

Turbo-Detected Unequal Error Protection Irregular Convolutional Codes Designed for the Wideband Advanced Multirate Speech Codec

J. Wang, N. S. Othman, J. Kliewer, L. L. Yang and L. Hanzo

School of ECS, University of Southampton, SO17 1BJ, UK

Tel: +44-23-8059 3125, Fax: +44-23-8059 4508

Email: {jw02r,nso01r,jk03v,lly,lh}@ecs.soton.ac.uk, http://www-mobile.ecs.soton.ac.uk

Abstract—since the different bits of multimedia information, such as speech and video, have different error sensitivity, efficient unequal-protection channel coding schemes have to be used to ensure that the perceptually more important bits benefit from more powerful protection. Furthermore, in the context of turbo detection the channel codes should also match the characteristics of the channel for the sake of attaining a good convergence performance. In this paper, we address this design dilemma by using irregular convolutional codes (IRCCs) which constitute a family of different-rate subcodes. We benefit from the high design flexibility of IRCCs and hence excellent convergence properties are maintained while having unequal error protection capabilities matched to the requirements of the source. An EXIT chart based design procedure is proposed and used in the context of protecting the different-sensitivity speech bits of the wideband AMR speech codec. As a benefit, the unequal-protection system using IRCCs exhibits an SNR advantage of about 0.4 dB over the equal-protection system employing regular convolutional codes, when communicating over a Gaussian channel.

I. MOTIVATION

Source encoded information sources, such as speech, audio or video, typically exhibit a non-uniform error sensitivity, where the effect of a channel error may significantly vary from one bit to another [1], [2]. Hence unequal error protection (UEP) is applied to ensure that the perceptually more important bits benefit from more powerful protection. In [3], the speech bits were protected by a family of Rate-Compatible Punctured Convolutional (RCPC) codes [4] whose error protection capabilities had been matched to the bit-sensitivity of the speech codec. Different-rate RCPC codes were obtained by puncturing the same mother code, while satisfying the rate-compatibility restriction. However, they were not designed in the context of turbo detection. Other schemes using a serially concatenated system and turbo processing were proposed in [5], [6], where the UEP was provided by two different-rate convolutional codes.

Recently, Tüchler *et al.* [7], [8] studied the construction of irregular convolutional codes (IRCCs) and proposed several design criteria. These IRCCs consisted of a family of convolutional codes having different code rates and were specifically designed with the aid of extrinsic information transfer (EXIT) charts [9] invoked, for the sake of improving the convergence behaviour of iteratively decoded serially concatenated systems. In general, EXIT chart analysis assumes having a long interleaver block lengths. However, it was shown in [8] that by using an appropriate optimization criterion, the concatenated system is capable of performing well even for short interleaver block lengths. Since the constituent codes have different coding rates,

The financial support of the European Union under the auspices of the Phoenix and Newcom projects and that of the EPSRC is gratefully acknowledged.

the resultant IRCC is capable of providing unequal error protection (UEP).

The novel contribution of the paper is that UEP and EXIT chart based code optimization can be jointly carried out and successfully applied to robust speech transmission. We propose a serially concatenated turbo transceiver using an IRCC as the outer code for the transmission of Adaptive Multi-Rate Wideband (AMR-WB) coded speech. Rather than being decoded separately, the constituent codes of the IRCC are decoded jointly and iteratively by exchanging extrinsic information with the inner code. The IRCC is optimized to match the characteristics of both the speech source codec and those of the channel, so that UEP is achieved while maximizing the iteration gain attained.

In Section II, the error sensitivity of the AMR-WB speech codec is characterized. The system model is introduced in Section III. Section IV describes the design procedure of IRCCs, followed by an example in Section V. Our simulation results are presented in Section VI, while Section VII concludes the paper.

