Mathematical Foundations of Data Sciences

Gabriel Peyré
CNRS & DMA
École Normale Supérieure
gabriel.peyre@ens.fr
https://mathematical-tours.github.io
www.numerical-tours.com

November 15, 2020

Chapter 11

Compressed Sensing

This chapter details an important class of inverse problems, which corresponds to using "random" forward operators Φ. This is interesting from an applicative point of view since it allows to model a novel class of imaging devices which can potentially have improved resolution with respect to traditional operators (e.g. low-pass filters for usual cameras) when using in conjunction with sparse regularization technics. This is also interesting from a theoretical point of view, since the mathematical analysis becomes much simpler than with deterministic operators, and one can have good recovery and stability performances. Let us however stress that the "physical" creation of hardware that fulfils the theoretical hypothesis, in particular in medical imaging, is still largely open (put aside some restricted areas), although the theory gives many insightful design guides to improve imaging devices.

The main references for this chapter are [17, 16, 23].

11.1   Motivation and Potential Applications

11.1.1   Single Pixel Camera

In order to illustrate the exposition, we will discuss the "single pixel camera" prototype developed at Rice University [?], and which is illustrated by the figure 11.1 (left). It is an important research problem of developing a new class of cameras allowing to obtain both the sampling and the compression of the image. Instead of first sampling very finely (ie with very large Q) the analog signal $\tilde{f}$ to obtain a $f \in \mathbb{R}^Q$ image then compressing enormously (ie with M small) using (??), we would like to dispose directly of an economic representation $y \in \mathbb{R}^P$ of the image, with a budget P as close to M and such that one is able to "decompress" y to obtain a good approximation of the image f_0.

The "single-pixel" hardware performs the compressed sampling of an observed scene $\tilde{f}_0$ (the letter "R" in Figure 11.1), which is a continuous function indicating the amount of light $\tilde{f}_0(s)$ reaching each point $s \in \mathbb{R}^2$ of the focal plane of the camera. To do this, the light is focused against a set of Q micro-mirrors aligned on the focal plane. These micro-mirrors are not sensors. Unlike conventional sampling (described in Section ??), they do not record any information, but they can each be positioned to reflect or absorb light. To obtain the complete sampling/compression process, one very quickly changes P times the configurations of the micro-mirrors. For $p = 1, \ldots, P$, one sets $\Phi_{p,q} \in \{0, 1\}$, depending on whether the micromirror at position q has been placed in the absorbing (0) or reflective (value 1) position at step p of the acquisition. The total light reflected at step p is then accumulated into a single sensor (hence the name "single pixel", in fact it is rather a "single sensor"), which achieves a linear sum of the reflected intensities to obtain the recorded $y_p \in \mathbb{R}$ value. In the end, if the light intensity arriving on the surface c_q of the mirror indexed by $f_q = \int_{c_q} \tilde{f}_0(s)\mathrm{d}s$ is denoted (as in the ?? section) as q, the equation that links the discrete image $f \in \mathbb{R}^Q$

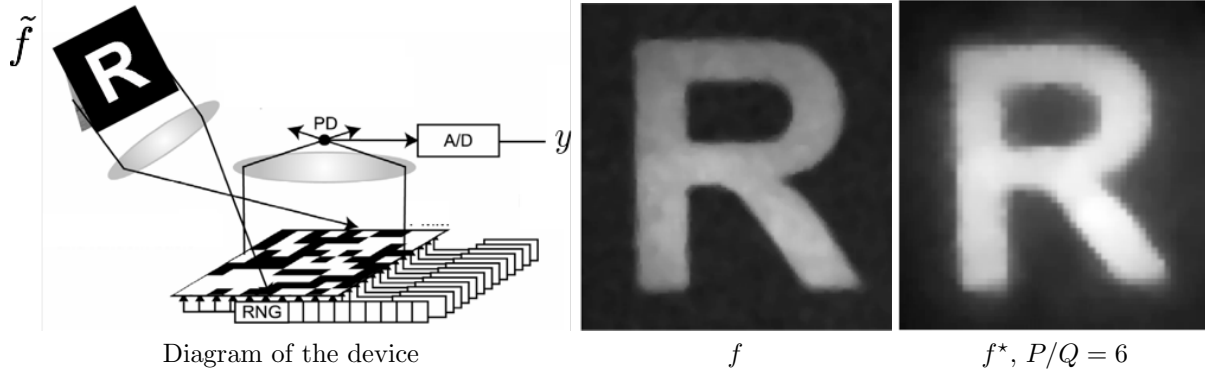| Diagram of the device | f | $f^\star$, $P/Q = 6$ |

Figure 11.1: Left: diagram of the single-pixel acquisition method. Center: image $f_0 \in \mathbb{R}^Q$ "ideal" observed in the focal plane of the micro-mirrors. Right: image $f_0^\star = \Psi x^\star$ reconstructed from observation $y \in \mathbb{R}^P$ with a compression factor $P/Q = 6$ using ℓ^1-type regularization.

"seen through the mirrors" to the P measures $y \in \mathbb{R}^P$ is

$$\forall p = 1, \ldots, P, \quad y_p \approx \sum_q \Phi_{p,n} \int_{c_n} \tilde{f}_0(s)\mathrm{d}s = (\Phi f_0)_p,$$

(here $\approx$ accounts for some noise), which corresponds to the usual forward model of inverse problems

$$y = \Phi f_0 + w \in \mathbb{R}^P$$

where w is the noise vector. It is important to note that the mirrors do not record anything, so in particular the f_0 discrete image is never calculated or recorded, since the device directly calculates the compressed representation y from the analog signal $\tilde{f}_0$. The term w models here the acquisition imperfections (measurement noise). The compressed sampling therefore corresponds to the transition from the observed scene $\tilde{f}_0$ to the compressed vector y. The "decompression" corresponds to the resolution of an inverse problem, whose goal is to find a good approximation of f_0 (the discrete image "ideal" as seen by the micro-mirrors) from y.

11.1.2 Sparse Recovery

In order to reconstruct an approximation of the (unknown) image f_0, following Section 9.2, we assume it is sparse in some dictionary Ψ. Denoting $A \stackrel{\text{def.}}{=} \Psi\Phi \in \mathbb{R}^{P \times N}$, this leads us to consider the usual ℓ^1 regularized poblem (9.10)

$$x_\lambda \in \operatorname*{argmin}_{x \in \mathbb{R}^N} \frac{1}{2\lambda}\|y - Ax\|^2 + \|x\|_1, \qquad (\mathcal{P}_\lambda(y))$$

so that the reconstructed image is $f_\lambda = \Psi x_\lambda$. We also sometimes consider the constraint problem

$$x_\varepsilon \in \operatorname*{argmin}_{\|Ax-y\| \leqslant \varepsilon} \|x\|_1, \qquad (\mathcal{P}^\varepsilon(y))$$

where, for the sake of simplicity, we set $\varepsilon = \|w\|$ (which we assume is known). From a mathematical point of view, these problem are equivalent in the sense that there exists a bijection between λ and ε which links it solution. But in practice, this bijection is not explicitly known and depends on y.

