Letter to the Editor

Chirp sensing codes: Deterministic compressed sensing measurements for fast recovery

Lorne Applebaum\textsuperscript{a}, Stephen D. Howard\textsuperscript{b}, Stephen Searle\textsuperscript{c}, Robert Calderbank\textsuperscript{a,}\textsuperscript{*}

\textsuperscript{a} Electrical Engineering, Princeton University, B318 Engineering Quad, Princeton, NJ 08544, USA
\textsuperscript{b} Defence Science & Technology Organisation, PO Box 1500, Edinburgh 5111, Australia
\textsuperscript{c} Electronic & Electrical Engineering, The University of Melbourne, Victoria 3010, Australia

A R T I C L E   I N F O

Article history:
Available online 2 September 2008
Communicated by Charles K. Chui on 5 April 2008

Keywords:
Compressed sensing
Deterministic measurement matrices
Chirp detection

A B S T R A C T

Compressed sensing is a novel technique to acquire sparse signals with few measurements. Normally, compressed sensing uses random projections as measurements. Here we design deterministic measurements and an algorithm to accomplish signal recovery with computational efficiency. A measurement matrix is designed with chirp sequences forming the columns. Chirps are used since an efficient method using FFTs can recover the parameters of a small superposition. We show that this type of matrix is valid as compressed sensing measurements. This is done by bounding the eigenvalues of sub-matrices, as well as an empirical comparison with random projections. Further, by implementing our algorithm, simulations show successful recovery of signals with sparsity levels similar to those possible by matching pursuit with random measurements. For sufficiently sparse signals, our algorithm recovers the signal with computational complexity $O(K \log K)$ for $K$ measurements. This is a significant improvement over existing algorithms.

Crown Copyright © 2008 Published by Elsevier Inc. All rights reserved.

1. Introduction

The sparsity of signals is a fact often exploited in signal processing. In particular, the common way to compress a signal is to transform it to the basis in which it is sparse and subsequently store only the locations and values of the few non-zero elements. Recently, it has been discovered that, in addition to storage, signal sparsity can be leveraged to reduce the number of measurements for signal acquisition and detection. It has been shown that, if a signal is sufficiently sparse, a small number of projections onto random vectors is enough to recover the signal [2,6]. This method has been called Compressed Sensing.

In compressed sensing, the use of randomly generated projections to make measurements has the useful consequence of sidestepping the computationally difficult task of checking whether the measurements allow for signal recovery. By considering recovery stochastically, it has been shown that measurements generated from Gaussian or Bernoulli random variables allow for signal recovery with high probability. In some ways, the use of random measurements may be viewed as an analogy to random codes used by Shannon to prove theorems in channel coding. Though useful in proofs, purely random channel codes are never used in practice because encoding and decoding would be far too computationally intensive. Instead, practical channel codes are developed with an efficient coding and decoding scheme in mind. We have a similar situation in compressed sensing. Though $\ell_1$ minimization has been shown to recover the signal from random projections [2], it is computationally expensive. The question arises as to whether we can design projections to facilitate the rapid recovery of signals.

* Corresponding author.
E-mail address: calderbk@math.princeton.edu (R. Calderbank).
of the signal. This is an issue of practical consequence. If compressed sensing is to be used in real-time systems, we must have a method which, in addition to reducing the number of measurements, is able to recover the signal quickly. Here we present a proof of concept scheme which accomplishes this.

A number of decoding schemes have been proposed that improve upon the ℓ_1 minimization signal recovery technique (also known as basis pursuit). However, most schemes presuppose random measurements. Examples include orthogonal matching pursuit [13] and its refinements [8]. In contrast, the scheme presented here exploits structure in deterministically designed measurements to make recovery much faster. There exists a small number of other schemes with less structurally random measurements [5,11,12]. The scheme presented here has lower recovery complexity.

The remainder of the paper is organized as follows. In Section 2 we provide necessary background and notation and in Section 3 we introduce our encoding scheme and the corresponding decoding algorithm. Section 4 provides analysis of our encoding matrix in terms of restricted isometry properties commonly employed in compressed sensing. In Section 5 we consider our scheme in the special case of Fourier signals and present a modification to improve the scheme’s robustness. In Sections 6 and 7 we examine our algorithm in terms of computational complexity, signal recovery and robustness to noise.

