Recover Sparse Signals from Under-Sampled Observations Lab Report

Yang Zhao

***Abstract*—Sampling acts as a bridge between analogue and digital worlds. The traditional Nyquist Theorem requires a sampling rate of at least twice the maximum frequency of the original signal, which is universal but inefficient in some cases. New cost-saving sampling and recovering approaches have been developed for signals with specific characteristics. This paper examines the reconstruction of sparse signals from incoherent under-sampled observations. We investigate three greedy algorithms, namely orthogonal matching pursuit (OMP), subspace pursuit (SP), and iterative hard thresholding (IHT) on a simplified model - an underdetermined linear system. It has been demonstrated that although these approaches handle the underdetermined equations well when sparse solutions exist, the complexity of OMP is high, the accuracy of IHT is low, and SP can overcome both problems and maintain a high success rate even if the solution is not very sparse.**

***Index Terms*—Compressed sensing, sparse signal reconstruction, sampling, reconstruction algorithms.**

I. Introduction

Sampling can be interpreted as projecting the original signal onto a subspace and represent it by the projection. If the error can be recovered or estimated with the properties of the subspace, the signal can be reconstructed without loss of information. Nevertheless, a perfect reconstruction on the conventional bandlimited subspace may require a large sampling rate, which can be either unavailable due to system restrictions or unworthy when the signal has only a few components (i.e. sparse). Fortunately, it is proved by [1] that the signal can be reconstructed by the samples with sampling rate far less than the Nyquist rate when it is sparse. The contribution is that as long as the target is sparse in any domain (i.e. after a specific transformation), it can be stored in a compact form and recovered from those under-sampled observations. Therefore, it is possible to use smaller storage to keep files while maintaining the resolution. Also, the dimension of data can be reduced significantly, and relevant analysis can be more efficient. Nevertheless, the new sampling scheme cannot be equally-spaced as conventional ones, and the reconstruction algorithms are not as accurate when the signal is not very sparse.

This report investigates the criteria of under-sampled observations, reconstructs the sparse signal from samples with OMP, SP, IHT algorithms, and compares the accuracy and complexity of those approaches on different sparsities. We utilise an underdetermined system as a simplified model. It is assumed that the sparse solution x is the representation of the target signal in a particular domain, the linear observation matrix A indicates the sampling process, and the sample y denotes the observed projection. First, the process and logic of three greedy algorithms are explained. Then, the outputs are compared with the exact solution to examine accuracy. Finally, the performance of those approaches is compared when the sparsity is relatively high.

II. Theory and Methods

A. Problem Formulation

A signal $x \in R^n$ of length n is defined as $S-sparse$ if the count of nonzero entries S is negligible compared to n [2]. It is mentioned in [3] and [4] that most natural signals can be transformed into a domain in which they are nearly sparse and thus can be approximated by sparse representations with more compact size. For instance, polynomial, exponential, and rational objects can be reproduced succinctly by the shifted version of scaling function of wavelets [5], while audio and image are typically sparse after Fourier transform and discrete cosine transform (DCT) respectively [2]. Therefore in most cases, start from a general signal s, it is possible to find a transformation that maps the original signal s onto another field such that the representation x in the new field is sparse (or nearly sparse) with sparsity S. We denote the transition matrix ψ such that:

$$x = \psi s$$

Notice that there should not be any distortion in the mapping, which implies the dimension of sparse signal x must be the same with the original signal s. In other words, the transition matrix ψ is a square matrix with size n^2.

