Compressed Spectrum Sensing for Whitespace Networking

Vishnu Katreddy
Masters Project Report

Abstract

We study the problem of wideband compressed spectrum sensing for detection of whitespaces when some prior knowledge about spectrum occupancy is available. Our experiments show that the primary user occupancy of the UHF band at a given location doesn't vary much with time. Thus, from initial sensing we estimate the support of spectrum occupancy and we use this knowledge in subsequent sensing. We explore two methods in this regard. However, the above methods still require spectrum recovery which has high complexity but since we are interested in only detection, we pose the detection problem directly on the compressed samples (thereby avoiding recovery) assuming i.i.d flat fading between the primary users, and analyse its performance.

I. INTRODUCTION

It is widely acknowledged that the occupancy of primary users is low in the UHF-VHF bands. The primary users consist of licensed digital television stations and wireless microphones. Cognitive radios enable efficient use of the spectrum by opportunistically utilizing these vacancies when they exist, which requires the need for efficient spectrum sensing mechanism. Though the TV band occupancies are static, the microphones can occupy and leave any vacant band at any time. Thus in a given location the spectrum occupancy is varying in time and the cognitive radios should not only be able to detect the vacant bands but also vacate them quickly when a primary user appears. Currently, the whitespaces are mostly unutilized by the cognitive radios, as a result of which one can find a vacant band by monitoring only few bands. But in the future, as the number of active cognitive radios in a given location increase, occupancy of whitespaces too increases. This makes it more difficult to find a vacant band by just monitoring few bands, which necessiates the need to monitor a wide band to find a vacancy. One approach to sense multiple bands is to sense them sequentially which increases the sensing time thus decreasing the time available for data transmission. The other approach is to sense a wide band using high rate ADC but it is prohibitive.

Compressed Sensing/Sampling (CS) has emerged as an exciting method for acquisition of sparse or compressible signals at rates significantly lower than Nyquist rates. The original signal can be recovered from the compressed samples by a convex optimization. [1] has shown that the signals of concern have sparse edge spectrum (differential of power spectral density) and propose a paradoxical compressed sensing scheme based on it. In [2] the authors propose a realisitic compressed spectrum sensing architecture exploiting the sparseness of edge spectrum. Both the above works assume that no prior knowledge of primary user occupancy is available. But, in reality, at a given location, the spectrum occupancy is slowly varying. So, initially when a cognitive radio boots up, it can perform normal spectrum sensing, identify the primary user occupancy and use this knowledge in subsequent sensing thereby further reducing the sampling rate.

II. COMPRESSED SENSING

A. Theory

In the standard CS framework, we acquire a signal $x \in R^n$ (the same theory can be extended to complex signals as well) via m linear measurements

$$y = \Phi x \tag{1}$$

where Φ is and $m \times n$ matrix representing the compressive sampling system and $y \in R^m$ is the vector of measurements. Nyquist sampling theory states that, in order to ensure that there is no loss of information, the number of samples m should be at least the signal dimension n. On the other hand, CS theory allows for $m << n$, as long as the signal x is sparse or compressible in some basis. Let Ψ be the $n \times n$ sparsity basis matrix of the signal x. Then x can be represented as $x = \Psi\beta$ where β is $n \times 1$ vector of coefficients. x is called s-sparse if β

has upto $s << n$ non-zero and large enough entries i.e. $\beta \in \Sigma_s$ where $\Sigma_s = \{\beta \in R^n : ||\beta||_0 := |supp(\beta)| \leq s\}$ and

$$y = \Phi x = \Phi\Psi\beta = \hat{\Phi}\beta \tag{2}$$

Since $m << n$ the CS scheme is vastly under determined. Without loss of generality, we assume $\Psi = I$. To understand the conditions under which one can stably recover the signal x, Tao and Candes introduced the restricted isometry property (RIP) of the matrix Φ in [3]. A matrix Φ is said to obey RIP of order-s if there exists a constant δ_s known as s-restricted isometry constant $\in (0,1)$ such that

$$(1-\delta_s)\|x\|^2 \leq \|\Phi x\| \leq (1+\delta_s)\|x\|^2 \tag{3}$$

holds for all $x \in \Sigma_s$. In other words Φ is a restricted isometry for vectors $\in \Sigma_s$.

They also established that if $\delta_{2s} < 1$ then x can be exactly recovered by solving

$$(P_0) \quad \hat{x} = \arg\min_x ||x||_0 \qquad s.t \quad y = \Phi x \tag{4}$$

(but the above optimization is NP-hard) and that if $\delta_s + \delta_{2s} + \delta_{3s} < 1$ which is stringent than $\delta_{2s} < 1$, then the optimization $P_1 \equiv P_0$

$$(P_1) \quad \hat{x} = \arg\min_x ||x||_1 \qquad s.t \quad y = \Phi x \tag{5}$$

It has been shown in [4] that if Φ is random gaussian matrix, then

$$m > c_1 s \left(\log\frac{n}{s} + 1 + \log\frac{12}{\delta}\right) \tag{6}$$

number of samples are needed for exact recovery using CS for a signal of sparsity-s.