2. Compressed sensing background and notation

We consider discrete signals of finite length. Let x be a length N signal which we would like to sense and recover. We assume that x is sparse in some orthonormal basis. Thus, we can write x as

\[ x = \Psi s \]  

where s is a length N vector with fewer than M non-zero elements. We measure x with K < N projections which have results given in the vector y. The vectors projected upon are set as the rows of the K × N matrix Φ which gives

\[ y = \Phi \Psi s = \Theta s \] 

where the second equality is by definition of Θ. We are free to design Φ and thus Θ. Though, if we design Θ we should remain aware that actual sensing of the signal is done with Φ.

Since Θ is a wide matrix, solving for s given y is ill posed. However, using non-linear methods, we can leverage the fact that s has at most M non-zero elements. It has been shown in [3] that if Θ satisfies certain restricted isometry properties (RIP), s can be recovered perfectly using an ℓ_1 minimization. An important example is randomly generated Θ. Several results exist showing that, when M satisfies

\[ M < cK / \log(N/K), \] (3)

with a known constant c, randomly generated matrices of various types satisfy RIP with high probability [1]. Thus, if a signal’s sparsity is bounded by (3), then it can be recovered from K random measurements with high probability.

We will consider our designed Θ more precisely in terms of RIP in Section 4. There, we will also give an empirical comparison of the eigenvalue statistics of our designed Θ with those of randomly generated measurements showing that (3) applies to recoverability from chirp sensing codes.

3. Chirped sensing codes

We approach the recovery problem by noting that finding s is equivalent to discovering which small linear combinations of the columns of Θ form y. We will design Θ to facilitate this. In particular, we will look at a Θ designed with chirp signals forming the columns.

A length K chirp signal has the form

\[ v_{m,r}(l) = \alpha \cdot e^{\frac{j2\pi ml}{K} + \frac{j2\pi rl}{K}}, \quad m, r \in \mathbb{Z}_K, \]  

where m is the base frequency and r is the chirp rate. For a length K signal, there are K^2 possible pairs (m, r). We will form a K × K^2 sized Θ which has columns filled with all K^2 uni-modular chirp signals (setting α = 1 for notational convenience, though in Section 4 α = \frac{1}{\sqrt{K}} is used).

Consider a vector y, indexed by l, formed from the linear combination of some chirp signals

\[ y(l) = s_1 e^{\frac{j2\pi m_1 l}{K} + \frac{j2\pi r_1 l}{K}} + s_2 e^{\frac{j2\pi m_2 l}{K} + \frac{j2\pi r_2 l}{K}} + \ldots \]  

which have base frequencies defined by m_l and chirp rates defined by r_l. The chirp rates can be recovered from y by looking at \( \tilde{y}(l)y(l + T) \), where the index l + T is taken mod K. This gives

\[ f(l) = \tilde{y}(l)y(l + T) = |s_1|^2 e^{\frac{j2\pi (m_1 T + r_1 T^2)}{K}} + |s_2|^2 e^{\frac{j2\pi (m_2 T + r_2 T^2)}{K}} + \ldots + \text{cross terms} \]  

where the cross terms are of the form

\[ s_p s_q e^{\frac{j2\pi}{K} (m_p T + r_p T^2)} e^{\frac{j2\pi}{K} (m_q - m_p + 2T r_p)} e^{\frac{j2\pi}{K} (r_p - r_q)} \]  

which gives

\[ s_p s_q e^{\frac{j2\pi}{K} (m_p T + r_p T^2)} e^{\frac{j2\pi}{K} (m_q - m_p + 2T r_p)} e^{\frac{j2\pi}{K} (r_p - r_q)} \]  

(7)
and are therefore chirps. We see that \( f(l) \) is a signal that has sinusoids at the discrete frequencies \( 2r_i T \mod K \). If \( K \) is prime, this is a bijection from chirp rates to FFT bins. Furthermore, the remainder of the signal consists of the cross terms. Since the cross terms are chirps, their energy is spread across all FFT bins.

As long as \( y \) consists of sufficiently few chirps (\( x \) is sparse), taking a FFT of \( f(l) \) results in a spectrum with significant peaks at locations corresponding to \( 2r_i T \mod K \) from which we can glean chirp rates.

