Postprint of: Czarnecki K., Fourer D., Auger F., Rojewski M., A fast time-frequency multi-window analysis using a tuning directional kernel, Signal Processing, Vol. 147 (2018), pp. 110-119, DOI: 10.1016/j.sigpro.2018.01.019

A fast time-frequency multi-window analysis using a tuning directional kernel

Krzysztof Czarnecki, Dominique Fourer, François Auger, Mirosław Rojewski

K. Czarnecki is with the Gdansk University of Technology,
Faculty of Electronics, Telecommunications and Informatics, Poland
krzycz@eti.pg.gda.pl, czarnecki.krzysiek@gmail.com

D. Fourer is with UMR STMS (IRCAM - CNRS - UPMC),
Paris, France

F. Auger is with LUNAM University, IREENA,
Saint-Nazaire, France

M. Rojewski is with the Gdansk University of Technology,
Faculty of Electronics, Telecommunications and Informatics, Poland

In this paper, a novel approach for time-frequency analysis and detection, based on the chirplet transform and dedicated to non-stationary as well as multi-component signals, is presented. Its main purpose is the estimation of spectral energy, instantaneous frequency (IF), spectral delay (SD), and chirp rate (CR) with a high time-frequency resolution (separation ability) achieved by adaptive fitting of the transform kernel. We propose two efficient implementations of this idea, which allow to use the fast Fourier transform (FFT). In the first one, referred to as "self-tuning", a previously proposed CR estimation is used for a local fitting of the chirplet kernel over time. For this purpose, we use the CR associated with the dominant (prominent) component. In the second one, we define a new measure for evaluating at each time-frequency point, how the used analyzing window is matched to the signal. This measure is defined as the absolute difference between the estimated CR and the CR parameter associated to the used analysis window. Our method is able to produce combined time-frequency distributions of the spectral energy, IF, SD, and CR. They are obtained using several classical chirplet transforms with analysis windows of various CRs. The compositions are made by finding the lowest fitting measure for every time-frequency points over all transforms. Finally, we assess the robustness of the

methods by a detection application and time-frequency localization, both in the presence of high additive white Gaussian noise (AWGN) as well as we present many TF images of synthetic and real-world signals.

Keywords: time-frequency analysis; chirp rate estimation; short-time Fourier transform; STFT; adaptation; reassignment; matching pursuit; flute sound; sonar signal.

1 Introduction

Our world is full of non-stationary and multicomponent signals. We say "everything flows" (gr. *panta rei*) after Heraclitus of Ephesus. In a signal processing context, this statement also concerns parameters which describe the signals that surround us. This is the reasons why the development of new methods dedicated to analyze non-stationary and multi-component signals, has a substantial practical importance. Simultaneously, it is commonly known, that time-frequency analysis is one of the most important and powerful approach designed to study real-world signals [1]. In fact, these signals are often non-stationary and consist of numerous components, including noise [2]. Despite many existing approaches [3–10], this paper focuses on well known tools such as the short-time Fourier transform (STFT) and one of its variants, the chirplet transform (CT) [11, 12]. This choice is motivated by results of our recent researches [13, 14] which allow to directly use the STFT in order to locally estimate the chirp rate (CR) in the time-frequency (TF) domain. The resulting distribution can shortly be called "TF phase accelerogram". Our main idea is to use the accelerogram to locally fit the kernel of the CT for each TF point. This adaptive approach can be considered as a contribution to the "matching pursuit" trend [15], since its main purpose is to locally improve the TF resolution (separation ability). Moreover, this approach can be applied iteratively in order to improve the estimation reliability.

Unfortunately, this idea has one fundamental drawback, which is a high computational cost, if the estimated phase accelerogram has different values – which must be assumed for non-stationary signals. Then, the FFT algorithm cannot be applied and the discrete Fourier transform (DFT) has to be directly computed according to its definition for every TF point independently. This can significantly increase the computational cost of the analysis. Therefore, we propose two efficient solutions, which allow a FFT-based implementation. The first one uses a CR-related analysis window adjusted over time. The considered CR should be prominent for a considered analyzed instant or/and can be established by a statistical analysis of a previously computed phase accelerogram with

its corresponding spectral energy distribution. Both are estimated using STFT whose analyzing window has CR equal to zero. This solution should be suitable for signals with a single dominant component or such, whose components are characterized by a similar time-varying CR, for example, speech. In a second solution, initially, several (or more) CTs, their energy distributions, and accelerograms using FFT for different window CRs have to be computed. Then, we define a simple indicator of "mismatch", and based on it, we introduce a new TF energy distribution directly computed from the CT. The final composed energy distribution can contain energy from the different transforms, however a single TF point represents only one, for which the measure is the smallest over all consider transforms in this point. These implementations can also be used to compute other signal parameters such as instantaneous frequency (IF), spectral delay (SD), and CR.

