Joint Turbo Coding and Source-Controlled Modulation of Cycle-Stationary Sources in the Bandwidth-Limited Regime

Idoia Ochoa[1], Pedro M. Crespo[1], Javier Del Ser[2], and Mikel Hernaez[1]

[1] CEIT and TECNUN (University of Navarra), 20009 San Sebastian, Spain
{iochoa,pcrespo,mhernaez}@ceit.es
[2] TECNALIA-TELECOM, 48170 Zamudio, Spain
jdelser@robotiker.es

Abstract. In this paper we propose a novel one-layer coding/shaping transmission system for the bandwidth-limited regime based on single-level codes and sigma-mapping [1]. Specifically, we focus on cycle-stationary information sources with independent symbols. High spectral efficiencies can be achieved by combine at the transmitter a Turbo code with a sigma-mapper. Furthermore, the encoded symbols are modulated by using an asymmetric energy allocation technique before entering the aforementioned sigma-mapper. The corresponding decoder iterates between the Turbo decoder and the sigma-demapper, which exchange progressively refined *extrinsic* probabilities of the encoded symbols. For the Additive White Gaussian Noise (AWGN) channel, simulation results obtained for very simple Turbo codes show that the proposed system attains low bit error rates at signal-to-noise ratios relatively close to the corresponding Shannon limit. These promising results pave the way for future investigations towards reducing the aforementioned energy gap, e.g. by utilizing more powerful Turbo codes.

Keywords: Turbo codes, sigma-mapping, bandwidth-limited regime, unequal energy allocation, cycle-stationary sources.

1 Introduction

We consider the transmission, over the Additive White Gaussian Noise (AWGN) channel, of binary symbols generated by cycle-stationary random processes, $\{T_k\}_{k=1}^{\infty}$, with independent symbols. This kind of processes may arise, for instance, when the output sequence generated by a binary stationary source with memory[1] is partitioned into blocks of K symbols, before being processed by the block-sorting Burrows-Wheeler Transform (BWT) [2] of length K. For large K, it can be shown [3] that the corresponding random process $\{T_k\}_{k=1}^{\infty}$ at the output of the BWT can be asymptotically approximated by a cycle-stationary

[1] A source with memory may be modeled by either a Markov Chain (MC) or a Hidden Markov Model (HMM).

P. Chatzimisios et al. (Eds.): MobiLight 2010, LNICST 45, pp. 618–625, 2010.

random process with time period K (i.e. the length of the BWT input block), and independent symbols inside each output block. Under these conditions, the output process is completely specified once the probability distribution of each of the symbols inside an arbitrary block, say the first block, are known, i.e., when $P_{T_k}(t)$, for $k = 1, \ldots, K$ are known. Notice that the random sequence $\{T_k\}_{k=1}^{K}$ is non-stationary.

The entropy rate of such a process can be computed as

$$\mathcal{H}(\mathcal{T}) = \lim_{n\to\infty} \frac{1}{nK} H(T_1, \ldots, T_{nK}) = \frac{1}{K} H(T_1, \ldots, T_K) = \frac{1}{K} \sum_{k=1}^{K} H(T_k), \quad (1)$$

where $H(T_1, \ldots, T_K)$ denotes the entropy of the random vector $\{T_k\}_{k=1}^{K}$ with joint distribution $P_{\mathbf{T}}(t_1, \ldots, t_K) = \prod_{k=1}^{K} P_{T_k}(t_k)$. In what follows, we will denote by $P^0(k) = P_{T_k}(t = 0)$, and the set of values $\{P^0(k)\}_{k=1}^{K}$ inside a non-stationary block will be referred to as the *zero probability profile*. Notice that by cycle-stationarity $P^0(lK + k) = P^0(k)$, $\forall l \in \mathbb{N}$.

