SIAM J. APPL. MATH.
Vol. 57, No. 2, pp. 518–540, April 1997

012

NONLINEAR APPROXIMATION OF RANDOM FUNCTIONS*

ALBERT COHEN† AND JEAN-PIERRE D'ALES‡

Abstract. Given an orthonormal basis and a certain class X of vectors in a Hilbert space H, consider the following nonlinear approximation process: approach a vector $x \in X$ by keeping only its N largest coordinates, and let N go to infinity. In this paper, we study the accuracy of this process in the case where $H = L^2(I)$, and we use either the trigonometric system or a wavelet basis to expand this space. The class of function that we are interested in is described by a stochastic process. We focus on the case of "piecewise stationary processes" that describe functions which are smooth except at isolated points. We show that the nonlinear wavelet approximation is optimal in terms of mean square error and that this optimality is lost either by using the trigonometric system or by using any type of linear approximation method, i.e., keeping the N first coordinates. The main motivation of this work is the search for a suitable mathematical model to study the compression of images and of certain types of signals.

Key words. nonlinear approximation, stochastic process, image compression, Fourier series, wavelet bases

AMS subject classifications. 42C05, 42A10, 60G12

PII. S0036139994279153

1. Introduction.

Let x be a vector in a finite- or infinite-dimensional space H. An approximation of x is usually defined by the action of a certain operator $\mathcal{A}_\mathcal{N}$ on x such that $\mathcal{A}_\mathcal{N}x$ is close to x in some sense (for example, in the sense of a norm defined on H) and can be characterized by N parameters.

The approximation is said to be nonlinear if $\mathcal{A}_\mathcal{N}$ is not a linear operator. Example of nonlinear approximation techniques such as "free knot splines" or best approximation with rational functions have been studied for several decades. It is well known that these methods outperform linear approximation in several situations, in particular when x belongs to the unit ball X of certain Besov spaces, the error being measured with a norm for which X is a compact set (see DeVore, Howard, and Micchelli (1989), Oswald (1990)).

More recently, nonlinear approximation methods based on wavelet and wavelet packet decompositions have been studied in the context of statistics (Donoho (1995)), compression (DeVore, Jawerth, and Popov (1992)), and signal processing algorithms such as best basis selection (Coifman, Meyer, and Wickerhauser (1992)) and adaptive time-frequency decomposition (Mallat and Zhang (1993)).

In this paper, we shall be interested in the following nonlinear approximation technique.

Suppose that H is a Hilbert space and $\{e_k\}_{k>0}$ is an orthonormal basis of H. We define for all $x \in H$ and $N > 0$

$$\mathcal{A}_N x = \sum_{k \in E_N} \langle x, e_k \rangle e_k, \tag{1.1}$$

*Received by the editors December 27, 1994; accepted for publication (in revised form) February 26, 1996.
http://www.siam.org/journals/siap/57-2/27915.html
†Laboratoire d'Analyse Numérique, Université Pierre et Marie Curie, 4 Place Jussieu, 75005 Paris, France (cohen@ann.jussieu.fr).
‡CEREMADE-Université Paris-IX Dauphine, Place du Maréchal de Lattre de Tassigny, 75775 Paris cedex 16, France (ales@ceremade.dauphine.fr).

where $E_N = E_N(x)$ represents the set of indices corresponding to the N largest coordinates of x, i.e., $\mathrm{Card}(E_N) = N$ and

$$k \in E_N, l \notin E_N \Rightarrow |\langle x, e_l\rangle| \leq |\langle x, e_k\rangle|. \tag{1.2}$$

Note the difference from the standard linear approximation defined by

$$\mathcal{L}_N x = \sum_{k \leq N} \langle x, e_k\rangle e_k = \mathrm{Proj}_{\{e_1,\ldots,e_N\}} x. \tag{1.3}$$

The results of nonlinear approximation that we shall prove in this paper are related to the problem of data compression and, in particular, image compression.

Over the past twenty years, the problem of data compression has become a central issue both in information theory and in signal processing. Signals such as still or animated images, speech, and music, in their digital form, represent a considerable amount of bits to be stored, transmitted, or processed. The goal of data compression is to reduce this amount significantly while keeping the essential information on the signal that will be necessary for a given application. Note that for a continuous (or analog) signal, this original amount is infinite. In the case of digitized signals, one can define the "compression ratio" as $c = N_0/N_1$, where N_0 and N_1 represent, respectively, the number of bits in the signal before and after compression. In the case of digitized signals, one classically distinguishes two types of compression:

- Lossless compression is performed by algorithms such as Huffman or Lempel–Ziv coding: a sequence of bits is transformed into a smaller one by an inversible operation. The signal can thus be recovered without any error.
- Lossy compression is performed by any algorithm that reduces the amount of bits and allows one to recover an approximation of the original signal. In that case, there is no theoretical limit to the compression ratio which is chosen accordingly to the approximation error that is acceptable for the given application.

A general technique consists in applying some lossless compression after a lossy compression algorithm. Here, we shall only be concerned with lossy compression techniques which have important connections with nonlinear approximation. These connections have been studied by DeVore, Jawerth, and Lucier (1992) for wavelet based image compression algorithms, using Besov spaces norms to characterize the properties of real images.

In this paper, we follow a different track: the signals are described as random functions $f(t, \omega)$, and their properties are expressed in a probabilistic sense. The precision of the approximation will be measured by the mean square error

$$\varepsilon_N = E(\|f - \mathcal{A}_N f\|_2^2). \tag{1.4}$$

The introduction of randomness to describe images and the use of the mean square error reflect the situation that is mostly encountered in image compression: one expects the compression algorithm to give a good result on most of the images and allows that it may be less efficient on some pathological images that may appear. In practice, one cannot obtain a full description of the probability distribution of f, because of its complexity: for example, in the case of a digital image of size 256×256 it is a distribution in $\mathbb{R}^{2^{16}}$. However, one can easily access some partial information expressed by statistical properties. We shall use this information to study ε_N for two different choices of basis: the trigonometric system and wavelet bases.

Note that choice of an L_2 error criterion is quite arbitrary. In the case of images it does not seem to fit with the measurement of the error that is done by the visual system. Although other norms have been proposed, we shall keep the L_2 norm, remarking that in the case of wavelet decompositions all the error computations that we shall present can be generalized to other norms, using the property that wavelet bases are unconditional bases for most function spaces. A more general remark is that it is not clear that the eye, at a low-level vision stage, uses a norm to measure the error. A norm that fits ideally with the eye's measurement of the error should both take into account the regularity of the image in the regions representing smooth objects and the existence of discontinuities representing the edges of these objects. We show in section 4 that the query for such a norm is faced with a completion problem.

Over the past decade, multiscale methods (and in particular wavelet bases) have become popular in image processing, because they yield sparse decompositions of the function $I(x, y)$ representing the light intensity of the image. Here "sparse" means that a small number of coefficients carry most of the significant information. As we shall see, this property can be described more precisely in terms of nonlinear approximation. However, compressing an image is a more complex process than approximating it with a small number of coefficients, since it also involves the quantization of these coefficients, i.e., their encoding on a finite number of bits. We shall devote the next section to a qualitative analysis of the relations between data compression and nonlinear approximation.

The rest of our paper is organized as follows: in section 3, we review linear approximation results that were obtained by Cohen, Froment, and Istas (1991), in the case where the only information available is the first- and second-order moments of the distribution. These results are based on a nonperiodic stationary statistical model for the signal. In that case the performance of Fourier series and wavelet bases on images are comparable. A more sophisticated model is proposed in section 4 to describe images. This model leads to nonlinear approximation results that are detailed in section 5. In that case, we show that a well-chosen wavelet basis outperforms Fourier series and any type of linear approximation.

For sake of simplicity, we shall consider unidimensional signals, described by a stochastic process $f(t, \omega)$, to establish our results, and we shall discuss how these results can be generalized to multidimensional signals and applied to images.

2. Nonlinear approximation and data compression. The goal of this section is to investigate, mostly in a qualitative way, the relations between data compression and nonlinear approximation.

The problem of data compression can be stated in the following general terms: a set $\mathcal{C}$ of vectors in a space V is given. This set may be finite or infinite as well as the dimension of V. For a fixed N, one tries to approximate $\mathcal{C}$ by a finite set $\mathcal{C}_N$ of cardinal N, in the sense that a "representative" $\rho(x) \in \mathcal{C}_N$ can be associated with every $x \in \mathcal{C}$. Each vector of $\mathcal{C}$ is thus "quantized" on N levels, and if $N = 2^b$, it can be encoded with b bits.

As we pointed out, in many practical situations, the set $\mathcal{C}$ is unknown and described by a probability distribution $p(x)$. The quality of the compression can be measured by the mean square error

$$\varepsilon(\mathcal{C}_N) = E(\|x - \rho(x)\|^2). \tag{2.1}$$

In many applications where the compression rate is preassigned, one looks for a set C_N that minimizes the above quantity for a given N. The complexity of this research essentially depends on the dimension of V.

