New Algorithms for Wideband Spectrum Sensing Via Compressive Sensing

Shwetank Mistry and Vinod Sharma
Department of Electrical and Communication Engineering
Indian Institute of Science, Bangalore, 560012, India.
Email: {shwetank,vinod}@ece.iisc.ernet.in

***Abstract*—We consider the problem of spectrum sensing in a Cognitive Radio (CR) system when the primaries can be occupying a few subbands in a wideband spectrum. Since the primary signal dimension is large, Nyquist rate can be very high. Compressive sensing (CS) can be useful in this setup. However a CR system needs to operate at a very low $SNR(\sim -20dB)$ where the compressive sensing techniques are usually not successful. Combining them with statistical techniques can be useful. But this has been difficult because the statistics of the parameters obtained from the recovery algorithms (e.g., OMP) are not available. We use a suboptimal recovery algorithm COR for which the statistics can be easily approximated. This allows us to use Neyman Pearson technique as well as sequential detection techniques with CS. The resulting algorithms provide satisfactory performance at -20 dB SNR. In fact COR's recovery performance is itself better than OMP at low SNR. We also modify the algorithm for the scenario when the channel gains and the noise variance may also not be available.**

keywords : Compressive Sensing, Cognitive Radio, Spectrum Sensing, SPRT.

I. INTRODUCTION

The need for higher data rate is increasing due to new multimedia applications. But the current static frequency allocation schemes can not accommodate these requirements. However at any time and place, the occupancy of the licensed spectrum is very low. Cognitive Radio (CR) technology ([12], [8]) allows opportunistic use of available spectrum by unlicensed users. The main task of CR is to sense the spectrum and then use it efficiently.

A number of different methods [27] have been proposed for identifying the presence of primary signal transmissions by CR users, e.g., Energy detection, Matched Filter detection and Cyclostationarity detection ([19], [24]) for narrow band sensing. But for sensing an ultra-wideband spectrum the analog front-end becomes very complex as many parallel filters are required to detect the spectrum using the narrow band sensing techniques. To sense the ultra-wideband spectrum directly the RF front-end should be able to sample at the Nyquist rate. As we are interested in the band of the order of a few GHz the analog to digital converter (ADC) should be able to sample at a very high rate. Clearly, this is a major implementation problem especially when one needs to implement it in relatively inexpensive CR units.

Often the occupancy by the primary users in the ultra-wideband spectrum is sparse [3]. For sparse signal recovery, compressive sensing (CS) at sub-Nyquist-rate sampling offers reliable signal recovery ([6], [3], [11]). Recently CS has been used for detection of wideband spectrum [15]$-$ [21] which use CS recovery algorithms to estimate the wideband spectrum and then use it to detect the spectrum occupancy. In [21] co-operative techniques are also used to improve the performance and they show Receiver Output Characteristics (ROC) at 0 dB via simulations but do not provide any theoretical results for the detector performance. Signal cyclostationarity which exploits the sparsity of the two-dimensional spectral correlation function (SCF) is proposed in [22], [18]. Performance is studied via simulations at SNR between 0-5 dB. For detecting a frequency hopping signal [26] uses CS techniques at $SNR \geq 0$. A Compressive Detector at $SNR \geq 10$ dB is studied in [5].

We use compressive sensing for detection of spectral holes. Our aim is to develop efficient, low complexity, practical algorithms for wideband spectrum sensing even at low $SNR(\leq -10dB)$. We develop a new simple algorithm called COR algorithm. We compare it to an orthogonal matching pursuit (OMP) based algorithm [2] and show that this computationally simpler algorithm actually performs better at low SNR. Next we obtain an improved sequential version of the algorithm. We theoretically analyze these algorithms and compare with simulations. We also consider the scenario when the channel gains, sparsity and noise variance are not known to the CR. To the best of our knowledge, these issues have not been addressed previously for CS based algorithms for spectrum sensing.

Rest of the paper is organized as follows. In Section II we present our system model. Section III discusses CS algorithms. In Section IV we present COR algorithm. In Section V we compare COR to OMP algorithm and show that this computationally simpler algorithm performs better at low SNR. We also consider the scenario when the channel gains, sparsity and noise variance are unknown for COR. In Section VI we obtain an improved sequential version of the algorithm. We theoretically analyze these algorithms and compare with simulations. Section VII concludes the paper.

We will use the following notation:

- $\mathbf{a}_i$: $i-$th column of a matrix $\mathbf{A}$.
- a_{ij} : $(i,j)-$th element of matrix $\mathbf{A}$.
- $\mathcal{N}(\mu, \sigma^2)$: real Gaussian distribution with mean μ and variance σ^2.
- $E\{\cdot\}$: expectation of the operand.

- $Var\{\cdot\}$: variance of the operand.

