THRESHOLDING ESTIMATORS FOR MINIMAX RESTORATION AND DECONVOLUTION

Jérôme Kalifa and Stéphane Mallat

Centre de Mathématiques Appliquées
Ecole Polytechnique, Plateau de Saclay
91128 Palaiseau Cedex, France

ABSTRACT

Inverting the distortion of signals and images in presence of additive noise is often numerically unstable. To solve these ill-posed inverse problems, we study linear and non-linear diagonal estimators in an orthogonal basis. General conditions are given to build nearly minimax optimal estimators with a thresholding in an orthogonal basis. As an application, we study the deconvolution of bounded variation signals, with numerical results on the deblurring of satellite images.

1. INTRODUCTION

We consider a measurement device that degrades a signal f of size N with a linear operator U and adds a Gaussian white noise W of variance σ^2. The measured signal Y is therefore related to the original signal f following :

$$Y = Uf + W\,, \tag{1}$$

We suppose that U and σ^2 have been calculated through a calibration procedure. Applying the inverse U^{-1} to Y yields an equivalent denoising problem

$$X = U^{-1}Y = f + U^{-1}W = f + Z. \tag{2}$$

The resulting noise Z is not white but remains Gaussian because U^{-1} is linear. Its covariance operator K is

$$K = \sigma^2\, U^{-1,*}\, U^{-1}\,, \tag{3}$$

where A^* is the adjoint of an operator A. When the inverse U^{-1} is not bounded, the resulting noise $Z = U^{-1}W$ is amplified by a factor that tends to infinity. Finding an estimate $\tilde{F}$ of the signal f is an *ill-posed* inverse problem.

To build efficient estimators, we need to introduce some prior information on our signals. A Bayes estimator supposes that we know the prior probability distribution of the signals to estimate and minimizes the average estimation error. However, it is rare that we know the probability distribution of complex signals such as natural images. The prior information often defines a set Θ where the signals are guaranteed to remain, without specifying their probability distribution in Θ. Minimax estimation tries to minimize the maximum estimation error for all signals in Θ. Donoho and Johnstone have obtained general minimax optimality results to estimate signals contaminated by white Gaussian noise with thresholding estimators in orthogonal bases [DJ94]. To obtain similar results when estimating signals contaminated by non white noises, one needs to adapt the basis to the covariance properties of the noise. Section 2 shows that thresholding estimators are quasi-minimax optimal if the basis nearly diagonalizes the covariance of the noise and if it concentrates the energy of the signal on a few coefficients. As an application, we study in section 3 the deconvolution of bounded variation signals, with an application to the deblurring of satellite images.

2. MINIMAX ESTIMATION IN GAUSSIAN NOISE

We consider the generic inverse problem of equation (1), equivalent to the estimation of a signal f contaminated by an additive Gaussian noise $Z = U^{-1}W$:

$$X = f + Z\,.$$

The random vector Z is characterized by its covariance operator K, and we suppose that $\mathsf{E}\{Z[n]\} = 0$.

The risk of an estimation $\tilde{F} = DX$ is

$$r(D, f) = \mathsf{E}\{\|DX - f\|^2\}.$$

The expected risk over a set Θ cannot be computed because we do not know the probability distribution of signals in Θ. To control the risk for any $f \in \Theta$, we try to minimize the maximum risk:

$$r(D, \Theta) = \sup_{f\in\Theta} \mathsf{E}\{\|DX - f\|^2\}.$$

Let $\mathcal{O}_n$ be the set of all linear and non-linear operators from $\mathbb{C}^N$ to $\mathbb{C}^N$. The *minimax risk* is the lower bound computed over all operators D:

$$r_n(\Theta) = \inf_{D\in\mathcal{O}_n} r(D, \Theta).$$

In practice, we must find D that is simple to implement and such that $r(D, \Theta)$ is close to the minimax risk $r_n(\Theta)$. As a first step, one can simplify this problem by restricting D to be a linear operator. Let $\mathcal{O}_l$ be the set of all linear operators from $\mathbb{C}^N$ to $\mathbb{C}^N$. The *linear minimax risk* over Θ is the lower bound:

$$r_l(\Theta) = \inf_{D\in\mathcal{O}_l} r(D, \Theta)$$

We shall see when this strategy is efficient, i.e. when $r_l(\Theta)$ is of the same order as $r_n(\Theta)$.

