

Analysis of Reduction in Area in MIMO Receivers Using SQRD Method and Unitary Transformation with Maximum Likelihood Estimation (MLE) and Minimum Mean Square Error Estimation (MMSE) Techniques

Sabitha Gauni¹, R.Kumar²

Department of Electronics and Communication, SRM University

¹sabitha.g@ktr.srmuniv.ac.in, ²kumar.r@ktr.srmuniv.ac.in

Abstract—In the field of Wireless Communication, there is always a demand for reliability, improved range and speed. Many wireless networks such as OFDM, CDMA2000, WCDMA etc., provide a solution to this problem when incorporated with Multiple input- multiple output (MIMO) technology. Due to the complexity in signal processing, MIMO is highly expensive in terms of area consumption. In this paper, a method of MIMO receiver design is proposed to reduce the area consumed by the processing elements involved in complex signal processing.

In this paper, a solution for area reduction in the Multiple input multiple output(MIMO) Maximum Likelihood Receiver(MLE) using Sorted QR Decomposition and Unitary transformation method is analyzed. It provides unified approach and also reduces ISI and provides better performance at low cost. The receiver pre-processor architecture based on Minimum Mean Square Error (MMSE) is compared while using Iterative SQRD and Unitary transformation method for vectoring. Unitary transformations are transformations of the matrices which maintain the Hermitian nature of the matrix, and the multiplication and addition relationship between the operators. This helps to reduce the computational complexity significantly. The dynamic range of all variables is tightly bound and the algorithm is well suited for fixed point arithmetic.

Keyword- Spatial Multiplexing, Detection, MIMO, MMSE, MLE, SQRD, Unitary Transformation, Log-likelihood ratio

I. INTRODUCTION

In wireless communications, the use of multiple antennas at both the transmitter and the receiver is a key technology to enable high data rate transmission without additional bandwidth or transmit power. Multiple-input multiple-output (MIMO) schemes are widely used in many wireless standards, allowing higher throughput using spatial multiplexing techniques. MIMO soft detection poses significant challenges to the MIMO receiver design as the detection complexity increases exponentially with the number of antennas. It offers increased data throughput and link range without additional bandwidth or the transmit power. MIMO technology leverages multipath behaviour with an added “spatial” dimension to dramatically increase range and performance in terms of minimization of errors and optimizing of data speed. Major function of MIMO is that it can be sub-divided into three main categories, precoding, spatial multiplexing or SM, and diversity coding.

The MMSE is an equalizer method which aims at minimizing the variance of the difference between the transmitted data and the signal at the equalizer output. QR decomposition is a linear algebraic method in which a matrix A is decomposed into product matrices $A=QR$, where Q is orthogonal matrix and R is the upper triangular matrix. This is a basis for a particular Eigen value algorithm called the QR algorithm. QR decomposition is based on modified Gram Schmidt algorithm. In MIMO systems QR decomposition is used in pre-processors for estimation of pseudo- or non-linear detection methods such as successive interference cancellation or sphere decoding of the channel matrix. SQRD helps in improving bit error rate (BER). SQRD is a highly efficient algorithm as it comes very close to the error performance. MMSE is a method of obtaining minimum mean square error of all the ISI terms and the noise power at the output of the equalizer. Iterative sorted Minimum mean square error QR decomposition further helps in obtaining better terms.

Maximum Likelihood Estimation (MLE) is an equalizer method estimating the parameters of a statistical model. MLE has many optimal properties in estimation: sufficiency (complete information about the parameter of interest contained in its MLE estimator); consistency (true parameter value that generated the data recovered

asymptotically, i.e. for data of sufficiently large samples); efficiency (lowest-possible variance of parameter estimates achieved asymptotically); and parameterization invariance (same MLE solution obtained independent of the parameterization used).

The two methods used in simulation of MIMO receiver:

- A. QR decomposition is a linear algebraic method in which a matrix A is decomposed into product matrices $A=QR$, where Q is orthogonal matrix and R is the upper triangular matrix. This is a basis for a particular Eigen value algorithm called the QR algorithm. QR decomposition is based on modified Gram Schmidt algorithm. In MIMO systems QR decomposition is used in pre-processors for estimation of pseudo- or non-linear detection methods such as successive interference cancellation or sphere decoding of the channel matrix. Sorted QR decomposition helps in improving bit error rate (BER). Sorted QR decomposition is a highly efficient algorithm as it comes very close to the error performance.
- B. Unitary transformation is defined as a transformation that preserves the inner product: the inner product of two vectors before the transformation is equal to their inner product after the transformation. More precisely, a unitary transformation is an isomorphism between two Hilbert spaces. In other words, a unitary transformation is a bijective function

$$U : H_1 \rightarrow H_2 \tag{1}$$

Where H_1 and H_2 are Hilbert spaces.

