Time-frequency and time-scale analysis of deformed stationary processes, with application to non-stationary sound modeling

Harold Omer, Bruno Torrésani

Time-frequency and time-scale analysis of deformed stationary processes, with application to non-stationary sound modeling

H. Omer, B. Torrésani*

Aix-Marseille Université, CNRS, Centrale Marseille, I2M, UMR 7373, 13453 Marseille, France

Abstract

A class of random non-stationary signals termed *timbre*×*dynamics* is introduced and studied. These signals are obtained by non-linear transformations of stationary random gaussian signals, in such a way that the transformation can be approximated by translations in an appropriate representation domain. In such situations, approximate maximum likelihood estimation techniques can be derived, which yield simultaneous estimation of the transformation and the power spectrum of the underlying stationary signal.

This paper focuses on the case of modulation and time warping of stationary signals, and proposes and studies estimation algorithms (based on time-frequency and time-scale representations respectively) for these quantities of interest.

The proposed approach is validated on numerical simulations on synthetic signals, and examples on real life car engine sounds.

Keywords: Stationary signals, Time-frequency, Time-scale, Warping, Modulation

1. Introduction

Very often, statistical signal processing approaches and algorithms rest on stationarity assumptions placed on the signals of interest and/or on the noise, stationarity being understood as some statistical form of translation invariance. While in many signal processing problems of interest such stationarity assumptions make perfect sense (at least within sufficiently small segments of the signal of interest), this is not always the case. There are situations where departure from stationarity carries essential information on the studied phenomena, and in which one needs to measure non-stationarity accurately. This is an old problem,

*Corresponding author; Phone: +33 4 13 55 14 12; Fax: +33 4 13 55 14 02
Email address: bruno.torresani@univ-amu.fr (B. Torrésani)

that has been considered in some specific situations. Most often, the reference stationary signals are modeled as (deterministic) sine waves or sums of sine waves (as is the case in modulation theory), and the object of interest takes the form of instantaneous frequency ou group delay (see e.g. [1] and references therein). Recent developments in this domain involve approaches such as Empirical Mode Decompositions (see [2] for a review) and various instances of reassignment and synchrosqueezing methods [3], up to the shape function analysis which allows one to step away from the sine wave model [4]. All these approaches assume underlying stationarity in a deterministic setting (sine waves, or shape functions) and do not involve stochastic modeling of stationarity. However, there are several cases of interest where one cannot assume that the underlying signal is a sine wave or a sum of sine waves, because it originates from complex systems that involve fluctuations.

A good example of such situations is provided by the classical *shape from texture* problem of image processing (see for example [5, 6]), where a regular (stationary) texture is projected on a curved surface, yielding images with non-stationary textures. When the underlying texture is periodic, deterministic approaches can be exploited, but this is no longer true when the underlying texture is random. Other examples, which we shall be interested in in this paper, come from audio signal processing, where non-stationary sounds can often be associated to non-stationary motions. Think for example of an accelerating car engine sound. Our ear is perfectly capable to detect regime changes from the sound, more precisely from the non-stationarity of the sound. That example is particularly interesting as in that case, the non-stationarity may be, in first approximation, modeled as a clock change (or time warping), i.e. a periodic system whose inner clock varies as a function of time. This example is the very motivation for the non-stationary sound models we are studying in this paper, which we term *timbre*$\times$*dynamics* models: models in which a stationary random signal (whose power spectrum is interpreted as *timbre*) is modified by some nonlinear function, which encodes the dynamics of some underlying systems.

Such models have already been considered by M. Clerc and S. Mallat in the context of the shape from texture problem (see [6], and [7] for a more complete mathematical and statistical analysis of the approach). One main contribution was the observation that the non-linear transformation involved in the non-stationarity can in some situations approximately translate into a transport in some appropriate representation domain (time-scale, time-frequency). In addition, it was shown that characterizing the transport in question leads to estimates for the deformation.

The current paper builds on this approach, exploiting a slightly different point of view, namely explicit modeling of the underlying stationary signal as a (possibly complex) Gaussian random signal. Under such assumptions, the deformed signals are also complex Gaussian random signals, and the transport property in an appropriate representation domain alluded to above can be made quantitative using simple approximation techniques. This yields an approximate Gaussian (or complex Gaussian) model from which maximum likelihood estimation can be formulated.

More precisely, we consider two deformation models: modulations (which are conveniently studied in the short time Fourier transform domain), and time warping (studied in the wavelet domain). In both cases, we provide approximate expressions for the transform of the corresponding deformed stationary process, together with error estimates. We also provide sufficient conditions for the invertibility of the covariance matrix of the so-obtained complex Gaussian models, and propose estimation algorithms. The theoretical results are complemented by numerical results, both on synthesized signals and on real signal (car engine sounds), that show that the proposed approach is much more accurate and robust than simpler approaches based upon local frequency (or scale) averages of Gabor (or wavelet) transforms. In comparison, the approach is also more accurate and robust than the algorithm of Clerc and Mallat, that doesn't fully exploit the stationarity of the underlying signal.

This paper is organized as follows. After this introduction, we briefly account in Section 2 for the mathematical background this work rests upon, and introduce our notations. Section 3 is devoted to the study of stationary random signals deformed by modulations, presented in the finite-dimensional, discrete case and Section 4 develops the case of time warping. Numerical results are provided and discussed in Section 5, before the conclusion. More technical proofs are given in Section 6.

Part of this paper (namely the frequency modulation estimation) is an elaborated version, of a short paper published in a conference proceedings [8] (where results were announced without proofs). The original contributions of the present paper consist in the proofs of the results of [8] (in the finite-dimensional case), a detailed treatment of the time warping estimation (which is presented in the continuous time setting, and requires slightly more sophisticated mathematical analysis), and extended numerical results, including validation on real signals.

2. Notations and background

We give in this section a short account of the mathematical tools we shall be using in the sequel. We shall treat both discrete time and continuous time signals, and to avoid inconsistencies use the notation $\boldsymbol{x} : t \to x(t)$ for continuous time signals and $\boldsymbol{x} : t \to x[t]$ in the discrete time case (functions and vectors are denoted by a bold symbol while their point values or components are denoted using a normal font).

2.1. Random signal models

We shall mostly be concerned with stationary stochastic signal models, which we emphasize in two different settings. A discrete, finite-dimensional setting, and a continuous-time, infinite-dimensional setting. This choice is motivated by the transforms we use, which turn out to take simpler forms in these settings.

2.1.1. Finite-dimensional models

We start with the simplest case, and consider real or complex random signals $\boldsymbol{X} = \{X[0], \dots X[L-1]\}$ of length L, and their infinite length extension obtained using periodic boundary conditions. We denote by I the integer interval $I = [\![0, L-1]\!]$, and by I^c the centered version obtained by $I^c =]\!] - L/2, L/2]\!]$ for L even and the analogous expression for L odd. Given such a random signal, we denote by $C_{\boldsymbol{X}}$ and $R_{\boldsymbol{X}}$ respectively the corresponding covariance and relation matrices, defined for $m, n \in [\![0, L-1]\!]$ by

$$C_{\boldsymbol{X}}[m,n] = \mathbb{E}\left\{X_m \overline{X}_m\right\}, \quad (1)$$

$$R_{\boldsymbol{X}}[m,n] = \mathbb{E}\left\{X_m X_m\right\}. \quad (2)$$

We shall be concerned with zero-mean (real) Gaussian and complex Gaussian signals. Zero-mean Gaussian signals are characterized by their covariance matrix. Zero-mean complex Gaussian signals are characterized by their covariance and relation matrices. A complex Gaussian centered signal is termed circular if its relation matrix vanishes (see e.g. [9]).

Given a real signal $\boldsymbol{X} \in \mathbb{R}^L$, the corresponding analytic signal $\boldsymbol{Z} \in \mathbb{C}^L$ is obtained by zeroing the negative frequencies of the discrete Fourier (DFT) transform of $\boldsymbol{X}$ and doubling it on positive frequencies. More precisely, $\boldsymbol{Z}$ is defined by its DFT as $\hat{Z}[k] = 0$ for $k \geq L/2$ and $\hat{Z}[k] = 2\hat{X}[k]$ otherwise. If $\boldsymbol{X}$ is a Gaussian, wide sense stationary random signal with power spectrum $\mathcal{S}_{\boldsymbol{X}}$, it is easily shown (see [9] for a proof in the continuous time setting, which can be adapted *mutatis mutandis*) that $\boldsymbol{Z}$ is a circular complex Gaussian wide sense stationary signal, therefore characterized by its covariance matrix $C_{\boldsymbol{Z}}[t,s] = \sum_{k \in I^+} \mathcal{S}_X[k] e^{2i\pi k[t-s]/L}$.

2.1.2. Continuous time models

Because we are also interested in time warping as a signal deformation model, we are naturally led to consider continuous time signal models, which we find convenient to address also in the infinite support case. Since we are interested in trajectories of random signals, the $L^2(\mathbb{R})$ model is not appropriate, and signals are therefore modeled as (random) distributions. Discussing mathematical aspects of such random distributions is far beyond the scope of the present paper, and we simply refer to [10] (in particular chapter 3) for a sketch of the theory.

For the present paper, we limit ourselves to a simple class of random distributions, termed filtered Gaussian white noise in [10]. Notions such as the generalized mean can be defined (in a distributional sense), and covariance and relation matrices are also well defined as (distributional) kernels of linear operators.

It will be sufficient for us to note that in such a context, measurements of the form $\langle \boldsymbol{X}, \varphi \rangle$ are well defined random variables, provided the measurement vectors belong to some appropriate test function space (here, the Schwartz class $\mathcal{S}(\mathbb{R})$). The main aspect of the approach we develop below is based upon transforms such as Gabor or wavelet transforms, both of which can be well defined using such a language.

