A Systematic Description of Source Significance Information

Norbert Goertz
Institute for Digital Communications
School of Engineering and Electronics
The University of Edinburgh
Mayfield Rd., Edinburgh EH9 3JL, Scotland, UK
Email: norbert.goertz@ed.ac.uk

***Abstract*— In this paper we introduce a universal bit-significance measure that allows to quantify the usually varying importance of each data bit at a particular bit position in successive transmission frames. The bit significances can be used for transmission energy allocation in a source-adaptive fashion.**

I. Introduction

In any digital communication system the data bits either stem from a quantiser (when discrete-time samples that describe a multimedia source signal are transmitted) or they are "true" data bits in the sense that they carry information which is discrete in time and amplitude by nature (e.g. a compressed text). Further, data transmission of any kind is usually carried out by frames which are used to split the whole amount of data into many relatively small portions that are separately encoded and transmitted. This concept is driven by limitations on the delay and on the complexity.

Even though the consecutive frames may be structurally similar, the data bit at a particular position within a frame might sometimes be very important while in other frames it is less significant. This means that the adverse effects of transmission errors at the receiver output are often not the same for a particular data bit position in different frames.

In present system implementations different bit sensitivities (or significances) are accounted for by *static* unequal error protection [1] allocating stronger channel coding to the bits that are "on average" more sensitive. As argued above, however, the error sensitivities might considerably change from one block to another for a particular bit position in a frame, so more flexible schemes for unequal error protection will obviously be beneficial. The latter was shown in [2] for digital modulation schemes that are used to directly transmit the output bits of a quantiser. The basic idea is to allocate different transmission energies to the individual bits, depending on their *current* significances. This can be done without any alterations at the receiver, i.e., a broadcast system or the downlink in mobile radio could be enhanced without the need to change the mobile terminals that might already be in use.

In this paper we extend the basic ideas given in [2] in that we introduce a universally applicable bit significance measure, the "S-value", and use it for the allocation of transmission energy in a binary modulation scheme.

II. Bit Significance Measure

A. Significance Measure for Quantiser Bits

The example in Fig. 1 illustrates a typical case in which the significances of the data bits change with the locations of the input source samples. We consider two realisations "A" and "B" of the source signal x, which is scalar quantised by the reproducer values $y_0, ..., y_7$; the mean squared error is used a quality criterion. The sample A is quantised by y_1

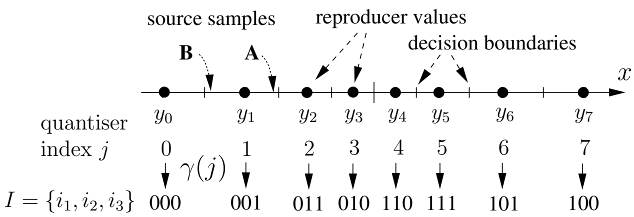

Fig. 1. Three-bit scalar quantiser with Gray bit-mapping $\gamma()$.

(the nearest neighbour of x among the reproducer values) and the corresponding bit combination "001" is transmitted. At the decoder output, an error in the middle bit will *not* lead to seriously larger distortion than the reproduction without any bit errors, because the reproducer value y_2 corresponding to the bit-combination "011" is also located close to the input sample A. The situation is different if we consider the sample B. Although B is also quantised by the reproducer value y_1, an error in the middle bit will now lead to much larger distortion. However, this time an error in the rightmost bit will cause only a small increase of the mean squared error. Hence, we have bits with strongly varying sensitivities, depending on the location of the unquantised source samples.[1]

Let us assume that $I = \{I_1, I_2, ..., I_K\} \in \mathcal{I}$ denotes a K-bit vector that represents the quantiser index j, and that $\mathcal{I} = \{0, 1\}^K$ is the set of all possible K-bit vectors. Moreover, we obtain the bit-vector I from the quantiser index j by the bit-mapping $\gamma()$, i.e., $I = \gamma(j)$; in Fig. 1 we use a Gray mapping as an example.

