An Improved Compressive Sensing Reconstruction Algorithm Using Linear/Non-Linear Mapping

Xinyu Zhang[*], Jiangtao Wen[*], Yuxing Han[‡] and John Villasenor[‡]

[*]Tsinghua University, Beijing, China

Email: {xy-zhang06, jtwen}@mails.tsinghua.edu.cn

[‡]Electrical Engineering Department, University of California, Los Angeles, CA 90095

Email: {ericahan, villa}@ee.ucla.edu

Abstract— We describe an improved algorithm for signal reconstruction based on the Orthogonal Matching Pursuit (OMP) algorithm. In contrast with the traditional implementation of OMP in compressive sensing (CS) we introduce a preprocessing step that converts the signal into a distribution that can be more easily reconstructed. This preprocessing introduces negligible additional complexity, but enables a significant performance improvement in the reconstruction accuracy.

I. Introduction

Compressive sensing (CS) refers to a growing body of techniques enable accurate recovery of sparsely sampled signal. Foundational contributions to CS include the work of Donoho, Candès, Romberg, Tao, and others [1], [2], [3], [4], [5]. The challenge of CS reconstruction, also referred to as the sparse approximation problem, is to solve an underdetermined system of linear equations using sparse priors. The Orthogonal Matching Pursuit (OMP) algorithm [6], [7], [8] and l_1-minimization (also called basis pursuit) [1], [2], [9] are two widely studied CS reconstruction algorithms. Orthogonal Matching Pursuit (OMP) solves the reconstruction problem by identifying the component of the sparse signal in each iteration which is most coincident with the sampling value. Stagewise OMP (StOMP) [10] and Regularized OMP (ROMP) [11] are two variants of the original OMP algorithm. The other commonly explored algorithm, l_1-minimization, replaces the original reconstruction problem with a linear programming problem. It then solves the linear programming problem using well established convex optimization approaches such as the primal-dual interior-point method [12]. l_1-minimization is generally believed to offer better reconstruction performance than OMP, while OMP has the advantage of simpler implementation and faster running speed [8]. Other reconstruction algorithms include iterative thresholding methods [13], [14] and various Bayesian methods [15], [16]. A more detailed summary of CS reconstruction algorithms is found in [17].

Papers on Bayesian Compressive Sensing [15] and optimally tuned reconstruction algorithms [18] have studied modeling of sparse signals and the worst amplitude distribution for non-zero components. However, compared with other aspects of compressive sensing, the impact of distribution and its potential usage for improving reconstruction performance is much less well-investigated. In an upcoming publication based on our earlier work in this area [19], through extensive experiments and heuristic analysis, we show that by introducing a preprocessing step D using either linear or non-linear mapping, the relative error of Dx and Dx^* is smaller than that of x and x^*, where x is the original signal and x^* is the reconstructed signal. In other words, one can convert a sparse signal with a distribution that is hard to reconstruct to another one that is easier. In [19], we study the impact of this method for l_1-minimization and iterative thresholding algorithms. In the present paper, we show that the relative error of x and $D^{-1}(Dx^*)$ is smaller than the relative error of x and x^* when the OMP algorithm is used. In what follows, we first provide analysis and experimental results about OMP reconstruction performance for sparse signals with different non-zero distributions. We then propose linear and non-linear mapping on top of the traditional OMP algorithm, and show that our improved OMP algorithm can offer better reconstruction performance without increasing the complexity and overhead of sampling and reconstruction.

II. Orthogonal Matching Pursuit Algorithm

A. Mathematical Formulations and Description of OMP

Assume a sparse signal $x \in R^n$ with k non-zero elements (called k-sparse), observed via an $m \times n$ measurement matrix A with $m < n$, producing measurement $Ax = y \in R^m$. Let a_i denotes the ith column of A, where $i \in [n]$ and $[n] := \{1, 2, ..., n\}$. Since the measurement y is a linear combination of k columns of A, the reconstruction of x can be recast as the problem of identifying the locations of these k columns. OMP solves this problem with a greedy approach. During each iteration, OMP selects the column of A which is mostly correlated with the residual of measurement y, and then it removes the contribution of this column to compute a new residual. Table I contains a description of OMP.

