Unified Multi-Mode Signal Detector for LTE-A Downlink MIMO System

Liang Liu, Johan Löfgren, and Peter Nilsson
Department of Electrical and Information Technology, Lund University, Sweden
Email: {Liang.Liu, Johan.Lofgren, Peter.Nilsson}@eit.lth.se

***Abstract*—This paper presents a low power and small area multi-mode MIMO detector targeting the LTE-A downlink. The detector supports the detection of spatial-multiplexing, spatial-diversity, and space-division-multiple-access MIMO signals of different antenna configurations and modulation schemes. Cost-efficiency is achieved by algorithm and architecture co-design where low-complexity, near-maximum-likelihood (ML) detection algorithms are proposed for these three MIMO technologies respectively with most of the mathematical operations being reused. A unified parallel VLSI architecture is accordingly developed that achieves high detection throughput and run-time reconfigurability to support different MIMO modes. The proposed detector has been designed using a 65-nm CMOS technology with 88.2K equivalent gate-count, representing a 22% less hardware-resource use than the state of art in the open literature. Operating at 1.2-V supply with 165-MHz clock, the detector achieves a 1.98Gb/s peak throughput with a energy consumption of 51.8 pJ/b, which is shown to be the most energy-efficient design compared with other 64-QAM detectors.**

I. Introduction

It has been a broad agreement that the enhanced multiple-input multiple-output (MIMO) technology plays an essential role in the emerging wireless standards, e.g., 3GPP Long Term Evolution Advanced (3GPP-LTE-A) [1], to achieve the International Mobile Telecommunications Advanced (IMT-A) requirement for the 4th Generation (4G) wireless network deployment where very high spectrum efficiencies are expected.

Cellular systems experience highly dynamic channel conditions, where the noise and channel response vary within huge ranges. To guarantee the quality of service (QoS), it is necessary that the system is equipped with multiple MIMO technologies to be adaptive to the fluctuating channels. Extensive discussions are ongoing about the multi-mode MIMO schemes in LTE-A [2]. Switching mechanism among different MIMO transmissions includes spatial multiplexing (SM), space division multiple access (SDMA), and spatial diversity (SD) [3]. For this multi-mode system, multiple detection modules must be configured at the receiver side with each dedicated module corresponding to the respective transmission mode. Such implementation strategy will nevertheless incur huge silicon area overhead and be immensely inefficient because most of the modules would remain in an idle state for a large part of the time. As a consequence, an efficient implementation is expected to integrate multiple MIMO detectors into a single module which can be reconfigured at run-time. Moreover, the power consumption should be low enough to be adopted in hand-held devices where the high performance and flexibility need to be combined with the energy efficiency.

To achieve this objective, this paper proposes a unified signal detector that supports 64-QAM modulated SM/SDMA/SD signals for up to 4×4 MIMO transmission and provides beyond-gigabit-per-second throughput to meet the LTE-A date rate requirement. The mode unification is mainly achieved by algorithm-level exploitation, where the developed algorithms comprise similar mathematical operations to enable substantial hardware reuse. Firstly, we develop detection algorithm for SM mode by introducing an early-pruned scheme to the fixed-complexity sphere decode (FSD) algorithm with a novel polygon-shaped constraint, which could greatly facilitate complexity saving without obvious performance degradation. On the basis of function analysis and partition to this early-pruned FSD, we then develop detection schemes for SD and SDMA modes. More specifically, we propose a real-valued successive interference cancelation (SIC) algorithm for SD system which reuses the QR/real-value decomposition at the pre-processing stage to notably simplify the detection. For the SDMA mode, we introduce a matrix permutation scheme that completely reuses the early-pruned FSD to realize an ML downlink multi-user detection. Taking advantage of these implementation-oriented algorithms, a unified VLSI architecture is subsequently developed, capable of being reconfigured at run-time to support three MIMO transmission modes of different antenna configurations, modulation schemes, as well as node-extension numbers. The effectiveness of the proposed design solutions have been validated by post-layout simulations with 65-nm CMOS standard digital cell.

II. System Model

We consider an LTE-A downlink MIMO system with one base station (BS) and K user equipments (UEs), where both the BS and the UE are equipped with N antennas. The received $N \times 1$ vector in the n^{th} subcarrier of user k is

$$\tilde{\boldsymbol{r}}_{\boldsymbol{k,n}} = \tilde{\boldsymbol{H}}^{\boldsymbol{c}}_{\boldsymbol{k,n}} \sum_{k=1}^{K} \tilde{\boldsymbol{W}}_{\boldsymbol{k,n}} \tilde{\boldsymbol{x}}_{\boldsymbol{k,n}} + \tilde{\boldsymbol{w}}_{\boldsymbol{k,n}}, n = 1, 2, ..., N_{sub}, \quad (1)$$

where N_{sub} is the number of sub-carriers, $\tilde{\boldsymbol{H}}^{\boldsymbol{c}}_{\boldsymbol{k,n}}$ is the $N \times N$ complex channel matrix between the BS and the k^{th} UE, $\tilde{\boldsymbol{x}}_{\boldsymbol{k,n}}$ is the $P \times 1$ transmitted vector at the n^{th} subcarrier for user k [1], $\tilde{\boldsymbol{w}}_{\boldsymbol{k,n}}$ is the vector of Gaussian noise samples, and $\tilde{\boldsymbol{W}}_{\boldsymbol{k,n}}$ is the $N \times P$ pre-coding matrix selected from a finite