To summarize this, we try to join two main trends of signal analysis: estimation of local CR in the TF domain and local fitting of the kernel of a time-frequency analyzer [8,16,17] which is the chirplet transformation [11]. Both topics are not new. Many authors proposed interesting methods of CR estimation. Here, we list only selected papers [18–23]. However, we base on our original study and fast estimators introduced in [13,14], which operates in the time-frequency domain directly on STFT (especially on CT) and have evolved since the reassignment approach [24–26]. We can also find some proposition to use the CR for adjusting the analysis window width [27,28]. However, the CR estimation based on CT is presented here the first time.

Hence, this paper is organized as follows. In Section 2, the chirplet transform is introduced with several local CR estimators. In this section, we also define the reassignment operators of the CT and its reassigned time-frequency representation. Then, the new proposed self-tuning chirplet method and multi-window chirplet-based TF distributions are introduced respectively, in Sections 3 and 4. Conclusions are then presented in Section 5.

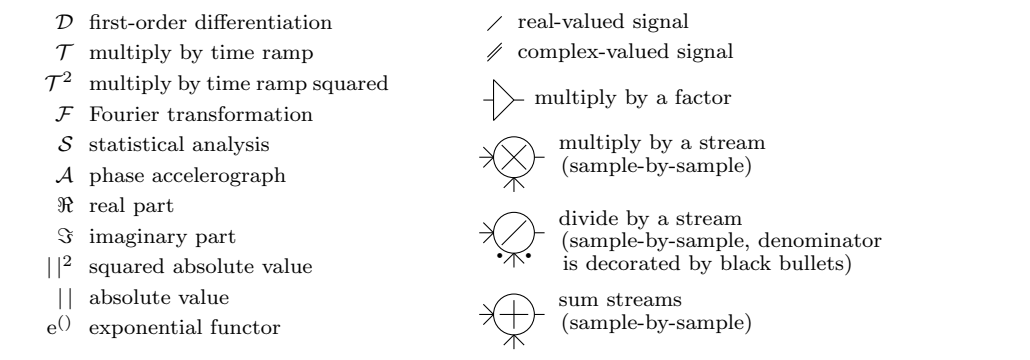

Figure 1: Notations for diagrams.

2 Chirplet-based analysis

The spectral parameters of an analyzed signal $x(t)$, such as IF and CR, can be estimated at any time-frequency coordinates (t, ω) through the STFT as proposed in [13,14]. Herein, we extend this approach to the CT which can be defined by a convolution product with a function g_α expressed as:

$$\begin{aligned} y_x^g(t,\omega) &= \int_{\mathbb{R}} x(\tau) g_\alpha(t-\tau,\omega) \mathrm{d}\tau = \int_{\mathbb{R}} x(\tau) h(t-\tau) \mathrm{e}^{\mathrm{j}\alpha\frac{(t-\tau)^2}{2}} \mathrm{e}^{\mathrm{j}\omega(t-\tau)} \mathrm{d}\tau \\ &= M_x^g(t,\omega) \mathrm{e}^{\mathrm{j}\phi_x^g(t,\omega)} \end{aligned} \tag{1}$$

where $M_x^g(t,\omega) = \left|y_x^g(t,\omega)\right|$, $\phi_x^g(t,\omega) = \arg\left(y_x^g(t,\omega)\right)$ stand respectively for the magnitude and the phase of the CT, j being the imaginary unit ($\mathrm{j}^2 = -1$) and e being the Euler's number. The kernel of this transformation corresponds to a linear frequency modulated chirp tapered by a real-valued and differentiable analysis window $h(t)$ as follows:

$$g_\alpha(t,\omega) = \underbrace{h(t)\mathrm{e}^{\mathrm{j}\alpha\frac{t^2}{2}}}_{h_\alpha(t)} \mathrm{e}^{\mathrm{j}\omega t} \tag{2}$$

α being the CR parameter of the analysis window. In this paper, we use the 4-term Blackman-Harris window [29] as $h(t)$, because its side lobes are strongly suppressed and even its high order derivatives are easy to compute.

