A Study using Hardware Accelerator to Meet the Engineering Requirements of 5G

Neevee C.Thomas
Communication Engineering
SCMS School of Engineering & Technology
Ernakulam, Kerala, India

Shibu R M, Haneesh Sankar
Strategic Electronics Group(SEG)
Centre for Development of Advanced Computing(C-DAC)
Trivandrum, Kerala, India

Abstract— **In this paper we focused on to reduce the latency time for 5G applications using hardware accelerator. In current 4G systems the roundtrip latency period is about 15mS, The 5G applications will need a roundtrip latency period less than about 1 ms, faster than 4G.The latency constraints may have major role in the design of a 5G communication system. 5G networks will require tight inter-coordination between multiple network elements and tight sharing of spectrum.OFDMA can only operate if strict time and frequency synchronization between users and a base station is achieved. There is a need for a filtered, multicarrier approach with reduced side-lobe levels of the waveform which could minimize inter-carrier interference (ICI) in 5G, FBMC (filter bank multi-carrier) generalizes traditional orthogonal frequency-division multiplexing (OFDM) schemes. The proposed work consists of latency analysis of filter bank based multicarrier (FBMC) transmitter and receiver with and without a hardware accelerator. This waveform is a possible candidate for 5G. The result of this experiment will be one of the references to finalize engineering requirements of 5G wireless communication system in terms of latency.**

Keywords—Filter Bank Multi Carrier; OFDM; Hardware Accelerator; FPGA; IP core generation

I. INTRODUCTION

The next generation of wireless networks is named as 5G. 5th generation mobile communication technology is represented as 5G technology.Mobile within high bandwidth can be made using 5G technology [2]. Due to the presence of extra features such as cognitive radio technology, high data rate, effective billing system, etc 5G technology will be the most dominant technology in the near future. The wirelesstechnology have been changed from 0G to 5G in the last few years. The advanced implementation of this technologiesare the major constraint in the execution work of upcoming 5G which is the wireless technology. When thinking of 5th generation wireless networks, most people think in terms of more bandwidth. However, this is only part of the story [1],[4]. The key technical advancements will come from two emerging concepts: cooperative networking and coexistence. Cooperative networking requires network elements to coordinate closely together, on a frame-by-frame basis and to transmit/receive data on shared resources. This will dramatically improve coverage and quality, especially in

difficult scenarios, which are currently underserved by modern networks. In a large number of cases, this coordination may involve small-cells. Coexistence requires the 5G cellular network to share spectrum with other radio access technologies in a common geographic area. Multiple radio access technologies (RATs) using common spectrum operate in coordination with each other.

Till now the standards of 5G is not set by the wireless industry, and 5G remains as concept, large ambivalence exists across 5G. The wireless industry is trying to achieve some key goals with 5G they are

- Appreciably faster data speeds: At present 4G networks have the capacity to achieve peak download speed of one gigabit per second, in actual practice it is never that fast. While we use 5G the speed can be improved to 10Gbps.

- Ultralow latency: Currently with 4G, the latency is 15 milliseconds, but 5G will reduce that to about one millisecond [2]. 'Latency' refers to the time it takes one device to send a packet data to another device. For industrial applications latency plays a greater role.

- A more "connected world": in the next 10 years the internet of things is expected to grow very fast and it will require a network that can coordinate billions of connected devices and gadgets. The major goal behind 5G is to provide such a capacity also it must be able to assign the bandwidth according to the application needs and user.