code-book that is known to both the BS and UE. Without loss in generality, we set $K = 1$ and $P = N$ in both SM and SD transmissions. $\tilde{\boldsymbol{W}}_{\boldsymbol{k}}$ is set to be an N-dimensional identity matrix in the SM system and an Alamouti space-frequency coding matrix [4] in the SD system. For the SDMA system, P is limited to one according to the 3GPP-LTE standard [5], K is set to N, and $\tilde{\boldsymbol{W}}_{\boldsymbol{k},\boldsymbol{n}}$ is chosen such that $\tilde{\boldsymbol{W}}_{\boldsymbol{k},\boldsymbol{n}}^{\boldsymbol{H}}\tilde{\boldsymbol{W}}_{\boldsymbol{k},\boldsymbol{n}} = 1$ and $\tilde{\boldsymbol{W}}_{\boldsymbol{k},\boldsymbol{n}}^{\boldsymbol{H}}\tilde{\boldsymbol{W}}_{\boldsymbol{l},\boldsymbol{n},\boldsymbol{l}\neq\boldsymbol{k}} = 0$. For simplicity, we will neglect the subcarrier index n selectively in the rest of the paper.

The N-dimensional complex system model can be transformed to its $2N$-dimensional real-valued representation as $\boldsymbol{r_k} = \boldsymbol{H}[\boldsymbol{x_1}, \ldots, \boldsymbol{x_k}, \ldots, \boldsymbol{x_K}]^T + \boldsymbol{w_k}$, where

$$\begin{aligned}\boldsymbol{r_k} &= [\Re(\tilde{r}_{k,1}), \Im(\tilde{r}_{k,1}), \ldots, \Re(\tilde{r}_{k,N}), \Im(\tilde{r}_{k,N})]^T \\ \boldsymbol{x_k} &= [\Re(\tilde{x}_{k,1}), \Im(\tilde{x}_{k,1}), \ldots, \Re(\tilde{x}_{k,P}), \Im(\tilde{x}_{k,P})] \\ \boldsymbol{w_k} &= [\Re(\tilde{w}_{k,1}), \Im(\tilde{w}_{k,1}), \ldots, \Re(\tilde{w}_{k,N}), \Im(\tilde{w}_{k,N})]^T,\end{aligned} \tag{2}$$

and

$$\boldsymbol{H} = \begin{bmatrix} \Re(\tilde{H}_{1,1}) & -\Im(\tilde{H}_{1,1}) & \ldots & \Re(\tilde{H}_{1,N}) & -\Im(\tilde{H}_{1,N}) \\ \Im(\tilde{H}_{1,1}) & \Re(\tilde{H}_{1,1}) & \ldots & \Im(\tilde{H}_{1,N}) & \Re(\tilde{H}_{1,N}) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \Re(\tilde{H}_{N,1}) & -\Im(\tilde{H}_{N,1}) & \ldots & \Re(\tilde{H}_{N,N}) & -\Im(\tilde{H}_{N,N}) \\ \Im(\tilde{H}_{N,1}) & \Re(\tilde{H}_{N,1}) & \ldots & \Im(\tilde{H}_{N,N}) & \Re(\tilde{H}_{N,N}) \end{bmatrix}. \tag{3}$$

In (2) and (3), $\tilde{\boldsymbol{H}} = \tilde{\boldsymbol{H}}_{\boldsymbol{k}}^{\boldsymbol{c}}[\tilde{\boldsymbol{W}}_{\boldsymbol{1}}, \ldots, \tilde{\boldsymbol{W}}_{\boldsymbol{K}}]$ is the equivalent complex channel matrix, $\Re(\cdot)$ and $\Im(\cdot)$ represent the real and imaginary parts of a complex number, respectively.

III. Multi-Mode Detection Algorithms

A. Spatial-Multiplexing Signal Detection

The objective of MIMO detection is to recover the original $\boldsymbol{x_k}$, given $\boldsymbol{r_k}$, $\boldsymbol{W_k}$, and $\boldsymbol{H_k}$. As previously mentioned, we start with the SM signal detection by exploiting the near-ML fixed-complexity sphere decoder (FSD), which has recently gained much attention because of its completely regular and feed-forward-only dataflow [6]. FSD transfers the closet-point search problem in ML detection to a tree-search problem by rewriting the Euclidean distance as $|\boldsymbol{y} - \boldsymbol{Q}\boldsymbol{s}|^2$, where $\boldsymbol{Q}$ is an unitary matrix and $\boldsymbol{R}$ is upper triangular obtained by $\boldsymbol{H} = \boldsymbol{Q}\boldsymbol{R}$, $\boldsymbol{y} = \boldsymbol{Q}^H\boldsymbol{r}$, and $\boldsymbol{v} = \boldsymbol{Q}^H\boldsymbol{n}$ is a noise vector with the same statistics as $\boldsymbol{n}$. Starting from the top ($2N^{th}$) layer, FSD carries out the tree search by

