# Blind quality estimation for corrupted source signals based on a-posteriori probabilities

Land, Ingmar; Thobaben, Ragnar

*Published in:*
Proc. IEEE Int. Symp. Inform. Theory (ISIT)

*Publication date:*
2004

*Document Version*
Publisher's PDF, also known as Version of record

# Blind Quality Estimation for Corrupted Source Signals Based on A-Posteriori Probabilities

Ragnar Thobaben
Institute for Circuits and Systems Theory
University of Kiel, Germany
rat@tf.uni-kiel.de

Ingmar Land
Information and Coding Theory Lab
University of Kiel, Germany
il@tf.uni-kiel.de

*Abstract* — **A novel approach is presented for assessing the quality of transmission systems, comprising quantized source signals and APP source decoders, via Monte-Carlo simulation. A-posteriori probabilities are exploited in order to obtain an unbiased estimate of both the symbol error probability and the expected distortion for the transmission system; knowledge of the transmitted source signal is not necessary. Compared to the conventional method this blind quality estimation has a smaller estimation variance.**

## Summary

The bit error rate estimation based on a-posteriori probabilities (APPs) was shown to be superior to the conventional one based on hard decisions [1]. In this paper this method is extended to symbol error rate (SER) and distortion estimation.

Let us assume a simplified transmission model, where a real-valued source signal[1] $U$ is quantized to quantization indices $I$, $i \in \mathcal{I}$, which are transmitted over a communication channel. Based on the channel observations $Y$ the receiver generates APPs $Pr(I\!=\!i|\boldsymbol{y})$ [2], which are exploited to obtain estimates $\hat{U}$ and $\hat{I}$ of $U$ and $I$.

Typically quality evaluation via Monte-Carlo simulation is based on a comparison of the transmitted source data $(u,\ i)$ to their reconstructed versions $(\hat{u},\ \hat{i})$ with respect to appropriate quality measures, such as the symbol error rate $P_s$ or distortion $D$, defined as $P_s := Pr(I \neq \hat{I})$ and $D := \mathrm{E}\{(U-\hat{U})^2\}$. Accordingly the conventional approach for measuring $P_s$ and $D$ can be summarized as follows:

**Method H:** Let us define the *hard SER sample* $z_H := \Pr(I \neq \hat{I}|I = i, \hat{I} = \hat{i})$, $z_H \in \{0,1\}$, indicating *whether* a symbol error occurred or not, and let us define the *hard distortion sample* $d_H := (u-\hat{u})^2$, taking into account the contribution to the reconstruction error due to estimate $\hat{u}$. For a transmission of $K$ source symbols, the corresponding quality samples can be used to compute the *hard SER estimate* $z_H^{(K)}$ and *hard distortion estimate* $d_H^{(K)}$ as

$$z_H^{(K)} := \frac{1}{K}\sum_{k=1}^{K} z_{H\,k} \quad \text{and} \quad d_H^{(K)} := \frac{1}{K}\sum_{k=1}^{K} d_{H\,k}.$$

Obviously, $z_H^{(K)}$ and $d_H^{(K)}$ rely on the knowledge of $u$ and $i$, from which it follows that the conventional Method H is not suitable for application in practical transmission systems.

Thus, we consider now the case, where knowledge of $u$ and $i$ is not available. I.e., only the source statistics, the estimates $\hat{u}$ and $\hat{i}$, and the set of APPs $p_{A_k} = \{Pr(I_k = i|\boldsymbol{y})\,|\, i \in \mathcal{I}\}$ may be used for quality estimation. These restrictions lead to a novel approach for the evaluation of $P_s$ and $D$, referred to as *Method S* in the following:

**Method S:** We define the *soft SER sample* as $z_S := Pr(I \neq \hat{I}|I = \hat{i}, P_A = p_A)$, which can be computed as $z_S = 1 - Pr(I = \hat{i}|\boldsymbol{y})$, and we define the *soft distortion sample* $d_S := \mathrm{E}\{(U-\hat{U})^2|P_A = p_A\}$, which is given by the *a-posteriori* expectation of the mean-squared error according to $d_S = \sum_{i\in\mathcal{I}} \mathrm{E}\{(u-\hat{u})^2|I=i\} \cdot Pr(I=i|\boldsymbol{y})$ for a given $\hat{u}$. Considering again the transmission of $K$ source symbols, the *soft SER estimate* $z_S^{(K)}$ and the *soft distortion estimate* $d_S^{(K)}$ for Method S are given by

$$z_S^{(K)} := \frac{1}{K}\sum_{k=1}^{K} z_{S\,k} \quad \text{and} \quad d_S^{(K)} := \frac{1}{K}\sum_{k=1}^{K} d_{S\,k}.$$

