CHAPTER

13 Inverse Problems

Recovering high-resolution and high-quality signals from partial and noisy measurements is the dream behind inverse problems. It is present in most signal processing, from medical imaging to analog and digital conversions, from seismic exploration to high-definition video display. Measurements are modeled with a linear operator applied to the input signal, but this operator is typically not invertible. Computing a precise signal estimation is thus not possible without strong a priori information on the signal.

The input data belong to a space of limited dimension that defines the measurement resolution. Estimating the signal at this resolution is already challenging because some signal components are attenuated and can thus barely be discriminated from the noise. Partially inverting the operator can considerably amplify the noise and do more harm than good.

Linear estimators implement a partial and regularized linear inversion that is related to a singular value decomposition. Nonlinear estimations improve these estimations by capturing more prior information on the signal. Sparse estimation algorithms incorporate this prior information in the design of a dictionary in which the estimated signal has a sparse representation. The estimation procedure and the resulting estimation risk depend on the dictionary property.

If there exists a basis providing a sparse signal representation and with vectors that nearly diagonalize the measurement operator, then thresholding estimators can have a nearly minimax risk. When diagonal thresholding estimators fail, super-resolution may have its chance. Super-resolution is more ambitious and computes a signal estimation at a resolution that is higher than the data resolution. Pursuit decompositions can compute sparse super-resolution estimations in redundant and incoherent dictionaries, but super-resolution is not always possible.

Compressive sensing gives a new perspective on inverse problems by stabilizing super-resolution with random measurements. Randomness is a powerful tool to build incoherent dictionaries. Compressive sensing suggests designing new signal-acquisition devices that recover high-resolution signals from lower-resolution randomized measurements.

Ending the book with a cocktail party leads us to blind source separation. Recovering simultaneously several conversations or signals from few mixed

measurements is another super-resolution problem where sparsity again plays a central role.

13.1 LINEAR INVERSE ESTIMATION

Let us consider measurements obtained with a linear operator applied to an incoming analog signal $\bar{f}(x)$ to which noise is added:

$$Y[q] = \overline{U}\bar{f}[q] + W[q], \tag{13.1}$$

where $\overline{U}\bar{f}[q] = \langle \bar{f}, \bar{u}_q \rangle$ is the measurement output of a sensor. The operator $\overline{U}$ and the noise variance are supposed to be known or measured with a calibration procedure.

To invert the degradation numerically, $\overline{U}$ is factorized into a stable sampling operator $\overline{\Phi}_s$ followed by a discrete operator U, which carries the degradation and may provide less than N measurements:

$$Y[q] = U\overline{\Phi}_s\bar{f}[q] + W[q]. \tag{13.2}$$

As explained in Section 3.1.3, the sampling operator $\overline{\Phi}_s\bar{f}$ projects $\bar{f}$ over a Riesz basis of an approximation space $\mathbf{U}_N$. It is partially inverted by a discrete-to-analog converter that recovers the orthogonal projection of $\bar{f}$ over $\mathbf{U}_N$.

The goal is to recover the best possible estimate of the high-resolution signal $f[n] = \overline{\Phi}_s\bar{f}[n]$ from

$$Y[q] = Uf[q] + W[q]. \tag{13.3}$$

Let Q be the dimension of the image space $\mathbf{ImU}$ of the operator U. Linear estimators recover the projection of f in a space of dimension at most Q and thus do not provide any super-resolution. They are introduced by imposing a regularity condition on the solution through a quadratic variational problem. This will lead us to regularized singular value decompositions.

13.1.1 Quadratic and Tikhonov Regularizations

Suppose that f has some form of regularity, expressed by a regularization operator Φ that yields small energy coefficients. A linear estimation $\tilde{F}$ of f is computed from $Y = Uf + W \in \mathbb{R}^N$ as a solution of a quadratic optimization

$$\tilde{F} = \underset{h \in \mathbb{R}^N}{\operatorname{argmin}} \|\Phi h\|^2 \quad \text{subject to} \quad \|Uh - Y\|^2 \leqslant \varepsilon, \tag{13.4}$$

where ε is of the order of the noise energy $\|W\|^2$. If W is a Gaussian white noise of variance σ^2, then $E\{\|W\|^2\} = N\sigma^2$. If the noise W is not white but has an invertible covariance K_W, then $\|Uh - Y\|^2$ is replaced by $\|K_W^{-1/2}(Uh - Y)\|^2$ where $K_W^{-1/2}$ performs a "whitening" of the error $Uh - Y$ before applying the Euclidean norm.

Minimizing $\|\Phi \tilde{F}\|^2$ yields coefficients $\Phi \tilde{F}[n]$ of small amplitude that are rather uniformly spread. A Tikhonov regularization corresponds to a finite different approximation of a first-order derivative or of a gradient operator $\Phi = \vec{\nabla}$ in multiple dimensions. In this case, the solution $\tilde{F}$ has a finite-energy derivative and is thus differentiable in the sense of Sobolev. Section 13.3 gives examples of linear Tikhonov regularizations to estimate missing image pixels.

Since (13.4) is a strictly convex minimization, its solution can be computed as a solution of a Lagrangian minimization

$$\tilde{F} = \operatorname*{argmin}_{h \in \mathbb{R}^N} \frac{1}{2} \|Uh - Y\|^2 + T^2 \|\Phi h\|^2, \tag{13.5}$$

where T is adjusted as a function of ε. The theorem 13.1 computes the resulting linear estimator.

Theorem 13.1. The solution of the quadratic minimization

$$\tilde{F} = \operatorname*{argmin}_{h \in \mathbb{R}^N} \|Y - Uh\|^2 + T^2 \|\Phi h\|^2 \tag{13.6}$$

is the linear estimator

$$\tilde{F} = (U^* U + T^2 \Phi^* \Phi)^{-1} U^* Y. \tag{13.7}$$

Proof. Since the Lagrangian (13.5) is quadratic relative to the signal coordinates, its minimum is obtained by setting its partial derivatives to zero. The partial derivative of $\|Y - Uh\|^2 + T^2 \|\Phi h\|^2$ with respect to $h[n]$ is $U^*(Uh - Y)[n] + T^2 \Phi^* \Phi h[n]$. Setting these derivatives to zero leads to the optimal solution (13.7). ■

The linear estimator (13.7) applies U^* to the data Y, which projects these data in $\mathbf{Im}U^* = (\mathbf{Null}U)^\perp$, which is a space of dimension Q. It results that $\tilde{F}$ remains in a space of dimension Q and thus does not provide any super-resolution. This optimal linear estimator can be interpreted as a pseudo inversion of U followed by a linear denoising estimator.

Let U^+ be the pseudo inverse, defined as the operator that inverts the restriction of U to $\mathbf{Im}U$ and that is equal to 0 on $(\mathbf{Im}U)^\perp$. The range of U^+ is $(\mathbf{Null}U)^\perp$ and $U^+ U f$ is the orthogonal projection of f in $\mathbf{Im}U^* = (\mathbf{Null}U)^\perp$. Applying this pseudo inverse on the data Y gives

$$X = U^+ Y = U^+ U f + U^+ W \in (\mathbf{Null}U)^\perp. \tag{13.8}$$

The resulting inverted noise $Z = U^+ W$ is typically amplified. The estimator (13.7) applies a linear denoising operator D that reduces this amplified noise

$$\tilde{F} = DX \quad \text{with} \quad D = (U^* U + T^2 \Phi^* \Phi)^{-1} U^* U. \tag{13.9}$$

This linear inverse estimator is thus a linear pseudo inverse followed by a linear denoising estimator.

13.1.2 Singular Value Decompositions

Linear estimators that are solutions of a quadratic regularization problem may be written as a diagonal singular value decomposition. For deconvolution problems, it defines diagonal estimators in a Fourier basis. The matrix U^*U is symmetric and can thus be diagonalized in an orthonormal basis $\mathcal{B}_S = \{e_k\}_{0 \leq k < N}$, which is called a basis of *singular vectors*

$$U^*Ue_k = \lambda_k^2 e_k \quad \text{for} \quad 0 \leqslant k < N.$$

The eigenvalues $\lambda_k = \|Ue_k\|$ of U^*U are called *singular values* and define the *singular spectrum* of U. Let $\{e_k\}_{k \in \Gamma_Q}$ be the set of Q basis vectors such that $Ue_k \neq 0$. It is an orthonormal basis of $(\mathbf{NullU})^{\perp}$ and one can verify that $\{Ue_k\}_{k \in \Gamma_Q}$ is an orthogonal basis of $\mathbf{ImU}$.

If W is a white noise of variance σ^2, the covariance of the inverted noise $Z = U^+W$ is $K_Z = \sigma^2 U^+ U^{+*} = \sigma^2 (U^*U)^+$. The covariance K_Z is thus also diagonalized by $\mathcal{B}_S$. This basis is therefore a Karhunen-Loève basis of the inverted noise Z. The variance of the inverted noise in this basis is

$$E\{|\langle Z, e_k\rangle|^2\} = \langle K_Z e_k, e_k\rangle = \sigma^2 (U^* U)^+ e_k = \sigma^2 \lambda_k^{-2} \quad \text{for} \quad k \in \Gamma_Q. \tag{13.10}$$

Suppose that $\mathcal{B}_S$ also diagonalizes $\Phi^*\Phi$. The eigenvalues of $\Phi^*\Phi$ are then $\|\Phi e_k\|^2$. Since $\langle U^*Y, e_k\rangle = \langle Y, Ue_k\rangle$, it results from (13.7) that the solution of the quadratic regularization (13.6) is diagonal in this basis and can be written as

$$\tilde{F} = \sum_{k \in \Gamma_Q} \frac{\langle Y, Ue_k\rangle}{\lambda_k^2 + \sigma^2 \|\Phi e_k\|^2} e_k. \tag{13.11}$$

This is called a *singular value decomposition* (SVD). The coefficients $\|\Phi e_k\|^2$ regularize this estimation when the singular spectrum λ_k^2 becomes too small. Appropriate choices for $\|\Phi e_k\|^2$ lead to efficient operators for a variety of applications [111], the simplest one being $\Phi = \mathrm{Id}$ so that $\|\Phi e_k\| = 1$.

Oracle SVD Risk

To understand how to adjust $\|\Phi e_k\|$ in order to minimize the risk, let us consider a particular signal f. An *oracle SVD* operator chooses the regularization operator Φ depending on f to minimize the risk $E\{\|\tilde{F} - f\|^2\}$. Inserting (13.11) in $E\{\|\tilde{F} - f\|^2\}$ and setting to zero partial derivatives relative to $\|\Phi e_k\|$ proves that the risk is minimized by $\|\Phi e_k\| = |\langle f, e_k\rangle|^{-1}$, and the resulting minimum oral risk is

$$r_{\inf}(f) = \sum_{k \in \Gamma_Q} \frac{|\langle f, e_k\rangle|^2 \sigma^2}{|\langle f, e_k\rangle|^2 \lambda_k^2 + \sigma^2} + \sum_{k \notin \Gamma_Q} |\langle f, e_k\rangle|^2. \tag{13.12}$$

For a linear operator, $\|\Phi e_k\|$ must remain constant for all f in a signal set Θ. The oracle choice shows that one can find $\|\Phi e_k\|$, which produces a small maximum error over Θ if the energy of all $f \in \Theta$ is concentrated over a small number of fixed-basis vectors, and thus if $\mathcal{B}_S$ provides efficient linear approximations of vectors in Θ.

Deconvolutions

Many inverse problems involve a convolution operator $Uf[n] = f \circledast u[n]$ that we suppose to be circular to simplify border problems. In this case, $U^*Uf = f \circledast u \circledast \tilde{u}[n]$ with $\tilde{u}[n] = u[-n]$. The singular basis that diagonalizes UU^* is therefore the discrete Fourier basis

$$\mathcal{B}_S = \{e_k[n] = N^{-1/2} e^{i2\pi kn/N}\}_{0 \leq k < N}.$$

The singular spectrum is $\lambda_k^2 = |\hat{u}[k]|^2$, and **NullU** is the space of signals f with a Fourier transform $\hat{f}[k]$ that is nonzero only when $\hat{u}[k] = 0$.

An SVD deconvolution is obtained with a regularization operator Φ that is also a convolution $\Phi f[n] = f \circledast \phi[n]$, so that $\Phi^*\Phi$ is also diagonalized in the discrete Fourier basis with eigenvalues $\|\Phi e_k\|^2 = |\hat{\phi}[k]|^2$. The resulting diagonal SVD operator (13.11) is a convolution $\tilde{f} = Y \circledast d$ with a transfer function that is

$$\hat{d}[k] = \frac{\hat{u}^*[k]}{|\hat{u}[k]|^2 + \sigma^2 |\hat{\phi}[k]|^2}. \tag{13.13}$$

When u is a low-pass filter, it typically restores the lower frequencies and sets to zero larger frequencies where $|\hat{u}[k]| \ll \sigma |\hat{\phi}[k]|$. There is no super-resolution since no signal component is restored when $\hat{u}[k] = 0$. A super-resolution estimator would also recover frequencies that have been totally removed by U, which is the case of the sparse spike deconvolutions in Section 13.3.2.

In a Tikhonov regularization, Φ is a finite-difference approximation of a first-order derivative $\phi[n] = \delta[n] - \delta[n-1]$, so $|\hat{\phi}[k]| = 2|\sin(\pi k/N)|$. The resulting estimator attenuates the signal high frequencies and thus restores a regular estimation. If f is not uniformly regular, then this regularization produces a large error. Section 13.2 shows that such estimators can be improved by nonlinear estimators that take advantage of a sparse representation in a different basis, such as a wavelet basis.

13.2 THRESHOLDING ESTIMATORS FOR INVERSE PROBLEMS

Linear inverse estimators can be factorized into a pseudo inverse followed by a linear denoising operator that attenuates the amplified noise. Replacing the linear denoising operator by a nonlinear thresholding estimator in an appropriate orthogonal basis can improve the estimation. Following the work of Donoho [214], conditions are given to obtain a nearly minimax risk. Section 13.2.2 studies applications to signal and image deconvolutions with wavelet and wavelet packet bases.

13.2.1 Thresholding in Bases of Almost Singular Vectors

Suppose that there exists an orthonormal basis $\mathcal{B} = \{g_m\}_{0 \leq m < N}$ in which f has a sparse representation. Its decomposition coefficients in $\mathcal{B}$,

$$f = \sum_{m=0}^{N-1} a[m] g_m,$$

are the decomposition coefficients of

$$Uf = \sum_{m=0}^{N-1} a[m]\, Ug_m$$

in $\{Ug_m\}_{0 \leqslant m < N}$. Let $U\mathcal{B} = \{Ug_m\}_{m \in \Gamma_Q}$ be the transformed basis of $Q \leqslant N$ vectors such that $Ug_m \neq 0$. If $U\mathcal{B}$ is a basis of **ImU**, then $a[m]$ for $m \in \Gamma_Q$ is calculated from the inner products of Uf with the vectors of a dual basis. An estimation of f from $Y = Uf + W$ is derived with a thresholding that reduces the noise. Since $a[m]$ is not estimated and thus set to zero if $Ug_m = 0$, the resulting estimator $\tilde{F}$ of f belongs to the space $(\mathbf{NullU})^{\perp}$ of dimension Q. It does not perform any super-resolution. Conditions are established on the basis $\mathcal{B}$ relative to the signal class and the operator U to obtain a nearly minimax estimator.

Thresholding Biorthogonal Bases

The transformed basis $U\mathcal{B} = \{Ug_m\}_{m \in \Gamma_Q}$ is supposed to be a basis of the finite-dimensional space **ImU**. Its dual basis is characterized by biorthogonality relations. Let us renormalize the transformed basis $\{\tilde{\lambda}_m^{-1} Ug_m\}_{m \in \Gamma_Q}$ so that the biorthogonal basis $\{\tilde{\phi}_m\}_{m \in \Gamma_Q}$ in **ImU** is normalized. This basis is characterized by the biorthogonality relations:

$$\forall (m,p) \in \Gamma_Q{}^2, \quad \langle \tilde{\phi}_p, \tilde{\lambda}_m^{-1} Ug_m \rangle = \delta[p-m], \tag{13.14}$$

and $\tilde{\lambda}_m$ is adjusted so that $\|\tilde{\phi}_m\| = 1$.

If $f = \sum_{m=0}^{N-1} a[m]\, g_m$, then $Uf = \sum_{m \in \Gamma_Q} \tilde{\lambda}_m a[m]\, (\tilde{\lambda}_m^{-1} Ug_m)$, so the coefficients $a[m]$ are obtained by decomposing Uf in the dual basis:

$$a[m] = \tilde{\lambda}_m^{-1} \langle Uf, \tilde{\phi}_m \rangle \quad \text{for} \quad m \in \Gamma_Q.$$

From $Y = Uf + W$, we get

$$\langle Y, \tilde{\lambda}_m^{-1} \tilde{\phi}_m \rangle = a[m] + \langle W, \tilde{\lambda}_m^{-1} \tilde{\phi}_m \rangle.$$

Since $\|\tilde{\phi}_m\| = 1$, if W is a Gaussian white noise of variance σ^2, then $\langle W, \tilde{\lambda}_m^{-1} \tilde{\phi}_m \rangle$ is a Gaussian random variable of variance $\sigma^2 \tilde{\lambda}_m^{-2}$. Donoho [214] proposed to estimate $a[m]$ and thus the projection of f in $(\mathbf{NullU})^{\perp}$ with a thresholding

$$\tilde{F} = \sum_{m \in \Gamma_Q} \rho_{T_m} \left(\langle Y, \tilde{\lambda}_m^{-1} \tilde{\phi}_m \rangle \right) g_m \in (\mathbf{NullU})^{\perp}, \tag{13.15}$$

where the thresholds are $T_m = \tilde{\lambda}_m^{-1}\, \sigma \sqrt{2 \log_e Q}$.

Diagonal Estimation with Amplified Noise

To better understand and compute the thresholding estimator (13.15), it is decomposed into a linear pseudo inverse followed by a thresholding denoising estimator. Let us write U_{Γ_Q} as the restriction of U to the space $(\mathbf{NullU})^{\perp}$ generated by the Q

vectors $\{g_m\}_{m\in\Gamma_Q}$. Like in (13.8), we apply the pseudo inverse U^+, which inverts the restriction U_{Γ_Q} of U to $(\mathbf{NullU})^\perp$,

$$X = U^+ Y = U^+ U f + U^+ W \in (\mathbf{NullU})^\perp,$$

where U^+U is the orthogonal projector on $(\mathbf{NullU})^\perp$. The inverted noise $Z = U^+W$ has a covariance on $(\mathbf{NullU})^\perp$ that is $K_Z = \sigma^2 U^+ U^{+*}$. Since U^+Uf has a sparse decomposition in the basis $\{g_m\}_{m\in\Gamma_Q}$ of $(\mathbf{NullU})^\perp$, it can be estimated by thresholding its coefficients in this basis, according to (11.66). Thresholds are proportional to the noise variance $\sigma_{\mathcal{B}}[m]^2 = E\{|\langle Z, g_m\rangle|^2\} = \langle g_m, K_Z g_m\rangle$.

Theorem 13.2 proves that the noise amplification of $\sigma_{\mathcal{B}}[m]$ relative to σ is specified by the normalization factors $\tilde{\lambda}_m$, and it derives a thresholding estimator of X.

Theorem 13.2. The renormalization factors satisfy

$$\tilde{\lambda}_m^{-2} = \langle (U_{\Gamma_Q}^* U_{\Gamma_Q})^{-1} g_m, g_m\rangle = \frac{\sigma_{\mathcal{B}}[m]^2}{\sigma^2}, \tag{13.16}$$

and the thresholding estimator

$$\tilde{F} = \sum_{m=0}^{Q-1} \rho_{T_m}(\langle X, g_m\rangle)\, g_m \in (\mathbf{NullU})^\perp \tag{13.17}$$

with $T_m = \sqrt{2\log_e Q}\,\sigma_{\mathcal{B}}[m]$ is equal to the thresholding inverse estimator (13.15).

Proof. The symmetric operator $U_{\Gamma_Q}^* U_{\Gamma_Q}$ is invertible over $(\mathbf{NullU})^\perp$. We derive from the biorthogonality relations that $U_{\Gamma_Q}^* \tilde{\phi}_m = \tilde{\lambda}_m g_m$ for $m\in\Gamma_Q$ and thus that $\tilde{\phi}_m = \tilde{\lambda}_m U_{\Gamma_Q}(U_{\Gamma_Q}^* U_{\Gamma_Q})^{-1} g_m$. Since $\|\tilde{\phi}_m\| = 1$, it results that $\tilde{\lambda}_m^{-1} = \|U_{\Gamma_Q}(U_{\Gamma_Q}^* U_{\Gamma_Q})^{-1} g_m\|$, so

$$\tilde{\lambda}_m^{-2} = \langle U_{\Gamma_Q}(U_{\Gamma_Q}^* U_{\Gamma_Q})^{-1} g_m,\, U_{\Gamma_Q}(U_{\Gamma_Q}^* U_{\Gamma_Q})^{-1} g_m\rangle = \langle (U_{\Gamma_Q}^* U_{\Gamma_Q})^{-1} g_m, g_m\rangle.$$

If $g_m \in (\mathbf{NullU})^\perp$, then $K_Z g_m = \sigma^2 (U_{\Gamma_Q}^* U_{\Gamma_Q})^{-1} g_m$, so $\sigma^2 \tilde{\lambda}_m^{-2} = \langle K_Z g_m, g_m\rangle = \sigma_{\mathcal{B}}[m]^2$, which finishes the proof of (13.16).

As a consequence,

$$\langle Y, \tilde{\lambda}_m^{-1}\tilde{\phi}_m\rangle = \langle (U_{\Gamma_Q}^* U_{\Gamma_Q})^{-1} U_{\Gamma_Q}^* Y, g_m\rangle = \langle X, g_m\rangle,$$

so (13.15) and (13.17) threshold the same coefficients. Since $\sigma\tilde{\lambda}_m^{-1} = \sigma_{\mathcal{B}}[m]$, the thresholds in (13.17) and (13.15) are also identical. Thus, both estimators are equal. ■

This theorem proves that the thresholding estimator (13.15) can be interpreted as a linear pseudo inverse of U followed by a thresholding estimator that reduces the amplified noise. We know from Chapter 11 that such estimators are highly efficient if f is sparse in $\mathcal{B}$ and if the amplified noise Z has nearly independent coefficients in the basis $\mathcal{B}$. It implies that $\mathcal{B}$ is a basis of "almost singular vectors" that "nearly diagonalizes" the covariance K_Z and thus U^*U.

