Worst-Case Mutual Information Trajectories in Concatenated Codes with Asymptotic Interleavers

D. Divsalar* and S. Shamai†

October 21, 2001

Abstract

In this summary we present extremal (worst and best) average mutual information values carried by the extrinsic loglikelihood under the constraint of a given mean and variance while accounting for the consistency feature of loglikelihoods. This is done in an effort to gain insight into the iterative decoding procedure without resorting to the classical "Gaussian approximation".

1. Introduction

In recent contributions [1]-[4] an approximate method based on a Gaussian approximation for the a posteriori extrinsic probability has been invoked in an effort to gain insight into the major parameters affecting the performance of iterative decoding of concatenated interleaved coding systems. While [1] focuses on signal-to-noise ratio transfer, the studies in [2]-[6] adopt the mutual information measure. All results in [1]-[4] rely on the Gaussian approximation in the sense of [7], that is the extrinsic [8] probability function modeled as Gaussian and hence characterized by its average and variance. See [7] for further details and justifications.

In this work, we wish to relax the Gaussian approximation. Specifically, we shall use the mutual information framework as in [3]-[4], but rather than invoking the Gaussian assumption, we shall introduce the worst case associated mutual information which still satisfies the discrimination (the expectation of the loglikelihood ratio) and a variance constraint as well as the consistency property [7], which is satisfied by the a posteriori extrinsic loglikelihood ratio (interpreted here as a random variable). To be addressed also the maximum of the associated mutual information under the same average and variance constraints of the extrinsic loglikelihood ratio.

It is emphasized that the parameters of the a posteriori extrinsic loglikelihood ratio variable, such as its average (discrimination) and variance should be evaluated via probability propagation [9], not adhering to any Gaussian approximations.

*Sequoia Communications and Jet Propulsion Laboratory. Corresponding address: Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109 USA. e-mail: `dariush@shannon.jpl.nasa.gov`

†Department of Electrical Engineering, Technion - Israel Institute of Technology, Haifa 32000, Israel. e-mail: `sshlomo@ee.technion.ac.il`

2. Preliminaries

Consider a random variable y which carries information on a binary symmetric random variable x taking on the values -1 and 1 with equal probability 1/2. Assume that y takes on values in some observation space $\mathcal{Y}$, and is characterized by the conditional probability measure $p(y|x)$.

We now define a symmetry property as follows:

Symmetry Definition [11]. For all $y \in \mathcal{Y}$, there corresponds a $y' \in \mathcal{Y}$, such that the Radon-Nykodim derivatives satisfy:

$$dp(y|1)/dp(y'|-1) = dp(y|-1))/dp(y'|1) = 1 \ . \tag{2.1}$$

A special case of this definition is the standard symmetric transition probability property where $y' = -y$.

In the following it is useful to introduce to loglikelihood variable

$$\lambda = \log \frac{dp(y|1)}{dp(y|-1)} \ , \tag{2.2}$$

which is in fact a transformation of the random variable y.

Following [9],[10] it is elementary to show the consistency of the random variable λ. Assume that the effective probability measure is $p(y|1)$ (interpreted as assuming that $x = 1$, was transmitted). This induces a probability measure on λ (2.2) designated by $q_1(\lambda)$, and respectively $q_{-1}(\lambda)$ is induced by $p(y|-1)$.

Consistency: For a symmetric conditional measure we have [11]

$$q_1(-\lambda) = e^{-\lambda} q_1(\lambda) \tag{2.3}$$

and correspondingly,

$$q_{-1}(\lambda) = e^{-\lambda} q_{-1}(-\lambda) \ . \tag{2.4}$$

2.1. Mutual Information

Consider now the average mutual information:

$$I = I(y;x) = \mathrm{E} \log \frac{dp(y|x)}{dp(y)} \ , \tag{2.5}$$

where E designates here the expectation operator. It follows straightforwardly that

$$I = \log 2 - \frac{1}{2} \int_{-\infty}^{\infty} dp(y|1) \log\left(1 + e^{-\lambda}\right) - \frac{1}{2} \int_{-\infty}^{\infty} dp(y|-1) \log\left(1 + e^{\lambda}\right) \ , \tag{2.6}$$

where λ and y are interrelated by (2.2).