For comparison of both methods we regard the hard and the soft SER and distortion samples as random variables $Z_H$, $Z_S$ and $D_H$, $D_S$. From their definitions and since the estimates are sample means, it follows that $\mu_Z = \mathrm{E}\{Z_H\} = \mathrm{E}\{Z_S\} = P_s$ and $\mu_D = \mathrm{E}\{D_H\} = \mathrm{E}\{D_S\} = D$. Thus, both estimates are unbiased for both the SER and the distortion estimation.

An appropriate figure-of-merit is the estimation variance. The variance of the hard SER sample $Z_H$ can be written as

$$\sigma_{Z_H}^2 = \mathrm{E}\{Z_H^2\} - \mu_Z^2 = \mathrm{E}\{Z_H\} - \mu_Z^2 = \mu_Z(1-\mu_Z), \qquad (1)$$

where the identity $Z_H^2 = Z_H$ was applied. The variance of the soft SER sample $Z_S$, respectively, can be written as

$$\sigma_{Z_S}^2 = \mathrm{E}\{Z_S^2\} - \mu_Z^2 \leq \mathrm{E}\{Z_S\} - \mu_Z^2 = \mu_Z(1-\mu_Z) \qquad (2)$$

and is upper bounded by $\mu_Z(1-\mu_Z)$, since from $Z_S \in [0,1]$ it follows that $Z_S^2 \leq Z_S$, and thus $\mathrm{E}\{Z_S^2\} \leq \mathrm{E}\{Z_S\}$. Equality holds for the uninteresting cases $Z_S\!=\!0$ and $Z_S\!=\!1$ ($\sigma_{Z_S}^2\!=\!0$). For all other cases we have a lower bound on the ratio of variances $\sigma_{Z_H}^2$ and $\sigma_{Z_S}^2$ of the SER samples:

$$\frac{\sigma_{Z_H}^2}{\sigma_{Z_S}^2} > 1, \qquad (3)$$

resulting directly from (1) and (2).

A similar bound on the ratio of variances $\sigma_{D_H}^2$ and $\sigma_{D_S}^2$ of the distortion samples can be derived by applying Jensen's inequality to the a-posteriori expectation of $D_H^2$:

$$\mathrm{E}\{D_H^2|P_A = p_A\} \geq \mathrm{E}\{D_H|P_A = p_A\}^2 = D_S^2, \qquad (4)$$

where the identity $D_H = (u-\hat{u})^2$ and the definition of the soft distortion sample $D_S$ is exploited. It follows from (4) that $\mathrm{E}\{D_H^2\} \geq \mathrm{E}\{D_S^2\}$, where again equality holds for $\sigma_{D_S}^2 = 0$ (see above), and otherwise

$$\frac{\sigma_{D_H}^2}{\sigma_{D_S}^2} > 1, \qquad (5)$$

which represents a lower bound on the ratio of variances $\sigma_{D_H}^2$ and $\sigma_{D_S}^2$ of the distortion samples.

The bounds in (3) and (5) prove that the hard SER sample as well as the hard distortion sample have always (except for $P_s = 0$) a larger variance than the soft SER sample and the soft distortion sample, respectively. This reveals the superiority of the proposed Method S to the conventional Method H. In numerical results for Gauss-Markov sources the gain with respect to the estimation variance turned out to be even larger than predicted.

## References

[1] I. Land and P. Hoeher, "New results on Monte-Carlo bit error simulation based on the a-posteriori log-likelihood ratio," in *Proc. Int. Symp. on Turbo Codes & Rel. Topics*, Brest, France, Sept. 2003, pp. 531–534.

[2] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, pp. 284–287, Mar. 1974.

[1] Random variables are denoted by uppercase letters, their realizations by lowercase letters. Indices are omitted for convenience, whenever this can be done without ambiguity.