Fixed-length lossy compression in the finite blocklength regime: Gaussian source

Victoria Kostina

Dept. of Electrical Engineering
Princeton University
NJ, 08544, USA
vkostina@princeton.edu

Sergio Verdú

Dept. of Electrical Engineering
Princeton University
NJ, 08544, USA
verdu@princeton.edu

***Abstract*—For an i.i.d. Gaussian source with variance σ^2, we show that it is necessary to spend $\frac{1}{2}\ln\frac{\sigma^2}{d}+\frac{1}{\sqrt{2n}}Q^{-1}(\epsilon)+O\left(\frac{\ln n}{n}\right)$ nats per sample in order to reproduce n source samples within mean-square error d with probability at least $1-\epsilon$, where $Q^{-1}(\cdot)$ is the inverse of the standard Gaussian complementary cdf. The first-order term is the rate-distortion function of the Gaussian source, while the second-order term measures its stochastic variability. We derive new achievability and converse bounds that are valid at any blocklength and show that the second-order approximation is tightly wedged between them, thus providing a concise and accurate approximation of the minimum achievable source coding rate at a given fixed blocklength (unless the blocklength is very small).**

***Index Terms*—Shannon theory, lossy source coding, rate distortion, memoryless sources, Gaussian source, finite blocklength regime, achievability, converse, sphere covering.**

I. Introduction

If the blocklength is permitted to grow without limit, the rate-distortion function characterizes the minimal source coding rate compatible with a given distortion level, either in average or excess distortion sense. However, as relatively short blocklengths are common in some applications due to both delay and coding complexity constraints, it is of critical practical interest to assess the unavoidable overhead over the rate-distortion function required to sustain the desired fidelity at a given fixed blocklength. Nether the strong version of Shannon's source coding theorem nor the reliability function, which gives the asymptotic exponential decay of the probability of exceeding a given distortion level when compressing at a fixed rate, supply an answer to that question.

The basic general achievability (Theorem 4 in Section IV) and converse (stating that the rate of the code that achieves average distortion d must be greater than the rate-distortion function) can be extracted from Shannon's coding theorem for memoryless sources [1]. Considerable attention has been paid to the asymptotic behavior of the redundancy, i.e. the difference between the average distortion $D(n,R)$ of the best $n-$dimensional quantizer and the distortion-rate function $D(R)$. Wyner [2] showed that for the i.i.d. Gaussian memoryless source with mean-square error distortion (GMS) the redundancy is upper-bounded by $O\left(\sqrt{\frac{\log n}{n}}\right)$. Note that Wyner obtained his result using a variant of Shannon's achievability bound that is rather loose in the finite blocklength regime. More recently, Ihara and Kubo [3] found the reliability function of the GMS. For variable-length quantization of the GMS at a guaranteed distortion level, Sakrison's achievability bound applies [4]. Kontoyiannis [5] studied second-order refinements of the lossy source coding theorem for the variable-length setup and also presented a nonasymptotic converse that parallels Barron's converse for lossless compression [6].

This paper presents new achievability (upper) and converse (lower) bounds to the minimum rate sustainable as a function of blocklength and excess probability for the GMS. In addition, we show that the minimum achievable finite blocklength coding rate of the GMS is well approximated by

$$R(n,d,\epsilon)\approx\frac{1}{2}\log\frac{\sigma^2}{d}+\frac{1}{\sqrt{2n}}Q^{-1}(\epsilon)\log e, \tag{1}$$

where n is the blocklength, σ^2 is the variance of the source, d is the distortion threshold, and ϵ is the probability of distortion exceeding d. Unless the blocklength is small, (1) is tightly sandwiched between our upper and lower bounds to $R(n,d,\epsilon)$, thus providing an excellent estimate of the minimum achievable finite blocklength coding rate.

Some of the results in this paper find counterparts for discrete memoryless sources with symbol error distortion, as we show in [7].

The rest of the paper is organized as follows. Section II introduces system model and definitions. Sections III, IV and V focus on the converse, the achievability and the second-order analysis, respectively. Section VI presents a numerical comparison of the new bounds, Shannon's achievability result and the second-order approximation. Section VII concludes the paper.

