

OPTIMIZING BINARY DECISION SYSTEMS BY MANIPULATING TRANSMISSION INTERVALS

Michael A. Lexa and Don H. Johnson

Rice University
Department of Electrical and Computer Engineering
Houston, TX 77005-1892
amlexa@rice.edu, dhj@rice.edu

ABSTRACT

We study the optimization of a binary decision system where quantized (soft) decisions are transmitted across an additive white Gaussian noise channel. We adjust the bit transmission intervals to maximize the Chernoff distance at the output of the channel. At low channel signal-to-noise ratios (when the probability of a bit error is higher), we find unequal transmission intervals yield significant gains in terms of Chernoff distance and the information transfer ratio over the equal transmission case. This paper is a companion paper to [1] wherein the gains of unequal bit transmissions are studied in terms of minimum squared error.

1. INTRODUCTION

Because of their close connection with the performance of optimal binary detectors and because of their computational tractability, measures of discrimination are often used in the design of binary decision systems. Invariably the problem boils down to the placement of the thresholds to which the likelihood ratio or sufficient statistic is compared. But in some applications of binary decision systems it is also necessary to transmit a decision (whether hard or soft) after the data is remotely collected. A natural question to ask is whether or not the transmission scheme can also be optimized so as to maintain the distributional distance characterizing the signals after transmission. In a recent paper, Johnson and Rodríguez-Díaz [1] showed that unequal bit transmission intervals can yield significant gains (several dB) over the equal duration case in terms of mean squared error at low channel signal-to-noise ratios.

Motivated by these results, we investigate the possible gains of unequal bit transmission intervals in terms of the Chernoff discrimination distance. Specifically, we analyze the binary decision system shown schematically in Figure 1. The local decision-quantizer collects data generated under one of two hypotheses and outputs a quantized version (soft decision) of the likelihood ratio. These decisions are then

transmitted digitally using a BPSK signal set through an additive white Gaussian noise channel. Based upon the received signal the receiver makes a final decision as to which hypothesis is true.

We maximize the Chernoff distance at the output of the channel over all allowable transmission intervals. Here, ‘allowable’ simply means we place a constraint on the total duration of the transmission. Mathematically, we seek the solution to

$$\begin{aligned} & \max_{t_l} d_C(P_0, P_1) \\ & \text{subject to } \sum_{l=1}^L t_l = T \end{aligned} \quad (1)$$

where L is the length (number of bits) of the soft decision, t_l is the duration of the l^{th} bit transmission interval, and

$$d_C(P_0, P_1) = \max_s -\log \sum P_0^{1-s} P_1^s, \quad 0 \leq s \leq 1.$$

We focus on the Chernoff distance because it is the asymptotic exponential error rate of a minimum probability of error detector. Thus the solution of (1) directly reflects the performance of the receiver, assuming it is optimally designed.

In our system we assume the design of the decision-quantizer closely follows the design proposed in [2]. It is a L -bit (or equivalently a 2^L level) quantizer of the likelihood ratio whose quantization thresholds are chosen to maximize the Chernoff distance at its output. Thus we take the approach of optimizing the individual elements of the binary decision system with respect to the Chernoff distance instead of jointly optimizing all the elements. Consequently, our proposed system may not achieve superior performance compared to a system in which the decision-quantizer and the transmission scheme are jointly optimized. But because our design is less complex it is perhaps more practical.

We assume the receiver to be a matched filter so that the probability of receiving a bit transmission in error is

$$P_e^{(l)} = Q\left(\sqrt{2A^2 t_l / N_0}\right), \quad l = 1, \dots, L \quad (2)$$

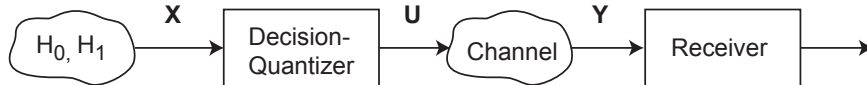


Figure 1: We assume the decision-quantizer collects data distributed either according to hypothesis H_0 or H_1 . It quantizes the likelihood ratio and then transmits its soft decision through an additive white Gaussian channel. The receiver makes a final decision based upon the received signal. We maximize the Chernoff distance at the input to the receiver by manipulating the transmission intervals of the soft decisions.

where A is the transmission signal amplitude and N_0 is the channel noise power. Note the dependence on the transmission intervals appears as an argument of the nonlinear $Q(\cdot)$ function. For our computations we resort to a standard numerical optimization routine, namely Matlab's `fmincon`.

2. RESULTS

Let X be the observation vector (data samples) collected by the local decision quantizer. Here, we assume X is normally distributed with mean m_0 if H_0 is true, or with mean m_1 if H_1 is true. Both hypotheses share a common variance σ^2 .

$$X|H_0 \sim \mathcal{N}(m_0, \sigma^2) \quad X|H_1 \sim \mathcal{N}(m_1, \sigma^2)$$

The outputs of the decision-quantizer are binary soft decisions characterized by two discrete distributions over the set of binary words of length L .

