# COMPRESSED SENSING USING REED-SOLOMON AND Q-ARY LDPC CODES

**Graduate Student: Kristin M. Jagiello**
**Faculty Advisors: William E. Ryan, Michael W. Marcellin and Nathan A. Goodman**
**ECE Dept., University of Arizona, Tucson, AZ 85721**
**(kristinj, ryan, marcellin)@ece.arizona.edu**

### ABSTRACT

We consider the use of Reed-Solomon (RS) and $q$-ary LDPC codes for compressed sensing of sparse signals. Signals sensed using the RS parity-check matrix are recovered using Berlekamp-Massey and those sensed using the LDPC parity-check matrix are recovered using majority-logic decoding. Results are presented for both types of sensing. In addition, a hardware architecture is discussed.

## 1. INTRODUCTION

In many applications the signal of interest is sparse in some domain. Examples include imaging, wideband communications, and radar. Traditional sensing requires the acquisition of signal samples corresponding to Nyquist-rate sampling. This is often followed by some type of compression. Compressive sensing (CS) is an emerging technology in which the signal is compressed during the acquisition stage, allowing sampling rates much slower than the Nyquist rate.

Recently there has been some research aimed at using forward error correction (FEC) decoders for compressed sensing (CS) reconstruction [1], [2], [3]. The main parallel between CS and FEC is that in both cases the decoder (reconstruction algorithm) seeks a sparse vector as its output. For CS reconstruction, this is the coefficient representation of the signal in the sparse domain. For FEC decoding, this is the error vector. A great deal of work has been done over the last several decades in the area of FEC. As a result, many practical FEC decoders exist. Clearly it would be beneficial to extend this work to the area of CS, particularly since the widely used CS reconstruction algorithms are very complex. For this reason, we examine the applicability of $q$-ary FEC decoders to CS.

Section 2 presents an overview of compressive sensing. Section 3 contains the system model used for the simulations and a discussion on a receiver architecture. Section 4 provides details of reconstruction methods used. Section 5 presents simulation results and Section 6 concludes the paper.

## 2. COMPRESSIVE SENSING OVERVIEW

CS is a method of acquiring signals at a much lower sampling rate than the Nyquist rate. It will work under the condition that the signals of interest are sparse in some domain, for example the Fourier domain or the wavelet domain. Most one-dimensional and two-dimensional signals are sparse or approximately sparse. The second requirement is that the sensing matrix or waveforms be incoherent with the signals. Given that these two conditions are satisfied, it is possible to recover the original signal from $m < n$ measurements, where $n$ is the dimensionality of the space in which the signal lies.

Let $\mathbf{f} \in \mathcal{C}^n$ be a complex signal that is sparse in some domain and let $\{\boldsymbol{\phi}_i\}$ be basis elements such that the columns of the matrix $\mathbf{\Phi} = [\boldsymbol{\phi}_0,\ \boldsymbol{\phi}_1,\ ...,\ \boldsymbol{\phi}_{n-1}]$ are an orthonormal basis. Then the signal can be expressed as

$$\mathbf{f} = \sum_{i=0}^{n-1} v_i \boldsymbol{\phi}_i \tag{1}$$

where $\mathbf{v}$ is the sparse (or coefficient) representation of the signal. We assume that the number of non-zero coefficients in $\mathbf{v}$ is at most $k$. The length-$m$ measurement vector is computed as $\mathbf{y} = \mathbf{Mf}$ where $\mathbf{M}$ is called a measurement matrix. Again, $m < n$ so in general solving for $\mathbf{f}$ from $\mathbf{y}$ is not possible because the system of equations is underdetermined. However, when $\mathbf{f}$ is sparse in some domain, $\mathbf{f} = \mathbf{\Phi v}$ where $\mathbf{v}$ has $k \ll n$ non-zero components. We then have $\mathbf{y} = \mathbf{Mf} = \mathbf{M\Phi v} = \mathbf{M'v}$. If $k \leq m$ and we know the locations of the non-zero coefficients of $\mathbf{v}$ we could solve for the values using classical techniques such as least squares. We do not, however, know the locations. All we have is the prior knowledge of the signals' sparsity.