B. Architecture Proposed in [2]

The compressed sensing approach proposed in [2] works as follows. The wideband analog signal $x(t)$ is sampled using an analog-to-information-converter (AIC). An AIC may be conceptually viewed as an ADC operating at Nyquist rate, followed by compressive sampler. Denote the output of ADC by x, $N \times 1$ vector and the output of compressive sampler by y, $M \times 1$ vector. Denote the respective $2N \times 1$ and $2M \times 1$ autocorrelation vectors of x and y as follows

$$\begin{aligned} r_x &= \begin{bmatrix} 0 & r_x(-N+1) & \dots & r_x(0) & r_x(N-1) \end{bmatrix}^T \\ r_y &= \begin{bmatrix} 0 & r_y(-M+1) & \dots & r_y(0) & r_y(M-1) \end{bmatrix}^T \end{aligned}$$

Using $y = \Phi x$ and above equations, after some algebraic manipulation one can obtain

$$r_y = \Phi_I r_x \tag{7}$$

where Φ_I can be expressed in terms of Φ (bit lengthy). From the compressive samples, y we need to recover the edge spectrum, z (sparse basis coefficients) of x. The edge spectrum, z of x can be expressed as

$$z = \Gamma F W r_x$$

where W, F are the $2N \times 2N$ matrices representing the wavelet smoothing operation and discrete fourier transform. Γ is the derivative operation, approximated by a first-order difference, given by the $2N \times 2N$ matrix,

$$\Gamma = \begin{bmatrix} 1 & 0 & \dots & 0 \\ -1 & 1 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ 0 & \dots & -1 & 1 \end{bmatrix}$$

Or alternatively we have $r_x = (\Gamma F W)^{-1} z$ and using (7) we obtain

$$r_y = \Phi_I \Psi z$$

where $\Psi = (\Gamma F W)^{-1}$ is the sparse basis.

The CS reconstruction of the edge spectrum can be formulated as an l_1 norm optimization problem as follows

$$\hat{z} = \arg\min_{z} ||z||_1 \qquad s.t \quad r_y = \Phi_I G z \tag{8}$$

PSD can be recovered from edge spectrum by $S_x(n) = \sum_{i=0}^{n} z(i)$

C. Our Architecture

The spectrum sensing device that we have samples the received signal, $x(t)$ at Nyquist rate, computes its FFT in hardware and gives the power spectral density (PSD) as output, denoted by S_x. The PSD is then smoothed using a Moving average filter which is linear and then compressive sampling is applied, the output of which is denoted by y.

Thus $y = \Phi S_m$ where S_m is the smoothed PSD. Using $S_m = \Gamma^{-1} z$ we obtain $y = \Phi \Psi z$ where $\Psi = \Gamma^{-1}$

Thus, we formulate the CS reconstruction of the edge spectrum as an l_1 norm optimization problem as follows

$$z = \arg\min_{z} ||z||_1 \qquad s.t \quad y = \Phi \Psi z \tag{9}$$

From z one can obtain the smoothed PSD.

One has to note that this is not a realistic architecture and a true architectue as in the above section does not perform as well as this. The reason as explained in [2] is because of the reduced incoherency between Φ_I and Ψ in the true architecture. Nevertheless, an architecture like this provides some useful insights as shown in [1] regarding the performance evaluation of various compressed sensing algorithms.

The compressive sampling rate m is restricted by the sparsity level of the signal and for $s-$sparse signal, it is given by (6).

III. COMPRESSED SENSING WITH PRIOR KNOWLEDGE

In this section we investigate if we can improve on m if a part of the support of x is known. The problem definition is the same as in section $2A$. We introduce some notation here. $S := support(x)$, i.e the set of all non-zero components of x and $|S| = s$. K is the 'known' part of S obtained from prior sensing and $|K| = k$. We assume that our knowledge has no errors i.e $K \subseteq S$. This framework can easily be extended to incorporate errors but we neglect it for time being to keep things simple. $U := S \backslash K$ is the unknown part of the support, which needs to be estimated and $|U| = u$. Since, $S = K \cup U$ and $K \cap U =$, $s = k + u$. We discuss the conditions for exact reconstruction of the original signal x from the compressed samples y, in this framework in the following section.

A. Signal Recovery

If $K = \{\}$, then one would use (5) to recover x from y and if one knew all of the support of x, i.e $K = S$ then one can use the least squares given below to obtain x from y. We call this 'LS-Oracle'.

$$x_s = (\Phi_s^T \Phi_s)^{-1} \Phi_s^T y, \qquad x_{s^\perp} = 0$$

where Φ_s is $\Phi \Psi_S$.

If $K \subseteq S$, then [5] have proved that if $\delta_{k+2u} < 1$ then x can be exactly recovered by solving

$$(P_0) \quad \hat{x} = \arg\min_{x} ||x_{K^c}||_0 \qquad s.t \quad y = \Phi x \tag{10}$$

(but the above optimization is NP-hard) and if a more stringent condition is satisfied, then the optimization $P_1 \equiv P_0$

$$(P_1) \quad \hat{x} = \arg\min_{x} ||x_{K^c}||_1 \qquad s.t \quad y = \Phi x \tag{11}$$

The above optimization is known as 'Modified Recovery'. Using the above conditions on δ the authors argued that the modified recovery performs better than the normal recovery for certain ranges of u, k and especially so when $u << s, k \approx s$ and showed experimental results. They do not give any bounds on the number of samples, m required for exact recovery.

We obtain bounds on m in this case, by using the approach in [4] used to establish (6).

