The plots clearly indicate that memoryless decoding is strictly suboptimal to IR-ARQ for all rates $R \geq 0.19$ (with $L = 2$) and $R \geq 0.46$ (with $L = 4$). Moreover, when $L = 4$, the proposed IR-ARQ scheme achieves the feedback exponent $E_F(R)$ for all rates below capacity (at the moment, we do not have a proof that this observation holds universally as in the case of BSCs).

V. Conclusion

We considered the error exponents of memoryless ARQ channels with an upper bound L on the maximum number of retransmission rounds. In this setup, we have established the superiority of IR-ARQ, as compared with Forney's memoryless decoding. For the BSC and VNC, our results show that choosing $L = 4$ is sufficient to ensure the achievability of Forney's feedback exponent, which is typically achievable with memoryless decoding in the asymptotic limit of large L. Finally, in the AWGN channel, numerical results also show the superiority of IR-ARQ over memoryless decoding, in terms of the achievable error exponent.

References

[1] M. V. Burnashev, "Data transmission over a discrete channel with feedback: Random transmission time," *Prob. Inf. Transmiss.*, vol. 12, no. 4, pp. 250–265, 1976.

[2] G. D. Forney, Jr., "Exponential error bounds for erasure, list, and decision feedback schemes," *IEEE Trans. Inf. Theory*, vol. 14, no. 2, Mar. 1968.

[3] R. G. Gallager*, Information Theory and Reliable Communication*. New York: Wiley, 1968.

[4] G. Caire and D. Tuninetti, "The throughput of hybrid-ARQ protocols for the gaussian collision channel," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 1971–1988, Jul. 2001.

[5] A. J. Viterbi, "Error bounds for the white gaussian and other very noisy memoryless channels with generalized decision regions," *IEEE Trans. Inf. Theory*, vol. 15, no. 2, pp. 279–287, Mar. 1969.

[6] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels I," *Inf. Control*, vol. 10, pp. 65–103, Jan. 1967.

[7] C. E. Shannon, "Probability of error for optimal codes in a gaussian channel," *Bell Syst. Tech. J.*, vol. 38, pp. 611–656, 1959.

[8] A. Valembois and M. P. C. Fossorier, "Sphere-packing bounds revisited for moderate block lengths," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 2998–3014, Dec. 2004.

[9] S. Dolinar, D. Divsalar, and F. Pollara, Code performance as a function of block size TMO Progress Rep., May 1998, pp. 42–133.

On the Maximal Diversity Order of Spatial Multiplexing With Transmit Antenna Selection

Joakim Jaldén*, Student Member, IEEE,* and Björn Ottersten*, Fellow, IEEE*

***Abstract*—Zhang *et al.* recently derived upper and lower bounds on the achievable diversity of an $N_R \times N_T$ i.i.d. Rayleigh fading multiple antenna system using transmit antenna selection, spatial multiplexing and a linear receiver structure. For the case of $L = 2$ transmitting (out of N_T available) antennas the bounds are tight and therefore specify the maximal diversity order. For the general case with $L \leq \min(N_R, N_T)$ transmitting antennas it was conjectured that the maximal diversity is $(N_T - L + 1)(N_R - L + 1)$ which coincides with the lower bound. Herein, we prove this conjecture for the zero forcing and zero forcing decision feedback (with optimal detection ordering) receiver structures.**

***Index Terms*—Antenna selection, diversity, spatial multiplexing, zero forcing receiver.**

I. Introduction

The multiple antennas in a multiple-input–multiple-output (MIMO) wireless system can be used either to increase the data rate or reliability (diversity) of the wireless link [1]. In order to capitalize on the benefits offered by the MIMO wireless link while maintaining manageable complexity and cost the use of antenna selection has been previously suggested [2]. In a system using antenna selection, only a small subset of the available antennas would typically be used, thereby limiting the number of RF chains required.