The relatively new theory of CS, [4], [5] allows us to solve this problem via $L_1$ optimization:

$$\min \parallel \mathbf{v} \parallel_1 \ \text{ subject to } \mathbf{y} = \mathbf{M'v}, \tag{2}$$

where $\parallel \mathbf{v} \parallel_1 = \sum_i |v_i|$. It can be shown that, provided $\mathbf{M'}$ satisfies the restricted isometry property (RIP) [5], $\mathbf{v}$ can be recovered exactly. Loosely, the RIP states that every possible combination of $k$ columns of $\mathbf{M'}$ will be an orthogonal set if $\mathbf{M'}$ satisfies the RIP. If the bases for $\mathbf{M}$ and $\mathbf{G}$ are *incoherent*, then the RIP will be satisfied. The *coherence* between $\mathbf{M}$ and $\mathbf{G}$ is defined to be the maximum-magnitude of the inner products between any two basis elements of the two matrices. Small coherence values correspond to greater incoherence.

## 3. SYSTEM MODEL

Figure 1 presents the system diagram we have adopted for our computer experiments involving the compressive sensing of a sum of $k$ complex sinusoids. This signal is expressed as

$$\mathbf{f} = \sum_{i=0}^{n-1} v_i \boldsymbol{\phi}_i = \sum_{j=0}^{k-1} a_j e^{j\omega_j t} \tag{3}$$

where $a_j$ are the non-zero elements of $\mathbf{v}$ and $\omega_j$ are the corresponding frequencies. The values of $a_j$ and $\omega_j$ are randomly selected and $t$ is a length-$n$ discrete-time vector.

The discrete-time, oversampled signal, $\mathbf{f}$, is the FFT processor input. The real and imaginary parts of the signal are then independently quantized to 8-bit integers. Finally, the quantized signal is sensed with a $q$-ary measurement matrix, $q = 256$. We consider two different measurement matrices: a 256-ary RS parity-check matrix and a 256-ary LDPC parity-check matrix. Reconstruction for signals sensed with the RS matrix is performed via the Berlekamp-Massey (BM) algorithm. Signals sensed with the LDPC parity-check matrix are reconstructed using the one-step majority-logic decoding algorithm from [6].

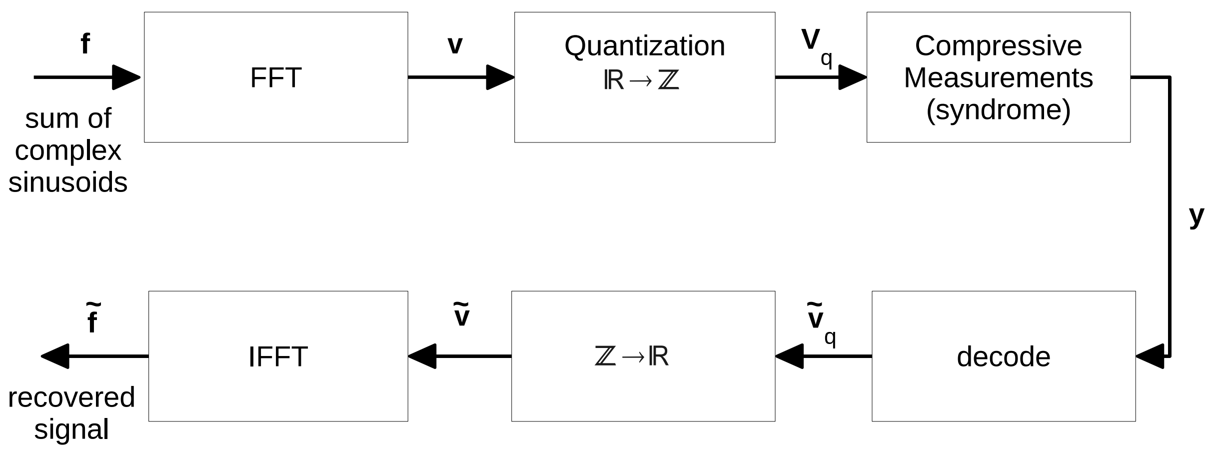


**Figure 1.** A block diagram of the system model.

In order for the system in Figure 1 to be practicable, it is necessary to project the original analog signal onto the Fourier basis, that is, to represent the signal in the sparse domain. One possible solution is presented in Figure 2. An analog signal $f(t)$ is the input to a bank of $n$ parallel correlators. After $T$ seconds, the output of each correlator is sampled, quantized, and then multiplied by the appropriate sensing matrix.