Theorem 1: Suppose m, n and $0 < \delta < 1$ are given and Φ is a random gaussian matrix, i.e $\Phi_{ij} \sim N(0, 1/m)$ then there exist constants c_1, c_2 depending only on δ such that RIP - holds with the prescribed δ with probability $\geq 1 - 2e^{-c_2 m}$ for any k, u obeying

$$c_1\left(u\left(\log\frac{n}{u} + 1 + \log\frac{9}{\delta}\right) + k\log\frac{12}{\delta}\right) < m \tag{12}$$

Proof: We know that for an $p-$dimensional subspace, the matrix Φ will fail to hold (RIP) with probability

$$\leq 2(9/\delta)^p e^{-c_0(\delta/2)m}$$

In our case, the dimension of each subspace that x can lie in is $k + u$, where $k = |K|$ and the total number of such subspaces are ${}^nC_u \leq (en/u)^u$. For all these subspaces, Φ will fail to hold RIP with probability

$$\begin{aligned} &\leq \quad 2(en/u)^u(9/\delta)^{(k+u)}e^{-c_0(\delta/2)m} \\ &= \quad 2e^{-c_0(\delta/2)m+u[\log(en/u)+\log(9/\delta)]+k\log(9/\delta)} \end{aligned}$$

Thus, for

$$m > c_1\left(u\left(\log\frac{n}{u} + 1 + \log\frac{9}{\delta}\right) + k\log\frac{9}{\delta}\right) \tag{13}$$

Φ will fail to hold RIP with probability $2e^{-c_2 m}$ ■

We now compare the bounds in (6) and (13). Noting that $s = k + u$, the above becomes

$$m > c_1\left(u\left(\log\frac{n}{u} + 1\right) + s\log\frac{9}{\delta}\right)$$

Letting δ be the same in both cases, the constants become equal, and it is a straightforward observation that RHS in (6) is greater that RHS in the above equation. Hence, when some knowledge about the support of x is available, the 'Modified Recovery' approach requires lesser number of samples than the normal recovery which is agnostic to any such knowledge.

The above approach still recovers both the known and unknown supports. If one were totally uninterested in recovering the known support then one can acheive a further reduction in sampling rate, which we discuss in next section.

B. Selective-CS

If some support of x denoted as K is always present, then we can reject the signal residing in this subspace and acquire only the signal of interest i.e. x residing in $K^{\perp}$, which is sparser than the original signal. This approach further decreases the number of compressive samples required. In practice K can thought of the support of TV bands in a given location which are known beforehand. To achieve this, we alter our sensing matrix Φ as follows. Let Ψ_K denote the subspace in which the known components reside. The projection matrix onto the space spanned by the columns of $\Psi_K$is given by

$$P_K = \Psi_K(\Psi_K'\Psi_K)^{-1}\Psi_K' \tag{14}$$

The projection matrix onto the space orthogonal to $\Psi_K$is given by the annihilator matrix, $P_{K^\perp} = I - P_K$. We choose our new sensing matrix as $\tilde{\Phi} = \Phi P_{K^\perp}$.

We now show that $\tilde{\Phi}$ preserves the structure in $K^{\perp}$ while simultaneously cancelling out signals from K given that Φ provides a stable embedding for $K^{\perp}$.

Theorem 2: Suppose that Φ provides a $\delta-$stable embedding of $(S_{K^\perp}, 0)$ where $S_{K^\perp}$ = set of all $u-$sparse signals residing in $K^{\perp}$. Define P_K and $P_{K^\perp}$ as above, then

$$\tilde{\Phi}x = \tilde{\Phi}x_{K^\perp} \qquad and \tag{15}$$

$$(1-\delta)||x_{K^\perp}||^2 \leq ||\tilde{\Phi}x_{K^\perp}||^2 \leq (1+\delta)||x_{K^\perp}||^2 \tag{16}$$

Proof: Observe that x can be decomposed into orthogonal components as $x = x_{K^\perp} + x_K$. Therefore,

$$\tilde{\Phi}x = \Phi P_{K^\perp}x = \Phi P_{K^\perp}(x_K + x_{K^\perp}) = \Phi P_{K^\perp}x_{K^\perp} = \tilde{\Phi}x_{K^\perp} \tag{17}$$

which establishes (15). Also,

$$\tilde{\Phi}x = \Phi P_{K^\perp}x = \Phi x_{K^\perp} \tag{18}$$

combining this with (17) we have

$$\tilde{\Phi}x_{K^\perp} = \Phi x_{K^\perp} \tag{19}$$

Since Φ provides a stable embedding for $x_{K^\perp}$ from (3), we have

$$(1-\delta)||x_{K^\perp}||^2 \le ||\Phi x_{K^\perp}||^2 \le (1+\delta)||x_{K^\perp}||^2$$

combining this with (19), we obtain (16) which completes the proof.

■

Thus the sensing matrix $\Phi P_{K^\perp}$ simultaneously acts as a stop band filter. The signal reconstruction remains same as in (5) but for Φ replaced with $\tilde{\Phi}$.

In [6] the authors propose Cancel-then-Recover approach to cancel out interference or unwanted signals when their subspace is known. In this approach, the received wideband signal is compressively sampled and then projected onto the subspace orthogonal to the unwanted signals inorder to cancel them i.e they use $P_{\Omega^\perp}\Phi$ as their sensing matrix, where $\Omega = \Phi\Psi_K$ annd $P_{\Omega^\perp}$ is an orthogonal projection operator onto the othogonal complement of rowspace(Ω) and is given by $P_{\Omega^\perp} = I - \Omega(\Omega^T\Omega)^{-1}\Omega^T$. For the $l_0 - norm$ optimization to yield unique solution, Cancel-then-Recover approach requires Φ to obey RIP of order $k+2u$, which is also the case in modified recovery which is stricter than $2u$, as required in our case. Thus 'cancel and recover' should perform as well as modified recovery but not as good as selective acquisition as shown later in results.

The number of samples, m required in this approach is given by (6) with s replaced by u and since $u < s$, it is lesser than that required in both normal CS and modified recovery.

