

Asymmetric Quantizers Are Better at Low SNR

Tobias Koch
University of Cambridge
Cambridge CB2 1PZ, UK
Email: tobi.koch@eng.cam.ac.uk

Amos Lapidoth
ETH Zurich
8092 Zurich, Switzerland
Email: lapidoth@isi.ee.ethz.ch

Abstract—We study the behavior of channel capacity when a one-bit quantizer is employed at the output of the discrete-time average-power-limited Gaussian channel. We focus on the low signal-to-noise ratio regime, where communication at very low spectral efficiencies takes place, as in Spread-Spectrum and Ultra-Wideband communications. It is well known that, in this regime, a symmetric one-bit quantizer reduces capacity by $2/\pi$, which translates to a power loss of approximately two decibels. Here we show that if an asymmetric one-bit quantizer is employed, and if asymmetric signal constellations are used, then these two decibels can be recovered in full.

I. INTRODUCTION

We study the effect on channel capacity of quantizing the output of the discrete-time average-power-limited Gaussian channel using a one-bit quantizer. We focus on the low signal-to-noise ratio regime, where communication at very low spectral efficiencies takes place (as in Spread-Spectrum and Ultra-Wideband communications). In this regime, a symmetric one-bit quantizer reduces the capacity by a factor of $2/\pi$, corresponding to a 2dB power loss [1]. Hence the rule of thumb that “hard decisions cause a 2dB power loss.” Here we demonstrate that if we allow for *asymmetric one-bit quantizers* with corresponding *asymmetric signal constellations*, these two decibels can be fully recovered. We further demonstrate that the capacity per unit-energy can be achieved by a simple pulse-position modulation (PPM) scheme.

The problem of output quantization is relevant for communication systems where the receiver uses digital signal processing techniques, which require the conversion of the analog received signal to a digital signal by means of an analog-to-digital converter (ADC). For ADCs with high resolution, the effects of quantization are negligible. However, using a high-resolution ADC may not be practical, especially when the bandwidth of the communication system is large and the ADC therefore needs to operate at a high sampling rate [2]. In this case a low-resolution ADC has to be employed. The capacity of the discrete-time Gaussian channel with one-bit output quantization indicates what communication rates can be achieved when the receiver employs a low-resolution ADC.

For a symmetric one-bit quantizer (which produces 1 if the quantizer input is nonnegative and -1 otherwise), the capacity $C_{\text{sym}}(P)$ under the average-power constraint P on the channel

The research leading to these results has received funding from the European Community’s Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 252663.

inputs is given by [1, (3.4.18)], [3, Th. 2]

$$C_{\text{sym}}(P) = \log 2 - H_b\left(Q\left(\sqrt{P/\sigma^2}\right)\right) \quad (1)$$

where $\log(\cdot)$ denotes the natural logarithm function; σ^2 the variance of the additive Gaussian noise;

$$H_b(p) \triangleq -p \log p - (1-p) \log(1-p), \quad 0 \leq p \leq 1$$

(with $0 \log 0 \triangleq 0$) the binary entropy function; and

$$Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-\xi^2/2) d\xi, \quad x \in \mathbb{R}$$

the Q -function. (Here \mathbb{R} denotes the set of real numbers.) This capacity can be achieved by transmitting $\pm\sqrt{P}$ equiprobably.

From (1), the capacity per unit-energy $\dot{C}_{\text{sym}}(0)$ for a symmetric one-bit quantizer can be computed as [1, (3.4.20)]

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

This is a factor of $\frac{2}{\pi}$ smaller than the capacity per-unit energy $\frac{1}{2\sigma^2}$ of the Gaussian channel without output quantization [4]. Thus, quantizing the channel output by a symmetric one-bit quantizer causes a loss of roughly 2dB.

It is tempting to attribute this loss to the fact that the quantizer discards information on the received signal’s magnitude and allows the decoder to perform only hard-decision decoding. However, we demonstrate that the loss of 2dB is not a consequence of the hard-decision decoder but of the suboptimal quantizer. In fact, with an asymmetric quantizer the loss vanishes.

