to demonstrate another technique.

We assume, without loss of generality, that the first codeword $(1,1)$ is transmitted, and thus fix $(\boldsymbol{U}_1, \mathfrak{C}_n(1)) = (\boldsymbol{u}_1, \mathcal{C}_n(1))$ as well as the output vector $\boldsymbol{Z} = \boldsymbol{z}$. The error probability conditioned on these RVs is given by

$$\Pr\left[\bigcup_{i=2}^{e^{nR_z}} \{W_z(\boldsymbol{z}|\mathfrak{C}_n(i)) \geq W_z(\boldsymbol{z}|\mathcal{C}_n(1))\}\right], \tag{4.76}$$

where $\mathfrak{C}_n(i)$ is the random code for the ith bin. Now, we use the fact that the truncated union bound is exponentially tight for pairwise independent events. That is, if $\{\mathcal{E}_m\}$ are *pairwise independent* events then (*e.g.*, [186, p. 109, Lemma A.2], or from the de Caen inequality [48])

$$\frac{1}{2} \cdot \min\left\{\sum_m \Pr\left[\mathcal{E}_m\right], 1\right\} \leq \Pr\left[\bigcup_m \mathcal{E}_m\right] \leq \min\left\{\sum_m \Pr\left[\mathcal{E}_m\right], 1\right\}. \tag{4.77}$$

Importantly, the number of events is arbitrary and could be exponential, while still preserving exponential tightness. Such a bound can be further generalized, as was shown in [175, Appendix A], to bounds on the probability of a multiply-indexed unions. Exploiting this result and the fact that the events in (4.76) are independent, we obtain that the probability is exponentially equal to

$$\min\left\{e^{nR_z} \cdot \Pr\left[W_z(\boldsymbol{z}|\mathfrak{C}_n(2)) \geq W_z(\boldsymbol{z}|\mathcal{C}_n(1))\right], 1\right\}. \tag{4.78}$$

We may now focus on the inner probability, and as $\boldsymbol{z}$ and $\mathcal{C}_n(1)$ are fixed at this moment, we set, for brevity, $s(\mathcal{C}_n(1), \boldsymbol{z}) = \frac{1}{n} \ln W_z(\boldsymbol{z}|\mathcal{C}_n(1))$. We evaluate this probability in two steps. First, we condition on a competing cloud-center codeword, say $\boldsymbol{U}_2 = \boldsymbol{u}_2$, and compute the conditional probability according to a random choice of $\mathfrak{C}_n(2)$. To this end, we define a suitable TCE.

Definition 4.3 (TCE for random-coding exponent of superposition codes)**.** For a superposition codebook $\mathcal{C}_n = \{\boldsymbol{x}_{i,j}\}$, a cloud center $\boldsymbol{u}$, an output vector $\boldsymbol{z}$ and a joint type Q_{UXZ} such that $\hat{Q}_{\boldsymbol{uz}} = Q_{UZ}$, let

$$N_{\boldsymbol{u},\boldsymbol{z}}(Q_{UXZ}, \mathcal{C}_n(i)) \triangleq \left|\left\{1 \leq j \leq e^{nR_y} \colon (\boldsymbol{u}, \boldsymbol{x}_{i,j}, \boldsymbol{z}) \in \mathcal{T}_n(Q_{UXZ})\right\}\right|. \quad (4.79)$$

This TCE counts the number of codewords in a single bin $\mathcal{C}_n(i)$ of a superposition code defined by the cloud center $\boldsymbol{u}$, which have a joint type Q_{UXZ} with $\boldsymbol{z}$. By the method of types (a generalization of (2.12) from Section (2.2.1) to conditional types, see [41, Lemma 2.13]), whenever $\boldsymbol{X}_j \sim \text{Uniform}[\mathcal{T}_n(Q_{X|U}|\boldsymbol{u})]$ it holds that

$$\Pr\left[(\boldsymbol{u}, \boldsymbol{X}_j, \boldsymbol{z}) \in \mathcal{T}_n(Q_{UXZ})\right] = k_n \cdot e^{-nI_Q(X;Z|U)} \quad (4.80)$$

for some $k_n \doteq 1$. So, for a random codebook $\mathfrak{C}_n(i) = \{\boldsymbol{X}_{i,j}\}$,

$$N_{\boldsymbol{u},\boldsymbol{z}}(Q_{UXZ}, \mathfrak{C}_n(i)) = \sum_{j=1}^{e^{nR_y}} \mathbb{1}\left\{(\boldsymbol{u}, \boldsymbol{X}_j, \boldsymbol{z}) \in \mathcal{T}_n(Q_{UXZ})\right\} \sim \text{Binomial}\left(e^{nR_y}, k_n \cdot e^{-nI_Q(X;Z|U)}\right). \quad (4.81)$$

As before, we will omit for brevity $\mathcal{C}_n(i)$ and $(\boldsymbol{u}, \boldsymbol{z})$ from the notation of the TCE, as it does not affect its probability distribution. With Definition 4.3, the probability of interest takes the form

$$\Pr\left[W_z(\boldsymbol{z}|\mathfrak{C}_n(2)) \geq e^{n\cdot s(\mathcal{C}_n(1),\boldsymbol{z})}\right]$$
$$= \Pr\left[\sum_{Q_{UXZ}} N(Q_{UXZ}) \geq e^{n\cdot s(\mathcal{C}_n(1),\boldsymbol{z})}\right] \quad (4.82)$$
$$\doteq \Pr\left[\max_{Q_{UXZ}} N(Q_{UXZ}) \geq e^{n\cdot s(\mathcal{C}_n(1),\boldsymbol{z})}\right] \quad (4.83)$$
$$\doteq \max_{Q_{UXZ}} \Pr\left[N(Q_{UXZ}) \geq e^{n\cdot s(\mathcal{C}_n(1),\boldsymbol{z})}\right], \quad (4.84)$$

where we have used a summation–maximization equivalence for probabilities, which can be derived analogously to (4.46). Without delving into details, we briefly mention that the sub-exponential factors that are hidden but the exponential equalities in (4.83) and (4.84) do not depend on $s(\mathcal{C}_n(1), \boldsymbol{z})$; a quantity which is a function of n on its own. Evidently, the next step requires evaluating the tail probability of $N(Q_{UXZ})$. This can be done, by generalizing Theorem 4.1 (Section 4.3) to the TCEs of superposition codes. Following this step, an exponentially-tight expression for the probability (4.84) is obtained, which implicitly depends on the choice of $\boldsymbol{U}_2$. The next step is thus to average this exponentially decaying probability over $\boldsymbol{U}_2$ by considering joint types Q_{UZ} agreeing with $\hat{Q}_{\boldsymbol{U}_2,\boldsymbol{z}}$. This average is dominated by the minimal exponent over all possible types.

The above evaluation of the error probability is for a fixed output vector $\boldsymbol{z}$ and the sub-codebook $\mathcal{C}_n(1)$. To obtain the average error probability, the next step is to average over $(\boldsymbol{Z}, \mathfrak{C}_n(1))$. Handling the randomness of $s(\mathfrak{C}_n(1), \boldsymbol{Z})$ is again obtained with similarly defined TCEs. However, at this step, the inequality defining the event of interest is in a reversed direction compared to (4.82). Thus, when proceeding this way, one encounters for some $t \in \mathbb{R}$ the following expression

$$\Pr\left[\sum_{Q_{UXZ}} N(Q_{UXZ}) e^{n\cdot\alpha_z(Q_{UXZ})} \le e^{n\cdot t}\right]$$

$$\stackrel{(a)}{\doteq} \Pr\left[\max_{Q_{UXZ}} N(Q_{UXZ}) e^{n\cdot\alpha_z(Q_{UXZ})} \le e^{n\cdot t}\right] \tag{4.85}$$

$$= \Pr\left[\bigcap_{Q_{UXZ}} \left\{N(Q_{UXZ}) \le e^{n\cdot[t-\alpha_z(Q_{UXZ})]}\right\}\right] \tag{4.86}$$

$$\stackrel{(b)}{\doteq} \mathbb{1}\left\{\min_{Q_{UXZ}} \left\{I_Q(X;Z|U) - R_y + [t-\alpha_z(Q_{UXZ})]_+\right\} > 0\right\}, \tag{4.87}$$

where here $N(Q_{UXZ})$ is defined as in (4.79), yet for $\mathfrak{C}_n(1)$, and excluding $\boldsymbol{x}_{1,1}$, where (a) follows from the summation–maximization equivalence for probabilities (as above), and (b) follows from Theorem 4.3. This exemplifies the necessity to evaluate the probability of the intersection of tail events of multiple enumerators in the course of the analysis, and also shows the possibility to evaluate this probability accurately on an exponential scale. The full details of such derivations can be found in, *e.g.*, [127, Proof of Theorem 1], [11, Section 5.1].

4.4.2 Distributed Compression and Random Binning

Our next setting pertains to a source coding problem, and specifically, the Slepian–Wolf (SW) problem of distributed lossless compression [189]. In this problem, a source X with a finite alphabet $\mathcal{X}$ is given at the encoder side, and side-information Y of a finite alphabet $\mathcal{Y}$ that is at the decoder side. The pair (X, Y) is correlated, and follows a joint distribution P_{XY}, and vectors $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{X}^n \times \mathcal{Y}^n$ are emitted from P_{XY}, with IID pairs of symbols. The source vector $\boldsymbol{X}$ is compressed by assigning it to an index $Z = f(\boldsymbol{X})$ of one of e^{nR} possible bins, where $f\colon \mathcal{X}^n \to \{1, 2, \ldots, e^{nR}\}$ is a called a *binning* rule. Given the side-information $\boldsymbol{Y} = \boldsymbol{y}$ and the bin $Z = z$, the decoder decides that the source vector is

$$\hat{\boldsymbol{x}}(\boldsymbol{y}, z) = \operatorname*{arg\,max}_{\boldsymbol{x} \in f^{-1}(z)} \Pr\left[\boldsymbol{X} = \boldsymbol{x} | \boldsymbol{Y} = \boldsymbol{y}\right] \tag{4.88}$$

$$\triangleq \operatorname*{arg\,max}_{\boldsymbol{x} \in f^{-1}(z)} \Pr\left[\boldsymbol{x} | \boldsymbol{y}\right]. \tag{4.89}$$

For a giving binning rule f, the error probability is then given by

$$P_{\mathsf{e}}(f) \triangleq \sum_{\boldsymbol{x} \in \mathcal{X}^n} \sum_{\boldsymbol{y} \in \mathcal{Y}^n} \Pr\left[\boldsymbol{X} = \boldsymbol{x}, \boldsymbol{Y} = \boldsymbol{y}\right] \times \mathbb{1}\left[\bigcup_{\boldsymbol{x}' \neq \boldsymbol{x}\colon \Pr[\boldsymbol{x}'|\boldsymbol{y}] \geq \Pr[\boldsymbol{x}|\boldsymbol{y}]} \{f(\boldsymbol{x}') = f(\boldsymbol{x})\}\right]. \tag{4.90}$$

For a joint type Q_{XY}, let us further denote the expected log-posterior as

$$g(Q_{XY}) \triangleq \mathbb{E}_Q \left[\ln P_{X|Y}(X|Y)\right], \tag{4.91}$$

so that for any $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{T}_n(Q_{XY})$ it holds that $\Pr[\boldsymbol{x}|\boldsymbol{y}] = e^{ng(Q_{XY})}$.

As expected, it is intractable to find the optimal binning rule f. However, whenever the compression rate R is above the minimal required rate $H(X|Y)$, the ensemble average of the error probability over random choice of binning functions decays exponentially [2], [30], [40], [74], [140], [141], [160], [217]. In fact, the optimal error exponent for the SW problem is directly related to the random-coding error exponent in channel coding (see, *e.g.*, [2], [30], [217]). Thus, we may evaluate the error probability averaged over random choice of binning rules, referred to as *random binning*. Accordingly, the exponential decay of the average error probability is called the *random-binning error exponent*. In simple random binning, the random rule F is such that the bin of any $\boldsymbol{x} \in \mathcal{X}^n$ is chosen uniformly at random from the e^{nR} possible bins. Analogously to TCEs for channel coding, we define the following TCE:

Definition 4.4 (TCE for random-binning exponent)**.** For a binning rule f, a side-information vector $\boldsymbol{y}$, an encoded index z, and a joint type Q'_{XY} such that $\hat{Q}_{\boldsymbol{y}} = Q'_Y$, let

$$\tilde{N}_{\boldsymbol{y},z}(Q'_{XY}, f) \triangleq \left|\left\{(\boldsymbol{x}', \boldsymbol{y}) \in \mathcal{T}_n(Q'_{XY}) \cap f^{-1}(z)\right\}\right|. \tag{4.92}$$

The TCE $\tilde{N}_{\boldsymbol{y},z}(Q_{XY}, f)$ counts the number of vectors in $\mathcal{T}_n(Q_X)$ except for $\boldsymbol{x}$, which have joint type Q'_{XY} with $\boldsymbol{y}$, and that are mapped to the index z. Intuitively, the sum of $\tilde{N}_{\boldsymbol{y},z}(Q_{XY}, f)$ over all joint types Q'_{XY} counts all the source vectors which may be erroneously decided to be the true source vector instead of $\boldsymbol{x}$, whenever the bin index is z and the side information is $\boldsymbol{y}$. By the method of types [(2.12) in Section 2.2.1]

$$|\mathcal{T}_n(Q'_{XY})| = k_n \cdot e^{nH(Q'_{XY})} \tag{4.93}$$

for some $k_n \doteq 1$. So, for a random binning rule

$$F(\boldsymbol{x}') \sim \text{Uniform}\left\{1, 2, \ldots, e^{nR}\right\},$$

it holds that

$$\begin{aligned}\tilde{N}_{\boldsymbol{y},z}(Q'_{XY}, F) &= \sum_{\boldsymbol{x}' \in \mathcal{T}_n(Q_X):\, \boldsymbol{x}' \neq \boldsymbol{x}} \mathbb{1}\left\{(\boldsymbol{x}', \boldsymbol{y}) \in \mathcal{T}_n(Q'_{XY}) \cap F^{-1}(z)\right\} \\ &\sim \text{Binomial}\left(k_n \cdot e^{nH_{Q''}(X|Y)}, e^{-nR}\right).\end{aligned} \tag{4.94}$$

This TCE displays a plausible *duality* with the TCE of channel coding: In random coding the number of trials is fixed and the success probability is type-dependent, whereas in random binning, it is the other way round.

As in channel-coding analysis, we will simplify the notation to $\tilde{N}_{\boldsymbol{y},z}(Q'_{XY})$ or even just $\tilde{N}(Q'_{XY})$ when the TCE is an RV. Given Definition 4.4, the random-binning error probability is then given by

$$\overline{P}_{\mathsf{e}} \triangleq \mathbb{E}\left[P_{\mathsf{e}}(F)\right]$$
$$= \sum_{\boldsymbol{x}\in\mathcal{X}^n} \sum_{\boldsymbol{y}\in\mathcal{Y}^n} \Pr\left[\boldsymbol{X}=\boldsymbol{x}, \boldsymbol{Y}=\boldsymbol{y}\right] \times$$
$$\Pr\left[\bigcup_{Q'_{XY}\colon Q'_Y=\hat{Q}_{\boldsymbol{y}},\, g(Q'_{XY})\geq g(\hat{Q}_{\boldsymbol{xy}})} \left\{\tilde{N}_{\boldsymbol{y},f(\boldsymbol{x})}(Q'_{XY}) \geq 1\right\}\right] \tag{4.95}$$
$$\stackrel{(a)}{\doteq} \sum_{Q_{XY}} \Pr\left[(\boldsymbol{X},\boldsymbol{Y}) \in \mathcal{T}_n(Q_{XY})\right] \times$$
$$\sum_{Q'\colon Q'_Y=Q_Y,\, g(Q')\geq g(Q)} \Pr\left[\tilde{N}(Q'_{XY}) \geq 1\right] \tag{4.96}$$
$$\stackrel{(b)}{\doteq} \max_{Q_{XY}} \max_{Q'_{XY}\colon Q'_Y=Q_Y,\, g(Q'_{XY})\geq g(Q_{XY})} e^{-nD(Q_{XY}||P_{XY})} \times$$
$$\Pr\left[\tilde{N}(Q'_{XY}) \geq 1\right], \tag{4.97}$$

where (a) follows from the summation–maximization equivalence for probabilities, and (b) follows from the method of types (Section 2.2.1). The tail probability $\Pr[\tilde{N}(Q'_{XY}) \geq 1]$ can be analyzed as in Section 4.3, and this results the exact random-binning exponent.

4.4.3 Generalized Decoders

Erasure/List Decoders. An erasure decoder may either decode a message or declare an *erasure*, that is, not to output any message. A list decoder may output multiple codewords, whose number is either fixed in advance, or varies according to the channel output. Forney showed in [68] that both Pareto-optimal erasure decoder and *variable* list-size decoder have a similar form, which uses the posterior probability, rather than the likelihood, to decide on its output. We begin by describing an erasure decoder. Concretely, consider a codebook $\mathcal{C}_n = \{\boldsymbol{x}_m\}$ from which a codeword $\boldsymbol{X}$ is chosen under the uniform distribution. The codeword is then transmitted over a channel W, and given an output vector $\boldsymbol{Y} = \boldsymbol{y}$, the posterior probability of the mth codeword is given by Bayes rule as

$$\Pr\left[\boldsymbol{X}=\boldsymbol{x}_m|\boldsymbol{Y}=\boldsymbol{y}\right] = \frac{W\left[\boldsymbol{y}|\boldsymbol{x}_m\right]}{\sum_{m'} W\left[\boldsymbol{y}|\boldsymbol{x}_{m'}\right]}. \tag{4.98}$$

If the maximal posterior over codewords is large enough, then the maximizing codeword is decoded. Otherwise, an erasure is declared. Equivalently, we may set a threshold parameter $T > 0$, so that the optimal erasure decoder outputs message m if $\boldsymbol{x}_m$ is the (unique) codeword such that

$$\frac{W\left[\boldsymbol{y}|\boldsymbol{x}_m\right]}{\sum_{m' \neq m} W\left[\boldsymbol{y}|\boldsymbol{x}_{m'}\right]} > e^{nT}. \tag{4.99}$$

The threshold parameter T determines the trade-off between two types of failure events: An *erasure event* $\mathcal{E}_1'(\mathcal{C}_n)$, in which no codeword is decoded, or an *undetected error event* $\mathcal{E}_2(\mathcal{C}_n)$, in which a wrong codeword is decoded. The event $\mathcal{E}_1(\mathcal{C}_n) = \mathcal{E}_1'(\mathcal{C}_n) \cup \mathcal{E}_2(\mathcal{C}_n)$ is then called the *total-error event*. As can be seen, the score of the decoder is a complicated function, since the denominator in (4.99) includes a summation over an exponential number of likelihoods.

To bound the probability of the total-error event, Forney has introduced a parameter $s \in [0, 1]$ and derived a Chernoff-style bound. As said, this bounding method is not guaranteed to be tight, and indeed leads to strictly sub-optimal exponents. The TCEM addresses the problem of evaluating the probability of the total-error event, by using the TCE $N(\tilde{Q}_{XY})$ of Definition 4.1. For a random codebook $\mathfrak{C}_n$,

$$\begin{aligned}
&\mathbb{E}\left\{\Pr\left[\mathcal{E}_1(\mathfrak{C}_n)\right]\right\} \\
&= \mathbb{E}\left\{\Pr\left[\mathcal{E}_1(\mathfrak{C}_n)|\boldsymbol{X}_1 \text{ transmitted}\right]\right\} && (4.100) \\
&= \Pr\left[\sum_{m'>1} W\left[\boldsymbol{Y}|\boldsymbol{X}_{m'}\right] \geq e^{-nT} \cdot W\left[\boldsymbol{Y}|\boldsymbol{X}_1\right]\right] && (4.101) \\
&= \sum_{Q_{XY}} \Pr\left[(\boldsymbol{X}_1, \boldsymbol{Y}) \in \mathcal{T}_n(Q_{XY})\right] \times \\
&\qquad \Pr\left[\sum_{Q'_{XY}\,:\,\tilde{Q}_Y = Q_Y} N(Q'_{XY}) e^{nf(Q'_{XY})} \geq e^{-nT} \cdot e^{-nf(Q_{XY})}\right]. && (4.102)
\end{aligned}$$

The first probability is given by the standard method of types [(2.12) in Section 2.2.1], and the second probability may be evaluated by the summation–maximization equivalence

$$\Pr\left[\sum_{Q'_{XY}:\, Q'_Y=Q_Y} N(Q'_{XY})e^{nf(Q'_{XY})} \geq e^{-nT}\cdot e^{-nf(Q_{XY})}\right]$$

$$\doteq \Pr\left[\max_{Q'_{XY}:\, Q'_Y=Q_Y} N(Q'_{XY})e^{nf(Q'_{XY})} \geq e^{-nT}\cdot e^{-nf(Q_{XY})}\right] \quad (4.103)$$

$$= \Pr\left[\bigcup_{Q'_{XY}:\, Q'_Y=Q_Y} \left\{N(Q'_{XY})e^{nf(Q'_{XY})} \geq e^{-nT}\cdot e^{-nf(Q_{XY})}\right\}\right] \quad (4.104)$$

$$\doteq \max_{Q'_{XY}:\, Q'_Y=Q_Y} \Pr\left[N(Q'_{XY}) \geq e^{-n[f(Q_{XY})-f(Q'_{XY})-T]}\right]. \quad (4.105)$$

The derivation is completed by the exact exponential analysis of the tail probability of $N(Q'_{XY})$ from Section 4.3. It can be shown that the resulting random-coding error exponent of $\mathbb{E}[\mathcal{E}_2(\mathfrak{C}_n)]$ is larger by exactly T than the exponent of the total-error event [91]. Thus, the exponent of the total-error event is equal to that of the erasure event.

In the setting of *variable* list size decoding, the Pareto optimal decoder also takes the form of (4.99), albeit with the threshold parameter set to some $T < 0$. With such a choice, the codeword that satisfy (4.99) may not be unique, and so the rule of (4.99) defines a variable list size decoder. In this regime, the trade-off is between the exponent of the error event (where the correct codeword is not in the output list), and the normalized logarithm of the expected list size. It can be shown that the expressions for these values are exactly as the ones of the total-error exponent and undetected error exponent in the erasure regime $T > 0$, and so the analysis is identical, while just allowing $T < 0$.