Given the input source sample x, we quantify the distortion if a particular bit I_l, $l = 1, 2, ..., K$, is reconstructed with some

[1]This can be extended to vector quantisers, i.e., the argumentation is true in a rather general context.

value $\tilde{I}_l = \xi \in \{0,1\}$, by

$$d'(\tilde{I}_l = \xi \mid x) = (x - y_{\gamma^{-1}(\tilde{I})})^2 \; , \quad (1)$$

with the quantiser reproducer value y_L, the inverse bit mapping $L = \gamma^{-1}(\tilde{I})$, and the bit vector $\tilde{I} = \{\tilde{I}_1, \tilde{I}_2, ..., \tilde{I}_K\} \in \mathcal{I}$, for which $\tilde{I}_l = \xi$ and $\tilde{I}_n = Q_n(x)$ for $n = 1, 2, ..., l-1, l+1, ..., K$. The notation $Q_n(x)$ corresponds to the output bit number n of the quantiser (including the bit mapping $\gamma()$) given the continuous-valued source sample x. In other words: in (1) we use the reproducer value y for reconstruction that has the same bit vector as the nearest neighbour of the input sample x, only excluding the bit position l for which we enforce the bit value $\xi \in \{0,1\}$ specified on the left-hand side of (1).

The definition in (1) will lead to a value of $d'(\tilde{I}_l = \xi \,|\, x) = 0$, if the input sample x is exactly identical to one of the reproducer values *and* if the specified bit value ξ is equal to the corresponding bit value at the output of the quantiser. This is not a problem as such, but below we will define the significance of a bit by the inverse of the (normalised) distortion, which would cause numerical problems if the distortion was close to zero. Therefore we modify (1) as follows:

$$d'(\tilde{I}_l = \xi \mid x) \doteq (x - y_{\gamma^{-1}(\tilde{I})})^2 + \beta \quad (2)$$

with the constant $0 < \beta \ll \sigma_x^2$, where σ_x^2 is the variance of the input samples (e.g. $\beta = 10^{-4} \cdot \sigma_x^2$).

Using (2) we define the normalised distortion:

$$d(\tilde{I}_l = \xi \mid x) \doteq \frac{d'(\tilde{I}_l = \xi \mid x)}{d'(\tilde{I}_l = 0 \mid x) + d'(\tilde{I}_l = 1 \mid x)} \; . \quad (3)$$

Hence, $d(\tilde{I}_l = 0 \mid x) + d(\tilde{I}_l = 1 \mid x) = 1$ and $0 < d(\tilde{I}_l = \xi \mid x) < 1$.

The significance of a reconstruction $\tilde{I}_l$ with some bit value ξ is inversely proportional to the distortion (3). Therefore, we define the *bit significance*

$$q(\tilde{I}_l = \xi|x) \doteq \frac{1}{d(\tilde{I}_l = \xi|x)} \; , \quad (4)$$

which will take on values in the open interval $(1, +\infty)$.

Now we define the *logarithmic bit significance measure* (S-value) as the logarithmic[2] ratio of the bit significances (4) for a "0" and a "1" bit:

$$S(\tilde{I}_l \mid x) \doteq \log \frac{q(\tilde{I}_l = 0 \mid x)}{q(\tilde{I}_l = 1 \mid x)} \; . \quad (5)$$

The S-value measures how sensitive the quality is in terms of a particular bit position. If the magnitude $|S(\tilde{I}_i \mid x)|$ is zero, this means it doesn't matter for the quality if the bit is reconstructed as "0" or "1". If the magnitude is large, the correct reconstruction of the bit is very important and the sign[3] of $S(\tilde{I}_i \mid x)$ determines whether the reconstruction of the bit should be "0" (when $S(\tilde{I}_i|x) > 0$) or "1" (when $S(\tilde{I}_i|x) < 0$).

