

EFFICIENT NEAR OPTIMAL JOINT MODULATION CLASSIFICATION AND DETECTION FOR MU-MIMO SYSTEMS

Hadi Sardeddeen* Mohammad M. Mansour* Louay M.A. Jalloul† Ali Chehab*

* Department of Electrical and Computer Engineering
American University of Beirut, Beirut, Lebanon
{has63, mmansour, chehab}@aub.edu.lb
† San Jose, CA 95110, USA, jalloul@ieee.org

ABSTRACT

Optimum data detection schemes for dual layer multi-user multiple-input multiple-output (MU-MIMO) systems are studied. A joint maximum likelihood (ML) modulation classification (MC) of the co-scheduled user and data detection receiver is developed. By expanding the max-log-maximum-a-posteriori MC approach to include distances of counter ML hypothesis symbols, the decision metric for MC is shown to be an accumulation over a set of tones of Euclidean distance computations also used by the ML detector for bit log-likelihood ratio soft decision generation. With a small complexity overhead, the proposed approach achieves near-optimal performance. An efficient hardware architecture is presented for the proposed approach.

Index Terms— MU-MIMO, Modulation Classification, Optimal Detection, Hardware Architecture

1. INTRODUCTION

Multiple-input multiple-output (MIMO) technology increases spectral efficiency by employing multiple antennas at the transmitter and the receiver [1]. And when used to allow simultaneous transmissions to multiple users over the same time and frequency resource elements, it forms an extended version of space-division multiple access called multi-user MIMO (MU-MIMO) [2, 3].

Several receiver designs exist for MU-MIMO. The interference rejection combining (IRC) and minimum mean-squared error (MMSE) receivers are linear processing schemes that only use the channel estimate of the co-scheduled user, without requiring knowledge of its modulation type [4]. The sub-optimal interference-aware (sub-IA) receiver does not estimate the interfering constellation; it assumes it to be fixed, and then applies maximum likelihood (ML) detection [5, 6]. An enhanced version of the sub-IA receiver adds an interference modulation classification (MC) routine, and feeds the knowledge about the interferer to an optimal ML interference-aware (IA) detector [7–9].

Modulation classification techniques can be classified into two categories: feature-based classification that depends on statistical properties, and ML classification that is based on likelihoods [10]. In this study, we consider the latter optimal approach [11]. The two main likelihood-based MC approaches [12] are the average likelihood ratio test (ALRT) and the generalized likelihood ratio test (GLRT). While ALRT treats the signal and channel parameters as unknown random variables with known distributions, GLRT treats them as deterministic but unknown. These approaches were extended to multiuser scenarios [13] and MIMO scenarios [14, 15].

We consider optimal detection methods for 2×2 MU-MIMO systems, which treat the interfering signal as a constrained unknown to be estimated. The receiver employs MC and ML detection, where the modulation type of the interferer is estimated before detection. We first study the optimal likelihood-based MC approach, and then introduce the max-log-maximum-a-posteriori (Max-Log-MAP) classifier and propose two extensions to it. The extensions consist of adding special distance metrics to the MC likelihood function. Following these extensions, we show that an optimized near-optimal MU-MIMO detector can be *efficiently* implemented with a slight modification to the soft-output ML MIMO detector.

Regarding notations, bold upper case, bold lower case, and lower case letters correspond to matrices, vectors, and scalars, respectively. Scalar norms, vector norms, conjugate transpose, and matrix inverse are represented by $|\cdot|$, $\|\cdot\|$, $(\cdot)^*$ and $(\cdot)^{-1}$, respectively. \mathbf{I}_n indicates an identity matrix of size n , $\mathbb{E}[\cdot]$ denotes the expected value, and $P(\cdot)$ denotes the probability density function.