$$\begin{aligned} T_i &= T_{i+1} + inc_i, \\ inc_i &= |y_i - \sum_{j=i+1}^{2N} R_{ij}x_j - R_{ii}x_i|^2 = |\tilde{y}_i - R_{ii}x_i|^2, \end{aligned} \tag{4}$$

where T_i is the partial Euclidean distance (PED) at the i^{th} layer. Uniquely characterized by a parameter p, an FSD algorithm instance performs a full-search in the first p layers, expanding all $\sqrt{M}$ branches per node, while a single-search in the remaining $(2N - p)$ layers, expanding only one branch per node. Although the use of single-search scheme reduces computations to a certain extent, there are still a large amount of wasted calculations induced by visiting too many nodes at upper layers, especially in high-order modulations. In tree-search algorithms, a sphere constraint is typically applied to reduce such unnecessary visits to some nodes [7], i.e., given a radius constraint r^2, we only consider those nodes that could possibly lead to a PED less than r^2. However, the reduction of calculations in radius-constrained algorithm is limited, because a node dissatisfying the constraint will not be pruned until its PED is completely calculated and compared with the radius.

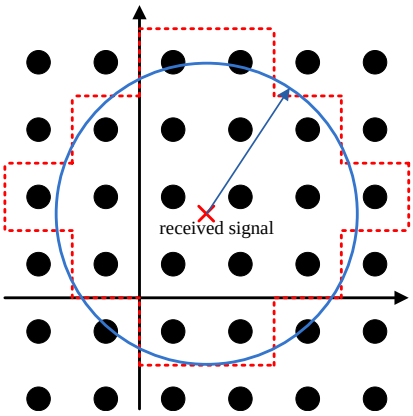

Fig. 1. Polygon-shaped constraint with $\boldsymbol{L}^{2i-1} = [5, 5, 3, 3, 1, 1, 0, 0]$.

To exploit the merit of the constraint technology while preventing its complexity overhead, we suggest an early-punning strategy by approximating the traditional circular-shaped admissible region with a polygon-shaped admissible region, as illustrated in Fig. 1. Practically, this polygon constraint is realized by putting a extension number limitation L to the node expansion with which only the L most promising child nodes are extended for further computation. In our real-valued tree search, L is characterized by a set of extension number limitations $\boldsymbol{L}^{2i-1} = [l_1^{2i-1}, \ldots, l_{\sqrt{M}}^{2i-1}]$, where the entries l_m^{2i-1} is responsible for imaginary nodes at the $(2i-1)^{th}$ layer, denoting that only l_m^{2i-1} best nodes are exploited from the m^{th} father node. On the other hand, the number of non-zero elements limits the extension of the real part at the $2i^{th}$ $(2N - 2i < p)$ layer. Specified in our considered 4×4 MIMO system, $\boldsymbol{L}^{2i-1}$ is only adopted to the $(2N-1)^{th}$ layer because the parameter p in FSD equals to 2 for 4×4 MIMO [6].

Different from the radius constraint, this proposed early-pruned scheme prunes less reliable paths before PED calculations by using the well-studied real-valued zigzag enumeration technology [8], thus can save a large number of PED computations. Moreover, the number of nodes to be extended is fixed in our algorithm with a pre-set constraint $\boldsymbol{L}^{2N-1}$, which is not the case in the radius-constrained algorithms where the nodes extension number is variable depending on the channel and the noise. Therefore, the polygon-constraint algorithm owns a very regular data flow and the corresponding control circuitry can be significantly simplified. Finally, the proposed scheme is also convenient in tuning the complexity-performance tradeoff by setting the total number of extended branches $L^{total} = \sum \boldsymbol{L}^{2N-1}$ to a smaller/larger number.

B. Application to Space-Division-Multiple-Access Mode

Detecting downlink SDMA signal is unique in that only the signals ($\tilde{x}_k$) dedicated to the k^{th} user are reserved. Obviously, the ML detection in [9] incurs great calculation waste by putting too much detecting effort on those to-be-discarded signals ($\tilde{x}_{l,l\neq k}$). To avoid this computation waste while maintaining an optimal performance, we propose to reuse the previously developed early-pruned FSD algorithm which conducts an ambitious search in the top layer while very simple single-node expansions in the remaining. To fit the original early-pruned FSD to the SDMA case, we make a slight modification such that the desired signal (i.e. $\tilde{x}_k$) is moved to the top layer of the search tree where near-ML result is guaranteed. The signal movement is accomplished by introducing a permutation matrix $\boldsymbol{P_k}$ as

$$\begin{aligned}\tilde{\boldsymbol{r}}_{\boldsymbol{k}} &= \tilde{\boldsymbol{H}}_{\boldsymbol{k}}^{c}[\tilde{\boldsymbol{W}}_{\boldsymbol{1}}, ..., \tilde{\boldsymbol{W}}_{\boldsymbol{K}}]\boldsymbol{P_k}\tilde{\boldsymbol{X}}_{\boldsymbol{k}} + \tilde{\boldsymbol{n}}_{\boldsymbol{k}} \\ &= \tilde{\boldsymbol{H}}_{\boldsymbol{k}}\tilde{\boldsymbol{X}}_{\boldsymbol{k}} + \tilde{\boldsymbol{n}}_{\boldsymbol{k}},\end{aligned} \tag{5}$$