IV. Results and Discussion

In this section, we first evaluate the performance of normal recovery, modified recovery and Selective-CS. In all our experiments, rather than l_1 minimization we use the iterative greedy CoSaMP (Compressive Sampling Matching Pursuit) algorithm proposed by [7] as it is one of the fastest algorithms. Its running time is $O(mn)$ whereas the running time form for convex l_1 optimization is $O(n^6)$ for $\Phi \sim n \times n$ matrix. Other greedy algorithms include orthogonal matching pursuit (OMP), tree based OMP (TOMP) etc. CoSaMP can be used without any modifications for normal recovery as well as Selective-CS but in the case of modified recovery, we had to modify it such that at each iteration we seek a sparse solution while simultaneously forcing it to contain the known set.

A. *Experiment 1 (Comparision of various CS schemes)*

1) We generate test signals of length $n = 500$. We fix $m = 100$ and the unknown sparsity, $u = 5$
2) Vary the known sparsity, k from 0 to 20 times u.
3) For each k, repeat the following 500 times
4)
   - Generate the support, S of size $s = k + u$, uniformly at random from $[1, n]$.
   - Generate the unknown support, U of size u, uniformly at random from S. The known support, S of size s is given by $S\backslash U$.
   - Generate the $m \times n$ random gaussian sampling matrix, Φ with $\Phi_{ij} \sim N(0, 1/m)$. Add gaussian noise to x such signal SNR is 15 dB and generate the compress samples, $y = \Phi x$.
   - From y, recover x using Normal recovery, Modified recovery, Selective Acquisition.Denote the recoverd signal as $\hat{x}$.
5) For each k, estimate the normalized mean square error(NMSE) and running time for the three methods averaged over 500 iterations. We also compute NMSE using the least squares assuming the entire support of x is known and also using Cancel-then-Recover approach.

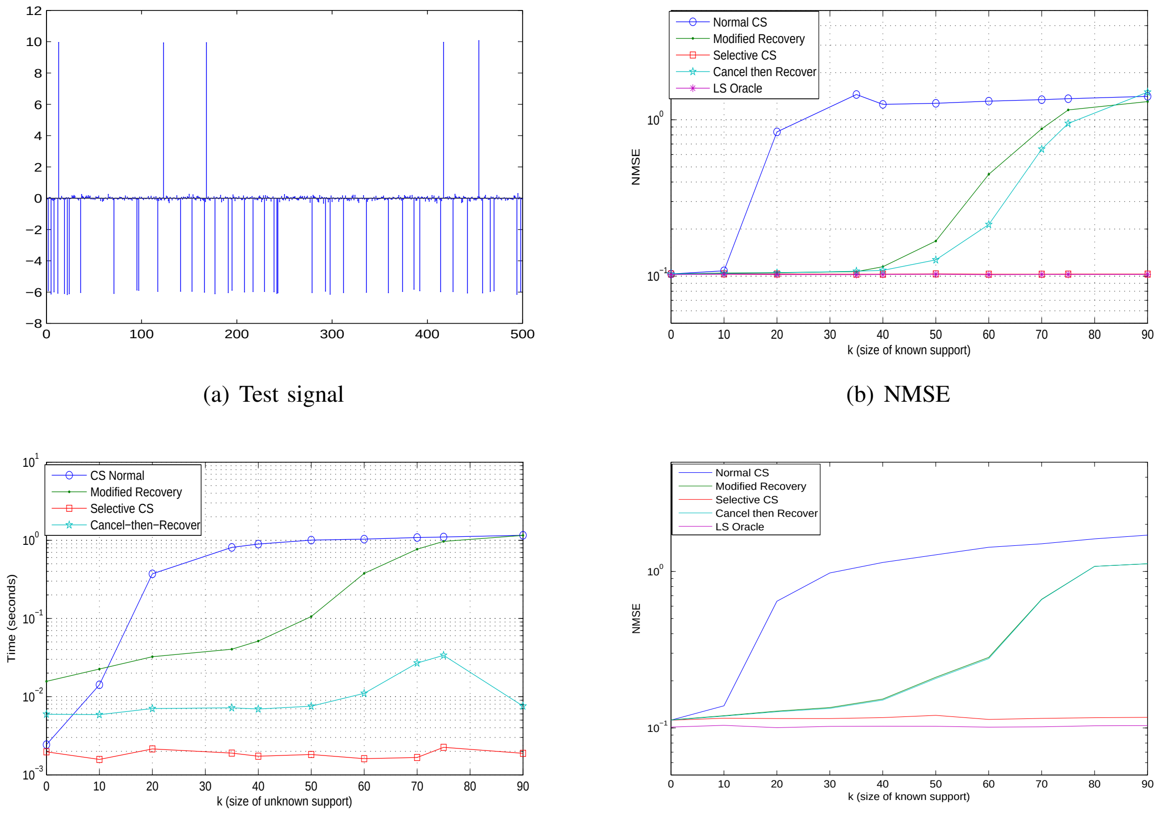
Normalized mean square error is defined as

$$NMSE = ||x - \hat{x}||/||x||$$

For fair comparision between the various methods, we use

$$NMSE = ||(x - \hat{x})_{K^\perp}||/||x_{K^\perp}||$$

i.e. we compute NMSE over unknown portion of the recovered signal.