II. SYSTEM MODEL

Consider an ultra-wide frequency band that is licensed to primary users. The entire wideband spectrum is divided into N non-overlapping narrowband subchannels centered at $\{f_m\}$, $m = 0, 1, .., N$. The locations of these bands are pre-defined and known, as in multi-band radios and OFDM systems. At any time some of these subbands are occupied by the primary users. The temporarily idle subbands are termed spectral holes and are available for opportunistic spectrum access by secondary users.

A CR node senses the wideband spectrum in order to identify the spectral holes. We assume that each subband is an additive white Gaussian noise (AWGN) channel. Our problem is to sense the primary spectrum at low $SNR(< -10$ dB) and detect if each subband is free or not. The received signal at the CR is

$$y(t) = h(t) * x(t) + w(t), \tag{1}$$

where $w(t)$ is white Gaussian noise with zero mean and power spectral density σ_w^2 and $h(t)$ is the channel impulse response.

We represent the continuous time primary signal $x(t)$ in discrete time as $\mathbf{x} = \{x(0), x(T_s), ..., x((N-1)T_s)\}$, which is an N length sequence sampled at uniform rate $W = 1/T_s$, where W is the baseband bandwidth. We denote by $\mathbf{X}$ the Discrete Fourier Transform (DFT) of vector $\mathbf{x}$. Thus $\mathbf{X} = \mathbf{F}\mathbf{x}$, where $\mathbf{F}$ is an $N \times N$ DFT matrix. The signal-to-noise-ratio SNR, E= signal energy/ noise energy, where in a time window of NT_s, where signal energy $= \sum_{n=0}^{N-1} |x[n]|^2 = \sum_{n=0}^{N-1} |X[n]|^2$ and noise energy is $N\sigma^2$. We assume that the signal is S sparse in DFT domain, i.e., out of N atmost $S \ll N$ entries of $\mathbf{X}$ are nonzero. We have the following two possibilities for any channel,

$$|X[n]| = \begin{cases} A & under\ H_0, \\ 0 & under\ H_1. \end{cases}$$

Our problem is to sense the noisy wideband primary signal received by the CR and identify spectral holes in the given wideband. We want to achieve robust detection even at $SNR = -20$ dB. However, for a wideband system, the sampling rate can be very high. To design a CR system with modest costs, this sampling rate can be prohibitive. Thus, we consider compressive sensing techniques that can reduce the sampling rate when the number of active primary users at a time is small.

III. COMPRESSIVE SENSING ALGORITHMS

Compressive sensing allows us to recover a signal completely even if it is sampled at a rate less than the Nyquist rate if the signal is sparse in some domain. In our case primary spectral usage is sparse in the frequency domain.

In the following we briefly present CS and then explain how we will use it in our context. Compressive sensing considers equation

$$\mathbf{v} = \mathbf{\Phi}\mathbf{x}, \tag{2}$$

where $\mathbf{x}$ is an N dimensional vector and $\mathbf{\Phi}$ is an $M \times N$ random projection matrix with $M < N$. CS theory [6], [11] implies that one can reconstruct $\mathbf{x}$ from $\mathbf{y}$ if $M \geq CS \log N$, where C is a constant.

In our case equation (2) is (we will assume $h(t) \equiv 1$ until further notice),

$$\mathbf{y} = \mathbf{\Phi}(\mathbf{x} + \mathbf{w}) = \mathbf{\Theta}(\mathbf{X} + \mathbf{W}), \tag{3}$$

where $\mathbf{X} = \mathbf{F_{inv}}\mathbf{x}$, $\mathbf{\Theta} = \mathbf{\Phi}\mathbf{F_{inv}}$, $\mathbf{W} = \mathbf{F_{inv}}\mathbf{w}$ and $\mathbf{F_{inv}}$ is the IDFT matrix.

The reconstruction requires a non-linear algorithm to find S-sparse signal $\mathbf{X}$ from $\mathbf{y}$. There are many such algorithms in literature survey for compressive sensing. Among the existing reconstruction algorithms, Basis Pursuit (BP) [6] uses ℓ_1 optimization. Another class of algorithms are iterative algorithms. These include matching pursuit (MP) and orthogonal matching pursuit (OMP) [23]. Their followers are stagewise OMP (StOMP) [7] and the regularized OMP (ROMP) [13]. The reconstruction complexity of these algorithms is much lower than ℓ_1 optimization. However these algorithms require more measurements for perfect recovery. There are algorithms such as subspace pursuit (SP) [4] and compressive sensing matching pursuit (CoSaMP) [14] that use the idea of backtracking.

