J Sign Process Syst
DOI 10.1007/s11265-015-1093-2

3.2 Gbps Channel-Adaptive Configurable MIMO Detector for Multi-Mode Wireless Communication

Farhana Sheikh[1] · **Chia-Hsiang Chen**[1] · **Dongmin Yoon**[1] · **Borislav Alexandrov**[1] · **Keith Bowman**[1] · **Anthony Chun**[2] · **Hossein Alavi**[1] · **Zhengya Zhang**[3]

Received: 24 February 2015 / Revised: 5 November 2015 / Accepted: 1 December 2015

Abstract A configurable, channel-adaptive K-best multiple-input, multiple-output (MIMO) detector for multi-mode wireless communications that adapts computation to varying channel conditions to achieve high energy-efficiency is presented. An 8-stage configurable MIMO detector supporting up to a 4 × 4 MIMO array and BPSK to 16-QAM modulation schemes has been implemented and simulated in 0.80 V, 22 nm Tri-gate CMOS process. Dynamic clock gating and power gating enable on-the-fly configuration and adaptive tuning of search radius K to channel response which results in 10 to 51 % energy-efficiency improvement over non-adaptive K-best MIMO detectors. During unfavorable channel conditions, the MIMO detector satisfies target bit error rate (BER) by setting K = 5. For favorable channel conditions, K is reduced to 1, where 22 nm circuit simulations show 68 % energy reduction. At 1.0GHz target frequency, the total power consumption is 15 mW (K = 1) to 35 mW (K = 5), resulting in energy-efficiency of 14.2pJ/bit (K = 1) to 44.7pJ/bit (K = 5) and 3.2Gbps throughput.

Keywords Multiple input · Multiple output (MIMO) detector · Channel-adaptive · LTE systems · IEEE 802.11n/ac · K-best · Sphere decoding · Multi-mode communication · Reconfigurable

✉ Farhana Sheikh
farhana.sheikh@intel.com

Chia-Hsiang Chen
chia-hsiang.chen@intel.com

Dongmin Yoon
dongmin.yoon@intel.com

Borislav Alexandrov
balexandrov3@gatech.edu

Keith Bowman
keithabowman@gmail.com

Anthony Chun
anthony.l.chun@intel.com

Hossein Alavi
hossein.alavi@intel.com

Zhengya Zhang
zhengya@umich.edu

[1] Intel Labs, Intel Corporation, 2111 NE 25th Ave, Mailstop: JF2-58, Hillsboro, OR 97124, USA

[2] Intel Labs, Intel Corporation, 3600 Juliette Lane, Santa Clara, CA 95054, USA

[3] Department of EECS, University of Michigan, Ann Arbor, 1301 Beal Ave, Ann Arbor, MI 48109-2122, USA

1 Introduction

New advanced wireless communication standards such as IEEE 802.11n/ac and 3GPP LTE Advanced Release 10/11 take advantage of a multiple-input, multiple-output (MIMO) communication schemes to enable increased spectral efficiency and high throughput data rates. For example, IEEE 802.11n allows for up to 4 × 4 antenna array (up to 4 streams with 4 transmit and 4 receive antennas); IEEE 802.11 ac calls for up to an 8 × 4 array (8 AP and 4 STA antennas); and 3GPP LTE Advanced release 10 [1] specifies up to an 8 × 8 antenna array. The enhancement in spectral efficiency and higher throughput comes at significant computational cost: work load profiling that we completed indicates that MIMO detection at the receiver can consume up to 42 % of the compute cycles in the physical layer baseband processing. The goal of this work is to present a configurable MIMO detector that can support multiple standards and adapt its computation to different channel conditions to improve energy-efficiency and reduce computational complexity.

Published online: 08 January 2016

A MIMO detector is used to determine the most likely data sequence transmitted by combining the multiple input streams at the receiver (see Fig. 1). As depicted in Fig. 1, the goal of the MIMO detector is to determine the transmitted vector which is a vector of all transmitted signals at each of the transmit antennas. The complexity of detection arises due to the fact that each received signal at each receive antenna is affected by all the transmitted signals and noise in the communication channel. There are a number of different algorithms used for MIMO detection as summarized in Table 1. Optimal Maximum Likelihood (ML) detection theoretically achieves the minimum bit-error rate (BER) but it is also the most expensive to implement since its complexity grows exponentially with increasing number of (transmit) antennas, resulting in large area and power requirements for large antenna array systems. Sub-optimal detectors, such as zero-forcing or minimum mean square error (MMSE), exhibit linear or polynomial computational complexity but behave poorly in terms of BER performance in low SNR conditions. Near-ML detectors, such as sphere decoders with termination criteria, can achieve significantly better BER performance in low SNR channels than linear detectors but at the cost of polynomial computational complexity.

In order to achieve high performance or low BER at reduced computational complexity relative to ML detection, the sphere decoding (SD) algorithm [2] is used to achieve near-ML detection performance by examining only a subset of solutions within a specified search radius [3]. Theoretically, the SD algorithm with termination criteria has cubic or polynomial computational complexity with respect to the number of antennas when the antenna array size is relatively small (less than 8×8) and the modulation order is not high (less than 64-QAM). Two types of algorithms are used to traverse the search space: depth-first search [4] and K-best breadth-first search [2]. In this work, we focus on K-best breadth-first search which enables regularized datapath interconnect and parallelism.

The rest of this paper is organized as follows. Section 2 reviews different MIMO detection algorithms, the SD algorithm, and advantages/disadvantages of different search criteria. Section 3 presents a summary of prior publications in this area. Section 4 and Section 5 detail the design and implementation of the channel-adaptive reconfigurable K-best MIMO detector. Section 6 presents simulation results in 22 nm CMOS and Section 7 concludes the paper.

1.1 Notation

Superscript T denotes the transpose. Real and imaginary parts of a complex number are denoted as $\Re[\cdot]$ and $\Im[\cdot]$. Upper- and lower-case bold letters indicate matrices and column vectors, respectively. $A_{i,\,k}$ indicates the (i,k) th entry of matrix $\boldsymbol{A}$. $\boldsymbol{I_N}$ denotes the $N \times N$ identity matrix, $0_{M \times N}$ denotes the complex valued $M \times N$ matrix with entries all identically zero, and $(1+\boldsymbol{j})_{M \times N}$ denotes the complex valued $M \times N$ matrix with entries all identically $(1+j)$. $\mathbb{C}$ is the set of complex numbers, $E\{\cdot\}$ denotes the statistical expectation, $\|\cdot\|$ denotes the 2-norm, and $[\cdot]$ denotes rounding to the nearest valid constellation point.

2 MIMO Systems and Design Challenges

Figure 1 shows the system architecture for a generic MIMO system. At the transmitter, a coded data stream is demultiplexed onto N_t antennas. The data streams may be modulated using different constellation sets of quadrature amplitude modulation (QAM) symbols. The use of multiple transmit channels allows the MIMO system to maximally use limited spectrum and deliver data at very high rates. At the receiver, the signals are processed by the analog front-end into

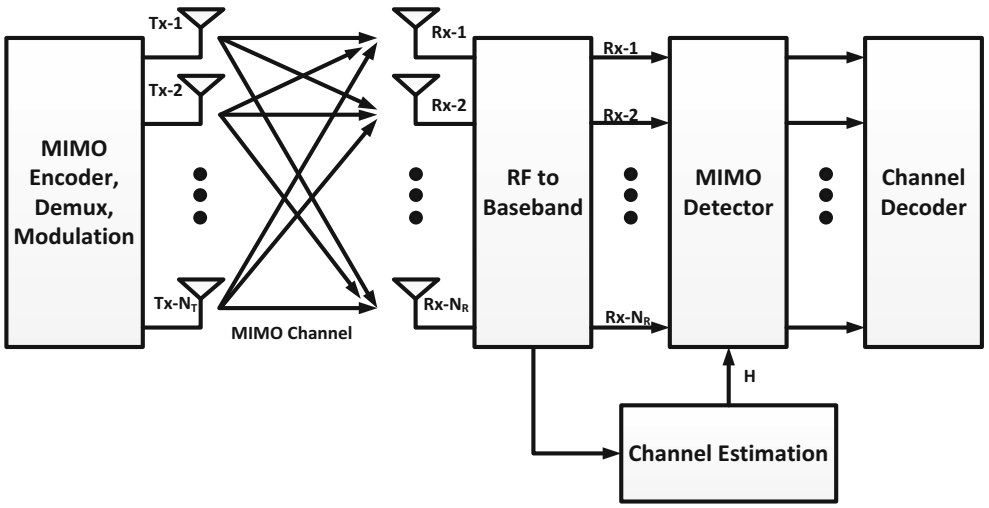


Figure 1 MIMO system overview.