In [3] Zhang *et al.* rigorously analyzed the maximal achievable diversity for a system transmitting L independent data-streams from L out of N_T possible transmit antennas in conjuncture with linear (decision feedback) processing at the receiver. In particular, for the case of a block independent and identically distributed (i.i.d.) Rayleigh-fading channel it was shown that the maximal diversity of such a system is bounded between $M_L \triangleq (N_T - L + 1)(N_R - L + 1)$ and $M_U \triangleq (N_T - L + 1)(N_R - 1)$ where N_R is the number of antennas at the receiver. Since $M_L = M_U$ for $L = 2$ these bounds uniquely determine the maximal diversity in the case of two transmitting antennas and thereby analytically prove some previous observations made in the literature [4], [5]. Further, for the general case where $2 < L < \min(N_R, N_T)$ it was in [3] conjectured that the maximal diversity coincides with the lower bound, M_L. Herein, we extend the analysis of [3] by proving this conjecture for the case of the zero forcing (ZF) and ZF-decision-feedback (DF) receivers (with optimal detection ordering). It should however be noted that the cases of the minimum mean square error (MMSE) and MMSE-DF receivers (although with a fixed detection ordering) also follow from our result by applying the analysis in [3].

The structure of this correspondence is as follows. The system model considered is covered in Section II, mainly in order to the introduce notation. The reader is referred to [3] for details regarding the systems model and a proper motivation of the problem considered. Our main contribution is then given in Section III in the form of Theorem 1.

Manuscript received December 19, 2006; revised June 10, 2007.

The authors are with the Signal Processing Lab, School of Electrical Engineering, KTH, Royal Institute of Technology, Stockholm, Sweden (e-mail: joakim.jalden@ee.kth.se; bjorn.ottersten@ee.kth.se)

Communicated by H. Boche, Associate Editor for Communications.

Digital Object Identifier 10.1109/TIT.2007.907455

II. Problem Statement

A. System Model

The case of an N_R by N_T frequency nonselective block Rayleigh-fading channel is considered. The channel matrix is denoted $\mathbf{H} \triangleq [\mathbf{h}_1 \quad \mathbf{h}_2 \quad \cdots \quad \mathbf{h}_{N_T}] \in \mathbb{C}^{N_R \times N_T}$ and is assumed constant over a block of T channel uses. Further, the elements of $\mathbf{H}$ are modeled as i.i.d. circularly symmetric complex Gaussian with zero mean and unit variance. The transmitter selects $L \leq \min(N_T, N_R)$ antennas (corresponding to columns of $\mathbf{H}$) and transmits independently coded data streams from each antenna. As in [3], let U_j denote the jth possible antenna subset where

$$\begin{aligned} U_1 &= \{\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_L\} \\ U_2 &= \{\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_{L-1}, \mathbf{h}_{L+1}\} \\ &\vdots \\ U_{N_U} &= \{\mathbf{h}_{N_T-L+1}, \ldots, \mathbf{h}_{N_T}\} \end{aligned} \tag{1}$$

and where $N_U = \binom{N_T}{L}$ is the total number of such subsets. The channel can then be modeled according to

$$\mathbf{y} = \sqrt{\frac{\rho_0}{L}} \mathbf{H}_j \mathbf{s} + \mathbf{n}. \tag{2}$$

where in the above; $\mathbf{H}_j \in \mathbb{C}^{N_R \times L}$ is the channel matrix containing the columns in the selected subset U_j; where $\mathbf{y} \in \mathbb{C}^{N_R \times T}$ is the signal block received during T channel uses; where $\mathbf{s} \in \mathbb{C}^{L \times T}$ is the transmitted signal block; and where $\mathbf{n} \in \mathbb{C}^{N_R \times T}$ is the circularly symmetric complex Gaussian noise which is assumed spatially and temporally white and of unit variance.

