Unequal Error Protection by Modulation with Unequal Power Allocation

Thomas Brüggen and Peter Vary

***Abstract*— In digital communication systems for speech, audio or video signals the individual bits of the transmitted parameters u exhibit different bit error sensitivities. Usually channel coding with unequal error protection (UEP) is applied. However, some transmission systems do not include channel coding for several reasons. For this situation, a novel concept is proposed which achieves UEP by allocating different transmission power to individual bits according to their bit error sensitivities. The optimization criterion for unequal power allocation is the mean square of the error between the original parameter u and the decoded parameter $\hat{u}$ which has a strong correlation with subjective perception.**

***Index Terms*— Unequal error protection, unequal power allocation.**

I. Introduction

IN MOST digital transmission systems for speech, audio, or video signals parameters u are produced by a source encoder. As the digital representation of these parameters exhibit different bit error sensitivities, channel coding with unequal error protection is often applied. However, some transmission systems, e.g., DECT and Bluetooth, include only weak or no channel coding for several reasons. For this situation a novel concept called *modulation with unequal power allocation* (MUPA) is proposed which achieves UEP by periodically allocating unequal transmission power to individual bits. In this contribution MUPA is explained in the context of BPSK modulation. The extension to other modulation schemes is possible. Before BPSK modulation the individual bits are weighted differently taking the channel SNR into account. The least significant bit (LSB) gets less energy than the most significant bit (MSB), while the average transmitted energy per bit remains unaffected.

The idea of differently weighted bits in PCM transmission has already been discussed in [1]. However, in this early proposal an approximate solution is derived which is based on the simplifying assumption that only single bit errors occur. In [2] UEP is achieved by bit-wise control of the BPSK modulation amplitudes. However, in contrast to our proposal neighborhood relations between the unquantized and quantized source parameters are exploited in combination with *Gray* or *Gray-like* bit mappings. Furthermore, the different amplitude levels are found by "systematically trying out a number of energy distributions" [2]. In contrast to [1] and [2], we present an analytical approach with numerical optimization for *arbitrary* bit mappings. As optimization criterion we consider the mean square error $\mathrm{E}\{(\mathrm{u}(\tau) - \hat{\mathrm{u}}(\tau))^2\}$ in the parameter domain. The corresponding *signal-to-noise ratio* (SNR)

$$\mathrm{SNR[dB]} = 10\log_{10}\left(\frac{\mathrm{E}\{\mathrm{u}^2(\tau)\}}{\mathrm{E}\{(\mathrm{u}(\tau) - \hat{\mathrm{u}}(\tau))^2\}}\right) \tag{1}$$

which has a strong correlation with subjective perception is used to compare the different solutions.

II. System Model

The model of the transmission system is shown in Fig. 1. Instead of any specific signals and source encoders, we use a zero-mean Gaussian source with variance $\sigma_u^2 = 1$ to model the parameters $u(\tau)$ which are delivered at time τ by the source encoder. This allows us to conduct precisely defined experiments. However, without loss of generality the results can be applied to any real source codecs. Each $u(\tau)$ is quantized to $\overline{u}(\tau) \in \{\overline{u}_\kappa | \kappa = 1, 2, ...2^M\}$ with the reproduction levels $\overline{u}_\kappa$. The index assignment (IA) maps the values $\overline{u}(\tau)$ to bit patterns $\underline{x}(\tau) \in \{\underline{x}_\kappa | \kappa = 1, 2, ...2^M\}$ with $\underline{x}_\kappa = (x_\kappa^{(1)}, x_\kappa^{(2)}, ...x_\kappa^{(M)})$. The bits $x_\kappa^{(i)} \in \{-1, +1\}$, $i = 1, 2, ...M$, with $x_\kappa^{(M)}$ being the LSB. Each value $x_\kappa^{(i)}$ is multiplied with the specific weight $w_i \in \mathbb{R}^+$ (Sec. III) of the diagonal matrix $\underline{W} = \mathrm{diag}(w_1, w_2, ...w_M)$. The obtained channel symbols (Fig. 2)

