EVALUATION OF MIMO SYMBOL DETECTORS FOR 3GPP LTE TERMINALS

Di Wu, Johan Eilert and Dake Liu

Department of Electrical Engineering, Linköping University, 58183 Linköping, Sweden
{diwu, je, dake}@isy.liu.se

ABSTRACT

This paper investigates various MIMO detection methods for 3GPP LTE open-loop downlink multi-antenna transmission. Targeting VLSI implementation, these detection methods are evaluated with respect to complexity and detection performance. A realistic 3GPP LTE simulation chain is developed for the evaluation. The result shows that with the aid of Hybrid Automatic Repeat reQuest (H-ARQ), a recently proposed reduced complexity close-ML detector called MFCSO achieves a good tradeoff between achievable throughput and complexity. An adaptive transmission and detection scheme is also proposed based on user scenarios.

1. INTRODUCTION

Multi-antenna or multi-input and multi-output (MIMO) technologies have been widely adopted by latest wireless standards. 3GPP Long-Term Evolution (LTE) is the 4th generation radio access technology which incorporates Orthogonal Frequency Division Multiple Access (OFDMA) as the multiple access scheme in downlink. MIMO technologies are also mandatory in LTE to achieve the LTE bit-rate targets (e.g 100 Mbit/s peak data rate for downlink). As part of the receiver chain depicted in Fig. 1, MIMO symbol detection is a significant challenge for VLSI implementation.

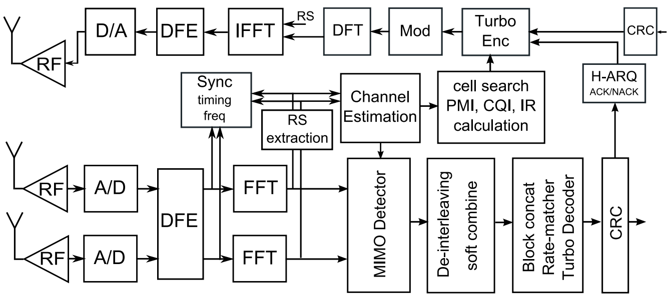

Figure 1: Baseband Chain of a 3GPP LTE Receiver

Various MIMO detection methods and their respective implementations have been proposed in literature such as [1], [2], [3] and [5]. However, none of them has taken the system specific features of LTE (e.g. OFDMA and H-ARQ) into consideration and are mostly based on very simple channel models (e.g. AWGN). In this paper, with the aid of a more realistic LTE simulation chain and 3GPP SCME channel model, several MIMO detection algorithms are applied to LTE system and with their performance quantitatively evaluated. Second, although the MFCSO detection algorithm proposed by the authors in [5] has a very low detection complexity, under random AWGN channels, it requires relatively strong channel coding to maintain a close-ML performance in frame-error-ratio [5]. In this paper, its performance with the aid of H-ARQ is investigated. Based on the performance and complexity analysis, an adaptive transmission and detection mechanism is proposed by the authors for different user scenarios.

The remainder of the paper is organized as follows. In Sec. 2, MIMO schemes in 3GPP LTE are presented in brief. Sec. 3 introduces several MIMO detection algorithms evaluated in this paper. Sec. 4 briefly describe the simulation chain and its configuration in this paper. Sec. 5 presents the simulation performance and Sec. 6 addresses the complexity issues. An adaptive transmission and detection scheme is proposed in Sec. 7. Finally, Sec. 8 concludes the paper.

2. MULTI-ANTENNA TRANSMISSION IN LTE

As defined in 3GPP LTE standard [8], the procedure to map modulated symbols to different antennas is called antenna mapping which in general supports up to two code streams and four transmitting antennas. As depicted in Fig. 2, antenna mapping consists of two parts namely layer mapping and precoding. The former multiplexes the modulated symbols belonging to one or two codewords into different number of layers (or codeblocks) to transmit. The later loads symbols from each layer and jointly process these symbols in time or frequency domain before mapping them to different antennas. In this paper, a configuration with only two transmitting antennas and two receiving antennas is considered. In orthogonal frequency multiplexing access (OFDMA) systems such as LTE, the general transmission model of each subcarrier is

$$r = \mathbf{H}s + n \tag{1}$$

where $\mathbf{H}$ is the frequency domain channel matrix, s and r are in respect the transmitted and received symbol vector.