2.1 Chirp rate estimation

As proposed in [13], a TF phase accelerogram can be estimated through the amplitude of the STFT. Following this idea and using notations consistent with [14], we can propose a new CR estimator based on the CT as defined by Eq. (1) using specific analysis windows $\mathcal{D}g_\alpha(t,\omega)$ and $\mathcal{T}g_\alpha(t,\omega)$ as:

$$\hat{R}_x^g(t,\omega) = \frac{\Re\left(\frac{y_x^{\mathcal{D}g}(t,\omega)}{y_x^g(t,\omega)}\right)}{\Im\left(\frac{y_x^{\mathcal{T}g}(t,\omega)}{y_x^g(t,\omega)}\right)}, \tag{3}$$

where $\Re()$ and $\Im()$ return respectively, the real and the imaginary parts of a complex number. In the corresponding software [30], we simply denote this estimator as "K". $\mathcal{D}g_\alpha(t,\omega)$ is the first-order derivative of the window $g_\alpha(t,\omega)$ with respect to time:

$$\mathcal{D}g_\alpha(t,\omega) = \frac{\partial g_\alpha(t,\omega)}{\partial t} = \left(\frac{dh_\alpha}{dt}(t) + \mathrm{j}\omega h_\alpha(t)\right) \mathrm{e}^{j\omega t} \tag{4}$$

as well as $\mathcal{T}g_\alpha(t,\omega)$ is the product of the window $g_\alpha(t,\omega)$ and a time ramp function which

is simply the linear odd function with a slope equal to one:

$$\mathcal{T}g_\alpha(t,\omega) = t\ g_\alpha(t,\omega). \tag{5}$$

Since we proposed in [14] enhanced CR estimators based on higher-order derivatives, we can also deduce similar ones by applying L'Hôpital's rule on Eq. (3). These new estimators can be expressed as follows:

$$\hat{\dot{R}}_x^g(t,\omega) = \frac{\frac{\partial}{\partial t}\left[\Re\left(\frac{y_x^{\mathcal{D}g}(t,\omega)}{y_x^g(t,\omega)}\right)\right]}{\frac{\partial}{\partial t}\left[\Im\left(\frac{y_x^{\mathcal{T}g}(t,\omega)}{y_x^g(t,\omega)}\right)\right]} = \frac{\Re\left(\frac{y_x^{\mathcal{D}^2 g}(t,\omega)}{y_x^g(t,\omega)} - \frac{y_x^{\mathcal{D}g}(t,\omega)^2}{y_x^g(t,\omega)^2}\right)}{\Im\left(\frac{y_x^{\mathcal{D}\mathcal{T}g}(t,\omega)}{y_x^g(t,\omega)} - \frac{y_x^{\mathcal{D}g}(t,\omega)}{y_x^g(t,\omega)}\frac{y_x^{\mathcal{T}g}(t,\omega)}{y_x^g(t,\omega)}\right)} \tag{6}$$

as well as

$$\hat{\acute{R}}_x^g(t,\omega) = \frac{\frac{\partial}{\partial \omega}\left[\Re\left(\frac{y_x^{\mathcal{D}g}(t,\omega)}{y_x^g(t,\omega)}\right)\right]}{\frac{\partial}{\partial \omega}\left[\Im\left(\frac{y_x^{\mathcal{T}g}(t,\omega)}{y_x^g(t,\omega)}\right)\right]} = -\frac{\Im\left(\frac{y_x^{\mathcal{D}\mathcal{T}g}(t,\omega)}{y_x^g(t,\omega)} - \frac{y_x^{\mathcal{D}g}(t,\omega)}{y_x^g(t,\omega)}\frac{y_x^{\mathcal{T}g}(t,\omega)}{y_x^g(t,\omega)}\right)}{\Re\left(\frac{y_x^{\mathcal{T}^2 g}(t,\omega)}{y_x^g(t,\omega)} - \frac{y_x^{\mathcal{T}g}(t,\omega)^2}{y_x^g(t,\omega)^2}\right)}, \tag{7}$$

where $\mathcal{D}^2 g_\alpha(t,\omega) = \frac{\partial^2 g_\alpha(t,\omega)}{\partial t^2}$, $\mathcal{T}^2 g_\alpha(t,\omega) = t^2 g_\alpha(t,\omega)$ and $\mathcal{D}\mathcal{T}g_\alpha(t,\omega) = g_\alpha(t,\omega) + \mathcal{D}g_\alpha(t,\omega)$. In the corresponding software [30], we simply denote this estimators, respectively, as "D" and "F". These estimators obtain better results when they are applied close to the component attractors identified in the TF domain by their instantaneous frequencies and their spectral delays [2, 14, 24, 25, 31, 32]. However, despite the spectral energy is often concentrated in these areas, both the numerator and denominator of Eq. (3) can go to zero causing numerical instabilities. Therefore, estimators (6) and (7) were shown to be more robust than (3) [14].

