On the Excess Distortion Exponent for Memoryless Gaussian Source-Channel Pairs*

Yangfan Zhong, Fady Alajaji and L. Lorne Campbell
Department of Mathematics and Statistics
Queen's University, Kingston, ON K7L 3N6, Canada
Email: {yangfan,fady,campblll}@mast.queensu.ca

***Abstract*— For a memoryless Gaussian source under the squared-error distortion fidelity criterion and a memoryless additive Gaussian noise channel with a quadratic power constraint at the channel input, upper and lower bounds for the joint source-channel coding excess distortion exponent (which is the exponent of the probability of excess distortion) are established. A necessary and sufficient condition for which the two bounds coincide is provided, thus exactly determining the exponent. This condition is observed to hold for a wide range of source-channel parameters.**

I. Introduction

In [5], Csiszár studies the joint source-channel coding (JSCC) excess distortion exponent under a fidelity criterion for discrete memoryless systems – i.e., the largest rate of asymptotic decay of the probability that the distortion resulting from transmitting the source over the channel via a joint source-channel (JSC) code exceeds a certain tolerated threshold. Specifically, given a discrete memoryless source (DMS) Q and a discrete memoryless channel (DMC) W (both with finite alphabets), a transmission rate t and a distortion measure, Csiszár shows that the lower (respectively upper) bound of the JSCC excess distortion exponent $E_J(Q,W,\Delta,t)$ under a distortion threshold Δ is given by the minimum of the sum of $tF(R/t,Q,\Delta)$ and $E_r(R,W)$ (respectively $E_{sp}(R,W)$) over R, where $F(R,Q,\Delta)$ is the source excess distortion exponent with distortion threshold Δ [10], and $E_r(R,W)$ and $E_{sp}(R,W)$ are respectively the random-coding and sphere-packing channel error exponents [7]. If the minimum of the lower (or upper) bound is attained for an R larger than the critical rate of the channel, then the two bounds coincide and E_J is determined exactly. The analytical computation of these bounds has been partially addressed in [12], where the authors use Fenchel duality [9] to provide equivalent bounds for a binary DMS and an arbitrary DMC under the Hamming distortion measure.

It is important to study the JSCC excess distortion exponent for the transmission of a continuous alphabet source over a channel with continuous input/output alphabets, since many real-world communication systems deal with the compression and transmission of analog signals. For instance, it is of interest to determine the best performance (in terms of the excess distortion probability) that a source-channel code can achieve if a stationary memoryless Gaussian source (MGS) is coded and transmitted through a stationary memoryless Gaussian channel (MGC), i.e., an additive white Gaussian noise channel. To the best of our knowledge, the JSCC excess distortion exponent for continuous-alphabet systems has not been addressed before. For a general class of stationary memoryless sources, the exponent of excess distortion was recently obtained in [8], where the exponent is expressed in Marton's form [10] (in terms of a minimized Kullback-Leibler divergence). The channel coding error probability exponent, on the other hand, is not yet fully studied even for the Gaussian channel with a quadratic power constraint. For a continuous channel with a transition probability density, Gallager [7, Chapter 7] derives a lower bound based on the random-coding argument subject to a power constraint. He shows that his lower bound, when specialized to the Gaussian channel, is identical to Shannon's classical sphere-packing upper bound [11] for high rates. Some recent works significantly improve Shannon's upper bound (for the Gaussian channel) at low rates (e.g., [3], [4]), but the determination of the error exponent (for all rates below channel capacity) still remains an open problem.

In this work, we study the JSCC exponent for a Gaussian communication system consisting of an MGS P_S with the squared-error distortion and an MGC W with additive noise P_Z and the quadratic power input constraint. We show that the JSCC excess distortion exponent $E_J(P_S,W,\Delta,\mathcal{E},t)$ with transmission rate t, under a distortion threshold Δ and power constraint $\mathcal{E}$, is upper bounded by the minimum of the sum of the Gaussian source excess distortion exponent $tF(R/t,P_S,\Delta)$ and the sphere-packing upper bound of the Gaussian channel error exponent $E_{sp}(R,W,\mathcal{E})$; see Theorem 2. The proof of the upper bound relies on a strong converse JSCC theorem (Theorem 1) and the judicious construction of an auxiliary MGS and an auxiliary MGC to lower bound the probability of excess distortion. We also establish a lower bound for $E_J(P_S,W,\Delta,\mathcal{E},t)$; see Theorem 3. To prove the lower bound, we employ a concatenated "quantization – lossless JSCC" scheme as in [2], use the type covering lemma [6] for the MGS [1], and then bound the probability of error for the lossless JSCC part, which involves a memoryless source with a countably infinite alphabet and the MGC, by using a modified version of Gallager's random-coding bound for the JSCC error exponent for DMS-DMC pairs [7, Problem 5.16] (the modification is made to allow for input power con-

*This work was supported in part by PREA and NSERC of Canada.

strained channels with countably-infinite input alphabets and continuous output alphabets). This lower bound is expressed by the maximum of the difference of Gallager's Gaussian-input channel function $\widetilde{E}_0(W,\mathcal{E},\rho)$ and the MGS guessing exponent $tE(P_S,\Delta,\rho)$ introduced by Arikan and Merhav in [1]. We next derive equivalent expressions (Theorem 4) for our lower and upper bounds by applying Fenchel's Duality Theorem [9] and obtain an explicit condition (Theorem 5) for which the two bounds coincide. Numerical results indicate that the exponent is exactly determined for a large class of source-channel conditions.