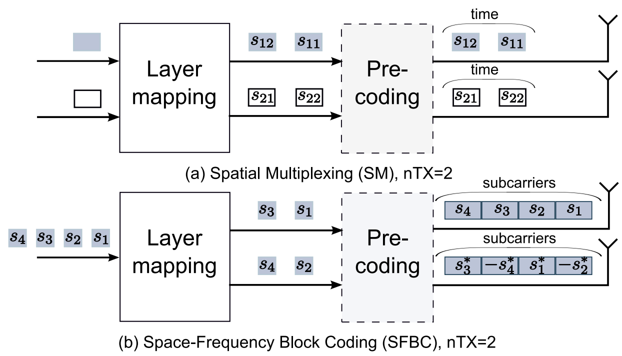

Figure 2: Downlink Multi-antenna Transmission Schemes

2.1 Spatial Multiplexing

Spatial multiplexing (SM) is a MIMO technique aimed at maximizing the data throughput by exploiting the degrees of freedom in MIMO channels. Since the multiplexing gain is only available for high SNR region, spatial multiplexing is usually used when high-SNR is available. As depicted in Fig. 2(a), for 2×2 SM in LTE, there are two codewords, with the first codeword is mapped to the first layer and the second codeword mapped to the second. In general, the degree of freedom (multiplexing gain) is determined by $min(n_t, n_r)$ which is the rank of the channel matrix

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \tag{2}$$

In case $\mathbf{H}$ is badly conditioned (e.g. when line-of-sight occurs), linear detection based on the pseudo-inversion of $\mathbf{H}$ in (6) will perform

© EURASIP, 2009

poorly. In other words, the gain of spatial multiplexing heavily depends on the multipath fading. To allow close-loop beamforming based on codebook, a pre-coding matrix W can be multiplied with the layer mapped symbols at the transmitter side. For downlink, W is usually computed at the basestation based on the codebook and UE feedback.

2.2 Space-Frequency Block Coding

Similar to Space-Time Block Coding (STBC), Space-Frequency Block Coding (SFBC) [8] is a technique to transmit data for guaranteed diversity with a low complexity symbol detector on the receiver side. Alamouti matrix [6] based orthogonal STBC has been widely adopted in latest wireless standards for the reason that it is the only full-rate linear STBC code with a diversity gain of 2. In other words, the SFBC considered in this paper is an Alamouti schemes in space and frequency domain. This assumes the channels of neighboring subcarriers are identical, so that when a single codeword is mapped to several neighboring subcarriers, frequency diversity is achieved. The basic 4×2 space-frequency channel matrix is defined as

$$\mathbf{H} = \begin{bmatrix} h_{11} & -h_{12} \\ h_{12} & -h_{22} \\ h_{12}^{*} & h_{11}^{*} \\ h_{22}^{*} & h_{12}^{*} \end{bmatrix} \quad (3)$$

3. MIMO DETECTION ALGORITHMS

For MIMO systems, a major challenge is the symbol detection at the receiver. As channel coding (e.g. Turbo) is used, soft-output, in effect the log-likelihood ratio (LLR), must be computed. Maximum Likelihood (ML) detection which is the optimum detector computes

$$L(b_i|r) = \log\left(\frac{\sum_{s:b_i(s)=1}\exp(-\frac{1}{\sigma^2}\|r-\mathbf{H}s\|^2)}{\sum_{s:b_i(s)=0}\exp(-\frac{1}{\sigma^2}\|r-\mathbf{H}s\|^2)}\right) \quad (4)$$

Here "$s : b_i(s) = \beta$" means all s for which the ith bit of s is equal to β. Computing (4) requires enumeration of the entire set of possible transmitted vectors. The complexity of doing this is usually not affordable for implementation in practice. However since ML provides the best theoretical performance, it is commonly used as a benchmark when comparing other algorithms.