2.1. Diagonal Estimation

When the additive noise is white, Donoho and Johnstone [DJ94] proved that non linear diagonal estimators in an orthonormal basis $\mathcal{B} = \{g_m\}_{0 \le m < N}$ are nearly minimax optimal if the basis provides a sparse signal representation, which means that the basis concentrates the energy of the signal on a few coefficients. When the noise is not white, the coefficients of the noise have a variance that depends upon each g_m:

$$\sigma_m^2 = \mathsf{E}\{|Z_{\mathcal{B}}[m]|^2\} = \langle K g_m, g_m \rangle .$$

The basis choice must therefore depend on the covariance K.

We study the risk of estimators that are diagonal in $\mathcal{B}$:

$$\tilde{F} = DX = \sum_{m=0}^{N-1} d_m(X_{\mathcal{B}}[m])\, g_m . \tag{4}$$

If $d_m(X_{\mathcal{B}}[m]) = a[m]\, X_{\mathcal{B}}[m]$, one can verify that the minimum risk $\mathsf{E}\{\|\tilde{F} - f\|^2\}$ is achieved by the following attenuation :

$$a[m] = \frac{|f_{\mathcal{B}}[m]|^2}{|f_{\mathcal{B}}[m]|^2 + \sigma_m^2}, \tag{5}$$

and

$$\mathsf{E}\{\|\tilde{F} - f\|^2\} = r_{\inf}(f) = \sum_{m=0}^{N-1} \frac{\sigma_m^2\, |f_{\mathcal{B}}[m]|^2}{\sigma_m^2 + |f_{\mathcal{B}}[m]|^2} . \tag{6}$$

Over a signal set Θ, the maximum risk of this attenuation is $r_{\inf}(\Theta) = \sup_{f\in\Theta} r_{\inf}(f)$. The attenuation (5) is called an *oracle attenuation* because it uses information normally not available, as $a[m]$ depends upon $|f_{\mathcal{B}}[m]|$ which is not known in practice. The risk $r_{\inf}(\Theta)$ is thus only a lower bound for the minimax risk of diagonal estimators. We shall see that a simple thresholding estimator has a maximum risk that is close to $r_{\inf}(\Theta)$.

A thresholding estimator is defined by

$$\tilde{F} = DX = \sum_{m=0}^{N-1} \rho_{T_m}(X_{\mathcal{B}}[m])\, g_m , \tag{7}$$

where $\rho_{T_m}(x)$ is for example a hard thresholding function

$$\rho_{T_m}(x) = \begin{cases} x & \text{if } |x| > T_m \\ 0 & \text{if } |x| \le T_m \end{cases} , \tag{8}$$

The risk of this thresholding estimator is

$$r_t(f) = r(D, f) = \sum_{m=0}^{N-1} \mathsf{E}\{|f_{\mathcal{B}}[m] - \rho_{T_m}(X_{\mathcal{B}}[m])|^2\} .$$

Donoho and Johnstone studied thresholding estimators when $T_m = \sigma_m \sqrt{2 \log_e N}$. If the signals belong to a set Θ, the threshold values are improved by considering the maximum of signal coefficients $s_{\mathcal{B}}[m] = \sup_{f\in\Theta} |f_{\mathcal{B}}[m]|$; if $s_{\mathcal{B}}[m] \le \sigma_m$ then setting $X_{\mathcal{B}}[m]$ to zero yields a risk $|f_{\mathcal{B}}[m]|^2$ that is always smaller than the risk σ_m^2 of keeping it. This is done by choosing $T_m = \infty$ to guarantee that $\rho_{T_m}(X_{\mathcal{B}}[m]) = 0$. Thresholds are thus defined by

$$T_m = \begin{cases} \sigma_m \sqrt{2 \log_e N} & \text{if } \sigma_m < s_{\mathcal{B}}[m] \\ \infty & \text{if } \sigma_m \ge s_{\mathcal{B}}[m] \end{cases} . \tag{9}$$