At the receiver, a ZF front-end is used to separate the transmitted data streams according to

$$\tilde{\mathbf{s}} = \mathbf{H}_j^{\dagger} \mathbf{y} = \sqrt{\frac{\rho_0}{L}} \mathbf{s} + \tilde{\mathbf{n}} \tag{3}$$

where $\mathbf{H}_j^{\dagger} = (\mathbf{H}_j^{\mathrm{H}} \mathbf{H}_j)^{-1} \mathbf{H}_j^{\mathrm{H}}$ is the pseudo-inverse of $\mathbf{H}_j$. Since $L \leq N_R$ by assumption it follows that $\mathbf{Q}_j \triangleq \mathbf{H}_j^{\mathrm{H}} \mathbf{H}_j$ is invertivble with probability one. The effective noise, $\tilde{\mathbf{n}}$, is spatially colored with covariance $\mathbf{Q}_j^{-1}$ and the effective postprocessing signal-to-noise ratio (SNR) of the kth data stream is given by

$$\rho_k^{(j)} = \left(\frac{\rho_0}{L}\right) / \left[\mathbf{Q}_j^{-1}\right]_{kk} \tag{4}$$

where $1 \leq k \leq L$ [1], [3]. A given data stream, k, is said to be in *outage* if the postprocessing SNR drops below a given threshold, $\gamma > 0$ and the diversity order, $d_k^{(j)}$, of this stream is defined according to

$$d_k^{(j)} \triangleq \lim_{\rho_0 \to \infty} \frac{\ln \mathrm{P}\left(\rho_k^{(j)} \leq \gamma\right)}{\ln \rho_0^{-1}}. \tag{5}$$

Similarly, let $\bar{\rho}^{(j)}$ and $\underline{\rho}^{(j)}$ denote the *maximal* and *minimal* postprocessing SNRs defined according to

$$\bar{\rho}^{(j)} \triangleq \max_{1 \leq k \leq L} \rho_k^{(j)} \quad \text{and} \quad \underline{\rho}^{(j)} \triangleq \min_{1 \leq k \leq L} \rho_k^{(j)}.$$

Note also that $\underline{\rho}^{(j)} \leq \bar{\rho}^{(j)}$ and that $\bar{\rho}^{(j)} \leq \gamma$ imply that all streams are simultaneously in outage. Thus

$$d_k^{(j)} \leq \bar{d}^{(j)} \triangleq \lim_{\rho_0 \to \infty} \frac{\ln \mathrm{P}\left(\bar{\rho}^{(j)} \leq \gamma\right)}{\ln \rho_0^{-1}} \tag{6}$$

and

$$d_k^{(j)} \geq \underline{d}^{(j)} \triangleq \lim_{\rho_0 \to \infty} \frac{\ln \mathrm{P}\left(\underline{\rho}^{(j)} \leq \gamma\right)}{\ln \rho_0^{-1}} \tag{7}$$

provides upper and lower bounds on the diversity order of the ZF receiver. Note that (6) and (7) also provide upper and lower bounds on the ZF-DF receiver with optimal ordering. This follows since if $\bar{\rho}^{(j)} \leq \gamma$ no data can be reliably decoded and the first data stream decoded is likely to be in error, regardless of the detection ordering policy applied. Similarly, $\underline{\rho}^{(j)} \geq \gamma$ implies that all streams can be reliably decoded by the ZF receiver and thus also by the ZF-DF receiver since the subtraction of correctly decoded streams can only increase the postprocessing SNR.

B. Problem Statement

In general terms, an *antenna selection policy* is characterized by some (measurable) function φ

$$\varphi : \mathbb{C}^{N_T \times N_R} \mapsto \{1, 2, \ldots, N_U\}$$

which selects an antenna subset, U_j, based on the channel matrix realization, $\mathbf{H}$, according to $j = \varphi(\mathbf{H})$. In [3] it is shown that there exists an antenna selection policy, $j = \varphi(\mathbf{H})$, for which

$$\underline{d}^{(j)} = (N_T - L + 1)(N_R - L + 1).$$

This bound in also shown to be tight in the case where $L = 2$ using a geometrical approach. Further, the bound is conjectured to be tight when $L > 2$. Herein, we confirm this conjecture in a positive sense by proving that

$$\bar{d}^{(j)} \leq (N_T - L + 1)(N_R - L + 1)$$

for *any* antenna selection policy, φ. The proof is given in the following section.

