

# Stability Of Efficient Deterministic Compressed Sensing For Images With Chirps And Reed-Muller Sequences

Submitted to the special issue on Numerical Analysis

S. Datta<sup>\*†</sup>, K. Ni<sup>\*‡</sup>, P. Mahanti<sup>§</sup>, and S. Roudenko<sup>‡</sup>

## Abstract

We explore the stability of image reconstruction algorithms under deterministic compressed sensing. Recently, we have proposed [1, 2, 3] deterministic compressed sensing algorithms for 2D images. These algorithms are suitable when Daubechies wavelets are used as the sparsifying basis. In the initial work, we have shown that the algorithms perform well for images with sparse wavelets coefficients. In this work, we address the question of robustness and stability of the algorithms, specifically, if the image is not sparse and/or if noise is present. We show that our algorithms perform very well in the presence of a certain degree of noise. This is especially important for MRI and other real world applications where some level of noise is always present.

*Key words and phrases* : Compressed Sensing, Reed-Muller Sequences, Chirps, Image Reconstruction

*2000 AMS Mathematics Subject Classification* — 94A08, 94A20.

## 1 Introduction

The theory of compressed sensing [4, 5, 6] states that it is possible to recover a sparse signal from a small number of measurements. A signal  $x_0 \in \mathbb{C}^N$  is  $k$ -sparse in a basis  $\Psi = \{\psi_j\}_{j=1}^N$  if  $x_0$  is a weighted superposition of at most  $k$  elements of  $\Psi$ . Compressed sensing broadly refers to the inverse problem of reconstructing such a signal  $x_0$  from linear measurements  $\{y_\ell = \langle x_0, \phi_\ell \rangle \mid \ell = 1, \dots, n\}$  with  $n < N$ , ideally with  $n \ll N$ . In the general setting, one has  $\Phi x_0 = y$ , where  $\Phi$  is an  $n \times N$  sensing matrix having the measurement vectors  $\phi_\ell$  as its rows,  $x_0$  is a length- $N$  signal and  $y$  is a length- $n$  measurement. By now, many authors have proposed different sensing matrices and reconstruction algorithms, establishing the feasibility of such reconstruction in practice. Applications have been shown for medical images [7], communications [8], analog-to-information conversion [9], geophysical data analysis [10], etc. The standard compressed sensing technique guarantees exact recovery of the original signal with overwhelmingly high probability if the sensing matrix satisfies the Restricted Isometry Property (RIP). This means that for a fixed  $k$ , there exists a small number  $\delta_k$ , such that

$$(1 - \delta_k) \|x_k\|_{\ell_2}^2 \leq \|\Phi x_k\|_{\ell_2}^2 \leq (1 + \delta_k) \|x_k\|_{\ell_2}^2, \quad (1.1)$$

for any  $k$ -sparse signal  $x_k$ . Solving for the original sparse signal with  $\ell_1$  minimization,

$$\min_x \|x\|_{\ell_1}, \quad \text{subject to } \Phi x = y, \quad (1.2)$$

guarantees successful recovery with a very high probability. This is a well-known theory and has been verified empirically in several papers, e.g., [5, 11]. In this paper, we explore the stability of deterministic compressed sensing when there is noise present in the signal or its measurements, or the signal is not sparse. Traditional

---

\* These authors contributed equally to this work

†Department of Mathematics, University of Idaho, Moscow ID 83844-1103, USA, [sdatta@uidaho.edu](mailto:sdatta@uidaho.edu)

‡HRL Laboratories, LLC, Malibu, CA 90265, USA, [kangyu.ni@gmail.com](mailto:kangyu.ni@gmail.com)

§School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85287, USA; [prasan.mahanti@gmail.com](mailto:prasan.mahanti@gmail.com)

‡ Department of Mathematics, George Washington University, Washington, DC 20052, USA, [roudenko@gwu.edu](mailto:roudenko@gwu.edu)

compressed sensing techniques use random projections for the sensing matrix as opposed to deterministic measurements. We discuss some comparisons between the two later in this section.

Signals that are *compressible* are not sparse in any transform domain. A signal  $x$  is said to be compressible [12] if its coefficients follow a decay rate ,

$$|x|_{(j)} \leq C_s(j + D)^{-s}, \quad (1.3)$$

where  $|x|_{(j)}$  is the  $j$ th largest coefficient ( $|x|_{(1)} > |x|_{(2)} > \dots$ ) of  $x$  with respect to some basis or transform  $\Psi$ ,  $s \geq 1$ , the constant  $C_s$  depends only on  $s$ , and  $D \geq 0$  is a shift. A compressible signal may not have any of its coefficients exactly equal to zero. The most relevant recent results on the reconstruction of compressible signals are summarized below.

In [13], Candès and Romberg demonstrate that empirically it is possible to recover a compressible signal  $x_0$  from about  $3k$  to  $5k$  projections with accuracy as good as the optimal  $k$ -term wavelet approximation of  $x_0$ . Their experiments are carried out for compressible 1D signals and 2D images. For images, they proposed a recovery algorithm that minimizes the total variation (TV) in the image domain (with  $\Psi$  being the inverse wavelet transform), with  $\ell_1$ -norm constraints on the wavelet coefficients:

$$\min \|\Psi x\|_{\text{TV}} \text{ such that } \Phi x = y, \text{ and } \|x_j\|_{\ell_1} \leq \|x_{0_j}\|_{\ell_1}. \quad (1.4)$$

The  $\ell_1$ -constraints restrict the locations where the large wavelet coefficients can appear, assuming the  $\ell_1$ -values  $\|x_{0_j}\|_{\ell_1}$  of each subband  $x_{0_j} = \{x_{0_{j,m}}, m = 0, \dots, 2^{j-1} - 1\}$  are known. In their experiments, random Fourier matrices were applied to the wavelet coefficients of the images. The number of measurements were taken to be 15%, 23%, 30%, and 38% of the total number of pixels, i.e.,  $256 \times 256$ . It was shown that the recovery from  $3k$  to  $5k$  random projections is comparable to the best  $k$ -term wavelet approximation [13].

The main results in [12] by Candès, Romberg and Tao are that the standard compressed sensing is able to stably recover *sparse* as well as *compressible* signals from noisy measurements. For the latter, if  $k$  is chosen such that  $\delta_{3k}$  and  $\delta_{4k}$  in (1.1) satisfy  $\delta_{3k} + 3\delta_{4k} < 2$ , the error bound for compressible signals with noise in the measurements is

$$\|x_0 - \hat{x}\|_{\ell_2} \leq C_{1,k} \cdot \|\mu\|_{\ell_2} + C_{2,k} \frac{\|x_0 - x_k\|_{\ell_1}}{\sqrt{k}}, \quad (1.5)$$

where  $x_0$  is the true signal,  $\hat{x}$  is the recovered signal,  $x_k$  is the best  $k$ -term approximation of  $x_0$ ,  $\mu$  is any perturbation, and constants  $C_{1,k}$  and  $C_{2,k}$  depend on  $\delta_{4k}$ . It is stated (Theorem 2, [12]) that  $\ell_1$  minimization stably recovers the  $k$ -largest entries. As the authors point out, this is a deterministic statement, i.e., there is *no probability of failure*. Their experiments on both 1D and 2D compressible signals with noise added to measurements confirm this theoretical result. It is also observed in their experiments that if the noise is Gaussian with small standard deviation, then the recovery error is dominated by the approximation error, i.e., by the last term in (1.5).