Implementing the thresholding estimator with (13.17) rather than (13.15) may require less operations if signal coefficients in $\mathcal{B}$ are computed with a fast algorithm and if the pseudo inverse U^+ is also implemented with a fast algorithm. This is the case for deconvolution estimators in wavelet packet bases.

It is proved in Theorem 11.8 that the risk produced by a thresholding estimator is of the same order as the oracle risk (11.65) obtained by setting to zero all coefficients $\langle X, g_m \rangle$ for which $|\langle f, g_m \rangle| \leqslant \sigma_{\mathcal{B}}[m] = \tilde{\lambda}_m^{-1} \sigma$. If over a signal class Θ we have $\sigma \tilde{\lambda}_m^{-1} \geqslant \sup_{f \in \Theta} |\langle f, g_m \rangle|$, then the oracle systematically sets $\langle X, g_m \rangle$ to zero because the amplified noise is too large relative to the signal. The estimation risk is thus reduced by doing the same. This is equivalent to reducing the set $\{g_m\}_{m \in \Gamma_Q}$ to a subset $\{g_m\}_{m \in \Gamma_{Q_0}}$ for which

$$\sigma \tilde{\lambda}_m^{-1} < \sup_{f \in \Theta} |\langle f, g_m \rangle|. \tag{13.18}$$

Suppressing the directions g_m corresponding to singular values $\tilde{\lambda}_m$ that are too small is important in numerical applications.

Nearly Minimax

It now remains to be understood under which conditions such a thresholding estimator is nearly optimal among all possible nonlinear estimators over a signal set Θ. Let $r_{\text{th}}(\Theta) = \sup_{f \in \Theta} E\{\|f - \tilde{f}\|^2\}$ be the maximum risk over Θ of thresholding estimators (13.15) and (13.17). Theorem 13.3 proves that $r_{\text{th}}(\Theta)$ is close to the nonlinear minimax risk $r_n(\Theta)$ if Θ is orthosymmetric in $\mathcal{B}$, and if the transformed basis $U\mathcal{B}$ satisfies a Riesz stability property, which implies that $\mathcal{B}$ "nearly diagonalizes" U^*U. Section 11.5.2 explains that Θ is orthosymmetric in $\mathcal{B}$ if any $f \in \Theta$ remains in Θ when reducing the amplitude of any of its decomposition coefficients in $\mathcal{B}$. Such a set is aligned with the vectors' directions in $\mathcal{B}$, which provides sparse signal approximations.

Theorem 13.3: *Donoho.* If Θ is orthosymmetric in $\mathcal{B}$ and there exists $B > 0$ such that

$$\forall a \in \mathbb{C}^Q, \quad \| \sum_{m \in \Gamma_Q} a[m] \tilde{\lambda}_m^{-1} U g_m \|^2 \leqslant B \|a\|^2$$

with

$$\tilde{\lambda}_m^{-2} = \langle (U_{\Gamma_Q}^* U_{\Gamma_Q})^{-1} g_m, g_m \rangle, \tag{13.19}$$

then for thresholds $T_m = \sigma \tilde{\lambda}_m^{-1} \sqrt{2 \log_{\text{e}} Q}$, the maximum thresholding risk $r_{\text{th}}(\Theta)$ satisfies for $N \geqslant 4$

$$r_n(\Theta) \leqslant r_{\text{th}}(\Theta) \leqslant (2 \log_{\text{e}} Q + 1) \left(\bar{\sigma}^2 + 1.25\, B\, r_n(\Theta) \right) \tag{13.20}$$

with $\bar{\sigma}^2 = Q^{-1} \sigma^2 \sum_{m \in \Gamma_Q} \tilde{\lambda}_m^{-2}$.

Proof. The main steps of the proof are given without detail. The thresholding risk $r_{\text{th}}(f) = E\{\|\tilde{f} - f\|^2\}$ is first compared to the minimum oracle risk $r_{\text{inf}}(f)$ of diagonal estimators.

Computing this oracle risk as in (11.31) gives

$$r_{\text{inf}}(f) = \sum_{m=0}^{Q-1} \frac{\sigma_m^2 \, |f_{\mathcal{B}}[m]|^2}{\sigma_m^2 + |f_{\mathcal{B}}[m]|^2} + \sum_{m=Q}^{N-1} |f_{\mathcal{B}}[m]|^2. \tag{13.21}$$

Since a thresholding estimator is a diagonal estimator, $r_{\text{th}}(f) \geqslant r_{\text{inf}}(f)$. The result (11.67) of Theorem 11.7 can be refined by replacing $r_{\text{pr}}(f)$ by $r_{\text{inf}}(f)$:

$$r_{\text{inf}}(f) \leqslant r_{\text{th}}(f) \leqslant (2 \log_{\text{e}} Q + 1) \left(\bar{\sigma}^2 + r_{\text{inf}}(f) \right) \quad \text{with} \quad \bar{\sigma}^2 = Q^{-1} \sum_{m \in \Gamma_Q} \sigma_{\mathcal{B}}^2[m], \tag{13.22}$$

and $\sigma_{\mathcal{B}}^2[m] = \langle K_Z g_m, g_m \rangle = \sigma^2 \tilde{\lambda}_m^{-2}$. If Θ is orthosymmetric in $\mathcal{B}$, then we prove that

$$r_{\text{inf}}(\Theta) \geqslant 1.25 \, B \, r_n(\Theta) \, . \tag{13.23}$$

The proof of this result considers first the particular case where $\mathbf{NullU} = \{0\}$ and $U^* U$ transforms $\mathcal{B}$ in an orthogonal basis. It implies that $U^* U$ is diagonal in $\mathcal{B}$ and thus that the covariance matrix K_Z of the inverted noise $Z = U^+ W$ is also diagonal in $\mathcal{B}$. As a result, the noise coefficients in $\mathcal{B}$ are independent. Since Θ is orthosymmetric in $\mathcal{B}$, renormalizing the noise gives a white noise, and Theorem 11.14 implies that diagonal estimators in $\mathcal{B}$ are nearly minimax, with

$$r_n(\Theta) \geqslant \frac{1}{1.25} \, r_{\text{inf}}(\Theta). \tag{13.24}$$

This result is then extended for a nondiagonal covariance K_Z. Let K_d be the diagonal matrix in $\mathcal{B}$ with a diagonal equal to the diagonal of K_Z. One can verify that

$$K_Z \geqslant B^{-1} K_d \quad \Longleftrightarrow \quad \forall f \in \mathbb{C}^N, \quad \langle K_Z f, f \rangle \geqslant B^{-1} \langle K_d f, f \rangle. \tag{13.25}$$

As a consequence of this inequality, a noise augmentation lemma proves that the minimax risk with a noise of covariance K_Z is necessarily larger than the minimax risk when the noise covariance is $B^{-1} K_d$. Using this result, (13.23) is derived from (13.24), which applies to $B^{-1} K_d$. When $\mathbf{NullU}$ is not empty, $U^* U$ is written as a limit of operators U_k with a null space $\mathbf{NullU}_k$ that is not empty and verifies (13.23) according to this proof. The result is then proved for U by taking the limit on k.

Inequalities (13.22) and (13.23) imply

$$\frac{B^{-1}}{1.25} \, r_{\text{inf}}(\Theta) \leqslant r_n(\Theta) \leqslant r_{\text{th}}(\Theta) \leqslant (2 \log_{\text{e}} Q + 1) \left(\bar{\sigma}^2 + r_{\text{inf}}(\Theta) \right), \tag{13.26}$$

which proves (13.20). ■

The constant B in (13.19) is the upper Riesz bound of the normalized basis $\{\tilde{\lambda}_m^{-1} U g_m\}_{m \in \Gamma_Q}$. Having a stable normalized basis is a requirement to stabilize the thresholding estimation. If B is not too large and if Θ is orthosymmetric, then this theorem proves that the maximum risk $r_{\text{th}}(\Theta)$ of a thresholding estimator is of the same order as the nonlinear minimax risk $r_n(\Theta)$. This result is further improved by reducing $\{g_m\}_{m \in \Gamma_Q}$ according to (13.18) into a subset $\{g_m\}_{m \in \Gamma_{Q_0}}$, which satisfies $\sigma \tilde{\lambda}_m^{-1} < \sup_{f \in \Theta} |\langle f, g_m \rangle|$. Indeed, one can verify [324] that it yields a tighter

inequality (13.20) where Q is replaced by Q_0 and $\bar{\sigma}^2$ by $\bar{\sigma}_0^2 = Q_0^{-1}\,\sigma^2\,\sum_{m\in\Gamma_{Q_0}} \tilde{\lambda}_m^{-2}$. This sum does not include the smallest $\tilde{\lambda}_m$ and is thus potentially much smaller.

When Θ is not orthosymmetric but can be embedded in two close orthosymmetric sets, then applying this theorem to each orthosymmetric set gives a similar result. In this case, since the thresholding estimator performs no super-resolution and is nearly minimax, it also implies that there is no super-resolution estimator that provides a significant improvement for all signals in Θ. Indeed, the basis $\mathcal{B}$ is optimal to represent f, and it includes $N-Q$ vectors that are completely cancelled by the operator U, along which the signal coefficients cannot be recovered.

Almost Singular Vectors with Narrow Spectrum

Section 5.1.2 proves that for a normalized Riesz basis the upper Riesz bound satisfies $B \geqslant 1$, and if $B = 1$, then the basis is orthonormal. In this case, g_m is an eigenvector of U^*U with an eigenvalue $\tilde{\lambda}_m^2 = \|U\,g_m\|^2$. This is generally not the case, but to get a small constant B, the basis $\mathcal{B}$ must nearly diagonalize U^*U, and the normalization constants $\tilde{\lambda}_m$ are then approximately equal to singular values.

Let $\mathcal{B}_s = \{e_k\}_{k\in\Gamma}$ be a singular basis that diagonalizes U^*U with singular values $\{\lambda_k^2\}_{k\in\Gamma}$. Each $g_m \in \mathcal{B}$ is a mix of singular values λ_k^2, which define its singular spectrum support. Theorem 13.4 relates B to the relative variations of the singular spectrum for vectors in $\mathcal{B}$.

Theorem 13.4. Let $\mathbf{C}^N = \oplus_{l=1}^L \mathbf{U}_l$ be a partition in orthogonal spaces $\mathbf{U}_l$ generated by families of singular vectors $\{e_k\}_{k\in\Gamma_l}$ with singular values $\{\lambda_k^2\}_{k\in\Gamma_l}$. Let $\mathcal{B} = \{g_m\}_{m\in\Gamma}$ be an orthonormal basis of $\mathbb{C}^N$ obtained as a union of orthonormal bases of each $\mathbf{U}_l$. If the constant B satisfies

$$B\,(\min_{k\in\Gamma_l} \lambda_k^2) \geqslant \max_{k\in\Gamma_l} \lambda_k^2 \quad \text{for} \quad 0 \leqslant l < L, \tag{13.27}$$

then

$$\forall a \in \mathbb{C}^Q, \quad \|\sum_{m\in\Gamma_Q} a[p]\,\tilde{\lambda}_m^{-1}\,U\,g_m\|^2 \leqslant B\,\|a\|^2$$

with

$$\tilde{\lambda}_m^{-2} = \langle (U_{\Gamma_Q}^* U_{\Gamma_Q})^{-1} g_m, g_m \rangle. \tag{13.28}$$

Proof. Let $\lambda_{l,\min} = \min_{k\in\Gamma_l} \lambda_k$ and $\lambda_{l,\max} = \max_{k\in\Gamma_l} \lambda_k$. We write U_{Γ_l} as the restriction of U to the space $\mathbf{U}_l$. Since each space $\mathbf{U}_l$ is generated by $\{e_k\}_{k\in\Gamma_l}$, it results that

$$\lambda_{l,\min}\,\mathrm{Id} \leqslant U_{\Gamma_l}^*\,U_{\Gamma_l} \leqslant \lambda_{l,\max}\,\mathrm{Id}.$$

For $g_m \in \mathbf{U}_l$, we thus have $\lambda_{l,\min} \leqslant \tilde{\lambda}_m \leqslant \lambda_{l,\max}$. Moreover,

$$\begin{aligned}\|\sum_{m\in\Gamma_Q} a[p]\,\tilde{\lambda}_m^{-1}\,U g_m\|^2 &= \sum_{l=1}^{L} \|\sum_{g_m\in\mathbf{U}_l} a[p]\,\tilde{\lambda}_m^{-1}\,U g_m\|^2 \leqslant \sum_{l=1}^{L} \lambda_{l,\min}^{-2}\,\|\sum_{g_m\in\mathbf{U}_l} a[p]\,U g_m\|^2 \\ &\leqslant \sum_{l=1}^{L} \lambda_{l,\min}^{-2}\,\lambda_{l,\max}^2\,\|\sum_{g_m\in\mathbf{U}_l} a[p]\,g_m\|^2 \leqslant B\,\|a\|^2,\end{aligned}$$

which proves (13.28). ■

The constant B in (13.27) is the maximum relative variation of singular values in the spaces $\mathbf{U}_l$ where the bases vectors g_m belong. In this case, each g_m has a decomposition over singular vectors with singular values that have relative variations bounded by B. If the vectors g_m have a *narrow singular spectrum*, then B gets close to 1.

The condition (13.27) is not strictly necessary and can be relaxed by just imposing that most of the energy of g_m is concentrated over singular values that have relative variation bounded by B. For deconvolutions, wavelets and wavelet packets are examples of bases providing sparse signal representations with a narrow singular spectrum in that sense. Other inverse problems have been more recently investigated with this approach [326]. When signals in Θ have a sparse representation in a basis with vectors that have a *spread spectrum*, then thresholding estimators are not optimal, but Section 13.3 shows that there may be an opportunity for a super-resolution estimation of f.

13.2.2 Thresholding Deconvolutions

The deconvolution estimation of f from $Y = Uf + W$ with $Uf[n] = f \circledast u[n]$ is studied in Section 13.1 with linear operators that are diagonal in the discrete Fourier basis

$$\mathcal{B}_S = \{ g_m[n] = N^{-1/2} e^{i2\pi mn/N} \}_{0 \leq m < N}.$$

The operator U^*U is a convolution diagonalized in this Fourier basis, and its transfer function is equal to the singular values $\lambda_k^2 = |\hat{u}[k]|^2$. Signals including singularities are not well approximated in a Fourier basis, and the resulting linear estimators produce a large risk. To reduce this risk with a thresholding estimator, one must find a basis $\mathcal{B}$ providing a sparse signal representation with vectors having a narrow spectrum. Theorem 13.4 shows that each vector $g_m \in \mathcal{B}$ must have a Fourier transform $\hat{g}_m[k]$ that as energy is concentrated over frequencies k for which $\lambda_k^2 = |\hat{u}[k]|^2$ has small relative variations. We consider two types of deconvolution problems where such bases can be constructed with wavelets or wavelet packets.

Homogeneous Deconvolutions with Wavelets

Derivative and integral operators are examples of convolution operators with transfer functions that vanish at the zero frequency or at infinity, with a homogeneous decay. After discretization, a homogeneous convolution operator $Uf[n] = f \circledast u[n]$ has by definition a transfer function that satisfies $|\hat{u}[k]| \sim |k|^p$. A first-order derivative $u[n] = \delta[n+1] - \delta[n-1]$ is homogeneous with $p = 1$: $|\hat{u}[k]| \sim |k|$. A derivative of order p yields $|\hat{u}[k]| \sim |k|^p$. Their inverse is singular at $k = 0$ and **NullU** is thus reduced to constant signals. Integrations are homogeneous convolutions with $p < 0$ and their inverse becomes singular at high frequencies.

Wavelet bases provide sparse representations of piecewise regular signals. Harmonic analysis results [44] also prove that singular homogeneous operators are "nearly diagonalized" in a wavelet basis. Indeed, a wavelet $\psi_{j,m}$ has a Fourier transform $\hat{\psi}_{j,m}[k]$ mostly concentrated on a dyadic frequency interval $k \in [2^{-j-1}, 2^{-j}]$.

Over such an interval, $|\hat{u}[k]| \sim |k|^p$ varies by a factor of the order of 2^p that does not depend on the scale. Suppose that the discrete wavelets have $q > p$ vanishing moments and correspond to the discretization of a regular wavelet $\psi(t)$ that is $\mathbf{C}^q$. For $L = -\log_2 N$, one can then verify that a periodic orthonormal wavelet family

$$\{\psi_{j,m}\}_{L<j\leqslant 0\,,\,0\leqslant m<2^{-j}} \tag{13.29}$$

is transformed into a Riesz basis by a homogeneous convolution operator U with an upper Riesz bound $B \sim 2^{2p}$ after renormalization. The constant scaling signal $\phi_0[n] = N^{-1/2}$ is not included because it is in **Null**U.

The transformed wavelets $U\psi_{j,m}$ are similar to wavelets and are called *vaguelettes* by Donoho [214]. The asymptotic minimax optimality of wavelet thresholding estimators is proved by Theorem 13.3 for homogeneous deconvolutions of signals that have a sparse signal representation in a wavelet basis. This includes bounded variation signals and images.

Mirror Wavelets Deconvolution

Analog-acquisition devices often remove high frequencies with a low-pass filter that vanishes at some maximum frequency. The sampling rate discretization is adjusted to this maximum frequency to avoid aliasing. If the low-pass transfer function has a smooth decay in the neighborhood of the maximum frequency, then the discretized signal is blurred. Optical systems often produce such a blur.

The maximum analog frequency is mapped by the discretization to the highest discrete frequency $2\pi k/N = \pm\pi/2$. The discrete signal blurring can thus be written as a discrete low-pass filter $Uf[n] = f \circledast u[n]$ with a transfer function $\hat{u}[k]$ that has a zero of order $p \geqslant 1$ at the maximum frequency index $k = N/2$:

$$|\hat{u}[k]| \sim |k - N/2|^p. \tag{13.30}$$

It results that **Null**U corresponds to signals $h[n]$ such that $\hat{h}[k] \neq 0$ only for $k = \pm N/2$, and thus $h[n] = c\,(-1)^n$. Since wavelet bases provide sparse representation of piecewise regular signals, they could be a good candidate to implement a thresholding deconvolution estimator. This requires to nearly diagonalize U^*U, and thus that the singular spectrum $|\hat{u}[k]|^2$ and its inverse $|\hat{u}[k]|^{-2}$ have small relative variations over the support of the Fourier transform of each basis vector.

At scales $2^j > 2N^{-1}$, $\hat{\psi}_{j,m}[k]$ has a frequency support nearly included in an interval $[2^{-j-1}, 2^{-j}]$ where $|\hat{u}[k]|^{-2}$ remains nearly constant. However, at the finest scale $2^{L+1} = 2N^{-1}$, wavelets $|\hat{\psi}_{L+1,m}[k]|$ have a spread spectrum because their energy is mainly concentrated in the higher-frequency band $[N/4, N/2]$, where $|\hat{u}[k]|$ varies by a huge factor on the order of N^{2p}. These fine-scale wavelets must therefore be replaced by wavelet packet vectors having a smaller-frequency support adjusted to the rapid relative variation of $|\hat{u}[k]|$.

To efficiently approximate piecewise regular signals, these wavelet packets must have the smallest possible spatial support, and thus the largest possible

frequency support. The optimal trade-off is obtained with wavelet packets that we denote $\tilde{\psi}_{j,m}$ that have a discrete Fourier transform $\widehat{\tilde{\psi}}_{j,m}[k]$ mostly concentrated in $[N/2 - 2^{-j}, N/2 - 2^{-j-1}]$, as illustrated by Figure 13.1. Over such intervals, $|\hat{u}[k]|^{-2}$ varies by a relative factor of 2^{2p} that does not depend on the scale 2^j. These particular wavelet packets, introduced by Kalifa and Mallat [323, 324], are called *mirror wavelets* because they are related to discrete wavelets by

$$|\widehat{\tilde{\psi}}_{j,m}[k]| = |\hat{\psi}_{j,m}[N/2 - k]| \quad \text{and} \quad \tilde{\psi}_{j,m}[n] = (-1)^{n-1}\psi_{j,m}[1-n].$$

A mirror wavelet basis is a wavelet packet basis composed of wavelets $\psi_{j,m}$ at scales $2^j > 2^{L+1} = 2N^{-1}$ and of mirror wavelets to replace wavelets at the finest scale 2^{L+1}:

$$\left\{\psi_{j,m}, \tilde{\psi}_{j,m}\right\}_{0 \leqslant m < 2^{-j}, L+1 < j \leqslant 0}.$$

The highest-frequency mirror wavelet $\tilde{\psi}_{0,0}[n] = N^{-1/2}(-1)^n$ belongs to **NullU** and is therefore not included in the estimation. The fast mirror wavelet transform studied in Exercise 8.10 is implemented with a wavelet packet filter bank, as described in Section 8.1.4.

If these wavelets and wavelet packets are constructed with conjugate mirror filters that define a continuous time wavelet $\psi(t)$ that is $\mathbf{C}^q$ with $q > p$ vanishing moments, then Kalifa and Mallat [323] prove that a thresholding estimator in a mirror wavelet basis yields a quasi-minimax deconvolution estimator for bounded variation signals. The resulting risk is then much smaller than the risk obtained by a linear singular value decomposition estimator.

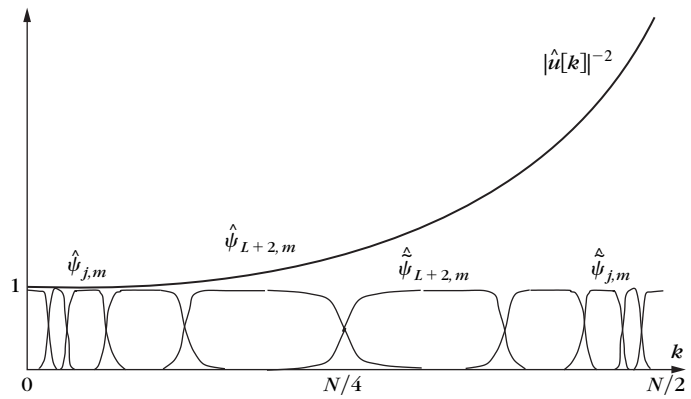


FIGURE 13.1

The singular spectrum $|\hat{u}[k]|^2$ of a low-pass filter decreases to zero at high frequencies. The support of mirror wavelets is reduced at high frequencies so that the relative variation of $|\hat{u}[k]|^{-2}$ remains uniformly bounded over their support.