where $\tilde{\boldsymbol{X}}_{\boldsymbol{k}} = [\tilde{x}_1, ...\tilde{x}_{k-1}, \tilde{x}_{k+1}, ..., \tilde{x}_K, \tilde{x}_k]^T$ is the transmit vector with signal $\tilde{x}_k$ being moved to the top layer and $\boldsymbol{P_k} = [\boldsymbol{p_1}, ...\boldsymbol{p_{k-1}}, \boldsymbol{p_{k+1}}, ..., \boldsymbol{p_K}, \boldsymbol{p_k}]$ is the permutation matrix, where $\boldsymbol{p_i}$ denotes an $N\times 1$ vector whose i^{th} element is one, but others are zeros. Taking $\tilde{\boldsymbol{H}}_{\boldsymbol{k}}$ as the equivalent channel matrix input, the matrix-permutated-FSD then conducts exactly the early-pruned FSD tree search to get the estimation result $\hat{\boldsymbol{X}}_{\boldsymbol{k}}$, in which $\hat{x}_k$ is the desired signal for user k, retained and outputted for further processing, while $\hat{x}_{l,l\neq k}$ are the signals intended for other users, discarded after detection.

C. Spatial-Diversity Signal Detection

LTE adopts the frequency domain version of the Alamouti code [2] as the basic spatial-diversity MIMO scheme. Let $\tilde{s}_n^i$ represent the complex signal transmitted from the i^{th} antenna in the n^{th} subcarrier, then:

$$\begin{bmatrix} \tilde{s}_1^1 & \tilde{s}_2^1 \\ \tilde{s}_1^2 & \tilde{s}_2^2 \end{bmatrix} = \begin{bmatrix} \tilde{x}_1 & \tilde{x}_2 \\ -\tilde{x}_2^* & \tilde{x}_1^* \end{bmatrix}, \tag{6}$$

where x^* represents the complex conjugate of x. Comparing to the SM mode, the detection of Alamouti signals is much simpler. For example, the complex-valued ML detection proposed in [10] decodes each transmitted symbol independently by means of maximal-ratio-combination (MRC), which owns a complexity linear to the constellation size M. Making use of the real-value/QR decomposition as well as the interference cancelation adopted in the aforementioned early-pruned FSD, we show in this paper that the complexity can be further reduced. In LTE-A, the application of the SFBC with more than two physical transmit antennas is converted, through precoding, into the case of two virtual antennas in order to provide robustness against spatial correlations [5]. In this condition, we derive our SD detection algorithm, as an example, for the 2×2 configuration. To treat $\tilde{x}_1$ and $\tilde{x}_2$ separately, we first perform element permutation to the signals in the second subcarrier as

$$[\tilde{r}_2^1, \tilde{r}_2^2]^T = \tilde{\boldsymbol{H}}_{\boldsymbol{2}} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \tilde{x}_1^* \\ \tilde{x}_2 \end{bmatrix} + \boldsymbol{w_2}, \tag{7}$$

where $\tilde{r}_n^i$ is the received complex signal at the i^{th} antenna in subcarrier n. The resulting real-valued representation of the QR-decomposed ($\boldsymbol{H_n} = \boldsymbol{Q_n}\boldsymbol{R_n}$) system model is then given as (the user index k is also neglected here in the SD mode)

$$\boldsymbol{y_n} = \boldsymbol{R_n}\boldsymbol{s_n} + \boldsymbol{v_n}, n \in [1, 2] \tag{8}$$

with

$$[\boldsymbol{s_1}, \boldsymbol{s_2}] = \begin{bmatrix} s_1^1 & s_2^1 \\ s_1^2 & s_2^2 \\ s_1^3 & s_2^3 \\ s_1^4 & s_2^4 \end{bmatrix} = \begin{bmatrix} \Re(\tilde{x}_1) & \Re(\tilde{x}_1) \\ \Im(\tilde{x}_1) & -\Im(\tilde{x}_1) \\ -\Re(\tilde{x}_2) & \Re(\tilde{x}_2) \\ \Im(\tilde{x}_2) & \Im(\tilde{x}_2) \end{bmatrix}, \tag{9}$$

$$\boldsymbol{y_n} = \begin{bmatrix} y_n^1 \\ y_n^2 \\ y_n^3 \\ y_n^4 \end{bmatrix} = \boldsymbol{Q}_n^H \begin{bmatrix} \Re(\tilde{r}_n^1) \\ \Im(\tilde{r}_n^1) \\ \Re(\tilde{r}_n^2) \\ \Im(\tilde{r}_n^2) \end{bmatrix}, \tag{10}$$

Taking advantage of the upper triangular feature of $\boldsymbol{R}$ and the property that $R_{n,(3,4)} = 0$ ($n \in [1, 2]$), we detect the real and imaginary parts of $\tilde{x}_2$ independently using the ML criteria after maximal-ratio combination of the received signal

$$\begin{aligned}\Im(\hat{x}_2) &= \arg \min_{\Im(\tilde{x}_2)\in\sqrt{M}} |(y_2^4 + y_1^4) - (R_{1,44} + R_{2,44})\Im(\tilde{x}_2)|^2, \\ \Re(\hat{x}_2) &= \arg \min_{\Re(\tilde{x}_2)\in\sqrt{M}} |(y_2^3 - y_1^3) - (R_{1,33} + R_{2,33})\Re(\tilde{x}_2)|^2.\end{aligned} \tag{11}$$

