Error Bounds for a Mixed Entropy Inequality

James Melbourne, Saurav Talukdar, Shreyas Bhaban and Murti Salapaka

Electrical and Computer Engineering,
University of Minnesota

May 30, 2018

Abstract

In this article, we investigate the entropy of a sum of a discrete and a continuous random variable. Bounds for estimating the entropy of the sum are obtained for the cases when the continuous random variable is Gaussian or log-concave. Bounds on the capacity of a channel where the discrete random variable is the input and the output is the input corrupted by additive noise modeled by the continuous random variable are obtained. The bounds are shown to be sharp in the case that the discrete variable is Bernoulli.

1 Introduction

We consider the entropy, $h(X+Z)$, of a discrete random random variable X taking values in $\mathbb{R}$ corrupted by continuous additive noise Z. Two models of the additive noise, Z, are investigated; the first assumes Z to be Gaussian and the other generalizes the study to log-concave descriptions. The estimates on the entropy, $h(X+Z)$, of the sum, $X+Z$, are employed to estimate the capacity of a channel with X as the input and $X+Z$ as the output. The special case of a Gaussian Z and Bernoulli X was relevant to the modeling of an experimental testing of Landauer's bound [11, 12], and will be the focus of section 3. In Section 4 we will approach the general log-concave case. The following section will introduce notations, background definitions, and some preliminary results.

2 Background and Preliminaries

Here we briefly summarize the notion of information entropy for discrete and continuous random variables. The reader can consult [4] for general background on information theory, and [8] for recent developments in entropic inequalities.

Definition 2.1. *For a random variable X taking discrete values $\{x_i\}_{i=1}^\infty \subset \mathbb{R}$ with the probability mass function, $P(X=x_i)=p_i$, we denote the Shannon entropy in "nats" as,*

$$H(Z) = -\sum_{i=1}^{\infty} p_i \ln p_i. \tag{1}$$

For a random variable Z with density $f_Z(z)$ on $\mathbb{R}$, whenever $f_Z \ln f_Z \in L^1$, the entropy is given by,

$$h(Z) = -\int_{-\infty}^{\infty} f_Z(z) \ln f_Z(z) dz. \tag{2}$$

In this article $\sum_{i=-\infty}^{\infty} a_i$ will be denoted as $\sum_i a_i$, $\sum_{i=-\infty, i\neq 0}^{\infty} b_i$ as $\sum_{i\neq 0} b_i$. We will suppress notation at times when the meaning of expressions is clear from context. We utilize $P(A)$ to denote the probability of an event A. We will use the notation $Y \sim f$ to indicate when Y is a continuous random variable that it has a density function f, and when Y is discrete to indicate that it has probability mass function f. For clarity we will also employ a subscript, and write f_Y for the density function or probability mass function of Y.

First a comment on the nature of $X + Z$ where $X \sim p$ and $Z \sim f$, with X and Z independent. Let $X + Z \sim f_{X+Z}$. Consider any Borel set A. Then

$$\begin{aligned} P(X + Z \in A) &= \textstyle\sum_k P(X = x_k, z \in -x_k + A) \\ &= \textstyle\sum_k P(X = x_k) P(z \in -x_k + A) \\ &= \textstyle\sum_k P(X = x_k) \int_{z\in -x_k + A} f_Z(z) dz \\ &= \textstyle\sum_k p_k \int_{y\in A} f_Z(y - x_k) dy \end{aligned}$$

Thus,

$$f_{X+Z}(y) = \sum_k p_k f_Z(y - x_k).$$

Proposition 2.2. *Consider a discrete $X \sim p$ and $Z \sim f_Z$, and $f_Z \log f_Z \in L^1(\mathbb{R})$, where X and Z are independent. Define*

$$\delta(X, Z) := H(X) + h(Z) - h(X + Z). \tag{3}$$

Then

$$\delta(X, Z) = \sum_k p_k \int f_Z(z - x_k) \ln \left(1 + \frac{\sum_{j\neq k} p_j f_Z(z - x_j)}{p_k f_Z(z - x_k)}\right) dz, \tag{4}$$

and satisfies

$$\delta(X, Z) \geq 0. \tag{5}$$

Proof. Note that

$$\begin{aligned} -h(X + Z) &= \textstyle\int f_{X+Z}(y) \ln f_{X+Z}(y) dy \\ &= \textstyle\int \sum_k p_k f_Z(y - x_k) \ln \left(\sum_k p_k f_Z(y - x_k)\right) dy \end{aligned}$$

$$
\begin{aligned}
-h(X+Z) &= \int f_{X+Z}(y)\ln f_{X+Z}(y)dy, \\
&= \int \sum_k p_k f_Z(y-x_k)\ln\left(\sum_j p_j f_Z(y-x_j)\right)dy, \\
&= \int \sum_k p_k f_Z(y-x_k)\ln\left(p_k f_Z(y-x_k)(1+\frac{\sum_{j\neq k} p_j f_Z(y-x_j)}{p_k f_Z(y-x_k)}\right))dy, \\
&= \sum_k p_k \ln p_k \int f_Z(y-x_k)dy + \sum_k p_k \int f_Z(y-x_k)\ln f_Z(y-x_k)dy + \\
&\sum_k p_k \int f_Z(y-x_k)\left(1+\frac{\sum_{j\neq k} p_j f_Z(y-x_j)}{p_k f_Z(y-x_k)}\right)dy, \\
&= -H(X) - h(Z) + \sum_k p_k \int f_Z(y-x_k)\left(1+\frac{\sum_{j\neq k} p_j f_Z(y-x_j)}{p_k f_Z(y-x_k)}\right)dy,
\end{aligned}
\tag{6}
$$