Likelihood Decoders. Similarly to an erasure/list decoder, a likelihood decoder [227] also uses the posterior probability. However, it outputs a *random* codeword from this posterior, so the decoded message is m with probability

$$\Pr[\boldsymbol{x}_m|\boldsymbol{y}] = \frac{W(\boldsymbol{y}|\boldsymbol{x}_m)}{\sum_{\tilde{m}=1}^{e^{nR}} W(\boldsymbol{y}|\boldsymbol{x}_{\tilde{m}})}. \quad (4.106)$$

More generally, one may choose a continuous function $g(Q_{XY})$ of joint types and an inverse-temperature $\beta > 0$, and consider a likelihood decoder that decodes message m with probability

$$\Pr\left[\boldsymbol{x}_m|\boldsymbol{y}\right] = \frac{\exp\left[n\beta g(\hat{Q}_{\boldsymbol{x}_m,\boldsymbol{y}})\right]}{\sum_{\tilde{m}=1}^{e^{nR}} \exp\left[n\beta g(\hat{Q}_{\boldsymbol{x}_{\tilde{m}},\boldsymbol{y}})\right]}. \tag{4.107}$$

When $\beta = 1$ and $g(Q_{XY}) = f(Q_{XY}) = \mathbb{E}_Q[\ln W(Y|X)]$ then the *ordinary likelihood decoder* (4.106) is reproduced. However, $g(\cdot)$ can be replaced by a choice that is mismatched to the channel, or even by a universal function such as $g(Q_{XY}) = I(Q_{XY})$. Similar to finite-temperature decoding [172], the parameter β controls the "amount of stochasticity" of the decoder: If $\beta \to \infty$ then the decoder becomes deterministic, reproducing the score-based decoder with $\alpha \equiv g$. As the temperature increases and β decreases, the decoder becomes more random (at the extreme $\beta = 0$, the output codeword is chosen uniformly at random). On the upside, for any fixed β, the ensemble average error probability follows a remarkably simple formula, given by

$$\overline{P_{\mathsf{e}}} = \mathbb{E}\left[P_{\mathsf{e}}(\mathfrak{C}_n)\right] = \mathbb{E}\left[P_{\mathsf{e}}\left(\mathfrak{C}_n|\boldsymbol{X}_1 \text{ transmitted}\right)\right] \tag{4.108}$$

$$\stackrel{(a)}{=} \mathbb{E}\left[1 - \Pr\left[\boldsymbol{X}_1|\boldsymbol{Y}\right]\right] \tag{4.109}$$

$$\stackrel{(b)}{=} \mathbb{E}\left[\frac{\sum_{m=2}^{e^{nR}} \exp\left[n\beta g(\hat{Q}_{\boldsymbol{X}_m,\boldsymbol{Y}})\right]}{\sum_{m=1}^{e^{nR}} \exp\left[n\beta g(\hat{Q}_{\boldsymbol{X}_m,\boldsymbol{Y}})\right]}\right], \tag{4.110}$$

where in (a) $\boldsymbol{Y}$ is the output the channel W when $\boldsymbol{X}_1$ is the input, and (b) follows from the decoding rule (4.107). On the downside, in this expression, *both* the numerator and denominator contain an exponential number of codewords, and this makes its analysis more challenging. Following the TCEM, let us condition on the joint type $(\boldsymbol{X}_1, \boldsymbol{Y}) \in \mathcal{T}_n(Q_{XY})$. Then, using the TCE of Definition 4.1, it holds that

$$\overline{P_{\mathsf{e}}} = \sum_{Q_{XY}} \Pr\left[(\boldsymbol{X}_1, \boldsymbol{Y}) \in \mathcal{T}_n(Q_{XY})\right] \cdot e(Q_{XY}) \tag{4.111}$$

where, for any given Q_{XY},

$$e(Q_{XY}) = \mathbb{E}\left[\frac{\sum_{\tilde{Q}_{XY}:\, \tilde{Q}_Y = Q_Y} N(\tilde{Q}_{XY}) e^{n\beta g(\tilde{Q}_{XY})}}{e^{n\beta g(Q_{XY})} + \sum_{\tilde{Q}_{XY}:\, \tilde{Q}_Y = Q_Y} N(\tilde{Q}_{XY}) e^{n\beta g(\tilde{Q}_{XY})}}\right]. \tag{4.112}$$

This expectation can be evaluated as:

$$
\begin{aligned}
&e(Q_{XY}) \\
&\doteq \mathbb{E}\left[\min\left\{\sum_{\tilde{Q}_{XY}} N(\tilde{Q}_{XY})e^{n[\beta g(\tilde{Q}_{XY})-\beta g(Q_{XY})]}, 1\right\}\right] && (4.113) \\
&\stackrel{(a)}{=} \int_0^\infty \Pr\left[\min\left\{\sum_{\tilde{Q}_{XY}} N(\tilde{Q}_{XY})e^{n[\beta g(\tilde{Q}_{XY})-\beta g(Q_{XY})]}, 1\right\} \geq t\right] \mathrm{d}t && (4.114) \\
&\stackrel{(b)}{=} \int_0^1 \Pr\left[\sum_{\tilde{Q}_{XY}} N(\tilde{Q}_{XY})e^{n[\beta g(\tilde{Q}_{XY})-\beta g(Q_{XY})]} \geq t\right] \mathrm{d}t && (4.115) \\
&= n\int_0^\infty e^{-n\theta}\cdot \Pr\left[\sum_{\tilde{Q}_{XY}} N(\tilde{Q}_{XY})e^{n[\beta g(\tilde{Q}_{XY})-\beta g(Q_{XY})]} \geq e^{-n\theta}\right] \mathrm{d}\theta && (4.116) \\
&\doteq \max_{\tilde{Q}_{XY}} n\int_0^\infty e^{-n\theta}\cdot \Pr\left[N(\tilde{Q}_{XY}) \geq e^{-n[\theta-\beta g(\tilde{Q}_{XY}+\beta g(Q_{XY}))]}\right] \mathrm{d}\theta, && (4.117)
\end{aligned}
$$

where in the summations above and in the maximization on the final line, $\tilde{Q}_{XY}$ is such that the constraint $\tilde{Q}_Y = Q_Y$ holds, (a) follows from the tail-integration identity $\mathbb{E}[X] = \int_0^\infty \Pr[X \geq t]\ \mathrm{d}t$, which holds for any non-negative RV X, and (b) follows from the summation–maximization equivalence (or, alternatively, as $\frac{1}{2}\min\{a,1\} \leq \frac{a}{a+1} \leq \min\{a,1\}$). The derivation continues by plugging into the integral the tight exponent of the tail probability of $N(\tilde{Q}_{XY})$ from Section 4.3. Then, the exponential decay rate of the integral can be determined using Laplace method from Section 3. After averaging WRT to $(\boldsymbol{X}_1, \boldsymbol{Y})$, the resulting expression may be minimized over Q_{XY} to obtain the exact exponent of the ensemble-average error probability. The full details of this derivation can be found in [131], [133].

4.4.4 Error Exponent of the Typical Random Code

As discussed above, the random-coding error exponent (4.9) may serve two purposes. First, it can be used to serve as performance measure for a communication system that actually randomly selects its codebook. Second, it is a lower bound on the error exponent of the optimal sequence of codes, *i.e.*, it is a technique to prove achievability bounds. Inspecting the definition of the random-coding error exponent (4.9), it is appears to be somewhat at odds with both goals. From the perspective of the first goal, a direct way is to evaluate the average error exponent, or the error exponent of the *typical random code* is

$$E_{\mathrm{trc}}(R) \triangleq \mathbb{E}\left[-\frac{1}{n}\ln P_{\mathsf{e}}(\mathfrak{C}_n)\right], \tag{4.118}$$

which is the ensemble average of the exponent, rather than the exponent of the average error probability, as in (4.9). From the perspective of the second goal, Jensen's inequality assures that $E_{\mathrm{trc}}(R) \geq E_{\mathrm{rc}}(R)$, and so this exponent leads to tighter achievability bounds than the standard random-coding exponent. Indeed, the root of this relation is that the exponent $E_{\mathrm{trc}}(R)$ is determined by *typical* codebooks, whereas the random-coding error exponent is actually dominated by *unlikely* poor codebooks. This can be seen from the following informal argument. Let $\mathcal{G}_E$ be the collection of codes $\{\mathcal{C}_n\}$ for which $P_{\mathsf{e}}(\mathcal{C}_n) \approx e^{-nE}$. Then, approximating the values of E by a discrete fine grid, results in

$$\mathbb{E}[P_{\mathsf{e}}(\mathfrak{C}_n)] \approx \sum_E \Pr\left[\mathfrak{C}_n \in \mathcal{G}_E\right] \cdot e^{-nE}. \tag{4.119}$$

This term is dominated by the largest term in the sum, yet the maximizer may occur for $\tilde{E}$ in which $\Pr[\mathfrak{C}_n \in \mathcal{G}_{\tilde{E}}]$ is exponentially small. Thus, codebooks with exponentially small low probability to occur may dominate the random-coding exponent. In contrast, it holds that

$$E_{\mathrm{trc}}(R) = \mathbb{E}\left[-\frac{1}{n}\ln P_{\mathsf{e}}(\mathcal{C}_n)\right] = \sum_E \Pr\left[\mathfrak{C}_n \in \mathcal{G}_E\right] \cdot E. \tag{4.120}$$

Thus, if there exists a value E_0 for which $\Pr[\mathfrak{C}_n \in \mathcal{G}_E] \to 1$ then $E_{\mathrm{trc}}(R) = E_0$ (and such E_0 does exist). Thus, the error exponent of the *typical random code* is determined by codebooks which are highly likely to occur. Evidently, the averaging of the normalized logarithm over the error probability mitigates the effect of high error-probability codebooks on the ensemble average.

Despite this obvious advantage, the error exponent of the typical random code was considered to be more difficult to evaluate than the random-coding error exponent, and thus was somewhat ignored in the traditional developments of bounds on the reliability function. In [14], Barg and Forney evaluated the error exponent of the typical random code for the BSC (and credit [15] for inspiration). The derivation is sufficiently simple to be done directly, and involves the typical *distance spectrum* of the code. The distance spectrum is defined as the number of pairs of codewords for each possible Hamming distance. Concretely, it is given by $\{\overline{N}(d)\}_{d=0}^{n}$ where

$$\overline{N}(d) \triangleq |\{m_1, m_2 \colon m_1 \neq m_2,\ d_{\mathrm{H}}(\boldsymbol{x}_{m_1}, \boldsymbol{x}_{m_2}) = d\}|\,, \tag{4.121}$$

and where $d_{\mathrm{H}}(\cdot,\cdot)$ is the Hamming distance. The error probability was then tightly bounded by the union bound as

$$P_{\mathsf{e}}(\mathcal{C}_n) \leq \sum_{d=0}^{n} \overline{N}(d) \cdot e^{-d\cdot Z(p)}, \tag{4.122}$$

where

$$Z(p) \triangleq D\left(\frac{1}{2}||p\right) = \frac{1}{2}\ln\left[4p(1-p)\right] \tag{4.123}$$

is the *Bhattacharyya distance*. The typical random exponent was determined by the typical behavior of $\overline{N}(d)$ over the ensemble, that is, a high-probability upper bound on its value (for the typical random exponent, the tightness of this bound is credited by Barg and Forney to [43]).

Evidently, the distance spectrum of the code depends on the distances between *pairs* of codewords. Therefore, the generalization of this expression to general DMCs, required for the derivation of the error exponent of the typical random code, involves a TCE that is determined by all pairs of codewords from the codebook. Specifically, it involves the following TCE:

Definition 4.5 (TCE for the error exponent of the typical random code)**.** For a codebook $\mathcal{C}_n$, and a joint type $Q_{X\tilde{X}}$, let

$$\begin{aligned}&\overline{N}(Q_{X\tilde{X}}, \mathcal{C}_n)\\ &= \left|\{(m,\tilde{m}) \colon m \neq \tilde{m},\ (\boldsymbol{x}_m, \boldsymbol{x}_{\tilde{m}}) \in \mathcal{T}_n(Q_{X\tilde{X}})\}\right|\end{aligned} \tag{4.124}$$

$$= \sum_{m=1}^{e^{nR}} \overline{N}_m(Q_{X\tilde{X}}, \mathcal{C}_n) \tag{4.125}$$

$$= \sum_{m=1}^{e^{nR}} \sum_{\tilde{m}=1}^{e^{nR}} \mathbb{1}\{\tilde{m} \neq m\} \cdot \mathbb{1}\left\{(\boldsymbol{X}_{\tilde{m}}, \boldsymbol{X}_m) \in \mathcal{T}_n(Q_{X\tilde{X}})\right\}, \tag{4.126}$$

where $\overline{N}_m(Q_{X\tilde{X}})$ is as defined in (4.35). The TCE $\overline{N}(Q_{X\tilde{X}}, \mathcal{C}_n)$ counts the number of pairs of codewords in the codebook which have a joint type $Q_{X\tilde{X}}$.

As for previous TCEs, we abbreviate the notation to $\overline{N}(Q_{X\tilde{X}})$ when it is an RV. Recall that the TCE for random-coding exponent (Definition 4.1) is a binomial RV when the codebook is random. In contrast, for a random codebook, the TCE $\overline{N}(Q_{X\tilde{X}})$ has no such simple probabilistic description, and so the derivation of the probabilistic properties $\overline{N}(Q_{X\tilde{X}})$ is more challenging. The root of this difficulty is the *dependencies* between pairs of codewords. Indeed, $\overline{N}(Q_{X\tilde{X}})$ counts the number of successes in $e^{nR}(e^{nR}-1) \doteq e^{2nR}$ trials, and the success probability of each trial is $\doteq e^{-nI(Q_{X\tilde{X}})}$. Had these trials were *statistically independent* then $\overline{N}(Q_{X\tilde{X}})$ was a binomial RV, as the TCE for random-coding exponent (Definition 4.1). However, these trials are *not* mutually independent. To explicitly see this dependence, consider the following extreme example: Let Q_X be uniform over $\mathcal{X}$ and let $Q_{X\tilde{X}}$ be the joint type that equals to $1/|\mathcal{X}|$ whenever $x = \tilde{x}$ and 0 otherwise. Then, without any prior knowledge, for every $\tilde{m} \neq m$,

$$\Pr\left[\boldsymbol{X}_m = \boldsymbol{X}_{\tilde{m}}\right] = \Pr\left[(\boldsymbol{X}_m, \boldsymbol{X}_{\tilde{m}}) \in \mathcal{T}_n(Q_{X\tilde{X}})\right] \tag{4.127}$$

$$\doteq \exp\left[-nI(Q_{X\tilde{X}})\right], \tag{4.128}$$

where $I(Q_{X\tilde{X}}) = \ln|\mathcal{X}|$. Now, conditioned on $\boldsymbol{X}_1 = \boldsymbol{X}_2$ and $\boldsymbol{X}_2 = \boldsymbol{X}_3$ it also holds with probability 1 that $\boldsymbol{X}_1 = \boldsymbol{X}_3$, thus showing their dependency. Nonetheless, this dependence is "weak", and as we will show, some of its asymptotic properties can be shown to be indifferent to this dependence, and match those of a regular Binomial$(e^{2nR}, e^{-nI(Q_{X\tilde{X}})})$ distribution.

The required moments and tail properties of $\overline{N}(Q_{X\tilde{X}})$ were evaluated as follows. In [197, Theorem 3], it was determined that for any $s \in \mathbb{R}$

$$\lim_{n\to\infty} -\frac{1}{n} \ln \Pr\left[\overline{N}(Q_{X\tilde{X}}) \geq e^{ns}\right] \geq \begin{cases} [I(Q_{X\tilde{X}}) - 2R]_+, & [2R - I(Q_{X\tilde{X}})]_+ > s \\ \infty, & [2R - I(Q_{X\tilde{X}})]_+ < s \end{cases}, \quad (4.129)$$

which is the same upper tail behavior as for a Binomial(e^{2nR}, $e^{-nI(Q_{X\tilde{X}})}$). This is intuitively justified because the events $\mathbb{1}\{(\boldsymbol{x}_{m_1}, \boldsymbol{x}_{m_2}) \in \mathcal{T}_n(Q_{X\tilde{X}})\}$ and $\mathbb{1}\{(\boldsymbol{x}_{m_3}, \boldsymbol{x}_{m_4}) \in \mathcal{T}_n(Q_{X\tilde{X}})\}$ are pairwise-independent, even if $m_1 = m_3$, and as the overall dependence between all events is fairly low. The proof of the first case in (4.129), *i.e.*, upper tail bound in the case of *populated types,* is based on bounding its integer moments, and showing that for any $k \in \mathbb{N}$

$$\mathbb{E}\left[\overline{N}^k(Q_{X\tilde{X}})\right] \dot{\leq} \begin{cases} e^{nk[2R - I(Q_{X\tilde{X}})]}, & I(Q_{X\tilde{X}}) < 2R \\ e^{n[2R - I(Q_{X\tilde{X}})]}, & I(Q_{X\tilde{X}}) > 2R \end{cases}. \quad (4.130)$$

For $k = 1$, the bound of (4.130) readily follows by the linearity of the expectation. Then, the proof of (4.130) for an arbitrary k follows by a careful induction argument over k. Once (4.130) is established, the proof of the first case in (4.129) is then completed by applying Markov's inequality with an arbitrarily large k. The proof of the second case in (4.129), *i.e.*, upper tail bound in the case of *typically empty types,* is based on *Janson's inequality* [95, Theorem 9] for the probability of the event that $\overline{N}(Q_{X\tilde{X}}) = 0$.

As for the lower tail that complements the upper tail in (4.129), it was determined in [197, Lemma 2] that given $\epsilon \in (0, 2R)$, if $I(Q_{X\tilde{X}}) \leq 2R - \epsilon$ then

$$\lim_{n\to\infty} -\frac{1}{n} \ln \Pr\left[\overline{N}(Q_{X\tilde{X}}) \leq e^{-n\epsilon} \cdot \mathbb{E}\left[\overline{N}(Q_{X\tilde{X}})\right]\right] = \infty \quad (4.131)$$

(see a more accurate statement therein). The proof is inspired by an investigation of *typicality graphs* by Nazari *et al.* [157]. For typically populated TCEs ($2R \geq I(Q_{X\tilde{X}})$), the analysis is based on a lower tail-bound form of Janson's inequality [95, Theorem 3]: For our analysis, this Janson's bound is a tail bound for the sum of possibly *dependent* Bernoulli RVs, which is suitable for settings in which each Bernoulli RV only depends on a small number of other Bernoulli RVs. Another useful moment result that was derived, is a bound on the correlation of TCE powers, given by [197, Proposition 4] as

$$\mathbb{E}\left[\overline{N}^k(Q_{X\tilde{X}})N^\ell(Q_{XY})\right] \dot{\leq} F(R, Q_{XY}, \ell) \cdot F(2R, Q_{X\tilde{X}}, k), \qquad (4.132)$$

where for a joint type Q_{UV}, $S \geq 0$ and $j \in \mathbb{N}$,

$$F(S, Q_{UV}, j) \triangleq \begin{cases} e^{nj[R-I(Q)]}, & I(Q) < S \\ e^{n[R-I(Q)]}, & I(Q) > S \end{cases}. \qquad (4.133)$$

The bound (4.132) again exhibits that asymptotic independence of TCEs. The proof of (4.132) generalizes the proof of (4.130) and utilizes a double-induction on both k and ℓ.

4.5 Further Applications

In the last 15 years, the TCEM has found extensive applications in diverse coding problems. We next briefly review these applications.

Expurgated exponents were considered in [123], [177] for both standard channel coding, as well as mismatched decoding and under input constraints, utilizing Definition 4.2 and an expurgation argument based on TCEs (see Appendix C).

The TCEM was widely used in multiuser and network problems [59]. For the broadcast channel [16], [33], [73], [106], [107], random-coding and expurgated error exponents were derived using the TCEM in [11], [12], [101], [127], for various decoders, including the optimal bin-index decoder. For the multiple-access channel (MAC), concurrently to the early development stages of the TCEM, Nazari *et al.* [156] also used TCEs (referred to as "packing functions"), but only obtained bounds, as their derivation was only based on the expectation and variance of the TCEs, rather than the their tail probabilities and higher moments. Scarlett, Martinez and Guillén i Fàbregas fully utilized the TCEM for the MAC in [175], [176]. For the interference channel, random-coding error exponents for the Han–Kobayashi scheme [80] and under the optimal ML decoder, were derived in [64], [89]. For the wiretap channel [226], assuming a multi-coding scheme [39], [111], [226], the correct-decoding exponent of the eavesdropper (as in [117]) was derived in

[126], and the exponential decay rate of the mutual information between the message and the eavesdropper output vector (*unnormalized* by the blocklength n) was derived in [164]; this later result refined previous bounds in [83], [84], [163]. The dirty-paper [32] and the Gel'fand–Pinsker [75] channels were analyzed using the TCEM in [196],[6] improving the bounds of Moulin and Wang [154]. The works above heavily rely on the idea of superposition coding and index-bin decoding, which evidently has wide applicability in multiuser problems [59]. Some settings in which it has not been applied yet include the relay channel [34], and channels with feedback [35], [161]. Similar derivations and results are therefore anticipated to these settings too.

For source coding problems, the TCEM was mainly used in distributed compression, and specifically for deriving exact random-binning exponents [129], [130]. For secure lossy compression, the optimal trade-off between the excess-distortion exponent of the legitimate receiver and the exiguous-distortion exponent of the eavesdropper was derived in [218]. For distributed hypothesis testing [1], [79], [81], [170], type-I and type II error exponents were derived for the *quantization-and-binning* scheme [79], [185] under optimal decision rule in [215].

Returning to channel coding, the TCEM was extensively used to derive error exponents for generalized decoders. For Forney's erasure/list decoder [68], random-coding and expurgated error exponents were derived in [76], [91], [118], [125], [190], [220] and by Cao and Tan [27] in a broadcast channel setting. Among other results, this has shown that the exact exponents can be arbitrarily large compared to Forney's bounds, and that, unlike for ordinary decoding [66], [145], they are not universally achievable. Hayashi and Tan [85] used the TCEM for erasure decoding in the *moderate deviation regime* [5], [167]. The exponents of a decoder with a fixed list size were derived in [128]. The result matches the celebrated converse bound of Shannon, Gallager and Berlekamp [183, Theorem 2] and improved the best lower bound previously known [41, p. 196, Problem 27, part (a)].

For asynchronous sparse communication, [28], [29], [202], [203], false-

[6] By means of source-channel duality [103], [168] these results are also applicable to the Wyner–Ziv distributed lossy compression problem [225].

alarm and mis-detection exponents were derived for joint codeword detection and decoding in [216], improving Wang *et al.* [213]. This result was generalized in [219] to joint channel detection and decoding. Exponents for generalized likelihood decoding [227] were derived in [131], [133]. Additional settings include decoding for biometric identification [93, Chapter 5], [208], [222] and content identification [44], [45], for which the exponents for vector-quantized codewords were derived in [132], [134], and error exponents for an alternative model for a biometric identification system, which is based on secret key generation and a helper messages during the enrollment phase [93, Chapter 2], in [134]. Error exponents for the bee identification problem [200] were derived in [193].

The error exponent of the typical random code was derived using the TCEM in [135], for a broad class of generalized likelihood decoders. One of the consequences of this analysis is that a general relation of the form $E_{\text{trc}}(R, P_X) \le E_{\text{ex}}(2R, P_X) + R$ holds for any R and generalized likelihood decoder (for ML decoding and the BSC, a similar relation with equality sign was shown in [14]). A Gallager–style exponent was developed in [137]. The results were then extended to the colored Gaussian channel in [138], to random time-varying trellis codes in [139], and to typical SW codes in [194]. In [195], the TCEM was used to establish that a stochastic MMI decoder, which is a universal decoder, achieves the exponent of the typical random code and the expurgated exponent. Finally, the concentration of the random error exponent to its mean value, the error exponent of the typical random code, was derived using the TCEM in [197], with refinements by Truong *et al.* in [205], and then by Truong and Guillén i Fàbregas [206]. In this last result, the TCEM was used for codewords that are drawn in a dependent manner, for an ensemble based on the Gilbert–Varshamov construction, previously suggested by Somekh-Baruch, Scarlett and Guillén i Fàbregas [191].

5

Manipulating Expectations of Nonlinear Functions of Random Variables

5.1 Introduction

An often-encountered challenge in information-theoretic analytical derivations involves the necessity to assess the expected value of a non-linear function applied to either an RV or a random vector. The conventional approach typically involves resorting to upper and lower bounds for the sought-after expectation, with the hope that these bounds are sufficiently accurate, at least for guiding us toward the correct behavior of the overall expression. When dealing with a non-linear function that exhibits convexity or concavity, it seems natural to employ Jensen's inequality, which yields an upper or lower bound, respectively. However, it is worth noting that this bound may not always prove precise enough to serve our intended purposes.

The primary aim of this section is to introduce a range of alternative tools that have proven their utility in prior research. These alternative tools can be broadly categorized into two main categories.

In the first category (Sections 5.2 and 5.3), the focus is on achieving *exact results*. Here, the fundamental approach involves leveraging *integral representations* of the non-linear function under consideration. In the second category (Sections 5.4, 5.5 and 5.6), we turn to bounding techniques, but these bounds are designed to be more refined and precise than what traditional applications of Jensen's inequality would typically yield. In some cases, these bounds even extend in the opposite direction, offering a comprehensive exploration of the problem at hand.

To provide the reader with a swift comprehension of the concept of an integral representation, as discussed in the first category mentioned earlier, let us delve into a straightforward example. Imagine we have a set of IID zero-mean Gaussian RVs, $X_1, X_2, \ldots, X_n$, each with a variance of σ^2. Our task is to compute the expected value of $\mathbb{E}\{1/\sum_{i=1}^n X_i^2\}$. At first glance, this expectation might appear insurmountable to compute precisely. However, let us consider the integral representation of the function $f(s) = 1/s$ as

$$\frac{1}{s} = \int_0^\infty e^{-st} \, \mathrm{d}t. \tag{5.1}$$

The concept is to employ this representation to tackle the current problem by rearranging the order between the expectation and the integration, much akin to our approach in Section 3:

$$\begin{aligned}
\mathbb{E}\left\{\frac{1}{\sum_{i=1}^n X_i^2}\right\} &\stackrel{(a)}{=} \mathbb{E}\left\{\int_0^\infty \exp\left(-t\sum_{i=1}^n X_i^2\right) \mathrm{d}t\right\} && (5.2)\\
&\stackrel{(b)}{=} \int_0^\infty \mathbb{E}\left\{\exp\left(-t\sum_{i=1}^n X_i^2\right)\right\} \mathrm{d}t && (5.3)\\
&\stackrel{(c)}{=} \int_0^\infty \left[\mathbb{E}\left\{\exp(-tX_1^2)\right\}\right]^n \mathrm{d}t && (5.4)\\
&\stackrel{(d)}{=} \int_0^\infty \frac{\mathrm{d}t}{(1+2\sigma^2 t)^{n/2}} && (5.5)\\
&= \begin{cases} \infty, & n \le 2 \\ \frac{1}{(n-2)\sigma^2}, & n > 2 \end{cases}, && (5.6)
\end{aligned}$$

where (a) is an application of (5.1), (b) amounts to interchanging integration and expectation order, (c) follows from the IID assumption, and (d) is due to the known expression for the MGF of X^2, which follows a chi-squared distribution.

Certainly, the example provided is quite elementary, but it is important to note that this concept can be applied in a wide range of scenarios, involving the presentation of the given function as the Laplace transform (or any other linear transform) of another function, and the expectation of the given non-linear function is represented as an integral of an expression that involves an expectation for which there is a closed-form expression, like the MGF.

Another family of integral representations relates to the following identity, which is applicable to any positive RV X (and can be readily extended to encompass any RV with a well-defined expectation):

$$\mathbb{E}\{X\} = \int_0^\infty \Pr\{X \geq t\} \, \mathrm{d}t. \tag{5.7}$$

In fact, this idea has already been used in Example 3.3 as well as in Section 4. Accordingly, if f is non-negative and monotonic, and hence invertible, we have

$$\mathbb{E}\{f(X)\} = \int_0^\infty \Pr\{f(X) \geq t\} \, \mathrm{d}t = \int_0^\infty \Pr\{X \geq f^{-1}(t)\} \, \mathrm{d}t, \tag{5.8}$$

which often lends itself to closed-form analysis.

In the first two upcoming sections, we will explore certain integral representations of two specific highly important functions in the context of information-theoretic analyses: The logarithmic function (in Section 5.2, which is based on [148]) and the power function (in Section 5.3, which is based on [149]). To the best of our knowledge, those integral representations are not very common in the information-theory literature, but the essence of the approach remains as described above: Substitute the expectation of the given non-linear function with an integral of a function for which a closed-form expectation exists.

