At Low SNR, Asymmetric Quantizers are Better

Tobias Koch*, Member, IEEE*, and Amos Lapidoth*, Fellow, IEEE*

Abstract—We study the capacity of the discrete-time Gaussian channel when its output is quantized with a 1-bit quantizer. We focus on the low signal-to-noise ratio (SNR) regime, where communication at very low spectral efficiencies takes place. In this regime, a symmetric threshold quantizer is known to reduce channel capacity by a factor of $2/\pi$, i.e., to cause an asymptotic power loss of approximately 2 dB. Here, it is shown that this power loss can be avoided by using asymmetric threshold quantizers and asymmetric signaling constellations. To avoid this power loss, flash-signaling input distributions are essential. Consequently, 1-bit output quantization of the Gaussian channel reduces spectral efficiency. Threshold quantizers are not only asymptotically optimal: at every fixed SNR, a threshold quantizer maximizes capacity among all 1-bit output quantizers. The picture changes on the Rayleigh-fading channel. In the noncoherent case, a 1-bit output quantizer causes an unavoidable low-SNR asymptotic power loss. In the coherent case, however, this power loss is avoidable provided that we allow the quantizer to depend on the fading level.

Index Terms—Capacity per unit energy, channel capacity, Gaussian channel, low signal-to-noise ratio (SNR), quantization.

I. INTRODUCTION

WE study the effect on channel capacity of quantizing the output of the discrete-time average-power-limited Gaussian channel using a 1-bit quantizer. This problem arises in communication systems where the receiver uses digital signal processing techniques, which require that the analog received signal be quantized using an analog-to-digital converter (ADC). For ADCs with high resolution, the effects of quantization are negligible. However, high-resolution ADCs may not be practical when the bandwidth of the communication system is large and the sampling rate high [1]. In such scenarios, low-resolution ADCs must be used. The capacity of the discrete-time Gaussian channel with 1-bit output quantization indicates what communication rates can be achieved when the receiver employs a low-resolution ADC.

We focus on the low-signal-to-noise-ratio (SNR) regime, where communication at low spectral efficiencies takes place, as in spread-spectrum and ultrawideband communications. In this regime, a symmetric threshold quantizer[1] reduces the capacity by a factor of $2/\pi$, corresponding to a 2 dB power loss [2]. Hence, the rule of thumb that "hard decisions cause a 2 dB power loss." Here, we demonstrate that if we allow for *asymmetric threshold quantizers* with corresponding *asymmetric signal constellations*, then the two decibels can be fully recovered.

This result shows that a threshold (but not necessarily symmetric) quantizer is asymptotically optimal as the SNR tends to zero. We further show that this is not only true asymptotically: for any fixed SNR, a threshold quantizer is optimal among all 1-bit output quantizers.

While quantizing the output of the Gaussian channel with a 1-bit quantizer does not cause a loss with respect to the low-SNR asymptotic capacity, it does cause a significant loss with respect to the spectral efficiency. Indeed, as we show, the low-SNR asymptotic capacity of the quantized Gaussian channel can only be achieved by *flash-signaling* input distributions [3, Def. 2]. For the Gaussian channel (even without output quantization), such input distributions result in poor spectral efficiency [3, Th. 16]: Gaussian inputs or (at low SNR) binary antipodal inputs yield much higher spectral efficiencies [3, Th. 11]. Since output quantization cannot increase the spectral efficiency, it follows that flash signaling results in poor spectral efficiency also on the quantized Gaussian channel. Thus, at low SNR, the Gaussian channel with optimal 1-bit output quantization has poor spectral efficiency.

It should be noted that the discrete-time channel model that we consider implicitly assumes that the channel output is sampled at Nyquist rate. While sampling the output at Nyquist rate incurs no loss in capacity for the additive white Gaussian noise (AWGN) channel [4], [5], it is not necessarily optimal (with respect to capacity) when the channel output is first quantized using a 1-bit quantizer. In fact, when a symmetric threshold quantizer is employed, sampling the output above the Nyquist rate increases the low-SNR asymptotic capacity [6], [7] and it increases the capacity in the noiseless case [8], [9].

The rest of this paper is organized as follows. Section II introduces the channel model and defines the capacity as well as the capacity per unit energy. Section III presents the paper's main results. Section IV demonstrates that the capacity per unit energy can be achieved by pulse-position modulation (PPM). Section V discusses the implications of our results on the spectral efficiency. Section VI studies the effect on the capacity per unit energy of quantizing the output of the Rayleigh-fading

---

[1]A threshold quantizer produces 1 if its input is above a threshold, and it produces 0 if it is not. A symmetric threshold quantizer is a threshold quantizer whose threshold is zero.
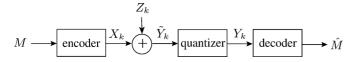
Fig. 1. System model.

channel using a 1-bit quantizer. Sections VII–X contain the proofs of our results: Section VII contains the proofs concerning channel capacity, Section VIII contains the proofs concerning the capacity per unit energy, Section IX contains the proofs concerning peak-power-limited channels, and Section X contains the proofs concerning Rayleigh-fading channels. Section XI concludes this paper with a summary and a discussion.

II. CHANNEL MODEL AND CAPACITY

We consider the discrete-time communication system depicted in Fig. 1. A message M, which is uniformly distributed over the set $\{1, 2, \ldots, \mathcal{M}\}$, is mapped by an encoder to the length-n real sequence $X_1, X_2, \ldots, X_n \in \mathbb{R}$ of channel inputs. (Here, $\mathbb{R}$ denotes the set of real numbers.) The channel corrupts this sequence by adding white Gaussian noise to produce the unquantized output sequence

$$\tilde{Y}_k = X_k + Z_k, \quad k \in \mathbb{Z} \tag{1}$$

where $\{Z_k, \; k \in \mathbb{Z}\}$ is a sequence of independent and identically distributed (i.i.d.) Gaussian random variables of zero mean and variance σ^2. (Here, $\mathbb{Z}$ denotes the set of integers.) The unquantized output sequence is then quantized using a quantizer that is specified by a Borel subset $\mathcal{D}$ of the reals: it produces 1 if $\tilde{Y}_k$ is in $\mathcal{D}$ and produces 0 if it is not. Denoting the time-k quantizer output by Y_k,

$$Y_k = \begin{cases} 1, & \text{if } \tilde{Y}_k \in \mathcal{D}, \\ 0, & \text{if } \tilde{Y}_k \notin \mathcal{D}. \end{cases}$$

While we only consider deterministic quantizers, it should be noted that our results continue to hold if we allow for randomized quantization rules, i.e., if the quantizer produces Y_k according to some probability distribution $P_{Y|\tilde{Y}}$ with binary Y. In view of the direct relationship between the set $\mathcal{D}$ and the quantizer it defines, we shall sometimes abuse notation and refer to $\mathcal{D}$ as the quantizer. An example of a 1-bit quantizer is the *threshold quantizer*, which corresponds to the set

$$\mathcal{D} = \{\tilde{y} \in \mathbb{R} : \tilde{y} \geq \Upsilon\}, \quad \Upsilon \in \mathbb{R}. \tag{2}$$

The decoder observes the quantizer's outputs $Y_1, Y_2, \ldots, Y_n$ and guesses which message was transmitted.

We impose an average-power constraint on the transmitted sequence: for every realization of the message M, the sequence $x_1, x_2, \ldots, x_n$ must satisfy

$$\frac{1}{n} \sum_{k=1}^{n} x_k^2 \leq \mathcal{P} \tag{3}$$

for some positive constant $\mathcal{P}$, which we call the *maximal allowed average power*.

For a fixed quantizer $\mathcal{D}$ and maximal allowed average power $\mathcal{P}$, the capacity $C(\mathcal{P}, \mathcal{D})$ is [5], [10]

$$C(\mathcal{P}, \mathcal{D}) = \sup_{\mathsf{E}[X^2] \leq \mathcal{P}} I(X; Y) \tag{4}$$

where the supremum is over all distributions of X under which the second moment of X does not exceed $\mathcal{P}$. Here, and throughout the paper, we omit the time indices where they are immaterial. We say that a *rate* R (in nats per channel use) is *achievable* using power $\mathcal{P}$ and 1-bit quantization if for every $\epsilon > 0$, there exists an encoder satisfying (3) and

$$\frac{\log \mathcal{M}}{n} > R - \epsilon \tag{5}$$

as well as a 1-bit quantizer and a decoder such that the probability of error $\Pr(\hat{M} \neq M)$ tends to zero as n tends to infinity. Here, $\log(\cdot)$ denotes the natural logarithm function. The *capacity* $C(\mathcal{P})$ is the supremum of all achievable rates and is given by

$$C(\mathcal{P}) = \sup_{\mathcal{D}} C(\mathcal{P}, \mathcal{D}) \tag{6}$$

$$= \sup_{\mathcal{D}, \mathsf{E}[X^2] \leq \mathcal{P}} I(X; Y) \tag{7}$$

where the first supremum is over all quantization regions $\mathcal{D}$, and the second supremum is over all quantization regions $\mathcal{D}$ and over all distributions of X satisfying $\mathsf{E}[X^2] \leq \mathcal{P}$.

Following [11], we define the capacity per unit energy of the quantizer $\mathcal{D}$ as follows: We say that a *rate per unit energy* $\dot{R}(0, \mathcal{D})$ (in nats per energy) is *achievable* with the quantizer $\mathcal{D}$ if for every $\epsilon > 0$, there exists an encoder satisfying

$$\sum_{k=1}^{n} x_k^2 \leq \mathcal{E}, \quad \text{for every realization of } M \tag{8}$$

and

$$\frac{\log \mathcal{M}}{\mathcal{E}} > \dot{R}(0, \mathcal{D}) - \epsilon \tag{9}$$

together with a decoder such that the probability of error $\Pr(\hat{M} \neq M)$ tends to zero as $\mathcal{E}$ tends to infinity. The *capacity per unit energy* $\dot{C}(0, \mathcal{D})$ is the supremum of all achievable rates per unit energy with the quantizer $\mathcal{D}$ and is given by [11, Th. 2]

$$\dot{C}(0, \mathcal{D}) = \sup_{\mathcal{P} > 0} \frac{C(\mathcal{P}, \mathcal{D})}{\mathcal{P}} \tag{10}$$

$$= \lim_{\mathcal{P} \downarrow 0} \frac{C(\mathcal{P}, \mathcal{D})}{\mathcal{P}} \tag{11}$$

where the second equation follows because, for every $\mathcal{D}$, the capacity $C(\mathcal{P}, \mathcal{D})$ is a concave function of $\mathcal{P}$.

The definition of capacity per unit energy using a 1-bit quantizer is analogous: We say that a *rate per unit energy* $\dot{R}(0)$ (in nats per energy) is *achievable* using a 1-bit quantizer if for every $\epsilon > 0$, there exists an encoder satisfying (8) and

$$\frac{\log \mathcal{M}}{\mathcal{E}} > \dot{R}(0) - \epsilon \tag{12}$$

as well as a 1-bit quantizer and a decoder such that the probability of error $\Pr(\hat{M} \neq M)$ tends to zero as $\mathcal{E}$ tends to infinity. The *capacity per unit energy* $\dot{C}(0)$ is the supremum of all achievable rates per unit energy.

Extending the proof of Theorem 2 in [11] to account for the additional maximization over all possible quantizers, we obtain

$$\dot{C}(0) = \sup_{\mathcal{P} > 0} \frac{C(\mathcal{P})}{\mathcal{P}} \tag{13}$$

which, by (6), can be expressed as

$$\dot{C}(0) = \sup_{\mathcal{P} > 0} \sup_{\mathcal{D}} \frac{C(\mathcal{P}, \mathcal{D})}{\mathcal{P}}. \tag{14}$$

Exchanging the order of the suprema and applying (10) yields

$$\dot{C}(0) = \sup_{\mathcal{D}} \dot{C}(0, \mathcal{D}) \tag{15}$$

$$= \sup_{\xi \neq 0, \mathcal{D}} \frac{D(P_{Y|X=\xi} \| P_{Y|X=0})}{\xi^2} \tag{16}$$

where the last step follows from [3, Th. 3]. Here, $D(\cdot\|\cdot)$ denotes relative entropy

$$D(P\|Q) \triangleq \begin{cases} \int \log\left(\frac{\mathrm{d}P}{\mathrm{d}Q}\right) \mathrm{d}P, & \text{if } P \ll Q \\ \infty, & \text{otherwise} \end{cases} \tag{17}$$

(where $P \ll Q$ indicates that P is absolutely continuous with respect to Q), and $P_{Y|X=x}$ denotes the output distribution corresponding to the input x. In our case, since the output of the quantizer is binary,

$$D(P_{Y|X=\xi} \| P_{Y|X=0})$$
$$= \Pr(\tilde{Y} \in \mathcal{D} \mid X = \xi) \log \frac{\Pr(\tilde{Y} \in \mathcal{D} \mid X = \xi)}{\Pr(\tilde{Y} \in \mathcal{D} \mid X = 0)}$$
$$+ \Pr(\tilde{Y} \notin \mathcal{D} \mid X = \xi) \log \frac{\Pr(\tilde{Y} \notin \mathcal{D} \mid X = \xi)}{\Pr(\tilde{Y} \notin \mathcal{D} \mid X = 0)}. \tag{18}$$

It follows from (6) and (11) that

$$\lim_{\mathcal{P} \downarrow 0} \frac{C(\mathcal{P})}{\mathcal{P}} = \lim_{\mathcal{P} \downarrow 0} \sup_{\mathcal{D}} \frac{C(\mathcal{P}, \mathcal{D})}{\mathcal{P}}$$
$$\geq \sup_{\mathcal{D}} \dot{C}(0, \mathcal{D}) \tag{19}$$

which, together with (13) and (15), yields

$$\dot{C}(0) = \lim_{\mathcal{P} \downarrow 0} \frac{C(\mathcal{P})}{\mathcal{P}}. \tag{20}$$

Thus, the capacity per unit energy is equal to the slope at zero of the capacity-versus-power curve.

By the Data Processing Inequality [10, Th. 2.8.1], $C(\mathcal{P}, \mathcal{D})$ is upper-bounded by the capacity of the unquantized channel [4]

$$C(\mathcal{P}, \mathcal{D}) \leq \frac{1}{2} \log\left(\frac{1 + \mathcal{P}}{\sigma^2}\right). \tag{21}$$

Consequently, by (11) and (15),

$$\dot{C}(0, \mathcal{D}) \leq \frac{1}{2\sigma^2} \quad \text{and} \quad \dot{C}(0) \leq \frac{1}{2\sigma^2}. \tag{22}$$

A ubiquitous quantizer is the *symmetric threshold quantizer*, for which $\mathcal{D} = \{\tilde{y} \in \mathbb{R} : \tilde{y} \geq 0\}$. For this quantizer, the capacity $C_{\text{sym}}(\mathcal{P})$ is given by [2, eq. (3.4.18)], [12, Th. 2]

$$C_{\text{sym}}(\mathcal{P}) = \log 2 - H_b\left(Q\left(\frac{\sqrt{\mathcal{P}}}{\sigma^2}\right)\right) \tag{23}$$

where $H_b(\cdot)$ denotes the binary entropy function

$$H_b(p) \triangleq -p \log p - (1-p) \log(1-p), \quad 0 \leq p \leq 1 \tag{24}$$

(where we define $0 \log 0 \triangleq 0$) and $Q(\cdot)$ denotes the Q-function

$$Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{t^2}{2}} \mathrm{d}t, \quad x \in \mathbb{R}. \tag{25}$$

The capacity $C_{\text{sym}}(\mathcal{P})$ can be achieved by transmitting $\sqrt{\mathcal{P}}$ and $-\sqrt{\mathcal{P}}$ equiprobably.

From (23), the capacity per unit energy $\dot{C}_{\text{sym}}(0)$ for a symmetric threshold quantizer is [2, eq. (3.4.20)]

$$\dot{C}_{\text{sym}}(0) = \lim_{\mathcal{P} \downarrow 0} \frac{C_{\text{sym}}(\mathcal{P})}{\mathcal{P}} = \frac{1}{\pi\sigma^2}. \tag{26}$$

This is a factor of $2/\pi$ smaller than the capacity per unit energy $1/(2\sigma^2)$ of the Gaussian channel without output quantization. Thus, quantizing the channel output using a symmetric threshold quantizer causes a loss of roughly 2 dB.

It is tempting to attribute this loss to the fact that the quantizer forces the decoder to perform only hard-decision decoding. However, as we shall see, the loss of 2 dB is not a consequence of the hard-decision decoder but of the suboptimal quantizer. In fact, with an asymmetric threshold quantizer, the loss vanishes (see Theorem 2).

III. MAIN RESULTS

Our main results are presented in the following two sections. Section III-A presents the results on channel capacity. We show that the capacity-achieving input distribution is discrete with at most three mass points and that threshold quantizers achieve capacity (see Theorem 1). Furthermore, we provide an expression for the capacity when the average-power constraint (3) is replaced by a peak-power constraint (see Proposition 1).

Section III-B presents the results on capacity per unit energy. We show that asymmetric threshold quantizers and asymmetric signal constellations can achieve the capacity per unit energy of the Gaussian channel (see Theorem 2), thus demonstrating that quantizing the output of the Gaussian channel with a 1-bit quantizer does not cause an asymptotic power loss. We further demonstrate that, in order to achieve this capacity per unit energy, *flash-signaling* input distributions [3, Def. 2] are required (see Theorem 3). Finally, we show that if the average-power constraint (3) is replaced by a peak-power constraint, then quantizing the output of the Gaussian channel with a 1-bit quantizer necessarily causes a 2 dB power loss (see Proposition 2).

A. Channel Capacity

Theorem 1 (Optimal Input Distribution and Quantizer):
1) For any given maximal allowed average power $\mathcal{P}$ and any Borel set $\mathcal{D}$, the supremum in (4) defining $C(\mathcal{P}, \mathcal{D})$ is achieved by some input distribution that is concentrated on at most three points.
2) For any given maximal allowed average power $\mathcal{P}$, the supremum in (7) is achieved by some threshold quantizer

$$\mathcal{D}^\star = \{\tilde{y} \in \mathbb{R} : \tilde{y} \geq \Upsilon\}$$

(where $\Upsilon \geq 0$ depends on $\mathcal{P}$ and σ^2) and by a zero-mean, variance-$\mathcal{P}$, input distribution that is concentrated on at most three points.

Proof: See Section VII.   ∎

The result that the capacity-achieving input distribution is concentrated on at most three mass points is consistent with [12, Th. 1], which shows that if the quantization regions of a K-bit quantizer partition the real line into 2^K intervals, then the capacity-achieving input distribution is concentrated on at most $2^K + 1$ points.

Proposition 1: If the average-power constraint (3) is replaced by the peak-power constraint

$$X_k^2 \leq \mathcal{P}, \quad k \in \mathbb{Z}, \quad \text{with probability one} \qquad (27)$$

then the capacity of the channel presented in Section II is given by

$$C_{\text{PP}}(\mathcal{P}) = \max_{\Upsilon \geq 0} \left\{ \log\left(1 + e^{-\Theta(\mathcal{P},\Upsilon)}\right) \right.$$
$$\left. + Q\left(\frac{\sqrt{\mathcal{P}}+\Upsilon}{\sigma}\right) \Theta(\mathcal{P},\Upsilon) - H_b\left(Q\left(\frac{\sqrt{\mathcal{P}}+\Upsilon}{\sigma}\right)\right) \right\} \qquad (28)$$

where

$$\Theta(\mathcal{P},\Upsilon) \triangleq \frac{H_b\left(Q\left(\frac{\sqrt{\mathcal{P}}-\Upsilon}{\sigma}\right)\right) - H_b\left(Q\left(\frac{\sqrt{\mathcal{P}}+\Upsilon}{\sigma}\right)\right)}{1 - Q\left(\frac{\sqrt{\mathcal{P}}-\Upsilon}{\sigma}\right) - Q\left(\frac{\sqrt{\mathcal{P}}+\Upsilon}{\sigma}\right)}. \qquad (29)$$

The capacity can be achieved by a binary input distribution with mass points at $\sqrt{\mathcal{P}}$ and $-\sqrt{\mathcal{P}}$ and by some threshold quantizer with threshold $\Upsilon \geq 0$.

Proof: See Section IX-A.   ∎

Numerical evaluation of (28) suggests that, for every maximal allowed peak power $\mathcal{P}$, the maximum is attained for $\Upsilon = 0$. In this case, $C_{\text{PP}}(\mathcal{P})$ would specialize to the capacity of the average-power-limited Gaussian channel with symmetric output quantization (23).

B. Capacity Per Unit Energy

Theorem 2 ($\dot{C}(0) = 1/(2\sigma^2)$): The capacity per unit energy of the channel presented in Section II is

$$\dot{C}(0) = \frac{1}{2\sigma^2}. \qquad (30)$$

Proof: See Section VIII-A.   ∎

Thus, if we allow for asymmetric threshold quantizers and asymmetric signal constellations, then quantizing the output of the average-power-limited Gaussian channel with an optimal 1-bit quantizer does not cause a loss with respect to the capacity per unit energy.

Considering the symmetry of the probability density function (PDF) of the Gaussian noise, it is perhaps surprising that an asymmetric quantizer yields a larger rate per unit energy than a symmetric one. However, the input distribution achieving (30) is asymmetric (see below). Hence, the PDF of the unquantized channel output is asymmetric, so it seems plausible that the capacity per unit energy is achieved by some asymmetric quantizer. In fact, even if the PDF of the unquantized channel

output were symmetric, this would not necessarily imply that the optimal quantizer is symmetric: There are examples in the source-coding literature of symmetric PDFs for which the optimal 1-bit quantizer with respect to the mean squared error is asymmetric (see, e.g., [13, Ex. 5.2, pp. 64-65]).

Theorem 2 is proved by analyzing (16) with a judicious choice of $\mathcal{D}$ and ξ. In Section IV, we provide an alternative proof by presenting a PPM scheme that achieves the capacity per unit energy (30). For this scheme, the error probability can be analyzed directly using the union bound and an upper bound on the Q-function: there is no need to resort to conventional methods used to prove coding theorems such as the method of types, information-spectrum methods, or random coding exponents.

The capacity per unit energy (30) can be achieved by binary on–off keying, i.e., by binary inputs of probability mass function

$$P(X = \xi) = 1 - P(X = 0) = \frac{\mathcal{P}}{\xi^2}, \quad \xi^2 \geq \mathcal{P} \qquad (31)$$

where the nonzero mass point ξ tends to infinity as $\mathcal{P}$ tends to zero. The distribution of such inputs belongs to the class of *flash-signaling* input distributions, which was defined by Verdú [3, Def. 2] as follows.

Definition 1 (Flash Signaling): A family of distributions of X parametrized by $\mathcal{P}$ is said to be *flash signaling* if it satisfies $\mathsf{E}\left[X^2\right] \leq \mathcal{P}$ and for every positive ν

$$\lim_{\mathcal{P}\downarrow 0} \frac{\mathsf{E}\left[X^2 \mathsf{I}\left\{X^2 > \nu\right\}\right]}{\mathcal{P}} = 1. \qquad (32)$$

Here, $\mathsf{I}\{\text{statement}\}$ denotes the indicator function: it is equal to one if the statement between the curly brackets is true and is equal to zero otherwise.