Thus (4) is established. (5) follows immediately from (3). □

Remark 2.3. *The inequality in the above Proposition is trivially sharp when one takes X to be a point mass. A more interesting case of equality is when Z is supported in the interval $(-\frac{1}{2}, \frac{1}{2})$, while X is supported on the integers.*

Lemma 2.4. *Consider a discrete $X \sim p$ and $Z \sim f$ both of bounded entropy and independent, then*

$$
\delta(X,Z) \leq \int f_Z(y)\ln\left(1+\frac{\sum_j p_j \sum_{k\neq j} f_Z(y+x_k-x_j)}{f(y)}\right)dy. \tag{7}
$$

Proof. We begin with the substitution $y = z - x_k$, in (4) resulting in,

$$
\begin{aligned}
\delta(X,Z) &= \sum_k p_k \int f_Z(y)\ln\left(1+\frac{\sum_{j\neq k} p_j f_Z(y+x_k-x_j)}{p_k f_Z(y)}\right)dy \\
&= \int f_Z(y)\left(\sum_k p_k \ln\left(1+\frac{\sum_{j\neq k} p_j f_Z(y+x_k-x_j)}{p_k f_Z(y)}\right)\right)dy
\end{aligned}
$$

It follows from Jensen's inequality and the concavity of logarithm that for every $y \in \mathbb{R}$,

$$
\begin{aligned}
&\sum_k p_k \ln\left(1+\frac{\sum_{j\neq k} p_j f_Z(y+x_k-x_j)}{p_k f_Z(y)}\right) \\
&\leq \ln\left(1+\sum_k p_k \frac{\sum_{j\neq k} p_j f_Z(y+x_k-x_j)}{p_k f_Z(y)}\right).
\end{aligned}
$$

Thus,

$$\delta(X, Z) \leq \int f_Z(y) \ln \left(1 + \frac{\sum_k \sum_{j \neq k} p_j f_Z(y + x_k - x_j)}{f_Z(y)}\right) dy.$$

Changing the order of summation leads to,

$$\int f_Z(y) \ln \left(1 + \frac{\sum_j p_j \sum_{k \neq j} f_Z(y + x_k - x_j)}{f_Z(y)}\right) dy.$$

This completes the proof.

□

2.1 Log-concave random variables

A function on $\mathbb{R}$, satisfying $f((1-t)x + ty) \geq f^{1-t}(x) f^t(y)$ for any $x, y \in \mathbb{R}$ and $t \in (0, 1)$ is called log-concave, as $\log f$ is a concave function. Equivalently, f can be expressed in the form e^{-V}, for some convex function V taking values in $(-\infty, \infty]$. A random variable will be called log-concave when it possess a density f, that is log-concave. Important examples of log-concave random variables, are the Gaussian, the Laplace and exponential distributions, uniform distributions on convex set. Log-concave distributions arise naturally in probability, statistics, and convex geometry. Recent research has shown that this class of measures has interesting interactions with information theoretic inequalities as well, see for example [1, 3, 5–7, 9, 13].

3 A single bit corrupted by Gaussian noise

In this section we analyze the case when Z is a continuous random variable with standard deviation σ and mean zero with X being a binary random variable taking values in the set $\{-\mu, \mu\}$.

We assume that X takes the value μ with probability $1 - p$ and $-\mu$ with probability p. Here,

$$f_Z(z) = \frac{e^{-z^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}$$

and $\delta(X, Z)$ is

$$\begin{aligned} p \int f_Z(z + \mu) \ln \left(1 + \frac{(1-p) f_Z(z - \mu)}{p f_Z(z + \mu)}\right) dz \\ + (1-p) \int f_Z(z - \mu) \ln \left(1 + \frac{p f_Z(z + \mu)}{(1-p) f_Z(z - \mu)}\right) dz. \end{aligned}$$

We first recall previous a result.

Theorem 3.1. *[10] If X is an integer valued random variable and Z is an independent centered Gaussian with standard deviation $\sigma < \frac{1}{2}$, then,*

$$\delta(X, Z) \leq \frac{e^{-1/8\sigma^2}}{\sqrt{2\pi}} \left(\frac{1}{\sigma} + 8\right).$$

In short, this shows that the deficit for Gaussians is controlled by Gaussian decay. In what follows we will show that the tails of log-concave Z control the deficit δ. We proceed by obtaining deficit bounds in terms of tails in the Gaussian case.

3.1 Bounds on the deficit

This subsection is dedicated to the proof of the following,

Theorem 3.2. *When X is distributed on $\pm\mu$ as above, and Z is an independent Gaussian random variable of standard deviation σ,*

$$H(X)\mathbb{P}(Z>\mu) \le \delta(X,Z) \le (4+2H(X))\mathbb{P}(Z>\mu). \tag{8}$$

In particular, for a fixed non-trivial X, $\delta(X,Z)$ is proportional to $\mathbb{P}(Z>\mu)$.

