A Low-Area Flexible MIMO Detector for WiFi/WiMAX Standards

Nariman Moezzi-Madani, Thorlindur Thorolfsson and William Rhett Davis

Electrical and Computer Engineering Department,
North Carolina State University
Raleigh, USA
{nmoezzi, trthorol, wdavis}@ncsu.edu

***Abstract*— MIMO wireless technology is required to increase the data rates for a broad range of applications, including low cost mobile devices. In this paper we present a very low area reconfigurable MIMO detector which achieves a high throughput of 103Mbps and uses 27 Kilo Gates when implemented in a commercial 180nm CMOS process. The low area is achieved by the proposed in-place architecture. This architecture implements the K-best algorithm and reduces area 4-fold compared to the widely used multi-stage architecture, while provides reconfigurability in terms of antenna configuration during real-time operation**.

I. INTRODUCTION

MIMO has recently garnered interest because it significantly increases data throughput without utilizing additional bandwidth or transmitting power. Novel wireless communication standards such as IEEE 802.11n (WiFi) and IEEE 802.16m (WiMAX) use MIMO to raise the data rate to hundreds of mega bits per second (Mbps). Sphere decoders (SD) are based on tree search and are very popular because of their near ML performance and relatively lower complexity when implement MIMO detectors. There are three main categories of SDs: depth first (DF) [1], FSD (fixed throughput SD) [2] and K-best (breadth-first) [3].

The K-best designs in the literature have considered just one type of antenna structure, mainly 4x4, while the standards use a variable number of antennas at both transmitter and receiver during run-time. Also, the demand for higher data rate systems is increasingly requiring a larger number of antennas. The conventional architecture for the K-best algorithm is the multi-stage architecture and results in huge area consumption for a large number of antennas or constellation sizes. Furthermore, this architecture does not provide reconfigurability for different antenna configurations.

We have designed the in-place architecture that uses the regularity of the K-best algorithm and reduces area 4-fold compared to the multi-stage architecture as well as provides the flexibility to operate with different antenna numbers. We have designed this architecture for the 802.11n standard that supports antenna configurations up to 4x4. Also, we reduced the area of the system by 55% by using the minimum mean squared error–sorted QR decomposition (MMSE-SQRD) technique in the pre-processing unit. This algorithm improves the BER performance by considering the noise level in the QR decomposition process.

II. SOFT-OUTPUT K-BEST SPHERE DECODER

The potential performance of the MIMO technology is not fully delivered with hard-output detectors; soft-output detectors are needed to obtain the maximum performance efficiency. The MIMO detector includes three units: The pre-processing unit provides the decomposed matrixes Q and R for the sphere decoder. The sphere decoder generates the candidate list for the LLR computation unit and this LLR computation unit provides the soft-information for the channel decoder.

A. Soft-Output Detection

Consider a MIMO system with M transmit and N receive antennas. The coded bit streams are modulated using a complex constellation O each point represented by M_c bits.

The modulated streams, S, are transmitted from the M antennas. The received signal at the baseband receiver is:

$$y = Hs + n \qquad (1)$$

where H is the NxM channel matrix and n is N dimensional i.i.d. Gaussian noise. A new vector set, X, which is the bit-level representation of the vector y can be defined as $X = [x_0, x_1, \ldots, x_{M.Mc}]$. The soft-output detector produces reliable information for each of the $M.M_c$ coded bits in X by calculating the extrinsic Log-Likelihood ratio (LLR) values. With the Max-log approximation, the LLR-value for a non-iterative decoding system denoted as $L(.)$ becomes:

$$L(x_j) = \frac{1}{2\sigma^2}\{\min_{x \in X_j^{+1}} \|y - Hs\|^2 - \min_{x \in X_j^{-1}} \|y - Hs\|^2\} \qquad (2)$$

where X_j^{+1} and X_j^{-1} represent two vectors from set X which have the jth bit equal to +1 and -1 respectively. The L-values in (2) need to be calculated for all bits in the received signal. One of the two terms in (2) is always the ML solution, which is also the solution to the hard-output detectors. The other term is created by a vector which its *jth* bit is the inverse of the ML

978-3-9810801-6-2/DATE10 © 2010 EDAA

solution bit, $s \in X_j^{\overline{X_{ML}}}$, the *counter hypothesis* vector. The LLR-values can be calculated from a sorted candidate list which includes the ML solution as well as candidates with small metrics. For more information regarding tree search we refer the readers to [5].