These algorithms give an estimate $\hat{\mathbf{X}}$ of $\mathbf{X}$. The performance of some these algorithms in a noisy setup has been studied in [2], [14] and it is found that to recover the signal they require $SNR \geq 10dB$. We now consider one commonly used recovery algorithm OMP. OMP works as follows: Initially it finds a column of $\mathbf{\Theta}$ which has maximum correlation with $\mathbf{y}$. Then from $\mathbf{y}$ it removes the orthogonal projection of $\mathbf{y}$ on that column. It repeats the above procedure until it finds S columns. Then $\hat{\mathbf{X}}$ is the projection of $\mathbf{y}$ on these columns.

We use a computationally simpler but sub-optimal algorithm. An advantage is that we can approximate distributions of its projections under different hypothesis and then use the statistical techniques to detect holes. We call this algorithm COR. In COR we find the correlation of $\mathbf{y}$ with each of the columns of $\mathbf{\Theta}$ and find the statistics of these correlations. Thus it is not an iterative procedure like OMP and hence it is simple to find the distribution of correlation for each column. As an $\mathbf{X}$ recovery algorithm COR is suboptimal. But along with statistical methods it can provide a better procedure at low SNR than OMP.

IV. COR ALGORITHM

We define

$$\mathbf{cor} = \mathbf{\Theta^h}\mathbf{y}. \tag{4}$$

We take each entry of $\mathbf{\Phi}$ as i.i.d. $\mathcal{N}(0, 1/M)$. Since $\mathbf{F_{inv}}$ is an unitary matrix, $\mathbf{\Theta}$ is a Complex Gaussian matrix. We will use $cor(i)$, the ith component of $\mathbf{cor}$, to detect if the ith channel is occupied or not.

Expanding the ith component of **cor**,

$$cor(i) = \theta_i^h \theta_i (X(i) + W(i)) + \sum_{j \neq i} \theta_i^h \theta_j (X(j) + W(j)) = z_{ii} X(i) + \sum_{j \neq i} z_{ij} X(j) + Z(i), \quad (5)$$

where θ_i^h denotes the conjugate transpose of ith column of matrix $\mathbf{\Theta}$ and $z_{i,j} = \theta_i^h \theta_j$. It can be easily shown that z_{ii} is a Chi-Squared distributed random variable with M Degrees of freedom. For a sufficiently large M, $z_{ii} \sim \mathcal{N}(0, 1/M)$. Similarly for large M, z_{ij} is a sum of a large number of i.i.d. random variables. Thus we can assume z_{ij} is distributed as $\mathcal{CN}(0, 1/M)$ if $j = N - i$ else $\mathcal{CN}(0, 2/M)$, where $\mathcal{CN}(0, \sigma^2)$ is circularly symmetric complex Gaussian distribution with variance σ^2. Therefore we can rewrite $cor(i)$ as, $cor(i) = z_{ii} X(i) + I(i) + Z(i)$, where $I(i)$ corresponds to the interference term present due to other existing primary channels. Thus, for any channel i we have the following Hypothesis

$$H_0 : \Re(cor(i)) \sim \mathcal{N}(0, \sigma^2), \Im(cor(i)) \sim \mathcal{N}(0, \sigma^2) \ versus$$
$$H_1 : \Re(cor(i)) \sim \mathcal{N}(\Re(X(i)), \sigma^2), \Im(cor(i)) \sim \mathcal{N}(\Im(X(i)), \sigma^2),$$

where $\Re(c)$ denotes the real part and $\Im(c)$ denotes the imaginary part of c.

Consider the case when $|X(i)| = A$, whenever channel i is being used by a primary. For this case $\sigma^2 = SA^2/2M + 1/2 + (N+2)/M$ and $A = \sqrt{EN/S}$.

All above approximations hold asymptotically, as $N \to \infty$, $M \to \infty$, $S/N \to e_1 < \infty$ and $M/N \to e_2 < \infty$, where e_1 is the sparsity and e_2 is the compression ratio. We are interested in the scenario where $e_1 < 0.02$ and $0.1 < e_2 < 0.5$.

To detect each band we use the Neyman-Pearson (NP) test for each channel i. Under H_0 we fix probability of False alarm $P_{FA} \triangleq P[Decision\ is\ H_1 | H_0]$ as α and we choose threshold η such that $P_{FA} = \alpha = exp(-\eta/2\sigma^2)$. Thus $\eta = 2\sigma^2 \log \frac{1}{\alpha}$ and probability of Detection $P_D \triangleq P[Decision\ is\ H_1 | H_1] = Q_{\chi_2'^2(\lambda)}(|cor(i)|^2/\sigma^2 > \eta^2/\sigma^2)$, because $|cor(i)|^2/\sigma^2$ is a noncentral chi-square random variable with degrees of freedom 2 and noncentrality parameter A^2. Thus we can find the Receiver operating characteristics (ROC) of our detector.

