HAL
open science

Spectral estimation for non-stationary signal classes

Adrien Meynard, Bruno Torrésani

To cite this version:

Adrien Meynard, Bruno Torrésani. Spectral estimation for non-stationary signal classes. 2017 International Conference on Sampling Theory and Applications (SampTA), Jul 2017, Tallinn, Estonia. pp.174-178, 10.1109/SAMPTA.2017.8024367 . hal-01475534v2

HAL Id: hal-01475534

https://hal.science/hal-01475534v2

Submitted on 25 Apr 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Spectral estimation for non-stationary signal classes

Adrien Meynard[1] and Bruno Torrésani[2]
1: Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France
2: CNRS, Université de Montréal, CRM, UMI 3457, Montréal, Canada
adrien.meynard@ens-cachan.fr bruno.torresani@univ-amu.fr

***Abstract*—An approach to the spectral estimation for some classes of non-stationary random signals is developed, that addresses stationary random processes deformed by a stationarity-breaking transformation. Examples include frequency modulation, time warping, non-stationary filtering and others. Under suitable smoothness assumptions on the transformation, approximate expressions are obtained in adapted representation spaces. In the Gaussian case, this leads to approximate maximum likelihood estimation algorithms, which are illustrated on synthetic as well as real signals.**

***Index Terms*—Non-stationary processes, deformation, time-frequency/scale analysis, local spectrum**

I. Introduction

Classical spectral theory for random processes concerns classes of stationary processes, i.e. processes that possess some generic shift invariance properties. As a convolution operator, the covariance operator of such processes is diagonalized by the Fourier transform, and the power spectrum thus coincides with the spectrum of the operator. Classes of non-stationary processes have also been considered by several authors in the literature. These include (among many others) works by Priestley and followers who introduced the notion of evolutionary spectrum [1], by Dahlhaus on locally stationary processes [2], Mallat et al on adaptive covariance estimation [3], time-frequency analysis based techniques [4].

We follow here another point of view, initiated by Clerc and Mallat in [5], [6] and reconsidered recently in [7], [8], which addresses stationary processes deformed by a stationarity breaking linear operator $\mathscr{T}$. This approach can be understood as modeling the covariance operator as the product $\mathscr{T}\mathscr{C}$ of a convolution operator $\mathscr{C}$ with $\mathscr{T}$. Spectral estimation can then be viewed as the joint estimation of $\mathscr{T}$ and the spectrum of $\mathscr{C}$. This problem is still far too general, but can be addressed if one restricts to semi-parametric classes $\mathscr{T}_\alpha$, where α is a function to be estimated together with the spectrum of $\mathscr{C}$.

Under some smoothness assumptions on the deformation $\mathscr{T}_\alpha$, joint estimation of α and the spectrum of $\mathscr{C}$ can be possible in a suitable representation space, where an approximate expression for the likelihood can be derived [7], [8], [9]. This includes time warpings, frequency modulations as well as generalized deformation models, for which adapted representation spaces are proposed. Besides these generalizations, the main contribution of this paper is a continuous parameter maximum likelihood approach, which allow us to derive corresponding Cramér-Rao bounds. The approach is illustrated on simulations as well as real non-stationary signals.

II. Non-stationary signals: models and approximations

A. Notations and background

Let us first introduce some notations. Throughout this paper, $\mathcal{S}$ and $\mathcal{S}'$ denote respectively the Schwartz space and the space of tempered distributions. By abuse of notations, we also denote by $\mathcal{S}'$ the space of generalized stochastic processes whose trajectories are almost surely tempered distributions. In what follows, we denote by $\langle .,. \rangle$ the L^2 inner product (or duality pairing), and reserve the symbol "$\cdot$" for the finite-dimensional Hermitian product.

Translation, modulation and rescaling operators are respectively denoted by T_τ, M_ν and D_s, with the conventions[1]

$$T_\tau x(t) = x(t-\tau),\ M_\nu x(t) = e^{2i\pi\nu t}x(t),\ D_s x(t) = q^{-s/2}x\big(q^{-s}t\big)$$

where $q > 0$ is a fixed number. With these notations, the short time Fourier transform and the continuous wavelet transform associate with any generalized random process $X \in \mathcal{S}'$ the random fields $\mathcal{G}_X$ and $\mathcal{W}_X$ defined respectively by

$$\mathcal{G}_X(\nu,\tau) = \langle X, M_\nu T_\tau \psi \rangle \ , \quad \mathcal{W}_X(s,\tau) = \langle X, T_\tau D_s \psi \rangle \ ,$$

where $\psi \in \mathcal{S}$ is the analysis waveform (window, or wavelet depending on the context).

