Vazquez-Vilar, G. (2021). Error Probability Bounds for Gaussian Channels Under Maximal and Average Power Constraints. *IEEE Transactions on Information Theory, 67*(6), pp. 3965–3985.

DOI: 10.1109/tit.2021.3063311

# Error Probability Bounds for Gaussian Channels under Maximal and Average Power Constraints

Gonzalo Vazquez-Vilar, *Member, IEEE*

***Abstract*—This paper studies the performance of block coding on an additive white Gaussian noise channel under different power limitations at the transmitter. New lower bounds are presented for the minimum error probability of codes satisfying maximal and average power constraints. These bounds are tighter than previous results in the finite blocklength regime, and yield a better understanding on the structure of good codes under an average power limitation. Evaluation of these bounds for short and moderate blocklengths is also discussed.**

***Index Terms*—Gaussian channel, channel coding, finite blocklength analysis, hypothesis testing, meta-converse, maximal power constraint, average power constraint, constellation design.**

## I. Introduction

We consider the problem of transmitting equiprobable messages over several uses of an additive white Gaussian noise (AWGN) channel using block codes. Given its practical importance, the AWGN channel has been widely studied assuming different power limitations at the transmitter:

(i) Equal power constraint, forcing every codeword in the transmission code to have equal energy.
(ii) Maximal power constraint, that requires that each of the codewords satisfies a certain energy threshold.
(iii) Average power constraint, for which the energy constraint is satisfied in average (thus allowing that some of the codewords violate the threshold).

In his seminal 1948 work [3], Shannon established the capacity of the power constrained AWGN channel, defined as the highest transmission rate under which reliable communication is possible with arbitrarily long codewords. A more refined asymptotic analysis follows from the study of the *reliability function*, which characterizes the exponential dependence between the error probability and the length of the codewords for a certain transmission rate. For the power-constrained AWGN channel, Shannon obtained the reliability function for rates close to the channel capacity [4]. Both the capacity [3] and the reliability function [4] of the AWGN channel do not depend on the specific power restriction considered at the transmitter. We conclude that equal, maximal and average power constraints can be cast as asymptotically equivalent.[1]

An alternative analysis of the reliability function is based on hypothesis testing. The channel coding error probability can be related to that of a surrogate binary hypothesis test between the distribution induced by the codebook and a certain auxiliary distribution [6]. An application of this technique was used in [7] to obtain the sphere-packing bound to the channel coding reliability function for general channels (see [8]–[12] for alternative derivations and refinements). To obtain the sphere-packing exponent, the hypothesis testing technique needs to be applied with an appropriately chosen auxiliary distribution, denoted as *exponent-achieving output distribution* (analogously to the *capacity-achieving output distribution* that follows from the channel capacity analysis). The sphere-packing exponent for the power-constrained AWGN channel was studied in [11, Sec. 4] and [12, Sec. 11].

While the focus of [4] is on the asymptotics of the power-constrained AWGN channel, Shannon also obtained upper and lower bounds in the finite blocklength regime [4, Eq. (20)]. His derivation follows from applying certain geometric arguments to codewords lying on the surface of an $n$-dimensional sphere, i.e., satisfying an equal power constraint, and then extending these results to maximal and average power limitations [4, Sec. XIII]. Following a different approach, Polyanskiy, Poor and Verdú established a fundamental lower bound to the error probability in the finite blocklength regime [13, Th. 27]. This result is usually referred to as *meta-converse* and corresponds to the error probability of a hypothesis test (for a formal definition see Sec. II-B). The standard application of the meta-converse bound for a specific channel requires either to solve a minimax optimization problem, or to make a lucky guess for the auxiliary distribution appearing in the hypothesis test. For the AWGN channel and an auxiliary distribution equal to the capacity achieving output distribution, the meta-converse particularizes to [13, Th. 41]. This bound is slightly weaker than Shannon's [4, Eq. (20)] for an equal power constraint and can be extended to maximal and average power constraints using the techniques in [4, Sec. XIII] (see [13, Lem. 39]). Polyanskiy also studied the exact solution of the meta-converse minimax optimization problem in [14]. Exploiting the existing symmetries in the AWGN channel with an equal power constraint, [14, Sec. VI] shows that, for a certain non-product auxiliary distribution, the meta-converse bound coincides with Shannon lower bound [4, Eq. (20)]. Therefore, Shannon lower

G. Vazquez-Vilar is with the Signal Theory and Communications Department, Universidad Carlos III de Madrid, 28911 Leganés, Spain, and with the Gregorio Marañón Health Research Institute, 28007 Madrid, Spain (e-mail: gvazquez@ieee.org).

This work was supported in part by the European Research Council (ERC) under grant 714161, and by the Spanish Ministry of Economy and Competitiveness under grant TEC2016-78434-C3 (AEI/FEDER, EU).

This work was presented in part at the 2019 IEEE International Symposium on Information Theory [1], Paris, France, and at the 2020 International Zürich Seminar on Communication and Information [2], Zürich, Switzerland.



[1]Note however that some asymptotic differences still exits. The strong-converse error exponent (relevant for rates above capacity) under equal and maximal power constraints is strictly positive, while it is zero under an average-power constraint [5, Sec. 4.3].

bound is still the tightest finite-length converse bound for the AWGN channel under an equal power constraint and it is often used as a benchmark for practical codes (see, e.g., [15]–[19]).

In this work, we complement the existing results in the literature for the AWGN channel with new lower bounds on the error probability of codes under maximal and average power limitations at the transmitter. In particular, our main contributions are the following:

- We provide an exhaustive characterization of the error probability of a binary hypothesis test between two Gaussian distributions. The error probability of this test corresponds to the meta-converse bound for an equal power constraint and an auxiliary independent and identically distributed (i.i.d.) zero-mean Gaussian distribution (not necessarily capacity achieving).
- Using this characterization, we optimize the meta-converse bound over input distributions satisfying maximal and average power constraints. We obtain that the error probability of a hypothesis test between two i.i.d. Gaussian distributions yields a lower bound that holds directly under a maximal power limitation. For an average power limitation, we obtain that this bound holds if the codebook size is below a certain threshold and, with a certain transformation, also above this threshold.
- We propose a saddlepoint expansion to estimate the error probability of a hypothesis test between two i.i.d. Gaussian distributions. This expansion yields a simple expression that can be used to evaluate [13, Th. 41] and the new bounds for maximal and average power constraints presented in this work.
- We provide several numerical examples and compare the new bounds with previous results in the literature showing their advantage in the finite-length regime. We show that considering an exponent-achieving auxiliary distribution under equal, maximal and average power constraints yields tighter bounds in general.

Given the difficulty of computing [4, eq. (20)] (see, e.g., [17], [20]–[22]), the bounds proposed here are not only tighter (for maximal and average power constraints) but also easier to evaluate than the original lower bound by Shannon. While the results obtained are specific for the AWGN channel, the techniques used in this work can in principle be extended to other scenarios in which the optimization of the meta-converse bound over input distributions is needed.

The organization of the manuscript is as follows. Section II presents the system model and a formal definition of the power constraints. Section III compares Shannon lower bound with the meta-converse for the AWGN channel with an equal power constraint. This section provides a geometric interpretation of [13, Th. 41] analogous to the one formulated in [4]. Sections IV and V introduce new bounds for maximal and average power constraints, respectively. The evaluation of the proposed bounds is studied in Section VI. Section VII presents a numerical comparison of the bounds with previous results and studies the effect of considering capacity and exponent achieving auxiliary distributions. Finally, Section VIII concludes the article discussing the results of this work.

## II. System Model and Preliminaries

We consider the problem of transmitting $M$ equiprobable messages over $n$ uses of an AWGN channel with noise power $\sigma^2$. Specifically, we consider a channel $W \triangleq P_{\boldsymbol{Y}|\boldsymbol{X}}$ which, for an input $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathcal{X}$ and output $\boldsymbol{y} = (y_1, y_2, \ldots, y_n) \in \mathcal{Y}$, with $\mathcal{X} = \mathcal{Y} = \mathbb{R}^n$, has a probability density function (pdf)

$$w(\boldsymbol{y}|\boldsymbol{x}) = \prod_{i=1}^{n} \varphi_{x_i,\sigma}(y_i), \tag{1}$$

where $\varphi_{\mu,\sigma}(\cdot)$ denotes the pdf of the Gaussian distribution,

$$\varphi_{\mu,\sigma}(y) \triangleq \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}}. \tag{2}$$

In our communications system, the source generates a message $v \in \{1, \ldots, M\}$ randomly with equal probability. This message is then mapped by the encoder to a codeword $\boldsymbol{c}_v$ using a codebook $\mathcal{C} \triangleq \{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_M\}$, and the sequence $\boldsymbol{x} = \boldsymbol{c}_v$ is transmitted over the channel. Then, based on the channel output $\boldsymbol{y}$, the decoder guesses the transmitted message $\hat{v} \in \{1, \ldots, M\}$. In the following we shall assume that maximum likelihood (ML) decoding is used at the receiver.[2] We define the average error probability of a codebook $\mathcal{C}$ as

$$\epsilon(\mathcal{C}) \triangleq \Pr\{\hat{V} \neq V\}, \tag{3}$$

where the underlying probability is induced by the chain of source, encoder, channel, and ML decoder.[3]

### A. Power constrained codebooks

The focus of this work is on obtaining lower bounds to the error probability $\epsilon(\mathcal{C})$ for codebooks $\mathcal{C} \triangleq \{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_M\}$ satisfying the following power constraints:

(i) Equal power constraint:

$$\mathcal{F}_{\mathrm{e}}(n, M, \Upsilon) \triangleq \Big\{\mathcal{C} \;\Big|\; \|\boldsymbol{c}_i\|^2 = n\Upsilon, \quad i = 1, \ldots, M\Big\}. \tag{4}$$

(ii) Maximal power constraint:

$$\mathcal{F}_{\mathrm{m}}(n, M, \Upsilon) \triangleq \Big\{\mathcal{C} \;\Big|\; \|\boldsymbol{c}_i\|^2 \leq n\Upsilon, \quad i = 1, \ldots, M\Big\}. \tag{5}$$

(iii) Average power constraint:

$$\mathcal{F}_{\mathrm{a}}(n, M, \Upsilon) \triangleq \Big\{\mathcal{C} \;\Big|\; \tfrac{1}{M}\sum\nolimits_{i=1}^{M} \|\boldsymbol{c}_i\|^2 \leq n\Upsilon\Big\}. \tag{6}$$

For fixed $n, M, \Upsilon$, we define the minimum error probability under a power constraint $i \in \{\mathrm{e}, \mathrm{m}, \mathrm{a}\}$ as

$$\epsilon_i^\star(n, M, \Upsilon) \triangleq \min_{\mathcal{C} \in \mathcal{F}_i(n,M,\Upsilon)} \epsilon(\mathcal{C}). \tag{7}$$

The next result relates the minimum error probability in the three scenarios considered via simple inequalities.

[2]Since the ML decoder minimizes the error probability, lower bounds to ML decoding error probability also apply to other decoding schemes.

[3]All the results in this article are derived under the average error probability formalism. For the maximal error probability, defined as $\epsilon_{\max}(\mathcal{C}) \triangleq \max_{v \in \{1,\ldots,M\}} \Pr\{\hat{V} \neq V \,|\, V = v\}$, it holds that $\epsilon_{\max}(\mathcal{C}) \geq \epsilon(\mathcal{C})$ and lower bounds on $\epsilon(\mathcal{C})$ also apply to $\epsilon_{\max}(\mathcal{C})$.

*Lemma 1 ([4, Sec. XIII], [5, Lemma 65]):* For any $n, M, \Upsilon > 0$, and $0 < s < 1$, the following inequalities hold:

$$\epsilon_{\mathrm{e}}^{\star}(n, M, \Upsilon) \;\geq\; \epsilon_{\mathrm{m}}^{\star}(n, M, \Upsilon) \;\geq \epsilon_{\mathrm{e}}^{\star}(n+1, M, \Upsilon), \tag{8}$$

$$\epsilon_{\mathrm{m}}^{\star}(n, M, \Upsilon) \;\geq\; \epsilon_{\mathrm{a}}^{\star}(n, M, \Upsilon) \;\geq\; s\epsilon_{\mathrm{m}}^{\star}\Big(n, sM, \tfrac{\Upsilon}{1-s}\Big). \tag{9}$$

*Remark:* The relations (8)-(9) were first proposed by Shannon in [4, Sec. XIII]. Nevertheless, there is a typo in the last equation of [4, Sec. XIII], which has been corrected in (9). In [4, Sec. XIII], Shannon states that "The probability of error for the new code [satisfying the maximal power constraint] cannot exceed $1/\alpha$ times that of the original code [satisfying the average power constraint]" (brackets added). While this reasoning is right, using his notation, this statement traslates to $P'_{\text{e opt}} \leq \frac{1}{\alpha} P''_{\text{e opt}}$ and hence $P''_{\text{e opt}} \geq \alpha P'_{\text{e opt}}$, which does not coincide with the last equation of [4, Sec. XIII]. These relations were rederived in [5, Lemma 65], where the statement of the bound corresponding to (9) is correct.

The relations from Lemma 1 show that lower and upper bounds on the error probability under a given power constraint can be adapted to other settings via simple transformations. Focusing on converse bounds, the analysis under an equal power constraint is usually simpler. However, the maximal power constraint and average power constraint are more relevant in practice, and therefore the transformations from Lemma 1 are often used to adapt converse bounds to this settings. While the loss incurred by using these transformations becomes negligible in the asymptotic regime, it can have a relevant impact at finite blocklengths.

In Sections IV and V we shall prove direct lower bounds in the finite blocklength regime for maximal and average power constraints. The new bounds are tighter than previous results resorting on the transformations from Lemma 1.

### B. Meta-converse bound

In [13], Polyanskiy *et al.* proved that the error probability of a binary hypothesis test with certain parameters can be used to lower bound the error probability $\epsilon(\mathcal{C})$ for a channel $W$. In particular, [13, Th. 27] establishes the *meta-converse bound*:

$$\epsilon(\mathcal{C}) \geq \inf_{P \in \mathcal{P}} \sup_{Q} \Big\{ \alpha_{\frac{1}{M}} \big(PW, P \times Q\big) \Big\}, \tag{10}$$

where $\mathcal{P}$ is the set of distributions over the input alphabet $\mathcal{X}$ satisfying a certain constraint, $Q$ is an auxiliary distribution over the output alphabet $\mathcal{Y}$ which is not allowed to depend on the input $\boldsymbol{x}$, and where $\alpha_\beta\left(PW, P \times Q\right)$ denotes the minimum type-I error for a maximum type-II error $\beta \in [0, 1]$ in a binary hypothesis testing problem between the distributions $PW$ and $P \times Q$. Formally, for two distributions $A$ and $B$ defined over an alphabet $\mathcal{Z}$, the minimum type-I error for a maximum type-II error $\beta \in [0, 1]$ is given by

$$\alpha_\beta(A, B) \triangleq \inf_{\substack{0 \leq T \leq 1: \\ \mathrm{E}_B[T(Z)] \leq \beta}} \Big\{ 1 - \mathrm{E}_A[T(Z)] \Big\}, \tag{11}$$

where $T : \mathcal{Z} \to [0, 1]$ and $\mathrm{E}_P[\cdot]$ denotes the expectation operator with respect to the random variable $Z \sim P$.

The results in this work follow from the following inequality chain, which always holds

$$\inf_{P \in \mathcal{P}} \sup_{Q} \Big\{ \alpha_{\frac{1}{M}} \big(PW, P \times Q\big) \Big\}$$

$$\geq \sup_{Q} \inf_{P \in \mathcal{P}} \Big\{ \alpha_{\frac{1}{M}} \big(PW, P \times Q\big) \Big\} \tag{12}$$

$$\geq \inf_{P \in \mathcal{P}} \Big\{ \alpha_{\frac{1}{M}} \big(PW, P \times Q\big) \Big\}. \tag{13}$$

Here, the first step follows from the max-min inequality, and the second is the result of fixing the auxiliary distribution $Q$. The properties of the exact minimax solution to the optimizations in the left-hand side of (12) are studied in [14]. Under mild assumptions, (12) is satisfied with equality and the saddle point property holds [14, Sec. V]. Therefore, in practice it is possible to fix the auxiliary distribution $Q$ in (10) and still obtain tight lower bounds. However, the minimization needs to be carried out over all the input probability distributions $P$ (non necessarily product) satisfying the constraint $P \in \mathcal{P}$.

In the following sections we consider the optimization of the meta-converse bound over input distributions for the AWGN channel under equal, maximal and average power constraints and for a certain auxiliary distribution $Q$.

## III. Lower Bounds for Equal Power Constraints

In this section we briefly discuss the results from [4], [13] and [14]. The bounds presented here apply for codes $\mathcal{C} \in \mathcal{F}_{\mathrm{e}}(n, M, \Upsilon)$ satisfying an equal power constraint, and they will be relevant in the sequel.

### A. Shannon cone-packing bound

Consider a $n$-dimensional cone with vertex at the origin, with axis going through the vector $\boldsymbol{x} = (1, \ldots, 1)$ and with half-angle $\theta$. Let $\Phi_n(\theta, \bar{\sigma}^2)$ denote the probability that the vector $\boldsymbol{x}$ is moved outside this cone by effect of the i.i.d. zero mean Gaussian noise with variance $\bar{\sigma}^2$ in each dimension.

*Theorem 1 ([4, Eq. (20)]):* Let $\theta_{n,M}$ be the half-angle of a cone with solid angle equal to $\Omega_n / M$, where $\Omega_n$ is the surface of the $n$-dimensional hypersphere. The error probability of an equal-power constrained code satisfies

$$\epsilon_{\mathrm{e}}^{\star}(n, M, \Upsilon) \geq \Phi_n\left(\theta_{n,M}, \frac{\sigma^2}{\Upsilon}\right). \tag{14}$$

The derivation of this bound follows from deforming the optimal decoding regions, which for codewords lying on the surface of an sphere correspond to pyramids, to cones of the same volume (see [4, Fig. 1]) and analyzing the resulting error probability. Given this geometric interpretation, Theorem 1 is often referred to as *cone-packing bound*. While the resulting expression is conceptually simple and accurate for low SNRs and relatively short codes [23], it is difficult to evaluate. Approximate and exact computation of this bound is treated, e.g., in [21], [22].

### B. Meta-converse bound for the AWGN channel

We consider now the meta-converse bound (10) for the equal power constrained AWGN channel. The exact solution of the minimax optimization in (10) was studied in [14, Sec. VI.F]. For an equal power constraint the codewords $\boldsymbol{c}_i$ are restricted to lie on the surface of a sphere of squared radius $n\Upsilon$. The optimal decoder does not depend on the norm of the received sequence and only on its direction, and therefore it is possible to define an equivalent channel $P_{\tilde{\boldsymbol{Y}}|\tilde{\boldsymbol{X}}}$ with

$$\tilde{\boldsymbol{x}} = \frac{\boldsymbol{x}}{\sqrt{n\Upsilon}} \in \tilde{\mathcal{X}}, \quad \tilde{\boldsymbol{y}} = \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|} \in \tilde{\mathcal{Y}}, \tag{15}$$

where $\tilde{\mathcal{X}} = \tilde{\mathcal{Y}} = \mathbb{S}^{n-1}$ corresponds to the $(n-1)$-dimensional sphere centered at the origin and with unit radius. Applying the meta-converse bound (10) to the random map $W = P_{\tilde{\boldsymbol{Y}}|\tilde{\boldsymbol{X}}}$, and mapping this result back to the original channel $P_{\boldsymbol{Y}|\boldsymbol{X}}$, it follows that the meta-converse (10) recovers Shannon cone-packing bound (14) [14, Sec. VI.F]. Analyzing the optimizing distributions in the original domain, we conclude that this bound is attained for a non-product auxiliary distribution $Q$.

Nevertheless, the meta-converse bound (10) can be weakened by fixing the auxiliary distribution $Q$. For any input distribution lying on the surface of a $n$-dimensional hypersphere of squared radius $n\Upsilon$ (equal power constraint) and an auxiliary distribution $Q$ that is invariant under rotations around the origin, it holds that [13, Lem. 29]

$$\alpha_{\frac{1}{M}}\big(PW, P\times Q\big) = \alpha_{\frac{1}{M}}\big(\varphi^n_{\sqrt{\Upsilon},\sigma}, Q\big). \tag{16}$$

In [13, Sec. III.J.2], Polyanskiy *et al.* fixed the auxiliary distribution $Q$ to be an i.i.d. zero-mean Gaussian distribution with variance $\theta^2$. The pdf of this auxiliary distribution is

$$q(\boldsymbol{y}) = \prod_{i=1}^{n} \varphi_{0,\theta}(y_i), \tag{17}$$

and its application yields the following lower bound.

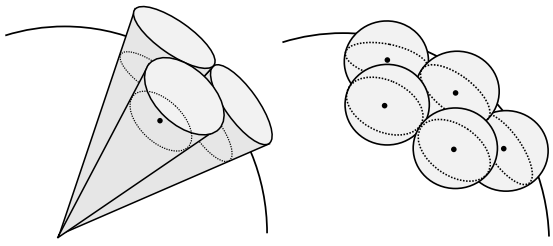
*Theorem 2 ([13, Th. 41]):* Let $\theta^2 = \Upsilon + \sigma^2$. The error probability of an equal-power constrained code satisfies

$$\epsilon^{\star}_{\mathrm{e}}(n, M, \Upsilon) \geq \alpha_{\frac{1}{M}}\big(\varphi^n_{\sqrt{\Upsilon},\sigma}, \varphi^n_{0,\theta}\big). \tag{18}$$

This expression admits a parametric form involving two Marcum-$Q$ functions (see Proposition 1 in Appendix A). However, for fixed rate $R \triangleq \frac{1}{n}\log_2 M$, the term $\frac{1}{M} = 2^{-nR}$ decreases exponentially with the blocklength and traditional series expansions of the Marcum-$Q$ function fail even for moderate values of $n$. Nevertheless, in contrast with the formulation in (14), the distributions appearing in (18) are i.i.d., and Laplace methods can be used to evaluate this bound. This point will be treated in detail in Section VI.

