On the Accuracy of the High SNR Approximation of the Differential Entropy of Signals in Additive Gaussian Noise

Ramy H. Gohary and Halim Yanikomeroglu
Department of Systems and Computer Engineering, Carleton University, Ottawa, ON, Canada

4–9 May 2014 | IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2014 | Florence, Italy

Introduction

- Capacity evaluation involves maximizing mutual information between transmitted and received signals; e.g., for point-to-point channels

$$C = \max_{p_X(x)} I(X;Y) = \max_{p_X(x)} h(Y) - h(Y|X).$$

- In many cases, including non-coherent communications, $h(Y|X)$ straigforward to compute, but $h(Y)$ difficult to compute, why?.
- Consider $Y = XH + Z$, when H is unknown. Computing pdf of Y is difficult, let alone computing $h(Y)$.
- Evaluating differential entropy of sum of random variables is generally difficult.

Available Results

- An expression for the entropy of the sum of n independent random variables with finite supports and distributions symmetric around zero (Ordentlich'06).
- Lower bounds on differential entropy of sum of independent random variables obtained using entropy power inequalities (Madiman et al'09).
- Upper bounds on the differential entropy of the sum of two statistically-dependent random variables with log-concave pdfs (Cover et al'94).
- Other bounds are available (Lapidoth et al'08, Madiman et al'10, Madiman '08).

Motivation and Goal

- Problem arises in evaluating capacity of noncoherent MIMO systems and systems with phase noise.
- Noise contribution to differential entropy of received signal is ignored.
- Goal: assess how the approximation error decays with the increase of the SNR.

Approach and Results

- Derive expression of the differential entropy of the sum of arbitrary and Gaussian-distributed random variable with variance $1/\mathsf{SNR}$.
- Derive expressions for first two non-constant terms of Taylor series expansion of the differential entropy of received signal.
- Show that the approximation error decays as $1/\mathsf{SNR}$.
- Case of received signal given by the product of a random variable and a Gaussian-distributed fading coefficient.
- Derive explicit expression for second order term of the Taylor series of differential entropy.
- Schwartz's inequality to obtain efficiently computable bound on second order term.

Preliminaries

- Exchanging operations of integration and taking limits.
- Lebesgue's dominated convergence theorem.
- Exchanging operations of differentiation and taking limits.
- Valid under (generally strict) uniform convergence conditions .

The Taylor series expansion of the sum of an arbitrary and a Gaussian-distributed random variable

- Let X be arbitrary with twice differentiable pdf $p_X(\cdot)$ with support $[a,b]$.
- Let Z be Gaussian-distributed, i.e., $p_Z(z) = \frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}$.
- Received signal $Y = X + tZ$, where $t = \frac{1}{\sqrt{\mathsf{SNR}}}$.
- What is the Taylor series expansion of

$$h_t(Y) = -\int_{-\infty}^{\infty} p_Y(y;t)\log p_Y(y;t)dy?$$

- For t>0, the pdf of tZ is given by $\frac{1}{\sqrt{2\pi}t}e^{-\frac{z^2}{2t^2}}$.
- The pdf of Y is $p_Y(y;t) = \frac{1}{\sqrt{2\pi}t}\int_{-\infty}^{\infty} p_X(y-u)e^{-\frac{u^2}{2t^2}}du$.
- When $h_t(Y)$ analytic in t, its Taylor series around 0 converges for all $t<1$.
- Note: $p_Y(y;0) = p_X(y)$ and thus $h_0(Y) = h(X)$. Hence

$$h_t(Y) = h(X) + \frac{dh_t(Y)}{dt}\Big|_{t=0} t + \frac{d^2h_t(Y)}{dt^2}\Big|_{t=0}\frac{t^2}{2!} + \mathcal{O}(t^3).$$

- How to evaluate $\frac{dh_t(Y)}{dt}\Big|_{t=0}$ and $\frac{d^2h_t(Y)}{dt^2}\Big|_{t=0}$?

Evaluating the First and Second Order Terms

- Under continuity assumptions, $\frac{dh_t(Y)}{dt}\Big|_{t=0} = \lim_{t\searrow 0}\frac{dh_t(Y)}{dt}$, and

$$\frac{dh_t(Y)}{dt}\Big|_{t=0} = -\int_{-\infty}^{\infty}\big(1+\log p_X(y)\big)\lim_{t\searrow 0}\frac{\partial p_Y(y;t)}{\partial t}dy.$$

- Using the expression of $p_Y(y;t)$

$$\frac{\partial p_Y(y;t)}{\partial t} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} p_X(y-u)\Big(\frac{u^2}{t^4}-\frac{1}{t^2}\Big)e^{-\frac{u^2}{2t^2}}du.$$