The inter-antenna interference introduced by $\tilde{x}_2$ is canceled from $y_n^i (n, i \in [1, 2])$ after we got the result $\hat{x}_2$

$$\tilde{y}_n^i = y_n^i - \sum_{l=3}^{4} R_{n,(i,l)}\hat{s}_n^l. \tag{12}$$

Since $R_{n,(1,2)} = 0$ ($n \in [1, 2]$), the symbol $\tilde{x}_1$ is detected with a similar method as that of (11)

$$\begin{aligned}\Im(\hat{x}_1) &= \arg \min_{\Im(\tilde{x}_1)\in\sqrt{M}} |(\tilde{y}_1^2 - \tilde{y}_2^2) - (R_{1,22} + R_{2,22})\Im(\tilde{x}_1)|^2, \\ \Re(\hat{x}_1) &= \arg \min_{\Re(\tilde{x}_1)\in\sqrt{M}} |(\tilde{y}_1^1 + \tilde{y}_2^1) - (R_{1,11} + R_{2,11})\Re(\tilde{x}_1)|^2.\end{aligned} \tag{13}$$

The detection results in the proposed algorithm are attained by decoding the real and imaginary parts of $\tilde{x}_1$ and $\tilde{x}_2$ separately. Apparently, this approach makes finding the results much simpler than the complex-valued ML algorithm, because the search zone has been reduced from M complex points to $\sqrt{M}$ real-valued points.

D. Simulation and Analysis

The detection performance of the proposed multi-mode MIMO algorithms is evaluated using computer simulations. The simulation setup is based on a simplified LTE-A downlink system with 64-QAM modulation and 2×2/4×4 MIMO configuration. The system bandwidth is assumed to be 5 MHz. The number of sub-carriers is 512, in which 300 sub-carriers are used for the information transmission. For the multi-path fading propagation, we applied the extended vehicular A (EVA) channel model described in [11]. There are 9 paths

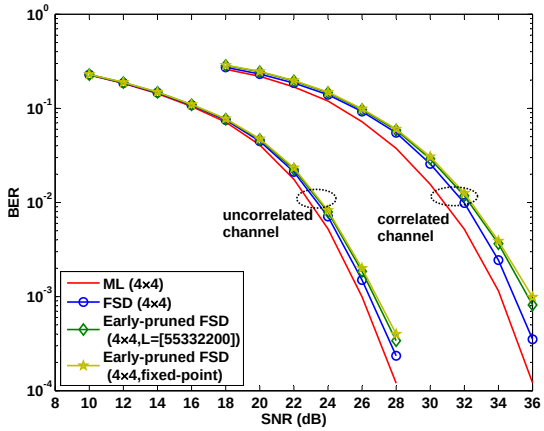

Fig. 2. Simulated BER performance in SM mode.

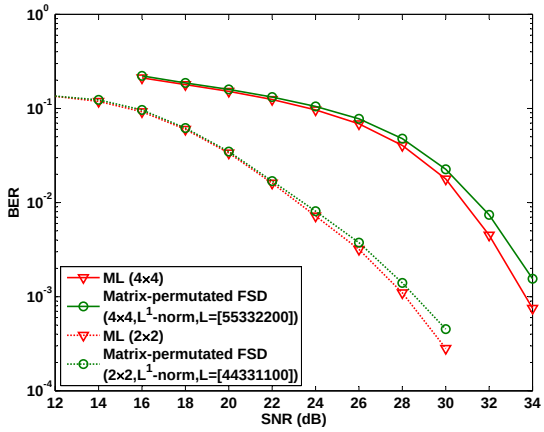

Fig. 3. Simulated BER performance in SDMA mode (moderate correlation).

in the channel with the largest path delay of 2510 ns. In addition, we considered two spatial correlation scenarios. In the moderately correlated channel, the antenna spacing at the BS is 4λ (the average correlation is 0.5) and the average spatial correlation at UE is 0.7. For the low correlation case, we assume a 10λ inter-antenna distance at the BS with an average correlation of 0.1 and the average correlation at the UE is 0.5. Finally, perfect channel estimation is assumed at the receiver.

It is shown in Fig. 2 that in uncorrelated channels, the proposed early-pruned FSD provides a performance very close to that of the original FSD for 4×4 MIMO when $\boldsymbol{L}^{2N-1} = [5, 5, 3, 3, 2, 2, 0, 0]$. In moderately correlated channels, the performance of the early-pruned FSD degrades by approximately 2dB compared to the ML detection at 10^{-3} BER. In this context, a larger L^{total} is suggested for performance enhancement. The fixed-point results in Fig. 2 are simulated by setting the data-path word-length (i.e., the word-length of PED) to 14 bits. The simulation results in Fig. 3 indicate that the correlation brings relatively smaller affection to the SDMA signal detection, where the proposed algorithm offers near-ML detection performance. This comes from the fact that the SDMA system exploits extra multi-user diversity by scheduling different UEs on different spatial streams over the same time-frequency resource. In addition, the unitary pre-coding in the SDMA mode also helps to combat correlations. Fig. 4 has also demonstrated that our proposed SD signal detection algorithm is capable of achieving near-optimum performance in low correlation channels.