3.1 Linear Detection

Linear detection schemes such as Zero-Forcing (ZF) and Minimum Mean-Square-Error (MMSE) have very low complexity. The only difference between ZF and LMMSE is the later one takes the noise power σ^2 into consideration while the former does not. The ZF and LMMSE detection is defined in the following

$$ZF : \hat{s}_{ZF} = (\mathbf{H^H H})^{-1}\mathbf{H}^H r \quad (5)$$

$$MMSE : \hat{s}_{MMSE} = (\mathbf{H^H H} + \sigma^2\mathbf{I})^{-1}\mathbf{H}^H r \quad (6)$$

The equation shows that matrix inversion is involved in the detection. The low complexity of linear detection makes them attractive for VLSI implementation, though they have relatively poor performance especially when the channel is slow-fading [3]. Fortunately, the "frequency hopping" of multiple users in OFDMA creates a fast fading channel for each individual user, which will to some extent improve the performance of linear detection.

3.2 Fixed-Complexity Soft-Output (FCSO) Detection

As a tradeoff between performance and complexity, sphere decoding such as [1] have been proposed to reach close-ML performance with lower complexity than ML. However, the complexity of sphere decoding grows exponentially with the number of transmit antennas and polynomially in the size of the signal constellation. More importantly, the tree search used in sphere decoding is in principle a sequential procedure which is difficult to parallelize. In [2], a fixed-throughput sphere detector was proposed with fixed-complexity and parallelism for hard-decision. A method namely layered orthogonal lattice detector (LORD) is presented in [4] to compute the soft-decision. Similarly, the FCSO detector [3] which computes soft-output, achieves close-ML detection performance via fully enumerating only one transmitted symbol and applying decision feedback equalization (DFE) to the rest of the symbols. However, the complexity of both FCSO and LORD will increase substantially as the constellation grows (e.g. from 16-QAM to 64-QAM).

3.3 Modified FCSO Detection

In [5], a reduced complexity variant of FCSO [3] for high-order modulation schemes is proposed called MFCSO for Modified FCSO. This section essentially repeats the algorithm description given in [5]. The approximation in MFCSO consists of only partially enumerating the symbols selected for exact marginalization. Taking a 2×2 MIMO system as an example, considering each complex-valued symbol as one layer, only one of them is exactly marginalized with the other approximately marginalized (using DFE hard-decision). The channel rate processing of MFCSO involves the QR decomposition (QRD) of two 2×2 channel matrices which are $\mathbf{H_1} = \mathbf{H}$ in (2) and

$$\mathbf{H_2} = \begin{bmatrix} h_{12} & h_{11} \\ h_{22} & h_{21} \end{bmatrix} \quad (7)$$

The QRD generates an upper triangular matrix R, and a unitary matrix Q so that

$$\mathbf{H_1} = \mathbf{Q_1 R_1} \qquad \mathbf{H_2} = \mathbf{Q_2 R_2} \quad (8)$$

Slightly different from the MFCSO presented in [5], the detection procedure for 2×2 SM is in the following

1. Linear detection in (6) or (5) is carried out to estimate the 2×1 initial symbol vector

$$\hat{s}_{init} = \min_{\hat{s}_{init,k}\in\mathscr{L}} \|\mathbf{H_1}s - r\|^2 \quad (9)$$

   Here s is the transmitted symbol vector, within which, s_k is the k^{th} symbol.
2. For each initially estimated symbol $\hat{s}_{init,k}$, $k\in\{1,2\}$, a candidate set $\mathscr{L}_k$ is created. $\mathscr{L}_k$ contains N lattice points close to $\hat{s}_{init,k}$. In this paper, it is decided that $N = 16$ for 64-QAM and $N = 9$ for 16-QAM.
3. First s_2 is chosen as the top-layer symbol. In order to perform DFE,

$$\widetilde{r} = \mathbf{Q_1}^H r \quad (10)$$

   needs to be computed. The same operation is needed once again when s_1 is chosen as the top layer later.
4. For the n^{th} constellation point $\zeta_n \in \mathscr{L}_2$, its effect on $\widetilde{r}_1$ will have to be canceled out.