By the Shannon Source-Channel Coding Theorem [4], the minimum average energy per source symbol E_{so} required for reliable communication of $\{T_k\}_{k=1}^{\infty}$ over an AWGN channel is given by

$$\frac{E_{so}}{N_0} > \frac{2^{2R\mathcal{H}(\mathcal{T})} - 1}{2R}, \quad (2)$$

where N_0 is the one-sided noise power spectral density, R denotes the transmission rate (source symbols per channel symbol), and $2R$ is the spectral efficiency (binary source symbols per two dimensions). When the system has a spectral efficiency equal or greater than 2, it operates in the bandwidth-limited regime. Otherwise, it is said that the system works in the power-limited regime.

By the Separation Theorem, the lower limit in expression (2) can be achieved by an ideal source encoder followed by a capacity achieving channel code. However, in this paper we propose a novel scheme, suitable for the bandwidth-limited regime, which do not use a source encoder but rather uses the distribution $P^0(k) \doteq P_{T_k}(t = 0)$ (*zero probability profile*) of the source symbols T_k to modify the BPSK constellation at the output of the Turbo encoder before entering the sigma-mapper [1]. In this context, the present paper can be viewed as an extension of the scheme proposed in [5], suitable in the power-limited regime, for the transmission of the symbols generated by a stationary source with memory over the AWGN channel. Preliminary simulation results with very simple Turbo codes show a Bit Error Rate performance relatively close to the associated Separation limit, which sets the scene for future research aimed at narrowing this performance gap.

The rest of the paper is organized as follows: Section 2 and 3 describe the encoding and decoding process, respectively. Simulation results are presented in Section 4, and finally, some concluding remarks are drawn in Section 5.

2 Encoding Process

Figure 1 shows the proposed system, which combines Turbo coding and shaping in a one-layer scheme. The cycle-stationary source $\mathcal{T}$ of period K, generates blocks of K independent symbols $\{T_k\}_{k=1}^{K}$ having a *zero probability profile* $\{P^0(k)\}_{k=1}^{K}$. The source symbols are first encoded by a Turbo code of rate $R_c = K/N$. The encoded sequence $\mathbf{c}$ of length N is next interleaved to form the sequence $\tilde{\mathbf{c}}$ of the same length. The interleaved block $\tilde{\mathbf{c}}$ is then transformed by a serial-to-parallel converter in $\mathcal{I}$ sequences $\mathbf{v}^{(i)}$ of length L, where $0 \leq i \leq (\mathcal{I}-1)$ and $N = L \cdot \mathcal{I}$.

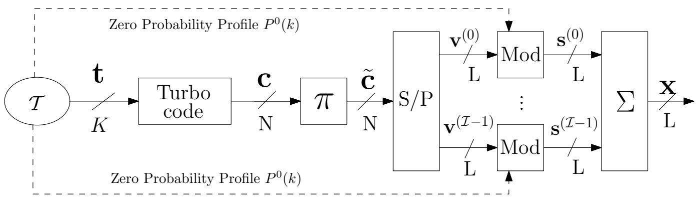


Fig. 1. Proposed transmission scheme

Before entering the sigma-mapper, the modulator assigns different amplitudes to the encoded symbols, depending on their position and their systematic or parity nature. The proposed energy allocation scheme (further detailed in Subsection 2.1) renders a set of $\mathcal{I}$ non-binary sequences $\mathbf{s}^{(i)}$, which are next fed to the sigma-mapper Σ [1]. The underlying idea of the sigma-mapper hinges on imposing a gaussian distribution on the output amplitude signal $\mathbf{x}$ of length L while satisfying, at the same time, the energy constraint[2] $(1/L) \cdot \sum_{l=1}^{L} \mathbb{E}\{|X_l|^2\} = E_c$. Finally, the destination receives a corrupted version of the amplitude sequence $\mathbf{x}$, denoted as $\mathbf{y} = \mathbf{x} + \mathbf{n}$, where $\mathbf{n}$ denotes a L-length sequence of i.i.d. Gaussian random variables with zero mean and variance per dimension $N_0/2$.