The rest of the paper is organized as follows. Section II describes the considered channel model, defines the capacity per unit-energy, and presents our main result. Section III provides the proof of this result. Section IV shows that the capacity per unit-energy can be achieved by a simple PPM scheme. Section V briefly discusses the capacity of the considered channel. And Section VI concludes the paper with a summary and discussion of our results.

II. CHANNEL MODEL AND CAPACITY PER UNIT-ENERGY

We consider the discrete-time communication system depicted in Figure 1. It is assumed that the message M is uniformly distributed over the set $\{1, 2, \dots, M\}$. The encoder maps M to the length- n sequence X_1, X_2, \dots, X_n , which is then corrupted by additive Gaussian noise and quantized.

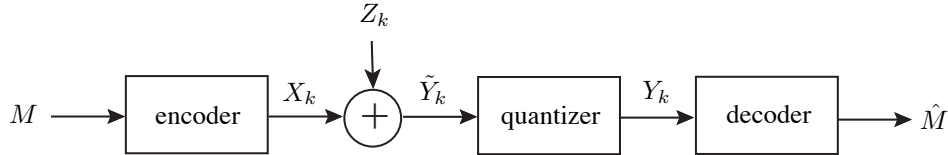


Figure 1. System model.

Thus, at every time instant $k \in \mathbb{Z}$ (where \mathbb{Z} denotes the set of integers), the received signal \tilde{Y}_k corresponding to the channel input $x_k \in \mathbb{R}$ is

$$\tilde{Y}_k = x_k + Z_k, \quad k \in \mathbb{Z} \quad (3)$$

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 . The quantizer produces 1 if \tilde{Y}_k is in the quantization region \mathcal{D} and -1 otherwise, for some Borel set $\mathcal{D} \subset \mathbb{R}$. For example, the quantization region of the symmetric quantizer is given by $\mathcal{D} = \{\tilde{y} \in \mathbb{R}: \tilde{y} \geq 0\}$. The decoder observes the quantizer's outputs Y_1, Y_2, \dots, Y_n and tries to guess which message was transmitted.

We assume that the energy in the transmitted sequence does not exceed E , i.e., that the encoder is such that, for every realization of M , the sequence x_1, x_2, \dots, x_n satisfies

$$\sum_{k=1}^n x_k^2 \leq E. \quad (4)$$

We define the capacity per unit-energy along the lines of [5]. We say that a *rate per unit-energy* $\dot{R}(0)$ (in nats per energy) is *achievable* if for every $\epsilon > 0$ there exists an encoder, a quantizer, and a decoder such that

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

and such that the probability of error $\Pr(\hat{M} \neq M)$ tends to zero as E tends to infinity. The *capacity per unit-energy* is the supremum of all achievable rates per unit-energy.

It follows from [5, Ths. 2 & 3] that the capacity per unit-energy is given by¹

$$\dot{C}(0) = \sup_{P>0} \frac{C(P)}{P} = \sup_{\xi \neq 0, \mathcal{D} \subset \mathbb{R}} \frac{D(P_{Y_1|X_1=\xi} \| P_{Y_1|X_1=0})}{\xi^2} \quad (5)$$

where $D(\cdot \| \cdot)$ denotes the relative entropy, i.e.,

$$D(P \| Q) = \begin{cases} \int \log \left(\frac{dP}{dQ} \right) dP, & \text{if } P \ll Q \\ \infty, & \text{otherwise} \end{cases}$$

(where $P \ll Q$ indicates that P is absolutely continuous with respect to Q); $P_{Y_1|X_1=x}$ denotes the output distribution given that the input is x ; and $C(P)$ is the capacity of the above channel under the average-power constraint P on the channel inputs, i.e., [6]

$$C(P) = \sup I(X_1; Y_1) \quad (6)$$

¹It is straightforward to incorporate the additional maximization over all possible quantizers into the proofs of [5, Ths. 2 & 3].

with the maximization being over all quantization regions \mathcal{D} and over all distributions on X_1 satisfying $E[X_1^2] \leq P$. Note that the first supremum in (5) is approached when P tends to zero. Thus, the capacity per unit-energy is equal to the slope of the capacity-vs-power curve at zero.

By the Data Processing Inequality [7, Th. 2.8.1], the capacity per unit-energy of the above channel is upper-bounded by the capacity per unit-energy of the Gaussian channel without output quantization, i.e., [4]

$$\dot{C}(0) \leq \frac{1}{2\sigma^2}. \quad (7)$$

In the following, we show that there exists a one-bit quantizer that achieves this upper bound.

