

---

## Chapter 10

# Signal Approximation and Compressed Sensing

---

### 10.1 Introduction

In this chapter, we discuss applications of  $\ell_1$ -based relaxation to problems of signal recovery and approximation. Our focus is the role played by sparsity in signal representation and approximation, and the use of  $\ell_1$ -methods for exploiting this sparsity for solving problems like signal denoising, compression, and approximation. We begin by illustrating that many classes of “natural” signals are sparse when represented in suitable bases, such as those afforded by wavelets and other multiscale transforms. We illustrate how such sparsity can be exploited for compression and denoising in orthogonal bases. Next we discuss the problem of signal approximation in overcomplete bases, and the role of  $\ell_1$ -relaxation in finding near-optimal approximations. Finally, we discuss the method of compressed sensing for recovering sparse signals. It is a combination of two ideas: taking measurements of signals via random projections, and solving a lasso-type problem for reconstruction.

### 10.2 Signals and Sparse Representations

Let us begin by providing some background on the role of sparse representations in signal processing. To be clear, our use of the term “signal” is general, including (among other examples) data such as sea water levels, seismic recordings, medical time series, audio recordings, photographic images, video data, and financial data. In all cases, we represent the signal by a vector  $\theta^* \in \mathbb{R}^p$ . (For two-dimensional signals such as images, the reader should think about a vectorized form of the image.)

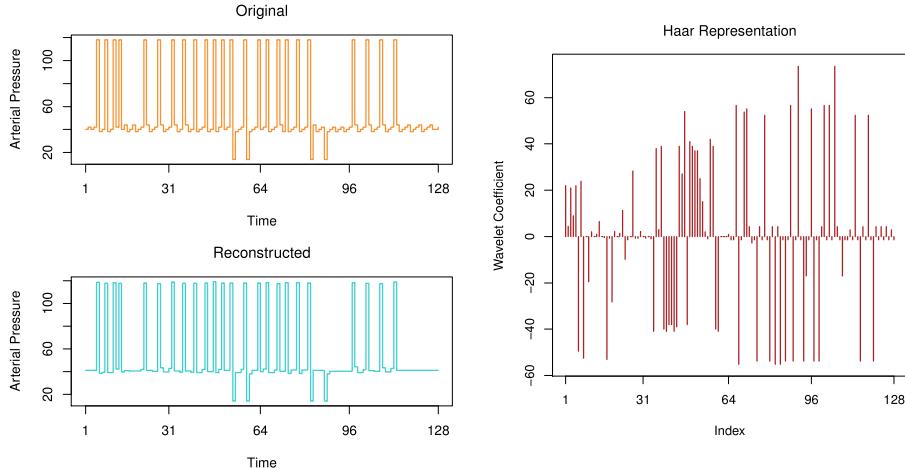
#### 10.2.1 Orthogonal Bases

In signal processing, it is frequently useful to represent signals in different types of bases. Examples include Fourier representations, useful for extracting periodic structure in time series, and multiscale representations such as wavelets. Such representations are described by a collection of vectors

$\{\psi_j\}_{j=1}^p$  that form an orthonormal basis of  $\mathbb{R}^p$ . If we define the  $p \times p$  matrix  $\Psi := [\psi_1 \ \psi_2 \ \dots \ \psi_p]$ , then the orthonormality condition guarantees that  $\Psi^T \Psi = I_{p \times p}$ . Given an orthonormal basis, any signal  $\theta^* \in \mathbb{R}^p$  can be expanded in the form

$$\theta^* := \sum_{j=1}^p \beta_j^* \psi_j, \quad (10.1)$$

where the  $j^{th}$  basis coefficient  $\beta_j^* := \langle \theta^*, \psi_j \rangle = \sum_{i=1}^p \theta_i^* \psi_{ij}$  is obtained by projecting the signal onto the  $j^{th}$  basis vector  $\psi_j$ . Equivalently, we can write the transformation from signal  $\theta^* \in \mathbb{R}^p$  to basis coefficient vector  $\beta^* \in \mathbb{R}^p$  as the matrix-vector product  $\beta^* = \Psi^T \theta^*$ .



**Figure 10.1** Illustration of sparsity in time series data. Left, top panel: Signal  $\theta^* \in \mathbb{R}^p$  of arterial pressure versus time over  $p = 128$  points. Left, bottom panel: Reconstruction  $\hat{\theta}^{128}$  based on retaining the largest (in absolute amplitude)  $k = 64$  coefficients from the Haar basis. Right: Haar basis coefficients  $\beta^* = \Psi^T \theta^*$  of the signal.

To give a simple example, consider the following matrix

$$\Psi := \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{-1}{\sqrt{2}} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{2} & 0 & \frac{-1}{\sqrt{2}} \end{bmatrix}. \quad (10.2)$$

It is an orthonormal matrix, satisfying  $\Psi^T \Psi = \mathbf{I}_{4 \times 4}$ , and corresponds to a two-level Haar transform for signal length  $p = 4$ . For any given signal  $\theta^* \in \mathbb{R}^4$ , the Haar basis coefficients  $\beta^* = \Psi^T \theta^*$  have a natural interpretation. The first

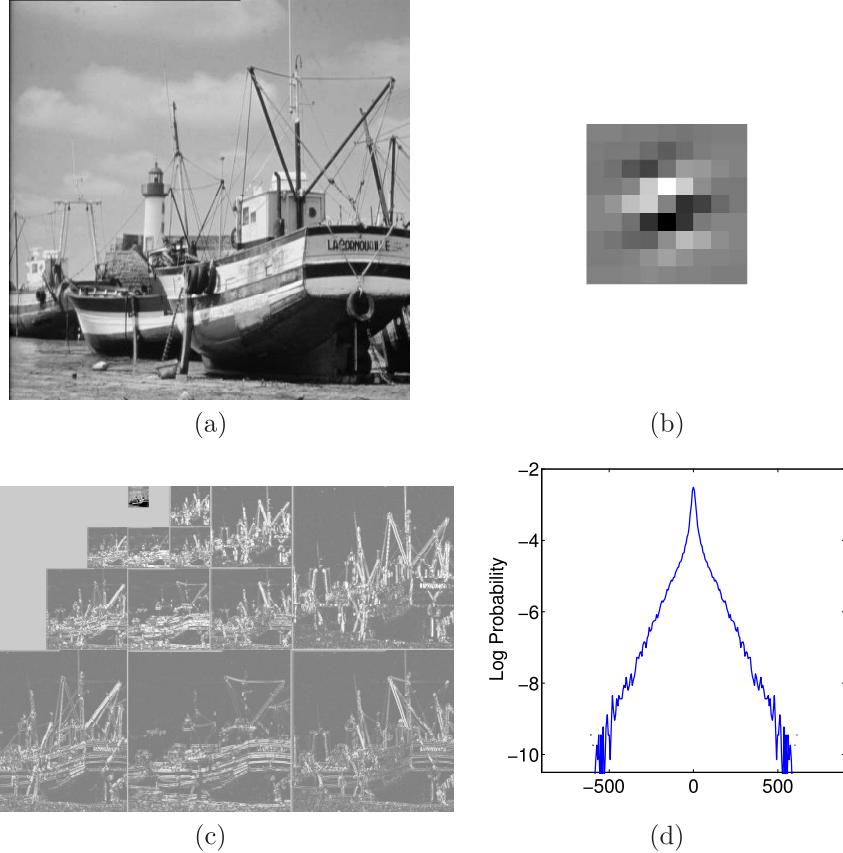
coefficient  $\beta_1^* = \langle \psi_1, \theta^* \rangle = \frac{1}{2} \sum_{j=1}^4 \theta_j^*$  is a rescaled version of the averaged signal. The second column  $\psi_2$  is a differencing operator on the full signal, whereas the third and fourth columns are local differencing operators on each half of the signal. This Haar transform is the simplest example of a wavelet transform.