Flash signaling is described in [3] as "the mixture of a probability distribution that asymptotically concentrates its mass at 0 and a probability distribution that migrates to infinity; the weight of the latter vanishes sufficiently fast to satisfy the vanishing power constraint." The next theorem shows that flash signaling is necessary to achieve (30).

Theorem 3 (Flash Signaling is Required to Achieve $\dot{C}(0)$): Every family of distributions of X parametrized by $\mathcal{P}$ that satisfies $\mathsf{E}\left[X^2\right] \leq \mathcal{P}$ and

$$\lim_{\mathcal{P}\downarrow 0} \frac{I(X;Y)}{\mathcal{P}} = \frac{1}{2\sigma^2} \qquad (33)$$

must be flash signaling.

Proof: See Section VIII-B.   ∎

It is easy to show that for flash-signaling input distributions, threshold quantizers with a bounded threshold give rise to zero rate per unit energy. We thus have the following corollary.

Corollary 1 (The Thresholds Must be Unbounded): If (33) holds for some family of threshold quantizers (parametrized by the average power), then the thresholds must be unbounded in the average power.

Proof: See Section VIII-C.   ∎

Intuitively, the power loss in quantizing the output of the Gaussian channel with a 1-bit quantizer can be avoided by using flash-signaling input distributions and asymmetric threshold

quantizers because for such input distributions and quantizers, the probability that the quantizer causes an error vanishes as the SNR tends to zero. Indeed, by using binary on–off keying (31) and threshold quantizers (2), and by cleverly choosing the rate at which ξ and Υ grow as $\mathcal{P}$ decreases, we can make the probabilities $\Pr(Y = 1|X = 0)$ and $\Pr(Y = 0|X = \xi)$ vanish as $\mathcal{P}$ tends to zero. This suggests that the loss caused by the quantizer disappears with decreasing $\mathcal{P}$. Note, however, that the same argument would also apply to the averaged-power-limited, noncoherent, Rayleigh-fading channel (see Section VI), but for this channel, quantizing the output with a 1-bit quantizer does cause a loss with respect to the capacity per unit energy (see Theorem 5).

As mentioned in Section II, the capacity per unit energy is equal to the slope at zero of the capacity-versus-power curve. Thus, Theorem 2 demonstrates that the first derivative of $C(\mathcal{P})$ at $\mathcal{P} = 0$ is equal to $1/(2\sigma^2)$. Theorem 3 implies that the second derivative of $C(\mathcal{P})$ at $\mathcal{P} = 0$ is $-\infty$.

Corollary 2 ($\ddot{C}(0) = -\infty$):

$$\ddot{C}(0) = 2 \lim_{\mathcal{P} \downarrow 0} \frac{C(\mathcal{P}) - \mathcal{P} \dot{C}(0)}{\mathcal{P}^2} = -\infty. \tag{34}$$

Proof: By the Data Processing Inequality, for every family of distributions of X parametrized by $\mathcal{P}$,

$$\lim_{\mathcal{P} \downarrow 0} \frac{I(X; Y) - \frac{\mathcal{P}}{2\sigma^2}}{\mathcal{P}^2} \leq \lim_{\mathcal{P} \downarrow 0} \frac{I(X; \tilde{Y}) - \frac{\mathcal{P}}{2\sigma^2}}{\mathcal{P}^2}. \tag{35}$$

To achieve $\dot{C}(0)$, it is necessary to use flash signaling (see Theorem 3). And for all flash-signaling input distributions, the right-hand side (RHS) of (35) is $-\infty$ (see [3, Th. 16]). Consequently, so is its left-hand side (LHS). ∎

Note that, for the Gaussian channel, the first and second derivatives of the capacity are [4]

$$\dot{C}_{\mathrm{G}}(0) = \frac{1}{2\sigma^2} \quad \text{and} \quad \ddot{C}_{\mathrm{G}}(0) = -\frac{1}{2\sigma^4} \tag{36}$$

where "G" stands for "Gaussian". Thus, while quantizing the output of the Gaussian channel with a 1-bit quantizer does not cause a loss with respect to the first derivative of the capacity-versus-power curve, it causes a substantial loss in terms of the second derivative. The implications on the spectral efficiency are discussed in Section V.

Proposition 2: If the average-power constraint (3) is replaced by the peak-power constraint

$$X_k^2 \leq \mathcal{P}, \quad k \in \mathbb{Z}, \quad \text{with probability one} \tag{37}$$

then the slope at zero of the capacity-versus-power curve is

$$\lim_{\mathcal{P} \downarrow 0} \frac{C_{\mathrm{PP}}(\mathcal{P})}{\mathcal{P}} = \frac{1}{\pi\sigma^2}. \tag{38}$$

Proof: See Section IX-B. ∎

As was shown by Shannon [4], the capacity of the peak-power-limited unquantized Gaussian channel satisfies

$$\lim_{\mathcal{P} \downarrow 0} \frac{C_{\mathrm{G,PP}}(\mathcal{P})}{\mathcal{P}} = \frac{1}{2\sigma^2}. \tag{39}$$

Thus, in contrast to the average-power-limited case, quantizing the output of the peak-power-limited Gaussian channel with a 1-bit quantizer does cause a 2-dB power loss.

IV. PULSE-POSITION MODULATION

We next demonstrate that the capacity per unit energy (30) can be achieved using a PPM scheme—no random-coding arguments are needed. For such a scheme, the encoder produces the $\mathcal{M}$ channel inputs $x_1(m), x_2(m), \ldots, x_{\mathcal{M}}(m)$ for each message m in $\{1, 2, \ldots, \mathcal{M}\}$, where

$$x_k(m) = \begin{cases} \xi, & \text{if } k = m, \\ 0, & \text{if } k \neq m, \end{cases} \quad k = 1, 2, \ldots, \mathcal{M} \tag{40}$$

and where $\xi^2 = \mathcal{E}$. For a fixed rate per unit energy

$$\dot{R}(0) = \frac{\log \mathcal{M}}{\mathcal{E}}$$

we have

$$\xi^2 = \mathcal{E} = \frac{\log \mathcal{M}}{\dot{R}(0)}. \tag{41}$$

Note that, while the *rate per unit energy* is fixed, the *rate* of this scheme is $\frac{\log \mathcal{M}}{\mathcal{M}}$ and tends to zero as $\mathcal{M}$ tends to infinity.

We employ a threshold quantizer (2) with the threshold Υ chosen so that for an arbitrary $0 < \epsilon < 1$, the probability that the quantizer produces 0 given that $X = \xi$ is equal to ϵ. Thus,

$$\Upsilon = \xi - \sigma Q^{-1}(\epsilon) \tag{42}$$

which yields

$$P(Y_k = 0 \mid X_k = \xi) = \epsilon \tag{43a}$$

$$P(Y_k = 1 \mid X_k = 0) = Q\left(\frac{\xi - \sigma Q^{-1}(\epsilon)}{\sigma}\right). \tag{43b}$$

In (42), $Q^{-1}(\cdot)$ denotes the inverse Q-function.

The decoder guesses "$\hat{M} = m$" provided that $Y_m = 1$ and that $Y_k = 0$ for all $k \neq m$. If $Y_k = 1$ for more than one k, or if $Y_k = 0$ for all $k = 1, 2, \ldots, \mathcal{M}$, then the decoder declares an error.

Suppose that message $M = m$ was transmitted. Then, the probability of an error is upper-bounded by

$$\Pr(\hat{M} \neq M \mid M = m)$$
$$= \Pr\left(\left.\bigcup_{k \neq m} (Y_k = 1) \cup (Y_m = 0) \right| M = m\right)$$
$$\leq \sum_{k \neq m} P(Y_k = 1 \mid X_k = 0) + P(Y_m = 0 \mid X_m = \xi)$$
$$= \sum_{k \neq m} P(Y_k = 1 \mid X_k = 0) + \epsilon$$
$$= (\mathcal{M} - 1) P(Y_1 = 1 \mid X_1 = 0) + \epsilon \tag{44}$$

where the second step follows from the Union Bound; the third step follows from (43a); and the fourth step follows because the channel is memoryless which implies that $\Pr(Y_k = 1|X_k = 0)$

does not depend on k. Since the RHS of (44) does not depend on m, it follows that also the probability of error

$$\Pr(\hat{M} \neq M) = \frac{1}{\mathcal{M}} \sum_{m=1}^{\mathcal{M}} \Pr(\hat{M} \neq M \mid M = m)$$

is upper-bounded by (44).

The first term on the RHS of (44) can be evaluated using (43b) and (41):

$$(\mathcal{M} - 1) P(Y_1 = 1 \mid X_1 = 0)$$
$$= (\mathcal{M} - 1) Q\left(\frac{\xi - \sigma Q^{-1}(\epsilon)}{\sigma}\right)$$
$$= (\mathcal{M} - 1) Q\left(\frac{\sqrt{\log \mathcal{M}} - \sigma Q^{-1}(\epsilon)\sqrt{\dot{R}(0)}}{\sigma\sqrt{\dot{R}(0)}}\right). \quad (45)$$

We continue by showing that if

$$\dot{R}(0) < \frac{1}{2\sigma^2}$$

then, for every fixed $0 < \epsilon < 1$, the RHS of (45) tends to zero as $\mathcal{M}$ tends to infinity. Indeed,

$$\lim_{\mathcal{M} \to \infty} (\mathcal{M} - 1) Q\left(\frac{\sqrt{\log \mathcal{M}} - \sigma Q^{-1}(\epsilon)\sqrt{\dot{R}(0)}}{\sigma\sqrt{\dot{R}(0)}}\right)$$
$$\leq \lim_{\alpha \to \infty} \exp\left(\sigma^2 \dot{R}(0)\left(\alpha + Q^{-1}(\epsilon)\right)^2\right) Q(\alpha)$$
$$\leq \lim_{\alpha \to \infty} \frac{1}{\sqrt{2\pi}\alpha} \exp\left(\sigma^2 \dot{R}(0)\left(\alpha + Q^{-1}(\epsilon)\right)^2 - \frac{1}{2}\alpha^2\right)$$
$$(46)$$

where the first step follows by upper-bounding $\mathcal{M} - 1 < \mathcal{M}$ and by substituting

$$\alpha = \frac{\sqrt{\log \mathcal{M}} - \sigma Q^{-1}(\epsilon)\sqrt{\dot{R}(0)}}{\sigma\sqrt{\dot{R}(0)}}$$

and the second step follows from the inequality [14, Prop. 19.4.2]

$$Q(\alpha) < \frac{1}{\sqrt{2\pi}\alpha} e^{-\alpha^2/2}, \quad \alpha > 0. \quad (47)$$

The RHS of (46) is zero for $\dot{R}(0) < \frac{1}{2\sigma^2}$.

Combining (46) with (44), we obtain that if $\dot{R}(0) < \frac{1}{2\sigma^2}$, then the probability of error tends to ϵ as $\mathcal{E}$—and hence, by (41), also $\mathcal{M}$—tends to infinity. Since ϵ can be chosen arbitrarily small, the probability of error can be made arbitrarily small, thus proving that the capacity per unit energy (30) is achievable with the above PPM scheme.

The fact that PPM achieves the capacity per unit energy of the Gaussian channel with a threshold quantizer follows also from the analysis of the probability of error for block orthogonal signals shown in [15, pp. 342–346]. The threshold $a \geq 0$ introduced to bound the RHS of (5.97 d) in [15] can be identified as the threshold Υ of the quantizer.

V. SPECTRAL EFFICIENCY

The discrete-time channel presented in Section II is closely related to the continuous-time AWGN channel with 1-bit output quantization. Indeed, suppose that the input to the latter channel is bandlimited to W Hz and that its average power is limited by $\mathcal{P}$, and suppose that the Gaussian noise is of double-sided power spectral density $\mathcal{N}_0/2$. Then, the discrete-time channel (1) with noise variance

$$\sigma^2 = W\mathcal{N}_0 \quad (48)$$

results from sampling the AWGN channel's output at the Nyquist rate $2W$. The capacity (in bits per second) of the AWGN channel with Nyquist sampling and 1-bit output quantization is given by

$$C_{\text{AWGN}}^{(2W)}(\mathcal{P}) = \frac{2W}{\log 2} C(\mathcal{P}) \quad (49)$$

where $C(\mathcal{P})$ is the capacity (7) of the discrete-time channel in nats per channel use. Note, however, that when the channel output is quantized, sampling at the Nyquist rate need not be optimal with respect to capacity: see, e.g., [6]–[9] for scenarios where sampling the quantizer's output above the Nyquist rate provides capacity gains. Consequently, $C_{\text{AWGN}}^{(2W)}(\mathcal{P})$ is, in general, a lower bound on the capacity of the AWGN channel with 1-bit output quantization.

The energy per information bit when communicating with power $\mathcal{P}$ at rate $C_{\text{AWGN}}^{(2W)}(\mathcal{P})$ is defined as

$$\frac{\mathcal{E}_{\text{b}}}{\mathcal{N}_0} \triangleq \frac{\mathcal{P}}{C_{\text{AWGN}}^{(2W)}(\mathcal{P})} \frac{1}{\mathcal{N}_0} \quad (50)$$

which, by (48) and (49), is equal to

$$\frac{\mathcal{E}_{\text{b}}}{\mathcal{N}_0} = \frac{\log 2}{2\sigma^2} \frac{\mathcal{P}}{C(\mathcal{P})}. \quad (51)$$

The spectral efficiency $\bar{C}(\cdot)$ (in bits per second per hertz) is defined as

$$\bar{C}\left(\frac{\mathcal{E}_{\text{b}}}{\mathcal{N}_0}\right) \triangleq \frac{C_{\text{AWGN}}^{(2W)}(\mathcal{P})}{W} \quad (52)$$

which, by (49), is

$$\bar{C}\left(\frac{\mathcal{E}_{\text{b}}}{\mathcal{N}_0}\right) = \frac{2}{\log 2} C(\mathcal{P}). \quad (53)$$

In (52) and (53), $\mathcal{P}$ is the solution to (50), namely,

$$\frac{\mathcal{E}_{\text{b}}}{\mathcal{N}_0} = \frac{\mathcal{P}}{C_{\text{AWGN}}^{(2W)}(\mathcal{P})} \frac{1}{\mathcal{N}_0}. \quad (54)$$

See [3] for a more thorough discussion of spectral efficiency. (Note that, in contrast to (1), the channel considered in [3] is complex-valued. Therefore, the expressions for $\mathcal{E}_{\text{b}}/\mathcal{N}_0$ and $\bar{C}(\mathcal{E}_{\text{b}}/\mathcal{N}_0)$ differ by a factor of two.)

The *minimum* $\mathcal{E}_b/\mathcal{N}_0$ required for reliable communication is determined by taking the infimum over $\mathcal{P}$ of the RHS of (51). By (13), this yields [3, eq. (35)]

$$\left(\frac{\mathcal{E}_b}{\mathcal{N}_0}\right)_{\min} = \frac{\log 2}{2\sigma^2} \frac{1}{\dot{C}(0)}. \tag{55}$$

Furthermore, the slope of $\mathcal{E}_b/\mathcal{N}_0 \mapsto \bar{C}(\mathcal{E}_b/\mathcal{N}_0)$ at $(\mathcal{E}_b/\mathcal{N}_0)_{\min}$ in bits per second per hertz per 3 dB is given by [3, Th. 9][2]

$$\mathcal{S}_0 = \frac{4\left[\dot{C}(0)\right]^2}{-\ddot{C}(0)}. \tag{56}$$

By (30) and (34), we have for the average-power-limited Gaussian channel with 1-bit output quantization

$$\dot{C}(0) = \frac{1}{2\sigma^2} \quad \text{and} \quad \ddot{C}(0) = -\infty \tag{57}$$

which yields

$$\left(\frac{\mathcal{E}_b}{\mathcal{N}_0}\right)_{\min} = \log 2 = -1.59 \text{ dB} \tag{58a}$$

$$\mathcal{S}_0 = 0 \ \frac{\text{bps/Hz}}{3 \text{ dB}}. \tag{58b}$$

In comparison, for the unquantized Gaussian channel (36)

$$\dot{C}_G(0) = \frac{1}{2\sigma^2} \quad \text{and} \quad \ddot{C}_G(0) = -\frac{1}{2\sigma^4} \tag{59}$$

and for the Gaussian channel with *symmetric* 1-bit output quantization (23)

$$\dot{C}_{\text{sym}}(0) = \frac{1}{\pi\sigma^2} \quad \text{and} \quad \ddot{C}_{\text{sym}}(0) = \frac{2}{3\pi\sigma^4}\left(\frac{1}{\pi} - 1\right). \tag{60}$$

This yields

$$\left(\frac{\mathcal{E}_b}{\mathcal{N}_0}\right)_{\min,G} = \log 2 = -1.59 \text{ dB} \tag{61a}$$

$$\mathcal{S}_{0,G} = 2 \ \frac{\text{bps/Hz}}{3 \text{ dB}} \tag{61b}$$

and

$$\left(\frac{\mathcal{E}_b}{\mathcal{N}_0}\right)_{\min,\text{sym}} = \frac{\pi}{2}\log 2 = 0.37 \text{ dB} \tag{62a}$$

$$\mathcal{S}_{0,\text{sym}} = \frac{6}{\pi - 1} = 2.8 \ \frac{\text{bps/Hz}}{3 \text{ dB}}. \tag{62b}$$

Comparing (62a) with (61a), we see once more that quantizing the output of the Gaussian channel with a symmetric threshold quantizer causes a power loss of roughly 2 dB. We further see that with an asymmetric threshold quantizer, we can recover the

[2]Again, the channel considered in [3] is complex-valued and the expressions for $(\mathcal{E}_b/\mathcal{N}_0)_{\min}$ and $\mathcal{S}_0$, therefore, differ by a factor of two. Nevertheless, since the capacity of the complex-valued channel is twice the capacity of the real-valued channel, it follows that the numerical values of $(\mathcal{E}_b/\mathcal{N}_0)_{\min}$ and $\mathcal{S}_0$ are the same as in [3].
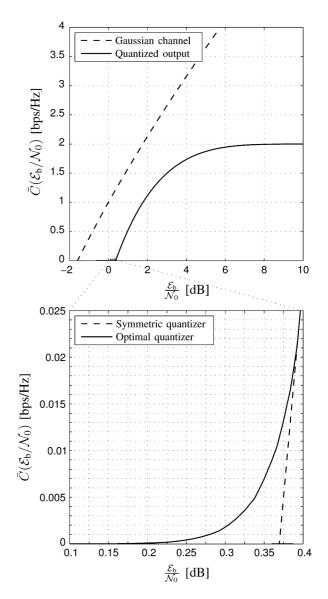

Fig. 2. Spectral efficiency versus energy per information bit. (Top) Spectral efficiencies of the Gaussian channel with and without 1-bit output quantization. (Bottom) Spectral efficiencies for the optimal 1-bit quantizer and for the symmetric threshold quantizer.

loss in terms of $(\mathcal{E}_b/\mathcal{N}_0)_{\min}$, but there is still a substantial loss in terms of spectral efficiency. Indeed, for the Gaussian channel with 1-bit output quantization, the wideband slope $\mathcal{S}_0$ is zero, whereas for the unquantized Gaussian channel it is 2 bits per second per Hz per 3 dB.

The above spectral efficiencies are shown in Fig. 2. The top subfigure shows the spectral efficiencies of the Gaussian channel with and without 1-bit output quantization. The bottom subfigure compares the spectral efficiency $\bar{C}(\cdot)$ for the optimal 1-bit quantizer with the spectral efficiency $\bar{C}_{\text{sym}}(\cdot)$ for the symmetric threshold quantizer. We observe that, even though the minimum energy per information bit is the same with and without 1-bit output quantization,[3] the corresponding spectral efficiencies differ substantially for all $\mathcal{E}_b/\mathcal{N}_0$. We further

[3]For numerical reasons, the spectral efficiency of the Gaussian channel with 1-bit output quantization can only be shown for $\mathcal{E}_b/\mathcal{N}_0$ above -0.5 dB.

observe that for spectral efficiencies above 0.02 bits/s/Hz, a symmetric threshold quantizer is nearly optimal.

We conclude that, for communication systems that operate at very low spectral efficiencies, asymmetric quantizers are beneficial, although for most practical scenarios, the potential power gain is significantly smaller than 2 dB. For example, at a spectral efficiency of 0.001 bits/s/Hz, allowing for asymmetric quantizers with corresponding asymmetric signal constellations provides a power gain of roughly 0.1 dB.

VI. ONE-BIT QUANTIZERS FOR FADING CHANNELS

For the average-power-limited (real-valued) Gaussian channel, we have demonstrated that by allowing for asymmetric threshold quantizers with corresponding asymmetric signal constellations, one can achieve the capacity per unit energy of the unquantized channel. The same holds for the average-power-limited *complex-valued* Gaussian channel [16]: using binary on–off keying (31) and a *radial quantizer* (which produces 1 if the magnitude of the channel output is above some threshold and produces 0 otherwise), one can achieve the capacity per unit energy of the unquantized channel by judiciously choosing the threshold and the nonzero mass point as functions of the SNR.

In this section, we briefly discuss the effect of 1-bit quantization on the capacity per unit energy of the discrete-time, average-power-limited, *Rayleigh-fading channel*. This channel's unquantized output $\tilde{Y}_k$ is given by

$$\tilde{Y}_k = H_k X_k + Z_k, \quad k \in \mathbb{Z} \tag{63}$$

where $\{H_k, k \in \mathbb{Z}\}$ and $\{Z_k, k \in \mathbb{Z}\}$ are independent sequences of i.i.d., zero-mean, circularly symmetric, complex Gaussian random variables, the former with unit variance and the latter with variance σ^2. We say that the channel is *coherent* if the receiver is cognizant of the realization of $\{H_k, k \in \mathbb{Z}\}$ and that it is *noncoherent* if the receiver is only cognizant of the statistics of $\{H_k, k \in \mathbb{Z}\}$. The unquantized output $\tilde{Y}_k$ is quantized using a 1-bit quantizer that is specified by a Borel subset $\mathcal{D}$ of the complex field $\mathbb{C}$: it produces 1 if $\tilde{Y}_k$ is in $\mathcal{D}$, and it produces 0 if it is not. The capacities $C(\mathcal{P}, \mathcal{D})$ and $C(\mathcal{P})$ are defined as in Section II but with the average-power constraint (3) replaced by

$$\frac{1}{n} \sum_{k=1}^{n} |x_k|^2 \leq \mathcal{P}. \tag{64}$$

Likewise, the capacities per unit energy $\dot{C}(0, \mathcal{D})$ and $\dot{C}(0)$ are defined as in Section II but with the energy constraint (8) replaced by

$$\sum_{k=1}^{n} |x_k|^2 \leq \mathcal{E}. \tag{65}$$

A. Coherent Fading Channels

Using the same arguments as in Section II, it can be shown that, for a fixed quantizer $\mathcal{D}$, we have for the coherent channel [11, Th. 3], [3]

$$\dot{C}(0, \mathcal{D}) = \sup_{\xi \neq 0} \frac{D\big(P_{Y|H,X=\xi} \,\|\, P_{Y|H,X=0} \,\big|\, P_H\big)}{|\xi|^2} \tag{66}$$

where $D(\cdot \| \cdot | \cdot)$ denotes conditional relative entropy

$$\begin{aligned} &D\big(P_{Y|H,X=\xi} \,\|\, P_{Y|H,X=0} \,\big|\, P_H\big) \\ &= \int D\big(P_{Y|H=h,X=\xi} \,\|\, P_{Y|H=h,X=0}\big) \mathrm{d}P_H(h) \end{aligned} \tag{67}$$

P_H denotes the distribution of the fading H, and $P_{Y|H=h,X=x}$ denotes the distribution of Y conditioned on $(H, X) = (h, x)$.[4] It can be further shown that

$$\dot{C}(0) = \sup_{\xi \neq 0, \mathcal{D}} \frac{D\big(P_{Y|H,X=\xi} \,\|\, P_{Y|H,X=0} \,\big|\, P_H\big)}{|\xi|^2}. \tag{68}$$

By the Data Processing Inequality, the capacity per unit energy is upper-bounded by that of the unquantized channel [3], [17]

$$\dot{C}(0) \leq \frac{1}{\sigma^2}. \tag{69}$$

We next show that, by choosing the 1-bit quantizer as a function of H and the SNR, this upper bound can be achieved.