We shall study in which case thresholding estimators are close to minimax optimality, and compare them with linear estimators. To analyse the properties of linear and non-linear estimators, we introduce orthosymmetric sets. Θ is *orthosymmetric* in $\mathcal{B}$ if for any $f \in \Theta$ and for any $a[m]$ with $|a[m]| \le 1$ then

$$\sum_{m=0}^{N-1} a[m]\, f_{\mathcal{B}}[m]\, g_m \in \Theta .$$

This means that the set Θ is elongated along the directions of the vectors g_m of $\mathcal{B}$. The "linear vs non-linear" diagonal estimation issue depends on the size of the orthosymmetric set Θ as compared to its *quadratic convex hull*, defined as following :

The "square" of a set Θ in the basis $\mathcal{B}$ is defined by

$$(\Theta)^2_{\mathcal{B}} = \{\tilde{f} \ : \ \tilde{f} = \sum_{m=0}^{N-1} |f_{\mathcal{B}}[m]|^2\, g_m \ \text{ with } f \in \Theta\} . \tag{10}$$

We say that Θ is *quadratically convex* in $\mathcal{B}$ if $(\Theta)^2_{\mathcal{B}}$ is a convex set. The *quadratic convex hull* $\mathrm{QH}[\Theta]$ of Θ in the basis $\mathcal{B}$ is defined by

$$\mathrm{QH}[\Theta] = \left\{ f \ : \ \sum_{m=0}^{N-1} |f_{\mathcal{B}}[m]|^2 \text{ is in the convex hull of } (\Theta)^2_{\mathcal{B}} \right\} . \tag{11}$$

It is the largest set whose square $(\mathrm{QH}[\Theta])^2_{\mathcal{B}}$ is equal to the convex hull of $(\Theta)^2_{\mathcal{B}}$.

2.2. Nearly Diagonal Covariance

Donoho and Johnstone [DJ94] obtained minimax estimation results on non linear thresholding estimators when the additive noise is white. To obtain similar results when the noise Z is not white, we need to find a basis $\mathcal{B}$ that transforms the noise into "nearly" independent coefficients. This approach was studied by Donoho for some specific deconvolution problems where wavelet bases are adapted [Don95], which is not the case for *hyperbolic deconvolution* such as deblurring in section 3. We give more general conditions on the orthogonal basis $\mathcal{B}$ to obtain nearly minimax thresholding estimators [KM99].

Since the noise Z is Gaussian, the coefficients $Z_{\mathcal{B}}[m]$ are nearly independent is they are nearly uncorrelated, which means that its covariance K is nearly diagonal in $\mathcal{B}$. This approximate diagonalization is measured by preconditioning K with its diagonal. We denote by K_d the diagonal operator in the basis $\mathcal{B}$, whose diagonal is equal to the diagonal of K. The diagonal coefficients of K and K_d are thus $\sigma_m^2 = \mathsf{E}\{|Z_{\mathcal{B}}[m]|^2\}$. Let K^{-1} be the inverse of K, and $K_d^{1/2}$ be the diagonal matrix whose coefficients are σ_m. Theorem 1 computes lower and upper bounds of the minimax risks with a conditioning factor defined with the operator sup norm $\|\,.\,\|_S$.