### C. Geometric interpretation of Theorem 2

The cone-packing bound from Theorem 1 corresponds to the probability that the additive Gaussian noise moves a given codeword out of the $n$-dimensional cone centered at the codeword that roughly covers $1/M$-th of the output space. We next show that the hypothesis-testing bound from Theorem 2 admits an analogous geometric interpretation.



(a) (b)

Fig. 1: Induced regions by (a) the Shannon cone-packing bound in (14), and (b) the hypothesis-testing bound in (18), for codewords (•) located on the shell of the sphere.

Let $\boldsymbol{x} = \big(\sqrt{\Upsilon}, \ldots, \sqrt{\Upsilon}\big)$ and let $\theta > \sigma$. For the hypothesis test on the right-hand side (RHS) of (18), the condition

$$\log \frac{\varphi^n_{\sqrt{\Upsilon},\sigma}(\boldsymbol{y})}{\varphi^n_{0,\theta}(\boldsymbol{y})} = n\log\frac{\theta}{\sigma} + \frac{\|\boldsymbol{y}\|^2}{2\theta^2} - \frac{\|\boldsymbol{y}-\boldsymbol{x}\|^2}{2\sigma^2} = t \tag{19}$$

defines the boundary of the decision region induced by the optimal Neyman-Pearson test for some $-\infty < t < \infty$. We characterize the shape of this region. To this end, we write

$$\begin{aligned}\frac{\|\boldsymbol{y}\|^2}{2\theta^2} - \frac{\|\boldsymbol{y}-\boldsymbol{x}\|^2}{2\sigma^2} &= -\frac{\theta^2-\sigma^2}{2\sigma^2\theta^2}\big(\|\boldsymbol{y}\|^2 - 2a\boldsymbol{x}^T\boldsymbol{y} + a\|\boldsymbol{x}\|^2\big) &(20)\\ &= -\frac{\theta^2-\sigma^2}{2\sigma^2\theta^2}\big(\|\boldsymbol{y}-a\boldsymbol{x}\|^2 + (a-a^2)\|\boldsymbol{x}\|^2\big), &(21)\end{aligned}$$

where we defined $a = \frac{\theta^2}{\theta^2-\sigma^2}$.

The boundary of the decision region induced by the optimal NP test, defined by (19) corresponds to (21) being equal to $t - n\log\frac{\theta}{\sigma}$. Using that, for an equal power constraint, $\|\boldsymbol{x}\|^2 = n\Upsilon$ and letting $\theta^2 = \Upsilon + \sigma^2$ (value considered in Theorem 2), this identity yields

$$\left\|\boldsymbol{y} - \Big(1+\frac{\sigma^2}{\Upsilon}\Big)\boldsymbol{x}\right\|^2 = r, \tag{22}$$

where $r = n\sigma^2\big(1+\frac{\sigma^2}{\Upsilon}\big)\big(1 - \frac{2t}{n} + \log\big(1+\frac{\Upsilon}{\sigma^2}\big)\big)$.

The region inside the boundary (22) corresponds to the interior of an $n$-dimensional sphere centered at $\big(1+\frac{\sigma^2}{\Upsilon}\big)\boldsymbol{x}$ with squared radius $r$. Then, we can describe the lower bound in Theorem 2 as the probability that the additive Gaussian noise moves the codeword $\boldsymbol{x}$ out of the $n$-dimensional sphere centered at $\big(1+\frac{\sigma^2}{\Upsilon}\big)\boldsymbol{x}$ that roughly covers $1/M$-th of the auxiliary measure $\varphi^n_{0,\theta}$.

The regions considered in Theorem 1 correspond to cones, while those induced by the hypothesis-testing bound in Theorem 2 are spheres (see Fig. 1). Cones are close to the optimal ML decoding regions for codewords evenly distributed on surface of an $n$-dimensional sphere with squared radius $n\Upsilon$. On the other hand, spheres would be close to the optimal ML decoding regions for other configurations of the codewords. This fact suggests that the hypothesis-testing bound in Theorem 2 may hold beyond the equal power constraint setting. This intuition is shown to be correct in the next sections.

## IV. Lower Bounds for Maximal Power Constraints

We consider now the family of codes satisfying a maximal power limitation, $\mathcal{C} \in \mathcal{F}_{\mathrm{m}}(n, M, \Upsilon)$. As discussed in Section II, Theorems 1 and 2 can be extended to the maximal power constraint via Lemma 1. Indeed, the second inequality in (8) can be slightly tightened to

$$\epsilon^{\star}_{\mathrm{m}}(n, M, \Upsilon) \geq \epsilon^{\star}_{\mathrm{e}}\Big(n+1, M, \tfrac{n\Upsilon}{n+1}\Big). \tag{23}$$

The proof of Lemma 1 (see [4, Sec. XIII] and [13, Lem. 39]) is based on extending a maximal power constrained codebook of length $n$ by adding an extra coordinate. The energy of the new codewords of length $n+1$ is then normalized to $(n+1)\Upsilon$, so that the equal power constraint is satisfied. The proof of (23) follows the same lines, but by normalizing the energy of the codewords to $n\Upsilon$, instead to $(n+1)\Upsilon$. Applying (23) to Theorem 1 we obtain the following result.

*Corollary 1:* Let $\theta_{n,M}$ denote the half-angle of a cone with solid angle equal to $\Omega_n/M$, where $\Omega_n$ is the surface of the $n$-dimensional hypersphere. Then,

$$\epsilon^{\star}_{\mathrm{m}}(n, M, \Upsilon) \geq \Phi_{n+1}\left(\theta_{n+1,M}, \frac{(n+1)\sigma^2}{n\Upsilon}\right). \tag{24}$$

An alternative lower bound to the error probability under maximal power constraint can be obtained via hypothesis testing. To this end, we consider the weakening of the meta-converse in (10) by fixing the auxiliary distribution $Q$ to be the zero-mean i.i.d. Gaussian distribution (17).

*Theorem 3 (Converse, maximal power constraint):* Let $\theta > \sigma$, $n \geq 1$. The error probability of a maximal-power constrained code satisfies

$$\epsilon^{\star}_{\mathrm{m}}(n, M, \Upsilon) \geq \alpha_{\frac{1}{M}}\big(\varphi^n_{\sqrt{\Upsilon},\sigma}, \varphi^n_{0,\theta}\big). \tag{25}$$

*Proof:* See Section IV-A. ∎

Setting $\theta^2 = \Upsilon + \sigma^2$ in (25), we recover the bound from Theorem 2. We conclude that the hypothesis-testing lower bound from Theorem 2 also holds under a maximal power constraint and not only for equal-power constrained codewords. This is not the case for the Shannon cone-packing bound from Theorem 1 as we show next with an example.

*Example:* We consider the problem of transmitting $M = 16$ codewords over $n = 2$ uses of an additive Gaussian noise channel. For $n = 2$, Shannon cone-packing bound (SCPB) from Theorem 1 coincides with the ML decoding error probability of a $M$-PSK constellation $\mathcal{C}_{M\text{-PSK}}$ satisfying the equal power constraint $\Upsilon$ (as 2-dimensional cones are precisely the ML decoding regions of the $M$-PSK constellation). For instance, for a 2-dimensional Gaussian channel with a signal-to-noise ratio (SNR) $\frac{\Upsilon}{\sigma^2} = 10$ and $M = 16$ codewords, we obtain SCPB $= \epsilon(\mathcal{C}_{16\text{-PSK}}) \approx 0.38$. Let us define a new code $\mathcal{C}_{M\text{-APSK}}$ composed by the points of an $(M-1)$-PSK constellation and an additional codeword located at $\boldsymbol{x} = (0, 0)$. While this code satisfies the maximal power constraint $\Upsilon$, its error probability violates SCPB for sufficiently large $M$. Indeed, the modified codebook attains $\epsilon(\mathcal{C}_{16\text{-APSK}}) \approx 0.34 < 0.38$. We conclude that, in general, Theorem 1 holds only under an equal power constraint. For a more detailed discussion comparing the bounds under different power constraints, see Section VII.

### A. Proof of Theorem 3

We consider the set of input distributions $\boldsymbol{X} \sim P$ satisfying the maximal power constraint

$$\mathcal{P}_{\mathrm{m}}(\Upsilon) \triangleq \Big\{P \;\Big|\; \Pr\big[\|\boldsymbol{X}\|^2 \leq n\Upsilon\big] = 1\Big\}. \tag{26}$$

Then, the meta-converse bound (10) for some fixed $Q$ becomes

$$\epsilon^{\star}_{\mathrm{m}}(n, M, \Upsilon) \geq \inf_{P \in \mathcal{P}_{\mathrm{m}}(\Upsilon)} \Big\{\alpha_{\frac{1}{M}}\big(PW, P \times Q\big)\Big\}. \tag{27}$$

In order to make the minimization over $P$ tractable, we shall use the following technical result.

*Lemma 2:* Let $\{P_\lambda\}$ be a family of probability measures defined over the input alphabet $\mathcal{X}$, parametrized by $\lambda \in \mathbb{R}$. Assume that the distributions $P_\lambda$ have pairwise disjoint supports and that there exists a probability distribution $S$ over the parameter $\lambda$ such that $P = \int P_\lambda S(\mathrm{d}\lambda)$. Then, the hypothesis testing error trade-off function satisfies

$$\begin{aligned}\alpha_\beta\big(PW, P \times Q\big) &\\ = \min_{\substack{\{\beta_\lambda\}:\\ \beta = \int \beta_\lambda S(\mathrm{d}\lambda)}} &\int \alpha_{\beta_\lambda}\big(P_\lambda W, P_\lambda \times Q\big) S(\mathrm{d}\lambda).\end{aligned} \tag{28}$$

*Proof:* This lemma is analogous to the second part of [14, Lem. 25]. Since here we require the family $P_\lambda$ to be parametrized by a continuous $\lambda$, for completeness, we reproduce the proof here for this case.

First, we observe that $\alpha_\beta\big(PW, P \times Q\big)$ is a jointly convex function on $(\beta, P)$ [14, Thm. 6]. Let $\{\beta_\lambda\}$ and $\{P_\lambda\}$ satisfy $\beta = \mathrm{E}_S[\beta_\lambda] = \int \beta_\lambda S(\mathrm{d}\lambda)$ and $P = \mathrm{E}_S[P_\lambda] = \int P_\lambda S(\mathrm{d}\lambda)$. Then, using Jensen's inequality it follows that

$$\mathrm{E}_S\big[\alpha_{\beta_\lambda}\big(P_\lambda W, P_\lambda \times Q\big)\big] \geq \alpha_{\mathrm{E}_S[\beta_\lambda]}\big(\mathrm{E}_S[P_\lambda]W, \mathrm{E}_S[P_\lambda] \times Q\big) \tag{29}$$

$$= \alpha_\beta\big(PW, P \times Q\big). \tag{30}$$

The RHS of (28) is thus an upper bound on $\alpha_\beta\big(PW, P \times Q\big)$.

To prove the identity (28), it remains to show that there exists $\{\beta_\lambda\}$ satisfying $\beta = \int \beta_\lambda S(\mathrm{d}\lambda)$ and such that (29)-(30) holds with equality. We consider the Neyman-Pearson test for the testing problem $\alpha_\beta\big(PW, P \times Q\big)$, which is given by

$$T(\boldsymbol{x}, \boldsymbol{y}) = \mathbb{1}\left[\log \frac{W(\boldsymbol{y}|\boldsymbol{x})}{Q(\boldsymbol{y})} > t'\right] + c\mathbb{1}\left[\log \frac{W(\boldsymbol{y}|\boldsymbol{x})}{Q(\boldsymbol{y})} = t'\right] \tag{31}$$

for $t' \geq 0$ and $c \in [0, 1]$ two parameters chosen to satisfy $\beta = \int T(\boldsymbol{x}, \boldsymbol{y}) Q(\mathrm{d}\boldsymbol{y}) P(\mathrm{d}\boldsymbol{x})$. We apply $T(\boldsymbol{x}, \boldsymbol{y})$ to the testing problem between $P_\lambda W$ and $P_\lambda \times Q$ and obtain the error probabilities

$$\epsilon_1(\lambda) = 1 - \int T(\boldsymbol{x}, \boldsymbol{y}) W(\mathrm{d}\boldsymbol{y}|\boldsymbol{x}) P_\lambda(\mathrm{d}\boldsymbol{x}), \tag{32}$$

$$\epsilon_2(\lambda) = \int T(\boldsymbol{x}, \boldsymbol{y}) Q(\mathrm{d}\boldsymbol{y}) P_\lambda(\mathrm{d}\boldsymbol{x}). \tag{33}$$

For the choice $\beta_\lambda = \epsilon_2(\lambda)$, the test (31) is precisely the Neyman-Pearson test of the problem $\alpha_{\beta_\lambda}\big(P_\lambda W, P_\lambda \times Q\big)$. Therefore,

$$\begin{aligned}
&\mathrm{E}_S\big[\alpha_{\beta_\lambda}\big(P_\lambda W, P_\lambda \times Q\big)\big] \\
&\quad = \int \epsilon_1(\lambda) S(\mathrm{d}\lambda) && (34)\\
&\quad = 1 - \int T(\boldsymbol{x},\boldsymbol{y}) W(\mathrm{d}\boldsymbol{y}|\boldsymbol{x}) P_\lambda(\mathrm{d}\boldsymbol{x}) S(\mathrm{d}\lambda) && (35)\\
&\quad = 1 - \int T(\boldsymbol{x},\boldsymbol{y}) W(\mathrm{d}\boldsymbol{y}|\boldsymbol{x}) P(\mathrm{d}\boldsymbol{x}) && (36)\\
&\quad = \alpha_\beta\big(PW, P\times Q\big). && (37)
\end{aligned}$$

Similarly, we can show that $\mathrm{E}_S\big[\beta_\lambda\big] = \mathrm{E}_S\big[\epsilon_2(\lambda)\big] = \beta$. We conclude that this choice of $\big\{\beta_\lambda\big\}$ yields equality in (30). Given the lower bound (29)-(30), it also attains the minimum in the RHS of (28) and the result follows. ∎

Lemma 2 asserts that it is possible to express a binary hypothesis test as a convex combination of disjoint sub-tests provided that the type-II error is optimally distributed among them. For our problem, we shall decompose the input distribution $P$ based on $\gamma = \|\boldsymbol{x}\|^2/n$ and apply this result.

For any $\gamma \geq 0$, we define the set $\mathcal{S}_\gamma \triangleq \big\{\boldsymbol{x} \,|\, \|\boldsymbol{x}\|^2 = n\gamma\big\}$. In words, the set $\mathcal{S}_\gamma$ corresponds to the spherical shell centered at the origin that contains all input sequences with energy $n\gamma$. Note that, whenever $\gamma_1 \neq \gamma_2$, the sets $\mathcal{S}_{\gamma_1}$ and $\mathcal{S}_{\gamma_2}$ are disjoint. We define the distribution $S(\gamma) \triangleq \Pr\{\boldsymbol{X} \in \mathcal{S}_\gamma\}$, and we let $P_\gamma$ be a distribution defined over $\mathcal{X}$ satisfying $P_\gamma(\boldsymbol{x}) = 0$ for any $\boldsymbol{x} \notin \mathcal{S}_\gamma$, and

$$P(\boldsymbol{x}) = \int P_\gamma(\boldsymbol{x}) S(\mathrm{d}\gamma). \tag{38}$$

This condition implies that $P_\gamma(\boldsymbol{x}) = \frac{P(\boldsymbol{x})}{S(\gamma)} \mathbb{1}[\boldsymbol{x} \in \mathcal{S}_\gamma]$ for $S(\gamma) > 0$, where $\mathbb{1}[\cdot]$ denotes the indicator function. When $S(\gamma) = 0$, then $P_\gamma$ can be an arbitrary distribution such that $P_\gamma(\boldsymbol{x}) = 0$ for any $\boldsymbol{x} \notin \mathcal{S}_\gamma$.

Given (38) and since the measures $P_\gamma$ have disjoint supports for different values of $\gamma$, the conditions in Lemma 2 hold for $\lambda \leftrightarrow \gamma$. Then, using (28), we obtain that the RHS of (27), for the auxiliary distribution $Q$ from (17), becomes

$$\begin{aligned}
&\inf_{P\in\mathcal{P}_{\mathrm{m}}(\Upsilon)} \Big\{\alpha_{\frac{1}{M}}\big(PW, P\times Q\big)\Big\} \\
&\quad = \inf_{\substack{\{S,\beta_\gamma\}\colon \gamma\leq\Upsilon,\\ \int \beta_\gamma S(\mathrm{d}\gamma)=\frac{1}{M}}} \left\{\int \alpha_{\beta_\gamma}\big(P_\gamma W, P_\gamma \times Q\big) S(\mathrm{d}\gamma)\right\} && (39)\\
&\quad = \inf_{\substack{\{S,\beta_\gamma\}\colon \gamma\leq\Upsilon,\\ \int \beta_\gamma S(\mathrm{d}\gamma)=\frac{1}{M}}} \left\{\int \alpha_{\beta_\gamma}\big(\varphi^n_{\sqrt{\gamma},\sigma}, \varphi^n_{0,\theta}\big) S(\mathrm{d}\gamma)\right\}, && (40)
\end{aligned}$$

where (40) follows given the spherical symmetry of each of the sub-tests in (39), since $\boldsymbol{x} = (\sqrt{\gamma},\ldots,\sqrt{\gamma}) \in \mathcal{S}_\gamma$.

In a nutshell, we transformed the original optimization over the $n$-dimensional distribution $P$ in the left-hand side of (39) into an optimization over a one-dimensional distribution $S$ and auxiliary function $\beta_\gamma$ in the RHS of (40). To obtain the lower bound stated in the theorem, we make use of the following properties of the function $\alpha_\beta\big(\varphi^n_{\sqrt{\gamma},\sigma}, \varphi^n_{0,\theta}\big)$.

*Lemma 3:* Let $0 < \sigma < \theta$, with $\sigma, \theta \in \mathbb{R}$ and $n \geq 1$. Then, the function

$$f(\beta,\gamma) \triangleq \alpha_\beta\big(\varphi^n_{\sqrt{\gamma},\sigma}, \varphi^n_{0,\theta}\big) \tag{41}$$

is non-increasing in $\gamma$ for any fixed $\beta \in [0,1]$, and convex non-increasing in $\beta$ for any fixed $\gamma > 0$.

*Proof:* The minimum type-I error $\alpha$ is a non-increasing convex function of the type-II error $\beta$ (see, *e.g.*, [14, Sec. I]). Since $f(\beta,\gamma)$ characterizes the trade-off between the type-I and type-II errors of a hypothesis test, for fixed $\gamma \geq 0$, the function $f(\beta,\gamma)$ is non-increasing and convex in $\beta \in [0,1]$.

To characterize the behavior of $f(\beta,\gamma)$ with respect to $\gamma$, in Appendix A we show that $f(\beta,\gamma)$ is differentiable with respect to both parameters and obtain the derivative of $f(\beta,\gamma)$ with respect to $\gamma$. In particular, it follows from (150) that

$$\frac{\partial f(\beta,\gamma)}{\partial \gamma} = -\frac{n}{2\delta}\left(\frac{t\delta}{\sigma^2\sqrt{n\gamma}}\right)^{\frac{n}{2}} e^{-\frac{1}{2}\left(\frac{n\gamma\sigma^2}{\delta^2}+\frac{t^2}{\sigma^2}\right)} I_{\frac{n}{2}}\left(\frac{t\sqrt{n\gamma}}{\delta}\right), \tag{42}$$

where $\delta = \theta^2 - \sigma^2 > 0$, $t$ satisfies $\beta(\gamma,t) = \beta$ for $\beta(\gamma,t)$ defined in (119) and $I_m(\cdot)$ denotes the $m$-th order modified Bessel function of the first kind.

For any $\gamma \geq 0$ and $\beta \in [0,1]$, the parameter $t$ that follows from the identity $\beta(\gamma,t) = \beta$ is non-negative. Then, using that $e^{-x/2} \geq 0$ and since $x \geq 0$ implies $I_m(x) \geq 0$, we conclude that (42) is non-positive for any $\delta = \theta^2 - \sigma^2 > 0$. As a result, the function $f(\beta,\gamma) = \alpha_\beta\big(\varphi^n_{\sqrt{\gamma},\sigma}, \varphi^n_{0,\theta}\big)$ is non-increasing in $\gamma$ for any fixed value of $\beta$, provided that the conditions in the lemma hold. ∎

According to Lemma 3, for any $0 \leq \gamma \leq \Upsilon$, it follows that $\alpha_\beta\big(\varphi^n_{\sqrt{\gamma},\sigma}, \varphi^n_{0,\theta}\big) = f(\beta,\gamma) \geq f(\beta,\Upsilon)$. As any maximal power constrained input distribution $P \in \mathcal{P}_{\mathrm{m}}(\Upsilon)$ satisfies $S(\gamma) = 0$ for $\gamma > \Upsilon$, we obtain

$$\begin{aligned}
&\inf_{\substack{\{S,\beta_\gamma\}\colon \gamma\leq\Upsilon,\\ \int \beta_\gamma S(\mathrm{d}\gamma)=\frac{1}{M}}} \left\{\int f\big(\beta_\gamma,\gamma\big) S(\mathrm{d}\gamma)\right\} \\
&\qquad \geq \inf_{\substack{\{S,\beta_\gamma\}\colon \gamma\leq\Upsilon,\\ \int \beta_\gamma S(\mathrm{d}\gamma)=\frac{1}{M}}} \left\{\int f\big(\beta_\gamma,\Upsilon\big) S(\mathrm{d}\gamma)\right\} && (43)\\
&\qquad \geq f\big(\tfrac{1}{M},\Upsilon\big), && (44)
\end{aligned}$$

where in (44) we used that the function $f(\beta,\Upsilon)$ is convex with respect to $\beta$ (Lemma 3); hence, by Jensen's inequality and using the constraint $\int \beta_\gamma S(\mathrm{d}\gamma) = \frac{1}{M}$, it yields

$$\int f\big(\beta_\gamma,\Upsilon\big) S(\mathrm{d}\gamma) \geq f\big(\textstyle\int \beta_\gamma S(\mathrm{d}\gamma),\Upsilon\big) = f\big(\frac{1}{M},\Upsilon\big). \tag{45}$$

Then, using (27), (40), (44), and since the function $f\big(\frac{1}{M},\Upsilon\big) = \alpha_{\frac{1}{M}}\big(\varphi^n_{\sqrt{\Upsilon},\sigma}, \varphi^n_{0,\theta}\big)$, the result follows.

