Memristor Crossbar-Based Ultra-Efficient Next-Generation Baseband Processors

Geng Yuan[1], Caiwen Ding[1], Ruizhe Cai[1], Xiaolong Ma[1], Ziyi Zhao[1], Ao Ren[1], Bo Yuan[2] and Yanzhi Wang[1]
[1]Dept. of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY
[2]Dept. of Electrical Engineering, City University of New York, City College, New York, NY
{geyuan, cading, rcai100, xma27, zzhao37, aren@syr.edu, ywang393}@syr.edu, byuan@ccny.cuny.edu

***Abstract*—As one of the most promising future fundamental devices, memristor has its unique advantage on implementing low-power high-speed matrix multiplication. Taking advantage of the high performance on basic matrix operation and flexibilitys of memristor crossbars, in this paper, we investigate both discrete Fourier transformation (DFT) and miltiple-input and multi-output (MIMO) detection unit in baseband processor. We reformulate the signal processing algorithms and model structures into a matrix-based framework, and present a memristor crossbar based DFT module design and MIMO detector module design. For both designs, experimental results demonstrate significant gains in speed and power efficiency compared with traditional CMOS-based designs.**

***Keywords*—Memristor crossbar, DFT, MIMO detector, baseband processor.**

I. Introduction

Within recent two decades, wireless communication has become one of the fastest growing information technologies driven by the rapidly increasing demands for reliable information transmission and mobile data access. In modern mobile device, to meet the requirements of executing massive signal processing tasks in baseband, the baseband processors become the key enabler for the rapid development of the modern wireless communication.

However, under the explosive development of mobile Internet and Internet of Things (IoT), the prior baseband processor designs are facing unprecedented severe challenges. For instance, in order to support the expansion and enhancement of IoT applications to realize the zero latency and GB experience [1], the processing latency and peak data rate of the communication link in the emerging 5G era are required to achieve at most 1ms and at least 10Gbps, respectively, with 1000× reduction in energy consumption [2], which indicates the urgent demands on ultra-efficient designs of high-speed, short-delay, and low-power wireless hardware components. Unfortunately, as Moores law is approaching its ultimate limit, the benefit offered by the technology scaling will inevitably diminish, whereas the demands on lower latency and higher speed of wireless ICs will never cease. Consequently, the escalating computation-intensive tasks over the entire physical layer turn the baseband processor into a huge bottleneck in the modern wireless communication hardware design.

Unquestionably, to investigate next-generation device/circuit technologies beyond the traditional bulk CMOS and novel computing architectures beyond the traditional Von-Neumann machine is more necessary than ever. Fortunately, it is found out that a crossbar array of the recently discovered memristor devices (i.e., a memristor crossbar) can be utilized to perform matrix-vector multiplication in the analog domain and solve systems of linear equations in $O(1)$ time complexity[3–5]. Furthermore, since the memristor crossbar is compatible with CMOS technology, it can achieve ultra-high layout density as well as being reconfigured for different applications by changing the resistances of memristors (i.e., the memristance) [6–10]. Because of the outstanding performance on basic matrix operation and flexibility, it is anticipated that memristor crossbar may enable significant acceleration on advanced matrix operations (e.g. matrix inversion) which are very resource-consuming in CMOS design, thereby potentially triggering a revolutionary reshaping of future baseband processors to achieve orders-of-magnitude improvement in energy and performance to fulfill the requirement of ubiquitous IoT applications.

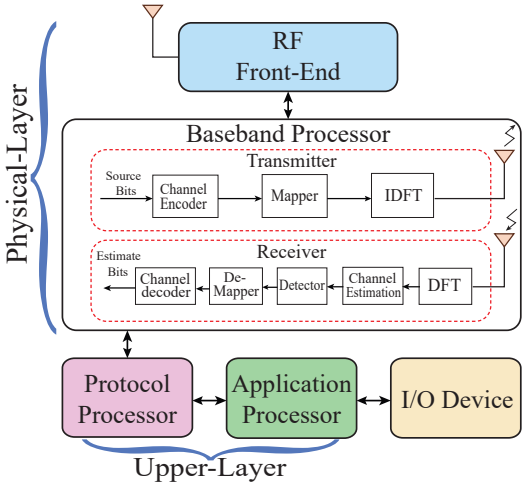

Fig. 1. Structure of wireless communication system

In this paper, we investigated ultra-efficient next-generation baseband processors which integrate CMOS circuits with memristor crossbars. We proposed the re-designs and optimizations of both signal processing algorithms and circuit implementations of the baseband processor. And our investigation mainly focus on the discrete Fourier transformation (DFT)/inverse DFT (IDFT), and multiple-input and multiple-output (MIMO) detection units, which are two critical blocks in baseband processor. We reformulated the signal processing algorithms and model structures into a matrix-based framework that can maximally exploit the advantage of memristor crossbars for computation.

