# CHARACTERIZATION OF SIGNALS BY THE RIDGES OF THEIR WAVELET TRANSFORMS

René A. Carmona * Wen L. Hwang*†
Department of Mathematics
University of California at Irvine
Irvine, CA 92717, USA

Bruno Torrésani
CPT, CNRS-Luminy, Case 907
13288 Marseille Cedex 09
FRANCE

March 22, 1995

**Abstract**

Many procedures have been developed to characterize a signal by some of the salient features of a specific transform, whether it is the Fourier, the Gabor, the Wigner or more recently the wavelet transform. Ridges in the modulus of the transform determine regions in the transform domain with a high concentration of energy. For this reason they are regarded as natural candidates for the characterization and the reconstruction of the original signal. We present a couple of new algorithmic procedures for the detection of ridges in the modulus of the (continuous) wavelet transform of one-dimensional signals. These detection procedures are shown to be robust to additive noise. We also derive and test a new reconstruction procedure. The latter uses only information of the restriction of the wavelet transform to a sample of points from the ridge. This provides with a very efficient way to code the information contained in the signal.

*Partially supported by ONR N00014-91-1010
†Supported by NSF IBN 9405146

## 1 Introduction and Notations

The characterization and the separation of amplitude and frequency modulated signals is a classical problem of signal analysis and signal processing. Applications can be found in many situations, such as for instance radar/sonar detection [2] and speech processing [14] [15]. It has been shown by various authors that time-frequency methods (linear methods such as wavelet or short time Fourier transforms, as well as bilinear methods such as Wigner distributions) can provide satisfactory answers, at least for large values of the signal to noise ratio. The main feature of these methods is the use of the time-frequency localization properties of the representations. They are mainly based on the following heuristics: *Let $f(x) = A(x)\cos\phi(x)$ be a real-valued function. Time-frequency representations of $f(x)$ happen to "concentrate" in regions of the time-frequency plane determined by the instantaneous frequency $\phi(x)$.* It is then natural to use the time-frequency localization of these representations to obtain a characterization of this class of signals. Many methods have been proposed in the past few years to analyze the time-frequency localization of signals. The most noticeable are the family of bilinear representations such as the Wigner representation and its generalizations (see Flandrin's book [9] for a review), and the linear representations such as wavelet or Gabor transforms.

An algorithm was proposed in 1990 by the group of Marseille (see [8] and [19] for a survey), based on the study of the phase of the wavelet (or Gabor) transform. The present work is an attempt to extend the "Marseille algorithm" to noisy situations. Numerous potential applications of such methods can be quoted, such as for instance radar and sonar detection, speech denoising or many problems of physical measurements. The main idea is here to use the localization properties of the modulus of the transform. In the case of frequency-modulated signals, the wavelet transform is "concentrated" in the neighborhood of curves (the ridges of the wavelet transform). In the presence of noise, the locations of the maxima of the modulus of the wavelet transform are more difficult to determine because of the presence of additional local maxima created by the noise. To determine the ridges, we propose two Monte-Carlo methods combining the wavelet transform with a simulated annealing algorithm. Since the pioneering work of the Geman brothers [10] simulated annealing has been used as a method of choice for the Bayesian reconstruction of signals when a Markov model is appropriate. Details on the convergence of the algorithm can be found in [11] and the interested reader is referred to the monograph [13] for an historical review, combinatorial optimization applications and more recent references. We also propose a stable method for signal reconstruction from the numerically computed ridges. This method is also based on an $L^2$-minimization procedure.

For the sake of simplicity, our discussion will be restricted to the case of the wavelet transform. But it is important to notice that since our algorithms deal only with post-processing of time-frequency transforms, they can be extended to any time-frequency energetic representations. The case of the Gabor transform will be considered in the companion paper [5] where still another stochastic search algorithm will be introduced.

We close this introduction with a short summary of the contents of the paper. The short

Section 2 below is devoted to the heuristic description of the signals we consider and to the mathematical definition of the ridges of their wavelet transform, Section 3 begins with the definition of our first penalization algorithm. For the sake of motivation we give a Bayesian interpretation of this approach to the problem of ridge estimation. Once this problem is reduced to a combinatorial optimization problem we review the principles of the stochastic relaxation approach to minimization problems and we give the details of the implementation of the simulated annealing algorithm to our ridge estimation problem. We also described how one can modify the penalization functional in order to accommodate the presence of an additive noise in the signal. This section ends with a discussion of an example of a bat sound signal which we embed in noise (other examples on academic signals are discussed in [6]). Section 4 is devoted to another ridge estimation procedure. The class of ridges which we consider is less restrictive than in the previous section. A new penalization method is proposed and the details of the minimization algorithm are given. The results are produced by this ridge estimation method are illustrated on the same numerical example. Section 5 deals with the reconstruction problem. Given a set of points in the time-scale domain, find the signal most likely to have a ridge going through these points. As before we give the details of the mathematical derivations and we illustrate the efficiency of the method on a numerical example. The paper ends with a short Section 6 which discuss briefly the advantages and the shortcomings of the procedures which we introduced.

## 2 Ridges

The first goal of this section is to set up an abstract formalism for the mathematical definition of the ridges of functions of two variables. Then, we propose two Monte Carlo algorithms to detect and identify these ridges. Applications of these ideas are given in the subsequent sections.

Let us first describe our notations and conventions. We shall work with the complex Hilbert space $L^2(\mathbb{R})$ of square-integrable functions. Our convention for the Fourier transform is the following:

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-i\xi x} dx \tag{1}$$

and consequently the Plancherel formula reads $||\hat{f}||^2 = 2\pi ||f||^2$. As explained earlier, we concentrate throughout this paper on the case of the continuous (complex-valued) wavelet transform. If $\psi \in L^1(\mathbb{R})$ is such that:

$$0 < c_\psi = \int_0^\infty |\hat{\psi}(\xi)|^2 \frac{d\xi}{\xi} < \infty, \tag{2}$$

*i.e.* fulfills the wavelet admissibility condition, then the wavelet transform of a signal $f(x)$ with respect to $\psi$ is given by:

$$T_f(b,a) = \frac{1}{a} \int_{-\infty}^{\infty} f(x) \overline{\psi\left(\frac{x-b}{a}\right)} dx. \tag{3}$$

We will be interested in complex-valued wavelets, providing description of the complex Hardy space $H^2(\mathbb{R}) = \left\{ f \in L^2(\mathbb{R}),\ \hat{f}(\xi) = 0\ \forall \xi \leq 0 \right\}$ and we will set:

$$\psi_{(b,a)}(x) = \frac{1}{a}\psi\left(\frac{x-b}{a}\right). \tag{4}$$

The wavelet transform can be inverted as follows. If $f \in H^2(\mathbb{R})$:

$$f = \frac{1}{c_\psi}\int_0^\infty \int_{\mathbb{R}} T_f(b,a)\psi_{(b,a)}\, db \frac{da}{a}. \tag{5}$$

The case of real-valued signals is handled similarly (a real-valued signal is completely characterized by its positive frequencies). If $\psi \in H^2(\mathbb{R})$, then we may write:

$$f = \mathrm{Re}\, \frac{1}{c_\psi}\int_0^\infty \int_{\mathbb{R}} T_f(b,a)\psi_{(b,a)}\, db \frac{da}{a}. \tag{6}$$

The following discussion deals only with a non-noisy situation for we can better describe the localization properties of the time-frequency representations of the signals in this case. Even though we shall only discuss the wavelet case here, we emphasize once more that all the arguments can be applied to the short-time Fourier case and to the case of any linear time-frequency distribution (such as for instance those described in [17] or in [20]). The only restriction (as far as the interpretation is concerned) is that the quantity under investigation must be thought of as a " time-frequency energy density". This is clearly true for the square-modulus of the wavelet transform, thanks to the "wavelet Plancherel formula" $||T_f||^2_{L^2(\mathbb{R}_+ \times \mathbb{R})} = c_\psi ||f||^2$ for $f \in L^2(\mathbb{R})$. It is also true for the square modulus of the Gabor transform (see [5]) and of the different types of Wigner representations.

Let us then consider a signal of the form:

$$f(x) = A(x)e^{i\phi(x)}, \tag{7}$$

and let us assume that its amplitude $A(x)$ and instantaneous frequency $\nu(x) = \phi'(x)/2\pi$ are slowly varying. At this point we restrict ourselves to complex valued analytic signals. We explain below why real-valued signals can be processed in the same way. Let $\psi(x)$ be an analytic complex valued wavelet, written in the form:

$$\psi(x) = A_\psi(x)e^{i\phi_\psi(x)}. \tag{8}$$

We assume that the wavelet $\psi(x)$ has "good" localization properties in the Fourier domain, for instance that $\hat{\psi}(\xi) \sim O(|\xi|^M)$ as $\xi \to 0$, and $\hat{\psi}(\xi) \sim o(|\xi|^{-K})$ as $\xi \to \pm\infty$ for some positive constants $K$ and $M$. In practice, we shall use wavelets whose Fourier transform $\hat{\psi}(\xi)$ is peaked near a particular value $\xi = \omega_0$ of the frequency.
It then follows from standard arguments (see [8]) that the wavelet transform may be approximated as

$$T_f(b,a) \approx A(b)e^{i\phi(b)}\hat{\psi}(a\phi'(b))^* + O(|A'|, |\phi''|). \tag{9}$$

From the localization properties of the wavelet in the Fourier domain one can see that the modulus $|T_f|$ of the wavelet transform is essentially maximum in the neighborhood of a curve $a = \varphi(b)$, called the *ridge of the wavelet transform*, which is related to the instantaneous frequency of the signal by the formula:

$$\varphi(b) = \frac{\omega_0}{\phi'(b)}. \tag{10}$$

Characterizing the signal's instantaneous frequency by the wavelet transform can then be achieved by extracting numerically the ridge as a set of local extrema of the modulus of the wavelet transform. Subsequently, whenever we want to raise the level of generality of our discussion, we consider a nonnegative function $M(b,a)$ of the couple $(b,a) \in \mathbb{R} \times [0,\infty)$. We should think of this function as the modulus of the wavelet transform of a finite energy one dimensional signal, i.e. $M(b,a) = |T_f(b,a)|^2$. We shall always assume that this function is smooth and that it has a single ridge in the sense that the local maxima of the functions of $a$ (when the variable $b$ is held fixed) $a \hookrightarrow M(b,a)$ are located on the graph of a smooth function $b \hookrightarrow \varphi(b)$ defined on a (possibly strict) subset of the domain of the variable $b$.