## V. Lower Bounds for Average Power Constraints

We now turn our attention to codes satisfying an average power limitation. To this end, we first introduce some useful concepts of convex analysis. The Legendre-Fenchel (LF) transform of a function $g$ is

$$g^*(b) = \max_{a\in\mathcal{A}} \big\{\langle a,b\rangle - g(a)\big\}, \tag{46}$$

where $\mathcal{A}$ is the domain of the function $g$ and $\langle a, b\rangle$ denotes the interior product between $a$ and $b$.

The function $g^*$ is usually referred to as Fenchel's conjugate (or convex conjugate) of $g$. If $g$ is a convex function with closed domain, applying the LF transform twice recovers the original function, *i.e.*, $g^{**} = g$. If $g$ is not convex, applying the LF transform twice returns the lower convex envelope of $g$, which is defined as the largest lower semi-continuous convex function majorized by $g$.

In our problem, for the 2-dimensional function $f(\beta,\gamma) = \alpha_\beta\big(\varphi^n_{\sqrt{\gamma},\sigma}, \varphi^n_{0,\theta}\big)$ with domain $\beta \in [0,1]$ and $\gamma \geq 0$, we define the lower convex envelope

$$\underline{f}(\beta,\gamma) \triangleq f^{**}(\beta,\gamma), \tag{47}$$

and note that $\underline{f}(\beta,\gamma) \leq f(\beta,\gamma)$.

The lower convex envelope (47) with certain parameters is a lower bound to the error probability in the average power constraint setting, as the next result shows.

*Theorem 4 (Converse, average power constraint):* Let $\theta > \sigma$, $n \geq 1$. The error probability of an average-power constrained code satisfies

$$\epsilon^\star_{\mathrm{a}}(n, M, \Upsilon) \geq \underline{f}\big(\tfrac{1}{M}, \Upsilon\big), \tag{48}$$

where $\underline{f}(\beta,\gamma)$ denotes the lower convex envelope of the function $f(\beta,\gamma) = \alpha_\beta\big(\varphi^n_{\sqrt{\gamma},\sigma}, \varphi^n_{0,\theta}\big)$ (see (47)).

*Proof:* We start by considering the general meta-converse bound in (10) where $\mathcal{P}$ is the set of distributions satisfying an average power constraint, i.e., $\mathcal{P} = \mathcal{P}_{\mathrm{a}}(\Upsilon)$ with

$$\mathcal{P}_{\mathrm{a}}(\Upsilon) \triangleq \Big\{ \boldsymbol{X} \sim P_{\boldsymbol{X}} \;\Big|\; \mathrm{E}\big[\|\boldsymbol{X}\|^2\big] \leq n\Upsilon \Big\}. \tag{49}$$

Proceeding analogously as in (39)-(40), but with the average power constraint $\int \gamma S(\mathrm{d}\gamma) \leq \Upsilon$ instead of the maximal power constraint, it follows that

$$\inf_{P\in\mathcal{P}_{\mathrm{a}}(\Upsilon)} \Big\{\alpha_{\frac{1}{M}}\big(PW, P\times Q\big)\Big\} = \inf_{\substack{\{S,\beta_\gamma\}:\\ \int \gamma S(\mathrm{d}\gamma)\leq\Upsilon\\ \int \beta_\gamma S(\mathrm{d}\gamma)=\frac{1}{M}}} \left\{\int \alpha_{\beta_\gamma}\big(\varphi^n_{\sqrt{\gamma},\sigma}, \varphi^n_{0,\theta}\big) S(\mathrm{d}\gamma)\right\}. \tag{50}$$

The function $f(\beta,\gamma) = \alpha_\beta\big(\varphi^n_{\sqrt{\gamma},\sigma}, \varphi^n_{0,\theta}\big)$ is non-increasing in $\gamma$ for any fixed $\beta \in [0,1]$ (see Lemma 3). Therefore,

$$\inf_{\substack{\{S,\beta_\gamma\}:\\ \int \gamma S(\mathrm{d}\gamma)\leq\Upsilon\\ \int \beta_\gamma S(\mathrm{d}\gamma)=\frac{1}{M}}} \left\{\int f\big(\beta_\gamma,\gamma\big) S(\mathrm{d}\gamma)\right\} \geq \inf_{\substack{\{S,\beta_\gamma\}:\\ \int \gamma S(\mathrm{d}\gamma)=\Upsilon\\ \int \beta_\gamma S(\mathrm{d}\gamma)=\frac{1}{M}}} \left\{\int f\big(\beta_\gamma,\gamma\big) S(\mathrm{d}\gamma)\right\}. \tag{51}$$

In words, the average power constraint can be assumed to hold with equality as this restriction does not increase the bound. Using (51) and since $f(\beta,\gamma) \geq \underline{f}(\beta,\gamma)$, we lower-bound the RHS of (50) as

$$\begin{aligned}
&\inf_{\substack{\{S,\beta_\gamma\}:\\ \int \gamma S(\mathrm{d}\gamma)\leq\Upsilon\\ \int \beta_\gamma S(\mathrm{d}\gamma)=\frac{1}{M}}} \left\{\int f\big(\beta_\gamma,\gamma\big) S(\mathrm{d}\gamma)\right\} \\
&\qquad\geq \inf_{\substack{\{S,\beta_\gamma\}:\\ \int \gamma S(\mathrm{d}\gamma)=\Upsilon\\ \int \beta_\gamma S(\mathrm{d}\gamma)=\frac{1}{M}}} \left\{\int \underline{f}\big(\beta_\gamma,\gamma\big) S(\mathrm{d}\gamma)\right\} &(52)\\
&\qquad\geq \inf_{\substack{\{S,\beta_\gamma\}:\\ \int \gamma S(\mathrm{d}\gamma)=\Upsilon\\ \int \beta_\gamma S(\mathrm{d}\gamma)=\frac{1}{M}}} \left\{\underline{f}\big(\textstyle\int \beta_\gamma S(\mathrm{d}\gamma), \int \gamma S(\mathrm{d}\gamma)\big)\right\} &(53)\\
&\qquad= \underline{f}\big(\tfrac{1}{M}, \Upsilon\big), &(54)
\end{aligned}$$

where (53) follows from applying Jensen's inequality since $\underline{f}(\cdot)$ is convex by definition, and (54) holds since, given the constraints $\int \beta_\gamma S(\mathrm{d}\gamma) = \frac{1}{M}$ and $\int \gamma S(\mathrm{d}\gamma) = \Upsilon$, the objective of the optimization does not depend on $\{S, \beta_\gamma\}$.

The lower bound (48) then follows from combining (10), (50) and the inequalities (52)-(54). ∎

The function $\underline{f}(\beta,\gamma)$ can be evaluated numerically by considering a 2-dimensional grid on the parameters $(\beta,\gamma)$, computing $f(\beta,\gamma)$ over this grid, and obtaining the corresponding convex envelope. Nevertheless, sometimes $\underline{f}(\beta,\gamma) = f(\beta,\gamma)$ and these steps can be avoided, as the next result shows.

*Lemma 4:* Let $\sigma, \theta, \gamma$ and $n$, be fixed parameters satisfying $0 < \sigma < \theta$, $\gamma > 0$, and $n \geq 1$; and let $\delta = \theta^2 - \sigma^2$. For $t \geq 0$, we define the auxiliary functions

$$\xi_1(t) \triangleq Q_{\frac{n}{2}}\left(\sqrt{n\gamma}\frac{\sigma}{\delta}, \frac{t}{\sigma}\right) - Q_{\frac{n}{2}}\left(0, \sqrt{\big(\tfrac{t^2}{\sigma^2} - n\gamma\tfrac{\theta^2}{\delta^2}\big)_+}\right), \tag{55}$$

$$\xi_2(t) \triangleq \frac{\theta^n}{\sigma^n} e^{\frac{1}{2}\left(\frac{n\gamma}{\delta} - \frac{\delta t^2}{\sigma^2\theta^2}\right)} \left(Q_{\frac{n}{2}}\left(0, \sqrt{\big(\tfrac{t^2}{\theta^2} - n\gamma\tfrac{\sigma^2}{\delta^2}\big)_+}\right) - Q_{\frac{n}{2}}\left(\sqrt{n\gamma}\frac{\theta}{\delta}, \frac{t}{\theta}\right)\right), \tag{56}$$

$$\xi_3(t) \triangleq \frac{n\gamma}{2\delta}\left(\frac{t\delta}{\sigma^2\sqrt{n\gamma}}\right)^{\frac{n}{2}} e^{-\frac{1}{2}\left(n\gamma\frac{\sigma^2}{\delta^2} + \frac{t^2}{\sigma^2}\right)} I_{\frac{n}{2}}\left(\sqrt{n\gamma}\frac{t}{\delta}\right), \tag{57}$$

where $(a)_+ = \max(0,a)$, $Q_m(a,b)$ is the Marcum $Q$-function and $I_m(\cdot)$ denotes the $m$-th order modified Bessel function of the first kind. Let $t_0$ be the solution to the implicit equation

$$\xi_1(t_0) + \xi_2(t_0) + \xi_3(t_0) = 0. \tag{58}$$

Then, for any $\beta$ satisfying $\big(1 - Q_{\frac{n}{2}}\big(\sqrt{n\gamma}\theta/\delta,\, t_0/\theta\big)\big) \leq \beta \leq 1$, it holds that

$$\underline{f}(\beta,\gamma) = f(\beta,\gamma). \tag{59}$$

*Proof:* See Appendix B. ∎

Combining Theorem 4 and Lemma 4 we obtain a simple lower bound on the error probability of any code satisfying an average-power constraint, provided that its cardinality is below a certain threshold.

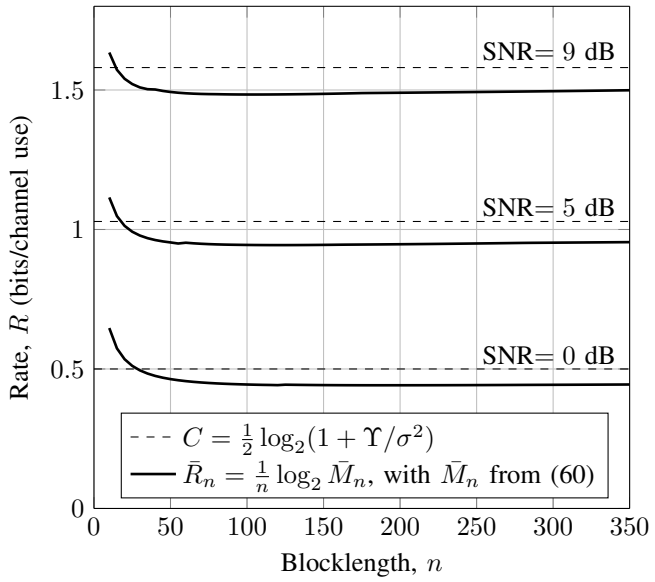



Fig. 2: Condition from Corollary 2 for AWGN channels with SNR of 0 dB, 5 dB and 9 dB compared with the channel capacity $C$.

*Corollary 2:* Let $\sigma, \theta > 0$ and $n \geq 1$, be fixed parameters, and $\delta = \theta^2 - \sigma^2$. Let $t_0$ be the solution to the implicit equation (58) with $\gamma = \Upsilon$ and define

$$\bar{M}_n \triangleq \left(1 - Q_{\frac{n}{2}}\left(\frac{\sqrt{n\Upsilon}\theta}{\delta}, \frac{t_0}{\theta}\right)\right)^{-1}. \tag{60}$$

Then, for any code $\mathcal{C} \in \mathcal{F}_{\mathrm{a}}(n, M, \Upsilon)$ with $M \leq \bar{M}_n$,

$$\epsilon^{\star}_{\mathrm{a}}(n, M, \Upsilon) \geq \alpha_{\frac{1}{M}}\left(\varphi^n_{\sqrt{\Upsilon},\sigma}, \varphi^n_{0,\theta}\right). \tag{61}$$

Figure 2 shows the condition $M \leq \bar{M}_n$ in Corollary 2 as an upper bound on the transmission rate of the system, defined as $R = \frac{1}{n}\log_2 M$. In this example, we let $\theta^2 = \Upsilon + \sigma^2$ and consider three different values of the signal-to-noise ratio, SNR=$10\log_{10}\frac{\Upsilon}{\sigma^2}$. The channel capacity $C = \frac{1}{2}\log_2\left(1 + \frac{\Upsilon}{\sigma^2}\right)$ for each of the SNRs is also included for reference. According to Corollary 2, for any code satisfying an average power constraint $\Upsilon$ with blocklength $n$ and rate $R \leq \bar{R}_n$, the lower bound (61) holds. Then, we can see in Fig. 2 that this condition holds except for rates very close to and above capacity (provided that the blocklegth $n$ is sufficiently large). For transmission rates above $\bar{R}_n$ in the plot, (61) does not apply and the lower convex envelope from Theorem 4 needs to be used instead.

The condition stated in Corollary 2 agrees with previous results in the literature. The asymptotic analysis of the RHS of (61) yields a strong converse behavior for rates above capacity [5, Th. 74]. That is, for any rate above capacity the RHS of (61) tends to one as the blocklength increases. However, [5, Th. 77] shows that for the AWGN channel under an average power constraint there exist no strong converse behavior in this regime. We conclude that the bound (61) cannot hold above capacity. The condition (60) thus corresponds to the finite-length equivalent of this asymptotic result.

### A. Optimal input distribution

The derivation of Theorem 4 and Corollary 2 can be used to characterize the structure of the optimal input distribution. Indeed, it follows that the tightest meta-converse bound for the auxiliary distribution $Q$ given in (17) corresponds precisely to the RHS of (48), i.e.,

$$\inf_{P \in \mathcal{P}_{\mathrm{a}}(\Upsilon)} \left\{\alpha_{\frac{1}{M}}\left(PW, P \times Q\right)\right\} = \underline{f}\left(\tfrac{1}{M}, \Upsilon\right). \tag{62}$$

To show this identity, we first note that the RHS of (62) corresponds to the value of the convex envelope $\underline{f}$ at the point $\left(\frac{1}{M}, \Upsilon\right)$. Using Carathéodory's theorem, it follows that any point on $\underline{f}$ can be written as a convex combination of (at most) 4 points in $f$.[4] Let us denote these 4 points as $\left(\beta_i, \gamma_i\right)$, $i = 1, \ldots, 4$. Then, for some $\lambda_i \geq 0$, $i = 1, \ldots, 4$, such that $\sum_{i=1}^4 \lambda_i = 1$, the following identities hold

$$\underline{f}\left(\tfrac{1}{M}, \Upsilon\right) = \sum_{i=1}^4 \lambda_i f\left(\beta_i, \gamma_i\right), \tag{63}$$

$$\frac{1}{M} = \sum_{i=1}^4 \lambda_i \beta_i, \tag{64}$$

$$\Upsilon = \sum_{i=1}^4 \lambda_i \gamma_i. \tag{65}$$

Let $S$ be the probability distribution that sets its mass points at $\gamma_i$, $i = 1, \ldots, 4$, with probabilities $S(\gamma_i) = \lambda_i$. Let $\beta_{\gamma_i} = \beta_i$, $i = 1, \ldots, 4$. This choice of $\{S, \beta_\gamma\}$ satisfies the constraints of the left-hand side of (52). Moreover, for this choice of $\{S, \beta_\gamma\}$ the left-hand side of (52) becomes $\sum_{i=1}^4 \lambda_i f\left(\beta_i, \gamma_i\right) = \underline{f}\left(\frac{1}{M}, \Upsilon\right)$ and therefore the inequality chain in (52)-(54) holds with equality. Also, increasing the power limit $\Upsilon$ yields a strictly lower error probability, and therefore (51) holds with equality. Then, according to this reasoning and using (50) we conclude that (62) holds.

To characterize the structure of the input distribution optimizing the left-hand side of (62), we recall (38). We conclude that the input distribution $P$ optimizing the left-hand side of (50) concentrates its mass on (at most) 4 spherical shells with squared radius $n\gamma_i$, $i = 1, \ldots, 4$. The probability of each of these shells is precisely $S(\gamma_i) = \lambda_i$ and the optimizing input distribution is uniformly distributed on the surface of each of the shells [14, Sec. VI.F].

In the discussion above we have not used specific properties of $f(\beta, \gamma)$ beyond that it is a function of two real variables and that it is strictly decreasing in $\gamma$. Studying the form of the function $f(\beta, \gamma)$ and its convex envelope we can describe the structure of the optimal input distribution more precisely.

The convex envelope $\underline{f}(\beta, \gamma)$ is analyzed in the proof of Lemma 4 in Appendix B. Let $(\beta_0, \gamma_0)$ denote the boundary described in Lemma 4, i.e., the points $(\beta_0, \gamma_0)$ satisfying

$$\beta_0 = 1 - Q_{\frac{n}{2}}\left(\sqrt{n\gamma_0}\theta/\delta,\, t_0/\theta\right). \tag{66}$$

where $t_0$ is the solution of (58) for $\gamma = \gamma_0$. This boundary defines two regions:

[4] For a 2-dimensional function its epigraph is a 3-dimensional set. Therefore, Carathéodory's theorem implies that at most $3 + 1$ points are needed to construct the convex hull of the epigraph, which corresponds to the convex envelope of the original function.

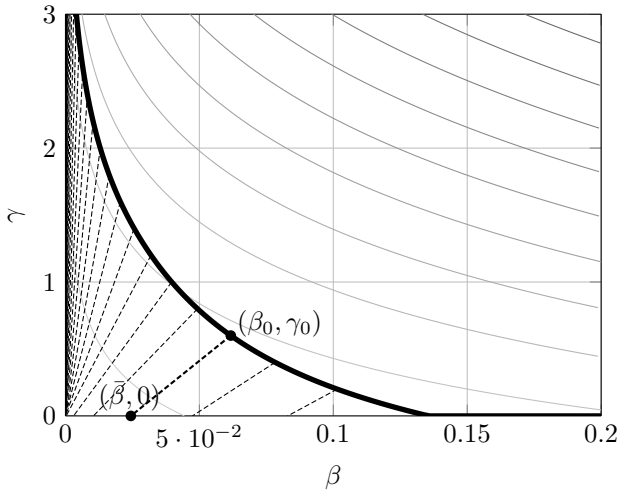



Fig. 3: In gray, the level curves of $f(\beta,\gamma)$ with parameters $n = 6$, $\sigma^2 = 1$, $\theta^2 = 2$. The bold line corresponds to the boundary (66). The dashed lines show the convex combinations that yield $\underline{f}(\beta,\gamma)$ below the boundary.

*1) Above the boundary (66):* This region corresponds to $\gamma \geq 0$ and $\big(1 - Q_{\frac{n}{2}}\big(\sqrt{n\gamma}\theta/\delta,\, t_0/\theta\big)\big) \leq \beta \leq 1$ where $t_0$ is the solution of (58). From Lemma 4 we know that, for this region, $\underline{f} = f$ and hence $\underline{f}$ is the convex combination of a single point of $f$.

*2) Below the boundary (66):* For $\gamma \geq 0$ and $0 \leq \beta < \big(1 - Q_{\frac{n}{2}}\big(\sqrt{n\gamma}\theta/\delta,\, t_0/\theta\big)\big)$, $\underline{f}$ and $f$ do not coincide. Instead, the convex envelope $\underline{f}$ evaluated at $(\beta,\gamma)$ corresponds to a convex combination of the function $f$ at the points $(\beta_0,\gamma_0)$ and $(\bar{\beta},0)$, where $(\beta_0,\gamma_0)$ satisfies (66), and $\bar{\beta} = 1 - Q_{\frac{n}{2}}\big(0,\, \bar{t}_\star/\theta\big)$ for $\bar{t}_\star$ in (185) (for details, see Appendix B).

An example of the construction of the convex envelope $\underline{f}$ is shown in Fig. 3. This figure shows in gray the level lines of $f\big(\beta,\gamma\big)$; the bold line corresponds to the boundary (66); and with dashed lines we show the convex combinations between $\big(\beta_0,\gamma_0\big)$ and $\big(\bar{\beta},0\big)$ that yield $\underline{f}(\beta,\gamma)$ for different values of $\gamma_0$. Above the boundary, the convex envelope $\underline{f}$ coincides with $f$, and therefore the input distribution $P$ optimizing the left-hand side of (62) is uniform over a spherical shell centered at the origin and with squared radius $n\Upsilon$. Below the boundary, the convex envelope $\underline{f}$ corresponds to the convex combination of two points of $f$ and, as a result, the optimal input distribution $P$ corresponds to a mass point at the origin (corresponding to the axis $\gamma = 0$) and a spherical shell centered at the origin and with squared radius $n\gamma_0$, where $\gamma_0 \geq \Upsilon$. The probability mass of each of these two components depends on the parameters of the system.