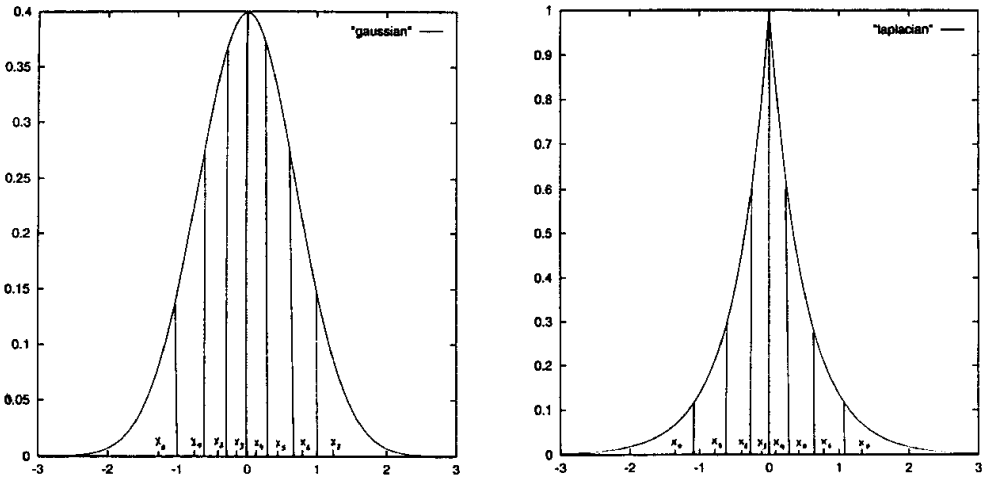

FIG. 1. *Optimal quantization in the* 1*D case.*

When dimV = 1, the situation is quite simple: x is a scalar random variable with probability law $p(x)$ (we assume here $E(|x|^2) < +\infty$), and one looks for a collection $\{x_k\}_{k=1,\dots,N}$ and a covering of $\mathbb{R}$ by disjoint intervals I_k, $x_k \in I_k$, that minimize the quantity

$$\varepsilon = \sum_{k=1}^{N} \int_{J_k} |x - y_k|^2 p(x) dx \tag{2.2}$$

over all such covering $\{J_k\}$ and collections $\{y_k\}$ of points. Examples of optimal solutions are represented in Figure 1.

When x is a vector random variable of finite dimension d, the optimal solution is given by a collection $\{x_k\}_{k=1,\dots,N}$ and a covering of $\mathbb{R}^d$ by disjoint connected sets I_k that minimize a quantity similar to (2.2).

In practice, if d is small (less than 10), these optimal solutions can be approached very efficiently by "vector quantizer design" algorithms (Linde, Buzo, and Gray (1980)), provided that $p(x)$ is regular enough that it can be estimated from a reasonable number of independent realizations of x.

Unfortunately, as d increases, this task becomes unrealistic for several reasons:

- The number of independent realizations of x that is necessary to estimate $p(x)$ grows exponentially with d.
- The shape of the optimal sets I_k may become complicated, as may the distribution of the $x_k \in \mathcal{C}_N$.
- Because of the complexity of the covering, the quantization $x \to \rho(x)$ may be difficult to execute numerically.

For vectors of very large dimension (such as images), of course it is possible to divide the vector into smaller blocks of coordinates and apply vector quantization independently on each of these blocks. This technique, however, generates "blocking effects" in images. Moreover, it is far from optimal in many situations.

Consider, for example, the simple case of a random bidimensional vector $z = (x, y)$ and suppose that one quantizes independently its coordinates x and y. The resulting two-dimensional (2D) quantizer is given by the tensor product of the optimal one-

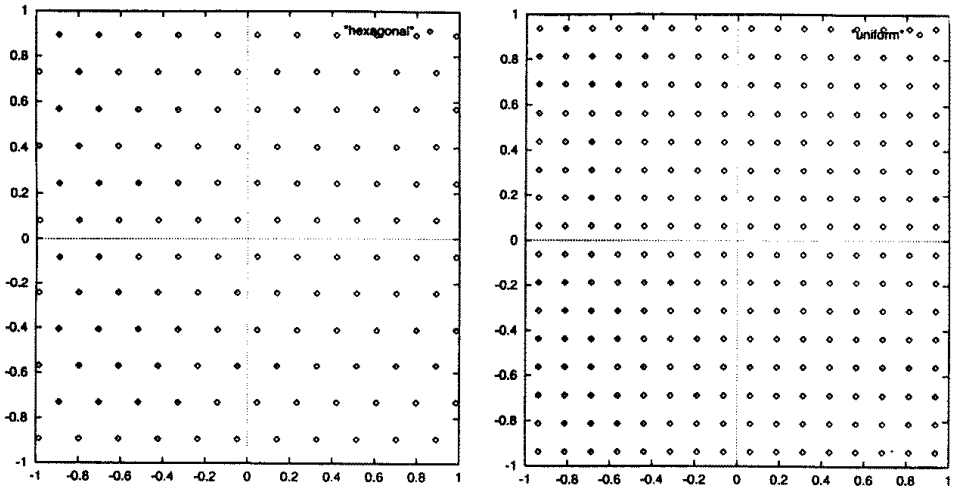

FIG. 2. *Optimal* (a) *and near optimal* (b) *quantization in the uniform* 2*D case.*

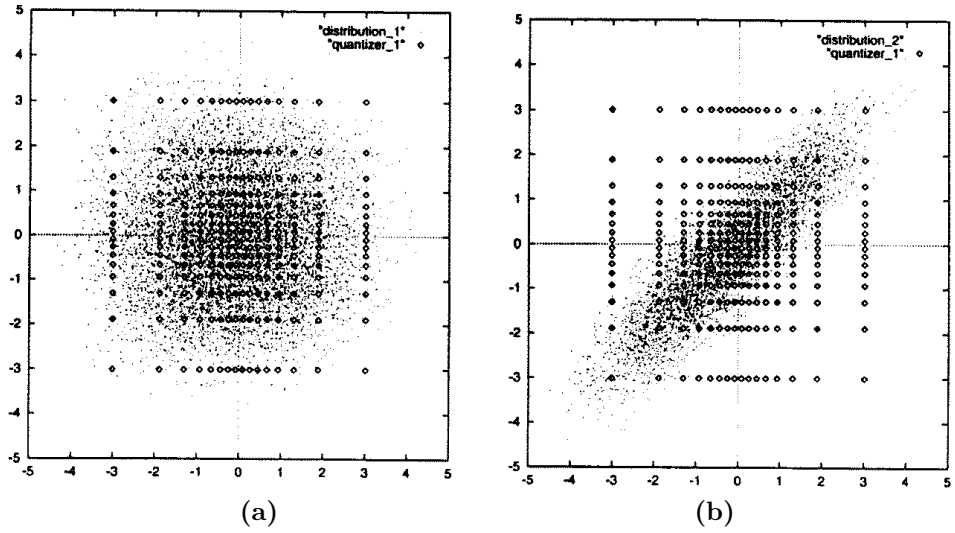

FIG. 3. *Scalar quantization of a separable* (a) *and a nonseparable* (b) *distribution.*

dimensional (1D) solutions associated with the marginal laws $p(x)$ and $p(y)$, i.e., a separable grid $z_{k,l} = (x_k, y_l)$, $k = 1, \ldots, N_x$, $l = 1, \ldots, N_y$, and the associated rectangular regions $R_{k,l} = I_k \times I_l$. The efficiency of this strategy highly depends on the property of the joint probability distribution $p(z) = p(x, y)$.

Figure 2 displays the optimal solution in the case of a uniform distribution: it is given by the hexagonal mesh that outperforms the square grid. The potential gain of using a mesh that differs from the square grid is small in this particular case, and more generally, it reaches a limit value as the dimension of V grows to $+\infty$.

More generally, when the distribution $p(x, y)$ is separable, i.e., x and y are independent variables, the tensor product strategy (or scalar quantization) leads to a good approximation of the optimal solution (Figure 3(a)). In contrast, when x and y are not independent, it is generally a bad solution, as shown in Figure 3(b).

