Time-scale synthesis for locally stationary signals

Adrien Meynard, Bruno Torrésani

TIME-SCALE SYNTHESIS FOR LOCALLY STATIONARY SIGNALS

Adrien Meynard[*]

Duke University,
Department of Mathematics,
Durham, NC, USA

Bruno Torrésani

Aix Marseille Univ, CNRS, Centrale Marseille,
I2M,
Marseille, France

ABSTRACT

We develop a time-scale synthesis-based probabilistic approach for the modeling of locally stationary signals. Inspired by our previous work, the model involves zero-mean, complex Gaussian wavelet coefficients, whose distribution varies as a function of time by time dependent translations on the scale axis. In a maximum *a posteriori* approach, we propose an estimator for the model parameters, namely the time-varying scale translation and an underlying power spectrum. The proposed approach is illustrated on a denoising example. It is also shown that the model can handle locally stationary signals with fast frequency variations, and provide in this case very sharp time-scale representations more concentrated than synchrosqueezed or reassigned wavelet transform.

***Index Terms*—** Wavelet transform, time warping, probabilistic synthesis model

1. INTRODUCTION

Classical time-frequency analysis is generally used for building signal representations from which relevant information can be extracted (see e.g. [1, 2, 3] for reviews). Under suitable assumptions, linear transforms such as the STFT, wavelet transform or generalizations are invertible, which also leads to so-called synthesis approaches [4]. The latter express signals as linear combinations of *time-frequency atoms*, and the corresponding time-frequency coefficients provide another type of time-frequency representation, which is less constrained by consistency requirements and uncertainty principles.

Statistical approaches to time-frequency analysis often rely on *ad hoc* statistical models for time-frequency transforms. Information extraction is then formulated as a statistical estimation problem. Examples include non-negative matrix factorization methods (see [5]), detection of time-frequency components [6, 7], and several other tasks. In most situations, modeling appears as a post-processing stage after computation of a time-frequency transform. However, statistical models are generally not compatible with consistency conditions satisfied by time-frequency transforms.

Recently, Turner and Sahani [8] introduced a new Bayesian paradigm, under the name of probabilistic time-frequency representations. The idea is to express signals as the result of a synthesis from a random time-frequency representation, for which a prior distribution is chosen. This is applied to various contexts, such as the synthesis of stationary signals, and several non-stationary situations, including NMF-based component estimation, non-stationary noise. A similar point of view has already been taken by various authors in the past, see for example [9] and references therein. We rely here on the Turner-Sahani model, which we revisit in a slightly different way, assuming a generative model of the form

$$y(t) = \mathrm{Re}\left(\sum_s (\psi_s * W_s)(t)\right) + \epsilon(t) \ , \qquad (1)$$

where $\{\psi_s, s\}$ is a filter bank, labeled by a frequency (or scale) index s, the $W_s(t)$ are random subband (time-frequency or time-scale) coefficients, and $\epsilon(t)$ is a noise.

Our focus is here on non-stationary signals, more precisely *locally stationary signals* for which time-dependent spectral characteristics can be defined. Following our earlier JEFAS approach [10, 11, 12], we investigate a class of non-stationarity generated by time-dependent shifts in the time-scale domain. JEFAS is analysis based, i.e. post-processing of time-scale representation. We introduce JEFAS-S, a Bayesian *synthesis approach* that provides adaptive time-scale signal representation, together with corresponding parameter estimation. While JEFAS based estimation was based on approximations of the time-scale transform, an exact estimation is possible here, and we provide a corresponding EM algorithm. In addition, in some situations, the model is flexible enough to provide extremely concentrated time-scale representations that can be sharper than reassigned transforms [13].

2. THE SYNTHESIS MODEL

In this paper, we limit to time-scale representation, i.e. wavelet synthesis. We denote by ψ the analysis wavelet, and by ψ_s scaled wavelets defined by $\psi_s(t) = q^{-s/2}\psi\left(q^{-s}t\right)$, for some constant $q > 1$. $s \in \{s_1, \dots s_M\}$ is a finite set of scales.

[*]The first author performed this work while at I2M, Aix-Marseille Université, France.

