# Chapter 2
# Compressive Sensing

Compressive sensing [47], [23] is a new concept in signal processing and information theory where one measures a small number of non-adaptive linear combinations of the signal. These measurements are usually much smaller than the number of samples that define the signal. From these small number of measurements, the signal is then reconstructed by a non-linear procedure. In what follows, we present some fundamental premises underlying CS: sparsity, incoherent sampling and non-linear recovery.

## 2.1 Sparsity

Let $\mathbf{x}$ be a discrete time signal which can be viewed as an $N \times 1$ column vector in $\mathbb{R}^N$. Given an orthonormal basis matrix $\mathbf{B} \in \mathbb{R}^{N \times N}$ whose columns are the basis elements $\{\mathbf{b}_i\}_{i=1}^N$, $\mathbf{x}$ can be represented in terms of this basis as

$$\mathbf{x} = \sum_{i=1}^{N} \alpha_i \mathbf{b}_i \tag{2.1}$$

or more compactly $\mathbf{x} = \mathbf{B}\alpha$, where $\alpha$ is an $N \times 1$ column vector of coefficients. These coefficients are given by $\alpha_i = \langle \mathbf{x}, \mathbf{b}_i \rangle = \mathbf{b}_i^T \mathbf{x}$ where $.^T$ denotes the transposition operation. If the basis $\mathbf{B}$ provides a $K$-sparse representation of $\mathbf{x}$, then (2.1) can be rewritten as

$$\mathbf{x} = \sum_{i=1}^{K} \alpha_{n_i} \mathbf{b}_{n_i},$$

where $\{n_i\}$ are the indices of the coefficients and the basis elements corresponding to the $K$ nonzero entries. In this case, $\alpha$ is an $N \times 1$ column vector with only $K$ nonzero elements. That is, $\|\alpha\|_0 = K$ where $\|.\|_p$ denotes the $\ell_p$-norm defined as

V.M. Patel and R. Chellappa, *Sparse Representations and Compressive Sensing for Imaging and Vision*, SpringerBriefs in Electrical and Computer Engineering,
DOI 10.1007/978-1-4614-6381-8_2, 

$$\|\mathbf{x}\|_p = \left( \sum_i | x_i |^p \right)^{\frac{1}{p}}$$

and the $\ell_0$-norm is defined as the limit as $p \to 0$ of the $\ell_p$-norm

$$\|\mathbf{x}\|_0 = \lim_{p\to 0} \|\mathbf{x}\|_p^p = \lim_{p\to 0} \sum_i | x_i |^p .$$

In general, the $\ell_0$-norm counts the number of non-zero elements in a vector

$$\|\mathbf{x}\|_0 = \sharp\{i : x_i \neq 0\}. \tag{2.2}$$

Typically, real-world signals are not exactly sparse in any orthogonal basis. Instead, they are *compressible*. A signal is said to be compressible if the magnitude of the coefficients, when sorted in a decreasing order, decays according to a power law [87],[19]. That is, when we rearrange the sequence in decreasing order of magnitude $\alpha_{(1)} \geq \alpha_{(2)} \geq \cdots \geq \alpha_{(N)}$, then the following holds

$$| \alpha |_{(n)} \leq C.n^{-s}, \tag{2.3}$$

where $| \alpha |_{(n)}$ is the $n$th largest entry of $\alpha$, $s \geq 1$ and $C$ is a constant. For a given $L$, the $L$-term linear combination of elements that best approximate $\mathbf{x}$ in an $L_2$-sense is obtained by keeping the $L$ largest terms in the expansion

$$\mathbf{x}_L = \sum_{n=0}^{L-1} \alpha_{(n)} \mathbf{b}_{(n)}.$$

If $\alpha$ obeys (2.3), then the error between $\mathbf{x}_L$ and $\mathbf{x}$ also obeys a power law as well [87], [19]

$$\|\mathbf{x}_L - \mathbf{x}\|_2 \leq CL^{-(s-\frac{1}{2})}.$$

In other words, a small number of vectors from $\mathbf{B}$ can provide accurate approximations to $\mathbf{x}$. This type of approximation is often known as the *non-linear approximation* [87].

Fig. 2.1 shows an example of the non-linear approximation of the Boats image using Daubechies 4 wavelet. The original Boats image is shown in Fig. 2.1(a). Two level Daubechies 4 wavelet coefficients are shown in Fig. 2.1(b). As can be seen from this figure, these coefficients are very sparse. The plot of the sorted absolute values of the coefficients of the image is shown in Fig. 2.1(c). The reconstructed image after keeping only 10% of the coefficients with the largest magnitude is shown in Fig. 2.1(d). This reconstruction provides a very good approximation to the original image. In fact, it is well known that wavelets provide the best representation for piecewise smooth images. Hence, in practice wavelets are often used to compressively represent images.

