A Ziv-Zakai-Rényi Lower Bound on Distortion at High Resolution

Itai Leibowitz
EE - Systems Department
Tel Aviv University
Tel Aviv, Israel
Email: leibo@eng.tau.ac.il

Ram Zamir
EE - Systems Department.
Tel Aviv University
Tel Aviv, Israel
Email: zamir@eng.tau.ac.il

***Abstract* – We follow a method introduced by Ziv and Zakai for finding 'informational' lower bounds on delay constrained joint source-channel coding. Their method uses the data processing theorem for generalized measures of information. We introduce the use of Rényi's information of order α in their framework, and use high-resolution approximations to find its rate distortion function for a source that possesses a smooth distribution with *r*th-power distortion. This allows us to present two new lower bounds, one on the distortion in fixed rate vector quantization, and the other on the transmission through low-dimensional modulo-lattice additive noise channels.**

I. Introduction

We consider a problem of length-constrained communication, where we wish to reproduce an approximation to a real-valued source at the far end of a channel, as appears in Fig. 1. We

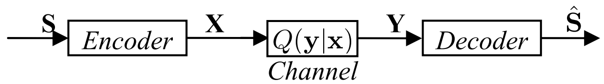


Fig. 1 - A length-constrained communication scheme

represent the constraint on the length of the code by using vector notation for the variables, where the source $\mathbf{S}$ and its reproduction $\hat{\mathbf{S}}$ are k-vectors, and the communication channel has input $\mathbf{X}$ and output $\mathbf{Y}$, which are both n-vectors. A distortion measure d measures the quality of the reproduction, where the overall distortion is measured by

$$D = \frac{1}{k} E\{d(\mathbf{S}, \hat{\mathbf{S}})\}. \tag{1}$$

The channel is described by its conditional distribution function $Q(\mathbf{y}|\mathbf{x})$.

In Shannon's original paper from 1948 [11], he presented a bound on the performance of such transmission, known also as the separation principle. This bound is the familiar

$$R(D) \le C, \tag{2}$$

where $R(D)$ is the rate distortion function (RDF) of the source under a specific distortion measure, and C is the capacity of the channel. The rate distortion function is defined by

$$R(D) = \inf I(\mathbf{S}; \hat{\mathbf{S}}), \tag{3}$$

where the infimum is taken over all conditional distributions $Q(\hat{\mathbf{s}}|\mathbf{s})$ that satisfy the distortion constraint (1). The channel capacity is defined by

$$C = \sup I(\mathbf{X}; \mathbf{Y}), \tag{4}$$

where the supremum is taken over all input distributions $f(\mathbf{x})$, or over a subset that satisfies a certain constraint (e.g. average or peak power constraint). Shannon's bound is then proved using the data processing theorem, which states that

$$\begin{array}{l} I(\mathbf{S}; \hat{\mathbf{S}}) \le I(\mathbf{S}; \mathbf{Y}) \le I(\mathbf{X}; \mathbf{Y}) \\ I(\mathbf{S}; \hat{\mathbf{S}}) \le I(\mathbf{X}; \hat{\mathbf{S}}) \le I(\mathbf{X}; \mathbf{Y}) \end{array}. \tag{5}$$

Shannon showed that in the limit of large block length, i.e. $k,n\to\infty$, A combination of the two separate solutions to the problems of lossy source coding and communication over the given channel can achieve the bound (2) asymptotically.

However, Shannon's result generally relies on the law of large numbers, and may break down whenever the length of the code is restricted, due to constraints on delay or complexity. In such cases, joint source/channel schemes may outperform the separate solution, and neither method is guaranteed to reach the bound (2). In fact, except for a few special cases, it is impossible to achieve Shannon's bound with a code of fixed length. In many cases of length- constrained communication, the optimal achievable performance is not known, and outer bounds that are stricter than Shannon's can serve as a goal and benchmark for the performance of communication schemes.

In their paper from 1973 [12] (also later generalized in [13]), Ziv and Zakai proposed a method for the calculation of bounds similar to (2) for communication schemes with finite block length. They showed that if we replace the $-\log(\cdot)$ function in the expression for the mutual information with a convex function $\Phi(\cdot)$ that satisfies a certain technical requirement, the functional $I_\Phi(\cdot)$ that results still obeys the data processing theorem. By defining $R_\Phi(D)$ as the infimum of $I_\Phi(\mathbf{S}; \hat{\mathbf{S}})$, and C_Φ as the supremum of $I_\Phi(\mathbf{X}; \mathbf{Y})$, we form a new inequality that states $R_\Phi(D) \le C_\Phi$.

Shannon's information is additive in the block length, and thus (2) is independent of n and k when dealing with memory-less sources and channels. The new information measures do not possess this quality, and thus the distortion bounds derived from them depend on the specific block length. This results in tighter bounds for coding with constrained length. We give a review of Ziv and Zakai's work in Section II.