II. Preliminaries

All logarithms and exponentials throughout this paper are in natural base. $\mathbb{E}(X)$ denotes the expectation of the random variable (RV) X.

A. MGS and MGC

We consider throughout this paper a communication system consisting of an MGS with alphabet $\mathcal{S}=\mathbb{R}$ and probability density function (pdf) $P_S(s) = (1/\sqrt{2\pi\sigma_S^2})\exp\left\{-\frac{s^2}{2\sigma_S^2}\right\}$, $s\in\mathcal{S}$, denoted by $P_S\sim\mathcal{N}(0,\sigma_S^2)$, and an MGC W with common input, output, and additive noise alphabets $\mathcal{X}=\mathcal{Y}=\mathcal{Z}=\mathbb{R}$ and described by $Y_i=X_i+Z_i$, where Y_i, X_i and Z_i are the channel's output, input and noise symbols at time i. We assume that X_i and Z_i are independent from each other. The transition pdf of the channel is given by

$$W(y|x)=P_Z(z)=\frac{1}{\sqrt{2\pi\sigma_Z^2}}e^{-\frac{z^2}{2\sigma_Z^2}}, \qquad z=y-x\in\mathcal{Z},$$

denoted by $P_Z\sim\mathcal{N}(0,\sigma_Z^2)$ for the noise distribution. We assume the squared-error distortion measure for the source given by $d(s,s')\triangleq(s-s')^2$ for any $s,s'\in\mathbb{R}$ and extended in the usual way for k-tuples as $d^{(k)}(\mathbf{s},\mathbf{s}')\triangleq\frac{1}{k}\sum_{i=1}^{k}(s_i-s_i')^2$ for any $\mathbf{s}\triangleq(s_1,...,s_k)\in\mathbb{R}^k$ and $\mathbf{s}'\triangleq(s_1',...,s_k')\in\mathbb{R}^k$. Given a distortion threshold $\Delta>0$, the rate-distortion function for MGS P_S is given by (e.g., [7])

$$\begin{aligned}R(P_S,\Delta) &= \inf_{P_{S'|S}:\mathbb{E}d(S,S')\le\Delta} I(S;S') \\ &= \begin{cases}\frac{1}{2}\ln\frac{\sigma_S^2}{\Delta}, & 0<\Delta<\sigma_S^2,\\ 0, & \sigma_S^2\le\Delta.\end{cases}\end{aligned}$$

where $I(S;S')$ is the mutual information between the source input and its representation. Given an input cost function $g:\mathcal{X}\to\mathbb{R}^+\triangleq[0,\infty)$ and a constraint $\mathcal{E}>0$, the channel capacity of MGC W is given by

$$C(W,\mathcal{E})\triangleq\sup_{P_X:\mathbb{E}g(X)\le\mathcal{E}} I(X;Y),$$

where $I(X;Y)$ (also denoted by $I(P_X;W)$) is the mutual information between the channel input and channel output. Throughout this paper we assume that $g(x)=x^2$. Under this assumption, the above supremum is achieved by the Gaussian distribution ([7]) and the channel capacity is equal to

$$C(W,\mathcal{E})=\frac{1}{2}\ln\left(1+\mathrm{SNR}\right),$$

where $\mathrm{SNR}\triangleq\mathcal{E}/\sigma_Z^2$ is the signal-to-noise ratio.

B. JSCC Excess Distortion Exponent

A JSC code $(f_n,\varphi_n,\Delta,\mathcal{E},t)$ with blocklength n and transmission rate t (source symbols/channel use) for the MGS P_S with a squared-error distortion measure and a distortion threshold Δ, and the MGC W with $g(x)=x^2$ is a pair of mappings: $f_n:\mathcal{S}^{tn}\longrightarrow\mathcal{X}^n$ and $\varphi_n:\mathcal{Y}^n\longrightarrow\mathcal{S}^{tn}$, where f_n is subject to an (arithmetic average) power constraint: $f_n\in\mathcal{F}_n^{\mathcal{E}}$, where $\mathcal{F}_n^{\mathcal{E}}\triangleq\left\{f_n:\frac{1}{n}\sum_{i=1}^{n}x_i^2\le\mathcal{E}\quad\text{for all }\mathbf{x}=f_n(\mathbf{s})\right\}$. Here $\mathbf{s}\triangleq(s_1,...,s_{tn})\in\mathcal{S}^{tn}$ is the transmitted source message and $\mathbf{x}\triangleq f_n(\mathbf{s})=(x_1,x_2,...,x_n)\in\mathcal{X}^n$ is the corresponding n-length codeword. The conditional pdf of receiving $\mathbf{y}\triangleq(y_1,y_2,...,y_n)\in\mathcal{Y}^n$ at the channel output given that the message $\mathbf{s}$ is transmitted is given by

$$P_{Y^n|X^n}(\mathbf{y}|f_n(\mathbf{s}))=\prod_{i=1}^{n}W(y_i|x_i)=\prod_{i=1}^{n}P_Z(y_i-x_i).$$

