Three-Stage Iterative Detection of a MIMO-aided Precoded AMR-WB Speech Transceiver

N. S. Othman, M. El-Hajjar, O. Alamri, S. X. Ng and L. Hanzo*

School of ECS, University of Southampton, SO17 1BJ, UK.
Tel: +44-23-8059 3125, Fax: +44-23-8059 4508
Email: lh@ecs.soton.ac.uk, http://www-mobile.ecs.soton.ac.uk

Abstract – **A jointly optimised iterative source and channel decoding (ISCD) scheme invoking the Adaptive Multi Rate Wideband (AMR-WB) speech codec is proposed. More explicitly, the transceiver investigated consists of serially concatenated Recursive Convolutional (RSC) codes, a Unity Rate Code (URC) referred to as a precoder and the AMR-WB speech codec, where the resultant bitstream is transmitted using Differential Space-Time Spreading (DSTS) and Sphere Packing (SP) modulation over narrowband temporally correlated Rayleigh fading channels. The convergence behaviour of the advocated scheme is investigated with the aid of Extrinsic Information Transfer (EXIT) charts, which suggests that potential performance improvements can be achieved by combining a serially concatenated precoder with the AMR-WB speech codec. The proposed system exhibits an E_b/N_0 gain of about 1.5 dB in comparison to the benchmark scheme carrying out iterative source- and channel-decoding as well as DSTS aided SP-demodulation, but dispensing with the precoder, when using $I_{system} = 4$ system iterations.**

1. MOTIVATION AND BACKGROUND

Recently, considerable research interest has been devoted to the employment of joint source and channel coding in both delay- and complexity-constrained speech transmission systems [1]. This is justified by the limited applicability of Shannon's classic source and channel coding separation theorem [2] in practical speech systems. A beneficial example of this technique exploits the residual redundancy found in the encoded bitstream of finite-delay lossy speech codecs as *a priori* information for supporting an iterative turbo decoding process, as detailed in [3]. More explicitly, the limited-complexity, limited-delay source encoders fail to remove all the redundancy from the correlated speech source signal, thus leaving some residual redundancy in the encoded parameters.

In [4, 5] the residual redundancy inherent in the source encoded bitstream was exploited using the innovative concept of soft speech bits, which was further developed to accept the *a priori* information passed to the speech decoder from the channel decoder as *extrinsic* information [6]. Hence, *extrinsic* information is iteratively exchanged between the source and channel decoders for the sake of improving the overall system performance. The convergence behaviour of this iterative decoding scheme can be studied using Extrinsic Information Transfer (EXIT) charts [7], which characterize the achievable performance of the ISCD scheme exploiting the residual redundancy inherent in the source encoder bitstream.

In this paper we propose the jointly optimised ISCD scheme of Figure 1 invoking the AMR-WB-speech codec [8], which employs a Unity Rate Code (URC) [9] referred to as a precoder. The precoder has a recursive structure, corresponding to an infinite impulse response invoked for the sake of achieving a precoding-aided decoding convergence enhancement. The resultant bitstream is protected by a Recursive Systematic Convolutional (RSC) code and transmitted using Differential Space-Time Spreading (DSTS) combined with Sphere Packing (SP) modulation [10]. The advocated scheme invokes DSTS employing two transmit and a single receive antenna, which provides a spatial diversity gain without the potentially excessive complexity of channel estimation required by coherently detected Multiple-Input Multiple-Output (MIMO) arrangements, which is an explicit benefit of using non-coherent detection, while tolerating 3 dB performance loss, when compared to the equivalent coherently detected scheme using perfect channel knowledge. We will refer to this three-stage system as the DSTS-SP-RSC-URC-AMRWB arrangement.

This paper is structured as follows. Section 2 provides an overview of the system considered, while in Section 3 we invoked three-dimensional (3D) EXIT charts and their two-dimensional (2D) projections for characterizing the iterative detection aided convergence behaviour of the advocated scheme. Section 4 quantifies the performance of our proposed three-stage scheme, while our conclusions are offered in Section 5.