Proof of the upper bound. Note that

$$\begin{aligned}
p\int_0^\infty f_Z(z+\mu)\ln\left(1+\frac{(1-p)f_Z(z-\mu)}{pf_Z(z+\mu)}\right)dz &= -p\ln p\int_0^\infty f_Z(z+\mu)dz \\
&\quad + p\int_0^\infty f_Z(z+\mu)\ln\left(p+\frac{(1-p)f_Z(z-\mu)}{f_Z(z+\mu)}\right)dz \\
&= -p\ln p\ \mathbb{P}(Z>\mu) \\
&\quad + p\int_0^\infty f_Z(z+\mu)\ln\left(p+(1-p)e^{2z\mu/\sigma^2}\right)dz
\end{aligned}$$

Using the bound $px+(1-p)y \le \max\{x,y\}$ we have,

$$\begin{aligned}
&\int_0^\infty f_Z(z+\mu)\ln\left(p+(1-p)e^{2z\mu/\sigma^2}\right)dz \\
&\qquad \le \int_0^\infty f_Z(z+\mu)\frac{2z\mu}{\sigma^2}dz \\
&\qquad = \int_0^\infty f_Z(z+\mu)\frac{2\mu(z+\mu)-2\mu^2}{\sigma^2}dz \\
&\qquad = 2\mu\int_0^\infty f_Z(z+\mu)\frac{z+\mu}{\sigma^2}dz - 2\frac{\mu^2}{\sigma^2}\mathbb{P}(Z>\mu).
\end{aligned}$$

Noting that f_Z is Gaussian that satisfies $\frac{d}{dz}f_Z(z+\mu) = -f_Z(z+\mu)\frac{(z+\mu)}{\sigma^2}$ we have

$$\int_0^\infty f_Z(z+\mu)\frac{z+\mu}{\sigma^2}dz = -\int_0^\infty \frac{d}{dz}f_Z(z+\mu)dz = f_Z(\mu).$$

Thus,

$$\int_0^\infty f_Z(z+\mu)\ln\left(p+(1-p)e^{2x\mu/\sigma^2}\right)dz \le 2\mu f_Z(\mu) - \frac{2\mu^2\mathbb{P}(Z>\mu)}{\sigma^2}.$$

It follows that

$$\begin{aligned}
&p\int_0^\infty f_Z(z+\mu)\ln\left(1+\frac{(1-p)f_Z(z-\mu)}{pf_Z(z+\mu)}\right)dz \\
&\qquad \le -\mathbb{P}(Z>\mu)p\ln p + p\left(2\mu f_Z(\mu) - \frac{2\mu^2\mathbb{P}(Z>\mu)}{\sigma^2}\right)
\end{aligned}$$

Writing $f_Z(z) = -\frac{\sigma^2 f_Z'(z)}{z}$ and then integrating by parts we have

$$\begin{aligned}
\mathbb{P}(Z > \mu) &= -\int_\mu^\infty \frac{\sigma^2 f'(z)}{z} dx \\
&= \frac{\sigma^2}{\mu} f(\mu) - \int_\mu^\infty \left(\frac{\sigma}{z}\right)^2 f(z) dz.
\end{aligned}$$

Thus,

$$\begin{aligned}
\left(2\mu f_Z(\mu) - \frac{2\mu^2 \mathbb{P}(Z > \mu)}{\sigma^2}\right) &= \frac{2\mu^2}{\sigma^2} \int_\mu^\infty \left(\frac{\sigma^2}{z}\right)^2 f_Z(z) dz \\
&\leq 2\mathbb{P}(Z > \mu).
\end{aligned}$$

$$\mathbb{P}(Z > \mu) = \sigma^2 f_Z(\mu)/\mu - \int_\mu^\infty \frac{\sigma^2}{z^2}$$

Compiling these computations, we have

$$p \int_0^\infty f_Z(z+\mu) \ln\left(1 + \frac{(1-p) f_Z(z-\mu)}{p f_Z(z+\mu)}\right) dz \leq (-p \ln p + 2p)\mathbb{P}(Z > \mu). \quad (9)$$

Observe that $\ln(1+z) \leq z$, implies

$$\begin{aligned}
p \int_{-2\mu}^0 f_Z(z+\mu) \ln\left(1 + \frac{(1-p) f_Z(z-\mu)}{p f_Z(z+\mu)}\right) dz &\leq (1-p) \int_{-2\mu}^0 f_Z(z-\mu) dz \\
&= (1-p)\mathbb{P}(Z \in (\mu, 3\mu)) \\
&\leq (1-p)\mathbb{P}(Z > \mu).
\end{aligned}$$

Finally, since $f_Z(z-\mu)/f_Z(z+\mu) \leq 1$ for $z \leq -\mu$

$$p \int_{-\infty}^{-2\mu} f_Z(z+\mu) \ln\left(1 + \frac{(1-p) f_Z(z-\mu)}{p f_Z(z+\mu)}\right) dz \leq -p \ln p \mathbb{P}(Z > \mu)$$

Compiling all of the above inequalities we obtain,

$$\begin{aligned}
& p \int f_Z(z+\mu) \ln\left(1 + \frac{(1-p) f_Z(z-\mu)}{p f_Z(z+\mu)}\right) dz \\
& \leq (1 + 2p - 2p \ln p)\mathbb{P}(Z > \mu)
\end{aligned}$$

Using similar arguments it follows that,

$$\begin{aligned}
& (1-p) \int f_Z(z-\mu) \ln\left(1 + \frac{p f_Z(z+\mu)}{(1-p) f_Z(z-\mu)}\right) dz \\
& (1 + 2(1-p) - 2(1-p)\ln(1-p))\mathbb{P}(Z > \mu)
\end{aligned}$$

Combining these results we have

$$\begin{aligned}
\delta(X, Z) &\leq (4 + 2H(X))\mathbb{P}(Z > \mu) \\
&\leq 5.39\, \mathbb{P}(Z > \mu).
\end{aligned}$$

Proof of the lower bound: Note that

$$\begin{aligned}
&p\int f_Z(z+\mu)\ln\left(1+\frac{(1-p)f_Z(z-\mu)}{pf_Z(z+\mu)}\right)dz\\
&=p\int f_Z(z+\mu)\ln\left(1+\frac{(1-p)}{p}e^{\frac{2z\mu}{\sigma^2}}\right)dz\\
&\geq p\int_0^\infty f_Z(z+\mu)\ln(1+\frac{(1-p)}{p})dz\\
&=-p\ln p\mathbb{P}(Z>\mu).
\end{aligned}$$

Analogously,

$$\begin{aligned}
&(1-p)\int f_Z(z-\mu)\ln\left(1+\frac{pf_Z(z+\mu)}{(1-p)f_Z(z-\mu)}\right)dz\\
&\geq -(1-p)\ln(1-p)\mathbb{P}(Z>\mu).
\end{aligned}$$

Adding the two inequalities delivers our result.

