IMPROVED DECODING OF WOVEN CONVOLUTIONAL CODES VIA HIERARCHICAL NEAR–EEP CODE STRUCTURE

Simon Huettinger and Johannes Huber

Telecommunications Institute II,
University Erlangen-Nürnberg
Cauerstr. 7, D-91058 Erlangen, Germany

***Abstract* — A hierarchically designed channel coding scheme protects a few number of input data streams so–called layers with decreasing amount of redundancy from layer 0 to layer L. Such a code design clearly leads to unequal error protection, but as shown in [9] it can be superior in the average bit error rate compared to an equal error protection (EEP) design, if the protection of layers is not too different. This is due to the fact, that a near–EEP hierarchical design can match the properties of the iterative decoder.**

In this paper we extend the result to Woven Convolutional Codes (WCC) [8]. Based on asymptotical analysis of the convergence of iterative decoding by means of EXIT charts [4] we show that for infinite interleaver size the hierarchical structure is superior. Furthermore, we show that even for practical block sizes used in modern communication systems a significant improvement is possible by hierarchical design.

1. Introduction

The class of Woven Convolutional Codes [8] is a promising alternative to classical Turbo–Codes [3]. Woven Convolutional Codes constitute a serial code concatenation within a more general concept as serially concatenated turbo codes [2] preserving the advantage of iterative decoding suitability and an extremely low error floor. But they also offer a higher degree of freedom in design, as they do not use a single outer code, but an array of codes. Hence, they can achieve a good power efficiency even for short block length and a very low residual error rate.

Up to now most attention is payed to extremely regular codes, as they protect all information symbols equally. Hence, codes are searched for, that have a high minimum hamming distance between the codewords. But as the suboptimal decoding of concatenated codes performed by iteratively using decoders for the component codes, can not realize the power efficiency that could be achieved if maximum likelihood decoding were applied, irregular code structures can be an advantage. Work on irregular Gallager codes (low density parity check codes) [10] and irregular Turbo–Codes [5] has shown that power efficiency can be improved significantly.

If the information bits are grouped into a small number of data streams, the redundancy can be distributed unequally by using a hierarchical coding scheme. By asymptotical analyses that are simplified by the hierarchical structure it is possible to calculate the maximally achievable gain in power efficiency. Furthermore it is possible to ensure, that the difference in protection is small enough to be negligible in practical systems.

In section 2 we will shortly review the construction of WCC with outer warp, which is the basis for the hierarchical design in section 4. Section 3 describes the iterative decoding algorithm used and the importance of extrinsic information. For the asymptotical analysis of classical and hierarchical WCC the concept of EXIT charts [4] is introduced in section 5. Simulation results in section 6 show, that the asymptotical result is also relevant for small block length.

2. Woven Convolutional Codes With Outer Warp

The construction of woven convolutional codes considered in this paper, is called *outer warp* with *row wise* interleaving. There are l_0 binary outer convolutional codes which have — in the original construction each — a mean rate R_o. The information sequence $\boldsymbol{u}$ is now divided into columns with height l_0. These columns form the sequences $\boldsymbol{u}_1^o \dots \boldsymbol{u}_{l_o}^o$ which are fed into the encoders in parallel. The code sequences $\boldsymbol{v}_1^o \dots \boldsymbol{v}_{l_o}^o$ — the *warp* — from each encoder are interleaved by l_0 parallel interleavers $\pi_1 \dots \pi_{l_o}$ and serialized to build the input sequence $\boldsymbol{u}_i$ of the inner binary convolutional encoder — the *weft* — of rate R_i. The woven convolutional code then has the overall rate $R = R_o R_i$.

Component codes of WCC can be any convolutional code. As in all serially concatenated coding schemes, encoding of outer codes does not have much impact on bit error performance, whereas the inner code has to be a recursive systematic convolutional code (RSC) to ensure convergency of iterative decoding. For simplicity we use RSCs for all component codes. Different rates are obtained by puncturing maximum free distance mother codes of rate $R = 1/2$ [7].