II. THE AMR-WB CODEC'S ERROR SENSITIVITY

The AMR-WB speech codec is capable of supporting bit rates varying from 6.6 to 23.85 kbit/s and it has become a 3GPP and ITU-T standard, which provides a superior speech quality in comparison to the conventional telephone-bandwidth voice codecs [10]. Each AMR-WB frame represents 20 ms of speech, producing 317 bits at a bitrate of 15.85 kbps plus 23 bits of header information per frame. The codec parameters in each frame include the so-called immittance spectrum pairs (ISPs), the adaptive codebook delay (pitch delay), the algebraic codebook excitation index and the jointly vector quantized pitch gains as well as algebraic codebook gains.

Most source coded bitstreams contain certain bits that are more sensitive to transmission errors than others. A common approach for quantifying the sensitivity of a given bit is to consistently invert this bit in every speech frame and evaluate the associated Segmental SNR (SegSNR) degradation [1]. The error sensitivity of the various encoded bits in the AMR-WB codec determined in this way is shown in Fig. 1. The results are based on speech samples taken from the EBU SQAM (Sound Quality Assessment Material) CD, sampled at 16 kHz and encoded at 15.85 kbps. It can be observed that the bits representing the ISPs, the adaptive codebook delay, the algebraic codebook index and the vector quantized gain are fairly error sensitive. By contrast, the least sensitive bits are related to the fixed codebook's excitation pulse positions. Statistically, about 10% (35/340) of the bits in a speech frame will cause a SegSNR degradation in excess of 10 dB, and about 8% (28/340) of the bits will inflict a degradation between 5 and 10 dB. Furthermore, the error-free reception of the 7% (23/340)

header information is in general crucial for the adequate detection of speech.

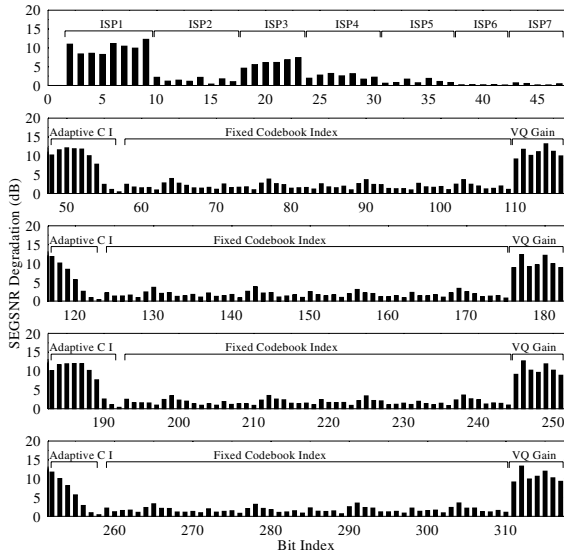


Fig. 1. SegSNR degradations versus bit index due to inflicting 100% BER in the 317-bit, 20 ms AMR-WB frame

III. SYSTEM MODEL

Fig. 2 shows the system's schematic diagram. At the transmitter, each of the K -bit speech frame is protected by a serially concatenated channel code consisting of an outer code (Encoder I) and an inner code (Encoder II) before transmission over the channel, resulting in an overall coding rate of R . At the receiver, iterative decoding is performed with advent of extrinsic information exchange between the inner code (Decoder II) and the outer code (Decoder I). Both decoders employ the *a-posteriori* probability (APP) decoding algorithm, e.g., the BCJR algorithm [11]. After F number of iterations, the speech decoder is invoked in order to reconstruct the speech frame.

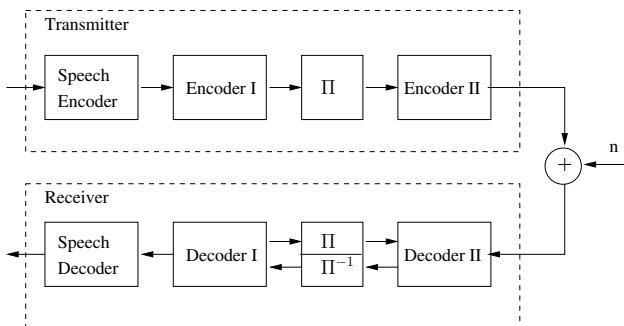


Fig. 2. System Model

According to the design rules of [12], the inner code of a serially concatenated system should be recursive to enable interleaver gain. Furthermore, it has been shown in [13] that for binary erasure channels (BECs) and block lengths tending to infinity the inner code should have rate-1 to achieve capacity. Experiments have shown that this approximately holds also for AWGN channels [7], [8]. For the sake of simplicity, we opted for employing a memory-1 recursive convolutional code having a generator polynomial of