2.1 Asymmetric Energy Allocation Scheme

From expression (2), the minimum average energy per channel symbol E_c for reliable communication of the data generated by the binary cycle-stationary source $\mathcal{T}$ is given by[3]

$$\frac{E_c}{N_0} > \frac{2^{2R\mathcal{H}(\mathcal{T})} - 1}{2}, \tag{3}$$

with

$$\mathcal{H}(\mathcal{T}) \doteq \frac{1}{K} \sum_{k=1}^{K} h_b(P^0(k)), \tag{4}$$

[2] $\mathbb{E}\{\cdot\}$ stands for *expectation*.

[3] Notice that in equation (3), R refers to the overall rate of the system, which may differ from R_c (coding rate).

where $h_b(p) \doteq -p\log_2 p - (1-p)\log_2(1-p)$. However, since in our case the symbols inside a block are non-stationary, each output symbol will require a different average energy $E_c(k)$ depending on its distribution $P^0(k)$. From expression (3), the corresponding lower limit will be given by

$$E_c^*(k) = (2^{2Rh_b(P^0(k))} - 1)\frac{N_0}{2}, \tag{5}$$

and the minimum average energy per block as

$$E_c^* = \frac{1}{K}\sum_{k=1}^{K} E_c^*(k). \tag{6}$$

The energies used to modulate the encoded symbols are now given by $E(k) = \beta E_c^*(k)$, for $k = 1, \ldots, K$ where $\beta > 1$ is a scaling factor. Following the scheme of [5], the amplitudes of the encoded systematic symbols depend on both their value and the associated *a priori* probability $P^0(k)$, whereas the amplitude of a given encoded parity symbol is driven by 1) its value and 2) the value and the *a priori* probability of the associated systematic bit. In particular, the amplitudes are set to:

$$\text{Systematic symbols:} \begin{cases} -\sqrt{\frac{1-P^0(k)}{P^0(k)}E(k)}, \text{ if } u_k = 0, \\ +\sqrt{\frac{P^0(k)}{1-P^0(k)}E(k)}, \text{ if } u_k = 1. \end{cases} \tag{7}$$

$$\text{Parity symbols:} \begin{cases} -\sqrt{\frac{\theta}{P^0(k)}E(k)}, \text{ if the parity symbol is 0 and } u_k = 0, \\ -\sqrt{\frac{1-\theta}{1-P^0(k)}E(k)}, \text{ if the parity symbol is 0 and } u_k = 1, \\ +\sqrt{\frac{\theta}{P^0(k)}E(k)}, \text{ if the parity symbol is 1 and } u_k = 0, \\ +\sqrt{\frac{1-\theta}{1-P^0(k)}E(k)}, \text{ if the parity symbol is 1 and } u_k = 1. \end{cases} \tag{8}$$

where the arbitrary parameter θ ($0 \leq \theta \leq 1$) is chosen to maximize the performance of the system, which is usually achieved when $\theta = 0.5$ [6]. In the simulations presented in this paper, θ was set to 0.5. Notice that the resulting constellation is not symmetric, since more energy is allocated to those symbols with less *a priori* probability. Also observe that for the sake of clarity, the above expressions (7) and (8) do not include the time index mappings due to the Turbo and π interleavers, which must be considered in the energy allocation procedure.

A better estimation of the energies $E_c^*(k)$ defined in equation (5) can be obtained by taking into account the loss in performance due to using a non-capacity-achieving channel code, i.e. the actual gap to the corresponding Shannon limit. This is done by introducing, into expression (5), a gap factor $\Gamma(P^0(k))$, i.e.

$$E_c^*(k) = (2^{2Rh_b(P_s^0(k))} - 1)\frac{N_0}{2}\Gamma(P^0(k)). \tag{9}$$

The gap $\Gamma(p)$, $p \leq 0.5$, is a function of the *a priori* probability and depends on the particular communication scheme being used. Its value should be computed off-line by Monte Carlo simulations. Once $\Gamma(\cdot)$ is known, the amplitudes of the encoded symbols are calculated by following expressions (9), (6), (7) and (8), with $E(k) = \beta E_c^*(k)$ for a given scaling factor β.