The multiple access and signal formats i.e., the waveformdesign, have changed significantly at each cellular generationand to a large extent they have been each generation's definingtechnical feature. They have also often been the subject of big intellectual and industrial disputes, which have playedout in the wider media and society. The 1G approach, which is based on analog frequency modulation with FDMA, transformed into more compact digitalformat for 2G and, although it employed both FDMA andTDMA for multiple access. It was generally known as "TDMA"due to the novelty of time multiplexing. In between, a niche read spectrum/CDMA standard which was developed by Qualcomm to compete for 2G became the authoritative approach to all global 3G standards. For the high speed data the limitations of CDMAbecame unavoidable , there was a discrete but distinct retreat back toward TDMA, with very low spectrum spreading

retained and with the important additionof channel-aware scheduling .Due to the increasing signalbandwidths needed to support data applications, orthogonalfrequency-division multiplexing (OFDM) was unanimouslyadopted for 4G in co-occurrence with scheduled FDMA/TDMAas the virtues of orthogonality were viewed with renewedappreciation.

OFDM is the unquestionable frontrunner for 5G. Multicarrier communication techniques are among the most successful air interface transmission techniques. Out of them orthogonal frequency division multiplexing (OFDM) is identified as most suitable technique for broadband wireless communication. Efficient use of spectrum, robustness against frequency selective fading, resistance to both inter symbol interference (ISI) and inter carrier interference (ICI) (with the aid of guard intervals and cyclic prefixes), capability of being implemented using FFT (Fast Fourier Transform) techniques can be highlighted as some of the key merits. However higher peak to average power ratio, higher sensitivity for carrier frequency offset (CFO) errors, higher amount of side lobes, inevitable overhead due to cyclic prefixes and additional time consumed due to insertion of guard intervals are among critical and limiting drawbacks.

However, some drawbacks do exist that could possibly become more pronounced in 5G networks. Due to the addition of uncorrelated inputs in the IFFT the envelops are nearly Gaussian due to this the peak-to-average-power ratio (PAPR) is higher in OFDM than in other techniques. Although a Gaussian signal distribution is capacityachieving under an average power constraint, in the face of an actual power amplifier a high PAPR sets up an unattractive tradeoff between the linearity of the transmitted signal and the cost of the amplifier.By pre-coding the OFDM signals at the cost of a slightly more involved equalization process at the receiver and a slight power penalty this problem can be solved; indeed, this technique is already being done in the LTE uplink.

The next is OFDM's spectral efficiency is satisfactory, but it could perhaps be further improved if the cyclic prefixes (CPs) that prevent inter block interference were smaller or discarded and if the requirements of strict orthogonality were relaxed. The major source of concerns or at least of open questions, is whether OFDM is applicable to millimeter Wave spectrumgiven the enormous bandwidths therein and the difficulty ofdeveloping efficient power amplifiers at those frequencies.

Concept of filter bank multicarrier (FBMC), transmission techniques is presented as a result of endeavors taken to overcome the inherent demerits associated with OFDM techniques.Interest in hardware accelerators comes to play due to the tightening power budgets and performance limits. Hardware accelerators are optimized functional blocks designed to offload specific tasks from general purpose CPUs. The analogous soft wares which are running in the CPU generally works in an sequential manner but due to the optimized and dedicated architectures these blocks can perform faster. Current 4G round-trip latencies are on the order of about 15 ms, and are based on the 1 ms sub frame time with necessary overheads for resource allocation and access. For most current services this latency is sufficient, anticipated 5G applications include two-way gaming, novel cloud-based technologies such as those that may be touchscreen activated, and virtual and enhanced reality. In order support these applications 5G must be able to support a roundtrip latency of about 1ms, which is an order greater than 4G.Here in this thesis we are considering FBMC as the reference study using this hardware accelerator.

II. THERORECTICAL BACKGROUND

A.FBMC

Traditional orthogonal frequency-division multiplexing (OFDM) schemes are generalized into FBMC (filter bank multi-carrier), allowing a nonrectangular sub-channel pulse shape in the time domain. The major advantage of this scheme a better spectral containment that improves interference mitigation in several time-variant environments. The basic idea underlying this technique is to perform a nonrectangular pulse-shaping as efficient as possible with the channel characteristics (time and frequency dispersion). In FBMC technique guard bands are reduced and cyclic prefix which is present in OFDM is absent due to this from a transmission perspective, the FBMC technique has the potential to increase bit rate. The possibility to allocate different subcarriers to different unsynchronized users in a spectrally efficient manner is an another feature of FBMC .Compared to OFDM out-of-band emission of FBMC is much lower.