B. Deformation models

We use here the formalism developed in [9], referring to [10] and [11] for more details on generalized random processes and signal processing applications. Let $X \in \mathcal{S}'$ be a stationary generalized random process, and let $\mathcal{C}_X : \mathcal{S} \to \mathcal{S}'$ and $\mathscr{S}_X \in \mathcal{S}'$ denote respectively the associated covariance operator and generalized power spectrum. Assume we are given observations of the form

$$X \longmapsto Y = \mathscr{T}(X)$$

where $\mathscr{T} : \mathcal{S}' \to \mathcal{S}'$ is a linear operator. We aim at estimating both $\mathscr{T}$ and $\mathscr{S}_X$ from Y, using solely the assumption of stationarity of X. To make the problem tractable, we actually formulate it as a semi-parametric problem, and limit ourselves to a semi-parametric family of unitary transformations $\mathscr{T}_\alpha$, whose functional parameter α we want to estimate. Examples of such families include frequency modulations

$$\mathscr{M}_\alpha : \qquad \mathscr{M}_\alpha x(t) = e^{2i\pi\alpha(t)}x(t) \ , \tag{1}$$

[1]Notice our convention for dilation differs from the standard one, as we use a logarithmic scale variable.

where $\alpha \in C^2$ is a smooth function, time warpings

$$\mathscr{D}_\gamma : \qquad \mathscr{D}_\gamma x(t) = \sqrt{\gamma'(t)} x(\gamma(t)) \tag{2}$$

where γ is a smooth, strictly increasing function, assumed to fulfill the control condition

$$0 < c_\gamma \leq \gamma'(t) \leq C_\gamma < \infty \qquad \forall t \ , \tag{3}$$

for some constants c_γ, C_γ (see [7] for stronger conditions that ensure continuity of $\mathscr{D}_\gamma$ on $\mathcal{S}'$). Other models of interest but not considered here include semi-parametric families of time-varying filters, i.e. convolution filters whose kernel varies smoothly as a function of time.

In these cases, approximations rely on vector fields of so-called tangent operators [9], denoted respectively by $\widetilde{\mathscr{M}}_\alpha^\tau$ and $\widetilde{\mathscr{D}}_\gamma^\tau$; given a smooth function g, supposed to be localized near a given location τ, one can write

$$\begin{aligned} \mathscr{M}_\alpha g(t) &\approx \widetilde{\mathscr{M}}_\alpha^\tau g(t) \ , \text{ with } \widetilde{\mathscr{M}}_\alpha^\tau = e^{2i\pi\alpha(\tau)} T_\tau M_{\alpha'(\tau)} T_{-\tau} \\ \mathscr{D}_\gamma g(t) &\approx \widetilde{\mathscr{D}}_\gamma^\tau g(t) \ , \quad \text{with } \widetilde{\mathscr{D}}_\gamma^\tau = T_\tau D_{-\log_q(\gamma'(\tau))} T_{-\gamma(\tau)} \ . \end{aligned}$$

This justifies the use of short time Fourier (resp. wavelet) transform to carry the estimation of α (resp. γ) in the frequency modulation (resp. time warping) model.

In the following we shall consider combinations of time warping and frequency modulations, which allow describing more general transformations, as a generic framework (more standard cases considered in [8] can be obtained as particular cases). This naturally leads to the extended transform (introduced in [12]) $\mathcal{V} : x \in \mathcal{S}' \to \mathcal{V}_x$:

$$\mathcal{V}_x(\nu, s, \tau) = \langle x, T_\tau M_\nu D_s \psi \rangle \ , \tag{4}$$

where $\psi \in \mathcal{S}$ is a fixed analyzing waveform (examples are given below).