We study the effect of parameters S, N, M, E on the performance of this detector. For simplicity we take $X(i) = A$, for all $i \epsilon S_{sup}$. where S_{sup} denotes the set of occupied channels. Then we have

$$P_{FA} = Q(\eta/\sigma^2), \quad (6)$$
$$P_D = Q(\frac{\eta - A}{\sigma}) = Q(f(S, N, M, E)). \quad (7)$$

where $A = \sqrt{EN/S}, \sigma^2 = SA^2/2M + 1/2 + (N+2)/M$, $c = \log(1/P_{FA})$ and $f(S, N, M, E) = c - A/\sigma$.

If we fix E, N, S and decrease M, $f(M, N, E, S)$ decreases and hence P_D decreases. This is verified in Fig 2. If we fix E, N, M and increase S, the argument in equation 7 decreases and hence the performance degrades with increase in S. This is also verified in Fig 3. In both the figures we plot theoretical and simulation results. There is an excellent match between them.

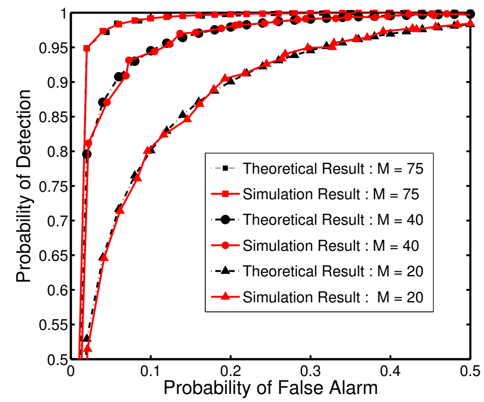

Fig. 1. Performance for given $N = 200, S = 4, SNR$=0 dB with variation in measurements M

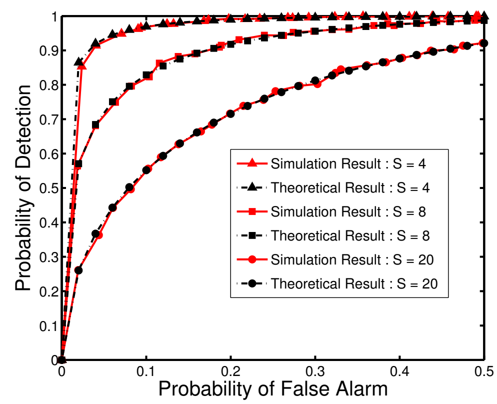

Fig. 2. Performance for given $N = 200, M = 100, H = 1, SNR$=0 dB with variation in sparsity S

Now we consider the case when we take K blocks of M samples of $\mathbf{y}$ and assume that the statistics of each channel doesn't change with block. Thus it is optimum to consider

$$\overline{cor}(i) = \frac{1}{K} \sum_{k=1}^{K} cor_k(i), \quad (8)$$

where $cor_k(i)$ is $cor(i)$ for the kth block. The effective variance for each channel now becomes $\sigma_{eff}^2 = \frac{SA^2}{2MK} + \frac{1}{2K} + \frac{N+2}{2MK}$. Now we show the trade off between (M, K) combination. If we have $2M$ measurements and K blocks then $\sigma_{eff}^2 = \frac{SA^2}{2(2M)K} + \frac{1}{2K} + \frac{N+2}{2(2M)K}$. Instead we can take M measurements and $2K$ blocks and then $\sigma_{eff}^2 = \frac{SA^2}{2M(2K)} + \frac{1}{2(2K)} + \frac{N+2}{2M(2K)}$. We see these results in lower effective variance than the former although the total number of observations is $2MK$ in both the cases.

Next we consider the case when the parameters S, N, M, E are unknown. We need to know $\sigma^2 = SA^2/2M + 1/2 + (N+2)/M$ and $A = \sqrt{EN/S}$ to find the P_{FA}, P_{MD} and hence the threshold η for appropriate performance guarantee. Thus we use Generalized Log-likelihood Ratio test (GLRT) to estimate the parameters and then detect the holes. We take K independent samples of $\mathbf{y}$ (i.e., K sets of M measurements).

The ML estimate of $\Re(X(i))$ and $\Im(X(i))$ are

$$\hat{X_R}(i)_{ML} = \frac{1}{K}\sum_{k=1}^{K}\Re(cor_k(i)), \quad (9)$$

$$\hat{X_I}(i)_{ML} = \frac{1}{K}\sum_{k=1}^{K}\Im(cor_k(i)). \quad (10)$$

The ML estimate of σ^2

$$\hat{\sigma^2}(i)_{ML} = \frac{1}{2K}\sum_{k=1}^{K}\left(\left\{\Re(cor_k(i)) - \hat{X_R}(i)_{ML}\right\}^2 + \left\{\Im(cor_k(i)) - \hat{X_I}(i)_{ML}\right\}^2\right), \quad (11)$$

Thus $\hat{X_R}(i) \sim \mathcal{N}(X_R(i), \sigma^2/K)$, $\hat{X_I}(i) \sim \mathcal{N}(X_I(i), \sigma^2/K)$. Also $E(\hat{\sigma^2}(i)) = \frac{K}{K-1}\sigma^2$ and $Var(\hat{\sigma^2}(i)) = 2\sigma^4/(K-1)$.