□

3.2 Bounds on Capacity of a Channel with discrete input with additive Gaussian noise

Here we consider a channel in which messages are modeled via X distributed on $\pm\mu$, and the receiver reads the transmitted message $X+Z$ after corruption by an additive Gaussian noise. We determine upper and lower bounds on the capacity $\mathcal{C}$ of the channel, defined as

$$\max_{p(X)} I(X;X+Z). \tag{10}$$

Theorem 3.3. *Consider a channel with input $X\in\{-\mu,\mu\}$ and output $X+Z$ where Z is zero mean Gaussian random variable. Then the capacity $\mathcal{C}$ admits the following bounds:*

$$\ln 2\mathbb{P}(Z>\mu)\leq \ln 2-\mathcal{C}\leq (3+2\ln 2)\mathbb{P}(Z>\mu).$$

Proof. By independence, the mutual information for the input and output of an additive channel depends only on $h(X+Z)$, since $I(X;X+Z)=h(X+Z)-h(X+Z|X)=h(X+Z)-h(Z)$. By our definition of $\delta(X,Z)$, we can write

$$I(X;X+Z)=H(X)-\delta(X,Z).$$

Thus invoking Theorem 3.2, we have

$$H(X)-H(X)\mathbb{P}(Z>\mu)\geq I(X;X+Z)\geq H(X)-(3+2H(X))\mathbb{P}(Z>\mu)$$

Note that as $\mu>0$, $\mathbb{P}(Z>\mu)<1/2$. Thus it follows that

$$\begin{aligned}
\mathcal{C}=\max_{p(X)} I(X;X+Z)&\geq \max_{p(X)}[H(X)-(3+2H(X))\mathbb{P}(Z>\mu)]\\
&=\max_{p(X)}[H(X)(1-2\mathbb{P}(Z>\mu))]-3\mathbb{P}(Z>\mu)\\
&=\ln 2(1-2\mathbb{P}(Z>\mu))]-3\mathbb{P}(Z>\mu)\\
&=\ln 2-(3+2\ln 2)\mathbb{P}(Z>\mu).
\end{aligned}$$

Also

$$
\begin{aligned}
I(X;X+Z) &\leq H(X) - H(X)\mathbb{P}(Z>\mu) \\
&= H(X)\mathbb{P}(Z\leq\mu) \\
&\leq \max_{p(X)} H(X)\mathbb{P}(Z\leq\mu) \\
&= \ln 2(1-\mathbb{P}(Z>\mu)).
\end{aligned}
$$

Thus

$$
\mathcal{C} = \max_{p(X)} I(X;X+Z) \leq \ln 2(1-\mathbb{P}(Z>\mu))
$$

Using the above inequalities we have,

$$
\ln 2\mathbb{P}(Z>\mu) \leq \ln 2 - \mathcal{C} \leq (3+2\ln 2)\mathbb{P}(Z>\mu).
$$

□

4 Discrete input corrupted by log-concave noise

In this section we consider the entropy, $h(X+Z)$, where X is a random variable that takes values in the set $x = \{x_i\}_{i\in\mathcal{I}}$ satisfying

$$
\mu := \mu_x := \inf_{i\neq j} \frac{|x_i - x_j|}{2} > 0. \tag{11}
$$

and Z is a symmetric log-concave continuous values random variable. We obtain upper bounds deficit $\delta(X,Z) := H(X) + h(Z) - h(X+Z)$, under the condition that $\varepsilon\|f_Z\|_\infty \geq \frac{1}{2\mu}$, for some $\varepsilon > 0$, and demonstrate the sharpness of the bounds by developing complementary lower bounds on δ in the case that X takes two values. We leverage these results toward bounds on the capacity of a channel with input modeled via X and output via $X+Z$.

Let us remark that the assumption $\varepsilon\|f_Z\|_\infty \geq \frac{1}{2\mu}$, though opaque, is of the same spirit as [10] where it was assumed that the spacing of the discrete random variable was compared to the standard deviation of the continuous (Gaussian) random variable, as for log-concave random variables, up to universal constants $\|f_Z\|_\infty^{-1}$ and σ are equivalent [2]. Some assumption on the spacing of X relative to the concentration of Z is necessary, as $\delta(X,Z)$ can generally be arbitrarily large. Indeed, take Z to be a uniform distribution on $(0,1)$ and X to be uniform on the the n points $\{0, \frac{1}{n}, \dots, \frac{n-1}{n}\}$, so that $\mu = \frac{1}{n}$. Then $h(Z) = 0$, $H(X) = \ln n = \ln\frac{1}{\mu}$, while $h(X+Z) \leq \ln 2$ since the support of $X+Z$ is contained in $(0,2)$. Thus

$$
\delta(X,Z) \geq \ln\frac{1}{2\mu}.
$$

Thus, the respective term in (16) below cannot be done away with in general. We first establish the following lemma for symmetric unimodal distributions.