## VI. Computation of $f(\beta,\gamma) = \alpha_\beta\big(\varphi^n_{\sqrt{\gamma},\sigma}, \varphi^n_{0,\theta}\big)$

In the previous sections we showed that both the function $f(\beta,\gamma)$ and its convex envelope $\underline{f}(\beta,\gamma)$ yield lower bounds to the error probability for Gaussian channels under different power constraints. In this section we provide several tools that can be used in the numerical evaluation of these functions.

### A. Exact computation of $f(\beta,\gamma)$

Proposition 1 in Appendix A provides a parametric formulation of the function $f(\beta,\gamma)$. A non-parametric expression for $f(\beta,\gamma)$ can be obtained by combining Proposition 1 and [24, Lem. 1] as shown next.

*Theorem 5 (Non-parametric formulation):* Let $\sigma,\theta > 0$ and $n \geq 1$, be fixed parameters. Then, it holds that

$$f(\beta,\gamma) = \max_{t\geq 0}\Bigg\{ Q_{\frac{n}{2}}\left(\sqrt{n\gamma}\frac{\sigma}{\delta}, \frac{t}{\sigma}\right) + \frac{\theta^n}{\sigma^n} e^{\frac{1}{2}\left(\frac{n\gamma}{\delta} - \frac{\delta t^2}{\sigma^2\theta^2}\right)} \times \left(1 - \beta - Q_{\frac{n}{2}}\left(\sqrt{n\gamma}\frac{\theta}{\delta}, \frac{t}{\theta}\right)\right)\Bigg\}. \tag{67}$$

*Proof:* Using [24, Lem. 1], we obtain the following alternative expression for $\alpha_\beta\big(\varphi^n_{\sqrt{\gamma},\sigma}, \varphi^n_{0,\theta}\big)$:

$$\alpha_\beta\big(\varphi^n_{\sqrt{\gamma},\sigma}, \varphi^n_{0,\theta}\big) = \max_{t'}\Big\{\Pr\big[j(\boldsymbol{Y}_0) \leq t'\big] + e^{t'}\big(\Pr\big[j(\boldsymbol{Y}_1) > t'\big] - \beta\big)\Big\}, \tag{68}$$

where $\boldsymbol{Y}_0 \sim \varphi^n_{\sqrt{\gamma},\sigma}$ and $\boldsymbol{Y}_1 \sim \varphi^n_{0,\theta}$ and where $j(\boldsymbol{y})$ denotes the log-likelihood ratio

$$j(\boldsymbol{y}) = \log \frac{\varphi^n_{\sqrt{\gamma},\sigma}(\boldsymbol{y})}{\varphi^n_{0,\theta}(\boldsymbol{y})} \tag{69}$$

$$= n\log\frac{\theta}{\sigma} - \frac{1}{2}\sum_{i=1}^{n}\frac{\theta^2(y_i - \sqrt{\gamma})^2 - \sigma^2 y_i^2}{\sigma^2\theta^2}. \tag{70}$$

Following analogous steps as in the proof of Proposition 1 in Appendix A, we obtain that

$$\Pr\big[j(\boldsymbol{Y}_0) \leq t'\big] = Q_{\frac{n}{2}}\left(\sqrt{n\gamma}\frac{\sigma}{\delta}, \frac{t}{\sigma}\right), \tag{71}$$

$$\Pr\big[j(\boldsymbol{Y}_1) > t'\big] = 1 - Q_{\frac{n}{2}}\left(\sqrt{n\gamma}\frac{\theta}{\delta}, \frac{t}{\theta}\right), \tag{72}$$

where the variables $t'$ and $t$ are related according to (131), c.f., $e^{t'} = \frac{\theta^n}{\sigma^n} e^{\frac{1}{2}\left(\frac{n\gamma}{\delta} - \frac{\delta t^2}{\sigma^2\theta^2}\right)}$. Then, the result follows from (68), (71) and (72) via the change of variable $t' \leftrightarrow t$. ∎

The formulation in Theorem 5 allows to obtain simple lower bounds on $f(\beta,\gamma)$ by fixing the value of $t$ in (67). Alternatively, Verdú-Han-type lower bounds can be obtained by using that $Q_{\frac{n}{2}}\big(\sqrt{n\gamma}\frac{\theta}{\delta}, \frac{t}{\theta}\big) \leq 1$, e.g.,

$$f(\beta,\gamma) \geq \max_{t\geq 0}\left\{ Q_{\frac{n}{2}}\left(\sqrt{n\gamma}\frac{\sigma}{\delta}, \frac{t}{\sigma}\right) - \frac{\theta^n}{\sigma^n} e^{\frac{1}{2}\left(\frac{n\gamma}{\delta} - \frac{\delta t^2}{\sigma^2\theta^2}\right)}\beta\right\}. \tag{73}$$

In order to evaluate (67) in Theorem 5 we need to solve a maximization over the scalar parameter $t \geq 0$ with an objective involving two Marcum-$Q$ functions. The computation of the bounds from Theorems 2, 3 and 4 for a fixed rate $R \triangleq \frac{1}{n}\log_2 M$, implies that the parameter $\beta = 2^{-nR}$ decreases exponentially with the blocklength $n$. Then, traditional Taylor series expansions of the Marcum-$Q$ function fail to achieve the required precision even for moderate values of $n$. In this regime, the following expansion yields an accurate approximation of $f(\beta,\gamma)$.

### B. Saddlepoint expansion of $f(\beta,\gamma)$

We define the (scalar) information density

$$j(y) \triangleq \log \frac{\varphi_{\sqrt{\gamma},\sigma}(y)}{\varphi_{0,\theta}(y)} = \log\frac{\theta}{\sigma} - \frac{1}{2}\frac{\theta^2(y-\sqrt{\gamma})^2 - \sigma^2 y^2}{\sigma^2\theta^2}, \tag{74}$$

and we consider the cumulant generating function of the random variable $j(Y)$, $Y \sim \varphi_{0,\theta}$, given by

$$\kappa_\gamma(s) \triangleq \log \int_{-\infty}^{\infty} \frac{\varphi_{\sqrt{\gamma},\sigma}(y)^s}{\varphi_{0,\theta}(y)^{s-1}}\,\mathrm{d}y \tag{75}$$

$$= \gamma\frac{s(s-1)}{2\eta(s)} + \log\frac{\theta^s\sigma^{1-s}}{\sqrt{\eta(s)}}, \tag{76}$$

where we defined $\eta(s) \triangleq s\theta^2 + (1-s)\sigma^2$. The first three derivatives of (76) with respect to $s$ are, respectively,

$$\kappa'_\gamma(s) = \gamma\frac{s^2\theta^2 - (1-s)^2\sigma^2}{2\eta(s)^2} - \frac{\theta^2-\sigma^2}{2\eta(s)} + \log\frac{\theta}{\sigma}, \tag{77}$$

$$\kappa''_\gamma(s) = \gamma\frac{\theta^2\sigma^2}{\eta(s)^3} + \frac{(\theta^2-\sigma^2)^2}{2\eta(s)^2}, \tag{78}$$

$$\kappa'''_\gamma(s) = -\left(3\gamma\frac{\theta^2\sigma^2(\theta^2-\sigma^2)}{\eta(s)^4} + \frac{(\theta^2-\sigma^2)^3}{\eta(s)^3}\right). \tag{79}$$

*Theorem 6 (Saddlepoint expansion):* We let $\sigma,\theta > 0$ and $n \geq 1$, be fixed parameters. Then,

$$f(\beta,\gamma) = \max_s\Big\{\big(a(s,\gamma)+b(s,\gamma)\big)e^{n(\kappa_\gamma(s)+(1-s)\kappa'_\gamma(s))} + \mathbb{1}[s>1] + \big(\mathbb{1}[s<0]-\beta\big)e^{n\kappa'_\gamma(s)}\Big\}, \tag{80}$$

where $\mathbb{1}[\cdot]$ denotes the indicator function and, for $\lambda_a(s) \triangleq |1-s|\sqrt{n\kappa''_\gamma(s)}$ and $\lambda_b(s) \triangleq |s|\sqrt{n\kappa''_\gamma(s)}$,

$$a(s,\gamma) = \operatorname{sgn}(1-s)\left(\Psi\big(\lambda_a(s)\big) + \frac{n(s-1)^3}{6} \times \left(\frac{\lambda_a(s)^{-1}-\lambda_a(s)^{-3}}{\sqrt{2\pi}} - \Psi\big(\lambda_a(s)\big)\right)\kappa'''_\gamma(s)\right) + o\big(n^{-\frac{1}{2}}\big), \tag{81}$$

$$b(s,\gamma) = \operatorname{sgn}(s)\left(\Psi\big(\lambda_b(s)\big) + \frac{ns^3}{6} \times \left(\frac{\lambda_b(s)^{-1}-\lambda_b(s)^{-3}}{\sqrt{2\pi}} - \Psi\big(\lambda_b(s)\big)\right)\kappa'''_\gamma(s)\right) + o\big(n^{-\frac{1}{2}}\big). \tag{82}$$

Here, $\operatorname{sgn}(\cdot)$ denotes the sign function, defined as $\operatorname{sgn}(s) = -1$ for $s < 0$ and $\operatorname{sgn}(s) = 1$ otherwise; the function $\Psi(\lambda)$ is defined as $\Psi(\lambda) \triangleq \mathsf{Q}(|\lambda|)e^{\frac{\lambda^2}{2}}$ where $\mathsf{Q}(\cdot)$ is the Gaussian Q-function; and $o\big(g(n)\big)$ summarizes the terms that approach zero faster than $g(n)$, *i.e.*, $\lim_{n\to\infty}\frac{o(g(n))}{g(n)} = 0$.

*Proof:* The proof follows the lines of [25, Th. 2] with a more refined expansion of $a(s,\gamma)$ and $b(s,\gamma)$.

We consider a sequence of i.i.d. non-lattice real-valued random variables with positive variance, $\{Z_\ell\}_{\ell=1}^n$, and we define their mean as

$$\bar{Z}_n \triangleq \frac{1}{n}\sum_{\ell=0}^{n} Z_\ell. \tag{83}$$

Let $\kappa_Z(s) \triangleq \log \mathrm{E}\big[e^{sZ_\ell}\big]$ denote the cumulant generating function of $Z_\ell$, and let $\kappa'_Z(s)$, $\kappa''_Z(s)$ and $\kappa'''_Z(s)$ denote its 1st, 2nd and 3rd derivatives with respect to $s$, respectively.

Assume that there exists $s$ in the region of convergence of $\kappa_Z(s)$ such that $\kappa'_Z(s) = t$ and $t \geq \mathrm{E}\big[\bar{Z}_n\big]$. Then, the tail probability $\Pr\big[\bar{Z}_n \geq t\big]$ satisfies [26, Prop. 1, Part 1]

$$\Pr\big[\bar{Z}_n \geq t\big] = e^{n(\kappa_Z(s)-s\kappa'_Z(s))} \times \left(\Psi(\lambda_{Z,s}) + \frac{ns^3}{6}\left(\frac{\lambda_{Z,s}^{-1}-\lambda_{Z,s}^{-3}}{\sqrt{2\pi}} - \Psi(\lambda_{Z,s})\right)\kappa'''_Z(s) + o\big(n^{-\frac{1}{2}}\big)\right), \tag{84}$$

where $\lambda_{Z,s} \triangleq |s|\sqrt{n\kappa''_Z(s)}$ and the error term $o\big(n^{-\frac{1}{2}}\big)$ holds uniformly in $s$.

We now apply this expansion to the probability terms $\Pr\big[j(\boldsymbol{Y}_0) \leq t'\big]$ and $\Pr\big[j(\boldsymbol{Y}_1) > t'\big]$ appearing in (68). To this end, we consider the random variables $Z_0 = -j(Y_0)$, $Y_0 \sim \varphi_{\sqrt{\gamma},\sigma}$, and $Z_1 = j(Y_1)$, $Y_1 \sim \varphi_{0,\theta}$, with $j(y)$ defined in (74). The cumulant generating functions of the random variables $Z_0$ and $Z_1$ are, respectively, given by

$$\kappa_{Z_0}(s) = \log \mathrm{E}\left[\left(\frac{\varphi_{\sqrt{\gamma},\sigma}(Y_0)}{\varphi_{0,\theta}(Y_0)}\right)^{-s}\right] = \kappa_\gamma(1-s), \tag{85}$$

$$\kappa_{Z_1}(s) = \log \mathrm{E}\left[\left(\frac{\varphi_{\sqrt{\gamma},\sigma}(Y_1)}{\varphi_{0,\theta}(Y_1)}\right)^{s}\right] = \kappa_\gamma(s), \tag{86}$$

for the function $\kappa_\gamma(s)$ defined in (76). The fact that the cumulant generating functions $\kappa_{Z_0}(s)$ and $\kappa_{Z_1}(s)$ are shifted and mirrored versions of each other will allow us to simplify the resulting expression.

The expansion in (84) requires that the threshold $t$ is above the average value $\mathrm{E}\big[\bar{Z}_n\big] = \mathrm{E}\big[Z_\ell\big]$. That is, this expansion is only accurate for evaluating the tail of the probability distribution. Whenever $t < \mathrm{E}\big[\bar{Z}_n\big]$, we shall use that

$$\Pr\big[\bar{Z}_n \geq t\big] = 1 - \Pr\big[\bar{Z}_n < t\big] \tag{87}$$

$$= 1 - \Pr\big[-\bar{Z}_n > -t\big]. \tag{88}$$

For non-lattice distributions, the expansion (84) coincides with that of $\Pr\big[\bar{Z}_n > t\big]$ and therefore, it can be used to estimate $\Pr\big[-\bar{Z}_n > -t\big]$. Moreover, given the mapping $\kappa'_Z(s) = t$, it can be checked that the condition $t \geq \mathrm{E}\big[\bar{Z}_n\big]$ corresponds to $s \geq 0$. Similarly, for $t < \mathrm{E}\big[\bar{Z}_n\big]$ and the mapping $\kappa'_Z(s) = t$, we obtain $s < 0$.

We use (84) for $\bar{Z}_{0,n} \triangleq -\frac{1}{n}j(\boldsymbol{Y}_0) = -\frac{1}{n}\sum_{\ell=1}^n j(Y_{0,\ell})$, and apply (87)-(88) whenever $t < \mathrm{E}\big[\bar{Z}_{0,n}\big]$, to obtain

$$\Pr\big[\bar{Z}_{0,n} \geq t\big] = \mathbb{1}[\bar{s}<0] + \operatorname{sgn}(\bar{s})e^{n(\kappa_{Z_0}(\bar{s})-\bar{s}\kappa'_{Z_0}(\bar{s}))} \times \left(\Psi(\lambda_{Z_0,\bar{s}}) + \frac{n\bar{s}^3}{6}\left(\frac{\lambda_{Z_0,\bar{s}}^{-1}-\lambda_{Z_0,\bar{s}}^{-3}}{\sqrt{2\pi}} - \Psi(\lambda_{Z_0,\bar{s}})\right)\kappa'''_{Z_0}(\bar{s}) + o\big(n^{-\frac{1}{2}}\big)\right), \tag{89}$$

where the value of $\bar{s}$ satisfies $\kappa'_{Z_0}(\bar{s}) = t$.

We consider the change of variable $\bar{s} \leftrightarrow s$ with $\bar{s} = 1 - s$ and $s = 1 - \bar{s}$. For this change of variables, from (85) we obtain the identities

$$\kappa_{Z_0}(\bar{s}) = \kappa_\gamma(1-\bar{s}) = \kappa_\gamma(s), \tag{90}$$
$$\kappa'_{Z_0}(\bar{s}) = -\kappa'_\gamma(1-\bar{s}) = -\kappa'_\gamma(s), \tag{91}$$
$$\kappa''_{Z_0}(\bar{s}) = \kappa''_\gamma(1-\bar{s}) = \kappa''_\gamma(s), \tag{92}$$
$$\kappa'''_{Z_0}(\bar{s}) = -\kappa'''_\gamma(1-\bar{s}) = -\kappa'''_\gamma(s). \tag{93}$$

Then, using $\bar{s} = 1 - s$ and (90)-(93) in (89), it yields

$$\Pr\big[j(\boldsymbol{Y}_0) \le t'\big] = \mathbb{1}[s > 1] + a(s,\gamma) e^{n(\kappa_\gamma(s) + (1-s)\kappa'_\gamma(s))} \tag{94}$$

where $s$ satisfies $\kappa'_\gamma(s) = t'/n$.

Proceeding analogously for the random variable $\bar{Z}_{1,n} = \frac{1}{n} j(\boldsymbol{Y}_1) = \frac{1}{n}\sum_{\ell=1}^n j(Y_{1,\ell})$, using (84) and (86), we obtain

$$\Pr\big[j(\boldsymbol{Y}_1) > t'\big] = \mathbb{1}[s < 0] + b(s,\gamma) e^{n(\kappa_\gamma(s) - s\kappa'_\gamma(s))} \tag{95}$$

where $s$ satisfies again $\kappa'_\gamma(s) = t'/n$.

We replace (94) and (95) in (68) and change the optimization variable from $t'$ to $s$ acording to the relation $t' = n\kappa'_\gamma(s)$. Noting that $e^{t'}$ becomes $e^{n\kappa'_\gamma(s)}$, then (68) becomes

$$\begin{aligned}
&\alpha_\beta\big(\varphi^n_{\sqrt{\gamma},\sigma}, \varphi^n_{0,\theta}\big) \\
&= \max_s \Big\{ \mathbb{1}[s > 1] + a(s,\gamma) e^{n(\kappa_\gamma(s) + (1-s)\kappa'_\gamma(s))} \\
&\qquad + e^{n\kappa'_\gamma(s)} \big( \mathbb{1}[s < 0] + b(s,\gamma) e^{n(\kappa_\gamma(s) - s\kappa'_\gamma(s))} - \beta \big) \Big\}.
\end{aligned} \tag{96}$$

The result follows from (96) by reorganizing terms. ∎

*Remark:* The refined expressions from (81)-(82) are needed to obtain an error term $o\big(n^{-\frac{1}{2}}\big)$ that is uniform on $s$. For practical purposes, however, the function $f(\beta,\gamma)$ can be approximated using (80) from Theorem 6 with $a(s,\gamma)$ and $b(s,\gamma)$ replaced by the simpler expressions

$$\hat{a}(s,\gamma) \triangleq \operatorname{sgn}(1-s)\Psi\big(\lambda_a(s)\big), \tag{97}$$
$$\hat{b}(s,\gamma) \triangleq \operatorname{sgn}(s)\Psi\big(\lambda_b(s)\big), \tag{98}$$

respectively. This approximation yields accurate results for blocklengths as short as $n = 20$ (see [25] for details), and we still obtain an approximation error of order $o\big(n^{-\frac{1}{2}}\big)$ for values of $s$ satisfying $s_0 \le s \le s_1$ for any $s_0 > 0$ and $s_1 < 1$, i.e., for $s$ bounded away from 0 and 1.

### C. Exponent-achieving output distribution

Often, the variance of the auxiliary distribution in Theorems 2, 3 and 4 is chosen to be the variance of the capacity-achieving output distribution, $\theta^2 = \Upsilon + \sigma^2$. While this choice of $\theta^2$ is adequate for rates approaching the capacity of the channel, it does not attain the sphere-packing exponent in general [7]. An auxiliary distribution that yields the right exponential behavior is the exponent-achieving output distribution.[5]

[5]If we restrict the auxiliary output distribution to be product, the exponent-achieving output distribution is unique in the sense that the hypothesis-testing bound with this auxiliary distribution attains the sphere-packing exponent [7]. Nevertheless, using other non-product auxiliary distributions can yield the same exponential behavior. One example is the optimizing distribution in (10) for the AWGN channel under an equal power constraint. As discussed in Section III-B this distribution recovers Shannon cone-packing bound with the sphere-packing exponent [14, Sec. VI.F].

The exponent-achieving output distribution for memoryless channels channels under input constraints was studied in [11], [27], among other works., The exponent-achieving output distribution for the power-constrained AWGN channel corresponds to be the zero-mean i.i.d. Gaussian distribution defined in (17), but with variance $\theta^2 = \tilde{\theta}_s^2$ where (see [27, Sec. 6, Example 4], [11, Sec. 4, Example 1])

$$\tilde{\theta}_s^2 \triangleq \sigma^2 + \frac{\gamma}{2} - \frac{\sigma^2}{2s} + \sqrt{\left(\frac{\gamma}{2} - \frac{\sigma^2}{2s}\right)^2 + \gamma\sigma^2}. \tag{99}$$

Here, $\gamma = \Upsilon$ is the power constraint and the parameter $s$ is the result of optimizing the sphere-packing exponent [11, Eq. (10)] for a transmission rate $R = \frac{1}{n}\log M$:

$$E_{\mathrm{sp}}(R) \triangleq \sup_{s\in(0,1)} \left\{ \frac{1-s}{s}\big(C_{s,W,\Upsilon} - R\big) \right\}, \tag{100}$$

where the units of the rate $R$ are nats/channel use and $C_{s,W,\Upsilon}$ is the so-called Augustin capacity. For the power-constrained AWGN channel, $C_{s,W,\Upsilon}$ is given by [11, Eq. (62)]

$$C_{s,W,\Upsilon} = \begin{cases} \frac{s\Upsilon}{2\tilde{\eta}(s)} + \frac{1}{s-1}\log\frac{\tilde{\theta}_s^s \sigma^{1-s}}{\sqrt{\tilde{\eta}(s)}}, & s \ge 0,\ s \ne 1, \\ \frac{1}{2}\log\big(1 + \frac{\Upsilon}{\sigma^2}\big), & s = 1, \end{cases} \tag{101}$$

where we defined $\tilde{\eta}(s) \triangleq s\tilde{\theta}_s^2 + (1-s)\sigma^2$. Note that $C_{s,W,\Upsilon}$ with $s = 1$ recovers the usual notion of channel capacity $C = \frac{1}{2}\log\big(1 + \frac{\Upsilon}{\sigma^2}\big)$. For any rate $R$ below capacity, i.e. $R < C$, the optimal value of $s$ in (100) is given by [11, Eq. (72)]

$$s = \frac{e^{2R} - 1}{2}\left( \sqrt{1 + \frac{4\sigma^2}{\Upsilon}\frac{e^{2R}}{e^{2R}-1}} - 1 \right). \tag{102}$$

For transmission rates approaching the channel capacity $C$, the optimal value of $s$ in (102) tends to 1 and, therefore, $\theta_s^2 \to \theta_1^2 = \Upsilon + \sigma^2$. Hence, in this case $\varphi^n_{0,\theta}(\boldsymbol{y})$ becomes the capacity-achieving output distribution used in [13, Th. 41].