Here, it is important to remember that A is drawn from a random matrix ensemble. For an arbitrary Ψ, it is hard to analyze this random distribution. If Ψ is orthogonal, and the distribution of the columns of Φ are invariant by rotation (which is the case if the entries are i.i.d. Gaussian), then A has the same distribution as Φ. In the following, we thus directly models the distribution of A and assumes it has some nice property (typically it is close to being Gaussian i.i.d.).

178

11.2 Dual Certificate Theory and Non-Uniform Guarantees

11.2.1 Random Projection of Polytopes

When there is no noise, $w = 0$ a way to tackle the problem is to use the caracterization of solutions of $(\mathcal{P}^0(Ax_0)) = (\mathcal{P}_0(Ax_0))$ given in Section 10.1.2. According to Proposition 28, identifiable vectors with sparsity $\|x_0\|_0 = s$ corresponds to s-dimensional faces of the ℓ^1 balls B_1 which are mapped to face of the projected polytope AB_1. This leads to a combinatorial problems to count the number of face of random polytope. Donoho and Tanner were able to perform a sharp analysis of this problem. They showed the existence of two regimes, using two functions C_A, C_M so that, with high probability (i.e. a probability converging exponentially fast to 1 with (n,p)) on the matrix A

− All x_0 so that $\|x_0\|_0 \leqslant C_A(P/N)P$ are identifiable.

− Most x_0 so that $\|x_0\|_0 \leqslant C_M(P/N)P$ are identifiable.

For instance, they show that $C_A(1/4) = 0.065$ and $C_M(1/4) = 0.25$. Figure 11.5 illustrates numerically these two phase transitions. This analysis can be shown to be sharp in high dimension, i.e. when $\|x_0\|_0 > C_M(P/N)P$, then x_0 is not identifiable with high probability (this corresponds to a phase transition phenomena). For large dimensions (N, P), the scaling given by C_M describe very well what one observe in practice. For $P = N/4$ (compression of a factor 4), one retrieve in practice all vector with sparsity smaller than P/N. The function C_M can be computed numerically, and it can be shown to have a logarithmic grows $C_M(r) \sim \log(r)$ for small r. This suggests that for high compression regime, one recovers with ℓ^1 minimization almost all vector with a sparsity $\|x_0\|_0$ proportional (up to log factor) to the number of measurements P.

11.2.2 Random Matrices

The analysis of the performance ℓ^1 minimization to solve compressed sensing problem is made possible because of the very precise understanding of the distribution of the singular values of certain random ensemble let us illustrate this in the Gaussian case, which is the simplest, and is very illustrative.

An important part of the recovery proof relies on controlling the correlation matrix $A_I^*A_I$ of selected columns, and even more importantly, its inverse $(A_I A_I)^{-1}$. These random matrices are called Wishart matrices and inverse Wishart matrices. Such a matrix $B = A_I$ is of size (P, s) and is also drawn from the Gaussian ensemble. Fixing s and letting $P \to +\infty$, one has thanks to the low of large numbers $B^*B \to \mathrm{Id}_s$ almost surely. This is however not a very realistic setting, since in general, one hope to have s almost equal, up to log factor, to P.

Linear growth $P = s/\beta$. A quite extreme setting is when s grows proportionally to P, and impose $s/P = \beta$. In this case, the eigenvalues of B^*B are, for large p, essentially contained in the interval $[\lambda_-, \lambda_+]$ where $\lambda_\pm = (1 \pm \sqrt{\beta})^2$, $\beta \overset{\text{def.}}{=} s/p$, in the sense that the probability distribution of eigenvalues converges (in the weak sense of measures) toward the Marcenko-Pastur law

$$\forall (u,v) \in \mathbb{R}_+^2, \quad \mathbb{P}(\mathrm{eig}(B^\top B) \in [u,v]) \overset{p \to +\infty}{\longrightarrow} \int_u^v f_\beta(\lambda)\mathrm{d}\lambda$$

where one fix the ratio $\beta = s/P$, and the Marcenko-Pastur law is

$$f_\beta(\lambda) \overset{\text{def.}}{=} \frac{1}{2\pi\beta\lambda}\sqrt{(\lambda - \lambda_-)_+(\lambda_+ - \lambda)}1_{[\lambda_-, \lambda_+]}(\lambda).$$

Figure (11.3) illustrates this convergence.

Super-linear grows $P = s\log(\ldots)$. In order to have a better concentration of the singular values of $A_I^*A_I$ around 1, one needs to have a slightly super-linear growth of P with s. In this setting one has that $A_I^*A_I$. In order to derive non-asymptotic (i.e. with explicit constants) results, one can use a celebrated
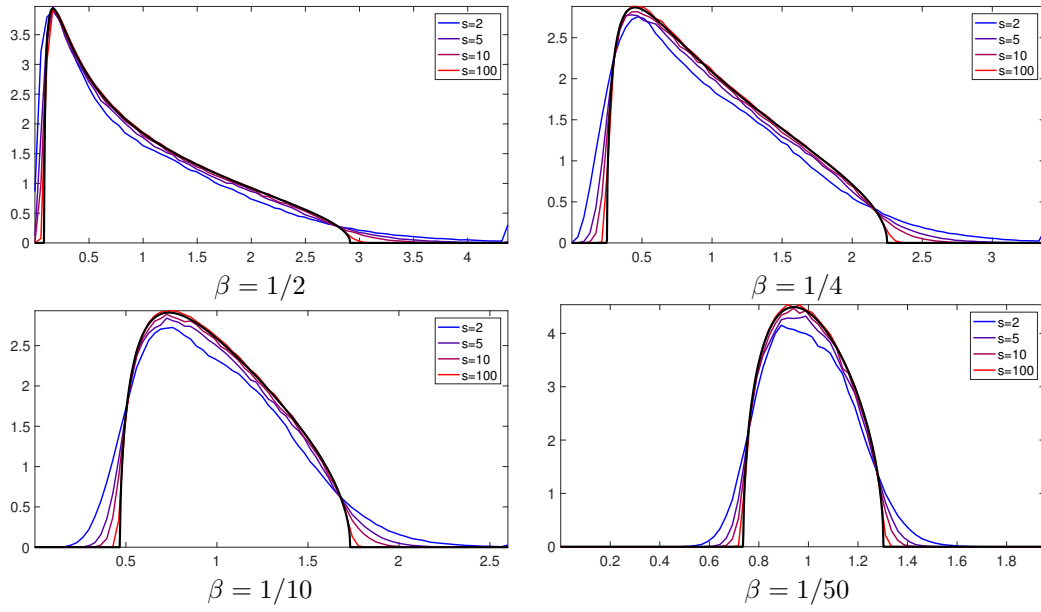
Figure 11.2:   Illustration of the convergence toward the Marcenko-Pastur law.
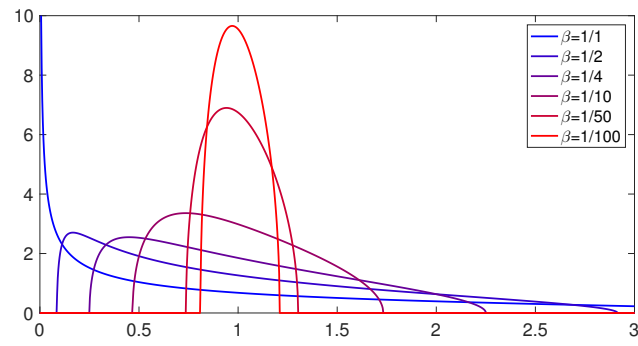