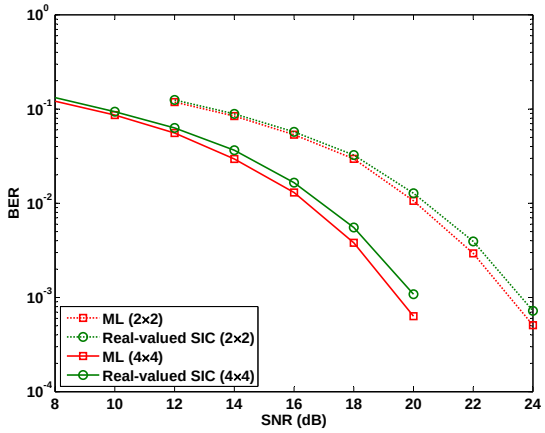

Fig. 4. Simulated BER performance in SD mode (low correlation).

IV. Reconfigurable VLSI Architecture

In this section, we present the architecture design of the proposed detector that supports the detection of SM, SD, and SDMA MIMO signals with up to 4×4 antenna and 64-QAM modulation. Configurations and/or schemes for different MIMO technologies, antenna numbers, constellations, as well as node-extension numbers (i.e., L^{total}) are demonstrated.

A. Overall Description

Exploiting the implementation-oriented detection algorithms, we design a multi-stage MIMO detector architecture that can be reconfigured on-the-fly to support multi-mode MIMO signal processing, which enables an efficient implementation. As shown in Fig. 5, the detector is partitioned into a pre-processing block and four stages of process elements (PEs), corresponding to the eight layers of the real-valued search tree in the case of 4×4 MIMO configuration.

Taking the channel matrix $\tilde{\boldsymbol{H}}^c$, the pre-coding matrix $\tilde{\boldsymbol{W}}$ and the received signal vector $\tilde{\boldsymbol{r}}$ as inputs, the pre-processing block executes matrix permutation, orthogonal real-value/QR decomposition, as well as the $\boldsymbol{y} = \boldsymbol{Q}^{\boldsymbol{H}}\boldsymbol{r}$ calculation. Each PE stage consists of three function blocks: an interference cancelation unit (ICU) that suppresses the inter-antenna interference introduced by the signals that previously have been detected, a node selection unit (NSU) that selects the L best nodes using the real-value zigzag enumeration method [8], and a PED calculation unit (PCU) which computes the accumulated partial Euclidean distance according to (4). In the detector, the PED calculation in two adjacent tree layers is carried out independently and simultaneously by one PE stage (e.g., the first PE stage is responsible for nodes at layer 8 and layer 7). This parallelism is due to the fact that $R_{i,i+1} = 0$ $(i = 1, 3, \ldots, 2N-1)$ when QR-decomposition is

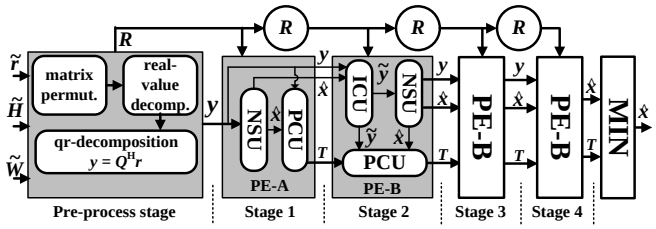

Fig. 5. Proposed VLSI architecture of the multi-mode MIMO detector.

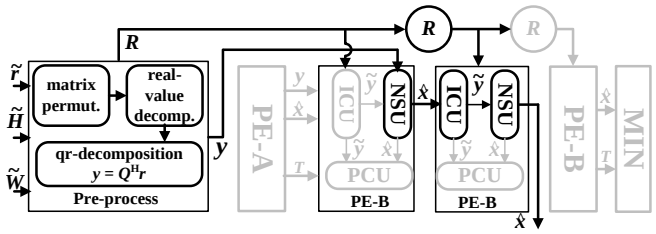

Fig. 6. Data path when the detector is configured to spatial-diversity mode.

performed to the orthogonal real-valued representation of the channel matrix $\boldsymbol{H}$ in (3) [12]. As a consequence, the number of PE stages is shorted to half of those pipelined detectors using traditional real-value decomposition [13], leading to an effective hardware saving by reducing the number of pipeline registers. Based on the number of branches to be expanded per father node and whether or not an ICU is applied, PEs in our design are divided into two types: PE-A and PE-B. PE-A is located at the first stage where multiple nodes are expanded according to the polygon-shaped constraint and interference cancelation is not employed, while PE-B performs the single-node expansion in the remaining 3 stages. Finally, the candidate-sharing technology [13] is also adopted in PE implementation to increase the hardware reuse ratio and ro support the detection of different modulation schemes.

In our design, the node computation at each stage is carried out in a parallel time-multiplexing fashion, in which the PE processes 10 father nodes simultaneously per iteration and finishes the computation of all the 20 nodes in two cycles when $\boldsymbol{L}^{2N-1} = [5, 5, 3, 3, 2, 2, 0, 0]$. Such architecture supports different extension number constraints by adjusting the iteration time. In case that a larger node extension number is needed for a better detection performance, we increase the iterations at the price of throughput reduction. The MIN block at the output stage generates the final detection result by selecting the branch with the smallest Euclidean distance.