B. The Approximation Error Bounds of OMP

Most research regarding OMP relies on the coherence statistic of a matrix A. The coherence statistic measures the correlation between different columns of A using the absolute value of inner product:

$$\mu \triangleq \max_{i \neq j} |\langle a_i, a_j \rangle|. \tag{1}$$

OMP generally requires μ to be small in order to avoid the relatively small non-zero components being masked by large

TABLE I

ORTHOGONAL MATCHING PURSUIT

Input: The measurement y and measurement matrix A **Output:** Reconstructed signal x^*
(1) Initialize the residual $r_0 = y$, index set $C_0 = \emptyset$ and counter k=1. (2) Find the column vector a_{c_k} of A that is mostly correlated with the residual: $$c_k = argmax_c|\langle r_{k-1}, a_c\rangle|, \quad c \in [n]$$ $$C_k = C_{k-1} \cup \{c_k\}$$ (3) Solve the least-square problem: $$x_k = argmin_x||y - A_{C_k}x||_2$$ where A_{C_k} denotes the columns of A indexed by C_k. (4) Update the residual to remove the contribution of a_{c_k} $$r_k = y - A_{C_k}x_k$$ (5) Increment k, and go back to step (2) until stopping criterion holds. (6) Return the output x^* with $x^*(i) = x_k(i)$ for $i \in C_k$ and $x^*(i) = 0$ otherwise.

Fig. 1. Phase transition line of the l_1-minimization algorithm.

ones. A well-known result is that $\mu < \frac{1}{2k-1}$ is sufficient for OMP to reconstruct a k-sparse signal x precisely [8].

In [8] Tropp also used the cumulative coherence function as a generalization of coherence statistic to study OMP. For a positive integer m, the cumulative coherence function is defined as

$$\mu_1(m) \triangleq \max_{l \in [n]} \max_{\substack{S \subset [n]\setminus\{l\} \\ |S| \leq m}} \sum_{j \in S} |\langle a_j, a_l\rangle|. \tag{2}$$

[8] then gave the following result about the approximation error bounds of OMP:

Theorem 1: Suppose that $\mu_1(m) \leq \frac{1}{3}$. For every input signal x, OMP will calculate an m-term approximant x^* that satisfies

$$||x - x^*||_2 \leq \sqrt{1 + 6m}||x - x_m||_2 \tag{3}$$

where x_m is an optimal m-term approximant of the input signal.

III. IMPROVED OMP ALGORITHM THROUGH LINEAR AND NON-LINEAR MAPPING

A. Phase Transition

Letting $\delta = m/n$, $\rho = k/m$, we can obtain a two-dimensional square $(\delta, \rho) \in [0,1]^2$. Each (δ, ρ) pair in this square describes the difficulty of reconstruction. The pairs in the upper left area, i.e. pairs with smaller δ and larger ρ, represent harder reconstruction challenges since the number of non-zero components is larger with fewer measurements. Theoretical analysis and empirical results have shown that for most reconstruction algorithms, there are two regions in the square: a success region, where the probability for the algorithm to precisely reconstruct the original sparse signal tends to be 1, and a failure region, where this probability tends to be 0. This phenomenon is referred to as phase transition in CS and the boundary of these two regions is called the phase
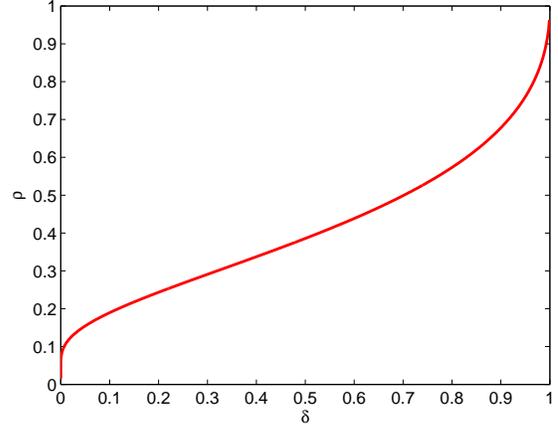
transition line [18]. A higher phase transition line indicates better performance of the reconstruction algorithm. To date only the phase transition line of l_1-minimization algorithm, as shown in Figure 1, has been rigorously proven and calculated with the help of computational geometry [20], [21], [22]. The cases of other algorithms are still under study and can only be empirically observed at present. Nevertheless, the evaluation by means of phase transition line is closer to the real performance of the reconstruction algorithm than the approximation error bounds [23], since the derivation of the latter often involves some loose scaling.