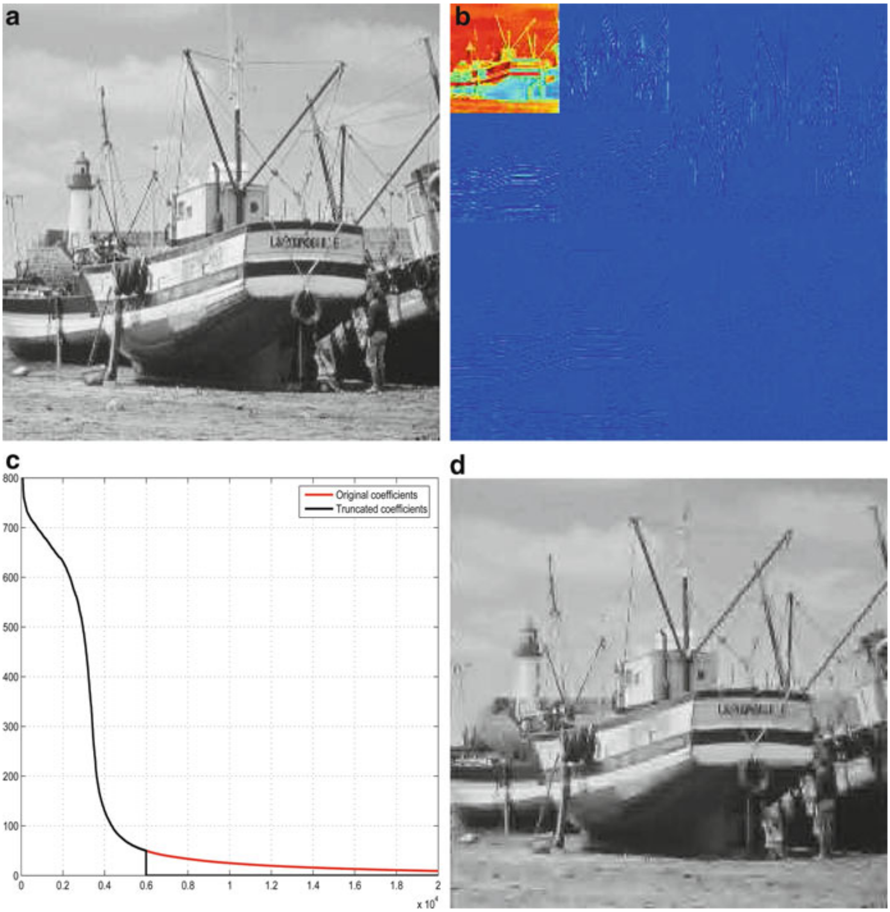


**Fig. 2.1** Compressibility of wavelets. **(a)** Original Boats image. **(b)** Wavelet coefficients. **(c)** The plot of the sorted absolute values of the coefficients. **(d)** Reconstructed image after keeping only 10% of the coefficients with the largest magnitude

## 2.2 Incoherent Sampling

In CS, the $K$ largest $\alpha_i$ in (2.1) are not measured directly. Instead, $M \ll N$ projections of the vector $\mathbf{x}$ with a collection of vectors $\{\phi_j\}_{j=1}^M$ are measured as in $y_j = \langle x, \phi_j \rangle$. Arranging the measurement vector $\phi_j^T$ as rows in an $M \times N$ matrix $\boldsymbol{\Phi}$ and using (2.1), the measurement process can be written as

$$\mathbf{y} = \boldsymbol{\Phi}\mathbf{x} = \boldsymbol{\Phi}\mathbf{B}\alpha = \mathbf{A}\alpha, \tag{2.4}$$

where $\mathbf{y}$ is an $M \times 1$ column vector of the compressive measurements and $\mathbf{A} = \boldsymbol{\Phi}\mathbf{B}$ is the measurement matrix or the sensing matrix. Given an $M \times N$ sensing matrix $\mathbf{A}$ and the observation vector $\mathbf{y}$, the general problem is to recover the sparse or compressible vector $\alpha$. To this end, the first question is to determine whether $\mathbf{A}$ is good for compressive sensing. Candés and Tao introduced a necessary condition on $\mathbf{A}$ that guarantees a stable solution for both $K$ sparse and compressible signals [26], [24].

**Definition 2.1.** A matrix $\mathbf{A}$ is said to satisfy the Restricted Isometry Property (RIP) of order $K$ with constants $\delta_K \in (0,1)$ if

$$(1-\delta_K)\|\mathbf{v}\|_2^2 \leq \|\mathbf{A}\mathbf{v}\|_2^2 \leq (1+\delta_K)\|\mathbf{v}\|_2^2$$

for any $\mathbf{v}$ such that $\|\mathbf{v}\|_0 \leq K$.