B. K-best Sphere Decoder

The K-best algorithm is a breadth-first SD search that expands only the K nodes (parents) that have the smallest PEDs instead of expanding all of the nodes in each layer of the tree. The number of PEDs in each layer after expansion will be K.Q, where Q is the number of constellation points. Only K smaller PEDs will survive after sorting. The node with the smallest PED in the last stage of the tree is most probably the ML solution. K-best SD has a fixed throughput and is simple to implement. K-best algorithm has the advantage that it can naturally generate the sorted candidate list. The last K survivors in the last stage of the tree generate a sorted candidate list. These candidates are used to compute the LLRs required for the soft-output detection. The performance of the system is highly dependent to the K parameter and the pre-processing techniques.

III. Pre-processing Techniques

The performance of the MIMO detector is highly dependent to the pre-processing unit where QR-decomposition takes place. Beside the SQRD (sorted QR decomposition) and RVD (real-value decomposition) there is another technique which can improve the performance. The majority of SDs [4]-[9] have used the popular Zero-Force QR decomposition. An extension to ZF-SQRD, MMSE-SQRD effectively improves the BER of the system compared to the ZF-SQRD.

A. MMSE-SQRD

Successive Interference Cancellation (SIC) is the base of BLAST detectors. SIC happens layer by layer so that after detecting a layer, the interference of the given layer is removed from all other layers before detecting the next layer. This technique needs QR decomposition to be able to detect the signals layer by layer. An extension to the ZF-QRD, MMSE-QRD was applied to V-BLAST receivers [11]. The difference between ZF-QRD and MMSE-QRD is that MMSE-QRD considers the received signal noise level in the calculations. Instead of the channel matrix H, a new channel matrix $\underline{H} = [H; \sigma_n I_M]$ needs to be used. This matrix decomposes into:

$$\underline{H} = \begin{bmatrix} H \\ \sigma_n I_M \end{bmatrix} = \underline{Q}\underline{R} = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} \underline{R} \quad (3)$$

Multiplying both sides of (1) gives:

$$\hat{y} = Q_1^H y = \underline{R}s + Q_1^H n - \sigma_n Q_2^H s \quad (4)$$

As shown in (4), the noise effect in the received signal is weakened. We have used MMSE-QRD in the K-best algorithm to improve the performance of the detection. The simulation results are shown in the next subsection.

B. Simulation Results

We simulated our detector in a MIMO-OFDM system with a 4x4 antenna structure and 16-QAM modulation. The OFDM frames include 64 subcarriers each. The code rate is 0.5 and a constraint length equal to seven is chosen for the convolutional encoder. The target BER for a coded MIMO system is 10^{-5}. At this BER our K-best detector with K = 5 using MMSE-SQRD technique achieves a better SNR gain than the K-best detectors using ZF-SQRD technique with K = 13 as shown in Fig. 1. The implementation results of the MIMO detector using K = 5 and 13 show that MMSE-SQRD reduces the area of the detector by 55% compared to the ZF-SQRD for the same data rate.

Although the new augmented channel matrix H adds about 50% more complexity to the decomposition process compared to the channel matrix H, but the channel decomposition rate is so much lower than the MIMO detection process. It states that the additional complexity in the channel processing unit is negligible compared to the complexity reduction in the sphere decoder.

IV. Multi-Stage Architecture

All K-best SD designers have used the popular multi-stage architecture for the K-best algorithm. The multi-stage architecture is shown in Fig. 2-a. In this architecture, two Processing Elements (PE) are used for each stage of the tree. All of the 8 PEs are connected sequentially (in a 4x4 system, RVD results in 8 stages). The significant difference in K-best designs is the sort algorithm. In [3] the bubble sort is used. The authors in [4] exploited the one-cycle merge algorithm. In [5] parallel merge algorithm, in [6] the relaxed distributed sort strategy and in [7] the winner path extension technique which is another sort strategy is used. Although these works include slightly modification to the K-best algorithm, the architecture remained unchanged.

