ITG-Fachtagung
Sprachkommunikation 2006

How To Improve the Speech Quality by Increasing the Average Bit Error Rate

Thomas Brüggen, Peter Vary
Institute of Communication Systems and Data Processing (ind), RWTH Aachen University, Germany
{brueggen,vary}@ind.rwth-aachen.de

Abstract

The speech quality is one of the most important features of a digital communication system. The intelligibility and the naturalness of the speech are decisive for the user's acceptance. In this paper we present a scheme for *unequal error protection* (UEP), which is based on *modulation with unequal power allocation* (MUPA). The MUPA technique allocates different transmission power to the modulation symbols according to the individual bit error sensitivities of the assigned bits. This corresponds to an intentional increase of the average bit error rate, but improves the speech quality, as the bit error rates of the most important bits are reduced. The average transmitted energy per bit remains unaffected, while the average channel capacity is decreased by MUPA.

1. Introduction

In digital transmission systems for speech, audio, and video signals source encoders extract parameters which are quantized and converted into a digital representation. As the individual bits of this representation exhibit different bit error sensitivities, usually channel coding with UEP is applied [1]. However, some transmission systems do not include channel coding for several reasons (e.g., DECT, Bluetooth). In this situation the MUPA concept [2], [3] can be applied to introduce implicit UEP. MUPA distributes the transmission energy unequally in a periodically time-varying manner to the modulation symbols by the weights w_i, $i = 1, 2, ...M$ with M being the number of bits of an encoded source parameter, e.g., the bits assigned to a quantizer reproduction level. The weights increase the distances between modulation symbols, which differ in the *most significant bit* (MSB), while other distances between modulation symbols differing, e.g., in the *least significant bit* (LSB) are made smaller. For each channel E_b/N_0 a set of weights is calculated once in advance.

UEP by differently weighted bits has been discussed, e.g., in [4] and [5]. In [4], differently weighted bits in *pulse code modulation* (PCM) transmission have been described. However, only for PCM an approximate solution is derived based on the simplifying assumption that only single bit errors occur. In [5] *discrete* symbol amplitudes are calculated for BPSK with a fixed step size. In contrast to [4] and [5], the MUPA approach allows rigorous optimization of the *continuous* symbol amplitudes [2], [3]. As optimization criterion we consider the

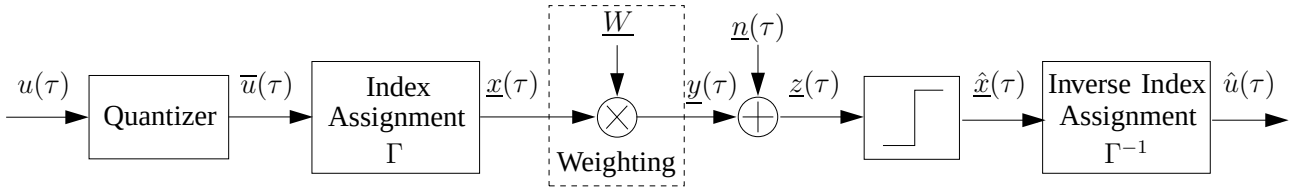

Fig. 1: Baseband system model of BPSK-MUPA with the weight matrix $\underline{W}$ and hard decision decoding.

mean square error (MSE) $\mathrm{E}\{(\mathrm{u}(\tau) - \hat{\mathrm{u}}(\tau))^2\}$ between the original source parameter $u(\tau)$, e.g., a prediction coefficient or a gain factor or even a PCM sample, and its reconstruction $\hat{u}(\tau)$ (Fig. 1). The corresponding *signal-to-noise ratio* is given by the parameter SNR

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

which especially has a strong correlation with subjective perception for most of the model based coding schemes such as CELP codecs.