An equivalent description of RIP is to say that all subsets of $K$ columns taken from $\mathbf{A}$ are nearly orthogonal. This in turn implies that $K$ sparse vectors cannot be in the null space of $\mathbf{A}$. When RIP holds, $\mathbf{A}$ approximately preserves the Euclidean length of $K$ sparse vectors. That is,

$$(1-\delta_{2K})\|\mathbf{v_1}-\mathbf{v_2}\|_2^2 \leq \|\mathbf{A}\mathbf{v_1}-\mathbf{A}\mathbf{v_2}\|_2^2 \leq (1+\delta_{2K})\|\mathbf{v_1}-\mathbf{v_2}\|_2^2$$

holds for all $K$ sparse vectors $\mathbf{v_1}$ and $\mathbf{v_2}$. A related condition known as incoherence, requires that the rows of $\boldsymbol{\Phi}$ can not sparsely represent the columns of $\mathbf{B}$ and vice versa.

**Definition 2.2.** The coherence between $\boldsymbol{\Phi}$ and the representation basis $\mathbf{B}$ is

$$\mu(\boldsymbol{\Phi},\mathbf{B}) = \sqrt{N} \max_{1\leq i,j\leq N} |\langle \phi_i, \mathbf{b}_j\rangle|, \tag{2.5}$$

where $\phi_i \in \boldsymbol{\Phi}$ and $\mathbf{b}_j \in \mathbf{B}$.

The number $\mu$ measures how much two vectors in $\mathbf{A} = \boldsymbol{\Phi}\mathbf{B}$ can look alike. The value of $\mu$ is between 1 and $\sqrt{N}$. We say that a matrix $\mathbf{A}$ is *incoherent* when $\mu$ is very small. The incoherence holds for many pairs of bases. For example, it holds for the delta spikes and the Fourier bases. Surprisingly, with high probability, incoherence holds between any arbitrary basis and a random matrix such as Gaussian or Bernoulli [6], [142].

## 2.3 Recovery

Since, $M \ll N$, we have an under-determined system of linear equations, which in general has infinitely many solutions. So our problem is ill-posed. If one desires to narrow the choice to a well-defined solution, additional constraints are needed.

One approach is to find the minimum-norm solution by solving the following optimization problem

$$\hat{\alpha} = \arg\min_{\alpha'} \|\alpha'\|_2 \text{ subject to } \mathbf{y} = \mathbf{A}\alpha'.$$

The solution to the above problem is explicitly given by

$$\hat{\alpha} = \mathbf{A}^\dagger \mathbf{y} = \mathbf{A}^*(\mathbf{A}\mathbf{A}^*)^{-1}\mathbf{y},$$

where $\mathbf{A}^*$ is the adjoint of $\mathbf{A}$ and $\mathbf{A}^\dagger = \mathbf{A}^*(\mathbf{A}\mathbf{A}^*)^{-1}$ is the pseudo-inverse of $\mathbf{A}$. This solution, however, yields a non-sparse vector. The approach taken in CS is to instead find the sparsest solution.

The problem of finding the sparsest solution can be reformulated as finding a vector $\alpha \in \mathbb{R}^N$ with a minimum possible number of nonzero entries. That is

$$\hat{\alpha} = \arg\min_{\alpha'} \|\alpha'\|_0 \text{ subject to } \mathbf{y} = \mathbf{A}\alpha'. \tag{2.6}$$

This problem can recover a $K$ sparse signal exactly. However, this is an NP-hard problem. It requires an exhaustive search of all $\binom{N}{K}$ possible locations of the nonzero entries in $\alpha$.

The main approach taken in CS is to minimize the $\ell_1$-norm instead

$$\hat{\alpha} = \arg\min_{\alpha'} \|\alpha'\|_1 \text{ subject to } \mathbf{y} = \mathbf{A}\alpha'. \tag{2.7}$$

Surprisingly, the $\ell_1$ minimization yields the same result as the $\ell_0$ minimization in many cases of practical interest. This program also approximates compressible signals. This convex optimization program is often known as Basis Pursuit (BP) [38]. The use of $\ell_1$ minimization for signal restoration was initially observed by engineers working in seismic exploration as early as 1970s [52]. In the last few years, a series of papers [47], [142], [21], [25], [19], [22], explained why $\ell_1$ minimization can recover sparse signals in various practical setups.

### 2.3.1 Robust CS

In this section we examine the case when there are noisy observations of the following form

$$\mathbf{y} = \mathbf{A}\alpha + \eta \tag{2.8}$$

where $\eta \in \mathbb{R}^M$ is the measurement noise or an error term. Note that $\eta$ can be stochastic or deterministic. Furthermore, let's assume that $\|\eta\|_2 \leq \varepsilon$. Then, $\mathbf{x}$ can be recovered from $\mathbf{y}$ via $\alpha$ by solving the following problem

$$\hat{\alpha} = \arg\min_{\alpha'} \|\alpha'\|_1 \text{ subject to } \|\mathbf{y} - \mathbf{A}\alpha'\| \leq \varepsilon. \tag{2.9}$$

The problem (2.9) is often known as Basis Pursuit DeNoising (BPDN) [38]. In [22], Candés *at. el.* showed that the solution to (2.9) recovers an unknown sparse signal with an error at most proportional to the noise level.