Theorem 1 *The conditioning factor satisfies*

$$\lambda_{\mathcal{B}} = \|K_d^{1/2}\, K^{-1}\, K_d^{1/2}\|_S \ge 1.$$

If Θ is orthosymmetric in $\mathcal{B}$ then

$$\frac{1}{\lambda_{\mathcal{B}}}\, r_{\inf}(\mathrm{QH}[\Theta]) \le r_l(\Theta) \le r_{\inf}(\mathrm{QH}[\Theta]). \tag{12}$$

and

$$\frac{1}{1.25\,\lambda_{\mathcal{B}}}\,r_{\inf}(\Theta) \leq r_n(\Theta) \leq r_t(\Theta) \leq (2\log_e N+1)\left(\bar{\sigma}^2 + r_{\inf}(\Theta)\right). \tag{13}$$

One can verify that $\lambda_{\mathcal{B}} = 1$ if and only if $K = K_d$ and is thus diagonal in $\mathcal{B}$. The closer $\lambda_{\mathcal{B}}$ is to 1 the more diagonal K. The main difficulty is to find a basis $\mathcal{B}$ which nearly diagonalizes the covariance of the noise and provides sparse signal representations so that Θ is orthosymmetric or can be embedded in two close orthosymmetric sets.

If the basis $\mathcal{B}$ nearly diagonalizes K so that $\lambda_{\mathcal{B}}$ is of the order of 1 then $r_l(\Theta)$ is of the order of $r_{\inf}(\mathrm{QH}[\Theta])$, whereas $r_n(\Theta)$ and $r_t(\Theta)$ are of the order of $r_{\inf}(\Theta)$. If Θ is quadratically convex then $\Theta = \mathrm{QH}[\Theta]$ so the linear and non-linear minimax risks are close. Otherwise its quadratic hull $\mathrm{QH}[\Theta]$ may be much bigger than Θ. When Θ is strongly elongated in the directions of the basis vectors g_m, a thresholding estimation in $\mathcal{B}$ may significantly outperform an optimal linear estimation.

3. DECONVOLUTION

The restoration of signals degraded by a convolution operator U is a generic inverse problem that is often encountered in signal processing. The convolution is supposed to be circular to avoid border problems. The goal is to estimate f from

$$Y = f \circledast u + W \,.$$

The circular convolution is diagonal in the discrete Fourier basis $\mathcal{B} = \{g_k[n]\}_{0\leq k<N}$. The inverse of U is $U^{-1}f = f \circledast u^{-1}$ where the discrete Fourier transform of u^{-1} is $\widehat{u^{-1}}[k] = \frac{1}{\hat{u}[k]}$. The deconvolved data is

$$X = U^{-1}Y = Y \circledast u^{-1} = f + Z.$$

The noise $Z = U^{-1}W$ is circular stationary. Its covariance K is a circular convolution with $\sigma^2\,u^{-1} \circledast \overline{u}^{-1}$, where $\overline{u}^{-1}[n] = u^{-1}[-n]$. The Karhunen-Loève basis which diagonalizes K is therefore the discrete Fourier basis $\mathcal{B}$. The eigenvalues of K are $\sigma_k^2 = \sigma^2\,|\hat{u}[k]|^{-2}$. When $\hat{u}[k] = 0$ we formally set $\sigma_k^2 = \infty$.

When the convolution filter is a low-pass filter with a zero at high frequency, the deconvolution problem is highly unstable. Suppose that the discrete Fourier transform $\hat{u}[k]$ has a zero of order $p \geq 1$ at the highest frequency $k = \pm N/2$

$$|\hat{u}[k]| \sim \left|\frac{2k}{N} - 1\right|^p. \tag{14}$$

The noise variance σ_k^2 has a hyperbolic growth when the frequency k is in the neighborhood of $\pm N/2$. This is called a *hyperbolic deconvolution* problem of degree p.

3.1. Linear Deconvolution

In many deconvolution problems the set Θ is translation invariant, which means that if $b \in \Theta$ then any translation of b modulo N also belongs to Θ. Since the amplified noise Z is circular stationary the whole estimation problem is translation invariant. In this case, the linear estimator that achieves the minimax linear risk is diagonal in the discrete Fourier basis. It is therefore a circular convolution. In the discrete Fourier basis, the oracle risk (6) is rewritten

$$r_{\inf}(f) = \sum_{k=0}^{N-1} \frac{\sigma_k^2\,N^{-1}\,|\hat{f}[k]|^2}{\sigma_k^2 + N^{-1}\,|\hat{f}[k]|^2}\,. \tag{15}$$

We denote by $\mathrm{QH}[\Theta]$ the quadratic convex hull of Θ in the discrete Fourier basis.