II. Definitions

In the standard model of fixed-to-fixed (block) compression, the output of a source with alphabet $\mathcal{A}^n$ and source distribution P_{X^n} is mapped to one of the M codewords from the reproduction alphabet $\mathcal{B}^n$. A lossy code is a pair of mappings $\mathsf{f}:\mathcal{A}^n\mapsto\{1,\ldots,M\}$ and $\mathsf{c}:\{1,\ldots,M\}\mapsto\mathcal{B}^n$. A distortion measure

This research was supported in part by NSF under grants CCF-1016625 and CCF 09-39370. The first author was supported in part by the Natural Sciences and Engineering Research Council of Canada.

$d : \mathcal{A}^n \times \mathcal{B}^n \mapsto \mathbb{R}_+$ is used to quantify the performance of a lossy code. In the case of i.i.d. Gaussian memoryless source with mean-square error distortion, $P_{X^n} = \mathcal{N}(0, \sigma^2 \mathbf{I})$, $\mathcal{A} = \mathcal{B} = \mathbb{R}$, and $d(x^n, y^n) = \frac{1}{n}\sum_{i=1}^n (x_i - y_i)^2$.

To judge the performance of a lossy code, its average distortion over the source statistics is often considered. A stronger criterion is also used, namely, the probability of exceeding a given distortion level (called *excess-distortion probability*). The following definitions abide by the excess distortion criterion.

Definition 1. *An (n, M, d, ϵ) code for $\{\mathcal{A}^n,\ \mathcal{B}^n,\ P_{X^n},\ d : \mathcal{A}^n \times \mathcal{B}^n \mapsto \mathbb{R}_+\}$ is a code with $|\mathsf{f}| = M$ such that $\mathbb{P}\left[d\left(X^n, \mathsf{c}(\mathsf{f}(X^n))\right) > d\right] \leq \epsilon$.*

Definition 2. *Fix ϵ, d and blocklength n. The minimum achievable code size and the finite blocklength rate-distortion function (excess distortion) are defined by, respectively*

$$M^\star(n, d, \epsilon) = \min\left\{M :\ \exists (n, M, d, \epsilon)\ \text{code}\right\} \tag{2}$$

$$R(n, d, \epsilon) = \frac{1}{n}\log M^\star(n, d, \epsilon) \tag{3}$$

III. Converse

Let $B_0(r)$ denote the n-dimensional ball of radius r with center at 0. Our first result is a simple sphere-covering converse.

Theorem 1. *Let $\{X_i\}$ be a Gaussian memoryless source with variance σ^2 and mean-square error distortion. Any (n, M, d, ϵ) code must satisfy*

$$M \geq \left(\frac{r^2}{d}\right)^{\frac{n}{2}}, \tag{4}$$

where r is the solution to

$$\mathbb{P}\left[\sigma^2 Z < n r^2\right] = 1 - \epsilon, \tag{5}$$

where $Z \sim \chi_n^2$ (i.e. chi square distributed with n degrees of freedom).

Proof: Inequality (4) simply states that the minimum number of n-dimensional balls of radius $\sqrt{nd}$ required to cover an n-dimensional ball of radius $\sqrt{n}r$ cannot be smaller than the ratio of their volumes. Without loss of generality, assume $\mathbb{E}[X_i] = 0$. Then $Z = \frac{1}{\sigma^2}\sum_{i=1}^n X_i^2$ is χ_n^2-distributed, and the left-side of (5) is the probability that the source produces a sequence that falls inside $B_0(\sqrt{n}r)$. But as follows from the spherical symmetry of the Gaussian distribution, $B_0(\sqrt{n}r)$ with r satisfying (5) has the smallest volume among all sets in $\mathbb{R}^n$ having probability $1-\epsilon$. Since any (n, M, d, ϵ)-code is a covering of a set that has total probability of at least $1-\epsilon$, the result follows. ■

IV. Achievability

The lower bound to the minimum number of $\sqrt{nd}$-balls required to cover $B_0(\sqrt{n}r)$ in Theorem 1 is the ratio of their volumes. This lower bound is not achievable because the spheres in such a covering will overlap. The following result studies the number of $\sqrt{nd}$-balls to cover $B_0(\sqrt{n}r)$ that is provably achievable.