Comparatively higher spectral efficiency and capability of having a continuous transmission without guard intervals or cyclic prefixes are among the key inherent attractions.[3],[5] Comparatively higher throughput is expected to be maintained with a continuous and efficient transmission. The other main advantage is much more efficient use of spectrum with lower spectral leakages or reduced amount of side lobes increasing the robustness against ICI. FBMC is supported by two other main categories of concepts viz. theories of multirate techniques and theories of well localized filer design.

In 5G we use FBMC technique. In orderto address the drawbacks of rectangular time windowing in OFDM, the need for large guard bands, shows that the use of filter bank multicarrier permits a robust estimation of very large propagation delays and of arbitrarily high carrier frequency offsets, whereas OFDM would have required a very long CP to attain the same performance levels.

Filter bank multicarrier is a development of OFDM.Using banks of filters that are implemented, typically using digital signal processing techniques. When carriers were modulated in an OFDM system, sidelobes spread out either side. With a filter bank system, the filters are used to remove these and therefore a much cleaner carrier results. The block diagram of FBMC consists of bandpass filters, mixers and an adder. The block diagram of FBMC is below in fig 1.
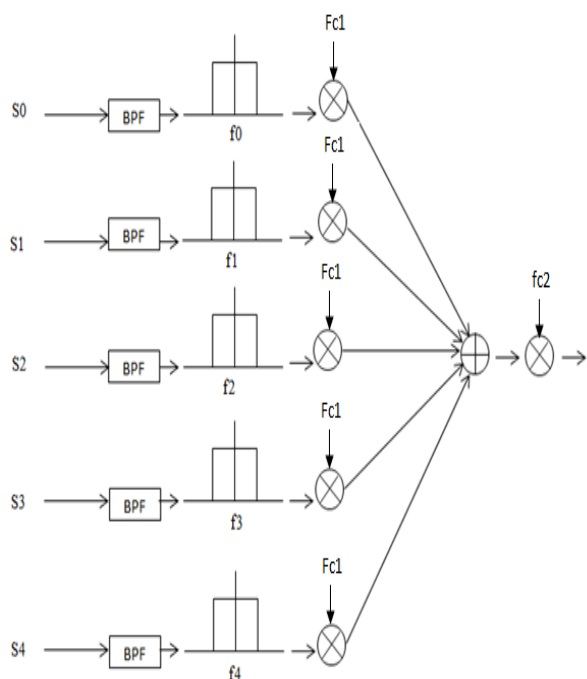
Fig 1. FBMC block diagram

A trans-multiplexer structure can be used to explain a multicarrier system, i.e.by a synthesis analysis filter bank. The analysis filter bank consists of all the matched receive filters [7] and synthesis filter bank is composed of all the parallel transmit filters.In order to better localize the sub carriers FBMC waveforms utilize a more advanced prototype filter design. Frequency sampling technique is used for the prototype filter used in this paper. This frequency sampling technique gives the advantage of using a closed-form representation that includes only a few adjustable design parameters.

In the case of filter bank-based systems, transmit pulses are localized in time and in frequency. By introducing half a symbol period delay between the in-phase and the quadrature components of every complex symbolorthogonality between the carriers is maintained. The well-adjusted frequency localization of the prototype filter guarantees that only adjacent carriers interfere with each other. This justifies the use of FBMC waveforms in a non-synchronous context and particularly for the fragmented scenario. Nevertheless, adjacent carriers significantly overlap with this kind of filtering. In order to keep adjacent carriers orthogonal, real and pure imaginary values alternate on successive carrier frequencies and on successive transmitted symbols for a given carrier at the transmitter side.FBMC is a design that features a near-ideal filtering property that is innate to its formulation. In cognitive radios, the filtering capability of FBMC systems makes them the perfect choice for filling in the spectrum holes.