B. Reconfigurability

This proposed detector architecture can efficiently realize various configurations of MIMO technologies. As described in Section III, the processes for the SM and the SDMA signals in our developed multi-mode detection algorithms are similar with a minor difference that matrix permutation is needed in SDMA to move the demanded signal to the first PE stage. Therefore, we simply by-pass the matrix-permutation in the pre-processing block in SM mode. The remaining function

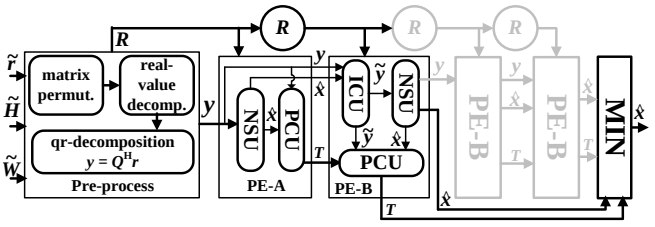

Fig. 7. Data path when the detector is configured to 2×2 mode.

blocks are completely reused in these two modes.

As a result, the key of the configurable detector design lies in how to reconfigure the architecture to the SD mode. According to (11)-(13), the tasks involved with the real-valued SIC SD signal detection can be decomposed into the following three major operations: channel matrix pre-processing, maximal-ratio combination (MRC)/single-node ML detection, and interference cancelation. Among these operation elements, the pre-processing block and the interference cancelation unit (ICU) already exist in the architecture and can thus be conveniently reused; the single-node ML detection can be accomplished by PE-B where the node with minimum partial Euclidean distance is found; the MRC block, however, is not included in the proposed architecture. Fortunately, unlike the complex ML algorithm [10], the maximal-ratio combination in our algorithm has been simplified to only consist of real additions, e.g., $(\pm y_1^i \pm y_2^i)$ and $(R_{1,ii} + R_{2,ii})$ in (11) and (13). Considering that only two nodes are calculated simultaneously in the SD operation, i.e., (11) or (13), and our parallel architecture is capable of multiple nodes processing, the MRC operation can be carried out by the adders assigned for the calculation of other nodes. Based on the above analysis, we give in Fig. 6 the detail data-path when the proposed detector is configured to the SD mode (the black parts are active data-paths, while the grey parts are disabled).

Furthermore, this multi-stage architecture is capable of supporting different antenna numbers by activating different PE stages. For example, the detector is configured to the 2×2 MIMO mode if we close the PEs in stage 3 and stage 4 and generate the detection result from stage 2 (shown in Fig. 7).

V. Implementations Results and Comparison

The designed multi-mode MIMO detector (without the pre-processing block) is modeled in Verilog Hardware Description Language (Verilog-HDL), synthesized using Synopsys Design Compiler with a 65-nm CMOS standard digital cell library, and routed using Cadence SoC Encounter/Silicon Ensemble. At the normal 1.2-V core power supply, the detector can work at a 165-MHz clock rate to support a 1.98-Gb/s peak throughput. The maximum allowable clock rate is obtained by post-layout simulation with delay information back annotated in SDF format. The peak throughput of our detector is achieved in the spatial-multiplexing mode and is formulated as

$$\text{Throughput} = f_c \times \frac{\log_2^M \times N}{C}, \tag{14}$$

TABLE I
COMPARISON OF THIS WORK TO PREVIOUS 64-QAM MIMO DETECTORS

	[14]	[15]	[13]	This Work
MIMO Modes	SM	SM	SM	SM/SD/ SDMA
Antenna Size	6×4	4×4	4×4	4×4
Modulation	64-QAM	64-QAM	64-QAM	64-QAM
Algorithm	OSIC with V-ML	K-Best	K-Best	Modified FSD
Process	0.13 μm	0.13 μm	0.13 μm	65 nm
Core Area	1.46 mm^2	0.9 mm^2	3.9 mm^2	0.25 mm^2
Gate Count	205 KG	114 KG	491 KG	88.2 KG
Clock Rate	71 MHz	282 MHz	137.5 MHz	165 MHz
Throughput	114 Mb/s	675 Mb/s	1.1 Gb/s	1.98 Gb/s
Power Consumption	30 mW @ 1.1 V	135 mW @ 1.3 V	127.2 mW @ 1.2 V	102.7 mW @ 1.2 V
Energy Consumption	263.2 pJ/b	200 pJ/b	115.6 pJ/b	51.8 pJ/b
Normalized Power	17.9 mW	57.5 mW	63.6 mW	102.7 mW
Normalized Energy	156.6 pJ/b	85.2 pJ/b	57.8 pJ/b	51.8 pJ/b

where M is the constellation size, N is the antenna number, f_c is the clock frequency, and C is the number of clock cycles needed for calculating the nodes in one layer. The parameter C equals to 2 in our design. Counting a two-input NAND gate as one equivalent gate, the multi-mode MIMO detector takes 88.2K equivalent gates and occupies a 0.25mm^2 core area (at 72% cell density). To investigate the power efficiency of the designed detector, power simulations are conducted for the post-layout design annotated with switching activities. Operating at 165-MHz clock rate with 1.2-V supply voltage and 25°C temperature, the detector consumes a maximal 102.7-mW core power when configured to the 4×4 spatial-multiplexing mode. The corresponding energy consumption to detect a bit is 51.8pJ/b.