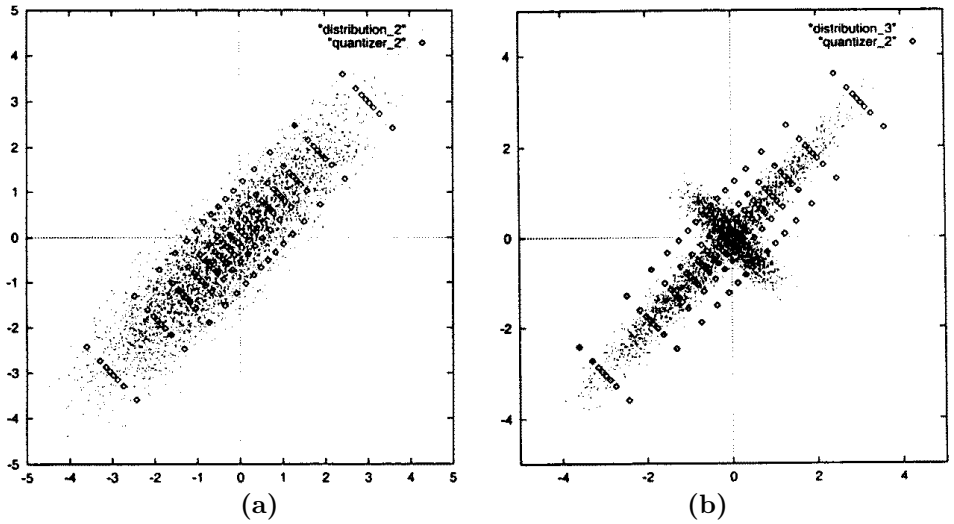

FIG. 4. *Scalar quantization in the Karhunen–Loève basis (Gaussian* (a) *and lacunary* (b) *distribution).*

In the case where x and y are not independent, one can try to recover a distribution that is "closer" to a separable function by a change of basis that decorrelates the coordinates. Such a transformation always exists and corresponds to the expansion in the Karhunen–Loève basis, i.e., the eigenvectors of the matrix $r_{i,j} = E(v_i v_j)$, $i = 1, 2$, $v_1 = x$, $v_2 = y$. As shown in Figure 4, the efficiency of this strategy again depends on the properties of $p(x, y)$. In particular, when $p(x, y)$ is Gaussian (4(a)), the Karhunen–Loève transform results in a separable distribution that is appropriate for scalar quantization. A typical example of such a situation is the decorrelation of the red, green, and blue pixels in a color image. In contrast, Figure 4(b) displays the example of a distribution that is essentially concentrated around its Karhunen–Loève directions. In that case, it is clear that a separable coding, even in the Karhunen–Loève expansion, is not adapted.

How should one quantize the lacunary distribution that is displayed in Figure 4(b)? A natural idea is to proceed in the following way.

- Describe z in the basis $\{u, v\}$ by its two coordinates $z_u = \langle z, u \rangle$ and $z_v = \langle z, v \rangle$.
- Encode on one bit the direction around which z is concentrated: 0 if $|z_u| \geq |z_v|$ (u dominates) and 1 if $|z_v| > |z_u|$ (v dominates).
- Apply a separable quantization in an adaptive way: if u (resp., v) dominates, encode z_u (resp., z_v) on b_1 bits and z_v (resp. z_u) on b_2 bits, where $b_2 << b_1$. The total number of bits is $b = b_1 + b_2 + 1$.

This technique results in a better adapted collection of quantization points, which is displayed in Figure 5(a). Note that the directions around which the distribution is concentrated need not be the Karhunen–Loève directions, as shown in Figure 5(b). In that case, it is clear that in the first step one should replace the Karhunen–Loève basis with $\{\tilde{u}, \tilde{v}\}$, which correspond to the directions around which the distribution is concentrated.

The idea of quantizing in an adaptive way is already present in Shapiro's EZW algorithm (Shapiro (1993)). It can be generalized to vector quantization: vectors with

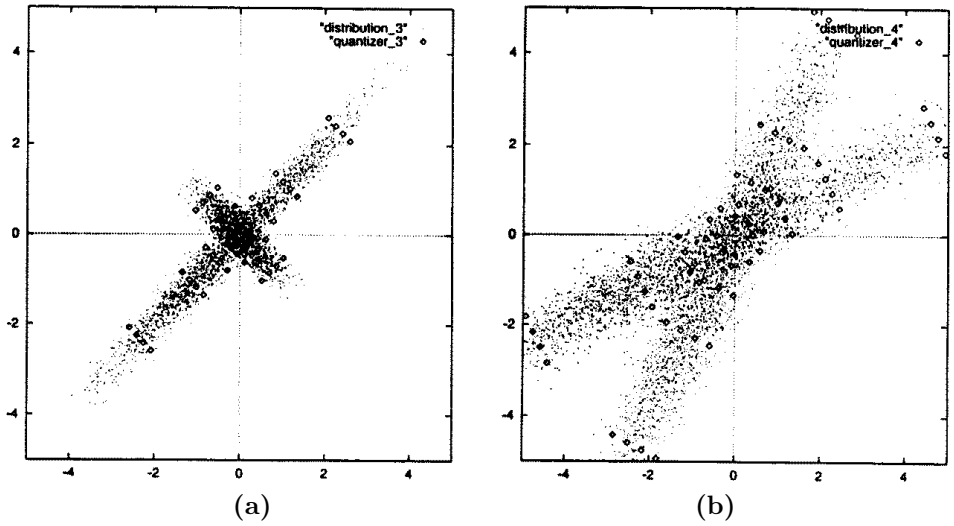

FIG. 5. *Adapted quantization of a lacunary distribution.*

a more important norm will be encoded on a higher amount of bits (Averbuch, Israeli, and Lazaar (1993)).

The bases (u, v) and $(\tilde{u}, \tilde{v})$ have different approximation properties. The Karhunen–Loève basis is optimal for linear approximation: it minimizes the error $\varepsilon = E(\|z - \langle z, e_0\rangle e_0\|^2)$ among all choices of (e_0, e_1). In contrast, the basis $(\tilde{u}, \tilde{v})$ is optimal for nonlinear approximation: it minimizes the error $\tilde{\varepsilon} = E(\|z - \langle z, e_s\rangle e_s\|^2)$ among all choices of (e_0, e_1), where $s = s(z) \in \{0, 1\}$ is such that $|\langle z, e_s\rangle| \geq |\langle z, e_{1-s}\rangle|$. Note that in the Gaussian case, both bases always coincide.

From these examples, we can formulate two "qualitative statements":

- In the case of a Gaussian distribution, the Karhunen–Loève basis is optimal for both linear and nonlinear approximation. The expansion in this basis can be quantized efficiently in a separable fashion.
- In the case of a lacunary distribution, the optimal basis for nonlinear approximation may differ from the Karhunen–Loève basis. There may be a substantial gain in using nonlinear approximation in this optimal basis rather than any linear approximation technique and quantizing the expansion in an adaptive rather than separable fashion.

Note that in both cases, the basis that is adapted to perform compression is the one that is optimal in terms of nonlinear approximation.

We claim that many important classes of signals—in particular, real images—have a lacunary structure and thus need to be quantized in an adaptive way. This structure cannot be directly identified on the probability law of these large-dimensional signals, since it is too complex and, as we pointed out, cannot be evaluated with a high precision in reasonable time. However, this structure will be revealed from the fact that nonlinear approximation with certain systems (in particular, wavelet bases) performs highly better than any linear approximation method. This fact will be analyzed quantitatively in the following sections, using a statistical modelization of these signals that does not require the full knowledge of their probability distribution.

3. Linear approximation. Let $s(t)$ be a stochastic process of the second order defined on $[0, 1]$. In this section, we shall assume that the sole information that we

have on the process $s(t)$ is its autocorrelation function, i.e.,

$$R(t_1, t_2) = E(s(t_1)\overline{s(t_2)}). \tag{3.1}$$

How can we use this information to compare the effect of the decomposition of the signal in different bases? Given an orthonormal basis $\{e_n\}_{n\geq 0}$, it is possible to compute the mean square value of the coordinates of $s(t)$ since we have

$$\begin{aligned} E\left(|\langle s, e_n\rangle|^2\right) &= E\left(\left|\int_0^1 s(t)\overline{e_n(t)}dt\right|^2\right) \\ &= E\left(\iint_{[0,1]^2} s(t)\overline{s(u)e_n(t)}e_n(u)dtdu\right) \\ &= \iint_{[0,1]^2} R(t,u)\overline{e_n(t)}e_n(u)dtdu \\ &= \langle R, e_{n,n}\rangle, \end{aligned}$$

with $e_{n,m}(x,y) = e_n(x)\overline{e_m(y)}$. Thus, the autocorrelation function allows one to estimate the mean square error between $s(t)$ and its linear approximation by its N first coordinates,

$$s_N(t) = \sum_{n=0}^{N-1} \langle s, e_n\rangle e_n(t), \tag{3.2}$$

since we have

$$\varepsilon(N) = E(\|s - s_N\|^2) = \sum_{n\geq N} E(|\langle s, e_n\rangle|^2). \tag{3.3}$$

The basis that minimizes this quantity for every N is the Karhunen–Loève basis. It is obtained by taking a complete orthonormal set of eigenfunctions $\{e_n\}_{n\geq 0}$ of the integral operator

$$\mathcal{R}f(t) = \int_0^1 R(t,u)f(u)du \tag{3.4}$$

and by rearranging them in such a way that the corresponding eigenvalues $\lambda_n \geq 0$ are nonincreasing with n.

