Sparse Representation for Blind Spectrum Sensing in Cognitive Radio: A Compressed Sensing Approach

Parthapratim De and Udit Satija

Institute for Infocomm Research, Singapore.

e-mail: pd4267@yahoo.com, Web-site: http://www.diat.ac.in/index.php?option=com_content\&view=article&id=176\&Itemid=400

School of Electrical Sciences, Indian Institute of Technology, Bhubaneswar, Orissa India.

e-mail: uditsatija007@gmail.com,

Abstract

Cognitive Radios dynamically sense the available unutilized spectrum for opportunistic transmission by secondary users (SU), when primary user(PU) is not transmitting. So rapid sensing is required under the consideration that the secondary users (SU) have to vacate the spectrum within very less time, whenever the primary user(PU) is active. Cyclo-stationary based spectrum sensing methods are better than energy detection methods in negative signal-to-noise (SNR) decibel (dB) regime, in which case the noise variance cannot be exactly estimated. However blind cyclo-stationary methods require a large number of symbols (and hence measurements). The paper aims to reduce the number of measurements in a blind sensing method (using a combination of linear prediction and QR decomposition), by employing compressed sensing at the receiver front-end, so as to reduce the A/D requirements needed with large number of measurements, along with multiple receive antennas. It is to be noted that compressed sensing has not been earlier investigated at extremely low like SNR of -15 dB. A comparison of few existing compressed sensing algorithms is undertaken, from the standpoint of spectrum sensing of extremely noisy signal, at highly negative SNR (dB). The proposed methods are also blind, and so do not require any information about the primary signal power, multipath channel distortions (between primary and secondary users) and band(s) occupied by the primary user and can be extended to the case when there are multiple primary users.

Index Terms

Cognitive Radio, Spectrum Sensing, Compressed Sensing, Orthogonal Matching Pursuit, Linear Prediction.

I. INTRODUCTION

This paper addresses a novel blind spectrum sensing method similar to [1], [3], [4], [5],[6], [7], but requiring fewer measurements, by employing the concept of compressed sensing (sub-Nyquist sampling) [8], [9], [10]. The compressed sensing approach is a consequence of the fact that the wireless signals (in licensed radio networks) are typically *sparse* in the frequency domain. For the task of spectrum sensing, the widely used energy detectors suffer from the *noise uncertainty* problem, below a *SNR wall*, which is the negative SNR value beyond which detection is not possible [3]. Cyclo-stationary detectors [3], [6] perform better in very low SNR situations (a reality in cognitive radio scenario), but have a huge computationally complexity and also require a large number of symbols, for signal detection in these extremely noisy situations. Moreover, most cyclo-stationary methods require information about the primary signal (modulation type, carrier frequency etc.) to operate. The novel algorithm in this paper is able to sense the primary signals at very low SNR (like -13 dB,when the noise power is much (about 20 times) higher than the signal power), by using much less symbols, and without requiring any such information about the primary signal or the multipath channel, shadowing between the primary and secondary users, and is thus truly blind in nature. Till date, integration of compressed sensing and blind spectrum sensing, has not been investigated in very low negative SNR regime.

The contributions of this paper may be given as

i) development of a model for blind spectrum sensing with multiple receive antennas (given in Section II),

ii) justification and development of compressive sensing model for *noiseless* signal (received at multiple receive antennas, at the SU), with different measurement (sensing) matrices at the multiple receive antennas, (in Section III), Also sensing model, with same measurement matrices at the multiple receive antennas, (in Section III.B),

iii) justification of compressive sensing model for *noisy* received signal (at multiple receive antennas, at the SU), with different measurement (sensing) matrices at the multiple receive antennas, (in Section III.A and III.A.1),

iv) The idea of joint recovery, where reconstruction and spectrum sensing of PU (at each receive antenna) signal is done from the measurements at all the receive antennas (in Section III.C).

v) Extension to the vector case of compressed sensing algorithms like a) Orthogonal Matching Pursuit (OMP), b) l_1 minimization (convex optimization) algorithm, c) Alternating Direction Method to obtain novel VOMP, VL1 and VAD algorithms, (in Section IV),

vi) Integration of the vector compressive algorithms with a modified blind spectrum sensing algorithm to obtain novel blind spectrum sensing algorithms (in Section V), which is able to detect PU signals at very negative SNR (dB) (when the noise power is much higher than the signal power), and the analysis of these algorithms, (in Section VI),

vii) Investigation of the suitability of the best compressed sensing algorithm, from the standpoint of blind spectrum sensing problem, with regards to a) probability of detection and probability of false alarms (for different channels, with different levels of measurement, b) comparative results with varying λ of VL1 to choose a suitable λ, c) computational times of different vector compressive sensing algorithms, d) determination of the best algorithm from comparative Receiver Operating Characteristics (ROCs) for different channels, e) Comparative results, with different measurement matrices (for different multiple receive antennas) or the same measurement matrix (for multiple receive antennas).

II. Signal Model

Let a vector $\tilde{\mathbf{x}}$, which is very sparse in frequency domain, be the transmitted signal of length $N \times 1$. Here $\tilde{a}$ refers to a (time-domain) signal a in the frequency domain. The N length received signal (in frequency domain), at ith receive antenna, is given by

$$\tilde{y}_i(k) = \tilde{h}_i(k)\tilde{x}(k) + \tilde{w}_i(k) \tag{1}$$

where $i = 1, 2.....M$, where M is the total number of receive antennas; $k = 1, 2, \cdots, N$ is the frequency index. $\tilde{h}_i(k)$ is the channel (of L taps length), in the frequency domain, from the

transmitter to the ith receive antenna, and $\tilde{w}_i$ is the discrete Fourier transform of the white gaussian noise at the ith receive antenna. The received signal at the ith antenna (over the N point frequency domain) is

$$\tilde{\mathbf{y}}_i = \left[\tilde{y}_i(0), \quad \tilde{y}_i(1), \quad \cdots, \quad \tilde{y}_i(N-1)\right]^T. \tag{2}$$

Using the $\tilde{\mathbf{y}}_i$'s, the received signal, over all the M receive antennas, is given by

$$\tilde{\mathbf{y}} = \left[\tilde{\mathbf{y}}_1^T, \quad \tilde{\mathbf{y}}_2^T, \quad \cdots, \quad \tilde{\mathbf{y}}_M^T\right]^T. \tag{3}$$

III. Compressed Sensing Model

In the noiseless case, (1) holds with the noise term $\tilde{w}_i(k) = 0$, which means that the received signal only contains the transmitted PU signal $\tilde{x}(k)$. The motivation of cognitive radio is that only a few PU's are active at a given time and location, over the entire radio spectrum, which means that the primary user wireless signals (in licensed radio networks) are typically *sparse* in the frequency domain [8]. This implies that compressed sensing can be applied to this problem and opportunistic transmission by SU is possible.

The frequency-domain received signal at the ith receive antenna of SU, $\tilde{\mathbf{y}}_i$, has few non-zero elements (i.e. is a sparse signal), because few of the PU signals, occupying different bands, are active (transmitting) at the same time/location. Since $\tilde{\mathbf{y}}_i$ is a sparse signal, for compressed sensing, the sampling process can be expressed (in frequency-domain) as

$$\tilde{\mathbf{v}}_i = \mathbf{A}_i \tilde{\mathbf{y}}_i, \;\; i = 1, 2, \ldots, M, \tag{4}$$

where a reduced dimension signal $\tilde{\mathbf{v}}_i$ can be generated by the above equation, by compressed sensing theory, due to the sparsity of $\tilde{\mathbf{y}}_i$, for each of the M receive antennas. $\tilde{\mathbf{v}}_i$ is a frequency-domain signal of size $d \times 1$ and $\mathbf{A}_i$ is a measurement matrix of size $d \times N$. Additionally, it is desired that $d << N$. The measurement matrix $\mathbf{A}_i$ may be different for each receive antenna (i.e.different $\mathbf{A}_i$'s) , or it may also be the same (i.e. $\mathbf{A}_i = \mathbf{A}, i = 1, 2, \cdots M$). Simulations are performed for both cases.

Now the frequency -domain received signal $\tilde{\mathbf{y}}_i$ (at ith receive antenna) can be expressed as

$$\tilde{\mathbf{y}}_i = \mathbf{F}_N \mathbf{y}_i, \tag{5}$$

where $\mathbf{F}_N$ is the $N \times N$ Discrete Fourier Transform (DFT) matrix. Under such sparsity conditions, one gets from (4) and (5),

$$\tilde{\mathbf{v}}_i = \mathbf{A}_i \tilde{\mathbf{y}}_i = \mathbf{A}_i(\mathbf{F}_N \mathbf{y}_i), \ \rightarrow \mathbf{v}_i = \mathbf{F}_d^{-1} \tilde{\mathbf{v}}_i = \mathbf{\Phi}_i \mathbf{y}_i \tag{6}$$

where $(\mathbf{\Phi}_i = \mathbf{F}_d^{-1} \mathbf{A}_i \mathbf{F}_N)$ is another measurement matrix of size $d \times N$. It is noted that the compressed sensing model in equation (6) is in terms of the time-domain signal $\mathbf{y}_i$.