An important fact is that many signal classes, while not sparse in the canonical basis, become sparse when represented in a different orthogonal basis. Figure 10.1 provides an illustration of this phenomenon for some medical time series data. The top-left panel shows  $p = 128$  samples of arterial pressure from a patient, showing that the signal  $\theta^*$  itself is not at all sparse. The right panel shows the Haar coefficient representation  $\beta^* = \Psi^T \theta^*$  of the signal; note how in contrast it is relatively sparse. Finally, the bottom-left panel shows a reconstruction  $\hat{\theta}$  of the original signal, based on discarding half of the Haar coefficients. Although not a perfect reconstruction, it captures the dominant features of the time series.

Figure 10.2 provides a second illustration of this sparsity phenomenon, this time for the class of photographic images and two-dimensional wavelet transforms. Panel (a) shows a  $512 \times 512$  portion of the “Boats” image; in our framework, we view this two-dimensional image as a vector in  $p = 512^2 = 262,144$  dimensions. Shown in panel (b) is the form of a particular two-dimensional wavelet; as can be discerned from the shape, it is designed to extract diagonally oriented structure at a particular scale. Taking inner products with this wavelet over all spatial positions of the image (a procedure known as convolution) yields a collection of wavelet coefficients at all spatial positions of the image. These coefficients are then sub-sampled, depending on the scale of the wavelet. Then we reconstruct the image from these coefficients. Doing so at multiple scales (three in this illustration) and orientations (four in this illustration) yields the multiscale pyramid shown in panel (c). Once again, although the original image is not a sparse signal, its representation in this multiscale basis is very sparse, with many coefficients either zero or very close to zero. As a demonstration of this sparsity, panel (d) shows a histogram of one of the wavelet coefficients, obtained by pooling its values over all spatial positions of the image. This histogram is plotted on the log scale, and the sharp peak around zero reveals the sparsity of the coefficient distribution.

### 10.2.2 Approximation in Orthogonal Bases

The goal of signal compression is to represent the signal  $\theta^* \in \mathbb{R}^p$ , typically in an approximate manner, using some number  $k \ll p$  of coefficients much smaller than the ambient dimension. In the setting of orthogonal bases, one method for doing so is based on using only a sparse subset of the orthogonal vectors  $\{\psi_j\}_{j=1}^p$ . In particular, for an integer  $k \in \{1, 2, \dots, p\}$  that characterizes the

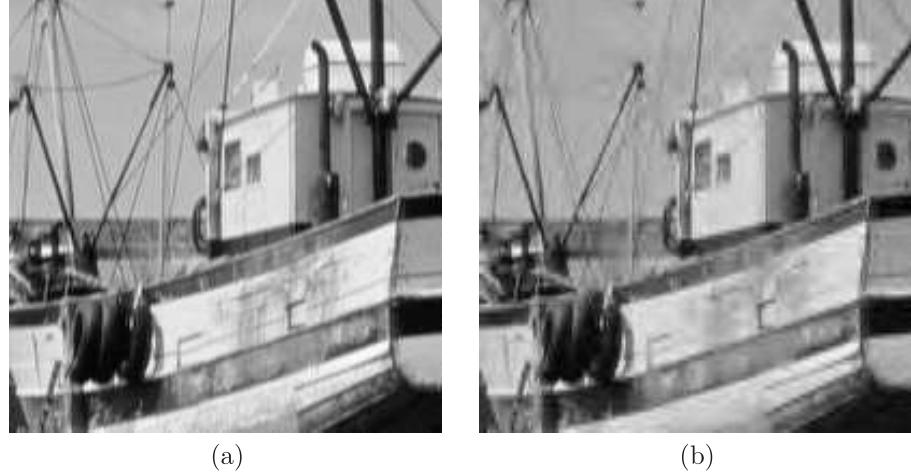


**Figure 10.2** Sparsity in wavelet-based representations of natural images. (a) “Boats” image. (b) Basis vector of a multiscale pyramid transform, drawn here as a 2-dimensional image. (c) Three levels of a multiscale representation of “Boats” image with four different orientations at each scale. (d) Log histogram of the amplitudes of a wavelet coefficient from a fixed scale and orientation, pooled over all pixels within the image. Note that the majority of coefficients are close to zero, with relatively few large in absolute value.

approximation accuracy, let us consider reconstructions of the form

$$\Psi\beta = \sum_{j=1}^p \beta_j \psi_j, \quad \text{such that } \|\beta\|_0 := \sum_{j=1}^p \mathbb{I}[\beta_j \neq 0] \leq k. \quad (10.3)$$

Here we have introduced the  $\ell_0$ -“norm,” which simply counts the number of nonzero elements in the vector  $\beta \in \mathbb{R}^p$ . We then consider the problem of



**Figure 10.3** Illustration of image compression based on wavelet thresholding. (a) Zoomed portion of the original “Boats” image from Figure 10.2(a). (b) Reconstruction based on retaining 5% of the wavelet coefficients largest in absolute magnitude. Note that the distortion is quite small, and concentrated mainly on the fine-scale features of the image.

optimal  $k$ -sparse approximation—namely, to compute

$$\hat{\beta}^k \in \arg \min_{\beta \in \mathbb{R}^p} \|\theta^* - \Psi^T \beta\|_2^2 \quad \text{such that } \|\beta\|_0 \leq k. \quad (10.4)$$

Given the optimal solution  $\hat{\beta}^k$  of this problem, the reconstruction

$$\theta^k := \sum_{j=1}^p \hat{\beta}_j^k \psi_j \quad (10.5)$$

defines the best least-squares approximation to  $\theta^*$  based on  $k$  terms. Figure 10.3 illustrates the idea.