Theorem 4 (Coherent Case): The capacity per unit energy of the coherent Rayleigh-fading channel is given by

$$\dot{C}(0) = \frac{1}{\sigma^2}. \tag{70}$$

It is achieved by a family of radial quantizers parametrized by $\mathcal{P}$ with thresholds that are proportional to $|H|$.

Proof: See Section X-A. ■

The assumption that the fading H is Gaussian is not essential. In fact, Theorem 4 holds for every fading distribution having unit variance.

B. Noncoherent Fading Channels

Using the same arguments as in Section II, it can be shown that in the noncoherent case

$$\dot{C}(0, \mathcal{D}) = \sup_{\xi \neq 0} \frac{D\big(P_{Y|X=\xi} \,\|\, P_{Y|X=0}\big)}{|\xi|^2} \tag{71}$$

and

$$\dot{C}(0) = \sup_{\xi \neq 0, \mathcal{D}} \frac{D\big(P_{Y|X=\xi} \,\|\, P_{Y|X=0}\big)}{|\xi|^2}. \tag{72}$$

---

[4]This can be shown along the lines of the proof of Theorem 3 in [11] but with the mutual information $I(X; Y)$ replaced by the conditional mutual information $I(X; Y | H)$. That the RHS of (66) is an upper bound on $\dot{C}(0, \mathcal{D})$ follows then immediately from [11, eq. (15)]. Showing that this holds with equality requires swapping the order of taking the limit as $\mathcal{P}$ tends to zero and of computing the expectation over the fading.

Since the capacity per unit energy of the unquantized Rayleigh-fading channel equals $1/\sigma^2$ irrespective of whether the channel is coherent or not [3], [17], it follows from the Data Processing Inequality that (69) holds also in the noncoherent case.

The capacity per unit energy (70) of the coherent channel with 1-bit output quantization is achieved using binary on–off keying where the nonzero mass point tends to infinity as the SNR tends to zero. This result might mislead one to think that (70) also holds in the noncoherent case. Indeed, in the absence of a quantizer, binary on–off keying with diverging nonzero mass point achieves the capacity per unit energy $1/\sigma^2$ irrespective of whether the receiver is cognizant of the fading realization or not [3], [17]. It might, therefore, seem plausible that also in the noncoherent case, quantizing the channel output with a 1-bit quantizer would cause no loss in the capacity per unit energy. But this is not the case:

Theorem 5 (Noncoherent Case): For the noncoherent Rayleigh-fading channel with 1-bit output quantization

$$\dot{C}(0) < \frac{1}{\sigma^2}. \tag{73}$$

Proof: See Section X-B. ∎

The case where the real and imaginary parts of the fading channel's output are quantized separately using a 1-bit quantizer for each was studied, e.g., in [18]–[22]. However, in [18]–[21], only symmetric threshold quantizers are considered.

VII. Proof of Theorem 1

We prove Theorem 1 in five steps.

1) We first show that for any given maximal allowed average power $\mathcal{P}$ and any Borel set $\mathcal{D}$, the supremum in (4) defining $C(\mathcal{P}, \mathcal{D})$ is achieved by some input distribution that is concentrated on at most three points (Section VII-A).
2) We next show that for every three-mass-point input distribution, the supremum over all quantizers can be replaced with the supremum over all threshold quantizers and all quantizers whose quantization region consists of a finite interval (see Section VII-B).
3) We continue by showing that the supremum in (7) defining $C(\mathcal{P})$ is achieved (see Section VII-C).
4) We then show that threshold quantizers are optimal by demonstrating that quantization regions consisting of a finite interval are suboptimal (see Section VII-D).
5) We finally show that the capacity-achieving input distribution must be centered and must satisfy the average-power constraint with equality (see Section VII-E).

A. Input Distributions Consisting of Three Mass Points

Generalizing the proof of Theorem 1 in [12] to arbitrary quantizers, we prove that for every fixed quantizer $\mathcal{D}$ and maximal allowed average power $\mathcal{P}$, the capacity $C(\mathcal{P}, \mathcal{D})$ is achieved by an input distribution consisting of three (or fewer) mass points. To this end, we first argue that we can introduce an additional peak-power constraint without reducing capacity, provided that we allow the maximal allowed peak power to tend to infinity.

Thus, we show that $C(\mathcal{P}, \mathcal{D})$, which is defined in (4) without a peak-power constraint, can also be expressed as

$$C(\mathcal{P}, \mathcal{D}) = \lim_{\mathcal{A} \to \infty} \sup_{\substack{\mathsf{E}[X^2] \leq \mathcal{P}, \\ |X| \leq \mathcal{A}}} I(P_X, W_\mathcal{D}) \tag{74}$$

where $W_\mathcal{D}$ denotes the channel law corresponding to the quantization region $\mathcal{D}$, and $I(P_X, W_\mathcal{D})$ denotes the mutual information of a channel with law $W_\mathcal{D}$ when its input is distributed according to P_X. Clearly, the RHS of (74) cannot exceed its LHS, because imposing an additional peak-power constraint cannot increase capacity. It remains to prove that the LHS cannot exceed the RHS.

By Fano's Inequality [10, Th. 2.11.1] and the Data Processing Inequality, we have that, for every blocklength n, every encoder $m \mapsto (x_1(m), \ldots, x_n(m))$ of rate $R = \frac{\log \mathcal{M}}{n}$ that satisfies the average-power constraint, and every quantization region $\mathcal{D}$, the probability of error is lower-bounded by [10, Sec. 8.9]

$$\Pr(\hat{M} \neq M) \geq 1 - \frac{1}{nR} \sum_{k=1}^{n} I(X_k(M); Y_k) - \frac{1}{nR}. \tag{75}$$

Let $\mathcal{A}_n$ be the largest magnitude of the symbols that the encoder can produce

$$\mathcal{A}_n \triangleq \max_{\substack{1 \leq k \leq n, \\ 1 \leq m \leq \mathcal{M}}} |x_k(m)| \tag{76}$$

so

$$|x_k(m)| \leq \mathcal{A}_n, \quad (k = 1, 2, \ldots, n, \; m = 1, 2, \ldots, \mathcal{M}). \tag{77}$$

With this notation, we have for every blocklength n and every quantizer $\mathcal{D}$,

$$\frac{1}{n} \sum_{k=1}^{n} I(X_k(M); Y_k) \leq \sup_{\substack{\mathsf{E}[X^2] \leq \mathcal{P}, \\ |X| \leq \mathcal{A}_n}} I(P_X, W_\mathcal{D})$$
$$\leq \sup_{\mathcal{A} > 0} \sup_{\substack{\mathsf{E}[X^2] \leq \mathcal{P}, \\ |X| \leq \mathcal{A}}} I(P_X, W_\mathcal{D}) \tag{78}$$

where the first inequality follows from (77) and by the concavity of

$$\mathcal{P} \mapsto \sup_{\substack{\mathsf{E}[X^2] \leq \mathcal{P}, \\ |X| \leq \mathcal{A}_n}} I(P_X, W_\mathcal{D}).$$

Thus, the RHS of (75) is bounded away from zero whenever R exceeds the RHS of (78), and the inequality

$$C(\mathcal{P}, \mathcal{D}) \leq \sup_{\mathcal{A} > 0} \sup_{\substack{\mathsf{E}[X^2] \leq \mathcal{P}, \\ |X| \leq \mathcal{A}}} I(P_X, W_\mathcal{D}) \tag{79}$$

is established. Since the inner supremum on the RHS of (79) is monotonically nondecreasing in $\mathcal{A}$, we can replace the outer supremum by a limit and thus establish (74).

Introducing a peak-power constraint in (74) allows us next to establish the existence of a capacity-achieving input distribution

of three mass points using Dubins's Theorem as follows. Recall that by (74)

$$C(\mathcal{P}, \mathcal{D}) = \lim_{\mathcal{A} \to \infty} C_{\mathcal{D}, \mathcal{A}}(\mathcal{P}) \qquad (80)$$

where $C_{\mathcal{D}, \mathcal{A}}(\mathcal{P})$ denotes the capacity of the memoryless channel $\Pr(\tilde{Y} \in \mathcal{D} \mid X = x)$ with the input X taking values in the interval $[-\mathcal{A}, \mathcal{A}]$ and with the binary output Y:

$$C_{\mathcal{D}, \mathcal{A}}(\mathcal{P}) \triangleq \sup_{\substack{\mathsf{E}[X^2] \leq \mathcal{P}, \\ |X| \leq \mathcal{A}}} I(P_X, W_{\mathcal{D}}). \qquad (81)$$

Proceeding along the lines of [23, Sec. II-C] but accounting for the additional average-power constraint, it can be shown that $C_{\mathcal{D}, \mathcal{A}}(\mathcal{P})$ is achieved by an input distribution consisting of three mass points. Indeed, since $\mathcal{P} \mapsto C_{\mathcal{D}, \mathcal{A}}(\mathcal{P})$ is concave, it is continuous, so there exists some $\mathcal{P}' \leq \mathcal{P}$ such that

$$C_{\mathcal{D}, \mathcal{A}}(\mathcal{P}) = \sup_{\substack{\mathsf{E}[X^2] = \mathcal{P}', \\ |X| \leq \mathcal{A}}} I(P_X, W_{\mathcal{D}}). \qquad (82)$$

The input distribution achieving $C_{\mathcal{D}, \mathcal{A}}(\mathcal{P})$ must be concentrated on the interval $[-\mathcal{A}, \mathcal{A}]$ and additionally satisfy

$$\int x^2 \mathrm{d}P_X(x) = \mathcal{P}'. \qquad (83)$$

The arguments in [23, Sec. II-C] thus go through with the set A in [23, Sec. II-C] replaced by the set of input distributions that induce the given output distribution and that additionally lie on the hyperplane (83).

Having established that under an additional peak-power constraint capacity is achieved by a three-mass-point input distribution, we now study what happens to these three mass points as the allowed peak power tends to infinity. We thus study how the three mass points at locations

$$\boldsymbol{\xi} = (\xi_{\mathrm{L}}, \xi_{\mathrm{M}}, \xi_{\mathrm{R}})$$

with corresponding masses

$$\mathbf{p} = (p_{\mathrm{L}}, p_{\mathrm{M}}, p_{\mathrm{R}})$$

behave as $\mathcal{A}$ tends to infinity.

By possibly considering a subsequence of peak powers, we can assume that, as $\mathcal{A}$ tends to infinity, $\boldsymbol{\xi}$ converges to some $\boldsymbol{\xi}^\star = (\xi_{\mathrm{L}}^\star, \xi_{\mathrm{M}}^\star, \xi_{\mathrm{R}}^\star)$ whose components are on the extended real line $\mathbb{R} \cup \{\pm\infty\}$. Likewise, we can assume that $\mathbf{p}$ converges to some probability vector $\mathbf{p}^\star$. Since the input distributions must satisfy the average-power constraint, if any of the components of $\boldsymbol{\xi}^\star$ is $\pm\infty$, then the corresponding component of $\mathbf{p}^\star$ must be zero. By Lemma 1 (Appendix I), for every $\ell = \{\mathrm{L}, \mathrm{M}, \mathrm{R}\}$ the probability $\Pr(\tilde{Y} \in \mathcal{D}|X = \xi_\ell)$ converges to $\Pr(\tilde{Y} \in \mathcal{D}|X = \xi_\ell^\star)$ whenever $\xi_\ell^\star$ is finite, so the continuity of

$$C_{\mathcal{D}, \mathcal{A}}(\mathcal{P}) = H_b\left(\sum_{\ell \in \{\mathrm{L}, \mathrm{M}, \mathrm{R}\}} p_\ell \Pr(\tilde{Y} \in \mathcal{D} \mid X = \xi_\ell) \right)$$
$$- \sum_{\ell \in \{\mathrm{L}, \mathrm{M}, \mathrm{R}\}} p_\ell H_b\left(\Pr(\tilde{Y} \in \mathcal{D} \mid X = \xi_\ell) \right)$$

demonstrates that $\lim_{\mathcal{A} \to \infty} C_{\mathcal{D}, \mathcal{A}}(\mathcal{P})$ [which equals $C(\mathcal{P}, \mathcal{D})$ by (74)] equals the mutual information corresponding to $(\mathbf{p}^\star, \boldsymbol{\xi}^\star)$ provided that in computing the latter the mass points of zero mass are ignored. Since the mass points at $\pm\infty$ are of zero mass (by the average-power constraint), those are ignored, and we conclude that $C(\mathcal{P}, \mathcal{D})$ is achieved by (at most) three *finite* mass point. For sufficiently large $\mathcal{A}$ (exceeding the largest of these mass points), the peak-power constraint is inactive.

B. Quantizers for Three-Mass-Point Input Distributions

Having established that for any quantizer $\mathcal{D}$, the capacity $C(\mathcal{P}, \mathcal{D})$ is achieved by a three-mass-point input distribution, we now fix some arbitrary three-mass-point input distribution[5] P_X concentrated at (ξ_1, ξ_2, ξ_3) and study the quantizer that maximizes the mutual information $I(P_X, W_{\mathcal{D}})$ corresponding to it. (Without loss of generality, we assume that $\xi_1 \neq \xi_2$, $\xi_1 \neq \xi_3$, and $\xi_2 \neq \xi_3$.) We will show that when P_X is a three-mass-point input distribution, we have

$$\sup_{\mathcal{D}} I(P_X, W_{\mathcal{D}}) = \sup_{\Upsilon_1 \leq \Upsilon_2} I(P_X, W_{\mathcal{D}(\Upsilon_1, \Upsilon_2)}) \qquad (84)$$

where the quantizer $\mathcal{D}(\Upsilon_1, \Upsilon_2)$ is defined as

$$\mathcal{D}(\Upsilon_1, \Upsilon_2) \triangleq \{\tilde{y} \in \mathbb{R} : \Upsilon_1 \leq \tilde{y} \leq \Upsilon_2\}, \quad \Upsilon_1 \leq \Upsilon_2 \qquad (85)$$

with

$$\mathcal{D}(-\infty, \Upsilon_2) \triangleq \{\tilde{y} \in \mathbb{R} : \tilde{y} \leq \Upsilon_2\}, \quad \Upsilon_2 \in \mathbb{R} \qquad (86\mathrm{a})$$
$$\mathcal{D}(\Upsilon_1, \infty) \triangleq \{\tilde{y} \in \mathbb{R} : \tilde{y} \geq \Upsilon_1\}, \quad \Upsilon_1 \in \mathbb{R} \qquad (86\mathrm{b})$$
$$\mathcal{D}(-\infty, \infty) \triangleq \mathbb{R} \qquad (86\mathrm{c})$$
$$\mathcal{D}(-\infty, -\infty) = \mathcal{D}(\infty, \infty) \triangleq \varnothing. \qquad (86\mathrm{d})$$

(Here, $\varnothing$ denotes the empty set.) Needless to say, the case $\Upsilon_1 = \Upsilon_2$ and the forms (86c) and (86d) yield zero mutual information and are thus uninteresting.

Define

$$\mathcal{W} \triangleq \Big\{ (\omega_1, \omega_2, \omega_3) \in [0, 1]^3 :$$
$$\omega_\ell = \Pr(\tilde{Y} \in \mathcal{D} \mid X = \xi_\ell), \mathcal{D} \subset \mathbb{R} \Big\} \qquad (87)$$

as the set of possible channel laws that different quantizers can induce for the inputs (ξ_1, ξ_2, ξ_3), and let $\overline{\mathcal{W}}$ denote the closure of the convex hull of $\mathcal{W}$. With this notation

$$\sup_{\mathcal{D}} I(P_X, W_{\mathcal{D}}) = \sup_{W \in \mathcal{W}} I(P_X, W)$$
$$\leq \sup_{W \in \overline{\mathcal{W}}} I(P_X, W) \qquad (88)$$

where the second step follows because $\mathcal{W} \subseteq \overline{\mathcal{W}}$. Recall that an extreme point of $\overline{\mathcal{W}}$ is a channel in $\overline{\mathcal{W}}$ that cannot be written as a convex combination of two different channels in $\overline{\mathcal{W}}$. By the Krein–Milman Theorem [24, Cor. 18.5.1], every channel law $W \in \overline{\mathcal{W}}$ can be written as a convex combination of extreme points of $\overline{\mathcal{W}}$. Since mutual information is convex in the channel law (when the input distribution is held fixed) [10, Th. 2.7.4], it

---

[5]Every two-mass-point distribution can be viewed as a three-mass-point distribution with one of the masses being zero.

follows that on the RHS of (88), we can replace the supremum over the set $\overline{\mathcal{W}}$ with the supremum over its extreme points.

We next show that the extreme points of $\overline{\mathcal{W}}$ correspond to quantizers of the form (85). Once we show this, it will follow that (88) holds with equality, because these extreme points of $\overline{\mathcal{W}}$ are in fact in $\mathcal{W}$. This will prove (84).[6]

To prove that the extreme points of $\overline{\mathcal{W}}$ are indeed the channel laws corresponding to quantizers of the form (85), we consider the *support function* of $\overline{\mathcal{W}}$ [24, Sec. 13]

$$f(\boldsymbol{\lambda}) \triangleq \sup_{(\omega_1, \omega_2, \omega_3) \in \overline{\mathcal{W}}} \{\lambda_1 \omega_1 + \lambda_2 \omega_2 + \lambda_3 \omega_3\} \qquad (89)$$

for $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3$. Since $\overline{\mathcal{W}}$ is the closure of all convex combinations of the elements of $\mathcal{W}$ [24, Th. 2.3], the support function of $\overline{\mathcal{W}}$ is the same as that of $\mathcal{W}$ and

$$f(\boldsymbol{\lambda}) = \sup_{\mathcal{D}} \{\lambda_1 \omega_1(\mathcal{D}) + \lambda_2 \omega_2(\mathcal{D}) + \lambda_3 \omega_3(\mathcal{D})\} \qquad (90)$$

where

$$\omega_\ell(\mathcal{D}) \triangleq \Pr(\tilde{Y} \in \mathcal{D} \mid X = \xi_\ell), \quad \ell = 1, 2, 3. \qquad (91)$$

We rewrite (90) as

$$f(\boldsymbol{\lambda}) = \sup_{\mathcal{D}} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathcal{D}} g_{\boldsymbol{\lambda}}(\tilde{y}) d\tilde{y} \qquad (92)$$

where

$$g_{\boldsymbol{\lambda}}(\tilde{y}) \triangleq \lambda_1 e^{-\frac{(\tilde{y}-\xi_1)^2}{2\sigma^2}} + \lambda_2 e^{-\frac{(\tilde{y}-\xi_2)^2}{2\sigma^2}} + \lambda_3 e^{-\frac{(\tilde{y}-\xi_3)^2}{2\sigma^2}}, \quad \tilde{y} \in \mathbb{R}. \qquad (93)$$

The integral on the RHS of (92) is maximized when $\mathcal{D}$ is the set

$$\mathcal{D}^\star(\boldsymbol{\lambda}) = \{\tilde{y} \in \mathbb{R} : g_{\boldsymbol{\lambda}}(\tilde{y}) \geq 0\}. \qquad (94)$$

The structure of $\mathcal{D}^\star(\boldsymbol{\lambda})$ depends on the zeros of $g_{\boldsymbol{\lambda}}(\cdot)$, which we study next.

Our study of the zeros of $g_{\boldsymbol{\lambda}}(\cdot)$ depends on the signs of $\lambda_1, \lambda_2, \lambda_3$ and on how many of them are zero. The case where $\lambda_1, \lambda_2,$ and λ_3 are all zero is trivial, because in this case, $f(\boldsymbol{\lambda})$ is zero irrespective of $\mathcal{D}$. We will see that in all other cases, the set $\mathcal{D}$ that achieves $f(\boldsymbol{\lambda})$ is unique up to Lebesgue measure zero. If exactly two λ's, say λ_1 and λ_2, are zero, then the set $\mathcal{D}$ that achieves $f(\boldsymbol{\lambda})$ is either $\mathbb{R}$ or $\varnothing$, depending on whether λ_3 is positive or negative. We next consider the case where exactly one of the λ's, say λ_3, is zero. In this case

$$g_{\boldsymbol{\lambda}}(\tilde{y}) = \lambda_1 e^{-\frac{(\tilde{y}-\xi_1)^2}{2\sigma^2}} + \lambda_2 e^{-\frac{(\tilde{y}-\xi_2)^2}{2\sigma^2}}, \quad \tilde{y} \in \mathbb{R} \qquad (95)$$

which is either positive (if $\lambda_1 > 0$ and $\lambda_2 > 0$), negative (if $\lambda_1 < 0$ and $\lambda_2 < 0$), or has a zero at

$$\tilde{y} = \frac{\xi_1 + \xi_2}{2} + \frac{\sigma^2}{\xi_2 - \xi_1} \log \left| \frac{\lambda_1}{\lambda_2} \right| \qquad (96)$$

---

[6]Note that $\overline{\mathcal{W}}$ is the set of possible channel laws that different quantizers can induce for the inputs (ξ_1, ξ_2, ξ_3), provided that we allow for randomized quantization rules. It thus follows that (84) continues to hold if on the LHS, instead of maximizing over all deterministic quantizers $\mathcal{D}$, we maximize over all probability distributions $P_{Y|\tilde{Y}}$ with Y binary.

(if λ_1 and λ_2 have opposite signs). Consequently, if exactly one of the λ's is zero, then the set $\mathcal{D}$ that achieves $f(\boldsymbol{\lambda})$ is either the entire real line, the empty set, or a ray, i.e., of the form $(-\infty, \Upsilon)$ or (Υ, ∞), where Υ is the RHS of (96).