2. SYSTEM MODEL

We consider a system model where the base station transmits to two users simultaneously on the same time-frequency resource ($N_t = 2$ transmit antennas and $N_r = 2$ receive antennas). The received signal at the user of interest over which another user is scheduled is expressed as $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$. Where $\mathbf{H} \triangleq [\mathbf{h}_1 \ \mathbf{h}_2]$ is the 2×2 channel matrix, $\mathbf{x} \triangleq [x_1 \ x_2]^T$ denotes the transmitted quadrature amplitude modulation (QAM) symbols, and \mathbf{n} is the complex additive white Gaussian noise vector with zero mean and variance σ_n^2 ($\mathbb{E}[\mathbf{n}\mathbf{n}^*] = \sigma_n^2 \mathbf{I}_2$). We denote by $\mathbf{x}_b = [b_i]_{i=1}^K$ the bit vector associated with the symbol x_1 of the user of interest, where $b_i \in \{0, 1\}$ denotes the i th bit of \mathbf{x}_b , and $K = \log_2(|\bar{\Lambda}|)$ ($\bar{\Lambda}$ is the constellation of the user of interest) is the number of bits per symbol.

The transmit power per antenna is normalized to unity, i.e., $\mathbb{E}[x_1 \cdot x_1^*] = \mathbb{E}[x_2 \cdot x_2^*] = 1$, and the noise variance is thus defined in terms of the number of transmit antennas and signal to noise ratio (SNR) as $\sigma_n^2 = N_t/\text{SNR}$. The received signal can also be written as:

$$\mathbf{y} = \mathbf{h}_1 x_1 + \mathbf{h}_2 x_2 + \mathbf{n} \quad (1)$$

where x_1 is drawn from the arbitrary, but known, constellation $\bar{\Lambda}$, that could be QPSK, 16-QAM or 64-QAM. However, x_2 is drawn from an unknown constellation Λ_j , $j \in \{0, 1, 2, 3\}$, where Λ_0 , Λ_1 , Λ_2 and Λ_3 correspond to the constellations \emptyset , QPSK, 16-QAM and 64-QAM, respectively, with \emptyset representing a constellation having one entry of zero power, corresponding to the case when there is no interferer.

3. INTERFERENCE REJECTION COMBINING (IRC)

Linear IRC detection is employed when estimates of both the desired and co-scheduled users' channels are available at the receiver, but knowledge of the modulation type of the co-scheduled user is *not*. IRC works as a linear MMSE receiver [4], performing whitening followed by maximum ratio combining:

$$\mathbf{h}_1^* \mathbf{R}^{-1} \mathbf{y} = \mathbf{h}_1^* \mathbf{R}^{-1} \mathbf{h}_1 x_1 + \mathbf{h}_1^* \mathbf{R}^{-1} (\mathbf{h}_2 x_2 + \mathbf{n}) \quad (2)$$

with $\mathbf{R} = \mathbf{h}_2 \mathbf{h}_2^* + \sigma_n^2 \mathbf{I}_n$ being the covariance matrix of the sum of interference and noise components. The resultant distance metric to be used in log likelihood ratio (LLR) computation is generated as:

$$\varphi_{\text{IRC}}(x_1) = \frac{1}{\sigma_{n,\text{IRC}}^2} |\mathbf{h}_1^* \mathbf{R}^{-1} \mathbf{y} - \mathbf{h}_1^* \mathbf{R}^{-1} \mathbf{h}_1 x_1|^2 \quad (3)$$

where unlike in [8], we have accounted for the variability of the variance from tone to tone by the scaling factor $\frac{1}{\sigma_{n,\text{IRC}}^2}$, with $\sigma_{n,\text{IRC}}^2 = \mathbf{h}_1^* \mathbf{R}^{-1} \mathbf{h}_1$. The LLR of the i th bit of \mathbf{x}_b can then be calculated using the Max-Log-MAP approximation as:

$$\lambda_{\text{IRC},i} = \min_{x_1 \in \bar{\Lambda}_{i,1}} \varphi_{\text{IRC}}(x_1) - \min_{x_1 \in \bar{\Lambda}_{i,0}} \varphi_{\text{IRC}}(x_1) \quad (4)$$

where $\bar{\Lambda}_{i,1}$ and $\bar{\Lambda}_{i,0}$ correspond to points in $\bar{\Lambda}$ having in the bit position i of \mathbf{x}_b a bit value of 1 and 0 respectively. Note that since the interference is discrete and not Gaussian, IRC is not the optimal detection strategy in MU-MIMO.