If l_o is chosen appropriately high, the woven convolu-

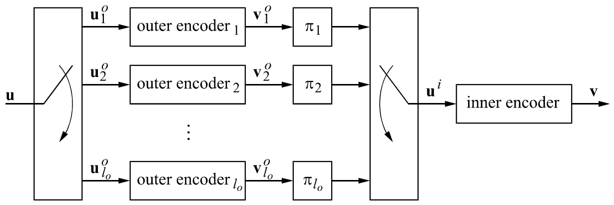

Figure 1: Woven convolutional code : *outer warp* with *row wise* interleaving.

tional code has the remarkable property, that its free distance can be lower bounded by the *product* of the free distances of the component codes.

$$d_{free} \geq d^o_{free} d^i_{free} \tag{1}$$

Hence, woven convolutional codes with outer warp provably have a high minimum distance, that can be achieved already for very small block length. Furthermore the structure can be regarded as a two step encoder consisting of the *warp* and the *weft* separated by an interleaver. Therefore, it is suitable to apply an iterative decoding scheme.

3. Iterative Decoding Of Woven Convolutional Codes

Maximum Likelihood decoding of the woven convolutional code is impossible in practice, as the memory grows with the number of outer codes l_o, the interleaver length N and the memory m of the constituent codes. Hence, we investigate woven convolutional code together with an iterative decoding algorithm, to find a performance of practical significance.

Therefore, we firstly decode *inner* code and warp separately. Within the decoders of the constituent codes of the warp we do not estimate the information bits — which would be the intermediate result of the first iteration — but the post decoding probabilities Pr^o_{pd} of the code symbols. As only extrinsic information, which refers to the incremental information about the current bit obtained from all other bits, can be used in further iterations, Pr^o_{pd} has to be normalized by the input to the outer decoders Pr^o_{apri} which is equal to the output of the inner decoder Pr^i_{pd}. The weighting is done according to [11]. Within the second and all further iterations the inner code then is decoded using the channel output sequence and the extrinsic information provided by the outer decoders to calculate Pr^i_{pd}.

Inner and outer decoder exchange only extrinsic information. Hence, for the asymptotical analysis in section 5 the decoders are characterized via their processing of a–priori information to extrinsic information.

After a number of iterations (typically 8, 16 or 32) the iterative process is stopped and the information sequence is estimated within the decoders of the warp.

To fix some parameters we restrict the analysis to the rate distribution $R_i = 2/3$, $R_o = 3/4$ that leads to an $R = 1/2$ code. Thus, the performance can easily be compared to *turbo–codes* [3]. Optimizing an ordinary woven convolutional code to achieve the best performance for long block length leads to the choice of memory $m = 2$ constituent code both in weft and warp. Fig. 2 shows simulation results with a block length of $K = 9900$ binary information symbols, corresponding to $N = 19800$ binary code symbols, $l_o = 10$ outer codes for 8, 16 and 32 iterations.

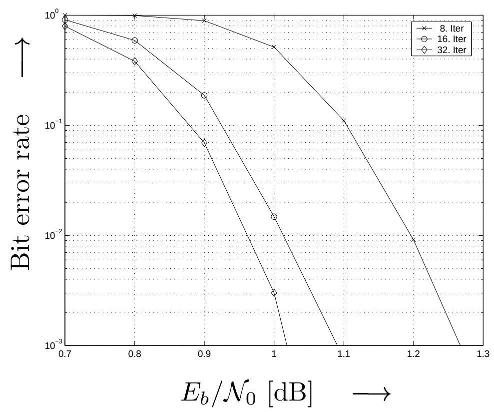

Figure 2: Performance of a classical woven convolutional code ($m = 2$, $R_i = 2/3$, $R_o = 3/4$, $K = 9900$, $N = 19800$, $l_o = 10$).

4. Hierarchical Design

To improve the performance of WCC we try to match the code structure to the properties of the iterative decoder. In [9] we have seen that unequally distributed redundancy within the data stream can significantly improve the convergence of iterative decoding. This is due to the fact, that iterative decoding is suboptimal and does not achieve the maximum likelihood decoding performance. Hence, not always the best code w.r.t. distance criterion will lead to the best performance, as the degree of sub-optimality depends on the code structure.

For multilevel codes it is known, that it is possible to employ low complex multistage decoding without performance loss compared to overall maximum likelihood decoding, if hierarchical structures are used [12]. But, also for concatenated coding schemes decoding complexity is a matter of rate design.