We finally turn to the case where all the λ's are nonzero. If they are all of equal sign, then $f(\boldsymbol{\lambda})$ has no zeros and the set $\mathcal{D}$ that maximizes $f(\boldsymbol{\lambda})$ is either the entire real line $\mathbb{R}$ or the empty set, depending on whether the λ's are all positive or all negative. It remains to study the case where the λ's are nonzero but not of equal sign. Changing the sign of all the λ's is tantamount to multiplying $g_{\boldsymbol{\lambda}}(\cdot)$ by -1 and, therefore, does not change the locations of the zeros, so we can assume without loss of generality that one of the λ's, say λ_1, is positive and that the remaining two λ_2, λ_3 are negative. In this case

$$g_{\boldsymbol{\lambda}}(\tilde{y}) = \lambda_1 e^{-\frac{(\tilde{y}-\xi_1)^2}{2\sigma^2}} h_{\boldsymbol{\lambda}}(\tilde{y}), \quad \tilde{y} \in \mathbb{R} \qquad (97)$$

where

$$h_{\boldsymbol{\lambda}}(\tilde{y}) \triangleq 1 - \left| \frac{\lambda_2}{\lambda_1} \right| e^{\frac{\xi_1^2 - \xi_2^2}{2\sigma^2}} e^{\tilde{y} \frac{\xi_2 - \xi_1}{\sigma^2}}$$
$$- \left| \frac{\lambda_3}{\lambda_1} \right| e^{\frac{\xi_1^2 - \xi_3^2}{2\sigma^2}} e^{\tilde{y} \frac{\xi_3 - \xi_1}{\sigma^2}}, \quad \tilde{y} \in \mathbb{R}. \qquad (98)$$

Note that the zeros of $g_{\boldsymbol{\lambda}}(\cdot)$ are the same as the zeros of $h_{\boldsymbol{\lambda}}(\cdot)$. Further note that $h_{\boldsymbol{\lambda}}(\cdot)$ is a nonzero analytic function whose second derivative

$$\frac{\partial^2}{\partial \tilde{y}^2} h_{\boldsymbol{\lambda}}(\tilde{y}) = -\frac{(\xi_2 - \xi_1)^2}{\sigma^4} \left| \frac{\lambda_2}{\lambda_1} \right| e^{\frac{\xi_1^2 - \xi_2^2}{2\sigma^2}} e^{\tilde{y} \frac{\xi_2 - \xi_1}{\sigma^2}}$$
$$- \frac{(\xi_3 - \xi_1)^2}{\sigma^4} \left| \frac{\lambda_3}{\lambda_1} \right| e^{\frac{\xi_1^2 - \xi_3^2}{2\sigma^2}} e^{\tilde{y} \frac{\xi_3 - \xi_1}{\sigma^2}}, \quad \tilde{y} \in \mathbb{R} \qquad (99)$$

is strictly negative. Consequently, $h_{\boldsymbol{\lambda}}(\cdot)$—and hence also $g_{\boldsymbol{\lambda}}(\cdot)$—can have at most two zeros. (If it had three or more, then by Rolle's Theorem, its derivative would have at least two zeros, and its second derivative would therefore have a zero in contradiction to (99).) If $h_{\boldsymbol{\lambda}}(\cdot)$ has at most one zero, then the set $\mathcal{D}$ achieving $f(\boldsymbol{\lambda})$ is either the entire real line, the empty set, or a ray. If it has two zeros, then $\mathcal{D}$ comprises two disjoint rays or else a finite interval—either way, $\mathcal{D}$ or its complement is a finite interval.

We next show that for every $\boldsymbol{\lambda} \neq \boldsymbol{0}$, the quantization region achieving $f(\boldsymbol{\lambda})$ is unique up to sets of Lebesgue measure zero. Let $\mathcal{D}^\star(\boldsymbol{\lambda})$ be the quantization region that achieves $f(\boldsymbol{\lambda})$, and let $\mathcal{D}_1$ be any other quantization region. Then

$$\int_{\mathcal{D}^\star(\boldsymbol{\lambda})} g_{\boldsymbol{\lambda}}(\tilde{y}) d\tilde{y} - \int_{\mathcal{D}_1} g_{\boldsymbol{\lambda}}(\tilde{y}) d\tilde{y}$$
$$= \int_{\mathcal{D}^\star(\boldsymbol{\lambda}) \cap \mathcal{D}_1^c} g_{\boldsymbol{\lambda}}(\tilde{y}) d\tilde{y} - \int_{\mathcal{D}_1 \cap \mathcal{D}^\star(\boldsymbol{\lambda})^c} g_{\boldsymbol{\lambda}}(\tilde{y}) d\tilde{y}$$
$$\geq \int_{\mathcal{D}^\star(\boldsymbol{\lambda}) \cap \mathcal{D}_1^c} g_{\boldsymbol{\lambda}}(\tilde{y}) d\tilde{y}$$
$$\geq 0 \qquad (100)$$

where the second step follows because for every $\tilde{y} \in \mathcal{D}^\star(\boldsymbol{\lambda})^c$, we have $g_{\boldsymbol{\lambda}}(\tilde{y}) < 0$; and the last step follows because for every $\tilde{y} \in \mathcal{D}^\star(\boldsymbol{\lambda})$, we have $g_{\boldsymbol{\lambda}}(\tilde{y}) \geq 0$. (Here, $\mathcal{A}^c$ denotes the complement of the set $\mathcal{A}$.) Furthermore, since the zeros of $g_{\boldsymbol{\lambda}}(\cdot)$ are isolated, it is nonzero almost everywhere, so the inequalities hold

with equality if, and only if, $\mathcal{D}^\star(\boldsymbol{\lambda}) \cap \mathcal{D}_1^c$ and $\mathcal{D}_1 \cap \mathcal{D}^\star(\boldsymbol{\lambda})^c$ have both Lebesgue measure zero.

Because quantizers that differ on a set of Lebesgue measure zero induce identical channel laws, the uniqueness (up to sets of Lebesgue measure zero) of the set $\mathcal{D}$ achieving $f(\boldsymbol{\lambda})$ (for $\boldsymbol{\lambda} \neq \mathbf{0}$) implies that for every $\boldsymbol{\lambda} \neq \mathbf{0}$, the tuple $(\omega_1^\star, \omega_2^\star, \omega_3^\star)$ that achieves $f(\boldsymbol{\lambda})$ is unique.

We next note that, by [24, Th. 13.1], every $(\omega_1, \omega_2, \omega_3) \in \overline{\mathcal{W}}$ satisfying

$$\lambda_1 \omega_1 + \lambda_2 \omega_2 + \lambda_3 \omega_3 < f(\boldsymbol{\lambda}), \quad \text{for every } \boldsymbol{\lambda} \neq 0$$

must be an interior point of $\overline{\mathcal{W}}$. Since an interior point cannot be an extreme point, it follows that every extreme point of a compact convex set achieves the supremum defining $f(\boldsymbol{\lambda})$ at some $\boldsymbol{\lambda} \neq 0$. Furthermore, since for a given $\boldsymbol{\lambda} \neq 0$ the support function $f(\boldsymbol{\lambda})$ is achieved uniquely by a channel law that is induced by a quantizer of the form (85) or their complement, it follows that the extreme points of $\overline{\mathcal{W}}$ are all achieved by quantizers of this form or their complement. Recalling that mutual information is maximized over $\overline{\mathcal{W}}$ (for a given input distribution) at an extreme point, and noting that the mutual information corresponding to the quantizer $\mathcal{D}$ is the same as that corresponding to its complement, we conclude that—for any fixed three-mass-point input distribution—the supremum over all quantizers can be replaced with the supremum over all quantizers of the form (85), thus proving (84).

C. Supremum Defining $C(\mathcal{P})$ is Achieved

Having established that to each quantizer, the optimal input distribution is of three mass points, and having established that to each three-mass-point input distribution, the optimal quantizer is of the form (85), we conclude that we can express $C(\mathcal{P})$ of (7) as

$$C(\mathcal{P}) = \sup_{\substack{(\mathbf{p}, \boldsymbol{\xi}) : \mathbb{E}[X^2] \leq \mathcal{P}, \\ \Upsilon_1 \leq \Upsilon_2}} I\big(\mathbf{p}, \mathbf{W}(\Upsilon_1, \Upsilon_2 | \boldsymbol{\xi})\big) \qquad (101)$$

where $(\mathbf{p}, \boldsymbol{\xi})$ denotes the three-mass-point distribution of masses

$$\mathbf{p} = (p_1, p_2, p_3) \in [0, 1]^3$$

and locations

$$\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3) \in \mathbb{R}^3$$

and where $\mathbf{W}(\Upsilon_1, \Upsilon_2 | \boldsymbol{\xi})$ denotes the channel law corresponding to the quantizer $\mathcal{D}(\Upsilon_1, \Upsilon_2)$ and to the mass points ξ_ℓ, $\ell = 1, 2, 3$:

$$W\big(\Upsilon_1, \Upsilon_2 \mid \xi_\ell\big) \triangleq \Pr\big(\tilde{Y} \in \mathcal{D}(\Upsilon_1, \Upsilon_2) \mid X = \xi_\ell\big). \quad (102)$$

We next show that this supremum is achieved.

By the definition of the supremum, there exists a sequence $\big\{(\mathbf{p}_i, \boldsymbol{\xi}_i, \Upsilon_{1,i}, \Upsilon_{2,i}), \, i \in \mathbb{N}\big\}$ (where $\mathbb{N}$ denotes the set of positive integers) such that

$$\lim_{i \to \infty} I\big(\mathbf{p}_i, \mathbf{W}(\Upsilon_{1,i}, \Upsilon_{2,i} | \boldsymbol{\xi}_i)\big) = C(\mathcal{P}). \qquad (103)$$

By taking a subsequence (if needed), we may assume without loss of generality that $\mathbf{p}_i$ converges to some $\mathbf{p}^\star$, that $\boldsymbol{\xi}_i$ converges to some $\boldsymbol{\xi}^\star$ (whose components may be $\pm\infty$) and that $\Upsilon_{1,i}$ and $\Upsilon_{2,i}$ converge to $\Upsilon_1^\star$ and $\Upsilon_2^\star$, both of which may be $\pm\infty$. From the continuity of the cumulative distribution function of the normal distribution, it follows that, whenever $\xi_\ell^\star$ is finite,

$$\lim_{i \to \infty} \Pr(\Upsilon_{1,i} \leq \xi_{\ell,i} + Z \leq \Upsilon_{2,i})$$
$$= \Pr(\Upsilon_1^\star \leq \xi_\ell^\star + Z \leq \Upsilon_2^\star) \qquad (104)$$

where we recall that Z is a centered Gaussian random variable of positive variance σ^2.

Since the mass $p_\ell^\star$ corresponding to nonfinite locations $\xi_\ell^\star$ is zero (by the average-power constraint), and since $p_{\ell,i}$ converges to $p_\ell^\star$, (104) and the continuity of the binary entropy function allow us to infer that

$$\lim_{i \to \infty} I\big(\mathbf{p}_i, \mathbf{W}(\Upsilon_{1,i}, \Upsilon_{2,i} | \boldsymbol{\xi}_i)\big)$$
$$= \lim_{i \to \infty} \left\{ H_b\left(\sum_{\ell=1}^{3} p_{\ell,i} \, W(\Upsilon_{1,i}, \Upsilon_{2,i} | \xi_{\ell,i}) \right) \right.$$
$$\left. - \sum_{\ell=1}^{3} p_{\ell,i} \, H_b\big(W(\Upsilon_{1,i}, \Upsilon_{2,i} | \xi_{\ell,i})\big) \right\}$$
$$= I\big(\mathbf{p}^\star, \mathbf{W}(\Upsilon_1^\star, \Upsilon_2^\star | \boldsymbol{\xi}^\star)\big) \qquad (105)$$

provided that in computing the mutual information on the LHS of (106) the mass points of zero mass are ignored. This combines with (103) to imply that

$$I\big(\mathbf{p}^\star, \mathbf{W}(\Upsilon_1^\star, \Upsilon_2^\star | \boldsymbol{\xi}^\star)\big) = C(\mathcal{P}). \qquad (106)$$

Noting that the mass points at $\pm\infty$ are of zero mass and therefore ignored, we conclude that $C(\mathcal{P})$ is achieved by an input distribution of (at most) three *finite* mass points and by a quantizer of the form (85).

D. Threshold Quantizer is Optimal

Having established that $C(\mathcal{P})$ is achieved by a three-mass-point input distribution and a quantizer of the form (85), we now prove that $C(\mathcal{P})$ is in fact achieved by a three-mass-point input distribution and a threshold quantizer, i.e., a quantizer of the form (86b). Clearly, Υ_1 and Υ_2 cannot be both nonfinite, as this would result in zero mutual information, whereas $C(\mathcal{P})$ is strictly positive whenever $\mathcal{P}$ is positive[7]

$$C(\mathcal{P}) > 0, \quad \mathcal{P} > 0. \qquad (107)$$

For the same reason, we can assume, without loss of optimality, that $\Upsilon_1 \neq \Upsilon_2$. Since (86a) is the complement of a set of the form (86b)—which gives rise to the same mutual information—it remains to rule out the case where Υ_1 and Υ_2 are both finite.

---

[7]This can be verified by noting that a symmetric threshold quantizer and an equiprobable $\pm\sqrt{\mathcal{P}}$ input distribution yield positive mutual information for every positive $\mathcal{P}$; cf. (23).

We shall prove this by contradiction. We shall assume that the quantization region $\mathcal{D}(\Upsilon_1, \Upsilon_2)$ for some finite $\Upsilon_1 < \Upsilon_2$ is optimal and derive a contradiction to optimality. Assume then that Υ_1 and Υ_2 are both finite with $\Upsilon_1 < \Upsilon_2$. Define

$$\theta \triangleq \frac{\Upsilon_1 + \Upsilon_2}{2}. \tag{108}$$

Let $\boldsymbol{\xi}$ be the mass points of the capacity-achieving input distribution, and let $\mathbf{p}$ be the corresponding probabilities. Note that there is no loss in optimality in assuming that θ is nonnegative

$$\theta \geq 0 \tag{109}$$

because if θ is negative, then we can consider the input $(\mathbf{p}, -\boldsymbol{\xi})$ [whose second moment is identical to that of $(\mathbf{p}, \boldsymbol{\xi})$] and the quantizer $\mathcal{D}(-\Upsilon_2, -\Upsilon_1)$ [whose midpoint is of opposite sign to that of $\mathcal{D}(\Upsilon_1, \Upsilon_2)$] which give rise to the same mutual information as the input $(\mathbf{p}, \boldsymbol{\xi})$ and the quantizer $\mathcal{D}(\Upsilon_1, \Upsilon_2)$.

Assume that the mass points are ordered, i.e., $\xi_1 < \xi_2 < \xi_3$. Since the locations of mass points of zero mass have no effect on the mutual information, there is no loss in optimality in assuming that the probability of the largest mass point satisfies $p_3 > 0$. Furthermore, $p_3 < 1$ since $p_3 = 1$ would imply that $C(\mathcal{P}) = 0$, $\mathcal{P} > 0$ in contradiction to (107).

We continue by noting that the symmetry of the normal distribution implies that

$$W(\Upsilon_1, \Upsilon_2 \mid \theta - \delta) = W(\Upsilon_1, \Upsilon_2 \mid \theta + \delta), \quad \delta \geq 0. \tag{110}$$

Indeed, defining $\Delta \triangleq (\Upsilon_2 - \Upsilon_1)/2$ (so $\Upsilon_1 = \theta - \Delta$ and $\Upsilon_2 = \theta + \Delta$), we have

$$W(\Upsilon_1, \Upsilon_2 \mid \theta - \delta) = \int_{\theta-\Delta}^{\theta+\Delta} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\tilde{y}-\theta+\delta)^2}{2\sigma^2}} d\tilde{y}$$
$$= \int_{\theta-\Delta}^{\theta+\Delta} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(-\tau+\theta+\delta)^2}{2\sigma^2}} d\tau$$
$$= W(\Upsilon_1, \Upsilon_2 \mid \theta + \delta) \tag{111}$$

where we made the substitution $\tau = -\tilde{y} + 2\theta$. Furthermore, since $\theta \geq 0$,

$$(\theta - \delta)^2 \leq (\theta + \delta)^2, \quad \delta \geq 0. \tag{112}$$

As we next argue, (110) and (112) imply that there is no loss in optimality in assuming that

$$\xi_1 < \xi_2 < \xi_3 \leq \theta. \tag{113}$$

Indeed, suppose $\xi_3 > \theta$. Then, ξ_3 can be written as $\theta + \delta$, for some $\delta > 0$. However, $\tilde{\xi}_3 = \theta - \delta$ gives rise to the same channel law (110) but has a smaller cost (112). Thus, for every $\xi_3 > \theta$, we can find a $\tilde{\xi}_3 < \theta$ satisfying the power constraint that achieves the same rate.

We next show that (113) leads to a contradiction by considering a perturbation of the quantizer. For every $\Gamma > \Upsilon_2$, define the perturbed quantization region

$$\tilde{\mathcal{D}} \triangleq (\Upsilon_1, \Upsilon_2) \cup [\Gamma, +\infty) \tag{114}$$

and denote the channel law corresponding to $\tilde{\mathcal{D}}$ and $\boldsymbol{\xi}$ by $\mathbf{W}(\tilde{\mathcal{D}}|\boldsymbol{\xi})$:

$$W(\tilde{\mathcal{D}} \mid \xi_\ell) \triangleq \Pr(\tilde{Y} \in \tilde{\mathcal{D}} \mid X = \xi_\ell)$$
$$= W(\Upsilon_1, \Upsilon_2 | \xi_\ell) + Q\left(\frac{\Gamma - \xi_\ell}{\sigma}\right) \tag{115}$$

for $\ell = 1, 2, 3$. We will contradict the optimality of the input $(\mathbf{p}, \boldsymbol{\xi})$ and the quantizer $\mathcal{D}(\Upsilon_1, \Upsilon_2)$ by showing that for $(\mathbf{p}, \boldsymbol{\xi})$ satisfying (113), we can find a sufficiently large Γ exceeding Υ_2 such that

$$I(\mathbf{p}, \mathbf{W}(\tilde{\mathcal{D}}|\boldsymbol{\xi})) > I(\mathbf{p}, \mathbf{W}(\Upsilon_1, \Upsilon_2|\boldsymbol{\xi})). \tag{116}$$

To show this, we use (115) to express the mutual information on the LHS of (116) as

$$I(\mathbf{p}, \mathbf{W}(\tilde{\mathcal{D}}|\boldsymbol{\xi})) = H_b(P(\Upsilon_1, \Upsilon_2) + P(\Gamma))$$
$$- \sum_{\ell=1}^{3} p_\ell H_b\left(W(\Upsilon_1, \Upsilon_2 \mid \xi_\ell) + Q\left(\frac{\Gamma - \xi_\ell}{\sigma}\right)\right) \tag{117}$$

where

$$P(\Upsilon_1, \Upsilon_2) \triangleq \sum_{\ell=1}^{3} p_\ell W(\Upsilon_1, \Upsilon_2|\xi_\ell) \tag{118a}$$

$$P(\Gamma) \triangleq \sum_{\ell=1}^{3} p_\ell Q\left(\frac{\Gamma - \xi_\ell}{\sigma}\right). \tag{118b}$$

A Taylor series expansion of $H_b(p + \epsilon)$ around p yields

$$H_b(p + \epsilon) = H_b(p) + \epsilon \log \frac{1-p}{p} + \mathsf{R}(p, \epsilon) \tag{119}$$

for $0 < p < 1 - \epsilon$ and some remainder $\mathsf{R}(p, \epsilon)$ satisfying

$$|\mathsf{R}(p, \epsilon)| \leq \frac{\epsilon^2}{2} \frac{1}{p(1 - p - \epsilon)}. \tag{120}$$

With this, we obtain

$$I(\mathbf{p}, \mathbf{W}(\tilde{\mathcal{D}}|\boldsymbol{\xi}))$$
$$= H_b(P(\Upsilon_1, \Upsilon_2)) + P(\Gamma) \log \frac{1 - P(\Upsilon_1, \Upsilon_2)}{P(\Upsilon_1, \Upsilon_2)}$$
$$- \sum_{\ell=1}^{3} p_\ell H_b\left(W(\Upsilon_1, \Upsilon_2 \mid \xi_\ell)\right)$$
$$- \sum_{\ell=1}^{3} p_\ell Q\left(\frac{\Gamma - \xi_\ell}{\sigma}\right) \log \frac{1 - W(\Upsilon_1, \Upsilon_2 \mid \xi_\ell)}{W(\Upsilon_1, \Upsilon_2 \mid \xi_\ell)}$$
$$+ \mathsf{K}(\mathbf{p}, \boldsymbol{\xi}, \Gamma)$$
$$= I(\mathbf{p}, \mathbf{W}(\Upsilon_1, \Upsilon_2|\boldsymbol{\xi})) + P(\Gamma) \log \frac{1 - P(\Upsilon_1, \Upsilon_2)}{P(\Upsilon_1, \Upsilon_2)}$$
$$- \sum_{\ell=1}^{3} p_\ell Q\left(\frac{\Gamma - \xi_\ell}{\sigma}\right) \log \frac{1 - W(\Upsilon_1, \Upsilon_2 \mid \xi_\ell)}{W(\Upsilon_1, \Upsilon_2 \mid \xi_\ell)}$$
$$+ \mathsf{K}(\mathbf{p}, \boldsymbol{\xi}, \Gamma) \tag{121}$$

where

$$\mathsf{K}(\mathbf{p}, \boldsymbol{\xi}, \Gamma) \triangleq \mathsf{R}\big(P(\Upsilon_1, \Upsilon_2), P(\Gamma)\big)$$
$$- \sum_{\ell=1}^{3} p_\ell \mathsf{R}\left(W(\Upsilon_1, \Upsilon_2 \mid \xi_\ell), Q\left(\frac{\Gamma - \xi_\ell}{\sigma}\right)\right). \quad (122)$$

Since the LHS of (115) is strictly smaller than 1 so is its RHS and it follows upon averaging over $\mathbf{p}$ that for every $\mathcal{P} > 0$ and every $\Upsilon_1 \leq \Upsilon_2 < \Gamma$

$$P(\Upsilon_1, \Upsilon_2) + P(\Gamma) < 1. \quad (123)$$

Furthermore, $P(\Upsilon_1, \Upsilon_2)$ is strictly positive since $W(\Upsilon_1, \Upsilon_2 | \xi_\ell) > 0$ for $\ell = 1, 2, 3$. Using (120), it thus follows that

$$\lim_{\Gamma \to \infty} \frac{\left|\mathsf{R}\big(P(\Upsilon_1, \Upsilon_2), P(\Gamma)\big)\right|}{Q\left(\frac{\Gamma - \xi_3}{\sigma}\right)}$$
$$\leq \lim_{\Gamma \to \infty} \frac{[P(\Gamma)]^2}{Q\left(\frac{\Gamma - \xi_3}{\sigma}\right)} \frac{1}{2\, P(\Upsilon_1, \Upsilon_2)\big(1 - P(\Upsilon_1, \Upsilon_2) - P(\Gamma)\big)}$$
$$\leq \lim_{\Gamma \to \infty} \frac{Q\left(\frac{\Gamma - \xi_3}{\sigma}\right)}{2\, P(\Upsilon_1, \Upsilon_2)\big(1 - P(\Upsilon_1, \Upsilon_2) - P(\Gamma)\big)}$$
$$= 0 \quad (124)$$

where the second step follows because $\xi_1 < \xi_2 < \xi_3$, which implies that

$$P(\Gamma) \leq Q\left(\frac{\Gamma - \xi_3}{\sigma}\right)$$

and where the last step follows because $P(\Gamma)$ and $Q\big((\Gamma - \xi_3)/\sigma\big)$ both tend to zero as Γ tends to infinity. Along the same lines, it can be shown that for $\ell = 1, 2, 3$

$$\lim_{\Gamma \to \infty} \frac{\left|\mathsf{R}\left(W(\Upsilon_1, \Upsilon_2|\xi_\ell), Q\left(\frac{\Gamma - \xi_\ell}{\sigma}\right)\right)\right|}{Q\left(\frac{\Gamma - \xi_3}{\sigma}\right)} = 0. \quad (125)$$