3 Decoding Process

The decoder is depicted in Figure 2. It iterates between the sigma-demapper, which introduces the *a priori* probabilities of the source symbols, and the Turbo decoder, which is based on applying the message passing Sum-Product Algorithm (SPA) over the factor graph that describes the Turbo code [7].

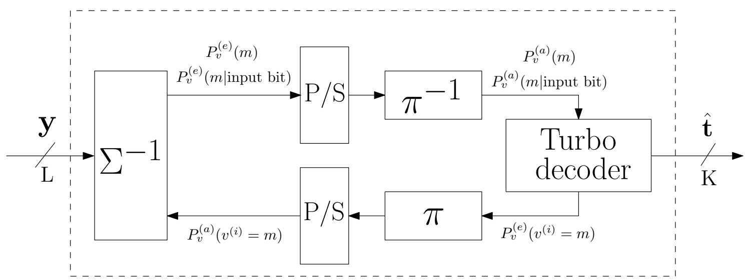
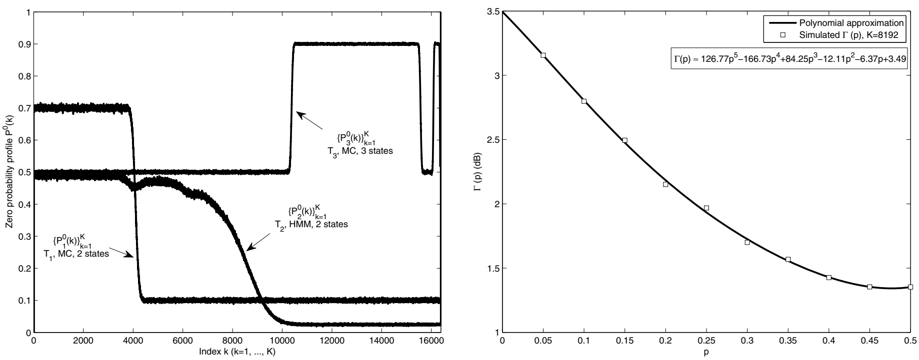


Fig. 2. Decoding scheme

The iterative decoding process starts from the sigma-demapper $\sum^{-1}$, which estimates the probabilities of the symbols contained in each of the $\mathcal{I}$ sequences $\mathbf{v}^{(i)}$ from the received sequence $\mathbf{y}$. These probabilities are denoted by $P_v^{(e)}(m)$ for systematic bits ($m \in \{0, 1\}$), and by $P_v^{(e)}(m|\text{input bit})$ for parity bits. Once these probabilities are computed, the Turbo decoder incorporates them (through the parallel-to-serial converter and the deinterleaver π^{-1}) as *a priori* information on the systematic and parity encoded symbols. Then, the SPA applied to the factor graph of the Turbo code produces a set of refined *a posteriori* and *extrinsic* probabilities[4]; the latter are then fed back to the sigma-demapper as *a priori* probabilities, giving rise to a new iteration. At each iteration, an estimation $\widehat{T}_k$ of T_k can be obtained for $k = 1, \ldots, K$ by performing a hard-decision over the *a posteriori* probabilities generated by the Turbo decoder. The decoding process is stopped after a fixed number of iterations Ψ. It is important to observe that, in our scheme, the SPA applied to the Turbo factor graph is extended with respect to the conventional forward and backward recursions (see [7, Section IV.A]) so as to incorporate the generation of the *extrinsic* probabilities corresponding to the parity bits.

[4] In the Turbo processing jargon, *extrinsic* refers to the fraction of the output probabilistic information that does not depend on any input *a priori* probability.



