

UNIVERSITY OF CALIFORNIA

Los Angeles

**Energy-Efficient VLSI Signal Processing for Wideband
Spectrum Sensing in Cognitive Radios**

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Electrical Engineering

by

Tsung-Han Yu

2013

© Copyright by

Tsung-Han Yu

2013

ABSTRACT OF THE DISSERTATION

Energy-Efficient VLSI Signal Processing for Wideband Spectrum Sensing in Cognitive Radios

by

Tsung-Han Yu

Doctor of Philosophy in Electrical Engineering

University of California, Los Angeles, 2013

Professor Dejan Marković, Co-chair

Professor Danijela Čabrić, Co-chair

With the rapid increases in the number of wireless devices, fixed spectrum allocation has shown to be a major limitation to the evolution of wireless technologies. Cognitive radio (CR) allows opportunistic spectrum access by searching and utilizing temporally and spatially unused spectrum, provided that CR users do not cause interference to the primary users of the spectrum. Spectrum sensing over a wide bandwidth increases the probability of finding under-utilized spectrum for cognitive radios. However, the realization of wideband sensing is challenging because strong primary users introduce large dynamic range and spectral leakage to adjacent unused bands.

This work presents an algorithm-architecture co-design framework for wideband spectrum sensing. The suppression of spectral leakage is achieved by multitap-windowed FFT processing, which also enables reduced sensing time. The sensing time and de-

tection threshold are adapted according to channel-specific spectral leakage, enabling reliable wideband detection within constrained sensing time. Power and area cost of the compute-intensive FFT block is minimized by using parallelism, radix factorization, and compact delay lines. A per-channel floating point data processing for large dynamic range signal is employed for power and area saving. A partial PSD estimation scheme that performs energy detection on only the band-of-interest further improves the energy efficiency. Two chips have been designed to demonstrate these concepts. These chips guarantee reliable weak signal detection with short sensing time, and outperform the prior work by at least 22x in power/bandwidth. Techniques developed in this dissertation enable energy-efficient chip implementation of advanced wideband signal processing for cognitive radios.

The dissertation of Tsung-Han Yu is approved.

Milos D. Ercegovać

Greg Pottie

Danijela Čabrić, Committee Co-chair

Dejan Marković, Committee Co-chair

University of California, Los Angeles

2013

To my family.

TABLE OF CONTENTS

1	Introduction	1
1.1	Cognitive Radios	1
1.2	Spectrum Sensing	3
1.3	Motivation for This Work	6
1.4	Dissertation Outline	8
2	Wideband Spectrum Sensing in Cognitive Radios	10
2.1	Frequency-Domain Wideband Spectrum Sensing	10
2.2	Overview of Conventional Channelization Schemes	13
2.3	Overview of Previous Work	14
2.4	Summary	16
3	Wideband Spectrum-Sensing Algorithms and Real-Time Testbed	17
3.1	Wideband Spectrum-Sensing Design Problem	17
3.1.1	System Model	18
3.1.2	Design Specification	18
3.1.3	Design Challenges	20
3.2	Proposed Wideband Spectrum-Sensing Algorithms	23
3.2.1	PSD Estimation Using Multitap-Windowed FFT Processing	25

3.2.2	Power Detector Matrix Model	28
3.2.3	Interfering Power Estimation	30
3.2.4	Detection-Threshold and Sensing-Time Adaptations	31
3.3	Numerical Simulations	33
3.3.1	Improvement in ROC	36
3.3.2	Improvement in Sensing Time	36
3.3.3	Improvement in False-Alarm Rate	37
3.4	Proposed VLSI Architecture and Experimental Results	38
3.4.1	Proposed VLSI Architecture	38
3.4.2	Fixed-Length vs. Reconfigurable FFT	39
3.4.3	Hardware Complexity	43
3.4.4	Implementation on CR Testbed	44
3.4.5	Experimental Results	46
3.5	Summary	49
4	Power and Area Minimization of Reconfigurable FFT Processor	51
4.1	Overview of FFT Designs	51
4.2	FFT Algorithms and Architectures	53
4.2.1	Radix-2-Butterfly Based Architecture	54
4.2.2	Reconfigurable Architecture	56

4.3	Power and Area Minimization	61
4.3.1	Parallel Architecture with FFT Decomposition	62
4.3.2	Estimation of Area and Power	64
4.3.3	Mixed-Radix Implementation	66
4.3.4	Delay-Buffer Architecture and Memory Element	70
4.3.5	Area-Efficient Twiddle-Factor Generator	76
4.4	Summary	77
5	Chip 1: 200MS/s Wideband Spectrum Sensing Processor	78
5.1	Design Challenges of Wideband Spectrum Sensing	78
5.2	Architecture and Circuit Design	80
5.2.1	Multi-path Pipelined MW-FPD	81
5.2.2	Floating-point Signal Processing	84
5.2.3	Processing Time Overhead Reduction	87
5.2.4	Summary of Design Optimization	91
5.3	Chip Implementation	93
5.3.1	FPGA-aided Verification	95
5.3.2	Measurement Results	97
5.3.3	Performance Comparison	99
5.4	Summary	101

6 Chip 2: 500MS/s Wideband Band Segmentation Processor for Blind Signal

Classification	102
6.1 Design Specifications	103
6.2 Energy Minimization Methodology	104
6.2.1 Partial PSD Estimation	105
6.2.2 Miss-detection Tolerant Signal Detection	115
6.2.3 Design Example	117
6.3 Proposed VLSI Architecture	119
6.3.1 Energy-Area Efficient Reconfigurable FFT	120
6.3.2 Energy Efficient CIC Filter	121
6.3.3 Summary of the Design Optimization	125
6.4 Chip Implementation	127
6.4.1 FPGA-aided Verification	127
6.4.2 Measurement Results	129
6.4.3 Performance Comparison	133
6.5 Summary	133
7 Conclusions	135
7.1 Research Contributions	136
7.2 Future Work	138

References 140

LIST OF FIGURES

1.1	Spectrum allocation of the United States.	2
1.2	Measurement of 0-6 GHz spectrum utilization at BWRC [4].	3
1.3	Interferer in-band power and spectral leakage.	8
2.1	Interferer in-band power and spectral leakage.	11
3.1	Maximum tolerable INR. For a 200 kHz wireless microphone primary user, the maximum INR is constrained by 30 dB.	20
3.2	Receiver operating characteristic of the conventional rectangular-windowed FFT power detector for sensing times of 0.5 ms (solid line) and 50 ms (dashed line). The strong adjacent-band interferers and the weak band-of-interest signal are 30-dB INR and -5 -dB SNR, respectively.	21
3.3	Block diagram of the proposed wideband spectrum sensing processor. The processor consists of a multitap-windowed frequency-domain power detector for PSD estimation, sensing-time adaptation for optimal sensing time, and detection-threshold adaptation for desired sensing rate.	24
3.4	Wideband spectrum sensing procedure.	25

3.5	Detection procedure for the windowed frequency-domain power detector (W-FPD), the windowed-overlapped frequency-domain power detector (WO-FPD), and the proposed multitap-windowed frequency domain power detector (MW-FDP).	27
3.6	Receiver operating characteristic curves for different power detectors. The multitap-windowed frequency-domain power detector achieves the best detection rate.	33
3.7	Required sensing time for conventional FFT power detector, windowed frequency-domain power detector, and multitap-windowed frequency-domain power detector as a function of interferer-to-signal spacing (a) and interferer-to-noise ratio (b). The proposed multitap-windowed power detector requires the least sensing time to reach a given detection rate. .	34
3.8	Influence of adaptation on P_{FA} . The adaptation threshold significantly lowers P_{FA} as compared to the use of conventional threshold. Both methods use MW-FPD for PSD estimation.	35
3.9	A VLSI architecture for the proposed spectrum sensing algorithms. (a) Multitap-windowed frequency-domain power detection block. (b) Sensing-time adaptation block. (c) Detection-threshold adaptation block.	40
3.10	Improvement in detection performance as a function of frequency resolution for a 6 MHz DTV signal. Larger FFT size with higher frequency resolution results in better sensing rates.	42

3.11	Comparison of relative hardware cost, sensing time, and computation complexity for W-FPD, WO-FPD, and MW-FDP.	45
3.12	P_D and number of averages versus INR.	48
3.13	P_{FA} and detection threshold versus INR.	48
4.1	(a) Signal flow graph of radix-2 butterfly for DIT and DIF structures. (b) Radix- 2^k single-path delay feedback (SDF) architecture. Output of stage k does not need twiddle-factor multiplication.	55
4.2	Radix-2, radix- 2^2 , radix- 2^3 , and radix- 2^4 butterfly operations.	58
4.3	Processing units (PUs) of a reconfigurable FFT processor that supports radix-2 to radix- 2^4 factorizations. (a) The architecture and (b) the constant multipliers of the PUs. The numbers in brackets next to each PU indicate relative cost (in terms of equivalent adders) of its constant multiplier.	60
4.4	(a) Reference N -point ($N = M \times L$) FFT architecture. (b) Parallel architecture. The level of parallelism = P . (c) A significant reduction in complexity is achieved when $P = L$	63
4.5	Delay and energy vs. supply voltage for a FO-4 inverter in TSMC 65nm CMOS.	65
4.6	Area-efficient implementation of constant multipliers.	67

4.7	Power and area minimization through FFT factorization and radix factorization.	70
4.8	Power and area for feasible radix factorizations of 128-point FFT. Architecture A10 has the lowest power-area product.	71
4.9	(a) Dual-port RF-based delay buffer and (b) pointer-based DFF-based delay buffer.	73
4.10	Power-area product of delay buffers: D flip-flops (DFFs) are used for delay lengths of 128 and 256, register files (RF) are used for delay lengths of 512 and 1024.	74
4.11	Power-area tradeoff for feasible memory partitions. Partitions 2×256 for length 512 and 4×256 for length 1024 delay buffers yield minimum power-area product.	74
4.12	Memory architecture for length 512 and 1024 delay buffers.	75
5.1	Proposed wideband spectrum sensing processor. An FFT-based multi-tap windowed frequency power detector estimates PSD, noise and interfering power. Sensing-time adaptation (STA) and detection-threshold adaptation (DTA) blocks adapt their parameters on a per-channel basis. .	80

5.2	Processing time for spectrum sensing consists of fixed time for MW- FPD calculation and variable time for STA and DTA blocks. The MW- FPD determines the sensing rate performance. The overhead intro- duced by the STA and DTA blocks degrades the sensing-time efficiency and has to be minimized.	81
5.3	Architecture of the frequency-domain power detector.	83
5.4	Comparison of 1024-point FFTs. Our chip that minimizes power-area product is measured at 200 MS/s. At 240 MS/s, our fully synthesized (# indicates synthesis estimate) mix-radix design optimized for minimum energy achieves lower energy and lower area than the custom radix-4 subthreshold design from [80].	83
5.5	Realization of floating-point operations: (a) addition, (b) multiplica- tion, (c) squaring by using 2's complement operators. M1 and M2 in- dicate mantissas, E1 and E2 indicate exponents. Blocks in the shaded area in (a) perform the exponent matching.	86
5.6	Architecture of the power estimation block.	87

5.7	(a) Architecture of fixed-point to floating-point conversion. The barrel shifter generates 10-bit mantissa, the priority encoder generates 5-bit exponent. (b) The plots of logic power and logic area of the power estimation block vs. mantissa wordlength show that the sensing performance spec is met with 10-bit mantissa. The floating-point format results in a 2x logic power reduction and a 2x logic area reduction as compared to the fixed-point reference design.	88
5.8	The impact of wordlength on the core power and area. A 20.1% of core power (left) and 27.5% of core area (right) are saved by wordlength reduction. The wordlength reduction is enabled by the use of floating-point arithmetic.	89
5.9	Architecture of the sensing-time-adaptation block. The initial-value estimation and the reciprocal calculation are performed using Newton-Raphson algorithm. Bounded sensing time allows for 10-bit output instead of full-precision 20-bit dictated by the input wordlength.	92
5.10	Power and sensing-time reduction of the STA block.	92

5.11	Summary of the optimization procedure. Parallelism with voltage scaling and radix factorization with power-area optimized delay lines are applied to the FFT block. Parallelism with voltage scaling is also applied at the system level. Floating-point signal processing with wordlength reduction is applied to STA and DTA blocks. An overall 17.9x reduction in power and 11.9x reduction in power-area product are achieved.	94
5.12	Chip micrograph of wideband spectrum sensing processor.	95
5.13	FPGA-aided I/O verification. Test vectors are loaded from Matlab into the FPGA board. The results are sampled in BRAM on the FPGA and read back by the terminal.	96
5.14	(a) Measured PSD estimated by MW-FPD. Two 30-dB interferers are present 2 bins away from the band of interest. The interfering power is suppressed by 20x using multitap windowing as compared to traditional FFT-based PSD estimation. (b) Measured P_D and P_{FA} as a function of adjacent-band interferer power ($0 \leq INR \leq 30dB$) for SNR = -5 dB. STA and DTA blocks effectively maintain $P_D \geq 0.9$ and $P_{FA} \leq 0.1$	98
6.1	System block diagram of blind signal classification.	103
6.2	Normalized energy with respect to different coarse-sensing and fine-sensing FFTs configuration. 64-point FFT for coarse-sensing results in minimum energy dissipation.	109
6.3	Filtering performance of a 3^{rd} -order CIC filter.	110

6.4	Required up-sampling ratio with respect to different band-of-interest bandwidths and cascaded stages.	112
6.5	Normalized energy with respect to band-of-interest bandwidth. Different fine-sensing FFT configurations are shown: a) 4096-point, b) 2048-point, c) 1024-point, d) 512-point.	114
6.6	Bandwidth and carrier frequency estimation by the binary decision results. The underestimated bandwidth makes the DSP processor unable to classify the signals.	116
6.7	The probability of having at most x consecutive zeros among b occupied channels.	117
6.8	A design example. (a) Power spectrum density of the wideband scenario. (b) Estimated PSD and the detection threshold of coarse sensing. (c) Binary decision result from coarse sensing. (c) Estimated PSD and the detection threshold of fine sensing. (e) Binary decision result from fine sensing.	118
6.9	Energy-efficient band segmentation, which consists of coarse sensing and fine sensing. Coarse sensing is to select band-of-interest, and fine sensing is to channelize the selected band with finer resolution.	120
6.10	A multi-path pipelined reconfigurable FFT with optimal parallelism and radix-factorization.	121
6.11	Comparison with the state-of-the-art FFTs.	122

6.12	4 th -order feed-forward CIC filter, supporting down-sampling ratio up to 1024, amenable to parallelism with smaller wordlength.	123
6.13	Power saving from the feed-forward architecture.	124
6.14	Summary of design optimization for signal bandwidth of 1 MHz, 10MHz, and 40 MHz. Four techniques are used to improve energy efficiency. 1) Miss-detection tolerant signal detection. 2) Partial PSD estimation. 3) 16-way parallelism and optimal radix factorization. 4) Feed-forward CIC filter architecture.	126
6.15	Chip micrograph and the layout of the band segmentation processor. . .	128
6.16	FPGA-aided I/O verification. Test vectors are stored on the FPGA board. The results are sampled and verified in real-time by using Xilinx ChipScope.	129
6.17	Measured energy with respect to different signal bandwidths. A maximum 3.2x saving in energy is achieved by using partial PSD estimation algorithms.	131
6.18	Measured energy with respect to different signal bandwidths. A maximum 3.2x saving in energy is achieved by using partial PSD estimation algorithms.	131

LIST OF TABLES

3.1	Wideband CR System Specification	22
3.2	Hardware Complexity Estimates of W-PFD, WO-FPD, and MW-FPD	44
3.3	FPGA Hardware Resource Breakdown for MW-FPD	46
3.4	Comparison with State-of-the-Art Power-Detector ASICs	46
4.1	Radix Implementations of Radix-2, Radix-2 ² , Radix-2 ³ , and Radix-2 ⁴ Processing Elements	57
4.2	The Number of Equivalent Adders Required to Implement FFT	69
5.1	Summary of Chip Features	99
5.2	Comparison with Prior Work	100
6.1	Wideband CR System Specification	104
6.2	Summary of Chip Features	130
6.3	Comparison with Prior Work	132

ACKNOWLEDGMENTS

My sincere gratitude goes to my advisors, Professor Dejan Marković and Professor Danijela Čabrić. Their stimulating guidance has sparked many ideas that carried my research. Dejan's constant drive for his students to succeed is truly exceptional, and I would not be typing this if it were not for him. The rightness and enthusiasm have been with me at every step along the way. Danijela's understanding and technical support has greatly transcended my knowledge about algorithm design. I am indebted to both of them for their faith in my research and their tolerance of my weaknesses. Their patient, trustful, and encouraging support has made me a better scholar and a better person. Professor Greg Pottie and Professor Milos D. Ercegovać provided thoughtful evaluation of my research proposal and review of this dissertation.

The multi-disciplinary, open minded setting of Parallel Data Architecture (PDA) Group and Cognitive Reconfigurable Embedded Systems (CORES) Lab created many opportunities to further develop my work in collaboration with others. Special gratitude goes to Chia-Hsiang Yang and Fang-Li Yuan. Chia-Hsiang has helped me get through pressures of graduate school and has been very supportive by sharing his wisdom and experience. I am thankful to Fang-Li for being a great roommate and for his support over year, and Fang-Li has been also a great friend since college. I am thankful to Oussama Sekkat for sharing the burden in setting up the BEE2 experiments. I am thankful to Henry Chen for sharing experience in using FPGA boards for ASIC test and debug. I am thankful to Paulo Isagani Urriza and Eric Rebeiz for

doing wonderful jobs in DARPA CLASIC project. I am thankful to Chun-Hao Li for being a great lab mate and roommate. I am thankful to Cheng C. Wang for dividing suffering during the tape-out time. I am also thankful to my PDA lab mates Sina Basir-Kazeruni, Chaitali Biswas, Hariprasad Chandrakumar, Sarah Gibson , Vahagn Hokhikyan, Richard Dorrance, Vaibhav Karkare, Rashmi Nanda, Fengbo Ren, Dejan Rozgic, and Yuta Toriyama, and my CORES lab mates Wesam Gabran, Ali Shahed, Jason Tran, and Jun Wang for stimulating discussions and friendship.

Personal gratitude goes to Jung-Mao Lin and Santiago Rodriguez Parera for many inspiring technical discussions. I am indebted to their contributions in my very first paper. I appreciate working with Shang-Kee Ting and Zaid Towfic in DARPA HEALICS project.

This research was supported by the HEALICS/DARPA, CLASIC/DARPA programs and UCLA graduate fellowship.

Finally, and most importantly, I would like to express gratitude to my parents, my brother, and my dear fiancée. This work is a product of their endless love and continued support.

VITA

- 2001–2005 B.S. in Electrical Engineering, National Taiwan University, Taipei, Taiwan.
- 2005–2007 M.S. in Electronics Engineering, National Taiwan University, Taipei, Taiwan.
- 2008–2013 Research Assistant, Electrical Engineering Department, University of California, Los Angeles, California, USA.

PUBLICATIONS

T.-H. Yu, C.-H. Yang, D. Čabrić, and D. Marković, “A 7.4mW 200MS/s Wideband Spectrum Sensing Digital Baseband Processor for Cognitive Radios,” *IEEE Journal of Solid-State Circuits (JSSC)*, vol. 47, no. 9, pp. 2235-2245, Sep. 2012.

C.-H. Yang, T.-H. Yu, and D. Marković, “Power and Area Minimization of Reconfigurable FFT Processors: A 3GPP-LTE Example,” *IEEE Journal of Solid-State Circuits (JSSC)*, vol. 47, no. 3, pp. 757-768, Mar. 2012.

T.-H. Yu, Oussama Sekkat, Santiago Rodriguez-Parera, D. Marković, and D. Čabrić, “A Wideband Spectrum-Sensing Processor with Adaptive Detection Threshold and Sens-

ing Time,” *IEEE Transaction on Circuit and System-I (TCAS-I)*, vol. 58, no. 11, pp. 2765-2775, Nov. 2011.

T.-H. Yu, C.-H. Yang, D. Marković, and D. Čabrić, “An Energy-Efficient VLSI Architecture for Wideband Spectrum Sensing for Cognitive Radios,” in *Proc. IEEE Global Telecommunications Conference (GLOBELCOM)*, Dec. 2011, pp. 1-6.

T.-H. Yu, C.-H. Yang, D. Čabrić, and D. Marković, “A 7.4mW 200MS/s Wideband Spectrum Sensing Digital Baseband Processor for Cognitive Radios,” in *Proc. IEEE International Symposium on VLSI circuit (VLSI)*, June 2011, pp. 254-255.

T.-H. Yu, S. Rodriguez-Parera, D. Marković, and D. Čabrić, “Cognitive Radio Wideband Spectrum Sensing Using Multitap Windowing and Power Detection with Threshold Adaptation,” in *Proc. IEEE International Conference on Communications (ICC)*, May 2010, pp. 1-6.

C.-H. Yang, T.-H. Yu, D. Marković, “A 5.8mW 3GPP-LTE Compliant 88 MIMO Sphere Decoder Chip with Soft-Outputs,” in *Proc. IEEE International Symposium on VLSI circuit (VLSI)*, June 2010, pp. 209-210.

CHAPTER 1

Introduction

1.1 Cognitive Radios

Wireless technology is rapidly proliferating into all aspects of daily life, and the demand for wireless connectivity is ever increasing. Higher and higher data rates are required for real-time and high-quality data transmission, with a rapid increase in the number of wireless devices. Till now, only specific users are allowed to access the wireless resources based on the spectrum polices. Figure 1.1 shows the spectrum allocation of the United States, where most of the radio spectrum is assigned to licensed users. This fixed spectrum allocation, however, does not guarantee the spectrum is efficiently utilized at all times. As indicated by numerous reports [1-4], the usage of the spectral resources is significantly under-utilized. A survey [1] reported in 2002 showed that, on average, only 2% of allocated spectrum in the United States is in use at any given moment. Another report [3], based on the measurement for the frequency bands below 3 GHz from January 2004 to August 2005, concluded that the averaged spectrum occupancy is only about 5.2%. Figure 1.2 [4], conducted by Berkeley Wireless Research Center (BWRC) over frequency bands from 0 to 6 GHz, illustrates the frequency bands

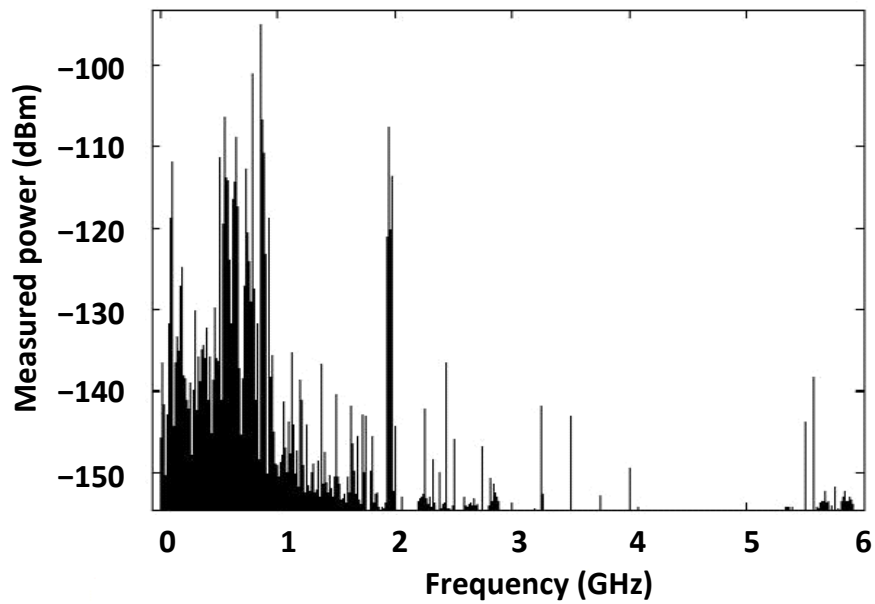


Figure 1.2: Measurement of 0-6 GHz spectrum utilization at BWRC [4].

Group (WG) [6]. The IEEE 802.22 WG allows the development of physical (PHY) and medium access control (MAC) layers for the use by license-exempt devices in TV bands. The 802.22 admits license-exempt devices to reuse vacant TV spectrum without introducing any harmful interference to primary users. The 802.22 has a leading and key role for facilitating the CR development and its outcome will serve as the basis for new and innovative research in this promising area.

1.2 Spectrum Sensing

The enabling technology for CR systems is spectrum sensing, in which the presence of primary users is detected in the band of interest to avoid harmful interference. The key requirement for spectrum sensing is reliable signal detection, with high sensitivity

within a constrained sensing time. The high sensitivity requirement prevents a CR system from causing interference to the primary users [7-13]. The sensitivity improves by increasing the sensing time, but a long sensing time reduces the effective time for communication on an unused primary channel, thus reducing the CR throughput. On the other hand, a short sensing time enhances the throughput but increases the chance of unintended collisions. The constrained sensing time balances the tradeoff between CR throughput and unintended collisions.

Wideband (> 100 MHz) sensing is a highly desirable feature of CR systems, since it allows for simultaneously sensing of multiple channels and thus increases the probability of finding available channel. Several algorithms have been proposed to resolve wideband spectrum sensing problem. Wideband sensing, however, imposes many design challenges at the physical layer [14-20]. The wideband front-end must have sufficient linearity to avoid mixing interferers into the band of interest [17-20]. The analog-to-digital converter (ADC) requires high resolution to support large-dynamic-range signals and high sampling rate to adequately sample wideband spectrum [20]. In the digital baseband, the sensing processor needs to provide a reliable signal detection in a negative SNR regime while operating in real time [14-16]. The DSP baseband, therefore, must accommodate advanced signal processing algorithms within limited power and area. This dissertation focuses on development of a DSP baseband processor for wideband spectrum sensing.

Generally, the spectrum-sensing algorithms can be categorized into two types: energy (power) detection and feature detection.

Energy detection identifies the presence of signals by comparing the measured energy level to a threshold. The processing gain is proportional to the observation time T . Longer observation time improves the signal detection accuracy that enhance the detection performance. Energy detection is widely used for spectrum sensing because it is easy to implement and is able to detect noise level and interferers, without the knowledge of electromagnetic environment. However, several drawbacks of energy detection might diminish its implementation simplicity. First, the detection performance, false-alarm probability and detection probability, are highly susceptible to the changing noise level and the activity of interferers. The threshold used for energy detection needs to be decided carefully to keep a balance of false-alarm probability and detection probability. This problem becomes more serious in large-dynamic-range wideband spectrum sensing. Second, energy detection is unable to differentiate between modulated signals, noise, and interferers. Adaptive signal processing cannot be applied for cancelling the interferers. In addition, when the system is required to treat noise and other secondary users differently, more sophisticated signal processing algorithms need to be developed.

Feature detection identifies the channel occupancy by using the cyclostationary property of a modulated signal [21-25]. The cyclostationary property is commonly characterized by spectral correlation function [21-22]. Since only a modulated signal has a specific feature at a nonzero cyclic frequency, this feature can be used to differentiate background noise and modulated signals. In addition, different types of modulated signals (such as BPSK, QAM, FSK), even having the same power spectrum density

function, have highly distinct spectral correlation functions. Thus, feature detection is able to distinguish modulated signals, noise, and interferers. Furthermore, signals with overlapping feature in their power spectrum have non-overlapping feature in their spectral correlation functions, indicating feature detection is able to differentiate overlapping signals as well. The computation complexity of feature detection, however, is much higher than energy detection [15]. As far as energy efficiency is concerned, feature detection is not feasible for real-time and low-cost implementation.

There are still other signal detection algorithms for wideband sensing, e.g. compressive spectrum sensing [26-32] or cooperative spectrum sensing [33-35]. Compressive sensing exploits spectrum scarcity and detects signals with sub-Nyquist sampling rates. Cooperative sensing improves detection performance by exploiting spatial diversity. [26-35] have shown these algorithms outperforming energy detection and feature detection. However, these algorithms are not applicable for low power implementation.

In this dissertation, we adopt energy detection with associated adaptation algorithms for reliable signal detection and energy-efficient implementation, making this design applicable to portable devices.

1.3 Motivation for This Work

The major challenge of wideband spectrum sensing comes from the large-dynamic-range spectrum. Without notch filters for the detected spectrum, it is inevitable having both weak signals and strong blockers. The blockers might be strong primary users,

or illegal or unexpected jammer. In the worst case, a weak signal with negative SNR, which is adjacent to two strong blockers, has to be detected. As shown in Fig. 1.3, under this scenario, an *interfering power* leaks into the band occupied by the weak signal. The interfering power comes from two sources. First, a modulated signal is shaped (by digital pulse-shaping and analog bandpass filters) in frequency according to a pre-defined frequency mask, which in reality is not a perfect brick-wall filter. As a result, the tail of the modulated signal spectrum might introduce significant interfering power in the adjacent band, provided that this modulated signal is very strong. Second, the time-domain received samples at the sensing processor are channelized with a non-ideal filter (e.g. FFT filterbank), which also causes spectral leakage. Both effects are represented by the shaded areas in Fig. 1.3. We call the combination of these two effects interfering power, which makes weak signal detection more difficult.

As energy detection is applied for wideband sensing, signal detection in real time is challenging because the detection performance is sensitive to the dynamic nature of electromagnetic environment. The threshold used for energy detection highly depends on the noise level and the activity of interferers. To accommodate the variation in background noise, an adaptive algorithm for the detection threshold needs to be designed. Also, with the presence of strong adjacent-band blockers, the detection threshold needs to be adapted to the channel condition. In addition to developing adaptive algorithms, the detection threshold needs to be channel-specific and adapted to real-time measured data.

Real-time and large-dynamic-range signal processing not only imposes challenges

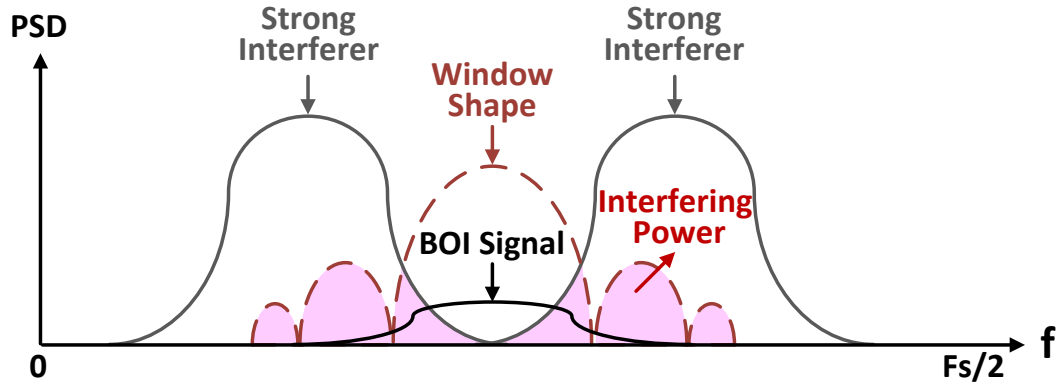


Figure 1.3: Interferer in-band power and spectral leakage.

on algorithmic development, but also makes low-power architectural design difficult. Wideband sensing in real time indicates high-throughput signal processing is required. A low-power and high-throughput VLSI design is not an easy task. Also, large-dynamic-range datapath dictates large wordlengths, which increases the cost of arithmetic operations and data storage. A low-power and high-throughput architecture for large-dynamic-range signal processing is the main focus in this dissertation.

1.4 Dissertation Outline

In response to the challenges of reliable weak signal detection in the presence of strong adjacent-band interferers, an algorithm-architecture co-design framework is developed. Chapter 2 reviews wideband spectrum sensing algorithms. The development of wideband spectrum sensing algorithms are presented in Chapter 3. The algorithms have been verified on a real-time FPGA hardware testbed. Fast Fourier Transform (FFT) is intensively involved in wideband spectrum sensing process. To enable energy-

efficient design, a design methodology for power-area minimized FFT processor is described in Chapter 4. The implementation of a wideband spectrum sensing processor and its measurement results are described in Chapter 5. Channelization of the entire spectrum is not always necessary, especially when the spectrum is scarce. A partial PSD estimation scheme, which dynamically channelizes only a portion of spectrum is proposed to enhance energy efficiency. A wideband band segmentation for blind signal classification is used as a design example. The associated algorithm-architecture co-design framework is demonstrated in Chapter 6. Chapter 7 concludes this work and discusses future research directions.

CHAPTER 2

Wideband Spectrum Sensing in Cognitive Radios

In order to define challenges addressed in this work, this chapter introduces energy detection algorithm and the associated conventional channelization scheme.

2.1 Frequency-Domain Wideband Spectrum Sensing

We first consider a sideband signal composed of several non-overlapping narrow-band signals (primary users), where each signal has the same bandwidth and modulation scheme, as shown in Fig. 2.1. Additive white Gaussian noise (AWGN) applies uniformly across the band. As a result, the noise power in all the individual narrowband channels is statistically equal.

Frequency-domain power detection (FPD) is adopted as the sensing algorithm for energy-efficient realization. In FPD, the Fast Fourier Transform is used to channelize the entire spectrum, where the model for each channel is similar to that in narrowband signal detection [35], [36-37]. The detection problem for each channel k can be modeled by a binary hypothesis test, where hypothesis $H_0(k)$ stands for noise only, and hypothesis $H_1(k)$ means that both noise and signal are present.

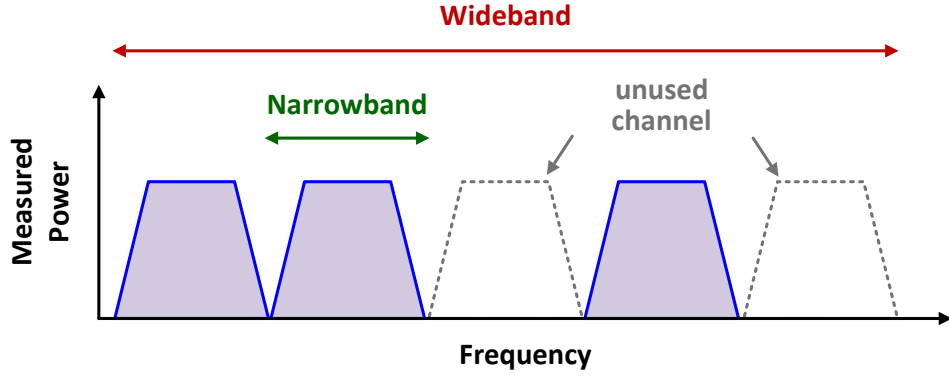


Figure 2.1: Interferer in-band power and spectral leakage.

Frequency-domain power detection involves power spectrum density (PSD) estimation of the entire band. The accuracy of the estimation is determined by the variance estimated PSD, which can be reduced by averaging (accumulating) the estimated channel power. Thus, for reliable signal detection, the power detection requires an adequate number of samples to obtain accurate PSD estimation. Then, the estimated PSD is compared to a detection threshold to decide between $H_0(k)$ and $H_1(k)$. The decision rule is represented by

$$T(k) \triangleq \sum_{m=0}^{M-1} |\hat{X}_m(k)|^2 \underset{H_0(k)}{\overset{H_1(k)}{\gtrless}} \gamma(k) \quad (2.1)$$

where the test statistic $T(k)$ indicates the channel power in one FFT bin and $\gamma(k)$ is the corresponding detection threshold in channel k . $X_m(k)$ is the FFT output from the bin (channel) index k and block index m . For the baseband sampling rate F_s , FFT size N and the number of FFT frames M , the corresponding sensing time is

$$T_{sensing} = M \times N \times 1/F_s \quad (2.2)$$

The probability of false-alarm (P_{FA}) and the probability of detection (P_D) are the metrics for the sensing algorithm performance. P_{FA} is the probability that a CR system fails to identify an unoccupied spectral segment; it measures the utilization of unused spectrum. P_D is the probability that a system detects the presence of a primary user, and measures the rate of avoiding interference. The probabilities of false-alarm and detection in channel k ($P_{FA}(k)$ and $P_D(k)$) are determined by the test statistic $T(k)$ and the channel-specific detection threshold $\gamma(k)$. According to the central limit theorem, $T(k)$ is asymptotically normally distributed if M is large enough ($M \geq 20$ is sufficient in practice). Using the mean and variance of $T(k)$, P_{FA} and P_D can be defined as

$$P_{FA}(k) = Q\left(\frac{\gamma(k) - \mu(T(k)|H_0(k))}{\sigma(T(k)|H_0(k))}\right), \quad (2.3)$$

and

$$P_D(k) = Q\left(\frac{\gamma(k) - \mu(T(k)|H_1(k))}{\sigma(T(k)|H_1(k))}\right), \quad (2.4)$$

respectively, where $Q(\cdot)$ is the tail probability of a zero-mean unit-variance Gaussian random variable.

The number of FFT frames (sensing time) M and detection threshold $\gamma(k)$ strongly affect P_D and P_{FA} . The sensing time dictates P_D and P_{FA} through the test statistic $T(k)$. A longer sensing time implies utilizing more samples for PSD estimation, which results in a smaller estimation error of $T(k)$ and thus higher P_D and lower P_{FA} . The detection threshold $\gamma(k)$ is strongly related to the sensing rates. The threshold determines the power-detector operating point, where $\gamma(k)$ is set usually according to a desired P_{FA} . Underestimating the threshold leads to a higher P_D but also higher P_{FA} ; overestimation

leads to a lower P_{FA} and lower P_D . If the sensing time is to be minimized for a given detection rate, an accurate formulation for the detection threshold is required.

According to [35] and [36-37], the sensing time and detection threshold are determined by the target SNR, P_D , P_{FA} , and the measured noise power, σ_{vf}^2 . The formulations of the sensing time M and detection threshold $\gamma(k)$ are given by

$$M = \left(\frac{Q^{-1}(P_{FA}) - Q^{-1}(P_D)}{SNR} - Q^{-1}(P_{FA}) \right)^2, \quad (2.5)$$

and

$$\gamma(k) = \left(Q^{-1}(P_{FA}) \cdot \sqrt{M} + M \right) \cdot \sigma_{vf}^2(c). \quad (2.6)$$

2.2 Overview of Conventional Channelization Schemes

Channelization of a wideband spectrum into several channels can be realized in several ways. Directly feeding time-domain samples into an FFT to decouple the channels in the frequency domain is the most straightforward method. This process can be also viewed as applying a rectangular window before the FFT, which only provides a 13.6 dB suppression of adjacent-band interferer [38]. For a 30 dB SNR sinusoidal signal that is 1.5 bin (300 kHz if the frequency resolution is 200 kHz) away from the band of interest, the detected-band interfering power is still 16.4 dB. It leads to a severe interference if the signal strength of the detected band is only -5 dB SNR. As a result, the filtering capability has to be improved to reduce the adjacent-band interfering power.