C. Approximation results

Suppose X is a zero mean, circular complex gaussian second order stationary generalized random process, with generalized spectrum $\mathscr{S}_X \in \mathcal{S}'$ (see e.g. [9] and references therein for details). The associated variance is denoted by σ_X^2. Suppose one is given observations of the form

$$Y = \mathscr{M}_\alpha \mathscr{D}_\gamma X + W \ , \tag{5}$$

where α and γ are smooth functions, and where W is a white noise generalized process. Assume the analyzing waveform $\psi \in \mathcal{S}$ is concentrated around the origin. Then, using the tangent operators defined above and standard commutation rules, we obtain the approximate expression

$$\begin{aligned} \mathcal{V}_Y(\nu, s, \tau) &\approx \left\langle \widetilde{\mathscr{M}}_\alpha^\tau \widetilde{\mathscr{D}}_\gamma^\tau X, T_\tau M_\nu D_s \psi \right\rangle := \widetilde{\mathcal{V}}_Y(\nu, s, \tau) \\ &= e^{2i\pi\alpha(\tau)} \mathcal{V}_X\left(\gamma(\tau), \frac{\nu - \alpha'(\tau)}{\gamma'(\tau)}, s + \log_q(\gamma'(\tau))\right) \end{aligned}$$

which indicates that the transformation (5) takes a rather simple form in the chosen representation domain. The approximation is made more precise by the following result, which uses simple decay assumption on the analyzing waveform ψ.

Theorem 1: With the above notations, let $\epsilon = \mathcal{V}_Y - \widetilde{\mathcal{V}}_Y$ denote the approximation error random field. Assume that the analyzing waveform ψ is localized in such a way that $|\psi(t)| \leq 1/(1 + |t|^\beta)$ for some $\beta > 2$, and that for all $u, v \in \mathbb{R}_+$,

$$I(u, v) := \sqrt{\left\langle \mathscr{S}_X, f_{u,v}^{(\beta)} \right\rangle} < \infty \ , \text{ with } f_{u,v}^{(\beta)}(\xi) = (u\xi + v)^{2\frac{\beta-1}{\beta+2}}.$$

Then

$$\mathbb{E}\left\{|\epsilon(s,\nu,\tau)|^2\right\} \leq q^{3s} \left(K_1 \|\gamma''\|_\infty + K_2 q^{s\frac{\beta-4}{\beta+2}} I(\|\gamma''\|_\infty, \|\alpha''\|_\infty)\right)^2$$

where

$$K_1 = \frac{\beta \sigma_X}{2(\beta - 2)\sqrt{c_\gamma}} \ , \quad K_2 = \left(\frac{\pi}{2}\right)^{\frac{\beta-1}{\beta+2}} \sqrt{C_\gamma} \frac{4(\beta+2)}{3(\beta-1)}$$

Remark 1:

1) The proof follows the same arguments as Theorem 2 in [8]: assuming ψ is localized around the origin, Taylor expand α and γ in the neighborhood of $t = \tau$ and approximate the frequency modulation and time warping operators by the corresponding tangent operators. The bound follows from upper bounds in the remainders of Taylor expansions, and the assumed decay of ψ. Similar bounds can be obtained as in [9] assuming different decay properties on ψ.
2) As expected, the bound depends on the smoothness of α' and γ'. It also depends on the power spectrum $\mathscr{S}_X$ in a fairly simple way, and on the waveform decay rate β; conditions on β given the decay of the power spectrum can be deduced.

If ψ is analytic (i.e. $\mathsf{Supp}(\hat{\psi}) \subset \mathbb{R}_+$), and if the analysis is limited to $\nu > \min \alpha'(\tau)$, then the random field $\widetilde{\mathcal{V}}_Y(\nu, s, \tau)$ is zero-mean, circular gaussian with covariance kernel

$$\mathbb{E}\left\{\widetilde{\mathcal{V}}_Y(\nu, s, \tau) \overline{\widetilde{\mathcal{V}}_Y}(\nu', s', \tau')\right\} = \left\langle \mathscr{S}_X, \overline{\widehat{\psi_{\nu,s}^{(\tau)}}} . \widehat{\psi_{\nu',s'}^{(\tau')}} \right\rangle \ , \tag{6}$$

where we have set

$$\begin{aligned} \psi_{\nu,s}^{(\tau)}(t) &= \left[T_{\gamma(\tau)} M_{(\nu - \alpha'(\tau))/\gamma'(\tau)} D_{s + \log_q(\gamma'(\tau))} \psi\right](t) \\ &= \frac{e^{2i\pi\alpha(\tau)}}{\sqrt{\gamma'(\tau)}} e^{2i\pi \frac{\nu - \alpha'(\tau))}{\gamma'(\tau)}(t - \gamma(\tau))} \psi\left(q^{-s} \frac{t - \gamma(\tau)}{\gamma'(\tau)}\right) \ . \end{aligned}$$

