Comparison of Compressive Sensing Algorithms.

By Asst.Lact. Ali Jalil Taher

Asst Prof. Dr.Thamer Rashed.
&Asst Prof.Dr .Ahmed abed Alsaheb. Hashem

Abstract. **T**he Shannon/Nyquist sampling theorem specifies that to avoid losing information when capturing a signal, one must sample $\mathbf{F_s}$ at least two times faster than the $\mathbf{F_m}$ signal bandwidth. $\mathbf{(F_s=2\ F_m)}$In many applications, including digital image and video cameras, the Nyquist rate is so high that too many samples result, making compression a necessity prior to storage or transmission. In other applications, including imaging systems (medical scanners and radars) and high-speed analog-to-digital converters, increasing the sampling rate is very expensive. This lecture note presents a new method to capture and represent compressible signals at a rate significantly below the Nyquist rate. This method, called compressive sensing, employs no adaptive linear projections that preserve the structure of the signal; the signal is then reconstructed from these projections using an optimization process. By using different method of Algorithms reconstruction and compression which one is the best?.

نظرية Nyquist (لشانون) معدل نايكوست أخذالعينات يحدد أنه لتجنب فقدان المعلومات عند التقاط إشارة ،يجب على المرءأن يجعل تردد العينةF_s على الأقل مرتين أسرع من تردد عرض الإشارة (النطاق الترددي)F_m. (F_s=2 F_m) في العديد من التطبيقات،بما في ذلك الصورالرقمية وكاميرات الفيديو، فإن تردد معدل نايكوست Nyquist نسبةعالية جدا أن الكثيرمن عينات لذلك،ممايجعل ضغط على ضرورة قبل التخزين أونقل في التطبيقات الأخرى،بمافي ذلك أنظمة التصوير (الماسحات الضوئية الطبية والرادارات) و محولات السرعة التناظرية الرقمية عالي ،زيادة معدلأت اخذالعينات مكلف جدا. هذه البحث يذكرطريقةجديدة لالتقاط وتمثيل القابل للانضغاط إشارات بمعدل بشكل كبيرتحت تردد Nyquist معدل نايكوست معدل هذاالأسلوب، الضغط الاستشعار compressive sensingعن بعد،توظف لا لتكيف خطيا لتوقعات أن تحافظ على بنية

الإشارة؛الإشارةيمكن بناؤها من هذه التوقع اثبت استخدام تحسين العملية. باستخدام مختلف الطرق الخوارزمية ومقارنة افضل خوارزمية .

Index Terms:Compressive Sensing(CS), Sparsity, Signal Recovery, minimization, Greedy Pursuit ,*Subspace Pursuit*(SP), Orthogonal Matching Pursuit (OMP), (***modifiedOMP***) ,Iterative Reweighted Least Square Algorithm (IRLS).and The Compresive Sampling Maching Parsuite Algorithm (CoSaMP).

I. INTRODUCTION .

Reconstruction is an inverse problem in which an image or signal can be recovered from the data given. The given data can be in the form of codes, from which data samples are required to be acquired and processed in order to recover the image or signal accurately. In this paper, the various techniques for image reconstruction using Compressive Sensing are discussed. Compressive Sensing is a highly efficient technique because of the high need for rapid, efficient and less expensive signal processing applications. Earlier, signal processing techniques employed Nyquist-Shannon theorem. This technique requires sampling rate to be at least two times the highest frequency in the signal. This requirement for the sampling frequency is called Nyquist condition. This theorem finds applications in many audio electronics, medical imaging devices, visual electronics, radio transmitters and receivers [1]. However, this technique requires larger storage space, running time and computations. Moreover, in some of the processes, there may be the only limited number of samples available or may be only limited data capturing devices or slow measurements. In such situation, Nyquist-Shannon theorem fails and Compressive Sensing (CS) is able to use the concept of sparsity using different transform domains and incoherent nature of these observations in the original domain. CS offers sampling and signal compression in one step and measures the only minimum number of samples that can carry the maximum set of information [2]. CS is one of the novel techniques that do not require the Nyquist sampling rate for image processing. The implementation of CS reduces the requirement of acquiring and preserving the large number of samples. It saves only those samples, which have significant value and samples with minimal value are discarded. CS offers various applications due to its property of sparsity and incoherence. This paper presents a brief historical description of CS and analyzes the various CS based image reconstruction algorithms. In this paper, section(1) presents the introduction of Compressive Sensing and image reconstruction. In Section(2), background and implementation of Compressive Sensing are discussed with certain properties. The Section (3) presents minimization recovery algorithm. Section (4) presents Matching Pursuit, Orthogonal Matching Pursuit (OMP) and its latter modification OMP

(OMPmod3). Iteratively Reweighted Algorithms for Compressive Sensing (IRLS).(5) Comparision of algorithms of for reconstraction Algorithms of Compressive sensing with number of measurement ***M*** .

Compressive Sensing. The field of CS has gained enormous interest recently. It is basically developed by D. Donoho [3], E. Candes and T. Tao [4]. It was used in Seismology for the first time in 1970 and then later l_1 on norm minimization was suggested by Santosa and Symes for recovery algorithms [5]. Then, total variation minimization was used by Ruden, Fatemi and Osher [6] in 1990s for image processing. The implementation of total variation minimization is somewhere identical to norm minimization. But, the idea of CS was actually taken into account by D. Donoho, E. Candes, Justin Romberg and T. Tao [3, 4].

A. Nyquist Sampling Theorem

Shannon proposed his famous Nyquist-Shannon Theorem in 1949. According to Shannon theorem, any band-limited signal, time varying in nature can be recovered perfectly if the sampling frequency is numerically equal to or greater than two times of the maximum frequency present in the signal itself. For a signal of frequency ***n*** Hertz, sampling rate requires to be 1/2***n*** seconds. For conventional processes, before transmission of the data or signal, it is sampled properly at Nyquist rate followed by compression. For example, if the data is sampled at ***n*** hertz and then compressed to ' ***s***' samples and ***n-s*** samples are discarded. At the recovery part, for most decompression is performed and ' ***n*** ' samples are extracted from compressed ' ***s*** ' data samples. In Fig. 1, ***x*** is the signal of interest and ***y*** is the compressed measurement vector. The figure depicts a model of Traditional Sampling Technique, in which acquisition and sampling are separate steps. In this approach, the number of captured samples remains greater than the information rate. This technique is computationally complex, especially for the high-dimensional signals, due to the requirement of the large storage space. Moreover, analog to digital conversion for high dimensional signals is expensive. The question that arises after studying the Nyquist-Shannon Theorem is that what is the
need of overall computation when only ' ***s*** ' data samples are sufficient.

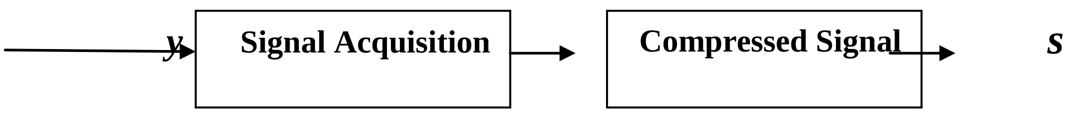


Figure.1-a Structure of Traditional Sampling Technique.