In addition, once the ridge is known, the instantaneous amplitude can be (approximately) recovered from the restriction of the wavelet transform to the ridge (the so-called "skeleton of the wavelet transform") via equation (9).

In the case of real-valued signals, the goal is to describe them in the form:

$$f(x) = A(x)\cos\phi(x). \tag{11}$$

In this case it is crucial to use the so-called "progressive wavelets", *i.e.* wavelets with vanishing negative frequencies. If $\psi(x)$ is such a wavelet, then the corresponding wavelet coefficients of $f(x)$ are given by:

$$T_f(b,a) = \langle f, \psi_{(b,a)}\rangle = \frac{1}{2}\langle Z_f, \psi_{(b,a)}\rangle \tag{12}$$

where $Z_f(x)$ is the "analytic signal" of $f(x)$, i.e. its orthogonal projection onto the complex Hardy space $H^2(\mathbb{R})$. It is given by:

$$Z_f(x) = \frac{1}{\pi}\,\mathrm{P.}\int f(x+y)\frac{dy}{y}. \tag{13}$$

It is well-known that a signal of the form (11) with $A(x)$ and $\phi'(x)$ slowly varying gives. This result is sometimes referred to as Bedrosian's theorem.

$$Z_f(x) \approx A(x)\cos\phi(x). \tag{14}$$

The main idea of the "Marseille" algorithm was to use the phase coherence of the wavelet transform to get a numerical estimate of the ridge (see [8] for more details). Since the phase can be somewhat difficult to control in noisy situations, we shall mainly focus here on the localization of the maxima of the modulus of the wavelet transform.

## 3 Ridge Detection by Annealing

### 3.1 General Description of the Algorithm

Throughout this section we assume that the wavelet transform of the signal $f$ which we analyze has only one ridge and that this ridge can be parametrized by a function $b \hookrightarrow \varphi(b)$ defined for all the values of $b$. For the sake of the present discussion we denote by $\Phi$ the space of all the twice differentiable functions with square integrable derivatives. We then define the penalty function $F_f$ on the set $\Phi$ of ridge candidates $\varphi$ by:

$$F_f(\varphi) = -\int |T_f(b, \varphi(b))|^2 db + \int \left[\lambda \varphi'(b)^2 + \mu \varphi''(b)^2\right] db \tag{15}$$

Our estimate of the unknown ridge of the wavelet transform of the signal $f$ will be the function $a = \varphi(b)$ which minimizes $F_f(\varphi)$. The Euler equation associated with this minimization problem reads:

$$\mu \varphi''''(b) - \lambda \varphi''(b) = 2\mathrm{Re}\left[T_f(b, \varphi(b)) \partial_a \overline{T_f(b, \varphi(b))}\right]. \tag{16}$$

Such an equation may be solved numerically: it is sufficient to discretize the $b$ variable, to replace derivatives by finite differences, and then to use standard numerical techniques. However, although in the non-noisy case one may expect to reach a global minimum or a reasonable local minimum, the noisy case is much more problematic. Indeed, the presence of noise implies the existence of many local extrema in which the algorithm may get trapped. We need a procedure which can jump over the local extrema to reach the global one(s). A natural candidate for this is the simulated annealing algorithm [10].

### 3.2 Bayesian Interpretation

It is possible to give a Bayesian interpretation of the estimation of the ridge based on the minimization of (15). We give it here for the sake of illustration. Our argument is perfectly rigorous in the case of (finite) discrete signals. It requires sophisticated measure theoretic arguments in the continuous case. We refrain from providing the details even though we opted to present our discussion with the convenient notations of integration and differentiation for continuous signals.

Let us introduce a probability measure $\mu_{prior}(d\varphi)$ to model the prior knowledge (or lack of knowledge) of the ridge in the space $\Phi$. We set:

$$\mu_{prior}(d\varphi) = \frac{1}{Z_1}\; e^{-\int\left[\mu_a|\varphi'(b)|^2 + \lambda_a|\varphi''(b)|^2\right]db} \quad {''}d\varphi{''} \tag{17}$$

where the normalization constant $Z_1$ is chosen so that $\mu_{prior}$ is a probability measure. The infinitesimal $''d\varphi''$ has an obvious (rigorous) meaning in the case of a discrete variable $b$. The measure $\mu_{prior}$ has to be defined as a Gaussian probability measure on an infinite dimensional space. Given an element $\varphi \in \Phi$, let us consider the probability $\mu_\varphi(df)$ defined by:

$$\mu_\varphi(df) = \frac{1}{Z_2}\; e^{\int |T_f(b,\varphi(b))|^2} \quad ''df''$$

on the space $\mathcal{S}_2$ of signals of finite energy. As before, $Z_2$ is a normalization constant whose role is to guarantee that the total mass of the measure is 1. Intuitively speaking, $\mu_\varphi(df)$ gives the probalility, knowing that the ridge is given by the function $\varphi$, that the true signal is in the infinitesimal $''df''$ in the space $\mathcal{S}_2$. Notice that the probability $\mu_\varphi(df)$ is large when the modulus of the wavelet transform of $f$ is large along the function $\varphi$. This is consistent with the intuitive interpretation of this probability.

Given an observation of a signal $f$ of finite energy, the posterior probability of the ridge is given by the classical Bayes formula:

$$\mu_{posterior}(d\varphi) = \frac{1}{Z_f}\; e^{-[\int\left(\mu_a|\varphi'(b)|^2+\lambda_a|\varphi''(b)|^2\right)db - \int |T_f(b,\varphi(b))|^2 db]} \quad ''d\varphi''$$

for some normalizing constant $Z_f$ depending on the observation $f$. The maximum aposteriori likelihood principle suggests to choose the value of $\varphi$ which maximizes the *density* of this distribution. In other words, given the observation $f$, the estimate $\hat{\varphi}$ of the ridge should minimize (15).

### 3.3 Combinatorial Optimization by Simulated Annealing

Let us briefly review the basics of the simulated annealing method (see [13] and [11] for more details).

We consider the problem of the minimization of a function $F$ over a finite set $\Phi$. We shall denote by $\varphi$ the elements of the set $\Phi$. We denote by $F_{min}$ the value of the minimum:

$$F_{min} = \inf_{\varphi\in\Phi} F(\varphi)$$

and by $\Phi_{min}$ the set of $\varphi \in \Phi$ at which the minimum is attained. We denote by $\nu(d\varphi)$ the normalized counting measure on $\Phi$. In other words, for each subset $A \subset \Phi$, $\nu(A)$ is given by $\nu(A) = |A|/|\Phi|$, the ratio of the number of elements of $A$ to the total number of elements in $\Phi$. We also assume that each $\varphi \in \Phi$ has a specific set $N(\varphi)$ of neighbors and we use the structure given by the neighborhood system and the function $F$ to define several random walks (Markov chains to be more specific) on the set $\Phi$. We assume (and this is our last assumption) that the system of neighborhoods is such that the steady state of the canonical random walk is the normalized counting measure. In other words, if $\{X_n\}_{n\geq 0}$ is a sequence of random variables such that, given $X_n$, $X_{n+1}$ is chosen independently of the entire past $X_0, X_1, \cdots, X_n$ uniformly (i.e. with probability $1/|N(X_n)|$) among the neighbors of $X_n$, then the distribution of $X_n$ converges (irrespectively of the initial value of $X_0$) toward the probability distribution $\nu(d\varphi)$.