II. Related Works

A. Wireless Baseband Processors

As the increasing demands for the crucial processing capacity to manage the communication and computational functions in the modern wireless devices, the performance of baseband processor is facing severe challenges. The baseband processor is a functional block that mainly responsible for processing

978-1-5090-6389-5/17/$31.00 ©2017 IEEE

the baseband signals in a wireless communication system including radio controlling, signal generation, modulation, and coding etc. Fig.1 demonstrates the functionality of the baseband processor in a wireless communication system. Due to its bi-directional transmission channel, the baseband processor can serve as the communication bridge between the Physical-layer and the Upper-layer. Moreover, it will also act as a pivotal role inside the physical-layer. The signal will be transferred between the baseband processor and the RF (radio-frequency) front-end for the signal processing purpose.

From the hardware design perspective, optimization of the key point of baseband processor has been applied in many up-to-dated studies [11–14]. Nevertheless, few of them focus on the baseband design by using the emerging technology, but the CMOS technology. Therefore, this paper presents a new viewpoint to reconsider the baseband processor design from the hardware level.

B. Memristor And Crossbar Memristor

A concatenation of memory resistors, has been proved as the fourth element in circuit by HP labs in 2008. It has many promising features, such as non-volatility, low-power, high integration density, and excellent scalability. More importantly, memristors have a unique property to record the historical profile of the excitations on the device [6–8]. The state of the memristor will remain unchanged, when the voltage equals to or is lower than the threshold voltage. Otherwise, when the voltage is higher than the threshold voltage, the state of the memristor will be changed. This critical feature makes memristor to be a desired candidate in matrix calculation field [10].

The Fig.2 shows the representative structure of the memristor. $\mathbf{WL}-$short for the Horizontal Word Line, and, $\mathbf{BL}-$short for the Vertical Bit Line, each pair of them will be connected by the memristor. By observing the crossbar, the entire structure can be treated as a matrix multiplication equation. The memristor behaves as a multiplication factor, the whole $\mathbf{WL}$ can be considered as the input vector, each $\mathbf{WL}$ is one of the dimensions in the input vector. Because the real input of the crossbar structure in $\mathbf{WL}$ is the voltage, and the output can be gained by collecting the current through each $\mathbf{BL}$, measured across register R_s with conductance of G_s. Assume that the memristor has a conductance of G_{ij}. The whole crossbar matrix multiplication process can be represented as the following equation: $\mathbf{V_0} = \mathbf{C} \cdot \mathbf{V_1}$ In the above equation, $\mathbf{V_1}$ stands for the input voltage vector in Horizontal Word Line, $\mathbf{C}$ stands for the connection matrix which can be constructed by the pro-programmed crossbar array, and Vo stands for the output vector which is collected from the Vertical Bit Line. By using this property, we can map the matrix multiplication onto memristor crossbar to significantly accelerate the multiplication operations.

III. Memristor Crossbar Based DFT module

In general, the majority of the wireless communication systems will contain the DFT/IDFT models which are responsible for transforming signals from time domain to frequency domain and vice versa. Naturally, there is no exception that the DFT/IDFT model becomes one of the critical parts in the baseband signal processor. However, despite the prior works [11], [12] of fast Fourier transformation(FFT) [15],

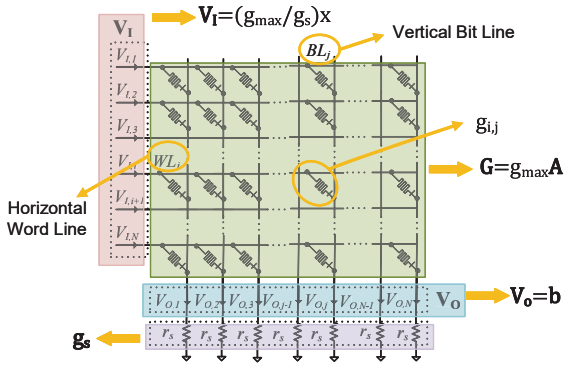

Fig. 2. Structure of Memristor Crossbar

which can reduce the computation complexity of the DFT from $O(N^2)$ down to $O(N \times logN)$ and the optimization works of FFT based on the selection of radix value, efficient pipelining etc., the COMS-based FFT designs cannot avoid the huge consumption on latency and power caused by the butterfly-like multi-stage structure.