Even though the above methods and results deal with non-sparse compressible signals they are not deterministic in the sense that the measurements are obtained via random projections. In practice, it is beneficial to use deterministic sensing matrices that are pre-determined in order to save computation time and memory. Work in this direction was done by DeVore ([14]), then followed by a few others. However, the recovery results were far inferior compared to that obtained by traditional random sensing. Recently, another deterministic compressed sensing method (mainly applicable for 1D signals) was proposed by Applebaum, Howard, Searle, and Calderbank [15] by using the chirp transform to create the sensing matrix. In their experiments (Figs. 2 and 3 in [15]), the full pass chirp decode algorithm results in smaller reconstruction error than using Gaussian random matrices with matching pursuit. They have also shown empirically (see Fig. 4 in [15]) that these two methods are comparable when there is noise in the measurements. The reconstruction algorithm is good for 1D signals but in experiments given in [15] the 1D signals are of size  $N = 41^2$  or  $67^2$  only, much smaller than the size encountered for 2D signals, such as images. A similar reconstruction algorithm was also proposed for a sensing matrix made from Reed-Muller sequences [16]. In [3], we proposed new reconstruction algorithms for deterministic sensing matrices made from chirp and Reed-Muller sequences that are suitable for 2D images.

The advantage of using deterministic matrices for compressed sensing is that reconstruction can be very efficient [15, 16]. The fast reconstruction algorithm [15, 16, 17] is called the Quadratic Reconstruction Algorithm. This algorithm takes advantage of the multivariable quadratic functions that appear as exponents

of the entries in the sensing matrix, and therefore, only requires vector-vector multiplication instead of matrix-vector multiplication required in Basis and Matching Pursuit algorithms that are used for reconstructions with random sensing matrices. In [17], Calderbank, Howard and Jafarpour set forth criteria on  $\Phi$  that ensure a high probability that the mapping taking the  $k$ -sparse signal vector  $x_k$  to the measurement vector  $y$  is injective, assuming a uniform probability distribution on the unit-magnitude  $k$ -sparse vectors in  $\mathbb{C}^N$ . They say that  $\Phi$  has the Statistical Restricted Isometry Property (StRIP) with respect to parameters  $\varepsilon$  and  $\delta$  if

$$(1 - \varepsilon)\|x_k\|_2^2 \leq \|\Phi x_k\|_2^2 \leq (1 + \varepsilon)\|x_k\|_2^2 \quad (1.6)$$

holds with probability exceeding  $1 - \delta$  when  $x_k$  is assumed to be uniformly distributed among  $k$ -sparse vectors in  $\mathbb{C}^N$  of some fixed norm (e.g., unit norm). They show that such deterministic sensing matrices satisfying StRIP can be constructed by chirps, Reed-Muller (RM) sequences, and BCH codes, as done in [15, 16, 17]. In the presence of noise, if  $\Phi$  satisfies the StRIP property with parameters  $\varepsilon$  and  $\delta$ , the reconstruction error due to the Quadratic Reconstruction Algorithm is given by

$$\|x_0 - \hat{x}\|_{\ell_2} \leq \frac{5 + \varepsilon}{1 - \varepsilon} \|x_0 - x_k\|_{\ell_2} + \frac{2}{1 - \varepsilon} \|\mu\|_{\ell_2}, \quad (1.7)$$

where  $x_0$  is the true signal,  $\hat{x}$  is the recovered signal,  $x_k$  is the best  $k$ -term approximation of  $x_0$ , and  $\mu$  is the measurement noise from a Gaussian distribution, see [17]. As mentioned there, this bound is tighter than bounds obtained by using random ensembles [12] and expander-based methods [18]. While the Quadratic Reconstruction Algorithm [15, 16] is shown to accurately reconstruct 1D sparse signals with nonzero locations chosen uniformly at random, from a small number of measurements, it does not work when it is directly applied to 2D signals, like images. This is because the locations of the nonzero coefficients, usually in the wavelet domain, are not distributed uniformly. Our proposed algorithm [3] takes this into account and we have shown that it outperforms a standard compressed sensing method, which takes random noiselet measurements and uses  $\ell_1$  minimization for reconstruction, for example, see [19].

Authors in [17] also derived that the deterministic sensing matrices are resilient to noise. Suppose the noise comes from the measurements, say  $y = \Phi x_k + \mu$ , where  $\mu$  are iid complex Gaussian random variables with zero mean and variance  $2\sigma^2$ . Suppose  $\Phi$  satisfies

$$(1 - \varepsilon')^2 \|x_k\|_{\ell_2}^2 \leq \|\Phi x_k\|_{\ell_2}^2 \leq (1 + \varepsilon')^2 \|x_k\|_{\ell_2}^2 \quad (1.8)$$

with probability exceeding  $\delta > 0$ . Then, for  $\gamma \geq 0$ ,

$$(1 - \varepsilon' - \gamma)^2 \|x_k\|_{\ell_2}^2 \leq \|y\|_{\ell_2}^2 \leq (1 + \varepsilon' + \gamma)^2 \|x_k\|_{\ell_2}^2, \quad (1.9)$$

with probability exceeding  $1 - 2(\delta + S[\frac{\gamma\|x\|}{\sigma}])$ , where

$$S(r) = \left( \int_r^\infty e^{-\frac{y^2}{2}} y^{(N-1)} dy \right) \left( \int_0^\infty e^{-\frac{y^2}{2}} y^{(N-1)} dy \right)^{-1}. \quad (1.10)$$

Therefore, we see that the noisy measurement is bounded in (1.9) by the signal in a similar way as the inequality (1.8) in the StRIP property.

Now, suppose that the noise comes from the signal, say  $y = \Phi(x_0 + \mu)$ , where  $\mu$  is a complex, multivariate, Gaussian distributed noise with zero mean and covariance  $E(\mu\mu^\dagger) = 2\sigma^2 I_{N \times N}$ . In this case, the estimates given in (1.9) can still be applied. The covariance of the measurement vector is  $E(\Phi\mu(\Phi\mu)^\dagger) = E(\Phi\mu\mu^\dagger\Phi^\dagger) = 2\sigma^2\Phi\Phi^\dagger$ . For deterministic matrices,  $\Phi\Phi^\dagger = \frac{N}{n}I_{n \times n}$  and this gives  $E(\Phi\mu(\Phi\mu)^\dagger) = 2\sigma^2\frac{N}{n}I_{n \times n}$ . The factor  $\frac{N}{n}$  that appears in  $E(\Phi\mu(\Phi\mu)^\dagger)$  is not present in  $E(\mu\mu^\dagger)$ , making the measurement variance in this case larger than the source noise variance  $2\sigma^2$ . This means that noise coming from the signal is harder to deal with than noise coming from the measurements.

One important and desirable task in reconstruction of signals via deterministic compressed sensing is to reconstruct signals efficiently. Moreover, such reconstruction should be stable under noise and also work when the signal is non-sparse. In this paper, we discuss the stability and robustness of the algorithms we introduced in [3]. We show that these algorithms are stable for non-sparse (compressible) signals and robust under various types of noise. In addition, we also indicate how, when using the sensing matrix composed

from Reed-Muller sequences, the algorithms in [3] may be modified to efficiently handle excessively large images or large 2D signals.

This paper is structured as follows: in Section 2, we review our reconstruction algorithm of [3] and comment on how to incorporate properties of the Reed-Muller sequences so that the algorithm can handle images of size much larger than considered so far. Then we discuss how to choose the proper wavelet domain and the level of decomposition of the image under the chosen wavelet basis. In Section 3, we discuss the stability and robustness under noise and for non-sparse signals. Finally, Section 4 contains the discussion of the results.

## 2 Choosing a Suitable Sparsifying Domain

We address the issue of choosing a suitable sparsifying domain for images, since the compressed sensing theory is based on the fact that the signal is sparse. Wavelets are commonly used for 2D images, but the choice of the specific wavelet to be used is not usually discussed much. In the next few subsections, we first review the fast reconstruction algorithm for 2D signals including comments on modifying the Reed-Muller sensing case based on certain properties of the Reed-Muller sequences. Then we discuss how to choose a suitable Daubechies DN wavelet basis and the optimal level of decomposition. The images used in this work are shown in Figure 1, where (a) is a  $512 \times 512$  MR image of brain, (b) is a  $512 \times 512$  MR angiogram image, and (c) is a natural image of  $1024 \times 1024$  resolution, which we refer to as the man image.

