

ASIC Implementation of a Computationally Efficient Compressive Sensing Detection Method using Least Squares Optimization in 45 nm CMOS Technology

Mohamed Shaban, Tarek Idriss, Haytham Idriss, Magdy Bayoumi, *Fellow IEEE*

The Center for Advanced Computer Studies
University of Louisiana at Lafayette
Lafayette, Louisiana, USA

Abstract—This paper presents a high speed architecture of a recently proposed compressive sensing detection method for wideband cognitive radios using least squares. Using least squares instead of the orthogonal matching pursuit for signal recovery reduces the computational complexity where, the index search and matrix inverse stages are avoided. The proposed architecture is fully pipelined where, 14 clock cycles are required to detect 1024-length signal occupying 8 channels from 16 measurements. The design is implemented in 45 nm CMOS operating at 165 MHz. Since, the sensing time for 1024-length signal is roughly 84.8 ns, the proposed design offers high speed signal detection compared with state of art orthogonal matching pursuit architecture.

Keywords—Cognitive radio; wideband spectrum sensing; compressive sensing; least squares; CMOS technology

I. INTRODUCTION

The cognitive radio (CR) was recently introduced as a way to enhance spectral efficiency by opportunistically accessing licensed frequency bands which are underutilized by licensed users. A major component of CR is spectrum sensing [1]. Spectrum sensing refers to the process of detecting a spectral opportunity in a defined licensed frequency band. A spectral opportunity is a frequency channel not being used by a licensed user.

Recent works have focused on techniques aimed at detecting a spectral opportunity within a wide spectrum band of interest [2]. However, it is infeasible to realize these techniques as they require high resolution analog-to-digital converters (ADCs) operating at extremely high sampling rates. In fact, the industry fastest 16 bit ADC offers a sampling rate of 250 MS/s which is far beyond the target sampling rates [3]. Furthermore, using commercial low resolution multi-GS/s ADCs for wideband spectrum sensing reduces the sensing accuracy while high design, computational complexities and large memory resources are required.

Fortunately, recent developments in mathematics and signal processing have contributed to overcoming this bottleneck by introducing compressed sensing theory [4]. The objective of compressed sensing is to sample an analog signal at a sub-Nyquist rate comparable to the signal information rate. The sampling operation is a trivial linear operation where M

measurements are acquired by projecting N time samples of the received signal into a random domain where $M \ll N$. Several architectures have been introduced to perform the sampling operation such as: random demodulation [6], random modulation preintegration [7], parallel segmented compressed sensing [8], and compressive channel multiplexing [9]. Recovery is realized using computationally complex non-linear convex optimization (i.e. basis pursuit (BP)) [4] or greedy iterative algorithms (i.e. orthogonal matching pursuit (OMP)) [5]. The complexity of the former and latter methods are $O(N^3)$ and $O(KMN)$ respectively where K is the sparsity of the received signal or the frequency occupation of the frequency band of interest [5]. Accordingly, the complexity of recovery introduces a bottleneck in the realization of compressed sensing method for wideband cognitive radios.

Furthermore, since the occupation of the frequency band of interest (K) is always unknown, it is impossible to assume an appropriate value for the minimum number of required measurements M . As a result, compressed sensing cannot be directly used when K is unknown. Sequential compressed sensing was then introduced where M is estimated [10]. However, this approach is computationally intensive as it requires the execution of the recovery algorithm several times increasing the computational complexity of the design. The computational complexity of the recovery is roughly $O(MN^3)$ and $O(KM^2N)$ when BP and OMP are used respectively.

In [11], we proposed a computationally efficient compressive sensing detection technique for wideband cognitive radios when the frequency occupation is unknown using the least squares optimization. The computational complexity of proposed method recovery is roughly $O(L^2)$ which is far lower than that of standard as well as sequential compressive sensing techniques where, L is the total number of channels [11].

In this paper, we propose high speed architectures for both recovery and detection operations of [11]. Furthermore, the design is implemented in 45 nm CMOS technology. Gate level net-list is obtained using Synopsis design vision. Moreover, the circuit layout is generated using Cadence encounter. The area, dissipated power, clock frequency, and sensing time are measured.

The remainder of this paper is organized as follows. Section II presents the proposed compressed sensing detection method.

Architectures for both recovery and detection are introduced in section III. Section IV presents the circuit layout. The conclusion and summary are discussed in section V.

