FIXED-POINT ARITHMETIC DETECTORS FOR MASSIVE MIMO-OFDM SYSTEMS

Ali J. Al-Askery, Charalampos C. Tsimenidis, and Said Boussakta*

School of Electrical and Electronic Engineering, Newcastle University, NE1 7RU, UK
Email: {a.al-askery, charalampos.tsimenidis, said.boussakta}@newcastle.ac.uk

ABSTRACT

In this paper, the performance of massive multiple input multiple output (MIMO) systems is investigated using reduced detection implementations for MIMO detectors. The motivation for this paper is the need for a reduced complexity detector to be implemented as an optimum massive MIMO detector with low precision. We used different decomposition schemes to build the linear detector based on the (IEEE 754) standard in addition to user-defined precision for selected detectors. Simulations are used to demonstrate the behaviour of several matrix inversion schemes under reduced bit resolution. The numerical results demonstrate improved performance when using QRD and pivoted LDLT decomposition schemes at reduced precision

Index Terms— Massive MIMO, OFDM, Multipath Fading, Low Precision Detection, Fixed Point Representation, Linear Receivers.

1. INTRODUCTION

Massive MIMO systems have become a key technology for future generations of wireless communications. Research in such systems is fuelled by the increased data rate requirements of modern multimedia applications. One of the major challenges in massive MIMO transmission is the increase in the computational complexity at the receiver due to the high number of receiving antennas, especially when using sophisticated non-linear demodulation schemes such as successive interference cancellation and sphere detectors. On the other hand, linear detectors require fewer operations without significantly compromising performance. Past researches [1, 2] have been conducted to show the behaviour of the MIMO system with few antennas at both sides (4x4) using MMSE utilizing a QR decomposition (QRD) detector in a coded system with a hardware implementation. On the other hand, a QRD based Vertical Bell Laboratories Layered Space Time (V-BLAST) detector has been implemented in [3], which takes the fixed point very large scale integration (VLSI) implementation into consideration with 4 antennas at both the transmitter and receiver. In other papers, [4] implemented the sphere detector (SD) without using any decomposition scheme, while [5, 6] used fixed point arithmetic with QRD and Cholesky decomposition. Furthermore, in [7, 8], different techniques have been used with fixed point arithmetic to implement the MIMO system as a field programmable gate array (FPGA) system or as VLSI. Finally, a large scale implementation for the massive MIMO receiver with fixed point representation [9] has used on FPGA system with reduced complexity Neumann series expansion to reduce the implementation complexity.

In this paper, different decomposition schemes are used for detection in a massive MIMO-OFDM system with fixed point arithmetic to simulate the hardware implementation. The standard IEEE 745 double and half precision with word length of 64 and 16 bits will be used in the simulations in addition to a user-defined precision of 12 and 10 bits to verify the ability of each detector.

The remainder of this paper is organized as follows. The system model is presented in Section 2 and the MIMO receivers used are illustrated along with the required algorithms in Section 3. The fixed-point representation is described in Section 4 and the complexity calculations in section 5. The hardware implementation for the best performance detector is presented in Section 6. The simulation results are then shown in Section 7 and conclusions are drawn in Section 8.

2. SYSTEM MODEL

In this paper, the uplink $M \times N$ MIMO-OFDM system depicted in Fig. 1 is considered, where N and M represent the number of transmitting and receiving antennas, respectively, with $M >> N$. After OFDM demodulation, i.e. removing the cyclic prefix (CP) and performing the FFT operation, the received signal can be given as

$$\mathbf{x}_n = \mathbf{H}_n \mathbf{s}_n + \mathbf{W}_n, \tag{1}$$

where $\mathbf{X}_n \in \mathbb{C}^{M\times 1}$ are the received signal samples in frequency domain, $\mathbf{s}_n \in \mathbb{C}^{N\times 1}$ are the transmitted information symbols that are modulated using 16 Quadrature Amplitude Modulation (QAM), $\mathbf{H}_n \in \mathbb{C}^{M\times N}$ is the channel matrix in frequency domain for the n-th FFT sub-carrier, where

*Ali J. Al-Askery is also a staff member with the Foundation of Technical Education in Iraq and sponsored by the MOHESR to complete his PhD.