(a) Test signal

(b) NMSE

(c) Recovery time

(d) l_1 minimization

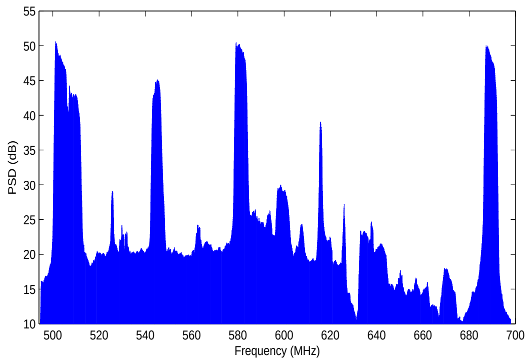
Fig. 1. Comparision of various CS schemes

Fig 1(a) shows a sample test signal ($k = 20$) used in the experiment. Fig 1(b) shows that the NMSE for normal CS increases very quickly with the size of known support k as it is agnostic to it. Modifed recovery and Cancel-then-Recovery approach do better than normal CS. The selective CS approach performs better than the above three and peforms nearly as well as LS-Oracle which is the best one can do. On a side note, cancel-then-recover seems to perform slightly better than modified recovery using CoSaMP, even though their performance should theoretically be identical. Hence, we repeated this experiment using convex l_1 optimization using 'cvx' toolbox [8] and present the results in fig 1(d) averaged over 10 instances. Fig 1(d) shows that cancel-then-recover and modified recovery have same NMSE.

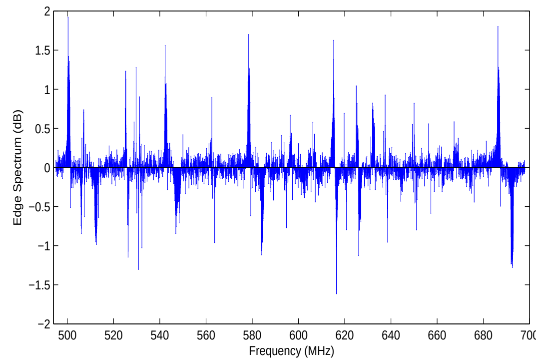
Fig 1(c) shows similar pattern in running times too. Selective CS performs faster than the other approaches as k increases. In case of running time, modified recovery takes more time than cancel-then-recover because of the increased complexity in forcing the sparse solution to contain the known set.

B. Experiment 2 (CS on Spectrum Data)

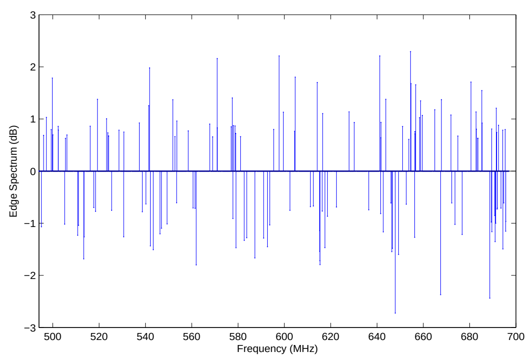
In this experiment we use CS to acquire the UHF band 494 to 698 MHz (bandwidth = 204 MHz) consisting of 34 channels ranging from channel number 18 to 51. Each channel is sensed individually and the PSD samples are then stacked up to form the samples of a single wideband on which CS is applied using our scheme described in $2C$. The frequency resolution is 50 kHz and the wideband signal length, $n = 4046$. We reconstruct back the edge spectrum from compressed samples using normal CS, modified recovery and selective CS for various levels of compression denoted m/n where m is length of compressed samples. From initial sensing (using energy detector on individual channels), we detect channels 19,20,26,32,50 to be occupied but we assume that only 19,20,26,32 to be occupied and use this as our knowledge in subsequent sensing. We compute NMSE over unknown portion of the recovered signal and the probability of detection for channel 50 (686-692 MHz) averaged over 100 instances using the three methods mentioned.



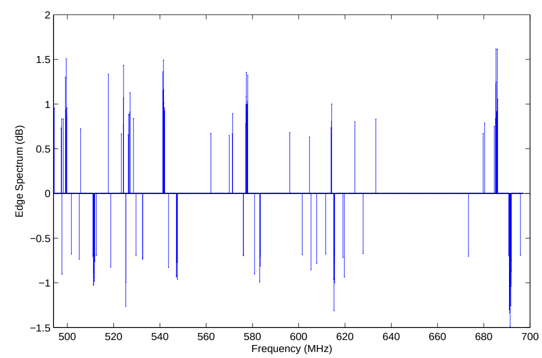
(a) Sample PSD



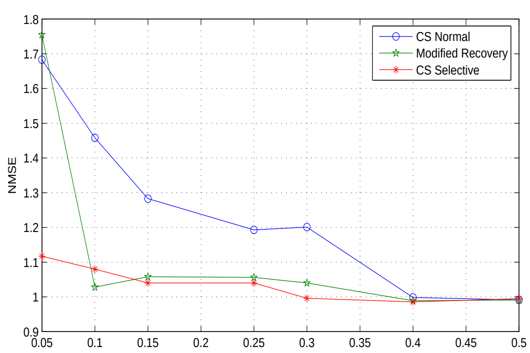
(b) Edge Spectrum



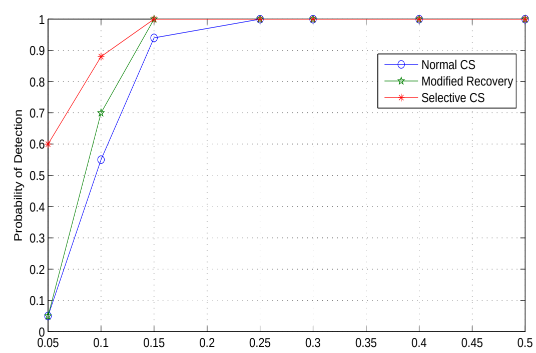
(c) CS, m/n = 0.1



(d) CS, m/n = 0.4



(e) NMSE



(f) Probability of Detection, Channel 50

Fig. 2. CS using Spectrum Data

Fig 2(a), 2(b) show the smoothed PSD and edge spectrum obtained from initial sensing, using moving average filter of order 21. Fig 2(c), 2(d) show typical edge spectrum reconstructions using normal CS for m/n = 0.1 and 0.4. Fig 1(c) shows that for higher m/n all the approaches perform identically and as m/n decreases, selective CS performs better than modified recovery which in turn does better than normal CS. The NMSE tends to be high because of compressive nature of edge spectrum.

