Turbo Codes for Non-Uniform Memoryless Sources over Noisy Channels

Guang-Chong Zhu and Fady Alajaji, *Senior Member, IEEE*

Abstract— **This work addresses the problem of designing Turbo codes for non-uniform binary memoryless or independent and identically distributed (i.i.d.) sources over noisy channels. The extrinsic information in the decoder is modified to exploit the source redundancy in the form of non-uniformity; furthermore, the constituent encoder structure is optimized for the considered non-uniform i.i.d. source to further enhance the system performance. Some constituent encoders are found to substantially outperform Berrou's (37,21) encoder. Indeed, it is shown that the bit error rate (BER) performance of the newly designed Turbo codes is greatly improved as significant coding gains are obtained.**

Index Terms — **Turbo codes, non-uniform i.i.d. sources, joint source-channel coding, AWGN and Rayleigh channels.**

I. Introduction

In channel coding, the source is usually assumed to be a binary memoryless Bernoulli (1/2) source, i.e., it generates uniform i.i.d. bit streams $\{d_k\}_{k=1}^{\infty}$, where $p_0 \triangleq Pr\{d_k = 0\} = 1/2$. In reality, however, natural sources (e.g., image and speech sources) often exhibit substantial amounts of redundancy in the form of memory and/or non-uniformity [1]; in this case, a source encoder would be used. An ideal source encoder would be able to eliminate all the source redundancy and hence produce a uniform i.i.d. sequence of bits for the channel encoder input. However, most existing source encoders (particularly fixed-length encoders) are suboptimal. As a result, the input to the channel encoder contains a certain amount of *residual* redundancy. For example, the 4.8 kbits/s FS1016 CELP speech vocoder produces line spectral parameters that contain 41.5% of residual redundancy due to non-uniformity and memory [2]. For uncompressed sources, many binary images (e.g., facsimile images) may contain up to 76% of redundancy due to non-uniformity (e.g. [6]), this corresponds to $p_0 = 0.96$. Therefore, transmission of sources with a considerable amount of residual or natural redundancy is an important issue. Several studies (e.g., [1], [9], etc.) have shown that appropriate use of the source redundancy can significantly improve the system performance.

Turbo codes have demonstrated excellent performance for uniform i.i.d. sources over additive white Gaussian noise (AWGN) [4] and Rayleigh fading [5] channels; the issue of using Turbo codes for non-uniform i.i.d. sources has not been systematically studied. In essence, this is a joint source-channel coding problem. In this work, we investigate the issue of designing Turbo codes for such sources sent over noisy channels. Our objective is to strive to come as close to the Shannon limit as possible. The extrinsic information in the Turbo decoder is modified to exploit the source redundancy. We also observe that while the original $(37, 21)$ Berrou code offers excellent ("waterfall") performance for uniform sources, it provides a relatively poor performance with a considerably high error floor when the source is non-uniform. Through a systematic search, we show that some constituent encoders substantially outperform Berrou's $(37, 21)$ encoders for a given non-uniform i.i.d. source. Significant coding gains are further achieved by combining this optimized encoder structure with the appropriately modified decoder.

Manuscript received September 28, 2001. The associate editor coordinating the review of this letter and approving it for publication was Prof. Marc Fossorier. This work was supported in part by NSERC of Canada and PREA of Ontario.

G.-C. Zhu is with Mathematics and Engineering, Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada (e-mail: zhugc@mast.queensu.ca).

F. Alajaji is with the Department of Mathematics and Statistics and with the Department of Electrical and Computer Engineering, Queen's University, Kingston, ON K7L 3N6, Canada (e-mail: fady@mast.queensu.ca).

II. Turbo Codes for Non-uniform I.I.D. Sources

A non-uniform i.i.d. binary source emits a sequence of bits $\{d_k\}_{k=1}^{\infty}$ with probability distribution described by p_0, $k = 1, 2, \cdots$. We consider designing Turbo codes for transmitting such a source over AWGN and Rayleigh fading channels. We modify the Turbo decoder to exploit the source redundancy. We also optimize the Turbo encoder structure with respect to the considered source.

A. Modifications of the Decoder Extrinsic Information

In the BCJR algorithm [3] used by the Turbo decoder, the soft output is decomposed into three terms:

$$\Lambda(d_k) = L_{ch}(d_k) + L_{ex}(d_k) + L_{ap}(d_k),$$

where $L_{ch}(d_k)$, $L_{ex}(d_k)$, and $L_{ap}(d_k)$ are the channel transition term, the extrinsic information term, and the *a priori* term, respectively, [4]. The extrinsic information produced by one constituent decoder is used as the *a priori* estimation for the other constituent decoder. At iteration one, for the first decoder, if $p_0 = 1/2$, which is considered in [4], we have $L_{ap}(d_k) = 0$.