We compare the proposed detector with recently reported 64-QAM MIMO detectors in Table I. Unlike previous works, our detector for the first time supports multiple MIMO signal transmission modes, namely, spatial-multiplexing (SM), spatial-diversity (SD), and space-division-multiple-access (SDMA). Despite the superiority of its reconfigurable multi-mode support property, the proposed detector consumes only 77.4% of hardware resources comparing to the best previous result [15]. From the detection throughput perspective, our detector is one of the very few detectors that achieve gigabit-per-second throughput to meet the requirement of next-generation cellular system. The energy consumption of our detector is also outstanding. Compared to the state-of-the-art in the open literature [13], 55.2% of the energy has been saved for single bit detection. To ensure a fair enough comparison, we consider the process factors. The power consumptions are normalized to the 65-nm technology and 1.2-V supply voltage as

$$\text{Power}_{norm} = \text{Power} \times (\frac{1.2V}{\text{Voltage}})^2 \times \frac{65nm}{\text{Technology}}. \quad (15)$$

The normalized energy consumption of our detector is still smaller than the state of art in open literature.

VI. CONCLUSION

This paper exploits the algorithm design and VLSI implementation of a multi-mode MIMO detector. At algorithm level, early-pruned FSD, real-valued SIC, and matrix-permutated-FSD are developed for detecting spatial-multiplexing, spatial-diversity, and space-division-multiple-access signals respectively. These proposed schemes reduce the complexity significantly and are still with the near-ML performance. In terms of the VLSI architecture design, a parallel multi-stage architecture has been developed for high-throughput and flexible reconfigurability. Implementation results have shown that the proposed detector outperforms the other reported works in terms of multi-mode support capability, hardware cost, and energy consumption.

REFERENCES

[1] Overview of 3GPP Release 10 V0.0.8 (2010-09) [Online]. Available: http://www.3gpp.org/ftp/Information/WORK_PLAN/Description_Releases/Rel-10_description_20100924.zip.

[2] Q. Li, *et al.*, "MIMO techniques in WiMAX and LTE: a feature overview," *IEEE Communications Magazine*, vol. 48, no. 5, pp. 86-92, May. 2010.

[3] C. Spiegel, *et al.*, "MIMO schemes in UTRA LTE, a comparison," *IEEE Vehicular Technology Conference (VTC Spring)*, pp. 2228-2232, May. 2008.

[4] S. Alamouti, "A Simple Transmit Diversity Technique for Wireless Communications," *IEEE J. on Selected Areas in Communications*, vol. 16, pp. 1451-1458, Oct. 1998.

[5] 3GPP Technical Specification 36.213 V9.1.0: Physical layer procedures (Release 9) [Online]. Available: http://www.3gpp.org/ftp/Specs/2010-03/Rel-9/36_series/36213-910.zip.

[6] L. G. Barbero and J. S. Thompson, "Fixing the complexity of the sphere decoder for MIMO detection," *IEEE Trans. on Wireless Communications*, vol. 7, no. 6, pp. 2131-2142, Jun. 2008.

[7] S. Chen, T. Zhang, and Y. Xin, "Relaxed K-Best MIMO signal detector design and VLSI implementation," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 15, no. 3, pp. 328-337, Mar. 2007.

[8] M. Wenk, *et al.*, "K-best MIMO detection VLSI architectures achieving up to 424 Mbps," *Proc. of IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1151-1154, May. 2006.

[9] S. Dragan, L. Angel, and B. P. Constantinos, "Design and experimental validation of MIMO multiuser detection for downlink packet data," *EURASIP J. on Applied Signal Processing*, vol. 11, pp. 1769-1777, 2005.

[10] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block coding for wireless communications: performance results," *IEEE J. on Selected Areas in Communications*, vol. 17, pp. 451-460, Mar. 1999.

[11] 3GPP Technical Specification 36.101 V10.1.1: User Equipment (UE) radio transmission and reception (Release 10)[Online]. Available: http://www.3gpp.org/ftp/Specs/archive/36_series/36.101/36101-a11.zip.

[12] J. Lofgren and P. Nilsson, "On MIMO k-best sphere detector architecture complexity reductions," *International Conference on Signal Processing and Communication Systems (ICSPCS)*, pp. 1-9, Dec. 2008.

[13] L. Liu, *et al.*, "A 1.1-Gb/s 115-pJ/bit configurable MIMO detector using 0.13-μm CMOS technology," *IEEE Trans. on Circuits and Systems-II*, vol. 57, no. 9, pp. 701-705, Sept. 2010.

[14] C.-J. Huang, C.-W. Yu, and H.-P. Ma, "A power-efficient configurable low-complexity MIMO detector," *IEEE Trans. Circuits Syst. I, Reg. Papers*, pp. 156-160, Jan. 2009.

[15] M. Shabany and P.G. Gulak, "A 0.13μm CMOS 655Mb/s 4×4 64-QAM K-Best MIMO detector," *Proc. of IEEE International Solid-State Circuits Conference (ISSCC)*, pp. 256-257/257a, Feb. 2009.