Reliability-Based Hybrid MMSE/Subspace-Max-Log-APP MIMO Detector

Andreas Ibing, David Kühling, and Holger Boche

Abstract—**We present an algorithm which is a flexible and scalable trade-off both in complexity and performance between unbiased MMSE and Max-Log-APP MIMO detection. In a first step, LLRs are generated based on linear detection (Max-Log per stream). A predefined number of LLRs is chosen as 'reliable' based on LLR magnitude. For the remaining 'unreliable' subspace we compute the joint Max-Log-APP solution. The algorithm contains unbiased MMSE detection as well as full Max-Log-APP detection as special cases and is suited for iterative detection-decoding.**

Index Terms—**MIMO, MMSE, Max-Log-APP.**

I. Introduction

OPTIMAL MIMO detection in the general formulation including iterative detection-decoding means computation of aposteriori bit probabilities (APP), given the received symbols and channel matrix and possibly apriori information about bit probabilities from the decoder. For practical implementation, optimal APP computation is considered too complex for the number of transmit antennas and modulation sizes used by modern standards.

There are several practical detector algorithms which approximate the APPs. They output log-likelihood-ratios (LLRs) of bit probabilities instead of the probabilities themselves, because multiplications can be conveniently computed as additions in the log domain, and the ratios avoid extra normalization effort. Computation is further reduced by using the Max-Log approximation [1]: $\ln \sum a_i \approx \max(\ln(a_i))$.

One standard approach is the separation of the transmit data streams at the receiver by use of unbiased LMMSE filtering. The separated data streams are then individually soft demodulated using the Max-Log approximation per stream. Complexity of this approach is low, but higher detection performance (LLR accuracy) is desirable.

Another approach is to jointly soft demodulate the data streams using the Max-Log approximation ('Max-Log-APP'). This means that each LLR is computed using the closest transmit vector candidate where the bit in question is positive and the closest vector vector candidate where it is negative (but not all possible vector candidates). In [2] it is shown analytically that Max-Log-APP MIMO demapping for QAM constellation converges with growing apriori information towards the performance of SIMO maximum ratio combining for a (shifted) BPSK constellation. For a small number of antennas and small modulation set the candidates can be found by 'brute-force' enumeration. The soft-output sphere decoder [3] finds the candidate vectors by depth-first tree search, which is enabled by QR-decomposition of the channel matrix (the order of node expansion is determined by Schnorr-Euchner enumeration). Unfortunately the Max-Log-APP detection problem is NP hard [4], resulting in non-constant and high complexity of Sphere decoding (depending on the channel matrix).

Manuscript received July 21, 2009. The associate editor coordinating the review of this letter and approving it for publication was B. Chen.

This work has been partially supported by BMBF under grant 01BU0652.

The authors are with the Fraunhofer Institute for Telecommunications, Heinrich-Hertz-Institut (e-mail: andreas.ibing@hhi.fraunhofer.de).

Digital Object Identifier 10.1109/LCOMM.2010.05.091523

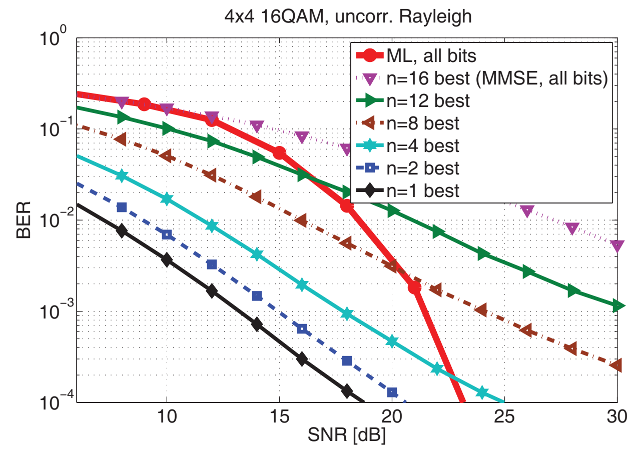

Fig. 1. Motivation: the partial bit error rate of linear MIMO detection is low for bit positions with relatively large LLR magnitude.

The soft-out M-algorithm (SOMA) [5] uses breadth first tree search to approximate the Max-Log-APP solution. In each tree level the so far best M candidates are expanded. The SOMA has a fixed complexity which makes it interesting for realtime implementation. By choice of M the performance can be scaled up to that of the Max-Log-APP solution. Major drawback of SOMA is the complexity of the sorting step in each tree level.

Successive interference cancellation (SIC) is the approach to successively demap one symbol (or one individually encoded symbol stream of a transmit antenna). Compared to the joint iterative demapping approach, the SIC approach suffers more interference when detecting the first data stream.