Table 1 MIMO detection algorithm tradeoffs.

Detector type	BER performance	Computational complexity	In a Large MIMO system
Optimal Detectors • Maximum Likelihood (ML) • Sphere decoder (SD)	Optimal	Exponential	Heavy computation (prohibitively high area/power)
Sub-optimal detectors • Zero Forcing (ZF) • Minimum Mean Square Error (MMSE) • V-BLAST • Successive Interference Cancellation (SIC)	Poor	Linear, Polynomial	Inaccurate (error propagation)
Near-ML detectors • SD with termination criteria • K-best SD, LR-aided • Markov Chain Monte Carlo	Near optimal	Polynomial	Feasible (potentially high power/area)

in-phase (I) and quadrature (Q) components for each received signal at each antenna. The digital front-end (DFE) performs carrier frequency offset compensation, synchronization, and equalization. The MIMO detector at the receiver combines each of the complex received signals before further processing in the channel decoder.

As antenna arrays increase to 8×8 and higher to meet increasing performance requirements, the complexity of MIMO detection in the physical downlink baseband processing grows exponentially, taking up to 42 % of compute cycles which results in prohibitively large area and power requirements for the receiver. This assessment is based on work load profiling carried out in a physical layer simulator developed in MATLAB. The ideal MIMO detector has low area/power costs but is able to provide consistent required BER performance based on application needs in different types of channel conditions and for different wireless standards. Next we mathematically model the MIMO system which sets the basis for the hardware optimization and implementation.

2.1 MIMO System Model

A MIMO system with N_t transmit antennas and N_t receive antennas, operating in a symmetric M-QAM scheme, with $\log_2 M$ bits per symbol is modeled by:

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{v}, \tag{1}$$

where $\boldsymbol{s} = [s_1, s_2, \ldots s_{N_t}]^T$, $(s_i \in \mathcal{S})$ is the N_t-dimensional complex information symbol vector transmitted. The set $\mathcal{S}$ is the constellation set of the QAM symbols, and $\boldsymbol{y} = [y_1, y_2, \ldots, y_{N_r}]^T$ is the N_r-dimensional complex information symbol vector received. The equivalent baseband model of the Rayleigh fading channel between the transmitter and receiver is described by a complex valued $N_r \times N_t$ channel matrix $\boldsymbol{H}$. The vector $\boldsymbol{v} = [v_1, v_2, \ldots, v_{N_r}]^T$ represents the N_r-dimensional complex zero-mean Gaussian noise vector with variance σ^2 [5].

The ML detector estimates the transmitted signal by solving:

$$\hat{s} = \arg\min_{\tilde{\boldsymbol{s}} \in \mathcal{S}^{N_t}} \left\| \boldsymbol{y} - \boldsymbol{H}\tilde{\boldsymbol{s}} \right\|^2, \tag{2}$$

where $\hat{s}$ represents the candidate complex information symbol vector and $\breve{s}$ is the estimated transmitted symbol vector. Equation (2) represents the closest point problem [6] which is an exhaustive search through the set of all possible lattice points in $\mathcal{S}^{N_t}$ for the global best in terms of Euclidean distance between the received signal $\boldsymbol{y}$ and $\boldsymbol{H}\tilde{\boldsymbol{s}}$. Effectively, Eq. (2) describes a tree search where each of the lattice points in $\mathcal{S}^{N_t}$ is evaluated to find the path through the tree that minimizes the global error. Each transmit antenna contributes two levels to the search tree for real-domain MIMO detection: one real and one imaginary. Only one level in the tree is required for each antenna if complex-domain detection is employed. The goal is to find the closest vector $\hat{s}$ to the original transmitted symbol vector s given vector y. Equation (2) is generally non-deterministic polynomial hard (NP-hard).

The channel matrix $\boldsymbol{H}$ can be factored into two simpler matrices using QR factorization, where $\boldsymbol{Q}$ is a $N_r \times N_t$ orthonormal unitary matrix and $\boldsymbol{R}$ is a $N_t \times N_t$ upper triangular matrix such that $\boldsymbol{H} = \boldsymbol{Q}\boldsymbol{R}$. Equation (2) is then reformulated as:

$$\hat{s} = \arg\min_{\tilde{\boldsymbol{s}} \in \mathcal{S}^{N_t}} \left\| \breve{y} - \boldsymbol{R}\tilde{\boldsymbol{s}} \right\|^2, \tag{3}$$

where $\breve{y} = \boldsymbol{Q}^T \boldsymbol{y}$. Equation (3) is obtained by multiplying $\boldsymbol{y}$ in Eq. (1) by $\boldsymbol{Q}^T$ which results in $\boldsymbol{H}\tilde{\boldsymbol{s}} = \boldsymbol{Q}^T\boldsymbol{Q}\boldsymbol{R}\tilde{\boldsymbol{s}} = \boldsymbol{R}\tilde{\boldsymbol{s}}$. The diagonal elements of $\boldsymbol{R}$ are real and since it is upper triangular, symbol detection begins from the last row (i.e., top node in the tree) and moves up through the matrix in several steps until the unknowns in the first row (i.e., last level in the tree) are determined [3]. Several non-optimal and near-optimal algorithms have been proposed in the literature to traverse the search space to determine $\hat{s}$ such that the total error is minimized. They are summarized in the next section and in Table 1.

2.2 MIMO Detection Algorithms

Several algorithms have been proposed to address the complexity of MIMO detection in the receiver, offering different tradeoffs between power and performance. Table 1 gives an overview of MIMO detection techniques and qualitative tradeoffs. Among the MIMO detection techniques listed in Table 1 and shown in Fig. 2, the ML detector minimizes the BER performance through exhaustive search, but complexity grows exponentially with increasing number of antennas [3]. In contrast, linear detectors (LDs) – such as zero-forcing and MMSE detectors – and successive interference cancellation (SIC) detectors that have polynomial complexity suffer from significantly higher BER for the same signal-to-noise ratio (SNR). Markov Chain Monte Carlo methods [7] perform well in low SNR channel conditions but exhibit poor performance in high SNR channels. The SD detector using depth-first search without termination criteria results in optimal BER as in the case of a ML detector but requires prohibitively large area and power, and may never reach its solution in bounded time. Sphere decoding using breadth-first search with termination criteria can achieve near-optimal performance with polynomial computational complexity in the case where antenna array sizes and modulation orders are relatively small, i.e., 8×8 antenna array size and 64-QAM modulation. However, such detectors may still require high area or power as antenna arrays increase in size. In this work, we focus our implementation on SD detection with breadth-first search and limit the number of candidates evaluated in each step to K candidates with minimum partial distance from $\breve{y}$. The next section details the K-Best SD detection algorithm with termination criteria.