To enhance the filtering, a time-domain window (other than a rectangular window) can be applied to the time-domain samples before feeding them into the FFT. This leads

to a degradation in the frequency resolution and SNR loss [38]. Alternatively, using a larger time-domain window with a larger FFT size could compensate for such degradation. Both the sensing time and the computational complexity would be increased due to the larger FFT size. Channelizing the overlapped time-domain samples could reduce the sensing time, but this requires multiple FFT processors operating in parallel and still needs a larger FFT size. Such an idea is not practical, considering the increased computational complexity and hardware cost.

Differences in the above methods can be explained by analyzing the test statistic expression (2.1). The same formula applies to all the methods, but the FFT output, $\hat{X}_m(k)$, has different expression for the different power detection methods, as given by

$$\hat{X}_m(k) = \sum_{n=0}^{N-1} w[n]x[n + m(1-d)N]e^{-j\frac{2\pi nk}{N}}, \quad (2.7)$$

where $x[n]$ is the time-domain sample with time index n , $w[n]$ is the normalized window coefficient, and d is the overlapping ratio between two time-domain N -sample blocks.

2.3 Overview of Previous Work

Wideband spectrum sensing based on power spectrum density (PSD) is generally used. In [39-44], wavelet transforms were used to estimate the PSD over a wide frequency range. The multi-resolution feature of wavelets is beneficial, but a wavelet function needs to be specifically designed for the signals to be detected, which is not feasible in real-time sensing and blind signal detection. [45-48] used filter bank with eigenvalue decomposition for PSD estimation. The computation of eigenvalue decom-

position, however, renders hardware realization impractical. PSD estimation through FFT with poly-phase filter bank [49-50] was applied for the simplicity in hardware implementation. However, [49-50] underestimated the detection threshold and sensing time, thus reducing the CR system throughput.

Applying a proper detection threshold is critically important for reliable detection when strong adjacent-band users are present. Methods to model detection threshold were given in [51-52]. Those models are inadequate as they underestimate the threshold in the presence of strong adjacent-band primary users. In fact, no existing literature considers the effects of adjacent-band interfering power.

Silicon realization of spectrum sensing is limited. [53-57] have demonstrated spectrum sensing for single narrow-bandwidth signal. [53-55] use analog correlators for the IEEE 802.22 standard with 1 MHz and 10 MHz sensing bandwidths with frequency resolution of 100 kHz. [56] proposes a multi-resolution spectrum-sensing processor with a reconfigurable filter for resolution from 12.5 kHz to 400 kHz. In [53-56], only power spectrum density (PSD) estimation is performed, while the sensing time and detection threshold are determined offline. [57] proposes a reconfigurable engine for multi-purpose sensing up to 6 GHz, but the instantaneous sensing bandwidth is only 20 MHz.

2.4 Summary

This chapter illustrates conventional wideband spectrum sensing algorithms for cognitive radios. However, with the presence of strong blockers, these algorithms are not energy efficient and unable to guarantee reliable weak signal detection within a limited processing time. In the next chapter, several energy-efficient wideband spectrum sensing algorithms will be presented, while these algorithms guarantee reliable weak signal detection with the presence of strong interferers.

CHAPTER 3

Wideband Speccturm-Sensing Algorithms and Real-Time Testbed

This chapter develops algorithms for weak signal detection in the presence of strong adjacent-band primary users. We propose a multitap-windowed frequency-domain power detector, which reduces the spectral leakage, maintains frequency resolution, and mitigates SNR loss simultaneously. Formulations of the sensing time and the detection threshold are presented to achieve reliable sensing rates with minimum sensing time. The proposed algorithms are experimentally validated on an FPGA-based radio testbed. Finally, we project the power and area requirements for a future ASIC implementation.

3.1 Wideband Spectrum-Sensing Design Problem

In order to define challenges addressed in this work, this section introduces the system model and the design specifications. The system model and these specification will be used in Chapter 3 and Chapter 5 as a design example.

3.1.1 System Model

We consider a wideband signal composed of several non-overlapping narrowband primary users, where each primary user has the same bandwidth and modulation scheme. Additive white Gaussian noise (AWGN) applies uniformly across the band. As a result, the noise power in all the individual narrowband channels is statistically equal.

3.1.2 Design Specification

The physical-layer design specifications are introduced here. The set of specifications consists of radio bandwidth, frequency resolution, detection sensitivity, sensing time, P_D and P_{FA} . The radio bandwidth dictates the minimum sampling frequency of the ADC and also limits the maximum number of channels that can be sensed at a time. The frequency resolution sets the minimum signal bandwidth of the detected primary user and the required FFT size. The detection sensitivity determines the minimum SNR for reliable detection specified by P_D and P_{FA} . The sensing time can be used to improve the reliability of signal detection but at the same time affects the cognitive radio system throughput. The detection sensitivity, sensing time, P_D and P_{FA} jointly depend on the specific signal detection algorithm.

Next, we define the hardware design constraints for our DSP baseband spectrum-sensing processor. State-of-the-art ADCs can operate at 500 MHz with 10-bit resolution [58]. Thus, we target a 200-MHz radio bandwidth for a feasible low-power and high-resolution wideband ADC. A 200-MHz specification relaxes the linearity require-

ment of the analog front end while still providing a considerable channel bandwidth. We target a 200-kHz frequency resolution, which means a 1024-point FFT is required to channelize the spectrum. This resolution would be sufficient to sense wireless microphones in the TV band or GSM signals.

Since power detection requires minimum a priori knowledge of the primary user signal, the sensitivity of power detection is drastically limited by noise uncertainty [59-64]. Thus, signals cannot be detected at arbitrarily low SNRs by increasing the sensing time. According to [59], the signal SNR should be at least -6 dB for a 0.5 dB noise uncertainty. We apply a 1 dB margin to accommodate for noise uncertainty, which brings the minimum SNR to -5 dB. The sensing time below 50 ms is our design choice for a good tradeoff between throughput loss due to sensing and reliability in primary user detection [65]. The target P_D and P_{FA} are 0.9 and 0.1 , respectively, based on the IEEE 802.22 draft standard specification for the detection of DTV signals and wireless microphones. The maximum tolerable interferer-to-noise ratio (INR) is defined by the transmitter frequency mask. As shown in Fig. 1, the interfering power is composed of leakage power and interferer in-band power. In this work, we only consider the spectral leakage by constraining INR to the values where interferer in-band power can be neglected. As a result (shown in Fig. 3.1), the adjacent-band signal power is constrained by:

$$INR_{max} = Attenuation(BW_{PU}) + SNR_{min} - Margin \quad (3.1)$$

Take the wireless microphone [66] as an example: the attenuation of the brick-wall

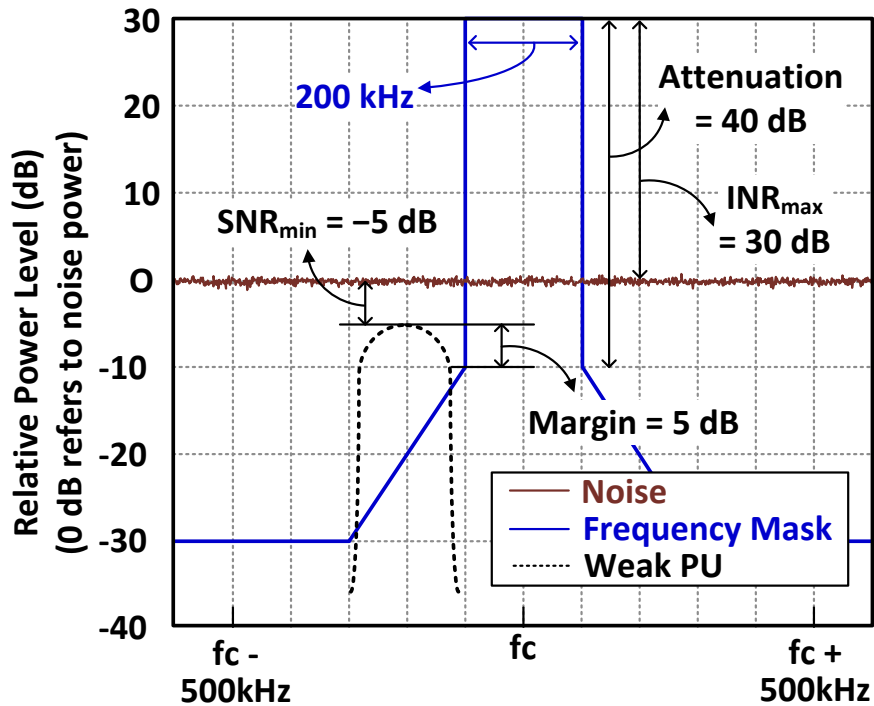


Figure 3.1: Maximum tolerable INR. For a 200 kHz wireless microphone primary user, the maximum INR is constrained by 30 dB.

portion is 40 dB, the minimum SNR is -5 dB, and we keep 5 dB for the margin. As a result, in our system the maximum INR for the wireless microphone primary user is constrained by 30 dB. The design specification is summarized in Table 3.1

3.1.3 Design Challenges

The development of proposed wideband sensing algorithms can be motivated by observing Fig. 3.2, which shows the receiver operating characteristic curves for the conventional rectangular-windowed FFT method. A -5 -dB SNR signal is detected between two adjacent-band primary users with a 30-dB INR. The solid line is the sensing-

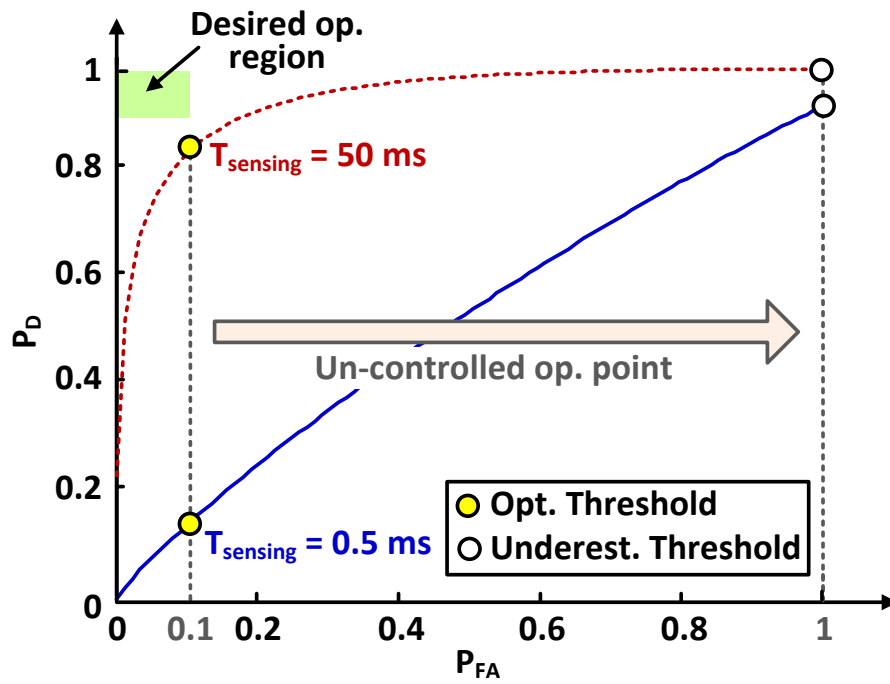


Figure 3.2: Receiver operating characteristic of the conventional rectangular-windowed FFT power detector for sensing times of 0.5 ms (solid line) and 50 ms (dashed line). The strong adjacent-band interferers and the weak band-of-interest signal are 30-dB INR and -5 -dB SNR, respectively.

Table 3.1: Wideband CR System Specification

Radio bandwidth	200 MHz
Frequency resolution	200 kHz
Sensitivity	$SNR \geq -5$ dB
Sensing time	$T_{sensing} \leq 50$ ms
False-alarm probability	$P_{FA} \leq 0.1$
Detection probability	$P_D \geq 0.9$
Interferer-to-noise ratio	$INR \leq 30$ dB

rate performance for a 0.5-ms sensing time. The dashed line shows the maximum attainable performance using this method, for a sensing time set to the maximum value of 50 ms. The results are still outside the desired operating region bounded by $P_D \geq 0.9$ and $P_{FA} \leq 0.1$. This is because the filtering capability of the rectangular window is not adequate for the case of adjacent-band primary users with $INR \geq 20$ dB. Clearly, the conventional FFT power detector cannot achieve reliable signal detection within constrained sensing time in the presence of strong interferers.

Adjusting the sensing time alone is sufficient to improve the detection rates. Allocating equal sensing time to all channels is also sub-optimal, even if the sensing time could be increased to accommodate the worst-case channel. Since different channels may experience different levels of the interfering power from the adjacent-band primary users, equal sensing time would be an overestimate for many channels. An overestimated sensing time not only degrades the access time to available channels, it also

increases the computational complexity and radio energy consumed in sensing. Therefore, every channel should adapt its own sensing time to the channel condition in order to maximize the throughput and minimize energy.

Conventional methods for the estimation of detection threshold are inadequate in the presence of strong interferers. The detection threshold based on (2.6) fails to reach $P_D \geq 0.9$, as shown in Fig. 3.2 by the gray circle marks. An underestimate of the threshold results in an increased P_{FA} , as shown by the white circle marks. Thus, the formulation of detection threshold in Eq.(2.6) must be modified to achieve $P_{FA} \leq 0.1$ and $P_D \geq 0.9$.

3.2 Proposed Wideband Spectrum-Sensing Algorithms

This section proposes algorithms that mitigate the design challenges associated with wideband spectrum sensing in the presence of strong adjacent-band interferers. Theoretical results are presented and verified by simulations to show that the proposed algorithms meet the system specifications defined in Table 3.1.

When strong primary users are present in adjacent bands, the resulting interfering power in the band of interest has to be reduced in order to reduce the required sensing time. The interfering power also has to be estimated to enable the adjustment in the sensing time and the detection threshold according to channel conditions. These two techniques combined ensure reliable detection within the P_D and P_{FA} constraints with minimum sensing time.

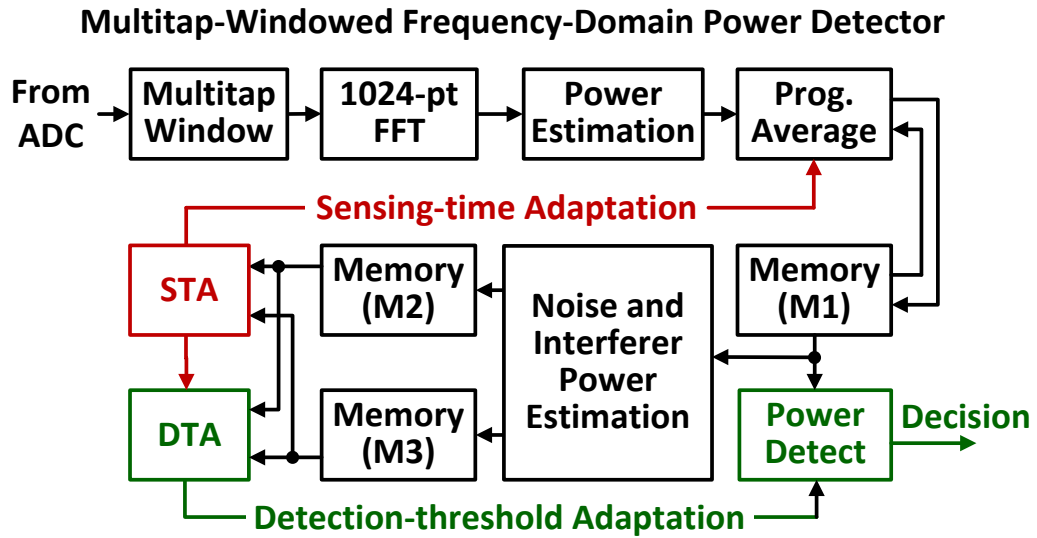


Figure 3.3: Block diagram of the proposed wideband spectrum sensing processor. The processor consists of a multitap-windowed frequency-domain power detector for PSD estimation, sensing-time adaptation for optimal sensing time, and detection-threshold adaptation for desired sensing rate.

Fig. 3.3 shows the block diagram of the proposed spectrum-sensing processor. It consists of a multitap-windowed frequency power detector, a sensing-time and detection-threshold adaptation blocks. The objectives of the algorithm design are to provide reliable detection rates with minimum sensing time and low hardware cost. The algorithmic steps are outlined next.

Fig. 3.4 is a high-level view of the proposed algorithm. The procedure starts by turning off the RF antenna to perform noise power calibration. Next, coarse sensing is performed using 64 averages (corresponding to 0.3 ms for a sampling frequency of 200 MHz). The coarse sensing is needed because not all channels need the same number of averages. The next two stages perform the estimation of in-band interfering power,

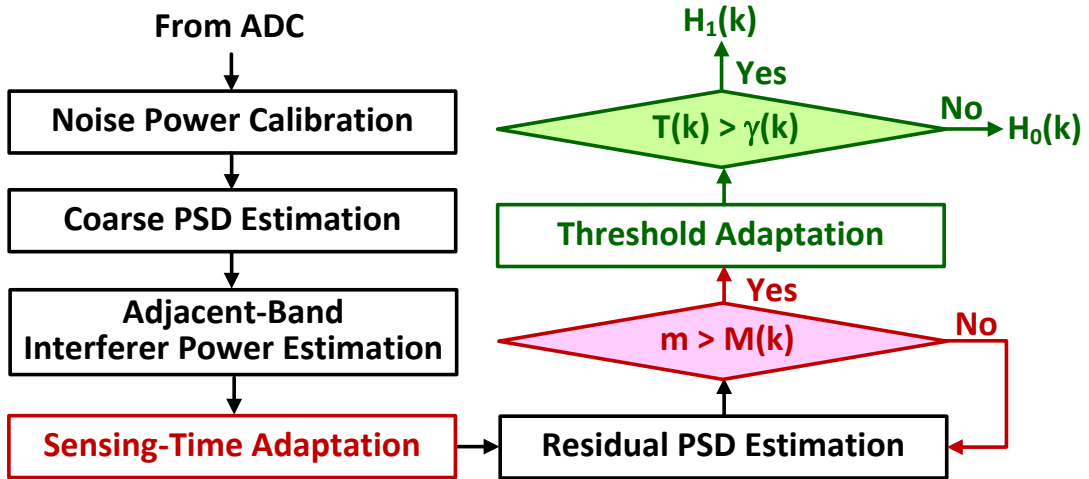


Figure 3.4: Wideband spectrum sensing procedure.

based on which sensing time is adapted for each channel. The in-band interfering power is projected by the corresponding estimated adjacent-band interferer powers [67]. The fifth step is to perform the PSD estimation. In this step, different sensing times from the previous step are applied to different channels to enhance the system throughput and reduce the power consumption. In the final stage, detection-threshold adaptation is performed. The detection threshold for each channel is adapted to the estimated interfering power and the corresponding number of averages. The decision about the presence of primary users is made right after the threshold is determined. The key algorithmic blocks are described below.

3.2.1 PSD Estimation Using Multitap-Windowed FFT Processing

In our proposed multitap-windowed frequency-domain power detector (MW-FPD), we apply a multitap window to overcome the spectral leakage problem of the FFT. The

traditional approach is to use either a windowed frequency-domain power detector (W-FPD) or a windowed-overlapped frequency-domain power detector (WO-FPD). Both methods require 2048-point FFT to channelize the time-domain samples. The W-FPD necessitates a long sensing time, which can be mitigated in the WO-FPD by using overlapped time-domain samples and two 2048-point FFT blocks. The idea in MW-FPD is to overlap and add the samples in time domain and channelize the resulting samples using a single 1024-point FFT. The overlap-and-add approach allows the use of a time-domain window with size that is larger than the FFT size, as shown in Fig. 3.5. The longer time-domain window simultaneously reduces the interfering power and compensates the degradation in frequency resolution.

The detection rule of MW-FPD can be still formulated by (2.1), but $\hat{X}_m(k)$ is given by

$$\hat{X}_m(k) = \sum_{n=0}^{N-1} \left(\sum_{p=0}^{P-1} w[n + pN]x[n + pN + mN] \right) e^{-j\frac{2\pi nk}{N}}, \quad (3.2)$$

where p is the tap index and P is the number of taps of the multitap window. (3.2) represents using $P \times$ longer time domain window and then overlapping-and-adding the windowed samples to keep the FFT size. A Gaussian assumption can be also applied to the test statistic. Therefore, (2.3) and (2.4) can be applied to model the sensing-rate performance in the MW-FPD method.

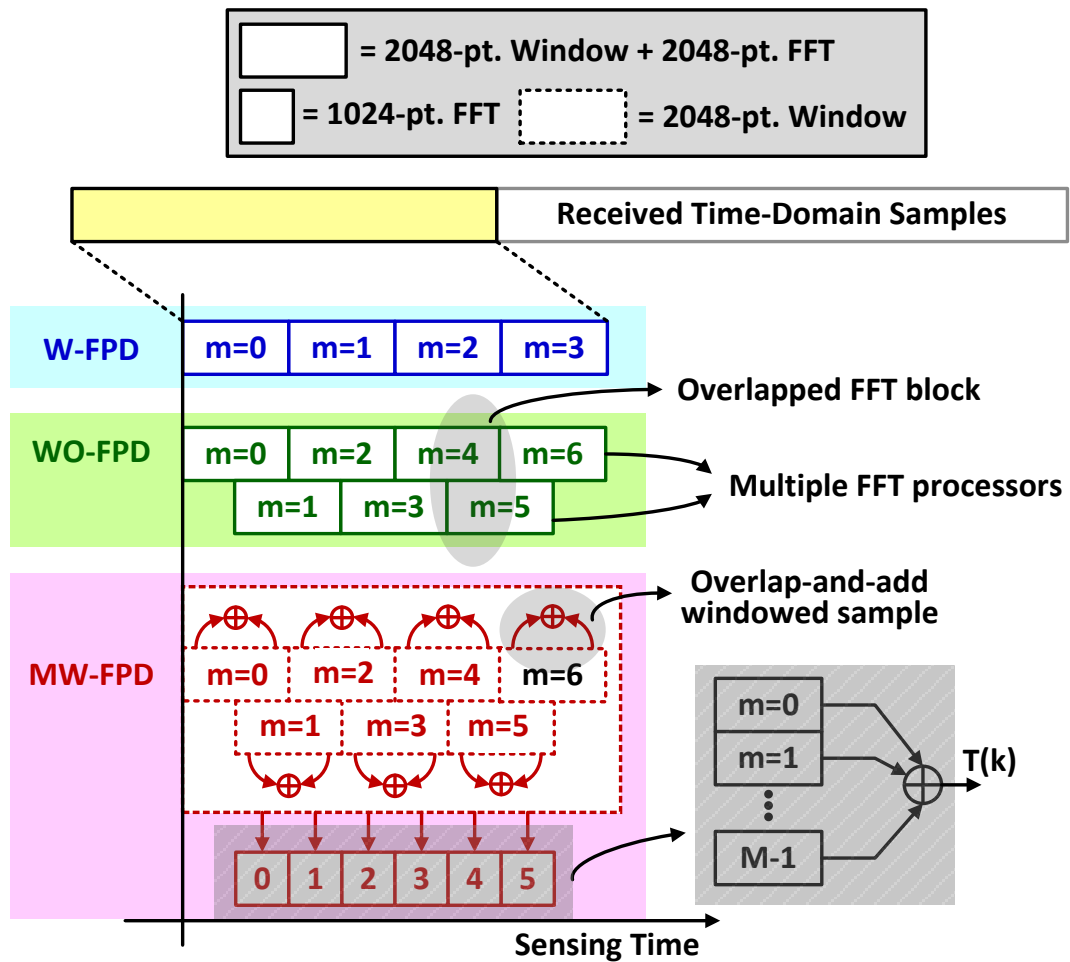


Figure 3.5: Detection procedure for the windowed frequency-domain power detector (W-FPD), the windowed-overlapped frequency-domain power detector (WO-FPD), and the proposed multitap-windowed frequency domain power detector (MW-FDP).

3.2.2 Power Detector Matrix Model

In this subsection, we generalize (2.1) to include conventional FFT frequency-domain power detector (C-FPD), windowed frequency-domain power detector (W-FPD), and multitap-windowed frequency-domain power detector (MW-FPD). A matrix formulation for the test statistic $T(k)$ facilitates a unified treatment of the performance of all three estimators.

In the following expressions, \mathbf{F} is an $N \times N$ FFT matrix. \mathbf{S}_k is a $N \times N$ selection matrix, which is used to select the k -th bin of the FFT. \mathbf{S}_c is given by,

$$\mathbf{S}_k = \text{diag}(\mathbf{e}_k). \quad (3.3)$$

\mathbf{e}_k is an $N \times 1$ vector, where the k -th entry is one and others are zeros. When a W-FPD is applied to reduce the spectral leakage between FFT bins, the matrix form for the time-domain window is

$$\mathbf{W} = \text{diag}([w_0, \dots, w_{N-1}]), \quad (3.4)$$

where w 's are the normalized window coefficients that satisfy $1/N \sum_{n=0}^{N-1} |w[n]|^2 = 1$. In Sec. 3.2.1, the MW-FPD uses a longer time-domain window (sample size $> N$) to improve leakage reduction and utilizes an overlap-and-add approach to form an N -point samples that allows using an N -point FFT for channelization. Such windowing and overlapping-and-adding processes are given by concatenating several $N \times N$ windowing matrices to form a $N \times NP$ block matrix. The formulation of this block matrix is as follows. First, we decompose the NP coefficients in (3.2) into P $N \times 1$ coefficient vectors,

where the p -th coefficient vector is $\mathbf{w}_p = [w_{pN}, \dots, w_{pN+N-1}]^T$. The NP coefficients are power normalized, such that $1/N \sum_{n=0}^{NP-1} |w[n]|^2 = 1$. Then, the P coefficient vectors form $P N \times N$ windowing matrices, with the p -th windowing matrix $\mathbf{M}_p = \text{diag}(\mathbf{w}_p)$. Finally, we combine the windowing matrices for the multitap-windowed process,

$$\mathbf{W}_{MW} = [\mathbf{W}_0, \dots, \mathbf{W}_{P-1}]. \quad (3.5)$$

Combining (3.3), (3.4), and (3.5), we represent the test statistic $T(k)$ for C-FPD, W-FPD, and MW-FPD respectively by

$$\begin{aligned} T_C(k) &= \frac{1}{M} \sum_{m=0}^{M-1} (\mathbf{S}_k \mathbf{F} \mathbf{x}_m)^H \mathbf{S}_k \mathbf{F} \mathbf{x}_m = \frac{1}{M} \sum_{m=0}^{M-1} \mathbf{x}_m^H \Lambda_{k,C} \mathbf{x}_m, \\ T_W(k) &= \frac{1}{M} \sum_{m=0}^{M-1} (\mathbf{S}_k \mathbf{F} \mathbf{W} \mathbf{x}_m)^H \mathbf{S}_k \mathbf{F} \mathbf{W} \mathbf{x}_m = \frac{1}{M} \sum_{m=0}^{M-1} \mathbf{x}_m^H \Lambda_{k,W} \mathbf{x}_m, \\ T_{MW}(k) &= \frac{1}{M} \sum_{m=0}^{M-1} (\mathbf{S}_k \mathbf{F} \mathbf{W}_{MW} \mathbf{x}_{m,MW})^H \mathbf{S}_k \mathbf{F} \mathbf{W}_{MW} \mathbf{x}_{m,MW} \\ &= \frac{1}{M} \sum_{m=0}^{M-1} \mathbf{x}_{m,MW}^H \Lambda_{k,MW} \mathbf{x}_{m,MW}, \end{aligned} \quad (3.6)$$

where the $N \times 1$ time-domain sample vector $\mathbf{x}_m = [x[mN], \dots, x[mN+N-1]]^T$. The sample vector $\mathbf{x}_{m,MW}$ in (3.6) concatenates $P N \times 1$ time-domain sample vector \mathbf{x}_m , in the form of $\mathbf{x}_{m,MW} = [\mathbf{x}_m^T, \dots, \mathbf{x}_{m+P-1}^T]^T$. $\Lambda_{k,C}$, $\Lambda_{k,W}$, and $\Lambda_{k,MW}$ combine the FFT, bin selection, and windowing operations. Each matrix weights the squared 2-norm of the time-domain samples to produce the test statistics. These matrix formulations are used to compute the detection threshold and required sensing time. More details are represented in the next subsection.

3.2.3 Interfering Power Estimation

This subsection quantifies the performance degradation caused by strong adjacent-band interferers. Assume the interferer i is allocated in bin l and MW-FPD is utilized for channelization, the expected interfering power in bin k is

$$\sigma_{if}^2(k) = E \left[\left| \sum_{n=0}^{N-1} \left(\sum_{p=0}^{P-1} w[n+pN]i[n+pN+mN] \right) e^{-j\frac{2\pi nk}{N}} \right|^2 \right]. \quad (3.7)$$

Since the interfering power $\sigma_{if}^2(k)$ is relatively small and co-exists with signal power $\sigma_{sf}^2(k)$ and noise power $\sigma_{vf}^2(k)$, σ_{if}^2 cannot be measured independently of $\sigma_{sf}^2(k)$ and $\sigma_{vf}^2(k)$. Since the ratio of main-lobe power and side-lobe power is constant regardless of the center frequency [67], the interfering power $\sigma_{if}^2(k)$ can be estimated from its main-lobe power $\sigma_{if}^2(l)$. Assume the interferer power $\sigma_{if}^2(l)$ dictates the power in bin l , the estimated interferer power in bin l is approximated by

$$\hat{\sigma}_{if}^2(l) = \frac{1}{M} \sum_{m=0}^{M-1} \mathbf{x}_{m,MW}^H \Lambda_{l,MW} \mathbf{x}_{m,MW}, \quad (3.8)$$

and the estimated interfering power in bin k is given by

$$\hat{\sigma}_{if}^2(k) = \hat{\sigma}_{if}^2(l) \cdot \beta, \quad (3.9)$$

where $\beta = \mathbf{g}_i^H \Lambda_{k,MW} \mathbf{g}_i$. The column vector \mathbf{g}_i represents the time-domain impulse response of the strong interferer, which is related to the modulation scheme of the interferers (primary users). In addition, this column vector is normalized to the band of detection, i.e. $\mathbf{g}_i^H \Lambda_{l,MW} \mathbf{g}_i = 1$.

3.2.4 Detection-Threshold and Sensing-Time Adaptations

The theoretical detection threshold can be derived from (2.3). In order to dynamically update the detection threshold, the threshold should be a function of measured powers, namely the noise power and interfering power (if adjacent-band interferers are present). We proposed a framework to derive the mean and variance of $T(k)$ under $H_0(k)$ and $H_1(k)$ [67]. For illustration, we only discuss the case of MW-FPD. Others can be derived using analogy.

We use a matrix formulation for (3.2), which is discussed in Sec. 3.2.2, to derive the mean, variance and covariance of $|\hat{X}_m(k)|^2$ as in [68]. The outer summation process of (2.1) is modeled by a P-dependent sequence [69]. The correlation between FFT frames $|\hat{X}_m(k)|^2$ is non-zero due to overlapping samples in the time domain. Combining the matrix formulation, the P-dependent sequence model, and the Gaussian assumption for (2.1), the mean and variance of the test statistic in H_1 and H_0 are

$$E[T(k)|H_0(k)] = \sigma_{vf}^2(k) + \sigma_{if}^2(k), \quad (3.10)$$

$$\text{Var}[T(k)|H_0(k)] = \frac{1}{M} \cdot \alpha \cdot (\sigma_{vf}^2(k) + \sigma_{if}^2(k))^2, \quad (3.11)$$

and

$$E[T(k)|H_1(k)] = \sigma_{sf}^2(k) + \sigma_{vf}^2(k) + \sigma_{if}^2(k), \quad (3.12)$$

$$\text{Var}[T(k)|H_1(k)] = \frac{1}{M} \cdot \alpha \cdot (\sigma_{sf}^2(k) + \sigma_{vf}^2(k) + \sigma_{if}^2(k))^2, \quad (3.13)$$

where $\alpha = \frac{\text{tr}(\Lambda_{k,MW} + \Gamma)}{\text{tr}(\Lambda_{k,MW})^2}$ is the fitting factor determined by the modulation scheme of the primary user and the time-domain multitap window. Γ is the covariance term that models the increased variance due to the correlation between FFT frames [67].

Using (3.10), (3.11), (3.12), (3.13), and the Gaussian assumption, the false-alarm and detection probabilities are

$$P_{FA}(k) = Q \left(\frac{\gamma(k) - \left(\sigma_{vf}^2(k) + \sigma_{if}^2(k) \right)}{\sqrt{1/M \cdot \alpha \cdot \left(\sigma_{vf}^2(k) + \sigma_{if}^2(k) \right)^2}} \right), \quad (3.14)$$

$$P_D(k) = Q \left(\frac{\gamma(k) - \left(\sigma_{sf}^2(k) + \sigma_{vf}^2(k) + \sigma_{if}^2(k) \right)}{\sqrt{1/M \cdot \alpha \cdot \left(\sigma_{sf}^2(k) + \sigma_{vf}^2(k) + \sigma_{if}^2(k) \right)^2}} \right). \quad (3.15)$$

From (3.14), the detection threshold is

$$\gamma(k) = \left(Q^{-1}(P_{FA}) \sqrt{\alpha \cdot M(k)} + M(k) \right) \cdot \Sigma_0(k). \quad (3.16)$$

$\Sigma_0(k) = \sigma_{vf}^2(k) + \sigma_{if}^2(k)$, the interfering power is estimated using the power measurements in the adjacent bands, as discussed in Sec. 3.2.3.

Using the methodology described above, the sensing time for each channel is obtained as

$$\begin{aligned} M(k) &= \alpha \left((1 + \psi(c)) \frac{Q^{-1}(P_{FA}) - Q^{-1}(P_D)}{SNR} - Q^{-1}(P_D) \right)^2 \\ &= \alpha \left(\frac{Q^{-1}(P_{FA}) - Q^{-1}(P_D)}{SINR} - Q^{-1}(P_D) \right)^2, \end{aligned} \quad (3.17)$$

where $\psi(k) = \sigma_{if}^2(k)/\sigma_{vf}^2(k)$. According to (3.17), the sensing time is approximately proportional to inverse square of $SINR$. In addition, correlation between FFT frames introduces $\alpha \times$ increase in sensing time.

Therefore, by adapting the detection threshold and sensing time to the adjacent-band interfering power, the proposed wideband spectrum-sensing processor is able to achieve the constrained sensing rates with minimum sensing time.

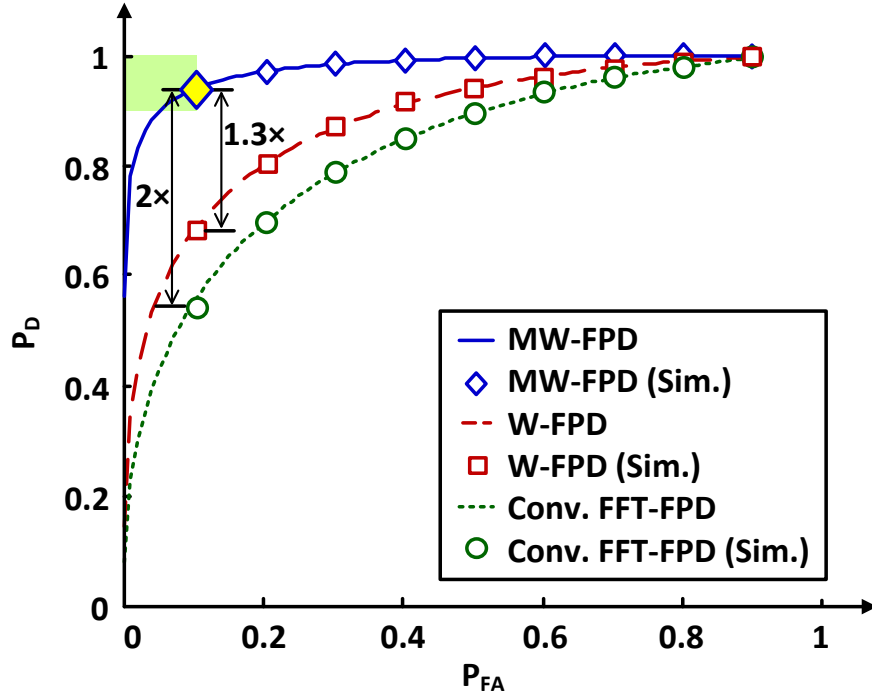
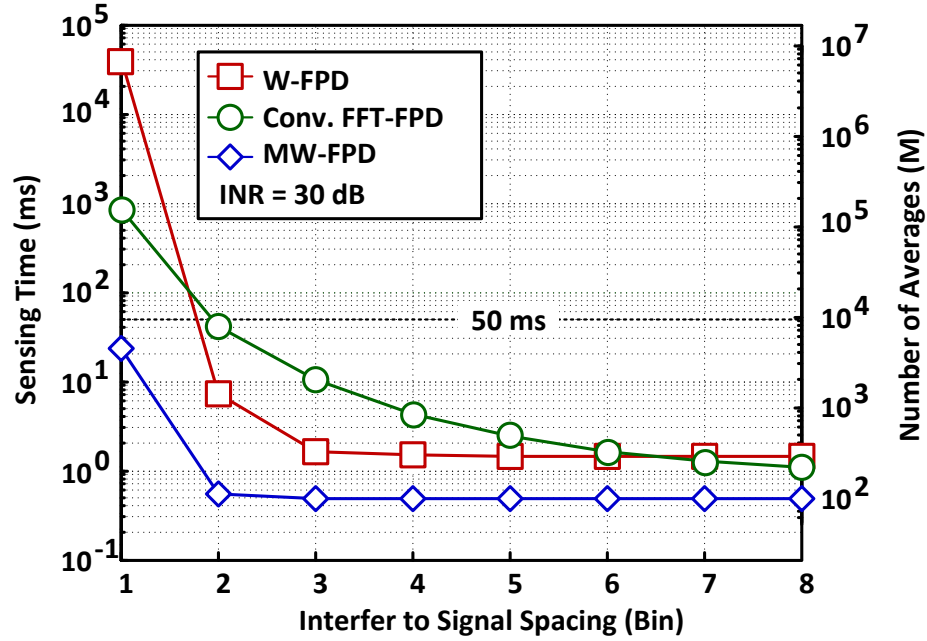


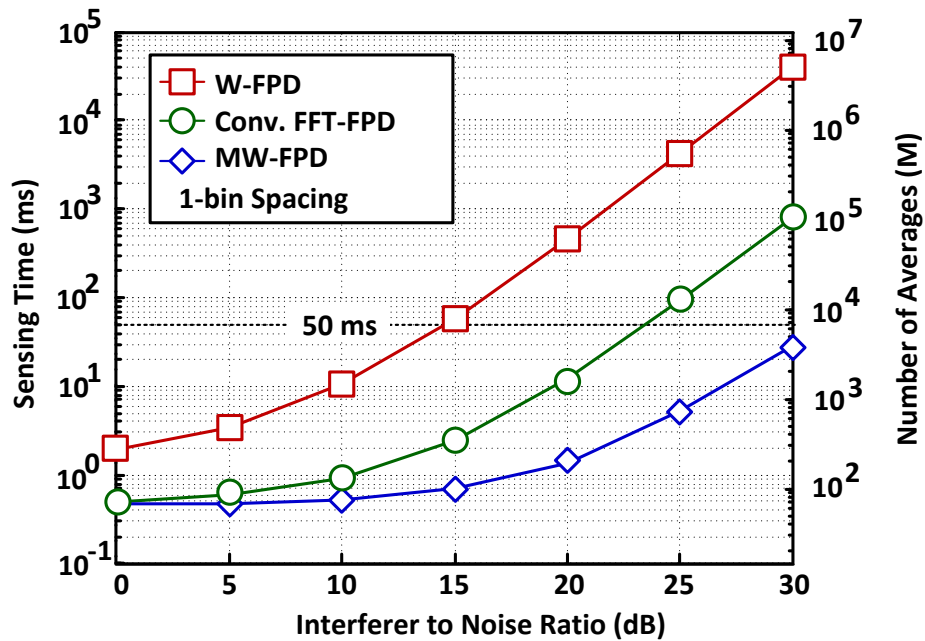
Figure 3.6: Receiver operating characteristic curves for different power detectors. The multitap-windowed frequency-domain power detector achieves the best detection rate.

