

Affine Scaling Transformation Algorithms for Harmonic Retrieval in a Compressive Sampling Framework.

Sergio D. Cabrera^a, Jose Gerardo Rosiles^a, and Alejandro E. Brito^b

^aDept. of Electrical and Comp. Eng., The University of Texas at El Paso, El Paso, TX 79968 USA

^bXerox Corporation, 3400 Hillview, Palo Alto, CA 94304 USA

ABSTRACT

In this paper we investigate the use of the Affine Scaling Transformation (AST) family of algorithms in solving the sparse signal recovery problem of harmonic retrieval for the DFT-grid frequencies case. We present the problem in the more general Compressive Sampling/Sensing (CS) framework where any set of incomplete, linearly independent measurements can be used to recover or approximate a sparse signal. The compressive sampling problem has been approached mostly as a problem of ℓ_1 norm minimization, which can be solved via an associated linear programming problem. More recently, attention has shifted to the random linear projection measurements case. For the harmonic retrieval problem, we focus on linear measurements in the form of: consecutively located time samples, randomly located time samples, and (Gaussian) random linear projections. We use the AST family of algorithms which is applicable to the more general problem of minimization of the ℓ_p p -norm-like diversity measure that includes the numerosity ($p=0$), and the ℓ_1 norm ($p=1$). Of particular interest in this paper is to experimentally find a relationship between the minimum number M of measurements needed for perfect recovery and the number of components K of the sparse signal, which is N samples long. Of further interest is the number of AST iterations required to converge to its solution for various values of the parameter p . In addition, we quantify the reconstruction error to assess the closeness of the AST solution to the original signal. Results show that the AST for $p=1$ requires 3-5 times more iterations to converge to its solution than AST for $p=0$. The minimum number of data measurements needed for perfect recovery is approximately the same on the average for all values of p , however, there is an increasing spread as p is reduced from $p=1$ to $p=0$. Finally, we briefly contrast the AST results with those obtained using another ℓ_1 minimization algorithm solver.

Keywords: Harmonic retrieval, compressed sensing, compressive sampling, compressive sensing, affine scaling transformation, sparse signals, random projections, extrapolation.

1. INTRODUCTION, BACKGROUND AND MOTIVATIONS

Compressive sampling, also known as compressive sensing or compressed sensing, is an exciting new set of techniques that can be used to recover sparse signals and images from a smaller number of measurements than the number of samples that the Shannon sampling theorem requires [1], [2]. This is possible because many signals and images of practical interest are sparse, or concentrated, in a known domain such as the Fourier domain or the wavelet domain. This is the case for signals that are subjected to compression based on transform coding where a discrete signal or image with N samples can be well approximated with $M \ll N$ quantized coefficients.

From the theoretical point of view, a K -sparse signal x has exactly K non-zero coefficients in some domain defined by an orthonormal transform. Assuming this signal lies in an N -dimensional signal space, we require N samples to represent it. In compressive sampling, we desire to represent the signal exactly or accurately with only $M \ll N$ coefficients. Equivalently, we desire to recover the signal from $M \ll N$ measurements instead of the N samples. This is intuitively possible for values of $M > K$ though it has only recently been shown that this recovery can take place efficiently using known optimization methods [2]. An interesting type of compressive sampling problem involves the use of random projection measurements and it is being approached mostly as an ℓ_1 minimization problem subject to the linear constraint produced by the known linear measurements [1], [2].

The harmonic retrieval problem, in its most well known form, involves the determination of a small number of unknown frequencies present in a signal using only a small set of noise-corrupted, consecutively-spaced time samples [3]. This problem in a restricted form can be cast as a compressive sampling recovery problem using the Discrete Fourier Transform (DFT) as the domain where the signal is known to be sparse [2]. In this paper, we focus on this simpler case (frequencies lie on the DFT-grid) with the additional restriction that the signal measurements are noise-free. We seek to contrast the recovery problem using consecutively spaced time samples vs. randomly selected time samples vs. random projection measurements (restricted to the case where the entries are values from a zero-mean, equal-variance Gaussian distribution).

The solution to the harmonic retrieval problem from time sample measurements is a problem of signal extrapolation and spectral estimation [3]. With some restrictions, it can be cast as a problem of sparse signal recovery, best basis selection, or atomic decomposition [4]. These methods are alternatives to traditional approaches such as AR modeling or principal components frequency estimation methods [5]. In [6] we consider methods that find the representation coefficients of a harmonic signal by minimizing the ℓ_1 norm. We compare two interior-point methods to solve the linear program when the basis pursuit principle [7] is implemented. The primal-dual method, which consists of the perturbed optimality conditions of the linear program, proves to be more robust than using the primal method associated with the logarithmic barrier formulation of the linear program. We also contrast these solutions with the solution of an iterative reweighted (IRW) algorithm to find a maximally sparse representation. More recently, motivated by the application to the SAR image formation problem, a more detailed comparison has appeared in [8] which includes an analysis of the relationship between the number of data samples and the number of harmonic components.

The IRW algorithm is shown in [7] to belong to a general class of algorithms known as FOCUSS (FOCal Underdetermined System Solver) which were later generalized and re-named the Affine Scaling Transformation (AST) methods [9, 10]. In particular, the IRW and its predecessor the Adaptive Weighted Norm Extrapolation (AWNE) method [3] were found to be very similar to the basic FOCUSS algorithm to minimize the ℓ_0 quasi-norm or 0-norm-like diversity measure (also known as “numerosity”) for the harmonic retrieval or Fourier dictionary problem. In the case of AWNE, the use of a time-domain windowing makes the algorithm converge to a spectrally concentrated solution rather than a maximally sparse solution. The AST algorithm is more general in that it minimizes an $\ell_{(p \leq 1)}$ p-norm-like diversity measure which covers the two particular cases $p=0$ and $p=1$ and the range in between. Using the previous solution in an iterative manner, the AST gradually concentrates the next solution until it converges to a small number of basis vectors that can linearly represent the given data vector. Equivalently, the problem can be viewed as finding minimum norm solutions using weighted ℓ_2 norms that are increasingly concentrated along a decreasing number of coordinate directions [3], [11], see figure 1. Even though the AST can solve the compressive sampling problem using any set of linear measurements including random projections, it has not received as much attention as other optimization methods. The use of ℓ_p norm minimization for compressed sensing is discussed in [12] for $0 < p \leq 1$ but the AST methods are not discussed.

We have already illustrated the application of the AST method to the more general harmonic retrieval problem where the frequencies are not on the DFT-grid, and the data is noise corrupted, for the problem of consecutively spaced samples [13]. In that paper, we describe a computationally efficient (adaptive) refinement of the frequency scale to improve resolution by strategically increasing the dictionary of basis vectors. We also assess the benefit of using *a priori* information regarding the number of sinusoidal components to choose the regularization parameter in the noisy data problem. Computer simulations are used to assess the behavior of various schemes and the suggested approach is to first perform adaptive refinement without regularization before incorporating the regularization using the expanded Fourier dictionary.