Equation (4) is a linear inverse problem with sparseness constraint, where the sparse vector $\tilde{\mathbf{y}}_i$ can be solved, using $\mathbf{A}$ as the measurement matrix. This converts the problem into a convex optimization problem

$$\hat{\tilde{\mathbf{y}}}_i = arg \ min_{\tilde{\mathbf{y}}_i} \|\tilde{\mathbf{y}}_i\|_1, \ s.t. \ \mathbf{A}_i \tilde{\mathbf{y}}_i = \tilde{\mathbf{v}}_i. \tag{7}$$

A. Compressed Sensing for Sparse Signal in Noisy Case

In the noisy signal case, the received signal is given by (1) which includes the noise term $\tilde{w}_i(k)$. Till now, compressed sensing and reconstruction for a noiseless PU signal, sparse in the frequency domain and relevant to the cognitive radio regime, has been explored. However, when the PU signal is received at the ith receive antenna at the SU, the PU signal has already been corrupted by additive white Gaussian noise $\mathbf{w}_i$ (at very low SNR's like -13 dB) (during transmission), which is a white noise process with zero mean and variance $\sigma_{0,i}^2 \mathbf{I}$, which is represented by $\tilde{\mathbf{w}}_i$ in frequency domain. After addition of measurement noise $\mathbf{z}_i$ (zero mean white noise with variance $\sigma_i^2 \mathbf{I}$, which is independent of $\mathbf{w}_i$) at the SU, (4) becomes (for M receive antennas)

$$\tilde{\mathbf{v}}_i = \mathbf{A}_i \tilde{\mathbf{y}}_i = \mathbf{A}_i(\tilde{\mathbf{h}}_i \tilde{\mathbf{x}} + \tilde{\mathbf{w}}_i) + \tilde{\mathbf{z}}_i, \ i = 1, 2, \ldots, M. \tag{8}$$

(8) can be re-written, for the ith receive antenna as

$$\tilde{\mathbf{v}}_i = \mathbf{A}_i(\tilde{\mathbf{h}}_i \tilde{\mathbf{x}}) + (\mathbf{A}_i \tilde{\mathbf{w}}_i + \tilde{\mathbf{z}}_i), \ i = 1, 2, \ldots, M, \tag{9}$$

where the second term $(\mathbf{A}_i\tilde{\mathbf{w}}_i+\tilde{\mathbf{z}}_i)$ is the frequency domain representation of the (non-white) noise. Then the time-domain signal $\mathbf{v}_i$ is given by (obtained by pre-multiplying (9) by $\mathbf{F}_d^{-1}$),

$$\mathbf{v}_i = \mathbf{\Phi}_i^1(\mathbf{h}_i * \mathbf{x}) + (\mathbf{\Phi}_i^1\mathbf{w}_i + \mathbf{z}_i),\ i = 1, 2, \ldots, M, \tag{10}$$

where $\mathbf{\Phi}_i^1 = \mathbf{F}_d^{-1}\mathbf{A}_i\mathbf{F}_d$. $*$ is the convolution operator. Then to whiten the noise, the whitening filter is defined, for ith receive antenna, $\mathbf{Q}_i = (\sigma_{0,i}^2(\mathbf{\Phi}_i^1\mathbf{\Phi}_i^{1H}) + \sigma_i^2\mathbf{I})\Gamma_i^{-1}$, where Γ_i is a $d \times d$ diagonal matrix whose diagonal entries are equal to $(\gamma_i == \frac{N}{d}\sigma_{0,i}^2 + \sigma_i^2)$ [16] (where the generic case of a signal, sparse in time, is considered, whereas in this paper, the specific case of spectrum sensing is considered, where the PU signal is sparse in the frequency domain). Then (10) becomes

$$\bar{\mathbf{v}}_i = \mathbf{B}_i(\mathbf{h}_i * \mathbf{x}) + \bar{\mathbf{u}}_i,\ i = 1, 2, \ldots, M, \tag{11}$$

where $\bar{\mathbf{v}}_i = \mathbf{Q}_i^{-1/2}\mathbf{v}_i$, $\mathbf{B}_i = \mathbf{Q}_i^{-1/2}\mathbf{\Phi}_i^1$ and $\bar{\mathbf{u}}_i = \mathbf{Q}_i^{-1/2}(\mathbf{\Phi}_i^1\mathbf{w}_i + \mathbf{z}_i)$. Then

$$\mathbf{F}_d\bar{\mathbf{v}}_i = \mathbf{F}_d\mathbf{B}_i\mathbf{F}_d^{-1}\mathbf{F}_d(\mathbf{h}_i * \mathbf{x}) + \mathbf{F}_d\bar{\mathbf{u}}_i,\ i = 1, 2, \ldots, M, \tag{12}$$

Thus,

$$\tilde{\bar{\mathbf{v}}}_i = \mathbf{\Phi}_i^2(\tilde{\mathbf{h}}_i\tilde{\mathbf{x}}) + \tilde{\bar{\mathbf{u}}}_i,\ i = 1, 2, \ldots, M, , \tag{13}$$

where $\mathbf{\Phi}_i^2 = (\mathbf{F}_d\mathbf{B}_i\mathbf{F}_d^{-1})$. The optimization problem then becomes

$$\hat{\tilde{\mathbf{x}}}_i = arg\ min_{\tilde{\mathbf{x}}_i}\|\tilde{\mathbf{x}}_i\|_1,\ s.t.\ \tilde{\bar{\mathbf{v}}}_i = \mathbf{\Phi}_i^2(\tilde{\mathbf{h}}_i\tilde{\mathbf{x}}) + \tilde{\bar{\mathbf{u}}}_i,\ i = 1, 2, \ldots, M. \tag{14}$$

Now since the AWGN noise is *non-sparse* in the frequency domain (in fact, it covers the entire frequency spectrum), the received signal is no longer sparse. Recovery of noise corrupted sparse signals have been investigated in [12], [14], [13], [15].

1. In particular, in [12], for the noisy received signal, the regularized version of OMP (ROMP) produces an approximation to the reconstructed (signal and noise, after OMP) $\hat{\mathbf{y}}$, satisfies

$$||\hat{\tilde{\mathbf{x}}}_i - \tilde{\mathbf{x}}_i||_2 \leq C||\tilde{\bar{\mathbf{u}}}_i||_2, \tag{15}$$

where C is a scalar constant. Thus ROMP does not require any knowledge about the error $\tilde{\tilde{\mathbf{u}}}_i$, the magnitude of $\tilde{\mathbf{u}}_i$ can be *arbitrary*, unlike convex methods, (pg 321, in [12]). Equation (15) indicates that, for a sparse PU signal corrupted by arbitrary large noise (even at SNR of -13 dB), the distortion in the reconstruction of the frequency domain (noise)-corrupted PU signal, (from reduced number of measurements), depends only (is bounded by) on the noise, and not on the PU signal. This means that, if two different sparse signals $s_1(n)$ and $s_2(n)$ are corrupted by the same amount of (even) large noise variance, the maximum degradation in the reconstruction of the noise-corrupted signal will be bounded by the same quantity, independent of whether the sparse signal is $s_1(n)$ or $s_2(n)$.

2. This leads to the conjecture, that reconstruction of noise corrupted PU signal (signal + noise), from limited number of samples (i.e. compressed sensing) is approximately similar to the average of the addition of the reconstructed noiseless PU signal (from the same number of PU signal samples), and the separately reconstructed noise (from same number of noise samples)

Simulation results in Figures 1 and 2 seem to support the conjecture.

3. **Idea Of Correlation In Novel Spectrum Sensing Algorithm**

To explain the ability of the novel spectrum sensing algorithm (discussed in Section V), to sense primary signals, even at SNR of -15 dB (when noise power is much higher than the signal power), as also in [3], [4], it is seen that the signal components (in the received signal and noise), on the multiple receive antennas, are correlated, as the same transmitted signal $x(n)$ travels through different multipath channels to be received on the multiple receive antennas. However, the (very large) noise components (of the received signals, on the multiple receive antennas) can be assumed to be uncorrelated [3], [4].

To explain the superior performance of the proposed (cyclo-stationary based) spectrum sensing algorithm (using full measurements, and not employing compressed sensing), a single-tap multipath channel (in time domain) i.e. number of channel taps $L = 1$ and number of receive antennas $M = 2$, is considered. Then the time-domain received signal (at the receive antennas), even at very low SNR

(like -13 dB), for the nth symbol, (by taking the IDFT of (1)),

$$\begin{bmatrix} y_0(n) \\ y_1(n) \end{bmatrix} = \begin{bmatrix} h_0(0) \\ h_1(0) \end{bmatrix} x(n) + \begin{bmatrix} w_0(n) \\ w_1(n) \end{bmatrix}$$
$$= \begin{bmatrix} d_0(n) \\ d_1(n) \end{bmatrix} + \begin{bmatrix} w_0(n) \\ w_1(n) \end{bmatrix} \quad (16)$$

When the noise power is much higher than the signal power (like -13 dB), the second term in the above equation is much larger than the first term.

However even then, the correlation between the noise components $w_0(n)$ and $w_1(n)$ can be considered to be very low [3], [4]. But the correlation between the signal components $d_0(n)$ and $d_1(n)$ is high, because of the presence of the common term $x(n)$ in both $d_0(n)$ and $d_1(n)$. This is $E[d_0(n)d_1^*(n)] = h_0(0)h_1^*(0)\sigma_x^2$. ($E[]$ is the expected value) Similar logic is provided by the author in [3] and in [4].

In Section V, more detailed analysis will be provided, with respect to the specific novel spectrum sensing method of this paper. It will be seen then that the value of the correlation (when signal is present) is not required, except that it is above a threshold, which does not depend on channel coefficients, nor on the SNR.

4. In the next Sub-subsection, it is seen that this reconstructed noise, at the multiple receive antennas, are still uncorrelated with each other.

1) Compressed Sensing for Spectrum Sensing of Noisy PU Signals: In the spectrum sensing problem investigated in this paper, reconstruction of noisy signal (from few measurements) is not the main goal. As mentioned above, the spectrum sensing method algorithm is based on the fact that the signal components (on the different receive antennas) are correlated, while the noise components, on the different receive antennas, are uncorrelated (see equation (16) above). It is now shown that even after the reduced number of measurements of the noise signals are reconstructed by the OMP and other recovery algorithms, the reconstructed noise signals on the different receive antennas) are still uncorrelated. This implies that in the novel spectrum sensing algorithm, which implicitly calculates the cross-correlation between noise corrupted signals, at the multiple receive antennas, the large noise terms (even at low SNR of -13 dB) will be removed, enabling signal detection.