Theorem 1 (Main Result): The capacity per unit-energy of the discrete-time Gaussian channel with one-bit output quantization is

$$\dot{C}(0) = \frac{1}{2\sigma^2}. \quad (8)$$

Moreover, the capacity per unit-energy can be achieved by a quantization region of the form

$$\mathcal{D} = \{\tilde{y} \in \mathbb{R}: \tilde{y} \geq \Upsilon\}, \quad \text{for some } \Upsilon \in \mathbb{R} \quad (9)$$

where Υ depends on the distribution of the channel input.

Proof: See Section III. ■

We prove Theorem 1 in Section III. A simple PPM scheme that achieves the capacity per unit-energy (8) is presented in Section IV.

III. PROOF OF THEOREM 1

We show that a one-bit quantizer with quantization region \mathcal{D} of the form (9) achieves the rate per unit-energy

$$\dot{R}(0) \geq \frac{1}{2\sigma^2}. \quad (10)$$

Together with (7), this proves Theorem 1.

To this end, we first note that, for the quantization region (9), the conditional probability of the output Y_1 given the input x is

$$P(Y_1 = 1 | X_1 = x) = Q\left(\frac{\Upsilon - x}{\sigma}\right), \quad x \in \mathbb{R} \quad (11)$$

and $P(Y_1 = -1 | X_1 = x) = 1 - P(Y_1 = 1 | X_1 = x)$.

Together with (5), this yields

$$\begin{aligned} \dot{R}(0) &= \sup_{\xi \neq 0, \Upsilon \in \mathbb{R}} \left\{ \frac{Q\left(\frac{\Upsilon-\xi}{\sigma}\right) \log \frac{Q\left(\frac{\Upsilon-\xi}{\sigma}\right)}{Q\left(\frac{\Upsilon}{\sigma}\right)}}{\xi^2} \right. \\ &\quad \left. + \frac{\left[1 - Q\left(\frac{\Upsilon-\xi}{\sigma}\right)\right] \log \frac{1-Q\left(\frac{\Upsilon-\xi}{\sigma}\right)}{1-Q\left(\frac{\Upsilon}{\sigma}\right)}}{\xi^2} \right\} \\ &= \sup_{\xi \neq 0, \Upsilon \in \mathbb{R}} \left\{ \frac{Q\left(\frac{\Upsilon-\xi}{\sigma}\right) \log \frac{1}{Q\left(\frac{\Upsilon}{\sigma}\right)}}{\xi^2} \right. \\ &\quad \left. + \frac{\left[1 - Q\left(\frac{\Upsilon-\xi}{\sigma}\right)\right] \log \frac{1}{1-Q\left(\frac{\Upsilon}{\sigma}\right)}}{\xi^2} \right. \\ &\quad \left. - \frac{H_b\left(Q\left(\frac{\Upsilon-\xi}{\sigma}\right)\right)}{\xi^2} \right\}. \quad (12) \end{aligned}$$

We choose $\Upsilon = \xi - \mu$ for some fixed $\mu \in \mathbb{R}$ and lower-bound the right-hand side (RHS) of (12) by letting ξ tend to infinity. This yields for the last two terms on the RHS of (12)

$$\lim_{\xi \rightarrow \infty} \frac{H_b\left(Q\left(-\frac{\mu}{\sigma}\right)\right)}{\xi^2} = 0 \quad (13)$$

and

$$\lim_{\xi \rightarrow \infty} \frac{\left[1 - Q\left(-\frac{\mu}{\sigma}\right)\right] \log \frac{1}{1-Q\left(\frac{\xi-\mu}{\sigma}\right)}}{\xi^2} = 0. \quad (14)$$