[2]Throughout the paper the natural log is used in this definition.

[3]The definition (5) is such that the interpretation of the sign is consistent with the well-established log-likelihood ratios [3] used for soft-in/soft-out *de*coding on the *receiver side*.

B. Significance Measure for Data Bits

In a "pure" data transmission scenario (without any quantiser), time-varying significances of a particular bit-position in a packet might arise from transmitting in-band signalisation (e.g. to initiate handover to another mobile radio cell) as in GSM. The signalisation bits will usually be extremely sensitive to bit errors, as their erroneous interpretation will cause severe system failures, while in "normal" traffic the bit positions are less important in which case we might be willing to accept higher error probabilities in the reconstructions of the bits. If a text is compressed and, e.g., a variable-length lossless coding scheme is applied, some bit errors might cause symbol insertions and deletions or complete loss of synchronisation, while an error in exactly the same bit position in another coded block might merely cause a single symbol error at the decoder output: again we observe strongly variable bit-error sensitivities of the data bits.

Let us assume we are willing to tolerate a bit-error probability $p_l < 0.5$ in the reconstruction $\tilde{I}_l$ of a bit I_l with the realisation $i_l \in \{0,1\}$ that is known at the transmitter side. Based on this we can define a "distortion measure" according to

$$d(\tilde{I}_l = \xi | i_l) = \begin{cases} p_l & i_l = \xi \\ 1 - p_l & i_l \neq \xi \end{cases} \quad \text{with} \quad \xi \in \{0,1\} \; . \quad (6)$$

The interpretation of (6) is as follows: the value of the distortion measure is $1 - p_l$ if the reconstruction $\tilde{I}_l$ is different from the bit-value i_l, i.e., it is smaller than the Hamming distance, as we are willing to tolerate some bit-error probability p_l. If the reconstruction $\tilde{I}_l$ equals the original i_l, the distortion measure equals p_l (and not zero); in this way, we make it less important for the reconstruction to be error-free. If the tolerable bit-error probability is set to $p_l = 0.5$, the distortion measure has the same value in both cases, i.e., the reconstruction with "1" is as good as the reconstruction with "0" in any case, which is consistent with our expectation for a tolerable error-probability of 0.5. Note that in this case the bit is completely irrelevant and one should not waste energy for its transmission.

The normalised significance $q(\tilde{I}_l = \xi | i_l)$ of a reconstruction $\tilde{I}_l$ with some bit value ξ is inversely proportional to the distortion (6):

$$q(\tilde{I}_l = \xi | i_l) \doteq \frac{1}{d(\tilde{I}_l = \xi | i_l)} \; . \quad (7)$$

Note that as in (4) we have $q(\tilde{I}_l = \xi | i_l) \in (1, +\infty)$.

Similar as in (5) we define the *logarithmic bit-significance measure* (S-value):

$$S(\tilde{I}_l \mid i_l) \doteq \log \frac{q(\tilde{I}_l = 0 \mid i_l)}{q(\tilde{I}_l = 1 \mid i_l)} \; . \quad (8)$$

If we insert (6) and (7) into (8) we obtain

$$S(\tilde{I}_l \mid i_l) = \begin{cases} \log(p_l/(1-p_l)), & i_l = 1 \\ \log((1-p_l)/p_l), & i_l = 0 \end{cases} \; , \quad (9)$$

which can equivalently be written as

$$S(\tilde{I}_l \mid i_l) = \mathrm{sign}(1 - 2 \cdot i_l) \cdot \log \frac{1 - p_l}{p_l} \quad \text{for} \quad p_l < 0.5 \quad (10)$$

with $i_l \in \{0,1\}$. Note that again the sign of $S(\tilde{I}_l \mid i_l)$ determines the bit-value a reconstruction should have, while the magnitude $|S(\tilde{I}_l \mid i_l)| = \log \frac{1-p_l}{p_l} > 0$ describes the significance of the reconstruction. If the tolerable error probability is $p_l = 0.5$, we obtain $\log \frac{1-p_l}{p_l} = 0$, i.e., the bit has no significance at all, while the value of $\log \frac{1-p_l}{p_l}$ is very large when p_l is close to zero.

