

Spectral analysis for nonstationary audio

Part II: Supplementary material

Adrien Meynard and Bruno Torr sani

Aix Marseille Univ, CNRS, Centrale Marseille, I2M, UMR 7373, Marseille, France.

I. INTRODUCTION

This note provides complementary information on [1]. In that manuscript, nonstationary signals of interest are obtained as linear deformations of stationary random signals, namely amplitude modulations and time warpings:

$$\mathcal{A}_a : \quad \mathcal{A}_a x(t) = a(t)x(t) , \quad \mathcal{D}_\gamma : \quad \mathcal{D}_\gamma x(t) = \sqrt{\gamma'(t)}x(\gamma(t)) , \quad (\text{I.1})$$

where $a \in C^1$ and $\gamma \in C^2$ are real valued functions, satisfying conditions

$$0 < c_a \leq a(t) \leq C_a < \infty , \quad 0 < c_\gamma \leq \gamma'(t) \leq C_\gamma < \infty , \quad \forall t , \quad (\text{I.2})$$

for some constants $c_a \leq C_a \in \mathbb{R}_+^*$ and $c_\gamma \leq C_\gamma \in \mathbb{R}_+^*$.

Given a (unique) realization of a random signal of the form

$$Y = \mathcal{A}_a \mathcal{D}_\gamma X \quad (\text{I.3})$$

where X is a stationary zero-mean real random process with (unknown) power spectrum \mathcal{S}_X , the goal is to estimate the deformation functions a and γ from this realization of Y , exploiting the assumed stationarity of X .

II. WAVELET TRANSFORM

Our analysis relies heavily on the continuous wavelet transform (and discretized versions). The wavelet transform of a signal is defined as:

$$\mathcal{W}_X(s, \tau) \triangleq \langle X, \psi_{s\tau} \rangle , \quad \text{with } \psi_{s\tau} = q^{-s/2} \psi(q^{-s}(t - \tau)) , \quad (\text{II.1})$$

$\psi \in L^2(\mathbb{R})$ being a smooth wavelet centered at the origin, and $q > 1$ a reference scale constant.

Part of this work was done when the two authors were at Centre de Recherches Math matiques, UMI 3457, CNRS and Universit  de Montr al, Canada.

Classical choices of wavelets in $H^2(\mathbb{R}) = \{\psi \in L^2(\mathbb{R}) : \text{supp}(\hat{\psi}) \subset \mathbb{R}^+\}$ are (analytic) derivative of Gaussian ψ_k (which has k vanishing moments), and the sharp wavelet ψ_{\sharp} (with infinitely many vanishing moments). These can be defined in the positive Fourier domain by

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

and 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 at some cutoff frequency ν_1), and the divergence δ is defined by $\delta(a, b) = \frac{1}{2} \left(\frac{a}{b} + \frac{b}{a} \right) - 1$.

In numerical applications, we choose $\nu_1 = F_s/2$ and ϵ small in order to prevent aliasing in the wavelet transform when $s = 0$. Nonetheless, it can be more meaningful to express $\hat{\psi}_{\sharp}$ in function of its quality factor Q . The quality factor is the ratio between the central frequency ν_0 of $\hat{\psi}_{\sharp}$ and its bandwidth $\Delta\nu = \nu_+ - \nu_-$ where $\hat{\psi}_{\sharp}(\nu_+) = \hat{\psi}_{\sharp}(\nu_-) = \frac{1}{\sqrt{2}}$. Then, one can show that

$$Q = \frac{1}{\sqrt{C(C+4)}} \quad \text{where} \quad C = -\frac{\ln 2}{\ln \epsilon} \delta(\nu_1, \nu_0) > 0.$$

Finally, we give the expression of $\hat{\psi}_{\sharp}$ in function of ν_0 and Q :

$$\hat{\psi}_{\sharp}(\nu) = 2^{-Q(2Q + \sqrt{1+4Q^2})\delta(\nu, \nu_0)}.$$

To show the influence of the parameter ϵ on the wavelet transform, we display on the left side of Fig. 1 the Fourier transform of the sharp wavelet obtained for two different values of ϵ . The smaller ϵ is, the sharper the Fourier transform of ψ_{\sharp} is. Furthermore, the right side of Fig. 1 displays the time representations of these two wavelets. It shows that when ϵ decreases, the time support of ψ_{\sharp} becomes larger.

Then, we compute the corresponding wavelet transforms. the scalograms are displayed on Fig. 2. The signal used to evaluate the effect of ϵ is a wolf howl recording. As a consequence of the previous result, a smaller value of ϵ improves the localization of the wavelet transform along the scale axis, while the localization along the time axis is degraded.

III. APPROXIMATION RESULT: PROOF OF THEOREM 1

We provide in this section more details on the approximation of the wavelet transform of time warped stationary signals discussed in Section III of [1].