2.2. Linear time-frequency and time-scale transforms

2.2.1. Gabor transform

We consider the finite-dimensional Gabor transform, as studied in e.g. [11]. Given a window function $\boldsymbol{g} \in \mathbb{C}^L$, the Gabor transform $\mathcal{G}_{\boldsymbol{x}}$ of $\boldsymbol{x} \in \mathbb{C}^L$ with window $\boldsymbol{g}$, hop size a and frequency step b (both a and b must be divisors of L) is defined for $m \in [\![0, M-1]\!]$ and $n \in [\![0 \dots N-1]\!]$ by

$$\mathcal{G}_{\boldsymbol{x}}[m,n] = \langle \boldsymbol{x}, \boldsymbol{g}_{mn} \rangle = \sum_{t=0}^{L-1} x[t] e^{-2i\pi mb[t-na]/L} \overline{g}[t-na] \tag{3}$$

with $M = L/b$ and $N = L/a$ and where $\boldsymbol{g}_{mn}$ denote Gabor atoms of the form $\boldsymbol{g}_{mn}(t) = e^{2i\pi mb[t-na]} g[t-na]$. $\mathcal{G}_{\boldsymbol{X}}$ is therefore an $M \times N$ array. For suitably chosen $\boldsymbol{g}$, and a and b small enough, the Gabor transform is invertible (see [12, 13]); in finite dimensional situations, efficient algorithms have been developed and implemented (see [14]).

Remark 1. Despite the fact that we are working here in a discrete time setting, notice that the Gabor transform can also be defined for any value of the frequency variable. We shall denote by $\mathcal{G}_{\boldsymbol{x}}(.,.)$ (with parentheses instead of brackets) the corresponding semi-continuous Gabor transform, defined as

$$\mathcal{G}_{\boldsymbol{x}}(\nu, n) = \sum_{t=0}^{L-1} x[t] e^{-2i\pi\nu[t-na]/L} \overline{g}[t-na] \ . \tag{4}$$

A simple way to obtain a local frequency estimation using Gabor transform is provided by the *local frequency*, which we define as the function

$$n \to \nu[n] = \frac{\sum_m mb |\mathcal{G}_{\boldsymbol{x}}[m,n]|^2}{\sum_m |\mathcal{G}_{\boldsymbol{x}}[m,n]|^2} \ , \tag{5}$$

i.e. the average value of the frequency variable, using the (suitably renormalized) spectrogram as probability density function. We shall use that quantity as baseline method for the sake of comparison.

2.2.2. Continuous wavelet transform and discretization

Wavelet transform is based upon the idea of dilation, which is intrinsically a continuous-time concept. Given the context of our work, and the fact that we are interested in arbitrary dilation factors, we are naturally led to continuous time models and the corresponding continuous wavelet transform (CWT for short), constructed as follows. Let $\boldsymbol{\psi} \in L^2(\mathbb{R})$. The corresponding continuous wavelet transform of $\boldsymbol{x} \in L^2(\mathbb{R})$ is the function $\mathcal{W}_{\boldsymbol{x}}$ defined as

$$\mathcal{W}_{\boldsymbol{x}}(s,u) = e^{-s/2} \int x(t) \overline{\psi}\left(e^{-s}(t-u)\right) dt = \langle \boldsymbol{x}, \boldsymbol{\psi}_{(s,u)} \rangle \, , \tag{6}$$

where $\boldsymbol{\psi}_{(s,u)}$, defined by $\psi_{(s,u)}(t) = e^{-s/2}\psi(e^{-s}(t-u))$ is a shifted and rescaled copy of the wavelet $\boldsymbol{\psi}$. Here, s and u take values in $\mathbb{R}$. Under suitable assumptions on $\boldsymbol{\psi}$ (the so-called admissibility condition, see e.g. [12]), the continuous

wavelet transform is invertible, a property which we will not exploit here. However, under some mild assumptions, the admissibility condition implies that $\hat{\psi}(0) = 0$, therefore the CWT may be viewed as a (continuous) filter bank involving band pass filters.

The CWT may be discretized in several different ways. We choose here the so-called dyadic wavelet transform [15] (also called stationary, or translation invariant, wavelet transform), which is obtained using a regular sampling of the log-scale variable $s = m \ln(q)$ and a regular, scale-independent sampling $u = na$ of the time variable. Given corresponding such sampling constants q and a, the corresponding wavelets and wavelet transform read

$$\begin{aligned} \psi_{mn}(t) &= q^{-m/2}\psi(q^{-m}(t - na)) \quad (7) \\ \mathcal{W}_{\boldsymbol{x}}[m, n] &= \mathcal{W}_{\boldsymbol{x}}(m \ln(q), na) = \langle \boldsymbol{x}, \boldsymbol{\psi}_{mn} \rangle \ . \quad (8) \end{aligned}$$

The CWT is readily extended to the analysis of distributions, provided the wavelet is chosen in a suitable test function spaces. Hereafter, we shall assume that $\boldsymbol{\psi} \in \mathcal{S}(\mathbb{R})$, which will allow us to analyze tempered distributions.

As in the Gabor case, a simple wavelet based time-dependent scale estimation can be obtained by the *local scale*, which we define as the function

$$n \to \sigma[n] = \frac{\sum_m q^m |\mathcal{W}_{\boldsymbol{x}}[m, n]|^2}{\sum_m |\mathcal{W}_{\boldsymbol{x}}[m, n]|^2} \ . \quad (9)$$

Again, we shall use that quantity for the sake of comparison.

3. Deformation by modulation and estimation

We first consider modulations, that may suitably be approximately described as frequency shifts in a joint time-frequency domain (see Theorem 1 below). Results of these section have been given (without proof) in [8]. We formulate the problem as follows.

3.1. Model and estimates

We limit ourselves to finite-dimensional situations, and use the notations of Section 2.1.1 (the infinite-dimensional case is developed in [16]). Let $\boldsymbol{X} \in \mathbb{R}^L$ be a zero-mean, wide sense stationary Gaussian random process, with covariance matrix $C_{\boldsymbol{X}}$, and let $\boldsymbol{Z} \in \mathbb{C}^L$ denote the associated analytic signal. Let $\mathcal{S}_{\boldsymbol{X}}$ be the power spectrum of $\boldsymbol{X}$, and assume that $\mathcal{S}_{\boldsymbol{X}}[0] = 0$, and if L is even, that $\mathcal{S}_{\boldsymbol{X}}[L/2] = 0$. Following the lines of [9], it is easily shown that under such an assumption, $\boldsymbol{Z}$ is a circular complex Gaussian random vector, therefore completely characterized by its covariance matrix. The latter reads

$$C_{\boldsymbol{Z}}[t, s] = \sum_{\nu \in I^+} \mathcal{S}_{\boldsymbol{X}}[\nu] e^{2i\pi\nu[t-s]/L} \ , \quad (10)$$

and $\boldsymbol{Z}$ is therefore wide-sense stationary.

For simplicity, we work with complex-valued signals. The observation is assumed to be an USB (upper sideband) modulated version $\boldsymbol{Y}$ of the reference stationary signal $\boldsymbol{X}$, namely

$$Y[t] = Z[t]e^{2i\pi\gamma(t)/L} + N[t] \ , \tag{11}$$

where $\gamma \in C^2([0, L])$ is an unknown smooth, slowly varying *modulation function*, and $\boldsymbol{N} = \{N[t], t \in I_L\}$ is a circular complex Gaussian white noise, with variance σ_0^2.

Remark 2. When γ is an affine function, $\boldsymbol{Y}$ is still a stationary random process. As the estimation procedure we shall be studying below is mainly based on a stationarity assumption placed in $\boldsymbol{Z}$ (in addition to a smoothness assumption on γ), this implies that the modulation γ can only be determined up to an affine function.

When γ is not an affine function, $\boldsymbol{Y}$ is not wide sense stationary any more. The problem at hand is to estimate the unknown modulation γ and the original power spectrum $\mathcal{S}_{\boldsymbol{X}}$ from a single realization of $\boldsymbol{Y}$.

We will base the frequency modulation estimation on a Gabor representation of the observed signal, and deliberately disregard correlations across time of the Gabor transform (hence focusing on time slices of the Gabor transform of the observation). The distribution of time slices of the Gabor transform $\mathcal{G}_{\boldsymbol{Z}}$ of the analytic signal $\boldsymbol{Z}$ associated with the original signal $\boldsymbol{X}$ is characterized in the following result, which result from direct calculations (see for example [12] for similar calculations)

Proposition 1. 1. *For fixed n, the Gabor transform $\mathcal{G}_{\boldsymbol{N}}[., n]$ of the Gaussian white noise is a stationary Gaussian random vector, with circular covariance matrix*

$$C_{\mathcal{G}_{\boldsymbol{N}}}[m, m'] = \sigma_0^2 \sum_{k=0}^{L-1} \overline{\hat{g}}[k]\hat{g}[k - (m' - m)b] \tag{12}$$

2. *For fixed time index n, the Gabor transform $\mathcal{G}_{\boldsymbol{Z}}[., n]$ of the analytic signal is a circular complex Gaussian random vector, with covariance matrix*

$$C_{\mathcal{G}_{\boldsymbol{Z}}}[m, m'] = \sum_{k \in I^+} \mathcal{S}_{\boldsymbol{X}}[k]\overline{\hat{g}}[k - mb]\hat{g}[k - m'b] \tag{13}$$

The estimation of the modulation will be based upon an approximation of the covariance matrix of the observed signal. In a few words, the Gabor transform of the frequency modulated signal can be approximated by a deformed version of the Gabor transform of the original signal. The deformation takes the form of a time-varying frequency shift, as follows from a first order Taylor expansion of the modulation function. Since the Gabor atom $\boldsymbol{g}_{mn}$ is localized around $t = na$, we just write $\gamma(t) \approx \gamma(na) + (t - na)\gamma'(na)$, and thus obtain

$$\mathcal{G}_{\boldsymbol{Z}}[m, n] \approx e^{2i\pi\gamma(na)/L}\mathcal{G}_{\boldsymbol{Z}}(mb + \gamma'(na), na) \ .$$

Notice that since $\gamma'(na)$ has no reason to be an integer multiple of the frequency step b, we are led to a version of the Gabor transform defined in the continuous-frequency case, thus the use of parentheses instead of brackets for the arguments of $\mathcal{G}_{\boldsymbol{Y}}$ (see Remark 1).