Therefore, the approximate covariance depends on the transformation functional parameters $(\alpha, \alpha', \gamma, \gamma')$. The latter are then amenable to (approximate) maximum likelihood estimation, which we describe in the next section.

III. Estimation procedure and algorithms

A. Estimation procedure

The generic estimation scheme is as follows. We describe the most general situation, keeping in mind that the estimation of a single transformation ($\mathscr{M}_\alpha$ or $\mathscr{D}_\gamma$) or a simple combination of both can be performed in a similar way (as discussed below). Suppose as before an observation model of the form (5), and let $y = \mathscr{M}_\alpha \mathscr{D}_\gamma x + w$ be a single realization. Denote by $\Lambda = \Lambda' \times \Lambda_\tau$, with $\Lambda_\tau = \{na, n = n_1 \ldots n_2\}$, $a \in \mathbb{R}^+$ and $\Lambda' \subset \mathbb{R}^2$ a finite sampling subset of the

frequency-scale-time space and by $\mathbf{V}_y = \{\mathcal{V}_y(\lambda),\, \lambda \in \Lambda\}$ the corresponding restriction of $\mathcal{V}_y$. Denote also by $\Theta = \{(\alpha(na), \alpha'(na), \gamma(na), \log_q(\gamma'(na))), n = n_1 \dots n_2\}$ the time-sampled functional parameters, and by $\mathbf{C}(\Theta)$ the restriction of the approximate covariance kernel (6) to the sampling set Λ. The latter can easily be evaluated numerically.

Given an estimate α, γ for the deformation functions, we will generically denote by $Z = \mathscr{D}_\gamma^{-1}\mathscr{M}_\alpha^{-1}Y \approx X + \mathscr{D}_\gamma^{-1}\mathscr{M}_\alpha^{-1}W$ a corresponding estimate for the deformation compensated signal. It is worth noticing that due to the unitarity of the considered deformation operators, $\mathscr{D}_\gamma^{-1}\mathscr{M}_\alpha^{-1}W$ is also a white noise process.

a) A global scheme. Start from the two following observations:

1) Assume that the power spectrum $\mathscr{S}_Z$ of Z is known. Then it follows from the above results that the log likelihood of the observations can be approximated by a smooth, generally non-convex function of Θ

$$\mathscr{L}(\Theta) = -\ln|\det \mathbf{C}(\Theta)| - \mathbf{C}(\Theta)^{-1}\mathbf{V}_y \cdot \mathbf{V}_y \ , \qquad (7)$$

that can be maximized numerically using standard tools, provided $\mathbf{C}(\Theta)$ is invertible.

2) Assume estimates α_e, γ_e for the functional parameters Θ are known. Then they can be used to "stationarize" the observation by computing

$$z_e := \mathscr{D}_{\gamma_e}^{-1}\mathscr{M}_{\alpha_e}^{-1}y \approx x + \mathscr{D}_{\gamma_e}^{-1}\mathscr{M}_{\alpha_e}^{-1}w \ , \qquad (8)$$

and perform spectral estimation on z_e using again standard tools (for example a Welch estimator).

These two steps can be performed iteratively, to yield a joint estimate for Θ and $\mathscr{S}_Z$, thus $\mathscr{S}_X$. This also requires an initial guess for Θ or $\mathscr{S}_Z$, and a stopping criterion. These two aspects will be discussed in the local approach described below.

Remark 2 (Sampling the representation space): For the scheme to make sense, the covariance matrix $\mathbf{C}(\Theta)$ must be invertible. This places serious restrictions on the size of the sampling set Λ, which must be coarse enough to allow invertibility.