4. ML MC FOR 2×2 MU-MIMO SYSTEMS

The optimal likelihood-based MC scheme decides on the modulation format that has the maximum likelihood within multiple hypotheses. Following the Bayesian formulation, hypothesis testing is performed on the possible modulation formats to estimate the constellation of the interferer. We consider four hypotheses: $\mathbf{y} \sim P(\mathbf{y}; x_1 \in \bar{\Lambda}, x_2 \in \Lambda_j)$, $j \in \{0, 1, 2, 3\}$, with likelihoods:

$$P(\mathbf{y}; \Lambda_j) = \sum_{x_1 \in \bar{\Lambda}, x_2 \in \Lambda_j} P(\mathbf{y}|x_1, x_2) P(x_1, x_2) \quad (5)$$

Under statistical independence between x_1 and x_2 , and assuming uniform priors, $P(x_1) = 1/|\bar{\Lambda}|$ and $P(x_2) = 1/|\Lambda_j|$, where $|\cdot|$ denotes the cardinality of the constellation ($P(x_1)$ is fixed over hypotheses and thus can be dropped), the ML MC decision metric can be derived as:

$$\hat{j} = \arg \max_{j \in \{0,1,2,3\}} \sum_{x_1 \in \bar{\Lambda}, x_2 \in \Lambda_j} P(\mathbf{y}|x_1, x_2) \frac{1}{|\Lambda_j|} \quad (6)$$

Noting that $P(\mathbf{y}|x_1, x_2) = \frac{1}{(\pi\sigma_n^2)^{Nr}} \exp(-\frac{1}{\sigma_n^2} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2)$, and neglecting the term $\frac{1}{(\pi\sigma_n^2)^{Nr}} = \frac{1}{(\pi\sigma_n^2)^2}$ which is assumed fixed over hypotheses, the resultant Log-MAP decision metric is:

$$\hat{j}_{\text{Log-MAP}} = \arg \max_{j \in \{0,1,2,3\}} \left(\log \frac{1}{|\Lambda_j|} + \log \sum_{x_1 \in \bar{\Lambda}, x_2 \in \Lambda_j} \exp\left(-\frac{1}{\sigma_n^2} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2\right) \right) \quad (7)$$

which is the optimal ALRT solution. Note that neglecting the correction term $\log(1/|\Lambda_j|)$ results in the GLRT solution [16].

Solving equation (7) is computationally intensive, because for each j we have to calculate $|\bar{\Lambda}| \times |\Lambda_j|$ exponential terms. However, one of these terms is dominant and corresponds to the ML distance:

$$d_{\text{ML},j} = \min_{x_1 \in \bar{\Lambda}, x_2 \in \Lambda_j} \varphi_{\text{ML}}(\mathbf{x}) \quad (8)$$

$$\varphi_{\text{ML}}(\mathbf{x}) = \frac{1}{\sigma_n^2} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \quad (9)$$

Hence, following the approximation $(\log \sum_r \exp(a_r) \approx \max_r a_r)$, we obtain:

$$\hat{j}_{\text{Max-Log-MAP}} = \arg \max_{j \in \{0,1,2,3\}} \left(\log \frac{1}{|\Lambda_j|} - d_{\text{ML},j} \right) \quad (10)$$

which is the sub-optimal Max-Log-MAP classifier [7] [8].

5. PROPOSED MU-MIMO RECEIVERS

Since the more distance metrics that get accumulated in equation (7), the better the approximation is, we can enhance the classifier by considering the most influential N distances that best minimize $\varphi_{\text{ML}}(\mathbf{x})$. We call this approach the Closest- N classifier, and we will use it as a reference to compare our second proposed approach to.

We next consider a special subset of distances, that consists of the counter ML distances corresponding to bits in \mathbf{x}_b in addition to the ML distance, and we call its corresponding classifier CMLD. Note that with CMLD, the distances considered are not the smallest, and hence not the most influential. The counter ML distance corresponding to a specific bit is defined as:

$$d_{\text{cML},j,i} = \begin{cases} \min_{x_1 \in \bar{\Lambda}, x_2 \in \Lambda_j | b_i=0} \varphi_{\text{ML}}(\mathbf{x}) & b_i^{(\text{ML})} = 1 \\ \min_{x_1 \in \bar{\Lambda}, x_2 \in \Lambda_j | b_i=1} \varphi_{\text{ML}}(\mathbf{x}) & b_i^{(\text{ML})} = 0 \end{cases} \quad (11)$$

with $b_i^{(\text{ML})}$ being the value of the i th bit in the bit vector of the ML solution.