It thus follows from (122), (124), (125), and the Triangle Inequality that

$$\lim_{\Gamma \to \infty} \frac{\left|\mathsf{K}(\mathbf{q}, \boldsymbol{\xi}, \Gamma)\right|}{Q\left(\frac{\Gamma - \xi_3}{\sigma}\right)}$$
$$\leq \lim_{\Gamma \to \infty} \frac{\left|\mathsf{R}\big(P(\Upsilon_1, \Upsilon_2), P(\Gamma)\big)\right|}{Q\left(\frac{\Gamma - \xi_3}{\sigma}\right)}$$
$$+ \lim_{\Gamma \to \infty} \sum_{\ell=1}^{3} p_\ell \frac{\left|\mathsf{R}\left(W(\Upsilon_1, \Upsilon_2|\xi_\ell), Q\left(\frac{\Gamma - \xi_\ell}{\sigma}\right)\right)\right|}{Q\left(\frac{\Gamma - \xi_3}{\sigma}\right)}$$
$$= 0. \quad (126)$$

We further have by [14, Prop. 19.4.2] that for $\ell = 1, 2$

$$\lim_{\Gamma \to \infty} \frac{Q\left(\frac{\Gamma - \xi_\ell}{\sigma}\right)}{Q\left(\frac{\Gamma - \xi_3}{\sigma}\right)} \leq \lim_{\Gamma \to \infty} \frac{\Gamma - \xi_3}{\Gamma - \xi_\ell} \frac{e^{\frac{\xi_3^2 - \xi_\ell^2}{2\sigma^2}}}{1 - \frac{\sigma^2}{(\Gamma - \xi_3)^2}} e^{-\Gamma \frac{\xi_3 - \xi_\ell}{\sigma^2}}$$
$$= 0. \quad (127)$$

We thus obtain from (118b), (121), (126), and (127) that

$$\lim_{\Gamma \to \infty} \frac{I\big(\mathbf{p}, \mathbf{W}(\tilde{\mathcal{D}}|\boldsymbol{\xi})\big) - I\big(\mathbf{p}, \mathbf{W}(\Upsilon_1, \Upsilon_2|\boldsymbol{\xi})\big)}{Q\left(\frac{\Gamma - \xi_3}{\sigma}\right)}$$
$$= p_3 \log \frac{1 - P(\Upsilon_1, \Upsilon_2)}{P(\Upsilon_1, \Upsilon_2)} - p_3 \log \frac{1 - W(\Upsilon_1, \Upsilon_2 \mid \xi_3)}{W(\Upsilon_1, \Upsilon_2 \mid \xi_3)}$$
$$= p_3 \left(\log \frac{1 - P(\Upsilon_1, \Upsilon_2)}{1 - W(\Upsilon_1, \Upsilon_2 \mid \xi_3)} + \log \frac{W(\Upsilon_1, \Upsilon_2 \mid \xi_3)}{P(\Upsilon_1, \Upsilon_2)} \right)$$
$$> 0 \quad (128)$$

where the inequality follows from the assumption $p_3 > 0$ and by noting that

$$\xi \mapsto W(\Upsilon_1, \Upsilon_2 \mid \xi)$$

is strictly increasing on $(-\infty, \theta)$ (see Appendix II), which together with $p_3 < 1$ implies that

$$W(\Upsilon_1, \Upsilon_2 \mid \xi_3) > P(\Upsilon_1, \Upsilon_2). \quad (129)$$

Consequently, for a sufficiently large Γ, $I\big(\mathbf{p}, \mathbf{W}(\tilde{\mathcal{D}}|\boldsymbol{\xi})\big)$ is strictly larger than $I\big(\mathbf{p}, \mathbf{W}(\Upsilon_1, \Upsilon_2|\boldsymbol{\xi})\big)$, contradicting the assumption that $\mathcal{D}(\Upsilon_1, \Upsilon_2)$ with finite $\Upsilon_1 \leq \Upsilon_2$ achieves $C(\mathcal{P})$.

E. Centered, Variance-$\mathcal{P}$ Input Distribution

We have shown that the supremum in (7) is achieved by some input distribution that is concentrated on at most three points and by some threshold quantizer:

$$C(\mathcal{P}) = I\big(\mathbf{p}^\star, \mathbf{W}(\Upsilon^\star|\boldsymbol{\xi}^\star)\big) \quad (130)$$

where $\boldsymbol{\xi}^\star \in \mathbb{R}^3$ is the location of the mass points, $\mathbf{p}^\star$ is their corresponding probabilities, $\Upsilon^\star$ is the threshold of the quantizer, and $\mathbf{W}(\Upsilon^\star|\boldsymbol{\xi}^\star)$ is the resulting channel law. We next show that the input distribution $(\mathbf{p}^\star, \boldsymbol{\xi}^\star)$ must be centered and must satisfy the average-power constraint with equality:

$$\sum_{\ell=1}^{3} p_\ell^\star \xi_\ell^\star = 0 \quad (131a)$$
$$\sum_{\ell=1}^{3} p_\ell^\star \big(\xi_\ell^\star\big)^2 = \mathcal{P}. \quad (131b)$$

To show this, we note that, for a fixed threshold quantizer $\Upsilon^\star$, the capacity as a function of the maximal allowed average power is a concave nondecreasing function that is strictly smaller than 1 bit per channel use, and that tends to 1 bit per channel use as the maximal allowed average power tends to infinity. Consequently, this capacity-cost function must be strictly increasing and the second moment of $(\mathbf{p}^\star, \boldsymbol{\xi}^\star)$ must therefore be $\mathcal{P}$. By noting that the capacity is achieved by some threshold quantizer, this argument also proves that $C(\mathcal{P})$ must be strictly increasing in $\mathcal{P}$. This further implies that $(\mathbf{p}^\star, \boldsymbol{\xi}^\star)$ must be centered because otherwise we could shift $\boldsymbol{\xi}^\star$ and $\Upsilon^\star$ by the mean and thus reduce the second moment without changing the mutual information.

VIII. PROOFS: CAPACITY PER UNIT ENERGY

A. Proof of Theorem 2

We will lower-bound the RHS of (16) by restricting the supremum to threshold quantizers (2) and thus demonstrate that

$$\dot{C}(0) \geq \frac{1}{2\sigma^2}. \tag{132}$$

Together with the upper bound (22), this will prove Theorem 2.

To prove (132), we first note that a threshold quantizer induces the channel

$$P\big(Y = 1 \mid X = x\big) = Q\left(\frac{\Upsilon - x}{\sigma}\right), \quad x \in \mathbb{R} \tag{133}$$

and $P\big(Y = 0 \mid X = x\big) = 1 - P\big(Y = 1 \mid X = x\big)$. By (16), we thus obtain

$$
\dot{C}(0) \geq \sup_{\xi \neq 0, \Upsilon \in \mathbb{R}} \left\{ \frac{Q\left(\frac{\Upsilon-\xi}{\sigma}\right) \log \frac{Q\left(\frac{\Upsilon-\xi}{\sigma}\right)}{Q\left(\frac{\Upsilon}{\sigma}\right)}}{\xi^2} \right.
$$

$$
\left. + \frac{\left[1 - Q\left(\frac{\Upsilon-\xi}{\sigma}\right)\right] \log \frac{1-Q\left(\frac{\Upsilon-\xi}{\sigma}\right)}{1-Q\left(\frac{\Upsilon}{\sigma}\right)}}{\xi^2} \right\}
$$

$$
= \sup_{\xi \neq 0, \Upsilon \in \mathbb{R}} \left\{ \frac{Q\left(\frac{\Upsilon-\xi}{\sigma}\right) \log \frac{1}{Q\left(\frac{\Upsilon}{\sigma}\right)}}{\xi^2} \right.
$$

$$
+ \frac{\left[1 - Q\left(\frac{\Upsilon-\xi}{\sigma}\right)\right] \log \frac{1}{1-Q\left(\frac{\Upsilon}{\sigma}\right)}}{\xi^2}
$$

$$
\left. - \frac{H_b\left(Q\left(\frac{\Upsilon-\xi}{\sigma}\right)\right)}{\xi^2} \right\}. \tag{134}
$$

We now change variables by defining $\mu \triangleq \xi - \Upsilon$ and by replacing the supremum over (ξ, Υ) with the supremum over (ξ, μ). This latter supremum we lower-bound by taking ξ to infinity while holding μ fixed. This yields for the last two terms on the RHS of (134)

$$\lim_{\xi \to \infty} \frac{H_b\left(Q\left(-\frac{\mu}{\sigma}\right)\right)}{\xi^2} = 0 \tag{135}$$

and

$$\lim_{\xi \to \infty} \frac{\left[1 - Q\left(-\frac{\mu}{\sigma}\right)\right] \log \frac{1}{1-Q\left(\frac{\xi-\mu}{\sigma}\right)}}{\xi^2} = 0. \tag{136}$$

We use the upper bound on the Q-function (47) to lower-bound the first term on the RHS of (134) as

$$\lim_{\xi \to \infty} \frac{Q\left(-\frac{\mu}{\sigma}\right) \log \frac{1}{Q\left(\frac{\xi-\mu}{\sigma}\right)}}{\xi^2}$$

$$\geq Q\left(-\frac{\mu}{\sigma}\right) \lim_{\xi \to \infty} \frac{\frac{1}{2}\log(2\pi) + \log \frac{\xi-\mu}{\sigma} + \frac{(\xi-\mu)^2}{2\sigma^2}}{\xi^2}$$

$$= Q\left(-\frac{\mu}{\sigma}\right) \frac{1}{2\sigma^2}. \tag{137}$$

Combining (135)–(137) with (134) yields

$$\dot{C}(0) \geq Q\left(-\frac{\mu}{\sigma}\right) \frac{1}{2\sigma^2} \tag{138}$$

from which we obtain (132) by letting μ tend to infinity. This proves Theorem 2.

Note that (16) is achieved by binary on–off keying [11]. By showing that (16) is lower-bounded by $1/(2\sigma^2)$ as we take ξ to infinity, we thus implicitly show that $\dot{C}(0)$ is achieved by binary on–off keying where the nonzero mass point tends to infinity as $\mathcal{P}$ tends to zero.

B. Proof of Theorem 3

We first argue that in order to prove Theorem 3, it suffices to show that for every fixed $\nu > 0$

$$\sup_{\xi^2 \leq \nu, \mathcal{D}} \frac{D\big(P_{Y|X=\xi} \,\big\|\, P_{Y|X=0}\big)}{\xi^2} < \frac{1}{2\sigma^2}. \tag{139}$$

Suppose then that this strict inequality holds for every $\nu > 0$. Consider a family of quantizers and input distributions parametrized by $\mathcal{P}$ with $\mathsf{E}\big[X^2\big] \leq \mathcal{P}$. By [11, eq. (15)], it follows that for every $\nu > 0$

$$\frac{I(X;Y)}{\mathcal{P}}$$

$$\leq \int \frac{D\big(P_{Y|X=x} \,\big\|\, P_{Y|X=0}\big)}{x^2} \frac{x^2}{\mathcal{P}} \mathrm{d}P_X(x)$$

$$= \int_{x^2 \leq \nu} \frac{D\big(P_{Y|X=x} \,\big\|\, P_{Y|X=0}\big)}{x^2} \frac{x^2}{\mathcal{P}} \mathrm{d}P_X(x)$$

$$+ \int_{x^2 > \nu} \frac{D\big(P_{Y|X=x} \,\big\|\, P_{Y|X=0}\big)}{x^2} \frac{x^2}{\mathcal{P}} \mathrm{d}P_X(x)$$

$$\leq \sup_{\xi^2 \leq \nu, \mathcal{D}} \left\{ \frac{D\big(P_{Y|X=\xi} \,\big\|\, P_{Y|X=0}\big)}{\xi^2} \right\} \frac{\mathsf{E}\big[X^2 \mathrm{I}\{X^2 \leq \nu\}\big]}{\mathcal{P}}$$

$$+ \sup_{\xi^2 > \nu, \mathcal{D}} \left\{ \frac{D\big(P_{Y|X=\xi} \,\big\|\, P_{Y|X=0}\big)}{\xi^2} \right\} \frac{\mathsf{E}\big[X^2 \mathrm{I}\{X^2 > \nu\}\big]}{\mathcal{P}}$$

$$= \sup_{\xi^2 \leq \nu, \mathcal{D}} \left\{ \frac{D\big(P_{Y|X=\xi} \,\big\|\, P_{Y|X=0}\big)}{\xi^2} \right\} \frac{\mathsf{E}\big[X^2 \mathrm{I}\{X^2 \leq \nu\}\big]}{\mathcal{P}}$$

$$+ \frac{1}{2\sigma^2} \frac{\mathsf{E}\big[X^2 \mathrm{I}\{X^2 > \nu\}\big]}{\mathcal{P}} \tag{140}$$

where the last step follows because the capacity per unit energy can be achieved by binary on–off keying where the nonzero mass point tends to infinity (see Section VII-A), so

$$\sup_{\xi^2 > \nu, \mathcal{D}} \frac{D\big(P_{Y|X=\xi} \,\big\|\, P_{Y|X=0}\big)}{\xi^2} = \frac{1}{2\sigma^2}. \tag{141}$$

Taking the limit as $\mathcal{P}$ tends to zero on both sides of (140) yields

$$\lim_{\mathcal{P} \downarrow 0} \frac{I(X;Y)}{\mathcal{P}}$$

$$\leq \lim_{\mathcal{P} \downarrow 0} \left(\frac{1}{2\sigma^2} \frac{\mathsf{E}\big[X^2 \mathrm{I}\{X^2 > \nu\}\big]}{\mathcal{P}} \right.$$

$$\left. + \sup_{\xi^2 \leq \nu, \mathcal{D}} \left\{ \frac{D\big(P_{Y|X=\xi} \| P_{Y|X=0}\big)}{\xi^2} \right\} \frac{\mathsf{E}\big[X^2 \mathrm{I}\{X^2 \leq \nu\}\big]}{\mathcal{P}} \right)$$

$$\leq \frac{1}{2\sigma^2} \tag{142}$$

where $\underline{\lim}$ denotes the *limit inferior*. Here, the last step follows from (139) and from the average-power constraint

$$\frac{\mathsf{E}\left[X^2 \mathrm{I}\left\{X^2 > \nu\right\}\right]}{\mathcal{P}} + \frac{\mathsf{E}\left[X^2 \mathrm{I}\left\{X^2 \leq \nu\right\}\right]}{\mathcal{P}} \leq 1. \quad (143)$$

Since the inequality in (139) is strict for every $\nu > 0$, it follows from (143) that the last line in (142) can hold with equality only if for every $\nu > 0$

$$\lim_{\mathcal{P}\downarrow 0} \frac{\mathsf{E}\left[X^2 \mathrm{I}\left\{X^2 > \nu\right\}\right]}{\mathcal{P}} = 1. \quad (144)$$

Thus, if (139) holds, then every family of distributions of X satisfying $\mathsf{E}\left[X^2\right] \leq \mathcal{P}$ that achieves

$$\lim_{\mathcal{P}\downarrow 0} \frac{I(X;Y)}{\mathcal{P}} = \frac{1}{2\sigma^2} \quad (145)$$

must be flash signaling, thus proving Theorem 3.

Having established that in order to prove Theorem 3, it suffices to show that (139) holds for every $\nu > 0$, we now proceed to do so. We first note that, for every $\xi \neq 0$, the supremum in (139) over all quantizers $\mathcal{D}$ can be replaced with the supremum over all threshold quantizers. Indeed, let

$$\begin{aligned} \mathcal{W} \triangleq \Big\{ (\omega_1, \omega_2) \in [0,1]^2 : \\ \omega_1 = \Pr\left(\tilde{Y} \in \mathcal{D} \mid X = \xi\right), \\ \omega_2 = \Pr\left(\tilde{Y} \in \mathcal{D} \mid X = 0\right), \mathcal{D} \subset \mathbb{R} \Big\} \end{aligned} \quad (146)$$

denote the set of possible conditional probability distributions $\left(P_{Y|X=\xi}, P_{Y|X=0}\right)$ that different quantizers can induce. Applying the methods of Section VII-B, it can be shown that the extreme points of $\overline{\mathcal{W}}$ correspond to threshold quantizers. (Recall that $\overline{\mathcal{W}}$ denotes the closure of the convex hull of $\mathcal{W}$.) Indeed, for binary inputs, the support function $f(\cdot)$ is given by (92) with $\lambda_3 = 0, \xi_1 = \xi$, and $\xi_2 = 0$. The quantization region $\mathcal{D}^\star(\boldsymbol{\lambda})$ that achieves the supremum in (92) consists of the set of $\tilde{y} \in \mathbb{R}$ for which $g_{\boldsymbol{\lambda}}(\tilde{y})$ in (95) is nonnegative. Since $g_{\boldsymbol{\lambda}}(\cdot)$ has at most one zero, it follows that $\mathcal{D}^\star(\boldsymbol{\lambda})$ consists of at most two regions, i.e., it is a threshold quantizer. Using that the relative entropy on the LHS of (139) is convex in $\left(P_{Y|X=\xi}, P_{Y|X=0}\right)$ [10, Th. 2.7.2], it follows by the same arguments as in Section VII-B that, for every $\xi \neq 0$, $D\left(P_{Y|X=\xi} \,\|\, P_{Y|X=0}\right)$ is maximized by some threshold quantizer.

We next note that we can assume, without loss of optimality, that the threshold Υ of the quantizer is nonnegative. Consequently, the supremum over $\mathcal{D}$ on the LHS of (139) can be replaced by a supremum over threshold quantizers of nonnegative thresholds $\Upsilon \geq 0$. Indeed, for $x \in \mathbb{R}$,

$$\Pr\left(\tilde{Y} \geq \Upsilon \mid X = x\right) = 1 - \Pr\left(\tilde{Y} \geq -\Upsilon \mid X = -x\right) \quad (147)$$

and consequently,

$$\begin{aligned} D\left(P_{Y|X=\xi} \,\|\, P_{Y|X=0}\right)\big|_{\mathcal{D}=\{\tilde{y} \in \mathbb{R}: \tilde{y} \geq \Upsilon\}} \\ = D\left(P_{Y|X=-\xi} \,\|\, P_{Y|X=0}\right)\big|_{\mathcal{D}=\{\tilde{y} \in \mathbb{R}: \tilde{y} \geq -\Upsilon\}}. \end{aligned} \quad (148)$$

Thus, to every pair (ξ, Υ) corresponds another pair $(-\xi, -\Upsilon)$ achieving the same relative entropy. Since ξ and $-\xi$ have the same magnitude, this implies that both pairs give rise to the same value for

$$\frac{D\left(P_{Y|X=\xi} \,\|\, P_{Y|X=0}\right)}{\xi^2}.$$

Hence, we can assume without loss of generality that $\Upsilon \geq 0$.

We continue by defining the random variable U as

$$U \triangleq \tilde{Y} \mathrm{I}\left\{\tilde{Y} \geq 0\right\}. \quad (149)$$

Note that, for $\Upsilon \geq 0$, the quantizer's output can be expressed as $Y = \mathrm{I}\left\{U \geq \Upsilon\right\}$. It thus follows from the Data Processing Inequality for Relative Entropy [10, Sec. 2.9] that

$$\begin{aligned} D\big(P_{Y|X=\xi} & \,\|\, P_{Y|X=0}\big) \\ & \leq D\big(P_{U|X=\xi} \,\|\, P_{U|X=0}\big) \\ & = \frac{1}{\sqrt{2\pi\sigma^2}} \int_0^\infty e^{-\frac{(\tilde{y}-\xi)^2}{2\sigma^2}} \log \frac{e^{-\frac{(\tilde{y}-\xi)^2}{2\sigma^2}}}{e^{-\frac{\tilde{y}^2}{2\sigma^2}}} \mathrm{d}\tilde{y} \\ & \quad + \frac{1}{\sqrt{2\pi\sigma^2}} \left(\int_{-\infty}^0 e^{-\frac{(\tilde{y}-\xi)^2}{2\sigma^2}} \mathrm{d}\tilde{y}\right) \log \frac{\int_{-\infty}^0 e^{-\frac{(\tilde{y}-\xi)^2}{2\sigma^2}} \mathrm{d}\tilde{y}}{\int_{-\infty}^0 e^{-\frac{\tilde{y}^2}{2\sigma^2}} \mathrm{d}\tilde{y}} \\ & \triangleq \Psi(\xi) \end{aligned} \quad (150)$$

irrespective of the threshold $\Upsilon \geq 0$. Here, the last equality should be viewed as the definition of $\Psi(\xi)$. By applying the Log-Sum Inequality [10, Th. 2.7.1] to $\Psi(\xi)$, we obtain

$$\begin{aligned} \Psi(\xi) & \leq \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty e^{-\frac{(\tilde{y}-\xi)^2}{2\sigma^2}} \log \frac{e^{-\frac{(\tilde{y}-\xi)^2}{2\sigma^2}}}{e^{-\frac{\tilde{y}^2}{2\sigma^2}}} \mathrm{d}\tilde{y} \\ & = \frac{\xi^2}{2\sigma^2} \end{aligned} \quad (151)$$

with equality if, and only if,

$$\frac{e^{-\frac{(\tilde{y}-\xi)^2}{2\sigma^2}}}{e^{-\frac{\tilde{y}^2}{2\sigma^2}}} = 2 Q\left(\frac{\xi}{\sigma}\right), \quad \text{for almost every } \tilde{y} \leq 0. \quad (152)$$

Since (152) holds only for $\xi = 0$, this yields

$$\Psi(\xi) < \frac{\xi^2}{2\sigma^2}, \quad \xi \neq 0. \quad (153)$$

Note that (153) and (151) give an upper bound on the relative entropy that does not depend on the threshold. By combining (150) and (153), and recalling that for every $\xi \neq 0$ the relative entropy in (139) is maximized by some threshold quantizer, we obtain

$$\sup_{\mathcal{D}} \frac{D\left(P_{Y|X=\xi} \,\|\, P_{Y|X=0}\right)}{\xi^2} \leq \frac{\Psi(\xi)}{\xi^2} < \frac{1}{2\sigma^2}, \quad \xi \neq 0. \quad (154)$$

Since the function $\xi \mapsto \xi^{-2}\Psi(\xi)$ is continuous on $\mathbb{R} \setminus \{0\}$ and, as shown in Appendix III, satisfies

$$\lim_{\xi \to 0} \frac{\Psi(\xi)}{\xi^2} = \frac{1}{2\sigma^2}\left(\frac{1}{2} + \frac{1}{\pi}\right) < \frac{1}{2\sigma^2} \quad (155)$$

we obtain (139) by maximizing (154) over $\xi^2 \leq \nu$. This proves Theorem 3.