Next, the sparse signal x is to be sampled with a rate much lower than Nyquist rate, which is equivalent as projecting the target onto subspaces with a compact dimension. The sampling process can be modelled as multiplying the signal x by an observation matrix A of size $m \times n (m \gg n)$:

$$y = Ax$$

where y is the observations of length m. Nevertheless, for the successful reconstruction of sparse signal x, the observation matrix A must satisfy the restricted isometry property (RIP) condition with parameters (k, δ) [6]. For a constant δ and all $\tau \in [n]$ such that $|\tau| \leq K$ and for all $x \in R^{|\tau|}$, it holds that:

$$(1-\delta)\left\|x\right\|_2^2 \leq \left\|A\tau x\right\|_2^2 \leq (1+\delta)\left\|x\right\|_2^2$$

Given a candidate sampling scheme A, it is difficult to verify whether the RIP condition is satisfied. Fortunately, [7] proved that the equivalent condition of RIP is that the observation matrix A and transition matrix ψ are incoherent. Therefore, we can choose a random Gaussian matrix as the observation matrix A which is incoherent with any transition matrix ψ [8]. It indicates that original signal s can always be reconstructed with an under-sampled pattern if the observation process follows a random Gaussian distribution. Notice that it is entirely different from the conventional sampling where the sampling space is fixed.

Now we have successfully interpreted the compressed sensing problem as a underdetermined linear system:

$$y = Ax = A\psi s$$

where y is the sample, A denotes the observation Gaussian matrix, x indicates the sparse representation of the original signal s after a certain transformation represented by transition matrix ψ.

B. Greedy Algorithms

We examine three conventional algorithms to find possible sparse solutions $\hat{x}$ with sparsity S of an underdetermined linear system. The aim of algorithms can be rendered as to approximate the observations linearly with the degree of freedom to be S. It can also be interpreted as to find exact solutions with the objective of maximising the most significant S entries to reduce the influence of the rest on observations, then to represent the original signal with those largest terms. In all cases, the sparsity S is assumed known.

1) The Orthogonal Matching Pursuit (OMP) Algorithm [9]: OMP is the orthogonal case of matching pursuit, an approximation on the subspace spanned by non-zero components of sparse solutions. Due to the randomness of observation, although there are aliasing and overlapping result from under-sampling, those essential components tend to retain traceable elements that can be detected through a threshold. When available kernels are infinite, the original signal can be entirely reconstructed by:

$$f(t) = \sum_{n=0}^{\infty} a_n g_{\gamma_n}(t)$$

where g_{γ_n} is the atom from the overcomplete dictionary D, n is the index of the atom, and a_n is the weight of the corresponding atom. The count of kernels available is the sparsity S. Therefore, those components with the largest contribution to the signal (i.e. largest weight a) should be used in approximation to minimise error. The logic of OMP is as below:

- Initialisation: set the initial solution x_0 to zero and the residue function y_r to y. No atoms are utilised, and the sparsity set S is ψ.
- Find the index of maximum remaining atom x_r from residue function y_r:
$$A^T y_r = A^T A x_r = cIx_r$$
which corresponds to the component that contributes most to the residue function, namely the most significant part to reduce error.
- Add that index to the sparsity set S. Span a new subspace by the updated set S:
$$\hat{S} = \hat{S} \cup \sup(H_1(A^T y_r))$$
- Derive the projection operator of the new subspace, and find the possible sparse solution $\widehat{x_s}$ in this subspace:
$$\widehat{x_{\widehat{S}}} = A^{\dagger}_{\widehat{S}} y$$
- Calculate the residue function y_r of the existing solution $x_{\widehat{S}}$:
$$y_r = y - A\hat{x}$$
then the normalised error:
$$\varepsilon = \frac{y_r}{y}$$
- Keep iterating until at least one of the three statements is satisfied:
  - the number of selected atoms equals the sparsity S;
  - the normalised error ε is smaller than the desired value;
  - the algorithm starts to diverge.

  then the current solution is regarded as valid.

OMP adds one more kernel to the solution in an iteration according to the degree of significance. In other words, the weight of existing atoms are unchanged once those values are derived, and the error is reduced by the latest component although the optimisation is narrower with the expansion of sparsity.