**Theorem 2.1.** *[22] Let* $\mathbf{A}$ *satisfy RIP of order* $4K$ *with* $\delta_{3K} + 3\delta_{4K} < 2$. *Then, for any* $K$ *sparse signal* $\alpha$ *and any perturbation* $\eta$ *with* $\|\eta\|_2 \leq \varepsilon$, *the solution* $\hat{\alpha}$ *to (2.9) obeys*

$$\|\hat{\alpha} - \alpha\|_2 \leq \varepsilon C_K$$

*with a well behaved constant* $C_K$.

Note that for $K$ obeying the condition of the theorem, the reconstruction from noiseless data is exact. A similar result also holds for stable recovery from imperfect measurements for approximately sparse signals (i.e compressible signals).

**Theorem 2.2.** *[22] Let* $\mathbf{A}$ *satisfy RIP of order* $4K$. *Suppose that* $\alpha$ *is an arbitrary vector in* $\mathbb{R}^N$ *and let* $\alpha_K$ *be the truncated vector corresponding to the* $K$ *largest values of* $\theta$ *in magnitude. Under the hypothesis of Theorem 2.1, the solution* $\hat{\alpha}$ *to (2.9) obeys*

$$\|\hat{\alpha} - \alpha\|_2 \leq \varepsilon C_{1,K} + C_{2,K} \frac{\|\alpha - \alpha_K\|_1}{\sqrt{K}}$$

*with well behaved constants* $C_{1,K}$ *and* $C_{2,K}$.

If $\alpha$ obeys (2.3), then

$$\frac{\|\hat{\alpha} - \alpha_K\|_1}{\sqrt{K}} \leq C' K^{-(s-\frac{1}{2})}.$$

So in this case

$$\|\hat{\alpha} - \alpha_K\|_2 \leq C'' K^{-(s-\frac{1}{2})},$$

and for signal obeying (2.3), there are fundamentally no better estimates available. This, in turn, means that with only $M$ measurements, one can achieve an approximation error which is almost as good as that one obtains by knowing everything about the signal $\alpha$ and selecting its $K$-largest elements [22].

#### 2.3.1.1 The Dantzig selector

In (2.8), if the noise is assumed to be Gaussian with mean zero and variance $\sigma^2$, $\eta \sim \mathcal{N}(0, \sigma^2)$, then the stable recovery of the signal is also possible by solving a modified optimization problem

$$\hat{\alpha} = \arg\min_{\alpha'} \|\alpha'\|_1 \text{ s. t. } \|\mathbf{A}^T(\mathbf{y} - \mathbf{A}\alpha')\|_\infty \leq \varepsilon' \tag{2.10}$$

where $\varepsilon' = \lambda_N \sigma$ for some $\lambda_N > 0$ and $\|.\|_\infty$ denotes the $\ell_\infty$ norm. For an $N$ dimensional vector $\mathbf{x}$, it is defined as $\|\mathbf{x}\|_\infty = \max(|x_1|, \cdots, |x_N|)$. The above program is known as the Dantzig Selector [28].

**Theorem 2.3.** *[28] Suppose $\alpha \in \mathbb{R}^N$ is any K-sparse vector obeying $\delta_{2K} + \vartheta_{K,2K} < 1$. Choose $\lambda_N = \sqrt{2\log(N)}$ in (2.10). Then, with large probability, the solution to (2.10), $\hat{\alpha}$ obeys*

$$\|\hat{\alpha} - \alpha\|_2^2 \leq C_1^2.(2\log(N)).K.\sigma^2, \tag{2.11}$$

*with*

$$C_1 = \frac{4}{1 - \delta_K - \vartheta_{K,2K}},$$

where $\vartheta_{K,2K}$ is the $K,2K$-*restricted orthogonal constant* defined as follows

**Definition 2.3.** The $K,K'$-restricted orthogonality constant $\vartheta_{K,K'}$ for $K + K' \leq N$ is defined to be the smallest quantity such that

$$|\langle \mathbf{A}_T \mathbf{v}, \mathbf{A}_{T'} \mathbf{v}' \rangle| \leq \vartheta_{K,K'} \|\mathbf{v}\|_2 \|\mathbf{v}'\|_2 \tag{2.12}$$

holds for all disjoint sets $T, T' \subseteq \{1,...,N\}$ of cardinality $|T| \leq K$ and $|T'| \leq K'$.

A similar result also exists for compressible signals (see [28] for more details).