## 4. RECONSTRUCTION METHODS

Let $\mathbf{v}$ be the length-$n$ sparse representation of the signal after quantization. We drop the subscript $q$ for convenience. Let $\mathbf{y}$ be the length-$m$ measurement vector computed as

$$\mathbf{y} = \mathbf{H}\mathbf{v} \tag{4}$$

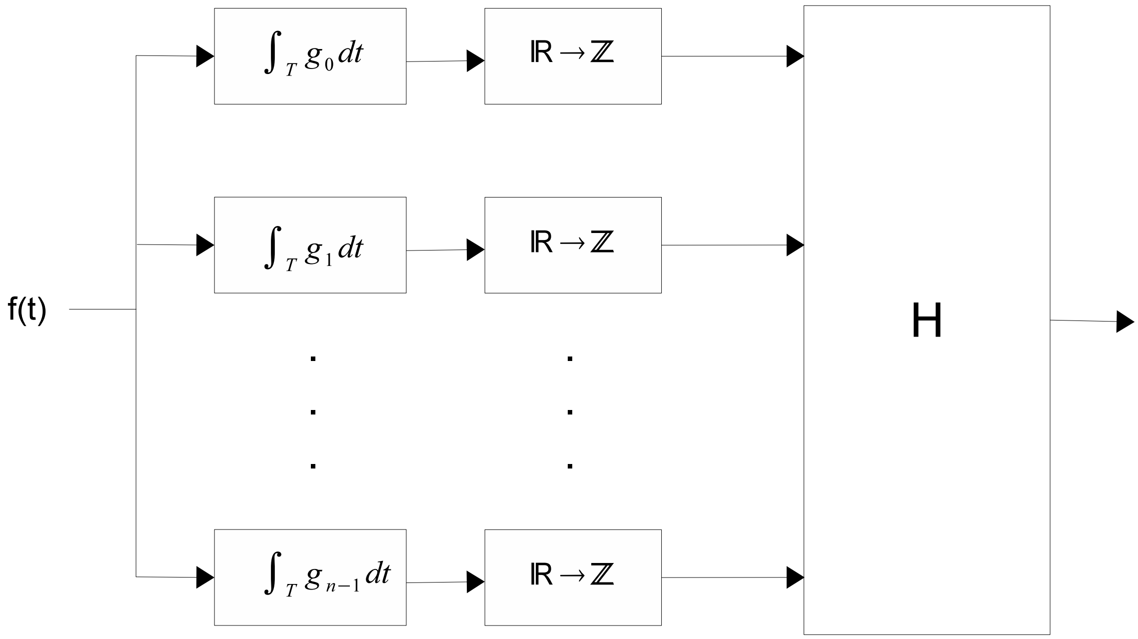


**Figure 2.** A block diagram of the parallel receiver configuration.

where $\mathbf{H}$ is the parity-check matrix of an $m \times n$ FEC code over GF(256), the size-256 Galois field. Note, in the context of FEC, $\mathbf{y}$ is a syndrome. Given $\mathbf{y}$, the objective is to recover $\mathbf{v}$.

### 4.1. Berlekamp-Massey Reconstruction

The RS parity-check matrix, $\mathbf{H}$, for a (255, 223) code was constructed as the $32 \times 255$ matrix over GF(256) given by

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & & 1 \\ \alpha^{m_0} & \alpha^{m_0+1} & \cdots & \alpha^{m_0+d-2} \\ \alpha^{2m_0} & \alpha^{2(m_0+1)} & & \alpha^{2(m_0+d-2)} \\ \vdots & \vdots & & \vdots \\ \alpha^{(n-1)m_0} & \alpha^{(n-1)(m_0+1)} & \cdots & \alpha^{(n-1)(m_0+d-2)} \end{bmatrix}, \tag{5}$$

where $\alpha$ is a primitive element in GF(256) [6]. The value of $m_0$ may be selected as any number from 0 to 254, but we chose it to be 112 (in coding this is the value corresponding to the generator polynomial that is its own reciprocal). The minimum distance, $d$, for this code is 33.

