I. INTRODUCTION

The seminal works in [1] and [2] on multiple antenna elements at the transmitter and receiver show a huge increase in the throughput of this point-to-point channel referred to also as multiple input multiple output (MIMO) system. These promising results shifted the focus of research on multi antenna communications and motivated the introduction of multiple antenna elements in the future communication systems as IEEE 802.11n [3], IEEE 802.16m [4] and 3GPP LTE [5]. Enhanced capacity (degrees of freedom) and improved reliability (diversity) are the two significant gains of MIMO communications which can not be realized simultaneously and need to be traded off [6].

Researchers persist to strive for finding space time codes (STC) with reduced decoding complexity. These codes take into account both the spatial and temporal dimensions of the MIMO channel. Orthogonal Space-Time Block Codes (OSTBCs) [7] are widely used because they are easy to encode and decode. For the case of two transmit antennas, the OSTBC is known as Alamouti code [8]. OSTBCs are repetition codes that only provide diversity gain. In order to approach the capacity limit they have to be used in concatenation with an outer code. Remarkable coding gains can be obtained if a capacity achieving temporal encoder, such as turbo or Low-Density Parity Check (LDPC) code is used in concatenation with a STC [9]. Recently it has been shown for the ergodic channels that the complex concatenation of the STC and the outer codes can be replaced with temporally coded and spatially multiplexed streams (coded spatial streams) for nearing capacity [10]. Each spatial stream can also be independently coded using temporal encoders as convolutional, turbo or LDPC codes where at the receiver, standard off-the-shelf decoders are used after the demodulator.

For coded spatial streams (also for the STC), the well-known data model after appropriate filtering and sampling is \( y = Hx + z \) (to be made precise in Section II) where \( y \) is the received data, \( H \) is the channel matrix, \( x \) is the symbol vector with the elements from a \( \mathcal{M} \)-ary finite constellation and \( z \) is the noise. The problem is then to detect \( x \) from \( y \). Essentially the same problem occurs in multiuser detection for CDMA [11] and for single-carrier transmission over channels that induce intersymbol interference. In these cases, the matrix \( H \) usually has a specific structure. We are interested in the fundamental aspects of the model \( y = Hx + z \) and we shall refer to it in the general terms as a MIMO model.

The problem of detecting \( x \) from \( y \) has stimulated a large body of research [11] and references therein. One can easily show that if the noise \( z \) is Gaussian then obtaining the maximum-likelihood (ML) solution is equivalent to minimizing the Euclidean distance \( \| y - Hx \|_2^2 \) with respect to \( x \) over the finite set spanned by all possible combinations of constellation points that can constitute the vector \( x \). For maximum likelihood (ML) soft MIMO detection, the demodulator calculates log-likelihood ratios (LLRs) for all the bits that constitute \( x \) by summing the Euclidean distances for the values of \( x \) for which that particular bit is one and zero thereby amounting to \( 2^{\log(M) \times n_t} \) terms where \( n_t \) is the number of transmit antennas [12]. In many cases of practical interest, one resorts to the approximation of replacing the sums with the largest term which is equivalent to minimizing the Euclidean distance and is termed as max log MAP approach. Unfortunately this problem is NP-hard for general \( H \) and \( y \) [13] which implies that there are no known efficient (i.e. polynomial-time) solutions. Many sophisticated methods as lattice reduction and sphere decoding [14] exist which find the ML solution with high probability, but these methods are
in general still computationally complex. This is true also in an average sense if $H$ is random (i.e. for a fading channel). The popular sphere decoding method is much more efficient than a brute-force search, but it still admits an average complexity that is exponential in the dimension of $x$.

Naive solutions, like neglecting the integer constraint coupled with the Gaussian assumption for the alphabets and then subsequently projecting the so-obtained solution onto the finite set of permissible $x$ [linear receivers as MMSE and zero forcing ZF], in general works poorly especially at lower SNRs. Ignoring the spatial color at the output of linear receivers results in the decoupling of spatial streams which fundamentally reduces the complexity of detection. These disregards proliferate the suboptimality of linear receivers which exhibit degraded performance especially at lower SNRs. Different spatial streams in a MIMO fading channel experience different SNRs as each stream sees a different realization of the channel and the error performance of the system is dominated by the spatial streams with the relatively lower SNRs. One can do somewhat better by using decision-feedback-equalization detection (nulling-and-canceling), whereby the elements of $x$ are detected one by one, and in an order that can be optimally chosen [15].

