Signal Reconstruction via Compressive Sampling

Anthony Ruggiero

1 Abstract

This report discusses the power of using an Orthogonal Matching Pursuit (OMP) algorithm to accurately recover a signal via ℓ-1 minimization. This demonstration is performed by taking K random linear projections of an N-dimensional signal where $K \ll N$. OMP recovers the signal iteratively selecting the most influential column of the measurement matrix at each step. Results for this implementation of OMP will be contrasted with those of the robust ℓ_1-Magic algorithm that uses a convex interior point method. Thus it can be concluded where OMP can be effectively implemented and where more rigorous approaches are necessary.

2 Introduction

Modern methods of signal recovery and image processing are governed by the Shannon-Nyquist Sampling Theorem. Which states that signals of length N, we must make exactly N measurements to reconstruct the signal without any loss of information. However, a recent development in the field of signal , termed "compressive sampling", suggests that signals and images can be reconstructed by taking less measurements than necessary. More specifically, the information would be observed in a format which would require no further compression ex post facto.

Current methods of data compression dictate that the information is sampled completely, which can lead to very high bandwidths. After which everything can be compressed into a smaller size by determining a suitable basis in which the information has a sparse representation and then discards the least significant coefficient. A common application of this is JPEG 2000, which represents data in the wavelet domain and determines which supports are influential in the reconstruction while discarding the rest. This process is very inefficient, as it demands a large amount of data acquisition up front but compression only identifies a small collection which are significant and tosses the rest. Such methods of massive data acquisition also demand hardware capable of supporting these bandwidths. For example, photos taken with a 10.2 MP SLR camera have a raw file size of almost 10 MB, but after JPEG compression, the same image is stored in a quarter of the original file size. Over 7 MB of the initial image are discarded, but the quality remains the same, meaning the bulk of the work the camera does goes to waste.

Compressive sampling aims to solve this problem of wasted effort. It has been shown, however, that N measurements may not be required for accurate reconstruction of signals and images. Rather this information can be recovered in much less than N measurements [1]. These methods of sub-Nyquist signal processing eliminate the need to acquire massive amounts of data before compression. If all of the information required to reconstruct the signal can obtained upon taking only a few measurements, then signal must have been measured in a compressed form, bypassing the need to compress it later. Methods of compression like JPEG 2000, MP3, and H.265 (standard for videos) are veritable proof that most information has sparse representations in some basis. As such, measurements can be taken in these sparse bases to capture the necessary supports without the need to discard any excess data.

By these means of measurement, the reconstruction of the original signal has turned massive data acquisition into solving a math problem. Standard methods of compression include their own decompresser to revert the information into the original form. Measurements in compressive sampling require a different approach to reconstruction than standard decompression. Acquiring the influential supports of the sparse representation cannot be done without knowing where to look *a priori*. Since sparsity is the only assumption made, locating the few supports in a different basis cannot be readily done. Instead, the measurements are multiplied by pseudorandom noise, taking the few sparse elements and smearing them across the entire spectrum while still preserving the structure of the signal. Figures 1 and 2 depict an example of the smearing. Now the signal can be sampled below the Nyquist rate from which recovery of the supports reduces to a least squares problem for an undetermined system [6].

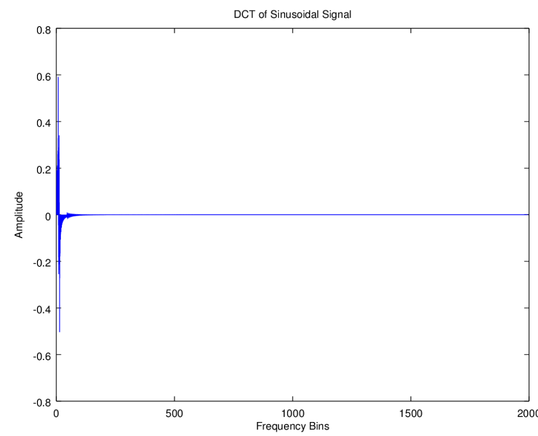

Figure 1: Sparse representation of a sinusoid in the DCT basis. Only a few nonzero elements.