A more precise argument, exploiting the smoothness of the modulation function γ, leads to the following result. Given a time index n and some number δ, denote by $\boldsymbol{G}^{(\delta;n)}$ the corresponding frequency-shifted copy of Gabor transform $\mathcal{G}_{\boldsymbol{Y}}$:

$$\boldsymbol{G}^{(\delta;n)}[m] \;=\; e^{2i\pi\gamma(na)/L}\sum_{t=0}^{L-1} Z_t\overline{g}[t-na]e^{-2i\pi[m-\delta][t-an]/M} \\ +\; \mathcal{G}_{\boldsymbol{N}}[m,n]\;. \tag{14}$$

Let us set $\delta=\gamma'(na)/b$, and introduce the following constants

$$\mu_1=\sum_{t\in I_1}|g[t]|\;,\;\mu_2=\sum_{t\in I_2}t^2|g[t]|\;,\;T=\sqrt{\frac{L}{\pi\|\gamma''\|_\infty}} \tag{15}$$

where $I_2=[\![-T,T]\!]$ and $I_1=I^c\backslash I_2$. Then we have

Theorem 1. *Assume that $L>4/\pi\|\gamma''\|_\infty$.*

1. *For fixed time n, the Gabor transform $m\to\mathcal{G}_{\boldsymbol{Y}}[m,n]$ may be approximated as*
$$\mathcal{G}_{\boldsymbol{Y}}[m,n]=\boldsymbol{G}^{(\delta;n)}[m]+R[m]\;, \tag{16}$$
and the remainder is bounded as follows: for all m,m',
$$\left|\mathbb{E}\left\{R[m]\overline{R}[m']\right\}\right|\leq\sigma_Z^2\left(2\mu_1+\frac{\pi e}{L}\|\gamma''\|_\infty\mu_2\right)^2, \tag{17}$$
where σ_Z^2 is the variance of Z.
2. *Given δ, and for fixed n, $\boldsymbol{G}^{(\delta;n)}$ is distributed following a circular multivariate complex Gaussian law, with covariance matrix*
$$C_{\boldsymbol{G}^{(\delta;n)}}[m,m']=C_{\mathcal{G}_{\boldsymbol{Z}}}[m-\delta,m'-\delta]+\;C_{\mathcal{G}_{\boldsymbol{N}}}[m,m']\;. \tag{18}$$

Proof: The proof is given in Section 6. □

The estimation procedure described below is a maximum likelihood approach, which requires inverting the covariance matrix of vectors $\boldsymbol{G}^{(\delta;n)}$. The latter is positive semi-definite by construction, but not necessarily definite. The result below provides a sufficient condition on $\boldsymbol{g}$ and the noise for invertibility.

Proposition 2. *Assume that the window $\boldsymbol{g}$ is such that*

$$K_{\boldsymbol{g}}:=\frac{1}{b}\min_{t\in[\![0,M-1]\!]}\left(\sum_{k=0}^{b-1}|g[t+kM]|^2\right)>0\;. \tag{19}$$

Then for all $\boldsymbol{x} \in \mathbb{C}^M$,

$$\langle C_{\boldsymbol{G}} \boldsymbol{x}, \boldsymbol{x} \rangle \geq \sigma_0^2 K_{\boldsymbol{g}} \left\| \boldsymbol{x} \right\|^2 , \tag{20}$$

and the covariance matrix is therefore boundedly invertible.

Proof: The proof is given in the Section 6. □

Remark 3. The condition may seem at first sight unnatural in terms of Gabor frame theory. However, it simply expresses that the number M of frequency bins should not be too large if one wants the covariance matrix to be invertible. Nevertheless, reducing M also reduces the precision of the estimate, and a trade-off has to be found, as discussed in the next section.

3.2. Estimation algorithm

We now describe in some details the estimation procedure corresponding to our problem. The estimation problem is the following: from a single realization of the signal model (11), estimate the modulation function γ and the original covariance matrix $C_{\boldsymbol{G}}$. We first notice the indeterminacy in the problem, namely the fact that adding an affine function to γ is equivalent to shifting $\mathcal{S}_{\boldsymbol{X}}$. This has to be fixed by adding an extra constraint in the estimation procedure.

The procedure is a two-step iterative approach: alternatively estimate γ given $C_{\boldsymbol{G}}$, and estimate $C_{\boldsymbol{G}}$ given γ.

3.2.1. Maximum likelihood modulation estimation

With the same notations as before, let n be a fixed value of the time index, and let $\boldsymbol{G} = \boldsymbol{G}^{(n)}$ be the corresponding fixed time slice of $\mathcal{G}_{\boldsymbol{Y}}$. As the signal and therefore the fixed time Gabor transform slices are distributed according to multivariate complex Gaussian laws, the log-likelihood of a slice takes the form

$$\mathcal{L}_\delta(\boldsymbol{G}) = -\left\langle C^{-1}_{\boldsymbol{G}^{(n;\delta)}} \boldsymbol{G}, \boldsymbol{G} \right\rangle - \ln\left(\pi^M \det(C_{\boldsymbol{G}^{(n;\delta)}})\right) . \tag{21}$$

Therefore, maximum likelihood estimation leads to the frequency shift estimate

$$\hat{\delta} = \underset{\delta}{\operatorname{argmin}} \left[\left\langle C^{-1}_{\boldsymbol{G}^{(\delta;n)}} \boldsymbol{G}, \boldsymbol{G} \right\rangle + \ln\left(\pi^M \det(C_{\boldsymbol{G}^{(\delta;n)}})\right) \right] . \tag{22}$$

Notice that $\det(C_{\boldsymbol{G}^{(\delta;n)}})$ actually does not depend on the modulation parameter δ. Hence the estimate reduces to

$$\hat{\delta} = \underset{\delta}{\operatorname{argmin}} \left\langle C^{-1}_{\boldsymbol{G}^{(\delta;n)}} \boldsymbol{G}, \boldsymbol{G} \right\rangle , \tag{23}$$

a problem to be solved numerically. For the numerical solution, we resort to a simple exhaustive search, since the search domain is small. As $\delta(n) \approx \gamma'(an)/b$, the estimates of δ for each n lead to an estimate of γ', which allows us to estimate γ using standard integration and interpolation techniques. Notice that the estimation of γ' is pointwise, no additional use of the smoothness assumption on γ was necessary in our experiments.

Notice that this requires the knowledge of the covariance matrix $C^{(0;n)}_{\boldsymbol{G}}$ corresponding to the Gabor transform of the noisy stationary signal. The latter is generally not available, and has to be estimated as well.

Remark 4. The minimisation problem (23) is solved by an exhaustive search on the δ, the estimate of γ' is thus quantized. In addition, according to Remark 3, the frequency step parameter b has to be large for the invertibility of the covariance matrix, and the quantization is therefore coarse. As described in [8], this problem is solved by computing a Gabor transform with fine frequency sampling, followed by a search on a sequence of shifted coarsely frequency subsampled transforms. We refer to [8] for more details, and to 8 below where a similar procedure is described in the case of time warping estimation.

As a result, the quantization effect on the modulation function estimate is significantly attenuated, as will be clear in the numerical simulations presented below in Section 5.

3.2.2. Covariance matrix estimation

We now describe a method for estimating the covariance matrix $C_{\boldsymbol{G}}^{(n;0)}$. Suppose that an estimate $\hat{\gamma}$ of the modulation function γ is available. Then the signal $\boldsymbol{Y}$ can be demodulated by setting

$$\boldsymbol{U} = \boldsymbol{Y} e^{-2i\pi\hat{\gamma}/L} \ , \tag{24}$$

Clearly, $\boldsymbol{U}$ is an estimator of $\boldsymbol{Z} + \boldsymbol{N}e^{-2i\pi\gamma/L}$, the noisy stationary signal. We can now compute the covariance matrix $C_{\mathcal{G}_{\boldsymbol{U}}}$ of the Gabor transform of $\boldsymbol{U}$, which is an estimator of $C_{\mathcal{G}_{\boldsymbol{Z}}} + C_{\mathcal{G}_{\boldsymbol{N}}}$. Comparing with equation (14) we finally obtain an estimator for the covariance matrix

$$C_{\mathcal{G}_{\boldsymbol{U}}} \approx C_{\boldsymbol{G}^{(0;n)}} \tag{25}$$

3.2.3. Power spectrum estimation

Suppose that an estimate $\hat{\gamma}$ of the modulation function γ is available. Then the signal $\boldsymbol{Y}$ can be demodulated as in (24), and an estimate for the power spectrum deduced from it. The power spectrum can be estimated using a standard Welch periodogram estimator, or from a Gabor transform as described in [12].

3.2.4. Summary of the estimation procedure

We now summarize an iterative algorithm to jointly estimate the covariance matrix $C_{\boldsymbol{G}}^{(0;n)}$ and the modulation function γ, that exploits alternatively the two procedures described above. The procedure is as follows, given a first estimation of the modulation function, we can perform a first estimation of the covariance matrix, which in turn allows us obtain a new estimation of the modulation function. The operation is repeated until the stopping criterion is satisfied.

For the initialization, we need a first modulation frequency estimate, for which we use the local frequency function defined in (5).

The stopping criterion is based upon the evolution of the frequency modulation along the iterations. More precisely, we use the empirical criterion

$$\frac{||\hat{\delta}^{(k)} - \hat{\delta}^{(k+1)}||_2}{||\hat{\delta}^{(k+1)}||_2} < \epsilon, \tag{26}$$

where $\delta^{(k)}$ is the estimation of δ at iteration number k. The pseudo-code of the algorithm has been given in an earlier paper [8]. We do not reproduce here as it is a mere modification of the time warping estimation algorithm to be described below.

4. Time warping and estimation

We now turn to the second class of deformations that is of interest to us, namely time warping. Again the assumption is that the observation is the image of a stationary random signal by some transformation, here a time warping, that breaks stationarity. To allow for arbitrary time changes, we have to consider here a continuous time model. We consider smooth, monotonically increasing *warping functions* $\gamma \in C^2$, and associate with them the corresponding time warping operator D_γ, defined by

$$[D_\gamma x](t) = \sqrt{\gamma'(t)}\, x(\gamma(t)) \ . \tag{27}$$

D_γ is a unitary operator on $L^2(\mathbb{R})$ (i.e. $\|D_\gamma \boldsymbol{x}\| = \|\boldsymbol{x}\|$ for all $\boldsymbol{x} \in L^2(\mathbb{R})$)

Remark 5. Notice that additional assumptions are needed for D_γ to be a well defined operator $\mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$. A sufficient condition is that $\gamma \in \mathcal{C}^\infty(\mathbb{R})$ be such that all its derivatives have at most polynomial growth, and satisfy the following condition

$$|\gamma(t)| > \alpha |t|^\beta \ , \quad \forall t, |t| > t_0$$

for some $\alpha \in \mathbb{R}_+^*$, $\beta \in \mathbb{R}_+^*$ and $t_0 \in \mathbb{R}^+$.