Remark 3 (Cramér-Rao bound): The observation being complex Gaussian, and the maximum likelihood estimator being asymptotically unbiased and consistent, the Cramér-Rao lower bound provides relevant informations regarding the achievable precision of the estimator. Since we are in the complex normal situation, the Slepian-Bangs formula [13] provides an explicit expression for the bound: for any component θ of the multivariate parameter Θ,

$$\mathsf{CRLB}(\theta) = \left[\mathsf{Trace}\left\{\left(\mathbf{C}(\Theta)^{-1}\frac{\partial \mathbf{C}(\Theta)}{\partial \theta}\right)^2\right\}\right]^{-1} . \qquad (9)$$

This quantity can be evaluated numerically from (6).

Remark 4 (Feasibility of the global scheme): When Λ is a large set, this procedure involves products of large matrices and optimization in large dimensional spaces, which make the whole approach computationally intensive. This is why a more local scheme has to be preferred in such cases.

b) Local scheme To make things simpler, the estimation of $\alpha(\tau)$ and/or $\gamma(\tau)$ can be performed independently for each value of τ. With the same notations as above, denote by $\mathbf{V}_y^n = \{\mathbf{V}_y[\lambda', na],\, \lambda' \in \Lambda'\}$. The above strategy simplifies by observing that for any n, the corresponding restricted covariance matrix $\mathbf{C}$ only depends on $\Theta_n = (\theta_{n,1}, \theta_{n,2}) := (\alpha'(na), \log_q(\gamma'(na)))$. More precisely, for any $\lambda_i' = (\nu_i, s_i)$, we have (in integral form)

$$(\mathbf{C}(\Theta_n))_{ij} = q^{(s_i+s_j)/2}\int_0^\infty \mathscr{S}_Z(q^{-\theta_{n,2}}u) \overline{\hat{\psi}}[q^{s_i}(u+\theta_{n,1}-\nu_i)]\,\hat{\psi}[q^{s_j}(u+\theta_{n,1}-\nu_j)]\,du$$

The estimation scheme is as before:

1) Suppose an estimate for the power spectrum $\mathscr{S}_Z$ is available; then for each value of the time index, the parameters Θ_n can be estimated by maximizing the approximate log-likelihood

$$\mathscr{L}(\Theta_n) = -\ln|\det \mathbf{C}(\Theta_n)| - \mathbf{C}(\Theta_n)^{-1}\mathbf{V}_y^n \cdot \mathbf{V}_y^n \ . \qquad (10)$$

2) Suppose estimates for $\alpha'(na), \gamma'(na)$ are available for all $n = n_1, \dots n_2$. Estimates for α and γ can be deduced, from which a "stationarized" signal z_e can be computed, as in (8). Spectral estimation on z_e yields an estimate for the power spectrum $\mathscr{S}_Z$.

The estimation procedure is based upon alternate estimations of the parameters Θ and power spectrum $\mathscr{S}_Z$ (and covariance matrix) as above, starting from an initial guess for the spectrum. The initialization is set by computing the Welch estimator of the spectrum on the original signal y.

The stopping criterion is defined in terms of the relative update: iterations stop when

$$\frac{\left\|\Theta^{(k)} - \Theta^{(k-1)}\right\|_2^2}{\left\|\Theta^{(k-1)}\right\|_2^2} < T \qquad (11)$$

A similar approach was developed in [8], where the likelihood maximization was performed by exhaustive grid search in a finite parameter space, which results in undesirable quantization effects. Here the parameters Θ_n are treated as continuous parameters, which allows the evaluation of the Cramér-Rao bound (9) for assessing the precision of the estimator.

The generic structure of the algorithm can be summarized as in Algorithm 1. In our current implementation, the maximization of $\mathscr{L}(\Theta_n)$ is done via simple gradient ascent, and the interpolation is based upon cubic splines.

Remark 5 (Sampling the transform space): A key ingredient of this algorithm is the invertibility of the covariance matrix. In [7] and [8], simple sufficient conditions involving the analyzing waveform ψ and the frequency (resp scale) sampling step were given in the case of frequency modulation (resp. time warping). The situation turns out to be far more complex in the mixed case and no simple closed form condition seem to be available. This point is currently under study.