Hence, we try to optimize the rate distribution between the outer codes grouped into levels to construct an unequal error protection code that is matched to the iterative decoder. This is easily possible, as the individual rates of the outer codes of a WCC can in general be different. With this construction we can improve the overall performance for some rate distributions, if the rates are chosen carefully. The changes of component code rates does not affect decoding complexity.

In the following we group the outer codes into two Levels ($L0$ and $L1$) by choosing R_{L0} for all codes with odd index and R_{L1} for all codes with even index. This ensures that the inner code strongly connects the levels. For even l_o the output block length $N_{o,L0}$ and $N_{o,L1}$ are equal, but the number of information bits on level $L0$ and level $L1$ differ.

The iterative decoder passes the information gained by decoding a trellis segment of a component code to the neighboring trellis segments during the iterations. Hence, it equalizes unequal error protection given by the code structure, as the less protected bits can get more help from the highly protected bits then vice versa. As our aim is to find a code/decoder pair with good performance in the average bit error rate, the difference in protection after iterative decoding between the best protected and the least protected bit within the code should be so small that the inequality does not occur in an practical system. Hence, the bit error curves of different protection levels must not vary in more than one order of magnitude in the bit error rate and in more than 0.1 dB in power efficiency.

For very long codes this can be ensured by a convergence analysis as performed in section 5, as convergence indicates that an arbitrary low probability of error can be achieved by increasing the block length N. Hence, this must be true even for the lowly protected bits. If the inequality of error protection is chosen to be too strong, convergence will not be achieved at a low signal–to–noise ratio. But, for shorter codes as investigated in section 6 it is more difficult and, hence, has to be verified by simulations.

5. Asymptotical Analysis

Firstly we investigate the behavior towards infinite interleaver length. In this case the problem of lower protected bit is diminishing, as long as all bits are protected by an outer and an inner code. Bits that are transmitted uncoded — even only through one stage — can not be used while decoding iteratively and hence lead to errors. Furthermore, long puncturing periods that lead to time variance of the convolutional codes, do not affect performance for infinitely long interleaver size, whereas they should be avoided for short block length.

For the analysis we use the EXIT charts [4]. In a simplified model for the iterative decoding process it is assumed that the a–priori and extrinsic information can be characterized by jointly independent Gaussian random variables. Then the decoder can be characterized by its *transfer function*, where at the input and output the mutual information between the Gaussian random variables and the transmitted sequence is measured. For large interleavers the extrinsic information is independent relative to decoding horizon over a large number of iterations, and hence, this model is in excellent agreement with measurements in concatenated coding schemes.

The transfer characteristics of inner and outer component codes are determined without any knowledge about the other component codes. Only the transfer characteristics of inner codes are parameterized by the signal–to–noise ratio $E_b/\mathcal{N}_0$ of the channel.

Previously determined transfer characteristics of inner and outer component codes then are plotted in a single figure, the so–called EXIT chart. EXIT charts provide a visualization of the exchange of extrinsic information between the two component decoders. For increased signal–to–noise ratio $E_b/\mathcal{N}_0$ the transfer characteristic of the inner codes moves upwards in the EXIT chart. Convergence of iterative decoding is possible if the transfer characteristics do no intersect. Hence, the signal–to–noise ratio $E_b/\mathcal{N}_0$ at which a turbo cliff occurs, i.e. a rapidly decreasing bit error rate, can be estimated very precisely from the EXIT chart.

Exemplary we want to construct a memory 2 WCC, which means, that all component codes are of memory 2. We fix the rate distribution between inner and outer codes to be $R_i = 2/3$, $R_o = 3/4$ which results in a total code rate $R = R_i \cdot R_o = 1/2$. With this parameters the EEP design (classical WCC) leads to a turbo cliff at about 0.8 dB. In the EXIT chart of Fig. 3 it can be observed that the transfer characteristics of the inner and outer codes do not intersect at Signal–to–Noise Ratios of $E_b/N_0 > 0.8$ dB, whereas no convergence of iterative decoding is possible for lower Signal–to–Noise Ratios.