In addition, elementary calculus shows that warping operators satisfy the composition property $D_\gamma D_\varphi = D_{\varphi \circ \gamma}$ for all $\varphi, \gamma \in C^2$, where the symbol $\circ$ denotes composition of functions. Similarly, $(D_\gamma)^{-1} = D_{\gamma^{-1}}$, γ^{-1} being the reciprocal function of γ.

A warping operator D_γ is said to be *controlled* if there exist two strictly positive constants c_γ and C_γ such that for all t,

$$0 < c_\gamma \leq \gamma'(t) \leq C_\gamma < \infty. \tag{28}$$

We shall need the following result, which follows from the unitarity of the warping operator.

Lemma 1. *Let $\boldsymbol{N} = \{N(t), t \in \mathbb{R}\}$ be a (complex) Gaussian white noise. Then for all monotonically increasing $\gamma \in C^2$, $D_\gamma \boldsymbol{N}$ is also a (complex) Gaussian white noise with the same variance.*

4.1. Model and estimates

We therefore consider a generative model of the form

$$Y(t) = [D_\gamma X](t) + N(t) \ , \tag{29}$$

where $\boldsymbol{X}$ is a second order stationary random process, $\boldsymbol{N}$ is a white noise with variance σ_0^2, and γ is a smooth, monotonically increasing function, such that the corresponding warping operator is controlled, as defined in (28). We denote again by $\mathcal{S}_{\boldsymbol{X}}$ the power spectrum of $\boldsymbol{X}$.

Let $\boldsymbol{\psi} \in \mathcal{S}(\mathbb{R})$ be a wavelet function; we are interested in the behavior of the discrete wavelet transform

$$\mathcal{W}_{\boldsymbol{Y}}[m,n] = \langle \boldsymbol{Y}, \boldsymbol{\psi}_{mn} \rangle = \langle D_\gamma \boldsymbol{X}, \boldsymbol{\psi}_{mn} \rangle + \langle \boldsymbol{N}, \boldsymbol{\psi}_{mn} \rangle \ , \tag{30}$$

within a given region of the time-scale plane $m, n \in \Lambda$. As remarked in [7], time warping can be approximated by translation in the wavelet domain. In the framework of the Gaussian model, we shall see that scale translations are enough to provide estimators for the time warping.

Remark 6. Clearly, when the warping function γ is an affine function, the warped signal $\boldsymbol{Y}$ is still second order stationary. Therefore, since our approach is based upon non-stationarity, the function γ can be determined only up to an affine function.

The behavior of the wavelet transform of stationary random signals is given by

Proposition 3. *Let $\boldsymbol{\psi} \in \mathcal{S}(\mathbb{R})$ be an analytic wavelet function. Let $\boldsymbol{X}$ denote a wide sense stationary circular generalized Gaussian process with power spectrum $\mathcal{S}_{\boldsymbol{X}}$, and let $\mathcal{W}_{\boldsymbol{X}}$ be its wavelet transform. For any value of the time index n, the corresponding wavelet time slice $\mathcal{W}_{\boldsymbol{X}}[,n]$ is a circular Gaussian random vector, with covariance*

$$C_{\mathcal{W}_{\boldsymbol{X}}}[m,m'] = q^{(m+m')/2} \langle \mathcal{S}_{\boldsymbol{X}}, \hat{\psi}^m \overline{\hat{\psi}^{m'}} \rangle,$$

where $\hat{\psi}^m(\nu) = \hat{\psi}(q^m \nu)$. The latter can be formally written

$$C_{\mathcal{W}_{\boldsymbol{X}}}[m,m'] = q^{(m+m')/2} \int_0^\infty \mathcal{S}_{\boldsymbol{X}}(\nu) \overline{\hat{\psi}}(q^m \nu) \hat{\psi}(q^{m'} \nu) d\nu \ .$$

In particular, if $\boldsymbol{X}$ is a white noise with variance σ_0^2,

$$C_{\mathcal{W}_{\boldsymbol{X}}}[m,m'] = \sigma_0^2 q^{(m+m')/2} \int_0^\infty \overline{\hat{\psi}}(q^m \nu) \hat{\psi}(q^{m'} \nu) d\nu \ .$$

We now turn to the analysis of time warped signals. As in the case of modulation, we shall see that if γ is smooth, the effect of warping on wavelet transform can be approximated by in the time-scale plane, as shown in [7]. Assuming that the wavelet is localized around the origin $t = 0$, and using a second order Taylor approximation for the warping function in the neighborhood of ψ_{mn}, namely $\gamma(t) \approx \gamma(na) + (t - na)\gamma'(na) + r(t)$ for some remainder r, we obtain the approximation

$$\mathcal{W}_{D_\gamma \boldsymbol{X}}[m,n] \approx \mathcal{W}_{\boldsymbol{X}}(m + \log_q(\gamma'(na), \gamma(na)) \ .$$

More precisely, let us set

$$\boldsymbol{W}^{(\delta;\tau)}[m] = \mathcal{W}_{\boldsymbol{X}}(m+\delta,\tau) \ + \mathcal{W}_{\boldsymbol{N}}(m,\tau) \ , \tag{31}$$

which involves samples of a deformed version of the continuous wavelet transform of $\boldsymbol{X}$, and set $\delta_n = \log_q(\gamma'(na))$ and $\tau_n = \gamma'(na)$. Notice that these samples need not (and will not in general) belong to the sampling grid of the discrete wavelet transform. The following result, whose elementary proof is left to the reader, will be useful in the sequel.

Lemma 2. *For fixed n, $m \to \boldsymbol{W}^{(\delta_n;\tau_n)}[m]$ is a circular Gaussian random vector with covariance matrix*

$$C^{(n)}_{\boldsymbol{W}}[m,m'] = C_{\mathcal{W}_{\boldsymbol{X}}}(m+\delta_n, m'+\delta_n) + C_{\mathcal{W}_{\boldsymbol{N}}}[m,m'] \ .$$

The approximation error can be controlled thanks to the following estimate

Theorem 2. *Let $\boldsymbol{\psi}$ be a wavelet function satisfying $|\psi(t)| \leq (1+|t|^\alpha)^{-1}$ with $\alpha > 2$. Let $\boldsymbol{X}$ be a stationary circular Gaussian generalized process, whose power spectrum $\mathcal{S}_{\boldsymbol{X}} \in L^1(\mathbb{R})$ is such that*

$$\rho_{\boldsymbol{X}}(\alpha) := \int_{-\infty}^{\infty} |\nu|^{2-\frac{6}{\alpha+2}} \mathcal{S}_{\boldsymbol{X}}(\nu) d\nu < \infty \ .$$

Let $\gamma \in C^2$ be a strictly increasing smooth function, satisfying (28), for some constants c_γ, C_γ, and let $\boldsymbol{Y}$ denote a noisy time-warped random signal, as defined in (29). Then with the above notations, for each fixed time index n the wavelet transform $\mathcal{W}_{\boldsymbol{Y}}$ of the time-warped signal $\boldsymbol{Y}$ may be approximated as

$$\mathcal{W}_{\boldsymbol{Y}}[m,n] = \boldsymbol{W}^{(\delta_n;\tau_n)}[m] + \mathcal{W}_{\boldsymbol{N}}[m,n] + \epsilon_{mn} \ ,$$

where the remainder ϵ_{mn} is circular Gaussian random vector, whose size can be bounded as

$$\mathbb{E}\left\{|\epsilon_{mn}|^2\right\} \leq \|\gamma''\|_\infty^2 q^{3m} \left(K_0 \sigma_{\boldsymbol{X}} + K_1 \|\gamma''\|_\infty^{\frac{-3}{\alpha+2}} q^m q^{\frac{-6m}{\alpha+2}} \sqrt{\rho_{\boldsymbol{X}}(\alpha)} \right) ,$$

where K_0 and K_1 are two constants that depend on α.

Remark 7. Similar estimates can be obtained starting from different decay assumptions on the analyzing wavelet ψ. We refer to [16] for detailed derivations.

Proof: The proof is mainly an adaptation of the proof of Theorem 1, that also takes into account decay assumptions on the wavelet, and is given in Section 6.4. □

We notice that for α large enough, the leading term of the error behaves as q^{3m}: the smoother the warping the better the estimate; also the error goes to

0 as the scale decreases. However, let us stress that the estimation algorithm is not based upon such an asymptotic analysis, but on an explicit modeling.

The estimate also seems to indicate (although the bound is not so easy to interpret) that large values of α should be preferred (i.e. well time localized wavelets). This is in accordance with the initial assumptions (i.e. $\psi \in \mathcal{S}(\mathbb{R})$), and seems to be confirmed by the numerical simulations presented below.

As before, the estimation of the warping function from a realization of the deformed signal will require inverting the covariance matrix of fixed n wavelet transform. The result below provides a sufficient condition for invertibility, and makes use of the so-called Fourier-Mellin transform. The latter associates with any analytic signal $\boldsymbol{z}$ (i.e. any $L^2(\mathbb{R})$ function whose Fourier transform vanishes on the negative half line) the function $\underline{\boldsymbol{z}}$ of the real variable s as

$$\underline{z}(s) = \int_0^\infty \hat{z}(\nu)\nu^{-2i\pi\nu s}\frac{d\nu}{\sqrt{\nu}} \ .$$

The Fourier-Mellin is a unitary transform (i.e. $\|\underline{\boldsymbol{z}}\| = \|\boldsymbol{z}\|$), and we have the following

Proposition 4. *With the same notations as before, assume that the analyzing wavelet $\boldsymbol{\psi}$ is such that its Fourier-Mellin transform $\underline{\boldsymbol{\psi}}$ satisfies*

$$K_{\boldsymbol{\psi}} := \frac{1}{\log(q)} \inf_{s\in[0,1/\ln(q)]} \sum_{\ell\in\mathbb{Z}} \left|\underline{\psi}\left(s+\frac{\ell}{\log(q)}\right)\right|^2 > 0$$

Then for all $\boldsymbol{x} \in \ell^2(\mathbb{Z})$,

$$\langle C_{\boldsymbol{W}}\boldsymbol{x}, \boldsymbol{x}\rangle \geq K_{\boldsymbol{\psi}}\sigma_0^2 \ \|\boldsymbol{x}\|^2 \ ,$$

and the covariance matrix $C_{\boldsymbol{X}}$ is therefore boundedly invertible.