Theorem 2. *For a Gaussian memoryless source with variance σ^2 and mean-square error distortion and $n \geq 2$, there exists an (n, M, d, ϵ) code such that*

$$M \leq \mathbb{M}\left(\frac{r}{\sqrt{d}}\right) \tag{6}$$

where $\mathbb{M}(\cdot)$ is specified in (7), and r is the solution to (5).

Proof: The classical result of Rogers [8] and a recent improvement by Verger-Gaugry [9] show that in $\mathbb{R}^n$, a ball of radius r can be covered by $\mathbb{M}(r)$ balls of radius 1. Thus, there exists a code with no more than $\mathbb{M}\left(\frac{r}{\sqrt{d}}\right)$ codewords such that all source sequences that fall inside $B_0(\sqrt{n}r)$ are reproduced within distortion d. The excess-distortion probability therefore does not exceed the probability that X^n falls outside $B_0(\sqrt{n}r)$. ■

Our next achievability result can be regarded as the rate-distortion counterpart to Shannon's geometric analysis of the Gaussian channel [10].

Theorem 3. *Let $\{X_i\}$ be a Gaussian memoryless source with variance σ^2 and mean-square error distortion. There exists an (n, M, d, ϵ) code with*

$$\epsilon \leq p_n + \mathbb{P}\left[Z < n a^2\right] + \mathbb{P}\left[Z > n b^2\right] \tag{8}$$

where $Z \sim \chi_n^2$,

$$a = \sqrt{1 - \frac{d}{\sigma^2}} - \sqrt{\frac{d}{\sigma^2}} \tag{9}$$

$$b = \sqrt{1 - \frac{d}{\sigma^2}} + \sqrt{\frac{d}{\sigma^2}} \tag{10}$$

$$p_n = \int_{a^2}^{b^2} \varphi_n(z) f_Z(nz)\, n dz \tag{11}$$

$$\mathbb{M}(r) = \begin{cases} e\left(n\ln n + n\ln\ln n + 5n\right) r^n & r \geq n \\ n\left(n\ln n + n\ln\ln n + 5n\right) r^n & \frac{n}{\ln n} \leq r < n \\ \frac{7^{4\ln 7/7}}{4}\sqrt{2\pi}\frac{n\sqrt{n}\left[(n-1)\ln rn + (n-1)\ln\ln n + \frac{1}{2}\ln n + \ln\frac{\pi\sqrt{2n}}{\sqrt{\pi n}-2}\right]}{r\left(1-\frac{2}{\ln n}\right)\left(1-\frac{2}{\sqrt{\pi n}}\right)\ln^2 n} r^n & 2 < r < \frac{n}{\ln n} \\ \sqrt{2\pi}\frac{\sqrt{n}\left[(n-1)\ln rn + (n-1)\ln\ln n + \frac{1}{2}\ln n + \ln\frac{\pi\sqrt{2n}}{\sqrt{\pi n}-2}\right]}{r\left(1-\frac{2}{\ln n}\right)\left(1-\frac{2}{\sqrt{\pi n}}\right)} r^n & 1 < r \leq 2 \end{cases} \tag{7}$$

$$\varphi_n(z) = \left[1 - \frac{\Gamma\left(\frac{n}{2}+1\right)}{\sqrt{\pi}n\Gamma\left(\frac{n-1}{2}+1\right)}\left(1 - \frac{\left(1+z-2\frac{d}{\sigma^2}\right)^2}{4\left(1-\frac{d}{\sigma^2}\right)z}\right)^{\frac{n-1}{2}}\right]^M$$

and $f_Z(\cdot)$ is the probability density function of Z.