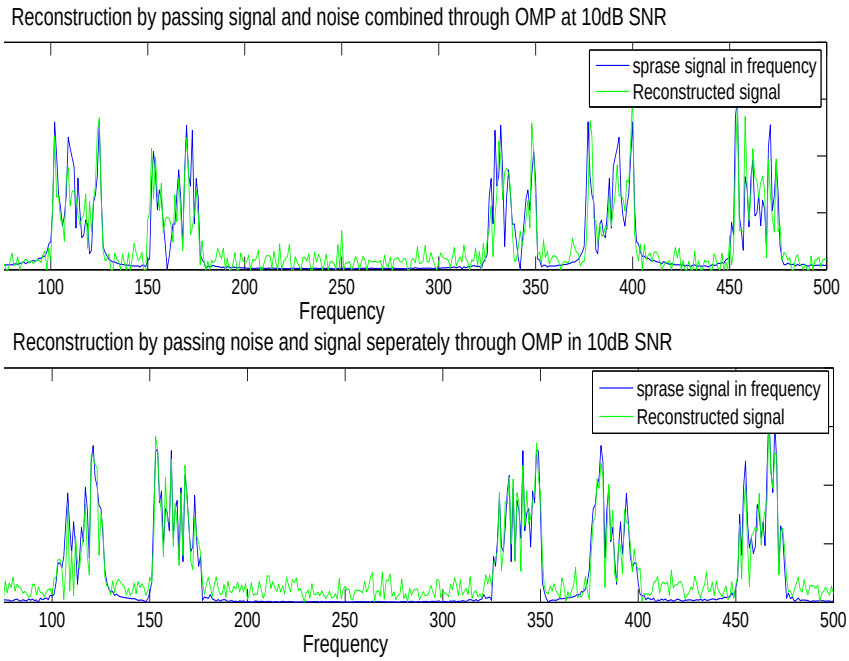

Fig. 1. Reconstruction through OMP in 10dB SNR by passing signal and noise combined and separately

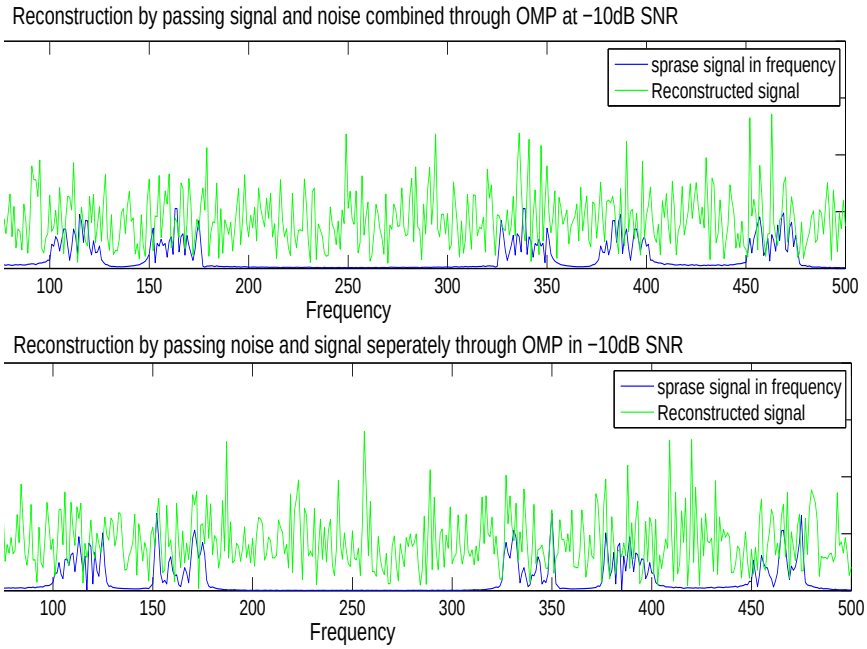

Fig. 2. Reconstruction through OMP in -10dB SNR by passing signal and noise combined and separately

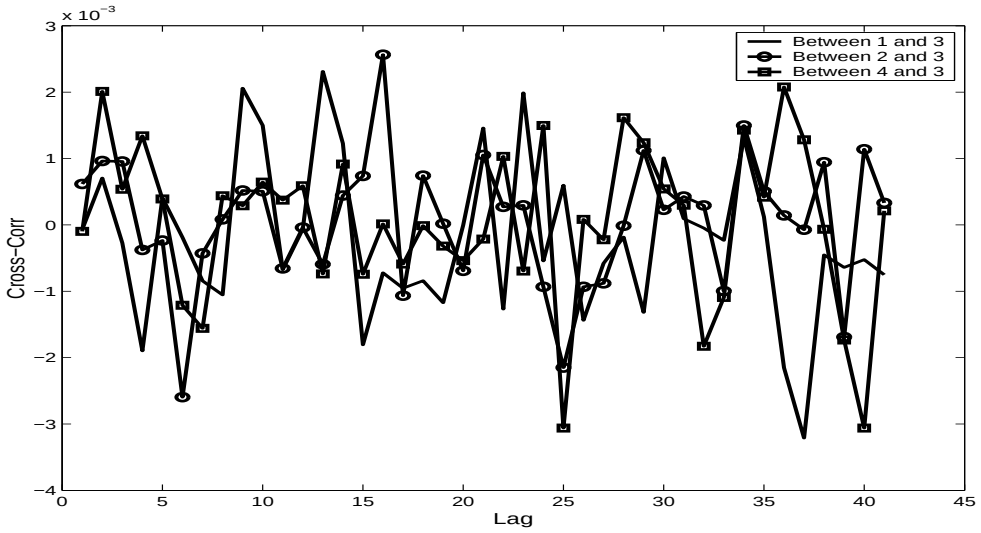

Fig. 3. Correlation for reconstructed noise at different receive antennas.

In the absence of PU signal, only noise $\mathbf{u}_i$ and $\mathbf{u}_j$ are received at the ith and jth receive antennas. They are then processed through measurement matrices $\mathbf{\Phi}_i$ and $\mathbf{\Phi}_j$ to produce reduced number of measurements d-length vectors $\mathbf{v}_i = \mathbf{\Phi}_i \mathbf{u}_i$ and $\mathbf{v}_j = \mathbf{\Phi}_j \mathbf{u}_j$. From these reduced vectors $\mathbf{v}_i$ and $\mathbf{v}_j$, the ROMP reconstructs the estimate to the original signals (in this case, noise), which are length N vectors $\hat{\mathbf{u}}_i = \mathbf{\Phi}_i^r \mathbf{\Phi}_i \mathbf{u}_i$ and $\hat{\mathbf{u}}_j = \mathbf{\Phi}_j^r \mathbf{\Phi}_j \mathbf{u}_j$, where $\mathbf{\Phi}_i^r$ is the reconstruction operator, performing the ROMP algorithm and may be thought of as computing (in a batch processing mode), $\mathbf{A}_i^H \tilde{\mathbf{v}}_i$ [15], which follows from (9).

Then

$$E[\hat{\mathbf{u}}_i \hat{\mathbf{u}}_j^H] = E[(\mathbf{\Phi}_i^r \mathbf{\Phi}_i \mathbf{u}_i)(\mathbf{\Phi}_j^r \mathbf{\Phi}_j \mathbf{u}_j)^H] = E[\mathbf{\Phi}_i^r \mathbf{\Phi}_i] E[(\mathbf{u}_i \mathbf{u}_j^H)] E[(\mathbf{\Phi}_j^r \mathbf{\Phi}_j)^H] = \mathbf{0} \tag{17}$$

. The above follows that $\mathbf{\Phi}_i$ (and $\mathbf{\Phi}_i^r$) are each uncorrelated with $\mathbf{\Phi}_j$ (and $\mathbf{\Phi}_j^r$) and also uncorrelated with $\mathbf{u}_i$ and $\mathbf{u}_j$. Also, $E[\mathbf{u}_i \mathbf{u}_j^H] = \mathbf{0},\ i \neq j$. To verify this statement, 4000 samples of white noise $\mathbf{u}_i,\ i = 1, 2, \ldots, M$ is pre-multiplied by a measurement matrix $\mathbf{\Phi}_i$ of size 1400×4000, to obtain $\mathbf{v}_i$, which has 1400 samples (i.e. only 35% of measurements). Then each of the $\hat{\mathbf{u}}_i$'s are reconstructed using ROMP algorithm. Then the cross-correlation between the reconstructed signals at antennas i and j, i.e. $\hat{\mathbf{u}}_i$ and $\hat{\mathbf{u}}_j$, for different pairs of i and j, are computed (and averaged over 100 computer experiments) and then plotted in Figure 3. Figure 3 confirms the result in equation (17).

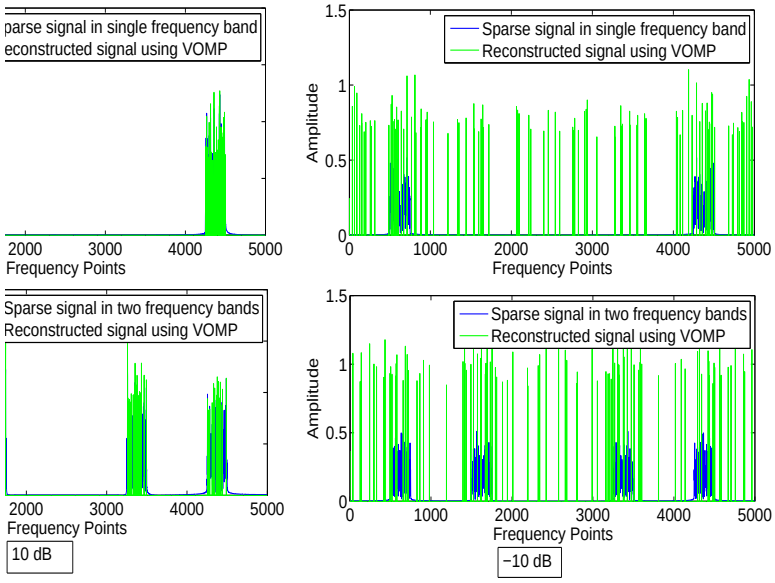

Fig. 4. Reconstruction using VOMP in single tap fixed channel using one antenna for one and two primary users

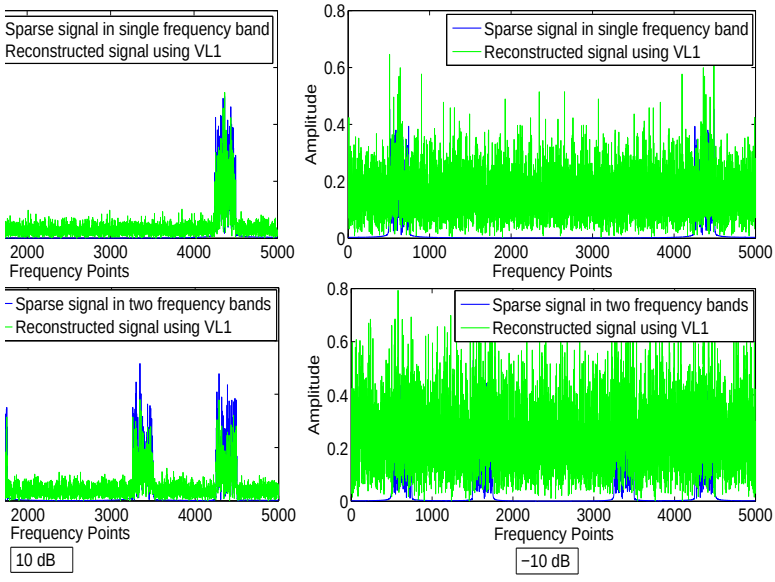

Fig. 5. Reconstruction using VL1 in single tap fixed channel using one antenna for one and two primary users