C. Conversion between S-values, Significances and Distortions

From the definitions (4), (7) and the fact that the distances (3), (6) are normalised (i.e., they sum up to one over all possible $\xi \in \{0,1\}$) we obtain

$$S(\tilde{I}_l \mid C) = \log \frac{d(\tilde{I}_l = 1 \mid C)}{1 - d(\tilde{I}_l = 1 \mid C)} = \log \frac{1 - d(\tilde{I}_l = 0 \mid C)}{d(\tilde{I}_l = 0 \mid C)} \tag{11}$$

where C denotes the condition either on a continuous-valued source sample x or on a data bit i_l. We can solve (11) either for $d(\tilde{I}_l = 1|C)$ or for $d(\tilde{I}_l = 0|C)$; the results can be written in compact form as follows:

$$q(\tilde{I}_l = \xi \mid C) = \frac{1}{d(\tilde{I}_l = \xi \mid C)} = \frac{1 + e^{S(\tilde{I}_l \mid C)}}{e^{S(\tilde{I}_l \mid C) \cdot \xi}} \tag{12}$$

with $\xi \in \{0,1\}$. Hence, we can uniquely convert the normalised significances and distortions into an S-values and vice versa.

III. Energy Allocation in Binary Transmission using S-Values

In what follows we will assume for simplicity that binary modulation is used although extensions to other modulation schemes are possible.

After computing the S-values by (5) or (10), we shall use them to allocate transmission energy[4] E_l to the individual bits. A simple approach is to linearly distribute the energy, with the S-Values as weighting factors, but one may also use more general rules such as

$$E_l = \frac{|S(i_l|x)|^\alpha}{\frac{1}{K}\sum_{l=1}^{K} |S(i_l|x)|^\alpha} \cdot E_s \tag{13}$$

with $\alpha > 0$; for $\alpha = 1$ this includes the linear case. In (13), E_s is the given *average* energy for each bit and K is the number of bits for which the joint energy allocation is to be applied.

An important property of the rule (13) is that $E_l = 0$ if $|S(I_l|x)| = 0$, i.e., no energy is allocated to insignificant bits. Another essential property of (13) is that the average energy stays the same, i.e., $\frac{1}{K}\sum_{l=1}^{K} E_l = E_s$.

It should be pointed out that this allocation of energies will not be "optimal" in any strict sense. If quantiser bits are transmitted, one could, in principle, state an optimisation problem and try to solve it e.g. by standard variational techniques such as Lagrange multipliers. The drawback of this approach is that the optimisation problem can usually only be solved numerically with high complexity [2]. Hence, our goal with the definition of S-values is to state a feasible general means of how to quantify bit significances and to conduct an energy allocation in the sense of a good practical solution that, once the S-values are known, no longer depends on the specific details of the source data. The most suitable allocation rule for a particular application (e.g., the parameter α in (13)) will be determined by simulations.

The problem of an "optimal" energy allocation becomes especially difficult if we consider the transmission of data bits with given tolerated error probabilities. If the number of data bits (i.e., the bit rate) is fixed and the channel cannot carry this amount of information, it is impossible to reconstruct the data bits at the receiver without errors. If the channel quality is very low it might even be impossible to reconstruct them with the given tolerated bit error rates. In the latter case the energy allocation rule will achieve a best-effort solution, which means that the very significant data bits might exhibit a larger bit error rate than tolerated but still it will be lower than for the less significant bits.