Upon discovering the chirp rate \( r_i \) we can "dechirp" the signal \( y(l) \) by multiplying by \( e^{-j2\pi r_i l^2/K} \). This converts only the chirps with rate \( r_i \), to sinusoids. Performing an FFT on the resulting signal can be used to retrieve the corresponding value(s) for \( m_i \) and \( s_i \).

Setting the elements of \( \Theta \) as

\[
[\Theta]_{ij} = e^{j2\pi r_i l} e^{j2\pi m l} \quad \text{with} \quad k = Kr + m \in \mathbb{Z}_K^2 \tag{8}
\]

we see that \( y = \Theta s \) will have the form (5). Given \( y \) formed using \( \Theta \), we summarize the algorithm described.

1. Choose a \( T \in \mathbb{Z}_K \), \( T \neq 0 \), and a stopping energy \( \epsilon \).
2. Form \( f(l) = y(l) y(l + T) \) and take length \( K \) FFT.
3. Find location of the peak in the FFT as \( 2r_i T \mod K \) and record the unique \( r_i \) corresponding to the location.
4. Multiply \( y(l) \) by \( e^{-j2\pi r_i l^2/K} \) and take length \( K \) FFT.
5. Find the location of the peak and record as \( m_i \). Use the value of the peak to recover \( s_i \).
6. Replace \( y \) with \( y - s_i e^{j2\pi r_i l^2/K} \).
7. Repeat steps (2)-(6) until \( \| y \|_2^2 < \epsilon \) or have iterated \( M \) times.

The recovery of \( s_i \) gives the value of an element in \( s \) while the pair \((m_i, r_i)\) gives its location in \( s \) as the index \( Kr_i + m_i \).

4. RIP analysis of \( \Theta \)

Having described the structure of \( \Theta \) in (8), here we consider the matrix in terms of restricted isometry properties giving signal recovery guarantees. Like much of the compressed sensing literature, the properties of interest here are defined in terms of restricted isometry constants [3]. For \( \Theta \) scaled such that its columns have unit norm, they are defined as the smallest \( \delta_M \) such that

\[
(1 - \delta_M)\|x\|^2 \leq \|\Theta x\|^2 \leq (1 + \delta_M)\|x\|^2 \quad \forall x, \quad |\Gamma| = M \tag{9}
\]

where \( \Theta_\Gamma \) is the sub-matrix of \( \Theta \) using the \( M \) columns specified in the set \( \Gamma \). The constants \( \delta_M \) are equivalently bounds on the eigenvalues of \( \Theta_\Gamma^H \Theta_\Gamma \) close to unity. Bounds on the restricted isometry constants can give guarantees on the recoverability. For example, from [3] we have the following lemma.

**Lemma 1.** Suppose that \( M \geq 1 \) is such that \( \delta_{2M} < 1 \), and let \( s \) be a \( M \)-sparse signal with measurements \( y \). Then \( s \) is the unique minimizer to

\[
\min \| d \|_{t_0} \quad \text{s.t.} \quad \Theta d = y \tag{10}
\]

where the \( t_0 \)-norm is the count of non-zero elements.

Using arguments similar to the uniqueness arguments in [7], here we calculate a bound on the sparsity \( M \) for \( \delta_{2M} < 1 \) to hold for \( \Theta \).

**Theorem 1.** Suppose that \( M < (\sqrt{K} + 1)/2 \) and let \( s \) be a \( M \)-sparse signal with chirp code measurements \( y \). Then \( s \) is the unique minimizer of (10).

**Proof.** Let \( |\Gamma| = 2M \) and \( \lambda \) be an arbitrary eigenvalue of \( G = \Theta_\Gamma^H \Theta_\Gamma \). By the Gershgorin circle theorem,

\[
|\lambda - 1| \leq \sum_{i \neq j} |G_{i,j}| \quad \forall j. \tag{11}
\]

The values \( G_{i,j}, i \neq j \), are the inner products between non-identical normalized chirps. From the work in [4], for prime \( K \) these inner products satisfy \( |G_{i,j}| \leq 1/\sqrt{K} \). Thus,

\[
|\lambda - 1| \leq \frac{1}{\sqrt{K}} (2M - 1). \tag{12}
\]

Since this is true for all eigenvalues of \( \Theta_\Gamma^H \Theta_\Gamma \), the upper-bound above also applies to \( \delta_{2M} \). Applying Lemma 1 upon the bound gives the result. \( \square \)
More restrictive conditions than Lemma 1 exist for recovery by $\ell_1$ minimization. These can be approached in a similar fashion to Theorem 1. However, rather than examine this directly, we will now show that chirp measurements can perform as well as randomly generated measurements. To do this, we consider the restricted isometry properties in a stochastic manner.