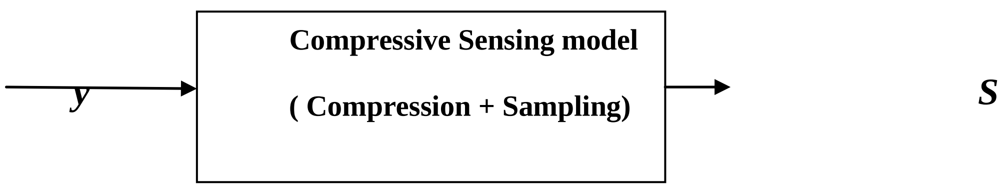


Figure.1-b. Structure of Compressive Sensing paradigm.

The compressed form contains major information; hence there is no need of computing rest of the non-significant samples. So Compressive Sensing is an alternative approach, which performs sampling and compression together as shown in Figure.1-a,b.

B. Compressive Sensing.

CS theory enables the implementation of the recovery of the high dimensional signal from lesser observations as a comparison to the actual number of measurements required in conventional techniques. The objective of CS recovery algorithms is to provide an estimate of the original signal from the captured measurements. It is based on the property of the signals to be able to offer their representation in the sparse domain with fewer numbers of nonzero coefficients. This property is called sparsity and the given signals as sparse signals. The reconstruction algorithm used with CS decides the number of samples needed for exact reconstruction. The model of reconstruction using CS depends on two properties:

1-Sparsity .

Many signals are capable to be stored in compressed form in terms of their projection in a suitable basis. The projected coefficients of these signals can be zero or a far lower value, if a suitable basis is used. For a signal having non-zero coefficients, it is called -sparse. As, these sparse signals may offer the larger number of smaller coefficients that can be ignored easily; hence a compressed signal can be obtained from the sparse form. For compressive Sensing, the suitable domains available are DCT, DWT, and Fourier Transform [7]. Discrete Wavelet Transform is usually preferred over Discrete Cosine Transform because it enables the removal of blocking artifacts [8]. Basically, Sparsity refers to the possibility of having a much smaller information rate for a continuous time signal as a comparison to the one depicted by its bandwidth. So, CS can use the advantage of using these natural signals with their compressed form in a particular domain. Suppose a signal can be represented in a suitable Orthonormal basis like wavelet, DCT. As in a signal, many coefficients are small and most of the important information lies in few larger coefficients. Hence, it can be expanded in an Orthonormal basis for sparse representation. Let x the given signal and $\psi=\{\psi_1\ \psi_2\ \psi_3\ \ldots..\psi_n\}$ represents the suitable basis, therefore, an image x in domain is given as ψ represents the suitable basis, therefore, an image in domain is given as:

$$s(t) = \sum_{i=1}^{n} x_i\, \psi_{i(t)} \quad \ldots\ldots\ldots\ldots\ldots(1)$$

Where S is the coefficients of the sparse form of x, $x_i = \langle S_i, \psi_i \rangle$. In a sparse representation of the signal, small coefficients in that signal can be neglected without much information loss. It's like considering the signal by keeping only the significant coefficients and discarding the smaller coefficients. Thus, the obtained vector is known as a sparse signal.

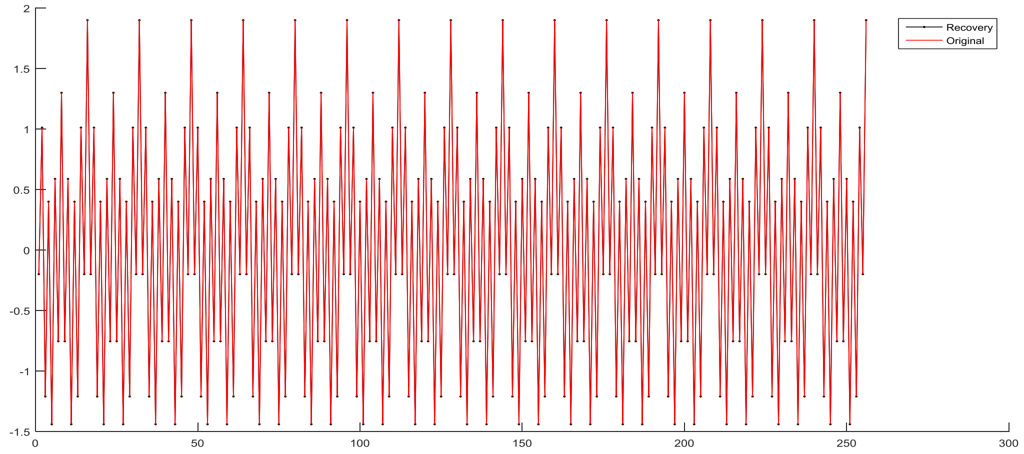


Figure.2. Recovery of the original signal by using Compressive sensing N=256.

II-Incoherence.

Incoherence shows that any signal with a sparse representation in a particular domain can be spread out in a domain in which it is actually captured. It enables the relationship of duality between time and frequency. It measures the maximum correlation for any two matrices. These two matrices give a form of different representation domains. For the measurement matrix $\boldsymbol{\Phi}$ with size $\boldsymbol{M}\times\boldsymbol{N}$ and the representation matrix $\boldsymbol{\Psi}$ of size $N\times N$, the representation matrix can be represented as $\Psi_1 \ \ldots\ \Psi_n$ as columns and measurement matrix as $\Phi_1 \ldots\ldots\ \Phi_n$ as rows. The coherence is given as.

$$\mu(\boldsymbol{\Phi}, \boldsymbol{\Psi}) = \sqrt{n}\ \max \left| \boldsymbol{\Phi}_k , \boldsymbol{\Psi}_j \right| \quad \ldots\ldots.(2)$$

for $1 \le \boldsymbol{j} \le n$, and $1 \le \boldsymbol{k} \le m$. Moreover, from linear algebra, for incoherence following result can be depicted.

$$1 \le \mu(\boldsymbol{\Phi}, \boldsymbol{\Psi}) \le \sqrt{n} \quad \ldots\ldots\ldots\ldots\ldots...(3)$$

In CS technology, the incoherence of two matrices is important. One is the Sensing matrix that is used to sense the significant columns of the signal of interest. The second one is the representation matrix $\boldsymbol{\Psi}$ in which the given signal is represented in the sparse form. The low value of incoherence for CS shows that the fewer measurements are required for reconstruction of the signal [9].

Coherence is able to measure the maximum correlation between the columns or elements of $\boldsymbol{\Psi}$ and $\boldsymbol{\Phi}$. Mostly low coherence pairs are considered in Compressive Sensing. The measurement matrix $\boldsymbol{\Phi}$ basically performs the function of sampling the coefficients. The measurement matrices like Fourier, Gaussian are able to satisfy the coherence property. The random matrices like i.i.d (independent identically distribution)Gaussian Matrix or binary ±1 matrix with fixed basis $\boldsymbol{\Psi}$ are mainly incoherent. These matrices are simple and possess lower convergence, which are required for recovery with Compressive Sensing.

III- <u>Restricted Isometry property (RIP)</u> .