Invoking the symmetry condition (2.1) yields immediately that

$$\begin{aligned} I &= \log 2 - \int_{-\infty}^{\infty} dp(y|1) \log\left(1+e^{-\lambda}\right) \\ &= \log 2 - \int_{-\infty}^{\infty} dq_1(\lambda) \log\left(1+e^{-\lambda}\right) . \end{aligned} \tag{2.7}$$

Since

$$I = H(x) - H(x|y) ,$$

where H stands for entropies or conditional entropies, by (2.7) it is clear that the equivocation

$$H(x|y) \triangleq H_{x|\lambda} \triangleq \int dq_1(\lambda) \log\left(1+e^{-\lambda}\right) . \tag{2.8}$$

3. Optimization Problem

We are now in the position to set up the optimization problem. We first wish to minimize I or to maximize the equivocation $H_{x|\lambda}$.

$$\max_{q_1(\lambda)} \int_{-\infty}^{\infty} dq_1(\lambda) \log\left(1+e^{-\lambda}\right) \tag{3.1}$$

over all distributions of λ that satisfy

$$\int_{-\infty}^{\infty} dq_1(\lambda) = 1 , \tag{3.2}$$

$$\int_{-\infty}^{\infty} \lambda \, dq_1(\lambda) = \bar{\lambda} , \tag{3.3}$$

$$\int_{-\infty}^{\infty} \lambda^2 \, dq_1(\lambda) = \sigma^2 + (\bar{\lambda})^2 , \tag{3.4}$$

$$\frac{dq_1(-\lambda)}{dq_1(\lambda)} = e^{-\lambda} . \tag{3.5}$$

Constraints (3.3),(3.4) are interpreted as a given predetermined mean ($\bar{\lambda}$) and variance (σ^2) while (3.5) designates the consistency property.

We now use the consistency property (3.4) and the evident equality

$$\int_{-\infty}^{0} dq_1(\lambda) \, G(\lambda) = \int_{0}^{\infty} dq_1(-\lambda) \, G(-\lambda) = \int_{0}^{\infty} dq_1(\lambda) \, e^{-\lambda} \, G(-\lambda) \tag{3.6}$$

for any $G(\lambda)$, to transform the optimization problem originally stated (3.1)-(3.5) in the double infinite interval $(-\infty, \infty)$ to the positive infinite interval $(0, \infty)$. This yields

$$\max_{\substack{q_1(\lambda) \\ \lambda \geq 0}} \int_0^\infty dq_1(\lambda) \left[\log\left(1 + e^{-\lambda}\right) + e^{-\lambda} \log\left(1 + e^{\lambda}\right)\right] , \tag{3.7}$$

with the constraints,

$$\int_0^\infty dq_1(\lambda) \left(1 + e^{-\lambda}\right) = 1 , \tag{3.8}$$

$$\int_0^\infty dq_1(\lambda)\, \lambda\left(1 - e^{-\lambda}\right) = \bar{\lambda} , \tag{3.9}$$

$$\int_0^\infty dq_1(\lambda)\, \lambda^2\left(1 + e^{-\lambda}\right) = \sigma^2 + (\bar{\lambda})^2 . \tag{3.10}$$