978-0-9928626-3-3/15/$31.00 ©2015 IEEE

$n = 1, 2, \cdots, N_{FFT}$, and finally, $\mathbf{W}_n \in \mathbb{C}^{M\times 1}$ is the FFT of the additive white Gaussian noise (AWGN) samples in time domain. The detected information symbols are obtained using

$$\hat{\mathbf{s}}_n = \mathbf{H}_n^{\dagger}\mathbf{X}_n, \tag{2}$$

where $\mathbf{H}^{\dagger}$ denotes the pseudoinverse of $\mathbf{H}$ if $M \neq N$ and $\mathbf{H}^{\dagger} = \mathbf{H}^{-1}$ if $M = N$. It is worth noting that the index n will be removed from the subsequent equations to maintain simplicity. We further assume that signals propagate through frequency selective fading channels that are not time selective over the OFDM symbol duration.

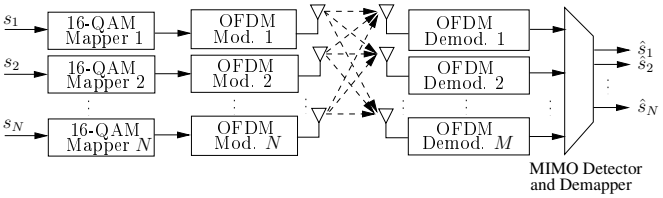

Fig. 1. Massive MIMO-OFDM Transceiver.

3. MIMO DETECTORS

The aim of the MIMO detector is to recover the transmitted symbols, $\hat{\mathbf{s}}$, with the lowest probability of error by utilizing the lowest level of precision in the receiver using different decomposition schemes. This has been motivated by the need to reduce the number of bit representations required in the detection of the massive MIMO system in order to reduce the hardware implementation and power consumption requirements. The number of operations required by these detectors can be very large and their computational complexity cost will be very high if implemented with double or single precision representation.

3.1. Zero Forcing (ZF) Detector

The matrix inversion method used here is the iterative Moore-Penrose pseudo inverse method [10] that has the advantage of reduced complexity detection compared to other types of MIMO detectors. This procedure is illustrated in Algorithm 1 and depends on successive steps to calculate the inverse of rank $N-1$ to the matrix of rank N. This in turn reduces the complexity of calculations as described later in this paper. The general ZF equation can be written as

$$\mathbf{H}_{\mathbf{ZF}}^{-\mathbf{1}} = \left(\mathbf{H}^{\mathbf{H}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathbf{H}}, \tag{3}$$

where the term $(\mathbf{H}^{\mathbf{H}}\mathbf{H})$ represents the Gram matrix, which is a symmetrical positive definite square matrix. Accordingly, Cholesky, LU and LDLT factorization can be used to implement the inverse of these matrices in addition to a Neumann approximation.

Algorithm 1 : Zero Forcing

procedure $A_{inv} = pinv(A)$
  $set\ \ k = 1$
    $A_k = a_k$
    $A_k^{\dagger} = (A_k^H A_k)^{-1} A_k^H$
  for $k \leftarrow 2\ \ to\ \ N$ **do**
    $c_k = (I - A_{k-1}A_{k-1}^{\dagger})a_k$
    $\gamma_k = a_k^H (A_{k-1}^H)^{\dagger} A_{k-1}^{\dagger} a_k$
    $b_k = \begin{cases} c_k^{\dagger}, & \texttt{if}\ \ c_k \neq 0, \\ (1+\gamma)^{-1} a_k^H (A_{k-1}^H)^{\dagger} A_{k-1}^{\dagger}, & \texttt{if}\ \ x = 0. \end{cases}$
    $A_k^{\dagger} = \begin{bmatrix} A_{k-1}^{\dagger} - A_{k-1}^{\dagger} a_k b_k \\ b_k \end{bmatrix}$
  end for
  $A_{inv} \leftarrow A_k^{\dagger}$

3.2. Gram Matrix Based Detector

In this MIMO detector, the LU, Cholesky, and LDLT factorization techniques are used to calculate the Gram matrix inverse. We will use this detector to investigate the effect of these decomposition schemes on the performance of the fixed point MIMO detector. Algorithm 2 was used in the simulation with fixed point design to compare the performance of these detectors. Since all of these detectors involve a triangular matrix inverse, a block matrix inverse procedure (Tri_inv) was used to reduce the complexity of inversion to the one-half of the full matrix inversion complexity.