C. Proof of Corollary 1

To prove Corollary 1, we need to show that for every $\nu > 0$ and every threshold quantizer with threshold $0 \leq \Upsilon \leq \nu$,

$$\sup_{\xi \neq 0, 0 \leq \Upsilon \leq \nu} \frac{D\left(P_{Y|X=\xi} \,\|\, P_{Y|X=0}\right)}{\xi^2} < \frac{1}{2\sigma^2}. \tag{156}$$

By (154), we have that for every $\xi \neq 0$ and every $\nu > 0$

$$\sup_{0 \leq \Upsilon \leq \nu} \frac{D\left(P_{Y|X=\xi} \,\|\, P_{Y|X=0}\right)}{\xi^2} \leq \frac{\Psi(\xi)}{\xi^2} < \frac{1}{2\sigma^2} \tag{157}$$

where $\xi \mapsto \xi^{-2} \Psi(\xi)$ is continuous on $\mathbb{R} \setminus \{0\}$ and satisfies (155). To conclude the proof of the corollary, it thus remains to show that for every $\nu > 0$

$$\varlimsup_{\xi^2 \to \infty} \sup_{0 \leq \Upsilon \leq \nu} \frac{D\left(P_{Y|X=\xi} \,\|\, P_{Y|X=0}\right)}{\xi^2} < \frac{1}{2\sigma^2} \tag{158}$$

where $\varlimsup$ denotes the *limit superior*. This can be done by noting that for $0 \leq \Upsilon \leq \nu$

$$D\left(P_{Y|X=\xi} \,\|\, P_{Y|X=0}\right)$$
$$= Q\left(\frac{\Upsilon - \xi}{\sigma}\right) \log \frac{1}{Q\left(\frac{\Upsilon}{\sigma}\right)} - H_b\left(Q\left(\frac{\Upsilon - \xi}{\sigma}\right)\right)$$
$$+ \left[1 - Q\left(\frac{\Upsilon - \xi}{\sigma}\right)\right] \log \frac{1}{1 - Q\left(\frac{\Upsilon}{\sigma}\right)}$$
$$\leq \log \frac{1}{Q\left(\frac{\Upsilon}{\sigma}\right)} + \log \frac{1}{1 - Q\left(\frac{\Upsilon}{\sigma}\right)}$$
$$\leq \log \frac{1}{Q\left(\frac{\nu}{\sigma}\right)} + \log 2 \tag{159}$$

where the second step follows because $0 \leq Q(x) \leq 1$, $x \in \mathbb{R}$ and $H_b(p) \geq 0$, $0 \leq p \leq 1$, and where the last step follows because $x \mapsto Q(x)$ is monotonically decreasing in $x \in \mathbb{R}$ and because $0 \leq \Upsilon \leq \nu$. Computing the limiting ratio of the RHS of (159) to ξ^2 as ξ^2 tends to infinity yields for every $\nu > 0$

$$\lim_{\xi^2 \to \infty} \sup_{0 \leq \Upsilon \leq \nu} \frac{D\left(P_{Y|X=\xi} \,\|\, P_{Y|X=0}\right)}{\xi^2} = 0 \tag{160}$$

thus establishing (158). This proves Corollary 1.

IX. PROOFS: PEAK-POWER-LIMITED CHANNELS

A. Proof of Proposition 1

The peak-power-limited Gaussian channel with 1-bit output quantization is a memoryless channel with a continuous input taking values in $\left[-\sqrt{\mathcal{P}}, \sqrt{\mathcal{P}}\right]$ and a binary output. It thus follows from Dubins's Theorem that, for every quantization region $\mathcal{D}$, the capacity-achieving input distribution is discrete with two mass points [23, Sec. II-C]. We shall denote these two mass points by ξ_1 and ξ_2.

We next argue that threshold quantizers are optimal. Let $\mathcal{W}$ denote the set of all possible channel laws, i.e.,

$$\mathcal{W} \triangleq \Big\{(\omega_1, \omega_2) \in [0, 1]^2 :$$
$$\omega_\ell = \Pr\left(\tilde{Y} \in \mathcal{D} \mid X = \xi_\ell\right), \mathcal{D} \subset \mathbb{R}\Big\}. \tag{161}$$

Applying the methods of Section VII-B to binary channel inputs, it can be shown that the extreme points of $\overline{\mathcal{W}}$ correspond to threshold quantizers (2) or complements thereof. (For more details, see also Section VIII-B.) By the same arguments as in Section VII-B, it follows that for every binary random variable X, the mutual information $I(X; Y)$ is maximized by some threshold quantizer.

The capacity of the peak-power-limited Gaussian channel with 1-bit output quantization is thus given by

$$C_{\text{PP}}(\mathcal{P}) = \sup_{(\mathbf{p}, \boldsymbol{\xi}), \Upsilon \in \mathbb{R}} I\left(\mathbf{p}, \mathbf{W}(\Upsilon|\boldsymbol{\xi})\right) \tag{162}$$

where $(\mathbf{p}, \boldsymbol{\xi})$ denotes the two-mass-point distribution with masses

$$\mathbf{p} = (p_1, p_2) \in [0, 1]^2$$

and locations

$$\boldsymbol{\xi} = (\xi_1, \xi_2) \in [-\sqrt{\mathcal{P}}, \sqrt{\mathcal{P}}]^2$$

and where $\mathbf{W}(\Upsilon|\boldsymbol{\xi})$ denotes the channel law corresponding to the threshold quantizer (2) and to the mass points (ξ_1, ξ_2):

$$W(\Upsilon|\xi_\ell) = \Pr\left(\tilde{Y} \geq \Upsilon \mid X = \xi_\ell\right), \quad \ell = 1, 2. \tag{163}$$

Following the steps in Section VII-C, it can be further shown that the supremum on the RHS of (162) is achieved.

In the following, we demonstrate that there is no loss in optimality in assuming that the mass points of the capacity-achieving input distribution are located at $-\sqrt{\mathcal{P}}$ and $\sqrt{\mathcal{P}}$. Indeed, suppose that the optimal mass points are located at

$$-\sqrt{\mathcal{P}} \leq \xi_1 < \xi_2 < \sqrt{\mathcal{P}}. \tag{164}$$

Then, it follows from the strict monotonicity of the Q-function that

$$Q\left(\frac{\Upsilon - \xi_1}{\sigma}\right) < Q\left(\frac{\Upsilon - \xi_2}{\sigma}\right) < Q\left(\frac{\Upsilon - \sqrt{\mathcal{P}}}{\sigma}\right). \tag{165}$$

Since $W(\Upsilon|\xi_1)$ does not depend on ξ_2, this implies that for every Υ and ξ_1, the channel law $\mathbf{W}(\Upsilon|\boldsymbol{\xi})$ can be written as a convex combination of $\mathbf{W}(\Upsilon|\boldsymbol{\psi})$ and $\mathbf{W}(\Upsilon|\boldsymbol{\zeta})$, where $\boldsymbol{\psi} = (\xi_1, \xi_1)$ and $\boldsymbol{\zeta} = (\xi_1, \sqrt{\mathcal{P}})$. By the convexity of mutual information in the channel law, and by noting that $I\left(\mathbf{p}, \mathbf{W}(\Upsilon|\boldsymbol{\psi})\right) = 0$, it follows that

$$I\left(\mathbf{p}, \mathbf{W}(\Upsilon|\boldsymbol{\xi})\right) \leq I\left(\mathbf{p}, \mathbf{W}(\Upsilon|\boldsymbol{\zeta})\right) \tag{166}$$

for every Υ and $(\mathbf{p}, \boldsymbol{\xi})$ satisfying (164). Thus, $\xi_2 = \sqrt{\mathcal{P}}$ achieves the capacity. By repeating the same arguments for ξ_1, we obtain that the mass points of the capacity-achieving input

distribution are located at $-\sqrt{\mathcal{P}}$ and $\sqrt{\mathcal{P}}$. It follows that the capacity can be expressed as

$$C_{\mathrm{PP}}(\mathcal{P}) = \max_{\Upsilon \in \mathbb{R}} C_{\Upsilon}(\mathcal{P}) \qquad (167)$$

where $C_{\Upsilon}(\mathcal{P})$ denotes the capacity of the binary asymmetric channel with crossover probabilities

$$W(0|1) = Q\left(\frac{\sqrt{\mathcal{P}} - \Upsilon}{\sigma}\right) \qquad (168a)$$

$$W(1|0) = Q\left(\frac{\sqrt{\mathcal{P}} + \Upsilon}{\sigma}\right). \qquad (168b)$$

For every $\Upsilon \in \mathbb{R}$, the capacity of the binary asymmetric channel can be computed as

$$C_{\Upsilon}(\mathcal{P}) = \log\left(1 + e^{-\theta}\right) + \theta\, W(1|0) - H_b\left(W(1|0)\right) \quad (169)$$

where

$$\theta \triangleq \frac{H_b\left(W(0|1)\right) - H_b\left(W(1|0)\right)}{1 - W(0|1) - W(1|0)}. \qquad (170)$$

Combining (169), (168a), and (168b) with (167) yields

$$C_{\mathrm{PP}}(\mathcal{P}) = \max_{\Upsilon \in \mathbb{R}} \left\{ \log\left(1 + e^{-\Theta(\mathcal{P}, \Upsilon)}\right) \right.$$
$$\left. + Q\left(\frac{\sqrt{\mathcal{P}} + \Upsilon}{\sigma}\right)\Theta(\mathcal{P}, \Upsilon) - H_b\left(Q\left(\frac{\sqrt{\mathcal{P}} + \Upsilon}{\sigma}\right)\right) \right\}$$
$$(171)$$

where

$$\Theta(\mathcal{P}, \Upsilon) \triangleq \frac{H_b\left(Q\left(\frac{\sqrt{\mathcal{P}} - \Upsilon}{\sigma}\right)\right) - H_b\left(Q\left(\frac{\sqrt{\mathcal{P}} + \Upsilon}{\sigma}\right)\right)}{1 - Q\left(\frac{\sqrt{\mathcal{P}} - \Upsilon}{\sigma}\right) - Q\left(\frac{\sqrt{\mathcal{P}} + \Upsilon}{\sigma}\right)}. \quad (172)$$

Proposition 1 follows then by noting that the RHS of (171) is symmetric in $\Upsilon \in \mathbb{R}$, so the maximization in (171) can be restricted to $\Upsilon \geq 0$ without reducing (171).

B. Proof of Proposition 2

It was shown in the previous section that the capacity is achieved with a threshold quantizer and a binary input distribution having mass points at $\sqrt{\mathcal{P}}$ and $-\sqrt{\mathcal{P}}$. Thus, the capacity can be expressed as

$$C_{\mathrm{PP}}(\mathcal{P}) = \max_{\Upsilon \geq 0} \left\{ H_b\left(p_+ Q\left(\frac{\Upsilon - \mathcal{A}}{\sigma}\right) + p_- Q\left(\frac{\Upsilon + \mathcal{A}}{\sigma}\right)\right) \right.$$
$$\left. - p_+ H_b\left(Q\left(\frac{\Upsilon - \mathcal{A}}{\sigma}\right)\right) - p_- H_b\left(Q\left(\frac{\Upsilon + \mathcal{A}}{\sigma}\right)\right) \right\} \quad (173)$$

for some probabilities $0 < p_+ < 1$ and $0 < p_- < 1$ satisfying $p_+ + p_- = 1$. To simplify notation, we have introduced $\mathcal{A} \triangleq \sqrt{\mathcal{P}}$ and we have made the dependence of p_+ and p_- on Υ implicit.

Expanding $H_b(\cdot)$ as a Taylor series around $Q(\Upsilon/\sigma)$, we obtain for the first term on the RHS of (173)

$$H_b\left(p_+ Q\left(\frac{\Upsilon - \mathcal{A}}{\sigma}\right) + p_- Q\left(\frac{\Upsilon + \mathcal{A}}{\sigma}\right)\right)$$
$$= H_b\left(Q\left(\frac{\Upsilon}{\sigma}\right)\right) + \log\frac{1 - Q\left(\frac{\Upsilon}{\sigma}\right)}{Q\left(\frac{\Upsilon}{\sigma}\right)}$$
$$\times \left[p_+ Q\left(\frac{\Upsilon - \mathcal{A}}{\sigma}\right) + p_- Q\left(\frac{\Upsilon + \mathcal{A}}{\sigma}\right) - Q\left(\frac{\Upsilon}{\sigma}\right)\right]$$
$$- \frac{1}{2Q\left(\frac{\Upsilon}{\sigma}\right)\left[1 - Q\left(\frac{\Upsilon}{\sigma}\right)\right]}$$
$$\times \left[p_+ Q\left(\frac{\Upsilon - \mathcal{A}}{\sigma}\right) + p_- Q\left(\frac{\Upsilon + \mathcal{A}}{\sigma}\right) - Q\left(\frac{\Upsilon}{\sigma}\right)\right]^2$$
$$+ \mathsf{R}_H(\mathcal{A}, \Upsilon, p_+) \qquad (174)$$

where

$$\mathsf{R}_H(\mathcal{A}, \Upsilon, p_+) \triangleq \frac{1 - 2\tilde{p}}{6\tilde{p}(1 - \tilde{p})}$$
$$\times \left[p_+ Q\left(\frac{\Upsilon - \mathcal{A}}{\sigma}\right) + p_- Q\left(\frac{\Upsilon + \mathcal{A}}{\sigma}\right) - Q\left(\frac{\Upsilon}{\sigma}\right)\right]^3$$
$$(175)$$

for some $\tilde{p} \in \left[Q((\Upsilon + \mathcal{A})/\sigma), Q((\Upsilon - \mathcal{A})/\sigma)\right]$. Expanding the Q-function as a Taylor series around Υ/σ yields

$$p_+ Q\left(\frac{\Upsilon - \mathcal{A}}{\sigma}\right) + p_- Q\left(\frac{\Upsilon + \mathcal{A}}{\sigma}\right) - Q\left(\frac{\Upsilon}{\sigma}\right)$$
$$= (p_+ - p_-)\frac{\mathcal{A}}{\sigma}\frac{1}{\sqrt{2\pi}}e^{-\frac{\Upsilon^2}{2\sigma^2}} + \mathsf{R}_Q(\mathcal{A}, \Upsilon, p_+) \quad (176)$$

where

$$\mathsf{R}_Q(\mathcal{A}, \Upsilon, p_+) \triangleq \frac{\mathcal{A}^2}{2\sigma^2}\frac{\tilde{x}}{\sqrt{2\pi\sigma^2}}e^{-\frac{\tilde{x}^2}{2\sigma^2}} \qquad (177)$$

for some $\tilde{x} \in [\Upsilon - \mathcal{A}, \Upsilon + \mathcal{A}]$. Note that

$$\left|\tilde{x}\exp\left(-\tilde{x}^2/(2\sigma^2)\right)\right| \leq \sigma/\sqrt{e} \qquad (178)$$

so $\mathsf{R}_Q(\mathcal{A}, \Upsilon, p_+)$ satisfies

$$\left|\mathsf{R}_Q(\mathcal{A}, \Upsilon, p_+)\right| \leq \frac{\mathcal{A}^2}{2\sigma^2\sqrt{2\pi e}}, \qquad 0 \leq p_+ \leq 1. \quad (179)$$

Combining (176) with (174), we obtain for the first term on the RHS of (173)

$$
\begin{aligned}
&H_b\left(p_+ Q\left(\frac{\Upsilon - \mathcal{A}}{\sigma}\right) + p_- Q\left(\frac{\Upsilon + \mathcal{A}}{\sigma}\right)\right) \\
&= H_b\left(Q\left(\frac{\Upsilon}{\sigma}\right)\right) + \log\frac{1 - Q\left(\frac{\Upsilon}{\sigma}\right)}{Q\left(\frac{\Upsilon}{\sigma}\right)} \\
&\quad \times \left[p_+ Q\left(\frac{\Upsilon - \mathcal{A}}{\sigma}\right) + p_- Q\left(\frac{\Upsilon + \mathcal{A}}{\sigma}\right) - Q\left(\frac{\Upsilon}{\sigma}\right)\right] \\
&\quad - \frac{1}{2 Q\left(\frac{\Upsilon}{\sigma}\right)\left[1 - Q\left(\frac{\Upsilon}{\sigma}\right)\right]} \\
&\quad \times \left[(p_+ - p_-)\frac{\mathcal{A}}{\sigma}\frac{1}{\sqrt{2\pi}}e^{-\frac{\Upsilon^2}{2\sigma^2}} + \mathsf{R}_Q(\mathcal{A}, \Upsilon, p_+)\right]^2 \\
&\quad + \mathsf{R}_H(\mathcal{A}, \Upsilon, p_+) \\
&= H_b\left(Q\left(\frac{\Upsilon}{\sigma}\right)\right) + \log\frac{1 - Q\left(\frac{\Upsilon}{\sigma}\right)}{Q\left(\frac{\Upsilon}{\sigma}\right)} \\
&\quad \times \left[p_+ Q\left(\frac{\Upsilon - \mathcal{A}}{\sigma}\right) + p_- Q\left(\frac{\Upsilon + \mathcal{A}}{\sigma}\right) - Q\left(\frac{\Upsilon}{\sigma}\right)\right] \\
&\quad - \frac{\mathcal{A}^2}{\sigma^2}\frac{e^{-\frac{\Upsilon^2}{\sigma^2}}}{4\pi Q\left(\frac{\Upsilon}{\sigma}\right)\left[1 - Q\left(\frac{\Upsilon}{\sigma}\right)\right]}(p_+ - p_-)^2 \\
&\quad + \mathsf{K}(\mathcal{A}, \Upsilon, p_+) + \mathsf{R}_H(\mathcal{A}, \Upsilon, p_+)
\end{aligned} \tag{180}
$$

where

$$
\begin{aligned}
\mathsf{K}(\mathcal{A}, \Upsilon, p_+) &\triangleq -\frac{2(p_+ - p_-)\frac{1}{\sqrt{2\pi}}e^{-\frac{\Upsilon^2}{2\sigma^2}}\frac{\mathcal{A}}{\sigma}\mathsf{R}_Q(\mathcal{A}, \Upsilon, p_+)}{2 Q\left(\frac{\Upsilon}{\sigma}\right)\left[1 - Q\left(\frac{\Upsilon}{\sigma}\right)\right]} \\
&\quad - \frac{\left|\mathsf{R}_Q(\mathcal{A}, \Upsilon, p_+)\right|^2}{2 Q\left(\frac{\Upsilon}{\sigma}\right)\left[1 - Q\left(\frac{\Upsilon}{\sigma}\right)\right]}.
\end{aligned} \tag{181}
$$

Taylor series expansions for the last two terms on the RHS of (173) follow directly from (180) by setting p_+ to 1 and to 0. Thus, by applying (180) to (173), and by using that $p_+ + p_- = 1$, we obtain

$$
\begin{aligned}
C_{\text{PP}}(\mathcal{P}) = \max_{\Upsilon \geq 0}\Bigg\{ &\frac{\mathcal{A}^2}{\sigma^2}\frac{e^{-\frac{\Upsilon^2}{\sigma^2}}}{4\pi Q\left(\frac{\Upsilon}{\sigma}\right)\left[1 - Q\left(\frac{\Upsilon}{\sigma}\right)\right]}\left[1 - (p_+ - p_-)^2\right] \\
&+ \mathsf{K}(\mathcal{A}, \Upsilon, p_+) + \mathsf{R}_H(\mathcal{A}, \Upsilon, p_+) \\
&- p_+\left[\mathsf{K}(\mathcal{A}, \Upsilon, 1) + \mathsf{R}_H(\mathcal{A}, \Upsilon, 1)\right] \\
&- p_-\left[\mathsf{K}(\mathcal{A}, \Upsilon, 0) + \mathsf{R}_H(\mathcal{A}, \Upsilon, 0)\right]\Bigg\}.
\end{aligned} \tag{182}
$$

As shown in Appendix IV,

$$
\limsup_{\mathcal{A}\downarrow 0}\sup_{\Upsilon \geq 0}\frac{\left|\mathsf{R}_H(\mathcal{A}, \Upsilon, p_+)\right|}{\mathcal{A}^2} = 0, \quad 0 \leq p_+ \leq 1 \tag{183a}
$$

$$
\limsup_{\mathcal{A}\downarrow 0}\sup_{\Upsilon \geq 0}\frac{\left|\mathsf{K}(\mathcal{A}, \Upsilon, p_+)\right|}{\mathcal{A}^2} = 0, \quad 0 \leq p_+ \leq 1. \tag{183b}
$$

Using (183a), (183b), and the Triangle Inequality, (182) can thus be upper-bounded by

$$
C_{\text{PP}}(\mathcal{P}) \leq \sup_{\Upsilon \geq 0}\frac{\mathcal{A}^2}{\sigma^2}\frac{e^{-\frac{\Upsilon^2}{\sigma^2}}\left[1 - (p_+ - p_-)^2\right]}{4\pi Q\left(\frac{\Upsilon}{\sigma}\right)\left[1 - Q\left(\frac{\Upsilon}{\sigma}\right)\right]} + o\left(\mathcal{A}^2\right) \tag{184}
$$

where $\lim_{\mathcal{A}\downarrow 0} o\left(\mathcal{A}^2\right)/\mathcal{A}^2 = 0$. Consequently, dividing (184) by $\mathcal{P} = \mathcal{A}^2$ and computing the limit as $\mathcal{P}$ tends to zero, yields

$$
\begin{aligned}
\lim_{\mathcal{P}\downarrow 0}\frac{C_{\text{PP}}(\mathcal{P})}{\mathcal{P}} &\leq \sup_{\Upsilon \geq 0}\frac{1}{\sigma^2}\frac{e^{-\frac{\Upsilon^2}{\sigma^2}}\left[1 - (p_+ - p_-)^2\right]}{4\pi Q\left(\frac{\Upsilon}{\sigma}\right)\left[1 - Q\left(\frac{\Upsilon}{\sigma}\right)\right]} \\
&\leq \sup_{\Upsilon \geq 0}\frac{e^{-\frac{\Upsilon^2}{\sigma^2}}}{4\pi Q\left(\frac{\Upsilon}{\sigma}\right)\left[1 - Q\left(\frac{\Upsilon}{\sigma}\right)\right]}\frac{1}{\sigma^2}
\end{aligned} \tag{185}
$$

where the second inequality holds with equality for $p_+ = p_- = 1/2$.

It remains to show that the maximum on the RHS of (185) is attained for $\Upsilon = 0$. To this end, we argue that the function

$$
f(\Upsilon) \triangleq \frac{e^{-\frac{\Upsilon^2}{\sigma^2}}}{Q\left(\frac{\Upsilon}{\sigma}\right)\left[1 - Q\left(\frac{\Upsilon}{\sigma}\right)\right]}, \quad \Upsilon \geq 0 \tag{186}
$$

is monotonically decreasing in $\Upsilon \geq 0$. Indeed, the first derivative of $f(\cdot)$ is given by

$$
f'(\Upsilon) = -\frac{\frac{1}{\sigma}e^{-\frac{\Upsilon^2}{\sigma^2}}}{\left[Q\left(\frac{\Upsilon}{\sigma}\right)\right]^2\left[1 - Q\left(\frac{\Upsilon}{\sigma}\right)\right]^2}g\left(\frac{\Upsilon}{\sigma}\right), \quad \Upsilon \geq 0 \tag{187}
$$

where

$$
g(u) \triangleq 2uQ(u)[1 - Q(u)] - \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}[1 - 2Q(u)] \tag{188}
$$

for $u \geq 0$. For $u \geq 2$, we lower-bound the Q-function as [14, Prop. 19.4.2]

$$
Q(u) > \frac{3}{4}\frac{1}{\sqrt{2\pi}u}e^{-\frac{u^2}{2}}, \quad u \geq 2 \tag{189}
$$

to obtain

$$
\begin{aligned}
g(u) &> \frac{3}{2}\frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}[1 - Q(u)] - \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}[1 - 2Q(u)] \\
&= \frac{e^{-\frac{u^2}{2}}}{\sqrt{8\pi}}[1 + Q(u)] \\
&> 0.
\end{aligned} \tag{190}
$$

For $0 \leq u \leq 2$, it can be shown numerically that $g(u) \geq 0$; see Fig. 3.