Note that the problem (10.4) is nonconvex and combinatorial, due to the  $\ell_0$ -norm constraint. Despite this fact, it is actually very easy to solve in this particular case, essentially due to the structure afforded by orthonormal transforms. In particular, suppose that we order the vector  $\beta^* \in \mathbb{R}^p$  of basis coefficients in terms of their absolute values, thereby defining the order statistics

$$|\beta_{(1)}^*| \geq |\beta_{(2)}^*| \geq \dots \geq |\beta_{(p)}^*|. \quad (10.6)$$

Then for any given integer  $k \in \{1, 2, \dots, p\}$ , it can be shown that the optimal  $k$ -term approximation is given by

$$\hat{\theta}^k := \sum_{j=1}^k \beta_{(j)}^* \psi_{\sigma(j)}, \quad (10.7)$$

where  $\sigma(j)$  denotes the basis vector associated with the  $j^{\text{th}}$  order statistic. In words, we retain only the basis vectors associated with the largest  $k$  coefficients in absolute value.

In summary, then, we have the following simple algorithm for computing optimal  $k$ -term approximations in an orthogonal basis:

1. Compute the basis coefficients  $\beta_j^* = \langle \theta^*, \psi_j \rangle$  for  $j = 1, 2, \dots, p$ . In matrix-vector notation, compute the vector  $\beta^* = \Psi^T \theta^*$ .
2. Sort the coefficients in terms of absolute values as in (10.6), and extract the top  $k$  coefficients.
3. Compute the best  $k$ -term approximation  $\hat{\theta}^k$  as in (10.7).

For any orthogonal basis, the computational complexity of this procedure is at most  $\mathcal{O}(p^2)$ , with the  $\mathcal{O}(p \log p)$  complexity of sorting in step 2 dominated by the complexity of computing the basis coefficients in step 1. An attractive feature of many orthogonal representations, including Fourier bases and discrete wavelets, is that the basis coefficients can be computed in time  $\mathcal{O}(p \log p)$ .

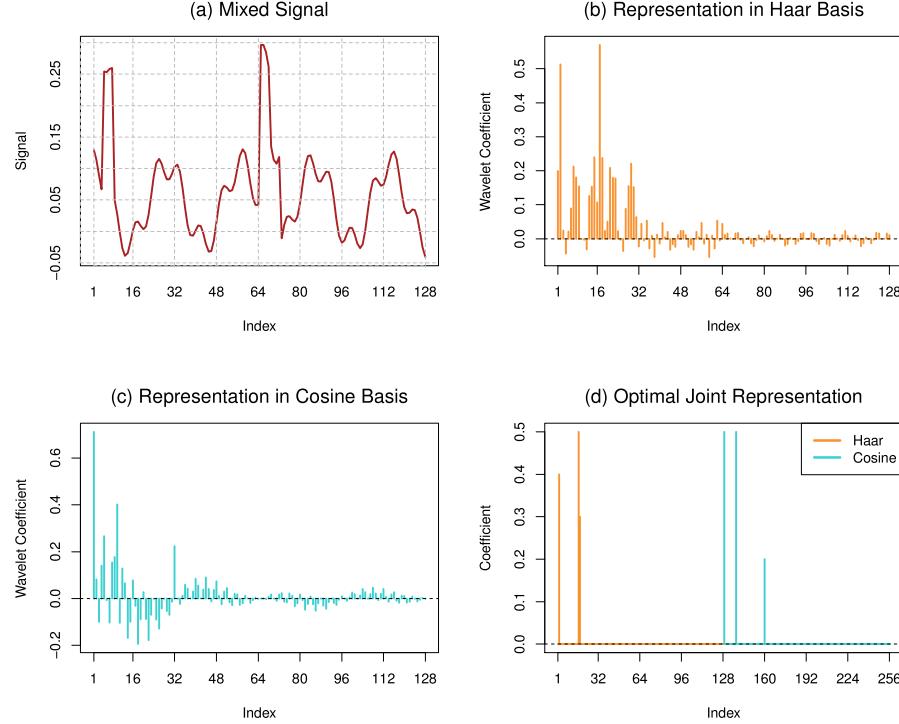
As discussed previously, Figure 10.1 provides one illustration of signal approximation within the Haar wavelet basis. In particular, the bottom-left panel shows the approximated signal  $\hat{\theta}^{64}$ , based on retaining only half of the Haar wavelet coefficients ( $k/p = 64/128 = 0.5$ ).

### 10.2.3 Reconstruction in Overcomplete Bases

Orthonormal bases, though useful in many ways, have a number of shortcomings. In particular, there is a limited class of signals that have sparse representations in any given orthonormal basis. For instance, Fourier bases are particularly well-suited to reconstructing signals with a globally periodic structure; in contrast, the Haar basis with its localized basis vectors is rather poor at capturing this kind of structure. On the other hand, the Haar basis excels at capturing step discontinuities, whereas such jumps have very nonsparse representations in the Fourier basis.

Based on this intuition, it is relatively straightforward to construct signals that are in some sense “simple,” but fail to have sparse representations in a classical orthonormal basis. As an illustration, panel (a) of Figure 10.4 shows a signal  $\theta^* \in \mathbb{R}^{128}$  that contains a mixture of both some globally periodic components, and some rapid (nearly discontinuous) transitions. As shown in panel (b), its Haar coefficients  $\beta^* = \Psi^T \theta^*$  are relatively dense, because many basis vectors are required to reconstruct the globally periodic portion of the signal. Similarly, as shown in panel (c), its representation  $\alpha^* = \Phi^T \theta^*$  in the discrete cosine basis (a type of Fourier representation) is also relatively dense. Due to this lack of sparsity, neither basis alone will provide a good sparse approximation to the original signal.

However, suppose that we allow the reconstruction to use subsets of vectors from *both bases* simultaneously; in this case, it might be possible to obtain a significantly more accurate, or even exact, sparse approximation. To set up



**Figure 10.4** (a) Original signal  $\theta^* \in \mathbb{R}^p$  with  $p = 128$ . (b) Representation  $\Psi^T \theta^*$  in the Haar basis. (c) Representation  $\Phi^T \theta^*$  in the discrete cosine basis. (d) Coefficients  $(\hat{\alpha}, \hat{\beta}) \in \mathbb{R}^p \times \mathbb{R}^p$  of the optimally sparse joint representation obtained by solving basis pursuit linear program (10.11).

the problem more precisely, given a pair of orthonormal bases  $\{\psi_j\}_{j=1}^p$  and  $\{\phi_j\}_{j=1}^p$ , let us consider reconstructions of the form

$$\underbrace{\sum_{j=1}^p \alpha_j \phi_j}_{\Phi \alpha} + \underbrace{\sum_{j=1}^p \beta_j \psi_j}_{\Psi \beta} \quad \text{such that } \|\alpha\|_0 + \|\beta\|_0 \leq k, \quad (10.8)$$

and the associated optimization problem

$$\underset{(\alpha, \beta) \in \mathbb{R}^p \times \mathbb{R}^p}{\text{minimize}} \|\theta^* - \Phi \alpha - \Psi \beta\|_2^2 \quad \text{such that } \|\alpha\|_0 + \|\beta\|_0 \leq k. \quad (10.9)$$

Despite its superficial similarity to our earlier  $k$ -term approximation problem (10.5), the optimization problem (10.9) is actually very difficult to solve. Unlike the earlier case, we are now working in an *overcomplete basis* described by the union of the two bases  $\Phi$  and  $\Psi$ .