$$\underline{y}(\tau) = \underline{W} \cdot \underline{x}(\tau) \tag{2}$$

are transmitted over a channel with additive white Gaussian noise (AWGN) $\underline{n}(\tau) = (n^{(1)}(\tau), n^{(2)}(\tau), ...n^{(M)}(\tau))$, $n^{(i)}(\tau) \in \mathbb{R}$, $i = 1, 2, ...M$, with $N(0, \sigma_n^2)$, the variance $\sigma_n^2 = N_0/2$, and the noise power spectral density N_0. We assume perfect knowledge of σ_n^2 at the transmitter.

In case of coherent BPSK modulation we obtain the disturbed vector $\underline{z}(\tau) = \underline{y}(\tau) + \underline{n}(\tau)$ at the receiver. After hard decision decoding, the estimate $\hat{u}(\tau)$ is produced by table lookup.

III. Optimization of Weights

A. Optimization Approach

We consider bipolar channel symbols $\pm\sqrt{E_b^{(i)}}$ and $E_b^{(i)}$ being the energy for each bit i. If bits in $\underline{x}(\tau)$ are inverted due to AWGN, a wrong decision $\hat{\underline{x}}(\tau)$ is made at the receiver. Thus, the distortion $d(\tau) = \overline{u}(\tau) - \hat{u}(\tau)$ occurs. We want to calculate optimal weights in the sense of a minimized expected value $\mathrm{E}\{d^2(\tau)\}$, i.e., the MMSE.

In the following we assume ergodicity. Thus, we calculate $\mathrm{E}\{d^2\}$ taking the different possible values

$$d_{\kappa,\eta} = \overline{u}_\kappa - \hat{u}_\eta \tag{3}$$

Manuscript received November 8, 2004. The associate editor coordinating the review of this letter and approving it for publication was Dr. Cha'o-Ming Chang.

The authors are with the Institute of Communication Systems and Data Processing, Aachen University (RWTH), 52074 Aachen, Germany (e-mail: brueggen@ind.rwth-aachen.de).

Digital Object Identifier 10.1109/LCOMM.2005.06016.

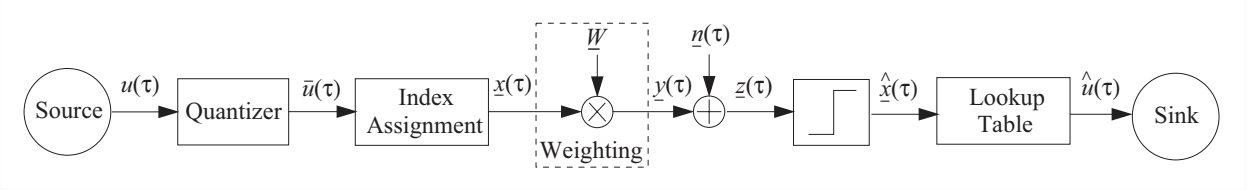

Fig. 1. Baseband system model of BPSK-MUPA with weight matrix __ and hard decision decoding.

and their probabilities of occurrence into account:

$$\mathrm{E}\{d^2\} = \sum_{\kappa=1}^{2^M}\sum_{\eta=1}^{2^M} d^2_{\kappa,\eta}\cdot P(\underline{x}_\kappa,\hat{\underline{x}}_\eta) \quad (4)$$

$$= \sum_{\kappa=1}^{2^M}\sum_{\eta=1}^{2^M} d^2_{\kappa,\eta}\cdot P(\underline{x}_\kappa)\cdot P(\hat{\underline{x}}_\eta|\underline{x}_\kappa) \quad (5)$$

The mean square error (MSE) $\mathrm{E}\{d^2\}$ depends on the probability of occurrence $P(\underline{x}_\kappa)$ of the reproduction levels $\overline{u}_\kappa$ and on the transition probabilities $P(\hat{\underline{x}}_\eta|\underline{x}_\kappa)$ between the transmitted and received bit patterns $\underline{x}_\kappa$ to $\hat{\underline{x}}_\eta$, respectively.