$$\widetilde{r_1} = \widetilde{r}_1 - \mathbf{R_1}(1,2)\zeta_n \quad (11)$$

   Based on ζ_n, the partial Euclidean distance

$$\delta_n = \|\mathbf{R_1}(2,2)\zeta_n - \widetilde{r}_2\|^2 \quad (12)$$

   computed for the top-layer.
5. DFE is applied to detect the other layer. Using back-substitution [7], $\hat{s}_1$ can be estimated from

$$\hat{s}_1 = \arg\min_{\hat{s}_1\in\mathscr{L}} \|\mathbf{R_1}(1,1)\hat{s}_1 - \widetilde{r}_1\|^2 \quad (13)$$

6. The estimated $\hat{s}_1$ together with $\hat{s}_2 = \zeta_n$ form a complete possible transmitted symbol vector $\hat{s}$, based on which, an accumulated full Euclidean distance

$$\delta_n = \delta_n + \|\mathbf{R_1}(1,1)\hat{s}_1 - \tilde{r}_1\|^2 \tag{14}$$

can be computed.

7. In total, there will be N different δ_n computed when s_2 is chosen as the top layer. Then s_1 is chosen as the top-layer symbol as well. Based on $\mathbf{Q_2}, \mathbf{R_2}$ and $\hat{s}_{init,1}$, the same procedure needs to be done once again to computed another N different δ_n. Hence for the 2×2 system, $2N$ different δ_n values need to be computed. They are used to update the LLR values in the end [5].

4. 3GPP LTE SIMULATION CHAIN

In order to carry out both fast prototyping and verification of the 3GPP LTE modems, a complete physical layer behavior model and simulation chain has been developed in Matlab and C. In combination to an LTE signal generator, it allows both quantitative performance evaluation and conformance test of the chip. The simulation chain includes a transmitter conforming to 3GPP technical spec [8][9] and [10], and a receiver which supports timing/frequency synchronization, channel estimation, subcarrier demapping, rate-matching, turbo decoding and Cyclic Redundancy Check (CRC). H-ARQ based on chase combining (CC) is included with up to three times retransmission allowed. The 3GPP SCME model [11] is used as the channel model. In the simulation done for this paper, 5000 subframes are simulated. Both 2×2 SM and 2×2 SFBC are chosen as the MIMO configuration. No close-loop precoding is assumed in this paper. Throughput is calculated based on the method in [12].

Channel Quality Indicator (CQI)	9, 15
Modulation	16-QAM/64-QAM
System bandwidth	5MHz
Num of UE	1
Num of BS	1
Channel model	Urban Micro
UE speed	3km/h
Channel estimation	Ideal
H-ARQ	Chase Combining
Turbo iterations	8

Table 1: Simulation Parameters

CQI	Modulation	Code rate
9	16-QAM	0.602
15	64-QAM	0.926

Table 2: CQI parameters in simulation [10]

5. PERFORMANCE ANALYSIS

Fig. 3, 4 and 5 show that in order to support CQI=15, relatively high SNR is required, which means the UE has to be close to the BS. Meanwhile, for 2×2 SM, FCSO achieves ML performance which is 7 dB better than MFCSO when reaching FER$= 0.01$ in Fig. 3 when the weakest code is used (0.926). MFCSO is around 10 dB better than MMSE to reach FER$= 0.01$ in the same criteria. Note that in wireless systems, compared to BER or FER, throughput is a more important performance factor (if not latency) which has direct effect on the user experience. Fig. 5 shows that the gain in throughput brought by MFCSO against MMSE is significant (up to 12Mbits/s, or 55% higher than the one achieved by MMSE). In comparison, the throughput gain brought by FCSO against MFCSO is much smaller (up to 2.5Mbits/s, or 7% higher than that achieved by MFCSO). The much smaller gap in throughput in comparison to that of FER mainly owes to the H-ARQ retransmission with chase combining.