3.3 Numerical Simulations

We verified the performance of the proposed algorithms by comparing the derived P_D and P_{FA} , and Monte Carlo simulations. The simulations are obtained assuming a signal strength in the band of interest to be -5 dB SNR. We also assume the worst-case wideband scenario, where two primary users are present in bands adjacent to the band of interest.



(a)



(b)

Figure 3.7: Required sensing time for conventional FFT power detector, windowed frequency-domain power detector, and multitap-windowed frequency-domain power detector as a function of interferer-to-signal spacing (a) and interferer-to-noise ratio (b). The proposed multitap-windowed power detector requires the least sensing time to reach a given detection rate.

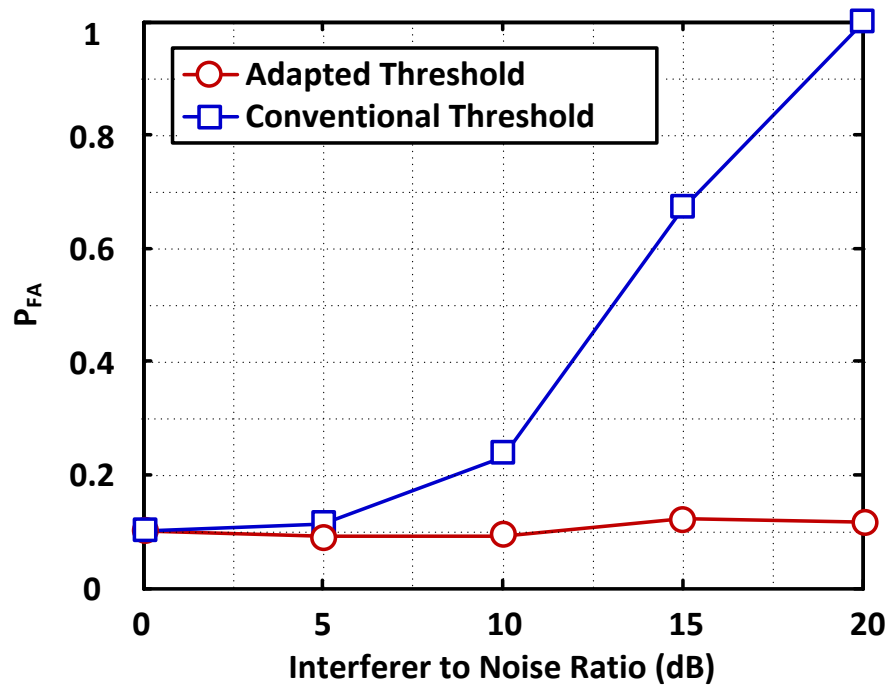


Figure 3.8: Influence of adaptation on P_{FA} . The adaptation threshold significantly lowers P_{FA} as compared to the use of conventional threshold. Both methods use MW-FPD for PSD estimation.

3.3.1 Improvement in ROC

Fig. 3.6 demonstrates the improvement in detection rate by means of MW-FPD as compared to conventional methods. To make a fair comparison, we assume a fixed hardware cost for the FFT processor, which is 1024 points for all methods. This is a good approximation of the overall hardware cost since the FFT block dominates the area. The WO-FPD requires two FFT cores, so we compare the conventional FFT, the W-FPD, and the proposed MW-FPD. A prolate-spheroidal window is applied as the time-domain window for both the W-FPD and the MW-FPD, but the W-FPD has a length of 1024 and the MW-FPD has a length of 2048. In this example, two primary users with 20 dB INR are placed two bins (400 kHz) away from the band of interest. The number of averages M in all three methods is 100, corresponding to 0.5 ms for a 200-MHz sampling frequency. The receiver operating characteristics in Fig. 3.6 indicate that only the proposed MW-FPD meets the sensing rate constraints within 0.5 ms. For a fixed $P_{FA} = 0.1$, MW-FPD achieves a 2x increase in P_D compared to the conventional FFT power detector due to better filtering, and a factor of 1.3 increase in P_D over the W-FPD due to the compensation of SNR loss.

3.3.2 Improvement in Sensing Time

The sensing-time improvement from the proposed MW-FPD detector is illustrated in Fig. 3.7. The impact of the frequency distance between the interfering primary users and band-of-interest signal is shown in Fig. 3.7 (a). The distance is 1 bin (200 kHz) to

8 bins (1.6 MHz), and the INR of the primary users is 30 dB. At 1-bin distance, only our proposed power detector meets the sensing time constraint (< 50 ms). Observe that the sensing time of the MW-FPD remains constant when the distance is larger than two bins (400 kHz). Therefore, when applying the multitap window, the spectral leakage is confined to the nearest bins for $\text{SNR} \geq -5$ dB and $\text{INR} \leq 30$ dB. Additionally, when the primary users are placed 8 bins away from the band of interest, the W-FPD requires more sensing time than the conventional FFT power detector due to the SNR loss inherent to the windowing process.

The impact of the interfering power on sensing time is shown in Fig. 3.7 (b), assuming that two primary users are placed one bin (200 kHz) away from the band of interest. All three power detectors adapt the sensing time to the corresponding in-band interfering power according to (3.17) to achieve fixed sensing rates ($P_{FA} = 0.1$ and $P_D = 0.9$). The proposed MW-FPD achieves more than a 10x reduction in sensing time compared to the conventional FFT power detector at an INR of 30 dB. The W-FPD requires the longest sensing time due to the frequency resolution that is larger than one bin. This highlights the necessity of using a multitap window to maintain a fine frequency resolution.

3.3.3 Improvement in False-Alarm Rate

Validation of the detection-threshold adaptation algorithm is illustrated in Fig. 3.8. The two primary users ($\text{INR} \leq 30$ dB) are placed one bin (200 kHz) away from the band of interest. Simulation results indicate that without threshold adaptation, P_{FA}

starts to rapidly increase once the INR exceeds 5 dB. The proposed threshold adaptation effectively satisfies the target false-alarm rate constraint ($P_{FA} \leq 0.1$) for $\text{INR} \leq 30$ dB (20 dB shown on the plot). For $\text{INR} \geq 10$ dB, the proposed adaptation algorithm achieves a decrease of P_{FA} by at least a factor of 2. Therefore, the algorithms proposed in this section demonstrate reliable signal detection in the presence of strong adjacent-band primary users.

3.4 Proposed VLSI Architecture and Experimental Results

In order to further validate the proposed algorithms, this section proposes a VLSI architecture and discusses its hardware complexity. The architecture is implemented on a real-time FPGA-based cognitive radio testbed to experimentally validate simulation results. Radio measurements in the ISM band show excellent matching of experimental and simulation results.

3.4.1 Proposed VLSI Architecture

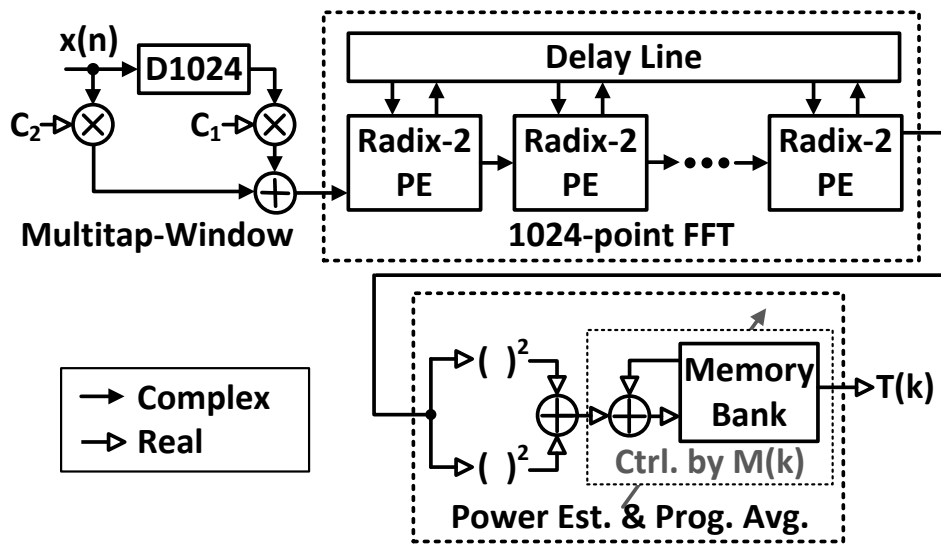
A VLSI architecture suitable for the proposed wideband spectrum-sensing processor is presented in this subsection. Fig. 3.9 (a) shows the architecture of the MW-FPD block. A pipelined FFT architecture provides the highest throughput with acceptable hardware cost. The window coefficients are programmed in a block random access memory (BRAM). This BRAM provides us with the flexibility to reconfigure the time-domain window. The block for PSD estimation and programmable averaging is con-

trolled by the number of averages $M(k)$. This block is disabled when the processing time exceeds the required sensing time for the corresponding channel.

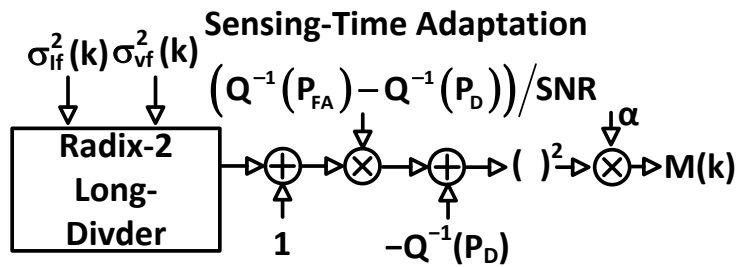
Architectures of the sensing-time adaptation (STA) and detection-threshold adaptation (DTA) blocks are shown in Figs. 3.9 (b) and (c), respectively. The pre-calculated parameters $Q^{-1}(P_{FA})$, $Q^{-1}(P_D)$, SNR, and α are programmed in a look-up table. This look-up table can be reprogrammed with other values of design parameters and sensing specifications. The DTA and STA blocks allow for dynamic tracking of channel conditions to maintain the P_D and P_{FA} specifications within the minimum sensing time. However, this adaptation adds processing-time latency, which degrades the overall sensing efficiency. To mitigate the processing-time latency overhead, a radix-2 long divider with 10 pipeline stages and a CORDIC-based square root with 8 pipeline stages are adopted in STA and DTA, respectively. A 10- μ s processing time overhead is achieved (for a 200-MHz sampling frequency) for both STA and DTA, which corresponds to 2% of the total sensing time if no adjacent-band interferer is present and 0.01% of the total sensing time if adjacent-band primary users with 30 dB INR are present.

3.4.2 Fixed-Length vs. Reconfigurable FFT

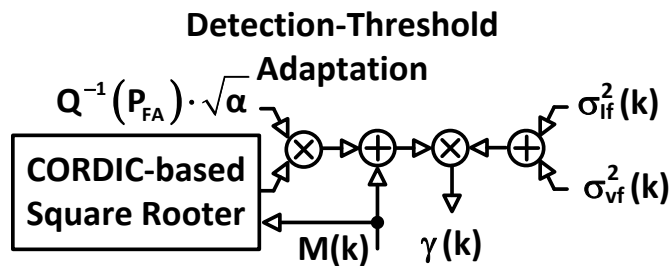
In practice, CR primary users can have different (distinct) signal bandwidths. Although we have assumed that all primary users have the same signal bandwidth in our system model, our proposed architecture can accommodate multiple frequency resolutions required by different communication standards. The 1024-point FFT processor channelizes the spectrum into several sub-bands with a resolution equal to the minimum



(a)



(b)



(c)

Figure 3.9: A VLSI architecture for the proposed spectrum sensing algorithms. (a) Multitap-windowed frequency-domain power detection block. (b) Sensing-time adaptation block. (c) Detection-threshold adaptation block.

signal bandwidth (200 kHz), and the channel-occupancy decision is made FFT-bin by FFT-bin. For a channel that occupies multiple FFT bins, the decision results of the corresponding bins are combined to identify the occupancy of the channel.

Another approach is to utilize a reconfigurable FFT processor that supports different FFT sizes to provide multiple frequency resolutions. In high-resolution mode, the FFT processor is reconfigured into a large FFT size to reduce the leakage power. In low-resolution mode, the FFT processor is reconfigured into a small FFT size to reduce the active area and power consumption.

A tradeoff between sensing performance (sensing rates and sensing time) and power consumption exists between these two approaches. For example, assume a 6 MHz DTV signal. In a 200 MHz bandwidth sensing, the required size of the FFT processor for the 6 MHz signal is 32 points. Using the reconfigurable architecture results in a 2x lower power consumption than using a 1024-point fixed-length FFT. This is due to the smaller FFT size and assuming the computational complexity is proportional to $\log_2(N)$. However, as shown in Fig. 3.10, the 1024-point fixed-length FFT provides higher spectral resolution and less interfering power. This higher resolution leads to better sensing rates and less sensing time, provided strong adjacent-band interferers are present. The reconfigurable FFT processor can therefore be used to explore the tradeoff between the sensing performance and power consumption. The fixed-length FFT always provides a higher frequency resolution and is less vulnerable to the environment with strong adjacent-band interferers. In this work, we target a wideband scenario with strong adjacent-band interferers, so we use the fixed-length FFT that meets our specifications.

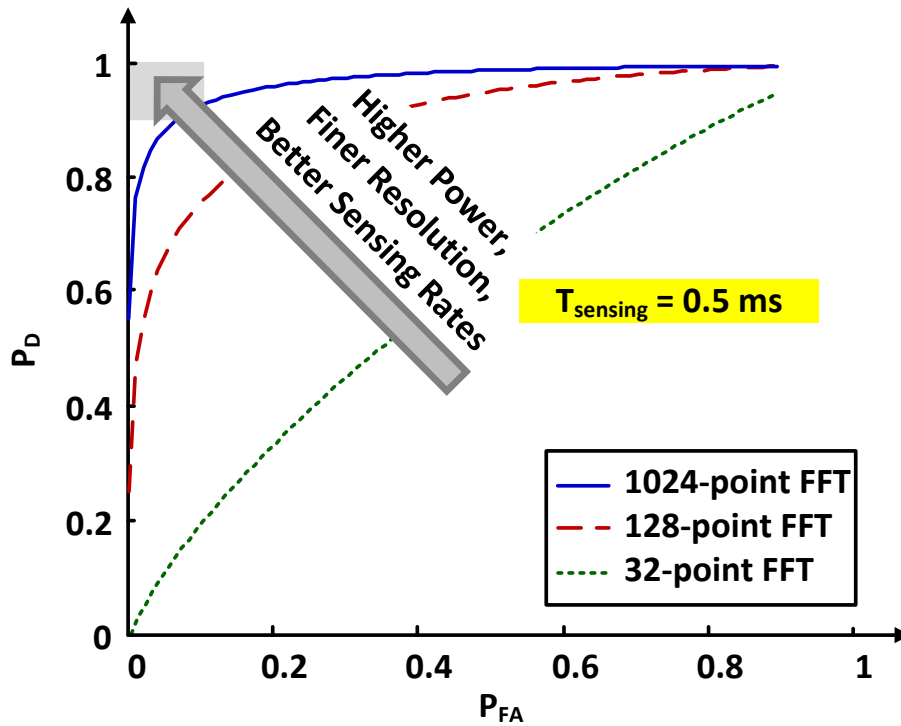


Figure 3.10: Improvement in detection performance as a function of frequency resolution for a 6 MHz DTV signal. Larger FFT size with higher frequency resolution results in better sensing rates.

3.4.3 Hardware Complexity

Comparisons of the relative hardware cost (area), sensing time and, computational complexity for several windowed power detector implementations are shown in Fig. 3.11 and Table 3.2. The conventional FFT power detector is not considered since it cannot satisfy the sensing-time requirement when strong adjacent-band primary users are present. Also note that the W-FPD and WO-FPD require a 2048-point FFT to meet the system specifications. Assuming that the same window is applied to the time-domain samples, the three power detectors would achieve the same sensing rate with the same M . The W-FPD requires the longest sensing time since a larger FFT processor is required to achieve the target frequency resolution. The WO-FPD reduces the sensing time by using two FFT processors, but has the highest hardware cost due to multiple FFT processors. The proposed MW-FPD achieves a 50% reduction in sensing time compared to the W-FPD and a 50% reduction in hardware cost compared to the WO-FPD. The computational complexity is calculated as $M \times N \times (C_{add} + C_{mult})$, where C_{add} and C_{mult} are the complexity of an adder and a multiplier, respectively. A 50% reduction in the computational complexity is achieved by using the MW-FPD compared to the W-FPD and WO-FPD. Therefore, the MW-FPD has the lowest sensing time, lowest computational complexity, and lowest hardware cost.

The hardware cost in terms of FPGA and ASIC resources was also estimated using chip synthesis. The FPGA hardware resource breakdown is summarized in Table 3.3. The proposed architecture requires 205 Kb of BRAM, 17.6 K slices and 39 embedded

Table 3.2: Hardware Complexity Estimates of W-FPD, WO-FPD, and MW-FPD

		W-FPD	WO-FPD	MW-FPD
Window	Memory	-	-	N
	C_{add}	-	-	1
	C_{mult}	1	1	2
FFT	Memory	$2N - 1$	$2 \cdot (2N - 1)$	$N - 1$
	C_{add}	$2\log_2(2N)$	$4\log_2(2N)$	$2\log_2(N)$
	C_{mult}	$\log_2(2N) - 1$	$2 \cdot (\log_2(2N) - 1)$	$\log_2(N) - 1$

multipliers. This is equivalent to a chip area of 0.98 mm^2 in a standard 65-nm CMOS process, as estimated from chip synthesis. The memory and logic blocks would occupy 0.82 mm^2 and 0.16 mm^2 , respectively. The estimated power consumption is 25.1 mW at 200 MHz from a 1-V supply. Table 5.2 compares the state-of-the-art processors [53-55] to our design. Our design can achieve a 200x wider detection bandwidth than [53-54] with a 2x higher power, and a 20x wider detection bandwidth than [55] with a 3x higher power. When optimized for low power using the methodology in [70], our design is estimated to operate below 8 mW. A low-power ASIC that integrates MW-FPD and sensing adaptation algorithms will be presented in Chapter 5.

3.4.4 Implementation on CR Testbed

The proposed algorithms were verified on a cognitive radio testbed, which consists of a Berkeley Emulation Engine 2 (BEE2) [71-73] and a custom RF front end [74]. The

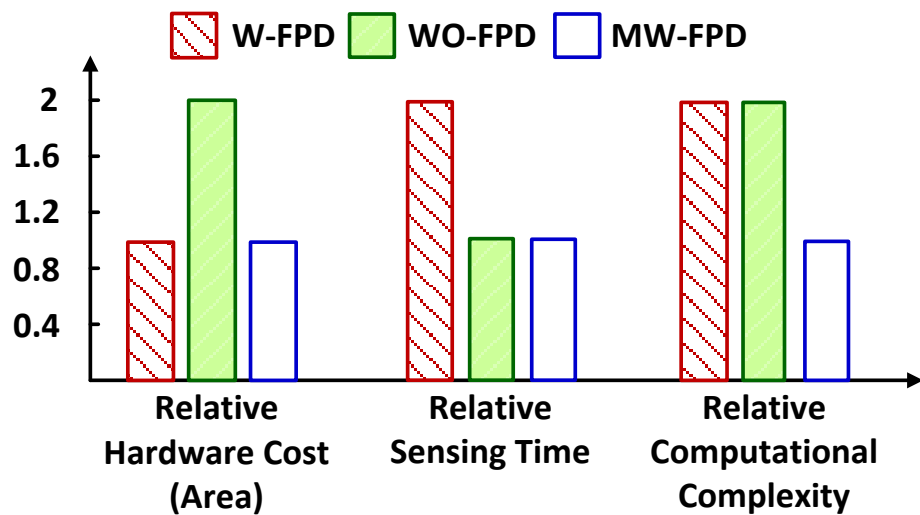


Figure 3.11: Comparison of relative hardware cost, sensing time, and computation complexity for W-FPD, WO-FPD, and MW-FDP.

BEE2 consists of five high-performance Xilinx Virtex-II FPGAs, and each FPGA has access to four XAUI multi-Gigabit interfaces that allow connectivity to external boards. We use those interfaces to connect the BEE2 to the RF front end, which allows us to verify our proposed algorithm in a true real-time radio. The radio front end operates in the unlicensed 2.4-GHz ISM band and can be tuned over 80 MHz. It also contains two 12-bit ADCs running at 64 MS/s for both the I and Q components giving a 64 MHz of bandwidth in the digital domain. Since the radio bandwidth in our testbed is limited by the ADC, we scale the target bandwidth from 200 MHz to 64 MHz. The resulting frequency resolution and sensing time constraint are therefore 62.5 kHz and 150 ms, respectively.

Table 3.3: FPGA Hardware Resource Breakdown for MW-FPD

	BRAM	Slices	Embedded Mults
Window	30 Kb	1.2 K	4
FFT Processor	80 Kb	13.5 K	27
Power Meas.	80 Kb	0.9 K	2
DTA + STA	15 Kb	2.0 K	6
Overall	205 Kb	17.6 K	39

Table 3.4: Comparison with State-of-the-Art Power-Detector ASICs

	[53-54]	[55]	This Work
Power	*11.7 mW	*8.4 mW	25.1 mW
Area	0.36 mm ²	0.36 mm ²	0.98 mm ²
Sensing Bandwidth	1 MHz	10 MHz	200 MHz

*Normalized to a 65-nm CMOS technology.

3.4.5 Experimental Results

FM-modulated signals in the ISM band were used to experimentally validate our simulation results on a radio testbed. A 50-kHz signal at a carrier frequency of 2.483 GHz was generated in the band of interest. The signal strength is -115 dBm, which corresponds to -5 dB SNR. Another signal generator provided the adjacent-band primary user signals, with the same bandwidth of 50 kHz and signal strength from -110 dBm to -80 dBm (corresponding to INR from 0 to 30 dB). The frequency spacing between

the two FM signals was 60 kHz, which corresponds to 1 FFT bin in the FPGA implementation. We measured the change in P_{FA} and detection threshold $\gamma(k)$, as well as the change in P_D and the number of averages $M(k)$, as we increased the signal strength of the adjacent-band primary users. The experiments were repeated 1000 times in order to get statistically relevant results for P_{FA} and P_D .

The proposed sensing-time adaptation algorithm is validated in Fig. 3.12. The lines with gray square markers represent the detection probability without sensing-time adaptation while the lines with gray circle markers correspond to the case with sensing-time adaptation. The lines with hollow circle markers show the adapted number of averages, which are related to sensing time as given by (2.2). Without sensing-time adaptation, P_D starts decreasing rapidly as the INR exceeds 15 dB. At 20 dB INR, P_D decreases by 25% when no adaptation is used. A 3x increase in the sensing time is required to raise P_D to the desired value of 0.9. The experimental results (solid lines) closely match the simulation results (dashed lines) discussed in Section 3.3.

Fig. 3.13 demonstrates the validity of the proposed detection-threshold adaptation scheme. The lines with gray square markers represents the false-alarm probability without threshold adaptation while the lines with gray circle markers illustrate the case with detection-threshold adaptation. The lines with hollow circle markers show the adapted detection threshold, which is normalized to the threshold when no adjacent-band primary user is present. Without threshold adaptation, P_{FA} shows a rapid increase as INR exceeds 5 dB. At an INR of 10 dB, P_{FA} increases by more than a factor of 2 when no adaptation is used. A 20% increase in the decision threshold is sufficient to compensate

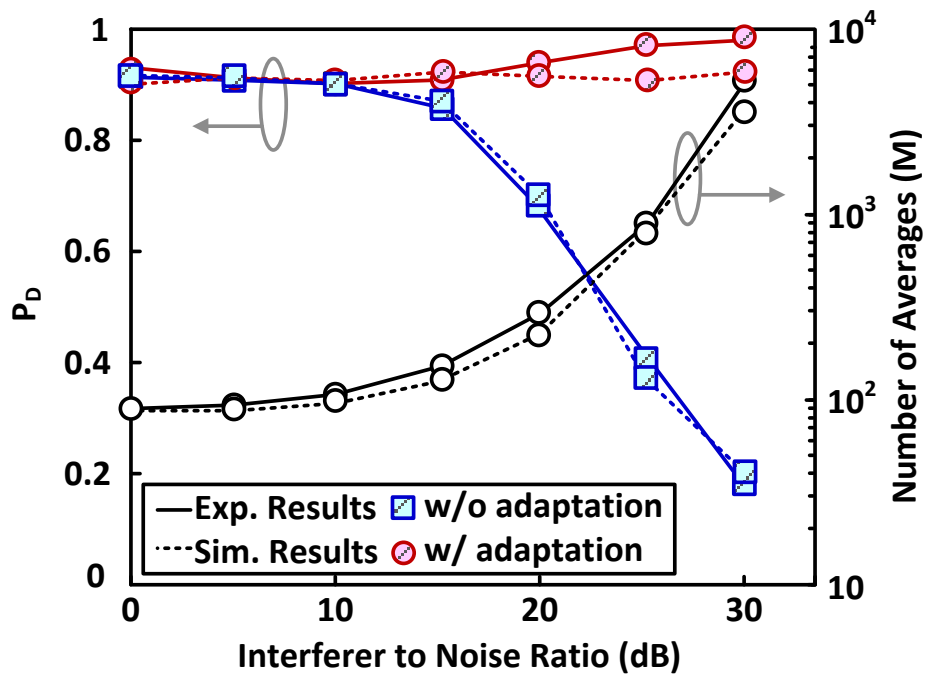


Figure 3.12: P_D and number of averages versus INR.

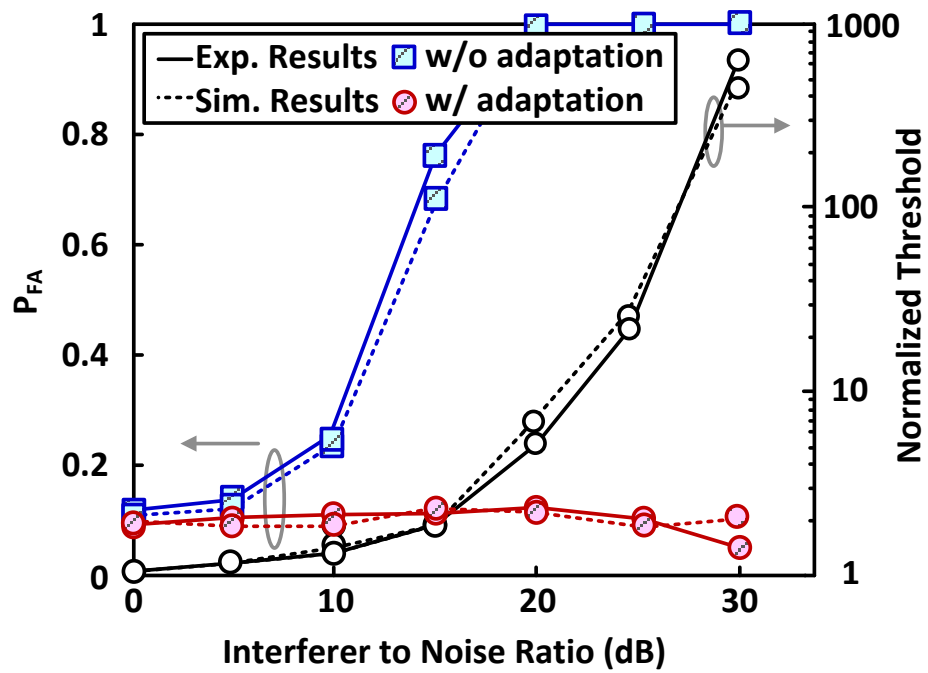


Figure 3.13: P_{FA} and detection threshold versus INR.

for this P_{FA} degradation, bringing it back to the desired value of 0.1. This validates that the P_{FA} is highly sensitive to the in-band interfering power when a signal detection is performed in a negative SNR regime. Again, the experimental results (solid lines) closely match the simulation results (dashed lines) discussed in Section 3.3.

3.5 Summary

In this chapter, we presented a wideband spectrum-sensing processor for weak-signal detection in the presence of strong adjacent-band primary users. We found that increasing the sensing time is not always a good way to achieve reliable detection rates when strong adjacent interferers are present. Instead, adaptive methods are needed. Our design consists of a multitap-windowed frequency-domain power detector, sensing-time, and decision-threshold adaptation algorithms.

The proposed design achieved minimum sensing time within constrained sensing rates. The adaptive algorithms, which adapt the sensing time and the detection threshold to the in-band interfering power, have demonstrated the power detector to maintain $P_{FA} = 0.1$ and $P_D = 0.9$ for -5 dB SNR and INR from 0 to 30 dB. The performance of the proposed algorithms is validated in simulation and real-time hardware radio testbed.

Simulation studies have shown that in the presence of adjacent-band interferers with $\text{INR} \geq 10$ dB, our system shows at least a 2x improvement in the P_{FA} compared to conventional methods. The proposed power detector is able to perform reliable signal detection within 50 ms for a -5 -dB SNR signal when primary users with $\text{INR} \leq 30$ dB

are present in the adjacent bands. An order-of-magnitude reduction in sensing time is achieved as compared to a conventional FFT power detector. The reduction in sensing time leads to a significant reduction in the number of operations required to reach a P_D of 0.9 and a P_{FA} of 0.1. The performance of our algorithms has been confirmed on FPGA-based real-time radio experiments in the ISM band. The obtained experimental results are in excellent agreement with the simulation estimates.

CHAPTER 4

Power and Area Minimization of Reconfigurable FFT

Processor

This chapter presents a design methodology for power and area minimization of flexible FFT processors. The proposed methodology provides fast evaluation in the power-area space that is obtained by adjusting algorithm, architecture, and circuit variables. Radix factorization is the key technique for achieving high flexibility and energy efficiency. Architecture parallelism and delay line implementation are next most significant design techniques. The reconfigurability is achieved by data-path inter-connection of atomic processing units (PUs) that explore radix-2 to radix-16 factorization.

4.1 Overview of FFT Designs

Fast Fourier transform is an important digital signal processing (DSP) technique to analyze the phase and frequency components of a time-domain signal. FFT processors have been widely used in various applications such as communications, image, and biomedical signal processing. For example, high-performance and low-power FFT processing is indispensable in orthogonal frequency-division multiplexing (OFDM) sys-

tems. Applications are now changing towards increasing diversity of features and standards that need to be supported on a single device. This change in applications greatly emphasizes the need for flexibility. At the same time, maintaining high levels of energy efficiency is of crucial importance for mobile terminals. We therefore investigate energy efficiency of *flexible* FFTs that be configured to a variety of FFT sizes and sampling rates.

FFT architectures have been extensively studied. Traditional architectures include memory-based [75], pipelined [76], array [77], and cached-memory architecture [78]. Advanced circuit techniques such as minimum-energy operation [79-80], dynamic voltage and frequency scaling (DVFS) [81], asynchronous logic [82], and deep pipelining [80] have also been used to enhance energy efficiency of FFT processors. The benefits of radix factorization for reduced hardware cost of custom FFTs have been largely unexplored. A ring-structured multiprocessor architecture was proposed in [83] to utilize mixed radix. A mixed-radix (radix 4 and radix 8) multipath delay feedback (MR-MDF) architecture and indexed-scaling pipelined architecture were introduced in [84] and [85], respectively. Prior work optimized various individual aspects of the FFT processors, which resulted in sub-optimal energy and area. A systematic design methodology that integrates parallelism, radix factorization, and memory parameters for flexible FFT design has not been thoroughly investigated.

We propose an FFT design methodology that jointly considers algorithm, architecture, and circuit parameters. We contribute with insights on how to use FFT radix structure for highly energy- and area-efficient implementations. Many combinations

of radix exist. For example, there are around a hundred architectures for a 2048-point FFT based on the degree of parallelism and radix factorization, as will be discussed in this chapter. Hardware estimates of energy and area are considered in the algorithm stage in order to choose FFT radix structure that best minimizes power/energy and area. Architecture parallelism is combined with FFT decomposition to explore power-area tradeoffs. Apart from parallelism and radix, delay buffers need to be efficiently implemented. Memory size partition and memory elements for delay lines of different lengths have to be evaluated. Our approach provides a cross-layered FFT design methodology to jointly optimize above parameters. For illustration, we will design for minimum power-area product.

4.2 FFT Algorithms and Architectures

The N -point discrete Fourier transform (DFT) of an input sequence $x[n]$ is defined as

$$X[k] = \sum_{n=0}^{N-1} x[n] W_N^{nk} \quad (4.1)$$

where $k = 0, 1, 2, \dots, N-1$ and $W_N = e^{-j2\pi/N}$ is known as the twiddle factor. Direct implementation of (4.1) requires N^2 complex multiplications and $N(N-1)$ complex additions. By taking advantage of symmetry ($W_N^{k+N/2} = -W_N^k$) and periodicity ($W_N^{k+N} = W_N^k$) of the twiddle factors, DFT can be computed more efficiently. The computationally efficient algorithms are collectively referred to as fast Fourier transform algorithms. The efficiency is further improved by using the divide-and-conquer tech-

nique that recursively breaks the N -point DFT into several smaller-size DFTs [86]. For example, for $N = M \times L$ and $k = Mp + q$, where $0 \leq p \leq L - 1$ and $0 \leq q \leq M - 1$, the N -point FFT can be represented in two-dimensional form as in Eq. (4.2).

$$X[k] = \underbrace{\sum_{l=0}^{L-1} \left(\underbrace{\sum_{m=0}^{M-1} x[l, m] W_M^{mq}}_{M-pt \text{ DFT}} W_N^{lq} \right) W_L^{lp}}_{L-pt \text{ DFT}} \quad (4.2)$$

The calculation of $X[k]$ is hence decomposed into three steps: 1) compute M -point DFT, 2) multiply the DFT outputs by twiddle factors, and 3) compute L -point DFT. Due to decomposition, the number of complex multiplications is reduced from N^2 to $N(M + L + 1)$. The number of additions is reduced from $N(N - 1)$ to $N(M + L - 2)$. To get a hardware cost for architecture exploration, individual blocks have to be examined. We start by looking at radix-2 butterfly based architecture to realize the M -point and L -point FFTs.

4.2.1 Radix-2-Butterfly Based Architecture

The radix-2 FFT algorithm decomposes the N -point DFT into 2-point DFT operations recursively. It requires $(N/2)\log_2 N$ multiplications and $N\log_2 N$ additions, leading to a significant saving for a large N compared to direct-mapped DFT. The basic arithmetic module of radix-2 FFT is the butterfly operation, as shown in Fig. 4.1(a). Decimation-in-time (DIT) and decimation-in-frequency (DIF) radix-2 butterfly operations have similar structures with different placements of the W_N multiplication. In this work, DIF structure is adopted, but the proposed design methodology is also ap-

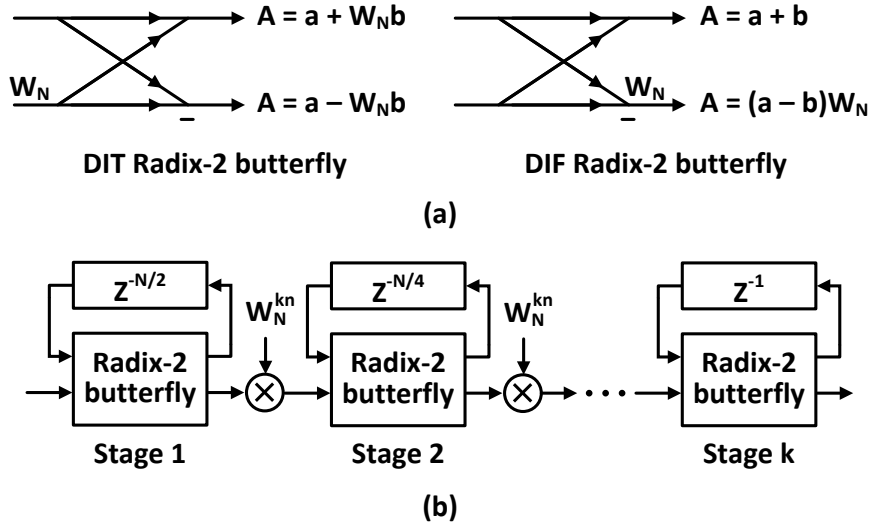


Figure 4.1: (a) Signal flow graph of radix-2 butterfly for DIT and DIF structures. (b) Radix- 2^k single-path delay feedback (SDF) architecture. Output of stage k does not need twiddle-factor multiplication.

plicable for DIT. N -point DFT can also be decomposed into r -point DFTs. This is known as radix- r FFT algorithm. There are $\log_r N$ stages and N/r butterflies per stage. Each butterfly requires $r - 1$ complex multiplications. When $r = 2^k$, the radix- r butterfly can be further decomposed by cascading k radix-2 stages, known as radix- 2^k algorithm [87]. The total number of complex multiplications of radix- r algorithm is $[N(r - 1)/r][(\log_r N) - 1]$ considering that the twiddle factors of the last stage are always equal to unity. As we will describe later, all multipliers internal to the radix- r butterfly ($r = 2^k$) can be implemented as constant multipliers in order to reduce area and power. The number of complex additions is $N \log_2 N$ regardless of r .