The probability of failing to decode the code $(f_n,\varphi_n,\Delta,\mathcal{E},t)$ within a prescribed distortion level Δ is called the probability of excess distortion and defined by

$$\begin{aligned}&P_\Delta^{(n)}(P_S,W,\mathcal{E},t)\\ \triangleq\ &\int_{\mathcal{S}^{tn}}P_{S^{tn}}(\mathbf{s})\int_{\mathbf{y}:d^{(tn)}(\mathbf{s},\varphi_n(\mathbf{y}))>\Delta}P_{Y^n|X^n}(\mathbf{y}|f_n(\mathbf{s}))d\mathbf{y}d\mathbf{s}.\end{aligned}$$

Definition 1: The JSCC excess distortion exponent $E_J(P_S,W,\Delta,\mathcal{E},t)$ for the above MGS P_S and MGC W is defined as the largest number E for which there exists a sequence of source-channel codes $(f_n,\varphi_n,\Delta,\mathcal{E},t)$ with blocklength n and transmission rate t such that

$$E\le\liminf_{n\to\infty}-\frac{1}{n}\ln P_\Delta^{(n)}(P_S,W,\mathcal{E},t).$$

III. Upper and Lower Bounds for E_J

We first derive a strong converse JSCC theorem under the probability of excess distortion criterion for the Gaussian system.

Theorem 1: (Strong Converse JSCC Theorem) For an MGS P_S and an MGC W, if $tR(P_S,\Delta)>C(W,\mathcal{E})$, then $\lim_{n\to\infty}P_\Delta^{(n)}(P_S,W,\mathcal{E},t)=1$ for any sequence of JSC codes $(f_n,\varphi_n,\Delta,\mathcal{E},t)$.

Note that the above theorem also holds for a slightly wider class of MGCs with scaled inputs, described by $Y_i=bX_i+Z_i$ (X_i and Z_i are independent from each other), and with transition pdf

$$W(y|x)=P_Z(y-bx)=\frac{1}{\sqrt{2\pi\sigma_Z^2}}e^{-\frac{(y-bx)^2}{2\sigma_Z^2}},$$

where b is a nonzero constant; we will apply this result to prove the upper bound of $E_J(P_S,W,\Delta,\mathcal{E},t)$. Meanwhile, it directly follows from Theorem 1 that the JSCC excess distortion exponent is 0 if the source rate-distortion function is larger than the channel capacity, i.e., $tR(P_S,\Delta)>C(W,\mathcal{E})$.

We thereby confine our attention to the case of $tR(P_S,\Delta) < C(W,\mathcal{E})$ in the following theorem.

Theorem 2: For an MGS P_S and an MGC W such that $tR(P_S,\Delta) < C(W,\mathcal{E})$, the JSCC excess distortion exponent satisfies

$$E_J(P_S,W,\Delta,\mathcal{E},t) \le \overline{E}_J(P_S,W,\Delta,\mathcal{E},t), \tag{1}$$

where

$$\begin{aligned} & \overline{E}_J(P_S,W,\Delta,\mathcal{E},t) \\ \triangleq \; & \min_{tR(P_S,\Delta)\le R\le C(W,\mathcal{E})} \left[tF\left(\frac{R}{t},P_S,\Delta\right)+E_{sp}(R,W,\mathcal{E})\right], \end{aligned}$$

in which

$$F(R,P_S,\Delta) \triangleq \begin{cases} \frac{1}{2}\left(\frac{\Delta\beta}{{\sigma_S}^2}-\ln\frac{\Delta\beta}{{\sigma_S}^2}-1\right) \\ \text{if} \quad R > R(P_S,\Delta)=\frac{1}{2}\ln\max\{\frac{{\sigma_S}^2}{\Delta},1\}, \\ 0 \quad \text{otherwise} \end{cases}$$

is the excess distortion exponent for an MGS P_S [8] and

$$\begin{aligned} & E_{sp}(R,W,\mathcal{E}) \\ \triangleq \; & \frac{\text{SNR}}{4\beta}\left[(\beta+1)-(\beta-1)\sqrt{1+\frac{4\beta}{\text{SNR}(\beta-1)}}\right] \\ + \; & \frac{1}{2}\ln\left\{\beta-\frac{\text{SNR}(\beta-1)}{2}\left[\sqrt{1+\frac{4\beta}{\text{SNR}(\beta-1)}}-1\right]\right\} \end{aligned}$$

is the sphere-packing bound of the channel error exponent for an MGC W ([7], [11]), where $\beta \triangleq e^{2R}$.