RIP is one of the most important properties of Compressive Sensing. It can be used as a tool in order to analyze the performance implementation of CS. RIP tool is able to ensure that the sensing matrix captures the significant columns from the given signal of interest. This property can give the condition for the reconstruction matrix for exact recovery. For any sparse $\boldsymbol{k}$ signal $\boldsymbol{S}$, restricted isometric constant $\boldsymbol{\delta_k}$ (***RIC***) ($0 < \boldsymbol{\delta_k} < 1$) is the minimum value that makes the inequality holds [10].

$$(1-\boldsymbol{\delta_k}) \| \boldsymbol{S} \|_2^2 \leq \| \boldsymbol{\Theta} \boldsymbol{S} \|_2^2 \leq (1+\boldsymbol{\delta_k}) \| \boldsymbol{S} \|_2^2 \quad \ldots\ldots(4)$$

When this property holds, then the matrix $\boldsymbol{\delta}$ can preserve the Euclidean length of - $\boldsymbol{k}$ sparse signals, which shows that the given sparse vectors are not in the null space. It can be considered as the sufficient condition for recovering the estimate of the support for the sparse signal. If ***RIC*** constant is not really close to 1, then the property ensures that all the columns selected are orthogonal and the sparse signal is not in the neglected part of the matrix which is sensed. Otherwise, it cannot be reconstructed. For the relation of CS and RIP property, it is shown that if -sparse signals are required to be recovered, then should be lesser than 1. This ensures the recovery of $\boldsymbol{k}$ sparse signals from observations vectors with compressive Sensing. The Restricted Isometry Property can be considered as the form of uniform uncertainty principle for many ensembles of random matrices like Fourier, Gaussian, and Bernoulli [4, 11].

These matrices satisfy the ***RIP*** condition with parameters $\boldsymbol{\delta} \in (0,1\backslash 2)$ which can be represented by Eq. (5)..

$$M = k \log^{o(1)} N \quad \ldots\ldots..(5)$$

Where ***M*** is the number of measurements and ***N*** be the dimension of the signal, ***k*** is the sparsity of the signal. Therefore, a signal can be recovered exactly for ***M*** measurements, ***k*** is the sparsity. In some applications, measurement matrices play an important role. Some of the measurement matrices include Fourier matrix having randomly selected rows. The complexity, in this case, can be reduced by using Fast Fourier transform. Bernoulli matrix which can also be used as measurement matrix has faster Sensing but storage complexity. The Gaussian random matrix is another measurement matrix which is derived from the normal distribution with zero mean and variance 1/***N***. Gaussian matrices are very simple and useful operators for Compressive Sensing. Gaussian matrix is mostly preferred because of its simplicity and easy implementation.

XI- Compressive Sensing Model .

Compressive Sensing model basically performs compression and sampling simultaneously. Considering an ***N*** dimensional signal , the sparse form of the signal can be constructed by representing it in any suitable basis like DCT, Fourier Transform, and wavelet Transform. The sparse form or the signal of interest can be given as:

$$\boldsymbol{x} = \boldsymbol{\Psi}\, \boldsymbol{s} \quad \ldots\ldots\ldots..(6)$$

Where $\boldsymbol{x}$ is the sparse form of $\boldsymbol{x}$ $\boldsymbol{s}$ and $\boldsymbol{\Psi}$ is the suitable basis that shows the projection coefficients of $\boldsymbol{x}$ on the given basis. The next step is to compute the measurement vector $\boldsymbol{y}$ with a suitable matrix either Gaussian [12] or Bernoulli [13]. The measured vector can be given as:

$$\boldsymbol{y} = \boldsymbol{\Phi}\, \boldsymbol{x} \quad \ldots\ldots\ldots..(7)$$

where $\boldsymbol{\Phi}$ is the measurement matrix of dimension $\boldsymbol{M} \times \boldsymbol{N}$. The overall Eq. can be represented as:

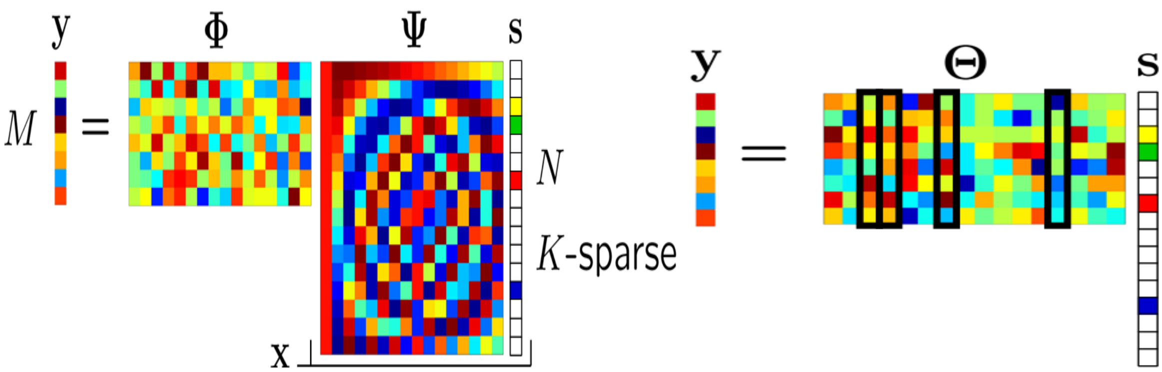
$$\boldsymbol{y} = \boldsymbol{\Theta}\, \boldsymbol{s} \quad \ldots\ldots\ldots..(8)$$

where Θ is the Sensing matrix and is depicted as $\Theta = \Phi\ \Psi$, it is also known as reconstruction matrix. In practical applications, the measurement or the random noises can also be considered. Equations (6) and (7) are reformulated as: Figure 3.

$$y = \Phi\, x + e \quad \text{..........(9)}$$

$$y = \Theta\, s + e \quad \text{..........(10)}$$

where e represents random noise vector. Hence, the primary objective of CS is to recover the signal from these captured measurements, under sparsifying conditions. Then, the recovery algorithms are applied on the given measurement vector. The recovery algorithms available are L_1 minimization, several Greedy algorithms which are discussed in the next section.



(a) (b)

Figure 3: (a) Compressive sensing measurement process with (random Gaussian) measurement matrix Φ and Discrete Cosine Transform (*DCT*) matrix Ψ. The coefficient vector s is sparse with K = 4. (b) Measurement process in terms of the matrix product $\Theta = \Phi\ \Psi$ with the four columns corresponding to nonzero s_i highlighted. The measurement vector y is a linear combination of these four columns.

X- L_1 Minimization.

