A 7.4-mW 200-MS/s Wideband Spectrum Sensing Digital Baseband Processor for Cognitive Radios

Tsung-Han Yu, *Student Member, IEEE*, Chia-Hsiang Yang, *Member, IEEE*, Danijela Čabrić, *Member, IEEE*, and Dejan Marković, *Member, IEEE*

***Abstract*—A digital baseband cognitive radio spectrum sensing processor with 200-kHz resolution over 200-MHz bandwidth is integrated in 1.64 mm^2 in 65-nm CMOS. The processor uses adaptive channel-specific threshold and sensing time to achieve detection probability $\geq$0.9 and false-alarm probability $\leq$0.1 for $-$5-dB SNR and adjacent-band interferers of 30-dB INR within a 50-ms sensing time. The chip power and area are minimized by jointly considering algorithm, architecture, and circuit parameters. The chip dissipates 7.4 mW for a 200-MHz sensing bandwidth, which is a 22$\times$ reduction in power per sensing bandwidth compared with prior work.**

***Index Terms*—CMOS digital integrated circuits, cognitive radio (CR), power and area minimization, wideband spectrum sensing.**

I. Introduction

COGNITIVE RADIO (CR) is an emerging technology for improving spectral utilization by taking advantage of temporally and spatially unused spectrum. The enabling technology for CR is spectrum sensing. The unused spectrum is sensed by detecting the presence of primary users in the band of interest to avoid interference. The key requirement for spectrum sensing is reliable signal detection, with high sensitivity within a constrained sensing time. Wideband (>100 MHz) sensing is a highly desirable feature for CR systems, since it allows for sensing multiple channels at a time and increases the probability of finding available spectrum.

Silicon realizations of spectrum sensing have been demonstrated in [1]–[5] for single narrow-bandwidth signals. The work in [1]–[3] use analog correlators for the IEEE 802.22 standard with 1- and 10-MHz sensing bandwidths with frequency resolution of 100 kHz. The work in [4] proposes a multiresolution spectrum-sensing processor with a reconfigurable filter for resolution from 12.5 to 400 kHz. In [1]–[4], only power spectrum density (PSD) estimation is performed, while the sensing time and detection threshold are determined offline. In [5], the authors proposed a reconfigurable engine for multipurpose sensing up to 6 GHz, but the instantaneous sensing bandwidth is only 20 MHz.

Apart from narrow sensing bandwidth, the main drawback of existing systems is that the sensing time and detection threshold cannot be adapted dynamically when strong adjacent-band primary users are present. Since sharp filtering is infeasible in wideband sensing, the strong adjacent-band primary users introduce spectral leakage to vacant bands. Without channel-specific sensing time and detection threshold, the strong adjacent-band primary users lead to degradation in spectrum sensing performance. Therefore, high-resolution wideband spectrum sensing in the presence of strong adjacent-band primary users is an open challenge.

This paper presents 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 [6]. The validity of the sensing algorithm has been verified on a real-time CR hardware testbed [6] prior to chip implementation of the spectrum sensing processor [7]. The chip, implemented in a 65-nm CMOS technology, outperforms the state of the art [1]–[4] in detection performance and power dissipation per bandwidth. Its high performance allows fast, reliable spectrum sensing for future cognitive radios. The low power dissipation of 7.4 mW makes it applicable to portable devices.

This paper is organized as follows. Section II reviews wideband spectrum sensing and introduces system specifications and proposed sensing algorithms. The hardware realization challenges, the corresponding architecture, and circuit solutions are presented in Sections III and IV. Chip measurements are discussed in Section V. Section VI concludes the paper.

II. Wideband Spectrum Sensing

A wideband signal composed of several nonoverlapping narrowband primary users is considered in our system model. Additive white Gaussian noise (AWGN) applies uniformly across the band. Frequency-domain power detection (FPD) is adopted as the sensing algorithm for energy-efficient realization [6]. In FPD, the FFT is used to channelize the entire spectrum, where the model for each channel is similar to that in narrowband signal detection [8]. 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.

FPD involves PSD estimation of the entire band. For reliable signal detection in the negative SNR regime, the power detector

Manuscript received November 01, 2011; revised March 25, 2012; accepted April 03, 2012. Date of publication May 30, 2012; date of current version August 21, 2012. This work was supported by the DARPA HEALICs program and UCLA Graduate Fellowship. This paper was approved by Associate Editor Stefan Rusu.

T.-H. Yu, D. Čabrić, and D. Marković are with the Electrical Engineering Department, University of California, Los Angeles, CA 94720 USA (e-mail: thyu0918@ee.ucla.edu).

C.-H. Yang is with the Electronics Engineering Department, National Chiao Tung University, 300 Hsinchu, Taiwan.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSSC.2012.2195933

TABLE I
WIDEBAND CR SYSTEM SPECIFICATIONS

Signal bandwidth	200 MHz
Frequency resolution	200 kHz
Signal-to-noise ratio	SNR ≥ -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 ≤ 30 dB

requires an adequate number of samples to obtain accurate PSD estimation. Then, the variance of the estimated PSD can be reduced by averaging (accumulation). Finally, the estimated PSD is compared with a detection threshold to decide between $H_0(k)$ and $H_1(k)$. The decision rule is represented by

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

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

$$T_{\text{sensing}} = M \times N \times 1/F_s. \tag{2}$$

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, which measures the utilization of unused spectrum, and P_D is the probability that a system detects the presence of a primary user and measures the rate of avoiding interference.

The main limitation of power detection is the SNR wall [9], especially in negative SNR regime. The SNR wall constrains the minimum SNR, below which the system fails to be robust, no matter how long it is allowed to observe the channel. The SNR wall is due to the uncertainty in the noise and fading processes. The more accurate estimation of noise and fading distribution allows detecting weaker signals. The noise uncertainty can be reduced by increasing the number of samples for noise power calibration [9], assuming that only thermal noise exists and the temperature holds constant during the sensing period. According to [9], assuming noise power calibration constrains the noise uncertainty under 0.1 dB, the system is able to maintain reliable detection of weak signals with -5-dB SNR.