2.3 K-Best Sphere Decoding with Termination Criteria

The main advantage of depth-first search is that ML performance can be achieved and radius shrinking can be used to prune the search to meet performance constraints. The advantages of the K-Best breadth-first search over the depth-first search is the regular datapath interconnect and potential for parallelization to improve throughput. The K-Best tree search for MIMO [2] is a particular subset of breadth-first tree search algorithms.

It represents a middle ground between the complexity of the ML detector and linear detectors, finding a better sub-optimal solution to the closest point search than the linear detectors, with less complexity than the ML detector as shown in Fig. 2. Figure 2 shows that the K-Best SD solution can span the BER vs. SNR tradeoff space by adjusting K based on application BER requirements and channel conditions. In this work, we take advantage of this property to create a configurable detector that behaves as a linear detector when channel SNR is high by setting $K = 1$ and when SNR is low, K is increased to a predetermined maximum value. This enables adaptation of computational complexity to channel

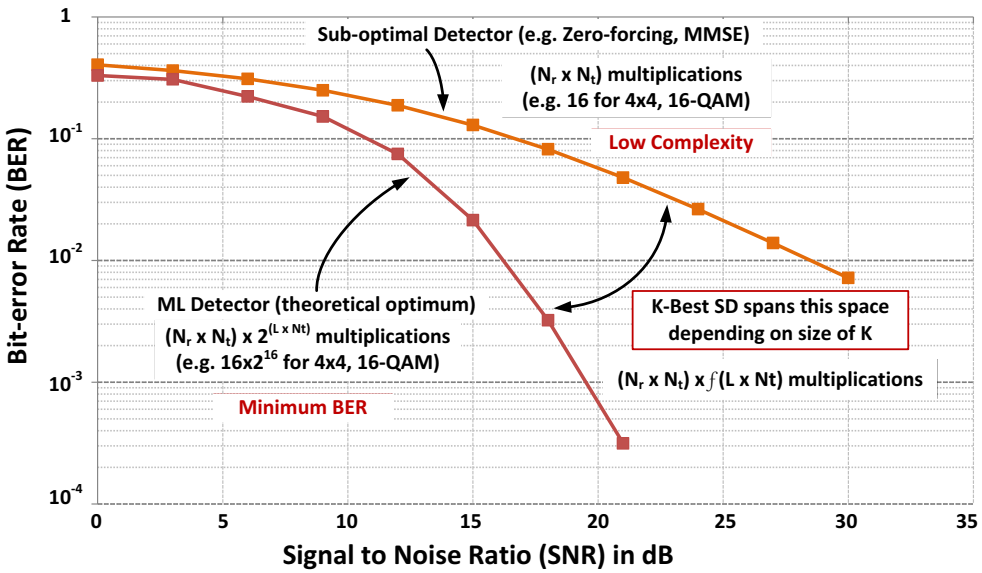


Figure 2 MIMO detection algorithm tradeoffs.

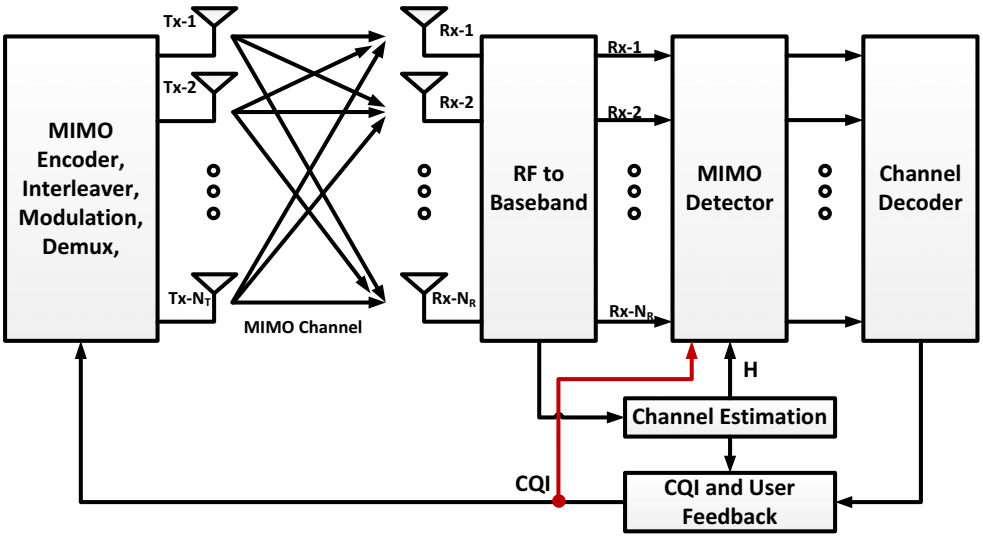


Figure 3 CQI feedback to channel-adaptive MIMO detector.

conditions, thereby having the ability to control power consumption of the MIMO detector in the receiver based on channel SNR. The adaptation to channel conditions is discussed after summarizing the *K*-Best SD algorithm as follows.

Since complex signals (i.e., I and Q) are being processed, a real domain realization of the tree search results in a single $2N_t$-dimensional search (i.e., one for the in-phase signal component and one for the quadrature component). The breadth first search begins with the N_t^{th} layer. For each nth layer, the algorithm computes the K best partial candidates $[\boldsymbol{s}_1^{(n)}, \boldsymbol{s}_2^{(n)}, \ldots, \boldsymbol{s}_K^{(n)}]$, where a partial candidate $\boldsymbol{s}_i^{(n)}$ represents the i^{th} path through the tree from the root node to level n, and is given by $\left[s_{i,n}{}^{(n)}, s_{i,n+1}^{n}, \ldots, s_{i,N_t}{}^{(n)}\right]^T$. The error at each step is measured by the partial Euclidean distance (PED). This is the accrued error at a given level of the tree, for a given path through the tree. Clearly, the K candidates at level n represent the K partial candidates with the minimum PED among all the children of the K candidates of the $(n+1$-st level, where the distance is calculated via:

$$PED_i^{(n)} = \sum_{j=n}^{N_t} \left[\left(\breve{y}_j - \sum_{k=j}^{N_t} R_{j,k} s_{i,k}^{(n)} \right)^2 \right]. \tag{4}$$

For an arbitrary level of the tree, the K best nodes are collected, and passed to the next level for consideration. In

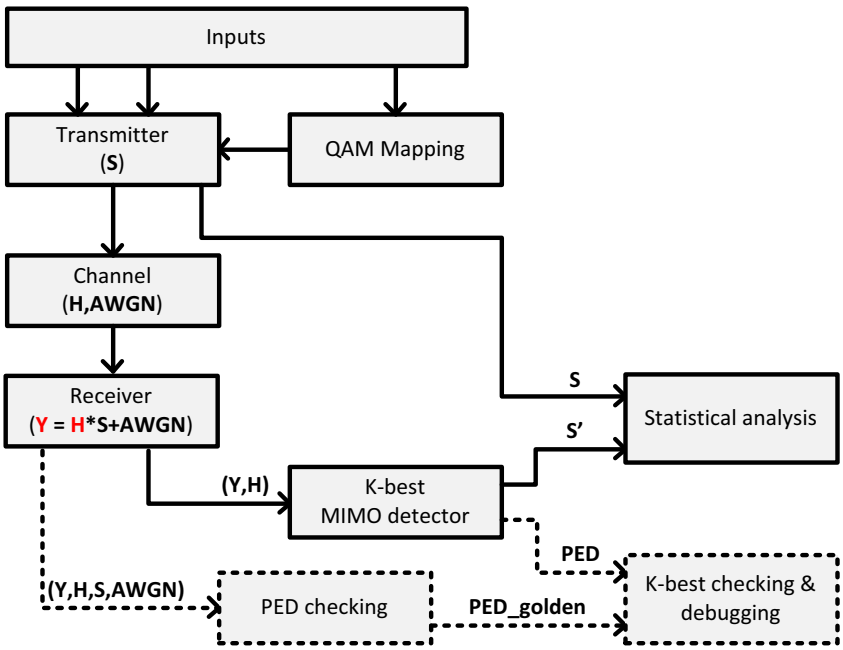


Figure 4 MATLAB model of channel-adaptive K-best MIMO detector.