To further optimize the DFT computation performance, a new structure needs to be found in order to break the limitations of FFT structure. From the perspective of matrix theory an N-point DFT output can be considered as a vector as $\mathbf{X} = [X_1, X_2, ...X_N]^\mathbf{T}$, which comes from $\mathbf{X} = \mathbf{W}\mathbf{x}$, where the N-point input signal $\mathbf{X} = [x_1, x_2, ...x_N]\mathbf{T}$, and an $N \times N$ dimensional transformation matrix $\mathbf{W}$. The matrix $\mathbf{W}$ can be expressed as:

$$\mathbf{W} = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & 1 & ... & 1 \\ 1 & \omega & \omega^2 & \omega^3 & ... & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \omega^6 & ... & \omega^{2(N-1)} \\ 1 & \omega^3 & \omega^6 & \omega^9 & ... & \omega^{3(N-1)} \\ ... & ... & ... & ... & ... & ... \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \omega^{3(N-1)} & ... & \omega^{(N-1)(N-1)} \end{bmatrix} \quad (1)$$

Where $\omega = e^{-2\pi i/N}$.

Base on the characteristics of DFT, as long as the number of input point N is fixed, the transformation matrix $\mathbf{W}$ is fixed as well. This indicates that we can map matrix $\mathbf{W}$ to memristor crossbar arrays to accelerate the DFT computation, and there is no need to upgrade the matrix for a certain point DFT during the entire process.However, the memristor crossbar arrays cannot be directly applied to the complex multiplications while all the $\mathbf{x}$, $\mathbf{X}$ and $\mathbf{W}$ consist of complex number in DFT computations, where the memristor crossbar can only represent real number. Moreover because of the resistance value is always larger than zero, so a memristor crossbar cannot represent the positive number and negative number at the same time. Thus, we proposed a new memristor crossbar-based DFT computing module which is an excellent solution for the complex multiplication and negative value issues. The proposed memristor crossbar-based DFT computing module is shown as Fig.3:

To calculate complex multiplication, we can always decompose them into real parts and imaginary parts. As shown in Fig.3, four groups of memristor crossbars represent real and imaginary parts $Re(\mathbf{W})$ and $Im(\mathbf{W})$ of transformation matrix W. Each group contains two memristor crossbars to represent the positive and negative terms, respectively, of the corresponding matrices ($Re(\mathbf{W})$ and $Im(\mathbf{W})$). By applying the input signal $Re(\mathbf{x})$ and $Im(\mathbf{x})$ to the memristor crossbar, we get

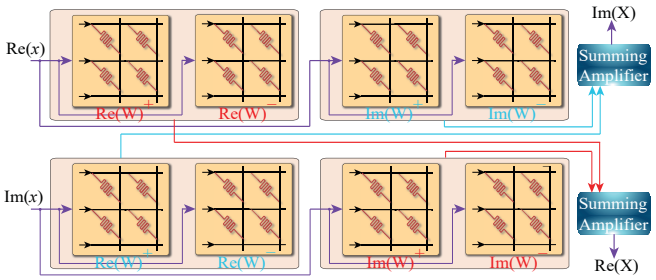

Fig. 3. Proposed memristor crossbar based DFT/IDFT implementation module

four product terms, which are $Re(\mathbf{W})Re(\mathbf{x})$, $Im(\mathbf{W})Im(\mathbf{x})$, $Im(\mathbf{W})Re(\mathbf{x})$, and $Re(\mathbf{W})$ $Im(\mathbf{x})$. To obtain the final output, summing amplifiers are applied to calculate $Re(\mathbf{X})$ and $Im(\mathbf{X})$. Note that the computation of IDFT has the similar way to the DFT, hence the IDFT module can be implemented in the similar memristor crossbar structure as the DFT module. Consequently, the DFT/IDFT computation are significantly accelerated by using memristor crossbar arrays.

