Efficient Implementation of QR Decomposition for Gigabit MIMO-OFDM Systems

Zheng-Yu Huang and Pei-Yun Tsai, *Member, IEEE*

***Abstract*—This paper presents a VLSI architecture of QR decomposition for 4×4 MIMO-OFDM systems. A real-value decomposed MIMO system model is handled and thus the channel matrix to be processed is extended to the size of 8×8. Instead of direct factorization, a QR decomposition scheme by cascading one complex-value and one real-value Givens rotation stages is proposed, which can save 44% hardware complexity. Besides, the requirement of skewed inputs in the conventional QR-decomposition systolic array is eliminated and 36% of delay elements are removed. The real-value Givens rotation stage is also constructed in a form of a stacked triangular systolic array to match with the throughput of the complex-value one. Hardware sharing is considered to enhance the utilization. The proposed design is implemented in 0.18-μm CMOS technology with 152K gates. From measurement, the maximum operating frequency is 100 MHz. It generates QR decomposition results every four clock cycles and accomplishes continuous projection every clock cycle to support MIMO detection up to 2.4 Gb/s. The measured power consumption is 318.6 mW and 219.6 mW for QR decomposition and projection, respectively, at the highest operating frequency. From the comparison, our proposed design achieves the highest throughput with high efficiency.**

***Index Terms*—QR decomposition, Givens rotation, systolic array, MIMO detection.**

I. Introduction

MULTIPLE-INPUT multiple-output (MIMO) techniques have been widely adopted to increase the data transmission rate or to improve the quality of services (QoS) in recent wireless communication systems [1]. Accompanying with the advances in VLSI technology, gigabit wireless transmission is gradually planned and developed such as IEEE 802.11ac/ad (802.11 very high throughput) [2] and IEEE 802.15.3c [3]. Thus, MIMO signal processing plays an important role and attracts much attention in system design, regarding both aspects of performance and implementation [4], [5]. To deal with the multi-dimensional signals, matrix inversion or triangularization is often required. Consequently, QR decomposition is an essential tool to achieve the goal. For example, QR decomposition has been utilized in the precoder of the transmitter to convert one MIMO-OFDM channel into layered subchannels [6]. It is also needed to pre-process the signal to be detected by MIMO sphere decoders [7], [8]. Besides, MIMO signal detection can be accomplished simply by QR decomposition [9]. Hence, to meet the demands of a high transmission rate, a high-throughput QR decomposition module is necessary.

Three known algorithms are widely used for decomposing a matrix into a unitary matrix $\mathbf{Q}$ and an upper triangular matrix $\mathbf{R}$. Gram–Schmidt algorithm obtains the orthogonal basis spanning the column space of the matrix by the orthogonality principle. Through a series of projection, subtraction, norm and division, the column vector of the unitary matrix containing the orthogonal basis can be acquired one by one. Meanwhile, the upper triangular matrix is also generated as a by-product [10]–[12]. Householder transformation (HT) tries to zero out the most elements of each column vector at a stroke by reflection operations. The upper triangular matrix is derived after each transformation matrix being applied to every column vector sequentially. The unitary matrix involves the multiplications of these Householder transformation matrixes and thus the complexity is much higher [13], [14]. On the other hand, Givens rotation (GR) zeros one element of the matrix at a time by two-dimensional rotation [15]–[17]. If an identity matrix is fed as an input, we can get the unitary matrix by using the same rotation sequence when the upper triangular matrix is obtained. In [16], it has been shown that without norm and division operations, Givens rotation is advantageous to perform QR decomposition by the CORDIC algorithm. In addition, with the systolic array architecture, parallelism is easily adopted for a large matrix being processed by Givens rotation [15], [18]. Consequently, Givens rotation is regarded as a possible solution in high-throughput design of QR decomposition.

In this paper, we design and implement a QR decomposition unit for real-value decomposed (RVD) MIMO-OFDM signals. We first analyze the required QR decomposition from the system-design point of view. According to the characteristics of the RVD channel matrix, a QR decomposition scheme that combines a complex-value Givens rotation stage and a real-value Givens rotation stage is proposed which can save 44% hardware complexity compared to the direct triangularization. We also alter the input sequences and thus can improve the requirement of skewed inputs in the conventional complex-value QR decomposition systolic array. As a result, 36% delay elements can be removed. The rate-matching real-value stage is also designed by the triangular systolic array but hardware sharing is adopted to upgrade its utilization. The proposed chip has 152K gate count. From measurement results, the highest operating frequency achieves 100 MHz at 1.8-V supply voltage with power

Manuscript received September 06, 2010; revised November 26, 2010; accepted February 01, 2011. Date of publication April 05, 2011; date of current version September 28, 2011. This work was supported in part by the National Science Council, Taiwan, R.O.C. under Grant NSC98-2220-E-008-001 and Grant NSC99-2220-E-008-007. This paper was recommended by Associate Editor G. Sobelman.

The authors are with the Department of Electrical Engineering, National Central University, Jhong-Li 32001, Taiwan (e-mail: cloud0037@hotmail.com; pytsai@ee.ncu.edu.tw).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCSI.2011.2123770

consumption of 318.6 mW and 219.6 mW for QR decomposition and projection, respectively. The features of the proposed QR decomposition architecture can be summarized as follows.

- The proposed architecture supports QR decomposition of a 4×4 complex channel matrix or an 8×8 RVD channel matrix.
- The proposed architecture can perform QR decomposition every four clock cycles, i.e., 40 ns.
- It also can compute signal projection every one clock cycle. Thus, it achieves high throughput of 2.4 Gb/s for MIMO detection.
- The hardware utilization is enhanced and thus it has good hardware efficiency compared to the conventional design.

This paper is organized as follows. In Section II, we illustrate the requirements of QR decomposition in MIMO-OFDM systems. Section III discusses conventional and proposed QR decomposition schemes. The hardware architecture is delineated in Section IV. The implementation results and comparisons are provided in Section V. Finally is the conclusion.

II. QR Decomposition in MIMO-OFDM Systems

Given a MIMO-OFDM system with N_t transmit antennas, N_r receive antennas, and K subcarriers, assume that the transmitted frequency-domain data at subcarrier k of antenna m is denoted as $X_k^{(m)}$ while the received frequency-domain data at antenna n is represented by $Y_k^{(n)}$. For some subcarrier k, the transmitted frequency-domain signal vector of size $N_t \times 1$ is denoted as $\mathbf{x}_k$ and

$$\mathbf{x}_k = \left[X_k^{(1)} \;\; X_k^{(2)} \;\; \cdots \;\; X_k^{(N_t)}\right]^T.$$

The received frequency-domain signal vector $\mathbf{y}_k$ is then expressed as

$$\begin{aligned}\mathbf{y}_k &= \left[Y_k^{(1)} \;\; Y_k^{(2)} \;\; \cdots \;\; Y_k^{(N_r)}\right]^T \\ &= \mathbf{H}_k\mathbf{x}_k + \mathbf{n}_k \end{aligned} \tag{1}$$

where $\mathbf{n}_k$ represents the noise vector and the complex channel matrix $\mathbf{H}_k$ of size $N_r \times N_t$ describes the channel frequency response at subcarrier k between each transmit and receive antenna pair.

Recently, sphere decoding is popular in MIMO detection. Due to easy enumeration of the possible child nodes without complicated sorting efforts, a real-valued decomposition (RVD) system model is preferred [8], [20]–[22]. The complex signal detection problem in (1) is then reformulated as the real signal detection problem as follows:

$$\begin{aligned}\tilde{\mathbf{y}}_k &= \begin{bmatrix}\mathrm{Re}\{\mathbf{y}_k\} \\ \mathrm{Im}\{\mathbf{y}_k\}\end{bmatrix} \\ &= \begin{bmatrix}\mathrm{Re}\{\mathbf{H}_k\} & -\mathrm{Im}\{\mathbf{H}_k\} \\ \mathrm{Im}\{\mathbf{H}_k\} & \mathrm{Re}\{\mathbf{H}_k\}\end{bmatrix} \cdot \begin{bmatrix}\mathrm{Re}\{\mathbf{x}_k\} \\ \mathrm{Im}\{\mathbf{x}_k\}\end{bmatrix} \\ &\quad + \begin{bmatrix}\mathrm{Re}\{\mathbf{n}_k\} \\ \mathrm{Im}\{\mathbf{n}_k\}\end{bmatrix} \\ &= \tilde{\mathbf{H}}_k\tilde{\mathbf{x}}_k + \tilde{\mathbf{n}}_k, \end{aligned} \tag{2}$$

where $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$ denote the real part and the imaginary part of their respective argument. The size of the real channel matrix $\tilde{\mathbf{H}}_k$ becomes $2N_r \times 2N_t$. The optimal solution to the MIMO detection is to find the symbol $\tilde{\mathbf{x}}_k$ that can minimize the Euclidean distance between received signal vector $\tilde{\mathbf{y}}_k$ and $\tilde{\mathbf{H}}_k\tilde{\mathbf{x}}_k$. By using QR decomposition of the channel matrix $\tilde{\mathbf{H}}_k$, i.e., $\tilde{\mathbf{H}}_k = \tilde{\mathbf{Q}}_k\tilde{\mathbf{R}}_k$ with a unitary matrix $\tilde{\mathbf{Q}}_k$ and an upper-triangular matrix $\tilde{\mathbf{R}}_k$, the optimal detection becomes

$$\begin{aligned}\hat{\mathbf{x}}_k &= \arg\min_{\tilde{\mathbf{x}}_k \in \Omega} \|\tilde{\mathbf{y}}_k - \tilde{\mathbf{H}}_k\tilde{\mathbf{x}}_k\| \\ &= \arg\min_{\tilde{\mathbf{x}}_k \in \Omega} \|\tilde{\mathbf{Q}}_k^H\tilde{\mathbf{y}}_k - \tilde{\mathbf{R}}_k\tilde{\mathbf{x}}_k\| \end{aligned} \tag{3}$$

where Ω is a set of the search space. In order to detect every transmitted signal $\tilde{\mathbf{x}}_k$, we need to obtain $\tilde{\mathbf{R}}_k$ and $\tilde{\mathbf{Q}}_k^H\tilde{\mathbf{y}}_k$ for every received signal, respectively. Note that for MIMO-OFDM systems operated in stationary environments, the channel matrix remains almost the same. Thus, QR decomposition of the channel matrix can be done only once to get matrix $\tilde{\mathbf{R}}_k$. On the other hand, the calculation of $\tilde{\mathbf{Q}}_k^H\tilde{\mathbf{y}}_k$ must be updated for every incoming signal $\tilde{\mathbf{y}}_k$. The MIMO detection throughput is highly related with the throughput of vector projection, namely $\tilde{\mathbf{Q}}_k^H\tilde{\mathbf{y}}_k$.

