Near-Capacity Iteratively Decoded Markov-Chain Monte-Carlo Aided BLAST System

Wei Liu, Lingkun Kong, Soon Xin Ng and Lajos Hanzo

***Abstract*— In this treatise, we propose an iteratively decoded Bell-labs LAyered Space-Time (BLAST) scheme, which serially concatenates an IRregular Convolutional Code (IRCC), a Unity-Rate Code (URC) and a BLAST transmitter. The proposed scheme is capable of achieving a near capacity performance with the aid of our EXtrinsic Information Transfer (EXIT) chart assisted design procedure. Furthermore, a Markov Chain Monte Carlo (MCMC) based BLAST scheme is employed, which is capable of significantly reducing the complexity imposed. For the sake of approaching the maximum achievable rate, iterative decoding is invoked to attain decoding convergence by exchanging extrinsic information among the three serial component decoders. Our simulation results show that the proposed MCMC-based iteratively detected IRCC-URC-BLAST scheme is capable of approaching the system capacity.**

I. Introduction

Multiple input multiple output (MIMO) systems are capable of supporting high-rate, high-integrity transmission [1]. In [2], Wolniansky *et al.* proposed the popular multilayer MIMO structure, referred to as the Vertical Bell-labs LAyered Space-Time (V-BLAST) scheme, which is capable of increasing the throughput without any increase in the transmitted power or the systems bandwidth.

For a coded system, in order to achieve decoding converge to an infinitesimally low bit error ratio (BER), the BLAST scheme is serially concatenated with outer codes for iteratively exchanging mutual information between the constituent decoders. The decoding convergence of iteratively decoded schemes can be analysed using EXtrinsic Information Transfer (EXIT) charts [3], [4]. Tüchler and Hagenauer [4], [5] proposed the employment of IRregular Convolutional Codes (IRCCs) in serial concatenated schemes, which are constituted by a family of convolutional codes having different rates, in order to design a near-capacity system. They were specifically designed with the aid of EXIT charts to improve the convergence behaviour of iteratively decoded systems. Furthermore, it was shown in [6], [7] that a recursive Unity-Rate Code (URC) should be employed as an intermediate code in order to improve the attainable decoding convergence.

W. Liu is with the Broadband Wireless Communications Laboratory, Information Science Institute, State Key Laboratory of ISN, Xidian University, Xi'an, China, 710071. E-Mail: liuweixd@mail.xidian.edu.cn

L. Kong, S. X. Ng and L. Hanzo are with the University of Southampton, Communications Research Group, Southampton, SO17 1BJ, U.K. E-Mail: {lk06r,sxn,lh}@ecs.soton.ac.uk

The financial support of the European Union under the auspices of the Optimix project, as well as that of the EPSRC UK is gratefully acknowledged. The financial support of the National Science Fund for Distinguished Young Scholars (60725105), of National Basic Research Program of China (973 Program) (2009CB320404), of the China 863 High Tech Key Project (2007AA01Z288), and the China 111 Project (B08038) are greatly acknowledged.

In MIMO schemes, the optimum performance can be achieved by the maximum likelihood (ML) soft demapper at the cost of a potentially high receiver complexity, especially for a large number of transmit antennas or for a high-order modulation scheme. In order to mitigate the complexity imposed, reduced complexity, suboptimal detection algorithms may be used, such as for example Sphere Decoding (SD), Markov Chain Monte Carlo (MCMC) detection [8], [9] etc. may achieve a near-optimal performance at a reasonable complexity. It was shown in [8] that the MCMC aided algorithm has the potential of outperforming the SD aided one, hence we opted for using the MCMC aided algorithm in this paper.

The novel contribution of this treatise is that we design an iteratively decoded reduced-complexity near-capacity three-stage IRCC-URC-BLAST scheme. Specifically, the computational complexity of this concatenated system is reduced by a factor of 256/50 or 256/60, at the cost of a modest reduction in the maximum achievable rate compared to ML detection, owing to the employment of the low-complexity, but near-optimum MCMC demapper in the BLAST detector.