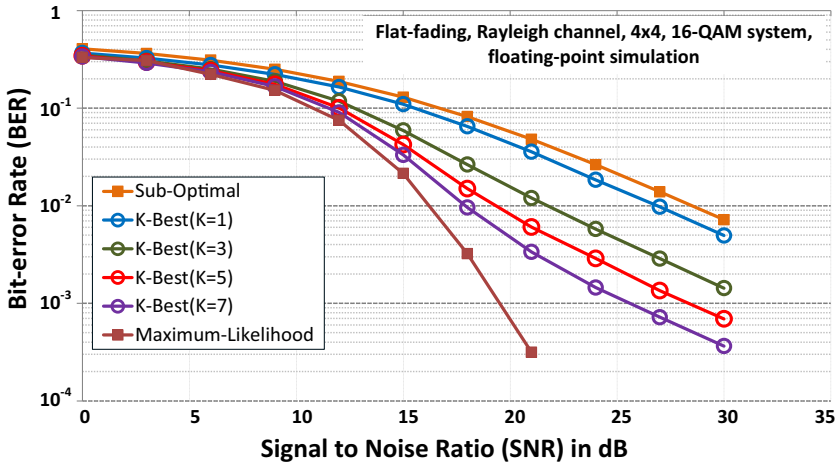


Figure 5 4 × 4, 16-QAM K-Best MIMO Simulation.

the last step, the K paths through the tree are evaluated to determine the path with the minimum total error.

3 Prior Art

Previous published work describes both K-best and depth-first search SD algorithms for MIMO detection. In [8], authors present a 192Mbps 4 × 4 MIMO for IEEE 802.11n supporting 16-QAM OFDM using Ordered Successive Interference Cancellation algorithm (OSIC). This design incorporated all of the baseband functions including AGC, Digital Down Conversion (DDC), IFFT, Channel Estimator, QR decomposition and MIMO detector for four RX antennas and was implemented in a 0.25 μm process.

The work in [4] presents two 4x4 depth-first search SD detectors implemented in a 0.25 μm process that achieve 73 and 169Mbps using a slightly a modified Schnorr-Euchner (SE) enumeration scheme that operates directly on complex constellations [9]. Both of these implementations support 16QAM and process one node per cycle. While the first implementation uses the l^2-norm, the second version uses the l^∞-norm that reduces the number of visited nodes and increases throughput substantially.

In [2], authors implement a K-best SD detector also using SE enumeration for 4 × 4 16QAM. Both hard and soft output decoders were implemented; the hard-output version was instantiated in a 0.35 μm process and consumed 626 mW at 100 MHz with at 2.8 V supply and provides a throughput of 53Mbps. The soft-output version was implemented in a 0.13 μm process and had a throughput of 106Mbps at a clock rate of 200 MHz; power numbers were not reported in this paper.

The most energy-efficient results published to date relating to SD detection are presented in [10] and [11] where a 16-core architecture using a hybrid K-best breadth-first, depth-first search approach results in a scalable MIMO detector with a hard-output SD kernel. The hybrid design combines the better decoding performance of the depth-first approach with the faster throughput of the breadth-first search method. In [10] the 90 nm design supports 2 × 2 to 16 × 16 configurations and BPSK to 64-QAM modulation schemes with estimated power consumption from synthesis of 33 mW and a peak data rate of 1.5Gbits/s with 16 × 16 64QAM in a 16 MHz bandwidth. In [11], the authors describe their implementation of the same kernel in 65 nm that dissipates 5.8 mW for the 4 × 4 LTE standard and also provides a peak data rate of 960Mbps for 8 × 8 64QAM and dissipates 13.93 mW with an energy efficiency of 15pJ/bit.

More recent work published in [12] presents a non-deterministic depth-first SD decoder in 65 nm supporting up to 4 × 4 64-QAM MIMO, resulting in 335Mbps throughput and power consumption of 36 mW at 333 MHz at 1.2 V. Note that this design enables the SD throughput to be adjusted based on SNR by tuning the parameter T corresponding to the number of leaves of the tree that are visited [13]. The throughput ranges from 296 Mb/s at 14.1 dB SNR to 807 Mb/s at 15.55 dB SNR.

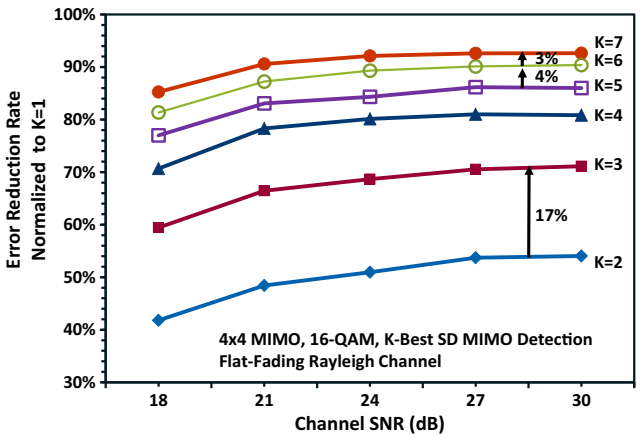


Figure 6 Error reduction rate for K-best MIMO.

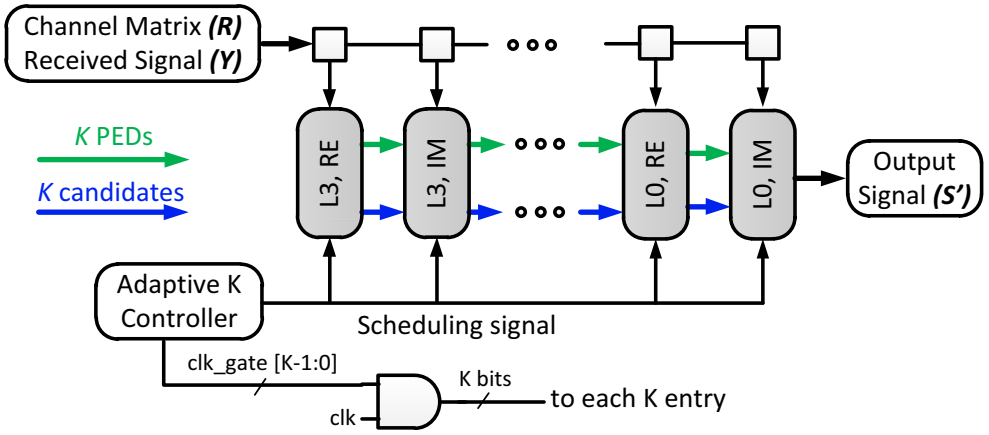


Figure 7 System overview of channel-adaptive K-best MIMO detector.

However, all these previously published works (except [12]) keep the search parameters fixed to some predetermined value and all computational units are active throughout the MIMO detection. The work presented in this paper distinguishes itself from previously published work in that our MIMO detector adapts its computation based on channel SNR conditions and application BER targets to reduce total energy and computational complexity by tuning *K*. In this work, a single detector is implemented that can approximate a linear detector when channel conditions are very good and a near-ML detector when channel conditions are poor.