III. Proof of Conjecture

In the proof, we let $\succeq$ denote the partial matrix ordering induced by the positive semidefinite (PSD) cone [6]. For Hermitian matrices $\mathbf{A}$ and $\mathbf{B}, \mathbf{A}, \mathbf{B} \in \mathbb{C}^{n \times n}$, we write $\mathbf{A} \succeq \mathbf{B}$ to denote that $\mathbf{A} - \mathbf{B}$ is PSD. In particular, we will use that $[\mathbf{A}]_{kk} \geq [\mathbf{B}]_{kk}$ whenever $\mathbf{A} \succeq \mathbf{B}$ and where $[\mathbf{A}]_{kk}$ and $[\mathbf{B}]_{kk}$ denotes the kth diagonal value of $\mathbf{A}$ and $\mathbf{B}$. Also, $\mathbf{A} \succeq \mathbf{B}$ is equivalent to $\mathbf{A}^{-1} \preceq \mathbf{B}^{-1}$ for strictly positive definite matrices $\mathbf{A}$ and $\mathbf{B}$ and $\mathbf{A} \succeq \mathbf{0}$ if and only if all principal sub-matrices of $\mathbf{A}$ are PSD [7].

We are now ready to state and prove the contribution of this work which is given by Theorem 1 below. Note also that the theorem yields the recently proved [8], [9] statement that detection ordering can not improve the ZF-DF diversity order as a special case by selecting $L = N_T \leq N_R$. It should also be noted that the proof of Theorem 1 is similar to a recently submitted proof [10] of this statement but that the antenna selection case represents a nontrivial extension.

Theorem 1: Given an arbitrary antenna selection policy $j = \varphi(\mathbf{H})$ let $\bar{d}^{(j)}$ be defined as in (6). Then

$$\bar{d}^{(j)} \leq (N_T - L + 1)(N_R - L + 1). \tag{8}$$

Proof: Let $\mathbf{Q} \triangleq \mathbf{H}^{\mathrm{H}} \mathbf{H}, \mathbf{Q}_j \triangleq \mathbf{H}_j{}^{\mathrm{H}} \mathbf{H}_j$ and note that $\mathbf{Q}_j$ is an $L \times L$ principal sub-matrix of $\mathbf{Q}$. Further, let the eigenvalue decomposition of $\mathbf{Q}$ be given by

$$\mathbf{Q} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^{\mathrm{H}}$$

where $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_{N_T})$ are the ordered eigenvalues, $\lambda_1 \geq \cdots \geq \lambda_{N_T}$, of $\mathbf{Q}$ and where $\mathbf{U} = [\mathbf{u}_1 \ \ldots \ \mathbf{u}_{N_T}]$ are the corresponding eigenvectors. Since $\mathbf{Q}$ is unitarily invariant it can be assumed that $\mathbf{U}$

is a Haar matrix and independent of $\mathbf{\Lambda}$ [11]. Let $\mathbf{V} \triangleq [\mathbf{u}_1 \ \ldots \ \mathbf{u}_{L-1}]$ and note that

$$\mathbf{Q} = \sum_{i=1}^{N_T} \lambda_i \mathbf{u}_i {\mathbf{u}_i}^{\mathrm{H}} \preceq \sum_{i=1}^{L-1} \lambda_1 \mathbf{u}_i \mathbf{u}_i^{\mathrm{H}} + \lambda_L \mathbf{I} = \lambda_1 \mathbf{V}\mathbf{V}^{\mathrm{H}} + \lambda_L \mathbf{I}.$$

Let

$$\mathbf{S} \triangleq \frac{\lambda_1}{\lambda_L} \mathbf{V}\mathbf{V}^{\mathrm{H}} + \mathbf{I}.$$

and let $\mathbf{S}_j$ be the $L \times L$ principal submatrix of $\mathbf{S}$ obtained by selecting the rows and columns corresponding to antenna subset j. Note that since $\mathbf{Q} \preceq \lambda_L \mathbf{S}$ it follows that $\mathbf{Q}_j \preceq \lambda_L \mathbf{S}_j$ and in particular $\mathbf{Q}_j^{-1} \succeq \lambda_L^{-1} \mathbf{S}_j^{-1}$ which implies that $[\mathbf{Q}_j^{-1}]_{kk} \geq \lambda_L^{-1} [\mathbf{S}_j^{-1}]_{kk}$ for $k = 1, \ldots, L$.