The canonical random walk was merely introduced to state our assumption on the neighborhood system. We shall in fact consider different "random walks".
For each $T > 0$ (one should think of $T$ as of a temperature) we consider the following perturbation of the canonical random walk. Given $X_n$ one defines the (random) transition to the next state $X_{n+1}$ in the following way. One considers first the state $X'_{n+1}$ uniformly distributed among the neighbors $N(X_n)$ of $X_n$ independently of the past. In other words $X'_{n+1}$ is generated by the random transition mechanism of the canonical random walk. If $F(X'_{n+1}) \leq F(X_n)$ we set $X_{n+1} = X'_{n+1}$. If $F(X'_{n+1}) > F(X_n)$ then we set $X_{n+1} = X'_{n+1}$ with probability $\exp[-(F(X'_{n+1}) - F(X_n))/T]$ and we set $X_{n+1} = X_n$ with probability $1 - \exp[(F(X_{n+1}) - F(X_n))/T]$. In words, the state changes according to the same random mechanism as for the canonical random walk as long as the function $F$ does not increase. Each time an adverse move is attempted to increase the value of $F$, it is accepted with a probability which depends upon the temperature $T$ and the amount by which the function $F$ would increase. $\{X_n\}_{n\geq 0}$ so defined is not a random walk because its increments are no longer independent of the past. But it is still a Markov chain and it is easy to check that its steady state is the probability measure $\nu_T(d\varphi)$ defined by the formula:

$$\nu_T(d\varphi) = \frac{1}{Z(T)} e^{-F(\varphi)/T} \nu(d\varphi) \tag{18}$$

where the constant $Z(T)$ is chosen so that $\nu_T(d\varphi)$ is a probability measure. The definition (18) of the limiting distribution $\nu_T$ is amenable to the asymptotic regime of the Laplace method. Indeed, a simple argument shows that, when the temperature parameter $T$ is decreased to 0 the measure $\nu_T(d\varphi)$ converges toward the measure $\nu_0(d\varphi)$ which is uniformly distributed over the set $\Phi_{min}$ of (global) minima of the function $F$. In particular, $\nu_T(d\varphi)$ converges toward the unit mass at the point $\varphi_{min}$ when the minimum is attained at a single point, say $\varphi_{min}$.

### 3.4 The Ridge Annealing Algorithm

We now come back to the set up of Subsection 3.1 and to the problem of the minimization of the functional $F_f(\varphi)$ defined in equation (15). The first step is a discretization of the problem. Assuming that we start from a wavelet transform defined for the discrete grid $[0, 1, ..., B-1] \times [0, 1, ..., A-1]$ in the time-scale plane, the ridge will take the form of a finite sequence $\{\varphi(0), ..., \varphi(\ell), ...\varphi(n-1)\}$. We need to specify a neighborhood system for our discretized ridges. The neighbors we shall consider are the simplest possible ones. The set $N(\varphi)$ of neighbors of $\varphi$ is the set of finite sequences $\{\varphi(0), ..., \varphi(\ell) \pm 1, ...\varphi(n-1)\}$ obtained by varying $\ell$ from 0 to $n-1$. In other words, two neighboring ridge $\varphi$ and $\varphi'$ can only differ at one point, say $\ell$, and $\varphi'(\ell) = \varphi(\ell) \pm 1$. The simulated annealing algorithm goes as follows.

1. *Initialization:* Fix a value for the temperature: $T_0 = C/\ln(2)$. Start from an initial guess for the ridge: $\{\varphi_0(0), \varphi_0(1), ...\varphi_0(n-1)\}$ and compute the value of the cost function $F_f(\varphi_0)$.

2. *Step k:* Suppose that the ridge at step $k-1$ is $\varphi_{k-1} = \{\varphi_{k-1}(0), ..., \varphi_{k-1}(n-1)\}$.
   - Update the temperature:
$$T_k = \frac{C}{\ln(1+k)} \tag{19}$$
   - Select randomly an integer $\ell \in [0, n-1]$ and a number $\varepsilon = \pm 1$. The candidate $\varphi_k^c$ for step $k$ is then $\{\varphi_k^c(0), ..., \varphi_k^c(\ell) + \varepsilon, ...\varphi_k^c(n-1)\}$.
   - Compute the value of the cost function $F_f(\varphi_k^c)$ and compare it with $F_f(\varphi_{k-1})$.
   - If $F_f(\varphi_k^c) \leq F_f(\varphi_{k-1})$ update the ridge: the new ridge at step $k$ is : $\varphi_k = \varphi_k^c$.
   - If $F_f(\varphi_k^c) > F_f(\varphi_{k-1})$ pick a random number $\sigma$ uniformly distributed between 0 and 1.
     - If $\sigma \leq \exp\{-\left[F_f(\varphi_k) - F_f(\varphi_{k-1})\right]/T_k\}$ update the ridge: $\varphi_k = \varphi_k^c$.
     - otherwise keep the previous ridge: $\varphi_k = \varphi_{k-1}$.
3. *Stopping criterion:* When the ridge has not changed for a certain number of steps (fixed in advance), stop the iteration.

**Remark:** It is important to keep in mind the complexity of the minimization problem. Let us consider for example the case of a wavelet transform with $B = 1024$ values of the position and $A = 50$ values of the scale. The search for a global minimum is done in a set with $50^{1024}$ elements. It may then be useful to reduce the size of such a space, for instance by "subsampling" the ridge. A way of achieving such a subsampling by first smoothing the wavelet transform modulus in the $b$-direction (or in the $b$ and $a$ directions if needed), and then running the algorithm on a subsampled version of the smoothed wavelet transform modulus. For instance, smoothing with a window of length 10 and subsampling by a factor of 10 reduces the size of the space to approximately $50^{103}$ elements, which very often makes a significant difference for the convergence of the algorithm.

### 3.5 Case of Noisy Signals

In many applications the observed signal $f(x)$ appears as the sum of a pure component $f_0(x)$ and a noise component $n(x)$. There are situations in which some "a-priori" knowledge of the noise is available. Let us quote for instance the cases where the power spectrum of the noise is known, or the cases where a piece of the signal is known to contain only noise (which gives us the chance to learn about the statistics of this noise). Then in order that the computed ridge is not attracted by regions where the noise dominates, one can "renormalize" the wavelet transform, by subtracting what is supposed to be the "typical" contribution of the noise. This contribution could be chosen to be the expectation $\mathbb{E}[|T_n(b,a)|^2]$ over all the possible realizations of the additive noise $n$. If an a-priori model for the noise is available, such a quantity may be estimated by Monte-Carlo simulations, or sometimes by a direct computation For instance, in the case of a second order stationary noise, with power

spectrum of the form $p(\xi) = \sigma^2\xi^\alpha$, it is easy to derive that $\mathbb{E}[|T_n(b,a)|^2] \sim K_\alpha \sigma^2 a^{-\alpha-1}$; for this the wavelet $\psi(x)$ has to be such that $K_\alpha = \int u^\alpha |\hat{\psi}(u)|^2 du < \infty$.
But in the general case, we only have one realization of the noise at hand, and it is impossible to compute directly this expectation. A simple ergodic argument justifies the use of the estimate:

$$V(a) = \frac{1}{B}\int_0^B |T_n(b,a)|^2 db. \tag{20}$$

Then the cost function given in equation (15) may be replaced by:

$$\tilde{F}_f[\varphi] = -\int \left[|T_f(b,\varphi(b))|^2 - V(\varphi(b))\right] db + \int \left[\lambda\varphi'(b)^2 + \mu\varphi''(b)^2\right] db \tag{21}$$

The effect of such a modification is to avoid "trapping" the ridge in regions dominated by the noise.

### 3.6 Numerical Results

The algorithm was tested on various frequency modulated (real or academic) signals embedded in noise. We present here results concerning the case of one type of (real) signal, namely the sonar signal emitted by a given species of bats (eptesicus fucus). The signal is presented in Figure 3 (actually only 500 samples at the beginning of the signal are displayed, in order to highlight the time-dependence of the frequency). The instantaneous frequency of such a signal is known to be well approximated by an hyperbolic function of time (i.e. a frequency of the form $\nu(t) = \nu_0 \log(1+\alpha t)$). We used Morlet's wavelet (i.e. $\psi(x) = \exp\{-x^2/2\}\exp\{i\omega_0 x\}$) with $\omega_0 = 2\pi$ with the scale varying over 3 octaves with 20 different scales per octave).
The annealing algorithm has been tested on pure and noisy forms of this signal; the noise was an additive Gaussian white noise, and the noise potential $V(a)$ was directly estimated from the flat spectrum of the noise ($V(a) = \sigma^2 a^{-1}$ in our case). The method was found efficient for values of the signal to noise ratio up to $-5dB$. See Figure 1.
The case of the pure signal was easy to process for the algorithm. In such a case, the result basically does not depend on the values of the $\lambda$ and $\mu$ parameters, since the localization properties of the modulus of the wavelet transform are sufficient to ensure the convergence to the correct ridge, for any reasonable value of $\lambda$ and $\mu$. In addition, no smoothing is necessary. However in the example described in Figure 1 the signal length was about 1000 samples, and a "smoothing+subsampling" with a window of length 10 was used (as stressed before, the "smoothing + subsampling" procedure reduces the complexity of the algorithm, and is then useful, even in a non-noisy situation); the same method without subsampling also converges to the same ridge, but requires more iterations. The penalty function is described in Figure 1. The number of iterations needed to converge was about 4000. The initial guess $\varphi_0$ was taken to be a linear function such that $\varphi_0(0) = 4$ and $\varphi_0(1000) = 5$ (see Figure 1 to compare the numerical results), i.e. close to the actual ridge. Different values for the initial guess yield the same ridge, but the number of iterations needed to converge

grows as a function of the distance between the guess and the actual ridge. For instance, starting from a $\varphi_0$ uniformly equal to 8 needs about 75000 iterations to converge.
In the noisy cases, the values of $\lambda$ and $\mu$ play an important role, and in the examples described here (as well as other examples on the same signal, with different values of the $SNR$), $\lambda \sim 100$ and $\mu \sim 50$ seemed to be good values. Let us stress at this point that the values of the parameters depend on the normalization of the signals. They have to be adjusted as functions of the variance of the signal. In our case, the variance of the pure signal equals 262970.