In most cases, these eigenfunctions have no explicit form and the computation of the coefficients of a discretized function f in the Karhunen–Loève basis requires a large number of operations. Fortunately, it is often possible to obtain a near optimal approximation with a simpler system that is better adapted for numerical computations.

Let us first consider the case of a wavelet basis. We recall that these bases, usually constructed on the whole real line (see Daubechies (1992) for a general introduction), can be adapted to the interval (Cohen, Daubechies, and Vial (1993)): $L^2[0,1]$ is approximated by a multiresolution analysis, i.e., a ladder of closed subspaces

$$V_{j_0} \subset V_{j_0+1} \subset V_{j_0+2} \subset \cdots \to L^2[0,1], \tag{3.5}$$

with $j_0 \geq 0$, where V_j is generated by 2^j orthonormal scaling functions $\varphi_{j,k}$, $k = 0, \ldots, 2^j - 1$, such that $\operatorname{supp}(\varphi_{j,k}) \subset [2^{-j}(k-c), 2^{-j}(k+c)]$ (c does not depend on

j). At each level, the orthonormal complement between V_j and V_{j+1} is generated by 2^j orthonormal wavelets $\psi_{j,k}$, $k = 0, \dots, 2^j - 1$, such that $\mathrm{supp}(\psi_{j,k}) \subset [2^{-j}(k - c), 2^{-j}(k + c)]$. As a consequence, the family

$$\bigcup_{j \geq j_0} \{\psi_{j,k}\}_{k=0,\dots,2^j-1}, \tag{3.6}$$

completed by $\{\varphi_{j_0,k}\}_{k=0,\dots,2^{j_0}-1}$, constitutes an orthonormal basis of $L^2[0,1]$. Although we shall not make use of this property, let us mention that when the support of the scaling functions and wavelets do not contain 0 or 1, these functions are simply defined from the standard scaling function and wavelet φ and ψ with the notation $f_{j,k} = 2^{j/2} f(2^j \cdot - k)$.

The results of linear approximation in the wavelet basis relate the regularity of $R(t,u)$ on the diagonal line $\{t = u\}$ and the cancellation properties of the wavelets, i.e., their number of vanishing moments. However, equivalent results could be stated in terms of linear approximation in the spaces V_j, ignoring the functions $\psi_{j,k}$ and considering the degree of polynomial reproduction satisfied by the scaling functions $\varphi_{j,k}$.

We give here the formulation using the cancellation properties of the wavelets that was given first in Cohen, Froment, and Istas (1991) in the case of a stationary process. For this, we reorder the wavelet basis (3.6) by $e_n = \varphi_{j_0,n}$ when $n < 2^{j_0}$ and $e_n = \psi_{j,k}$ when $n = 2^j + k$, $0 \leq k < 2^j$, $j \geq j_0$.

In the paper of Cohen, Froment, and Istas (1991) the regularity of R was controlled with the help of the speed of decrease at infinity of its Fourier transform. We use a simpler condition which emphasizes the fact that only the regularity of R along the diagonal is of importance here. For $\alpha > 0$, we say that a real function $F(x)$, $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ is C^α at $z \in \mathbb{R}^d$ if and only if there exists a polynomial $P(x) = \sum_{|m|<\alpha} a_m x^m$ (with $|m| = m_1 + \cdots + m_d$ and $x^m = x_1^{m_1}, \dots, x_d^{m_d}$) such that for all x in a neighborhood V_z of z

$$|P(x) - F(x)| \leq C\|x - z\|^\alpha. \tag{3.7}$$

Note that for $\alpha \in \mathbb{N}$ this is weaker than α differentiability. If $D \subset \mathbb{R}^d$, the function F is said to be C^α on D if it is C^α at each point $z \in D$ and if the constant C in (3.7) is uniformly bounded on D.

THEOREM 1. *Suppose that $R(t,u)$ is C^α on the diagonal line $\{(t,t) | t \in [0,1]\}$ for a certain $\alpha > 0$ and that for all $m \in \mathbb{N} \cap [0, \alpha[$, $j \geq j_0$ and $k = 0, \dots, 2^j - 1$, $\int_0^1 x^m \psi_{j,k}(x) dx = 0$. Then*

$$\varepsilon(N) \leq C N^{-\alpha}. \tag{3.8}$$

Proof. It is clearly sufficient to prove (3.8) for $N = 2^p$, $p \geq j_0$. The property that $R(t,u)$ is C^α on the diagonal means that there exists $C > 0$ such that for all $v \in [0,1]$, there is a polynomial $P_v(t,u)$ of global degree strictly less than α with

$$|R(t,u) - P_v(t,u)| \leq C(|t - v| + |u - v|)^\alpha \tag{3.9}$$

for all $(t,u) \in [0,1]^2$. Using this estimate at $v = 2^{-j}k$, we get

$$\begin{aligned} E(|\langle s, \psi_{j,k}\rangle|^2) &= \iint_{[0,1]^2} R(t,u)\overline{\psi_{j,k}(t)}\psi_{j,k}(u)dtdu \\ &= \iint_{[0,1]^2} (R(t,u) - P_{2^{-j}k}(t,u))\overline{\psi_{j,k}(t)}\psi_{j,k}(u)dtdu \\ &\leq C \iint_{[0,1]^2} (|t-v| + |u-v|)^\alpha |\psi_{j,k}(t)\psi_{j,k}(u)|dtdu. \end{aligned}$$

Using the support properties of the functions $\psi_{j,k}$ and Schwarz inequality, we derive

$$\begin{aligned} E(|\langle s, \psi_{j,k}\rangle|^2) &\leq C(2c)^\alpha 2^{-\alpha j} \iint_{[0,1]^2} |\psi_{j,k}(t)\psi_{j,k}(u)|dtdu \\ &\leq C(2c)^{\alpha+1} 2^{-(\alpha+1)j}. \end{aligned}$$

Summing on all $k = 0, \ldots, 2^j - 1$, then on all $j \geq p$, we finally obtain the desired estimate. □

Remark. This result is in fact a simple rephrasing, in the stochastic framework, of the deterministic results on the multiresolution approximation of functions in Sobolev spaces, the error being measured in L^2 norm (see Meyer (1990) and Daubechies (1992)): $\alpha/2$ represents the degree of differentiability in the mean square sense of the process $s(t)$.

Let us now consider the trigonometric system $e_n(t) = e^{i2\pi nt}$, $n \in \mathbb{Z}$. In that case, an estimate on the quantity

$$E(|\langle s, e_n\rangle|^2) = \iint_{[0,1]^2} R(t,u)e^{i2\pi n(u-t)}dtdu \tag{3.10}$$

can be easily obtained from the regularity of $R(t,u)$ only when this function is the restriction of a regular $\mathbb{Z}^2$-periodic function. However, most signals—in particular, images—do not satisfy this property. To estimate $\varepsilon(N)$, one thus needs more information on $R(t,u)$.

At this point, let us remark that an important class of 1D signals satisfies the property of stationarity, i.e., $R(t,u) = r(|t-u|)$. Note that in the case where $r(t)$ is a 1-periodic function, the Karhunen–Loève basis is given by the trigonometric system. As was mentioned, the signals that we have in mind do not satisfy this property: an image, a video sequence, or a piece of speech is the restriction on a finite domain of a nonperiodic function. For all these signals, $r(t)$ is typically an even function that decreases on $[0, +\infty[$.

In the case of black and white pictures, a commonly used approximation of the autocorrelation function of real images is given by

$$E(I(x,y)I(x',y')) = R(x,x',y,y') = R_0 e^{-\mu(|x-x'|+|y-y'|)}, \tag{3.11}$$

where R_0 and λ are constants that depend on the normalization of the light intensity function and the size of the image. In particular, $I(x,y)$ is supposed to be centered around zero: it takes the values $I_{\max}$ (resp., $-I_{\max}$) in the white (resp., black) regions. We shall thus focus on the 1D stationary processes with autocorrelation function

$$R(t,u) = e^{-|t-u|}. \tag{3.12}$$

Note that the exponential decay of the correlation as a function of the distance is also a typical feature of Markov processes.

In that particular case, Theorem 1 indicates that the linear approximation error $\varepsilon(N)$ with a wavelet basis is majorized by CN^{-1}.