$1/(1 + D)$, which is actually a simple accumulator. Hence the decoding complexity of the inner code is extremely low. In the proposed system, we use an IRCC as the outer code, while in the benchmarker system, we use a regular non-systematic convolutional (NSC) code as the outer code. BPSK modulation and encountering an AWGN channel are assumed.

IV. DESIGN OF IRREGULAR CONVOLUTIONAL CODES

An IRCC is constructed from a family of P subcodes. First, a rate- r convolutional mother code C_1 is selected and the $(P - 1)$ other subcodes C_k of rate $r_k > r$ are obtained by puncturing. Let L denote the total number of encoded bits generated from the K input information bits. Each subcode encodes a fraction of $\alpha_k r_k L$ information bits and generates $\alpha_k L$ encoded bits. Given the target code rate of $R \in [0, 1]$, the weighting coefficient α_k has to satisfy:

$$1 = \sum_{k=1}^P \alpha_k, \quad R = \sum_{k=1}^P \alpha_k r_k, \quad \text{and } \alpha_k \in [0, 1], \quad \forall k. \quad (1)$$

For example, in [8] a family of $P = 17$ subcodes were constructed from a systematic, rate-1/2, memory-4 mother code defined by the generator polynomial $(1, g_1/g_0)$, where $g_0 = 1 + D + D^4$ is the feedback polynomial and $g_1 = 1 + D^2 + D^3 + D^4$ is the feedforward one. Higher code rates may be obtained by puncturing, while lower rates are created by adding more generators and by puncturing under the constraint of maximizing the achievable free distance. The two additional generators used are $g_2 = 1 + D + D^2 + D^4$ and $g_3 = 1 + D + D^3 + D^4$. The resultant 17 subcodes have coding rates spanning from 0.1, 0.15, 0.2, \dots , to 0.9.

The IRCC constructed has the advantage that the decoding of all subcodes may be performed using the same mother code trellis, except that at the beginning of each block of $\alpha_k r_k L$ trellis sections corresponding to the subcode C_k , the puncturing pattern has to be restarted. Trellis termination is necessary only after all of the K information bits have been encoded.

We now optimize the iterative receiver by means of EXIT charts [9], which is capable of predicting the performance of an iterative receiver by examining the extrinsic information transfer function of each of the component devices independently.

For the outer decoder (Decoder I), denote the mutual information between the *a priori* input A and the transmitted code bits C as $I_{A1} = I(C; A)$, while the mutual information between the extrinsic output E and the transmitted code bits C is denoted as $I_{E1} = I(C; E)$. Then the transfer function of Decoder I can be defined as:

$$I_{E1} = T_I(I_{A1}), \quad (2)$$

which maps the input variable I_{A1} to the output variable I_{E1} . Similarly, for the inner decoder (Decoder II), we denote the mutual information between the *a priori* input A and the transmitted information bits X as $I_{A2} = I(X; A)$. Furthermore, we denote the mutual information between the extrinsic output E and the transmitted information bits X as $I_{E2} = I(X; E)$. Note that the extrinsic output of the inner code also depends on the channel SNR or E_b/N_0 . Hence the transfer function of the inner code is defined as

$$I_{E2} = T_{II}(I_{A2}, E_b/N_0). \quad (3)$$

The transfer functions can be obtained by using the histogram-based LLR measurements as proposed in [9] or the simplified method as proposed in [14].