2) The Subspace Pursuit (SP) Algorithm [10]: SP utilises a different scheme to generate the support set $\hat{S}$. Instead of expanding the subspace every time, it refines the estimated set $\hat{S}$ during each iteration. In this recursive filtering method, the candidate set $\widetilde{S}$ is produced based on the estimated set $\hat{S}$ at the current stage and possible competitors from the residue function. It introduces the possibility to remove the indexes which are considered as reliable at previous stages but proved to be wrong with the refinement of iterations. A typical process of SP is:

- Initialisation: determine the initial support set $\hat{S}$ by the largest S entries of $A^T y$, which represent the most significant components of the observation y. Then calculate the corresponding residue function y_r;
- Combine the current support set $\hat{S}$ and the largest S atoms that contribute to the residue function y_r to create a candidate set $\widetilde{S}$ of size $2S$:
$$\widetilde{S} = \widehat{S} \cup \sup(H_{\widehat{S}}(A^T y_r))$$
- Solve the equation on the projection subspace of dimension $2S$. $b_{\widetilde{S}} = A^{-1}_{\widetilde{S}} y$ is the solution with sparsity $2S$;
- Truncate $b_{\widetilde{S}}$ and keep the largest S entries in the estimated set $\widehat{S}$:
$$\widehat{S} = \sup(H_{\widehat{S}}(b))$$

- Find the sparse solution in the new subspace of dimension S spanned by the estimated set $\widehat{S}$:

$$\widehat{x} = A^{\dagger}_{\widehat{S}} y$$

- Calculate the residue function y_r of the existing solution $x_{\widehat{S}}$:

$$y_r = y - A\hat{x}$$

  then the normalised error:

$$\varepsilon = \frac{y_r}{y}$$

- Keep iterating until at least one of the three statements is satisfied:
  - the normalised error ε is smaller than the desired value;
  - the algorithm starts to diverge.

  sparse solutions satisfy the requirements above are acceptable.

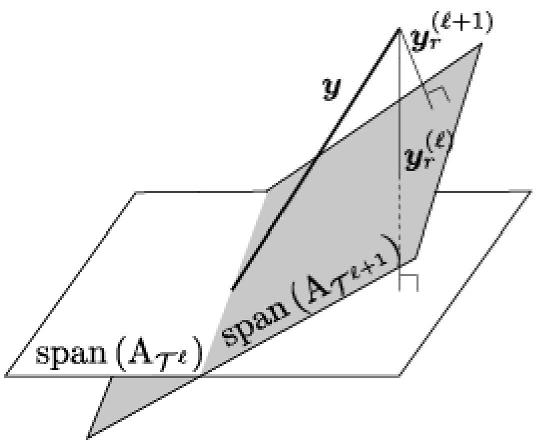

Fig. 1. Illustration of the subspace pursuit algorithm [3]

As Figure 1 indicates, instead of reserving the existing components and introduce another one in each iteration as OMP, SP analyses the problem in a global range by choosing the most substantial S entries out of the candidate set $\widetilde{S}$ of size $2S$. Therefore, the complexity is significantly reduced when the solution is not very sparse, and the accuracy is improved since those indexes wrongly introduced in early stages can be filtered out with more iterations. The statements are demonstrated by simulations in the next section.

3) The Iterative Hardthresholding (IHT) Algorithm [11]: IHT starts from an initial solution to approach a feasible sparse one by iterations. In each loop, a candidate is derived by combining existing solution $\widehat{S}$ and x residue function $A^T y_r$ then optimised by truncating to keep the most significant S entries. The process can be expressed as:

- Initialisation: set the existing solution $\widehat{S}$ as $\mathbf{0}$ and the residue function y_r as y;
- Add the x residue function $A^T y_r$ to the current sparse solution $\widehat{S}$; the sum indicates a candidate $\widetilde{S}$ which is less sparse but more accurate;
- Keep the most significant S elements of the candidate $\widetilde{S}$ and set other entries to zero. It reduces the accuracy but the overall error after an iteration tends to be smaller;
- Keep iterating until at least one of the three statements is satisfied:
  - the normalised error ε is smaller than the desired value;
  - the algorithm starts to diverge.

  a sparse solution can be obtained in the end.

