Direct Spectrum Sensing from Compressed Measurements

Steven Hong
Dept of Electrical Engineering
Stanford University, Stanford, CA, 94305-9505
Email: hsiying@stanford.edu

***Abstract*—Because current Cognitive Radios are limited in their operational bandwidth by existing hardware devices, much of the extensive theoretical work on spectrum sensing is impossible to realize in practice over a wide frequency band. To solve this problem, many have used Compressive Sensing (CS) in sequence with CRs: first acquiring compressed samples, then reconstructing the Nyquist Rate signal, and lastly performing spectrum sensing on the reconstructed signal. While CS alleviates the bandwidth constraints imposed by front-end ADCs, the resulting increase in computation/complexity is non-trivial, especially in a power-constrained mobile CR. This motivates us to look at different ways to reduce computational complexity while achieving the same goals.**

In this paper, we will demonstrate how directly performing spectrum sensing from the compressed measurements can achieve the sampling reduction advantage of Compressive Sensing with significantly less computational complexity. Our key observation is that the CR does not have to reconstruct the entire signal because it is only interested in detecting the presence of Primary Users. Our algorithm takes advantage of this observation by estimating signal parameters directly from the compressed signal, thereby eliminating the reconstruction stage and reducing the computational complexity. In addition, our framework provides a measure of the quality of estimation allowing the system to optimize its data acquisition process to always acquire the minimum number of compressed measurements, even in a dynamic spectral environment.

I. Introduction

As the number of wireless services devices and services has grown exponentially the past few years, wireless spectrum has become an increasingly valuable commodity. The traditional method of statically allocating blocks of spectrum is fast becoming inadequate as there is no longer any bandwidth left to allocate. While at first glance the shrinking amount of licensed spectrum appears to be a major issue, upon close inspection it becomes apparent that most of the allocated bands are unoccupied more than 90% of the time. However, because these bands have already been allocated to Primary Users (PU) who pay a hefty fee for the exclusive right to the spectrum, the band cannot be used by another user who may cause interference to the PU.

Cognitive Radio (CR) is a spectrally efficient communications paradigm which promises to maximize the efficiency of a wireless system by detecting and transmitting on underused frequency bands while avoiding interference with the Primary User. The key enabling technologies which make a CR unique are its sensing and adaptive capabilities, which allow it to operate over a broad swath of spectrum and adapt its transmissions to fill spectral holes. Because the CR does not need any pre-assigned frequencies to operate, there are numerous commercial and military benefits which can be attained by such a system.

However, due to hardware and sampling constraints, CRs in practice are often limited to a restricted frequency range which severely limits its usefulness. If a CR wishes to detect multiple communication signals falling in a very wide band (HF-VHF), the Nyquist sampling rates exceed the limitations of current state of the art analog-to-digital converters by several orders of magnitude. As a result, most of the CR results derived in literature [1][2][3] are impossible to realize in practice over a wide frequency band.

Recently, an area that has demonstrated a considerable amount of potential to alleviate the sampling bottleneck in wideband communications is Compressive Sensing (CS), which asserts that the one can recover certain signals from far fewer samples than traditionally used, conditioned on the fact that the signal is sufficiently sparse in a particular domain [4][5]. CS relies on this principle of sparsity, which means that the information rate of a continuous time signal is much smaller than suggested by its bandwidths, so that a concise representation of the signal is possible when expressed in the proper basis. As discussed before, due to the low percentage of spectrum occupancy by active radios, wireless communication signals in open-spectrum networks are typically sparse in the frequency domain, allowing us to use compressive sensing to alleviate the sampling bottleneck.

While CS is a powerful technique, it does not allow the CR to sample at low rates for free; the resulting increase in computation/complexity is non-trivial, especially in a power-constrained mobile Cognitive Radio. This motivates our development of a scheme which can perform the functions of a CR (detect primary users in a wide frequency range) with as few samples as possible while incurring minimum computation/complexity. Previous works [6][7] have also realized the synergies between Compressive Sensing and Cognitive Radios and attempted to capitalize by combining both processes in a sequential fashion. In both works, the sensing bottleneck was alleviated with compressive sensing, which was used to capture the signal at a reduced rate. The entire signal was reconstructed from the compressed samples to the Nyquist number of samples, from which the spectrum sensing was then performed. Both papers failed to take into account the significant computational complexity necessitated by the reconstruction technique, an issue which we will directly address and solve in this paper. In addition, [6][7] have to have *a priori* knowledge of the sparsity of the signal in order to determine how many measurements (M) are required to ensure perfect reconstruction. Without this knowledge, they had no metric to assess the quality of their reconstruction. In practice, this knowledge is generally unavailable and the sparsity of the signal will in fact change significantly over time.