IV. Simulation Results

Some simulations results are given in this section for the quantisation and transmission of uncorrelated Gaussian source samples x with the variance $\sigma_x^2 = 1$. The source samples are optimally scalar quantised with fixed rate (Lloyd-Max quantiser [4]) by $K = 3$ bits. The transmission of the quantiser bits is carried out over a binary-input (BPSK modulation) Additive White Gaussian Noise (AWGN) channel with the average "channel SNR" E_s/N_0.

Fig. 2 shows, for each bit of a Gray bit mapping (see illustration in Fig. 1), the S-values that are obtained from (5) for $-3 < x < 3$ with $\beta = 10^{-4}\sigma_x^2$ in (2). A comparison with the natural binary bit mapping is given in Fig. 3. In both figures the reproducer values are indicated by "×" and the bit mappings used for transmission are given as well.

Figs. 2 and 3 confirm that the S-values indeed measure the significances of the bits: the S-values are large in magnitude if the x-value is close to a reproducer value and the bits that change at the decision boundaries in the middle between two reproducer values have S-values close to zero. A positive sign of the S-value indicates that the reconstruction of this bit should be "0" while a negative sign indicates a "1".

For the two bit mappings and based on the S-values in Figs. 2 and 3, the Figs. 4 and 5 show the resulting energy distribution according to (13) for $\alpha = 1.00$. As required, the bits with S-values close to zero are allocated very small energy, while the signficant bits will get up to twice as much energy than spent "on average".

In Figs. 6 and 7 we present some simulation results. As a performance measure we use the signal-to-noise ratio (SNR) of the source samples x and their reconstructions $\tilde{x}$ at the decoder output. We compare the conventional system (without source-adaptive energy allocation) with three schemes using S-values for energy allocation with $\alpha = \{0.5, 1.0, 2.0\}$ in (13). A simple hard-decision decoder was used, assuming coherent bit detection. In the region of interest, i.e., for SNRs of 10 dB

[4]The energy E_l for a bit is related to the power P_l by $P_l = E_l/T$, with T the constant bit transmission period.

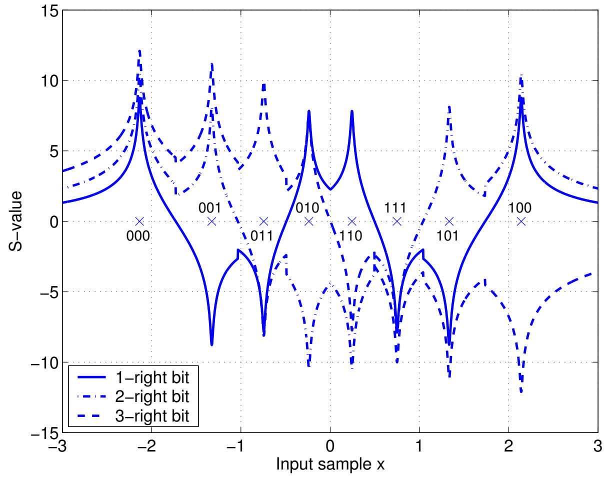

Fig. 2. S-values for a three-bit scalar Lloyd-Max quantiser for a Gaussian source; Gray bit mapping; $\beta = 10^{-4}\sigma_x^2$ in (2); the quantiser reproducer values are marked by "×", the bit mapping is included as well.

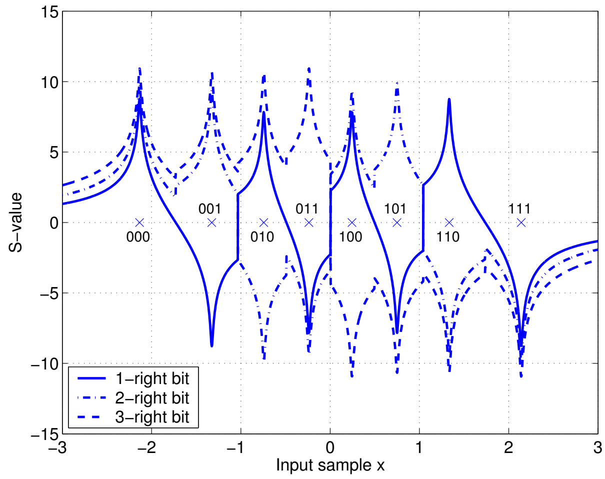

Fig. 3. S-values for a three-bit scalar Lloyd-Max quantiser for a Gaussian source; natural binary bit mapping; $\beta = 10^{-4}\sigma_x^2$ in (2); the quantiser reproducer values are marked by "×", the bit mapping is included as well.