When using IRCCs, the transfer function of an IRCC can be obtained from those of its subcodes. Denote the transfer function

of the subcode k as $T_{I,k}(i)$. Assuming that the trellis fractions of the subcodes do not significantly interfere with each other, which might change the associated transfer characteristics, the transfer function $T_I(i)$ of the target IRCC is the weighted superposition of the transfer function $T_{I,k}(i)$ [8], yielding,

$$T_I(i) = \sum_{k=1}^P \alpha_k T_{I,k}(i). \quad (4)$$

Note that in iterative decoding, the extrinsic output $E2$ of Decoder II becomes the *priori* input $A1$ of Decoder I and vice versa. Given the transfer function, $T_{II}(i, E_b/N_0)$, of the inner code, and that of the outer code $T_I(i)$, the extrinsic information I_{E1} at the output of Decoder I after the i th iteration can be calculated using the recursion of:

$$\mu_i = T_I(T_{II}(\mu_{i-1}, E_b/N_0)), \quad i = 1, 2, \dots, \quad (5)$$

with $\mu_0 = 0$, i.e., assuming the absence of a *priori* input for Decoder II at the commencement of iterations.

Generally, interactive speech communication systems require a low delay, and hence a short interleaver block length. And the number of iterations for the iterative decoder is also limited due to the constraint of complexity. It has been found [8] that EXIT charts may provide a reasonable convergence prediction for the first couple of iterations even in the case of short block lengths. Hence, we fixed the transfer function of the inner code for a given E_b/N_0 value yielding $T_{II}(i) = T_{II}(i, E_b/N_0)$, and optimized the weighting coefficients $\{\alpha_k\}$ of the outer IRCC for the sake of obtaining a transfer function $T_I(i)$ that specifically maximizes the extrinsic output after exactly F number of iterations [8], which is formulated as:

$$\text{maximize } \mu_i = T_I(T_{II}(\mu_{i-1})), \quad i = 1, 2, \dots, F, \quad (6)$$

with $\mu_0 = 0$.

Additionally, considering the non-uniform error sensitivity of the speech source bits, we may intentionally enhance the protection of the more sensitive source data bits by using strong subcodes, thus imposing the source constraints of:

$$\sum_{k=k_1}^{k_2} \alpha_k r_k / R \geq x\%, \quad 1 \leq k_1 \leq k_2 \leq P, \quad 0 \leq x \leq 100, \quad (7)$$

which implies that the percentage of the speech source bits protected by the subcodes k_1 to k_2 is at least $x\%$.

Finally, our task is to find a weight vector $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_P]^T$, so that eq. (6) is maximized, while satisfying the constraints of eq. (1) and eq. (7). This optimization problem can be solved by slightly modifying the procedure proposed in [8], as it will be illustrated by the following example.

V. AN EXAMPLE IRREGULAR CONVOLUTIONAL CODE

We assume the overall system coding rate to be $R = 0.5$. As stated in Section III, the inner code has a unitary code rate, hence all the redundancy is assigned to the outer code. We use a half-rate, memory-4, maximum free distance NSC code having the generator polynomials of $g_0 = 1 + D + D^2 + D^4$, and $g_1 = 1 + D^3 + D^4$. The extrinsic information transfer functions of the inner code and the outer NSC code are shown in Fig. 3. It can be seen that the minimum convergence SNR threshold for the benchmarker system using the NSC outer code is about 1.2 dB, although we note that these curves are based on the assumption of having an infinite interleaver length and a Gaussian Log Likelihood Ratio (LLR) distribution. In the case of short block lengths, the actual SNR convergence threshold might be higher.

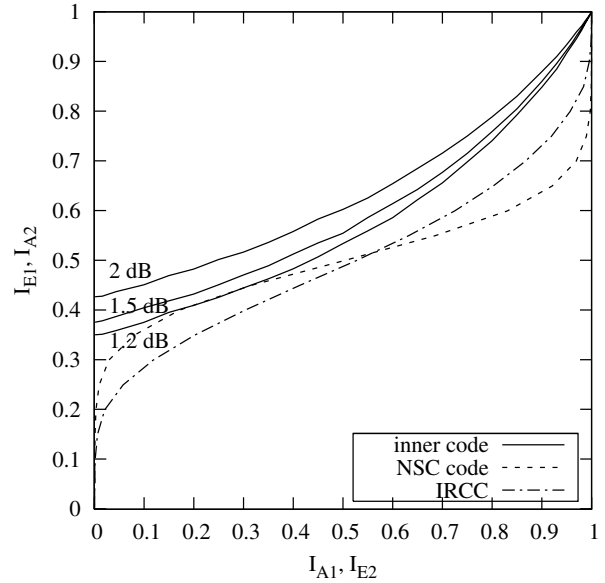


Fig. 3. Extrinsic information transfer functions of the outer NSC code and the designed IRCC, as well as those of the inner code at $E_b/N_0 = 1.2, 1.5$ and 2 dB.