Proof: the proof follows the line of the proof of Proposition 2, (moving from discrete to continuous, and replacing DFT by Mellin's transformation) and is detailed in Section 6.5. □

Notice that the condition given in Proposition 4 places restrictions on the wavelet and the scale variable discretization. When q is too small, the periodized Fourier-Mellin transform of $\boldsymbol{\psi}$ with period $1/\ln(q)$ can approach 0, and the covariance matrix may not be invertible. This will lead us to stick to large enough values of the scale sampling parameter q. Consequences are discussed in Remark 8 below.

4.2. Estimation algorithm

The algorithm is essentially the same as the algorithm described in Section 3.2: it disregards time correlations, and searches for a logarithmic scale translation for each fixed-time slice of a wavelet transform.

4.2.1. Maximum likelihood modulation estimation

Let n be a fixed value of the time index, and let $\boldsymbol{W} = \mathcal{W}[\cdot, n]$ be the corresponding fixed-time slice of $\mathcal{W}_{\boldsymbol{Y}}$.

With the same assumptions as in Section 3.2 the maximum likelihood estimation for the scale shift takes the form

$$\hat{\delta} = \underset{\delta}{\operatorname{argmin}} \left\langle C^{-1}_{\boldsymbol{W}^{(\delta;\tau_n)}} \boldsymbol{W}, \boldsymbol{W} \right\rangle \ , \tag{32}$$

where $\delta \approx \log_q(\gamma'(na))$. Notice that, as in the modulation problem, the knowledge of the covariance matrix $C^{(0;\tau_n)}_{\boldsymbol{W}}$ is required. The latter matrix has to be estimated.

4.2.2. Covariance matrix estimation

The method for estimating the covariance matrix $C^{(n;0)}_{\boldsymbol{W}}$ is very close to what we described for the modulation estimation, and mainly differs in the way we invert the time warping on $\boldsymbol{Y}$. Suppose that an estimate $\hat{\gamma}$ of the warping function γ is available, $\boldsymbol{Y}$ can be time-unwarped by setting

$$\mathcal{W}_{D_{\gamma^{-1}}\boldsymbol{Y}}[m, n] = \langle \boldsymbol{Y}, D_\gamma \boldsymbol{\psi}_{mn} \rangle \ . \tag{33}$$

From this we can estimate the needed covariance matrix using a standard sample estimator.

4.2.3. Summary of the estimation procedure

We summarize an iterative algorithm to jointly estimate the covariance matrix $C^{(0;\tau_n)}_{\boldsymbol{W}}$ and the warping function γ. Algorithm is substantially the same as in Section 3.2: given a first estimation of the warping function, the covariance matrix is estimated using sample estimate from the unwarped signal as in (33); this in turn allows us obtain a new estimation of the warping function. The operation is repeated until the stopping criterion (26) is satisfied. For the initialization, we need a first warping function estimate, for which we use the local scale function defined in (9).

The pseudo-code of the algorithm can be found in Algorithm 1.

Remark 8. As a consequence of Proposition 4, the scale sampling step must be large enough to ensure the invertibility of the wavelet covariance matrix $\boldsymbol{W}^{(0;\tau_n)}$. The optimization being performed using exhaustive search, this results in a quantization effect that can be quite significant. To reduce such an effect, we use the following procedure. The wavelet transform is computed on a fine scale sampling grid (i.e. with a small value of the scale sampling step q), but the optimization (32) is performed independently on sub-grids of the form $Q_i = \{q_k = q^{i+kj}, k = k_{min} \dots k_{max}\}$, where i runs from 0 to $j-1$ and j is large enough to insure invertibility for the covariance matrix. Then the δ estimate from the best subgrid (i.e. yielding the maximal likelihood value) is retained. This procedure can be seen as a regularization that ignores correlations between close scales.

Algorithm 1: Joint covariance and warping function estimation

Initialize as in (9)
while criterion (26) is false **do**
- Compute $\hat{\gamma}^{(k)}$ by interpolation and integration from $\delta^{(k)}$.
- Unwarp $\boldsymbol{Y}$ using $\hat{\gamma}^{(k)}$ following (33)
- Compute the wavelet transform of the unwarped signal $\hat{\boldsymbol{W}}^{(k)}[m] = \mathcal{W}^{(k)}_{D_{\gamma^{-1}}\boldsymbol{Y}}[m,n]$
- Given the covariance matrix of $\hat{\boldsymbol{W}}^{(k;n)}$, estimate $\hat{\delta}^{(k+1)}$ using (32).
- $k := k+1$

end while

5. Numerical Results

We illustrate in this section the results obtained from the proposed approaches in both cases, namely estimation of modulation or time warping. In the modulation case, we limit to results on synthetic signals for simplicity. We notice however that this case could be of practical interest, for instance for micro-Doppler estimation, of BLU demodulation. For the case of time warping, numerical results include both simulations and results on real signals (car engine sounds).

5.1. Estimation of modulation

The algorithm has been implemented using the Matlab scientific environment, and the time-frequency tools implemented in the toolbox LTFAT [17]. Since corresponding numerical results have already been published elsewhere, we simply illustrate here the method on a synthetic signal, and refer to [8] for more details.

We display in Fig. 1 the results obtained using the proposed algorithm on a synthetic random signal. The original stationary signal $\boldsymbol{X}$ was a band-pass filtered Gaussian white noise, and the synthesized signal was obtained by frequency modulation with a given modulation function (displayed with a yellow curve on the time-frequency plot). Gabor transform was computed using the LTFAT toolbox [17], using a periodized Gaussian window function. The squared modulus of the Gabor transform is displayed in FIG 1, and the original frequency modulation (yellow) and the estimated one (red) are superimposed. The algorithm is clearly able to reproduce the frequency modulation curve with good accuracy.

We refer to [8] for more complete results. Since we aim at studying real signals, for which the modulation model is often not suitable, we limit the illustrations of the latter to that example, and focus on the time warping model for which we present more complete results below.

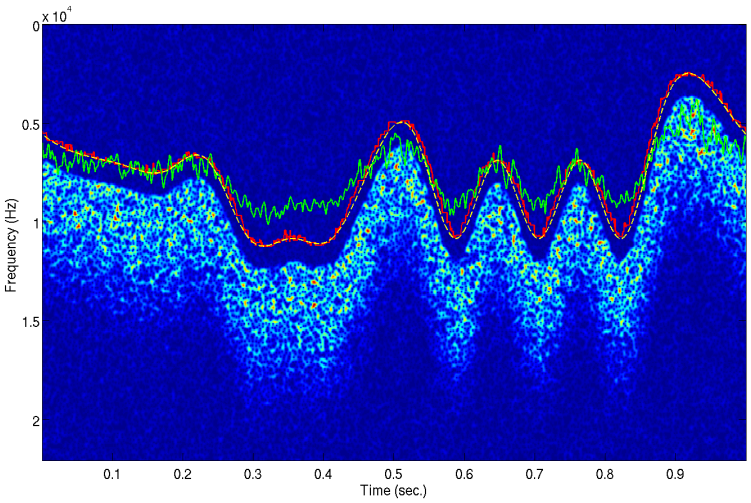

Figure 1: Gabor transform of a synthetic signal and frequency modulation estimation: ground truth (yellow, dotted) and estimates using the local frequency function (green) and the proposed algorithm (red).

5.2. Estimation of time warping

We now illustrate results obtained with the time warping estimation algorithm, starting with the validation of the method on synthetic signals. A synthetic time warped signal was generated, by warping a band pass random stationary signal using a prescribed warping function. The warping function derivative γ' was then estimated (up to a constant function) using the proposed method, as well as the local scale function defined in (9) (which was also used for initializing the proposed approach). The wavelet transform was computed using an analytic derivative of Gaussian wavelet, defined by its Fourier transform $\hat{\boldsymbol{\psi}}$ which vanishes for negative frequencies and reads on the positive half axis

$$\hat{\psi}(\nu) = \nu^k e^{-\alpha\nu^2} \text{ for } \nu \in \mathbb{R}^+ , \qquad \hat{\psi}(\nu) = 0 \text{ for } \nu \in \mathbb{R}^-$$

with α a positive integer number, tuned so that the mode of $\hat{\boldsymbol{\psi}}$ is located at $\mathsf{fs}/4$, fs being the sampling frequency. Examples of such wavelets are displayed in Fig 2. We notice that the parameter α controls the smoothness of $\hat{\boldsymbol{\psi}}$ at the origin of frequencies, and thus the decay of $\boldsymbol{\psi}$, which play an important role in the error estimates in Theorem 2. The parameters of the wavelet analysis were set as follows: we used a very fine scale sampling, $q = 2^{-70}$, and m was chosen the interval $m \in [\![0; 279]\!]$.

We display in Fig 3 the wavelet transform of a synthetic time warped stationary signal, superimposed with the warping function derivative, together with two

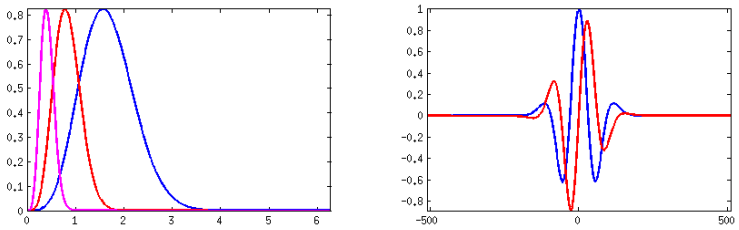

Figure 2: Analytic derivative of Gaussian wavelets ($k = 4$). Left: Fourier transforms, with three different values of the scale parameter. Right: real part (blue) and imaginary part (red) of a wavelet in the time domain.

estimates: the local scale function and the proposed method. As can be seen, similar to what was shown in the case of modulation, the estimate provided by the proposed method turns out to be quite accurate.