Figure 11.3:   Display of the Marcenko-Pastur distribution f_β for various β.

concentration inequality due to Talagrand, which assert that one has a fast concentration of this randomized covariance $A_I^* A_I$ toward its expectation Id_s.

$$\mathbb{P}\left(\|A_I^* A_I - \mathrm{Id}_s\|_{\mathrm{op}} \geqslant t + \sqrt{\frac{s}{P}}\right) \leqslant e^{-\frac{t^2 s}{2}}. \tag{11.1}$$

11.2.3 Dual Certificates

In order to analyze recovery performance, one can looks not only for ℓ^2 stability ($\|x_\lambda - x_0\| \sim \|w\|$) but also that x_λ has the same support as x_0 when $\|w\|$ is small. As detailed in Section 10.2.3, this requires to ensure that the pre-certificate η_F defined in (10.24) is non-degenerated, i.e.

$$\|\eta_F\|_\infty \leqslant 1 \quad \text{where} \quad \eta_F = A^* A_I (A_I^* A_I)^{-1} \operatorname{sign}(x_{0,I}). \tag{11.2}$$

Figure 10.9 suggests that this should be the case if P is large enough with respect to $\|x_0\|$. This theorem backup this observation.

Coherence-based analysis. We first perform a crude analysis using the so-called coherence of the matrix $A = (a_j)_{j=1}^N$ where the $a_j \in \mathbb{R}^P$ are the columns of A, which we assume to be normalized $\|a_j\| = 1$

$$\mu \overset{\mathrm{def.}}{=} \max_{i \neq j} |\langle a_i, a_j \rangle| = \|A^* A - \mathrm{Id}_N\|_\infty \tag{11.3}$$

where $\|C\|_\infty = \max_{i,j} |C_{i,j}|$. The coherence is 0 for an orthogonal matrix, and is always smaller than 1, $\mu \in [0, 1]$. The smaller the coherence, the better conditioned the inverse problem $Ax = y$ is, and the more likely is the certificate η_F to be non-degenerate, as shown by the following proposition.

Proposition 36. *One has, denoting $s = \|x_0\|_0 = |I|$ where $I = \operatorname{supp}(x_0)$, for $\mu < \frac{1}{s-1}$,*

$$\|\eta_{F,I^c}\|_\infty \leqslant \frac{s\mu}{1 - (s-1)\mu} \quad \text{and} \quad \|p_F\|^2 \leqslant \frac{s}{1 - (s-1)\mu}. \tag{11.4}$$

In particular, if $s < \frac{1}{2}\left(1 + \frac{1}{\mu}\right)$, $\|\eta_{F,I^c}\|_\infty < 1$ and one can thus apply the recovery Theorem 14.

Proof. We recall that the ℓ^∞ operator norm (see Remark 1) is

$$\|B\|_\infty = \max_i \sum_j |B_{i,j}|.$$

We denote $C = A^* A$. One has

$$\|A_{I^c}^* A_I\|_\infty = \max_{j \in I^c} \sum_{i \in I} C_{i,j} \leqslant s\mu \quad \text{and} \quad \|\mathrm{Id}_s - A_I^* A_I\|_\infty = \max_{j \in I} \sum_{i \in I, i \neq j} C_{i,j} \leqslant (s-1)\mu$$

One also has

$$\|(A_I^* A_I)^{-1}\|_\infty = \|((\mathrm{Id}_s - A_I^* A_I) - \mathrm{Id}_s)^{-1}\|_\infty = \|\sum_{k \geqslant 0} (\mathrm{Id}_s - A_I^* A_I)^k\|_\infty$$

$$\leqslant \sum_{k \geqslant 0} \|\mathrm{Id}_s - A_I^* A_I\|_\infty^k \leqslant \sum_{k \geqslant 0} ((s-1)\mu)^k = \frac{1}{1 - (s-1)\mu}$$

which is legit because the matrix series indeed converge since $(s-1)\mu < 1$. Using these two bounds, one has

$$\|\eta_{F,I^c}\|_\infty = \|A_{I^c}^* A_I (A_I^* A_I)^{-1} \operatorname{sign}(x_{0,I})\|_\infty \leqslant \|A_{I^c}^* A_I\|_\infty \|(A_I^* A_I)^{-1}\|_\infty \|\operatorname{sign}(x_{0,I})\|_\infty \leqslant (s\mu) \times \frac{1}{1 - (s-1)\mu} \times 1.$$

One has
$$\frac{s\mu}{1-(s-1)\mu} \iff 2s\mu < 1+\mu$$

which gives the last statement. One also has

$$\|p_F\|^2 = \langle (A_I^* A_I)^{-1} s_I, \, s_I \rangle \leqslant \|(A_I^* A_I)^{-1}\|_\infty \|s\|_1 \leqslant \frac{s}{1-(s-1)\mu}$$

<div style="text-align:right">□</div>

Note that this proof actually shows that if $s < \frac{1}{2}\left(1+\frac{1}{\mu}\right)$, all certificate η_F are valid, for any sign pattern $\mathrm{sign}(x_0)$. This actually implies a much stronger stability in the sense that whatever the noise w (which might not be small), the support of x_λ is included (not necessarily equal) in the one of x_0.