In this paper we continue the exploration of the capabilities of the family of AST algorithms. The novel framework of compressive sampling covers the cases of randomly located time samples and random projection measurements. This paper is organized as follows. In section 2, the AST algorithm is presented for application to the harmonic retrieval problem. Conversion of the problem to a special case of compressive sampling is presented in section 3. An extensive computational evaluation of the capabilities of the AST for three different types of measurement schemes and for various algorithm parameters is presented in section 4. Results using other ℓ_1 norm minimization methods are shown in Section 5. In Section 6, the conclusions are presented in section 6.

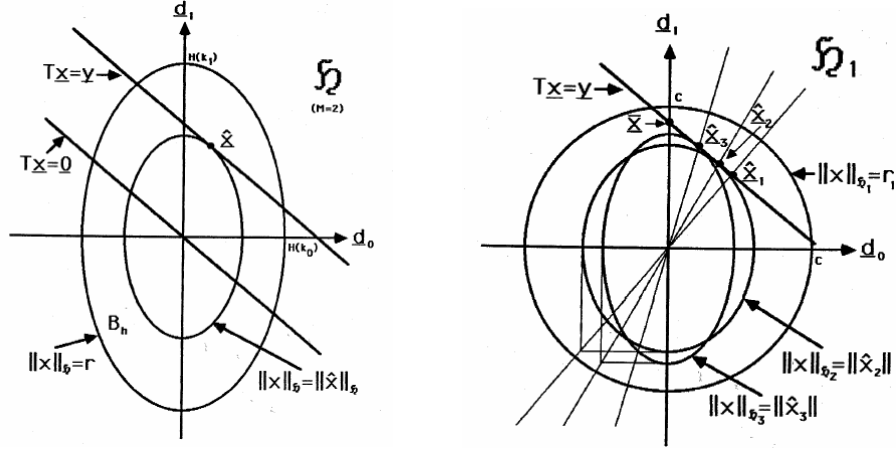


Figure 1: The minimum norm solution to $T\hat{x}=y$ is \hat{x} using a weighted ℓ_2 norm in 2-D space (left). The use of iterative reweighting achieves increasingly concentrated solutions \hat{x}_k , $k=1,2,3,\dots$ converging to a sparse solution \bar{x} (right). The figures are taken from [11].

2. THE AST ALGORITHM FOR DFT-GRID HARMONIC RETRIEVAL.

In the general form of harmonic retrieval problem, we seek to find or estimate the frequencies of a signal consisting of a sum of K complex sinusoids,

$$x(n) = \sum_{r=1}^K C_r e^{j2\pi f_r n} ,$$

from a finite set of M , possibly non-uniformly spaced, time samples $\{x(n): n=m_1, m_2, \dots, m_M\}$. The DFT-grid version of the problem assumes that $x(n)$ is of duration N and that its frequencies lie on a finite grid $f_r = k_r/N$, and therefore, its DFT is a sparse sequence with only K out of N non-zero coefficients as shown below:

$$x(n) = \sum_{r=1}^K C_r e^{j\frac{2\pi}{N}k_r n} = \frac{1}{N} \sum_{k=0}^{N-1} X_k e^{j\frac{2\pi}{N}kn} . \quad (1)$$

If the frequencies are unknown, we can get this information by recovering the complete signal by solving for the C_r 's and k_r 's or equivalently for the sparse DFT X_k on the right side of the above equation. Thus, if the number of measurements $M < N$, we can consider the problem as being that of determining a sparse solution to an underdetermined system

$$A\mathbf{X} = \mathbf{b}$$

where the vector \mathbf{b} consists of the M available measurement samples and the $M \times N$ matrix A is wide and short with entries $a_k(m_i)$. The columns of A are truncated or subsampled complex sinusoidal signals of frequencies $f_n = k/N$ which are evaluated on the index locations $n=m_1, m_2, \dots, m_M$, and \mathbf{X} is the vector of unknown DFT coefficients. Thus the set of equations to be solved are (note the change in notation from Eq. (1) to Eq. (2) from "k" to "n"):

$$\sum_{n=0}^{N-1} X_n a_n(m_i) = \sum_{n=0}^{N-1} X_n \frac{e^{j\frac{2\pi}{N}nm_i}}{N} = b_i = x(m_i); \quad i = 1, 2, \dots, M . \quad (2)$$

Using an orthonormal DFT instead, we have a new set of unknowns $\tilde{X}_n = X_n / \sqrt{N}$ and A is scaled up by \sqrt{N} .

At each iteration of the AST, see figure 2, a pseudoinverse is computed using a modified A matrix whose columns are weighted by the magnitude of the previous solution. Thus, columns of A that were significant in the previous

solution are favored to be present in the next solution. Successive weighting produces a successive de-emphasis of components that are not present in most previous solutions and in the limit, the number of surviving components is a minimum (at most $M=\text{rank}(\mathbf{A})$). The procedure is equivalent to solving a series of underdetermined problems $\mathbf{A}\mathbf{W}^{(k-1)}\underline{q}^{(k-1)} = \underline{b}$ by minimizing a weighted ℓ_2 -norm per iteration to obtain $\mathbf{X}^{(k)} = \mathbf{W}^{(k-1)}\underline{q}^{(k-1)}$ [3], [7]. Noting that the weighted \mathbf{A} matrix will become increasingly ill-conditioned, it is necessary to perform a regularization of the pseudoinverse process. This is very simple to incorporate in the AST algorithm by introducing the regularization parameter λ and replacing $(\mathbf{A}^{(k)}\mathbf{A}^{(k)H})^{-1}$ with $(\mathbf{A}^{(k)}\mathbf{A}^{(k)H} + \lambda\mathbf{I})^{-1}$ in step (3) of Fig. 2, see also [14].

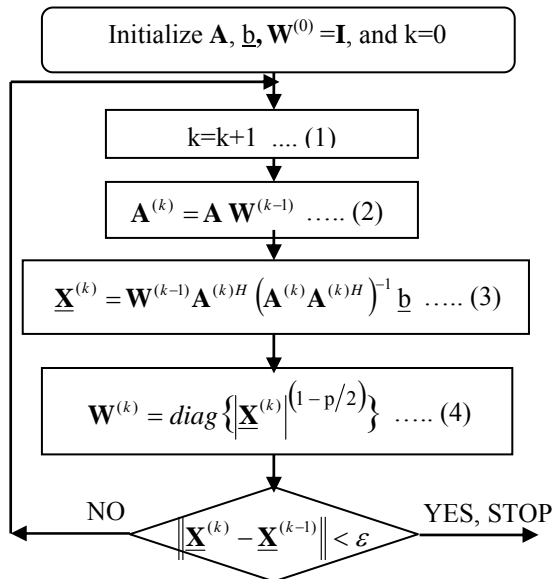


Figure 2: Flow Chart of the Affine Scaling Transformation (AST) Algorithm for a given p .