For the trigonometric system, we obtain

$$\begin{aligned} E(|\langle s, e_n\rangle|^2) &= \iint_{[0,1]^2} e^{-|t-u|} e^{i2\pi n(u-t)} dt du \\ &= \int_{-1}^{1} e^{-|t|}(1-|t|)e^{i2\pi nt} dt \\ &= \int_0^1 (1-t)e^{-t}(e^{i2\pi nt} + e^{-i2\pi nt}) dt \\ &= \frac{2}{1+4\pi^2 n^2} + \frac{(8\pi^2 n^2 - 2)(1-1/e)}{(1+4\pi^2 n^2)^2}. \end{aligned}$$

As N goes to infinity, we thus obtain the estimate

$$\varepsilon(N) \approx \sum_{|n|\geq N/2} E(|\langle s, e_n\rangle|^2) \approx \frac{2(2-1/e)}{\pi^2 N}, \tag{3.13}$$

which shows that the trigonometric system performs as well as a wavelet basis (for these particular processes). In fact, all these systems are near optimal for linear approximation, in the sense that they perform as well as the Karhunen–Loève basis. This can be checked by determining explicitly the Karhunen–Loève functions: for a given eigenvalue $\lambda > 0$ of $\mathcal{R}$ (defined by (3.4)), the associated eigenfunction is C^∞ and satisfies

$$\begin{aligned} \lambda f(t) &= \int_0^1 e^{-|t-u|} f(u) du \\ &= e^{-t} \int_0^t e^u f(u) du + e^t \int_t^1 e^{-u} f(u) du. \end{aligned}$$

Differentiating once, we obtain

$$\lambda f'(t) = -e^{-t} \int_0^t e^u f(u) du + e^t \int_t^1 e^{-u} f(u) du, \tag{3.14}$$

which shows that necessarily

$$f'(0) = f(0) \quad \text{and} \quad f'(1) = -f(1). \tag{3.15}$$

After a second differentiation, we obtain the equation

$$\lambda f''(t) = (\lambda - 2) f(t). \tag{3.16}$$

Now it is clear that $\|\mathcal{R}\| \leq 1$ so that necessarily $(\lambda-2)/\lambda \leq 0$. The solutions of (3.16) are thus of the form $a\cos(\omega t) + b\sin(\omega t)$ with $\omega = \sqrt{2/\lambda - 1}$. From the boundary conditions (3.15), we finally obtain the family of orthogonal eigenfunctions

$$e_n(t) = \pi n \cos(\pi n t) + \sin(\pi n t), \quad n \in \mathbb{Z} - \{0\}, \tag{3.17}$$

with the associated eigenvalues

$$\lambda_n = (1+\pi^2 n^2)^{-1}. \tag{3.18}$$

It follows that the approximation error in the Karhunen–Loève basis satisfies

$$\varepsilon(N) \approx 2(\pi^2 N)^{-1}. \tag{3.19}$$

The conclusion of this section is that, when the autocorrelation is given by (3.12), the asymptotic performances of the Karhunen–Loève basis, the trigonometric system and wavelet bases are equivalent for linear approximation.

In the bidimensional case (3.11), it is immediate to derive an analogous estimate because of the separable structure of $R(x,x',y,y')$. One obtains in that case an approximation error $\varepsilon(N)$ of order $N^{-1}(\log N)^2$ for wavelet bases and $N^{-1}\log N$ for trigonometric and Karhunen–Loève bases.

In order to investigate nonlinear approximation, we shall now introduce more information on the process $s(t)$. In particular, in the case of images, the autocorrelation function averages the smooth regions (that correspond to homogeneous objects) and the isolated discontinuities (that correspond to sharp edges). The model that we present in the next section is an attempt to describe this piecewise smooth aspect on stochastic processes.

4. Piecewise stationary processes. In order to reflect the "piecewise smooth" property in a stochastic framework, we shall describe our process $s(t)$ as follows.

A finite set of discontinuities $D = \{d_1, d_2, \dots, d_L\} \subset [0,1]$, $d_i \le d_{i+1}$, is obtained as the realization on $[0,1]$ of a Poisson process of parameter $\mu > 0$. This means that the number of discontinuity is a random number with probability law

$$P(|D| = L) = e^{-\mu}\frac{\mu^L}{L!} \tag{4.1}$$

and that conditionally to the event $|D| = L$ the distribution of $(d_1, \dots, d_L)$ is uniform over the simplex $\{0 \le x_1 \le \cdots \le x_L \le 1\}$.

Conditionally to the data of such a set $\{d_1, \dots, d_L\}$, we set $d_0 = 0$ and $d_{L+1} = 1$ and define $s(t)$ on $[d_i, d_{i+1}[$, $i = 0, \dots, L$, by

$$s(t) = s_i(t), \tag{4.2}$$

where the functions s_i are independent realizations of a stationary process with autocorrelation function $R(t,u) = r(|t-u|)$ and mean M. We shall assume in the rest of this section that $r(t)$ is C^α on $[-1,1]$ for some $\alpha > 3/2$.

At this point, one should remark that the global process $s(t)$ is also stationary: its autocorrelation function is given by

$$R_s(t,u) = P(t,u)R(t,u) + (1-P(t,u))M^2, \tag{4.3}$$

where $P(t,u) = e^{-\mu|t-u|}$ is the probability that no discontinuity d_i lies between t and u. We thus have

$$R_s(t,u) = M^2 + e^{-\mu|t-u|}(r(|t-u|) - M^2) = r_s(|t-u|). \tag{4.4}$$

Note that in the simplest case where $s_i(t) = s_i$ are independent realizations of a constant centered process, the autocorrelation function is given by

$$R_s(t,u) = r_s(0)e^{-\mu|t-u|}. \tag{4.5}$$

This shows that the parameter μ in the model (3.11) can be interpreted as the Poisson density of discontinuities on a line in a real image.

Remarks.

- These processes are also a good model for the evolution in time of the intensity of a fixed pixel in a video sequence: the discontinuities correspond to an edge in motion that crosses the pixel at a given time or to a brutal change of image.
- The generalization of this model to bidimensional signals is not straightforward: the discontinuities are no longer isolated points but curves. So far, we have investigated only the unidimensional situation, which is nevertheless revealing about the results that we can expect in the multidimensional case.
- In contrast with the results of wavelet linear approximation in the previous section, the result of nonlinear approximation that we shall prove in the next section seems difficult to identify with deterministic approximation in function spaces. Indeed, a proper function space that describes piecewise smooth functions should be equipped with a norm that measures both the height of the jumps at the discontinuities and the smoothness—for example, in the Sobolev H^1 norm—between these discontinuities. For a function f with a finite number of discontinuities, this norm would be given by

$$\|f\|_{\mathrm{ps}} = \sum_{i=1}^{L} |f(d_i)_+ - f(d_i)_-| + \left(\sum_{i=0}^{L} \int_{d_i}^{d_{i+1}} |f(x)|^2 + |f'(x)|^2 dx \right)^{1/2}. \tag{4.6}$$

  It is clear that the space of piecewise H^1 functions with a finite number of discontinuities is not complete for the above norm. The completion process leads to the space $BV[0,1]$, which cannot be well approximated linearly or nonlinearly and does not describe the property of piecewise smoothness: the smoothness term has been "swallowed" by the jump term.

We end this section with a result that shows that the piecewise smooth processes that we have introduced cannot be well approximated if one proceeds linearly.

THEOREM 2. *Let λ_n, $n \geq 0$, be the sequence of eigenvalues of the integral operator $\mathcal{R}_s$ associated with the kernel $R_s(t,u)$. We assume that λ_n is nonincreasing and that the eigenvalues are repeated according to their multiplicity. There exists $C > 0$ such that for all $N \geq 0$*

$$\varepsilon(N) = \sum_{n \geq N} \lambda_n \geq C N^{-1}. \tag{4.7}$$

Consequently any linear approximation of $s(t)$ cannot achieve better than CN^{-1} in the mean square sense.

Proof. From (4.4), we can write

$$R_s(t,u) = K_A(t,u) + K_B(t,u), \tag{4.8}$$

where

$$K_A(t,u) = e^{-\mu|t-u|}(r(0) - M^2) \tag{4.9}$$

and

$$K_B(t,u) = M^2 + e^{-\mu|t-u|}(r(|t-u|) - r(0)). \tag{4.10}$$

Note that $K_B(t,u)$ belongs to C^2. Denote by

$$\mathcal{R}_s = A + B \tag{4.11}$$

the associated decomposition of the autocorrelation operator.

It is clear that A is a positive operator. From the discussion in section 3, we know that its eigenvalues $\{a_n\}_{n\geq 0}$, ordered as a nonincreasing sequence, satisfy

$$\lim_{n\to+\infty} n^2 a_n = C, \tag{4.12}$$

where C depends on the parameters M, $r(0)$, and μ. In contrast, B need not be a positive operator. We denote by $\{b_n\}_{n\geq 0}$ its singular values, i.e., the eigenvalues of $|B|$, ordered as a nonincreasing sequence. To prove (4.7), we shall use the following result due to Ky Fan (1951) (see also Gohberg and Krein (1965)).