In the second category of tools explored in this section, which focuses on modified versions of the Jensen inequality, we delve into three distinct types of bounding techniques:

1. Jensen's inequality combined with a change of measure (Section 5.4), where our exposition relies strongly on [9], [121] and [136].

2. Reverse Jensen inequalities (Section 5.5), which summarizes the main findings on [142].

3. Jensen-like inequalities, where the convex/concave function is only part of the expression and the supporting line is re-optimized (Section 5.6), which is based on [144].

While these techniques provide bounds rather than exact results, as seen in the first category, their applicability extends across a broader range of scenarios. Furthermore, they often yield substantial improvements compared to the bounds derived from the conventional Jensen inequality.

5.2 An Integral Representation of the Logarithmic Function

In this section, we will explore the following integral representation of the logarithmic function, which states that for $x > 0$

$$\ln x = \int_0^\infty \frac{e^{-u} - e^{-ux}}{u} \, \mathrm{d}u, \tag{5.9}$$

and can be easily proved by substituting $\int_0^\infty e^{-ut} \, \mathrm{d}t$ for $1/u$ on the RHS and interchanging the order of the integration. This representation finds its immediate utility in scenarios where the argument of the logarithmic function is a positive-valued RV denoted as X, and our goal is to compute the expectation, denoted as $\mathbb{E}\{\ln X\}$. By assuming the validity of interchanging the expectation operator with the integration over the variable u, we can simplify the calculation of $\mathbb{E}\{\ln X\}$ into evaluating the MGF of X, which is often a more straightforward task. This transformation allows us to express it as:

$$\mathbb{E}\left\{\ln X\right\} = \int_0^\infty \left[e^{-u} - \mathbb{E}\{e^{-uX}\}\right] \frac{\mathrm{d}u}{u}. \tag{5.10}$$

In particular, if $X_1, \ldots, X_n$ are positive IID RVs, then

$$\mathbb{E}\left\{\ln(X_1 + \ldots + X_n)\right\} = \int_0^\infty \left(e^{-u} - \left[\mathbb{E}\{e^{-uX_1}\}\right]^n\right) \frac{\mathrm{d}u}{u}. \tag{5.11}$$

This concept is not entirely novel, as it has been previously applied in physics, as evidenced in sources such as [63, Eq. (2.4) and beyond], [151, Exercise 7.6, p. 140], and [192, Eq. (12) and beyond]. However, in the field of information theory, this approach is seldom utilized, despite its potential significance. This significance arises from the frequent requirement to compute logarithmic expectations – a common occurrence in numerous problem areas within information theory. Furthermore, the integral representation (5.9) extends its utility beyond mere expectation calculations; it also proves invaluable in evaluating higher moments of $\ln X$, most notably, the second moment or variance. This added functionality allows us to assess statistical fluctuations around the mean, enhancing our analytical capabilities in the field.

In [148], the practicality of this approach was effectively showcased across various application domains. These applications encompassed areas such as entropy and differential entropy assessments, performance analysis of universal lossless source codes, and the determination of ergodic capacity for the Rayleigh SIMO channel, with AWGN. It is worth noting that within some of these examples, we successfully computed variances related to the pertinent RVs. In particular, in [148, Proposition 2], the following result is stated and proved: For an RV X and $s \in \mathbb{R}$ let

$$M_X(s) \triangleq \mathbb{E}\left\{e^{sX}\right\}, \tag{5.12}$$

be the MGF of X. If $X \geq 0$ with probability one, then

$$\mathbb{E}\{\ln(1+X)\} = \int_0^\infty \frac{e^{-u} \cdot [1 - M_X(-u)]}{u} \, \mathrm{d}u, \tag{5.13}$$

and

$$\begin{aligned} &\mathrm{Var}\{\ln(1+X)\} \\ &= \int_0^\infty \int_0^\infty \frac{e^{-(u+v)}}{uv} \Big[M_X(-u-v) - M_X(-u)\, M_X(-v)\Big] \, \mathrm{d}u\mathrm{d}v. \end{aligned} \tag{5.14}$$

It is worth highlighting an intriguing consequence of the integral representation (5.10). It transforms the calculation of the expectation of the logarithm of X into the expectation of an exponential function of X. This transformation has an added benefit: It simplifies expressions involving quantities like $\ln(n!)$ into the integral of a summation of a geometric series, a form that is readily expressible in closed form. Specifically,

$$\ln(n!) = \sum_{k=1}^{n} \ln k \tag{5.15}$$

$$= \sum_{k=1}^{n} \int_0^\infty (e^{-u} - e^{-uk}) \frac{\mathrm{d}u}{u} \tag{5.16}$$

$$= \int_0^\infty \left(ne^{-u} - \sum_{k=1}^{n} e^{-uk} \right) \frac{\mathrm{d}u}{u} \tag{5.17}$$

$$= \int_0^\infty e^{-u} \left(n - \frac{1 - e^{-un}}{1 - e^{-u}} \right) \frac{\mathrm{d}u}{u}. \tag{5.18}$$

For a positive integer-valued RV, denoted as N, the computation of $\mathbb{E}\{\ln N!\}$ becomes a straightforward task, requiring only the calculation of $\mathbb{E}\{N\}$ and the MGF, $\mathbb{E}\{e^{-uN}\}$. This is useful, for example, when N follows a Poisson distribution, as shown in [148] in detail.

In [148], the usefulness of the integral representation of the logarithmic function is illustrated in several problem areas in information theory, including graphs of numerical results. Here, we briefly summarize two of the examples provided therein.

5.2.1 Differential Entropy for Generalized Multivariate Cauchy Densities

Let $(X_1, \ldots, X_n)$ be a random vector whose PDF is of the form

$$f(x_1, \ldots, x_n) = \frac{C_n}{[1 + \sum_{i=1}^{n} g(x_i)]^q}, \tag{5.19}$$

for $(x_1, \ldots, x_n) \in \mathbb{R}^n$, a given non-negative function g, and a real $q > 0$ such that

$$\int_{\mathbb{R}^n} \frac{\mathrm{d}\boldsymbol{x}}{[1 + \sum_{i=1}^{n} g(x_i)]^q} < \infty. \tag{5.20}$$

We term this category of density as *generalized multivariate Cauchy*, primarily because the multivariate Cauchy density arises as a specific instance when $g(x) = x^2$ and $q = \frac{1}{2}(n + 1)$. Employing the Laplace transform relation,

$$\frac{1}{s^q} = \frac{1}{\Gamma(q)} \int_0^\infty t^{q-1} e^{-st} \, \mathrm{d}t, \tag{5.21}$$

which holds for $q \geq 1$ and $\mathrm{Re}(s) > 0$, f can be displayed as a mixture of product-form PDFs:

$$f(x_1, \ldots, x_n) = \frac{C_n}{[1 + \sum_{i=1}^{n} g(x_i)]^q} \tag{5.22}$$

$$= \frac{C_n}{\Gamma(q)} \int_0^\infty t^{q-1} e^{-t} \cdot \exp\left\{-t \sum_{i=1}^n g(x_i)\right\} \, \mathrm{d}t. \tag{5.23}$$

Defining for $t > 0$

$$Z(t) \triangleq \int_{-\infty}^{\infty} e^{-tg(x)} \, \mathrm{d}x \tag{5.24}$$

we obtain from (5.23),

$$1 = \frac{C_n}{\Gamma(q)} \int_0^\infty t^{q-1} e^{-t} \int_{\mathbb{R}^n} \exp\left\{-t \sum_{i=1}^n g(x_i)\right\} \, \mathrm{d}x_1 \ldots \mathrm{d}x_n \mathrm{d}t \tag{5.25}$$

$$= \frac{C_n}{\Gamma(q)} \int_0^\infty t^{q-1} e^{-t} \left(\int_{-\infty}^{\infty} e^{-tg(x)} \mathrm{d}x\right)^n \mathrm{d}t \tag{5.26}$$

$$= \frac{C_n}{\Gamma(q)} \int_0^\infty t^{q-1} e^{-t} Z^n(t) \, \mathrm{d}t, \tag{5.27}$$

and so,

$$C_n = \frac{\Gamma(q)}{\displaystyle\int_0^\infty t^{q-1} e^{-t} Z^n(t) \, \mathrm{d}t}. \tag{5.28}$$

Evaluating the differential entropy of f involves deriving $\mathbb{E}\{\ln[1 + \sum_{i=1}^n g(X_i)]\}$. Using (5.13)

$$\mathbb{E}\left\{\ln\left[1 + \sum_{i=1}^n g(X_i)\right]\right\} = \int_0^\infty \frac{e^{-u}}{u} \left(1 - \mathbb{E}\left\{\exp\left[-u \sum_{i=1}^n g(X_i)\right]\right\}\right) \mathrm{d}u, \tag{5.29}$$

and

$$\mathbb{E}\left\{\exp\left[-u \sum_{i=1}^n g(X_i)\right]\right\}$$

$$= \frac{C_n}{\Gamma(q)} \int_0^\infty t^{q-1} e^{-t} \int_{\mathbb{R}^n} \exp\left\{-(t+u) \sum_{i=1}^n g(x_i)\right\} \mathrm{d}x_1 \ldots \mathrm{d}x_n \mathrm{d}t \tag{5.30}$$

$$= \frac{C_n}{\Gamma(q)} \int_0^\infty t^{q-1} e^{-t} Z^n(t+u) \, \mathrm{d}t. \tag{5.31}$$

Thus, the joint differential entropy is given by

$$h(X_1, \ldots, X_n)$$

$$= q \cdot \mathbb{E}\left\{\ln\left[1+\sum_{i=1}^{n} g(X_i)\right]\right\} - \ln C_n \tag{5.32}$$

$$= q\int_0^\infty \frac{e^{-u}}{u}\left(1-\frac{C_n}{\Gamma(q)}\int_0^\infty t^{q-1}e^{-t}Z^n(t+u)\,\mathrm{d}t\right)\mathrm{d}u - \ln C_n \tag{5.33}$$

$$= \frac{qC_n}{\Gamma(q)}\int_0^\infty\int_0^\infty \frac{t^{q-1}e^{-(t+u)}}{u}\left[Z^n(t)-Z^n(t+u)\right]\mathrm{d}t\mathrm{d}u - \ln C_n. \tag{5.34}$$

For $g(x) = |x|^\theta$, with an arbitrary $\theta > 0$, we obtain from (5.24) that

$$Z(t) = \frac{2\cdot\Gamma(1/\theta)}{\theta\cdot t^{1/\theta}}. \tag{5.35}$$

In particular, for $\theta = 2$ and $q = \frac{1}{2}(n+1)$, we get the multivariate Cauchy density from (5.19). In this case, since $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, it follows from (5.35) that $Z(t) = \sqrt{\frac{\pi}{t}}$ for $t > 0$, and from (5.28)

$$C_n = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{n/2}\displaystyle\int_0^\infty t^{(n+1)/2-1}e^{-t}\,t^{-n/2}\,\mathrm{d}t} = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{n/2}\,\Gamma\left(\frac{1}{2}\right)} = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{(n+1)/2}}. \tag{5.36}$$

Combining (5.34), (5.35) and (5.36) gives

$$h(X_1,\ldots,X_n) = \frac{n+1}{2\pi^{(n+1)/2}}\int_0^\infty\int_0^\infty \frac{e^{-(t+u)}}{u\sqrt{t}}\left[1-\left(\frac{t}{t+u}\right)^{n/2}\right]\mathrm{d}t\mathrm{d}u + \frac{(n+1)\ln\pi}{2} - \ln\Gamma\left(\frac{n+1}{2}\right). \tag{5.37}$$

In this application example, we find it intriguing that (5.34) offers what can be considered a "single-letter expression." Remarkably, the n-dimensional integral tied to the original expression of the differential entropy $h(X_1,\ldots,X_n)$ is effectively replaced by the two-dimensional integral in (5.34), and notably, this replacement remains independent of the value of n.

5.2.2 Ergodic Capacity of the Rayleigh SIMO Channel

Let us consider the SIMO channel with L receive antennas and AWGN. We make the assumption that the channel transfer coefficients, denoted as $h_1, h_2, \ldots, h_L$, are independent and follow a zero-mean, circularly symmetric complex Gaussian distribution with variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_L^2$. In this context, the ergodic capacity of the SIMO channel, measured in nats per channel use, is expressed as an expected value:

$$C = \mathbb{E}\left\{\ln\left(1+\rho\sum_{\ell=1}^{L}|h_\ell|^2\right)\right\} = \mathbb{E}\left\{\ln\left(1+\rho\sum_{\ell=1}^{L}(f_\ell^2+g_\ell^2)\right)\right\}, \quad (5.38)$$

where $f_\ell \triangleq \mathrm{Re}\{h_\ell\}$, $g_\ell \triangleq \mathrm{Im}\{h_\ell\}$, and $\rho \triangleq \frac{P}{N_0}$ is the signal-to-noise ratio (SNR). In view of (5.13), let

$$X \triangleq \rho\sum_{\ell=1}^{L}(f_\ell^2+g_\ell^2). \quad (5.39)$$

For all $u > 0$,

$$\begin{aligned} M_X(-u) &= \mathbb{E}\left\{\exp\left(-\rho u\sum_{\ell=1}^{L}(f_\ell^2+g_\ell^2)\right)\right\} && (5.40)\\ &= \prod_{\ell=1}^{L}\left\{\mathbb{E}\left\{e^{-u\rho f_\ell^2}\right\}\cdot\mathbb{E}\left\{e^{-u\rho g_\ell^2}\right\}\right\} && (5.41)\\ &= \prod_{\ell=1}^{L}\frac{1}{1+u\rho\sigma_\ell^2}, && (5.42)\end{aligned}$$

where (5.42) holds since

$$\begin{aligned} \mathbb{E}\left\{e^{-u\rho f_\ell^2}\right\} &= \mathbb{E}\left\{e^{-u\rho g_\ell^2}\right\} && (5.43)\\ &= \int_{-\infty}^{\infty}\frac{\mathrm{d}w}{\sqrt{\pi\sigma_\ell^2}}\cdot e^{-w^2/\sigma_\ell^2}\cdot e^{-u\rho w^2} && (5.44)\\ &= \frac{1}{\sqrt{1+u\rho\sigma_\ell^2}}. && (5.45)\end{aligned}$$

From (5.13), (5.38) and (5.42), the ergodic capacity (in nats per channel use) is given by

$$C = \mathbb{E}\left\{\ln\left(1+\rho\sum_{\ell=1}^{L}(f_\ell^2+g_\ell^2)\right)\right\} \quad (5.46)$$

$$= \int_0^\infty \frac{e^{-u}}{u} \left(1 - \prod_{\ell=1}^{L} \frac{1}{1+u\rho\sigma_\ell^2}\right) \mathrm{d}u \tag{5.47}$$

$$= \int_0^\infty \frac{e^{-x/\rho}}{x} \left(1 - \prod_{\ell=1}^{L} \frac{1}{1+\sigma_\ell^2 x}\right) \mathrm{d}x. \tag{5.48}$$

We next turn to the variance of $\ln[1+\rho\sum_{\ell=1}^{L}[f_\ell^2+g_\ell^2]]$. In the context of a fading channel, the randomness of this RV is related to random fluctuations in the channel quality, related to the so-called *instantaneous capacity*, and thus to the *outage probability* [207]. Concerning the variance, owing to (5.14) and (5.42), we have:

$$\begin{aligned}
&\mathrm{Var}\left\{\ln\left(1+\rho\sum_{\ell=1}^{L}[f_\ell^2+g_\ell^2]\right)\right\} \\
&= \int_0^\infty \int_0^\infty \frac{e^{-(x+y)/\rho}}{xy} \times \\
&\qquad \left\{\prod_{\ell=1}^{L} \frac{1}{1+\sigma_\ell^2(x+y)} - \prod_{\ell=1}^{L}\left[\frac{1}{(1+\sigma_\ell^2 x)(1+\sigma_\ell^2 y)}\right]\right\} \mathrm{d}x\mathrm{d}y.
\end{aligned} \tag{5.49}$$

The capacity C can be expressed as a linear combination of integrals of the form

$$\int_0^\infty \frac{e^{-x/\rho}\,\mathrm{d}x}{1+\sigma_\ell^2 x} = \frac{1}{\sigma_\ell^2}\int_0^\infty \frac{e^{-t}\,\mathrm{d}t}{t+1/(\sigma_\ell^2\rho)} \tag{5.50}$$

$$= \frac{e^{1/(\sigma_\ell^2\rho)}}{\sigma_\ell^2}\int_{1/(\sigma_\ell^2\rho)}^\infty \frac{e^{-s}}{s}\,\mathrm{d}s \tag{5.51}$$

$$= \frac{1}{\sigma_\ell^2}\cdot e^{1/(\sigma_\ell^2\rho)}\cdot E_1\left(\frac{1}{\sigma_\ell^2\rho}\right), \tag{5.52}$$

where $E_1(\cdot)$ is the (modified) exponential integral function, defined for $x > 0$ as

$$E_1(x) \triangleq \int_x^\infty \frac{e^{-s}}{s}\,\mathrm{d}s. \tag{5.53}$$

5.3 An Integral Representation of the Power Function

In this section, we build upon the same approach as in Section 5.2, expanding the scope to introduce an integral representation for a general moment of a non-negative RV, X. Specifically, we aim to find an expression for $\mathbb{E}\{X^\rho\}$ where $\rho > 0$. When ρ is an integer, it is well-known that this moment can be computed as the ρ-th order derivative of the MGF of X, evaluated at the origin. However, our proposed integral representation, presented in this work, applies to any non-integer positive value of ρ. Here as well, it replaces the direct calculation of $\mathbb{E}\{X^\rho\}$ with the integration of an expression involving the MGF of X. We refer to this representation as an "extension" of the integral representation of the logarithmic function discussed in Section 5.2. This is because the latter can be derived as a special case of the formula for $\mathbb{E}\{X^\rho\}$ by employing the identity:

$$\mathbb{E}\left\{\ln X\right\} = \lim_{\rho \to 0} \frac{\mathbb{E}\{X^\rho\} - 1}{\rho}, \tag{5.54}$$

or alternatively, the identity,

$$\mathbb{E}\left\{\ln X\right\} = \lim_{\rho \to 0} \frac{\ln\left[\mathbb{E}\{X^\rho\}\right]}{\rho}. \tag{5.55}$$

As in the previous section, here too, the proposed integral representation is applied to a range of examples motivated by information theory [149]. This application showcases how the representation streamlines numerical evaluations. In particular, much like the case of the logarithmic function, when employed to compute a moment of the sum of a large number, denoted as n, of non-negative RVs, it becomes evident that integration over one or two dimensions, as suggested by our integral representation, is notably simpler than the alternative of integrating over n dimensions, as required in the direct calculation of the desired moment. Additionally, single or double-dimensional integrals can be promptly and accurately computed using built-in numerical integration techniques.

In order to present the integral representation, we commence by recalling the definition of the Gamma function and defining the Beta function, as follows:

$$\Gamma(u) \triangleq \int_0^\infty t^{u-1} e^{-t} \, \mathrm{d}t, \tag{5.56}$$

for $u > 0$, and

$$B(u, v) \triangleq \int_0^1 t^{u-1} (1-t)^{v-1} \, \mathrm{d}t = \frac{\Gamma(u)\Gamma(v)}{\Gamma(u+v)}, \tag{5.57}$$

for $u, v > 0$. Let X be a non-negative RV with an MGF $M_X(\cdot)$, and let $\rho > 0$ be a non-integer real. Then, as shown in [149],

$$\mathbb{E}\{X^\rho\} = \frac{1}{1+\rho}\sum_{\ell=0}^{\lfloor\rho\rfloor}\frac{\alpha_\ell}{B(\ell+1,\rho+1-\ell)} + \frac{\rho\cdot\sin(\pi\rho)\cdot\Gamma(\rho)}{\pi}\times \int_0^\infty \frac{1}{u^{\rho+1}}\left(\sum_{j=0}^{\lfloor\rho\rfloor}\left\{\frac{(-1)^j\cdot\alpha_j}{j!}\cdot u^j\right\}e^{-u} - M_X(-u)\right)\mathrm{d}u, \quad (5.58)$$

where for all $j \in \{0, 1, \dots, \lfloor\rho\rfloor\}$

$$\alpha_j \triangleq \mathbb{E}\left\{(X-1)^j\right\} \quad (5.59)$$

$$= \frac{1}{j+1}\sum_{\ell=0}^{j}\frac{(-1)^{j-\ell}\cdot M_X^{(\ell)}(0)}{B(\ell+1,j-\ell+1)}. \quad (5.60)$$

The proof of (5.58) in [149] does not apply to natural values ρ (see [149, Appendix A], where the denominators vanish). However, taking a limit in (5.58) where we let ρ tend to an integer, and applying L'Hôpital's rule, one can reproduce the well-known result for integer ρ, which is given in terms of the ρ-th order derivative of the MGF at the origin. For $\rho \in (0,1)$, the above simplifies to:

$$\mathbb{E}\{X^\rho\} = 1 + \frac{\rho}{\Gamma(1-\rho)}\int_0^\infty \frac{e^{-u} - M_X(-u)}{u^{1+\rho}}\,\mathrm{d}u. \quad (5.61)$$

In [149], the profound utility of the integral representation shines through in a comprehensive exploration across various domains within information theory and statistics. These applications include detailed investigations accompanied by graphical illustrations. The showcased instances span a range of analytical inquiries, encompassing randomized guessing, estimation errors, the Rényi entropy of n-dimensional generalized Cauchy distributions, and mutual information calculations for channels featuring a specific jammer model. Here, we will provide a succinct overview of one of these application examples, focusing primarily on the easier scenario where $\rho \in (0,1)$ for clarity and simplicity of exposition.

Example 5.1 (Moments of Guesswork). Suppose we have an RV X that assumes values from a finite alphabet $\mathcal{X}$. Let us explore a random guessing strategy wherein the guesser submits a sequence of independent random guesses, drawn from a specific probability distribution denoted as $\widetilde{P}(\cdot)$, defined over $\mathcal{X}$. Consider any instance where $x \in \mathcal{X}$ represents a realization of X, and we have the guessing distribution $\widetilde{P}$ at our disposal. In such a scenario, the RV G, representing the number of independent guesses required to achieve success, follows a geometric distribution:

$$\Pr\{G = k|x\} = \left[1 - \widetilde{P}(x)\right]^{k-1} \cdot \widetilde{P}(x). \tag{5.62}$$

Hence, for $u < \ln \frac{1}{1-\widetilde{P}(x)}$ the corresponding MGF is equal to

$$M_G(u|x) = \sum_{k=1}^{\infty} e^{ku} \cdot \Pr\{G = k|x\} \tag{5.63}$$

$$= \frac{\widetilde{P}(x)}{e^{-u} - (1 - \widetilde{P}(x))}. \tag{5.64}$$

For $\rho \in (0, 1)$, it is shown in [149] that

$$\mathbb{E}\left\{G^\rho|x\right\} = 1 + \frac{\rho}{\Gamma(1-\rho)} \int_0^\infty \frac{e^{-u} - e^{-2u}}{u^{\rho+1}\left[\left(1 - \widetilde{P}(x)\right)^{-1} - e^{-u}\right]} \, \mathrm{d}u. \tag{5.65}$$

Consider the distribution of the RV X, denoted as P. To compute the unconditional ρ-th moment using (5.65), we average over all possible values of X. This yields the following result for all ρ in the open interval $(0, 1)$:

$$\mathbb{E}\left\{G^\rho\right\} = 1 + \frac{\rho}{\Gamma(1-\rho)} \int_0^1 \frac{1-z}{(-\ln z)^{\rho+1}} \sum_{x\in\mathcal{X}} \frac{P(x)(1 - \widetilde{P}(x))}{1 - z(1 - \widetilde{P}(x))} \, \mathrm{d}z, \tag{5.66}$$

where (5.66) is by changing the integration variable according to $z = e^{-u}$. In conclusion, equation (5.65) provides a computable one-dimensional integral expression for the ρ-th guessing moment for any $\rho > 0$ (at least numerically). This eliminates the necessity for numerical computations involving infinite sums.