3. HARMONIC RETRIEVAL IN A COMPRESSIVE SAMPLING FRAMEWORK.

In order to set up the more general compressive sampling framework, we desire to change notation to that in [2] by changing the \mathbf{A} matrix to Θ , the measurements vector \underline{b} to \underline{y} and the desired solution for the orthonormal DFT vector $\tilde{\underline{X}}$ to \underline{x} so that we have $\Theta\underline{x}=\underline{y}$ instead of $\mathbf{A}\underline{X}=\underline{b}$. Using a linear measurements matrix Φ of dimension $M \times N$, and the orthonormal $N \times N$ DFT matrix Ψ , we have the entries in each matrix taking the form:

$$\Phi_{r,c} = \begin{cases} 1, & c = m_r + 1 \\ 0, & \text{else} \end{cases} ; \quad (\Psi^H)_{r,c} = \frac{1}{\sqrt{N}} e^{j\frac{2\pi}{N}(r-1)(c-1)} \quad (3)$$

In this case, the matrix/vector equation to solve is:

$$\Phi\Psi^H \underline{x} = \underline{y} \quad \text{or} \quad \Theta\underline{x} = \underline{y}.$$

As mentioned in [2], this is an underdetermined linear inverse problem where there are an infinite number of solutions for \underline{x} . Looking for sparse or the sparsest solution, reduces the set of solutions of interest. The approach used here is to find sparse solutions by minimize the ℓ_1 norm, the ℓ_0 numerosity, or an ℓ_p norm-like diversity measure for $0 < p < 1$, subject to $\Theta\underline{x}=\underline{y}$. In the more general Compressive Sampling (CS) problem, the matrix Φ is the measurements matrix which can be of a more general form than that in Eq. (3). For example, one can use a matrix of zero-mean Gaussian random values, random binary patterns, etc. as mentioned in [1], [2], which are the novel aspects of CS.

4. EXPERIMENTAL EVALUATIONS USING THE AST ALGORITHM

In this section, we will illustrate the various approaches to the DFT-grid or discrete harmonic retrieval problem when the Affine Scaling Transformation method is used to solve the necessary optimization problems. We evaluate the impact of the following options or choices in the signal recovery problem:

- The type of measurements matrix Φ .
- The number M of linear measurements used to attempt the recovery of a signal, N samples long, with K non-zero DFT coefficients.
- The parameter p in the range $[0,1]$ for the p norm or p norm-like like diversity measure to be minimized by AST.

Of particular interest in this paper is to find the minimum value of M which is needed in order to produce exact recovery of the test signal for the various choices of measurement matrices and algorithm parameters. We will attempt to provide experimentally derived relationships among K , M and N for various choices of Φ and p . In addition, for the cases where we cannot recover the signal exactly, we will quantify the reconstruction error that measures the closeness of the AST optimal solution \hat{x} to the original signal x using the percent RMSE metric defined using the energy of the difference signal as follows:

$$\%RMSE = 100 \|x - \hat{x}\|_2 / \|x\|_2$$

The three types of measurement matrices of interest in this paper correspond to different measurement vectors y obtained from the test signal $x(n)$ as follows:

- The matrix Φ_a selects the first M out of N time samples: $y_m=x(m-1)$ for $m=1,2,\dots,M$.
- The matrix Φ_b selects a set of M randomly located time samples $\{x(n): n=m_1, m_2, \dots, m_M\}$ in the range $[0,N-1]$.
- The matrix Φ_c computes a set of M linear random projections from $x(n)$ and the entries in the matrix are zero-mean Gaussian random numbers with variance 1 or $1/N$ [2] which scales the Φ matrix by a constant factor .

The other parameters in the AST algorithm are chosen to be the same for all examples. In particular, the tolerance for convergence is chosen to be $\varepsilon=0.001\|x\|_2$ which is a stopping condition that is always used to terminate the process before a secondary condition (limit of 1000 iterations) is reached in all the examples presented in this paper. The regularization parameter is chosen to be $\lambda=10^{-6}$ in all cases independently of the Φ matrix scaling since the results are not affected.

4.1- Selected Fixed Test Signals.

Test signals are generated having K non-zero DFT coefficients that have random amplitudes in the range $[0.1,2.0]$, random phases in $[0,\pi]$, and randomly chosen distinct, non-consecutive DFT index locations for DFT length $N=256$. The standard DFT magnitudes of these signals are shown in Figure 3 a)-c) with the frequency variable increasing from left to right. The test signals are described in detail in Table 1 in terms of DFT index (zero-based), magnitude, and phase of each coefficient.

Table 1: Description of the 3 test signals with $K=4$, $K=7$ and $K=10$ non-zero DFT components.

K=4			K=7			K=10		
index	Magnitude/N	Phase/ π	index	Magnitude/N	Phase/ π	index	Magnitude/N	Phase/ π
23	0.5871	0.4549	92	1.6262	0.2877	3	0.4756	0.0034
108	0.4810	0.3794	150	0.7462	0.2207	21	0.6630	0.4737
171	1.9004	0.3560	180	1.7750	0.3921	86	0.5578	0.7728
223	1.7940	0.1170	182	1.3244	0.1485	95	1.4233	0.4903
			187	1.4485	0.7305	109	0.7754	0.6531
			234	0.8416	0.7919	130	1.6256	0.1851
			245	1.9697	0.2625	162	0.7060	0.6781
						176	0.1279	0.8649
						185	1.6398	0.3738
						201	0.5187	0.1501

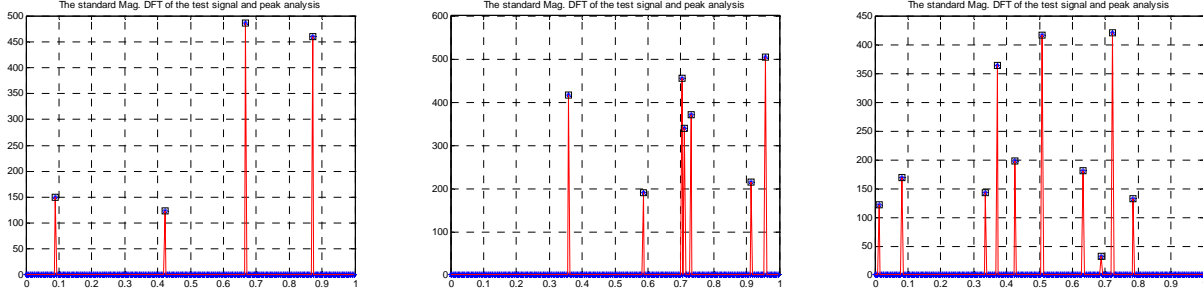


Figure 3: DFT magnitude plots for the test signals having $K=4$, $K=7$, and $K=10$ non-zero DFT coefficients respectively from left to right.