Let $\mathbf{V}_j \in \mathbb{C}^{L \times (L-1)}$ be the matrix consisting of the L *rows* of $\mathbf{V}$ corresponding to antenna subset j. Note also that $\mathbf{S}_j = \frac{\lambda_1}{\lambda_L} \mathbf{V}_j \mathbf{V}_j^{\mathrm{H}} + \mathbf{I}$. By the matrix inversion lemma it follows that

$$\begin{aligned} \mathbf{S}_j^{-1} &= \left(\frac{\lambda_1}{\lambda_L} \mathbf{V}_j {\mathbf{V}_j}^{\mathrm{H}} + \mathbf{I} \right)^{-1} \\ &= \mathbf{I} - \mathbf{V}_j \left(\frac{\lambda_L}{\lambda_1} \mathbf{I} + \mathbf{V}_j^{\mathrm{H}} \mathbf{V}_j \right)^{-1} {\mathbf{V}_j}^{\mathrm{H}}. \end{aligned} \tag{9}$$

As $\lambda_1, \lambda_L \geq 0$ it follows that

$$\frac{\lambda_L}{\lambda_1} \mathbf{I} + \mathbf{V}_j^{\mathrm{H}} \mathbf{V}_j \succeq \mathbf{V}_j^{\mathrm{H}} \mathbf{V}_j$$

and therefore

$$\left(\frac{\lambda_1}{\lambda_L} \mathbf{I} + \mathbf{V}_j^{\mathrm{H}} \mathbf{V}_j \right)^{-1} \preceq \left(\mathbf{V}_j^{\mathrm{H}} \mathbf{V}_j \right)^{-1} \tag{10}$$

which is equivalent to

$$-\left(\frac{\lambda_1}{\lambda_L} \mathbf{I} + \mathbf{V}_j^{\mathrm{H}} \mathbf{V}_j \right)^{-1} \succeq -\left(\mathbf{V}_j^{\mathrm{H}} \mathbf{V}_j \right)^{-1}. \tag{11}$$

Note also that the inverse on the right hand side of (10) exists with probability one due the unitary invariance of $\mathbf{V}$ (the probability that any L rows are linearly dependent is zero). Now, inserting (11) into (9) yields

$$\mathbf{S}_j^{-1} \succeq \mathbf{I} - \mathbf{V}_j \left(\mathbf{V}_j^{\mathrm{H}} \mathbf{V}_j \right)^{-1} \mathbf{V}_j^{\mathrm{H}} \triangleq \mathbf{P}_j. \tag{12}$$

In the above, $\mathbf{P}_j$ corresponds to a projection onto the null-space of $\mathbf{V}_j^{\mathrm{H}}$ (which has dimension one since $\mathbf{V}_j \in \mathbb{C}^{L \times L-1}$). Note also that for a fixed j (independent of $\mathbf{H}$) the distribution of $\mathbf{V}_j$ is invariant to multiplication from the right by $L \times L$ unitary matrices. This follows from the unitary invariance of $\mathbf{U}$ (and $\mathbf{V}$) [11]. Therefore, the null-space of $\mathbf{V}_j^{\mathrm{H}}$ is unitarily invariant and

$$P([\mathbf{P}_j]_{kk} = 0) = 0$$

for fixed j and k since $[\mathbf{P}_j]_{kk} = \mathbf{e}_k^{\mathrm{H}} \mathbf{P}_j \mathbf{e}_k$ is the squared length of the projection of the kth natural basis vector, $\mathbf{e}_k$, onto the null-space of $\mathbf{V}_j$ (the probability that $\mathbf{e}_k$ is completely orthogonal to the null-space is zero). Since there are a finite number of possible k and j it follows that