Since the noise for each channel is almost same for S >5, we can further improve the estimate of variance as $\hat{\sigma^2} = \frac{1}{N}\sum_{i=1}^{N}\hat{\sigma^2}(i)$. Then, $E\{\hat{\sigma^2}\} = \frac{K\sigma^2}{(K-1)}$ and $Var\{\hat{\sigma^2}\} \geq \frac{2\sigma^4}{N(K-1)}$.

For $SNR = -15dB, N = 200, S = 4$, $\sigma^2 = 1.5416$, $K = 10$, $E\{\hat{\sigma^2}\} = 1.387$, $Var\{\hat{\sigma^2}\} = 0.0337$, $\frac{K}{K-1}E\{\hat{\sigma^2}\} = 1.5411$. For $K = 5$, $E\{\hat{\sigma^2}\} = 1.2327$, $Var\{\hat{\sigma^2}\} = 0.1257$, $\frac{K}{K-1}E\{\hat{\sigma^2}\} = 1.5409$. Thus even if we do not know the parameters we can achieve satisfactory performance because of the good estimate of the variance.

The GLRT can also be used when the channel gains h are unknown. We assume slow fading and hence for a given detection interval channel gains are constant. Here we assume σ^2 to be unknown, therefore S, E can be unknown parameters. Some results are reported in Table I for COR algorithm with GLRT at -15 dB.

h	P_{MDsim}	$\hat{P}_{MDsim}$	$P_{MDtheory}$
1.00	0.0480	0.0481	0.0490
0.90	0.1130	0.1120	0.1107
0.80	0.2300	0.2300	0.2300
0.70	0.3890	0.3880	0.3880
0.60	0.5500	0.5530	0.5512
0.55	0.6480	0.6400	0.6403

TABLE I
COMPARISONS OF THEORETICAL AND SIMULATION RESULTS OF P_{MD} FOR GIVEN P_{FA} = 0.01, $N = 1000$, $M = 500$, $S = 10$, $SNR = -15$ dB, $K = 10$, AT DIFFERENT CHANNEL GAINS

In the Table I, h column denotes the channel gain, P_{MDsim} denotes the probability of miss detection found by simulation using COR-GLRT and $\hat{P}_{MDsim}$ denotes the probability of miss detection found theoretically from $\hat{\sigma^2}$ and $P_{MDtheory}$ when all the channel gains are known a priori to the CR. From these results we conclude that the performance of the COR under the unknown parameters case is almost the same as that of the COR algorithm with the known parameters.

From the above two results we conclude that even when we do not know channel gains, σ^2 etc we can achieve satisfactory performance even at low SNR for COR with GLRT.

We can show that correlation between $cor(i)$ and $cor(j)$, $i \neq j$ $\rho = \frac{1}{2M}$. Thus if $M > 100$ one can assume $cor(i)$ and $cor(j)$ are roughly independent. But if $M < 100$ then joint detection can improve performance. We do ML detection from observed cor. As there are $\binom{N}{S}$ possible hypothesis, for large N even for small S, the number of possible hypothesis is very large. Thus we search for busy channels over the channels with aS largest $cor(i)$ coefficients, where $1 < a < 2$. Here we assume S is known a priori to the detector. We have the following results for parameters $N = 200, M = 100, S = 4$ and SNR = 0 dB. For P_{FA} = 0.01, P_{MD} are 0.02, 0.06 via OMP and COR respectively, but via COR-JOINT P_{MD} = 0.03.

V. COMPARISON WITH OMP ALGORITHMS

We compare COR with OMP. For OMP the statistics of the channel coefficients are not known. Thus for OMP we obtained results via estimating $\hat{\mathbf{X}}$ and comparing each entry of $\hat{X}(i)$ with optimum threshold chosen using simulations. We take $N = 200, S = 4$ and we compare OMP with COR for different values of SNR and M. CS theory recommends M $\geq CS\log N$, where C is between 2−4.