2.1. The discrete model

We consider the finite periodic case: assume we have N time values $\boldsymbol{\tau} = (\tau_1 \cdots \tau_N)^T$ and the corresponding sampled signal $\mathbf{y} = (y(\tau_1) \cdots y(\tau_N))^T$ with sampling frequency F_s. We then focus on a corresponding discretized wavelet transform. For $n \in \{1, ... N\}, m \in \{1, ... M\}$, denote by $\psi_{nm} \in \mathbb{C}^N$ the vector $(\psi_{s_m}(\tau_{1-n}) \cdots \psi_{s_m}(\tau_{N-n}))^T$, and by $\boldsymbol{\Psi}_n \in \mathbb{C}^{N\times M}$ the matrix obtained by concatenation of vectors $\psi_{nm}, m \in \{1, ... M\}$. The observation equation (1) then reads

$$\mathbf{y} = \mathbf{y}_0 + \boldsymbol{\epsilon} = \mathrm{Re}\left(\sum_{n=1}^{N} \boldsymbol{\Psi}_n \mathbf{w}_n\right) + \boldsymbol{\epsilon}\ , \tag{2}$$

where the $\mathbf{w}_n \in \mathbb{C}^M$, $n = 1, \dots N$ are vectors of synthesis coefficients. This model can also be written in matrix form as $\mathbf{y} = \mathrm{Re}(\mathbf{DW} + \boldsymbol{\epsilon})$, where the dictionary matrix $\mathbf{D}$ is the concatenation of matrices $\boldsymbol{\Psi}_n$, and $\mathbf{W} = \mathsf{vec}(\mathbf{w}_1, \dots \mathbf{w}_N)$.

In this paper, $\boldsymbol{\epsilon}$ will be modeled as a Gaussian white noise, with variance σ^2, as in [8]. Non-stationarity will be introduced via a suitable prior on $\mathbf{W}$, that intends to describe locally time-warped situations as introduced in [14].

2.2. A class of non-stationary priors: time warping

When all subband signals W_s in (1) are stationary, the resulting signal y is stationary. We are interested here in a specific situation where non-stationarity induces a time-dependent shift on the scale axis, as studied in [14, 12, 10]. It was shown there that such a model can account for signals obtained by time warping stationary signals, namely signals of the form

$$y(t) = (\mathcal{D}_\gamma x)(t) \triangleq \sqrt{\gamma'(t)}x(\gamma(t))\ , \tag{3}$$

where x is a wide sense stationary random signal, and γ is a smooth, strictly increasing function.

To build the prior distribution on discrete subband coefficients, we make the following assumptions

- The vectors $\mathbf{w}_n$ are decorrelated, zero-mean, circular complex Gaussian vectors: $\mathbf{w}_n \sim \mathscr{CN}_c(\mathbf{0}, \mathbf{C}_n)$
- The corresponding covariance matrices $\mathbf{C}_n$ are translates of a fixed function f as shown in [10], namely

  $$[\mathbf{C}_n]_{mm'} \triangleq [\mathbf{C}(\theta_n)]_{mm'} = f(s_m{+}\theta_n, s_{m'}{+}\theta_n)\ , \tag{4}$$

  where $f : \left(\mathbb{R}_+^*\right)^2 \to \mathbb{C}$ is Hermitian and positive-semidefinite, and $\theta_n \in \mathbb{R}$ is the shift parameter.

In [10], it was shown that the wavelet coefficients of a stationary random signal modified by time warping transform can be approximated by random vectors satisfying the above assumptions. There, the parameter θ_n represents a local dilation factor at time τ_n (derivative $\gamma'(\tau_n)$ of the time warping function at τ_n), and f involves the power spectrum $\mathscr{S}$ of the underlying signal and the Fourier transform of the wavelet:

$$f(s, s') = q^{\frac{s+s'}{2}} \int_0^\infty \mathscr{S}(\xi)\overline{\hat{\psi}(q^s\xi)}\hat{\psi}(q^{s'}\xi)\, d\xi\ . \tag{5}$$