FBMC systems can be designed to be equally robust to channel time and frequency spreading. A perfect match to the applications such as PLC and DSL communication systems, where the channel is subject to a number of high-power interfering narrow-band signals the well-designed prototype filters in FBMC make this modulation

B. Hardware accelerator

The hardware that performs the acceleration, when in a separate unit from the CPU, is referred to as a hardware accelerator.The acceleration design flow consists of three main tasks (Fig. 2). [8]The first and important step is finding the time critical function which is to be accelerated i.e. profiling the application running in the original system. The next step after a time-critical function is identified, it is necessary to execute the same function at gate level it is necessary to create an IP (intellectual propriety) core. There are two different ways to create hardware accelerators, first one is using HDL manually describing the target function, the other is using tools that automatically generate the HDL block directly from a C/C++ code or a Simulink model.

In the next step inorder to connect the accelerator to the CPU an interface design is needed.[8]hardware accelerators can be hosted using Field Programmable Gate Arrays since they can perform highly optimized functions concurrently at gate level.

Compared to general purpose cores designers rely on specialization to increase logic efficiency, which improves energy consumption and performance by 100-500x [9],[11].By limiting each accelerator to the workloads it was designed for accelerators achieve these gains at the expense of flexibility.Processors handle diverse set of workloads, accelerated systems must use a number of hardware accelerators to target these workloads. In order to prevent I/O complexity and other bottlenecks, each accelerator must be small since many accelerators will be needed, and must communicate with other accelerators. The acceleration design flow is as shown below fig.2
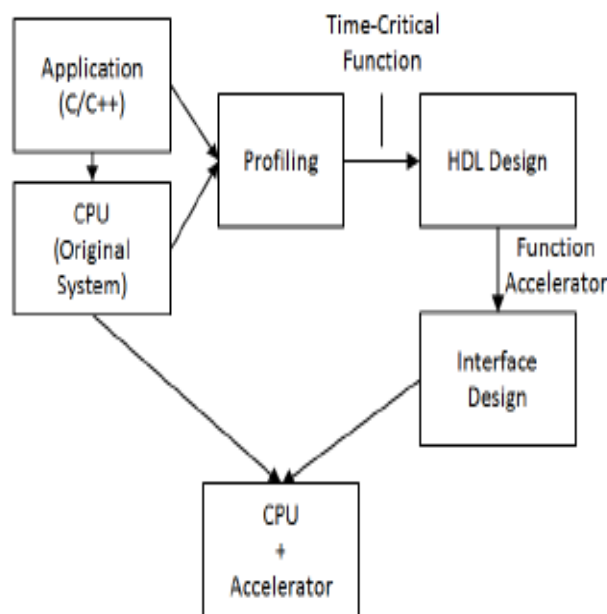

