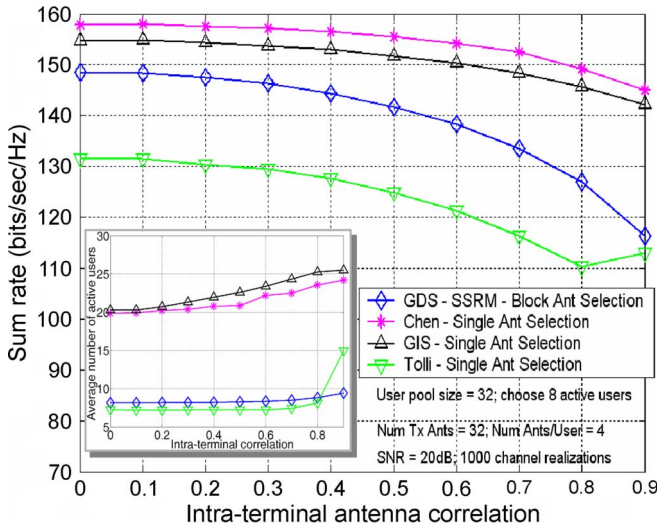

Fig. 4. Comparing RAS algorithms from [9], [12], and [13] with decoupled user–antenna selection with additional scheduling (GDS-MDR-SSRM).

cases where all receive antennas are RF chain equipped. Alternatively, a strategy to schedule the maximum of M channels using the released degrees of freedom may further yield sum rate gains, particularly for large user pools where multiuser diversity is significant. It was shown, however, that RAS/SMS helps narrow the performance gap from sum capacity even for small user pools because the gap is narrowed even when no additional user scheduling is possible. A block antenna/mode selection approach is introduced to address the shortcomings of existing RAS algorithms. A systematic means for resource allocation with rate loss minimization has also been proposed, and a streamlined process that encompasses user selection, RAS/SMS and resource allocation has been developed with the objective of BD-SDM sum rate maximization.

REFERENCES

[1] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.
[2] M. Sharif and B. Hassibi, "A comparison of time-sharing, DPC, and beamforming for MIMO broadcast channels with many users," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 11–15, Jan. 2007.
[3] Q. H. Spencer, A. L. Swindlehurst, and M. H. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.
[4] Z. Pan, K. K. Wong, and T. S. Ng, "Generalized multiuser orthogonal space-division multiplexing," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1969–1973, Nov. 2004.
[5] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 528–541, Mar. 2006.
[6] Z. Shen *et al.*, "Sum capacity of multiuser MIMO broadcast channels with block diagonalization," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2040–2045, Jun. 2007.
[7] Z. Shen *et al.*, "Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization," in *Proc. 39th Asilomar Conf. Signals, Syst., Comput.*, Oct. 2005, pp. 628–632.
[8] B. C. Lim, W. A. Krzymień, and C. Schlegel, "User selection methods for multi-user downlinks with multiple-antenna terminals," in *Proc. ICCS*, Singapore, Oct. 2006, pp. 1–5.
[9] A. Gorokhov, D. Gore, and A. Paulraj, "Receive antenna selection for MIMO spatial multiplexing: Theory and algorithms," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2796–2807, Nov. 2003.
[10] Y. Wu *et al.*, "Receive antenna selection in the downlink of multiuser MIMO systems," in *Proc. VTC—Fall*, Sep. 2005, pp. 477–481.
[11] B. C. Lim, C. Schlegel, and W. A. Krzymień, "Sum rate maximization and transmit power minimization for multi-user orthogonal space division multiplexing," in *Proc. Globecom*, Nov. 2006, pp. 1–5.
[12] R. Chen, J. Andrews, R. Heath, and Z. Shen, "Low-complexity user and antenna selection for multiuser MIMO systems with block diagonalization," in *Proc. ICASSP*, Apr. 2007, pp. III-613–III-616.
[13] A. Tolli and M. Juntti, "Scheduling for multiuser MIMO downlink with linear processing," in *Proc. PIMRC*, Sep. 2005, pp. 156–160.
[14] B. C. Lim, W. A. Krzymień, and C. Schlegel, "Impact of receive antenna selection on scheduling for orthogonal space division multiplexing," in *Proc. WPMC*, Sep. 2006, pp. 840–844.
[15] H. Lütkepohl, *Handbook of Matrices*. Chichester, U.K.: Wiley, 1996.
[16] B. C. Lim, C. Schlegel, and W. A. Krzymień, "Efficient receive antenna selection algorithms and framework for transmit zero-forcing beamforming," in *Proc. VTC—Spring*, May 2006, pp. 2241–2245.

Iterative AMR-WB Source and Channel Decoding Using Differential Space–Time Spreading-Assisted Sphere-Packing Modulation

Noor Shamsiah Othman, *Graduate Student Member, IEEE*, Mohammed El-Hajjar, *Graduate Student Member, IEEE*, Osamah Alamri, Soon Xin Ng, *Senior Member, IEEE*, and Lajos Hanzo, *Fellow, IEEE*

***Abstract*—In this paper, we present a novel system that invokes jointly optimized iterative source and channel decoding for enhancing the error resilience of the adaptive multirate wideband (AMR-WB) speech codec. The resultant AMR-WB-coded speech signal is protected by a recursive systematic convolutional (RSC) code and transmitted using a noncoherently detected multiple-input–multiple-output (MIMO) differential space–time spreading (DSTS) scheme. To further enhance the attainable system performance and to maximize the coding advantage of the proposed transmission scheme, the system is also combined with multidimensional sphere-packing (SP) modulation. Furthermore, the convergence behavior of the proposed scheme is evaluated with the aid of extrinsic information transfer (EXIT) charts. The proposed system exhibits an E_b/N_0 gain of about 1 dB, as compared with the benchmark scheme carrying out joint channel decoding and DSTS-aided SP demodulation in conjunction with separate AMR-WB decoding, when using only $I_{\text{system}} = 2$ system iterations and when communicating over narrow-band correlated Rayleigh fading channels.**

***Index Terms*—Adaptive multirate wideband (AMR-WB) speech codec, differential space–time spreading (DSTS), soft-bit-based iterative speech decoding, sphere decoder, sphere-packing (SP) modulation, three-stage iterative detection, turbo.**

I. MOTIVATION AND BACKGROUND

The classic Shannonian source and channel coding separation theorem [1] has limited applicability in the context of finite-complexity, finite-delay lossy speech [2]. These arguments are particularly valid when the limited-complexity, limited-delay source encoders fail to remove all the redundancy from the correlated speech source signal.