Sketch of Proof: For any sufficiently small $\varepsilon > 0$, fix an $R \in [tR(P_S,\Delta)+\varepsilon, C(W,\mathcal{E})]$. Define an auxiliary MGS for this R with alphabet $\mathcal{S}=\mathbb{R}$ and distribution $\widetilde{P}_S \sim \mathcal{N}(0,\widetilde{\sigma}_S^2)$, where $\widetilde{\sigma}_S^2 \triangleq \Delta e^{2R/t}$, so that the rate-distortion function of $\widetilde{P}_S$ is given by $R(\widetilde{P}_S,\Delta) = R/t$. Also, it can be easily verified that the Kullback-Leibler divergence between the auxiliary MGS $\widetilde{P}_S$ and the original source P_S is $D(\widetilde{P}_S \parallel P_S) = F\left(\frac{R}{t},P_S,\Delta\right)$.

Next we define for $R' \triangleq R-\varepsilon$ an auxiliary MGC with scaled inputs $\widetilde{W}$ associated with the original MGC W with the alphabets $\mathcal{X}=\mathcal{Y}=\mathbb{R}$ and transition pdf

$$\widetilde{P}_{Y|X}(y|x) \triangleq \frac{1}{\sqrt{2\pi\widetilde{\sigma}_Z^2}}e^{-\frac{(y+ax)^2}{2\widetilde{\sigma}_Z^2}}$$

where the parameter a is uniquely determined by β' ($\beta' = e^{2R'}$) and SNR as follows

$$a \triangleq \frac{-\text{SNR}(\beta'-1)-\sqrt{\text{SNR}^2(\beta'-1)^2+4\text{SNR}\beta'}}{2\text{SNR}\beta'} < 0,$$

and $\widetilde{\sigma}_Z^2 \triangleq a^2\mathcal{E}/(\beta'-1)$. It can be verified that the capacity of the MGC $\widetilde{W}$ is given by

$$C(\widetilde{W},\mathcal{E}) = \sup_{P_X:\mathbb{E}x^2\le\mathcal{E}} I(P_X;\widetilde{W}) = \frac{1}{2}\ln\left(1+\frac{a^2\mathcal{E}}{\widetilde{\sigma}_Z^2}\right) = R',$$

where the supremum is achieved by the Gaussian distribution, and under the above choice of a and $\widetilde{\sigma}_Z^2$,

$$\sup_{P_X:\mathbb{E}X^2\le\mathcal{E}} D(\widetilde{W}\parallel W|P_X) = E_{sp}(R',W,\mathcal{E}),$$

where the supremum is achieved by any distribution satisfying $\mathbb{E}X^2=\mathcal{E}$. For some $\delta>0$, define the set

$$\begin{aligned} \widehat{A} \;\triangleq\; & \left\{(\mathbf{s},\mathbf{y}) : \ln\frac{\widetilde{P}_{S^{tn}}(\mathbf{s})\widetilde{P}_{Y^n|X^n}(\mathbf{y}|f_n(\mathbf{s}))}{P_{S^{tn}}(\mathbf{s})P_{Y^n|X^n}(\mathbf{y}|f_n(\mathbf{s}))}\right. \\ & \left.\le n\left(tF\left(\frac{R}{t},P_S,\Delta\right)+E_{sp}(R',W,\mathcal{E})+\delta\right)\right\}. \end{aligned}$$

Consequently, we can use $\widehat{A}$ to lower bound the probability of excess distortion of any sequence of JSC codes $(f_n,\varphi_n,\Delta,\mathcal{E},t)$,

$$\begin{aligned} & P_\Delta^{(n)}(P_S,W,\mathcal{E},t) \\ \ge & \int_{\left\{(\mathbf{s},\mathbf{y}):d^{(tn)}(\mathbf{s},\varphi_n(\mathbf{y}))>\Delta)\right\}\cap\widehat{A}} P_{S^{tn}}(\mathbf{s})P_{Y^n|X^n}(\mathbf{y}|f_n(\mathbf{s}))d\mathbf{s}d\mathbf{y} \\ \ge & e^{-n\left(tF\left(\frac{R}{t},P_S,\Delta\right)+E_{sp}(R',W,\mathcal{E})+\delta\right)} \\ & \int_{\left\{(\mathbf{s},\mathbf{y}):d^{(tn)}(\mathbf{s},\varphi_n(\mathbf{y}))>\Delta)\right\}\cap\widehat{A}} \widetilde{P}_{S^{tn}}(\mathbf{s})\widetilde{P}_{Y^n|X^n}(\mathbf{y}|f_n(\mathbf{s}))d\mathbf{s}d\mathbf{y}, \end{aligned}$$

