

Compressive Sensing with Directly Recoverable
Optimal Basis and Applications in Spectrum Sensing

Youngjune Gwon, H. T. Kung, and Dario Vlah

TR-08-11



Computer Science Group
Harvard University
Cambridge, Massachusetts

Compressive Sensing with Directly Recoverable Optimal Basis and Applications in Spectrum Sensing

Youngjune Gwon, H. T. Kung, and Dario Vlah
Harvard University

Abstract—We describe a method of integrating Karhunen-Loève Transform (KLT) into compressive sensing, which can as a result leverage KLT’s optimality in revealing the sparsity of a signal. We present two complementary results: (1) by using the KLT to find the optimal basis for decoding we can drastically reduce the number of measurements for compressive sensing used in applications such as spectrum sensing; (2) by using compressive sensing we can compute the KLT basis directly from measurements of the input signal, with substantially fewer samples than the Nyquist rate. For a non-stationary signal, we suggest strategies in addressing the trade-off of incurring additional measurements for updating a KLT basis or compensating an obsolete KLT basis in signal recovery. We validate our results with field experiments to detect multiple radio transmitters and sense the UHF spectrums using software-defined radios.

I. INTRODUCTION

Compressive sensing [1] is an emerging technology that has drawn considerable attention recently for its capability to acquire and extract critical information efficiently. It has found applications in various fields such as medical imaging, cognitive radio, wireless communication, sensor networks [2], cloud computing [3], and IC Trojan detection [4]. Three features of compressive sensing are worth noting: First, the encoding of compressive measurements is blind to the content of the signal and has low computational complexity. Secondly, the number of measurements required to decode a signal is approximately proportional to its information content, *i.e.*, the sparsity of the signal, not its length. Lastly, optimal bases or transforms custom-designed at the decoding time can be used for decoding to reveal the sparsity of the signal of present interest. These properties have made compressive sensing especially attractive in sensing applications involving large amounts of measurement data where effective compression is critical, and important information is sparse.

The Karhunen-Loève Transform (KLT) [5] is a classical transform for a signal. KLT is optimal in the sense that it completely decorrelates the signal and maximally compacts the information contained in the signal. Given its optimality in revealing the sparsity of a signal, it should be natural to explore the use of the KLT basis in compressive sensing decoding. However, KLT has a well-known drawback in that it is dependent on the signal. That is, the KLT basis is determined by the eigenvalue decomposition of the autocorrelation matrix of the signal so far. Thus, when the signal changes so also KLT. Note that it would be inappropriate to compute the KLT basis based on the recovered signal. This is because the latter is only approximately exact, the resulting KLT basis would have built-in errors making it ineffective in improving the

accuracy of future recovery. Because of this signal dependence of KLT, other discrete transforms such as cosine transform (DCT), even though suboptimal, have been extremely popular in compressive sensing.

In this paper, we propose an approach to fix this undesirable situation where optimal KLT is not being used in compressive sensing decoding. Our key insight is that for a sparse signal we can recover its KLT basis directly from random projections of the signal, just like we normally do in recovering the signal itself with compressive sensing. This allows us to update the KLT basis with many fewer samples than when sampling at the Nyquist rate, while using measurements based on the original input signal rather than the recovered one.

We summarize the main contributions of this paper concerning the use of optimal bases in compressive sensing: (1) We show drastic reduction of required measurements when an optimal KLT basis is used in decoding; (2) We describe a compressive sensing method of updating the optimal KLT basis directly from measurements of the input signal, which only requires a non-adaptive basis for decoding such as the Fourier basis. Although the Fourier basis is suboptimal due to its inability to adapt to data, for those signals such as the UHF channels which exhibit sparsity in the Fourier domain, the recovery is still substantially less expensive in measurements than an uncompressed approach; (3) We show how the above two results can work together. That is, we use an adaptive optimal basis in recovering the input signal, and use a non-adaptive suboptimal basis in recovering an updated optimal basis. Although the latter requires more measurements, it need not be performed as frequently as the former. This leads to an overall gain in performance; (4) We demonstrate the performance gains with real-world sensing data measured in the field by software-defined radios.

To our knowledge, we are the first in the literature to suggest the approach of compressive sensing with directly recoverable optimal basis, as outlined in (1), (2), and (3) above.

II. BACKGROUND

In this section, we briefly review how compressive sensing works and discuss the important properties of the compressive sensing decoding related to the use of appropriate transform or basis to reveal sparsity. We also review the Karhunen-Loève Transform, which is known in the literature to be *mean squared error* (MSE) optimal.

A. Compressive Sensing

Compressive sensing [1] exploits a *sparse* structure present in data. Some data naturally exhibit sparsity while others can be made sparse in some domains such that only a few significant coefficients can represent them. For example, after applying the Fourier Transform to decompose sinusoids oscillating in the time domain into their frequency components, we often find that just a few significant frequency components will be sufficient to represent the signal.

The exploitation of sparsity is in fact the common theme behind most compression schemes. The fundamental premise of compressive sensing is that we can *directly* capture K significant coefficients without completely analyzing the data of size N where $K \ll N$. Traditional compression schemes, however, would require statistical analysis of the entire data. Compressive sensing achieves non-analytical, low-complexity encoding with an $M \times N$ measurement matrix Φ , which can be generated randomly according to some known distributions (*e.g.*, Gaussian or Bernoulli).

Now consider a real-valued (for simplicity) data \mathbf{x} of size N , which has a transform Ψ such that in $\mathbf{s} = \Psi\mathbf{x}$ there are only K *nonzero* elements. We say such signal K -sparse. Compressive sensing encodes \mathbf{x} in compressive measurements $\mathbf{y} = \Phi\mathbf{x}$, a reduction of size from N to $M \ll N$ where $M \geq cK \log \frac{N}{K}$ for some small constant $c > 0$. The compressive measurements form an *underdetermined* linear system of equations, *i.e.*, there are more unknowns (N) than equations (M), thereby subject to many solutions for \mathbf{x} .

According to compressive sensing theory, however, *Restricted Isometry Property* (RIP) [6] of $\Phi\Psi^{-1}$ guarantees the exact recovery of \mathbf{x} with high probability if $M \geq cK \log \frac{N}{K}$ by choosing the solution that minimizes ℓ_1 -norm of \mathbf{s} :

$$\min_{\mathbf{s} \in \mathbb{R}^N} \|\mathbf{s}\|_{\ell_1} \quad \text{subject to} \quad \mathbf{y} = (\Phi\Psi^{-1})\mathbf{s} \quad (1)$$

Linear programming can solve (1). (\mathbf{s} is recovered first, and from $\mathbf{x} = \Psi^{-1}\mathbf{s}$, \mathbf{x} can be recovered.)