Let A and B be compact operators in a Hilbert space, and let $C = A + B$. Let a_n, b_n, and c_n be the associated sequences of singular values, and assume that, for some $r > 0$,

$$\lim_{n\to+\infty} n^r a_n = C \tag{4.13}$$

and

$$\lim_{n\to+\infty} n^r b_n = 0. \tag{4.14}$$

Then

$$\lim_{n\to+\infty} n^r c_n = C. \tag{4.15}$$

We will thus concentrate on proving that

$$\lim_{n\to+\infty} n^2 b_n = 0. \tag{4.16}$$

By the Karhunen–Loève theorem, we remark that we have

$$\sum_{n\geq N} b_n = \min_{\{e_k\} \text{ o.b.}} \sum_{n\geq N} \langle e_n, |B| e_n\rangle, \tag{4.17}$$

the minimum being taken over all orthonormal bases. Consider a wavelet basis of type (3.6), reordered as in section 3, and such that the functions $\psi_{j,k}$ have two vanishing moments. We thus have, for $p \geq j_0$,

$$\sum_{n\geq 2^p} b_n \leq \sum_{j\geq p,k} \langle \psi_{j,k}, |B| \psi_{j,k}\rangle. \tag{4.18}$$

We define $\tilde{\psi}_{j,k} = U\psi_{j,k}$ where U is a unitary operator such that $B = U|B|$. We thus have, using the Schwarz inequality,

$$\begin{aligned}
\langle \psi_{j,k}, |B| \psi_{j,k}\rangle &= \langle \tilde{\psi}_{j,k}, B\psi_{j,k}\rangle \\
&= \iint_{[0,1]^2} K_B(t,u)\psi_{j,k}(u)\tilde{\psi}_{j,k}(t)dtdu \\
&\leq \left(\int_0^1 \left| \int_0^1 K_B(t,u)\psi_{j,k}(u)du \right|^2 dt \right)^{-1/2} \\
&\leq \sup_{t\in[0,1]} \left| \int_0^1 K_B(t,u)\psi_{j,k}(u)du \right|.
\end{aligned}$$

Since $K_B(t,u)$ is C^β for $\beta = \min(\alpha, 2)$ and the functions $\psi_{j,k}$ have two vanishing moments, we have

$$\begin{aligned}\sup_{t\in[0,1]} \left| \int_0^1 K_B(t,u)\psi_{j,k}(u)du \right| &\le C \int_0^1 |u - 2^{-j}k|^\beta |\psi_{j,k}(u)|du \\ &\le C\|\psi_{j,k}\|_1 \int_{|2^j u - k|\le c} |u - 2^{-j}k|^\beta du \\ &\le C 2^{-(\beta+1/2)j}.\end{aligned}$$

We thus have

$$\langle \psi_{j,k}, |B|\psi_{j,k}\rangle \le C 2^{-(\beta+1/2)j} \tag{4.19}$$

and

$$\sum_{n\ge N} b_n \le C N^{-\beta+1/2}. \tag{4.20}$$

Finally, since the b_n are positive and nonincreasing, (4.20) implies

$$b_n \le C n^{-(\beta+1/2)}, \tag{4.21}$$

which allows one to apply Ky Fan's theorem. □

Remark. Note that to estimate the decay of b_n one can use

$$b_n = \inf_{T\in\mathcal{T}_n} \|B - T\|, \tag{4.22}$$

where $\mathcal{T}_n$ is the space of operators with rank at most equal to n. To majorize this infimum, it seems natural to choose a kernel K_T that is an approximation of K_B by a sum of n separable functions, typically piecewise polynomial functions. Then one can relate the decreasing order of the sequence (b_n) with the smoothness of R_s in the Besov scale (see Birman and Solomyak (1977)). However, this does not lead to any improvement of the theorem.

5. Nonlinear approximation. We now turn to the nonlinear approximation

$$\mathcal{A}_N s = \sum_{k\in E_N} \langle s, e_k\rangle e_k, \tag{5.1}$$

where $E_N = E_N(s)$ is the set of indices of the N largest coordinates of the process $s(t)$ that has been described in the previous section. In this section, the mean square error will be defined by

$$\varepsilon(N) = E(\|\mathcal{A}_N s - s\|_2^2) = E\left(\sum_{n\notin E_N(s)} |\langle s, e_n\rangle|^2 \right). \tag{5.2}$$

We first consider the case where $\{e_n\}_{n\ge 0}$ is a wavelet basis of the type (3.6), i.e., $\{\varphi_{j_0,k}\}_k \cup \{\psi_{j,k}\}_{j\ge j_0,k}$ reordered as in section 3. The following theorem shows that nonlinear wavelet approximation performs as well as if there were no discontinuities in the process $s(t)$.

THEOREM 3. *Suppose that $r(t)$ is in C^α for some $\alpha > 0$ and that for all $m \in \mathbb{N} \cap [0, \alpha[$, $j \geq j_0$, and $k = 0, \ldots, 2^j - 1$, $\int_0^1 x^m \psi_{j,k}(x)dx = 0$. Then*

$$\varepsilon(N) \leq CN^{-\alpha}. \tag{5.3}$$

Proof. As with Theorem 1, it is clearly sufficient to prove (5.3) for $N = 2^p$, $p \geq j_0$. Remark that if we define

$$\mathcal{A}'_N s = \sum_{k \in E'_N(s)} \langle s, e_k \rangle e_k, \tag{5.4}$$

where $E'_N(s) \subset \mathbb{N}$ has cardinal $|E'_N| \leq N$, we always have

$$\varepsilon(N) \leq \varepsilon'(N) = E(\|\mathcal{A}'_N s - s\|_2^2). \tag{5.5}$$

By this remark, it is sufficient to build a near optimal nonlinear approximation $\mathcal{A}'_N$ such that

$$\varepsilon'(2^p) \leq C2^{-\alpha p}. \tag{5.6}$$

For $N = 2^p$, we shall define the set $E'_N(s)$ by

$$E'_N = \{0, 1, \ldots, 2^{p-1} - 1\} \cup E''_N(s), \tag{5.7}$$

where $E''_N(s)$ is a set of cardinal $|E''_N| \leq 2^{p-1}$ that depends on the locations of the discontinuities $\{d_1, d_2, \ldots, d_L\}$ in $s(t)$: recalling from section 3 that $\mathrm{supp}(\psi_{j,k}) \subset [2^{-j}(k-c), 2^{-j}(k+c)]$, we consider the subset of discontinuities $\{d_1, d_2, \ldots, d_{L(p)}\}$, where

$$L(p) = \max \left\{ m \in \{1, \ldots, L\};\ m \leq \frac{2^{p-1}}{2c\alpha p} \right\}, \tag{5.8}$$

and we define E''_N as the set of indices n of all wavelets $\psi_{j,k} = e_n$ such that $p - 1 \leq j < (\alpha + 1)p - 1$ and $d_i \in \mathrm{supp}(\psi_{j,k})$ for some $i \in \{1, \ldots, L(p)\}$. From the definition of $L(p)$, it is clear that $|E''_N| \leq 2^{p-1}$.

We now evaluate $\varepsilon'(2^p)$. Two events will be considered: $L(p) = L$ and $L(p) < L$.

In the event where $L(p) = L$, we remark that if the index n of a wavelet $\psi_{j,k} = e_n$ is not in E'_N, two situations are possible:

• $d_i \notin \mathrm{supp}(\psi_{j,k})$ for all $i \in \{1, \ldots, L(p)\}$. This means that the support of $\psi_{j,k}$ is fully contained between d_i and d_{i+1} for some $i \in \{1, \ldots, L\}$, and we obtain from Theorem 1

$$E(|\langle s, \psi_{j,k} \rangle|^2) = E(|\langle s_i, \psi_{j,k} \rangle|^2) \leq C2^{-(\alpha+1)j}. \tag{5.9}$$

• $d_i \in \mathrm{supp}(\psi_{j,k})$ for some $i \in \{1, \ldots, L(p)\}$. In that case, we use Schwarz inequality to obtain a crude estimate,

$$E(|\langle s, \psi_{j,k} \rangle|^2) \leq E\left(\int_{(k-c)2^{-j}}^{(k+c)2^{-j}} |s(t)|^2 dt \right) \leq 2cr(0)2^{-j}. \tag{5.10}$$

However, in view of the definition of E''_N, one necessarily has $j \geq (\alpha + 1)p - 1$ in that case.