and the last integration can be decomposed as

$$\begin{aligned} & \int_{\left\{(\mathbf{s},\mathbf{y}):d^{(tn)}(\mathbf{s},\varphi_n(\mathbf{y}))>\Delta)\right\}\cap\widehat{A}} \widetilde{P}_{S^{tn}}(\mathbf{s})\widetilde{P}_{Y^n|X^n}(\mathbf{y}|f_n(\mathbf{s}))d\mathbf{s}d\mathbf{y} \\ \ge & \int_{(\mathbf{s},\mathbf{y}):d^{(tn)}(\mathbf{s},\varphi_n(\mathbf{y}))>\Delta)} \widetilde{P}_{S^{tn}}(\mathbf{s})\widetilde{P}_{Y^n|X^n}(\mathbf{y}|f_n(\mathbf{s}))d\mathbf{s}d\mathbf{y} \\ & \quad - \int_{\widehat{A}^c} \widetilde{P}_{S^{tn}}(\mathbf{s})\widetilde{P}_{Y^n|X^n}(\mathbf{y}|f_n(\mathbf{s}))d\mathbf{s}d\mathbf{y} \\ = & P_\Delta^{(n)}(\widetilde{P}_S,\widetilde{W},t) - P\left(\widehat{A}^c\right), \end{aligned} \tag{2}$$

where the probabilities are with respect to the joint distribution $\widetilde{P}_{S^{tn}}(\cdot)\widetilde{P}_{Y^n|X^n}(\cdot|\cdot)$. Note that the first term in the right-hand side of (2) is exactly the probability of excess distortion for the joint source-channel system consisting of the auxiliary MGS $\widetilde{P}_S$ and the auxiliary MGC $\widetilde{W}$ with transmission t, and, according to our setting, with

$$tR(\widetilde{P}_S,\Delta) = R > R' = C(\widetilde{W},\mathcal{E}).$$

Thus, this quantity converges to 1 as n goes to infinity according to the strong converse JSCC theorem. The proof is then completed by justifying that $P\left(\widehat{A}^c\right) \to 0$ as $n \to 0$. ■

Since the MGS excess distortion exponent $tF(R/t,P_S,\Delta)$ is convex increasing for $R \ge tR(P_S,\Delta)$ and the sphere-packing bound $E_{sp}(R,W,\mathcal{E})$ is convex decreasing in $R \le C(W,\mathcal{E})$, their sum is also convex and there exists a global minimum in the interval $[tR(P_S,\Delta), C(W,\mathcal{E})]$ for the upper bound given in (1).

Theorem 3: For an MGS P_S and an MGC W, the JSCC excess distortion exponent satisfies

$$E_J(P_S,W,\Delta,\mathcal{E},t) \ge \underline{E}_J(P_S,W,\Delta,\mathcal{E},t), \tag{3}$$

where

$$\underline{E}_J(P_S, W, \Delta, \mathcal{E}, t) \triangleq \max_{0\le\rho\le 1} [\widetilde{E}_o(W, \mathcal{E}, \rho) - tE(P_S, \Delta, \rho)],$$

in which

$$\begin{aligned}\widetilde{E}_o(W, \mathcal{E}, \rho) \quad &\triangleq \quad \max_{r\ge 0}\Bigg\{r(1+\rho)\mathcal{E} + \frac{1}{2}\ln(1-2r\mathcal{E}) \\ &\qquad + \frac{\rho}{2}\ln\left[1 - 2r\mathcal{E} + \frac{\mathcal{E}}{(1+\rho)\sigma_Z^2}\right]\Bigg\}\end{aligned}$$

is Gallager's Gaussian-input channel function [7] and

$$E(P_S, \Delta, \rho) = \sup_{\widetilde{P}_S\sim\mathcal{N}(0,\widetilde{\sigma}_S^2)} [\rho R(\widetilde{P}_S, \Delta) - D(\widetilde{P}_S \parallel P_S)]$$

is the guessing exponent [1] of MGS P_S.

Sketch of Proof: Fix $t > 0$. In the sequel we let $k = tn$ and assume that k (and hence n) is sufficiently large. We also let $o(k)$ and $\zeta(\epsilon)$ stand for some terms such that $o(k)/k \to 0$ as $n \to \infty$ and $\zeta(\epsilon) \to 0$ as $\epsilon \to 0$. For a given $\epsilon > 0$ small enough, we partition the whole source space $\mathbb{R}^k$ by a sequence of Gaussian type classes [1] $\mathcal{T}_i$ given by

$$\mathcal{T}_0 \triangleq \{\mathbf{s} : \mathbf{s}^T\mathbf{s} \le k\epsilon\}$$

and

$$\mathcal{T}_i \triangleq \left\{\mathbf{s} : (2i-1)k\epsilon \le \mathbf{s}^T\mathbf{s} \le (2i+1)k\epsilon\right\},$$