II. System Overview

The schematic of the proposed serially concatenated system is illustrated in Fig. 1. The transmitter consists of three components, an IRCC encoder, a URC encoder and a BLAST. Furthermore, two different high-length bit interleavers are introduced between the three encoder components to guarantees that the assumptions facilitating the application of EXIT charts are complied with [3].

The IRCC encoder takes the information bits u_1 and outputs the coded bits c_1, where each input-stream fraction's code rate was designed for achieving a near-capacity performance with the aid of EXIT charts [3]. 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 N denote the total number of encoded bits generated from the K uncoded information bits. Each subcode encodes a fraction of $\alpha_k r_k N$ of the original uncoded information bits and generates $\alpha_k N$ encoded bits. Given the overall average code rate target of $R \in [0,1]$, the weighting coefficient α_k has to satisfy:

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

Clearly, the individual code rates r_k and the weighting coefficients α_k play a crucial role in shaping the EXIT function of the resultant IRCC. For example, in [5] 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 maximising the achievable free distance. In the proposed system the two additional generators 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 $0.1, 0.15, 0.2, \cdots,$ to 0.9.

The EXIT function of an IRCC can be obtained from those of its subcodes. More specifically, the EXIT function of the target IRCC is the weighted superposition of the EXIT functions of its subcodes [5]. Hence, a careful selection of the weighting coefficients α_k could produce an outer code EXIT curve that closely matches the shape of the inner code EXIT curve. When the area between the two EXIT curves is minimized, decoding convergence would be achieved at the lowest possible Signal-to-Noise Ratios (SNR).

Following the IRCC encoder, a recursive URC was employed to encode the information bits u_2 and output coded bits c_2. It was shown in [6], [7] that a recursive code is needed as an intermediate code, when the inner code is non-recursive, in order to achieve decoding convergence at a low SNR. The URC employed has a generator polynomial of $\frac{1}{1+D}$ and it is used as an intermediate code between the IRCC and BLAST schemes.

Assume that the number of transmit antennas is N_t for an M-ary modulation scheme. At time instant t, the BLAST encoder maps $(N_t \cdot B)$ bits of the information bit stream u_3, expressed as a vector $\boldsymbol{b} = [b_1, b_2, \cdots, b_{N_t \cdot B}]$, where $B = \log_2 M$, into an N_t-component transmitted symbol vector $\boldsymbol{x}$ expressed as $\boldsymbol{x} = [x_1, x_2, \cdots, x_{N_t}]^T$. Furthermore, assume that the number of receive antennas is N_r. Then the received length-(N_r) observation vector $\boldsymbol{y}$ at time instant t can be expressed with the aid of the channel impulse response (CIR) matrix $\boldsymbol{H}$ connecting the N_t transmit antennas with the N_r receive antennas at time instant t as

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}, \qquad (2)$$

where, again, $\boldsymbol{H}$ is the $(N_r \times N_t)$-component CIR matrix. Specifically, flat fading is assumed. Furthermore, $\boldsymbol{n}$ is a length-N_r noise observation vector, which is assumed to be Gaussian distributed with a zero mean and a covariance matrix given by $\sigma^2 \boldsymbol{I}_{N_r}$.

According to Fig. 1, an iterative decoding procedure is operated at the receiver side, which employs three A Posteriori Probability (APP)-based decoders. The received signals of Fig. 1 are first detected by the APP-based BLAST detector in order to produce the *a posteriori* log-Likelihood Ratio (LLR) values $L_{3,p}(u_3)$ of the information bits u_3. Extrinsic information is iteratively exchanged between the BLAST detector, URC decoder and the IRCC decoder.