Fig 1(d) shows the probability of detection of channel 50 (SNR is about 25-30 dB) using its edge energy. The probability of false alarm, P_f is fixed at 0.05. The detection procedure is as follow. Let $\hat{z_{50}}$ correspond to the samples of channel 50 in the reconstructed edge spectrum (we used about 20 samples for rising edge and 20 for falling edge). We compare the edge energy against a threshold. To compute probability of false alarm, one needs a vacant channel 50 but since it is always occupied, we use channel 51 as it unoccupied template. Similar results as in the case of NMSE can be seen here. Normal CS achieves P_d of 1 for $m/n = 0.25$ which is encouraging. Since P_d degrades with SNR, one needs to repeat this for lower SNRs but practical problem is that all the channels occupied in Madison have SNR > 20 dB and have similar performance. Nevertheless, results are still encouraging. Also, the time taken for edge spectrum reconstruction using CoSaMP (using reasonable tolerance) for $m \times n$ = 1000×4000 is about a second which is reasonable.

V. COMPRESSED DETECTION

In this section we pose the primary user detection problem directly on the compressed samples and thereby avoid the recovery of edge spectrum.

A. CS Binary Detection

[4] addresses the problem of binary signal detection in CS framework when the signal is known. We discuss their main results briefly. They aim to distinguish between two hypotheses:

$$\begin{aligned} H_0 : y &= \Phi n \\ H_1 : y &= \Phi(s + n) \end{aligned}$$

where $s \in R^N, n \sim N(0, \sigma^2 I_N)$ is i.i.d Gaussian noise and Φ is random real gaussian matrix as usual and s is known. Upon taking log likelihood ratio of y, the decision rule obtained is

$$t := y^T (\Phi\Phi^T)^{-1} \Phi s \quad \underset{H_0}{\overset{H_1}{\gtrless}} \quad \gamma \tag{20}$$

where t is the test statistic and its pdf is

$$\begin{aligned} H_0 : t &\sim N(0, \sigma^2 ||P_{\Phi^T} s||^2) \\ H_1 : t &\sim N(||P_{\Phi^T} s||^2, \sigma^2 ||P_{\Phi^T} s||^2) \end{aligned}$$

where $P_{\Phi^T} = \Phi^T (\Phi\Phi^T)^{-1} \Phi$

Setting the probability of false alarm $P_F = \alpha$ yields the probability of detection as

$$\begin{aligned} P_D(\alpha) &= Q(Q^{-1}(\alpha) - ||P_{\Phi^T} s||/\sigma) \\ &\approx Q(Q^{-1}(\alpha) - \sqrt{M/N} ||s||/\sigma) \\ &= Q(Q^{-1}(\alpha) - \sqrt{M/N} \sqrt{SNR}) \end{aligned}$$

The above equation tells us that as the number of samples decrease, the probability of detection too decreases. For example, if $M = 0.5N$, then the signal must have twice the original SNR to achieve same $P_D(\alpha)$ i.e the performance loss is 3 dB in terms of SNR.

B. CS Multi-User Detection

Our problem consists of determining if a particular channel is occupied by a primary user or not. We assume that the primary users undergo i.i.d flat rayleigh fading and that their signal energy is known. Let the bandwidth of each channel be B, number of channels be C and number of primary users be U where $U < C$. The primary users are orthogonal to each other in frequency. The recieved baseband (which is wideband) signal can be represented in discrete domain as

$$x = \sum_{u=1}^{U} h_u s_u + n \tag{21}$$

where $s_u \in C^n$ is the signal corresponding to u^{th} user, $n \sim CN(0, \sigma^2 I_N)$ and h_u's are i.i.d, $\sim CN(0, I_N)$ The compressed samples are given by

$$y = \Phi x = \Phi(\sum_{u=1}^{U} h_u s_u + n) \tag{22}$$

where Φ is random real gaussian matrix with $\Phi_{ij} \sim N(0, 1/M)$

When $U = C = 1$, i.e we are sensing a single channel, the problem is to distinguish between the hyptheses

$$\begin{aligned} H_0 : y &= \Phi n \\ H_1 : y &= \Phi(hs + n) \end{aligned}$$

The pdf of y under the hypotheses is

$$\begin{aligned} H_0 : y &\sim N(0, \sigma^2 \Phi \Phi^T) \\ H_1 : y &\sim N(0, \Sigma_s + \sigma^2 \Phi \Phi^T) \end{aligned}$$

where $\Sigma_s = E[(\Phi hs)(\Phi hs)^H] = \Phi E[(hs)(hs)^H]\Phi^T = \Phi diag(|s_1|^2, \ldots, |s_n|^2)\Phi^T = \Phi \Lambda_s \Phi^T$

Upon taking log likelihood ratio of y, we obtain the decision rule as

$$t := y^H ((\sigma^2 \Phi \Phi^T)^{-1} - (\Sigma_s + \sigma^2 \Phi \Phi^T)^{-1}) y \underset{H_0}{\overset{H_1}{\gtrless}} \gamma \tag{23}$$

where t is the test statistic. The above is a compressed energy detector.

