



**IRWIN AND JOAN JACOBS
CENTER FOR COMMUNICATION AND INFORMATION TECHNOLOGIES**

Trade-offs Between Weak-Noise Estimation Performance and Outage Exponents in Nonlinear Modulation

Neri Merhav

**CCIT Report #920
February 2018**

 Electronics
Computers
Communications

**THE ANDREW & ERNA VITERBI FACULTY OF ELECTRICAL ENGINEERING
TECHNION—ISRAEL INSTITUTE OF TECHNOLOGY, HAIFA 3200003, ISRAEL**



Trade-offs Between Weak–Noise Estimation Performance and Outage Exponents in Nonlinear Modulation

Neri Merhav

The Andrew & Erna Viterbi Faculty of Electrical Engineering
Technion - Israel Institute of Technology
Technion City, Haifa 32000, ISRAEL
E-mail: merhav@ee.technion.ac.il

Abstract

We focus on the problem of modulating a parameter onto a power-limited signal transmitted over a discrete-time Gaussian channel and estimating this parameter at the receiver. Considering the well-known threshold effect in non-linear modulation systems, our approach is the following: instead of deriving upper and lower bounds on the total estimation error, which weigh both weak-noise errors and anomalous errors beyond the threshold, we separate the two kinds of errors. In particular, we derive upper and lower bounds on the best achievable trade-off between the exponential decay rate of the weak-noise expected error cost and the exponential decay rate of the probability of the anomalous error event, also referred to as the *outage event*. This outage event is left to be defined as part of the communication system design problem. Our achievability scheme, which is based on lattice codes, meets the lower bound at the high signal-to-noise (SNR) limit and for a certain range of trade-offs between the weak-noise error cost and the outage exponent.

Index Terms: modulation, parameter estimation, joint source-channel coding, Shannon-Kotelnikov mappings, lattice codes, error exponents.

I. Introduction

Consider the problem of conveying a real-valued parameter u by means of n uses of the additive white Gaussian noise (AWGN) channel,

$$y_i = x_i + z_i, \quad i = 1, 2, \dots, n, \quad (1)$$

where x_i is the i -th component of a channel input vector, $\mathbf{x} = (x_1, x_2, \dots, x_n) = f_n(u)$, that depends on the parameter u , and that is subjected to a power constraint, $\|\mathbf{x}\|^2 \leq nP$, $\{z_i\}$ are

independent, zero-mean, Gaussian random variables with variance σ^2 , and y_i is the i -th component of the channel output vector, $\mathbf{y} = (y_1, y_2, \dots, y_n)$. In a nutshell, our central interest, in this work, is to address the following question: how well can one estimate u based on \mathbf{y} when one is allowed to optimize, not only the estimator, but also the modulator, that is, the function $f_n(\cdot)$ that maps u into a channel input vector? How fast does the estimation error decay as a function of n when the best modulator and estimator are used?

In principle, this problem, which is the discrete-time analogue of the classical problem of “waveform communication” (in the terminology of [10, Chap. 8]), can be viewed both from the information-theoretic and the estimation-theoretic perspectives. In the former, this is an instance of the joint source-channel coding problem (see, e.g., [7] and references therein), where a single source symbol u is transmitted by using the channel n times, and it also falls within the framework of Shannon-Kotel’nikov mappings (see, e.g., [4], [5], [6], [8] and references therein). From the estimation-theoretic point of view, every given modulator induces a certain parametric family of conditional densities of \mathbf{y} given u , and estimation theory provides a plethora of both Bayesian and non-Bayesian lower bounds (depending on whether u is a deterministic parameter or a random variable) on the estimation performance, as well as useful estimators that perform well, e.g., the maximum likelihood (ML) estimator in the non-Bayesian case, the maximum a-posteriori (MAP) estimator in the Bayesian case, and others.

A well known inherent problem in the design of non-linear modulators is the *threshold effect* (see, e.g., [10, Chap. 8]). The threshold effect means a sharp transition between two modes of behavior when the signal-to-noise ratio (SNR) crosses a certain critical value. For high SNR, a.k.a. the *weak-noise* regime, the estimation error behaves similarly as the one attained for a linear modulator, where it roughly achieves the Cramér-Rao lower bound (CRLB). But beyond a certain noise level, the estimation performance completely breaks down rather abruptly. As explained in [10, Chap. 8], for a given non-linear modulator, one can identify a certain anomaly event, or *outage event*, whose probability becomes (quite abruptly) appreciably large as the threshold noise level is crossed.

The literature on the problem of non-linear modulation and estimation contains plenty of results in the form of performance bounds (see, e.g., [1], [2], [3] and many references therein), in

which there is no distinction between weak-noise errors and anomalous errors, i.e., both types of errors weigh in the evaluation of the total mean square error (MSE). A little thought, however, suggests that it makes a lot of sense to separate the two kinds of errors because estimation in the presence of an outage event is rather meaningless. Indeed, separate treatments of the two kinds of errors appear already in [10, pp. 661–674], but not quite in a formal manner. A more systematic approach in this direction, of separating weak-noise errors from anomalous errors, appears in [7, Section IV.A], where the problem was posed in terms of designing a communication system, along with a definition of an outage event (albeit with a different motivation in mind), with the target of minimizing the MSE given that the outage event has not occurred, subject to the constraint that the outage probability would not exceed a given (small) constant. It was shown in [7], that the data processing lower bound is asymptotically achieved, in this setup, by a conceptually simple scheme, that first quantizes the parameter and then maps the set of quantized parameter values into a good digital channel code for the Gaussian channel. The receiver first decodes the digital message and then maps it back to the corresponding quantized parameter value. The outage event is then the error event in the digital part and then the weak-noise MSE is simply the quantization error. The missing link in this result, however, is that the data processing lower bound is not quite compatible to this setting, as it corresponds to a situation where there is no freedom to allow an outage event.

In this paper, we sharpen the approach of [7] in two ways: firstly, both in the lower bound and in the upper bound, we allow an outage event, whose definition is subject to optimization, depending on the communication system itself. Therefore, the lower bound and the upper bound refer to the same setting. Secondly, we somewhat refine the “resolution” in the quantitative aspect of the problem being addressed: instead of constraining the outage probability to be upper bounded by a small constant, we impose the constraint that the outage probability would not exceed a given exponential function of n , that is, $e^{-\lambda n}$ for some prescribed positive constant λ . Under such a constraint, we seek the fastest possible decay rate of the MSE (as a function of λ and the SNR, P/σ^2), or more generally, the expectation of an arbitrary symmetric, convex function of the estimation error. More precisely, we derive an upper bound and a lower bound, which coincide in the limit of high SNR for a certain range of values of λ . Our proposed achievability scheme is based on quantization and modulation, as described before, but since the optimal outage event turns out

to be the complement of a sphere in the space of noise vectors (due to the Gaussianity of the noise), this suggests the use of good *lattice codes*, whose Voronoi cells can indeed be shaped arbitrarily closely to n -dimensional spheres [12, Chap. 7], for large n .