Hence, when constructing the IRCC, we choose the target inner code transfer function $T_{II}(i)$ at $E_b/N_0 = 1.5$ dB, and the number of iterations $F = 6$. For the constituent subcodes, we use those proposed in [8] except that code rates of $r_k > 0.75$ are excluded from our design for the sake of avoiding significant error floors. The resultant code rates of the subcodes span the range of $r_1 = 0.1, r_2 = 0.15, \dots, r_{14} = 0.75$.

Initially the source constraint of eq. (7) was not imposed. By using the optimization procedure of [8], we arrive at the weight vector of $\alpha_0 = [0 \ 0 \ 0 \ 0 \ 0.01 \ 0.13 \ 0.18 \ 0.19 \ 0.14 \ 0.12 \ 0.10 \ 0.01 \ 0.03 \ 0.10]^T$, and the percentage of the input speech data bits protected by the different subcodes becomes $[0, 0, 0, 0, 0.6\%, 9.0\%, 14.4\%, 16.7\%, 14.0\%, 13.0\%, 11.5\%, 1.6\%, 4.2\%, 15.0\%]^T$. The extrinsic output of Decoder I after 6 iterations becomes $\mu_6 = 0.98$.

Observe in the context of the vector containing the corresponding speech bit fractions that only 0.6% of the source bits are protected by the $r_5 = 0.3$ -rate subcode, whereas a total of 23.4% of the speech bits is protected by the $r_6 = 0.35$ and $r_7 = 0.4$ -rate subcodes. In order to enhance the protection of the more sensitive speech bits, we impose now the source constraint of eq. (7) by requiring all the header information bits in a speech frame to be protected by the relatively strong $r_5 = 0.3$ -rate subcode. More explicitly, we impose the constraint of $\alpha_5 r_5 / 0.5 \geq 7\%$, resulting in a new weight vector of $\alpha_1 = [0 \ 0 \ 0 \ 0 \ 0.12 \ 0.06 \ 0.14 \ 0.16 \ 0.13 \ 0.12 \ 0.10 \ 0.02 \ 0.04 \ 0.11]^T$, and the new vector of speech bit fractions becomes $[0, 0, 0, 0, 7.1\%, 4.0\%, 10.9\%, 14.8\%, 13.5\%, 13.3\%, 12.2\%, 2.7\%, 5.5\%, 16\%]^T$. The extrinsic output after 6 iterations is now slightly reduced to $\mu_6 = 0.97$, which is close to the maximum value of 0.98. Furthermore, now, 14.9% of the speech bits is protected by the $r_6 = 0.35$ and $r_7 = 0.4$ -rate subcodes.

The extrinsic information transfer function of this IRCC is also shown in Fig. 3. As seen from the EXIT chart, the convergence SNR threshold for the system using the IRCC is lower than 1.2 dB and there is a wider EXIT chart tunnel between the inner code's curve and

the outer code's curve which is particularly so at the low I_A values routinely encountered during the first couple of iterations. Hence, given a limited number of iterations, we would predict that the system using the IRCC may be expected to perform better than that using the NSC outer code in the range of $E_b/N_0 = 1.5 \sim 2$ dB.

VI. SIMULATION RESULTS

Finally, the achievable system performance was evaluated for a $K = 340$ speech bit per 20 ms transmission frame, resulting in an interleaver length of $L = 688$ bits, including 8 tail bits. This wideband-AMR speech coded [10] frame was generated at a bit rate of 15.85 kbps in the codec's mode 4. Before channel encoding, each frame of speech bits is rearranged according to the descending order of the error sensitivity of the bits, so that the more important data bits are protected by stronger IRCC subcodes. An S-random interleaver [15] was employed with $S = 15$, where all of the subcodes' bits are interleaved together, and 10 iterations were performed by the iterative decoder.