The result shows us that even with a sub-optimal detector (which also implies much lower complexity), a throughput that is close to the one achievable by ML detectors can be reached when H-ARQ is presented.

Simulation result of CQI=9 are depicted in Fig. 6, 7 and 8. The result shows that 16-QAM only requires moderate SNR which will be available in most part of the cell range. It also shows that MFCSO ($N = 9$) achieve the same performance as FCSO and ML detectors. It has a throughput that up to 68% higher than the one achieved by MMSE.

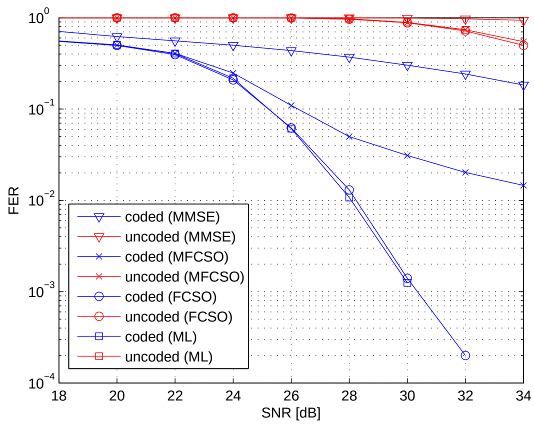

Figure 3: Frame-Error-Ratio (2×2 SM, CQI=15)

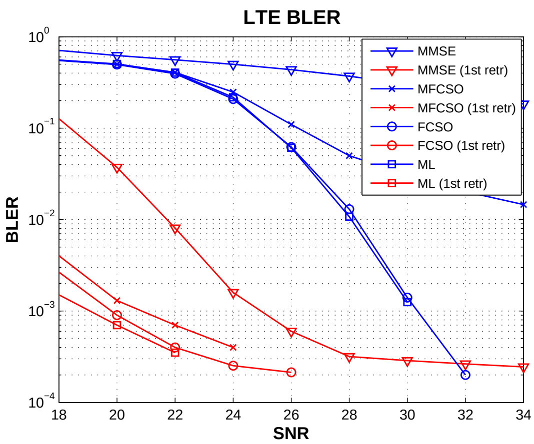

Figure 4: BLock-Error-Ratio (2×2 SM, CQI=15), red curves are the BLER of the 1st retransmission of H-ARQ

Fig. 10 and 9 show the BLER and throughput of 2×2 SFBC with two different CQI values (9 and 15). The simulation shows that SFBC reaches FER=0.01 at much lower SNR than SM as depicted in Tab. 3, though the throughput is half.

CQI	SFBC (MMSE)	SM (MFCSO)	SM (MMSE)
9	10 dB	17 dB	24 dB
15	24 dB	36 dB	N/A

Table 3: Minimum SNR to reach FER=0.01

Fig. 11 depicts the achievable throughput using two-level adaptive modulation and coding (AMC). The result shows that when SNR is worse than 10 dB, SFBC achieves both higher throughput and lower BLER than SM even if ML detector is used.

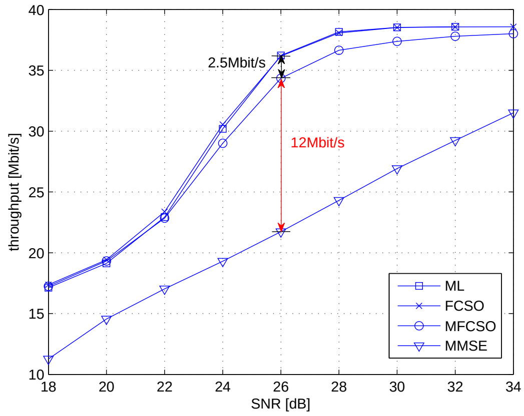

Figure 5: Coded Throughput (2×2 SM, CQI=15)