II. MIMO SYSTEM

The Multiple-Input Multiple-Output (MIMO) systems have been widely considered a viable solution to overcome the current limits in wireless communication. The application of Ultra-wideband (UWB) to indoor environments, with the rich energy scattering and large angular spreads of the multipath channel, provides an ideal scenario for MIMO. By unclosing additional degree of freedom for communication, multiple antennas can effectively turn multipath propagation, considered initially a drawback in wireless communications, into an advantage so as to linearly increase the capacity of the system, or improve its coverage and robustness in terms of error probability.

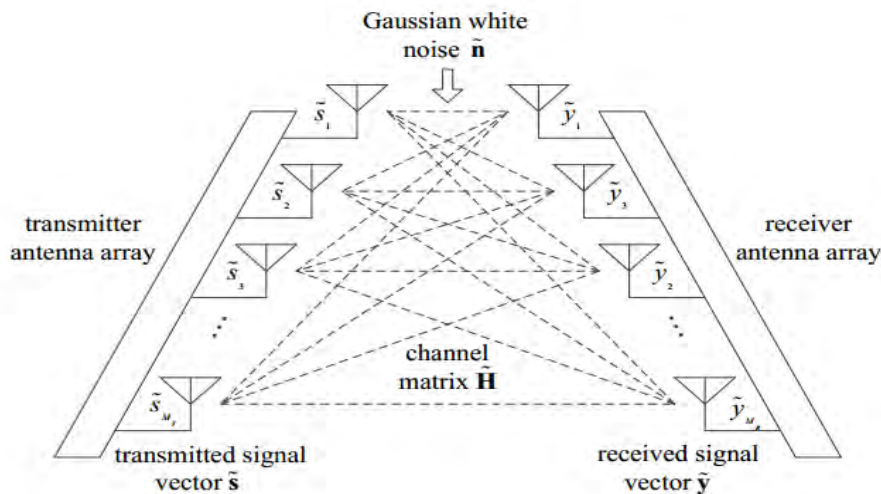


Fig 1: MIMO system model

III. MIMO CHANNEL MODEL

We consider a MIMO channel with M_T transmit and M_R receive antennas. The time-varying channel impulse response between the j th ($j=1,2,\dots,M_T$) transmit antenna and the i th ($i=1,2,\dots,M_R$) receive antenna is denoted as $h_{i,j}(\tau, t)$. This is the response at time t to an impulse applied at time τ . The composite MIMO channel response is given by the matrix with

$$H(\tau, t) = \begin{bmatrix} h_{1,1}(\tau, t) & h_{1,2}(\tau, t) & \dots & h_{1,M_T}(\tau, t) \\ h_{2,1}(\tau, t) & h_{2,2}(\tau, t) & \dots & h_{2,M_T}(\tau, t) \\ \vdots & \vdots & \ddots & \vdots \\ h_{M_R,1}(\tau, t) & h_{M_R,2}(\tau, t) & \dots & h_{M_R,M_T}(\tau, t) \end{bmatrix} \tag{2}$$

The vector is referred to as the spatio-temporal signature induced by the j th transmit antenna across the receive antenna array. Furthermore, given that the signal $s_j(t)$ is launched from the j th transmit antenna, the signal received at the i th receive antenna is given by which can be expressed as follows :-

$$y_i(t) = \sum h_{i,j}(\tau, t) \otimes s_j(t) + n_i(t), i = 1, 2, \dots, M_R \tag{3}$$

IV. MINIMUM MEAN SQUARE ERROR (MMSE)

The MMSE detection criteria reduce the expected value of the difference between the transmitted signal and a linear combination of the received signals [7].

The QR matrix decomposition also known as orthogonal matrix triangularisation is the decomposition of the given matrix A into an orthogonal matrix (Q) and an upper triangular matrix (R). The QR decomposition can be done using any one of the three methods:

1. Gram Schmidt process
2. Householder reflections
3. Givens rotations

The two basic operations for Givens rotations, vectoring and rotation, can both be implemented using conventional arithmetic or dedicated CORDIC circuits [8-9]. CORDICs are a well established method to implement Givens rotations in hardware. In short, the concept of the CORDIC algorithm is to decompose the rotation of a vector into a series of micro rotations by applying shift and add operations. This sequence of shift and add operations is first determined by the vectoring block, and afterwards executed similarly by the rotation block.