In this paper, we measure the location of the phase transition line in a similar way to that used in [18]. For each (δ, ρ) pair we carry out M Monte Carlo trials of sampling and reconstruction. Using the definition below of Signal-to-Noise Ratio (SNR), we regard a trial as successful if

$$SNR = 10\log_{10}\frac{||x||_2^2}{||x - x^*||_2^2} \geq 40dB \tag{4}$$

where x and x^* represent the original sparse signal and the output of reconstruction algorithm respectively. The number of successful trials, denoted by S, is distributed binomial $Bin(\pi, M)$ with the success probability $\pi \in [0, 1]$. The finite-N phase transition is the value of ρ where the success probability crosses *50%*:

$$\pi(\rho|\delta; N) = \frac{1}{2} \qquad at \qquad \rho = \rho^*(\delta; \theta)$$

where θ refers to the set of parameters of the algorithm. In the experiments below, we set $M = 20$ and $\Delta\delta = 0.035, \Delta\rho = 0.04$ for (δ, ρ) to find the approximate phase transition location on the lattice of the square, and then subdivide ρ near the location to find a more precise transition point.

B. Motivation

It has been theoretically proven that the phase transition line of l_1-minimization algorithm is independent of non-zero distribution of the sparse signals [20], [23], which is not the case of other algorithms. Empirical results in [18], [24], [10] have
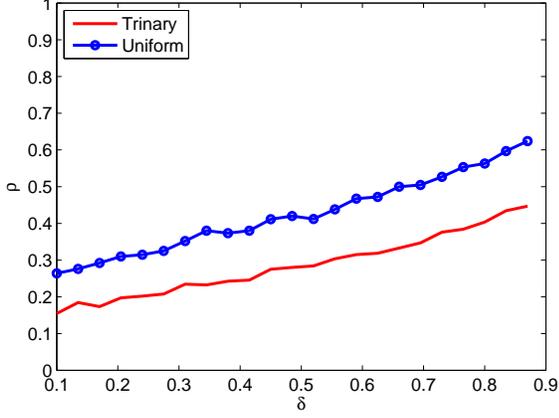
Fig. 2. Phase transition lines of trinary and uniform signals whose non-zero components $x_i = \pm 1$ and $x_i \sim U[-1, 1]$ respectively.

shown that for iterative algorithms, Sparse Bayesian Learning and StOMP, the reconstruction performance of trinary signals, i.e. signals whose non-zero components have same amplitude with random sign, is worse than that of signals with other non-zero distributions, e.g. uniform or Gaussian distribution. We focus on the case of OMP algorithm and the results are shown in Figure 2. As can be seen from the figure, uniformly distributed signals are much easier to recover than trinary signals.

This observation can be heuristically justified as below. If the non-zero components of the k-sparse signal x are ordered as $\{x_{(1)}, x_{(2)}, ..., x_{(k)}\}$ with $|x_{(1)}| \geq |x_{(2)}| \geq ... \geq |x_{(k)}|$, then from Theorem 1 and the definition of SNR, the following holds if $\mu_1(m) \leq \frac{1}{3}$:

$$
\begin{aligned}
SNR &\propto \frac{||x||_2^2}{||x - x^*||_2^2} \\
&\geq C \frac{||x||_2^2}{||x - x_m||_2^2} \\
&= C \frac{\sum_{i=1}^k |x_{(i)}|^2}{\sum_{i=m+1}^k |x_{(i)}|^2} \geq C \frac{k}{k - m}
\end{aligned} \tag{5}
$$

where $C = \frac{1}{\sqrt{1+6m}}$ and this constant only depends on the measurement matrix A. The last inequality is satisfied with equality if and only if x is trinary.