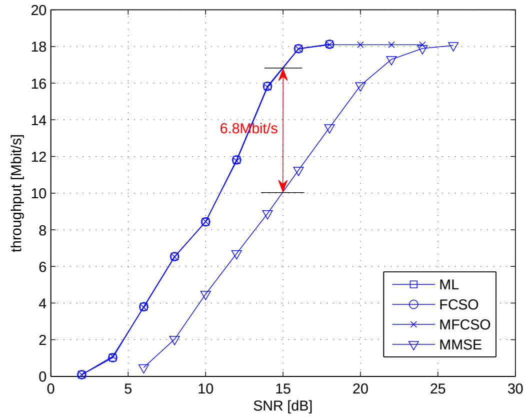

Figure 8: Coded Throughput (2×2 SM, CQI=9)

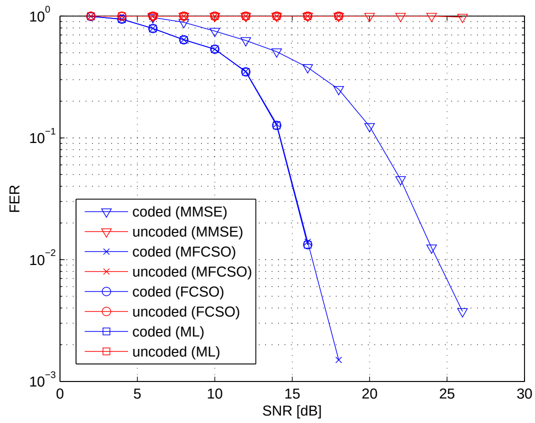

Figure 6: Frame-Error-Ratio (2×2 SM, CQI=9)

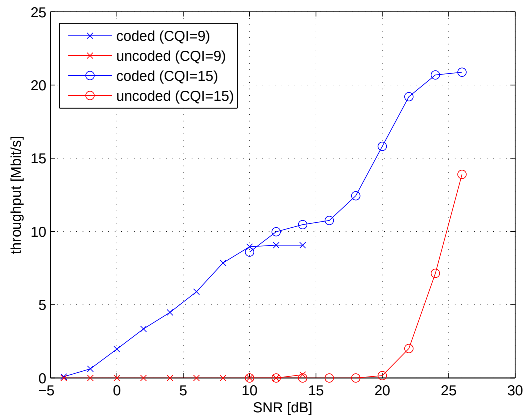

Figure 9: Througput (2×2 SFBC, MMSE)

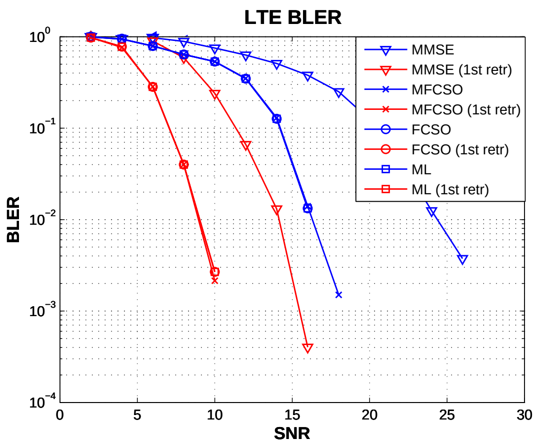

Figure 7: BLock-Error-Ratio (2×2 SM, CQI=9)

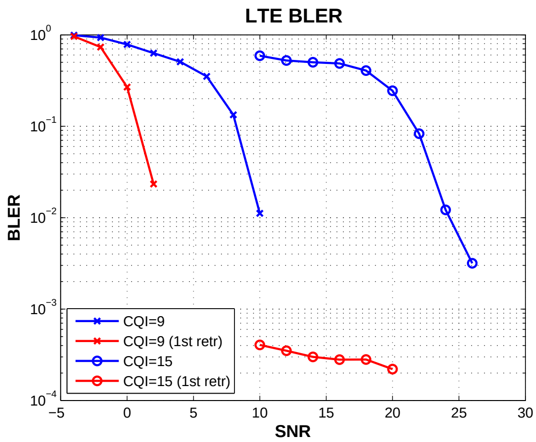

Figure 10: BLock-Error-Ratio (2×2 SFBC, MMSE)

6. IMPLEMENTATION CONSIDERATIONS

In LTE [8], taking a 5 MHz bandwidth LTE system as an example, up to 7 OFDM symbols need to be processed within one slot (0.5 ms) which contain 1900 data subcarriers. This means that there will be no more than $0.26\mu s$ to finish the detection of each subcarrier in average. Therefore, proper detection methods have to be chosen in order to maximize the data rate at reasonable implementation cost.