In Section III, we present a variation on Rényi's information of order α, and show that it satisfies the Ziv-Zakai (ZZ) conditions. We then continue to find the Rényi RDF at high-resolution conditions, for a squared distance distortion measure, and for a source that possesses a smooth probability density function (pdf) $f(s)$ with finite Rényi entropy.

In Section IV, we illustrate the ZZ-Rényi method by finding

This research was partially supported by Israel Science Foundation grant #1259/07, and by a grant from the Advanced Communication Center at Tel-Aviv University.

lower bound on the distortion in fixed-rate scalar quantization. In Section V, we extend the derivation to the vector case and to *r*th power distortion. We compare our bounds to the well-known results of Bennett and Zador. The quantization case provides insight on the tightness of the new ZZ-Rényi bounds.

In section VI, We find the ZZ-Rényi lower bound for delay-constrained transmission through a low dimensional modulo-lattice additive Gaussian noise (MLAN) channel. This gives a bound on the performance of the 'analog matching' scheme [5] at low dimensions.

II. The Ziv-Zakai Method

Shannon's information can be written as

$$I(\mathbf{X};\mathbf{Y}) = -\iint f(\mathbf{x},\mathbf{y})\log\frac{f(\mathbf{x})q(\mathbf{y})}{f(\mathbf{x},\mathbf{y})}d\mathbf{x}d\mathbf{y}, \quad (6)$$

where $f(\mathbf{x},\mathbf{y})$ is the joint pdf of $\mathbf{X}$ and $\mathbf{Y}$, and $f(\mathbf{x})$ and $q(\mathbf{y})$ are its marginals. Throughout this work, we refer to the natural logarithm. Let $\Phi(t)$ be a real-valued function defined on $[0,\infty)$, where $\Phi(0)$ may be ∞, that satisfies:

$$\Phi(t) \text{ is convex,} \quad (7a)$$

$$\lim_{t\to 0} t\cdot\Phi(1/t) = 0. \quad (7b)$$

These conditions also imply that $\Phi(t)$ is non-increasing (see [12] for the proof). The generalized mutual information relative to the function $\Phi(t)$ is defined by

$$I_\Phi(\mathbf{X};\mathbf{Y}) = \iint f(\mathbf{x},\mathbf{y})\Phi\left(\frac{f(\mathbf{x})q(\mathbf{y})}{f(\mathbf{x},\mathbf{y})}\right)d\mathbf{x}d\mathbf{y}. \quad (8)$$

Referring to the setup of Fig. 1, The *generalized data processing theorem* ([12], Thm. 3) states:

$$\begin{aligned} I_\Phi(\mathbf{S};\hat{\mathbf{S}}) &\le I_\Phi(\mathbf{S};\mathbf{Y}) \\ I_\Phi(\mathbf{S};\mathbf{Y}) &= I_\Phi(\mathbf{X};\mathbf{Y}) \end{aligned}, \quad (9)$$

and, consequently,

$$I_\Phi(\mathbf{S};\hat{\mathbf{S}}) \le I_\Phi(\mathbf{X};\mathbf{Y}). \quad (10)$$

The *generalized RDF* of the source S with respect to a distortion measure $d(\mathbf{S},\hat{\mathbf{S}})$ is defined as

$$R_\Phi(D) = \inf I_\Phi(\mathbf{S};\hat{\mathbf{S}}), \quad (11)$$

where the infimum is taken over all the conditional distributions $Q(\hat{\mathbf{s}}|\mathbf{s})$ that satisfy the distortion condition (1). For a channel defined by $Q(\mathbf{y}|\mathbf{x})$ and by a constraint on its input $\mathbf{X}$, The *generalized capacity* is defined as

$$C_\Phi = \sup I_\Phi(\mathbf{X};\mathbf{Y}), \quad (12)$$

where the supremum is taken over all input distributions that satisfy the given constraint. Combining (10)-(12), we get

$$\begin{aligned} R_\Phi(D) = \inf I_\Phi(\mathbf{S};\hat{\mathbf{S}}) &\le I_\Phi(\mathbf{S};\hat{\mathbf{S}}) \\ \le I_\Phi(\mathbf{X};\mathbf{Y}) &\le \sup I_\Phi(\mathbf{X};\mathbf{Y}) = C_\Phi \end{aligned}. \quad (13)$$

We now note a significant difference between Shannon's information and the generalized measures. Let the source and channel be memoryless, which means that $(s_1,x_1,y_1,\hat{s}_1)$ and $(s_2,x_2,y_2,\hat{s}_2)$ are independent, and let *k*=*n*. For a block of length *n*, The additivity property of the -log(·) function in Shannon's information leads to the same inequality:

$$\begin{aligned} &\inf I(\mathbf{S};\hat{\mathbf{S}}) = \inf I(S_1,S_2,...,S_n;\hat{S}_1,\hat{S}_2,...,\hat{S}_n) \\ &= n\cdot\inf I(S_1;\hat{S}_1) = n\cdot R_1(D) \end{aligned}, \quad (14)$$

and similarly for the channel capacity, which together lead to:

$$n\cdot R_1(D) \le n\cdot C_1. \quad (15)$$

This essentially means that Shannon's bound does not depend on the block length. On the other hand, for any other function $\Phi(\cdot)$, we have

$$I_\Phi(X_1,X_2;Y_1,Y_2) \ne I_\Phi(X_1;Y_1) + I_\Phi(X_2;Y_2). \quad (16)$$

Thus, the bound for a different block length will also be different. For a good choice of function $\Phi(\cdot)$, this bound may be stricter than Shannon's.