Fig 2. Acceleration design flow

III .VERIFICATION &IMPLEMENTATION

A. Verification

FBMC consists of a set of parallel filters. The first step is verifying the FBMC block implementation. FBMC block is designed and verified using Simulink. Here we are considering a bandwidth of 1MHz and we are splitting it into five different frequency bands using five Band Pass filters (BPF), each having a frequency range of 200 KHz. i.e 0-200KHz, 200 KHz-400 KHz, 400 KHz-600 KHz, 600 KHz-800 KHz, 800 KHz – 1MHz. Here we are using five input signals 100 KHz, 300 KHz, 500 KHz, 700 KHz and 900 KHzof equal amplitude. These signals are passed through the corresponding band pass filter banks. The output of the BPF is fed to corresponding mixer blocks which is used to modulate (up convert) the input signal. Here the carrier frequency used for modulation is 32MH and is fed as the local oscillator signal to the mixers. Each mixer will modulate (up convert) the input signal and produce two different frequencies, one is the sum frequency (Fc+fm) and the other is the difference frequency (fc-fm). Then all the mixer outputs are combined using an adder block. The band width of the adder output will be 2MHz (32MHz +/- 1MHz, ie. 31MHz – 33 MHz).The output of the adder block is further up converted to 600MHz using another mixer at its output. 600MHz is applied to the mixer as the local oscillator frequency. The final output will have these frequency components. The block diagram in Simulink is as below
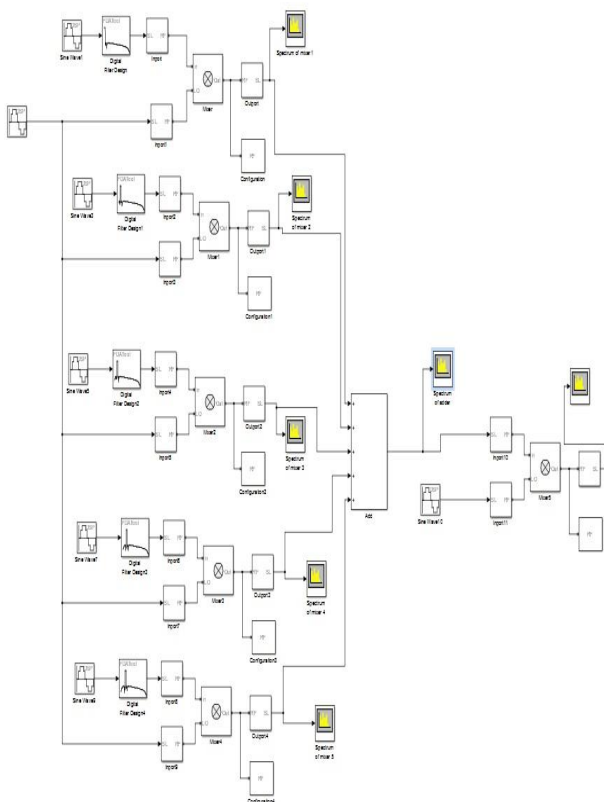

Fig 3. FBMC block diagram in simulink

One input to the final mixer is the adder output which consists of frequencies 31.1MHz,   31.3MHz, 31.5MHz, 31.7MHz, 31.9MHz, 32.1MHz, 32.3MHz, 32.5MHz, 32.7MHz and 32.9MHz   and the LO frequency is 600MHz. The sum output consists of the frequencies at 631.1MHz, 631.3MHz, 631.5MHz, 631.7MHz, 631.9MHz, 632.1MHZ, 632.3MHz, 632.5MHz, 632.7MHz and 632.9MHz. Similarly will obtain difference frequencies also.

Using the Simulink software we designed and verified the FBMC modulation, here only considered the transmission part.

B. Implementation

1) General purpose processors

General purpose processors are designed for a wide variety of computational tasks. While considering the performance of a system latency is the major concern. Typically latency is undesirable. Communication standards specify the delay tolerance between the components in the system.Additional latency occurred while we implement the physical layer on GPP is mainly due to two reasons they are

- The latency introduced by the links which are used to pass the data between the radio DACs/ADCs and the GPP.

- The latency due to the scheduling and buffering of the individual processing blocks in the physical layer.

The process under filtering is nothing but convolution. In order to find out the processing time for comparison we are considering two different coding one is in c code and the other is Mat lab code

a) C code

Filter is coded using the c code and it is verified using c free software. The processor used is Intel core I3 and the CPU clock speed is 2.4GHz. The execution time obtained is 31 ms.

b) Mat lab code

Filtering is nothing but convolution operation. Mat lab code for convolution is executed in Intel core I5 processor with a CPU clock speed of 2.6GHz.The execution time for Mat lab code is 200ms.

2) Hardware accelerator

a) Vivado

The most important stage in improving the performance of a system is finding the time critical function and minimizing it. In the case of FBMC technique which we are introducing in 5G the time critical function is the filtering part, so it can be improved using the hardware accelerator. Hardware accelerator is nothing but a set of hardware components which are specially designed to do a set of functions. So in the case of FBMC we can use the hardware accelerator to reduce the latency in the filter part.