3. ESTIMATION PROCEDURE

3.1. Bayesian inference

The estimation of the subband coefficient matrix $\mathbf{W}$ relies on the evaluation of the corresponding posterior distribution. The latter depends on the following parameters, which are supposed to be known at this point: the dilation factors θ_n, and the covariance function f. Let $\boldsymbol{\Gamma}_0 \in \mathbb{C}^{MN\times MN}$ be the block diagonal matrix with blocks $\mathbf{C}_1, \dots \mathbf{C}_N$. The posterior distribution of the subband coefficient is a complex Gaussian law $p(\mathbf{W}|\mathbf{y}) \sim \mathscr{CN}(\boldsymbol{\mu}, \boldsymbol{\Gamma}, \mathbf{R})$, with mean and covariance

$$\boldsymbol{\mu} = \boldsymbol{\Gamma}_0\mathbf{D}^H\mathbf{C}_y^{-1}\mathbf{y}\ , \qquad \boldsymbol{\Gamma} = \boldsymbol{\Gamma}_0 - \frac{1}{4}\boldsymbol{\Gamma}_0\mathbf{D}^H\mathbf{C}_y^{-1}\mathbf{D}\boldsymbol{\Gamma}_0\ ,$$

(the relation matrix $\mathbf{R}$, not useful here is not provided) where

$$\mathbf{C}_y = \sigma^2\mathbf{I} + \frac{1}{2}\mathrm{Re}\left(\mathbf{D}\boldsymbol{\Gamma}_0\mathbf{D}^H\right)\ . \tag{6}$$

Therefore the posterior expectation $\tilde{\mathbf{w}}_n$ of $\mathbf{w}_n$ reads

$$\tilde{\mathbf{w}}_n = \frac{1}{2}\mathbf{C}_n\boldsymbol{\Psi}_n^H\mathbf{C}_y^{-1}\mathbf{y}\ , \tag{7}$$

where the matrix $\mathbf{C}_y$ can be expressed as

$$\mathbf{C}_y = \sigma^2\mathbf{I} + \frac{1}{2}\mathrm{Re}\left(\sum_{n=1}^{N}\boldsymbol{\Psi}_n\mathbf{C}_n\boldsymbol{\Psi}_n^H\right)\ . \tag{8}$$

Remark 1. *It is worth mentioning that unlike the prior distribution, the posterior distribution of subband coefficients involves time correlations. Indeed, given any $n, n' = 1 \dots N$,*

$$\mathbb{E}\{\tilde{\mathbf{w}}_n\tilde{\mathbf{w}}_{n'}^H\} = \delta_{nn'}\mathbf{C}_n - \frac{1}{4}\mathbf{C}_n\boldsymbol{\Psi}_n^H\mathbf{C}_y^{-1}\boldsymbol{\Psi}_{n'}\mathbf{C}_{n'}\ ,$$

which generally does not vanish when $n' \neq n$.

3.2. Parameter selection and reconstruction

We now discuss the choice of the two model parameters, i.e. the scaling factors θ_n and the covariance function f. We first notice that the expression in equation (5) provides a natural choice for the covariance function f. The latter involves the wavelet, which is known, and the power spectrum $\mathscr{S}$ of the underlying stationary process, which is unknown. In this setting, we then have to provide the vector $\boldsymbol{\theta} = (\theta_1, \dots \theta_N)$ of scaling factors and the power spectrum.

The JEFAS algorithm introduced in [10] provides a valuable, analysis-based approach for estimating $\boldsymbol{\theta}$ and $\mathscr{S}$. We now describe an alternative algorithm, based on the EM (Expectation Maximization) principle [15], especially tailored for the synthesis approach developed in this paper. Here, $\mathbf{y}$ is the observation, $\boldsymbol{\theta}$ the parameter, and $\mathbf{W}$ the latent variable.

Proposition 1 (EM steps)**.** *Denote by $\tilde{\boldsymbol{\theta}}^{(k-1)}$ the vector of dilation factors at iteration $k-1$ of the algorithm. Let $\tilde{\mathbf{W}}^{(k-1)}$ be the matrix of subband coefficients at iteration $k-1$. Then the update at iteration k relies on the following two steps:*

1. *For $n \in \{1, \ldots, N\}$, the estimate (7) at time τ_n reads*

$$\tilde{\mathbf{w}}_n^{(k)} = \frac{1}{2}\mathbf{C}\left(\tilde{\theta}_n^{(k-1)}\right)\mathbf{\Psi}_n^H \mathbf{C}_y\left(\tilde{\boldsymbol{\theta}}^{(k-1)}\right)^{-1}\mathbf{y} \ . \quad (9)$$