The complexity of the algorithms for QR decomposition has been shown in [16]. However, from (3), it is clear that instead of the unitary matrix $\tilde{\mathbf{Q}}_k$, the product of $\tilde{\mathbf{Q}}_k^H$ and $\tilde{\mathbf{y}}_k$ is what we want for MIMO detection. In the following, we will examine the complexity of obtaining $\tilde{\mathbf{Q}}_k^H\tilde{\mathbf{y}}_k$ and $\tilde{\mathbf{R}}_k$ from received signal $\tilde{\mathbf{y}}_k$ and channel matrix $\tilde{\mathbf{H}}_k$ by different QR decomposition algorithms.

III. Algorithms and Complexity

To select a proper algorithm for hardware implementation, we will first examine the arithmetic complexity of the required operations. Then, we will describe our proposed scheme for further complexity reduction.

A. QR Decomposition Algorithms

Generally given an $N \times N$ real or complex matrix $\mathbf{A}$, to decompose the matrix $\mathbf{A}$ into an upper triangular matrix $\mathbf{R}$, the Gram–Schmidt algorithm projects the ith column vector $\mathbf{a}_i$ of matrix $\mathbf{A}$ onto the subspace spanned by $\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}$ to obtain the element $r_{j,i}$ of $\mathbf{R}$ for $1 \le j < i$. Then, the unit column vector $\mathbf{q}_i$ of the unitary matrix $\mathbf{Q}$ is derived from normalizing the new orthogonal basis $\mathbf{u}_i$ that is formed by removing these linearly dependent components from $\mathbf{a}_i$. The pseudo codes of the Gram–Schmidt algorithm for processing both real and complex matrix $\mathbf{A}$ are given in Table I. For an $N \times 1$ vector $\mathbf{b}$, to perform $\mathbf{z} = \mathbf{Q}^H\mathbf{b}$, additional matrix multiplication is needed. The complexity corresponding to each step is evaluated by its arithmetic operations and is also given in the table. For example, the number of multiplications for $N \times N$ real and complex matrix decomposition can be estimated, respectively, as

$$\sum_{i=1}^{N}\left[\left(\sum_{j=1}^{i-1} 2N\right) + 2N\right] = N^2(N+1) \tag{4}$$

and

$$\sum_{i=1}^{N}\left[\left(\sum_{j=1}^{i-1} 8N\right) + 6N\right] = 2N^2(2N+1). \tag{5}$$

TABLE I
GRAM–SCHMIDT ALGORITHM

Real matrix	Add.	Mul.	Div.	Sqrt.
for i=1:N				
$\mathbf{u}_i = \mathbf{a}_i$				
for j=1:i-1				
$r_{j,i} = \mathbf{q}_j^T \mathbf{a}_i$	$N-1$	N		
$\mathbf{u}_i = \mathbf{u}_i - r_{j,i}\mathbf{q}_j$	N	N		
end				
$\cdot\ r_{i,i} = \|\mathbf{u}_i\|$	$N-1$	N		1
$\mathbf{q}_i = \mathbf{u}_i / r_{i,i}$			N	
$\cdot\ z_i = \mathbf{q}_i^T \mathbf{b}$	$N-1$	N		
end				
Complex matrix	**Add.**	**Mul.**	**Div.**	**Sqrt.**
for i=1:N				
$\mathbf{u}_i = \mathbf{a}_i$				
for j=1:i-1				
$r_{j,i} = \mathbf{q}_j^H \mathbf{a}_i$	$4N-2$	$4N$		
$\mathbf{u}_i = \mathbf{u}_i - r_{j,i}\mathbf{q}_j$	$4N$	$4N$		
end				
$\cdot\ r_{i,i} = \|\mathbf{u}_i\|$	$2N-1$	$2N$		1
$\mathbf{q}_i = \mathbf{u}_i / r_{i,i}$			$2N$	
$\cdot\ z_i = \mathbf{q}_i^H \mathbf{b}$	$4N-2$	$4N$		
end				

TABLE II
HOUSEHOLDER TRANSFORMATION

Real matrix	Add.	Mul.	Div.	Sqrt.
$\mathbf{R} = \mathbf{A}$, $\mathbf{z} = \mathbf{b}$				
for i=1:N-1				
$\alpha = \text{sign}(\mathbf{R}_{i,i})\|\mathbf{R}_{i:N,i}\|$	$\gamma - 1$	γ		1
$\mathbf{v} = \mathbf{R}_{i:N,i} - \alpha \mathbf{e}_1$	1			
$\mathbf{w} = \mathbf{v}\mathbf{v}^T / (\mathbf{v}^T \mathbf{v})$	$\gamma - 1$	$\gamma^2 + \gamma$	γ^2	
$\cdot\ \mathbf{R}_{i:N,i:N} = \mathbf{R}_{i:N,i:N} - 2\mathbf{w}\mathbf{R}_{i:N,i:N}$	γ^3	γ^3		
$\cdot\ \mathbf{z}_{i:N} = \mathbf{z}_{i:N} - 2\mathbf{w}\mathbf{z}_{i:N}$	γ^2	γ^2		
end				
Complex matrix	**Add.**	**Mul.**	**Div.**	**Sqrt.**
$\mathbf{R} = \mathbf{A}$, $\mathbf{z} = \mathbf{b}$				
for i=1:N-1				
$\alpha = \|\mathbf{R}_{i:N,i}\|$	$2\gamma - 1$	2γ		1
$\mathbf{v} = \mathbf{R}_{i:N,i} - \alpha \mathbf{e}_1$	1			
$\mathbf{w} = \mathbf{v}\mathbf{v}^H / \mathbf{v}^H \mathbf{R}_{i:N,i}$	$5\gamma^2 + 4\gamma - 2$	$10\gamma^2 + 4\gamma$	$2\gamma^2$	
$\cdot\ \mathbf{R}_{i:N,i:N} = \mathbf{R}_{i:N,i:N} - \mathbf{w}\mathbf{R}_{i:N,i:N}$	$4\gamma^3$	$4\gamma^3$		
$\cdot\ \mathbf{z}_{i:N} = \mathbf{z}_{i:N} - \mathbf{w}\mathbf{z}_{i:N}$	$4\gamma^2$	$4\gamma^2$		
end				

Although there exists the modified Gram–Schmidt algorithm for numerical stability consideration, the arithmetic complexity is basically the same.

The Householder transformation reflects a vector cross a hyperplane, called Householder plane, to obtain a mirror vector $\alpha\mathbf{e}_1$ where $\mathbf{e}_1$ is a unit vector with one 1's in the first entry and α is related with the norm of the original vector as listed in Table II. The vector $\mathbf{v}$ denotes the difference between the vector to be processed and the mirror vector. The subscript $(i:N)$ indicates the ith entry to the Nth entry in the vector or matrix. Unlike the Gram–Schmidt algorithm, it is not necessary to calculate the matrix $\mathbf{Q}$ explicitly in order to obtain $\mathbf{z} = \mathbf{Q}^H\mathbf{b}$. The complexity is evaluated in terms of $\gamma = N - i + 1$.

The Givens rotation tries to eliminate the undesired elements one by one to get an upper triangular matrix. Consider a 4×4 complex matrix:

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} \end{bmatrix}$$

The Givens rotation matrix $\mathbf{G}$ which intends to cancel $a_{3,1}$ by $a_{1,1}$ can be given by

$$\mathbf{G} = \begin{bmatrix} c & 0 & s & 0 \\ 0 & 1 & 0 & 0 \\ -s^H & 0 & c^H & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where

$$c = \frac{a_{1,1}^H}{\sqrt{|a_{1,1}|^2 + |a_{3,1}|^2}}, s = \frac{a_{3,1}^H}{\sqrt{|a_{1,1}|^2 + |a_{3,1}|^2}}.$$

Thus,

$$\mathbf{GA} = \begin{bmatrix} a_{1,1}^{(1)} & a_{1,2}^{(1)} & a_{1,3}^{(1)} & a_{1,4}^{(1)} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ 0 & a_{3,2}^{(1)} & a_{3,3}^{(1)} & a_{3,4}^{(1)} \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} \end{bmatrix} \quad (6)$$

where the superscript represents the count of rotations and $a_{1,1}^{(1)}$ becomes a real number. The operations of Givens rotation to derive the upper triangular matrix $\mathbf{R}$ are also summarized in Table III. Obviously, the same strategy can be employed to get vector $\mathbf{z} = \mathbf{Q}^H\mathbf{b}$.