The BER performance of the UEP system using IRCCs and that of the Equal Error Protection (EEP) benchmarker system using the NSC code are depicted in Fig. 4. It can be seen that the UEP system outperforms the EEP system in the range of $E_b/N_0 = 1.5 \sim 2.5$ dB, which matches our performance prediction inferred from the EXIT chart analysis of Section IV.

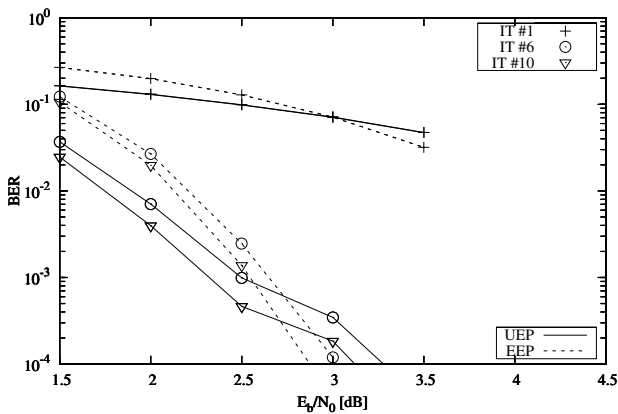


Fig. 4. BER performance of both the UEP system employing the IRCC and the EEP system using the NSC code

The actual decoding trajectories of both the UEP system and the EEP system recorded at $E_b/N_0=1.5$ and 2 dB are shown in Fig. 5 and Fig. 6, respectively. These are obtained by measuring the evolution of mutual information at the input and output of both the inner decoder and the outer decoder as the iterative decoding algorithm is simulated. Due to the relatively short interleaver block length of 688 bits, the actual decoding trajectories do not closely follow the transfer functions especially when increasing the number of iterations. Nonetheless, the UEP system does benefit from having a wider open tunnel during the first couple of iterations and hence it is capable of reaching a higher extrinsic output in the end, resulting in a lower BER.

The BER profiles of the UEP system at $E_b/N_0=1.5, 2$ and 2.5 dB are plotted in Fig. 7. As intended, different fractions of the speech frame benefitted from different degrees of IRCC-aided protection. The first 60 bits represent the header information bits and the most sensitive speech bits, which require the lowest BER.

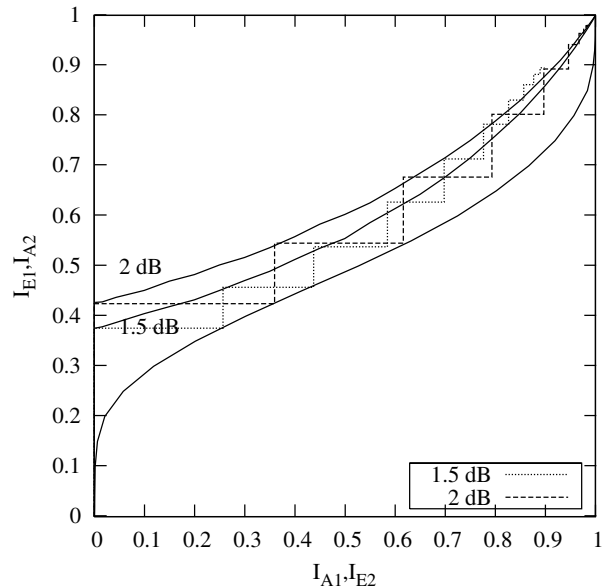


Fig. 5. The EXIT chart and the simulated decoding trajectories of the UEP system using our IRCC as the outer code and a rate-1 recursive code as the inner code at both $E_b/N_0=1.5$ and 2 dB