For recovering the sparse signal from its observed measurements, recovery algorithms are required. Convex relaxation is the class of algorithms which can solve reconstruction issues through linear programming [14]. The number of measurements needed for these techniques is small, but the methods are complex in computational terms. The recovery problem needs to solve the highly convex problem according to (11).

$$\min \|s\|_0 \ \text{subject to} \quad y = \Phi\, s \quad \text{..........(11)}$$

It was later suggested by Donoho and his associates that for given measurement matrix, L_0 as an NP- hard problem can be considered equivalent to its convex relaxation i.e. L_1 minimization.

$$\min \| s \|_1 \text{ subject to } y = \Phi s \quad \ldots\ldots\ldots.(12)$$

Where $\| s \|_1 = \Sigma_i | s_i |$ denotes the L_1- norm. This kind of convex reconstruction problem can be solved by linear programming. For minimization, the measurement matrix must satisfy ***RIP*** property for a small value of Restricted Isometric Constant (***RIC***). ***BP*** (Basis Pursuit) [15], BP de-noising are some of the examples of minimization techniques. It allows the recovery of norm minimization instead of recovering the low-rank matrix from the compressed form of the signal. This algorithm reduces the complexity, dimensions and order of the system because low-rank matrix signifies low order and less complex system. Basis pursuit works as an optimization problem and minimizes the cost function constructed by the Lagrange multipliers.

$$\min \| y - \Phi \check{v} \|_2 + \lambda \| s \|_1 \quad \ldots\ldots\ldots.(13)$$

Where λ is a positive constant. The algorithm works in simple steps by first initializing the sparse signal and providing the cost function. The Sensing matrix is computed by selecting the required number of measurements and then the coefficients of the sparse image are modified in order to obtain the minimized form. The iterations work until the number of total significant components remains less than k , sparsity of the signal.

Broad categorisation of reconstruction approaches

Greedy Pursuits

Convex Optimization (L_1 minimization)

Nonconvex Minimization (L_P minimization, $0<p<1$)

Bayesian Methods

1- <u>Iterative Reweighted Least Square Algorithm (IRLS).</u>

NONCONVEX COMPRESSED SENSING BASED ON IRLS

An easy way to comply with the conference paper formatting requirements is to use this document as a template and simply type your text into it. Replacing L_1 norm by L_p norm with the value of $p < 1$ exact reconstruction is possible with much less measurements. When we replace L_1 norm by L_p norm with p lying in the range $0 < p < \infty$, then we have $\|S\|_p^p = \Sigma_j^n \mid S_j \mid^p$, Unlike $\|.\|_0$, $\|.\|_p$ is not a norm when $0 < p < 1$, though $\|.\|_p$ satisfies triangle inequality and induces a metric [10]. The resulting optimization problem is nonconvex. L_p norm is a much simpler way of finding a local minimizer to produce exact reconstruction of sparse signals with many fewer measurements than when $p = 1$. The recovery of signal from fewer measurements is possible if the number of measurement is greater than or equal to the value $m \geq a$ for a constant c. The incoherence μ is kept as low as possible if the incoherence is 1 we don't need more than k log n samples for a constant c. L_p minimization can further reduce the value of the constant. The objective is to compute a global minimizer of a nonconvex functional. On the other hand, any local optimization method can do so if initialized by a point sufficiently close to the global optimum. Nonconvex approach increases robustness to noise and image nonsparsity. Several methods were being proposed as an alternative to linear programming method, Iterative Reweighted Least Square method is one of them. An IRLS algorithm appeared for the first time in the works of Lawson [11] in the form of algorithm for solving uniform approximation problems. It is seen that IRLS obeys a linear convergence rate. It started as if the solution of L_1 norm has no nonzero coordinates then the solution of the weighted least squares problem is

$$S^{X} = argmin\|S\|_{L2}n(x), \quad x=(x_1, x_2, .., x_n) \qquad \dots (14)$$

For the weights $x_j = \mid s_j' \mid^{-1}$ coincides with x'. Since we do not know s' this observation cannot be used directly. So we choose the first weight x_0. after which (1) is solved. We then use this solution to define a new weight x_1 and repeat this process [12]. We will be using IRLS algorithm that gives reweighting and avoids infinite components in the weight with the help of regularization [14], [15], [16].

The iterative way of solving L_p minimization problem:

$$\min \|s\|_p , \quad \Phi s = y$$

for $0 < p < 1$. IRLS iteratively minimizes a weighted L_2 function to generate $\min_x \Sigma^n_{i=1} x_i s_i^2$. The solution of this equation is given as

$$s^n = Q_n \Phi^T (\Phi Q_n \Phi^T)^{-1} y$$

Where Q_n is a diagonal matrix with entries $1/x_i = \mid s_i^{n-1} \mid^{(2-p)}$, this is found by Euler-Lagrange for (2).If we set p to 0 the objective function becomes $\Sigma \log(\mid s_i^2 \mid)$.

As (p – 2) will be negative the weight x_1 become undefined whenever s_i^{n-1} becomes 0. To deal with such situation we regularize the problem by introducing a factor $\epsilon > 0$. So now the weights become $x_i = (|s_i^{n-1}|^2 + \epsilon)^{(p/2-1)}$ [14].

Algorithm of IRLS .

Step1. Initialization

Step2. Input-k sparse noisy data s

Step3. Precomputation of y.

$$y = \Phi s \quad .$$

Step4. Update of weights for each iterations.

$$x = (s_{old}^2 + \epsilon)^{(p/2-1)}$$

$$t = 1/x.$$

Step5. Solving the weighted L_2 minimization

$$Q = diagonal\ (t).$$

$$r = inverse(\Phi * Q * \Phi^T)$$

$$s_{new} = Q * \Phi^T * r * y$$

2-The Compresive Sampling Maching Parsuite Algorithm (CoSaMP).

The most difficult part of signal reconstruction is to identify the locations of the largest components in the target signal. CoSaMP uses an approach inspired by the restricted isometry property. Suppose that the sampling matrix Φ has Restricted Isometric Constant $\delta_s \ll 1$. For an s-sparse signal s, the vector $y = \Phi^* \Phi s$ can serve as a proxy for the signal because the energy in each set of s components of y approximates the energy in the corresponding x components of s. In particular, the largest s entries of the proxy

y point toward the largest s entries of the signal s. Since the samples have the form $u = \Phi x$, we can obtain the proxy just by applying the matrix Φ^* to the samples. The algorithm invokes this idea iteratively to approximate the target signal. At each iteration, the current approximation induces a residual, the part of the target signal that has not been approximated. As the algorithm progresses, the samples are updated so that they reflect the current residual. These samples are used to construct a proxy for the residual, which permits us to identify the large components in the residual. This step yields a tentative support for the next approximation. We use the samples to estimate the approximation on this support set using least squares. This process is repeated until we have found the recoverable energy in the signal.

Overview. As input, the CoSaMP algorithm requires four pieces of information: • Access to the sampling operator via matrix–vector multiplication. • A vector of (noisy) samples of the unknown signal. • The sparsity of the approximation to be produced. • A halting criterion. The algorithm is initialized with a trivial signal approximation, which means that the initial residual equals the unknown target signal. During each iteration, CoSaMP performs five major steps:

(1) Identification. The algorithm forms a proxy of the residual from the current samples and locates the largest components of the proxy.

(2) Support Merger. The set of newly identified components is united with the set of components that appear in the current approximation.

(3) Estimation. The algorithm solves a least-squares problem to approximate the target signal on the merged set of components.

(4) Pruning. The algorithm produces a new approximation by retaining only the largest entries in this least-squares signal approximation.