An interesting property of the ℓ_1 -minimization is that the quality of the decoding is a function of M . The larger M is, the more accurate the reconstruction becomes. Furthermore, recovery is incremental—using small M we can recover the largest components of \mathbf{s} that M corresponds to, and the error in the recovered components is related to the total magnitude of unrecovered components. If we wish to recover more components or reduce errors, we can increase M accordingly.

However, it is important to note the following observation: fixing M by no means fixes the decoding accuracy, because by finding another basis that can better pronounce sparsity of \mathbf{x} and using such basis for decoding also results in improving the accuracy. This reasoning has guided the overall approach of this paper, namely, its focus on the use of the optimal KLT basis in decoding.

B. Karhunen-Loève Transform (KLT)

Consider \mathbf{x} , an $N \times 1$ vector of complex-valued, wide-sense stationary signal with zero mean (*e.g.*, sinusoids) data, and its

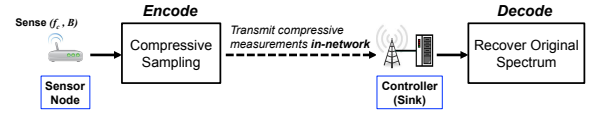


Fig. 1. In this simple model, a sensor node (*e.g.*, wireless user devices distributed across the network) performs compressive sampling on a given spectrum (f_c, B) and transmits the compressive measurements in-network. Controller acts as the data sink collecting compressive measurements from the sensor and recovers the original spectrum by compressive sensing decoding

autocorrelation matrix \mathbf{R}_x :

$$\mathbf{R}_x = \mathbb{E}[\mathbf{x}\mathbf{x}^H] \quad (2)$$

($\mathbb{E}[\cdot]$ is the expectation operator, and the superscript \mathbf{H} denotes the Hermitian transpose, *i.e.*, $\mathbf{x}^H = \mathbf{x}^{*T}$.) This means

$$\mathbf{R}_x = \mathbb{E} \left[\begin{array}{c} \left(\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_N \end{array} \right) \left(\begin{array}{cccc} x_1^* & x_2^* & \dots & x_N^* \end{array} \right) \end{array} \right] \quad (3)$$

Note that \mathbf{R}_x is the second-order statistics of \mathbf{x} . Since \mathbf{R}_x is a symmetric matrix, by eigenvalue decomposition we can write:

$$\mathbf{R}_x = \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^H \quad (4)$$

where \mathbf{Q} is a matrix containing the eigenvectors of \mathbf{R}_x , and $\mathbf{\Lambda}$ the diagonal matrix of eigenvalues. Note also that \mathbf{Q} is orthogonal, thus $\mathbf{Q}^{-1} = \mathbf{Q}^H$. Consider:

$$\mathbf{s} = \mathbf{Q}^H \mathbf{x} \quad (5)$$

We can find: $\mathbf{R}_s = \mathbb{E}[\mathbf{s}\mathbf{s}^H] = \mathbb{E}[\mathbf{Q}^H \mathbf{x} \mathbf{x}^H \mathbf{Q}] = \mathbf{Q}^H \mathbf{R}_x \mathbf{Q} = \mathbf{\Lambda}$. This signifies that \mathbf{s} is an uncorrelated representation of \mathbf{x} such that there is no statistical redundancy in \mathbf{s} , *i.e.*, it is the most compact representation possible for \mathbf{x} . This process is known as the *Karhunen-Loève Transform* (KLT), and we call \mathbf{Q} the KLT basis or matrix. Throughout the rest of this paper, we will use the term “KLT basis” and “KLT matrix” interchangeably.

III. MOTIVATING APPLICATION—SPECTRUM SENSING AND RECOVERY VIA COMPRESSIVE SENSING

We present a compressive sensing framework for spectrum sensing and recovery and illustrate the gain of using KLT in decoding.

A. Overview of Application System

As depicted in Fig. 1, our spectrum sensing example leverages wireless user nodes distributed across the network (*i.e.*, spatially varied). For simplicity, we focus on their role as spectrum sensors performing compressive measurements and making *in-network* transmission to *Controller* or the system’s data sink as denoted in Fig. 1. That is, the nodes form random linear combinations of sampled data, and the resulting compressive measurements are sent to Controller for the recovery of the original spectrum via compressive sensing decoding. We explain the details of each step as follows.

1) *Compressive sensing encoding*: For sampling in the time domain, we assume conversion of the analog RF/IF into the complex digital in-phase and quadrature (I-Q) samples. Note that we use the notation (f_c, B) to specify a spectrum sensing assignment such that a node should tune to f_c , the center frequency of the spectrum, and configure the sampling rate according to the assigned spectrum bandwidth B . These can be achieved by *coherent detection* [7] or by *direct bandpass sampling* [7]. Let each sampling instance yield $\mathbf{x} \in \mathbb{C}^N$ containing N I-Q samples from (f_c, B) . Finally, the encoding uses an $M \times N$ matrix Φ with $M \ll N$, measuring M random linear combinations of \mathbf{x} :

$$\mathbf{y} = \Phi \mathbf{x} \quad (6)$$

2) *Transmitting compressive measurements*: In-network transmissions occur over-the-air using bandwidth-limited channels, thus should be done efficiently. The Nyquist sampling of the (f_c, B) spectrum incurs $2 \times B$ samples per second. Since a time-domain sample is represented using a complex float data type, it may be as small as 20 compressed bits (*e.g.*, GNU Radio for USRP [8]) but could go up to 64 bits. This means that the Nyquist sampling of a 1-MHz spectrum alone can generate $2 \times 20 = 40$ Mbps to $2 \times 32 = 64$ Mbps data rate! Compressive sensing can reduce the data rate by a factor of $\frac{N}{M}$, with $M = cK \log \frac{N}{K}$ where K is the sparsity of the sensed spectrum.

3) *Compressive sensing decoding*: Ψ , the transform or basis used for decoding will determine the sparsity of \mathbf{x} to be exploited. Decoding should be successful with a relatively small number of measurements as long as we can find a suitable basis, Ψ . Thus, we can *optimistically* treat the compressive sensing decoding by choosing an optimal Ψ and defer the finding of such basis that gives the minimum sparsity by the decoding time.