The standard compressed sensing formulation described in the literature considers the signal $s$ as fixed and examines recovery using of random measurements. The probability of recovery is inherited from the randomness in the measurements. However, here we have developed deterministic measurements $\Theta$. This makes it difficult to find the values of $\delta_M$ for non-random $\Theta$ and, in general, requires the computationally difficult problem of checking all $\binom{N}{M}$ possible $\Gamma$. However, we conjecture that, considering the signal rather than the measurements stochastically, a randomly generated sparse $s$ can be recovered from $\Theta$ with high probability. Empirical results presented here support this claim.

Recall that $\Theta$ satisfies RIP when $\Theta^H \Theta \Gamma \Theta^T \Gamma$ has eigenvalues sufficiently close to 1 for appropriately sized $\Gamma$. These sets relate to the possible supports of an $M$-sparse $s$. Considering these sets randomly, we consider the probability that $\Theta^H \Theta \Gamma \Theta^T \Gamma$ has eigenvalues appropriately close to 1.

After scaling $\Theta$ so that its columns have unit norm, we compare the statistics of eigenvalues of its Gram matrices to those of a matrix with Gaussian entries of zero mean and variance $\frac{1}{\pi}$. From well-known compressed sensing results, this Gaussian matrix is known to satisfy RIP with high probability when (3) is satisfied [3].

Fig. 1 shows the sample means and standard deviations of the maximum and minimum eigenvalues of $\Theta^H \Theta \Gamma \Theta^T \Gamma$ for varying $M$. For every value $M$, sets $\Gamma$ are generated uniformly random over all sets and the statistics are accumulated from 10,000 samples. A value of $K = 67$ was used for the simulation. For comparison, the sample means of the maximum and minimum eigenvalues of the Gram matrices of the Gaussian measurements are also shown.

As a specific example, if we consider Fig. 1 in terms of Lemma 1 we see that, with $|\Gamma| \leq 16$, the eigenvalues of $\Theta^H \Theta \Gamma$ within 1 of unity with high probability. This implies $\ell_0$ minimization can recover a random $8$-sparse signal with high probability.

More generally we note that from Fig. 1 we see that the eigenvalues of $\Theta^H \Theta \Gamma$ are, on average, closer to 1 by more than a standard deviation compared to the corresponding eigenvalues of Gaussian measurements. Thus, if Gaussian measurements satisfy any condition based on the RIP constants, then our $\Theta$ will also be able to recover a random $M$-sparse signal with high probability. Results are similar for other values of $K$.

It is important to note that here we have analyzed the measurement matrix $\Theta$ in isolation of the algorithm presented in Section 3. In this section, we have shown that $\Theta$ is suitable as compressed sensing measurements in general. In Section 7, we examine how the measurements perform jointly with our corresponding algorithm.

5. Specializations

Here we consider the use of Chirp codes in the special case of sparse Fourier signals as well as a modification to the algorithm to mitigate cross-term interference and noise.
5.1. **Φ for sparse Fourier signals**

Though we are interested in being able to determine the combination of columns of Θ, measurements are taken upon x and thus are made with the matrix Φ = ΘΨ⁻¹. We are therefore concerned with Φ for implementation. When Ψ is the Fourier matrix (if x is a sparse superposition of sinusoids), we can find the structure of Φ directly.

As described above, we set Θ to have the l,k entry of the form

\[ [Θ]_{lk} = e^{\frac{j2\pi l^2}{K}} e^{\frac{j2\pi lm}{K}} \quad \text{with} \quad k = Kr + m \in \mathbb{Z}_K^2. \]

This construction of Θ groups the columns of chirps in blocks of chirp rates.