III. Rényi's Information and RDF

In 1961, Alfréd Rényi introduced an "information of order α" as a generalization of Shannon's information measure [10]. *Rényi's entropy of order α* is given by

$$H_\alpha(\mathbf{X}) = \frac{1}{1-\alpha}\log\int f(\mathbf{x})^\alpha d\mathbf{x}. \quad (17)$$

Rényi's information of order α is given by

$$I_\alpha(\mathbf{X};\mathbf{Y}) = \frac{1}{\alpha-1}\log\iint f(\mathbf{x},\mathbf{y})\cdot\left(\frac{f(\mathbf{x},\mathbf{y})}{f(\mathbf{x})q(\mathbf{y})}\right)^{\alpha-1}d\mathbf{x}d\mathbf{y}. \quad (18)$$

In both definitions, $\alpha > 0$ and $\alpha \ne 1$. In the limit when α approaches 1, Rényi's entropy and information become identical to Shannon's original definition. While the above definition is not in the structure of the Ziv-Zakai formulation, we define the *Rényi information power of order α* as:

$$\begin{aligned} IP_\alpha(\mathbf{X};\mathbf{Y}) &= e^{(\alpha-1)I_\alpha(\mathbf{X};\mathbf{Y})} \\ &= \iint f(\mathbf{x},\mathbf{y})\cdot\left(\frac{f(\mathbf{x},\mathbf{y})}{f(\mathbf{x})q(\mathbf{y})}\right)^{\alpha-1}d\mathbf{x}d\mathbf{y} \end{aligned} \quad (19)$$

This definition fits in the formulation of (8) with the function $\Phi_\alpha(t) = t^{1-\alpha}$. For any $\alpha > 1$, we have that $\Phi_\alpha(t)$ is a convex function. We now show that it also satisfies (7b):

$$\lim_{t\to 0} t\Phi_\alpha(1/t) = \lim_{t\to 0} t(1/t)^{1-\alpha} = \lim_{t\to 0} t^\alpha = 0 \quad (20)$$

We therefore have that for $\alpha > 1$, Rényi's information power of order α follows the data processing theorem (9)-(10), and it is thus possible to write the appropriate bound:

$$R_{IP_\alpha}(D) = \inf IP_\alpha(\mathbf{S};\hat{\mathbf{S}}) \le \sup IP_\alpha(\mathbf{X};\mathbf{Y}) = C_{IP_\alpha} \quad (21)$$

Since log(·) is a monotonically increasing function, and since α > 1, we can take the logarithm of the previous equation and divide both sides by $(\alpha-1)$ to arrive at the equivalent bound on Rényi's information of order α:

$$R_\alpha(D) = \inf I_\alpha(\mathbf{S};\hat{\mathbf{S}}) \le \sup I_\alpha(\mathbf{X};\mathbf{Y}) = C_\alpha \quad (22)$$

In their papers, Ziv and Zakai gave several examples of lower bounds using a version of Rényi information power (without referring to it as such). In [13], they used convexity arguments to show that using α=2 must result in a stricter bound on fixed-rate quantization than Shannon's. In the following, we present an explicit calculation for this stricter bound in the limit of high resolution.

We now wish to find the Rényi rate distortion function for a source $\mathbf{X}$ under *r*th power distortion, i.e.

$$d(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|^r. \quad (23)$$

For a given source pdf $f(\mathbf{x})$, we are looking for the conditional distribution function $Q(\mathbf{y}|\mathbf{x})$ that minimizes (18). As we noted, IP_α and I_α are related by a monotonically increasing function, and therefore they are minimized and maximized simultaneously. We minimize IP_α for a simpler derivation. We begin with a derivation for the scalar case,

k=n=1, and for α=2, r=2, and generalize it in Section V.

In our derivation, we use an extension of the *moment-entropy inequality* to Rényi entropy, as appears in [7]. When dealing with Shannon's entropy, the Gaussian distribution maximizes entropy under a power constraint. In [7], the authors present the distribution that maximizes Rényi's entropy under an rth power constraint, and call it a *generalized Gaussian distribution.* They present an inequality of the moments and Rényi entropy of this distribution. We restate their definitions using our terminology.