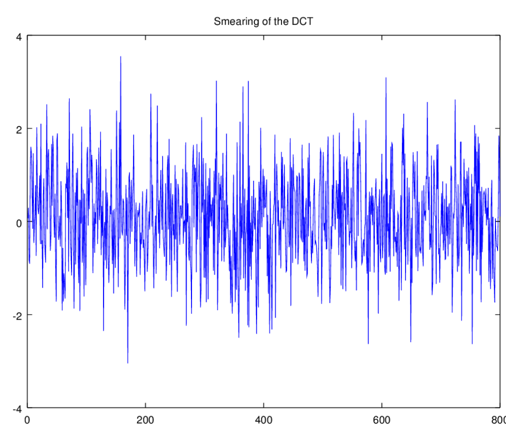

Figure 2: Smearing of the DCT in figure 1 by multiplying it by a matrix of random numbers.

3 The Problem

In practice, compressive sampling takes the form of solving undetermined linear systems with the added constraint of sparsity. Consider an N-dimensional signal, $\boldsymbol{s}$ in $\mathbb{R}^N$ that is K-sparse in some domain, be it time, Fourier, wavelet, etc, then $\boldsymbol{s}$ can be exactly recovered by only taking m random linear projections where $m \geq K$ and $m \ll N$. The sparse representation of $\boldsymbol{s}$ is then $z = \boldsymbol{\Psi s}$, where $\boldsymbol{\Psi}$ is an $N \times N$ basis transformation matrix in which $\boldsymbol{x}$ is sparse in. In order to successfully recover the signal, we measure the sparse representation with a Gaussian measurement matrix, $\boldsymbol{\Phi}$, which is an $m \times N$ matrix with randomly chosen indices. Mathematically, this is represented as $\boldsymbol{b} = \boldsymbol{\Phi} z = \boldsymbol{\Phi\Psi s} = \boldsymbol{As}$ where $\boldsymbol{A} = \boldsymbol{\Phi\Psi}$. In order to recover $\boldsymbol{s}$, we must solve the following minimization problem,

$$\hat{\boldsymbol{s}} = \text{argmin}||\boldsymbol{b}||_1 \quad \text{subject to} \quad \boldsymbol{As} = \boldsymbol{b}. \tag{1}$$

The method of solving (1) that this paper is focused on is a greedy algorithm, called Orthogonal Matching Pursuit (OMP). This method has been shown to effectively recover signals with very little computational effort [5].

4 Orthogonal Matching Pursuit

Greedy algorithms, like Orthogonal Matching Pursuit, are very popular in sparse recovery and compressive sensing problems. They can provide quick and effective solutions to (1) with a high probability of success. OMP is an iterative algorithm that selects the most influential column of $\boldsymbol{A}$ in $\boldsymbol{b}$. Denote the columns of $\boldsymbol{A}$ as atoms, $\langle \boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \boldsymbol{\phi}_3, ..., \boldsymbol{\phi}_N \rangle$ It then projects the observations $\boldsymbol{b}$ onto a subspace in which its orthogonal to the set of columns selected from $\boldsymbol{A}$ via modified Gram-Schmidt, call this the residual, $\boldsymbol{r}$. Given that $\boldsymbol{s}$ is K sparse, there should only be at most K columns in $\boldsymbol{A}$ that contribute to the measurement, therefore $\hat{\boldsymbol{s}}$ should be recovered in K iterations or less.