Assuming smoothness for a and γ a local analysis is performed using wavelet transform $\mathcal{W} : Y \rightarrow \mathcal{W}_Y$ defined by $\mathcal{W}_Y(s, \tau) = \langle Y, \psi_{s\tau} \rangle$. We then prove that

$$\mathcal{W}_Y(s, \tau) \approx \widetilde{\mathcal{W}}_Y(s, \tau) \triangleq a(\tau) \mathcal{W}_X \left(s + \log_q(\gamma'(\tau)), \gamma(\tau) \right). \quad (\text{III.1})$$

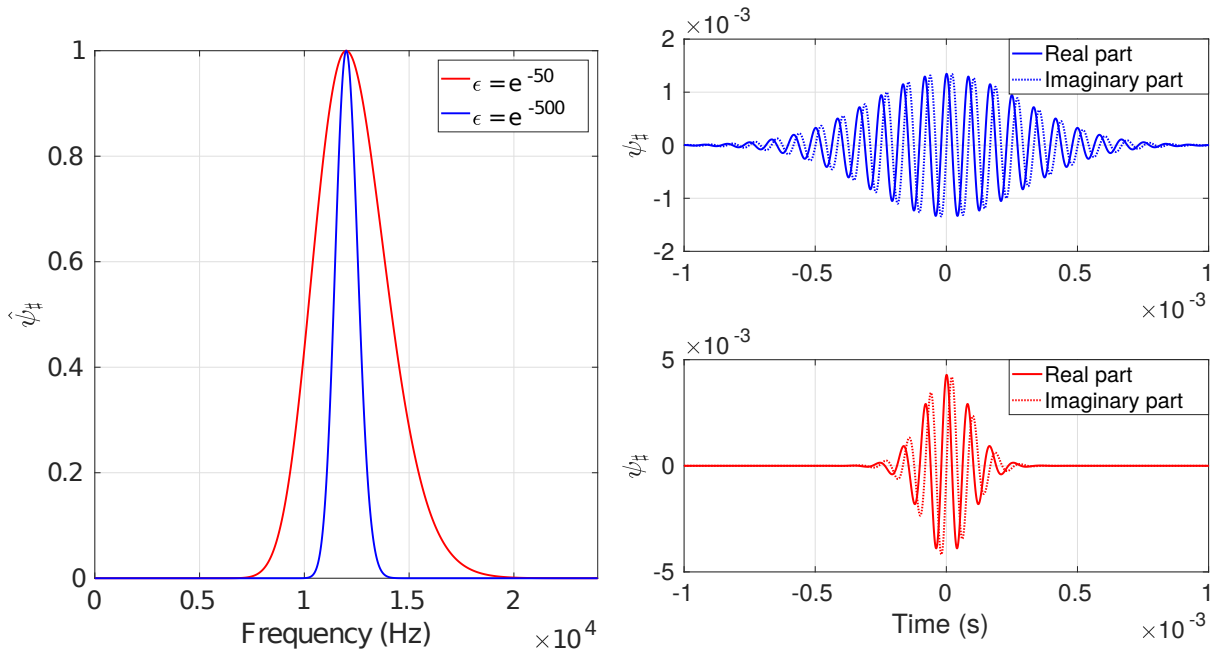


Fig. 1. Left: Fourier transform of the sharp wavelet ψ_\sharp for two different values of ϵ . Right: Corresponding wavelets in the time domain: $\epsilon = e^{-500}$ i.e. $Q = 13.43$ (top) and $\epsilon = e^{-50}$ i.e. $Q = 4.24$ (bottom). Here, $\nu_0 = F_s/4 = 11025$ Hz and $\nu_1 = F_s/2 = 22050$ Hz.

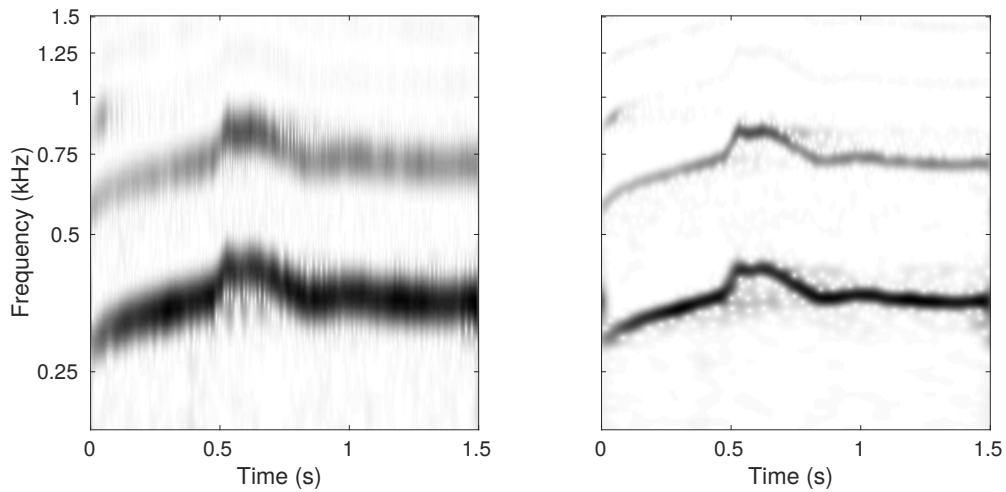


Fig. 2. Scalograms of a wolf howl recording using the sharp wavelet with parameter $\epsilon = e^{-50}$ (left), and $\epsilon = e^{-500}$ (right).