The generalized Gaussian distribution that maximizes Rényi entropy of order α under rth power constraint, for $r > 0$, is:

$$G_\alpha(x) = \begin{cases} a_{1,r,\alpha}(1+(1-\alpha)\,|x|^r)_+^{1/(\alpha-1)} & \alpha \neq 1 \\ a_{1,r,1}e^{-|x|^r} & \alpha = 1 \end{cases}, \quad (24)$$

where

$$a_{k,r,\alpha} = \begin{cases} \frac{(1-\alpha)^{k/r+1}\Gamma(\frac{k}{2}+1)}{\pi^{k/2}\beta\left(\frac{k}{r}+1,\frac{1}{1-\alpha}-\frac{k}{r}\right)} & \alpha < 1 \\ \frac{\Gamma(\frac{k}{2}+1)}{\pi^{k/2}\Gamma\left(\frac{k}{r}+1\right)} & \alpha = 1, \\ \frac{(\alpha-1)^{k/r+1}\Gamma(\frac{k}{2}+1)}{\pi^{k/2}\beta\left(\frac{k}{r}+1,\frac{1}{\alpha-1}\right)} & \alpha > 1 \end{cases} \quad (25)$$

Γ(x) denotes the gamma function, $\beta(a,b)$ denotes the beta function, and t_+= max(t,0). They define a scaled version, for t>0:

$$G_{\alpha,t}(x) = G_\alpha(x/t)/t \quad (26)$$

We note that for any $\alpha > 1$, $G_{\alpha,t}(x)$ has bounded support. The rth deviation is defined for 0<r<∞ as the rth root of the rth moment of a pdf f:

$$\sigma_r[f] = \left(\int |x|^r f(x)dx\right)^{1/r} \quad (27)$$

The Rényi moment-power inequality follows:

Theorem 1- Scalar Rényi moment-power inequality [7]: Let $f:\mathbb{R}\to\mathbb{R}$ be a pdf. If $0 \le r \le \infty, \alpha > 1/(1+r)$, and, by the definitions in (17), (27), $H_\alpha[f], \sigma_r[f] < \infty$, then

$$\frac{\sigma_r[f]}{e^{H_\alpha[f]}} \ge \frac{\sigma_r[G_\alpha]}{e^{H_\alpha[G_\alpha]}}. \quad (28)$$

In the above, G_α is given by (24). Equality holds iff $f = G_{\alpha,t}$ of (26) for some $t \in (0,\infty)$. The theorem is a statement of the maximum entropy quality of G_α, where for any f that shares the same rth deviation as G_α, the latter has higher Rényi entropy.

We also use the following modification of Hölder's inequality. Let $f, g : \mathbb{R} \to \mathbb{R}$ be two functions, then:

$$\int f^{k/r} g^{-k/r} \ge \|f\|_{\frac{k}{k+r}}^{k/r} \cdot \|g\|_1^{-k/r} \quad (29)$$

$$\|f\|_p = \left(\int f^p\right)^{1/p} \quad (30)$$

Equality in (29) is achieved iff $g = c \cdot f^{k/k(k+r)}$.

We now state our high-resolution result for α=2, r=2, k=1.

Theorem 2 – High-resolution Rényi$_2$ RDF for squared distortion: For a source S with pdf $f(s)$ that can be approximated by a piecewise-flat function, and possesses finite Rényi entropy (17) of order 2 and finite second moment (27), the high resolution Rényi RDF of order 2 with squared distortion is:

$$R_2(D) \cong \frac{1}{2}\log\left(\frac{9}{125}\|f\|_{1/3}\cdot\frac{1}{D}\right) \quad (31)$$

where ≅ is taken to mean that the difference between the two goes to 0 in the limit $D \to 0$.

A rigorous proof can be obtained following an approach similar to the derivation of high-resolution quantization in [1]. We present here a heuristic approach, using high-resolution assumptions about the densities of x and y, which parallel the approximations used in high-resolution quantization:

High Resolution Assumptions (HRA): in the limit $D \to 0$,

(A) For any x, $Q(y|x)$ is localized within a small neighborhood $\|y-x\| \le \Delta_x$.

(B) For any y in this neighborhood of x, $q(y) \approx f(x)$.

Rényi's information power (19) of order α=2 now becomes:

$$\begin{aligned} IP_2(X;Y) &= \int f(x)\int \tfrac{Q^2(y|x)}{q(y)}dydx \\ &\overset{(A)}{\approx} \int f(x)\int_{x-\Delta_x}^{x+\Delta_x} \tfrac{Q^2(y|x)}{q(y)}dydx \overset{(B)}{\approx} \iint Q^2(y|x)dydx \end{aligned} \quad (32)$$

where we used HRA (A) and (B). For each x, we define y_x=y-x, and $Q_x(y_x)$=$Q(y|x)$. We note that for any x, Q_x is a probability distribution function of y_x. The last expression has the form of Rényi's entropy power of order 2 of $Q_x(y_x)$:

$$\begin{aligned} IP_2(X;Y) &\approx \iint Q_x^2(y_x)dy_xdx = \int e^{-H_2[Q_x(y_x)]}dx \\ &\overset{(a)}{\ge} \int \tfrac{3}{5^{3/2}}\left(\int y_x^2 Q_x(y_x)dy\right)^{-1/2}dx \\ &= \tfrac{3}{5^{3/2}}\int f^{1/2}(x)\cdot\left(f(x)\int y_x^2 Q_x(y_x)dy\right)^{-1/2}dx \\ &\overset{(b)}{\ge} \tfrac{3}{5^{3/2}}\tfrac{\left(\int f^{1/3}(x)dx\right)^{3/2}}{\left(\int f(x)\int y_x^2 Q_x(y_x)dydx\right)^{1/2}} \overset{(c)}{=} \sqrt{\tfrac{9}{125}\|f\|_{1/3}\cdot\tfrac{1}{D}} \end{aligned} \quad (33)$$

In the above, (a) is due to Theorem 1, and a direct calculation on G_2, and is achieved for Q_x that satisfies

$$Q_x(y_x) = (1-y_x^2/\Delta_x^2)_+/\Delta_x, \quad (34)$$

where Δ_x is the radius of the neighborhood of x where $Q_x(y_x)$ is positive. In (b) we used (29), with equality when Δ_x solves:

$$\int_{-\Delta_x}^{\Delta_x} y_x^2 Q_x(y_x)dy_x = c\cdot f^{-2/3}(x) \quad (35)$$

In (c), we used the definition of the overall distortion:

$$D = \int f(x,y)(y-x)^2dydx = \int f(x)\int Q_x(y_x)y_x^2dydx \quad (36)$$

We note that following this derivation for α=1 amounts to the asymptotical tightness of Shannon's lower bound, which is a lower bound at any distortion level. Unfortunately, in the case of α>1, we will later show that (33) is not a lower bound at all distortion levels. Taking the logarithm of (33) gives (31).

IV. Fixed-Rate Scalar Quantization

We now use the result of Section III to find an 'informational' lower bound on the distortion achievable in fixed-rate high-resolution scalar quantization (k=n=1). We use a noiseless N-ary channel in the setup of Fig. 1. This requires a representation of the source S by one of N possible channel inputs, and a reconstruction of $\hat{S}$ out of the same output. This is equivalent to the representation of S by an N-level quantizer.

We first find the Rényi$_2$ capacity of the N-ary noiseless channel. Let p_i be the probability of the ith input, and q_j the probability of the jth output. The noiseless transition matrix is $Q(j|i)$=$\delta_{i,j}$, and thus q_j is equal to p_j. We find:

$$e^{(\alpha-1)C_\alpha} = \max_{p_i}\sum_{i,j} p_i\frac{\delta_{ij}^\alpha}{q_j^{\alpha-1}} = \max_{p_i}\sum_i p_i^{2-\alpha} = N^{\alpha-1} \quad (37)$$

The maximum above is achieved for $0 \le \alpha \le 2$ by the uniform distribution and gives C_α=log$N$. We note that for $\alpha > 2$, setting any single input to have 0 probability results in infinite Rényi information, which means that the N-ary channel has infinite capacity, and thus using $\alpha > 2$ does not lead to a useful bound. Combining (31) and (37) in the manner of (13), we have that in the limit of high resolution,

$$\frac{1}{2}\log\left(\frac{9}{125}\left\|f\right\|_{1/3}\cdot\frac{1}{D}\right)\leq\frac{1}{2}\log N^{2}, \tag{38}$$

which gives the ZZ-Rényi lower bound on distortion:

$$D_{ZZ-Rényi}\geq\frac{1}{N^{2}}\cdot\frac{9}{125}\left\|f\right\|_{1/3}. \tag{39}$$

Bennett's classic result about the optimum distortion achievable in fixed rate scalar quantization is (see [4]):

$$D_{Bennett}=\frac{1}{N^{2}}\cdot\frac{1}{12}\left\|f\right\|_{1/3}. \tag{40}$$

It is apparent that our lower bound has a similar form to Bennett's result, and it is lower by a factor of ~1.16, or 0.63dB, for any source pdf. We remark on this gap in Section V.

Repeating (38) with α=1 results in Shannon's lower bound:

$$D_{Shannon}\geq\frac{1}{N^{2}}\cdot\frac{1}{2\pi e}\int f\log f. \tag{41}$$

For a Gaussian source, for example, calculation shows that Shannon's bound is ~4.3dB below Bennett's result, making our new bound tighter by ~3.7dB.