Theorem 2 *Let Θ be a translation invariant set. The minimax linear risk for estimating f from $X = f + Z$ is reached by circular convolutions and*

$$r_l(\Theta) = r_{\inf}(\mathrm{QH}[\Theta])\,. \tag{16}$$

If Θ is closed and bounded, then there exists $x \in \mathrm{QH}[\Theta]$ such that $r_{\inf}(x) = r_{\inf}(\mathrm{QH}[\Theta])$. One can verify that the minimax linear estimator is $\tilde{F} = DY = d \circledast Y$, with

$$\hat{d}[k] = \frac{N^{-1}\,|\hat{x}[k]|^2\,\hat{u}^*[k]}{\sigma^2 + N^{-1}\,|\hat{x}[k]|^2\,|\hat{u}[k]|^2}\,. \tag{17}$$

If $\sigma_k^2 = \sigma^2\,|\hat{u}[k]|^{-2} \ll N^{-1}\,|\hat{x}[k]|^2$ then $\hat{d}[k] \approx \hat{u}^{-1}[k]$, but if $\sigma_k^2 \gg N^{-1}\,|\hat{x}[k]|^2$ then $\hat{d}[k] \approx 0$. The filter d is thus a regularized inverse of u.

The total variation of a discrete signal f of size N is defined with

$$\|f\|_V = \sum_{n=0}^{N-1} |f[n] - f[n-1]|\,. \tag{18}$$

The total variation measures the amplitude of all signal oscillations and is well suited to model the spatial inhomogeneity of piecewise regular signals. Bounded variation signals may include sharp transitions such as discontinuities. A set Θ_V of bounded variation signals of period N is defined by:

$$\Theta_V = \left\{ f \;:\; \|f\|_V = \sum_{n=0}^{N-1}\Big|f[n] - f[n-1]\Big| \leq C \right\}.$$

Theorem 2 can be applied to the set Θ_V which is indeed translation invariant [KM99].

Theorem 3 *For a hyperbolic deconvolution of degree p, if $1 \leq C/\sigma \leq N$ then*

$$\frac{r_l(\Theta_V)}{N\sigma^2} \sim \left(\frac{C}{N^{1/2}\,\sigma}\right)^{(2p-1)/p}. \tag{19}$$

For a constant signal to noise ratio $C^2/(N\,\sigma^2) \sim 1$, (19) implies that

$$\frac{r_l(\Theta_V)}{N\sigma^2} \sim 1\,. \tag{20}$$

Despite the fact that σ decreases and N increases the normalized linear minimax risk remains of the order of 1.

3.2. Thresholding Deconvolution

An efficient thresholding estimator is implemented in a basis $\mathcal{B}$ which defines a sparse representation of signals in Θ_V and which nearly diagonalizes K. The covariance operator K is diagonalized in the discrete Fourier basis and its eigenvalues are

$$\sigma_k^2 = \frac{\sigma^2}{|\hat{u}[k]|^2} \sim \sigma^2 \left| \frac{2k}{N} - 1 \right|^{-2p} . \qquad (21)$$