The paper is organized as follows. In Section II, we define notation conventions and formalize the problem being addressed along with the main assumptions. In Section III, we provide the converse bound. In Section IV, we describe the achievability scheme, and finally, in Section V, we summarize and conclude.

II. Notation, Problem Formulation and Assumptions

Consider the following communication system. The transmitter wishes to convey to the receiver, a real-valued parameter u , taking on values in a finite interval, say, without essential loss of generality, the interval $[0, 1]$. To this end, the transmitter uses the channel n times, subject to a given power constraint. The receiver has to estimate u from the n noisy channel outputs. More precisely, given $u \in [0, 1]$, the transmitter sends a vector $\mathbf{x} = f_n(u) \in \mathbb{R}^n$, whose power is limited by $\|\mathbf{x}\|^2 \leq nP$, where P is the maximum allowed average power per channel use. The received vector is

$$\mathbf{Y} = f_n(u) + \mathbf{Z}, \quad (2)$$

where $\mathbf{Z} \in \mathbb{R}^n$ is a zero-mean Gaussian noise vector with covariance matrix $\sigma^2 \cdot I$, I being the $n \times n$ identity matrix. The receiver implements an estimator $\hat{u} = g_n[\mathbf{y}]$ (with \mathbf{y} designating a realization of the random vector \mathbf{Y}) of the parameter u , where $g_n : \mathbb{R}^n \rightarrow [0, 1]$. For every $u \in [0, 1]$, let $\mathcal{O}_n(u) \subset \mathbb{R}^n$ designate an event defined in the space of noise vectors, $\{\mathbf{z}\}$, which is henceforth referred to as the *outage event* (or the *anomalous error event*) given u .

Our general objective is to design a communication system, defined by a transmitter (which is a modulator), $f_n(\cdot)$ that satisfies the aforementioned power constraint, and a receiver (which is an estimator), $g_n[\cdot]$, along with a family of outage events, $\mathcal{O}_n = \{\mathcal{O}_n(u), 0 \leq u \leq 1\}$, so as to minimize

$$\sup_{u \in [0,1]} \mathbf{E} \left\{ \rho(g_n[f_n(u) + \mathbf{Z}] - u) \middle| \mathcal{O}_n^c(u) \right\} \quad (3)$$

subject to the constraint that

$$\sup_{u \in [0,1]} \Pr\{\mathcal{O}_n(u)\} \leq e^{-\lambda n}, \quad (4)$$

where the expectation $\mathbf{E}\{\cdot\}$ in (3) and the probability $\Pr\{\cdot\}$ in (4) are with respect to (w.r.t.) the randomness of the noise vector \mathbf{Z} . Here, $\lambda > 0$ is a prescribed constant (independent of n), henceforth referred to as the *outage exponent*, $\mathcal{O}_n^c(u)$ is the complement of $\mathcal{O}_n(u)$, and $\rho : \mathbb{R} \rightarrow \mathbb{R}^+$ is referred to as the *error cost function* (ECF), which is assumed to have the following properties: (i) symmetry: $\rho(t) = \rho(-t)$, (ii) convexity, (iii) increasing monotonicity for $t \geq 0$ (and hence decreasing monotonicity for $t \leq 0$), and (iv) $\rho(0) = 0$. Let $\mathcal{C}_n(\lambda)$ denote the class of all families of outage events, $\{\mathcal{O}_n\}$, that satisfy (4).

For certain ECFs, like the squared error, $\rho(t) = t^2$, and more generally, for $\rho(t) = |t|^\alpha$, $\alpha \geq 1$, it is well known, from classical communication theory (see, e.g., [10, Chap. 8]), that there exist modulators, receivers and families of outage events for which (3) decays exponentially with n . Since the above defined constrained optimization problem is difficult to solve in a precise closed form, we will adopt the customary information-theoretic approach of characterizing the fastest possible exponential decay rate of (3) subject to (4). More formally, for given sequences of modulators $F = \{f_n(\cdot)\}_{n \geq 1}$ (all satisfying the power constraint), estimators $G = \{g_n[\cdot]\}_{n \geq 1}$, and families of outage events, $\mathcal{O} = \{\mathcal{O}_n\}_{n \geq 1}$ (with $\mathcal{O}_n \in \mathcal{C}_n(\lambda)$), let

$$\mathcal{E}(F, G, \mathcal{O}) = \liminf_{n \rightarrow \infty} \left[-\frac{1}{n} \ln \left(\sup_{u \in [0,1]} \mathbf{E} \left\{ \rho(g_n[f_n(u) + \mathbf{Z}] - u) \middle| \mathcal{O}_n^c(u) \right\} \right) \right]. \quad (5)$$

Our purpose is to derive upper and lower bounds to the best achievable value of $\mathcal{E}(F, G, \mathcal{O})$ as functions of the outage exponent, λ , and the SNR, $\gamma = P/\sigma^2$. In particular, we derive simple formulas for upper and lower bounds to the best achievable high-SNR error-cost exponent $\mathcal{E}(F, G, \mathcal{O})$, denoted $E_U(\lambda, \gamma)$ and $E_L(\lambda, \gamma)$, respectively, which are asymptotically compatible in the sense that $\lim_{\gamma \rightarrow \infty} [E_U(\lambda, \gamma) - E_L(\lambda, \gamma)] = 0$, for a certain range of values of the outage exponent, λ .

To this end, we will make three additional assumptions, the first one concerns the ECF, the second is about the modulator, and third one is regarding the outage events.

A.1 For any given constant $c > 0$, we assume that $\rho(e^{-nc}) \doteq e^{-n\zeta(c)}$, where $\zeta(\cdot)$ is some continuous function with the property that $c > 0$ implies $\zeta(c) > 0$, and where \doteq denotes equivalence in the exponential scale: the notation $a_n \doteq b_n$, for two positive sequences, $\{a_n\}$ and $\{b_n\}$, means that $\frac{1}{n} \log \frac{a_n}{b_n} \rightarrow 0$, as $n \rightarrow \infty$. Moreover, we will assume that $\alpha_n \doteq e^{-nc}$ implies $\rho(\alpha_n) \doteq e^{-n\zeta(c)}$.¹

¹The function $\zeta(\cdot)$ is, of course, induced by the function ρ . For example, if $\rho(t) = |t|^\alpha$, then obviously, $\zeta(c) = \alpha c$.