Theorem 1 in [1], which is recalled below, provides a quantitative assessment of the quality of the approximation.

Theorem 1: Let X be a second order zero-mean stationary random process, let Y be the nonstationary process defined in (I.3). Let ψ be a smooth test function, localized in such a way that $|\psi(t)| \leq 1/(1+|t|^\beta)$ for some $\beta > 2$. Let \mathcal{W}_Y denote the wavelet transform of Y , $\widetilde{\mathcal{W}}_Y$ its approximation given in (III.1), and let $\varepsilon = \mathcal{W}_Y - \widetilde{\mathcal{W}}_Y$ denote the approximation error. Assume ψ and \mathcal{S}_X are such that

$$I_X^{(\rho)} \triangleq \sqrt{\int_0^\infty \xi^{2\rho} \mathcal{S}_X(\xi) d\xi} < \infty, \text{ where } \rho = \frac{\beta - 1}{\beta + 2}.$$

Then the approximation error ε is a second order, two-dimensional complex random field, and

$$\mathbb{E} \left\{ |\varepsilon(s, \tau)|^2 \right\} \leq C_a^2 q^{3s} \left(K_1 \|\gamma''\|_\infty + K_2 q^{\mu s} \|\gamma''\|_\infty^\rho + K_3 \|a'\|_\infty \right)^2$$

where

$$K_1 = \frac{\beta \sigma_X}{2(\beta - 2)\sqrt{c_\gamma}}, \quad K_2 = I_X^{(\rho)} \left(\frac{\pi}{2} \right)^\rho \sqrt{C_\gamma} \frac{4}{3\rho}, \quad K_3 = \frac{\sqrt{C_\gamma} \beta \sigma_X}{(\beta - 2)c_a}, \quad \mu = \frac{\beta - 4}{\beta + 2},$$

σ_X^2 being the variance of X .

Proof of the theorem: To simplify notations, let \mathcal{B}_γ denote the operator $\mathcal{D}_\gamma / \sqrt{\gamma'}$. We split the approximation error as follows

$$\begin{aligned} \varepsilon(s, \tau) &= \langle \mathcal{A}_a \mathcal{D}_\gamma X, \psi_{s\tau} \rangle - \langle \widetilde{\mathcal{A}}_a^\tau \widetilde{\mathcal{D}}_\gamma^\tau X, \psi_{s\tau} \rangle \\ &= \widetilde{\mathcal{A}}_a^\tau \left(\langle \widetilde{\mathcal{A}}_a^{\tau^{-1}} \mathcal{A}_a \mathcal{D}_\gamma X, \psi_{s\tau} \rangle - \langle \widetilde{\mathcal{D}}_\gamma^\tau X, \psi_{s\tau} \rangle \right) \\ &= a(\tau) \left(\langle \sqrt{\gamma'} \mathcal{B}_\gamma X, \psi_{s\tau} \rangle - \langle \sqrt{\gamma'(\tau)} \widetilde{\mathcal{B}}_\gamma^\tau X, \psi_{s\tau} \rangle + \langle (\widetilde{\mathcal{A}}_a^{\tau^{-1}} \mathcal{A}_a - 1) \mathcal{D}_\gamma X, \psi_{s\tau} \rangle \right) \\ &= a(\tau) \left(\langle (\sqrt{\gamma'} - \sqrt{\gamma'(\tau)}) \mathcal{B}_\gamma X, \psi_{s\tau} \rangle + \sqrt{\gamma'(\tau)} \langle (\mathcal{B}_\gamma - \widetilde{\mathcal{B}}_\gamma^\tau) X, \psi_{s\tau} \rangle \right. \\ &\quad \left. + \langle (\widetilde{\mathcal{A}}_a^{\tau^{-1}} \mathcal{A}_a - 1) \mathcal{D}_\gamma X, \psi_{s\tau} \rangle \right) \\ &= a(\tau) \left(\varepsilon^{(1)}(s, \tau) + \sqrt{\gamma'(\tau)} \varepsilon^{(2)}(s, \tau) + \varepsilon^{(3)}(s, \tau) \right), \end{aligned}$$

where

$$\begin{aligned} \varepsilon^{(1)}(s, \tau) &\triangleq \left\langle (\sqrt{\gamma'} - \sqrt{\gamma'(\tau)}) \mathcal{B}_\gamma X, \psi_{s\tau} \right\rangle = \left\langle X, \mathcal{B}_{\gamma^{-1}} \left(\sqrt{\gamma'} - \sqrt{\gamma'(\tau)} \right) \psi_{s\tau} \right\rangle, \\ \varepsilon^{(2)}(s, \tau) &\triangleq \left\langle (\mathcal{B}_\gamma - \widetilde{\mathcal{B}}_\gamma^\tau) X, \psi_{s\tau} \right\rangle = \left\langle X, \left(\mathcal{B}_{\gamma^{-1}} - \widetilde{\mathcal{B}}_\gamma^{\tau^{-1}} \right) \psi_{s\tau} \right\rangle, \\ \varepsilon^{(3)}(s, \tau) &\triangleq \left\langle (\widetilde{\mathcal{A}}_a \widetilde{\mathcal{A}}_a^{\tau^{-1}} - 1) \mathcal{D}_\gamma X, \psi_{s\tau} \right\rangle = \left\langle X, \mathcal{D}_{\gamma^{-1}} \left(\widetilde{\mathcal{A}}_a \widetilde{\mathcal{A}}_a^{\tau^{-1}} - 1 \right) \psi_{s\tau} \right\rangle. \end{aligned}$$