B. Decoding with Fourier Basis

Typically, the carrier-modulated RF signals show no sparsity in the time domain, thus it is customary to look for their frequency domain representation by taking the Fourier Transform. There are numerous variants of the Fourier Transform that can be used in compressive sensing decoding, but since we operate on discrete data sampled from the analog domain, we choose the Discrete Fourier Transform (DFT). The DFT takes N samples in the time domain and produce their frequency domain representation in N frequency components. Given the time-sampled input vector \mathbf{x} of size N , the DFT computes inner products of \mathbf{x} with sampled complex sinusoids; writing in matrix form,

$$\mathbf{X} = \mathcal{F} \mathbf{x} \quad (7)$$

with matrix coefficients $f_{nk} = e^{-j2\pi nk/N}$ for $n, k = 0, 1, \dots, N-1$. The matrix \mathcal{F} is *invertible*, so we also have

$$\mathbf{x} = \mathcal{F}^{-1} \mathbf{X} \quad (8)$$

Using (8), we can rewrite (6) to be used in the compressive decoding:

$$\mathbf{y} = \Phi (\mathcal{F}^{-1} \mathbf{X}) = (\Phi \mathcal{F}^{-1}) \mathbf{X} \quad (9)$$

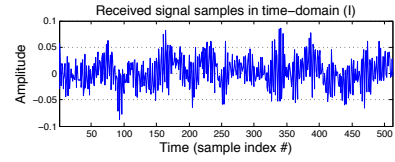


Fig. 2. Actual example of the time-domain representation of a wireless spectrum (source: 512 I-Q samples from our measurement on UHF Channel 21 at $f_c = 515$ MHz with 6-MHz bandwidth in Cambridge, MA on Saturday May 21, 2011 11:03pm EST using USRP2)

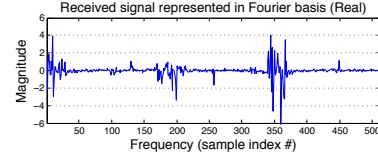


Fig. 3. The samples from Fig. 2 represented using the Fourier basis

So, to recover the original spectrum from the compressive measurements, we input the matrix product $\Phi \mathcal{F}^{-1}$ to the ℓ_1 -min decoder to solve for \mathbf{X} , then recover \mathbf{x} from \mathbf{X} . Decoding with the Fourier basis forms the baseline approach against which our proposed KLT-based approach will compare.

C. How Good Is Fourier Basis?

Fig. 2 illustrates the samples acquired from a real spectrum. We have sampled a 6-MHz UHF channel at its Nyquist rate using GNU Radio on USRP2 [8], [9] and plotted the I-Q amplitude in the time domain. Evidently, from the plots we cannot conclude any sparsity. In Fig. 3, however, we depict the same samples in the frequency domain after taking the FFT. The Fourier basis unveils the sparse structure present in the samples.

We have not yet presented any empirical results on decoding with the Fourier basis, but the use of the Fourier basis seems plausible at least for this example. From Fig. 3, we observe that about 21% to 24% of the samples (*i.e.*, ≈ 110 out of 512 samples) can be said *nonzero* due to their relatively large magnitudes. (The sparsity for the UHF case actually varies over a long time, but is stable in second ranges. This is sufficiently long for our optimal basis update scheme, as supported by our performance results of Section VI.) One benefit about using the Fourier basis is that the DFT matrix is fixed, independent of signals present in a spectrum. But can we do better than Fourier?

D. Decoding with Optimal KLT Basis

Before talking about finding any optimal basis for decoding, we point out the drawback of using a fixed basis such as Fourier. Since not all spectrums are expected to have approximately 20% nonzero elements in their frequency domain, the decoding accuracy would vary for a given number of compressive measurements (M). In other words, sparsity can be quite different for one spectrum to another. As a consequence, we may have to take some M_1 measurements from the spectrum (f_1, B_1) while we need to take M_7 measurements from (f_7, B_7) . For the worst, some spectrums are not at all

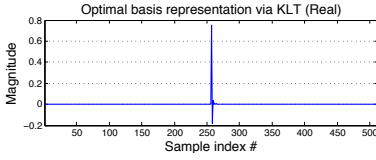


Fig. 4. The samples from Fig. 2 represented using the optimal basis found via KLT

sparse in the frequency domain (*e.g.*, spectrum with broadband multi-carrier transmissions), which would make decoding with Fourier basis fail if M does not increase appropriately. Adapting M would require prior knowledge about the signal under sensing, but our spectrum sensing task is supposed to find this knowledge.

This motivates us to use the KLT, as it can adapt to signals by using optimal bases for them. To illustrate the advantage of the KLT-based approach, we apply the KLT on the same data from Fig. 2 to find the optimal basis \mathbf{Q} , the KLT matrix. This yields a new representation, $\mathbf{s} = \mathbf{Q}^H \mathbf{x}$. In compressive sensing decoding, we will use $\mathbf{\Psi} = \mathbf{Q}$ as in (1) opposed to $\mathbf{\Psi} = \mathcal{F}$ when the Fourier basis is used.

We plot \mathbf{s} in Fig. 4. The result is astonishing—we can represent the signal sampled over the 6-MHz UHF channel in the KLT basis using just a few significant coefficients. This corresponds to the gain of about 50 in sparsity compared to the Fourier representation discussed earlier. As a result, the required measurements to achieve the same accuracy in the recovered solution of compressive sensing will be drastically reduced by a similar factor.

We stress that our KLT example here is idealistic, for we had the deterministic knowledge on the second-order statistics of the samples. In Section IV, we will present the key technical contribution of this paper introducing our method to estimate the KLT matrix directly from the measurements of the original signal via compressive sensing.

E. Sparsifying a Vector via Its KLT Matrix

Before closing Section III, we give a mathematical proof of the remarkable sparsity exhibited in Fig. 4. Theorem 1 shows that, for deterministic \mathbf{x} , $\mathbf{Q}^H \mathbf{x}$ is a sparse vector with only one nonzero component. That is, a vector can be made sparse by its KLT matrix. Since deterministic \mathbf{x} approximates a stationary signal over a short time interval, this result provides the insight into why the use of the KLT basis in compressive sensing decoding can promote sparsity and therefore reduce the required measurements.

Theorem 1: Suppose that \mathbf{x} is deterministic, *i.e.*, \mathbf{x} is a nonzero vector of constant components. Let \mathbf{Q} be the KLT matrix of \mathbf{x} , that is, $\mathbf{x}\mathbf{x}^H = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H$ where $\mathbf{\Lambda}$ is diagonal and \mathbf{Q} is orthogonal. Then, the vector $\mathbf{Q}^H \mathbf{x}$ has an exactly one nonzero component.

Proof: Since $\mathbf{x}\mathbf{x}^H$ is a symmetric matrix, we can write:

$$\mathbf{x}\mathbf{x}^H = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H \quad (10)$$

where $\mathbf{\Lambda}$ is diagonal and \mathbf{Q} is orthogonal. Note that $\mathbf{x}\mathbf{x}^H$ is a rank-1 matrix, so is $\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H$. This means that $\mathbf{\Lambda}$ has only one

nonzero entry on the diagonal. Without loss of generality we assume that the nonzero entry is the first entry on the diagonal, and denote it λ_1 .