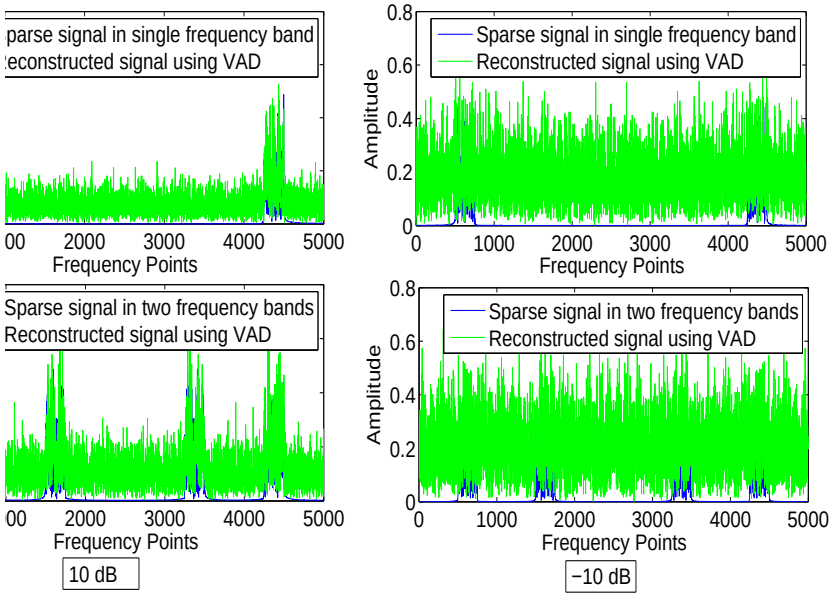

Fig. 6. Reconstruction using VAD in single tap fixed channel using one antenna for one and two primary users

2) Definition of SNR in Compressed Sensing: [16] considers the case where the sparse signal is subjected to random noise, prior to measurements (as opposed to the case when signal is noiseless but the measuring device adds noise to signal). In the spectrum sensing problem discussed in this paper, first case is relevant since the channel adds noise to the transmitted signal, before it is measured at the receiver, where measurement noise is added. From the discussion in Section III.A, the equivalent noise-variance will be given by

$$\mathbf{Q} = (\frac{N}{d}\sigma_{0,i}^2 + \sigma_i^2)\mathbf{I}. \tag{18}$$

IV. Measurement Matrices

Till now, different measurement matrices $\mathbf{\Phi}_i$ for the different receive antennas, was considered. Now, some other cases are considered.

A. Same Measurement Matrix

In the case when the measurement matrix is the same for all M receive antennas (i.e. $\mathbf{\Phi}_i = \mathbf{\Phi}$, $i = 1, 2, \cdots, M$), then the equation (6) becomes

$$\begin{bmatrix}\mathbf{v}_1^T & \mathbf{v}_2^T & \dots & \mathbf{v}_M^T\end{bmatrix}^T = \begin{bmatrix}(\mathbf{\Phi}\mathbf{y}_1)^T & (\mathbf{\Phi}\mathbf{y}_2)^T & \dots & (\mathbf{\Phi}\mathbf{y}_M)^T\end{bmatrix}^T \tag{19}$$

This can be rewritten as

$$\mathbf{v} = diag(\mathbf{\Phi}, \mathbf{\Phi}, \ldots, \mathbf{\Phi}) \begin{bmatrix} \mathbf{y}_1^T & \mathbf{y}_2^T & \ldots & \mathbf{y}_M \end{bmatrix}^T = \tilde{\mathbf{\Phi}}\mathbf{y}, \tag{20}$$

where the $(Md \times 1)$ vector $\mathbf{v}$ is the vector on the left hand side of (20), $(Md \times MN)$ matrix $\tilde{\mathbf{\Phi}}$ is the matrix on the right hand side of (20) and $\mathbf{y}$ is the $(NM \times 1)$ vector on the right-hand side of (20).

Measurement matrix $\mathbf{\Phi}_i$, being different for each receive antenna (i.e. different $\mathbf{\Phi}_i$'s), or being the same (i.e. $\mathbf{\Phi}_i = \mathbf{\Phi}, i = 1, 2, \cdots, M$) are both considered in the novel compressed sensing based spectrum sensing method. Since the PU signal is detected after reconstruction, hence as far as spectrum sensing is concerned, having the same measurement matrix $\mathbf{\Phi}$ or different measurement matrices, for different receive antennas, may not make much difference.

B. Joint Signal Recovery

In equation (20), since $\tilde{\mathbf{\Phi}}$ is a block-diagonal matrix (with M blocks, each of size $d \times N$), the measurement and reconstruction of the sparse signals (on each receive antenna) is done independently. Next the case of joint signal recovery is considered. Starting with a measurement matrix (of size $(Md \times MN)$) $\tilde{\tilde{\mathbf{\Phi}}}$, which is a non-diagonal matrix (as opposed to $\tilde{\mathbf{\Phi}}$ in (20)), which is generated randomly (as in [8]), one has (similar to the time-domain compressed sensing model, in the second equation in (6)),

$$\mathbf{v} = \tilde{\tilde{\mathbf{\Phi}}}\mathbf{y}, \tag{21}$$

where $\mathbf{y}$ is the received data on the the M receive antennas, concatenated one below the other, while the $((j-1)d+1)$ to jd elements of $\mathbf{v}$ correspond to the compressed (received) data for the jth receive antenna, i.e. $\mathbf{v}_j$.

Expressing the time-domain received signal $\mathbf{y}$ as $\mathbf{y} = \mathbf{F}_{NM}^{-1}\tilde{\mathbf{y}}$, where $\mathbf{F}_{NM}$ is the $(NM \times NM)$ Discrete Fourier Transform (DFT) matrix. $\tilde{\mathbf{y}}$, the DFT of $\mathbf{y}$, (combined over all the M receive antennas) has few non-zero elements (i.e. is a sparse signal), due to the same reasons as above. This

gives from (20),

$$\mathbf{v} = \tilde{\tilde{\mathbf{\Phi}}}\mathbf{y} = \tilde{\tilde{\mathbf{\Phi}}}\mathbf{F}_{NM}^{-1}\tilde{\mathbf{y}},\ \rightarrow \tilde{\mathbf{v}} = (\mathbf{F}_d\tilde{\tilde{\mathbf{\Phi}}}\mathbf{F}_{NM}^{-1}\tilde{\mathbf{y}}) = \tilde{\mathbf{A}}\tilde{\mathbf{y}} \tag{22}$$

where $\tilde{\mathbf{A}} = (\mathbf{F}_d\tilde{\tilde{\mathbf{\Phi}}}\mathbf{F}_{NM}^{-1})$ is the overall measurement matrix, obtained by jointly solving the compressed sensing problem over all the M receive antennas, rather than solving the compressed sensing problem (4) individually for each received antenna.

Similar to (7), we have

$$\hat{\tilde{\mathbf{y}}} = arg\ min_{\tilde{\mathbf{y}}}\|\tilde{\mathbf{y}}\|_1,\ s.t.\ \tilde{\mathbf{A}}\tilde{\mathbf{y}} = \tilde{\mathbf{v}}, \tag{23}$$

except that here $\tilde{\mathbf{v}}$ is of size $(Md\times 1)$, (while $\mathbf{v}_i$ in (7) is of size $(d\times 1)$), $\tilde{\mathbf{A}}$ is of size $Md\times NM$ (while $\mathbf{A}$ in (7) is of size $d\times N$) and $\mathbf{y}$ is of size $(MN\times 1)$, (while $\mathbf{y}_i$ in (7) is of size $(N\times 1)$). Thus by jointly performing the compressed sensing and then re-construction, it becomes a much larger scale problem, with much higher computational complexity. When, as in Section VIII, $M = 4$, $N = 5000$, with 50% measurements $d = 2500$, sparsity level $m = 0.6\%$ of total values, which are non-zero (on each receive antenna), i.e. $m = 30$, the overall matrix $\tilde{\mathbf{A}}$ in (23) is of size $(4\times 2500)\times(4\times 5000) = (10000\times 2000)$, $\mathbf{v}$ is a (10000×1) sized vector, $\mathbf{y}$ is a (20000×1) sized vector, the total number of non-zero entries in $\mathbf{y} = Mm = 120$.

However, joint signal recovery, from the combined compressed sampling model in (21) and (23), has some advantages over individual compressed sensing and reconstruction at each receive antenna in (6). By (21), each $\mathbf{v}_j$ includes contributions from different $\mathbf{y}_i$'s, since $\tilde{\tilde{\mathbf{\Phi}}}$ is a full (non-block diagonal) matrix. Then if the received signal on one of the antennas (say $\mathbf{y}_l$) is extremely weak as compared to that on other receive antennas, (due to fading channel paths), even then $\mathbf{v}_l$ has contributions from other $\mathbf{y}_i$'s, each of which contains the transmitted PU signal, making its detection easier. Simulation results in Section VIII confirm this observation (Fig 13).

V. Vector Compressed Sensing Algorithms

In this paper, equation (7) is solved using three different approaches: 1. Vector Regularized Orthogonal Matching Pursuit (VOMP) algorithm, for M receive antennas, as shown below. 2. Fast

Vector l_1 minimization (VL1) (convex optimization) algorithm. 3. Vector version of alternating direction method (VAD).

A. Vector Regularized Orthogonal Matching Pursuit (VOMP)

The scalar Orthogonal Matching Pursuit in [12], [13] is extended to the vector case for M receive antennas. Though the extension is simple, the algorithm is enumerated in terms of the novel compressed sensing based blind spectrum sensing model in Sections II, III and IV. Starting from the optimization problem in (7), d is the number of measurements and let m be the sparsity level, i.e. m is the number of zero elements in $\tilde{\mathbf{y}}_i$. Since the sparsity is in the frequency domain, the frequency domain signal $\tilde{\mathbf{y}}_i$ is reconstructed from $\tilde{\mathbf{v}}_i$.