In the following, the right-hand term in the scalar product defining $\varepsilon^{(k)}(s, \tau)$ will be denoted by $f_{s\tau}^{(k)}$ so that $\varepsilon^{(k)}(s, \tau) = \langle X, f_{s\tau}^{(k)} \rangle$ (with $k \in \{1, 2, 3\}$).

Besides, the triangle inequality gives:

$$\mathbb{E} \left\{ |\varepsilon(s, \tau)|^2 \right\} \leq C_a^2 \left(\sqrt{\mathbb{E} \left\{ |\varepsilon^{(1)}(s, \tau)|^2 \right\}} + \sqrt{C_\gamma \mathbb{E} \left\{ |\varepsilon^{(2)}(s, \tau)|^2 \right\}} + \sqrt{\mathbb{E} \left\{ |\varepsilon^{(3)}(s, \tau)|^2 \right\}} \right)^2. \quad (\text{III.2})$$

Let us now determine an upper bound for each error term. To this end, by definition of the spectrum, they are written as follows:

$$\mathbb{E} \left\{ \left| \varepsilon^{(k)}(s, \tau) \right|^2 \right\} = \mathbb{E} \left\{ \left| \langle X, f_{s\tau}^{(k)} \rangle \right|^2 \right\} = \left\langle \mathcal{S}_X, \left| f_{s\tau}^{(k)} \right|^2 \right\rangle = \int_0^\infty \mathcal{S}_X(\xi) \left| \hat{f}_{s\tau}^{(k)}(\xi) \right|^2 d\xi,$$

with $k \in \{1, 2, 3\}$.

Concerning the first error term, a Taylor expansion of $\sqrt{\gamma'}$ around τ gives

$$\begin{aligned} \left| \hat{f}_{s\tau}^{(1)}(\xi) \right| &= \left| \int_{\mathbb{R}} \left(\sqrt{\gamma'(t)} - \sqrt{\gamma'(\tau)} \right) \psi_{s\tau}(t) e^{-2i\pi\gamma(t)\xi} dt \right| \\ &\leq \int_{\mathbb{R}} \left\| \frac{\gamma''}{2\sqrt{\gamma'}} \right\|_\infty |t - \tau| q^{-s/2} \left| \psi \left(\frac{t - \tau}{q^s} \right) \right| dt \\ &\leq q^{\frac{3s}{2}} \frac{\|\gamma''\|_\infty}{2\sqrt{c_\gamma}} I_\psi, \end{aligned}$$

where $I_\psi = \int_{\mathbb{R}} |t\psi(t)| dt$. Furthermore, the localization assumption on ψ allows us to write

$$I_\psi \leq 2 \int_0^\infty \frac{t}{1+t^\beta} dt \leq 2 \left(\int_0^1 t dt + \int_1^\infty \frac{1}{t^{\beta-1}} dt \right) = \frac{\beta}{\beta-2}.$$

Finally, we can control the first error term as follows:

$$\mathbb{E} \left\{ \left| \varepsilon^{(1)}(s, \tau) \right|^2 \right\} \leq \left(q^{3s/2} \frac{\|\gamma''\|_\infty}{\sqrt{c_\gamma}} \frac{\beta \sigma_X}{2(\beta-2)} \right)^2.$$

Concerning the second error term, the successive applications of the tangent half-angle trigonometric identity and the Taylor-Lagrange formula give

$$\begin{aligned} \left| \hat{f}_{s\tau}^{(2)}(\xi) \right| &= \left| \int_{\mathbb{R}} \left(e^{-2i\pi\gamma(t)\xi} - e^{-2i\pi(\gamma(\tau)+(t-\tau)\gamma'(\tau))\xi} \right) \psi_{s\tau}(t) dt \right| \\ &\leq \int_{\mathbb{R}} \left| 1 - e^{-2i\pi(\gamma(\tau)+(t-\tau)\gamma'(\tau)-\gamma(t))\xi} \right| |\psi_{s\tau}(t)| dt \\ &\leq \int_{\mathbb{R}} 2 \left| \sin \left(\pi (\gamma(\tau) + (t - \tau)\gamma'(\tau) - \gamma(t)) \xi \right) \right| |\psi_{s\tau}(t)| dt \\ &\leq \int_{\mathbb{R}} 2 \left| \sin \left(\frac{\pi}{2} \xi (t - \tau)^2 \gamma''(t_*) \right) \right| |\psi_{s\tau}(t)| dt, \end{aligned}$$