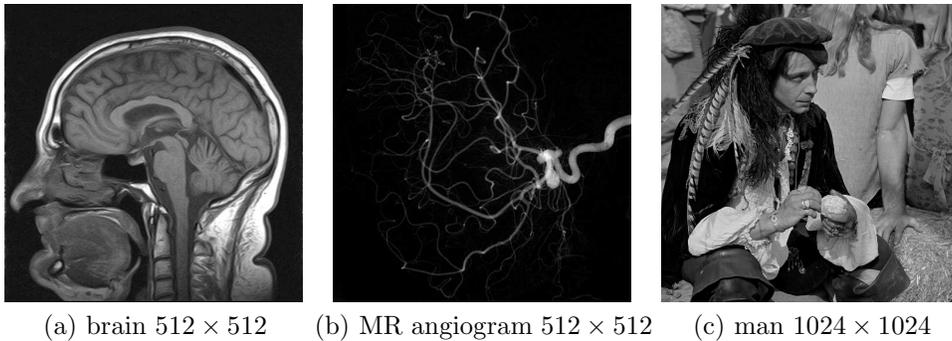


Figure 1: 2D images used in this paper

### 2.1 Reconstruction algorithm for deterministic sensing matrices

We start with a discussion of deterministic compressed sensing using chirps, i.e., frequency modulated discrete sinusoids. Our reconstruction algorithm using Reed-Muller sequences is similar and is briefly discussed here as well. The detailed reconstruction technique for both can be found in [3].

#### 2.1.1 Compressed sensing with chirps

A discrete chirp of length  $n$  with *chirp rate*  $r$  and *base frequency*  $m$  has the form

$$\phi_{r,m}(\ell) = \frac{1}{\sqrt{n}} e^{\frac{2\pi i}{n} r \ell^2 + \frac{2\pi i}{n} m \ell}, \quad r, m, \ell \in \mathbb{Z}_n. \quad (2.1)$$

Note that the coefficient  $1/\sqrt{n}$  is present in order for the vector to have a unit  $\ell_2$  norm. For a fixed  $n$ , there are  $n^2$  possible pairs  $(r, m)$ . The full chirp sensing matrix  $\Phi$  thus has size  $n \times n^2$  and can be written as

$$\Phi_{\text{chirp}} = [U_{r_1} \ U_{r_2} \ \cdots \ U_{r_t} \ \cdots \ U_{r_n}], \quad 1 \leq t \leq n. \quad (2.2)$$

Each  $U_{r_t}$  is an  $n \times n$  matrix with columns given by chirp signals having a fixed chirp rate  $r_t$  with base frequency  $m$  varying from 0 to  $n - 1$ . The chirp rate  $r$  also varies from 0 to  $n - 1$ . Therefore, column  $j = m + rn + 1$  of  $\Phi_{\text{chirp}}$  is a discrete chirp with chirp rate  $r$  and base frequency  $m$ .

**Example 2.1.** Let  $n = 2$ ;  $r, m, \ell \in \{0, 1\}$ . Then

$$U_0 = \begin{bmatrix} 1 & 1 \\ 1 & e^{i\pi} \end{bmatrix}, U_1 = \begin{bmatrix} 1 & 1 \\ e^{i\pi} & e^{i\pi+i2\pi} \end{bmatrix}.$$

The  $2 \times 4$  sensing matrix is then given by

$$\Phi_{\text{chirp}} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & e^{i\pi} & e^{i\pi} & e^{i\pi+i2\pi} \end{bmatrix}.$$

Note that  $U_0$ , the  $n \times n$  matrix corresponding to chirp rate  $r = r_1 = 0$ , is the Discrete Fourier Transform (DFT) matrix. The Fast Fourier Transform (FFT), which implements the DFT, is used in the algorithm for recovering a sparse signal from chirp measurements, see Section 2.1.5.

### 2.1.2 The chirp sensing matrix for 2D signals

Despite the success in accurate reconstruction of very sparse one-dimensional (1D) signals with the algorithm described in [16], applying it to real two-dimensional (2D) images is impractical. This is because, in general, real images are not as sparse in any transform domain as the one-dimensional signals used in [15] or [16].

A good approximation of a  $256 \times 256$  pixel image is typically obtained by retaining the largest 10% wavelet coefficients in some suitably chosen wavelet domain. In particular, many medical images are well approximated by transform coding using 10–20% of their wavelet coefficients, but begin to show appreciable degradation as the percentage of coefficients retained falls below these levels. However, a  $256 \times 256$  image with 10% sparsity has 6,554 nonzero coefficients, which is much larger than the sparsity considered for the 1D signals in [15] and [16]. A rule of thumb, see [20, Theorem 1], for the number of measurements in the standard compressed sensing using the Gaussian random matrices with  $\ell_1$  minimization is given by

$$n > k \log_2(1 + N/k). \quad (2.3)$$

This rule guarantees successful reconstruction with high probability if the number of measurements  $n$  is large compared to the sparsity and signal size. Using (2.3) in the above example, at least 22,672 measurements are needed for the correct reconstruction. The ratio  $N/n$  is 2.89, and this implies that roughly only three chirp rates are needed to form the sensing matrix.

As explained above, due to the nature of sparsity of images and the rule of thumb (2.3), a few submatrices of  $\Phi_{\text{chirp}}$  can be used to make the sensing matrix, with the ratio  $N/n = 2.89$  for 10%-sparse images. In practice, a larger ratio can be used, such as 4 for 10%-sparse images, which will be analyzed later. Consequently, there is more freedom in the choice of the chirp rates  $r$  when constructing the sensing matrix.

The inner product of any pair of distinct chirp vectors is as follows:

$$|\langle \phi_{r_t, m_t}, \phi_{r_{t'}, m_{t'}} \rangle| = \begin{cases} \frac{1}{\sqrt{n}}, & \text{if } r_t \neq r_{t'} \\ 0, & \text{if } r_t = r_{t'} \text{ and } m_t \neq m_{t'}. \end{cases} \quad (2.4)$$

Therefore, a submatrix should use as few chirp rates as possible and the choice of the chirp rates can be arbitrary. For example, the submatrix can be

$$\Phi_{\text{chirp}} = \begin{bmatrix} U_{r_1} & U_{r_2} & U_{r_3} & \tilde{U}_{r_4} \end{bmatrix}, \quad (2.5)$$

where  $r_1 = 0, r_2 = 1, r_3 = 2, r_4 = 3$ , and  $\tilde{U}_{r_4}$  denotes a submatrix of  $U_{r_4}$  so that the number of columns of  $\Phi_{\text{chirp}}$  matches the signal size. When only the first  $J$  submatrices  $U_{r_1}, \dots, U_{r_J}$  are used to form the sensing matrix, it follows from an argument given by Alltop [21, Section IV] that  $n$  need not necessarily be prime. Rather, the crucial condition for unique identification of the chirp rate  $r_t$  in the reconstruction algorithm is that the smallest prime divisor of  $n$  is greater than  $J$ . For instance, the sensing matrix for a  $256 \times 256$  ( $N = 65536$ ) image may be taken to be of size  $n \times N = 16385 \times 65536$ . Note that  $16385 = 5 \times 29 \times 113$  is closest to and larger than  $25\% \times 65536 = 16384$  whose smallest prime divisor is greater than 4. Here, the 25% ratio comes from the fact that four chirp rates are used. In this example,  $\tilde{U}_{r_4}$  is of size  $16385 \times 16381$  and can be  $U_{r_4}$  without the last four columns.

### 2.1.3 Reed-Muller sequences in deterministic compressed sensing

The set of real valued second-order Reed-Muller (RM) codes or sequences with length  $2^p$  is parameterized by  $p \times p$  binary symmetric matrices  $P$  and binary  $p$ -vectors  $b \in \mathbb{Z}_2^p$ . In terms of these parameters, a second-order RM code is given by

$$\phi_{P,b}(a) = \frac{1}{\sqrt{2^p}} i^{(2b+Pa)^\top a}. \quad (2.6)$$

In analogy with the chirps, the vector  $b$  in the linear term of (2.6) and the matrix  $P$  in the quadratic term may be regarded as the ‘‘frequency’’ and ‘‘chirp rate’’ of the code, respectively. In the expression (2.6),  $a \in \mathbb{Z}_2^p$  indexes the  $2^p$  components of the code  $\phi_{P,b}$ . So, for given parameters  $P$  and  $b$ , the code is a vector of length  $2^p$ . These vectors will serve as the columns of the sensing matrix  $\Phi_{\text{RM}}$ . In addition, during implementation,  $P$  is taken to be zero on its main diagonal to ensure that the codes generated are real valued. This implies that the components of these codes are all  $\pm 1$ .