A.2 Denoting $M_n = e^{nE_U(\lambda, \gamma)}$ (with $E_U(\lambda, \gamma)$ to be defined in Section III), consider the partition of the unit interval into M_n non-overlapping sub-intervals of length $1/M_n$. Then, it will be assumed that for all sufficiently large n , the number, M_n^c , of sub-intervals in which $f_n(\cdot)$ is continuous, is of the exponential order of M_n , i.e., $M_n^c \doteq e^{nE_U(\lambda, \gamma)}$. Also, in each sub-interval $[a, b]$ of continuity, the integral $\int_{f_n(a)}^{f_n(b)} \|df_n(u)\|$ exists and is finite.² Henceforth, we define $\mathcal{F}_n(P)$ as the class of modulators that satisfy these conditions, in addition to the power constraint, $\|f_n(u)\|^2 \leq nP$.

A.3 If, for a certain vector $\mathbf{z} \in \mathcal{O}_n(u)$, one of the components vanishes, then upon replacing this component by any non-zero number, the resulting vector remains in $\mathcal{O}_n(u)$. In addition, it is assumed that as noise variance tends to zero, the covering radius of $\mathcal{O}_n^c(u)$ tends to zero as well.³

The lower bound, $E_L(\lambda, \gamma)$, will be achieved by a conceptually simple achievability scheme, described as follows: we first uniformly quantize the parameter into a grid of $M = e^{nR}$ ($R > 0$) points, and then map this set of points into a good rate- R lattice code of dimension n . The coding rate R will be used to control the trade-off between the outage exponent, λ , and the error-cost exponent, $E_L(\lambda, \gamma)$. The reason for the choice of lattice codes will become apparent from the derivation of the upper bound, $E_U(\lambda, \gamma)$.

A few comments are in order concerning the above problem formulation.

1. The assumption that u takes on values in the unit interval is made largely for the sake of convenience. The extension to any other finite interval, $[a, b]$, will be straightforward. Concerning the extension to the entire real line, the converse part may remain intact, but for the achievability part, the supremum over u in (5) would have to be confined to an interval that grows with n slowly enough, i.e., the interval $[-A_n, +A_n]$ where A_n grows at a sub-exponential rate. Anyway, with this formulation, the entire real line is eventually covered in

As a side remark, note that the function $\rho(\cdot)$ is allowed to depend on n .

²Note that continuity alone does not guarantee that this first order variation integral is finite. A simple counter-example is Brownian motion.

³The first point is reasonable since the new vector designates “stronger noise” than the original one. The second requirement also makes sense, because as the noise becomes weaker, the non-outage events may shrink too (and in all directions on \mathbb{R}^n), in order to improve the weak-noise estimation performance without violating the outage constraint.

the limit $n \rightarrow \infty$.

2. Here, we adopt the minimax approach, a.k.a., the worst-case approach, of minimizing the estimation performance for the worst-case value of u in $[0, 1]$. Owing to the fact that we focus on exponential error bounds, this is asymptotically equivalent to the Bayesian setting as long as the prior of u is bounded away from zero and infinity. This is because given a function $s(u)$ (that does not depend on n), $\int_0^1 p(u)e^{-ns(u)}du$ is dominated by $\sup_{u \in [0,1]} e^{-ns(u)} = \exp\{-n \inf_{u \in [0,1]} s(u)\}$.
3. Since the estimation errors considered are in an exponentially small scale, it is actually only important how the function $\rho(\cdot)$ behaves in the vicinity of the origin. For most of the conceivable convex symmetric cost functions, near the origin, $\rho(t)$ is proportional to $|t|^\alpha$ for some $\alpha \geq 1$. Thus, power cost functions are essentially as general as any convex cost function for our purposes.
4. While this point is well known, we nonetheless feel compelled to emphasize that the index i of each component x_i of the vector $\mathbf{x} = f_n(u)$ (as well as the those of \mathbf{y} and \mathbf{z}) need not necessarily designate discrete time. More generally, \mathbf{x} should be thought of as a vector of coefficients that represent the transmitted signal as a linear combination of n arbitrary orthonormal basis functions (in either discrete time or continuous time). In particular, if these basis functions are taken to be sinusoids (or complex exponentials), then the index i of each component x_i designates frequency. For example, the space of time-limited and (approximately) band-limited signals, of duration T and bandwidth W , are well known to be spanned by $n \approx 2WT$ basis functions for $WT \gg 1$ (see, e.g., [11]). Thus, our setup can easily be translated into the framework of the continuous-time, band-limited AWGN channel, with σ^2 being replaced by the noise spectral density $N_0/2$, the discrete block-length n will be replaced by the time duration T , and all the exponents should be multiplied by the factor $2W$, owing to the substitution of n by $2WT$ in all places (see also [9]).
5. The achievability part in [7, Section IV.A], where the outage probability was kept below a small constant $\delta > 0$, is clearly a special case of our problem formulation with $\lambda = \frac{1}{n} \ln \frac{1}{\delta} \rightarrow 0$. The performance of the achievability scheme in [7], however, was contrasted with the traditional data processing converse bound, that allows no outage at all, namely, $\mathcal{O}(u)$ must

be an empty set for all u . But this is equivalent to the special case of our framework with $\lambda = \infty$. This observation displays a considerable mismatch between the settings of the achievability and the converse parts in [7], when formalized in our framework. As a side remark, we should point out that in [7], the Bayesian approach was adopted, where the parameter u was assumed to be a Gaussian random variable.

III. Converse Bound

Given a positive real λ , let $w(\lambda)$ be the unique solution θ to the equation

$$\theta - \ln(1 + \theta) = 2\lambda. \quad (6)$$

Let

$$E_{\text{v}}(\lambda, \gamma) = \zeta\left(\frac{1}{2} \ln \frac{\gamma}{1 + w(\lambda)}\right), \quad (7)$$

where $\zeta(\cdot)$ is as defined in Assumption A.1, in Section II. Our first main result is the following.