Proof: Consider a code (f, c) with M representation points drawn independently from the uniform distribution on the surface of the sphere with center at $\mathbf{0}$ and radius $r_0 = \sqrt{n}\sigma\sqrt{1-\frac{d}{\sigma^2}}$. Such positioning of representation points is optimal in the limit of large n, see Fig. 1(a). Indeed, for large n, most source sequences will be concentrated within a thin shell near the surface of the sphere of radius $\sqrt{n}\sigma$. The center of the sphere of radius $\sqrt{nd}$ must be at distance r_0 from the origin in order to cover the largest area of the surface of the sphere of radius $\sqrt{n}\sigma$. Average-case analysis of a variable-length version of this code was performed by Sakrison [4]. In (8), the last two terms represent the probabilities that the source produces an atypical sequence that lies too far from the surface of $B_0(r_0)$ to be covered by any of the spheres of radius $\sqrt{nd}$ around representation points. It remains to show that the probability that a source sequence falls in-between the $\sqrt{nd}$-spheres around the representation points is upper bounded by p_n, that is (we write C in place of c to emphasize averaging over the decoder; we write $\mathbf{X}$ for the vector X^n),

$$\mathbb{P}\left[d\left(\mathbf{X}, \mathsf{C}\left(f(\mathbf{X})\right)\right) > d,\ a^2\sigma^2 < d\left(\mathbf{X}, \mathbf{0}\right) < b^2\sigma^2\right] \leq p_n$$

The result will then follow by the random coding argument.

It is convenient to introduce the following notation.

- $S_n(r) = \frac{n\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)} r^{n-1}$ - surface area of an n-dimensional sphere of radius r;
- $S_n(r,\theta)$ - surface area of an n-dimensional polar cap of radius r and polar angle θ.

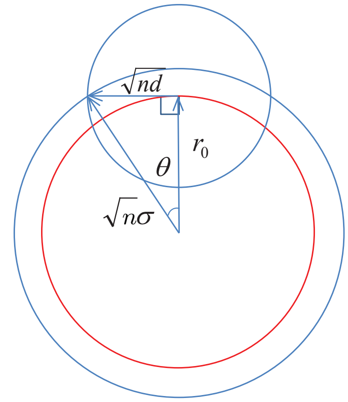
Similar to [4], [10], from Fig. 1(b),

$$S_n(r,\theta) \geq \text{area of an } n-1\text{-dimensional disc of radius } r\sin\theta = \frac{\pi^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2}+1\right)}(r\sin\theta)^{n-1}$$

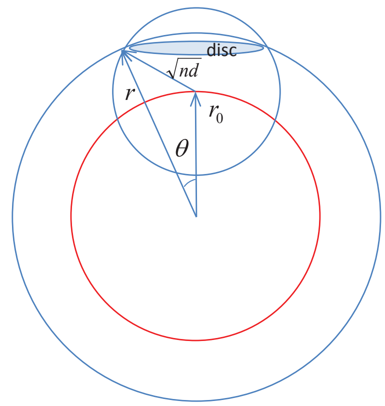
Consider the probability that $\mathbf{X}$ is not reproduced within distortion d given that $\mathbf{X}$ is at distance r from the origin, $\mathbb{P}\left[d\left(\mathbf{X}, \mathsf{C}\left(f(\mathbf{X})\right)\right) > d \mid d\left(\mathbf{X}, \mathbf{0}\right) = r\right]$. From the spherical symmetry of the Gaussian distribution and the fact that M representation points are drawn independently from the uniform distribution and Fig. 1(b), this probability is given by

$$\mathbb{P}\left[d\left(\mathbf{X}, \mathsf{C}\left(f(\mathbf{X})\right)\right) > d \mid d\left(\mathbf{X}, \mathbf{0}\right) = r\right] = \left[1 - \frac{S_n(r,\theta)}{S_n(r)}\right]^M$$
$$\leq \left[1 - \frac{\Gamma\left(\frac{n}{2}+1\right)}{\sqrt{\pi}n\Gamma\left(\frac{n-1}{2}+1\right)}(\sin\theta)^{n-1}\right]^M \tag{12}$$

By the law of cosines, $\cos\theta = \frac{r^2+r_0^2-nd}{2rr_0}$, and, substituting $\sin^2\theta = 1-\cos^2\theta$ into (12), one obtains $\varphi_n(z)$. Since the squared-norm of the Gaussian vector $\mathbf{X}$ is χ_n^2-distributed, the integration of (12) with respect to such a measure yields p_n. ■



(a)



(b)

Fig. 1. Optimum positioning of the representation sphere (a) and the geometry of the excess-distortion probability calculation (b).