2. *The scaling factor $\tilde{\boldsymbol{\theta}}$ is re-estimated by solving*

$$\tilde{\theta}_n^{(k)} = \arg\min_{\theta} \ Q_{kn}(\theta) \ , \quad (10)$$

$$Q_{kn}(\theta) = \Big[\log|\det(\mathbf{C}(\theta))| + \tilde{\mathbf{w}}_n^{(k)H}\mathbf{C}(\theta)^{-1}\tilde{\mathbf{w}}_n^{(k)} + \mathrm{Trace}\left(\mathbf{C}(\theta)^{-1}\mathbf{\Gamma}_n\left(\tilde{\boldsymbol{\theta}}^{(k-1)}\right)\right)\Big] \ ,$$

$\mathbf{\Gamma}_n\left(\tilde{\boldsymbol{\theta}}^{(k-1)}\right) \in \mathbb{C}^{M\times M}$ *being the n-th diagonal block of the posterior covariance matrix* $\mathbf{\Gamma}\left(\tilde{\boldsymbol{\theta}}^{(k-1)}\right)$.

After running the corresponding algorithm (described in more details below), an estimate for the time-scale coefficients $\tilde{\mathbf{w}}$ is available, and a corresponding estimate $\tilde{\mathbf{y}}_0$ for the signal $\mathbf{y}_0$ can be obtained as

$$\tilde{\mathbf{y}}_0 = \mathrm{Re}\left(\sum_{n=1}^{N}\mathbf{\Psi}_n\tilde{\mathbf{w}}_n\right) \ . \quad (11)$$

Notice that the reconstruction expression (11) combined with (7) can be interpreted as a Wiener filtering. The bias and variance of the estimator can be evaluated.

Proposition 2. *With the above notation, the bias of the estimator $\tilde{\mathbf{y}}_0$ is given by*

$$\mathbf{B} \triangleq \mathbb{E}\left\{\tilde{\mathbf{y}}_0|\mathbf{y}_0\right\} - \mathbf{y}_0 = -\sigma^2\mathbf{C}_y^{-1}\mathbf{y}_0 \ , \quad (12)$$

and the corresponding error variance reads

$$\mathbf{R}(\tilde{\mathbf{y}}_0|\mathbf{y}_0) \triangleq \mathbb{E}\left\{\left(\tilde{\mathbf{y}}_0 - \mathbb{E}\{\tilde{\mathbf{y}}_0|\mathbf{y}_0\}\right)\left(\tilde{\mathbf{y}}_0 - \mathbb{E}\{\tilde{\mathbf{y}}_0|\mathbf{y}_0\}\right)^T \Big| \mathbf{y}_0\right\} = \sigma^2\left(\mathbf{I} - \sigma^2\mathbf{C}_y^{-1}\right)^2 \ . \quad (13)$$

3.3. Algorithm: JEFAS-Synthesis

The steps of the estimation algorithm are given in Algorithm 1. The latter takes as input the signal $\mathbf{y}$, the noise variance σ^2, a precision parameter Λ for the stopping criterion and a bandwidth parameter N' (see below).

Initialization. The algorithm requires initial estimates $\boldsymbol{\theta}^{(0)}$ for the parameters, and the function f in (4). In JEFAS-S, we use the expression (5), for which an initial estimate of $\mathscr{S}$ has to be provided. When successful, JEFAS [10] provides such an estimate. Otherwise, a rough estimate can be obtained from the Welch periodogram of the input signal $\mathbf{y}$.

Stopping criterion. EM guarantees the monotonicity of the Likelihood function $\mathcal{L}(\boldsymbol{\theta})$. The increment of the latter is used as a stopping criterion: EM will stop when the condition

$$\mathcal{L}(\boldsymbol{\theta}^{(k)}) - \mathcal{L}(\boldsymbol{\theta}^{(k-1)}) < \Lambda \quad (14)$$

is true. Here $\Lambda > 0$ is a parameter fixed by the user.