We propose an algorithm which has constant complexity (different from sphere decoder) and does not need QR-decomposition and sorting steps like SOMA. It can be applied for iterative APP computation in stochastic inference frameworks like Bayesian inference (Belief propagation). Performance (and complexity) is scalable by parameter choice between that of unbiased MMSE and Max-Log-APP.

II. Algorithm Motivation

Unbiased MMSE detection gives reliability information about bits (in form of LLRs) at low complexity. Fig. 1 illustrates the uncoded error rates in 4x4 16QAM transmission (assuming uncorrelated Rayleigh fading) for the subset of the

n LLRs with largest magnitude. For the SNR range of interest, the error rate e.g. of the 4 LLRs with largest magnitude (per MIMO vector) is far lower than that of maximum likelihood (ML) detection of all bits. The idea is to reduce the candidate search space for the Max-Log-APP solution to a fixed predefined size, based on reliability: we assume the n bit positions with largest LLR magnitude as correct, and the remaining subspace of size $2^{N_T k-n}$ is 'most likely' to contain the ML solution and Max-Log-APP counter-hypotheses (N_T transmit antennas and k bit per transmit antenna, gives a sum of $N_T k$ LLRs to generate). For this remaining (smaller) 'unreliable' subspace we can perform Max-Log-APP postprocessing.

III. Algorithm Description

The detection algorithm consists of three steps. The first step is to perform unbiased MMSE detection. MIMO transmission over the channel matrix $\mathbf{H}$ is denoted as

$$\mathbf{y} = \mathbf{H} \cdot \mathbf{x}(\mathbf{c}) + \mathbf{n}, \tag{1}$$

where $\mathbf{c}$ is the bit vector which is mapped to the symbol vector $\mathbf{x}$. With noise variance σ^2, the MMSE receiver filter matrix for separation of the streams is given by:

$$G_{\text{bias}} = (H^H H + \sigma^2 I)^{-1} H^H$$

The unbiased version is $G_{\text{unb}} = S G_{\text{bias}}$, where S is the diagonal matrix which removes the bias introduced by the MMSE criterion [6]:

$$S = \text{diag}\Big(\frac{1}{(G_{\text{bias}} H)_{1,1}}, \ \ldots, \ \frac{1}{(G_{\text{bias}} H)_{N_T,N_T}}\Big)$$

The equalized symbol vector is:

$$\hat{\mathbf{x}} = G_{\text{unb}} \mathbf{y}$$

The receiver then computes an LLR for each transmit bit c:

$$L(c) = \ln \frac{P(c=+1)}{P(c=-1)} \tag{2}$$

The LLR for transmit antenna i and bit position j is (under assumption of Gaussian noise and Max-Log approximation per stream):

$$L(c_{i,j}) = -\frac{1}{2\sigma_{eq}^2}\Big(\min_{x_i \in \mathcal{X}^1_{j=1}} |\hat{x}_i - x_i|^2 - \min_{x_i \in \mathcal{X}^1_{j=-1}} |\hat{x}_i - x_i|^2 \Big)$$

where ${\sigma_{eq}}^2$ is the noise variance on the stream after filtering. $x_i \in \mathcal{X}^1_{j=1}$ means the set of symbols where the bit whose LLR is to be computed has the value 1, $\mathcal{X}^1_{j=-1}$ is the complement set.

In the second step the n LLRs with largest magnitude $|L(c_{i,j})|$ are selected as 'reliable', where n is a predefined constant. The bit values are the signs of the LLRs. The remaining $N_T k - n$ bit positions are the 'unreliable' subspace.

In the third and last step the algorithm computes the joint Max-Log-APP solution for the 'unreliable' candidate subspace. Since this subspace is relatively small, we perform binary enumeration of the candidates. For each bit position, the subspace always contains hypothesis and counter-hypothesis vectors, so that clipping operations like in the list sphere decoder [7] are not necessary. The extrinsic (Max-Log) LLR for bit position i is [8]:

$$L_e(c_i) = \max_{\mathbf{x} \in \mathcal{X}^{N_T}_{i=1}} \{-\frac{1}{2\sigma^2}||\mathbf{y} - \mathbf{H}\mathbf{x}(\mathbf{c})||^2 + \frac{1}{2} L_a(\mathbf{c})^T \mathbf{c}\}$$
$$- \max_{\mathbf{x} \in \mathcal{X}^{N_T}_{i=-1}} \{-\frac{1}{2\sigma^2}||\mathbf{y} - \mathbf{H}\mathbf{x}(\mathbf{c})||^2 + \frac{1}{2} L_a(\mathbf{c})^T \mathbf{c}\} - L_a(c_i)$$

The notation $\mathbf{x} \in \mathcal{X}^{N_T}_{i=1}$ means the set of all (remaining) symbol vectors of dimension N_T with the bit value 1 in position i of the corresponding bit vector. The computation uses apriori information (from the decoder) in the form of the apriori LLRs $L_a(\mathbf{c})$. The LLR vector which the detector outputs consists of n values generated by linear detection and $N_T k - n$ values generated by subspace-Max-Log-APP computation. For iterative detection-decoding only step 3 needs to be repeated after running the decoder. A pseudo-code formulation of the algorithm is given in Alg. 1.