One can show that one always has

$$\mu \geqslant \sqrt{\frac{N-P}{P(N-1)}} \tag{11.5}$$

which is equivalent to $1/\sqrt{P}$ for $N \gg P$. For Gaussian matrix $A \in \mathbb{R}^{P\times N}$, one has for large $(N,P) \to +\infty$

$$\mu \sim \sqrt{\log(PN)/P}$$

which shows that Gaussian matrix are close to being optimal for the bound (11.5) if $N \gg P$. Applying Proposition 36 thus shows that ℓ^1 regularization is able to recover with a stable support vector with less than $s \sim O(\sqrt{P})$ (ignoring log terms) non-zero coefficients. In face, we will show now that it does much better and recover a proportional number $s \sim O(P)$. This is because the coherence bound (11.4) is very pessimistic.

Randomized analysis of the Fuchs certificate. We consider here a class of sufficiently "random" matrices.

Definition 1 (sub-Gaussian random matrix). *A random matrix $\sqrt{P}\,A$ is said to be sub-Gaussian if its entries are independent such that $\mathbb{E}(A_{i,j}) = 0$ (zero mean) $\mathbb{E}(A_{i,j}^2) = 1/P$ and*

$$\mathbb{P}(|\sqrt{P}\,A_{i,j}| \geqslant t) \leqslant \beta e^{-\kappa t^2}.$$

Note that its entries does not needs to be identically distributed, but the sub-Gaussiannity parameter (β,κ) should not depend on (i,j). Note also the presence of the normalization factor $\sqrt{P}$, which is here to ensure $\mathbb{E}(\|a_j\|^2) = 1$ where a_j are the columns of A.

Typical example of sub-Gaussian random matrix are Gaussian or Bernoulli matrices.

Theorem 15. *For a given $x_0 \in \mathbb{R}^N$, denoting $s = \|x_0\|_0$, and assuming A is sub-Gaussian, then for for any $0 < \varepsilon < 1$ provided that*

$$P \geqslant \frac{4c}{1-\delta} s \log(2N/\varepsilon) \quad where \quad \delta^2 \overset{\mathrm{def.}}{=} \frac{C}{4c}\left(\frac{7}{\log(2N/\varepsilon)} + \frac{2}{s}\right)$$

condition (11.2) holds with probability $1-\varepsilon$, so that x_λ has the same support and sign as x_0 when $(\|w\|, \|w\|/\lambda)$ is small enough. The constant $C, c, C = \frac{2}{3\tilde c}$ where $\tilde c \overset{\mathrm{def.}}{=} \frac{\kappa^2}{4\beta+2\kappa}$ only depends on the sub-Gaussianity parameter (β,κ) appearing (1), and for Gaussian or Benoulli matrices, $c = 1/2$.

For a Gaussian matrix, the scaling is that one should have $P \geqslant (2+\delta)s\log(N)$ with a high probability.
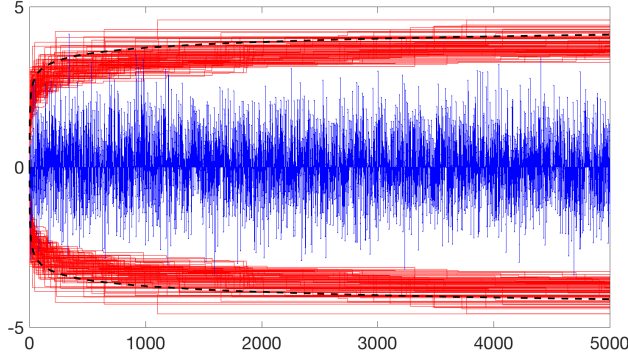
Figure 11.4: Graphical display of how the maximum of N i.i.d. gaussians concentrates tightly just below the $\sqrt{2\log(N)}$ dashed curve.

Proof. We only give the main insight for the proof. Its crux relies on the fact that $\|A_{I^c}^* p_F\|_\infty \leqslant 1$ reads

$$\max_{j \notin I} |\langle a_j, p_F \rangle| \leqslant 1$$

where $p_F = A_I^{*+} s_{0,I}$ is *independent* from the vectors $(a_j)_{j \notin I}$ it is correlated against. This allows one to check this condition by first controlling $\|p_F\|$ and then making as if p_F was a fixed deterministic vector. Noticing

$$\|p_F\|^2 = \langle A_I (A_I^* A_I)^{-1} s_{0,I}, A_I (A_I^* A_I)^{-1} s_{0,I} \rangle = \langle (A_I^* A_I)^{-1} s_{0,I}, s_{0,I} \rangle,$$

the heuristic reasoning is that, following what we said in Section (11.2.2), if P grows slightly (logarithmically) faster than s, $A_I^* A_I$ is close to Id_s (see Talagrand inequality (11.1)), so that

$$\|p_F\|^2 \sim \|s_{0,I}\|^2 = s \tag{11.6}$$

For a *fixed* p_F, one has that $\langle a_j, p_F \rangle$ is Gaussian distributed with variance $\|p_F\|^2/P$, and we use the well known fact (already encountered for denoising using thresholding) that the maximum of $P - s$ such vectors concentrates just below the universal threshold $\|p_F\|\sqrt{2\log(N - s)/P}$. Using the estimate (11.6), one sees that $\|p_F\|_\infty \leqslant 1$ is implied by $2\log(N)s/P \leqslant 1$, which gives the sharp scaling $P \geqslant 2\log(N)s$.

In order to get robustness to noise, one needs to impose that $\|A_{I^c}^* p_F\| < 1$, which can be achieve by using a slightly stronger scaling $P \geqslant 2(1 + \delta)\log(N)s$ for a small δ.