We use the upper bound on the Q -function [8, Prop. 19.4.2]

$$Q(\alpha) < \frac{1}{\sqrt{2\pi}\alpha} e^{-\alpha^2/2}, \quad \alpha > 0 \quad (15)$$

to lower-bound the first term on the RHS of (12) as

$$\begin{aligned} &\lim_{\xi \rightarrow \infty} \frac{Q\left(-\frac{\mu}{\sigma}\right) \log \frac{1}{Q\left(\frac{\xi-\mu}{\sigma}\right)}}{\xi^2} \\ &\geq Q\left(-\frac{\mu}{\sigma}\right) \lim_{\xi \rightarrow \infty} \frac{\frac{1}{2} \log(2\pi) + \log \frac{\xi-\mu}{\sigma} + \frac{(\xi-\mu)^2}{2\sigma^2}}{\xi^2} \\ &= Q\left(-\frac{\mu}{\sigma}\right) \frac{1}{2\sigma^2}. \quad (16) \end{aligned}$$

Combining (13), (14), and (16) with (12) yields

$$\dot{R}(0) \geq Q\left(-\frac{\mu}{\sigma}\right) \frac{1}{2\sigma^2} \quad (17)$$

from which we obtain (10) by letting μ tend to infinity. This proves Theorem 1.

IV. PULSE-POSITION MODULATION

The capacity per unit-energy (8) can be achieved by the following PPM scheme. For each message $m \in \{1, 2, \dots, M\}$, the encoder produces the sequence $x_1(m), x_2(m), \dots, x_M(m)$, where

$$x_k(m) = \begin{cases} \xi, & k = m \\ 0, & k \neq m \end{cases} \quad (18)$$

and where ξ satisfies the energy constraint (4) with equality. Thus, we have $\xi^2 = E$, which for a fixed rate per unit-energy $\dot{R}(0) = \frac{\log M}{E}$ is equal to

$$\xi^2 = \frac{\log M}{\dot{R}(0)}. \quad (19)$$

Note that, while the *rate per unit-energy* is fixed, the *rate* of this scheme is $\frac{\log M}{M}$ and tends to zero as M tends to infinity. Nevertheless, by (5) the capacity per unit-energy is equal to the slope of the capacity-vs-power curve at zero. It thus follows from Theorem 1 that there also exists a transmission scheme of nonzero rate that achieves (8).

We employ a quantizer with quantization region

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

i.e., at every time instant k the quantizer produces 1 if $\tilde{Y}_k \geq \Upsilon$ and -1 otherwise. We choose the threshold Υ such that the probability that the quantizer produces -1 given that the transmitter sends ξ is equal to some arbitrary $0 < \epsilon < 1$, i.e.,

$$\Upsilon = \xi - \sigma Q^{-1}(\epsilon) \quad (20)$$

which yields

$$P(Y_k = -1 \mid X_k = \xi) = Q\left(\frac{\xi - \Upsilon}{\sigma}\right) = \epsilon.$$

Here $Q^{-1}(\cdot)$ denotes the inverse Q -function. Note that this threshold is of the same form as the threshold we chose in Section III to prove Theorem 1.

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

Suppose that message $M = m$ was transmitted. The probability of an error is given by

$$\begin{aligned} &\Pr(\text{error} \mid M = m) \\ &= \Pr\left(\bigcup_{k \neq m} (Y_k = 1) \cup (Y_m = -1) \mid M = m\right) \\ &\leq \sum_{k \neq m} P(Y_k = 1 \mid X_k = 0) + P(Y_m = -1 \mid X_m = \xi) \\ &= \sum_{k \neq m} P(Y_k = 1 \mid X_k = 0) + \epsilon \\ &= (M-1)P(Y_1 = 1 \mid X_1 = 0) + \epsilon \quad (21) \end{aligned}$$

where the second step follows from the union bound; the third step follows from our choice of Υ ; and the fourth step follows because the channel is memoryless which implies that the probability $\Pr(Y_k = 1 \mid X_k = 0)$ does not depend on k . Since the RHS of (21) does not depend on m , it follows that the probability of error

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

is also upper-bounded by (21).