A more detailed analysis shows that CORDICs are particularly well suited for the area-efficient implementation of vectoring using iterative decomposition, while fast rotation can be realized more efficient by using conventional complex valued multipliers, but this implies the availability of the corresponding complex-valued rotation coefficients. A solution to this problem is to attach an area-optimized slave CORDIC in rotation mode to the vectoring CORDIC as shown in fig.1

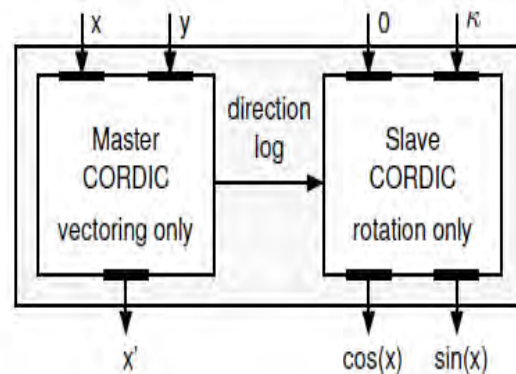


Fig.2: To compute $\cos(x)$ and $\sin(x)$ directly using the Enhanced vectoring CORDIC for subsequent vector rotations using standard multiplications

The input to this slave CORDIC is a unit vector, prescaled by the CORDICs constant scaling factor. The corresponding output values are the coefficients required for the multipliers which carry out the vector rotation.

The overall VLSI architecture of the QR decomposition unit is shown in below Fig. The dedicated vectoring and rotation circuits (using CORDIC and conventional arithmetic, respectively) are extended to handle complex valued matrix entries. The memory which stores the original and intermediate matrices $Z(i)$ is shown as QR Cache and is realized using RAMs with a dedicated read and write port. To satisfy the high memory bandwidth requirements of the rotation block (two read and two write accesses per cycle), the cache is split into two independent memory banks. One bank holds the even rows, the other holds the odd rows of $Z(i)$. Since the rotation block always requires the full memory bandwidth, the vectoring block is fed by a separate FIFO and an additional shadow memory. This solution prevents the rotation block from being stalled by memory access conflicts.

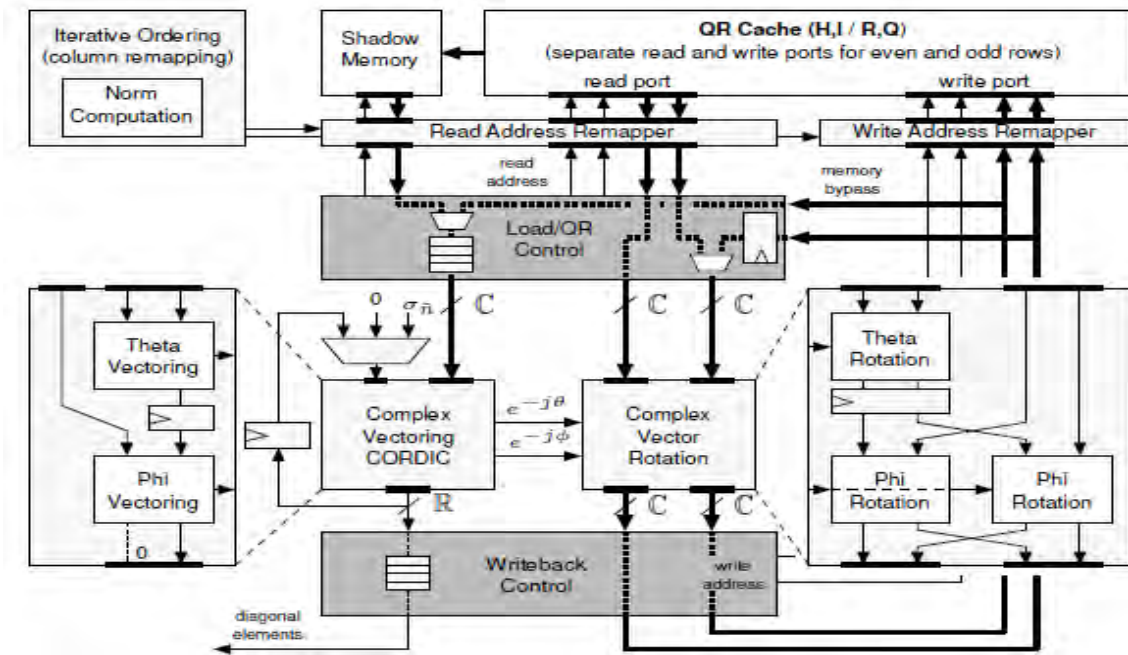


Fig.3: Low latency VLSI architecture for iterative sorted MMSE QR decomposition with super-scalar vectoring and vector rotation.