5.4 Jensen's Inequality with a Change of Measure

In this section, we advocate for the practical utility of combining Jensen's inequality with a change of measure, effectively introducing an additional degree of freedom for optimization. To illustrate this concept concretely, let us consider a concave function f and an RV X characterized by a PDF p, with its support set in $\mathcal{X}$. Additionally, let q represent another PDF, also with support in $\mathcal{X}$. We will use $\mathbb{E}_p\{\cdot\}$ and $\mathbb{E}_q\{\cdot\}$ to denote expectation operators WRT p and q, respectively. Now, let us delve into the following chain of simple inequalities:

$$\begin{aligned} f\left(\mathbb{E}_p\{X\}\right) &= f\left(\int_{\mathcal{X}} p(x)x \,\mathrm{d}x\right) && (5.67)\\ &= f\left(\int_{\mathcal{X}} q(x)\cdot\frac{xp(x)}{q(x)}\,\mathrm{d}x\right) && (5.68)\\ &\geq \int_{\mathcal{X}} q(x) f\left(\frac{xp(x)}{q(x)}\right)\mathrm{d}x && (5.69)\\ &= \mathbb{E}_q\left\{f\left(\frac{Xp(X)}{q(X)}\right)\right\}. && (5.70)\end{aligned}$$

Given that the inequalities mentioned above are valid for any PDF q supported by $\mathcal{X}$, we have the flexibility to maximize the rightmost side of this chain, which can be expressed as:

$$f\left(\mathbb{E}_p\{X\}\right) \geq \sup_{q\in\mathcal{Q}} \mathbb{E}_q\left\{f\left(\frac{Xp(X)}{q(X)}\right)\right\}, \tag{5.71}$$

where $\mathcal{Q}$ is any class of PDFs with this support. Clearly, when p belongs to the set $\mathcal{Q}$, the choice of $q = p$ reduces the inequality in (5.71) to the standard Jensen's inequality. On the opposite end of the spectrum, if $\mathcal{Q}$ encompasses the entire collection of PDFs over $\mathcal{X}$, and if X is a positive RV with $\mathbb{E}_p\{X\} < \infty$, then selecting $q(x) = \frac{xp(x)}{\mathbb{E}_p\{X\}}$ results in a trivial and uninformative identity. However, this highlights that in such a scenario, the inequality in (5.71) essentially becomes an equality, depicted as:

$$f\left(\mathbb{E}_p\{X\}\right) = \sup_{\{q:\ \mathrm{supp}\{q\}=\mathcal{X}\}} \mathbb{E}_q\left\{f\left(\frac{Xp(X)}{q(X)}\right)\right\}. \tag{5.72}$$

In the sequel, we abbreviate suprema and infima over $\{q\colon \mathrm{supp}\{q\} = \mathcal{X}\}$ simply by writing $\sup_q$ and $\inf_q$, respectively. Likewise, if f is convex, we have

$$f\left(\mathbb{E}_p\{X\}\right) \leq \inf_{q\in\mathcal{Q}} \mathbb{E}_q\left\{f\left(\frac{Xp(X)}{q(X)}\right)\right\}. \tag{5.73}$$

The effectiveness of these inequalities hinges on our judicious selection of $\mathcal{Q}$. As we have observed, $\mathcal{Q}$ should encompass p to ensure that the resultant bound, after optimizing over $q \in \mathcal{Q}$, does not fall short of the standard Jensen's inequality. Conversely, $\mathcal{Q}$ should exclude the choice $q(x) = \frac{xp(x)}{\mathbb{E}_p\{X\}}$, which renders the inequality uninformative. Ideally, the class $\mathcal{Q}$ should be well-suited for practical use, allowing for closed-form optimization. This convenience would enable us to derive bounds that significantly improve upon the standard Jensen's inequality, making the approach both mathematically tractable and practically useful.

Perhaps the most important special case of (5.72) pertains to the case of $f(x) = \ln x$, where it becomes

$$\ln\left(\mathbb{E}_p\{X\}\right) = \sup_q \mathbb{E}_q\left\{\ln\left(\frac{Xp(X)}{q(X)}\right)\right\} \tag{5.74}$$

$$= \sup_q \left\{\mathbb{E}_q\{f(X)\} - D(q\|p)\right\}, \tag{5.75}$$

which is intimately related to the Laplace principle [58] in large-deviations theory, or more generally, to Varadhan's integral lemma [49, Section 4.3] or the Donsker-Varadhan variational principle.

Example 5.2 (Exponential moments of codeword lengths in lossless compression). Let $U_1, \ldots, U_n$ be drawn from a finite-alphabet memoryless source P and let $\ell(U_1, \ldots, U_n)$ be the length (in nats) of the compressed version of $(U_1, \ldots, U_n)$ under some given fixed-to-variable length lossless source code. We consider the exponents of $\ell(U_1, \ldots, U_n)$ given by $X = \exp\{\alpha\ell(U_1, \ldots, U_n)\}$ where $\alpha > 0$ is a given real parameter. These moments form the MGF, and as such they provide the full information on their probability distribution. In addition, they are relevant for the large-deviations behavior, *e.g.*, assessing the probability of codeword length buffer overflow [92], [97], [217]. Now, a naive application of Jensen's inequality yields

$$\mathbb{E}\{X\} = \mathbb{E}\left\{\exp[\alpha\ell(U_1, \ldots, U_n)]\right\} \tag{5.76}$$

$$\geq \exp\left\{\alpha\mathbb{E}\{\ell(U_1, \ldots, U_n)\right\} \tag{5.77}$$

$$\geq e^{\alpha n H(P)}, \tag{5.78}$$

where $H(P)$ is the per-symbol entropy of the source P. On the other hand, considering P^n and Q^n as probability distributions of n-vectors from the source, we have

$$\ln\left(\mathbb{E}_{P^n}\{X\}\right) \geq \sup_{Q^n\in\mathcal{Q}}\left[\mathbb{E}_{Q^n}\{\ln X\} - D(Q^n\|P^n)\right] \tag{5.79}$$

$$= \sup_{Q^n\in\mathcal{Q}}\left[\alpha\mathbb{E}_{Q^n}\{\ell(U_1,\ldots,U_n)\} - D(Q^n\|P^n)\right] \tag{5.80}$$

$$\geq \sup_{Q^n\in\mathcal{Q}}\left[\alpha H(Q^n) - D(Q^n\|P^n)\right]. \tag{5.81}$$

Now, rather than taking $\mathcal{Q}$ to be the class of all probability distributions of n-vectors, let us take it to be the class of all product form distributions, *i.e.*, $Q^n(u_1,\ldots,u_n) = \prod_{i=1}^n Q(u_i)$. Since P^n has a product form too, *i.e.*, $P^n(u_1,\ldots,u_n) = \prod_{i=1}^n P(u_i)$, we readily obtain that the last expression reads

$$\sup_{Q^n\in\mathcal{Q}}\left[\alpha H(Q^n) - D(Q^n\|P^n)\right] = n\cdot\sup_{Q}[\alpha H(Q) - D(Q\|P)], \tag{5.82}$$

that yields the Rényi entropy of order α pertaining to P, which is an attainable lower bound to the exponential moment of $\ell(U_1,\ldots,U_n)$, unlike the lower bound obtained from the naive use of Jensen's inequality above. In other words, rather than maximizing over the entire class of all probability distributions of n-vectors, $\{Q^n\}$, we observe that the much smaller class of memoryless probability distributions is large enough to obtain a tight result. The same idea was used also in the converse part of [9] in the context of guessing, which is strongly related to source coding.

The identity (5.75) has found extensive utility, not only in this context but also in previous works such as [121], where it was applied to exponential moments of various loss functions, and [136], where it played a crucial role in establishing lower bounds on exponential moments of estimation errors. Numerous references within these two articles further emphasize the importance of (5.75). However, it is vital to highlight a key takeaway message from this section: The relation (5.71) is not limited to the logarithmic function alone; it holds true for any concave function (or convex function with appropriate adjustments) and extends its applicability beyond just the logarithmic case.

Example 5.3. To demonstrate another special case of combining Jensen's inequality with a change of measure, consider the example of deriving an upper bound to the expectation of the harmonic mean of n positive RVs, $X_1, \ldots, X_n$, *i.e.*,

$$\mathbb{E}\left\{\frac{n}{\sum_{i=1}^n 1/X_i}\right\}. \tag{5.83}$$

This expectation cannot be upper bounded by a direct application of Jensen's inequality, because it provides a lower bound,

$$\mathbb{E}\left\{\frac{n}{\sum_{i=1}^n 1/X_i}\right\} \geq \frac{n}{\sum_{i=1}^n \mathbb{E}\{1/X_i\}}, \tag{5.84}$$

rather than an upper bound, and moreover, it requires the expectations of $1/X_i$ rather than those of X_i. However, consider the following approach: Let $q = (q_1, \ldots q_n)$ be an arbitrary probability vector, *i.e.*, a set of n positive numbers summing to unity. Then,

$$\sum_{i=1}^n \frac{1}{X_i} = \sum_{i=1}^n q_i \cdot \frac{1}{q_i X_i} \tag{5.85}$$

$$\stackrel{(*)}{\geq} \frac{1}{\sum_{i=1}^n q_i \cdot (q_i X_i)} \tag{5.86}$$

$$= \frac{1}{\sum_{i=1}^n q_i^2 X_i}, \tag{5.87}$$

where $(*)$ follows from the (ordinary) Jensen inequality applied to the convex function $f(u) = 1/u$. Equivalently,

$$\frac{n}{\sum_{i=1}^n 1/X_i} \leq n \cdot \sum_{i=1}^n q_i^2 X_i. \tag{5.88}$$

Since this inequality holds for every probability vector q, we may minimize the RHS over q, to obtain

$$\frac{n}{\sum_{i=1}^n 1/X_i} \leq n \cdot \min_q \sum_{i=1}^n q_i^2 X_i. \tag{5.89}$$

Taking the expectations of both sides, we get:

$$\mathbb{E}\left\{\frac{n}{\sum_{i=1}^n 1/X_i}\right\} \leq n \cdot \mathbb{E}\left\{\min_q \sum_{i=1}^n q_i^2 X_i\right\} \tag{5.90}$$

$$\leq n \cdot \min_q \mathbb{E}\left\{\sum_{i=1}^n q_i^2 X_i\right\} \tag{5.91}$$

$$= n \cdot \min_q \sum_{i=1}^n q_i^2 \mathbb{E}\{X_i\} \tag{5.92}$$

$$\stackrel{(*)}{=} \frac{n}{\sum_{i=1}^n 1/\mathbb{E}\{X_i\}}, \tag{5.93}$$

where $(*)$ follows from the optimal choice of q, which is according to

$$q_i = \frac{1/\mathbb{E}\{X_i\}}{\sum_{j=1}^n 1/\mathbb{E}\{X_j\}}. \tag{5.94}$$

More generally, whenever the function $f(u) = u^\rho$ is convex (namely, for $\rho \notin (0,1)$), we can similarly obtain the inequality

$$\mathbb{E}\left\{\left(\sum_{i=1}^n X_i\right)^\rho\right\} \le \left[\sum_{i=1}^n \left(\mathbb{E}\{X_i^\rho\}\right)^{1/\rho}\right]^\rho. \tag{5.95}$$

Note that no assumptions were imposed on the dependence/independence among the RVs $\{X_i\}$.

5.5 Reverse Jensen Inequalities

The widely used Jensen inequality states that

$$\mathbb{E}\{f(X)\} \ge f(\mathbb{E}\{X\}) \tag{5.96}$$

for any convex function f and an RV X. Frequently, however, applied mathematicians, and especially information-theorists, encounter a rather vexing situation where Jensen's inequality seems to operate in the opposite direction of their desired results. This observation has spurred significant research efforts aimed at developing various versions of the so-called *reverse Jensen inequality* (RJI). A myriad of articles, including, but not limited to, [3], [24], [25], [54], [55], [96], [104], [105], [187], [224], have delved into this topic, showcasing its rich and evolving landscape. In the majority of these works, the derived inequalities find practical applications in diverse fields. Examples include establishing valuable relationships between arithmetic and geometric means, deriving reverse bounds on entropy, KL divergence, and more generally, Csiszár's f-divergence [4], [37]. Additionally, these inequalities have been extended to reverse versions of the Hölder inequality, among other applications.

In many of the aforementioned papers, the primary results manifest in the form of an upper bound on the difference $\mathbb{E}\{f(X)\} - f(\mathbb{E}\{X\})$, where f denotes a convex function and X is an RV. It is worth noting that these upper bounds predominantly rely on global properties of the function f, such as its range and domain, rather than on the underlying PDF of X or its probability mass function in the discrete case. Ideally, a desirable characteristic of an RJI would be its ability to provide tight bounds when the PDF of X is highly concentrated around its mean, akin to the well-known property of the standard Jensen inequality (5.96). Such tightness in the presence of concentration around the mean is a hallmark of the ordinary Jensen inequality, and it would be advantageous for RJIs to exhibit a similar behavior under such conditions.

In [142], we extend the concepts introduced in [224], providing a fresh perspective on the RJI landscape. Our contributions encompass several novel variants of RJI, and what sets these apart is their ability to exhibit the desired property of tightness in cases of measure concentration, a characteristic we consistently emphasize.

Our journey in this section begins from the same foundational point as found in the proof of [224, Lemma 1], but we continue by taking a significantly different path. As we amply demonstrate, this novel approach leads to notably tighter bounds, which prove to be eminently tractable and analyzable in a multitude of scenarios. We then expand upon these ideas, and venture into multivariate functions that exhibit convexity (or concavity) in each variable individually, but may not possess this property jointly across all variables. Furthermore, building upon similar underlying principles, we extend our investigations to derive upper and lower bounds on the expectations of functions that do not necessarily exhibit convexity or concavity across their entire domain. These diverse contributions collectively enrich the toolbox of RJIs and broaden their potential utility in a wide array of practical and theoretical contexts.

We commence our exploration from a foundational point that closely resembles [224, Lemma 1]. Let $f : \mathbb{R}^+ \to \mathbb{R}$ be a non-decreasing concave function, and $\mathbb{1}[X > a]$ denote the indicator function of event $\{X > a\}$. Then,

$$f(x) = \frac{f(x) - f(0)}{x} \cdot x + f(0) \tag{5.97}$$
$$\geq \frac{f(x) - f(0)}{x} \cdot x \cdot \mathbb{1}\{x \leq a\} + f(0) \tag{5.98}$$
$$\geq \frac{f(a) - f(0)}{a} \cdot x \cdot \mathbb{1}\{x \leq a\} + f(0) \tag{5.99}$$
$$= \frac{f(a) - f(0)}{a} \cdot x \cdot [1 - \mathbb{1}\{x > a\}] + f(0). \tag{5.100}$$

Letting now X be a non-negative RV with a finite mean, $\mathbb{E}\{X\} = \mu$, it is readily seen that by taking expectations of both sides that

$$\mathbb{E}\{f(X)\} \geq \sup_{a>0} \left[\frac{\mu}{a} \cdot f(a) + \left(1 - \frac{\mu}{a}\right) \cdot f(0) - \frac{f(a) - f(0)}{a} \cdot \mathbb{E}\{X \cdot \mathbb{1}[X > a]\}\right]. \tag{5.101}$$

This foundational inequality sets the stage for our subsequent derivations. The primary challenge at this juncture is to evaluate the term:

$$q(a) \equiv \mathbb{E}\{X \cdot \mathbb{1}[X > a]\}. \tag{5.102}$$

In straightforward cases, the exact calculation of $q(a)$ is achievable through closed-form expressions. Examples include scenarios where the PDF of X follows uniform, triangular, or exponential distributions, among others. However, for the majority of cases that pique our interest, obtaining an exact, closed-form expression for $q(a)$ becomes a formidable task, if not an impossibility. Consequently, we must rely on upper bounds to further constrain the RHS of (5.101).

In situations where the computation of $q(a)$ eludes an exact closed-form expression, we introduce two fundamental alternative approaches for bounding $q(a)$. Both approaches share a common feature: When the RV X tightly concentrates around its mean μ, even slight deviations of a from μ result in small values for $q(a)$. This characteristic ensures that our bounds closely approach the value of $f(\mu)$. The selection between these two approaches depends on the specific problem under consideration and the feasibility of obtaining closed-form expressions for the moments involved, if such expressions exist at all.

1. *The Chernoff approach.* The first approach is to upper bound the indicator function, $\mathbb{1}\{x > a\}$ by the exponential function $e^{s(x-a)}$ $(s \geq 0)$, akin to Chernoff's bound. This results in

$$q(a) \leq \inf_{s \geq 0} \mathbb{E}\{Xe^{s(X-a)}\} \tag{5.103}$$

$$= \inf_{s \geq 0} \left[e^{-as}\mathbb{E}\{Xe^{sX}\}\right] \tag{5.104}$$

$$= \inf_{s \geq 0} \left[e^{-as}\Phi'(s)\right] \tag{5.105}$$

$$\triangleq q_{\text{Chernoff}}(a), \tag{5.106}$$

where $\Phi'(s)$ is the derivative of the MGF, $\Phi(s) \triangleq \mathbb{E}\{e^{sX}\}$. Thus, (5.101) is further lower bounded as

$$\mathbb{E}\{f(X)\} \geq \sup_{a>0} \left[\frac{\mu}{a} \cdot f(a) + \left(1 - \frac{\mu}{a}\right) \cdot f(0) - \frac{f(a) - f(0)}{a} \cdot q_{\text{Chernoff}}(a)\right]. \tag{5.107}$$

This bound proves to be particularly valuable when the RV X possesses a finite MGF, denoted as $\Phi(s)$, within a certain range of positive s values. Furthermore, it is essential that $\Phi(s)$ is differentiable within this range. To ensure the practicality of this bound, it is crucial that $q_{\text{Chernoff}}(a)$ can be expressed in a reasonably straightforward closed-form manner. A slight variation of the Chernoff approach involves bounding not just the indicator function factor but the entire function $x \cdot \mathbb{1}[x > a]$ by an exponential function of the form $a \cdot e^{s(x-a)}$. To ensure the effectiveness of this approach, we choose s such that the derivative WRT x at $x = a$ is not less than 1. This ensures that the exponential function is at least tangential to the function $x \cdot \mathbb{1}[x > a]$ as x approaches a from above. Mathematically, this condition can be expressed as $as \geq 1$, which implies that s should be greater than or equal to $1/a$. Thus,

$$q(a) \leq a \cdot \inf_{s \geq 1/a} \{e^{-as}\Phi(s)\} \triangleq \tilde{q}_{\text{Chernoff}}(a), \tag{5.108}$$

which, of course, may replace $q_{\text{Chernoff}}(a)$ in (5.107). This bound can be applied in the same cases as $q_{\text{Chernoff}}(a)$. It has the small advantage that there is no need to differentiate $\Phi(s)$, but the range of the optimization over s is somewhat smaller.

2. *The Chebychev–Cantelli approach.* According to this approach, the function $x \cdot \mathbb{1}[x > a]$ is upper bounded by a quadratic function, in the spirit of the Chebychev–Cantelli inequality, *i.e.*,

$$x \cdot \mathbb{1}[x > a] \leq \frac{a(x+s)^2}{(a+s)^2}, \tag{5.109}$$

where the parameter $s \geq 0$ is optimized under the constraint that the derivative at $x = a$, which is $2a/(a+s)$, is at least 1 (again, to be at least tangential to the function itself at $x \downarrow a$), which is equivalent to the requirement, $s \leq a$. In this case, denoting $\sigma^2 = \text{Var}\{X\}$, we get

$$q(a) \leq \frac{a\mathbb{E}\left\{(X+s)^2\right\}}{(a+s)^2} = \frac{a\left[\sigma^2 + (\mu+s)^2\right]}{(a+s)^2}, \tag{5.110}$$

which, when minimized over $s \in [0, a]$, yields

$$s^* = \min\left\{a, \frac{\sigma^2}{a-\mu} - \mu\right\}, \tag{5.111}$$

and then the best bound is given by

$$q(a) \leq q_{\text{Cheb-Cant}}(a) \triangleq \begin{cases} \frac{\sigma^2+(a+\mu)^2}{4a}, & a < a_{\text{c}} \\ \frac{a\sigma^2}{\sigma^2+(a-\mu)^2}, & a \geq a_{\text{c}} \end{cases}, \tag{5.112}$$

where $a_{\text{c}} \triangleq \sqrt{\sigma^2 + \mu^2}$.

The Chernoff approach often outperforms the Chebychev–Cantelli approach in many scenarios. Let us consider an example to illustrate this point. Suppose we have an RV X expressed as the sum of n IID RVs, $Y_1, Y_2, \ldots, Y_n$, all with the same mean μ_Y, variance σ_Y^2, and MGF $\Phi_Y(s)$. In this case, we can readily calculate that $\mu = n\mu_Y$, $\sigma^2 = n\sigma_Y^2$, and $\Phi(s) = [\Phi_Y(s)]^n$. Additionally, for the sake of simplicity, let us assume that $f(0) = 0$. Now, if we aim to apply the Chebychev–Cantelli approach, we typically end up with a bound that relies on the variance of X and its mean, which are both multiplied by n. This often results in a relatively loose bound due to the dependence on the sample size n. On the other hand, when we employ the Chernoff approach, we leverage the MGF of X and, consequently, the MGF of Y_i, which remains unchanged as n grows. This approach frequently yields tighter bounds, even when n is substantial. Thus, in cases like this, the Chernoff approach tends to be more effective in providing more accurate and meaningful bounds. In particular, Chernoff's bounds yields

$$\begin{aligned}
&\mathbb{E}\left\{f\left(\sum_{i=1}^{n} Y_i\right)\right\} \\
&\geq \frac{n\mu_Y}{a} \cdot f(a) - \frac{f(a)}{a} \inf_{s\geq 0}\left\{e^{-sa}\frac{\mathrm{d}}{\mathrm{d}s}[\Phi_Y(s)]^n\right\} && (5.113)\\
&= \frac{n\mu_Y}{a} \cdot f(a) - \frac{nf(a)}{a} \inf_{s\geq 0}\left\{e^{-sa}[\Phi_Y(s)]^{n-1}\Phi_Y'(s)\right\} && (5.114)\\
&= \frac{nf(a)}{a}\left[\mu_Y - \inf_{s\geq 0}\left\{e^{-sa}[\Phi_Y(s)]^n \cdot \frac{\mathrm{d}\ln\Phi_Y(s)}{\mathrm{d}s}\right\}\right]. && (5.115)
\end{aligned}$$

Now, if $Y_1, Y_2, \ldots$ obey a large-deviations principle, the second term in the square brackets tends to zero exponentially for the choice $a = n(\mu_Y + \epsilon)$ with arbitrarily small $\epsilon > 0$. In this case, let $s^* > 0$ be the maximizer of $[s(\mu+\epsilon) - \ln\Phi_Y(s)]$, and denote $I(\epsilon) = s^*(\mu+\epsilon) - \ln\Phi_Y(s^*)$. Then,

$$\mathbb{E}\left\{f\left(\sum_{i=1}^{n} Y_i\right)\right\} \geq \frac{f[(\mu_Y+\epsilon)n]}{\mu_Y+\epsilon}\left[\mu_Y - e^{-nI(\epsilon)}\frac{\mathrm{d}\ln\Phi_Y(s)}{\mathrm{d}s}\bigg|_{s=s^*}\right]. \quad (5.116)$$

For large enough n, the second term in the square brackets becomes negligible, and the lower bound becomes arbitrarily close to $f[(\mu_Y + \epsilon)n] \cdot \mu_Y/(\mu_Y+\epsilon)$. On the other hand, Jensen's upper bound is $f(\mu_Y n)$. In some cases, the difference is not very large, at least for asymptotic evaluations. For example, if $f(x) = \ln(1+x)$, which is a frequently encountered concave function in information theory, $\ln[1+n(\mu_Y+\epsilon)] \geq \ln n + \ln(\mu_Y+\epsilon)$, whereas $\ln(1+n\mu_Y) \leq \ln n + \ln(\mu_Y + 1/n)$, which are very close for large n and small $\epsilon > 0$.

In the Chebychev–Cantelli approach, on the other hand, we have $a_{\rm c} = \sqrt{n^2\mu_Y^2 + n\sigma_Y^2} \sim n\mu_Y$ for large n. Thus, if we take $a = n(\mu_Y + \epsilon) > a_{\rm c}$, we have

$$q_{\text{Cheb-Cant}}\left[n(\mu_Y + \epsilon)\right] = \frac{n\sigma_Y^2}{n\sigma_Y^2 + n^2\epsilon^2} = \frac{\sigma_Y^2}{\sigma_Y^2 + n\epsilon^2}, \tag{5.117}$$

which tends to zero, but only at the rate of $1/n$, as opposed to the exponential decay in the Chernoff approach. Still, for large n, the main term of the bound becomes asymptotically tight, as before.

In spite of the superiority of the Chernoff approach relative to the Chebychev–Cantelli approach, as we now demonstrated, one should keep in mind that there are also situations where the RV X does not have an MGF (*i.e.*, when the PDF of X has a heavy tail), yet it does have a mean and a variance. In such cases, the Chebychev–Cantelli approach is applicable while the Chernoff approach is not. But even when the MGF exists, in certain cases, the calculation of the first and the second moment are easier than the calculation of the exponential moment.

We summarize our main finding this section so far in the following inequality:

$$\mathbb{E}\{f(X)\} \geq \sup_{a>0}\left[\frac{\mu}{a}\cdot f(a) + \left(1 - \frac{\mu}{a}\right)\cdot f(0) - \frac{f(a) - f(0)}{a}\cdot q_{\min}(a)\right], \tag{5.118}$$

where

$$q_{\min}(a) \triangleq \min\left\{q_{\text{Chernoff}}(a), \tilde{q}_{\text{Chernoff}}(a), q_{\text{Cheb-Cant}}(a)\right\}. \tag{5.119}$$

We now demonstrate the lower bound in two information-theoretic application examples. More examples can be found in [142].

Example 5.4 (Capacity of the Gaussian channel with random SNR)**.** Consider a zero-mean, circularly symmetric complex Gaussian channel whose SNR, Z, is an RV (*e.g.*, due to fading), known to both the transmitter and the receiver. The capacity is given by $C = \mathbb{E}\{\ln(1+gZ)\}$, where g is a certain deterministic gain factor and the expectation is

WRT the randomness of Z. For simplicity, let us assume that Z is distributed exponentially, *i.e.*,

$$p_Z(z) = \theta e^{-\theta z}, \tag{5.120}$$

for $z \geq 0$, where the parameter $\theta > 0$ is given. In this case, $f(x) = \ln(1 + gx)$, $\mu = 1/\theta$ and $q(a)$ can be easily derived in closed form, to obtain

$$q(a) = \theta \cdot \int_a^\infty z e^{-\theta z}\,\mathrm{d}z = \left(a + \frac{1}{\theta}\right) \cdot e^{-\theta a}. \tag{5.121}$$

Consequently,

$$C \geq \sup_{a \geq 1/\theta} \frac{\ln(1+ga)}{a}\left[\frac{1}{\theta} - \left(a + \frac{1}{\theta}\right) \cdot e^{-a\theta}\right] \tag{5.122}$$

$$= \sup_{s \geq 1} \left[\frac{1-(s+1)e^{-s}}{s}\right] \cdot \ln\left(1 + \frac{gs}{\theta}\right), \tag{5.123}$$

whereas the Jensen upper bound is $C \leq \ln(1 + g/\theta)$. A plot of the bounds can be found in [142], which shows that the bounds become tight for large θ (high SNR).