III. MCMC Aided BLAST Detector

At the BLAST detector, in order to obtain the LLR values $L_{3,p}(u_3)$, the *a posteriori* probability of each bit b_k of $\boldsymbol{b}$ is required. To this end, the optimal performance is achieved by the ML soft demapper, which can be expressed as

$$P[b_k = +1|\boldsymbol{y}, L_a(\boldsymbol{b})] = \sum_{\boldsymbol{b}_{-k}} P[b_k = +1, \boldsymbol{b}_{-k}, \boldsymbol{y}, L_a(\boldsymbol{b})], \quad (3)$$

where the $(N_t \cdot B)$-component vector $\boldsymbol{b}_{-k} = [b_1, \cdots, b_{k-1}, b_{k+1}, \cdots, b_{N_t \cdot B}]$ is obtained by removing the kth bit b_k from the transmitted bit vector $\boldsymbol{b}$ and the summation is over all possible combinations of $\boldsymbol{b}_{-k}$. Still referring to Eq. (3), $L_a(\boldsymbol{b})$ represents the *a priori* LLR values of the information bits $\boldsymbol{b}$. For large number of transmit antennas N_t or for a high number of modulation levels M, the number of combinations grows exponentially, which makes the employment of the ML detector prohibitive for practical application. Instead of evaluating all the combinations, the MCMC algorithm was shown to succeed in selecting only the influential combinations, resulting in a low complexity, but still approaching the optimal performance.

In the context of the MCMC algorithm, the Gibbs sampler was employed to generate the Markov Chain, which can be described as follows [8], [9]:

1) Initialize $\boldsymbol{b}$ randomly;
2) for $i = 1$ to N_{MC} do
   draw sample from $P[b_1|\boldsymbol{b}^{i-1}_{-1}, \boldsymbol{y}, L_a(\boldsymbol{b})]$;
   draw sample from $P[b_2|\boldsymbol{b}^{i}_{-2}, \boldsymbol{y}, L_a(\boldsymbol{b})]$;
   $\vdots$
   draw sample from $P[b_{N_t \cdot B}|\boldsymbol{b}^{i}_{-N_t \cdot B}, \boldsymbol{y}, L_a(\boldsymbol{b})]$;
   if $i \geq 0$ add sample $\boldsymbol{b}^i$ to
   i++
   end

where N_{MC} is the length of a single Markov Chain.

Since the samples generated from a single Markov Chain are correlated, this may result in insufficiently important diverse samples. As an alternative solution, L number of parallel Markov Chains can be generated, resulting in $L \cdot N_{MC}$ samples. Afterwards, the repetition of identical samples is removed and only N_s different samples are retained [8], [9]. Upon using these N_s different samples, the *a posteriori* probability of b_k is approximated as [8], [9]:

$$P(b_k = +1|\boldsymbol{y}, L_a(\boldsymbol{b})) \approx \frac{\sum_{i=1}^{N_s} P[b_k = +1|\boldsymbol{y}, \boldsymbol{b}^i_{-k}, L_a(\boldsymbol{b})] P[\boldsymbol{b}^i_{-k}|\boldsymbol{y}, L_a(\boldsymbol{b})]}{\sum_{i=1}^{N_s} P[\boldsymbol{b}^i_{-k}|\boldsymbol{y}, L_a(\boldsymbol{b})]}. \quad (4)$$

IV. System Design

For discrete-amplitude QAM or PSK [10] modulation, we encounter a Discrete-input Continuous-output Memoryless Channel (DCMC) [10]. In order to design a near-capacity coding scheme, we have to derive the bandwidth efficiency η of various BLAST schemes for transmission over the DCMC. This will be achieved based on the properties of EXIT charts [11] as detailed in the next paragraph. In this contribution, both the full-rank scenario of $N_t = 4$ transmit and $N_r = 4$ receive antennas and the rank-deficient scenario of $N_t = 4$ transmit and $N_r = 2$ receive antennas are considered.