Nonetheless, we can resort to our usual relaxation of the  $\ell_0$ -“norm,” and consider the following convex program

$$\underset{(\alpha, \beta) \in \mathbb{R}^p \times \mathbb{R}^p}{\text{minimize}} \|\theta^* - \Phi\alpha - \Psi\beta\|_2^2 \quad \text{such that } \|\alpha\|_1 + \|\beta\|_1 \leq R, \quad (10.10)$$

where  $R > 0$  is a user-defined radius. This program is a constrained version of the lasso program, also referred to as the relaxed basis-pursuit program. When seeking a perfect reconstruction, we can also consider the even simpler problem

$$\underset{(\alpha, \beta) \in \mathbb{R}^p \times \mathbb{R}^p}{\text{minimize}} \|\alpha\|_1 + \|\beta\|_1 \quad \text{such that } \theta^* = [\Phi \quad \Psi] \begin{bmatrix} \alpha \\ \beta \end{bmatrix}. \quad (10.11)$$

This problem is a linear program (LP), often referred to as the basis-pursuit linear program.

Returning to the example discussed in Figure 10.4, panel (d) shows the optimal coefficients  $(\hat{\alpha}, \hat{\beta}) \in \mathbb{R}^p \times \mathbb{R}^p$  obtained by solving the basis pursuit LP (10.11). We thus find that the original signal in panel (a) can be generated by an extremely sparse combination, with only six nonzero coefficients, in the overcomplete basis formed by combining the Haar and discrete cosine representations. In fact, this is the sparsest possible representation of the signal, so that in this case, solving the basis pursuit LP (10.11) is equivalent to solving the  $\ell_0$ -constrained problem (10.9).

Naturally, the reader might wonder about the generality of this phenomenon—namely, when does the solution to the basis pursuit LP coincide with the computationally difficult  $\ell_0$ -problem (10.9)? As it turns out, the answer to this question depends on the degree of incoherence between the two bases  $\Phi$  and  $\Psi$ , as we explore at more length in Section 10.4.

### 10.3 Random Projection and Approximation

In the previous sections, we discussed approximating a signal by computing its projection onto each of a fixed set of basis functions. We now turn to the use of random projections in signal approximation. This allows one to use a smaller number of (random) basis functions than is required under a fixed basis. We will combine this with an  $\ell_1$ -penalty on the coefficient of each projection, leading to the idea of *compressed sensing*.

A random projection of a signal  $\theta^*$  is a measurement of the form

$$y_i = \langle z_i, \theta^* \rangle = \sum_{j=1}^p z_{ij} \theta_j^*, \quad (10.12)$$

where  $z_i \in \mathbb{R}^p$  is a random vector. The idea of using random projections for dimensionality reduction and approximation is an old one, dating back (at least) to classical work on metric embedding and spherical sections of convex

bodies (see the bibliographic section for more details). We begin by describing a classical use of random projection, namely for embedding data while preserving distances between points, and then move on to discuss compressed sensing, which combines random projections with  $\ell_1$ -relaxation.

### 10.3.1 Johnson–Lindenstrauss Approximation

As one application of random projection, let us consider how they can be used to approximate a finite collection of vectors, say representing some dataset. The technique that we describe is often known as Johnson–Lindenstrauss embedding, based on the authors who pioneered its use in studying the more general problem of metric embedding (see the bibliographic section for more details). Suppose that we are given  $M$  data points  $\{u_1, \dots, u_M\}$  lying in  $\mathbb{R}^p$ . If the data dimension  $p$  is large, then it might be too expensive to store the dataset. In this setting, one approach is to design a dimension-reducing mapping  $F : \mathbb{R}^p \rightarrow \mathbb{R}^N$  with  $N \ll p$  that preserves some “essential” features of the dataset, and then store only the projected dataset  $\{F(u_1), \dots, F(u_M)\}$ . For example, since many algorithms operate on datasets by computing pairwise distances, we might be interested in a mapping  $F$  with the guarantee that for some tolerance  $\delta \in (0, 1)$ , we have

$$(1-\delta) \|u_i - u_{i'}\|_2^2 \leq \|F(u_i) - F(u_{i'})\|_2^2 \leq (1+\delta) \|u_i - u_{i'}\|_2^2 \text{ for all pairs } i \neq i'. \quad (10.13)$$

Of course, this is always possible if the projected dimension  $N$  is large enough, but the goal is to do it with relatively small  $N$ .

As shown in the seminal work of Johnson and Lindenstrauss, random projections provide one method for designing such approximate distance-preserving embeddings. The construction is straightforward:

- (a) Form a random matrix  $\mathbf{Z} \in \mathbb{R}^{N \times p}$  with each  $Z_{ij} \sim N(0, 1)$ , i.i.d., and define the linear mapping  $F : \mathbb{R}^p \rightarrow \mathbb{R}^N$  via

$$F(u) := \frac{1}{\sqrt{N}} \mathbf{Z} u. \quad (10.14)$$

- (b) Compute the projected dataset  $\{F(u_1), F(u_2), \dots, F(u_M)\}$ .

An interesting question is the following: for a given tolerance  $\delta \in (0, 1)$  and number of data points  $M$ , how large should we choose the projected dimension  $N$  to ensure that approximate distance-preserving property (10.13) holds with high probability? In Exercises 10.1 and 10.2, we show that this property holds with high probability as long as  $N > \frac{c}{\delta^2} \log M$  for some universal constant  $c$ . Thus, the dependence on the number  $M$  of data points scales logarithmically, and hence is very mild.

As a particular example, suppose that our goal is to obtain a compressed representation of all Boolean vectors  $u \in \{-1, 1\}^p$  that are  $k$ -sparse.<sup>1</sup> By a

---

<sup>1</sup>A vector  $u \in \mathbb{R}^p$  is  $k$ -sparse if only  $k \leq p$  elements are nonzero.

simple counting argument, there are  $M = 2^k \binom{p}{k}$  such vectors. Noting that  $\log M \leq k \log \left(\frac{e^2 p}{k}\right)$ , we see that a projection dimension  $N > \frac{c}{\delta^2} k \log \left(\frac{e^2 p}{k}\right)$  suffices to preserve pairwise distances up to  $\delta$ -accuracy between all  $k$ -sparse Boolean vectors. This example provides a natural segue to the method of compressed sensing, which combines random projections with  $\ell_1$ -relaxation.