Theorem 1 *Under the assumptions of Section II, for every sequence F of modulators ($f_n \in \mathcal{F}_n(P)$), every sequence G of estimators, and every sequence \mathcal{O} of families of outage events,*

$$\limsup_{\gamma \rightarrow \infty} [\mathcal{E}(F, G, \mathcal{O}) - E_{\text{v}}(\lambda, \gamma)] \leq 0. \quad (8)$$

Before we prove Theorem 1, a brief informal discussion about the plan of the proof is in order. Consider first the relatively simple case where the modulator, $f_n(\cdot)$, is continuous across the entire unit interval, $[0, 1]$ (the idea can then be extended also to the case where there are discontinuities). As u exhausts the unit interval, $\mathbf{x} = f_n(u)$ draws a curve inside the sphere, $\mathcal{S}_n(P) \triangleq \{\mathbf{x} : \|\mathbf{x}\|^2 \leq nP\}$. We refer to this curve as the *signal locus* associated with the modulator f_n . The length of the signal locus is given by

$$L(f_n) = \int_{f_n(0)}^{f_n(1)} \|df_n(u)\|, \quad (9)$$

which is an integral that exists according to Assumption A.2. If, in addition, $f_n(\cdot)$ is differentiable for all $u \in (0, 1)$, we may write

$$L(f_n) = \int_0^1 \|\dot{f}_n(u)\| du, \quad (10)$$

where $\dot{f}_n(u)$ is the vector of derivatives of the components of $f_n(u)$ w.r.t. u . Our proof plan is as follows: we first derive a lower bound on the estimation performance, $\mathbf{E}\{\rho(g_n[f_n(u)+\mathbf{Z}]-u)|\mathcal{O}_n^c(u)\}$, which is a monotonically non-increasing function of $L(f_n)$. Then, we derive an upper bound, L_n^* , on $L(f_n)$, that must apply (at least in the exponential scale) to every $f_n \in \mathcal{F}_n(P)$ in the high SNR regime. This is carried out using certain volume considerations, a.k.a. “tube packing” (in analogy to sphere packing in channel coding), which will also imply that for all u , the best choice of the outage event, $\mathcal{O}_n(u)$, would be the complement of the sphere, centered at the origin, whose radius is exactly large enough to satisfy the outage constraint. Finally, upon substituting $L(f_n)$ by L_n^* in the lower bound to the estimation performance, we arrive at our ultimate lower bound that no longer depends on the specific modulator, f_n . The proof of Theorem 1 is then completed by assessing the exponential rate of the lower bound in the high SNR regime.

We should point that generally speaking, the derivations in [10, Chap. 8] are guided by essentially the same considerations, but there are some important differences. The main difference is that in [10], there is no really a derivation of a lower bound on the estimation performance, but a rather informal argument that concerns the weak-noise MSE of the ML estimator. Now, the weak-noise MSE of the ML estimator is inversely proportional to the energy of the signal derivative, $\|\dot{f}_n(u)\|^2$. Since this L_2 norm is not related to the length, $L(f_n)$ in a unique manner, a minimax consideration is invoked in [10, p. 620] in order to convince the reader that it is best to confine attention to modulators for which $\|\dot{f}_n(u)\|$ is identical to a constant S (independent of u) and then $L(f_n) = S$ and the energy of the derivative is simply S^2 , and so, they now become related in simple manner. There are at least three weaknesses in this kind of argument: (i) it assumes that f_n is differentiable, (ii) it applies to the ML estimator, but it is not a lower bound for an arbitrary estimator, and (iii) it is not quite clear that the confinement to modulators with the property $\|\dot{f}_n(u)\| \equiv S$, does not harm the trade-off with the outage performance. Our derivation, on the other hand, does not suffer from these weaknesses, because, as said, the estimation performance is bounded directly in terms of the signal locus length, not in terms of the energy of the derivative, which may not even exist for a general modulator.

Proof of Theorem 1. Let f_n, g_n and \mathcal{O}_n be given and let Assumptions A.1–A.3 be satisfied. Let M be a positive integer, which will be specified later. For simplicity, we assume first that $f_n(\cdot)$ is continuous along the entire interval (and describe the modifications needed when this is not the

case, in short discussion in the sequel). Then, denoting the density of noise vector, \mathbf{Z} , by $p(\mathbf{z})$, and the indicator function of the event $\mathcal{O}_n^c(u)$, by $\mathcal{I}\{\mathcal{O}_n^c(u)\}$, we have:

$$\begin{aligned}
& \sup_{u \in [0,1]} \mathbf{E} \left\{ \rho(g_n[\mathbf{Y}] - u) \middle| \mathcal{O}_n^c(u) \right\} \\
\geq & \frac{1}{2M} \sum_{i=0}^{M-1} [\mathbf{E}(\rho(g_n[\mathbf{Y}] - u_i) \cdot \mathcal{I}\{\mathcal{O}_n^c(u_i)\}) + \mathbf{E}(\rho(g_n[\mathbf{Y}] - u_{i+1}) \cdot \mathcal{I}\{\mathcal{O}_n^c(u_{i+1})\})] \\
\stackrel{(a)}{\geq} & \frac{1}{2M} \sum_{i=0}^{M-1} \int_{\mathcal{O}_n^c(u_i) \cap \mathcal{O}_n^c(u_{i+1})} [p(\mathbf{y} - f_n(u_i))\rho(g_n[\mathbf{y}] - u_i) + \\
& p(\mathbf{y} - f_n(u_{i+1}))\rho(u - g_n[\mathbf{y}])] d\mathbf{y} \\
\geq & \frac{1}{M} \sum_{i=0}^{M-1} \int_{\mathcal{O}_n^c(u_i) \cap \mathcal{O}_n^c(u_{i+1})} \min\{p(\mathbf{y} - f_n(u_i)), p(\mathbf{y} - f_n(u_{i+1}))\} \times \\
& \left[\frac{1}{2}\rho(g_n[\mathbf{y}] - u_i) + \frac{1}{2}\rho(u_{i+1} - g_n[\mathbf{y}]) \right] d\mathbf{y} \\
\stackrel{(b)}{\geq} & \frac{1}{M} \sum_{i=0}^{M-1} \int_{\mathcal{O}_n^c(u_i) \cap \mathcal{O}_n^c(u_{i+1})} \min\{p(\mathbf{y} - f_n(u_i)), p(\mathbf{y} - f_n(u_{i+1}))\} \times \\
& \rho\left(\frac{g_n[\mathbf{y}] - u_i}{2} + \frac{u_{i+1} - g_n[\mathbf{y}]}{2}\right) d\mathbf{y} \\
= & \frac{1}{M} \sum_{i=0}^{M-1} \rho\left(\frac{u_{i+1} - u_i}{2}\right) \int_{\mathcal{O}_n^c(u_i) \cap \mathcal{O}_n^c(u_{i+1})} \min\{p(\mathbf{y} - f_n(u_i)), p(\mathbf{y} - f_n(u_{i+1}))\} d\mathbf{y} \\
= & \rho\left(\frac{1}{2M}\right) \cdot \frac{1}{M} \sum_{i=0}^{M-1} \int_{\mathcal{O}_n^c(u_i) \cap \mathcal{O}_n^c(u_{i+1})} \min\{p(\mathbf{y} - f_n(u_i)), p(\mathbf{y} - f_n(u_{i+1}))\} d\mathbf{y} \\
= & 2\rho\left(\frac{1}{2M}\right) \cdot \frac{1}{M} \sum_{i=0}^{M-1} \left[\frac{1}{2} \int_{\mathbb{R}^n} \min\{p(\mathbf{y} - f_n(u_i)), p(\mathbf{y} - f_n(u_{i+1}))\} d\mathbf{y} - \right. \\
& \left. - \frac{1}{2} \int_{\mathcal{O}_n(u_i) \cup \mathcal{O}_n(u_{i+1})} \min\{p(\mathbf{y} - f_n(u_i)), p(\mathbf{y} - f_n(u_{i+1}))\} d\mathbf{y} \right] \\
\stackrel{(c)}{\geq} & 2\rho\left(\frac{1}{2M}\right) \cdot \frac{1}{M} \sum_{i=0}^{M-1} \left[Q\left(\frac{\|f_n(u_{i+1}) - f_n(u_i)\|}{2\sigma}\right) - \right. \\
& \left. \frac{1}{2} \int_{\mathcal{O}_n(u_i)} p(\mathbf{y} - f_n(u_i)) d\mathbf{y} - \frac{1}{2} \int_{\mathcal{O}_n(u_{i+1})} p(\mathbf{y} - f_n(u_{i+1})) d\mathbf{y} \right] \\
\stackrel{(d)}{\geq} & 2\rho\left(\frac{1}{2M}\right) \cdot \frac{1}{M} \sum_{i=0}^{M-1} \left[Q\left(\frac{\|f_n(u_{i+1}) - f_n(u_i)\|}{2\sigma}\right) - e^{-\lambda n} \right] \\
= & 2\rho\left(\frac{1}{2M}\right) \cdot \left[\frac{1}{M} \sum_{i=0}^{M-1} Q\left(\frac{\|f_n(u_{i+1}) - f_n(u_i)\|}{2\sigma}\right) - e^{-\lambda n} \right] \\
\stackrel{(e)}{\geq} & 2\rho\left(\frac{1}{2M}\right) \cdot \left[Q\left(\frac{1}{2\sigma M} \sum_{i=0}^{M-1} \|f_n(u_{i+1}) - f_n(u_i)\|\right) - e^{-\lambda n} \right]
\end{aligned}$$

$$\stackrel{(f)}{\geq} 2\rho\left(\frac{1}{2M}\right) \cdot \left(Q\left[\frac{L(f_n)}{2\sigma M}\right] - e^{-\lambda n}\right), \quad (11)$$

where the labeled inequalities are explained as follows: (a) is due to the assumed symmetry of $\rho(\cdot)$; (b) is by its assumed convexity; (c) follows from the union bound and by identifying the first integral of the preceding line as the error probability of the ML decision rule in distinguishing between the two equiprobable hypotheses: $\mathcal{H}_0 : \mathbf{Y} = f_n(u_i) + \mathbf{Z}$ and $\mathcal{H}_1 : \mathbf{Y} = f_n(u_{i+1}) + \mathbf{Z}$, and by using the fact that this error probability is given by $Q(\|f_n(u_{i+1}) - f_n(u_i)\|/2\sigma)$, where $Q(s) = \frac{1}{\sqrt{2\pi}} \int_s^\infty e^{-t^2/2} dt$; (d) is by the outage constraint (4); (e) is due to the convexity of the function $Q(s)$ for $s \geq 0$ (which can easily be verified from its second derivative), and finally, (f) is because $Q(\cdot)$ is monotonically decreasing and because $\sum_i \|f_n(u_{i+1}) - f_n(u_i)\| \leq L(f_n)$, which in turn follows from the fact that the Euclidean norm is a metric (and so, a straight line between any two points is always shorter than any other curve connecting them).

Our next step would be to derive an upper bound on $L(f_n)$. As a preparatory step toward this end, we present the following consideration. For the given $u \in [0, 1]$, let us represent the noise vector \mathbf{z} as (z_0, \mathbf{z}') , where z_0 is the component of \mathbf{z} in the direction of $df_n(u)$ and $\mathbf{z}' = (z_1, \dots, z_{n-1})$ is the orthogonal component of \mathbf{z} . From the outage constraint (4), we have that

$$e^{-\lambda n} \geq \sup_{u \in [0, 1]} \Pr\{\mathcal{O}_n(u)\} \quad (12)$$

$$= \sup_{u \in [0, 1]} \int_{\mathcal{O}_n(u)} p(\mathbf{z}) d\mathbf{z} \quad (13)$$

$$\geq \sup_{u \in [0, 1]} \int_{-\infty}^{+\infty} p(z_0) dz_0 \int_{\mathcal{O}_n(u, z_0)} p(\mathbf{z}') d\mathbf{z}', \quad (14)$$

where we have defined $\mathcal{O}_n(u, z_0) \triangleq \{\mathbf{z}' : (z_0, \mathbf{z}') \in \mathcal{O}_n(u)\}$. By Assumption A.3, $\mathcal{O}_n(u, 0) \subseteq \mathcal{O}_n(u, z_0)$, for every real z_0 , and so, we can continue the above chain of inequalities as follows:

$$e^{-\lambda n} \geq \sup_{x \in [0, 1]} \int_{-\infty}^{+\infty} p(z_0) dz_0 \int_{\mathcal{O}_n(u, 0)} p(\mathbf{z}') d\mathbf{z}' \quad (15)$$

$$\begin{aligned} &= \sup_{u \in [0, 1]} \int_{\mathcal{O}_n(u, 0)} p(\mathbf{z}') d\mathbf{z}' \\ &= \sup_{u \in [0, 1]} \Pr\{\mathcal{O}_n(u, 0)\}. \end{aligned} \quad (16)$$

Thus, for every $u \in [0, 1]$, $\mathcal{O}_n(u, 0)$ must satisfy the same outage constraint as $\mathcal{O}_n(u)$, with exponent λ . Now, due to the Neyman–Pearson theorem and due to the Gaussianity of the noise vector, the

minimum volume of $\mathcal{O}_n^c(u, 0)$, subject to the constraint, $\Pr\{\mathcal{O}_n^c(u, 0)\} \leq e^{-\lambda n}$, is attained when $\mathcal{O}_n^c(u, 0)$ is an $(n - 1)$ -dimensional Euclidean sphere centered at the origin.⁴ Using the Chernoff bound (which is well-known to be exponentially tight) and the given outage constraint, the radius of this sphere is easily found to be $\sigma\sqrt{n[1 + w(\lambda)]}$.