for some t_* between t and τ . Besides, we have $|\sin(u)| \leq |u|$ and $|\sin(u)| \leq 1$ so that:

$$\left| \hat{f}_{s\tau}^{(2)}(\xi) \right| \leq 2q^{s/2} \left(\int_J \frac{\pi}{2} \xi \|\gamma''\|_\infty q^{2s} t^2 |\psi(t)| dt + \int_{\mathbb{R} \setminus J} |\psi(t)| dt \right).$$

where $J = [-v, v]$. Finding the zeros of the derivative of this expression with respect to v , one can show that the value of v minimizing the right-hand side of the latter equation is $v = \left(\frac{\pi}{2}\xi\|\gamma''\|_\infty q^{2s}\right)^{-1/(\beta+2)}$. Therefore, we have:

$$\left|\hat{f}_{s\tau}^{(2)}(\xi)\right| \leq q^{\frac{5\beta-2}{2(\beta+2)}} \frac{4(\beta+2)}{3(\beta-1)} \left(\frac{\pi}{2}\xi\|\gamma''\|_\infty\right)^{\frac{\beta+2}{\beta-1}}.$$

Finally, we can control the second error term as follows:

$$\mathbb{E} \left\{ \left| \varepsilon^{(2)}(s, \tau) \right|^2 \right\} \leq \left(q^{\frac{5\beta-2}{2(\beta+2)}} \frac{4(\beta+2)}{3(\beta-1)} \left(\frac{\pi}{2}\|\gamma''\|_\infty\right)^{\frac{\beta+2}{\beta-1}} I_X^{(\rho)} \right)^2.$$

Concerning the third error term, the application of the Taylor-Lagrange inequality gives:

$$\begin{aligned} \left|\hat{f}_{s\tau}^{(3)}(\xi)\right| &= \left| \int_{\mathbb{R}} \sqrt{\gamma'(t)} \left(\frac{a(t)}{a(\tau)} - 1 \right) \psi_{s\tau}(t) e^{-2i\pi\gamma(t)\xi} dt \right| \\ &\leq \sqrt{C_\gamma} \int_{\mathbb{R}} \frac{\|a'\|_\infty}{c_a} |t - \tau| |\psi_{s\tau}(t)| dt = q^{3s/2} \sqrt{C_\gamma} \frac{\|a'\|_\infty}{c_a} I_\psi, \end{aligned}$$

Finally, we can control the third error term as follows:

$$\mathbb{E} \left\{ \left| \varepsilon^{(3)}(s, \tau) \right|^2 \right\} \leq \left(q^{3s/2} \frac{\sqrt{C_\gamma}}{c_a} \frac{\beta\sigma_X}{\beta-2} \|a'\|_\infty \right)^2.$$

To conclude the proof, the three errors terms in equation (III.2) are replaced by their upper bounds to obtain the approximation error given in the theorem. \blacksquare

IV. ESTIMATION ALGORITHM: PROOF OF PROPOSITION 1

The goal of JEFAS is to estimate both deformation functions γ and a from the approximated wavelet transform $\widetilde{\mathcal{W}}_y$ of a realization y of Y , assuming the latter is a reliable approximation of the true wavelet transform. Under Gaussianity assumption, the algorithm relies on an iterative alternate optimization of (an approximate) likelihood of parameters with respect to deformation function and power spectrum respectively.

The wavelet transform is computed on a regular grid $\Lambda = \mathbf{s} \times \boldsymbol{\tau} \in \mathbb{R}^{M_s} \times \mathbb{R}^{N_\tau}$. The goal is to estimate the vector of parameters $\Theta = (\theta_1, \theta_2, \theta_3) \triangleq (a(\boldsymbol{\tau})^2, \log_q(\gamma'(\boldsymbol{\tau})), \gamma(\boldsymbol{\tau}))$. Let $\mathbf{W}_y = \widetilde{\mathcal{W}}_y(\Lambda)$ denote the corresponding discretized transform. The algorithm ignores time correlations and performs warping and amplitude modulation parameters estimation for discrete times $\tau_n \in \boldsymbol{\tau}$, and power spectrum estimations on \mathbf{s} . Proposition 1 provides an upper bound for the bias of the power spectrum estimate.