We propose a detection algorithm which exploits the low complexity of linear detectors and the optimal performance of max log MAP detectors. In the proposed algorithm, the spatial streams which have experienced higher SNRs are detected using linear detectors as ZF or MMSE and are subsequently stripped off while the spatial streams which have experienced lower SNRs are detected using max log MAP algorithm. The algorithm therefore successfully escapes the exponential complexity of MIMO detection and the degraded performance of linear receivers. This leads to a performance (error rate) better than the standard linear detector based approaches. Our proposed algorithm can also be extended to the systems with the spatial streams of equal-rate and non-uniform power distribution in view of successive stripping [17].

Keeping in view the existing and forthcoming wireless standards as IEEE 802.11n, IEEE 802.16m and 3GPP LTE, we consider in this paper bit interleaved coded modulation (BICM) MIMO OFDM system. The rationales for the selection of this system for next-generation wireless networks are the enhanced spectral efficiency of MIMO, reduced complexity of equalization of OFDM and improved code diversity of BICM [18] [19] for fading channels. Regarding notations, we will use lowercase or uppercase letters for scalars, lowercase boldface letters for vectors and uppercase boldface letters for matrices. $\cdot$ and $||\cdot||$ indicate norm of scalar and vector while $(\cdot)^T$ and $(\cdot)^\dagger$ indicate transpose and conjugate transpose respectively. The paper is organized as follows. In section II we provide the system model while section III gives a review of the existing approaches. In section IV we propose low complexity detection algorithm and section V describes simulation results which is followed by the conclusions.

**II. System Model**

We consider a MIMO system which is a $n_t \times n_r$ ($n_r \geq n_t$) BICM MIMO OFDM system with $n_t$ spatial streams. The detection is based on segregating the streams into two groups based on their post detection SNR. It is followed by decoding of the first group of streams which is subsequently stripped off leading to the detection of the second group of streams. Block diagrams of the transmitter and receiver are shown in the figures 1 and 2 respectively. We consider slow fading and assume channel to remain constant for the transmission of a frame. Well known baseband model of the system at the $n$-th frequency tone for one channel use is given as:-

$$y_n = h_{1,n}x_1 + h_{2,n}x_2 + \cdots + h_{n_t,n}x_{n_t} + z_n, \quad n = 1, 2, \cdots, N$$

where $N$ is the total number of frequency tones and $n_t$ is the total number of spatial streams/transmit antennas. The channel would remain constant during the transmission of a frame while the symbols would change for every channel use. We can conveniently drop the frequency index and can rewrite the
where $y, z \in \mathbb{C}^{n_r}$ are the vectors of the received symbols and circularly symmetric complex white Gaussian noise of variance $N_0$ at the $n_r$ receive antennas. $h_k \in \mathbb{C}^{n_r}$ is the vector characterizing flat fading channel response from the $k$-th transmitting antenna to $n_r$ receive antennas and $x_k$ is the complex symbol of the $k$-th stream transmitted by the $k$-th transmit antenna with $E[|x_k|^2] = \sigma_k^2$. It is assumed that each channel path between the transmitter and the receiver is independent. The complex symbols $x_1, \ldots, x_{n_t}$ of $n_t$ streams are also assumed independent.

$$H = [h_1, h_2, \ldots, h_{n_t}]$$ is the channel matrix with complex channel gains $E[|h_{ij}|^2] = 1$. Under the power constraint $P_T$, the average SNR at each receiver branch is $\frac{P_T}{N_0}$.

III. REVIEW OF EXISTING SCHEMES

We now review some existing schemes for MIMO detection.