In this paper, we make the key observation that because the fundamental task of a CR is not reconstructing the signal, rather it is estimating the presence of Primary Users, the reconstruction stage could be completely eliminated. In order to accomplish this, we propose to use a Bayesian formulation [8] to estimate the parameters of the underlying signal from compressed measurements. We will demonstrate that a Bayesian approach to signal parameter estimation has the following advantages:

- It becomes possible to estimate the parameters of the signal (Carrier Frequency of Primary Users, Bandwidth, Power, etc.) directly from the compressed measurements. The most important thing to note, is that the Bayesian Compressive Sensing (BCS) directly performs the function of a Cognitive Radio Spectrum Detector, estimating the parameters of the primary users.

- Not only are the signal parameters estimated, but "error bars" are also easily estimated, providing a measure of confidence in the estimated parameters; using these error bars, we will design an iterative optimization strategy to acquire the minimum number of samples in a dynamic sparsity model without any *a priori* assumptions.

A BCS framework lends itself well to an interative optimization strategy to acquire the minimum number of samples in a dynamic sparsity model without any *a priori* assumptions. We will demonstrate through simulations how this approach to Primary User detection achieves our primary objective: reducing the computational complexity in Compressive Spectrum Sensing through direct estimation from compressed measurements.

The remainder of the paper is organized as follows. In Sec. II, we'll begin by defining a signal model which demonstrates sparsity so that compressive sensing can be applied. This analysis is extended in Sec. III to demonstrate how a Bayesian signal inversion framework essentially functions as a wideband CR, performing signal reconstruction and spectrum sensing in one step, reducing complexity. In Sec. IV, we will demonstrate how this framework lends itself well to an iterative refining algorithm which can acquire the fewest number of samples without any *a priori* assumptions. Finally, example results for our method, along with comparisons to other algorithms, will be presented in Sec. V with an emphasis on computation complexity, speed, and accuracy.

II. Signal Model

We'll begin by defining the signal model, which demonstrates sparsity so that compressive sensing can be applied [9]. Suppose that a Cognitive Radio wishes to analyze a signal $x(t)$ that spans a total of B_{tot} Hz in the frequency range $[f_0, f_N]$. The range is occupied by $j \in [0, K]$ users, each occupying arbitrary disjoint bands of equal bandwidth B. In order to sample below the Nyquist rate, we assume that $2jB < f_{nyq}$ (if this condition is not satisfied, CS cannot be used), so the signal is sparse in the frequency domain. Suppose that the time window for sensing is $t \in [0, NT_{nyq}]$, where T_{nyq} is the Nyquist sampling period and assuming Nyquist theory, N samples are needed to recover $x(t)$ perfectly without aliasing. A digital receiver may not be able to sample at this rate for a very wideband and so we apply Compressive Sensing to obtain low-rate samples in the following form.

$$\mathbf{y}_t = \mathbf{S}\mathbf{x}_t = \mathbf{S}\mathbf{F}_N^{-1}\mathbf{X}_f = \mathbf{A}\mathbf{X}_f \tag{1}$$

where $\mathbf{S}$ is an MxN projection matrix and $\mathbf{x}_t$ represent the Nyquist rate Nx1 vector with $x_t[n] = x(t)|_{t=nT_0}, n = 1, ..., N$. The columns of $\mathbf{S}$ are a set of matched filters or set of basis signals and the compressed measurements $y_t[m]|_{m=1}^{M}$ are essentially the projection of $\mathbf{x}_t$ onto the basis. The relationship between the measurement lengths is $j << M < N$. $\mathbf{F}_N$ denotes an N point unitary DFT matrix mapping the time domain signal to the frequency domain where $\mathbf{X}_f$ is sparse. $\mathbf{A} = \mathbf{S}\mathbf{F}_N^{-1}$, which is MxN, allows us to express the relationship between the compressed time domain signal, $\mathbf{y}_t$, to the Nyquist rate frequency domain sampled signal $\mathbf{X}_f$.