Manuscript received August 7, 2007; revised February 1, 2008, April 14, 2008, and April 15, 2008. First published May 7, 2008; current version published January 16, 2009. This work was supported in part by the Universiti Tenaga Nasional, Kajang, Malaysia; by Vodafone under the auspices of the Dorothy Hodgkin Postgraduate Award; by the Ministry of Higher Education of Saudi Arabia; by the Engineering and Physical Sciences Research Council, U.K.; and by the European Union in the framework of the Phoenix and Newcom projects. The Associate Editor coordinating the review of this paper was Dr. T. J. Lim.

The authors are with the University of Southampton, S017 1BJ Southampton, U.K.

Digital Object Identifier 10.1109/TVT.2008.924977

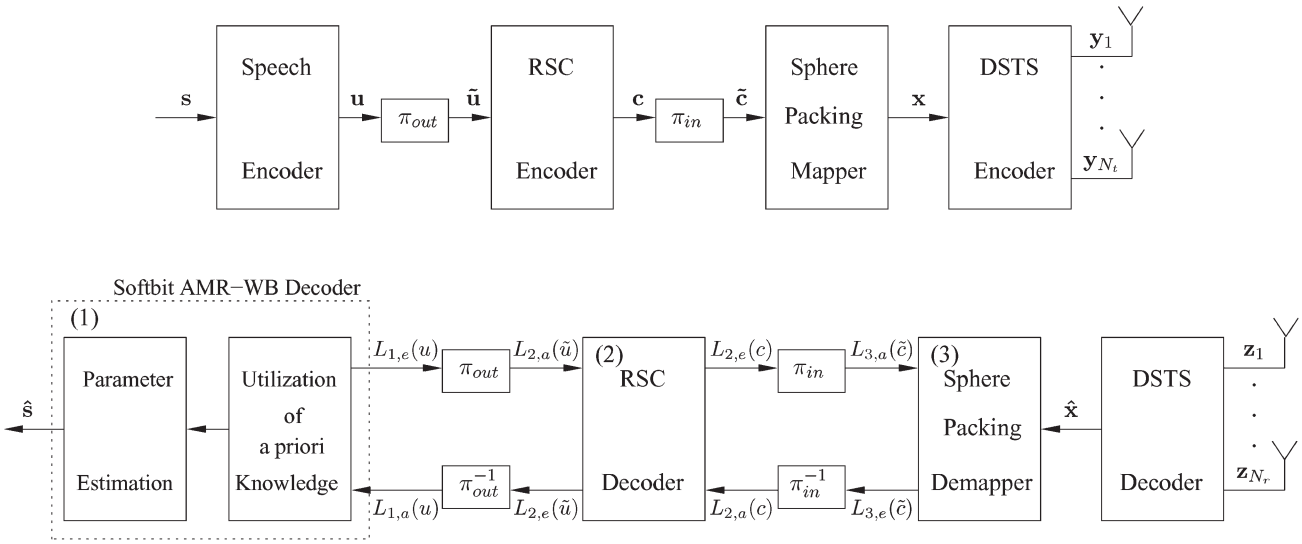

Fig. 1. Block diagram of the DSTS–SP–RSC–AMRWB scheme.

Fortunately, this residual redundancy may beneficially be exploited for error protection by intelligently exchanging soft information among the various receiver components.

The innovative concept of soft speech bits was developed by Fingscheidt and Vary [3], which culminated in the formulation of iterative source and channel decoding (ISCD) [4]. More explicitly, in ISCD, the source and channel decoders iteratively exchange *extrinsic* information for the sake of improving the overall system performance. As a further development, ISCD may beneficially be combined with iterative soft demapping in the context of multilevel modulation and may also be amalgamated with a number of other sophisticated wireless transceiver components. In the resultant multistage scheme, *extrinsic* information is exchanged among three receiver components, namely, the demodulator, the channel decoder, and the soft-input source decoder, in the spirit of [5] and [6]. Explicitly, we propose and investigate the jointly optimized ISCD scheme of Fig. 1, invoking the adaptive multirate wideband (AMR-WB) speech codec [7], which is protected by a recursive systematic convolutional (RSC) code. The resultant bit stream is transmitted using differential space–time spreading (DSTS) amalgamated with sphere-packing (SP) modulation [8] over a narrow-band temporally correlated Rayleigh fading channel. An efficient iterative turbo detection scheme is utilized for exchanging *extrinsic* information between the constituent decoders. In an effort to mitigate the effects of the hostile Rayleigh fading channel, DSTS [8] employing two transmit and one receive antennas was invoked for the sake of providing a spatial diversity gain. This powerful wireless transceiver is advocated here in conjunction with SP modulation, since it was demonstrated in [9] that the employment of SP modulation combined with the orthogonal transmit diversity designs outperformed its conventional counterpart of [10] and [11]. We will refer to this three-stage iteratively detected scheme as the DSTS–SP–RSC–AMRWB scheme.

The convergence behavior of this iterative process is studied using extrinsic information transfer (EXIT) charts [12] by visualizing the input/output mutual information (MI) exchange of the individual constituent of the soft-input–soft-output (SISO) decoders.

The novelty and rationale of the proposed system can be summarized in the list that follows.

1) A SISO AMR-WB decoder is contrived, which is capable of accepting the extrinsic information passed to it from the channel decoder, and subsequently exchanges its extrinsic information with the channel decoder. More explicitly, the residual redundancy inherent in the AMR-WB speech codec parameters is quantified, and this redundancy is exploited as *a priori* knowledge for achieving further performance gains when compared with the less-sophisticated receiver dispensing with this *a priori* knowledge.
2) EXIT chart analysis has been used to design the optimum combination of receiver components. More explicitly, the EXIT curve of the AMR-WB source decoder never reaches the $(1, 1)$ point of perfect convergence, and hence, the achievable bit error rate (BER) remains high. However, if the intermediate RSC decoder and the AMR-WB decoder are viewed as a combined outer SISO module, then the joint EXIT function of this module becomes capable of reaching the convergence point of $(1, 1)$.
3) Conventional coherent space–time spreading requires the estimation of the channel impulse responses of all the multiple-antenna links. For the sake of eliminating the potentially high complexity of MIMO channel estimation in the proposed scheme, the employment of noncoherently detected DSTS using two transmit antennas and a single receive antenna is advocated to achieve a transmit diversity gain. The employment of SP modulation facilitates the joint design of several DSTS time-slot signals, which allows the direct minimization of the DSTS symbol error probability.

This paper is structured as follows: In Section II, the overall system model is described. In Section III, the residual redundancy inherent in the AMR-WB-encoded parameters is quantified, whereas the system's convergence behavior is analyzed in Section IV with the aid of EXIT charts. Section V quantifies the performance of our proposed three-stage scheme, whereas our conclusions are offered in Section VI.