Due to its regular structure, the FFT can be realized using radix- r ($r = 2^k$) butterflies, delay lines, and complex multipliers. Radix- r can be realized using several archi-

tectures considering a tradeoff between silicon area and execution time. The memory-based time-multiplexed architecture [75] has only one radix- r butterfly and $r - 1$ complex multipliers. The inputs, outputs, and interim results are read from and written back to memory for complete FFT operation. This architecture has the lowest area and longest execution time. Another extreme is the direct-mapped fully parallel architecture, which requires $[N(r - 1)/r][(\log_r N) - 1]$ complex multiplications. Between these two extreme cases, there is a pipelined architecture which provides a balanced tradeoff between area and execution time. Two major types of pipelined architectures are multi-path delay commutator (MDC) and single-path delay feedback (SDF) [87]. For higher-radix algorithms, the SDF architecture is preferred since it requires less memory and fewer complex multipliers than the MDC architecture [87].

The radix-2 SDF architecture for 2^k -point FFT is shown in Fig. 4.1(b). In each stage, the required $N/2$ butterfly operations are time-multiplexed onto one butterfly operator. Delay buffers are used to support the time-multiplexing. The inter-stage multipliers are used to multiply stage outputs by twiddle factors W_N^{kn} . The radix-2 SDF architecture requires $\log_2 N - 1$ inter-stage complex multipliers, $2\log_2 N$ complex adders, and $N - 1$ delay buffers.

4.2.2 Reconfigurable Architecture

Based on the pipelined SDF architecture, a reconfigurable FFT architecture can be implemented by cascading several radix- 2^k stages in order to accommodate different FFT sizes. The signal-flow graphs for radix-2 to radix- 2^4 butterflies are shown in Fig.

Table 4.1: Radix Implementations of Radix-2, Radix-2², Radix-2³, and Radix-2⁴ Processing Elements

Type	Implementation	Butterfly adders	Constant multipliers
Radix-2	PU1	4	0
Radix-2 ²	PU2 + PU1	8	1
Radix-2 ³	PU2 + PU3 + PU1	12	13
Radix-2 ⁴	PU2 + PU3 + PU4 + PU1	16	43

4.2. Note that the minus signs in the butterfly modules are ignored for simplicity. The highlighted inter-stage multiplications are all implemented as constant (complex) multipliers as opposed to full multipliers. The radix-2^k can be realized by cascading several atomic processing units (PUs) as in Table 4.1. The PUs are shown in Fig. 4.3. The cost of the constant multipliers in terms of equivalent adders is shown in brackets. Radices beyond 2⁴ are impractical, because the increasing number and complexity of required constant multipliers makes them no longer advantageous over full multipliers. In addition, full multipliers need extra ROMs to store coefficients, but dedicated constant multipliers are all hard-wired and their coefficients can be locally computed. Therefore, radix-2 to radix-2⁴ is the right level of granularity for mixed-radix FFT implementation.

As shown in Fig. 4.3(a), each PU contains a basic butterfly module and dedicated constant multipliers. The butterfly module is initialized to the data-switch mode until

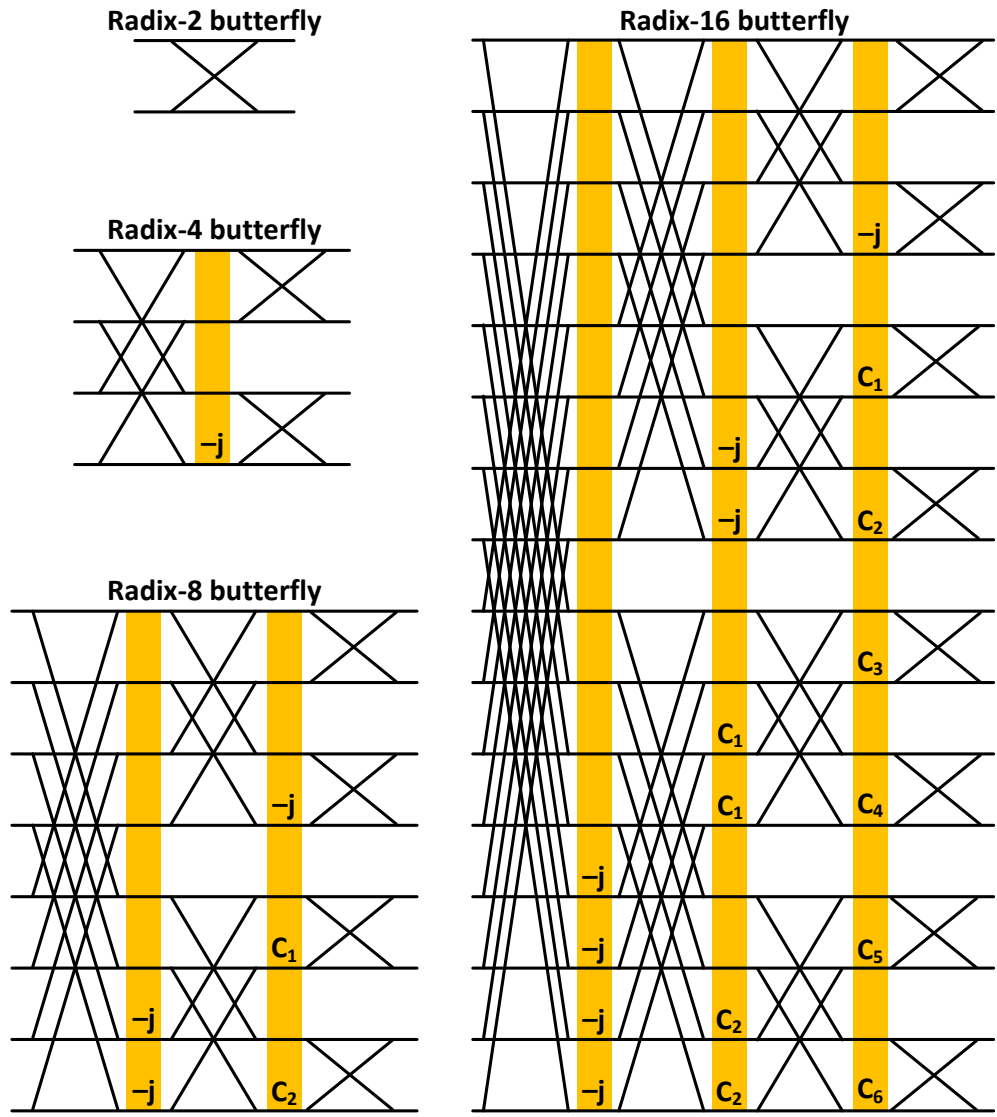


Figure 4.2: Radix-2, radix- 2^2 , radix- 2^3 , and radix- 2^4 butterfly operations.

the delay buffers are filled by the valid inputs and then switched to the butterfly mode for FFT operation. The required constant multipliers for PUs 1-4 are shown in Fig. 4.3(b). Only half the twiddle factors (dark-filled circles) are generated in the PUs, the other half (gray-filled circles) are created using the symmetry property. PUs with lower index can be deduced from the PUs with higher index. This *back-compatibility* feature will be leveraged for reconfigurable designs. All the multipliers inside the PUs for a 2^k -point FFT ($k \leq 4$) are constant multipliers. Full multipliers are used for the multiplications for the inter-stage twiddle factors. Since the inter-stage full multipliers have higher hardware cost than the intra-stage constant multipliers, radix factorization should be done in such a way as to minimize the number of full multipliers.

Radix factorization was utilized in prior work with a limited success. A variable-length FFT processor that integrates two radix-2 stages and three radix- 2^3 stages for FFT sizes 512, 1024 and 2048 was proposed in [88]. A single architecture is considered without investigating other possible factorizations. We consider all feasible solutions to select the final architecture. Also, the use of high-radix is commonly believed to be more area-efficient than low-radix algorithms [81] [84]. While true for generic hardware, this observation is inaccurate for dedicated designs that consider the use of constant multiplications and variable wordlengths of twiddle operators. The complexity analysis only considering the asymptotic trend $O(N \log_r N)$ may lead to an inefficient architecture for dedicated designs.

Apart from radix factorization, architecture- and circuit-level parameters have to be considered. Although many point-wise solutions have been proposed at the architecture

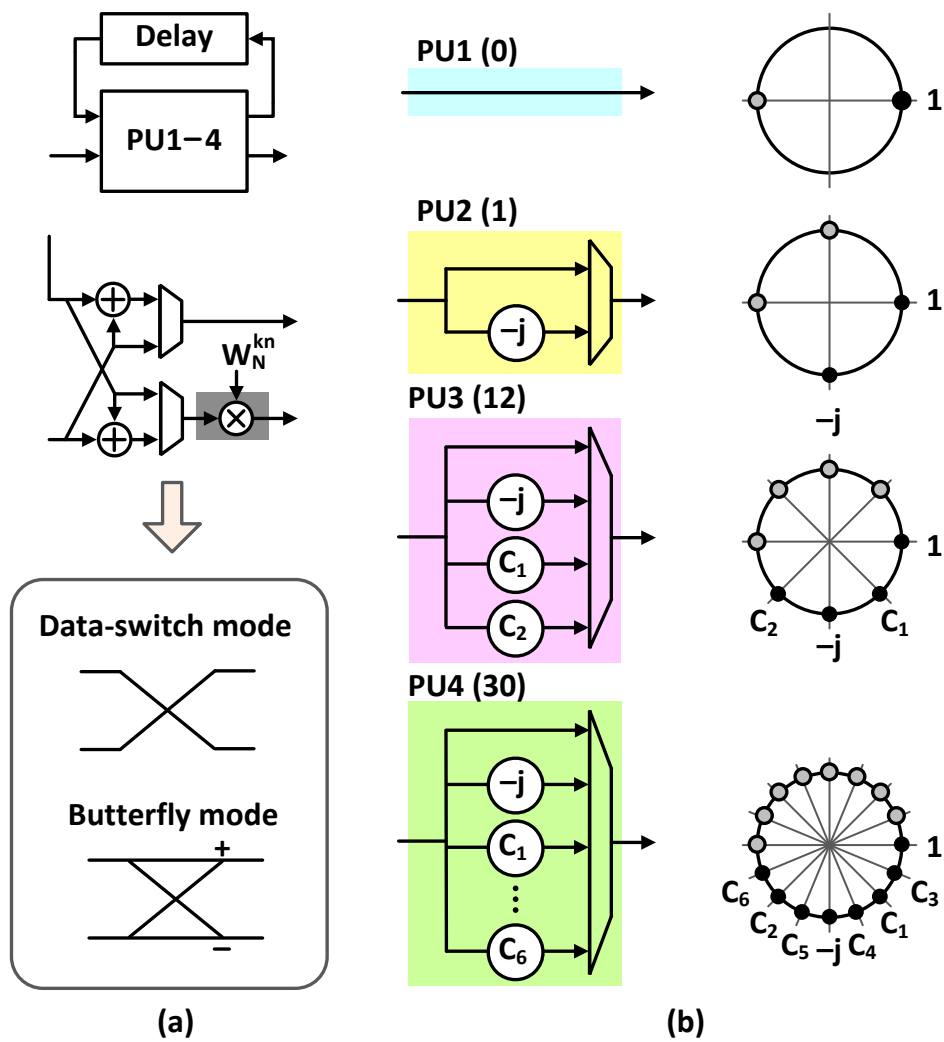


Figure 4.3: Processing units (PUs) of a reconfigurable FFT processor that supports radix-2 to radix- 2^4 factorizations. (a) The architecture and (b) the constant multipliers of the PUs. The numbers in brackets next to each PU indicate relative cost (in terms of equivalent adders) of its constant multiplier.

level, the tradeoff between sequential and parallel architectures has not been thoroughly investigated. At the circuit level, memory cells and logic elements have to be carefully chosen to minimize the power and area. Our methodology for power and area minimization jointly considers radix factorization, architecture and circuit techniques.

4.3 Power and Area Minimization

We propose a systematic methodology to explore feasible FFT realizations and select those that best minimize power and area. Different realizations will be made using architecture parallelism and FFT decomposition, radix factorization, and memory implementation. Given many feasible solutions, the optimal design can be selected from the power-area tradeoff. To evaluate feasible realizations early in the algorithm development stage, high-level area and power estimation are developed. Area is estimated as the total area of multipliers, adders, and memory. Power is estimated from the area, switching activity, operating frequency and voltage.

FFT realizations are systematically explored in three steps. First, architecture parallelism combined with FFT decomposition is used to explore the power-area space through voltage scaling. Next, radix factorization is explored for a given FFT size. The third step consists of memory partition, selection of memory cells and logic operators. To support multiple FFT sizes, optimal mapping of processing units that considers all required FFT sizes is performed.

4.3.1 Parallel Architecture with FFT Decomposition

Parallelism is an effective technique to increase throughput [81] or to reduce power consumption [89-90] of an FFT processor. For a fixed throughput, voltage scaling and a lower clock frequency improve the energy efficiency of a parallel architecture. Since time-multiplexing is inherently applied to SDF architecture, parallelism is used to move the design point in the area-energy-delay plane [70]. The arithmetic property of FFT allows for an area-efficient parallel architecture by leveraging FFT decomposition. As shown in Fig. 4.4(a), an N -point FFT is decomposed into M -point FFT and L -point FFT. The direct-mapped architecture by the level of parallelism P is described in Fig. 4.4(b). The resulting area of logic and memory is increased by P (ignoring the overhead of the serial-to-parallel and parallel-to-serial circuitry). For the case where $P = L$, the P identical L -point SDF FFT can be combined into a single L -point parallel FFT, as in Fig. 4.4(c). This architecture simplification is possible since the M -point FFTs can be computed first and combined into the L -point ($L = N/M$) output stage to compute an N -point FFT. Instead of distributing multiple data streams into parallel branches as in Fig. 4.4(b), one data stream is processed by a single L -point FFT unit as in Fig. 4.4(c).

The architecture in Fig. 4.4(c) contains $L[(\log_2 M) - 1] + L + L[(\log_2 L) - 1]/2 = L[(\log_2 NM) - 1]/2$ complex multipliers and $L(2\log_2 M) + L\log_2 L = L(\log_2 NM)$ complex adders. Besides reduced arithmetic complexity compared to Fig. 4.4(b), effective implementation of memory is required. $L - 1$ delay buffers used for scheduling of the L -point FFT in Fig. 4.4(b) can be removed since the inputs of L paths are available

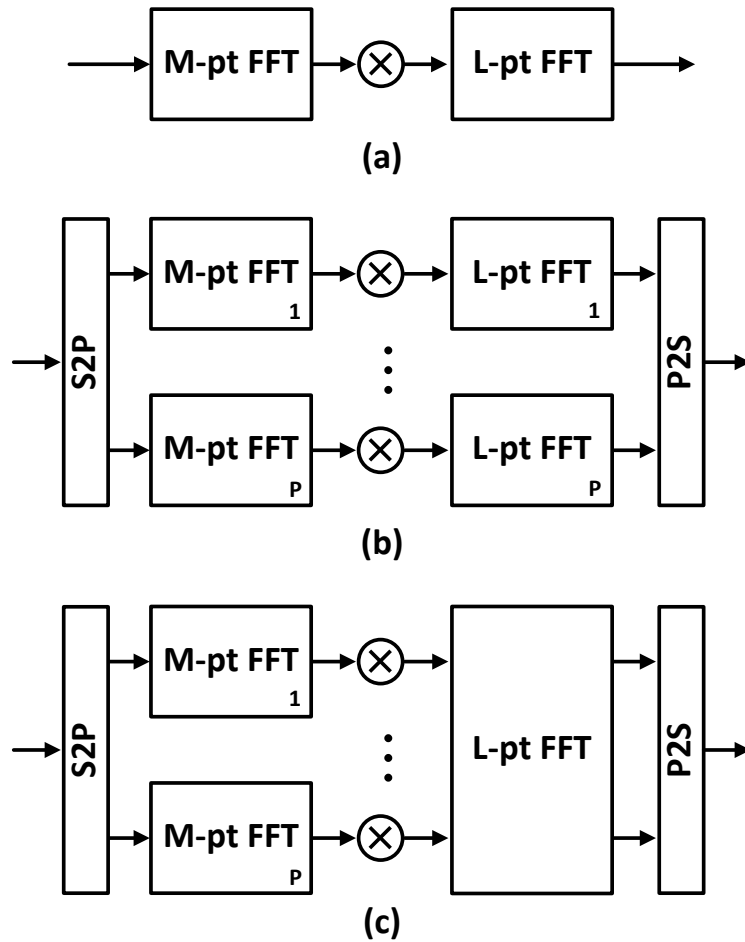


Figure 4.4: (a) Reference N -point ($N = M \times L$) FFT architecture. (b) Parallel architecture. The level of parallelism = P . (c) A significant reduction in complexity is achieved when $P = L$.

and aligned in time in Fig. 4.4(c). The total number of delay buffers is reduced from $L(N - 1)$ to $L(N - 1) - L(L - 1) = L(N - L)$.

Finally, the hardware complexity of the M -point FFT can be reduced. Combining L streams in Fig. 4.4(b) necessitates zero-padding of data and the length of delay buffers in the M -point FFTs to be a multiple of L . When $L = P$, as in Fig. 4.4(c), the length- L delay buffers in each of the M -point FFTs can be replaced by length-1 buffers at L times lower rate to match the delay. This reduces the number of delay buffers in Fig. 4.4(c) from $L(N - L)$ to $L(N - L)/L = N - L = L(M - 1)$, which is the minimal number of delay buffers required for an L -path M -point FFT.

4.3.2 Estimation of Area and Power

To evaluate the area and power for different levels of parallelism, high-level area and power models are required for the FFT building blocks. The area of the L -path SDF architecture in Fig. 4.4(c) is estimated as

$$Area = A_{mult} \cdot L[(\log_2 NM) - 1]/2 + A_{add} \cdot L(\log_2 NM) + A_{mem} \cdot L(M - 1), \quad (4.3)$$

where A_{mult} , A_{add} , and A_{mem} represent the area of multiplier, adder, and delay buffer, respectively. These three parameters can be estimated from synthesis. Without loss of generality, we use 12-bit arithmetic and DFF-based delay buffers in our analysis.

Delay of a FO4 inverter is used to estimate the critical-path delay as a function of supply voltage. Figure 4.5 shows the delay vs. supply voltage curve in the typical-typical (TT) corner in a standard- V_T TSMC 65nm CMOS. The delay-voltage curve is

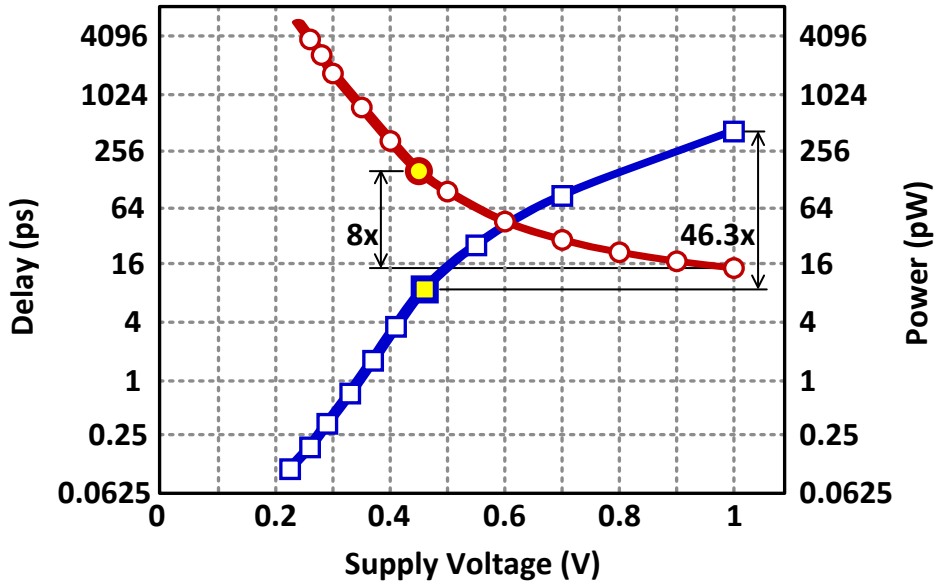


Figure 4.5: Delay and energy vs. supply voltage for a FO-4 inverter in TSMC 65nm CMOS.

used to predict the amount of voltage scaling, analyze high-level architecture retiming and fine-grain circuit pipelining. FO4 inverters are also used for power estimation. The use of Fig. 4.5 for low-voltage design can be exemplified on a design required to run at 20 MHz. Since digital libraries are characterized for nominal voltage, designing for 20 MHz at 0.45 V in the TT corner translates to 160 MHz at 1 V for logic synthesis. By reducing V_{DD} from 1 V to 0.45 V and frequency from 160 MHz to 20 MHz, the power consumption is reduced by 46.3x. The normalized power of the FFT can be therefore estimated by

$$Power = Area \cdot P(V_{DD}, f), \quad (4.4)$$

where f is the switching frequency.

Power cost is estimated by considering both switching (P_{sw}) and leakage (P_{leak})

components. A ratio $P_{sw}/P_{leak} = 13.5$ is estimated from a FO4 inverter chain at 0.45 V for 20 MHz (around 135x FO4 delay). The switching power P_{sw} of the processing units can be estimated as the product of utilization rate α and active area A_{total} per cycle since circuits are operated at the same voltage and frequency, as indicated by

$$P_{sw} = CV^2 f \propto \alpha A_{total}, \quad (4.5)$$

where A_{total} is the area cost. Inactive blocks are disabled by using clock gating or wired-AND circuits. As shown in Fig. 4.2, the two complex adders (implemented as four real adders) in the butterfly module of PUs have utilization rate of 1/2 because the periods for data-switch mode and butterfly mode are the same, resulting in 2 real additions per cycle, on average. According to the signal-flow graph in Fig. 4.1, the utilization rate of each constant multiplier is 1/4, 1/8, and 1/16 per clock cycle for PU2, PU3, and PU4, respectively. The utilization rate of the inter-stage full-precision multipliers depends on the preceding PUs and it is accounted for in our analysis.

4.3.3 Mixed-Radix Implementation

As mentioned in Section 4.2.1, higher-radix structures could be made more area efficient by replacing full multipliers with constant multipliers. These constant multipliers are implemented using canonic signed digit (CSD) representation [91], as shown in Fig. 4.6. A 10-bit precision is assumed. Each of the constant factors requires no more than 4 simple additions, which leads to a large area reduction. Due to the symmetry property of twiddle factors, the area of constant multipliers is minimized through hard-

Factor	CSD realization	Adders
0.7071	$2^{-1} + 2^{-3} + 2^{-4} + 2^{-6} + 2^{-8}$	4
0.9238	$2^0 + 2^{-3} + 2^{-5} + 2^{-6} + 2^{-9}$	4
0.3838	$2^{-2} + 2^{-3} + 2^{-7}$	2

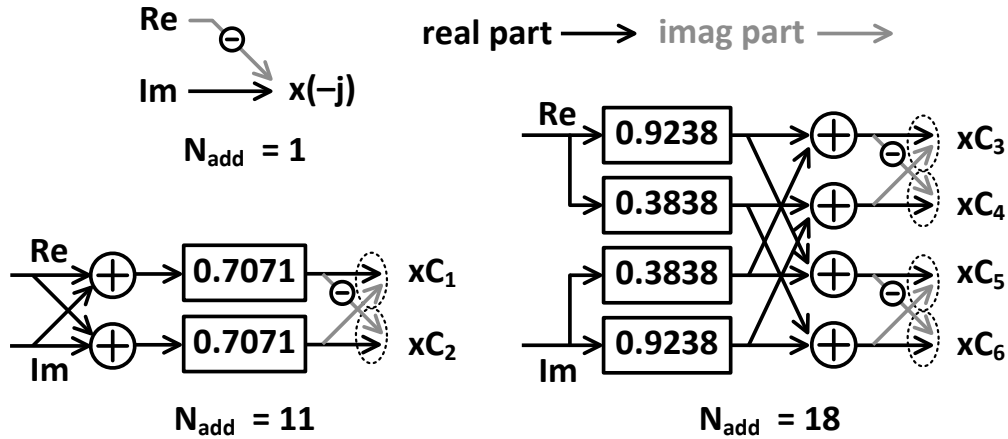


Figure 4.6: Area-efficient implementation of constant multipliers.

ware sharing for multiple multiplicands. Only 30 adders are need for all twiddle-factor multiplications as shown in Fig. 4.2. The cost of implementing radix-2 to radix-2⁴ operations, which will be used in radix factorizations, is summarized in Table 4.1.

To minimize the hardware cost of inter-stage full multipliers (between radix-2^k blocks), full complex multipliers are implemented by using 3 real multipliers and 5 real adders (rather than 4 real multipliers and 2 adders) [92] as follows:

$$(a + jb)(c + jd) = [c(a - b) + b(c - d)] + j[d(a + b) + d(c - d)]. \quad (4.6)$$

The equivalent number of adders in a full real multiplier is estimated as the wordlength of twiddle factors. Since the adders for butterfly operations exist in all radix structures, only the equivalent adders for (constant vs. full) complex multiplications can be used

for quick comparison of different FFT architectures. The number of equivalent adders for 4- to 16-point FFTs with twiddle factor wordlengths 8b-16b is summarized in Table 4.2. As an example, a 16-point FFT can be implemented by two radix- 2^2 stages, each having 1 adder (Table 4.1), along with an inter-stage full multiplier (with 3 multipliers and 5 adders). For 8-bit twiddle factors, the number of equivalent real adders is therefore $2 + (38 + 5) = 31$. FFTs up to 16 points are considered, because this is adequate level of granularity for radix factorization. The use of radix depends on the FFT size and twiddle-factor wordlength. Radix structures with minimum area are highlighted. Radix- 2^2 architecture needs only one adder for 4-point FFT, but the radix-2 architecture needs 29-53 adders due to the use of full multipliers. Using higher radix makes sense here, because radix- 2^2 exactly matches 4 FFT points. As the FFT size increases, a larger number of constant multipliers required to support higher radix diminishes the area advantage over full multipliers. For 16 points, the radix- 2^4 architecture needs 37-49 adders while the radix-2 architecture needs 87-159 adders (a 2.53-3.24x larger area). The area saving for the 16-point FFT is not as large as that for the 4-point FFT. The wordlength of twiddle factors also affects the area. Radix- 2^2 architecture is the most area efficient for twiddle factors below 14 bits, else radix- 2^4 should be used. Therefore, radix-2 is the least area efficient. Higher radix (up to 2^4) is generally better unless wordlength of twiddle factors is small (less than 14 bits).

Based on the area and power models from Section 4.3.2, we can quickly evaluate area-power tradeoff for various radix factorizations and degrees of parallelism. As an example, Fig. 4.7 shows the area-power tradeoff space for 1024-point FFT imple-

Table 4.2: The Number of Equivalent Adders Required to Implement FFT

FFT size	4 points					8 points				
	8	10	12	14	16	8	10	12	14	16
TW bits	8	10	12	14	16	8	10	12	14	16
Radix-2	29	35	41	47	53	58	70	82	94	106
Radix-2 ²	1	1	1	1	1	-	-	-	-	-
Radix-2 ³	-	-	-	-	-	11	13	13	13	15
Radix-2 ⁴	-	-	-	-	-	-	-	-	-	-

FFT size	16 points				
TW bits	8	10	12	14	16
Radix-2	87	105	123	141	159
Radix-2 ²	31	37	43	49	55
Radix-2 ³	-	-	-	-	-
Radix-2 ⁴	37	43	43	43	49

mented using architecture from Fig. 4.4(c). Partitioning with $M = 128$ and $L = P = 8$ gives the lowest power-area product (PAP) for 1024 points. It achieves a 4.3x lower PAP than $M = 1024$, $L = 1$ design. Next, since $M > 2^4$, we have to examine radix factorization of 128-point FFT for further area and power reduction. Radix structure of 128-point FFT is shown in Fig. 4.8. The architecture with seven radix-2 stages (A1) occupies the largest area, as expected from prior analysis. Architecture A8 (consisting of two radix-2² stages and one radix-2³ stage) has the lowest PAP. A 4x reduction is achieved from radix-2 only (A1) to radix-2²/2²/2³ (A8) due to radix factorization. An

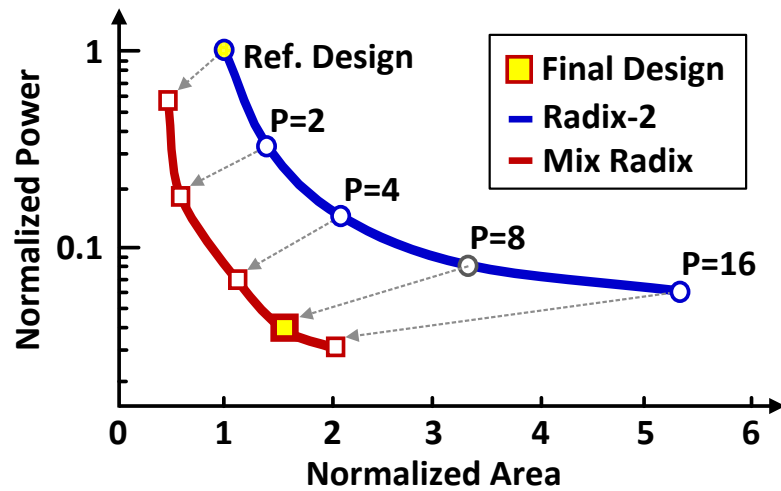


Figure 4.7: Power and area minimization through FFT factorization and radix factorization.

overall 17.2x PAP reduction compared to the direct mapped architecture ($M = 1024$, $L = 1$) is achieved through FFT decomposition (4.3x) and radix factorization (4x).

4.3.4 Delay-Buffer Architecture and Memory Element

After FFT decomposition and radix factorization, the next step is the power-area optimization of delay buffers. Three options are considered for delay buffer implementation: 1) DFF-based shift register, 2) register file (RF)-based, and 3) SRAM-based delay buffer [87]. Architecture and memory cells for various delay lengths have to be evaluated. To illustrate our methodology, we assume delay lengths up to 1024.

We start by comparing RF-based and SRAM-based delay buffers, which have different memory cells and peripheral circuits. A straightforward implementation is a dual-port (DP) RF/SRAM-based architecture, as shown in Fig. 4.9(a). If the RF/SRAM

Arch	R2	R4	R8	R16
A1	7			
A2	5	1		
A3	3	2		
A4	1	3		
A5	4		1	
A6	2	1	1	
A7	1		2	
A8		2	1	
A9	3			1
A10	1	1		1
A11			1	1

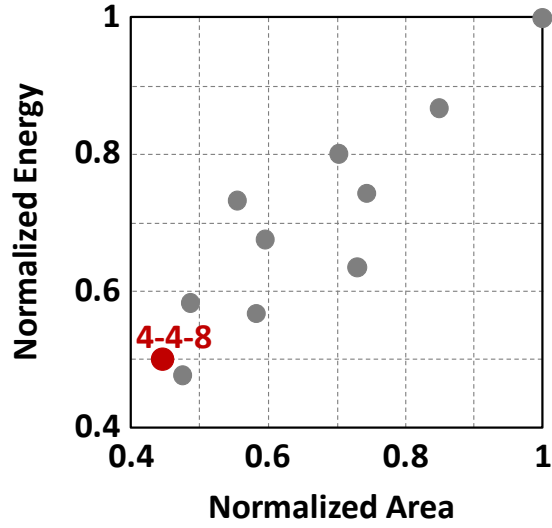


Figure 4.8: Power and area for feasible radix factorizations of 128-point FFT. Architecture A10 has the lowest power-area product.

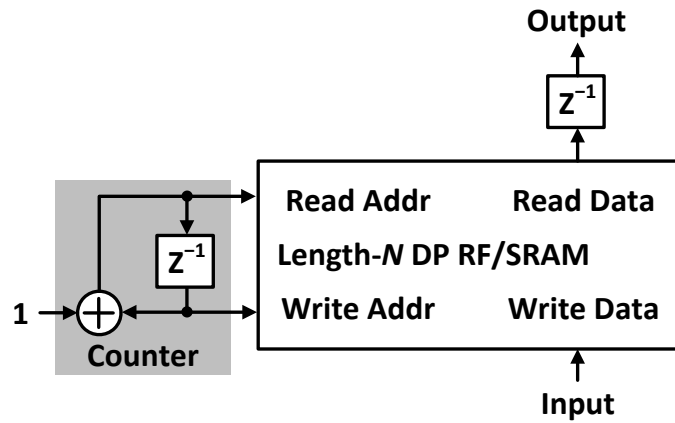
memory size is N , the output of the dual-port RF/SRAM is read after $N - 1$ cycles, so an overall N -cycle delay is achieved. Area and power of RF and SRAM designs are evaluated using commercial memory compilers for the target 65-nm technology. For 32-bit (for complex-valued input) delay buffers, RF-based design is superior since it consumes 41-49% power with a 66-82% silicon area compared to the SRAM-based counterpart.

Next, we compare RF-based and DFF-based designs. RF-based delay buffers are more area efficient than DFF-based designs due to the compact 6T/bitcell structure but the peripheral circuitry contributes considerable area overhead for smaller memory size. However, RF-based designs cannot be operated at very low voltage due to cell read margin and sense amplifier operation. Since DFF-based designs can operate in a low-voltage regime, they are more energy efficient despite their area disadvantage.

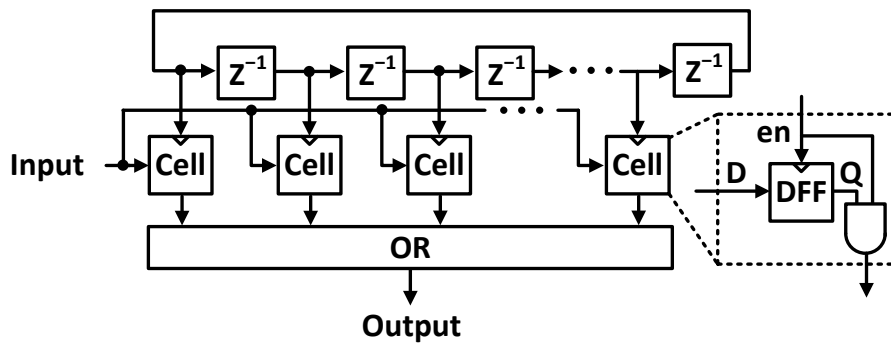
Besides voltage scaling, the power consumption of the DFF-based delay buffers can be reduced through a pointer-based architecture shown in Fig. 4.9(b). Instead of shifting the data, we use a shift-delay-line to choose the corresponding DFF to read and write. The remaining DFFs are clock gated when they are not activated by the shift-delay-line, eliminating significant dynamic power. To evaluate the tradeoff between RF-based and DFF-based delay buffers, power and area for the delay line lengths of interest are shown in Fig. 4.10. For the delay buffers longer than 256, the RF-based design operated at 0.9 V has a smaller PAP compared to the DFF-based design operated at 0.4 V at 20 MHz. Therefore, the delay buffer of length 512 and 1024 are implemented using RFs.

The power of RF can be further improved through memory partitioning. Fig. 4.11 shows the possible memory partition schemes and the optimal partition for lengths 512 and 1024. We minimize PAP, but the architecture with lower power is chosen if multiple instances have the same PAP. The instances with two and four 256x32b RF banks are chosen for the delay buffers of length 512 and 1024, respectively.

Finally, one length-256 dual-port (DP) RF module can be replaced with two length-128 single-port (SP) RF modules to achieve even lower PAP: 0.63x and 0.58x lower PAP for lengths 512 and 1024, respectively. The final memory structure for delay buffers of lengths 512 and 1024 is shown in Fig. 4.12. In the final design, eight 128x32b and four 128x32b register files are used to implement length 1024 and 512 delay buffers. Input/output of two SP RF modules are written/read alternatively to create adequate delays. For the remaining delay buffers, DFF-based registers with aggressive voltage scaling are used instead.



(a)



(b)

Figure 4.9: (a) Dual-port RF-based delay buffer and (b) pointer-based DFF-based delay buffer.

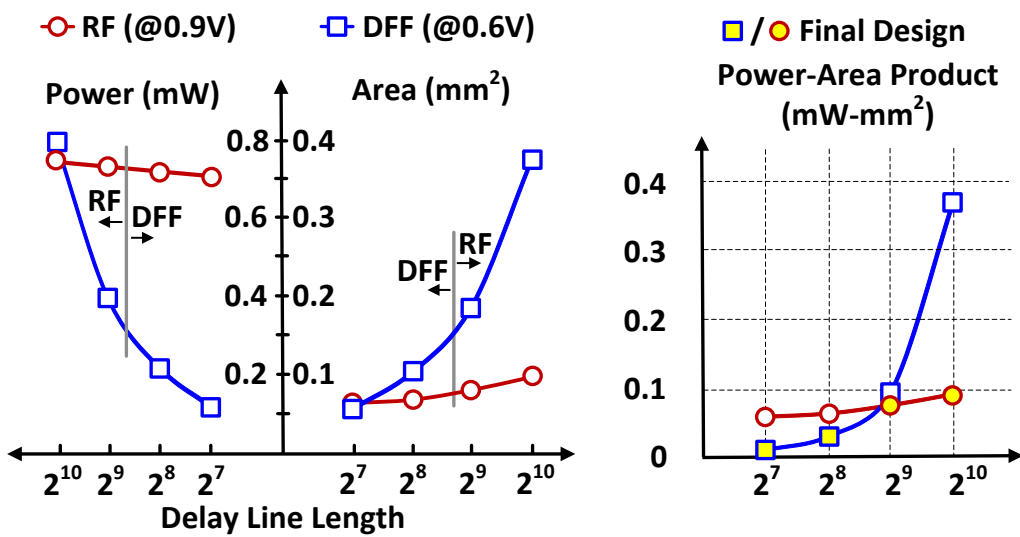


Figure 4.10: Power-area product of delay buffers: D flip-flops (DFFs) are used for delay lengths of 128 and 256, register files (RF) are used for delay lengths of 512 and 1024.