For the minimization of (5) we start with the bit error rate (BER) P_b for the uniform case (i.e., $w_i = 1.0, i = 1,...M$) [3]

$$P_b = \frac{1}{2}\,\mathrm{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) = \int_{\sqrt{E_b}}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_n}\cdot\exp\left(\frac{-\xi^2}{2\sigma_n^2}\right)\,\mathrm{d}\xi\,. \quad (6)$$

P_b depends on the bit energy E_b. With the MUPA concept, the M different bits are transmitted with different bit energies $E_b^{(i)} = w_i^2\cdot E_b$, $i = 1,...M$. For $w_i > 1$, the bit error probability is reduced, if $w_i < 1$, the BER is increased. This different weighting is illustrated in Fig. 2.

In a frame of M bits the total energy is:

$$\sum_{i=1}^{M} w_i^2\cdot E_b = M\cdot E_b \quad (7)$$

resulting in the energy constraint

$$\sum_{i=1}^{M} w_i^2 = M. \quad (8)$$

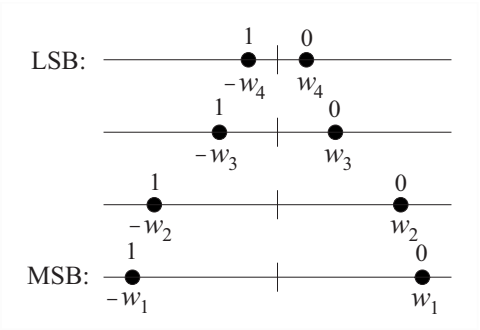

Fig. 2. Channel symbols _( ) with their amplitudes $_i$ for = 4 bits, binary bit representation.

Using the normalization $E_b = 1$ we get the different bit error probabilities

$$P_b^{(i)} = \int_{w_i}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_n}\cdot\exp\left(\frac{-\xi^2}{2\sigma_n^2}\right)\,\mathrm{d}\xi\,. \quad (9)$$

Our goal is to minimize $\mathrm{E}\{d^2\}$ by optimizing the weights w_i. As a consequence, e.g., the number of bit errors in the MSB positions is reduced, such that the overall quality (parameter SNR) can be increased. Due to the statistical independence of the noise samples, the transition probabilities can be calculated as [3][4]

$$P\left(\hat{\underline{x}}_\eta|\underline{x}_\kappa\right) = \left(\prod_{\substack{i=1\\ x_\kappa^{(i)}\neq\hat{x}_\eta^{(i)}}}^{M} P_b^{(i)}\right)\cdot\left(\prod_{\substack{i=1\\ x_\kappa^{(i)}=\hat{x}_\eta^{(i)}}}^{M}\left(1-P_b^{(i)}\right)\right). \quad (10)$$

B. Lagrange Multiplier Method (LMM)

The optimization problem is the minimization of (5) with the constraint (8). An appropriate approach is to use the *Langrange Multiplier* Method (LMM) with the *Lagrange Multiplier* λ as follows:

$$L(w_1,...w_M,\lambda) = \left(\sum_{\kappa=1}^{2^M}\sum_{\eta=1}^{2^M} d^2_{\kappa,\eta}\cdot P(\underline{x}_\kappa)\cdot\left(\prod_{\substack{i=1\\ x_\kappa^{(i)}\neq\hat{x}_\eta^{(i)}}}^{M} P_b^{(i)}\right)\cdot\left(\prod_{\substack{i=1\\ x_\kappa^{(i)}=\hat{x}_\eta^{(i)}}}^{M}\left(1-P_b^{(i)}\right)\right)\right) - \lambda\left(M-\sum_{i=1}^{M} w_i^2\right)\,. \quad (11)$$