Multiplying the two sides of (10) by \mathbf{Q}^H on the left and \mathbf{x} on the right, we get:

$$\mathbf{Q}^H \mathbf{x}\mathbf{x}^H \mathbf{x} = \mathbf{Q}^H \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H \mathbf{x} \quad (11)$$

This means:

$$u\mathbf{Q}^H \mathbf{x} = \mathbf{\Lambda}\mathbf{Q}^H \mathbf{x} \quad (12)$$

with $u = \|\mathbf{x}\|^2$. Given the property of $\mathbf{\Lambda}$ described above, the right-hand side is a vector that has only one nonzero component. It is the first component that is nonzero. In addition, we note that $\lambda_1 = u$. ■

IV. COMPUTING KLT BASIS WITH COMPRESSIVE SENSING

So, how does compressive sensing help compute the KLT matrix? To answer this question, we need to revisit the compressive measurement operation. Recall from (4) that the KLT matrix \mathbf{Q} is computed from the original data \mathbf{x} 's autocorrelation matrix \mathbf{R}_x . Actually, the same compressive measurements, after applied to a function, encode the autocorrelation as well. Consider:

$$\mathbf{R}_y = \mathbb{E}[\mathbf{y}\mathbf{y}^H] \quad (13)$$

By (6), we note that

$$\mathbf{R}_y = \mathbb{E}[\mathbf{\Phi}\mathbf{x}\mathbf{x}^H\mathbf{\Phi}^T] = \mathbf{\Phi}\mathbb{E}[\mathbf{x}\mathbf{x}^H]\mathbf{\Phi}^T \quad (14)$$

So, $\mathbf{R}_y = \mathbf{\Phi}\mathbf{R}_x\mathbf{\Phi}^T$. Since $\mathbf{\Phi}$ is not a square matrix, multiplying each side by $(\mathbf{\Phi}^T)^\dagger$, which is the pseudo-inverse of $\mathbf{\Phi}^T$, we obtain:

$$\mathbf{R}_y(\mathbf{\Phi}^T)^\dagger = \mathbf{\Phi}\mathbf{R}_x \quad (15)$$

Here, we just discovered that we have been encoding the autocorrelation matrix \mathbf{R}_x in $\mathbf{R}_y(\mathbf{\Phi}^T)^\dagger$, which is captured by the same compressive measurements, $\mathbf{y} = \mathbf{\Phi}\mathbf{x}$, that we used to encode the original data \mathbf{x} !

Following the above reasoning, we now describe a procedure to completely update the KLT basis. Computing an updated KLT matrix \mathbf{Q} requires the recovery of \mathbf{R}_x based on (15). The procedure comprises four steps:

- 1) Decode \mathbf{X} from $\mathbf{y} = (\mathbf{\Phi}\mathcal{F}^{-1})\mathbf{X}$ where \mathcal{F}^{-1} is the inverse DFT matrix as in (9);
- 2) Recover \mathbf{x} by computing $\mathbf{x} = \mathcal{F}^{-1}\mathbf{X}$;
- 3) Obtain \mathbf{Q} by performing the eigenvalue decomposition $\mathbf{R}_x = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H$ where $\mathbf{R}_x = \mathbb{E}[\mathbf{x}\mathbf{x}^H]$.

We note that Step 1) employs the Fourier basis rather than an optimal KLT basis. As discussed in Sections III.C and III.D, the Fourier basis reveals certain sparsity of the signal, but yields more significant nonzero coefficients than KLT. We cannot use an updated KLT basis in Step 1), as it is the information sought by this procedure. This means that, for Step 1), we will need to devote more measurements in decoding \mathbf{X} due to its use of the Fourier basis rather than an optimal KLT basis.

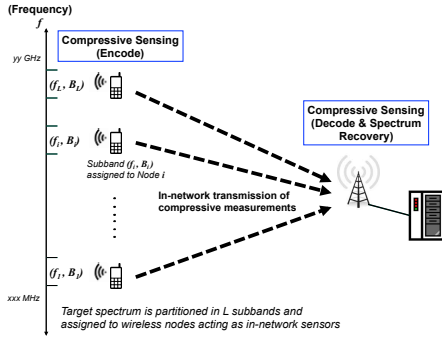


Fig. 5. System model incorporates sampling, compressive measurements, and transmission from distributed wireless nodes acting as in-network sensors

In addition, we note that, for Step 4), $O(N^3)$ computation is normally required for the eigenvalue decomposition. Fortunately, as will be discussed in Sections V.D and V.E, we only need to perform the procedure of complete KLT update infrequently.

V. COMPRESSIVE SENSING SYSTEM WITH OPTIMAL BASIS RECOVERY VIA KLT

We briefly describe our compressive sensing system that uses the KLT to compute optimal basis for decoding. We also present three algorithms to estimate and update the KLT basis online.

A. System Overview

Fig. 5 describes a compressive sensing system incorporating L distributed sensors transmitting their compressive spectrum measurements in-network. The objective of the system is to build fine-grained spectral information from distributed, *compressive* measurements on a wideband spectrum. The system partitions the target spectrum into L subbands and dispatches an assignment, denoted (f_i, B_i) for sensor i that should tune to f_i and start its measurement according to the channel bandwidth B_i . Note the total sensing bandwidth for the system, $B_{total} = \sum_{i=1}^L B_i$.

Each sensor node performs its task periodically and asynchronously without knowledge on how other sensors operate. The sensors act as compressive sensing encoders and transmit the encoded data to a system base station. The system collects the data and performs compressive sensing decoding to recover the spectral data encoded in the compressive measurements. Assume that at the system back-end there is computing infrastructure that can run the ℓ_1 -min decoders and other computing tasks to make sound recovery of the entire spectrum.

B. Workflow and Subsystems

Fig. 6 outlines workflow of the system. First of all, we leverage the low-complexity encoding of compressive sensing—our approach implements it faithfully without any modification. The novelty is in the decoding side. We add a new component that estimates the KLT basis from the compressive measurements \mathbf{y} as covered in Section IV. This additional component is labeled *CS-KLT* (Compressive Sensing-KLT) in the figure. CS-KLT is essentially a compressive sensing

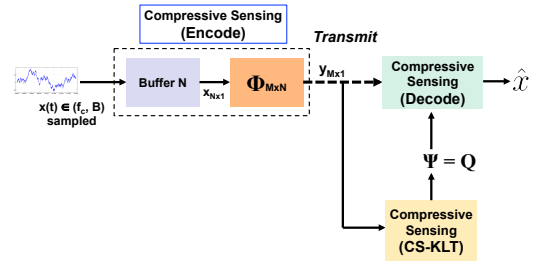


Fig. 6. Our approach implements an *unmodified* compressive sensing encoder, but features an extra block in the decoder, namely CS-KLT that estimates the KLT basis \mathbf{Q} directly from the compressive measurements and is used as the basis of decoding to recover the original data \mathbf{x}