RS codes can be decoded using the Berlekamp-Massey algorithm. It is an iterative algorithm capable of decoding to the minimum distance of the RS code, that is, any pattern of $t$ or fewer errors are correctable, where $t$ is the largest integer satisfying $2t < d$. If the measurement vector (syndrome) is zero, then the recovered signal is zero and decoding stops. If the measurement vector is non-zero, the Berlekamp-Massey algorithm can be used to find coefficient locations and values provided the number of non-zero elements does not exceed $t$.

Details of the BM algorithm can be found in many textbooks; see [6]. Briefly, the algorithm uses the syndrome vector to find an error locator polynomial and from this the error locations. Once the locations have been determined, the error values are found.

### 4.2. One-Step Majority-Logic Reconstruction

The LDPC parity-check matrix **H** was obtained by selecting a column weight $\gamma$ and then randomly generating non-zero elements from GF(256) to place along the columns of **H**. An additional constraint placed on the matrix was that no two columns could overlap by more than one non-zero element. This constraint ensures that all measurements are orthogonal to each other [6]. The lower bound for the number of recoverable non-zero coefficients is $\lfloor \gamma/2 \rfloor$, where $\lfloor x \rfloor$ is the integer part of $x$.

The one-step majority-logic decoding (OSMLGD) algorithm was used to recover the sparse signal. A thorough description of majority-logic decoding for Euclidean geometry (EG) codes can be found in [6]. (EG codes are a class of structured LDPC codes.) The algorithm is briefly described below.

The measurement vector $\mathbf{y}$ is computed from (4). Let $j = 1, 2, ..., n-1$ and let $i = 1, 2, ..., m-1$. Then let $N_i$ be the set of indices indicating the positions of the non-zero elements in row $i$. To reconstruct $\mathbf{v}$ from $\mathbf{y}$ it is necessary to construct $n$ sets of measurements (or syndrome elements), one set for each $v_j$. $S_j$ is the set of $\gamma$ measurements that are dependent on $v_j$. Each measurement set is then normalized as

$$\tilde{s}_i = h_{i,j}^{-1} s_i = x_j + h_{i,j}^{-1} \sum_{l \in N_i, l \neq j} h_{i,l} x_l, \tag{6}$$

where $h_{i,l}$ are the non-zero elements from the **H** matrix used to measure the signal. The normalized measurement sets, $\tilde{S}_j$ are used to decode $v_j$. Assume that there are no more than $\lfloor \gamma/2 \rfloor$ non-zero coefficients in $\mathbf{v}$. If there is a clear majority of normalized measurements in $\tilde{S}_j$ that are identical and non-zero, then $v_j$ is assumed to be non-zero and is decoded to the value of $\tilde{s}_i$ that has the majority. If $\lfloor \gamma/2 \rfloor$ normalized check-sums in $\tilde{S}_j$ are zero, then $v_j$ is assumed to be zero [6].

## 5. RESULTS

### 5.1. RS Measurement Matrix

The RS parity-check matrix was used to sense signals having 16 or fewer non-zero frequency-domain components. Figure 3 is a plot of the signal in the frequency domain. The RS parity-check matrix resulted in a compression ratio of eight as the number of samples was reduced from 255 to 32 because the measurment matrix, **H**, is $32 \times 255$. The Berlekamp-Massey algorithm was used to recover the sparse signal. Figure 4 is a plot of the real part of a Berlekamp-Massey reconstructed signal along with the original signal. The only error present in the recovered signal is due to quantization. It is difficult to distinguish between the original signal and its reconstruction.