Since Ψ is the Fourier matrix, Φ = ΘΨ⁻¹ is a matrix with rows formed by the \( K^2 \) length Fourier transform of the rows of Θ. The Fourier transform of the \( l \)th row of Θ is given by

\[
\sum_{k=0}^{K-1} [Θ]_{lk} e^{-\frac{j2\pi kr}{K}} = \sum_{r=0}^{K-1} \sum_{m=0}^{K-1} e^{-\frac{j2\pi kr}{K}} e^{-\frac{j2\pi rm}{K}} e^{-\frac{j2\pi ml}{K}}
\]

\[
= \sum_{r=0}^{K-1} e^{-\frac{j2\pi r(\omega)}{K}} \left[ \sum_{m=0}^{K-1} e^{-\frac{j2\pi m(Kl-\omega)}{K}} \right]
\]

\[
= K \delta_{l(\omega)\omega} e^{j\pi(\omega)\omega(1/K-1/K^2)} \frac{\sin(\pi(\omega)/K)}{\sin(\pi(\omega)/K)}
\]

where \( \delta_i \) is the Kronecker delta with \( i \) taken mod \( K \). Thus, the rows of Φ are periodic trains of delta functions modulated by a sinc function. This means that the measurements \( y \) can be formed simply as a weighted sum of a sparse number of samples of \( x \). Thus, the encoding of \( y \) has a relatively low computational cost.

This formulation of \( Φ \), along with the scenario of sparse signals in the Fourier domain, is used in the simulations illustrated later.

5.2. **Interference and noise mitigation**

At the expense of more computation, we can improve the performance of the algorithm by exploiting the availability of \( f_T(l) \) for different delays \( T \). By adding the FFT bins of each chirp rate \( r \) with those of the FFTs formed from the other delays, we can mitigate the effect of noise and any significant values from cross terms in (6). A bin with a chirp at \( r \) is correlated across different \( T \) while with a white noise approximation of the cross term interference (or simply white noise), other bins are not correlated and have zero mean. In the extreme case, we take a FFT for all \( K-1 \) possible shifts \( T \).

The modified algorithm is summarized here.

1. Choose a stopping energy \( \epsilon \).
2. Form \( y_T = \hat{y}(l)y(l + T) \) for every \( T \in \mathbb{Z}_K, T \neq 0 \) and take length \( K \) FFT of each.
3. Using \( 2r_T \mod K \), reorganize the output of each FFT such that the bins are in order of increasing \( r_T \).
4. Sum the absolute value of the reorganized FFTs and record the peak \( r_T \).
5. Multiply \( y(l) \) by \( e^{-\frac{j2\pi r_T^2}{K}} \) and take length \( K \) FFT.
6. Find the location of the peak and record as \( m_T \). Use the value of the peak to recover \( s_T \).
7. Replace \( y \) with \( y - s_T e^{-\frac{j2\pi m_T r_T^2}{K}} \).
8. Repeat steps (2)–(7) until \( \| y \|^2 < \epsilon \) or have iterated \( M \) times.

We compare the performance of original algorithm with this modified algorithm in Sections 7.1 and 7.2.

6. **Computational complexity**

Reconstruction in compressed sensing is normally done by solving a linear program minimizing \( \| s \|_1 \). As remarked earlier, this method is computationally intensive and has complexity \( O(N^3) \). An alternative scheme is the greedy matching pursuit algorithm with complexity \( O(KMN) \), which we use in simulations for comparison. Here, we consider the complexity of the chirp code algorithm.

By using the chirp decoding algorithm we leverage the efficiency of the FFT. Similar to matching pursuit, the algorithm iteratively pulls out the strongest signals. In this algorithm, each “peel” requires the computation of two FFTs of length \( K \): a first, to extract the coded chirp rate, and a second to extract the coded frequency. Since, for a \( M \)-sparse signal approximately \( M \) peels are required, the complexity of the computation is

\[ O(MK \log K). \]
As noted in Section 5.2 we can trade additional computation for improved performance by using multiple delays $T$ to form $f(l)$. In the most computationally intensive case, we perform $K$ length $K$ FFTs for each peel which gives an overall complexity of

$$O(MK^2 \log K).$$

In terms of computation, the algorithm is a significant improvement upon $\ell_1$ minimization and matching pursuit. For comparison, a table of various algorithms and their complexities can be found in [12].

7. Performance

Here we present some simulation results characterizing the performance of the algorithm. The simulations were produced using a measurement matrix formed as described in Section 5.1 acting upon a signal sparse in the Fourier domain. We compare its performance against the matching pursuit algorithm using a Gaussian random $\Phi$.