V. Fixed-Rate Vector Quantization

The vector definition of $G_{\alpha}(\mathbf{x})$, along with the vector Rényi moment-entropy inequality, appear in [8]:

Theorem 3 - Vector Rényi moment-power inequality [8]: For 0<r<∞, $\alpha > k/(k+r)$, and a random vector $\mathbf{X}$ in $\mathbb{R}^{k}$ with finite Rényi entropy of order α (17) and rth deviation (27),

$$\frac{E\left\{\left\|\mathbf{X}\right\|^{r}\right\}^{1/r}}{e^{\frac{1}{k}H_{\alpha}(\mathbf{X})}}\geq c_{k,r,\alpha}, \tag{42}$$

where for $a_{k,r,\alpha}$ as given in (25), $c_{k,r,\alpha}$ is given by:

$$c_{k,r,\alpha}=a_{k,r,\alpha}^{1/k}\left[\alpha\left(1+\tfrac{r}{k}\right)-1\right]^{-\frac{1}{r}}b_{k,r,\alpha}$$
$$b_{k,r,\alpha}=\begin{cases}\left(1-\frac{k(1-\alpha)}{r\alpha}\right)^{\frac{1}{k(1-\alpha)}} & \alpha\neq 1\\ e^{-1/r} & \alpha=1\end{cases} \tag{43}$$

The derivation in the vector case follows the same method of section III, using the HRA with the neighborhoods of **y**=**x** interpreted as k-dimensional neighborhoods. Following the same derivation as in (33) results in the following:

Theorem 4 – High-resolution vector Rényi RDF for rth power distortion: For a k-dimensional source S with pdf $f(\mathbf{s})$ that can be approximated by a cubewise-flat function, and that possesses finite Rényi entropy (17) of order α and finite rth deviation (27), the Rényi RDF at high resolution is:

$$R_{\alpha}(D)\cong\frac{k}{r}\log\left(\frac{1}{k}c_{k,r,\alpha}^{r}\cdot\left\|f^{\frac{(2-\alpha)r+k(\alpha-1)}{k(\alpha-1)}}\right\|_{\frac{k(\alpha-1)}{k(\alpha-1)+r}}\cdot\frac{1}{D}\right) \tag{44}$$

where ≅ has the same meaning as in Theorem 2.

We can now examine the dependency on α of the above. In Fig. 3(a), we plot the argument in the parentheses for D=1 and r=2, for different α and k. The highest $R_{\alpha}(D)$ is at α=2. Setting **Y** to **0** for any **X** has 0 information, and a distortion equal to the source deviation, thus R(1)=0. Fig. 3(a) demonstrates that for any α>1, setting D=1 in (44) results in an expression greater than 0, and thus (44) is not a lower bound for all D.

We now apply the Ziv-Zakai method using the same N-ary noiseless channel of section IV, with a single channel use for each source sample (n=1). For $0\leq\alpha\leq 2$, this channel has Rényi capacity of $\log N$, as stated in (37). We find the following lower bound on fixed-rate vector quantization:

$$D_{ZZ-Rényi}\geq\frac{1}{N^{r/k}}\cdot\frac{1}{k}c_{k,r,\alpha}^{r}\cdot\left\|f^{\frac{(2-\alpha)r+k(\alpha-1)}{k(\alpha-1)}}\right\|_{\frac{k(\alpha-1)}{k(\alpha-1)+r}} \tag{45}$$

While the general solution for the vector case is unknown, Zador has shown (see [3], [4]) that

$$D_{Zador}=\frac{A(k,r)}{N^{r/k}}\cdot\left\|f\right\|_{\frac{k}{k+r}}, \tag{46}$$

where $A(k,r)$ is independent of the source distribution. For α=2, the ZZ-Rényi bound becomes a parallel:

$$D_{ZZ-Rényi}\geq\frac{1}{N^{r/k}}\cdot\frac{1}{k}c_{k,r,2}^{r}\cdot\left\|f\right\|_{\frac{k}{k+r}}, \tag{47}$$

and $\frac{1}{k}c_{k,r,\alpha}^{r}$ becomes a lower bound on $A(k,r)$.

Gersho has conjectured that $A(k,r)$ is the least normalized moment of inertia of k-dimensional tessellating polytopes [3], a result which has been proven for k=1 and k=2, r=2. Even if his conjecture is true, the best tessellating polytopes are only known for low dimensions, and mostly for r=2. Several authors have given lower and upper bounds to $A(k,r)$, and we use Zador's original bound in order to compare to our new bound:

$$\Gamma\left(1+\frac{r}{k}\right)\cdot V_{k}^{-r/k}\geq A(k,r)\geq\frac{1}{k+r}\cdot V_{k}^{-r/k} \tag{48}$$

where V_k is the volume of a k-dimensional unit sphere. A direct calculation gives:

$$\frac{1}{k}c_{k,r,2}^{r}=\frac{1}{k+2r}V_{k}^{-r/k}\cdot\left(1+\frac{k}{k+2r}\right)^{r/k} \tag{49}$$

Although our bound is always lower than Zador's operational lower bound, it comes close, as demonstrated in Fig. 2.