$$\mathrm{P}(\exists\, k, j [\mathbf{P}_j]_{kk} = 0) = 0. \tag{13}$$

From (13) it follows that there is some constant, $\kappa > 0$, for which

$$\mathrm{P}(\exists\, k, j [\mathbf{P}_j]_{kk} < \kappa) < 1 \tag{14}$$

or equivalently for which

$$\mathrm{P}(\forall\, k, j [\mathbf{P}_j]_{kk} \geq \kappa) > 0. \tag{15}$$

In particular, for $j = \varphi(\mathbf{H})$, it follows that

$$\mathrm{P}([\mathbf{P}_j]_{kk} \geq \kappa, k = 1, \ldots, L) > 0$$

which states that the probability that all diagonal values of $\mathbf{P}_j$ are simultaneously large (in the sense that they are bounded away from zero) is strictly positive.

For notational convenience in the following, let

$$\tau \triangleq \min_{k,j} [\mathbf{P}_j]_{kk}.$$

Since

$$\left[\mathbf{Q}_j^{-1}\right]_{kk} \geq \lambda_L^{-1} \left[\mathbf{S}_j^{-1}\right]_{kk} \geq \lambda_L^{-1} [\mathbf{P}_j]_{kk} \geq \lambda_L^{-1} \tau$$

it follows, by (4), that

$$\rho_k^{(j)} = \left(\frac{\rho_0}{L} \right) / \left[\mathbf{Q}_j^{-1}\right]_{kk} \leq \frac{\lambda_L \rho_0}{\tau L}$$

for $k = 1, \ldots, L$. Let $\kappa > 0$ be given as in (14) and note that if $\tau \geq \kappa$ and $\lambda_L \leq \kappa \gamma L \rho_0^{-1}$ it follows that $\rho_k^{(j)} \leq \gamma$ for $k = 1, \ldots, L$. This implies

$$\begin{aligned} \mathrm{P}\left(\bar{\rho}^{(j)} < \gamma \right) &\leq \mathrm{P}\left(\lambda_L \leq \kappa \gamma L \rho_0^{-1} \cap \kappa \leq \tau \right) \\ &= \mathrm{P}\left(\lambda_L \leq \kappa \gamma L \rho_0^{-1} \right) \mathrm{P}\left(\kappa \leq \tau \right) \end{aligned}$$

where the last equality follows by the independence of τ (which is a function of $\mathbf{U}$) and λ_L. This implies

$$\frac{\ln \mathrm{P}\left(\bar{\rho}^{(j)} < \gamma \right)}{\ln \rho_0^{-1}} \leq \frac{\ln \mathrm{P}\left(\lambda_L \leq \kappa \gamma L \rho_0^{-1} \right)}{\ln \rho_0^{-1}} + \frac{\ln \mathrm{P}\left(\kappa \leq \tau \right)}{\ln \rho_0^{-1}}$$

where

$$\lim_{\rho_0 \to \infty} \frac{\ln \mathrm{P}\left(\kappa \leq \tau \right)}{\ln \rho_0^{-1}} = 0$$

due to (15) and since $\mathrm{P}(\kappa \leq \tau) > 0$ does not depend on ρ_0. Thus

$$\begin{aligned} \bar{d}^{(j)} &\triangleq \lim_{\rho_0 \to \infty} \frac{\ln \mathrm{P}\left(\bar{\rho}^{(j)} < \gamma \right)}{\ln \rho_0^{-1}} \\ &\leq \lim_{\rho_0 \to \infty} \frac{\ln \mathrm{P}\left(\lambda_L \leq \kappa \gamma L \rho_0^{-1} \right)}{\ln \rho_0^{-1}} + \lim_{\rho_0 \to \infty} \frac{\ln \mathrm{P}\left(\kappa \leq \tau \right)}{\ln \rho_0^{-1}} \\ &= \lim_{\rho_0 \to \infty} \frac{\ln \mathrm{P}\left(\lambda_L \leq \rho_0^{-1} \right)}{\ln \rho_0^{-1}} \\ &= (N_T - L + 1)(N_R - L + 1) \end{aligned}$$

where the last equality follows from [12, Equation (17)] or as a special case of [13, Equation (15)]. This completes the proof and established the assertion made by the theorem. □