One can actually make this reasoning very precise, because quite surprisingly, it turns out that $Ps/\|p_F\|^2$ is actually distributed according to a χ^2 variable with $P - s + 1$ degrees of freedom (i.e. the sum of $P - s + 1$ squares of Gaussian variables). Indeed, for $s = 1$, one immediately sees that $P/\|p_F\|^2$ is χ^2 with P degrees of freedom. The general case is more involved, and the proof relies on the fact that the isotropy of the distribution implies that $p - P/\|p_F\|^2$ is the square of the distance between the first column of A_I and the linear space spanned by the other columns (hence $P - s + 1$ degrees of freedom). Since one has a very good understanding of the clustering of such a χ^2 variable around its means, one can thus show that $\|p_F\| \leqslant (1 - \delta)\sqrt{s}$ with high precision for arbitrary small $\delta > 0$. $\qquad \square$

For $(N, s) \to +\infty$, one has $\delta \to 0$, so that informally, the scaling is

$$P \geqslant 2s \log(2N/\varepsilon). \tag{11.7}$$

This theorem states a non-uniform recovery guarantee, in the sense that one first choose a vector x_0, then draws the matrix A, and the recovery results holds with high probability. This should be contrasted with the RIP theory developed in Section 11.3 which provides stronger uniform guarantees.
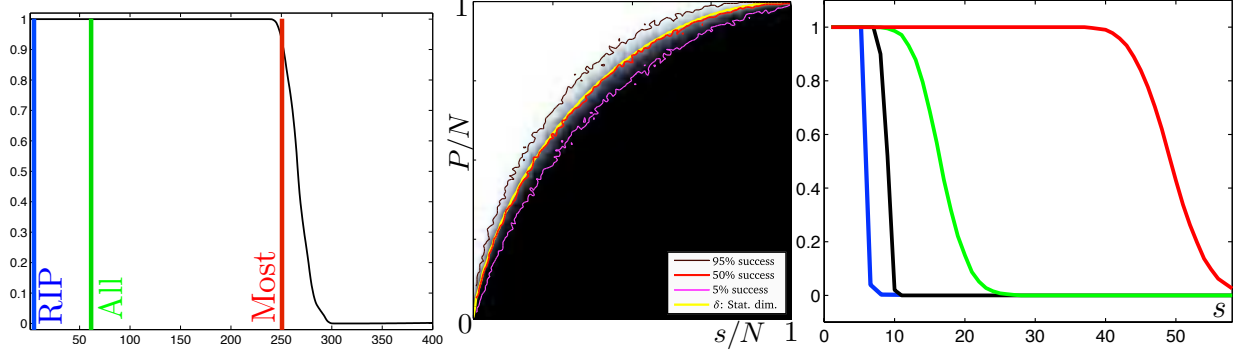
183

Figure 11.5: Phase transitions. For the figure on the right shows probability as function of sparsity that certain criteria hold true, blue: w-ERC, black: ERC, green $|\eta_F| \leqslant 1$, red: identifiability.

11.3 RIP Theory for Uniform Guarantees

11.3.1 Restricted Isometry Constants

The Restricted Isometry constant δ_s of a matrix $A \in \mathbb{R}^{P \times N}$ is defined as

$$\forall z \in \mathbb{R}^N, \quad \|z\|_0 \leqslant s \quad \Longrightarrow \quad (1 - \delta_s)\|z\|^2 \leqslant \|Az\|^2 \leqslant (1 + \delta_s)\|z\|^2, \tag{11.8}$$

and one usually chose the smallest δ_s so that these relation hold.

A related concept is the (s, s') restricted orthogonality constant $\theta_{s,s'}$, which is such that for all (x, x') with $\|x\|_0 \leqslant s$, $\|x'\|_0 \leqslant s'$ and disjoint support , one has

$$|\langle Ax,\, Ax' \rangle| \leqslant \theta_{s,s'} \|x\| \|x'\|$$

The following lemma shows that RI and RO constants are tightly related.

Lemma 5. *One has*

$$\theta_{s,s'} \leqslant \delta_{s+s'} \leqslant \theta_{s,s'} + \max(\delta_s, \delta_{s'}).$$

Proof. We prove the first inequality (which is the most important). We prove that if z and z' have disjoints supports and $\|z\| \leqslant s$ and $\|z'\|_0 \leqslant s$,

$$|\langle Az,\, Az' \rangle| \leqslant \delta_{2s} \|z\| \|z'\|.$$

Using the RIP (11.8) since $z \pm z'$ has support of size $s + s'$ and the fact that $\|z \pm z'\|^2 = \|z\|^2 + \|z'\|^2$, one has

$$(1 - \delta_{s+s'}) \left(\|z\|^2 + \|z'\|^2 \right) \leqslant \|Az \pm Az'\|^2 \leqslant (1 + \delta_{s+s'}) \left(\|z\|^2 + \|z'\|^2 \right).$$

One thus has using the parallelogram equality

$$|\langle Az,\, Az' \rangle| = \frac{1}{4}|\|Az + Az'\|^2 - \|Az - Az'\|^2| \leqslant \delta_{s+s'} \|z\| \|z'\|.$$

$\square$

The following theorem states that for a sub-Gaussian random matrix, these RIP constants grow slowly with s. Let us stress that, although this is an abuse of notation, here we assume that A is a random matrix, and not a deterministic one as previously considered.

Theorem 16. *If A is a sub-Gaussian random matrix, then provided*

$$P \geqslant C\delta^{-2} s \log(eN/s) \tag{11.9}$$

it satisfies $\delta_s \leqslant \delta$ with probability $1 - 2e^{-\delta^2 \frac{m}{2C}}$, where C only depends on the sub-Gaussianity parameters appearing in (1).
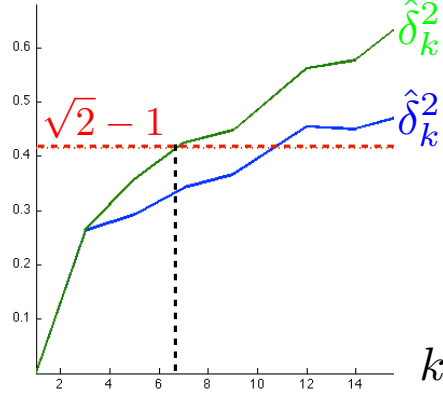
184

Figure 11.6: Left: evolution of lower bounds $\hat{\delta}_k$ on the RIP constant.

We do not prove this Theorem, and simply give the main intuition. The proof of this theorem relies on results regarding the distribution of the singular values of Gaussian matrices. Indeed, the RIP condition (11.8) is equivalent to having the bound $\operatorname{eig}(A_I^* A_I) \subset [1 - \delta_s, 1 + \delta_s]$ for all Gaussian matrices A_I extracted from A. In the Gaussian case (actually this holds for any random matrix with i.i.d. entries and unit covariance), one has a very good understanding of the distribution of the singular values of covariance matrices $B^* B \in \mathbb{R}^{s \times s}$ of a Gaussian matrix B of size (P, s), $B \sim \texttt{randn}(P, s)/\sqrt{P}$, as detailed in Section 11.2.2. In particular, using Talagrand concentration inequality (11.1), one obtains the desired controls over the δ_s constants. The intuition is that, if we assume that $s/P = \beta$ is constant and P is large, then one has that the eigenvalue of $A_I^* A_I$, and for instance its smaller one should be of the order of $2\sqrt{s/P} - s/P$, so that δ_s should be a function of s/P, and hence P should scale proportionally to the s. The log term comes from the exponential number of such matrix to control, and one need to use a non-asymptotic lower bound in place of the Marcenko-Pastur asymptotic law.