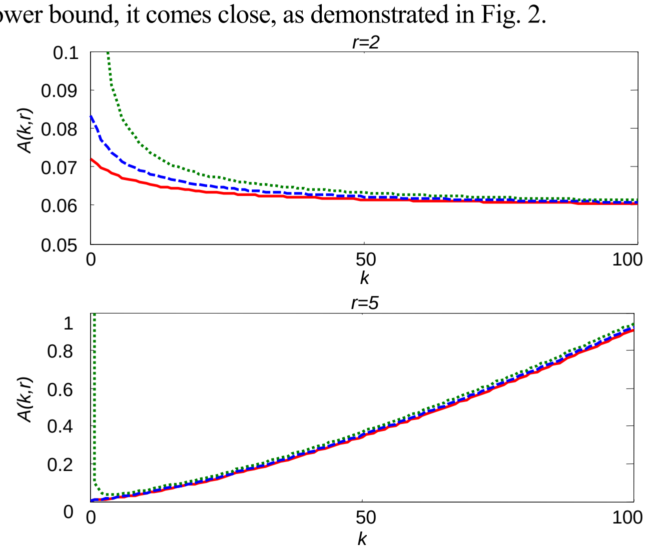


Fig. 2 - Bounds for $A(k,r)$ for different power r and dimension k . The short dashed lines (---) are Zador's upper bound, and his lower bound appears in the long dashed lines (– –). The full line (—) is the new ZZ-Rényi lower bound

Here, as in the scalar case, there is a gap between the bound and the optimum achievable results. We attribute this gap to a mismatch in the distribution of the source given its reconstruction, or the 'backward channel'. In the case of quantization, this distribution is uniform on a quantization cell. In Shannon's case, this distribution is the Gaussian distribution, and in the Rényi case, this is the generalized

Gaussian distribution G_α. The latter has finite support and is closer to the uniform distribution, thereby achieving a closer bound. As k grows, and the Gaussian distribution becomes closer to a uniform distribution on a multi-dimensional cell, so does the generalized Gaussian distribution, and they both give bounds that come closer to the actual achievable performance.

VI. MODULO LATTICE ADDITIVE CHANNEL

Recently, modulo-lattice techniques have been used to cancel interference known at the transmitter [2], utilize side information known at the receiver, or perform a combination of both [5], [9]. In all these schemes, transmission is equivalent to transmission through a modulo lattice additive noise channel, in which $\mathbf{y} = \mathbf{x} + \mathbf{z} \bmod \mathbf{\Lambda}$.(for details on the MLAN channel, see[2] or [6].) In the limit of infinitely many dimensions, these lattice techniques attain the capacity of the AWGN channel, while solving the problem of memory in the source and channel. The use of high dimensional lattices is, however, very complex, and a lower dimensional lattice may serve as a feasible compromise.

One drawback to the use of low dimensional lattices is the loss of shaping gain, which reduces the modulo-lattice channel's capacity in comparison to an AWGN channel (see [2]). For a scalar lattice, for example, this loss is ~0.254 bit/channel use, or ~1.53dB in SNR. However, even in order to achieve this lowered capacity, it is generally necessary to use an additional channel-coding scheme (as in [2]). When complexity or delay is restricted, so that only a finite block length may be used, we can again use the Ziv-Zakai method to find a tighter lower bound.

We present the case where the source and channel bandwidths are equal, i.e. $k=n$, easily extendable to $k\neq n$ (bandwidth reduction/expansion). Rényi's information of order α=2 again gives the tightest bound. In the MLAN channel, a uniform input distribution then achieves capacity, equal to:

$$\begin{aligned} C_2 &= \log\left(\int_{\nu_0} \frac{f(\mathbf{x})}{q(\mathbf{y})} Q^2(\mathbf{y}-\mathbf{x} \bmod \Lambda) d\mathbf{x} d\mathbf{y}\right) \\ &= \log\left(V(\nu_0)\cdot\int_{\nu_0} Q^2(\zeta \bmod \Lambda) d\zeta\right) \end{aligned}, \quad (50)$$

where ν_0 is the basic lattice cell. In the most common case, the noise is Gaussian. Since we performed our derivation for the RDF under high-resolution conditions, we make high-SNR simplifications in the channel as well. These amount to the noise being contained within the basic lattice cell ν_0. Thus:

$$\begin{aligned} C_2 &= \log V(\nu_0) + \log \int_{\nu_0} \frac{1}{(2\pi N)^k} \exp -\frac{\|\zeta\|^2}{N} d\zeta \\ &= \frac{k}{2}\log\left[\frac{1}{4\pi N}\cdot V^{\frac{2}{k}}(\nu_0)\right] \end{aligned} \quad (51)$$

Combining now with (44) at $r=2$, we have :

$$\frac{1}{k} c_{k,2,2}^2 \cdot \|f\|_{\frac{k}{k+2}} \cdot \frac{1}{D} \le \left[\frac{1}{4\pi G(\Lambda)}\cdot\frac{P}{N}\right], \quad (52)$$