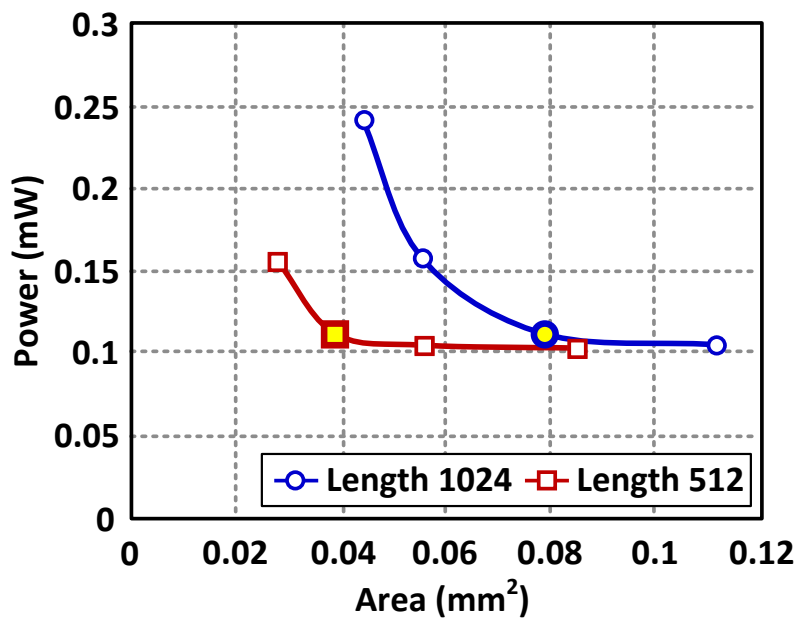


Figure 4.11: Power-area tradeoff for feasible memory partitions. Partitions 2×256 for length 512 and 4×256 for length 1024 delay buffers yield minimum power-area product.

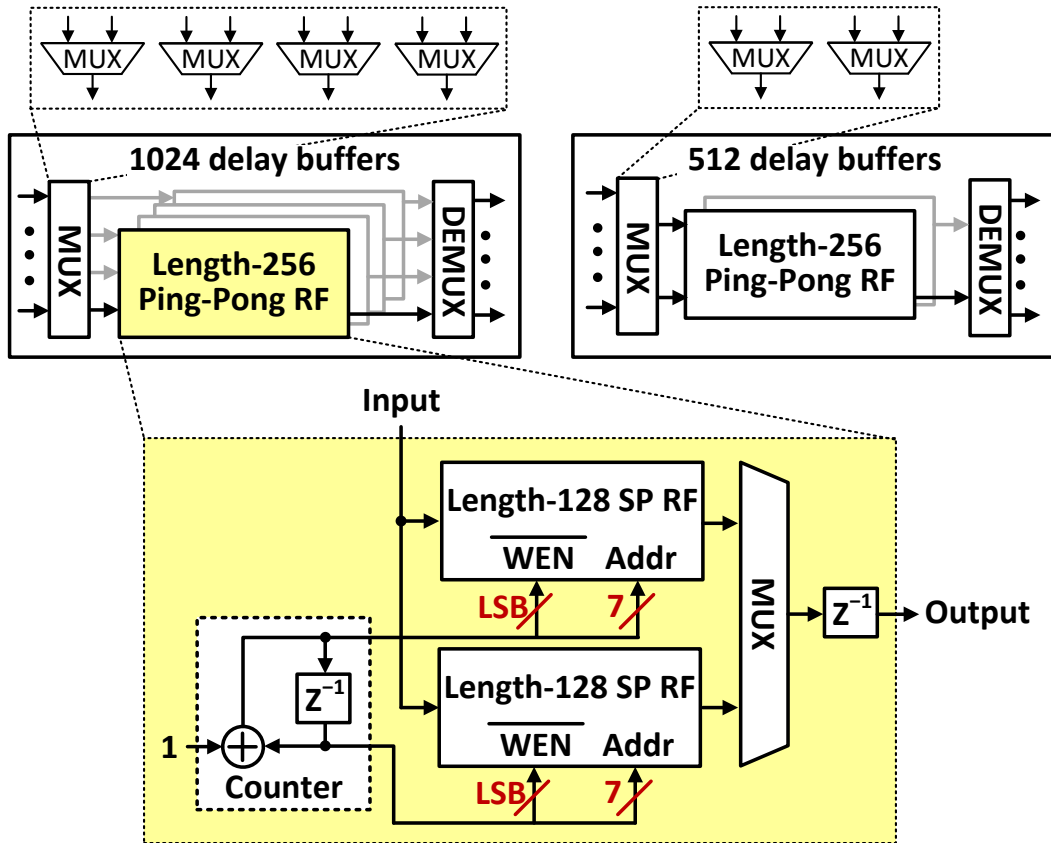


Figure 4.12: Memory architecture for length 512 and 1024 delay buffers.

4.3.5 Area-Efficient Twiddle-Factor Generator

Twiddle factors are generated in an area-efficient way by trigonometric approximation [87] instead of fetching coefficients from ROMs. The trigonometric approximation is realized by a first-order linear approximation [93], which can be described as follows. First, by means of the symmetry of sine/cosine values, angles $[0, \pi/4)$ can be used to construct the whole sine/cosine space. Second, sine values can be approximated through piecewise-linear approximation, as given by

$$\sin(2\pi\alpha) \approx \begin{cases} 25/4\alpha & 0 \leq \alpha < 1/16 \\ 41\alpha/8 + 9/128 & 1/16 \leq \alpha \leq 1/8. \end{cases} \quad (4.7)$$

Third, $\cos(2\pi\alpha) = \sqrt{1 - \sin^2(2\pi\alpha)}$ for $0 \leq \sin(2\pi\alpha) \leq 0.7071$ is approximated by another linear approximation as

$$\cos(2\pi\alpha) \approx \begin{cases} 1 - \sin(2\pi\alpha)/16 & 0 \leq \sin(2\pi\alpha) < 1/8 \\ 65/64 - 3\sin(2\pi\alpha)/16 & 1/8 \leq \sin(2\pi\alpha) < 1/4 \\ 17/16 - 3\sin(2\pi\alpha)/8 & 1/4 \leq \sin(2\pi\alpha) < 1/2 \\ 5/4 - 3\sin(2\pi\alpha)/4 & \text{otherwise.} \end{cases} \quad (4.8)$$

Based on these approximations, only shift and add operations are needed to generate twiddle factors (TFs). The overall number of intra-stage TF generators is $(L - 1) + (F - 1)$, where F is the number of stages of PUs in the M -point FFT block. Since TFs for the multipliers in the same stage of the L M -point FFT blocks can be shared, only $F - 1$ unique TF generators are necessary. In addition, there are $L - 1$ TF generators for the multipliers between M -point and L -point FFTs.

4.4 Summary

An FFT processor design methodology yielding optimal power-area tradeoff is explored by examining feasible parallel architectures and radix factorizations. The use of constant multipliers for intra-stage twiddle factors enables substantial area and power savings compared to the use of full multipliers. Radix factors up to 16 should be used. Radices beyond 16 are ineffective due to a large number of constant multipliers required. Short delay line buffers (up to 256) are best realized in D-flip-flops, medium buffers (length 512 and 1024) are the most energy and area efficient when realized with register files. Twiddle factors are generated using trigonometric approximations as opposed to ROM memories. More detailed design example will be illustrated in Chapters 5 and 6.

CHAPTER 5

Chip 1: 200MS/s Wideband Spectrum Sensing

Processor

This chapter presents our first chip: a practical solution to the wideband sensing problem with adjacent-band interferers. An FFT-based multitap windowing power detector architecture is used for wideband (200 MHz) sensing; sensing time and detection threshold are dynamically adjusted to remove the impact of adjacent-band interferers [94]. The chip power and area are minimized by jointly considering algorithm, architecture, and circuit parameters. The chip, implemented in a 65-nm CMOS technology, outperforms state-of-the-art [53-56] in detection performance and power dissipation per bandwidth. Its high performance allows fast, reliable spectrum sensing for future cognitive radios. The low power dissipation makes it applicable to portable devices.

5.1 Design Challenges of Wideband Spectrum Sensing

Figure 5.1 shows the block diagram of the proposed spectrum-sensing processor. It consists of a multitap-windowed frequency power detector and two sensing adaptation

blocks, sensing-time adaptation (STA) and detection-threshold adaptation (DTA). As shown in Fig. 5.1, key design challenges for power- and area-efficient implementation are high-speed FFT-based power measurement and large dynamic range datapaths. Two adaptation blocks, STA and DTA, dynamically track channel conditions to maintain the specification, but this adaptation adds extra processing-time latency. The latency overhead should be minimized to increase sensing-time efficiency.

The FFT is the most hardware-intensive block in our architecture. Its throughput and size dictate the sensing time and frequency resolution. To support real-time sensing of 200 MHz, the FFT processor requires a throughput of 200 MS/s. In our application, 1024 points are required to sense 200 MHz with 200 kHz resolution. Thus, a power-area efficient FFT processor is required.

The large dynamic range, required for precise and simultaneous detection of both strong interferers and weak primary users, dictates large wordlengths, which increase the cost of arithmetic operations and storage of the estimated PSD, noise power, and interferer power. Since the increased wordlength is used to represent the large dynamic range signal rather than enhance the precision, fixed-point arithmetic is inefficient. Instead, floating-point arithmetic will be used.

The adaptation blocks, STA and DTA, dynamically track channel condition to maintain system performance. As shown in Fig. 5.2, the shorter the processing time of PSD estimation, STA, and DTA, the longer the time the system can utilize available spectrum. The processing time of PSD is the sensing time needed to meet the sensing rates (P_D and P_{FA}) and this processing time is fixed. The processing time of STA and DTA

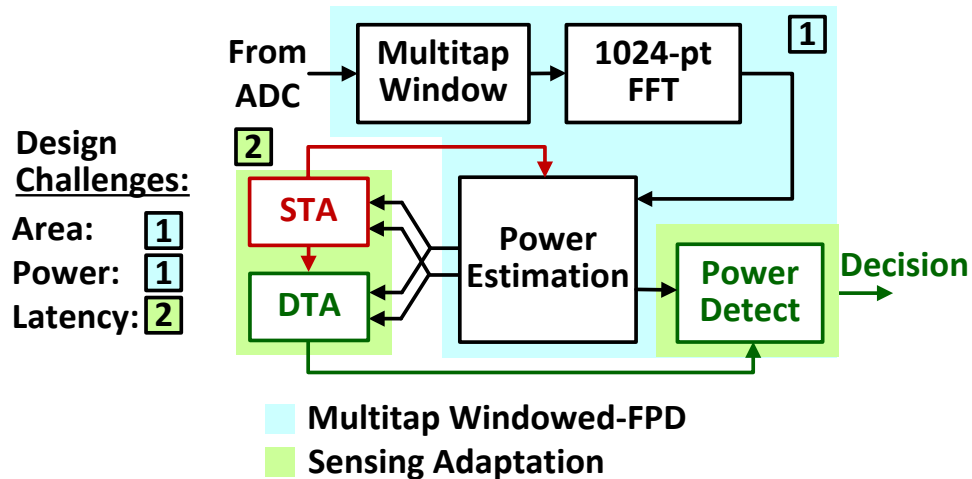


Figure 5.1: Proposed wideband spectrum sensing processor. An FFT-based multitap windowed frequency power detector estimates PSD, noise and interfering power. Sensing-time adaptation (STA) and detection-threshold adaptation (DTA) blocks adapt their parameters on a per-channel basis.

blocks doesn't influence the sensing rates. These blocks affect the sensing time; hence, their processing time overhead has to be minimized.

5.2 Architecture and Circuit Design

Several architectural and signal processing techniques are proposed to minimize power and area. The FFT processor uses parallelism and radix factorization to achieve minimum power-area product (PAP). A fixed-to-floating point converter is designed to minimize wordlengths of the mantissa and exponent, which helps reduce memory size. The power and latency of STA and DTA blocks are reduced by utilizing the Newton-Raphson division and square-root algorithms.

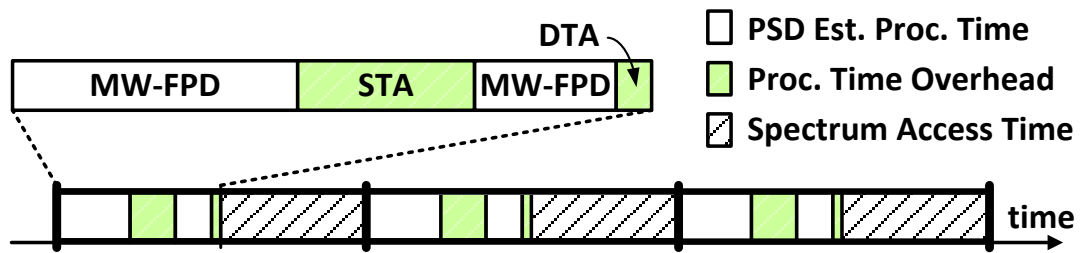


Figure 5.2: Processing time for spectrum sensing consists of fixed time for MW-FPD calculation and variable time for STA and DTA blocks. The MW-FPD determines the sensing rate performance. The overhead introduced by the STA and DTA blocks degrades the sensing-time efficiency and has to be minimized.

5.2.1 Multi-path Pipelined MW-FPD

A multi-path pipelined FFT is adopted to achieve high throughput and low power. The power efficiency of the pipelined architecture is improved through architecture parallelism, which allows for voltage scaling due to a lower operating frequency. Also, a radix-factorization methodology [95] is used to determine the optimal radix structure. Fig. 5.3 shows the architecture of the 1024-point FFT, which is decomposed into eight banks of 128-point pipelined FFTs followed by an 8-point parallel FFT to minimize power-area product.

The 1024-point FFT is designed by using the methodology as described in Chapter 4. The amount of parallelism is determined by choosing the minimum power-area product point. The 8-path single-delay-feedback (SDF) architecture operated at 0.6 V is chosen as the optimal design, resulting in a 4.3x PAP reduction compared to the reference design with no parallelism. The optimal radix factorization is selected from

all feasible combinations of radix-2/2²/2³/2⁴ processing elements (PEs). A 4x PAP reduction is achieved from radix-2 only to the final design radix-2²/2²/2³. An overall 17.2x PAP reduction is achieved through FFT decomposition (parallelism) and radix factorization.

The delay lines are designed for minimum PAP by exploiting the power-area trade-off between register files (RFs) and D flip-flops (DFFs) [96]. RFs (6T/bitcell) occupy less area compared to DFFs while DFFs allow lower V_{DD} for power reduction. Delay lines of lengths 512 and 1024 are implemented using register files (RFs) that operate at 0.9 V; else DFFs operating at 0.6 V are used. Level-shifters are placed between low- V_{DD} (0.6 V) and high- V_{DD} (0.9 V) domains. Such partitioning reduces the core area by 0.4 mm² (20%) with power increased by 0.3 mW (4%) as compared to a DFF-only implementation.

The energy and area of the 200 MS/s 1024-point FFT measured at 0.7 V are 20 nJ/FFT and 0.75 mm², respectively. Our chip measurements at 200 MS/s were consistent with synthesis estimates. Next, we re-synthesize for 240 MS/s and compare our mix-radix design with the 240 MS/s 1024-point radix-4 FFT [80] that was designed for minimum energy using circuit-level techniques for subthreshold operation. The comparison in Fig. 5.4 shows that synthesized mix-radix design dissipates lower energy and achieves smaller area than custom design from [80]. The energy and area gains from radix factorization exceed those available from circuit-level tuning. At 240 MS/s, synthesized design operates at 0.46 V as compared to 0.27 V in [80], which leaves ample room for circuit-level optimizations if desired.

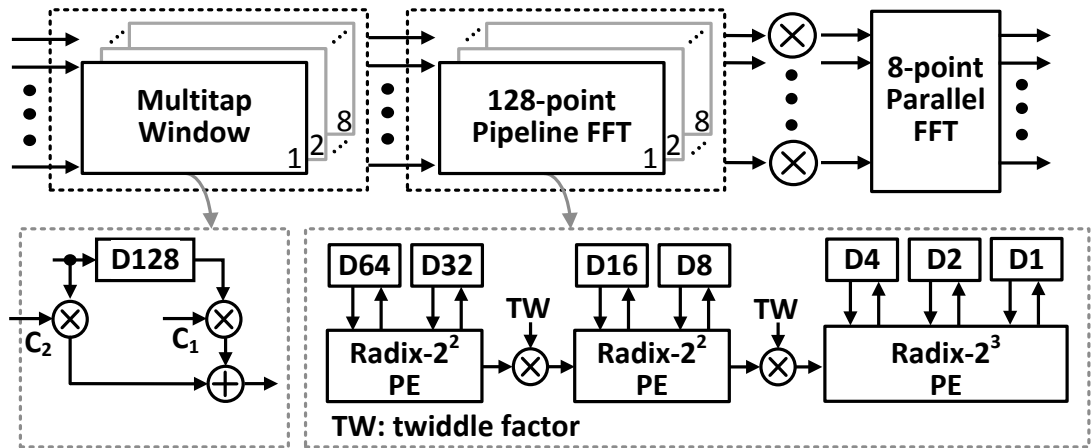


Figure 5.3: Architecture of the frequency-domain power detector.

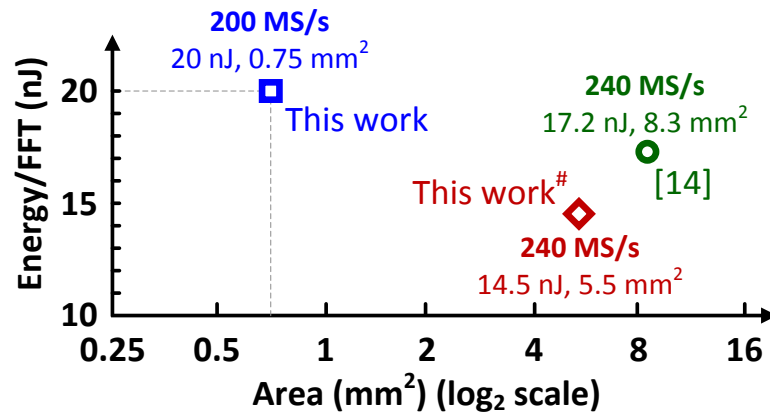


Figure 5.4: Comparison of 1024-point FFTs. Our chip that minimizes power-area product is measured at 200 MS/s. At 240 MS/s, our fully synthesized (# indicates synthesis estimate) mix-radix design optimized for minimum energy achieves lower energy and lower area than the custom radix-4 subthreshold design from [80].

5.2.2 Floating-point Signal Processing

The power estimation block uses multipliers and memory banks to perform squaring and accumulation for the 1024 FFT outputs. In addition, channel-specific STA and DTA blocks consist of multipliers and dividers, as described by (3.16) and (3.17). These operators and storage elements occupy large area and power due to the increased wordlength for large dynamic range signal processing. Using the fixed-point wordlength reduction [97], the PSD estimation requires at least 30 bits for INR of 30 dB and SNR of -5 dB. The hardware area is reduced by exploiting two data properties. First, we assume the spectrum does not change significantly during the sensing period of 50 ms. The per-channel variation is several orders of magnitude smaller than the variation across the entire spectrum. That is, the required dynamic range for each channel is much smaller than the dynamic range of the entire spectrum. Second, signal processing involved in power estimation and programmable averaging as well as sensing adaptation is performed channel by channel. When the magnitude of a channel is large, calculations can be performed using only MSBs. When a channel has a small magnitude, only LSBs contain signal information. In other words, not all 30 bits are necessary for arithmetic. These two observations suggest a floating-point data format, which is composed of mantissa for precision and exponent for magnitude, as the best candidate for signal processing in the adaptation blocks [98-100].

The implementation of required floating-point operations is shown in Fig. 5.5. Two inputs $M_1 2^{E_1}$ and $M_2 2^{E_2}$ are assumed, where M_1 and M_2 are the mantissas and E_1 and

E_2 are the exponents. Both mantissas and exponents are in 2's complement format. Assuming $E_1 \geq E_2$, the addition of $M_1 2^{E_1}$ and $M_2 2^{E_2}$ can be represented by $M_1 2^{E_1} + M_2 2^{E_2} = (M_1 + M_2 \cdot 2^{(E_2-E_1)}) 2^{E_1}$, as in Fig. 5.5(a). The mantissa and exponents of the sum can be calculated by conventional 2's complement operators. The shaded area shows circuitry needed to match the two exponents. The MUXes are used to swap $M_1 2^{E_1}$ and $M_2 2^{E_2}$ if $E_1 < E_2$, and $M_2 \cdot 2^{(E_2-E_1)}$ is realized by the barrel shifter. The floating-point format is inherently suitable for multiplication, $(M_1 2^{E_1}) \cdot (M_2 2^{E_2}) = (M_1 \cdot M_2) 2^{(E_1+E_2)}$, Fig. 5.5(b). No additional block is needed to match the exponents E_1 and E_2 . Squaring is a special case of multiplication, $(M_1 2^{E_1})^2 = M_1^2 2^{2E_1}$, Fig. 5.5(c). All the above floating-point operations are realized by using standard 2's complement arithmetic.

Fixed-point to floating-point conversion is made by removing the consecutive zeros in MSBs. The resulting MSBs are set as mantissa; the number of zeros is accounted as bias in the exponent. The fixed-to-floating point converter is realized by a barrel shifter and a priority encoder, as shown in Fig. 5.7(a). The barrel shifter and priority encoder are used to generate the mantissa and exponent, respectively. The use of floating-point data for large-dynamic range processing allows the use of only mantissa for arithmetic, which drastically reduces the area of arithmetic blocks. Fig. 5.6 shows the architecture of the power estimation block. The wordlength reduction not only saves the area of arithmetic blocks, but also saves the required memory for storing averaged PSD (M1), noise power (M2), and interfering power (M3). The wordlengths of the mantissa and exponent are designed to have minimum area without affecting P_D and P_{FA} . Beyond a

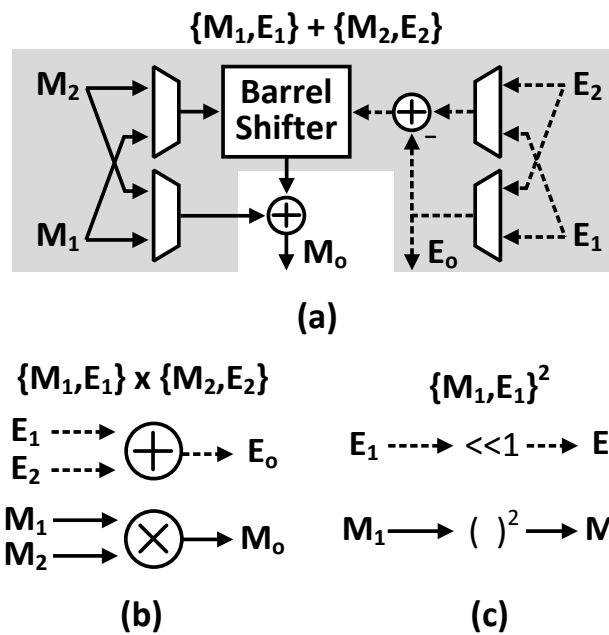


Figure 5.5: Realization of floating-point operations: (a) addition, (b) multiplication, (c) squaring by using 2's complement operators. M1 and M2 indicate mantissas, E1 and E2 indicate exponents. Blocks in the shaded area in (a) perform the exponent matching.

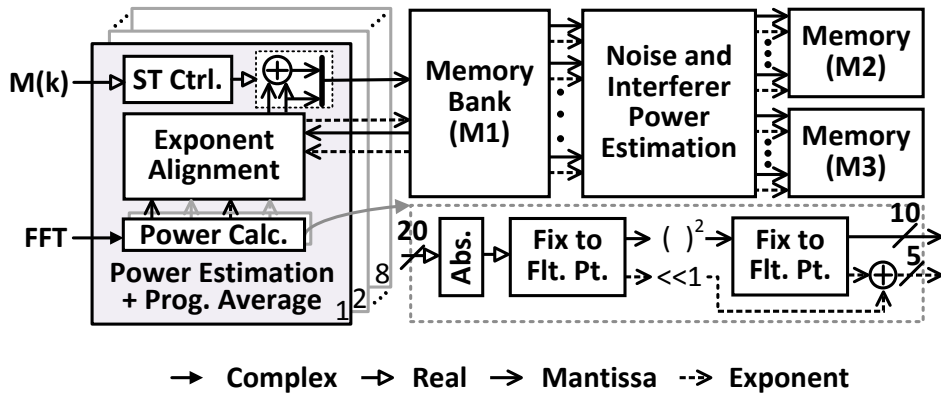


Figure 5.6: Architecture of the power estimation block.

10-bit mantissa and 5-bit exponent, further increase in wordlength does not influence the sensing rates. This holds for the entire sensing range down to -5 dB SNR. As shown in Fig. 5.7(b), the wordlength reduction from floating point results in a 2x less logic area and a 2x less logic power of the power estimation block. The required memory size for M1, M2, and M3 is reduced from 105 kb to 60 kb, a 35% saving. As shown in Fig. 5.8, the use of floating-point arithmetic and the associated wordlength pruning reduces the overall core area and core power by 27.5% and 20.1%, respectively.

5.2.3 Processing Time Overhead Reduction

STA and DTA blocks, Eqs. (3.16) and (3.17), use multiply, divide, and square-root operations. Realization of the divide and square-root using conventional radix-2 long division [101-104] and CORDIC [105-107] would result in processing time of 0.14 ms [94]. When a -5 dB SNR weak primary user without adjacent-band interferers is

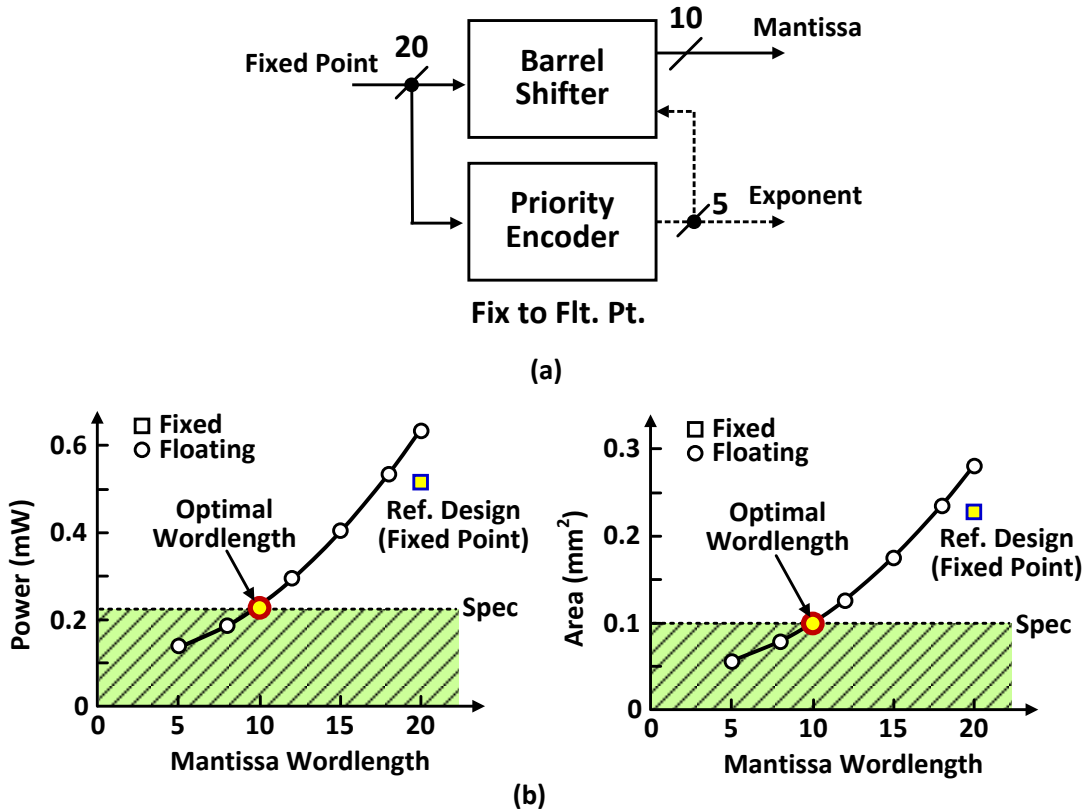


Figure 5.7: (a) Architecture of fixed-point to floating-point conversion. The barrel shifter generates 10-bit mantissa, the priority encoder generates 5-bit exponent. (b) The plots of logic power and logic area of the power estimation block vs. mantissa wordlength show that the sensing performance spec is met with 10-bit mantissa. The floating-point format results in a 2x logic power reduction and a 2x logic area reduction as compared to the fixed-point reference design.

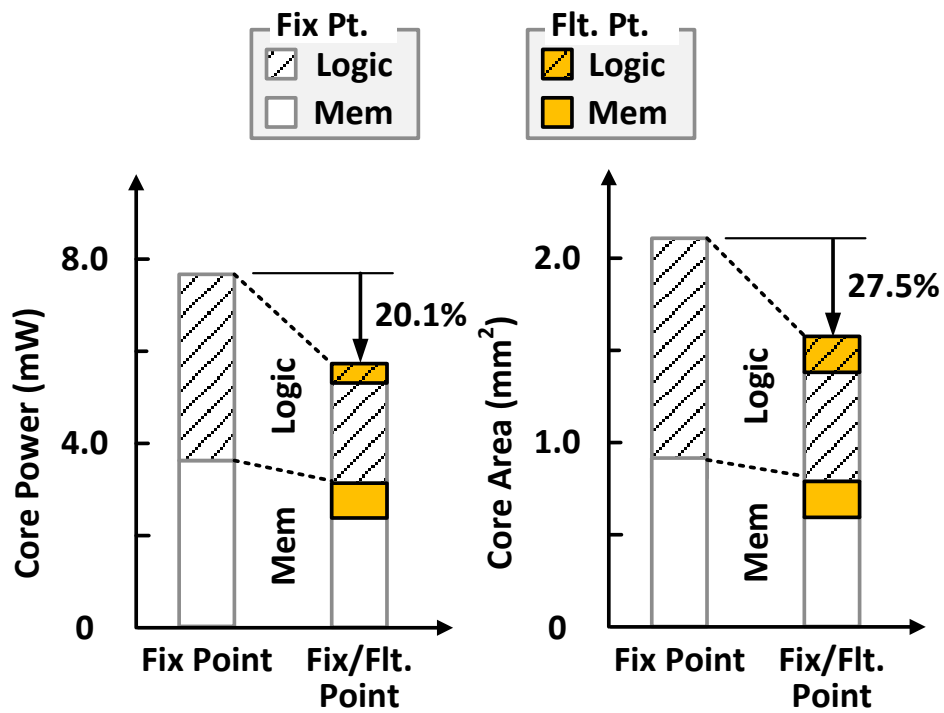


Figure 5.8: The impact of wordlength on the core power and area. A 20.1% of core power (left) and 27.5% of core area (right) are saved by wordlength reduction. The wordlength reduction is enabled by the use of floating-point arithmetic.

detected, the required processing time for PSD estimation is only 0.5 ms, indicating that the STA and DTA blocks occupy a 30% of the entire processing time. This processing time overhead has to be minimized to enhance throughput.

Several design techniques are combined to reduce power and processing time overhead of the STA and DTA blocks. The first step is to reduce power of the multipliers. The calculation of the number of averages, $M(k)$, can be simplified by leveraging constant factors α , $Q^{-1}(P_{FA})$, $Q^{-1}(P_D)$ and SNR [94], shown by

$$M(k) = \alpha \left((1 + \psi(k)) \frac{Q^{-1}(P_{FA}) - Q^{-1}(P_D)}{SNR} - Q^{-1}(P_D) \right)^2 = 74.25 \cdot (1.15 + \psi(k))^2. \quad (5.1)$$

This simplification is possible since the specs for P_D , P_{FA} and SNR are known. The constant multipliers make use of CSD representation leading to a 50% lower power in the STA block.

The second step is to reduce processing time. The bottleneck of processing time is the calculation of $\psi(k) = \sigma_{if}^2(k)/\sigma_{vf}^2(k)$. The realization of $\psi(k)$ using radix-2 long division achieves throughput of 1 b/cycle making the processing time proportional to the wordlength of $\psi(k)$. To reduce processing time, the wordlength of $\psi(k)$ is reduced by exploiting two data properties. First, the number of averages, $M(k)$, is constrained by the 50 ms sensing time. Thus, the dynamic range of $\psi(k)$ can be determined by the output wordlength instead of input wordlengths, resulting in a lower dynamic range and fewer integer bits: only 4 integer bits are needed. Second, $M(k)$ is a positive integer with precision of only 20, indicating the precision requirement of $\psi(k)$ is dictated by $M(k)$, not by $\sigma_{if}^2(k)$ and $\sigma_{vf}^2(k)$. Based on these two observations and applying [97]

to optimize the wordlength, 10-bit output is allowed instead of a full 20-bit precision. This leads to a 2x lower processing time overhead and a 30% lower power in the STA block. The latency of division is further reduced by using the iterative Newton-Raphson algorithm, which achieves quadratic convergence (2 b/cycle). Since $\sigma_{v_f}^2(k)$ has a small variation during the sensing period, $1/\sigma_{v_f}^2(k)$ can be updated in one clock cycle instead of ten clock cycles after the first reciprocal value is calculated.

Figure 5.9 is the architecture of the STA block incorporating the above techniques. The initial-value-calculation and iterative-reciprocal-calculation calculate $1/\sigma_{v_f}^2(k)$ by using the Newton-Raphson algorithm. The sensing time calculation realizes Eq. (5.1). The initial-value-calculation constrains the mantissa between 384 and 738, so we set the initial values of $1/\sigma_{v_f}^2(k)$ to 1/512. Since one pipeline register is inserted in the reciprocal calculation loop, data interleaving is used to reduce area. Namely, a 4-way parallelism is applied in the STA block instead of 8-way parallelism. Applying the above techniques in the STA and DTA blocks, an overall 7x reduction in processing time overhead is achieved. The processing time is reduced to only 0.02 ms, representing a negligible overhead in spectrum sensing.

5.2.4 Summary of Design Optimization

The power and area of the baseband wideband spectrum sensing processor are minimized by architecture-circuit co-design. The optimization framework includes power-area optimized FFT processor, wordlength pruning, and voltage scaling.

Figure 5.11 summarizes our optimization procedure. First, an 8-way parallel 1024-

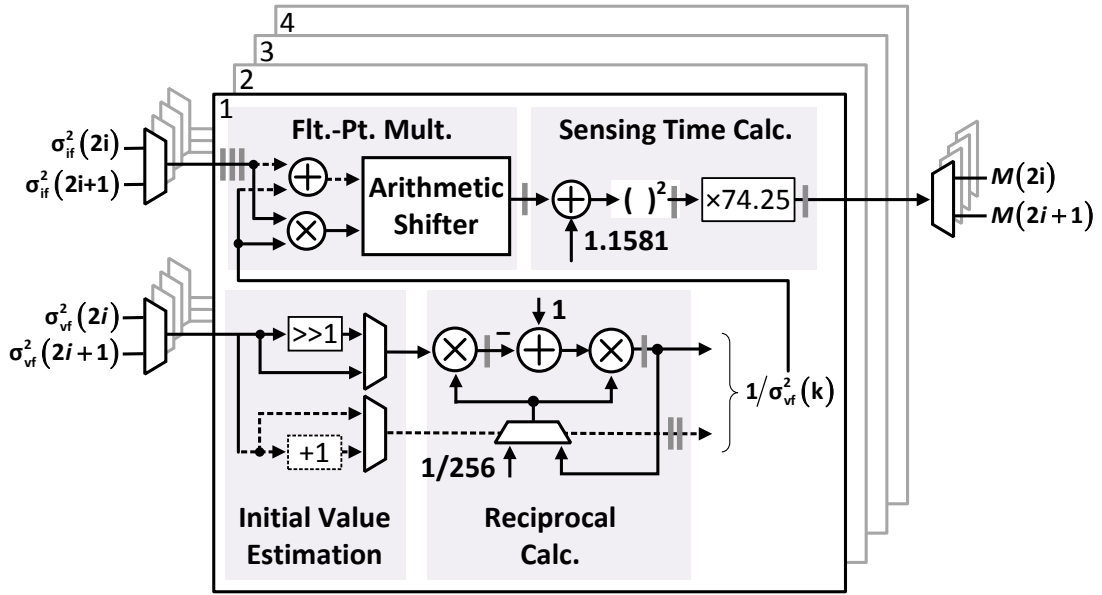


Figure 5.9: Architecture of the sensing-time-adaptation block. The initial-value estimation and the reciprocal calculation are performed using Newton-Raphson algorithm. Bounded sensing time allows for 10-bit output instead of full-precision 20-bit dictated by the input wordlength.