Example 5.5 (Universal source coding)**.** Let us delve into the evaluation of the expected code length linked with the universal lossless source code developed by Krichevsky and Trofimov [108]. In essence, this code serves as a universal solution for encoding memoryless sources. In the binary context, at each time step t, it systematically assigns probabilities to the next binary symbol based on a biased version of the empirical distribution derived from the source data observed up to that point, denoted as $s_1, s_2, \ldots, s_t$. To be more specific, let us examine the ideal code-length function (measured in nats):

$$L(s^n) = -\sum_{t=0}^{n-1} \ln Q(s_{t+1}|s_1, \ldots, s_t), \tag{5.124}$$

where

$$Q(s_{t+1} = s|s_1, \ldots, s_t) = \frac{N_t(s) + 1}{t + 2}, \tag{5.125}$$

and $N_t(s)$, $s \in \{0, 1\}$, is the number of occurrences of the symbol s in $(s_1, \ldots, s_t)$. Therefore,

$$\mathbb{E}\left\{L(S^n)\right\} = \sum_{t=0}^{n-1} \ln(t+2) - \sum_{t=0}^{n-1} \mathbb{E}\{\ln[N_t(S_{t+1}) + 1]\} \tag{5.126}$$

$$= \ln[(n+1)!] - \sum_{t=0}^{n-1} \mathbb{E}\left\{\ln\left(1 + \sum_{i=0}^{t} \mathbb{1}[S_i = S_{t+1}]\right)\right\} \quad (5.127)$$

$$= \ln[(n+1)!] - p \cdot \sum_{t=0}^{n-1} \mathbb{E}\left\{\ln\left(1 + \sum_{i=0}^{t} \mathbb{1}[S_i = 1]\right)\right\}$$

$$- (1-p) \cdot \sum_{t=0}^{n-1} \mathbb{E}\left\{\ln\left(1 + \sum_{i=0}^{t} \mathbb{1}[S_i = 0]\right)\right\}, \quad (5.128)$$

where $\mathbb{1}[\cdot]$ are indicator functions of the corresponding events and where p and $1-p$ are the probabilities of '1' and '0', respectively. To establish an upper bound for $\mathbb{E}\{L(S^n)\}$, one can now use (5.116) for lower bounds for each of the terms: $\mathbb{E}\{\ln(1 + \sum_{i=0}^{t} \mathbb{1}[S_i = 1])\}$ and $\mathbb{E}\{\ln(1 + \sum_{i=0}^{t} \mathbb{1}[S_i = 0])\}$, which are approximately $\ln(1+np)$ and $\ln[1+n(1-p)]$, respectively.

5.6 Jensen-Like Inequalities

In this section, which summarizes the main findings of [144], we consider inequalities that are founded upon a fundamental insight closely tied to the derivation of the ordinary Jensen inequality. This insight revolves around the relationship between a given convex function, denoted as $f(x)$, and the tangential affine function, $\ell(x) = f(a) + f'(a)(x-a)$. Here, a is an arbitrary value within the domain of x, and $f'(a)$ represents the derivative of f at the point $x = a$ (assuming the differentiability of f at that point). By strategically choosing a to be $\mathbb{E}\{X\}$ (the expected value of the RV X) and subsequently taking expectations of both sides of the inequality $f(X) \geq \ell(X)$, we can effortlessly establish the traditional Jensen inequality (5.96). This crucially hinges on the fact that $a_* = \mathbb{E}\{X\}$ constitutes the optimal selection of a in the context of maximizing $\mathbb{E}\{\ell(X)\}$ across all potential values of a. This, in turn, furnishes us with the most stringent lower bound within the scope of lower bounds for $\mathbb{E}\{f(X)\}$. However, it is worth noting that the optimal choice of a may differ when we are dealing with more intricate expressions where the expectation needs to be lower bounded. For instance, one might seek to establish a lower bound for $\mathbb{E}\{g[f(X)]\}$, where g is a monotonically non-decreasing function, or $\mathbb{E}\{f(X)g(X)\}$, where g is a non-negative and/or convex function, or perhaps a combination of these conditions and more. In such cases, the optimal choice of a could deviate from $\mathbb{E}\{X\}$.

To illustrate this point, let g be a non-negative function, and let us examine the lower bound of $\mathbb{E}\{f(X)g(X)\}$. In this scenario, we can establish the following inequality:

$$\mathbb{E}\left\{f(X)g(X)\right\} \geq \mathbb{E}\left\{[f(a) + f'(a)(X-a)]g(X)\right\}. \tag{5.129}$$

By optimizing the RHS over the parameter a, we can easily determine the optimal choice for a, denoted as a_*:

$$a_* = \frac{\mathbb{E}\{Xg(X)\}}{\mathbb{E}\{g(X)\}}. \tag{5.130}$$

This result leads to the inequality:

$$\mathbb{E}\left\{f(X)g(X)\right\} \geq f\left(\frac{\mathbb{E}\{Xg(X)\}}{\mathbb{E}\{g(X)\}}\right) \cdot \mathbb{E}\left\{g(X)\right\}. \tag{5.131}$$

This inequality proves valuable, provided that we can readily compute both $\mathbb{E}\{g(X)\}$ and $\mathbb{E}\{Xg(X)\}$ for the given function g. Our first example concerns a function that is intimately related to the Shannon entropy.

Example 5.6 (An entropy-related function). Letting $f(x) = -\ln x$ and $g(x) = x$ for $x > 0$, we obtain

$$\mathbb{E}\left\{-X\ln X\right\} \geq -\mathbb{E}\{X\} \cdot \ln\frac{\mathbb{E}\{X^2\}}{\mathbb{E}\{X\}} \tag{5.132}$$

$$= -\mathbb{E}\{X\} \cdot \ln(\mathbb{E}\{X\}) - \mathbb{E}\{X\} \cdot \ln\left(1 + \frac{\mathrm{Var}\{X\}}{[\mathbb{E}\{X\}]^2}\right). \tag{5.133}$$

Notice that the function $-x\ln x$ exhibits concavity, rather than convexity. Nevertheless, we establish a lower bound, not an upper one, on its expectation, thereby unveiling a RJI. The right-most side of the expression comprises two components: The initial term represents the standard Jensen upper bound for $\mathbb{E}\{-X\ln X\}$, while the second term accounts for the gap. This gap is contingent not only upon the expectation of X but also on its variance, reflecting the fluctuations around $\mathbb{E}\{X\}$. Clearly, in scenarios where $\text{Var}\{X\} = 0$, the second term disappears — a logical outcome, as a degenerate RV causes Jensen's inequality to hold with equality, eliminating any gap. This inequality promptly finds application in deriving a lower bound for the expected empirical entropy of a sequence generated by a memoryless source. Such an application is significant to universal source coding, as detailed in [108] (see more details in [144]).

Another important example is associated with moments.

Example 5.7 (Bounds on moments)**.** Let s and t be two real numbers whose difference, $s-t$, is either negative or larger than unity. Now, let $g(x) = x^t$, and $f(x) = x^{s-t}$. Then,

$$\mathbb{E}\left\{X^s\right\} = \mathbb{E}\left\{X^t X^{s-t}\right\} \tag{5.134}$$

$$\geq \left(\frac{\mathbb{E}\{X^{t+1}\}}{\mathbb{E}\{X^t\}}\right)^{s-t} \cdot \mathbb{E}\{X^t\} \tag{5.135}$$

$$= \frac{(\mathbb{E}\{X^{t+1}\})^{s-t}}{(\mathbb{E}\{X^t\})^{s-t-1}}. \tag{5.136}$$

In particular, for $t = 1$ and $s \notin (1,2)$, this becomes

$$\mathbb{E}\left\{X^s\right\} \geq \frac{\left(\mathbb{E}\{X^2\}\right)^{s-1}}{\left(\mathbb{E}\{X\}\right)^{s-2}} = [\mathbb{E}\{X\}]^s \cdot \left(1 + \frac{\text{Var}\{X\}}{[\mathbb{E}\{X\}]^2}\right)^{s-1}, \tag{5.137}$$

which is, once again, a bound that depends only on the first two moments of X. For $s \in (0,1)$, the function x^s exhibits concavity, resulting in a RJI. Conversely, when $s \leq 0$ or $s \geq 2$, the function x^s is convex, giving rise to an enhanced version of Jensen's inequality. In this enhanced version, the first term, $[\mathbb{E}\{X\}]^s$, corresponds to the standard Jensen inequality, while the second factor quantifies the degree of enhancement. This enhancement is contingent on the relative fluctuation term, $\text{Var}\{X\}/[\mathbb{E}\{X\}]^2$. Naturally, the extent of improvement hinges on the variance of X. When the variance tends to zero, there is no room for improvement since the standard Jensen inequality attains equality. In contrast, a larger variance results in a wider gap between the conventional Jensen bound, $[\mathbb{E}\{X\}]^s$, and the enhanced counterpart. This underscores the importance of optimizing the parameter a, as opposed to the default choice of $a = \mathbb{E}\{X\}$ in the standard Jensen inequality.

Another family of Jensen-like bounds is associated with the product of two non-negative convex functions. Let f and g be non-negative, differentiable convex functions of $x \geq 0$, where f is monotonically non-decreasing. Then,

$$\begin{aligned}
&\mathbb{E}\left\{f(X)g(X)\right\} \\
&\geq \mathbb{E}\left\{[f(a) + f'(a)(X-a)] \cdot g(X)\right\} && (5.138)\\
&= [f(a) - af'(a)] \cdot \mathbb{E}\left\{g(X)\right\} + f'(a)\mathbb{E}\left\{Xg(X))\right\} && (5.139)\\
&\stackrel{(*)}{\geq} [f(a) - af'(a)]\, \mathbb{E}\left\{[g(b) + g'(b)(X-b)]\right\} + \\
&\qquad f'(a)\mathbb{E}\left\{X\left[g(c) + g'(c)(X-c)\right]\right\} && (5.140)\\
&= [f(a) - af'(a)] \cdot [g(b) - bg'(b) + g'(b)\mathbb{E}\{X\}] + \\
&\qquad f'(a)\left[(g(c) - cg'(c))\mathbb{E}\{X\} + g'(c)\mathbb{E}\left\{X^2\right\}\right], && (5.141)
\end{aligned}$$

where $(*)$ follows since the convexity of f implies that $f(a) \geq af'(a) \geq 0$. Maximizing the right-most side over a, b and c, one obtains the inequality:

$$\mathbb{E}\left\{f(X)g(X)\right\} \geq f\left(\frac{\mathbb{E}\{X\} \cdot g(\mathbb{E}\{X^2\}/\mathbb{E}\{X\})}{g(\mathbb{E}\{X\})}\right) \cdot g\left(\mathbb{E}\{X\}\right). \quad (5.142)$$

Example 5.8 (Second moment of Gaussian capacity)**.** Consider the example of the AWGN channel with a random SNR, denoted as Z. In this context, we aim to bound the variance of the (instantaneous) capacity, denoted as $c(Z)$. This variance is important in order to assess the random fluctuations of the quality of the channel [207]. Indeed, if Z is deterministic then the variance is zero, and $c(Z)$ is constant, trivially equals to $\mathbb{E}[c(Z)]$. So, reliable information can be sent over the channel at rate $\mathbb{E}[c(Z)]$ at arbitrarily small error probability, using a capacity achieving code [171]. If, however, Z is random, then the system may deploy a capacity achieving code at some chosen rate R, *e.g.*, the expected capacity $\mathbb{E}[c(Z)]$. In this case, if the variance of $c(Z)$ is large, then there is a high probability for an outage event, to wit, the event $c(Z) \leq R$ in which the instantaneous capacity of the system does not suffice to support the coding rate R, and the decoding error probability is high [10], [57]. Thus, for such channels, the error probability is dominated by the outage probability, which in turn, is directly related to the variance of $c(Z)$.

The variance of $c(Z)$ can be expressed as follows:

$$\text{Var}\{c(Z)\} = \mathbb{E}\left\{c^2(Z)\right\} - [\mathbb{E}\{c(Z)\}]^2 \tag{5.143}$$

$$= \mathbb{E}\left\{\ln^2(1+gZ)\right\} - [\mathbb{E}\{\ln(1+gZ)\}]^2. \tag{5.144}$$

To establish an upper bound for $\text{Var}\{c(Z)\}$, we can derive upper bounds for both $\mathbb{E}\{\ln^2(1+gZ)\}$ and a lower bound for $\mathbb{E}\{\ln(1+gZ)\}$. For the former, we can utilize the inequality presented here, employing $f(z) = g(z) = \ln(1+gz)$. This yields the following upper bound, relying solely on the first two moments of Z:

$$\mathbb{E}\left\{\ln^2(1+gZ)\right\} \leq \ln(1+g\mathbb{E}\{Z\}) \cdot \ln\left(1 + \frac{g\mathbb{E}\{Z\}\ln(1+g\mathbb{E}\{Z^2\}/\mathbb{E}\{Z\})}{\ln(1+g\mathbb{E}\{Z\})}\right). \tag{5.145}$$

Notably, the function $\ln^2(1+gz)$ is neither convex nor concave. Nevertheless, our approach provides an upper bound that can be easily computed, given the ability to calculate the first two moments of Z.

These are just a few out of many more examples provided in [144]. The main features of the results on Jensen-like inequalities in general, are the following. Firstly, in many instances, such as the one mentioned above, we can analytically determine the optimal value of a parameter (*e.g.*, a in the preceding discussion). However, in cases where closed-form optimization is not feasible, we have two viable options: (i) Perform numerical optimization or (ii) select an arbitrary value for a and derive a valid lower bound. It is important to note that a well-informed choice for a can potentially yield a robust lower bound. Secondly, these inequalities offer two distinct types of bounds: (i) Bounds that necessitate computing the first two moments (or equivalently, the first two cumulants) of the RV X, and (ii) bounds that require calculating the MGF of X and its derivative, or equivalently, the cumulant generating function of X and its derivative. These moment calculations are often straightforward, especially in scenarios where X is represented as the sum of IID RVs — a common occurrence in information-theoretic applications.

It should also be noted that the classes of Jensen-like inequalities provide ample flexibility for deriving lower bounds on functions that may not be inherently convex, some may even be concave. This opens the door to an alternative approach for RJIs, different than those discussed in Section 5.5. This can be achieved by representing the given function within one of the discussed categories, such as a product of a convex function and a non-negative function, a product of two non-negative convex functions, or a composition of a monotone function and a convex function. Finally, the Jensen-like inequalities possess the desirable property of tightening as the RV X becomes increasingly concentrated around its mean, akin to the conventional Jensen inequality.

6

Summary, Outlook and Open Issues

In this monograph, we have provided an analytical toolbox for information-theoretic analysis. We have described a generalization of the method of types, which allows to address settings that go beyond the finite alphabet case, including the prominent example of Gaussian sources and channels, possibly with memory. This allows to evaluate the volumes of various high-dimensional sets, and thus also their probability. We have also described a generalization of this method to distributions from exponential families. Further generalizing and refining such extensions to broader classes of distributions is an interesting path for future research. We have then described the saddle-point method for integration, which not only allows to evaluate the pre-exponent of volumes or probabilities, it is also necessary in the evaluation of redundancy rates, and may provide solutions in settings for which the method of types fails.

We then continued to present the TCEM, for evaluating the exponential behavior of random codes. The method is principled, allows to analyze optimal decoders, and is guaranteed to provide exponentially tight results. It also provides the best known exponents in diverse problem settings. Future research may further explore additional settings, *e.g.*, the error exponent of the typical random code in multi-user configurations [59]. An additional important future research direction is to consider structured random-ensembles. The TCEM relies on the assumption that the codewords in the ensemble are drawn at random, IID (or some variant of such a random ensemble). For practical decoding algorithms, codes must have some structure, *e.g.*, linear codes over finite fields, lattice codes for real/complex-input channels [229], convolutional codes or trellis-codes, or even well-defined structure such as turbo-codes [17], LDPC codes [171], polar codes [8], and so on. It is of interest to develop methods, akin to the TCEM, to accurately analyze the error exponents of such codes. In addition, it is also of interest to explore methods inspired by the TCEM in derivation of converse results, in the finite-blocklength regime [166], in the moderate-deviations regime [5], [167] and so on. We have briefly mentioned a few such initial results, which hints at the possibility of enriching this direction. Finally, it is also of interest to further delve into the optimization problems involved in the computation of exponents obtained by the TCEM, and develop efficient, and perhaps "general-purpose", solvers, to solve them.

We then considered the tight evaluation of expectations of non-linear functions of RVs, including integral representations and a few variants of Jensen's inequality. These techniques are highly useful in information theory, as information measures typically involve such expectations. For RJI, we have emphasized that it approaches the standard Jensen inequality, when the RV of interest is tightly concentrated around its mean value. It is thus of interest to relate the RJI we considered to concentration-of-measure ideas [20].

Acknowledgements

We thank the anonymous reviewers for their very useful comments, and the editor, Albert Guillén i Fàbregas, for his many insightful comments and suggestions. We also thank Igal Sason for his comments on the manuscript. We believe that the comments have helped to significantly improve the manuscript. The work of N. W. was supported by the Israel Science Foundation (ISF), grant no. 1782/22.

Appendices

A

On the Tightness of Chernoff's Bound via the Method of Types

Let P be a memoryless source over an alphabet $\mathcal{X}$. For simplicity, we focus on finite-alphabet sources, though a similar derivation can be carried out using the extended method of types developed in Section 2 for more general sources. Let f be a real function of probability distributions over $\mathcal{X}$, and $\alpha \in \mathbb{R}$. Then,

$$\Pr\left[f(\hat{P}_{\boldsymbol{x}}) \geq \alpha\right]$$

$$= \sum_{\boldsymbol{x}\in\mathcal{X}^n} P(\boldsymbol{x}) \cdot \mathbb{1}\left[f(\hat{P}_{\boldsymbol{x}}) \geq \alpha\right] \tag{A.1}$$

$$\stackrel{(a)}{=} \sum_{\boldsymbol{x}\in\mathcal{X}^n} P(\boldsymbol{x}) \cdot \inf_{s\geq 0} e^{ns[f(\hat{P}_{\boldsymbol{x}})-\alpha]} \tag{A.2}$$

$$\stackrel{(b)}{=} \sum_{Q} e^{-n\cdot D(Q||P)} \cdot \inf_{s\geq 0} e^{ns[f(\hat{P}_{\boldsymbol{x}})-\alpha]} \tag{A.3}$$

$$\stackrel{(c)}{\doteq} \exp\left[-n \cdot \min_{Q}\left\{D(Q||P) - \inf_{s\geq 0} s\left[f(\hat{P}_{\boldsymbol{x}}) - \alpha\right]\right\}\right] \tag{A.4}$$

$$= \exp\left[-n \cdot \min_{Q} \sup_{s\geq 0}\left\{D(Q||P) - s\left[f(\hat{P}_{\boldsymbol{x}}) - \alpha\right]\right\}\right] \tag{A.5}$$

$$\stackrel{(d)}{\leq} \exp\left[-n \cdot \sup_{s\geq 0} \min_{Q}\left\{D(Q||P) - s\left[f(\hat{P}_{\boldsymbol{x}}) - \alpha\right]\right\}\right] \tag{A.6}$$

$$= \inf_{s\geq 0} \exp\left[-n \cdot \min_{Q}\left\{D(Q||P) - s\left[f(\hat{P}_{\boldsymbol{x}}) - \alpha\right]\right\}\right] \tag{A.7}$$

$$\overset{(e)}{\doteq} \inf_{s\geq 0} \sum_{\boldsymbol{x}\in\mathcal{X}^n} P(\boldsymbol{x}) \cdot e^{ns[f(\hat{P}_{\boldsymbol{x}})-\alpha]} \tag{A.8}$$

$$= \inf_{s\geq 0} \mathbb{E}\left[e^{ns[f(\hat{P}_{\boldsymbol{x}})-\alpha]}\right], \tag{A.9}$$

where (a) follows from the elementary bound $\mathbb{1}\{t \geq \alpha\} \leq e^{ns(t-\alpha)}$ that holds for any $s \geq 0$, (b) follows from the probability of a type class [(2.12) in Section 2.2.1], and where the summation is over all possible types, (c) follows since the number of possible types is polynomial in n [(2.2) in Section 2.2.1], and so the sum is exponentially on the same scale as the maximum element, (d) follows since maximin is always less or equal than the minimax, and (e) follows again from the method of types, reversing the reasoning above.

The final term in (A.9) is exactly Chernoff's bound for the event $\{f(\hat{P}_{\boldsymbol{x}}) \geq \alpha\}$. Importantly, if f is concave then the *minimax theorem* [188] implies the inequality in (d) above is, in fact, an equality, and so the chain of passages is exponentially tight. In many applications, f is affine (*e.g.*, the empirical mean of some cost) and thus concave, and so Chernoff's bound is assured to be *tight*. See [49] for a thorough discussion.

B

Computation of Exponents

In this appendix, we describe two possible approaches to efficiently compute or bound the exponents obtained using the TCEM. This aspect is an indispensable part of the TCEM, since it is possible for an error exponent to take a rather intricate formula. Indeed, recall that the TCEM exponents are given by Csiszár–Körner-style formulas, *e.g.*, as in (4.10). Thus, they involve a constrained optimization problem over joint distributions, and the dimensionality of the optimized joint distributions increases with the alphabet sizes of the problem (*e.g.*, input and output alphabets of the channel). Thus, a direct optimization, using an exhaustive search or "general-purpose" global optimization over the probability simplex may be prohibitively complex.

The first approach we consider is based on *Lagrange duality* [21] (see also [180, Appendix]), in which the original exponent optimization problem is considered to be the *primal* optimization problem. When deriving instead the *dual* optimization problem of the exponent, the result is a Gallager-style bound [71, Chapter 5], which is often rather easy to compute and plot for an entire range of rates, rather than for a specific rate; see (B.19) in what follows for a typical formula. This is especially useful in multiuser problems [59], for which even problem instances with binary alphabets lead to optimization problems in non-trivial dimensions. For example, for a broadcast channel problem with input alphabet $\mathcal{X}$ and two receivers, each with an alphabet $\mathcal{Y}$, a joint distribution of the input and the two outputs has dimensionality $|\mathcal{X}| \cdot |\mathcal{Y}|^2 - 1$, which is at least 7. In some of the problems, the number of optimization variables for the Gallager-style bound does not increase with the alphabet size of the source or channel. The downside is that, as we shall see, the derivation might include the utilization of bounds that may sacrifice tightness. Indeed, in minimization optimization problems, the value of the dual problem is a *lower bound* on the value of the primal problem, and if the primal optimization problem is convex then *strong duality* holds (under typically mild conditions) [21, Chapter 5], and both values are equal. However, there is no guarantee that the primal optimization problem of the exponent is convex, and sometimes obtaining reasonably simple dual problems requires additional steps, which may also sacrifice tightness.

The second approach is based on utilization of convex optimization solvers. While the optimization problem involved in the computation of the exponent may not be convex as is, in many cases it is possible to develop a procedure that allows to compute it by only solving convex optimization problems.

Moreover, typically, the primal problem involves mostly *minimization* operators (over joint types), while the dual problem involves *maximization* operators (over scalar parameters). From this aspect, the dual exponent is preferable, because even a sub-optimal choice of the dual variables leads to a valid bound on the exponent. Thus, *e.g.*, a coarse exhaustive search on the dual variables may be performed and still lead to a tight bound. In contrast, the minimization in the primal problem must be performed accurately in order to obtain a valid numerical value of the exponent. Nonetheless, it also possible for the primal problem to include a maximization operator (possibly intertwined between minimization operators), and the same holds for such maximization problems — any sub-optimal choice leads to a valid bound. In fact, in some cases, an educated guess for the maximizing primal variable may be proposed, and in some settings it is possible to show that this choice is actually optimal.

B.1 Exponent Computation by Lagrange Duality

Lagrange duality is based on the *minimax theorem* [188], stating the minimax value of a functional convex in the minimization variable and concave in the maximization variable equals to the maximin value. We will next exemplify this technique on the random-coding error exponent $E_{\text{rc},\alpha}(R, P_X)$ from (4.27), and derive a Lagrange dual lower bound on its value. As we have seen, if we consider the MMI rule, then the random-coding error exponent is greatly simplified to the standard random-coding error exponent in (4.10), which only contains a minimization over $Q_{Y|X}$ (with the minimization over $\tilde{Q}_{Y|X}$ removed). In accordance, it is not very difficult to obtain a dual Lagrange form of this exponent. In order to demonstrate a few other techniques that are generally useful for the TCE-based exponents, we will next let $\alpha(\cdot)$ be general, yet restricted to be a linear function of Q_{XY}, given by $\alpha(Q_{XY}) \triangleq \sum_{x\in\mathcal{X},y\in\mathcal{Y}} \alpha(x, y) \cdot Q(x, y)$ (this includes, *e.g.*, the ML decoder).