The approximation error bounds for signals whose non-zero components conform to uniform distribution can also be analyzed heuristically. In this case for a k-sparse signal x and $k \to +\infty, m \to +\infty, \frac{m}{k} \to \tau$ with $0 < \tau < 1$,

$$
\lim_{k \to +\infty} E\left(\frac{\sum_{i=1}^k x_{(i)}^2}{k} \right) = 1/3,
$$

$$
\lim_{k \to +\infty} Var\left(\frac{\sum_{i=1}^k x_{(i)}^2}{k} \right) = 0;
$$

$$
\lim_{\substack{k \to +\infty, m \to +\infty \\ m/k \to \tau}} E\left(\frac{\sum_{i=m+1}^k x_{(i)}^2}{k - m} \right) = \frac{(1 - \tau)^2}{3},
$$

$$
\lim_{\substack{k \to +\infty, m \to +\infty \\ m/k \to \tau}} Var\left(\frac{\sum_{i=m+1}^k x_{(i)}^2}{k - m} \right) = 0.
$$

By the Law of Large Numbers (LLN), $\forall \epsilon > 0$, we have:

$$
P(|\frac{||x||_2^2}{k} - 1/3| < \epsilon) = 1,
$$

$$
P(|\frac{||x - x_m||_2^2}{k - m} - \frac{(1 - \tau)^2}{3}| < \epsilon) = 1.
$$

Therefore, the lower bound of SNR can be derived:

$$
\begin{aligned}
\lim_{\substack{k \to +\infty, m \to +\infty \\ m/k \to \tau}} \frac{||x||_2^2}{||x - x^*||_2^2} &\geq \lim_{\substack{k \to +\infty, m \to +\infty \\ m/k \to \tau}} C \frac{||x||_2^2}{||x - x_m||_2^2} \\
&= \frac{C}{(1 - \tau)^3}.
\end{aligned} \tag{6}
$$

As shown in (6), the approximation error bound of sparse signals whose non-zero components are uniformly distributed is lower than that of trinary sparse signals with identical measurement matrix and sparseness. Furthermore, although the derivation here is about the infinite case, figure 2 indicates that the situation is similar for the finite case.

C. Improved OMP using Pseudo-random Linear Mapping

The heuristics and experimental results above indicate that the performance of OMP for signals of the same sparseness but different distributions can be significantly different, with trinary signals harder to reconstruct than uniform ones. One way to make use of this is to perform pre-processing for the input signal before sampling, with the purpose of converting it from a distribution that is difficult to reconstruct to another one that is easier. This idea underlies the improved OMP approach using pseudo-random linear mapping that is described in the present paper. The linear mapping is applied to the trinary input x, creating a new input Dx for sampling, where

$$
D = Diag\{d_1, ..., d_n\} + \epsilon \times Diag\{sgn(d_1), ..., sgn(d_n)\} \tag{7}
$$

is a diagonal matrix. Here $d_i \sim U[0, 1]$ and $sgn(.)$ represents the sign function. ϵ will be explained in the next paragraph. As for reconstruction, the scaled input Dx is first reconstructed using traditional OMP, and then the approximation of original input x is resolved from Dx^* by $D^{-1}(Dx^*)$.

In practice, while inverting D after Dx^* has been reconstructed, if some elements of d_i are too small, any error in the corresponding components of Dx^* will be amplified by $1/d_i$. To prevent this from happening, in the definition of D, we introduce a small fixed nonzero number ϵ to each element of d_i. Although setting $\epsilon = 0$ will lead to higher probability of successful reconstruction of Dx, the SNR of x is in general lower. This is due to the large distortions at positions of Dx^* whose corresponding d_i's are small. By contrast, when $\epsilon > 0$ the reconstruction of Dx is less likely to succeed, but when it does, the SNR of x is usually higher. Obviously, when $\epsilon \to +\infty$, the mapping becomes virtual "trinarization" of x.
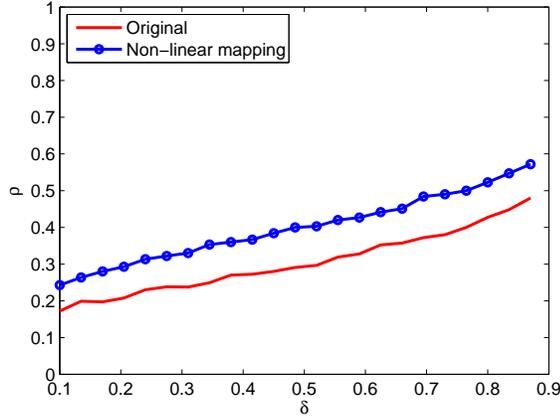
Fig. 3. Phase transition lines of original uniform signals whose non-zero components $x_i \sim U[1,2]$ and signals after non-linear mapping $\Phi(x) = x^3$.