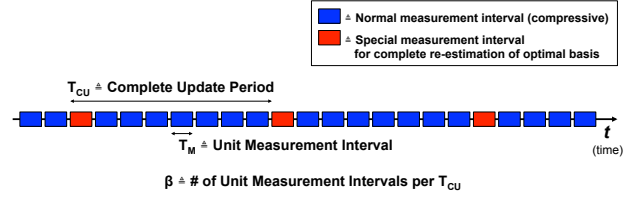


Fig. 7. Measurements for compressive sensing encoding occur in intervals that are further classified to two different categories

decoder that recovers an estimate of \mathbf{R}_x , the autocorrelation matrix of the original data \mathbf{x} , *directly* from \mathbf{y} , the compressive measurements on \mathbf{x} . We run the KLT on this estimate to obtain the *data-dependent*, dynamically changing optimal basis \mathbf{Q} . In recovering \mathbf{x} , the decoder takes $\Psi = \mathbf{Q}$ where Ψ is the transform described in Section II.A.

It is important to distinguish our approach from the baseline approach of Section III. The baseline approach does not require any block related to the KLT, and its compressive sensing decoder works with the fixed Fourier basis independent of data, *i.e.*, $\Psi = \mathcal{F}$.

C. Summary of Our Approach

We summarize our approach by comparing a decoder using the Fourier basis \mathcal{F} and that using a KLT basis \mathbf{Q} . Given a typical signal \mathbf{x} such as the signal received from a UHF channel, $\mathcal{F}\mathbf{x}$ is expected to be sparse. In contrast, if \mathbf{Q} is the KLT basis of \mathbf{x} , then $\mathbf{Q}\mathbf{x}$ is guaranteed to be highly sparse for any \mathbf{x} (as shown by Theorem 1). Decoding with the KLT basis in general can use fewer measurements than that with the Fourier basis to achieve the same recovery accuracy, and more importantly it can work with any signal. However, unlike the Fourier basis, the KLT basis is data-dependent. This means that the latter needs to be updated from time to time if the signal is non-stationary. To reduce required measurements in recovering an updated KLT basis, we shall still exploit whatever sparsity the signal may possess in some domains. For this purpose, a data-independent basis such as the Fourier basis, must be used, given that any previously computed data-dependent basis must have been obsoleted. This implied a larger number of required measurements. Fortunately, we can update the optimal basis relatively infrequently, that is, we only do this after the signal has changed substantially.

D. Online Algorithms to Estimate Optimal Basis

To formulate an algorithm that works on the fly, we first explain the low-level details of the measurement process. Sensor nodes sample their assigned spectrum in intervals. In Fig. 7, we depict the measurement intervals. *Unit measurement interval* denoted as T_M in the figure spans the time to perform a single instance of compressive sensing encoding. T_M is a configurable parameter. Note that compressive sensing encoding may not be the only task for a sensor node; in other words, it can also have other networking and computing tasks for its own. Thus, T_M is more of a *duty cycle* during some portion of which the sensor devotes to perform the compressive sensing task for the system. In fact, T_M could be substantially longer than the time for acquiring required number of samples to compress.

Given a discrete sequence of data, computing the autocorrelation matrix is relatively a straightforward task. We can align this computation with the duty-cycled measurements of Fig. 7, add the new matrix to the sum of all the past matrices, and take the expectation. This is a simple, numerical method that estimates \mathbf{R}_x online. We formulate two basic methods based on simple averaging schemes and introduce an advanced method that can directly recover the KLT basis within our compressive sensing framework.

1) *Finite averaging*: This is a simple method that keeps a finite number of most recent historical $\mathbf{x}\mathbf{x}^H$ values (say, depth- n) and computes their average. The matrix computed using the new block of data obsoletes the oldest.

2) *IIR-style weighted averaging*: We can compute the correlation matrix at k -th interval using $\mathbf{R}_x^{(k)} = \alpha(\mathbf{x}\mathbf{x}^H) + (1 - \alpha)\mathbf{R}_x^{(k-1)}$ with $0 < \alpha < 1$. We note that the weight α near 1 expresses the estimation bias toward the most recently acquired sample.

3) *Periodic updates with CS-KLT algorithm*: We present our pseudo-code implementation in Algorithm 1, which features 4 subroutines. Referring back to Fig. 7, we note that there are two types of measurement intervals: normal compressive measurement interval (marked in *blue*); and the special interval (in *red*) that the global variable *completeUpdate* is set `true`, and **procedure** CSKLT calls **procedure** CSDECODERX to completely reestimate \mathbf{R}_x . This reestimate is necessary because in reality wireless channels are time-varying (and stationarity does not hold up). It works on lower compression as noted in Section IV. As illustrated in Fig. 7, we schedule the complete update once every β unit measurement intervals spanning T_{CU} . On a normal measurement interval, CSKLT invokes **procedure** DOINCUPDATERX to update \mathbf{R}_x incrementally. The incremental update is based on weighted averaging between the two successive intervals as implemented in lines 20–22 of Algorithm 1.

E. Optimal Interval to Update KLT Basis

Recovering \mathbf{R} costs more measurements than recovering the input signal. So we should not do complete update too frequently. But incremental update will incur increased measurements per time step due to use of un-updated KLT basis.

Algorithm 1 The CS-KLT Online Optimal Basis Estimation

```

1: procedure CSKLT(curRx,  $\mathbf{y}$ ,  $\alpha$ ,  $\Phi_{CU}$ ,  $\Phi_{normal}$ ,  $\mathbf{Q}$ )
2:   if completeUpdate == True then
3:     newRx = CSDECODERX( $\mathbf{y}$ ,  $\Phi_{CU}$ ,  $\mathcal{F}^{-1}$ )
4:   else
5:     newRx =
6:       DOINCUPDATERX(curRx,  $\mathbf{y}$ ,  $\alpha$ ,  $\Phi_{normal}$ ,  $\mathbf{Q}$ )
7:   end if
8:   newQ = EIG(newRx)
9:   return newQ
10: end procedure
11: procedure CSDECODERX( $\mathbf{y}$ ,  $\Phi$ ,  $\Psi^{-1}$ )
12:    $a \leftarrow$  MTXMULTIPLY( $\Phi$ ,  $\Psi^{-1}$ )
13:    $b \leftarrow \ell_1$ -MINSOLVE( $\mathbf{y}$ ,  $a$ )
14:   return  $b$ 
15: end procedure
16: procedure DOINCUPDATERX( $\mathbf{R}_x$ ,  $\mathbf{y}$ ,  $\alpha$ ,  $\Phi$ ,  $\Psi$ )
17:    $x \leftarrow \ell_1$ -MINSOLVE( $\mathbf{y}$ ,  $\Phi$ ,  $\Psi$ )
18:    $a \leftarrow$  CMLXCONJ( $x$ )
19:    $b \leftarrow$  VTRTRANPOSE( $a$ )
20:    $c \leftarrow$  VTRMULTIPLY( $x$ ,  $b$ )
21:    $d \leftarrow \alpha \times c$ 
22:    $e \leftarrow (1 - \alpha) \times \mathbf{R}_x$ 
23:    $f \leftarrow d + e$ 
24:   return  $f$ 
25: end procedure
26: procedure CSENCODECH( $f$ ,  $bw$ ,  $\Phi$ )
27:   ( $M$ ,  $N$ )  $\leftarrow$  SIZEOF( $\Phi$ )
28:   TUNERF( $f$ )
29:    $t_s \leftarrow \frac{1}{2 \cdot bw}$ 
30:    $a \leftarrow$  BANDPASSSAMPLE( $t_s$ )
31:    $b \leftarrow$  MTXMULTIPLY( $\Phi$ ,  $a$ )
32:   return  $b$ 
33: end procedure