II. System Overview

Fig. 1 shows the iterative decoder structure of the DSTS–SP–RSC–AMRWB scheme, where the *extrinsic* information gleaned is exchanged among all three constituent decoders, namely, the AMR-WB decoder, the RSC decoder, and the SP demapper. The AMR-WB speech codec is capable of supporting nine different bit rates [13], each of which may be activated in conjunction with different-rate

TABLE I
BIT ALLOCATION OF THE AMR-WB SPEECH CODEC AT 23.05 kb/s [13]

Parameter	1st subfr	2nd subfr	3rd subfr	4th subfr	Total per frame
VAD-flag	1				1
ISPs	46				46
LTP Lag	9	6	9	6	30
LTP-filt.-flag	1	1	1	1	4
Fixed Index	88	88	88	88	352
CB Gains	7	7	7	7	28
Total					461

channel codecs and different-throughput adaptive modem modes [14]. Similar near-instantaneously adaptive speech systems were designed in [2]. In our prototype system investigated here, the AMR-WB codec operated at 23.05 kb/s, generating a set of speech parameters encoded by a total of 461 bits per 20-ms frame that represents the 8-kHz bandwidth speech signal sampled at 16 kHz. Similar to most code-excited linear prediction (CELP)-based codecs [2], it performs short-term prediction and long-term prediction (LTP) and generates the excitation codebook (CB) parameters [7]. The resultant bit-allocation scheme is summarized in Table I.

A. Transmitter

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 $\kappa = 1, \ldots, K$ denoting the index of each parameter in the encoded speech frame and $K = 52$, whereas τ denotes the time index referring to the current encoded frame index. Then, $\mathbf{v}_{\kappa,\tau}$ is quantized and mapped to the bit sequence $\mathbf{u}_{\kappa,\tau} = [\, u(1)_{\kappa,\tau} \quad u(2)_{\kappa,\tau} \quad \cdots \quad u(M)_{\kappa,\tau}\,]$, where M is the total number of bits assigned to the κth parameter. Then, the outer interleaver π_{out} permutes the bits of the sequence $\mathbf{u}$, yielding $\tilde{\mathbf{u}}$ of Fig. 1.

The bit sequence c of Fig. 1 is generated by a 1/2-rate RSC code having a code memory of 3 and octally represented generator polynomials of $(G_1, G_2) = (13, 6)$. The DSTS–SP modulator of Fig. 1 first maps B number of channel-coded bits $\tilde{\mathbf{c}} = [\, \tilde{c}_0 \quad \tilde{c}_1 \quad \cdots \quad \tilde{c}_{B-1}\,] \in 0, 1$ to an SP symbol $x \in X$, using the mapping function $x = \text{map}_{\text{sp}}(\tilde{\mathbf{c}})$. Furthermore, we have $B = \log_2(L_{\text{SP}}) = \log_2(16) = 4$, where L_{SP} represents the set of legitimate SP constellation points, as outlined in [9]. This set of SP symbols is transmitted using DSTS in conjunction with two transmit antennas, as detailed in [8]. In this paper, we consider transmissions over a narrow-band temporally correlated Rayleigh fading channel associated with a normalized Doppler frequency of $f_D = 0.01$, whereas the spatial channel coefficients are independent. The notations $L(\cdot)$ in Fig. 1 denote the logarithmic-likelihood ratios (LLRs) of the bit probabilities. The notations $\tilde{c}$, c, $\tilde{u}$, and u in the round brackets $(\cdot)$ in Fig. 1 denote the SP bits, RSC-coded bits, RSC data bits, and AMR-WB-encoded bits, respectively. The specific nature of the LLRs is represented by the subscripts of $L_{\cdot,a}$, $L_{\cdot,p}$ and $L_{\cdot,e}$, which denote the *a priori*, *a posteriori*, and *extrinsic* information, respectively, in Fig. 1. The LLRs associated with one of the three constituent decoders having a label of $\{1, 2, 3\}$ are differentiated by the corresponding subscripts $(\cdot)$ of $\{1, 2, 3\}$. Note that the subscript 2 is used for representing the RSC decoder of Fig. 1.

B. Receiver

Inner Iterations: The complex-valued received symbols $\mathbf{z}$ are demapped to their LLR [15] representation for each of the B number of RSC-encoded bits per DSTS–SP symbol. As shown in Fig. 1, the *a priori* LLR values $L_{3,a}(\tilde{c})$ provided by the RSC decoder are subtracted from the *a posteriori* LLR values $L_{3,p}(\tilde{c})$ at the output of the SP demapper for the sake of generating the *extrinsic* LLR values $L_{3,e}(\tilde{c})$. Then, the LLRs $L_{3,e}(\tilde{c})$ are deinterleaved by a soft-bit deinterleaver. Next, the deinterleaved soft bits $L_{2,a}(c)$ of Fig. 1 are passed to the RSC decoder to compute the *a posteriori* LLR values $L_{2,p}(c)$ provided by the log MAP algorithm [16] for all the RSC-encoded bits. The *extrinsic* information $L_{2,e}(c)$ shown in Fig. 1 is generated by subtracting the *a priori* information $L_{2,a}(c)$ from the *a posteriori* information $L_{2,p}(c)$, which is then fed back to the SP demapper as the *a priori* information $L_{3,a}(\tilde{c})$ after appropriately reordering them using the inner soft-value interleaver. The SP demapper of Fig. 1 exploits the *a priori* information $L_{3,a}(\tilde{c})$ for the sake of providing improved *a posteriori* LLR values $L_{3,p}(\tilde{c})$, which are then passed to the RSC decoder and, in turn, back to the SP demapper for further iterations.

Outer Iterations: As shown in Fig. 1, the *extrinsic* LLR values $L_{2,e}(\tilde{u})$ are generated by subtracting the *a priori* LLR values $L_{2,a}(\tilde{u})$ of the RSC decoder from the LLR values $L_{2,p}(\tilde{u})$. Then, the LLRs $L_{2,e}(\tilde{u})$ are deinterleaved by the outer soft-bit deinterleaver. The resultant soft bits $L_{1,a}(u)$ are passed to the AMR-WB decoder that was further developed for handling soft input bits to compute the *extrinsic* LLR values $L_{1,e}(u)$ with the aid of soft-bit source decoding, as proposed in [4] and detailed during our further discourse. These *extrinsic* LLR values are then fed back to the RSC decoder after appropriately reordering them in the specific order required by the RSC decoder for the sake of completing an outer iteration.