2. BPSK-MUPA System Model

The baseband model of the BPSK transmission system is shown in Fig. 1. The parameters $u(\tau)$ delivered at time τ by the source encoder are modeled here by a zero-mean Gaussian source with the variance $\sigma_u^2 = 1$. This approach allows to conduct investigations under precisely defined conditions and independent from any specific codec standard. 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) utilizes the one-to-one mapping function Γ and maps the values $\overline{u}(\tau)$ to the bit patterns $\underline{x}(\tau) \in \{\underline{x}_\kappa | \kappa = 1, 2, ...2^M\}$ with $\underline{x}_\kappa = (x_\kappa^{(1)}, x_\kappa^{(2)}, ...x_\kappa^{(M)})$: $\underline{x}(\tau) = \Gamma(\overline{u}(\tau))$. The bits $x_\kappa^{(i)} \in \{-1, +1\}$, $i = 1, 2, ...M$, are ordered from the MSB $x_\kappa^{(1)}$ to the LSB $x_\kappa^{(M)}$. Each value $x_\kappa^{(i)}$ is multiplied with the specific weight $w_i \in \mathbb{R}^+$ (Sec. 3) of the diagonal matrix $\underline{W} = \mathrm{diag}(w_1, w_2, ...w_M)$. The obtained channel symbols $\underline{y}(\tau) = \underline{W} \cdot \underline{x}(\tau)$ are transmitted over a channel with additive white Gaussian noise (AWGN) $\underline{n}(\tau)$ having the variance $\sigma_n^2 = N_0/2$ with the noise power spectral density N_0. In the 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 $\hat{u}(\tau) \in \{\overline{u}_\kappa | \kappa = 1, 2, ...2^M\}$ is produced by inverse IA (table lookup).

3. Unequal Power Allocation

We consider bipolar channel symbols $\pm\sqrt{E_b^{(i)}}$ with $E_b^{(i)}$ being the energy for each bit i. If bits in the transmitted bit pattern $\underline{x}_\kappa$ are inverted due to AWGN, an erroneous bit pattern $\hat{\underline{x}}_\eta$ is received. Thus, the distortion $d_{\kappa,\eta} = \overline{u}_\kappa - \hat{u}_\eta$ occurs. We want to calculate optimal weights w_i in the sense of a minimized expected value

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

The MSE $\mathrm{E}\{d^2\}$ depends on the probability of occurrence $P(\underline{x}_\kappa)$ of the reproduction levels $\overline{u}_\kappa = \Gamma^{-1}(\underline{x}_\kappa)$ 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. The transmission energy is normalized to the average bit energy E_b to 1. With $E_b^{(i)} = w_i^2 \cdot E_b$ the energy constraint is given by

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

The minimization problem (1) with the energy constraint (2) is solved by the *Lagrange* and *Newton* algorithms [2], [6].

4. Analysis of BPSK-MUPA

4.1 Bit Error Rate (BER)

Due to the variation of the bit energy $E_b^{(i)}$ by MUPA the different bits experience different channel qualities and bit error rates characterized by the $E_b^{(i)}/N_0$ ratio. According to the significance of the bits, the bit specific BERs are distributed above (LSB) and below (MSB) the average BER curve (Fig. 2). Due to the high BER of the LSB weighted by w_4, the average BER is worse for MUPA than for conventional BPSK, i.e., $w_i = 1.0 \ \forall \ i$. However, for the typical model based speech codecs, the hearing impression becomes better with MUPA (Fig. 3: compare the dashed line and the line marked with "$\circ$"). Even if PCM audio samples are considered as parameters $\overline{u}(\tau)$, MUPA gives a significant perceptual improvement. If the transmitter has imprecise estimation of the E_b/N_0 (channel mismatch, Fig. 3), the weights of a higher E_b/N_0 should be chosen to reduce the performance loss.

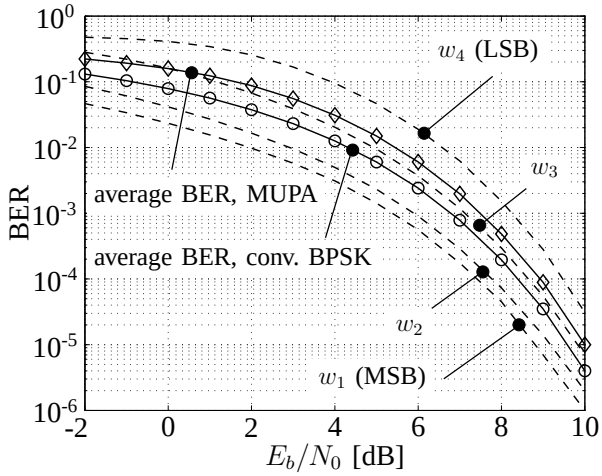

Fig. 2: Bit error rates of BPSK-MUPA (Γ: Gray) compared to a conventional BPSK system.