4.1. Unitary Transformation

Matrix methods that are based upon unitary transformations have been proven to be both useful and important in many application areas including signal processing. They offer stability and are able to handle ill-conditioned problems gracefully [10]. Unitary matrices, they comprise a class of matrices that have the remarkable properties that as transformations they preserve length, and preserve the angle between vectors. For real matrices, unitary is the same as orthogonal. In fact, there are some similarities between orthogonal matrices and unitary matrices. The rows of a unitary matrix are a unitary basis. That is, each row has length one, and their Hermitian inner product is zero. Similarly, the columns are also a unitary basis [11]. In fact, given any unitary basis, the matrix whose rows are that basis is a unitary matrix. It is automatically the case that the columns are another unitary basis. The definition of a unitary matrix guarantees that

$$U^H U = I \tag{4}$$

where I is the identity matrix.

Unitary transformations are transformations of the matrices which maintain the Hermitian nature of the matrix, and the multiplication and addition relationship between the operators. They also maintain the eigen values of the matrix. Due to these advantages benefits of unitary matrices unitary transformation help in processing faster with more accuracy.

4.2. Procedure of Unitary Transformation

Consider a general Hermitean matrix A. This matrix has eigen values a_i and eigen vectors $|A, i\rangle$. Now we want to do some transformation of the matrix A such that the new matrix \tilde{A} is Hermitean and has the same eigen values as A does. This new matrix will have eigenvectors $|\tilde{A}, a_i\rangle$. Let us assume that the transformation is linear i.e, the transformation of a sum is the sum of the transformed vectors. Then it can write the transformation as

$$|\tilde{A}, a_i\rangle = U|A, a_i\rangle \tag{5}$$

where U is some matrix. Now, we want that a unit vector be taken to a unit vector. Thus,

$$\langle \tilde{A}, a_i | \tilde{A}, a_j \rangle = \langle A, a_i | U^\dagger U | A, a_j \rangle \tag{6}$$

But we want

$$\langle \tilde{A}, a_i | \tilde{A}, a_j \rangle = \langle A, a_i | A, a_j \rangle \tag{7}$$

This implies that

$$U^H U = I \tag{8}$$

There is an argument that the requirement that U leave the eigen values of all Hermitean operators the same gives us in addition that

$$U U^H = I \tag{9}$$

Thus, in order to preserve the Hermitian character and the eigen values of an arbitrary matrix A, we need that the transformation be of the form UAU^\dagger and that U be Unitary i.e.,

$$U^\dagger U = 1.$$

Heisenberg's dynamic equations should preserve the eigen values of the matrix A since we are simply looking at the same attribute at different times. Thus, we should have that

$$A(t) = U(t)A_0U^\dagger(t) \tag{10}$$

There is an alternative way of finding the time dependence. Instead of having the matrices change with time, one can have the state change with time.

In the Heisenberg representation, the matrices and in particular their eigenvectors change in time

$$|A(t), a_i\rangle = U(t)|A_0, a_i\rangle \tag{11}$$

The inner product between the state of the system $|\Psi\rangle$ and any eigenvector, which determines the probabilities is given by $\langle A(t), a_i | \Psi \rangle$. However, we get the same amplitude if instead of having the eigenvectors evolve, we have the state evolve.

$$\langle A(t), a_i | \Psi \rangle = \langle U(t)A_0, a_i | \Psi \rangle = \langle A_0, a_i | U(t)^\dagger | \Psi \rangle \tag{12}$$

i.e., all of the amplitudes and probabilities remain the same if, instead of having the operators depend on time, we instead have the state evolve as $U(t)^\dagger | \Psi \rangle$.

The term $U^\dagger H U$ is the Schrodinger Hamiltonian. If H and U commute, which they will do if H is independent of time, then the Heisenberg and Schrodinger Hamiltonian's are the same.

This equation is the Schrodinger for for the evolution of the system. In this case the matrices representing the attributes of the system remain constant, and the state changes with time. Especially for complex systems, this equation is often more easily solved than are the Heisenberg equations of motion.