The same experiment was performed with other values of the $SNR$. In all cases, a logarithmic temperature schedule was chosen as in (19) and we used $C = 3$. The number of iterations was about 1500 in all cases (less than in the non-noisy case, because the subsampling was more important), and a "smoothing + subsampling" over windows of 50 samples was chosen (of course, the larger the $SNR$, the smaller the window needed; we have chosen to present here results obtained with a constant window size). As before, the initial guess $\varphi_0$ was also taken to be a linear function such that $\varphi_0(0) = 4$ and $\varphi_0(1000) = 5$. Finally, we used $\lambda = 100$ and $\mu = 50$.
In Figure 1, we also compare the ridges obtained with the annealing approach for the 800 intermediate samples of the signal, together with that obtained in the non-noisy case with the Marseille algorithm. The agreement is found to be very good.
The case of signals associated with multiple ridges is more difficult to process, the function $\varphi(b)$ being single-valued by construction. However, in such cases, several runs of the algorithm, starting from different initial guesses $\varphi_0(b)$ may yield the different ridges. In the case of the bat signal, the wavelet transform modulus is highly concentrated in a neighborhood of the ridge we obtained, but there is another region in which the modulus is relatively large. This region corresponds to the first harmonic of the signal (the previous ridge corresponding to the fundamental mode). Running our algorithm in a smaller region of the signal where the first harmonic is significantly large, we can detect this secondary ridge with a different initial guess. We used again a linear function $\varphi_0(b)$ with $\varphi_0(350) = 2.37$ and $\varphi_0(1000) = 3.36$) and the same values as before for the parameters. The convergence occurred after about 2000 iterations.

# 4 Snake Ridges

Let us first point out some of the limitations of the annealing approach described above:

- The procedure does not allow for "vertical ridges" such as the ridges occuring in the case of singular signals. Although such signals are not in the class of signals we are interested in, one still may want to generalize the approach to include such cases. Unfortunately, the ridge annealing algorithm is not suitable for such a purpose.
- The boundaries of the ridge are not handled in a systematic way: since one is looking for ridges of the form $a = \varphi(b)$, the boundary values of $b$ have to be fixed in advance (at least in the current form of the algorithm). This implies that one has to know

in advance the time duration of the signal under consideration (which is not a trivial problem in the noisy situation), or to introduce in the algorithm a post-processing stage to determine the actual location of the amplitude and frequency component under consideration.

- The procedure is ill suited for signals having multiple ridges, such as for instance speech signal (the function $\varphi$ must of course be single-valued).

For these reasons, we present an alternative algorithm based on the ideas of "snakes" developed in [12]. In this second method the boundaries of the ridge adjust dynamically to the boundaries of the unknown ridge of the signal. Also this new algorithm can handle more complicated ridges. In addition, when combined with segmentation methods such as those proposed in [1] it could also handle multiple ridges. The new algorithm is more adaptive, but the price to pay is a loss of stability, and the latter can be significant in very noisy situations.

### 4.1 Snake Energy Penalization

We now consider a ridge as a parametrized curve $r(s) = [\rho_1(s), \rho_2(s)]$ in the time-frequency plane where $s \in [0,1]$ is a dummy variable parametrizing the ridge. The ridge then takes the form of a "snake" in the time-scale plane. To localize the snake near the location of the maxima of the modulus of the wavelet transform, we introduce a new cost function. The latter takes into account the modulus of the wavelet transform, as well as some additional terms needed in order to ensure the smoothness of the ridge (now understood as a smoothness requirement both in the $b$ and $a$ directions). We set:

$$\begin{aligned} F_f(r) = \quad & -\int |T_f\,(\rho_1(s), \rho_2(s))\,|^2 ds \\ & + \int [\lambda_a \rho_2'(s)^2 + \mu_a \rho_2''(s)^2 + \lambda_b \rho_1'(s)^2 + \mu_b \rho_1''(s)^2]\ ds, \end{aligned} \quad . \tag{22}$$

where $\lambda_a$, $\lambda_b$, $\mu_a$ and $\mu_b$ are positive constants. In the "snake terminology" of [12] the second term (containing the derivatives of $\rho_1$ and $\rho_2$) is the "internal energy" of the snake. Its role is to control the smoothness and the rigidity of the snake. The first term is the "external energy" of the snake. It accounts for the interaction of the snake with the wavelet transform modulus.

**Remark:** As before, when some spectral information on the noise is available, it can be incorporated into the cost function in the form of a renormalization potential $V(a)$. We used the form:

$$\begin{aligned} \tilde{F}_f(r) = \quad & -\int [|T_f\,(\rho_1(s), \rho_2(s))\,|^2 - V(\rho_2(s))]\,ds \\ & + \int [\lambda_a \rho_2'(s)^2 + \mu_a \rho_2''(s)^2 + \lambda_b \rho_1'(s)^2 + \mu_b \rho_1''(s)^2]\ ds. \end{aligned} \tag{23}$$

The Euler equations (there are now two such equations) read:

$$\left\{ \begin{aligned} \mu_b \rho_1''''(s) - \lambda_b \rho_1''(s) \quad &= 2\mathrm{Re}\left[T_f(\rho_1(s), \rho_2(s)).\partial_b \overline{T_f(\rho_1(s), \rho_2(s))}\right] \\ \mu_a \rho_2''''(s) - \lambda_a \rho_2''(s) \quad &= 2\mathrm{Re}\left[T_f(\rho_1(s), \rho_2(s)).\partial_a \overline{T_f(\rho_1(s), \rho_2(s))}\right] \end{aligned} \right. \tag{24}$$

An iterative scheme was proposed in [12] for the numerical solution of such a system. It has been widely used since then (see for example [1]). The problem of such a scheme is the same as before: it is efficient for smooth data, but not in the case of noisy signals. For that reason we turn again to simulated annealing procedures.

### 4.2 Cost Minimization by Simulated Annealing

As before, we rely on a discretization procedure. A discrete snake is a collection of two-dimensional vectors:

$$r = (\rho_1, \rho_2) = \{(\rho_1(0), \rho_2(0)), \cdots, (\rho_1(n-1), \rho_2(n-1))\}. \tag{25}$$

The integral in the cost function $F_f(r)$ is naturally replaced by a finite sum, taking into account the values of the snake at the nodes $0, 1, \ldots n-1$. Starting from an initial guess $r_0 = (\rho_{1,0}, \rho_{2,0})$, we update the ridge using the neighborhood system defined as follows: two neighboring ridges $r = (\rho_2, \rho_2)$ and $r_1' = (\rho_1', \rho_2')$ can only differ at one point $\ell$ at most and in such a way that:

- $\rho_2'(\ell) = \rho_2(\ell) \pm \delta_v$ and $\rho_1'(\ell) = \rho_1(\ell)$, or
- $\rho_2'(\ell) = \rho_2(\ell)$ and $\rho_1'(\ell) = \rho_1(\ell) \pm \delta_h$

where $v$ and $h$ are positive constants giving the sized of the steps. More precisely, the general algorithm goes as follows. We give the details of the algorithm in the case of the cost function $F_f(r)$ defined in (22). The same procedure can be used with the cost function $\tilde{F}_f(r)$ defined in formula (23).

1. *Initialization:* Fix an initial value for the temperature, say $T_0 = C/\ln(2)$. Start from an initial guess for the ridge:

   $$r_0 = \{(\rho_{1,0}(0), \rho_{2,0}(0)), (\rho_{1,0}(1), \rho_{2,0}(1)), \cdots, (\rho_{1,0}(n-1), \rho_{2,0}(n-1))\}$$

   and compute the corresponding value $F_f(r_0)$ of the cost function.

2. *Step k:* Suppose that the estimate of the ridge:

   $$r_{k-1} = \{(\rho_{1,k-1}(0), \rho_{2,k-1}(0)), (\rho_{1,k-1}(1), \rho_{2,k-1}(1)), \cdots, (\rho_{1,k-1}(n-1), \rho_{2,k-1}(n-1))\}.$$

   at step $k-1$ has been constructed,

   - Update the temperature as in (19).
   - Select randomly an integer $\ell \in [0, n-1]$ and a possible jump: left,right,up or down. Depending on the result, define the increments $\varepsilon_v = 0$ or $\pm\,\delta_v$, $\varepsilon_h = 0$ or $\pm\,\delta_h$. For instance, in the case of a left jump, the vertical increment $\varepsilon_v = 0$ and the horizontal increment $\varepsilon_h = -1$, while $\varepsilon_h = 1$ for a right jump. The candidate for the ridge at step $k$ is then:

     $$r_k^c = \left\{(\rho_{1,k}^c(0), \rho_{2,k}^c(0)), (\rho_{1,k}^c(1), \rho_{2,k}^c(1)), \cdots, (\rho_{1,k}^c(n-1), \rho_{2,k}^c(n-1))\right\}$$

where $\rho^c_{1,k}(j) = \rho_{1,k-1}(j)$ and $\rho^c_{2,k}(j) = \rho_{2,k-1}(j)$ for all $j \neq \ell$, and $\rho^c_{1,k}(\ell) = \rho_{1,k-1}(\ell) + \varepsilon_v$ and $\rho^c_{2,k}(\ell) = \rho_{2,k-1}(\ell) + \varepsilon_h$.