The first step is implementing FBMC in the system generator as shown in the fig.4. FBMC system is designed and implemented in system generator and verified in Simulink.

The next step is converting this design to gate level ie. IP core generation. While converting the system generator block to an IP we can directly call that IP in the designing step and we can complete the implementation. After converting it in to IP we can generate the block design in Vivado. Then run the block diagram and synthesizing it, in the next step bit stream is generated. Finally compiling into hardware is done by using SDK. The output signal is obtained as in fig.5. The hardware used here is ZYNQ 7020.
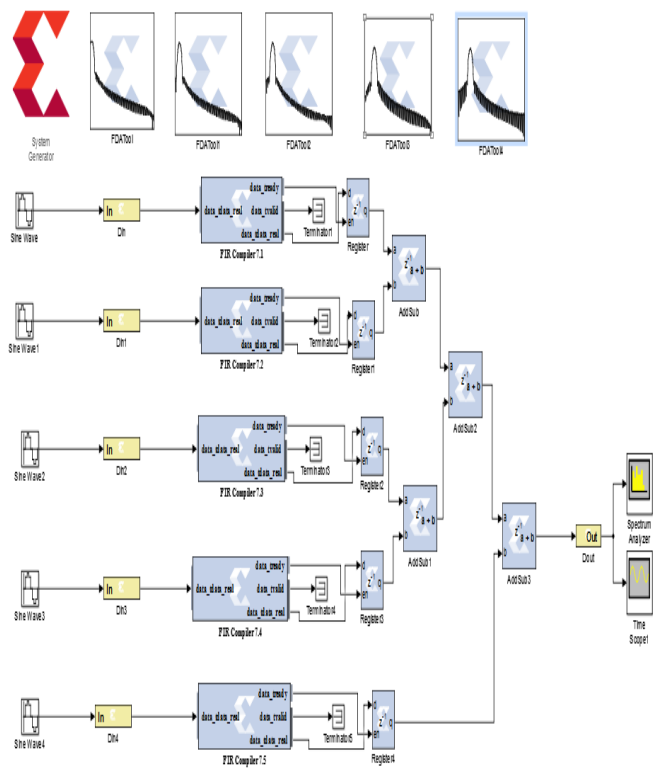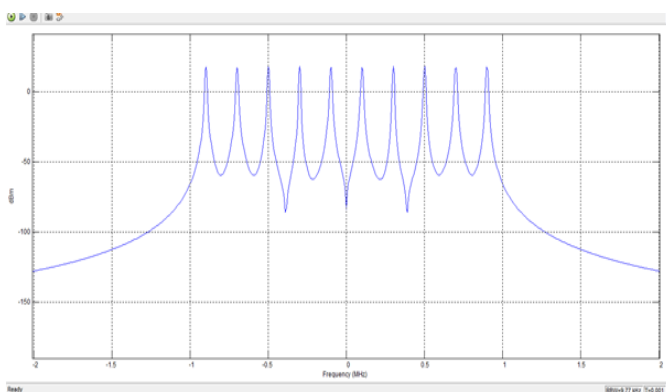

Fig 4. FBMC in system generator

Fig 5. Output signal of FBMC

3) Comparison

In the case of general purpose processors, blocks are processed in a sequential way ie. each sequential process requires clock cycles. Because of that the total process time will be more compared to parallel processing. In the case of hardware accelerators they posses dedicated hardware for doing the parallel processing in a single clock, so the time for processing will be less. Such systems are more efficient. Here in this thesis work filter is implemented in c-code, Mat lab code and also using hardware accelerator. In the case of c code it takes 31 ms and in Mat lab code it takes 200ms and in the case of this hardware accelerator it takes 1.9μs. This difference occurs due to the parallel processing in the hardware accelerator which takes only fewer amounts of clock cycles to do the process.