We next invoke a “tube-packing” argument in the spirit of [10, pp. 672–673] (see also [4, Subsection 2.3.1], [8]). To comply with the outage constraint, the signal locus curve must have the property that $\mathcal{O}_n^c(u, 0)$ and $\mathcal{O}_n^c(u', 0)$ are disjoint whenever the u and u' are far apart in the sense of being points that belong to different folds of the signal locus curve. This is because no point in space can possibly belong simultaneously to the non-outage regions of two remote values of the parameter. In mathematical terms, this means that in the high SNR limit,⁵ the volume of the body $\mathcal{T} \triangleq \{\mathbf{y} = f_n(u) + (0, \mathbf{z}') : u \in [0, 1], \mathbf{z}' \in \mathcal{O}_n^c(u, 0)\}$ is of the exponential order of $\int_{f_n(0)}^{f_n(1)} \|df_n(u)\| \cdot \text{Vol}\{\mathcal{O}_n^c(u, 0)\}$. We next derive an upper bound and a lower bound to $\text{Vol}\{\mathcal{T}\}$, so that by comparing these two bounds, we can find an upper bound to $L(f_n)$.

As for the upper bound to $\text{Vol}\{\mathcal{T}\}$, due to the power constraint, it cannot exceed the volume of a sphere whose radius is upper bounded by $\sqrt{nP} + r_{\text{cov}}\{\mathcal{O}_n^c(u)\}$, where $r_{\text{cov}}\{\mathcal{O}_n^c(u)\}$ is the covering radius of $\mathcal{O}_n^c(u)$, which in the high SNR limit, is negligible compared to \sqrt{nP} (due to Assumption A.3), that is, $\sqrt{n[P + o(\gamma)]}$, where $o(\gamma)$ is a term that tends to zero as $\gamma \rightarrow \infty$. Thus,

$$\text{Vol}\{\mathcal{T}\} \leq (2\pi e[P + o(\gamma)])^{n/2}, \quad (17)$$

where we have used the fact that the volume of an n -dimensional sphere of radius \sqrt{nS} is of the exponential order of $(2\pi eS)^{n/2}$ (see, e.g., [12, eq. (7.30)]).

For the lower bound to $\text{Vol}\{\mathcal{T}\}$, we have by the aforementioned Neyman–Pearson/iso-perimetric consideration and by the outage constraint,

$$\text{Vol}\{\mathcal{T}\} \geq L(f_n) \cdot \min_u \text{Vol}\{\mathcal{O}_n^c(u, 0)\} \geq L(f_n) \cdot (2\pi e\sigma^2[1 + w(\lambda)])^{(n-1)/2}. \quad (18)$$

Before we proceed, an important comment is now in order: the r.h.s. is of the same exponential order as the volume of the union of n -dimensional spheres of radius $\sigma\sqrt{n[1 + w(\lambda)]}$ (as opposed to the $(n - 1)$ -dimensional spheres described above), centered at $f_n(u)$, where the union runs from

⁴This argument follows also from one of the iso-perimetric inequalities, see, e.g., [12, Theorem 7.1.1, eq. (7.12d)].

⁵Here the high SNR limit means that the volumes of $\mathcal{O}_n^c(u, 0)$ are relatively very small: in particular, they shrink when σ is small, and σ is assumed very small compared to the radius of curvature of the signal locus curve.

$u = 0$ to $u = 1$. This means that in the high SNR limit, the optimal choice of $\mathcal{O}_n(u)$ is the complement of a sphere of radius $\sigma\sqrt{n[1+w(\lambda)]}$, independently of u . This observation will be important for the achievability part (Section IV).

Comparing the upper and the lower bounds to $\text{Vol}\{\mathcal{T}\}$, we obtain

$$\begin{aligned} L(f_n) &\stackrel{\cdot}{\leq} \frac{(2\pi e[P+o(\gamma)])^{n/2}}{(2\pi e\sigma^2[1+w(\lambda)])^{(n-1)/2}} \\ &\doteq \sigma \cdot \exp\left\{\frac{1}{2} \ln \frac{P+o(\gamma)}{\sigma^2[1+w(\lambda)]}\right\} \\ &= \sigma e^{n[E_{\text{V}}(\lambda,\gamma)+o(\gamma)]} \triangleq L_n^*. \end{aligned} \tag{19}$$

Returning to (11), and choosing $M = M_n = L_n^*/(2\sigma s)$ for some arbitrary constant⁶ $s > 0$, we have

$$\begin{aligned} \sup_{u \in [0,1]} \mathbf{E} \left\{ \rho(g_n[\mathbf{Y}] - u) \middle| \mathcal{O}_n^c(u) \right\} &\stackrel{\cdot}{\geq} \rho\left(\frac{s}{L_n^*}\right) \cdot [Q(s) - e^{-\lambda n}] \\ &\doteq \exp\{-n(\zeta[E_{\text{V}}(\lambda,\gamma)] + o(\gamma))\}, \end{aligned} \tag{20}$$

completing the proof of Theorem 1 for the case where $f_n(\cdot)$ is continuous.

Relaxing now the continuity assumption, and referring to Assumption A.2, we may allow discontinuities of $f_n(\cdot)$, even at majority of the sub-intervals $[0, 1/M_n), [1/M_n, 2/M_n), \dots, [(M_n - 1)/M_n, 1]$, provided that the number M_n^c of intervals of continuity is still of the exponential order of M_n defined above. In such a case, the modification of the above derivation would have the following ingredients: (i) the summation over i in the first steps of (11) should be further lower bounded by excluding terms, indexed by i , for which there are discontinuities of $f_n(u)$ in the interval $i/M_n \leq u < (i+1)/M_n$; (ii) the application of the Jensen inequality to the function Q (step (e) in (11)) would be limited to the remaining terms; (iii) consequently, the lower bound above would be multiplied by a factor of M_n^c/M_n^* (which will not affect the exponent, thanks to Assumption A.2), and (iv) $L(f_n)$, in the argument of the Q -function, would mean the sum of lengths of the continuous parts of the signal locus curve. The tube packing argument will continue to apply to the total volume of the disjoint union all tubes formed by the separate continuous parts of the signal locus curve. This completes the proof of Theorem 1.