In the vector regularized OMP (ROMP) algorithm, let it be initialized that the residual $\mathbf{r}_{0,i} = \mathbf{v}_i$, and the index set (of chosen atoms) $\mathbf{\Lambda}_{0,i}$ to zero and the matrix (of chosen atoms) $\mathbf{A}_{0,i}$ as an empty matrix at iteration $t = 0$. The following steps are repeated m times. First, a set J_t of the m largest values of $\mathbf{b} = | < \mathbf{A}_{t,i}, \mathbf{r}_{t-1,i} > |$, where each element of the vector $\mathbf{b}$ is the correlation of a column of $\mathbf{A}_{t,i}$ with $\mathbf{r}_{t-1,i}$, is chosen. Then in the regularization step, among all subsets $\tilde{J}_t \subset J_t$ with comparable coordinates $|b(i)| \leq 2|b(j)|$, for all $i, j \epsilon \tilde{J}_t$, choose $\tilde{J}_t$ with maximal energy $||\mathbf{b}|_{\tilde{J}_t}||_2$. Then

$$\mathbf{\Lambda}_{t,i} = \mathbf{\Lambda}_{t-1,i} \bigcup \{\tilde{J}_t\}, \ \mathbf{A}_{t,i} = \left[\mathbf{A}_{t-1,i} \quad | \quad \mathbf{a}_{\tilde{J}_t}\right], \tag{24}$$

where $\mathbf{a}_{\tilde{J}_t}$ contains the columns of $\mathbf{A}_{t,i}$, indexed by $\tilde{J}_t$. Next, the linear system of equations $\mathbf{v}_i = \mathbf{A}_{t,i}\mathbf{z}_{t,i}$ has to be solved at iteration t. At iteration t, the matrix $\mathbf{A}_{t,i}$ is of size up to $d \times 2m$. Since d is chosen such that $d >> m$ (in [9], d can be $\approx m^2$), $\mathbf{A}_{t,i}$ can be assumed to be of full column rank and thus the above equation can be solved (in the least squares sense) by $\mathbf{z}_{t,i} = [\mathbf{A}_{t,i}^H \mathbf{A}_{t,i}]^{-1} \mathbf{A}_{t,i}^H \mathbf{v}_{t,i}$. The new approximation to the data and the new residual are calculated as $\mathbf{a}_{t,i} = \mathbf{A}_{t,i}\mathbf{z}_{t,i}, \ \mathbf{r}_{t,i} = \mathbf{v}_i - \mathbf{a}_{t,i}$. Then the ROMP, in at most m iterations, outputs the set $\mathbf{\Lambda}_i$, such that $|\mathbf{\Lambda}_i| \leq 2m$ [13].

For the joint recovery problem in Section IV.B, the compressed sensing model is given by (23), as opposed to individual measurement matrices (for each receive antennas) given by (7), except that the dimensions of the vectors and matrices in (23) are much larger than those in (7). In the version

of VOMP, for joint signal recovery, the residual is initialized, at time $t = 0$, as $\mathbf{r}_0 = \mathbf{v}$, where $\mathbf{r}_0$ is a ($dM \times 1$) vector; the index set (of chosen atoms), $\mathbf{\Lambda}_0$ is initialized to zero and the matrix (of chosen atoms) $\tilde{\mathbf{A}}_0$ as an empty matrix. At time t, again a set J_t of the (mM) largest values of $\mathbf{b} = | < \tilde{\mathbf{A}}_{t,i}, \mathbf{r}_{t-1,i} > |$, which is the correlation of MN columns of $\tilde{\mathbf{A}}_{t,i}$ with $\mathbf{r}_{t-1,i}$, is chosen. Then the regularization step is same as above. Also $\mathbf{\Lambda}_t = \mathbf{\Lambda}_{t-1} \bigcup \{\tilde{J}_t\}$, $\tilde{\mathbf{A}}_{t,i} = \left[\tilde{\mathbf{A}}_{t-1,i} \quad | \quad \tilde{\mathbf{a}}_{\tilde{J}_t}\right]$. Since the matrix $\tilde{\mathbf{A}}_t$ is of size up to $(Md \times 2Mm)$, and $d \geq m$, as in [9], $\tilde{\mathbf{A}}_t$ is a tall matrix and assumed to be of full column rank and thus the system of equations $\mathbf{v}_t = \tilde{\mathbf{A}}_t \mathbf{y}_t$ has a least squares solution $\mathbf{y}_t = [\tilde{\mathbf{A}}_t^H \tilde{\mathbf{A}}_t]^{-1} \tilde{\mathbf{A}}_t^H \mathbf{v}_t$. The new approximation to the data and the new residual are calculated as $\mathbf{a}_t = \tilde{\mathbf{A}}_t \mathbf{y}_t$ and $\mathbf{r}_t = \mathbf{v} - \mathbf{a}_t$.

B. Fast Vector l_1 minimization (VL1) (convex optimization) algorithm

In this algorithm, compressed sensing problem is solved as convex optimization problem

$$\hat{\tilde{\mathbf{y}}}_i = arg\ min_{\tilde{\mathbf{y}}_i} \lambda \|\tilde{\mathbf{y}}_i\|_1 + \|(\mathbf{A}\tilde{\mathbf{y}}_i - \mathbf{v}_i)\|_2^2 \tag{25}$$

where λ is the Lagrangian multiplier. A gradient projection algorithm is used to solve (25), which is a combined l_1 and least squares method. It is based on a *truncated Newton interior-point algorithm* [18], [17]. It transforms the cost function in (25) to a quadratic program with inequality constraints

$$\begin{aligned} & min \|(\mathbf{A}\tilde{\mathbf{y}}_i - \mathbf{v}_i)\|_2^2 + \lambda \sum_{j=1}^{n} u_i(j), \\ & s.t. - u_i(j) < y_i(j) \leq u_i(j), \end{aligned} \tag{26}$$

where $u_i(j)$, $y_i(j)$ are the jth components of $\mathbf{u}_i, \mathbf{y}_i$. A logarithmic barrier for the constraints $-u_i(j) < y_i(j) \leq u_i(j)$ can be constructed as [18]

$$\Psi(\mathbf{y}_i, \mathbf{u}_i) = -\sum_j log(u_i(j) + y_i(j)) - \sum_j log(u_i(j) - y_i(j)). \tag{27}$$

With the above formulation of the combined l_1 and least squares method, for the vector case of M receive antennas, the gradient projection algorithm based on a truncated Newton interior-point algorithm [18], [17], developed for the scalar case, is extended to the vector case, in a similar manner to the VOMP algorithm in Section IV.A, to obtain the novel fast vector VL1 algorithm. Measurement

matrix $\mathbf{\Phi}_i$, different for each receive antenna (i.e.different $\mathbf{\Phi}_i$'s) , or same (i.e. $\mathbf{\Phi}_i = \mathbf{\Phi}, i = 1, 2, \cdots M$ are both considered for this algorithm. The VL1 algorithm can be extended, to the case of joint signal recovery, similar to VOMP case in Section IV.C.

C. Vector Alternating Direction Method (VAD).

Alternating Direction Method is an algorithm that alternates between optimizing the sparse signal and the residual error $\mathbf{e}_i$, for ith receive antenna.

$$min_{\tilde{\mathbf{y}}_i, \mathbf{e}_i} \|\tilde{\mathbf{y}}_i\|_1 + \frac{1}{2\mu}\|\mathbf{e}_i\|_2^2, \tag{28}$$

subject to $\mathbf{v}_i = (\mathbf{A}\tilde{\mathbf{y}}_i + \mathbf{e}_i)$. Alternating Direction Method can be applied to the dual problems of l_1-min, which iterates in both the primal and dual spaces to converge to the optimal solution [19]. The VAD can be extended to the case of joint signal recovery, similar to Section IV.C.

VI. Novel Blind Spectrum Sensing Algorithms:

The novel algorithm (in this paper) is a combination of compressed sensing (VOMP and VL1), with reduced number of measurements at multiple receive antennas, and a modified (blind) spectrum sensing algorithm to reliably detect PU signals up to extremely low SNR of -15 dB, using only 50% of measurements in [3]. Now, starting from (7), the (frequency domain) signal $\tilde{\mathbf{y}}_i$ is reconstructed, (for M receive antennas), using both VOMP and VL1. From the estimate of $\tilde{\mathbf{y}}_i$ above, using an inverse DFT, an estimate of (time-domain signal) $\mathbf{y}_i$, denoted by $\mathbf{z}_i$, is obtained. Then the blind method of [3] (which does *not* suffer from noise uncertainty problem, at very negative SNR) is modified and applied to the vital spectrum sensing problem. The advantage of the novel method proposed in this paper is to overcome the large number of symbols (measurements) required in [3], by using few measurements and reconstructing the large number of symbols and then using the large number of symbols in the blind detection algorithm of [3], which is able to use correlations between signal components (and *uncorrelatedness of noise* components) on the receive antennas. Let $\mathbf{z}_i$ be the estimate of $\mathbf{y}_i$. Then defining a vector by stacking the signal on the M receive antennas,

$$\mathbf{z} = \begin{bmatrix} \mathbf{z}_1; & \mathbf{z}_2; & \cdots; & \mathbf{z}_M \end{bmatrix} \tag{29}$$

and at symbol interval n, we define

$$\mathbf{z}(n) = \Big[z_1(n); \quad z_2(n); \quad \cdots ; \quad z_M(n) \Big]. \tag{30}$$

Stacking P successive symbols on top of each other, we get

$$\mathbf{z}_P(n) = \Big[\mathbf{z}^T(n), \mathbf{z}^T(n-1), \cdots, \mathbf{z}^T(n-P+1) \Big]^T. \tag{31}$$