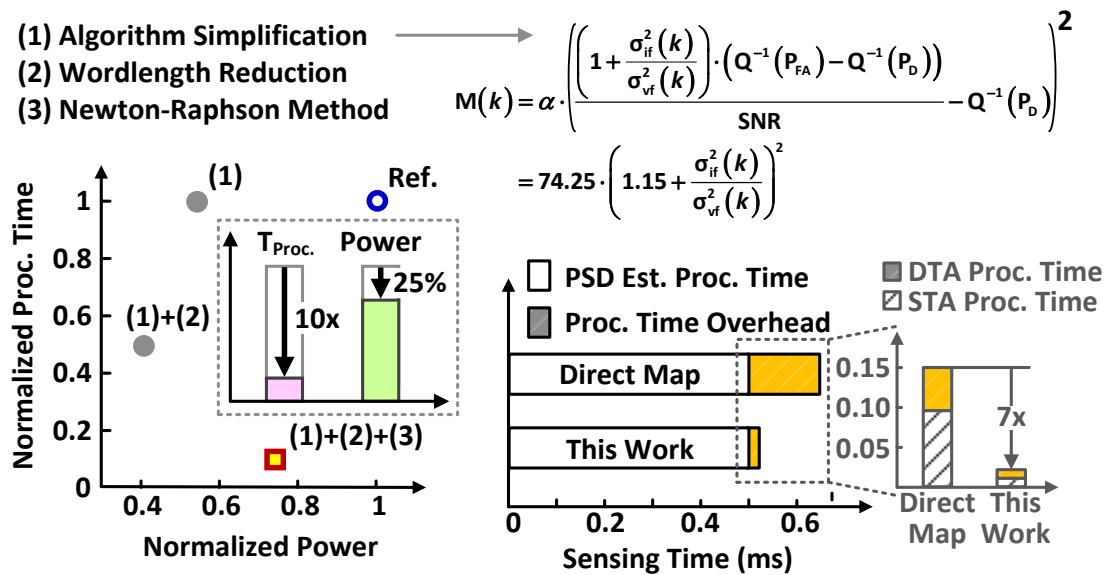


Figure 5.10: Power and sensing-time reduction of the STA block.

point FFT with 25 MHz clock and logic voltage of 0.6 V results in minimum PAP, which represents a 2.3x reduction. Then, applying the radix-factorization with radix- $2^2/2^2/2^3$ and power-area optimized delay lines to the eight banks of 128-point FFTs leads to another 2x reduction in PAP. Next, parallelism along with 25 MHz clock and logic voltage of 0.6 V is applied at the top level. Memory voltage of 0.9 V is chosen to minimize power consumption of the memory bank, resulting in a 1.7x reduction in PAP. Finally, floating-point signal processing is utilized to reduce the power and area of the large-dynamic-range datapath. Floating point with 10-bit mantissa and 5-bit exponent reduces PAP by 1.6x without introducing performance loss. An overall 17.9x reduction in power and 11.9x reduction in PAP is achieved.

5.3 Chip Implementation

Figure 5.12 shows the die photo of the wideband spectrum sensing processor that supports PSD estimation with sensing-time and detection-threshold adaptations. The core area is $1.14 \times 1.44 \text{ mm}^2$. The logic and memory are placed and routed first as hard macros, then carefully placed to facilitate global clock tree synthesis and signal routing. To guarantee timing and support multiple voltage domains, the buffers and cross-coupled level shifters are placed and routed in the remaining area. The total chip area with I/O pads is 2.77 mm^2 . The core supply voltage is tunable between 0.6 V to 1.0 V, and the supply voltage for I/O pads is 1.0/2.5 V. Measurements at 0.7 V confirmed real-time 200 MS/s operation. At the core-I/O boundary, the level shifters are also

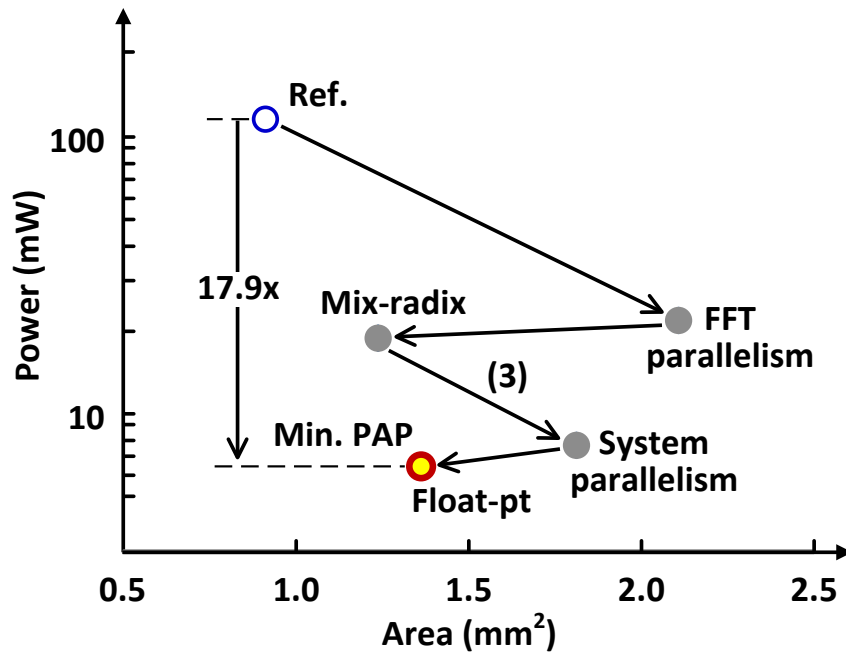


Figure 5.11: Summary of the optimization procedure. Parallelism with voltage scaling and radix factorization with power-area optimized delay lines are applied to the FFT block. Parallelism with voltage scaling is also applied at the system level. Floating-point signal processing with wordlength reduction is applied to STA and DTA blocks. An overall 17.9x reduction in power and 11.9x reduction in power-area product are achieved.

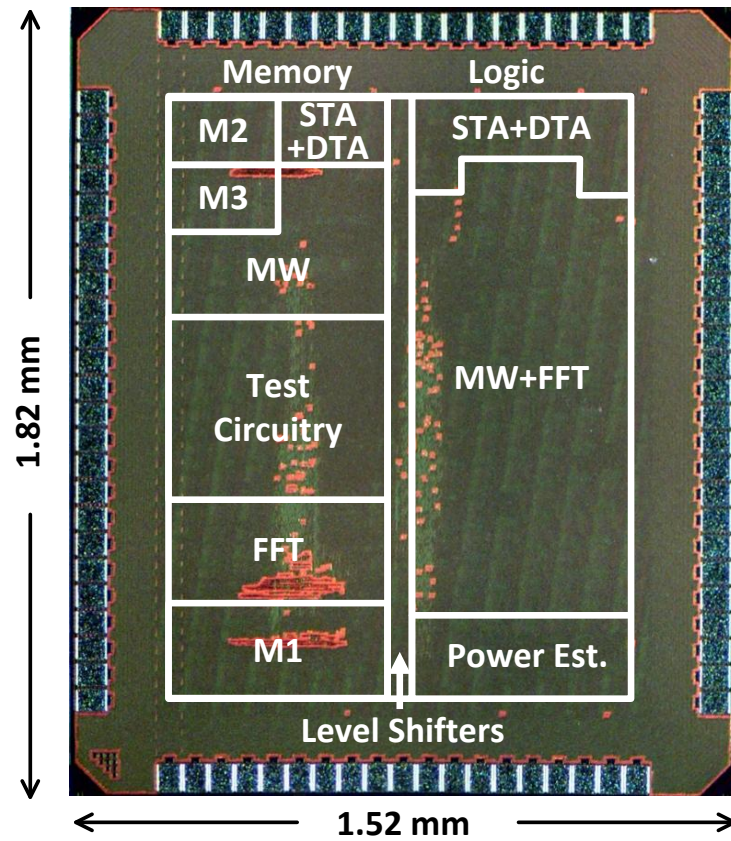


Figure 5.12: Chip micrograph of wideband spectrum sensing processor.

inserted to drive output pads for the entire range of supply voltage.

5.3.1 FPGA-aided Verification

ASIC verification is performed with the use of an FPGA board [108] for pattern generation and data analysis, as shown in Fig. 5.13. The I/O interface between the PC terminal and the FPGA board is built on the BPS environment [71]. Test vectors are stored on the FPGA board, which stimulates the ASIC over two high-speed Z-DOK+ connectors. An external clock source is used to provide a wide range of operating frequency. The outputs of the ASIC are captured into block RAMs for analysis.

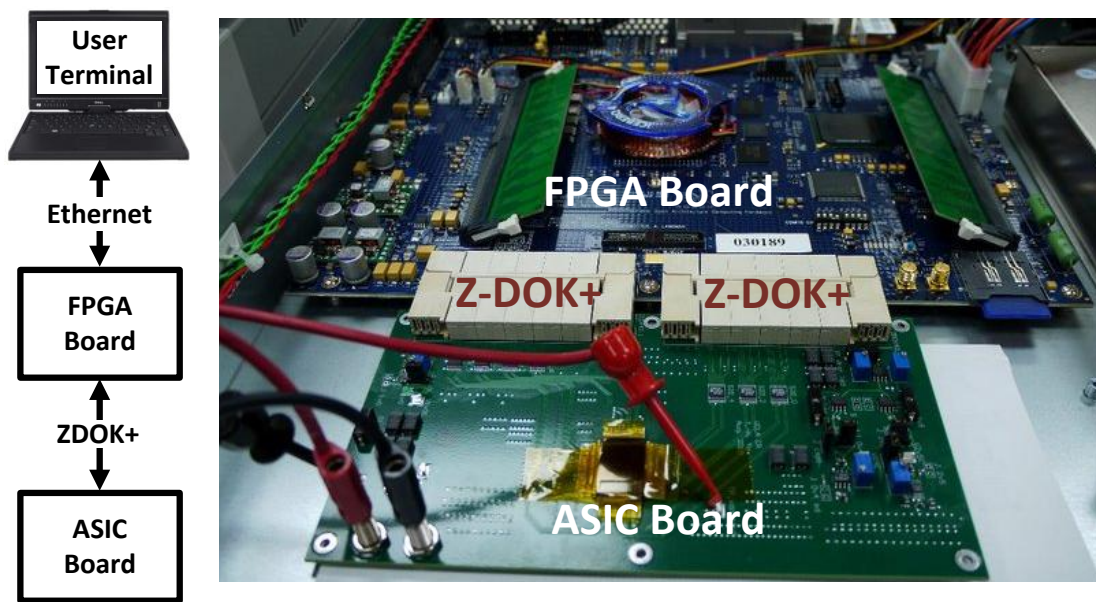


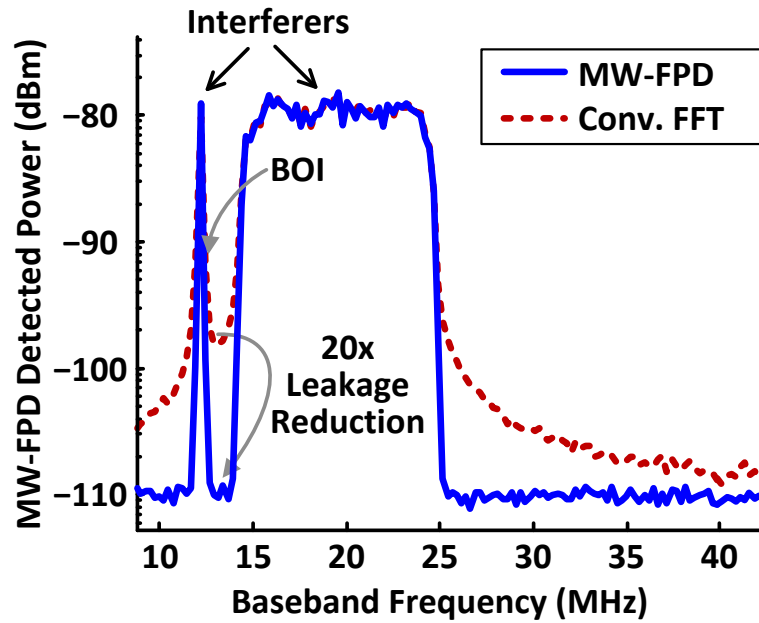
Figure 5.13: FPGA-aided I/O verification. Test vectors are loaded from Matlab into the FPGA board. The results are sampled in BRAM on the FPGA and read back by the terminal.

The FPGA board also contains a XAUI multi-Gigabit interface that allows connectivity to external components. A reconfigurable RF front-end processor with two 12-bit ADCs was connected to the FPGA board through this interface to verify the functionality in a true real-time radio. The whole system, including antenna, analog front-end and ADC, has been verified on this FPGA platform [94] before chip implementation.

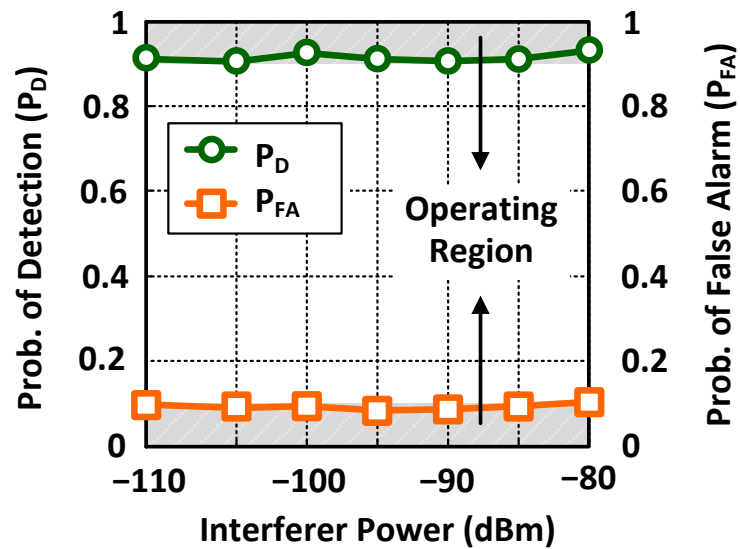
5.3.2 Measurement Results

The chip measurements are shown in Fig. 5.14. Fig. 5.14(a) is a snapshot of spectrum sensing from a 200 MHz bandwidth, with two 30-dB INR interferers and sensing time of 30 ms. A -5 dB SNR primary user is 200 kHz away from the narrowband interferer. A 20x reduction in spectral leakage is achieved by using MW-FPD architecture compared to the conventional FFT-based spectrum sensor. In addition, two bins away from a strong interferer, the interfering power is completely removed by the multitap windowing. The effect of residual interfering power is compensated by the STA and DTA blocks. Fig. 5.14(b) shows that these two blocks maintain P_D and P_{FA} at 0.9 and 0.1, respectively, for a -115 dBm signal in the band of interest (BOI) with adjacent-band interferer power increasing from -80 dBm to -110 dBm. Resolution of 200 kHz is demonstrated in the presence of adjacent-band interferers, with sensing times < 1 ms for 20 dB and < 50 ms for 30 dB INR.

The chip features are summarized in Table 5.1. To minimize power, the design is partitioned into voltage islands consisting of logic and memory blocks, respectively. Clock gating is applied to reduce switching activity of inactive paths. The logic and



(a)



(b)

Figure 5.14: (a) Measured PSD estimated by MW-FPD. Two 30-dB interferers are present 2 bins away from the band of interest. The interfering power is suppressed by 20x using multitap windowing as compared to traditional FFT-based PSD estimation. (b) Measured P_D and P_{FA} as a function of adjacent-band interferer power ($0 \leq INR \leq 30dB$) for $SNR = -5$ dB. STA and DTA blocks effectively maintain $P_D \geq 0.9$ and $P_{FA} \leq 0.1$.

Table 5.1: Summary of Chip Features

Technology	65nm 1P9M CMOS
Supply Voltage	Core: 0.7 to 1.0 V, I/O: 2.5 V
Core Size	$1.44 \times 1.14 \text{ mm}^2$
Gate Count	1.46M
On-chip Memory	106 Kb
Peak Performance	37.5 GOPS @ 200 MS/s
Power Efficiency	5.1 GOPS/mW
Area Efficiency	22.9 GOPS/mm ²

memory blocks occupy 0.82 mm^2 each and operate at 0.7 V and 0.9 V, respectively. The chip dissipates 7.4 mW (3.3 mW by logic, 4.1 mW by memory). The power and area efficiency of the chip are 5.1 GOPS/mW and 22.9 GOPS/mm^2 , respectively (OP = 16-bit add).

5.3.3 Performance Comparison

A comparison of the chip with state-of-the-art is listed in Table 5.2. In spectrum sensing performance, this work supports a 20x wider bandwidth than prior work. It can sense a 200 MHz bandwidth with SNR of -5 dB , $P_D \geq 0.9$, $P_{FA} \leq 0.1$, and sensing time $< 50 \text{ ms}$, with 30 dB adjacent-band INR. Adaptation of sensing time and decision threshold enabled meeting a short sensing time and low SNR requirements in the presence of strong adjacent-band primary users. On the contrary, the detection and

Table 5.2: Comparison with Prior Work

	[53-54]	[55]	[56]	Chip #1 [96] [111]
Radio Bandwidth (MHz)	1	10	0.4	200
P_D	-	-	-	90%
P_{FA}	-	-	-	10%
Processing Time	-	-	-	< 50 ms
Technology	0.18 μm	0.18 μm	65 nm	65 nm
Area (mm^2)	*0.36	*0.36	#0.16	1.64
Power (mW)	*11.7	*8.4	#3.45	7.4
P/BW (mW/MHz)	*11.7	*0.84	#8.62	0.037

*Normalized to 65 nm # Synthesis estimate

sensing performance of prior work is not guaranteed since only power spectrum density estimation was performed. The larger area of our chip is needed to accommodate wideband sensing algorithms, but the power is comparable with other chips. To make a fair comparison, power per bandwidth is shown. With 7.4 mW to sense 200 MHz, our design outperforms prior work in power dissipation per bandwidth by at least 22x despite adding the on-chip sensing-time and detection-threshold adaptations. The lower power dissipation per bandwidth mainly comes from digital as opposed to analog implementation. The digital approach allows voltage scaling and sophisticated sensing algorithms for wideband detection to support simultaneous multiple-channel detection with low power. The power and area efficiency are further improved by the power-

area optimization methodology that incorporates power-area efficient FFT processor, floating-point signal processing, and clock gating.

5.4 Summary

A wideband spectrum sensing processor ASIC is realized in 1.64 mm^2 in a 65-nm CMOS technology. Parallelism is used extensively along with frequency scaling, voltage scaling, and clock gating for high-throughput and low-power wideband signal processing. The power and area of computation-intensive FFT block are minimized by considering parallelism, radix factorization, and the implementation of delay lines. A per-channel floating-point data processing scheme for large dynamic range signal effectively reduces the core area. The sensing time efficiency is improved by applying coefficient lookup, wordlength reduction, and Newton-Raphson algorithms in STA and DTA blocks. Compared to the prior state-of-the-art, our chip dissipates 22x less power per bandwidth for a 20x wider sensing bandwidth. Therefore, a wideband (> 100 MHz) spectrum sensing should leverage multitap windowed frequency power detection, adaptive sensing time and detection threshold estimations.

CHAPTER 6

Chip 2: 500MS/s Wideband Band Segmentation

Processor for Blind Signal Classification

This chapter presents our second chip to demonstrate an energy-efficient channelization scheme. As we found in Chapter 5, channelization by the computation-intensive FFT processor contributes more than 50% of the overall power consumption (4 mW out of 7.4 mW), after design optimizations. The high power comes from the large sensing bandwidth and large FFT size for the required frequency resolution. Channelizing the entire spectrum, however, is not necessary, provided that the spectrum is under-utilized. Instead, we should only channelize a portion of spectrum for low power. We adopt a wideband band segmentation processor for blind signal classification application as a design example. In this design, a partial PSD estimation scheme, which channelizes only the band of interest, is developed for low power. The band of interest is detected via a smaller FFT, which size is determined by minimizing the overall energy dissipation. A miss-detection tolerant signal detection algorithm is proposed for reliable real-time signal detection as well as energy saving. These algorithms and the associated VLSI architectures are implemented in 40-nm CMOS technology.

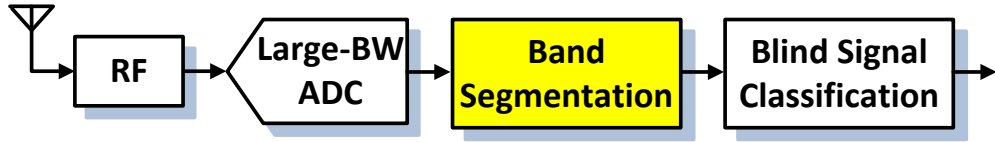


Figure 6.1: System block diagram of blind signal classification.

6.1 Design Specifications

Band segmentation is the key module for energy-efficient blind signal classification application. Figure 6.1 shows a high level system block diagram of blind signal classification. The main objective of band segmentation is to detect *only* signals of interest from a wideband spectrum and then reconstruct the signals in a lower sampling rate to blind signal classifier for low power. In addition, band segmentation estimate signal parameters (bandwidth and center frequency) in order to reduce the processing time of blind signal classifier. Table 6.1 summarizes the design specifications of the proposed band segmentation processor. We consider wideband of 500 MHz with a 100 kHz frequency resolution in order to allow a fine frequency resolution to detect narrowband interferers. The minimum signal strength is 0-dB SNR, and the signal bandwidth is arbitrary from 100 kHz to 500 MHz. The band segmentation processor should identify channel occupancy with a detection probability P_D of 95% and a false-alarm probability P_{FA} of 0.5%. The proposed processor needs to meet an energy constraint of 50 μJ and a processing time of 1 ms. An algorithm-architecture *co-design* framework is developed for an energy-efficient wideband band segmentation processor. The energy minimization methodology will be presented in the next section.

Table 6.1: Wideband CR System Specification

Radio bandwidth	500 MHz
Frequency resolution	100 kHz
Sensitivity	$SNR \geq 0$ dB
False-alarm probability	$P_{FA} \leq 0.5\%$
Detection probability	$P_D \geq 95\%$
Processing time	$T_{sensing} \leq 1$ ms
Energy Consumption	$50 \mu J$

6.2 Energy Minimization Methodology

An algorithmic design framework for energy-efficient signal processing is presented in this section. Key algorithms for energy minimization are: 1) partial PSD estimation algorithm that saves computation for channelization and 2) miss-detection tolerant signal detection algorithm that reduces processing time for parameter estimation. The partial PSD estimation algorithm decomposes the signal detection procedure into multiple steps for low power. The miss-detection tolerant signal detection algorithm applies a probabilistic approach to combines multiple binary decision results to enhance the signal detection performance. Based on these two algorithms, the band segmentation procedure has three steps: 1) coarse sensing, 2) band-pass filtering and fine sensing, and 3) time-domain signal reconstruction. Coarse sensing exploits the channel scarcity and estimates the bandwidth and center frequency of the band of interest with a smaller size FFT. Then, the band of interest, which is only a portion of the spectrum, is selected

and down-sampled by a band-pass filter, and fine sensing channelizes this selected band by a larger size FFT with frequency resolution of 100 kHz. The passband bandwidth, passband center frequency, and the FFT size are programmable, and the configuration is controlled by the estimated bandwidth and center frequency from coarse sensing. Fine sensing finally identifies the channel occupancies and estimates the signal parameters (bandwidths and center frequencies) using the proposed signal detection algorithm. These signal parameters are used to configure the band-pass filter to reconstruct the time-domain signal of interest with lower sampling rate.

The development of the proposed channelization and signal detection algorithms is detailed in the following.

6.2.1 Partial PSD Estimation

The configuration of coarse sensing and fine sensing is presented in this subsection. Coarse sensing is the enabling technique for partial PSD estimation, which estimates the band-of-interest parameters for fine sensing, based on the binary decision results. The number of the consecutive H_1 's represents the normalized bandwidth \hat{b} , and the FFT index of the middle of the consecutive H_1 's stands for the normalized center frequency \hat{c} . Assume a N_{coarse} -point FFT is utilized for coarse sensing, the coarse-estimated normalized bandwidth and center frequency are respectively given by $\hat{b} = \lceil BW_{BOI} \cdot N_{coarse} / Fs \rceil$ and $\hat{c} = \lceil f_{BOI} \cdot N_{coarse} / Fs \rceil$. BW_{BOI} and f_{BOI} are the band-of-interest bandwidth and center frequency, respectively. The knowledge of \hat{b} and \hat{c} allows fine sensing to channelize only the band of interest. Therefore, the FFT size for fine

sensing is reduced, which also enables reduced sampling rate. The minimum FFT size used for fine sensing is $N_{fine,min} = 2 \cdot (N/N_{coarse})$, where we assume fine sensing has $2\times$ over-sampling to prevent aliasing. N is the FFT size for entire spectrum estimation with frequency resolution of 100 kHz, given by $N = 2^{\lceil \log_2(F_s/100kHz) \rceil}$.

Determining the optimal coarse-sensing FFT size N_{coarse} is challenging. A larger coarse-sensing FFT is able to detect narrower band-of-interest bandwidth, which allows a smaller fine-sensing FFT for power saving. However, the coarse-sensing FFT itself introduces overhead in both processing and power consumption. The optimal N_{coarse} balances the energy overhead from coarse sensing and the energy saving from fine sensing. In order to quantify the energy dissipated by different coarse-sensing and fine-sensing configurations, we decompose the overall energy into three parts: 1) coarse sensing, 2) band-pass filtering, and 3) fine sensing. Each part is formulated by the product of power and processing time. A high-level model for the power and processing are presented next.

The power is modeled by the required multiply-accumulate (MAC) operations per second. Coarse sensing involves an N_{coarse} -point FFT for channelization and a MAC operation for power estimation with sampling rate of F_s . Fine sensing performs the same operation as coarse sensing, but with N/N_{fine} times lower sampling rate. The high-level model of power consumption for coarse and fine sensing are $(1 + \log_2 N_{coarse}) \cdot F_s$ and $(1 + \log_2 N_{fine}) \cdot F_s \cdot N_{fine}/N$, respectively.

The processing time determines signal detection performance. The required processing time that guarantees reliable signal detection can be found by using the con-

strained sensing probabilities ($P_{FA}(k)$ and $P_D(k)$) and the test statistic $T(k)$.

Conventionally, $T(k)$ is approximated by Gaussian distribution, and $P_D(k)$ and $P_{FA}(k)$ are expressed by a $Q(\cdot)$ function, according to (2.4) and (2.3). This Gaussian approximation, however, is no longer valid when the constrained P_D and P_{FA} are ≤ 0.005 or ≥ 0.995 . Under this circumstance, $P_D(k)$ and $P_{FA}(k)$ should be modeled by their *true* probability density functions. Theoretically, $T(k)$ is chi-square distributed with degree of freedom $2M$ [38]. According to [38], the *true* expressions for $P_D(k)$ and $P_{FA}(k)$ are

$$P_D(k) = \int_{\gamma_0(k)}^{\infty} f(T(k); 2M|H_1(k))dT(k), \quad (6.1)$$

and

$$P_{FA}(k) = \int_{\gamma_0(k)}^{\infty} f(T(k); 2M|H_0(k))dT(k), \quad (6.2)$$

where M is the number of FFT frames and $f(T(k); 2M)$ represents a chi-square-distributed probability density function with degree of freedom $2M$ and unit noise power. Using (6.2) and (6.1), the required FFT frames M and normalized detection threshold $\gamma_0(k)$ can be computed off-line. The detection threshold $\gamma(k)$ for signal detection should be updated in real-time by scaling the pre-computed $\gamma_0(k)$ with the measured noise power. Let M_{coarse} and M_{fine} be the FFT frames for coarse and fine sensing. The required processing time for the coarse and fine sensing are $M_{coarse} \cdot N_{coarse}/F_s$ and $M_{fine} \cdot N/F_s$, respectively.

Using the model developed in this section, the normalized energy by using partial PSD estimation is expressed by (6.3). E_{BPF} represented the dissipated energy from band-pass filtering. Since the configurations of coarse sensing and fine sensing has only

limited possible values, N_{coarse} and N_{fine} can be obtained through enumeration. Assume 0-dB SNR, the required number of FFT frames for coarse sensing and fine sensing are 160 and 40, respectively. Let $N = 8192$, the normalized energy with respect to different coarse-sensing and fine-sensing configurations is shown in Fig. 6.2. a 64-point FFT for coarse sensing results in minimum energy dissipation. Coarse sensing that uses a FFT with size smaller than 64 points is unable to provide adequate resolution for fine sensing, leading the fine-sensing FFT fail to reconfigure to a smaller size for energy saving. On the other hand, provided that coarse sensing uses a FFT with size larger than 64 points, the sensitivity of energy overhead is larger than the sensitive of energy reduction. It indicates that we can't get further improvement in energy efficiency when the coarse-sensing FFT size is beyond 64 points. The 64-point FFT balances the energy overhead from coarse sensing and the energy saving from fine sensing.

$$\begin{aligned}
\text{Normalized Energy} &= \underbrace{M_{coarse} \cdot N_{coarse} \cdot 1/F_s}_{\text{Coarse } T_{Processing}} \times \underbrace{(1 + \log_2 N_{coarse}) \cdot F_s}_{\text{Coarse Norm. Power}} \\
&\quad + \underbrace{M_{fine} \cdot N \cdot 1/F_s}_{\text{Fine } T_{Processing}} \times \underbrace{(1 + \log_2 N_{fine}) \cdot F_s \cdot N_{fine}/N}_{\text{Fine Norm. Power}} + E_{BPF}. \\
&= \underbrace{M_{coarse} \cdot (N_{coarse} + N_{coarse} \cdot \log_2 N_{coarse})}_{\text{FineEnergy}} \\
&\quad + \underbrace{M_{fine} \cdot (N_{fine} + N_{fine} \cdot \log_2 N_{fine})}_{\text{FineEnergy}} + E_{BPF}
\end{aligned} \tag{6.3}$$

The energy dissipation of the band-pass filter is temporally ignored, because it doesn't affect the energy tradeoff from the configurations of coarse sensing and fine

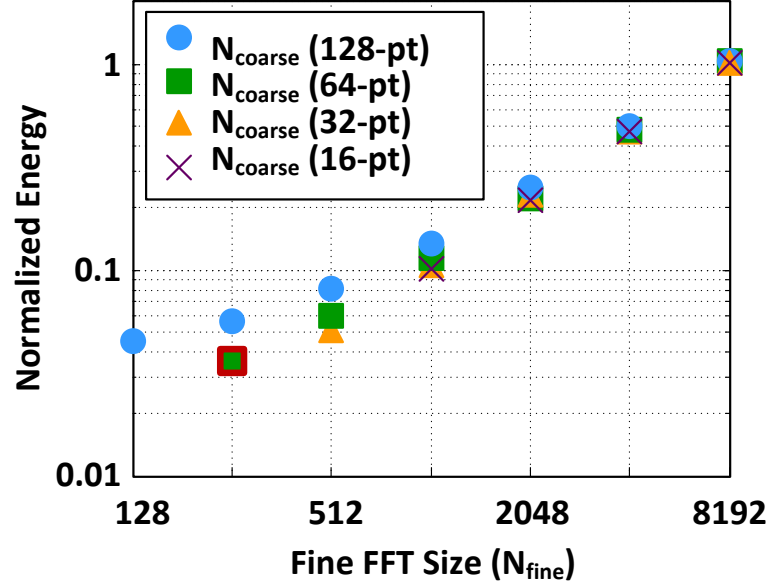


Figure 6.2: Normalized energy with respect to different coarse-sensing and fine-sensing FFTs configuration. 64-point FFT for coarse-sensing results in minimum energy dissipation.

sensing. The specification of the band-pass filter, however, influences the configuration of fine sensing. The design of the band-pass filter is presented in the following.

The band-pass decimation filter selects the band of interest and down-samples signals for low power. The band-pass filter is implemented by down-converting the band of interest into the baseband and low-pass filtering the down-converted signals. Ideally, provided that the low-pass filter has brick-wall frequency response, the down-sampling ratio should be $R_{ideal} = \lfloor F_s/BW_{BOI} \rfloor = \lfloor N_{coarse}/\hat{b} \rfloor$. However, because a CIC filter with brick-wall frequency response is infeasible, the down-sampling ratio needs to be reduced for anti-aliasing. In other words, the down-sampling ratio is related to not only the band-of-interest bandwidth, but also the filtering performance. Fig. 6.3 gives an

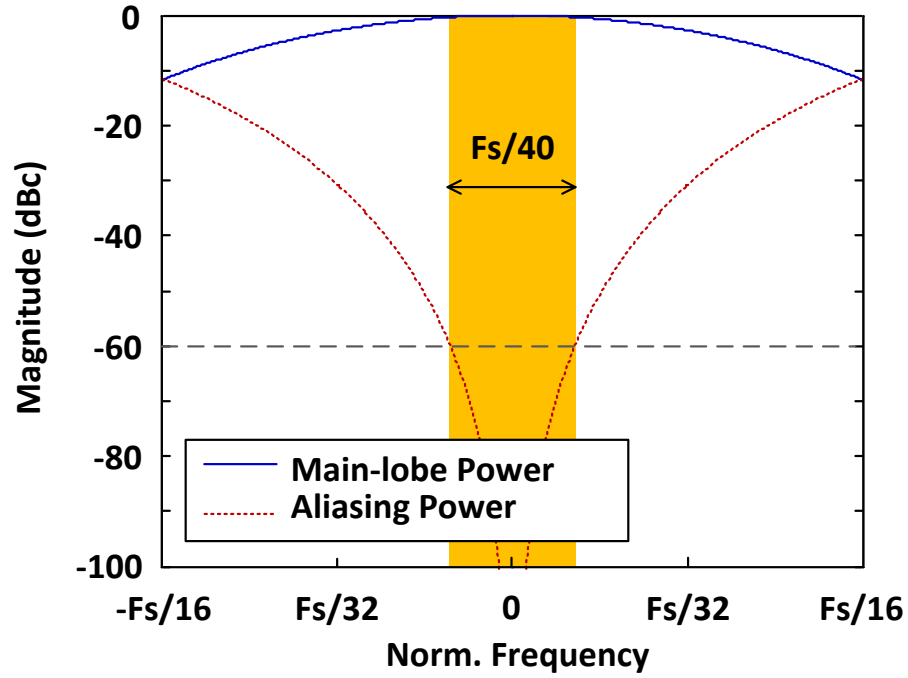


Figure 6.3: Filtering performance of a 3^{rd} -order CIC filter.

example of the filtering performance of a third-order CIC filter. As shown in Fig. 6.3, given a band-of-interest bandwidth of $F_s/40$ and a filtering performance of 60 dB, the maximum down-sampling ratio for a third-order CIC filter is 8. This indicates that the sampling rate for fine sensing is $5\times$ faster than the band-of-interest bandwidth.

The fine-sensing FFT size is $N_{fine} = N/R$, where R is the maximum down-sampling ratio that prevents the band of interest aliased by undetected blockers. According to [112-114], the frequency response of a CIC filter is characterized by down-sampling ratio Φ and the cascade order S . Given a band-of-interest bandwidth \hat{b} and a filter-

performance constraint A, R is given by

$$\begin{aligned}
R &= \arg \max_{\Phi} \Phi \\
\text{subject to } f\left(\frac{BW_{BOI}/2}{F_s}, \Phi\right) &\geq f\left(\frac{1}{\Phi} - \frac{BW_{BOI}/2}{F_s}, \Phi\right) \\
\text{where } f(x, \Phi) &= \left| \frac{\sin(\pi \cdot x \cdot \Phi)}{\sin(\pi \cdot x)} \right|^{2S}, \Phi = 2^i, i = 0, \dots, 6.
\end{aligned} \tag{6.4}$$

We define the over-sampling ratio for fine sensing as $F_s/R/BW_{BOI}$. Smaller over-sampling ratio indicates smaller fine-sensing FFT. Fig. 6.4 illustrates the the over-sampling ratio for different band-of-interest bandwidth. As shown in Fig. 6.4, a higher-order CIC filter, which has sharper frequency response and thus better filtering performance, allows using a smaller FFT. The increased order, however, imposes larger dynamic range on intra-stage datapath, leading to increased wordlength and energy.

The CIC filter specification is designed by investigating the tradeoff between energy consumption and filtering performance. The optimal cascade order S minimizes the overall energy consumption (6.3). A high-level model for power and processing time to quantify E_{BPF} is presented in the following.