II. COMPRESSIVE SENSING DETECTION METHOD

In this section we review, the proposed compressive wideband spectrum sensing detection method mentioned in [11]. The purpose of the proposed method is to compressively sense licensed signals within a wide frequency band of interest when the frequency occupation is unknown. The proposed method offers a reduced computational complexity compared to standard as well as sequential compressed sensing methods.

The proposed method directly estimates the average of individual frequency components for each frequency sub-band of the frequency band of interest, instead of estimating the nonzero frequency components (K) within the band. For example, if the frequency band of interest is divided into L frequency sub-bands, L samples are required to be estimated where, $L \ll K$. As a result, a reduction in the number of required measurements M as well as the design complexity of the sampling stage is achieved [11]. Moreover, the proposed method offers a solution for sensing the frequency band of interest when the frequency occupation is unknown.

Furthermore, in order to estimate the L samples using given M measurements where, $M > L$ ($M \cong CL$ where, $C > 1$), least squares optimization was introduced to solve the overdetermined system with M linear equations and L unknown variables. As a result, the computational complexity of the recovery stage is reduced; a reduction related to the use of fewer number of measurements for recovery, as well as a reduction related to the replacement of standard recovery algorithms (i.e. BP and OMP) by the least squares optimization. Computational complexity of the recovery stage is roughly $O(L^2)$ [11]. Comparing the recovery complexities of the proposed and standard (or sequential) compressive sensing methods, the proposed method offers a computationally efficient high speed recovery.

Moreover, detection was introduced such that the squared magnitudes of estimated averages are compared to predefined thresholds [11]. Fig. 1 shows the block diagram of the proposed method which consists of three stages: sampling stage, recovery stage and detection stage.

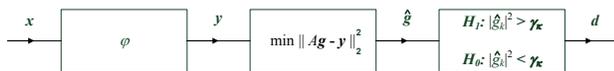


Fig. 1: Proposed compressive sensing detection block diagram

A. Sampling stage

Let $x(t)$ be the received signal by the cognitive radio. Generally, $x(t)$ represents the contributions of licensed signals transmitted within L frequency sub-bands. Let $N \times 1$ vector \mathbf{x} be the discrete time representation of $x(t)$ when sampled at the Nyquist rate. Vector \mathbf{x} can then be written as:

$$\mathbf{x} = \boldsymbol{\psi} \mathbf{z} \quad (1)$$

where, \mathbf{z} is $N \times 1$ frequency domain components vector and $\boldsymbol{\psi}$ is $N \times N$ discrete Fourier transform coefficients basis matrix. Furthermore, we define an average vector \mathbf{g} where,

$$\mathbf{g} = \Gamma \mathbf{z} \quad (2)$$

where, Γ is $L \times N$ matrix with L orthogonal rows such that each row consists of N/L ones corresponding to the location of each frequency sub-band within the entire frequency band of interest.

$$\Gamma = L/N \begin{bmatrix} 1 & 1 \dots 1 & 0 & 0 \dots 0 & \dots & 0 & 0 \dots 0 \\ 0 & 0 \dots 0 & 1 & 1 \dots 1 & \dots & 0 & 0 \dots 0 \\ 0 & 0 \dots 0 & \ddots & \ddots & \ddots & \vdots & \\ 0 & 0 \dots 0 & \dots & 0 & 0 \dots 0 & 1 & 1 \dots 1 \end{bmatrix} \quad (3)$$

\mathbf{x} is then rewritten as follows:

$$\mathbf{x} = \boldsymbol{\psi} \Gamma^+ \mathbf{g} \quad (4)$$

where, Γ^+ is the pseudo inverse (right inverse) of Γ (i.e. $\Gamma^+ = \Gamma^T (\Gamma \Gamma^T)^{-1}$). Furthermore, sampling is done using the compressed sensing theory [4] where, the output measurements vector \mathbf{y} is given as follows:

$$\mathbf{y} = \boldsymbol{\varphi} \mathbf{x} = \Pi \mathbf{g} \quad (5)$$

where, $\Pi = \boldsymbol{\varphi} \boldsymbol{\psi} \Gamma^+$ and, $\boldsymbol{\varphi}$ is $M \times N$ sensing matrix with Bernoulli independent identically distributed (i.i.d.) entries.