Similar to [3],

$$\mathbf{z}_P(n) = \mathbf{H}_P \mathbf{x}(n) + \mathbf{w}_P(n) \tag{32}$$

where $\mathbf{H}_P$ and $\mathbf{w}_P(n)$ are the channel matrix and the stacked noise vector as in [3]. Assuming that $\mathbf{H}_P$ is of full column rank and defining the data matrix (over S symbols),

$$\mathbf{Z}(n) = \Big[\mathbf{z}_P(n), \mathbf{z}_P(n-1), \cdots, \mathbf{z}_P(n-S+1) \Big], \tag{33}$$

and defining the prediction error vector $\mathbf{e}(n) = \Big[-\mathbf{P}_{P-1} \quad \mathbf{I}_M \Big] \mathbf{z}_P(n)$ ($\mathbf{P}_{P-1}$ is the backward linear prediction filter of length $(P-1)M$) and minimizing the prediction error energy and assuming that the channel matrix $\mathbf{H}_P$ is of full column rank, we get (as in [3]),

$$\Big[-\mathbf{P}_{P-1} \quad \mathbf{I}_M \Big] \mathbf{Z}(n)\mathbf{Q} = \Big[\mathbf{0}_{M,r} \quad \mathbf{B}(n)[\mathbf{Q}]_{r+1:S} \Big], \tag{34}$$

where $\mathbf{0}_{M,r}$ is a zero matrix of size $M \times r$ and the received data matrix is partitioned as

$$\mathbf{Z}(n) \underset{=}{\Delta} \begin{bmatrix} \tilde{\mathbf{Z}}(n) \\ \mathbf{B}(n) \end{bmatrix} \tag{35}$$

where $\tilde{\mathbf{Z}}(n)$ is the top $(P-1)M \times S$ block-matrix, at the top of $\mathbf{Z}(n)$, while $\mathbf{B}(n)$ is the bottom block-row (of size $M \times S$) of $\mathbf{Z}(n)$. Also the rank-revealing QR decomposition of $\tilde{\mathbf{Z}}(n)$ is given by (as in [3]),

$$\tilde{\boldsymbol{Z}}^H(n)\boldsymbol{\Pi} = \boldsymbol{Q} \begin{bmatrix} \boldsymbol{R_1} & \boldsymbol{R_2} \\ \mathbf{0}_{S-r,r} & \mathbf{0}_{S-r,(P-1)M-r} \end{bmatrix} \tag{36}$$

and $r = P + L_h - 1$, where L_h is the multipath channel delay spread (in symbols), r is also the rank of $\tilde{\boldsymbol{Z}}(n)$. See [3], by the author, for derivation and more details.

Then based on the analysis in [3], on (signal and noise) case and (noise only) case, first the backward prediction parameters are determined (in least squares sense) by

$$\mathbf{P}_{P-1} = [\mathbf{B}(n)\tilde{\mathbf{Z}}^H(n)][\tilde{\mathbf{Z}}(n)\tilde{\mathbf{Z}}^H(n)]^\dagger, \tag{37}$$

where $\dagger$ is the pseudo-inverse operation. Once the rank-revealing QR decomposition of $\tilde{\mathbf{Z}}^H(n)$ is calculated, two signal statistics are computed as

$$\boldsymbol{S}_1 = \mathbf{B}(n)[\mathbf{Q}]_{1:r}, \tag{38}$$

and

$$\boldsymbol{S}_2 = \begin{bmatrix} -\mathbf{P}_{P-1} & \mathbf{I}_M \end{bmatrix} \mathbf{Z}(n)[\mathbf{Q}]_{1:r}. \tag{39}$$

The Frobenius norm of $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ are calculated and then if $norm(\boldsymbol{S}_1) \geq \gamma norm(\boldsymbol{S}_1)$, (where γ is the threshold), then we decide that the primary signal is present. The threshold γ is chosen such that the probability of false alarm is 0.12.

With primary signal transmitting, signal statistic $\boldsymbol{S}_1$ will have a large value, since the received signal at a secondary user, $\mathbf{B}(n)$, (at time $(n-N+1)$), will have non-zero projection along $[\mathbf{Q}]_{1:r}$ (the signal space of $\tilde{\mathbf{Z}}(n)$). However, even then signal $\boldsymbol{S}_2$ statistic will have a small value (theoretically zero), due to (34) (i.e. even with primary signal transmitting). In the absence of the primary signal, it can be shown by an extension of [3], both signal statistics $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ will theoretically be close to zero.

VII. Analysis of the Novel Algorithm

Moreover, unlike [3], we utilize the fact that very few PUs are active in their bands (at the same time/location) to allow the rich theory of compressed sensing theory [9], [17] (i.e. superior reconstruction of sparse signals from limited number of measurements) to be combined with the

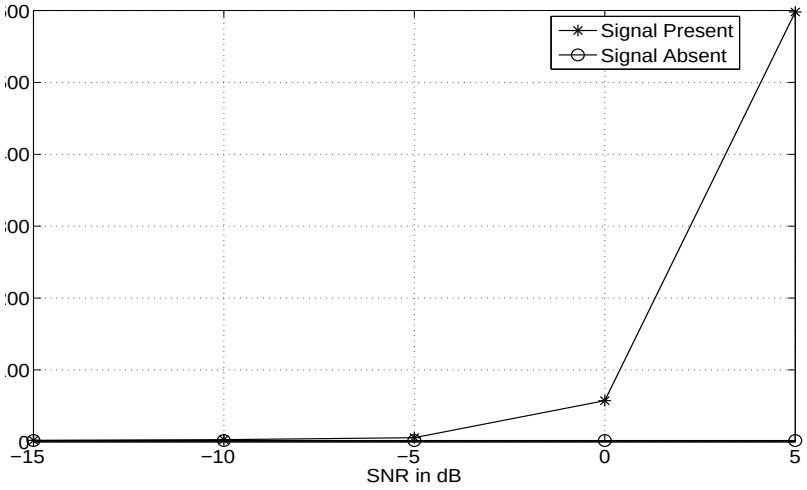

Fig. 7. Ratio of Signal Statistics

modified multi-antennas based blind spectrum sensing algorithms to obtain a practical solution to the challenging problem of blindly detecting PU signals at even SNR of -15 dB, with even as little as 50% of the measurements in [3].

VIII. Simulation Results and Comparison of Different Algorithms

In this section, extensive simulations are provided for the various novel compressed spectrum sensing algorithms, at very negative SNR in dB. Moreover, analytical comparison of the various facets of the different algorithms is undertaken, supported by simulation results.

For all simulations SNR is defined as, i.e.

$$SNR = \frac{E(|\tilde{\mathbf{y}}(n)|^2)}{E(|\tilde{\mathbf{w}}(n)|^2)} \tag{40}$$

where $\tilde{\mathbf{y}}(n)$ is the received signal vector (in time domain, for the multiple antennas) after passing through the channel and $\tilde{\mathbf{w}}(n)$ is the white gaussian noise (in time domain), at varying SNRs. All simulations were carried out over 100 computer experiments. The original sparse signal is simulated (in time domain) with five thousand data symbols, but having 0.6% non-zero samples (in frequency domain, i.e. sparsity level m). The number of iterations is 30 in VOMP, linear prediction order $P = 15$ and $M = 4$ receive antennas are employed.

TABLE I

COMPUTATIONAL TIME (IN SEC.)

Method		**Fixed channel**	**5 taps random channel with one primary user**	**5 taps random channel with two primary users**
VOMP	80% measurements	16.9421	16.9555	16.7288
	50% measurements	14.7458	14.7719	14.8405
VAD	80% measurements	58.123	59.3402	57.6383
	50% measurements	31.8374	29.7203	29.833

A. Reconstruction of noisy sparse signals

Figure 1 shows the sparse signal (in frequency domain) and its reconstruction (from 80% measurements) using one receive antenna by VOMP algorithm, for one and two PU's respectively, first at a SNR of 10 dB, and then at a SNR of -10 dB. Figure 2 shows the sparse signal (in frequency domain) and its reconstruction (from 80% measurements) using one receive antenna by VL1 algorithm, for one and two PU's respectively, first at a SNR of 10 dB, and then at a SNR of -10 dB. Figure 3 shows the sparse signal (in frequency domain) and its reconstruction (from 80% measurements) using one receive antenna by VAD algorithm, for one and two PU's respectively, first at a SNR of 10 dB, and then at a SNR of -10 dB. At SNR of 10 dB, when the reconstruction (to a sparse signal, using the known sparsity level m of PU signal in VOMP) is very good and then at a very noisy SNR of -10 dB, when (as expected), the reconstruction (to a sparse signal) is completely non-sparse and is very different from the original sparse signal. It is to be noted that while the PU signal is sparse, the AWGN noise is not sparse. Reconstruction is best in the case of VOMP, because the VOMP algorithm uses the value of m explicitly, by stopping its iteration at $t = m$, where m is the sparsity level (number of non-zero elements in PU signal), and thus rejecting some of the noise (in other bands).

Yet it will be seen from Figures 4, 5 and 6, that from the reconstructed (signal + noise) at the multiple receive antennas and by employing the proposed blind sensing algorithm (which implicitly utilizes the cross-correlation between the signal components on the multiple receive antennas), detection of primary signals, even at SNR of -15 dB, is possible; this is due to cyclo-stationarity.

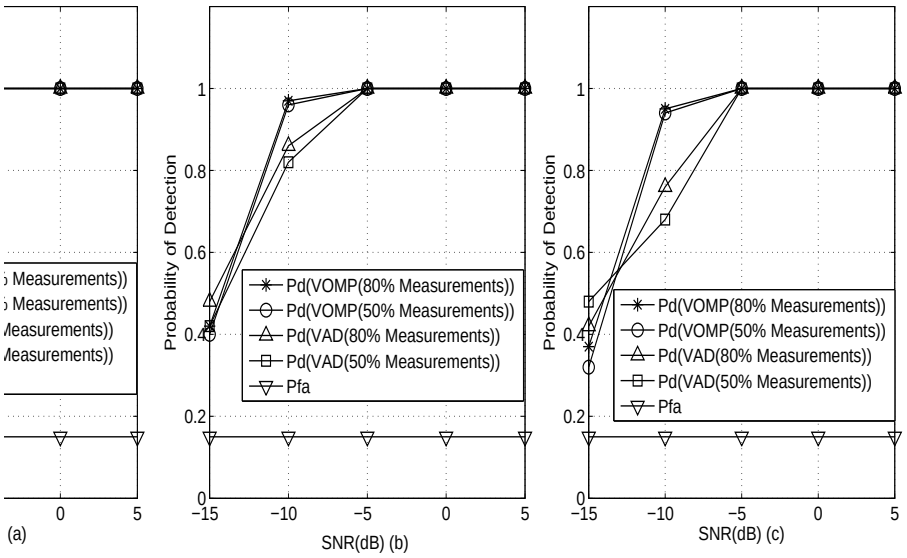

Fig. 8. Probability of Detection at different measurements using VOMP and VAD for (a) single tap fixed channel (having channel coefficients $\mathbf{h}_0$=[0.8,0.3,0.55,0.15] (for the four receive antennas) (b) 5 taps random channel for one primary user (c) 5 taps random channel for two primary users