**Example 2.2.** Let  $p = 2$ . Then  $\mathbb{Z}_2^2 = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$  and  $a, b \in \mathbb{Z}_2^2$ . There are  $2^{2(2-1)/2} = 2$  zero diagonal symmetric matrices of size  $2 \times 2$ . These are

$$P_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

and

$$P_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Set  $b = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ . Then if  $a = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ , we get for  $P_1$

$$\phi_{P_1,b}(a) = \phi_{P_1,b} \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right) = -1.$$

Similarly for other values of  $a$  in  $\mathbb{Z}_2^2$  one gets

$$\phi_{P_1,b} \left( \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) = 1, \quad \phi_{P_1,b} \left( \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) = -1, \quad \phi_{P_1,b} \left( \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) = 1.$$

The vector  $\begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}$  will be one of the columns of the sensing matrix.

In general, the compressed sensing matrix proposed in [16] has the form

$$\Phi_{\text{RM}} = [U_{P_1} \quad U_{P_2} \quad \cdots \quad U_{P_t} \quad \cdots \quad U_{P_{2^{p(p-1)/2}}}], \quad 1 \leq t \leq 2^{p(p-1)/2}, \quad (2.7)$$

where each  $U_{P_t}$  is a  $2^p \times 2^p$  orthogonal matrix whose columns are  $\phi_{P_t,b}$  with  $b$  going through all binary  $p$ -vectors. In addition, each  $\phi_{P_t,b}$  is multiplied with a phase factor  $(-1)^{\text{wt}(b)}$ , where  $\text{wt}(b)$  is the Hamming weight of  $b$ , i.e., the number of ones in  $b$ . The extra phase factor ensures that the total number of plus and minus signs of the inner products of any two columns are the same. For convenience,  $P_1$  is chosen to be the zero matrix, and therefore, without the phase factor,  $U_{P_1}$  is a Hadamard matrix up to a scaling [22]. Consequently, multiplication by  $U_{P_1}$  is the Walsh-Hadamard transform (DHT) which, up to a scaling, is its own inverse [22]. In the reconstruction of a sparse signal using the Reed-Muller sensing matrix the algorithm uses the DHT, see Section 2.1.5. This is analogous to the DFT that is used with the chirp sensing matrix.

**Example 2.3.** In Example 2.2,

$$U_{P_1} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

In this case, the first column corresponds to  $b = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ , the second column corresponds to  $b = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$  (as shown in Example 2.2), and so on. Similarly, for the matrix  $P_2$  we get

$$U_{P_2} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ -1 & -1 & -1 & -1 \end{bmatrix}.$$

Together, we get the  $2^2 \times 2^3$  Reed-Muller sensing matrix, for  $p = 2$ , as

$$\Phi_{\text{RM}} = [U_{P_1} \ U_{P_2}] = \begin{bmatrix} 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \end{bmatrix}.$$

As noted earlier, the vector  $b$  in the linear term of (2.6) and the matrix  $P$  in the quadratic term may be regarded as the binary ‘‘frequency’’ and ‘‘chirp rate’’ of the code, respectively. There are  $2^{p(p-1)/2}$  such  $P$  matrices with zero-diagonal, so the maximum size of the RM sensing matrix in (2.7) is  $2^p \times 2^{p(p+1)/2}$ . For a given  $P$ ,  $2^p$  codes are obtained by varying  $b$ , yielding the columns of a  $2^p \times 2^p$  orthogonal matrix  $U_P$ .

The Delsarte-Goethals set  $DG(p, r)$  is the binary vector space of  $p \times p$  binary symmetric matrices with the property that the difference between any two distinct matrices has rank greater than or equal to  $p - 2r$  [23]. Evidently, these sets are nested

$$DG(p, 0) \subset DG(p, 1) \subset \dots \subset DG\left(p, \frac{p-1}{2}\right). \quad (2.8)$$

**Example 2.4.** Let  $p = 3$ . Then  $DG(3, 0)$ , consisting of matrices whose differences have rank at least 3, is spanned by the set [24]

$$\left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \right\}.$$

The set  $DG(3, 1)$ , consisting of matrices whose differences have rank at least  $3 - 2 = 1$ , is spanned by [24]

$$DG(3, 0) \cup \left\{ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \right\}.$$

Note that these 6 matrices generate all the  $2^6$  binary symmetric matrices of size  $3 \times 3$ .

The set of all  $P$  matrices that reside in  $DG(p, 0)$  is called the *Kerdock* set. This means that the difference between any two distinct matrices in the Kerdock set has full rank. The Reed-Muller codes made from the matrices  $P$  in the Kerdock set produces the Kerdock codes [25]. Two distinct Kerdock codes, normalized to unit length, have inner product modulus that is either zero (if they correspond to the same  $P$ ) or  $1/\sqrt{2^q}$  (if they correspond to distinct  $P$ s). More generally,

$$|\langle \phi_{P,b}, \phi_{P',b'} \rangle| = \begin{cases} \frac{1}{\sqrt{2^q}}, & 2^q \text{ times,} \\ 0, & 2^p - 2^q \text{ times,} \end{cases} \quad (2.9)$$

where  $q = \text{rank}(P - P')$ . So, if the domain of  $P$  is  $DG(p, q)$ , the set of possible inner product modulus values for distinct normalized codes is  $\{0, 2^{-p/2}, \dots, 2^{-(p-2r)/2}\}$ . Allowing  $P$  to range over all of  $DG(p, (p-1)/2)$ , (2.6) gives the full set of second-order RM codes.

Defining  $N = 2^{p(p+1)/2}$  and  $n = 2^p$ , a  $k$ -sparse signal  $x \in \mathbb{C}^N$  yields a measurement  $y = \Phi_{\text{RM}}x \in \mathbb{C}^n$ , which is the superposition of  $k$  RM functions

$$y(a) = z_1 \phi_{P_1, b_1}(a) + z_2 \phi_{P_2, b_2}(a) + \dots + z_k \phi_{P_k, b_k}(a) = \sum_{t=1}^k z_t \phi_{P_t, b_t}(a). \quad (2.10)$$

In (2.10),  $z_t$  are used instead of  $x$  in order to only write the nonzero terms, and  $P_t$  and  $b_t$  may individually repeat in the equation.

### 2.1.4 The Reed-Muller sensing matrix for 2D signals

When forming a sensing matrix from submatrices of the RM matrix given in (2.7), the choice of the submatrix cannot be arbitrary. The inner product of two columns of  $\Phi_{\text{RM}}$ , one taken from  $U_{P_t}$  and another from  $U_{P_{t'}}$ ,  $t \neq t'$ , is given by (2.9) with  $q = \text{rank}(P_t - P_{t'})$ . If  $q = p$ , the inner product is always  $1/\sqrt{2^p}$ , which is smaller than the inner product in other cases,  $q < p$ . Since the nonzero locations of the signal are unknown, it is desirable that the inner products between any two columns are as small as possible, thus making the columns of the resulting sensing matrix close to orthogonal. Taking  $q = p$ , and thus drawing  $P$  matrices only from  $DG(p, 0)$  (i.e., the Kerdock set) gives the best situation. For a given  $p$ , there are  $2^{p-1}$  zero-diagonal matrices in the Kerdock set [24]. A sensing matrix can be constructed in the form

$$\Phi_{\text{RM}} = [U_{P_1} \quad U_{P_2} \quad U_{P_3} \quad U_{P_4}], \quad (2.11)$$

where  $P_1, P_2, P_3$ , and  $P_4$  are matrices from the Kerdock set. The idea behind choosing four  $P$  matrices or chirp rates is the same as that discussed in Section 2.1.2 for the chirp sensing matrix. For example, the sensing matrix for a  $256 \times 256$  ( $N = 2^{16}$ ) image with 10% sparsity is of size  $n \times N = 2^{14} \times 2^{16}$ , which means that only 25% of the signal entries are sampled. Note that for images with sparsity much smaller than 10%, fewer measurements are needed, and therefore, more  $P$  matrices can be used, since the ratio  $N/n$  becomes larger.