### 10.3.2 Compressed Sensing

Compressed sensing is a combination of random projection and  $\ell_1$ -regularization that was introduced in independent work by Candes and Tao (2005) and Donoho (2006); since this pioneering work, an extensive literature on the topic has developed, with numerous applications including medical imaging and single-pixel cameras, among others. In this section, we provide a brief introduction to the basic ideas.

The motivation for compressed sensing is the inherent wastefulness of the standard method for compressing signals in an orthogonal basis. As described in Section 10.2.2, this approach involves first computing the full vector  $\beta^* \in \mathbb{R}^p$  of basis coefficients (step 1 on page 274), and then discarding a large fraction of them in order to obtain the  $k$ -sparse approximation  $\widehat{\theta}^k$  of the underlying signal  $\theta^*$  (step 2). Given that we end up discarding most of the basis coefficients, is it really necessary to compute all of them? Of course, if one knew *a priori* which subset of  $k$  coefficients were to be retained for the sparse approximation  $\widehat{\theta}^k$ , then one could simply compute this subset of basis coefficients. We refer to this approach as the oracle technique. Of course, it is unimplementable in practice, since we don't know *a priori* which coefficients are the most relevant for a given signal.

The power of compressed sensing is that it enables one to mimic the behavior of the oracle with very little computational overhead. It combines random projection with  $\ell_1$ -minimization in the following way. Instead of pre-computing all of the basis coefficients  $\beta^* = \Psi^T \theta^*$ , suppose that we compute some number  $N$  of random projections, say of the form  $y_i = \langle z_i, \theta^* \rangle$ , for  $i = 1, 2, \dots, N$ . We are free to choose the form of the random projection vectors  $z_i \in \mathbb{R}^p$ , and we discuss a number of reasonable choices shortly.

Thus, the setup of our problem is as follows: we are given an  $N$ -vector  $\mathbf{y}$  of random projections of the signal  $\theta^*$ . Also known to us is the  $N \times p$  random matrix  $\mathbf{Z}$  with  $i^{th}$  row  $z_i$ , used to compute the random projections; we refer to  $\mathbf{Z}$  as the *design matrix* or measurement matrix. The observation vector  $\mathbf{y}$  and design matrix  $\mathbf{Z}$  are linked to the unknown signal  $\theta^* \in \mathbb{R}^N$  by the matrix-vector equation  $\mathbf{y} = \mathbf{Z}\theta^*$ , and our goal is to recover (exactly or approximately) the signal  $\theta^* \in \mathbb{R}^p$ . See Figure 10.5(a) for an illustration of this setup.

At first sight, the problem seems very simple, since determining  $\theta^*$  amounts to solving a linear system. However, for this method to be cheaper than the standard approach (and therefore of practical interest), it is essential that the number of projections (or sample size)  $N$  be much smaller than the ambient dimension  $p$ . For this reason, the linear system  $\mathbf{y} = \mathbf{Z}\theta^*$  is highly under-

Figure 10.5 consists of two parts, (a) and (b), illustrating linear systems.

(a) An under-determined linear system  $\mathbf{y} = \mathbf{Z}\theta^*$ . On the left, a green vertical bar labeled  $\mathbf{y}$  is followed by an equals sign. To the right of the equals sign is a gray rectangular matrix labeled  $\mathbf{Z}$  with the text  $N \times p$  inside. To the right of  $\mathbf{Z}$  is a purple vertical bar labeled  $\theta^*$ .

(b) Equivalent representation of the linear system. On the left, a green vertical bar labeled  $\mathbf{y}$  is followed by an equals sign. To the right of the equals sign is a gray rectangular matrix labeled  $\mathbf{Z}$  with the text  $N \times p$  inside. To the right of  $\mathbf{Z}$  is an orange rectangular matrix labeled  $\Psi$ . To the right of  $\Psi$  is a vertical bar labeled  $\beta^*$ , which is divided into two colored segments: red at the top labeled  $k$  and blue at the bottom labeled  $p - k$ .

**Figure 10.5** (a) An under-determined linear system  $\mathbf{y} = \mathbf{Z}\theta^*$ : Each row  $z_i$  of the  $N \times p$  measurement matrix  $\mathbf{Z}$  defines the random projection  $y_i = \langle z_i, \theta^* \rangle$ . The signal  $\theta^* \in \mathbb{R}^p$  need not be sparse in the canonical basis. (b) Equivalent representation of the linear system: Basis coefficients  $\beta^* = \Psi^T \theta^*$  are assumed to be  $k$ -sparse. This transformation defines an equivalent linear system  $\mathbf{y} = \tilde{\mathbf{Z}}\beta^*$  with sparsity that can be exploited.

determined: there are many signals  $\theta$  that are consistent with the observed random projections.

However, if we also have the additional side-information that  $\Psi^T \theta^*$  is sparse, then it could be possible to recover  $\theta^*$  exactly, even though the linear system on its own is under-determined. In an ideal world, we would like to exploit this sparsity by solving the  $\ell_0$ -based problem

$$\underset{\theta \in \mathbb{R}^p}{\text{minimize}} \|\Psi^T \theta\|_0 \quad \text{such that } \mathbf{y} = \mathbf{Z}\theta. \quad (10.15)$$

The  $\ell_0$ -problem is combinatorial, and known to be computationally intractable (NP-hard) in general; thus, we are led to consider the  $\ell_1$ -relaxation

$$\underset{\theta \in \mathbb{R}^p}{\text{minimize}} \|\Psi^T \theta\|_1 \quad \text{such that } \mathbf{y} = \mathbf{Z}\theta. \quad (10.16)$$

Equivalently, we can write this problem in terms of the transform coefficient vector  $\beta \in \mathbb{R}^p$ , namely as

$$\underset{\beta \in \mathbb{R}^p}{\text{minimize}} \|\beta\|_1 \quad \text{such that } \mathbf{y} = \tilde{\mathbf{Z}}\beta, \quad (10.17)$$

where we have defined the transformed matrix  $\tilde{\mathbf{Z}} := \mathbf{Z}\Psi \in \mathbb{R}^{N \times p}$ . See Figure 10.5(b) for an illustration of this transformed linear system.

In summary, then, the method of compressed sensing operates as follows:

1. For a given sample size  $N$ , compute the random projections  $y_i = \langle z_i, \theta^* \rangle$  for  $i = 1, 2, \dots, N$ .
2. Estimate the signal  $\theta^*$  by solving the linear program (10.16) to obtain  $\hat{\theta}$ . (Equivalently, solve the linear program (10.17) to obtain  $\hat{\beta}$ , and set  $\hat{\theta} = \Psi\hat{\beta}$ .)