The first term on the RHS of (21) can be evaluated as

$$\begin{aligned} & (M-1)P(Y_1 = 1 \mid X_1 = 0) \\ &= (M-1)Q\left(\frac{\xi - \sigma Q^{-1}(\epsilon)}{\sigma}\right) \\ &= (M-1)Q\left(\frac{\sqrt{\log M} - \sigma Q^{-1}(\epsilon)\sqrt{\dot{R}(0)}}{\sigma\sqrt{\dot{R}(0)}}\right) \end{aligned} \quad (22)$$

where the second step follows from (19). We continue by showing that if

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

then, for every fixed $0 < \epsilon < 1$, the RHS of (22) tends to zero as M tends to infinity. Indeed, we have

$$\begin{aligned} & \lim_{M \rightarrow \infty} (M-1)Q\left(\frac{\sqrt{\log M} - \sigma Q^{-1}(\epsilon)\sqrt{\dot{R}(0)}}{\sigma\sqrt{\dot{R}(0)}}\right) \\ & \leq \lim_{\alpha \rightarrow \infty} \exp\left(\sigma^2 \dot{R}(0) (\alpha + Q^{-1}(\epsilon))^2\right) Q(\alpha) \\ & \leq \lim_{\alpha \rightarrow \infty} \frac{1}{\sqrt{2\pi\alpha}} \exp\left(\sigma^2 \dot{R}(0) (\alpha + Q^{-1}(\epsilon))^2 - \frac{1}{2}\alpha^2\right) \end{aligned} \quad (23)$$

where the first step follows by upper-bounding $M-1 < M$ and by substituting

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

and the second step follows from (15). The limit on the RHS of (23) vanishes for $\dot{R}(0) < \frac{1}{2\sigma^2}$.

Combining (23) with (21), we obtain that if $\dot{R}(0) < \frac{1}{2\sigma^2}$, then the probability of error tends to ϵ as E —and hence also $M = \exp(E\dot{R}(0))$ —tends to infinity. Since ϵ is arbitrary, it follows that the probability of error can be made arbitrarily small by choosing Υ sufficiently large, thus proving that the capacity per unit-energy (8) is achievable by the above PPM scheme.

V. CHANNEL CAPACITY

The definition of channel capacity is analog to that of capacity per unit-energy. A rate $R(P)$ is said to be achievable if for every $\epsilon > 0$ there exists an encoder, a quantizer, and a decoder satisfying

$$\frac{1}{n} \sum_{k=1}^n x_k^2 \leq P, \quad \text{for every realization of } M$$

such that

$$\frac{\log M}{n} > R(P) - \epsilon$$

and such that the probability of error $\Pr(\hat{M} \neq M)$ tends to zero as n tends to infinity. The capacity $C(P)$ is the supremum of all achievable rates. For the above channel, the capacity is given by [6]

$$C(P) = \sup I(X_1; Y_1) \quad (24)$$

where the supremum is over all quantization regions $\mathcal{D} \subset \mathbb{R}$ and over all distributions on X_1 satisfying $E[X_1^2] \leq P$.

If we do not maximize over the quantization region but assume a symmetric quantizer, i.e., $\mathcal{D} = \{\tilde{y} \in \mathbb{R} : \tilde{y} \geq 0\}$, then the capacity is given by [1, (3.4.18)], [3, Th. 2]

$$C_{\text{sym}}(P) = \log 2 - H_b\left(Q\left(\sqrt{P/\sigma^2}\right)\right) \quad (25)$$

where $H_b(\cdot)$ denotes the binary entropy function and $Q(\cdot)$ denotes the Q -function, see Section I. In this case the capacity-achieving input distribution is binary with equiprobable mass points at \sqrt{P} and $-\sqrt{P}$.

To the best of our knowledge, the capacity of the above channel (maximized over the input distribution and the quantization region) as well as the capacity-achieving input distribution and the optimal quantization region are unknown. The following two propositions present results on the latter two problems.

Proposition 2: The capacity-achieving input distribution is discrete with at most three mass points $\xi_\ell \in \mathbb{R}$, $\ell = 1, 2, 3$ satisfying

$$\sum_{\ell=1}^3 p_\ell \xi_\ell = 0 \quad \text{and} \quad \sum_{\ell=1}^3 p_\ell \xi_\ell^2 = P \quad (26)$$

where p_ℓ denotes the probability corresponding to mass point ξ_ℓ , i.e., $0 \leq p_\ell \leq 1$, $\ell = 1, 2, 3$ and $p_1 + p_2 + p_3 = 1$.

Proof: Omitted. ■

This result is consistent with [3, Th. 1], which shows that if the quantization regions of a K -bit quantizer partition the real line into 2^K intervals, then a discrete input distribution with not more than $2^K + 1$ mass points achieves the capacity.