Finally, for completeness and because we will compare our results against it in Section VI, we list the only existing finite blocklength bound known to us that is applicable to the block compression of the Gaussian source. It can be distilled [11] from Shannon's coding theorem for memoryless sources.

Theorem 4 (Shannon's random coding bound [1])**.** *Fix P_{X^n}, a positive integer M and $d \geq 0$. There exists an (n, M, d, ϵ) code such that*

$$\epsilon \leq \inf_{P_{Y^n|X^n}} \mathbb{P}\left[d\left(X^n, Y^n\right) > d\right] + \inf_{\gamma>0}\left\{\mathbb{P}\left[\imath_{X^n;Y^n}\left(X^n;Y^n\right) > \log M - \gamma\right] + e^{-\exp(\gamma)}\right\} \tag{13}$$

where

$$\imath_{U;V}(a;b) = \log\frac{dP_{UV}}{d(P_U \times P_V)}(a,b) \tag{14}$$

denotes the information density of the joint distribution P_{UV} at (a,b).

V. Second-Order Analysis

While Theorems 1, 2 and 3 provide firm lower and upper bounds to the function $R(n,d,\epsilon)$ valid for any blocklength

n, a compact and accurate approximation of $R(n,d,\epsilon)$ for moderate to large values of n is supplied by the second-order asymptotic analysis of $R(n,d,\epsilon)$. Toward this end, in the spirit of [12] we introduce the following definition.

Definition 3 (rate dispersion [7])**.** *Fix $d \geq 0$. The rate-dispersion function (squared information units per source symbol) is defined as*

$$V(d) = \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{n\left(R(n,d,\epsilon) - R(d)\right)^2}{2 \ln \frac{1}{\epsilon}} \tag{15}$$

Suppose the target is to sustain the probability of exceeding distortion d bounded by ϵ at rate $R = (1+\eta)R(d)$ for some $\eta > 0$. As (1) implies, the required blocklength scales linearly with rate dispersion:

$$n(d,\eta,\epsilon) \approx \frac{V(d)}{R^2(d)} \left(\frac{Q^{-1}(\epsilon)}{\eta}\right)^2 \tag{16}$$

where note that only the first factor depends on the source, while the second depends only on the design specifications.

The rate-dispersion function of the Gaussian source with mean-square error distortion can be found either directly from our achievability and converse results, or as a particular case of the following general result.

Theorem 5 (Second-order approximation [7])**.** *Fix stationary memoryless source $\{X_i\}$ with alphabet $\mathcal{A}$ and separable distortion measure. Under the following technical conditions,*

(i) For all $\mathsf{x} \in \mathcal{A}$ with positive probability, $\min_{\mathsf{y} \in \mathcal{B}} d(\mathsf{x},\mathsf{y}) = 0$, and the acceptable distortion level satisfies $0 < d < d_0$, where $d_0 = \min\{d : R(d) = 0\}$.

(ii) The random variable $\log \mathbb{E}\left[\exp\{\lambda^\star d(\mathsf{X},\mathsf{Y})\}|\mathsf{X}\right]$ has finite third moment, where Y is independent of X, and its distribution is the marginal of $P_{\mathsf{X}} P^\star_{\mathsf{Y}|\mathsf{X}}$, where $P^\star_{\mathsf{Y}|\mathsf{X}}$ achieves

$$I(P_{\mathsf{X}}, P^\star_{\mathsf{Y}|\mathsf{X}}) = R(d) = \min_{\substack{P_{\mathsf{Y}|\mathsf{X}}: \\ \mathbb{E}[d(\mathsf{X},\mathsf{Y})] \leq d}} I(\mathsf{X};\mathsf{Y}), \tag{17}$$