The Givens rotation can also be done by the famous coordinate rotation digital computer (CORDIC) algorithm. By using polar representation, if $a_{1,1} = |a_{1,1}|e^{j\theta_x}$ and $a_{3,1} = |a_{3,1}|e^{j\theta_y}$, then

$$c = \frac{a_{1,1}^H}{\sqrt{|a_{1,1}|^2 + |a_{3,1}|^2}} = \cos(\theta_z)e^{-j\theta_x}$$
$$s = \frac{a_{3,1}^H}{\sqrt{|a_{1,1}|^2 + |a_{3,1}|^2}} = \sin(\theta_z)e^{-j\theta_y} \quad (7)$$

where $\tan(\theta_z) = |a_{3,1}|/|a_{1,1}|$. When both inputs are real numbers, i.e. $c = \cos(\theta_z)$ and $s = \sin(\theta_z)$, the real Givens rotation can be done by one CORDIC operation.

The architecture of the complex Givens rotation can be implemented as shown in Fig. 1 [15]. To gain further insight, we discuss its operations in details. In vectoring mode, the lower element in the leading pair is intended to be vanished. Two complex inputs are compensated for $e^{-j\theta_x}$ and $e^{-j\theta_y}$ by the left two CORDIC modules and their imaginary-part outputs then both become zeros. Hence, only three CORDIC operations are required and the lower-right CORDIC module is useless. In addition, the micro-rotation sequences related with θ_x, θ_y and θ_z are saved. On the other hand, in the rotation mode for the subsequent data, four CORDIC operations are necessary and the micro-rotation sequences are retrieved from the storages. If the upper element of the leading pair is a real number, only two and three CORDIC operations are required in vectoring mode and in rotation mode, respectively, because the upper left CORDIC can be bypassed. Note that in the complex Givens rotation algorithm in Table III, the last entry $a_{N,N}$ is not processed in the FOR loop. Thus, in order to make the diagonal terms of the matrix $\mathbf{R}$ to become all real, additional rotation is necessary. In

TABLE III
GIVENS ROTATION

Real matrix	Add.	Mul.	Div.	Sqrt.	CORDIC operations
$\mathbf{R}=\mathbf{A},\mathbf{z}=\mathbf{b}$					
for i=1:N-1					
for $j=i+1:N$					
$\alpha=\|\|[\mathbf{R}_{i,i}\ \mathbf{R}_{j,i}]\|\|$	1	2		1	
$c=\mathbf{R}_{i,i}/\alpha$			1		
$s=\mathbf{R}_{j,i}/\alpha$			1		
$\cdot\ \mathbf{R}_{i,i:N}=c\mathbf{R}_{i,i:N}+s\mathbf{R}_{j,i:N}$	$(N-i+1)$	$2(N-i+1)$			$N-i+1$
$\cdot\ \mathbf{R}_{j,i:N}=-s\mathbf{R}_{i,i:N}+c\mathbf{R}_{j,i:N}$	$(N-i+1)$	$2(N-i+1)$			
$\cdot\ z_i=cz_i+sz_j$	1	2			1
$\cdot\ z_j=-sz_i+cz_j$	1	2			
end					
end					
Complex matrix	**Add.**	**Mul.**	**Div.**	**Sqrt.**	**CORDIC operations**
$\mathbf{R}=\mathbf{A},\mathbf{z}=\mathbf{b}$					
for i=1:N-1					
for $j=i+1:N$					
$\alpha=\|\|[\mathbf{R}_{i,i}\ \mathbf{R}_{j,i}]\|\|$	3	4		1	
$c=\mathbf{R}_{i,i}^H/\alpha$			2		
$s=\mathbf{R}_{j,i}^H/\alpha$			2		
$\cdot\ \mathbf{R}_{i,i:N}=c\mathbf{R}_{i,i:N}+s\mathbf{R}_{j,i:N}$	$6(N-i+1)$	$8(N-i+1)$			$4(N-i)+3$ or
$\cdot\ \mathbf{R}_{j,i:N}=-s^H\mathbf{R}_{i,i:N}+c^H\mathbf{R}_{j,i:N}$	$6(N-i+1)$	$8(N-i+1)$			$3(N-i)+2^\dagger$
$\cdot\ z_i=cz_i+sz_j$	6	8			4 or $3^\dagger$
$\cdot\ z_j=-s^Hz_i+c^Hz_j$	6	8			
end					
end					
$\cdot\ \mathbf{R}_{N,N}=\mathbf{R}_{N,N}\mathbf{R}_{N,N}^H/\|\mathbf{R}_{N,N}\|$	1	2		1	1
$\cdot\ z_N=z_N\mathbf{R}_{N,N}^H/\|\mathbf{R}_{N,N}\|$	2	4	2		1

† if $R_{i,i}$ is real.

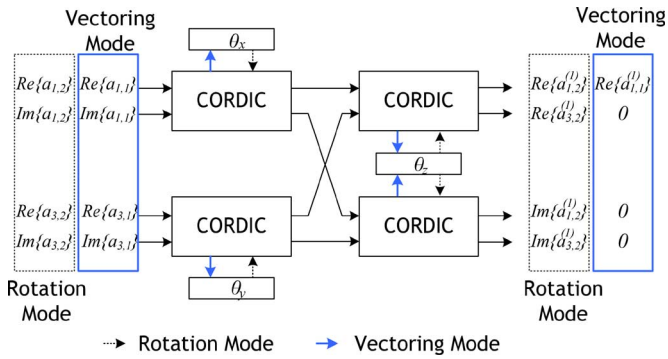

Fig. 1. Architecture of complex Givens rotation.

Table III, we provide the complexity evaluation of Givens rotation algorithm both in arithmetic operations and in CORDIC operations.

As the antenna configurations in advanced MIMO-OFDM systems may vary from 2×2 to 8×8 [19], we then evaluated the complexity of decomposing 2×2, 4×4 and 8×8 complex channel matrixes and 4×4, 8 × 8 and 16 × 16 RVD channel matrixes as shown in Fig. 2. Fig. 2(a), (b), and (c) show the complexity in arithmetic operations of three algorithms. Fig. 2(d) indicates the number of CORDIC operations of Givens rotation algorithm for decomposing the channel matrix under the same antenna configurations. From the figure, it is clear that as the size of the matrix grows, the complexity also increases rapidly. For the same antenna configuration, it takes more efforts to process the RVD channel matrix than to decompose the complex channel matrix. Moreover, the Householder transformation is hardware consuming compared to the others. For the Gram–Schmidt algorithm, a lot of multiplications, divisions and square root operations are needed when a large-size matrix is decomposed. If Givens rotation algorithm is implemented by the arithmetic operations such as additions and multiplications, no obvious saving in complexity is shown. However, CORDIC-based Givens rotation algorithm with only shifters and adders avoids the demands of a large number of complicated arithmetic units. Furthermore, it also has the advantages in two aspects. On one hand, both matrix $\mathbf{R}$ and vector $\mathbf{z}$ can be obtained by using the same hardware, which is good for hardware efficiency. On the other hand, the systolic array with parallelism facilitates the throughput enhancement. Thus, CORDIC-based Givens rotation algorithm is employed in our design. In the following, we will show our additional improvement in QR decomposition of an RVD channel matrix.

B. Proposed Decomposition Scheme

From the previous discussion, it is clear that direct decomposition of an RVD channel matrix is quite complicated compared to decomposition of a complex channel matrix. In [16], the symmetry property of four sub-matrixes of the RVD channel matrix in (2) is kept so as to reduce the complexity. Here, we adopt a two-stage decomposition scheme by exploiting the dependency of the real-value channel matrix and the complex-value channel matrix as much as possible at the first stage. A second stage is then designed to derive the upper triangular matrix with minimum CORDIC operations [23], and thus hardware cost can be saved.

As shown in Fig. 3, instead of RVD done before triangularization, the $N_t \times N_t$ complex matrix $\mathbf{H}_k$ is first converted into an upper triangular matrix $\mathbf{P}_k$ by the complex Givens rotation