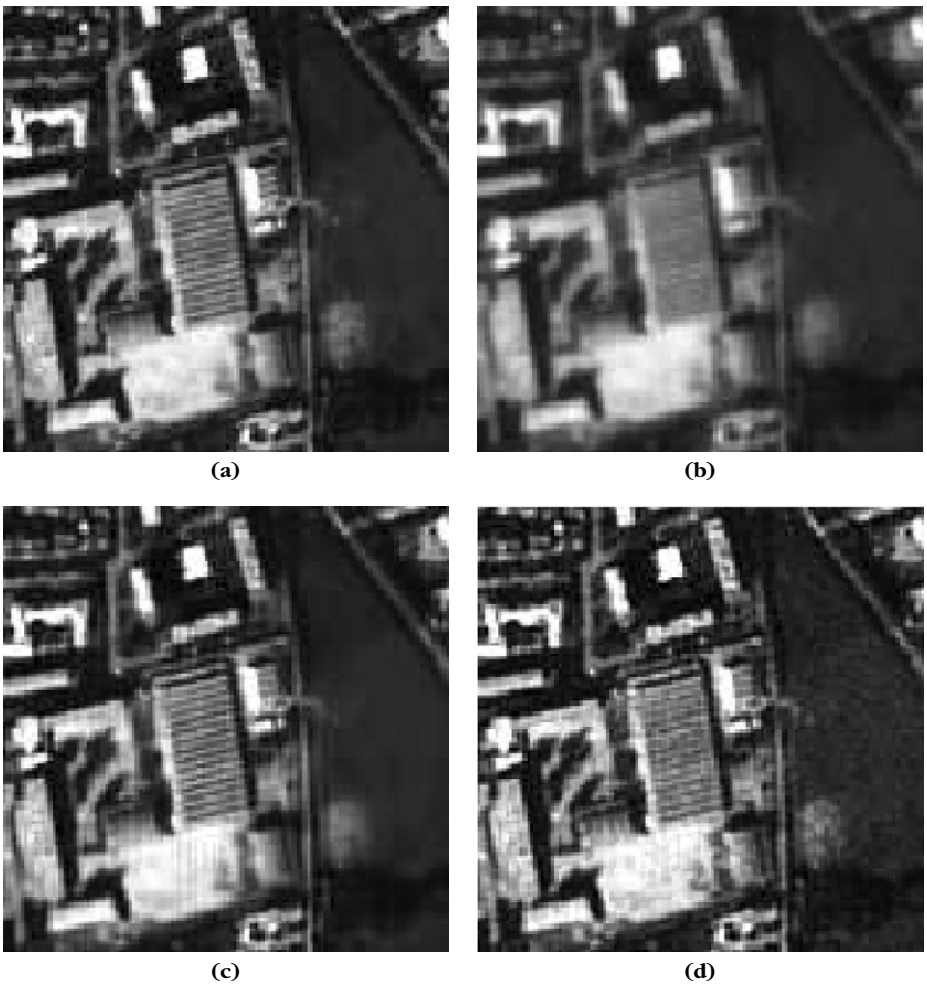


FIGURE 13.2

(a) Original airplane image. **(b)** Simulation of a satellite image provided by CNES (SNR = 31.1 db). **(c)** Deconvolution with a translation-invariant thresholding in a mirror wavelet basis (SNR = 34.1 db). **(d)** Deconvolution calculated with a circular convolution, which yields a nearly minimax risk for bounded variation images (SNR = 32.7 db).

Deconvolution of Images

For separable low-pass filters that vanish at the highest frequencies, nearly optimal deconvolution of bounded variation images is calculated with a separable extension of the deconvolution estimator in a mirror wavelet basis. Such restoration algorithms are used in wavelet packet and mirror wavelet bases [324, 417, 418] for deblurring satellite images. The exposition time of the satellite photoreceptors cannot be

reduced too much because the light intensity reaching the satellite is small and must not be dominated by electronic noises. The satellite movement thus produces a blur, which is aggravated by the imperfection of the optics. The electronics of the photoreceptors add a Gaussian white noise. The Figure 13.2(b), provided by the French spatial agency CNES, is a simulated satellite image calculated from the airplane image shown in Figure 13.2(a).

Figure 13.2(c) shows an example of deconvolution calculated in the mirror wavelet basis. The thresholding is performed with a translation-invariant algorithm. This can be compared with the linear estimation in Figure 13.2(d), calculated with a circular convolution estimator that has a maximum risk over bounded variation images close to the minimax linear risk. The linear deconvolution sharpens the image but leaves a visible noise in the regular parts of the image. The thresholding algorithm mostly removes the noise in these regions while improving the restoration of edges and oscillatory parts. Algorithms alternating between linear estimations and thresholding estimations in a wavelet basis can also provide efficient deconvolutions of such images [385].

13.3 SUPER-RESOLUTION

Numerically increasing the resolution of measured data has major industrial applications when data acquisition is difficult or costly. In geophysics, the highest possible resolution must be recovered from relatively few measurements obtained by sending waves underground and measuring reflections with sensors distributed on the sea or on the ground. In X-ray imaging, the radiation time of a patient and thus the data acquisition are also limited. For Earth observation, improving resolution usually means sending a new satellite, which is not a light project. On the consumer front, the resolution of videos in standard-definition television formats (PAL or NTSC) must be numerically increased to match the larger resolution of high-definition flat panel televisions. Many more examples can be found.

The linear and thresholding estimators in Sections 13.1 and 13.2 estimate the projection of $f \in \mathbb{C}^N$ in a space of dimension $Q_0 \leqslant Q = \dim(\mathbf{ImU}) \leqslant N$, which provides no super-resolution. Given $Q < N$ independent measurements, super-resolution aims at estimating the projection of f in a space of dimension larger than Q, and if possible of dimension N.

13.3.1 Sparse Super-resolution Estimation

Similar to Section 13.2, we suppose that f has a sparse approximation in a dictionary $\mathcal{D} = \{g_p\}_{p\in\Gamma}$ of size $|\Gamma| = P > Q$, but with $Ug_p \neq 0$ for all $p \in \Gamma$. This dictionary may not generate whole space $\mathbb{C}^N$. The projection of f is estimated from $Y = Uf + W$ over the space generated by this dictionary, which has a dimension larger than Q and is adjusted depending on super-resolution capabilities.

Sparsity means that f is precisely approximated by its orthogonal projection f_Λ over a subspace generated by a small number $|\Lambda|$ of vectors $\{g_p\}_{p\in\Lambda}$ chosen in $\mathcal{D}$:

$$f_\Lambda = \sum_{p\in\Lambda} a[p]\, g_p. \tag{13.31}$$

The error $w_\Lambda = f - f_\Lambda$ thus has a small norm. The approximation vectors $\{g_p\}_{p\in\Lambda}$ are not restricted a priori to a space of dimension Q. Since $Y = Uf + W$, it results from (13.31) that

$$Y = \sum_{p\in\Lambda} a[p]\, Ug_p + W' \quad \text{with} \quad W' = Uw_\Lambda + W. \tag{13.32}$$

The coefficients $a[p]$ can be estimated with a sparse denoising estimation of Y in the transformed and normalized dictionary

$$\mathcal{D}_U = \left\{ \frac{Ug_p}{\|Ug_p\|} \right\}_{p\in\Gamma}. \tag{13.33}$$

A major difficulty is that $\mathcal{D}_U$ is a redundant dictionary with $P > Q$ vectors that are in the operator image space **ImU** of dimension Q.

Let $\tilde{Y}$ be a sparse approximation of Y computed with an algorithm that selects a subset of dictionary vectors $\{Ug_p/\|Ug_p\|\}_{p\in\tilde{\Lambda}}$:

$$\tilde{Y} = \sum_{p\in\tilde{\Lambda}} \tilde{a}[p]\, \frac{Ug_p}{\|Ug_p\|}. \tag{13.34}$$

An $\mathbf{l}^1$ Lagrangian pursuit (12.89) computes such an approximation with:

$$\tilde{a} = \underset{a\in\mathbb{C}^P}{\operatorname{argmin}} \frac{1}{2}\|Y - \sum_{p\in\Gamma} a[p]\, \frac{Ug_p}{\|Ug_p\|}\|^2 + T\|a\|_1. \tag{13.35}$$

This minimization can be solved by the iterative thresholding algorithm in Theorem 12.9. A matching pursuit or an orthogonal matching pursuit can also compute the coefficients $\tilde{a}$ of a sparse approximation $\tilde{Y}$ in $\mathcal{D}_U$. An estimation $\tilde{F}$ of f is derived by inverting U on the decomposition (13.34) of $\tilde{Y}$:

$$\tilde{F} = \sum_{p\in\tilde{\Lambda}} \frac{\tilde{a}[p]}{\|Ug_p\|}\, g_p. \tag{13.36}$$

If $\mathcal{D}$ is an orthonormal basis that diagonalizes the operator U^*U, then one can verify that an $\mathbf{l}^1$ Lagrangian pursuit computes an estimator $\tilde{F}$ that is identical to the soft-thresholding inverse estimator (13.15), but it provides no super-resolution.

Several conditions are necessary to recover a super-resolution estimation of f:

- *Stability*. $\{Ug_p/\|Ug_p\|\}_{p\in\Lambda}$ must be a Riesz basis.
- *Support recovery*. Decomposing Y in $\mathcal{D}_U$ must recover a support $\tilde{\Lambda}$ that closely approximates the approximation support Λ of f.
- *Spread singular spectrum*. Vectors in $\mathcal{D}$ must mix singular values of different amplitudes so that each $\|Ug_p\|$ is not too small.

Stability with Incoherence and Sparsity

Suppose that the approximation support Λ is given by some oracle. When there is no noise $W = 0$, the decomposition coefficients $a[p]$ of f_Λ in $\{g_p\}_{p\in\Lambda}$ can be recovered from the decomposition of $Y = Uf$ in $\{Ug_p/\|Ug_p\|\}_{p\in\Lambda}$ only if this family is linearly independent. If this is the case, in presence of noise, a stable computation also requires a nonzero lower Riesz bound $A_\Lambda > 0$:

$$\forall a \in \mathbb{C}^{|\Lambda|}, \quad A_\Lambda \|a\|^2 \leq \| \sum_{p\in\Lambda} a[p] \frac{Ug_p}{\|Ug_p\|} \|^2. \tag{13.37}$$

This is a nontrivial condition because $\mathcal{D}_U$ is a redundant dictionary of $P > Q$ vectors in a space of dimension Q, and the support Λ may a priori recombine any subset of vectors in $\mathcal{D}_U$. A vector ϕ_p for $p \in \Lambda$ should not be closely approximated by a linear combination of few other vectors in Λ, which is an incoherence property. Section 12.5.1 shows that the condition $A_\Lambda > 0$ is less difficult to obtain if the number of vectors $|\Lambda|$ is small relative to Q, and thus if the signal approximation is sparse.

Support Recovery

The computed support $\tilde{\Lambda}$ must provide a good estimation of Λ in the sense that the projection $f_{\tilde{\Lambda}}$ of f in the space generated by $\{g_p\}_{p\in\tilde{\Lambda}}$ should have an error $\|f - f_{\tilde{\Lambda}}\|$ comparable to $\|f - f_\Lambda\|$. If $\mathcal{D}$ is an orthonormal basis, it implies recovering a subset of the support of Λ, which carries the coefficients $\langle f, g_p \rangle$ of large amplitude. Decomposing Y in $\mathcal{D}_U$ may not recover this support because Y does not have a unique decomposition in this redundant dictionary. If there is no noise and $f = f_\Lambda$ then $Y = Uf_\Lambda$ belongs to the space $\mathbf{V}_\Lambda$ generated by $\{Ug_p\}_{p\in\Lambda}$. Let Λ^c be the complement of Λ. The *exact recovery criteria* (ERC) by Tropp [461] imposes that

$$\mathrm{ERC}(\Lambda) = \sup_{h\in\mathbf{V}_\Lambda} \frac{\max_{q\in\Lambda^c} |\langle h, Ug_q \rangle| / \|Ug_q\|}{\max_{p\in\Lambda} |\langle h, Ug_p \rangle| / \|Ug_p\|} < 1.$$

Theorems 12.11 and 12.15 then prove that the support Λ of f is recovered by decomposing Y with a matching pursuit or an $\mathbf{l}^1$ basis pursuit in $\mathcal{D}$. Theorem 12.12 also proves that the condition $\mathrm{ERC}(\Lambda) < 1$ requires that vectors in Λ have a small mutual correlation and a small correlation with any vector in the complement Λ^c. This is again an incoherence property over dictionary vectors. In presence of noise, the Riesz stability (13.37) is crucial to partially recover this support with an orthogonal matching pursuit or an $\mathbf{l}^1$ Lagrangian pursuit, as proved by Theorems 12.13 and 12.15.

Spread Singular Spectrum

The coefficients $\tilde{a}$ calculated from $Y = Uf + W$ carry the projection of the noise over the space generated by $\{Ug_p/\|Ug_p\|\}_{p\in\tilde{\Lambda}}$. This noise is amplified by the normalization factors $1/\|Ug_p\|$ in $\tilde{F}$, which should not be too large. Recovering a signal coordinate in the direction of $g_p \in \mathcal{D}$ must not amplify the noise above the maximum

signal coefficient that can be recovered. If W is a white noise of variance σ^2, it implies that

$$\forall p \in \Gamma, \quad \sigma \|U g_p\|^{-1} < \sup_{f \in \Theta} |\langle f, g_p \rangle|, \tag{13.38}$$

where Θ is the set of all possible signals. The same condition appears in (13.18) for thresholding inverse estimators.

Let $\mathcal{B}_S = \{e_k\}_{0 \leq k < N}$ be a singular vector basis that diagonalizes $U^* U$ with singular values $\{\lambda_k^2\}_{0 \leq k < N}$. Since

$$\|U g_p\|^2 = \langle U^* U g_p, g_p \rangle = \sum_{k=0}^{N-1} \lambda_k^2 \, |\langle g_p, e_k \rangle|^2,$$

to guarantee that $\|U g_p\|$ is not too small, g_p must have part of its energy spread over singular vectors $\{e_k\}_k$ having relatively large singular values $\{\lambda_k^2\}_k$. However, to recover a super-resolution estimation of f in **NullU** or in directions e_k where λ_k^2 is small, the vectors g_p must also be spread over these directions. Each g_p should thus have a spread spectrum that mixes small and large singular spectrum values. Ideally, $|\langle g_p, e_k \rangle| = N^{-1/2}$ and $\|U g_p\|^2 = N^{-1} \sum_{k=0}^{N-1} \lambda_k^2$. This condition is opposite to the narrow spectrum condition in Theorem 13.4 for thresholding estimators, which cannot perform any super-resolution.

If all vectors g_p have a fully spread spectrum, then $\|U g_p\|$ is approximately constant for all $p \in \Gamma$. The normalization then has a more marginal impact and Y can be decomposed in a nonnormalized transformed dictionary $\mathcal{D}_U = \{U g_p\}_{p \in \Gamma}$. The normalized $\mathbf{l}^1$ Lagrangian minimization (13.35) and (13.36) are then replaced by

$$\tilde{F} = \sum_{p \in \tilde{\Lambda}} \tilde{a}[p] \, g_p \quad \text{with} \quad \tilde{a} = \operatorname*{argmin}_{a \in \mathbb{C}^P} \frac{1}{2} \|Y - \sum_{p \in \Lambda} a[p] U g_p\|^2 + T \|a\|_1, \tag{13.39}$$

which can simplify computations.

Suppose that some vectors $g_p \in \mathcal{D}$ do not have a sufficiently spread spectrum and do not even satisfy the maximum noise amplification condition (13.38). If these vectors are not removed from the dictionary, the transformed dictionary should not be normalized to avoid numerical instabilities. This is equivalent to solve the nonnormalized $\mathbf{l}^1$ Lagrangian minimization (13.39). Directions g_p for which $\|U g_p\|$ is small are then barely recovered, which is indirectly equivalent to removing these vectors from the dictionary, but the lack of normalization penalizes the recovery of other directions.

Super-Resolution Recovery

Let us consider a normalized sparse super-resolution estimation calculated with an $\mathbf{l}^1$ Lagrangian pursuit:

$$\tilde{F} = \sum_{p \in \tilde{\Lambda}} \tilde{a}[p] \, \frac{g_p}{\|U g_p\|} \quad \text{with} \quad \tilde{a} = \operatorname*{argmin}_{a \in \mathbb{C}^P} \frac{1}{2} \|Y - \Phi_U^* a\|^2 + T \|a\|_1, \tag{13.40}$$

where $\Phi_U^* a = \sum_{p\in\Gamma} a[p]\, Ug_p/\|Ug_p\|$. Theorem 13.5 computes a conservative upper bound of the estimation error by setting a threshold T large enough so that the support $\tilde{\Lambda}$ of $\tilde{a}$ satisfies $\tilde{\Lambda}\subset\Lambda$ with a high probability. The theorem assumes that the approximation family $\{g_p\}_{p\in\Lambda}$ of f is a Riesz basis of the space it generates, and we write $\bar{B}_\Lambda$, the upper Riesz bound. The theorem also assumes that the transformed vectors $\{Ug_p/\|Ug_p\|\}_{p\in\Lambda}$ define a Riesz basis with lower Riesz bounds $A_\Lambda > 0$, and that the exact recovery condition $\mathrm{ERC}(\Lambda) < 1$ is satisfied.

Theorem 13.5. If $\mathrm{ERC}(\Lambda) < 1$ and

$$T = \lambda\, \frac{\|U\|_S\, \|f - f_\Lambda\| + \sigma\sqrt{2\log_{\mathrm{e}} P}}{1 - \mathrm{ERC}(\Lambda)} \quad \text{with} \quad \lambda > 1, \tag{13.41}$$

then there exists a unique $\mathbf{l}^1$ pursuit solution $\tilde{a}$ with a support that satisfies $\tilde{\Lambda}\subset\Lambda$ and the estimator $\tilde{F} = \sum_{p\in\tilde{\Lambda}} \tilde{a}[p]\, g_p/\|Ug_p\|$ has an error

$$\|\tilde{F} - f\|^2 \leqslant \|f - f_\Lambda\|^2 + \frac{\bar{B}_\Lambda\, (\lambda+2)^2\, |\Lambda|\, \left(\|U\|_S\, \|f - f_\Lambda\| + \sigma\sqrt{2\log_{\mathrm{e}} P}\right)^2}{(\min_{p\in\Lambda} \|Ug_p\|^2)\, A_\Lambda^2\, (1 - \mathrm{ERC}(\Lambda))^2}, \tag{13.42}$$

with a probability that tends to 1 as P increases.

Proof. The proof is derived from the proof of Theorem 12.15. To compute a solution with a support in Λ, we also consider a solution $\tilde{a}_\Lambda$ of the $\mathbf{l}^1$ Lagrangian minimization over Λ:

$$\tilde{a}_\Lambda = \underset{a_\Lambda\in\mathbb{C}^{|\Lambda|}}{\mathrm{argmin}} \frac{1}{2}\|Y - \Phi_{U\Lambda}^* a_\Lambda\|^2 + T\|a_\Lambda\|_1,$$

and $\tilde{a}$ is defined by $\tilde{a}[p] = \tilde{a}_\Lambda[p]$ for $p\in\Lambda$ and $\tilde{a}[p] = 0$ for $p\in\Lambda^c$. Let h be defined by

$$Th = \Phi_U(Y - \Phi_U^*\tilde{a}_\Lambda) = \Phi_U(Y - \Phi_{U\Lambda}^*\tilde{a}_\Lambda). \tag{13.43}$$

To prove that $\tilde{a}$ is a solution of the $\mathbf{l}^1$ Lagrangian pursuit (13.40), according to Theorem 12.8, we must verify that $\|h_\Lambda^c\|_\infty \leqslant 1$. Like in (12.150), we prove that

$$\|h_{\Lambda^c}\|_\infty \leqslant T^{-1} \max_{q\in\Lambda^c} |\langle \phi_q,\ Y - Y_\Lambda\rangle| + \mathrm{ERC}(\Lambda), \tag{13.44}$$

where $Y_\Lambda = P_{\mathbf{V}_\Lambda} Y$ is the orthogonal projection of Y on the space $\mathbf{V}_\Lambda$ generated by $\{Ug_p/\|Ug_p\|\}_{p\in\Lambda}$. Since $Y = Uf + W$,

$$\langle Y - Y_\Lambda, \phi_q\rangle = \langle Uf - P_{\mathbf{V}_\Lambda} Uf, \phi_q\rangle + \langle W - P_{V_\Lambda} W, \phi_q\rangle.$$

Since there are P vectors in the dictionary, and W is a white noise of variance σ^2,

$$\max_{q\in\Gamma} |\langle W - P_{V_\Lambda} W, \phi_q\rangle| \leqslant \sigma\sqrt{2\log_{\mathrm{e}} P},$$

with a probability that tends to 1 as P increases. It results that

$$\max_{q\in\Gamma} |\langle Y - Y_\Lambda, \phi_q\rangle| \leqslant \|Uf - P_{\mathbf{V}_\Lambda} Uf\| + \sigma\sqrt{2\log_{\mathrm{e}} P}. \tag{13.45}$$

Since $Uf_\Lambda \in \mathbf{V}_\Lambda$,

$$\|Uf - P_{\mathbf{V}_\Lambda} Uf\| \leq \|Uf - Uf_\Lambda\| \leq \|U\|_S \, \|f - f_\Lambda\|.$$

Equation (13.44) together with (13.45) implies that

$$\|h_{\Lambda^c}\|_\infty \leq \frac{\|U\|_S \, \|f - f_\Lambda\| + \sigma\sqrt{2\log_e P}}{T} + \mathrm{ERC}(\Lambda). \tag{13.46}$$

If

$$T = \frac{\lambda \, (\|U\|_S \|f - f_\Lambda\| + \sigma\sqrt{2\log_e P})}{1 - \mathrm{ERC}(\Lambda)} \text{ with } \lambda > 1, \tag{13.47}$$

then $\|h_{\Lambda^c}\|_\infty < 1$ and Theorem 12.8 prove that $\tilde{a}$ is indeed an $\mathbf{l}^1$ Lagrangian pursuit solution of (13.40). The same argument as in the proof of Theorem 13.5 shows that this solution is unique.