### *2.3.2 CS Recovery Algorithms*

The $\ell_1$ minimization problem (2.10) is a linear program [28] while (2.9) is a second-order cone program (SOCP) [38]. SOCPs can be solved using interior point methods [74]. Log-barrier and primal dual methods can also be used [15], [3] to solve SOCPs. Note, the optimization problems (2.7), (2.9), and (2.10) minimize convex functionals, hence a global minimum is guaranteed.

In what follows, we describe other CS related reconstruction algorithms.

#### 2.3.2.1 Iterative Thresholding Algorithms

A Lagrangian formulation of the problem (2.9) is the following

$$\hat{\alpha} = \arg\min_{\alpha'} \|\mathbf{y} - \mathbf{A}\alpha'\|_2^2 + \lambda \|\alpha'\|_1. \tag{2.13}$$

There exists a mapping between $\lambda$ from (2.13) and $\varepsilon$ from (2.9) so that both problems (2.9) and (2.13) are equivalent. Several authors have proposed to solve (2.13) iteratively [12], [45], [11], [9]. This algorithm iteratively performs a soft-thresholding to decrease the $\ell_1$ norm of the coefficients $\alpha$ and a gradient descent to decrease the value of $\|\mathbf{y} - \mathbf{A}\alpha\|_2^2$. The following iteration is usually used

$$\mathbf{y}^{n+1} = T_\lambda(\mathbf{y}^n + \mathbf{A}^*(\alpha - \mathbf{A}\mathbf{y}^n)), \tag{2.14}$$

where $T_\lambda$ is the element wise soft-thresholding operator

$$T_\lambda(a) = \begin{cases} a + \frac{\lambda}{2}, & \text{if } a \le \frac{-\lambda}{2} \\ 0, & \text{if } |a| < \frac{\lambda}{2} \\ a - \frac{\lambda}{2}, & \text{if } a \ge \frac{\lambda}{2}. \end{cases}$$

The iterates $\mathbf{y}^{n+1}$ converge to the solution of (2.9), $\hat{\alpha}$ if $\|\mathbf{A}\|_2 < 1$ [45]. Similar results can also be obtained using the hard-thresholding instead of the soft-thresholding in (2.14) [11].

Other methods for solving (2.13) have also been proposed. See for instance GPSR [61], SPGL1 [8], Bregman iterations [159], split Bregman iterations [65], SpaRSA [157], and references therein.

#### 2.3.2.2 Greedy Pursuits

In certain conditions, greedy algorithms such as matching pursuit [88], orthogonal matching pursuit [109], [138], gradient pursuits [13], regularized orthogonal matching pursuit [94] and Stagewise Orthogonal Matching Pursuit [49] can also be used to recover sparse (or in some cases compressible) $\alpha$ from (2.8). In particular, a greedy algorithm known as, CoSaMP, is well supported by theoretical analysis and provides the same guarantees as some of the optimization based approaches [93].

Let $T$ be a subset of $\{1, 2, \cdots, N\}$ and define the restriction of the signal $\mathbf{x}$ to the set $T$ as

$$\mathbf{x}_{|T} = \begin{cases} x_i, & i \in T \\ 0, & \text{otherwise} \end{cases}$$

Let $\mathbf{A}_T$ be the column submatrix of $\mathbf{A}$ whose columns are listed in the set $T$ and define the pseudoinverse of a tall, full-rank matrix $\mathbf{C}$ by the formula $\mathbf{C}^\ddagger = (\mathbf{C}^*\mathbf{C})^{-1}\mathbf{C}^*$. Let $supp(\mathbf{x}) = \{x_j : j \neq 0\}$. Using this notation, the pseudo-code for CoSaMP is given in Algorithm 1 which can be used to solve the under-determined system of linear equations (2.4).

#### 2.3.2.3 Other Algorithms

Recently, there has been a great interest in using $\ell_p$ minimization with $p < 1$ for compressive sensing [37]. It has been observed that the minimization of such a nonconvex problem leads to recovery of signals that are much less sparse than required by the traditional methods [37].

Other related algorithms such as FOCUSS and reweighted $\ell_1$ have also been proposed in [68] and [29], respectively.

**Algorithm 1:** Compressive Sampling Matching Pursuit (CoSaMP)

**Input:** $\mathbf{A}$, $\mathbf{y}$, sparsity level $K$.
**Initialize:** $\alpha_0 = \mathbf{0}$ and the current residual $\mathbf{r} = \mathbf{y}$.
**While not converged do**
1. Compute the current error:
$\mathbf{v} = \mathbf{A}^*\mathbf{r}$.
2. Compute the best 2K support set of the error:
$\Omega = \mathbf{v}_{2K}$.
3. Merge the the strongest support sets:
$T = \Omega \cup supp(\alpha_{J-1})$.
4. Perform a least-squares signal estimation:
$\mathbf{b}_{|_T} = \mathbf{A}^{\ddagger}_{|_T}\mathbf{y}$, $\mathbf{b}_{T^c} = \mathbf{0}$.
5. Prune $\alpha_{J-1}$ and compute $\mathbf{r}$ for the next round:
$\alpha_J = \mathbf{b}_k$, $\mathbf{r} = \mathbf{y} - \mathbf{A}\alpha_J$.