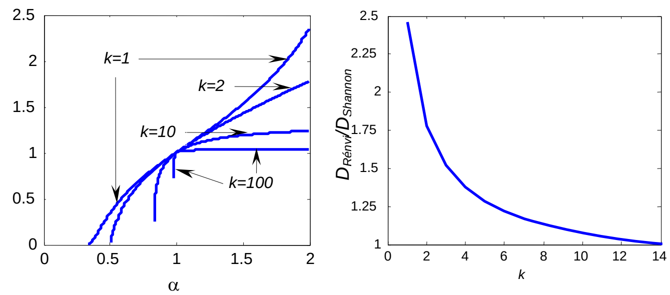
where we have used G(Λ), the normalized second moment of the lattice (see [2]). We arrive at:

$$D_{ZZ-Rényi} \ge \frac{4\pi G(\Lambda)\cdot \frac{1}{k} c_{k,2,2}^2 \cdot \|f\|_{\frac{k}{k+2}}}{SNR} \quad (53)$$

For a Gaussian source, Shannon's bound at high SNR is:

$$\frac{1}{2}\log\frac{1}{D} \le \frac{1}{2}\log SNR \qquad \Rightarrow \quad D_{Shannon} \ge \frac{1}{SNR} \quad (54)$$

We plot the ratio between (53) and (54) in Fig. 3(b) to see how the use of a delay limited modulo-Λ channel deteriorates distortion performance. Since G(Λ) is not known for all k, we replace it with Zador's lower bound, which is exact at k=1 and at $k\rightarrow\infty$. We can see that the ZZ-Rényi lower bound is stricter than Shannon's for k<14, but becomes lower at higher dimensions. For k=1, the bound is $D\geq 2.42/\text{SNR}$. The actual achievable distortion for a scalar MLAN channel is found using a compander [6], which compresses the possibly infinite support of the source into the finite lattice cell, and prevents undesired modulo shifts. The compander follows equations similar to Bennett's error in high-rate quantization, and has the same form as the familiar result of Panter and Dite (see [4]). The actual achievable distortion using this compander is $D_{opt}\approx 2.72/\text{SNR}$. This means that the ZZ-Rényi lower bound is closer to the true distortion than Shannon's bound, but as in the case of quantization, there still is a small gap between the bound and achievable results.



(a) (b)

Fig. 3 – (a) - the dependency of the Rényi RDF on α. (b) - The ratio between the Rényi and Shannon lower bounds on distortion in the MLAN channel at

REFERENCES

[1] J. Bucklew and G. Wise, "Multidimensional asymptotic quantization theory with rth power distortion measures," *IEEE Transactions on Information Theory*, vol.28, no.2, pp. 239-247, Mar 1982

[2] U. Erez, S. Shamai and R. Zamir, "Capacity and lattice-strategies for cancelling known interference", *IEEE Tr. Info. Th.*, IT-51:3820--3833 Nov. 2005.

[3] A. Gersho, "Asymptotically optimal block quantization," *IEEE Trans. Information Theory*, vol.25, no.4, pp. 373-380, Jul 1979

[4] R.M. Gray and D.L. Neuhoff, "Quantization," *IEEE Transactions on Information Theory,* , vol.44, no.6, pp.2325-2383, Oct 1998

[5] Y. Kochman and R.Zamir, "Analog matching of colored sources to colored channels," *2006 IEEE Intl. Symp Info. Theory,* pp.1539-1543.

[6] I. Leibowitz, "The Ziv-Zakai bound at high-resolution, analog matching, and companding," M.Sc. thesis, Dept. Elec. Eng., Tel-Aviv Univ., Tel-Aviv, Israel, 2007.

[7] E. Lutwak; D. Yang; G. Zhang, "Cramér-Rao and moment-entropy inequalities for Rényi entropy and generalized Fisher information,", *IEEE Trans. Information Theory*, vol.51, no.2, pp. 473-478, Feb. 2005.

[8] E. Lutwak; D. Yang; G. Zhang, "Moment–entropy inequalities," *Annals of Probability*, vol.32, pp. 757–774, 2004

[9] N. Merhav and S. Shamai. On Joint Source-Channel Coding for the Wyner-Ziv Source and the Gel'fand-Pinsker Channel. *IEEE Trans. Info. Theory*, vol.40, pp. 2844-2855, Nov. 2003.

[10] A. Rényi, "On Measures of Entropy and Information." *Proc.4th Berk. Symp. Math., Stat. and Prob.*, v1. U. of Calif. Press. pp. 547-561, 1961.

[11] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol.27, pt. I, pp. 379–423, 1948; pt. II, pp. 623–656, 1948.

[12] J. Ziv and M. Zakai, "On functionals satisfying a data-processing theorem," *IEEE Trans. Info. Theory*, vol.19, pp. 275-283, May 1973.

[13] J. Ziv and M. Zakai, ``A Generalization of the Rate-Distortion Theory and Applications", in *Information Theory - New Trends and Open Problems*, edited by G. Longo, Springer-Verlag, pp. 88–125, Wien, 1976