III. Bayesian Modeling of CS Inversion

With Bayesian Modeling, all of the unknowns are treated as stochastic quantities with a probability distribution function. The unknown signal, $\mathbf{X}_f$, is assigned a *prior* distribution $p(\mathbf{X}_f|\boldsymbol{\beta})$, modeling its nature. The observation of the compressed samples is also a random process with a conditional distribution $p(\boldsymbol{y}_t|\boldsymbol{X}_f, \beta_0)$. We also note that the measurements $\boldsymbol{y}_t$ may be noisy and so we add measurement noise, $\boldsymbol{n}_f$, to our model, giving us $\boldsymbol{y}_t = \boldsymbol{A}\boldsymbol{X}_f + \boldsymbol{n}_f$ where $\boldsymbol{n}_f$ is zero mean uncorrelated Gaussian noise with variance σ^2. $\boldsymbol{\beta}$ and β_0 are hyperparameters associated with the weight of each estimate which dictate the prior; β is the inverse variance (precision) of a Gaussian pdf and $\beta_0 = 1/\sigma^2$ is the noise precision. The conditional distribution of the observation thus becomes a Gaussian liklihood model

$$p(\boldsymbol{y}_t|\boldsymbol{X}_f, \beta_0) = (2\pi/\beta_0)^{-M/2} exp(-\frac{\beta_0}{2}||\boldsymbol{y}_t - \boldsymbol{A}\boldsymbol{X}_f||^2) \tag{2}$$

which we seek to estimate. This converts the CS problem of reconstructing the sparse frequency domain samples $\boldsymbol{X}_f$ into a linear regression problem with the constraint that $\boldsymbol{X}_f$ is sparse. Assuming knowledge of $\boldsymbol{A}$, the goal is to estimate the parameters of $\boldsymbol{X}_f$ and σ^2. We begin by looking at point modeling, which is the traditional BCS method of estimating parameters [8]. We first define a zero-mean Gaussian prior on each element of $\boldsymbol{X}_f$.

$$\begin{aligned} p(\boldsymbol{X}_f|\boldsymbol{\beta}) &= \prod_{n=1}^{N} N(X_n|0, \beta_n^{-1}) & (3) \\ &= (2\pi)^{-N/2} \prod_{n=1}^{N} \beta_n^{1/2} exp(-\frac{\beta_n X_{f_n}^2}{2}) & (4) \end{aligned}$$

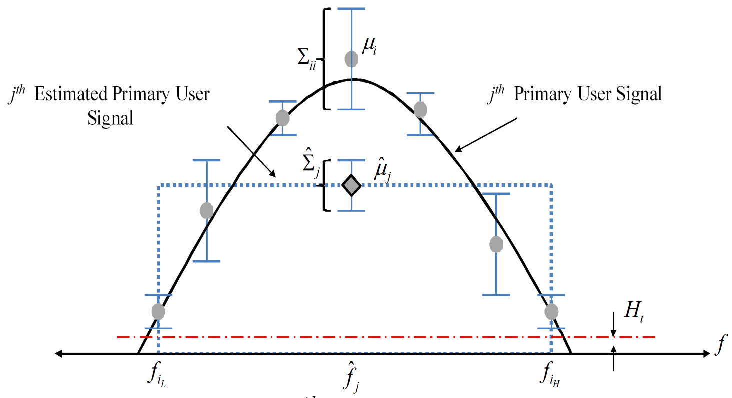

Fig. 1. The signal of the j^{th} Primary User. The gray points represent the weights, μ_i, of the point estimates and the blue "error bars" represent the variance of the estimate, Σ_{ii}. Note that closer the estimate is to the true signal, the smaller the "error bars". $\hat{\Sigma}_j$ and $\hat{\mu}_j$ represent the fused parameters ("error bar" and average magnitude respectively) describing the user's signal and the width of the dotted blue box is the bandwidth of the estimated signal.

Conditioning over the unknown precision and the known compressed samples, the full parameter distribution condition is obtained by combining the liklihood with Bayes' rule

$$p(\boldsymbol{X}_f|\boldsymbol{y}_t, \boldsymbol{\beta}, \beta_0) = \frac{p(\boldsymbol{y}_t|\boldsymbol{X}_f, \beta_0)p(\boldsymbol{X}_f|\boldsymbol{\beta})}{p(\boldsymbol{y}_t|\boldsymbol{\beta}, \beta_0)} \quad (5)$$

$$= (2\pi)^{-\frac{N}{2}}|\boldsymbol{\Sigma}|^{-\frac{1}{2}} exp\left\{-\frac{1}{2}(\boldsymbol{X}_f - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{X}_f - \boldsymbol{\mu})\right\} \quad (6)$$

which is a Gaussian distribution $N(\mu, \Sigma)$ with mean and covariance of

$$\mu = \beta_0 \Sigma \boldsymbol{A}^T \boldsymbol{y}_t \quad (7)$$
$$\boldsymbol{\Sigma} = (\boldsymbol{B} + \beta_0 \boldsymbol{A}^T \boldsymbol{A})^{-1} \quad (8)$$

where $\boldsymbol{B} = diag\{\beta_1, ..., \beta_N\}$. μ is the point estimate of the signal while the diagonal elements of the covariance matrix in (8) provide "error bars" on the accuracy of the point estimate μ.