If $\Phi = I_N$ i.e. we use nyquist rate sampling instead of compressive sampling, then $y = \Phi x = Ix = x$ and the detection rule would take the form,

$$\begin{aligned} x^H ((\sigma^2 I_N)^{-1} - (\Lambda_s + \sigma^2 I_N)^{-1}) x &\underset{H_0}{\overset{H_1}{\gtrless}} \gamma \\ x^H diag\left(\frac{|s_1|^2}{|s_1|^2 + \sigma^2}, \ldots, \frac{|s_n|^2}{|s_n|^2 + \sigma^2}\right) x &\underset{H_0}{\overset{H_1}{\gtrless}} \gamma \\ x^H \tilde{x} &\underset{H_0}{\overset{H_1}{\gtrless}} \gamma' \end{aligned}$$

When s is unknown, we assume $s_1 = \ldots = s_N = \sqrt{\mathcal{E}}$ and the above detection rule takes the form

$$x^H x \underset{H_0}{\overset{H_1}{\gtrless}} \gamma'' \tag{24}$$

which is the normal energy detector.

Under unknown s, $\Lambda_s = \mathcal{E} I$ and (23) takes the form

$$\begin{aligned} y^H ((\sigma^2 \Phi \Phi^T)^{-1} - (\mathcal{E} \Phi \Phi^T + \sigma^2 \Phi \Phi^T)^{-1}) y &\underset{H_0}{\overset{H_1}{\gtrless}} \gamma \\ y^H \left(\frac{\mathcal{E}}{\sigma^2 + \mathcal{E}} (\Phi \Phi^T)^{-1}\right) y &\underset{H_0}{\overset{H_1}{\gtrless}} \gamma \\ y^H (\Phi \Phi^T)^{-1} y &\underset{H_0}{\overset{H_1}{\gtrless}} \gamma' \end{aligned}$$

Thus the test statistic is $t := y^T (\Phi \Phi^T)^{-1} y$, where $(\Phi \Phi^T)^{-1}$ is non-diagonal. t is distributed as χ^2 under either hypothesis. I'm not sure how to obtain the actual pdf of t under either hypothesis using which the expressions for probability of false alarm, P_f and probability of detection, P_d can be obtained.

When $U, C > 1$, in the case of uncompressed detection, to determine whether a particular channel,c is occupied or not, one normally uses a pass band filter tuned to its bandwidth and uses the energy detector. But in case of compressed sensing, we acquire the entire wideband signal. To use the above energy detector for say channel u, one has to first obtain the signal component Φx_u corresponding to channel u. This can be achieved using a projection operation as follows.

Let Ψ_u be an $N \times K_u$ matrix whose columns form an orthonormal basis for channel u, and define the $M \times K_u$ matrix $\Omega = \Phi\Psi_u$. Thus the desired projection operator is given by

$$P_\Omega = \Omega(\Omega^H\Omega)^{-1}\Omega^H \tag{25}$$

We can decompose x into orthogonal components as x as $x = x_u + x_{u^\perp}$ where x_u is the signal correspong to user u. Then

$$\begin{aligned} P_\Omega\Phi x &= P_\Omega\Phi(x_u + x_{u^\perp}) \\ &= \Phi x_u + P_\Omega\Phi x_{u^\perp} \end{aligned}$$

where Φx_u is the desired signal and $P_\Omega\Phi x_{u^\perp}$ is the interference caused by other signals. From the above equation, it can be seen that such a projection operation preserves the structure in desired signal while not completely cancelling out undesired signals which causes interference. [6] have obtained the bound on the second term for such a projection operation as

$$||P_\Omega\Phi x_{u^\perp}|| \leq \frac{\delta}{1-\delta}\sqrt{1+\delta}||x_{u^\perp}|| \tag{26}$$

Denoting $y_u = P_\Omega\Phi x$, one can now use the following subspace compressed energy detector

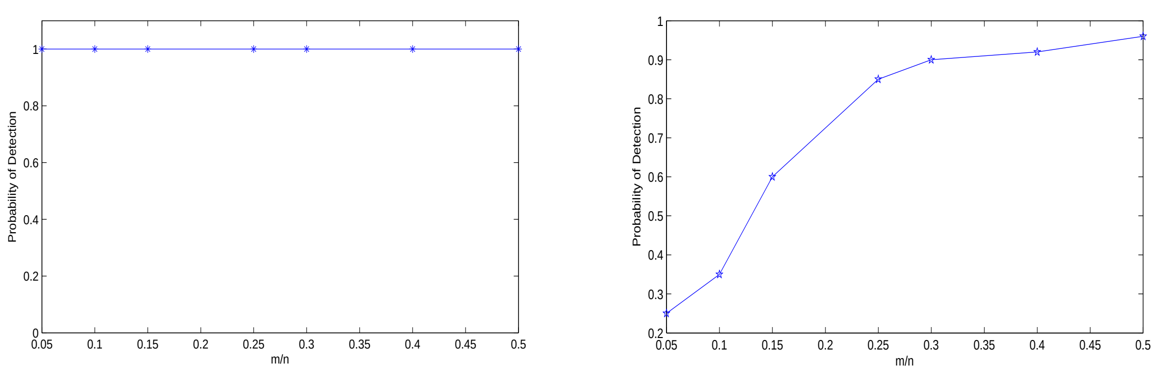
$$t := y_u^H(\Phi\Phi^T)^{-1}\Phi y_u \quad \underset{H_0}{\overset{H_1}{\gtrless}} \quad \gamma \tag{27}$$

to determine if the channel u is occupied or not. But the presence of interference in y_u, degrades the performance of this subspace detector. One approach to mitigate the interference to some extent is to employ selective acquisition as described earlier but it in the case of compressive signals, it may not do that well.