Note that there is no classical analog for these two approaches. In classical physics, the state, the values which are associated with some attribute of the system are assumed to change in exactly the same way as do the variables which represent those attributes. Thus $x(t) = A \cos(\omega t)$ for a harmonic oscillator means that the attribute which is the position is a function of time in the same way as are the values which are actually ascribed to an attribute.

V. MAXIMUM LIKELIHOOD ESTIMATION

The method of maximum likelihood selects the set of values of the model parameters that maximizes the likelihood function and gives a unified approach to estimation, which is well-defined in the case of the normal distribution and many other problems. For example, one may be interested in the heights of adult female penguins, but be unable to measure the height of every single penguin in a population due to cost or time constraints. Assuming that the heights are normally (Gaussian) distributed with some unknown mean and variance, the mean and variance can be estimated with MLE while only knowing the heights of some sample of the overall population. MLE would accomplish this by taking the mean and variance as parameters and finding particular parametric values that make the observed results the most probable (given the model).

A. The principle of maximum likelihood estimation (MLE), states that the desired probability distribution is the one that makes the observed data “most likely”, which means that one must seek the value of the parameter vector that maximizes the likelihood function.

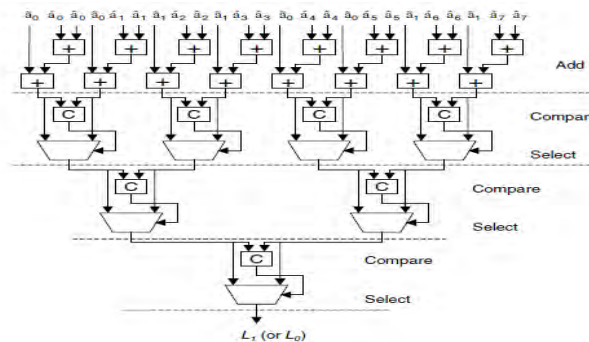


Fig 4: MLE Approach

The MIMO system model we consider here is with M_T transmit and M_R receive antennas. The matrix H describes the MIMO channel. The M_T -dimensional transmit signal vector is denoted by $s = [s_1 \dots s_{M_T}]^T$, and the M_R -dimensional vector n represents the additive zero-mean, Gaussian noise with variance σ^2 per complex dimension. The energy of the transmitted symbol vector is normalized such that $E\{ss^H\} = I_{M_T}$, where I_{M_T}

is $M_T \times M_T$ – dimensional Identity matrix. The corresponding M_R –dimensional receive signal vector is given by

$$y = [y_1 \dots y_{M_R}]^T \text{ is given by } y = Hs + n \tag{13}$$

and the signal to noise ratio is defined as $SNR = M_T/\sigma^2$.

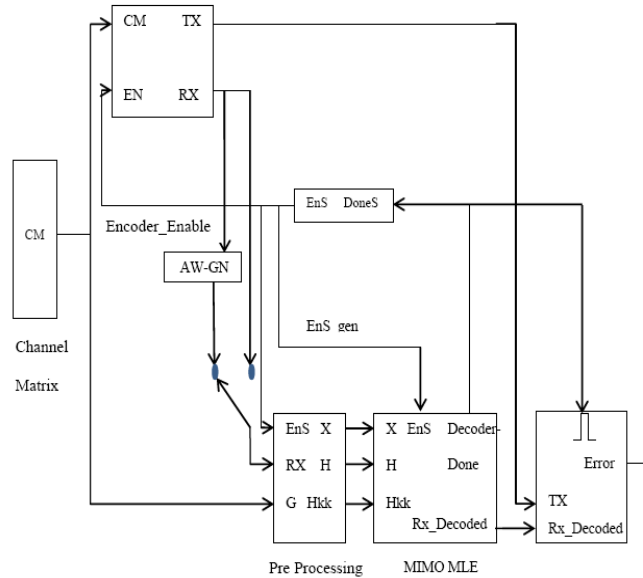


Fig 5: Proposed Model architecture

5.1 MIMO DETECTION BASED ON MLE SQRD

The QRD for MIMO detection starts by decomposing H into a unitary matrix Q and an upper-triangular matrix R with real-valued non-negative elements on the main diagonal. In order to improve the detection performance, the SQRD algorithm efficiently computes $H = QRP^T$ such that the sorting $R_{i,i} \leq R_{j,j}$ for $i < j$ is approximated. The $M_T \times M_T$ -dimensional permutation matrix P^T accounts for the sorting induced by the SQRD algorithm. The basic idea underlying regularized SQRD is to reduce the probability of ill-conditioned channel matrices by computing the SQRD of the matrix.

$$\bar{H} = \begin{bmatrix} H \\ \sigma I_{M_T} \end{bmatrix} = \begin{bmatrix} Q_a \\ Q_b \end{bmatrix} \bar{R} P^T \tag{14}$$

Where $\bar{Q} = [Q_a^T \ Q_b^T]$ and \bar{R} is upper-triangular with real-valued elements on the main diagonal. The dimensions of matrices $\bar{Q}_a, \bar{Q}_b, \bar{R}$ are $M_R \times M_T, M_T \times M_T$, and $M_T \times M_T$ respectively.