11.3.2 RIP implies dual certificates

The following theorem ensures that having small enough restricted isometry constant implies the existence of a valid dual certificates. This means that one can apply Theorem 13, which in turn ensures that one has a stable recovery of sparse signals.

Theorem 17. *If $\delta_s + \theta_{s,s} + \theta_{s,2s} < 1$, then for any x_0 with $\|x_0\| \leqslant s$, there exists $\eta \in \mathcal{D}_0(Ax_0, x_0)$, i.e. $\eta \in \operatorname{Im}(A^*) \cap \partial \|x_0\|_1$. More precisely, one has, denoting $I = \operatorname{supp}(x_0)$, $\eta_I = \operatorname{sign}(x_{0,I})$ and*

$$\|\eta_{I^c}\|_\infty \leqslant \frac{\theta_{s,s}}{1 - \delta_s - \delta_{s,2s}} < 1 \quad and \quad \|p\| \leqslant \frac{\delta_{s,s}}{1 - \delta_s - \theta_{s,2s}} \frac{\sqrt{s}}{\sqrt{1 - \delta_s}}.$$

Note that thanks to Lemma 5, condition

$$\delta_s + \theta_{s,s} + \theta_{s,2s} \leqslant \delta_s + \delta_{2s} + \delta_{3s} \leqslant 3\delta_{3s}$$

so that condition $\delta_s + \theta_{s,s} + \theta_{s,2s} < 1$ is implied by $\delta_{3s} < 1/3$. It is furthermore possible to refine the proof of this theorem to obtain alternative (often much sharper) sufficient condition such as $\delta_{2d} \leqslant \sqrt{2} - 1$. These sufficient conditions involving RI constants are often called "Restricted Isometry Properties" (RIP). As we will illustrate next, however, the constant involved on the sufficient sparsity to guarantee that such uniform RIP conditions holds are large, of the order of a few hundred.

To prove this theorem, it is not enough to directly consider the pre-certificate η_F defined in (11.2). Indeed, using such a certificate leads to slightly suboptimal log-factors in the asymptotic. The proof strategy, developed by Candes and Tao (in "Decoding with Linear Programming") consists in iteratively "improving"

185

this certificate by removing a vector interpolating the largest s entries outside I. In order to study η_F and to perform the improvement, it is crucial the behavior of least square solution of interpolation problems using random matrices.

Lemma 6. *We assume $\delta_s < 1$. Let $c \in \mathbb{R}^n$ with $\|c\|_0 \leqslant s$ and $\mathrm{supp}(c) = J$. Let $\eta = A^*p$ be the least square solution of $\eta_J = A_J^*p = c_J$, i.e.*

$$p \overset{\text{def.}}{=} A_J^{*,+}c_J = A_J(A_J^*A_J)^{-1}c_J$$

(note that A_J is full rank because $\delta_s < 1$). Then, denoting $K \subset J^c$ the s' largest entries in magnitude of η_{J^c}, one has

$$\|\eta_{(J\cup K)^c}\|_\infty \leqslant \frac{\theta_{s,s'}\|c\|}{(1-\delta_s)\sqrt{s}} \quad and \quad \|\eta_K\| \leqslant \frac{\theta_{s,s'}\|c\|}{1-\delta_s} \quad and \quad \|p\| \leqslant \frac{\|c\|}{\sqrt{1-\delta_s}}.$$

Proof. Since $|J| \leqslant s$, one has that $\lambda_{\min}(A_J^*A_J) \geqslant 1 - \delta_s$ and hence

$$\|(A_JA_J^*)^{-1}\| \leqslant \frac{1}{1-\delta_s}$$

One has

$$\|p\|^2 = \langle A_J(A_J^*A_J)^{-1}c_J, A_J(A_J^*A_J)^{-1}c_J \rangle = \langle (A_J^*A_J)^{-1}c_J, c_j \rangle \leqslant \frac{\|c\|^2}{1-\delta_s}.$$

Let L be any set with $L \cap J = \emptyset$ and $|L| \leqslant s'$. One has

$$\|\eta_L\|^2 = |\langle \eta_L, \eta_L \rangle| = |\langle A_J(A_J^*A_J)^{-1}c_J, A_L\eta_L \rangle| \leqslant \theta_{s,s'}\|(A_J^*A_J)^{-1}c_J\|\|\eta_L\| \leqslant \frac{\theta_{s,s'}}{1-\delta_s}\|c_J\|\|\eta_L\|.$$

so that this gives

$$\|\eta_L\| \leqslant \frac{\theta_{s,s'}}{1-\delta_s}\|c_J\|. \tag{11.10}$$

Let us denote $\bar{K} = \{k \in J^c \ ; \ |\eta_k| > T\}$ where $T \overset{\text{def.}}{=} \frac{\theta_{s,s'}}{(1-\delta_s)\sqrt{s'}}\|c_J\|$. One necessarily has $|\bar{K}| \leqslant s'$, otherwise one would have, taking $K \subset \bar{K}$ the s' largest entries (in fact any s' entries)

$$\|\eta_K\| > \sqrt{s'T^2} = \frac{\theta_{s,s'}}{1-\delta_s}\|c_J\|$$

which contradicts (11.10). This shows that the entries in η_{J^c} after the rank s' are smaller than T. $\square$

We can now prove the Theorem 17.