To be clear, we have actually described a family of procedures, depending on our choice of the random projection vectors  $\{z_i\}_{i=1}^N$ , or equivalently the transformed design matrix  $\tilde{\mathbf{Z}}$ . A variety of different design matrices  $\mathbf{Z}$  have been studied for the purposes of compressed sensing. Perhaps the simplest choice is to choose its entries  $z_{ij} \sim N(0, 1)$  in an i.i.d. manner, leading to a standard Gaussian random matrix. Other choices of matrices for compressed sensing include random Bernoulli matrices formed with i.i.d. entries drawn as  $z_{ij} \in \{-1, +1\}$  with equal probability, as well as random submatrices of Fourier matrices.

When can compressed sensing succeed using a number of projections  $N$  less than the signal dimension  $p$ ? As we discuss in Section 10.4.2, it is sufficient that the columns of the transformed design matrix  $\tilde{\mathbf{Z}}$  be “incoherent”, and there are different measures of such incoherence. The simplest measure of incoherence is pairwise, based on the inner products between the columns of  $\tilde{\mathbf{Z}}$ . A more sophisticated notion of incoherence is the restricted isometry property (RIP), based on looking on the conditioning of submatrices of  $\tilde{\mathbf{Z}}$  consisting of up to  $k$  columns. An important fact is that the random design matrices discussed above satisfy RIP with high probability using a relatively small number of projections  $N$ . For instance, for the standard Gaussian or Bernoulli cases, it can be shown that RIP holds with high probability with as few as  $N = \Omega(k \log \frac{p}{k})$  samples, where  $k < p$  is the sparsity of the basis coefficient vector  $\beta^*$ . Note that any method—even the unimplementable oracle that already knew the support of  $\beta^*$ —would require at least  $N = k$  random projections for exact recovery. Thus, compressed sensing incurs a multiplicative overhead of only  $\mathcal{O}(\log(p/k))$  relative to oracle performance.

#### 10.4 Equivalence between $\ell_0$ and $\ell_1$ Recovery

Thus far, we have discussed a number of applications of  $\ell_1$ -norm regularization in signal processing, including sparse approximation in overcomplete bases (Section 10.2.3), and compressed sensing (Section 10.3.2). In both cases, the  $\ell_1$ -norm is introduced as a computationally tractable surrogate to optimization problems involving the intractable  $\ell_0$ -“norm.” Up to this point, we have not addressed in any depth an important question: when is solving the  $\ell_1$ -relaxation equivalent to solving the original  $\ell_0$ -problem?

More precisely, given an observation vector  $\mathbf{y} \in \mathbb{R}^p$  and a design matrix

$\mathbf{X} \in \mathbb{R}^{N \times p}$ , let us consider the two problems

$$\underset{\beta \in \mathbb{R}^p}{\text{minimize}} \|\beta\|_0 \quad \text{such that } \mathbf{X}\beta = \mathbf{y}, \quad (10.18)$$

and

$$\underset{\beta \in \mathbb{R}^p}{\text{minimize}} \|\beta\|_1 \quad \text{such that } \mathbf{X}\beta = \mathbf{y}. \quad (10.19)$$

This setup includes as a special case the problem of sparse approximation in an overcomplete basis, as discussed in Section 10.2.3; in this case, the observation  $\mathbf{y}$  is equal to the signal  $\theta^*$  to be approximated, and the design matrix  $\mathbf{X} = [\Phi \ \Psi]$ . It also includes the case of compressed sensing, where  $\mathbf{X}$  is the transformed version of the random projection matrix (namely,  $\tilde{\mathbf{Z}}$  in our earlier notation).

#### 10.4.1 Restricted Nullspace Property

Suppose that the  $\ell_0$ -based problem (10.18) has a unique optimal solution, say  $\beta^* \in \mathbb{R}^p$ . Our interest is in understanding when  $\beta^*$  is also the unique optimal solution of the  $\ell_1$ -based problem (10.19), in which case we say that the basis pursuit LP is *equivalent* to  $\ell_0$ -recovery. Remarkably, there exists a very simple necessary and sufficient condition on the design matrix  $\mathbf{X}$  for this equivalence to hold. For a given subset  $S \subseteq \{1, 2, \dots, p\}$ , it is stated in terms of the set

$$\mathbb{C}(S) := \{\beta \in \mathbb{R}^p \mid \|\beta_{S^c}\|_1 \leq \|\beta_S\|_1\}. \quad (10.20)$$

The set  $\mathbb{C}(S)$  is a convex cone, containing all vectors that are supported on  $S$ , and other vectors as well. Roughly, it corresponds to the cone of vectors that have most of their mass allocated to  $S$ . Given a matrix  $\mathbf{X} \in \mathbb{R}^{N \times p}$ , its nullspace is given by  $\text{null}(\mathbf{X}) = \{\beta \in \mathbb{R}^p \mid \mathbf{X}\beta = \mathbf{0}\}$ .

*Definition 10.1. Restricted nullspace property.* For a given subset  $S \subseteq \{1, 2, \dots, p\}$ , we say that the design matrix  $\mathbf{X} \in \mathbb{R}^{N \times p}$  satisfies the *restricted nullspace property* over  $S$ , denoted by  $\text{RN}(S)$ , if

$$\text{null}(\mathbf{X}) \cap \mathbb{C}(S) = \{0\}. \quad (10.21)$$

In words, the  $\text{RN}(S)$  property holds when the only element of the cone  $\mathbb{C}(S)$  that lies within the nullspace of  $\mathbf{X}$  is the all-zeroes vector. The following theorem highlights the significance of this property:

*Theorem 10.1.  $\ell_0$  and  $\ell_1$  equivalence.* Suppose that  $\beta^* \in \mathbb{R}^p$  is the unique solution to the  $\ell_0$  problem (10.18), and has support  $S$ . Then the basis pursuit relaxation (10.19) has a unique solution equal to  $\beta^*$  if and only if  $\mathbf{X}$  satisfies the  $\text{RN}(S)$  property.

The proof of Theorem 10.1 is relatively short, and is provided in Section 10.4.3.

Since the subset  $S$  is not known in advance—indeed, it is usually what we

are trying to determine—it is natural to seek matrices that satisfy a uniform version of the restricted nullspace property. For instance, we say that the uniform RN property of order  $k$  holds if  $\text{RN}(S)$  holds for all subsets of size at most  $k$ . In this case, we are guaranteed that the  $\ell_1$ -relaxation succeeds for any vector supported on any subset of size at most  $k$ .