Lemma 4.1. *For a symmetric unimodal density function f, and a sequence $\{x_n\}_n$ such that $|x_i - x_j| \geq 2\mu > 0$ for $i\neq j$,*

$$
\sum_n f(x_n) \leq 2\|f\|_\infty + \frac{1}{2\mu}.
$$

Proof. Observe that f being unimodal and symmetric takes it's unique maximum at 0. Let $I_i^+ := (x_i - 2\mu, x_i)$. Let $x_i \geq 2\mu$.

$$\int_{I_i} f(x)dx \geq 2\mu f(x_i)$$

which holds as $f(x) \geq f(x_i)$ for all $0 \leq x \leq x_i$. Similarly, let $I_j^- := (x_j, x_j + 2\mu)$ where $x_j \leq -2\mu$. Then

$$\int_{I_j} f(x)dx \geq 2\mu f(x_j),$$

which follows as $f(x) \geq f(x_j)$ for $0 > x > x_j$. Notice that by the inequality $|x_i - x_j| \geq 2\mu > 0$, the intervals I_j^- and I_i^+ are disjoint. Thus, by identifying $I_i = I_i^+$ if $x_i \geq 2\mu$ and $I_i = I_i^-$ if $x_i \leq -2\mu$ we have

$$\begin{aligned}
\sum_{i:|x_i|\geq 2\mu} f(x_i) &= \sum_{i:x_i \geq 2\mu} f(x_i) + \sum_{j:x_j \leq -2\mu} f(x_i) \\
&\leq \frac{1}{2\mu} \sum_{|x_i|\geq 2\mu} \left(\int_{I_i} f(x)dx \right) \\
&= \frac{1}{2\mu} \int_{\mathbb{R}} \left(\sum_{|x_i|\geq 2\mu} \mathbb{1}_{I_i} \right) f dx \\
&\leq \frac{1}{2\mu} \int_{\mathbb{R}} f dx \\
&= \frac{1}{2\mu}.
\end{aligned}$$

By the definition of μ, at most two elements of the sequence $\{x_i\}$ belong to the set $(-2\mu, 2\mu)$; both satisfy $f(x_i) \leq \|f\|_\infty$. Thus

$$\sum_n^\infty f(x_n) < 2\|f\|_\infty + \frac{1}{2\mu}$$

The result follows. $\square$

We will establish the following upper bound through a series of intermediate results.

Theorem 4.1. *Let X be a discrete valued random variable taking values in $x = \{x_i\}_i$ where $P(X = x_i) = p_i$. Furthermore, let $\mu = \mu_x = \min_{i\neq j} \frac{|x_i - x_j|}{2} > 0$. Let Z be a symmetric log-concave real-valued random variable with $Z \sim f_Z$ with Z independent of X. Then, given $\lambda \in (0,1)$ there exists an $\tilde{M} = \tilde{M}(\lambda, X, Z)$ such that*

$$\delta(X, Z) \leq \tilde{M}\, \mathbb{P}(Z > \lambda\, \mu_X), \tag{12}$$

where

$$\tilde{M}(\lambda, X, Z) = \tilde{M} := \left(2 + 2\ln\left(3\|f_Z\|_\infty + \frac{1}{2\mu_x} \right) + \frac{4}{e(1-\lambda)\lambda} \|f_Z\|_\infty^{\lambda-1} \right) \tag{13}$$

and

Further, when $\varepsilon \|f_Z\|_\infty \geq \frac{1}{2\mu}$, *we can sharpen to*

$$\delta(X,Z) \leq M\,\mathbb{P}(Z > \lambda\,\mu_X),$$

$$M = M(\lambda,\varepsilon) = 2 + 2\ln(3+\varepsilon) + 2\frac{\ln(2/\lambda)}{1-\lambda} \tag{14}$$

Lemma 4.2. *For a symmetric log-concave random variable* Z *with density* f_Z*, and a sequence* $\{x_i\}_{i=1}^\infty$ *in* $\mathbb{R}$ *such that* $|x_k - x_j| \geq 2\mu$ *for* $k \neq j$*, the following holds:*

$$\int_{-\mu}^{\mu} f_Z(y) \ln\left(1 + \frac{\sum_j p_j \sum_{k\neq j} f_Z(y + x_k - x_j)}{f_Z(y)}\right) dy \\ \leq 2\int_{\mu}^{\infty} f_Z(y)dy.$$

Proof. Using the inequality $\ln(1+x) \leq x$ for $x \geq 0$, we have,

$$\int_{-\mu}^{\mu} f_Z(y) \ln\left(1 + \frac{\sum_j p_j \sum_{k\neq j} f_Z(y + x_k - x_j)}{f_Z(y)}\right) dy \\ \leq \int_{-\mu}^{\mu} \sum_j p_j \sum_{k\neq j} f_Z(y + x_k - x_j)dy.$$

After exchanging the summation and integral, and changing variables the right hand side is,

$$\sum_j p_j \sum_{k\neq j} \int_{x_k - x_j - \mu}^{x_k - x_j + \mu} f_Z(y)dy = \sum_j p_j \int_{-\infty}^{\infty} f_Z(y) \sum_{k\neq j} \mathbb{1}_{I_{k,j}}(y)dy, \tag{15}$$

where $I_{k,j} := (x_k - x_j - \mu, x_k - x_j + \mu)$ denotes the interval of radius μ centered at $x_k - x_j$ and $\mathbb{1}_{I_S}$ denotes the indicator function where $\mathbb{1}_S(s) = 1$ if $s \in S$ else it is zero. Notice that by the assumed separation of the sequence for a fixed j, the intervals $I_{k,j}$ are disjoint and since k sums only over $k \neq j$,

$$\bigcup_k I_{k,j} \subseteq \{|x| \geq \mu\}.$$

It follows from this and symmetry that

$$\int_{-\infty}^{\infty} f_Z(y) \sum_{k\neq j} \mathbb{1}_{I_{k,j}}(y)dy \leq 2\int_{\mu}^{\infty} f_Z(y)dy.$$

Taking a weighted sum of the above inequality over j with weights p_j (see (15) we have our result. □

We would like to exert similar control over the same integrand, when integrated over the domain $\{|y| \geq \mu\}$. To this end we present the following lemma.