The discrete Fourier basis is not appropriate for the thresholding algorithm because it does not approximate efficiently bounded variation signals. Periodic wavelet bases provide efficient approximations of bounded variation signals, but a wavelet basis fails to approximatively diagonalize K. The discrete Fourier transforms of these wavelets have an energy mostly concentrated on dyadic intervals, as illustrated by Figure 1. On all scales but the finest, (21) shows that the eigenvalues σ_k^2 remain of the order of σ^2. These wavelets are therefore approximate eigenvectors of K. At the finest scale, the wavelets have an energy mainly concentrated in the higher frequency band $[N/4, N/2]$, where σ_k^2 varies by a huge factor of the order of N^{2r}. To construct a basis of approximate eigenvectors of K, the finest scale wavelets must be replaced by vectors that have a Fourier transform concentrated in subintervals of $[N/4, N/2]$ where σ_k^2 varies by a factor that does not grow with N. We replace the finest scale wavelets by wavelet packets [Wic94] whose discrete Fourier transform support decrease exponentially at high frequencies, while keeping a small spatial support (and hence the largest possible frequency support) to efficiently approximate piecewise regular signals. The optimal tradeoff is obtained by particular wavelet packets illustrated in figure 1, called *mirror wavelets* because of their frequency distribution symmetric with respect to wavelets. More details can be found in [KM99]. To prove that the covariance K is "almost diagonalized" in $\mathcal{B}$ for

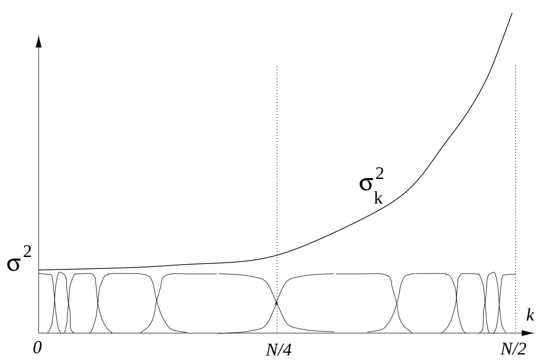


FIG. 1 – *Frequency decomposition induced by a mirror wavelet basis. The variance σ_k^2 of the noise has a hyperbolic growth but varies by a bounded factor on the frequency support of each mirror wavelet.*

all N, the asymptotic behavior of the discrete wavelets and mirror wavelets must be controlled. The following theorem thus supposes that these wavelets and wavelet packets are constructed with a conjugate mirror filter which yields a continuous time wavelet that has $q > p$ vanishing moments and which is $\mathbf{C}^q$. The near diagonalization is verified to prove that a thresholding estimator in a mirror wavelet basis has a risk whose decay is equivalent to the non-linear minimax risk.

Theorem 4 *Let $\mathcal{B}$ a mirror wavelet basis constructed with a conjugate mirror filter that defines a wavelet that is $\mathbf{C}^q$ with q vanishing moments. For a hyperbolic deconvolution of degree $p < q$, if $1 \leq C/\sigma \leq N^{p+\frac{1}{2}}$ then*

$$\frac{r_n(\Theta_V)}{N\sigma^2} \sim \frac{r_t(\Theta_V)}{N\sigma^2} \sim \left(\frac{C}{\sigma}\right)^{4p/(2p+1)} \frac{(\log_e N)^{1/(2p+1)}}{N} . \qquad (22)$$

This theorem proves that a thresholding estimator in a mirror wavelet basis yields a quasi-minimax deconvolution estimator for bounded variation signals. If we suppose that the signal to noise ratio $C^2/(N\sigma^2) \sim 1$ then

$$\frac{r_n(\Theta_V)}{N\sigma^2} \sim \frac{r_t(\Theta_V)}{N\sigma^2} \sim \left(\frac{\log_e N}{N}\right)^{1/(2p+1)} . \qquad (23)$$

As opposed to the normalized linear minimax risk (20) which remains of the order of 1, the thresholding risk in a mirror wavelet basis converges to zero as N increases. The larger the number p of zeros of the low-pass filter $\hat{u}[k]$ at $k = \pm N/2$ the slower the risk decay.