We define one inner iteration followed by two outer iterations as having one "system iteration," which is denoted as $I_{\text{system}} = 1$. The residual redundancy quantified in Section III is exploited as *a priori* information to compute the *extrinsic* LLR values and to estimate the speech parameters. More explicitly, the details of the algorithm used for computing the *extrinsic* LLR values $L_{1,e}(u)$ of the speech parameters can be found in [4] and [17], which are briefly reviewed in the following discussion. First, the channel decoder's output information related to each speech parameter is given by the product of each of the constituent bits, as follows:

$$p(\hat{\mathbf{u}}_{\kappa,\tau}|\mathbf{u}_{\kappa,\tau}) = \prod_{m=1}^{M} p\,[\hat{u}_{\kappa,\tau}(m)|u_{\kappa,\tau}(m)] \tag{1}$$

where $\hat{\mathbf{u}}_{\kappa,\tau} = [\, \hat{u}(1)_{\kappa,\tau} \quad \hat{u}(2)_{\kappa,\tau} \quad \cdots \quad \hat{u}(M)_{\kappa,\tau}\,]$ represents the received bit sequence of the κth parameter, whereas $\mathbf{u}_{\kappa,\tau}$ is the corresponding transmitted bit sequence, provided that all these bits are independent of each other, although in reality, the M bits of the κth parameter are not entirely independent of each other at the output of a practical source codec. The effects of this approximation are eliminated by the iterative detector during its consecutive iterations.

Extrinsic LLR of Soft Speech Bit Generation for Exploiting the Parameters' Unequal Probability: As usual, we exclude the bit under consideration from the present bit sequence within each of the κth parameters, where $\kappa = 1, \ldots, K$, and $K = 52$, namely, from $\mathbf{u}_{\kappa,\tau} = [u_{\kappa,\tau}(\lambda)\mathbf{u}_{\kappa,\tau}^{[\text{ext}]}]$. The *extrinsic* channel decoder output information $u_{\kappa,\tau}(\lambda)$ of each desired bit is expressed as

$$p\left(\hat{\mathbf{u}}_{\kappa,\tau}^{[\text{ext}]}|\mathbf{u}_{\kappa,\tau}^{[\text{ext}]}\right) = \prod_{m\neq\lambda, m=1}^{M} p\,[\hat{u}_{\kappa,\tau}(m)|u_{\kappa,\tau}(m)] \tag{2}$$

where the term $\mathbf{u}_{\kappa,\tau}^{[\text{ext}]}$ denotes all elements of the bit pattern $\mathbf{u}_{\kappa,\tau}$ but excludes the desired bit $u_{\kappa,\tau}(\lambda)$ itself. Finally, the *extrinsic* LLR value $L_{1,e}(u)$ generated for each bit can be obtained by combining the corresponding channel decoder output information and the *a priori*

knowledge $p(\mathbf{u}_{\kappa,\tau})$ concerning the κth parameter, which is given by [4], [18]

$$L_{1,e}\left(u_{\kappa,\tau}(\lambda)\right) = \log \frac{\sum\limits_{\mathbf{u}_{\kappa,\tau}^{[\mathrm{ext}]}} p\left(\mathbf{u}_{\kappa,\tau}^{[\mathrm{ext}]} \middle| u_{\kappa,\tau}(\lambda) = +1\right) \cdot \exp A}{\sum\limits_{\mathbf{u}_{\kappa,\tau}^{[\mathrm{ext}]}} p\left(\mathbf{u}_{\kappa,\tau}^{[\mathrm{ext}]} \middle| u_{\kappa,\tau}(\lambda) = -1\right) \cdot \exp A} \quad (3)$$

where

$$A = \sum_{u_{\kappa,\tau}(l) \text{ of } \mathbf{u}_{\kappa,\tau}^{[\mathrm{ext}]}} \frac{u_{\kappa,\tau}(l)}{2} \left(L_{1,a}\left[u_{\kappa,\tau}(l)\right]\right)$$

and $L_{1,a}$ represents the *a priori* LLR values of the AMR-WB decoder, which is the deinterleaved counterpart of $L_{2,e}$ generated by the RSC decoder.

Extrinsic LLR of Soft Speech Bit Generation for Exploiting the First-Order Interframe Correlation: As shown in Fig. 1, the *extrinsic* LLRs $L_{1,e}(u)$ can be generated by subtracting the *a priori* information $L_{1,a}(u)$ from the *a posteriori* information $L_{1,p}(u)$. Again, to realize a transmission scheme imposing no extra latency, we generate and exploit only the forward *a posteriori* probability (APP) by exploiting the *a priori* knowledge expressed in terms of $p(\mathbf{u}_{\kappa,\tau}|\mathbf{u}_{\kappa,\tau-1})$, yielding

$$\alpha_{\tau-1}(\mathbf{u}_{\kappa,\tau-1}) = Cp(\hat{\mathbf{u}}_{\kappa,\tau-1}|\mathbf{u}_{\kappa,\tau-1}) \times \sum_{\mathbf{u}_{\kappa,\tau-2}} p(\mathbf{u}_{\kappa,\tau-1}|\mathbf{u}_{\kappa,\tau-2}) \cdot \alpha_{\tau-2}(\mathbf{u}_{\kappa,\tau-2}) \quad (4)$$

where $\alpha_{\tau-1}(\mathbf{u}_{\kappa,\tau-1})$ represents a forward recursive value, and C represents a normalization constant. Finally, the *a posteriori* LLR value $L_{1,p}(u)$ generated for each bit is given in (5), shown at the bottom of the page.

III. Residual Redundancy in the AMR-WB Speech Codec

The ideal Shannonian entropy coding-based source encoder would produce a stream of independent identically distributed equiprobable bits. However, since the AMR-WB encoder is not an ideal high-delay lossless entropy encoder but a realistic finite-delay lossy CELP codec [13], it leaves some residual redundancy in the encoded parameters.

First, this residual redundancy manifests itself in terms of the unequal probability of occurrence of the different values of a specific parameter in each 20-ms AMR-WB-encoded frame, which we may refer to as unequal-probability-related redundancy. The second manifestation of the residual redundancy is constituted by the similarity of the parameters in the current and the immediately preceding 20-ms AMR-WB-encoded frames, which may be referred to as first-order interframe correlation. The parameters extracted from speech, such as the linear prediction coefficients (LPCs), show a significant correlation between successive frames, particularly for voiced segments. The two popular LPC parameter representations used in wideband speech codecs are the line spectral frequencies (LSFs) [19] and immittance spectral pairs (ISPs) [20]. It was shown in [2] for the LSFs plotted as a function of time that the correlation between successive frames is high. The AMR-WB codec uses the so-called ISP [13] to represent the LPC parameters. Thus, it is expected that the correlation of ISPs between successive frames is also high. The remaining AMR-WB-encoded parameters, namely, the LTP lags, the CB gains, and the fixed CB indexes, are encoded on a per-subframe or a 5-ms basis. These parameters of the subsequent subframes within a frame also exhibit correlations. Nonetheless, the odd subframes' LTP lags and the even subframes' LTP lags have differently been encoded. Intuitively, the fixed CB indexes representing the random excitation vectors are expected to exhibit no significant first-order interframe correlation. Hence, it is justifiable that we only quantify the first-order interframe correlation of the ISP and the CB gain parameters.