In principle, we can use the (asymptotically) optimal value of $s$ from (102) in the variance (99). Nevertheless, the saddlepoint expansion from Theorem 6 allows to introduce a dependence of $\theta^2$ with the auxiliary parameter $s$ without incurring in an extra computational cost, as we show next.

For $\tilde{\eta}(s) = s\tilde{\theta}_s^2 + (1-s)\sigma^2$ we define

$$\tilde{\kappa}_\gamma(s) \triangleq \gamma\frac{s(s-1)}{2\tilde{\eta}(s)} + \log\frac{\tilde{\theta}_s^s\sigma^{1-s}}{\sqrt{\tilde{\eta}(s)}}, \tag{103}$$
$$\tilde{\kappa}^{(1)}_\gamma(s) \triangleq \gamma\frac{s^2\tilde{\theta}_s^2 - (1-s)^2\sigma^2}{2\tilde{\eta}(s)^2} - \frac{\tilde{\theta}_s^2 - \sigma^2}{2\tilde{\eta}(s)} + \log\frac{\tilde{\theta}_s}{\sigma}, \tag{104}$$
$$\tilde{\kappa}^{(2)}_\gamma(s) \triangleq \gamma\frac{\tilde{\theta}_s^2\sigma^2}{\tilde{\eta}(s)^3} + \frac{(\tilde{\theta}_s^2 - \sigma^2)^2}{2\tilde{\eta}(s)^2}, \tag{105}$$
$$\tilde{\kappa}^{(3)}_\gamma(s) \triangleq -\left( 3\gamma\frac{\tilde{\theta}_s^2\sigma^2(\tilde{\theta}_s^2 - \sigma^2)}{\tilde{\eta}(s)^4} + \frac{(\tilde{\theta}_s^2 - \sigma^2)^3}{\tilde{\eta}(s)^3} \right). \tag{106}$$

The definitions (103)-(106) correspond to (76)-(79) after setting $\theta = \tilde{\theta}_s$. Note however that (104)-(106) do not coincide with the derivatives of (103) in general, since in obtaining (77)-(79) we have considered $\theta$ to be independent of $s$.

*Corollary 3 (Exponent-achieving saddlepoint expansion):* Let $\sigma > 0$ and $n \geq 1$, be fixed parameters. Then,

$$\max_{\theta \geq \sigma} \Big\{ \alpha_\beta\big(\varphi^n_{\sqrt{\gamma},\sigma}, \varphi^n_{0,\theta}\big) \Big\} \geq \tilde{f}(\beta,\gamma), \tag{107}$$

where the function $\tilde{f}(\beta,\gamma)$ is defined as

$$\tilde{f}(\beta,\gamma) \triangleq \max_s \Big\{ \big(\tilde{a}(s,\gamma) + \tilde{b}(s,\gamma)\big) e^{n(\tilde{\kappa}_\gamma(s) + (1-s)\tilde{\kappa}^{(1)}_\gamma(s))} + \mathbb{1}[s>1] + \big(\mathbb{1}[s<0] - \beta\big) e^{n\tilde{\kappa}^{(1)}_\gamma(s)} \Big\}, \tag{108}$$

$$\tilde{a}(s,\gamma) = \operatorname{sgn}(1-s) \Bigg( \Psi\big(\tilde{\lambda}_a(s)\big) + \frac{n(s-1)^3}{6} \times \Bigg( \frac{\tilde{\lambda}_a(s)^{-1} - \tilde{\lambda}_a(s)^{-3}}{\sqrt{2\pi}} - \Psi\big(\tilde{\lambda}_a(s)\big) \Bigg) \tilde{\kappa}^{(3)}_\gamma(s) \Bigg) + o\big(n^{-\frac{1}{2}}\big), \tag{109}$$

$$\tilde{b}(s,\gamma) = \operatorname{sgn}(s) \Bigg( \Psi\big(\tilde{\lambda}_b(s)\big) + \frac{ns^3}{6} \times \Bigg( \frac{\tilde{\lambda}_b(s)^{-1} - \tilde{\lambda}_b(s)^{-3}}{\sqrt{2\pi}} - \Psi\big(\tilde{\lambda}_b(s)\big) \Bigg) \tilde{\kappa}^{(3)}_\gamma(s) \Bigg) + o\big(n^{-\frac{1}{2}}\big). \tag{110}$$

with $\tilde{\lambda}_a(s) \triangleq |1-s|\sqrt{n\tilde{\kappa}^{(2)}_\gamma(s)}$ and $\tilde{\lambda}_b(s) \triangleq |s|\sqrt{n\tilde{\kappa}^{(2)}_\gamma(s)}$.

*Proof:* We follow the steps in the proof of Theorem 6, but we let the variance $\theta^2$ depend on $s$. To this end, we note that [26, Prop. 1, Part 1] establishes the saddlepoint expansion for a family of non-lattice random variables parametrized by some parameter $\theta \in \Theta$.

Let $\{Z_\ell\}_{\ell=1}^n$ be a sequence of i.i.d. non-lattice real-valued random variables parametrized by $\theta \in \Theta$ and let $\bar{Z}_n$ denote their average (see (83)). Let $m_{Z,\theta}(s) \triangleq \mathrm{E}\big[e^{sZ_\ell}\big]$ denote the moment generating function of $Z_\ell$, and let $\kappa_{Z,\theta}(s) \triangleq \log m_{Z,\theta}(s)$ denote its cumulant generating function. Suppose that

$$\sup_{\theta\in\Theta} |m^{\mathrm{iv}}_{Z,\theta}(s)| < \infty \tag{111}$$

where $m^{\mathrm{iv}}_{Z,\theta}(s)$ denotes the 4th derivative of the moment generating function, and

$$\inf_{\theta\in\Theta} |\kappa''_{Z,\theta}(s)| > 0. \tag{112}$$

Then, if there exists $s$ in the region of convergence of $\kappa_{Z,\theta}(s)$ such that $\kappa'_{Z,\theta}(s) = t$, with $t \geq \mathrm{E}\big[\bar{Z}_n\big]$, the tail probability $\Pr\big[\bar{Z}_n \geq t\big]$ satisfies (84) with $\kappa_Z$ replaced by $\kappa_{Z,\theta}$ and with an error term $o\big(n^{-\frac{1}{2}}\big)$ that holds uniformly in $s$ and $\theta$.

Both the random variables $Z_0 = -j(Y_0)$, $Y_0 \sim \varphi_{\sqrt{\gamma},\sigma}$, and $Z_1 = j(Y_1)$, $Y_1 \sim \varphi_{0,\theta}$, with $j(y)$ defined in (74), are parametrized by the variance $\theta^2$. We recall that their cumulant generating functions (85)-(86) are shifted and mirrored versions of $\kappa_\gamma(s)$. Moreover, the 4th derivative of the moment generating function only depends on the cumulant generating function and its 4 first derivatives as

$$m^{\mathrm{iv}}_{Z,\theta}(s) = e^{\kappa_{Z,\theta}(s)} \Big( \big(\kappa'_{Z,\theta}(s)\big)^4 + 6\big(\kappa'_{Z,\theta}(s)\big)^2 \kappa''_{Z,\theta}(s) + 4\kappa'_{Z,\theta}(s)\kappa'''_{Z,\theta}(s) + 3\big(\kappa''_{Z,\theta}(s)\big)^2 + \kappa^{\mathrm{iv}}_{Z,\theta}(s) \Big). \tag{113}$$

Then, to apply [26, Prop. 1, Part 1], it suffices to show that $\kappa_\gamma(s)$, $\kappa'_\gamma(s)$, $\kappa''_\gamma(s)$, $\kappa'''_\gamma(s)$, and $\kappa^{\mathrm{iv}}_\gamma(s)$ are bounded for the range of interest of $\theta$. The cumulant generating function $\kappa_\gamma(s)$ and its 3 first derivatives are given in (76)-(79), and its 4th derivative is

$$\kappa^{\mathrm{iv}}_\gamma(s) = 12\gamma \frac{\theta^2\sigma^2(\theta^2-\sigma^2)^2}{\eta(s)^5} + 3\frac{(\theta^2-\sigma^2)^4}{\eta(s)^4}, \tag{114}$$

with $\eta(s) = s\theta^2 + (1-s)\sigma^2$. We define the interval $\Theta = \big\{\theta \,|\, \sigma < \theta \leq \bar{\theta}\big\}$ for an arbitrary $\bar{\theta} < \infty$. For any $\theta \in \Theta$, (76)-(79) and (114) are bounded. Therefore, the conditions of [26, Prop. 1, Part 1] are satisfied.

Following analogous steps as in the proof of Theorem 6, we conclude that (96) holds with the error terms $o\big(n^{-\frac{1}{2}}\big)$ uniform in $s$ and $\theta \in \Theta$. Then, it follows that

$$\max_{\theta\geq\sigma} \big\{ \alpha_\beta\big(\varphi^n_{\sqrt{\gamma},\sigma}, \varphi^n_{0,\theta}\big) \big\} \geq \max_{\theta\in\Theta} \big\{ \alpha_\beta\big(\varphi^n_{\sqrt{\gamma},\sigma}, \varphi^n_{0,\theta}\big) \big\} \tag{115}$$

$$= \max_{\theta\in\Theta} \max_s \Big\{ \big(a(s,\gamma) + b(s,\gamma)\big) e^{n(\kappa_\gamma(s) + (1-s)\kappa'_\gamma(s))} + \mathbb{1}[s>1] + \big(\mathbb{1}[s<0] - \beta\big) e^{n\kappa'_\gamma(s)} \Big\} \tag{116}$$

$$= \max_s \max_{\theta\in\Theta} \Big\{ \big(a(s,\gamma) + b(s,\gamma)\big) e^{n(\kappa_\gamma(s) + (1-s)\kappa'_\gamma(s))} + \mathbb{1}[s>1] + \big(\mathbb{1}[s<0] - \beta\big) e^{n\kappa'_\gamma(s)} \Big\}, \tag{117}$$

where in (115) we restricted the range over which the maximization is performed, in (116) we used (96) with the error terms $o\big(n^{-\frac{1}{2}}\big)$ uniform in $s$ and $\theta \in \Theta$, and in (117) we interchanged the two maximizations.

In (117), we can fix a value $\theta \in \Theta$ and still obtain a lower bound to $\max_{\theta\geq\sigma} \big\{ \alpha_\beta\big(\varphi^n_{\sqrt{\gamma},\sigma}, \varphi^n_{0,\theta}\big) \big\}$. Moreover, as the maximization over $\theta$ is inside the maximization over $s$, the chosen value for $\theta$ may depend on $s$. Then, letting $\theta = \tilde{\theta}_s$ in the inner maximization we obtain the desired result. ∎

According to Corollary 3, we can use $\tilde{f}\big(\frac{1}{M},\gamma\big)$ instead of $f\big(\frac{1}{M},\gamma\big)$ in Theorems 2, 3 and in Corollary 2 and obtain a lower bound to the error probability. In Theorem 4 however, the convex hull of $f\big(\beta,\gamma\big)$ needs to be evaluated for a fixed variance $\theta^2$, which can be (i) the capacity-achieving $\theta^2 = \Upsilon + \sigma^2$, (ii) the exponent-achieving $\theta^2 = \tilde{\theta}^2_s$ for the value (102) optimizing the sphere-packing exponent (100), or (iii) $\theta^2 = \tilde{\theta}^2_s$ for the value of $s$ optimizing the finite-length expression (108) disregarding the $o\big(n^{-\frac{1}{2}}\big)$ terms.

## VII. Numerical Examples

### A. Comparison under the different power constraints

We consider the transmission of $M = 2^{nR}$ codewords over $n$ uses of an AWGN channel with $R = 1.5$ bits/channel use and $\mathrm{SNR} = 10\log_{10}\frac{\Upsilon}{\sigma^2} = 10$ dB. The channel capacity is $C = \frac{1}{2}\log_2\big(1+\frac{\Upsilon}{\sigma^2}\big) \approx 1.73$ bits/channel use under the three power constraints considered.

*1) Equal power constraint:* In Fig. 4, we compare the lower bounds discussed in Section III for the AWGN channel under an equal power constraint. For reference, we include the achievability part of [4, Eq. (20)], which was derived for an average power limitation and that, therefore, applies under

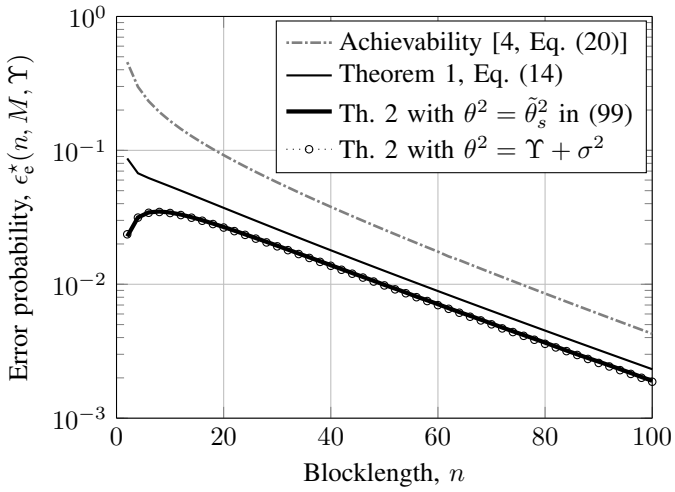



Fig. 4: Upper and lower bounds to the error probability for an AWGN channel under an equal power constraint. System parameters: SNR $= 10$ dB, and rate $R = 1.5$ bits/channel use.

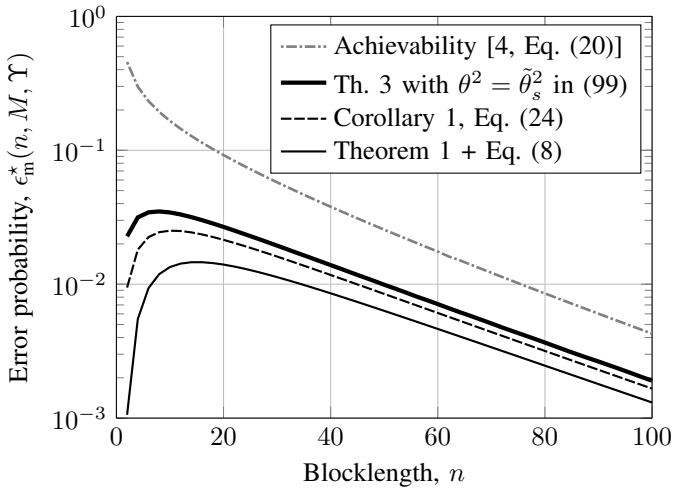



Fig. 5: Upper and lower bounds to the error probability for an AWGN channel under a maximal power constraint. System parameters: SNR $= 10$ dB, and rate $R = 1.5$ bits/channel use.

equal, maximal and average power constraints. We observe that Shannon cone-packing bound from Theorem 1 is the tightest lower bound in this setting. As discussed in Section III-A, by considering the optimal auxiliary (non-product) distribution, the meta-converse bound (10) recovers the cone-packing bound and therefore, it coincides with the curve from Theorem 1. The hypothesis testing bound from Theorem 2, with an auxiliary i.i.d. Gaussian distribution, is slightly weaker. Since the rate of the system $R = 1.5$ bits/channel use is relatively close to channel capacity $C \approx 1.8$ bits/channel use, the gain by using an auxiliary distribution equal to the exponent achieving output distribution is negligible in this example.

*2) Maximal power constraint:* For the same system parameters as in the previous example, Fig. 5 compares the bounds derived under a maximal power constraint. In particular, we

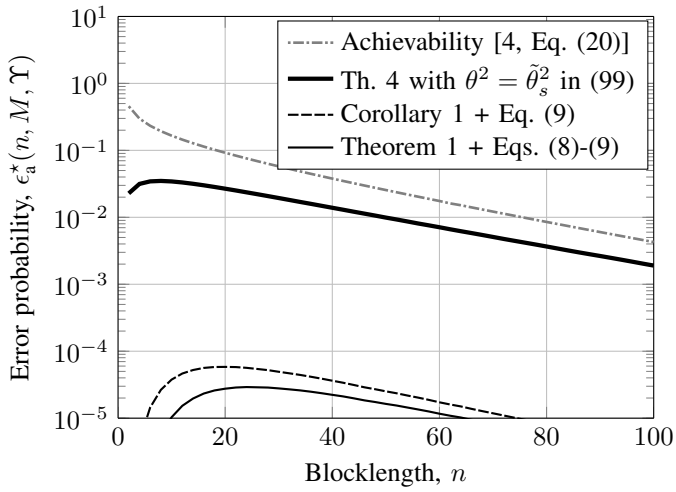



Fig. 6: Upper and lower bounds to the error probability for an AWGN channel under an average power constraint. System parameters: SNR $= 10$ dB, and rate $R = 1.5$ bits/channel use.

consider the combination of Theorem 1 with (8), the slightly sharper Corollary 1 and the hypothesis testing bound from Theorem 3 with $\theta = \theta_s$ as defined in (99). In the figure, we can see that the tightest bound in this setting corresponds to the new hypothesis testing bound from Theorem 3. Applying the relation (8) to extend the cone-packing bound from Theorem 1 to a maximal power constraint incurs in a certain loss, which is slightly tightened in Corollary 1. In the figure we observe that the four curves present the same asymptotic slope as they feature the same error exponent.

*3) Average power constraint:* We compare now the bounds for an average power constraint. In particular, we consider the combination of Theorem 1 with (8)-(9), the combination of Corollary 1 with (9) and the hypothesis testing bound from Theorem 4 with $\theta = \theta_s$ as defined in (99). For this set of system parameters, the condition in Corollary 2 is satisfied for all $n$, and the simplified bound (61) can be used to evaluate Theorem 4. The proposed hypothesis testing bound from Theorem 4 is the tightest lower bound in this setting, as shown in Fig. 6. The application of (9) to obtain bounds for an average power constraint incurs in a large loss with respect to the bound from Theorem 4.[6]

For this choice of system parameters, the condition in Corollary 2 is satisfied for all $n$. Then, it follows that the bounds from Theorems 2, 3 and 4 coincide in Figs. 4, 5 and 6. While in the equal power constraint setting, Shannon cone-packing bound is still the best lower bound, under both maximal and average power constraints the new hypothesis testing bounds yield tighter results. Indeed, for an average power constraint the advantage of Theorem 4 over previous results is significant in the finite blocklength regime, as shown in Fig. 6.

[6]All the bounds presented here hold under the average probability of error formalism. While the counterpart of (9) for maximal error probability is tighter (see [5, Lem. 65]), the finite-length gap to Theorem 4 is still significant.

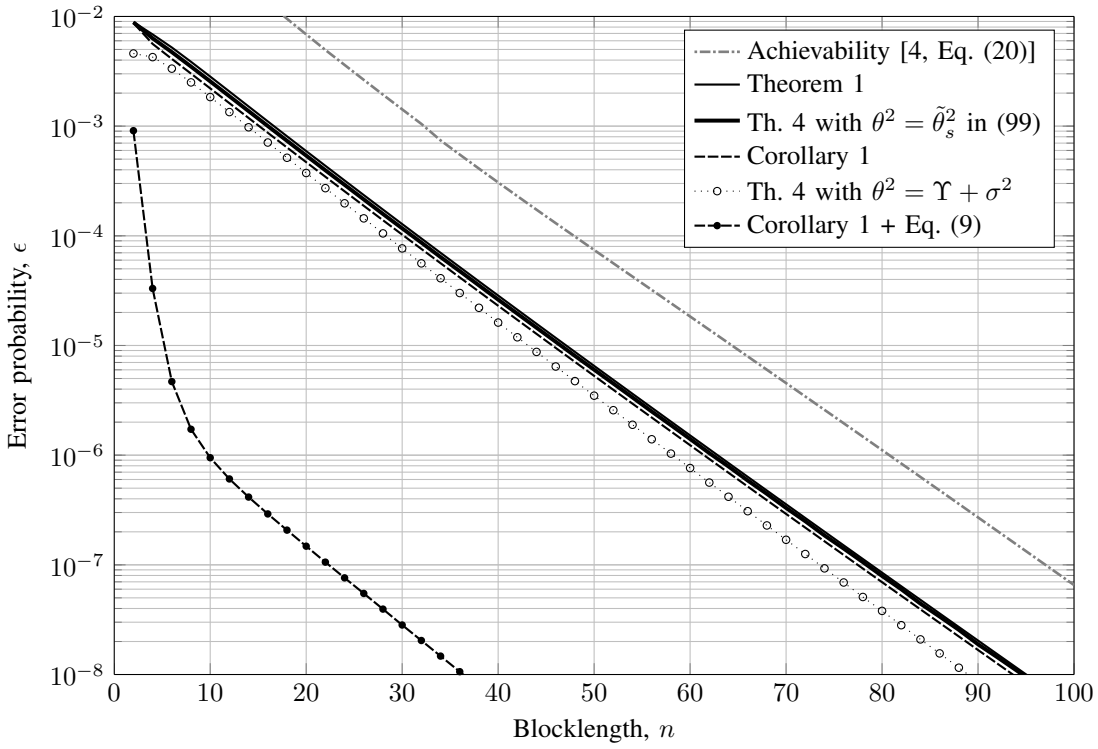



Fig. 7: Upper and lower bounds to the channel coding error probability for an AWGN channel with SNR $= 5$ dB and rate $R = 0.58$ bits/channel use.