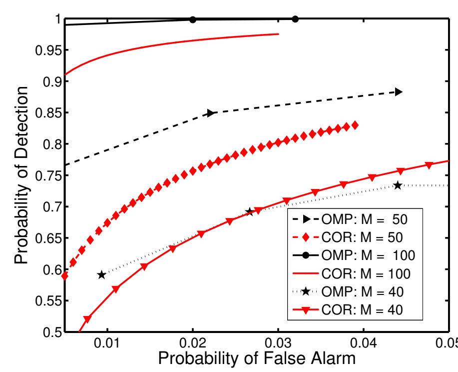

Fig. 3. OMP vs COR: ROC for given $N = 200, S = 4$ and SNR = 0 dB

We see that at 0 dB (Fig 4) OMP provides much better performance than COR when $M \geq 50$ but not for $M = 40$. However at -10 dB (Fig 5) their performance are close at $M = 100$ but COR outperforms OMP at $M = 40$. This can be explained as follows. We have $\sigma^2 = SA^2/2M + 1/2 + (N+2)/2M$. As $A = \sqrt{EN/S}$, for higher SNR A is also large. Thus σ^2 is dominated by A at each iteration of OMP. In the next iteration σ^2 will decrease as we remove the orthogonal projection of the higher energy channel. Thus OMP improves performance. But as M decreases σ^2 is dominated by N/M and therefore OMP does not perform well. At low SNR, $\sigma^2 \approx$

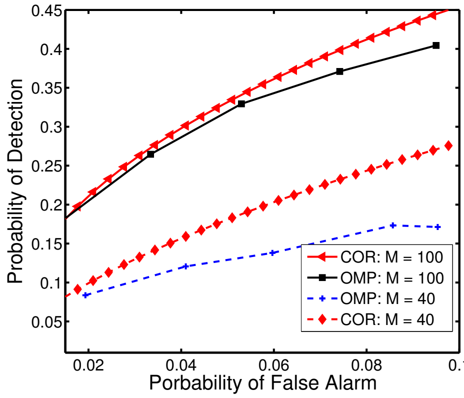

Fig. 4. OMP vs COR: ROC for given $N = 200, S = 4$ and $SNR = -10$ dB

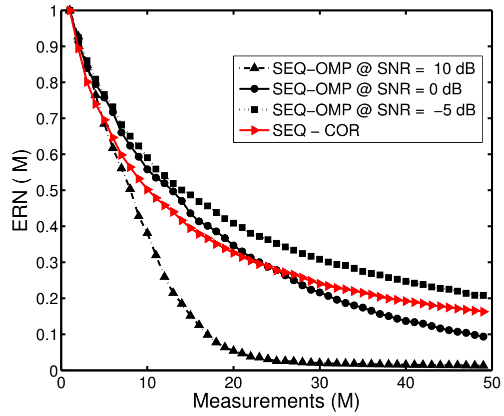

Fig. 5. Performance for given $N = 200, S = 2, H = 1$

$1/2 + (N + 2)/M$. Thus at low SNR COR performs better than OMP for small M. As we want to detect at low SNR we should prefer COR algorithm for detection of a wideband signal.

Now consider the fact that $M \geq M_{min} = CS \log N$, where C is a constant. The minimum SNR requirement for OMP during the first iteration can be found by using this. Since $\sigma^2 \leq SA^2/2M_{min}+1/2+(N+2)/(2M_{min}) = EN/(2SC \log N)+1/2+(N+2)/(2CS \log N)$, minimum SNR requirement with $C = 2$ is

$$E_m = \frac{SA^2}{E_m N/(4S \log N) + 1/2 + (N+2)/(4S \log N)} = \frac{1}{1/(4S \log N) + 1/2E_m N + 1/(4E_m S \log N)}.$$

For $N = 200, S = 4$ $E_{min} = 15dB$.

Now we consider the algorithm in [10] called Sequential OMP. We know that $M \geq S \log N$ measurements are required for signal recovery even for the noiseless case. Thus to fix M we need S. If knowledge of S is unknown then sequential OMP has been suggested to find M. The algorithm starts with some arbitrary M measurements and we keep on increasing M by 1 until $||\hat{X}_{M+1} - \hat{X}_M||_2 < e$, where e is a small number.

We use this concept on COR algorithm also. We compare its performance with Sequential OMP.

We find $E\{||\mathbf{\hat{X}_M} - \mathbf{\hat{X}_{M+1}}||_2\}$ for both the algorithms via simulation. In Fig 5 we plot

$$ERN(M) = \frac{E\{||\hat{X_M} - \hat{X_{M+1}}||_2\}}{E\{||\hat{X}_1 - \hat{X}_2||_2\}} \tag{12}$$

We conclude that the recovery of Sequential-COR is same as Sequential-OMP at low $SNR \leq -5$ dB.