IV. Primary User Detection using BCS

BCS inversion in essence allows the mapping of the compressed time domain measurements $\boldsymbol{y}_t$ to the Nyquist rate frequency domain sampled signal $\boldsymbol{X}_f$ through the entries in the projection matrix $\mathbf{A}$. Using the point wise estimates in (7) and the corresponding "error bars" in (8), we can reconstruct our signal $\boldsymbol{X}_f$ with probability overwhelmingly close to 1. However, as we mentioned earlier, *reconstructing the signal entirely is not our desired goal; our goal is to estimate the parameters of the primary user signals*. To do so, we must fuse the parameters describing the primary user signals.

In (7) and (8), all of the parameters are known except $\boldsymbol{\beta}$ and β_0 which can be estimated by marginal liklihood maximization techniques such as the relevance vector machine (RVM)[10]. The RVM is is a type-II maximum-liklihood (ML) technique which estimates the most probable point estimate $\beta_{MP(i)} \ \ \forall\ i = 1, ..., N$ for each desired point. The advantage of using RVM over other reconstruction methods is that the RVM enables the sequential addition or deletion of candidate basis functions (columns of $\mathbf{A}$) to find $\beta_{MP(i)}$. This allows the algorithm to acquire more samples to mitigate errors or acquire fewer samples if the error is small enough, enabling the reconstruction to always operate with the minimum number of samples, even in changing environments without *a priori* knowledge of the sparsity. Using RVM in the traditional sense [10] will enable us to find the most probable point estimate $\beta_{MP(i)} \ \ \forall\ i = 1, ..., N$ for the entire signal. But estimating $\beta_{MP(i)}$ for the points in $\boldsymbol{X}_f$ where one of the j Primary Users signals is not present is a waste of time, since we are *not interested in reconstructing the entire signal*. By fusing the parameters describing the primary user signals, we can utilize RVM to estimate only the points describing primary user signals with the fewest number of samples.

A. Primary User Parameter Fusion

Most of the computational complexity from the RVM comes from the iterative optimization to obtain $\beta_{MP(i)}$. So to avoid this expensive optimization on empty parts of the spectrum, we set a threshold H_t. If $\mu_i > H_t$, we optimize the weight and if $\mu_i < H_t$, we will not optimize the weight. After we've set this threshold, we group contiguous point estimates together and use those contiguous points to represent the j^{th} user.

As you can see from Fig. 1, the signal of each user is described with multiple point estimates. We want to find a composite feature point to describe the entire user, since our goal is to estimate the parameters of these users. The i^{th} point estimate can be described by 3 parameters

$$\{[f_i, \mu_i], \Sigma_{ii}\} \quad (9)$$

where f_i represents the spectral location, μ_i represents the magnitude, and Σ_{ii} represents the estimation error. Using contiguous point estimates whose weights satisfy the threshold, H_t, we want to fuse their information into one composite feature point, describing the bandwidth, the carrier frequency, the average magnitude, and the estimation error of the j^{th} user. The bandwidth, B_j, for the j^{th} user is estimated as follows:

$$\forall\ k_c \ \ s.t. \ \ \mu_k > H_t, \ \ f_k \in [f_{i_L}, f_{i_H}] \quad (10)$$
$$B_j = f_{i_H} - f_{i_L}$$

where k_c is the index for point estimates contiguous in the frequency domain. To find the average magnitude, carrier frequency, and estimation error, we form a composite feature point to describe the j^{th} user

$$\left\{[\hat{f}_j, \hat{\mu}_j], \hat{\Sigma}_j\right\} \quad (11)$$

which is a fusion of the point estimate parameters $[f_k, \mu_k]$ weighted by the precision Σ_{kk}^{-1} satisfying $k_c \ \ s.t. \ \ \mu_k > H_t$

$$\hat{\Sigma}_j^{-1} \begin{bmatrix} \hat{f}_j \\ \hat{\mu}_j \end{bmatrix} = \sum_{k \in \{i_L, i_H\}} \Sigma_{kk}^{-1} \begin{bmatrix} f_k \\ \mu_k \end{bmatrix} \quad (12)$$

where $\hat{\Sigma}_j^{-1} = \sum_{k \in \{i_L, i_H\}} \Sigma_{kk}^{-1}$. So inverting $\hat{\Sigma}_j^{-1}$ to find the average magnitude and carrier frequency of the j^{th} signal we get

$$\begin{bmatrix}\hat{f}_j\\ \hat{\mu}_j\end{bmatrix} = \hat{\Sigma}_j \sum_{k\in\{i_L,i_H\}} \Sigma_{kk}^{-1}\begin{bmatrix}f_k\\ \mu_k\end{bmatrix} \tag{13}$$