Notice that there is no subspace projection in the IHT algorithm. It is trying to derive a more accurate solution based on the previous result and the x residue, then truncate it for sparsity. The instability in truncation is more likely to produce a larger error than the previous result, which means the divergence is possible to happen before it provides a precise sparse solution.

III. Result and Analysis

A. Exercise 1

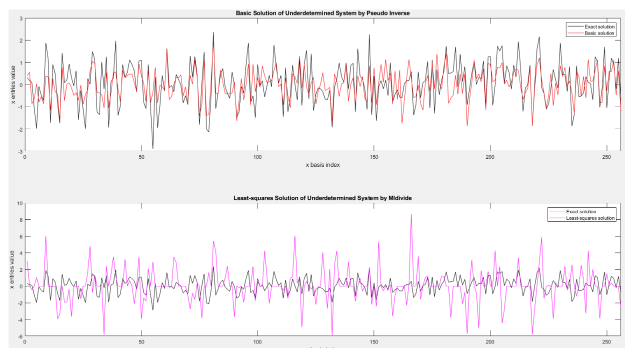

Fig. 2. Unsparse solutions of an underdetermined system by pseudo inverse and least square error approaches

Figure 2 compares the solution of an underdetermined system where the number of observations m is 128 while unknowns n is 256. It can be observed that although pseudo inverse provides a better approximation to the original signal than matrix left divide, both strategies cannot recover the original dense signal from under-sampled observations. Also, the normalised error is 1.2517×10^{-16} for pseudo inverse and 9.7036×10^{-18} for matrix left divide, which means both solutions are valid. The reason is that an underdetermined system has infinite possible solutions. Therefore, it is impossible to recover all the information hidden on observations with an insufficient sample rate if the signal is not sparse. Nevertheless, as mentioned above, if we can find a transformation that maps this dense signal into a sparse domain, the representation in that field can be almost fully reconstructed from an underdetermined system, then the signal can be retrieved by inverse transformation. In following exercises we assume the signal has been successfully transformed into the sparse domain.

B. Exercise 2

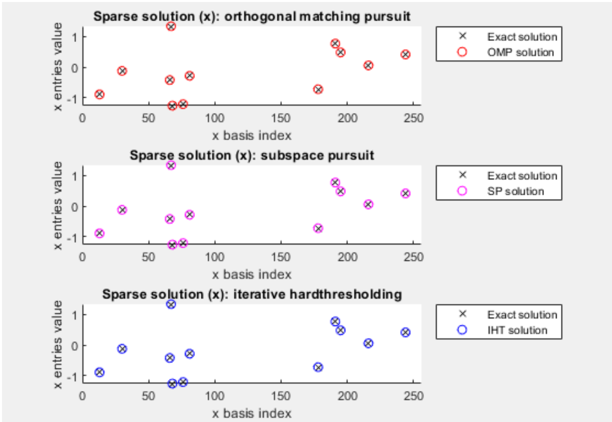

Fig. 3. Sparse solutions of an underdetermined system by OMP, SP, IHT algorithms compared with ground truth: case 1

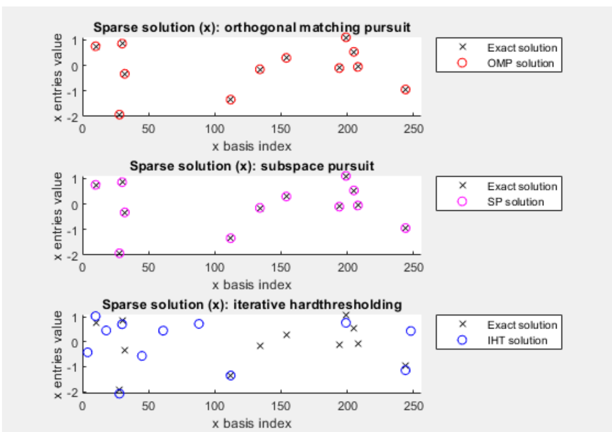

Fig. 4. Sparse solutions of an underdetermined system by OMP, SP, IHT algorithms compared with ground truth: case 2

Similarly, the number of observations m is 128, and the number of unknowns n is 256. Notice that the sparsity S should be known or estimated in advance, which is 12 in this case.