VI. SEQUENTIAL COR

In the previous analysis we assumed a fixed number K of sample $\mathbf{y}$. But at low SNR to achieve satisfactory performance we need more samples. We now use sequential techniques for COR algorithm to minimize K for a given target, say $P_{FA} < 0.1$ and $1 - P_D \triangleq P_{MD} < 0.1$. It is known that Sequential Probability Ratio Test (SPRT) [10] requires minimum mean number of samples for K given P_{FA} and P_{MD}. The test samples till $\bar{K} = inf\left\{k : L_k \geq \beta \text{ } or \text{ } L_k \leq \alpha\right\}$ where L_k is the likelihood ratio of k samples and α and β are chosen to satisfy P_{FA} and P_{MD}. At time $\bar{K}$, the decision is H_1 if $L_{\bar{K}} \geq \beta$ and H_0 if $L_{\bar{K}} \leq \alpha$. When applied on COR we call it COR-SPRT while the algorithm with NP as COR-NP. Since for OMP we do not know the statistics of the projections, we are unable to use SPRT with OMP. For $H_0 : \mathcal{N}(0, \sigma^2)$ vs $H_1 : \mathcal{N}(A, \sigma^2)$, it has been shown [16] that $K_{NP}/E\{\bar{K}\} = 4$ when $P_{FA}, P_{MD} \to 0$, where K_{NP} is the number of samples for NP.

SNR	P_{FA}	P_{MD}	K_{NP}	K_{OMP}	$E\{K \mid H_0\}$	$E\{K \mid H_1\}$
-15dB	0.05	0.01	20	23	09.35	14.14
-20dB	0.10	0.05	40	44	21.50	27.50

TABLE II
COMPARISON AMONG COR-NP, COR-SPRT AND OMP

We compare the number of samples K required by COR-NP, OMP and COR-SPRT for $SNR = -15$ dB in Table II where K_{NP} and K_{OMP} are the number of samples needed for COR-NP and OMP; while $E\{\bar{K}|H_j\}$ is the mean number of samples required by COR-SPRT under hypothesis $i = 0, 1$. We see that COR-SPRT needs about half the number of observations required for K_{NP} and K_{OMP}.

Next we discuss COR-SPRT when parameters A and σ^2 may be unknown. For A unknown Lai [9] proposed the following stopping rule

$$\tilde{K} = \inf\left\{k : max\left[\sum_{i=1}^{k} log\frac{f_{\hat{\theta}k}(X(i))}{f_{\theta 0}(X(i))}, \sum_{i=1}^{k} log\frac{f_{\hat{\theta}k}(X(i))}{f_{\theta 1}(X(i))}\right] \geq g(ck)\right\}. \tag{13}$$

where $\hat{\theta_k}$ is the ML estimate of θ using k samples and $g(t) \sim log\, t^{-1}$ as $t \to 0$.

We compare the performance of COR-SPRT when the parameters are known vs when they are not known at -20 dB in Table III. We denote by $E\{\bar{K}|H_j\}$ the expected number

c	0.01	0.003	0.001
P_{FA}	0.014	0.062	0.110
P_{MD}	0.014	0.066	0.100
$E\{\bar{K}\|H_0\}$	11.60	13.92	18.58
$E\{\bar{K}\|H_1\}$	11.20	14.00	18.88
$E\{\hat{K}\|H_0\}$	11.87	17.58	23.58
$E\{\hat{K}\|H_1\}$	11.90	17.94	22.76
$E\{\tilde{K}, \hat{\sigma^2}\|H_0\}$	13.50	22.92	26.55
$E\{\tilde{K}, \hat{\sigma^2}\|H_1\}$	13.80	23.50	26.58

TABLE III
COMPARISONS OF SAMPLES REQUIRED FOR SPRT AND COMPOSITE-SPRT ALGORITHMS FOR $N = 200, M = 100, S = 4, SNR = -20$ dB

of samples required under Hypothesis j when parameter A is known and $E\{\hat{K}|H_j\}$ denotes the same when A is unknown. For both cases we assume that σ^2 is known. For the general case when both A and σ^2 are unknown then these are denoted by $E\{\tilde{K}, \hat{\sigma^2}|H_j\}$. We observe that the performance degradation due to not knowing the parameters is marginal.

VII. CONCLUSIONS

We have developed a simple COR algorithm which can be used with compressive sensing data for detection from noisy observations. We have also analyzed it theoretically. COR is a suboptimal algorithm and not designed to be used as a data recovery algorithm in CS. Instead, because of its simplicity, its statistics can be easily approximated and hence it allows us to use it with standard detection algorithms available in literature. This combination allows us to develop powerful algorithms for spectrum sensing. We have compared COR with OMP algorithm (for which statistics are intractable and hence statistical techniques are difficult to combine). We conclude that if we are operating at a low $SNR(\leq -5$ db) then the COR algorithm is a good choice for detection as compared to the existing algorithms, for fixed size sample case. We have also improved over COR by developing its sequential version. Finally we have extended these algorithms to the case when the parameters are not known. This work seems to be one of the first in this direction.

REFERENCES

[1] D. D. Ariananda, M. K. Lakshmanan, H. Nikookar, “A Survey on Spectrum Sensing Techniques for Cognitive Radio,” International Research Center for Telecommunications and Radar, Wireless VITAE 2009, Aalborg, Denmark.

[2] T. Cai, L. Wang “Orthogonal Matching Pursuit for Sparse Signal Recovery With Noise,” , IEEE Transactions on Information Theory, 2011 , pp. 4680-4688.