Algorithm 1 Joint covariance and modulation estimation

Initialization: Compute an estimate $\hat{\mathscr{S}}_Y$ of the power spectrum of Y as an initial guess $\hat{\mathscr{S}}_Z^{(0)}$ for $\mathscr{S}_Z$.
Compute the transform $\mathcal{V}_y$ of y.
while criterion (11) is false and $k \leq k_{max}$ **do**
- For each n, estimate $\Theta_n^{(k+1)}$ by maximizing the approximate log-likelihood $\mathscr{L}(\Theta_n)$ in (10)
- Compute $\hat{\alpha}^{(k)}$ and $\hat{\gamma}^{(k)}$ by interpolation from $\Theta^{(k)}$.
- Construct a "stationarized" signal $z_e^{(k)}$ from y using $\hat{\alpha}^{(k)}$ and $\hat{\gamma}^{(k)}$, and estimate the corresponding power spectrum $\hat{\mathscr{S}}_Z^{(k+1)}$.
- $k := k + 1$

end while

B. Numerical examples

We show here examples of joint estimation of transformation and power spectrum. Generic choices for the analyzing waveform are (analytic) derivative of gaussian ψ_k (which has k vanishing moments), and the sharp wavelet $\psi_\sharp$ (with infinitely many vanishing moments), defined in the positive Fourier domain by

$$\hat{\psi}_k(\nu) = \nu^k e^{-k\nu^2/2\nu_0^2}\,,\ \hat{\psi}_\sharp(\nu) = \epsilon^{\frac{\delta(\nu,\nu_0)}{\delta(\nu_1,\nu_0)}}\,,\ \nu > 0 \qquad (12)$$

and which vanish on the negative Fourier half axis. Here ν_0 is the mode of $\hat{\psi}$. In the expression of $\hat{\psi}_\sharp$, ν_1 is chosen so that $\hat{\psi}_\sharp(\nu_1) = \epsilon$ (a prescribed numerical tolerance), and the divergence δ is defined by $\delta(a,b) = \frac{1}{2}\left(\frac{a}{b} + \frac{b}{a}\right) - 1$.

1) Time warping estimation: Consider first the simple case considered in [8], i.e. the case $\alpha = 0$. Since no modulation is present in the deformation model, it is enough to limit to the wavelet transform (i.e. $\nu_i = 0$ for all i), and limit the estimation of parameters $\theta_{n,2} = \log_q(\gamma'(na))$. In the examples presented below, the signals are uniformly sampled with a sampling rate denoted by F_s. We used the wavelet $\psi_\sharp$ given in (12), where $\nu_1 = 2\nu_0 = F_s/2$ and $\epsilon = 10^{-6}$.

As a first example, we consider a synthetic signal, generated using a band pass spectrum and a warping function γ whose derivative is a sinusoidal function of time. The additive Gaussian white noise is generated so that the signal-to-noise ratio is equal to 8 dB. The results of the estimation are also displayed in figure 1, superimposed with the ground truth. The algorithm was able to estimate the power spectrum $\mathscr{S}_Z$ (right plot), as well as the time warping derivative (left plot). The fluctuations of the latter estimate $\hat{\gamma}'(t)$ are consistent with the Cramér-Rao bound $\mathsf{CRLB}(\hat{\gamma}(t))$, evaluated at the estimated parameter value (superimposed on the left plot).

The second example is a real signal, namely an audio recording of a Formula one racing car engine during an acceleration stage. Given the underlying physics, time warping seems to be a sensible deformation model. We display in figure 2 (top) the wavelet transforms of the original signal and of the estimated underlying stationary signal (i.e. after time unwarping). It clearly appears that the speed non-stationarity has been nicely corrected. A closer look at the right image

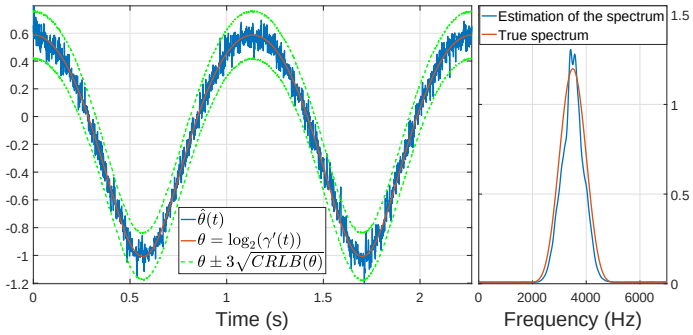

Figure 1: Joint time warping/spectrum estimation on a synthetic signal. Left: time warping function estimate (full, blue), ground truth (full, red) and Cramér-Rao bound (dotted, green); Right: spectrum of the underlying stationary signal (right).