IV. Conclusion

We have proved the conjecture of Zhang *et al.* in [3] regarding the diversity order of spatial multiplexing systems with transmit antenna selection.

References

[1] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. : Cambridge University Press, 2005.

[2] A. F. Molisch and M. Z. Win, "Mimo systems with antenna selection," *IEEE Microw. Mag.*, vol. 5, no. 1, pp. 46–56, Mar. 2004.

[3] H. Zhang, H. Dai, Q. Zhou, and B. L. Hughes, "On the diversity order of spatial multiplexing systems with transmit antenna selection: A geometrical approach," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5297–5311, Dec. 2006.

[4] R. W. Heath, Jr., S. Sandhu, and A. Paulraj, "Antenna selection for spatial multiplexing systems with linear receivers," *IEEE Commun. Lett.*, vol. 5, no. 4, pp. 142–144, Apr. 2001.

[5] R. W. Heath Jr. and A. Paulraj, "Antenna selection for spatial multiplexing systems based on minimum error rate," in *Proc. IEEE Int. Conf. Commun., ICC*, 2001.

[6] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge University Press, 2004.

[7] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge University Press, 1985.
[8] Y. Jiang, X. Zheng, and J. Li, "Asymptotic performance analysis of V-BLAST," in *Proc. IEEE Global Telecommun. Conf., GLOBECOM*, 2005.
[9] H. Zhang, H. Dai, and B. L. Hughes, "On the diversity-multiplexing tradeoff for ordered SIC receivers over MIMO channels," in *Proc. IEEE Int. Conf. Commun., ICC*, 2006.
[10] Y. Jiang, M. K. Varanasi, X. Zheng, and J. Li, "Performance analysis of V-BLAST at high SNR regime," *IEEE Trans. Wireless Commun.*, submitted for publication.
[11] A. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Found. Trends in Commun. Inf. Theory*, vol. 1, no. 1, June 2004.
[12] E. Sengul, E. Akay, and E. Ayanoglu, "Diversity analysis of single and multiple beamforming," in *Proc. IEEE Veh. Technol. Conf.*, 2005, VTC, Spring.
[13] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

Iterative List Decoding of Some LDPC Codes

Jørn Justesen, *Member, IEEE*, Tom Høholdt, *Fellow, IEEE*, and Johann Hjaltason

***Abstract*—We present an iterative list decoding algorithm for low-density parity-check (LDPC) codes. In particular we apply this decoder to a class of LDPC codes from finite geometries and show that the $(73, 45, 10)$ projective geometry code can be maximum-likelihood (ML) decoded with low complexity. Moreover, the list decoding approach enables us to give a theoretical analysis of the performance. We also consider list bit-flipping (BF) decoding of longer LDPC Codes.**

***Index Terms*—Bit flipping (BF), finite-geometry codes, iterative decoding, list decoding.**

I. Introduction

In the last decade two old methods for decoding linear block codes have gained considerable interest, *iterative decoding* as first described by Gallager in [1] and *list decoding* as introduced by Elias [2]. In particular, iterative decoding of low-density parity-check (LDPC) codes, has been an important subject of research, see, e.g., [3] and the references therein. "Good" LDPC codes are often randomly generated by computer, but recently codes with an algebraic or geometric structure have also been considered, e.g., [3] and [4]. The performance of the iterative decoder is typically studied by simulations and a theoretical analysis is more difficult.