For the classical WCC it can be observed (see Fig. 3), that at $E_b/\mathcal{N}_0 \approx 0.8$ dB a tunnel opens that is narrow only in a small interval. This indicates, that inner and outer code do not perfectly fit. Codes that show a more uniformly wide open tunnel generally achieve a higher power efficiency.

With a hierarchical design we try to improve the matching of outer and inner transfer characteristics. In Fig. 4 the transfer characteristics of combinations of punctured versions of the same mother code that fulfill $(R_{L0}+R_{L1})/2 \approx 3/4$ within a puncturing period of maximally 8 are shown.

In Fig. 4 we observe, that the transfer characteristics

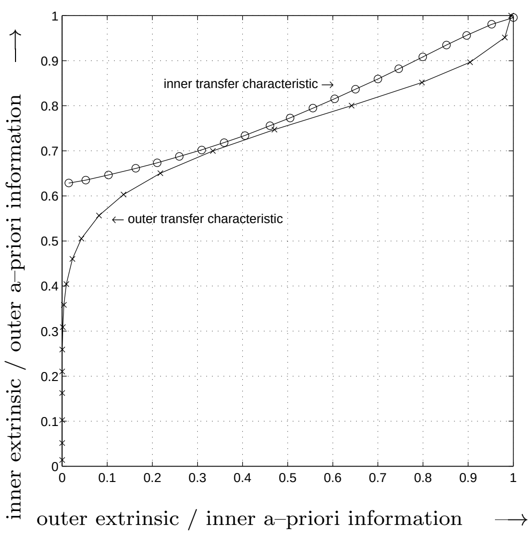

Figure 3: EXIT–Chart for a classical woven convolutional code at $E_b/\mathcal{N}_0 = 0.8$ dB ($m = 2$, $R_i = 2/3$, $R_o = 3/4$).

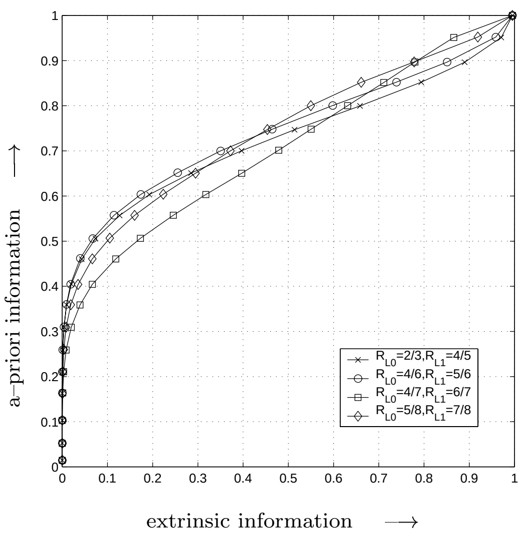

Figure 4: Transfer characteristics of hierarchically constructed outer codes.

of the different hierarchical constructions of outer codes significantly differ from each other. Obviously the one with the best performance at the beginning (a–priori information $I_A \approx 0$) is the worst at the end (a–priori information $I_A \approx 1$).

If we compare the hierarchical transfer characteristics to the transfer characteristic of the inner code, the rate distribution $R_{L0} = 2/3$, $R_{L1} = 4/5$ is the best matching one. Fig. 5 shows the EXIT chart of this hierarchical WCC at $E_b/\mathcal{N}_0 = 0.8$ dB. The tunnel now is much more wide open, which indicates that convergency of iterative decoding is possible for even lower signal–to–noise ratios, and varies less in diameter.

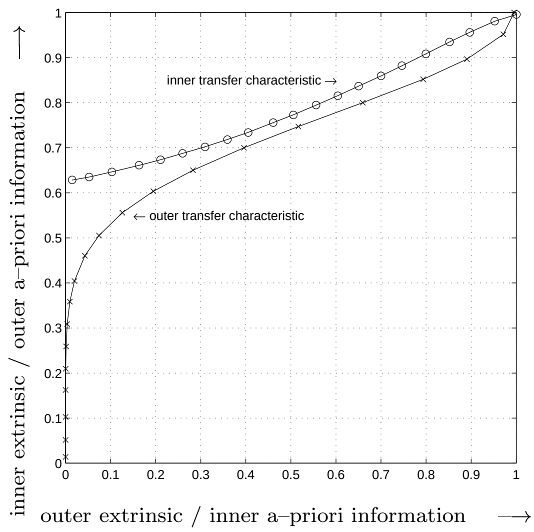

Figure 5: EXIT–Chart for a hierarchical woven convolutional code at $E_b/\mathcal{N}_0 = 0.8$ dB ($m = 2$, $R_i = 2/3$, $R_{L0} = 2/3$, $R_{L1} = 4/5$).