IV. CONCLUSION

Till 4G OFDM techniques are used in the communication field. The upcoming 5G technology requires more performance and efficiency since it provides wireless World Wide Web. Performance depends upon latency in the case of a communication system. If latency is less we can say that the system is more efficient. OFDM requires more bandwidth so we go for FBMC technique in 5G. The bandwidth of FBMC is very much improved since it uses bank of filters. The filter used is band pass. The guard band can be avoided in FBMC so it is more efficient. This FBMC is designed and implemented using Simulink. The next stage was finding the time critical function. The filtering part is the time critical function. First the filter part was implemented in GPP ie. in c code and in Mat lab code and the corresponding execution times were obtained. Then by using the system generator in Vivado software, an IP core is generated and compiled to hardware and the time required for filtering was obtained. The time required will be reduced by using the hardware accelerator i.e. the latency can be improved. Hardware accelerator is nothing but a set of hardware components which are dedicated to do different tasks. By doing the filtering part in GPPs and hardware accelerator, we can understand that the processing time can be reduced to μs from ms. Major goal of 5G is, reducing the delay to 1ms. By doing the filter part in HA the processing time can be improved and the latency can be reduced.

REFERENCES

[1] ArmanFarhang_, Nicola Marchetti_, FabrıcioFigueiredo "Massive MIMO and Waveform Design for 5th Generation Wireless Communication Systems" IEEE 1 Jan 2015

[2] Jeffrey G. Andrews, Fellow, IEEE, Stefano Buzzi, Senior Member, IEEE, Wan Choi, Senior Member, IEEE, "What Will 5G Be?" IEEE Journal On Selected Areas In Communications, Vol. 32, No. 6, June 2014

[3] Jean-Baptiste Dore, Vincent Berg, Nicolas Cassiau and DimitriKtenas "FBMC receiver for multi-user asynchronous transmission on fragmented spectrum" EURASIP Journal on Advances in Signal Processing 2014

[4] Rita C. Nilawar, D.M. Bhalerao"review on a new generation wireless mobile network - 5g" IJRET: International Journal of Research in Engineering and Technology, june 2014

[5] Nicolas Cassiau, DimitriKténas and Jean-Baptiste Doré"Time and frequency synchronization for CoMP with FBMC" IEEE ISWCS 2013

[6] Vida Vakilian, Thorsten Wild†, Frank Schaich "Universal-Filtered Multi-Carrier Technique for Wireless Systems Beyond LTE" IEEE 2013

[7] Martin Kasparick "System-level interfaces and performance evaluation methodology for 5G physical layer based on non-orthogonal waveforms" Asilomar Conference, Nov. 5 th, 2013

[8] Paulo Possa, David Schaillie, and Carlos Valderrama "FPGA-based Hardware Acceleration: A CPU/Accelerator Interface Exploration" IEEE 2011

[9] BehrouzFarhang-Boroujeny "OFDM versus filter bank multicarrier" IEEE signal processing magazine may 2011

[10] Maurice Bellanger, Didier Le Ruyet "Filter bank based multicarrier for cognitive radio" presentation GDR-ISIS, 2011

[11] M. J. Lyons, M. Hempstead, G-Y.Wei, and D. Brooks, "The Accelerator Store framework for high-performance, low-power accelerator-based systems," in IEEE Computer Architecture Letters, July-Dec 2010.

[12] A.Kennedy, X. Wang, and B. Liu, "Energy efficient packet classification hardware accelerator," in IEEE International Symposiumon Parallel and Distributed Processing, Miami, FL, EUA, 2008.

[13] Altera, "AN 531: Reducing power with hardware accelerators,"2008

[14] SebastienLafond, Johan Lilius, "Interrupt Costs in Embedded System with Short Latency Hardware Accelerators" 15th Annual IEEE International Conference and Workshop on the Engineering of Computer Based Systems, 2008

[15] Minho Shin, Arunesh Mishra, William A. Arbaugh "Improving the Latency of 802.11 Hand-offs using Neighbor Graphs" IEEE MobiSys'04, June 6–9, 2004