It was claimed in [5], [11] that the maximum achievable rate of the system is the same as the area under the EXIT curve of the inner code, when the channel's input is independently and uniformly distributed. Furthermore, the area under the EXIT curve of the outer code is approximately equal to (1-R), where R is the outer code rate. Assuming that the area under the EXIT curve of the inner decoder, i.e. the BLAST detector, is represented by A_E, the maximum achievable rate curves of two BLAST schemes are shown in Fig. 2 together with the capacity curves of the unrestricted Continuous-input Continuous-output Memoryless Channel (CCMC) [10], [12] for comparison. It can be seen from Fig. 2 that the more receive antennas are used, the higher the capacity in both the CCMC and DCMC scenarios. Furthermore, the capacity of the DCMC scenario was upper-bounded by that of the CCMC for both full-rank and rank-deficient scenarios.

The main objective of employing EXIT charts [3] is to analyse the convergence behaviour of iterative decoders by examining the evolution of the input/output mutual information exchange between the inner and outer decoders during the consecutive iterations. As mentioned above, the area under the EXIT curve of the inner decoder is approximately equal to the channel capacity, when the channel's input is independently and uniformly distributed. Similarly, the area under the EXIT curve of the outer code is approximately equal to (1-R), where R is the outer code rate. Furthermore, our experimental results show that an intermediate URC changes only the shape but not the area under the EXIT curve of the inner code. A narrow, but marginally open EXIT-tunnel in an EXIT chart indicates the possibility of achieving a near-capacity performance. Therefore, we invoke IRCCs for the sake of appropriately shaping the EXIT curves by minimizing the area in the EXIT-tunnel using the procedure of [4], [5].

Again, the EXIT function of an IRCC can be obtained by superimposing those of its subcodes. More specifically, the EXIT function of the target IRCC is the weighted superposition of the EXIT functions of its subcodes [5]. Hence, a careful selection of the weighting coefficients could produce an outer code EXIT curve that matches closely the EXIT curve of the inner code. When the area between the two EXIT curves is minimized, decoding convergence to an infinitesimally low BER would be achieved at the lowest possible SNR.

V. Simulation Results

In this section, numerical results are provided in order to characterize the proposed scheme. Specifically, $N_t = 4$ transmit and $N_r = 4$ or $N_r = 2$ receive antennas are employed and 4QAM was adopted. Furthermore, we invoked $L = 10$ parallel Gibbs samplers of length $N_{MC} = 5$ in the full-rank (4×4) system and $L = 15$ parallel Gibbs samplers of length $N_{MC} = 4$ in the rank-deficient (4×2) system. Hence, we have 50 and 60 samples out of 256 possible combinations for the approximation of the extrinsic information in Eq. (3) in the full-rank and rank-deficient scenarios, respectively. Moveover, the IRCC outer code having an average coding rate of $R = 0.5$ was employed, resulting in the effective throughput of $\eta = 4 \cdot 2 \cdot \frac{1}{2} = 4$ bit/s/Hz for 4QAM, while the channel capacity and the maximum achievable rate computed according to the properties of EXIT charts [5], [11] at a throughput of 4 bit/s/Hz are depicted in Fig. 2.

In Fig. 3 the exchange of extrinsic information in the schematic of Fig. 1 is characterized by an EXIT chart for the full-rank system. The EXIT curve of the MCMC-BLAST scheme is a slanted line, which crosses the EXIT curve of the outer code and hence prevents us from reaching the (1.0, 1.0) point of perfect convergence. By contrast, when relying on the extrinsic information exchange between the URC decoder and the BLAST detector, the curve reaches the (1.0, 1.0) point and hence becomes capable of achieving a near-capacity performance. When there is no iteration between the URC decoder and the MCMC detector, the EXIT curve shape of the URC decoder depends on the initial I_E value provided by the BLAST detector at $I_A = 0$. Hence, the BLAST-URC scheme requires a higher E_b/N_0 value in order to maintain an area of $A_E = 0.5$ than the scheme having iterations between the URC decoder and the MCMC detector, as shown in Fig. 1. In other words, a throughput loss will occur, if there is no iteration between the URC decoder and the BLAST detector.