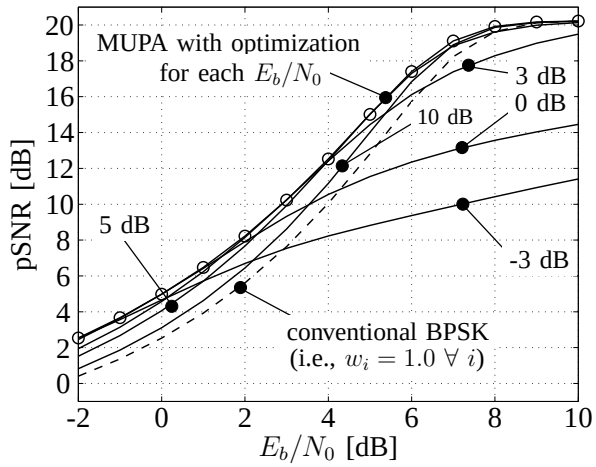

Fig. 3: Comparison of BPSK (Γ: Gray) with and without MUPA, and cases of weight mismatch (weights optimized for the E_b/N_0 values -3, 0, 3, 5, and 10 dB are used for all channel qualities).

4.2 Channel Capacity

The different bit energies according to the significance of the bits also influence the channel capacity $\mathcal{C}$ [7]. A BPSK-MUPA transmission of M bits can be modeled by subdividing the channel into M different subchannels. The bit at position i, $i = 1, 2, ...M$, in $\underline{x}(\tau)$ is transmitted over the subchannel i with the specific capacity $\mathcal{C}^{(i)}$ and the channel quality $E_b^{(i)}/N_0$. The following derivation of the channel capacities $\mathcal{C}^{(i)}$ assumes a BPSK transmission system with optimized MUPA weights w_i and a memoryless and symmetric AWGN channel with binary input $y \in \{-w_i, +w_i\}$ and the random variable Y, and the continuous output z with the random variable Z. The average mutual information $\mathcal{I}(Y;Z)$ is defined as [7], [8]

$$\mathcal{I}(Y;Z) = \mathcal{H}(Y) - \mathcal{H}(Y|Z). \tag{3}$$

with the entropy $\mathcal{H}(Y)$ of Y and the conditional entropy $\mathcal{H}(Y|Z)$. In general, the channel capacity $\mathcal{C}$ is defined as the maximum of the average mutual information $\mathcal{I}(Y;Z)$

$$\mathcal{C} = \max_{P(y)} \mathcal{I}(Y;Z). \tag{4}$$

with respect to all possible choices of the input distribution $P(y)$. The equal distribution $P(y = -w_i) = P(y = +w_i) = 1/2$ results in $\mathcal{H}(Y) = 1$, i.e., the maximum value of the first part of (3). This is fulfilled, e.g., by the natural binary index assignment with the *Lloyd-Max* quantizer (LMQ) and a symmetric probability distribution of $\overline{u}(\tau)$. The second term in (3) is the conditional entropy $\mathcal{H}(Y|Z)$. In the special case of a *discrete* channel input value y and a *continuous* output value z it holds

$$\mathcal{H}(Y|Z) = -\sum_y P(y) \cdot \int_{-\infty}^{\infty} p(z|y) \cdot \log_2(P(y|z))\ \mathrm{d}z \tag{5}$$

Next, the *a posteriori* probabilities $P(y|z)$ are replaced by applying the Bayes theorem in mixed form taking the two possible BPSK symbols $-w_i$ and $+w_i$ for y into account:

$$P(y = +w_i|z) = \frac{p(z|y = +w_i) \cdot P(y = +w_i)}{p(z)} \tag{6}$$

and analog for $P(y = -w_i|z)$. With equal distribution of $P(y)$, (6) can be written as

$$P(y|z) = \left(1 + \left(\frac{p(z|y = -w_i)}{p(z|y = +w_i)}\right)^{\mathrm{sgn}(y)}\right)^{-1} \tag{7}$$

for $y \in \{-w_i, +w_i\}$ with the sign function sgn. Since an AWGN channel with zero-mean white Gaussian noise and the variance $\sigma_n^2 = N_0/2$ is

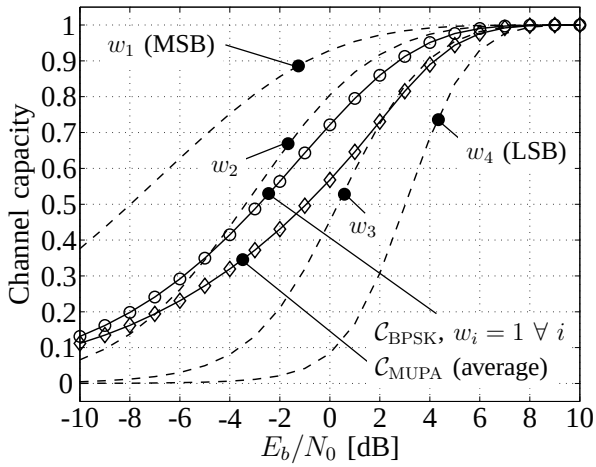

Fig. 4: Channel capacity $\mathcal{C}_{\mathrm{MUPA}}$ of BPSK-MUPA compared to the conventional BPSK system with $\mathcal{C}_{\mathrm{BPSK}}$ (M = 4, Γ: natural binary).