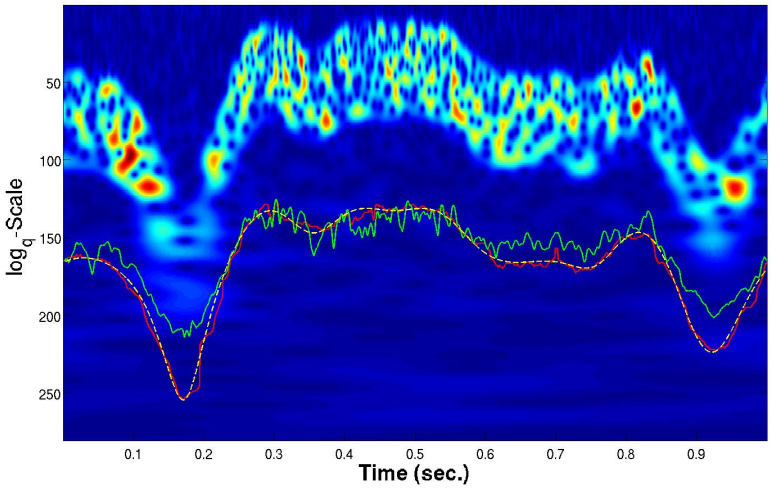

Figure 3: Scalogram of a synthetic time warped random signal, and warping function estimation: ground truth (yellow, dotted) and estimates using the local scale function (green) and the proposed algorithm (red).

To assess quantitatively the accuracy of the proposed algorithm and compare with alternative approaches, we performed a simulation study: 500 realizations of a time-warped random signal were generated following (29), using the same warping function, and the latter was estimated using the proposed approach and the local scale function in (9). We provide in TABLE 1 the normalized average error (i.e. the euclidean norm of $\gamma' - \tilde{\gamma}'$, normalized by the euclidean

	Simulation 1		Simulation 2		Simulation 3	
	err.	var.	err.	var.	err.	var.
$\alpha = 70$: Local scale	0.394	32.08	0.388	32.05	0.391	32.04
$\alpha = 70$: Proposed method	0.115	6.61	0.126	7.76	0.151	10.44
$\alpha = 25$: Local scale	0.248	19.53	0.259	21.56	0.358	28.55
$\alpha = 25$: Proposed method	0.129	7.91	0.165	11.25	0.257	20.54
$\alpha = 7$: Local scale	0.296	33.47	0.294	34.29	0.315	37.90
$\alpha = 7$: Proposed method	0.226	24,75	0.236	26.47	0.275	33.03

Table 1: Performance analysis of time warping estimation methods for three different analyzing wavelets (analytic derivative of Gaussian function of respective degree $\alpha = 70$, top, $\alpha = 25$, middle and $\alpha = 7$, bottom): average error and variance of the estimate.

norm of γ', where $\tilde{\gamma}'$ is the estimate), averaged over all realizations. We also provide in the same table the variances of the estimate. The time-warping functions under consideration were sine waves with period equal to the signal length (simulation 1), half the signal length (simulation 2) and one fourth of the signal length (simulation 3). The numerical results clearly show that the proposed approach outperforms very significantly the results obtained using local scale, both in terms of error and variance. Three different wavelets were tested, with respectively $\alpha = 70$, $\alpha = 25$ and $\alpha = 7$. Numerical results show that better localized wavelets (here $\alpha = 70$) yield more accurate estimates, which is in accordance with Theorem 2. We didn't reproduce here results obtained with the Mallat & Clerc algorithm, because the latter requires averagin over a large number of realizations of the random signal to yield results comparable to the ones described here. This originates from the fact that the estimate obtained in [7] is very local and therefore sensitive to fluctuations. In addition, the iterative nature of our approach and the corresponding covariance re-estimation at each iteration seems to better exploit the underlying stationarity.

Finally, the algorithm was also tested on real signals. We display in Fig. 4 (top) the result obtained on an accelerating car engine. As before, the estimate provided by the proposed algorithm turns out to be significantly smoother than the local scale function estimate. After estimation, the warping function can be inverted to produce a "stationarized" signal, the scalogram of which is also displayed in Fig. 4 (bottom). Corresponding sound files are available (together with additional examples) at the web site

`https://www.i2m.univ-amu.fr/~omer/SounDef/`

that show that the stationarized signal indeed captured the main aspects of the timbre of the original.

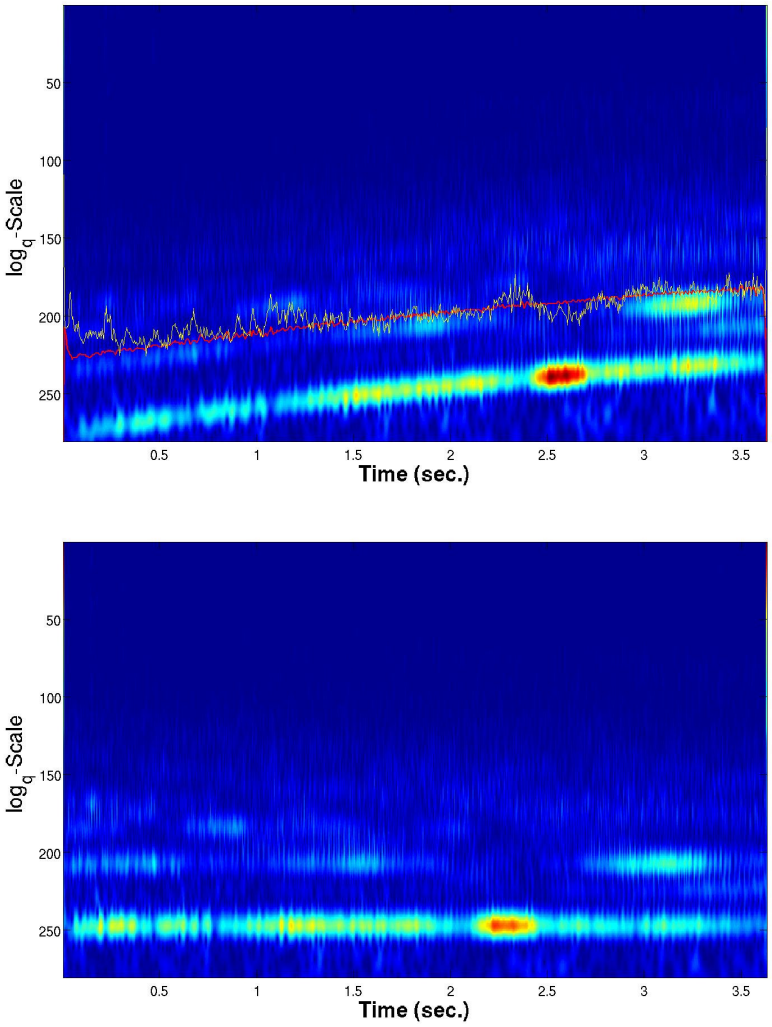

Figure 4: Top: Scalogram of an accelerating car engine sound, and time warping function estimation: estimates using the local scale function (yellow) and the proposed algorithm (red). Bottom: Scalogram of the "stationarized" signal.

6. Proofs of the main results

6.1. Proof of Theorem 1.

The idea is to use the localization of the window function g around the origin, and "freeze" the modulation function γ at the location $t = na$ where the

Gabor function is centered. For the sake of simplicity, let us set

$$\begin{aligned} G &= \sum_{t=0}^{L-1} Z_t e^{2i\pi\gamma(t)/L} \overline{g}[t-na] e^{-2i\pi mb[t-na]/L} , \\ \tilde{G} &= \sum_{t=0}^{L-1} Z_t \overline{g}[t-na] e^{2i\pi\gamma(na)/L} e^{-2i\pi[t-na][mb-\gamma'(na)]/L} , \end{aligned}$$

and set $R = G - \tilde{G}$. Then

$$R = \sum_{t=0}^{L-1} Z_t \overline{g}[t-na] e^{2i\pi\gamma(t)} e^{-2i\pi mb[t-na]/L} [1 - e^{2i\pi r(t)}] ,$$

where $r(t) = \gamma(t) - \gamma(na) - [t-na]\gamma'(na)$; clearly, $|r| \leq \frac{1}{2}[t-na] \sup(|\gamma''|)$. Let us now consider the size estimate for the remainder $\mathbb{E}\left\{|R|^2\right\}$. We consider the centered interval $I^c =]\!] - L/2, L/2]\!]$, which we split into two parts, namely small values $I_1 = \{t \in I^c : |t| \leq T\}$ and large values $I_2 = \{t \in I^c : |t| > T\}$, with $T = \sqrt{L/\pi\|\gamma''\|_\infty}$. Then

$$\begin{aligned} \mathbb{E}\left\{|R|^2\right\} &= \sigma_Z^2 \left| \sum_t \overline{g}[t-na] e^{2i\pi\gamma(t)/L} \left[1 - e^{2i\pi r(t)/L}\right] \right|^2 \\ &\leq \sigma_Z^2 \Big(\sum |g[t]| \left|1 - e^{2i\pi r[t+na]}\right| \Big)^2 \\ &\leq \sigma_Z^2 \left(2\sum_{t\in I_1} |g[t]| + 2\pi \sum_{t\in I_2} |g[t]||r[t]| e^{2\pi|r[t]|/L} \right)^2 \\ &\leq \sigma_Z^2 \left(2\mu_1 + \frac{\pi\|\gamma''\|_\infty}{L} \sum_{t\in I_2} t^2 |g[t]| e^{\pi t^2 \|\gamma''\|_\infty/L} \right)^2 \\ &\leq \sigma_Z^2 \left(2\mu_1 + \frac{\pi e\|\gamma''\|_\infty}{L} \mu_2 \right)^2 , \end{aligned}$$

where we have used the inequality $|e^z - 1| \leq |z|e^{|z|}$, and where $\mu_1 = \sum_{t\in I_1} |g[t]|$ and $\mu_2 = \sum_{t\in I_2} t^2 |g[t]|$ characterize respectively the behavior of the window $\boldsymbol{g}$ on its tail and near the origin.