Proof. We denote $I_0 = \mathrm{supp}(x_0)$. We first consider $\eta_1 = \eta_F$, $p_1 = p_F$, and we use (6) with $c_I = \sigma_I \overset{\text{def.}}{=} \mathrm{sign}(x_{0,I})$, $J = I$, $s' = s$, to control this first pre-certificate. This lemma defines a second set I_1 (denoted K in the lemma) which are the s largest entries of η_{1,I_1^c}, with

$$I_0 \cap I_1 = \emptyset, \quad |I_1| \leqslant s, \quad \eta_{1,I_0} = s_I, \quad \|\eta_{1,(I_0\cup I_1)^c}\|_\infty \leqslant \frac{\theta_{s,s}}{1-\delta_s}, \quad \|\eta_{I_1}\| \leqslant \frac{\theta_{s,s}\sqrt{s}}{1-\delta_s}, \quad \|p_1\| \leqslant \frac{\sqrt{s}}{\sqrt{1-\delta_s}}.$$

Now we proceed recursively. Having defined the vectors $(\eta_1 = A^*p_1, \ldots, \eta_n = A^*p_n)$ with associated sets $(I_1, \ldots, I_n)$, we define $\eta_{n+1} = A^*p_{n+1}$ and I_{n+1} by applying (6) to $J = I_0 \cup I_n$ with $c = (0_{I_0}, \eta_{n,I_n})$ (and thus $(2s, s)$ in place of (s, s')), which hence satisfy

$$I_{n+1} \cap (I_0 \cup I_n) = \emptyset, \quad |I_{n+1}| \leqslant s, \quad \eta_{n+1,I_0\cup I_n} = (0_{I_0}, \eta_{n,I_n}), \quad \text{and}$$

$$\|\eta_{n+1,(I_0\cup I_n\cup I_{n+1})^c}\|_\infty \leqslant \frac{\theta_{s,s}}{1-\delta_s}Q^n, \quad \|\eta_{n+1,I_{n+1}}\| \leqslant \frac{\theta_{s,s}\sqrt{s}}{1-\delta_s}Q^n, \tag{11.11}$$

186

where we denoted $Q \stackrel{\text{def.}}{=} \frac{\theta_{s,2s}}{1-\delta_s}$, and we have

$$\|p_{n+1}\| \leqslant \frac{\|\eta_{n,I_n}\|}{\sqrt{1-\delta_s}} \leqslant \frac{1}{\sqrt{1-\delta_s}} \frac{\theta_{s,s}\sqrt{s}}{1-\delta_s} Q^{n-1}.$$

Since $\delta_s + \theta_{s,s} < 1$, one has that $Q < 1$, and thus setting

$$p = \sum_{n=1}^{+\infty} (-1)^{n-1} p_n \quad \text{and} \quad \eta = A^* p$$

defines a convergent series. By construction, since $\eta_{1,I} = \sigma$ and $\eta_{n,I} = 0$ for $n > 1$, one has $\eta_I = \sigma_I$, thus this vector interpolates the sign vector σ. Now consider $j \in I^c$ and define $E_j \stackrel{\text{def.}}{=} \{n \geqslant 1 \; ; \; j \in I_n\} = \{n_1 \leqslant n_2 \leqslant n_3 \leqslant \ldots\}$. Since $I_n \cap I_{n+1} = \emptyset$, necessary $n_{k+1} \geqslant n_k + 2$ (j cannot belong to two consecutive I_n). Furthermore, if $n \in E_j \Leftrightarrow j \in I_n$, then by construction

$$\eta_{n,j} = \eta_{n+1,j}$$

so that these two consecutive terms cancels out in the sum defining η [**ToDo: make a drawing**], which in turn can thus be written in the form

$$\eta = \sum_{n \in H} (-1)^{n-1} \eta_n.$$

The index set H is composed of $n \notin E_j$ such that $n - 1 \notin E_j$ (because otherwise one could cancel η_n from the sum). So this means that for $n \in H$, one has $j \notin (I_0 \cup I_n \cup I_{n+1})$, thus applying the property (11.11), one has

$$\forall\, n \in H, \quad |\eta_{j,n}| \leqslant \frac{\theta_{s,s}}{1-\delta_s} Q^{n-1},$$

so that

$$|\eta_j| \leqslant \sum_{n \in H} |\eta_{j,n}| \leqslant \sum_{n=1}^{+\infty} \frac{\theta_{s,s}}{1-\delta_s} Q^{n-1} = \frac{\theta_{s,s}}{1-\delta_s} \frac{1}{1-\theta_{s,2s}(1-\delta_s)^{-1}} = \frac{\theta_{s,s}}{1-\delta_s-\theta_{s,2s}}.$$

Note that one also has the bound

$$\|p\| \leqslant \sum_n \|p_n\| \leqslant \sum_n \frac{1}{\sqrt{1-\delta_s}} \frac{\theta_{s,s}\sqrt{s}}{1-\delta_s} Q^{n-1} = \frac{\delta_{s,s}}{1-\delta_s-\theta_{s,2s}} \frac{\sqrt{s}}{\sqrt{1-\delta_s}}.$$

$\square$

11.3.3 RIP implies stable recovery

Putting together Theorems 16 and **??**, and using the general inverse prolem stability theorem **??**, one obtains the following recovery guarantee.

Theorem 18. *If A is a sub-Gaussian random matrix, then there exists constants (C, C') such that provided*

$$P \geqslant Cs \log(N/s) \tag{11.12}$$

with probability $1 - 2e^{-C'P}$ on the draw of A, one has that for every s-sparse signal x_0,

$$\|x_\lambda - x_0\| = O(\|w\|)$$

where x_λ is the unique solution of (9.10) with measurements $y = Ax_0 + w$ when choosing $\lambda \sim \|w\|$.

It is possible to extend this theorem when x_0 is not exactly s-sparse but only approximately. Defining $x_{0,s}$ the best s-term approximation, the easiest way to go is to write $y = Ax_{0,s} + \tilde{w}$ where $\tilde{w} = w + A(x_0 - x_{0,s})$ and applying the previous result to obtain

$$\|x_\lambda - x_{0,s}\| = O(\|w\| + \|A\|\|x_0 - x_{0,s}\|).$$

It is possible to obtain better scaling in term of the non-linear approximation error $\|x_0 - x_{0,s}\|$ by doing a more careful proof.

This theorem provides a uniform recovery guarantee, in the sense that it means

$$\mathbb{P}(\forall\, s - \text{sparse } x_0, x^\star = x_0) \text{ goes fast to 1 when } P \to +\infty.$$

In contrast, theorem 15 proves a weaker non-uniform guarantee, in the sense that it means

$$\forall\, s - \text{sparse } x_0, \mathbb{P}(x^\star = x_0) \text{ goes fast to 1 when } P \to +\infty.$$

The recovery performance analysis based on RIP constants proves a better scaling in term of log-factors. This is because the analysis using η_F does not only imply stable recovery, it also provides stability of the support (sparsistency). Note however that the constants involved in the RIP analysis are very large (of the order of a few hundreds, as highlighted by Figure 11.6, left, and by Figure 11.5, right), while the constant appearing in (11.7) are small and known to be sharp.

Also, one can show that having small enough RIP constant implies the existence of a valid dual certificate (but this certificate is not necessarily η_F).