It thus follows that $g\left(\Upsilon/\sigma\right) \geq 0$, $\Upsilon/\sigma \geq 0$ and hence, by (187), $f'(\Upsilon) \leq 0$, $\Upsilon \geq 0$. Consequently,

$$
\max_{\Upsilon \geq 0}f(\Upsilon) = f(0) = 4 \tag{191}
$$

which together with (185) yields

$$
\lim_{\mathcal{P}\downarrow 0}\frac{C_{\text{PP}}(\mathcal{P})}{\mathcal{P}} \leq \frac{1}{\pi\sigma^2}. \tag{192}
$$

Noting that the RHS of (192) is achieved for $p_+ = p_- = 1/2$ and a symmetric threshold quantizer [cf. (26)], this proves Proposition 2.
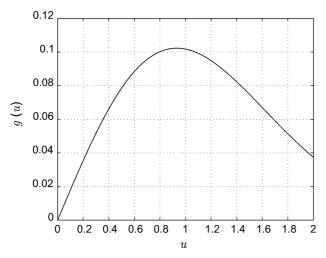
Fig. 3. Function $u \mapsto g(u)$ for $0 \leq u \leq 2$.

X. PROOFS: FADING CHANNELS

A. Proof of Theorem 4

We will lower-bound the RHS of (68) by restricting the supremum to radial quantizers

$$\mathcal{D} = \{\tilde{y} \in \mathbb{C} : |\tilde{y}| \geq \Upsilon\}, \quad \Upsilon > 0 \qquad (193)$$

and thus demonstrate that

$$\dot{C}(0) \geq \frac{1}{\sigma^2}. \qquad (194)$$

Together with the upper bound (69), this will prove Theorem 4.

To prove (194), note that, conditioned on $(H, X) = (h, x)$, the squared magnitude of $\sqrt{2/\sigma^2}\tilde{Y}$ has a noncentral chi-square distribution with two degrees of freedom and noncentrality parameter $\frac{2}{\sigma^2}|h|^2|x|^2$ [25, p. 8]. Consequently, a radial quantizer induces the channel [25, Sec. 2-E]

$$\Pr(Y = 1 \mid H = h, X = x)$$
$$= Q_1\left(\sqrt{\frac{2}{\sigma^2}}|h||x|, \sqrt{\frac{2}{\sigma^2}}\Upsilon\right) \qquad (195)$$

for $h \in \mathbb{C}$, $x \in \mathbb{C}$, and $\Upsilon > 0$, where $Q_1(\cdot, \cdot)$ denotes the first-order Marcum Q-function [25, eq. (2.20)]. For $x = 0$, this becomes
$$\Pr(Y = 1 \mid H = h, X = 0) = e^{-\frac{\Upsilon^2}{\sigma^2}} \qquad (196)$$

for $h \in \mathbb{C}$ and $\Upsilon > 0$. This yields

$$D\left(P_{Y|H,X=\xi} \,\middle\|\, P_{Y|H,X=0} \,\middle|\, P_H\right)$$
$$= \mathsf{E}\left[Q_1\left(\sqrt{\frac{2}{\sigma^2}}|H||\xi|, \sqrt{\frac{2}{\sigma^2}}\Upsilon\right) \log \frac{1}{e^{-\frac{\Upsilon^2}{\sigma^2}}}\right]$$
$$+ \mathsf{E}\left[\left\{1 - Q_1\left(\sqrt{\frac{2}{\sigma^2}}|H||\xi|, \sqrt{\frac{2}{\sigma^2}}\Upsilon\right)\right\} \log \frac{1}{1 - e^{-\frac{\Upsilon^2}{\sigma^2}}}\right]$$
$$- \mathsf{E}\left[H_b\left(Q_1\left(\sqrt{\frac{2}{\sigma^2}}|H||\xi|, \sqrt{\frac{2}{\sigma^2}}\Upsilon\right)\right)\right] \qquad (197a)$$
$$\geq \mathsf{E}\left[Q_1\left(\sqrt{\frac{2}{\sigma^2}}|H||\xi|, \sqrt{\frac{2}{\sigma^2}}\Upsilon\right)\frac{\Upsilon^2}{\sigma^2}\right] - \log 2 \qquad (197b)$$

where (197b) follows because the second term in (197a) is nonnegative, and because the binary entropy function is upper-bounded by $\log 2$.

By applying (197b) to (68), we obtain

$$\dot{C}(0) \geq \sup_{\substack{\xi \neq 0, \\ \Upsilon > 0}} \left\{ \mathsf{E}\left[Q_1\left(\sqrt{\frac{2}{\sigma^2}}|H||\xi|, \sqrt{\frac{2}{\sigma^2}}\Upsilon\right)\frac{\Upsilon^2}{|\xi|^2\sigma^2}\right] \right.$$
$$\left. - \frac{1}{|\xi|^2}\log 2\right\}. \qquad (198)$$

We lower-bound the supremum on the RHS of (198) by choosing $\Upsilon = \mu|h||\xi|$ for some fixed $0 < \mu < 1$ and by taking $|\xi|$ to infinity. We then lower-bound the first-order Marcum Q-function using [25, Sec. C-2, eq. (C.24)]

$$Q_1(\alpha, \beta)$$
$$\geq 1 - \frac{1}{2}\left[\exp\left(-\frac{(\alpha - \beta)^2}{2}\right) - \exp\left(-\frac{(\alpha + \beta)^2}{2}\right)\right] \qquad (199)$$

for $\alpha > \beta \geq 0$. This yields

$$\dot{C}(0)$$
$$\geq \frac{\mu^2 \mathsf{E}[|H|^2]}{\sigma^2}$$
$$- \lim_{|\xi|\to\infty} \frac{1}{2|\xi|^2} \mathsf{E}\left[\exp\left(-\frac{|H|^2|\xi|^2}{\sigma^2}(1-\mu)^2\right)\frac{\mu^2|H|^2|\xi|^2}{\sigma^2}\right]$$
$$+ \lim_{|\xi|\to\infty} \frac{1}{2|\xi|^2} \mathsf{E}\left[\exp\left(-\frac{|H|^2|\xi|^2}{\sigma^2}(1+\mu)^2\right)\frac{\mu^2|H|^2|\xi|^2}{\sigma^2}\right]$$
$$\geq \frac{\mu^2 \mathsf{E}[|H|^2]}{\sigma^2} - \lim_{|\xi|\to\infty} \frac{\mu^2}{2|\xi|^2 e\,(1-\mu)^2}$$
$$= \frac{\mu^2 \mathsf{E}[|H|^2]}{\sigma^2} \qquad (200)$$

where the second step follows because $0 \leq xe^{-\alpha x} \leq 1/(e\alpha)$ for every $x \geq 0$ and $\alpha > 0$. This establishes (194) because H is of unit variance and μ can be arbitrarily close to 1.

B. Proof of Theorem 5

By the Data Processing Inequality for Relative Entropy, the relative entropy on the RHS of (72) is upper-bounded by the relative entropy corresponding to the unquantized channel, i.e., [3, eq. (64)]

$$\frac{D\left(P_{Y|X=\xi} \,\middle\|\, P_{Y|X=0}\right)}{|\xi|^2} \leq \frac{1}{\sigma^2} - \frac{\log\left(1 + \frac{|\xi|^2}{\sigma^2}\right)}{|\xi|^2}. \qquad (201)$$

Consequently, the capacity per unit energy (72) is strictly smaller than $1/\sigma^2$ unless the supremum on the RHS of (72) is approached as $|\xi|$ tends to infinity. It thus remains to show that

$$\varlimsup_{|\xi|\to\infty} \sup_{\mathcal{D}} \frac{D\left(P_{Y|X=\xi} \,\middle\|\, P_{Y|X=0}\right)}{|\xi|^2} < \frac{1}{\sigma^2}. \qquad (202)$$

To this end, we first note that, for every $\xi \neq 0$, the supremum in (202) over all quantizers $\mathcal{D}$ can be replaced with the supremum over all *radial* quantizers (193). Indeed, for every quantization region satisfying

$$\Pr(Y = 1 \mid X = \xi) = \beta, \quad 0 < \beta < 1$$

the relative entropy

$$D\big(P_{Y|X=\xi} \,\|\, P_{Y|X=0}\big)$$
$$= \beta \log \frac{1}{\Pr\big(Y = 1 \mid X = 0\big)}$$
$$+ (1 - \beta) \log \frac{1}{1 - \Pr\big(Y = 1|X = 0\big)} - H_b(\beta)$$
$$(203)$$

is a convex function of $\Pr\big(Y = 1 \mid X = 0\big)$. Thus, for every $0 < \beta < 1$, the RHS of (203) is maximized for the quantization region that minimizes (or maximizes) $\Pr\big(Y = 1 \mid X = 0\big)$ while holding $\Pr\big(Y = 1 \mid X = \xi\big) = \beta$ fixed. By the Neyman–Pearson Lemma [26], such a quantization region has the form

$$\mathcal{D}^\star = \left\{ \tilde{y} \in \mathbb{C} : \frac{f(\tilde{y}|0)}{f(\tilde{y}|\xi)} \leq \Lambda \right\}, \quad \Lambda > 0 \qquad (204)$$

(or the complement thereof), where $f(\tilde{y}|x)$ denotes the conditional density of $\tilde{Y}$, conditioned on $X = x$, and where Λ is such that $\Pr\big(\tilde{Y} \in \mathcal{D}^\star \mid X = \xi\big) = \beta$. (Note that for every $0 < \beta < 1$, there exists such a Λ since, for the channel model (63), $\Pr\big(\tilde{Y} \in \mathcal{D}^\star \mid X = \xi\big)$ is a continuous, strictly increasing function of $\Lambda > 0$.) The likelihood ratio on the RHS of (204) is given by

$$\frac{f(\tilde{y}|0)}{f(\tilde{y}|\xi)} = \left(1 + \frac{|\xi|^2}{\sigma^2}\right) e^{-\frac{|\tilde{y}|^2}{\sigma^2} \frac{|\xi|^2}{\sigma^2 + |\xi|^2}}, \quad \tilde{y} \in \mathbb{C} \qquad (205)$$

so (204) is a radial quantizer with threshold

$$\Upsilon = \sigma \sqrt{\left(1 + \frac{\sigma^2}{|\xi|^2}\right) \log\left(\frac{1 + \frac{|\xi|^2}{\sigma^2}}{\Lambda}\right)}. \qquad (206)$$

Thus, for every $0 < \beta < 1$, the RHS of (203) is maximized by a radial quantizer whose threshold is a function of β. This implies that, for every nonzero ξ, the relative entropy $D(P_{Y|X=\xi}\|P_{Y|X=0})$ is maximized by a radial quantizer. Such a quantizer induces the channel

$$\Pr\big(Y = 1 \mid X = x\big) = \exp\left(-\frac{\Upsilon^2}{|x|^2 + \sigma^2}\right) \qquad (207)$$

for $x \in \mathbb{C}$ and $\Upsilon > 0$. Consequently,

$$D\big(P_{Y|X=\xi} \,\|\, P_{Y|X=0}\big)$$
$$= e^{-\frac{\Upsilon^2}{|\xi|^2 + \sigma^2}} \log \frac{1}{e^{-\frac{\Upsilon^2}{\sigma^2}}}$$
$$+ \left[1 - e^{-\frac{\Upsilon^2}{|\xi|^2 + \sigma^2}}\right] \log \frac{1}{1 - e^{-\frac{\Upsilon^2}{\sigma^2}}} - H_b\left(e^{-\frac{\Upsilon^2}{|\xi|^2 + \sigma^2}}\right)$$
$$\leq \frac{\Upsilon^2}{\sigma^2} e^{-\frac{\Upsilon^2}{|\xi|^2 + \sigma^2}} - \left[1 - e^{-\frac{\Upsilon^2}{\sigma^2}}\right] \log\left(1 - e^{-\frac{\Upsilon^2}{\sigma^2}}\right)$$
$$\leq \frac{\Upsilon^2}{\sigma^2} e^{-\frac{\Upsilon^2}{|\xi|^2 + \sigma^2}} + \frac{1}{e} \qquad (208)$$

where the second step follows because $H_b(\cdot) \geq 0$ and $\exp\big(-\Upsilon^2/(|\xi|^2 + \sigma^2)\big) \geq \exp\big(-\Upsilon^2/\sigma^2\big)$, and the third step follows because $-x \log x \leq \frac{1}{e}, 0 < x < 1$.

The first term on the RHS of (208) is maximized for $\Upsilon^2 = |\xi|^2 + \sigma^2$, which yields

$$\frac{\Upsilon^2}{\sigma^2} e^{-\frac{\Upsilon^2}{|\xi|^2 + \sigma^2}} \leq \frac{|\xi|^2}{e\,\sigma^2} + \frac{1}{e}, \quad \Upsilon > 0. \qquad (209)$$

The RHS of (208) is thus upper-bounded by

$$D\big(P_{Y|X=\xi} \,\|\, P_{Y|X=0}\big) \leq \frac{|\xi|^2}{e\,\sigma^2} + \frac{2}{e}. \qquad (210)$$

Dividing both sides of (210) by $|\xi|^2$, and computing the limit as $|\xi|$ tends to infinity, yields

$$\varlimsup_{|\xi| \to \infty} \sup_{\mathcal{D}} \frac{D\big(P_{Y|X=\xi} \,\|\, P_{Y|X=0}\big)}{|\xi|^2} \leq \frac{1}{e\,\sigma^2} < \frac{1}{\sigma^2}. \qquad (211)$$

This proves Theorem 5.

XI. SUMMARY AND CONCLUSION

It is well known that quantizing the output of the discrete-time, average-power-limited, Gaussian channel using a symmetric threshold quantizer reduces the capacity per unit energy by a factor of $2/\pi$, a loss which translates to a power loss of approximately 2 dB. We have shown that this loss can be avoided by using asymmetric threshold quantizers with corresponding asymmetric signal constellations. Moreover, the capacity per unit energy can be achieved by a PPM scheme. For this scheme, the error probability can be analyzed directly using the Union Bound and the standard upper bound on the Q-function (47). There is no need to resort to conventional methods used to prove coding theorems such as the method of types, information-spectrum methods, or random coding exponents.

The above results demonstrate that the 2-dB power loss incurred on the Gaussian channel with symmetric 1-bit output quantization is not due to the hard decisions but due to the suboptimal quantizer. In fact, if we employ an asymmetric threshold quantizer, and if we use asymmetric signal constellations, then hard-decision decoding achieves the capacity per unit energy of the Gaussian channel.

The above results also demonstrate that a threshold quantizer is asymptotically optimal as the SNR tends to zero. This is not only true asymptotically: for every fixed SNR, we have shown that, among all 1-bit quantizers, a threshold quantizer is optimal.

We have also shown that the capacity per unit energy can only be achieved by flash-signaling input distributions. Since such signaling leads to poor spectral efficiencies, a significant loss in spectral efficiency is unavoidable. Thus, while 1-bit output quantization does not reduce the capacity per unit energy, it does reduce the spectral efficiency.

For Rayleigh-fading channels, we have shown that, in the coherent case, a 1-bit quantizer does not reduce the capacity per unit energy, provided that we allow the quantizer to depend on the fading level. This is no longer true in the noncoherent case: here all 1-bit output quantizers reduce the capacity per unit energy.

APPENDIX I

Lemma 1: Let $\mathcal{D}$ be a Borel subset of the reals, and let the sequence of real numbers $\{x_k\}$ converge to ξ. Let Z be a zero-mean Gaussian random variable of positive variance σ^2. Then

$$\lim_{k\to\infty} \Pr\big(x_k + Z \in \mathcal{D}\big) = \Pr\big(\xi + Z \in \mathcal{D}\big). \qquad (212)$$

Proof: Let $f(\cdot)$ denote the density of a zero-mean, variance-σ^2 Gaussian random variable, so

$$\Pr\big(x_k + Z \in \mathcal{D}\big) = \int_{\mathcal{D}} f(\tilde{y} - x_k)\mathrm{d}\tilde{y}.$$

Since $f(\cdot)$ is continuous, and since the sequence $\{x_k\}$ converges to ξ, it follows that the sequence of densities $\tilde{y} \mapsto f(\tilde{y} - x_k)$ converges to $\tilde{y} \mapsto f(\tilde{y} - \xi)$. The result follows then by noting that, for every k,

$$\Pr\big(x_k + Z \in \mathbb{R}\big) = \Pr\big(\xi + Z \in \mathbb{R}\big) = 1 \qquad (213)$$

and from Scheffe's Theorem [27, Th. 16.12]. ∎

From Lemma 1, we conclude that $x \mapsto \Pr(Y = 1|X = x)$ is continuous. Since it also bounded, it follows that $\Pr(Y = 1)$ is continuous in the input distribution under the weak topology. Since the binary entropy function is a continuous bounded function, this implies that $H(Y)$ is continuous in the input distribution. By the same lemma, it follows that also the mapping $x \mapsto H_b\big(\Pr(Y = 1|X = x)\big)$ is continuous and bounded, so $H(Y|X)$ is also continuous in the input distribution. We thus have the following lemma.

Lemma 2: For every fixed quantizer $\mathcal{D}$, the functionals $H(Y)$, $H(Y|X)$, and $I(X;Y)$ are continuous in the input distribution under the weak topology.

For proving the existence of a capacity-achieving input distribution, we need a compactness result:

Lemma 3: Let $\mathcal{A} > 0$ be fixed. Every sequence of probability measures on the interval $[-\mathcal{A}, \mathcal{A}]$ of second moment not exceeding $\mathcal{P}$ has a subsequence that converges weakly to a probability distribution on the interval $[-\mathcal{A}, \mathcal{A}]$ of second moment not exceeding $\mathcal{P}$.

Proof: By Prokhorov's Theorem, every sequence of probability measures on $[-\mathcal{A}, \mathcal{A}]$ has a subsequence that converges weakly to some probability measure on $[-\mathcal{A}, \mathcal{A}]$. The second moment of this limiting probability measure cannot exceed $\mathcal{P}$ because the function $x \mapsto x^2$ is a continuous bounded function on the interval $[-\mathcal{A}, \mathcal{A}]$. ∎

Note that Lemma 3 continues to hold for sequences of probability measures on $\mathbb{R}$ of second moment not exceeding $\mathcal{P}$, albeit with a slightly different proof. Thus, the amplitude constraint $\mathcal{A}$ is not essential.

It follows from Lemmas 1–3 that the supremum in (81) defining $C_{\mathcal{D},\mathcal{A}}(\mathcal{P})$ is achieved.

APPENDIX II

We show that, for $\xi < \theta$, the function $\xi \mapsto W\big(\Upsilon_1, \Upsilon_2 \mid \xi\big)$ is strictly increasing. To this end, we note that

$$W\big(\Upsilon_1, \Upsilon_2 \mid \xi\big) = Q\left(\frac{\theta - \Delta - \xi}{\sigma}\right) - Q\left(\frac{\theta + \Delta - \xi}{\sigma}\right) \qquad (214)$$

and take the derivative with respect to ξ. (Recall that $\theta = (\Upsilon_1 + \Upsilon_2)/2$ and $\Delta = (\Upsilon_2 - \Upsilon_1)/2$.) This yields

$$\begin{aligned}
&\frac{\partial}{\partial \xi} W\big(\Upsilon_1, \Upsilon_2 \mid \xi\big) \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\theta - \Delta - \xi)^2}{2\sigma^2}} - \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\theta + \Delta - \xi)^2}{2\sigma^2}} \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\theta - \xi)^2 + \Delta^2}{2\sigma^2}} \left[e^{\Delta \frac{\theta - \xi}{\sigma^2}} - e^{-\Delta \frac{\theta - \xi}{\sigma^2}} \right] \\
&> 0, \qquad \xi < \theta
\end{aligned} \qquad (215)$$

thus proving the claim.

APPENDIX III

To show that

$$\lim_{\xi \to 0} \frac{\Psi(\xi)}{\xi^2} = \frac{1}{2\sigma^2}\left(\frac{1}{2} + \frac{1}{\pi}\right) \qquad (216)$$

we write $\Psi(\xi)$ as

$$\begin{aligned}
\Psi(\xi) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_0^\infty e^{-\frac{(\tilde{y} - \xi)^2}{2\sigma^2}} \left(\frac{\tilde{y}\xi}{\sigma^2} - \frac{\xi^2}{2\sigma^2}\right) \mathrm{d}\tilde{y} \\
&\quad + Q\left(\frac{\xi}{\sigma}\right) \log\left(2Q\left(\frac{\xi}{\sigma}\right)\right) \\
&= \frac{\xi^2}{2\sigma^2} Q\left(-\frac{\xi}{\sigma}\right) + \frac{\xi}{\sqrt{2\pi\sigma^2}}\left(e^{-\frac{\xi^2}{2\sigma^2}} - 1\right) \\
&\quad + \left[Q\left(\frac{\xi}{\sigma}\right) \log\left(2Q\left(\frac{\xi}{\sigma}\right)\right) + \frac{\xi}{\sqrt{2\pi\sigma^2}} \right] \qquad (217)
\end{aligned}$$

and compute the limiting ratio of each term on the RHS of (217) to ξ^2 as ξ tends to zero. For the first two terms, we have

$$\lim_{\xi \to 0} \frac{\frac{\xi^2}{2\sigma^2} Q\left(-\frac{\xi}{\sigma}\right)}{\xi^2} = \frac{1}{4\sigma^2} \qquad (218)$$

and

$$\lim_{\xi \to 0} \frac{\frac{\xi}{\sqrt{2\pi\sigma^2}}\left(e^{-\frac{\xi^2}{2\sigma^2}} - 1\right)}{\xi^2} = 0. \qquad (219)$$

To evaluate the last term on the RHS of (217), we express $\xi \mapsto Q\big(\xi/\sigma\big)$ as a Taylor series around zero

$$Q\left(\frac{\xi}{\sigma}\right) = \frac{1}{2} - \frac{\xi}{\sqrt{2\pi\sigma^2}} + o\big(\xi^2\big). \qquad (220)$$

With this, we obtain

$$
\left[Q\!\left(\frac{\xi}{\sigma}\right) \log\!\left(2Q\!\left(\frac{\xi}{\sigma}\right)\right) + \frac{\xi}{\sqrt{2\pi\sigma^2}} \right]
$$

$$
= \left(\frac{1}{2} - \frac{\xi}{\sqrt{2\pi\sigma^2}} + o(\xi^2) \right) \log\!\left(1 - \frac{\xi}{\sigma}\sqrt{\frac{2}{\pi}} + o(\xi^2) \right)
$$

$$
\quad + \frac{\xi}{\sqrt{2\pi\sigma^2}}
$$

$$
= \left(\frac{1}{2} - \frac{\xi}{\sqrt{2\pi\sigma^2}} + o(\xi^2) \right)\!\left(-\frac{\xi}{\sigma}\sqrt{\frac{2}{\pi}} - \frac{\xi^2}{\sigma^2}\frac{1}{\pi} + o(\xi^2) \right)
$$

$$
\quad + \frac{\xi}{\sqrt{2\pi\sigma^2}}
$$

$$
= \frac{\xi^2}{2\sigma^2}\frac{1}{\pi} + o(\xi^2) \tag{221}
$$

where the second step follows because

$$
\log(1+x) = x - \frac{1}{2}x^2 + o(x^2). \tag{222}
$$

Consequently,

$$
\lim_{\xi\to 0} \frac{Q\!\left(\frac{\xi}{\sigma}\right)\log\!\left(2Q\!\left(\frac{\xi}{\sigma}\right)\right) + \frac{\xi}{\sqrt{2\pi\sigma^2}}}{\xi^2} = \frac{1}{2\sigma^2}\frac{1}{\pi}. \tag{223}
$$

The claim follows by combining (218)–(223) with (217).