When the source is non-uniform, $\log((1 - p_0)/p_0) \neq 0$ is used as the initial *a priori* input for the first decoder at the first iteration. As a result, in the output $\Lambda(d_k)$ produced by the first decoder, $L_{ap}(d_k) = \log((1 - p_0)/p_0)$. By passing this term together with the extrinsic information from the first decoder to the second, a BER performance gain is observed[1]. It can be shown via the BCJR algo-

[1]This simple modification was also briefly mentioned in [8] (Remark 4.d, p. 433) but not explicitly studied and evaluated. One of our goals in this paper is to explicitly assess the gains achieved by this method and examine how close we can come to the Shannon limit.

rithm's derivation that $\log((1-p_0)/p_0)$ will appear in the output $\Lambda(d_k)$ as an extra term. In our design, we then use $L_{ex} + \log((1-p_0)/p_0)$ as the new extrinsic information for both decoders at each iteration. With this simple procedure, the performance is greatly improved.

B. Optimizing the Encoder Structure

The original Turbo codes by Berrou *et al.* used a 16-state $(37, 21)$ recursive systematic convolutional (RSC) code in both constituent encoders. Our simulations showed that it performs poorly when $p_0 \neq 1/2$ over a wide range of E_b/N_0 values, where E_b is the average energy per information bit, and $N_0/2$ is the noise variance. An analysis of the encoder reveals that for non-uniform sources, especially when p_0 is heavily biased, many possible states of this encoder may rarely be reached. For example, when $p_0 = 0.9$, suppose the $(37, 21)$ encoder starts at state 0000, where each digit represents the content of each shift register, then the encoder would remain in this state until a "1" arrives, which would cause a transition to state 1000. Since $p_0 = 0.9$, with high probability, a state transition would occur among very limited number of states, such as 1000, 1100, 0110, 0011, and 0001, etc. As a result, this drawback would cause performance degradations in the decoder. Thus, when $p_0 \neq 1/2$, it is important to find appropriate encoders that can overcome this problem.

With this motivation, we performed a systematic search for better constituent encoders for a given p_0. In our simulations, we only focused on 16-state encoders. Denote the coefficients of the feedback and feed-forward polynomials of a 16-state RSC encoder in binary form as $\{f_0, f_1, f_2, f_3, f_4\}$ and $\{g_0, g_1, g_2, g_3, g_4\}$, respectively, where $f_i, g_j = 0$ or 1, $i, j = 0, 1, \cdots, 4$. Altogether there are $2^4 \times (2^5 - 1) = 496$ combinations; therefore, a full search is impractical. Instead, we choose to determine the sub-optimal encoder among those with $f_0 = f_4 = g_0 = g_4 = 1$. The total number of such encoders is 64. Again, to avoid an exhaustive search, the sub-optimal encoders are obtained through the following iterative steps: 1) Fix the feed-forward polynomial, (e.g., $\{g_0, g_1, g_2, g_3, g_4\} = \{1, 1, 1, 1, 1\}$), find (by simulation) the best feedback polynomial among the remaining possible choices; 2) Fix the feedback polynomial to the one found in step 1), find the best feed-forward polynomial among all remaining possible choices; 3) Fix the feed-forward polynomial as found in step 2), go back to step 1), if the feedback polynomial coincides with the one obtained in step 1), stop (otherwise, go to the next step).

Via this procedure, for a given p_0, several encoders were found to outperform Berrou's $(37, 21)$ encoder significantly; they also have considerably lower error floors. Among them, the $(35, 23)$ encoder gives the best performance for $p_0 = 0.7$ and 0.8; when $p_0 = 0.9$, the best encoder is $(31, 23)$. However, when $p_0 = 1/2$, Berrou's $(37, 21)$ encoder gives a better performance in the waterfall region than the above encoders, though its error floor is higher. For uniform sources, improving the error floor region of Turbo codes is usually achieved at the expense of the waterfall region; however, for the non-uniform case, this tendency seems to decrease as p_0 increases.

III. Simulation Results and Discussions

In this section, we present simulation results of Turbo codes for non-uniform i.i.d. sources over BPSK-modulated AWGN and Rayleigh fading channels. The pseudo-random interleaver of [4] is used. The sequence length is $N = 512 \times 512 = 262144$ and 200 blocks are used; this would guarantee a reliable BER estimation at the 10^{-5} level with 524 errors. The number of iterations used in the decoder is 20. Simulations for our selected codes and the Berrou code are performed for rates R_c=1/3, 1/2, and for p_0=0.5, 0.7, 0.8 and 0.9. To show the performance gains due to using the modified extrinsic information, simulations are also performed by using the unchanged extrinsic information; i.e., the decoder has no knowledge of p_0 and assumes the source is uniform[2].