also reveals a non-stationary behavior in the amplitude, that has not been corrected (since it was not present in the model). Accounting for such amplitude non-stationarities will require a more sophisticated model, for which one may try to use time-frequency or time-scale multipliers (as in [14]) for the estimation. The spectrum estimation (figure 2, bottom) of the underlying stationary signal shows the harmonic nature of the sound produced by the engine. The spectrum of the original is provided with sound files on the companion web site[2].

2) Time warping and modulation estimations: We illustrate here a simple situation in which time-warping and frequency modulations can be combined, and analyzed in a suitable representation space. We consider a deformation model in which a relation between frequency modulation and time warping is introduced, of the form

$$\alpha(t) = \eta(t - \gamma(t))$$

where $\eta > 0$ is assumed to be known. Thus, considering the extended transform $\mathcal{V}$, limited to $\nu = \eta$ yields a covariance

[2] https://www.i2m.univ-amu.fr/~torresan/papers/SAMPTA17/MT17.html

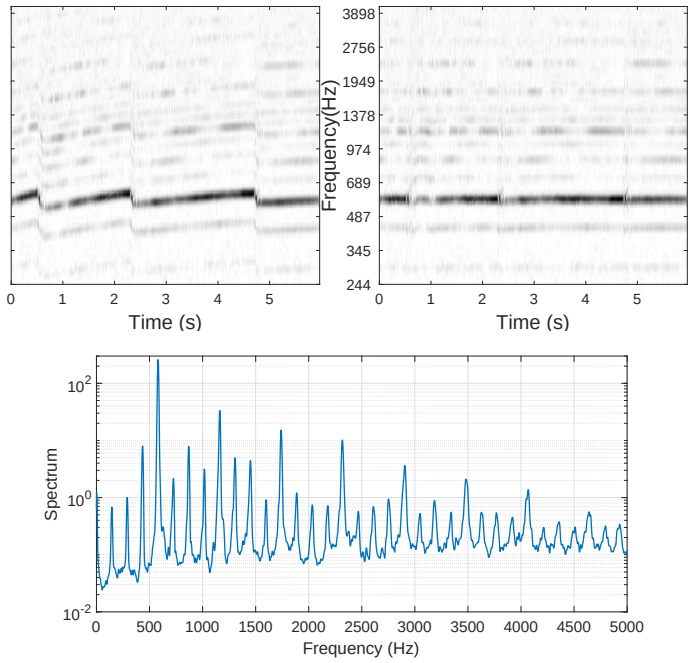

Figure 2: Joint time warping/spectrum estimation on an accelerating car engine: wavelet transforms of the original signal (top left) and the estimated underlying stationary signal (top right), and its spectrum (bottom).

matrix of the form

$$(\mathbf{C}(\theta_{n,2}))_{ij} = q^{\theta_{n,2}+\frac{s_i+s_j}{2}} \times \int_0^\infty \mathscr{S}_{\widetilde{Z}}(u)\overline{\hat{\psi}}\left(q^{s_i+\theta_{n,2}}u\right)\hat{\psi}\left(q^{s_j+\theta_{n,2}}u\right)du$$

where $\widetilde{Z} = M_{-\eta}Z$ is a stationary random process. This adapted transform allows the estimation of the deformation parameters $\theta_{n,2}$ as a translation on the scale axis, similar to the wavelet estimation of time warping. This clearly appears in figure 3, which displays the ordinary wavelet transform and the transform adapted to this model. The synthetic signal is constructed using the same warping function γ and spectrum $\mathscr{S}_Z$ as the synthetic example of the previous section. Here, the estimation of the deformation (not shown here due to the lack of space) can be done along the same lines as before using the adapted transform.

This example is nothing more than a toy example, aiming to illustrate the adaptation of the transform to the deformation. Applications to more complex examples are in progress.