for $i = 1, 2, \cdots$. It can be shown that for all types $i = 0, 1, 2, \cdots$, the probability of $\mathcal{T}_i$ under the Gaussian distribution P_S, denoted by $P_S(\mathcal{T}_i)$, decays exponentially at the rate of $D(P_S^{(i)} \parallel P_S) + \zeta(\epsilon)$ in k, where $P_S^{(i)}$ is a zero-mean Gaussian source with variance $\sigma^2(i)$, i.e., $P_S^{(i)} \sim N(0, \sigma^2(i))$, and $\sigma^2(0) = \epsilon$ and $\sigma^2(i) = 2i\epsilon$ for $i \ge 1$. The type covering lemma [1] is applicable for all these types, i.e., for each type $\mathcal{T}_i$ there exists a code $\mathcal{C}_i \in \mathbb{R}^k$ of size $|\mathcal{C}_i| \le \exp\{k[R(P_S^{(i)}, \Delta) + \zeta(\epsilon)] + o(k)\}$ that covers it (in the sense that every sequence $\mathbf{s} \in \mathcal{T}_i$ is contained, for some $\mathbf{c} \in \mathcal{C}_i$, in the sphere $B(\mathbf{c}, \Delta) \triangleq \left\{\mathbf{s} : d^{(k)}(\mathbf{s}, \mathbf{c}) \le \Delta\right\}$ for k sufficiently large). We next employ a concatenated "quantization – lossless JSCC" scheme [2], which is described as follows.

First Coding Stage: Δ-admissible Quantization.

For each type $\mathcal{T}_i$ ($i = 0, 1, 2, \cdots$), we set a rate-distortion encoder $\gamma_k^{(i)} : \mathcal{T}_i \to \mathcal{C}_i$ to encode each source message $\mathbf{s} \in \mathcal{T}_i$ into $\mathbf{c} \triangleq \gamma_k^{(i)}(\mathbf{s})$ such that $d^{(k)}(\mathbf{s}, \mathbf{c}) \le \Delta$, where $\mathcal{C}_i \in \mathbb{R}^k$ is the codebook of $\gamma_k^{(i)}$ (associated with the type $\mathcal{T}_i$). As mentioned before, the type-covering lemma asserts that there exists such an encoder for each $\mathcal{T}_i$ with codebook size $|\mathcal{C}_i| \le \exp\{k[R(P_S^{(i)}, \Delta) + \zeta(\epsilon)] + o(k)\}$. Since the entire Euclidean space $\mathbb{R}^k$ is partitioned by these types $\{\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_2, \cdots\}$, the first stage encoder is regarded as a quantizer on $\mathbb{R}^k$:$f_k^{(1)} : \mathbb{R}^k \longrightarrow \bigcup_{i=0}^{\infty}\mathcal{C}_i$.

Second Coding Stage: Lossless JSCC with Power Constraint.

In the second stage, a lossless joint source-channel encoder $f_n^{(2)} : \bigcup_{i=0}^{\infty}\mathcal{C}_i \longrightarrow \mathcal{X}^n$ is applied on each output of the quantizer $\mathbf{c}$ and we send the codeword $\mathbf{x} \in \mathcal{X}^n$ over the channel. The second stage encoding $f_n^{(2)}$ is subject to a power constraint $f_n^{(2)} \in \mathcal{F}_n^{\mathcal{E}}$, i.e., for each codeword $\mathbf{x} \in \mathbb{R}^n$ (recalling that $\mathcal{X} = \mathbb{R}$ for the MGC), we have $\frac{1}{n}\sum_{j=1}^{n} x_j^2 \le \mathcal{E}$.

Lossless Joint Source-Channel Decoding.

At the channel output, we employ a lossless joint source-channel decoder $\varphi_n : \mathcal{Y}^n \longrightarrow \bigcup_{i=0}^{\infty}\mathcal{C}_i$ to each received codeword $\mathbf{y} \in \mathcal{Y}^n$ and create an approximation $\widehat{\mathbf{c}} = \varphi_n(\mathbf{y})$.

Probability of Excess Distortion.

Under the two-stage coding scheme, the probability of excess distortion Δ can be rewritten as

$$\begin{aligned}&P_\Delta^{(n)}(P_S, W, \mathcal{E}, t) \\ &= \sum_{i=0}^{\infty}\int_{\mathbf{s}\in\mathcal{T}_i} P_{S^k}(\mathbf{s})\int_{\mathbf{y}:d^{(k)}(\mathbf{s},\widehat{\mathbf{c}})>\Delta} P_{Y^n|X^n}(\mathbf{y}|f_n(\mathbf{s}))\,d\mathbf{y}d\mathbf{s},\end{aligned}$$

where f_n here is the composition of $f_k^{(1)}$ and $f_n^{(2)}$, denoted by $f_n = f_n^{(2)} \circ f_k^{(1)}$. With the Δ-admissible quantization, the event that the distortion between the source message $\mathbf{s}$ and the reproduced sequence $\widehat{\mathbf{c}}$ is larger than Δ implies that $\widehat{\mathbf{c}} \ne \mathbf{c}$. Thus, the above probability can be upper bounded by