Proposition 1: Let $\psi \in H^2(\mathbb{R})$ be an analytic wavelet such that $\hat{\psi}$ is bounded and $|\hat{\psi}(u)| = \mathcal{O}_{u \rightarrow \infty}(u^{-\eta})$ with $\eta > 2$. Let φ_1 and φ_2 be bounded functions defined on \mathbb{R}^+ by $\varphi_1(u) = u |\hat{\psi}(u)|^2$ and $\varphi_2(u) = u^2 |\hat{\psi}(u)|$. Assume \mathcal{S}_X is such that

$$J_X = \int_0^\infty \xi^{-1} \mathcal{S}_X(\xi) d\xi < \infty.$$

Let $\mathcal{S}_X^{(k)}$ denote the estimation of the spectrum after k iterations of the algorithm. Let $b_{\mathcal{S}_X}^{(k)}$ denote the bias defined for all $m \in \llbracket 1, M_s \rrbracket$ by

$$b_{\mathcal{S}_X}^{(k)}(m) = \mathbb{E} \left\{ \mathcal{S}_X^{(k)}(q^{-s_m} \omega_0) \right\} - \mathcal{S}_{X,\psi}(q^{-s_m} \omega_0).$$

Assume there exists a constant $c_{\theta_1} > 0$ such that $\theta_{n,1}^{(k)} > c_{\theta_1}, \forall n, k$. Then

$$\left\| b_{\mathcal{S}_X}^{(k)} \right\|_\infty \leq \frac{J_X}{\|\psi\|_2^2} \left(K'_1 \left\| \theta_1 - \tilde{\theta}_1^{(k)} \right\|_\infty + K'_2 \left\| \tilde{\theta}_2^{(k)} - \theta_2 \right\|_\infty \right), \quad (\text{IV.1})$$

where

$$K'_1 \triangleq \frac{\|\varphi_1\|_\infty}{c_{\theta_1}} < \infty, \quad K'_2 \triangleq \ln(q) (\|\varphi_1\|_\infty + 2\|\hat{\psi}'\|_\infty \|\varphi_2\|_\infty) < \infty.$$

Proof of the proposition. Notations are as follows. Assume parameter estimates $\tilde{\theta}_1^{(k)}$ and $\tilde{\theta}_2^{(k)}$ are available. Let $\mathbf{w}_{x,s_m} \in \mathbb{C}^{N_\tau}$ be the vector of wavelet coefficients of the original signal x at scale s_m , and let $\tilde{\mathbf{w}}_{x,s_m}^{(k)} \in \mathbb{C}^{N_\tau}$ denote the estimation of \mathbf{w}_{x,s_m} after k iterations of the algorithm. Considering equation (18) in [1], we have $\tilde{\mathbf{w}}_{x,s_m}^{(k)} = \frac{1}{\sqrt{\tilde{\theta}_1^{(k)}}} \tilde{\mathcal{W}}_y \left(s_m - \tilde{\theta}_2^{(k)}, \boldsymbol{\tau} \right)$. Besides, starting from the definition of the spectrum given by equation (21) in [1] we obtain:

$$\mathbb{E} \left\{ \mathcal{S}_X^{(k)}(q^{-s_m} \omega_0) \right\} = \frac{1}{N_\tau \|\psi\|_2^2} \mathbb{E} \left\{ \left\| \tilde{\mathbf{w}}_{x,s_m}^{(k)} \right\|^2 \right\}.$$

To simplify notations, let us introduce some variables. We define $s_m^{(\xi)} = s_m + \log_q(\xi)$ and $h(x) = \varphi_1(q^x) = q^x |\hat{\psi}(q^x)|^2$ for $x \in \mathbb{R}$.

By means of the covariance expression given in equation (12) in [1] we can write

$$\begin{aligned} \mathbb{E} \left\{ \left\| \tilde{\mathbf{w}}_{x,s_m}^{(k)} \right\|^2 \right\} &= \sum_{n=1}^{N_\tau} \frac{1}{\tilde{\theta}_{n,1}^{(k)}} \mathbb{E} \left\{ \tilde{\mathcal{W}}_y \left(s_m - \tilde{\theta}_{n,2}^{(k)}, \tau_n \right) \overline{\tilde{\mathcal{W}}_y \left(s_m - \tilde{\theta}_{n,2}^{(k)}, \tau_n \right)} \right\} \\ &= \sum_{n=1}^{N_\tau} \frac{\theta_{n,1}}{\tilde{\theta}_{n,1}^{(k)}} \int_0^\infty \frac{\mathcal{S}_X(\xi)}{\xi} \xi q^{s_m - \tilde{\theta}_{n,2}^{(k)} + \theta_{n,2}} \left| \hat{\psi} \left(\xi q^{s_m - \tilde{\theta}_{n,2}^{(k)} + \theta_{n,2}} \right) \right|^2 d\xi \\ &= \sum_{n=1}^{N_\tau} \frac{\theta_{n,1}}{\tilde{\theta}_{n,1}^{(k)}} \int_0^\infty \frac{\mathcal{S}_X(\xi)}{\xi} h \left(s_m^{(\xi)} + \theta_{n,2} - \tilde{\theta}_{n,2}^{(k)} \right) d\xi \\ &= \int_0^\infty \frac{\mathcal{S}_X(\xi)}{\xi} \left(\sum_{n=1}^{N_\tau} \frac{\theta_{n,1}}{\tilde{\theta}_{n,1}^{(k)}} h \left(s_m^{(\xi)} + \theta_{n,2} - \tilde{\theta}_{n,2}^{(k)} \right) \right) d\xi. \end{aligned}$$