This equation is a non-linear function $L:\mathbb{R}^{M+1}\rightarrow\mathbb{R}$ with $M+1$ unknown variables, i.e., the M weights w_i and the *Lagrange Multiplier* λ. Next, we need the $M+1$ partial derivatives $\partial L(w_1,...w_M,\lambda)/\partial w_j$, $j = 1,2,...M$, and $\partial L(w_1,...w_M,\lambda)/\partial\lambda$ of (11). Two cases can occur: if $x_\kappa^{(j)}\neq\hat{x}_\eta^{(j)}$, then the first product contains the factor $P_b^{(j)}$ and therefore has to be differentiated with respect to w_j according to (9). Otherwise $(1-P_b^{(j)})$ has to be differentiated.

C. Newton Algorithm

In the next LMM step the roots w_j and λ of the $M+1$ derivatives have to be calculated. Due to the transcendent function $\exp(x)$, we employ the $(M+1)$-dimensional *Newton* algorithm [5] to calculate the roots. For convenience, we insert λ into the set of variables by renaming $(w_1,...w_M,\lambda)^T$ to $\underline{v} = (v_1,...v_{M+1})^T$ with $v_i = w_i$ for $i = 1,2,...M$, and $v_{M+1} = \lambda$. We define $L'_j(\underline{v}) = \partial L(\underline{v})/\partial v_j$,

$j = 1, 2, ...M + 1$, and $\underline{L}'(\underline{v}) = (L'_1(\underline{v}), ...L'_{M+1}(\underline{v}))^T$ for the vector of the first partial derivatives. With this notation, the *Newton* iteration rule is given by

$$\underline{v}_{k+1} = \underline{v}_k - J_L^{-1}(\underline{v}_k) \cdot \underline{L}'(\underline{v}_k) \tag{12}$$

with the iteration counter k and the inverse *Jacobian* matrix $J_L^{-1}(\underline{v}_k)$ containing the first partial derivatives of $L'(\underline{v})$, i.e., the second partial derivatives of $L(\underline{v})$, $L''(\underline{v}) = \partial L'_g(\underline{v})/\partial v_j$, for all combinations of $g, j = 1, 2, ...M + 1$. When the exit condition $||\underline{v}_{k+1} - \underline{v}_k|| < \epsilon$ is fulfilled, the LMM supplies the extremum $\underline{v}$ which is a local or a global minimum.

IV. Simulation Results: Example

The performance of MUPA is confirmed by simulations with M = 4 bits, natural binary IA, and a source signal $u(\tau)$ (Fig. 1) with the Gaussian probability density function. An optimal (non-uniform) Lloyd-Max quantizer (LMQ) [6] is used.

Fig. 3 shows that for very bad channels with $E_b/N_0 < -5$ dB the weight w_1 for the MSB is emphasized at the cost of all the other ones. Above $E_b/N_0 = -5$ dB also w_2 becomes slightly larger than 1. With further increasing E_b/N_0 all weights approach towards $w_i = 1$. This can be explained with the extreme sensitivity (gradient) of the bit error probability in that range with respect to weight variations. Consequently, the different weights are still different but close to 1.

In Fig. 4 the *horizontal* coding gains in terms of E_b/N_0 for transmission with and without MUPA are approximately 1 to 1.5 dB for E_b/N_0 values in the range of 3 to 6 dB.

By MUPA, the parameter SNR is improved by the *vertical* gains or unchanged but never impaired for all E_b/N_0 values. The maximum parameter SNR gain of 3.27 dB is achieved at E_b/N_0 = 3 dB. MUPA also outperforms a r = 1/2, (7,5)-convolutional code with hard decision Viterbi decoding at most channel qualities and is slightly worse only for E_b/N_0 values in the range of 5.2 to 8 dB. The hard decision Viterbi decoder is much more complex than MUPA which requires only table lookup decoding.

The results in Fig. 4 are based on weights w_i which are optimized under the assumption that the E_b/N_0 is perfectly known. However, MUPA is quite insensitive to an E_b/N_0 mismatch of 3 dB or even more. Therefore a few different sets of optimal weights, precalculated for different classes of E_b/N_0 ratios, could be applied.