$$\begin{aligned}&P_\Delta^{(n)}(P_S, W, \mathcal{E}, t) \\ &\le \sum_{i=0}^{\infty}\int_{\mathbf{s}\in\mathcal{T}_i} P_{S^k}(\mathbf{s})\int_{\mathbf{y}:\widehat{\mathbf{c}}\ne\mathbf{c}} P_{Y^n|X^n}(\mathbf{y}|f_n(\mathbf{s}))\,d\mathbf{y}d\mathbf{s} \\ &= \sum_{i=0}^{\infty}\sum_{\mathbf{c}\in\mathcal{C}_i}\underbrace{P_{S^k}(\mathcal{T}_i)P_{S^k}^{(i)}(\mathbf{c})}_{P(\mathbf{c})}\int_{\mathbf{y}:\widehat{\mathbf{c}}\ne\mathbf{c}} P_{Y^n|X^n}(\mathbf{y}|\,\mathbf{c})\,d\mathbf{y}, \quad (4)\end{aligned}$$

where

$$P_{S^k}^{(i)}(\mathbf{c}) \triangleq \frac{1}{P_{S^k}(\mathcal{T}_i)}\int_{\mathbf{s}\in\in\mathcal{T}_i:\gamma_k^{(i)}(\mathbf{s})=\mathbf{c}} P_{S^k}(\mathbf{s})d\mathbf{s}.$$

We note that (4) is exactly the lossless JSCC probability of error for the system consisting of a memoryless source with countable alphabet $\bigcup_{i=0}^{\infty}\mathcal{C}_i$ and distribution $P(\mathbf{c}) = P_{S^k}(\mathcal{T}_i)P_{S^k}^{(i)}(\mathbf{c})$ (the distribution of the output of the Δ-admissible quantizer) and the MGC W with power constraint. Thus, there exists a sequence of lossless JSC codes $\left(f_n^{(2)}, \varphi_n, \mathcal{E}\right)$ such that $P_\Delta^{(n)}(P_S, W, \Delta, \mathcal{E}, t)$ is upper bounded by a modified version of Gallager's JSSC random-coding bound [7, Problem 5.24] for the discrete source $\{P : \bigcup_{i=0}^{\infty}\mathcal{C}_i\}$ and the MGC W. Using the type covering lemma for each type $\mathcal{T}_i$ and the fact that $P_S(\mathcal{T}_i)$ decays exponentially in k at rate $D(P_S^{(i)} \parallel P_S) + \zeta(\epsilon)$, it can be shown that the excess distortion exponent is lower bounded by $\underline{E}_J(P_S, W, \Delta, \mathcal{E}, t)$. Thus, we have demonstrated the existence of a sequence of concatenated JSC codes $\left(f_n^{(2)} \circ f^{(1)}, \varphi_n, \Delta, \mathcal{E}, t\right)$ such that

$$P_\Delta^{(n)}(P_S, W, \mathcal{E}, t) \le \exp\left[-n\underline{E}_J(P_S, W, \Delta, \mathcal{E}, t) + o(n)\right].$$

for n sufficiently large. ■

It can be shown that the lower bound is positive if $tR(P_S, \Delta) < C(W, \mathcal{E})$; it is 0 otherwise. Since the lower bound is expressed by a maximum of a function of ρ over the closed interval $[0, 1]$, the maximum and the maximizing ρ can be numerically obtained.

IV. When does $\overline{E}_J = \underline{E}_J$?

The following theorem illustrates the relation between the upper and lower bounds to E_J.

Theorem 4: Let $tR(P_S, \Delta) < C(W, \mathcal{E})$. Then

$$\min_{tR(P_S,\Delta)\le R\le C(W,\mathcal{E})} \left[tF\left(\frac{R}{t}, P_S, \Delta\right) + E_{sp}(R, W, \mathcal{E}) \right] = \max_{0\le\rho<\infty} [\widetilde{E}_0(W, \mathcal{E}, \rho) - tE(P_S, \Delta, \rho)], \quad (5)$$

$$\min_{tR(P_S,\Delta)\le R\le C(W,\mathcal{E})} \left[tF\left(\frac{R}{t}, P_S, \Delta\right) + E_{\dagger}(R, W, \mathcal{E}) \right] = \max_{0\le\rho<1} [\widetilde{E}_0(W, \mathcal{E}, \rho) - tE(P_S, \Delta, \rho)], \quad (6)$$

where

$$E_{\dagger}(R, W, \mathcal{E}) = \max_{0\le\rho\le 1} [-\rho R + \widetilde{E}_o(W, \mathcal{E}, \rho)] = \begin{cases} E_{sp}(R, W, \mathcal{E}), \\ \quad R_{cr}(W) \le R \le C(W, \mathcal{E}), \\ 1 - \beta + \frac{\text{SNR}}{2} + \frac{1}{2}\ln\left(\beta - \frac{\text{SNR}}{2}\right) + \frac{1}{2}\ln\beta - R, \\ \quad 0 \le R \le R_{cr}(W), \end{cases}$$

is convex strictly decreasing in $0 < R \le C(W, \mathcal{E})$ with a straight-line section of slope -1 for $R \le R_{cr}(W)$, where

$$R_{cr}(W) \triangleq \frac{1}{2}\ln\left[\frac{1}{2} + \frac{\text{SNR}}{4} + \frac{1}{2}\sqrt{1 + \frac{\text{SNR}^2}{4}}\right]$$

is the critical rate of the MGC.