Algorithm 2 : Gram matrix inverse

procedure $A_{inv} = Gram\_inv(A,\ 'option')$
  if $(LU \leftarrow option)$ **then**
    $\mathbf{A} = \mathbf{LU}$,
    $\mathbf{L_{inv}} = Tri\_inv(\mathbf{L},\ 'Lower')$,
    $\mathbf{U_{inv}} = Tri\_inv(\mathbf{U},\ 'Upper')$,
    $\mathbf{A_{inv}} = \mathbf{U_{inv}}\mathbf{L_{inv}}$,
  else if $(Cholesky \leftarrow option)$ **then**
    $\mathbf{A} = \mathbf{L}\mathbf{L}^{\mathbf{H}}$,
    $\mathbf{L_{inv}} = Tri\_inv(\mathbf{L},\ 'Lower')$,
    $\mathbf{A_{inv}} = (\mathbf{L_{inv}})^{\mathbf{H}}\mathbf{L_{inv}}$,
  else if $(LDLT \leftarrow option)$ **then**
    $\mathbf{A} = \mathbf{PLDL}^{\mathbf{H}}\mathbf{P}^{\mathbf{H}}$,
    $\mathbf{L_{inv}} = Tri\_inv(\mathbf{L},\ 'Lower')$,
    $\mathbf{D_{inv}} = diag(1./diag(\mathbf{D}))$,
    $\mathbf{A_{inv}} = \mathbf{P}^{\mathbf{H}}\mathbf{L}_{\mathbf{inv}}^{\mathbf{H}}\mathbf{D_{inv}}\mathbf{L_{inv}}\mathbf{P}$,
  end if
  Return $\mathbf{A_{inv}}$

3.3. Neumann series expansion

One of the most well-known method that is used to determine the matrix inversion with reduced complexity is the Neumann series expansion [11]. If matrix $\mathbf{A}$ can be written such that $\lim_{n\to\infty}(\mathbf{I}-\mathbf{A})^n = 0$, then it can be decomposed into a diagonal matrix $\mathbf{D}$, which represents the main diagonal of matrix $\mathbf{A}$, and matrix $\mathbf{E} = \mathbf{A} - \mathbf{D}$ with the rest of the matrix $\mathbf{A}$

elements. To illustrate this, matrix $\mathbf{A}$ can be written as

$$\mathbf{A} = \mathbf{D} + \mathbf{E}, \tag{4}$$

$$\mathbf{A}^{-1} = \sum_{n=0}^{\infty} \left(-\mathbf{D}^{-1}\mathbf{E}\right)^n \mathbf{D}^{-1}. \tag{5}$$

The inverse using this method will depend mainly on the choice of n, which represents a factor used to control the complexity of inversion. In previous studies [9, 11, 12], the limit in selecting n was $n = 1, 2$, and 3 with large matrix size. The increase in the receiver diversity in the aforementioned system improves the system performance at low values of n.

3.4. QR Factorization based Detector (SIC)

The QR decomposition is used here as a successive interference cancellation (SIC) with the modified Gram-Schmidt procedure, in which matrix $Q \in \mathbb{C}^{M\times N}$ and the upper triangular matrix $R \in \mathbb{C}^{N\times N}$ [13]. The transmitted signal here can be recovered by using Algorithm 3 which will first multiply the received signal by the Hermitian of the orthonormal matrix Q. Then, the back subsitution procedure reconstructs the transmitted streams completely in N steps.

Algorithm 3 : QR-SIC

1: **procedure** $\hat{\mathbf{s}} = SIC(\mathbf{H}, \mathbf{x})$
2: $\mathbf{QR} \leftarrow \mathbf{H}$
3: $\mathbf{y} = \mathbf{Q^H x} = \mathbf{Rs} + \mathbf{Q^H n}$
4: $$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \cong \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1N} \\ 0 & r_{22} & \cdots & r_{2N} \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & r_{NN} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_N \end{bmatrix},$$
5: $\hat{s}_N = \dfrac{y_N}{r_{NN}},$
6: $\hat{s}_{N-1} = \dfrac{y_{N-1} - r_{N-1N}\hat{s}_N}{r_{N-1N-1}},$
7: $\vdots$
8: $\hat{s}_1 = \dfrac{y_1 - \cdots - r_{1N}\hat{s}_N}{r_{11}}.$
9: **Return** $\hat{\mathbf{s}}$.

4. FIXED POINT REPRESENTATION

The aim of this paper is to show the performance of different decomposition schemes used in MIMO detectors to equalize the channel effects. The fixed point calculations are applied to the output of the FFT of the channel matrix and the received signal in addition to the decomposition schemes above to simulate the behaviour of the implemented MIMO detector.