2. SYSTEM OVERVIEW

2.1. Transmitter

The schematic of the DSTS-SP-RSC-URC-AMRWB scheme is portrayed in Figure 1. As shown in Figure 1, *extrinsic* information is exchanged amongst the AMR-WB decoder, the URC decoder and the RSC decoder. In the advocated scheme, the soft-bit assisted AMR-WB codec of [11] is employed. The AMR-WB speech codec is capable of supporting nine different

The financial support of the Universiti Tenaga Nasional Malaysia, of Vodafone under the auspices of the Dorothy Hodgkin Postgraduate Award and that of the EPSRC, UK as well as that of the European Community under Seventh Framework Programme grant agreement ICT OPTIMIX nINFSO-ICT-214625 is gratefully acknowledged.

978-1-4244-2515-0/09/$25.00 ©2009 IEEE

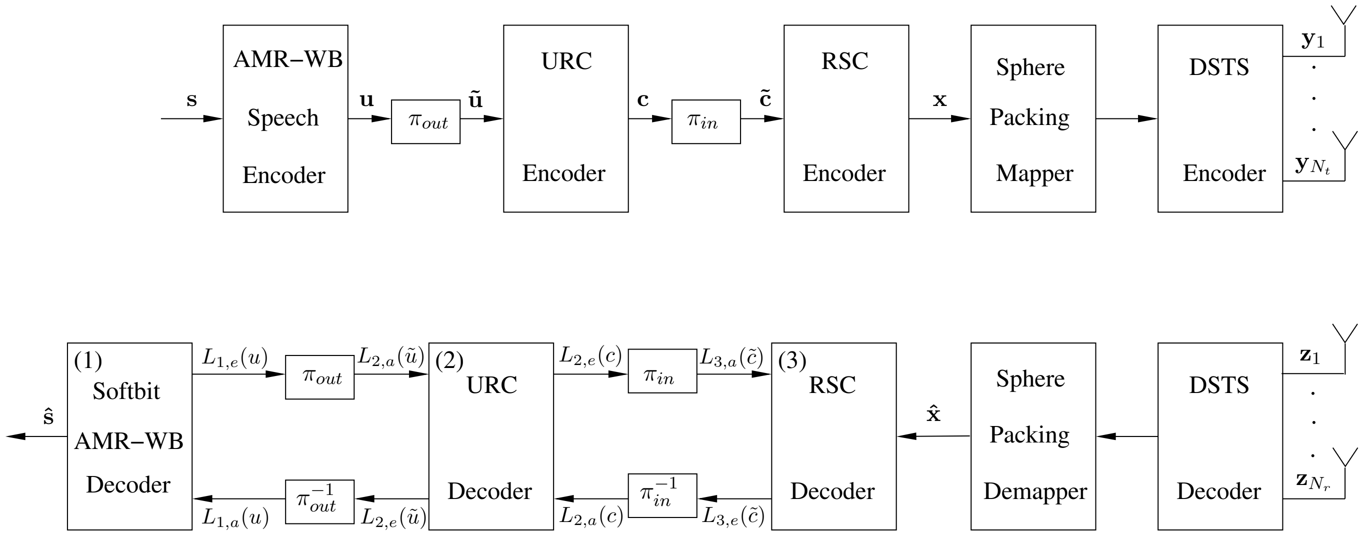

Figure 1: Block diagram of the DSTS-SP-RSC-URC-AMRWB scheme.

speech codec modes having bit rates of 23.85, 23.05, 19.85, 18.25, 15.85, 14.25, 12.65, 8.55 and 6.6 kbit/s [12], each of which may be activated in conjunction with different-rate channel codecs and different-throughput adaptive modem modes [13]. Similar near-instantaneously adaptive speech and video systems were designed in [1, 14]. In our prototype system investigated here the AMR-WB codec is used for encoding the speech signal at a bit rate of 15.85 kbps, which extracts a set of speech parameters per 20 ms frame for representing the speech signal.

The AMR-WB speech encoder produces a frame of speech codec parameters, namely $\{\mathbf{v}_{1,\tau},\mathbf{v}_{2,\tau},\ldots,\mathbf{v}_{\kappa,\tau},\ldots,\mathbf{v}_{52,\tau}\}$, where $\mathbf{v}_{\kappa,\tau}$ denotes an encoded parameter, with $\kappa = 1, ..., K$ denoting the index of each parameter in the encoded speech frame and $K = 52$, whilst τ denotes the time index referring to the current encoded frame's index. Then, $\mathbf{v}_{\kappa,\tau}$ is quantised and mapped to the bit sequence $\mathbf{u}_{\kappa,\tau} = [u(1)_{\kappa,\tau}\ u(2)_{\kappa,\tau}\ \ldots\ u(M)_{\kappa,\tau}]$, where M is the total number of bits assigned to the κth parameter. Then, the bits of the sequence $\mathbf{u}$ is permuted by the outer interleaver π_{out}, yielding $\tilde{\mathbf{u}}$ of Figure 1.