Stepwise, this algorithm is as follows,

- INPUT
  - The measurement matrix, $\boldsymbol{A} \in \mathbb{R}^{m \times N}$
  - The observed measurement vector, $\boldsymbol{b}$
  - The sparsity level K, which is not always known but can be *cheated* by being defining it as the number of measurements taken.
- OUTPUT
  - $\hat{\boldsymbol{s}}$, an estimate for the original signal
  - A set of the residuals removed from $\boldsymbol{b}$
- ITERATION STEP
  - Initialize the step variable $t = 1$ and the residual to the input measurement, $\boldsymbol{r} = \boldsymbol{b}$
  - Locate the most influential atom in $\boldsymbol{A}$ subject to

    $$\gamma_t = \text{argmax}_{j=1:N} |\langle \boldsymbol{r}_{j-1} \mid \boldsymbol{\phi}_j \rangle|.$$

  - Use modified Gram-Schmidt to make γ_t orthogonal to the new entry,

    $$\gamma_t = \gamma_t - \sum_{j=1}^{t} \boldsymbol{\phi}_t \langle \gamma_t \mid \boldsymbol{\phi}_j \rangle$$

    and add it atom to the set of selected atoms, $\boldsymbol{\Gamma}$.

– Solve the following least squares problem,

$$\boldsymbol{\alpha}_t = \text{argmin}\, |\boldsymbol{b} - \boldsymbol{\Gamma}^T \boldsymbol{s}|$$

which is fast since $\boldsymbol{\Gamma}$ is an orthogonal matrix.

– Finally calculate the new estimate for the original signal and update the residuals.

$$\hat{\boldsymbol{s}} = \boldsymbol{\Gamma}_T \boldsymbol{\alpha}_t$$
$$\boldsymbol{r} = \boldsymbol{b} - \boldsymbol{\alpha}_t$$

– Repeat for t ≤ K.

It is worth noting that making use of modified Gram-Schmidt in step 3 vastly improves the speed at which OMP recovers the signal, but may not be as efficient as other implementations [5].

5 Results

5.1 Sparse in Time Signals

Signals that are sparse in the time domain are the simplest examples for proving the capabilities of compressive sampling. Random samples of the signal do not need to be immediately taken, since $\boldsymbol{s}$ is highly sparse, $\boldsymbol{A}\boldsymbol{s}$ will pick out the most influential atoms of the measurement. It is also worth pointing out that the complexity of the recovery problem is greatly reduced by allowing $\boldsymbol{A} = \boldsymbol{\Phi}$ (thus $\boldsymbol{\Psi}$ is the identity) from $\boldsymbol{s}$'s sparsity in its observed domain.

The first signal tested for reconstruction is a 1D signal of $N = 500$ elements with only $K = 7$

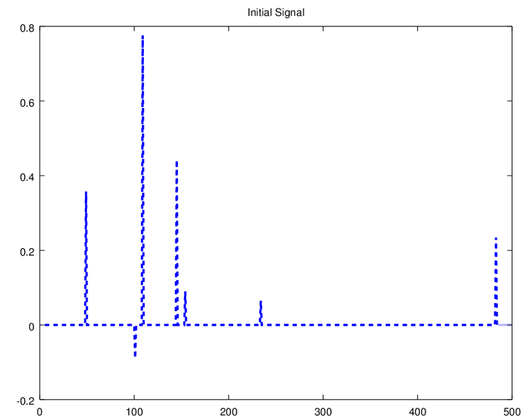

Figure 3: Original signal, constructed by $K = 7$ random permutations of N

nonzero hits, with added Gaussian noise. With this trial, I inteded to test out Candès' theorem 2.1 [1],

$$m \geq C \cdot K \cdot \log N \tag{2}$$

in which C is a constant that varies between 4 and 20, depending on the sparsity.

By inspection of figures 3 and 4, it is clear that OMP successfully chose the proper columns of the original signal. However, looking closely at the rightmost point of the reconstruction, the amplitude fell a little short of actual value in the original signal. This deviation from the exact signal is likely due to the use of modified Gram-Schmidt from step 3 above. Orthogonalization of the chosen atoms likely reduces the true support of the intended solution to the least squares problem.