Figure 3 and 4 are two typical cases of sparse solutions derived by OMP, SP, and IHT algorithms denoted by circles. Also, the normalised error of the two cases are respectively 2.7033×10^{-16}, 2.9703×10^{-16} for OMP, 2.7033×10^{-16}, 2.4581×10^{-16} for SP, and 9.6995×10^{-7}, 0.6655 for IHT. It can be observed that OMP and SP handled the underdetermined system quite well, but the solution suggested by IHT can have large deviation from the ground truth signal that indicated by crosses. Overall, it suggests that these algorithms can be employed to recover sparse signals even if the sampling frequency is far lower than the Nyquist rate.

The result is as anticipated with the analysis in the previous section. OMP seeks to find one component with the most significant distribution to the error at a time. If the current solution is more sparse than desired, the precision can be improved

TABLE I
PARAMETERS IN SUCCESS RATE SIMULATION

Number of observations	$m = 128$
Number of unknowns	$n = 256$
Sparsity	$S = 3 - 63$
Number of tests per sparsity	$p = 500$
Normalised error bound	$\varepsilon_{thr} = 10^{-6}$

by spanning the solution subspace by combining the existing one and the dimension of the most-contributing component, and thus updating the result on that space. Therefore, the error is reduced in each iteration, and the sparse solution tends to be accurate. In contrast, SP amends the estimated set $\widehat{S}$ out of the enlarged subspace $\widetilde{S}$. This strategy enables the possibility to remove those kernels considered essential in the early stage but proved wrong during refinements. Thus, the error drops rapidly and the solution proposed is accurate. Nevertheless, there are no subspace projections in IHT, and it merely attempts to approximate the solution with the previous result, and x residue then truncate for the sparsity. The truncation is more likely to produce a significant error, which means this algorithm is not very stable and thus with more possibility to diverge before a precise result is obtained.

C. Exercise 3

We desire to test the feasibility of greedy algorithms in recovering the signal when it is not very sparse. In this section, the complexity of those approaches are compared, and the accuracy is examined when the sparsity decreases. It is determined that the solution is valid if the normalised error is smaller than a given bound, and the success rate is calculated from numerous independent tests for each value S. Table I describes the parameter in this simulation.

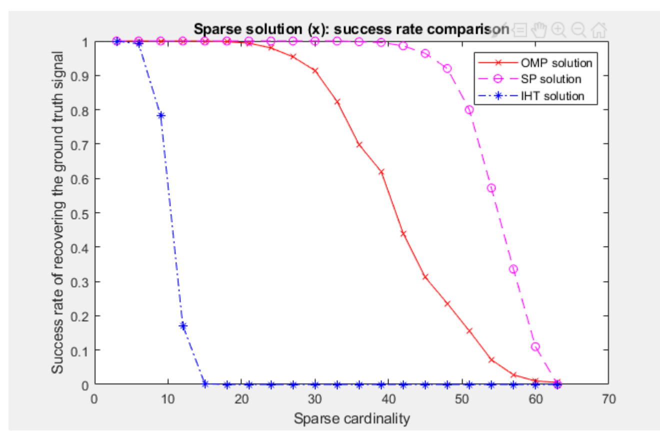

Fig. 5. Success rate comparison of OMP, SP, IHT for sparse cardinality $S = 3 - 63$

Figure 5 presents the success rate of those algorithms for signals with different sparsities. When the sparsity S is smaller than 6 (i.e. five percents non-zero coefficients), all approaches worked quite well. IHT starts to make mistakes frequently when $S = 9$, and the success rate declines dramatically and reaches zero when $S = 15$. In contrast, OMP maintains a perfect reconstruction when $S \leq 18$ (i.e. 14 percent non-zero kernels). The steady decrease of success rate lasts until $S = 63$ where it touches the bottom. SP outperforms the other two strategies by ensuring zero failure until $S = 36$ (i.e. 28 percents non-zero kernels). It drops sharply after that turning point and reaches zero when $S = 63$.