### *B. Exponent-achieving output distribution*

In the previous examples, the transmission rate was very close to channel capacity. Therefore, using the exponent achieving or the capacity achieving output distributions in the hypothesis-testing bounds did not result in significant differences. We consider now an power-constrained AWGN channel with SNR $= 10 \log_{10} \frac{\Upsilon}{\sigma^2} = 5$ dB. The asymptotic capacity of this channel is $C \approx 1.03$ bits/channel use and its critical rate is $R_{\text{cr}} \approx 0.577$ bits/channel use.[7]

Figure 7 shows several of the bounds from previous examples for $M = \lceil 2^{nR} \rceil$ messages with $R = 0.58$ bits/channel use. This transmission rate is slightly above the critical rate of the channel. We first note that the gap between the achievability part of [4, Eq. (20)] and the converse bounds is larger than in the Figures 4, 5 and 6. We can see that Shannon cone-packing bound Theorem 1 is still the tightest lower bound. Nevertheless, the gap between Theorem 1 and the curve of Theorem 4 with $\theta^2 = \tilde{\theta}_s^2$ in (99) for the value of $s$ optimizing (108) is very small, and both bounds are almost indistinguishable in the plot. Moreover, while Theorem 1 holds only under equal power constraints, the bound from Theorem 4 holds under the three power constraints considered. In this example, Corollary 1 and Corollary 1 + Eq. (9) yield weaker and much weaker bounds, respectively. An error exponent analysis shows that the asymptotic slope of the hypothesis-testing lower bound from Theorem 4 with $\tilde{\theta}_s^2$ coincides with that of Shannon cone-packing lower bound from Theorem 1, hence both curves are parallel. In contrast, the curve with $\theta^2 = \Upsilon + \sigma^2$ presents a (slightly) larger error exponent and hence this bound will diverge as $n$ grows and become increasingly weaker. As a final remark on this results, note that by using the value $\theta^2 = \tilde{\theta}_s^2$, we obtain not only the sphere-packing exponent but also tighter finite-length bounds.

The observations from Fig. 7 are complemented in Fig. 8. In this figure, we analyze the highest transmission rate versus the blocklength for a given error probability $\epsilon = 10^{-6}$. We include the bounds from Fig. 7 and, for reference, the asymptotic channel capacity $C$ and the condition from Corollary 2 as an upper bound on the transmission rate of the system $R \leq \bar{R}_n = \frac{1}{n} \log_2 \bar{M}_n$. The upper bounds from Theorem 1, Corollary 1 and Theorem 4 with $\theta^2 = \tilde{\theta}_s^2$ are almost indistinguishable from each other. Note however that Theorem 1 was derived for an equal power constraint, Corollary 1 for a maximal power constraint and Theorem 4 for an average power constraint. Comparing these bounds with the achievability bound [4, Eq. (20)], we observe that the behavior of the transmission rate approaching capacity is precisely characterized for blocklengths $n \geq 30$. The upper bound from Theorem 4 with $\theta^2 = \Upsilon + \sigma^2$ is slightly weaker than that considering the exponent achieving output distribution and Corollary 1 + Eq. (9) yields a much weaker bound for average power constraints. The condition from Corollary 2 shows that the simpler bound (61) can be used to evaluate Theorem 4 in the range of values of $n$ considered, as $\bar{R}_n = \frac{1}{n} \log_2 \bar{M}_n$ is

[7]The critical rate of a channel is defined as the point below which the sphere-packing exponent and the random-coding exponent start to diverge [28]. For the power-constrained AWGN channel this point corresponds to the rate at which the maximum in (100) is attained for $s = \frac{1}{2}$.

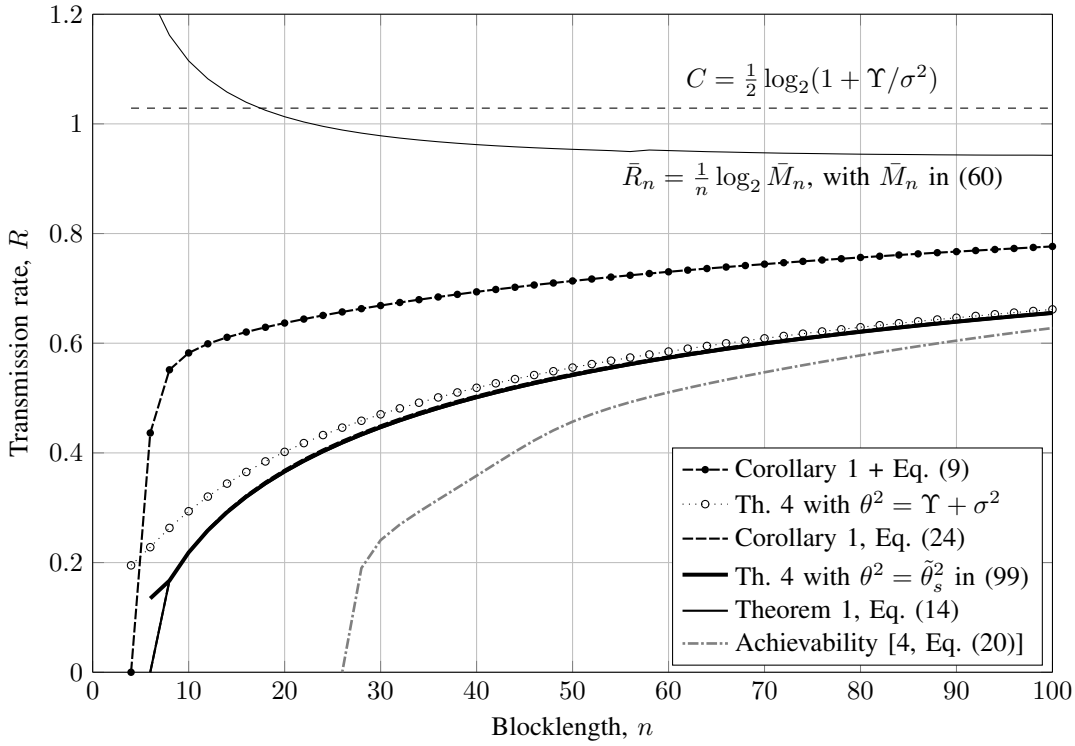



Fig. 8: Bounds to the transmission rate for an AWGN channel with SNR $= 5$ dB and error probability $\epsilon = 10^{-6}$.

well-above of the curve corresponding to Theorem 4.

### C. Numerical evaluation via the saddlepoint expansion

We now evaluate the accuracy of the saddlepoint expansion introduced in Theorem 6. To this end, in Fig. 9 we show with lines the exact hypothesis-testing bound evaluated using Theorem 5, and with markers (•) the approximation that follows from disregarding the $o\big(n^{-\frac{1}{2}}\big)$ terms in Theorem 6. We observe that, both for the capacity-achieving variance $\theta^2$ and for the exponent-achieving variance $\tilde{\theta}_s^2$, the approximation is accurate for blocklengths as short as $n = 10$. This is also be true for larger values of $n$, for which numerical evaluation of the Marcum-$Q$ functions appearing in Theorem 5 becomes unfeasible using traditional methods. In this scenario, the saddlepoint expansion from Theorem 6 is an useful tool to evaluate the hypothesis-testing bounds presented in this work.

Moreover, if we compare the curves of Shannon cone-packing bound from Theorem 1 and the hypothesis-testing bounds from Theorems 2, 3 and 4 in Figs. 4-9, we only observe a small difference. Then, regardless of the power constraint considered, it may be sufficient to use Theorem 4 as lower bound –since it was derived under an average power constraint, it applies for all equal, maximal and average power limitations– and the achievability part of [4, Eq. (20)] as an upper bound –this bound was derived assuming an equal-power constraint, and since it is an achievability result, it applies for all equal, maximal and average power limitations. The saddlepoint approximation from Theorem 6 is accurate for values of $n \geq 10$ and it can be safely applied in the evaluation of Theorem 4.

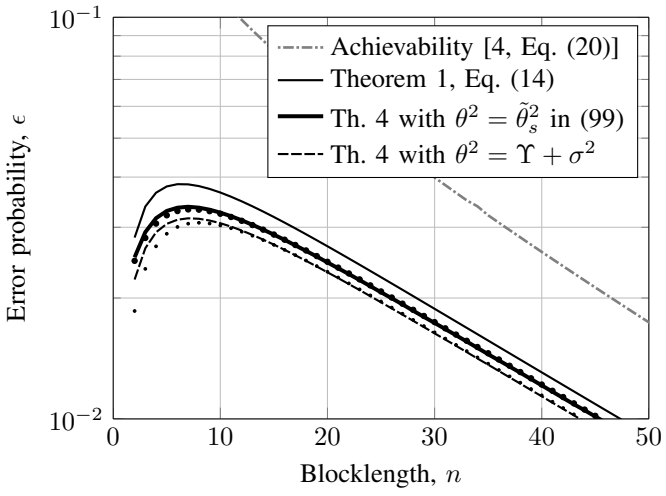



Fig. 9: Bounds to the error probability for an AWGN channel with SNR $= 5$ dB and $R = 0.8$ bits/channel use. The bounds from Theorem 4 have been evaluated exactly using Theorem 5 (lines) and using the approximation that follows from Theorem 6 disregarding the small-o terms (markers •).

### D. Constellation design for uncoded transmission ($n = 2$)

In the last example of this section, we consider the problem of transmitting $M \geq 2$ codewords over $n = 2$ uses of an AWGN channel with SNR $= 10$ dB. This problem corresponds to finding the best constellation for an uncoded quadrature communication system.

Figure 10 depicts cone-packing bound from Theorem 1,

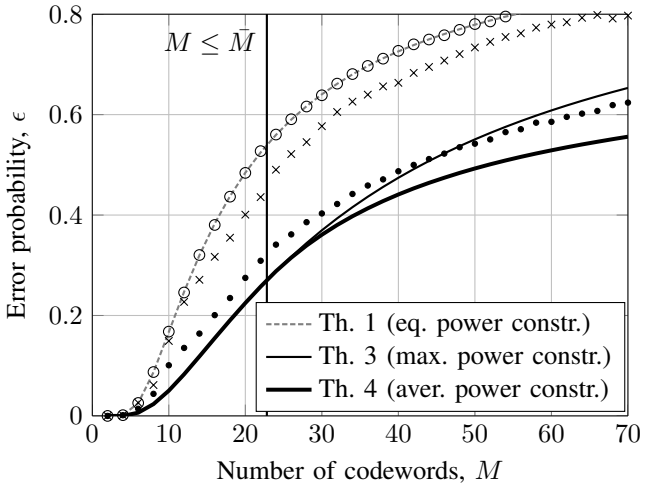



Fig. 10: Lower bounds to the channel coding error probability over an AWGN channel with $n = 2$ and SNR $= 10$ dB. Markers show the simulated error probability of a sequence of codes satisfying an equal ($\circ$), maximal ($\times$) and average ($\bullet$) power constraints. Vertical line corresponds to the boundary $M \leq \bar{M}$ from Corollary 2.

valid for equal power-constraints, the bound from Theorem 3 for $\theta^2 = \Upsilon + \sigma^2$, which is valid for maximal power-constraint, and that from Theorem 4, valid for average power-constraint. The vertical line shows the boundary of the region $M \leq \bar{M}$ defined in Corollary 2 where the bounds from Theorems 3 and 4 coincide. With markers, we show the simulated ML decoding error probability of a sequence of $M$-PSK (phase-shift keying) constellations satisfying the equal power constraint ($\circ$) and that of a sequence of $M$-APSK (amplitude-phase-shift keying) constellations satisfying maximal ($\times$) and average ($\bullet$) power constraints.[8]

Since the ML decoding regions of an $M$-PSK constellation are precisely 2-dimensional cones, Shannon'59 lower bound coincides with the corresponding simulated probability ($\circ$). However, Shannon'59 lower bound does not apply to general $M$-APSK constellations satisfying maximal ($\times$) and average ($\bullet$) power constraints, as discussed in Section IV.

We discuss now the results observed for codes satisfying maximal and average power constraints. We can see that while Theorem 4 applies in both of these settings, this is not the case for Theorem 3, that in general only applies under maximal power constraint. As stated in Corollary 2, the bounds from Theorems 3 and 4 coincide for $M \leq \bar{M} \approx 22.8$. Above this point, the two bounds diverge, and we can see from the figure that the average power constrained code ($\bullet$) violates the bound from Theorem 3 for $M > 45$.

Analyzing the constellations that violate Theorem 3, we found that they present several symbols concentrated at the origin of coordinates $(0, 0)$. As these symbols coincide, it is not possible to distinguish between them and they will often yield a decoding error. However, since the symbol $(0, 0)$ does not require any energy for its transmission, the average power for the remaining constellation points is increased and this code yields an overall smaller error probability. This effect was also observed in [5, Sec. 4.3.3], where a code with several codewords concentrated at the origin was used to study the asymptotics of the error probability in an average power constrained AWGN channel.

Interestingly, this optimal structure is also suggested by the input distribution that follows from the derivation of Theorem 4. The lower bound (48) in Theorem 4 corresponds to the value of the convex envelope $\underline{f}$ at the point $\left(\frac{1}{M}, \Upsilon\right)$. Whenever $M \leq \bar{M}$, this convex envelope corresponds to a convex combination of the functions $f\left(\bar{\beta}, 0\right)$ and $f\left(\beta_0, \gamma_0\right)$ with $\gamma_0 > \Upsilon$. Therefore, the input distribution induced by Theorem 4 is composed by a mass point at the origin and by a uniform distribution over the spherical shell with squared radius $\gamma_0 > \Upsilon$ (see Section V-A). While this distribution does not characterize how the codewords are distributed over the space, it suggest that several codewords could be concentrated at the symbol $(0, 0)$.

## VIII. Discussion

We studied the performance of block coding on an AWGN channel under different power limitations at the transmitter. In particular, we showed that the hypothesis-testing bound, [13, Th. 41] which was originally derived under an equal power limitation, also holds under maximal power constraints (Theorem 3), and, for rates below a given threshold, under average power constraints (Corollary 2). For rates close and above capacity, we proposed a new bound using the convex envelope of the error probability function of a certain binary hypothesis test (Theorem 4).

The performance bounds described above follow from the analysis of the meta-converse bound [13, Th. 27], which corresponds to the error probability a surrogate hypothesis test between the distribution induced by the channel and a certain auxiliary distribution. For the optimal auxiliary distribution and an equal power-constraint, Polyanskiy showed in [14, Sec. VI.F] that the meta-converse bound recovers Shannon cone-packing bound [4, Eq. (20)]. In this work, however, we chose the auxiliary distribution to be an i.i.d. Gaussian distribution with zero-mean and certain variance. If the variance is chosen to be capacity achieving, the resulting bound has a sub-optimal error exponent [11], [27]. Considering the variance of the exponent-achieving output distribution yields tighter finite-length bounds in general, which feature the sphere-packing exponent. Moreover, using an accurate asymptotic expansion (Theorem 6), it is possible to evaluate the bounds for the exponent-achieving output distribution without incurring in extra computational cost (Corollary 3).

While the numerical advantage of the new finite-length bounds compared to previous results in the literature is small for a maximal power constraint, it is significant for an average power constraint, as shown in Figures 5-8. Additionally, several of the theoretical contributions are of independent interest:

[8]The parameters of the $M$-APSK constellations (number of rings, number of points, amplitude and phase of each ring) have been optimized to minimize the error probability $\epsilon$ for each value of $M$. To this end, the constellation parameters are randomly chosen around their best known values, and only the constellations with lower error probability are used in the next iteration of the stochastic optimization algorithm.

- A new geometric interpretation of [13, Th. 41] which is analogous to the one in [4]. The hypothesis testing bound [13, Th. 41] can then be described as the probability of the noise moving a codeword $\boldsymbol{x}$ out of an $n$-dimensional sphere that roughly covers $1/M$-th of the output space. Interestingly, this sphere is not centered at the codeword $\boldsymbol{x}$ but at $\left(1+\frac{\sigma^2}{\Upsilon}\right)\boldsymbol{x}$.
- Optimization of the meta-converse bound over input distributions. While the results obtained are specific for an additive Gaussian noise channel, the techniques used can in principle be applied to more complicated channels, e.g., via the analysis of the saddlepoint expansion of the meta-converse bound [25].
- For an average power constraint and rates close to and above capacity, the input distribution that optimizing the meta-converse bound presents a mass point at the origin. This suggest that the optimal codes in this region must have several all-zeros codewords (as it occurs for the APSK constellations studied in Section VII-D) and motivates the fact that no strong-converse exists for an average power limitation at the transmitter [5, Th. 77].
- In Appendix A, we provide an exhaustive characterization of the error probability of a binary hypothesis test between two Gaussian distributions, which may be of interest in related problems.

In our derivations, we did not impose any structure to the codebooks beyond the corresponding power limitation. Then, the results obtained are general and do not require the codes to belong to a certain family, to use a specific modulation, or to satisfy minimum distance constraints. Nevertheless, the study of lower bounds for structured codes remains an active area of research (see, e.g., [22]). Tight lower bounds for BPSK modulations (or general $M$-PSK modulations), can be obtained from the meta-converse bound using the results from [25]. Evaluation of the meta-converse bound for general modulations is still an open problem due to the combinatorial nature of the optimization over input distributions.

## Acknowledgment

Fruitful discussions with Barış Nakiboğlu, Tobias Koch and David Morales-Jimenez are gratefully acknowledged.

## Appendix A
## Analysis of $f(\beta,\gamma)=\alpha_\beta\big(\varphi^n_{\sqrt{\gamma},\sigma},\varphi^n_{0,\theta}\big)$

### A. Parametric computation of $f(\beta,\gamma)$

*Proposition 1:* Let $\sigma,\theta>0$ and $n\geq 1$, be fixed parameters, and define $\delta\triangleq\theta^2-\sigma^2$. The function trade-off between $\alpha$ and $\beta$ in $\alpha_\beta\big(\varphi^n_{\sqrt{\gamma},\sigma},\varphi^n_{0,\theta}\big)$ admits the following parametric formulation as a function of the auxiliary parameter $t\geq 0$,

$$\alpha(\gamma,t)=Q_{\frac{n}{2}}\left(\sqrt{n\gamma}\frac{\sigma}{\delta},\frac{t}{\sigma}\right), \tag{118}$$

$$\beta(\gamma,t)=1-Q_{\frac{n}{2}}\left(\sqrt{n\gamma}\frac{\theta}{\delta},\frac{t}{\theta}\right), \tag{119}$$

where $Q_m(a,b)$ denotes the Marcum $Q$-function

$$Q_m(a,b)\triangleq\int_b^\infty\frac{t^m}{a^{m-1}}e^{-\frac{a^2+t^2}{2}}I_{m-1}(at)\,\mathrm{d}t. \tag{120}$$

To evaluate $f(\beta,\gamma)=\alpha_\beta\big(\varphi^n_{\sqrt{\gamma},\sigma},\varphi^n_{0,\theta}\big)$, we let $t_\star$ be the solution to the implicit equation $\beta(\gamma,t_\star)=\beta$ for $\beta(\gamma,t)$ defined in (119). Then, for $\alpha(\gamma,t)$ defined in (118), it follows that

$$f(\beta,\gamma)=\alpha(\gamma,t_\star). \tag{121}$$

*Proof:* The proof follows the lines of that of [13, Th. 41], and it is included here for completeness.[9]

Let $\sigma,\theta>0$ and $n\geq 1$, be fixed parameters. We define the log-likelihood ratio

$$j(\boldsymbol{y})\triangleq\log\frac{\varphi^n_{\sqrt{\gamma},\sigma}(\boldsymbol{y})}{\varphi^n_{0,\theta}(\boldsymbol{y})} \tag{122}$$

$$=n\log\frac{\theta}{\sigma}-\frac{1}{2}\sum_{i=1}^n\frac{\theta^2(y_i-\sqrt{\gamma})^2-\sigma^2y_i^2}{\sigma^2\theta^2}. \tag{123}$$

According to the Neyman-Pearson lemma, the trade-off $\alpha_\beta\big(\varphi^n_{\sqrt{\gamma},\sigma},\varphi^n_{0,\theta}\big)$ admits the parametric form

$$\alpha(t')=\Pr\big[j(\boldsymbol{Y}_0)\leq t'\big], \tag{124}$$

$$\beta(t')=\Pr\big[j(\boldsymbol{Y}_1)>t'\big], \tag{125}$$

in terms of the auxiliary parameter $t'\in\mathbb{R}$ and where $\boldsymbol{Y}_0\sim\varphi^n_{\sqrt{\gamma},\sigma}$, $\boldsymbol{Y}_1\sim\varphi^n_{0,\theta}$.