Equation (12) generalizes the likelihood function assuming T observations (tones) are accumulated under a constant interfering modulation type before deciding on a winning hypothesis, where S corresponds to the subset of lattice points to consider.

$$\hat{j} = \arg \max_{j \in \{0,1,2,3\}} \sum_{t=1}^T \left(\log \frac{1}{|\Lambda_j|} + \log \sum_{\mathbf{x} \in S} \exp(-\varphi_{\text{ML}}(\mathbf{x})) \right) \quad (12)$$

After the classifier decides on \hat{j} , a ML soft-output detector generates the bit LLRs as follows:

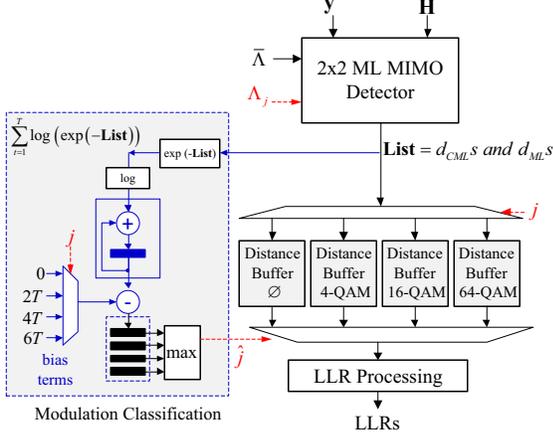
$$\lambda_{\text{ML},i} = \min_{x_1 \in \bar{\Lambda}_{i,1}, x_2 \in \Lambda_{\hat{j}}} \varphi_{\text{ML}}(\mathbf{x}) - \min_{x_1 \in \bar{\Lambda}_{i,0}, x_2 \in \Lambda_{\hat{j}}} \varphi_{\text{ML}}(\mathbf{x}) \quad (13)$$

Hence, the main component of the decision metric for MC is found to be an accumulation over a set of tones of Euclidean distance computations, which are also used by the ML detector for bit LLR soft decision generation. Combining MC and detection routines is thus computationally efficient.

The joint MC and detection setup is described as follows: After observing T vectors, and for each of the four possible hypotheses, the detection routine is called T times and the outputs are stored in memory. Concurrently, the likelihood for each hypothesis gets computed. Eventually, the hypothesis that gets the maximum likelihood is declared a winner and the corresponding output is retrieved. However, these computational savings come at the expense of higher space complexity.

Table 1. Computational Complexity of MC Schemes

	Approach	Lattice Points (S)	Log.	Exp.	Distance Computations
Non-Joint	Log-MAP	All	T	$T \times \bar{\Lambda} \times (\Lambda_0 + \Lambda_1 + \Lambda_2 + \Lambda_3)$	$ \bar{\Lambda} \times (\Lambda_0 + \Lambda_1 + \Lambda_2 + \Lambda_3)$
	Closest- N	Closest N	T	$4 \times T \times N$	$ \bar{\Lambda} \times (\Lambda_0 + \Lambda_1 + \Lambda_2 + \Lambda_3)$
	CMLD	ML + Counter MLs of x_1	T	$4 \times T \times (K + 1)$	$ \bar{\Lambda} \times (\Lambda_0 + \Lambda_1 + \Lambda_2 + \Lambda_3)$
	Max-Log-MAP	ML	T	$4 \times T$	$ \bar{\Lambda} \times (\Lambda_0 + \Lambda_1 + \Lambda_2 + \Lambda_3)$
Joint	Log-MAP	All	T	$T \times \bar{\Lambda} \times (\Lambda_0 + \Lambda_1 + \Lambda_2 + \Lambda_3)$	0
	Closest- N	Closest N	T	$4 \times T \times N$	0
	CMLD	ML + Counter MLs of x_1	T	$4 \times T \times (K + 1)$	0
	Max-Log-MAP	ML	T	$4 \times T$	0