The interleaved AMR-WB-encoded bit sequence $\tilde{\mathbf{u}}$ of Figure 1 is then encoded by a unity-rate recursive precoder having a code memory of 5. The URC coded bit sequence $\mathbf{c}$ is then permuted by the inner interleaver π_{in}, before it is fed to the $\frac{1}{2}$-rate RSC encoder having a code memory of 3. The RSC encoded bits are transmitted by using the DSTS-SP scheme of [10]. The SP mapper maps B number of channel-coded bits $\tilde{\mathbf{x}}=[\tilde{x}_0\ \tilde{x}_1\ \ldots\ \tilde{x}_{B-1}] \in \{0,1\}$ to a SP symbol, as detailed in [10]. Therefore, we have $B = \log_2(L_{SP}) = \log_2(16) = 4$, where L_{SP} represents the set of legitimate SP constellation points. In this investigation, we consider transmissions over a narrowband temporally correlated Rayleigh fading channel, associated with a normalised Doppler frequency of $f_D = 0.01$.

2.2. Receiver

At the receiver, the *extrinsic* information gleaned is exchanged amongst all three constituent decoders of Figure 1, namely the AMR-WB decoder, the URC decoder and the RSC decoder in a number of consecutive iterations. The *inner* iterative loop corresponds to the iterative RSC and URC decoders, while the *outer* iterative loop represents the *extrinsic* information exchange between the AMR-WB speech decoder and the URC decoder.

The notations $L(.)$ in Figure 1 represent the LLRs of the bit probabilities, while $\tilde{c}$, c, $\tilde{u}$ and u in the round brackets $(.)$ of Figure 1 denote the RSC data bits, the URC coded bits, the URC data bits and the AMR-WB encoded bits, respectively. The specific nature of the LLRs is represented by the subscripts of $L_{.,a}$, $L_{.,p}$ and $L_{.,e}$, which denote in Figure 1 the *a priori*, *a posteriori* and *extrinsic* information, respectively. The LLRs associated with one of the three constituent decoders having a label of $\{1,2,3\}$ are differentiated by the corresponding subscripts $(.)$ of $\{1,2,3\}$. Note that the subscript 2 is used for representing the URC decoder of Figure 1.

We define an inner iteration exchanging *extrinsic* information between the RSC and URC decoders followed by two outer iterations between the URC and AMR-WB decoders as having one "system iteration" denoted as $I_{system} = 1$. The advocated scheme invokes the soft-bit assisted AMR-WB speech decoder of [11], which exploits the natural residual redundancy inherent in the AMR-WB speech codec parameters. The details of the algorithm used for computing the *extrinsic* LLR values $L_{1,e}(u)$ of the speech parameters can be found in [6, 15].

The proposed scheme's performance was studied against its benchmark scheme, which does not employ the URC. We will refer to the benchmark scheme as the DSTS-SP-RSC-AMRWB arrangement.

3. THREE-DIMENSIONAL EXIT CHART ANALYSIS

The employment of EXIT charts [7] in the design of the proposed scheme facilitates the prediction of its convergence behaviour, based on the mutual information exchange amongst the constituent receiver components.

As seen from Figure 1, the URC decoder has two *extrinsic* mutual information (MI) [2] outputs, namely $I_{2,E}(\tilde{u})$ and $I_{2,E}(c)$. Both of the *extrinsic* MI outputs are functions of the *a priori* MI inputs of the URC decoder, namely of $I_{2,A}(\tilde{u})$ and $I_{2,A}(c)$, which corresponds to the data bits $\tilde{u}$ originating from the *extrinsic* output of the RSC decoder and the coded bits c generated from the *extrinsic* output of the AMR-WB decoder, respectively. Note that the MI between the *extrinsic* value $E(x)$ and the symbol x is denoted by $I_{\cdot,E}(x)$, whilst $I_{\cdot,A}(x)$ denotes the *a priori* value $A(x)$ and the symbol x. Therefore, the EXIT characteristic of the URC decoder can be described by the following two EXIT functions [16]:

$$I_{2,E}(\tilde{u}) = T_{\tilde{u}}[I_{2,A}(\tilde{u}), I_{2,A}(c)], \quad (1)$$

$$I_{2,E}(c) = T_{c}[I_{2,A}(\tilde{u}), I_{2,A}(c)], \quad (2)$$

which are illustrated by the 3D surfaces seen in Figures 2 and 3, respectively.