Let U represent a particular soft decision and let $\Pr[U_k|H_j]$ denote the probability that $U = k$ under H_j where $k = 0, \dots, 2^L - 1$ and $j = 0, 1$. These probabilities directly depend upon the quantizer's thresholds, which as mentioned above, are chosen to maximize the Chernoff distance between the two discrete distributions $P_{U|H_0}, P_{U|H_1}$. The probabilities at the output of the channel characterizing the received signals can be expressed as

$$\Pr[Y_m|H_j] = \sum_{k=0}^{2^L-1} \prod_{l=1}^L [1 - P_e^{(l)}]^{1-b_l^{(m)} \oplus b_l^{(k)}} [P_e^{(l)}]^{b_l^{(m)} \oplus b_l^{(k)}} \Pr[U_k|H_j] \quad (3)$$

where $b_l^{(m)}$ denotes the l^{th} bit in binary expansion of the number m and $b_l^{(k)}$ likewise denotes the l^{th} bit in the binary expansion of the number k . The notation $b_l^{(m)} \oplus b_l^{(k)}$ means the modulo-2 sum of the bits, which equals zero when the bits agree and one when they differ. $P_e^{(l)}$ is the channel error probability for the l^{th} bit (equation (2)).

Using equation (3) we can write an explicit expression for the Chernoff distance between the distributions at the

output of the channel. Our optimization problem now becomes

$$\max_{t_l, s} -\log \sum_{m=0}^{2^L-1} \Pr[Y_m|H_0]^{1-s} \Pr[Y_m|H_1]^s$$

subject to $\sum_{l=1}^L t_l = T; \quad 0 \leq s \leq 1. \quad (4)$

Note we simultaneously maximize over the transmission intervals t_l and the parameter s .

Figure 2 shows the gain in Chernoff distance when the decision-quantizer makes soft decisions of varying lengths and plots the corresponding optimized bit interval durations. For these examples the total time allotted to transmit L bits equals one ($T = 1$), $m_1 = 1$, $m_0 = 0$, and $\sigma^2 = 1$. The gain and bit intervals are plotted against the parameter $a = 2A^2/N_0$. Here, a replaces what normally is the SNR per bit. With varying bit intervals t_l the true SNR per bit changes. Thus we opt to use a parameter directly related to SNR but is independent of t_l .

Figure 2 shows that significant gains (factors of 2-4) can be made in the Chernoff distance for low signal-to-noise ratios. These gains can greatly reduce the probability of error of the receiver because the Chernoff distance an *exponential* error rate. These results also indicate that when the channel is very noisy it is optimal to expend most of the energy transmitting what amounts to a *hard decision*. In other words, the transmission of a coarser quantized version of the likelihood ratio leads to better performance than the transmission of a finely quantized version when the channel is poor. Beyond $a = 25$ there is no gain for the cases studied. These results are consistent with those found in [1] where the optimized criteria is mean-squared error.

The reason that the stacked bar graph does not show equal duration bit intervals for values above $a = 25$ demonstrates the non-uniqueness of the solutions to this problem. A certain value of the Chernoff distance can be obtained by several discrete distributions with differing number of elements.

Complimenting the previous results, we examine a quantity introduced in [3] which quantities how well the in-

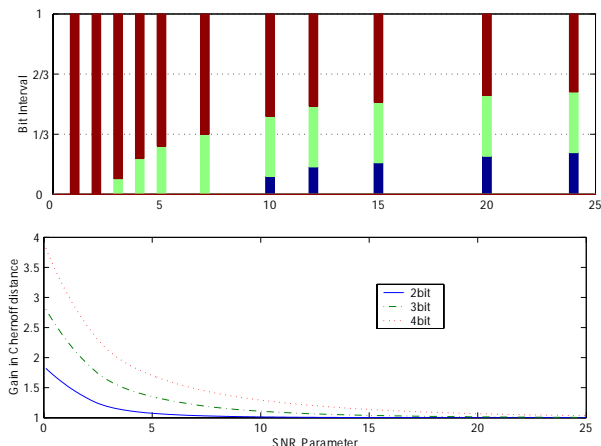


Figure 2: The top panel shows a stacked bar histogram of the bit transmission intervals that maximize the Chernoff distance at the output of the channel when the decision-quantizer makes 3-bit soft decisions. At low channel signal-noise ratios unequal bit duration are optimal but as the channel improves there is no gain over the equal duration case. The lower panel plots the gain in Chernoff distance for three different values of L , the number of bits in the soft decision. The horizontal axis in both plots is the SNR parameter $a = 2A^2/N_0$.

formation contained in X is preserved in Y . The information transfer ratio, γ , is defined as the ratio of the Kullback-Leibler distances at the input and output of any system. It is a number between zero and one and measures the degree to which a system maintains or degrades the Kullback-Leibler discrimination distance presented at its input. One indicates the system perfectly preserves the discriminatory information present at X in Y and zero means the system destroys all ability to distinguish the distributions in question apart. The complete theoretical basis of this criterion can be found in [3]. For our purposes here we substitute the Chernoff distance for the Kullback-Leibler distance and define

$$\gamma = \frac{d_C(P_{Y|H_0}, P_{Y|H_1})}{d_C(P_{X|H_0}, P_{X|H_1})}. \quad (5)$$

This substitution poses no theoretical problems in terms of the meaning of the information transfer ratio because the Kullback-Leibler and Chernoff distances share many of the same properties, of which the Data Processing Theorem [4] is the most important.