## 2.4 Sensing Matrices

Most of the sensing matrices in CS are produced by taking i.i.d. random variables with some given probability distribution and then normalizing their columns. These matrices are guaranteed to perform well with high probability. In what follows, we present some commonly used sensing matrices in CS [22], [142], [26].

- *Random matrices with i.i.d. entries:* Consider a matrix $\mathbf{A}$ with entries drawn independently from the Gaussian probability distribution with mean zero and variance $1/M$. Then the conditions for Theorem 2.1 hold with overwhelming probability when
$$K \le CM/\log(N/M).$$
- *Fourier ensemble:* Let $\mathbf{A}$ be an $M \times N$ matrix obtained by selecting $M$ rows, at random, from the $N \times N$ discrete Fourier transform matrix and renormalizing the columns. Then with overwhelming probability, the conditions for Theorem 2.1 holds provided that
$$K \le C\frac{M}{(\log(N))^6}.$$
- *General orthogonal ensembles:* Suppose $\mathbf{A}$ is obtained by selecting $M$ rows from an $N \times N$ orthonormal matrix $\Theta$ and renormalizing the columns. If the rows are selected at random, then the conditions for Theorem 2.1 hold with overwhelming probability when
$$K \le C\frac{1}{\mu^2}\frac{M}{(\log(N))^6},$$
where $\mu$ is defined in (2.5).

## 2.5 Phase Transition Diagrams

The performance of a CS system can be evaluated by generating phase transition diagrams [86], [48], [10], [51]. Given a particular CS system, governed by the sensing matrix $\mathbf{A} = \boldsymbol{\Phi}\mathbf{B}$, let $\delta = \frac{M}{N}$ be a normalized measure of undersampling factor and $\rho = \frac{K}{M}$ be a normalized measure of sparsity. A plot of the pairing of the variables $\delta$ and $\rho$ describes a 2-D phase space $(\delta, \rho) \in [0,1]^2$. It has been shown that for many practical CS matrices, there exist sharp boundaries in this phase space that clearly divide the solvable from unsolvable problems in the noiseless case. In other words, a phase transition diagram provides a way of checking $\ell_0/\ell_1$ equivalence, indicating how sparsity and indeterminacy affect the success of $\ell_1$ minimization [86], [48], [51]. Fig. 2.2 shows an example of a phase transition diagram which is obtained when a random Gaussian matrix is used as $\mathbf{A}$. Below the boundary, $\ell_0/\ell_1$ equivalence holds and above the boundary, the system lacks sparsity and/or too few measurements are obtained to solve the problem correctly.

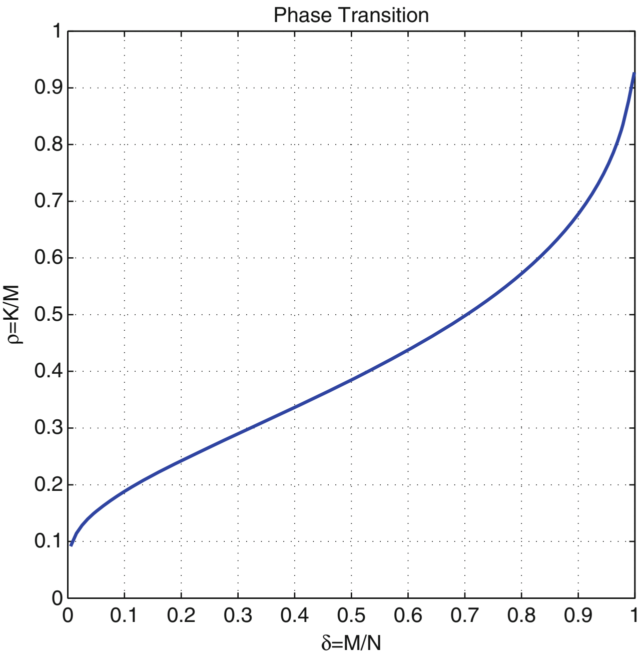


**Fig. 2.2** Phase transition diagram corresponding to a CS system where **A** is the random Gaussian matrix. The boundary separates regions in the problem space where (2.7) can and cannot be solved. Below the curve solutions can be obtained and above the curve solutions can not be obtained