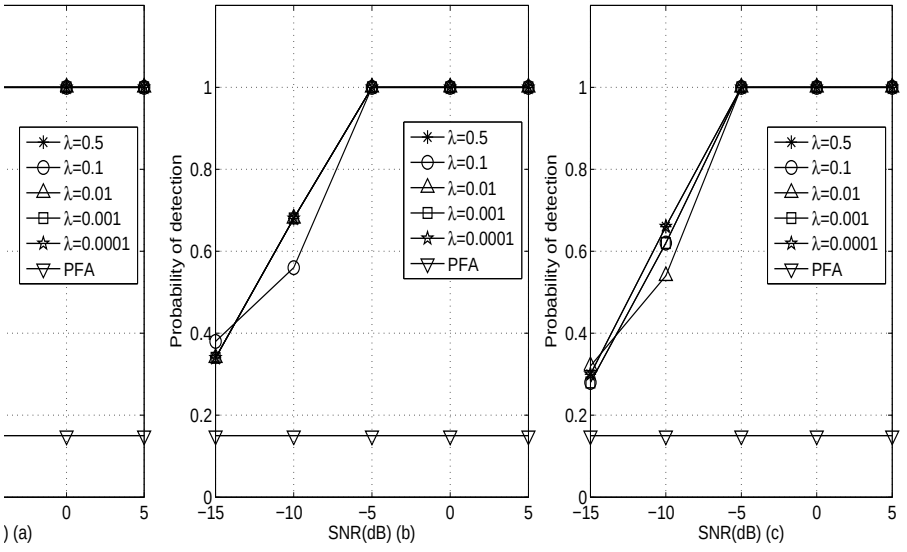

Fig. 9. Probability of Detection using VL1 for different λ in (a) single tap fixed channel (having channel coefficients $\mathbf{h}_0$=[0.8,0.3,0.55,0.15] (for the four receive antennas) (b) 5 taps random channel for one primary user (c) 5 taps random channel for two primary users

TABLE II

COMPUTATIONAL TIME (IN SEC.) FOR VL1 WITH 50% MEASUREMENTS

Method		Fixed channel	5 tap random channel with one primary user
VL1	λ = 0.5	120.32	109.16
VL1	λ = 0.1	66.69	58.39
VL1	λ = 0.01	29.16	27.12

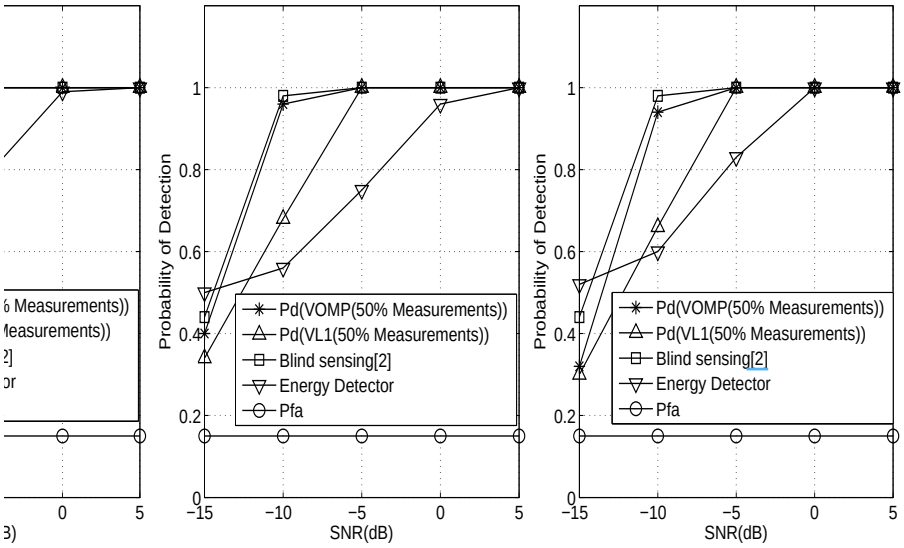

Fig. 10. Comparison between proposed method using VOMP, VL1, VAD with Blind Spectrum Sensing in [2] and Energy Detector for (a) single tap fixed channel (having channel coefficients $\mathbf{h}_0$=[0.8,0.3,0.55,0.15] (for the four receive antennas) (b) 5 taps random channel for one primary user (c) 5 taps random channel for two primary users

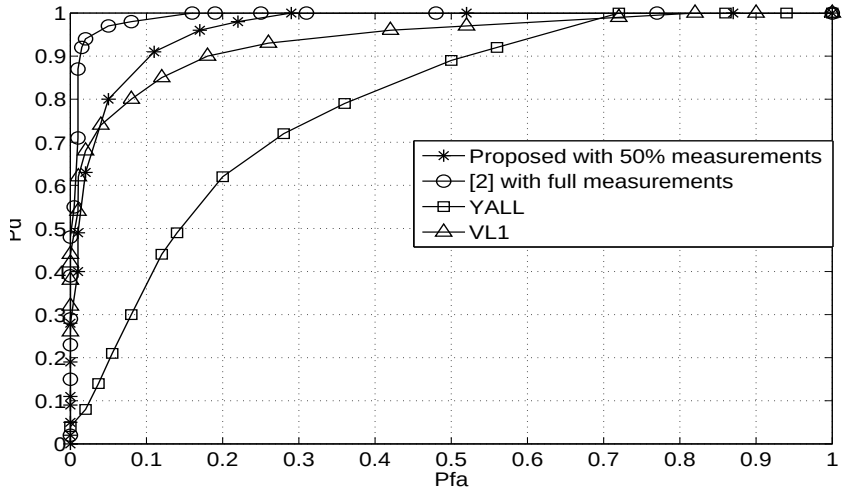

Fig. 11. ROC curve for single tap fixed channel (having channel coefficients $\mathbf{h}_0$=[0.8,0.3,0.55,0.15] (for the four receive antennas)

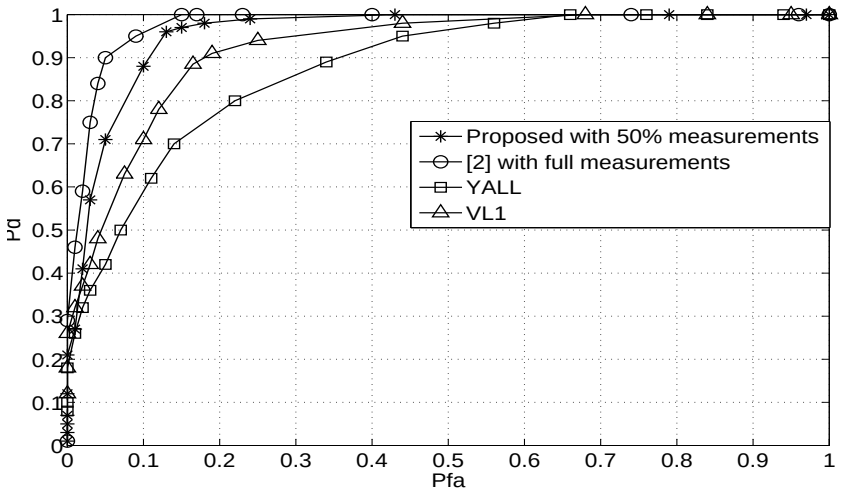

Fig. 12. ROC curve for 5 taps random channel for one primary user

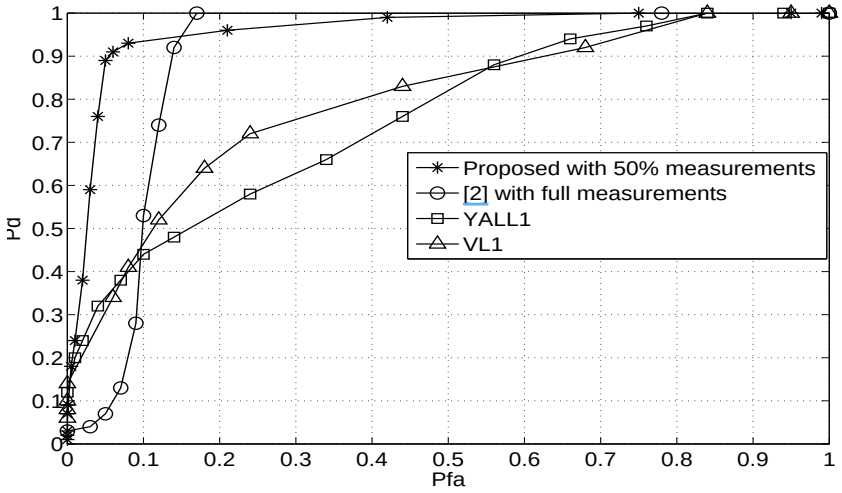

Fig. 13. ROC curve for 5 tap random channel for two primary users

B. Probability of detection and false alarm

In this section, comparative plots of probabilities of detection and false alarm of the proposed (blind) spectrum sensing method (including detecting PU signal at -15 dB), utilizing different compressed sensing algorithms for reconstruction of the weak PU signal, from a varying number of sparse measurements.