1) MMSE: Detection based on MMSE equalization [16] involves linear MMSE preprocessing i.e. applying a spatial filter $M$ to the received signal vector $y$ i.e. $\hat{x} = M y$ where $\hat{x}$ is the biased estimate of $x = [x_1, x_2, \ldots, x_{n_t}]^T$. $M$ is given as

$$M = (N_0 P^{-1} + H^H H)^{-1} H^T$$

where $P$ is the diagonal power distribution matrix with the diagonal as $[\sigma_1^2, \sigma_2^2, \ldots, \sigma_{n_t}^2]$. It is followed by an unbiasing operation i.e. $\hat{x} = \Gamma^{-1} \hat{x}$ where $\Gamma = \text{diag}(M H)$. Based on the Gaussian assumption of the post detection interference, MMSE preprocessing decouples the spatial streams and the bit metric for the $i$-th bit for bit value $b$ of the symbol $x_k$ on the $k$-th stream is given as

$$\lambda^i_k(y, b) \approx \max_{x_k \in \mathcal{X}_k} \left[ -\frac{\gamma^2}{N_0} |\hat{x}_k - x_k|^2 \right]$$

for $k = 1, 2, \ldots, n_t$ where $\chi_k$ is the $k$-th diagonal element of $\Gamma$. $\chi_{k,b}^i$ denotes the subset of the signal set $x_k \in \chi_k$ whose labels have the value $b \in \{0, 1\}$ in the position $i$.

2) Max Log MAP: In an $n_t \times n_r$ system, the ML bit metric for bit $b$ at the $i$-th location of the $k$-th stream $x_k$ is given as

$$\lambda^i_k(y, b) = \log \sum_{x_i \in \chi_i} \ldots \sum_{x_{n_t} \in \chi_{n_t}} \exp \left[ -\frac{1}{N_0} ||y - Hx||^2 \right]$$

Applying log sum approximation [18] we have

$$\lambda^i_k(y, b) \approx \max_{x_i \in \chi_i \ldots x_{n_t} \in \chi_{n_t}} \left[ -\frac{1}{N_0} ||y - Hx||^2 \right]$$

Finding the maximum term in this sum is equivalent to minimizing $||y - Hx||^2$ subject to the constellation constraint on $x$, and this is an NP-hard optimization problem. Therefore the max-log approximation, while it is conceptually simple, still suffers from being computationally very intensive. Minimizing $||y - Hx||^2$ by hard decision sphere decoding [14] is feasible, but incurs exponential average complexity. This limits the range of problem sizes which may be addressed. Moreover, the complexity of this approach is random (i.e. it depends on the realizations of $H$ and $x$) which may introduce random decoding delays. Additionally sphere decoding is based on searching through a tree in a sequential manner and therefore is difficult to execute efficiently on parallel hardware architectures.

IV. PROPOSED DETECTION ALGORITHM

Proposed algorithm exploits the low complexity of linear detectors and the improved performance of max log MAP detectors. The fact that the linear receivers as MMSE have significantly lower complexity as compared to the brute-force search but exhibit degraded performance particularly at lower SNRs is exploited by detecting the spatial streams which have seen good realizations of the channel and therefore enjoy relatively higher SNR using these linear filters. These detected streams are subsequently stripped off leading to a group of yet undetected streams which have seen comparatively poor realizations of the channel and therefore have relatively lower SNR. This group of streams is then detected using log MAP detectors. So the proposed detection algorithm can be divided into 3 stages.

A. Optimal Detection Ordering

The system performance is affected by the order in which the spatial streams are detected. We consider the V-BLAST criteria by simply ordering the streams as per the decreasing post detection SNR. This approach is termed as the “best first” cancellation approach within the multi-user community. The spatial streams to be detected by MMSE filter are decided by choosing the rows of $M$ with the smallest norms. Depending on the system parameters, a threshold for the row norm of the MMSE filter $M$ can be decided as a criteria for detection using MMSE filters.

B. Detection based on linear MMSE filters

Once the spatial stream to be detected by MMSE filter are decided, these streams are detected using MMSE filters. The MMSE filter for detecting $k$-th stream which is the $k$-th row of $M$ is given as

$$m_k = (h_k^H R_k^{-1} h_k + 1/\sigma_k^2)^{-1} h_k^H R_k^{-1}$$

where

$$R_k = \sigma_1^2 h_1 h_1^H + \cdots + \sigma_k^2 h_k h_k^H + \cdots + \sigma_{n_t}^2 h_{n_t} h_{n_t}^H$$

Application of MMSE filter on the received vector $y$ yields

$$m_k y = y_k = \alpha_k x_k + z_k$$

So the bit metric for bit $b$ at $i$-th location of $k$-th spatial stream is given as

$$\lambda^i_k(y, b) \approx \max_{x_k \in \chi_k} \left[ -\frac{1}{N_k} |y_k - \alpha_k x_k|^2 \right]$$
where \( N_k = m_k^T R_k m_k \). Bit LLRs are calculated using these bit metrics which are then deinterleaved and are subsequently fed to the decoders for the detection of \( k \)-th stream. The detected stream is then again encoded, interleaved, modulated and convolved with the channel and is consequently subtracted from \( y \).