It can be concluded that although these greedy algorithms can recover sparse signals, IHT is only suitable for very sparse cases, OMP can handle more dense signal with acceptable success rate, and SP performs the best which can guarantee successful recovery for the signal with 30 percents non-zero coefficients.

IHT diverges with a higher probability when the sparse cardinality increases since there are more entries associated before truncation, thus the error is more likely to increase compared with small S cases. OMP fails when the kernels selected cannot satisfy the error bound, and there is no more available space for atoms, or if the algorithm starts to diverge. The result suggests the previous is the main reason since the decomposition in the early stage can be inaccurate especially when the sparse cardinality S is large. OMP regards all existing kernels as part of the result without checking the validity, and therefore this estimation-based approach corrects error slowly, which can produce a more substantial error than expected after using up all freedoms. SP improves that by introducing the expanded candidate set $\widetilde{S}$ spanned by current solution and most significant part of the residue, and minimising the error within this subspace to find a proper solution subspace of size S. In this way, the sparse solution approaches the ground truth vertically, and it can manage the reconstruction even if the signal is not very sparse.

OMP approaches the ground truth sparse signal in each iteration by adding one more kernel to the existing sparse solution. In other words, it takes S steps to derive the solution for the original signal with sparsity S, and the complexity is $O(S)$. SP and IHT update the whole solution at a time. Hence the complexity is $O(\log(S))$. Overall, SP is the optimal algorithm among the three on both accuracy and complexity.

IV. Conclusion

In this experiment, we explored a new signal sampling, and recovering approach called compressed sensing by focusing on the accuracy and complexity of reconstruction algorithms. It is proved that as modelled by underdetermined systems, signals that are sparse or nearly sparse after a particular transformation can be recovered from random under-sampled observation satisfying the RIP condition. Also, SP is the optimal sparse signal recovering strategy regarding accuracy and complexity, compared with the inefficient OMP and the unstable IHT algorithm. Moreover, SP is also reliable when the sparse cardinality is large (perfect reconstruction of the signal with 28 percents non-zero kernels by half sampling rate). Nevertheless, the simulation only considers the hard thresholding method of the iterative shrinkage-thresholding algorithm due to the limitation of time. Soft thresholding that is more smooth and stable in iterations can produce a more precise result compared with IHT and deserves attention in the further research. It can be developed based on the code provided in the appendix.

V. Appendix: MATLAB code

The source code can be accessed via https://github.com/SnowzTail/.

References

[1] E. J. Candes, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, vol. 59, no. 8, pp. 1207–1223, 2006.

[2] W. Dai, "Lab note on recover sparse signals from under-sampled observations," September 2018.

[3] ——, "Topics in large dimensional data processing," September 2018.

[4] P. L. Dragotti, "Wavelets and applications," September 2018.

[5] P. L. Dragotti, M. Vetterli, and T. Blu, "Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets strang–fix," *IEEE Transactions on signal processing*, vol. 55, no. 5, pp. 1741–1757, 2007.

[6] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE transactions on information theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[7] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, 2008.

[8] E. J. Candes, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Applied and Computational Harmonic Analysis*, vol. 31, no. 1, pp. 59–73, 2011.

[9] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on*. IEEE, 1993, pp. 40–44.

[10] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE transactions on Information Theory*, vol. 55, no. 5, pp. 2230–2249, 2009.

[11] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Applied and computational harmonic analysis*, vol. 27, no. 3, pp. 265–274, 2009.