```

Thus there is a tradeoff in determining the length of the interval between two consecutive complete updates. The choice will depend on the expected stability of the channel and relative measurement costs between complete and incremental updates. In any case, there is an optimal KLT basis update interval.

VI. EVALUATION

We evaluate the performance of our KLT-based compressive sensing in two spectrum sensing applications: (1) sensing 200-MHz UHF whitespace spectrum; (2) detecting/identifying the presence of one or more radio transmitters. We obtained all of the data used in the empirical evaluation of this paper in field and laboratory experiments; we collected them using commercially available software-defined radio products.

A. Experiments

1) *Sensing UHF spectrum*: We prepared 4 USRP2 radios [9] in an indoor lab for compressive sensing of the 200-MHz UHF spectrum partitioned in $L = 8$ subbands (*i.e.*, each with 25 MHz bandwidth) with center frequencies, $f_c \in \{512.5, 537.5, 562.5, 587.5, 612.5, 637.5, 662.5, 687.5\}$ (MHz). Note that our wideband spectrum spans about 33 UHF channels (6-MHz each) between Ch. 19 and Ch. 51. We configured each USRP2 sensor responsible for measuring two subbands. We used $T_M = 1$ msec for the Unit Measurement Interval. The Nyquist sampling rate of each subband is: $f_s =$

25 MHz \times 2 = 50 million samples/sec. Based on this rate, it is possible to obtain up to 50,000 samples for a single measurement interval (*i.e.*, during $T_M = 1$ msec), and we required each sensor acquire at least 2,048 samples during T_M . We ran a number of experimental sessions over different days at various hours, with each session lasting up to 5 seconds consisting of thousands of consecutive measurement intervals.

2) *Detecting radio transmitters*: This is our field experiment conducted in an outdoor wireless test field. We again used 4 USRP2 radios as sensors for compressive sensing of approximately 100-MHz spectrum partitioned in $L = 4$ subbands of 25-MHz bandwidth each. The center frequencies for this experiment were 363, 386, 409, and 432 MHz, meaning there is 2-MHz overlap between any two adjacent sensors. We placed 4 Harris walkie-talkie transmitters [10] operating in the first, second, and fourth subbands. Two of them were engaged in live voice communications over the first subband while the other two were turned on and off randomly operated by us manually each on one of the last two subbands. All sensors worked in parallel on their subbands without collaborating with one another. We assumed the same measurement parameters as in Experiment 1) but used longer sessions that lasted over 10 seconds.

B. Recovery of Wideband Spectrums under Periodically Updated KLT Basis

1) *UHF spectrum*: We implemented all three online algorithms described in Section V.D that estimate \mathbf{R}_x and compute a KLT matrix used for compressive sensing decoding. We also evaluated the baseline approach relying on the Fourier basis and compared the decoding errors against our approach. The error metric we adopted is the L2-norm of the difference between the original, uncompressed samples and the recovered signal: $e_{\ell_2} = \|\mathbf{x} - \hat{\mathbf{x}}\|_{\ell_2}$. Fig. 11 presents the error performances. As for our periodic updating algorithm, we used $\beta = 10$ and $N = 128 \times 8 = 1024$, tested it with and without incremental updates under various compression ratios (M/N). We first observe the clear performance gains on using the optimal basis via KLT. Given a fixed error budget, our approach can result in the compression gain of 2 to 4.3, that is, we can approximately quadruple the rate of transmitted information in the best scenarios. For illustrative purposes, we plot the original frequency response and the recovered frequency responses under compression ratios ($\frac{M}{N}$, thus smaller the higher the compression gain) from .1 to .8 for using the Fourier basis in Fig. 9 and for using the KLT basis (found from our best online algorithm, *i.e.*, periodic and incremental updates) in Fig. 10.

2) *Harris radio detection*: Again, we observe clear benefits of using the optimal basis for the Harris radio experiment. We similarly illustrate the original and recovered frequency responses in Figs. 12 and 13. Note that the combined frequency response of 4 relatively narrowband pulses resulted from Harris radios is much sparser than the UHF spectrum, which makes the detection problem easier for the baseline approach with the Fourier basis. The error performance on detecting Harris radios is presented in Fig. 14 where we only

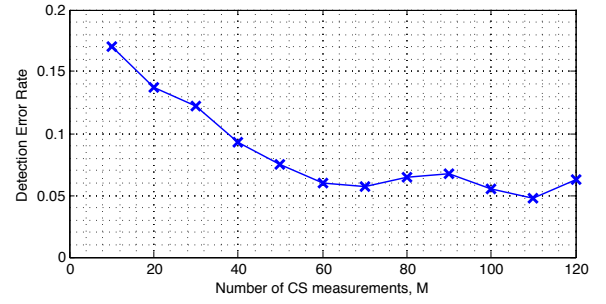


Fig. 8. Error performance of CS-KLT in a real-world experiment

depict the error curve for our best online algorithm against the baseline.

C. Dynamic Identification of Subsets of Pre-trained KLT Patterns

In this Section we examine a preliminary application of detecting the ON/OFF state of combinations of multiple transmitters using previously trained KLT matrices. For this purpose, we reused the experimental data collected from the 4 Harris walkie-talkies, as described in the previous section. Suppose we wish to detect which of the targets is in the ON state; to do this, we propose the following target detection strategy based on CS-KLT.

First, we perform a training phase, where we precompute the KLT bases for individual walkie-talkie radios by activating them one at a time. Let us call the 4 KLT matrices obtained in our field experiment $\mathbf{Q}_1, \dots, \mathbf{Q}_4$. There may be other ways to obtain the KLT characterizations of the targets requiring less access; for example, through training on reference signals, or from known signal models, such as the power-spectral densities of specific modulation schemes.