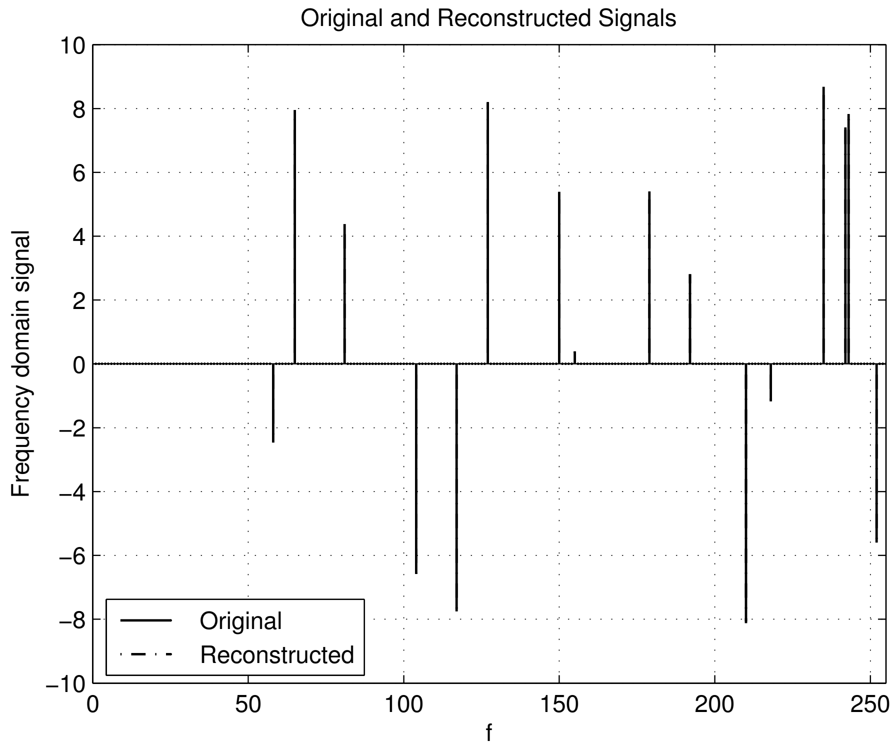


**Figure 3.** The original signal in the frequency domain and its reconstruction via the BM algorithm.

Let $k$ denote the number of non-zero coefficients. Figure 5 is a plot of the average mean-squared difference (MSD) between the original and reconstructed signals as a function of $k$. We define MSD as

$$\text{MSD} = \| \mathbf{f} - \tilde{\mathbf{f}} \|_2^2 = \sum_{i=0}^{n-1} \Big( (\Re\{f_i\} - \Re\{\tilde{f}_i\})^2 + (\Im\{f_i\} - \Im\{\tilde{f}_i\})^2 \Big), \tag{7}$$

where $\Re$ denotes the real part of the signal and $\Im$ denotes the imaginary part. For each choice of $k$, 10,000 signals were simulated. In this plot, we see that the MSD increases slightly for $4 \leq k \leq 16$ corresponding to slightly increasing quantization error. At $k = 16$ the decoder fails completely and the MSD increases by four orders of magnitude to 1.7e-2. The gradual slope seen in the plot is due to the fact that quantization error is only present in the non-zero coefficients. As the number of non-zero coefficients increases, the potential for quantization error increases as well.

The RS code results in good compression as the measured signal is eight times smaller than it would need to be for traditional acquisition. Further, the Berlekamp-Massey algorithm is easy to implement and its complexity is considered low.