The significance of such a problem can be relaxed when trying to resolve points of interest in images. Figure 4.3, depicts an image that is immediately sparse in the time domain. The im-

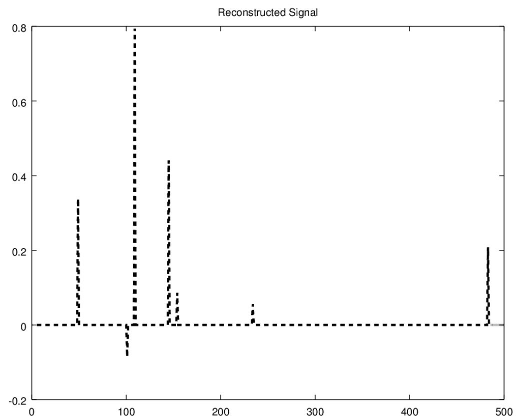

Figure 4: Reconstruction of Figure 5.1 from $m = C \cdot K \cdot \log N$ where $C = 6$ and $m = 261$ random linear projections.

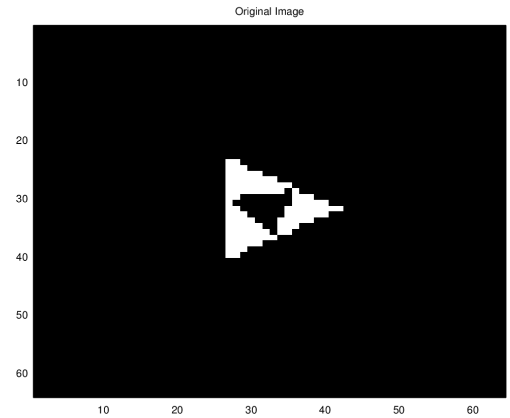

Figure 5: Original polygon image, $N = 4096$ (64×64) containing $K = 107$ nonzero coefficients.

age is rastered into a 1D vector to test it against OMP. This increases the complexity of the recovery problem to $O(N^2)$, but it has been shown by Fang, Huang, and Wu that a 2D OMP algorithm is possible and has significant performance over the 1D method [7]. For small N the increase in performance is negligible enough to forgo introducing the 2D case.

Reconstruction of the image was exact on the location of the nonzero coefficients in the original image. Successful reconstruction of the image occured with $C = 11$, else the algorithm would not correctly solve the ℓ_1 minimization problem. The goal of this trial was to test compressive sampling's applications to feature specific imaging. Amplitudes aside, OMP correctly selected the nonzero elements of the original image.

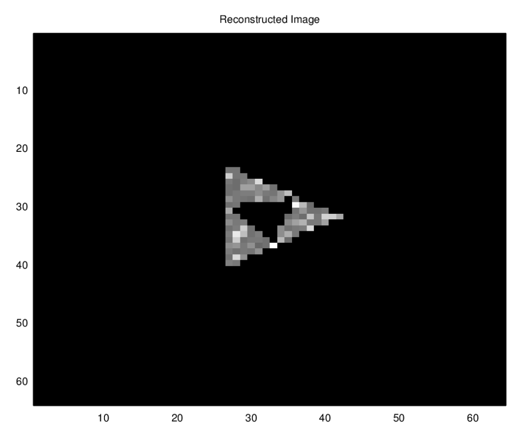

Figure 6: Reconstruction of figure 4.3, $m = 1177$ significant pixels of image were recovered but with varying amplitudes.

5.2 Sparse in Frequency Signals

Signals that are not sparse in their measured domains can still be recovered via OMP. However the probability of recovering the original signal is not as high as those that are sparse in time. In practice, the value of K is not readily known and as such recovery success is more dependent on the number of random projections taken as well as how diverse the random measurements of actually are.

The proposed signal in figure 7 is given by the following,

$$\boldsymbol{s} = \sin 6\pi t + \sin 4\pi t \tag{3}$$

which is sampled for 2 s at a rate of 1000 Hz, or simply $N = 2000$. Random Gaussian noise is then added to the signal with amplitude 10^{-5}. The random measurements of the signal are taken with $\boldsymbol{\Psi}$ which is a modified DCT matrix in $\mathbb{R}^{N \times N}$ with only $K = 250$ nonzero columns, proposed by Tropp and Gilbert [5]. In this case, nothing is assumed of the original signal other than knowledge of the basis in which it is sparse, as such K is not known. Figure 4.6 shows the reconstruction of the discrete representation of (3). Since nothing is assumed of the signal, theorem 2.1 (2) cannot be applied for the number of random projections to take with $\boldsymbol{A}$. An arbitrary number of $m = 2.5K$ was chosen carefully as to not exceed the total number of points in the original signal.