Secondly, we perform a separate CS-KLT decoding operation using each KLT matrix \mathbf{Q}_i as the sparse transform. Each decoding operation using \mathbf{Q}_i detects the presence of target i , as desired; clearly, this decoding process can detect any combination of active transmitters. Furthermore, since the KLT bases of individual transmitters do not change, there is no need for periodic updating.

To compute statistics on the detection success performance, we took from each of the 4 USRP2 devices 4 seconds worth of spectrum samples, encompassing 400 blocks of 2048 samples each. We added the samples together into a single stream of measurements, and encoded them using a standard compressive sensing process. On the decoding side, we decoded the combined samples using the KLT matrix \mathbf{Q}_3 , and used a simple threshold to decide whether transmitter 3 was ON or OFF. As the performance metric, we computed the detection error rate—the fraction of the 400 CS-KLT instances in which the decoding decision was wrong. We present the results in Figure 8 for varying M . We can see that even for small M , the error rate is relatively small; this is encouraging, considering that the raw number of samples per measurement interval is close to 10^4 .

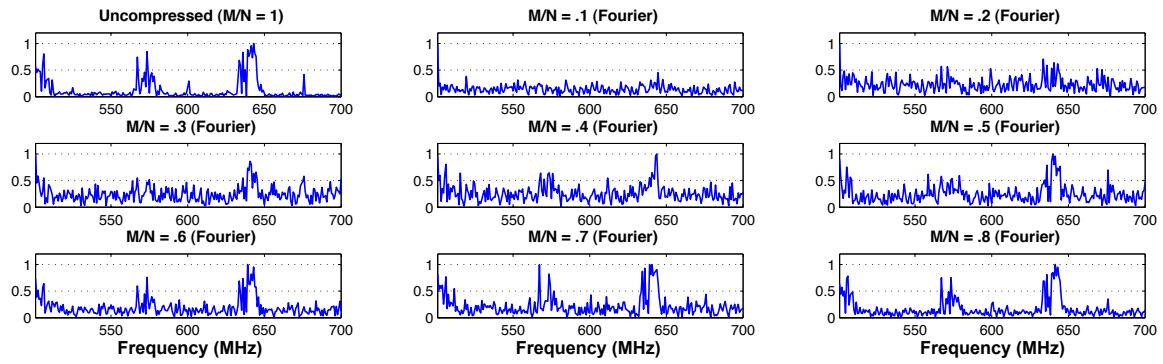


Fig. 9. Recovering 200-MHz UHF spectrum with the Fourier basis under various compression ratios ($\frac{M}{N}$)

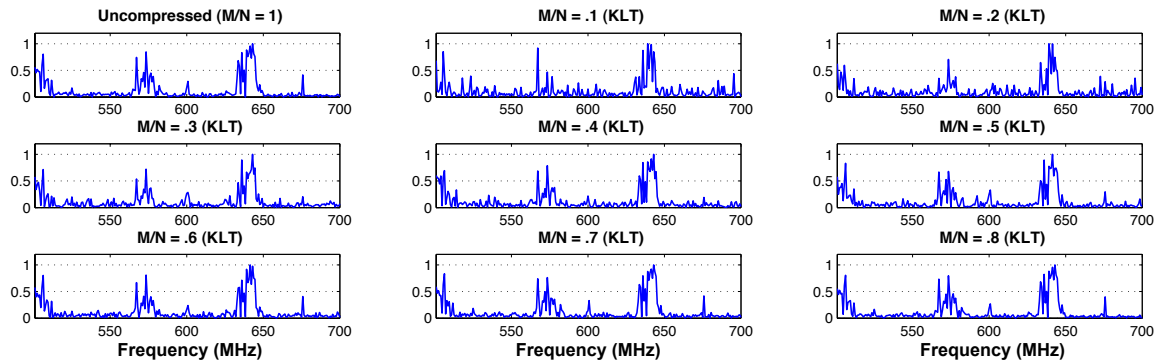


Fig. 10. Recovering 200-MHz UHF spectrum with the KLT basis under various compression ratios ($\frac{M}{N}$)

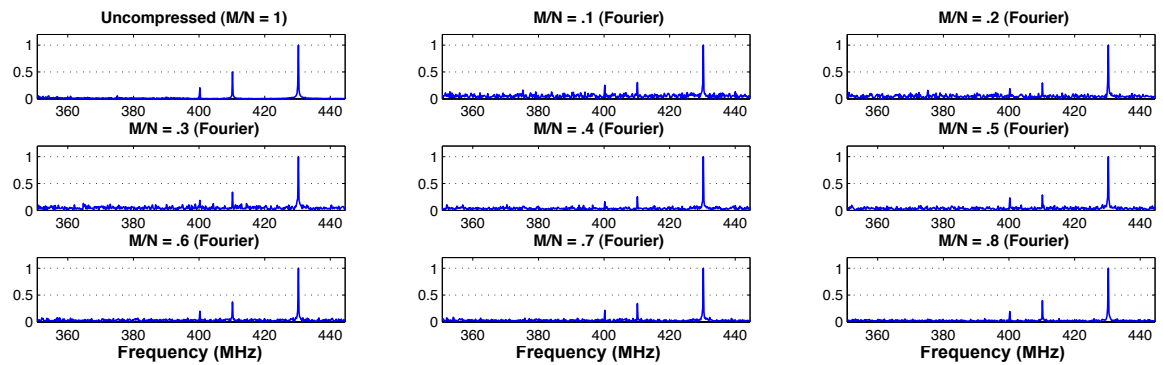


Fig. 12. Recovering 100-MHz spectrum featuring 4 Harris radios with the Fourier basis under various compression ratios ($\frac{M}{N}$)

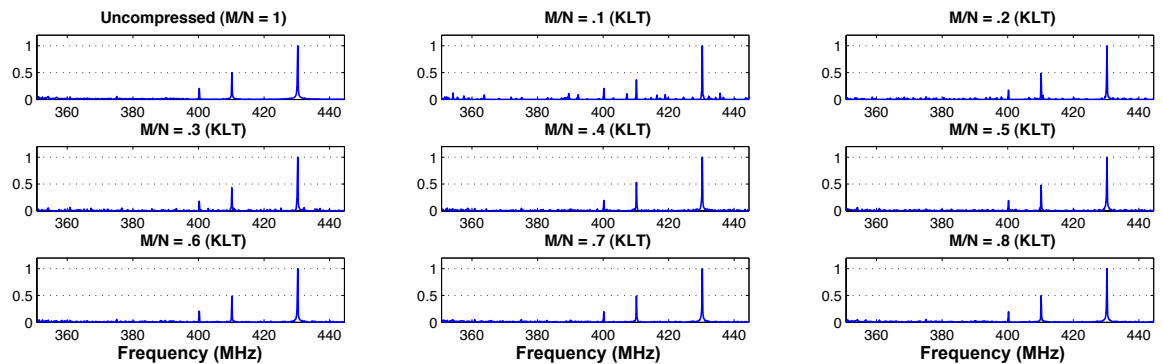


Fig. 13. Recovering 100-MHz spectrum featuring 4 Harris radios with the KLT basis under various compression ratios ($\frac{M}{N}$)