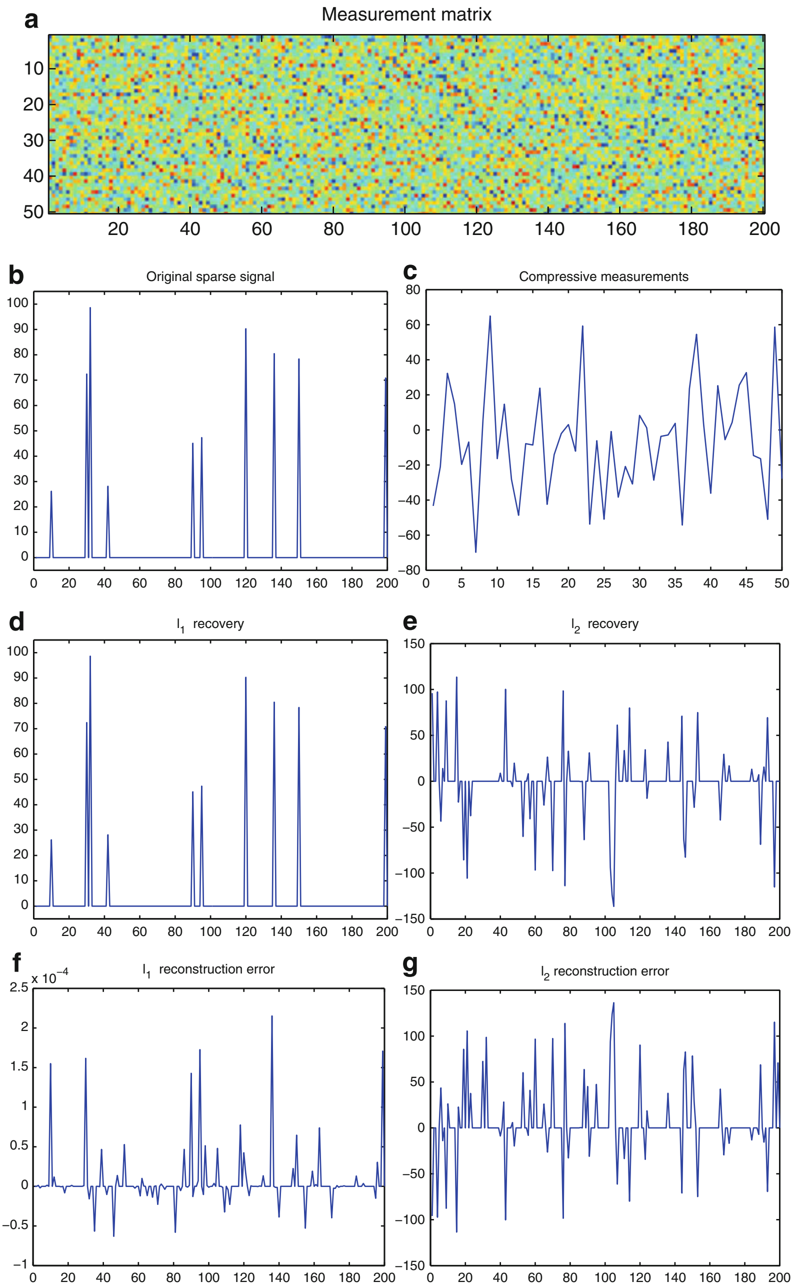


**Fig. 2.3** 1D sparse signal recovery example from random Gaussian measurements. **(a)** Compressive measurement matrix. **(b)** Original sparse signal. **(c)** Compressive measurements. **(d)** $\ell_1$ recovery. **(e)** $\ell_2$ recovery. **(f)** $\ell_1$ reconstruction error. **(g)** $\ell_2$ reconstruction error