Figure 4.6 shows the reconstruction of the signal in (3) with the added noise. At first glance, the end points were not recovered to their zeroed values, and there is a clear distinction in the magnitudes of first peaks in the first and second periods. Upon closer inspection, the effects of the noise can be seen in the peaks and the troughs of the reconstruction. The difference in amplitude of the two points is a consequence of the random measurements of the signal not containing enough of the information about their relative amplitudes.

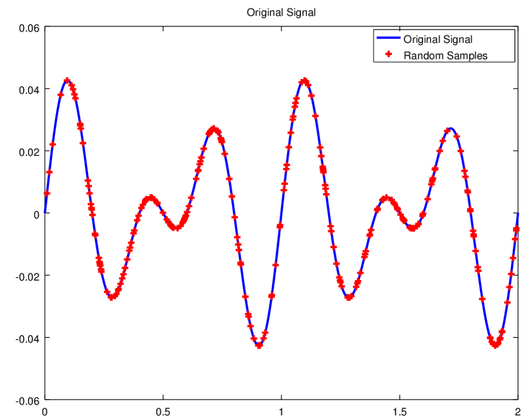

Figure 7: Proposed signal given by (3) which has 2 distinct periods in the given time frame. Random measurements of the signal are shown in red, totaling 250 points out of 2000.

If more than just sparsity is assumed, recovery error can be minimized further. In practice these assumptions can only be made if knowledge of the signal's origin is known beforehand. Additional assumptions can be made by imposing Bayesian analysis on the observations, a topic that will be omitted in this paper.

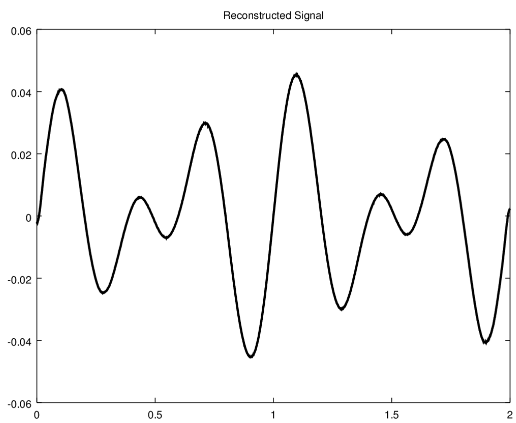

Figure 8: Reconstruction of figure 7, by only taking 250 distinct measurements. At first glance, it is clear that the end points were not recovered exactly, and the first peak of the second period clearly has a stronger amplitude than the first peak.

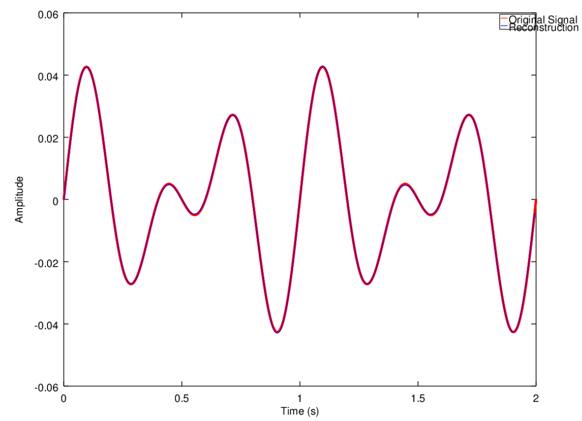

Figure 9: Zero padding the ends after the measurements lead to well behaved tails in the reconstruction.