Algorithm 1 $(\tilde{\mathbf{W}}, \tilde{\boldsymbol{\theta}}, \tilde{\mathscr{S}}_X) = \text{JEFAS-S}(\mathbf{y}, \sigma^2, \Lambda, N')$

- **Initialization:** estimate $\tilde{\boldsymbol{\theta}}^{(0)}$ and $\tilde{\mathscr{S}}^{(0)}$ using JEFAS.
- $k \leftarrow 1$.

while stopping criterion (14) = FALSE **do**
  for all $n \in \{1, \ldots, N\}$ **do**
    - Restrict $\mathbf{\Psi}_n$, $\mathbf{C}_y\left(\tilde{\boldsymbol{\theta}}^{(k-1)}\right)$ and $\mathbf{y}$ to the interval $[n - N'/2, n + N'/2]$.
    - Compute $\tilde{\mathbf{w}}_n^{(k)}$ using (9).

  end for
  - Estimate $\tilde{\boldsymbol{\theta}}^{(k)}$ by solving (10).
  - Estimate $\tilde{\mathscr{S}}^{(k)}$ using the wavelet based estimate.
  - $k \leftarrow k + 1$.

end while

Dimension reduction. The matrix $\mathbf{C}_y$ of dimension $NM \times NM$ can be extremely large. However, it generally has fast off-diagonal decay. This can be exploited to speed up the evaluation of $\tilde{\mathbf{w}}_n$ in (9) by restricting to a neighborhood $[n - N'/2, n + N'/2]$ of n of given bandwidth N'.

Optimization. The optimization problem (10) is solved using a standard quasi-Newton scheme.

Spectrum estimate update. The spectrum update from the current estimate of $\mathbf{W}$ is performed in two steps: first correct for the translation by θ_n, to obtain an approximately stationary subband transform, then average over time to obtain a wavelet based spectral estimate as in [10].

Remark 2. *Other choices can be made for the function f, which can lead to different estimates for subband coefficients, while preserving reconstruction (see section 4.2).*

4. NUMERICAL RESULTS

4.1. Illustration on denoising of a synthetic signal

We first evaluate the performances of JEFAS-S on a denoising problem. A synthetic non-stationary signal $\mathbf{y}$ is built as follows: start from a stationary signal $\mathbf{x}$, with power spectrum $\mathscr{S}$ equal to the sum of two non-overlapping Hann windows, and apply the time warping deformation $\mathcal{D}_\gamma$ to $\mathbf{x}$, with γ' an exponentially damped sine wave. Here, $\mathbf{x}$ is one second long, sampled at $F_s = 8192$ Hz.

We denote by $\mathsf{SNR}_{\mathbf{y}}$ and $\mathsf{SNR}_{\tilde{\mathbf{y}}_0}$ the input and output signal-to-noise ratios. Numerical results show that $\mathsf{SNR}_{\tilde{\mathbf{y}}_0}$ is larger than $\mathsf{SNR}_{\mathbf{y}}$ as long as $\mathsf{SNR}_{\mathbf{y}}$ is in the range $[2\text{ dB}, 25\text{ dB}]$, with maximal improvement of 8 dB. The 25 dB upper limit for $\mathsf{SNR}_{\tilde{\mathbf{y}}_0}$ is presumably due the distortion intrinsically introduced by the reconstruction formula (11): bias and variability in the time warping estimation.

In the specific case where the input SNR is 16 dB, and after initializing with the output of JEFAS, JEFAS-S converges in 3 iterations (CPU time: 347 seconds on a computer running

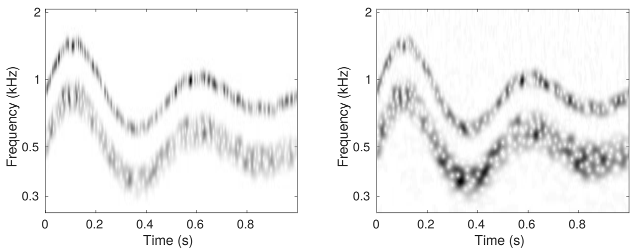

Fig. 1. Synthetic signal. Left: representation given by JEFAS-S. Right: scalogram (wavelet transform).

an Intel Xeon E5-2680 v4 processor). JEFAS-S does not significantly improve the quality of the estimated time warping function. Indeed, the mean square error on the time warping function estimation decreases by about 0.5% from JEFAS to JEFAS-S. We display the estimated adapted time-scale representation $\tilde{\mathbf{W}}$ in Fig. 1 (left). As expected, it is very similar to the wavelet transform (right), though a bit sharper. Indeed, the choice of the expression (5) for the covariance function f yields a wavelet-like representation. The main visible difference concerns the temporal oscillations of $\tilde{\mathbf{W}}$, due to the prior assumption of temporal decorrelation between $\mathbf{w}_n$.