As we can see from Fig. 3 the Monte-Carlo simulation based decoding trajectory of the (MCMC-BLAST)-URC-IRCC scheme only has slight mismatches in comparison to the corresponding EXIT curves. This is due to the reduced-complexity approximation of the extrinsic information provided by the MCMC demappers, because we use 50 out of 256 possible samples in the computation of Eq. (3). We apply the same technique for a rank-deficient (4×2) system, where only 60 out of 256 samples are used in Eq. (3). Fig. 4 displays the BER performance of the (MCMC-BLAST)-URC-IRCC schemes. As we can see from Fig. 4, the (MCMC-BLAST)-URC-IRCC scheme employing $N_t = 4$ transmit and $N_r = 4$ receive antennas is capable of working within 0.3-0.4 dB of the corresponding maximum achievable rate obtained with the aid of our EXIT chart assisted technique, while the (MCMC-BLAST)-URC-IRCC scheme invoked in the rank-deficient scenario performs within 1.0 dB of the corresponding DCMC capacity.

VI. Conclusions

In this paper, we investigated a MCMC aided iteratively decoded BLAST-URC-IRCC scheme with the aid of EXIT chart analysis. The simulation results show that the proposed scheme is capable of achieving a near-capacity performance at a reduced complexity, when using 50 and 60 out of 256 samples in Eq. (3) to approximate the *a posteriori* probability in the full-rank and rank-deficient scenarios, respectively.

References

[1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, May 1999.

[2] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *International Symposium on Signals, Systems, and Electronics*, Pisa, Italy, 29 September - 2 October 1998, pp. 295–300.

[3] S. T. Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Transactions on Communications*, vol. 49, no. 10, pp. 1727 – 1737, October 2001.

[4] M. Tüchler and J. Hagenauer, "EXIT charts of irregular codes," in *Proceeding of the 36th Annual Conference on Information Sciences and Systems*, Princeton, NJ, USA, 4-7 March 2002, p. CDROM.

[5] 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.

[6] ——, "Convergence prediction for iterative decoding of threefold concatenated systems," in *IEEE Global Telecommunications Conference*, Taipei, Taiwan, 17 - 21 November 2002, pp. 1358– 1362.

[7] S. X. Ng, J. Wang, M. Tao, L. L. Yang, and L. Hanzo, "Iteratively decoded variable length space-time coded modulation: code construction and convergence analysis," *IEEE Transactions on Wireless Communications*, vol. 6, no. 5, pp. 1953 – 1963, May 2007.

[8] H. Zhu, B. Farhang-Boroujeny, and R. R. Chen, "On performance of sphere decoding and Markov Chain Monte Carlo detection methods," *IEEE Signal Processing Letters*, vol. 12, no. 10, pp. 669 – 672, October 2005.

[9] B. Farhang-Boroujeny, H. Zhu, and Z. Shi, "Markov chain Monte Carlo algorithms for CDMA and MIMO communication systems," *IEEE Transactions on Signal Processing*, vol. 54, no. 5, pp. 1896 – 1909, May 2006.

[10] J. G. Proakis, *Digital Communications*, 4th ed. McGraw Hill, 2000.

[11] A. Ashikhmin, G. Kramer, and S. T. Brink, "Extrinsic information transfer functions: model and erasure channel properties," *IEEE Transactions on Information Theory*, vol. 50, no. 11, pp. 2657 – 2673, November 2004.

[12] S. X. Ng and Hanzo, "On the MIMO Channel Capacity of Multi-Dimensional Signal Sets," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 2, pp. 528–536, March 2006.

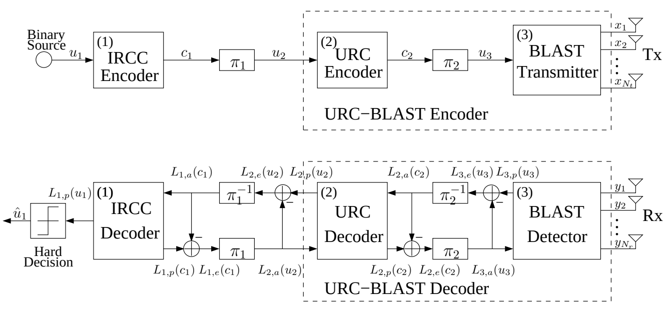

Fig. 1. Schematic of the proposed IRCC-URC-BLAST scheme.