The multi-stage architecture has three drawbacks: first, the architecture is not flexible enough to work with a variable number of antennas. Second, the area increases proportionally with the antenna sizes and results in huge area for large antenna sizes such as 12x12. The last problem is the large area consumption for large constellation sizes like 64-QAM. For example, the 4x4 64-QAM detector using multi-stage architecture consumes a few million gates [6].

V. In-Place Architecture

A. Architecture

One particular benefit of the K-best algorithm is its regularity which has not been exploited efficiently yet. It is possible to implement the algorithm by using one core within a loop and reuse the core in consecutive cycles. The new architecture called “in-place” is shown in Fig. 2-b and its circuitry is shown in Fig. 4. The advantages of this architecture are: flexibility to work with variable number of antennas, significant reduction in area, and providing the basic core for multi-core architectures. This architecture is also a good candidate for large constellation sizes like 64-QAM because it reduces area effectively.

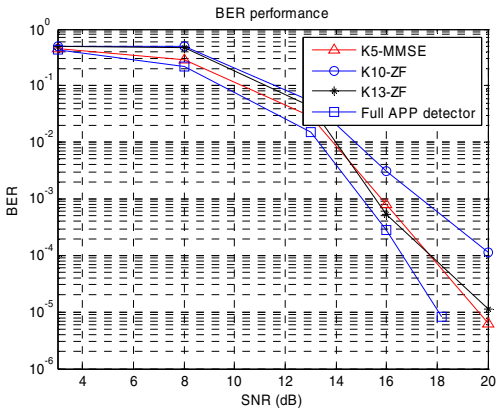

Figure 1. BER comparison for different algorithms.

B. *Reconfigurability*

One of the characteristics of the emerging wireless standards is the switching between different numbers of transmitter and receiver antennas. To address this feature, we used multiplexers and enabled flip-flops along with pre-defined controlling signals stored in the controller to organize the data flow for reconfigurability. Although this work is designed for the 802.11n standard which covers antenna configurations from 2x2 to 4x4, it can support larger numbers of antennas such as 12x12.

C. *Multi-Core Architecture*

The current design projects are defined based on a very small number of antennas while large MIMO systems with larger number of transmit and receive antennas will be of immense interest because of the very high spectral efficiencies possible in such systems. The multi-stage decoder results in a huge area for the large antenna structures because two PEs are needed for each antenna. The detection process for a large number of antennas will be more efficient by parallelizing in-place cores. The in-place architecture has two advantages for multi-core design: It has a small area and makes it easy for the multi-core design. Also each core works independently on separate OFDM subcarriers without performance degradation. Figure 3 shows the multi-core overhead in gray.

The multi-core architecture in [9] uses the depth-first algorithm. This work uses 16 cores and divides the tree to sub-trees and assigns each sub-tree to one core. In the depth-first algorithm the process of each node is highly correlated to the previous layers. In this algorithm, using multiple cores working independently on different sub-trees eliminates the correlation and reduces the BER performance, while in our architecture there is no BER performance degradation.

D. *Area Reduction*

Synthesis results in a commercial 180nm CMOS process show that the in-place architecture provides the throughput of 103Mbps consuming 27Kilo gates (KG) which is about half of the area reported by the smallest sphere decoder in the literature [1] and 75% smaller than a regular K-best design which consumes about 110KG.

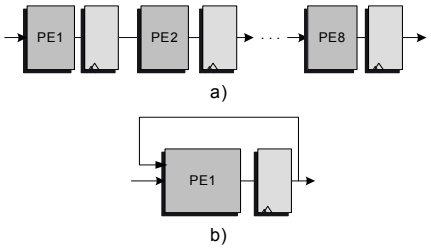

Figure 2. a) Multi-stage architecture b) in-place architecture.

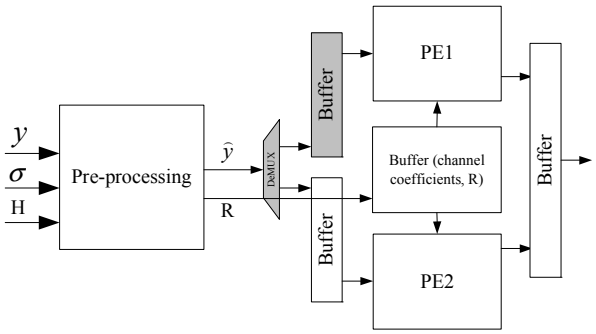

Figure 3. Multi-core structure.