Since $\tilde{\Lambda} \subset \Lambda$, the error bound (13.42) can be computed from the coefficients restricted to Λ. Similar to (12.146), we prove that the coefficients $\tilde{a}_\Lambda$ of $\tilde{Y}$ satisfy

$$\tilde{a}_\Lambda = \Phi_{U\Lambda}^{*+} Y - T(\Phi_{U\Lambda}\Phi_{U\Lambda}^*)^{-1} h_\Lambda. \tag{13.48}$$

Writing $f_\Lambda = \sum_{p\in\Lambda} a_\Lambda[p] g_p / \|Ug_p\|$, we get

$$Y = \sum_{p\in\Lambda} a_\Lambda[p] \frac{Ug_p}{\|Ug_p\|} + U(f - f_\Lambda) + W.$$

It results that

$$\Phi_{U\Lambda}^{*+} Y = a_\Lambda + \Phi_{U\Lambda}^{*+} U(f - f_\Lambda) + \Phi_{U\Lambda}^{*+} W.$$

We derive from (13.48) that

$$\begin{aligned}\|\tilde{a}_\Lambda - a_\Lambda\| &\leq \|\Phi_{U\Lambda}^{*+} U(f - f_\Lambda)\| + \|\Phi_{U\Lambda}^{*+} W\| + \|T(\Phi_{U\Lambda}\Phi_{U\Lambda}^*)^{-1} h_\Lambda\| \\ &\leq \frac{1}{\sqrt{A_\Lambda}}(\|P_{\mathbf{V}_\Lambda} W\| + \|U\|_S \|f - f_\Lambda\|) + \frac{T\sqrt{|\Lambda|}}{A_\Lambda},\end{aligned}$$

where we used that $\|h_\Lambda\| \leq \sqrt{|\Lambda|}$. In a dictionary of size P, Lemma 12.1 proves that the energy of the noise projected in any space generated by $|\Lambda|$ dictionary vectors satisfies

$$\|W_\Lambda\| \leq 2\sigma\sqrt{2\,|\Lambda|\,\log_e P} \tag{13.49}$$

with a probability that tends to 1 as P increases. Inserting the value of T in (13.47) gives

$$\begin{aligned}\|\tilde{a}_\Lambda - a_\Lambda\| &\leq \frac{2\sigma\sqrt{|\Lambda|\,2\,\log_e P} + \|U\|_S \, \|f - f_\Lambda\|}{\sqrt{A_\Lambda}} \\ &\quad + \frac{\sqrt{|\Lambda|}\lambda(\|U\|_S \, \|f - f_\Lambda\| + \sigma\sqrt{2\log_e P})}{A_\Lambda(1 - \mathrm{ERC}(\Lambda))} \\ \|\tilde{a}_\Lambda - a_\Lambda\| &\leq \frac{(\lambda+2)\sqrt{|\Lambda|}\,(\|U\|_S \, \|f - f_\Lambda\| + \sigma\sqrt{2\log_e P})}{A_\Lambda\,(1 - \mathrm{ERC}(\Lambda))}.\end{aligned} \tag{13.50}$$

Since

$$\|\tilde{F}-f_\Lambda\| = \|\sum_{p\in\Lambda}(\tilde{a}_\Lambda[p]-a_\Lambda[p])\frac{g_p}{\|Ug_p\|}\| \leqslant \frac{\sqrt{\bar{B}_\Lambda}\,\|\tilde{a}_\Lambda - a_\Lambda\|}{\min_{p\in\Lambda}\|Ug_p\|},$$

it results that

$$\|\tilde{F}-f_\Lambda\| \leqslant \frac{\sqrt{\bar{B}_\Lambda}\,(\lambda+2)\sqrt{|\Lambda|}\,(\|U\|_S\,\|f-f_\Lambda\| + \sigma\sqrt{2\log_e P})}{\min_{p\in\Lambda}\|Ug_p\|\,A_\Lambda\,(1-\mathrm{ERC}(\Lambda))}. \tag{13.51}$$

Since $\tilde{F}-f_\Lambda \in \mathbf{V}_\Lambda$, we have $\|\tilde{F}-f\|^2 = \|\tilde{F}-f_\Lambda\|^2 + \|f-f_\Lambda\|^2$. Inserting (13.51) proves (13.42). ■

The result of this theorem is conservative but shows the main sources of instabilities of sparse super-resolution algorithms. It proves that part of the approximation support of Λ can be recovered by approximating the noisy data Y if $\mathrm{ERC}(\Lambda) < 1$. The multiplier T behaves as a soft threshold and must be above the noise, thus the term $\sigma\sqrt{2\log_e P}$. The normalization factors $\|Ug_p\|$ cannot be too small to avoid amplifying the noise.

13.3.2 Sparse Spike Deconvolution

Seismic sparse spike deconvolution is probably the first super-resolution algorithm used in industry for seismic exploration. This problem perfectly illustrates the main super-resolution ideas and difficulties. Mineral and oil seismic explorations measure underground reflectivity by sending pressure waves. The reflected pressure waves are recorded at the surface as a function of time and spatial position. Seismic inversion includes different steps such as migration and stacking to invert the wave propagation equation. After these inversions, at a given position of the surface, the resulting seismic data Y are approximately related to the underground reflectivity f through a convolution equation $Y[n] = u \star f[n] + W[n]$ where n is a time variable that is related to depth.

The convolution kernel u is called a *seismic wavelet* in geophysics, which is the origin of this name chosen by the geophysicist Morlet [276]. It depends on the pressure wave sent underground but also on the subsequent inversion operations. These seismic wavelets are calibrated from reflectivity and seismic data measured along wells that are drilled in the ground. The singular basis $\mathcal{B}_S$ that diagonalizes U^*U is the Fourier basis, and the singular values are given by the transfer function $|\hat{u}[k]|^2$, which is a band-pass filter. The noise W includes not only random measurement noise but also model errors, for example, neglecting multiple reflections in the wave propagation equation.

In a simple model, the underground impedance is approximated by piecewise constant functions corresponding to layers of homogeneous rocks. The reflectivity f is then a set of Diracs corresponding to the difference of impedance at the interfaces between different geophysical layers:

$$f[n] = \sum_{p\in\Lambda} a[p]\,\delta[n-p\,\Delta].$$

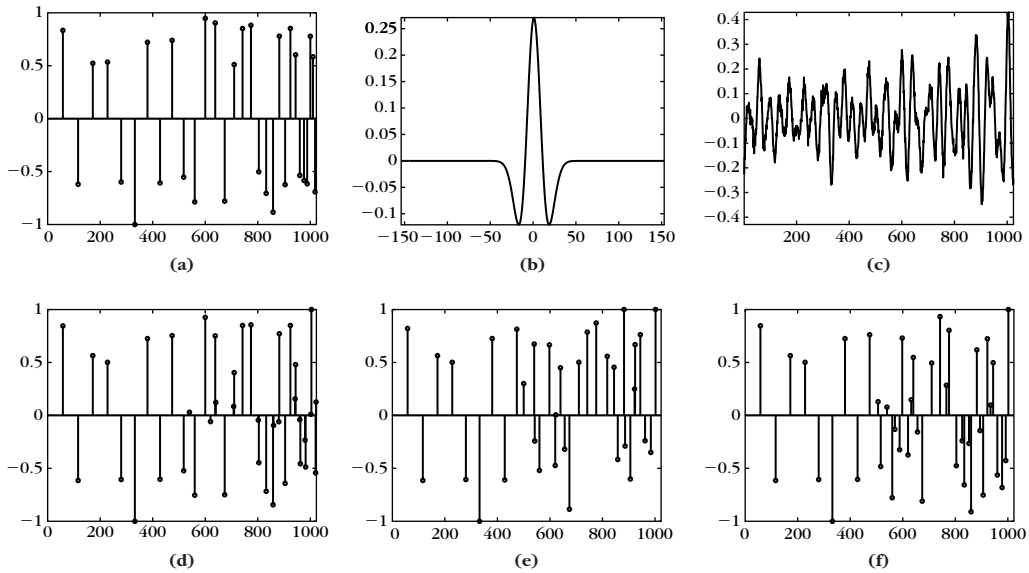


FIGURE 13.3

(a) Sparse spikes signal f; the distance between spikes decreases from left to right, from 30Δ to 5Δ. **(b)** Seismic wavelet u (zoom). **(c)** Measured seismic data y. **(d)** Sparse spikes deconvolution with an $\mathbf{l}^1$ pursuit, **(e)** a matching pursuit, and **(f)** an orthogonal matching pursuit.

Thus, it has a sparse representation in a Dirac orthonormal dictionary

$$\mathcal{D} = \{ g_p[n] = \delta[n - p\Delta] \}_{0 \leqslant p < P},$$

translated on a grid of interval $\Delta = N/P$ that defines the resolution of the sparse spike deconvolution. This dictionary fully satisfies the spread spectrum hypothesis since each Dirac has a flat spectrum in a Fourier basis.

Figure 13.3 shows a synthetic example of a sparse spike signal f of size $N = 1024$ and the resulting noisy observation $Y[n]$. The seismic wavelet $u[n]$ is the second derivative of a Gaussian. Without noise, the dimension $Q = \dim(\mathbf{ImU})$ of the observation space is the number of frequencies such that $|\hat{u}[k]| \neq 0$. In presence of noise, a linear or thresholding estimator can recover an estimation of f in a space of lower dimension Q_0 for which the amplified noise $|\hat{u}[k]|^{-1}\sigma$ is not above the maximum amplitude of signal coefficients. In this case, $Q_0 \approx 100$.

If f only includes Diracs that are far away so that $u[n - p\Delta]$ barely overlaps with $u[n - q\Delta]$ for $(p, q) \in \Lambda^2$, then the locations $p\Delta$ can be detected with a "match filtering," which is equivalent to a matching pursuit in the dictionary $\mathcal{D}_U$. An accurate estimation $\tilde{a}[p]$ of each $a[p]$ is then derived, which yields an estimation $\tilde{F}$ of f. This is a super-resolution estimation since the low and high frequencies of f are restored although they were fully removed by the band-pass filter U. The main difficulty of sparse spike estimation is to recover thin geophysical layers corresponding to closely located Diracs, producing overlapping seismic wavelets $u[n - p\Delta]$.

To identify these closely located *spikes*, in 1973 Clearbout and Muir [166] proposed to use an $\mathbf{l}^1$ minimization in the Dirac basis. Since $U g_p[n] = u \star \delta[n - p\Delta] = u[n - p\Delta]$, the transformed dictionary is a family of translated seismic wavelets:

$$\mathcal{D}_U = \left\{ \frac{u[n - p\Delta]}{\|u\|} \right\}_{0 \leqslant p < P}. \tag{13.52}$$

In 1986, Santosa and Symes [424] implemented this idea with an $\mathbf{l}^1$ relaxed minimization, which is a Lagrangian basis pursuit (13.35) of the seismic signal $Y[n]$ in the transformed dictionary of translated wavelets (13.52). It yields a sparse set of coefficients $\tilde{a}[p]$ from which a sparse spike estimation of f is derived according to (13.36):

$$\tilde{F} = \sum_{p \in \Gamma} \frac{\tilde{a}[p]}{\|u\|} \delta[n - p\Delta]. \tag{13.53}$$

The resolution Δ is set relative to the scale s of the wavelet. Increasing Δ reduces the maximum resolution of the sparse spike, but Section 12.5 shows that when Δ is too small, computations become unstable. Close wavelets $u[n - p\Delta]$ and $u[n - (p+1)\Delta]$ become too similar to choose between them, and the Lagrangian pursuit algorithm converges more slowly.

If $\mathrm{ERC}(\Lambda) < 1$ and the noise is small, Theorems 12.13 and 12.15 prove that an orthogonal matching pursuit as well as an $\mathbf{l}^1$ Lagrangian pursuit recover the spikes in the support Λ. Theorem 12.12 shows that $\mathrm{ERC}(\Lambda)$ decreases as the distance between spikes in Λ increases [232]. If u is the second-order derivative of a Gaussian

$$u[n] = \lambda \, (1 - s^{-2} n^2) \, \mathrm{e}^{-s^{-2} n^2 / 2},$$

then a numerical calculation shows that $\mathrm{ERC}(\Lambda) < 1$ if the distance between any two consecutive spikes in p and q satisfies $|p\Delta - q\Delta| \geqslant 5\, s$.

The sparse spike deconvolutions in Figure 13.3 are calculated with a wavelet scaled by $s = 10$ that corresponds to $Q_0 \approx 100$ frequency measurements over a signal of size $N = 1024$. The transformed dictionary includes $P = 512$ waveforms separated by $\Delta = 2$. If all spikes have a distance of 25Δ, then $\mathrm{ERC}(\Lambda) < 1$, and they can thus be recovered by an $\mathbf{l}^1$ Lagrangian pursuit or a matching pursuit. Figure 13.3 shows a sparse spike signal f and the measurement Y with noise. The $\mathbf{l}^1$ pursuit and matching pursuits are computed with a backprojection to restore the amplitude of spikes. The spikes have a spacing that decreases nonlinearly from 30Δ to 5Δ from left to right. The three algorithms recover all spikes up to a spacing of 22Δ (middle of the figure), whereas $\mathrm{ERC}(\Lambda) < 1$ is only for a spacing of 25Δ or larger. The three algorithms begin to fail below 22Δ but the $\mathbf{l}^1$ pursuit yields a higher SNR. In this example, $\|\tilde{F} - f\| / \|f\|$ is 0.45 for an $\mathbf{l}^1$ pursuit, and 0.9 for a matching pursuit and for an orthogonal matching pursuit. A matching pursuit and an orthogonal matching pursuit selects first the same "coherent structures" corresponding to the spikes on the left that have a distance larger than 22Δ.

Although slightly pessimistic, the $\mathrm{ERC}(\Lambda) < 1$ gives a good prediction for the recovery of signal components, but the $\mathbf{l}^1$ pursuit can still recover information below this limit. It improves the result of matching pursuit algorithms at a computational cost. Donoho et al. [225–227] prove that the $\mathbf{l}^1$ Lagrangian pursuit has the ability to recover closer spikes if they are not too numerous, but computations become unstable when their distance is reduced.

13.3.3 Recovery of Missing Data

Applications of super-resolution are studied for image zooming, Radon transform inversion in medical imaging, and image restoration with missing pixels. In missing data problems, partial observations are specified by a set of noisy measurements

$$Y[q] = Uf[q] + W[q] = \langle f, u_q \rangle + W[q] \quad \text{with} \quad q \in \Omega \quad \text{and} \quad |\Omega| = Q < N.$$

The family $\{u_q\}_{q\in\Omega}$ is a basis of a subspace $\mathbf{V}$ of dimension Q. A linear estimation of the orthogonal projection $P_{\mathbf{V}} f$ of f in $\mathbf{V}$ can be computed with a dual basis:

$$\tilde{F}_l[n] = \sum_{q\in\Omega} Y[q]\, \tilde{u}_q[n] = P_{\mathbf{V}} f[n] + \sum_{q\in\Omega} W[q]\, \tilde{u}_q[n] \in \mathbf{V}. \tag{13.54}$$

For super-resolution, the dictionary $\mathcal{D}$ must include vectors g_p with a spread spectrum. If $\{u_q\}_{q\in\Omega}$ is an orthonormal family, then $U^* U = P_{\mathbf{V}}$. So $\mathbf{V}$ is an eigenspace with singular value 1 and $\mathbf{V}^\perp$ is the other eigenspace with singular value 0. Each g_p should thus have orthogonal projections in $\mathbf{V}$ and in $\mathbf{V}^\perp$ that are relatively large.

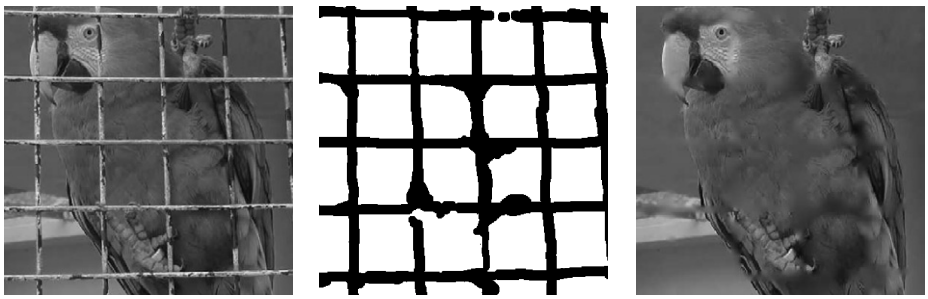
Image Inpainting

Inpainting is an example of missing data recovery for damaged images where pixel values are available in a known region Ω, and missing in its complement Ω^c:

$$Uf[q] = f[q] \quad \text{for} \quad q \in \Omega \quad \text{with} \quad |\Omega| = Q < N. \tag{13.55}$$

Elad et al. [244] as well as Fadili, Starck, and Murtagh [250] studied inpainting solutions with the sparse Lagrangian $\mathbf{l}^1$ pursuit minimization. This dictionary must include vectors with restrictions to Ω and Ω^c that have sufficiently large energy, while providing a sparse image representation. Figure 13.4 gives an example similar to [250], where the grid of the bird cage is removed from Ω. The interpolation $\tilde{F}$ in Figure 13.4(c) is computed with a translation-invariant dyadic wavelet dictionary $\mathcal{D}$. Fine-scale wavelets with a support nearly inside Ω^c must be removed from the dictionary to compute a normalized Lagrangian estimation (13.35). To simplify computations, they are kept in the dictionary and the super-resolution estimation $\tilde{F}$ is calculated with a nonnormalized $\mathbf{l}^1$ pursuit (13.39).

Figure 13.5(c) shows a second inpainting example in a dictionary $\mathcal{D}$ that is the union of a translation-invariant wavelet dictionary and a tight frame of local cosine vectors with a redundancy factor of 4. The restriction of the original image to Ω is shown in Figure 13.5(a). Local cosine vectors have a support larger than the size of the holes and satisfy the noise-amplification conditions, but this is not the



(a) (b) (c)

FIGURE 13.4

(a) Original image f. **(b)** Available pixels in Ω are shown in white. **(c)** Estimation $\tilde{F}$ computed with an $\mathbf{l}^1$ Lagrangian pursuit in a translation-invariant wavelet dictionary.

(a) (b) (c)

FIGURE 13.5

(a) Observed image restricted to Ω. **(b)** Linear Tikhonov interpolation (SNR = 16.2db). **(c)** Interpolation with an $\mathbf{l}^1$ Lagrangian pursuit in a wavelet and local cosine dictionary (SNR = 18.2 db).

case of fine-scale wavelets inside Ω_c. The estimation $\tilde{F}$ is thus also calculated with a nonnormalized $\mathbf{l}^1$ pursuit. The resulting SNR calculated relative to the original image without a hole is 18.2 db. It improves by 2 db the SNR obtained with a linear Tikhonov regularization in Figure 13.4(b).

Linear Tikhonov Regularization

As any inverse problem, missing data can be computed with a linear inverse estimator, studied in Section 13.1, which recovers

$$\tilde{F} = \underset{h \in \mathbb{R}^N}{\operatorname{argmin}} \|\Phi h\|^2 \quad \text{with} \quad \|Y - U\tilde{F}\|^2 \leqslant \varepsilon, \tag{13.56}$$

where Φ is a regularization operator. A Tikhonov minimization regularizes the solution with a gradient operator $\Phi h = \vec{\nabla} h$, and $\|\vec{\nabla} h\|^2$ is then a Sobolev norm that tends

to recover a uniformly regular image. The Lagrangian formula of this convex problem gives

$$\tilde{F} = \underset{h \in \mathbb{R}^N}{\operatorname{argmin}} \frac{1}{2} \|Uh - y\|^2 + T^2 \|\Phi h\|^2 \quad \text{with} \quad \Phi h = \vec{\nabla} \Phi, \tag{13.57}$$

where T is adjusted as a function of ε. The solution computed in (13.7) is

$$\tilde{F} = (U^* U + T^2 \Phi^* \Phi)^{-1} U^* Y.$$

It is calculated by inverting the symmetric operator $L = U^* U + T^2 \Phi^* \Phi$ with the conjugate-gradient algorithm or the Richardson gradient descent (see Section 5.1.3). As explained in Section 13.1, these linear estimators compute the solution in a space of dimension Q and thus do not perform any super-resolution.

For an inpainting problem, where image values are known in Ω, one can verify that the resulting solution satisfies $\triangle \tilde{F}[n] = 0$ for $n \in \Omega^c$ with boundary conditions specified by image values in Ω. The Tikhonov regularization thus diffuses the image values in Ω^c with an isotropic heat-diffusion equation. If the noise is neglected and thus $\varepsilon = 0$, then image values in Ω are not regularized and the boundary values of Ω^c are the values of f at the boundary of Ω. Figure 13.4(b) is an example of inpainting computed with a Tikhonov regularization. The SNR computed relative to the original image (without holes) is 16.2 db.

Total variation regularizations often do not outperform a linear Tikhonov regularization for image inpainting. Masnou and Morel [372] improved total variation regularization algorithms by also minimizing the $\mathbf{l}^1$ norm of the curvature of level sets. The solution is obtained with a nonlinear partial differential equation, which performs an anisotropic diffusion of the image values in the holes. Other partial differential equations that impose more geometric regularity have also been studied [94, 110, 154, 249, 465]. The algorithms give good results but can have instabilities when the domain Ω^c is nonconvex and complex, as in Figure 13.4.

Image Scaling and Deinterlacing

Image and video screens often have more pixels than the images that are displayed. To fit the whole screen, images must be scaled, while restoring as many details as possible and minimizing artifacts. This is a major challenge for videos, in particular for high-definition television (HDTV). Indeed, most current television images are in an interlaced standard-definition television (SDTV) format (PAL or NTSC). Interlacing means that one image out of two carries only the even rows and the next one carries only the odd rows, which is adapted to CRT television displays. Flat HDTV screens simultaneously display the even and odd rows of each image. The number of rows and columns of high-definition images is also at least twice as large as SDTV formats [358]. Thus, to display SDTV interlaced images on HDTV screens requires us to increase the number of pixels by at least 8 for each image. Moreover, recent screens display 120 or 100 images per second as opposed to 60 or 50, which also requires us to double the number of images in time. In such scaling applications, the image is known over a coarse regular spatial or space–time

grid Ω having Q pixels, and image values must be interpolated on a finer grid with N pixels. For SDTV to HDTV conversion, $N \geqslant 16\,Q$. The noise $W[n]$ is often complex because it is dominated by compression artifacts as opposed to camera noise.

If the noise can be neglected, then known image values are preserved and a linear interpolation computes

$$\tilde{F}[n] = \sum_{m \in \Omega} Y[m]\, \theta[n-m].$$

The interpolation kernel satisfies $\theta[n-m] = \delta[n-m]$ for any $(n, m) \in \Omega^2$, so that $\tilde{F}[m] = Y[m]$ for $m \in \Omega$. A quadratic minimization (13.56) with $\varepsilon = 0$ implements such an interpolation. For a Tikhonov regularization with $\Phi = \vec{\nabla}$, the kernel θ computes a linear interpolation. Cubic spline interpolation kernels θ are most often used in image processing and correspond to a third-order differential operator Φ [458]. Figure 13.6(a) shows an example of a linear scaling by four along the image rows and columns with a cubic spline interpolation. The image is blurred and oscillations appear along directional structures such as contours.