4.2- Example 1: AST from time samples vs. random projections for variable p -norm.

The test signal of length $N=256$ with $K=4$ non-zero DFT coefficients is used here. Figure 4 shows the %RMSE and No. of iterations for convergence using $M=15$ consecutive time samples as p varies from 0 to 1. Figure 5 shows the same results using $M=15$ random projection measurements. Figure 6 compares the resulting spectra for all 21 signals for each of the two types of measurements used to produce Figs. 4 and 5. In figure 4, the AST algorithm gives essentially perfect recovery up to $p=0.8$. The plots are both monotonically increasing in terms of %RMSE error and iterations vs. p norm with both metrics increasing significantly as $p=1$ is approached. In figure 5 we can see that the recoverability from random projections is very poor with a much larger and erratic %RMSE error except in the range of p in $[0.8,0.9]$. Figure 6, where the frequency variable increases from right to left, confirms the finding deduced from Figs 4 and 5. Finally, figure 7 shows the corresponding example for the case where the given measurements are randomly located time samples in the range $[0,N-1]$. The behavior of AST with these type of measurements is similar to that with random projections, therefore, we will focus on random projections for the rest of the paper.

4.3- Example 2: AST from time samples vs. random projections for variable M .

The same test signal of length $N=256$ with $K=4$ non-zero DFT coefficients is used here. Figure 8 shows for increasing M the %RMSE error metric for AST with $p=0$ using consecutive time samples and random projection measurements. In both cases, there is a distinct minimum value of M beyond which we obtain perfect recovery. It is clear that this minimum M is much smaller for the consecutive samples (left side) case (7 vs. 19). Figure 9 shows the corresponding results for $p=1$ where it is clear that an obvious minimum is not easily identified for the case of consecutive samples (left). The DFT of the recovered signals for all values of M are shown in figure 10 where it is clear that the consecutive samples case produces spectra with the correct number of peaks (4) even when perfect recovery is far from being achieved.

4.4- Example 3: AST from random projections for variable M and multiple random matrices.

The same test signal of length $N=256$ with $K=4$ non-zero DFT coefficients is used here. Figure 11 shows the %RMSE for a set of 520 examples obtained from random projection measurements for $p=0$ and $p=1$ by varying the random matrix Φ_c 20 times. For each matrix, we also sweep over all values of M from 10 to 35. This way, we can look for consistency (or lack thereof) in the recovery process. For these same cases, we monitor in Fig. 12 the number of iteration until convergence for all 520 cases for $p=0$ and $p=1$. It is clear that the average number of iterations is much higher for the $p=1$ case (33) than for the $p=0$ case (10).

4.5- Example 4: AST from random projections finding the minimum M for multiple random matrices.

The next set of examples is also for the random projections case. This time we sweep over a range of values of M until we find perfect recovery which is defined to be those cases where the %RMSE error is below 0.5. First, using the test signal with $K=4$ components, we obtain the histograms in Fig. 13 for $p=0$, $p=0.5$, and $p=1$ with corresponding average minimum values 23.8, 19.73, and 19.74 respectively. For the test signal with $K=7$ components, we repeat the experiment using $p=0$, $p=0.5$ and $p=1.0$ for which we obtain averages for the minimum value of M of 28.3, 25.7, and 31.2

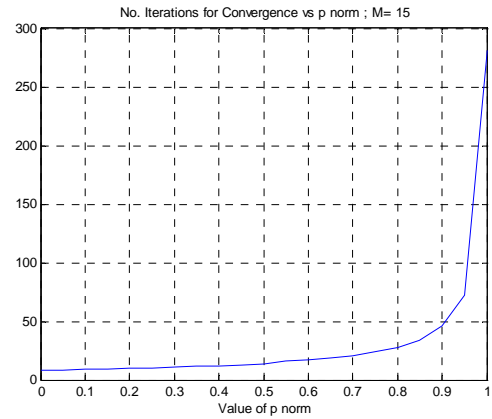
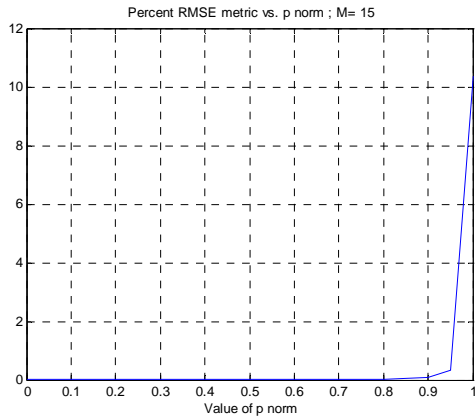


Figure 4: Error and number of iterations for convergence of AST vs. p norm ($[0,1]$ in steps of 0.05) using $M=15$ consecutive samples for the test signal with $K=4$ components.

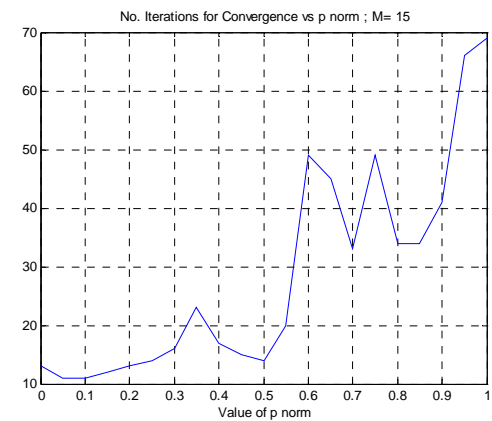
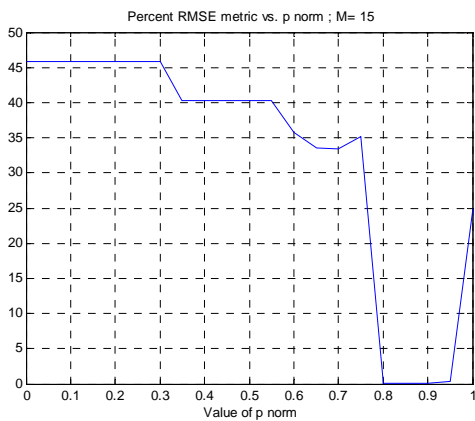


Figure 5: Error and number of iterations for convergence of AST vs. p norm using $M=15$ random projections for the test signal with $K=4$ components.