We can further transform the optimization into an optimization with respect to probability measure $\mu(\lambda)$ where

$$d\mu(\lambda) = dq_1(\lambda)\left(1 + e^{-\lambda}\right) \quad , \quad \lambda \geq 0 , \tag{3.11}$$

which reads,

$$\begin{aligned} \max_{\mu(\lambda), \lambda \geq 0} \quad & \int_0^\infty d\mu(\lambda) \left[\frac{\log\left(1 + e^{-\lambda}\right) + e^{-\lambda} \log\left(1 + e^{\lambda}\right)}{1 + e^{-\lambda}}\right] \\ = \quad & \int_0^\infty d\mu(\lambda) \left[\log\left(1 + e^{-\lambda}\right) + \frac{\lambda e^{-\lambda}}{1 + e^{-\lambda}}\right] , \end{aligned} \tag{3.12}$$

$$\int_0^\infty d\mu(\lambda) = 1 , \tag{3.13}$$

$$\int_0^\infty d\mu(\lambda)\, \lambda \left(\frac{1 - e^{-\lambda}}{1 + e^{-\lambda}}\right) = \bar{\lambda} , \tag{3.14}$$

$$\int_0^\infty \lambda^2\, d\mu(\lambda) = \sigma^2 + (\bar{\lambda})^2 . \tag{3.15}$$

Assertion (tentative solution):

The maximizing distribution for (3.12) under constraints (3.13)-(3.15) is given by:

$$d\mu(\lambda) = (1 - a)\, \delta(0)\, d\lambda + a\delta(\lambda - \lambda_0)\, d\lambda , \tag{3.16}$$

or equivalently in terms of:

$$dq_1(\lambda) = (1 - a)\, \delta(0)\, d\lambda + a \left(\frac{\delta(\lambda - \lambda_0)}{1 + e^{-\lambda_0}} + \frac{e^{-\lambda_0}\, \delta(\lambda + \lambda_0)}{1 + e^{-\lambda_0}}\right) d\lambda , \tag{3.17}$$

where

$$\lambda_0 = \left(\sigma^2 + (\overline{\lambda})^2\right)/x \tag{3.18}$$

and where

$$x \triangleq a\lambda_0 \tag{3.19}$$

is the solution of the equation

$$x\left(\frac{1 - e^{-\frac{(\sigma^2+\overline{\lambda}^2)}{x}}}{1 + e^{-\frac{(\sigma^2+\overline{\lambda}^2)}{x}}}\right) = \overline{\lambda}\,. \tag{3.20}$$

By (3.18), (3.20) we have,

$$a = \frac{x^2}{\left(\sigma^2 + (\overline{\lambda})^2\right)}\,. \tag{3.21}$$

The corresponding average mutual information is

$$I = \frac{-x^2}{\sigma^2 + \overline{\lambda}^2}\log\left(\frac{1 + e^{-\left(\frac{\sigma^2+\overline{\lambda}^2}{x}\right)}}{2}\right) - \frac{x\,e^{-\left(\frac{\sigma^2+\overline{\lambda}^2}{x}\right)}}{1 + e^{-\left(\frac{\sigma^2+\overline{\lambda}^2}{x}\right)}}\,. \tag{3.22}$$

Results (3.20),(3.22) are presented in a parametric form. That is for a fixed value of $\frac{\overline{\lambda}^2}{\sigma^2}$, $\overline{\lambda}$ and $I_{\min}$ are given in a parametric form as a function of $x \in \left(0, \sqrt{\sigma^2 + (\overline{\lambda})^2}\right)$. It might be useful to introduce another parameter a (3.21) which takes on values in $a \in (0, 1]$.

Assertion (tentative solution):

The minimization in (3.12) over $\mu(\lambda)$ under constraints (3.13)-(3.15) is achieved by

$$d\mu(\lambda) = \lim_{\varepsilon\to 0}\left\{(1-\varepsilon)\delta(\lambda - \lambda_0)\,d\lambda + \varepsilon\delta\left(\lambda - \frac{\sqrt{\sigma^2 + \overline{\lambda}^2 - (1-\varepsilon)\lambda_0^2}}{\sqrt{\varepsilon}}\right)\right\}d\lambda\,, \tag{3.23}$$

where

$$\lambda_0\left(\frac{1 - e^{-\lambda_0}}{1 + e^{-\lambda_0}}\right) = \overline{\lambda}\,.$$