- Compute the value of the cost function $F_f(r^c_k)$ and compare it with $F_f(r_{k-1})$.
- If $F_f(r^c_k) \leq F_f(r_{k-1})$ update the ridge: the new ridge at step $k$ is : $r_k = r^c_k$.
- Otherwise, pick a random number $\sigma$ uniformly distributed between 0 and 1 and:
  - if $\sigma \leq \exp\{-\left[F_f(r^c_k) - F_f(r_{k-1})\right]/T_k\}$ update the ridge: $r_k = r^c_k$.
  - otherwise keep the previous ridge: $r_k = r_{k-1}$.

3. *Stopping criterion:* When the ridge hasn't changed for a certain number of steps (fixed in advance), stop the iteration.

**Remark:** In the previous section, we introduced a "smoothing + subsampling" procedure, in order to stabilize the algorithm by reducing its complexity and smoothing the data. In the case of the "snake" algorithm, the complexity reduction is not needed any more, since it is automatically done by choosing a small number of nodes to describe the snake. However, it is still useful to smooth the wavelet transform modulus, to improve the stability of the algorithm.

### 4.3 Results and Discussion

The snake algorithm was tested on several amplitude and frequency modulated signals. The results are similar to those of the previous annealing algorithm. As before, we used Morlet's wavelet with the scale varying over 3 octaves with 20 different scales per octave. Figure 2 shows the snake estimate obtained for the pitch of the bat sonar signal. The snake is a 20 nodes one, an the initial guess was a linear snake (the same as the one we considered before), with $\varphi_0(0) = 4$ and $\varphi_0(1) = 5$, and $\rho(0) = 1$ and $\rho(1) = 1000$. About 1000 iterations were necessary to converge (the cost function is not displayed here for it has the same behavior as the cost function of the previous algorithm. Other tests with an initial guess further away from the true ridge yield the same result as long as the number of iterations is large. The parameters in the penalty function were set to $\mu = 50$ and $\lambda = 100$, the temperature constant to $C = 3$. The numerical values for the ridge function are in good agreement with those obtained with the other methods. As in the previous algorithm, the ridge estimate depends weakly on the values of the $\lambda$ and $\mu$ "rigidity parameters" (as long as they take "reasonable values"), since the localization properties of the wavelet transform modulus are good enough to ensure the convergence to a correct ridge. In addition, basically no smoothing was necessary.
Figure 2 also gives the snake estimate for the pitch of a noisy bat sonar signal. The noise is again a Gaussian white noise and the SNR is equal to $-5dB$. We used the pre-estimator $V(a) = \sigma^2 a^{-1}$ to introduce some "a-priori" information about the noise in the cost function. Approximately 1200 iterations were needed for the convergence (starting from a uniformly sampled linear initial guess $(\varphi_0, \rho_0)$, analogous to the one used in the annealing case). The parameters in the penalty function were set to $\mu = 100$ and $\lambda = 200$ (the variance of the

pure signal equals 262970). Their role now becomes crucial, since the wavelet transform modulus localization is not good enough to ensure the convergence. The numerical values for the ridge function are in good agreement with those obtained in the non-noisy situation.

# 5 Reconstruction from a Ridge

The purpose of this section is to present a new algorithmic reconstruction of a signal from the knowledge of sample values of its wavelet transform on the ridges of its modulus. For the sake of simplicity we restrict ourselves to the case of a single ridge. See nevertheless [5] for the analysis (in the case of the Gabor transform) of the more general case of finitely many arbitrary ridges. We assume that the ridge can be parametrized by a continuous function $[b_{min}, b_{max}] \ni b \hookrightarrow \varphi(b) \in (0, \infty)$. Such a ridge function is usually difficult to determine. A good starting point could be the result of a ridge estimation procedure such as those presented in Section 3 and 4. In any case, one will presumably have to construct a smooth ridge function from the knowledge of a few points. Indeed, in practical applications one only knows sample points $(b_1, \varphi_1), \cdots, (b_n, \varphi_n)$ and the smooth function $b \hookrightarrow \varphi(b)$ which we use in lieu of the true (unknown) ridge function is merely a guess which one constructs from the sample points. We use a smoothing spline (but any other kind of nonlinear regression curve would do as well). From now on, $\varphi(b)$ is a smooth ridge function which is constructed from the $n$ sample data points.

## 5.1 Statement of the Problem

We are concerned with the implementation of the folk belief that *a signal can be characterized by the values of the restriction of its wavelet transform to its ridges.* Illustrations can be found in [8] and [19], where it is shown that in the case of signals of the form $f(x) = A(x)\cos\phi(x)$ the restriction of the wavelet transform to its ridge (the so-called *skeleton* of the wavelet transform) essentially behaves as $A(x)\exp[i\phi(x)]$. Similar remarks (for the Gabor transform) also yield the so-called *sinusoidal model* for speech (see e.g. [14], [15] and [18]) which achieves good quality reconstruction with high compression rate. Such an approach can be used in non-noisy situations, but it does fail in the presence of a significant noise.

In order to consider a result in this spirit we assume that the values of the wavelet transform of an unknown signal of finite energy $f_0(t)$ are known at sample points $(b_j, \varphi_j)$ which are tentatively regarded as representative of the ridge of the modulus of the wavelet transform of the (unknown) signal $f_0(t)$. We use the notation $z_j$ for the value of the wavelet transform of $f_0$ at the point $(b_j, \varphi_j)$. The set of sample points $(b_j, \varphi_j)$ together with the values $z_j$ constitutes what we call the skeleton of the wavelet transform of the signal to be reconstructed.

As we already explained, we use a smooth function $b \hookrightarrow \varphi(b)$ which *interpolates* the sample points and we look at the graph of this function as our best guess for the ridge of the modulus of the wavelet transform of $f_0$.

We look for a signal $f(t)$ of finite energy whose wavelet transform $T_f(b,a)$ satisfies:

$$T_f(b_j, \varphi_j) = z_j, \qquad\qquad j = 1, \cdots, n \tag{26}$$

and has the graph of the function $\varphi(b)$ as ridge. Recall that this last statement means that for each $b$, the point $(b, \varphi(b))$ of the time-scale plane is a local maximum of the function $a \hookrightarrow |T_f(b,a)|$. We are about to show how to construct such a signal. We will also show that it is a very good approximation of the original signal $f_0(t)$.

### 5.2 The Penalization Approach

Solving the reconstruction problem via the solution of an optimization problem is a time honored approach in statistical data analysis. In the case of the wavelet analysis of signals this approach was successfully used by Mallat and Zhong to reconstruct of a signal from the extrema of its dyadic wavelet transform (see [17] and [3]).

In the present setting, we look for a signal $f(t)$ whose wavelet transform has given values at the sample points of the ridge, i.e. which satisfies the constraints (26) while the $L^2$-norm in the scale variable $a$ of the modulus is kept to a minimum for each $b$. In other words we search for a signal $f(t)$ which minimizes the functional:

$$F_1(f) = \frac{1}{c_\psi} \int db \int \frac{da}{|a|} |T_f(b,a)|^2 \tag{27}$$

while its wavelet transform has the prescribed values at the sample points of the ridge. Notice that when the integration is performed on the whole half-plane, $F_1(f)$ is nothing but $||f||^2$. Solving such a constrained optimization problem is easy. Since the cost function $F_1(f)$ is a quadratic form in the unknown function $f$, the solution is easily computed by means of the use of Lagrange multipliers. See [3] or the discussion below for details. A solution can be constructed as a linear combination of the wavelets $\psi_{(b_j,\varphi_j)}$ at the sample points of the ridge, the coefficients being given by the solution of a $n \times n$ linear system.

This solution is not completely satisfactory, especially when the number of sample points is small. In particular, it ignores the empirical fact that (in most of the practical applications) the restriction of the modulus $|T_f(b,a)|$ to the ridge, i.e. the function $b \hookrightarrow |T_f(b, \varphi(b))|$, is smooth and slowly varying. In order to force the solution of the constrained optimization problem to respect this requirement, we add the following contribution to the cost function:

$$\tilde{F}_2(f) = \int_{b_{min}}^{b_{max}} \left| \frac{d}{db} |T_f(b, \varphi(b))| \right|^2 \, db, \tag{28}$$

and consider the minimization of the cost function:

$$\tilde{F}(f) = F_1(f) + \epsilon \tilde{F}_2(f) \tag{29}$$

where the free parameter $\epsilon > 0$ can be chosen to control the relative importance of the two contributions to the penalty. Unfortunately, $\tilde{F}_2(f)$ is not quadratic in $f$. This makes the problem more difficult to solve. Furthermore, the solution is not given by a linear combination of *reconstruction* wavelets placed at each of the sample points of the ridge. In order to remedy this problem we replace the contribution $\tilde{F}_2(f)$ by a quadratic form which gives a good approximation to $\tilde{F}_2(f)$ when the signal belongs to one of the classes of signals likely to be characterized by their skeletons on their ridges (see for example [8] and [19]). Let us denote by $\Omega_f(b,a)$ the phase of the wavelet transform $T_f(b,a)$ (i.e. the argument of the complex number $T_f(b,a)$). We have:

$$T_f(b,a) = M_f(b,a)e^{i\Omega(b,a)}.$$

For the purpose of the present discussion we assume that a smooth version of the phase exists and that such a choice was made for $\Omega_f(b,a)$. Notice that, with these notations we have:

$$\frac{d}{db}T_f(b,\varphi(b)) = \left(\frac{d}{db}M_f(b,\varphi(b))\right)e^{i\Omega(b,\varphi(b))} + i\left(\frac{d}{db}\Omega_f(b,\varphi(b))\right)T_f(b,\varphi(b))$$

so that:

$$|\frac{d}{db}T_f(b,\varphi(b))|^2 = |\frac{d}{db}M_f(b,\varphi(b))|^2 + \left(\frac{d}{db}\Omega_f(b,\varphi(b))\right)^2 M_f(b,\varphi(b))^2.$$