Proposition 3: A threshold quantizer is optimal, i.e., the capacity-achieving quantizer is of the form

$$\mathcal{D} = \{\tilde{y} \in \mathbb{R} : \tilde{y} \geq \Upsilon\}, \quad \text{for some } \Upsilon \in \mathbb{R} \quad (27)$$

where Υ is determined by P and σ^2 .

Proof: Omitted. ■

Propositions 2 and 3 demonstrate that the capacity of the above channel is equal to the capacity of a discrete memoryless channel with input alphabet $\{\xi_1, \xi_2, \xi_3\}$, output alphabet $\{1, -1\}$, and channel law

$$\Pr(Y = 1 \mid X = \xi_\ell) = Q\left(\frac{\Upsilon - \xi_\ell}{\sigma}\right), \quad \ell = 1, 2, 3. \quad (28)$$

It can be further shown that the supremum on the RHS of (24) is achieved, so

$$C(P) = \max_{\substack{(\mathbf{p}, \boldsymbol{\xi}) \in \mathcal{P}(P), \\ \Upsilon \in \mathbb{R}}} I(\mathbf{p}, W(\Upsilon, \boldsymbol{\xi})) \quad (29)$$

where $I(\mathbf{p}, W)$ denotes the mutual information of a channel with channel law W when the channel input is distributed according to \mathbf{p} ; $\mathcal{P}(P)$ denotes the set of probability vectors $\mathbf{p} \in [0, 1]^3$ and mass points $\boldsymbol{\xi} \in \mathbb{R}^3$ satisfying

$$\sum_{\ell=1}^3 p_\ell = 1, \quad \sum_{\ell=1}^3 p_\ell \xi_\ell = 0, \quad \text{and} \quad \sum_{\ell=1}^3 p_\ell \xi_\ell^2 = P; \quad (30)$$

and $W(\Upsilon, \xi)$ denotes the channel law given by (28). Thus, instead of maximizing the mutual information over all Borel sets \mathcal{D} and all probability distributions on X_1 satisfying $E[X_1^2] \leq P$ (24), it suffices to maximize the mutual information over the four real numbers ξ_1, ξ_2, ξ_3 , and Υ and the three-dimensional probability vector \mathbf{p} satisfying (30).

VI. SUMMARY AND CONCLUSION

It is well-known that quantizing the output of the discrete-time average-power-limited Gaussian channel using a symmetric one-bit quantizer reduces the capacity per unit-energy by a factor of $2/\pi$. We have shown that this loss can be fully recovered by allowing for asymmetric one-bit quantizers with corresponding asymmetric signal constellations. We have further shown that the capacity per unit-energy can be achieved by a simple PPM scheme. For this scheme, the error probability can be analyzed directly using the union bound and the upper bound (15) on the Q -function. We thus need not 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 2dB power loss incurred on the Gaussian channel with symmetric one-bit output quantization is not due to the hard-decision decoder, but due to the symmetric quantizer. In fact, if we employ an asymmetric 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 signal-to-noise ratio tends to zero. We have further shown that this is true not only asymptotically: for a fixed signal-to-noise ratio, we have shown that, among all one-bit output quantizers, a threshold quantizer is optimal.

ACKNOWLEDGMENT

The authors would like to thank P. Sotiriadis, who sparked their interest in the problem of quantization.

REFERENCES

- [1] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. McGraw-Hill, 1979.
- [2] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 4, pp. 539–550, April 1999.
- [3] J. Singh, O. Dabeer, and U. Madhow, "On the limits of communication with low-precision analog-to-digital conversion at the receiver," *IEEE Transactions on Communications*, vol. 57, no. 12, pp. 3629–3639, December 2009.
- [4] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423 and 623–656, July and October 1948.
- [5] S. Verdú, "On channel capacity per unit cost," *IEEE Transactions on Information Theory*, vol. 36, pp. 1019–1030, September 1990.
- [6] R. G. Gallager, *Information Theory and Reliable Communication*. John Wiley & Sons, 1968.
- [7] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 1st ed. John Wiley & Sons, 1991.
- [8] A. Lapidoth, *A Foundation in Digital Communication*. Cambridge University Press, 2009.