(a) (b)

Fig. 3. (a) Zero probability profiles of the considered MC and HMM sources concatenated with the BWT, and $K = 16384$. (b) Polynomial approximation of the gap margin functions for the proposed scheme and $K = 8192$.

4 Simulation Results

In order to study the performance of the proposed system, we have considered three different cycle-stationary sources $\mathcal{T}_i$ with *zero probability profiles* $\{P_i^0(k)\}_{k=1}^{K}$, with $i \in \{1, 2, 3\}$ and $K = 16384$. They have been obtained by estimating the probability distribution at the output of the BWT for an input source with memory following a Markov Chain (MC) for $i \in \{1, 3\}$, and a Hidden Markov Model (HMM) for $i = 2$. For $i = 1$ and $i = 2$, we have utilized the MC and the HMM employed in [5], both having two states and entropy rates 0.58 and 0.62 bits per source symbol, respectively. On the other hand, for $i = 3$ we have selected a MC with 3 states and entropy rate 0.83 bits per source symbol. The corresponding *zero probability profiles* at the output of the BWT are shown in Figure 3.a, which have been estimated by using frequency of occurrence as in [8, Expression (57)]. The Turbo code from [1, Example A] with rate $R_c = 1/3$ and generator polynomial $G(D) = 1/(1 + D)$ has been adopted for our simulations. The serial-to-parallel converter outputs $\mathcal{I} = 3$ sequences $\mathbf{v}^{(i)}$, and therefore the overall transmission rate R (source symbol per dimension) is 1, i.e. we are working in the bandwidth-limited regime (spectral efficiency of 2). The corresponding Shannon limits E_{so}/N_0 when $R = 1$ are -2.13 dB ($\mathcal{T}_1$), -1.69 dB ($\mathcal{T}_2$) and 0.293 dB ($\mathcal{T}_3$).

Figure 3.b depicts the gap factor $\Gamma(p)$, $p \leq 0.5$, used in expression (9). This gap has been calculated by simulating the proposed scheme for a stationary i.i.d. source with symbol distribution $(p, 1-p)$ and blocklength $K = 8192$ ($\Box$ markers). The figure also includes a 5^{th}-order polynomial approximation for the simulated $\Gamma(p)$, which can be easily programmed beforehand to produce the gamma function for any arbitrary value of p. Observe that $\Gamma(p)$ is a monotonically decreasing function of p, which indicates that the performance of the proposed setup with

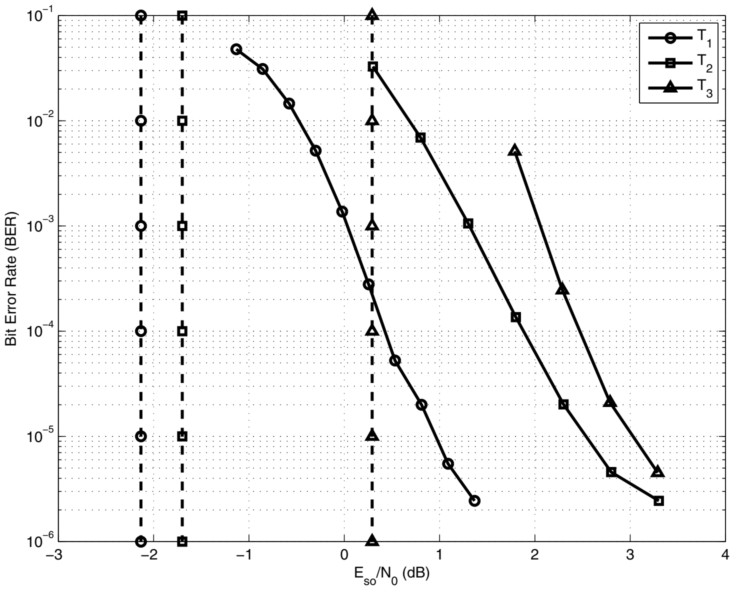


Fig. 4. BER vs E_{so}/N_0 of the proposed scheme for $\{\mathcal{T}_i\}_{i=1}^{3}$

i.i.d. stationary sources degrades as the distribution of the source symbols is more asymmetric.