D. Extension to Non-linear Mapping

The linear mapping can be seen as a type of "stretch" operation applied to the non-zero components of the signal. It can be extended to non-linear mapping for signals that conform to general distributions, which is necessary for universal application of this method since there are many types of signals for which prior information is lacking. The non-linear mapping function has the form

$$\Phi(x) = x^\alpha \qquad (8)$$

where the non-linear mapping parameter α is an odd number to prevent two components of original signal being mapped to the same value.

Figure 3 shows experimental results regarding the effect of non-linear mapping. Here the non-zero components of original sparse signals conform to a uniform distribution on the interval $[1,2]$. It is clear form the figure that the performance of OMP to reconstruct x^3 is much better than that of reconstructing x directly.

In accordance with the above analysis, we can also give a heuristic justification for using the lower bound of the SNR:

$$
\begin{aligned}
SNR_x &\propto \frac{||x||_2^2}{||x - x^*||_2^2} \\
&\geq C\frac{||x||_2^2}{||x - x_m||_2^2} = C\frac{\sum_{i=1}^{k} x_{(i)}^2}{\sum_{j=m+1}^{k} x_{(j)}^2}, \quad (9)
\end{aligned}
$$

while the case of reconstructing $\phi(x) = x^3$ is

$$
\begin{aligned}
SNR_{\phi(x)} &\propto \frac{||\phi(x)||_2^2}{||\phi(x) - \phi(x^*)||_2^2} \\
&\geq C\frac{||\phi(x)||_2^2}{||\phi(x) - \phi(x)_m||_2^2} \\
&= C\frac{\sum_{i=1}^{k} x_{(i)}^6}{\sum_{j=m+1}^{k} x_{(j)}^6}. \qquad (10)
\end{aligned}
$$

Because $|x_{(1)}| \geq |x_{(2)}| \geq ... \geq |x_{(n)}| \geq 0$, we know

$$\sum_{i=1}^{m} \sum_{j=m+1}^{k} x_{(i)}^2 x_{(j)}^2 \left(x_{(i)}^4 - x_{(j)}^4 \right) \geq 0$$

or

$$\left(\sum_{i=1}^{m} x_{(i)}^6 \right)\left(\sum_{j=m+1}^{k} x_{(j)}^2 \right) \geq \left(\sum_{i=m+1}^{k} x_{(i)}^6 \right)\left(\sum_{j=1}^{m} x_{(j)}^2 \right),$$

$$\frac{\sum_{i=1}^{k} x_{(i)}^6}{\sum_{j=s+1}^{k} x_{(j)}^6} \geq \frac{\sum_{i=1}^{k} x_{(i)}^2}{\sum_{j=s+1}^{k} x_{(j)}^2}.$$

Therefore, the lower bound of SNR to reconstruct $\phi(x)$ is higher, which is consistent with the results in Figure 3. Although the above analysis (including the case of linear mapping) is based on SNR bounds, and only applies to the performance of recovering mapped signals instead of the case after inverse mapping, experiments in the next section show that the improved OMP algorithm using linear and non-linear mapping does lead to significantly better performance.

From the analysis above, it can be seen that the SNR bound gets higher with larger α, theoretically leading to better reconstruction performance. However, it is worth noting that the analysis is based on the performance of the mapped signals $\Phi(x)$, rather than that of the reconstructed signals x^* after inverse mapping $\Phi^{-1}(x)$. In fact, based on extensive experimental results as shown in Section IV, it is observed that there is a trade-off between the reconstructed mapped signal $\Phi(x)$ and reconstructed original signal x. This can be explained as: with larger α, the small non-zero components of the original signal get relatively smaller after non-linear mapping x^α. This may make the reconstruction algorithm unable to separate these small components from the interference caused by the other nonzero components of x^α with the coherence of the columns of A, as OMP fails to recover these components [25]. Therefore, although the performance of reconstructing $\Phi(x)$ gets better in the sense that the important components are reconstructed more precisely, after the inverse mapping, the effect caused by failing to reconstruct small components becomes larger, resulting in worse overall reconstruction performance.