Let us now split the bias into two terms such that $b_{\mathcal{S}_X}^{(k)}(m) = g_1(m) + g_2(m)$, where g_1 and g_2 are defined as

$$g_1(m) = \frac{N_\tau^{-1}}{\|\psi\|_2^2} \int_0^\infty \frac{\mathcal{S}_X(\xi)}{\xi} \left(\sum_{n=1}^{N_\tau} \left(\frac{\theta_{n,1}}{\tilde{\theta}_{n,1}^{(k)}} - 1 \right) h \left(s_m^{(\xi)} + \theta_{n,2} - \tilde{\theta}_{n,2}^{(k)} \right) \right) d\xi,$$

$$g_2(m) = \frac{N_\tau^{-1}}{\|\psi\|_2^2} \int_0^\infty \frac{\mathcal{S}_X(\xi)}{\xi} \left(\sum_{n=1}^{N_\tau} \left(h \left(s_m^{(\xi)} + \theta_{n,2} - \tilde{\theta}_{n,2}^{(k)} \right) - h \left(s_m^{(\xi)} \right) \right) \right) d\xi.$$

Regarding the first term, we directly have

$$\begin{aligned} |g_1(m)| &\leq \frac{\|h\|_\infty}{\|\psi\|_2^2} \frac{1}{N_\tau} \sum_{n=1}^{N_\tau} \frac{|\theta_{n,1} - \tilde{\theta}_{n,1}^{(k)}|}{\tilde{\theta}_{n,1}^{(k)}} \int_0^\infty \frac{\mathcal{S}_X(\xi)}{\xi} d\xi \\ &\leq \frac{\|h\|_\infty}{\|\psi\|_2^2 c_{\theta_1}} \|\theta_1 - \tilde{\theta}_1^{(k)}\|_\infty J_X. \end{aligned}$$

Besides, we have $\|h\|_\infty = \|\varphi_1\|_\infty$ and the smoothness and decay assumptions on $\hat{\psi}$ allow us to write $\varphi_1(u) = \mathcal{O}_{u \rightarrow \infty}(u^{1-2\eta}) \xrightarrow{u \rightarrow \infty} 0$. Then φ_1 is bounded and $K'_1 < \infty$. This yields

$$|g_1(m)| \leq \frac{J_X K'_1}{\|\psi\|_2^2} \|\theta_1 - \tilde{\theta}_1^{(k)}\|_\infty.$$

Regarding the second term, a Taylor expansion of h around $s_m^{(\xi)}$ gives

$$\begin{aligned} |g_2(m)| &\leq \frac{N_\tau^{-1}}{\|\psi\|_2^2} \int_0^\infty \frac{\mathcal{S}_X(\xi)}{\xi} \sum_{n=1}^{N_\tau} \left| h \left(s_m^{(\xi)} + \theta_{n,2} - \tilde{\theta}_{n,2}^{(k)} \right) - h \left(s_m^{(\xi)} \right) \right| d\xi \\ &\leq \frac{\|h'\|_\infty}{\|\psi\|_2^2} \frac{1}{N_\tau} \sum_{n=1}^{N_\tau} |\theta_{n,2} - \tilde{\theta}_{n,2}^{(k)}| \int_0^\infty \frac{\mathcal{S}_X(\xi)}{\xi} d\xi \\ &\leq \frac{\|h'\|_\infty}{\|\psi\|_2^2} \|\theta_2 - \tilde{\theta}_2^{(k)}\|_\infty J_X. \end{aligned}$$

Furthermore, $\forall x \in \mathbb{R}$

$$\begin{aligned} |h'(x)| &= |\ln(q) q^x \varphi_1'(q^x)| \leq \ln(q) \left(q^x |\hat{\psi}(q^x)|^2 + 2q^{2x} |\hat{\psi}(q^x) \hat{\psi}'(q^x)| \right) \\ &\leq \ln(q) (\|\varphi_1\|_\infty + 2\|\hat{\psi}'\|_\infty \|\varphi_2\|_\infty) = K'_2. \end{aligned}$$

Besides, the decay assumption on $\hat{\psi}$ gives $|\varphi_2(u)| = \mathcal{O}_{u \rightarrow \infty}(u^{2-\eta}) \xrightarrow{u \rightarrow \infty} 0$ because $\eta > 2$. Then φ_2 is bounded and $K'_2 < \infty$. This yields

$$|g_2(m)| \leq \frac{J_X K'_2}{\|\psi\|_2^2} \|\theta_2 - \tilde{\theta}_2^{(k)}\|_\infty.$$

The proof is concluded by summing up the upper bounds of $|g_1|$ and $|g_2|$ to obtain the upper bound of $b_{\mathcal{S}_X}^{(k)}$. Notice that this bound does not depend on m . \blacksquare

V. APPLICATION TO SPECTRAL ANALYSIS OF A WIDEBAND WIND SOUND

We illustrate here the ability of JEFAS to extract power spectrum information from a real, non-stationary audio signal. The power spectrum is there the power spectrum of a “stationarized” version of the signal, i.e. a version of the amplitude modulation and time warping estimated by JEFAS have been corrected.