Table 1 summarizes the achievable power efficiency for the studied rate distributions of hierarchical WCC. $E_b/\mathcal{N}_0|_{cliff}$ denotes the signal–to–noise ratio where the turbo–cliff, i.e. a rapidly decreasing bit error rate for increased $E_b/\mathcal{N}_0$, can be expected if the block length is large enough.

Table 1: Convergence of iterative decoding of WCC

Rate level 0	Rate level 1	approx. turbo cliff
$R_{L0} = 3/4$	$R_{L1} = 3/4$	$E_b/\mathcal{N}_0|_{cliff} \approx 0.8$dB
$R_{L0} = 2/3$	$R_{L1} = 4/5$	$E_b/\mathcal{N}_0|_{cliff} \approx 0.6$dB
$R_{L0} = 4/6$	$R_{L1} = 5/6$	$E_b/\mathcal{N}_0|_{cliff} \approx 0.8$dB
$R_{L0} = 4/7$	$R_{L1} = 6/7$	$E_b/\mathcal{N}_0|_{cliff} \approx 1.3$dB
$R_{L0} = 5/8$	$R_{L1} = 7/8$	$E_b/\mathcal{N}_0|_{cliff} \approx 1.0$dB

The small difference in total rate given by the combination $R_{L0} = 2/3$, $R_{L1} = 4/5$ is equalized by puncturing 2% of the output bits of the inner encoder.

Simulations show that for a block length $K = 897600$, $N = 1795200$ there is a bit error rate $BER \approx 10^{-1}$ i.e. a word error rate $WER \approx 1$ for $E_b/\mathcal{N}_0 = 0.75$ dB for the classical WCC respectively for $E_b/\mathcal{N}_0 = 0.55$ dB for the hierarchical WCC ($R_{L0} = 2/3, R_{L1} = 4/5$), whereas the residual error rates for $E_b/\mathcal{N}_0 = 0.85$ dB respectively $E_b/\mathcal{N}_0 = 0.65$ dB are such low, that it can not be measured in simulations with sufficient accuracy. Hence, the hierarchical WCC with $R_{L0} = 2/3$, $R_{L1} = 4/5$ leads to a substantially lower SNR requirement for convergence as it has the best matching transfer characteristic to the inner code. The gain of $\Delta E_b/\mathcal{N}_0 = 0.2$ dB is already one third of the distance to the Shannon–bound which is at $E_b/\mathcal{N}_0 = 0.19$ dB for $R = 1/2$ codes signaled with binary antipodal modulation over the additive white Gaussian noise channel.

6. Simulation Results for Hierarchical WCC of small block length

To show, that the asymptotical superiority of the hierarchical design compared to the original EEP construction of woven convolutional codes is preserved even for short block length, several simulations have been performed. The used block length are $K = 660, 1980, 4620, 9900$ and 897600, respectively $N = 1320, 3960, 9240, 19800$ and 1795200.

The ensemble performance is obtained by creating a new random interleaver for every transmitted block, which is an approximation of uniform interleaving [1]. The distance between the hierarchical and the classical design grows steadily from no gain for very short block length to the asymptotic gain of 0.2 dB. But surprisingly there was no cross over point observed. The hierarchical design is at least as good as the classical design for any block length.

Fig. 6 shows the $E_b/\mathcal{N}_0$ required for a target word error rate $WER = 10^{-3}$ in dependence of the block length N. For comparison a bound on the required $E_b/\mathcal{N}_0$ calculated from the Gallager Error Exponent [6] for random codes of rate $R = 1/2$ is also shown in Fig. 6. This bound is tight for large block length.