By contrast, the AMR-WB decoder and the SP demapper only receive input from and provide output for the URC decoder. However, the EXIT characteristic of the SP demapper is also dependent on the E_b/N_0 value. Therefore, the corresponding EXIT functions are:

$$I_{1,E}(u) = T_{u}[I_{1,A}(u)], \quad (3)$$

for the AMR-WB decoder and

$$I_{3,E}(\tilde{c}) = T_{\tilde{c}}[I_{3,A}(\tilde{c}), E_b/N_0], \quad (4)$$

for the RSC decoder of the two-transmit-antenna aided DSTS-SP system associated with a single receive antenna. Equations (3) and (4) are illustrated in Figures 2 and 3, respectively.

As seen in Figure 2, the intersection of the surfaces characterizes the best attainable performance, when exchanging information between the URC decoder and the AMR-WB decoder of Figure 1 for different fixed values of $I_{2,A}(c)$, which is shown as a bold line. For each point $[I_{2,A}(\tilde{u}), I_{2,A}(c), I_{2,E}(\tilde{u})]$ of this line on the 3D surface of Figure 3, there is a specific value of $I_{2,E}(c)$ determined by $I_{2,A}(\tilde{u})$ and $I_{2,A}(c)$ according to the EXIT function of Equation (2). Therefore, the solid line on the surface of the EXIT function of the URC decoder seen in Figure 2 is mapped to the solid line shown in Figure 3.

The 3D EXIT charts of Figures 2 and 3 are somewhat cumbersome to interpret. Therefore, we project the bold EXIT curve of Figure 2 onto the 2D plane at $I_{2,A}(\tilde{u}) = 0$, yielding the line indicated by the squares in Figure 4. Also shown is the EXIT curve of the RSC decoder in the two-transmit antenna aided DSTS-SP system associated with a single receive antenna for various E_b/N_0 values. By contrast, in the DSTS-SP-RSC-AMRWB benchmark scheme, the soft-bit assisted AMR-WB decoder dispensing with the URC is denoted by the dotted line marked with triangles in Figure 4.

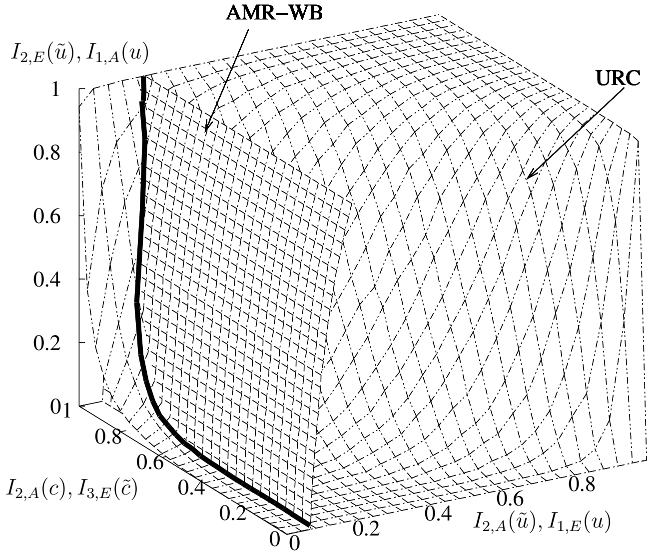

Figure 2: 3D EXIT chart of the URC decoder and of the AMR-WB decoder.

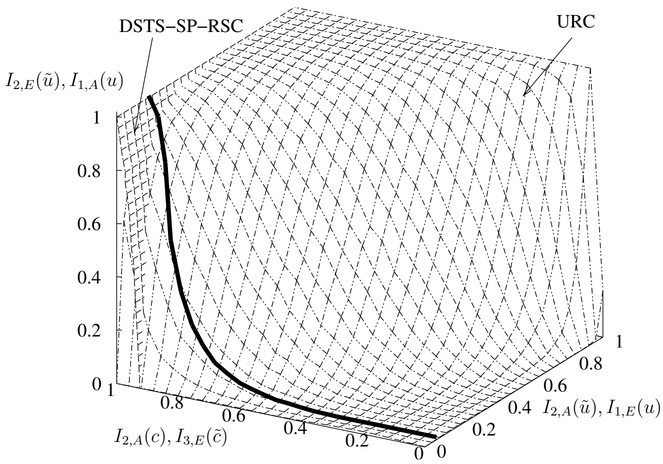

Figure 3: 3D EXIT chart of the URC and the RSC decoders for the two-transmit-antenna aided DSTS-SP system at E_b/N_0=10.0 dB with projection from Figure 2.