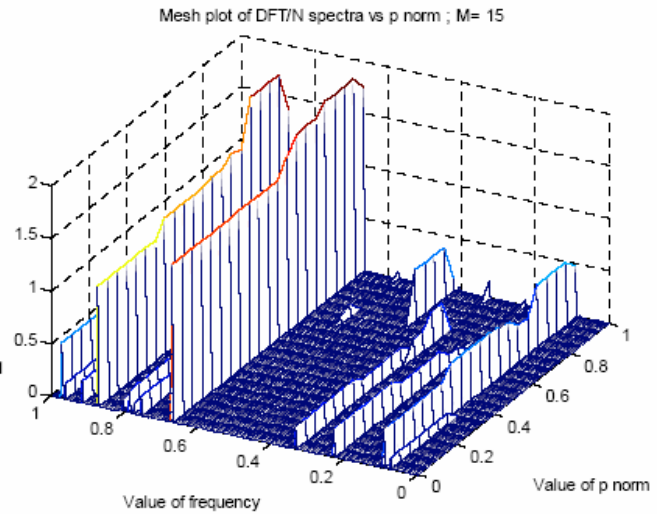
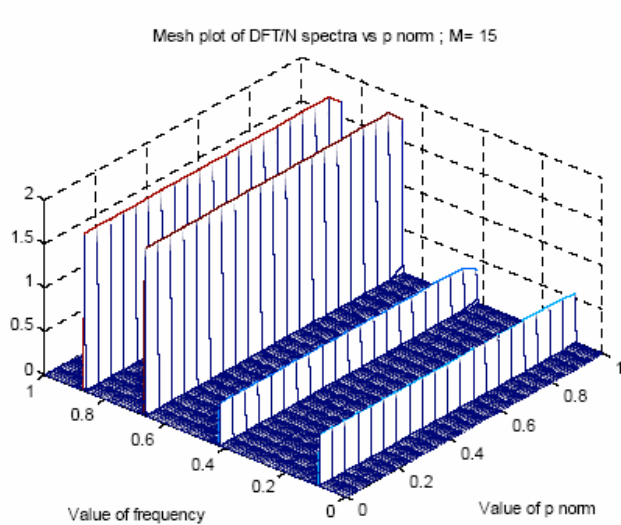


Figure 6: Plot of all 21 spectra for the AST examples in figures 4 (consecutive samples, left) and 5 (random projections, right) where p varies from 0 to 1.0 in steps of 0.05.

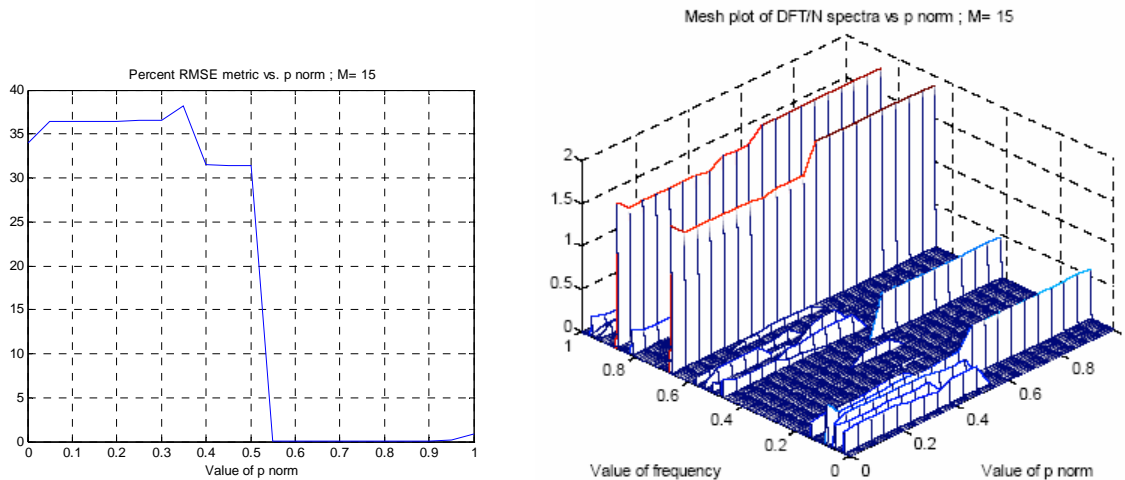


Figure 7: Same as in figures 4-6 except that we use randomly located time samples at time indices: 4, 44, 58, 113, 116, 123, 154, 156, 188, 194, 201, 209, 227, 235, and 242.

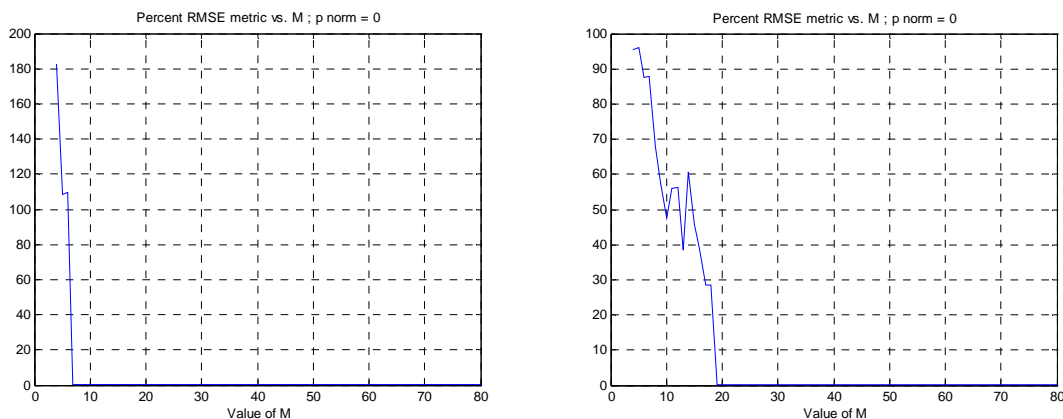


Figure 8: Error in recovery using AST with $p=0$ for an increasing number of measurements M (4 to 80): consecutive samples (left) vs. random projections (right).

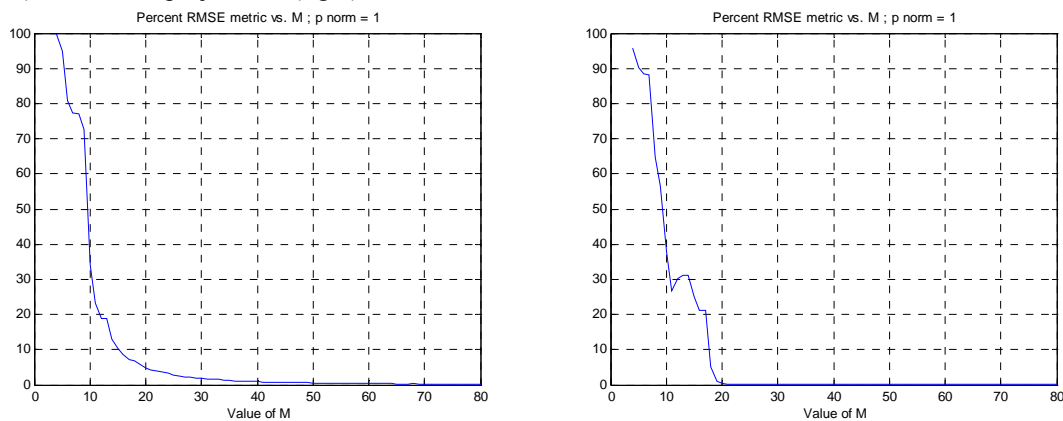


Figure 9: Error in recovery using AST with $p=1$ for an increasing number of measurements M (4 to 80): consecutive samples (left) vs. random projections (right).