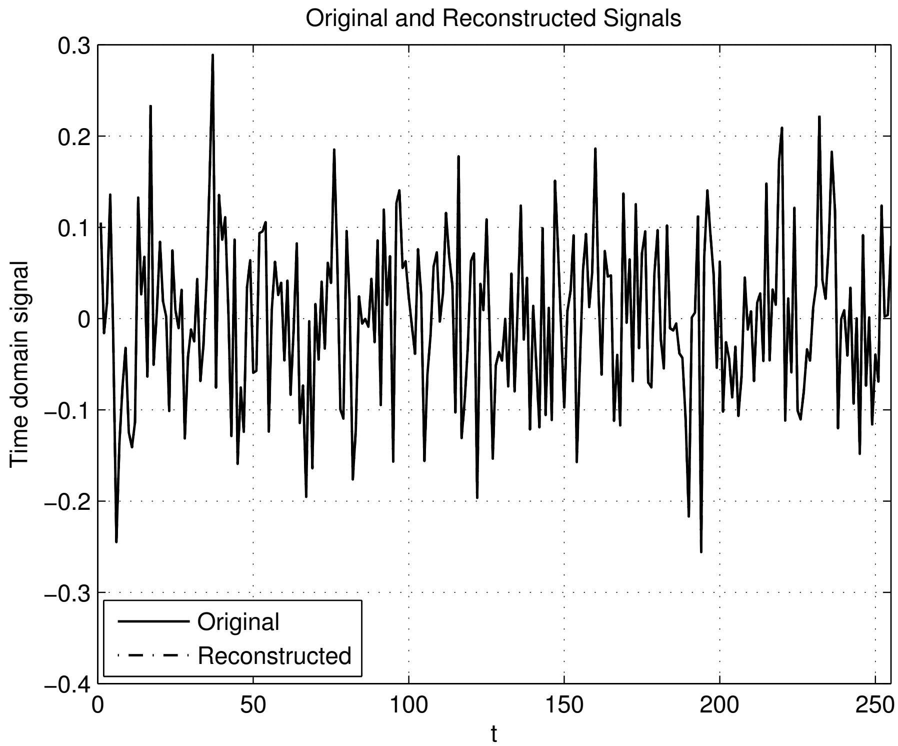


**Figure 4.** The real part of the original signal and its reconstruction via the BM algorithm.

### 5.2. 256-ary LDPC Measurement Matrix

Next we examined a 256-ary LDPC code parity-check matrix to determine its suitability for CS applications. Specifically, we designed a $500 \times 2000$ parity-check matrix with a column weight of eight. The number of samples was reduced from 2000 to 500. For this code the lower bound on the number of recoverable coefficients is $\lfloor 8/2 \rfloor = 4$.

The OSMLGD algorithm was used for recovery. We simulated the recovery capability for varying quantities of non-zero coefficients. 10,000 signals were simulated for each choice of $k$. Figure 6 is a plot of the average number of errors present in the recovered signal as a function of $k$. The lower bound on the number of recoverable coefficients is four, but it can be seen from Figure 6 that the code is able to do better than the lower bound for many signals. For example, signals having six non-zero coefficients results in an error only 0.21% of the time and signals having eight non-zero coefficients result in an error 1.9% of the time.

Figure 7 is a plot of the mean-squared difference as a function of $k$. It can be seen that as the number of coefficients increases past four, the number of errors and, thus, the MSD increases. This is to be expected as the code is only guaranteed to recover four non-zero coefficients. The MSD for the OSMLGD reconstructed signal is lower than that of the BM reconstructed signal. This is due to the fact that the OSMLGD simulations used a length-2000 signal whereas the BM simulations used a length-255 signal. As mentioned previously,

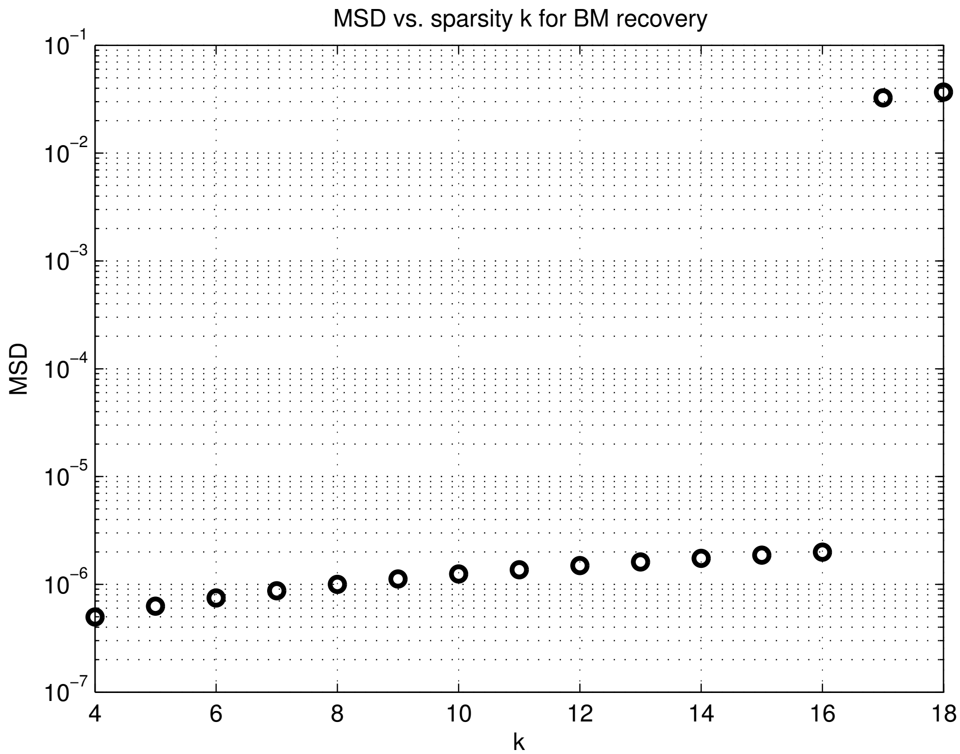


**Figure 5.** A plot of the mean-squared difference as a function of the number of non-zero coefficients in the sparse representation of the signal for BM recovery.

quantization error is only present in the non-zero coefficients. Because of this, the longer code has an average quantization error that is lower.