A. Performance Evaluations

Fig. 1 shows the performance comparison of Berrou's rate-1/3 $(37, 21)$ Turbo code and our selected codes for transmitting uniform and non-uniform i.i.d. sources (with p_0=0.5, 0.7, 0.8 and 0.9) over AWGN channels by using the modified decoder extrinsic information. The $(35, 23)$ code offers the best performance from p_0=0.7 to 0.8; when $p_0 = 0.9$, the $(31, 23)$ seems to be the best choice. At the 10^{-5} BER level, the gains over the Berrou $(37, 21)$ code due to optimization of the encoder for p_0=0.7, 0.8 and 0.9 are 0.09, 0.51 and 1.18 dB, respectively; the gains due to using the modified extrinsic information for p_0=0.7 and 0.8 (with the $(35, 23)$ code) are 0.43 dB and 1.08 dB, respectively; for p_0=0.9 (with the $(31, 23)$ code) it is 2.46 dB.

In Fig. 2, we provide the performance comparison for the Rayleigh fading channel with known channel state information. When p_0=0.7, the gain due to encoder optimization is visible only below the 10^{-5} BER level, because the performance of Berrou's code has lower error floor than that in the AWGN channel case. When p_0=0.8, our $(35, 23)$ Turbo code offers a 0.33 dB gain at a BER of 10^{-5} over its $(37, 21)$ peer; when p_0=0.9, the gain is further increased up to 1.08 dB by using the $(31, 23)$ encoder. Furthermore, the gains due to exploiting the source redundancy become bigger than those obtained in the AWGN channel case: when p_0=0.7 and 0.8, the gains at a BER of 10^{-5} produced by the $(35, 23)$ Turbo code are 0.54 dB and 1.36 dB, respectively; when p_0=0.9, the $(31, 23)$ Turbo code realizes a significant 3.01 dB gain. Similar results are also obtained for rate 1/2 for both AWGN and Rayleigh fading channels.

B. Shannon Limit

Shannon's *Lossy Information Transmission Theorem* [7] states that a memoryless source can be transmitted via a source-channel code (for sufficiently large source block lengths) over a memoryless channel with capacity C at a

[2]In this case, regardless of the source generated at the encoder ($p_0 = 0.9$, 0.8, 0.7 or 0.5), the system performance with unchanged extrinsic information at the decoder varies very slightly with p_0; so we can assume that $p_0 = 0.5$ at both the encoder and the decoder.

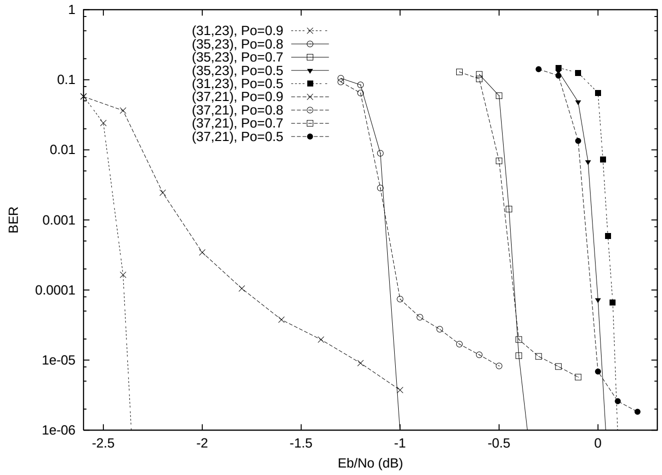

Fig. 1. Turbo codes for non-uniform i.i.d. sources, R_c=1/3, N=262144, AWGN channel.

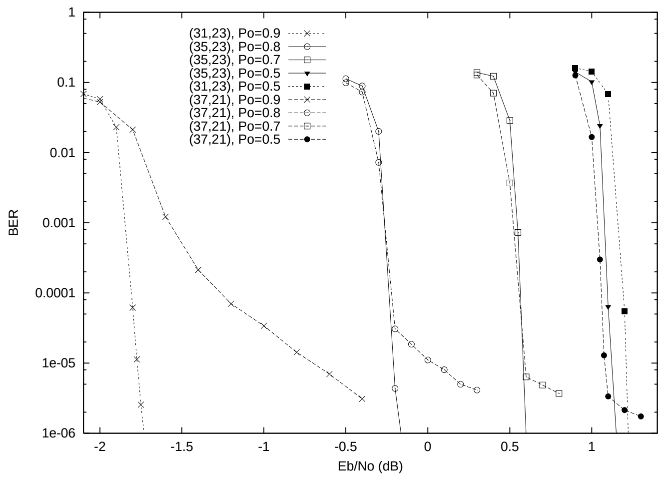

Fig. 2. Turbo codes for non-uniform i.i.d. sources, R_c=1/3, N=262144, Rayleigh fading channel.