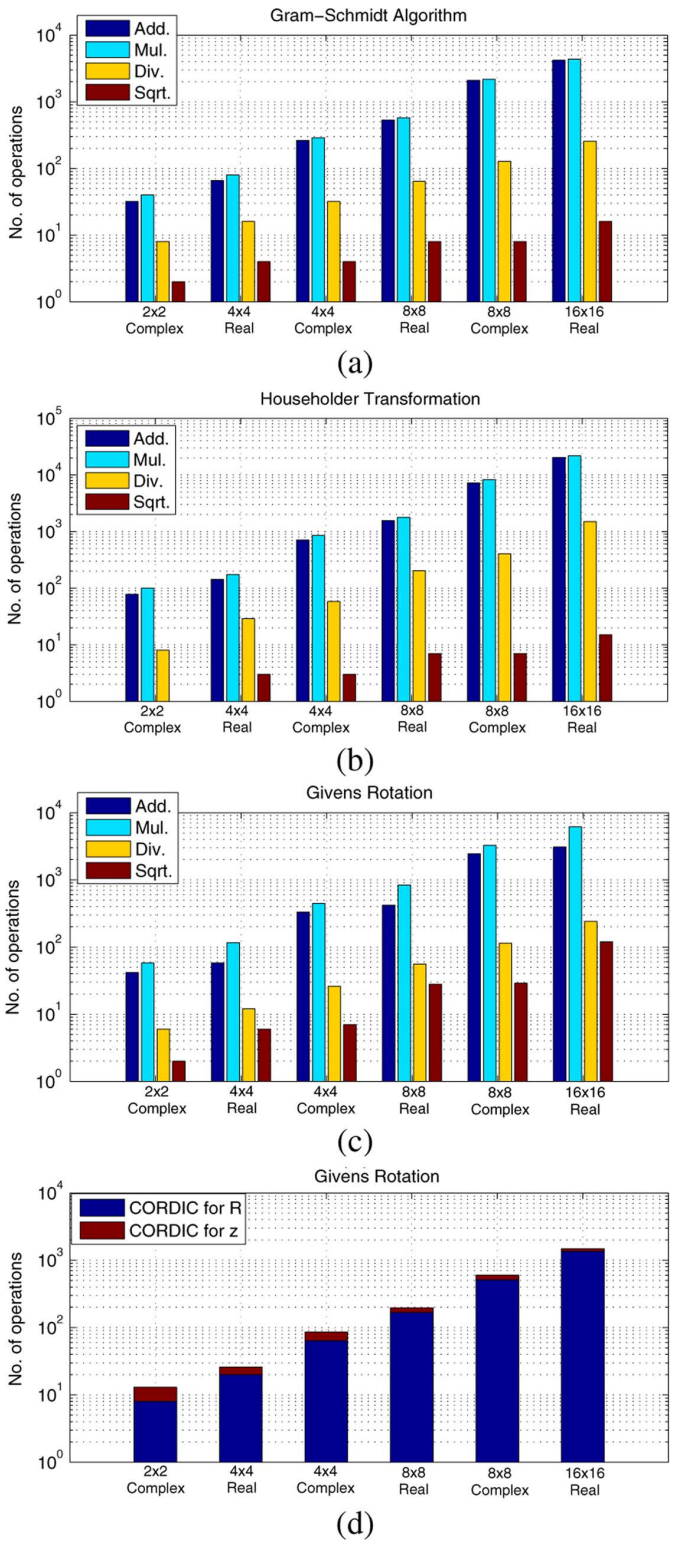

Fig. 2. Complexity evaluation of decomposing complex or real channel matrixes of various sizes by (a) Gram–Schmidt algorithm, (b) Householder transformation, (c) Givens rotation with arithmetic operations, and (d) Givens rotation with CORDIC operations.

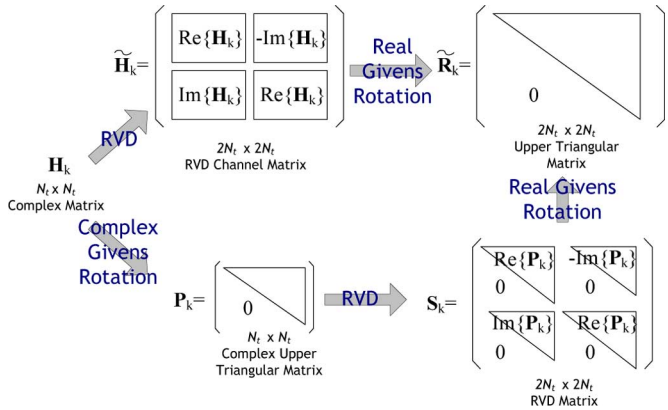

Fig. 3. Proposed decomposition scheme.

algorithm. Note that the diagonal terms of matrix $\mathbf{P}_k$ become real numbers, but its off-diagonal terms are still complex numbers. Thereafter, $\mathbf{P}_k$ is expanded to a real matrix $\mathbf{S}_k$ of size $2N_t \times 2N_t$. Finally, the non-zero elements in the lower-left corner of $\mathbf{S}_k$ are eliminated by the real Givens rotation algorithm.

To guarantee the minimum operations, one constraint should be set that extra non-zero term can not be generated in the lower right corner during the annihilation process of the non-zero terms in the lower left corner. Consequently, we must pay attention to the processing sequence. From Table III, it is clear that if two elements to be processed are both zero, the results after Givens rotation will be still zeros. Hence, in matrix $\mathbf{S}_k$, we must select the row i for $1 \le i \le N_t$ to cancel the undesired terms in row j for $N_t + 1 \le j \le 2N_t$, where the number of 0's in row i should be greater than or equal to the number of 0's in row j so as to retain these null elements in row j. For instance, with $N_t = 4$ in Fig. 4, $s_{2,2}$, $s_{3,3}$ and $s_{4,4}$ in row 2, 3, 4 are used to cancel $s_{5,2}$, $s_{6,3}$, and $s_{7,4}$ in row 5, 6, 7 in the first round. Row 3 and row 4 subsequently are utilized to eliminate the (5,3)th and (6,4)th entries. Finally, we zero the (5,4)th entry by using row 4.

The complexity reduction is then evaluated. From Table III, for direct triangularization of a $2N_t \times 2N_t$ real matrix, the required CORDIC operations are

$$\sum_{i=1}^{2N_t-1} \sum_{j=i+1}^{2N_t} (N_t - i + 1) = \frac{(2N_t)^3 - (2N_t)}{3}. \quad (8)$$

With the proposed scheme, decomposition a complex $N_t \times N_t$ matrix needs CORDIC operations given by

$$\sum_{i=1}^{N_t-1} \{3 + 4(N_t - i) + \sum_{j=i+2}^{N_t} (2 + 3(N_t - i))\} + 1 = N_t^3. \quad (9)$$

Thereafter, the real Givens rotation for $2N_t \times 2N_t$ matrix $\mathbf{S}_k$ requests CORDIC operations as follows:

$$\sum_{i=1}^{N_t-1} \sum_{j=i+1}^{N_t} [(N_t - j + 1) + (N_t - (j - i) + 1)] = \frac{N_t^3 - N_t}{2} \quad (10)$$

where the first term and the second term are related with the possible rotations incurred by the non-zero terms in the upper left corner and lower right corner, respectively. In Fig. 5, we show the reduction of our proposed scheme. As N_t increases, namely more antennas are allocated in both the transmitter and receiver,

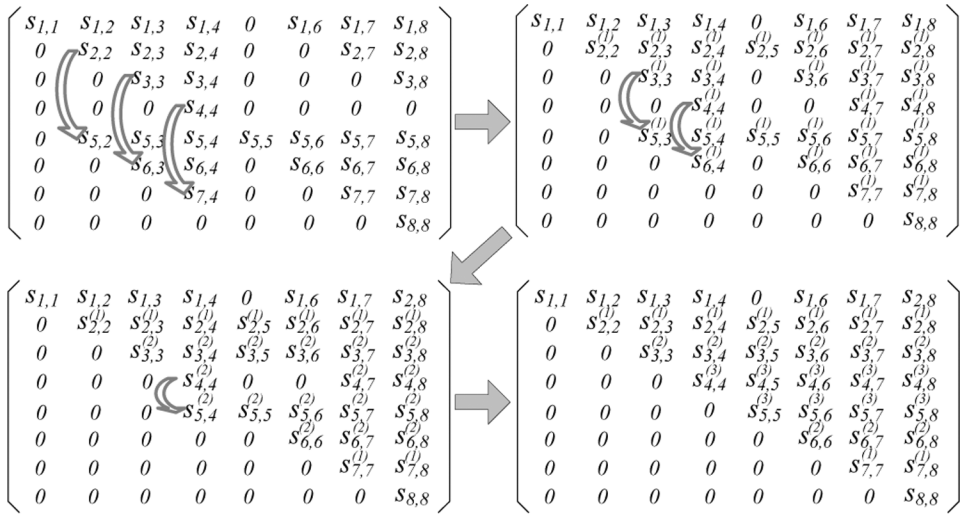

Fig. 4. Processing sequence of real Givens rotation.

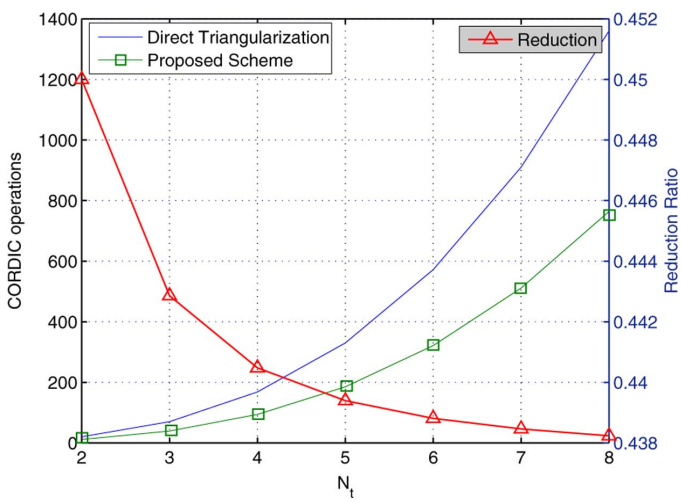

Fig. 5. Required CORDIC operations in the proposed scheme as well as the direct triangularization for various antenna configurations and the corresponding hardware reduction ratio.

our proposed scheme can save up to 44% arithmetic operations compared to the conventional direct triangularization.

IV. Architecture Design

Based on the proposed decomposition scheme, the hardware architecture is then designed with $N_t = 4$. Conventionally, the triangular systolic array has been employed to implement the hardware of the complex QR decomposition in [15]. Although the triangular systolic array is still utilized as the basic architecture, we remove the redundant components to improve the hardware efficiency, alter the processing sequences to decrease the delay elements, and propose a rate-matching real Givens rotation stage. The entire block diagram of the proposed QR decomposition architecture is shown in Fig. 6. The scheduling block

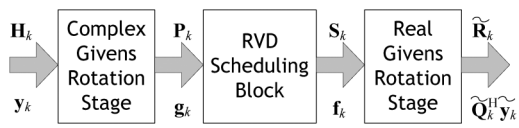

Fig. 6. Block diagram of the proposed QR decomposition architecture.

is inserted between the complex stage and the real stage for adjusting the data sequences.