B. Optimizing the Hyperparameters

Now that we've modeled each user's signal parameters in terms of the point parameters, we must still estimate the most probable point estimates $\beta_{MP(i)} \quad \forall\ i \in \boldsymbol{J}$ where $\mathbf{J}$ consists of the detected users. Because the optimization of the hyperparameters $\boldsymbol{\beta}$ and β_0 is the most computationally complex component, we save tremendously by optimizing the hyperparameters based on the estimates of the Primary User signal and not on the signal in its entirety.

Our the goal is to maximize our confidence in our estimate, (5), but as it was shown in [10], rather than evaluating (5) explicitly, the more effective way to obtain (6) is to write

$$p(\boldsymbol{X}_f|\boldsymbol{y}_t,\boldsymbol{\beta},\beta_0)p(\boldsymbol{y}_t|\boldsymbol{\beta},\beta_0) = p(\boldsymbol{y}_t|\boldsymbol{X}_f,\beta_0)p(\boldsymbol{X}_f|\boldsymbol{\beta}) \tag{14}$$

and integrate over $\boldsymbol{X}_f$. By expanding the right-hand side, gathering together all terms in $\boldsymbol{X}_f$ appearing within the exponential, and completing the square we get the following equation which is now completely in terms of our hyperparameters $\boldsymbol{\beta}$ and β_0:

$$\begin{aligned} p(\boldsymbol{y}_t|\boldsymbol{\beta},\beta_0) = &\quad \int_{-\infty}^{\infty} p(\boldsymbol{y}_t|\boldsymbol{X}_f,\beta_0)p(\boldsymbol{X}_f|\boldsymbol{\beta})d\boldsymbol{X}_f & (15)\\ = &\quad (2\pi)^{-\frac{M}{2}}|\mathcal{Z}|^{-\frac{1}{2}}exp\left\{-\tfrac{1}{2}\boldsymbol{y}_t^T\boldsymbol{\mathcal{Z}}^{-1}\boldsymbol{y}_t\right\} & (16)\end{aligned}$$

where

$$\boldsymbol{\mathcal{Z}} = \boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^T + \beta_0^{-1}\boldsymbol{I} \tag{17}$$

Now we wish to find the values of $\boldsymbol{\beta}$ and β_0 which maximize (15). We do so by first differentiating (16) w.r.t. $\boldsymbol{\beta}$ and setting the equation equal to 0, giving us

$$\beta_i^{update} = \gamma_i/\mu_i^2 \qquad i \in \boldsymbol{J} \tag{18}$$

where μ_i is the i^{th} posterior mean weight and γ_i is defined as

$$\gamma_i = 1 - \beta_i\Sigma_{ii} \tag{19}$$

where Σ_{ii} is the i^{th} diagonal element of the posterior signal covariance in (8) computed with the updated $\boldsymbol{\beta}$ and β_0 values. Following the same process for the noise precision β_0 we get the following updated value

$$\beta_0^{update} = \frac{M - \sum_i \gamma_i}{||\boldsymbol{y}_t - \boldsymbol{A}\boldsymbol{\mu}||^2} \tag{20}$$

To maximize the marginal liklihood, the calculation of (18) and (20) is repeated while iteratively updating the composite feature points of the Primary Users, $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$, until a certain convergence criterion have been satisfied. We define this convergence criterion as the average precision of each user's signal, which must satisfy some threshold, Γ:

$$\frac{1}{|\boldsymbol{J}|}\sum_{k\in\boldsymbol{J}} \hat{\Sigma}_k < \Gamma \tag{21}$$

Until (21) has been satisfied, the system iterates between the process in Sec. IV.A and Sec. IV.B. In each time step, the system augment $\mathbf{A}$ by adding a $(M+1)^{th}$ row represented by $\boldsymbol{a}_{M+1}^T$. We can either select this new projection vector randomly or select it to optimize $\boldsymbol{a}_{M+1}^T\boldsymbol{\Sigma}\boldsymbol{a}_{M+1}$. We omit discussion of this optimization and refer the reader to [8] which studied different ways to optimize this additional projection vector. This augmented projection matrix $\mathbf{A}$ is then used to refine the composite feature point for each user (11), which is then optimized again. If (21) still has not been satisfied, then the process is repeated by further augmenting $\mathbf{A}$.