Based on the above definitions, we can formulate physical-layer design specifications. The specifications consist of signal bandwidth, frequency resolution, minimum SNR, sensing time, P_D, and P_{FA}, and maximum tolerable interferer-to-noise ratio (INR) as summarized in Table I. The details of these specifications can be found in [6]. The signal bandwidth dictates the minimum sampling frequency of the ADC and 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 minimum SNR is constrained by the noise uncertainty for reliable detection specified by P_D and P_{FA}. The sensing time trades off the reliability of signal detection and the system throughput. INR is the power ratio of strong primary user to band-of-interest weak signal. The maximum tolerable INR is constrained by the primary user transmitter frequency mask.

A. Proposed Sensing Algorithms

When strong primary users are present in adjacent bands, the in-band interfering power is introduced due to the imperfect in-band filtering of the FFT. The effective SNR is degraded by this in-band interfering power, which needs to be mitigated to reduce the required sensing time. The interfering power has to be estimated to enable the adjustment in the sensing time and 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.

The sensing procedure starts by turning off the RF antenna to perform noise power calibration. A processing time of 0.5 ms is required to achieve noise uncertainty of 0.1 dB for -5-dB SNR. Next, coarse sensing is performed for the power estimation of adjacent-band interferers. After this step, in-band interfering power is estimated, based on which sensing time is adapted for each channel. The in-band interfering power is obtained from the estimated power of the adjacent-band interferers [6]. Now, the residual PSD estimation is performed. In this step, a channel-specific sensing time is set to enhance the system throughput and reduce power consumption. The final stage does detection-threshold adaptation. The threshold for each channel is adapted to the estimated interfering power and the corresponding number of averages, i.e., sensing time. The decision about the presence of primary users is made right after the threshold is determined. The key algorithms in this procedure are multitap windowed frequency power detector (MW-FPD) for PSD estimation and adaptive-threshold and sensing-time estimation.

In our proposed frequency-domain power detection, we apply a multitap window to overcome the spectral leakage problem of the FFT [6]. 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 larger (2048-point) FFTs to channelize the time-domain samples. The proposed MW-FPD overlaps and adds the samples in time domain and channelizes 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. The longer time-domain window simultaneously reduces the interfering power and compensates the degradation in frequency resolution.

Necessary algorithmic features will be presented next in order to explain implementation challenges. The detection rule of MW-FPD can be formulated by (1), and $\hat{X}_m(k)$ is given by

$$\hat{X}_m(k) = \sum_{n=0}^{N-1} \left(\sum_{p=0}^{P-1} w[n+pN] x[n+pN+mN] \right) e^{-j\frac{2\pi nk}{N}} \tag{3}$$

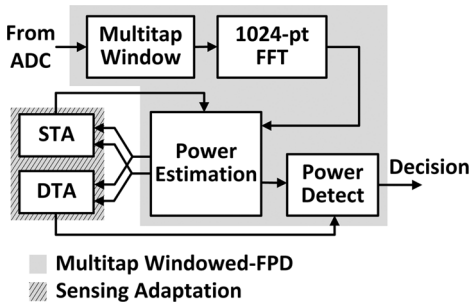


Fig. 1. Proposed wideband spectrum sensing processor. An FFT-based MW-FDP estimates PSD, noise, and interfering power. STA and DTA blocks adapt their parameters on a per-channel basis.

where p is the tap index P is the number of taps of the multitap window, and $w[n]$ is the normalized window coefficient. The channel specific sensing time $M(k)$ and detection threshold $\gamma(k)$ are modeled by

$$M(k) = \alpha \left((1+\psi(k)) \frac{Q^{-1}(P_{\rm FA}) - Q^{-1}(P_D)}{\rm SNR} - Q^{-1}(P_D) \right)^2 \tag{4}$$

$$\gamma(k) = \left(Q^{-1}(P_{\rm FA}) \sqrt{\alpha \cdot M(k)} + M(k) \right) \cdot \Sigma_0(k) \tag{5}$$

where α is the fitting factor determined by the modulation scheme of the primary user and the time-domain multitap window. $\psi(k) = \sigma_{if}^2(k)/\sigma_{nf}^2(k)$ and $\Sigma_0(k) = \sigma_{nf}^2(k) + \sigma_{if}^2(k)$, where $\sigma_{nf}^2(k)$ and $\sigma_{if}^2(k)$ are the in-band noise power and interfering power, respectively. Algorithmic details can be found in [6]. By adapting the detection threshold and sensing time to the adjacent-band interfering power, as described, the proposed wideband spectrum-sensing processor is able to achieve the constrained sensing rates $P_{\rm FA}$ and P_D with minimum sensing time.

III. Design Challenges of Wideband Spectrum Sensing

Fig. 1 shows the block diagram of the proposed spectrum-sensing processor. It consists of a MW-FDP and two sensing adaptation blocks, sensing-time adaptation (STA) and detection-threshold adaptation (DTA). 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 P_D and $P_{\rm FA}$ 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 interferes 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.

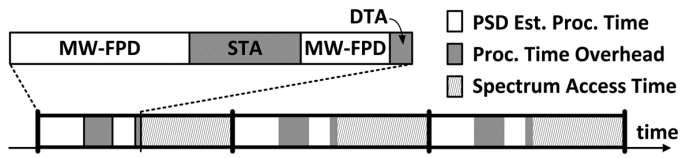


Fig. 2. Processing time for spectrum sensing consists of fixed time for MW-FPD calculation and variable time for STA and DTA blocks. The MW-FPD determines the sensing rate performance. The overhead introduced by the STA and DTA blocks degrades the sensing-time efficiency and has to be minimized.

The adaptation blocks, STA and DTA, dynamically track the channel condition to maintain system performance. As shown in Fig. 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 $P_{\rm FA}$), and this processing time is fixed. The processing time of STA and DTA blocks does not influence the sensing rates. These blocks affect the sensing time; hence, their processing time overhead has to be minimized.