For the sake of quantifying the residual redundancy inherent in the bit stream, a large training sequence of 2 133 035 20-ms frames was applied to the AMR-WB encoder, which produces 52 different encoded parameters for each 20-ms frame. The relative frequency of each individual legitimate AMR-WB-encoded parameter transition was computed for the sake of estimating the transition probabilities of those parameters, which did exhibit nonnegligible first-order interframe correlation in two consecutive 20-ms frames. Similarly, the probability of occurrence was recorded for each of the AMR-WB-encoded parameters, which did not exhibit exploitable first-order interframe correlation. The resultant residual redundancy recorded for the interframe correlated parameters, such as the ISP and the CB gain parameters, are summarized in Table II in terms of their MI $\mathcal{R}_M$ between the corresponding parameters of two consecutive 20-ms-spaced speech frames. By contrast, the residual unequal-probability-related redundancy of each AMR-WB codec parameter was quantified in terms of $\mathcal{R}_D = \mathcal{B} - H(\mathcal{U})$, where $H(\mathcal{U})$ is the entropy of the quantized parameter $\mathcal{U}$, and $\mathcal{B}$ is the number of bits actually used for quantizing the parameter $\mathcal{U}$. However, for simplicity, not all AMR-WB-encoded parameters' residual redundancy is shown in Table II. As an example, we can observe in Table II that the speech-energy-related CB gain parameters have high first-order interframe MI, where the CB gain parameters of the first subframe contain 0.97 bits/parameter information, i.e., 13.8% of residual redundancy. By contrast, the efficient employment of the S-MSQV in the encoding process of the ISP parameters removes most of the redundancy, and hence, only the first two ISP parameters have relatively high first-order interframe MI, where they contain 20.9% and 15.8% redundancy, respectively. These results suggest that the high-correlation CB gain parameters and the first two ISP parameters would benefit from exploiting the nonnegligible first-order interframe correlation-based extrinsic information, whereas the rest of the parameters would benefit from exploiting the nonnegligible unequal-probability-based extrinsic information. To realize a transmission scheme that does not introduce any extra interframe coding-induced delay, we only exploit the APP gleaned from the previously received speech frames. The algorithm used for computing the APPs quantified in terms of their LLRs and their MAP decoding was briefly reviewed in Section II. It will be shown in Section V that the exploitation of this residual redundancy at the decoder has the potential to provide useful performance gains when compared to the less-sophisticated receiver dispensing with this *a priori* knowledge.

$$L_{1,p}\left(u_{\kappa,\tau}(\lambda)\right) = \log \frac{\sum\limits_{u_{\kappa,\tau}(\lambda)=+1} p(\hat{\mathbf{u}}_{\kappa,\tau}|\mathbf{u}_{\kappa,\tau}) \sum\limits_{\mathbf{u}_{\kappa,\tau-1}} p(\mathbf{u}_{\kappa,\tau}|\mathbf{u}_{\kappa,\tau-1})\alpha_{\tau-1}(\mathbf{u}_{\kappa,\tau-1})}{\sum\limits_{u_{\kappa,\tau}(\lambda)=-1} p(\hat{\mathbf{u}}_{\kappa,\tau}|\mathbf{u}_{\kappa,\tau}) \sum\limits_{\mathbf{u}_{\kappa,\tau-1}} p(\mathbf{u}_{\kappa,\tau}|\mathbf{u}_{\kappa,\tau-1})\alpha_{\tau-1}(\mathbf{u}_{\kappa,\tau-1})} \quad (5)$$

TABLE II
RESIDUAL REDUNDANCY IN THE AMR-WB CODEC PARAMETERS

First-Order Inter-frame Correlation		
Parameter	$\mathcal{R}_M = \mathcal{I}(x;y)$	Residual Redundancy Percentage
$ISP_1(t); ISP_1(t-1)$	1.67	20.9
$ISP_2(t); ISP_2(t-1)$	1.26	15.8
$ISP_3(t); ISP_3(t-1)$	0.29	4.8
$ISP_4(t); ISP_4(t-1)$	0.12	1.7
$ISP_5(t); ISP_5(t-1)$	0.12	1.7
$ISP_6(t); ISP_6(t-1)$	0.05	1.0
$ISP_7(t); ISP_7(t-1)$	0.06	1.0
$CBGain_{sub_1}(t); CBGain_{sub_1}(t-1)$	0.97	13.9
$CBGain_{sub_2}(t); CBGain_{sub_2}(t-1)$	1.00	14.3
$CBGain_{sub_3}(t); CBGain_{sub_3}(t-1)$	1.05	13.1
$CBGain_{sub_4}(t); CBGain_{sub_4}(t-1)$	1.04	14.9
Inter-frame Unequal Redundancy		
Parameter	$\mathcal{R}_D = \mathcal{B} - H(\mathcal{U})$	Residual Redundancy Percentage
$ISP_1(t)$	0.57	7.1
$ISP_2(t)$	0.45	5.6
$ISP_3(t)$	0.10	1.7
$ISP_4(t)$	0.06	0.8
$ISP_5(t)$	0.14	2.0
$ISP_6(t)$	0.03	0.6
$ISP_7(t)$	0.07	1.4
$LTPLag_{sub_1}(t)$	0.23	2.6
$LTPLag_{sub_2}(t)$	0.40	6.6
$LTPLag_{sub_3}(t)$	0.20	2.2
$LTPLag_{sub_4}(t)$	0.39	6.5
$FixedInd._{sub_1}(t)$	0.33	3.0
$FixedInd._{sub_2}(t)$	0.31	2.8
$FixedInd._{sub_3}(t)$	0.31	2.8
$FixedInd._{sub_4}(t)$	0.31	2.8
$CBGain_{sub_1}(t)$	0.12	1.7
$CBGain_{sub_2}(t)$	0.12	1.7
$CBGain_{sub_3}(t)$	0.13	1.9
$CBGain_{sub_4}(t)$	0.14	2.0

IV. EXIT CHART ANALYSIS

EXIT charts have widely been used in the design of iterative schemes, which facilitate the prediction of the associated convergence behavior, based on the exchange of MI among the constituent receiver components.