Figure 4 shows that the EXIT curve of the soft-bit assisted AMR-WB decoder intersects with that of the RSC decoder. This implies that residual errors persist, despite increasing the number of iterations used and the size of the interleaver. On the other hand, the precoder-aided AMR-WB decoder is capable of closely approaching the point of perfect convergence at (1,1). Thus, there is an open tunnel between the EXIT curve of the RSC decoder and that of the precoder-aided AMR-WB decoder at E_b/N_0=9.0 dB, as seen in Figure 4. Thus, accord-

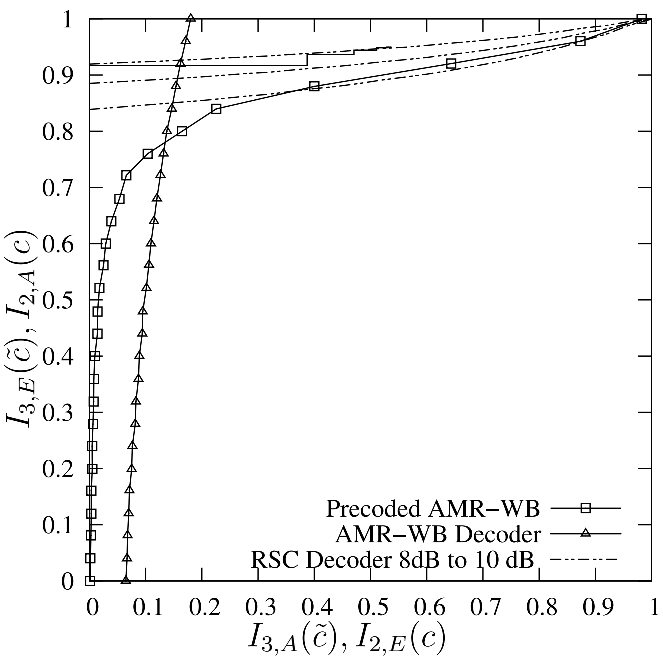

Figure 4: 2D projection of the EXIT chart of the three-stage DSTS-SP-RSC-URC-AMRWB scheme and the 2D EXIT chart of the two-stage DSTS-SP-RSC-AMRWB benchmarker scheme.

ing to the EXIT chart predictions, the proposed system outperforms its benchmark scheme.

4. PERFORMANCE RESULTS

In this section we evaluate the attainable overall system performance of the proposed scheme using both the Bit Error Ratio (BER) and the Segmental Signal to Noise Ratio (SegSNR) [1]. The simulation parameters were described in Section 2. In our simulations each inner iteration between the RSC and URC decoders was followed by two outer iterations between the URC and AMR-WB decoders of Figure 1, which together formed a three-stage "system iteration". Having two outer iterations implied that the *extrinsic* information was exchanged twice between the intermediate URC decoder and the outer AMR-WB decoder. As characterized by the intersection of the surfaces seen in Figure 2 of Section 3, this resulted in the best attainable *extrinsic* LLR values $L_{2,e}(c)$ before being interleaved and fed back to the URC decoder as the *a priori* information $L_{3,a}(\tilde{c})$. Any further increase of the number of outer iterations resulted in an increased complexity, but only marginally increased the attainable *extrinsic* LLR values.

Figure 5 compares the BER performance of the DSTS-SP-RSC-URC-AMRWB scheme and that of its corresponding DSTS-SP-RSC-AMRWB benchmark scheme, when communicating over narrowband temporally correlated Rayleigh fading channels. It can be seen from Figure 5 that the DSTS-SP-RSC-URC-AMRWB scheme outperforms the DSTS-SP-RSC-