IV. Memristor Crossbar Based MIMO Detector Model

Combining with spatial multiplexing, the MIMO technology has been integrated into most modern wireless communication standards, such as IEEE 802.11n [16], 3GPP-LTE [17] etc. By utilizing multiple antennas at both transmitter and receiver ends, the MIMO technology offers significantly higher data rate and spectral efficiency than the conventional single-antenna system with the cost of dramatic increase in the computational complexity of signal detection. Consequently, the design of low-complexity high-speed detector is a key challenge for MIMO-based wireless system. Some prior works have presented multiple digital CMOS-based implementations of various MIMO detecting algorithms, including maximum-likelihood (ML) [18], V-BLAST [19] and soft output-supported sphere decoding algorithms [20]. However, these small-scale MIMO-oriented designs have low hardware performance in the presence of emerging massive MIMO systems [21] which require a large number of antennas at the both sides of wireless link. In [22], [23], FPGA and ASIC designs of linear detectors have been reported and showed their potential suitability for massive MIMO systems. However, because these large-scale linear detecting algorithms involve extensive matrix operations that are inherently highly resource-consuming, the conventional digital CMOS design methodology may not be a perfect choice for such matrix-rich applications.

Inspired by the outstanding performance of memristor array on matrix computations, we propose to develop memristor-based MIMO detector. Here the linear detection with *minimum mean squared error* (MMSE) criterion, which can achieve near-optimal error performance in massive MIMO scenario, is adopted as detection algorithm. For a massive MIMO system with T and R antennas at the transmitter and receiver ends, respectively, the R-dimensional received signal vector $\mathbf{y}$ can be modeled as:

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n} \tag{2}$$

where $\mathbf{H}$ is the $R-by-T$ channel matrix that is obtained from channel estimation. $\mathbf{s} = [s_1, s_2, s_T]^T$ with $E[s_i] = E_s$

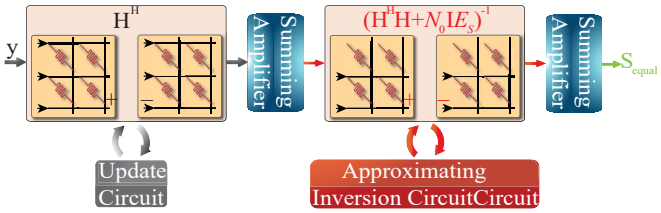

Fig. 4. Memristor crossbar-based equalization module for MIMO detector.

is the T-dimensional transmit signal vector, and $\mathbf{n}$ is the R-dimensional independent and identically distributed (i.i.d.) zero-mean white noise vector with variance N_0. Accordingly, the channel equalization, which is the most important step in the entire detection process, can be performed as:

$$\mathbf{s_{equal}} = (\mathbf{H^H}\mathbf{H} + N_0\mathbf{I}/E_s)^{-1}\mathbf{H^H}\mathbf{y} \tag{3}$$

where $\mathbf{s_{equal}}$ is the target symbol vector after equalization.

Since the digital CMOS-based design of equalization module demands numerous iterations and huge hardware cost, we propose a memristor crossbar-based equalization module for massive MIMO systems, with the over-all hardware structure shown in Fig.4. The input analog signal vector (from the DFT module) $\mathbf{y}$ will be applied on a memristor crossbar group (with two crossbars to represent positive and negative terms) that represents $\mathbf{H^H}$, and the output signal vector (after summing amplifier) is $\mathbf{H^H}\mathbf{y}$. $\mathbf{H^H}\mathbf{y}$ will be further applied on a memristor crossbar group representing $(\mathbf{H^H}\mathbf{H} + N_0\mathbf{I}/E_s)^{-1}$, and output will be the required $\mathbf{s_{equal}}$ vector.

It is important to note that matrices $\mathbf{H}$ and $(\mathbf{H^H}\mathbf{H} + N_0\mathbf{I}/E_s)^{-1}$ need to be updated when channel estimation is performed, which is different from the DFT module in which $\mathbf{W}$ is a constant matrix. However channel estimation and updating of matrices $\mathbf{H}$ and $(\mathbf{H^H}\mathbf{H} + N_0\mathbf{I}/E_s)^{-1}$ are performed with much less frequency compared with the computation of $\mathbf{s_{equal}}$ in MIMO detector, and hence will not incur significant power/performance overheads. Based on our proposed memristor crossbar based MIMO detector module, the computation complexity was significantly reduced by using the matrix multiplication property of memristor crossbar arrays comparing to the CMOS-based designs. Furthermore, by eliminating the demands of numerous iterations and huge hardware cost in CMOS-based MIMO detector designs, this memristor crossbar-based module can also achieve conspicuous low power consumption.