## 6. CONCLUSION

Traditional signal acquisition methods use Nyquist sampling followed by compression. In [7], Baraniuk names three disadvantages of this method. First, the initial set of samples is of size $n$ even if the signal is sparse and contains a small amount of information. This means a great deal of storage is utilized for a small amount of information. Second, all transform coefficients must be computed for compression. Last, it is necessary to encode the locations of the largest coefficients. CS techniques provide a solution to these inefficiencies.

The most obvious benefit of the FEC recovery algorithms presented here is their simplicity. In addition, they allow for non-adaptive measurements, thus simplifying the encoding process as compared to the traditional method of encoding the locations and values of the largest coefficients. The necessity of incorporating the FFT into the receiver, however, increases the hardware complexity of the system. Also, it is still necessary to acquire $n$ samples. Another drawback is that the FEC methods considered here will not work for approximately sparse signals or noisy signals. Our further research will consider soft decoding algorithms which

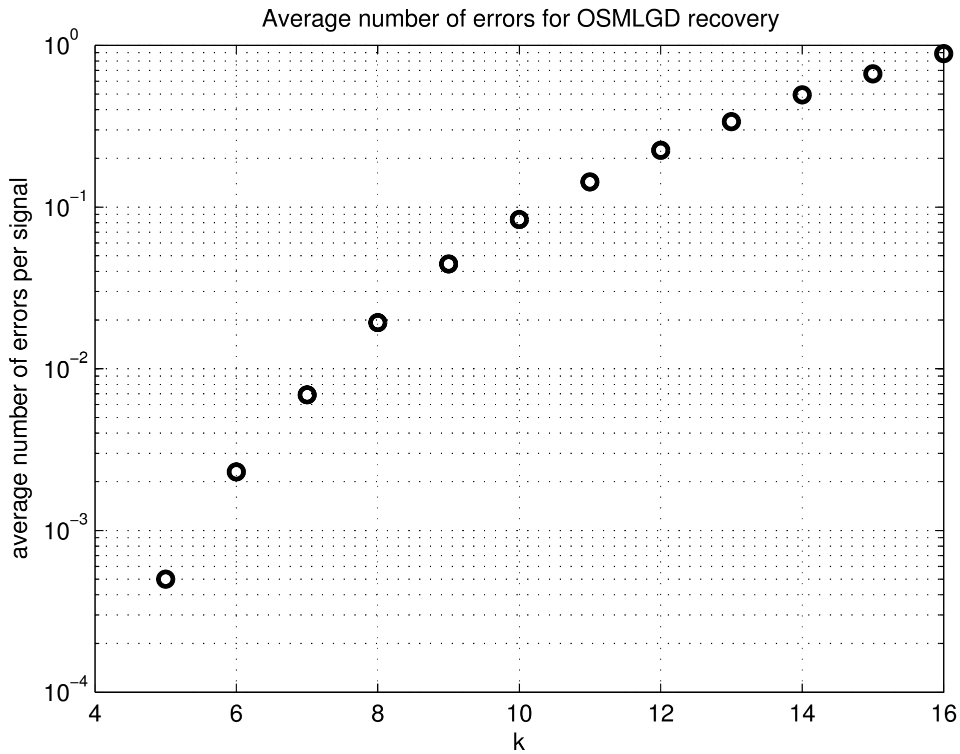


**Figure 6.** A plot of the average number of errors as a function of the number of non-zero coefficients in the sparse representation of the signal for OSMLGD recovery. The average number is zero for $k \leq 4$.

are expected to perform better under these circumstances.

The RS type sensing resulted in a compression ratio of eight and was able to recover all signals having 16 or fewer frequency components. As was seen in the results, the only error present was due to quantization. A major drawback with Berlekamp-Massey decoding, however, is that recovery fails completely for signals that contain more than 16 frequency components.