### 2.1.5 The reconstruction algorithm

Here we outline the reconstruction algorithm. For more details, we refer the reader to [3].

Input:  $y$  and  $\Phi = [U_1 \quad -U_2 \quad U_3 \quad -U_4]$ .

Output:  $\tilde{z}$

1. Approximation: Perform hard-thresholding  $U_1^*y$  to obtain a set of nonzero locations, denoted by  $\Gamma$ . Let  $A = U_1|_{\Gamma}$  be a submatrix of  $U_1$  restricted on the set  $\Gamma$ . Then, the initial approximation is  $\tilde{z} = A^*y$  and the residual is obtained by  $y_0 = y - A\tilde{z}$ .
2. Detection<sup>1</sup>: From  $w(t, \ell) = \text{DHT}_n \left\{ y_0(\ell) \overline{v_t(\ell)} \right\}$  or  $\text{DFT}_n \left\{ y_0(\ell) \overline{v_t(\ell)} \right\}$ ,  $t = 1, 2, 3, 4$ , where  $v_t$  is the first column of  $U_t$ , update  $\Gamma = \Gamma \cup \left\{ \text{locations associated with } d \text{ largest } |w(t, \ell)| \right\}$ . Let  $A = \Phi|_{\Gamma}$ .
3. Least Squares:  $\tilde{z} = \underset{z}{\text{argmin}} \|y - Az\|_2$ .
4. Define  $y_0 = y - A\tilde{z}$ . Repeat 2.-4. until  $\|y_0\|_2$  is sufficiently small.

If all the non-zero terms are caught by the set  $\Gamma$  then the precision of reconstruction is dependent on the LSQR algorithm [26] that is used in step 3. above.

## 2.2 Using the hierarchy of the Delsarte-Goethals (DG) sets for larger signals

So far while using the RM sensing matrix the size of the images was such that it was possible to pick all the  $P$  matrices, needed to construct the sensing matrix, from the Kerdock set. Sometimes one has to deal with signals whose sizes are such that the required sensing matrix has to be larger than what can be obtained only from the Kerdock set. Then we have to move on to  $P$  matrices from the DG sets that do not have the full rank property of the Kerdock matrices. This leads us to look at the rank and other properties, see, for example, (2.8), of matrices from the DG sets as described in Section 2.1.3. In this case some assumptions on the signal are useful. Compressed sensing for signals with prior information has been addressed by other authors [27, 28, 29, 30]. These authors design appropriate reconstruction algorithm but not sensing matrices as we do here.

We assume that there is some a priori knowledge of the behavior of the signal coefficients. Let us suppose that the signal is sparse and the coefficients of the signal (with respect to some sparsity domain, for example,

---

<sup>1</sup>DHT is used for the Reed-Muller sensing matrix while DFT is used in the case of the chirp sensing matrix.

some wavelet basis) follow a decay pattern. This means that for each location in the signal we know the probability of having a non-zero value. In this case the sensing matrix from the second order Reed-Muller sequences can be constructed in the following way. We start assigning vectors generated from  $P$  matrices in the Kerdock or  $DG(p, 0)$  set to the columns of the sensing matrix that correspond to the signal locations that are most likely to be non-empty. Once the Kerdock set is exhausted we take vectors that come from  $DG(p, 1)$ ,  $DG(p, 2)$ , and so on following the hierarchy given in (2.8). Alternatively, without any loss of generality, one can assume that the signal elements are arranged in ascending order. Then the columns of the sensing matrix appear in the DG order, i.e., the columns from the Kerdock set come first, followed by the  $DG(p, 1)$  and so on. So

$$\Phi = [\phi_{K_1} \cdots \phi_{K_n} \phi_{DG(p,1)_1} \phi_{DG(p,1)_2} \cdots \phi_{DG(p,2)_1} \cdots].$$

This is outlined in the following algorithm. Note that in the algorithm given below,  $f$  is the probability density function of the nonzero (or active) locations. It is assumed to be a decreasing function, which corresponds to the design of the DG matrix.

Input:  $y, f, \Phi = [\phi_{K_1} \cdots \phi_{K_n} \phi_{DG(p,1)_1} \phi_{DG(p,1)_2} \cdots \phi_{DG(p,2)_1} \cdots]$

Output:  $\tilde{z}$

Initially, the residual  $y_0 = y$ .

1. Detection: From  $w(t, \ell) = (f(\ell))^{1/3} \text{DFT}_n \left\{ y_0(\ell) \overline{v_t(\ell)} \right\}$ ,  $t = 1, 2, \dots$ , where DFT is DHT in the RM case and  $v_t$  is the first column of  $U_t$ , update  $\Gamma = \Gamma \cup \{\text{locations associated with } d \text{ largest } |w(t, \ell)|\}$ . Let  $A = \Phi \Big|_{\Gamma}$ .
2. Least Squares:  $\tilde{z} = \underset{z}{\text{argmin}} \|y - Az\|_2$ .
3. Define  $y_0 = y - A\tilde{z}$ . Repeat 1.-3. until  $\|y_0\|_2$  is sufficiently small.

Some preliminary results have been obtained by using the above method and are given in Table 1. According to the rule of thumb, as given in (2.3), if we have a signal of size  $2^{21}$  and we wish to recover this using  $2^6$  measurements, then the sparsity should satisfy  $k \leq 3$ . However, our results in Table 1 indicate that with our algorithm, using the DG order, we are able to recover signals with sparsity greater than 3. This is part of ongoing and future work.

Table 1: The density function used is  $f(j) = \frac{1}{1-a^N} (-\ln(a) a^j)$ . Signal size is  $N = 2^{21}$ , sparsity  $k = 6, 7$ , measurement size or no. of rows of the sensing matrix =  $n = 2^6 = 64$ . The magnitude of the signal entries are either ones or randn (normally distributed pseudorandom).

No. of nonzero entries	$a$	magnitude	No. of successes out of 100 independent trials
6	0.98	ones	99/100
6	0.99	ones	98/100
7	0.98	ones	96/100
7	0.99	ones	90/100
6	0.98	randn	97/100
6	0.99	randn	96/100
7	0.98	randn	99/100
7	0.99	randn	97/100

### 2.3 Choosing the best Daubechies wavelet basis

It is commonly believed that the Haar wavelet (D2) works best for 2D images in compressed sensing. This is because many images chosen for experiments are relatively small, i.e., at most  $128 \times 128$ . In this paper, we are interested in higher-resolution images, such as  $512 \times 512$  and  $1024 \times 1024$ . The degree of smoothness in an image loosely determines the support width of polynomials. For an image scene, if the image resolution is large, the degree of smoothness becomes large. Therefore, an appropriate support width should be large.

Figure 2 shows experiments of chirp compressed sensing using different Daubechies wavelet transforms. The vertical axis is the reconstruction SNR in dB using chirp compressed sensing. The horizontal axis shows which Daubechies wavelet  $DN$  (or  $db\frac{N}{2}$  in Matlab) is used as the sparsifying domain. The level of decompositions is fixed here and is equal to 4 levels (see explanation about the levels in Section 2.2). In particular, D2 (or db1 in Matlab) is the Haar wavelet. The 2D images are not sparsified. The number of measurements taken is 25% of the number of pixels. In all three plots, the reconstruction SNR is optimal around D16. More strictly, the largest reconstruction SNR is within 0.5 dB from the reconstruction error at D16. The decay is faster in the vessel image (b) than in the images (a) and (c), since the vessel image is sharper than the brain and man images and contains more edges. See Table A.1 in Appendix for the precise values. Therefore, we suggest using D16 wavelet basis in deterministic compressed sensing. In particular, instead of the Haar basis our results are obtained using D16 basis.