This concludes the proof. □

6.2. Proof of Proposition 2

The covariance matrix of the Gabor transform (see for example [12] for the general case) is given by

$$C_{\boldsymbol{G}}[m, m'] = \sum_{\nu\in I_c} (\mathcal{S}_{\boldsymbol{Z}}[\nu - \delta] + \sigma_0^2) \overline{\hat{g}}[\nu - mb] \hat{g}[\nu - m'b] .$$

Let $\boldsymbol{x} \in \mathbb{C}^M$, and compute

$$\begin{aligned}\langle C_{\boldsymbol{G}}\boldsymbol{x}, \boldsymbol{x}\rangle &= \sum_{m,m'=0}^{M-1} x[m]\overline{x}[m'] \sum_{\nu\in I_c} (\mathcal{S}_{\boldsymbol{Z}}[\nu-\delta]+\sigma_0^2)\hat{g}[\nu-mb]\overline{\hat{g}}[\nu-m'b] \\ &= \sum_{\nu\in I_c} (\mathcal{S}_{\boldsymbol{Z}}[\nu-\delta]+\sigma_0^2)|\hat{h}[\nu]|^2 \geq \sigma_0^2\|\hat{\mathbf{h}}\|^2 \ ,\end{aligned}$$

where we have set $\hat{h}[\nu] = \sum_{m=0}^{M-1} \hat{g}[\nu-mb]x[m]$. Now, $h[t] = (\sum_{m=0}^{M-1} x[m]e^{2i\pi mt/M})g[t]/L$ and the standard periodization argument yields

$$\begin{aligned}\|\hat{\mathbf{h}}\|^2 &= L\|\mathbf{h}\|^2 \\ &= \frac{L}{L^2}\sum_{t=0}^{L-1}|g[t]|^2\left|\sum_{m=0}^{M-1} x[m]e^{2i\pi mt/M}\right|^2 \\ &= \frac{1}{L}\sum_{u=0}^{M-1}\sum_{k=0}^{b-1}|g[u-kM]|^2\left|\sum_{m=0}^{M-1} x[m]e^{2i\pi mu/M}\right|^2 \\ &\geq \frac{1}{L}\min_{u\in[\![0,M-1]\!]}\left(\sum_{k=0}^{b-1}|g[u-kM]|^2\right)\|\hat{\boldsymbol{x}}\|^2 \\ &\geq \frac{1}{b}\min_{u\in[\![0,M-1]\!]}\left(\sum_{k=0}^{b-1}|g[u-kM]|^2\right)\|\boldsymbol{x}\|^2\end{aligned}$$

This concludes the proof of the proposition. □

It is worth noticing that this proof, as often in the case of Gabor analysis, carries through *mutatis mutandis* to the case of Gabor analysis for continuous time signals.

6.3. Proof of Lemma 1.

The proof follows from the unitarity of the warping operator. The Gaussian white noise with variance σ_0^2 is characterized by its characteristic function $P_{\boldsymbol{N}}(\varphi) = \mathbb{E}\left\{e^{i\langle \boldsymbol{N},\varphi\rangle}\right\} = e^{-\|\varphi\|^2/2\sigma_0^2}$, for all test function in the Schwartz space $\varphi \in \mathcal{S}(\mathbb{R})$. Similarly, the warped white noise's characteristic function reads

$$P_{D_\gamma \boldsymbol{N}}(\varphi) = \mathbb{E}\left\{e^{i\langle D_\gamma \boldsymbol{N},\varphi\rangle}\right\} = \mathbb{E}\left\{e^{i\langle \boldsymbol{N},{D_\gamma}^{-1}\varphi\rangle}\right\} = e^{-\|{D_\gamma}^{-1}\varphi\|^2/2\sigma_0^2} = e^{-\|\varphi\|^2/2\sigma_0^2}$$

which implies that $D_\gamma \boldsymbol{N}$ is a Gaussian white noise, with the same variance. □

6.4. Proof of Theorem 2

Let us compute the difference between the actual wavelet transform of the deformed process and the approximation. Fix a time index n, and denote by

$\tilde{\gamma}_n(t) = \gamma(na) + (t - na)\gamma'(na)$ the first order approximation of γ in the neighborhood of the center $t = na$ of the wavelet ψ_{mn}. Set also $D^\circ_\gamma = D_\gamma/\sqrt{\gamma'}$ and define $D^\circ_{\tilde{\gamma}_n}$ similarly. Write

$$\begin{aligned}\epsilon_{mn} &= \mathcal{W}_{\boldsymbol{Y}}[m,n] - \mathcal{W}_{\boldsymbol{X}}(m + \log_q(\gamma'(an)), \gamma(an)) \\ &= \langle (D_\gamma - D_{\tilde{\gamma}_n})\boldsymbol{X}, \psi_{mn}\rangle = \epsilon^{(0)}_{mn} + \epsilon^{(1)}_{mn}\end{aligned}$$

where we have set

$$\begin{aligned}\epsilon^{(0)}_{mn} &= \left\langle \left(\sqrt{\gamma'} - \sqrt{\gamma'(na)}\right) D^\circ_\gamma \boldsymbol{X}, \psi_{mn}\right\rangle \\ \epsilon^{(1)}_{mn} &= \sqrt{\gamma'(na)}\,\langle (D^\circ_\gamma - D^\circ_{\tilde{\gamma}_n})\boldsymbol{X}, \psi_{mn}\rangle\end{aligned}$$

We bound those two terms independently. Consider first $\epsilon^{(0)}_{mn}$, and write

$$\mathbb{E}\left\{|\epsilon^{(0)}_{mn}|^2\right\} = \int_{-\infty}^{\infty} \mathcal{S}_{\boldsymbol{X}}(\nu)\,|g(\nu)|^2\,d\nu\ ,$$

where

$$\begin{aligned}g(\nu) &= \left|\int_{-\infty}^{\infty} e^{2i\pi\nu\gamma(t)}\left(\sqrt{\gamma'(t)} - \sqrt{\gamma'(na)}\right)\psi_{mn}(t)dt\right| \\ &\le \frac{1}{2}\left\|\frac{\gamma''}{\sqrt{\gamma'}}\right\|_\infty q^{-m/2}\int_{-\infty}^{\infty}|t-na|\,\left|\psi\left(q^{-m}(t-na)\right)\right|dt \\ &\le \frac{\mu}{2}\left\|\frac{\gamma''}{\sqrt{\gamma'}}\right\|_\infty q^{3m/2}\ ,\ \text{where } \mu = \int_{-\infty}^{\infty}|s|\,|\psi(s)|\,ds\ .\end{aligned}$$

Assume now that $|\psi(t)| \le (1 + |t|^\alpha)^{-1}$. Then

$$\mu \le 2\int_0^\infty \frac{s\,ds}{1+|s|^\alpha} \le 2\int_0^1 s ds + 2\int_1^\infty s^{1-\alpha}ds \le \frac{\alpha}{2(\alpha-2)}$$

All together, letting $\sigma^2_{\boldsymbol{X}} = \|\mathcal{S}_{\boldsymbol{X}}\|_1$ be the variance of $\boldsymbol{X}$, we obtain

$$\mathbb{E}\left\{|\epsilon^{(0)}_{mn}|^2\right\} \le \frac{\alpha^2\sigma^2_{\boldsymbol{X}}}{16(\alpha-2)^2}\left\|\frac{\gamma''}{\sqrt{\gamma'}}\right\|^2_\infty q^{3m}\ .$$

Consider now $\epsilon^{(1)}_{mn} = \sqrt{\gamma'(na)}\,\langle (D^\circ_\gamma - D^\circ_{\tilde{\gamma}_n})\boldsymbol{X}, \psi_{mn}\rangle$. Then

$$\begin{aligned}\mathbb{E}\left\{\left|\epsilon^{(1)}_{mn}\right|^2\right\} &= q^{-m}\gamma'(na)\int_{-\infty}^{\infty}\mathcal{S}(\nu)\left|\int e^{2i\pi\nu\tilde{\gamma}_n(t)}\ \times\right. \\ &\qquad \left.\left[1 - e^{i\pi\nu(t-na)^2 r(t)}\right]\overline{\psi}\left((q^{-m}(t-na)\right)dt\right|^2 d\nu \\ &\le 4q^{-m}\gamma'(na)\int_{-\infty}^{\infty}\mathcal{S}_{\boldsymbol{X}}(\nu)f(\nu)^2 d\nu\ ,\end{aligned}$$

where f (that actually also depends on n and m) is bounded as

$$\begin{aligned} f(\nu) &= \frac{1}{2}\left| \int e^{2i\pi\nu\tilde{\gamma}_n(t)} \left[1-e^{i\pi\nu(t-na)^2 r(t)}\right] \overline{\psi}\big((q^{-m}(t-na)\big)\, dt \right| \\ &\leq \int_{-\infty}^{\infty} \left|\sin\left(\frac{\pi}{2}\nu(t-na)^2 r(t)\right)\right| |\psi\left(q^{-m}(t-na)\right)| dt \\ &\leq q^m \int_{-\infty}^{\infty} |\psi(s)| \left|\sin\left(\frac{\pi}{2}\nu q^{2m} s^2 r(t)\right)\right| ds \\ &\leq q^m \left[\int_{I_1(\nu)} |\psi(s)| ds + \frac{\pi\|\gamma''\|_\infty}{2}|\nu| q^{2m} \int_{I_2(\nu)} s^2 |\psi(s)| ds\right] . \end{aligned}$$

Here we have introduced a splitting of the real axis as the union of two domains (that depend on the frequency ν)

$$I_1(\nu) = \{s : |s| \geq u_0\}, \quad I_2(\nu) = \mathbb{R}\backslash I_1(\nu)$$

where u_0 is a free parameter that can be adapted. We have also used the bounds $|\sin(u)| \leq 1$ and $|\sin(u)| \leq |u|$ within those two domains.