11.3.4 Fourier sampling RIP

A practical issue is that doing hardware implementing random operators A is very difficult, specially if this operator is "fully" random, i.e. if its entries are i.i.d. A more practical option is to use structured sampling operator, which are in some sense "less random". A possibility is to consider a random sub-sampling of orthogonal projection of the signal in some ortho-basis $\Xi = (\xi_\omega)_{\omega=1}^N$ of $\mathbb{R}^N$, so that

$$Ax \stackrel{\text{def.}}{=} (\langle x, \xi_\omega \rangle)_{\omega \in \Omega} \in \mathbb{R}^P \tag{11.13}$$

where $\Omega \subset \{1, \ldots, N\}$ is drawn uniformly at random among all sets of size P. The following theorem ensure that such an operator satisfies the RIP properties for large s (proportional to P up to log factors) is the atomes φ_ω are "spread enough", i.e. have a small magnitude, as measured by

$$\rho(\Xi) \stackrel{\text{def.}}{=} \sqrt{N} \max_{1 \leqslant \omega \leqslant N} \|\xi_\omega\|_\infty.$$

Theorem 19 (Rudelson-Vershinyn). *For any $0 < c < 1$, there exists C, such that provided that*

$$P \geqslant C\rho(\Xi)^2 s \log(N)^4 \tag{11.14}$$

with high probability on Ω, then A defined as in (11.13) satisfies $\delta_{2s} \leqslant c$.

One always has $1 \leqslant \rho(\Xi)^2 \leqslant N$. The worse case is ξ_ω to be Sirac atoms (i.e. $\Xi = \text{Id}_N$), having $\rho(\Xi)^2 = N$, in which case P needs to be as large as N. In sharp contrast, optimal sampling bases are for instance Fourier atoms $\xi_\omega = (e^{\frac{2i\pi}{N}\omega n})_{n=1}^N \in \mathbb{C}^N$, for which $\rho(\Xi) = 1$ (it is also possible to consider a Hadamard basis for instance). In this case, up to log-factors, the scaling (11.14) is similar to the one for sub-Gaussian matrices (11.12).

Theorem 19 deals with cases where the data x_0 to recover is sparse in the Dirac basis. If the data f_0 is sparse in another basis $\Psi = (\psi_m)_{m=1}^N$, one can do a change of variable $x = \Psi^* f$ (x being the coefficients of f in the basis Ψ), in which case $\rho(\Xi)$ appearing in (11.14) should be replaced by the mutual coherence between the sampling and the sparsity bases

$$\rho(\Psi^*\Xi) \stackrel{\text{def.}}{=} \sqrt{N} \max_{1 \leqslant \omega, m \leqslant N} |\langle \psi_m, \xi_\omega \rangle|.$$

Good recovery performances are thus reached by (sampling,sparsity) pairs which are incoherent. The (Fourier,Dirac) pair is maximally incoherent. In contrast, Wavelet and Fourier are highly coherent. There exists explicit construction of "noiselets" bases which are almost maximally incoherent with wavelets. Note however that in contrast to Gaussian matrices, these structured measurement matrices are not universal, in the sense that there compressed sensing recovery performances depend on the sparsity basis Ψ which is used.

Bibliography

[1] Amir Beck. *Introduction to Nonlinear Optimization: Theory, Algorithms, and Applications with MAT-LAB.* SIAM, 2014.

[2] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*, 3(1):1–122, 2011.

[3] Stephen Boyd and Lieven Vandenberghe. *Convex optimization.* Cambridge university press, 2004.

[4] E. Candès and D. Donoho. New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. *Commun. on Pure and Appl. Math.*, 57(2):219–266, 2004.

[5] E. J. Candès, L. Demanet, D. L. Donoho, and L. Ying. Fast discrete curvelet transforms. *SIAM Multiscale Modeling and Simulation*, 5:861–899, 2005.

[6] A. Chambolle. An algorithm for total variation minimization and applications. *J. Math. Imaging Vis.*, 20:89–97, 2004.

[7] Antonin Chambolle, Vicent Caselles, Daniel Cremers, Matteo Novaga, and Thomas Pock. An introduction to total variation for image analysis. *Theoretical foundations and numerical methods for sparse recovery*, 9(263-340):227, 2010.

[8] Antonin Chambolle and Thomas Pock. An introduction to continuous optimization for imaging. *Acta Numerica*, 25:161–319, 2016.

[9] S.S. Chen, D.L. Donoho, and M.A. Saunders. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, 1999.

[10] Philippe G Ciarlet. Introduction à l'analyse numérique matricielle et à l'optimisation. 1982.

[11] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *SIAM Multiscale Modeling and Simulation*, 4(4), 2005.

[12] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun. on Pure and Appl. Math.*, 57:1413–1541, 2004.

[13] D. Donoho and I. Johnstone. Ideal spatial adaptation via wavelet shrinkage. *Biometrika*, 81:425–455, Dec 1994.

[14] Heinz Werner Engl, Martin Hanke, and Andreas Neubauer. *Regularization of inverse problems*, volume 375. Springer Science & Business Media, 1996.

[15] M. Figueiredo and R. Nowak. An EM Algorithm for Wavelet-Based Image Restoration. *IEEE Trans. Image Proc.*, 12(8):906–916, 2003.

[16] Simon Foucart and Holger Rauhut. *A mathematical introduction to compressive sensing*, volume 1. Birkhäuser Basel, 2013.

[17] Stephane Mallat. *A wavelet tour of signal processing: the sparse way.* Academic press, 2008.

[18] D. Mumford and J. Shah. Optimal approximation by piecewise smooth functions and associated variational problems. *Commun. on Pure and Appl. Math.*, 42:577–685, 1989.

[19] Neal Parikh, Stephen Boyd, et al. Proximal algorithms. *Foundations and Trends® in Optimization*, 1(3):127–239, 2014.

[20] Gabriel Peyré. *L'algèbre discrète de la transformée de Fourier.* Ellipses, 2004.

[21] J. Portilla, V. Strela, M.J. Wainwright, and Simoncelli E.P. Image denoising using scale mixtures of Gaussians in the wavelet domain. *IEEE Trans. Image Proc.*, 12(11):1338–1351, November 2003.

[22] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4):259–268, 1992.

[23] Otmar Scherzer, Markus Grasmair, Harald Grossauer, Markus Haltmeier, Frank Lenzen, and L Sirovich. *Variational methods in imaging.* Springer, 2009.

[24] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948.

[25] Jean-Luc Starck, Fionn Murtagh, and Jalal Fadili. *Sparse image and signal processing: Wavelets and related geometric multiscale analysis.* Cambridge university press, 2015.