For the examples we have in mind we expect that, at least along the ridge, the wavelet transform of the signal can be approximated by the wavelet transform of a function of the form $A(x)\cos\phi(x)$. This essentially means that the signal itself can be modeled as a superposition of such components, and in addition, that the wavelet "separates the components". More precisely, the distance between two consecutive components has to be larger than the wavelet's bandwidth. In this case $\Omega(b,\varphi(b)) \approx K + F(b)$ so that:

$$\frac{d}{db}\Omega(b,\varphi(b)) \approx F'(b) = \frac{\omega_0}{\varphi(b)}.$$

As a consequence, we replace the natural cost function $\tilde{F}_2(f)$ defined in equation (28) by the approximation:

$$F_2(f) = \int_{b_{min}}^{b_{max}} \left(|\frac{d}{db}T_f(b,\varphi(b))|^2 - \frac{\omega_0^2}{\varphi(b)^2}|T_f(b,\varphi(b))|^2\right)db \tag{30}$$

and the global cost function $\tilde{F}$ defined in equation (29) by $F(f) = F_1(f) + F_2(f)$. But the interpretation remains the same: one considers a smoothness penalty in the spirit of the $H^1$-norm of the restriction of the modulus of the wavelet transform to the ridge. Consequently our reconstruction is given by the solution of the minimization of $F(f)$ subject to the linear constraints (26) which we rewrite in the form $L_j(f) = z_j$ where the linear functionals $L_j$ are defined by:

$$L_j(f) = \frac{1}{\varphi_j}\int f(t)\overline{\psi\left(\frac{t-b_j}{\varphi_j}\right)}dt. \tag{31}$$

### 5.3 Solution of the Optimization Problem

A simple computation shows that:

$$F_2(f) = \int\int Q_2(x,y)f(x)f(y)\,dx dy$$

where the kernel $Q_2(x,y)$ is defined by the formula:

$$Q_2(x,y) = \int \frac{db}{\varphi(b)^4}\left(\overline{\psi(\frac{x-b}{\varphi(b)})}\psi(\frac{y-b}{\varphi(b)})[\varphi'(b)^2-\omega_0^2]\right.$$
$$+\overline{\psi'(\frac{x-b}{\varphi(b)})}\psi'(\frac{y-b}{\varphi(b)})[\frac{(x-b)(y-b)}{\varphi(b)^2}+1+\frac{x-2b+y}{\varphi(b)}]$$
$$\left.+\overline{\psi(\frac{x-b}{\varphi(b)})}\psi'(\frac{y-b}{\varphi(b)})\varphi'(b)[1+\frac{y-b}{\varphi(b)}]+\overline{\psi'(\frac{x-b}{\varphi(b)})}\psi(\frac{y-b}{\varphi(b)})\varphi'(b)[1+\frac{x-b}{\varphi(b)}]\right). \quad (32)$$

Using the definition of the total cost function $F(f)$ and the conservation of energy (to compute the first term $F_1(f)$) we see that:

$$F(f) = \int\int Q(x,y)f(x)f(y)\,dx dy$$

is a quadratic functional given by the kernel:

$$Q(x,y) = \delta(x-y) + \epsilon Q_2(x,y). \quad (33)$$

Notice that the kernel $Q(x,y)$ becomes a finite matrix for the purpose of practical applications. Notice further that formulas (33) and (32) give a practical way to compute the entries of the matrix $Q$.
From this point on the procedure introduced in [3] can be used to reconstruct the signal. We merely review the main steps of the algorithm.

The optimization problem can be reformulated as a minimization problem in the real domain rather than the complex domain by noticing the two obvious facts:

- the $n$ complex constraints (26) can be replaced by the $2n$ real constraints:

  $$R_j(f) = r_j, \qquad j = 1,\cdots,2n, \quad (34)$$

  where:

  $$R_j(f) = \frac{1}{\varphi_j}\int f(x)\Re\psi\left(\frac{x-b_j}{\varphi_j}\right)dx \text{ and } R_{n+j}(f) = -\frac{1}{\varphi_j}\int f(x)\Im\psi\left(\frac{x-b_j}{\varphi_j}\right)dx \quad (35)$$

  for $j = 1,\cdots,n$ and $r_j = \Re z_j$ and $r_{n+j} = \Im z_j$.

- The kernel $Q(x,y)$ defining the quadratic functional $F(f)$ can be replaced by its real part $\Re Q(x,y)$. Indeed, the kernel $Q(x,y)$ is sesquilinear and we are only computing $F(f)$ for real signals.

Consequently, there exist real numbers $\lambda_1, \cdots, \lambda_n$, $\lambda_{n+1}, \cdots, \lambda_{2n}$ (the Lagrange multipliers of the problem) such that the solution $\hat{f}(x)$ of the optimization problem is given by the formula:

$$\hat{f}(x) = \sum_{j=1}^{2n} \lambda_j \hat{\psi}_j(x) \tag{36}$$

where the functions $\hat{\psi}_j$ are defined by:

$$\hat{\psi}_j(x) = Q^{-1}\psi_{(b_j,\varphi_j)}, \qquad\qquad j = 1, \cdots, 2n, \tag{37}$$

The Lagrange multipliers are determined by requiring that the constraints (34) be satisfied. In other word, by demanding that the wavelet transform of the function $\hat{f}$ given in (36) be equal to the $z_j$'s at the sample points $(b_j, \varphi_j)$ of the time-scale plane. This gives a system of $(2n) \times (2n)$ linear equations from which the Lagrange multipliers $\lambda_j$'s can be computed.

**Remark:** It is instructive to consider the following particular case: assume that we consider for $b_j$ all the possible values of $b$, and we set $\epsilon = 0$. Then $F = F_1$, $Q$ is a multiple of the identity and it is easy to see that the solution of the constrained optimization problem yields a reconstructed signal of the form:

$$\hat{f}(x) = \int T_f(b, \varphi(b)) \overline{\psi\left(\frac{x-b}{\varphi(b)}\right)} \frac{db}{\varphi(b)}$$

This is a superposition of wavelets located on the ridge. We recover here a reconstruction formula close to those used in references [8] or [14]. However, when the sampling of the ridge becomes coarser, this reconstruction procedure is not satisfactory any longer. This lack of smoothness justifies (once more) the introduction of the second term $F_2$ of the cost function.

### 5.4 The Reconstruction Algorithm

We summarize the results of the discussion of this section in an algorithmic walk through our solution to the reconstruction problem.

- determination of a finite set of sample points $(b_1, \varphi_1), \cdots, (b_n, \varphi_n)$ on the ridge
- Construction of a smooth estimate $b \hookrightarrow \varphi(b)$ of the ridge from the knowledge of the sample points
- computation of the matrix $Q(x,y)$ of the smoothness penalty along the ridge estimate

- computation of the reconstruction wavelets $\hat{\psi}_j = Q^{-1}\psi_{(b_j,\varphi_j)}$ at the ridge sample points of the time-scale plane
- computation of the coefficients $\lambda_j$

The solution $\hat{f}$ of the reconstruction problem is then given by formula (36).

### 5.5 Supplementary Remarks Concerning Noisy Signals

In most practical applications, the input signal is corrupted by noise and the identification and reconstruction problems are significantly more difficult. We already explained how the presence of an additive noise can be handled by the ridge detection algorithms we presented in this paper. But even when sample points of a ridge estimate have been reliably determined, there are still several possible variations on the theme presented in this section when it comes to the reconstruction problem. We mention a few of them for the record.

• In practical applications the exact values the wavelet transform at the sample points of the ridge are not known exactly. The $z_j$'s are merely noisy perturbations of the true values. The reconstruction problem can be solved in a very similar way. See for example [3] and [21] for the modifications which are necessary in order to get the solution of this seemingly more general problem.

• There is an instructive Bayesian interpretation of the penalization problem. The details of this interpretation are given in [21] in the case of the smoothing splines. The argument can easily be adapted to the present situation.

### 5.6 An Example

We selected a subset of $n = 500$ consecutive samples from the bat signal to illustrate the reconstruction procedure described above. This portion of the bat signal can be seen at the top of Figure 3. We used 40 sample points on the estimate of the ridge and the value $\epsilon = .5$ to reconstruct the signal. The result of the reconstruction is given at the bottom of Figure 3. Figure 4 compares the modulus of the wavelet transform of the reconstructed signal to the modulus of the wavelet transform of the original signal. This figure is included not only to show that a very good approximation of the signal can be recovered with a small number of sample points, but also to show that, because we chose a ridge estimate which ignored the existence of a secondary (high frequency) ridge, the latter is not present in the reconstruction. A similar comparison for the arguments instead of the moduli would show that the argument of the wavelet transform is reconstructed *perfectly* in the neighborhood of the ridge. This is a striking feature of the reconstruction which justifies the approximation we made in the definition of the quadratic penalty function which we choose.