In terms of $q_1(\lambda)$ (3.23) at the limit $\varepsilon = 0$ transforms into

$$dq_1(\lambda) = \frac{1}{1 + e^{-\lambda_0}}\,\delta(\lambda - \lambda_0)\,d\lambda + \frac{e^{-\lambda_0}}{1 + e^{-\lambda_0}}\,\delta(\lambda + \lambda_0)\,d\lambda\,. \tag{3.24}$$

The corresponding mutual information is given by

$$I_{\max} = -\log\left(\frac{1 + e^{-\lambda_0}}{2}\right) - \frac{\lambda_0\,e^{-\lambda_0}}{1 + e^{-\lambda_0}}\,. \tag{3.25}$$

Note that $I_{\max}$ is independent of the value of σ^2 and is fully determined by the sole constraint $\overline{\lambda}$.

4. Conclusions

In this summary we have identified the minimum (and maximum) average mutual information associated with the log likelihood consistent random variable, standing for the extrinsic information. This is done in an effort to relax the Gaussian approximation while propagating mutual information expression through the iteration process in a similar manner as done in [2]-[4]. To that end the a posteriori average $\bar{\lambda}$ and variance σ^2 should be exactly evaluated [10] without resorting to a Gaussian approximation [7]. The minimum mutual information is then, a function of $\bar{\lambda}$ and σ^2. The maximal value of the mutual information depends on $\bar{\lambda}$ only. Using the proposed methods we obtained numerical examples that provide the upper bound and lower bound on the signal to noise ratio thresholds of concatenated codes with asymptotic interleavers.

References

[1] D. Divsalar, S. Dolinar and F. Pollara, "Iterative Turbo Decoder Analysis Based on Density Evolution," IEEE JSAC, May 2001.

[2] S. ten Brink, "Convergence Behavior of iterative Decoded Parallel Concatenated Codes," IEEE Trans. Commun. Vol. 49, No. 10, pp. 1727-1737, October 2001.

[3] S. ten Brink, "Design of Serially Concatenated Codes Based on Iterative Decoding Convergence," *Proc. 2nd and International Symposium on Turbo Codes*, pp. 319–322, Brest, France, Sept. 4–7, 2000.

[4] S. ten Brink, "Designing Iterative Decoding Schemes with Extrinsic Information Transfer Chart," AEU-International Journal of Electronics and Communications, Vol. 54, No. 6, pp. 389-398, 2000.

[5] S. ten Brink, "Exploiting the Chain Rule of Mutual Information for the Design of Iterative Decoding Schemes," 39th Ann. Allerton Conf. Com. Contr. and Computing, Monticello, Illinois, October 2001.

[6] A. Ashimkin, G. Kramer and S. ten Brink, "Code Rate the Area under Extrinsic Information Transfer Curves," submission to: IEEE Trans. Inform. Theory, Lausanne, Switzerland, June 30-July 5, 2002.

[7] S.-Y. Chung, T. Richardson and R. Urbanke, "Analysis of Sum Product Decoding of Low Density Parity Check Codes using Gaussian Approximations," *IEEE Trans. Inform. Theory*,Feb. 2001.

[8] J. Hagenauer, E. Offer and L. Papke, "Iterative Decoding of Binary Block and Convolutional Codes," *IEEE Trans. Inform. Theory*, Vol. 42, No. 2, pp. 429–445.

[9] T. Richardson and R. Urbanke, "Analysis and Design of Iterative Decoding Systems," *1999 IMA Summer Program: Codes Systems and Graphical Models*, Minnesota, USA, August 2–6, 1999.

[10] T. Richardson and R. Urbanke, "An Introduction to the Analysis of Iterative Coding Systems," preprint.

[11] M. Peleg, "Consistency with General Symmetry," *Personal Communication.*