Going back to the CS architecture in [2], in which the edge spectrum is sparse, we note that the above detection scheme is infeasible because y doesn't have linear representation interms of z. The relation between y and z is $r_y = \Phi_I\Psi z$ where r_y is autocorrelation of y as described in section $2B$. Thus a projection operation such as P_Ω is infeasible.

VI. Results and Discussion

We provide some experimental results for unknown signal detection in CS framework. We use a paradoxical CS scheme i.e we have the PSD samples, whose differential gives the edge spectrum,z which we compress. From the compressed samples, we try to detect the presence of primary users. In this scheme, $\Psi = I$ and $y = \Phi z$. Number of simulations used is 300.



(a) Single Band (b) Wideband

Fig. 3. CS detection

A. Compressed Detection, Single User

We sense channel 50 using CS. The detector used is (27) where $y_u = y$. For $P_f = 0.05$ the probability of detection as observed in 3(a) is 1 even for m/n as low as 0.05, which is beacuse it has high SNR.

B. Compressed Detection, Multi User

We sense all the channels simultaneously using CS, from which we detect whether channel 50 is occupied. The detector used is (27) with y_u = projection of y onto subspace of channel 50. The detector performance for $P_f = 0.05$ as observed in 3(b) has degraded. Even for $m = 0.5n$, P_d is only around 0.95 even when the SNR is high. Naturally, it gets worse for lower SNRs.

Normal CS followed by edge spectrum reconstruction and 'edge energy' detection as in fig 2(f) performs much better than this. Also the running time using compressed detector is

$$T_2 = c \times t_2$$

and for the former is

$$T_1 = T_{rec} + c \times t_1$$

where c =number of channels to detect, t_2 is time taken to detect a channel using compressed detector (it is hampered by the projection operation and matrix multiplications needed to compute the test statistic (27)), t_1 is time taken to detect a channel using normal detector, T_{rec} is the time taken for CS reconstruction. For $m \times n$ = 1000×4000 and $c = 100$, $T_1 \approx 1 + 0.2 = 1.2s$ and $T_2 \approx 3s$.

In fact, to detect microphones signals (bandwidth = 200 kHz), one needs to first detect TV transmissions and then divide the unused spectrum into narrow chunks to detect the presence of microphones, in which case c can be even higher and T_2 increases rapidly compared to T_1.

Thus from both the perspectives of Detector Performance and Running time, the former approach does better than compressed detector approach.

VII. Conclusion

In this work, we take the first steps in determining the practical feasibility of CS for detection of spectrum holes based on sparseness of edge spectrum. Results show that using normal CS, for $P_f = 0.05$, P_d of an occupied channel approaches 1 using compressive sampling rate, $m = 0.25n$ which is encouraging although it is for high SNRs. When some knowledge of spectrum occupancy is available, depending on the amount of knowledge, P_d can be improved as shown using modified recovery and selective CS. Reconstruction of edge spectrum using greedy CoSaMP takes about a second (using reasonable tolerance), for Φ of size 1000×4000 in the case of normal CS which is acceptable. Literature has shown that further improvements in P_d can be obtained using collabarative sensing which could be explored further.

In the later part of work, we study the feasibility of the detection problem directly from the compressed samples and thereby avoid the reconstruction phase. The problem seems to be infeasible based on the assumption that edge spectrum is sparse because the compressed signal, y doesn't have a linear representation in terms of its sparsity basis. Even if there exists such a sparse basis, multi user detection suffers from interference and incase the signal is rather compressible but not sparse in its basis, interference can be severe as shown in results. Also, even if y is linear in terms of its basis, the detection time is hampered by the complexity of the test statistic.

Thus the former approach provides a practical CS scheme for the detection of spectrum holes but at low SNRs, the edges tend to be weaker, and this approach may not scale well.

VIII. Acknowledgements

The author would like to thank Sayandeep Sen and Suman Banerjee for their support and feedback on the project.

REFERENCES

[1] Z. Tian and G. Giannakis, "Compressed sensing for wideband cognitive radios," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, vol. 4, Honolulu, Hawaii, USA, Apr. 2007, pp. 1357-1360*.

[2] Yvan Polo, Ying Wang, Ashish Pandharipande, Geert Leus, "Compressive wide-band spectrum sensing," *pp.2337-2340, 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, 2009*.

[3] E. J. Candes and T. Tao. "Decoding by linear programming". *IEEE Trans. Inform. Theory, 51 4203-4215, 2004*.

[4] R. G. Baraniuk, M. Davenport, R. A. DeVore, and M. B. Wakin, "A Simple Proof of the Restricted Isometry Property for Random Matrices (aka "The Johnson-Lindenstrauss Lemma Meets Compressed Sensing,"), *Constructive Approximation, January 2008*.

[5] N. Vaswani and W. Lu, "Modified-CS: Modifying compressive sensing for problems with partially known support," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT), Seoul, Korea, Jun 2009*.

[6] Mark A. Davenport, Petros T. Boufounos, Michael B. Wakin, and Richard G. Baraniuk, "Signal processing with compressive measurements", to appear in *Journal of Selected Topics in Signal Processing, 2010*.

[7] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples", *Appl. Comp. Harmonic Anal., Vol. 26, pp. 301–321, 2008*.

[8] M. Grant and S. Boyd, CVX: Matlab Software for Disciplined Convex Programming, version 1.21, http://cvxr.com/cvx.