considered, the channel transition probability density $p(z|y)$ in (7) can be calculated by

$$p(z|y) = \frac{1}{\sqrt{2\pi}\sigma_n} \cdot e^{-\frac{(z-y)^2}{2\sigma_n^2}} \tag{8}$$

leading to the *a posteriori* probabilities

$$P(y|z) \quad = \quad \left(1 + e^{\frac{-4zy\ \cdot\ \mathrm{sgn}(y)}{2\sigma_n^2}}\right)^{-1}. \tag{9}$$

The sum over y in (5) consists of two addends due to $y \in \{-w_i, +w_i\}$. Since the channel probability density function is symmetrical, i.e., $p(z|y) = p(-z|-y)$, these two addends are equal. Thus, the sum can be eliminated by taking the addend with $y = -w_i$ twice. With (8) and (9) in (5), and with (3) in (4), the capacity $\mathcal{C}^{(i)}$ of the subchannel i becomes

$$\mathcal{C}^{(i)} = 1 - \frac{1}{\sqrt{2\pi}\sigma_n} \cdot \int_{-\infty}^{\infty} e^{-\frac{(z+w_i)^2}{2\sigma_n^2}} \cdot \log_2\left(1 + e^{\frac{2zw_i}{\sigma_n^2}}\right)\ \mathrm{d}z. \tag{10}$$

In the baseband BPSK-MUPA transmission system (Fig. 1), the bit at position i is transmitted only in every M^{th} pulse. Consequently, in the case of MUPA the channel capacity $\mathcal{C}_{\mathrm{MUPA}}$ is the average value of the M capacities $\mathcal{C}^{(i)}$

$$\mathcal{C}_{\mathrm{MUPA}} = \frac{1}{M} \cdot \sum_{i=1}^{M} \mathcal{C}^{(i)}, \tag{11}$$

which can be compared to the channel capacity $\mathcal{C}_{\mathrm{BPSK}}$ of conventional BPSK, i.e., $w_i = 1.0\ \forall\ i$. Fig. 4 depicts the channel capacity $\mathcal{C}$ for a conventional BPSK system and BPSK-MUPA with M = 4 bits and a natural binary index assignment. In the case of MUPA, the channel capacity $\mathcal{C}_{\mathrm{MUPA}}$ is reduced compared to $\mathcal{C}_{\mathrm{BPSK}}$ for the unweighted conventional BPSK system.

5. Conclusions

Although the average BER is increased and the channel capacity is reduced, the speech quality in terms of the parameter SNR is improved by up to 2.6 dB for BPSK-MUPA compared to systems with constant symbol energy. In this paper we consider hard decision decoding, but MUPA can also be extended with soft demodulation using *a priori* knowledge 1st order [9].

References

[1] J. Hagenauer. Rate-Compatible Punctured Convolutional Code (RCPC Codes) and their Applications. *IEEE Trans. Commun.*, April 1988.

[2] T. Brüggen and P. Vary. Unequal Error Protection by Modulation with Unequal Power Allocation. *IEEE Communications Letters*, June 2005.

[3] T. Brüggen, C. Schulte-Hillen, and P. Vary. Soft Demodulation and Unequal Error Protection for Digital Modulation Schemes. In *Proc. of IEEE ICASSP*, Philadelphia, PA, USA, March 2005.

[4] E. Bedrosian. Weighted PCM. *IRE Transactions on Information Theory*, March 1958.

[5] N. Görtz and E. Bresch. Source-Adaptive Power Allocation for Digital Modulation. *IEEE Communications Letters*, 7:569–571, December 2003.

[6] T. Brüggen and P. Vary. Analysis of Digital Modulation with Unequal Power Allocation. In *Proc. of IEEE VTC Spring*, Stockholm, Sweden, May 2005.

[7] C. E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, July 1948.

[8] T. Berger. *Rate Distortion Theory - A Mathematical Basis for Data Compression*. Prentice Hall Inc., 1971.

[9] T. Brüggen, M. Adrat, and P. Vary. BPSK Modulation with Unequal Power Allocation Considering A Priori Knowledge. In *Proc. of IEEE PIMRC*, Berlin, Germany, September 2005.