IV. Architecture and Circuit Design

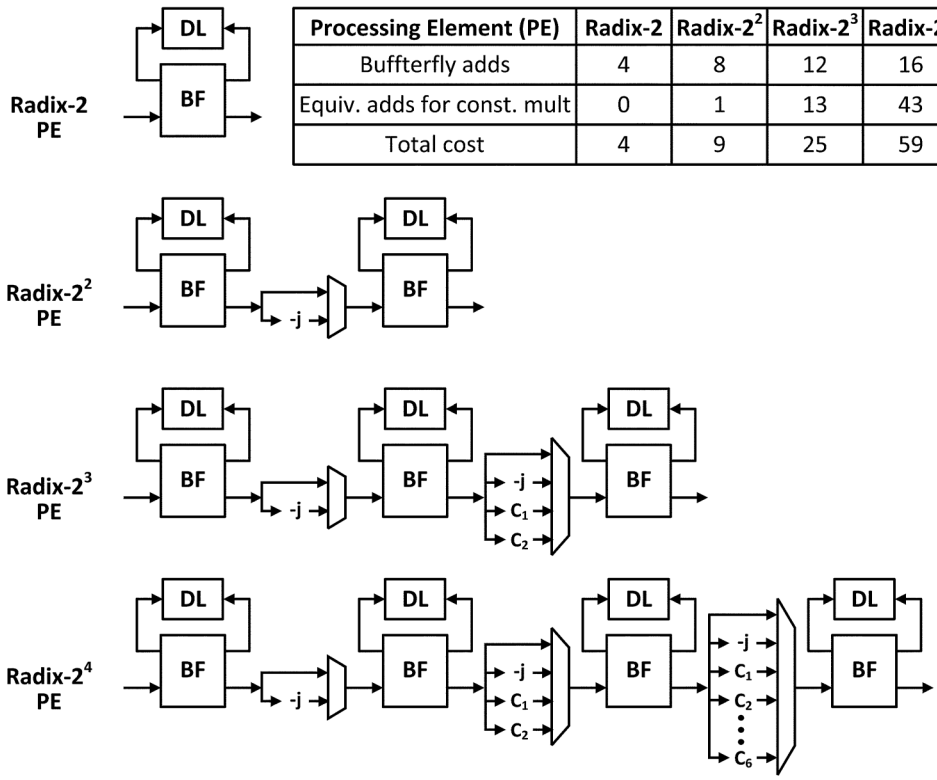
Several architectural and signal processing techniques are proposed to minimize power and area. The FFT processor uses parallelism and radix factorization to achieve minimum power–area product (PAP). A fixed-to-floating point converter is designed to minimize wordlengths of the mantissa and exponent, 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.

A. Multi-Path Pipelined MW-FPD

A multipath 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 [10] is used to determine the optimal radix structure. The 1024-point FFT is decomposed into eight banks of 128-point pipelined FFTs followed by an eight-point parallel FFT to minimize PAP.

The area for different levels of parallelism is estimated from hardware synthesis. Delay of a fanout-of-4 (FO4) inverter is used to estimate the critical-path delay as a function of supply voltage. The delay–voltage curve is 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 that includes both switching and leakage components. The eight-path single-delay-feedback (SDF) architecture operated at 0.6 V is chosen as the optimal design, resulting in a $4\times$ PAP reduction compared to the reference design with no parallelism.

The optimal radix factorization is selected from all feasible combinations of radix-$2/2^2/2^3/2^4$ processing elements (PEs) shown in Fig. 3. Canonic signed digit (CSD) representation



Processing Element (PE)	Radix-2	Radix-2^2	Radix-2^3	Radix-2^4
Buffterfly adds	4	8	12	16
Equiv. adds for const. mult	0	1	13	43
Total cost	4	9	25	59

Fig. 3. FFT building blocks. All intra-stage multipliers for radix $2/2^2/2^3/2^4$ are realized as constant multipliers. The complexity of PEs in terms of the number of real-valued adders for various radices is indicated in the table.

TABLE II
HARDWARE COMPLEXITY ANALYSIS FOR 128-POINT FFT

Architecture	Radix-2	Radix-2^2	Radix-2^3	Radix-2^4	Area cost
A1	7				238
A2	5	1			204
A3	4		1		181
A4	3	2			170
A5	3			1	176
A6	2	1	1		147
A7	1	3			136
A8	1		2		124
A9	1	1		1	142
A10*		**2**	**1**		**113**
A11			1	1	119

* final design

[13] is used to implement constant intra-stage multipliers. The table in Fig. 3 shows the normalized hardware complexity of the PEs in terms of the number of real-valued adders. To quantify the impact of radix changes on power and area, hardware estimates of various radix factorizations are explored as shown in Table II. The area cost is evaluated by the total number of real-valued adders that include adders for butterfly operation, constant-factor intra-stage multipliers and full-precision inter-stage multipliers [10]. A $4.2\times$ PAP reduction is achieved from radix-2 (A1) to the final design radix-$2^2/2^2/2^3$ (A10). An overall $17.2\times$ PAP reduction is achieved through FFT decomposition (parallelism) and radix factorization.

The delay lines are designed for minimum PAP by exploiting the power-area tradeoff between register files (RFs) and D flip-flops (DFFs) [7]. RFs (6 T/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; otherwise, 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 mm^2 (20%) with power increased by 0.3 mW (4%) as compared with 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 mm^2, respectively. Our chip measurements at 200 MS/s were consistent with synthesis estimates. Next, we resynthesize for 240 MS/s and compare our mix-radix design with the 240-MS/s 1024-point radix-4 FFT [14] that was designed for minimum energy using circuit-level techniques for subthreshold operation. The comparison in Fig. 4 shows that synthesized mix-radix design dissipates lower energy and achieves smaller area than custom design from [14]. 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 [14], which leaves ample room for circuit-level optimizations if desired.

B. 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 (4) and (5). These operators and storage elements occupy large area and

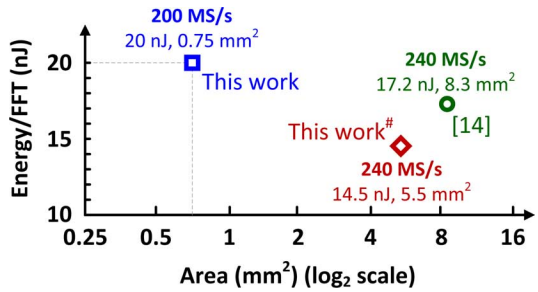


Fig. 4. Comparison of 1024-point FFTs. Our chip that minimizes power-area product is measured at 200 MS/s. At 240 MS/s, our fully synthesized (# indicates synthesis estimate) mix-radix design optimized for minimum energy achieves lower energy and lower area than the custom radix-4 subthreshold design from [14].