A. Complex Givens Rotation Stage

Our complex Givens rotation stage is shown in Fig. 7. The complex Givens rotation stage has two operation schemes, scheme A for obtaining the upper triangular matrix $\mathbf{P}_k$ and scheme B for acquiring projection vector $\mathbf{g}_k$. In scheme A, signals of the complex channel matrix are fed into the systolic array. When the signal marked by a red rectangle enters into the processing element (PE) from the left, that processing element is operated in the vectoring mode, which generates θ_x, θ_y, and θ_z as given in (7). Otherwise, the processing element is configured in the rotation mode, in which the input signal pair is simply rotated according to the previous generated angles. In scheme B, received complex signal vector $\mathbf{y}_k$ goes into the processing elements, which are controlled to be operated in the rotation mode at all times.

In the vectoring mode, for two complex inputs, the south output of the complex PE (CPE) becomes zero while the east output becomes a real number. Consequently, $\mathbf{H}_{k,21}$ and $\mathbf{H}_{k,31}$ vanish after being processed by U1 and U2. Meanwhile, real $\mathbf{H}^{(1)}_{k,41}$ and $\mathbf{H}^{(1)}_{k,11}$ are produced. Since both inputs of U4 in the vectoring mode are real, we use CPE3 with the left two CORDIC modules for computing θ_x and θ_y being removed. Then, $\mathbf{H}^{(1)}_{k,41}$, output from U2, is zeroed after U4 and $\mathbf{P}_{k,11}$ is obtained from the east output of U4. $\mathbf{H}^{(1)}_{k,22}$ cancels $\mathbf{H}^{(1)}_{k,32}$ in U3 and thus real $\mathbf{H}^{(2)}_{k,22}$ enters into U5 for nulling $\mathbf{H}^{(2)}_{k,42}$. CPE2 which has three CORDIC modules is employed in U5 because again it is not necessary to calculate θ_x. In U6, we use $\mathbf{H}^{(2)}_{k,33}$ to cancel $\mathbf{H}^{(3)}_{k,43}$. A delay element (DE) is inserted in the west data

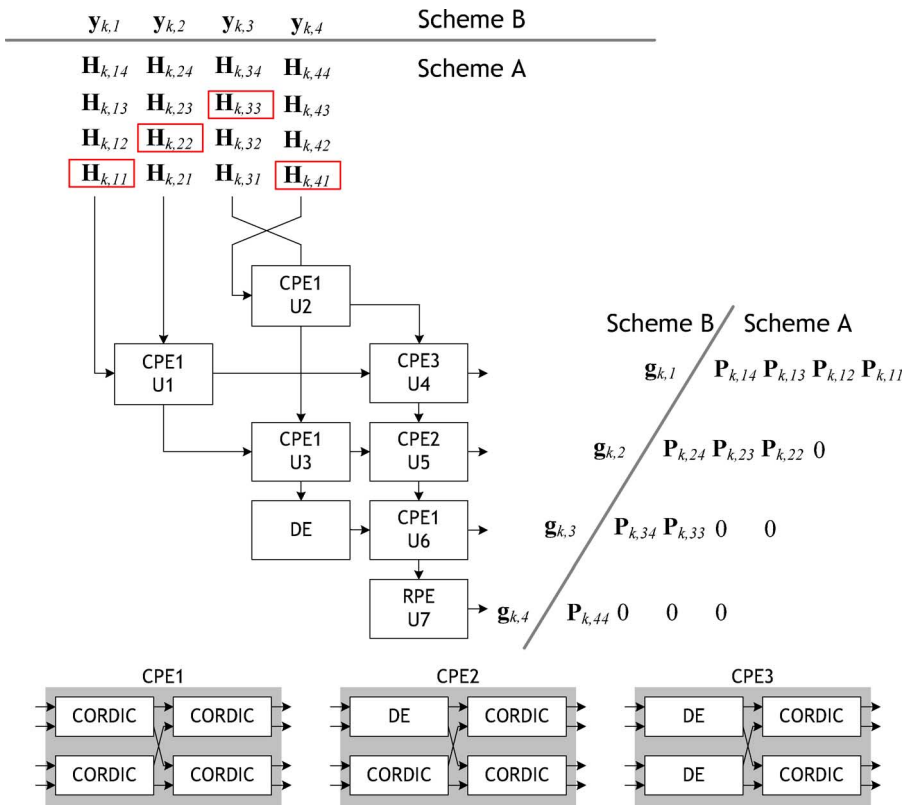

Fig. 7. Proposed systolic array for the complex Givens rotation stage.

path of U6 to adjust the same arrival time of $\mathbf{H}_{k,33}^{(2)}$ and $\mathbf{H}_{k,43}^{(3)}$. Finally, $\mathbf{H}_{k,44}^{(4)}$ is processed by the real PE (RPE) with only one CORDIC module so as to get real $\mathbf{P}_{k,44}$.

Annihilation re-ordering has been adopted in [18] to enhance the throughput. Here, our proposed re-ordering method in this triangular systolic array not only eliminates those delay elements used for skewed data inputs but also matches the data flows in the subsequent CPEs. Note that to keep the same latency for signal passing through the processing element, the same pipeline stages as the one in CPE1 are inserted in CPE2 and CPE3, indicated as the DE. Skewed outputs are naturally formed because signal of each row goes through different numbers of PE blocks. In scheme B, skewed output vector $\mathbf{g}_k = [\mathbf{g}_{k,1} \ \cdots \ \mathbf{g}_{k,4}]^T$ is derived from the input vector $\mathbf{y}_k$.

B. Real Givens Rotation Stage

At the real Givens rotation stage, we aim to eliminate the lower left non-zero elements in matrix $\mathbf{S}_k$ as illustrated in Fig. 4. If we simply use a triangular systolic array to implement the real Givens rotation stage, we will obtain the one as shown in Fig. 8(a). The RPE consists of one CORDIC module. Similarly in scheme A, when the input marked by a red square enters from west, the RPE operates in the vectoring mode. Otherwise, it works in the rotation mode. In this architecture, it takes eight clock cycles to complete triangularization and as a result becomes the bottleneck of the whole design. Furthermore, we can observe that a lot of 0's are fed as inputs, which cause inefficient utilization, both in hardware complexity and in processing time. Besides, as triangularization goes on, more 0's will be generated at each PE's south output. Consequently, we try to improve the throughput and the utilization by removing the occupation of these 0's spatially and temporally.

A rate-matching real Givens rotation stage is then designed as in Fig. 8(b), where the triangular systolic array is still employed at the bottom layer to process the elements in columns 4, 6, 7, and 8 of matrix $\mathbf{S}_k$. On top of it, additional RPEs are allocated to accomplish the remaining operations. Note that the vectoring mode must be occurred before its subsequent rotation mode so as to provide the necessary angles. Also, the elements in the same column must be aligned to ensure encountering the correct counterpart in each RPE. Thus, for matrix $\mathbf{S}_k$, s_{22} and s_{33} cancel s_{52} and s_{63} in U15 and U13 at time slot $4n$. Meanwhile, s_{23} and s_{53} processed in U14 follow the same micro-rotations as the ones generated by s_{22} and s_{52}. At time slot $4n+1$, two multiplexers with outputs connecting to U13 are switched and thus we deliver s_{44} and s_{74} into U13 set in the vectoring mode. On the other hand, $s_{53}^{(1)}$ from U14 and $s_{33}^{(1)}$ from U13 are sent into U15 for further processing. At time slot $4n+2$, $s_{55}^{(1)}$ and 0 are encountered in U15 for the second rotation. Subsequently at time $4n+3$, $s_{55}^{(2)}$ loops back and re-enters U15 for the third rotation with 0. The remaining all work the same as in Fig. 8(a).

By adding only two more RPEs and some multiplexers, we reduce the required clock cycles for triangularization of matrix $\mathbf{S}_k$ from eight to four. Obviously, with appropriate scheduling, the utilization at the real Givens rotation stage raises up to 94% in scheme A. Only U14 is idle for two clock cycles among four. Assume $\mathbf{f}_k = [f_1 \ f_2 \ \cdots \ f_7 \ f_8]^T$. In scheme B, elements of the vector $\mathbf{f}_k$ from row 2 to row 7 are needed to pass through the triangular systolic array at the bottom layer, all programmed in the rotation mode. The systolic array allows streaming inputs

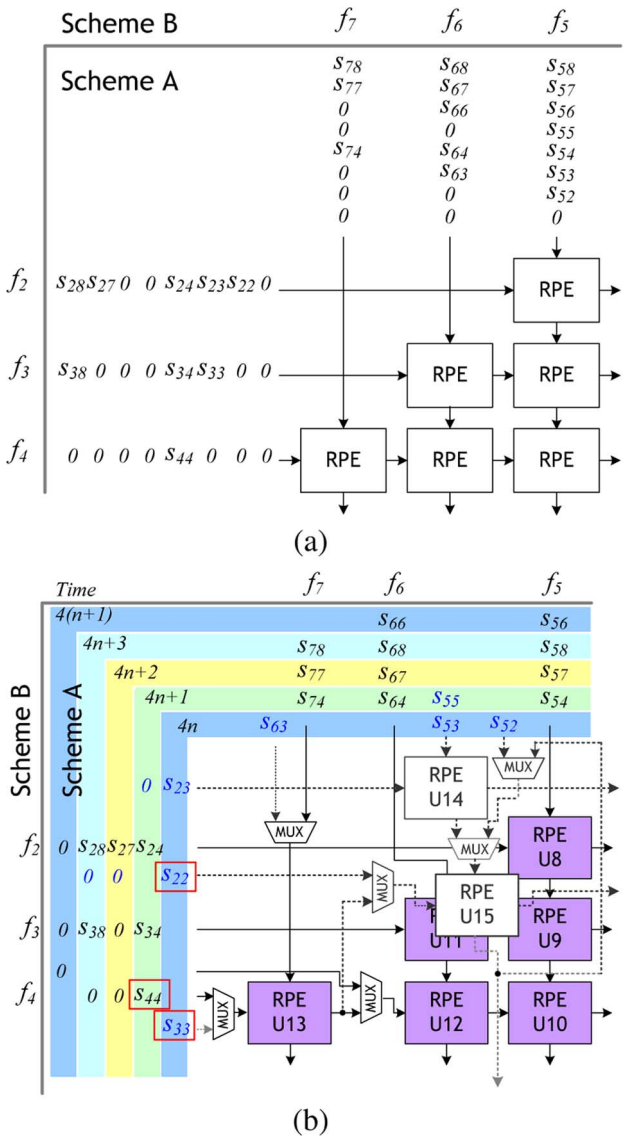

Fig. 8. Real Givens rotation stage (a) without rate-matching, and (b) with rate-matching.

and outputs. Consequently, the throughput of the real Givens rotation stage matches with the one of the complex Givens rotation stage.