To summarize, the proposed improved OMP algorithm works in the following manner: apply a linear or non-linear mapping to the original sparse signals; perform sampling and reconstruction using the traditional OMP algorithm; after the mapped signals have been recovered, use an inverse mapping to get the reconstructed signals. From the viewpoint of practical application, for the linear-mapping case, we are simply sampling using AD, instead of with A after linear mapping of x, so the complexity and overhead of sampling is comparable with that of typical CS approahces. In addition, the non-linear mapping could potentially be achieved by analog devices in applications such as MRI.
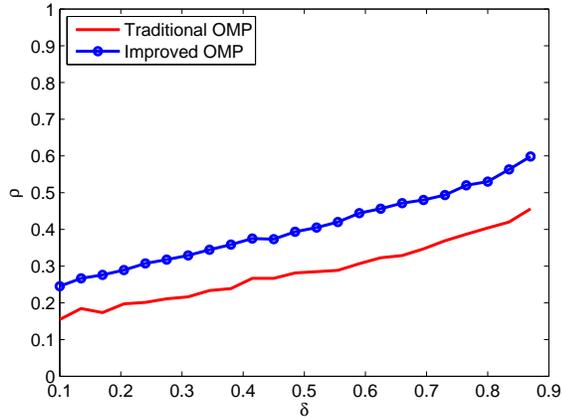
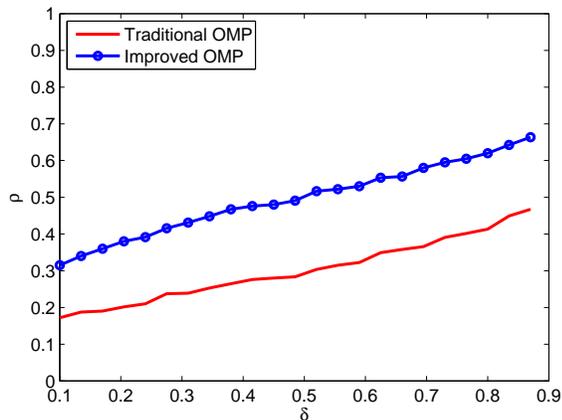Fig. 4. Phase transition lines of traditional OMP algorithm and our improved OMP using linear mapping.

(a) Performance of reconstructing mapped signals with different α

Fig. 5. Phase transition lines of traditional OMP algorithm and our improved OMP using non-linear mapping.

(b) Performance of reconstructing original signals with different α

Fig. 6. The trade-off between α and reconstruction performance.
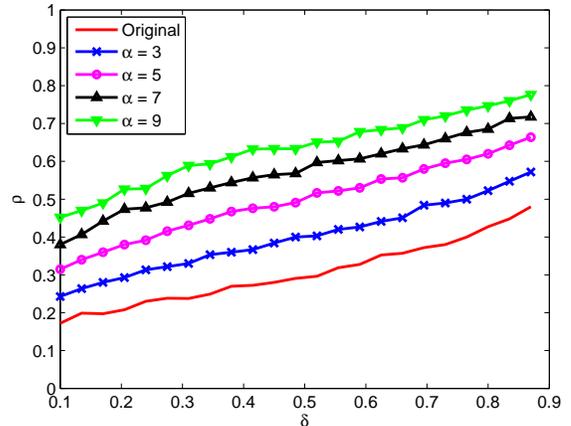
IV. EXPERIMENTAL RESULTS

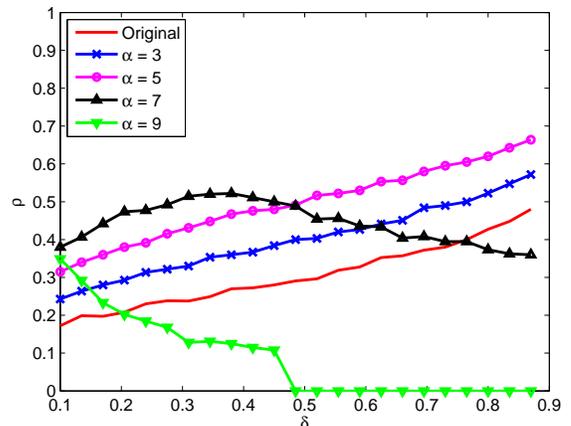### A. Improved OMP Using Linear and Non-linear Mapping