On the other hand, if $\frac{1}{|\boldsymbol{J}|}\sum_{k\in\boldsymbol{J}}\hat{\Sigma}_k << \Gamma$ due to a decrease in the number of users in the system, the algorithm will subtract a projection vector (remove a column from $\mathbf{A}$). It will continue to reduce the number of measurements that it acquires until $\frac{1}{|\boldsymbol{J}|}\sum_{k\in\boldsymbol{J}}\hat{\Sigma}_k \approx \Gamma$. This adaptive process allows the system to always acquire the fewest samples to estimate the parameters of the primary users, without any *a priori* knowledge of the signal sparsity.

In Sec. VI we'll show numerically that when this convergence criterion is met, the parameters of the signal will have been estimated accurately.

V. Simulations

In this section, we present some simulation results to demonstrate the advantages of directly estimating PU signals from compressed measurements. We consider a radio scene occupied by $j \in [0, K]$ users, each occupying arbitrary disjoint bands of equal bandwidth $B = 10Hz$. Each user employs Binary Phase Shift Keying (BPSK) and the carrier frequency is randomly generated between 0-500 Hz. We model measurement noise as a zero-mean Gaussian with $\sigma = 0.005$. In each of our simulations, the total occupied bandwidth is $2jB << f_{nyq}$ and so the received signal is sparse in the frequency domain. We use a 512 point unitary DFT matrix mapping the time domain signal to the frequency domain where $\mathbf{X}_f$ is sparse, so the Nyquist rate signal has length N=512. We demonstrated the algorithm on a 512 Hz signal for simplicity but the same principles apply for a wideband signal.

We compare the BCS signal detection in this paper against a commonly used signal reconstruction method called Basis Pursuit (BP) [5], an inversion method which uses linear programming. The three metrics we will use to compare them are 1) Computational Time, which measures the computational complexity of the algorithm, 2) Reconstruction Error, which measures the accuracy of the inversion 3) Minimum number of samples to satisfy Γ, which measures how much compression can be achieved.

A. Simulation 1 - PU Detection with BCS

The first simulation in Fig. 2 illustrates how the parameters of the primary users can be estimated directly from compressed measurements. We show snapshots at different time intervals to demonstrate how the estimation of the primary user parameters is refined over time. In Fig.2(b), there are several point estimates which do not match the original signal, but

TABLE I
COMPARISON OF INVERSION TECHNIQUES

Inversion Technique	Computation Time (s)	Reconstruction Error 200 samples: $\frac{\|\|\hat{\boldsymbol{X}}_f-\boldsymbol{X}_f\|\|}{\|\|\boldsymbol{X}_f\|\|}$	Sample $_{min}$ satisfying Γ
Basis Pursuit (BP)	1.3482	.2381	223
Bayesian Compressive Sensing (BCS)	.3102	.0891	164

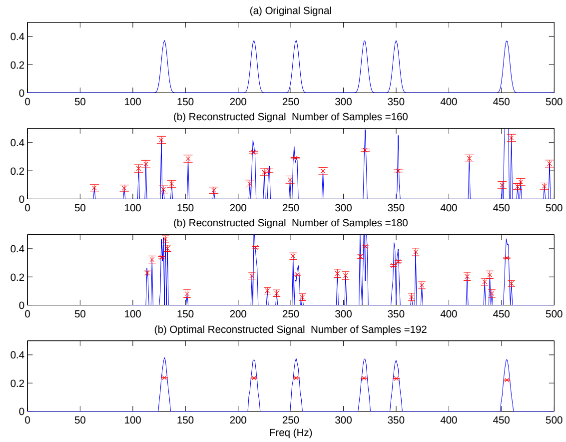

Fig. 2. Detection of Primary Users using BPSK modulation for K = 6, M = 160, N = 512, $\Gamma = 2 * 10^{-4}$. (a) Original signal; (b) Reconstructed signal with 160 samples (Error = $4.9913 * 10^{-4}$); (c) Reconstructed signal with 180 samples (Error = $3.1008 * 10^{-4}$); (d) Reconstructed signal with 192 samples (Error = $1.9983 * 10^{-4}$)