4 Channel-Adaptive *K*-Best Sphere Decoder

This section details the new MIMO detecting approach based on adaptation to channel conditions to achieve high energy-efficiency. Channel-adaptive *K*-best SD detection automatically adjusts computation to fluctuating factors in the MIMO channel, such as wide dynamic SNR range, different fading paths, and varying BER target requirements driven by applications or user experience. In contrast to previous solutions, we implement feedback from channel and CQI estimation (see Fig. 3) to adapt the *K*-best SD MIMO detector to reduce computation and power during favorable channel conditions and application BER targets. In previous publications, the number of search paths are fixed throughout the entire lifetime of the *K*-best SD detection.

In the LTE standard, the channel quality indicator (CQI) is a 4-bit value that the user equipment must estimate which indicates channel quality which is sent to the transmitter to adapt the modulation scheme to channel conditions to maximize channel efficiency [1]. Reference [1] provides more details on how the 4-bit CQI value is determined and how often it must be relayed to the transmitter. In this work we leverage the CQI so that we can dynamically vary *K* depending on channel quality as shown in Fig. 3. This creates a closed feedback loop inside the receiver, where channel quality is fed back to the MIMO detector at frequent intervals to adapt the *K*-Best SD MIMO detection to channel conditions. If *K* is arbitrarily large (e.g., to positive infinity), the search becomes exhaustive as in ML detection. If $K=1$, the detector reduces its computation to linear MMSE detection. Tuning *K* between these two extreme limits allows traversal of the BER vs. SNR vs. computational complexity tradeoff space as shown in Fig. 2.

The control algorithm adjusts the search radius *K* based on whether channel SNR meets certain threshold requirements. A

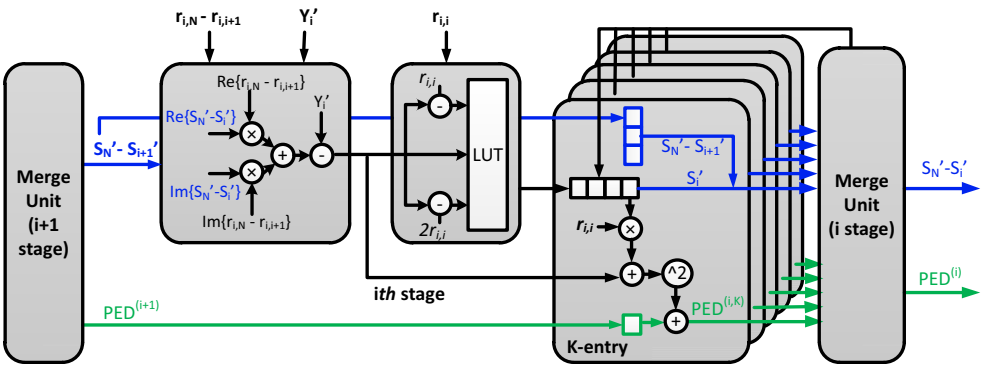


Figure 8 Detailed implementation of single stage.

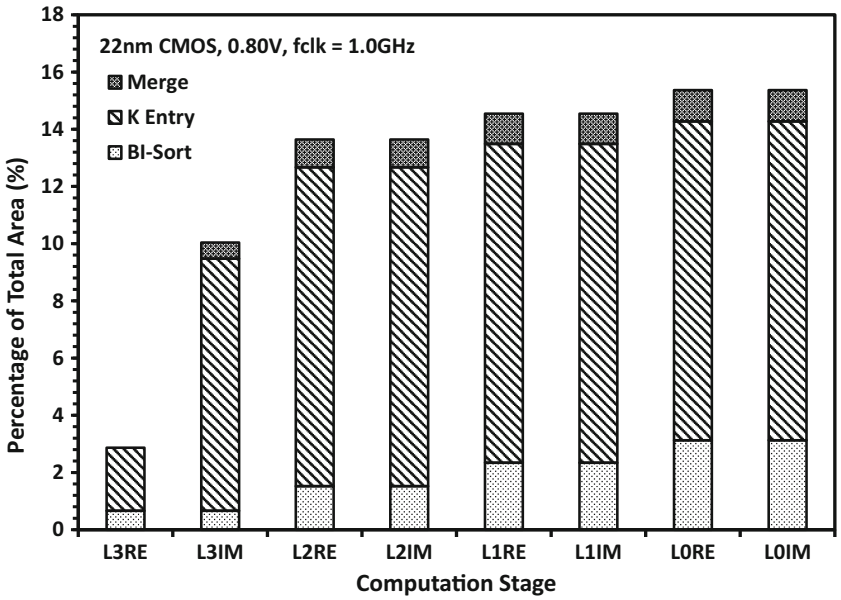


Figure 9 Percentage distribution of total area across computation stages.

small look-up table is used for this purpose where the 16 possible CQI values map to K values and this value is used to turn on/off computational units. For example in a 4×4 16-QAM MIMO system, when the channel has high SNR (>25 dB), then the search radius is set to $K=1$ and additional compute elements are power-gated off. In low SNR conditions (<10 dB), the search radius is set to $K=5$ and all compute elements are active.

The adaptive MIMO detector, supporting up to 4×4 array and from BPSK to 16-QAM modulation, targets a maximum $K=5$ based on the worst-case channel response. Figure 4 shows the Matlab model for the K-best MIMO detector which is used to evaluate its performance as shown in Fig. 5 as compared to a ML detector and a sub-optimal linear MMSE detector. The model allows users to change design parameters including size of the antenna array (e.g., 2×2, 2×4, 4×4, 8×8, and so on), the modulation scheme from BPSK to 256-QAM, and the search radius K. Simulation results shown in Fig. 5 for a 4×4 MIMO system with 16-QAM modulation under a standard Rayleigh fading channel with white Gaussian noise show that when $K=1$, the K-Best algorithm provides better performance than a sub-optimal detector.

Figure 5 also shows that by adjusting K the entire SNR vs. BER tradeoff space can be traversed using a single K-best SD MIMO detector. However, for the hardware implementation, the maximum K needs to be determined for the largest supported configuration. This simulation is carried out using the MATLAB model described in Fig. 4 where K is swept from $K=1$ to $K=7$ and the impact on BER is determined by how much the error decreases as K is increased. Simulation results in Fig. 6 show that when $K>5$, the incremental error reduction rate becomes marginal (<5 %) for the K-best SD detector supporting up to 4×4, 16-QAM, indicating that near-optimum results are possible without exhaustive search. Hence, for the detector presented in this work, maximum K is chosen to be 5. The hardware implementation in 22 nm CMOS is presented next.

5 Adaptive *K*-Best SD Detector Implementation

K-best SD MIMO detection is typically realized as a multi-stage pipeline since no trace-back is required. Using a breadth-first search implementation, pre-determined maximum K paths through the search tree are processed in parallel using a parallel architecture [10]. In contrast to previously published work, the MIMO detector presented here can vary K from 1 to 5 depending on channel SNR using the CQI as a controlling

Table 2 Summary of 22 nm CMOS synthesis results (4×4, 16-QAM).