A digital mixer is realized by one multiplier operated at F_s . An S^{th} -order CIC filter involves S addition running at F_s , and another S addition running at F_s/R . The required wordlength of an S th-order CIC filter to support the intr-stage dynamic range is

$$WL_{CIC} = S \cdot \log_2 R + WL_{FFT}, \tag{6.5}$$

where WL_{FFT} and WL_{CIC} represent the wordlength of the FFT processor and CIC filter, respectively. The required processing time is the same as for fine sensing. Combining

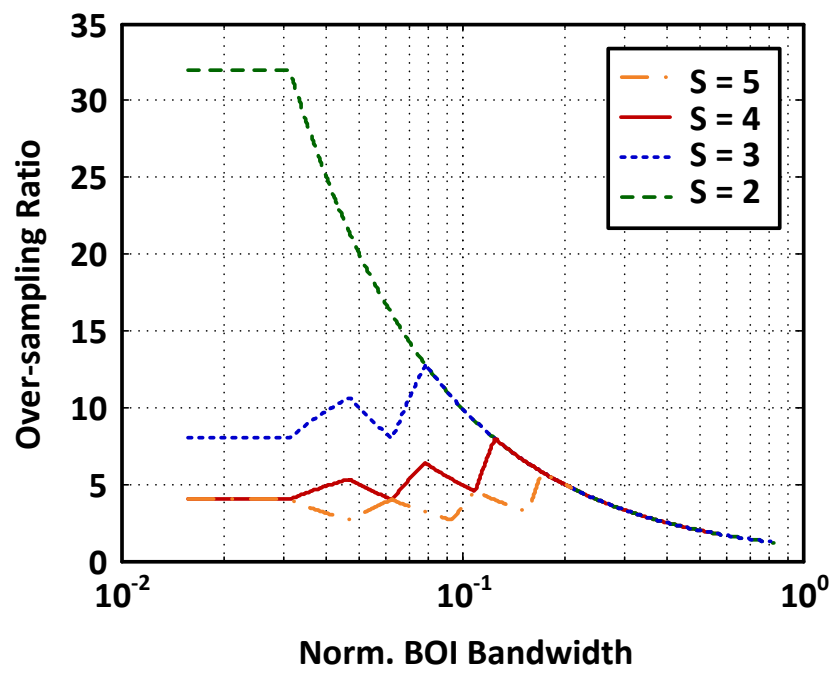


Figure 6.4: Required up-sampling ratio with respect to different band-of-interest bandwidths and cascaded stages.

the power consumption and the processing time, the energy consumption of this band-pass filtering module is expressed in (6.6). WL_{CIC}/WL_{FFT}^2 is the scaling factor to normalize a WL_{CIC} addition to a $WL_{FFT} \times WL_{FFT}$ multiplication.

$$\begin{aligned}
E_{BPF} &= \underbrace{M_{fine} \cdot N \cdot 1/F_s}_{TP_{processing}} \times \underbrace{(F_s + WL_{CIC}/WL_{FFT}^2 \cdot (S \cdot F_s + S \cdot F_s \cdot N_{fine}/N))}_{Norm.Power} \\
&= M_{fine} \cdot (N + WL_{CIC}/WL_{FFT}^2 \cdot S \cdot (N + N_{fine})).
\end{aligned} \tag{6.6}$$

Combining (6.3), (6.4), (6.5), and (6.6), the energy dissipation for different band-of-interest bandwidth and cascade order is quantified. Let the filtering performance as 60-dB. The normalized energy dissipation for different band-of-interest bandwidth with respect to different cascade order is shown in Fig. 6.5. This figure illustrates the impact of filtering performance on the fine-sensing FFT configuration and energy saving. Smaller S has poor filtering performance, which limits the configuration of fine-sensing FFT. Take $S = 2$ for example, only two configurations ($N_{fine} = 8192$ and 4096) are allowed. As $S = 4$, the number of configurations of fine-sensing FFT becomes five ($N_{fine} = 8192, 4096, 2048, 1024,$ and 512). In addition, for the same band-of-interest bandwidth, larger cascade order allows utilizing smaller fine-sensing FFT. Take a band-of-interest of 0.016 for example, the fine-sensing FFT size changes from 4096-point to 512-point, provided that S increases from 2 to 4. This also indicates larger S allows larger range of band-of-interest bandwidth channelized by a smaller FFT. Arbitrarily increasing S , however, doesn't improve energy efficiency further. When $S > 4$, the overhead in energy dictates the saving from the performance gain. Keeping increasing

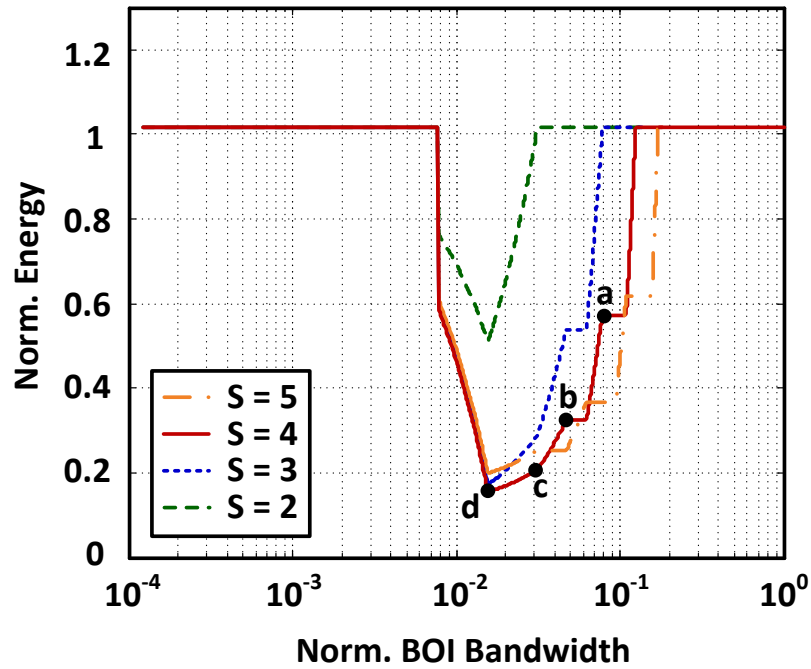


Figure 6.5: Normalized energy with respect to band-of-interest bandwidth. Different fine-sensing FFT configurations are shown: a) 4096-point, b) 2048-point, c) 1024-point, d) 512-point.

S , however, doesn't guarantee lower achievable energy dissipation due to the larger energy overhead. Therefore, $S = 4$ is selected for the CIC filter specification. Our partial PSD estimation, which uses a 64-point FFT for coarse sensing and a fourth-order CIC filter for band-pass filtering, achieves a 6.25x saving in energy as compared the conventional entire PSD estimation.

6.2.2 Miss-detection Tolerant Signal Detection

As mentioned in Sec. 6.2.1, the channel occupancy is identified by comparing the test statistics to detection thresholds. Based on the binary decision, the number of the consecutive H_1 's represents the normalized bandwidth, and the FFT index of the middle of the consecutive H_1 represents the normalized center frequency. Using the consecutive H_1 's for parameter estimation, however, is not robust when the bandwidth is large. As shown in Fig .6.6, a signal of interest is detected on when all the occupied channels are H_1 . In other words, one of the channels has miss detection leads the signal of interest to be regarded as multiple signals. In this scenario, the signal of interest will no longer be classified. Conventionally, to guarantee all occupied channels to be detected, the detection probability of each channel needs to increase, especially for the signal with large bandwidth. Provided that the signal of interest occupies b channels and the required detection probability for this signal is $P_D(SOI)$, the required detection probability of these channels is $P_D(k) = \sqrt[b]{P_D(SOI)}$. Assume a signal occupies 1,000 channels and the signal detection probability is 95%, the detection probability of each channel $P_D(k)$ is approximately 1. This detection probability can be achieved by increasing the processing time by $7\times$. However, longer processing time results in higher energy consumption.

Instead of increasing processing time to enhance $P_D(k)$, an miss-detection tolerant signal detection scheme is proposed to improve $P_D(SOI)$. In the proposed algorithm, an empty channel is found only when the adjacent few channels are all H_0 . When the

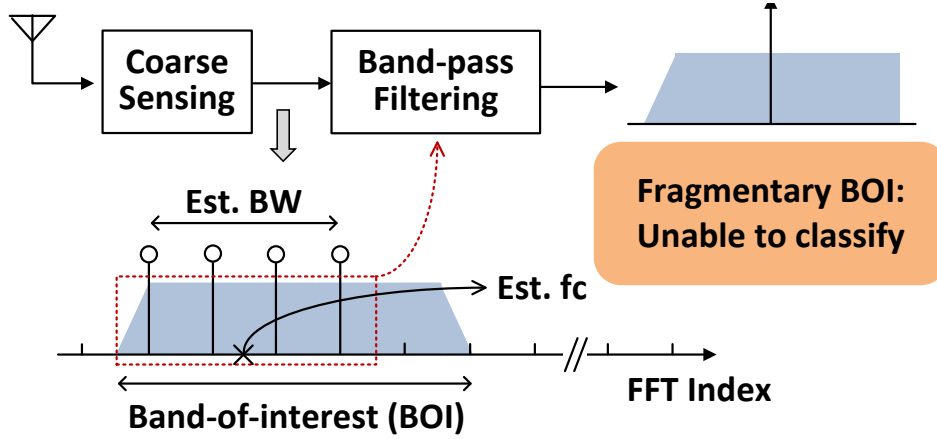


Figure 6.6: Bandwidth and carrier frequency estimation by the binary decision results.

The underestimated bandwidth makes the DSP processor unable to classify the signals.

number of those consecutive H_0 is smaller than or equal to a threshold, we claim those H_0 as H_1 and use the updated binary decision results for parameter estimation. As a result, $P_D(SOI)$ obtained from the updated binary decision results becomes $P_D(SOI) = P_D(k) \cdot \Pr(b - 2|x) \cdot P_D(k)$. $\Pr(b - 2|x)$ represents a detection probability of having at most x consecutive H_0 's among $b - 2$ channels. Assume each channel is statistically independent with detection probability of $P_D(k)$, $\Pr(b|x)$ is given by (6.7), which can be computed recursively with initial condition of $\Pr(b|x) = 1, \forall b \leq x$.

$$\begin{aligned} \Pr(b + 1|x) &= P_D(k) \cdot \Pr(b|x) \\ &+ (1 - P_D(k)) \cdot (\Pr(b|x) - (1 - P_D(k))^x \cdot P_D(k) \cdot \Pr(b - x - 2|x)). \end{aligned} \quad (6.7)$$

Fig. 6.7 shows $P_D(SOI)$ under different bandwidth b and threshold x . Given $P_D(k) = 95\%$, for $b < 8192$, $x = 3$ guarantees signal detection probability larger than 95%. This approach maintains the reliable real-time signal detection without increasing processing

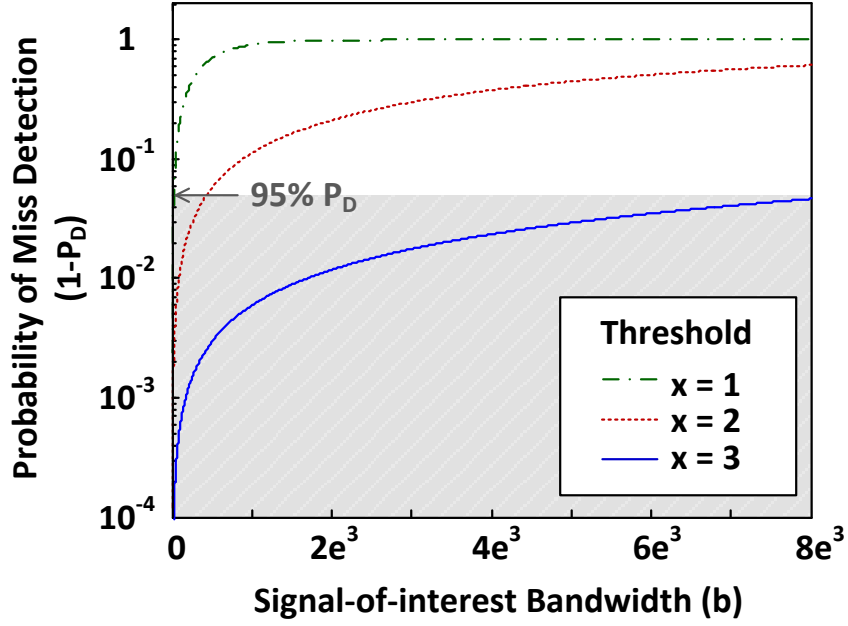


Figure 6.7: The probability of having at most x consecutive zeros among b occupied channels.

time, resulting in a $7 \times$ saving in energy.

6.2.3 Design Example

A design example is shown in Fig. 6.8. We generate a 10 MHz QPSK signal in a 500 MHz spectrum. The signal is centered at 137.5 MHz at SNR of 0 dB, as shown in Fig. 6.8(a). A 64-point FFT is used in coarse sensing to exploit the channel scarcity. The estimated PSD and the detection threshold are depicted in Fig. 6.8(b). The detection threshold is constant due to the AWGN channel. Fig. 6.8(c) illustrates the binary detection results from coarse sensing, by which \hat{b} and \hat{c} are estimated. The digital mixer shifts the band-of-interest into baseband according to $\hat{c} = 16.5$, and the fine-sensing

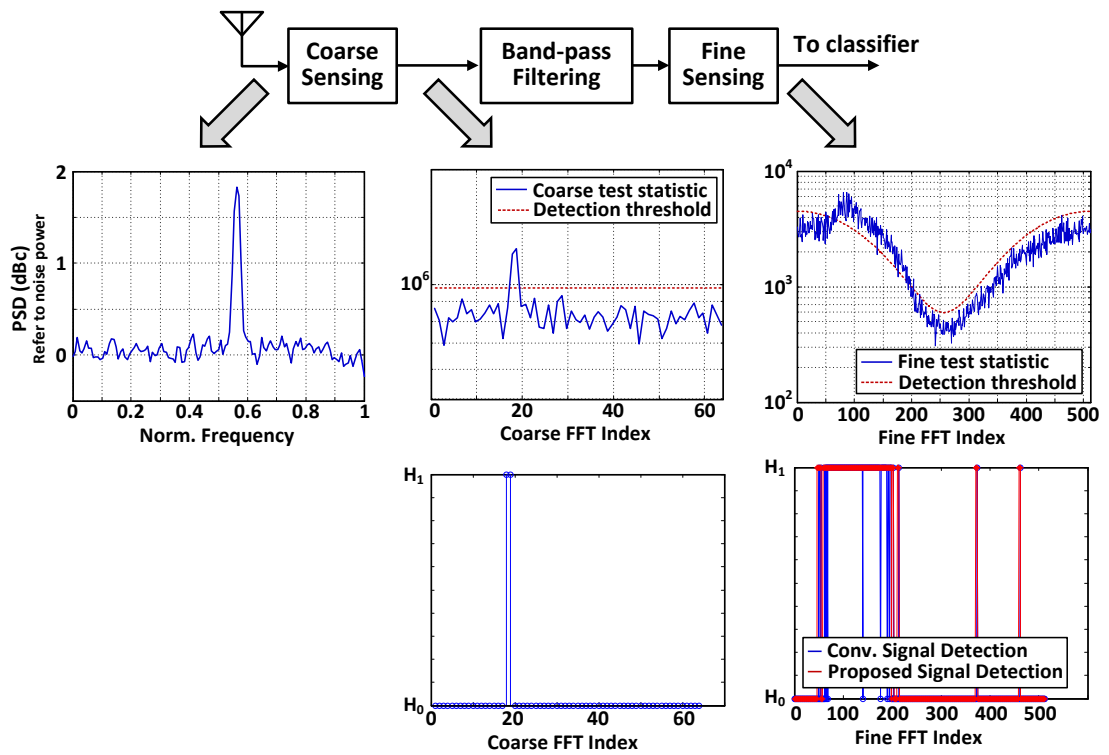


Figure 6.8: A design example. (a) Power spectrum density of the wideband scenario. (b) Estimated PSD and the detection threshold of coarse sensing. (c) Binary decision result from coarse sensing. (d) Estimated PSD and the detection threshold of fine sensing. (e) Binary decision result from fine sensing.

FFT is reconfigured to 512 points according to $\hat{b} = 2$. Fig. 6.8(d) demonstrates the estimated PSD and the detection threshold in fine-sensing stage. The detection threshold is shaped by the frequency response of the 4-order CIC filter. The binary detection results from fine sensing are shown in Fig. 6.8(e). The false-alarm probability is $< 0.5\%$. The fine estimated bandwidth $\hat{b} = 170$ and carrier frequency $\hat{c} = 132$ are used to reconstruct the down-sampled time-domain signals. These two parameters are also passing for signal classification to reduce computational complexity [110]. Resolution of 100 kHz is demonstrated, with sensing times $< 1\text{ms}$ for 0 dB SNR.

6.3 Proposed VLSI Architecture

Figure 6.9 shows the block diagram of the proposed band segmentation processor. It consists of a digital mixer, a CIC filter, a reconfigurable FFT processor, a power estimation module, and a parameter estimation module. The digital mixer and the CIC filter are used for computationally efficient band-pass filter. The reconfigurable FFT processor, the power estimation module, and the parameter estimation module are shared for coarse sensing and fine sensing. The FFT processor and CIC filter are the two key modules that need to be carefully designed to improve energy efficiency. The high-throughput reconfigurable FFT processor is computational-intensive, and the intra-stage datapath of the CIC filter requires large dynamic range. Several architectural and signal processing techniques are proposed to minimize power and area. The FFT processor uses parallelism and radix factorization to achieve minimum

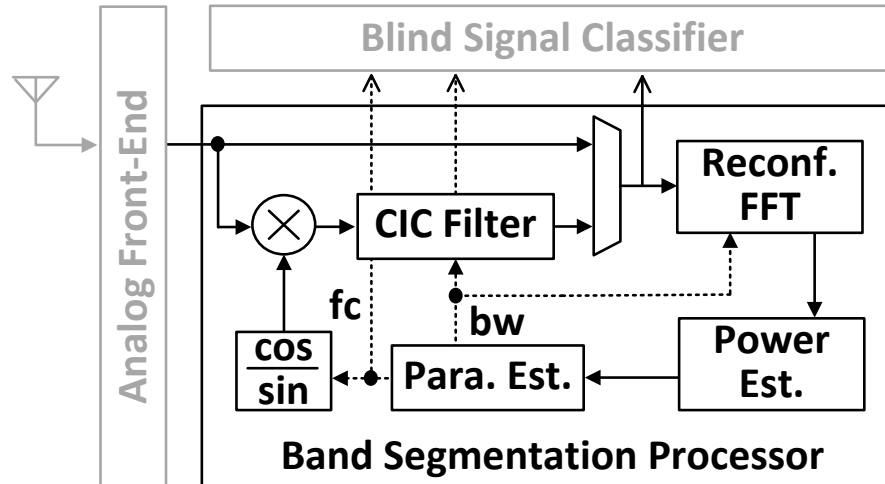


Figure 6.9: Energy-efficient band segmentation, which consists of coarse sensing and fine sensing. Coarse sensing is to select band-of-interest, and fine sensing is to channelize the selected band with finer resolution.

energy-area product. A feed-forward CIC architecture is designed to reduce the required wordlengths and enable parallelism for low power.

6.3.1 Energy-Area Efficient Reconfigurable FFT

A 64-8192 point FFT is implemented using a multi-path pipelined architecture, which provides high throughput and low power, thus maximizing energy efficiency. The energy efficiency of the pipelined architecture is further improved via architecture parallelism [89-90], which allows for voltage scaling due to a lower operating frequency. Also, a radix-factorization methodology [95] is used to determine the optimal radix structure. This reconfigurable FFT is decomposed into 16 banks of 4-512-point pipelined FFTs followed by a 16-point parallel FFT to minimize power-area product.

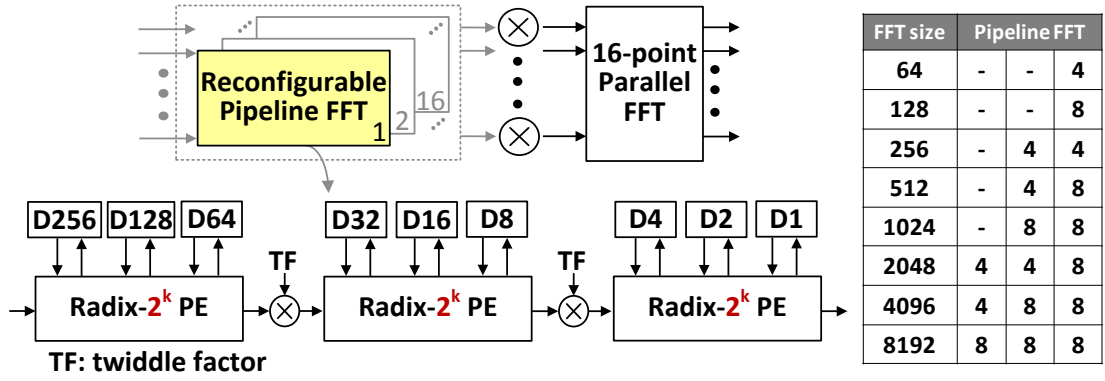


Figure 6.10: A multi-path pipelined reconfigurable FFT with optimal parallelism and radix-factorization.

Fig. 6.10 presents the proposed architecture and the corresponding reconfiguration.

Fig. 6.11 shows the comparison with the state-of-the-art FFTs. Number of 1024-point FFTs per $\mu J \cdot mm^2$ is adopted to evaluate the performance. Compared to [80], which was designed for minimum energy using circuit-level techniques for subthreshold operation, we demonstrate a 4.1x improvement in energy-area product. However, compared to our dedicated 1024-point FFT design [111], this work sacrifices a 2.3x higher energy-area product due to the flexibility.

6.3.2 Energy Efficient CIC Filter

The conventional S^{th} -order CIC decimation filter with down-sampling ratio of R consists of a cascade of S integrator operating at the input sampling rate F_s , and R -fold decimator, and a cascade of S comb filters operating at the reduced sampling rate F_s/R . This architecture, however, has two major challenges for low power design: 1)

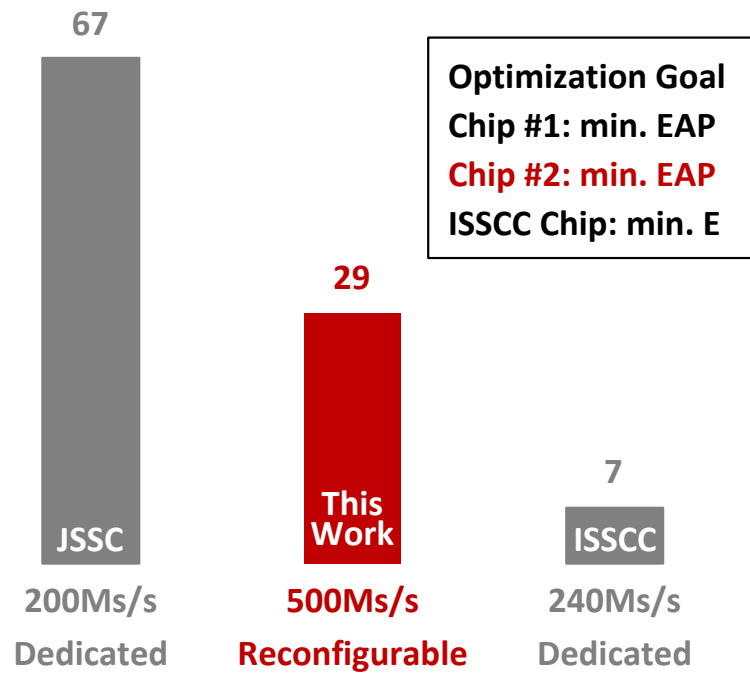


Figure 6.11: Comparison with the state-of-the-art FFTs.

The intra-stage adders of the CIC filter requires long wordlength to support the large dynamic range. According to (6.5), to support down-sampling ratio up to 1024, extra 40 bits are required for the intra-stage adders. 2) The recursive nature of the CIC filter makes the integrators unable to be pipelined or parallelized, indicating that aggressive voltage scaling is not allowed for this module.

Several signal processing techniques are applied to reduce the power of the CIC filter. By exploring the down-sampling ratios for fine sensing are only power of 2, the transfer function of the filter can be represented as a purely feed-forward function [115],

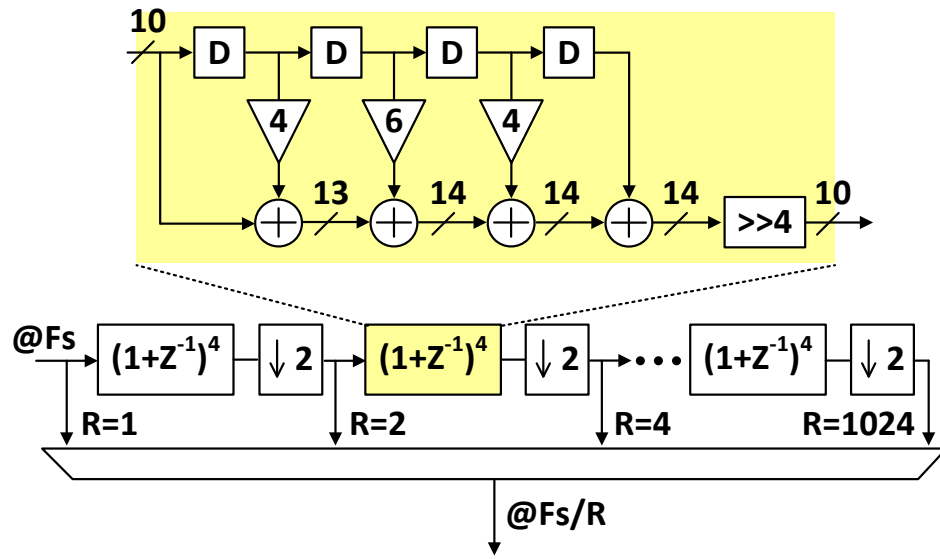


Figure 6.12: 4th-order feed-forward CIC filter, supporting down-sampling ratio up to 1024, amenable to parallelism with smaller wordlength.

as shown in (6.8).

$$\begin{aligned}
 H(z) &= \frac{(1 - Z^{-R})^S}{(1 - Z^{-1})^S} \\
 &= (1 + Z^{-1})^S (1 + Z^{-2})^S \dots (1 + Z^{-R/2})^S.
 \end{aligned} \tag{6.8}$$

The modified transfer function is realized by the architecture shown in Fig. 6.12. This feed-forward architecture allows the wordlength of the intra-stage adders increase by only 4 bits, which significantly reduces the hardware cost and shortens the critical path delay of each intra-stage adder. Ten cascaded stages of $(1 + Z^{-1})^4$ are needed for the maximum down-sampling ratio of 1024, and each stage operates at 2x lower sampling rate than the previous stage. Only the first i stages turn on for down-sampling ratio of 2^i . The last $10 - i$ stages are clock-gated for power saving. This feed-forward architecture also allows parallelism for voltage scaling. A 16-way parallelism is chosen

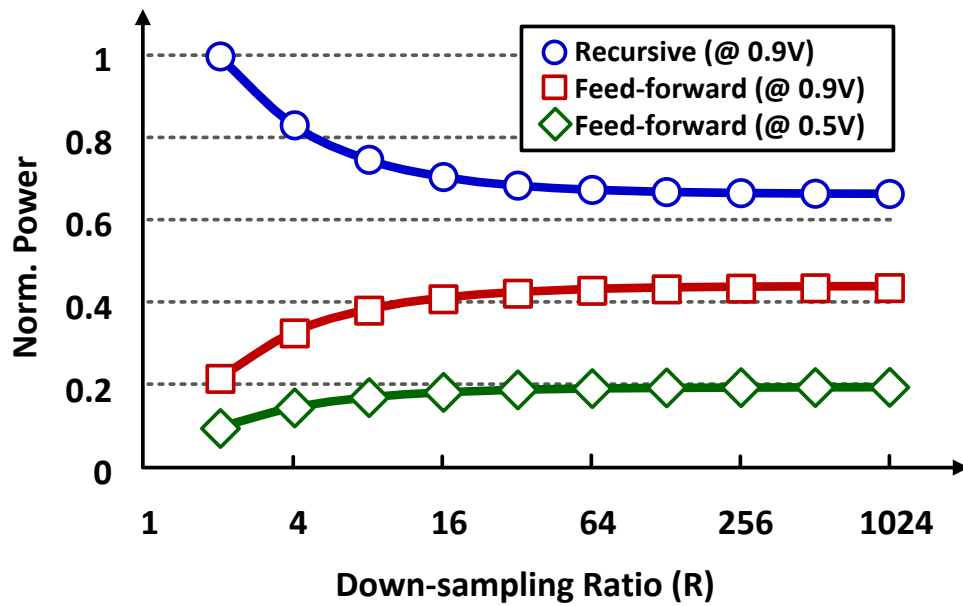


Figure 6.13: Power saving from the feed-forward architecture.

to synchronize CIC filter and FFT processor, which reduces the clocking power and improves the system robustness.

The impact of the feed-forward architecture on power saving is quantified. The number of 1-b full-adder operation per cycle scaled by the supply voltage is used to represent the normalized power. Fig. 6.13 illustrates the power saving with respect to different optimization techniques. As shown in Fig 6.13, at least 2.5x saving is achieved from the wordlength reduction. The 16-way parallelism allows supply voltage of 0.5 V, resulting in an additional 4x saving in power.

6.3.3 Summary of the Design Optimization

The energy consumption of the band segmentation are minimized by algorithm-architecture co-design. The optimization framework includes partial PSD estimation scheme, real-time signal detection algorithm, power-area optimized reconfigurable FFT processor, wordlength pruning, and voltage scaling.

Figure 6.14 summarizes our optimization scheme with respect to different signal bandwidths. Different signal bandwidth has different energy saving. Take signal bandwidth of 10 MHz for example. First, a miss-detection tolerant signal detection scheme enables reliable real-time signal detection, leading to a 7x reduction in processing time and energy consumption. Then, using a 64-point FFT to detect the band-of-interest and partially channelizing the band-of-interest via a smaller FFT result in a 6.25x saving in energy. Then, In addition, a 16-way parallel 64-8192-point FFT with optimal radix factorization enables energy-efficient VLSI signal processing, a 10x low energy is achieved compared to state-of-the-arts [95]. Finally, the feed-forward CIC filter allows smaller wordlength and voltage scaling, reducing the energy by 3.75x. In summary, algorithmic optimization achieves 7x–45x reduction in energy consumption, and architectural optimization contributes to another 5x–10x energy saving. An overall 70x–210x saving in energy is achieved by the algorithm-architecture co-design.

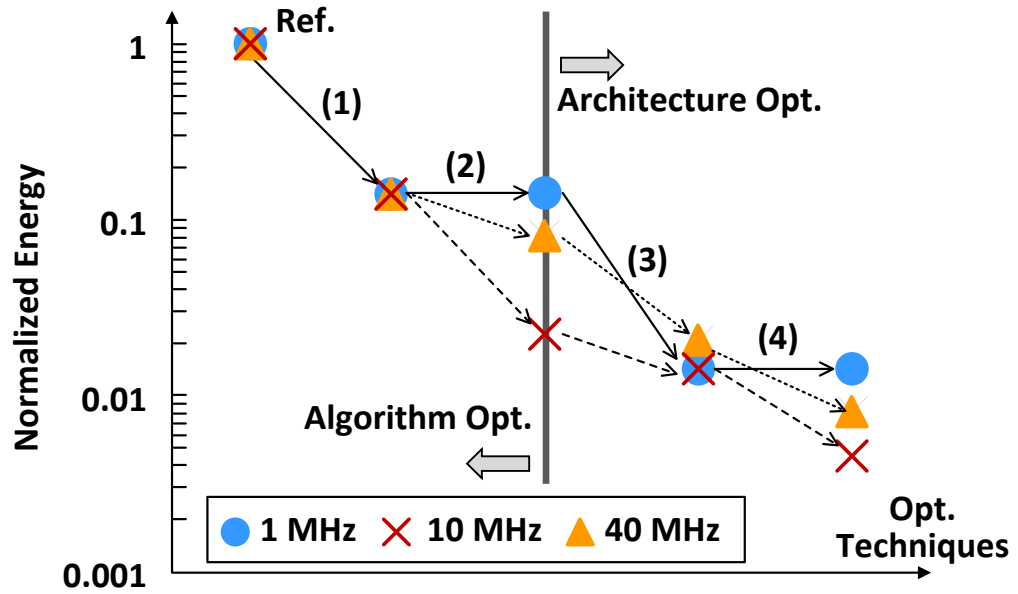


Figure 6.14: Summary of design optimization for signal bandwidth of 1 MHz, 10MHz, and 40 MHz. Four techniques are used to improve energy efficiency. 1) Miss-detection tolerant signal detection. 2) Partial PSD estimation. 3) 16-way parallelism and optimal radix factorization. 4) Feed-forward CIC filter architecture.

6.4 Chip Implementation

The proposed algorithms and the associated optimized architectures are implemented in a standard 40-nm CMOS process. Fig. 6.15 shows the blind signal classification processor, where the wideband band segmentation processor is on the bottom right corner. The layout view of the band segmentation processor is shown on the bottom, including a 64-8192-point reconfigurable FFT, a digital mixer, and a CIC filter for partial PSD estimation algorithm. The core area is $1.19 \times 1.08 \text{ mm}^2$. Different from the first chip, only the logic is placed and routed first as a hard macro. Then, this logic macro, memory banks are carefully placed to facilitate global clock tree synthesis and signal routing. To guarantee timing and support multiple voltage domains, the buffers and cross-coupled level shifters are placed and routed in the remaining area. The core supply voltage is tunable between 0.65 V to 0.9 V, and the supply voltage for I/O pads is 0.9/1.8 V. Measurements at 0.65 V confirmed real-time 500 MS/s operation. At the core-I/O boundary, the level shifters are also inserted to drive output pads for the entire range of supply voltage.

6.4.1 FPGA-aided Verification

ASIC verification is performed with the use of a Xilinx Kintex-7 board for pattern generation and data analysis, as shown in Fig. 6.16. The I/O interface between the PC terminal and the FPGA board is build on Xilinx ChipScope Pro through Joint Test Action Group (JTAG) cable. Test vectors are stored on the FPGA board, which stim-

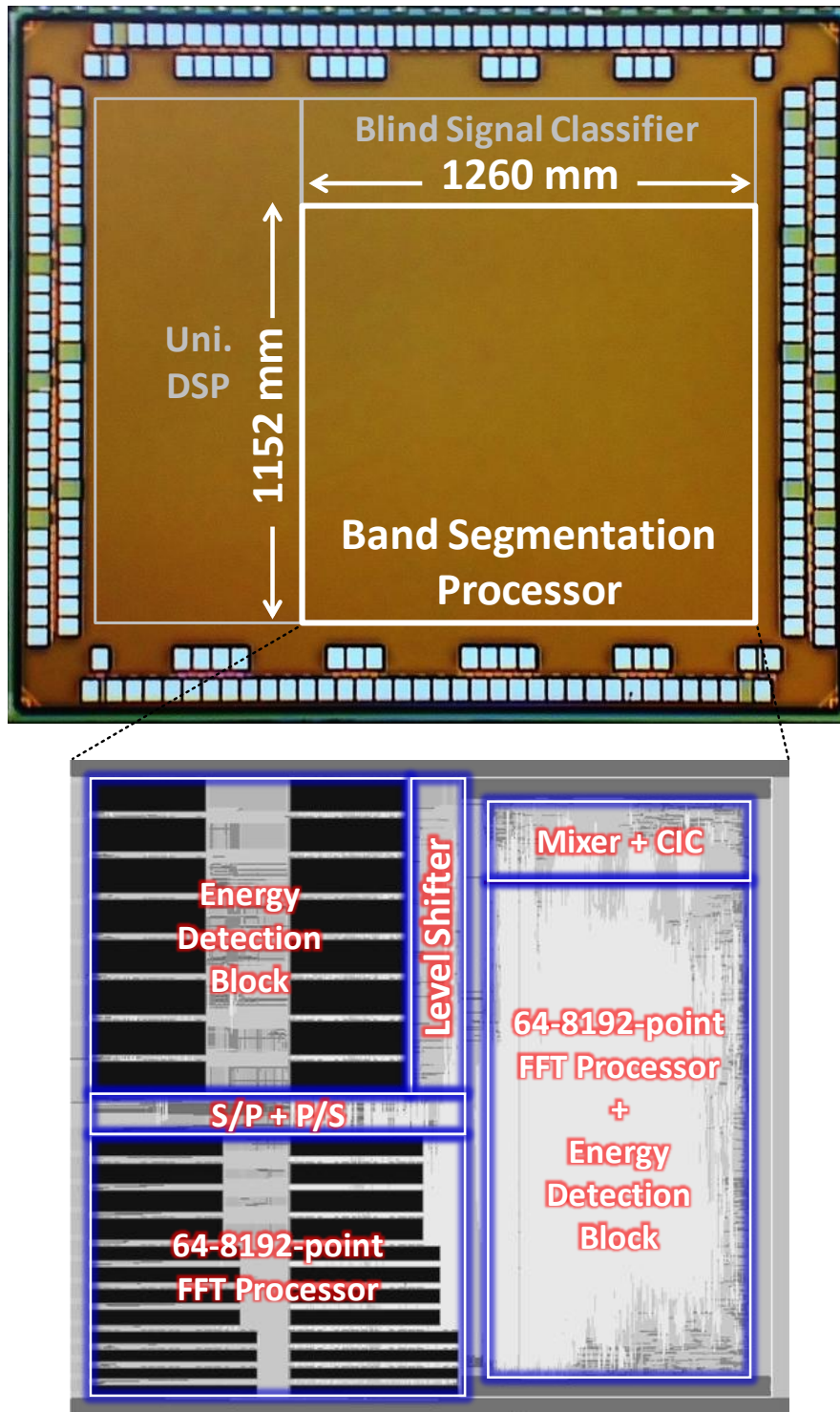


Figure 6.15: Chip micrograph and the layout of the band segmentation processor.

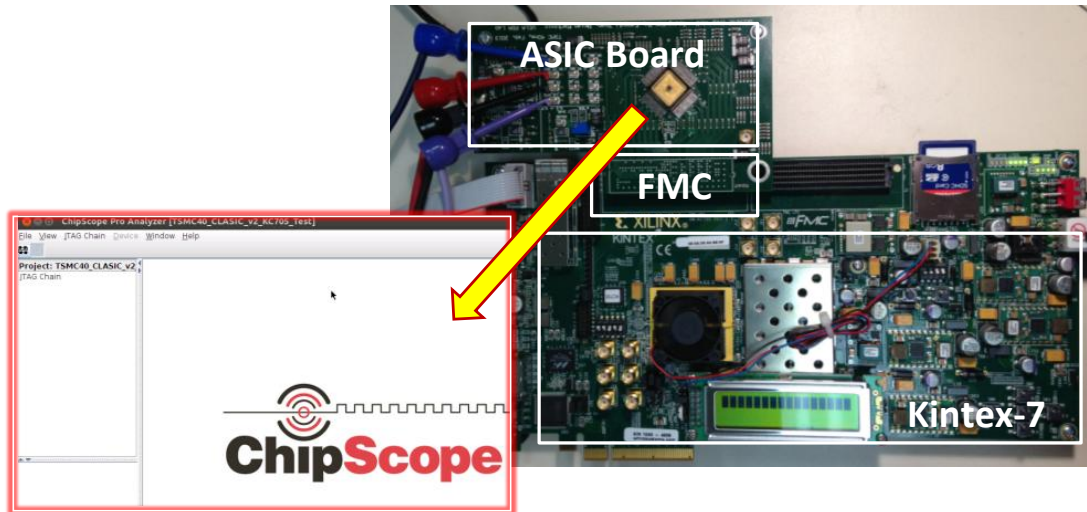


Figure 6.16: FPGA-aided I/O verification. Test vectors are stored on the FPGA board. The results are sampled and verified in real-time by using Xilinx ChipScope.

ulates the ASIC over one high-speed FMC connector. The nominal voltage for FMC pins is 2.5 V, which can be real-time adjusted and monitored via Fusion Digital Power Designer from Texas Instruments. The clock signal is generated by the Xilinx Digital Clock Manager (DCM) module, providing a wide range of operating frequency. The outputs of the ASIC are captured by the ChipScope Integrated Logic Analyzer (ILA), allowing real-time monitor and analysis.