5.2 MIMO DETECTION BASED ON MLE UNITARY TRANSFORMATION

Matrix methods that are based upon unitary transformations have been proven to be both useful and important in many application areas including signal processing. They offer stability and are able to handle ill-conditioned problems gracefully. Unitary matrices, they comprise a class of matrices that have the remarkable properties that as transformations they preserve length, and preserve the angle between vectors. For real matrices, unitary is the same as orthogonal. In fact, there are some similarities between orthogonal matrices and unitary matrices. The rows of a unitary matrix are a unitary basis. That is, each row has length one, and their Hermitian inner product is zero. Similarly, the columns are also a unitary basis. In fact, given any unitary basis, the matrix whose rows are that basis is a unitary matrix. It is automatically the case that the columns are another unitary basis.

An important example of these matrices, the rotations, have already been considered. In case of unitary matrices it is been proved for its various properties as

- i. Every set of orthonormal vectors is linearly independent.
- ii. Let $U \in M_n$. The following are equivalent.
- iii. If $U \in M_n$ is unitary, then it is diagonalizable.
- iv. If B and A are unitarily equivalent.

$$\sum_{i,j=1}^n |b_{i,j}|^2 = \sum_{i,j=1}^n |a_{i,j}|^2 \tag{15}$$

Unitary transformations are transformations of the matrices which maintain the Hermitian nature of the matrix, and the multiplication and addition relationship between the operators. They also maintain the eigenvalues of the

matrix. Due to these advantages benefits of unitary matrices unitary transformation help in processing faster with more accuracy.

STEPS FOLLOWED:

1. While receiving input from the Transmitter, Error Component will check possibilities of error using greater component method.
2. Using Error Component we construct the internals of unitary matrix combination. So that Unitary matrix should compute the value using only correct set of inputs.
3. Log-likelihood ratio (LLR) computation for MLE is achieved by instantiating the unitary transform in a parallel manner to get 8 bit inputs at single clock edge.

VI. SIMULATION RESULTS

The two methods have been coded using Verilog hardware descriptive language and simulated using Altera Quartus II for Area Simulation and block synthesis.

6.1 MIMO MMSE SQR DECOMPOSITION

The two MMSE CORDIC methods have been coded using Verilog hardware descriptive language and simulated using ModelSim 6.6 and Altera Simulation using Quartus II for Area simulation and block synthesis. The QR decomposition reference design includes the following key features such as Runtime parameterizable support for QRD and QRD-RLS implementation—uses systolic array, with each cell in the array performing Givens rotations, different input matrix size decompositions, Single matrix decomposition or parallel multiple matrix decompositions, one or more output columns of data—can output the inverse *Q* matrix explicitly.

This also supports several formats such as Fixed-point mode (data in and out) for real or complex data when data is always real, the synthesis time parameter allows optimization of resources and throughput and Floating-point mode. This method gives a highly optimized solution.

TABLE 1
Comparison of Area by the MMSE blocks

List of parameters	MMSE SQRD block	MMSE Unitary transformation block
Top-level entity name	QRD_MMSE	MMSE_Unitary
Family	Cyclone IV GX	Cyclone IV GX
Total logic elements	2,108/21,280 (10%)	694/21,280 (3%)
Total combinational functions	2,080/21,280 (10%)	693/21,280 (3%)
Dedicated Logic registers	136/21,280 (<1%)	137/21,280 (<1%)

6.2.MIMO MLE SQR DECOMPOSITION:

The SQR decomposition reference design includes the following key features:

Runtime parameterizable support for:

- QRD and QRD-RLS implementation—uses systolic array, with each cell in the array performing Givens rotations.
- Different input matrix size decompositions.
- Single matrix decomposition or parallel multiple matrix decompositions.
- One or more output columns of data—can output the inverse *Q* matrix explicitly.