Detecting Method	Channel-Adaptive K-best SD
Technology	22 nm CMOS, 0.80 V
Clock frequency	1.0GHz
Total core area	30,655 μm^2
Latency (K = 1 to K = 5)	42–70 cycles
Total power (K = 1 to K = 5)	15.1–35.5 mW
Energy per bit (K = 1 to K = 5)	14.2–44.7 pJ/bit
Throughput	3.2Gbps

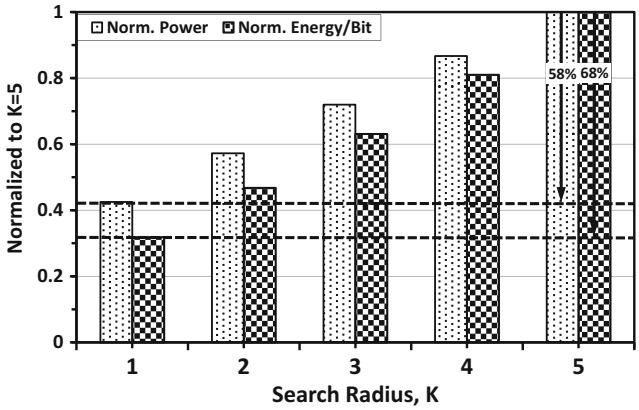


Figure 10 Effect of adapting K on power and energy efficiency.

input. We implement an 8-stage channel-adaptive K-best SD detector in System Verilog which supports up to a 4×4 antenna configuration and from BPSK to 16-QAM modulation. The design is implemented in 22 nm CMOS [14] and simulated for different configurations.

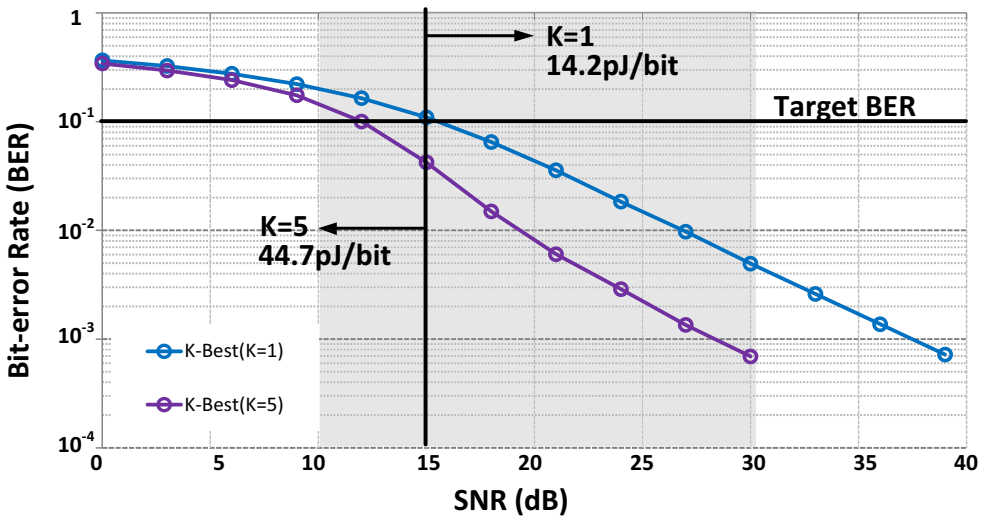
An overview block diagram of the channel-adaptive K-best MIMO detector is shown in Fig. 7. There are 8 stages in the pipeline to support both I and Q processing for each of the maximum 4 received signals at each of the 4 receive antennas. The CQI is used as an input to a look-up table to generate control signals that control the power-gating and clock-gating of compute elements. At the input of the MIMO detector, it is assumed that QR decomposition as described in Section 2 has already been completed, such that the R matrix is an input to the MIMO detector along with the received signal vector $\breve{y}$. The adaptive K controller takes the CQI as input to generate the clk_gate and scheduling signals used for power gating.

Each stage of the pipeline consists of four main blocks: Branch Interference (BI) cancellation Unit, a sorting circuit, storage of K-best candidates, and a Merge Unit. The details of these blocks are shown in Fig. 8. The BI unit calculates the interference from detected signals on the symbol candidate and mitigates the interference using multipliers and adders as shown in the first block. The BI unit constitutes the first part of the calculation necessary for SE enumeration and the second block which constitutes a simple sorting circuit saves the enumeration decision in the LUT which saves the current candidate list. The sorting circuit sorts the potential symbol candidates based on the SE enumeration scheme and PED. The K-entry storage block stores the K-best candidates that have been found so far and updates the PEDs based on Eq. (4) for each of the candidate paths through the tree. The Merge Unit selects the K-best symbol candidates for each cycle based on the updated PEDs for the path traversed through the tree and propagates them to the next stage.

An adaptive-K control unit uses a simple look-up-table (LUT) based on the 4-bit CQI value to output power-gating and clock-gating signals such as clk_gate to turn off the K-entry storage blocks that are not required in good channel conditions. For example, if the CQI is between 1101 and 1111 (indicating good channel conditions), then clk_gate signal is 00001 indicating that $K = 1$ and all but one of the K-entry storage blocks is required. The BI unit, the sorting circuit, the Merge Unit in each stage remain inactive for $(K_{max} - K)$ cycles when $K < K_{max}$. These units are clock-gated off during inactive cycles. Using the CQI signal from CQI estimation in the receiver as input to the MIMO detector creates closed-loop feedback inside the receiver as shown earlier in Figure.

Pipeline scheduling and time interleaving are implemented to eliminate stalls and improve throughput. Analysis shows that the K-entry block can introduce stalls and reduce the utilization of other three blocks if implemented without time-interleaving. For example, the Merge Unit takes K cycles to select K candidates, causing stalls for K cycles and resulting

Figure 11 High BER application channel adaptation case study.



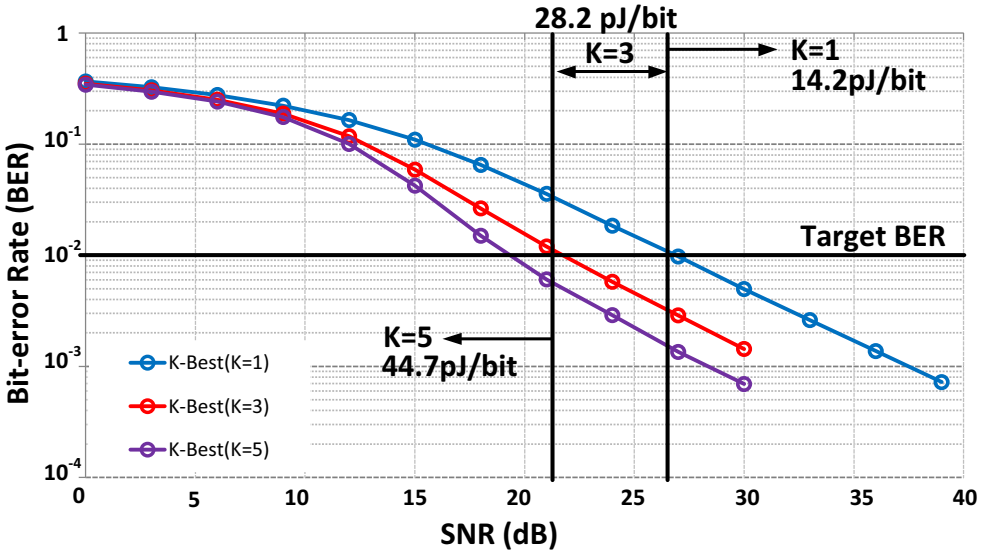


Figure 12 Low BER application channel adaptation case study.

in a throughput penalty. These stalls are prevented and resource utilization is improved by doubling the size of the *K*-entry block to accommodate time-interleaving. The adaptive *K*-Best SD MIMO detector described above is simulated in 22 nm CMOS and the results are presented in the following section.

