Title	Error Resistant Lossless Data Compression with Equal Length Coding using Fine Tuned Multiple Label Mapping
Author(s)	Tervo, Valtteri; Fukawa, Kisho; Matsumoto, Tad
Citation	2012 IEEE International Conference on Communications (ICC): 4707-4711
Issue Date	2012-06
Type	Conference Paper
Text version	author
URL	http://hdl.handle.net/10119/10890
Rights	This is the author's version of the work. Copyright (C) 2012 IEEE. 2012 IEEE International Conference on Communications (ICC), 2012, 4707-4711. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
Description	

Error Resistant Lossless Data Compression with Equal Length Coding using Fine Tuned Multiple Label Mapping

Valtteri Tervo*+, K. Fukawa+, Tad Matsumoto*+
{wade, tadashi.matsumoto}@ee.oulu.fi
*Centre for Wireless Communications, University of Oulu
P.O. Box 4500, 90014 University of Oulu, Finland.
+Japan Advanced Institute of Science and Technology
1-1 Asahi-Dai, Nomi, Ishikawa, 923-1292 Japan.

***Abstract*—Data inherently including memory can be compressed using either equal or variable length coding. This paper proposes a novel method to compress the data with memory by using equal length code words while maintaining the error correction capability. The main drawback of variable length coding (VLC) is that because of the boundary problem, it is very sensitive to errors, which results in many cases in long burst errors due to error propagation. However, VLC is optimal in the sense of minimizing the expected length of the codeword. To reduce the compression rate using equal length code, a technique called multiple label mapping (MLM) is employed in this paper. Furthermore, MLM is extended with a fine tuning block to achieve compression rates as close to the Shannon limit as possible. Simulation results for symbol error rate (SER) evaluation as well as EXIT analysis indicate that the proposed technique can achieve compression rates very close to the VLC performance while maintaining the error correction capability. Furthermore, it is shown that the SER performance of the proposed joint source-channel coding scheme is very close to the Shannon limit.**

I. Introduction

Lossless source codes, e.g., Huffman codes [1] or arithmetic codes [2], can compress data very efficiently, and reach very close to the theoretical limit of the compression rate, i.e., the entropy of the source. However, Huffman and arithmetic codes utilizes the appearance probabilities of the symbols to be encoded, and thereby the code words have different lengths, depending on the appearance probabilities of the source symbols. The codes of this category is referred to as variable length entropy encoding. However, the main drawback of variable length coding (VLC) is that because of boundary problems, it is very sensitive to the bit errors happening in the channel, which results in many cases in long burst errors due to error propagation.

Shannon's source-channel separation theorem states that as long as the entropy of the source is less than the channel capacity, there exists a separable source and channel coding (SCC) scheme which allows transmission over the channel with an arbitrarily low probability of error. However, this theorem assumes that the source can be treated as a stationary stochastic process which satisfies the asymptotic equipartition property (AEP). Moreover, Shannon's source compression limit assumes the use of infinitely long sequences.

The mobile access of multimedia data, such as video and audio, is one of the key applications in the current and future generation wireless communications. In such applications, source compression is usually performed using well established techniques, which, in order to achieve high compression gains, often employ VLC. In delay- and/or complexity-constrained transmission scenarios, joint optimization of SCC techniques are often more advantageous than the classical separation based source and channel encoding techniques. Hence, joint SCC (JSCC) schemes has been a core issue during the past several decades [3], [4].

In many approaches to JSCC, the implicit residual source redundancy after source encoding is additionally used for error protection in the decoder. This is useful to reduce the bandwidth and latency for the overall transmission system because very powerful forward error correction coding (FEC), requiring heavy computational burden for decoding, is not necessary [5]. Hagenauer [3] proposes JSCC by combining the trellis diagrams of the source model and the channel code for Viterbi decoding. In [6], [7], residual redundancy of the source is used in JSCC by exploiting the memory structure of hidden Markov models (HMMs). Recently, a technique called *over-complete source-mapping* is proposed for joint optimization of iterative source and channel decoding (SCD) in [8]. VLC is employed on JSCC in [5], [9]. A combined utilization of JSCC and SCD for multiple correlated sources is investigated in [10]–[12].