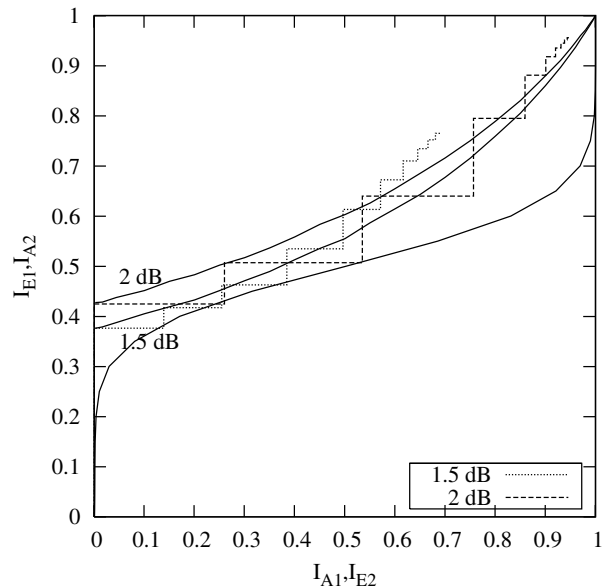


Fig. 6. The EXIT chart and the simulated decoding trajectories of the EEP system using our NSC code as the outer code and a rate-1 recursive code as the inner code at both $E_b/N_0=1.5$ and 2 dB

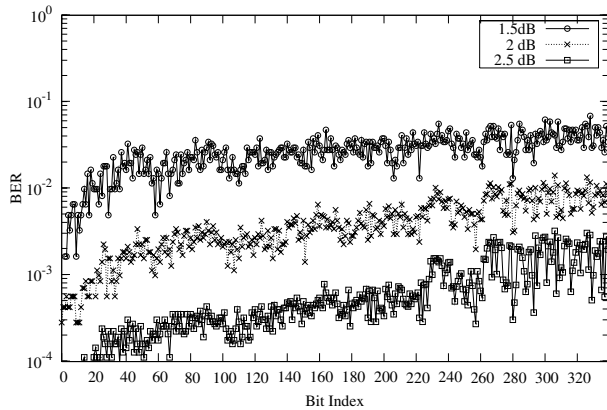


Fig. 7. Bit error rate of the different speech bits after ten iterations at both $E_b/N_0=1.5, 2$ and 2.5 dB recorded by transmitting 10^5 speech frames.

The SegSNR performances of both the UEP and EEP system are depicted in Fig. 8. The UEP system is seen to outperform the EEP system at $E_b/N_0 \leq 2.5$ dB. Above this E_b/N_0 point, the two systems attained almost the same SegSNRs. To achieve a good speech quality associated with $\text{SegSNR} > 9$ dB, the UEP system requires $E_b/N_0 \geq 2$ dB, about 0.3 dB less than the EEP system.

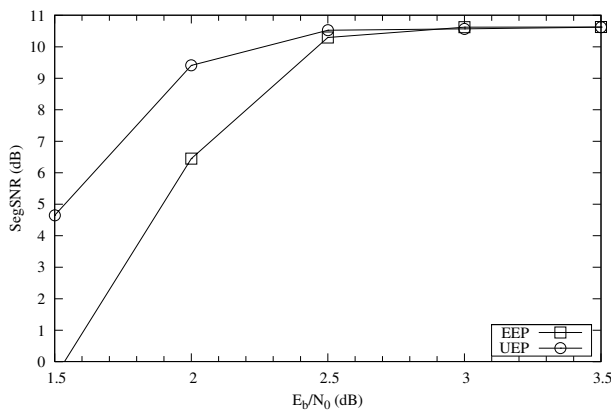


Fig. 8. Comparison of SegSNRs of the AMR-WB speech codec using both EEP and UEP

VII. CONCLUSIONS

We investigated the application of IRCCs for the sake of providing UEP for the AMR-WB speech codec. The IRCCs were optimized with the aid of EXIT charts and the design procedure used was illustrated with the aid of an example.

In the design of IRCCs, we aimed for matching the extrinsic information transfer function of the outer IRCC to that of the inner code, where that of the latter is largely determined by the channel SNR. At the same time, we imposed certain source constraints determined by the error sensitivity of the AMR-WB source bits. Hence the design method proposed here may be viewed as an attractive joint source/channel codec optimization.