APPENDIX IV

A. Proof of (183a)

To prove (183a), namely

$$
\limsup_{\mathcal{A}\downarrow 0}\ \sup_{\Upsilon\geq 0}\ \frac{|\mathsf{R}_H(\mathcal{A},\Upsilon,p_+)|}{\mathcal{A}^2} = 0, \quad 0 \leq p_+ \leq 1
$$

we fix some $\nu \geq 1$ and analyze the cases $0 \leq \Upsilon \leq \nu$ and $\Upsilon > \nu$ separately. Since we are interested in the limit as $\mathcal{A}$ tends to zero, there is no loss in generality in assuming that $\mathcal{A} \leq 1$.

If $0 \leq \Upsilon \leq \nu$, then $\tilde{p}$ in (175) is bounded by

$$
Q\!\left(\frac{\nu+\mathcal{A}}{\sigma}\right) \leq \tilde{p} \leq Q\!\left(-\frac{\mathcal{A}}{\sigma}\right) \tag{224}
$$

which, by the assumption $\mathcal{A} \leq 1$, implies that $\tilde{p}$ is bounded away from 0 and 1:

$$
Q\!\left(\frac{\nu+1}{\sigma}\right) \leq \tilde{p} \leq Q\!\left(-\frac{1}{\sigma}\right). \tag{225}
$$

Consequently, combining (176) with (175) and using the Triangle Inequality yields for $0 \leq \Upsilon \leq \nu$

$$
\left|\mathsf{R}_H(\mathcal{A},\Upsilon,p_+)\right|
$$

$$
\leq \left[\frac{\mathcal{A}}{2\sigma}\frac{|p_+ - p_-|}{\sqrt{2\pi}} e^{-\frac{\Upsilon^2}{2\sigma^2}} + \left|\mathsf{R}_Q(\mathcal{A},\Upsilon,p_+)\right| \right]^3 \frac{|1-2\tilde{p}|}{\tilde{p}^2(1-\tilde{p})^2}
$$

$$
\leq \left[\frac{\mathcal{A}}{2\sigma}\frac{1}{\sqrt{2\pi}} e^{-\frac{\Upsilon^2}{2\sigma^2}} + \left|\mathsf{R}_Q(\mathcal{A},\Upsilon,p_+)\right| \right]^3 \frac{1}{\tilde{p}^2(1-\tilde{p})^2}
$$

$$
\leq \frac{\mathcal{A}^3\left[\frac{1}{2\sigma\sqrt{2\pi}} + \frac{\mathcal{A}}{2\sigma^2\sqrt{2\pi e}} \right]^3}{\left[Q\!\left(\frac{\nu+1}{\sigma}\right)\!\left(1 - Q\!\left(-\frac{1}{\sigma}\right)\right) \right]^2}. \tag{226}
$$

Here, the second step follows by upper-bounding $|1 - 2\tilde{p}| \leq 1$ and $|p_+ - p_-| \leq 1$; and the third step follows from (179) and (225) and by upper-bounding $\exp(-\Upsilon^2/(2\sigma^2)) \leq 1$. Since the RHS of (226) does not depend on Υ, this yields

$$
\limsup_{\mathcal{A}\downarrow 0}\ \sup_{0\leq\Upsilon\leq\nu}\ \frac{|\mathsf{R}_H(\mathcal{A},\Upsilon,p_+)|}{\mathcal{A}^2} = 0, \quad 0 \leq p_+ \leq 1. \tag{227}
$$

For $\Upsilon > \nu$, we first upper-bound (177) as

$$
\left|\mathsf{R}_Q(\mathcal{A},\Upsilon,p_+)\right| \leq \frac{\mathcal{A}^2}{2\sigma^2}\frac{\Upsilon+\mathcal{A}}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\Upsilon-\mathcal{A})^2}{2\sigma^2}}
$$

$$
\leq \frac{\mathcal{A}^2}{\sigma^2}\frac{\Upsilon}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\Upsilon-1)^2}{2\sigma^2}} \tag{228}
$$

where the first step follows by upper-bounding $\tilde{x} \leq \Upsilon + \mathcal{A}$ and $\exp(-\tilde{x}^2/(2\sigma^2)) \leq \exp(-(\Upsilon-\mathcal{A})^2/(2\sigma^2))$, and the second step follows because $\Upsilon > \nu$ and $\mathcal{A} \leq 1$, so $\mathcal{A} < \Upsilon$. Combining (228) with (175) yields for $\Upsilon > \nu$

$$
\left|\mathsf{R}_H(\mathcal{A},\Upsilon,p_+)\right|
$$

$$
\leq \left[\frac{\mathcal{A}}{2\sigma}\frac{|p_+ - p_-|}{\sqrt{2\pi}} e^{-\frac{\Upsilon^2}{2\sigma^2}} + \left|\mathsf{R}_Q(\mathcal{A},\Upsilon,p_+)\right| \right]^3 \frac{|1-2\tilde{p}|}{\tilde{p}^2(1-\tilde{p})^2}
$$

$$
\leq \left[\frac{\mathcal{A}}{2\sigma}\frac{1}{\sqrt{2\pi}} e^{-\frac{\Upsilon^2}{2\sigma^2}} + \frac{\mathcal{A}^2}{\sigma^2}\frac{\Upsilon}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\Upsilon-1)^2}{2\sigma^2}} \right]^3 \frac{1}{\tilde{p}^2(1-\tilde{p})^2}
$$

$$
\leq \frac{\Upsilon^3}{(2\pi\sigma^2)^{\frac{3}{2}}} e^{-\frac{3(\Upsilon-1)^2}{2\sigma^2}} \left[\frac{\mathcal{A}}{2} + \frac{\mathcal{A}^2}{\sigma^2} \right]^3 \frac{1}{\tilde{p}^2(1-\tilde{p})^2} \tag{229}
$$

where the first step follows from the Triangle Inequality; the second step follows from (228) and because $|p_+ - p_-| \leq 1$ and $|1 - 2\tilde{p}| \leq 1$, and the last step follows because $\exp(-\Upsilon^2/(2\sigma^2)) \leq \Upsilon \exp(-(\Upsilon-1)^2/(2\sigma^2))$ for $\Upsilon > 1$.

We next note that, since $\Upsilon > \nu > \mathcal{A}$ and $0 \leq \mathcal{A} \leq 1$, we have

$$
\tilde{p} \leq Q\!\left(\frac{\Upsilon-\mathcal{A}}{\sigma}\right) < \frac{1}{2} \tag{230}
$$

and

$$
\tilde{p} \geq Q\!\left(\frac{\Upsilon+\mathcal{A}}{\sigma}\right)
$$

$$
> \left(1 - \frac{\sigma^2}{(\Upsilon+\mathcal{A})^2} \right) \frac{\sigma}{\sqrt{2\pi}(\Upsilon+\mathcal{A})} e^{-\frac{(\Upsilon+\mathcal{A})^2}{2\sigma^2}}
$$

$$
> \left(1 - \frac{\sigma^2}{\nu^2} \right) \frac{\sigma}{\sqrt{2\pi}(\Upsilon+1)} e^{-\frac{(\Upsilon+1)^2}{2\sigma^2}}, \quad \Upsilon > \nu \tag{231}
$$

where the second step follows from [14, Prop. 19.4.2]. Consequently, using (230) and (231), the RHS of (229) can be upper-bounded by

$$
\left|\mathsf{R}_H(\mathcal{A},\Upsilon,p_+)\right|
$$

$$
\leq \frac{\Upsilon^3}{(2\pi\sigma^2)^{\frac{3}{2}}} e^{-\frac{3(\Upsilon-1)^2}{2\sigma^2}} \frac{4\left[\frac{\mathcal{A}}{2} + \frac{\mathcal{A}^2}{\sigma^2} \right]^3}{\left(1 - \frac{\sigma^2}{\nu^2} \right)^2 \frac{\sigma^2}{2\pi(\Upsilon+1)^2} e^{-\frac{(\Upsilon+1)^2}{\sigma^2}}}
$$

$$
= \frac{4\left[\frac{\mathcal{A}}{2} + \frac{\mathcal{A}^2}{\sigma^2} \right]^3}{\sqrt{2\pi}\sigma^5 \left(1 - \frac{\sigma^2}{\nu^2} \right)^2} \Upsilon^3(\Upsilon+1)^2
$$

$$
\times \exp\!\left(-\frac{3(\Upsilon-1)^2}{2\sigma^2} + \frac{(\Upsilon+1)^2}{\sigma^2} \right), \quad \Upsilon > \nu. \tag{232}
$$

Since the function

$$\Upsilon \mapsto \Upsilon^3 (\Upsilon + 1)^2 \exp\left(-\frac{3(\Upsilon-1)^2}{2\sigma^2} + \frac{(\Upsilon+1)^2}{\sigma^2}\right)$$

is bounded in $\Upsilon > \nu$, this yields

$$\limsup_{\substack{\mathcal{A}\downarrow 0 \\ \Upsilon > \nu}} \frac{|\mathsf{R}_H(\mathcal{A},\Upsilon,p_+)|}{\mathcal{A}^2} = 0, \quad 0 \le p_+ \le 1. \qquad (233)$$

Combining (227) and (223) proves (183a).

B. Proof of (183b)

To prove (183b), namely

$$\limsup_{\substack{\mathcal{A}\downarrow 0 \\ \Upsilon \ge 0}} \frac{|\mathsf{K}(\mathcal{A},\Upsilon,p_+)|}{\mathcal{A}^2} = 0, \quad 0 \le p_+ \le 1$$

we fix some $\nu \ge 1$ and analyze the cases $0 \le \Upsilon \le \nu$ and $\Upsilon > \nu$ separately. Without loss of generality, we assume that $\mathcal{A} \le 1$. If $0 \le \Upsilon \le \nu$, then we have

$$Q\left(\frac{\nu}{\sigma}\right) \le Q\left(\frac{\Upsilon}{\sigma}\right) \le \frac{1}{2} \qquad (234)$$

which yields for every $0 \le p_+ \le 1$ and every $\mathcal{A} \le 1$

$$\begin{aligned}
&|\mathsf{K}(\mathcal{A},\Upsilon,p_+)| \\
&= \frac{\left|\frac{\mathcal{A}}{\sigma}\frac{2(p_+ - p_-)}{\sqrt{2\pi}} e^{-\frac{\Upsilon^2}{2\sigma^2}}\mathsf{R}_Q(\mathcal{A},\Upsilon,p_+) + |\mathsf{R}_Q(\mathcal{A},\Upsilon,p_+)|^2\right|}{2Q\left(\frac{\Upsilon}{\sigma}\right)\left[1 - Q\left(\frac{\Upsilon}{\sigma}\right)\right]} \\
&\le \frac{\frac{\mathcal{A}}{\sigma}\frac{2|p_+ - p_-|}{\sqrt{2\pi}}|\mathsf{R}_Q(\mathcal{A},\Upsilon,p_+)| + |\mathsf{R}_Q(\mathcal{A},\Upsilon,p_+)|^2}{Q\left(\frac{\nu}{\sigma}\right)} \\
&\le \frac{1}{Q\left(\frac{\nu}{\sigma}\right)}\left[\frac{\mathcal{A}^3}{\sigma^3}\frac{1}{2\pi\sqrt{e}} + \frac{\mathcal{A}^4}{4\sigma^4 2\pi e}\right], \quad 0 \le \Upsilon \le \nu. \quad (235)
\end{aligned}$$

Here, the second step follows from (234), from the upper bound $\exp\left(-\Upsilon^2/(2\sigma^2)\right) \le 1$, $\Upsilon \in \mathbb{R}$, and from the Triangle Inequality; and the third step follows from (179) and because $|p_+ - p_-| \le 1$. Consequently,

$$\lim_{\substack{\mathcal{A}\downarrow 0 \\ 0 \le \Upsilon \le \nu}} \sup \frac{|\mathsf{K}(\mathcal{A},\Upsilon,p_+)|}{\mathcal{A}^2} = 0, \quad 0 \le p_+ \le 1. \qquad (236)$$

If $\Upsilon > \nu$, then we have [14, Prop. 19.4.2]

$$\frac{\sigma}{\sqrt{2\pi}\Upsilon}\left(1 - \frac{\sigma^2}{\nu^2}\right)e^{-\frac{\Upsilon^2}{2\sigma^2}} < Q\left(\frac{\Upsilon}{\sigma}\right) < \frac{1}{2} \qquad (237)$$

and, by (228),

$$|\mathsf{R}_Q(\mathcal{A},\Upsilon,p_+)| \le \frac{\mathcal{A}^2}{\sigma^2}\frac{\Upsilon}{\sqrt{2\pi\sigma^2}}e^{-\frac{(\Upsilon-1)^2}{2\sigma^2}}, \quad \Upsilon > \nu. \quad (238)$$

We thus obtain for $\Upsilon > \nu$

$$\begin{aligned}
&|\mathsf{K}(\mathcal{A},\Upsilon,p_+)| \\
&= \frac{\left|\frac{\mathcal{A}}{\sigma}\frac{2(p_+ - p_-)}{\sqrt{2\pi}} e^{-\frac{\Upsilon^2}{2\sigma^2}}\mathsf{R}_Q(\mathcal{A},\Upsilon,p_+) + |\mathsf{R}_Q(\mathcal{A},\Upsilon,p_+)|^2\right|}{2Q\left(\frac{\Upsilon}{\sigma}\right)\left[1 - Q\left(\frac{\Upsilon}{\sigma}\right)\right]} \\
&\le \frac{\sqrt{2\pi}\Upsilon e^{\frac{\Upsilon^2}{2\sigma^2}}}{\sigma\left(1 - \frac{\sigma^2}{\nu^2}\right)}\left[\frac{\mathcal{A}}{\sigma}\frac{2|p_+ - p_-|}{\sqrt{2\pi}}e^{-\frac{\Upsilon^2}{2\sigma^2}}|\mathsf{R}_Q(\mathcal{A},\Upsilon,p_+)| \right. \\
&\qquad\qquad\qquad\qquad \left. + |\mathsf{R}_Q(\mathcal{A},\Upsilon,p_+)|^2\right] \\
&\le \frac{\sqrt{2\pi}\Upsilon e^{\frac{\Upsilon^2}{2\sigma^2}}}{\sigma\left(1 - \frac{\sigma^2}{\nu^2}\right)}\left[\frac{\mathcal{A}^3}{\sigma^3 \pi}\frac{\Upsilon}{\sigma}e^{-\frac{\Upsilon^2}{2\sigma^2} - \frac{(\Upsilon-1)^2}{2\sigma^2}} + \frac{\mathcal{A}^4}{\sigma^4 2\pi}\frac{\Upsilon^2}{\sigma^2}e^{-\frac{(\Upsilon-1)^2}{\sigma^2}}\right] \\
&\le \sqrt{\frac{2}{\pi}}\frac{1}{\left(1 - \frac{\sigma^2}{\nu^2}\right)}\frac{\Upsilon^2}{\sigma^2}e^{\frac{\Upsilon^2}{2\sigma^2} - \frac{(\Upsilon-1)^2}{\sigma^2}}\left[1 + \frac{\mathcal{A}}{2\sigma}\frac{\Upsilon}{\sigma}\right]\frac{\mathcal{A}^3}{\sigma^3} \\
&\le \sqrt{\frac{2}{\pi}}\frac{1}{\left(1 - \frac{\sigma^2}{\nu^2}\right)}\frac{\Upsilon^2}{\sigma^2}e^{\frac{\Upsilon^2}{2\sigma^2} - \frac{(\Upsilon-1)^2}{\sigma^2}}\left[1 + \frac{\Upsilon^2}{2\sigma^2}\right]\frac{\mathcal{A}^3}{\sigma^3} \qquad (239)
\end{aligned}$$

where the second step follows from (237) and from the Triangle Inequality; the third step follows from (238) and because $|p_+ - p_-| \le 1$; the fourth step follows by upper-bounding $\exp\left(-\Upsilon^2/(2\sigma^2)\right) \le \exp\left((\Upsilon-1)^2/(2\sigma^2)\right)$; and the last step follows because $\Upsilon > \nu$ and $\mathcal{A} \le 1$, so $\mathcal{A} \le \Upsilon$.

Since the function

$$\Upsilon \mapsto \frac{\Upsilon^2}{\sigma^2}e^{\frac{\Upsilon^2}{2\sigma^2} - \frac{(\Upsilon-1)^2}{\sigma^2}}\left[1 + \frac{\Upsilon^2}{2\sigma^2}\right]$$

is bounded in $\Upsilon > \nu$, this yields

$$\limsup_{\substack{\mathcal{A}\downarrow 0 \\ \Upsilon > \nu}} \frac{|\mathsf{K}(\mathcal{A},\Upsilon,p_+)|}{\mathcal{A}^2} = 0, \quad 0 \le p_+ \le 1. \qquad (240)$$

Combining (236) and (240) proves (183b).

REFERENCES

[1] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE J. Select. Areas Commun.*, vol. 17, no. 4, pp. 539–550, Apr. 1999.

[2] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. New York, NY, USA: McGraw-Hill, 1979.

[3] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.

[4] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, Oct. 1948.

[5] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY, USA: Wiley, 1968.

[6] T. Koch and A. Lapidoth, "Increased capacity per unit-cost by oversampling," in *Proc. IEEE 26th Conv. Elect. Electron. Eng. Israel*, Eilat, Israel, Nov. 17–20, 2010, pp. 684–688.

[7] T. Koch and A. Lapidoth, "Increased capacity per unit-cost by over-sampling," Sep. 2010 [Online]. Available: http://arxiv.org/abs/1008.5393

[8] E. N. Gilbert, "Increased information rate by oversampling," *IEEE Trans. Inf. Theory*, vol. 39, no. 6, pp. 1973–1976, Nov. 1993.

[9] S. Shamai, "Information rates by oversampling the sign of a bandlimited process," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1230–1236, Jul. 1994.

[10] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 1st ed. New York, NY, USA: Wiley, 1991.

[11] S. Verdú, "On channel capacity per unit cost," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 1019–1030, Sep. 1990.

[12] J. Singh, O. Dabeer, and U. Madhow, "On the limits of communication with low-precision analog-to-digital conversion at the receiver," *IEEE Trans. Commun.*, vol. 57, no. 12, pp. 3629–3639, Dec. 2009.

[13] S. Graf and H. Luschgy, *Foundations of Quantization for Probability Distributions*, ser. Lecture Notes in Mathematics. New York, NY, USA: Springer-Verlag, 2000, vol. 1730.

[14] A. Lapidoth, *A Foundation in Digital Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2009.

[15] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. New York, NY, USA: Wiley, 1965.

[16] P. Zhang, F. M. J. Willems, and L. Huang, "Investigations of noncoherent OOK based schemes with soft and hard decisions for WSNs," in *Proc. 49th Allerton Conf. Commun., Control Comput.*, Allerton House, Monticello, IL, USA, Sep. 28–30, 2011, pp. 1702–1709.

[17] A. Lapidoth and S. Shamai, "Fading channels: how perfect need 'perfect side-information' be?," *IEEE Trans. Inf. Theory*, vol. 48, no. 5, pp. 1118–1134, May 2002.

[18] A. Mezghani and J. A. Nossek, "On ultra-wideband MIMO systems with 1-bit quantized outputs: Performance analysis and input optimization," in *Proc. IEEE Int. Symp. Inf. Theory*, Nice, France, Jun. 24–29, 2007, pp. 1286–1289.

[19] A. Mezghani and J. A. Nossek, "Analysis of Rayleigh-fading channels with 1-bit quantized output," in *Proc. IEEE Int. Symp. Inf. Theory*, Toronto, ON, Canada, Jul. 6–11, 2008, pp. 260–264.

[20] A. Mezghani and J. A. Nossek, "Analysis of 1-bit output noncoherent fading channels in the low SNR regime," in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, Korea, Jun. 28–Jul. 3, 2009, pp. 1080–1084.

[21] S. Krone and G. Fettweis, "Fading channels with 1-bit output quantization: Optimal modulation, ergodic capacity and outage probability," in *Proc. Inf. Theory Workshop*, Dublin, Ireland, Aug. 30–Sep. 3, 2010, pp. 1–5.

[22] T. Koch and A. Lapidoth, "One-bit quantizers for fading channels," in *Proc. IZS*, Zurich, Switzerland, Feb. 29–Mar. 2, 2012, pp. 36–39.

[23] H. S. Witsenhausen, "Some aspects of convexity useful in information theory," *IEEE Trans. Inf. Theory*, vol. IT-26, no. 3, pp. 265–271, May 1980.

[24] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ, USA: Princeton Univ. Press, 1970.

[25] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables: A Handbook for Engineers and Scientists*. Norwell, MA, USA: Kluwer, 2002.

[26] J. Neyman and E. Pearson, "On the problem of the most efficient test of statistical hypotheses," *Phil. Trans. R. Soc. Lond. A*, vol. 231, no. 694–706, pp. 289–337, Jan. 1932.

[27] P. Billingsley, *Probability and Measure*, ser. Probability and Mathematical Statistics, 3rd ed. New York, NY, USA: Wiley, 1995.

Tobias Koch (S'02–M'09) is a Visiting Professor with the Signal Theory and Communications Department of Universidad Carlos III de Madrid (UC3M), Spain. He received the M.Sc. degree in electrical engineering (with distinction) in 2004 and the Ph.D. degree in electrical engineering in 2009, both from ETH Zurich, Switzerland. From June 2010 until May 2012 he was a Marie Curie Intra-European Research Fellow with the University of Cambridge, UK. He was also a research intern at Bell Labs, Murray Hill, NJ in 2004 and at Universitat Pompeu Fabra (UPF), Spain, in 2007. He joined UC3M in 2012. His research interests include digital communication theory and information theory.

Dr. Koch is serving as Vice Chair of the Spain Chapter of the IEEE Information Theory Society in 2013–2014.

Amos Lapidoth (S'89–M'95–SM'00–F'04) received the B.A. degree in mathematics (*summa cum laude*, 1986), the B.Sc. degree in electrical engineering (*summa cum laude*, 1986), and the M.Sc. degree in electrical engineering (1990) all from the Technion—Israel Institute of Technology. He received the Ph.D. degree in electrical engineering from Stanford University in 1995.

In the years 1995–1999 he was an Assistant and Associate Professor at the Department of Electrical Engineering and Computer Science at the Massachusetts Institute of Technology, and was the KDD Career Development Associate Professor in Communications and Technology. He is now Professor of Information Theory at ETH Zurich in Switzerland. He is the author of the book *A Foundation in Digital Communication*, published by Cambridge University Press in 2009. His research interests are in digital communications and information theory.

Dr. Lapidoth served in the years 2003–2004 and 2009 as Associate Editor for Shannon Theory for the IEEE TRANSACTIONS ON INFORMATION THEORY.