In this correspondence, we combine the two decoding methods and present an iterative list decoding algorithm. In particular, we apply this decoder to a class of LDPC codes from finite geometries and show that the $(73, 45, 10)$ projective geometry code can be maximum-likelihood (ML) decoded with low complexity. Moreover, the list decoding approach enables us to give a complete analysis of the performance in this case. We also discuss the performance of the list bit-flipping algorithm for longer LDPC codes.

Manuscript received November 20, 2005; revised January 9, 2007. the material in this correspondence was presented in part at the IEEE International Symposium on Information Theory, Seattle, WA, July 2006.

J. Justesen is with COM, The Technical University of Denmark, DK 2800 Lyngby, Denmark (e-mail: jju@com.dtu.dk).

T. Høholdt is with the Department of Mathematics, The Technical University of Denmark, DK 2800 Lyngby, Denmark (e-mail: T. Hoeholdt@mat.dtu.dk).

J. Hjaltason is with INOPI, DK 2630 Tästrup, Denmark (e-mail:jh@emergens.dk).

Communicated by T. J. Richardson, Associate Editor for Coding Theory.
Digital Object Identifier 10.1109/TIT.2007.907485

We consider hard-decision iterative decoding of a binary (n, k, d) code. For a received vector y, we calculate an extended syndrome $s = Hy'$, where H is a parity-check matrix, but usually has more than $n-k$ rows. Let r denote the length of the syndrome. The idea of using extended syndromes was also used in [5]. Our approach is based on one of the common versions of bit flipping (BF) [3], where the schedule is such that the syndrome is updated after each flip. In each step, we flip a symbol chosen among those positions that reduce the weight of the extended syndrome, which we refer to briefly as the syndrome weight u. A decoded word is reached when $u = 0$. In this correspondence, we consider a variation of the common algorithm in the form of a tree-structured search. Whenever there is a choice between several bits, all possibilities are tried in succession. The result of the decoding algorithm is, in general, a list of codewords, obtained as leaves of the search tree. This form of the BF algorithm leads naturally to a solution in the form of a list of codewords at the same smallest distance from y [6]. This list decoding concept is somewhat different from list decoding in the usual sense of all codewords within a certain distance from y.

The outline of the correspondence is as follows. In Section II, we consider regular LDPC codes and discuss the relations between the weight of an error pattern and the weight of the corresponding syndrome. The section also contains results on the performance of the standard BF algorithm. Section III contains the new iterative list decoding algorithm. In Section IV, we recall the LDPC codes from finite geometries and give some results on the number of minimum-weight words of these codes. In Section V, we give a complete analysis of the performance of the list BF algorithm used on the $(73, 45, 10)$ projective geometry code and show that the decoder is ML. Since so little is known on complete decoding (of for instance Bose–Chaudhuri–Hocquenghem (BCH) codes) we felt it appropriate to do this by using the introduced methods. Section VI contains results on the complexity of the decoder from Section V and in Section VII, we treat the $(63, 37, 9)$ Euclidean geometry code. In Section VIII, we discuss the performance of the list BF algorithm for longer regular LDPC codes.

II. Error Patterns and Syndromes

We consider a regular LDPC code given by a parity check matrix, possibly with more than $n-k$ rows, and the corresponding syndromes. All syndromes associated with single errors, i.e., the columns of the parity-check matrix, have the same low weight γ. For any pair of columns, there is at most one row where both have a one, so the girth of the Tanner graph for the code is at least 6, and any two rows have at most one 1 in common.

An error pattern of weight w has syndrome weight u where

$$w\gamma - w(w-1) \leq u \leq w\gamma. \tag{1}$$

It follows that the minimum distance of the code is lower-bounded by

$$d \geq \gamma + 1. \tag{2}$$

If we approximate the distribution of ones in a row by independent variables with probability $\frac{\gamma}{r}$ we get the average syndome weight to be

$$\bar{u} = r \sum_{i\,\mathrm{odd}} \binom{w}{i} \left(\frac{\gamma}{r}\right)^i \left(1 - \frac{\gamma}{r}\right)^{w-i}. \tag{3}$$