Let us start by writing the objective function of $E_{\text{rc},\alpha}(R, P_X)$ using a dual variable $\rho \in \mathbb{R}$ as

$$
\begin{aligned}
&E_{\text{rc},\alpha}(R, P_X) \\
&= \min_{Q_{Y|X},\tilde{Q}_{Y|X}} D(Q_{Y|X}||W|P_X) + \left[I(P_X \times \tilde{Q}_{Y|X}) - R\right]_+ &\text{(B.1)}\\
&= \min_{Q_{Y|X},\tilde{Q}_{Y|X}} D(Q_{Y|X}||W|P_X) + \max\left\{I(P_X \times \tilde{Q}_{Y|X}) - R, 0\right\} &\text{(B.2)}\\
&\overset{(*)}{=} \min_{Q_{Y|X},\tilde{Q}_{Y|X}} D(Q_{Y|X}||W|P_X) + \max_{\rho\in[0,1]} \rho \cdot \left[I(P_X \times \tilde{Q}_{Y|X}) - R\right] &\text{(B.3)}\\
&= \min_{Q_{Y|X},\tilde{Q}_{Y|X}} \max_{\rho\in[0,1]} D(Q_{Y|X}||W|P_X) + \rho \cdot \left[I(P_X \times \tilde{Q}_{Y|X}) - R\right], &\text{(B.4)}
\end{aligned}
$$

where $(*)$ follows from the identity $\max\{t, 0\} = \max_{\rho\in[0,1]} \rho t$. Now, the objective function is linear, and hence concave, in the maximizing variable ρ, and the interval $[0, 1]$ is convex. Moreover, $D(Q_{Y|X}||W|P_X)$ is convex in $Q_{Y|X}$ and $\rho \cdot I(P_X \times \tilde{Q}_{Y|X})$ is convex in $\tilde{Q}_{Y|X}$ (for $\rho \geq 0$), hence the objective functional is jointly convex in $(Q_{Y|X}, \tilde{Q}_{Y|X})$. The constraint set for $(Q_{Y|X}, \tilde{Q}_{Y|X})$, given by

$$\Big\{ Q_{Y|X}, \tilde{Q}_{Y|X} \colon (P_X \times Q_{Y|X})_Y = (P_X \times \tilde{Q}_{Y|X})_Y, \\ \alpha(P_X \times \tilde{Q}_{Y|X}) \geq \alpha(P_X \times Q_{Y|X}) \Big\}, \quad \text{(B.5)}$$

is the intersection of a hyperplane and a half space. We also note the implicit constraint that $Q_{Y|X}$ and $\tilde{Q}_{Y|X}$ are conditional probabilities, *i.e.*, $\sum_{y\in\mathcal{Y}} Q_{Y|X}(y|x) = \sum_{y\in\mathcal{Y}} \tilde{Q}_{Y|X}(y|x) = 1$ for all $x \in \mathcal{X}$ and $Q_{Y|X}(y|x), \tilde{Q}_{Y|X}(y|x) \geq 0$ for all $x \in \mathcal{X}, y \in \mathcal{Y}$. These are also convex constraints, and since the intersection of convex sets is convex, the constraint set for $(Q_{Y|X}, \tilde{Q}_{Y|X})$ is convex. So, the minimax theorem [188] implies that

$$E_{\mathrm{rc},\alpha}(R, P_X) = \\ \max_{\rho\in[0,1]} \min_{Q_{Y|X}, \tilde{Q}_{Y|X}} D(Q_{Y|X}||W|P_X) + \rho \cdot \Big[I(P_X \times \tilde{Q}_{Y|X}) - R \Big] \quad \text{(B.6)}$$

over the constraint set. We next focus on the inner minimization for a given $\rho \in [0,1]$. Following Lagrange duality [21, Chapter 5], we introduce dual variables $\lambda \geq 0$ and $\{\nu(y)\}_{y\in\mathcal{Y}} \subset \mathbb{R}$. The variable λ is for the inequality constraint $\alpha(P_X \times \tilde{Q}_{Y|X}) \geq \alpha(P_X \times Q_{Y|X})$, whereas the variables $\{\nu(y)\}_{y\in\mathcal{Y}}$ are for the constraint of equal output marginals, that is, the $|\mathcal{Y}|$ constraints $(P_X \times Q_{Y|X})_Y = (P_X \times \tilde{Q}_{Y|X})_Y$. Note that the constraint that $Q_{Y|X}$ and $\tilde{Q}_{Y|X}$ are conditional probability distributions is kept implicit. Hence, the minimization of interest is

$$\min_{Q_{Y|X}, \tilde{Q}_{Y|X}} \max_{\lambda\geq 0} \max_{\{\nu(y)\}_{y\in\mathcal{Y}}} D(Q_{Y|X}||W|P_X) + \rho \cdot \Big[I(P_X \times \tilde{Q}_{Y|X}) - R \Big] \\ + \sum_{y\in\mathcal{Y}} \nu(y) \cdot \left[\sum_{x\in\mathcal{X}} P_X(x) \left(\tilde{Q}_{Y|X}(y|x) - Q_{Y|X}(y|x) \right) \right] \\ + \lambda \cdot \left[\sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} \alpha(x,y) \cdot P_X(x) \left(Q_{Y|X}(y|x) - \tilde{Q}_{Y|X}(y|x) \right) \right]. \quad \text{(B.7)}$$

The minimax theorem now implies that we may interchange the minimization and maximization order. We next focus on the minimization, and begin by expressing the mutual information term via the *golden formula* using an arbitrary probability distribution S_Y on $\mathcal{Y}$, as

$$I(P_X \times \tilde{Q}_{Y|X}) = D(\tilde{Q}_{Y|X}||\tilde{Q}_Y|P_X) - D(\tilde{Q}_Y||S_Y) \tag{B.8}$$

$$= \min_{S_Y} D(\tilde{Q}_{Y|X}||S_Y|P_X). \tag{B.9}$$

Using this relation and slightly re-organizing the objective function, we are left with the minimization of the functional

$$\begin{aligned}\min_{S_Y} D(Q_{Y|X}||W|P_X)+ \\ &\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} P_X(x)Q_{Y|X}(y|x)\cdot[-\nu(y)+\lambda\cdot\alpha(x,y)] \\ &+\rho D(\tilde{Q}_{Y|X}||S_Y|P_X) \\ &+\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} P_X(x)\tilde{Q}_{Y|X}(y|x)\cdot[\nu(y)-\lambda\cdot\alpha(x,y)]\end{aligned} \tag{B.10}$$

over $(Q_{Y|X}, \tilde{Q}_{Y|X})$. It can be noticed that the minimization over $Q_{Y|X}$ is decoupled from the minimization over $\tilde{Q}_{Y|X}$, and each of them can be solved directly. Alternatively, we may use the *Donsker–Varadhan* variational formula [20, Corollary 4.15], [53], stating that for any two probability measures P_1 and P_2 on $\mathcal{Z}$ and a function $f\colon \mathcal{Z} \to \mathbb{R}$ that does not depend on P_1

$$\min_{P_2} \{D(P_2||P_1) + \mathbb{E}_{P_2}[f(Z)]\} = -\ln \mathbb{E}_{P_1}\left[e^{-f(Z)}\right]. \tag{B.11}$$

Let $W(\cdot|x)$ denote the conditional output of the channel given $x \in \mathcal{X}$. By employing (B.11) separately for each $x \in \mathcal{X}$ we get

$$\begin{aligned}&\min_{Q_{Y|X}} D(Q_{Y|X}||W|P_X) + \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} P_X(x)Q_{Y|X}(y|x)\cdot[-\nu(y)+\lambda\cdot\alpha(x,y)] \\ &= \sum_{x\in\mathcal{X}} P_X(x)\cdot\Bigg\{\min_{Q_{Y|X=x}} D(Q_{Y|X=x}||W(\cdot|x)) \\ &\qquad + \sum_{y\in\mathcal{Y}} Q_{Y|X}(y|x)\cdot[-\nu(y)+\lambda\cdot\alpha(x,y)]\Bigg\} \qquad (B.12) \\ &= -\sum_{x\in\mathcal{X}} P_X(x)\cdot\ln\left(\sum_{y\in\mathcal{Y}} W(y|x)\cdot e^{\nu(y)-\lambda\cdot\alpha(x,y)}\right). \qquad (B.13)\end{aligned}$$

Similarly, the minimization over $\tilde{Q}_{Y|X}$ leads to

$$\begin{aligned}
&\sum_{x\in\mathcal{X}} P_X(x)\cdot\Bigg\{\min_{\tilde{Q}_{Y|X=x}} \rho D(\tilde{Q}_{Y|X=x}||S_Y) \\
&\qquad\qquad + \sum_{y\in\mathcal{Y}} \tilde{Q}_{Y|X}(y|x)\cdot[\nu(y)-\lambda\cdot\alpha(x,y)]\Bigg\} \\
&= \min_{S_Y} -\rho \sum_{x\in\mathcal{X}} P_X(x)\cdot\ln\left(\sum_{y\in\mathcal{Y}} S_Y(y)\cdot e^{-[\nu(y)+\lambda\cdot\alpha(x,y)]/\rho}\right) \qquad \text{(B.14)}\\
&\geq \min_{S_Y} -\rho\ln\left(\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} P_X(x)S_Y(y)\cdot e^{-[\nu(y)+\lambda\cdot\alpha(x,y)]/\rho}\right), \qquad \text{(B.15)}
\end{aligned}$$

where the inequality follows from convexity and Jensen inequality, yet is *not* guaranteed to be tight. Since $\rho\in[0,1]$, minimizing this last term over S_Y corresponds to maximizing

$$\sum_{y\in\mathcal{Y}} S_Y(y)\sum_{x\in\mathcal{X}} P_X(x)\cdot e^{-[\nu(y)+\lambda\cdot\alpha(x,y)]/\rho}, \qquad \text{(B.16)}$$

which, due to Schwarz–Cauchy inequality, occurs when

$$S_Y(y) = \frac{\sum_{x\in\mathcal{X}} P_X(x)\cdot e^{-[\nu(y)+\lambda\cdot\alpha(x,y)]/\rho}}{\sum_{y\in\mathcal{Y}}\sum_{x\in\mathcal{X}} P_X(x)\cdot e^{-[\nu(y)+\lambda\cdot\alpha(x,y)]/\rho}}. \qquad \text{(B.17)}$$

The minimal value over S_Y is then

$$\begin{aligned}
&\min_{S_Y} -\rho\ln\left(\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} P_X(x)S_Y(y)\cdot e^{-[\nu(y)+\lambda\cdot\alpha(x,y)]/\rho}\right) \\
&= -\rho\ln\left(\frac{\sum_{y\in\mathcal{Y}}\left(\sum_{x\in\mathcal{X}} P_X(x)e^{-[\nu(y)+\lambda\cdot\alpha(x,y)]/\rho}\right)^2}{\sum_{y\in\mathcal{Y}}\sum_{x\in\mathcal{X}} P_X(x)\cdot e^{-[\nu(y)+\lambda\cdot\alpha(x,y)]/\rho}}\right). \qquad \text{(B.18)}
\end{aligned}$$

We thus conclude the dual lower bound

$$\begin{aligned}
&E_{\text{rc},\alpha}(R,P_X) \\
&\geq -\sum_{x\in\mathcal{X}} P_X(x)\cdot\ln\left(\sum_{y\in\mathcal{Y}} W(y|x)\cdot e^{\nu(y)-\lambda\cdot\alpha(x,y)}\right)
\end{aligned}$$

$$- \rho \ln \left(\frac{\sum_{y\in\mathcal{Y}} \left(\sum_{x\in\mathcal{X}} P_X(x) e^{-[\nu(y)+\lambda\cdot\alpha(x,y)]/\rho} \right)^2}{\sum_{y\in\mathcal{Y}} \sum_{x\in\mathcal{X}} P_X(x) \cdot e^{-[\nu(y)+\lambda\cdot\alpha(x,y)]/\rho}} \right), \quad \text{(B.19)}$$

for any choice of $\rho \in [0,1]$, $\lambda \geq 0$ and $\{\nu(y)\}_{y\in\mathcal{Y}} \subset \mathbb{R}$.

Let us compare the primal optimization in (B.1), with the dual lower bound (B.19). The primal problem is a minimization problem of dimension $2|\mathcal{X}|(|\mathcal{Y}|-1)$ over a constrained set $(Q_{Y|X}, \tilde{Q}_{Y|X})$ (the constraints further reduce the dimension by $|\mathcal{Y}|+1$). For the exact exponent, this minimization must be accurately solved. By comparison, the dual exponent is a lower bound on the exact exponent [recall (B.15)], and can be maximized over dimension $|\mathcal{Y}|+2$. Nonetheless, this maximization can be performed in a crude manner, since any choice of the dual parameters leads to a valid lower bound on the exponent.

For additional derivations of dual Lagrange exponents formulations and Gallager-style bounds, see [41, Exercise 10.24] and [165] (in Russian), and in the context of the TCEM, see [11], [137], [177].

B.2 Exponent Computation Procedures with Convex Optimization Solvers

As we have seen, we may write

$$E_{\mathrm{rc},\alpha}(R, P_X) = \max_{\rho\in[0,1]} \min_{Q_{Y|X}, \tilde{Q}_{Y|X}} D(Q_{Y|X}||W|P_X) + \rho \cdot \left[I(P_X \times \tilde{Q}_{Y|X}) - R \right], \quad \text{(B.20)}$$

and when $\alpha(Q_{XY})$ is a linear function of Q_{XY}, then the feasible set of $(Q_{Y|X}, \tilde{Q}_{Y|X})$ is convex. Hence, the inner minimization problem is a convex optimization problem that can be efficiently solved. However, in principle, it should be solved for the continuous set of values $\rho \in [0,1]$. We next describe an alternative method to evaluate $E_{\mathrm{rc},\alpha}(R, P_X)$.

Let us write $E_{\mathrm{rc},\alpha}(R, P_X) = \min\{E_-(R), E_+(R)\}$ where[1]

$$E_-(R) = \min_{Q_{Y|X}, \tilde{Q}_{Y|X}} D(Q_{Y|X}||W|P_X), \quad \text{(B.21)}$$

[1]For brevity, we omit the explicit dependence on the score α and the input distribution P_X.

where the minimization is over the set

$$\Big\{Q_{Y|X}, \tilde{Q}_{Y|X} \colon (P_X \times Q_{Y|X})_Y = (P_X \times \tilde{Q}_{Y|X})_Y,$$

$$\alpha(P_X \times \tilde{Q}_{Y|X}) \geq \alpha(P_X \times Q_{Y|X}),\ I(P_X \times \tilde{Q}_{Y|X}) \leq R\Big\}, \quad \text{(B.22)}$$

and where

$$E_+(R) = \min_{Q_{Y|X}, \tilde{Q}_{Y|X}} D(Q_{Y|X}||W|P_X) + I(P_X \times \tilde{Q}_{Y|X}) - R, \quad \text{(B.23)}$$

where the minimization over the set

$$\Big\{Q_{Y|X}, \tilde{Q}_{Y|X} \colon (P_X \times Q_{Y|X})_Y = (P_X \times \tilde{Q}_{Y|X})_Y,$$

$$\alpha(P_X \times \tilde{Q}_{Y|X}) \geq \alpha(P_X \times Q_{Y|X}),\ I(P_X \times \tilde{Q}_{Y|X}) \geq R\Big\}. \quad \text{(B.24)}$$

Note that the only difference between $E_-(R)$ and $E_+(R)$ is the constraint $I(P_X \times \tilde{Q}_{Y|X}) \gtreqless R$, and due to the continuity of the objective function, we have included the points $\{I(P_X \times \tilde{Q}_{Y|X}) = R\}$ in both problems. Now, since the KL divergence is also a convex function of $Q_{Y|X}$, it can be seen that the objective function is jointly convex in $\{Q_{Y|X}, \tilde{Q}_{Y|X}\}$ for both optimization problems. Since $\alpha(Q_{XY})$ is a linear function of Q_{XY}, the set $\{Q_Y = \tilde{Q}_Y,\ \alpha(P_X \times \tilde{Q}_{Y|X}) \geq \alpha(P_X \times Q_{Y|X})\}$ is a convex set. Furthermore, the set $\{I(P_X \times \tilde{Q}_{Y|X}) \leq R\}$ is also a convex set, and thus so is its intersection with the previous set. Consequently, the minimization problem of $E_-(R)$ is a convex optimization problem [21] (of dimension $2|\mathcal{X}| \times (|\mathcal{Y}| - 1)$), which can be efficiently solved, *e.g.*, using software packages such as `CVX` [78]. In contrast, the minimization problem of $E_+(R)$ involves the set $\{I(P_X \times \tilde{Q}_{Y|X}) \geq R\}$, which is *not* a convex set.

We thus proceed as follows. First, let us solve $E_+(R)$ for $R = 0$. In this case, the constraint $I(P_X \times Q_{Y|X}) \geq R$ is idle, and so

$$E_+(0) = \min_{Q_{Y|X},\tilde{Q}_{Y|X}:\, \alpha(P_X\times\tilde{Q}_{Y|X})\geq\alpha(P_X\times Q_{Y|X})} D(Q_{Y|X}||W|P_X)+I(P_X\times\tilde{Q}_{Y|X}). \tag{B.25}$$

This is a convex optimization problem, which can be efficiently solved. Let us denote the solution of this problem as $(Q^{(0)}_{Y|X},\tilde{Q}^{(0)}_{Y|X})$. Now, as long as $R \leq R_{\mathrm{cr}} \triangleq I(\tilde{Q}^{(0)}_{Y|X})$, then the objective function in $E_+(R)$ is minimized by the unconstrained solution $(Q^{(0)}_{Y|X},\tilde{Q}^{(0)}_{Y|X})$, even if the constraint $I(P_X \times Q_{Y|X}) \geq R$ is imposed. For these rates it thus holds that $E_+(R) = E_+(0) - R$. Now, if $R \geq R_{\mathrm{cr}}$ then the unconstrained solution $(Q^{(0)}_{Y|X},\tilde{Q}^{(0)}_{Y|X})$ does not solve $E_+(R)$, and so the solution must be obtained on the boundary $\{I(P_X \times \tilde{Q}_{Y|X}) = R\}$. However, for such rates

$$\begin{aligned}
E_+(R) &\\
&= \min_{Q_{Y|X},\tilde{Q}_{Y|X}:\, I(P_X\times\tilde{Q}_{Y|X})=R} D(Q_{Y|X}||W|P_X) + I(P_X \times \tilde{Q}_{Y|X}) - R \qquad \text{(B.26)}\\
&= \min_{Q_{Y|X},\tilde{Q}_{Y|X}:\, I(P_X\times\tilde{Q}_{Y|X})=R} D(Q_{Y|X}||W|P_X) \qquad \text{(B.27)}\\
&\geq \min_{Q_{Y|X},\tilde{Q}_{Y|X}:\, I(P_X\times\tilde{Q}_{Y|X})\leq R} D(Q_{Y|X}||W|P_X) \qquad \text{(B.28)}\\
&= E_-(R), \qquad \text{(B.29)}
\end{aligned}$$

where all the above minimization operators are under the constraint $\alpha(P_X \times \tilde{Q}_{Y|X}) \geq \alpha(P_X \times Q_{Y|X})$, and the inequality holds since the feasible set is larger for $E_-(R)$. Consequently, for rates $R \geq R_{\mathrm{cr}}$, the exponent is given by $\min\{E_-(R), E_+(R)\} = E_-(R)$.

To conclude, despite the fact that the minimization problem of $E_+(R)$ is not a convex optimization problem, the exponent can be computed for all rates by only solving convex optimization problems. To summarize, this is done by the following procedure: (1) Solve the optimization problem for $E_+(0)$, and compute the critical rate R_{cr}. (2) Solve the optimization problem $E_-(R)$ for any $R > R_{\mathrm{cr}}$. The exponent is

$$\begin{cases} E_+(0) - R, & 0 \le R \le R_{\text{cr}} \\ E_-(R), & R > R_{\text{cr}} \end{cases}. \tag{B.30}$$

Note that this method requires solving two convex optimization problems at most for each rate, and the first one for finding $E_+(0)$ one is common to all rates.

For additional computational algorithms, see, for example, [64, Section V] for the computation of the exponent of the interference channel, [216, Appendix A] for the exponents of joint detection and decoding, and [215, Section VI] for exponents of distributed hypothesis testing.

C

The Derivation of the Expurgated Exponent

In this appendix, we outline the expurgation argument that follows the TCEM method. The proof follows [128, Appendix]. Let us focus on a specific codeword index m. We showed in Section 4.3 that, effectively, $\overline{N}_m(Q_{X\tilde{X}}) \sim \text{Binomial}(e^{nR}, e^{-nI(Q_{X\tilde{X}})})$. Thus, we separate between *typically populated* joint types ($I(Q_{X\tilde{X}}) \leq R$) and *typically empty* joint types ($I(Q_{X\tilde{X}}) > R$). First, for the populated types, for any $\epsilon > 0$, it holds by (4.66) that

$$\Pr\left[\overline{N}_m(Q_{X\tilde{X}}) \geq e^{n(R-I(Q_{X\tilde{X}})+\epsilon)}\right] \doteq e^{-n\infty}. \tag{C.1}$$

Taking the union over an exponentially number of codewords e^{nR} and a polynomial number of joint types, it follows from the union bound that

$$\mathcal{F} \triangleq \bigcup_{m=1}^{e^{nR}} \bigcup_{Q_{X\tilde{X}}:\, Q_X = Q_{\tilde{X}} = P_X,\, I(Q_{X\tilde{X}}) \geq R} \left\{\overline{N}_m(Q_{X\tilde{X}}) \geq e^{n(R-I(Q_{X\tilde{X}})+\epsilon)}\right\} \tag{C.2}$$

satisfies $\Pr[\mathcal{F}] \doteq e^{-n\infty}$. Since by (4.67) the lower tail also similarly decays double-exponentially, for the sake of exponent analysis, the TCE are *effectively* deterministic, for all codewords in the codebook and all joint types with $I(Q_{X\tilde{X}}) \leq R$, and is given by

$$\overline{N}_m(Q_{X\tilde{X}}) \doteq e^{n[R-I(Q_{X\tilde{X}})]}. \tag{C.3}$$

Second, for the empty types for which $I(Q_{X\tilde{X}}) > R$, it holds by (4.66) that

$$\Pr\left[\overline{N}_m(Q_{X\tilde{X}}) \geq 1\right] \doteq e^{-n[I(Q_{X\tilde{X}})-R]}, \tag{C.4}$$

which is exponentially small. Thus, we do not expect to observe other codewords $\tilde{m} \neq m$ which have joint type $Q_{X\tilde{X}}$ with $\boldsymbol{X}_m$. Indeed, the event

$$\mathcal{E}_m \triangleq \left\{ \bigcup_{Q_{X\tilde{X}}:\, Q_X = Q_{\tilde{X}} = P_X,\, I(Q_{X\tilde{X}})>R} \left\{\overline{N}_m(Q_{X\tilde{X}}) \geq 1\right\} \right\} \tag{C.5}$$

is the event that the mth codeword is a a-typical neighboring codeword, in the sense that there exists a $Q_{X\tilde{X}}$ with $I(Q_{X\tilde{X}}) > R$ and at least one neighboring codeword $\boldsymbol{X}_{\tilde{m}}$ so that $\hat{Q}_{\boldsymbol{X}_m\boldsymbol{X}_{\tilde{m}}} = Q_{X\tilde{X}}$. By the union bound, since the number of joint types increases polynomially with n, $p_n \triangleq \Pr[\mathcal{E}_m] \doteq e^{-n(I(Q_{X\tilde{X}})-R)}$. Thus, on the average, we expect that $p_n e^{nR}$ codewords will have such a-typical neighboring codewords. So, the event

$$\mathcal{E}^* \triangleq \left\{ \frac{1}{e^{nR}} \sum_{m=1}^{e^{nR}} \mathbb{1}\{\mathcal{E}_m\} \geq 2p_n \right\}, \tag{C.6}$$

in which more than $2p_n e^{nR}$ have such a-typical neighboring codeword has low probability. Indeed, Markov's inequality, which does not require independence of the events $\{\mathcal{E}_m\}$, implies that $\Pr[\mathcal{E}^*] \leq \frac{1}{2}$. Hence, with probability larger than $1/2 - \Pr[\mathcal{F}] \geq 1/3$, both $\mathcal{F}^c$ and $[\mathcal{E}^*]^c$ hold. We thus may choose a codebook $\mathcal{C}_n$ that belongs to the event $\mathcal{F}^c \cap [\mathcal{E}^*]^c$. The number of codewords in this codebook for which $\mathbb{1}\{\mathcal{E}_m\} = 1$ is less than $3p_n e^{nR}$. Thus, we can *expurgate* those codewords from the codebook, and obtain a new codebook $\mathcal{C}_n^*$ which satisfies: (1) Its size is larger than $|\mathcal{C}_n^*| \geq e^{nR}(1-3p_n) \doteq e^{nR}$. (2) Its TCEs $\overline{N}_m^*(Q_{X\tilde{X}})$ are only smaller than those of the original codebook, and specifically, $\overline{N}_m^*(Q_{X\tilde{X}}) = 0$ for all $Q_{X\tilde{X}}$ with $I(Q_{X\tilde{X}}) > R$. (3) $\overline{N}_m^*(Q_{X\tilde{X}}) \leq e^{n(R-I(Q_{X\tilde{X}})+\epsilon)}$ for all $Q_{X\tilde{X}}$ with $I(Q_{X\tilde{X}}) \leq R$.

For such a codebook, and after taking $\epsilon \downarrow 0$, the error probability bound in (4.38) is given by

$$P_{\mathsf{e}} \leq \exp\left[-n \cdot E_{\text{ex}}(R, P_X)\right], \tag{C.7}$$

where $E_{\text{ex}}(R, P_X)$ is as defined in (4.14).