Lemma 4.3. *Consider a symmetric log-concave random variable Z with density $f_Z(z) = e^{-V(z)}$. Given $\varepsilon > 0$, for $\lambda \in (0,1)$, there exists a constant $C(\lambda) > 0$ such that,*

$$-\mathbb{E}\left[\ln f_Z(Z)\mathbb{1}_{\{|Z|>\varepsilon\}}\right] \leq C(\lambda)\|f_Z\|_\infty^{\lambda-1}\mathbb{P}(Z > \lambda\varepsilon).$$

Explicitly, one can take $C(\lambda) = \frac{2}{e(1-\lambda)\lambda}$.

Proof. Note that $xe^{-x} \leq \frac{e^{-\lambda x}}{e(1-\lambda)}$ for $x \geq 0$. It follows that,

$$\begin{aligned}
-\mathbb{E}[\ln f_Z(z)\mathbb{1}_{\{|Z|>\varepsilon\}}] &= 2\int_\varepsilon^\infty V(z)e^{-V(z)}dz \\
&\leq \frac{2}{e(1-\lambda)}\int_\varepsilon^\infty e^{-\lambda V(z)}dz \\
&= \frac{2}{e(1-\lambda)}\int_\varepsilon^\infty f_Z^\lambda(z)dz,
\end{aligned}$$

where we have also tacitly used $V(z) = V(-z)$. By log-concavity (where for every $0 < t < 1$, x and y, $f(tx + (1-t)y) \geq f^t(x)f^{1-t}(y)$) it follows that $f_Z^\lambda(z) \leq f_Z(\lambda z)f_Z^{\lambda-1}(0)$. Thus,

$$\begin{aligned}
\int_\varepsilon^\infty f_Z^\lambda(z)dz &\leq f_Z^{\lambda-1}(0)\int_\varepsilon^\infty f_Z(\lambda z)dz \\
&= \|f_Z\|_\infty^{\lambda-1}\lambda^{-1}\int_{\lambda\varepsilon} f_Z(z)dz.
\end{aligned}$$

The result follows with $C(\lambda) = \frac{2}{e(1-\lambda)\lambda}$. □

Lemma 4.4. *Consider a symmetric log-concave random variable Z with density $f_Z(z)$. Let $\{x_i\}$ be a $\mathbb{R}$-valued sequence. Then,*

$$\begin{aligned}
&\int_{\{|y|>\mu\}} f_Z(y)\ln\left(1 + \frac{\sum_j p_j\sum_{k\neq j} f_Z(y + x_k - x_j)}{f_Z(y)}\right)dy \\
&\leq \ln\left(3\|f_Z\|_\infty + \frac{1}{2\mu}\right)\mathbb{P}(Z > \mu) - \mathbb{E}\left[\ln f_Z(Z)\mathbb{1}_{\{|Z|>\mu\}}\right].
\end{aligned}$$

Proof. Let

$$a(y) := 1 + \frac{\sum_j p_j\sum_{k\neq j} f_Z(y + z_k - z_j)}{f_Z(y)}.$$

By Lemma 4.1,

$$\begin{aligned}
a(y) \leq 1 + \sum_j \frac{p_j(2\|f\|_\infty + \frac{1}{2\mu})}{f_Z(y)} &= 1 + \frac{2\|f\|_\infty + \frac{1}{2\mu}}{f_Z(y)} \\
&\leq \frac{1}{f_Z(y)}\left(3\|f\|_\infty + \frac{1}{2\mu}\right).
\end{aligned}$$

Thus,

$$\begin{aligned}
\int_\mu^\infty f_Z(y)\ln(a(y))\,dy &\leq \ln\left(3\|f\|_\infty + \frac{1}{2\mu}\right)\mathbb{P}(Z > \mu) - \int_\mu^\infty f_Z(y)\ln f_Z(y)dy \\
&= \ln\left(3\|f\|_\infty + \frac{1}{2\mu}\right)\mathbb{P}(Z > \mu) - \mathbb{E}\left[\ln f_Z(Z)\mathbb{1}_{\{|Z|>\mu\}}\right].
\end{aligned}$$

□

We can now prove our first main result.