Conventional source coding under a distortion constraint and channel coding under resource constraints of the channels, such as power and/or spectrum availability, have long been considered as information-theoretic duals of each other, starting from Shannon's 1959 paper [13]. In [14], an exact

This research was carried out in the framework of the project Distributed Decision Making for Future Wireless Communication Systems (DIDES) which was funded by the Finnish Funding Agency for Technology and Innovation (TEKES). This work was also in part supported by the Japanese government funding program, Grant-in-Aid for Scientific Research (B), No. 23360170. This research has also been supported by the Nokia Foundation.

characterization of the Shannon duality between data transmission and compression through Lagrange duality in convex optimization is provided. This suggests that excellent channel codes may also be excellent source codes. Garcia-Frias *et al.* proposes the use of punctured turbo codes for compression of binary sources in [15]. Since then, a lot of work has been conducted related to the utilization of channel codes in source compression. For example in [16]–[22] channel codes are applied in source compression by puncturing, i.e., eliminating bits after the channel encoder.

In this paper, a novel method to remove the error propagation problem in source coding without imposing information loss is proposed. The idea of multiple label mapping (MLM) [23] is extended by introducing a fine tuning (FT) block, which allows the proposed technique to achieve compression rates arbitrarily close to the limit. FT consists of variable nodes [24], which enable redundancy increment, and check nodes [24], which enable redundancy decrement. Furthermore, an accumulator (ACC), i.e., rate one recursive convolutional code [25] with polynomial $1/(1+D)$, is employed before MLM to make the binary bit appearance probabilities equal. The bit appearance has to be equiprobable in order to produce entropy achieving equal length codes. In this paper, the combined block of ACC, FT and MLM is referred to as FT-MLM for notational convenience. Usually, entropy achieving source codes employ VLC, but in contrast to the previous work, FT-MLM uses equal length code words, and hence no code word boundary problem occurs. FT-MLM is actually a form of joint source-channel coding because it can tolerate and even correct errors by utilizing the memory structure of the source in the joint decoding process.

Determining the labeling map of MLM has a crucial role on the performance of the proposed scheme. Adaptive binary switching algorithm (ABSA) [26] is employed for finding the most suitable mapping rule for MLM. It utilizes binary switching algorithm (BSA) [27] with constraints on extrinsic mutual information provided by outer decoder representing the memory structure of the source. The outer decoder in this paper is BCJR decoder [28] of the source, which allows the exploitation of the correlation of a Markov source.

The rest of this paper is organized as follows. The whole system model as well as detailed description of FT-MLM is presented in Section II. In Section III, a cost function for ABSA is derived for finding the most suitable mapping rule for MLM. Numerical results are given in Section IV. Section V concludes the paper with some concluding statements.

II. System model

Fig. 1 presents a block diagram of the system considered in this paper. A binary transition emitting Markov (TEM) source is assumed in this paper, however, extension to non-binary TEM within the same theoretical framework is straightforward. In Fig. 1, source parameters p and q are the state transition probabilities; the state changes from S_0 to S_1 with probability p and from S_1 to S_0 with probability q. At every state transition, the source emits a symbol comprised of two bits.

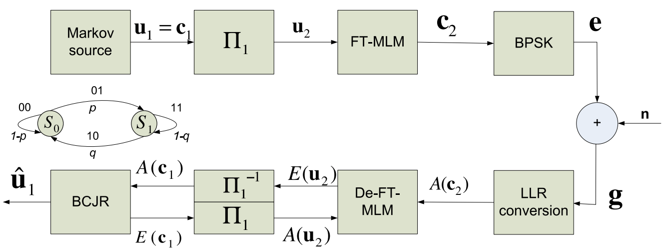

Fig. 1. The block diagram of the system model.