Fig. 1. Architecture for a 2×2 MU-MIMO detector

6. PROPOSED ARCHITECTURE

An optimized architecture for a 2×2 MU-MIMO detector following the CMLD approach is shown in Fig. 1. At the core of this architecture is a ML MIMO detector, that detects \mathbf{x} assuming all possible choices of the interferer's modulation type, and generates the corresponding lists of d_{CMLs} and d_{MLS} for all T vectors. These distances and symbols are stored in buffers of size $T \times (K + 1)$. The sum of the logarithm of the exponential of the distance metrics are passed to an adder that accumulates them over a span of T tones, during which the interferer modulation is assumed to be static. The resulting accumulated distances for each interferer hypothesis are stored in a buffer, and after deciding on a winning hypothesis, the corresponding stored distances are forwarded for LLR processing.

This algorithm can be used in 802.11ac (WiFi) [17], which supports 80MHz of bandwidth with 242 usable tones, 8 of which are reserved for pilots and 234 data tones (worst case is 20MHz, 4 pilots, and 52 data tones). The length of the data field in a WiFi frame can be a very large number of OFDM symbols (L). Since the interferer's modulation constellation remains static over T tones and L symbols, the particular choice of $T = 234$ results in substantial savings in computations. The detector only needs to run in the above mode to identify the interferer's constellation for one OFDM symbol in the frame. It can then switch back to normal ML detection mode (without MC) to generate the LLRs for the remaining OFDM symbols. Moreover, the algorithm can be used in long term evolution (LTE) [18] transmission modes 7, 8, and 9, where estimates of desired and co-scheduled users channels are available at the receiver.

The total number of distance computations needed to generate the LLRs from the $234 \times L$ data tones is $234 \times L \times |\bar{\Lambda}| \times |\Lambda_j|$, and the average overhead of the MC routine is $234 \times 3 \times |\bar{\Lambda}| \times |\Lambda_j|$. This corresponds to an increase of only $3/L\%$, compared to distances

computed by an ML detector with perfect knowledge of the interferer. The size L of the data field can take a range of values from 8 to more than 1024, hence, the increase in distance computations ranges between 37.5% and less than 0.3%. However, an additional burden of the MC routine is the number of exponential and logarithmic operations it requires (blue box in Fig. 1). Table 1 compares the computational complexity of different MC routines, both when applied solely or in a joint setup, in terms of the number of logarithmic, exponential, and Euclidean distance computations.

7. SIMULATION RESULTS

Joint MC and ML detection was implemented following the system model described in Sec. 2. The decision on hypothesis is done after receiving $T = 52$ or $T = 12$ tones, with the first corresponding to WiFi's worst case, and the second to an LTE scenario. Turbo coding is used, with a code rate of $1/3$ and number of decoding iterations equal to 4. We considered the scenario where the user of interest uses 16-QAM, while the interferer hops over the four hypotheses with equal probability on every new frame. This scenario reflects the realistic situation where the interferer changes from frame to frame. We consider two channel types, the first is zero-mean complex Gaussian and circularly symmetric with unit variance, that is independent and identically distributed from tone-to-tone (rich scattering). The second assumes high antenna correlations, with transmit and receive correlation coefficients of 0.9.

Several receiver types are simulated. Four of these receivers are assisted by the MC schemes studied in sections 4 and 5: Log-MAP, Max-Log-Map, Closest-5 (Closest- N with $N = 5$) and CMLD. In addition, we included the receiver that always assumes the interferer to be 16-QAM, as well as the ideal IA receiver. Figure 2 shows, for highly correlated channels, the correct classification ratio (CCR) for the MC approaches with $T = 12$. The CCR gaps are remarkable, but all approaches converge to unity at high SNR. Figures 3 and 4 show the coded frame error rate (CFER) plots with high channel correlation, for $T = 52$ and $T = 12$, respectively. The choice $T = 52$ made the Log-MAP MC-based receiver approach the performance of the IA receiver. On average, compared to Max-Log-MAP, CMLD resulted in a CFER SNR gain of 0.6 dB, Closest-5 a gain of 1.1 dB, and Log-MAP a gain of 2.2 dB. Moreover, the IRC receiver and the receiver that assumes the interferer to be 16-QAM performed badly under high channel correlation.