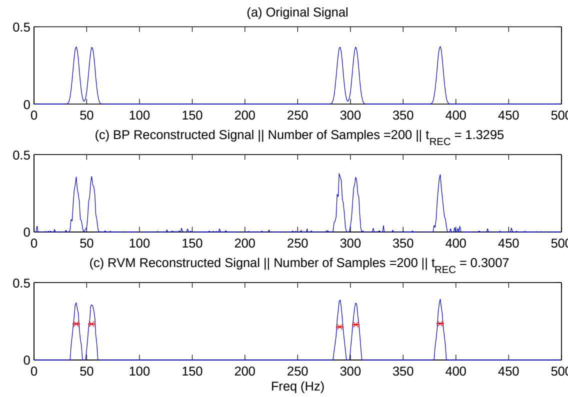

Fig. 3. Comparison of signal estimation using BP, BCS PU detection, and MRBCS PU detection for K = 6, M = 160, N = 512, $\Gamma = 2{*}10^{-4}$. (a) Original signal, (b) Reconstruction with BP, $t_{REC} = 1.3482$ secs (c) Reconstruction with BCS PU detection, $t_{REC} = 0.3102$ secs (d) Reconstruction with MRBCS PU detection, $t_{REC} = 0.0551$ secs

these have large error bars and will be discarded in subsequent iterations. In Fig.2(d), it is shown that $Sample_{min}$ (the minimum number of samples to satisfy the hyperparameter convergence criterion) was 192 samples. Sampling at only 37.5% of the Nyquist rate, we were still able to estimate of the bandwidth, carrier frequency, and magnitude of each user's signal very accurately.

Next we make the comparison between our BCS PU detection algorithm and the traditional BP reconstruction [5] shown in Fig. 3. As you can see from the figure, BCS PU detection is much cleaner than BP, as the relative reconstruction error of BCS is less than a third of the error incurred by BP as shown in Table I. In addition, because BCS PU detection refines only the Primary User signals, and not the entire communications signal as BP does, it achieves this improved result in less than 20% of the time that it takes BP to reconstruct the entire signal. Finally, BCS PU detection also yields a confidence metric in the form of "error bars", this allow it to decide in the next instance whether to increase or decrease the number of samples to meet certain application requirements. This is a feature that BP does not have.

B. Simulation 2 - Impact of Dynamics

The samping rate reduction is a function of how sparse the signal is, which in our setting is a function of how many users there are. Our fourth simulation highlights how the Bayesian modeling of the CS inversion requires no *a priori* knowledge and responds well in changing environments. In this simulation, 4 of the 8 users will stop transmitting at a certain time, $T_0 = 100$ after the optimum number of samples satisfying (21) has been found, they will begin transmitting again at time $2T_0 = 200$. Fig. 4 demonstrates that the signal detection when there are fewer Primary Users requires less samples than when there are more Primary Users. Fig. 5 demonstrates how our adaptive algorithm always achieves $Sample_{min}$ by measuring the error bars of the detected Primary Users. Using our BCS PU detection, the CR is able to always acquire the minimum number of samples and adapt to a dynamic spectral environment. Other reconstruction techniques would require prior knowledge of this change to adapt the measurement procedure but using Bayesian modeling makes this unnecessary. Once the system notices that the error is significantly below the threshold, it will reduce the number of measurements it makes until it finds $Sample_{min}$. At the initial time $T = 0$, with 8 users the algorithm finds that $Sample_{min}$ is 224 samples, 21.9% of the Nyquist rate. At time $T = T_0$, the number of users has decreased and so the algorithm begins to reduce its sampling rate until finding the new $Sample_{min}$ is 151 samples, 14.8% of the Nyquist rate. At time $T = T_0$, the algorithm finds that $Sample_{min}$ has once again become 224 samples, 21.9% of the Nyquist rate. We demonstrate our adaptive algorithm with a Nyquist Rate signal of only 512 samples for ease of implementation, but from these results the complexity savings at high Nyquist Rates (several GHz) would be immense.

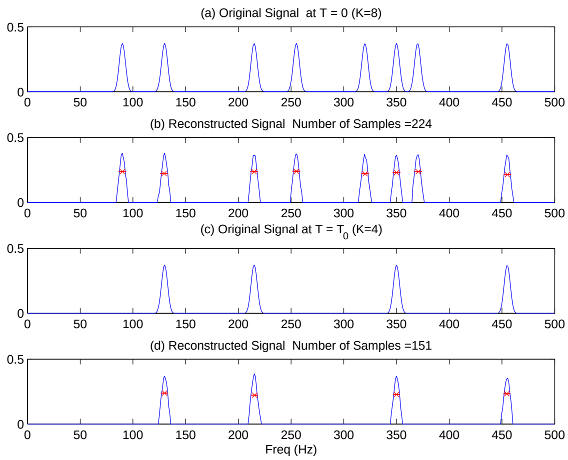

Fig. 4. Adaptive detection as the number of Primary Users decreases and the signal becomes more sparse, K = 8 at T=0, K = 4 at T = T_0, M = 160, N = 512, $\Gamma = 2 * 10^{-4}$. a) Original signal at T=0; (b) Reconstructed signal at T=0; (c)Original signal at $T = T_0$; (d) Reconstructed signal at $T = T_0$