The bit sequence $\mathbf{c}_1$ emitted from the Markov source is passed through the random interleaver Π_1, which shuffles the bits randomly and breaks the memory of the TEM source. After Π_1 the corresponding bit sequence $u_{2,1}, u_{2,2}, \ldots, u_{2,L}$ is compressed by FT-MLM, resulting in compressed sequence $c_{2,1}, c_{2,2}, \ldots, c_{2,R_sL}$, where R_s is the compression rate of the FT-MLM algorithm. The compressed sequence is modulated with binary phase shift keying (BPSK), resulting a symbol sequence $\mathbf{e}$ and sent through an additive white Gaussian noise (AWGN) channel. The receiver receives the noisy bit sequence $g_j = e_j + n_j$, $j = 1, 2, \ldots, R_sL$, where $e_j \in \{-1, 1\}$ is the transmitted sequence and $n_j \in \mathbb{R}$ is zero mean AWGN component with variance σ_n^2. In the receiver, the channel values are converted to LLRs and the inverse operation of FT-MLM, referred as De-FT-MLM is performed. De-FT-MLM produces extrinsic LLR sequence $E(\mathbf{u}_2)$, which is deinterleaved and provided to the BCJR source decoder as *a priori* information $A(\mathbf{c}_1)$. The BCJR decoder exploits the memory structure of the source and produces the extrinsic LLRs $E(\mathbf{c}_1)$, which are interleaved and fed back to De-FT-MLM as *a priori* information $A(\mathbf{u}_2)$. Since iterative structure is employed, matching of the mutual information transfer characteristics between the constituent soft input / soft output (SISO) decoders plays crucial roles to achieve arbitrary low SER as well as to minimize the bit rate after compression.

A. Fine Tuned Multiple Label Mapping

The block diagram of FT-MLM is shown in Fig. 2. First, the incoming bit sequence $\mathbf{u}_2$ is passed through ACC which makes the bit appearance probabilities equal and hence, allows us to design entropy achieving equal length codes. Proof for equal bit appearance probabilities after ACC is given in [29]. ACC produces a bit sequence $\mathbf{c}_2^{\text{ACC}}$ which is interleaved and compressed by using MLM with a compression rate R_s^{MLM}. This is defined as

$$R_s^{\text{MLM}} = \arg\min_{(R_s^{\text{MLM}})'} \{|(R_s^{\text{MLM}})' - R_s|\}, \tag{1}$$

where $(R_s^{\text{MLM}})' \in \{(R_s^{\text{MLM}})'_j\}$, $j = 1, 2, \ldots, o$, are the compression rates achieved with MLM[1] and R_s is the final target compression rate of the system. In (1), we just take the $(R_s^{\text{MLM}})'$, which is closest to a given R_s. MLM produces a bit

[1] To achieve all compression rates with MLM one needs to determine infinite number of mapping schemes, such as it used in extended mapping [23]. For practical purposes we can choose fixed number of compression rates in the range from 0 to 1.

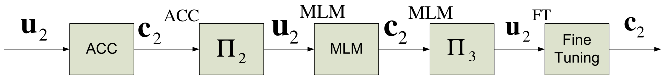

Fig. 2. Block diagram of FT-MLM.

sequence $\mathbf{c}_2^{\mathrm{MLM}}$, which is followed by interleaver Π_3 resulting in a bit sequence $\mathbf{u}_2^{\mathrm{FT}}$, which is input to the FT block.

After MLM, the data may still have some redundancy, or the rate after the compression may be excessively low compared with the target compression rate R_{s}. FT is used to make the subtle adjustment of the rate such that the compression rate, as a whole, becomes as close to the entropy as possible: on one hand, when incremental redundancy is needed, FT works as a variable node encoder (VNE) [24] with the average variable node degree $\bar{d}_{\mathrm{v}}$; on the other hand, when the objective is to decrease the redundancy, FT works as a check node encoder (CNE) [24] with the average check node degree $\bar{d}_{\mathrm{c}}$. The compression rate of the FT block is now given by

$$R_{\mathrm{s}}^{\mathrm{FT}} = \frac{R_{\mathrm{s}}}{R_{\mathrm{s}}^{\mathrm{MLM}}}. \quad (2)$$