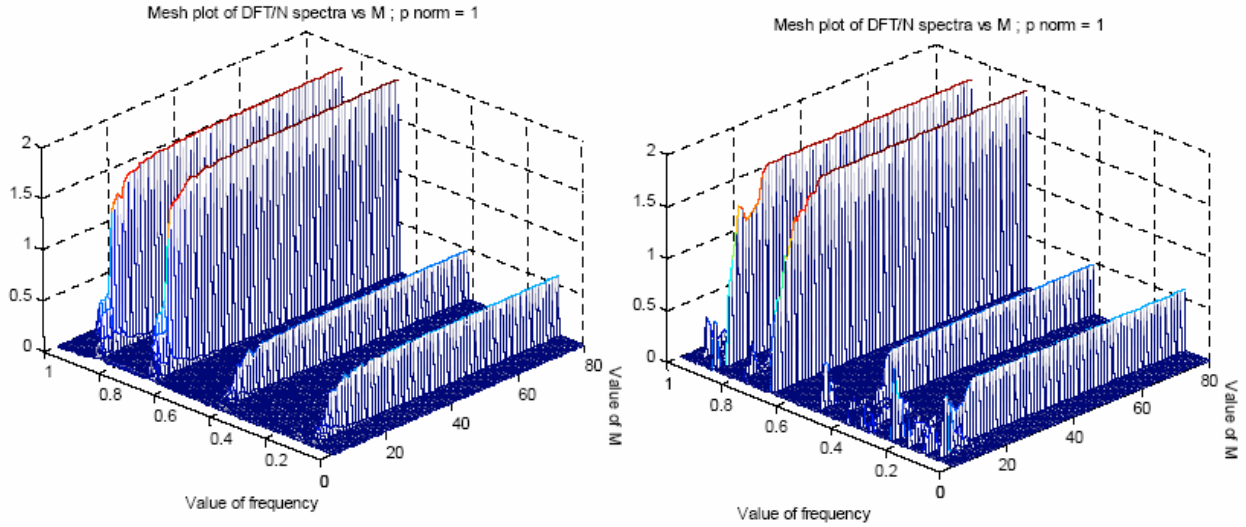


Figure 10: Plot of all 77 spectra from AST with $p=1$ in figure 9 where M varies from 4 to 80 for consecutive samples (left) and random projections (right). For the consecutive samples and small M , the spectrum is concentrated at the correct peak locations, a desirable property in spectral analysis.

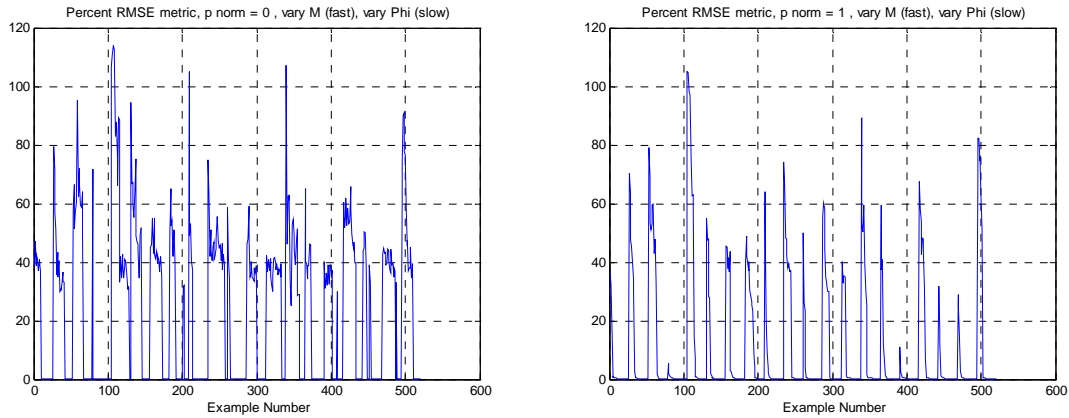


Figure 11: Illustration of the %RMSE recovery error using multiple values of M (from 10 to 35) and 20 different random projection measurements matrices in order to identify that perfect recovery is possible in this range of values of M for $p=0$ (left) and $p=1$ (right). An error of 0.5 %RMSE is always used to declare perfect recovery in this paper.

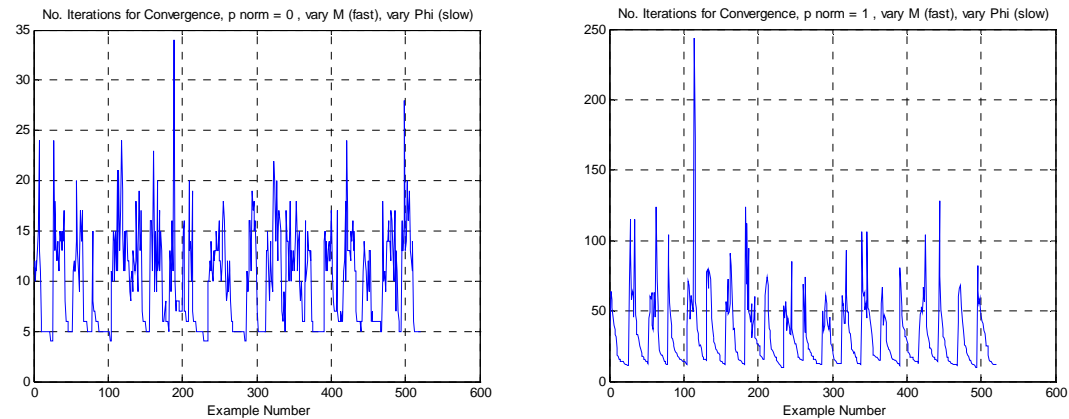


Figure 12: For the same 520 examples as in figure 11, we illustrate here the number of iterations for convergence of AST in order to compare the cases for $p=0$ (left) and $p=1$ (right). The averages are 10 for the $p=0$ and 33 for the $p=1$ case.

respectively and the histogram plots are shown in Fig. 14. Next, in figure 15 we show the results using the test signal with $K=10$ components for $p=0$, $p=0.5$, and $p=1$ to obtain average of 49.6, 42, and 41.1 respectively for the minimum value of M for perfect recovery. The results for the minimum value analysis are summarized in table 2. The trend seen in this example is that the average minimum value of M is less spread out for $p=1$, and that this value of M increases roughly linearly with K value.