4.2. Locally harmonic signal with fast varying frequency

We now consider a locally harmonic signal, of the form

$$y(t) = A(t)\cos(2\pi\phi(t)) \ ,$$

where the instantaneous frequency ϕ' is a fast varying function chosen as the measurement of the heart rate of a person suffering from atrial fibrillation (real data). The synthetic instantaneous amplitude A is a slowly varying function, the signal is termed "semi-real". Apart from the amplitude modulation, this signal follows the model (3): the time warping function derivative is the instantaneous frequency, and the underlying stationary signal x is sinusoidal. The signal duration is 83.1 seconds, sampled at $F_{\mathrm{s}} = 10$ Hz ($N = 832$ samples).

Because of the fast instantaneous frequency variations, the wavelet transform of y (not shown here) contains interference patterns, the model in [10] is not adequate and JEFAS does not converge. We initialized JEFAS-S to $\boldsymbol{\theta} = \mathbf{0}$, and a constant function for $\mathscr{S}$. Given these initial values (far from actual values), JEFAS-S converges slowly (72 iterations). Results are displayed in Fig. 2, where the estimated instantaneous frequency is superimposed on the ground-truth. This shows that JEFAS-S is indeed able to estimate fast varying frequency modulations.

In addition, one can take advantage of this result to obtain a sharper time-scale representation. To that end, we choose a sharply concentrated prior covariance function $f_\sharp$, of the form $f_\sharp(s,s') = \exp\{-(s-\varsigma)^2/\sigma_s^2\}\delta_{ss'}$, where ν_1 denotes the central frequency of the sine wave, $\varsigma = \log_q(\xi_0/\nu_1)$, and σ_s is a tuning parameter for the scale concentration. We display in the top of Fig. 3 the covariance matrices $\mathbf{C}(0)$ corresponding to the expression (5) (left) and $f_\sharp$ (right), which is indeed

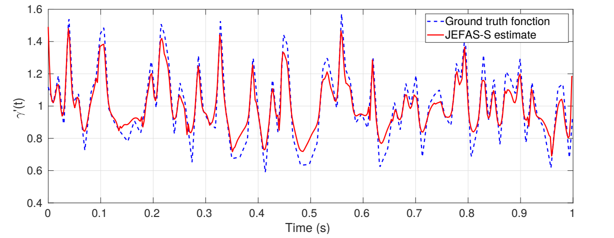

Fig. 2. Semi-real signal. Estimated time warping function compared with the normalized instantaneous frequency ϕ'.

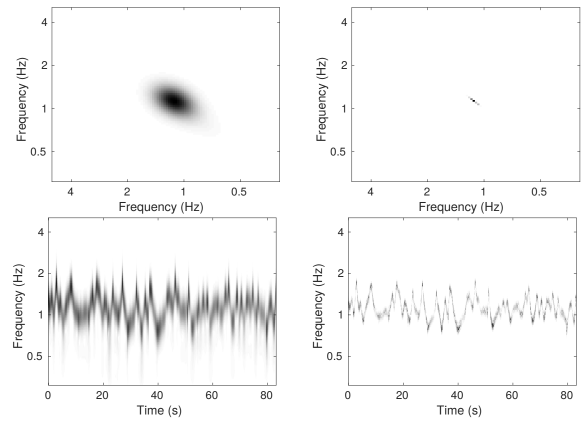

Fig. 3. Semi-real signal. Top: two priors for the covariance matrix. Bottom: associated time-scale representations.

very sharp. The corresponding estimated time-scale representations are displayed on the the bottom images of Fig. 3. The new prior is clearly adapted to locally harmonic signals, i.e. signals with a sparse underlying spectrum. Thus, in such situations, JEFAS-S enables the construction of sharp time-scale representations, competing with standards techniques such as synchrosqueezing. Furthermore, we stress that the quality of the reconstruction is not degraded.