If $R_{\mathrm{s}}^{\mathrm{FT}} < 1$ FT is CNE and $\bar{d}_{\mathrm{c}} = \frac{1}{R_{\mathrm{s}}^{\mathrm{FT}}}$. The check node degrees are specified as follows: let D_{c} be the number of different check node degrees, and denote their degrees by $\tilde{d}_{\mathrm{c},i}$, $i = 1, \ldots, D_{\mathrm{c}}$. The average check node degree is calculated as [24]

$$\bar{d}_{\mathrm{c}} = \sum_{i=1}^{D_{\mathrm{c}}} a_{\mathrm{c},i} \tilde{d}_{\mathrm{c},i}, \quad (3)$$

where $a_{\mathrm{c},i}$ is the fraction of nodes having degree $\tilde{d}_{\mathrm{c},i}$. Since some of the check nodes has to have degree 1 [30], let $\tilde{d}_{\mathrm{c},1} = 1$. Otherwise, the convergence does not start, because there is zero *a priori* information to be provided to the De-MLM. The fraction $a_{\mathrm{c},1} \in [0, 1]$ can be chosen arbitrarily, however, it is reasonable to set $a_{\mathrm{c},1}$ at a relatively small value because it only determines the left-most point in the EXIT curve of FT. Given $\tilde{d}_{\mathrm{c},1}$ and $a_{\mathrm{c},1}$, one has to choose $D_{\mathrm{c}} \geq 3$ in order to achieve arbitrary compression rate. In order to keep the complexity smallest, we choose $D_{\mathrm{c}} = 3$. Moreover, $\tilde{d}_{\mathrm{c},3}$ is chosen to be the smallest integer satisfying $a_{\mathrm{c},1} + (1 - a_{\mathrm{c},1})\tilde{d}_{\mathrm{c},3} \geq \bar{d}_{\mathrm{c}}$, i.e.,

$$\tilde{d}_{\mathrm{c},3} = \left\lceil \frac{\bar{d}_{\mathrm{c}} - a_{\mathrm{c},1}}{1 - a_{\mathrm{c},1}} \right\rceil, \quad (4)$$

where $\lceil \cdot \rceil$ indicates the smallest integer larger than its argument. To achieve a desired value of $\bar{d}_{\mathrm{c}}$, let $\tilde{d}_{\mathrm{c},2} = \tilde{d}_{\mathrm{c},3} - 1$. With (3) and $\sum_{i=1}^{D_{\mathrm{c}}} a_{\mathrm{c},i} = 1$,

$$a_{\mathrm{c},2} = a_{\mathrm{c},1} + (1 - a_{\mathrm{c},1})\tilde{d}_{\mathrm{c},3} - \bar{d}_{\mathrm{c}}, \quad (5)$$

and

$$a_{\mathrm{c},3} = 1 - a_{\mathrm{c},1} - a_{\mathrm{c},2}. \quad (6)$$

have to be satisfied.

If $R_{\mathrm{s}}^{\mathrm{FT}} > 1$ FT is VNE and $\bar{d}_{\mathrm{v}} = R_{\mathrm{s}}^{\mathrm{FT}}$. The procedure for choosing the check node degrees and their distribution can also be used in determining the variable node degrees and their distribution. It should be noticed that if the equality $R_{\mathrm{s}}^{\mathrm{FT}} = 1$ holds, there is no need to use FT.

III. Mapping rule optimization

Determining the labeling map of MLM has a significant impact on the performance of the proposed scheme. We use ABSA to determine the labeling map. Initially, BSA was introduced for index optimization in vector quantization [27]. In [31], BSA was applied for finding the most suitable index assignments to arbitrary, high order signal constellations. However, BSA aims to find the optimal mapping pattern assuming that full *a priori* information is available. ABSA extends the BSA's optimization criterion such that *a priori* information is partially available. Moreover, it introduces a weight vector that takes into account which part of the demapper EXIT curve should be modified to achieve the desired result. EXIT constrained BSA (EBSA) [32] further extends the idea by introducing a constraint ϵ' which controls the gap between the two EXIT curves. Fundamental difference between EBSA and ABSA is that EBSA aims to minimize the gap between the two EXIT curves while ABSA only aims to push upwards the demapper EXIT curve. Hence, ABSA makes the convergence to the target MI point possible, given the decoder curve. In EBSA, code parameter optimization follows the BSA process, which means that the both decoder and demapper curves are changed wit the aim of rate loss minimization. In this paper, fixed decoder curve is assumed and hence, ABSA is used for labeling optimization.