or higher, we observe a gain of about 1 dB in E_s/N_0 (which directly turns into a gain in average transmission power) for the Gray mapping; the gain for the natural binary mapping is somewhat smaller. For both mappings $\alpha = 1.0$ in (13), i.e., linear averaging, turns out to be a good choice.

In Figs. 8 and 9 we present performance results for the same transmission schemes as in Figs. 6 and 7, but now with a conventional soft-decision decoder. Such a receiver, for an uncorrelated source and an AWGN channel model with binary input, coherent detection and without knowledge of source-adaptive energy allocation is known (e.g. [2]) to be given by

$$\tilde{x} = \sum_{I \in \mathcal{I}} y_{\gamma^{-1}(I)} \cdot \Pr(I \mid \tilde{z}_1, ..., \tilde{z}_K) \; . \qquad (14)$$

The quantity $\tilde{x}$ represents the reconstruction of the input sample x at the decoder output and $y_{\gamma^{-1}(I)}$ denotes the quantiser reproducer value that corresponds — via the bit mapping $\gamma()$

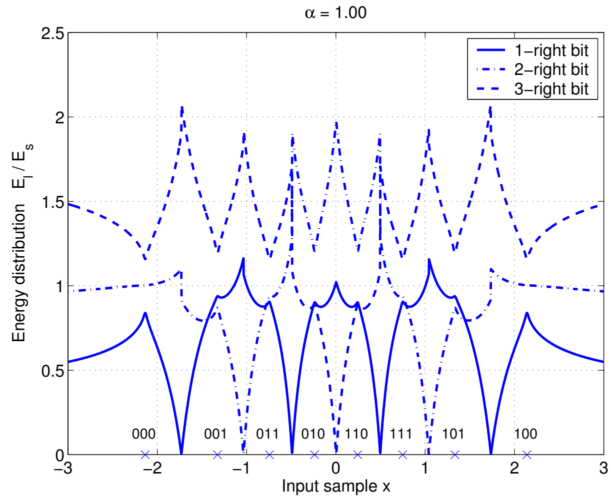

Fig. 4. Energy distribution according to (13) with $\alpha = 1.00$ based on the S-values in Fig. 2 (Gray bit mapping).

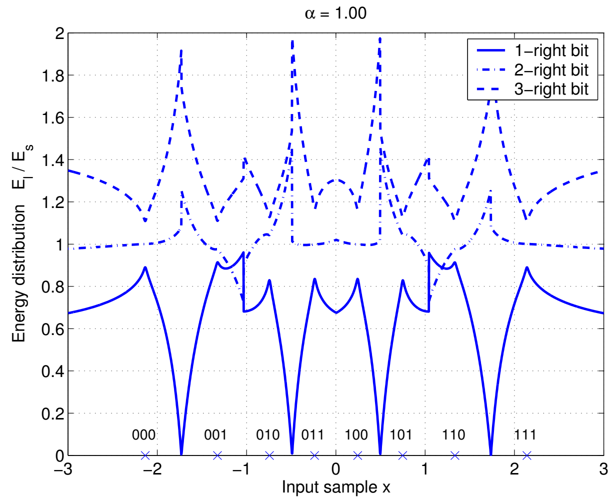

Fig. 5. Energy distribution according to (13) with $\alpha = 1.00$ based on the S-values in Fig. 3 (natural binary bit mapping).