With ideal channel conditions, the total gap between the Log-MAP and Max-Log-MAP MC-based receivers does not exceed 0.7 dB, as shown in Fig. 5 for $T = 12$. Compared to Max-Log-MAP, CMLD resulted in a CFER SNR gain of 0.15 dB and Closest-5 a gain of 0.3 dB. Note that the Log-Map MC-based detection did not approach the optimal ML IA receiver here, because the CCR values are far from unity over low SNR range.

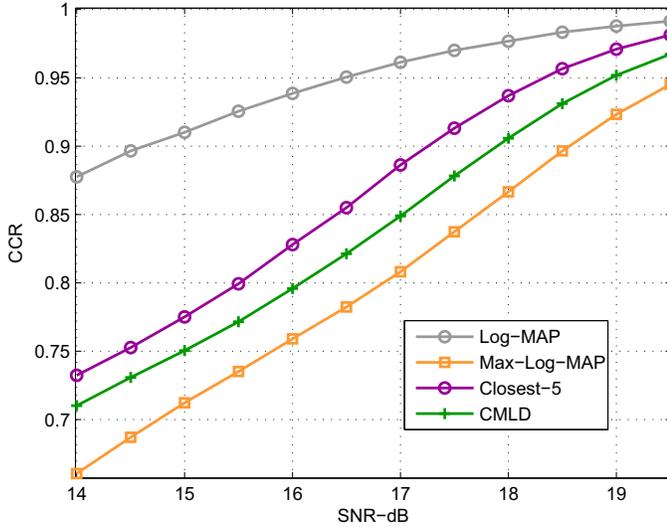


Fig. 2. CCR Performance - $T = 12$ - Correlated Channels

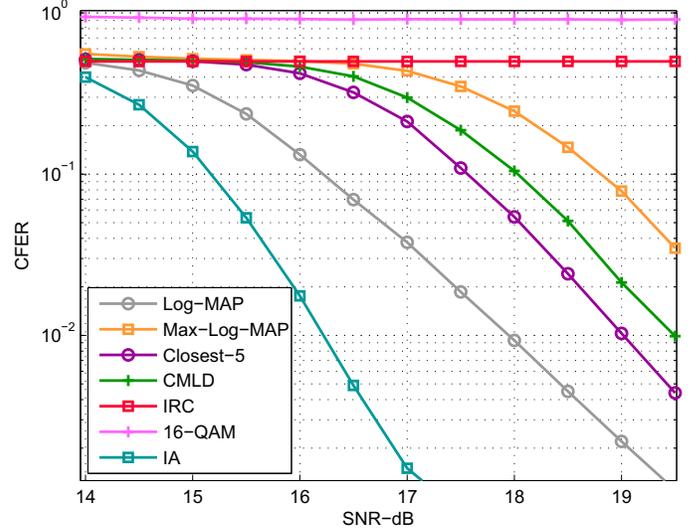


Fig. 4. CFER Performance - $T = 12$ - Correlated Channels

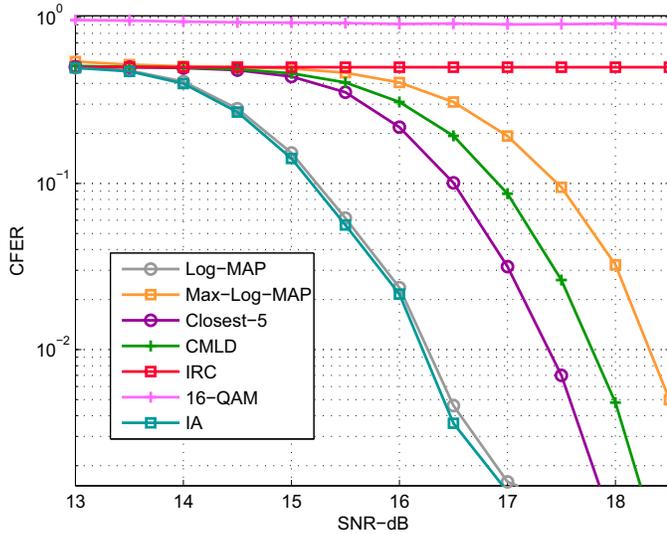


Fig. 3. CFER Performance - $T = 52$ - Correlated Channels

The benefits of CMLD become clearer in the context of joint MC and soft-output sphere decoding [13]. A sphere decoder reduces the number of visited lattice points, however, points leading to counter ML distances are never omitted and efficient joint CMLD MC and detection is maintained. On the other hand, the Closest- N classifier is more attractive if applied jointly with list sphere decoding [19], which keeps track of distances to closest neighbouring symbols, or with sphere decoding with adaptive radius pruning [20, 21], both of which do not guarantee the inclusion of counter ML symbols. Finally, this work can be combined with [22, 23], using constant and linear Max-Log-MAP to enhance the approximation performance.