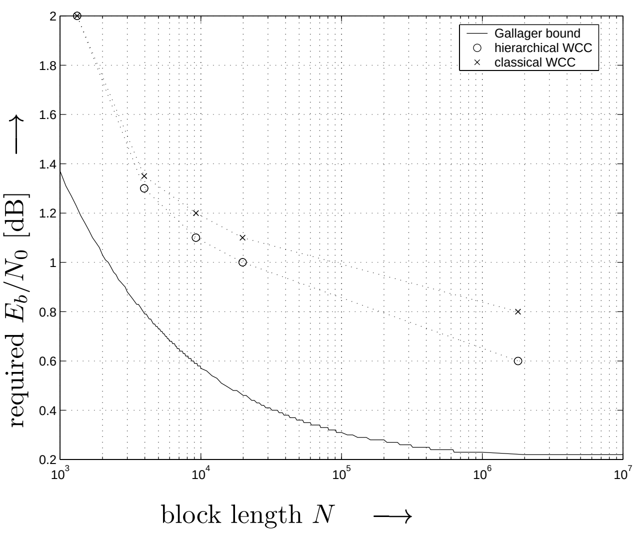

Figure 6: Spectral efficiency of Woven Convolutional Codes.

7. Conclusion

Woven Convolutional Codes are iterative decoding suitable codes that can achieve a good power efficiency even for short block length and a low residual error rate. We have shown, that they still can be improved by choosing the *outer* codes appropriately. Iterative decoding can converge at a lower $E_b/\mathcal{N}_0$, if the outer codes are of different rates, as it can approach the maximum likelihood decoding performance more closely.

Without introducing additional decoding complexity, as always the same mother code is used and only the puncturing patterns are changed, we were able to improve the power efficiency of the memory 2 WCC, which is one of the most power efficient WCC, by approximately 0.2 dB.

References

[1] S. Benedetto and G. Montorsi, "Average performance of parallel concatenated block codes", *IEE Electronics Letters*, Vol.31, No.3, pp. 156–158, Feb. 1995.

[2] S. Benedetto, G. Montorsi, "Iterative decoding of serially concatenated convolutional codes". *Electronics Letters*, vol. 32, no. 13, pp. 1186–1188, June 1996.

[3] C. Berrou, A. Glavieux, P. Thitimajshima, "Near Shannon Limit Error–Correcting Coding and Decoding: Turbo–Codes". *Proc. of ICC '93*, pp. 1064-1070, 1993.

[4] S. ten Brink, "Convergence of iterative decoding", *IEE Electronics Letters*, Vol.35, No.10, pp. 806–808, May 1999.

[5] B. Frey, D. MacKay, "Irregular turbocodes", *Proc. of the 37th Allerton Conference on Communication, Control and Computiong*, Allerton House, Illinois, 1999.

[6] R. G. Gallager, "Information Theory and Reliable Communication", *John Wiley & Sons, Inc.*, New–York, 1968.

[7] D. Haccoun, G. Begin, "High–Rate Punctured Convolutional Codes for Viterbi and Sequential Decoding", *IEEE Transactions on Communications*, vol. 37, no. 11, pp. 1113–1125, Nov. 1989.

[8] S. Höst, R. Johannesson, V. Zyablov, "A first encounter with binary woven convolutional codes", In *Proceedings of the 4th Intern. Symposium on Comm. Theory and Appl.*, pp. 13–18, Lake District, UK, July 1997.

[9] S. Huettinger, J. Huber, "Hierarchical near–EEP code design for improved decoding performance", *Winter School on Coding and Information Theory*, Reisensburg, December 2000.

[10] M. Luby, M. Mitzenmacher, M. Shokrollahi, D. Spielmann, "Improved low–density parity–check codes using irregular graphs and belief propagation", *Proceedings of International Symposium on Information Theory*, 1998.

[11] U. Wachsmann, J. Huber, "Power and Bandwidth Efficient Digital Communication Using Turbo Codes in Multilevel Codes". *European Transactions on Telecommunications (ETT)*, vol. 6, no. 5, pp. 557–567, Sept.–Oct. 1995.

[12] U. Wachsmann, R. Fischer, J. Huber, "Multilevel codes: Theoretical concepts and practical design rules". *IEEE Transactions on Information Theory*, vol. IT-45, pp. 1361–1391, July 1999.