Compared to the TCEM, the properties of codebook $\mathcal{C}_n^*$ traditionally follow from the *packing lemma* [41, Exercise 10.2], [42] (which is somewhat similar) or from a *graph decomposition lemma* [40, Corollary to Lemma 2]. In the latter case, equipped with the existence of such a codebook, [40] derived a bound for decoders with general score $\alpha(\cdot)$, and when $\alpha(\cdot)$ is set to be the ML decoder, then this exponent is shown to be at least as high as both the random-coding error exponent and the expurgated exponent.

D

Proofs for Section 4.3

Before proving Theorems 4.1, 4.2 and 4.3, we recall the following Chernoff tail bounds of a binomial RV $X \sim \text{Binomial}(m, p)$. If $r > p$ then $rm > \mathbb{E}[X] = pm$ and so the probability of the upper tail is

$$e^{-m \cdot D(r||p) - o(m)} \leq \Pr[X > rm] \leq e^{-m \cdot D(r||p)}, \tag{D.1}$$

where $D(r||p) \triangleq r \ln \frac{r}{p} + (1-r) \ln \frac{(1-r)}{(1-p)}$ is the binary KL divergence. If $r < p$ then this probability $\Pr[X > rm] \geq \Pr[X > \lfloor \mathbb{E}[X] \rfloor] \geq 1/2$, and the so the exponent is zero. Similarly, if $r < p$ then the probability of the lower tail is

$$e^{-m \cdot D(r||p) - o(m)} \leq \Pr[X < rm] \leq e^{-m \cdot D(r||p)}, \tag{D.2}$$

and if $r > p$ then the exponent is zero.

We will also need the following simple lemma regarding the KL divergence.

Lemma D.1. *Let $\{a_n, b_n\}$ be sequences in $(0, 1)$ such that $a_n = o(1)$ and $b_n = o(1)$. Then,*

$$D(a_n||b_n) \sim \begin{cases} b_n & \frac{a_n}{b_n} = o(1) \\ a_n \ln \frac{a_n}{b_n}, & \frac{a_n}{b_n} = \omega(1) \end{cases}, \tag{D.3}$$

where for a sequence $\{c_n\}$, *the notation* $c_n = o(1)$ *means that* $\lim_{n\to\infty} c_n = 0$ *and the notation* $c_n = \omega(1)$ *means that* $\lim_{n\to\infty} c_n = \infty$.

Proof. We use the expansion $\ln(1+x) = x + \Theta(x^2)$ throughout. If $\frac{a_n}{b_n} = o(1)$ then it holds that

$$(1-a_n)\ln\left[\frac{1-a_n}{1-b_n}\right]$$
$$= (1-a_n)\ln(1-a_n) - (1-a_n)\ln(1-b_n) \tag{D.4}$$
$$= -a_n(1-a_n) + \Theta(a_n^2) + b_n(1-a_n) + \Theta(b_n^2) \tag{D.5}$$
$$= (b_n - a_n)(1-a_n) + \Theta(b_n^2) \tag{D.6}$$
$$= b_n \cdot \left[\left(1 - \frac{a_n}{b_n}\right) - a_n(1-a_n) + \Theta(b_n^2)\right] \tag{D.7}$$
$$\sim b_n, \tag{D.8}$$

and so for all n large enough

$$\left|a_n \ln \frac{a_n}{b_n}\right| = a_n \ln \frac{b_n}{a_n} = -b_n \cdot \frac{a_n}{b_n} \ln \frac{a_n}{b_n} = -o(b_n) \tag{D.9}$$

since $\lim_{t\downarrow 0} t \ln t = 0$. This is negligible compared to the first term.

If $\frac{a_n}{b_n} = \omega(1)$ then

$$\left|(1-a_n)\ln\left(\frac{1-a_n}{1-b_n}\right)\right|$$
$$= |(1-a_n)\ln(1-a_n) - (1-a_n)\ln(1-b_n)| \tag{D.10}$$
$$= \left|(1-a_n)\left[-a_n + \Theta(a_n^2) + b_n + \Theta(b_n^2)\right]\right| \tag{D.11}$$
$$= \Theta(a_n), \tag{D.12}$$

which is negligible compared to $a_n \ln \frac{a_n}{b_n} = \omega(a_n)$. □

We are now ready to prove Theorem 4.1, which provides exact exponents of the tail probabilities of the TCE N.

Proof of Theorem 4.1. In the case of a TCE, we are dealing with both an exponential number of trials and an exponentially decaying success probability, and so we consider the events $\{N > e^{n\lambda}\}$ and $\{N < e^{n\lambda}\}$

for some $\lambda \in \mathbb{R}$. Throughout, we will use the asymptotic expansion of the binary KL divergence in Lemma D.1.

We distinguish between two cases:

1. If $A > B$ then the mean value $\mathbb{E}[N] = e^{n(A-B)}$ is exponentially large. For the upper tail, we assume $\lambda > A - B$, for which
$$\Pr\left[N > e^{n\lambda}\right] \leq \exp\left[-e^{nA} \cdot D(e^{-n(A-\lambda)}||e^{-nB})\right]. \tag{D.13}$$
Since $A - B < \lambda$ then $e^{-n(A-\lambda)}/e^{-nB} = \omega(1)$ and the exponent is
$$e^{nA} \cdot D(e^{-n(A-\lambda)}||e^{-nB}) \sim e^{nA}e^{-n(A-\lambda)} \ln \frac{e^{-n(A-\lambda)}}{e^{-nB}} \tag{D.14}$$
$$= n(\lambda - (A-B))e^{n\lambda}. \tag{D.15}$$
Thus, the right-tail probability decays double-exponentially. Similarly, for the lower tail, we assume $\lambda < A - B$, for which
$$\Pr\left[N < e^{n\lambda}\right] \leq \exp\left[-e^{nA} \cdot D(e^{-n(A-\lambda)}||e^{-nB})\right]. \tag{D.16}$$
Since $A - B > \lambda$ then $e^{-n(A-\lambda)}/e^{-nB} = o(1)$ and the exponent is
$$e^{nA} \cdot D(e^{-n(A-\lambda)}||e^{-nB}) \sim e^{n(A-B)}. \tag{D.17}$$
Thus, the lower-tail probability also decays double-exponentially.

2. If $B > A$ then the mean value $\mathbb{E}[N] = e^{-n(B-A)} \leq 1$ is exponentially small. For the upper tail, we set $\lambda > 0 > A - B$ and obtain a double-exponentially decay, exactly as in the previous case. Next, as N is integer, for $\lambda \leq 0$, Markov's inequality implies that
$$\Pr\left[N > e^{n\lambda}\right] = \Pr\left[N \geq 1\right] \leq \mathbb{E}[N] = \exp\left[-n(B-A)\right]. \tag{D.18}$$
On the other hand,
$$\Pr\left[N > e^{n\lambda}\right] \geq \Pr\left[N = 1\right] = \binom{e^{nA}}{1} \cdot e^{-nB} \cdot (1 - e^{-nB})^{e^{nA}-1} \tag{D.19}$$
$$= e^{-n(B-A)} \cdot (1 - e^{-nB})^{e^{nA}-1} \tag{D.20}$$
$$\sim \exp\left[-n(B-A)\right], \tag{D.21}$$
which shows that Markov's inequality is exponentially tight in this case, and hence $\Pr[N > e^{n\lambda}] \doteq e^{-n(B-A)}$. The variable N has no lower tail since the above implies that $\Pr[N = 0] \geq 1 - e^{-n(B-A)}$.

Combining the two cases leads to the claimed result. □

We next prove Theorem 4.2, which states the exponent of $\mathbb{E}[N^s]$.

Proof of Theorem 4.2. We separate again between two cases, depending on the sign of $A - B$.

1. If $A > B$ then we know that any exponential deviation from the mean leads to a double-exponentially decay. Hence, for any $\lambda > A - B$

$$\mathbb{E}\left[N^s\right] = \Pr[N \le e^{n\lambda}] \cdot \mathbb{E}\left[N^s | N \le e^{n\lambda}\right] + \Pr[N > e^{n\lambda}] \cdot \mathbb{E}\left[N^s | N \ge e^{n\lambda}\right] \tag{D.22}$$

$$\dot{\le} e^{n\lambda s} + e^{-n\infty} \cdot e^{nsA} \tag{D.23}$$

$$\doteq e^{n\lambda s}, \tag{D.24}$$

where we have used the fact that $N \le e^{nA}$ with probability 1, and write $e^{-n\infty}$ for a probability that decays super-exponentially. Taking the limit $\lambda \downarrow A - B$ shows that

$$\mathbb{E}\left[N^s\right] \dot{\le} e^{n(A-B)s}. \tag{D.25}$$

A matching lower bound can be derived in an analogous way: For any $\lambda < A - B$

$$\mathbb{E}\left[N^s\right] = \Pr[N \ge e^{n\lambda}] \cdot \mathbb{E}\left[N^s | N \ge e^{n\lambda}\right] + \Pr[N < e^{n\lambda}] \cdot \mathbb{E}\left[N^s | N < e^{n\lambda}\right] \tag{D.26}$$

$$\ge \left[1 - \Pr[N < e^{n\lambda}]\right] \cdot e^{n\lambda s} \tag{D.27}$$

$$\sim e^{n\lambda s}, \tag{D.28}$$

after taking the limit $\lambda \uparrow A - B$. Hence,

$$\mathbb{E}\left[N^s\right] \doteq e^{n(A-B)s}. \tag{D.29}$$

2. If $A < B$ then we take $\lambda > 0$ to obtain

$$\mathbb{E}\left[N^s\right] = \Pr[1 \le N \le e^{n\lambda}] \cdot \mathbb{E}\left[N^s | 1 \le N \le e^{n\lambda}\right]$$

$$+ \Pr[N > e^{n\lambda}] \cdot \mathbb{E}\left[N^s | N \geq e^{n\lambda}\right] \quad \text{(D.30)}$$

$$\dot{\leq} \Pr[N \geq 1] \cdot e^{n\lambda} + e^{-n\infty} \cdot e^{nsA} \quad \text{(D.31)}$$

$$\dot{\leq} e^{-n(B-A)} \cdot e^{n\lambda}. \quad \text{(D.32)}$$

Taking the limit $\lambda \downarrow 0$ shows that

$$\mathbb{E}\left[N^s\right] \dot{\leq} e^{-n(B-A)}. \quad \text{(D.33)}$$

A lower bound is obtained by

$$\mathbb{E}\left[N^s\right] \geq \Pr[N = 1] \cdot 1^s \geq [1 + o(1)] \cdot e^{-n(B-A)}, \quad \text{(D.34)}$$

which shows that the upper bound is tight.

Combining the two cases leads to the claimed result. □

We finally prove Theorem 4.3, which states that the probability of an intersection of lower tail events of a set of TCEs is exponentially equivalent to either 0 or 1.

Proof of Theorem 4.3. If there is a $j^* \in [k_n]$ so that $B_{j^*} < A_{j^*}$ and $\lambda < A_{j^*} - B_{j^*}$ then $\Pr[N_{j^*} < e^{n\lambda}] \doteq e^{-n\infty}$. So,

$$\Pr\left[\bigcap_{j=1}^{k_n} \left\{N_j < e^{n\lambda}\right\}\right] \leq \min_{1 \leq j \leq k_n} \Pr\left[N_j < e^{n\lambda}\right] \doteq e^{-n\infty}. \quad \text{(D.35)}$$

Otherwise, if all $j = 1, \ldots, k_n$ it holds that either $B_j > A_j$ or $\lambda > A_j - B_j$ then (4.66) implies that $\Pr[N_j > e^{n\lambda}] \dot{\leq} e^{-n\infty}$ for all $j = 1, \ldots, k_n$. Thus, from the union bound, as $n \to \infty$

$$\Pr\left[\bigcap_{j=1}^{k_n} \left\{N_j \leq e^{n\lambda}\right\}\right] = 1 - \Pr\left[\bigcup_{j=1}^{k_n} \left\{N_j > e^{n\lambda}\right\}\right] \quad \text{(D.36)}$$

$$\geq 1 - \sum_{j=1}^{k_n} \Pr\left[N_j > e^{n\lambda}\right] \quad \text{(D.37)}$$

$$\geq 1 - k_n \cdot \max_{1 \leq j \leq k_n} \Pr\left[N_j > e^{n\lambda}\right] \quad \text{(D.38)}$$

$$\geq 1 - k_n \cdot e^{-\min_{1 \leq j \leq k_n} E_j} \quad \text{(D.39)}$$

$$\to 1. \quad \text{(D.40)}$$

Combining (D.35) and (D.40) leads to the stated claim. □

References

[1] R. Ahlswede and I. Csiszár, “Hypothesis testing with communication constraints,” *IEEE Trans. Inf. Theory*, vol. 32, no. 4, 1986, pp. 533–542.

[2] R. Ahlswede and G. Dueck, “Good codes can be produced by a few permutations,” *IEEE Trans. Inf. Theory*, vol. 28, no. 3, 1982, pp. 430–443.

[3] M. A. Ali, H. Budak, and Z. Zhang, “A new extension of quantum Simpson’s and quantum Newton’s inequalities for quantum differentiable convex functions,” *Mathematical Methods in the Applied Sciences*, 2021. DOI: 10.1002/mma.7889.

[4] S. M. Ali and S. D. Silvey, “A general class of coefficients of divergence of one distribution from another,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 28, no. 1, 1966, pp. 131–142.

[5] Y. Altuğ and A. B. Wagner, “Moderate deviations in channel coding,” *IEEE Trans. Inf. Theory*, vol. 60, no. 8, 2014, pp. 4417–4426.

[6] Y. Altuğ and A. B. Wagner, “Refinement of the random coding bound,” *IEEE Trans. Inf. Theory*, vol. 60, no. 10, 2014, pp. 6005–6023.

[7] D. Anade, J.-M. Gorce, P. Mary, and S. M. Perlaza, "An upper bound on the error induced by saddlepoint approximations – applications to information theory," *Entropy*, vol. 22, no. 6, 2020, p. 690. DOI: 10.3390/e22060690.

[8] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, 2009, pp. 3051–3073.

[9] E. Arikan and N. Merhav, "Guessing subject to distortion," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, 1998, pp. 1041–1056.

[10] S. Arimoto, "On the converse to the coding theorem for discrete memoryless channels (corresp.)," *IEEE Trans. Inf. Theory*, vol. 19, no. 3, 1973, pp. 357–359.

[11] R. Averbuch and N. Merhav, "Exact random coding exponents and universal decoders for the asymmetric broadcast channel," *IEEE Trans. Inf. Theory*, vol. 64, no. 7, 2018, pp. 5070–5086.

[12] R. Averbuch, N. Weinberger, and N. Merhav, "Expurgated bounds for the asymmetric broadcast channel," *IEEE Trans. Inf. Theory*, vol. 65, no. 6, 2019, pp. 3412–3435.

[13] R. R. Bahadur and R. R. Rao, "On deviations of the sample mean," *Ann. Math. Statist.*, vol. 31, no. 4, 1960, pp. 1015–1027.

[14] A. Barg and G. D. Forney, "Random codes: Minimum distances and error exponents," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, 2002, pp. 2568–2573.

[15] L. A. Bassalygo, S. I. Gel'fand, and M. S. Pinsker, "Simple methods for deriving lower bounds in the theory of codes," *Probl. Pered. Inform*, vol. 27, no. 4, 1991, pp. 3–8.

[16] P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Trans. Inf. Theory*, vol. 19, no. 2, 1973, pp. 197–207.

[17] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes. 1," in *Proceedings of ICC'93-IEEE International Conference on Communications*, IEEE, vol. 2, pp. 1064–1070, 1993.

[18] P. Billingsley, "Statistical methods in Markov chains," *Ann. Math. Statist.*, vol. 32, 1961, pp. 12–40.

[19] R. E. Blahut, "Hypothesis testing and information theory," *IEEE Trans. Inf. Theory*, vol. 20, no. 4, 1974, pp. 405–417.

[20] S. Boucheron, G. Lugosi, and P. Massart, *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, 2013.

[21] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[22] L. B. Boza, "Asymptotically optimal tests for finite Markov chains," *Ann. Math. Stat.*, vol. 42, 1971, pp. 1992–2007.

[23] N. G. de Bruijn, *Asymptotic Methods in Analysis*. Dover Publications, 1981.

[24] H. Budak, M. A. Ali, and M. Tarhanaci, "Some new quantum Hermite–Hadamard like inequalities for coordinated convex functions," *Journal of Optimization Theory and Applications*, vol. 186, no. 3, 2020, pp. 899–910.

[25] I. Budimir, S. S. Dragomir, and J. Pečari, "Further reverse results for Jensen's discrete inequality and applications in information theory," *Journal of Inequalities in Pure and Applied Mathematics*, vol. 2, 2001, pp. 1–14.

[26] M. V. Burnashev, "Data transmission over a discrete channel with feedback. random transmission time," *Problemy peredachi informatsii*, vol. 12, no. 4, 1976, pp. 10–30.

[27] D. Cao and V. Y. F. Tan, "Exact error and erasure exponents for the asymmetric broadcast channel," *IEEE Trans. Inf. Theory*, vol. 66, no. 2, 2019, pp. 865–885.

[28] V. Chandar, A. Tchamkerten, and D. Tse, "Asynchronous capacity per unit cost," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, 2012, pp. 1213–1226.

[29] V. Chandar, A. Tchamkerten, and G. W. Wornell, "Optimal sequential frame synchronization," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, 2008, pp. 3725–3728.

[30] J. Chen, D.-k. He, A. Jagmohan, and L. A. Lastras-Montaño, "On universal variable-rate Slepian-Wolf coding," in *Proc. of IEEE International Conference on Communications*, IEEE, pp. 1426–1430, 2008.

[31] S.-Y. Chung, G. D. Forney, T. J. Richardson, and R. Urbanke, "On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit," *IEEE Communications letters*, vol. 5, no. 2, 2001, pp. 58–60.

[32] M. Costa, "Writing on dirty paper (corresp.)," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, 1983, pp. 439–441.

[33] T. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, 1972, pp. 2–14.

[34] T. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, 1979, pp. 572–584.

[35] T. Cover and C. Leung, "An achievable rate region for the multiple-access channel with feedback," *IEEE Trans. Inf. Theory*, vol. 27, no. 3, 1981, pp. 292–298.

[36] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, 2006.

[37] I. Csiszár, "Eine informationstheoretische ungleichung und ihre anwendung auf den beweis der ergodizität von markoffschen ketten," *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei*, vol. 8, no. 1-2, 1963, pp. 85–108.

[38] I. Csiszár, "The method of types," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, 1998, pp. 2505–2523.

[39] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, 1978, pp. 339–348.

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

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

[42] I. Csiszár, J. Körner, and K. Marton, "A new look at the error exponent of discrete memoryless channels," in *Proc. of International Symposium on Information Theory*, 107 (abstract), 1977.

[43] A. G. D'yachkov, "Bounds on the average error probability for a code ensemble with fixed composition," *Problemy Peredachi Informatsii*, vol. 16, no. 4, 1980, pp. 3–8.

[44] G. Dasarathy and S. C. Draper, "On reliability of content identification from databases based on noisy queries," in *2011 IEEE International Symposium on Information Theory Proceedings*, IEEE, pp. 1066–1070, 2011.

[45] G. Dasarathy and S. C. Draper, "Upper and lower bounds on the reliability of content identification," in *23th International Zurich Seminar on Communications (IZS 2014)*, ETH-Zürich, 2014.

[46] L. D. Davisson, "Universal noiseless coding," *IEEE Trans. Inf. Theory*, vol. 29, no. 6, 1973, pp. 783–795.

[47] L. D. Davisson, G. Longo, and A. Sgarro, "The error exponent for noiseless encoding of finite ergodic Markov sources," *IEEE Trans. Inf. Theory*, vol. 27, no. 4, 1981, pp. 431–438.

[48] D. De Caen, "A lower bound on the probability of a union," *Discrete mathematics*, vol. 169, no. 1-3, 1997, pp. 217–220.

[49] A. Dembo and O. Zeitouni, *Large Deviations and Applications*. Jones and Bartlett Publishers, 1993.

[50] B. Derrida, "Random-energy model: Limit of a family of disordered models," *Physical Review Letters*, vol. 45, no. 2, 1980, p. 79.

[51] B. Derrida, "The random energy model," CEA Centre d'Etudes Nucleaires de Saclay, Tech. Rep., 1980.

[52] B. Derrida, "Random-energy model: An exactly solvable model of disordered systems," *Physical Review B*, vol. 24, no. 5, 1981, p. 2613.

[53] M. D. Donsker and S. R. S. Varadhan, "Asymptotic evaluation of certain Markov process expectations for large time. iv," *Communications on pure and applied mathematics*, vol. 36, no. 2, 1983, pp. 183–212.

[54] S. S. Dragomir, "Some reverses of the Jensen inequality for functions of selfadjoint operators in Hilbert spaces," *Journal of Inequalities and Applications*, 2010. DOI: https://doi.org/10.1155/2010/496821.

[55] S. S. Dragomir, “Some reverses of the Jensen inequality with applications,” *Bulletin of the Australian Mathematical Society*, vol. 87, 2013, pp. 177–194.

[56] M. Drmota and W. Szpankowski, *Analytic Information Theory: From Compression to Learning*. Cambridge University Press, 2023.

[57] G. Dueck and J. Körner, “Reliability function of a discrete memoryless channel at rates above capacity (corresp.),” *IEEE Trans. Inf. Theory*, vol. 25, no. 1, 1979, pp. 82–85.

[58] P. Dupuis and R. S. Ellis, *A Weak Convergence Approach to the Theory of Large Deviations*. John Wiley & Sons, 1997.

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

[60] P. Elias, “Coding for noisy channels,” in *IRE Conv. Rec.*, vol. 3, pp. 37–46, 1955.

[61] P. Elias, “List decoding for noisy channels,” in *IRE WESCON Conf. Rec.*, vol. 2, pp. 94–104, 1957.

[62] T. Erseghe, “Coding in the finite-blocklength regime: Bounds based on Laplace integrals and their asymptotic approximations,” *IEEE Trans. Inf. Theory*, vol. 62, no. 12, 2016, pp. 6854–6883.

[63] S. E. Esipov and T. J. Newman, “Interface growth and Burgers turbulence: The problem of random initial conditions,” *Physical Review E*, vol. 48, no. 2, 1993, pp. 1046–1050.

[64] R. H. Etkin, N. Merhav, and E. Ordentlich, “Error exponents of optimum decoding for the interference channel,” *IEEE Trans. Inf. Theory*, vol. 56, no. 1, 2009, pp. 40–56.

[65] R. M. Fano, *Transmission of Information: A Statistical Theory of Communications*. M.I.T. Press, 1961.

[66] M. Feder and N. Merhav, “Universal composite hypothesis testing: A competitive minimax approach,” *IEEE Trans. Inf. Theory*, vol. 48, no. 6, 2002, pp. 1504–1517.

[67] J. Font-Segura, G. Vázquez-Vilar, A. Martinez, A. Guillén i Fàbregas, and A. Lancho, “Saddlepoint approximations of lower and upper bounds to the error probability in channel coding,” in *Proc. 2018 52nd Annual Conference on Information Sciences and Systems (CISS 2018)*, Princeton, NJ, USA, 2018.

[68] G. D. Forney, "Exponential error bounds for erasure, list, and decision feedback schemes," *IEEE Trans. Inf. Theory*, vol. 14, no. 2, 1968, pp. 206–220.

[69] R. G. Gallager, "Low-density parity-check codes," *IRE Transactions on information theory*, vol. 8, no. 1, 1962, pp. 21–28.

[70] R. G. Gallager, "A simple derivation of the coding theorem and some applications," *IEEE Trans. Inf. Theory*, vol. 11, no. 1, 1965, pp. 3–18.

[71] R. G. Gallager, *Information Theory and Reliable Communication*, vol. 588. Springer, 1968.

[72] R. G. Gallager, "The random coding bound is tight for the average code (corresp.)," *IEEE Trans. Inf. Theory*, vol. 19, no. 2, 1973, pp. 244–246.

[73] R. G. Gallager, "Capacity and coding for degraded broadcast channels," *Problemy Peredachi Informatsii*, vol. 10, no. 3, 1974, pp. 3–14.

[74] R. G. Gallager, "Source coding with side information and universal coding," M.I.T., LIDS- P-937, 1976.

[75] S. I. Gel'fand and M. S. Pinsker, "Coding for channels with random parameters," *Probl. Contr. Inform. Theory*, vol. 9, no. 1, 1980, pp. 19–31.

[76] S. Ginzach, N. Merhav, and I. Sason, "Random-coding error exponent of variable-length codes with a single-bit noiseless feedback," in *2017 IEEE Information Theory Workshop (ITW)*, IEEE, pp. 584–588, 2017.

[77] V. D. Goppa, "Nonprobabilitistic mutual information without memory," *Problems of Control and Inform., Theory*, vol. 4, 1975, pp. 97–102.

[78] M. Grant, S. P. Boyd, and Y. Ye, "CVX users' guide," 2009. URL: http://www.%20stanford.%20edu/boyd/software.%20html.

[79] T. S. Han and S. Amari, "Statistical inference under multiterminal data compression," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, 1998, pp. 2300–2324.

[80] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, 1981, pp. 49–60.

[81] T. S. Han and K. Kobayashi, “Exponential-type error probabilities for multiterminal hypothesis testing,” *IEEE Trans. Inf. Theory*, vol. 35, no. 1, 2006, pp. 2–14.

[82] E. A. Haroutunian, “Bounds for the exponent of the probability of error for a semicontinuous memoryless channel,” *Problemy Peredachi Informatsii*, vol. 4, no. 4, 1968, pp. 37–48.

[83] M. Hayashi, “General nonasymptotic and asymptotic formulas in channel resolvability and identification capacity and their application to the wiretap channel,” *IEEE Trans. Inf. Theory*, vol. 52, no. 4, 2006, pp. 1562–1575.