Let l_{map} denote the length of the labeling map and let the cost of the labeling be denoted by $Z_{l_{\mathrm{ap}}}$, where $l_{\mathrm{ap}} = 0, 1, \ldots, l_{\mathrm{map}} - 1$ is the number of known bits. Let $\mathbf{u}_i \in \mathcal{U}$, $i = 1, 2, \ldots, 2^{l_{\mathrm{map}}}$, be the input binary vector and $\mathbf{s}_i \in \mathcal{S}$, $i = 1, 2, \ldots, n$ be the code word in MLM. The length of a binary code word is $m = \log_2 n$. Let $\mathrm{Prob}(\mathbf{s}_v \to \hat{\mathbf{s}}_v)$ denote the pair-wise probability that the code word $\hat{\mathbf{s}}_v$ is chosen instead of the transmitted code word $\mathbf{s}_v$. In order to calculate the Euclidean distance between the code words, the following transformation is performed:

$$\gamma_{k,i} = \eta(s_{k,i}) = 2s_{k,i} - 1, \quad (7)$$

$\forall i = 1, 2, \ldots, m$ and $\forall k = 1, 2, \ldots, n$, where $s_{k,i}$ is the i^{th} element of the code word $\mathbf{s}_k$. Now γ_k is located in the n-dimensional space $\mathcal{V}$, of which coordinates are the elements of γ_k. The labeling cost can now be expressed as

$$Z_{l_{\mathrm{ap}}} = \frac{1}{l_{\mathrm{map}} 2^{l_{\mathrm{map}}-1} 2^{l_{\mathrm{map}}-l_{\mathrm{ap}}-1}} \sum_{v=1}^{l_{\mathrm{map}}} \sum_{\mathbf{s}|u_v=0} \sum_{\hat{\mathbf{s}}|\hat{u}_v=1} exp(-SNR_{\mathrm{FT}} ||\eta(\mathbf{s}) - \eta(\hat{\mathbf{s}})||_2^2), \quad (8)$$

where the function $\eta(\mathbf{s})$ returns the point γ in the space $\mathcal{V}$ and v denotes the v^{th} bit of an input word for MLM. SNR_{FT} denotes the signal to noise power ratio (SNR) after FT decoder. If $R_{\mathrm{s}}^{\mathrm{FT}} > 1$,

$$SNR_{\mathrm{FT}} = \bar{d}_{\mathrm{v}} \cdot SNR_{\mathrm{ch}}, \quad (9)$$

where SNR_{ch} is the channel SNR. If $R_s^{\text{FT}} < 1$, the standard deviation of the LLRs provided by the channel is calculated as

$$\sigma_{\text{ch}} = \sqrt{4 \cdot SNR_{\text{ch}}}. \tag{10}$$

The extrinsic information provided by check node decoder (CND) can be calculated as [24]

$$I_{E(\mathbf{u}_3)} = 1 - J\left(\sqrt{\bar{d}_c} \cdot J^{-1}\left(1 - I_{A(\mathbf{g})}\right)\right), \tag{11}$$

where $I_{A(\mathbf{g})}$ is the MI provided by the channel and $J(\cdot)$ is the so called J-function detailed for example in [33]. In this paper the J-function and its inverse J^{-1} is calculated according to [34]. Now, the SNR after FT can be written as

$$SNR_{\text{FT}} = \frac{(J^{-1}(I_{E(\mathbf{u}_3)}))^2}{4}. \tag{12}$$

Mapping rule which enables the convergence can be found by performing ABSA with the cost function (8).

IV. Simulation Results

In this section, the performance of the proposed method is evaluated through simulations. For FT, $a_{c,1} = 0.1$ was used in the simulations. The block length was set at $L = 18000$ for convergence property evaluations through EXIT analysis, and $L = 250000$ for the trajectory and SER simulations.

We use the ABSA algorithm with a cost function derived in Section III to enable the convergence of the system. An example is shown in Fig. 3. If the labeling is optimized only with full *a priori* information the EXIT curves intersect. However, ABSA lifts up the parts of the De-FT-MLM curve under the BCJR curve, and thereby the intersection can be avoided.

SER results for the same parameters as in Fig. 3 is presented in Fig. 4. The labeling pattern is optimized for $E_b/N_0 = -1.5$ dB. It can be seen that our technique achieves SER$\leq 10^{-4}$ about 2.9 dB away in E_b/N_0 from the Shannon limit.