Support for several formats:

- Fixed-point mode (data in and out) for real or complex data when data is always real, the synthesis time parameter allows optimization of resources and throughput.
- Floating-point mode for Time-shared, single processing element, which processes all cells in the systolic array. This gives a highly optimized solution.

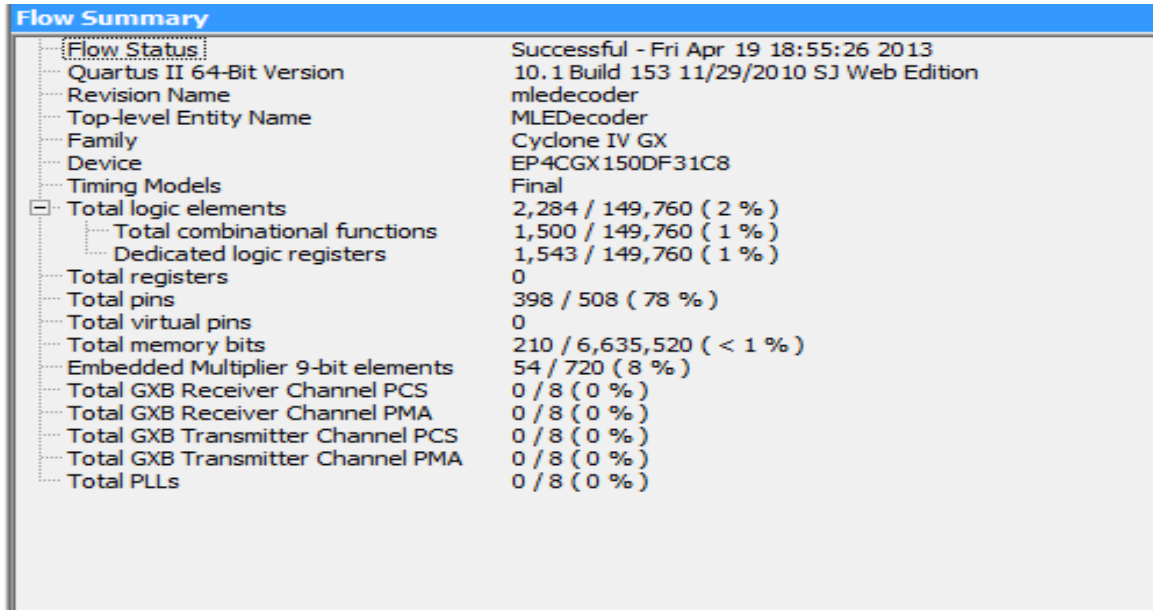


Fig 8: Altera Simulation flow Summary of MLE receiver using Sorted QR Decomposition

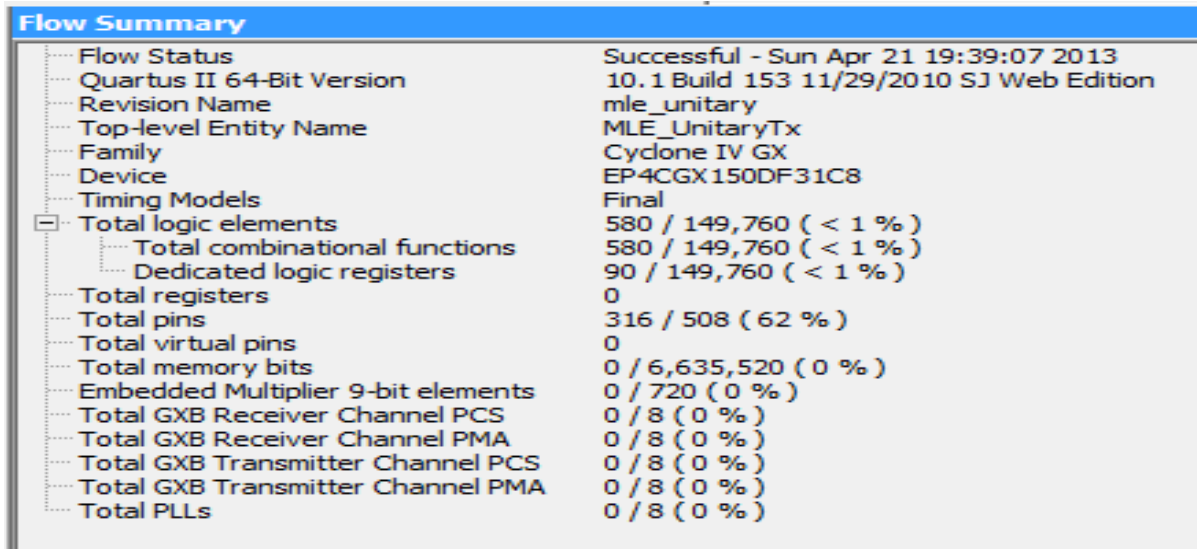


Fig 9: Altera Simulation Flow Summary of MLE receiver using Unitary Transformation