Instead of interpolating all pixels with a predefined linear kernel, adaptive directional interpolations adapt the interpolation kernel for each missing pixel, depending on the observed image regularity. If the image regularity is not isotropic, an elongated kernel is used to perform the interpolation along a direction where the image is locally the most regular. Along edges, the interpolation kernel is typically elongated in the direction of the edge tangent. Such techniques are used in industry

(a) (b)

FIGURE 13.6

(a) Separable linear interpolation by a factor of four along rows and columns with cubic splines. **(b)** Nonlinear directional interpolation with the same factors.

for video deinterlacing and scaling. Finding locally the best interpolation directions and optimizing the shapes of the kernels are difficult problems, most often solved with ad hoc algorithms. Yet, good results are obtained even on complex images, as shown in Figure 13.6(b).

Nonlinear scaling can be computed with a sparse super-resolution estimator in a dictionary $\mathcal{D}$ with vectors g_p that intersect both Ω and Ω^c, while providing a sparse image representation. Such dictionaries must include elongated directional waveforms of large support such as curvelets or bandlets in order to take advantage of the directional image regularity [290]. A Lagrangian pursuit estimator can then be interpreted as an adaptive image interpolation. The interpolation directions and the size of the scale of the interpolation kernels correspond to the direction and size of the reconstructing dictionary waveforms Ug_p for $p \in \tilde{\Lambda}$, computed to recover a sparse representation of the observed image Y on Ω.

The ability to achieve some super-resolution depends on the geometry of f relative to Ω. If Ω is a square subsampled grid, then one can verify that no super-resolution is possible along a strictly horizontal or vertical edge. When edge angles are very close to horizontal and vertical, some super-resolution is possible but limited by instabilities. Similarly, when video images do not move, constant values in time provide no information to increase the spatial resolution.

Tomography Inversion

A two-dimensional X-ray tomographic imaging system measures the Radon transform of body slices $\bar{f}(x)$ along a limited number of angles $\{\theta_1, \ldots, \theta_L\}$ in order to reduce the exposition time of patients. The Radon transform of $\bar{f}(x)$ along a ray parameterized by $x_1 \cos\theta + x_2 \sin\theta = \tau$ is

$$\forall\, \tau \in \mathbb{R}, \quad \overline{U}\bar{f}(\theta, \tau) = p_\theta(\tau) = \int\!\!\int \bar{f}(x)\, \delta(x_1 \cos\theta + x_2 \sin\theta - \tau)\, dx.$$

The Fourier slice theorem (2.10) proves that

$$\hat{p}_\theta(\omega) = \widehat{\bar{f}}(\omega \cos\theta, \omega \sin\theta). \tag{13.58}$$

A linear orthogonal projection of f can be computed from these observations over the space $\mathbf{V}$ generated by Fourier vectors at the available frequencies. It is obtained with the backprojection theorem (2.11) as a partial sum:

$$P_{\mathbf{V}}\bar{f}(x) = \frac{1}{L}\sum_{\ell=1}^{L} p_{\theta_\ell} \star h(x_1 \cos\theta_\ell + x_2 \sin\theta_\ell) \quad \text{with} \quad \hat{h}(\xi) = |\xi|. \tag{13.59}$$

Tomographic measurement systems provide discrete measurements with noise

$$Y[l, q] = Uf[l, q] + W[l, q] = p_{\theta_l}[q] + W[l, q] \quad \text{for} \quad 1 \leqslant l \leqslant L,$$

from which we want to recover an estimation $\tilde{F}$ of a high-resolution image $f \in \mathbb{C}^N$. According to the Fourier slice theorem, the Fourier transform of $Y[l, q]$ along q gives

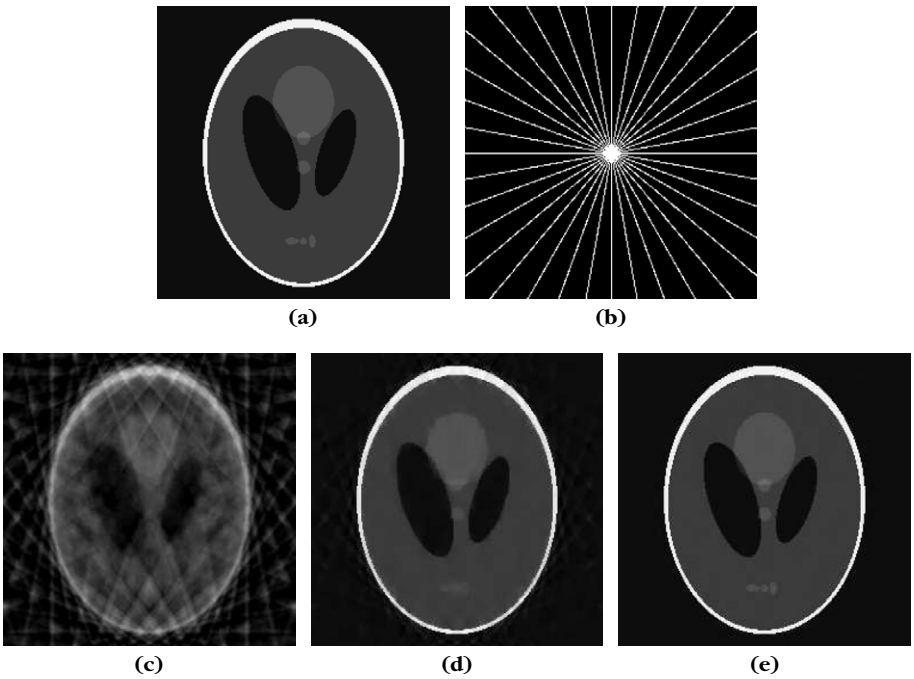


FIGURE 13.7

(a) Original phantom image. **(b)** Frequency plane showing in white the frequency rays in Ω. **(c)** Reconstruction with a linear orthogonal projection computed with a backprojection. **(d)** Lagrangian pursuit estimation in a Haar translation-invariant wavelet dictionary. **(e)** Inversion with a total variation regularization.

noisy measurements of $\hat{f}[m]$ along rays $m \in \Omega$, as illustrated in Figure 13.7. Inverse tomography is thus a missing Fourier data recovery problem. The Fourier basis is a basis of singular vectors that diagonalize U^*U with singular values equal to 1 on Ω. A super-resolution recovery requires using a dictionary of spread spectrum vectors, which have a Fourier transform as delocalized as possible.

Medical image models are often piecewise regular, and thus have sparse wavelet approximations. Simple piecewise constant phantom images, as shown in Figure 13.7, are often used to evaluate inversion algorithms. In this case, the most sparse wavelet representation is obtained with Haar wavelets, which are discontinuous and thus have a Fourier transform that is quite spread out. Figure 13.7(c) shows the image reconstructed with a linear backprojection. The resulting image is highly oscillatory because of the missing frequencies. Figure 13.7(d) shows a nonnormalized Lagrangian $\mathbf{l}^1$ pursuit estimation (13.39) with a Haar translation-invariant wavelet dictionary. It recovers a much more precise and sharp piecewise constant estimation.

Total Variation Regularization

Instead of supposing that the solution has a sparse synthesis in a dictionary as we did in this section, a sparse analysis assumes that a particular linear image transform Φf is sparse. As explained in Section 12.4.4, an estimator $\tilde{F}$ of f can be defined as

$$\tilde{F} = \underset{h \in \mathbb{R}^N}{\operatorname{argmin}} \|\Phi h\|_1 \quad \text{with} \quad \|Y - U\tilde{F}\|^2 \leqslant \varepsilon. \tag{13.60}$$

For images, Rudin, Osher, and Fatemi [420] introduced this approach with $\Phi f = \vec{\nabla} f$, in which case $\|\Phi f\|_1 = \|\vec{\nabla} f\|_1 = \|f\|_V$ is the total image variation. The Lagrangian formulation then computes

$$\tilde{F} = \underset{h \in \mathbb{R}^N}{\operatorname{argmin}} \frac{1}{2}\|y - Uh\|^2 + T\|\Phi h\|_1 \quad \text{with} \quad \Phi h = \vec{\nabla} h. \tag{13.61}$$

Section 12.4.4 describes an iterative algorithm solving this minimization.

This minimization (13.60) looks similar to the Tikhonov regularization (13.56), where the $\mathbf{l}^2$ norm $\|\Phi h\|$ is replaced by a $\mathbf{l}^1$ norm $\|\Phi h\|_1$, but the estimator properties are completely different. A $\mathbf{l}^2$ norm is minimized by maintaining small-amplitude coefficients distributed uniformly, which yields a uniformly regular signal with a Tikhonov regularization computed with $\Phi = \vec{\nabla}$. As explained in Section 12.4.1, the minimization of a $\mathbf{l}^1$ norm tends to produce many zero- or small-amplitude coefficients and few large-amplitude ones. For $\Phi = \vec{\nabla}$, the coarea theorem (2.9) proves that the total image variation $\|\vec{\nabla} f\|_1 = \|f\|_V$ is the average length of the level sets of f.

The phantom image of Figure 13.7(a) is ideal for total variation estimation. Indeed, the gradient is zero everywhere outside the edges of the image objects, which have a length that is not too large. Figure 13.7(e) is obtained by minimizing the Lagrangian formulation (13.61) of the total variation minimization with the Radon transform operator U. Without noise, this total variation regularization performs an almost exact recovery of the input image f, which is not the case of the Lagrangian pursuit with Haar wavelets. Indeed, the gradient field is more sparse than with a multiscale Haar wavelet transform. Haar wavelets do not restore the boundaries of the image phantoms as precisely as a total variation regularization, which minimizes the length of restored contours.

Real medical images are not piecewise constant and include much more complex structures. Total variation estimations are therefore not as spectacular on real images. They have a tendency to remove textures and oscillatory structures by producing flat image areas, which can reduce the SNR.

13.4 COMPRESSIVE SENSING

Super-resolution is not always possible and often unstable for usual measurement operators U. Candès and Tao [138, 139] as well as Donoho [217, 218] observed that sparse super-resolution becomes stable for all sufficiently sparse signals when

U is a random measurement operator that computes random linear combinations of all signal values. This remarkable result opens the door to compressive sensing strategies, where randomized linear measurements can recover a higher-resolution approximation of signals that have a sparse approximation in some dictionary. It also gives a conceptual probabilistic framework where random mixing appears as an efficient information acquisition strategy for structured information.

13.4.1 Incoherence with Random Measurements

As opposed to previously studied inverse problems, in compressive sensing, we have the luxury to design the measurement operator that is not imposed. Random sensing operators create highly incoherent transformed dictionaries where sufficiently sparse signals have a stable recovery.

Compressive Sensing Acquisition and Recovery

The compressive sensing acquisition of an analog signal $\bar{f}(x)$ is implemented with a continuous sensing operator $\overline{U}$ that provides Q measurements

$$Y[q] = \overline{U}\bar{f}[q] + W[q] = \langle \overline{u}_q, \bar{f} \rangle + W[q]. \tag{13.62}$$

In analog compressive sensing, the hardware device outputs randomized analog measurements with transfer functions $\bar{u}_q(x)$ that are realizations of a random process. This acquisition is modeled with a stable high-resolution analog-to-digital converter $\overline{\Phi}_s$ followed by a discrete operator U, which outputs Q random combinations of these high-resolution measurements: $\overline{U} = U\overline{\Phi}_s$. Measurements can thus be rewritten as

$$Y[q] = Uf[q] + W[q] \quad \text{for} \quad 0 \leqslant q < Q,$$

where $f = \overline{\Phi}_s \bar{f} \in \mathbb{C}^N$ is a high-resolution discretization of $\bar{f}(x)$ with $N \gg Q$. A super-resolution operator computes an estimation $\tilde{F} \in \mathbb{C}^N$ of f from the vector Y of Q measurements.

Suppose that f has a sparse approximation in a dictionary $\mathcal{D} = \{g_p\}_{p\in\Gamma}$. According to the sparse super-resolution algorithm described in Section 13.3, an estimation of $f \in \mathbb{C}^N$ is computed by finding a sparse approximation of Y in the transformed dictionary $\mathcal{D}_U = \{Ug_p\}_{p\in\Gamma}$. Suppose that U is a random matrix with coefficients that are obtained with independent random variables of the same distribution and variance Q^{-1}. For large N, the law of large numbers guarantees with high probability that $\|Ug_p\|^2$ is close to 1. It implies with a high probability that any $g_p \in \mathcal{D}$ has a spread singular spectrum relative to the singular vectors and singular values of U^*U. Each Ug_p is thus already nearly normalized.

Let $\Phi f[p] = \langle f, g_p \rangle$. The super-resolution estimator is

$$\tilde{F} = \sum_{p\in\tilde{\Lambda}} \tilde{a}[p]\, g_p = \Phi^* \tilde{a},$$

where $\tilde{a}$ are sparse approximation coefficients of Y in the transformed dictionary $\mathcal{D}_U$. They can be computed with an $\mathbf{l}^1$ pursuit minimization

$$\tilde{a} = \text{argmin}_{a \in \mathbb{C}^P} \|a\|_1 \text{ subject to } \|\sum_{p \in \Gamma} a[p]\, Ug_p - Y\| \leqslant \varepsilon, \tag{13.63}$$

or as a solution of an $\mathbf{l}^1$ Lagrangian pursuit

$$\tilde{a} = \underset{a \in \mathbb{C}^P}{\text{argmin}} \frac{1}{2} \|\sum_{p \in \Gamma} a[p] Ug_p - Y\|^2 + T\|a\|_1. \tag{13.64}$$

The sparse approximation $\tilde{a}$ can also be computed with a matching pursuit decomposition of Y in $\mathcal{D}_U$.

Restricted Isometry and Incoherence

Section 12.5 explains that an estimation of a sparse approximation f_Λ is possible by decomposing Y in the transformed dictionary $\mathcal{D}_U = \{Ug_p\}_{p \in \Gamma}$ only if $\{Ug_p\}_{p \in \Lambda}$ is a frame with a frame bound ratio A_Λ / B_Λ that is not close to 0. If the vectors are normalized vectors, then $A_\Lambda \leqslant 1 \leqslant B_\Lambda$, and it is equivalent to impose that $\delta_\Lambda = \max(1 - A_\Lambda, B_\Lambda - 1)$ is not too small. To get a stable recovery of all sparse signals, compressive sensing imposes a uniform bound on all sufficiently sparse sets Λ:

$$\delta_\Lambda \geqslant \delta_M(\mathcal{D}_U) > 0 \text{ if } |\Lambda| \leqslant M,$$

where $\delta_M(\mathcal{D}_U)$ is called an *M-restricted isometry bound*. It results that for all Λ with $|\Lambda| \leqslant M$,

$$\forall a \in \mathbb{C}^{|\Lambda|}, \quad (1 - \delta_M(\mathcal{D}_U)) \sum_{p \in \Lambda} |a[p]|^2 \leqslant \|\sum_{p \in \Lambda} a[p] Ug_p\|^2 \leqslant (1 + \delta_M(\mathcal{D}_U)) \sum_{p \in \Lambda} |a[p]|^2. \tag{13.65}$$

Theorem 12.10 relates $\delta_M(\mathcal{D}_U)$ to the dictionary mutual coherence

$$\delta_M(\mathcal{D}_U) \leqslant (M-1)\, \mu(\mathcal{D}_U) \quad \text{with} \quad \mu(\mathcal{D}_U) = \max_{(p,q) \in \Gamma^2, p \neq q} \langle Ug_p, Ug_q \rangle. \tag{13.66}$$

However, the mutual coherence $\mu(\mathcal{D}_U)$ does not provide a tight upper bound of $\delta_M(\mathcal{D}_U)$. Indeed, it depends on the correlation of pairs of dictionary vectors, whereas $\delta_M(\mathcal{D}_U) > 0$ measures the stability of potentially much larger groups of M dictionary vectors. Restricted isometry bounds are stronger measures of the dictionary incoherence. A simple geometric interpretation explains why random measurement operators define incoherent dictionaries with $\delta_M(\mathcal{D}_U) > 0$ for relatively large M.

We know that all dictionary vectors Ug_p belong to the space **ImU** of dimension Q. An orthonormal basis is a stable family of Q vectors that are perfectly spread on the unit sphere of **ImU**. A family $\{Ug_p\}_{p \in \Lambda}$ is a stable Riesz basis of a subspace if these points remain well distributed on this sphere. For this result to be valid for any collection of less than M vectors in a dictionary of size $P > Q$, we need to distribute as uniformly as possible these P vectors on the unit sphere of **ImU**. A natural

idea is to define such vectors as P realizations of a Gaussian white noise. Indeed, a Gaussian white noise vector of dimension Q and variance Q^{-1} has a probability density that is constant on all spheres of $\mathbb{R}^Q$, and each realization has a norm that is close to 1 when Q is large. It is thus highly likely that these P realizations are well spread on the unit sphere. If $\mathcal{D}$ is an orthonormal basis and U is a Gaussian random matrix with coefficients that are independent Gaussian random variables, then $\{Ug_p\}_{p\in\Lambda}$ are also Q independent Gaussian random variables of variance Q^{-1}. Random matrix operators are thus good candidates to build a transformed dictionary that satisfies the M-restricted isometry condition (13.65) for a relatively large M. For the mathematical analysis, we shall suppose in the following that the dictionary $\mathcal{D}$ is an orthonormal basis.

Gaussian and Bernouilli Random Sensing Matrices

Up to now, random matrices are the only universal large-size matrices that ensure that the vectors $\{Ug_p\}_p$ are nearly uniformly spread around the unit sphere of $\mathbb{C}^Q$ with a high probability for any fixed orthogonal basis $\mathcal{D}$. This is necessary to guarantee that any collection of less than M vectors defines a Riesz basis for M relatively large. All known deterministic sensing matrices U have some regularity that prevents the set of vectors $\{Ug_p\}_{p\in\Gamma}$ to be sufficiently well distributed.

A Gaussian random matrix U has coefficients that are realizations of independent Gaussian random variables of mean 0 and variance Q^{-1}. Its rows and columns are thus realizations of Gaussian white noise random vectors. The mutual coherence $\mu(\mathcal{D}_U)$ of $\mathcal{D}_U = \{Ug_p\}_p$ can be shown to be $O(\sqrt{(\log N)/Q})$ with a high probability [230]. The inequality (13.66) derives that $\delta_M(\mathcal{D}_U) < 1$ for $M = O(\sqrt{Q/(\log N)})$. Theorems 12.14 and 12.15 also prove that the approximation support of any signal with $M = O(\sqrt{Q/(\log N)})$ nonzero coefficients is recovered by an orthogonal matching pursuit or $\mathbf{l}^1$ Lagrangian pursuit. Candès and Tao [139] as well as Donoho [217] proved that this result can be considerably improved.

Theorem 13.6: *Candès, Tao, Donoho.* Let U be a Gaussian matrix and $\mathcal{D}$ be an orthonormal basis. For any $\delta < 1$, there exists a constant β such that for

$$M \leqslant \frac{\beta Q}{\log(N/Q)}, \tag{13.67}$$

the dictionary $\mathcal{D}_U = U\mathcal{D}$ satisfies $\delta_M(\mathcal{D}_U) \leqslant \delta$ with a probability that increases toward 1 exponentially fast with N.

Proof. Let Φ be the analysis matrix associated to $\mathcal{D}$: $\Phi f = \langle f, g_p\rangle$. The matrix Φ_U is associated to the transformed dictionary $\mathcal{D}_U = U\mathcal{D}$ is $\Phi_U = \Phi U^*$. Indeed, $\Phi_U f = \langle f, Ug_p\rangle$. If U is a Gaussian random matrix, then its columns $\{u_q[p]\}_{p\in\Gamma, 0\leqslant q<Q}$ are realizations of Q independent Gaussian white noise. It results that $\Phi_U = \{\langle g_p, u_q\rangle\}_{p\in\Gamma, 0\leqslant q<Q}$ are the decomposition coefficients of these Q white noises in an orthonormal basis that remain independent Gaussian random variables. The matrix Φ_U is thus also a Gaussian random matrix with Q columns that are realizations of independent Gaussian white noises. Its restriction Φ_Λ to $|\Lambda|$ rows indexed in Λ is also a Gaussian random matrix with Q columns.

Since $\Phi^*_{U\Lambda} a = \sum_{p\in\Lambda} a[p]\, U g_p$,

$$\langle \Phi_{U\Lambda}\, \Phi^*_{U\Lambda} a, a\rangle = \| \sum_{p\in\Lambda} a[p] U g_p \|^2 .$$

Proving the M-restricted isometry bound (13.65) satisfies $\delta_M(\mathcal{D}_U) \leqslant \delta$ is thus equivalent to proving that the largest and smallest eigenvalues of $\Phi_{U\Lambda}\, \Phi^*_{U\Lambda}$ are between $1-\delta$ and $1+\delta$ for any set of $|\Lambda| \leqslant M$ rows chosen among N. This means finding an upper and lower bound of the maximum and minimum singular value of a Gaussian random matrix having $|\Lambda|$ rows and Q columns with a probability that tends to 1 when N increases.

The matrix $\Phi_{U\Lambda}\, \Phi^*_{U\Lambda}$ is an $|\Lambda|$ by $|\Lambda|$ symmetric matrix with coefficients that are inner products between Λ realizations of independent Gaussian white noises of size Q chosen among N. Upper bounds of Gaussian matrix singular values are relatively classic results [245] and recent concentration inequalities have been proved on the smallest singular value [353]. Candès and Tao [139] as well Donoho [217] prove that these concentration inequalities imply that for any $\delta < 1$, there exists a constant β such that for all $M \leqslant \beta Q/(\log N/Q)$, the maximum and minimum eigenvalues are bounded by $1-\delta$ and $1+\delta$. ■

This theorem proves that up to a logarithmic factor, $\delta_M < 1$ for M proportional to Q, as opposed to $\sqrt{Q}$ as in the result obtained with the dictionary mutual coherence. However, the constant β derived from upper and lower bounds of Gaussian matrix singular values is very small. Gaussian random matrices are universal in the sense that this result does not depend on the orthonormal basis $\mathcal{D}$.

Implementing in hardware an operator $\bar{U}\bar{f}$ that projects signals over independent Gaussian white noise processes can be difficult. The numerical super-resolution estimation $\tilde{F}$ of $f \in \mathbb{C}^N$ also requires us to store the transformed dictionary $\{U g_p\}_{p\in\mathbb{Z}}$ in a memory of size $O(QN)$, which is huge when Q and N are large. The estimation with an $\mathbf{l}^1$ Lagrangian pursuit or with an orthogonal matching pursuit then iteratively decomposes signals in this unstructured dictionary. It requires QN additions and multiplications each time, which is again too much when Q and N are large. This suggests finding other random matrices providing M-restricted isometry bounds $\delta_M(\mathcal{D}_U)$ similar to Gaussian random matrices, but with less memory and computational requirements.