⁶In principle, the lower bound can be maximized over s , but such a maximization will not affect the exponential order of the bound, at least not in the important case of $\rho(t) = |t|^\alpha$.

IV. Achievability

Upon careful inspection of the proof of Theorem 1, one may find some suggestive indications what would be a good achievability scheme.

First, let us look at the last line of eq. (11): it is natural to think of the term $\rho(1/2M)$ as one that can be achieved by uniform quantization: if we quantize the unit interval uniformly and create a grid of M points with equal spacings of $1/M$ in between, then upon quantizing an arbitrary $u \in [0, 1]$ into the nearest grid point, the absolute value of the quantization error cannot exceed $1/2M$, and so, the error cost will never be above $\rho(1/2M)$. Thus, if we can come up with an achievability scheme for which the weak-noise error is exactly this quantization error, we can achieve the weak-noise error bound up to a constant factor (given by the other multiplicative terms in the last line of (11)). But once we have quantized u into one out of M quantization points, it only remains to map each one of these points into a corresponding channel input vector, namely, to apply channel coding. The error event of this channel code will then have to be the outage event. But according to the comment that follows eq. (18), the best choice of the outage event (at least in the high SNR limit) is the complement of a sphere, independently of the transmitted code vector. It is well known that for good lattice codes in high dimension, the Voronoi cells become closer and closer to spheres [12, Chap. 7], and so, the error event (which is independent of the transmitted code vector) is roughly the event that the noise vector \mathbf{Z} would fall outside a sphere. The outage exponent λ must then be the error exponent of such a lattice code. The coding rate $R = \frac{1}{n} \ln M$ would then be a free parameter that controls the trade-off between the weak-noise (quantization) error and the outage (decoding error) exponent.

More precisely, let $M = e^{nR}$ be a positive integer, where $R \geq 0$ is a design parameter to be selected later. Consider the uniform quantization of the the unit interval into M non-overlapping bins, each of size $\Delta = \frac{1}{M} = e^{-nR}$, and let $q(u)$ be the midpoint of the bin $i(u)$ to which u belongs, i.e.,

$$i(u) = \lfloor u/\Delta \rfloor \tag{21}$$

$$q(u) = \Delta \left[i(u) + \frac{1}{2} \right]. \tag{22}$$

Let us define the modulator according to

$$f_n(u) = \mathbf{x}[i(u)], \quad (23)$$

where $\{\mathbf{x}[0], \mathbf{x}[1], \dots, \mathbf{x}[M-1]\}$ is (part of) a channel lattice code whose members all lie within the sphere of radius \sqrt{nP} around the origin, so as to comply with the power constraint. At the receiver side, we first decode $i(u)$ from the channel output, and then u is estimated according to

$$g_n[\mathbf{y}] = \Delta \left[\widehat{i(u)} + \frac{1}{2} \right], \quad (24)$$

where $\widehat{i(u)}$ is the decoded version of $i(u)$ (which is, of course, a function of \mathbf{y} only). Thus, if $i(u)$ is decoded correctly, the (weak-noise) estimation error is simply the quantization error, whose cost is of the exponential order of

$$\mathbf{E} \left\{ \rho(g_n[\mathbf{Y}] - u) \middle| \mathcal{O}_n^c(u) \right\} \leq \rho(e^{-nR}) \doteq e^{-n\zeta(R)}. \quad (25)$$

Referring to converse bound in Theorem 1 (see eq. (7)), it is suggestive that, in the quest for asymptotically optimal performance, the coding rate R of the lattice code would be set to

$$R = R_\lambda \triangleq \frac{1}{2} \ln \frac{\gamma}{1 + w(\lambda)} = \frac{1}{2} \ln \gamma - \frac{1}{2} \ln[1 + w(\lambda)]. \quad (26)$$

The first term on the right-most side is obviously identified as the high-SNR approximation to the channel capacity, $C = \frac{1}{2} \ln(1 + \gamma)$, and the second term is interpreted as the inevitable gap to capacity that must be suffered in order to comply with the outage constraint.

It now remains to examine the achievable error exponents of lattice codes at rate R_λ and to compare to λ . According to Zamir [12, Theorem 13.4.1], the random coding exponent of lattice codes (w.r.t. the MHS ensemble) is given by $E_r(\mu/2\pi e)$, where

$$E_r(x) \triangleq \begin{cases} 0 & x \leq 1 \\ \frac{1}{2}(x - \ln x - 1) & 1 \leq x \leq 2 \\ \frac{1}{2} \ln \frac{ex}{4} & x \geq 2 \end{cases} \quad (27)$$

and μ is the normalized volume-to-noise ratio (NVNR), a constant that depends on the lattice Λ and the noise variance σ^2 , defined as

$$\mu(\Lambda, \sigma^2(P_e)) = \frac{[V(\Lambda)]^{2/n}}{\sigma^2(P_e)}, \quad (28)$$

where $V(\Lambda)$ is the volume of the Voronoi cell of the lattice, and $\sigma^2(P_e)$ is the value of the noise variance such that $\Pr\{\mathbf{Z}$ falls outside the Voronoi cell} is equal to a prescribed value of $P_e \in (0, 1)$. For such a fixed P_e , we have $\mu(\Lambda, \sigma^2(P_e)) > 2\pi e$, but the lower bound of $2\pi e$ can be approached arbitrarily closely by some sequence of lattices [12, Theorem 7.7.1]. Now, if one enlarges those lattice cells by a factor of $\sqrt{1+w(\lambda)}$ in each dimension (by means of reducing the rate $R \approx \frac{1}{2} \ln[(2\pi e P)/V^{2/n}(\Lambda)]$ from C to R_λ), keeping the noise variance intact, this would increase μ by a factor of $1+w(\lambda)$, and then the error exponent would be $E_r[1+w(\lambda)]$, which for $w(\lambda) \in [0, 1]$, gives

$$E_r[1+w(\lambda)] = \frac{1}{2} \{1+w(\lambda) - \ln[1+w(\lambda)] - 1\} = \frac{1}{2} \cdot 2\lambda = \lambda, \quad (29)$$

where the second equality is by the definition of the function $w(\lambda)$ (see eq. (6)). Thus, the outage constraint is met in the range where $w(\lambda) \in [0, 1]$, and the converse bound is asymptotically achieved.