The standard IEEE 754 precision can be divided according to Fig. 2 into half, single, double and quadruple precision [14]. The first three of the latter are the most popular types and can be represented using

$$X = (-1)^s(1 + \sum_{i=1}^{f} b_{f-i}2^{-i})2^{e-z}. \tag{6}$$

where s is the sign, and e, f and z are the exponent, fraction length and the zero-offset for that number, respectively. The zero-offset equals to $z = 2^{e-1} - 1$, which is 1023 and 15 for double and half precision, respectively. The user-defined precision enables the use of different levels of accuracy depending on the required word length to be used. A word length of 12 bits and 10 bits were used to verify the performance of each detector at a reduced precision detection.

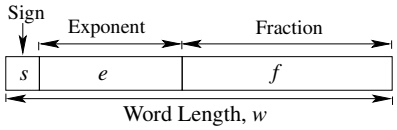

Fig. 2. Numbers representation with fixed point arithmetic

5. COMPLEXITY ANALYSIS

An approximate calculation that depends on *gaxpy* algorithm [15] is introduced here to calculate the complexity required by each MIMO detector. According to this algorithm, the number of operations is a general expression used to identify any mathematical operation. Referring back to the methods of

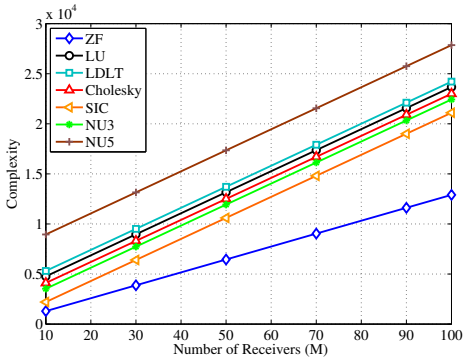

Fig. 3. Complexity calculations required by each method.

MIMO detection illustrated in Section 3, the implementation of the ZF-MIMO detector used in this paper has a complexity of $O(9M + 2MN(N-4))$ operations, which is thus a reduced complexity approach to find the Moore-Penrose matrix inversion compared to the traditional ZF implementation. The MIMO detector based on the Gram matrix inverse were implemented using three different decomposition schemes to compare their individual performance. Firstly, the Cholesky implementation requires in total $O(2N^2M + 2N^3 + NM)$ operations to implement the MIMO detector including calculations of the triangular matrix inverse. The second is the LDLT-based MIMO detector requiring $O(N^2(2M + 2) + N(M + 1) + 3N^3)$ operations to be implemented as it requires a diagonal matrix inverse in addition to the lower triangular inverse. The third Gram matrix-based MIMO detector is LU-factorization, which requires calculating the triangular

Method	$M = 20$	$M = 50$	$M = 100$
ZF	2580	6450	12900
Cholesky	6200	12500	23000
LDLT	7410	13710	24210
LU	6867	13167	23667
QR	4300	10600	21100
Neumann, $n = 3$	5630	11930	22430
Neumann, $n = 5$	11050	17350	27850

Table 1. Operations required by each method at $N = 10$ transmitters and $M = 20$, 50 and 100 receivers.

inverse twice and takes $O(N^2(2M + 16N/6) + NM)$ operations. The QRD has been used as a successive interference cancellation procedure with backward substitution, and this requires $O(N^2(2M + 1) + NM)$ operations. Finally, matrix inversion with the Neumann series expansion needs $O(N(3+M)+2N^2(2+M)+N^3)$ operations, when $n = 3$ and requires $O(5N + 8N^2 + 6N^3 + 2N^2M + NM)$ operations for $n = 5$. The numbers of operations required in the simulations are presented in Table. 1 and Fig. 3 with different receive antennas and at $N = 10$ transmitters.