As shown in Fig. 1, the RSC decoder receives inputs from and provides outputs for both the SP and AMR-WB decoders. More explicitly, let $I_{\cdot,A}(x)$ denote the MI [1] between the *a priori* value $A(x)$ and the symbol x, while $I_{\cdot,E}(x)$ denote the MI between the *extrinsic* value $E(x)$ and the symbol x. The MI associated with one of the three constituent decoders having a label of $\{1,2,3\}$ is differentiated by the corresponding subscripts $(\cdot)$ of $\{1,2,3\}$. Thus, the input of the RSC decoder is constituted by the *a priori* input $I_{2,A}(c)$, corresponding to the coded bits c originating from the *extrinsic* output of the SP decoder, as well as the *a priori* input $I_{2,A}(\tilde{u})$ available for the data bits $\tilde{u}$, which was generated from the *extrinsic* output of the AMR-WB decoder. Note that the subscript 2 is used for representing the RSC decoder of Fig. 1.

Correspondingly, the RSC decoder generates both the *extrinsic* output $I_{2,E}(\tilde{u})$, which represents the data bits $\tilde{u}$, as well as the *extrinsic* output $I_{2,E}(c)$, representing the coded bits c, where the corresponding EXIT functions are $T_{\tilde{u}}[I_{2,A}(\tilde{u}), I_{2,A}(c)]$ and $T_c[I_{2,A}(\tilde{u}), I_{2,A}(c)]$, respectively. By contrast, the AMR-WB and SP decoders only receive input from and provide output for the RSC decoder. Thus, the corresponding EXIT functions are $T_u[I_{1,A}(u)]$ for the AMR-WB decoder and $T_{\tilde{c}}[I_{3,A}(\tilde{c}), E_b/N_0]$ for the DSTS–SP decoder.

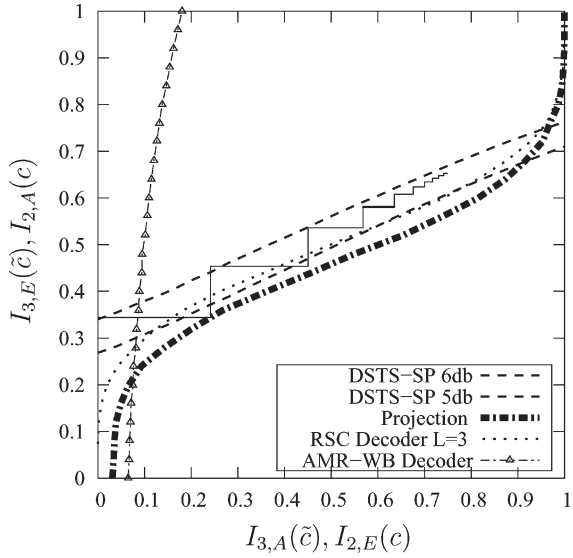

Fig. 2. EXIT charts of the three-stage DSTS–SP–RSC–AMRWB scheme and the two-stage benchmarker scheme.

The EXIT chart [12] analysis of the iterative decoding scheme's convergence behavior indicates that an infinitesimally low BER may only be achieved by an iterative receiver if an open tunnel exists between the EXIT curves of the two SISO components.

The EXIT chart of the advocated system of Fig. 1 is shown in Fig. 2 when using the system parameters described in Section II. This scheme is compared with the benchmark scheme carrying out joint channel decoding and DSTS-aided SP demodulation in conjunction with separate AMR-WB decoding. As shown in Fig. 2, the EXIT curve of the AMR-WB decoder, which is denoted by the line marked with triangles, cannot reach the convergence point of $(1,1)$ and intersects with the EXIT curve of the inner SP demapper, which implies that residual errors persist. On the other hand, if the intermediate RSC decoder and the AMR-WB decoder are viewed as a combined outer SISO module, then the joint EXIT function of this module is capable of reaching the convergence point of $(1,1)$, which can be described by $T_c^p[I_{3,E}(\tilde{c}), I_{2,A}(c)]$. The joint EXIT function is denoted by the dotted bold line in Fig. 2. Furthermore, this joint EXIT function characterizes the best possible attainable performance when exchanging information between the RSC decoder and the AMR-WB decoder of Fig. 1 for different fixed values of $I_{2,A}(c)$. The EXIT curves of the SP demapper for various E_b/N_0 values and the EXIT curve of the RSC decoder used in the DSTS–SP–RSC benchmarker scheme are also shown. This EXIT chart can therefore be used to determine the convergence threshold in terms of the minimum E_b/N_0 value required. It can be shown in Fig. 2 that there is an open tunnel between the joint EXIT curve and that of the SP demapper at $E_b/N_0 = 5.0$ dB. By contrast, the EXIT curve of the SP demapper and that of the RSC decoder of the benchmarker scheme employing no outer iterations exhibit an open tunnel at $E_b/N_0 = 6.0$ dB. Thus, according to the EXIT chart predictions, the three-stage system outperforms its benchmark scheme.

V. RESULTS AND DISCUSSION

In this section, we characterize the attainable performance of the proposed scheme using both the BER and the segmental signal-to-noise ratio (SegSNR) [2] evaluated at the speech decoder's output as a function of the channel SNR. We consider a two-transmit-antenna-aided DSTS–SP system associated with $L = 16$ and a single receiver antenna, whereas the remaining simulation parameters were described in Section II. Fig. 3 depicts the BER versus SNR per bit, namely,

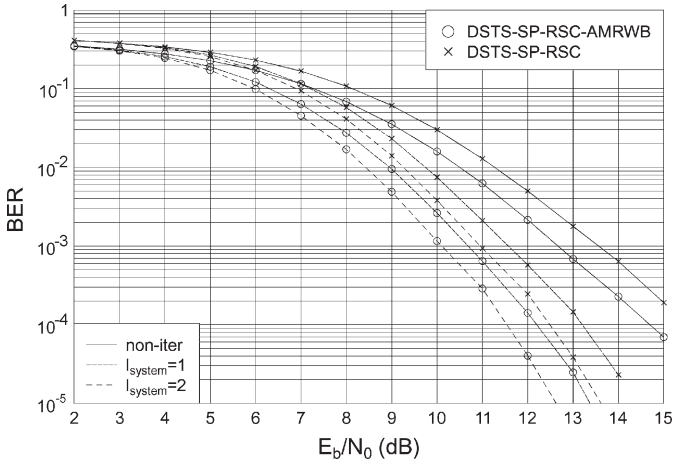

Fig. 3. BER versus E_b/N_0 performance of the jointly optimized DSTS–SP–RSC–AMRWB scheme of Fig. 1 when communicating over correlated nondispersive Rayleigh fading channels.