The proof follows from the Fenchel's Duality Theorem and a convexity argument as in [12]. Indeed we observe that the upper bound, though proved in the form of a minimum of the sum of source and channel exponents, can also be represented as a (dual) maximum of the difference of Gallager's channel function and the source guessing exponent. Symmetrically, the lower bound, which is established in Gallager's form, can also be represented in Csiszár's form, as the minimum of the sum of the source exponent and the lower bound of the channel exponent. In this regard, our result is a natural extension of Csiszár's upper and lower bounds from the case of (finite alphabet) discrete memoryless systems to the case of memoryless Gaussian systems.

Setting STR $\triangleq \sigma_S^2/\Delta$ to be the source-to-threshold ratio (i.e., the source variance to distortion threshold ratio), then the upper bound $\overline{E}_J(P_S, W, \Delta, \mathcal{E}, t)$ and the lower bound $\underline{E}_J(P_S, W, \Delta, \mathcal{E}, t)$ are only functions of STR and SNR. We then compare the upper and lower bounds using their equivalent forms and derive an explicit necessary and sufficient analytical condition for which the two bounds coincide.

Theorem 5: Let $tR(P_S, \Delta) < C(W, \mathcal{E})$. The upper and lower bounds for $E_J(P_S, W, \Delta, \mathcal{E}, t)$ given in Theorems 2 and 3 are equal iff

$$2(2\text{STR})^t - \frac{2(2\text{STR})^t}{2(2\text{STR})^t - 1} \ge \text{SNR}. \quad (7)$$

Remark: For $tR(P_S, \Delta) \ge C(W, \mathcal{E})$, $E_J(P_S, W, \Delta, \mathcal{E}, t) = 0$.

In Fig. 1 we plot the two bounds according to the formulas given in Theorems 2 and 3 for different STR-SNR pairs and transmission rate $t = 1$ (in the figure, the SNR and the STR are expressed in dB). We note that the two bounds coincide for a large class of (STR, SNR) pairs. For example, when STR $= 9$ dB, the two bounds are observed to coincide for SNR ≤ 14.88 dB, as predicted by (7).

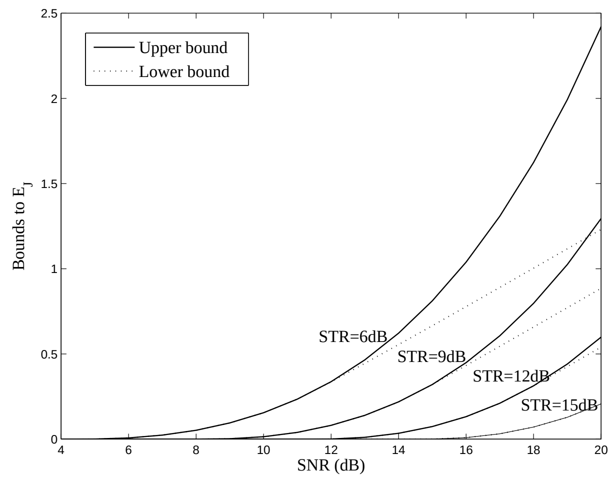

Fig. 1. The upper and lower bounds for $E_J(P_S, W, \Delta, \mathcal{E}, t)$ with $t = 1$.

References

[1] E. Arikan and N. Merhav, "Guessing subject to distortion," *IEEE Trans. Inform. Theory*, vol. 44, no. 3, pp. 1041–1056, May 1998.

[2] E. Arikan and N. Merhav, "Joint source-channel coding and guessing with application to sequential decoding," *IEEE Trans. Inform. Theory*, vol. 44, no. 5, pp. 1756–1769, Sep. 1998.

[3] A. E. Ashikhmin, A. Barg, and S. N. Litsyn, "A new upper bound on the reliability function of the Gaussian channel,"*IEEE Trans. Inform. Theory*, vol. 46, no. 6, pp. 1945–1961, Sep. 2000.

[4] A. Barg and A. McGregor, "Distance distribution of binary codes and the error probability of decoding," *IEEE Trans. Inform. Theory*, vol. 51, pp. 4237-4246, Dec. 2005.

[5] I. Csisz´ar, "On the error exponent of source-channel transmission with a distortion threshold," *IEEE Trans. Inform. Theory*, vol. 28, pp. 823–828, Nov. 1982.

[6] I. Csisz´ar and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic, 1981.

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

[8] S. Ihara and M. Kubo, "Error exponent of coding for stationary memoryless sources with a fidelity criterion," *IEICE Tran. Fundamentals*, vol. E88-A, no. 5, pp 1339-1345, May. 2005.

[9] D. G. Luenberger, *Optimization by Vector Space Methods*, Wiley, 1969.

[10] K. Marton, "Error exponent for source coding with a fidelity criterion," *IEEE Trans. Inform. Theory*, vol. 20, pp. 197–199, Mar. 1974.

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

[12] Y. Zhong, F. Alajaji, and L. L. Campbell, "On the joint source-channel coding error exponent for discrete memoryless systems," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1450–1468, April 2006.