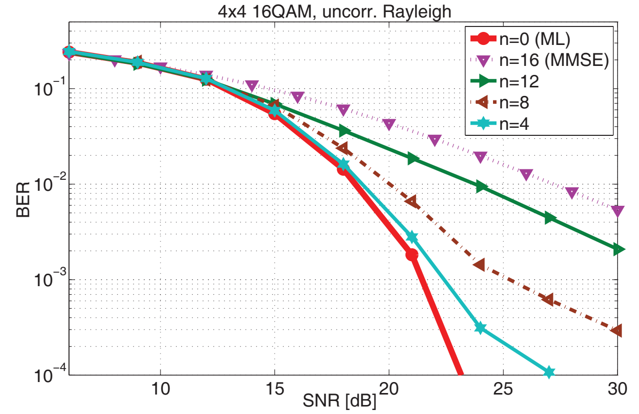

Fig. 2. Error rate of proposed algorithm for different number of linearly detected bit positions.

Algorithm 1 Subspace-Max-Log by enumeration

$\hat{\mathbf{x}} = G_{unb}\mathbf{y}$ {MMSE stream separation}
for all tx-antennas **do**
  for all bit-positions **do** {bits of this antenna}
    $L_e(c_j) = (\min_{x_i \in \mathcal{X}^1_+} \Delta - \min_{x_i \in \mathcal{X}^1_-} \Delta)$ {max-log per stream}
select n positions of largest $|L(c_i)|$ {'reliable'}
for all $\mathbf{c}^{(sub)} \in 2^{N_T k - n}$ **do** {'unreliable' subspace}
  $metric(\mathbf{x}(\mathbf{c})) = \frac{-1}{2\sigma^2}||\mathbf{y} - \mathbf{H}\mathbf{x}||^2 + \frac{1}{2} L_a(\mathbf{c})^T \mathbf{c}$
for $i = 1$ to $N_T k - n$ **do** {'unreliable' bit positions}
  $max\_p = \max_{\mathbf{c}^{(sub)}|c_j=+1} (metric(\mathbf{x}(\mathbf{c})))$
  $max\_m = \max_{\mathbf{c}^{(sub)}|c_j=-1} (metric(\mathbf{x}(\mathbf{c})))$
  $L_e(c_j) = max\_p - max\_m - L_a(c_j)$

IV. Performance

This section illustrates that the performance in terms of error rates and mutual information varies between that of unbiased MMSE and Max-Log-APP. Both are included in the algorithm as special cases for $n = N_T k$ and $n = 0$ respectively. Simulation assumes 16QAM transmission over 4x4 uncorrelated Rayleigh fading, and perfect channel estimation at the receiver.

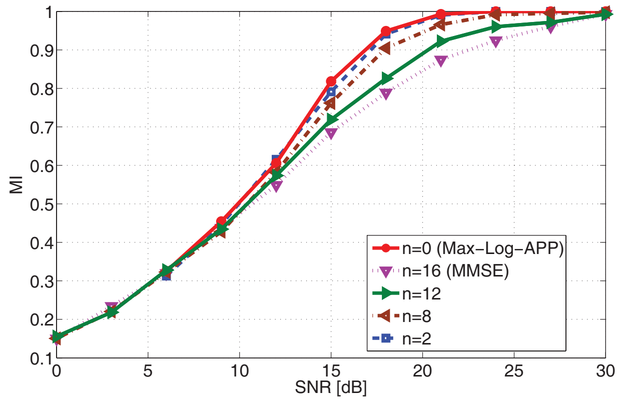

Fig. 3. Mutual information for different number of linearly detected bit positions.

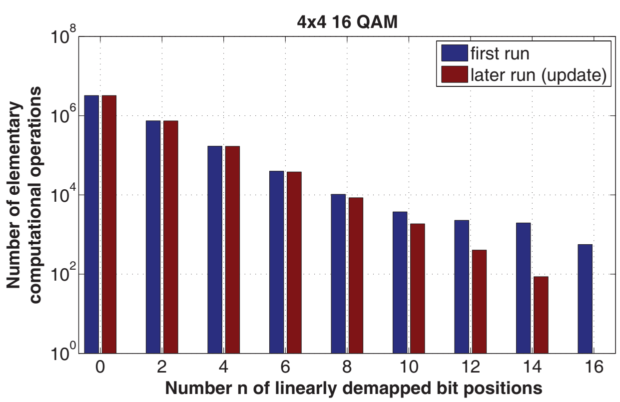

Fig. 4. Complexity scales exponentially with parameter n.