In [12, Theorem 13.7.1], an expurgated error exponent for lattice codes is presented, which is given by $E_x(\mu/2\pi e)$, where

$$E_x(x) = \begin{cases} 0 & x \leq \frac{4}{e} \\ \frac{1}{2} \ln \frac{ex}{4} & \frac{4}{e} \leq x \leq 4 \\ \frac{x}{8} & x \geq 4 \end{cases} \quad (30)$$

The best achievable exponent is then obtained by combining these two error exponents, which results in the *Polytyrev error exponent*, $E_P(x)$ [12, Theorem 13.7.1 and subsequent discussion].

For a general λ , let us define $E_P^{-1}(\cdot)$ as the inverse function of $E_P(\cdot)$. In order to comply with the outage constraint, one must have $\mu = 2\pi e E_P^{-1}(\lambda)$, which means a coding rate whose high-SNR approximation is given by

$$\begin{aligned} R &\approx \frac{1}{2} \ln \frac{2\pi e P}{V^{2/n}(\Lambda)} \\ &= \frac{1}{2} \ln \frac{2\pi e P}{\mu \sigma^2} \\ &= \frac{1}{2} \ln \frac{2\pi e P}{2\pi e \sigma^2 E_P^{-1}(\lambda)} \\ &= \frac{1}{2} \ln \gamma - \frac{1}{2} \ln E_P^{-1}(\lambda). \end{aligned} \quad (31)$$

Thus, for a general value of λ , the weak-noise error cost exponent is of the proposed communication system is given by

$$E_L(\lambda, \gamma) = \zeta \left[\frac{1}{2} \ln \gamma - \frac{1}{2} \ln E_P^{-1}(\lambda) \right]. \quad (32)$$

Strictly speaking, the above described quantization–and–coding scheme does not comply with Assumption A.2 since $f_n(u)$ contains discontinuities at *all* quantization boundaries. However, a small formal modification makes it compliant with Assumption A.2: let every two consecutive points, of the form (u_{2i}, u_{2i+1}) , be mapped to the same codeword, so that along the interval between them, the signal locus is a fixed point in \mathbb{R}^n (which is that codeword), and so, it is trivially continuous. This way, M_n^c is half of the total number of intervals, and hence of the same exponential order, complying with Assumption A.2. At the decoder, a random selection would be made between the two parameter values that pertain to the resulting decoded codeword. In this case, the quantization error (or the weak–noise error) would not exceed $3/2M$, whose cost is $\rho(3/2M)$, which is still of the same exponential order.

We have just proved the following theorem.

Theorem 2 *There exists a sequence F of modulators ($f_n \in \mathcal{F}_n(P)$), a sequence G of estimators, and a sequence \mathcal{O} of families of outage events, such that*

$$\liminf_{\gamma \rightarrow \infty} [\mathcal{E}(F, G, \mathcal{O}) - E_L(\lambda, \gamma)] \geq 0, \quad (33)$$

where $E_L(\lambda, \gamma)$ is defined as in eq. (32). Moreover, $E_L(\lambda, \gamma) = E_U(\lambda, \gamma)$ in the range of λ where $w(\lambda) \in [0, 1]$.

Finally, we point out that there are improved coding schemes, such as the one in [12, Theorem 13.8.3], where the Voronoi modulation is based on lattice shaping with dither and minimum mean square error estimation prior to nearest–neighbor decoding. However, in the high–SNR regime, considered in this work, their performance is essentially the same (see, in particular, eq. (13.68) in [12]).

V. Summary and Conclusion

In this paper, we addressed the problem of modulation and estimation under outage constraints. For the high–SNR regime, we have derived upper and lower bounds on the exponential order of the weak–noise error cost for a given outage exponent. The bounds coincide for a certain range of outage exponents. The proposed achievability scheme naturally employs lattice codes due to the spherical shape of the optimal non–outage events in the space of noise vectors.

Beyond the obvious quest for closing, or at least shrinking, the gaps for every λ and γ , some interesting future directions may include extensions in ways: (i) the case where the parameter is a vector rather than a scalar, (ii) the use of feedback, (iii) multi-user situations like the multiple access channel and the interference channel, (iv) non-Gaussian memoryless channels, and (v) Gaussian channels with correlated noise components (colored noise).

Acknowledgment

Interesting discussions with Ertem Tuncel are greatly appreciated.

References

- [1] M. V. Burnashev, "A new lower bound for the a-mean error of parameter transmission over the white Gaussian channel, *IEEE Trans. Inform. Theory*, vol. 30, no. 1, pp. 23–34, January 1984.
- [2] M. V. Burnashev, "On minimum attainable mean-square error in transmission of a parameter over a channel with white Gaussian noise, *Problemy Peredachi Informatsii*, vol. 21, no. 4, pp. 3–16, 1985.
- [3] D. L. Cohn, *Minimum Mean Square Error Without Coding*, Ph.D. dissertation, Massachusetts Institute of Technology, July 1970.
- [4] P. A. Floor, *On the Theory of Shannon-Kotelnikov Mappings in Joint Source Channel Coding*, Ph.D. dissertation, Faculty of Information Technology, Mathematics and Electrical Engineering, Department of Electronics and Telecommunications, Norwegian University of Science and Technology, May 2008.
- [5] F. Hekland, *On the design and analysis of Shannon-Kotelnikov Mappings for Joint Source Channel Coding*, Ph.D. dissertation, Faculty of Information Technology, Mathematics and Electrical Engineering, Department of Electronics and Telecommunications, Norwegian University of Science and Technology, May 2007.

- [6] F. Hekland, P. A. Floor, and T. A. Ramstad, “Shannon-Kotelnikov mappings in joint source-channel coding, *IEEE Trans. Comm.*, vol. 57, no. 1, pp. 94–105, January 2009.
- [7] E. Köken, D. Günduz and E. Tuncel, “Energy–distortion exponents in lossy transmission of Gaussian sources over Gaussian channels, *IEEE Trans. on Inform. Theory*, vol. 63, no. 2, pp. 1227–1236, February 2017.
- [8] B. Kravtsov and D. Raphaeli, “Multidimensional analog modulation and optimal mappings, *Proc. IEEE Globecom 2007*, pp. 1426–1430, 2007.
- [9] N. Weinberger and N. Merhav, “Lower bounds on parameter modulation–estimation under bandwidth constraints,” *IEEE Trans. Inform. Theory*, vol. 63, no. 6, pp. 3854–3874, June 2017.
- [10] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*, John Wiley & Sons Inc., 1965.
- [11] A. D. Wyner, “A bound on the number of distinguishable functions which are time-limited and approximately band-limited”, *SIAM Journal on Applied Mathematics*, vol. 24, no. 3, pp. 289–297, May 1973.
- [12] R. Zamir, *Lattice Coding for Signals and Networks*, Cambridge University Press, United Kingdom, 2014.