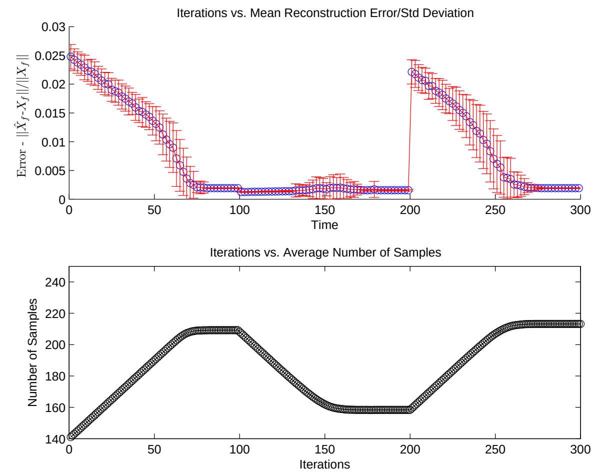

Fig. 5. Adaptive detection: K = 8 at T=0, K = 4 at T = T_0, K = 4 at T = $2T_0$, M = 140, N = 512, $\Gamma = 2 * 10^{-4}$. a) Reconstruction Error - Ran scenario 300x, blue dots are the mean error at each time step and red bars are the standard deviation of the error. (b) Average Number of Samples

VI. Conclusions

In this paper, we highlighted the potential for CS in CRs to alleviate the sampling limitations of modern analog-to-digital hardware. We then discussed how CS is a powerful technique, but the resulting increase in computation/complexity is non-trivial, especially in a power-constrained mobile Cognitive Radio. This motivated our development of the BCS PU detection scheme which directly estimated the Primary User signal from compressed measurements.

One of the key observations of this paper was the fact that when applying CS to CRs, the fundamental task of the CR remains the same: detect the presence of Primary Users. At this point, our work diverges from the prior applications of CS to Primary User Detection and establishes a new precedent. Because the goal of the CR is not to reconstruct the signal, rather it is to estimate the presence of Primary Users, the reconstruction stage can be completely eliminated. In order to accomplish this, we used a Bayesian formulation to estimate the parameters of the underlying signal from compressed measurements. We demonstrated in multiple simulations how the error bars provided by such a Bayesian formulation could be used to design an iterative optimization strategy to acquire the minimum number of samples in a dynamic sparsity model without any *a priori* assumptions. When compared with traditional signal inversion strategies such as BP, the BCS PU detection is less computationally expensive, more accurate, and achieves a higher compression rate. In summary, the BCS PU detection allows us to achieve our primary goal: significantly reducing the computational complexity associated with Compressive Sensing in Cognitive Radios.

References

[1] S. Haykin, "Cognitive Radio: Brain-Empowered Wireless Communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, 2005.

[2] Y. Liang, Y. Zeng, E. Peh, A. Hoang, "Sensing-Throughput Tradeoff for Cognitive Radio Networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 4, 2008.

[3] Z. Quan, S. Cui, A. Sayed, "Optimal Linear Cooperation for Spectrum Sensing in COgnitive Radio Networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, 2008.

[4] E. Candes, Compressive Sampling *Proceedings of the International Congress of Mathematicians*, Madrid, Spain, 2006.

[5] S. Chen, D.L. Donoho, M.A. Saunders, "Atomic Decomposition by Basis Pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, 1999

[6] Z. Tian, G. Giannakis,"Compressed Sensing for Wideband Cognitive Radios", *International Conference on Acoustics, Speech, and Signal Processing*, Honolulu, Hawaii, 2007.

[7] Y. Wang, A. Pandharipande, Y. Polo, G. Leus,"Distributed Compressive Wide-Band Spectrum Sensing" *Information Theory and Applications Workshop*, 2009.

[8] S. Ji, Y. Xue, L. Carin, "Bayesian Compressive Sensing", *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2346-2356, June 2008.

[9] M. Mishali, Y. Eldar, "From Theory to Practice: Sub-Nyquist Sampling of Sparse Wideband Analog Signals, *IEEE Selected Topics on Signal Processing*, arXiv:0902.4291v3, Feb. 2009

[10] M.E. Tipping, "Sparse Bayesian Learning and the Relevance Vector Machine," *Journal of Machine Learning Research*, vol. 1, pp. 211-244, 2001.