Finally, Monte Carlo simulations have been performed for a blocklength of $K = 16384$ source symbols, $\Psi = 50$ decoding iterations and 1000 source sequences per simulated point. Figure 4 plots the Bit Error Rate (BER) versus E_{so}/N_0 for the three probability profiles and by using the energy allocation technique introduced in Section 2.1. However, we have modified the way the set of energies $E_c^*(k)$ is calculated, because simulations have empirically shown a performance improvement when R is replaced by R_c in expression (9). When the source is generated with the first probability profile $\{P_1^0(k)\}_{k=1}^{K}$ ($\mathcal{T}_1$), the proposed system is 2.55dB away of the theoretical limit for a BER of 10^{-4}. For the second probability profile ($\mathcal{T}_2$), the system is 3.57 dB away of its corresponding Shannon limit. Finally, the system applied to the third source $\mathcal{T}_3$ performs at 2.17 dB away from the corresponding Shannon limit at the same BER level. Notice that these results have been obtained by means of a very simple Turbo code. We believe that by optimizing the Turbo code better results (i.e. closer to the Shannon limit) could be obtained.

5 Concluding Remarks

We have proposed a novel scheme for the transmission of cycle-stationary sources over AWGN channels in the bandwidth-limited regime. The novel scheme is based on the combination of a Turbo encoder and a sigma-mapper that jointly perform the source and channel coding task, and on the use of an asymmetric waterfilling energy allocation technique. The simulation results obtained for very simple Turbo codes state that, for a variety of cycle-stationary sources, the BER performance of our proposed scheme gets close to the corresponding Shannon separation limit. These promising results motivate further research towards narrowing the aforementioned gap, e.g. by utilizing Turbo codes with enhanced BER waterfall performance.

Acknowledgments

This work was supported in part by the Spanish Ministry of Science and Innovation through the CONSOLIDER-INGENIO 2010 programme (CSD200800010, www.comonsens.org), and through the Torres-Quevedo funding program (PTQ-09-01-00740).

References

1. Ma, X., Ping, L.: Coded Modulation Using Superimposed Binary Codes. IEEE Transactions on Information Theory 50(12) (2004)
2. Burrows, M., Wheeler, D.: A Block Sorting Lossless Data Compression Algorithm, Digital Systems Center, Research Report 124 (1994)
3. Visweswariah, K., Kulkarni, S., Verdu, S.: Output Distribution of the Burrows-Wheeler transform. In: IEEE International Symposium on Information Theory, Sorrento, Italy (2000)
4. Shannon, C.E.: A Mathematical Theory of Communication. Bell System Technical Journal 27, 379–423, 623–656 (1948)
5. Del Ser, J., Crespo, P.M., Esnaola, I., Garcia-Frias, J.: Source-Controlled Turbo Coding of Sources with Memory using the Burrows-Wheeler Transform. In: IEEE International Symposium on Turbo Codes, Munich, Germany (2006)
6. Cabarcas, F., Souza, R.D., Garcia-Frias, J.: Turbo Coding of Strongly Non-Uniform Memoryless Sources with Unequal Energy Allocation and PAM Signaling. IEEE Transactions on Signal Processing 54(5), 1942–1946 (2006)
7. Kschischang, F.R., Frey, B.J., Loeliger, H.-A.: Factor Graphs and the Sum-Product Algorithm. IEEE Transactions on Information Theory 47(2), 498–519 (2001)
8. Crespo, P.M., Loyo, E., Del Ser, J.: Uncoded Optimal Binary Communication for Sources with Memory. Elsevier AEU International Journal of Electronics and Communications 62(8), 597–609 (2007)