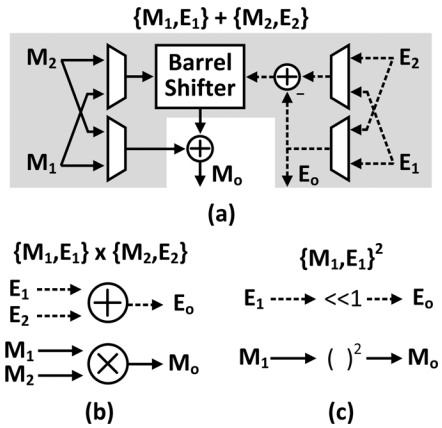


Fig. 5. Realization of floating-point operations: (a) addition, (b) multiplication, and (c) squaring by using 2's complement operators. M_1 and M_2 indicate mantissas, E_1 and E_2 indicate exponents. Blocks in the shaded area in (a) perform the exponent matching.

power due to the increased wordlength for large dynamic range signal processing. Using the fixed-point wordlength reduction [15], the PSD estimation requires at least 30 b for INR of 30 dB and SNR of -5 dB. The hardware area is reduced by exploiting two data properties. First, we assume that 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 b 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 [16]–[18].

The implementation of required floating-point operations is shown in Fig. 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(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(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(c). All of 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. 6(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. The wordlength reduction 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 10-b mantissa and 5-b 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. 6(b), the wordlength reduction from floating point results in a $2\times$ less logic area and a $2\times$ 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, which is a 35% saving. As shown in Fig. 7, 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.

C. Processing Time Overhead Reduction

STA and DTA blocks, given by (4) and (5), use multiply, divide, and square-root operations. Realization of the divide and square-root using conventional radix-2 long division [19] and CORDIC [20] would result in processing time of 0.14 ms [6]. When a -5-dB SNR weak primary user without adjacent-band interferers is 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, $Q^{-1}(P_{\mathrm{FA}})$, $Q^{-1}(P_D)$, and SNR [6], shown by

$$M(k) = \alpha \left((1+\psi(k)) \frac{Q^{-1}(P_{\mathrm{FA}}) - Q^{-1}(P_D)}{\mathrm{SNR}} - Q^{-1}(P_D) \right)^2 = 74.25 \cdot (1.15 + \psi(k))^2. \quad (6)$$

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.

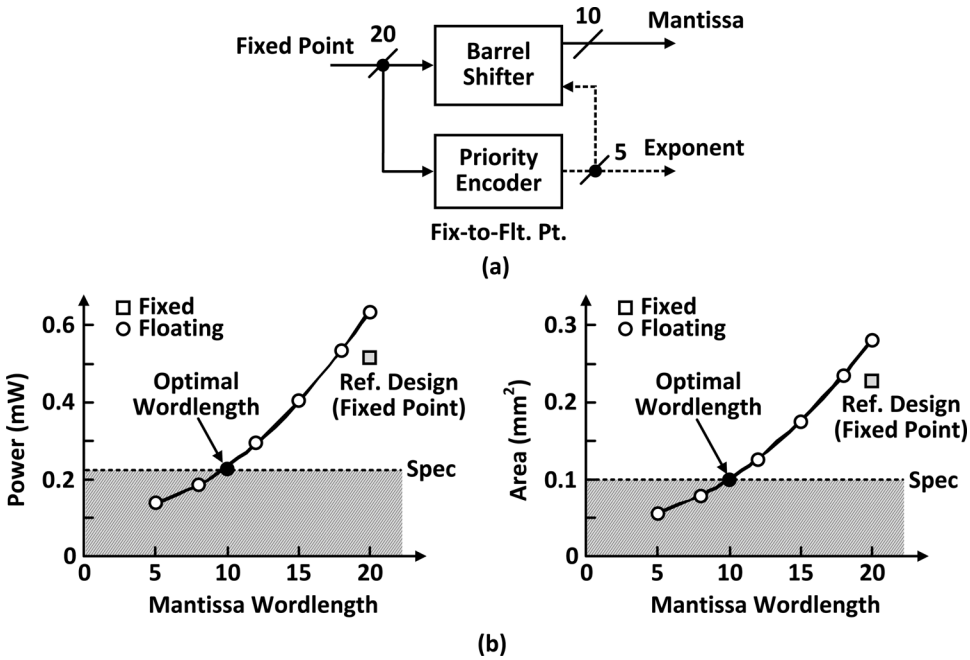


Fig. 6. (a) Architecture of fixed-point to floating-point conversion. The barrel shifter generates 10-b mantissa, the priority encoder generates 5-b exponent. (b) The plots of logic power and logic area of the power estimation block versus mantissa wordlength show that the sensing performance spec is met with 10-b mantissa. The floating-point format results in a 2× logic power reduction and a 2× logic area reduction as compared with the fixed-point reference design.

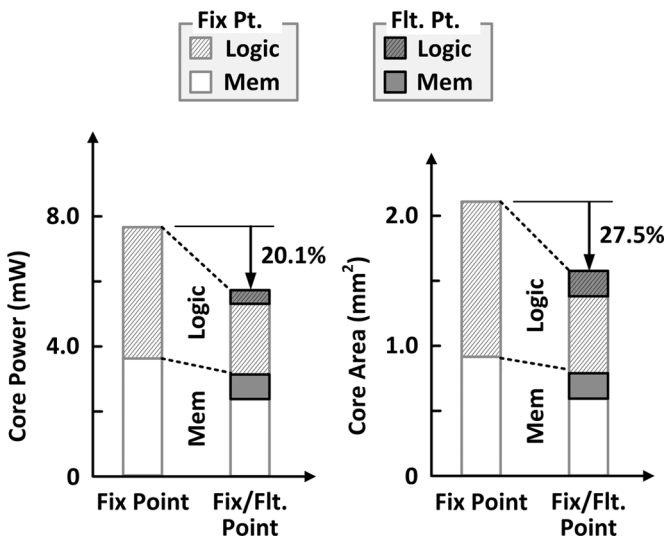


Fig. 7. Impact of wordlength on the core power and area. A 20.1% of core power (left) and 27.5% of core area (right) are saved by wordlength reduction. The wordlength reduction is enabled by the use of floating-point arithmetic.