Summing (5.9) on k and on $j \geq p-1$ and (5.10) on k and $j \geq (\alpha+1)p-1$, we finally obtain the desired estimate for the conditional expectation:

$$e_p^1 = E(\|\mathcal{A}'_N s - s\|_2^2 \,;\, L = L(p)) \leq C 2^{-\alpha p}. \tag{5.11}$$

In the event where $L(p) < L$, i.e., $L > \frac{2^{p-1}}{2c\alpha p}$, a very crude estimate will be sufficient (because of the very small probability of this event as p goes to $+\infty$). We simply use

$$e_p^2 = E(\|\mathcal{A}'_N s - s\|_2^2;\ L > L(p)) \leq E(\|s\|_2^2; L > L(p)) = r(0). \tag{5.12}$$

Combining these estimates, we obtain

$$\begin{aligned}
\varepsilon'(2^p) &= P(L = L(p)) e_p^1 + P(L > L(p)) e_p^2 \\
&\leq C\left(2^{-\alpha p} + P\left(L > \frac{2^{p-1}}{2c\alpha p}\right)\right) \\
&= C\left(2^{-\alpha p} + e^{-\mu} \sum_{l > \frac{2^{p-1}}{2c\alpha p}} \frac{\mu^l}{l!}\right).
\end{aligned}$$

One easily checks that the second term decreases exponentially faster than the first one so that we finally have (5.6). This concludes the proof of the theorem. □

We now turn to the trigonometric system $e_n(x) = e^{i2\pi nx}$, $n \in \mathbb{Z}$. We shall assume that r is twice differentiable; i.e., the process that describes $s(t)$ between the discontinuities is differentiable in the mean square sense. In that case, we shall prove that nonlinear approximation does not perform substantially better than linear projection: in contrast with the wavelet basis that sparsifies the process $s(t)$, the best Fourier coefficients essentially coincide with the first Fourier coefficients. This fact is specific to the type of processes that we are considering. Other types of signals, such as velocity fields in turbulent flows (see Farge et al. (1992)), may present a more lacunary structure in the Fourier domain.

THEOREM 4. *Assume $r(0) > r(1)$. Then the nonlinear approximation error in the trigonometric system satisfies*

$$\varepsilon(N) \geq C N^{-1}. \tag{5.13}$$

Proof. We define

$$s_k = \langle s, e_k \rangle = \int_0^1 s(t) e^{-i2\pi kt} dt. \tag{5.14}$$

We shall prove that there exists $K, D > 0$ such that the event

$$|k| \geq K \to |s_k| \geq D|k|^{-1} \tag{5.15}$$

has a probability $p > 0$. It is clear that this property will imply the estimate (5.13): in the event (5.15), we have indeed, for any set F_N of cardinal $N > 0$,

$$\sum_{k \notin F_N} |s_k|^2 \geq C N^{-1}. \tag{5.16}$$

Applying (5.6) to E_N, we finally obtain

$$\varepsilon(N) \geq pCN^{-1}, \tag{5.17}$$

which is equivalent to (5.13), up to a change in the constant C.

To show that (5.15) occurs with strictly positive probability, we consider the event where there is no discontinuity, i.e., $L = 0$. This event has a probability $e^{-\mu} > 0$. In that situation we decompose $s(t)$ into

$$s(t) = a(t) + b(t), \tag{5.18}$$

where

$$a(t) = (1-t)s(0) + ts(1) \text{ and } b(t) = s(t) - (1-t)s(0) - ts(1). \tag{5.19}$$

We denote by a_k and b_k, $k \in \mathbb{Z}$, the Fourier coefficients of $a(t)$ and $b(t)$ so that $s_k = a_k + b_k$.

From the assumption $r(0) > r(1)$, we get

$$E(|s(0) - s(1)|^2;\ L = 0) = 2(r(0) - r(1)) > 0 \tag{5.20}$$

and thus the event

$$|s(0) - s(1)| \geq \sqrt{r(0) - r(1)} \tag{5.21}$$

occurs with probability $p' > 0$. It is clear that (5.21) implies

$$|a_k| \geq 2D|k|^{-1} \tag{5.22}$$

for some $D > 0$ related to $\sqrt{r(0) - r(1)}$.

We now turn to the coefficients b_k: from (5.19), it is clear that $b(0) = b(1)$. Since $s(t)$ is differentiable in the mean square sense (in the case where $L = 0$), we have for all $k \in \mathbb{Z} - \{0\}$,

$$\begin{aligned} b_k &= \int_0^1 b(t) e^{-i2\pi kt} dt \\ &= (-i2\pi k)^{-1} \int_0^1 b'(t) e^{-i2\pi kt} dt \\ &= (-i2\pi k)^{-1} \int_0^1 s'(t) e^{-i2\pi kt} dt. \end{aligned}$$

This leads to

$$E(|b_k|^2) = \frac{E(|\langle s', e_k \rangle|^2)}{4\pi^2 k^2}. \tag{5.23}$$

Since $E(\|s'\|_2^2) < +\infty$, it follows that

$$\sum_{k \neq 0} k^2 E(|b_k|^2) < +\infty, \tag{5.24}$$

which implies, by the Chebyshev inequality,

$$\sum_{k \neq 0} P(|b_k| > D|k|^{-1}) < +\infty. \tag{5.25}$$

From (5.25), we can apply the Borel–Cantelli theorem to conclude that for any $\rho > 0$, there exists K_ρ such that the event

$$|k| > K_\rho \to |b_k| \leq D|k|^{-1} \tag{5.26}$$

has probability $1 - \rho$. We choose $\rho = p'/2$ so that (5.22) and (5.26) occur simultaneously with a probability larger than $p'/2$.

Consequently, (5.15) is satisfied with $K = K_{p'/2}$ and $p \geq e^{-\mu} p'/2 > 0$. □

Remarks.

- To obtain the estimate (5.15), we have used the regularity of $r(t)$. In particular, it seems difficult to avoid the assumption that r is twice differentiable.
- In contrast, the assumption $r(0) > r(1)$ is not strictly necessary. It allows one to consider only the event $L = 0$. If this assumption is removed, the process could be periodic in the case $L = 0$. In this case, one still obtains (5.15) by working on the event $L = 1$. We kept the assumption $r(0) > r(1)$ since in many practical situations $r(t)$ reaches its maximum only at the origin.

6. Application to images. This part is devoted to the application of the different algorithms that were studied previously in this paper to numerical images. More precisely, we want to check the relevance of the model introduced in section 4 for representing images and to test whether the discrepancies still occur for images between linear and nonlinear approximation with wavelet bases and between nonlinear approximation with Fourier and wavelet bases.

Since our model is 1D and aims to represent lines and columns in images, we consider approximation of 1D decompositions (wavelet and Fourier) applied separately to lines and columns of images. But our goal is to evaluate algorithms based on 2D decompositions of whole images. Therefore we also consider approximation of 2D decompositions.

The main problem that we have to tackle is one of discretization. Namely, we want to check the consistency of a continuous-time model with the help of discrete data. To simulate the decomposition of the image over the continuous-time wavelet and Fourier bases, we used, respectively, the pyramidal algorithm of Mallat (see Mallat (1989)) and the discrete Fourier transform, applied directly to the pixel values.

To relate the discrete- and continuous-time transform in the case of wavelets, one can for example, assimilate the pixel values in the image to the approximation coefficients of some continuous time image f (see Mallat (1989)):

$$p_{k,l} = \langle f, \varphi(. - k, . - l) \rangle.$$

However, this does not seem relevant to us, since the pixel values were not computed by performing a scalar product of the light intensity function I with shifted versions of the scaling function φ, i.e.,

$$p_{k,l} \neq \langle I, \varphi(. - k, . - l) \rangle.$$

As a consequence the pixel values do not have the same properties as these scalar products. Nevertheless the approximation coefficients that are obtained after applying successive iterations of the pyramidal algorithm to the pixel values come closer to the corresponding approximation coefficients of the light intensity function (see Daubechies (1992)),

$$\langle f, \varphi(2^j . - k, 2^j . - l) \rangle \sim \langle I, \varphi(2^j . - k, 2^j . - l) \rangle,$$

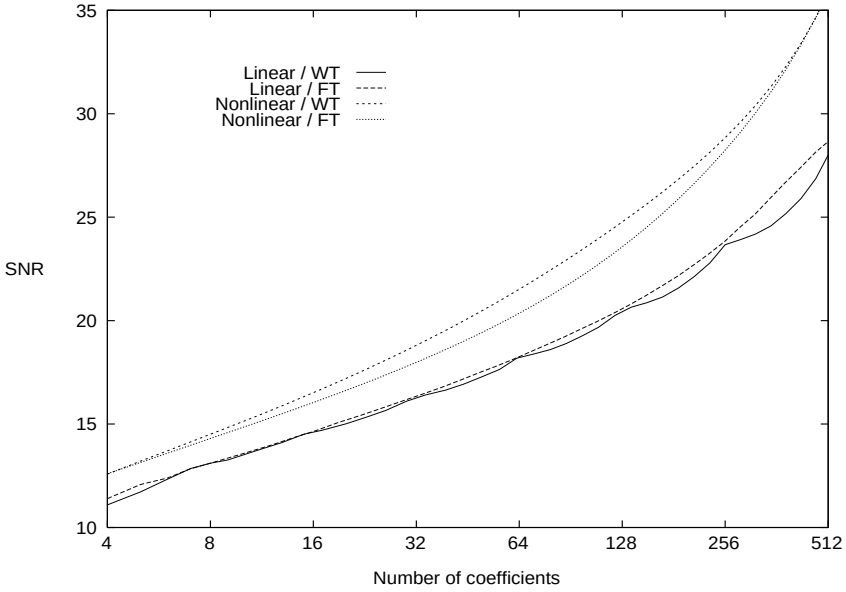

FIG. 6. *Linear/nonlinear approximation error, averaged on rows and columns of height* 1024 × 1024 *images, using* 1*D wavelet/Fourier transform.*