The concatenated system using an IRCC benefits from having a low convergence SNR threshold. Owing to its design flexibility, various

transfer functions can be obtained for an IRCC. We have shown that our IRCC was capable of achieving better convergence than a regular NSC code having the same constraint length and code rate. Hence the system using IRCCs has the potential of outperforming the corresponding arrangement using regular NSC codes in the low SNR region.

Furthermore, IRCCs are capable of providing UEP, since it is constituted by various subcodes having different code rates and hence different error protection capabilities. Multimedia source information, such as speech, audio and video source can benefit from this property, when carefully designing the IRCC to match the source's bit sensitivity. Our future research aims for exchanging soft speech bits between the speech and channel decoders.

REFERENCES

- [1] L. Hanzo, F. C. A. Somerville, and J. P. Woodard, *Voice compression and communications: principles and applications for fixed and wireless channels*. Chichester, UK: John Wiley-IEEE Press, 2001.
- [2] L. Hanzo, P. J. Cherriman, and J. Streit, *Wireless Video Communications: Second to Third Generation System and Beyond*. Piscataway, NJ: IEEE Press, 2001.
- [3] R. V. Cox, J. Hagenauer, N. Seshadri, and C.-E. Sundberg, "Subband speech coding and matched convolutional channel coding for mobile radio channels," *IEEE Transactions on signal processing*, vol. 39, no. 8, pp. 1717–1731, August 1991.
- [4] J. Hagenauer, "Rate-compatible punctured convolutional codes (rccp codes) and their applications," *IEEE Transactions on Communications*, vol. 36, no. 4, pp. 389–400, April 1988.
- [5] N. S. Othman, S. X. Ng, and L. Hanzo, "Turbo-decoded unequal protection audio and speech transceivers using serially concatenated convolutional codes, trellis coded modulation and space-time trellis coding," to appear in *Proceedings of the IEEE Vehicular Technology Conference*, September, 2005.
- [6] S. X. Ng, J. Y. Chung, and L. Hanzo, "Turbo-decoded unequal protection MPEG-4 telephony using trellis coded modulation and space-time trellis coding," in *Proceedings of IEE International Conference on 3G mobile communication Technologies (3G 2004)*, London, UK, 18-20 October 2004, pp. 416–420.
- [7] M. Tüchler and J. Hagenauer, "Exit charts of irregular codes," in *Proceedings of Conference on Information Science and Systems [CDROM]*, Princeton University, 20-22 March 2002.
- [8] M. Tüchler, "Design of serially concatenated systems depending on the block length," *IEEE Transactions on Communications*, vol. 52, no. 2, pp. 209–218, February 2004.
- [9] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Transactions on Communications*, vol. 49, no. 10, pp. 1727–1737, October 2001.
- [10] B. Bessette, R. Salami, R. Lefebvre, M. Jelinek, J. Rotola-Pukkila, J. Vainio, H. Mikkola, and K. Jarvinen, "The adaptive multirate wide-band speech codec," *IEEE Transaction on Speech and Audio Processing*, vol. 10, no. 2, pp. 620–636, November 2002.
- [11] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimal symbol error rate," *IEEE Transactions on Information Theory*, vol. 20, pp. 284–287, March 1974.
- [12] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: Performance analysis, design, and iterative decoding," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 909–926, May 1998.
- [13] A. Ashikhmin, G. Kramer, and S. ten Brink, "Extrinsic information transfer functions: model and erasure channel properties," *IEEE Transactions on Information Theory*, vol. 50, no. 11, pp. 2657–2673, November 2004.
- [14] I. Land, P. Hoeher, and S. Gligorević, "Computation of symbol-wise mutual information in transmission systems with logAPP decoders and application to EXIT charts," in *Proceedings of International ITG Conference on Source and Channel Coding (SCC)*, Erlangen, Germany, January 2004, pp. 195–202.
- [15] S. Dolinar and D. Divsalar, "Weight distributions for turbo codes using random and nonrandom permutations," *JPL-TDA Progress Report 42-122*, pp. 56–65, August 1995.