Proof of Theorem 4.1. Let us first show that (13) implies (14) when $\varepsilon\|f_Z\|_\infty \geq 1/2\mu$. Observe that our two main terms are scale invariant in the following manner, for $c > 0$

$$\begin{aligned}
\delta(cX, cZ) &= H(cX) + h(cZ) - h(cX + cZ) \\
&= H(X) + h(Z) + \ln c - h(X + Z) - \ln c \\
&= \delta(X, Z).
\end{aligned}$$

Recall that for a general continuous random variable, $Y \sim f_Y(y)$ then $cY \sim \frac{1}{c} f_Y(y/c)$ resulting in $h(cY) = h(Y) + \ln c$. Also, with $\mu_x = \inf_{i \neq j} \frac{|x_i - x_j|}{2}$. that satisfies $\mu_{cx} = c\mu_x$ we have

$$\mathbb{P}(cZ > \lambda\mu_{cx}) = \mathbb{P}(Z > \lambda\mu_x).$$

Also observe that X and Z satisfy $\varepsilon\|f_Z\|_\infty \geq 1/2\mu_x$ then $\varepsilon\|f_{cZ}\|_\infty \geq 1/2\mu_{cx}$, thus we can apply (13) and our hypothesis $\varepsilon\|f_Z\|_\infty \geq 1/2\mu_x$ to $\delta(cX, cZ)$, to obtain

$$\begin{aligned}
\delta(X, Z) &= \delta(cX, cZ) \\
&\leq \tilde{M}(\lambda, cZ, cX) P(cZ \geq \lambda\mu_{cx}) \\
&= \tilde{M}(\lambda, cZ, cX)\mathbb{P}(Z \geq \lambda\mu_x).
\end{aligned}$$

Since,

$$\begin{aligned}
\tilde{M}(\lambda, cZ, cX) &= \left(2 + 2\ln\left(3\|f_{cZ}\|_\infty + \frac{1}{2\mu_{cx}}\right) + 2C(\lambda)\|f_{cZ}\|_\infty^{\lambda-1}\right) \\
&\leq 2 + 2\ln(3 + \varepsilon) + 2\left(\ln(\|f_{cZ}\|_\infty) + C(\lambda)\|f_{cZ}\|_\infty^{\lambda-1}\right).
\end{aligned}$$

and $\|f_{cZ}\|$ can take any value $y \in (0, \infty)$, we can strengthen our inequality to

$$\delta(X, Z) \leq \min_{y>0}\left(2 + 2\ln(3 + \varepsilon) + 2\left(\ln(y) + C(\lambda)y^{\lambda-1}\right)\right) P(Z \geq \lambda\mu_x).$$

Through routine calculus,

$$\min_{y>0}\left(2 + 2\ln(3 + \varepsilon) + 2\left(\ln(y) + C(\lambda)y^{\lambda-1}\right)\right) = 2 + 2\ln(3 + \varepsilon) + 2\frac{\ln(2/\lambda)}{1 - \lambda},$$

and thus (14) follows.

Now let us combine our previous results to obtain (13). We have

$$\delta(X, Z) \leq 2\left(\int_0^\mu + \int_\mu^\infty f_Z(y) \ln a(y) dy\right).$$

Using Lemma 4.2 we have

$$2\int_0^\mu f_Z(y) a(y) dy \leq 2\mathbb{P}(Z > \mu),$$

and by Lemma 4.4

$$2\int_\mu^\infty f_Z(y) \leq \ln\left(3\|f_Z\|_\infty + \frac{1}{2\mu}\right)\mathbb{P}(Z>\mu) - \mathbb{E}\left[\ln f_Z(Z)\mathbb{1}_{\{|Z|>\mu\}}\right],$$

and by Lemma 4.3

$$-\mathbb{E}\left[\ln f_Z(Z)\mathbb{1}_{\{|Z|>\mu\}}\right] \leq C(\lambda)\|f\|_\infty^{\lambda-1}\mathbb{P}(Z>\lambda\mu).$$

Since $\mathbb{P}(Z>\mu) \leq \mathbb{P}(Z>\lambda\mu)$ for $\lambda \in (0,1)$ we have

$$\delta(X,Z) \leq \left(2 + 2\ln\left(3\|f_Z\|_\infty + \frac{1}{2\mu}\right) + 2C(\lambda)\|f_Z\|_\infty^{\lambda-1}\right)\mathbb{P}(Z>\lambda\mu). \tag{16}$$

□

4.1 Sharpness of Bounds

We now show that the bound derived in the previous section is tight and cannot be improved significantly. Consider the discrete random variable X to be a Bernoulli(p) which we denote by B, then,

$$f_{X+B}(x) = (1-p)e^{-V(x+\mu)} + pe^{-V(x-\mu)}.$$

What follows generalizes the result of [10] (section IV) which handled the case that Z is Gaussian and $p = \frac{1}{2}$

Theorem 4.2. *When Z is a symmetric log-concave random variable with density function $f = e^{-V}$ and B independent and taking only two values,*

$$\delta(B,Z) \geq H(B)\mathbb{P}(Z>\mu)$$

Proof. Writing out the deficit in this case, and without loss of generality assuming that B takes the values $\pm\mu$,

$$\delta(Z,B) = (1-p)\int_{\mathbb{R}} e^{-V(x+\mu)}\ln\left(1 + \frac{p}{1-p}e^{V(x+\mu)-V(x-\mu)}\right)dx$$
$$+ p\int_{\mathbb{R}} e^{-V(x-\mu)}\ln\left(1 + \frac{(1-p)}{p}e^{V(x-\mu)-V(x+\mu)}\right)dx.$$

Let us investigate the first integral carefully, the second is similar. Since V is convex,

$$x \mapsto V(x+\mu) - V(x-\mu)$$

is increasing in x. Hence for $x \geq 0$ we have

$$e^{V(x+\mu)-V(x-\mu)} \geq e^{V(\mu)-V(-\mu)} = 1,$$

where the equality is a consequence of the assumed symmetry of Z. By the same logic, for $x < 0$

$$e^{V(x-\mu)-V(x+\mu)} \geq e^{V(-\mu)-V(\mu)} = 1.$$

Thus,

$$\delta(Z,B) \geq (1-p)\int_0^\infty e^{-V(x+\mu)}\ln\left(1 + \frac{p}{1-p}\right)dx$$
$$+ p\int_{-\infty}^0 e^{-V(x-\mu)}\ln\left(1 + \frac{(1-p)}{p}\right)dx$$