C. Scheduling Block

Scheduling block is needed to perform RVD of matrix $\mathbf{P}_k$ and to allocate the signal in the correct time slot for the real Givens rotation stage. Observing Figs. 7 and 8(b), we can see that among one set of the skewed outputs, $\mathbf{P}_{k,44}$ is the last signal indicating the completion of one complex Givens rotation, while $s_{44} = \mathrm{Re}\{\mathbf{P}_{k,44}\}$ must be scheduled at time slot $(4n+1)$. Consequently, the necessary delay for each input signal can be derived by comparing their relative temporal spacing.

The delay buffer is designed as shown in Fig. 9. We first align $\mathrm{Re}\{\mathbf{P}_{k,44}\}$ and s_{44} at time slot $(4n+1)$ for both input and output sequences as depicted in Fig. 9(c). Instead of elaborating one delay buffer for each signal, we try to merge those delay buffers. The steps are given in the following.

- List the input signal generated from the complex Givens rotation stage to the delay buffer from time slot $4n-5$ to $4n$.
- Observe the required buffer depth and width for total signals at each time slot as shown in Fig. 9(a). The maximum buffer depth is decided by the signal with the longest delay time. The buffer width varies and depends on the largest number of the signals to be stored for certain memory depth. Consequently, if we use vector $\mathbf{m}$ with the ith entry denoting the required memory width for depth i, then we get $\mathbf{m} = [4\ 4\ 4\ 3\ 2\ 1\ 1\ 1]$.
- To avoid commutators/switches demanded in between every delay element, adjust the signals in every column vertically so that they are kept moving in the same row. Besides, to allow for streaming input/output, each row contains at most four consecutive signals. Regard them as one group as depicted in Fig. 9(b).
- To increase the utilization, signals of group 4 are inserted to the end of the second delay buffer. $\mathrm{Re}\{\mathbf{P}_{k,12}\}$ for s_{56} with the longest delay of 8 units stays at the last delay element for three clock cycles in the first delay buffer because no more desired signal enters. Thus, the number of delay elements is further improved from 20 to 17 with $\mathbf{m}' = [3\ 3\ 3\ 3\ 2\ 2\ 1]$ as illustrated in Fig. 9(c).

In scheme B, the required delay buffer is much simpler because all the signals are synchronized to a single time slot as indicated in Fig. 8(b).

V. Implementation Results and Comparisons

The proposed QR decomposition is then implemented in 0.18-μm CMOS technology. In the following, we will first demonstrate its performance and measurement results. Subsequently, comparisons and discussions will be provided to show the advantages of our design.

A. Implementation Results

To examine the performance of the hardware, a cycle-accurate and bit-true model is built. We then simulate its functional capability considering the finite precision effect to support MIMO detection with constellation of 64-QAM. Two commonly-used sphere decoders (SD) are considered here. One is K-best SD to offer constant detection throughput. The other is depth-first SD, which can achieve maximum likelihood (ML) detection performance. In Fig. 10, we scan various word-length settings in the fractional part of data-path corresponding to the precision to see its influence in bit-error-rate performance. The SNR is set to 35 dB and the K-best SD is used. The CORDIC module has nine micro-rotation stages plus one initial stage. According to the simulation, the word-length of the data path is selected to be 16 bits, including 11 bits for the fractional part and 5 bits for the integral part.

In Fig. 11, we provide the bit-error-rate performance in the interested SNR range. From the figure, we can see that even the proposed fixed-point design is used as a pre-processing block of an ML MIMO detector, the degradation is almost negligible.

To achieve high-speed operation, three pipeline stages $(N_{pp} = 3)$ are inserted in each CORDIC module and the

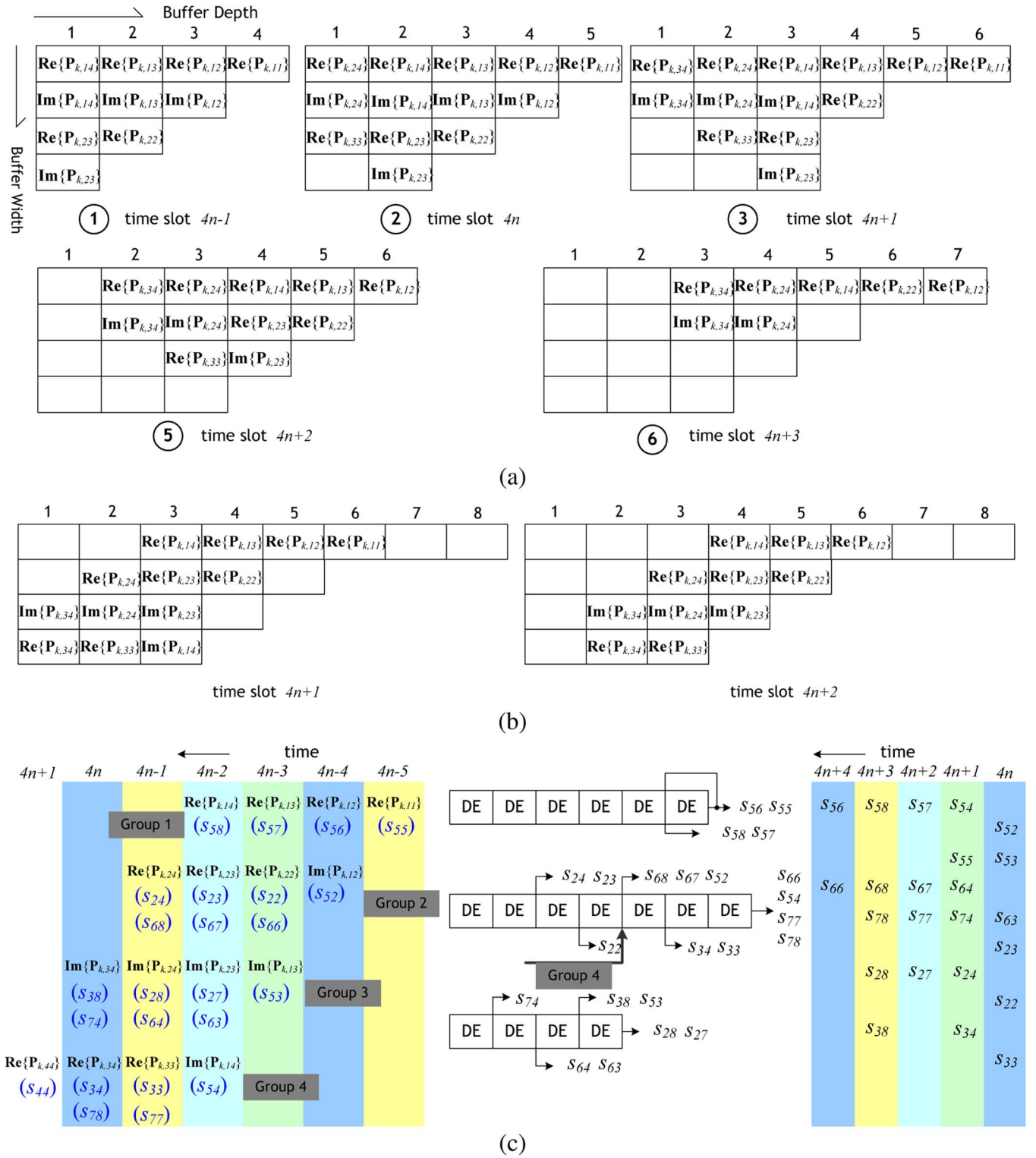

Fig. 9. (a) Observe the required depth and width of the delay buffer. (b) Adjust the signal group to allow streaming signal flow. (c) Design of proposed delay buffer.

critical path delay is reduced to about 10 ns. Hence, each delay element consists of three D-flip-flops. The chip photo is shown in Fig. 12 with essential blocks indicated. The gate count is about 152K, and the operating frequency can be driven up to 100 MHz according to the measurement results. The complex Givens rotation stage and the real Givens rotation stage occupy 73% and 20% gate counts, respectively. The remaining 7% gates are delay elements in the scheduling block. The chip is packaged in CQFP with 160 I/O pins. Among them, 106 pins are reserved as the inputs to ensure all the necessary input signals can be updated simultaneously. Also, one pin is defined as scheme selection, which can configure the chip to switch between two schemes. However, to avoid a pad-limited design, the outputs are multiplexed in 12 pins. We use a 4-bit control

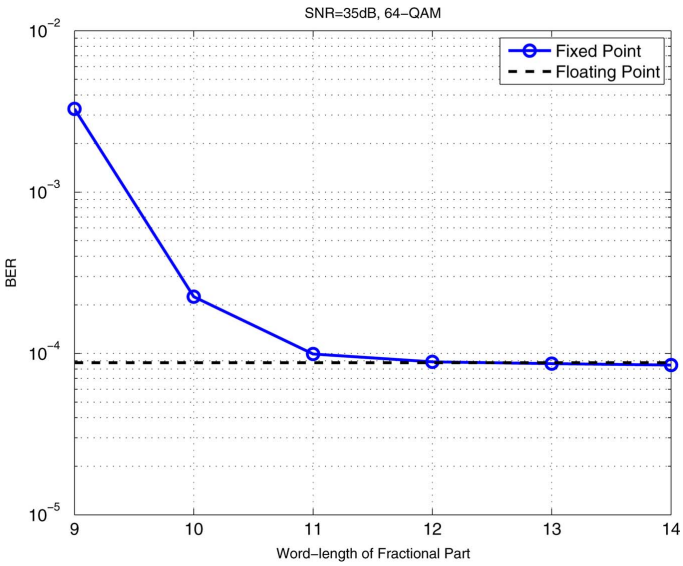

Fig. 10. Bit-error-rate performance versus word-length settings.