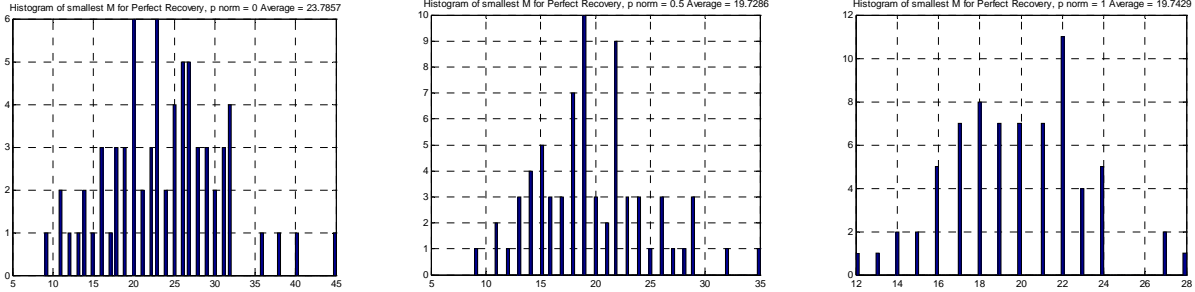


Figure 13: Histograms of minimum value of M required to achieve perfect recovery (below 0.5% RMSE) of the test signal with $K=4$ components using 70 different random projection measurement matrices for $p=0$ (left, average is 23.8), $p=0.5$ (center, average is 19.73), and $p=1$ (right, average is 19.74).

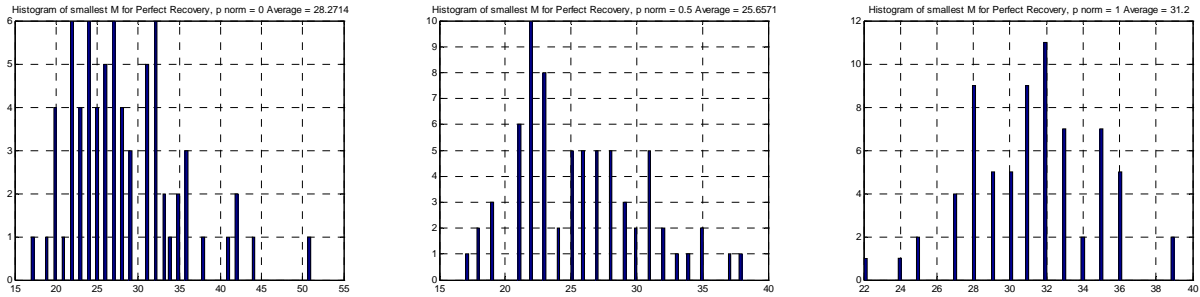


Figure 14: Histograms of minimum value of M required to achieve perfect recovery (below 0.5% RMSE) of the test signal with $K=7$ components using 70 different random projection measurement matrices for $p=0$ (left, average is 28.3), $p=0.5$ (center, average is 25.7), and $p=1$ (right, average is 31.2).

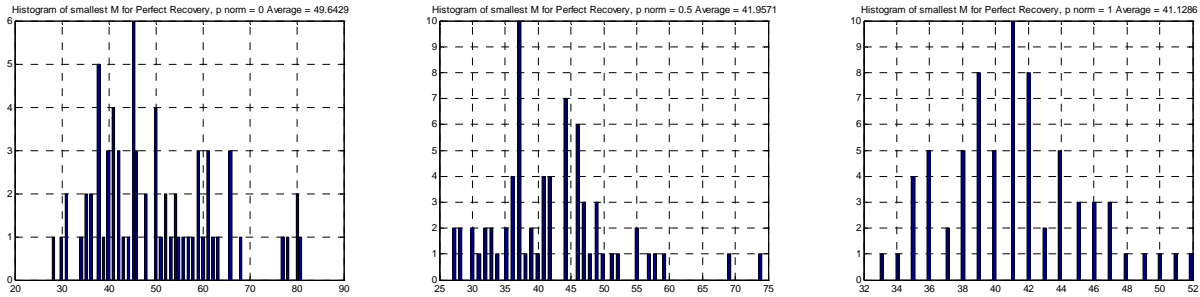


Figure 15: Histograms of minimum value of M required to achieve perfect recovery (below 0.5% RMSE) of the test signal with $K=10$ components using 70 different random projection measurement matrices for $p=0$ (left, average is 49.6), $p=0.5$ (center, average is 42), and $p=1$ (right, average is 41.1).

Table 2: Summary of the average minimum number of measurements needed to perfectly recover a signal with K components using AST with various values of diversity measure p .

p norm	0	0.5	1	0	0.5	1	0	0.5	1
K	4	4	4	7	7	7	10	10	10
Ave. M	23.8	19.73	19.74	28.3	25.7	31.2	49.6	42	41.1

5. EXPERIMENTAL EVALUATION USING OTHER OPTIMIZATION METHODS

In this section, we briefly present results to solve the same compressive sampling problem that we have illustrated in the previous section (subsection 4.5) for random projection measurements and $p=1$ norm. Here we use a different ℓ_1 norm minimization algorithm in the form that appears in the compressive sampling problem. We use the L1-MAGIC MATLAB code available at <http://www.acm.caltech.edu/l1magic/> to minimize the ℓ_1 norm with equality constraints. This solver re-casts the problem as a linear program and uses the primal-dual interior-point method as in [8]. The only change made to the L1-MAGIC code is that the random projections measurements matrix is left unchanged with zero mean and variance $1/N$. Results in figure 16 show that the average minimum M is larger for the alternative solver, furthermore, the recovery error is also larger for this solver as shown in Fig. 17.

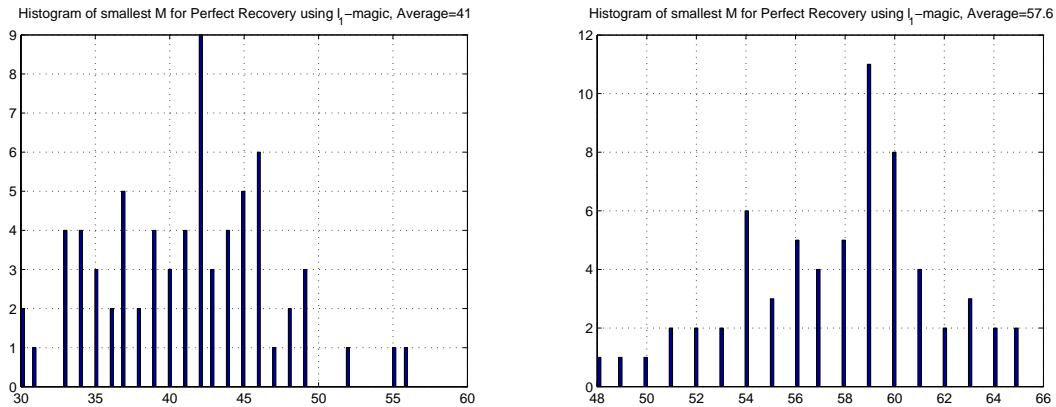


Figure 16: Histograms of minimum value of M required to achieve perfect recovery (below 0.5% RMSE) of the test signal using another ℓ_1 norm minimization algorithm (L1-MAGIC, see <http://www.acm.caltech.edu/l1magic/>) for $K=4$ (left) and $K=7$ (right) producing averages of 41 and 57.6 which can be compared to averages of 19.74 and 31.2 for AST with $p=1$.