The right hand side reduces to our result. □

4.2 Capacity Bounds for a finite input, continuous output channel

Here we consider a channel in which messages are modeled via X distributed on n-elements, $x = \{x_i\}$, and the receiver reads the transmitted message $X+Z$ after corruption by an additive symmetric log-concave noise. We determine upper and lower bounds on the capacity $\mathcal{C}$ of the channel, defined as

$$\max_{p(X)} I(X; X+Z). \tag{17}$$

Theorem 4.3. *Consider a channel with input $X \in \{x_i\}$ and output $X+Z$ where Z is symmetric log-concave random variable. Let and $\mu = \sup_{i \neq j} |x_i - x_j|/2$ satisfying $\varepsilon \|f\|_\infty \geq \frac{1}{2\mu}$. Then, given any $\lambda \in (0,1)$, the capacity $\mathcal{C}$ admits the following bounds:*

$$\mathcal{C} \geq \log n - M\mathbb{P}(Z > \lambda\mu),$$

with $M = M(\lambda, \varepsilon)$ the universal constant defined as in (14).

The bound $\mathcal{C} \leq \log n$ is a consequence of $\delta(X,Z) \geq 0$. Thus the qualitative statement of the the theorem is that the tails of the log-concave noise control the capapcity.

Proof. Writing $I(X+Z; X) = H(X) - \delta(X,Z)$ and then utilizing the bound $\delta(X,Z) \leq M\mathbb{P}(Z > \lambda\mu)$ we have

$$I(X+Z; X) \geq H(X) - M\mathbb{P}(Z > \lambda\mu).$$

Taking supremums we have

$$\begin{aligned} \mathcal{C} &= \sup_{P(X)} I(X+Z; X) \\ &\geq \sup_{P(X)} H(X) - M\mathbb{P}(Z > \lambda\mu) \\ &\geq \log n - M\mathbb{P}(Z > \lambda\mu). \end{aligned}$$

□

Here we consider a channel in which messages are modeled via X distributed on $\pm\mu$, and the receiver reads the transmitted message $X+Z$ after corruption by an additive symmetric log-concave noise. We derive upper bounds on the capacity

$$\max_{p(X)} I(X; X+Z). \tag{18}$$

Theorem 4.4. *A binary channel with input modeled by X taking values on $\pm\mu$, subject to additive symmetric log-concave noise Z admits the following capacity bounds*

$$\mathcal{C} \leq \ln(2)\mathbb{P}(Z \leq \mu).$$

Proof. Applying Theorem (4.2) to the expression $I(X, X+Z) = H(X)-\delta(X, Z)$, we have

$$\begin{aligned}
\mathcal{C} &= \sup_{p(X)} H(X) - \delta(X, Z) \\
&\leq \sup_{p(X)} H(X) - H(X)\mathbb{P}(Z > \mu) \\
&= \ln(2)\mathbb{P}(Z \leq \mu).
\end{aligned}$$

$\square$

5 Acknowledgement

The authors acknowledge the support of the National Science Foundation for funding the research under Grant No. CMMI-1462862, CNS 1544721, and thank Mokshay Madiman for valuable suggestions and comments.

References

[1] S. Bobkov and M. Madiman. Concentration of the information in data with log-concave distributions. *Ann. Probab.*, 39(4):1528–1543, 2011.

[2] S.G. Bobkov. Isoperimetric and analytic inequalities for log-concave probability measures. *Ann. Prob.*, 27(4):1903–1921, 1999.

[3] Thomas A Courtade, Max Fathi, and Ashwin Pananjady. Wasserstein stability of the entropy power inequality for log-concave densities. *arXiv preprint arXiv:1610.07969*, 2016.

[4] T.M. Cover and J.A. Thomas. *Elements of Information Theory.* J. Wiley, New York, 1991.

[5] M. Fradelizi, J. Li, and M. Madiman. Concentration of information content for convex measures. *Preprint,* `arXiv:1512.01490`, 2015.

[6] Varun Jog and Venkat Anantharam. On the geometry of convex typical sets. In *Information Theory (ISIT), 2015 IEEE International Symposium on*, pages 829–833. IEEE, 2015.

[7] Varun Jog and Venkat Anantharam. Intrinsic entropies of log-concave distributions. *IEEE Transactions on Information Theory*, 64(1):93–108, 2018.

[8] M. Madiman, J. Melbourne, and P. Xu. Forward and reverse entropy power inequalities in convex geometry. *Convexity and Concentration*, pages 427–485, 2017.

[9] Arnaud Marsiglietti and James Melbourne. On the entropy power inequality for the rényi entropy of order [0, 1]. *arXiv preprint arXiv:1710.00800*, 2017.

[10] James Melbourne, Saurav Talukdar, Shreyas Bhaban, and Murti Salapaka. Error bounds for a mixed entropy inequality. In *Information Theory (ISIT), 2018 IEEE International Symposium on.* IEEE, 2018.

[11] Saurav Talukdar, Shreyas Bhaban, James Melbourne, and Murti V Salapaka. Analyzing effect of imperfections on the landauer's bound. *arXiv preprint arXiv:1802.01511*, 2018.

[12] Saurav Talukdar, Shreyas Bhaban, and Murti V Salapaka. Memory erasure using time-multiplexed potentials. *Physical Review E*, 95(6):062121, 2017.

[13] G. Toscani. A Strengthened Entropy Power Inequality for Log-Concave Densities. *IEEE Trans. Inform. Theory*, 61(12):6550–6559, 2015.