- Using $\lim_{t\searrow 0}\frac{1}{t^2}e^{-\frac{u^2}{2t^2}} = \frac{1}{2}\lim_{t\searrow 0}\frac{u^2}{t^4}e^{-\frac{u^2}{2t^2}}$ yields

$$\lim_{t\searrow 0}\frac{\partial p_Y(y;t)}{\partial t} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} p_X(y-u)\lim_{t\searrow 0}\frac{1}{t^2}e^{-\frac{u^2}{2t^2}}du.$$

- To evaluate limit, use $\lim_{t\searrow 0}\frac{1}{t^2}e^{-\frac{u^2}{2t^2}}U(u) = -\delta'(u)$. Because $\delta'(u)$ is odd

$$\frac{dh_t(Y)}{dt}\Big|_{t=0} = 0.$$

- A similar approach yields

$$\lim_{t\searrow 0}\frac{d^2h_t(Y)}{dt^2} = \int_a^b \frac{1}{p_X(y)}\Big(\frac{dp_X(y)}{dy}\Big)^2 dy - \big(1+\log p_X(y)\big)\frac{dp_X(y)}{dy}\Big|_a^b$$

The Leading Terms of Taylor Series—The General Case

$$h_t(Y) = h(X) + \frac{t^2}{2}\Big(\int_a^b \frac{1}{p_X(y)}\Big(\frac{dp_X(y)}{dy}\Big)^2 dy - \big(1+\log p_X(y)\big)\frac{dp_X(y)}{dy}\Big|_a^b\Big) + \mathcal{O}(t^3).$$

A Special Case—X=RS

- For general S and R

$$p_X(x) = \int_0^{\infty}\frac{1}{r}\Big[p_{R,S}\Big(r,\frac{x}{r}\Big) + p_{R,S}\Big(-r,-\frac{x}{r}\Big)\Big]dr.$$

- When S and R statistically independent and R Gaussian

$$p_X(x) = \frac{1}{\sqrt{2\pi}}\int_0^{\infty}\frac{1}{r}\big(p_R(r)+p_R(-r)\big)e^{-\frac{x^2}{2r^2}}dr.$$

- For latter case: $h_t(Y) = h(X) + \frac{t^2}{2}\int_{-\infty}^{\infty}\frac{1}{p_X(y)}\Big(\frac{dp_X(y)}{dy}\Big)^2 dy + \mathcal{O}(t^3)$.

An Upper Bound on the Second Order Error Term

- Evaluating the coefficient of $t^2/2$ can be intractable.
- Bound $\int_{-\infty}^{\infty}\frac{1}{p_X(y)}\Big(\frac{dp_X(y)}{dy}\Big)^2 dy$.
- Using Schwartz's inequality, we have

$$\frac{1}{p_X(y)}\Big(\frac{dp_X(y)}{dy}\Big)^2 \le \frac{y^2}{\sqrt{2\pi}}\int_0^{\infty}\frac{1}{r^5}\big(p_R(r)+p_R(-r)\big)e^{-\frac{y^2}{2r^2}}dr.$$

- Hence,

$$\int_{-\infty}^{\infty}\frac{1}{p_X(y)}\Big(\frac{dp_X(y)}{dy}\Big)^2 dy \le \int_0^{\infty}\frac{1}{r^2}\big(p_R(r)+p_R(-r)\big)dr.$$

- Equality holds if r assumes a deterministic value in $[0,\infty)$.

Example 1

- Input signal, R, antipodal, i.e., $R = \pm a$ with equal probability.
- $p_R(r) = \frac{1}{2}\big(\delta(r+a)+\delta(r-a)\big)$
- $|h_t(Y) - h(X)| \le \frac{t^2}{2a^2} + \mathcal{O}(t^3)$.
- Bound derived from Schwartz's inequality is tight

Example 2

- Input signal, R, χ-distributed with m degrees of freedom, its pdf is given by

$$p_R(r) = \frac{2}{2^{m/2}\sigma^m\Gamma(m/2)}r^{m-1}e^{-r^2/2\sigma^2}U(r).$$

- For $m \ge 3$

$$|h_t(Y)-h(X)| \le \left(\frac{2^{-(m/2-1)}\rho_0}{\sigma^2\Gamma(m/2)}(m-4)(m-6)\cdots\right)t^2 + \mathcal{O}(t^3).$$

where $\rho_0 = \begin{cases}1 & m \text{ even}\\ \sqrt{\frac{\pi}{2}} & m \text{ odd}\end{cases}$.

- Coefficient of $t^2 \to 0$ as $m \to \infty$.

Conclusions

- Dominant error resulting from ignoring noise in computing entropy of received signal of non-coherent single-input single-output communication systems
- Error decays with $1/\mathsf{SNR}$.
- Explicit expression and upper bound for dominant error when input signal is transmitted over a channel with Gaussian fading coefficients.
- Considered real case only; complex case more work, but $h_t(Y)$ also decays as $1/SNR$.
- Extension: for MIMO systems with equal number of transmit and receive antennas, $h_t(Y)$ also decays as $1/SNR$.