— to the K-bit vector $I = \{I_1, ..., I_K\}$. Moreover, $\tilde{z}_l$ denotes the received sampled matched-filter output for the transmitted data bit $I_l \in \{0, 1\}$, $l = 1, ..., K$. The a-posteriori probability $\Pr(I \mid \tilde{z}_1, ..., \tilde{z}_K)$ can be computed by (Bayes rule)

$$\Pr(\underbrace{I_1, ..., I_K}_{=I} \mid \tilde{z}_1, ..., \tilde{z}_K) = A \cdot \Pr(I) \cdot \prod_{l=1}^{K} p(\tilde{z}_l \mid I_l) \; , \qquad (15)$$

where the factor A is a constant that ensures that the probabilities $\Pr(I \mid \tilde{z}_1, ..., \tilde{z}_K)$ sum up to one over all possible $I \in \mathcal{I}$; $\Pr(I)$ is the known (easily measurable) probability of the reproducer value which is transmitted by the bit vector I and $p(\tilde{z}_l \mid I_l)$ is the probability density function of the channel which for the binary-input AWGN model is given by

$$p(\tilde{z}_l \mid I_l) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp(-\frac{1}{2\sigma_n^2} (\tilde{z}_l - z_l)^2) \Big|_{z_l = 1 - 2 \cdot I_l} \; . \qquad (16)$$

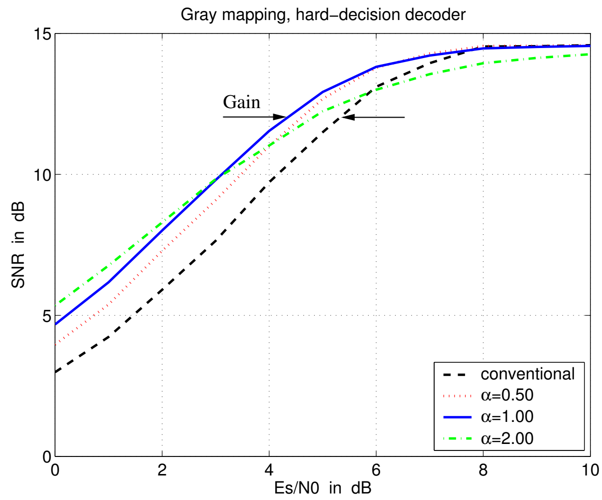

Fig. 6. System performance in terms of source SNR using energy allocation by S-values with $\alpha = \{0.5, 1.0, 2.0\}$; comparison with a conventional system; Gray mapping. A conventional hard-decision receiver was always used.

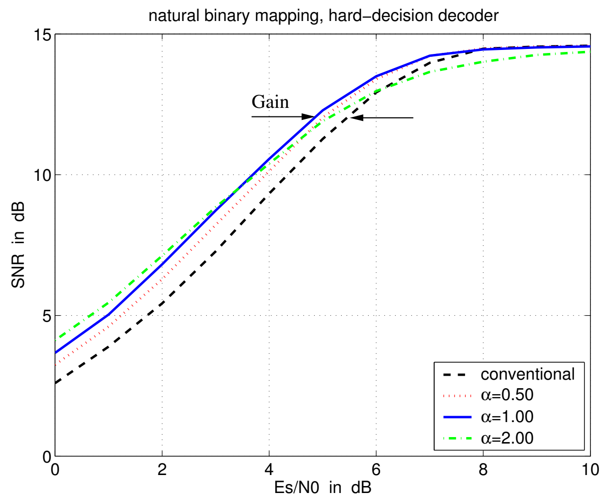

Fig. 7. System performance in terms of source SNR using energy allocation by S-values with $\alpha = \{0.5, 1.0, 2.0\}$; comparison with a conventional system; natural binary mapping. A conventional hard-decision receiver was used.