Fig. 5 shows the EXIT behavior of the system with a parameter set $p = 0.1549$, $q = 0.9121$, $R_s = 0.4970$, and $l_{\text{map}} = 4$ optimized for $E_b/N_0 = -2.4$ dB. It can be seen that the trajectory cannot reach the convergence point when $E_b/N_0 = -2.4$ dB due to the too small gap between the EXIT curves. However, as it can be seen in Fig. 6 the convergence threshold is between -2.1 dB and -2 dB.

V. Conclusions

In this paper a near capacity achieving joint source-channel coding (JSCC) scheme by using iterative decoding has been proposed for binary Markov source. A new compression technique, fine tuned multiple label mapping (FT-MLM) technique, was proposed which produces equal length code words. This completely avoids the error propagation problem which, with VLC, is likely to happen due to code word synchronization error.

A conclusion of [23] was that in theory the MLM can achieve the entropy of the source, if probability grouping is performed so that the appearance probability of each code

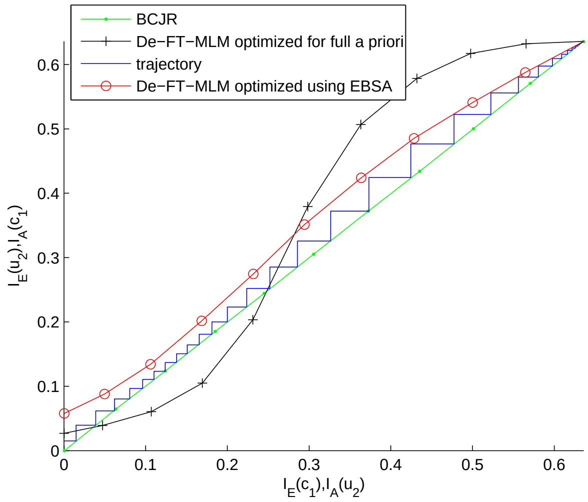

Fig. 3. EXIT chart of the system with $p = 0.8741$, $q = 0.1674$, $R_s = 0.4174$, E_b/N_0 =0 dB and $l_{\text{map}} = 5$.

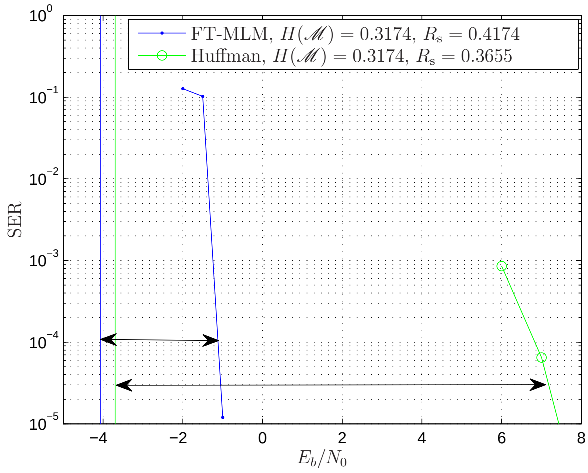

Fig. 4. SER of the proposed scheme and Huffman code when $l_{\text{map}} = 5$.

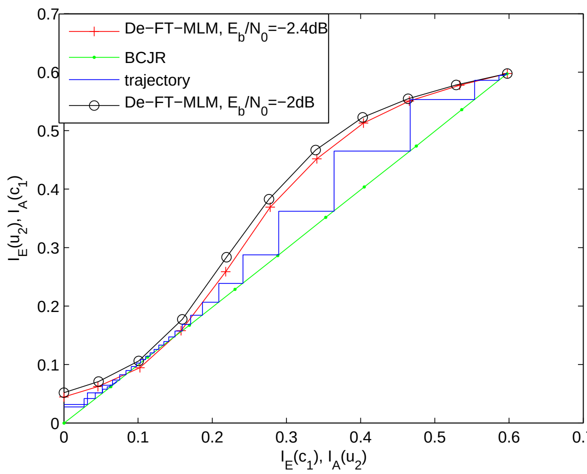

Fig. 5. EXIT chart of the system with $p = 0.1549$, $q = 0.9121$, $R_s = 0.4970$, and $l_{\text{map}} = 4$ optimized for $E_b/N_0 = -2.4$ dB.