V. Conclusions and Outlook

The digital transmission of parameters over a noisy channel can be improved by allocating unequal power to individual bits according to their significance. This MUPA approach is of special interest if the transmission system does not include channel coding for some reason.

The optimization criterion is the minimum mean square error of the decoded parameters. By the example of a 4 bit quantizer and BPSK modulation it is shown that E_b/N_0 coding gains (*horizontal gains*) of up to 1.5 dB, and parameter SNR gains (*vertical gains*) of up to 3.27 dB can be achieved in comparison to transmission systems with constant bit energy. In comparison to a standard r = 1/2, memory 2, convolutional encoder with hard decision Viterbi decoding, the MUPA approach achieves almost comparable or even better results. However, the MUPA decoder is just a simple table lookup decoder.

If some additional complexity is allowed, further performance improvements can be achieved by combining MUPA with softbit source decoding (SBSD) [7]. In this case the hard bit table lookup decoding of MUPA is replaced by parameter estimation. The MUPA concept can be applied to higher order modulation schemes.

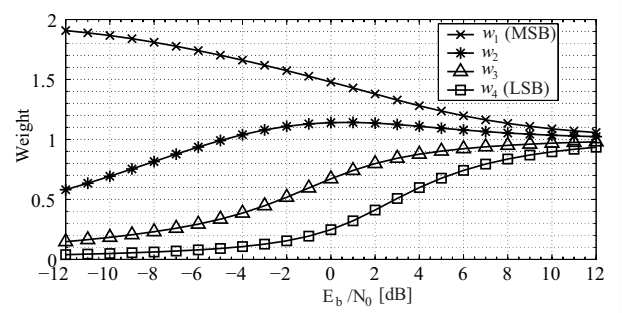

Fig. 3. Weight values $_i$ (LMQ, natural binary).

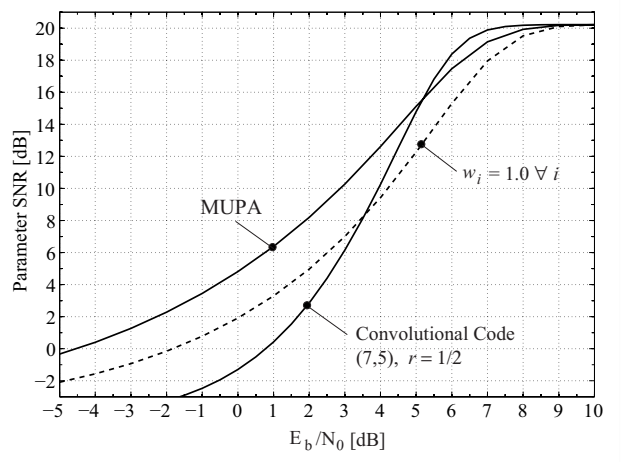

Fig. 4. Comparison of the parameter SNR.

VI. Acknowledgement

The authors appreciate the valuable comments of the editor and the reviewers which helped to improve the paper.

References

[1] E. Bedrosian, "Weighted PCM," *IRE Trans. Inform. Theory*, vol. IT-4, pp. 45–49, Mar. 1958.
[2] N. Görtz and E. Bresch, "Source-Adaptive Power Allocation for Digital Modulation," *IEEE Commun. Lett.*, vol. 7, pp. 569–571, Dec. 2003.
[3] J. G. Proakis, *Digital Communications.* New York: McGraw-Hill Inc., 1995.
[4] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes.* New York: McGraw-Hill Inc., 2002.
[5] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*. Cambridge University Press, 1992.
[6] N.S. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice Hall, 1984.
[7] T. Fingscheidt and P. Vary, "Softbit speech decoding: a new approach to error concealment," *IEEE Trans. Speech Audio Processing*, vol. 9, pp. 240–251, Mar. 2001.