6. HARDWARE VERIFICATION

The hardware implementation is the final step in every system design to verify the results obtained. Different methods have been used to generate the hardware description language (HDL) code are used in the design of the field-programmable gate arrays (FPGA). The system generator with logical operations is one method that can represent any system as a combination of logic gates. For example, the 2×2 MIMO receiver design using Cholesky inversion and the sphere decoder has been implemented using the system generator method [16]. Another high level design method to generate the HDL code

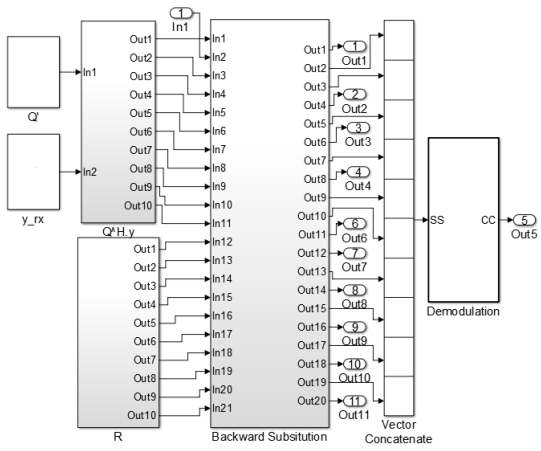

Fig. 4. Top level block diagram of the QRD-SIC massive MIMO receiver.

uses mathematical instead of logical operations such as multiplication, division, subtraction and addition to design the system. The SIC design for the 50×10 MIMO receiver can be seen in Fig. 4 and the expanded version inside one arbitrarily chosen unit of the SIC procedure is presented in Fig. 5.

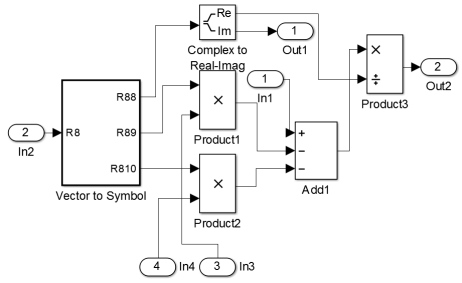

Fig. 5. In depth sub-unit of the back substitution block.

7. SIMULATION AND RESULTS

The performance of the MIMO detectors will be affected by the level of error resulting from calculating the matrix inversion in the methods above with reduced precision. This will result in degradation in the bit error rate (BER) with respect to the signal to noise ratio (SNR) as the precision decreases. The simulations here assume $N = 10$ transmitting and $M = 100$ receiving antennas as shown in Fig. 6 and Fig. 7 to simulate a massive MIMO scenario. Comparing the performance of these detectors at double and single precision will give similar performance since the error resulting from the calculations remains small. As the precision of the calculations decreases to half precision, the performance of the Neumann approximation will no longer be useful in calculating the matrix inversion. The performance of the LU detector will degrade enormously due to the large number of operations required by this detector. In comparison, the other detectors will have no major effect on performance at half precision detection, as illustrated in Fig. 6. To fulfil the requirements of hardware

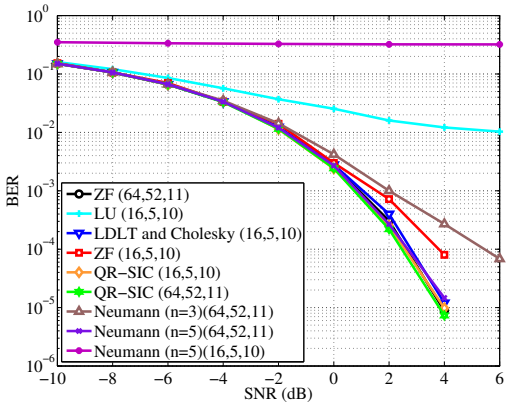

Fig. 6. System performance at double and half precision with $N = 10, M = 100$.

implementation, the calculations of the detector are made at below the standard IEEE 754 representation in order to minimize the required bit representation. In Fig. 7, the performance of the best detectors is presented utilizing user-defined representations with word length $w = 12$ and $w = 10$. According to this simulation, the QRD-SIC detectors exhibit the best performance compared to the other detectors, followed by the LDLT detector. It is worth noting that in the legend, the triplet (w, e, f) indicates the word length, exponent and fraction, respectively.

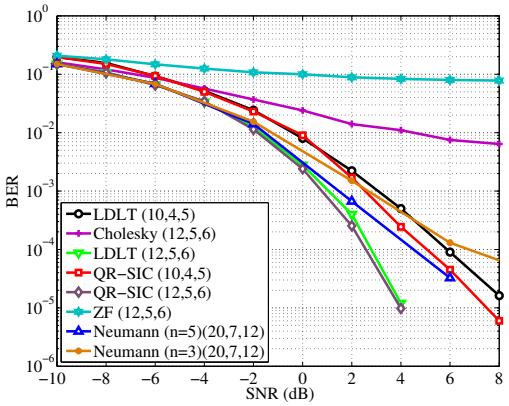

Fig. 7. System performance at a reduced precision with $N = 10, M = 100$.