The LDPC-type sensing matrix we generated was able to compress the signal by a factor of four. However, it was necessary to have a fairly long signal to accomplish this. The main problem is that the lower bound on the number of recoverable non-zero coefficients is dependent on the column weight of the code. There is a constraint on the code construction that prevents any two columns from overlapping by more than one non-zero element. This constraint prevents 4-cycles and ensures that the check-sums are orthogonal, however, it limits the column weight of a given matrix from growing very large.

Belief propagation or other soft decoding algorithms appear to be more promising than algebraic techniques [8]. Further work should determine if belief propagation is capable of addressing all three inefficiencies of traditional acquisition techniques.

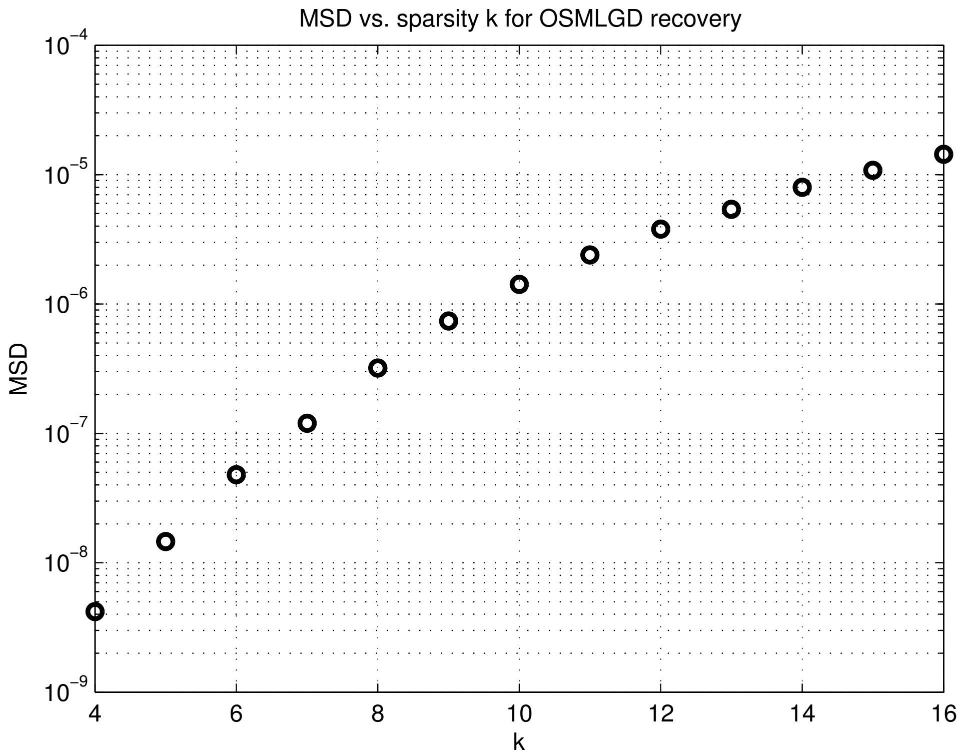


**Figure 7.** A plot of the mean-squared difference as a function of the number of non-zero coefficients in the sparse representation of the signal for OSMLGD recovery.

## REFERENCES


1. E. J. Candés and M. B. Wakin, "An introduction to compressive sampling", *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21-30, March 2008.
2. M. Akcakaya and V. Tarokh, "A frame construction and a universal distortion bound for sparse representations", *IEEE Trans. on Signal Proc.*, vol. 56, pp. 2443-2450, June 2008.
3. E. J. Candés and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4203-4215, December 2005.
4. D. L. Donoho, "Compressed sensing," *IEEE Trans. on Information Theory*, vol. 52, no. 4, pp. 1289-1306, April 2006.
5. E. J. Candés, "The restricted isometry property and its implications for compressed sensing," *Compte Rendus de l'Academie des Sciences*, vol. 346, pp. 589-592, 2008.
6. S. Lin and D. J. Costello, *Error Control Coding*, 2nd ed., Pearson Prentice Hall, New Jersey, 2004.
7. R. B. Baraniuk, "Compressive Sensing," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118-124, July 2007.
8. S. Sarvotham, D. Baron and R. Baraniuk, "Compressed sensing reconstruction via belief propagation," 2006.