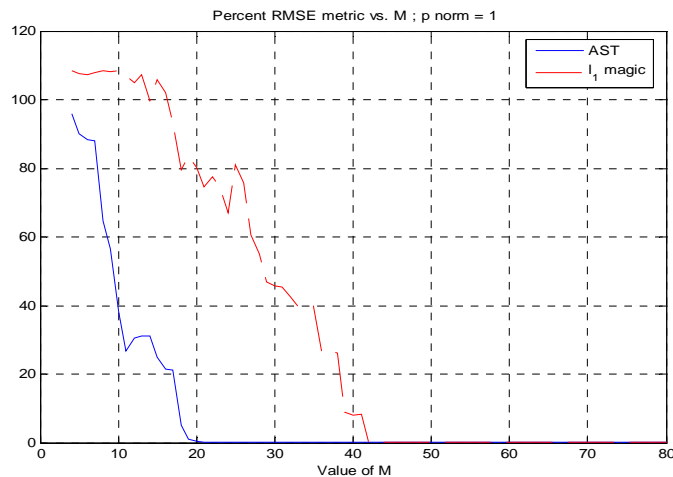


Figure 17: Further comparison of AST with the L1-MAGIC: %RMSE error for increasing values of M for the test signal with $K=4$ components.

6. CONCLUSIONS.

For the harmonic retrieval problem, we have considered the recovery problem of a sparse signal using measurements in the form of time samples and contrasted this case with that of random projection measurements, which is a novel Compressive Sampling framework. We have illustrated the use of the Affine Scaling Transformation (AST) approach

to solve this sparse signal recovery problem. Since the AST allows for the minimization of a p -norm like diversity measure including the $p=0$ case corresponding to numerosity and the $p=1$ case corresponding to ℓ_1 norm, we have contrasted these two cases as well. We found that AST with $p=0$ (vs. $p=1$) is faster in terms of iterations in all cases by a factor of 3-5 times. The minimum value of measurements M needed for perfect recovery of a signal with K components is approximately the same on the average for a given test signal for any p in $[0,1]$. On the other hand, the distribution of these values is more spread out for $p=0$ compared to $p=1$ indicating that the $p=0$ case is less predictable in this sense. In most cases, the recovery with AST is easier (requires less measurements) with consecutive samples than with random projections. For the random projections case, a brief comparison to other well-known ℓ_1 norm minimization solvers (available in the L1-MAGIC software) indicates that the AST with $p=1$ produces lower reconstruction error and a smaller average minimum M value for perfect recovery.

7. ACKNOWLEDGMENTS

We would like to thank Dr. Guoqing Liu of Nanjing University of Technology in China for the many valuable discussions of this topic over the years.

8. REFERENCES

- [1] Emmanuel J. Candes, "Compressive Sampling," in Proc. of Intl. Congress of Mathematicians, pp. 1433-1452, Madrid, Spain, 2006.
- [2] R. Baraniuk, "Compressive Sensing", in *IEEE Signal Processing Magazine*, Vol. 24, No. 4, pp. 118, 119, 120, 124, July 2007.
- [3] S. D. Cabrera and T. W. Parks, "Extrapolation and Spectral Estimation with Iterative Weighted Norm Modification," *IEEE Trans. on Signal Processing*, Vol. 39, No. 4, pp. 842-851, April 1991.
- [4] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. on Information Theory*, Vol. 47, No. 7, pp 2845-2862, Nov. 2001.
- [5] M. H. Hayes. *Statistical Digital Signal Processing and Modeling*. John Wiley & Sons, Inc., New York, NY, 1996.
- [6] A. E. Brito, S. D. Cabrera, and C. Villalobos, "Optimal Sparse Representation Algorithms for Harmonic Retrieval," in Proc. of Thirty-Fifth Asilomar Conference on Signals, Systems and Computers, pp. 1407-1411, Pacific Grove, CA, Nov. 2001.
- [7] Alejandro E. Brito. *Iterative Adaptive Extrapolation Applied to SAR Image Formation and Sinusoidal Recovery*. PhD dissertation, Dept of Electrical and Computer Engineering, The University of Texas at El Paso, El Paso, TX, July 2001.
- [8] Alejandro E. Brito, Cristina Villalobos, and Sergio D. Cabrera, "Interior-Point Methods in ℓ_1 Optimal Sparse Representation Algorithms for Harmonic Retrieval," *Journal of Optimization and Engineering*, Vol. 5, pp. 503-531, Dec. 2004.
- [9] I. F. Gorodnitsky and B. D. Rao. "Sparse signal reconstruction from limited data using focuss: A re-weighted minimum norm algorithm," *IEEE Trans. on Signal Processing*, Vol. 45, No. 3, pp. 600-616, March 1997.
- [10] B. D. Rao and K. Kreutz-Delgado, "An affine scaling methodology for best basis selection," *IEEE Trans. Signal Processing*, Vol. 47, No. 1, pp. 187-200, Jan. 1999.
- [11] S. D. Cabrera, Optimal Recovery of Signals from Linear Measurements and Prior Knowledge, Ph. D. dissertation, Dept. of Electrical and Computer Eng., Rice University, Houston, TX, May 1985.
- [12] Yaakov Tsaig, and David L. Donoho, "Extensions of Compressed Sensing," in *Signal Processing*, Vol. 86, No. 3, pp. 549-571, March 2006.
- [13] Sergio D. Cabrera, Suresh Malladi, Rama Mulpuri, and Alejandro E. Brito, "Adaptive Refinement in Maximally Sparse Harmonic Signal Retrieval," in Proc. of Eleventh IEEE Digital Signal Processing Workshop, pp 231-235, Taos Sky Valley, NM, Aug. 2004.
- [14] B. D. Rao, K. Engan, S. F. Cotter, J. Palmer and K. Kreutz-Delgado, "Subset selection in noise based on diversity measure minimization," *IEEE Trans. on Signal Processing*, Vol. 51, No. 3., pp. 760-770, March 2003.