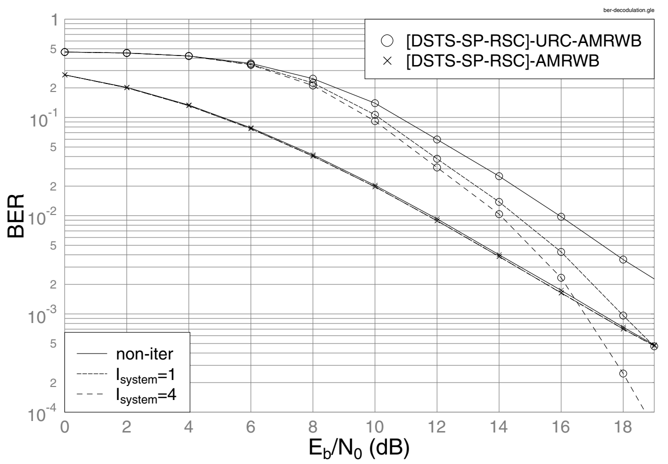

Figure 5: BER versus E_b/N_0 performance of the jointly optimised DSTS-SP-RSC-URC-AMRWB scheme of Figure 1, when communicating over narrowband temporally correlated Rayleigh fading channels.

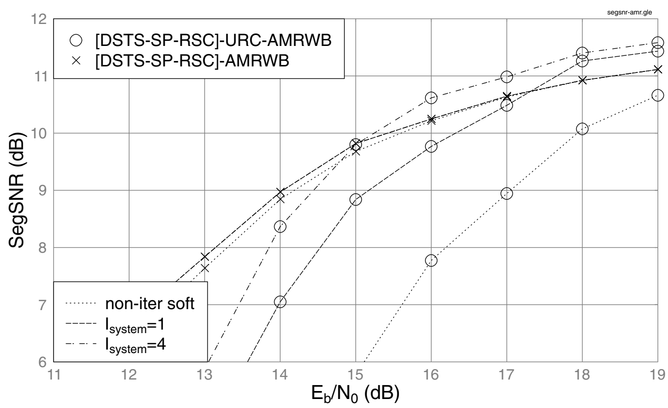

Figure 6: Average SegSNR versus E_b/N_0 performance of the jointly optimised DSTS-SP-RSC-URC-AMRWB scheme of Figure 1 in comparison to the DSTS-SP-RSC-AMRWB benchmark scheme, when communicating over narrowband temporally correlated Rayleigh fading channels.

AMRWB benchmarker by about 1.5 dB at BER= 4×10^{-4} after $I_{system} = 4$ iterations, where again, we define a "system iteration" I_{system} as having an inner iteration followed by two outer-iterations, as mentioned in Section 2. Observe in Figure 4 that with the advent of introducing a recursive precoder, the EXIT characteristic of the AMR-WB decoder is enhanced, which resulted in an open EXIT tunnel for the DSTS-SP-RSC-URC-AMRWB scheme at E_b/N_0=9.0 dB. Thus, the system is expected to achieve a low BER. However, due to the short interleaver length of 462 bits, the actual decoding trajectory of Figure 4 recorded for $I_{system} = 4$ iterations at E_b/N_0=9.0 dB was unable to reach $I_{2,E}(c)$=1.0, and hence the precoded system's actual BER failed to reach an infinitesimally low value.

Figure 6 illustrates the speech SegSNR performance of both the proposed scheme and of the benchmark scheme versus the Signal to Noise Ratio (SNR) per bit, namely E_b/N_0. It is shown in Figure 6 that the precoded scheme performs by 1 dB better in terms of the required channel E_b/N_0 value than its corresponding benchmark scheme, when tolerating a SegSNR degradation of 1 dB.

5. CONCLUSIONS

In this section, we proposed the three-stage serially concantenated DSTS-SP-RSC-URC-AMRWB scheme of Figure 1, where the iterative receiver was constituted by three SISO modules, namely the RSC decoder, the URC decoder and the AMR-WB decoder. Furthermore, the convergence behaviour of our advocated scheme was analyzed using 3D EXIT charts and their 2D projections. It has been demonstrated that the DSTS-SP-RSC-URC-AMRWB scheme achieves a significant BER performance improvement compared to that of the conventional two-stage DSTS-SP-RSC-AMRWB scheme, when communicating over narrowband temporally correlated Rayleigh fading channels. This is achieved by the employment of the recursive URC, which enhanced the EXIT characteristics of the AMR-WB decoder, resulting in an improved BER performance. Explicitly, at BER=4×10^{-4} the precoder was capable of enhancing the achievable E_b/N_0 performance by 1.5 dB, when communicating over narrowband temporally correlated Rayleigh fading channels. Our future research aims for combining the benefits of a recursive URC and the so-called over-complete-mapping scheme of [17], both of which are capable of improving the EXIT characteristics of the AMR-WB decoder in order to approach the point of perfect convergence at (1,1).