We first consider the performance improvement using linear and non-linear mapping. Figure 4 compares the performance of traditional OMP and the improved OMP using linear mapping. The original signals are trinary and ϵ in Equation (7) is set to 0.05. Figure 5 shows the case of non-linear mapping, where the non-zero components of original signals are uniformly distributed on [1,2] and α in Equation (8) is set to 5. The higher phase transition lines after using linear and non-linear mapping in both figures indicate the better performance of the improved OMP algorithm.

B. The Trade-off between α and Reconstruction Performance

Figure 6 confirms the analytical results regarding the trade-off for selection of parameter α. Here the non-zero components of the original signals are uniformly distributed on [1,2]. Figure 6(a) shows that the performance of reconstructing x^{α} does indeed become better with larger α, while figure 6(b)

indicates that when $\alpha = 7$, the performance becomes unstable for larger δ and the situation is even worse for $\alpha = 9$. This trade-off for α plays an important role in the performance of non-linear mapping.

C. Applications

We consider two applications used in [10] to test the performance of the improved OMP algorithm. The first uses the object *Bumps* from the Wavelab package [26], rendered with $N = 4096$ samples. Such signals are known to have wavelet expansions with relatively few significant coefficients, as shown in Figure 7(a). We used hybrid CS in [27] for reconstruction and set the number of measurement $m = 360$. The results in figure 7 show that the improved OMP offers better reconstruction performance than traditional OMP in terms of the relative error.

We further test the improved algorithm on an image (512×512 pixels) as displayed in 8(a). We used this image and a Multiscale CS scheme [27] with a total number of $m = 69834$ measurements. Here we applied our non-linear
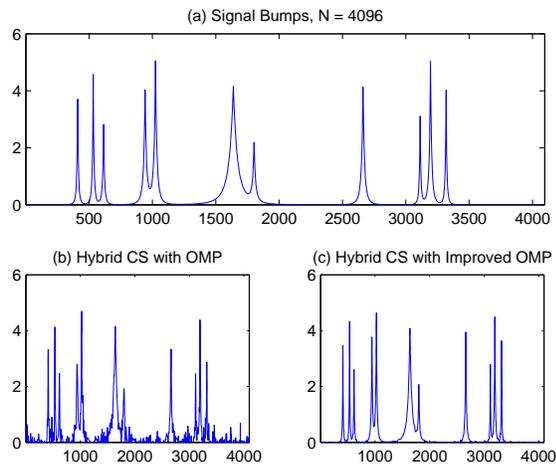
Fig. 7. Reconstruction of *Bumps* with hybrid CS. (a): Signal *Bumps*, with 4096 samples; (b): Reconstruction with traditional OMP, $||x^* - x||_2/||x||_2 = 0.247$; (c): Reconstruction with improved OMP using non-linear mapping, $||x^* - x||_2/||x||_2 = 0.0428$.
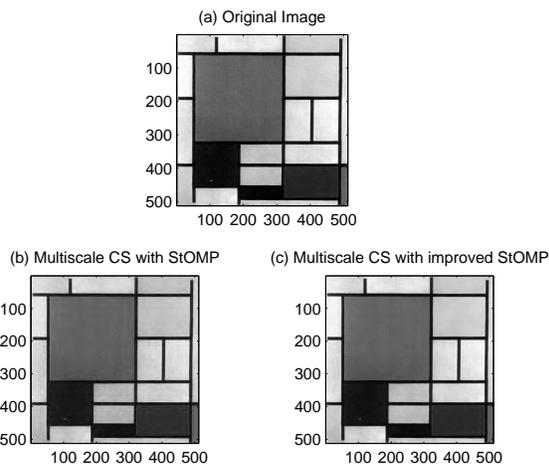

Fig. 8. Reconstruction of *Mondrian* with Multiscale CS. (a): *Mondrian* painting, 512×512 pixels; (b): Reconstruction with traditional StOMP, $||x^* - x||_2/||x||_2 = 0.0361$; (c): Reconstruction with StOMP using non-linear mapping, $||x^* - x||_2/||x||_2 = 0.0231$.

mapping to the StOMP algorithm, a variant of OMP described in [10]. We tested the performance of StOMP with and without non-linear mapping. We used the CFAR threshold selection rule and set the stage number to 10. The results in figure 8 show that using non-linear mapping improves reconstruction performance for StOMP in terms of the relative error.