IV. Conclusions

We have discussed in this paper extensions of the methods and algorithms described earlier in [7], [9], [8] for the joint estimation of deformation operator and power spectrum for deformed stationary signals. The main improvements described in this paper concern the following two points

1) formulation of the problem as a continuous parameter estimation problem, which yields both more precise estimates and a better control through a Cramér-Rao bound, and

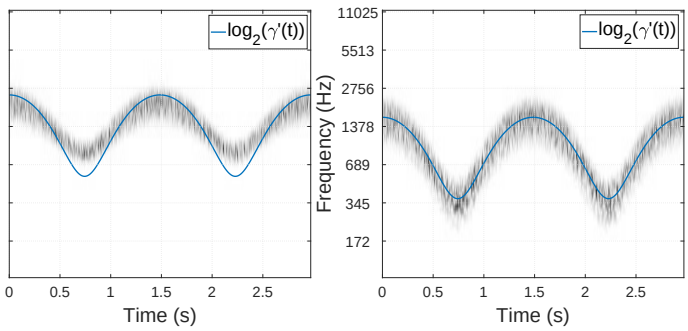

Figure 3: Wavelet transform (left) and adapted transform (right) superimposed with the function γ' to be estimated.

2) extension to more general and complex deformation models, by combining frequency modulation and time warping.

While the standard deformation models (frequency modulation and time warping) can already address interesting real world applications (variable speed car engine sounds are prototypical examples), they are not sufficient to describe many other situations of practical interest, such as the sound deformation models discussed in e.g. [15]. Here, we simply illustrated non-standard deformations on a simple toy model, that shows the benefit of adapting the representation to the deformation.

The study of combined modulations and warpings is currently in progress, as well as the integration of amplitude modulations in the model. The latter turns out to be more difficult than expected, the corresponding Cramér-Rao bound

being quite bad. The problem of sampling the extended transform mentioned in Remark 5 is also of great interest.

ACKNOWLEDGEMENTS

B. Torrésani wishes to thank CNRS and Centre de Recherches Mathématiques de Montréal, where this work was completed. A. Meynard wishes to thank Centre de Recherches Mathématiques de Montréal for hospitality, and GDR ISIS for financial support.

REFERENCES

[1] M. Priestley, *Spectral analysis and time series*, ser. Probability and mathematical statistics. Academic Press, 1982, no. v. 1-2.

[2] R. Dahlhaus, "A likelihood approximation for locally stationary processes," *Ann. Statist.*, vol. 28, no. 6, pp. 1762–1794, 12 2000. http://dx.doi.org/10.1214/aos/1015957480

[3] S. Mallat, G. Papanicolaou, and Z. Zhang, "Adaptive covariance estimation of locally stationary processes," *Ann. Statist.*, vol. 26, no. 1, pp. 1–47, 02 1998. http://dx.doi.org/10.1214/aos/1030563977

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

[5] 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.

[6] ——, "Estimating deformations of stationary processes," *Ann. Statist.*, vol. 31, no. 6, pp. 1772–1821, 12 2003. http://dx.doi.org/10.1214/aos/1074290327

[7] 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. http://hal.archives-ouvertes.fr/hal-00822186

[8] H. Omer and B. Torrésani, "Time-frequency and time-scale analysis of deformed stationary processes, with application to non-stationary sound modeling," *Applied and Computational Harmonic Analysis*, 2016. https://hal.archives-ouvertes.fr/hal-01094835

[9] 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.

[10] I. Gel'fand and N. Y. Vilenkin, *Generalized Functions, Volume 4: Applications of Harmonic Analysis*. American Mathematical Society, 1964.

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

[12] B. Torrésani, "Wavelets associated with representations of the affine Weyl–Heisenberg group," *Journal of Mathematical Physics*, vol. 32, no. 5, pp. 1273–1279, 1991. http://dx.doi.org/10.1063/1.529325

[13] P. Stoica and R. Moses, *Introduction to Spectral Analysis*. Prentice Hall, 1997.

[14] 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.

[15] S. Conan, E. Thoret, M. Aramaki, O. Derrien, C. Gondre, S. Ystad, and R. Kronland-Martinet, "An intuitive synthesizer of continuous-interaction sounds: Rubbing, scratching, and rolling," *Computer Music Journal*, vol. 38, no. 4, pp. 24–37, Dec 2014.