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 four integer bits are needed. Second, $M(k)$ is a positive integer with precision of only 2°, 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 [15] to optimize the wordlength, 10-b output is allowed instead of a full 20-b precision. This leads to a 2× 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 $\sigma_{vf}^2(k)$ has a small variation during the sensing period, $1/\sigma_{vf}^2(k)$ can be updated in one clock cycle instead of ten clock cycles after the first reciprocal value is calculated.

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

D. 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 an power-area optimized FFT processor, wordlength pruning, and voltage scaling.

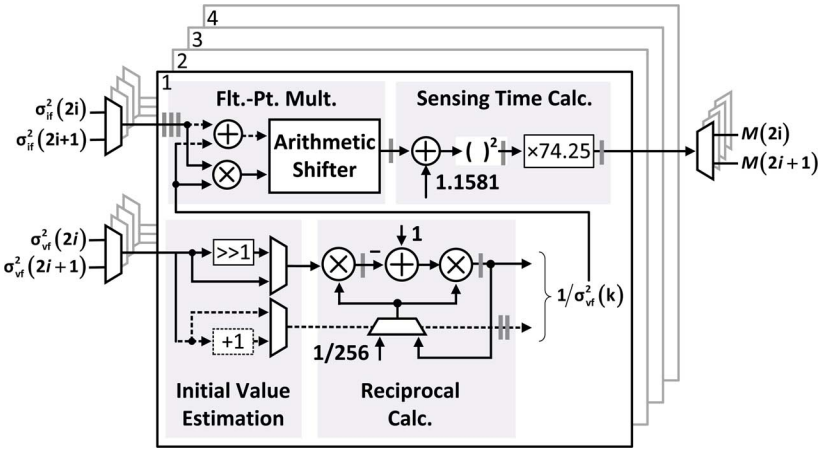


Fig. 8. Architecture of the sensing-time-adaptation block. The initial-value estimation and the reciprocal calculation are performed using Newton–Raphson algorithm. Bounded sensing time allows for 10-b output instead of full-precision 20-b dictated by the input wordlength.

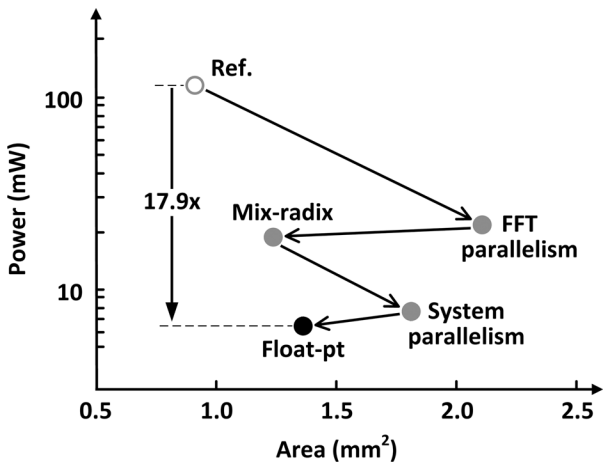


Fig. 9. Summary of the optimization procedure. Parallelism with voltage scaling and radix factorization with power-area optimized delay lines are applied to the FFT block. Parallelism with voltage scaling is also applied at the system level. Floating-point signal processing with wordlength reduction is applied to STA and DTA blocks. An overall $17.9\times$ reduction in power and $11.9\times$ reduction in power-area product are achieved.

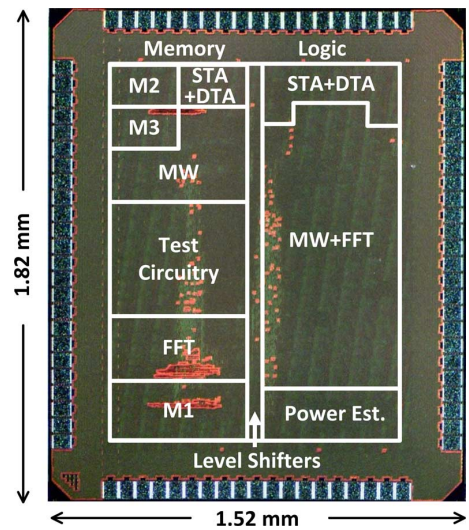


Fig. 10. Chip micrograph of wideband spectrum sensing processor.

Fig. 9 summarizes our optimization procedure. First, an 8-way parallel 1024-point FFT with 25-MHz clock and logic voltage of 0.6 V results in minimum PAP, which represents a $2.3\times$ 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 $2\times$ 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 $1.7\times$ 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-b mantissa and 5-b exponent reduces PAP by $1.6\times$ without introducing performance loss. An overall $17.9\times$ reduction in power and $11.9\times$ reduction in PAP is achieved.

V. Chip Implementation

Fig. 10 shows the die photograph of the wideband spectrum sensing processor that supports PSD estimation with sensing-time and detection-threshold adaptations. The core area is 1.14×1.44 mm^2. The logic and memory are placed and routed first as hard macros and 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 0.6–1.0 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 inserted to drive output pads for the entire range of supply voltage.