## 6 Conclusions

We have described in this paper a new approach to the problem of instantaneous frequency characterization for amplitude and frequency modulated signals embedded in noise. Starting from any energetic distribution of the signal in a time-frequency plane, the problem amounts to finding a curve in the time-scale plane. Our approach is based on the minimization of a penalty function on the set of all possible such candidates. The penalty function takes into account a-priori information on the signal (namely the time-frequency representation of the signal which is essentially localized around the curve, the curve is smooth,...) and possibly on the noise (through an a-priori noise model, or simulations). The minimization is achieved through Monte-Carlo type methods.
We have focused here on the case where the time-frequency representation is given by the square modulus of the wavelet transform (the scale variable being interpreted as an inverse frequency variable). But we would like to stress that any time-frequency representation can be used as well, for instance the square modulus of the Gabor transform, or any (generalized) Wigner distribution (see e.g. [9]). The performances of the algorithms will even improve when the time-frequency representation is adapted to the frequency modulation of the signal.
Let us now point out some of the possible improvements of the method:

- Most of the computational burden of the reconstruction procedure is due to the computation of the reconstructing dual wavelets:

  $$\hat{\psi}_j = Q^{-1}\psi_{(b_j,\varphi_j)}.$$

  In previous implementations of this reconstruction algorithm [3] [4], the reconstructing wavelets did not depend upon the signals to be analyzed. As a consequence, they could be computed once for all. This not the case in the present situation since the penalty function and the matrix $Q$ depend on the signal. In other words, a new matrix $Q$ has to be computed and inverted for each new signal reconstruction. This matrix is band-diagonal, and its bandwidth at position $b$ is proportional to the value of the function $\varphi(b)$, i.e. for instance small bandwidth at the beginning of the signal corresponds to smaller scales. This remark suggests that its inversion should be tackled with specific fast algorithm, for example by rewriting it in an appropriate wavelet basis. This will certainly speed up the reconstruction of signals for which we want to (or need to) use a large number of sample points in the ridge estimate.

- Our approach is not adapted to the case of signals associated with multiple ridges (such as voiced speech signals for instance, see e.g. [14] [15]). A possible solution is to run many times the same algorithm to catch the different ridges. A possible alternative would be to include into the snake algorithm a snake segmentation procedure, such as the one described in [1]. Another possible alternative is described in [5].

**Acknowledgements:** This work was done while the third named author (B.T.) was visiting the Department of Mathematics of the University of California at Irvine. Its warm hospitality should be acknowledged. We would also like to thank P. Flandrin for providing us with the bat sonar signal we used as a main example throughout this paper.

## References

[1] M.-O. Berger (1993): Towards Dynamic Adaptation of Snake Contours. In *International Conference on Image Analysis and Processing, Como (Italy)*. IAPR, September 1991.

[2] P. Bertrand and J. Bertrand (1985): Représentation Temps-Fréquence des Signaux à Large Bande. *La Recherche Aérospatiale (in French)* ***5*** 277–283

[3] R. Carmona (1992): Spline Smoothing & Extrema Representation: Variations on a Reconstruction Algorithm of Mallat and Zhong. to appear in the proceedings of the conference *Wavelets and Statistics*, Villard de Lans, France (1994), A. Antoniadis & G. Oppenheim Eds, Lecture Notes in Statistics.

[4] R. Carmona (1993): Wavelet Identification of Transients in Noisy Signals. Proc. SPIE 15-16 June 1993, San Diego, California. Mathematical Imaging: Wavelet Applications in Signal and Image Processing. 392-400.

[5] R. Carmona, W.L. Hwang and B. Torrésani (1994): Ridge Detection and Speech Reconstruction. (preprint)

[6] R. Carmona, W.L. Hwang and B. Torrésani (1994): Identification of Chirps with Continuous Wavelet Transform. Preprint CPT.93/P.3163, to appear in the proceedings of the conference *Wavelets and Statistics*, Villard de Lans, France (1994), A. Antoniadis & G. Oppenheim Eds, Lecture Notes in Statistics.

[7] I. Daubechies (1992): Ten Lectures on Wavelets. *SIAM.*

[8] N. Delprat, B. Escudié, P. Guillemain, R. Kronland-Martinet, Ph. Tchamitchian, B. Torrésani (1992): Asymptotic wavelet and Gabor analysis: extraction of instantaneous frequencies. *IEEE Trans. Inf. Th.* ***38****, special issue on Wavelet and Multiresolution Analysis* 644-664.

[9] P. Flandrin (1993): Temps-Fréquence. *Traité des Nouvelles Technologies, série Traitement du Signal, Hermès.*

[10] S. Geman and D. Geman (1984): Stochastic Relaxation, Gibbs Distributions and Bayesian Restoration of Images. *IEEE Proc. Pattern Ana. Mach. Intell.* ***6***, 721-741.

[11] B. Gidas (1985): Nonstationary Markov Chains and Convergence of the Annealing Algorithm. *J. Statist. Phys.* ***39***, 73-131.

[12] M. Kass, A. Witkin and D.Terzopoulos (1988): Snakes: Active Contour Models, Int. J. of Computer Vision, 321-331.

[13] P.J.M. van Laarhoven and E.H.L. Aarts (1987): Simulated Annealing: Theory and Applications. Reidel Pub. Co.

[14] R.J. McAulay and T.F. Quatieri (1986): Speech Analysis/Synthesis Based on a Sinusoidal Representation. *IEEE Trans. on Audio, Speech and Sign. Proc.* ***34 #4*** 744–754.

[15] R.J. McAulay and T.F. Quatieri (1992): Low Rate Speech Coding Based on the Sinusoidal Model. In *Advances in Speech Signal Processing*, Edited by S. Furui and M. Mohan Sondui.

[16] S. Mallat and W.L. Hwang (1992): Singularities Detection and Processing with Wavelets. *IEEE Trans. Info. Theory* ***38#2***, 617–643.

[17] S. Mallat and S. Zhong (1992): Characterization of Signals from Multiscale Edges. *IEEE Trans. Pattern Anal. Machine Intel.* ***14***, #7.

[18] T.F. Quatieri and R.J. McAulay (1989): Phase coherence in speech reconstruction for enhancement and coding applications. *Proc. IEEE Int. Conf. Audio, Speech and Sig. Proc.*, Glasgow, 207-209.

[19] Ph. Tchamitchian and B. Torrésani (1992): Ridge and Skeleton Extraction from the Wavelet Transform. In *Wavelets and Applications, M.B. Ruskai & al Eds, Jones & Bartlett Pub. Comp. Boston,* 123–151.

[20] B. Torrésani (1992): Time-Frequency Distributions: Wavelet Packets and Optimal Decompositions. *Annales de l'Institut Henri Poincaré* ***56#2***, 215–234.

[21] G. Wahba (1988): Spline Models for Observational Data. CBMS-NSF Reg. Conf. Ser. in Applied Math. # 59. *SIAM.*