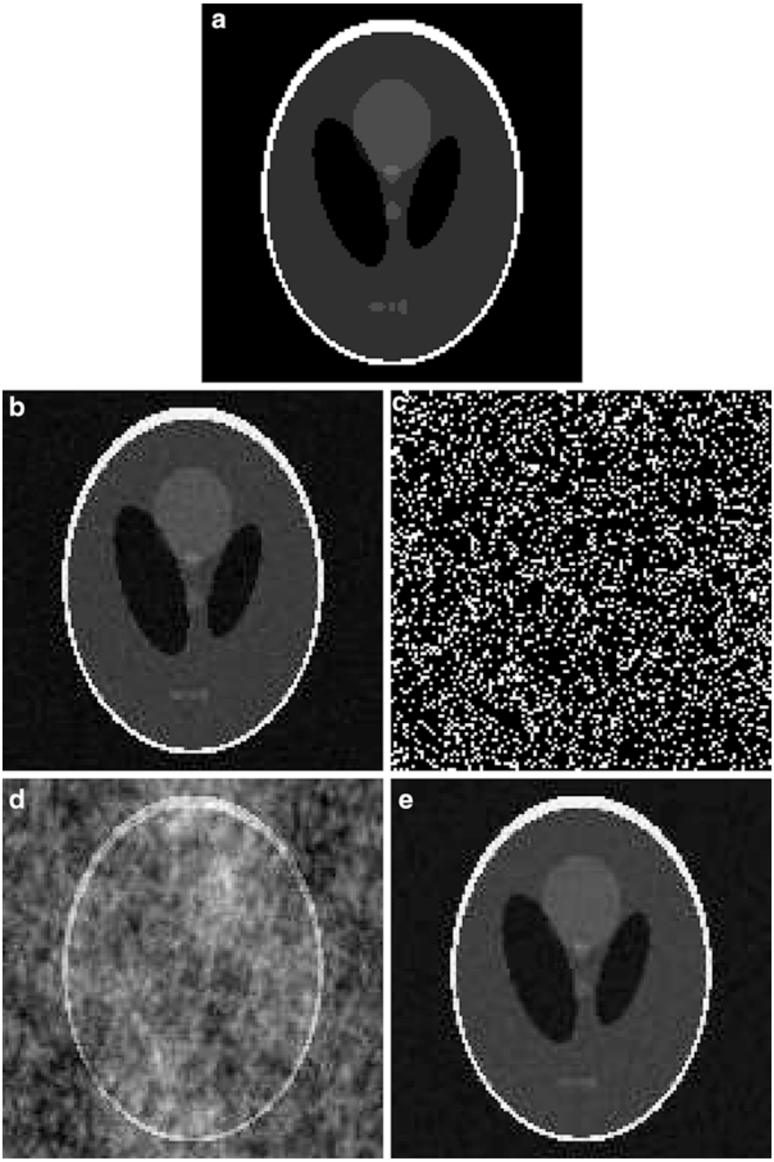


**Fig. 2.4** 2D sparse image recovery example from random Fourier measurements. **(a)** Original image. **(b)** Original image contaminated by additive white Gaussian noise with signal-to-noise ratio of 20 dB. **(c)** Sampling mask in the Fourier domain. **(d)** $\ell_2$ recovery. **(e)** $\ell_1$ recovery

## 2.6 Numerical Examples

We end this section by considering the following two examples. In the first example, a 1D signal $\mathbf{x}$ of length 200 with only 10 nonzero elements is undersampled using a random Gaussian matrix $\boldsymbol{\Phi}$ of size $50 \times 200$ as shown in Fig. 2.3(a). Here, the sparsifying transform $\mathbf{B}$ is simply the identity matrix and the observation vector $\mathbf{y}$ is of length 50. Having observed $\mathbf{y}$ and knowing $\mathbf{A} = \boldsymbol{\Phi}$ the signal $\mathbf{x}$ is then recovered by solving the following optimization problem

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}' \in \mathbb{R}^N} \|\mathbf{x}'\|_1 \text{ subject to } \mathbf{y} = \mathbf{A}\mathbf{x}'. \tag{2.15}$$

As can be seen from Fig. 2.3(d), indeed the solution to the above optimization problem recovers the sparse signal exactly from highly undersampled observations. Whereas, the minimum norm solution (i.e. by minimizing the $\ell_2$ norm), as shown in Fig. 2.3(e), fails to recover the sparse signal. The errors corresponding the $\ell_1$ and $\ell_2$ recovery are shown in Fig. 2.3(f) and Fig. 2.3(g), respectively.

In the second example, we reconstructed an undersampled Shepp-Logan phantom image of size $128 \times 128$ in the presence of additive white Gaussian noise with signal-to-noise ratio of 30 dB. For this example, we used only 15% of the random Fourier measurements and Haar wavelets as a sparsifying transform. So the observations can be written as $\mathbf{y} = \mathbf{MFB}\alpha + \eta$, where $\mathbf{y}, \mathbf{M}, \mathbf{F}, \mathbf{B}, \alpha$ and $\eta$ are the noisy compressive measurements, the restriction operator, Fourier transform operator, the Haar transform operator, the sparse coefficient vector and the noise vector with $\|\eta\|_2 \leq \varepsilon$, respectively. The image was reconstructed via $\alpha$ estimated by solving the following optimization problem

$$\hat{\alpha} = \arg\min_{\alpha'} \|\alpha'\|_1 \text{ subject to } \|\mathbf{y} - \mathbf{MFB}\alpha'\| \leq \varepsilon.$$

The reconstruction from $\ell_2$ and $\ell_1$ minimization is shown in Fig. 2.4(d) and Fig. 2.3(e), respectively. This example shows that, it is possible to obtain a stable reconstruction from the compressive measurements in the presence of noise. For both of the above examples we used SPGL1 [8] algorithm for solving the $\ell_1$ minimization problems.

In [23], [47], a theoretical bound on the number of samples that need to be measured for a good reconstruction has been derived. However, it has been observed by many researchers [79], [22], [142], [19], [26] that in practice samples in the order of two to five times the number of sparse coefficients suffice for a good reconstruction. Our experiments also support this claim.

http://www.springer.com/978-1-4614-6380-1

Sparse Representations and Compressive Sensing for Imaging and Vision
Patel, V.M.; Chellappa, R.
2013, X, 102 p. 41 illus., Softcover
ISBN: 978-1-4614-6380-1