A. FPGA-Aided Verification

ASIC verification is performed with the use of an FPGA board [21] for pattern generation and data analysis, as shown in Fig. 11. The I/O interface between the PC terminal and the FPGA board is built on the BPS environment [22]. Test vectors are stored on the FPGA board, which stimulates the ASIC over two high-speed Z-DOK+ connectors. An external clock source

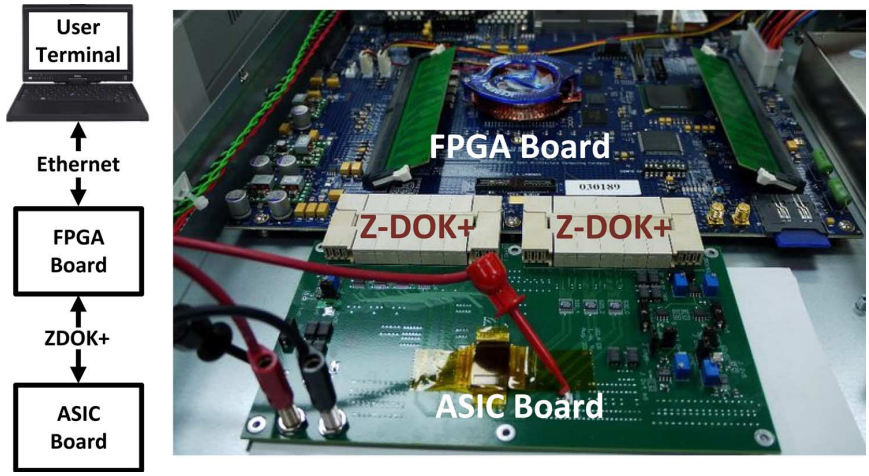


Fig. 11. FPGA-aided I/O verification. Test vectors are loaded from MATLAB into the FPGA board. The results are sampled in BRAM on the FPGA and read back by the terminal.

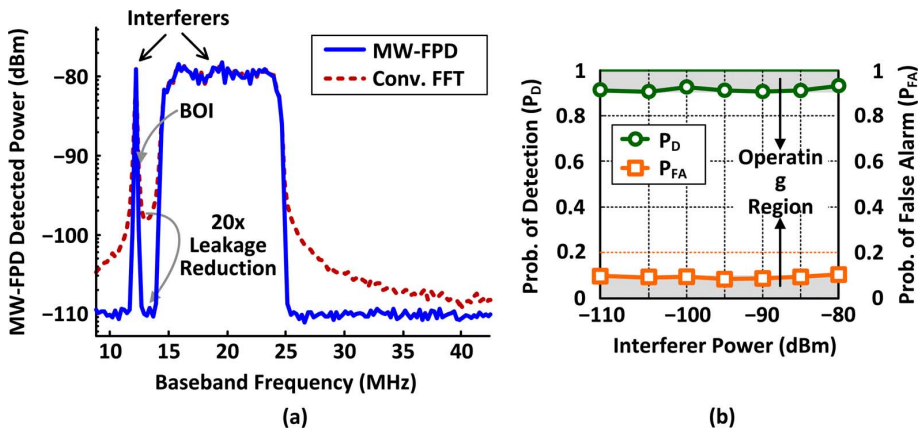


Fig. 12. (a) Measured PSD estimated by MW-FPD. Two 30-dB interferers are present two bins away from the BOI. The interfering power is suppressed by $20\times$ using multitap windowing as compared with traditional FFT-based PSD estimation. (b) Measured P_D and $P_{\rm FA}$ as a function of adjacent-band interferer power ($0 \leq$ INR $\leq$ 30 dB) for SNR $= -5$ dB. STA and DTA blocks effectively maintain $P_D \geq 0.9$ and $P_{\rm FA} \leq 0.1$.

is used to provide a wide range of operating frequency. The outputs of the ASIC are captured into block RAMs for analysis.

The FPGA board also contains a XAUI multigigabit interface that allows connectivity to external components. A reconfigurable RF front-end processor with two 12-b 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 [6] before chip implementation.

B. Measurement Results

The chip measurements are shown in Fig. 12. Fig. 12(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 $20\times$ 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. 12(b) shows that these two blocks maintain P_D and $P_{\rm FA}$ 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 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.

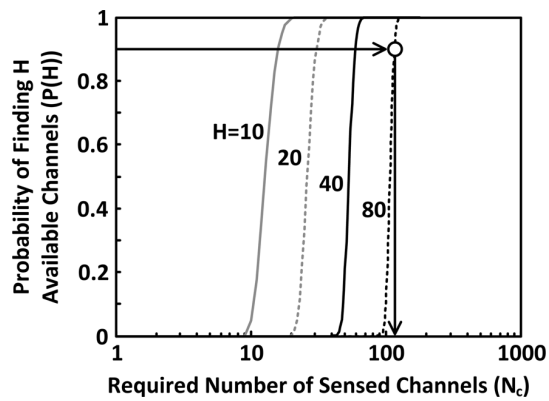


Fig. 13. Probability of finding H available channels versus the required number of sensed channels N_c. $P_{\rm FA} = 0.1$, $P_D = 0.9$, and $P_{\rm hold} = 0.8$. Detecting 128 channels is sufficient to achieve $P(H = 80) = 0.9$.

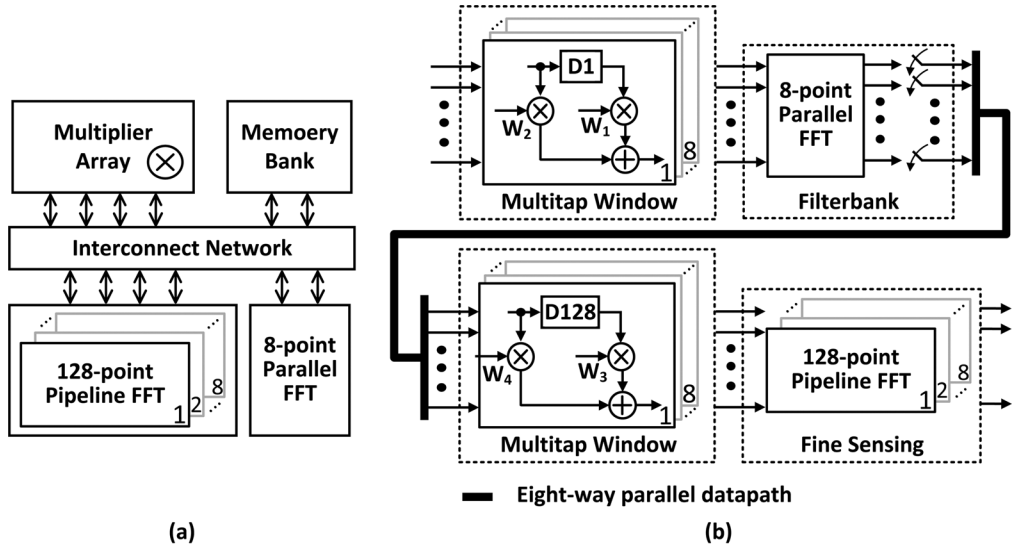


Fig. 14. (a) Architecture of the energy-efficient channelization, which supports both full and partial PSD estimation. (b) Hardware reconfiguration in partial PSD estimation mode.