B. Recovery stage

The recovery stage is introduced to estimate the average of frequency components for each frequency sub-band \mathbf{g} . Since, Π is $M \times L$ matrix where, $M > L$, then " $\mathbf{y} = \Pi \mathbf{g}$ " represents an overdetermined system of linear equations [11]. Least squares method is then introduced to obtain an approximate estimate of \mathbf{g} defined as follows,

$$\hat{\mathbf{g}} = \arg_{\mathbf{g}} \min \|\Pi \mathbf{g} - \mathbf{y}\|_2^2 \quad (6)$$

If Π is well conditioned, (i.e. L columns of Π are linearly independent), least squares will be reduced to,

$$\hat{\mathbf{g}} = \Pi^+ \mathbf{y} \quad (7)$$

where, Π^+ is the pseudo inverse (left inverse) of Π (i.e. $\Pi^+ = (\Pi^H \Pi)^{-1} \Pi^H$).

C. Detection stage

Since, the purpose of wideband spectrum sensing is to decide whether licensed signals are present within the L frequency sub-bands or not, a detection stage is introduced following the recovery stage where, hypothesis H_0 and H_1 are introduced such that,

$$H_1 : |\hat{g}_k|^2 > \gamma_k \quad \text{“Signal is present”}$$

$$H_0 : |\hat{g}_k|^2 < \gamma_k \quad \text{“Signal is absent”}$$

where, $|\hat{g}_k|^2$ is the squared magnitude of the k^{th} element of \hat{g} and γ_k is a predefined detection threshold for the k^{th} frequency sub-band where, $k = 1, 2, \dots, L$ [11].

III. PROPOSED ARCHITECTURE

Several implementations for compressed sensing recovery using the OMP algorithm were proposed recently in both FPGA [12][13] and ASIC (65 nm CMOS technology) [14]. One drawback of the OMP implementations [12-14] is that, they require two computationally expensive operations (i.e. matrix vector multiplication in the index search stage, and matrix inverse in the least squares stage). Another drawback is that OMP implementations [12-14] assume a real representation basis matrix ψ , therefore, these implementations are incapable of recovering compressively sampled frequency sparse licensed signals.

In this paper, we introduce high speed architectures for the recovery and detection stages of the proposed compressive sensing detection method [11]. First, the proposed design is fully pipelined and optimized to reduce the sensing time. Second, ψ is assumed as a complex matrix where frequency sparse signals are considered. Third, recovery is realized using the least squares method. However, since, matrix Π is fixed for the design, Π^+ is computed offline and stored in $2L$ registers. As a result, the least squares method is reduced to a matrix vector multiplication where, Π^+ is multiplied by y , therefore, $2ML$ multiplications and $2(M-1)L$ additions are required. A reduction in the circuit complexity as well as enhancement in the design speed is achieved via eliminating the index search stage and matrix inverse operation required by OMP implementations [14-16].

In this paper, a hardware is implemented for $N = 1024$, $M = 16$, and $L = 8$. Each sample is represented in 24 bits fixed point arithmetic (8 integer bits and 16 fractional bits). Fig. 2 shows the proposed architecture for the recovery operation.

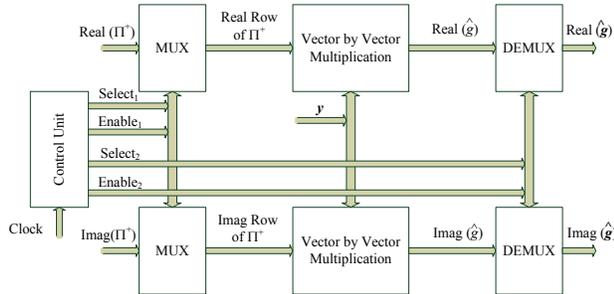


Fig. 2: Proposed compressive sensing recovery architecture using least squares

The proposed architecture realizes the recovery operation defined by (7) where, a complex matrix Π^+ is multiplied by y . The architecture consists of a control unit, two 8×1 multiplexers (MUXs), two 1×8 demultiplexers (DEMUXs) and

two vector by vector multiplication modules. Since, Π^+ is 8×16 complex matrix, Π^+ is then stored in 16 registers (384 bits wide) where real or imaginary entries of each row are concatenated and stored in each register. Also, since y is 16×1 real vector, then y is also stored in 16 registers (24 bits wide).

Each MUX has 8 input ports, each is 384 bits wide. The output of each MUX is driven once the multiplexer is triggered by the positive edge of the clock signal and the enable as well as select signals are set. Each MUX is cascaded by a vector by vector multiplication module such that, each MUX output is stored into a 384-bit register which is then split into 16 components (24 bits wide each) in order to be fed as inputs to the module.