6.4.2 Measurement Results

The chip features are summarized in Table 6.2. To minimize power, the design is partitioned into voltage islands consisting of logic and memory blocks, respectively. Clock gating is applied to reduce switching activity of inactive paths. The logic and memory blocks occupy 0.58 mm^2 and 0.70 mm^2 and operate at 0.65 V and 0.75 V,

Table 6.2: Summary of Chip Features

Technology	40nm 1P9M CMOS
Supply Voltage	Core: 0.65 to 0.9 V, I/O: 2.5 V
Core Size	$1.19 \times 1.08 \text{ mm}^2$
Gate Count	2.30M
On-chip Memory	298 Kb
Peak Performance	138 GOPS @ 500 MS/s
Power Efficiency	3.9 GOPS/mW
Area Efficiency	107.8 GOPS/mm ²

respectively. For entire spectrum estimation, the chip dissipates 35.4 mW. For partial PSD estimation, the chip dissipates only 10.2 mW. The peak power and area efficiency of the chip are 3.9 GOPS/mW and 107.8 GOPS/mm², respectively (OP = 16-bit add).

Figure 6.17 shows the measured energy with respect to different signal bandwidths. In order to save the energy, the fine-sensing FFT configurations depend on channel scarcity and signal bandwidths. The measurement results validate that the partial PSD estimation algorithm effectively improves the energy efficiency. The maximum energy is 23.7uJ for detecting the entire spectrum, and the minimum energy is 7.2uJ for detecting only 1/16 spectrum. Without partial PSD estimation, the energy dissipation is 23.7uJ regardless the signal bandwidth, indicating that a maximum 3.2x saving in energy is achieved.

The impact of the band segmentation to blind signal classification is shown in Fig.

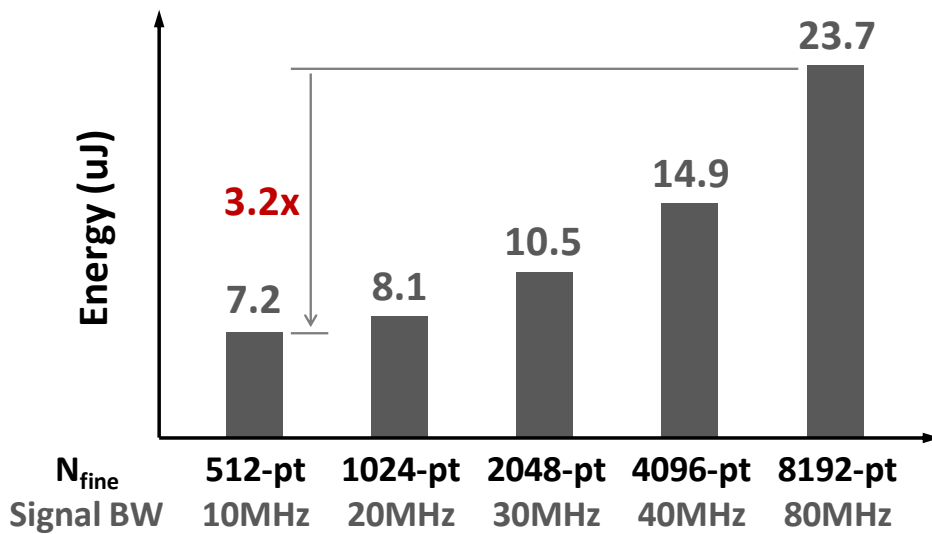


Figure 6.17: Measured energy with respect to different signal bandwidths. A maximum 3.2x saving in energy is achieved by using partial PSD estimation algorithms.

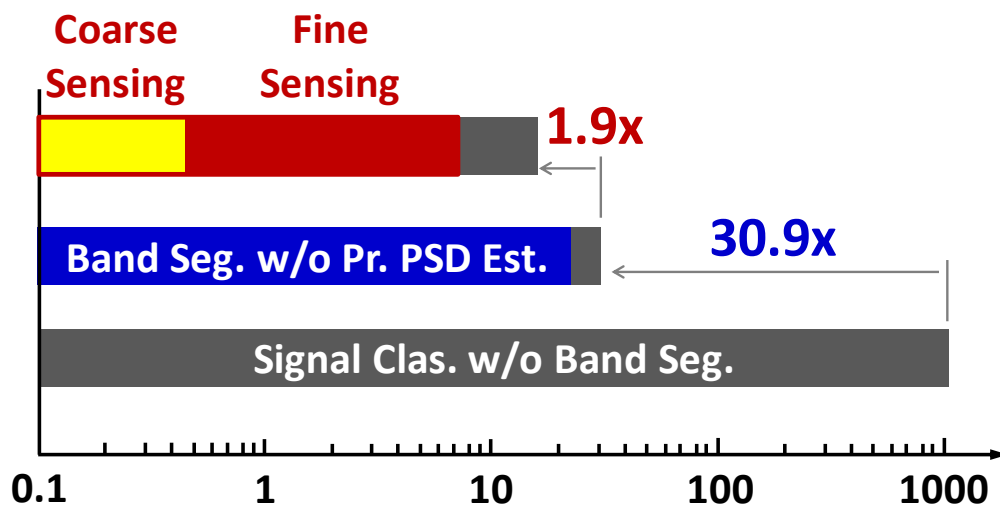


Figure 6.18: Measured energy with respect to different signal bandwidths. A maximum 3.2x saving in energy is achieved by using partial PSD estimation algorithms.

Table 6.3: Comparison with Prior Work

	[53-54]	[55]	[56]	Chip #1	Chip #2
Radio Bandwidth (MHz)	1	10	0.4	200	500
P_D	-	-	-	90%	95%
P_{FA}	-	-	-	10%	0.5%
Processing Time	-	-	-	< 50 ms	< 1 ms
Technology	0.18 μm	0.18 μm	65 nm	65 nm	40 nm
Area (mm^2)	*0.18	*0.18	*#0.08	*0.82	1.28
Power (mW)	*8.3	*5.9	*#2.4	*5.2	10.2
P/BW (mW/MHz)	*8.3	*0.59	*#6	*0.026	0.02

*Normalized to a 40-nm CMOS technology. #Synthesis estimate

6.18. The gray bar represents the required energy for classifying a 10-MHz signal from a 500-MHz spectrum without band segmentation. It dissipates 1mJ. Since band segmentation estimates signal bandwidths and carrier frequencies for the blind signal classifier, the processing time of the blind signal classifier is reduced by more than 100x, resulting in a 30.9x saving in energy. The partial PSD estimation provides additional 1.9x reduction in energy. Therefore, an overall 58.7x reduction in energy is achieved. This demonstrates the value of the band segmentation for energy-efficient blind signal classification.

6.4.3 Performance Comparison

A comparison of this work with the state of the art is listed in Table 6.3. This work supports a 2.5x wider bandwidth than our first chip, and a 50x wider bandwidth than the state-of-the-art. It can sense a 500-MHz bandwidth with an SNR of 0 dB, $P_D \geq 95\%$, $P_{FA} \leq 0.5\%$, and processing time < 1 ms. Only our chips guarantee real-time reliable signal detection, since the prior work only performed power spectrum density estimation. We use power per bandwidth as the metric for a fair comparison. With 10.2 mW to sense 500 MHz, our second chip outperforms [53-56] in power dissipation per bandwidth by at least 29.5x. As compared to our first chip, we demonstrate a 20% improvement in power per bandwidth, even with penalty from flexibility, larger FFT size, and finer frequency resolution. This low power consumption come from the partial PSD estimation algorithm and the optimization framework for the reconfigurable FFT processor and the large-dynamic-range CIC filter.

6.5 Summary

In the second chip, we presented an energy-efficient band segmentation processor used for blind signal classification. The band segmentation is the key function to reduce the dissipated energy from signal classification. First, the band segmentation selects only signal-of-interest to the signal classifier that allows the classifier using lower sampling rate for signal processing. Also, the band segmentation provides estimated signal parameters that reduce the processing time from the signal classifier. However,

the band segmentation is energy consuming due to high throughput and fine frequency resolution requirement. An algorithm-architecture co-design framework is proposed to improve the energy efficiency.

In the algorithmic development, the partial PSD estimation scheme saves the energy consumption by detecting only band-of-interest and thus reducing the computational complexity. A miss-detection tolerant signal detection algorithm is proposed to reduce the processing time. The threshold is probabilistically defined, and it guarantees reliable real-time signal detection. In the architectural design, the power and area of computation-intensive reconfigurable FFT block are minimized by parallel architecture and radix factorization. The low-power feed-forward CIC architecture results from wordlength pruning and voltage scaling. This algorithm-architecture co-design yields a 70x–220x saving in energy dissipation, where 7x–45x saving comes from algorithmic optimization, and 5x–10x comes from architectural optimization.

As a proof of concept, the proposed algorithms and architectures are implemented in a standard 40-nm CMOS technology. The area is 1.28 mm², and the power with throughput of 500 MS/s is 35.4 mW and 10.2 mW for entire spectrum estimation and partial spectrum estimation, respectively. Compared to the prior state-of-the-art, our second chip dissipates 29x less power per bandwidth for a 50x wider sensing bandwidth. Also, our second chip achieves 20% lower power per bandwidth, even with penalty from flexibility, larger FFT size, and finer frequency resolution. Therefore, an energy-efficient wideband band segmentation (spectrum sensing) should leverage partial PSD estimation and miss-detection tolerant signal detection algorithms.

CHAPTER 7

Conclusions

This dissertation presents an energy-efficient digital-circuit design framework for cognitive radio wideband signal processing, while guarantees reliable real-time signal detection. The energy efficiency is improved in both algorithmic and architectural approaches. In the algorithmic development, the main challenges lies in weak-signal detection in the presence of strong adjacent-band interferers. A low-complexity multitap-windowed frequency-domain power detector, which mitigates spectral leakage without scarifying frequency resolution, is the solution for energy-efficient wideband channelization. The channel-specific sensing time and detection threshold are necessary to guarantee reliable signal detection. The sensing time and detection threshold are adapted to the measured noise power and interfering power, hereby allowing real-time signal detection. When the spectrum is scarce, instead of channelizing the entire spectrum, channelization of only a portion of spectrum further improves the energy efficiency. In the architectural design, high-throughput power/area-efficient FFT processor and the associated large-dynamic-range signal processing are the main challenges. Parallelism, radix factorization, and DFF/RF-mixed delay line are essential to minimize the power and area of a computation-intensive FFT processor. A mini floating-point

signal processing is area- and energy-efficient for large-dynamic-range datapath. The proposed algorithms and architecture is validated in simulation and real-time hardware radio testbed.

For the proof of concept, these algorithms and architecture are implemented in two chips. The first chip is a wideband spectrum sensing processor, with radio bandwidth of 200 MHz and fixed-and-known signal bandwidth of 200 kHz. The multitap-windowed frequency-domain power detector with the sensing adaptation schemes is demonstrated in this baseband processor. Compared with the prior state of the art, this chip dissipates 22x less power per bandwidth for a 20x wider radio bandwidth. The second chip is a wideband band segmentation processor, with radio bandwidth of 500 MHz and arbitrary signal bandwidth. This chip demonstrates the energy saving from partial PSD estimation with the associated signal processing techniques. An additional 6.25x saving in energy dissipation is achieved as compared to channelizing the entire spectrum. The proposed synthesis-based methodology is robust to process scaling and can quickly port across technologies. The methodology can be further refined with circuit-level customizations if that is available to the designer. It is also applicable to digital filters and general DSP architecture optimization.

7.1 Research Contributions

Major contributions:

- Development of a spectrum sensing algorithms, including multitap-windowed

frequency-domain power detector (MW-FPD), sensing-time adaptation (STA) algorithm, and detection-threshold adaptation (DTA) algorithm. Simulations with a 20-dB INR indicate a 2x improvement in detection compared to conventional power detectors. An order-of-magnitude improvement in sensing time is achieved in the presence of 30-dB interferers, while maintaining a false-alarm rate of 0.1 and a detection rate of 0.9.

- Development of partial PSD estimation algorithm, including coarse sensing, band-pass filtering, and fine sensing, to replace entire PSD estimation. The configurations of the coarse-sensing and fine-sensing FFTs is determined by minimizing the overall energy dissipation, a 6.25x saving in energy dissipation is achieved.
- Development of a methodology for power-area-minimized FFT processor by exploring parallelism, radix factorization, and delay line implementation. Our FFT processor achieves 67 1024-point FFTs per $\mu\text{J} \cdot \text{mm}^2$.
- Development of floating point signal processing for large dynamic range arithmetics. The use of floating-point arithmetic and the associated wordlength pruning reduces the overall core area and core power by 27.5% and 20.1%, respectively.

Minor contributions:

- Development of a framework to evaluate the detection performance of MW-FPD. The detection performance is characterized by the measured noise power and interfering power, allowing real-time sensing adaptation.

- Implementation of STA and DTA by using Newton-Raphson algorithms for high throughput. An overall 7x reduction in processing time is achieved.
- Development of a miss-detection tolerant signal detection algorithm for real-time signal detection. A 7x reduction in processing time is achieved.
- Implementation of a CIC decimation filter by using feed-forward architecture for low power. The uses of voltage scaling and wordlength pruning result in a 2.5x reduction in power as compared to the conventional recursive architecture.

7.2 Future Work

The research on wideband spectrum sensing is still in its infancy. This work demonstrates a methodology of an energy-efficient wideband spectrum sensing processor that allows reliably detecting a -5 -dB SNR weak signal in the presence of 30-dB INR adjacent-band interferers. However, there are still some challenges remaining: 1) improvement of sensitivity, and 2) improvement of interference cancellation, for examples. As far as the sensitivity is concerned, in the presence of shadowing, the spectrum sensing processor needs to identify signals that are as much as 20-dB below noise. When the noise uncertainty issue is taken into account, due to the SNR wall [59], applying energy detection only and purely increasing sensing time is unable to improve detection performance. As for interferers, the 30-dB INR limitation is too small for future devices. According to standards of 802.22 [6], the white space for CR in the future will operate in a variety of bands from 470 MHz to ISM bands up to 5 GHz. Traffic

from broadcast television, existing cellular, and WiFi devices will coexist in this white space, indicating that the dynamic range requirement might exceed 100 dB, so as the INR requirement. In this scenario, adaptation to interfering power only will no longer be adequate to maintain sensing performance with constrained sensing time.

To address these challenges, more sophisticated digital signal processing algorithms and the associated circuit architecture should be developed. One research direction is to build a hybrid spectrum sensing scheme, including both energy detection and feature detection to improve the sensitivity. The energy detection could be utilized as coarse sensing to detect the band of interest for low power. Then, the feature detection, which is more robust to noise uncertainty, could detect only the band of interest to achieve reliable signal detection with SNR of -20 dB. Investigation of interference cancellation algorithm is another research direction. The strong interferers could be suppressed in both analog and digital domain. A simple notch filter with reconfigurable bandwidth could be used to mitigate interferers in analog domain to improve the resolution of the analog-to-digital converters. In digital domain, more complex DSP algorithms could be developed for mitigation of the residual interferers. It is also interesting to look into the reconfiguration and the trade-off of these blocks.

REFERENCES

- [1] Federal Communications Commission “Spectrum Policy Task Force Report, ET Docket No. 02-155,” Nov. 2002.
- [2] Federal Communications Commission “Notice of Proposed Rule Making and Order, ET Docket 0.-108” Dec. 2003.
- [3] “The XG Vision. [Online]. <http://www.darpa.mil/ato/programs/XG/>,” 2008.
- [4] D. Cabric *et al.*, “Implementation issues in spectrum sensing for cognitive radios,” in *Proc. 38th Asilomar Conf. Signals, Systems, and Computers (ASILOMAR)*, vol. 1, Nov. 2004, pp. 772-776.
- [5] J. Mitola III, “Cognitive radio an integrated agent architecture for software defined radio,” *PHD thesis, KTH Royal Institute of Technology, Stockholm, Sweden,*, 2000.
- [6] “[Online]. <http://www.ieee802.org/22/>, IEEE 802.22 WRAN WG website.”
- [7] F. F. Digham *et al.*, “On the Energy Detection of Unknown Signals Over Fading Channels,” *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 21-24, Jan. 2007.
- [8] A. Taherpour *et al.*, “Wideband spectrum sensing in unknown white Gaussian noise,” *IET Commun.*, vol. 2, no. 6, pp. 763-771, Jul. 2008.
- [9] H. Sun *et al.*, “Wideband spectrum sensing for cognitive radio networks: a survey,” *IEEE Wireless Commun.*, vol. 20, no. 2, pp. 74-81, Apr. 2013.
- [10] H. Sun *et al.*, “How much time is needed for wideband spectrum sensing?,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 11, pp. 5466-5471, Nov. 2009.
- [11] K. Hossain *et al.*, “Wideband Spectrum Sensing for Cognitive Radios With Correlated Subband Occupancy,” *IEEE Signal Process. Lett.*, vol. 18, no. 1, pp. 35-38, Jan. 2011.
- [12] S. Rodriguez-Parera *et al.*, “Spectrum Sensing over SIMO Multi-Path Fading Channels Based on Energy Detection,” in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov. 2008, pp. 1-6.
- [13] P. Paysarvi-Hoseini *et al.*, “Optimal Wideband Spectrum Sensing Framework for Cognitive Radio Systems,” *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1170-1182, Mar. 2011.
- [14] D. Cabric *et al.*, “Physical Layer Design Issues Unique to Cognitive Radio Systems,” in *Proc. IEEE 16th Int. Symp. Personal, Indoor and Mobile Radio Commun. (PIMRC)*, vol. 2, Sep. 2005, pp. 759-763.

- [15] D. Cabric *et al.*, "Spectrum Sharing Radio," *IEEE Circuits Syst. Mag.*, vol. 6, no. 2, pp. 30-45, Jul. 2006.
- [16] T. Yucek *et al.*, "A Survey of Spectrum Sensing Algorithms for Cognitive Radio Applications," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 1, pp. 116-130, First Quarter 2009.
- [17] R. Mahesh *et al.*, "A Low-Complexity Flexible Spectrum-Sensing Scheme for Mobile Cognitive Radio Terminals," *IEEE Trans. Circuits Syst. (II)*, vol. 58, no. 6, pp. 371-375, Jun. 2011.
- [18] M. Kitsunezuka *et al.*, "A 30-MHz2.4-GHz CMOS Receiver With Integrated RF Filter and Dynamic-Range-Scalable Energy Detector for Cognitive Radio Systems," *IEEE J. Solid-State Circuits*, vol. 47, no. 5, pp. 1084-1093, May 2012.
- [19] T. Rapinoja, "A Digital Frequency Synthesizer for Cognitive Radio Spectrum Sensing Applications," *IEEE Trans. Microw. Theory Tech.*, vol. 58, no. 5, pp. 1339-1348, May 2010.
- [20] B. Razavi, "Cognitive Radio Design Challenges and Techniques," *IEEE J. Solid-State Circuits*, vol. 45, no. 8, pp. 1542-1553, Aug. 2010.
- [21] W. A. Gardner *et al.*, "Characterization of cyclostationary random signal processes," *IEEE Trans. Inf. Theory*, vol. 21, no. 1, pp. 4-14, Jan. 1975.
- [22] W. A. Gardner, "Signal interception: a unifying theoretical framework for feature detection," *IEEE Trans. Commun.*, vol. 36, no. 8, pp. 897-906, Aug. 1988.
- [23] Z. Tian *et al.*, "Cyclic Feature Detection With Sub-Nyquist Sampling for Wideband Spectrum Sensing," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 1, pp. 58-69, Feb. 2012.
- [24] M. Bkassiny *et al.*, "Wideband Spectrum Sensing and Non-Parametric Signal Classification for Autonomous Self-Learning Cognitive Radios," *IEEE Trans. Wireless Commun.*, vol. 11, no. 7, pp. 2596-2605, Jul. 2012.
- [25] W.-L. Chin *et al.*, "Features detection assisted spectrum sensing in wireless regional area network cognitive radio systems," *IET Commun.*, vol. 6, no. 8, pp. 810-818, May 2012.
- [26] M. Mishali *et al.*, "Wideband Spectrum Sensing at Sub-Nyquist Rates [Applications Corner]," *IEEE Signal Process. Mag.*, vol. 28, no. 4, pp. 102-135, Jul. 2011.
- [27] H. Sun *et al.*, "Wideband Spectrum Sensing With Sub-Nyquist Sampling in Cognitive Radios," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 6068-6073, Nov. 2012.

- [28] Y. Wang *et al.*, “Sparsity Order Estimation and its Application in Compressive Spectrum Sensing for Cognitive Radios,” *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2116-2125, Jun. 2012.
- [29] H. Sun *et al.*, “Adaptive Compressive Spectrum Sensing for Wideband Cognitive Radios,” *IEEE Commun. Lett.*, vol. 16, no. 11, pp. 1812-1815, Nov. 2012.
- [30] F. Zeng *et al.*, “Distributed Compressive Spectrum Sensing in Cooperative Multihop Cognitive Networks,” *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 1, pp. 37-48, Feb. 2011.
- [31] L. Bai *et al.*, “Compressive Spectrum Sensing Using a Bandpass Sampling Architecture,” *IEEE Emerging Sel. Topics Circuits Syst.*, vol. 2, no. 3, pp. 433-442, Sep. 2012.
- [32] D.D. Ariananda *et al.*, “Compressive Wideband Power Spectrum Estimation,” *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4775-4789, Sep. 2012.
- [33] Z. Zhang *et al.*, “Belief Propagation Based Cooperative Compressed Spectrum Sensing in Wideband Cognitive Radio Networks,” *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 3020-3031, Sep. 2012.
- [34] M. Derakhtian *et al.*, “Cooperative wideband spectrum sensing for cognitive radio networks in fading channels,” *IET Signal Process.*, vol. 6, no. 3, pp. 227-238, May. 2012.
- [35] Z. Quan *et al.*, “Collaborative Wideband Sensing for Cognitive Radio,” *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 60-73, Nov. 2008.
- [36] Z. Quan *et al.*, “Optimal Multiband Joint Detection for Spectrum Sensing in Cognitive Radio Networks,” *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 1128-1140, Mar. 2009.
- [37] Z. Quan *et al.*, “Wideband Spectrum Sensing in Cognitive Radio Networks,” in *IEEE Proc. Int. Communications Conf. (ICC)*, May 2008, pp. 901-906
- [38] J. G. Proakis *et al.*, “Digital Signal Processing,” *Pearson Prentice Hall*, 2007.
- [39] Z. Tian *et al.*, “A Wavelet Approach to Wideband Spectrum Sensing for Cognitive Radios,” in *Proc. Int. Conf. Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, Jun. 2006, pp. 1-5.
- [40] T.-W. Chiang *et al.*, “Optimal Detector for Multitaper Spectrum Estimator in Cognitive Radios,” in *Proc. IEEE Global Telecommunications Conf. (GLOBECOM)*, Nov. 2009, pp. 1-6.
- [41] C.K. Tan *et al.*, “Reliable and low-complexity wavelet-based spectrum sensing for cognitive radio systems at low SNR regimes,” *IEEE Electron. Lett.*, vol. 48, no. 24, pp. 1565-1567, Nov. 2012.

- [42] “Improved wideband spectrum sensing techniques using wavelet-based edge detection for cognitive radio,” in *Proc. Int. Conf. Networking and Commun. (ICNC)*, Jan. 2013, pp. 418-423.
- [43] S. Chantaraskul *et al.*, “Implementation of wavelet analysis for spectrum opportunity detection,” in *Proc. IEEE 20th Int. Symp. Personal, Indoor and Mobile Radio Commun. (PIMRC)*, Sep. 2009, pp. 2310-2314.
- [44] D. Liu *et al.*, “A Novel Signal Separation Algorithm for Wideband Spectrum Sensing in Cognitive Networks,” in *Proc. IEEE Global Telecommunications Conf. (GLOBECOM)*, Dec. 2010, pp. 1-6.
- [45] B. Farhang-Boroujeny, “Filter Bank Spectrum Sensing for Cognitive Radios,” *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1801-1811, May 2008.
- [46] N. Pillay, “Blind eigenvalue-based spectrum sensing for cognitive radio networks,” *IET Commun.*, vol. 6, no. 11, pp. 1388-1396, Nov. 12.
- [47] Y. Zeng, “Eigenvalue-based spectrum sensing algorithms for cognitive radio,” *IEEE Tran. Commun.*, vol. 57, no. 6, pp. 1784-1793, Jun. 09.
- [48] A. Taherpour, “Multiple antenna spectrum sensing in cognitive radios,” *IEEE Tran. Wireless Commun.*, vol. 9, no. 2, pp. 814-823, Feb. 10.
- [49] F. Sheikh *et al.*, “Cognitive Spectrum Sensing and Detection Using Polyphase DFT Filter Banks,” in *Proc. Consumer Communications and Networking Conf. (CCNC)*, Jan. 2008, pp. 973-977.
- [50] F. Sheikh, “Harmonic power detection in wideband cognitive radios,” *IET Signal Process.*, vol. 3, no. 1, pp. 40-50, Jan. 09.
- [51] J.J. Lehtomaki, “CFAR Outlier Detection With Forward Methods,” *IEEE Trans. Signal Process.*, vol. 55, no. 9, pp. 4702-4706, Sep. 2007.
- [52] D. Kun *et al.*, “A New Low-Cost CFAR Detector for Spectrum Sensing with Cognitive Radio Systems,” in *Proc. IEEE Aerospace Conf.*, Mar. 2009, pp. 1-8.
- [53] J. Park *et al.*, “A Fully-Integrated UHF Receiver with Multi-Resolution Spectrum-Sensing (MRSS) Functionality for IEEE 802.22 Cognitive-Radio Applications,” in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, Feb. 2008, pp. 526-633.
- [54] J. Park *et al.*, “A Fully Integrated UHF-Band CMOS Receiver With Multi-Resolution Spectrum Sensing (MRSS) Functionality for IEEE 802.22 Cognitive Radio Applications,” *IEEE J. Solid-State Circuits*, vol. 44, no. 1, pp. 258-268, Jan. 2009.
- [55] T. Song *et al.*, “A 122-mW Low-Power Multiresolution Spectrum-Sensing IC With Self-Deactivated Partial Swing Techniques,” *IEEE Trans. Circuits Syst. (II)*, vol. 57, no. 3, pp. 188-192, Mar. 2010.

- [56] S. Izumi *et al.*, “A Low-Power Multi Resolution Spectrum Sensing (MRSS) Architecture for a Wireless Sensor Network with Cognitive Radio,” in *Proc. IEEE 4th Int. Conf. Sensor Technologies and Applications (SENSORCOMM)*, Jul. 2010, pp. 39-44.
- [57] S. Pollin *et al.*, “An Integrated Reconfigurable Engine for Multi-purpose Sensing up to 6 GHz,” in *IEEE Proc. Int. Symp. Dynamic Spectrum Access Networks (DySPAN)*, May. 2011, pp. 656-657.
- [58] A. Verma *et al.*, “A 10-Bit 500-MS/s 55-mW CMOS ADC,” *IEEE J. Solid-State Circuits*, vol. 44, no. 11, pp. 3039-3050, Nov. 2009.
- [59] R. Tandra *et al.*, “SNR Walls for Signal Detection,” in *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 4-17, Feb. 2008.
- [60] D. Hamza *et al.*, “Wideband Spectrum Sensing Order for Cognitive Radios with Sensing Errors and Channel SNR Probing Uncertainty,” *IEEE Wireless Commun. Lett.*, vol. 2, no. 2, pp. 151-154, Apr. 2013.
- [61] L. Shen *et al.*, “Blind Spectrum Sensing for Cognitive Radio Channels with Noise Uncertainty,” *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1721-1724, Jun. 2011.
- [62] M.S.O. Alink *et al.*, “Lowering the SNR Wall for Energy Detection Using Cross-Correlation,” *IEEE Trans. Veh Commun.*, vol. 60, no. 8, pp. 3748-3757, Oct. 2011.
- [63] P. De *et al.*, “Blind Spectrum Sensing Algorithms for Cognitive Radio Networks,” *IEEE Trans. Veh Commun.*, vol. 57, no. 5, pp. 2834-2842, Sep. 2008.
- [64] A. Mariani *et al.*, “Effects of Noise Power Estimation on Energy Detection for Cognitive Radio Applications,” *IEEE Trans. Commun.*, vol. 59, no. 12, pp. 3410-3420, Dec. 2011.
- [65] Y.-C. Liang *et al.*, “Sensing-Throughput Tradeoff for Cognitive Radio Networks,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326-1337, Apr. 2008.
- [66] “Wireless Microphone in the 25 MHz to 3 GHz Frequency Range,” in *ETSI EN 300 422-1 v1.3.2*, pp. 22-24, Mar. 2008.
- [67] T.-H. Yu *et al.*, “Cognitive Radio Wideband Spectrum Sensing Using Multitap Windowing and Power Detection with Threshold Adaptation,” in *IEEE Proc. Int. Communications Conf. (ICC)*, May 2010, pp. 1-6.
- [68] S. R. Searle, “Linear models,” *Wiley*, 1971.
- [69] T. S. Ferguson, “A Course in Large Sample Theory,” *CRC Press*, 1996.
- [70] D. Markovic *et al.*, “Power and Area Minimization for Multidimensional Signal Processing,” *IEEE J. Solid-State Circuits*, vol. 42, no. 4, pp. 922-934, Apr. 2007.

- [71] C. Chang *et al.*, "BEE2: A High-End Reconfigurable Computing System," *IEEE Des. Test. Comput.*, vol. 22, no. 2, pp. 114-125, Mar. 2005.
- [72] C.-H. Yang *et al.*, "A Flexible DSP Architecture for MIMO Sphere Decoding," *IEEE Trans. Circuits Syst. (I)*, vol. 56, no. 10, pp. 2301-2314, Oct. 2009.
- [73] W.R. Davis *et al.*, "A Design Environment for High Throughput, Low Power Dedicated Signal Processing Systems," *IEEE J. Solid-State Circuits*, vol. 37, no. 3, pp. 420-431, Mar. 2002.
- [74] S.M. Mishra *et al.*, "A Real Time Cognitive Radio Testbed for Physical and Link Layer Experiments," in *IEEE Proc. 1st Int. Symp. Dynamic Spectrum Access Networks (DySPAN)*, Nov 2005, pp. 562-567.
- [75] S. Magar *et al.*, "An Application Specific DSP Chip Set for 100 MHz Data Rates," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 4, Apr 1988, pp. 1989-1992.
- [76] S. He *et al.*, "Design and Implementation of a 1024-Point Pipeline FFT Processor," in *Proc. IEEE Custom Integrated Circuits Conf. (CICC)*, May 1998, pp. 131-134.
- [77] J. OBrien *et al.*, "A 200 MIPS Single-Chip 1k FFT Processor," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, Feb. 1989, pp. 166-167.
- [78] B. M. Bass, "A Low-Power High-Performance, 1024-Point FFT Processor," *IEEE J. Solid-State Circuits*, vol. 34, no. 3, pp. 380-387, Mar. 1999.
- [79] A. Wang *et al.*, "A 180-mV Subthreshold FFT Processor Using a Minimum Energy Design Methodology," *IEEE J. Solid-State Circuits*, vol. 40, no. 1, pp. 310-319, Jan. 2005.
- [80] M. Seok *et al.*, "A 0.27V 30MHz 17.7nJ/transform 1024-pt Complex FFT Core with Super-pipelining," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, Feb. 2011, pp. 342-343.
- [81] Y. Chen *et al.*, "A 2.4-Gsample/s DVFS FFT Processor for MIMO OFDM Communication Systems," *IEEE J. Solid-State Circuits*, vol. 43, no. 5, pp. 1260-1273, May. 2008.
- [82] K.-S. Chong *et al.*, "Energy-Efficient Synchronous-Logic and Asynchronous-Logic FFT/IFFT Processors," *IEEE J. Solid-State Circuits*, vol. 42, no. 9, pp. 2034-2045, Sep. 2007.
- [83] G. Zhong *et al.*, "A Power-Scalable Reconfigurable FFT/IFFT IC Based on Multi-Processor Ring," *IEEE J. Solid-State Circuits*, vol. 41, no. 2, pp. 483-495, Feb. 2006.

- [84] Y.-W. Lin *et al.*, "A 1-GS/s FFT/IFFT Processor for UWB Applications," *IEEE J. Solid-State Circuits*, vol. 40, no. 8, pp. 1726-1753, Aug. 2005.
- [85] Y. Chen *et al.*, "An indexed-scaling pipelined FFT processor for OFDM-based WPAN applications," *IEEE Trans. Circuits Syst. (II)*, vol. 55, no. 2, pp. 146-150, Feb. 2008.
- [86] J. G. Proakis *et al.*, "Digital Signal Processing: Principles, Algorithms, and Applications," *Prentice Hall, 3rd ed.*, 1996.
- [87] T.-D. Chiueh *et al.*, "OFDM Baseband Receiver Design for Wireless Communications," *Wiley*, 2007.
- [88] Y.-T. Lin *et al.*, "Low-power variable-length fast Fourier transform processor," in *IEE Proc. Comput. Digit. Tech.*, vol. 152, no. 4, Jul. 2005, 499-506.
- [89] A. P. Chandrakasan *et al.*, "Low-power CMOS digital design," *IEEE J. Solid-State Circuits*, vol. 27, no. 4, pp. 473-484, Apr. 1992.
- [90] K. K. Parhi, "VLSI Digital Signal Processing Systems," *Wiley*, 1999.
- [91] S. W. Reitwiesner, "Binary arithmetic," *Advances in Computers*, pp. 231-308, 1966.
- [92] A. Wenzler *et al.*, "New structures for complex multipliers and their noise analysis," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, vol. 2, May 1995, 1432-1435.
- [93] K.-H. Lin, "Design of a Baseband Receiver for DVB-T Standard," *MS thesis, National Taiwan University, Taiwan*, Sep. 2004.
- [94] T.-H. Yu *et al.*, "A Wideband Spectrum-Sensing Processor with Adaptive Detection Threshold and Sensing Time," *IEEE Trans. Circuits Syst. (I)*, vol. 58, no. 11, pp. 2765-2775, Nov. 2011.
- [95] C.-H. Yang *et al.*, "Power and Area Minimization of Reconfigurable FFT Processors: A 3GPP-LTE Example," *IEEE J. Solid-State Circuits*, vol. 47, no. 3, pp. 757-768, Mar. 2012.
- [96] T.-H. Yu *et al.*, "A 7.4mW 200MS/s Wideband Spectrum Sensing Digital Baseband Processor for Cognitive Radios," in *Proc. Int. Symp. VLSI Circuits (VLSI)*, Jun. 2011, pp. 254-255.
- [97] C. Shi *et al.*, "Automated Fixed-point Data-type Optimization Tool for Signal Processing and Communication Systems," in *Proc. Design Automation Conf.*, Jun. 2004, pp. 478-483.

- [98] K. R. Ralev *et al.*, “Realization of Block Floating-point Digital Filters and Application to Block Implementations,” *IEEE Trans. Signal Process.*, vol. 47, no. 4, pp. 1076-1086, Apr. 1999.
- [99] A. Mitra *et al.*, “A Block Floating-point Treatment to the LMS Algorithm: Efficient Realization and a Roundoff Error Analysis,” *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4536-4544, Dec. 2005.
- [100] M. Chakraborty *et al.*, “A Block-Floating-Point-Based Realization of the Block LMS Algorithm,” *IEEE Trans. Circuits Syst. (II)*, vol. 53, no. 9, pp. 812-816, Sep. 2006.
- [101] M. D. Ercegovic *et al.*, “Digital Arithmetic,” *Morgan Kaufmann Publishers*, 2004.
- [102] H.R. Srinivas *et al.*, “A fast radix-4 division algorithm and its architecture,” *IEEE Trans. Comput.*, vol. 44, no. 6, pp. 826-831, Jun. 1995.
- [103] A.E. Bashagha *et al.*, “Digit serial division algorithm,” *IEEE Electron. Lett.*, vol. 31, no. 12, pp. 985-959, Jun. 1995.
- [104] S.F. Oberman *et al.*, “Division algorithms and implementations,” *IEEE Trans. Comput.*, vol. 46, no. 8, pp. 833-854, Aug. 1997.
- [105] Y. H. Hu, “CORDIC-based VLSI Architectures for Digital Signal Processing,” *IEEE Signal Process. Mag.*, vol. 9, no. 3, pp. 16-35, Mar. 1992.
- [106] J. Duprat *et al.*, “The CORDIC algorithm: new results for fast VLSI implementation,” *IEEE Trans. Comput.*, vol. 42, no. 2, pp. 168-178, Feb. 1993.
- [107] P.K. Meher *et al.*, “50 Years of CORDIC: Algorithms, Architectures, and Applications,” *IEEE Trans. Circuits Syst. (I)*, vol. 56, no. 9, pp. 1893-1907, Sep. 2009.
- [108] “ROACH: Reconfigurable Open Architecture Computing Hardware. [Online]. <http://casper.ssl.berkeley.edu/wiki/ROACH>”
- [109] T.-H. Yu *et al.*, “An Energy-Efficient VLSI architecture for Cognitive Radio Wideband Spectrum Sensing,” in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec. 2011, pp. 1-6.
- [110] E. Rebeiz *et al.*, “Energy-Efficient Processor for Blind Signal Classification in Cognitive Radio Networks,” in *submission to IEEE Trans. on Circuits and Syst. (I)*.
- [111] T.-H. Yu *et al.*, “A 7.4 mW 200 MS/s Wideband Spectrum Sensing Digital Baseband Processor for Cognitive Radios,” *IEEE J. Solid-State Circuits*, vol. 47, no. 9, pp. 2235-2245, Sep. 2012.

- [112] Y. Khoo *et al.*, “Efficient High-Speed CIC Decimation Filter,” in *Proc. IEEE Int. ASIC Conf.*, Sep. 1998, pp. 251-254.
- [113] D. Babic *et al.*, “Power efficient structure for conversion between arbitrary sampling rates,” in *IEEE Signal Process. Lett.*, vol. 12, no. 1, pp. 1-4, Jan. 2005.
- [114] C. Riley *et al.*, “High-decimation digital filters,” in *Proc. Int. Conf. Acoustics, Speech, and Signal Process. (ICASSP)*, vol. 3, Apr. 1991, pp. 1613-1616.
- [115] R. Nanda *et al.*, “A Low-Power Digital Front-end Direct-sampling Receiver for Flexible Radios,” in *Proc. Asian Solid-State Circuits Conf. (ASSCC)*, Nov. 2011, pp. 377-380.