TABLE III
SUMMARY OF CHIP FEATURES

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

TABLE IV
COMPARISON WITH PRIOR WORK

	[1-2]	[3]	[4]	This Work
Sensing BW (MHz)	1	10	0.4	200
Sensitivity	-	-	-	−5 dB
P_D	-	-	-	91%
P_{FA}	-	-	-	9%
Sensing 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 65nm # Synthesis estimate

The chip features are summarized in Table III. 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 0.82 mm^2 each and operate at 0.7 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-b add).

C. Performance Comparison

A comparison of the chip with the state of the art is listed in Table IV. In spectrum sensing performance, this work supports a $20\times$ wider bandwidth than prior work. It can sense a 200-MHz bandwidth with an SNR of -5 dB, $P_D \geq 0.9$, $P_{\rm FA} \leq 0.1$, and sensing time $<$50 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 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 to that of 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 $22\times$ 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.

VI. ENERGY-EFFICIENT CHANNELIZATION

Channelization by the computation-intensive FFT processor contributes more than 50% of the overall power consumption, after design optimizations described previously. The high power comes from the large sensing bandwidth of 200 MHz and the 1024-point FFT for the required frequency resolution of 200 kHz. When the spectrum is sparse, channelizing the entire 200-MHz spectrum, however, is not energy-efficient. Assume that the probability of the appearance of the spectral hole $P_{\rm hole}$

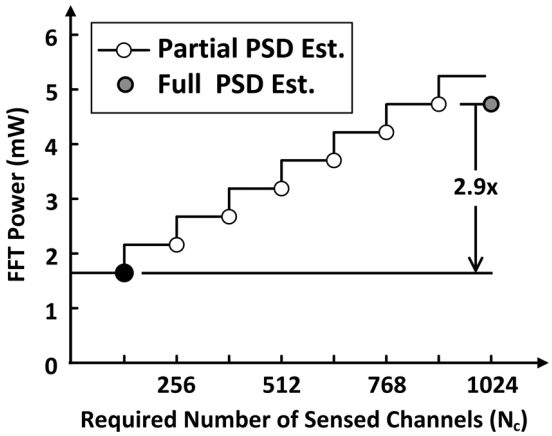


Fig. 15. Power benefit of partial PSD estimation. A $2.9\times$ lower power is achieved for $N_c \leq 128$.

is equal in each channel. The probability of finding at least H unused channels out of N_c channels can be modeled as

$$P(H) = 1 - \sum_{h=0}^{H-1} \binom{N_c}{h} (1-P_{\text{find hole}})^{N_c-h} \cdot (P_{\text{find hole}})^h \tag{7}$$

where H is the required number of spectral hole for secondary users, and $P_{\text{find hole}} = (1-P_{\text{FA}})\cdot P_{\text{hold}} + (1-P_D)\cdot(1-P_{\text{hold}})$ is the probability of find a spectrum hold [23]. Given $P_{\text{FA}} = 0.1$, $P_D = 0.9$, and a sparse spectrum with $P_{\text{hole}} = 0.8$, for example, the $P(H)$ curve with respect to N_c is shown in Fig. 13. For $H = 80$ and $P(H) \geq 0.9$, detecting 128 channels instead of 1024 is sufficient. Therefore, the FFT power can be further reduced by adapting the sensing bandwidth and the FFT size to the channel scarcity.

Fig. 14 shows the architecture of the energy-efficient channelization, which supports both full and partial PSD estimation. In the full PSD estimation mode, the eight-way parallel 128-point FFTs and multipliers concatenated with the eight-point FFT execute 1024-point FFT operation. In partial PSD estimation mode, the eight-point FFT performs filterbank to select the band of interest and to reduce sensing bandwidth. The 128-point FFTs channelize the spectrum to achieve frequency resolution of 200 kHz. The switches between the eight-point FFT and the 128-point FFTs are controlled according to the required number of sensed channels N_c. The hardware reconfiguration is made by changing the datapath interconnect due to the parallel architecture. Fig. 15 estimates the power saving with respect to N_c using hardware synthesis. Given a sparse spectrum and $N_c = 128$, a $2.9\times$ saving in FFT power that corresponds to a $1.7\times$ saving in total processor power can be achieved.

VII. Conclusion

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 with the prior state of the art, our chip dissipates $22\times$ less power per bandwidth for a $20\times$ wider sensing bandwidth. Therefore, a wideband ($>$100 MHz) spectrum sensing should leverage multitap windowed frequency power detection, adaptive sensing time and detection threshold estimations.

Acknowledgment

The authors would like to thank H. Chen for help with chip testing.

References

[1] 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 ISSCC Dig. Tech. Papers*, Feb. 2008, pp. 526–633.

[2] 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.

[3] T. Song *et al.*, "A 122-mW low-power multiresolution spectrum-sensing IC with self-deactivated partial swing techniques," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 57, no. 3, pp. 188–192, Mar. 2010.

[4] S. Izumi *et al.*, "A low-power multi resolution spectrum sensing (MRSS) architecture for a wireless sensor network with cognitive radio," in *Proc. IEEE SENSORCOMM*, Jul. 2010, pp. 39–44.

[5] S. Pollin *et al.*, "An integrated reconfigurable engine for multi-purpose sensing up to 6 GHz," in *Proc. Int. Symp. New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, May 2011, pp. 656–657.

[6] T.-H. Yu, O. Sekkat, S. Rodriguez-Parera, D. Marković, and D. Čabrić, "A wideband spectrum-sensing processor with adaptive detection threshold and sensing time," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 58, no. 11, pp. 2765–2775, Nov. 2011.

[7] T.-H. Yu, C. H. Yang, D. Čabrić, and D. Marković, "A 7.4 mW 200 MS/s wideband spectrum sensing digital baseband processor for cognitive radios," in *Proc. Int. Symp. VLSI Circuits*, Jun. 2011, pp. 254–255.