3.3. Deconvolution of Satellite Images

Nearly optimal deconvolution of bounded variation images can be calculated with a separable extension of the deconvolution estimator in a mirror wavelet basis. Such a restoration algorithm is used by the French Spatial Agency (CNES) for the production of satellite images. The satellite movement and the imperfection of the optics produces a blur, to which is added a Gaussian white noise due to the electronic of the photoreceptors. A calibration procedure measures the impulse response u of the blur and the noise variance σ^2. The image 2(b) is a simulated satellite image provided by the CNES, which is calculated from an airplane image shown in Figure 2(a). The impulse response is a separable low-pass filter

$$Uf[n_1, n_2] = f \circledast u[n_1, n_2] \text{ with } u[n_1, n_2] = u_1[n_1]\, u_2[n_2] .$$

The discrete Fourier transform of u_1 and u_2 have respectively a zero of order p_1 and p_2 at $\pm N/2$

$$\hat{u}_1[k_1] \sim \left| \frac{2k_1}{N} - 1 \right|^{p_1} \text{ and } \hat{u}_2[k_2] \sim \left| \frac{2k_2}{N} - 1 \right|^{p_2} .$$

Most satellite images are well modeled by bounded variation images. For a square discrete image of N^2 pixels, the total variation is defined by

$$\|f\|_V = \frac{1}{N} \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} \Big(\Big| f[n_1, n_2] - f[n_1 - 1, n_2] \Big|^2 + \Big| f[n_1, n_2] - f[n_1, n_2 - 1] \Big|^2 \Big)^{\frac{1}{2}} .$$

We say that an image has a bounded variation if $\|f\|_V$ is bounded by a constant independent of the resolution N. Let Θ_V be the set of images that have a total variation bounded by C

$$\Theta_V = \Big\{ f \;:\; \|f\|_V \leq C \Big\} .$$

Bounded variation plays an important role in image processing, where its value depends on the length of the image level sets.

The deconvolved noise has a covariance K that is diagonalized in a two-dimensional discrete Fourier basis. The eigenvalues are

$$\sigma^2_{k_1,k_2} = \frac{\sigma^2}{|\hat{u}_1[k_1]|^2 \, |\hat{u}_2[k_2]|^2} \sim \sigma^2 \left|\frac{2k_1}{N} - 1\right|^{-2p_1} \left|\frac{2k_2}{N} - 1\right|^{-2p_2} . \tag{24}$$

The main difficulty is again to find an orthonormal basis which provides a sparse representation of bounded variation images and which nearly diagonalizes the noise covariance K. Each vector of such a basis should have a Fourier transform whose energy is concentrated in a frequency domain where the eigenvectors $\sigma^2_{k_1,k_2}$ vary at most by a constant factor. Rougé [Rou97] has demonstrated numerically that efficient deconvolution estimations can be performed with a thresholding in a wavelet packet basis. This algorithm is inspired by his approach although the chosen basis is different.

At low frequencies $(k_1, k_2) \in [0, N/4]^2$ the eigenvalues remain approximatively constant $\sigma^2_{k_1,k_2} \sim \sigma^2$. This frequency square can be covered with a separable discrete wavelet basis. The remaining high frequency annulus is covered by two-dimensional mirror wavelets that are separable products of two one-dimensional mirror wavelets. One can verify that the union of these two families define an orthonormal basis of images of N^2 pixels. This two-dimensional mirror wavelet basis is an anisotropic wavelet packet basis, in which decomposing a signal with a filter bank requires $O(N^2)$ operations [Wic94]. One can prove that there exists λ such that $\|K_d^{1/2}\, K^{-1}\, K_d^{1/2}\|_S \leq \lambda$.

A thresholding estimator in $\mathcal{B}$ has a risk $r_t(\Theta_V)$ close to the non-linear minimax risk $r_n(\Theta_V)$ and that converges to zero as N increases, whereas a linear minimax estimator does not reduce the original noise energy $N^2\sigma^2$ by more than a constant.