As depicted in Eq. (6), for 2×2 SM, the MMSE detector needs to compute the inverse of a 2×2 matrix. It has been presented in [13] that the inversion of small matrices can be done using direct inversion which supplies sufficient precision for most of the channels. The FCSO and MFCSO detector involves the search of a number of trellis nodes as depicted in Tab. 4. The FCSO detector always visits the complete constellation (e.g. 16 for 16-QAM and 64 for 64-QAM) while MFCSO only visits a subset of it (e.g. 9 for 16-QAM and 16 for 64-QAM). Note that MFCSO requires MMSE detection to compute the inital estimate (9) which is an extra cost compared to FCSO.

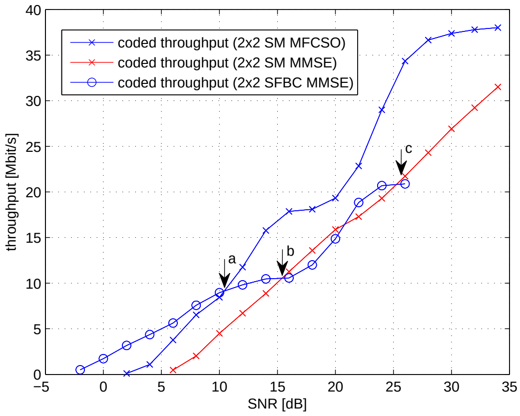

Figure 11: Coded Throughput with 2-level AMC (CQI 15 and 9)

		MMSE	MFCSO	FCSO	ML
Num nodes	16-QAM	1	18	32	256
	64-QAM	1	32	128	4096
Area Estimate (mm^2)	64-QAM	0.08	0.2	0.6	20

Table 4: Complexity Analysis for ASIC Implementation (65 nm)

In practice, the hardware is usually implemented taking both the cost and performance issues into consideration. Based on the complexity analysis in Tab. 4 and the performance analysis in Sec. 5, MFCSO falls into the favor of the authors to be chosen as the target algorithm for ASIC implementation. Using ST 65nm CMOS process, while meeting the $0.26\mu s$ constraint, the implemented detector supporting both MMSE and MFCSO for 2×2 SM and up to 64-QAM modulation occupies less than 0.25 mm^2.

7. ADAPTIVE TRANSMISSION AND DETECTION

As depicted in Tab. 4, a detector supporting MFCSO/MMSE consumes 2.5 times the area of the one only supporting MMSE. Hence the former one is assumed to target high-end users willing to pay more in area and power for performance (e.g. laptops). The MMSE single-mode detector is in favor of low-end users for connectivity with minimum cost (e.g. smartphones). Note that the user cares about latency as well as throughput, and latency is partly determined by the number of retransmissions. Hence it is also important to keep the retransmissions to a minimum (which requires low FER). Fig. 11 shows that with AMC, SM using MFCSO detector always brings higher throughput when SNR is greater than 10 dB. For both types of users, when SNR is worse than 10 dB (a in Fig. 11), SFBC is preferred instead of SM. For low-end users, SM can be used when $SNR \geq 25\ dB$ while SFBC is still preferable (due to the low FER thus fewer retransmissions resulting in low latency) used from 10 to 25 dB. For high-end users, SM is preferred when SNR is at least higher than 10 dB. On the other hand, the MMSE-mode will consume substantiately lower power than the MFSCO-mode, the high-end users might only want to switch to MFCSO-mode when there is enough battery power and high SNR (e.g. $\geq 25\ dB$). When SNR is very low, SFBC is also preferred due to its robustness (as depicted in Fig. 11). The SNR ranges suggested for the mode-switching of two types of detector hardware are shown in Tab. 5. The adaptive scheme brings power efficiency and can supply best-effort performance in an economic way.