(5) Sample Update. Finally, the samples are updated so that they reflect the residual, the part of the signal that has not been approximated. These steps are repeated until the halting criterion is triggered. In the body of this work, we concentrate on methods that use a fixed number of iterations. Appendix A discusses some other simple stopping rules that may also be useful in practice. Pseudocode for CoSaMP appears as Algorithm 6. This code describes the version of the algorithm that we analyze in this paper. Nevertheless, there are several adjustable parameters that may improve performance: the number of components selected in the identification step and the number of components retained in the pruning step.

Algorithm 2: CoSaMP Recovery Algorithm.

CoSaMP(Φ,u,S)

Input :Sampling Φ, noise sample vector u, sparsity level S.

Output : An ***s***-sparse approximation ***a*** of the target signal .

$\boldsymbol{a^0} \leftarrow \mathbf{0}$ {Trivial initial approximation}

$\boldsymbol{v} \leftarrow \boldsymbol{u}$ { Current sample= input sample }

repeat

k ← k+1

$\boldsymbol{y} \leftarrow \boldsymbol{\Phi^* v}$ { form signal proxy }

$\boldsymbol{\Omega} \leftarrow \mathbf{supp}(\boldsymbol{y_{2s}})$ { identify large components }

$\mathrm{T} \leftarrow \boldsymbol{\Omega}\ \mathbf{U}\ \mathbf{supp}(\mathbf{a^{k-1}})$ { Merge supports }

$\boldsymbol{b} \,|\, \boldsymbol{T} \leftarrow \boldsymbol{\Phi}^{\dagger}_{\boldsymbol{T}}\, \boldsymbol{u}$ { Signal estimation by least squares }

$\boldsymbol{b} \,|\, \boldsymbol{T}^{c} \leftarrow \mathbf{0}$

$\boldsymbol{a^k} \leftarrow \boldsymbol{b}_{\mathrm{s}}$ { prune to obtain next approximation }

$\boldsymbol{v} \leftarrow \boldsymbol{u} - \boldsymbol{\Phi a^k}$ { Update current samples }

Until halting criterion true.

CoSaMP Hypotheses

- The sparsity level ***s*** is fixed.
- The ***M***× *N* sampling operator **Φ** has restricted isometry constant $\boldsymbol{\delta_{4s}} \leq 0.1$.
- The signal $\boldsymbol{x} \in \mathrm{C}^{N}$ is arbitrary, except where noted. [15]
- The noise vector $\boldsymbol{e} \in \mathrm{C}^{\boldsymbol{m}}$ is arbitrary.
- The vector of samples $\boldsymbol{u} = \boldsymbol{\Phi}\, \boldsymbol{x} + \boldsymbol{e}$.

3-. <u>Orthignal Matching Pursuit (OMP)</u>.

The greedy algorithms are iterative approaches; capable of recovering the images using Compressive Sensing. Block based Compressive Sensing enables the use of Greedy algorithms for recovering thermal images also [16]. It works in an iterative fashion in order to recover the sparse signal. Orthogonal matching pursuit (OMP), a greedy algorithm, basically comes into account in the form of a variant of MP

(Matching Pursuit) algorithm. Earlier, the MP algorithm, an iterative greedy method, was introduced in order to approximate the decomposition [17]. It identifies those bases and their coefficients; that can construct the input signal on combining. It starts with the assumption that all the bases are orthogonal to each other i.e. independent of each other. The value calculated by correlation of the given signal with the basis gives the influence of basis on the signal. For the basis to be the important part of the signal; correlation value should be high and for a lower value of the correlation, the basis has a negligible contribution. It works by selecting the elements having the maximum correlation with the residual vector throughout the algorithm at each step.

Let $\boldsymbol{\Phi_i} < \boldsymbol{i} < \boldsymbol{N}$ be the element having the strongest correlation with residual $\boldsymbol{r}$ denoting ***ith*** the element of measurement matrix $\boldsymbol{\Phi}$. All atoms are assumed to be normalized with value unity. The selection of elements by the algorithm, at each iteration, is represented as

$$\boldsymbol{\Phi_k} = \boldsymbol{argmax}_{1<i<N} \left| \langle \boldsymbol{r_{k-1}} , \boldsymbol{\Phi_i} \rangle \right| \ldots\ldots\ldots\ldots(15)$$

where the inner product is denoted by $\langle . \rangle$ and $\boldsymbol{r_{k-1}}$ shows the residual at $\boldsymbol{k}$**-1** iteration. The algorithm then updates the residual depending on the column selected till the termination criterion occurs. When the halting condition occurs, the algorithm stops i.e. when the value of the norm of residual becomes lesser than the predefined threshold or error bound. The algorithm may stops even when the number of distinctly selected elements in the approximation set becomes equal to the desired limit. The Matching Pursuit algorithm is used mostly because of its simplicity. But it suffers from the drawback of slow convergence and poor sparse results.

The Orthogonal Matching Pursuit (OMP) [17, 18] has the capability to remove this drawback by projecting orthogonally the signal on the subspace corresponding to the selected set of columns. The method of selecting the elements remains the same in both the algorithm. As OMP is based on Orthogonalization, an atom is selected only once throughout the algorithm. Y. C. Pati *et al.*, demonstrated OMP as the recursive algorithm, to calculate the functions in the form of the non-orthogonal basis using wavelet frames.

OMP constructs the orthogonal projection onto the observation vector. Before the orthogonal projection, the algorithm computes the inner product of the residue and the measurement matrix. Then the coordinate of the highest magnitude is selected and the column corresponding to that coordinate is extracted. These columns are then embedded into the selected set. It offers better asymptotic convergence as comparison to the conventional MP algorithm. OMP algorithm is considered as a powerful and fast recovery algorithm for the sparse signal from the random measurements. It can provide recovery results for M measurements and N dimension with $O(M \ln N)$ random measurements. It is a huge improvement over the conventional recovery algorithms. OMP algorithm has an evident place because of its speed and easier implementation. Consider a given signal x of dimension N. Select the number of measurements required i.e. M. Construct measurement matrix Φ of dimension M x N. Then, obtain the sparse form of the signal and construct the observation vector $y = \Phi x$, this measurement vector is of dimension M, hence obtaining the compression and sampling in one step. OMP algorithm works step by step. The OMP algorithm searches for the significant column in the inner product of the residue and the measurement matrix. Then, the orthogonal projection is computed onto the subspace of the measurement vector. This orthogonal projection gives the estimate of the support set. The residue is updated according as the estimate. The iterations continued until the residue is lesser than the specified threshold. Finally, the recovered signal is obtained in terms of the estimate of the support system. For the termination of the algorithm, the norm of the residue can be checked with respect to the threshold. Here, the residual r_t is always orthogonal to the columns of the measurement matrix. The conditions for termination of the algorithm can be based on cumulative coherence property [19]. This algorithm selects a new significant column at each step. The column selected corresponds to the index having maximum inner-product. Then, this selected column is augmented with the initialized set. The residual is updated with each step for the new estimate of the support set. Finally, the estimate of the sparse signal is computed with the updated residual. The running time of the algorithm is depicted by the step of identifying the new indices. A prototype of OMP algorithm first came into account in 1950 [20, 21]; later on it was investigated further for its implementation in recovering the sparse image from random measurements with noise in year 2011 by T. Tony Cai and Lie Wang [22]. The OMP algorithm was studied with bounded noise and its recovery performance is checked under such condition. The implementation of OMP provides an effective way of recovering the sparse signal from random observations even in the presence of noise with appropriate measurement matrix. The steps of the OMP algorithm are given as follows:

4-*Modifying the OMP Algorithm.* (*modifiedOMP*)

We have introduced two BP based reconstruction algorithms considering prior knowledge and their limitations. To overcome these, we propose a modified OMP algorithm, which will be able to deal with the three categories separately. OMP is a greedy algorithm. It reconstructs original signal by iterative search for non-zero indices and performs leastsquares estimation of the values on the non-zero indices. The original OMP algorithm [13] is as follows: The iteration operation gives us the freedom to consider three categories separately. Since there must be non-zeros in $\boldsymbol{S}_1$, we can initialize the original $\boldsymbol{\Lambda_0}$ with $\boldsymbol{S}_1$. This modification benefits the reconstruction accuracy because occupied indices will always be counted. Then, during each iteration, if index λ_t is found and if $\lambda_t \in \boldsymbol{u}_i \subset \boldsymbol{S}_2$ as well, all elements in $\boldsymbol{u}_i$ will be added to $\boldsymbol{\Lambda}_t$. This modification benefits the reconstruction accuracy because only selected $\{\boldsymbol{u}_i\}$ are counted. If index $\lambda_t \in \boldsymbol{u}_i \subset \boldsymbol{S}_3$, only λ_t will be added, similar to the original OMP process. As a result, all three categories are treated separately. The modified OMP algorithm is summarized as:

THE ORIGINAL OMP ALGORITHM.

INPUT:

- **An $M \times N$ matrix $\Theta = \Phi$**
- **An $M \times 1$ sample vector y**
- **Maximum number of iterations M.**
- **Error tolerance η**

OUTPUT:

- **An estimate $N \times 1$ vector f̂ for the ideal signal**
- **An index set Λ_t containing t elements from $\{1, ...,N\}$**
- **An $M \times 1$ residual vector res_t**

PROCEDURE:

1) Initialize the residual $res_0 = y$, the index set $\Lambda_0 = \phi$ and the iteration counter t = 1.

2) Find the index λ_t that satisfies the following equation

$$\lambda_t = \arg_{max\ j=1,...,N} |\langle res_{t-1}, \theta_j \rangle| \quad (16)$$

Where θ_j denotes the j-th column vector of Θ, $\langle a, b \rangle$ denotes the inner product of two vectors a and b. If the maximum occurs for multiple indices, break the tie deterministically.

3) Augment the index set $\Lambda_t = \Lambda_{t-1} \cup \{\lambda_t\}$. $\Theta_1 = \Theta_{\Lambda t}$ and the matrix of chosen atoms $\Theta_1 = \Theta_{\Lambda t}$. where $\Theta_{\Lambda t}$ is a sub-matrix of Θ containing columns with indices in Λ_t.

4)Solve the least-squares problem to obtain a new signal estimate:

$x_t = arg_{\, x\, min} \| \Theta_t x - y \|_2$........ (17)

5) Calculate the new residual:

$res_t = y - \Theta_t x_t$ ....(18)

6) If $t < m$ or or $\| res_t \|_2 > \eta$, increment t, and return to Step 2.

7) x_t is the estimated signal f ^, with non-zero indices at components listed in Λ_t.

4-*THE MODIFIED OMP ALGORITHM*.(ModifiedOMP).

INPUT:

- **An $M \times N$ matrix $\Theta = \Phi$**
- **An $M \times 1$ sample vector y**
- **Boundary information $\{b_1, b_2, ..., b_K\}$, $\{u_1, u_2, ..., u_K\}$ and $\{S_1, S_2, S_3\}$**
- **Maximum number of iterations M.**
- **Error tolerance η**

OUTPUT:

- **An estimate $N \times 1$ vector f ^ for the ideal signal**
- **An index set Λ_t containing t elements from $\{1, ...,N\}$**
- **An $M \times 1$ residual vector res_t**

PROCEDURE:

1) Initialize the residual $res_0 = y$, the index set $\Lambda_0 = S_1$, the matrix of chosen atoms $\Theta_1 = \Theta_{S1}$ and the iteration counter $t = 1$.

2) Solve the least-squares problem in Equation (17) to obtain a new signal estimate.

3) Calculate the new residual using Equation (18).

4) Increment t.

5) Find the index λ_t that satisfies Equation (16).

6) Augment the index set $\Lambda_t = \Lambda_{t-1} U \{ \lambda_t \}$. If $\lambda_t \in u_i \subset S_2$, let $\Lambda_t = \Lambda_t U u_i$.

7) Set the matrix of chosen atoms $\Theta_1 = \Theta_{\Lambda t}$.

8) Solve the least-squares problem in Equation (15) to obtain a new signal estimate.

9) Calculate the new residual using Equation (18).

10) Return to Step 4 if $t <$ m or $\| res_t \|_2 > \eta$.

11) x_t is the estimated signal f ^, with non-zero indices at components listed in Λ_t.

5-*Subspace Pursuit*(*Algorithm* SP) .

The subspace pursuit (SP) algorithm was developed by Dai and Milenkovic [16] and published in 2009. This algorithm is a ***p***-pursuit algorithm and according to the original paper with the underlying algorithmic principle borrowed from the A∗ order-statistic algorithm [18]. Another algorithm, called CoSaMP [17] closely resembles SP and was developed simultaneously as the latter. The difference between these two algorithms lies in the extension phase, which we will get back to in the description of SP. Each iteration in SP requires more computational effort than each iteration in OMP, but it turns out that in practice SP iterates fewer times. By experimental evaluations it can be observed that the two(Greedy Pursuit) GP algorithms SP and OMP provide for similar execution times and similar performance. We show the SP algorithm in Algorithm *Subspace Pursuit(SP)*. In the initialization phase of Algorithm *Subspace Pursuit(SP)*, SP uses the K_{max} largest in-amplitude components from the matched filter as a first estimate of the support-set T_0. Also the iteration counter k is reset and the first residual r_0 is calculated. In the k'$_{th}$ iteration, SP evaluates the correlation filter, $A^T r_{k-1}$, identifies the indices corresponding to the K_{max} largest amplitudes and stores this in a temporary set, T_Δ (step 3). Then, the union between the old support-set and T_Δ is formed in step 4. We call this the extension phase and notice that the support-set T' is at most of size $2K_{max\ 2}$. The Algorithm solves a ls problem with the selected indices of T' (requirement: $|T'| \leq M$) and identifies a new support-set corresponding to the K_{max} largest amplitudes in $\tilde{x}$ (step 5 and 6). The algorithm then computes the residual

(step 7), which is used as a stopping-criterion for the iteration phase.