Figure 4 shows comparative plots of probability of detection (P_d) and probability of false alarm (P_{fa}), for 80% and 30% measurements (corresponding to 4000 and 1500 data symbols out of total

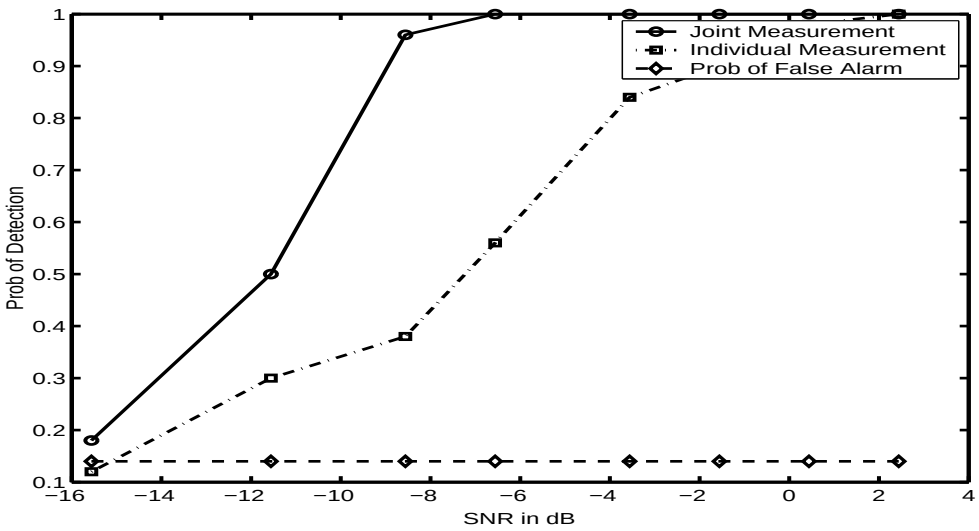

Fig. 14. Probability of Detection using VOMP at 50% measurements for same and different Φ in (a) single tap fixed channel (having channel coefficients $\mathbf{h}_0$=[0.8,0.3,0.55,0.15] (for the four receive antennas) (b) 5 taps random channel for one primary user

of 5000 data symbols), using VOMP and VAD, with SNR ranging from -15 dB to 5 dB. Figure 4 (a) shows the plots of probability of detection and false alarm in a fixed channel, which is a is a single-tap channel with channel coefficient vector $\mathbf{h}_0$=[0.8,0.3,0.55,0.15] (from the PU to the four receive antennas at SU). Figure 4(b) shows the plots of P_d and P_{fa} in a five tap random channel for one PU; figure 4(c) shows the plots of P_d and P_{fa} in five tap random channel for two PUs. It is especially seen in the random channel with one or two PUs, the VOMP algorithm behaves exactly both for 80% and 50% measurements and its performance is much better than that of the VAD, whose performance deteriorates when the number of measurements decrease from 80% to 50% of the total number of measurements.

C. VL1 algorithm with different λ

The compressed sensing problem (for spectrum sensing) in (26), is solved using VL1, for different values of λ to find the optimal value of λ for this problem, plots of P_d of which are shown in Figure 5. It is seen that best performance is achieved for $\lambda = 0.5$. This may be because in equation (25), a larger value of λ imposes the sparsity constraint more strongly.

In Figure 6 (a), (b), (c), comparison is made with the author's algorithm in [3], which uses full measurements (i.e. 5000 data symbols), with in fixed channel, 5 taps random channel with one

PU, 5 taps random channel with two PUs respectively, with the methods in this paper. Even with 50% measurements in the algorithm (with VOMP), P_d is very close to [3]. For VL1 and VAD, the performance of blind spectrum sensing deteriorates, with VL1 performing even worse than the VAD algorithm. While P_d close to 1 is achieved at SNR of -10 dB (for both [2] and VOMP with 50% measurements), for VAD and VL1, it is achieved at SNR of -5 dB.

From Sections VII.B and VIII.C (Figures 4 and 6), it is seen (as far as the plots of P_d versus SNR), the VOMP performs the best as a compressed sensing algorithm for the vital problem of spectrum sensing in cognitive radio.

D. Comparative Computational Speeds

Table I shows the computational time for VOMP, VAD and VL1 (for different values of λ), with varying levels of measurements, for each computer experiment (realization). First, it is seen that for both VOMP and VAD algorithms, for the same number of measurements, each algorithm takes the same time over all the three channel profiles (fixed channel, random channel with one PU, random channel with two PUs). Moreover, as expected, processing 80% of total measurements takes more time than than processing 50% of total measurements. Moreover, the computational speed of VOMP (for this problem) is much faster than of VAD; also, with increase in the number of measurements, the computational time of VOMP does not increase as much as in the VAD algorithm.

It is seen that the computational time of the VL1 algorithm increases with an increase in the value of λ, maybe because increasing value of λ enforces the sparsity constraint more rigorously. Though the best PU signal detection of VL1 is achieved for $\lambda = 0.5$, the computational time (or computational complexity) is very high. Thus the value of λ is determined, according to the situation. However, the performance of the VL1 method is inferior to that of VOMP and VAD based sensing method, as witnessed from Figure 6, as VL1 is not well suited to the problem with large dimension of $\mathbf{A}$ in (7). Moreover, VL1 is of the slowest algorithms, for very sparse signal,

as is the case in this example, with the PU signal having 0.6% non-zero samples in frequency domain. This has also been noted in the paper on face recognition case in [17].

Again, VOMP turns out to be the compressed sensing algorithm of choice for the spectrum sensing problem in cognitive radio, from the computational speed standpoint.

E. Comparative Receiver Operating Characteristics (ROC)s

To compare the different compressed sensing algorithms, from the standpoint of its suitability for blind spectrum sensing (at very negative SNR (dB)), receiver operating characteristics of each method is obtained (ie which illustrates the performance of a binary classifier), by plotting P_d versus P_{fa}, as the decision threshold is varied. ROC curves are plotted for VOMP, VL1 and VAD based blind spectrum sensing algorithm at -10 dB. Since the best compressed sensing method would provide a curve which is the uppermost and left-most part of the ROC curve, next to [3] with full measurements, is the VOMP algorithm blind spectrum sensing method, which is superior to the the VL1 method, which is better than the VAD method,in Figure 8. Simulation results with two PU signals,also vindicates the same results in Figure 9.

F. Joint Sparsity Recovery

The results allow spectrum sensing over a wide band, by using a limited number of measurements (of the received signal- sub-Nyquist sampling), and thus satisfying the timing requirements for rapid sensing. This paper is also a feasibility study of detecting primary signals from limited number of measurements (like -13 dB) (a very important criterion in cognitive radio systems like IEEE 802.22), which distinguishes it from [8], which considers only the noiseless case.

IX. Conclusion

In this paper, utilizing the sparse-ness of the radio spectrum at any given time/location, a limited number of measurements of the received signal (over the wide spectrum) at different multiple antennas are used to reconstruct the complete received signal, using VOMP, VL1 and VAD, after which a novel (modified) blind spectrum sensing is applied to the reconstructed signal.This blind spectrum sensing is accomplished by employing a combination of linear prediction and rank-revealing QR decomposition, of the reconstructed received signal matrix. Till date, integration of compressed sensing and blind spectrum sensing, has not been investigated in very low negative SNR regime. The novel method performs similar to [3], which requires very large numbers of measurements (symbols), similar to other cyclo-stationary based spectrum methods in extremely low SNR conditions. The

proposed method is also blind so does not require any information about the primary signal power, multipath channel distortions (between primary and secondary users) and band(s) occupied by the primary user and can be extended to the case when there are multiple primary users. Simulations show that the novel spectrum sensing method is feasible at SNR of -15 dB, with even 50% measurements, (unlike [8], which does not consider negative SNR (dB) case); thus satisfying the timing requirements for rapid sensing.

REFERENCES

[1] Ghasemi, A. and Sousa, E. S. "Spectrum sensing in cognitive radio networks: requirements, challenges and design trade-offs", *IEEE Commun. Mag.*, 2008, **46**, pp. 32-39

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

[3] De, P. and Liang, Y. C. "Blind spectrum sensing algorithms for cognitive radio networks", *IEEE Trans. Veh. Technol.*, Sep. 2008, **58**, pp. 2834-2842.

[4] M. S. Oude Alink, A. B. Kokkeler, E. A. Klumperink, G . J. Smit, B. Nauta, "Lowering the SNR wall for energy detection using cross-correlation," *IEEE Trans. Veh. Technol.*, vol. 60, no. 8, Oct. 2011, pp. 3748-3757.

[5] J. Font-Segura and X. Wang, "GLRT-based spectrum sensing for cognitive radio with prior information," *IEEE Trans. Communications*, vol. 58, no. 7, July 2010, pp. 2137-2146.

[6] A. Sahai and D. Cabric, "Cyclostationary feature detection," *DySPAN Part II*, 2005.

[7] P. De,"New methods for blind sensing of bandlimited signals in cognitive radio," *IEEE Vehicular Technology Conference*, Dublin, Ireland, Apr. 2007.

[8] Tian, Z. and Giannakis, G. B. "Compressed sensing for wideband cognitive radios", *IEEE Int. Conf. Acoust., Speech, Signal Process.(ICASSP)*, 2007, **4**, pp. IV-1357

[9] Tropp, J. A. and Gilbert, A. C.:"Signal recovery from random measurements via orthogonal matching pursuit", *IEEE Trans. Inf. Theory*, 2007, **53**, pp. 4655-4666

[10] Donoho, D. L.: 'Compressed sensing', *IEEE Trans. Inf. Theory*, 2006, **52**, pp. 1289-1306

[11] J. Meng, W. Yin, H. Li, E. Hossain and Z. Han, "Collaborative spectrum sensing from spares observations in cognitive radio networks," *IEEE J. Sel. Topics Signal Process.*, vol. 29, no. 2, Feb 2011 2010, pp. 327-337.

[12] D. Needell, R. Vershynin, "Signal recovery fromincompleteand inaccurate measurements cvia regularized orthogonal matching pursuit," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, Apr. 2010, pp. 310-316.

[13] D. Needell, R. Vershynin, "Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit," *Found. Comput. Math.*, 2007, DOI: 10.1007/s10208-008-9031-3.

[14] D. Needell and J.A. Tropp, "CoSAMP: Iterative signal recovery from incomplete and inaccurate samples," *Communications of the ACM*, Dec. 2010, pp. 93-100.

[15] S.Mistry and V. Sharma,"New algorithms for wideband spectrum sensing, *IEEE International Conference on Communications*, 2013, pp. 2595-2600.

[16] E.-Castro and Y.C. Eldar, "Noise folding in compressed sensing," *IEEE Signal Processing Letters*, vol. 18, no. 8, Aug 2011, pp. 478-481.

[17] Yang, A. et. al.:'Fast L-1 minimization algorithms and an application in robust face recognition: a review', *Technical Report UCB/EECS-2010-13*, Univ. California Berkeley, Feb. 2010.

[18] S.Kim, K.Koh, M.Lustig, S. Boyd and D. Gorinevsky, "An iterative interior-point method for large-scale l_1 -regularized least squares," *IEEE Journal of Selected Topics in Signal Processing,* 1(4):606-617, 2007.

[19] J. Yang and Y. Zhang, "Alternating direction algorithms for l_1 problems in compressive sensing," *SIAM Journal on Scientific Computing*, 33(1), 250-278, 2011.