V. Simulation and Results

The simulation results of the proposed memristor crossbar-based DFT module and MIMO detector module. Our results on memristor crossbar-based DFT modules are accurately estimated using memristor models and those CMOS based FFT modules are from actual implementations. And from our prior work [24], it is demonstrated that, due to process variations, the actual memristance matrix of a memristor crossbar may be different from the theoretical values, and the relative calculation inaccuracy at the output is closely interrelated to signal point size. And the prior results show that when the signal size is larger than 256-point, the relative inaccuracy can be reduced below than 0.23%. From Fig.5 and Fig.6 the latency and power comparison of CMOS-based FFT module and memristor crossbar-based DFT module can be observed. Considering the inaccuracy issue, we choose three

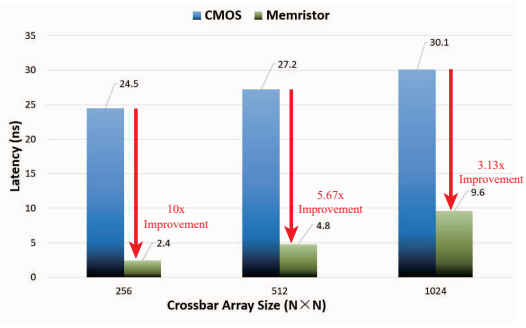

Fig. 5. Latency comparison between CMOS based FFT/IFFT implementation module and memristor crossbar DFT/IDFT based implementation module

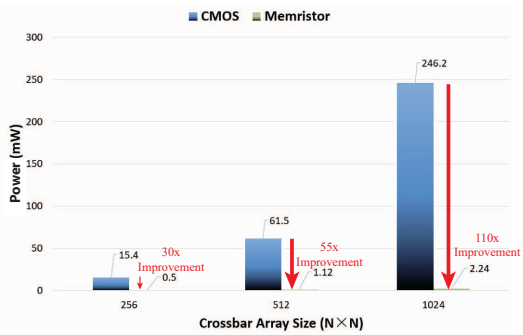

Fig. 6. Power efficiency comparison between CMOS based FFT/IFFT implementation module and memristor crossbar DFT/IDFT based implementation module

different cases starting from 256-point in our simulation. As shown in Fig.5 all the 256-point, 512-point, and 1024-point cases can significantly improve the latency performance. In power efficiency aspect, though the CMOS-based design used the FFT algorism, the memristor crossbar-based DFT module can still show its outstanding performance (up to 110×).

In our proposed MIMO detector module, the calculation of $(\mathbf{H^H H}+N_0\mathbf{I}/E_s)^{-1}$ motivated by the approximate computing approach for matrix inversion using Neumann series. And in Table.I our Matlab results illustrated that, as we estimated, because of the reduction of computation complexity, the proposed memristor crossbar-based MIMO detector can significantly reduce the latency by 102×. Moreover, in the MIMO detector module design, the memristor crossbar architecture showed its absolute advantage in high power efficiency. And there is a 144× energy efficiency improvement compared with the CMOS-based MIMO design.

TABLE I
COMPARISON RESULTS BETWEEN CMOS BASED MIMO DETECTOR MODULE AND MEMRISTOR CROSSBAR-BASED MIMO DETECTOR MODULE

	Power(mW)	Latency(ns)
Memristor	6.52	30.22
CMOS	938.06	3085.48
Improvement	143.79×	102.10×

VI. CONCLUSION AND FUTURE WORKS

This paper discussed new hardware designs of baseband processor, which is a critical block in wireless communication system by using the emerging memristor technology. An implementation of memristor crossbar-based DFT module design have been presented and compared with the CMOS-based FFT module in computing speed and power efficiency. Furthermore, we presented a MIMO detector module based on memristor crossbar architecture to challenge the CMOS-based module. Experimental results demonstrate significant gain in speed and power efficiency compared to traditional CMOS-based (FFT) design in both the DFT module and MIMO detector module.