Uncoded error rates are shown for different n in Fig. 2. The detector curve for $n = 0$ is denoted ML, since Max-Log-APP without apriori information and with hard output (bits instead of LLRs) reduces to searching the most likely candidate bit vector. Postprocessing for a small subspace of four bit positions (16 candidate vectors) already results in more than 1dB improvement for the practically interesting uncoded BER range around 10^{-1}.

To assess the performance of soft-output detectors, mutual information (MI) is a more suitable measure than uncoded BER (which only evaluates the signs of LLRs and not their magnitude). MI in dependence on SNR is shown in Fig. 3. By this measure, Max-Log processing only brings gains for SNR larger than 10dB in this scenario.

For enhanced detector performance also at low SNR, iterative detection-decoding can be applied. Detector performance is then measured by an EXIT chart, given in Fig. 5. While the linear detector does not benefit from apriori information, the proposed algorithm increasingly exploits it with decreasing parameter n.

V. Complexity

The number of elementary real-valued computational operations for different n is illustrated in Fig. 4. Operations like *Multiply-Accumulate* and *Compare-Select* are counted as the same unit. For hardware independence we assume the possibility of reuse of intermediate results ('infinite' memory) and cost-free *Load/Store* operations. The figure shows that

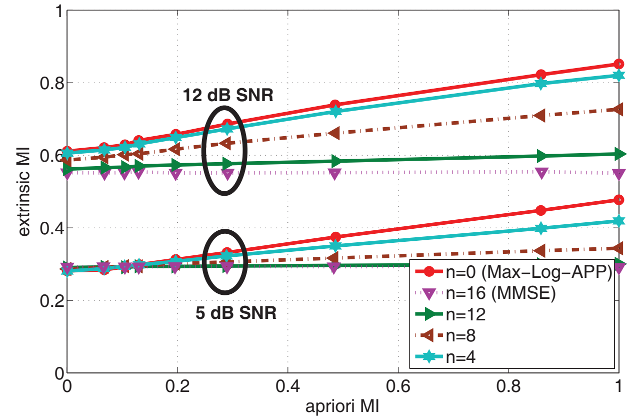

Fig. 5. EXIT chart: usage of apriori knowledge from the channel decoder.

complexity scales exponentially with n (apart from MMSE preprocessing), which is due to the problem being NP hard.

VI. Discussion

The presented algorithm reduces the search space for Max-Log-APP detection based on the magnitude of MMSE generated LLRs (reliability information). For iterative detection-decoding we used a constant bit position selection. A modification of the algorithm is also possible: the selection of bit positions in step 2 of the algorithm can be based on the sum of MMSE LLRs and apriori LLRs from the decoder (this sum is the current APP estimate). For the remaining 'unreliable' subspace we computed the Max-Log-APP solution by candidate enumeration, based on the assumption that the subspace is small. For a larger subspace (smaller parameter n) the sphere decoder could also be applied (tradeoff between vectorized implementation on the one hand and sorting effort for reducing the number of computed metrics on the other). The parameter n can be chosen according to an overall receiver optimization including channel estimation and decoding: maximization of mutual information for minimal or predefined channel SNR and complexity.

References

[1] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and suboptimal MAP decoding algorithms operating in the log domain," in *IEEE ICC*, 1995.

[2] A. Ibing, D. Kühling, and H. Boche, "On the relation of MIMO APP detection and SIMO maximum ratio combining," in *IEEE Globecom*, 2009, in press.

[3] C. Studer, A. Burg, and H. Bölcskei, "Soft-output sphere decoding: algorithms and VLSI implementation," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 2, pp. 290–300, Feb. 2008.

[4] S. Arora, L. Babai, J. Stern, and Z. Sweedyk, "The hardness of approximate optima in lattices, codes, and systems of linear equations," *J. Computer and Syst. Sciences*, vol. 54, pp. 317–331, 1997.

[5] K. Wong, "The soft-output M-algorithm and its applications," Ph.D. dissertation, Queen's University, Ontario, Canada, Aug. 2006.

[6] E. Zimmermann, "Complexity aspects in near-capacity MIMO detection-decoding," Ph.D. dissertation, TU Dresden, 2007.

[7] B. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, Mar. 2003.

[8] J. Hagenauer, "The turbo principle in mobile communications," in *IEEE ISITA*, 2002.