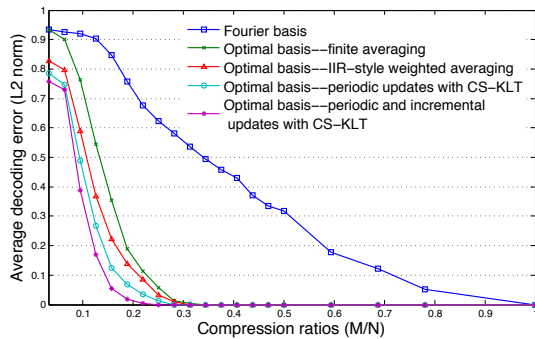


Fig. 11. Error performance on recovering 200-MHz UHF spectrum using 8 USRP2 sensors

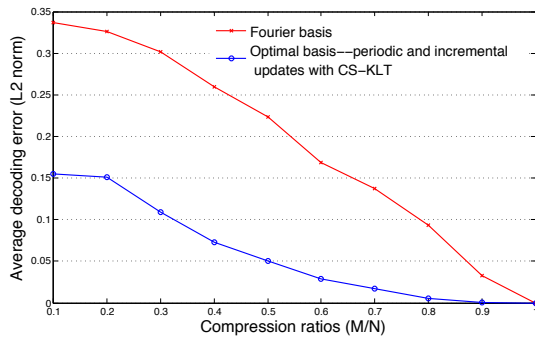


Fig. 14. Error performance on recovering 100-MHz spectrum featuring 4 Harris radios using 4 USRP2 sensors

VII. RELATED WORK

Considering general applicability of compressive sensing and KLT, we do not attempt to summarize all of its related work here. Instead, we focus on compressive data gathering and spectrum sensing. Luo *et al.* [11] presents a framework that integrates compressive sensing for gathering data in-network data for sensor networks. Haupt *et al.* [12] discusses early work in data aggregation using compressive sensing techniques over multi-hop wireless networks.

Cognitive radio systems attracted many research projects in spectrum sensing. Polo *et al.* [13] introduced an early spectrum sensing framework that features cognitive radios and uses distributed, compressive measurements from them. Wang *et al.* [14] extends this framework presenting a system more close to the baseline approach that we described in Sections III and V. Their decoding scheme relies on the combined use of analog-to-information conversion (AIC) hardware and the SOMP [15] algorithm. More interestingly, they discuss the recovery of wideband PSD using compressive sensing. Duarte *et al.* [15] leverages model-based joint sparsity of wireless signals for compressive sensing decoding.

VIII. CONCLUSION

In this paper we have described how compressive sensing can take advantage of optimal KLT basis. By using KLT in decoding we can drastically reduce the number of required measurements for achieving the same recovery accuracy. By using compressive sensing we can dynamically update

the KLT basis directly from measurements of the signal, with substantially fewer samples than the Nyquist rate. We have demonstrated the resulting performance gains in UHF spectrum sensing and radio transmitters detection, based on measurement data gathered in the field by software-defined radios.

The results of this paper suggest that we can make effective use of optimal KLT basis in compressive sensing decoding, in spite of the fact that KLT is signal dependent and needs to be updated from time to time. This represents a significant departure from the current practice in compressive sensing applications which almost always uses fixed transforms such as DCT for decoding.

While encouraged by the performance of KLT-based compressive sensing, we realize that much further work is needed. It would be useful to develop a comprehensive optimality theory on the efficiency limit of compressive sensing under decoding with adaptive optimal basis. In the present paper, for simplicity we decode the correlation matrix with the Fourier basis. More optimized bases could be used for this purpose. Lastly, we should explore a wide set of applications to leverage the new opportunity enabled by compressive sensing decoding with KLT-like optimal bases.

REFERENCES

- [1] D. Donoho, "Compressed Sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [2] H.-C. Chen, H. T. Kung, D. Vlah, D. Hague, M. Muccio, and B. Poland, "Collaborative Compressive Spectrum Sensing in a UAV Environment," in *MILCOM*, 2011.
- [3] H. T. Kung, C.-K. Lin, and D. Vlah, "CloudSense: Continuous Fine-grain Cloud Monitoring with Compressive Sensing," in *USENIX HotCloud*, 2011.
- [4] Y. L. Gwon, H. T. Kung, and D. Vlah, "DISTROY: Detecting Integrated Circuit Trojans with Compressive Measurements," in *USENIX HotSec (to appear)*, 2011.
- [5] K. Karhunen, "Über Lineare Methoden in der Wahrscheinlichkeitsrechnung (On Linear Methods in Probability and Statistics)," PhD Dissertation, University of Helsinki, 1947.
- [6] E. J. Candès, "Compressive Sampling," in *Proc. of the International Congress of Mathematicians*, 2006.
- [7] W. Waters and B. Jarrett, "Bandpass Signal Sampling and Coherent Detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 18, no. 6, pp. 731–736, Nov. 1982.
- [8] GNU Radio. [Online]. Available: <http://gnuradio.org>
- [9] The Universal Software Radio Peripheral (USRP) Products, Ettus Research LLC. [Online]. Available: <http://www.ettus.com/>
- [10] Harris Radio. [Online]. Available: <http://www.harris.com/>
- [11] C. Luo, F. Wu, J. Sun, and C. W. Chen, "Compressive data gathering for large-scale wireless sensor networks," in *ACM MobiCom*, 2009.
- [12] J. Haupt, W. Bajwa, M. Rabbat, and R. Nowak, "Compressed Sensing for Networked Data," *Signal Processing Magazine*, vol. 25, 2008.
- [13] Y. Polo, Y. Wang, A. Pandharipande, and G. Leus, "Compressive Wideband Spectrum Sensing," in *IEEE ICASSP*, 2009.
- [14] Y. Wang, A. Pandharipande, Y. Polo, and G. Leus, "Distributed Compressive Wide-band Spectrum Sensing," in *Information Theory and Applications Workshop*, 2009.
- [15] M. Duarte, S. Sarvotham, D. Baron, M. Wakin, and R. Baraniuk, "Distributed Compressed Sensing of Jointly Sparse Signals," in *Signals, Systems and Computers*, 2005.