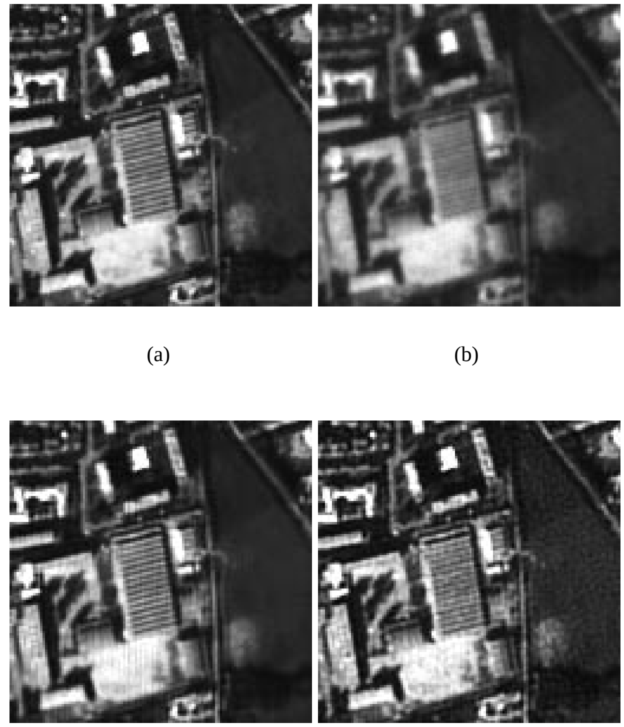
Theorem 5 *For a separable hyperbolic deconvolution of degree $p = \max(p_1, p_2) \geq 3/2$, if $C^2/(N^2\,\sigma^2) \sim 1$ then*

$$\frac{r_l(\Theta_V)}{N^2\sigma^2} \sim 1 \;\; \textit{and} \;\; \frac{r_n(\Theta_V)}{N^2\sigma^2} \sim \frac{r_t(\Theta_V)}{N^2\sigma^2} \sim \left(\frac{\log_e N}{N^2}\right)^{\frac{1}{2p+1}} .$$

Figure 2(c) shows an example of deconvolution calculated in the mirror wavelet basis. This can be compared with the linear estimation in Figure 2(d), calculated with a circular convolution estimator whose maximum risk over bounded variation images is 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 removes completely the noise in these regions while improving the restoration of edges and oscillatory parts. This algorithm was chosen among several competing algorithms by photointerpreters of the French spatial agency (CNES) to perform the deconvolution of satellite images, and it is now integrated in the CNES satellite image acquisition channel.

4. CONCLUSION

We have built a theoretical framework for minimax optimal restoration of signals and images in the case of ill-posed inverse problems. One can perform an optimal restoration if one can find an orthogonal basis which can both compress the signal to estimate on a few coefficients and nearly diagonalize the covariance of the non-white Gaussian noise obtained after applying the inverse of the degradation operator. The use of this approach to solve hyperbolic deconvolution of signals and images leads to the creation of mirror wavelet bases in which a simple thresholding procedure on the coefficients of the decomposition yields previously unobtained minimax optimality results. A competition set by the French spatial agency showed that this type of algorithms gives the best numerical results among all competing algorithms.



(a) (b) (c) (d)

FIG. 2 – *(a): Original airplane image. (b): Simulation of a satellite image (SNR = 31.1db). (c): Deconvolution with a thresholding in a mirror wavelet basis (34.1db). (d): Nearly minimax optimal linear deconvolution calculated with a circular convolution (32.7db).*

5. REFERENCES

[DJ94] D. Donoho and I. Johnstone. Ideal spatial adaptation via wavelet shrinkage. *Biometrika*, 81:425–455, December 1994.

[Don95] D. Donoho. Nonlinear solution of linear inverse problems by wavelets-vaguelette decompositions. *J. of Appl. and Comput. Harmonic Analysis*, 2(2):101–126, 1995.

[KM99] J. Kalifa and S. Mallat. *Bayesian Inference in Wavelet Based Models*, chapter ”Minimax restoration and deconvolution”. Springer-Verlag, 1999. P. Müller and B. Vidakovic, eds.

[Rou97] B. Rougé. *Théorie de la chaine image optique et restauration*. PhD thesis, Université Paris-Dauphine, 1997. Thèse d'habilitation.

[Wic94] M. V. Wickerhauser. *Adapted Wavelet Analysis from Theory to Software*. AK Peters, 1994.