Bernouilli random matrices have random entries that are independent Bernouilli random variables, thus taking values ± 1 with a probability $1/2$. These matrices are renormalized by $Q^{-1/2}$ so that $\|U g_p\| = 1$. Candès and Tao [139] prove that Theorem 13.6 still holds for Bernouilli matrices. The constant β is smaller, however, because realizations of Bernouilli processes do not have a uniform probability density over spheres of $\mathbb{R}^Q$ as Gaussian white noises do. The vectors $U g_p$ of the transformed dictionary are therefore not as uniformly distributed as with a Gaussian process. Bernouilli random matrices replace all multiplications by additions and subtractions, which requires less operations and storage, but the computational complexity and memory storage remain large when N is large.

Random Projectors

More structured random operators are constructed with a random subsampling of the decomposition coefficients of a signal in an orthonormal basis $\mathcal{B} = \{u_m\}_{0 \leq m < N}$. The operator U is then an orthogonal projector on a family of Q vectors $\{u_q\}_{q \in \Gamma_Q}$ where the set Γ_Q is randomly chosen with a uniform probability distribution among all subsets of size Q in an index set of size N. Since U is an orthogonal projection, decomposing coefficients in the transformed dictionary can be written as $\langle f, U g_p \rangle = \langle U f, g_p \rangle$. They are computed with a fast algorithm if the projector U is implemented with a fast algorithm as well as signal decompositions in the orthonormal basis $\mathcal{D} = \{g_p\}_{p \in \Gamma}$. If U is a projection over Fourier basis vectors of random frequencies, Uf is computed with $O(N \log N)$ operations with an FFT.

Theorem 13.7, proved by Candès, Romberg, and Tao [139] and Rudelson and Vershynin [419], shows that randomized subsampled orthogonal transforms have low restricted isometry constants if the mutual coherence $\mu(\mathcal{B} \cup \mathcal{D})$ of the union of bases is small.

Theorem 13.7: *Candès, Romberg, Tao, Rudelson, Vershynin.* Let $\mathcal{D}$ be an orthonormal basis and U be a projector on a randomly chosen subset of vectors in an orthonormal basis $\mathcal{B}$. For any $\delta < 1$, there exists β such that for all M satisfying

$$M \leq \frac{\beta Q}{N \mu(\mathcal{B} \cup \mathcal{D})^2 (\log N)^5}, \tag{13.68}$$

the dictionary $\mathcal{D}_U = U\mathcal{D}$ satisfies $\delta_M(\mathcal{D}_U) \leq \delta$ with a probability that increases toward 1 like $1 - N^{-c}$ where c is a constant.

A randomized subsampled orthogonal transform is not universal as opposed to Gaussian or Bernouilli random matrices. The upper bound depends on the mutual coherence of the random sampling orthonormal basis $\mathcal{B}$ with the basis $\mathcal{D}$ that provides a sparse signal representation. If $\mathcal{B}$ is a discrete Fourier basis and $\mathcal{D}$ is a Dirac basis then $\mu(\mathcal{B} \cup \mathcal{D}) = N^{-1/2}$, and (13.68) becomes $M \leq CQ/(\log N)^5$. In applications, random subsampling projectors are often used because they are computationally more efficient than Gaussian or Bernouilli random matrices [95].

Randomized Sparse Spike Deconvolutions

A random Fourier sampling for sparse spike signals can be interpreted as a randomized sparse spike deconvolution. A comparison with standard sparse spike deconvolutions shows the importance of randomization to improve the signal recovery.

In a sparse spike deconvolution problem, $Y = f \star u + W$ and $f = \sum_{p \in \Lambda} a[p]\, \delta[n - p\Delta]$ is sparse in the Dirac basis $\mathcal{D}$. As explained in Section 13.3.2, sparse spike deconvolutions estimate the coefficients $a[p]$ by decomposing Y in the transformed dictionary $\mathcal{D}_U = \{u[n - \Delta p]\}_p$ with u normalized $\|u\| = 1$. Since $\langle u[n - p\Delta], u[n - q\Delta] \rangle = u \star \bar{u}[\Delta(p - q)]$ with $\bar{u}[n] = u[-n]$, the dictionary mutual coherence is

$$\mu(\mathcal{D}_U) = \max_{p \neq 0} |\bar{u} \star u[p\Delta]|. \tag{13.69}$$

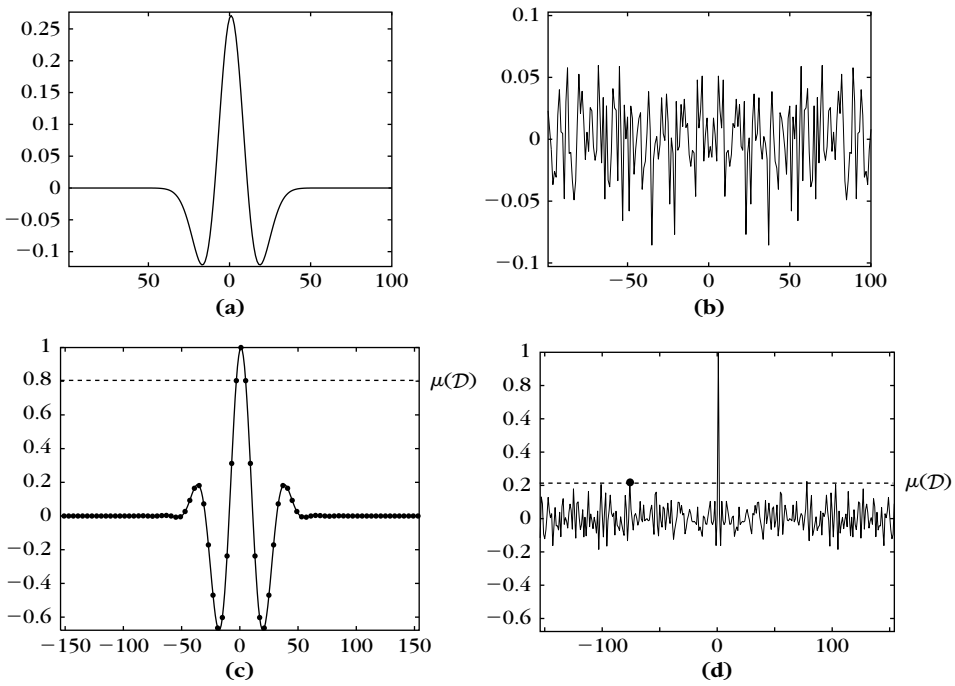


FIGURE 13.8

(a) Gaussian second-derivative wavelet $u_1[n]$. **(b)** Filter $u_2[n]$ with $\hat{u}_2[k]$ equal to 1 over Q random symmetric frequencies. **(c)** Value of $u_1 \star \bar{u}_1[n]$ with $\mu(\mathcal{D}_U) = 0.8$. The dots have a spacing of Δ. **(d)** Value of $u_2 \star \bar{u}_2[n]$ with $\mu(\mathcal{D}_U) = 0.2$.

A seismic wavelet is a band-pass filter. Figure 13.8(a) gives an example of filter u_1, which is the second derivative of a Gaussian scaled by s. Since it is regular, the resulting dictionary coherence is close to 1 if Δ/s is small, as illustrated in Figure 13.8(c) where $\mu(\mathcal{D}_U) = 0.8$.

A random Fourier sampling with $Q/2$ random positive frequencies and $Q/2$ symmetric negative frequencies is a convolution with a real filter u_2 that has a Fourier transform $\hat{u}_2[k]$ that is 1 over Q frequencies. Figure 13.8(b) shows that $u_2[n]$ is highly irregular with a uniformly spread energy. As a result, $\mathcal{D}_U$ has a low mutual coherence $\mu(\mathcal{D}_U) = \max_{p>0} |u_2 \star \bar{u}_2[p\Delta]| = 0.2$, shown in Figure 13.8(d).

The Gaussian derivative filter u_1 has a Fourier transform that is nonnegligible over $Q = 100$ frequencies like the Fourier random sampling filter u_2. Both filters thus provide Q frequency measurements. Figure 13.9(a) shows an example of sparse spike signal. Figures 13.9(b, c) give the estimated sparse spike signal $\tilde{F}_1$ and $\tilde{F}_2$ recovered from $Y_1 = f \star u_1 + W$ and $Y_2 = f \star u_2 + W$ with an $\mathbf{l}^1$ pursuit. As expected, close spikes are not recovered with the Gaussian derivative u_1, but they are recovered with u_2. This random Fourier sampling filter is able to recover the location and sign of 19 Diracs, including very close ones, from $Q = 100$ Fourier frequencies.

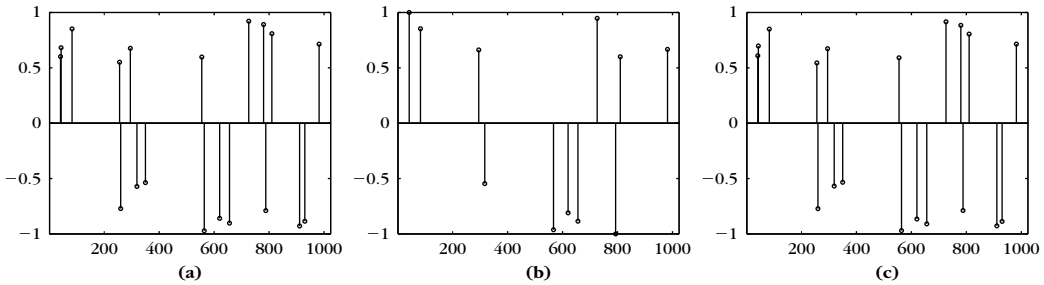


FIGURE 13.9

(a) Sparse spike f. **(b)** Estimation $\tilde{F}_1$ from $Y_1 = u_1 \star f + W$. **(c)** Estimation $\tilde{F}_2$ from $Y_2 = u_2 \star f + W$.

The dictionary incoherence not only improves the stability and precision of the signal recovery but also the convergence of the iterative thresholding algorithm detailed in Theorem 12.9. This could suggest sending random waves for seismic exploration as opposed to usual "wavelets," but regretfully seismic wavelet design must also take into account geophysical constraints and the attenuation of underground wave propagation.

13.4.2 Approximations with Compressive Sensing

Compressive sensing provides an alternative to linear and nonlinear approximation strategies. Previously described approximation schemes first compute a linear approximation in a high-resolution space $\mathbf{U}_N$ with an error ε. A sparse support of size $M \ll N$ is then calculated in some dictionary with an error that remains of the order of ε. Compressive sensing suggests to directly perform a sparse signal measurement with Q measurements while restoring a signal with the same approximation error. It is proved that compressive sensing recovery can indeed have an approximation error that has the same asymptotic decay as a nonlinear approximation error, but the devil is in the constants.

Approximation Error

If f is measured at a high resolution N, a nonlinear approximation is computed in an orthonormal basis $\mathcal{D} = \{g_p\}_{p \in \Gamma}$ of $\mathbb{C}^N$ by finding the M-largest coefficients. Let $\{f_{\mathcal{D}}^r[k]\}_{1 \leqslant k \leqslant N}$ be the ranked coefficients $\langle f, g_p \rangle$ in $\mathcal{D}$: $|f_{\mathcal{D}}^r[k]| \geqslant |f_{\mathcal{D}}^r[k+1]|$. The M-term approximation error is the energy of leftover smaller coefficients

$$\|f - f_M\|^2 = \sum_{k > M} |f_{\mathcal{D}}^r[k]|^2.$$

Instead of measuring f at a high resolution N, computing N coefficients $\langle f, g_p \rangle$, and throwing away the $N - M$ smallest ones, we want to acquire a smaller number of Q measurements with the same error. Since we do not know where the M-largest coefficients of f in $\mathcal{D}$ are, compressed sensing performs $Q \geqslant M$

spread random measurements. Recovering an estimation nearly as precise as this M-term approximation strongly relies on M-restricted isometry conditions in the transformed dictionary

$$\forall a \in \mathbb{C}^{|\Lambda|}, \quad (1-\delta_M(\mathcal{D}_U)) \sum_{p\in\Lambda} |a[p]|^2 \leqslant \| \sum_{p\in\Lambda} a[p] U g_p \|^2 \leqslant (1+\delta_M(\mathcal{D}_U)) \sum_{p\in\Lambda} |a[p]|^2 \tag{13.70}$$

for all $\Lambda \subset \Gamma$ with $|\Lambda| \leqslant M$.

Theorem 13.8 by Candès, Romberg, and Tao [137] gives an error bound to compare the nonlinear approximation error with the compressive sensing estimation $\tilde{F}$ computed from $Y = Uf + W$. In this context, the noise W has a small energy that only limits the computational precision. A similar theorem is given by Donoho [219]. We write $\Phi f = \langle f, g_p \rangle$ as the decomposition operator of $\mathcal{D}$.

Theorem 13.8: *Candès, Romberg, Tao.* Let U be a sensing matrix and $\mathcal{D} = \{g_p\}_{p\in\Gamma}$ be an orthonormal basis. Suppose that M satisfies $\delta_{3M}(\mathcal{D}_U) \leqslant \delta < 1/3$ with $\mathcal{D}_U = \{U g_p\}_{p\in\Gamma}$. Let $\tilde{F} = \sum_{p\in\Gamma} \tilde{a}[p] g_p$ be defined by

$$\tilde{a} = \operatorname*{argmin}_{b\in\mathbb{C}^N} \|b\|_1 \quad \text{subject to} \quad \| \sum_{p\in\Gamma} b[p]\, U g_p - Y \| \leqslant \|W\|.$$

There exists a constant C that only depends on δ such that

$$\|f - \tilde{F}\| \leqslant \frac{C}{\sqrt{M}} \sum_{k>M} |f_{\mathcal{D}}^r[k]| + C\, \|W\|, \tag{13.71}$$

where $f_{\mathcal{D}}^r[k]$ are the ranked coefficients $f_{\mathcal{D}}$ ordered by decaying magnitude.

Proof. The following proof uses arguments provided in [169]. We write $a[p] = \langle f, g_p \rangle$ and $\tilde{a} = \langle \tilde{f}, g_p \rangle$ as the decomposition coefficients in the orthonormal basis $\mathcal{D}$. Let $\Lambda \subset \Gamma$ be the indexes p of the M-largest coefficients $|a[p]|$. The coefficient error $b = a - \tilde{a}$ is evaluated on the complement Λ^c of Λ in Γ. This complement is subdivided into $\Lambda^c = \Lambda_1 \cup \ldots \cup \Lambda_k$, where Λ_1 are the indexes p of the $2M$ largest coefficients $|b_{\Lambda^c}[p]|$, Λ_2 are the indexes of the next $2M$ largest coefficients, and so on until Λ_k, which may contain fewer than $2M$ elements. We also write $\Lambda_{01} = \Lambda \cup \Lambda_1$ and $\Lambda_{01}^c = \Lambda_2 \cup \cdots \cup \Lambda_k$.

Since for any $p \in \Lambda_{j+1}$ and $p' \in \Lambda_j$, $|b[p]| \leqslant |b[p']|$, so $|b[p]| \leqslant (2M)^{-1} \|b_{\Lambda_j}\|_1$ and thus

$$\|b_{\Lambda_{j+1}}\| \leqslant (2M)^{-1/2} \|b_{\Lambda_j}\|_1.$$

Since $\|b_{\Lambda_{01}^c}\| \leqslant \sum_{j\geqslant 2} \|b_{\Lambda_j}\|$, it implies

$$\|b_{\Lambda_{01}^c}\| \leqslant (2M)^{-1/2} \sum_{j\geqslant 1} \|b_{\Lambda_j}\|_1 = (2M)^{-1/2} \|b_{\Lambda^c}\|_1. \tag{13.72}$$

Since $\|\Phi^* a - Y\| = \|W\|$, the definition of $\tilde{a}$ implies $\|\tilde{a}\|_1 \leqslant \|a\|_1$. Triangular inequalities on the restriction of $b = a - \tilde{a}$ to Λ and Λ^c give

$$\begin{cases} \|\tilde{b}_{\Lambda^c}\|_1 \leqslant \|\tilde{a}_{\Lambda^c}\|_1 + \|a_{\Lambda^c}\|_1, \\ \|\tilde{b}_{\Lambda}\|_1 \geqslant \|a_{\Lambda}\|_1 - \|\tilde{a}_{\Lambda}\|_1. \end{cases}$$

Together with $\|\tilde{a}\|_1 \leqslant \|a\|_1$, it implies that

$$\|b_{\Lambda^c}\|_1 \leqslant \|b_\Lambda\|_1 + 2\|a_{\Lambda^c}\|_1 = \|b_\Lambda\|_1 + 2\eta \quad \text{where} \quad \eta = \sum_{k>M} |f^r_{\mathcal{D}}[k]|. \tag{13.73}$$

Equations (13.72) and (13.73) lead to

$$\|b_{\Lambda^c_{01}}\| \leqslant \frac{1}{\sqrt{2M}}\|b_\Lambda\|_1 + \frac{2}{\sqrt{2M}}\eta \leqslant \frac{1}{\sqrt{2}}\|b_\Lambda\| + \frac{2}{\sqrt{2M}}\eta \leqslant \frac{1}{\sqrt{2}}\|b_{\Lambda_{01}}\| + \frac{2}{\sqrt{2M}}\eta.$$

The recovery error is thus written as a function of $\|b_{\Lambda_{01}}\|$ as follows:

$$\|a-\tilde{a}\| = \|b\| \leqslant \|b_{\Lambda_{01}}\| + \|b_{\Lambda^c_{01}}\| \leqslant \left(1+\frac{1}{\sqrt{2}}\right)\|b_{\Lambda_{01}}\| + \frac{2}{\sqrt{2M}}\eta. \tag{13.74}$$

Since $\|\Phi^*\tilde{a} - Y\| \leqslant \|W\|$, the vector b satisfies

$$\|\Phi^* b\| \leqslant \|\Phi^* a - Y\| + \|\Phi^*\tilde{a} - Y\| \leqslant 2\|W\|. \tag{13.75}$$

The reversed triangular inequality gives

$$2\|W\| \geqslant \|\Phi^* b\| \geqslant \|\Phi^* b_{\Lambda_{01}}\| - \|\Phi^* b_{\Lambda^c_{01}}\|. \tag{13.76}$$

The restricted isometry inequality (13.70) applies to b_{Λ_j} which is a vector with less than $2M$ coefficients, and $\|\Phi^* b_{\Lambda^c_{01}}\|$ is bounded with the same argument as in (13.72),

$$\|\Phi^* b_{\Lambda^c_{01}}\| \leqslant \sqrt{1+\delta_{2M}(\mathcal{D}_U)} \sum_{j\geqslant 2} \|b_{\Lambda_j}\| \leqslant \frac{\sqrt{1+\delta_{2M}(\mathcal{D}_U)}}{\sqrt{2M}}\|b_{\Lambda^c}\|_1,$$

which, together with (13.73), leads to

$$\begin{aligned} \|\Phi^* b_{\Lambda^c_{01}}\| &\leqslant \tfrac{\sqrt{1+\delta_{2M}(\mathcal{D}_U)}}{\sqrt{2M}}\left(\|b_\Lambda\|_1 + 2\eta\right) \\ &\leqslant \tfrac{\sqrt{1+\delta_{2M}(\mathcal{D}_U)}}{\sqrt{2}}\|b_\Lambda\| + 2\tfrac{\sqrt{1+\delta_{2M}(\mathcal{D}_U)}}{\sqrt{2M}}\eta. \end{aligned} \tag{13.77}$$

The restricted isometry inequality (13.70) applied to $b_{\Lambda_{01}}$, which has at most $3M$ nonzero coefficients, gives $\|\Phi^* b_{\Lambda_{01}}\| \geqslant \sqrt{1-\delta_{3M}(\mathcal{D}_U)}\,\|b_{\Lambda_{01}}\|$. Inserting this inequality and the upper bound (13.77) on $\|\Phi^* b_{\Lambda^c_{01}}\|$ in (13.76) with $\|b_\Lambda\| \leqslant \|b_{\Lambda_{01}}\|$ gives

$$2\|W\| \geqslant A\,\|b_{\Lambda_{01}}\| - 2\frac{\sqrt{1+\delta_{2M}(\mathcal{D}_U)}}{\sqrt{2M}}\eta \quad \text{where} \quad A = \sqrt{1-\delta_{3M}(\mathcal{D}_U)} - \frac{1}{\sqrt{2}}\sqrt{1+\delta_{2M}(\mathcal{D}_U)}.$$

Since $\delta_{2M}(\mathcal{D}_U) \leqslant \delta_{3M}(\mathcal{D}_U)$, the condition $\delta_{3M}(\mathcal{D}_U) < 1/3$ implies that $A > 0$, so that $\|b_{\Lambda_{01}}\|$ is bounded as follows:

$$\|b_{\Lambda_{01}}\| \leqslant 2\frac{\sqrt{1+\delta_{2M}(\mathcal{D}_U)}}{A\sqrt{2M}}\eta + \frac{2}{A}\|W\|.$$

Using this bound on $\|b_{\Lambda_{01}}\|$ in the recovery error $\|a-\tilde{a}\|$ of (13.74) gives the result of the theorem. ■

The theorem implies in particular that if $\delta_{3M}(\mathcal{D}_U) < 1/3$ and if there is no noise $W = 0$, then a signal $f = \sum_{p\in\Lambda} a[p]\, g_p$ that has $|\Lambda| \leqslant M$ nonzero coefficients is exactly recovered by a basis pursuit. However, the theorem is much more powerful since it proves stability of compressed sensing approximations by relating the error to the decay of smaller-amplitude signal coefficients. The proof does not rely on the orthogonality or the independence of vectors in $\mathcal{D}$ and thus can be extended to redundant frames. Similar results hold for modified orthogonal matching pursuit algorithms [383, 384], and other algorithmic approaches have also been developed to recover sparse Fourier expansions from few randomized point-wise evaluations [272].

When W is a Gaussian white noise of variance σ^2, following the work of Candès and Tao [143], Bickel, Ritov, and Tsybakov [113] proved that the noise term $\|W\|$ that appears in (13.71) can be reduced to its projection over the space generated by the vectors in the approximation support of $\tilde{F}$. For a dictionary of size P, they proved that under a similar M-restricted isometry condition, solving an $\mathbf{l}^1$ Lagrangian pursuit with a threshold $T = \lambda\, \sigma \sqrt{2 \log_e P}$ yields almost the same result as the model selection theorem (12.3), which minimizes an $\mathbf{l}^0$ Lagrangian. If $\tilde{\Lambda}$ is the support of the solution $\tilde{a}$, then the noise term $\|W\|$ is replaced by $|\tilde{\Lambda}|\, T$ up to a multiplicative factor.