REFERENCES

[1] "5G use cases and requirements," *Nokia White Paper*, May 2015.
[2] "5G: New air interface and radio access virtualization," *Huawei White Paper*, Aprial 2015.
[3] M. Hu, H. Li, Q. Wu, and D. Rose, "Hardware realization of neuromorphic bsb model with memristor crossbar network," *IEEE Design Automation Conference (DAC)*, pp. 554–559, 2012.
[4] M. Sharad, D. Fan, and K. Roy, "Ultra low power associative computing with spin neurons and resistive crossbar memory," *Proc. of Design Automation Conference (DAC)*, 2013.
[5] S. Yu and Y. Cao, "On-chip sparse learning with resistive cross-point array architecture," *ACM Great Lakes Symposium on VLSI (GLSVLSI)*, 2015.
[6] B. Liu, Y. Chen, B. Wysocki, and T. Huang, "The circuit realization of a neuromorphic computing system with memristor-based synapse design," *Neural Information Processing*, pp. 357–365, 2012.
[7] D. B. Strukov, G. S. Snider, D. R. Stewart, and R. S. Williams, "The missing memristor found," *Nature*, vol. 453, pp. 80–83, 2008.
[8] J. Liang, S. Yeh, S. Wong, and H.-S. P. Wong, "Effect of wordline/bitline scaling on the performance, energy consumption, and reliability of cross-point memory array," *ACM Journal on Emerging Technologies in Computing Systems*, vol. 9, no. 1, 2013.
[9] S. Kvatinsky, D. Belousov, S. Liman, G. Satat, N. Wald, E. Friedman, A. Kolodny, and C. Weiser, U, "Magic-memristoraided logic," *IEEE Trans. On Circuits and Systems II*, 2014.
[10] S. Jo, T. Chang, I. Ebong, B. Bhadviya, P. Mazumder, and W. Lu, "Nanoscale memristor device as synapse in neuromorphic systems," *Nano letters*, vol. 10, no. 4, pp. 1297–1301, 2010.
[11] Y. Lin, H. Liu, and C. Lee, "A 1-GS/s FFT/IFFT processor for UWB applications," *IEEE J. of Solid-State Circuits*, vol. 40, no. 8, pp. 1726–1735, 2005.
[12] S. Tang, J. Tsai, and T. Chang, "A 2.4-GS/s FFT processor for OFDM-based WPAN applications," *IEEE Trans. on Circuits and Systems II: Express Briefs*, vol. 57, no. 6, pp. 451–455, June 2010.
[13] C. Studer, S. Fateh, and D. Seethaler, "ASIC implementation of softinput soft-output MIMO detection using MMSE parallel interference cancellation," *IEEE J. Solid- State Circuits*, vol. 46, no. 7, pp. 1754–1765, July 2011.
[14] S. Chen, T. Zhang, and Y. Xin, "Relaxed k-best mimo signal detector design and vlsi implementation," *IEEE Trans. on VLSI System*, vol. 15, pp. 328–337, Mar 2007.
[15] A. Oppenheim, R. Schafer, and J. Buck, *Discrete-Time Signal Processing, 2nd ed.* NJ: Prentice Hall, 1998.
[16] T. Paul and T. Ogunfrunmiri, "Wireless lan comes of age: Understanding the ieee 802.11 n amendment," *IEEE Circuits and Systems Magazine*, vol. 8, no. 1, pp. 28–54, 2008.
[17] J. Lee, J.-K. Han, and J. C. Zhang, "Mimo technologies in 3gpp lte and lte-advanced," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, no. 1, pp. 1–10, 2009.
[18] D. Wubben, R. Bohnke, V. Kuhn, and K.-D. Kammeyer, "Near-maximum-likelihood detection of mimo systems using mmse-based lattice reduction," in *Communications, 2004 IEEE International Conference on*, vol. 2. IEEE, 2004, pp. 798–802.
[19] G. Ginis and J. M. Cioffi, "On the relation between v-blast and the gdfe," *IEEE Communications letters*, vol. 5, no. 9, pp. 364–366, 2001.
[20] Z. Guo and P. Nilsson, "Algorithm and implementation of the k-best sphere decoding for mimo detection," *IEEE Journal on selected areas in communications*, vol. 24, no. 3, pp. 491–503, 2006.
[21] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive mimo for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, 2014.
[22] P. Murphy, F. Lou, A. Sabharwal, and J. P. Frantz, "An fpga based rapid prototyping platform for mimo systems," in *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on*, vol. 1. IEEE, 2003, pp. 900–904.
[23] A. Burg, M. Borgmann, M. Wenk, M. Zellweger, W. Fichtner, and H. Bolcskei, "VLSI implementation of MIMO detection using the sphere decoding algorithm," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 7, pp. 1566–1577, 2005.
[24] R. Cai, A. Ren, and Y. Wang, "Memristor-based discrete fourier transform for improving performance and power efficiency," *IEEE Computer Society Annual Symposium on VLSI(ISVLSI)*, July 2016.