The support-set T *is a set of indices corresponding to the non-zero (active) components in the sparse vector* x*:*

$T \overset{\Delta}{\equiv} \{i : x_i \neq 0\}$. *Its complement is denoted by* $\bar{T}$

$\bar{T} \overset{\Delta}{\equiv}, \{i : x_i = 0\}$.

$$resid(y,A) \overset{\Delta}{\equiv} y - AA^\dagger y,$$

Clearly, $T \cup \bar{T} = \{1, 2, ...,N\}$ and $T \cap \bar{T} = \emptyset$. Furthermore, we can pick all the active components in x and place them one-by-one in a column-vector $x_T = \{x_i : x_i \neq 0\}$ which has size $\| x \|_0 = |T| \leq K$. Similarly, we can take the active column vectors in A and form the matrix

$A_T = \{a_i : x_i \neq 0\}$.

$$A = \Phi$$

Φ= Measurement Matrix.
the support-set T' is at most of size 3 K_{max}.

Algorithm : Subspace Pursuit (SP).

Input: $A = \Phi$, $K_{max} = M$, y
K_{max}= maximum number of iterations.
Φ= Measurement Matrix.
Initialization:
1: $T_0 \leftarrow$ *max indices* $(A^T y, K_{max})$
2: $r_0 \leftarrow resid(y, A_{T0})$
3: $k \leftarrow 0$
***Iteration*:**
1: repeat
2: $k \leftarrow k + 1$
3: $T_\Delta \leftarrow$ *max indices* $(A^T r_{k-1}, K_{max})$
4: $T' \leftarrow T_{k-1} \cup T_\Delta$
5: $\tilde{x}$ such that $\tilde{x}_{T'} = A^\dagger_{T'} y$, and $\tilde{x}_{\bar{T'}} = 0$.
6: $T_k \leftarrow$ max indices$(\tilde{x}, K_{max})$
7: $r_k \leftarrow resid(y, A_{Tk})$
8: until $(\|r_k\|_2 \geq \|r_{k-1}\|_2)$
9: $k \leftarrow k - 1$ ('Previous iteration count')
Output:
1: $\hat{T} \leftarrow T_k$
2: $\hat{x}$ such that $\hat{x}_{Tk} = A^\dagger_{Tk} y$ and $\hat{x}_{\bar{T}k} = 0$
3: $n_r \leftarrow \|r_k\|_2$

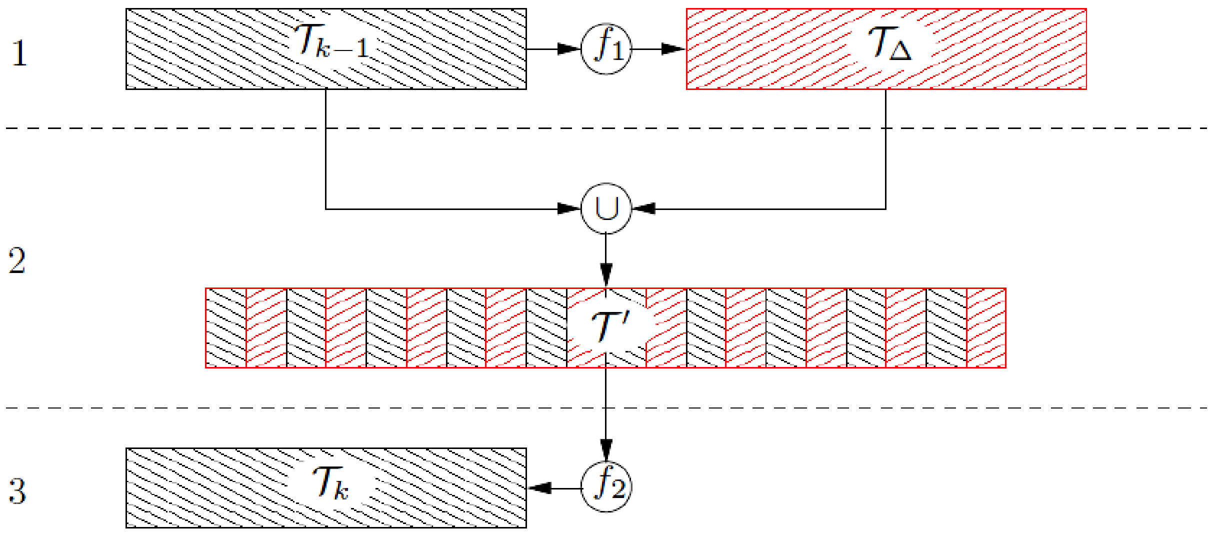


Figure 4. Support-set evolution in Subspace Pursuit(SP)

Simulation Result.

In this study we consider the Compression of five Recovery Algorithms (IRLS¸OMP modified***OMP*** ¸CoSaMP and SP). Figure 5and Figure 6.
To get the Best One with less number of computations.(***M***) With Empirical probability of Reconstruction at level 10^{-8} ,***N***=64¸ and 10^{-6} , ***N***=81 .

the theoretical worst-case performances of the different algorithms:
$SP: O(K(MN + KM)) \sim O(KMN)$.

$OMP: O(K(MN + K_2 + KM) \sim O(KMN)$.

Where the ***O***( ) Mean Optimization operations.

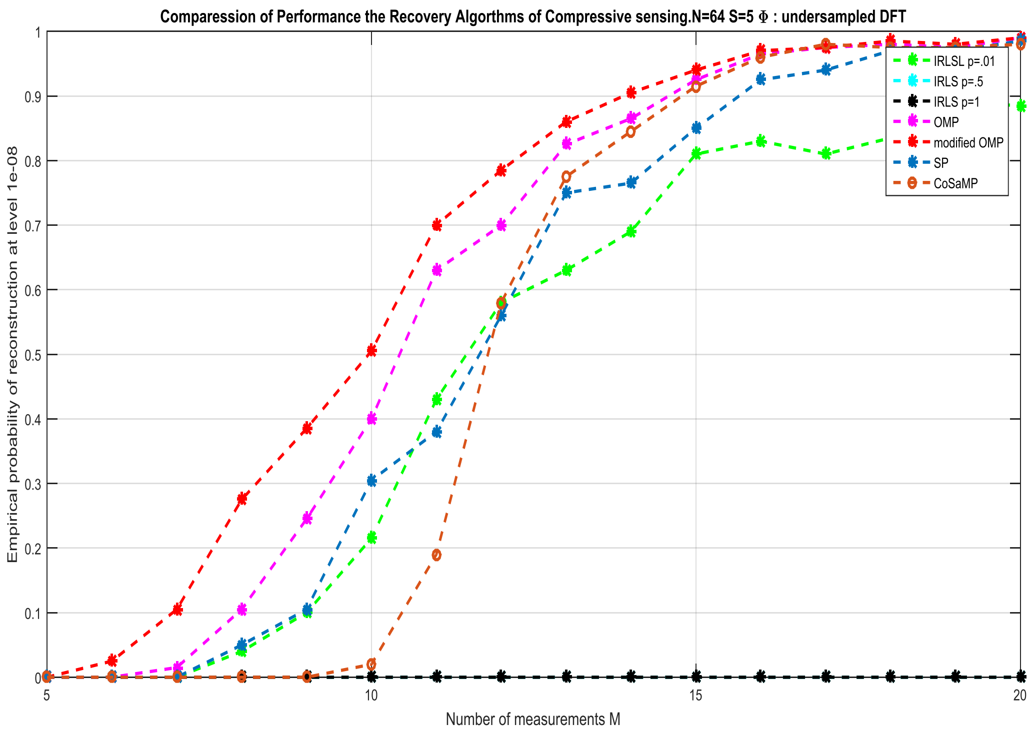


Figure 5. compression of performance the recovery Algorithms of compressive sensing *N*=64,*S*=5.with number of measurements *M*=20 and Empirical probability of reconstruction At level 1*10^{-8} .