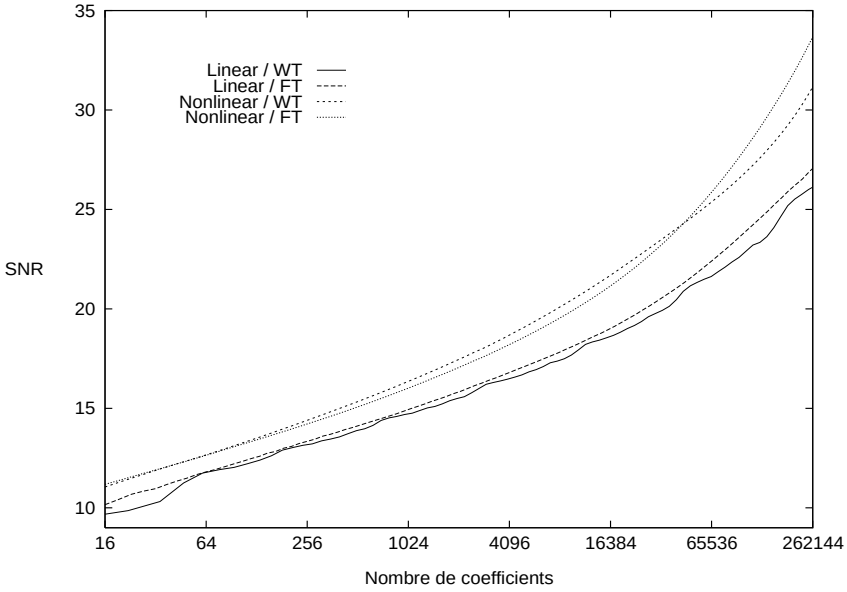

FIG. 7. *Linear/nonlinear approximation error averaged on height* 1024 × 1024 *images with* 2*D wavelet/Fourier transform.*

when j decreases to $-\infty$. Therefore we can hope to check the consistency of our model as long as fine details (which correspond to small values of $|j|$) in the wavelet transform are not taken into account. This means that we have to look for values of N (the number of coefficients that are kept to reconstruct the approximation) that are small compared to the total number of coefficients in the transform. (This number is equal to the number of pixels in a line or a column in the 1D case and to the number of pixels in the whole image in the 2D case.)

FIG. 8. *Original* 512×512 *image.*

FIG. 9. *Linear approximation with* 16384 *wavelet coefficients. PSNR* $= 28.2$.

FIG. 10. *Linear approximation with* 16384 *Fourier coefficients. PSNR* $= 28.3$.

FIG. 11. *Nonlinear approximation with* 4096 *wavelet coefficients. PSNR* $= 28.6$.

FIG. 12. *Nonlinear approximation with* 4096 *Fourier coefficients. PSNR* $= 26.3$.

Since our theoretical results are essentially asymptotic and thus concern large values of N, we are faced with two contradictory requirements. N should not be too big or too small. The same conclusion applies to Fourier transform: the high frequencies' coefficients are not to be taken into account because of the discretization problem, and our theoretical results are asymptotic too. To partially overcome this difficulty, we considered relatively big images (1024×1024).

In Figures 6 and 7, we plotted in the log-log scale the approximation error as a function of the number of coefficients that are kept in the decomposition, respectively for 1D transforms applied separately to rows and columns and for 2D transforms applied to the whole images. We used the Daubechies wavelet basis adapted to the interval with a cancellation degree of 4, which is closest to a symmetric function (see Cohen, Daubechies, and Vial (1993) and Daubechies (1992)).

We observe in Figure 6 that the curves are rectilinear for $N \leq 128$ in the case of linear approximation and $N \leq 64$ in the case of nonlinear approximation. The same is true in Figure 7 for $N \leq 128^2 = 16384$ with linear approximation and $N \leq 64^2 = 4096$ with nonlinear approximation. Thus it appears that the asymptotic results that we presented in this paper are also true for small values of N in the case of images. Note that the images that we used contain a lot of fine detail. For images containing, on the other hand, large smooth regions like that in Figure 8, one obtains curves that are rectilinear over a wider range of values of N.

We observe also that the slope of the curves in the 2D case is roughly half the slope in the 1D case. This suggests that it is sufficient to consider 1D approximation and models to get an insight into what happens with 2D approximation.

One can estimate the empirical value of α from the slope of these curves. In Figure 6 we obtain $\alpha \sim 0.5$. Since this is less than 1, it is quite normal, according to Theorems 1 and 3, to get roughly the same slope for linear and nonlinear approximation. Thus the empirical discrepancy observed between linear and nonlinear approximation corresponds to a difference between the constants C appearing in (3.8) and (5.3): the vertical shifts of these curves are proportional to $-\log C$. It would be interesting to precisely estimate these constants from a theoretical point of view. Note that in the case of 1D nonlinear wavelet approximation, the slope of the curve is slightly higher (corresponding to $\alpha \sim 0.6$). This is negligible when compared to the difference of vertical shifts with linear approximation. However, this is quite significant when compared to nonlinear Fourier approximation, since for $N \sim 128$ it implies a difference of 1 dB between the two curves. Our model is still too rough to explain this discrepancy.

Finally we present some examples of approximating images using the different algorithms presented in this paper. We see that for linear approximation, Fourier and wavelet bases give similar results. (The mean square errors are also roughly equal for Figures 9 and 10.) In contrast, for nonlinear approximation, wavelet bases outperform trigonometric bases. (There is a difference of 2.3 dB between Figures 11 and 12.)

REFERENCES

A. Averbuch, M. Israeli, and D. Lazaar (1993), *Image compression using wavelet transforms and multiresolution decompositions*, IEEE Trans. Signal Processing, to appear.

M. S. Birman and M. Z. Solomyak (1977), *Estimates of singular values of integral operators*, Uspekhi Mat. Nauk., 32, pp. 15–89 (translated in Russian Math. Surveys 32, (1977), pp. 17–84).

A. Cohen, I. Daubechies, and P. Vial (1993), *Wavelets and fast wavelet transforms on an interval*, Applied and Computational Harmonic Analysis, 1, pp. 54–81.

A. Cohen, J. Froment, and J. Istas (1991), *Analyse multirésolution de signaux aléatoires*, C. R. Acad. Sci. Paris Sér. I Math., 312, pp. 567–570.

R. R. Coifman, Y. Meyer, and V. M. Wickerhauser (1992), *Entropy-based algorithms for best basis selection*, IEEE Trans. Inform. Theory, IT-38, pp. 713–718.

I. Daubechies (1992), *Ten Lectures on Wavelets*, SIAM, Philadelphia.

R. DeVore, R. Howard, and C. A. Micchelli (1992), *Optimal nonlinear approximation*, Manuscripta Math., 63, pp. 469–478.

R. DeVore, B. Jawerth, and B. J. Lucier (1992), *Image compression through transform coding*, IEEE Trans. Inform. Theory, IT-38, pp. 719–746.

R. DeVore, B. Jawerth, and V. A. Popov (1992), *Compression of wavelet decompositions*, Amer. J. Math., 114, pp. 737–785.

D. Donoho, I. Johnstone, G. Kerkyacharian, and D. Picard (1995), *Wavelet shrinkage: Asymptotia*, J. Roy. Statist. Soc. Ser. B, 57, pp. 301–369.

M. Farge, E. Goirand, Y. Meyer, F. Pascal, and V. Wickerhauser (1992), *Improved predictability of two-dimensional turbulent flows using wavelet packet compression*, Fluid Dynamics Research 10, pp. 229–250.

I. C. Gohberg and M. G. Krein (1965), *Opérateurs linéaires non auto-adjoints dans un espace hilbertien*, Dunod, Paris.

Ky Fan (1951), *Maximum properties and inequalities for the eigenvalues of completely continuous operators*, Proc. Nat. Acad. Sci. USA, 37, pp. 760–766.

Y. Linde, A. Buzo, and R. M. Gray (1980), *An algorithm for vector quantizer design*, IEEE Transactions on Communications, COM-28, pp. 84–95.

S. Mallat (1989), *A theory for multiresolution signal decomposition: The wavelet representation*, IEEE Trans. on Pattern Analysis and Machine Intelligence, PAMI-11, pp. 674–693.

S. Mallat and Z. Zhang (1993), *Nonlinear adaptive time-frequency decomposition*, in Progress in Wavelet Analysis and Applications, Y. Meyer and S. Roques, eds, Frontières, Paris.

Y. Meyer (1990), *Ondelettes et opérateurs*, Hermann, Paris.

P. Oswald (1990), *On the degree of nonlinear spline approximation in Besov-Sobolev spaces*, J. Approx. Theory, 61, pp. 131–157.

J. Shapiro (1993), *Embedded image coding using zerotrees of wavelet coefficients*, IEEE Transactions on Signal Processing, SP-41, pp. 3445–3462.