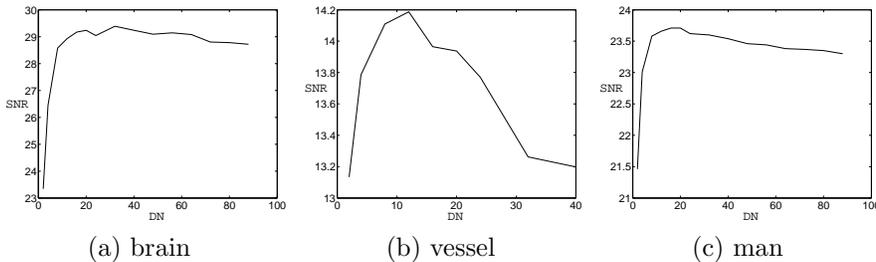


Figure 2: Reconstruction SNR (in dB) of chirp compressed sensing with different Daubechies wavelet  $DN$  (or  $db\frac{N}{2}$  in Matlab) and 4 levels of wavelet decomposition. The 2D images from Figure 1 are not sparsified. The number of measurements taken is 25% of the number of pixels.

### 2.4 Number of levels for wavelet decomposition

In this subsection we investigate how many levels of decomposition should be used in the reconstruction of large images. Figure 3 shows experiments of chirp compressed sensing using D16 with different levels of decomposition. The vertical axis is the reconstruction SNR in dB using chirp compressed sensing. The horizontal axis is the number of levels. In (a) and (c), 3 levels give the best reconstruction SNR. However, we observe that the sparsified images with 3 levels lost a lot more details than with higher levels, when compared to the original un-sparsified images. Therefore, we do not choose 3 levels. The reconstruction SNR from level 4 to 7 is within 0.1 dB range. For experiments in the next section, we thus choose 4 levels. See Table A.2 in Appendix for the precise values.

## 3 Stability of the Deterministic Compressed Sensing Algorithms

By the observation in the previous section, we use Daubechies D16 with 4 levels of wavelet decomposition in all of the following experiments. To study the stability of deterministic compressed sensing using chirp and Reed-Muller sequences for 2D images under noise, we consider five cases. In the first three cases, the noise is in the sparse signal, before the transform  $\Phi$  is taken. In the last two cases, the noise is added after the transform. In what follows we choose noise to be iid Gaussian with mean zero and variance  $2\sigma^2$ . The following are the cases:

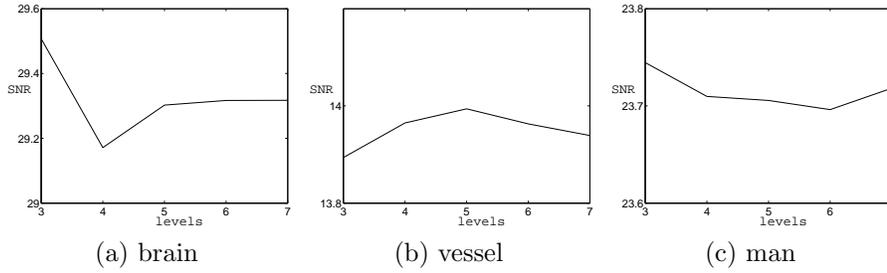


Figure 3: Reconstruction SNR (in dB) of chirp compressed sensing with Daubechies D16 wavelet and different number of levels of wavelet decomposition. The 2D images are not sparsified. The number of measurements taken is 25% of the number of pixels.

1. The 2D images are sparsified with pre-defined sparsity in the wavelet domain. The wavelet coefficients  $x_k$  is  $k$ -sparse and there is a noise in the wavelet domain. The measurement is

$$y = \Phi(x_k + \mu). \quad (3.1)$$

2. The wavelet coefficients, denoted by  $x_0$ , is not sparsified, but assumed to be compressible. The noise is again in the wavelet domain. The measurement is

$$y = \Phi(x_0 + \mu). \quad (3.2)$$

3. The wavelet coefficients  $x_k$  is  $k$ -sparse and the noise  $\mu_{k^c}$ , in the wavelet domain, is only non-zero outside the support of  $x_k$ . Here  $k^c$  stands for the complement locations of the non-zero wavelet coefficients, i.e., the noise is added to the zero wavelet coefficients. Therefore, the nonzero coefficients of the signal are not perturbed. The measurement is

$$y = \Phi(x_k + \mu_{k^c}). \quad (3.3)$$

4. Noise occurs at the step of measurements and is added after the signal  $x_k$  is transformed by the CS matrix  $\Phi$ :

$$y = \Phi x_k + \mu. \quad (3.4)$$

5. Noise occurs at the step of measurements and is added after the compressible non-sparse signal  $x_0$  is transformed by  $\Phi$ :

$$y = \Phi x_0 + \mu. \quad (3.5)$$

We test our algorithm on 3 images, see Figure 1, which we refer to as brain, vessel, and man. The first two are typical MRI images used in the medical community, and the third is a high resolution natural image. We reconstruct these images using chirp and Reed-Muller sensing matrices and compare with the reconstruction by noiselets (for example, see [31]). The noiselet method takes random noiselet measurements and then reconstructs by  $\ell_1$  minimization whereas our chirp and Reed-Muller algorithms use deterministic measurements and the reconstruction is by a least squares method.

The reconstruction error is defined as:

$$\text{Error(dB)} = 10 \log_{10} \left[ \frac{\|x_{\text{actual}} - x_{\text{reconstructed}}\|^2}{\|x_{\text{actual}}\|^2} \right].$$

The negative of the above error is known as the *signal-to-noise ratio* (SNR). Tables 2-6 show the reconstructed SNR of measurements taken according to (3.1)-(3.5), respectively. Several noise levels were chosen for the experiment; and specifically, the standard deviation  $\sigma$  of Gaussian noise with zero mean are  $\sigma = 0, 0.01, 0.05, 0.1, \text{ and } 0.2$ . These values are chosen for comparison purposes, for example, with Tables 1 and 2 in [12].

For Tables 1-3, the third columns are the SNR of the noisy images compared to the images before noise is added to the wavelet coefficients. If noise is present in the signal, all reconstructed SNR are smaller than the reference SNR via noiselet measurements and  $\ell_1$  minimization. The reconstructed SNR by chirp has *the highest value (best)* in all cases and stays very close to the reference SNR. The reconstructed SNR by RM is higher than noiselets when the noise level is small.

When the noise level is large, it seems that the noiselets slightly outperform RM even though the discussion in Section 1 assures that the deterministic compressed sensing matrices comprised of chirps and Reed-Muller sequences have advantage over random matrices. The reason is that the deterministic sensing matrices here satisfy StRIP for sparse signals whose nonzero locations follow a uniform distribution, and the wavelet coefficients do not follow a uniform distribution. Therefore, better performance is not guaranteed anymore. Nevertheless, all results show that the deterministic algorithms are as stable as the reconstruction using noiselets and  $\ell_1$  minimization.

For tables 5 and 6, the third column gives the SNR of the noisy measurement compared to the clean measurement. The reconstructed SNR is calculated the same way as in Tables 1-3, comparing the reconstructed image to the reference image. Therefore, the SNR in the third columns is in a different domain, or differs by the  $\Phi$  transform. All results show that the reconstruction by chirp or RM performs *significantly better* than noiselets. Note that in Tables 1, 3, and 4, where there is no noise, the reconstruction is very accurate, above 100 dB.