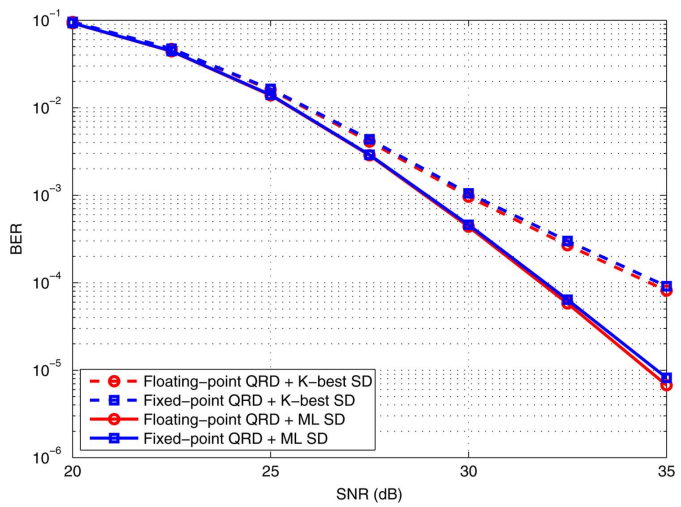

Fig. 11. Bit-error-rate performance of our proposed QR decomposition module used for pre-processing in MIMO detection.

signal to choose the interested observation data paths. With the capability of streaming I/O, it takes 40 ns to complete the QR-decomposition of an 8×8 RVD channel matrix or a 4×4 complex channel matrix and 10 ns to compute vector projection of $\mathbf{Q}^H\mathbf{y}$ at 100-MHz operating frequency and 1.8-V supply voltage, satisfying the requirement for gigabit MIMO detection. Fig. 13(a) depicts the measured maximum operating frequency at different supply voltages. Fig. 13(b) shows its power consumption for two operation schemes at the maximum operating frequency under several supply voltages. As expected, QR decomposition/triangularization is more power-consuming than vector projection since almost all the modules are active and the dynamic range is larger. From the figure, it demonstrates that our design not only offers high throughput but also can be used in power-saving condition. The chip summary is given in Table IV.

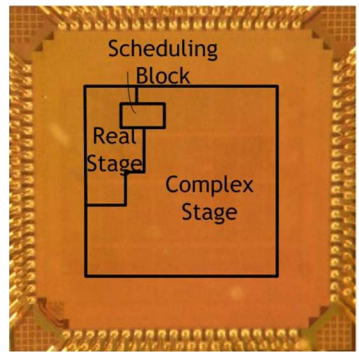

Fig. 12. Chip photo.

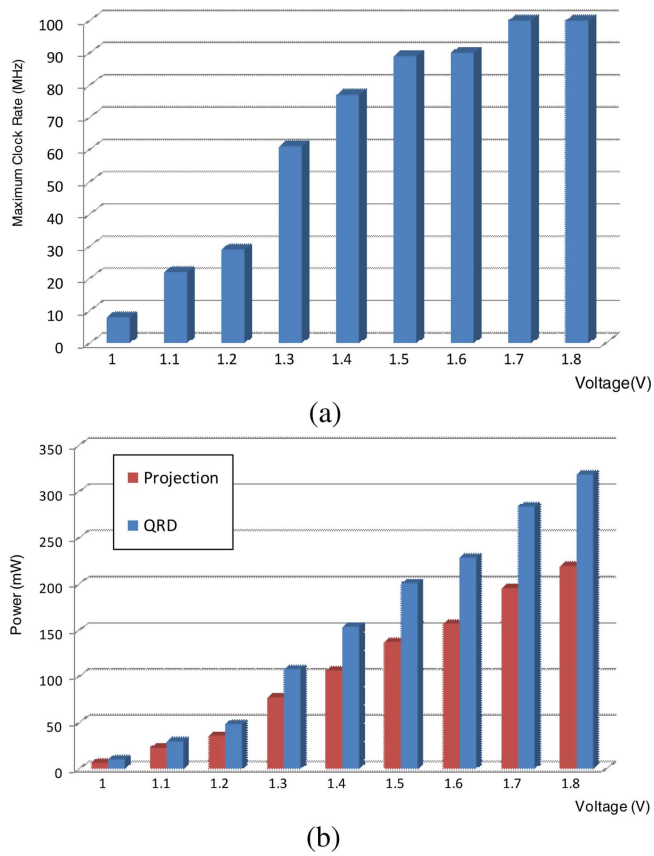

Fig. 13. Measurement results of (a) maximum operating frequency and (b) power consumption at maximum operating frequency under different supply voltages.

TABLE IV
CHIP SUMMARY

Technology	TSMC 0.18 μm CMOS	
Package	CQFP160	
Voltage (Core/Pad)	1.8V/3.3V	
Core Size	1.677 × 1.673 mm^2	
Max. Freq.	100 MHz	
	QRD	Projection
Max. Throughput	1/40ns	2.4 Gbps
Power Consumption	319 mW @100MHz, 1.8V 10 mW @8MHz, 1V	220 mW @100MHz,1.8V 6 mW @8MHz, 1V

TABLE VI
COMPARISON OF THE CONVENTIONAL WORKS AND THE PROPOSED DESIGN

Items	Huang [16]	Singh [10]	Chang [11]	Salmela [24]	Patel [17]	This Work	
Algorithm	Givens Rotation	Gram-Schmidt	Gram-Schmidt	Gram-Schmidt	Givens Rotation	Givens Rotation	
Complex/Real	Complex	Real	Real	Complex	Complex	Complex	Real
Matrix Size ($N \times N$)	2×2	4×4	4×4	4×4	4×4	4×4	8×8
Technology	0.18 μm	0.18 μm	0.18 μm	0.13 μm	0.13 μm	0.18 μm	
Max. Freq.	202 MHz	277 MHz	400 MHz	269 MHz	270 MHz	100 MHz†	
Gate Count	17K	72K	32.6K	23K	36K	111K	152K
Processing Cycles (QRD)	-	67	35	139	40	4	
QRD Rate (QRD/s)	-	1/242n	1/87.5n	1/517n	1/148n	1/40n	
N. T. of QRD (QRD/s)	-	1M	2.9M	0.7M	2.4M	12.5M	25M
N.H.E	-	14.4	87.6	30.3	67.7	112.6	164.5
Power(mW)	-	-	-	-	-	-	319
Processing Cycles (Projection)	8	67	35	139	10	1	
Projection Rate (Projection/s)	1/39.6n	1/242n	1/87.5n	1/517n	1/37n	1/10n	
MIMO Detection Throughput	303 Mbps	50 Mbps	137 Mbps	46 Mbps	648 Mbps	2.4 Gbps	
Power (mW)	-	-	-	-	-	220	

†: Measurement result.

TABLE V
SUMMARY OF HARDWARE SAVING

Items	Original	Improved	Saving
Decomposition Strategy	Direct triangularization	Cascading two stages	44%
Delay Elements at Complex Stage	$88 \times N_{pp} \times N_{wl}$	$56 \times N_{pp} \times N_{wl}$	36%
CORDIC Modules at Complex Stage	25	22	12%
CORDIC Modules at Real Stage	12	8	33%
Delay Elements in Scheduling Block	$66 \times N_{pp} \times N_{wl}$	$17 \times N_{pp} \times N_{wl}$	74%

B. Discussions and Comparisons

The complexity saving that has been attained through the proposed techniques is summarized in Table V. By cascading complex Givens rotation stage and the real Givens rotation stage, arithmetic operations are decreased. At the complex Givens rotation stage, we remove the necessity of the delay elements for skewed inputs. For N_{pp} pipeline stages in one CORDIC module and the word-length of N_{wl} bits, it saves $12 \times N_{pp} \times N_{wl} \times 2$ delay elements for complex signals. Besides, delay elements between CPEs are also decreased by $4 \times N_{pp} \times N_{wl} \times 2$. Furthermore, we use only 22 rather than $4 \times 6 + 1$ CORDIC modules. Compared to [15], 36% improvement in delay elements and 12% saving in arithmetic logics at the complex Givens rotation stage are achieved. At the real Givens rotation stage, instead of 12, we utilize eight CORDIC modules to meet the throughput requirement, a reduction ratio of 33% in complexity. Finally, by sharing and reusing techniques, we attain 74% decrease in delay elements that originally may be incurred between the complex and real Givens rotation stages. With these design efforts, hardware utilization is upgraded.