SNR range	SFBC	SM
High-end Detector (MFCSO/MMSE)	$-2\ dB \rightarrow 10\ dB$	$\geq 10\ dB$
Low-end Detector (MMSE only)	$-2\ dB \rightarrow 26\ dB$	$\geq 26\ dB$

Table 5: Adaptive Transmission and Detection

8. CONCLUSION

In this paper, the result shows that MFCSO [5] detector achieves close-ML throughput in LTE, even with a relatively weak channel code and with high order modulation (e.g. CQI=15). Furthermore, since the algorithm has sufficiently low complexity [5], it is chosen over FCSO [3] and other close-ML detection schemes for VLSI implementation. Based on the adaptive scheme proposed in Sec. 7, a good performance and cost tradeoff can be achieved. The result also emphasizes the need of a configurable detector to enable the adaptive scheme in real-time.

9. ACKNOWLEDGEMENT

The work of D. Wu, J. Eilert and D. Liu was supported in part by the European Commission through the EU-FP7 Multi-base project with Ericsson AB, Infineon Austria AG, IMEC, Lund University and KU-Leuven. The authors would like to thank Christian Mehlführer and the Christian Doppler Laboratory for Design Methodology of Signal Processing Algorithms at Vienna University of Technology, for contributions on the LTE simulation chain.

REFERENCES

[1] A. Burg, M. Borgmann, M. Wenk, M. Zellweger, W. Fichtner, and H. Bölcskei, "VLSI implementation of MIMO detection using the sphere decoding algorithm", *IEEE Journal of Solid-State Circuits*, vol. 40, no. 7, pp. 1566-1577, July 2005.

[2] L. G. Barbero, J. S. Thompson, "Rapid prototyping of a fixed-throughput sphere decoder for MIMO systems", in *Proc. of IEEE ICC*, Jun. 2006.

[3] E. G. Larsson and J. Jalden, "Soft MIMO detection at fixed complexity via partial marginalization", *IEEE Transactions on Signal Processing*, vol. 56, pp. 3397-3407, Aug. 2008.

[4] M. Siti and M. P. Fitz, *"A Novel Soft-Output Layered Orthogonal Lattice Detector for Multiple Antenna Communications"*, in Proc. of IEEE ICC, Jun. 2006.

[5] D. Wu, E. G. Larsson and D. Liu, "Implementation Aspects of Fixed-Complexity Soft-Output MIMO Detection", in *Proc. of IEEE VTC-Spring*, April 2009.

[6] S. M. Alamouti, "A Simple Transmit Diversity Technique for Wireless Communications", *IEEE J. Select. Areas Commun*, vol. 16, no. 8, pp. 1451-1458, 1998

[7] G. H. Golub, C. F. Van Loan, "Matrix Computations, Third Edition", The Johns Hopkins University Press, 1996.

[8] 3GPP, "Tech. Spec. 36.211 V8.4.0, E-UTRA; Physical Channels and Modulation", Sept 2008.

[9] 3GPP, "Tech. Spec. 36.212 V8.4.0, E-UTRA; Multiplexing and Channel Coding", Sept 2008.

[10] 3GPP, "Tech. Spec. 36.213 V8.4.0, E-UTRA; Physical Layer Procedure", Sept 2008.

[11] D. S. Baum, J. Salo, M. Milojevic, P. Kyöti, and J. Hansen *"MATLAB implementation of the interim channel model for beyond-3G systems (SCME)"*, May 2005.

[12] C. Mehlführer, S. Caban, M. Rupp, "Experimental Evaluation of Adaptive Modulation and Coding in MIMO WiMAX with Limited Feedback" *EURASIP Journal on Advances in Signal Processing*, Vol. 2008.

[13] J. Eilert, D. Wu, D. Liu, "Efficient Complex Matrix Inversion for MIMO Software Defined Radio", in *Proc. IEEE ISCAS*, 2007.