[3] E. J. Candes, J. Romberg, T. Tao,“Robust Uncertainty Principles: Exact Signal Reconstruction From Highly Incomplete Frequency Information,” IEEE Transactions on Information Theory, vol. 52, no. 2, Feb. 2006, pp. 489-502.

[4] W. Dai and O. Milenkovic, “Subspace pursuit for compressive sensing recovery: closing the gap between performance and complexity,” in Queen’s University, Kingston, ON, Canada, March, 2008.

[5] M. A. Davenport, M. B. Wakin, and R. G. Baraniuk, “Detection and estimation with compressive measurements,” Tech. Rep., Dept. of ECE, Rice University, 2006.

[6] D. L. Donoho, “Compressed sensing,” IEEE Trans. on Information Theory, vol. 52, Apr. 2006, pp. 1289-1306.

[7] D. L. Donoho, Y. Tsaig, and Jean-Luc Starck, “Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit,” Technical Report, Mar. 2006.

[8] Federal Communications Commission,“Notice of proposed rule making and order: Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies,” Feb. 2005.

[9] T. L. Lai, “Nearly optimal sequential tests of composite hypotheses”, The Annals of Statistics, vol.16, no.2, 1988, pp. 856-886.

[10] D. Malioutov, S Sanghavi, and S. Willsky, “Sequential Compressed Sensing,” IEEE Journal of selected topics in signal processing, vol. 4, no. 2, April 2010.

[11] ℓ_1 Magic, a collection of MATLAB routines for solving the convex optimization programs central to compressive sampling. http://www.acm.caltech.edu/l1magic.

[12] I. Mitola, J. Maguire, G. Q.,“ Cognitive radio: making software radios more personal,” IEEE Personal Commun. Mag., vol. 6, no. 4, Aug. 1999, pp. 13-18.

[13] D. Needell and R. Vershynin, “Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit,” IEEE Journal of selected topics in signal processing , Vol. 4, no. 2, April 2010, pp. 310-316.

[14] D. Needell and J. A. Tropp, “Cosamp: Iterative signal recovery from incomplete and inaccurate samples,” Communications of the ACM, ”Research Highlights” section, Dec. 2010.

[15] Y. L. Polo, Y. Wang, A. Pandharipande, and G. Leus,“Compressive wideband spectrum sensing,” IEEE International Conference on Acoustics, Speech and Signal Processing, Apr. 2009, pp. 2337-2340.

[16] H. V. Poor, “An Introduction to Signal Detection and Estimation,” in Springer-Verlag: New York, 1988; Second Edition, 1994, pp. 108-109.

[17] Z. Quan, S. Cui, H. V. Poor, and A. H. Sayed, “Collaborative wideband sensing for cognitive radios,” IEEE Signal Processing Magazine, vol. 25, no. 6, 2008, pp. 6073.

[18] E. Rebeiz, V. Jain, D. Cabric, “Cyclostationarity-Based Low Complexity Wideband Spectrum Sensing using Compressive Sampling”IEEE International Conference on Communications (ICC 2012), Ottawa, Canada, June 10-15, 2012.

[19] S. Shankar, C. Cordeiro, and K. Challapali, “Spectrum agile radios: utilization and sensing architectures,” in Proc. IEEE Int. Symposium on New Frontiers in Dynamic Spectrum Access Networks, Baltimore,Maryland, USA, Nov. 2005, pp. 160-169.

[20] Z. Tian and G.B. Giannakis, “Compressed sensing for wideband cognitive radios,” IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 4, Apr. 2007, pp. 1357-1360.

[21] Z. Tian, “Compressed Wideband Sensing in Cooperative Cognitive Radio Networks”, IEEE Global Communications Conference (GLOBECOM), Nov. 30-Dec. 4, 2008.

[22] Z. Tian, “Cyclic feature based wideband spectrum sensing using compressive sampling,” in Proc. IEEE ICC, 2011.

[23] J. Tropp and A. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit.,” IEEE Trans. Info. Theory, vol. 53, Dec 2007, pp. 4655-4666.

[24] H. Urkowitz, “Energy detection of unknown deterministic signals,” Proc. IEEE, vol. 55, Apr. 1967, pp. 523-531.

[25] A. Wald and J. Wolfowitz, “Optimum character of the sequential probability ratio test,” Annals of Statistics, vol.19, 1948, pp. 326-339.

[26] J. Yuan, “The Detection of Frequency Hopping Signal Using Compressive Sensing,” International Conference on Information Engineering and Computer Science (ICIECS), 2009, pp. 1-4.

[27] T. Yucek, H. Arslan, “A Survey of Spectrum Sensing Algorithms for Cognitive Radio Applications,” IEEE Communications Surveys and Tutorials, 2009 , pp. 116-130.