#### 10.4.2 Sufficient Conditions for Restricted Nullspace

Of course, in order for Theorem 10.1 to be useful in practice, we need to verify the restricted nullspace property. A line of work has developed various conditions for certifying the uniform RN property. The simplest and historically earliest condition is based on the *pairwise incoherence*

$$\nu(\mathbf{X}) := \max_{j,j'=1,2,\dots,p} \frac{|\langle \mathbf{x}_j, \mathbf{x}_{j'} \rangle|}{\|\mathbf{x}_j\|_2 \|\mathbf{x}_{j'}\|_2}. \quad (10.22)$$

For centered  $\mathbf{x}_j$  this is the maximal absolute pairwise correlation. When  $\mathbf{X}$  is rescaled to have unit-norm columns, an equivalent representation is given by  $\nu(\mathbf{X}) = \max_{j \neq j'} |\langle \mathbf{x}_j, \mathbf{x}_{j'} \rangle|$ , which illustrates that the pairwise incoherence measures how close the Gram matrix  $\mathbf{X}^T \mathbf{X}$  is to the  $p$ -dimensional identity matrix in an element-wise sense.

The following result shows that having a low pairwise incoherence is sufficient to guarantee exactness of the basis pursuit LP:

**Proposition 10.1. Pairwise incoherence implies RN.** Suppose that for some integer  $k \in \{1, 2, \dots, p\}$ , the pairwise incoherence satisfies the bound  $\nu(\mathbf{X}) < \frac{1}{3k}$ . Then  $\mathbf{X}$  satisfies the uniform RN property of order  $k$ , and hence, the basis pursuit LP is exact for all vectors with support at most  $k$ .

See Section 10.4.3 for the proof of this claim.

An attractive feature of the pairwise incoherence is that it is easily computed; in particular, in  $\mathcal{O}(Np^2)$  time. A disadvantage is that it provides very conservative bounds that do not always capture the actual performance of  $\ell_1$ -relaxation in practice. For instance, consider the matrix  $\mathbf{X} = [\Phi \ \Psi]$ , as arises in the overcomplete basis problem (10.11). We can numerically compute the incoherence, say for the discrete cosine and Haar bases in dimension  $p = 128$ , as illustrated in Figure 10.4. We find that Proposition 10.1 guarantees exact recovery of all signals with sparsity  $k = 1$ , whereas in practice, the  $\ell_1$ -relaxation works for much larger values of  $k$ .

For random design matrices, such as those that arise in compressed sensing, one can use probabilistic methods to bound the incoherence. For instance, consider a random matrix  $\mathbf{X} \in \mathbb{R}^{N \times p}$  with i.i.d.  $N(0, 1/N)$  entries. Here we have rescaled the variance so that the columns of  $\mathbf{X}$  have expected norm equal to one. For such a matrix, one can show that  $\nu(\mathbf{X}) \lesssim \sqrt{\frac{\log p}{N}}$  with high probability as  $(N, p)$  tend to infinity (see Exercise 10.5). Combined with Proposition 10.1, we conclude that the  $\ell_1$ -relaxation (10.16) will exactly recover all

signals with sparsity at most  $k$  as long as the number of projections scales as  $N \gtrsim k^2 \log p$ .

In fact, for random designs and compressed sensing, this scaling can be sharpened using the *restricted isometry property* (RIP). Recall that the incoherence condition (10.22) is a measure of the orthonormality of pairs of columns of the design matrix  $\mathbf{X}$ . The notion of restricted isometry is to constrain much larger submatrices of  $\mathbf{X}$  to have nearly orthogonal columns.

**Definition 10.2. Restricted isometry property.** For a tolerance  $\delta \in (0, 1)$  and integer  $k \in \{1, 2, \dots, p\}$ , we say that  $\text{RIP}(k, \delta)$  holds if

$$\|\mathbf{X}_S^T \mathbf{X}_S - \mathbf{I}_{k \times k}\|_{\text{op}} \leq \delta \quad (10.23)$$

for all subsets  $S \subset \{1, 2, \dots, p\}$  of cardinality  $k$ .

We recall here that  $\|\cdot\|_{\text{op}}$  denotes the operator norm, or maximal singular value of a matrix. Due to the symmetry of  $\mathbf{X}_S^T \mathbf{X}_S$ , we have the equivalent representation

$$\|\mathbf{X}_S^T \mathbf{X}_S - \mathbf{I}_{k \times k}\|_{\text{op}} = \sup_{\|u\|_2=1} |u^T (\mathbf{X}_S^T \mathbf{X}_S - \mathbf{I}_{k \times k}) u| = \sup_{\|u\|_2=1} |\|\mathbf{X}_S u\|_2^2 - 1|.$$

Thus, we see that  $\text{RIP}(k, \delta)$  holds if and only if for all subsets  $S$  of cardinality  $k$ , we have

$$\frac{\|\mathbf{X}_S u\|_2^2}{\|u\|_2^2} \in [1 - \delta, 1 + \delta] \quad \text{for all } u \in \mathbb{R}^k \setminus \{0\},$$

hence the terminology of restricted isometry.

The following result shows that RIP is a sufficient condition for the restricted nullspace to hold:

**Proposition 10.2. RIP implies restricted nullspace.** If  $\text{RIP}(2k, \delta)$  holds with  $\delta < 1/3$ , then the uniform RN property of order  $k$  holds, and hence the  $\ell_1$ -relaxation is exact for all vectors supported on at most  $k$  elements.

We work through the proof of a slightly weaker version of this claim in Exercise 10.8. Observe that the  $\text{RIP}(2k, \delta)$  condition imposes constraints on a huge number of submatrices, namely  $\binom{p}{2k}$  in total. On the other hand, as opposed to the pairwise incoherence condition, the actual RIP constant  $\delta$  has no dependence on  $k$ .

From known results in random matrix theory, various choices of random projection matrices  $\mathbf{X}$  satisfy RIP with high probability as long as  $N \gtrsim k \log \frac{ep}{k}$ . Among other matrix ensembles, this statement applies to a standard Gaussian random matrix  $\mathbf{X}$  with i.i.d.  $N(0, \frac{1}{N})$  entries; see Exercise 10.6 for details. Thus, we see that the RIP-based approach provides a certificate for exact recovery based on far fewer samples than pairwise incoherence, which as previously discussed, provides guarantees when  $N \gtrsim k^2 \log p$ . On the other hand, a major drawback of RIP is that—in sharp contrast to the pairwise incoherence—it is very difficult to verify in practice due to the number  $\binom{p}{2k}$  of submatrices.

### 10.4.3 Proofs

We conclude the chapter by providing proofs of the claims given in the preceding section.