and $\lambda^\star < 0$ is the slope of the rate-distortion function at d,

$$\lambda^\star = R'(d) \tag{18}$$

it holds that

$$R(n,d,\epsilon) = R(d) + \sqrt{\frac{V(d)}{n}} Q^{-1}(\epsilon) + \theta\left(\frac{\log n}{n}\right) \tag{19}$$

$$V(d) = \mathrm{Var}\left[\log \frac{1}{\mathbb{E}\left[\exp\{\lambda^\star [d(\mathsf{X},\mathsf{Y}) - d]\}|\mathsf{X}\right]}\right] \tag{20}$$

where Y, $\lambda^\star$ are as in (ii)*, and the remainder term in* (19) *satisfies*

$$-\frac{\log n}{n} + O\left(\frac{1}{n}\right) \leq \theta\left(\frac{\log n}{n}\right) \leq \frac{1}{2}\frac{\log n}{n} + O\left(\frac{\log \log n}{n}\right) \tag{21}$$

Note that $\lambda^\star$ can be alternatively characterized as the unique solution to [13] [14]

$$d = \frac{d}{d\lambda} \mathbb{E}\left[\log \mathbb{E}\left[\exp\{\lambda d(\mathsf{X},\mathsf{Y})\}|\mathsf{X}\right]\right] \tag{22}$$

Since the rate-distortion function can be expressed as [13]

$$R(d) = \mathbb{E}\left[\log \frac{1}{\mathbb{E}\left[\exp\{\lambda^\star [d(\mathsf{X},\mathsf{Y}) - d]\}|\mathsf{X}\right]}\right] \tag{23}$$

the rate-distortion function is equal to the expectation of the random variable whose variance we take in (20), thereby drawing a pleasing parallel with the channel coding results in [12].

Theorem 6. *Let X be a stationary Gaussian memoryless source with variance σ^2 and mean-square error distortion. The minimum achievable rate at blocklength n satisfies*

$$R(n,d,\epsilon) = \frac{1}{2}\log\frac{\sigma^2}{d} + \sqrt{\frac{1}{2n}} Q^{-1}(\epsilon) \log e + \theta\left(\frac{\log n}{n}\right) \tag{24}$$

where the remainder term satisfies

$$O\left(\frac{1}{n}\right) \leq \theta\left(\frac{\log n}{n}\right) \leq \frac{1}{2}\frac{\log n}{n} + O\left(\frac{\log\log n}{n}\right) \tag{25}$$

Proof: The achievability part follows from Theorem 5: if $\mathsf{X} \sim (0, \sigma^2)$, then $\mathsf{Y} \sim (0, \sigma^2 - d)$, $d(\mathsf{x}, \mathsf{Y}) = (\mathsf{Y} - \mathsf{x})^2$, so

$$\lambda^\star = R'(d) = -\frac{1}{2d} \tag{26}$$

$$\mathbb{E}\left[e^{\lambda^\star d(\mathsf{x},\mathsf{Y})}\right] = \sqrt{\frac{\sigma^2}{d}} e^{-\frac{\mathsf{x}^2}{2\sigma^2}} \tag{27}$$

$$V(d) = \mathrm{Var}\left[\ln \mathbb{E}\left[e^{\lambda^\star d(\mathsf{X},\mathsf{Y})}|\mathsf{X}\right]\right] = \frac{1}{2} \tag{28}$$

where we performed all calculations in nats for simplicity.

Alternatively, the achievability part can be proven using Theorem 3. Theorem 2 also leads to the correct rate-dispersion term, but produces a weaker $\frac{\log n}{n}$ term.

Let us show the converse part, which has a stronger $\frac{\log n}{n}$ term than that in Theorem 5.