6. REFERENCES

[1] L. Hanzo, F. C. A. Somerville and J. P. Woodard, *Voice and Audio Compression for Wireless Communications, 2nd Edition*. Chichester, UK: John Wiley-Sons Inc., 2007.

[2] C. E. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423,623–656, July, October 1948.

[3] N. S. Othman, M. El-Hajjar, O. Alamri and L. Hanzo, "Iterative AMR-WB Source and Channel-Decoding Using Differential Space-Time Spreading Assisted Sphere Packing Modulation," *to appear in IEEE Transactions on Vehicular Technology*.

[4] T. Fingscheidt and P. Vary, "Speech Decoding with Error Concealment Using Residual Source Redundancy," *IEEE Workshop on Speech Coding for Telecommunication*, pp. 91–92, 7-10 Sept 1997.

[5] T. Fingscheidt and P. Vary, "Softbit Speech Decoding: A New Approach to Error Concealment," *IEEE Transactions on Speech and Audio Processing*, vol. 9, pp. 240–251, March 2001.

[6] M. Adrat, U. von Agris and P. Vary, "Convergence Behavior of Iterative Source-Channel Decoding," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 269–272, 6-10 April 2003.

[7] S. ten Brink, "Convergence Behaviour of Iteratively Decoded Parallel Concatenated Codes," *IEEE Transactions on Communications*, vol. 49, pp. 1727–1737, October 2001.

[8] B. Bessette, R. Salami, R. Lefebvre, M. Jelinek, J. Rotola-Pukkila, J. Vainio, H. Mikkola and K. Jarvinen, "The Adaptive Multirate Wideband Speech Codec (AMR-WB)," *IEEE Transactions on Speech and Audio Processing*, vol. 10, pp. 620–636, November 2002.

[9] D. Divsalar, S. Dolinar and F. Pollara, "Serial Concatenated Trellis Coded Modulation with Rate-1 Inner Code," *IEEE Global Telecommunications Conference*, vol. 2, pp. 1168–1175, 27November-1 December 2000.

[10] M. El-Hajjar, O. Alamri, S. X. Ng and L. Hanzo, "Turbo Detection of Precoded Sphere Packing Modulation Using Four Transmit Antennas for Differential Space-Time Spreading," *IEEE Transactions on Wireless Communications*, vol. 7, pp. 943–952, March 2008.

[11] N. S. Othman, M. El-Hajjar, O. Alamri and L. Hanzo, "Soft-Bit Assisted Iterative AMR-WB Source-Decoding and Turbo-Detection of Channel-Coded Differential Space-Time Spreading Using Sphere Packing Modulation," *IEEE 65th Vehicular Technology Conference*, pp. 2010–2014, 22-25 April 2007.

[12] "AMR Wideband Speech Codec: Transcoding Functions," in *3GPP TS 26.190 V5*, December 2001.

[13] L. Hanzo, C. H. Wong and M. S. Yee, *Adaptive Wireless Transceivers: Turbo-Coded, Turbo-Equalized and Space-Time Coded TDMA, CDMA, and OFDM Systems*. New York, USA : John Wiley and Sons, 2002.

[14] L. Hanzo, P. J. Cherriman and J. Streit, *Video Compression and Communications: H.261, H.263, H.264, MPEG4 and Proprietary Codecs*. New York: John Wiley-Sons Inc., 2007.

[15] M. Adrat, P. Vary, and J. Spittka, "Iterative Source-Channel Decoder Using Extrinsic Information from Softbit-Source Decoding," *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 4, pp. 2653–2656, 7-11 May 2001.

[16] F. Brännström, L. K. Rasmussen and A. J. Grant, "Convergence Analysis and Optimal Scheduling for Multiple Concatenated Codes," *IEEE Transactions on Information Theory*, vol. 51, pp. 3354–3364, September 2005.

[17] A. Q. Pham, L. Hanzo and L. -L. Yang, "Joint Optimization of Iterative Source and Channel Decoding Using Over-Complete Source-Mapping," in *IEEE 66th Vehicular Technology Conference*, pp. 1072 – 1076, 30 Sept.-3 Oct. 2007.