8. CONCLUSIONS AND DISCUSSION

In this paper, near-ML receivers for MU-MIMO systems have been proposed, which are based on joint MC and detection, and have been compared to other state of the art receivers. Two novel low complexity MC approaches have been proposed, that offer optimized low complexity realizations. The performance of the proposed receivers has been shown to lie between that of MC-based receivers

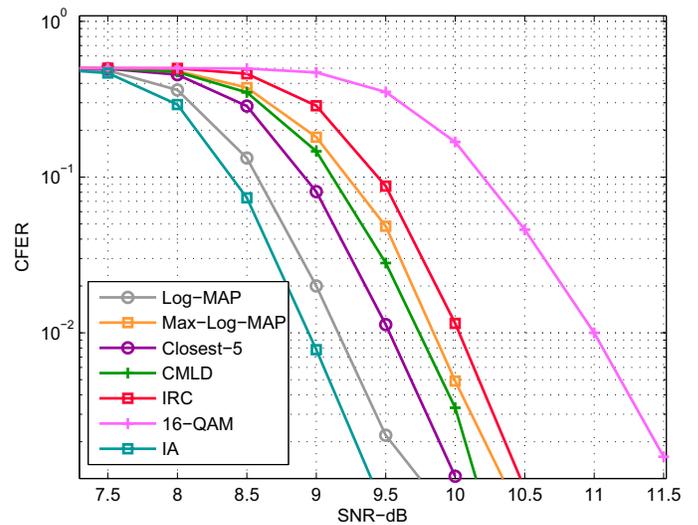


Fig. 5. CFER Performance - $T = 12$ - Uncorrelated Channels

that employ Log-MAP and Max-Log-MAP MC schemes, remarkably beating the latter, especially in case of high channel correlation. An efficient hardware implementation has been proposed. Finally, it has been shown that the proposed approaches can be used in various current communication standards, where in the special case of WiFi they result in a negligible complexity overhead.

Having only optimized a 2×2 MIMO detection core does not mean that the corresponding optimizations will not scale up with higher order MIMO. To support more layers, an efficient channel decomposition scheme exists [9] [24], that reduces the detection problem of multiple layers into multiple 2-layer detection subproblems, that map onto the 2-layer core with a slight modification.

Moreover, the increase in constellation size also threatens the scalability of the proposed schemes. An intuitive approach in this scenario is to limit the search regions to special subsets of the constellations [25]. However, if the search region shrinks, there will be no guarantee on finding the ML symbol vector, nor the counter MLs. Nevertheless, we can choose the symbol vectors with the best distance metrics within a search region to be the pseudo-ML and pseudo-cML points, and use them in our proposed scheme.