Compressive Sensing versus Nonlinear Approximations

The nonlinear approximation error $\|f - f_M\|^2$ in a basis $\mathcal{D}$ is typically computed by assuming that the sorted coefficients have a decay $|f_{\mathcal{D}}[k]| = O(k^{-s})$ in which case Theorem 9.9 proves that $\|f - f_M\| = O(M^{-s+1/2})$. Theorem 13.9 proves that for $s > 1$, the compressive sensing error $\|\tilde{F} - f\|$ has nearly the same asymptotic decay rate.

Theorem 13.9. Suppose that $|f^r_{\mathcal{D}}[k]| = O(k^{-s})$ with $s > 1$. Let U be a Gaussian or a Bernouilli random matrix. Any vector Y of Q measurements yields an approximation

$$\tilde{F} = \sum_{p\in\Gamma} \tilde{a}[p]\, g_p \quad \text{with} \quad \tilde{a} = \underset{b\in\mathbb{C}^N}{\operatorname{argmin}} \|b\|_1 \quad \text{subject to} \quad \Big\| \sum_{p\in\Gamma} \tilde{b}[p]\, U g_p - Y \Big\| \leqslant \|W\|,$$

which satisfies

$$\|f - \tilde{F}\| = O\Big((\beta\, Q)^{1/2-s} |\log(N/Q)|^{s-1/2} + \|W\| \Big), \tag{13.78}$$

with a probability that increases exponentially to 1 with N.

Proof. Theorem 13.6 proves in (13.67) that the condition $\delta_{3M} < 1$ of Theorem 13.8 is achieved for $M \leqslant \beta Q/(\log N/Q)$ if U is a Gaussian random matrix. This result is also valid for Bernouilli matrices. Since $|f^r_{\mathcal{D}}[k]| = O(k^{-s})$, there exists $C > 0$ such that

$$\frac{1}{\sqrt{M}} \sum_{k>M} |f^r_{\mathcal{D}}[k]| \leqslant \frac{1}{\sqrt{M}} \sum_{k>M} C\, k^{-s} \leqslant \frac{C}{M^{1/2}}\, \frac{M^{1-s}}{1-s},$$

and inserting this in (13.71) implies (13.78). ■

If U computes a random sampling in a Fourier basis and if $\mathcal{D}$ is a Dirac basis, then Theorem 13.7 implies that the upper bound (13.78) is satisfied if $|\log(N/Q)|^{1/2-s}$ is replaced by $|\log N|^{5(1/2-s)}$. This theorem proves that with Q measurements a compressed sensing recovers an approximation $\tilde{F}$ with an error having the same asymptotic decays $Q^{1/2-s}$ as a nonlinear approximation, up to a logarithmic factor $\log(N/Q)$. This is a spectacular result because a standard linear approximation can have a much slower decay depending on the distribution of large coefficients. However, the applicability of this result depends on the constants that are involved, and evaluating them is a delicate topic.

Perfect Recovery Constants

Theorem 13.8 proves that an M sparse signal having only M nonzero coefficients is exactly reconstructed if $\delta_{3M}(\mathcal{D}_U) < 1/3$. Theorem 13.6 implies that this is valid if $M = \beta Q/(\log N/Q)$, but lower bounds of β computed so far are very small, which gives very pessimistic upper bounds of Q/M.

A lower bound of Q/M can be computed by looking for "bad" sparse signals f with M coefficients and that are not recovered exactly with $Q = C\,M$ measurements $Y = Uf$ without noise. For $N = 1024$ and $Q = 100$, one can find such signals with $M = 6$, which proves that $Q/M \geqslant 100/6$. This is still a large constant, which a priori should increase with N because of the $\log(N/Q)$ factor.

More encouraging evaluations of Q/M are computed "on average" with Monte-Carlo simulations. Results are slightly improved by renormalizing the vectors Ug_m of a Gaussian random matrix U in order to adjust their norm exactly to 1. A random M sparse signal F is defined with M nonzero coefficient positions that are randomly distributed in $\{0, \ldots, N-1\}$ with amplitudes that are independent Gaussian random variables of unit variance. For realizations f of F, a basis pursuit approximation $\tilde{f}$ is computed from the Q measurements $Y = Uf$ with no noise. The ratio of perfect recovery is the probability that an M sparse signal f drawn from this distribution is exactly recovered by the basis pursuit. It is evaluated numerically by Monte-Carlo sampling.

Figure 13.10 shows the recovery performance of basis pursuit for $Q = 100$ and for several values of N. For $Q = 100$ and $N = 1024$, the average perfect recovery ratio reaches 1 for $M = 13$, which is much better than the worse case $M = 6$ previously mentioned. The recovery performance deteriorates when N increases, which is consistent with Theorem 13.6. The perfect recovery ratio reaches 1 for $Q/M = 6.2$ when $N = 512$, for $Q/M = 7.7$ when $N = 1024$, for $Q/M = 11$ when $N = 2048$, and for $Q/M = 16$ when $N = 4096$. Bernouilli matrices yield nearly the same recovery ratio as Gaussian matrices. If the orthogonal basis $\mathcal{D}$ is a Dirac basis and U is a random Fourier sampling matrix, then the ratio of perfect recovery remains nearly the same. These results do not prevent us from obtaining better results in examples with large-amplitude coefficients, such as in Figure 13.9 where $M = 19$ elements are recovered from $Q = 100$ measures for $N = 1024$.

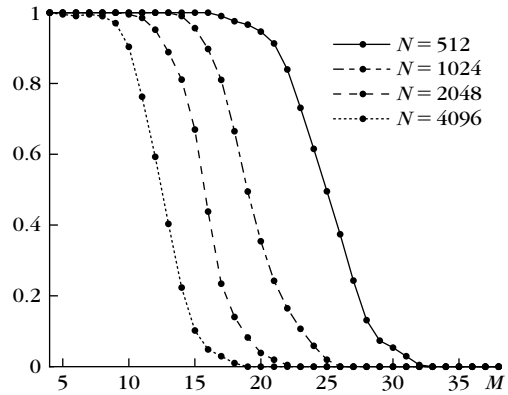


FIGURE 13.10

Perfect recovery ratio calculated with a basis pursuit, for $Q = 100$ random Gaussian measurements for several signal size values N.

Approximation Constants

Perfect recovery experiments give very partial information because most signals are not M sparse. To better understand the range of applicability of compressed sensing, we give a numerical indication of the ratio Q/M for which the compressed sensing approximation produces an error equal to an M-term nonlinear approximation. This is again performed with Monte-Carlo simulations that compute an average case over a particular signal model. It is therefore not computed from worst cases.

Large signal classes can be modeled by the decay of their sorted coefficients in an orthonormal basis $\mathcal{D}$. Bounded variation images have sorted wavelet coefficients that have a decay that is $O(k^{-s})$ for $s = 1$. For the boat, Lena, and Barbara images previously shown, the exponent is indeed close to 1. For the mandrill image in Figure 10.7, $s = 0.7$ because of the irregular textures. Let $F[n]$ be a random vector with coefficients $\{\langle F, g_p \rangle\}_{0 \leq p < N}$ in $\mathcal{D}$ that are a random permutation of the values $\{(-1)^k k^{-\alpha}\}_{0 \leq k < N}$. The ranked coefficients of F always satisfy $|F_{\mathcal{D}}^r[k]| = k^{-s}$, but there is no prior information on the location of large versus small coefficients. For all realizations, the M-term approximation error is therefore

$$\varepsilon_n[M] = \sum_{|k| > M} k^{-2s} \approx (2s-1)^{-1/2} M^{-2s+1}. \tag{13.79}$$

We study the average value of Q/M as a function of the number Q of measurements needed to have a compressed sensing error that is equal to the M term approximation error $\varepsilon_n[M]$. It compares the relative approximation efficiency of compressed sensing and nonlinear approximations.

Figure 13.11(a) gives the average value of Q/M, computed for $s = 3/2$ in a Dirac basis $\mathcal{D}$, with a normalized Gaussian random measurement matrix and a random Fourier sampling. This ratio has a small relative variation that verifies that the compressed sensing error has the same decay rate as the nonlinear approximation

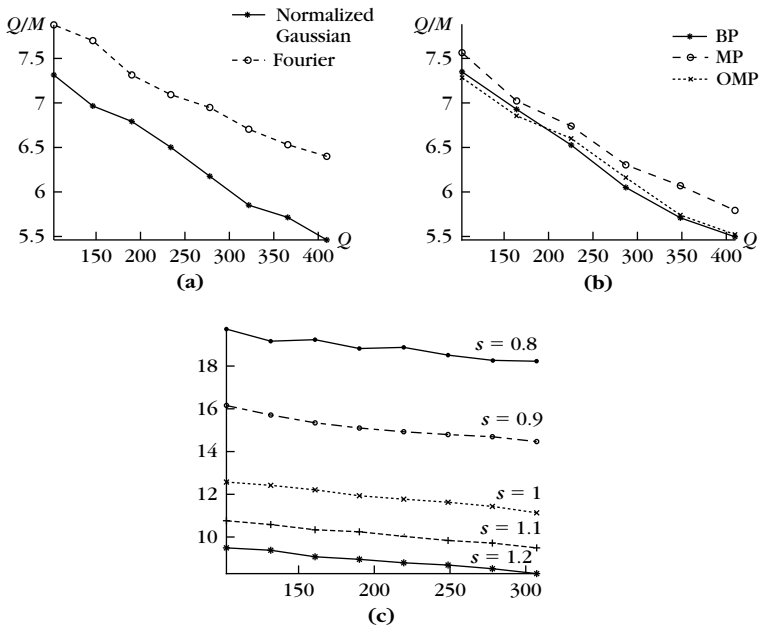


FIGURE 13.11

Ratio Q/M as a function of Q to obtain the same error with Q compressed-sensing measurements and an M-term nonlinear approximation. **(a)** Results for a normalized random Gaussian matrix and a Fourier random sampling with $N = 1024$ and $s = 3/2$. **(b)** Comparison between a matching pursuit with backprojection (MP), an orthogonal matching pursuit, and a basis pursuit with backprojection for $N = 1024$ and $s = 3/2$. **(c)** Evolution of Q/M for $N = 1024$ and s varying.

error, as predicted by Theorem 13.9. The slow decay is partially explained by the $\log(Q/N)$ factor for Gaussian random matrices. A Fourier random sampling is not universal, but in this favorable case where the signal is sparse in a Dirac basis, it gives a slightly lower Q/M ratio. For $s = 3/2$, numerical experiments for several signal sizes N verify that $Q/M \approx \beta\,(\log_2 N)$, which is coherent with the theorem statement. In these experiments, $\beta \approx 0.75$ for $s = 3/2$.

Figure 13.11(b) gives the evolution of the ratio depending on the algorithm used to compute the sparse approximation of Y in the transformed dictionary $\mathcal{D}_U$. An $\mathbf{l}^1$ Lagrangian pursuit slightly outperforms an orthogonal matching pursuit that is slightly better than a nonorthogonal matching pursuit, but in this case, the difference of efficiency between these algorithms is not so large relative to the difference of computational complexity.

Theorem 13.9 is valid only for $s > 1$. Figure 13.11(c) gives the value of Q/M when s decreases with a random Fourier sampling. Estimations are computed with an orthogonal matching pursuit. For each s the ratio remains nearly constant, but when s goes below 1 this ratio increases very significantly—up to 19 for $s = 0.8$.

Compressed Sensing versus Linear Image Approximations

Numerical experiments verify that the error of a compressed sensing has the same decay rate as a nonlinear approximation error, which is by itself remarkable since all measurements are linear. However, for images the constant factors are quite large and in the same range as standard linear approximations.

In a wavelet orthonormal basis, the sorted coefficients of most images have a decay exponent $s \leqslant 1$. For images of $N = 512^2$ pixels, the ratio Q/M is typically below 5 to obtain the same error with an M-term wavelet approximation and a linear approximation with Q low frequencies provided by a uniform sampling. This is below the compressed sensing ratio previously computed. However, Q/M does not remain constant for linear approximations, which have errors that decay more slowly than M-term approximation errors.

Compressed sensing is improved by using more prior information on the image. Donoho and Tsaig [224] use the distribution of wavelet coefficients across scales by computing a scale-by-scale compressed sensing of wavelet coefficients. The number of wavelet coefficients at a scale 2^j is proportional to 2^{-2j}, but a constant number of random measurements is performed at each scale. This is coherent with image properties, where an edge produces the same number of large-amplitude coefficients at each scale. Incorporating such prior information lowers the ratio Q/M between 4 and 5 but it depends on the image size and its content. Further reductions are possible by mixing linear and compressed sensing measurements, which can outperform linear approximations [136, 415], depending on the images.

Improving constants is a central challenge for compressive sensing applications that will influence the range of its applications. These algorithms can indeed reduce approximation errors by taking better advantage of new representations or prior information on signal coefficients [96].

13.4.3 Compressive Sensing Applications

Randomized data acquisition offers the possibility to improve the resolution of measurement devices. For analog signal acquisition, the measurement operator $\overline{U}$ provides randomized linear measurements of an analog signal $\bar{f}(x)$. One can build hardware that implements this measurement randomization. Examples are given with fully randomized sensors such as a single-pixel camera. Applications can also involve a mixture of randomized and structured acquisitions, which both play a role in the signal reconstruction.

Compressive acquisition is a democratic acquisition process where all measurements are equally important. The loss of a particular coefficient introduces an error that is diluted over the whole signal reconstruction and is thus less visible than a localized error created by a dysfunctional sensor such as a camera pixel. This robustness can also be important for signal acquisition with unreliable sensors.

Analog-to-Digital Converters

Two compressed sensing strategies are studied by Candès and Wakin [144] to improve current analog-to-digital converters. For signals that are sparse in a Fourier

basis, a uniform sampling at a Nyquist rate can be replaced by a random sampling in time. Indeed, the Fourier basis and Dirac bases have a mutual low coherence and Theorem 13.7 proves that a random time sampling yields an incoherent dictionary. If the signal is sufficiently sparse in frequency, the average sampling rate becomes lower than the uniform Nyquist rate.

Ultrawide-band signals in communication have a bandwidth that is limited by hardware analog-to-digital conversion [294]. Whereas it can be difficult to increase the sampling rate, changing the signal polarization at a very high rate may be possible. A random modulation multiplies the signal at a very high rate with $+1$ or -1 and performs an integration over a time window, which is digitized at regular time intervals. This is implemented over multiple parallel channels that modulate the signal with different random sequences of $+1$ and -1 in order to provide enough Bernouilli random measurements. This random sampling operator is universal and can thus be applied to any signal having a sparse representation in some dictionary, such as a time-frequency Gabor dictionary.

Single-Pixel Camera

Takhar et al. [453] built a compressive sensing camera that uses a single pixel photodetector to compute inner products with Q measurement vectors $\{\overline{u}_q\}_q$. A micromirror array located on the focal plane of the camera multiplies the image of the scene $\bar{f}(x)$ with a pseudo-random mask $\overline{u}_q$ that has constant values $+1$ or -1 on a regular lattice of N squares. A photoreceptor sums this randomly modulated signal, which computates $Y[q] = \langle \bar{f}, \overline{u}_q \rangle + W[q]$. It implements a random Bernouilli measurement with random signs. The Bernouilli random measurement matrix is universal and can thus take advantage of any dictionary $\mathcal{D}$ providing a sparse image representation in order to recover a high-resolution image $f \in \mathbb{C}^N$.

Tomography and MRI Imaging

Tomography imaging acquires integrals $\overline{U}\bar{f}$ of the analog signal $\bar{f}(x)$ along rays, as illustrated in Figure 13.7. It provides a subsampling of the signal Fourier transform $\widehat{\bar{f}}(\omega)$ for $\omega \in \Omega$. Angles are usually chosen to be uniformly distributed so that Ω is located along evenly distributed rays. Randomizing the ray locations can increase the incoherence of the tomography inversion problem, but the ray integration considerably limits the level of incoherence.

Medical resonance imaging (MRI) is another example of acquisition that subsamples Fourier frequencies $\widehat{\bar{f}}(\omega)$ for $\omega \in \Omega$. The excitation of atoms with a spatially varying magnetic field can select an arbitrary $\omega \in \Omega$. In contrast to tomography, Ω is therefore not restricted to be located along rays. In principle, a pseudo-random set Ω could be chosen to obtain an incoherent inverse problem. However, physical and physiological limitations enforce the magnetic field direction to follow smooth sampling curves so that Ω cannot be fully random. For applications to whole-heart coronary MRIs, Lustig et al. [356] designed a set Ω composed of spirals distributed radially with a pseudo-random density. This reduces the number of

required measurements and makes it possible to acquire an image of the entire heart in a single held breath of the patient.

Error Correction

Channel coding adds some information to a message f to build a longer encoded message Y that is robust against noisy transmissions. Classical error-correcting methods [39] consider a message f over a finite alphabet, and construct Y from f using arithmetics over finite fields. Compressed sensing provides strategies to design error-correcting codes over the real numbers.

Wyner [491] considers a real-valued message $f \in \mathbb{R}^N$, and performs the coding as $Y = Af \in \mathbb{R}^{N_0}$ where $A \in \mathbb{R}^{N_0 \times N}$ is a random matrix. The additional $Q = N_0 - N$ dimensions encode redundant information that is used to detect errors e of an unreliable transmission $Y = Af + e$.

Since most of the entries of the error $e = Y - Af$ are expected to be zero, the signal f is recovered using an $\mathbf{l}^1$ optimization

$$\tilde{f} = \underset{h \in \mathbb{R}^N}{\operatorname{argmin}} \|Y - Ah\|_1. \tag{13.80}$$

The vectors $d = Y - Ah$ is the set of vectors such that

$$Ud = U(Y - Ah) = UY,$$

where $U \in \mathbb{R}^{Q \times N_0}$ in any annihilating matrix that satisfies $\mathbf{Null}\mathbf{U} = \mathbf{Im}A$, and thus $UA = 0$. Thanks to this change of variables, the error is recovered with basis pursuit

$$\tilde{e} = \underset{d \in \mathbb{R}^{N_0}}{\operatorname{argmin}} \|d\|_1 \quad \text{subject to} \quad Ud = UY, \tag{13.81}$$

which can be solved by linear programming, as explained in Section 12.4.1. The signal f is estimated by the solution $\tilde{F}$ of the equation $A\tilde{F}f = Y - \tilde{e}$.

Donoho and Huo [230] made a first analysis of this algorithm with a mutual incoherence argument. This result was refined by Candès and Tao [138] with a compressive algorithm. If the annihilating matrix U has a small restricted isometry constant, $\delta_{3M}(U) < 1/3$, and if $\|e\|_0 \leqslant M$, then Theorem 13.8 proves that we have an exact signal recovery $\tilde{F} = f$. If U is a Gaussian random matrix, then Theorem 13.6 proves with a large probability that $\delta_{3M}(U) < 1/3$ for $M = \beta Q/(\log N_0/Q)$. If A is an encoding matrix that is annihilated by U, then the $\mathbf{l}^1$ basis pursuit (13.81) can thus recover $CQ/\log(N_0/Q)$ transmission errors.

13.5 BLIND SOURCE SEPARATION

Signals are sometimes recorded as a mixture, from which the original sources must be separated. Separating the sounds of S musical instruments recorded with K microphones is an example, with $K = 2$ for stereo recordings. Discriminating the heartbeat of a fetus from the heartbeat of his or her mother with K electrocardiogram signals is another source separation example with $S = 2$. In these *blind source separation* problems, the mixing parameters of the sources are unknown.

Linear mixtures of sources is a particularly appropriate model for sound acquisition. The K channel measurements of S sources $\{f_s\}_{0 \leq s < S}$ can then be written as

$$Y_k[n] = \sum_{s=0}^{S-1} u_{k,s} f_s[n] + W_k[n] \quad \text{for} \quad 0 \leq k < K, \tag{13.82}$$

where $U = \{u_{k,s}\}_{0 \leq k < K, 0 \leq s < S}$ is the mixing matrix and W_k are measurement noises.

The number of measurements K is often smaller than the number S of sources. Knowing the mixing matrix is then not enough to recover the sources f_s from the measurements Y_k. This source separation can be interpreted as a super-resolution problem where S sources of size N and thus SN data values must be recovered from $Q = KN < SN$ measurements. The situation looks worse than in previous inverse problems since the operator U to invert is not even known.

A successful method for source separation is based on stochastic source models, which are supposed to be independent. The pioneer work of Herault and Jutten [298] and Comon [187] has established the principles of independent component analysis for source separation. Efficient procedures such as the JADE algorithm of Cardoso [148] are separating sources by optimizing functionals that promote the independence of the recovered sources. As previously explained, it can be difficult to define a stochastic model of complex signals and thus verify that they are independent. This strong independence assumption is also not always valid, for instance, in the recording of musical instruments playing together.

Sparse blind source separation is based on weaker deterministic models. Prior information is used to define a dictionary where the different sources have a sparse representation. Jourjine, Rickard, and Yilmaz [321, 493] as well as Zibulevsky et al. [117, 498, 499] have developed algorithms that estimate the mixing operators and all the sources under the hypothesis that sources have approximation supports that do not overlap too much in an appropriate dictionary. These algorithms have then been further refined by a number of authors [103, 116, 270, 308, 343, 349, 350, 457]. The smaller the approximation support of the sources, the more likely they are to be separated. Support separation is even stronger if the dictionary takes into account differences between the sources to guarantee that the chosen dictionary vectors are different. Following a French discussion is indeed much easier in a cocktail of English speakers.

13.5.1 Blind Mixing Matrix Estimation

For blind source separation, the mixing coefficients $u_{k,s}$ are estimated by constructing a sparse representation of the multichannel measurements. As explained in Section 12.6, a whitening operator may first be applied to the measurement vectors in order to decorrelate and renormalize all channels before further processing.