Figure 9 depicts an overlay of a reconstruction under the assumption that the signal vanished at the ends points. This was performed by zero padding the end points after the measurements. This method tricks the algorithm into believing that the signal dies off rapidly, forcing the reconstruction to zero at those points. Knowledge of where these zeros were placed means those locations can be simply omitted when viewing the reconstruction. In the case of 5.7, the reconstruction was slightly oscillatory around zers in locations of the padding. This occurred because applying $\mathbf{\Phi}$ to the measurements only smears the observed tones, leaving OMP to fill in the missing pieces. Zeros of the initial signal can be observed, but the smearing leads OMP to believe that they are points that weren't observed.

5.3 Sparse in Wavelet

Trouble was encountered in exploring wavelet transformation matrices. Only sources on generating the Haar matrix could be found. The forward transformations were generating sparse representations of the image in figure 5. But following the OMP procedure for reconstruction, the inverse wavelet transform was not yielding appropriate answers. Upon inspection, omitting the compressive sampling procedures, applying the forward and then reverse transforms completely lost the original signal. It is possible that the sources for creating the Haar matrix were not reliable. Octave has a native discrete wavelet transform function built into its source code but is not fully implemented yet and has no available wavelet bases to choose form.

6 Error Analysis

As proposed by Tropp and Gilbert [5], optimizing the error in the reconstructed signal can be explored by minimizing the following,

$$\frac{||\hat{\boldsymbol{s}} - \boldsymbol{s}||_2}{||\boldsymbol{s}||_2}. \tag{4}$$

But (4) is not very practical when exploring error in feature specific imaging. However, such error can be determined visually rather than numerically.

Accuracy of reconstructions can also be explored by performing a chi-squared test for goodness of fit. At first glance this approach may appear straightforward, but it is unclear how many degrees freedom to use. It may be appropriate to use the number of measurements taken, however that parameter is a single variable that determines how many random projections to make.

Returning to (4) and applying it to the frequency sparse signal, recovery accuracy can be explored in both number of measurements taken and the signal in the presence of noise.

In the presence of low noise, OMP performs exceptionally well, with an average error less than 0.01. The signal-to-noise (SNR) was calculated with the following expression,

$$SNR = \frac{RMS_{signal}}{RMS_{noise}}. \tag{5}$$

For figure 10, Gaussian noise was added to the signal, but at low amplitudes, the smearing of the measurement appears to mask it and still recover a fairly accurate signal. For low SNR, OMP fails to distinguish between what is noise and where the original signal lies. Cai and Wang explore the performance of OMP in the presence of noise further in [2].

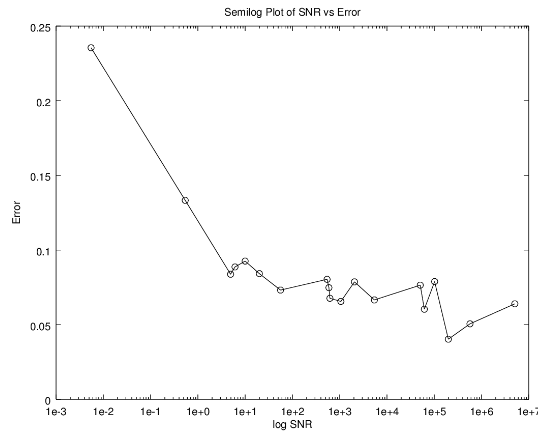

Figure 10: Error computed with (4), which falls off linearly for high SNR.

Another important aspect of the signal recovery problem is *howmany* measurements are necessary to produce and accurate reconstruction. Unlike the case with noise, OMP does not perform as well when the sampling is driven down. Figure 11 shows the performance of OMP via (4), successful reconstructions can be considered for errors below 0.1. It is evident that OMP becomes sensitive to measurements that capture only 10% of the original signal. Reconstruction error diverges for measurements that capture 5% or less of the original signal. For even lower sampling, an even sparser representation of the input would be required.