We now apply the change of variable $\boldsymbol{z}=(\boldsymbol{y}_0-\sqrt{\gamma})/\sigma$ such that for $\boldsymbol{Y}_0\sim\varphi^n_{\sqrt{\gamma},\sigma}$ we have that $\boldsymbol{Z}\sim\varphi^n_{0,1}$. It can be checked that the distribution of the random variable $j(\boldsymbol{Y}_0)$, $\boldsymbol{Y}_0\sim\varphi^n_{\sqrt{\gamma},\sigma}$ coincides with that of $j_0(\boldsymbol{Z})$, $\boldsymbol{Z}\sim\varphi^n_{0,1}$, where

$$j_0(\boldsymbol{z})\triangleq n\log\frac{\theta}{\sigma}+\frac{n}{2}\frac{\gamma}{\delta}-\frac{1}{2}\frac{\delta}{\theta^2}\sum_{i=1}^n\left(z_i-\frac{\sigma\sqrt{\gamma}}{\delta}\right)^2. \tag{126}$$

Analogously, if we define

$$j_1(\boldsymbol{z})\triangleq n\log\frac{\theta}{\sigma}+\frac{n}{2}\frac{\gamma}{\delta}-\frac{1}{2}\frac{\delta}{\sigma^2}\sum_{i=1}^n\left(z_i-\frac{\theta\sqrt{\gamma}}{\delta}\right)^2, \tag{127}$$

it follows that the distributions of $j(\boldsymbol{Y}_1)$, $\boldsymbol{Y}_1\sim\varphi^n_{0,\theta}$, and that of $j_1(\boldsymbol{Z})$, $\boldsymbol{Z}\sim\varphi^n_{0,1}$ coincide.

Then, we may rewrite (124)-(125) as

$$\alpha(t')=\Pr\big[j_0(\boldsymbol{Z})\leq t'\big], \tag{128}$$

$$\beta(t')=\Pr\big[j_1(\boldsymbol{Z})>t'\big], \tag{129}$$

where $\boldsymbol{Z}\sim\varphi^n_{0,1}$. Using (126) in (128), we obtain

$$\alpha(t')=\Pr\left[n\log\frac{\theta}{\sigma}+\frac{n}{2}\frac{\gamma}{\delta}-\frac{1}{2}\frac{\delta}{\theta^2}\sum_{i=1}^n\left(Z_i-\frac{\sigma\sqrt{\gamma}}{\delta}\right)^2\leq t'\right]. \tag{130}$$

We consider the change of variable $t'\leftrightarrow t$ such that

$$t'=n\log\frac{\theta}{\sigma}+\frac{n}{2}\frac{\gamma}{\delta}-\frac{\delta t^2}{2\sigma^2\theta^2}. \tag{131}$$

Using (131) in (130) and making the dependence on the parameter $\gamma$ explicit, we obtain

$$\alpha(\gamma,t)=\Pr\left[\sum_{i=1}^n\left(Z_i-\frac{\sigma\sqrt{\gamma}}{\delta}\right)^2\geq\left(\frac{t}{\sigma}\right)^2\right]. \tag{132}$$

[9]Note that the resulting trade-off (118)-(119) is scale invariant provided that $\sigma^2$, $\theta^2$ and $\gamma$ are scaled by the same quantity. Therefore, Proposition 1 is not more general than [13, Th. 41] by allowing $\sigma^2\neq 1$.

Proceeding analogously for (129) yields

$$\beta(\gamma,t)=\Pr\left[\sum_{i=1}^{n}\left(Z_i-\frac{\theta\sqrt{\gamma}}{\delta}\right)^2<\left(\frac{t}{\theta}\right)^2\right]. \tag{133}$$

Given (132) and (133), we conclude that $j_0(\boldsymbol{Z})$ and $j_1(\boldsymbol{Z})$ follow a (shifted and scaled) noncentral $\chi^2$ distribution with $n$ degrees of freedom and non-centrality parameters $n\gamma\sigma^2/\delta^2$ and $n\gamma\theta^2/\delta^2$, respectively. The cumulative density function of a non-central $\chi^2$ distribution with $n$ degrees of freedom and non-centrality parameter $\nu$ can be written in terms of the generalized Marcum $Q$-function $Q_m(a,b)$ as [29]

$$F_{n,\nu}(x)=1-Q_{\frac{n}{2}}\left(\sqrt{\nu},\sqrt{x}\right). \tag{134}$$

Noting that $F_{n,\nu}(x)$ is continuous, using (134) in (132) and (133), we obtain the desired result. ∎

### B. Derivatives of $f(\beta,\gamma)$

Let $\sigma,\theta>0$ and $n\geq 1$, be fixed parameters, and let $\delta=\theta^2-\sigma^2$. To obtain the derivatives of $f(\beta,\gamma)$ with respect to $\beta$ and $\gamma$, we start from the parametric formulation from Proposition 1 and use the following result.

*Proposition 2:* The derivatives of $Q_m(a,b)$ defined in (120) with respect to its parameters $a>0$ and $b>0$ are

$$\frac{\partial Q_m(a,b)}{\partial a}=\frac{b^m}{a^{m-1}}e^{-\frac{a^2+b^2}{2}}I_m(ab), \tag{135}$$

$$\frac{\partial Q_m(a,b)}{\partial b}=-\frac{b^m}{a^{m-1}}e^{-\frac{a^2+b^2}{2}}I_{m-1}(ab), \tag{136}$$

where $I_m(\cdot)$ denotes the $m$-th order modified Bessel function of the first kind.

*Proof:* The derivative (136) follows since the variable $b$ appears only in the lower limit of the definite integral in (120), then the derivative corresponds to the integrand evaluated at $t=b$.

To prove (135), let $n=m+\ell$ for some $\ell\in\mathbb{Z}^+$, and define the truncated sum

$$\tilde{Q}_m^{(n)}(a,b)\triangleq 1-e^{-\frac{a^2+b^2}{2}}\sum_{r=m}^{n}\left(\frac{b}{a}\right)^r I_r(ab). \tag{137}$$

The sequence of functions $\tilde{Q}_m^{(n)}(a,b)$ converges to $Q_m(a,b)$ as $n\to\infty$ [30, Eq. (4.63)]. The partial derivate of the truncated function (137) with respect to $a$ is given by

$$\frac{\partial\tilde{Q}_m^{(n)}(a,b)}{\partial a}=e^{-\frac{a^2+b^2}{2}}\sum_{r=m}^{n}\left(\frac{b}{a}\right)^r\left(\left(a+\frac{r}{a}\right)I_r(ab)-bI_r'(ab)\right). \tag{138}$$

Using the identity $I_m'(x)=\frac{m}{x}I_m(x)+I_{m+1}(x)$ [31, Sec. 8.486] and canceling terms, we obtain

$$\frac{\partial\tilde{Q}_m^{(n)}(a,b)}{\partial a}=\frac{b^m}{a^{m-1}}e^{-\frac{a^2+b^2}{2}}I_m(ab)-\frac{b^{n+1}}{a^n}e^{-\frac{a^2+b^2}{2}}I_{n+1}(ab). \tag{139}$$

Interchanging summation and differentiation is possible if the derivatives of the summands present uniform convergence [31, Sec. 0.307]. We next show that the sequence (139) presents uniform convergence to the RHS of (135). Then, noting that the sequence of functions $\tilde{Q}_m^{(n)}(a,b)$ converges to $Q_m(a,b)$ as $n\to\infty$, and since the sequence $\partial\tilde{Q}_m^{(n)}(a,b)/\partial a$ converges uniformly to $\partial Q_m(a,b)/\partial a$, we obtain the desired identity (135).

Indeed, using [31, Sec. 8.431] it follows that, for $n\geq 2$

$$\left(\frac{b}{a}\right)^{n+1}I_{n+1}(ab)=\frac{(b^2/2)^{n+1}}{\Gamma\left(n+\frac{3}{2}\right)\Gamma\left(\frac{1}{2}\right)}\int_{-1}^{1}\left(1-t^2\right)^{n+\frac{1}{2}}e^{abt}\,\mathrm{d}t \tag{140}$$

$$\leq\frac{(b^2/2)^{n+1}e^{ab}}{\Gamma\left(n+\frac{3}{2}\right)\Gamma\left(\frac{1}{2}\right)}, \tag{141}$$

where in the last step we used that $e^{abt}\leq e^{ab}$ for $t\in[-1,1]$ and that $\int_{-1}^{1}\left(1-t^2\right)^{n+\frac{1}{2}}\mathrm{d}t<1$ for $n\geq 2$.

Then, from (139) and (141) we obtain

$$\left|\frac{\partial\tilde{Q}_m^{(n)}(a,b)}{\partial a}-\frac{b^m}{a^{m-1}}e^{-\frac{a^2+b^2}{2}}I_m(ab)\right|=ae^{-\frac{a^2-2ab+b^2}{2}}\frac{(b^2/2)^{n+1}}{\Gamma\left(n+\frac{3}{2}\right)\Gamma\left(\frac{1}{2}\right)} \tag{142}$$

which, for $n$ sufficiently large, is uniformly bounded for any $0<a\leq\bar{a}$ and $0<b\leq\bar{b}$ with $\bar{a},\bar{b}<\infty$, since the growth of $\Gamma\left(n+\frac{3}{2}\right)$ is asymptotically faster than that of $(b^2/2)^{n+1}$. ∎

Using the derivatives of the Marcum-$Q$ function (135) and (136), we obtain that the derivatives of (118) are

$$\frac{\partial\alpha(\gamma,t)}{\partial\gamma}=\frac{1}{2}\frac{\sigma\sqrt{n/\gamma}}{\delta}\frac{b^{\frac{n}{2}}}{a^{\frac{n}{2}-1}}e^{-\frac{a^2+b^2}{2}}I_{\frac{n}{2}}(ab), \tag{143}$$

$$\frac{\partial\alpha(\gamma,t)}{\partial t}=-\frac{1}{\sigma}\frac{b^{\frac{n}{2}}}{a^{\frac{n}{2}-1}}e^{-\frac{a^2+b^2}{2}}I_{\frac{n}{2}-1}(ab), \tag{144}$$

with $a=\sqrt{n\gamma}\frac{\sigma}{\delta}$ and $b=\frac{t}{\sigma}$. Proceeding analogously, for the derivatives of (119) we obtain

$$\frac{\partial\beta(\gamma,t)}{\partial\gamma}=-\frac{1}{2}\frac{\theta\sqrt{n/\gamma}}{\delta}\frac{\bar{b}^{\frac{n}{2}}}{\bar{a}^{\frac{n}{2}-1}}e^{-\frac{\bar{a}^2+\bar{b}^2}{2}}I_{\frac{n}{2}}(\bar{a}\bar{b}), \tag{145}$$

$$\frac{\partial\beta(\gamma,t)}{\partial t}=\frac{1}{\theta}\frac{\bar{b}^{\frac{n}{2}}}{\bar{a}^{\frac{n}{2}-1}}e^{-\frac{\bar{a}^2+\bar{b}^2}{2}}I_{\frac{n}{2}-1}(\bar{a}\bar{b}), \tag{146}$$

where $\bar{a}=\sqrt{n\gamma}\frac{\theta}{\delta}$ and $\bar{b}=\frac{t}{\theta}$. Note that $ab=\bar{a}\bar{b}$, hence, $I_{\frac{n}{2}}(ab)=I_{\frac{n}{2}}(\bar{a}\bar{b})$ and $I_{\frac{n}{2}-1}(ab)=I_{\frac{n}{2}-1}(\bar{a}\bar{b})$.

We now obtain the derivatives of $f(\beta,\gamma)$ with respect to the parameters $\beta$ and $\gamma$:

*1) Derivative $\partial f(\beta,\gamma)/\partial\gamma$ for fixed $\beta$:* Let $\beta\in[0,1]$ be fixed and let $t(\gamma)$ be such that $\beta\big(\gamma,t(\gamma)\big)=\beta$ from (119). We apply the chain rule for total derivatives to write

$$\frac{\partial\beta\big(\gamma,t(\gamma)\big)}{\partial\gamma}=\left(\frac{\partial\beta(\gamma,t)}{\partial\gamma}+\frac{\partial\beta(\gamma,t)}{\partial t}\frac{\partial t(\gamma)}{\partial\gamma}\right)\Bigg|_{t=t(\gamma)}. \tag{147}$$

As $\beta\big(\gamma,t(\gamma)\big)=\beta$ is fixed, then (147) must be equal to 0. Then, identifying (147) to 0 and solving for $\frac{\partial t(\gamma)}{\partial\gamma}$ yields

$$\frac{\partial t(\gamma)}{\partial\gamma}=-\frac{\frac{\partial}{\partial\gamma}\beta(\gamma,t)}{\frac{\partial}{\partial t}\beta(\gamma,t)}=\frac{\theta^2}{2\delta}\sqrt{\frac{n}{\gamma}}\frac{I_{\frac{n}{2}}\left(\sqrt{n\gamma}\frac{t}{\delta}\right)}{I_{\frac{n}{2}-1}\left(\sqrt{n\gamma}\frac{t}{\delta}\right)}, \tag{148}$$

where $t=t(\gamma)$, and where we used (145) and (146). Note that we obtained an expression for $\frac{\partial t(\gamma)}{\partial\gamma}$ without computing $t(\gamma)$

explicitly, as doing this would require to invert (119) which is not analytically tractable.

We apply now the chain rule for total derivatives to $\alpha\big(\gamma,t(\gamma)\big)$ to write

$$\frac{\partial\alpha\big(\gamma,t(\gamma)\big)}{\partial\gamma}=\left(\frac{\partial\alpha(\gamma,t)}{\partial\gamma}+\frac{\partial\alpha(\gamma,t)}{\partial t}\frac{\partial t(\gamma)}{\partial\gamma}\right)\Bigg|_{t=t(\gamma)} \tag{149}$$

Note that, for fixed $\beta$, $\frac{\partial f(\beta,\gamma)}{\partial\gamma}=\frac{\partial\alpha(\gamma,t(\gamma))}{\partial\gamma}$. Hence, using (143), (144) and (148) in (149) we finally obtain

$$\frac{\partial f(\beta,\gamma)}{\partial\gamma}=-\frac{n}{2\delta}\left(\frac{t\delta}{\sigma^2\sqrt{n\gamma}}\right)^{\frac{n}{2}}e^{-\frac{1}{2}\left(\frac{n\gamma\sigma^2}{\delta^2}+\frac{t^2}{\sigma^2}\right)}I_{\frac{n}{2}}\left(\sqrt{n\gamma}\frac{t}{\delta}\right), \tag{150}$$

where $t$ satisfies $\beta(\gamma,t)=\beta$ with $\beta(\gamma,t)$ given in (119).

*2) Derivative $\partial f(\beta,\gamma)/\partial\beta$ for fixed $\gamma$:* In this case we use (144) and (146) to obtain

$$\frac{\partial f(\beta,\gamma)}{\partial\beta}=\frac{\frac{\partial}{\partial t}\alpha(\gamma,t)}{\frac{\partial}{\partial t}\beta(\gamma,t)}=-\frac{\theta^n}{\sigma^n}e^{\frac{1}{2}\left(\frac{n\gamma}{\delta}-t^2\left(\frac{1}{\sigma^2}-\frac{1}{\theta^2}\right)\right)} \tag{151}$$

where $t$ satisfies $\beta(\gamma,t)=\beta$ with $\beta(\gamma,t)$ given in (119).

*3) Derivative $\partial^2 f(\beta,\gamma)/(\partial\beta\partial\gamma)$:* Taking the derivative of (151) with respect to $\gamma$ yields

$$\frac{\partial^2 f(\beta,\gamma)}{\partial\beta\partial\gamma}=-\frac{\theta^n}{\sigma^n}e^{\frac{1}{2}\left(\frac{n\gamma}{\delta}-t^2\left(\frac{1}{\sigma^2}-\frac{1}{\theta^2}\right)\right)}\times\left(\frac{n}{2\delta}-t\left(\frac{1}{\sigma^2}-\frac{1}{\theta^2}\right)\frac{\partial t(\gamma)}{\partial\gamma}\right), \tag{152}$$

where $t$ satisfies $\beta(\gamma,t)=\beta$ with $\beta(\gamma,t)$ given in (119), and where $\frac{\partial t(\gamma)}{\partial\gamma}$ is given in (148).

*4) Derivative $\partial^2 f(\beta,\gamma)/(\partial\beta)^2$:* Taking the derivative of (151) with respect to $\beta$ yields

$$\frac{\partial^2 f(\beta,\gamma)}{(\partial\beta)^2}=t\frac{\theta^n}{\sigma^n}e^{\frac{1}{2}\left(\frac{n\gamma}{\delta}-t^2\left(\frac{1}{\sigma^2}-\frac{1}{\theta^2}\right)\right)}\left(\frac{1}{\sigma^2}-\frac{1}{\theta^2}\right)\frac{\partial t}{\partial\beta} \tag{153}$$

where $t$ satisfies $\beta(\gamma,t)=\beta$ with $\beta(\gamma,t)$ given in (119), and where the term $\frac{\partial t}{\partial\beta}$ can be obtained from (146),

$$\frac{\partial t}{\partial\beta}=\left(\frac{\partial\beta(\gamma,t)}{\partial t}\right)^{-1} \tag{154}$$

$$=\frac{\delta}{\sqrt{n\gamma}}\left(\frac{\theta^2\sqrt{n\gamma}}{t\delta}\right)^{\frac{n}{2}}e^{\frac{1}{2}\left(\frac{n\gamma\theta^2}{\delta^2}+\frac{t^2}{\theta^2}\right)}\times\left(I_{\frac{n}{2}-1}\left(\sqrt{n\gamma}\frac{t}{\delta}\right)\right)^{-1}. \tag{155}$$

*5) Derivative $\partial^2 f(\beta,\gamma)/(\partial\gamma)^2$:* Taking the derivative of (150) with respect to $\gamma$, straightforward but tedious algebra yields

$$\frac{\partial^2 f(\beta,\gamma)}{(\partial\gamma)^2}=-\frac{n}{4\delta}\left(\frac{t\delta}{\sigma^2\sqrt{n\gamma}}\right)^{\frac{n}{2}}e^{-\frac{1}{2}\left(\frac{n\gamma\sigma^2}{\delta^2}+\frac{t^2}{\sigma^2}\right)}I_{\frac{n}{2}}\left(\sqrt{n\gamma}\frac{t}{\delta}\right)\times\left(\frac{n}{\delta}-\frac{n}{\gamma}+\sqrt{\frac{n}{\gamma}}\frac{t}{\delta}\left(\frac{I_{\frac{n}{2}-1}\left(\sqrt{n\gamma}\frac{t}{\delta}\right)}{I_{\frac{n}{2}}\left(\sqrt{n\gamma}\frac{t}{\delta}\right)}-\frac{\theta^2}{\sigma^2}\frac{I_{\frac{n}{2}}\left(\sqrt{n\gamma}\frac{t}{\delta}\right)}{I_{\frac{n}{2}-1}\left(\sqrt{n\gamma}\frac{t}{\delta}\right)}\right)\right), \tag{156}$$

where $t$ satisfies $\beta(\gamma,t)=\beta$ with $\beta(\gamma,t)$ given in (119). Here, we used the identity $I'_m(x)=I_{m-1}(x)-\frac{m}{x}I_m(x)$ [31, Sec. 8.486].

### C. Derivatives of $f(\beta,\gamma)$ at $\gamma=0$

The function $f(\beta,0)$ can be evaluated by setting $\gamma=0$ and using (118)-(119). However, the preceding expressions for the derivatives of $f(\beta,\gamma)$ often yield an indeterminacy in this case. This can be avoided by taking the limit as $\gamma\to 0$ and using that [31, Sec. 8.445]

$$I_m(x)=\frac{\left(\frac{x}{2}\right)^m}{\Gamma\big(m+1\big)}+o(x^m), \tag{157}$$

where $\Gamma(\cdot)$ denotes the gamma function and $o\big(g(x)\big)$ summarizes the terms that approach zero faster than $g(x)$, *i.e.*, $\lim_{x\to 0}\frac{o(g(x))}{g(x)}=0$. For example, using (157) and $\frac{\Gamma(m+1)}{\Gamma(m)}=m$ we obtain from (148) that

$$\left.\frac{\partial t(\gamma)}{\partial\gamma}\right|_{\gamma=0}=\frac{t}{2}\frac{\theta^2}{\delta^2}. \tag{158}$$

Proceeding analogously for the derivatives of $f(\beta,\gamma)$, it follows that

$$\left.\frac{\partial f(\beta,\gamma)}{\partial\gamma}\right|_{\gamma=0}=-\frac{1}{\delta}\frac{t_0^n}{\sigma^n}\frac{e^{-\frac{1}{2}\frac{t_0^2}{\sigma^2}}}{\Gamma\big(\frac{n}{2}\big)2^{\frac{n}{2}}}, \tag{159}$$

$$\left.\frac{\partial f(\beta,\gamma)}{\partial\beta}\right|_{\gamma=0}=-\frac{\theta^n}{\sigma^n}e^{-\frac{1}{2}t_0^2\left(\frac{1}{\sigma^2}-\frac{1}{\theta^2}\right)}, \tag{160}$$

$$\left.\frac{\partial^2 f(\beta,\gamma)}{\partial\beta\partial\gamma}\right|_{\gamma=0}=-\frac{\theta^n}{\sigma^n}\left(\frac{n}{2\delta}-\frac{t_0^2}{2\delta\sigma^2}\right)e^{-\frac{1}{2}t_0^2\left(\frac{1}{\sigma^2}-\frac{1}{\theta^2}\right)}, \tag{161}$$

$$\left.\frac{\partial^2 f(\beta,\gamma)}{(\partial\beta)^2}\right|_{\gamma=0}=\frac{\theta^n}{\sigma^n}\left(\frac{\theta\sqrt{2}}{t_0}\right)^{n-2}\frac{\delta}{\sigma^2}\Gamma\big(\tfrac{n}{2}\big)e^{-\frac{1}{2}t_0^2\frac{\delta-\sigma^2}{\theta^2\sigma^2}}, \tag{162}$$

$$\left.\frac{\partial^2 f(\beta,\gamma)}{(\partial\gamma)^2}\right|_{\gamma=0}=-\frac{n}{4\delta}\frac{t_0^n}{\sigma^n 2^{\frac{n}{2}}}\left(\frac{n}{\delta}+\left(\frac{n}{n+2}-\frac{\theta^2}{\sigma^2}\right)\frac{t_0^2}{\delta^2}\right)\times\frac{e^{-\frac{1}{2}\frac{t_0^2}{\sigma^2}}}{\Gamma(\frac{n}{2}+1)}, \tag{163}$$

where in all cases $t_0$ satisfies $\beta(0,t_0)=\beta$ with $\beta(\gamma,t)$ given in (119). To obtain (163) from (156) we used (157) and the expansions

$$\frac{I_{m-1}(x)}{I_m(x)}=\frac{2m}{x}+\frac{x}{2(m+1)}+o(x), \tag{164}$$

$$\frac{I_m(x)}{I_{m-1}(x)}=\frac{x}{2m}+o(x). \tag{165}$$

## Appendix B
## Proof of Lemma 4

We characterize the region where $f(\beta,\gamma)$ and its convex envelope $\underline{f}(\beta,\gamma)$ coincide using the following result.