[8] Z. Quan, S. Cui, H. V. Poor, and A. Sayed, "Collaborative wideband sensing for cognitive radio," *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 60–73, Nov. 2008.

[9] R. Tandra and A. Sahai, "SNR walls for signal detection," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 4–17, Feb. 2008.

[10] C.-H. Yang, T.-H. Yu, and D. Marković, "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.

[11] Y.-W. Lin, H.-Y. Liu, and C.-Y. Lee, "A 1-GS/s FFT/IFFT processor for UWB applications," *IEEE J. Solid-State Circuits*, vol. 40, no. 8, pp. 1726–1735, Aug. 2005.

[12] Y. Chen, Y.-W. Lin, Y.-C. Tsao, and C.-Y. Lee, "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.

[13] K. K. Parhi, *VLSI Digital Signal Processing Systems*. New York: Wiley, 1999.

[14] M. Seok, D. Jeon, C. Chakrabarti, D. Blaauw, and D. Sylvester, "A 0.27 V 30 MHz 17.7 nJ/transform 1024-pt complex FFT core with super-pipelining," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2011, pp. 342–344.

[15] C. Shi and R. W. Brodersen, "Automated fixed-point data-type optimization tool for signal processing and communication systems," in *Proc. Design Autom. Conf.*, Jun. 2004, pp. 478–483.

[16] K. R. Ralev and P. H. Bauer, "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.

[17] A. Mitra, M. Chakraborty, and M. H. Sakai, "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.

[18] M. Chakraborty, R. Shaik, and M. H. Lee, "A block-floating-point-based realization of the block LMS algorithm," *IEEE Trans. Circuits Syst. II, Exp Briefs*, vol. 53, no. 9, pp. 812–816, Sep. 2006.
[19] M. D. Ercegovac and T. Lang, *Digital Arithmetic*. San Mateo, CA: Morgan Kaufmann, 2004.
[20] Y. H. Hu, "CORDIC-based VLSI architectures for digital signal processing," *IEEE Signal Process. Mag.*, vol. 9, no. 3, pp. 16–35, Mar. 1992.
[21] "ROACH: Reconfigurable Open Architecture Computing Hardware." [Online]. Available: http://casper.ssl.berkeley.edu/wiki/ROACH
[22] C. Chang, J. Wawrzynek, and R. Brodersen, "BEE2: A high-end reconfigurable computing system," *IEEE Des. Test Comput.*, vol. 22, no. 2, pp. 114–125, Mar. 2005.
[23] T.-H. Yu, C.-H. Yang, D. Marković, and D. Čabrić, "An energy-efficient VLSI architecture for cognitive radio wideband spectrum sensing," in *Proc. IEEE Global Telecommun. Conf.*, Dec. 2011, pp. 1–6.

Tsung-Han Yu (S'10) received the B.S. and M.S. degrees in electrical engineering from National Taiwan University, Taipei, Taiwan, in 2005 and 2007, respectively. He is currently working toward the Ph.D. degree in the Department of Electrical Engineering, University of California, Los Angeles.

His research is focused on digital integrated circuits and architectures for communication signal processing. He is currently is working on development of algorithm and circuit architecture on wideband spectrum sensing for cognitive radios.

Chia-Hsiang Yang (S'07–M'10) received the B.S. and M.S. degrees from the National Taiwan University, Taiwan, in 2002 and 2004, respectively, and the Ph.D. degree from the University of California, Los Angeles, in 2010. , all in electrical engineering.

He then joined the faculty of the Electronics Engineering Department, National Chiao Tung University, Taiwan, as an Assistant Professor. His current research is focused on energy-efficient integrated circuits and architectures for biomedical and communication signal processing.

Dr. Yang won the DAC/ISSCC Student Design Contest in 2010. In 2011, he received the Distinguished Ph.D. Dissertation in Circuits & Embedded Systems Award from the Department of Electrical Engineering, University of California, Los Angeles.

Danijela Čabrić (M'12) received the Dipl.Ing. degree from the University of Belgrade, Belgrade, Serbia, in 1998, the M.S. degree in electrical engineering from the University of California, Los Angeles, in 2001, and the Ph.D. degree in electrical engineering from the University of California, Berkeley, in 2007.

She was a member of the Berkeley Wireless Research Center, University of California, Berkeley. In 2008, she joined the faculty of the Electrical Engineering Department, University of California, Los Angeles, as an Assistant Professor. Her key contributions involve the novel radio architecture, signal processing, and networking techniques to implement spectrum sensing functionality in cognitive radios. She has written three book chapters and over 25 major journals and conferences in the fields of wireless communications and circuits and embedded systems.

Prof. Čabrić received the Samueli Fellowship in 2008, the Okawa Foundation Research Grant in 2009, and the National Science Foundation CAREER Award in 2012.

Dejan Marković (S'96–M'06) received the Dipl.Ing. degree from the University of Belgrade, Belgrade, Serbia, in 1998, and the M.S. and Ph.D. degrees from the University of California, Berkeley, in 2000 and 2006, respectively, all in electrical engineering.

In 2006, he joined the faculty of the Electrical Engineering Department, University of California, Los Angeles (UCLA), as an Assistant Professor. Since 2009, he has been affiliated with the Biomedical Engineering Interdepartmental Program, UCLA, as a cochair of the Neuroengineering field. He is also a director of the Integrated Circuits track within the UCLA Master of Science in Engineering Online Program. His current research is focused on integrated circuits for emerging radio and healthcare systems, programmable ICs, design with post-CMOS devices, optimization methods, and CAD flows.

Dr. Marković received the CalVIEW Fellow Award in 2001 and 2002 for excellence in teaching and mentoring of industry engineers through the University of California, Berkeley, distance learning program. In 2004, he was a corecipient of the Best Paper Award at the IEEE International Symposium on Quality Electronic Design. In recognition of the impact of his Ph.D. work, he was awarded 2007 David J. Sakrison Memorial Prize at UC Berkeley. He received the National Science Foundation CAREER Award in 2009. In 2010, he was a corecipient of ISSCC Jack Raper Award for Outstanding Technology Directions and a winner of the DAC/ISSCC Student Design Contest.