9. REFERENCES

- [1] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
- [2] J. Lee, J.-K. Huan, and J. Zhang, "MIMO technologies in 3GPP LTE and LTE-Advanced," *EURASIP J. on Wireless Commun. and Netw.*, vol. 2009, no. 1, pp. 1-10, 2009.
- [3] J. Duplity *et al.*, "MU-MIMO in LTE systems," *EURASIP J. on Wireless Commun. and Netw.*, vol. 2011, no. 1, pp. 1-13, 2011.
- [4] Z. Bai *et al.*, "On the equivalence of MMSE and IRC receiver in MU-MIMO systems," *IEEE Commun. Lett.*, vol. 15, no. 12, pp. 1288-1290, Dec. 2011.
- [5] R. Ghaffar and R. Knopp, "Interference sensitivity for multiuser mimo in lte," *IEEE Workshop on Sig. Proc. Advances in Wireless Commun. (SPAWC)*, pp. 506-510, Jun. 2011.
- [6] —, "Interference-aware receiver structure for multiuser MIMO and LTE," *EURASIP J. on Wireless Commun. and Netw.*, vol. 40, pp. 1-17, 2011.
- [7] A. Gomma *et al.*, "Multi-User MIMO Receivers With Partial State Information," *arXiv preprint arXiv:1502.00212*, 2015.
- [8] J. H. Bae *et al.*, "Advanced downlink MU-MIMO receiver for 3GPP LTE-A," *IEEE ICC Workshop on Telecommunications: from Research to Standards*, 2012.
- [9] M. M. Mansour and L. Jalloul, "Optimized configurable architectures for scalable soft-input soft-output MIMO detectors with 256-QAM," *IEEE Trans. Signal Process.*, vol. 22, no. 4, pp. 408-412, Apr. 2015.
- [10] O. Dobre *et al.*, "Survey of Automatic Modulation Classification Techniques: Classical Approaches and New Trends," *IET Communications*, vol. 1, no. 2, pp. 137-156, Apr. 2007.
- [11] W. Wei and J. Mendel, "Maximum-likelihood classification for digital amplitude-phase modulations," *IEEE Trans. Commun.*, vol. 48, no. 2, pp. 189-193, Feb. 2000.
- [12] P. Panagiotou, A. Anastopoulos and A. Polydoros, "Likelihood ratio tests for modulation classification," *Proc. IEEE MILCOM*, vol.2, pp.670-674 vol.2, 2000.
- [13] B. Shim and I. Kang, "Joint modulation classification and detection using sphere decoding," *IEEE Signal Proc. Lett.* 16, vol. 16, no. 9, pp. 778-781, Jun. 2009.
- [14] V. Choqueuse *et al.*, "Blind modulation recognition for MIMO systems," *J. Military Technical Academy Review*, vol. XIX, pp. 183-196, Jun. 2009.
- [15] M.S. Muhlhous *et al.*, "Automatic Modulation Classification for MIMO Systems Using Fourth-Order Cumulants," *IEEE Vehicular Technology Conf. (VTC Fall)*, pp.1-5, 3-6 Sept. 2012
- [16] T. Yucek and H. Arslan, "A novel sub-optimum maximum-likelihood modulation classification algorithm for adaptive OFDM systems," *Proc. IEEE Wireless Commun. Netw. Conf.*, vol. 2, pp. 739-744, Mar. 2004.
- [17] *IEEE Standard for local and metropolitan area networks – Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY)*, IEEE Std. 802.11, 2012. [Online]. Available: <http://standards.ieee.org/getieee802/802.11.html>
- [18] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation*, 3GPP Std. TS 36.211. [Online]. Available: <http://www.3gpp.org>
- [19] B. Hochwald and S. ten Brink, "Achieving nearcapacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389-399, Mar. 2003.
- [20] M. M. Mansour, S. Alex, and L. Jalloul, "Reduced complexity soft-output MIMO sphere detectors Part I. Algorithmic optimizations," *IEEE Trans. Signal Proc.*, vol. 62, no. 21, Nov. 2014, pp. 5505-5520.
- [21] M. M. Mansour, S. Alex, and L. Jalloul, "Reduced complexity soft-output MIMO sphere detectors Part II. Architectural optimizations," *IEEE Trans. Signal Proc.*, vol. 62, no. 21, Nov. 2014, pp. 5521-5535.
- [22] B. Classon, K. Blankenship, and V. Desai, "Turbo decoding with the constant-log-MAP algorithm," *Proc. Int. Sympos. Turbo Codes and Related Topics*, pp. 467-470, Sep. 2000.
- [23] M. Valenti and J. Sun, "The UMTS turbo code and an efficient decoder implementation suitable for software-defined radios," *Int. J. Wireless Inform. Netw.*, vol. 8, no. 4, pp. 203-214, Oct. 2001.
- [24] M. M. Mansour, "A Low-Complexity MIMO Subspace Detection Algorithm," *EURASIP Journal on Wireless Communications and Networking*, Mar. 2015.
- [25] A. Tomasoni *et al.*, "Low Complexity, Quasi-Optimal MIMO Detectors for Iterative Receivers," *IEEE Trans. on Wireless Commun.*, vol.9, no.10, pp.3166-3177, Oct. 2010.