Since in Theorem 1 $\sigma^2 Z = \sum_{i=1}^n X_i^2$, $X_i \sim \mathcal{N}(0,\sigma^2)$, we apply the Berry-Esseen central limit theorem (CLT) to $Z_i = \frac{1}{\sigma^2} X_i^2$. Each Z_i has mean, second and third central moments equal to 1, 2 and 8, respectively. Let

$$\bar{r}^2 = \sigma^2\left(1 + \sqrt{\frac{2}{n}} Q^{-1}\left(\epsilon + \frac{12\sqrt{2}}{\sqrt{n}}\right)\right) \tag{29}$$

$$= \sigma^2 + \sqrt{\frac{2}{n}} \sigma^2 Q^{-1}(\epsilon) + O\left(\frac{1}{n}\right) \tag{30}$$

Then by the Berry-Esseen CLT

$$\mathbb{P}\left[\sum_{i=1}^n X_i^2 > n\bar{r}^2\right] \geq \epsilon \tag{31}$$

and therefore r that achieves the equality in (5) must satisfy $r^2 \geq \bar{r}^2$. Weakening (4) by plugging $\bar{r}$ instead of r and taking the logarithm of both sides therein, one obtains:

$$\log M \geq \frac{n}{2}\log\frac{\bar{r}^2}{d} \tag{32}$$

$$= \frac{n}{2}\log\frac{\sigma^2}{d} + \sqrt{\frac{n}{2}} Q^{-1}(\epsilon) \log e + O(1) \tag{33}$$

■

Theorem 6 agrees with the recent result of Ingber and Kochman [15], who observed independently of our work [16] that the rate-dispersion function of the Gaussian source with mean-square distortion is equal to $\frac{1}{2}$ nats^2 per source output.

VI. Numerical comparison

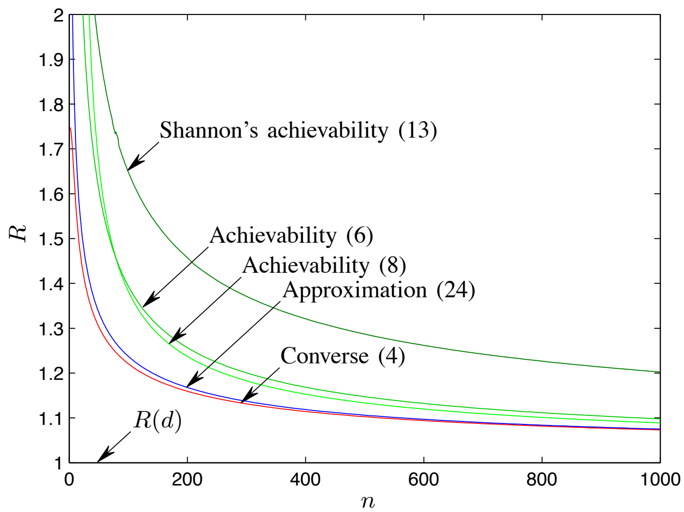


Fig. 2. Bounds to $R(n, d, \epsilon)$ for the Gaussian memoryless source with mean-square error distortion, $\sigma = 1$, $d = \frac{1}{4}$, $\epsilon = 10^{-2}$.

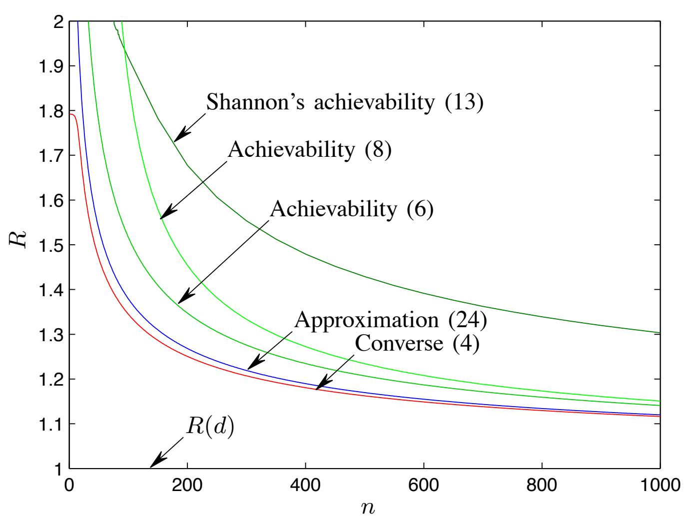


Fig. 3. Bounds to $R(n, d, \epsilon)$ for the Gaussian memoryless source with mean-square error distortion, $\sigma = 1$, $d = \frac{1}{4}$, $\epsilon = 10^{-4}$.