7.1. Sparsity requirements

An important examination is whether the algorithm’s improved computation complexity degrades the sparsity level at which signals can be recovered. We look at the sparsity requirements for various signal lengths while using $K = \sqrt{N}$ measurements.

Figs. 2 and 3 compare the reconstruction error of the chirp sensing code algorithm with that of matching pursuit for signal lengths $N = 41^2$ and $N = 67^2$. A signal comprised of a small number of sinusoids was measured and reconstructed. We include two chirp sensing code algorithms: using a single shift as well as using all possible shifts. We see that when all shifts are utilized, the chirp sensing algorithm is able to outperform matching pursuit, successfully reconstructing signals containing more sinusoids.

Results for other values of $N$ show that the sparsity levels $M$ required by the algorithm for signal recovery follow (3) with different values of $c$ for single delays and multiple delays.

7.2. Detection in noise

Strictly speaking, $s$ is sparse if it has very few non-zero elements. However, this is not a good model of practical signals. Practical signals will have small values in all elements of $s$ either due to noise or components that can be discarded. A practical recovery algorithm must be able to work under these circumstances. Further, it is important to know at which noise levels the algorithm can operate.

Fig. 4 compares the performance of the chirped sensing code algorithm with matching pursuit. In the figure, we examine the detection of a single sinusoid in noise. We consider a correct detection if the first “peel” of the signal corresponds to
the sinusoid. The figure was generated by simulation using $K = 41$ measurements and length $N = 41^2$ signals. Probabilities were estimated using 1000 samples. Both the single shift $O(MK \log K)$ algorithm and the $O(MK^2 \log K)$ algorithm using all shifts are included in the comparison.

We see that the algorithm that uses all the shifts achieves the performance of matching pursuit. These results can be compared to those in [9].
8. Conclusions and extensions

The chirped sensing codes we introduced here are an illustration of how, by particularly selecting measurements in $\Phi$, we can utilize a computationally efficient reconstruction scheme. The choice of the measurements in $\Phi$ were made such that from $\Theta$, its form in the sparse basis, we can recover small linear combinations of columns. In particular, we designed $\Theta$ filled with columns of chirps since we have an efficient method to recover chirp rates and frequencies from a small superposition. Further, with this design of $\Theta$, $\Phi$ has a convenient form in the case of sparse Fourier signals.

Unlike most compressed sensing literature, we used deterministic measurements. By finding bound on restricted isometry constants $\delta_M$ we can provide guarantees on recoverability from our designed $\Theta$. In addition, we considered a modified version of the RIP which regards the signal, rather than the measurements, stochastically. Empirical evidence showed that the majority of Gram matrices of $\Theta$ have eigenvalues closer to 1 than correspondingly sized Gram matrices of random Gaussian measurements. This, in turn showed that a signal recoverable from the random measurements is very likely recoverable from measurements made with $\Theta$.

Signal recovery from our measurements was also shown by the implementation of our decoding algorithm. The recovery exploited the efficiency of the FFT in each of two steps: the first to recover the chirp rates and second to recover the chirp frequency. By identifying the chirp rates and frequencies, the superimposed columns of $\Theta$ are determined. In simulation the algorithm was shown to equal the performance of matching pursuit in noise resilience and exceed matching pursuit’s performance in signal sparsity requirements.

A limitation of the chirp code algorithm is the restriction $K \geq \sqrt{N}$. This derives from the size of the family of length $K$ chirps which necessitates $\Theta$ be $K \times K^2$ or narrower. As a result, this limits the algorithm’s abilities in situations where $K$ must be small. Stemming from this work, a similar algorithm based on second-order Reed–Muller codes found in [10] addresses this. Second-order Reed–Muller codes can be viewed analogously to chirps and decoding can be done using the Fast Hadamard Transform in place of the FFTs. The class of length $K$ second-order Reed–Muller codes is very large, essentially removing this lower bound on the number of measurements. This paper can serve as a preliminary read to [10] for those less familiar with Reed–Muller codes.

Regardless of the above limitation, chirp sensing codes excel as method for fast signal recovery with compressed sensing.

Acknowledgment

The authors would like to thank Jarvis Haupt, Waheed Bajwa and Rob Nowak for their useful discussions on RIP recovery conditions.

References