Let us set for simplicity $C_0 = \frac{\pi\|\gamma''\|_\infty}{2}|\nu| q^{2m}$. Using the decay of the wavelet ψ, namely $|\psi(t)| \leq (1+|t|^\alpha)^{-1}$, we obtain

$$\int_{I_1(\nu)} |\psi(s)| ds \leq \frac{2}{\alpha-1} u_0^{1-\alpha} , \quad \int_{I_2(\nu)} s^2 |\psi(s)| ds \leq \frac{2}{3} u_0^3 .$$

Then

$$f(\nu) \leq 2q^m F(u_0) ,$$

where

$$F(u_0) = \frac{u_0^{1-\alpha}}{\alpha-1} + C_0 \frac{u_0^3}{3} .$$

The value $u_{0\mathsf{opt}}$ of u_0 that minimizes the latter expression is easily found to be $u_{0\mathsf{opt}} = C_0^{-\frac{1}{\alpha+2}}$, which leads to the bound

$$F(u_{0\mathsf{opt}}) = \frac{\alpha+2}{3(\alpha-1)} C_0^{1-\frac{3}{\alpha+2}}.$$

Therefore,

$$f(\nu) \leq 2q^m \frac{\alpha+2}{3(\alpha-1)} \left(\frac{\pi\|\gamma''\|_\infty}{2}|\nu| q^{2m}\right)^{1-\frac{3}{\alpha+2}}$$

Putting things together, we finally obtain

$$\begin{aligned} \mathbb{E}\left\{\left|\epsilon_{mn}^{(1)}\right|^2\right\} &\leq \frac{16(\alpha+2)^2}{9(\alpha-1)^2}\gamma'(an)\left(\frac{\pi\|\gamma''\|_\infty}{2}\right)^{2-\frac{6}{\alpha+2}} \\ &\quad \times (q^m)^{5-\frac{12}{\alpha+2}} \int_{-\infty}^{\infty} |\nu|^{2-\frac{6}{\alpha+2}} \mathcal{S}_{\boldsymbol{X}}(\nu) d\nu . \end{aligned}$$

Finally, bound the total error by

$$\mathbb{E}\left\{|\epsilon_{mn}|^2\right\} \leq \sqrt{\mathbb{E}\left\{|\epsilon_{mn}^{(0)}|^2\right\}} + \sqrt{\mathbb{E}\left\{|\epsilon_{mn}^{(1)}|^2\right\}},$$

concludes the proof. □

6.5. Proof of Proposition 4.

Let n be a fixed time index. We consider the covariance matrix of the approximate wavelet transform $\boldsymbol{W} = \boldsymbol{W}^{(\gamma(n);\log_q(\gamma'(n)))}$, which takes the form

$$C_{\boldsymbol{W}}^{(n)}[m,m'] = \mathbb{E}\left\{\boldsymbol{W}[m]\overline{\boldsymbol{W}}[m']\right\} = q^{(m+m')/2}\left\langle (\mathcal{S}_{\boldsymbol{X}} + \sigma_0^2), \hat{\psi}^m\overline{\hat{\psi}^{m'}}\right\rangle.$$

Let $\boldsymbol{x} \in \ell^2(\mathbb{Z})$, and set $y(\nu) = \sum_{m=-\infty}^{\infty} q^{m/2}x[m]\hat{\psi}\left(q^m\nu\right)$, for $\nu \in \mathbb{R}^+$. Then

$$\langle C_{\boldsymbol{W}}^{(n)}\boldsymbol{x}, \boldsymbol{x}\rangle = \langle (\mathcal{S}_{\boldsymbol{X}} + \sigma_0^2), |y|^2\rangle \geq \sigma_0^2\|\boldsymbol{y}\|^2 \geq \sigma_0^2\|\underline{\boldsymbol{y}}\|^2 \ ,$$

with $\underline{\boldsymbol{y}}$ the Mellin transform of $\boldsymbol{y}$. An explicit calculation shows that

$$\underline{y}(s) = \underline{\psi}(s)\,\hat{x}(s\ln(q)) \ ,$$

and therefore a standard periodization argument yields

$$\begin{aligned}
\|\underline{\boldsymbol{y}}\|^2 &= \int_{-\infty}^{\infty} |\underline{\psi}(s)|^2|\hat{x}(s\ln(q))|^2\,ds \\
&= \int_0^{1/\ln(q)} |\hat{x}(s\ln(q))|^2 \sum_{\ell=-\infty}^{\infty}\left|\underline{\psi}\left(s + \frac{\ell}{\ln(q)}\right)\right|^2 ds \\
&\geq K_{\boldsymbol{\psi}}\|\boldsymbol{x}\|^2 \ .
\end{aligned}$$

Putting all estimates together yields the result. □

7. Discussion and conclusions

We have presented in this paper a new approach for the estimation of smooth deformations of Gaussian stationary random signals, that only exploit the stationarity and smoothness assumptions. Two types of deformations were considered, namely smooth modulation and smooth time warping, for which the transformation approximately manifests itself by translations in a suitable representation space. In both cases, an iterative algorithm was proposed, that estimates alternately the deformation and the covariance matrix of the signal. The analysis was very similar in both cases, though the time warping case presents additional mathematical difficulties. The modulation case was presented in a discrete, finite-dimensional setting, but can also be developed in the continuous time setting, following the lines of what was done for time warping.

The presented numerical simulations show the reliability of the proposed approach, and results on real world signals, though still preliminary, are quite

encouraging. Further tests will involve validation on a database of real sounds, and may require perceptive tests. Among the possible improvements of our approach, a more thorough study of the optimization algorithms (for which we have chosen to stick to a simple and robust approach) will be a next goal.

Let us stress that the motivation of this work was mainly in the field of audio signal processing, namely problems where an underlying transformation (related to a clock change, as is the case for sounds produced by rotating engines with variable rotation speed) is to be estimated. Nevertheless, the problem is fairly general and can be transposed to different contexts, such as the shape from texture problem in image processing (see [6]).

Let us also notice that the models introduced here form a new family of models for linear systems (or operators) represented directly in time-frequency or time scale space. Contrary to multiplier models (see for example [18, 19] that are defined by pointwise multiplications in the representation domains, and that can be estimated (see for example [20]), these transformation models implement translations in that space. The proposed approach may therefore represent an interesting starting point for system estimation in the case of such categories of systems.

Acknowledgments

This work was supported by Agence Nationale de la Recherche (ANR), in the framework of the Metason project (ANR-10-CORD-010). H. Omer's is supported by Agence Nationale de la Recherche (same project).

References

[1] P. Flandrin, Time-Frequency/Time-Scale Analysis. Academic Press, 1999.

[2] N. E. Huang and S. S. P. Shen, Hilbert-Huang Transform and its Applications. World Scientific, 2005.

[3] F. Auger, P. Flandrin, Y.-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, and H.-T. Wu, "Time-frequency reassignment and synchrosqueezing: An overview," IEEE Signal Processing Magazine, vol. 30, no. 6, pp. 32–41, 2013.

[4] H.-T. Wu, "Instantaneous frequency and wave shape function (i)," Applied and Computational Harmonic Analysis, vol. 35, pp. 181–199, 2013.

[5] J. Malik and R. Rosenholtz, "Computing local surface orientation and shape from texture for curved surfaces," International Journal of Computer Vision, vol. 23, no. 2, pp. 149–168, 1997.

[6] M. Clerc and S. Mallat, "The texture gradient equation for recovering shape from texture," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 4, pp. 536–549, 2002.

[7] ——, "Estimating deformations of stationary processes," Annals of Statistics, vol. 31, no. 6, pp. 1772–1821, 2003.

[8] H. Omer and B. Torrésani, "Estimation of frequency modulations on wideband signals; applications to audio signal analysis," in Proceedings of the 10th International Conference on Sampling Theory and Applications (SampTA), G. Pfander, Ed. Eurasip Open Library, 2013, pp. 29–32. [Online]. Available: http://hal.archives-ouvertes.fr/hal-00822186

[9] B. Picinbono, "On circularity," IEEE Transactions on Signal Processing, vol. 42, no. 12, pp. 3473–3482, 1994.

[10] M. Unser and P. Tafti, Sparse stochastic processes. Cambridge University Press, 2013.

[11] P. Soendergaard, "Finite discrete Gabor analysis," Ph.D. dissertation, Institut for Matematik Denmark Technical University, 2007.

[12] R. Carmona, W. L. Hwang, and B. Torrésani, Practical time-frequency analysis: Gabor and Wavelet Transforms With an Implementation in S, C. K. Chui, Ed. Academic Press, 1998.

[13] K. Gröchenig, Foundations of time-frequency analysis, ser. Applied and Numerical Harmonic Analysis. Boston, MA: Birkhäuser Inc., 2001.

[14] P. L. Søndergaard, "Efficient Algorithms for the Discrete Gabor Transform with a long FIR window," J. Fourier Anal. Appl., vol. 18, no. 3, pp. 456–470, 2012.

[15] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation." IEEE Trans. Pat. Anal. Mach. Intell., vol. 11, pp. 674–693, 1989.

[16] H. Omer, "Modèles de déformation de processus stochastiques généralisés. application à l'estimation des non stationnarités dans les signaux audio." Ph.D. dissertation, Aix-Marseille Université, 2015, in preparation.

[17] P. Soendergaard, B. Torrésani, and P. Balazs, "The linear time frequency analysis toolbox," International Journal of Wavelets and Multiresolution Information Processing, vol. 10, no. 4, pp. 1 250 032–1 – 1 250 032–27, 2012.

[18] H. Feichtinger, M. Hampejs, and G. Kracher, "Approximation of matrices by gabor multipliers," IEEE Signal Processing Letters, vol. 11, no. 11, pp. 883 – 886, 2004.

[19] M. Dörfler and B. Torrésani, "On the time-frequency representation of operators and generalized Gabor multiplier approximations," Journal of Fourier Analysis and Applications, vol. 16, pp. 261–293, 2010.

[20] A. Olivero, B. Torrésani, and R. Kronland-Martinet, "A class of algorithms for time-frequency multiplier estimation," IEEE Trans. Audio, Speech and Language Processing, vol. 21, no. 8, pp. 1550 – 1559, 2013.