*Proposition 3:* Suppose $g$ is differentiable with gradient $\nabla g$. Let $\mathcal{A}$ denote the domain of $g$, and let $a_0\in\mathcal{A}$. If the inequality

$$g(\bar{a})\geq g(a_0)+\nabla g(a_0)^T(\bar{a}-a_0), \tag{166}$$

is satisfied for all $\bar{a}\in\mathcal{A}$, then, $g(a_0)=g^{**}(a_0)$ holds.

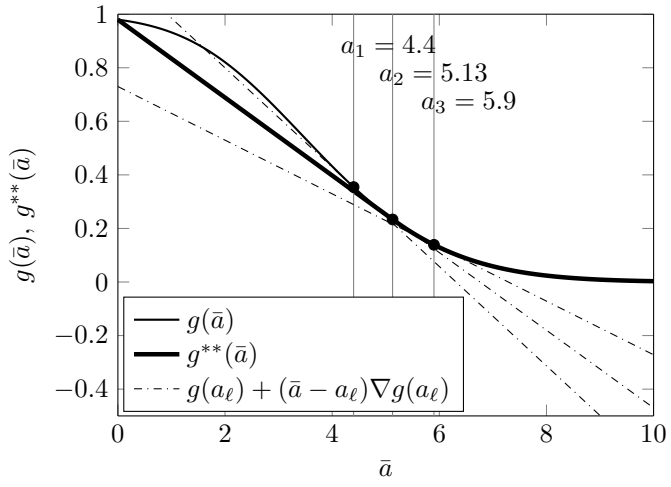



Fig. 11: Example of Proposition 3 for the one-dimensional function $g(\bar{a}) = f(\beta, \bar{a})$ with $\beta = 0.001$, $n = 6$, $\sigma^2 = 1$ and $\theta^2 = 3$, which is defined for $\bar{a} \geq 0$.

*Proof:* As $g^{**}$ is the lower convex envelope of $g$, then $g(a_0) \geq g^{**}(a_0)$ trivially. It remains to show that (166) implies $g(a_0) \leq g^{**}(a_0)$. Fenchel's inequality [32, Sec. 3.3.2] yields

$$g^{**}(a_0) \geq \langle a_0, b\rangle - g^*(b), \tag{167}$$

for any $b$ in the domain of $g^*$.

Setting $b = \nabla g(a_0)$ and using (46) in (167), we obtain

$$\begin{aligned} g^{**}(a_0) &\geq \nabla g(a_0)^T a_0 - \max_{\bar{a}\in\mathcal{A}}\big\{\nabla g(a_0)^T \bar{a} - g(\bar{a})\big\} && (168)\\ &= \min_{\bar{a}\in\mathcal{A}}\big\{\nabla g(a_0)^T (a_0 - \bar{a}) + g(\bar{a})\big\} && (169)\\ &\geq \min_{\bar{a}\in\mathcal{A}}\big\{g(a_0)\big\}, && (170)\end{aligned}$$

where in the last step we used (166) to lower bound $g(\bar{a})$. Since the objective of (170) does not depend on $\bar{a}$, we conclude from (168)-(170) that $g(a_0) \leq g^{**}(a_0)$ and the result follows. ■

Figure 11 shows an example of Proposition 3 for a certain one-dimensional function $g$. When $a_0 = a_1$, we can see in the figure that (173) is violated as the corresponding first order Taylor approximation (dash-dotted line) is larger than $g(\bar{a})$ (thin solid line) for small values of $\bar{a}$. Then, the convex envelope $g^{**}$ (thick solid line) is strictly smaller than the function $g$ at the point $a_0 = a_1$. In contrast, when $a_0 = a_2$ (173) is satisfied for all values of $\bar{a} \geq 0$. Therefore $g$ coincides with its convex envelope $g^{**}$ at $a_0 = a_2$. This is also true for any $a_0 > a_2$ (e.g., $a_0 = a_3$), and therefore $g$ and its convex envelope $g^{**}$ coincide for $a_0 \geq a_2$.

We apply Proposition 3 to the function $f(\beta, \gamma)$. We recall that $f(\beta, \gamma)$ is differentiable for $\beta \in [0, 1]$ and $\gamma \geq 0$ with derivatives given in Appendix A. We define the gradients

$$\nabla_\beta f(b, g) \triangleq \frac{\partial f(\beta, \gamma)}{\partial \beta}\Big|_{\beta=b,\gamma=g}, \tag{171}$$

$$\nabla_\gamma f(b, g) \triangleq \frac{\partial f(\beta, \gamma)}{\partial \gamma}\Big|_{\beta=b,\gamma=g}. \tag{172}$$

According to Proposition 3, the function $f(\beta_0, \gamma_0)$ and its convex envelope $\underline{f}(\beta_0, \gamma_0)$ coincide if

$$\begin{aligned} f(\bar{\beta}, \bar{\gamma}) \;\geq\; & f(\beta_0, \gamma_0) + (\bar{\beta} - \beta_0)\nabla_\beta f(\beta_0, \gamma_0) \\ & + (\bar{\gamma} - \gamma_0)\nabla_\gamma f(\beta_0, \gamma_0). \end{aligned} \tag{173}$$

is satisfied for all $\beta \in [0, 1]$ and $\gamma \geq 0$. This condition implies that the first-order Taylor approximation of $f(\beta, \gamma)$ at $(\beta_0, \gamma_0)$ is a global under-estimator of the original function $f$.

The derivatives of $f(\beta, \gamma)$, given in Appendix A, imply that the function is decreasing in both parameters, convex with respect to $\beta \in [0, 1]$, and jointly convex with respect to $(\beta, \gamma)$ except for a neighborhood near the axis $\gamma = 0$. Using these properties, it follows that the condition (173) only needs to be verified along the axis $\bar{\gamma} = 0$. For example, for the one-dimensional function $g$ shown in Fig. 11, we can see that if the first-order condition is satisfied at $\bar{a} = 0$, it is also satisfied for any $\bar{a} \geq 0$.

Then, we conclude that $f(\beta_0, \gamma_0) = \underline{f}(\beta_0, \gamma_0)$ if (173) holds for every $\bar{\beta} \in [0, 1]$ and $\bar{\gamma} = 0$, i.e., if

$$\begin{aligned} f(\beta_0, \gamma_0) - f(\bar{\beta}, 0) \geq\; & (\beta_0 - \bar{\beta})\nabla_\beta f(\beta_0, \gamma_0) \\ & + \gamma_0 \nabla_\gamma f(\beta_0, \gamma_0). \end{aligned} \tag{174}$$

Let $\theta \geq \sigma > 0$, $n \geq 1$. Let $t_0$ be the value such that $\beta(\gamma_0, t_0) = \beta_0$ and let $\bar{t}$ satisfy $\beta(0, \bar{t}) = \bar{\beta}$, for $\beta(\gamma, t)$ defined in (119). Using (118) and the derivatives (150) and (151) from Appendix A, we obtain the identities

$$f(\beta_0, \gamma_0) - f(\bar{\beta}, 0) = Q_{\frac{n}{2}}\left(\sqrt{n\gamma_0}\frac{\sigma}{\delta}, \frac{t_0}{\sigma}\right) - Q_{\frac{n}{2}}\left(0, \frac{\bar{t}}{\sigma}\right), \tag{175}$$

$$\nabla_\beta f(\beta_0, \gamma_0) = -\frac{\theta^n}{\sigma^n} e^{\frac{1}{2}\left(\frac{n\gamma_0}{\delta} - t_0^2\left(\frac{1}{\sigma^2} - \frac{1}{\theta^2}\right)\right)}, \tag{176}$$

$$\begin{aligned} \nabla_\gamma f(\beta_0, \gamma_0) = & -\frac{n}{2\delta}\left(\frac{t_0\delta}{\sigma^2\sqrt{n\gamma_0}}\right)^{\frac{n}{2}} e^{-\frac{1}{2}\left(n\gamma_0\frac{\sigma^2}{\delta^2} + \frac{t_0^2}{\sigma^2}\right)} \\ & \times I_{\frac{n}{2}}\left(\sqrt{n\gamma_0}\frac{t_0}{\delta}\right). \end{aligned} \tag{177}$$

As $\beta(\gamma_0, t_0) = \beta_0$ and $\beta(0, \bar{t}) = \bar{\beta}$, using (119), it follows that

$$\beta_0 - \bar{\beta} = Q_{\frac{n}{2}}\left(0, \frac{\bar{t}}{\theta}\right) - Q_{\frac{n}{2}}\left(\sqrt{n\gamma_0}\frac{\theta}{\delta}, \frac{t_0}{\theta}\right). \tag{178}$$

Then, substituting (175) and (178) in (174), reorganizing terms, it yields

$$\begin{aligned} Q_{\frac{n}{2}}\left(\sqrt{n\gamma_0}\frac{\sigma}{\delta}, \frac{t_0}{\sigma}\right) + \nabla_\beta f(\beta_0, \gamma_0) Q_{\frac{n}{2}}\left(\sqrt{n\gamma_0}\frac{\theta}{\delta}, \frac{t_0}{\theta}\right) \\ - \gamma_0 \nabla_\gamma f(\beta_0, \gamma_0) \;\geq\; h(\bar{t}), \end{aligned} \tag{179}$$

where $h(t)$ is given by

$$h(t) \triangleq Q_{\frac{n}{2}}\left(0, \frac{t}{\sigma}\right) + \nabla_\beta f(\beta_0, \gamma_0) Q_{\frac{n}{2}}\left(0, \frac{t}{\theta}\right). \tag{180}$$

The interval $\bar{\beta} \in [0, 1]$ corresponds to $\bar{t} \geq 0$. We maximize (180) over $t = \bar{t} \geq 0$ and we only verify the condition (179) for this maximum value.

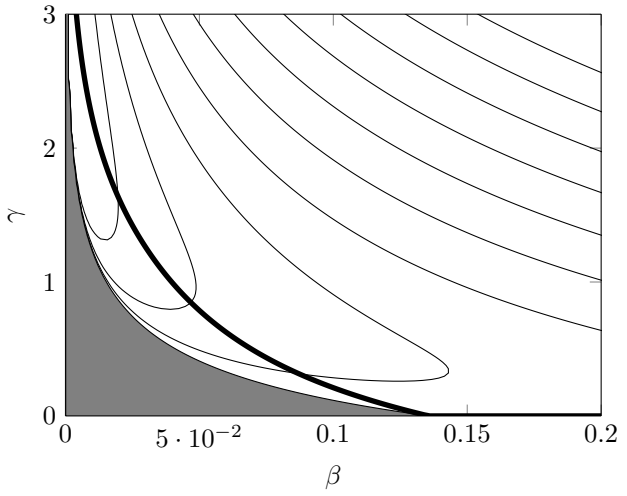



Fig. 12: Level curves of $\det \nabla^2 f(\beta, \gamma)$ for $n = 6$, $\sigma^2 = 1$, $\theta^2 = 2$. The region where $\det \nabla^2 f(\beta, \gamma) < 0$ is shaded in gray. The bold line corresponds to the points where $\beta = 1 - Q_{\frac{n}{2}}\big(\sqrt{n\gamma}\theta/\delta,\, t_0/\theta\big)$ as described in Lemma 4.

To this end, we find the derivative of (180) with respect to $t$, we identify the resulting expression with zero and solve for $t$. Using (136) and (157) it follows that

$$\frac{\partial}{\partial b} Q_m\left(0, b\right) = -\frac{b^{2m-1}}{2^{m-1}} \frac{e^{-\frac{b^2}{2}}}{\Gamma(m)} \tag{181}$$

and therefore

$$\frac{\partial}{\partial t} h(t)$$

$$= -\frac{1}{\sigma} \frac{(t/\sigma)^{n-1}}{2^{\frac{n}{2}-1}} \frac{e^{-\frac{t^2}{2\sigma^2}}}{\Gamma(n/2)} - \frac{\nabla_\beta f(\beta_0, \gamma_0)}{\theta} \frac{(t/\theta)^{n-1}}{2^{\frac{n}{2}-1}} \frac{e^{-\frac{t^2}{2\theta^2}}}{\Gamma(n/2)} \tag{182}$$

$$= -\frac{t^{n-1}}{\sigma^n 2^{\frac{n}{2}-1}} \frac{e^{-\frac{t^2}{2\sigma^2}}}{\Gamma(n/2)} + \frac{t^{n-1}}{\sigma^n 2^{\frac{n}{2}-1}} \frac{e^{-\frac{t^2}{2\theta^2}}}{\Gamma(n/2)} e^{\frac{1}{2}\left(\frac{n\gamma_0}{\delta} - t_0^2\left(\frac{1}{\sigma^2} - \frac{1}{\theta^2}\right)\right)} \tag{183}$$

where in the second step we used (176). Identifying (183) with zero, we obtain the root $t = 0$ and (after some algebra)

$$t^2 = t_0^2 - n\gamma \frac{\sigma^2 \theta^2}{\delta^2}. \tag{184}$$

By evaluating the second derivative of (180), we verify that (184) corresponds to a maximum of $h(t)$. Therefore, we conclude that the RHS of (179) is maximized for

$$\bar{t}_\star = \sqrt{\left(t_0^2 - n\gamma\sigma^2\theta^2/\delta^2\right)_+} \tag{185}$$

where the threshold $(a)_+ = \max(0, a)$ follows from the constraint $\bar{t} \geq 0$.

Using (176), (177) and (185) in (179) we obtain the desired characterization for the region of interest. For the statement of the result in Lemma 4, we select the smallest $t_0$ that fulfills (179) (which satisfies the condition with equality) and we simplify the notation by using $(\beta, \gamma)$ instead of $(\beta_0, \gamma_0)$.

We emphasize that the condition for Lemma 4 derived in this appendix does not correspond to the region where $f(\beta, \gamma)$ is locally convex, but it precisely characterizes the region where $f(\beta, \gamma) = \underline{f}(\beta, \gamma)$. Figure 12 shows the difference between these two regions for a given set of parameters: the shaded area shows the points where $f(\beta, \gamma)$ is locally non-convex, while the bold line corresponds to the lower-boundary of the region where $f(\beta, \gamma) = \underline{f}(\beta, \gamma)$.

## REFERENCES

[1] G. Vazquez-Vilar, "On the error probability of optimal codes in Gaussian channels under maximal power constraint," in *2019 IEEE Int. Symp. on Inf. Theory*, Paris, France, July 7-12 2019.

[2] ——, "On the error probability of optimal codes in Gaussian channels under average power constraint," in *2020 Int. Zurich Seminar on Inf. and Commun.*, Zurich, Switzerland, Feb. 26-28 2020.

[3] C. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423 and 623–656, July and Oct. 1948.

[4] ——, "Probability of error for optimal codes in a Gaussian channel," *Bell System Technical Journal*, vol. 38, p. 611656, 1959.

[5] Y. Polyanskiy, "Channel coding: non-asymptotic fundamental limits," Ph.D. dissertation, Princeton University, Sep. 2010.

[6] G. Vazquez-Vilar, A. Tauste Campo, A. Guillén i Fàbregas, and A. Martinez, "Bayesian M-ary hypothesis testing: The meta-converse and Verdú-Han bounds are tight," *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2324–2333, May 2016.

[7] C. Shannon, R. Gallager, and E. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels. I," *Information and Control*, vol. 10, no. 1, pp. 65 – 103, 1967.

[8] E. Haroutunian, "Estimates of the error exponents for the semi-continuous memoryless channel (in Russian)," *Probl. Peredachi Inform.*, vol. 4, no. 37-48, 1968.

[9] G. Wiechman and I. Sason, "An improved sphere-packing bound for finite-length codes over symmetric memoryless channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1962–1990, May 2008.

[10] Y. Altuğ and A. B. Wagner, "Refinement of the sphere-packing bound: Asymmetric channels," *IEEE Trans. Inf. Theory*, vol. 60, no. 3, pp. 1592–1614, March 2014.

[11] B. Nakiboğlu, "The sphere packing bound for memoryless channels," *Problems Inf. Trans.*, vol. 56, no. 3, pp. 201–244, July 2020.

[12] S. Verdú, "Error exponents and $\alpha$-mutual information," *MDPI Entropy*, vol. 23, no. 2, 2021.

[13] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.

[14] Y. Polyanskiy, "Saddle point in the minimax converse for channel coding," *IEEE Trans. Inf. Theory*, vol. 59, no. 5, pp. 2576–2595, May 2013.

[15] D. E. Lazic, T. Beth, and S. Egner, "Constrained capacity of the AWGN channel," in *1998 IEEE Int. Symp. on Inf. Theory*, Aug. 1998, p. 237.

[16] S. Dolinar, D. Divsalar, and F. Pollara, "Code performance as a function of block size," *JPL TDA Progress Report*, vol. 42, no. 133, 1998, jet Propulsion Laboratory, Pasadena, CA.

[17] S. Vialle and J. Boutros, "Performance of optimal codes on Gaussian and Rayleigh fading channels: A geometrical approach," in *Proc. 37th Annu. Allerton Conf. Commun., Control Comp.*, Allerton, IL, USA, Sep. 1999, pp. 515–524.

[18] J. Shi and R. D. Wesel, "A study on universal codes with finite block lengths," *IEEE Trans. Inf. Theory*, vol. 53, no. 9, pp. 3066–3074, Sep. 2007.

[19] I. E. Bocharova, A. Guillén i Fàbregas, B. D. Kudryashov, A. Martinez, A. Tauste Campo, and G. Vazquez-Vilar, "Multi-class source-channel coding," *IEEE Trans. Inf. Theory*, vol. 62, no. 9, pp. 5093–5104, Sep. 2016.

[20] D. Slepian, "Bounds on communication," *Bell Syst. Tech. J.*, vol. 42, p. 681707, 1963.

[21] A. Valembois and M. Fossorier, "Sphere-packing bounds revisited for moderate block lengths," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 2998–3014, 2004.

[22] G. Wiechman and I. Sason, "An improved sphere-packing bound for finite-length codes over symmetric memoryless channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1962–1990, May 2008.

[23] I. Sason and S. Shamai (Shitz), *Performance analysis of linear codes under maximum-likelihood decoding: a tutorial*. Foundations and Trends Commun. and Inf. Theory, now Publishers, 2006.

[24] G. Vazquez-Vilar, A. Guillén i Fàbregas, and S. Verdú, "The error probability of generalized perfect codes via the meta-converse," *IEEE Trans. Inf. Theory*, pp. 1–1, 2019.
[25] G. Vazquez-Vilar, A. Guillén i Fàbregas, T. Koch, and A. Lancho, "Saddlepoint approximation of the error probability of binary hypothesis testing," in *2018 IEEE Int. Symp. on Inf. Theory*, June 2018, pp. 2306–2310.
[26] A. Lancho, J. stman, G. Durisi, T. Koch, and G. Vazquez-Vilar, "Saddlepoint approximations for short-packet wireless communications," *IEEE Trans. Wireless Comm.*, vol. 19, no. 7, pp. 4831–4846, 2020.
[27] B. Nakiboğlu, "The Augustin capacity and center," *Problems Inf. Trans.*, vol. 55, no. 4, pp. 299–342, Oct 2019.
[28] R. G. Gallager, *Information Theory and Reliable Communication*. New York: John Wiley & Sons, Inc., 1968.
[29] A. Nuttall, "Some integrals involving the $Q_M$ function," *IEEE Trans. Inf. Theory*, vol. 21, no. 1, pp. 95–96, 1975.
[30] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels*, 2nd ed. New Jersey: Wiley-IEEE Press, 2004.
[31] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. London: Elsevier, 2007.
[32] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, US: Cambridge University Press, 2004.

**Gonzalo Vazquez-Vilar** (S'08–M'12) received the Telecommunication Engineering degree from the University of Vigo, Spain, in 2004, the Master of Science degree from Stanford University, U.S., in 2008 and the Ph.D. in Communication Systems from the University of Vigo, Spain, in 2011.

In 2011-2014 he was a post-doctoral fellow in the Department of Information and Communication Technologies, Universitat Pompeu Fabra, Spain and since 2014 he has been with the Department of Signal Theory and Communications, Universidad Carlos III de Madrid, Spain. He has held appointments as visiting researcher at Stanford University, U.S., University of Cambridge, U.K., and Princeton University, U.S. His research interests lie in the field of Shannon theory, with emphasis on finite-length information theory and communications.