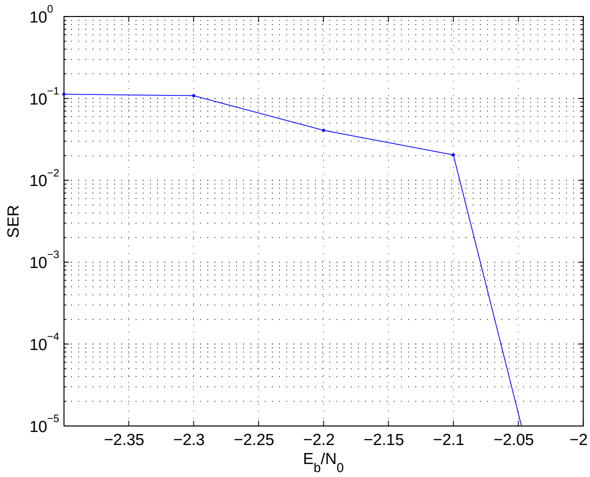

Fig. 6. SER of the proposed scheme with $p = 0.1549$, $q = 0.9121$, $R_\mathrm{s} = 0.4970$, and $l_\mathrm{map} = 4$ optimized for $E_b/N_0 = -2.4$ dB.

word after MLM is equal. In this paper, accumulator (ACC) is used to equalize the appearance probabilities of the code words. It was also indicated in [23] that the most crucial process of MLM is to find the best grouping of the source alphabet for a given source model. This problem has been solved in this paper by using the adaptive binary switching algorithm (ABSA). MLM can also be applied for quantization which is, however, left as future study.

REFERENCES

[1] D. A. Huffman, "A method for the construction of minimum-redundancy codes," in *Proc. Institute of Radio Engineers*, Sep. 1952.

[2] J. J. Rissanen, "Generalized kraft inequality and arithmetic coding," *IBM Journal of Research and Development*, vol. 20, no. 3, pp. 198–203, May 1976.

[3] J. Hagenauer, "Source-controlled channel decoding," *IEEE Trans. Commun.*, vol. 43, no. 9, pp. 2449–2457, 1995.

[4] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, USA, 542 p: John Wiley, 1991.

[5] R. Thobaben and J. Kliewer, "Low-complexity iterative joint source-channel decoding for variable-length encoded Markov sources," *IEEE Trans. Commun.*, vol. 53, no. 12, pp. 2054–2064, 2005.

[6] J. Garcia-Frias and J. D. Villasenor, "Joint turbo decoding and estimation of hidden Markov sources," *IEEE J. Select. Areas Commun.*, vol. 19, no. 9, pp. 1671–1679, 2001.

[7] N. Dütsch, S. Graf, J. Garcia-Frias, and J. Hagenauer, "Source model aided lossless turbo source coding," in *In: Proc. 4th Int. Symp. on Turbo Codes and Related Topics*, Munich, Germany, Apr. 3–7 2006, pp. 1–6.

[8] A. Q. Pham, L. L. Yang, and L. Hanzo, "Joint optimization of iterative source and channel decoding using over-complete source-mapping," in *Proc. IEEE Veh. Technol. Conf.*, Baltimore, MD, USA, Sep. 30–Oct. 3 2007.

[9] R. G. Maunder, J. Wang, S. X. Ng, L.-L. Yang, and L. Hanzo, "On the performance and complexity of irregular variable length codes for near-capacity joint source and channel coding," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1338–1347, 2008.

[10] J. Garcia-Fras, "Joint source-channel decoding of correlated sources over noisy channels," in *Proc. IEEE Data Compression Conf.*, Snowbird, UT, USA, Mar. 27-29 2001.

[11] J. D. Ser, P. M. Crespo, and O. Galdos, "Asymmetric joint source-channel coding for correlated sources with blind HMM estimation at the receiver," *EURASIP J. Wireless Comm. and Netw.*, vol. 2005, pp. 483–492, 2005.

[12] J. D. Ser, P. Crespo, and A. Muoz, "Joint source-channel decoding of correlated sources over ISI channels," in *Proc. IEEE Veh. Technol. Conf.*, Stockholm, Sweden, May 30–Jun. 1.

[13] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *IRE Nat. Conv. Rec.*, pp. 142–163, Mar. 1959.

[14] M. Chiang and S. Boyd, "Geometric programming duals of channel capacity and rate distortion," *IEEE Trans. Inform. Theory*, vol. 50, no. 2, pp. 245–258, Feb. 2004.