Each vector by vector multiplication module consists of three sub-modules, each corresponding to a different pipeline stage. The first sub-module performs multiplication via 16 array multipliers. The original architecture of an array multiplier is composed of 22 carry save adders and a single carry look ahead adder. In fact, the use of a carry look ahead adder for each array multiplier results in a slower multiplication stage, requiring much longer clock period than the rest of the stages. However, since the outputs of array multipliers will be added together to get the final output of the vector by vector multiplication module, we propose an enhancement to the array multiplier architecture where, the carries of each array multiplier is forwarded to the next stage, therefore, the carry look ahead adder is no longer needed. Accordingly, area, power consumption and clock delay is reduced. The semi products and carries obtained are then stored into 32 registers (24 bits wide).

In the second sub-module, semi products and carries are added using 30 carry save adders to obtain a partial sum and carry which are then stored into 2 registers (24 bits wide). The partial sum and carry are then added using a single carry look ahead adder in the third sub-module.

Table 1: Time schedule for recovery operation

Clock Cycle	Operation
1 st	Enable and select signals of each MUX and DEMUX are set.
2 nd	Each MUX delivers either a row of real entries or a row of imaginary entries of Π^+ to the corresponding vector by vector multiplication module.
3 rd	Semi products and carries are computed by the first vector by vector multiplication sub-module and then stored into 32 24-bit registers.
4 th	Partial sum and carry are computed by the second vector by vector multiplication sub-module and then stored into 2 24-bit registers.
5 th	Final product is computed by the third vector by vector multiplication sub-module and then stored into 16 24-bit registers via a DEMUX.

The outputs of the vector by vector multiplication modules Π^+ are stored into 16 registers (24 bits wide) using a DEMUX which is triggered by the positive edge of the clock signal. Each DEMUX has 8 output ports of 24 bits width. A control unit is designed to schedule the transmission of data from each

MUX to the corresponding vector by vector multiplication module and from each vector by vector multiplication module to the corresponding DEMUX. Time schedule of the recovery operation is then illustrated in table 1.

From table 1, the entire recovery operation will take 40 clock cycles till the 8 rows of Π^+ are multiplied by y . In order to reduce the number of clock cycles and therefore increase the speed of the design, the proposed architecture is fully pipelined where, rows of Π^+ are processed at consecutive clock cycles. As a result, the duration of the recovery process is reduced to 12 clock cycles.

Fig. 3 shows the proposed architecture for a single channel of the detection operation. Generally, the proposed detection architecture consists of L channels. In the l^{th} channel, $|\hat{g}_l|^2$ is computed where, 2 array multipliers, 2 carry save adders, a single carry look ahead adder are employed and then compared with a predefined threshold η using a comparator. Only two pipeline stages are required. The output of the first pipeline stage (i.e. semi products and carries of array multipliers) is stored into 32 registers (24 bits wide) to be further processed by the carry save and carry look ahead adders. Furthermore, the output of the second pipeline stage (i.e. decision of the L detectors d_l where, $l=1, 2, \dots, L$) is stored into 8 flip-flops. The whole operation takes 2 clock cycles.

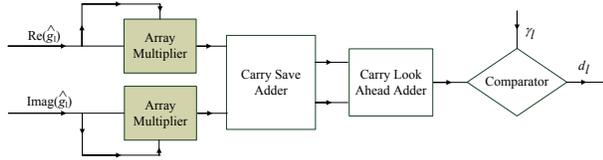


Fig. 3: Proposed detection architecture for a single channel l

IV. CIRCUIT IMPLEMENTATION

Both architectures are implemented in 45 nm CMOS technology operating at 1.1 V supply voltage. Architectures were synthesized using synopsis design vision while circuit layout was generated using Cadence encounter. Fig. 4 shows the circuit layout generated for the recovery and detection architectures.

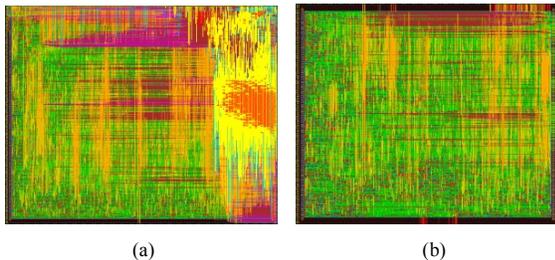


Fig. 4: Circuit layout in 45 nm CMOS technology for (a) proposed recovery architecture (b) proposed detection architecture

Table 2 shows the layout summary for both recovery and detection architectures. The total area of the design is 0.61 mm^2 while, the total dissipated power is 163.14 mW. Since, the

circuit operates at a clock frequency of 165 MHz, the total sensing time (recovery and detection time) is 84.8 ns.