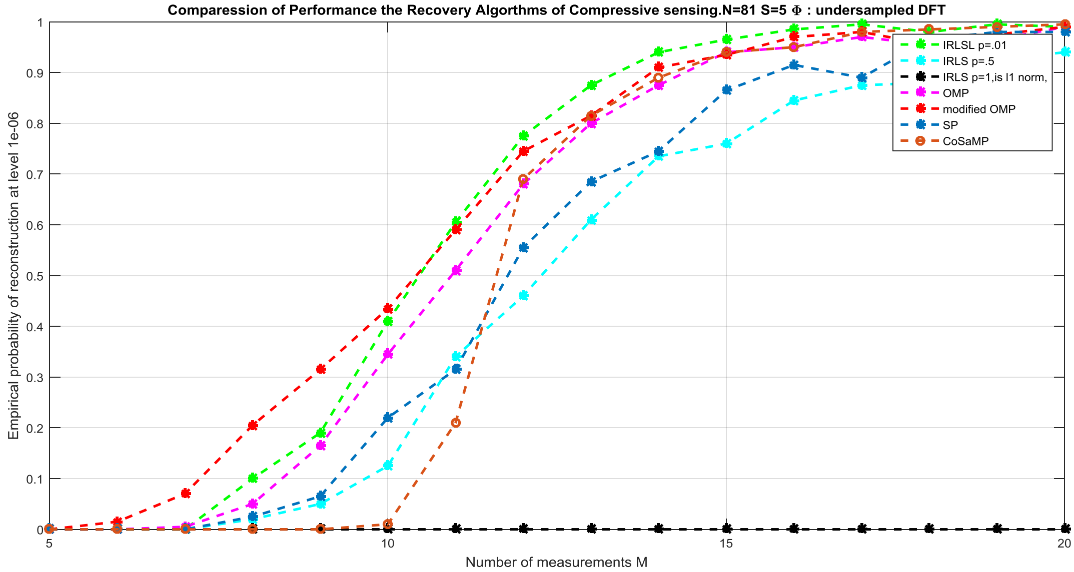


Figure 6. compression of performance the recovery Algorithms of compressive sensing *N*=81,*S*=5.With number of measurements *M*=20 and Empirical probability of reconstruction At level 1*10^{-6}.

References.

[1] D. L. Donoho, "Compressed Sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289-1306, 2006.

[2] E. J. Candes and M. Wakin, "An Introduction to Compressive Sampling," *IEEE Signal Process. Magazine*, vol. 25, no. 2, pp.21- 30, March 2008.

[3] E. J. Candes, "Compressive Sampling," *In Proceedings of the International Congress of Mathematicians*, pp. 1433-1452, Madrid, Spain, 2006.

[4] E. J. Candes, J. Romberg and T. Tao, "Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489-509, February 2006.

[5] E. J. Candes, J. Romberg and T. Tao, "Stable Signal Recovery from Incomplete and Inaccurate Measurements," *Comm. Pure Appl. Math*., vol. 59, no.8, pp. 1207-1223, 2006.

[6] E. J. Candes and J. Romberg, "Sparsity and Incoherence in Compressive Sampling," *IEEE signal processing magazine*, vol. 23, no. 3, pp. 969-985, 10th April 2007.

[7] E. J. Candes and T. Tao, "Reflection on Compressive Sensing," *IEEE signal processing magazine*, vol. 58, no. 4, pp. 14-18, December 2008.

[8] R. Baraniuk, "Compressive Sensing," *IEEE Signal Processing magazine*, vol. 24, no. 4, pp. 118-121, July 2007.

[9] E. J. Candes and T. Tao, "Near Optimal Signal Recovery from Random Projections: Universal Encoding Strategies?," *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5406-5425, 2006.
[10] R. Chartrand, "Exact Reconstruction of Sparse Signals via Nonconvex Minimization," *Signal Process. Lett.*, vol. 14, no. 10, pp. 707-710, 2007.
[11] C. L Lawson, *Contributions to the Theory of Linear Least Maximum Approximation*, PhD thesis, University of California, Los Angeles, 1961.
[12] I. Daubechies, R. DeVore, M. Fornasier, and S. Gunturk, "Iteratively Reweighted Least Squares minimization: Proof of Faster than Linear Rate for Sparse Recovery," *42nd Annual Conference on Information Sciences and Systems, CISS 2008*, pp. 26-29,March 2008.
[13] M. R. Osborne, *Finite Algorithms in Optimization and Data Analysis*, John Wiley and Sons Ltd., Chichester, 1985.
[14] R. Chartrand and W. Yin, "Iteratively Reweighted Algorithms for Compressive Sensing," *Proceedings of the 33rd International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 3869-3872, 2008.
[15] D. NEEDELL AND J. A. TROPP COSAMP: ITERATIVE SIGNAL RECOVERY FROM INCOMPLETE AND INACCURATE SAMPLES. Date: 16 March 2008. Revised 16 April 2008 and 11 July 2008. Initially presented at Information Theory and Applications, 31 January 2008, San Diego. 2000 Mathematics Subject Classification. 41A46, 68Q25, 68W20, 90C27.
[16] Wei Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.
[17] G. Reeves and M. Gastpar, "A note on optimal support recovery in compressed sensing," *Proc. Asilomar Conf. Signals, Sys., and Comp.*, pp. 1576 – 1580, nov. 2009.
[18] D. Needell and R. Vershynin, "Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 310–316, Apr. 2010.
[19] Wei Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230 –2249, May 2009.
[20] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. and Comp. Harm. Analysis*, vol. 26, pp. 301–321, May 2009.
[21] S. Chatterjee, D. Sundman, M. Vehkaper¨a, and M. Skoglund, "Projection-based and look ahead strategies for atom selection," *IEEE Trans. Signal Processing*, vol. PP, no. 99, pp. 1, 2011.
[22] David L. Donoho, Yaakov Tsaig, Iddo Drori, and Jean luc Starck, "Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit," Tech. Rep., 2006.

[23] H. Huang and A. Makur, “Backtracking-based matching pursuit method for sparse signal reconstruction,” *IEEE Signal Processing Letters*, vol. 18, no. 7, pp. 391 –394, july 2011.
[24] D. Sundman, S. Chatterjee, and M. Skoglund, “Look ahead parallel pursuit,” in *2011 IEEE Swedish Communication Technologies Workshop (Swe-CTW)*, oct. 2011, pp. 114 –117.
[25] D. Needell and R. Vershynin, “Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 310–316, Apr. 2010.
[26] G. Reeves and M. Gastpar, “A note on optimal support recovery in compressed sensing,” *Proc. Asilomar Conf. Signals, Sys., and Comp.*, pp. 1576 –1580, nov. 2009.
[27] E.J. Candes and M.B. Wakin, “An introduction to compressive sampling,” *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21 –30, Mar. 2008.