The sound we analyze is a 12 seconds long audio sample of howling wind (the full audio recording is available at *Freesound*, *sample #144083*). Unlike the sounds analyzed in the paper [1], this sound does not display a clear harmonic structure. Indeed, as can be seen on the wavelet transform of the original sound displayed on the top-left of Fig. 3), two ridges are present but they are quite large, and the frequency content is really wideband, with a strong low frequency component. The spectrum of the underlying stationary sound estimated by JEFAS is displayed on the bottom-right of Fig. 3; this confirms the wideband nature of the signal, and the fact that the two frequency-localized components are indeed poorly frequency localized, and couldn't be well accounted for by sine waves. Besides, the estimated time-warping function displayed on the bottom-left gives a representation of the instantaneous wind speed.

VI. APPLICATION TO CROSS-SYNTHESIS

The ability to model non-stationary sounds using spectral and warping informations opens interesting possibilities for sound synthesis. Cross-synthesis, which uses the spectral information of a signal and warping information of a second one, is a simple example. Examples of such synthetic sounds are available on GitHub (together with source code that allows reproducing the numerical examples)

<https://github.com/AdMeynard/JEFAS>

and on the website

<https://meynard.perso.math.cnrs.fr/paperJEFAS/NonStationaryAudio.html>

REFERENCES

- [1] A. Meynard and B. Torr sani, Spectral analysis for nonstationary audio, submitted.

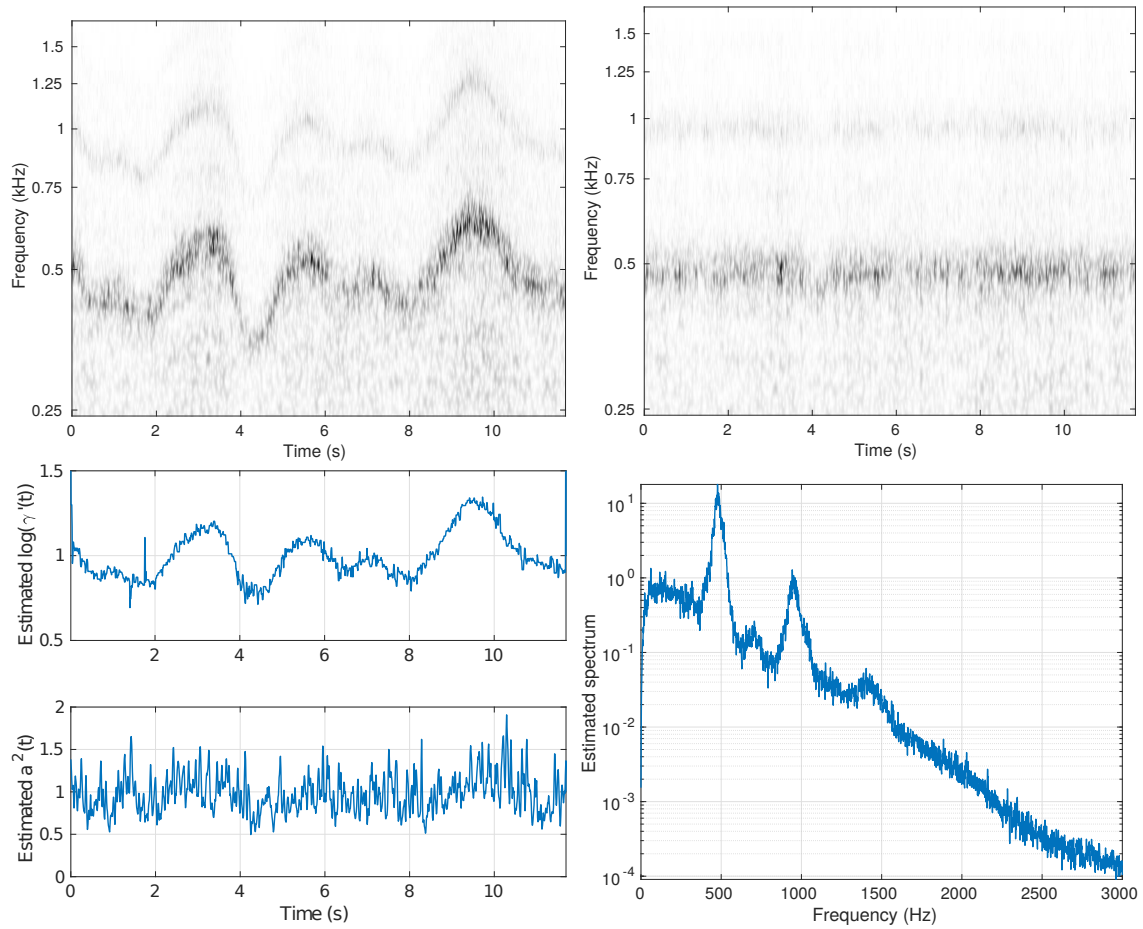


Fig. 3. Wind sound analysis. Top left: Scalogram of the original signal. Top right: Scalogram of the unwarped and unmodulated signal. Bottom left: Estimated time-warping and amplitude modulation functions. Bottom right: Estimated spectrum of the underlying stationary signal.