6 Simulation Results in 22 nm CMOS

The configurable, channel-adaptive *K*-best SD MIMO detector, supporting up to 4 × 4 array configuration and modulation schemes up to 16-QAM, was simulated in 22 nm Tri-gate CMOS [14] for area and power estimates at 1.0GHz target frequency and 0.80 V supply voltage. The total power consumed in the MIMO detector varies from 15 mW ($K=1$) to 35 mW ($K=5$) with energy-efficiency of 14.2pJ/bit ($K=1$) to 44.7pJ/bit ($K=5$). The latency per process array is dependent on K and is given by $(35+7\times K)$: 42 cycles when $K=1$ and 70 cycles when $K=5$. At 1.0GHz, the simulated throughput is 3.2Gbps. The area distribution across stages and compute blocks is given in Fig. 9. The *K*-entry blocks across all stages consume 78 % of the total area whereas the BI unit, sorting circuits, and merge units make up the rest of the area. A summary of the simulation results using default activity and input toggle factors is given in Table 2.

The effect of adapting K to channel conditions is shown in Fig. 10. As K is reduced from 5 to 1, there is a 68 % reduction in energy per bit and 58 % reduction in total power. This simulation includes the reduced activity of blocks that are inactive when $K<K_{max}$. When K is dynamically adjusted to provide a constant BER over a varying SNR conditions, gains in energy-efficiency are realized by tuning K to channel conditions as outlined in the next section.

6.1 High BER Case Study

Consider a high BER application where the target BER is set to 10^{-1}. Assuming typical range of wireless received signal power is -70 to -90dBm and the thermal noise floor is set to -111dBm, we consider the 10 to 30 dB SNR range as shown in

Table 3 Performance comparison with published works.

ML and Near-ML MIMO detectors	Garrett ISSCC 04 [16]	Burg JSSC 2005 [4]	Shabany ISSCC 2009 [17]	Yang VLSI-C 2010 [11]	Winter ISSCC 2012 [12]	This work
Array size	4 × 4	4 × 4	4 × 4	8 × 8	4 × 4	4 × 4
Modulation	QPSK	16-QAM	64-QAM	64-QAM	64-QAM	16-QAM
Method	Exhaustive search	Depth first search	K-best	Hybrid breadth-first/depth-first	N/A	K-Best
Technology	180 nm	250 nm	130 nm	65 nm	65 nm	22 nm
Throughput (Mb/s)	28	73 (ML) 169 (N-ML)	655	960	335	3200
Energy/bit (pJ/bit)	N/A	N/A	200	15	55	14.2–44.7

Fig. 11 which displays BER curves for the channel adaptive K-best MIMO detector operating with $K=5$ and $K=1$. At an SNR ≤15 dB, we are forced to operate the detector with $K=5$, dissipating 44.7pJ/bit. When channel SNR is greater than15dB, K is adapted to one and energy dissipation reduces to 14.2pJ/bit, resulting in 51 % energy savings assuming that channel SNR varies uniformly from 10 dB to 30 dB.

6.2 Low BER Case Study

Figure 12 shows an application with a target BER lower than 10^{-2}. In this case, when channel SNR is less than 21 dB, the MIMO detector must operate with $K=5$ to achieve the desired BER target. When channel conditions improve slightly (21 dB < SNR ≤26 dB) then K can be reduced to 3, resulting in 28.2pJ/bit energy dissipation. If the channel SNR improves such that SNR > 26 dB, then K can be reduced to 1. The total energy reduction in this case is 23 % assuming channel SNR varies uniformly from 10 to 30 dB. Over a range of different target BER and channel conditions, an average of 10 to 51 % energy reduction can be achieved using channel-adaptive K-best MIMO detection.

7 Conclusions and Future Work

We have presented a unique approach to MIMO detection that uses channel information to adapt the search radius of a K-best SD detector, thereby enabling linear complexity in high SNR channel conditions, while maintaining BER performance during low SNR channel conditions. The estimated energy-reduction through simulations ranges from 10 to 51 %, depending on target BER and channel SNR. During unfavorable SNR channel conditions, the MIMO detector satisfies the target BER by operating at $K=5$. When channel SNR is high, 22 nm circuit simulations demonstrate 68 % reduction in energy-per-bit by setting $K=1$, while maintaining target BER. The channel-adaptive MIMO detector achieves 3.2Gbps throughput at a target frequency of 1.0GHz at 0.80 V supply and dissipates maximum 44.7pJ/bit when all compute elements are active. Comparison to prior work is given in Table 3.

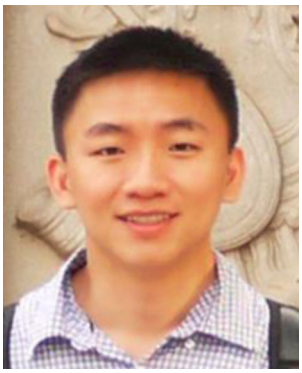
In future, we plan to extend and optimize the channel-adaptive MIMO detector to support up to 8×8 array configurations and 64-QAM modulation schemes. We also plan to modify this architecture to support soft-output MIMO detection. In [15] the authors develop a method to modify a hard-output SD MIMO detector to include soft-outputs at an area overhead of 58 %. In our proposed architecture we would also need to modify the number of times that a node is visited and the leaf enumeration procedure. As in [15] the PED calculation would remain the same and some small amount of additional memory would be required to track soft output metrics.

Acknowledgments The authors thank M. Wu, R. Iyer, K. Stewart, V. De, R. Forand, G. Taylor, V. Ilderem, W. H. Wang, I. Perez-Gonzalez, M. Jorgovanovic, M. Weiner, and G. Chen for encouragement and discussions.

References

1. 3GPP standards website: www.3gpp.org
2. Guo, Z., & Nilsson, P. (2006). Algorithm and implementation of the K-best sphere decoding for MIMO detection. *IEEE Journal on Selected Areas of Communications, 24*(3), 491–503.
3. Chiueh, T. D., Tsai, P. Y., & Lai, I. W. (2012). *Baseband Receiver Design for Wireless MIMO-OFDM Communications 2nd Edition*. Singapore: Wiley.
4. Burg, A. (2005). VLSI Implementation of MIMO Detection Using the Sphere Decoding Algorithm. *IEEE Journal of Solid-State Circuits, 40*(7), 1566–1577.
5. Mahdavi, M., & Shabany, M. (2013). "Novel MIMO detection algorithm for high-order constellations in the complex domain", in *IEEE Transactions on VLSI Systems*, 12:5.
6. Agrell, E., Eirksson, T., Vardy, A., & Zeger, K. (2002). Closest point search in lattices. *IEEE Transactions on Information Theory, 48*(8), 2201–2214.
7. Laraway, S. A., & Farhang-Boroujeny, B. (2009). Implementation of a Markov Chain Monte Carlo Based Multiuser/MIMO Detector. *IEEE Transactions on Circuits and Systems – I, 56*(1), 246–255.
8. Perels, D. et. al. (2005). "ASIC Implementation of a MIMO-OFDM Transceiver for 192Mbps WLANs", in *ESSCIRC Digest of Technical Papers*, Grenoble, France.
9. Schnorr, C. P., & Euchner, M. (1994). Lattice basis reduction: Improved practical algorithms and solving subset sum problems. *Math. Programming, 66*, 181–191.
10. Yang, C. H., & Markovic, D. (2009). A Flexible DSP Architecture for MIMO Sphere Decoding. *IEEE Transactions on Circuits and Systems-I, 56*(10), 2301–2314.
11. Yang, C. H., Yu, T. H., & Markovic, D. (2010). "A 5.8 mW 3GPP-LTE compliant 8x8 MIMO sphere decoder chip with soft-outputs", in *Technical Digest of Technical Papers for IEEE 2010 Symposium on VLSI Circuits*, Hawaii USA.
12. Winter, M. (2012). "A 335 Mb/s 3.9 mm2 65nm CMOS flexible MIMO detection-decoding engine achieving 4G wireless data rates", in *2012 ISSCC Digest of Technical Papers*, San Francisco, USA.
13. Adeva, E. P. et al. (2011). "VLSI Architecture for soft-output tuple search sphere decoding", IEEE Workshop on Signal Processing Systems.
14. Jan, C. H. et. al. (2012). "A 22 nm SoC platform technology featuring 3-D tri-gate and High-k/Metal gate, optimized for ultra low power, high performance and high density SoC applications", in *Proceedings of IEDM*, 4–7.
15. Studer, C., Burg, A., & Bolcskei, H. (2008). Soft-Output Sphere Decoding: Algorithms and VLSI Implementation. *IEEE Journal of Selected Areas in Communications, 26*(2), 290–300.
16. Garret, D. et. al. (2004). "A 28.8 Mb/s 4x4 MIMO 3G high-speed downlink packet access receiver with normalized least mean square equalization", in *Proceedings of 2004 I.E. International Solid-State Circuits Conference*, 15–19 February 2004, San Francisco, CA, pp. 420–536.
17. Shabany, M. & Gulak, P. G. (2009). "A 0.13µm CMOS 655Mb/s 4x4 64-QAM K-Best MIMO Detector", in *Proceedings of 2009 I.E. International Solid-State Circuits Conference*, 8–12 February 2009, San Francisco, CA, pp. 256–257.