Table 2: The 2D images are first sparsified in the domain of Daubechies wavelet D16 (db8 in Matlab) with 4 levels of decomposition. The image, its size, and the pre-determined sparsity are shown in the first column. Gaussian noise with zero mean and variance  $\sigma^2$  (or standard deviation  $\sigma$ ) is added to the wavelet coefficients of the sparsified image; see (3.1). The third column shows the SNR in dB compared to the noise-free sparsified image. Columns 4 - 6 show the reconstruction SNR in dB using noiselets, chirp, and Reed-Muller CS algorithms, respectively. The number of measurements taken is 25% of the number of pixels.

image, size, sparsity	stan. dev. $\sigma$	SNR	noiselets	Chirp	RM
brain, $512 \times 512$ , 7%	0	–	29.8	127.0	121.6
	0.01	31.6	24.0	31.0	27.4
	0.05	17.7	15.4	17.1	16.1
	0.1	11.9	10.8	11.6	10.4
	0.2	6.5	6.5	6.8	5.6
vessel, $512 \times 512$ , 5%	0	–	37.1	126.2	121.0
	0.01	21.6	15.7	18.1	15.2
	0.05	8.9	7.0	7.7	6.8
	0.1	3.2	3.6	4.2	3.0
	0.2	-2.5	1.2	2.1	0.9
man, $1024 \times 1024$ , 2.38%	0	–	43.3	122.6	118.1
	0.01	32.5	30.0	32.2	30.1
	0.05	18.7	16.8	17.9	16.7
	0.1	12.9	11.8	12.5	11.3
	0.2	7.2	7.2	7.5	6.3

## 4 Discussions

In this work, we show that the algorithms introduced in [3] are stable even if the signal is not exactly sparse and there is some noise in the measurements or in the signal. We also address several practical concerns about compressed sensing with images.

First of all, the Haar wavelet is commonly used as the sparsifying transform domain. We compare the performance of standard Daubechies DN wavelets and suggest using Daubechies D16 as it performs the best for larger resolutions images, such as  $512 \times 512$  or  $1024 \times 1024$  pixels.

Secondly, we explain the proper selection of the number of levels for wavelet decomposition. As discussed in Section 2, levels 4-7 are optimal for wavelet decomposition, and for optimal computation time we work with level 4.

Finally, our experiments show that the algorithms are *indeed stable* if the signal is compressible and there is noise in either measurements or the signal. Furthermore, our chirp and RM compressed sensing

Table 3: In this table, the 2D images are not sparsified. Gaussian noise with zero mean and standard deviation  $\sigma$  is added to all 4 levels of D16 wavelet coefficients decomposition of the non-sparsified image; see (3.2). The third column shows the SNR in dB of the noisy images compared to the original image. Columns 4 - 6 show the reconstruction SNR in dB using noiselets, chirp, and Reed-Muller CS algorithms, respectively. The number of measurements taken is 25% of the number of pixels.

image, size	stan. dev. $\sigma$	SNR	noiselets	Chirp	RM
brain, $512 \times 512$	0	-	23.4	28.4	25.7
	0.01	31.6	22.2	25.7	24.6
	0.05	17.7	15.2	17.0	15.9
	0.1	12.0	10.8	11.6	10.4
	0.2	6.5	6.5	6.8	5.6
vessel, $512 \times 512$	0	-	12.0	14.1	13.4
	0.01	21.8	11.4	12.8	12.7
	0.05	9.0	6.7	7.5	6.6
	0.1	3.3	3.6	4.1	3.0
	0.2	-2.4	1.2	2.0	0.8
man, $1024 \times 1024$	0	-	20.0	23.2	22.6
	0.01	32.5	19.8	22.8	22.0
	0.05	18.8	15.2	17.3	16.1
	0.1	12.9	11.3	12.4	11.1
	0.2	7.3	7.1	7.5	6.3

Table 4: The 2D images are first sparsified in the Daubechies wavelet D16 domain with 4 levels of decomposition. The pre-determined sparsity for each image is shown in the first column. The Gaussian noise with zero mean and standard deviation  $\sigma$  is added only to the non-zero wavelet coefficients of the sparsified image; see (3.3). The third column shows the SNR in dB of the images with added noise compared to the noise-free sparsified image. Columns 4 - 6 show the reconstruction SNR in dB using noiselets, chirp, and Reed-Muller CS algorithms, respectively. The number of the measurements taken is 25% of the number of pixels.

image, size, sparsity	stan. dev. $\sigma$	SNR	noiselets	Chirp	RM
brain, $512 \times 512$ , 7%	0	-	29.7	127.0	121.6
	0.01	31.9	24.2	31.1	27.1
	0.05	18.0	15.5	17.1	16.1
	0.1	12.3	11.0	11.6	10.4
	0.2	6.8	6.7	6.8	5.6
vessel, $512 \times 512$ , 5%	0	-	31.7	126.2	121.0
	0.01	21.3	15.7	17.9	15.2
	0.05	8.6	7.0	7.7	6.7
	0.1	3.0	3.6	4.1	3.0
	0.2	-2.8	1.2	1.9	0.8
man, $1024 \times 1024$ , 2.38%	0	-	43.1	122.6	118.1
	0.01	35.5	29.9	32.2	30.1
	0.05	18.7	16.8	17.9	16.8
	0.1	12.8	11.8	12.6	11.3
	0.2	7.2	7.2	7.5	6.3

Table 5: The 2D images are first sparsified in the Daubechies wavelet D16 domain with 4 levels of decomposition. The pre-determined sparsity for each image is shown in the first column. Measurements are taken from the D16 domain. The number of measurement taken is 25% of the number of pixels. Gaussian noise with zero mean and standard deviation  $\sigma$  is added to the measurements of the sparsified image; see (3.4). The third column shows the SNR in dB of the noisy measurements compared to the noise-free measurements. Columns 4 - 6 show the reconstruction SNR in dB using noiselets, chirp, and Reed-Muller CS algorithms, respectively.

image, size, sparsity	stan. dev. $\sigma$	SNR	noiselets	Chirp	RM
brain $512 \times 512$ , 7%	0	-	29.7	127.0	121.6
	0.01	31.6	24.7	46.4	45.6
	0.05	17.6	16.8	27.5	26.8
	0.1	11.5	12.6	22.0	21.4
	0.2	5.5	8.1	15.5	14.9
vessel $512 \times 512$ , 5%	0	-	37.1	126.2	121.0
	0.01	20.6	16.4	32.8	32.7
	0.05	6.6	7.5	14.6	15.0
	0.1	0.5	2.8	10.9	10.0
	0.2	-5.5	-2.6	5.2	4.5
man $1024 \times 1024$ , 2.38%	0	-	43.1	122.6	118.1
	0.01	32.4	31.4	46.9	46.5
	0.05	18.4	18.3	27.0	26.0
	0.1	12.4	13.5	22.0	21.0
	0.2	6.4	8.6	16.0	15.5

Table 6: The 2D images are not sparsified in this experiment. Measurements are taken from the D16 domain with 4 levels of decomposition. The number of measurements taken is 25% of the number of pixels. Then, Gaussian noise with zero mean and standard deviation  $\sigma$  is added to the measurements of non-sparsified image; see (3.5). The third column shows the SNR in dB of the noisy measurements compared to the noise-free measurements. Columns 4 - 6 show the reconstruction SNR in dB using noiselets, chirp, and Reed-Muller CS algorithms, respectively.

image, size	stan. dev. $\sigma$	SNR	noiselets	Chirp	RM
brain $512 \times 512$	0	-	23.4	28.4	25.7
	0.01	31.6	22.4	28.0	25.7
	0.05	17.6	16.5	25.2	24.9
	0.1	11.5	12.5	21.3	20.9
	0.2	5.5	8.0	15.3	14.7
vessel $512 \times 512$	0	-	12.0	14.1	13.4
	0.01	20.6	11.5	13.8	13.4
	0.05	6.7	6.9	12.4	12.3
	0.1	0.7	2.6	9.9	9.1
	0.2	-5.3	-2.6	4.9	4.3
man $1024 \times 1024$	0	-	20.0	23.2	22.6
	0.01	32.4	19.6	23.2	22.5
	0.05	18.4	16.2	22.5	21.7
	0.1	12.4	12.7	20.2	19.4
	0.2	6.4	8.4	15.6	15.1

reconstruction algorithms outperform the standard compressed sensing reconstruction using random matrices and  $\ell_1$  minimization. We have also done some analysis on how our efficient reconstruction algorithm can be adapted by using the nesting of the Delsarte-Goethals sets of the Reed-Muller codes to deal with very large images. Currently, we are trying to use this to improve the performance of our deterministic compressed sensing algorithms.