8. CONCLUSION

This paper has presented a comparison of different linear massive MIMO detectors implemented using matrix decomposition schemes. The simulation results have suggested that at reduced precision with word length less than 12 bits, the performance of QRD and LDLT decompositions outperform that of the other schemes such as Cholesky, LU, and ZF techniques. Furthermore, the complexity calculations of the QRD-SIC detector along with its enhanced performance at fixed point arithmetic have promoted this detector to be used in the hardware implementation of the MIMO detector. Finally, matrix inversion using the Neumann expansion has a limited application in fixed point expression since it shows reduced performance at IEEE 754 half precision and below.

REFERENCES

[1] J. Janhunen, T. Pitkanen, O. Silven, and M. Juntti, "Fixed- and floating-point processor comparison for mimo-ofdm detector," *IEEE J. Sel. Topics in Signal Process.*, vol. 5, no. 8, pp. 1588–1598, Dec 2011.

[2] J. Janhunen, P. Salmela, O. Silven, and M. Juntti, "Fixed- versus floating-point implementation of mimo-ofdm detector," in *Acoustics, Speech and Signal Process. (ICASSP), 2011 IEEE Int. Conf.*, May 2011, pp. 3276–3279.

[3] F. Sobhanmanesh, S. Nooshabadi, and D. Habibi, "A robust qr-based detector for v-blast and its efficient hardware implementation," in *Commun., 2005 Asia-Pacific Conf.*, Oct 2005, pp. 421–424.

[4] N. Tax and B. Lankl, "Fixed effort sphere decoder for mimo ofdm systems," in *Wireless Commun. Signal Process. (WCSP), 2013 Int. Conf.*, Oct 2013, pp. 1–6.

[5] L.M. Davis, "Scaled and decoupled cholesky and qr decompositions with application to spherical mimo detection," in *Wireless Commun. and Network., 2003. WCNC 2003. 2003 IEEE*, March 2003, vol. 1, pp. 326–331 vol.1.

[6] T.E. Schmuland and M.M. Jamali, "Generation of fixed-point vhdl mimo-ofdm qr pre-processor for spherical detectors," in *Circuits and Systems (ISCAS), 2014 IEEE Int. Symp.*, June 2014, pp. 1227–1230.

[7] Matthias Mehlhose and Stefan Schiffermüller, "Efficient fixed-point implementation of linear equalization for cooperative mimo systems," in *EUSIPCO 17th European Signal Process. Conf.*, 2009.

[8] J. Eilert, Di Wu, and D. Liu, "Implementation of a programmable linear mmse detector for mimo-ofdm," in *Acoustics, Speech and Signal Process., 2008. ICASSP 2008. IEEE Int. Conf.*, March 2008, pp. 5396–5399.

[9] M. Wu, Bei Yin, Guohui Wang, C. Dick, J.R. Cavallaro, and C. Studer, "Large-scale mimo detection for 3gpp lte: Algorithms and fpga implementations," *IEEE J. Sel. Topics in Signal Process.,*, vol. 8, no. 5, pp. 916–929, Oct 2014.

[10] Randall E Cline, *Elements of the theory of generalized inverses for matrices*, Springer, 1979.

[11] Lin Bai and Jinho Choi, *Low complexity MIMO detection*, Springer Science & Business Media, 2012.

[12] Jae Kun Lim, "Neumann series expansion of the inverse of a frame operator," *Commun. Korean Math. Soc*, vol. 13, no. 4, pp. 791–800, 1998.

[13] Bei Yin, M. Wu, Guohui Wang, C. Dick, J.R. Cavallaro, and C. Studer, "A 3.8gb/s large-scale mimo detector for 3gpp lte-advanced," in *Acoustics, Speech and Signal Process. (ICASSP), 2014 IEEE Int. Conf.*, May 2014, pp. 3879–3883.

[14] Vinay Ingle and John Proakis, *Digital signal processing using MATLAB*, Cengage Learning, 2011.

[15] Gene H Golub and Charles F Van Loan, *Matrix computations*, vol. 3, JHU Press, 2012.

[16] Mikel Mendicute, Luis G Barbero, Gorka Landaburu, John S Thompson, Jon Altuna, and Vicente Atxa, "Real-time implementation of a sphere decoder-based mimo wireless system," in *European Signal Process. Conf. (EUSIPCO06)*. Citeseer, 2006.