Farhana Sheikh (IEEE SM'14, IEEE M'93) received the B.Eng. degree in Systems and Computer Engineering (with high distinction, Chancellor's Medal) from Carleton University, Ottawa, Canada, in 1993 and the M.Sc. and Ph.D. degrees in Electrical Engineering and Computer Sciences from the University of California, Berkeley, in 1996 and 2008, respectively. In 1993, Farhana was selected as one of 100 Canadians to Watch in Canada's national magazine, Macleans. From 1993 to 1994 she worked at Nortel Networks as a software engineer in firmware and embedded system design. From 1996 to 2001, she was at Cadabra Design Automation, now part of Synopsys as software engineer and senior manager. Her research interests include low-power digital CMOS design, energy-efficient high-performance circuits next generation wireless systems, wireless security, cryptography, and signal/image processing. She has co-authored several publications in VLSI circuit design and has filed six patents on accelerators for wireless communication, cryptography, and 3-D graphics at Intel Corporation. She worked as a Staff Research Scientist in the Circuit Research Lab, part of Intel Labs from 2008 to 2014. In 2014, she became Senior Staff Scientist and Technical Manager of Digital Communications Lab, in Wireless Communications Research, Intel Labs. Dr. Sheikh was a recipient of the Association of Professional Engineers of Ontario Gold Medal for Academic Achievement, the NSERC'67 scholarship for graduate studies, and the Intel Foundation Ph.D. Fellowship. In 2012, she received the ISSCC Distinguished Technical Paper award.



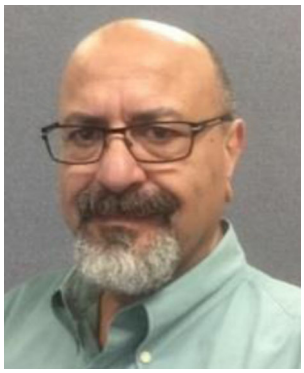
Chia-Hsiang Chen (IEEE S'10-M'14) received the B.S. degree in electrical engineering and computer science honor program from the National Chiao Tung University, Hsinchu, Taiwan, in 2008, and the M.S. and Ph.D. degrees in electrical engineering from the University of Michigan, Ann Arbor, in 2012 and 2014, respectively. He was a recipient of the Intel Ph.D. Fellowship in 2014. In 2015, he joined Intel Corporation as a senior research scientist in Digital Communication Lab in Wireless Communication Research, Intel Labs. His work has focused on error-resilient, low-power and high-performance VLSI circuits and systems for computing, communications and signal processing to leverage circuits, architectures, and algorithms.

Dongmin Yoon received the B.S. degree in electrical engineering from Seoul National University, Seoul, Korea, in 2009, and the M.S. and Ph.D. degree in electrical engineering from the University of Michigan, Ann Arbor, in 2011 and 2015, respectively. Since 2015, he has been with Intel Corporation where he is an analog engineer. His research interests include ultra-low power circuit design and wireless transceiver.

Borislav Alexandrov received his B.S. in Electrical and Computer Engineering from Cornell University in May 2010. Since August 2010 he has been working towards the PhD degree at the department of Electrical and Computer Engineering at the Georgia Institute of Technology. He has held several undergraduate and graduate internships with Intel, Qualcomm, and AMD. His research interests are in digital and analog integrated circuit design and the thermal and power management of next generation chips.

Keith A. Bowman (S'97-M'02) received the B.S. degree in electrical engineering from North Carolina State University, Raleigh, NC in 1994 and the M.S. and Ph.D. degrees in electrical engineering from the Georgia Institute of Technology, Atlanta, GA in 1995 and 2001, respectively. He is currently a Senior Staff Engineer in the Processor Research Team at Qualcomm in Raleigh, NC. He is responsible for researching and developing variation-tolerant circuit design solutions for enhancing the performance and energy efficiency of Qualcomm Snapdragon processors. From 2001 to 2004, he worked as a Senior Computer-Aided Design (CAD) Engineer in the Technology-CAD Group at Intel in Hillsboro, OR to develop and support statistical-based models, methodologies, and software tools to predict the performance and power variability of Intel microprocessors. From 2004 to 2012, he was a Staff Research Scientist in the Circuit Research Lab at Intel in Hillsboro, OR. In this role, he developed circuit techniques to mitigate the impact of device and circuit parameter variations on the performance and energy efficiency of Intel microprocessors. He has published over 70 technical papers in refereed conferences and journals and presented over 25 tutorials on variation-tolerant circuit designs.

Anthony Chun (IEEE M'90) received the BS, MS and PhD degrees from Stanford University in 1983, 1984 and 1990, respectively. Since 2000 he has been a Research Scientist at Intel Corporation. His research interests include wireless communication and speech processing algorithms and architectures.



Hossein Alavi As a senior research scientist in Intel Labs, is leading research on RF front end for cellular and connectivity radios. He joined Intel Corporation in 2004 as the Director of Radio Integration Lab (RIL) where he directed research on wireless transceiver SoCs on Intel's leading edge process technology. Prior to Intel, he was a university professor and founder and director of Wireless and Radar Research Center in Isfahan University of Technology in Iran, visiting professor at University of Toronto, founder of two startup companies developing broadband powerline modem technology, and senior scientist in Com Dev space communication in Canada. He received his Ph.D. in Electrical Engineering from Michigan State University in 1984. His fundamental work on Optimum Power Control for Cellular Radios created the foundation for power control techniques used extensively in Cellular Wireless protocols. Alavi has published over 30 research papers, holds 10 US patents, and is a reviewer for multiple IEEE Transactions.

Zhengya Zhang (S'02–M'09) received the B.A.Sc. degree in computer engineering from the University of Waterloo, Canada, in 2003, and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 2005 and 2009, respectively. Since September 2009, he has been with the Department of Electrical Engineering and Computer Science at the University of Michigan, Ann Arbor, where he is currently an Associate Professor. His research is in the area of low-power and high-performance VLSI circuits and systems for computing, communications and signal processing. Dr. Zhang was a recipient of the National Science Foundation CAREER Award in 2011, the Intel Early Career Faculty Honor Program Award in 2013, the David J. Sakrison Memorial Prize for outstanding doctoral research in EECS at UC Berkeley, and the Best Student Paper Award at the Symposium on VLSI Circuits. He serves an Associate Editor for the IEEE Transactions on Circuits and Systems–I: Regular Papers, the IEEE Transactions on Circuits and Systems–II: Express Briefs, and the IEEE Transactions on Very Large Scale Integration (VLSI) Systems.