Sparse Multichannel Signal Decomposition

Let us represent multichannel measurements and the mixing matrix as vectors in $\mathbb{R}^K$:

$$\vec{Y}[n] = (Y_k[n])_{0 \leq k < K}, \quad \vec{W}[n] = (W_k[n])_{0 \leq k < K}, \quad \text{and} \quad \vec{u}_s = (u_{k,s})_{0 \leq k < K}.$$

The measurement equation:

$$Y_k[n] = \sum_{s=0}^{S-1} u_{k,s} f_s[n] + W_k[n] \quad \text{for} \quad 0 \leqslant k \leqslant K \tag{13.83}$$

is rewritten as a signal vector equation:

$$\vec{Y}[n] = \sum_{s=0}^{S-1} \vec{u}_s f_s[n] + \vec{W}[n]. \tag{13.84}$$

Let $\mathcal{D} = \{\phi_p\}_{p\in\Gamma}$ be a dictionary of unit vectors in which each source f_s has a sparse approximation. The measurement vector $\vec{Y}$ also has a sparse approximation support, which is the union of the supports of all f_s. As explained in Section 12.6, the multichannel signal $\vec{Y}$ can be decomposed in $\mathcal{D}$ by calculating the inner product vectors:

$$\langle \vec{Y}, \phi_p \rangle = \Big(\langle Y_k, \phi_p \rangle \Big)_{0 \leqslant k < K} \in \mathbb{R}^K.$$

The Euclidean norm of a vector $\vec{a} \in \mathbb{R}^K$ is written as $\|\vec{a}\| = \sum_{k=0}^{K-1} |a[k]|^2$. The noise is reduced by thresholding the inner product norms $\|\langle \vec{Y}, \phi_p \rangle\|$. The resulting approximation support is

$$\tilde{\Lambda} = \Big\{ p \in \Gamma : \|\langle \vec{Y}, \phi_p \rangle\| \geqslant T \Big\}.$$

If the noise is not white, then T depends on p and is proportional to the noise variance in the direction of ϕ_p with a large multiplicative factor.

Let us write

$$\vec{b}[p] = \langle \vec{Y}, \phi_p \rangle, \quad \vec{w}[p] = \langle \vec{W}, \phi_p \rangle \quad \text{and} \quad a_s[p] = \langle f_s, \phi_p \rangle. \tag{13.85}$$

Computing the inner product of the measurement vector equation (13.83) with ϕ_p gives

$$\vec{b}[p] = \sum_{s=0}^{S-1} a_s[p]\, \vec{u}_s + \vec{w}[p] \quad \text{for} \quad p \in \tilde{\Lambda}. \tag{13.86}$$

The vectors $\vec{b}[p]$ thus define a cloud of points in $\mathbb{R}^K$ that are a combination of the S mixing vectors $\vec{u}_s$ plus a noise perturbation.

Support Separation

If the source supports are mostly disjoint, then the mixing vectors can be identified. Indeed, suppose that the supports of each $a_s[p]$ are strictly disjoint:

$$a_s[p] \neq 0 \implies a_{s'}[p] = 0 \quad \text{for} \quad s' \neq s.$$

Since for each $p \in \tilde{\Lambda}$ there exists a single s for which $a_s[p] \neq 0$, (13.86) implies

$$\vec{b}[p] = a_s[p]\, \vec{u}_s + \vec{w}[p] \in \mathbb{R}^K, \tag{13.87}$$

which provides the direction of the mixing vector up to a noise perturbation

$$\frac{\vec{b}[p]}{\|\vec{b}[p]\|} = \frac{\vec{u}_s}{\|\vec{u}_s\|}(1+\varepsilon_1) + \frac{\vec{w}[p]}{\|\vec{b}[p]\|} \quad \text{with} \quad |\varepsilon_1| \leqslant \frac{\|\vec{w}[p]\|}{\|\vec{b}[p]\|}.$$

The noise perturbation is small if we only keep coefficients $\vec{b}[p]$ with norms that are larger than a threshold that is well above the noise variance, so that

$$\frac{\|\vec{w}[p]\|}{\|\vec{b}[p]\|} \leqslant \varepsilon \ll 1 \tag{13.88}$$

with a high probability.

The strict support disjoint hypothesis is typically not satisfied. However, if a signal coefficient is much larger than the others and satisfies

$$\varepsilon\, |a_s[p]|\, \|\vec{u}_s\| \geqslant \sum_{s \neq s'} \|\vec{u}_{s'}\|\, |a_{s'}[p]|, \tag{13.89}$$

then we verify with (13.86) by inserting (13.88) and (13.89) that the normalized coefficients give the mixing direction with a small error that is of the order of ε:

$$\frac{\vec{b}[p]}{\|\vec{b}[p]\|} = \frac{\vec{u}_s}{\|\vec{u}_s\|} + \vec{\varepsilon}_2 \quad \text{with} \quad \|\vec{\varepsilon}_2\| \leqslant 3\,\varepsilon. \tag{13.90}$$

If there are enough such coefficients for all s, then the mixing direction is identified with a voting procedure using a histogram. Norm $\|\vec{u}_s\|$ is not recovered, which means that ultimately sources are computed up to a multiplicative factor. However, this global amplitude is most often an arbitrary factor that is normalized afterward.

The accuracy of this algorithm relies on the near separation of the source supports. Source supports are more likely to be separated if each f_s has a sparse approximation in $\mathcal{D}$, with a support Λ_s that has size M_s, which is small relative to N. Indeed, these sets are then unlikely to intersect often. Figure 13.12 gives a synthetic simulation illustrating the impact of sparsity. We consider three sources having $M_s = M$ nonzero coefficients in an orthonormal basis $\mathcal{D}$. In this simulation, these M coefficients are randomly distributed among the N vectors of the basis. Figure 13.12 shows that for a relatively low sparsity $M/N = 0.4$, the three mixing vector directions can still be identified.

Identification of Mixing Directions

The source directions $\vec{u}_s/\|\vec{u}_s\|$ are identified with a voting implemented by local maxima detection in a histogram of directions. The source directions belong to the unit sphere of $\mathbb{R}^K$ and are thus characterized by $K-1$ parameters that can be a vector $\vec{\theta}_s$ of $K-1$ angles. For $K=2$ measurements, there is a single angle.

According to (13.90), a normalized coefficient $\vec{b}[p]/\|\vec{b}[p]\|$ gives the direction of a dominating source $\vec{u}_s$ up to an error that depends on the noise and the relative energy of other sources in this direction. Let $\vec{\theta}(p)$ be the angle vector of $\vec{b}[p]/\|\vec{b}[p]\|$.

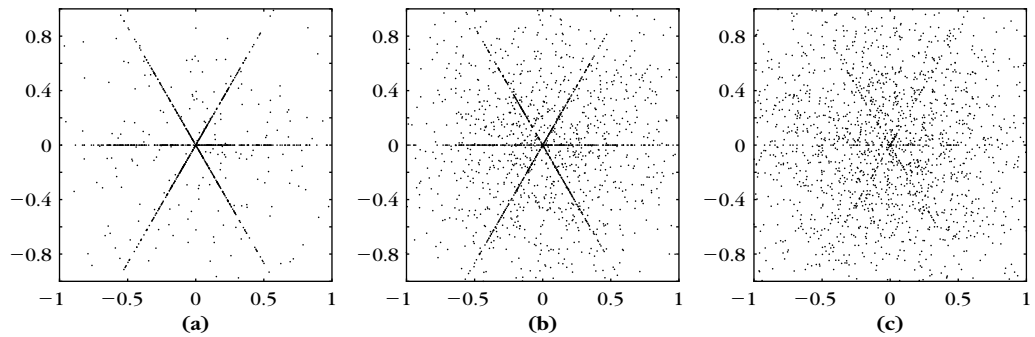


FIGURE 13.12

Distribution of synthetic coefficients $\{\vec{b}[p]\}_p$ computed with $S = 3$ sources having M nonzero coefficients each, and that are randomly distributed among the N coordinates of a basis: **(a)** $M/N = 0.1$, **(b)** $M/N = 0.4$, and **(c)** $M/N = 0.7$.

A histogram is defined over all angles with a weighting function that depends on $\|\vec{b}[p]\|$:

$$H(\vec{\theta}) = \sum_{p \in \tilde{\Lambda}} \rho(\|\vec{b}[p]\|)\, P(\vec{\theta} - \vec{\theta}[p]), \tag{13.91}$$

where $P(\vec{x})$ is a Parzen window that regularizes the histogram. The weighting function $\rho(\|\vec{b}[p]\|)$ reduces the influence of smaller-amplitude coefficients that are more affected by the noise. An appropriate weighting is

$$\rho(\|\vec{b}[p]\|) = \|\vec{b}[p]\|^2,$$

but other weighting schemes are also possible [493]. The Parzen window $P(\vec{x})$ is typically separable along the $K-1$ directions. The histogram $H(\vec{\theta})$ is sampled along a $K-1$ dimensional array at intervals proportional to the window size in each direction. The number of sources and directions of the sources are identified as local maxima of the histogram (13.91). Other classification algorithms such as K-mean algorithms may also be used to identify the mixing directions $\vec{\theta}_s$.

Figure 13.13 shows an example of stereo sound separation with $K = 2$ measurements for $S = 3$ audio sources $f_s[n]$ that are shown in the top now. The cloud of N vectors $\vec{Y}[n] \in \mathbb{R}^2$ is shown in the figure. Since the sources $f_s[n]$ are not sparse in time, this cloud has no preferential direction. In this example, the dictionary $\mathcal{D}$ is a local cosine orthonormal basis over windows of constant time duration, defined in Section 8.4.3. Figure 13.13(g) gives the cloud of local cosine coefficient vectors $\{\vec{b}[p]\}_{p \in \Gamma}$ in $\mathbb{R}^2$. This cloud is clearly elongated along 3 preferential directions corresponding to the directions of the $S = 3$ mixing vectors $\vec{u}_s$. All cosine coefficients have been kept. As a consequence, a large group of many small noisy vectors is at the center of the cloud. Figure 13.13(h) shows an angle histogram $H(\vec{\theta})$ computed with a Parzen window $P(\vec{\theta})$, which is the indicator of an interval, with

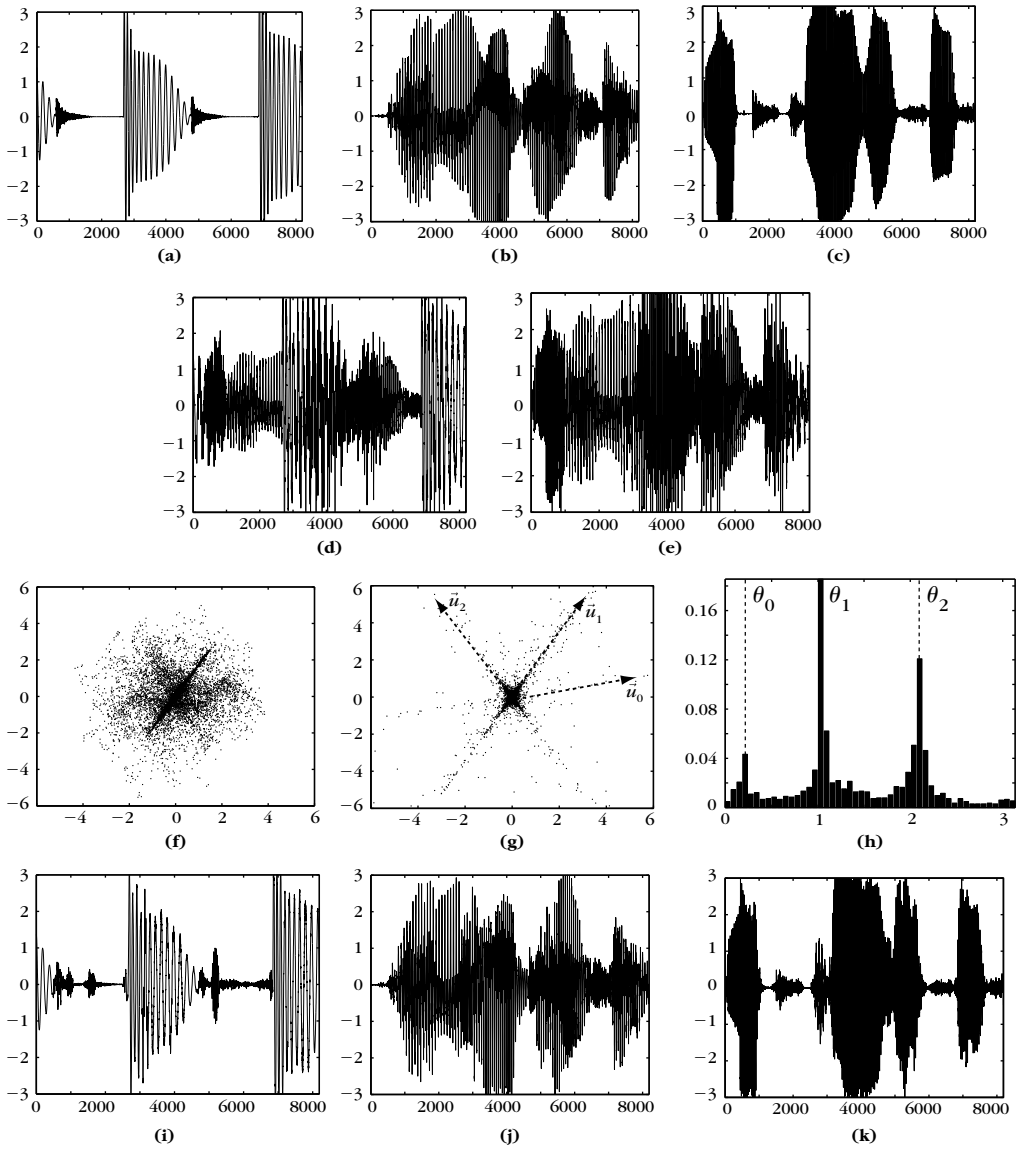


FIGURE 13.13

Example of blind-source separation of $S = 3$ sound sources from $K = 2$ measurements. **(a, b, c)** Sources of f_0, f_1, and f_2. **(d, e)** Measurements Y_0 and Y_1. **(f)** Cloud$\{\vec{Y}[n]\}_n$ in $\mathbb{R}^2$. **(g)** Cloud $\{\vec{b}[p]\}_p$ in $\mathbb{R}^2$. **(h)** Angle histogram $H(\vec{\theta})$. **(i, j, k)** Estimations $\tilde{F}_0$, $\tilde{F}_1$, and $\tilde{F}_2$.

a weighting $\rho(\|\vec{b}[p]\|)$ that is a thresholding that eliminates the smallest-amplitude coefficients and keeps the 5% vectors of largest norm $\|\vec{b}[p]\|$. This histogram exhibits three local maxima that correspond to the $S = 3$ mixing directions $\vec{\theta}_s$. The identification of mixing directions is clearly more difficult when they are close. It is then particularly important to improve the separation of the source approximation supports by decomposing the measurements over richer redundant dictionaries.

Improved Separation with Pursuits in Redundant Dictionaries

Making sure that coefficient vectors $\vec{b}[p]$ of a large norm are mostly influenced by a single source requires us to build sparse source representations or to construct a dictionary where different sources have a tendency to choose different approximation vectors. When a drum plays with a guitar, the impulsive sounds are easy to discriminate from the guitar's narrow harmonics. In a windowed Fourier dictionary or a local cosine basis where windows sizes are chosen a priori, these sounds may be difficult to discriminate. Indeed, they both occur at the same time and the impulsive sounds of the drum have a spread frequency that overlaps the guitar frequencies. To clearly separate these sounds, it is necessary to use a larger and redundant dictionary including time-frequency atoms of different scales. In the multiscale Gabor dictionary $\mathcal{D}_\Delta$ in (12.78), impulsive sounds are better represented by narrow Gabor atoms, where the guitar harmonics have a more sparse representation with elongated atoms having a better frequency resolution. Redundant dictionaries improve source separation by providing more sparse approximations and by approximating the sources with different types of atoms depending on their properties [101]. Section 12.5.3 gives an example where edges and textures are separated in an image by separating the wavelet and local cosine vectors selected in a large dictionary.

A sparse representation of $\vec{Y}$ is computed in a redundant dictionary $\mathcal{D}$ with multichannel extensions of pursuit algorithms, described in Section 12.6. We want to define a sparse representation of the measurement vector $\vec{Y}$ with coefficients $\vec{b}[p]$, which are related to the mixing directions by the same equation as (13.86):

$$\vec{b}[p] = \sum_{s=0}^{S-1} a_s[p]\, \vec{u}_s + \vec{w}[p] \quad \text{for} \quad p \in \tilde{\Lambda}. \tag{13.92}$$

For this purpose, after selecting the dictionary vectors $\{\phi_p\}_{p\in\tilde{\Lambda}}$ with a multichannel pursuit on $\vec{Y}$, pursuit coefficients are computed with an orthogonal projection. A matching pursuit or an $\mathbf{l}^1$ pursuit are therefore followed by a backprojection, described in Section 12.6. It computes the orthogonal projection of $\vec{Y}$ on the space $\mathbf{V}_{\tilde{\Lambda}}$ generated by $\{\phi_p\}_{p\in\tilde{\Lambda}}$.

The resulting coefficients are inner products with the dual frame $\{\tilde{\phi}_{p,\tilde{\Lambda}}\}_{p\in\tilde{\Lambda}}$:

$$\vec{b}[p] = \langle \vec{Y}, \tilde{\phi}_{p,\tilde{\Lambda}} \rangle, \quad \vec{w}[p] = \langle \vec{W}, \tilde{\phi}_{p,\tilde{\Lambda}} \rangle, \quad \text{and} \quad a_s[p] = \langle f_s, \tilde{\phi}_{p,\tilde{\Lambda}} \rangle \quad \text{for} \quad p \in \tilde{\Lambda}. \tag{13.93}$$

Equation (13.92) then results from the inner product of the measurement vector equation (13.84) with $\phi_{p,\tilde{\Lambda}}$. The identification of the source directions $\vec{u}_s / \|\vec{u}_s\|$ proceeds as previously described, by using the angle histogram (13.91).

13.5.2 Source Separation

Let us suppose that we know the mixing vectors $\vec{u}_s$. They are either provided by some a priori knowledge on the mixing system, or they are calculated with the identification algorithm previously described. The measurement vector $\vec{Y}[n]$ is represented by coefficient vectors $\vec{b}[p]$ that are either computed by projecting $\vec{Y}$ on dictionary elements as in (13.86) or with a pursuit in a redundant dictionary as in (13.92):

$$\vec{b}[p] = \sum_{s=0}^{S-1} a_s[p]\,\vec{u}_s + \vec{w}[p] \quad \text{for} \quad p \in \tilde{\Lambda}. \tag{13.94}$$

Cone Classification

A simple but effective masking algorithm introduced by Jourjine, Rickard, and Yilmaz [321, 493] divides the space $\mathbb{R}^K$ into cones C_s corresponding to vectors that have directions that are the closest to each mixing direction $\vec{u}_s/\|\vec{u}_s\|$,

$$C_s = \left\{ \vec{c} \in \mathbb{R}^K : s = \underset{0 \leqslant s' < S}{\operatorname{argmax}} \frac{|\langle \vec{c}, \vec{u}_{s'} \rangle|}{\|\vec{u}_{s'}\|} \right\},$$

where $\langle \vec{a}, \vec{b} \rangle$ is the usual inner product in $\mathbb{R}^K$.

Each source is estimated by projecting $\vec{b}[p]$ over the mixing direction of the cone C_{s_0} where it belongs:

$$\tilde{a}_s[p] = \begin{cases} \langle \vec{b}[p], \vec{u}_{s_0} \rangle / \|\vec{u}_{s_0}\| & \text{if} \quad s_0 = \underset{0 \leqslant s' < S}{\operatorname{argmax}} |\langle \vec{b}[p], \vec{u}_{s'} \rangle| / \|\vec{u}_{s'}\| \\ 0 \quad \text{otherwise} \end{cases}. \tag{13.95}$$

If the dictionary $\mathcal{D}$ is an orthogonal basis or a tight frame and the $\vec{b}[p]$ have been computed with a decomposition operator (13.85), then a signal estimation is recovered with

$$\tilde{F}_s = \sum_{p \in \Lambda} \tilde{a}_s[p]\,\phi_p. \tag{13.96}$$

If $\mathcal{D}$ is a redundant dictionary and the $\vec{b}[p]$ are calculated with a dual family in (13.93), then the reconstruction formula (13.96) remains valid.

As previously explained, $\tilde{F}_s$ is an estimator of f_s up to the unknown multiplicative constant $\|\vec{u}_s\|$. Figure 13.13 (i, j, k) display the $S = 3$ estimated sources $\tilde{F}_s$ computed from orthogonal local cosine coefficients $\vec{b}[n]$ with the cone classification (13.96). The source directions are calculated with the histogram shown in Figure 13.13(h).

Source Demixing with Pursuits

When the number of sources S becomes relatively large, it is more likely that a coefficient $\vec{b}[p]$ is the superposition of several nonnegligible source coefficients $a_s[p]$. As shown by Zibulevsky et al. [117, 498, 499] and analyzed by Gribonval and

Nielsen [280], these sources can still be identified with a pursuit algorithm that finds a sparse approximation of $\vec{b}[p]$ in the dictionary of normalized mixing directions $\mathcal{D}_u = \{\vec{u}_s / \|\vec{u}_s\|\}_{0 \leq s < S}$:

$$\vec{b}_{\Lambda_p}[p] = \sum_{s \in \Lambda_p} \tilde{a}_s[p] \, \frac{\vec{u}_s}{\|\vec{u}_s\|}.$$

For a fixed p, $\tilde{a}_s[p] \neq 0$ only if $s \in \Lambda_p$ and a source estimation is recovered with (13.96).

The coefficient classification (13.95) can be interpreted as a first step of a matching pursuit that projects $\vec{b}[p]$ on the direction $\vec{u}_{s_0} / \|\vec{u}_{s_0}\|$ of the best match. If the residue

$$R^1 \vec{b}[p] = \vec{b}[p] - \frac{\langle \vec{b}[p], \vec{u}_{s_0} \rangle}{\|\vec{u}_{s_0}\|}$$

is large $\|R^1 \vec{b}[p]\| \geqslant T$, then it is further decomposed by finding a next direction $\vec{u}_{s_1} / \|\vec{u}_{s_1}\|$ in the dictionary $\mathcal{D}_u$ of mixing directions, which best matches $R^1 \vec{b}[p]$, and so on. It is preferable to implement an orthogonal matching pursuit that orthogonalizes the projection directions and computes the decomposition coefficients $\tilde{a}_{s_m}[p]$ from the orthogonalized residues $R^m \vec{b}[p]$, as explained in Section 12.3.2. The iterations can be stopped with a threshold T on the norm of the residue, which is typically proportional to the noise standard deviation $(E\{\|\vec{w}[p]\|^2\})^{1/2}$.

The sparse decomposition of $\vec{b}[p]$ in the dictionary of mixing directions can also be implemented with an $\mathbf{l}^1$ pursuit that computes

$$(\tilde{a}_s[p])_{0 \leq s < S} = \underset{(a_s)_s \in \mathbb{R}^S}{\operatorname{argmin}} \frac{1}{2} \|\vec{b}[p] - \sum_{s=0}^{S-1} a_s \, \frac{\vec{u}_s}{\|\vec{u}_s\|}\|^2 + T \sum_{s=0}^{S-1} |a_s|.$$

This minimization can be solved with the iterative thresholding algorithm in Section 12.4.3.

13.6 EXERCISES

For this chapter's exercises, see the Web site at *http://wavelet-tour.com*.