Also in-place architecture adds flexibility when used to implement the multi-core architecture in smaller technologies. For example in 180nm technology process, four in-place cores working in parallel and independently on separate subcarriers of the OFDM frame provide the data rate of 360Mbps which is the maximum throughput required by the 4x4 16-QAM configuration of IEEE 802.11n standard. While in 65nm technology, two cores are enough to provide the throughput. Therefore, the implemented multi-core in-place architecture for the 802.11n standard saves up to 51% of area compared to the multi-stage architecture, in 65nm technology.

VI. In-Place Architecture Realization

There are three processes which happen in all layers of the tree search: calculation of PEDs, finding the *K* smallest PEDs and removing the interference of the current layer from the next layers, as shown in Fig. 4. Another process is performed just once, which calculates the LLRs after finding the candidate list. In the proposed architecture there are four units that implement these functions. The metric computation unit (MCU) calculates the PEDs of each parent. By using real value decomposition (RVD) in the pre-processing unit, the number of stages in the tree will be doubled, and the search process will happen in one dimensional real plain instead of two dimensional complex plain. Therefore, instead of sixteen children (16-QAM) for each parent, four PEDs are generated.

The sort unit finds the K smallest PEDs out of $K.\sqrt{Q}$ PEDs. We replaced the computation expensive sort operation by the merge operation using a circuit level modification: we pre-sorted the PEDs of the parent in MCU in order to be able to use the merge unit. The odd-even merge algorithm is chosen for the merge unit, because of the balanced delay-complexity tradeoff the algorithm provides for hardware implementation.

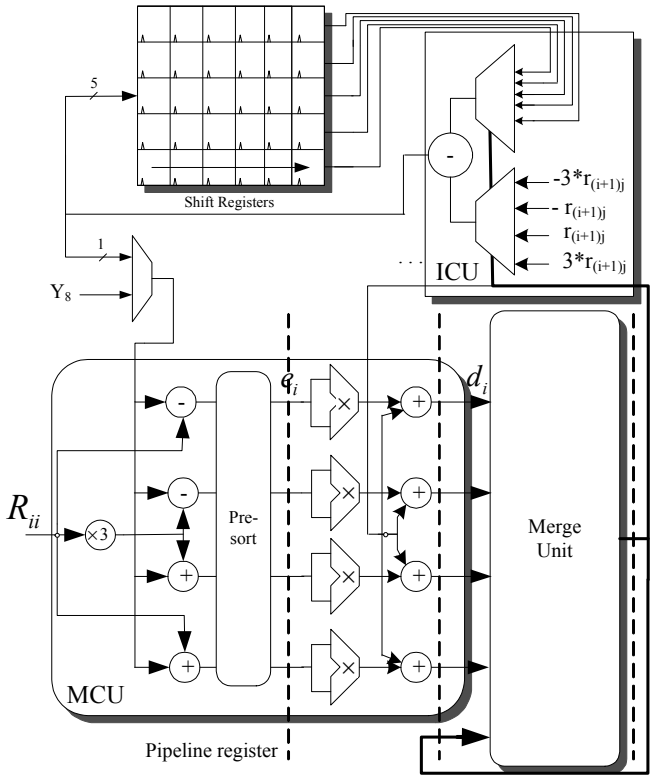

Figure 4. The simplified circuit diagram of the in-place architecture.

Odd-even merge unit provides a smaller area than PMA introduced in [5] and a higher throughput circuit than one-cycle merge introduced in [4]. Increment calculator unit (ICU) removes the interference of the current layer from next layers. Shift registers are used to transfer the computed data from ICU to the next level of the tree. The last unit is the LLR calculator that computes the LLR values for the channel decoder.

VII. Implementation Results

To evaluate the silicon complexity of the circuit accurately, we implemented the soft-output 16-QAM MIMO detector which supports antenna configurations up to 4x4, using a commercial 180nm CMOS technology. The results are generated by the placed and routed simulations including parasitics using Synopsys Design Compiler and Encounter. By using the in-place architecture along with the MMSE-SQRD technique we designed a reconfigurable detector which has the smallest area compared to the other SDs in the literature as shown in Table I.