According to our observations, the gate count and power consumption per CORDIC module do not vary too much. The required CORDIC operations under various antenna configurations are given by (9) and (10). The efficient hardware architecture in terms of actual CORDIC modules can be derived by the acceptable processing cycles constrained by respective system specification. Hence, given that the scheduling block adds minor overhead, design scaling concerning power and area with the number of transmit antennas can be approximately estimated.

In Table VI, we list several works for QR decomposition in the order of complexity. Synthesis results are provided in other works. It is clear that we support the more complicated QR decomposition for 8×8 RVD channel matrixes for MIMO detection problems. In [16], their design outputs MIMO detection results directly without $\mathbf{R}$ and $\mathbf{Q}$ matrixes, and thus the related data for QR decomposition are not available. Note that the processing cycles refer to the number of cycles that a new matrix can enter for decomposition or projection. With streaming I/O, our design can process QR decomposition every four clock cycles. The QRD rate defines how many QR decompositions per second the work is able to perform. According to the complexity analysis in Section III and the result in Fig. 2, the complexity almost grows linearly in log scale for decomposition of a matrix in different sizes with real or complex entries. As a result, for a fair comparison the normalized throughput (N.T.) of QR decomposition is defined as

$$\text{N.T.} = (\text{QRD rate}) \cdot \eta \cdot \frac{N \times N}{8 \times 8} \cdot \frac{\text{Technology}}{0.18\ \mu\text{m}}. \quad (11)$$

If a complex matrix is decomposed, $\eta = 2$ and the normalized throughput is doubled to count for the real part and the imaginary part and thus $\eta = 2$. Otherwise, $\eta = 1$. We can see our work has the highest QRD throughput. The metric of normalized hardware efficiency (NHE) is defined as the normalized throughput divided by the gate counts. From the table, undoubtedly our design also has the higher hardware efficiency than the conventional works for QR decomposition.

As to the projection of $\mathbf{Q}^H\mathbf{y}$ for MIMO detection, assume that in [10], [11], [24] the complex multiplication is done right after the column vector of $\mathbf{Q}$ is acquired. (However, extra complexity of multiplication is not included here.) For those works which decompose a 4×4 real matrix, $N_t = 2$

equivalently. The MIMO detection throughput is computed as $6N_t/(\text{Projection rate})$, where 6 corresponds the number of bits in a symbol if 64-QAM constellation is used. Consequently, we can support the highest MIMO detection rate up to 2.4 Gb/s, which is essential to the development for gigabit wireless systems.

VI. Conclusion

In this paper, we have implemented a high-throughput QR decomposition module for MIMO-OFDM systems. We combine both complex Givens rotation algorithm and real Givens rotation algorithm to reduce the required CORDIC operations. Moreover, we eliminate the delay elements required in the conventional complex Givens rotations systolic array and improve hardware utilization at the real Givens rotation stage by the time-sharing technique. We have shown that the proposed architecture has good hardware efficiency to achieve the more complicated QR decomposition of the 8×8 RVD channel matrix. According to the measurement results, the operating frequency can achieve 100 MHz in 0.18-μm CMOS technology. It can accomplish the QR decomposition every 40 ns as well as vector projection every 10 ns to support MIMO detection up to 2.4 Gb/s with power consumption of 319 and 220 mW, respectively. The proposed design definitely plays an important role in advance gigabit wireless communication systems.

Acknowledgment

The authors greatly appreciate the Chip Implementation Center (CIC) of the National Science Council, Taiwan, R.O.C., for chip implementation.

References

[1] A. J. Paulray, D. A. Gore, R. U. Nabar, and H. Bolcskei, "An overview of MIMO communications—A key to gigabit wireless," *Proc. IEEE*, vol. 92, no. 2, pp. 198–218, Feb. 2004.

[2] H. Jeon, IEEE P802.11 Wireless LANs—802.11 TGac Functional Requirements. IEEE 802 committee, doc.: IEEE 802.11-08/1285r0, Nov. 2008.

[3] WPAN Millimeter Wave Alternative PHY Task Group 3c (TG3c), IEEE 802.15. [Online]. Available: http://www.ieee802.org/15/pub/TG3c.html

[4] J. Liu, Y. V. Zakharov, and B. Weaver, "Architecture and FPGA design of dichotomous coordinate descent algorithms," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, no. 11, pp. 2425–2438, Nov. 2009.

[5] T. H. Im, I. Park, J. Kim, J. Yi, J. Kim, S. Yu, and Y. S. Cho, "A new signal detection method for spatially multiplexed MIMO systems and its VLSI implementation," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 56, no. 5, pp. 399–403, May 2009.

[6] K. J. Kim, M. O. Pun, and R. A. Iltis, "QRD-based precoded MIMO-OFDM systems with reduced feedback," *IEEE Trans. Commun.*, vol. 58, pp. 394–398, Feb. 2010.

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

[8] M. Shabany and P. G. Gulak, "A 0.13 mm CMOS 655 Mb/s 4x4 64-QAM K-best MIMO detector," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, 2009, pp. 256–257.

[9] Y. Kim, J. H. Park, and J. W. Kim, "Hybrid MIMO receiver using QR-MLD and QR-MMSE," in *Proc. IEEE GLOBECOM*, 2009, pp. 1–5.

[10] C. K. Singh, S. H. Prasad, and P. T. Balsara, "VLSI architecture for matrix inversion using modified Gram–Schmidt based QR decomposition," in *Proc. Int. Conf. VLSI Design*, 2007, pp. 836–841.

[11] R. C.-H. Chang, C. H. Lin, K. H. Lin, C. L. Huang, and F. C. Chen, "Iterative decomposition architecture using the modified Gram–Schmidt algorithm for MIMO systems," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 57, no. 5, pp. 1095–1102, May 2010.

[12] P. Luethi, C. Studer, S. Duetsch, E. Zgraggen, H. Kaeslin, N. Felber, and W. Fichtner, "Gram–Schmidt-based QR decomposition for MIMO detection: VLSI implementation and comparison," in *Proc. IEEE Asia Pacific Conf. Circuits and Systems*, 2008, pp. 830–833.

[13] C. F. T. Tang, K. J. R. Liu, and S. A. Tretter, "On systolic arrays for recursive complex Householder transformations with applications to array processing," in *Proc. Int. Conf. Acoustics, Speech, and Signal Process.*, 1991, pp. 1033–1036.

[14] K.-L. Chung and W.-M. Yan, "The complex Householder transform," *IEEE Trans. Signal Process.*, vol. 45, no. 9, pp. 2374–2376, Sep. 1997.

[15] A. Maltsev, V. Pestretsov, R. Maslennikov, and A. Khoryaev, "Triangular systolic array with reduced latency for QR-decomposition of complex matrices," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, 2006, pp. 385–388.

[16] Y. T. Hwang and W. D. Chen, "A low complexity complex QR factorization design for signal detection in MIMO OFDM systems," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, 2008, pp. 932–935.

[17] D. Patel, M. Shabany, and P. G. Gulak, "A low-complexity high-speed QR decomposition implementation for MIMO receivers," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, 2009, pp. 1409–1412.

[18] Z. Liu, K. Dickson, and J. V. McCanny, "Application-specific instruction set processor for SoC implementation of modern signal processing algorithms," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 52, no. 4, pp. 755–765, Apr. 2006.

[19] IEEE project 802.16m as an IMT-Advanced Technology. IEEE 802.16 Working Group on Broadband Wireless Access. [Online]. Available: http://www.ieee802.org/16/liaison/docs/L80216-08_057r2.pdf

[20] P. Y. Tsai, W. T. Chen, X.-C. Lin, and M. Y. Huang, "A 4x4 64-QAM reduced-complexity K-best MIMO detector up to 1.5 Gb/s," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, 2010, pp. 3953–3956.

[21] L. Liu, F. Ye, X. Ma, T. Zhang, and J. Ren, "A 1.1-Gb/s 115-pJ/bit configurable MIMO detector using 0.13 μm CMOS technology," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 57, no. 9, pp. 701–705, Sep. 2010.

[22] C. A. Shen and A. M. Eltawil, "A radius adaptive K-best decoder with early termination: Algorithm and VLSI architecture," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 57, no. 9, pp. 2476–2486, Sep. 2010.

[23] Z. Y. Huang and P. Y. Tsai, "High-throughput QR decomposition for MIMO detection in OFDM systems," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, 2010, pp. 1492–1495.

[24] P. Salmela, A. Burian, H. Sorokin, and J. Takala, "Complex-valued QR decomposition implementation for MIMO receivers," in *Proc. Int. Conf. Acoustics, Speech, and Signal Process.*, 2008, pp. 1433–1436.

Zheng-Yu Huang received the B.S. degree in electronics engineering from Da-Yeh University, Changhua, Taiwan, R.O.C., in 2007, and the M.S. degree in electrical engineering from National Central University, Taoyuan, Taiwan, R.O.C., in 2010.

His research interests include VLSI architectures, digital signal processors, wireless communication systems, especially in MIMO-OFDM baseband receivers.

Mr. Huang received the MXIC Golden Silicon Award in 2010.

Pei-Yun Tsai (S'02–M'06) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the National Taiwan University, Taipei, Taiwan, in 1994, 1996, and 2005, respectively.

From 1996 to 2000, she worked at ASUStek and participated in the team of optical storage systems. She is now an Assistant Professor of electrical engineering at National Central University, Taoyuan, Taiwan. Her research interests include baseband signal processing algorithms and VLSI design for digital communication systems.

Prof. Tsai received the Acer Longtern Award and the First Asian Solid-State Circuit Conference Student Design Contest Outstanding Award in 2005. She also received the MXIC Golden Silicon Award in 2005 and 2010.