[84] M. Hayashi, “Exponential decreasing rate of leaked information in universal random privacy amplification,” *IEEE Trans. Inf. Theory*, vol. 57, no. 6, 2011, pp. 3989–4001.

[85] M. Hayashi and V. Y. F. Tan, “Asymmetric evaluations of erasure and undetected error probabilities,” *IEEE Trans. Inf. Theory*, vol. 61, no. 12, 2015, pp. 6560–6577.

[86] E. Hille, *Analytic Function Theory*, vol. 2. American Mathematical Soc., 1959.

[87] J. Honda, “Exact asymptotics of random coding error probability for general memoryless channels,” in *Proc. 2018 IEEE International Symposium on Information Theory (ISIT 2018)*, pp. 1844–1848, Vail, CO, U.S.A, 2018.

[88] W. Huleihel and N. Merhav, “Universal decoding for Gaussian intersymbol interference channels,” *IEEE Trans. Inf. Theory*, vol. 61, no. 4, 2015, pp. 1606–1618.

[89] W. Huleihel and N. Merhav, “Random coding error exponents for the two-user interference channel,” *IEEE Trans. Inf. Theory*, vol. 63, no. 2, 2016, pp. 1019–1042.

[90] W. Huleihel, S. Salamatian, N. Merhav, and M. Médard, “Gaussian intersymbol interference channels with mismatch,” *IEEE Trans. Inf. Theory*, vol. 65, no. 7, 2019, pp. 4499–4517.

[91] W. Huleihel, N. Weinberger, and N. Merhav, “Erasure/list random coding error exponents are not universally achievable,” *IEEE Trans. Inf. Theory*, vol. 62, no. 10, 2016, pp. 5403–5421.

[92] P. A. Humblet, "Generalization of Huffman coding to minimize the probability of buffer overflow," *IEEE Trans. Inf. Theory*, vol. 27, 1981, pp. 230–232.

[93] T. Ignatenko and F. M. J. Willems, "Biometric security from an information-theoretical perspective," *Foundations and Trends® in Communications and Information Theory*, vol. 7, no. 2–3, 2012, pp. 135–316.

[94] A. Ingber, T. Courtade, and T. Weissman, "Compression for quadratic similarity queries," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, 2015, pp. 2729–2747.

[95] S. Janson, "New versions of Suen's correlation inequality," *Random Structures and Algorithms*, vol. 13, no. 3-4, 1998, pp. 467–483.

[96] T. Jebara and A. Pentland, "On reversing Jensen's inequality," in *Proc. 13th International Conference on Neural Information Processing Systems (NIPS 2000)*, pp. 213–219, Denver, CO, U.S.A., 2000.

[97] F. Jelinek, "Buffer overflow in variable length coding of fixed rate sources," *IEEE Trans. Inf. Theory*, vol. 14, no. 3, 1968, pp. 490–501.

[98] F. Jelinek, "Evaluation of expurgated bound exponents," *IEEE Trans. Inf. Theory*, vol. 14, no. 3, 1968, pp. 501–505.

[99] F. Jelinek, *Probabilistic Information Theory.* McGraw-Hill, 1968.

[100] R. Johannesson and K. S. Zigangirov, *Fundamentals of Convolutional Coding.* John Wiley & Sons, 2015.

[101] Y. Kaspi and N. Merhav, "Error exponents for broadcast channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, 2010, pp. 101–123.

[102] B. G. Kelly and A. B. Wagner, "Reliability in source coding with side information," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, 2012, pp. 5086–5111.

[103] G. Keshet, Y. Steinberg, and N. Merhav, "Channel coding in the presence of side information," *Foundations and Trends® in Communications and Information Theory*, vol. 4, no. 6, 2008, pp. 445–586.

[104] S. Khan, M. A. Khan, and Y.-M. Chu, "Converses of Jensen inequality derived from the Green functions with applications in information theory," *Mathematical Methods in Applied Sciences*, vol. 43, 2020, pp. 2577–2587.

[105] S. Khan, M. A. Khan, and Y.-M. Chu, "New converses of Jensen inequality via Green functions with applications," *Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A, Matematicas*, vol. 114, no. 3, 2020.

[106] J. Körner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. 23, no. 1, 1977, pp. 60–64.

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

[108] R. E. Krichevsky and V. K. Trofimov, "The performance of universal encoding," *IEEE Trans. Inf. Theory*, vol. 27, no. 2, 1981, pp. 199–207.

[109] A. Lancho, J. Östman, G. Durisi, T. Koch, and G. Vázquez-Vilar, "Saddlepoint approximations for short-packet wireless communications," *IEEE Trans. on Wireless Communications*, vol. 19, no. 7, 2020, pp. 4831–4846.

[110] E. L. Lehmann, *Theory of Point Estimation.* New York: John Wiley & Sons, 1983.

[111] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Information theoretic security," *Foundations and Trends® in Communications and Information Theory*, vol. 5, no. 4–5, 2009, pp. 355–580.

[112] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, vol. 16. Elsevier, 1977.

[113] A. Martinez and A. Guillén i Fàbregas, "Random-coding bounds for threshold decoders: Error exponent and saddlepoint approximation," in *Proc. 2011 IEEE International Symposium on Information Theory (ISIT 2011)*, pp. 2905–2909, St. Petersburg, Russia, 2011.

[114] A. Martinez and A. Guillén i Fàbregas, "Saddlepoint approximation of random-coding bounds," in *The 2011 Information Theory and Applications Workshop (ITA 2011)*, La Jolla, CA, USA, 2011.

[115] N. Merhav, "On the estimation of the model order in exponential families," *IEEE Trans. Inf. Theory*, vol. 35, no. 5, 1989, pp. 1109–1114.

[116] N. Merhav, "Universal decoding for memoryless Gaussian channels with a deterministic interference," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, 1993, pp. 1261–1269.

[117] N. Merhav, "A large-deviations notion of perfect secrecy," *IEEE Trans. Inf. Theory*, vol. 49, no. 2, 2003, pp. 506–508.

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

[119] N. Merhav, "Relations between random coding exponents and the statistical physics of random codes," *IEEE Trans. Inf. Theory*, vol. 55, no. 1, 2008, pp. 83–92.

[120] N. Merhav, "Statistical physics and information theory," *Foundations and Trends in Communications and Information Theory*, vol. 6, no. 1-2, 2009.

[121] N. Merhav, "On optimum strategies for minimizing exponential moments of a given cost function," *Communications in Information and Systems*, vol. 11, no. 4, 2011, pp. 343–368.

[122] N. Merhav, "Subset-sum phase transitions and data compression," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2011, 2011, P09017.

[123] N. Merhav, "Another look at expurgated bounds and their statistical-mechanical interpretation," *arXiv preprint arXiv:1301.4117*, 2013.

[124] N. Merhav, "Universal decoding for arbitrary channels relative to a given class of decoding metrics," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, 2013, pp. 5566–5576.

[125] N. Merhav, "Erasure/list exponents for Slepian–Wolf decoding," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, 2014, pp. 4463–4471.

[126] N. Merhav, “Exact correct-decoding exponent of the wiretap channel decoder,” *IEEE Trans. Inf. Theory*, vol. 60, no. 12, 2014, pp. 7606–7615.

[127] N. Merhav, “Exact random coding error exponents of optimal bin index decoding,” *IEEE Trans. Inf. Theory*, vol. 60, no. 10, 2014, pp. 6024–6031.

[128] N. Merhav, “List decoding – random coding exponents and expurgated exponents,” *IEEE Trans. Inf. Theory*, vol. 60, no. 11, 2014, pp. 6749–6759.

[129] N. Merhav, “Statistical physics of random binning,” *IEEE Trans. Inf. Theory*, vol. 61, no. 5, 2015, pp. 2454–2464.

[130] N. Merhav, “Universal decoding for source–channel coding with side information,” *Communications in Information and Systems*, vol. 16, no. 1, 2016, pp. 17–58.

[131] N. Merhav, “Correction to "the generalized stochastic likelihood decoder: Random coding and expurgated bounds",” *IEEE Trans. Inf. Theory*, vol. 63, no. 10, 2017, pp. 6827–6829.

[132] N. Merhav, “Reliability of universal decoding based on vector-quantized codewords,” *IEEE Trans. Inf. Theory*, vol. 63, no. 5, 2017, pp. 2696–2709.

[133] N. Merhav, “The generalized stochastic likelihood decoder: Random coding and expurgated bounds,” *IEEE Trans. Inf. Theory*, vol. 63, no. 8, 2017, pp. 5039–5051.

[134] N. Merhav, “Ensemble performance of biometric authentication systems based on secret key generation,” *IEEE Trans. Inf. Theory*, vol. 65, no. 4, 2018, pp. 2477–2491.

[135] N. Merhav, “Error exponents of typical random codes,” *IEEE Trans. Inf. Theory*, vol. 64, no. 9, 2018, pp. 6223–6235.

[136] N. Merhav, “Lower bounds on exponential moments of the quadratic error in parameter estimation,” *IEEE Trans. Inf. Theory*, vol. 64, no. 12, 2018, pp. 7636–7648.

[137] N. Merhav, “A Lagrange-dual lower bound to the error exponent of the typical random code,” *IEEE Trans. Inf. Theory*, vol. 66, no. 6, 2019, pp. 3456–3464.

[138] N. Merhav, “Error exponents of typical random codes for the colored Gaussian channel,” *IEEE Trans. Inf. Theory*, vol. 65, no. 12, 2019, pp. 8164–8179.

[139] N. Merhav, “Error exponents of typical random trellis codes,” *IEEE Trans. Inf. Theory*, vol. 66, no. 4, 2019, pp. 2067–2077.

[140] N. Merhav, “Universal decoding for asynchronous Slepian–Wolf encoding,” *IEEE Trans. Inf. Theory*, vol. 67, no. 5, 2020, pp. 2652–2662.

[141] N. Merhav, “On more general distributions of random binning for Slepian–Wolf encoding,” *IEEE Trans. Inf. Theory*, vol. 68, no. 2, 2021, pp. 737–751.

[142] N. Merhav, “Reversing Jensen’s inequality for information-theoretic analyses,” *Information*, vol. 13, no. 1, 2022, p. 39. DOI: 10.3390/info13010039.

[143] N. Merhav, “D-semifaithful codes that are universal over both memoryless sources and distortion measures,” *IEEE Trans. Inf. Theory*, vol. 69, no. 7, 2023, pp. 4746–4757.

[144] N. Merhav, “Some families of Jensen-like inequalities with application to information theory,” *Entropy*, vol. 25, no. 5, 2023, p. 752. DOI: 10.3390/e25050752.

[145] N. Merhav and M. Feder, “Minimax universal decoding with an erasure option,” *IEEE Trans. Inf. Theory*, vol. 53, no. 5, 2007, pp. 1664–1675.

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

[147] N. Merhav and E. Sabbag, “Optimal watermark embedding and detection strategies under limited detection resources,” *IEEE Trans. Inf. Theory*, vol. 54, no. 1, 2008, pp. 255–274.

[148] N. Merhav and I. Sason, “An integral representation of the logarithmic function with applications in information theory,” *Entropy*, vol. 22, no. 1, 2020, p. 51. DOI: 10.3390/e22010051.

[149] N. Merhav and I. Sason, “Some useful integral representations for information-theoretic analyses,” *Entropy*, vol. 22, no. 6, 2020, p. 707. DOI: 10.3390/e22060707.

[150] N. Merhav and M. J. Weinberger, "On universal simulation of information sources using training data," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, 2004, pp. 5–20.

[151] M. Mézard and A. Montanari, *Information, Physics and Computation.* Oxford University Press, 2009.

[152] M. Mitzenmacher and E. Upfal, *Probability and Computing: Randomization and Probabilistic Techniques in Algorithms and Data Analysis.* Cambridge University Press, 2017.

[153] P. Moulin, "The log-volume of optimal codes for memoryless channels, asymptotically within a few nats," *IEEE Trans. Inf. Theory*, vol. 63, no. 4, 2017, pp. 2278–2313.

[154] P. Moulin and Y. Wang, "Capacity and random-coding exponents for channel coding with side information," *IEEE Trans. Inf. Theory*, vol. 53, no. 4, 2007, pp. 1326–1347.

[155] S. Natarajan, "Large deviations, hypotheses testing, and source coding for finite Markov chains," *IEEE Trans. Inf. Theory*, vol. 31, no. 3, 1985, pp. 360–365.

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

[157] A. Nazari, R. Venkataramanan, D. Krithivasan, S. S. Pradhan, and A. Anastasopoulos, "Typicality graphs: Large deviation analysis," *arXiv preprint arXiv:1010.1317*, 2010.

[158] T. Neuschel, "Apéry polynomials and the multivariate saddle point method," *Contr. Approx.*, vol. 40, 2014, pp. 487–507.

[159] R. Nevanlinna and V. Paatero, *Introduction to Complex Analysis*, vol. 310. American Mathematical Society, 2007.

[160] Y. Oohama and T. S. Han, "Universal coding for the Slepian-Wolf data compression system and the strong converse theorem," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, 1994, pp. 1908–1919.

[161] L. Ozarow, "The capacity of the white Gaussian multiple access channel with feedback," *IEEE Trans. Inf. Theory*, vol. 30, no. 4, 1984, pp. 623–629.

[162] A. Papoulis, *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill, 1991.

[163] M. B. Parizi and E. Telatar, "On the secrecy exponent of the wire-tap channel," in *Proc. 2015-Fall IEEE Information Theory Workshop (ITW 2015-Fall)*, IEEE, pp. 287–291, 2015.
[164] M. B. Parizi, E. Telatar, and N. Merhav, "Exact random coding secrecy exponents for the wiretap channel," *IEEE Trans. Inf. Theory*, vol. 63, no. 1, 2016, pp. 509–531.
[165] G. S. Poltyrev, "Random coding bounds for discrete memoryless channels," *Problemy Peredachi Informatsii*, vol. 18, no. 1, 1982, pp. 12–26.
[166] Y. Polyanskiy, "Channel coding: Non-asymptotic fundamental limits," Ph.D. dissertation, Department of Electrical Engineering, Princeton University, 2010.
[167] Y. Polyanskiy and S. Verdú, "Channel dispersion and moderate deviations limits for memoryless channels," in *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, IEEE, pp. 1334–1339, 2010.
[168] S. S. Pradhan, J. Chou, and K. Ramchandran, "Duality between source coding and channel coding and its extension to the side information case," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, 2003, pp. 1181–1203.
[169] M. Raginsky and I. Sason, "Concentration of measure inequalities in information theory, communications, and coding," *Foundations and Trends® in Communications and Information Theory*, vol. 10, no. 1-2, 2013, pp. 1–246.
[170] M. S. Rahman and A. B. Wagner, "On the optimality of binning for distributed hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 58, no. 10, 2012, pp. 6282–6303.
[171] T. Richardson and R. Urbanke, *Modern Coding Theory.* Cambridge University Press, 2008.
[172] P. Ruján, "Finite temperature error-correcting codes," *Physical review letters*, vol. 70, no. 19, 1993, p. 2968.
[173] J. Scarlett, "On the dispersions of the Gel'fand–Pinsker channel and dirty paper coding," *IEEE Trans. Inf. Theory*, vol. 61, no. 9, 2015, pp. 4569–4586.

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

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

[176] J. Scarlett, A. Martinez, and A. Guillén i Fàbregas, "Mismatched multi-letter successive decoding for the multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 64, no. 4, 2017, pp. 2253–2266.

[177] J. Scarlett, L. Peng, N. Merhav, A. Martinez, and A. Guillén i Fàbregas, "Expurgated random-coding ensembles: Exponents, refinements, and connections," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, 2014, pp. 4449–4462.

[178] J. Scarlett and V. Y. F. Tan, "Second-order asymptotics for the Gaussian MAC with degraded message sets," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, 2015, pp. 6700–6718.

[179] J. Scarlett, A. G. i Fàbregas, A. Somekh-Baruch, and A. Martinez, "Information-theoretic foundations of mismatched decoding," *Foundations and Trends® in Communications and Information Theory*, vol. 17, no. 2–3, 2020, pp. 149–401.

[180] J. M. Scarlett, "Reliable communication under mismatched decoding," Ph.D. dissertation, University of Cambridge, 2014.

[181] G. Schwartz, "Estimating the dimension of a model," *Ann. Stat.*, vol. 6, 1978, pp. 461–464.

[182] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, 1948, pp. 379–423.

[183] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels. i," *Information and Control*, vol. 10, no. 1, 1967, pp. 65–103.

[184] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels. ii," *Information and Control*, vol. 10, no. 5, 1967, pp. 522–552.

[185] H. Shimokawa, T. S. Han, and S. Amari, "Error bound of hypothesis testing with data compression," in *Proceedings of 1994 IEEE International Symposium on Information Theory*, IEEE, p. 114, 1994.

[186] N. Shulman, "Communication over an unknown channel via common broadcasting," Ph.D. dissertation, Tel-Aviv Univ., Tel-Aviv, Israel, 2003.

[187] S. Simić, "On a new converse of Jensen's inequality," *Publications de l'Institut Mathématique*, vol. 85, no. 99, 2009, pp. 107–110.

[188] M. Sion, "On general minimax theorems," *Pacific Journal of Mathematics*, vol. 8, no. 1, 1958, pp. 171–176.

[189] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, 1973, pp. 471–480.

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

[191] A. Somekh-Baruch, J. Scarlett, and A. Guillén i Fàbregas, "Generalized random Gilbert–Varshamov codes," *IEEE Trans. Inf. Theory*, vol. 65, no. 6, 2019, pp. 3452–3469.

[192] J. Song, S. Still, R. D. H. Rojas, I. P. Castillo, and M. Marsili, "Optimal work extraction and mutual information in a generalized Szilárd engine," *Physical Review E.*, vol. 103, 2021, p. 052 121.

[193] R. Tamir and N. Merhav, "Error exponents in the bee identification problem," *IEEE Trans. Inf. Theory*, vol. 67, no. 10, 2021, pp. 6564–6582.

[194] R. Tamir and N. Merhav, "Trade-offs between error exponents and excess-rate exponents of typical Slepian–Wolf codes," *Entropy*, vol. 23, no. 3, 2021, p. 265.

[195] R. Tamir and N. Merhav, "Universal decoding for the typical random code and for the expurgated code," *IEEE Trans. Inf. Theory*, vol. 68, no. 4, 2021, pp. 2156–2168.

[196] R. Tamir and N. Merhav, "Error exponents of the dirty-paper and Gel'fand–Pinsker channels," *IEEE Trans. Inf. Theory*, vol. 69, no. 12, 2023, pp. 7479–7498.

[197] R. Tamir, N. Merhav, N. Weinberger, and A. Guillén i Fàbregas, "Large deviations behavior of the logarithmic error probability of random codes," *IEEE Trans. Inf. Theory*, vol. 66, no. 11, 2020, pp. 6635–6659.

[198] V. Y. F. Tan, "Asymptotic estimates in information theory with non-vanishing error probabilities," *Foundations and Trends in Communications and Information Theory*, vol. 11, no. 1-2, 2014, pp. 1–184.

[199] V. Y. F. Tan and M. Tomamichel, "The third-order term in the normal approximation for the AWGN channel," in *Proc. 2014 IEEE International Symposium on Information Theory (ISIT 2014)*, pp. 2077–2081, Honolulu, HI, U.S.A, 2014.

[200] A. Tandon, V. Y. F. Tan, and L. R. Varshney, "The bee-identification problem: Bounds on the error exponent," *IEEE Transactions on Communications*, vol. 67, no. 11, 2019, pp. 7405–7416.

[201] A. E. Taylor, *General Theory of Functions and Integration.* Courier Corporation, 1985.

[202] A. Tchamkerten, V. Chandar, and G. W. Wornell, "Communication under strong asynchronism," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, 2009, pp. 4508–4528.

[203] A. Tchamkerten, V. Chandar, and G. W. Wornell, "Asynchronous communication: Capacity bounds and suboptimality of training," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, 2012, pp. 1227–1255.

[204] S. Tridenski and A. Somekh-Baruch, "The method of types for the AWGN channel," *arXiv preprint arXiv:2307.13322*, 2023.

[205] L. V. Truong, G. Cocco, J. Font-Segura, and A. Guillén i Fàbregas, "Concentration properties of random codes," *IEEE Trans. Inf. Theory*, vol. 69, no. 12, 2023, pp. 7499–7537.

[206] L. V. Truong and A. Guillén i Fàbregas, "Generalized random Gilbert-Varshamov codes: Typical error exponent and concentration properties," *IEEE Trans. Inf. Theory*, 2023.

[207] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* Cambridge university press, 2005.

[208] E. Tuncel, “Capacity/storage tradeoff in high-dimensional identification systems,” *IEEE Trans. Inf. Theory*, vol. 55, no. 5, 2009, pp. 2097–2106.

[209] G. Ungerboeck, “Channel coding with multilevel/phase signals,” *IEEE Trans. Inf. Theory*, vol. 28, no. 1, 1982, pp. 55–67.

[210] A. Vamvatsikos, *On the Wagner–Anantharam outer bound and achievable Gaussian source coding exponents*, 2007.

[211] G. Vázquez-Vilar, A. Guillén i Fàbregas, T. Koch, and A. Lancho, “Saddlepoint approximation of the error probability of binary hypothesis testing,” in *Proc. 2018 IEEE International Symposium on Information Theory (ISIT 2018)*, pp. 2306–2310, Vail, CO, USA, 2018.

[212] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding.* Dover Publications, 2009.

[213] D. Wang, V. Chandar, S.-Y. Chung, and G. W. Wornell, “Error exponents in asynchronous communication,” in *2011 IEEE International Symposium on Information Theory Proceedings*, IEEE, pp. 1071–1075, 2011.

[214] M. J. Weinberger, N. Merhav, and M. Feder, “Optimal sequential probability assignment for individual sequences,” *IEEE Trans. Inf. Theory*, vol. 40, no. 2, 1994, pp. 384–396.

[215] N. Weinberger and Y. Kochman, “On the reliability function of distributed hypothesis testing under optimal detection,” *IEEE Trans. Inf. Theory*, vol. 65, no. 8, 2019, pp. 4940–4965.

[216] N. Weinberger and N. Merhav, “Codeword or noise? Exact random coding exponents for joint detection and decoding,” *IEEE Trans. Inf. Theory*, vol. 60, no. 9, 2014, pp. 5077–5094.

[217] N. Weinberger and N. Merhav, “Optimum tradeoffs between the error exponent and the excess-rate exponent of variable-rate Slepian–Wolf coding,” *IEEE Trans. Inf. Theory*, vol. 61, no. 4, 2015, pp. 2165–2190.

[218] N. Weinberger and N. Merhav, “A large deviations approach to secure lossy compression,” *IEEE Trans. Inf. Theory*, vol. 63, no. 4, 2017, pp. 2533–2559.

[219] N. Weinberger and N. Merhav, "Channel detection in coded communication," *IEEE Trans. Inf. Theory*, vol. 63, no. 10, 2017, pp. 6364–6392.

[220] N. Weinberger and N. Merhav, "Simplified erasure/list decoding," *IEEE Trans. Inf. Theory*, vol. 63, no. 7, 2017, pp. 4218–4239.

[221] P. Whittle, "Some distributions and moment formulae for the Markov chain," *J. Roy. Stat. Soc.*, B, vol. 17, 1955, pp. 235–242.

[222] F. M. J. Willems, T. Kalker, J. Goseling, and J.-P. Linnartz, "On the capacity of a biometrical identification system," in *IEEE International Symposium on Information Theory*, pp. 82–82, 2003.

[223] J. M. Wozencraft, "List decoding," *Quarterly Progress Report*, vol. 48, 1958, pp. 90–95.

[224] G. Wunder, B. Groß, R. Fritschek, and R. F. Schaefer, "A reverse Jensen inequality result with application to mutual information estimation," in *Proc. 2021 IEEE Information Theory Workshop (ITW 2021)*, Kanazawa, Japan, 2021.

[225] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 22, no. 1, 1976, pp. 1–10.

[226] A. D. Wyner, "The wire-tap channel," *Bell System Technical Journal*, vol. 54, no. 8, 1975, pp. 1355–1387.

[227] M. H. Yassaee, M. R. Aref, and A. Gohari, "A technique for deriving one-shot achievability results in network information theory," in *2013 IEEE International Symposium on Information Theory*, IEEE, pp. 1287–1291, 2013.

[228] R. C. Yavas, V. Kostina, and M. Effros, "Third-order analysis of channel coding in the small-to-moderate deviations regime," in *Proc. 2022 IEEE International Symposium on Inform. Theory (ISIT 2022)*, pp. 2309–2314, Espoo, Finland, 2022.

[229] R. Zamir, *Lattice Coding for Signals and Networks: A Structured Coding Approach to Quantization, Modulation, and Multiuser Information Theory.* Cambridge University Press, 2014.

[230] Y. Zhong, F. Alajaji, and L. L. Campbell, "A type covering lemma and the excess distortion exponent for coding memoryless Laplacian sources," in *Proc. 23rd Biennial Symposium on Communications*, pp. 100–103, Kingston, ON, Canada, 2006.

[231] Y. Zhong, F. Alajaji, and L. L. Campbell, "Joint source-channel coding excess distortion exponent for some memoryless continuous-alphabet systems," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, 2009, pp. 1296–1319.

[232] L. Zhou, V. Y. F. Tan, and M. Motani, "Second-order and moderate deviations asymptotics for successive refinement," *IEEE Trans. Inf. Theory*, vol. 63, no. 5, 2017, p. 2921.