TABLE 2
Comparison of MIMO MLE Receiver using SQRD method and Unitary Transformation

PARAMETERS	MIMO MLE receiver using SQRD method	MIMO MLE receiver using Unitary Transformation
AREA	2284/149760 (2%) logic elements used	580/149760 (<1%) logic elements used
THERMAL POWER DISSIPATION	276.39 mW	244.02 mW
Fmax	69.95MHz	171.29MHz

VII.CONCLUSION

The main objective of an area efficient, high throughput VLSI architecture involves both algorithmic and VLSI implementation aspects. The MMSE SQRD algorithm employs CORDIC circuits to evaluate the matrices rotations using Givens rotations and applied through complex multipliers. This complexity is reduced by using the Unitary transformation to substitute the SQRD in the MIMO MMSE block as the area is reduced almost by 7% as shown in the above results.

Unitary Transformation can be used for SQRD as they have an advantage that vector rotations can be employed as atomic operations which preserve the total power of operations. Hence the dynamic range of all variables is tightly bounded and the algorithm is well suited for fixed point arithmetic. Hence by using Unitary Transformation in place of SQRD in the MIMO MLE block helps in area reduction shown in the above results successfully.

REFERENCES

- [1] Khan, Z.; Arslan, T.; Thompson, J. S.; Erdogan, A.T.; "Analysis and Implementation of Multiple-Input, Multiple-Output VBLAST Receiver from Area and Power Efficiency Perspective", IEEE transaction on VLSI systems, pp. 1281-1286, Dec 2006.
- [2] P. Luethi, A. Brug, S. Haene, D.Perels, N. Felber and W. Fichtner; "VLSI Implementation of a High-Speed Iterative Sorted MMSE QR Decomposition", IEEE international symposium on circuits and systems, pp 1421-1424, 2007.
- [3] G. J. Foschini, "Layered space-time architecture for wireless communication in fading environments when using multiple antennas," Bell Labs Tech. J., vol. 2, Autumn, 1996.
- [4] Adjoudani, E. C. Beck, A. P. Burg, G. M. Djuknic, T. G. Gvoth, D. Haessig, S. Manji, M. A. Milbrodt, M. Rupp, and D. Samardzija, "Prototype experience for MIMO BLAST over third-generation wireless system," IEEE J. Selected Areas Commun., vol. 21, no. 3, pp. 440-451, Apr. 2003.
- [5] D. Garrett, L. Davis, S. Brink, B. Hochwald, and G. Knagge, "Silicon complexity for maximum likelihood MIMO detection using spherical decoding," IEEE J. Solid-State Circuits, vol. 39, no. 9, pp. 1544-1552 Sep. 2004.
- [6] A. Burg, M. Borgmann, M. Wenk, M. Zellweger, W. Fichtner, and H. B'oleskei, "VLSI implementation of MIMO detection using the sphere decoder algorithm," IEEE Journal of Solid-State Circuits, 2005.
- [7] K. Kusume, M. Joham, and W. Utschick, "MMSE Block Decision Feedback Equalizer for Spatial Multiplexing with Reduced Complexity," in Proc. IEEE Global Telecommunications Conference (GLOBECOM 2004), vol. 4, pp. 2540-2544, (Dallas, Texas, USA), November 2004.
- [8] G. H. Golub and C. F. Van Loan, Matrix Computations. John Hopkins Univ. Press, 1996.
- [9] J. Volder, "The CORDIC trigonometric computing technique," IRE Transaction on Electronic Computers, vol. EC-8, no. 3, pp. 330-334, Sept. 1959.
- [10] Z. Guo and P. Nilsson, "A VLSI implementation of MIMO detection for future wireless communications," in Proc. IEEE PIMRC, vol. 3, pp. 2852-2856, 2003.
- [11] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, "VBLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in Proc. IEEE ISSSE, Oct. 1998, pp. 295-300.
- [12] "MMSE Block Decision-Feedback Equalizer for Spatial Multiplexing with Reduced Complexity"
- [13] In Jay Myung, "Tutorial on Maximum Likelihood Estimation".
- [14] Chen Shen, "Design and Implementation of an Improved Soft-output MIMO Detector".