V. CONCLUSIONS

In this paper, we have shown the relationship of the distribution of sparse signals on the performance of the OMP algorithm both heuristically and experimentally. We then proposed a method that uses this relationship by applying linear and non-linear mapping to convert the signal from a distribution that is hard to reconstruct to a signal that is easier to reconstruct.

Experimental results have shown that our improved OMP algorithm can offer better reconstruction performance than traditional OMP algorithm.

REFERENCES

[1] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Review*, vol. 43, no. 1, pp. 129–159, 2001.

[2] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[3] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[4] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, pp. 1289–1306, 2006.

[5] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9-10, pp. 589–592, 2008.

[6] Y. Pati, R. Rezaiifar, and P. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," *the twenty seventh Asilomar Conference on Signals, Systems and Computers*, no. 1, pp. 40–44, 1993.

[7] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[8] J. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2231–2242, 2004.

[9] D. L. Donoho and Y. Tsaig, "Fast solution of l_1-norm minimization problems when the solution may be sparse," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 4789–4812, 2008.

[10] D. Donoho, Y. Tsaig, I. Drori, and J.-L. Starck, "Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit," Stanford Statistics Department, Tech. Rep. 2006-02, Mar. 2006.

[11] D. Needell and R. Vershynin, "Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit," *Foundations of Computational Mathematics*, vol. 9, no. 3, pp. 317–334, 2009.

[12] S. Wright, *Primal-dual interior-point methods*. Society for Industrial Mathematics, 1997.

[13] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, 2004.

[14] T. Blumensath and M. Davies, "Iterative thresholding for sparse approximations," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5, pp. 629–654, 2008.

[15] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2346–2356, 2008.

[16] D. Wipf and B. Rao, "Sparse Bayesian learning for basis selection," *IEEE Transactions on Signal Processing*, vol. 52, no. 8, pp. 2153–2164, 2004.

[17] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 948–958, 2010.

[18] A. Maleki and D. Donoho, "Optimally tuned iterative reconstruction algorithms for compressed sensing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 330–341, 2010.

[19] X. Zhang, Z. Chen, J. Wen, J. Ma, Y. Han, and J. Villasenor, "A compressive sensing reconstruction algorithm for trinary and binary sparse signals using pre-mapping," *to appear in Proc. 2011 IEEE Data Compression Conference*, March.

[20] D. L. Donoho and J. Tanner, "Counting faces of randomly projected polytopes when the projection radically lowers dimension," *AMERICAN MATHEMATICAL SOCIETY*, vol. 22, no. 1, pp. 1–53, 2009.

[21] ——, "Neighborliness of randomly-projected simplices in high dimensions," *Proceedings of the National Academy of Sciences*, vol. 102, no. 27, pp. 9452–9457, 2005.

[22] D. L. Donoho, "High-dimensional centrally symmetric polytopes with neighborliness proportional to dimension," *Discrete and Computational Geometry*, vol. 35, no. 4, pp. 617–652, 2006.

[23] D. Donoho and J. Tanner, "Precise undersampling theorems," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 913–924, 2010.

[24] D. Wipf and B. Rao, "Comparing the effects of different weight distributions on finding sparse representations," in *Advances in Neural Information Processing Systems 18*, 2006, pp. 1521–1528.

[25] Z. Ben-Haim, Y. Eldar, and M. Elad, "Coherence-based performance guarantees for estimating a sparse vector under random noise," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5030–5043, 2010.

[26] J. Buckheit and D. Donoho, "Wavelab and reproducible research," in *Wavelets and Statistics*, A. Antoniadis, Ed. Springer, 1995.

[27] Y. Tsaig and D. Donoho, "Extensions of compressed sensing," *Signal processing*, vol. 86, no. 3, pp. 549–571, 2006.