rate of R_c source symbols/channel symbols and reproduced at the receiver within an end-to-end distortion D if [7]:

$$R_c \cdot R(D) < C, \tag{1}$$

where $R(D)$ is the rate-distortion function. Under the Hamming distortion measure, $D = P_e$ (BER); then for a binary i.i.d. source with distribution p_0, $R(D)$ is given by

$$R(P_e) = \begin{cases} h_b(p_0) - h_b(P_e), & 0 \le P_e \le \min\{p_0, 1-p_0\} \\ 0, & P_e > \min\{p_0, 1-p_0\} \end{cases}$$

where $h_b(\cdot)$ is the binary entropy function.

For AWGN and Rayleigh channels, the capacity is a function of E_b/N_0. Therefore, the optimum E_b/N_0 to guarantee a BER of P_e can be solved using (1) assuming equality. This value of E_b/N_0 is called the *Shannon limit*, or the *optimal performance theoretically achievable (OPTA)*. The Shannon limit cannot be explicitly solved for a BPSK-modulated input due to the lack of a closed form expression; so it is computed via numerical integration.

For our simulations, we computed the OPTA at the 10^{-5} BER level for $R_c = 1/2$ and $1/3$, and for p_0=0.7, 0.8 and

Turbo Coding System	AWGN		Rayleigh	
	$R_c = 1/2$	$R_c = 1/3$	$R_c = 1/2$	$R_c = 1/3$
(37,21), $p_0 = 0.7$	1.07	0.89	1.17	0.91
(35,23), $p_0 = 0.7$	0.87	0.80	1.16	0.91
(37,21), $p_0 = 0.8$	2.02	1.70	2.18	1.61
(35,23), $p_0 = 0.8$	1.56	1.19	1.88	1.28
(37,21), $p_0 = 0.9$	3.69	3.20	4.12	3.26
(31,23), $p_0 = 0.9$	2.61	2.02	2.99	2.18

TABLE I

OPTA GAPS IN E_b/N_0 AT BER=10^{-5} LEVEL (IN DB).

0.9. Table 1 provides the gaps in E_b/N_0 between the system performance and the corresponding OPTA values. The performances of our Turbo codes are significantly closer to OPTA than those offered by their $(37, 21)$ peer. When p_0 increases, the gaps become wider; meanwhile, the gains achieved by using our encoders over the $(37, 21)$ encoder become more significant. Designing more sophisticated Turbo-based joint source-channel codes that further fill the OPTA gap for heavily biased sources (e.g., with $p_0 = 0.9$) is a challenging and interesting future work.

IV. Conclusion

In this work, we investigate the joint source-channel coding issue of designing Turbo codes for non-uniform i.i.d. sources over AWGN and Rayleigh fading channels. The source redundancy in the form of non-uniformity is exploited by the decoder via a modified extrinsic information, and the constituent encoders are optimized to further enhance the performance. Our results are sub-optimal; further gains might be achieved by a more complete encoder optimization and the use of asymmetric Turbo codes.

Acknowledgment

The authors would like to thank Prof. Marc Fossorier and the reviewers for their valuable comments.

References

[1] F. Alajaji, N. Phamdo, N. Farvardin and T. Fuja, "Detection of binary Markov sources over channels with additive Markov noise," *IEEE Trans. Inform. Theory,* vol. 42, pp. 230–239, Jan. 1996.

[2] F. Alajaji, N. Phamdo, T. Fuja, "Channel codes that exploits the residual redundancy in CELP-encoded speech," *IEEE Trans. Speech and Audio Processing*, vol. 4, pp. 325–336, Sept. 1996.

[3] 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, pp. 248-287, Mar. 1974.

[4] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," *IEEE Trans. Commun.*, vol. 44, pp. 1261–1271, Oct. 1996.

[5] E. Hall and S. Wilson, "Design and analysis of Turbo codes on Rayleigh fading channels," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 160–174, Feb. 1998.

[6] J. Kroll and N. Phamdo, "Source-channel optimized trellis codes for bitonal image transmission over AWGN channels," *IEEE Trans. Image Processing,* vol. 8, pp. 899–912, Jul. 1999.

[7] R. J. McEliece, *The Theory of Information and Coding,* Addison-Wesley, 1977.

[8] J. Hagenauer, E. Offer and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Trans. Inform. Theory,* vol. 42, pp. 429–445, Mar. 1996.

[9] W. Xu, J. Hagenauer and J. Hollmann, "Joint source-channel decoding using the residual redundancy in compressed images," *Proc. Int. Conf. Commun.*, pp. 142–148, Dallas, TX, Jun. 1996.