Without loss of generality, let the last \( K \) spatial streams \((n_t - K + 1, \cdots, n_t)\) are detected using MMSE filters which are then subsequently stripped off.

### C. Max Log MAP

For the remaining \( n_t - K \) spatial streams which have experienced poor channel realizations, max log MAP detector is used. The new system equation after nulling the last \( K \) detected streams is

\[
y' = H' \chi + z' \quad (7)
\]

where \( y', z' \in \mathbb{C}^{n_t} \) and \( H' = n_t \times (n_t - K) \) complex matrix and \( \chi' \in \mathbb{C}^{n_t-K} \). The bit metric for bit \( b \) at the \( i \)-th location of the \( k \)-th stream \( x_k \) is given as

\[
\lambda_k (y, b) \approx \max_{x_i \in \chi_1^i, \cdots, x_K \in \chi_K^i, x_{n_t-K} = k} \left[ -\frac{1}{N_0} \left\| y' - H' \chi \right\|^2 \right]
\]

The computational complexity of this detection process is \( O(|\chi|^{n_t-K}) \). The closer \( K \) is to \( n_t \), lower is the complexity of detection. Bit LLRs are calculated using these bit metrics and consequently the streams are detected using the corresponding channel decoders.

### V. Simulation Results

We now assess the performance of the proposed approach of soft detection of spatially multiplexed MIMO system by means of simulation of the frame-error rates (FER). We focus on target error-rates of the order of \( 10^{-2} \). We consider \( 3 \times 3 \) and \( 4 \times 4 \) BICM MIMO OFDM systems using the de facto standard, 64 state (133, 171) rate-1/2 convolutional encoder of 802.11n standard and the punctured rate 1/2 turbo code\(^1\) of 3GPP LTE [5]. The upcoming WLAN standard 802.11n [3] supports the codeword sizes of 648, 1296, and 1944 bits. For our purposes, we selected the codeword size of 1296 bits. We consider slow fading and the MIMO channel remains constant for the duration of one frame. The channel has iid Gaussian matrix entries with unit variance. The channel is independently generated for each frame and perfect CSI at the receiver is assumed. Furthermore, all mappings of coded symbols use Gray encoding. We consider the max log MAP detector, MMSE based detector and the proposed detector. Spatial streams of equal rates and equal power are transmitted in a \( 3 \times 3 \) and \( 4 \times 4 \) system. For the proposed detector, one stream is detected by MMSE in \( 3 \times 3 \) system while remaining two streams are detected by max log MAP detector. In \( 4 \times 4 \) system, two streams are detected by MMSE detector while the remaining two are detected using max log MAP detector. Figures 3, 4, 5 and 6 indicate improved performance of the proposed approach with respect to that of MMSE based approach in \( 3 \times 3 \) and \( 4 \times 4 \) system. We observe not only better performance but also an increase in slope of FER curves which indicates improved diversity of the proposed approach as compared to that of MMSE. It is well known that MMSE loses one order of diversity while max log MAP approach has full diversity [20]. The proposed approach being the fusion of MMSE and the max log MAP approach therefore has improved diversity than that of MMSE detector while diversity order is less than that of max log MAP detector.

### VI. Conclusions

We have presented a novel detection scheme for BICM MIMO OFDM system. Our proposed detector is based on a combination of linear and non linear detectors and has

\(^{1}\)The LTE turbo decoder design was performed using the coded modulation library www.iterativesolutions.com
reduced computational complexity as compared to the brute force search and better performance than the standard linear equalizer based solutions as MMSE filter. The proposed detector exploits the low complexity of linear detectors and better performance of max log MAP detectors. We have used convolutional codes because of its widespread application in existing MIMO systems like IEEE 802.11n [3].

ACKNOWLEDGMENTS

Eurecom’s research is partially supported by its industrial partners: BMW, Bouygues Telecom, Cisco Systems, France Télécom, Hitachi Europe, SFR, Sharp, ST Microelectronics, Swisscom, Thales. The research work leading to this paper has also been partially supported by the European Commission under the IST FP7 research network of excellence NEWCOM++.  

REFERENCES