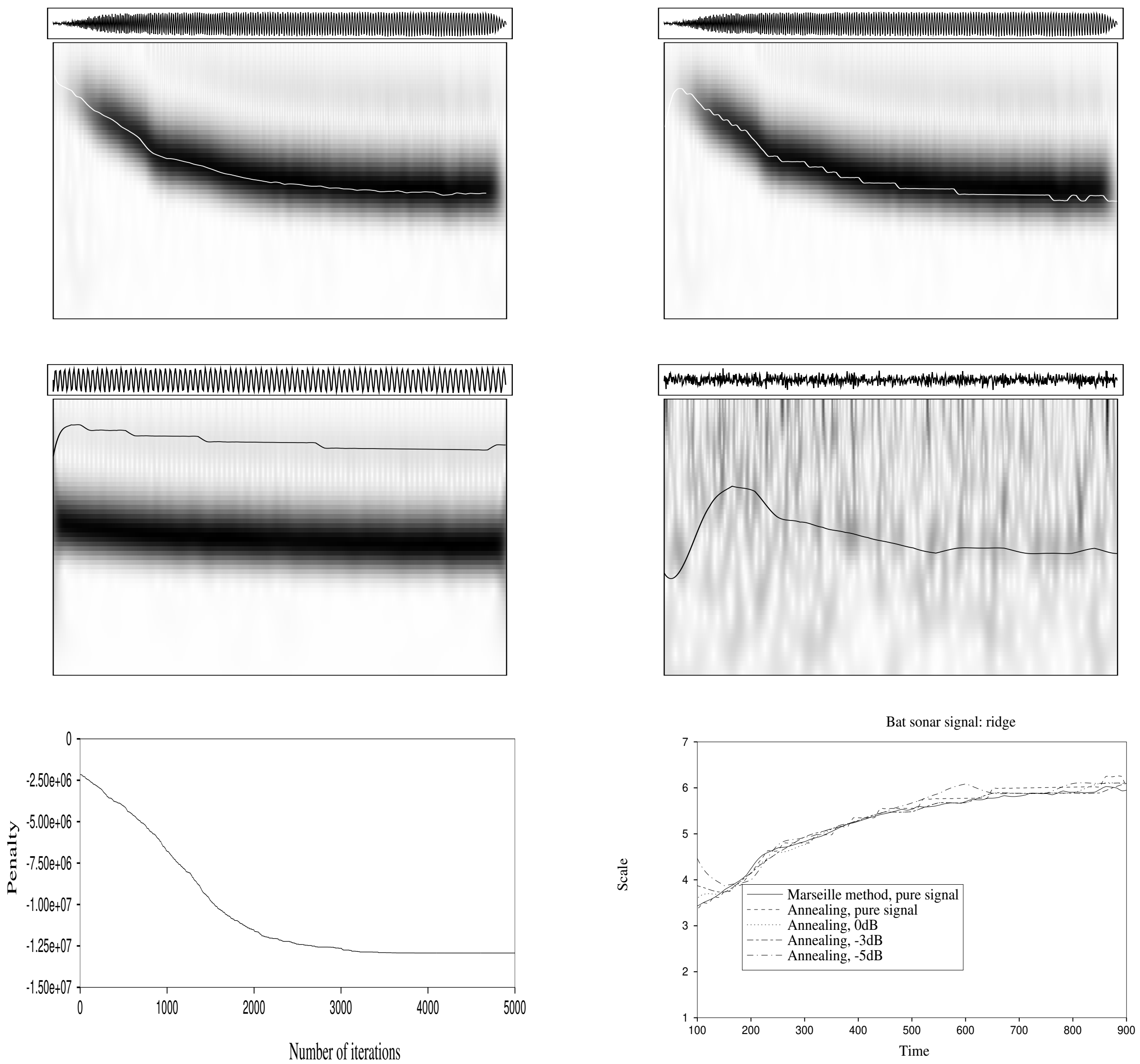


Figure 1: Top Row: Left) Intensity plot of the modulus of the wavelet transform of the bat signal. The exact ridge superimposed. Right) Same plot of the modulus of the wavelet transform of the bat signal with the estimate of the ridge given by the annealing method. Middle Row: Left) Estimate of the first harmonic of the bat sonar signal as given by the annealing method. Right) Ridge estimate given by the annealing method when the signal to noise ratio is $-5dB$. Bottom Row: Left) Graph of the value of the annealing penalty as a function of the number of iterations. Right) Comparison of several ridge estimates for the bat sonar signal. Notice that the scale is now increasing upward as opposed to all the other plots where it is increasing downward.

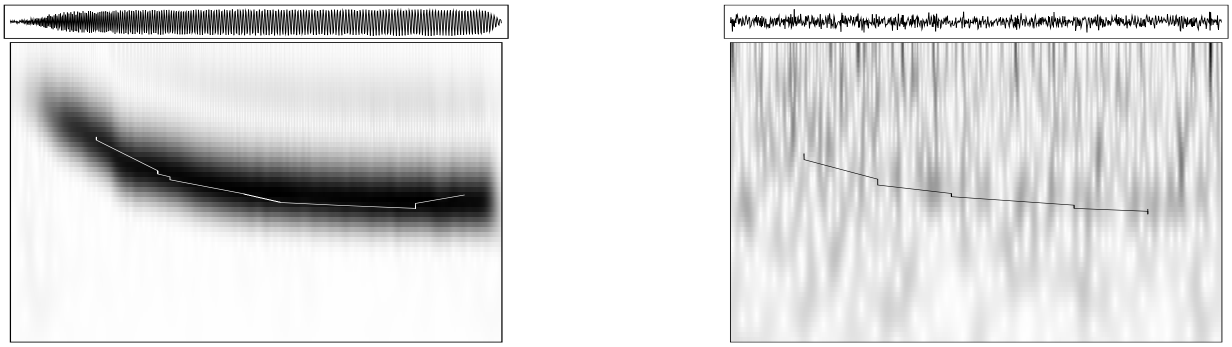

Figure 2: Ridge estimates using the snake algorithm. The bat signal was pure on the left and embedded in a Gaussian white noise (with signal to noise ratio $-5dB$) on the right.

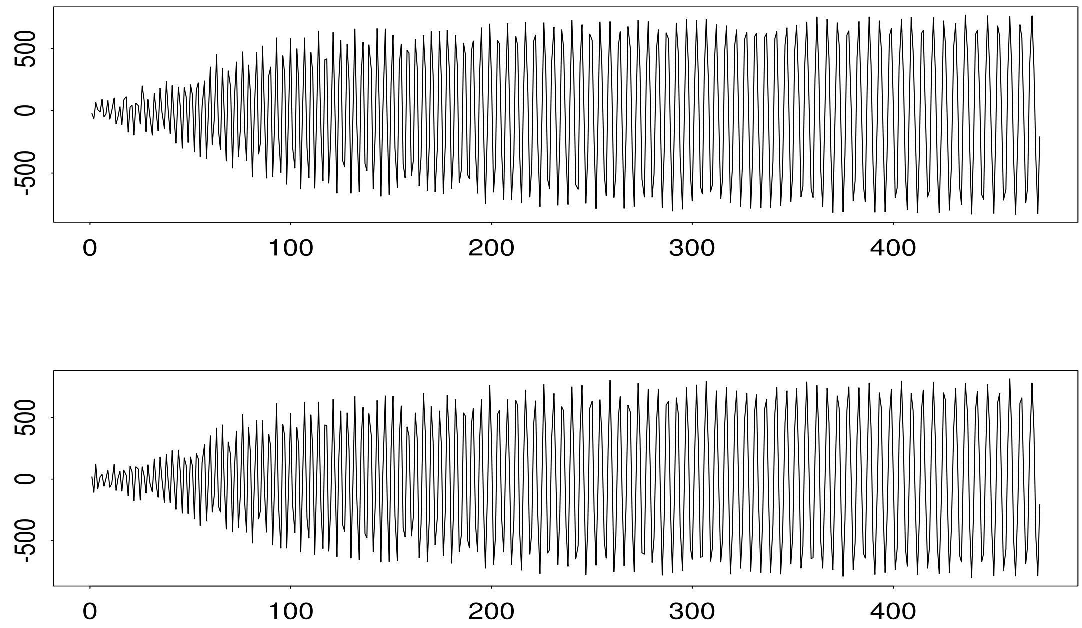


Figure 3: Top) Bat signal used to illustrate the reconstruction procedure. Bottom) Result of the reconstruction from the values of the wavelet transform at the 40 points of the estimate of the ridge.

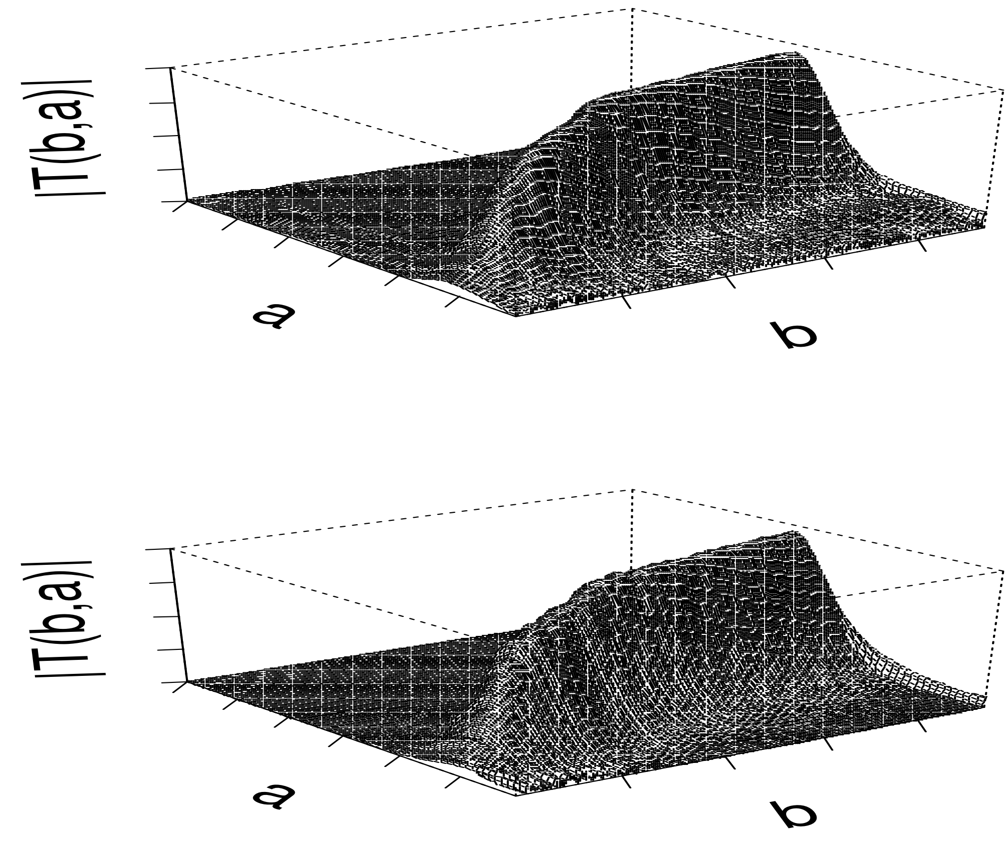


Figure 4: Surface plot of the modulus of the wavelet transform of the 500 samples of the original bat signal (top) and of the reconstructed signal (bottom) shown at the bottom of Figure 3. Notice that the secondary ridge is not reconstructed.