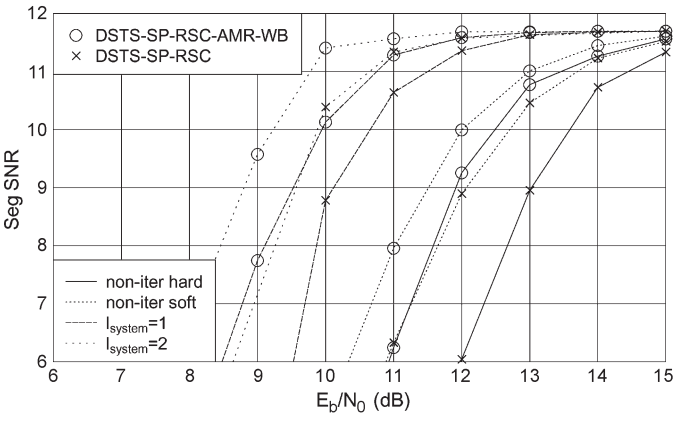

Fig. 4. Average SegSNR versus E_b/N_0 performance of the jointly optimized DSTS–SP–RSC–AMRWB scheme of Fig. 1 when communicating over correlated nondispersive Rayleigh fading channels.

versus E_b/N_0 performance of the DSTS–SP–RSC–AMRWB scheme and that of its corresponding DSTS–SP–RSC benchmarker when communicating over narrow-band correlated Rayleigh fading channels. It can be shown in Fig. 3 that the DSTS–SP–RSC–AMRWB scheme outperforms the DSTS–SP–RSC benchmarker scheme by about 1 dB at BER $= 4.0 \times 10^{-5}$ after $I_{\text{system}} = 2$ iterations. The AMR-WB-decoded scheme has a lower BER at its speech-decoded output than its benchmarker dispensing with iterative speech decoding, because the *extrinsic* information exchange between the AMR-WB decoder and the RSC decoder has the potential of improving the attainable BER. In Fig. 2, it is expected that the DSTS–SP–RSC–AMRWB scheme outperforms the DSTS–SP–RSC benchmarker scheme at $E_b/N_0 = 5.0$ dB. This is indeed expected, since there is an open EXIT tunnel for the DSTS–SP–RSC–AMRWB scheme at $E_b/N_0 = 5.0$ dB, which is expected to lead to a low BER. However, due to the short interleaver length of 461 bits, the actual iterative decoding trajectories do not closely follow the EXIT characteristics, particularly when increasing the number of iterations [21], as shown in Fig. 2. More explicitly, the actual decoding trajectory of Fig. 2. recorded for $I_{\text{system}} = 10$ iterations at $E_b/N_0 = 5.0$ dB was unable to reach $I_{2,E}(c) = 1.0$, and hence, the combined system's actual BER failed to reach an infinitesimally low value. Let us now study the speech SegSNR performance of the proposed scheme in Fig. 4. It can be shown in Fig. 4 that the exploitation of the residual source redundancy during the parameter estimation in soft-bit speech decoding [3] provides valuable *a priori* information. More explicitly, at the point of tolerating a SegSNR degradation of 1 dB, the employment of soft-bit-assisted AMR-WB decoder performs approximately 0.5 dB better in terms of the required channel E_b/N_0 value than its corresponding hard speech decoding-based counterpart. Additionally, iteratively exchanging the soft information among the three receiver components of the amalgamated DSTS–SP–RSC–AMRWB scheme has resulted in a further E_b/N_0 gain of about 3.0 dB after $I_{\text{system}} = 2$ iterations, again to the point of tolerating a SegSNR degradation of 1 dB.

VI. Conclusion

In this paper, the three-stage turbo-detection-aided DSTS–SP–RSC–AMRWB scheme of Fig. 1 was proposed for transmission over a temporally correlated narrow-band Rayleigh fading channel. The employment of the soft-output AMR-WB speech codec, which exploits the residual redundancy inherent in the encoded bit stream, demonstrates a significant improvement in terms of the average SegSNR versus channel E_b/N_0 performance compared with its corresponding hard decoding-based benchmarker. The performance of the three-component turbo receiver is about 1 dB better in terms of the E_b/N_0 required, as compared with the benchmarker scheme also employing joint iterative channel decoding and DSTS-aided SP demodulation but using separate noniterative AMR-WB decoding.

References

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 623–656, Jul./Oct. 1948.

[2] L. Hanzo, F. C. A. Somerville, and J. P. Woodard, *Voice and Audio Compression for Wireless Communications*, 2nd ed. Chichester, U.K.: Wiley, 2007.

[3] T. Fingscheidt and P. Vary, "Softbit speech decoding: A new approach to error concealment," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 3, pp. 240–251, Mar. 2001.

[4] M. Adrat, P. Vary, and J. Spittka, "Iterative source–channel decoder using extrinsic information from softbit-source decoding," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 7–11, 2001, pp. 2653–2656.

[5] T. Clevorn, J. Brauers, M. Adrat, and P. Vary, "Turbo decodulation: Iterative combined demodulation and source–channel decoding," *IEEE Commun. Lett.*, vol. 9, no. 9, pp. 820–822, Sep. 2005.

[6] R. Perkert, M. Kaindl, and T. Hindelang, "Iterative source and channel decoding for GSM," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 7–11, 2001, vol. 4, pp. 2649–2652.

[7] 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 Trans. Speech Audio Process.*, vol. 10, no. 8, pp. 620–636, Nov. 2002.

[8] M. El-Hajjar, O. Alamri, and L. Hanzo, "Differential space–time spreading using iteratively detected sphere packing modulation and two transmit antennas," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Apr. 2006, vol. 3, pp. 1664–1668.

[9] W. Su, Z. Safar, and K. J. R. Liu, "Space–time signal design for time-correlated Rayleigh fading channels," in *Proc. IEEE Int. Conf. Commun.*, May 2003, vol. 5, pp. 3175–3179.

[10] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[11] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space–time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.

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

[13] *AMR Wideband Speech Codec: Transcoding Functions*, Dec. 2001. 3GPP TS 26.190 V5.

[14] 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: Wiley, 2002.