Table 2: Layout summary for proposed recovery and detection architectures

	Recovery	Detection
Technology	45 nm, 1.1 V	
Area	0.42 mm ²	0.19 mm ²
Power Dissipation	107.79 mW	55.35 mW
Clock Frequency	165 MHz	
No. of Clock Cycles	12	2
Delay	72.7 ns	12.1 ns

Comparing our results with those of the state of art OMP architecture [13], we observe that, using proposed recovery and detection architectures for the wideband spectrum sensing of 1024-length signals with sparsity (K) varying from 0 to 1024 is roughly 7.3×10^3 times faster than using the state of art OMP architecture for mere recovery of same length signals with K varying from 0 to 36 [13]. Furthermore, the maximum clock frequency achieved using proposed architectures (i.e. 165 MHz) is higher than that achieved by the state of art OMP architecture (i.e. 100 MHz) [13]. Accordingly, the proposed recovery and detection architectures offer a high speed signal detection for wideband cognitive radios when compared with the state of art compressed sensing OMP recovery architecture.

V. CONCLUSIONS

In this paper, we introduced a high speed architecture and ASIC implementation (45 nm CMOS technology) for a recently proposed computationally efficient compressive sensing detection method for wideband cognitive radios that employs the least squares optimization for the recovery stage. The design was fully pipelined and the sensing time for 1024-length signals with unknown frequency occupation was measured. It was found that, the proposed method offers a fast signal recovery and detection for licensed signals within a wide frequency band of interest with unknown frequency occupation with respect to signal recovery using state of art OMP architecture which assumes prior knowledge of licensed signals frequency occupation.

REFERENCES

- [1] T. Yucek, and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications", *IEEE Communications Surveys & Tutorials*, vol. 11, no. 1, March 2009.
- [2] H. Sun, A. Nallanathan, C. Wang, and Y. Chen, "Wideband spectrum sensing for cognitive radio networks: a survey", *IEEE Wireless Communications Magazine*, vol. 20, no. 2, April 2013.
- [3] "AD9467: 16-Bit, 200 MSPS/250 MSPS Analog-to-Digital Converter", Analog Devices, 2010, <http://www.analog.com/en/analog-to-digital-converters/ad-converters/ad9467/products/product.html>.
- [4] E. Candés, "Compressive sampling", *Proceedings of the International Congress of Mathematicians*, Madrid, Spain, August 2006.
- [5] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit", *IEEE Transactions on Information Theory*, vol. 53, no. 12, December 2007.

- [6] S. Kirolos, J. Laska, M. Wakin, M. Duarte, D. Baron, T. Ragheb, Y. Massoud, R. Baranuik, "Analog-to-information conversion via random demodulation", *IEEE Dallas/CAS Workshop on Design, Applications, Integration and Software*, Richardson, Texas, October 2006.
- [7] J. Yoo et. al., "Design and implementation of a fully integrated compressed sensing signal acquisition system", *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Kyoto, Japan, March 2012.
- [8] Z. Yu, S. Hoyos, and B. M. Sadler, "Mixed-signal parallel compressed sensing and reception for cognitive radio", *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Las Vegas, Nevada, April 2008.
- [9] J. Slavinsky, J. Laska, M. Davenport and R. Baranuik, "The compressive multiplexer for multi-channel compressive sensing", *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Prague, Czech Republic, May 2011.
- [10] D. Malioutov, S. Sanghavi, and A. Willsky, "Sequential compressed sensing", *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, April 2010.
- [11] M. Shaban, D. Perkins, M. Bayoumi, "Application of compressed sensing in wideband cognitive radios when sparsity is unknown", *IEEE Wireless and Microwave Technology Conference*, Tampa, Florida, June 2014.
- [12] A. Septimus, R. Steinberg, "Compressive sampling hardware reconstruction", *IEEE International Symposium on Circuits and Systems*, Paris, France, May 2010.
- [13] L. Bai, P. Maechler, M. Muehlberghuber, H. Kaeslin, "High-speed compressed sensing reconstruction on FPGA using OMP and AMP", *IEEE International Conference on Electronics, Circuits and Systems*, Seville, Spain, December 2012.
- [14] J. Stanislaus, T. Mohsenin, "High performance compressive sensing reconstruction hardware with QRD process", *IEEE International Symposium on Circuits and Systems*, Seoul, South Korea, May 2012.