The measured area as the total equivalent NAND gates is 23Kilo gates (KG) for the sphere decoder and 4KG for the memory required storing the channel coefficients and the vector $\hat{y}$. To have a better comparison we have scaled the throughput of different designs to 180nm technology. This evaluates the design better in terms of throughput. As the table shows, our design outperforms the other designs in area while provide a high data rate.

Table I. Implementation Results for 4*4 16-QAM MIMO Detectors.

Reference	[1]	[10]	[3]	[4]	This work
Technology (µm)	0.25	0.18	0.18	0.25	0.18
Algorithm	DF	DF	K-best	K-best	K-best
Soft-Hard Decision	Soft	Hard	Soft	Hard	Soft
Gate Count (KG)	56.8	175	97	115	27
Throughput (Mbps)	100 (max)	520 (max)	65.5	376	103
Max. clock freq.	71	230	123	117	226
Scaled Throughput	138	520	65.5	522	103

VIII. Conclusion

We proposed the in-place architecture for the K-best algorithm. This architecture exploits the regularity of the K-best algorithm and uses the core cyclically to process the input vector. This architecture provides reconfigurability in terms of antenna configuration, and consumes the smallest area compared to other sphere decoders without sacrificing performance. Also, we exploited the MMSE-SQRD pre-processing technique to achieve a BER performance close to the optimum detector. The design consumes 27KG at throughput of 103Mbps in 180nm CMOS technology.

References

[1] C. Studer, A. Burg and H. Bolcskei , "Soft-output sphere decoding: Algorithms and VLSI implementation," *IEEE Journal on Selected Areas in Communications*, Vol. 26, No. 2, pp. 290-300, Feb. 2008.

[2] C. Hess, M. Wenk, A. Burg, P. Luethi, C. Studer, N. Felber, and W. Fichtner, "Reduced-complexity MIMO detector with close-to ML error rate performance," G*LSVLSI*, pp. 200-203, 2007, Italy.

[3] Z. Guo and P. Nilsson, "Algorithm and implementation of the K-best sphere decoding for MIMO detection", *IEEE Journal on Selected Areas in Communications*, vol.24, no.3, pp. 491- 503, March 2006.

[4] M. Wenk et al.,"K-Best MIMO Detection VLSI architectures achieving up to 424 Mbps," *IEEE ISCAS*, vol. 3, pp. 1151-1154, May 2006.

[5] N. Moezzi-Madani and W. R. Davis, "High-throughput low-complexity MIMO detector based on K-best algorithm," *GLSVLSI'09*, Pages 451-456, boston, May 2009.

[6] S. Chen, T. Zhang, and Y. Xin, "Relaxed K-best MIMO Signal Detector Design and VLSI Implementation," *IEEE Trans. on VLSI (TVLSI)*, vol. 15, issue 3, pp. 328-337, March 2007.

[7] S. Mondal, Ahmed M. Eltawil, and Khaled N. Salama, "Architectural Optimizations for Low-Power K-Best MIMO Decoders," *Accepted to IEEE Transactions on Vehicular Technologies*, Feb. 2009.

[8] H. Kim, D. Lee, and J. D. Villasenor, "Design Tradeoffs and Hardware Architecture for Real-Time Iterative MIMO Detection Using Sphere Decoding and LDPC Coding," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 6, pp. 1003-1014, August 2008.

[9] C. Yang and D. Markovic, "A Multi-Core Sphere Decoder VLSI Architecture for MIMO Communications," *IEEE GLOBECOM'08*, pp. 3297-3301, Dec. 2008.

[10] R. S. Jenkal and W. R. Davis, "An architecture for energy efficient sphere decoding," In Proc. of *ISLPED'07*, pp. 244 - 249, Portland.

[11] D. Wubben, R. Bohnke, V. Kuhn, and K. D. Kammeyer "MMSE Extension of V-BLAST based on Sorted QR Decomposition," in *IEEE Proc. VTCl*, Orlando, Oct 2003.

[12] B. M. Hochwald and S. ten Brink, "Achieving Near-Capacity *on a.* Multiple-Antenna Channel," *IEEE Trans. on Communications,* vol. 51, no. 3, pp. 389-399, March 2003.