A numerical comparison of Shannon's achievability bound (13) and the new bounds in Theorems 1, 2 and 3 as well as the second-order approximation in Theorem 6 (obtained by letting $\theta\left(\frac{\log n}{n}\right) = 0$) is presented in Figures 2 and 3 for two values of excess-distortion probability ϵ. Shannon's achievability bound, calculated with a stationary memoryless $P_{X^n|Y^n}$, while asymptotically tight, is rather loose in the displayed range of blocklengths. The achievability bound in (6) is tighter than the one in (8) at shorter blocklengths, which should be intuitively expected as the former spreads the codewords inside a ball while the latter places them on the surface of a sphere. As long as the blocklength is not too short, the second-order approximation is tightly sandwiched between our converse and achievability bounds, thus providing an accurate estimate of $R(n, d, \epsilon)$.

VII. Conclusion

To estimate the minimum rate required to sustain a given mean-square error at a given blocklength when compressing a Gaussian i.i.d. source, we showed new achievability and converse bounds that are tighter than existing bounds. As in the discrete case [7], the rate dispersion (along with the rate-distortion function) serves to give tight approximations to the fundamental fidelity-rate tradeoff unless the blocklength is small.

References

[1] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *IRE Int. Conv. Rec.*, vol. 7, pp. 142–163, Mar. 1959.

[2] A. D. Wyner, "Communication of analog data from a Gaussian source over a noisy channel," *Bell Syst. Tech. J*, vol. 47, pp. 801–812, May/June 1968.

[3] S. Ihara and M. Kubo, "Error exponent for coding of memoryless Gaussian sources with a fidelity criterion," *IEICE Transactions On Fundamentals Of Electronics Communications And Computer Sciences E Series A*, vol. 83, pp. 1891–1897, Oct. 2000.

[4] D. Sakrison, "A geometric treatment of the source encoding of a Gaussian random variable," *IEEE Transactions on Information Theory*, vol. 14, pp. 481–486, May 1968.

[5] I. Kontoyiannis, "Pointwise redundancy in lossy data compression and universal lossy data compression," *Information Theory, IEEE Transactions on*, vol. 46, pp. 136–152, Jan. 2000.

[6] A. Barron, *Logically smooth density estimation*. PhD thesis, Yale University, 1985.

[7] V. Kostina and S. Verdú, "Fixed-length lossy compression in the finite blocklength regime: discrete memoryless sources," in *IEEE International Symposium on Information Theory*, (Saint-Petersburg, Russia), Aug. 2011.

[8] C. A. Rogers, "Covering a sphere with spheres," *Mathematika*, vol. 10, no. 02, pp. 157–164, 1963.

[9] J. L. Verger-Gaugry, "Covering a ball with smaller equal balls in R n," *Discrete and Computational Geometry*, vol. 33, no. 1, pp. 143–155, 2005.

[10] C. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell Syst. Tech. J*, vol. 38, no. 3, pp. 611–656, 1959.

[11] S. Verdú, "ELE528: Information theory lecture notes," *Princeton University*, 2009.

[12] Y. Polyanskiy, H. Poor, and S. Verdú, "Channel coding rate in finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, pp. 2307–2359, May 2010.

[13] T. Berger, *Rate distortion theory*. Prentice-Hall Englewood Cliffs, NJ., 1971.

[14] I. Csiszár, "On an extremum problem of information theory," *Studia Scientiarum Mathematicarum Hungarica*, vol. 9, no. 1, pp. 57–71, 1974.

[15] A. Ingber and Y. Kochman, "The Dispersion of Lossy Source Coding," *Arxiv preprint arXiv:1102.2598*, Feb. 2011.

[16] V. Kostina and S. Verdú, "Fixed-length lossy compression in the finite blocklength regime," *Arxiv preprint arXiv:1102.3944*, Feb. 2011.