[15] J. Garcia-Frias and Y. Zhao, "Compression of binary memoryless sources using punctured turbo codes," *IEEE Commun. Lett.*, vol. 6, no. 9, pp. 394–396, Sep. 2002.

[16] G. Caire, S. Shamai, and S. Verd, "Noiseless data compression with low-density parity-check codes," *DIMACS Series in Discrete Mathematics ans Theoretical Computer Science*, pp. 1–22, Sep. 2004.

[17] J. Hagenauer, J. Barros, and A. Schaefer, "Lossless turbo source coding with decremental redundancy," in *Proc. Fifth Conf. Source and Channel Coding*, Erlangen, Germany, 2004.

[18] N. Dütch, "Code optimization for lossless turbo source coding," in *Proc. of 14th IST Mobile & Wireless Commun. Summit*, Dresden, Germany, Jun. 2005.

[19] M. Adrat, P. Vary, and T. Clevorn, "Optimized bit rate allocation for iterative source-channel decoding and its extension towards multi-mode transmission," in *Proc. 14th IST Mobile & Wireless Commun. Summit*, Dresden, Germany, Jun. 2005.

[20] N. Dütch, "Code optimization for lossless compression of binary memoryless sources based on fec codes," *European Trans. Telecommun.*, no. 17, pp. 219–229, 2006.

[21] J. Haghighat, M. R. Soleymani, and W. Hamouda, "Lossless source coding using repeat-accumulate codes," in *Proc. IEEE Int. Symp. Signal Processing and Its Applications*, Sharjah, United Arab Emirates, Feb. 12–15 2007.

[22] L. Schmalen, P. Vary, T. Clevorn, and M. Adrat, "Turbo source compression with jointly optimized inner irregular and outer irregular codes," in *Proc. IEEE Veh. Technol. Conf.*, Ottawa, Canada, Sep. 6–9 2010.

[23] V. Tervo, T. Matsumoto, and J. Karjalainen, "Joint source-channel coding using multiple label mapping," in *Proc. IEEE Veh. Technol. Conf.*, Ottawa, Canada, Sep.6–9 2010, pp. 1–6.

[24] S. ten Brink and G. Kramer, "Design of repeat-accumulate codes for iterative detection and decoding," *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2764–2772, Nov. 2003.

[25] P. Elias, "Coding for noisy channels," in *IRE Convention Record 3(4):37-46. (Reprinted in Key Papers in the Development of Coding Theory, ed. E. Berlekamp*, New York: IEEE Press, 1955, pp. 48–55).

[26] Z. Yang, Q. Xie, K. Peng, and J. Song, "Labeling optimization for bicm-id systems," *IEEE Commun. Lett.*, vol. 14, no. 11, pp. 1047–1049, 2010.

[27] K. Zeger and A. Gersho, "Pseudo-gray coding," *IEEE Trans. Commun.*, vol. 38, no. 12, pp. 2147–2158, Dec. 1990.

[28] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. 20, no. 2, pp. 284–287, 1974.

[29] V. Tervo, K. Fukawa, K. Tervo, and T. Matsumoto, "Error resistant lossless data copmpression with equal length coding using fine tuned multiple label mapping," *IEEE Trans. Commun.*, 2011 (under review).

[30] T. Yano and T. Matsumoto, "Arithmetic extended mapping for BICM-ID with repetition codes," in *Proc. ITG Workshop Smart Antennas*, Berlin, Germany, Feb.16–18 2009, pp. 1–8.

[31] F. Schreckenbach, N. Görtz, and J. Hagenauer, "Optimized symbol mappings for bit-interleaved coded modulation with iterative decoding," in *Proc. IEEE Global Telecommun. Conf.*, San Francisco, USA, Dec.1–5 2003, pp. 3316–3320.

[32] K. Fukawa, S. Ormsub, K. Anwar, A. Tölli, and T. Matsumoto, "EXIT-constrained BICM-ID design using extended mapping," *EURASIP J. Wireless Comm. and Netw.*, (to be submitted).

[33] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.

[34] F. Brännström, "Convergence analysis and design of multiple concatenated codes," Ph.D. dissertation, Chalmers University of Technology, Gothenburg, Sweden, 2004.