5. CONCLUSION

We have described an alternative to the JEFAS model of [10] for locally deformed signals. Unlike JEFAS, which is an analysis based approach (i.e. post-processing of wavelet transform), JEFAS-S is synthesis-based and therefore less constrained by uncertainty principles. We illustrated the JEFAS-S on a denoising example. Our numerical results also show that JEFAS-S is able to handle locally stationary signals with fast varying instantaneous frequency, and can provide very sharp time-scale representations.

While the current paper was focused on wavelet transform, the JEFAS-S model can handle arbitrary subband decompositions (such as the NSDGT [16]). Such extensions will be discussed in a forthcoming publication, together with additional numerical results and complete proofs. JEFAS-S can also be extended to more general transformations, for example involving amplitude modulations or filtering posterior to time warping. This is an ongoing work.

6. REFERENCES

[1] René Carmona, Wen-Liang Hwang, and Bruno Torrésani, *Practical time-frequency analysis: Gabor and Wavelet Transforms With an Implementation in S*, Academic Press, 1998.

[2] Patrick Flandrin, *Explorations in Time-Frequency Analysis*, Cambridge University Press, 2018.

[3] Karlheinz Gröchenig, *Foundations of time-frequency analysis*, Applied and Numerical Harmonic Analysis. Birkhäuser Inc., Boston, MA, 2001.

[4] Peter Balazs, Monika Dörfler, Matthieu Kowalski, and Bruno Torrésani, "Adapted and adaptive linear time-frequency representations: a synthesis point of view," *IEEE Signal Processing Magazine*, vol. 30, no. 6, pp. 20–31, Nov. 2013.

[5] Cédric Févotte and Ali T. Cemgil, "Nonnegative matrix factorisations as probabilistic inference in composite models," in *Proc. 17th European Signal Processing Conference (EUSIPCO)*, Glasgow, Scotland, Aug. 2009, pp. 1913–1917.

[6] Isabella Reinhold, Maria Sandsten, and Josefin Starkhammar, "Objective detection and time-frequency localization of components within transient signals," *The Journal of the Acoustical Society of America*, vol. 143, no. 4, pp. 2368–2378, 2018.

[7] Julien Huillery, Fabien Millioz, and Nadine Martin, "On the description of spectrogram probabilities with a chi-squared law," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2249 – 2258, June 2008.

[8] Richard E. Turner and Maneesh Sahani, "Time-frequency analysis as probabilistic inference," *IEEE Transactions on Signal Processing*, vol. 62, no. 23, pp. 6171–6183, Dec. 2014.

[9] Manuel Davy, Simon J. Godsill, and Jérôme Idier, "Bayesian Analysis of Polyphonic Western Tonal Music," *Journal of the Acoustical Society of America*, vol. 119, no. 4, pp. 2498–2517, 2006.

[10] Adrien Meynard and Bruno Torrésani, "Spectral Analysis for Nonstationary Audio," *IEEE/ACM Transactions on Audio, Speech and Language Processing*, vol. 26, no. 12, pp. 2371 – 2380, Dec. 2018.

[11] Adrien Meynard, "Séparation de sources doublement non stationnaire," in *GRETSI 2019 - XXVIIème Colloque francophone de traitement du signal et des images*, Lille, France, Aug. 2019.

[12] Harold Omer and Bruno Torrésani, "Time-frequency and time-scale analysis of deformed stationary processes, with application to non-stationary sound modeling," *Applied and Computational Harmonic Analysis*, vol. 43, no. 1, pp. 1 – 22, 2017.

[13] François Auger and Patrick Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Transactions on Signal Processing*, vol. 43, no. 5, pp. 1068–1089, May 1995.

[14] Maureen Clerc and Stéphane Mallat, "Estimating deformations of stationary processes," *Ann. Statist.*, vol. 31, no. 6, pp. 1772–1821, Dec. 2003.

[15] Arthur P. Dempster, Nan M. Laird, and Donald B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.

[16] Peter Balazs, Monika Dörfler, Florent Jaillet, Nicki Holighaus, and Gino Velasco, "Theory, implementation and applications of nonstationary Gabor frames," *Journal of Computational and Applied Mathematics*, vol. 236, no. 6, pp. 1481 – 1496, 2011.