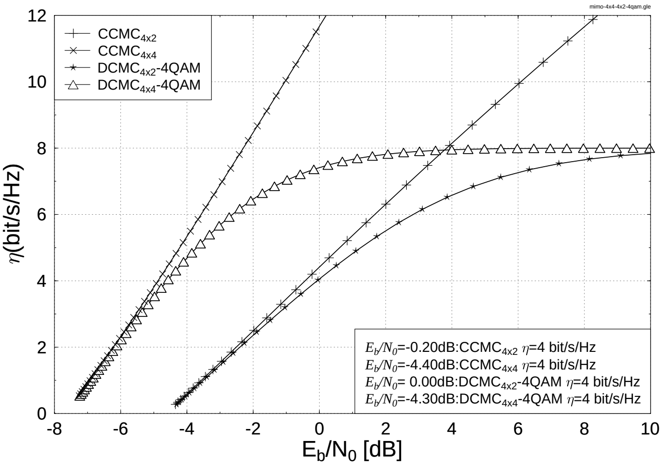

Fig. 2. The capacity of BLAST schemes when communicating over uncorrelated flat Rayleigh fading channels for both $Nt = 4$ transmit and $Nr = 4$ receive antennas and $N_t = 4$ transmit and $N_r = 2$ receive antennas.

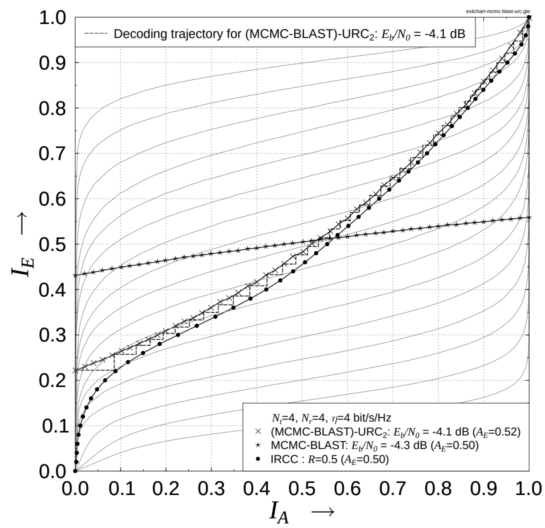

Fig. 3. The EXIT chart curves for the (MCMC-BLAST)-URC, IRCC and the IRCC subcodes, when communicating over uncorrelated flat Rayleigh fading channels using $N_t = 4$ transmit and $N_r = 4$ receive antennas. The subscript of URC denotes the number of iterations between the BLAST and the URC decoders.

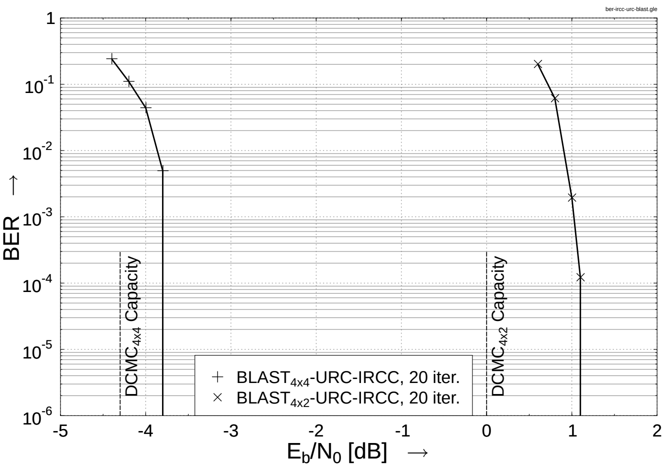

Fig. 4. The BER performance of the proposed (MCMC-BLAST)-URC-IRCC scheme, when communicating over uncorrelated flat Rayleigh fading channels using $N_t = 4$ transmit and $N_r = 4$ receive antennas.