#### 10.4.3.1 Proof of Theorem 10.1

First, suppose that  $\mathbf{X}$  satisfies the  $\text{RN}(S)$  property. Let  $\widehat{\beta} \in \mathbb{R}^p$  be any optimal solution to the basis pursuit LP (10.19), and define the error vector  $\Delta := \widehat{\beta} - \beta^*$ . Our goal is to show that  $\Delta = 0$ , and in order to do so, it suffices to show that  $\Delta \in \text{null}(\mathbf{X}) \cap \mathbb{C}(S)$ . On the one hand, since  $\beta^*$  and  $\widehat{\beta}$  are optimal (and hence feasible) solutions to the  $\ell_0$  and  $\ell_1$  problems, respectively, we are guaranteed that  $\mathbf{X}\beta^* = \mathbf{y} = \mathbf{X}\widehat{\beta}$ , showing that  $\mathbf{X}\Delta = 0$ . On the other hand, since  $\beta^*$  is also feasible for the  $\ell_1$ -based problem (10.19), the optimality of  $\widehat{\beta}$  implies that  $\|\widehat{\beta}\|_1 \leq \|\beta^*\|_1 = \|\beta_S^*\|_1$ . Writing  $\widehat{\beta} = \beta^* + \Delta$ , we have

$$\begin{aligned} \|\beta_S^*\|_1 &\geq \|\widehat{\beta}\|_1 = \|\beta_S^* + \Delta_S\|_1 + \|\Delta_{S^c}\|_1 \\ &\geq \|\beta_S^*\|_1 - \|\Delta_S\|_1 + \|\Delta_{S^c}\|_1, \end{aligned}$$

where the final bound follows by triangle inequality. Rearranging terms, we find that  $\Delta \in \mathbb{C}(S)$ ; since  $\mathbf{X}$  satisfies the  $\text{RN}(S)$  condition by assumption, we conclude that  $\Delta = 0$  as required.

We lead the reader through a proof of the converse in Exercise 10.4.

#### 10.4.3.2 Proof of Proposition 10.1

We may assume without loss of generality (rescaling as needed) that  $\|\mathbf{x}_j\|_2 = 1$  for all  $j = 1, 2, \dots, p$ . To simplify notation, let us assume an incoherence condition of the form  $\nu(\mathbf{X}) < \frac{\delta}{k}$  for some  $\delta > 0$ , and verify the sufficiency of  $\delta = 1/3$  in the course of the argument.

For an arbitrary subset  $S$  of cardinality  $k$ , suppose that  $\beta \in \mathbb{C}(S) \setminus \{0\}$ . It suffices to show that  $\|\mathbf{X}\beta\|_2^2 > 0$ , and so we begin with the lower bound

$$\|\mathbf{X}\beta\|_2^2 \geq \|\mathbf{X}_S \beta_S\|_2^2 + 2\beta_S^T \mathbf{X}_S^T \mathbf{X}_{S^c} \beta_{S^c}. \quad (10.24)$$

On one hand, we have

$$\begin{aligned} 2 \left| \beta_S^T \mathbf{X}_S^T \mathbf{X}_{S^c} \beta_{S^c} \right| &\leq 2 \left| \sum_{i \in S} \sum_{j \in S^c} |\beta_i| |\beta_j| |\langle \mathbf{x}_i, \mathbf{x}_j \rangle| \right| \\ &\stackrel{(i)}{\leq} 2 \|\beta_S\|_1 \|\beta_{S^c}\|_1 \nu(\mathbf{X}) \\ &\stackrel{(ii)}{\leq} \frac{2\delta \|\beta_S\|_1^2}{k} \\ &\stackrel{(iii)}{\leq} 2\delta \|\beta_S\|_2^2, \end{aligned}$$

where inequality (i) uses the definition (10.22) of the pairwise incoherence;

inequality (ii) exploits the assumed bound on  $\nu(\mathbf{X})$  combined with the fact that  $\beta \in \mathbb{C}(S)$ ; and inequality (iii) uses the fact that  $\|\beta_S\|_1^2 \leq k\|\beta_S\|_2^2$ , by Cauchy–Schwarz, since the cardinality of  $S$  is at most  $k$ . Consequently, we have established that

$$\|\mathbf{X}\beta\|_2^2 \geq \|\mathbf{X}_S\beta_S\|_2^2 - 2\delta\|\beta_S\|_2^2. \quad (10.25)$$

In order to complete the proof, it remains to lower bound  $\|\mathbf{X}_S\beta_S\|_2^2$ . Letting  $\|\cdot\|_{\text{op}}$  denote the operator norm (maximum singular value) of a matrix, we have

$$\|\mathbf{X}_S^T \mathbf{X}_S - \mathbf{I}_{k \times k}\|_{\text{op}} \leq \max_{i \in S} \sum_{j \in S \setminus \{i\}} |\langle x_i, x_j \rangle| \leq k \frac{\delta}{k} = \delta.$$

Consequently,  $\|\mathbf{X}_S\beta_S\|_2^2 \geq (1-\delta)\|\beta_S\|_2^2$ , and combined with the bound (10.25), we conclude that  $\|\mathbf{X}\beta\|_2^2 > (1 - 3\delta)\|\beta_S\|_2^2$ , so that  $\delta = 1/3$  is sufficient as claimed.

### Bibliographic Notes

There is an extensive literature on the sparsity of images and other signal classes when represented in wavelet and other multiscale bases (Field 1987, Ruderman 1994, Wainwright, Simoncelli and Willsky 2001, Simoncelli 2005). Sparse approximation in overcomplete bases is discussed in various papers (Donoho and Stark 1989, Chen et al. 1998, Donoho and Huo 2001, Elad and Bruckstein 2002, Feuer and Nemirovski 2003). The multiscale basis illustrated in Figure 10.2 is known as the steerable pyramid (Simoncelli and Freeman 1995). Random projection is a widely used technique in computer science and numerical linear algebra (Vempala 2004, Mahoney 2011, Pilanci and Wainwright 2014, e.g.). Johnson and Lindenstrauss (1984) proved the lemma that now bears their name in the context of establishing the existence of metric embeddings, using random projection as a proof technique. Compressed sensing was introduced independently by Candès, Romberg and Tao (2006) and Donoho (2006). Lustig, Donoho, Santos and Pauly (2008) discuss the applications of compressed sensing to medical imaging, whereas Candès and Wakin (2008) discuss various applications in signal processing.

The restricted nullspace property is discussed in Donoho and Huo (2001), Feuer and Nemirovski (2003), and Cohen, Dahmen and DeVore (2009). Various authors (Donoho and Huo 2001, Elad and Bruckstein 2002, Feuer and Nemirovski 2003) have studied the pairwise incoherence of overcomplete bases and other design matrices, as a sufficient condition for the restricted nullspace property. Candès and Tao (2005) introduced the restricted isometry property as a milder sufficient condition for the restricted nullspace property. For random matrices with i.i.d. sub-Gaussian rows, it follows from a combination of union bound and standard results in random matrix theory (Davidson and Szarek 2001, Vershynin 2012) that a sample size  $N > ck \log(\frac{ep}{k})$  suffices to