Analysis of the measurement matrix $\boldsymbol{\Phi}$, would require more time. Candès and Tao introduced the Restricted Isometry Property (RIP) [4] for $\boldsymbol{\Phi}$ that can be minimized in order to max-

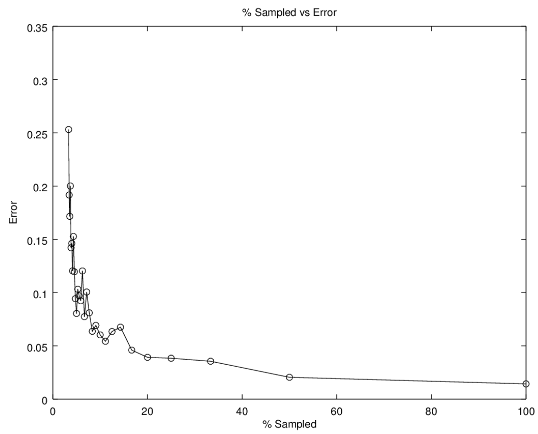

Figure 11: Average error of reconstructions as the sampling is varied. Due to the randomness in using measurements with $\mathbf{\Phi}$, OMP cannot exactly recover the signal even from N observations.

imize the probability for exact recovery,

$$(1 - \delta_s)\|\boldsymbol{s}\|_{\ell_2}^2 \leq \|\mathbf{\Phi}_s \boldsymbol{s}\|_{\ell_2}^2 \leq (1 + \delta_s)\|\boldsymbol{s}\|_{\ell_2}^2 \qquad (6)$$

for $\delta \in (0, 1]$. Additionally, their report discusses how Gaussian ensembles satisfy the RIP well with high probability. For exactly orthogonal $\mathbf{\Phi}$, perfect recovery of $\boldsymbol{s}$ occurs. Solving the minimization problem for δ in (6) is work to be left for a different project.

7 Conclusion

The work done in this paper demonstrates that OMP is an effective tool for signal recovery via random measurements. Before any optimization of the algorithm is applied, recovery accuracy performs relatively well. Even in the presence of noise, the spectral smearing manages to mask its effects quite. However, due to the "randomness" of the measurement matrix, it is possible for OMP to fail to recognize the original signal.

In practical applications of compressive sensing, the sampling rates can be driven so low that multiple trials of recovery can be performed and average while still remaining more efficient than sampling at the Nyquist rate. This would lead to consistent reconstructions as the true error would not be known *a priori*. There are, however, more robust methods than OMP, such as l_1-magic, a method developed specifically for the recovery problem [3].

While much work is left to be done to optimize the algorithm for consistent reconstructions, it remains an effective tool in compressive sampling. Being on of the easiest algorithms to implement in compressive sampling, OMP performs very efficiently and that can increased even further if knowledge of the signals origin is known. Its a fast and loose approach that you can trust.

References

[1] *Compressive Sampling*, Proceedings of the International Congress of Mathematicians, 2006. Emamnuel J, Candès.

[2] T. Tony Cai and Lie Wang. Orthogonal matching pursuit for sparse signal recovery with noise. *IEE Transactions on Information Theory*, 57:4680–4688, 2011.

[3] Emmanuel Candès and Justin Rhomberg. ℓ_1-magic: Recovery of sparse signals via convex programming. *URL: www. acm. cal-*

tech. edu/l1magic/downloads/l1magic. pdf, 4:46, 2005.

[4] Emmanuel Candès and Terence Tao. Decoding by linear programming. *IEEE Transactions on Information Theory*, 51:4203–4215, 2005.

[5] Anna C. Gilbert Joel A. Tropp. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on Information Theory*, 53, 2007.

[6] Marco F. Duarte Justin K. Romberg Joel A. Tropp, Jason N. Laska and Richard G. Baraniuk. Beyond nyquist: Efficient sampling of sparse bandlimited signals. *IEEE Transactions on Information Theory*, 56, 2010.

[7] Jianji Wu Yong Fang, Bormin Huang. $2d$ sparse signal recovery via $2d$ orthogonal matching pursuit. *Science China, Information Science*, 55, 2011.