[15] L. Hanzo, T. H. Liew, and B. L. Yeap, *Turbo Coding, Turbo Equalization and Space Time Coding for Transmission Over Wireless Channels*. New York: Wiley, 2002.
[16] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain," in *Proc. IEEE Int. Conf. Commun.*, Jun. 18–22, 1995, vol. 2, pp. 1009–1013.
[17] M. Adrat, U. von Agris, and P. Vary, "Convergence behavior of iterative source–channel decoding," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Apr. 6–10, 2003, pp. 269–272.
[18] M. Adrat, "Iterative source–channel decoding for digital mobile communications," Ph.D. dissertation, RWTH Aachen Univ., Aachen, Germany, 2003.
[19] F. Itakura, "Line spectrum representation of linear predictive coefficients of speech signals," *J. Acoust. Soc. Amer.*, vol. 57, p. S35, Apr. 1975.
[20] Y. Bistritz and S. Peller, "Immittance spectral pairs (ISP) for speech encoding," in *Proc. IEEE ICASSP*, Apr. 27–30, 1993, vol. 2, pp. 9–12.
[21] M. Tüchler, "Design of serially concatenated systems depending on the block length," *IEEE Trans. Commun.*, vol. 52, no. 2, pp. 209–218, Feb. 2004.

Generalized Performance Analysis of Group-Orthogonal Multicarrier Systems

Felip Riera-Palou, *Member, IEEE*, and
Guillem Femenias, *Member, IEEE*

***Abstract*—In this paper, we present a new unified bit-error-rate analysis of group-orthogonal multicarrier (GO-MC) systems for both uplink and downlink segments when using a maximum-likelihood (ML) multiuser and multisymbol detector, respectively. The proposed analysis is based on the union bound, and it is general enough to allow each of the transmitted symbols in the group to come from a different modulation alphabet and have different received power. Monte Carlo simulation results are also presented to assess the accuracy of the derived analytical expressions. A very good agreement between analytical and simulation results can be appreciated, highlighting the usefulness of our approach in the assessment and planning of generalized GO-MC systems.**

***Index Terms*—Group orthogonal, M-ary modulation, maximum likelihood (ML), multicarrier.**

I. Introduction

Group-orthogonal multicarrier code-division multiple access (GO-MC-CDMA) has recently been proposed and analyzed in [1] as a combination of multicarrier code-division multiple access (MC-CDMA) [2] and orthogonal frequency-division multiple access (OFDMA), suitable for the uplink segment of a wireless system. The basic idea behind GO-MC-CDMA is to partition the available (orthogonal) subcarriers into (orthogonal) groups and distribute users among the groups. The main advantage of this system is that each group functions as an independent MC-CDMA system with a small number of users. This makes the use of maximum-likelihood multiuser detection (ML-MUD) within each group computationally feasible. Although low-complexity versions of the ML algorithm (e.g., sphere detection) relax the computational constraints, employing too many subcarriers per group can even degrade the bit-error-rate (BER) performance. In line with this, a new analysis of ML-MUD for GO-MC-CDMA has been presented in [3] and [4], identifying the group size as a crucial parameter that serves to balance diversity and multiuser interference and which, therefore, should be carefully chosen. Group-orthogonal multicarrier (GO-MC) has also been proposed and analyzed as a suitable technique for the downlink segment [5], [6]. In the resulting scheme, termed group-orthogonal multicarrier code-division multiplex (GO-MC-CDM), the groups are used to code-multiplex symbols from the same user who, at reception, can apply joint detection (multisymbol rather than multiuser) using also ML. A particularly limiting characteristic of all previous BER analyses of GO-MC systems [1], [3], [4], [6] was the assumption that all transmitted symbols within a group use the same modulation format and also presuppose that they are all received with equal power. This paper expands previous analyses of ML detection in GO-MC systems. As in previous works, the new analysis is based on the union bound; however, it is applicable to the cases where modulation order and received power might be different for each user/symbol in a group. Relaxation of both restrictions leads to a general closed-form expression, which can simplify network management and resource allocation. In the uplink, the differences in modulation order and power requirements among users in the group might be due to distinct quality of service (QoS) constraints or to the different operating conditions (e.g., particular channel realization and distance from the base station) of the diverse users. In the downlink, the symbols in the group, despite directed to the same user, may correspond to different services and, therefore, be also subject to different QoS constraints.

Manuscript received November 21, 2007; revised February 20, 2008 and April 7, 2008. First published April 18, 2008; current version published January 16, 2009. This work was supported in part by the Ministerio de Educacion y Ciencia (MEC) and Fondos Europeos para el Desarrollo Regional (FEDER) under Project MARIMBA (TEC2005-0997), by Govern de les Illes Balears under Project XISPES (PROGECIB-23A) and Grant PCTIB-2005GC1-09, and by a Ramon y Cajal fellowship (partially funded by the European Social Fund), Spain. The review of this paper was coordinated by Prof. Y. Ma.

The authors are with the Mobile Communications Group, Department of Mathematics and Informatics, University of the Balearic Islands, 07122 Palma, Spain (e-mail: felip.riera@uib.es; guillem.femenias@uib.es).

Digital Object Identifier 10.1109/TVT.2008.923692

II. System Model

We focus on a single group of a GO-MC system (either uplink or downlink) using N subcarriers, where $K \leq N$ symbols are multiplexed for transmission. In the case of the downlink, transmission is inherently synchronous, whereas in the uplink, users are assumed to operate quasi-synchronously. Quasi-synchronicity refers to the fact that the delay between the first and last users in the group to reach the base station is less than the duration of the cyclic prefix (CP), thus allowing the relative delays of the users to be absorbed in the random phases of the subcarriers [1].

Let $\mathbf{a} = [a^0 \ \cdots \ a^{K-1}]^T$ denote the transmitted symbol block in a group, with each symbol a^i potentially drawn from a different M-ary normalized complex-valued symbol alphabet $\mathcal{A}^i$ satisfying $E\{|a^i|^2\} = 1$. Without loss of generality, each symbol a^i is allocated a transmission power p_i, and it is multiplied by a different spreading code of the form $\mathbf{c}^i = [c_0^i \ \cdots \ c_{N-1}^i]^T$ with $E\{|c_n^i|^2\} = 1/N$. The resulting spread symbols are added up and modulated, typically using the inverse fast Fourier transform, onto the group of N orthogonal subcarriers assigned to the symbol block $\mathbf{a}$. A CP is appended to the resulting signal to fight against the effects of the channel dispersion. In the analysis presented in this paper, we will assume that the CP is of sufficient length to completely eliminate any interblock interference.

The transmitted signal propagates through a frequency-selective channel with a scenario-dependent power delay profile $\mathcal{P}(\tau)$ given by

$$\mathcal{P}(\tau) = \sum_{l=0}^{P-1} \phi(l)\delta(\tau - \tau_l) \tag{1}$$