Note that the source-adaptive energy allocation based on the S-values by (13) will lead to different transmission energies E_l for the bits. As this energy allocation is unknown at the receiver, we use a conventional soft-decision decoder which has only knowledge about the *average* channel quality given by E_s/N_0. This leads to the noise-variance $\sigma_n^2 = \frac{N_0}{2E_s}$ in the channel model (16).

Figs. 8 and 9 indicate that the conventional soft-decision decoder (14) gives a gain of somewhat less than 1 dB over the hard-decision decoder for all transmission systems. The relative gain by source-adaptive energy allocation using S-values is again around 1 dB in E_s/N_0 (average transmission power).

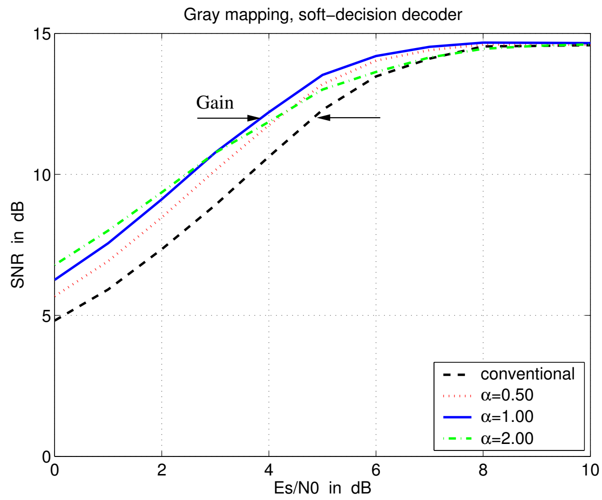

Fig. 8. System performance in terms of source SNR using energy allocation by S-values with $\alpha = \{0.5, 1.0, 2.0\}$; comparison with a conventional system; Gray mapping. A conventional soft-decision receiver was always used.

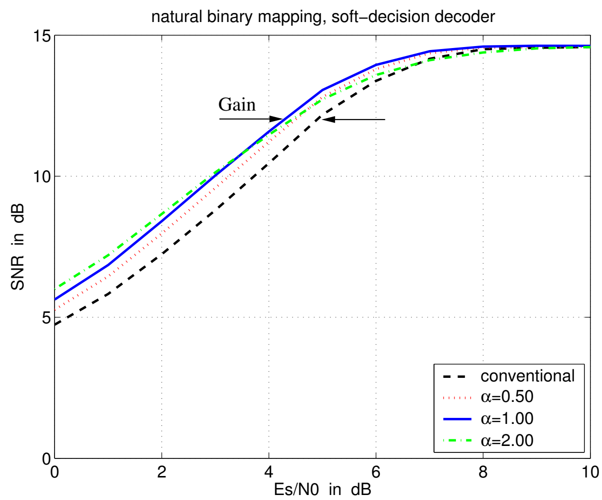

Fig. 9. System performance in terms of source SNR using energy allocation by S-values with $\alpha = \{0.5, 1.0, 2.0\}$; comparison with a conventional system; natural binary mapping. A conventional soft-decision receiver was used.

V. Conclusions

In this paper we have introduced S-values as a new measure for the signficances of data bits. The S-values can be used for bitwise allocation of transmission energy. We presented simulation results for a simple system model which show that significant gains can be achieved by the new concept.

References

[1] B. Masnick and J. Wolf, "On linear unequal error protection codes," *IEEE Transactions on Information Theory*, vol. IT-3, pp. 600–607, Oct. 1967.

[2] N. Görtz and E. Bresch, "Source-adaptive power allocation for digital modulation," *IEEE Communications Letters*, vol. 7, pp. 569–571, Dec. 2003.

[3] J. Hagenauer, "The turbo principle: Tutorial introduction and state of the art," in *Proceedings International Symposium on Turbo Codes & Related Topics*, pp. 1–11, Sept. 1997.

[4] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*. New Jersey: Prentice-Hall, Inc., Englewood Cliffs, 1984.