Figure 3 plots the information transfer ratios for the three cases considered above. The Chernoff distance at the input of the system (denominator of equation (2)) remains fixed by assumption. Thus, the variation of γ is solely due to the changes in the Chernoff distance at the output of the

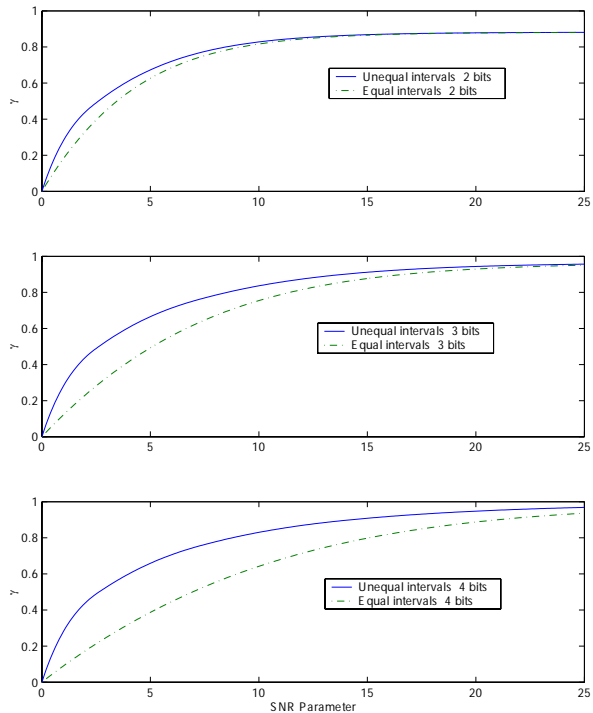


Figure 3: The performance of the binary decision system in terms of the information transfer ratio for both the unequal and equal bit intervals schemes are shown for various values of L . The horizontal axis in all three plots is again the SNR parameter a . Unequal bit intervals lift the system's performance as compared to equal bit intervals. At higher SNRs the gain decreases.

channel. Because we assume X is Gaussian distributed under each hypothesis, it equals

$$d_C(P_{X|H_0}, P_{X|H_1}) = \frac{(m_1 - m_0)^2}{8\sigma^2} = \frac{1}{8}.$$

We see the largest gains at low signal-to-noise ratios. At high signal-to-noise ratios, the equal and unequal bit interval curves become coincident indicating once again that when the channel is good there is no advantage to unequal bit intervals. These plots also clearly show that as L increases the system better preserves Chernoff distance. In fact, the 4-bit quantizer achieves an information transfer ratio of .97 for high signal-to-noise ratios. This implies further increases in L will not result in large improvements in the output Chernoff distance.

3. CONCLUSION

The binary decision system in Figure 1 can easily be expanded into a type of distributed detection system. Imagine

a bank of N decision-quantizers all collecting data from a common phenomenon and transmitting their soft decisions to a centralized receiver. If the inputs X_1, \dots, X_N are statistically independent and if the above optimization is generalized to maximizing the Chernoff distance between the joint distributions $P_{Y|H_0}$ and $P_{Y|H_1}$, the optimization reduces to optimizing each individual branch in the bank of decision-quantizers. Thus the results of this paper are applicable to certain distributed detection problems. This is not surprising because discrimination distances are commonly used in the analysis and design of distributed detection systems (e.g see [5]). However, to the authors' knowledge, the Chernoff distance has not been maximized by manipulating transmission intervals.

As is stated in [1] unequal bit intervals can be interpreted as a physical layer method of power control aimed at improving performance of the entire binary decision system. Given a channel state, it tells you how to allocate power on a bit-by-bit basis and tells you number of bits to transmit.

In summary, our results show unequal bit transmission intervals can yield significant gains in preserving the Chernoff distance and thus can lead to significant reduction in the receiver's probability of error.

4. REFERENCES

- [1] D.H. Johnson and H. Rodríguez-Díaz, "Optimizing physical layer data transmission for minimal signal distortion," *Accepted to IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2003.
- [2] H.V. Poor and J.B. Thomas, "Applications of ali-silvey distance measures in the design of generalized quantizers for binary decision systems," *IEEE Trans. on Communications*, vol. 25, no. 9, pp. 893–900, Sep 1977.
- [3] S. Sinanovic and D.H. Johnson, "Towards a theory of information processing," *Submitted to IEEE Trans. Signal Processing*, 2002.
- [4] T.M. Cover and J.A. Thomas, *Elements of Information Theory*, Wiley, 1991.
- [5] M. Longo, T.D. Lookagaugh, and R.M. Gray, "Quantization for decentralized hypothesis testing under communication constraints," *IEEE Trans. on Information Theory*, vol. 36, no. 2, pp. 241–255, Mar 1990.