#### ACKNOWLEDGEMENT

This work is partially supported by NSF-DMS grant #0652833, NSF-DUE #0633033, and ONR-BRC grant #N00014-08-1-1110. The authors are immensely grateful to Douglas Cochran for his invaluable guidance during the course of this work. The authors also wish to express their gratitude to Stephen Howard for many useful discussions.

## References

- [1] K. Ni, P. Mahanti, S. Datta, S. Roudenko, D. Cochran, Image reconstruction by deterministic compressive sensing with chirp matrices, in: Proceedings of SPIE, Vol. 7497, 2009.
- [2] K. Ni, S. Datta, P. Mahanti, S. Roudenko, D. Cochran, Using reed-muller codes as deterministic compressive sensing matrices for image reconstruction, in: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Dallas, 2010.
- [3] K. Ni, S. Datta, P. Mahanti, S. Roudenko, D. Cochran, Efficient deterministic compressed sensing for images with chirps and reed-muller sequences, *SIAM Journal in Imaging and Sciences* 4 (3).
- [4] E. Candès, Compressive sampling, in: Proceedings of the International Congress of Mathematicians, Vol. III, European Mathematical Society, Madrid, 2006, pp. 1433–1452.
- [5] E. Candès, J. Romberg, T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, *IEEE Transactions on Information Theory* 52 (2) (2006) 489 – 509.
- [6] D. L. Donoho, Compressed sensing, *IEEE Transactions on Information Theory* 52 (4) (2006) 1289 – 1306.
- [7] M. Lustig, D. Donoho, J. M. Pauly, Sparse MRI: The application of compressed sensing for rapid MR imaging, *Magnetic Resonance in Medicine* 58 (6) (2007) 1182 – 1195.
- [8] W. U. Bajwa, J. Haupt, A. M. Sayeed, R. Nowak, Compressed channel sensing: A new approach to estimating sparse multipath channels, in: Proc. IEEE, 2010.
- [9] M. Mishali, Y. C. Eldar, Blind multiband signal reconstruction: Compressed sensing for analog signals, *IEEE Trans. on Signal Processing* 57 (30) 993–1009.
- [10] T. Lin, F. J. Herrmann, Compressed wavefield extrapolation, *Geophysics* 72 (5) (2007) SM77–SM93.
- [11] E. Candès, T. Tao, Near optimal signal recovery from random projections: Universal encoding strategies?, *IEEE Transactions on Information Theory* 52 (12) (2006) 5406 – 5425.
- [12] E. Candès, J. Romberg, T. Tao, Stable signal recovery from incomplete and inaccurate measurements, *Communications on Pure and Applied Mathematics* 59 (8) (2006) 1207 – 1223.
- [13] E. J. Candès, J. K. Romberg, Signal recovery from random projections, in: Proceedings of SPIE, Vol. 5674, 2005, pp. 76 – 86.
- [14] R. DeVore, Deterministic constructions of compressed sensing matrices, *Journal of Complexity* 23 (2007) 918 – 925.
- [15] L. Applebaum, S. D. Howard, S. Searle, R. Calderbank, Chirp sensing codes: Deterministic compressed sensing measurements for fast recovery, *Applied and Computational Harmonic Analysis* 26 (2009) 283 – 290.
- [16] S. D. Howard, A. R. Calderbank, S. J. Searle, A fast reconstruction algorithm for deterministic compressive sensing using second order Reed-Muller codes, in: Proceedings of the Conference on Information Sciences and Systems, Princeton, New Jersey, 2008, pp. 11 – 15.
- [17] R. Calderbank, S. Howard, S. Jafarpour, Construction of a large class of deterministic matrices that satisfy a statistical isometry property, *IEEE Journal on Selected Topics in Signal Processing* 29 (4).
- [18] P. Indyk, Near-optimal sparse recovery in the  $\ell_1$  norm, in: 49th Annual IEEE Symposium on Foundations of Computer Science, 2008, pp. 199–207.

- [19] E. Candès, J. Romberg, Sparsity and incoherence in compressive sampling, *Inverse Problems* 23 (3) (2007) 969 – 985.
- [20] D. Baron, M. F. Duarte, S. Sarvotham, M. B. Wakin, R. G. Baraniuk, An information-theoretic approach to distributed compressed sensing, in: *Proceedings of the 43rd Allerton Conference on Communications, Control, and Computing*, 2005.
- [21] W. O. Alltop, Complex sequences with low periodic correlations, *IEEE Transactions on Information Theory* 26 (3) (1980) 350 – 354.
- [22] K. J. Horadam, *Hadamard Matrices and their Applications*, Princeton University Press, 2007.
- [23] J. A. Roger Hammons, P. V. Kumar, A. R. Calderbank, N. J. A. Sloane, P. Solé, The  $\mathbb{Z}_4$ -linearity of Kerdock, Preparata, Goethals, and related codes, *IEEE Transactions on Information Theory* 40 (2) (1994) 301 – 319.
- [24] R. Calderbank, personal communication (2009).
- [25] A. M. Kerdock, A class of low rate nonlinear binary codes, *Information and Control* 20 (2) (1972) 182 – 187.
- [26] C. C. Paige, M. A. Saunders, LSQR: An algorithm for sparse linear equations and sparse least squares, *ACM Transactions on Mathematical Software* 8 (1) (1982) 43 – 71.
- [27] T. Faktor, Y. C. Eldar, M. Elad, Exploiting statistical dependencies in sparse representations for signal recovery, preprint (2010).
- [28] C. La, M. N. Do, Tree-based orthogonal matching pursuit algorithm for signal reconstruction, in: *ICIP, Atlanta, GA*, 2006.
- [29] Y. C. Eldar, M. Mishali, Robust recovery of signals from a structured union of subspaces, *IEEE Transactions on Information Theory* 55 (2009) 5302 – 5316.
- [30] Y. C. Eldar, P. Kuppinger, H. Blcskei, Compressed sensing of block-sparse signals: Uncertainty relations and efficient recovery (June 2010).
- [31] J. Romberg, Imaging via compressive sampling, *IEEE Signal Processing Magazine* 25 (2) (2008) 14 – 20.

## Appendix A: SNR values for Figures 2 and 3

Please see Tables A.1 and A.2.

Table A.1: The table below shows reconstruction SNR of chirp compressed sensing with different Daubechies wavelet  $DN$  ( $=\text{db}\frac{N}{2}$  in matlab) and 4 levels of decomposition. Specifically, D2 is the Haar wavelet. The 2D images are not sparsified. Each column shows the reconstruction SNR in dB with the indicated  $DN$ . The number of measurements taken is 25% of the number of pixels. This table is plotted in Figure 2.

	D2	D4	D8	D12	D16	D20	D24	D32	D40
brain	23.34	26.45	28.58	28.93	29.17	29.24	29.04	29.39	29.24
vessel	13.13	13.79	14.11	14.19	13.97	13.94	13.77	13.26	13.12
man	21.46	23.01	23.58	23.66	23.71	23.71	23.62	23.60	23.54

	D48	D56	D64	D72	D80	D88
brain	29.09	29.14	29.08	28.80	28.78	28.72
man	23.46	23.44	23.38	23.37	23.35	23.30

Table A.2: The table below shows reconstruction error of chirp compressed sensing with Daubechies D16 wavelet and various levels of decomposition ranging from 3 to 7. The 2D images are not sparsified. Each column shows the reconstruction SNR in dB with the indicated number of levels. The number measurement taken is 25% of the number of pixels. This table is plotted in Figure 3.

	3	4	5	6	7
brain	29.51	29.17	29.30	29.31	29.32
vessel	13.89	13.97	13.99	13.96	13.94
man	23.74	23.71	23.71	23.70	23.72