2.2 Directional kernel

The concept of the directional kernel (TF-oriented) in the context of adaptive TF analysis is still developed [8]. In very few words, this approach is closely associated with the ambiguity function of the analysis window used by CT which is defined by:

$$A(\tau,\nu) = \int_{\mathbb{R}} h_\alpha(t) h_\alpha(t-\tau)^* \mathrm{e}^{j\nu t}\, dt. \tag{8}$$

Significant values of this function are distributed into the ambiguity area according to the window envelope $h(t)$ and its CR α as illustrated in Fig. 2. Especially, the CR indicates a direction in the TF domain. The energy of each component, whose attractor is distributed along to this direction, is well concentrated in the spectrogram obtained using this window. If the CR is exactly equal to this direction, the window is locally

matched to this component. This aspect highlights the importance of a proper selection of the window and of its parameters, especially its CR α.

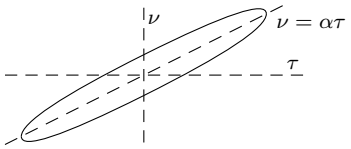

Figure 2: Illustration of the ambiguity area (represented in the time-delay / Doppler-frequency domain) of a chirped analysis window whose envelope is defined by the Gaussian-like function and where CR α is constant.

As CT is related to the short-time fractional Fourier transform for which the theory was further investigated in [33], it can be shown that CT divides the TF plane into parallelograms which have the same resolution in time as the STFT. However, the CT's frequency resolution depends on the parameter α. For a given STFT's frequency resolution $\sigma_\omega^2 = \frac{\int_{\mathbb{R}}(\omega-\bar{\omega})^2|F_h(\omega)|^2 d\omega}{\int_{\mathbb{R}}|F_h(t,\omega)|^2 d\omega}$ corresponding to the squared bandwidth around a central frequency $\bar{\omega} = \frac{\int_{\mathbb{R}}\omega|F_h(\omega)|^2 d\omega}{\int_{\mathbb{R}}|F_h(t,\omega)|^2 d\omega}$, where $F_h(\omega)$ is the Fourier transform of window $h(t)$, thus CT modifies the frequency resolution such as [33]:

$$\sigma_{\omega,\alpha}^2 = \sigma_\omega^2 \sin(\theta)^2, \quad \text{with} \;\; \theta = \tan^{-1}(\alpha)\,. \tag{9}$$

2.3 Time-frequency reassignment

IF and SD (or group delay) point component attractors in the time-frequency domain. In TF reassignment, they are used to improve the energy concentration, which is relocated close to these attractors [2, 12, 24, 25, 31, 32]. IF and SD can also be estimated using the chirplet transform in the TF domain. IF corresponds to the partial derivative of the CT's phase, with respect to time and can be obtained by the following manner [26, 34]:

$$\hat{\Omega}_x^g(t,\omega) = \Im\left(\frac{y_x^{\mathcal{D}g}(t,\omega)}{y_x^g(t,\omega)}\right). \tag{10}$$

SD is defined as the partial derivative of the CT's phase with respect to frequency which can be computed as:

$$\hat{D}_x^g(t,\omega) = -\Re\left(\frac{y_x^{\mathcal{T}g}(t,\omega)}{y_x^g(t,\omega)}\right). \tag{11}$$

Then, the new TF localization of spectral energy is indicated by $\hat{\Omega}_x^g(t,\omega)$ and, dually, by corrected time defined as:

$$\hat{\Gamma}_x^g(t,\omega) = t + D_x^g(t,\omega) = t - \Re\left(\frac{y_x^{\mathcal{T}g}(t,\omega)}{y_x^g(t,\omega)}\right). \tag{12}$$

Then, both IF and corrected time can be used for energy reassignment following [35]:

$$\Sigma_x^g(t,\omega) = \iint_{\mathbb{R}^2} E_x^g(\tau,\upsilon)\delta\big(t - \hat{\Gamma}_x^g(\tau,\upsilon)\big)\delta\big(\omega - \hat{\Omega}_x^g(\tau,\upsilon)\big)\mathrm{d}\tau\mathrm{d}\upsilon, \tag{13}$$

with

$$E_x^g(t,\omega) = \left|y_x^g(t,\omega)\right|^2 = M_x^g(t,\omega)^2, \tag{14}$$

where $\delta(t)$ denotes the Dirac distribution, $E_x^g(t,\omega)$ and $\Sigma_x^g(t,\omega)$ denote respectively, the classical and the reassigned energy distributions. A variant of the reassignment method which admits a reconstruction formula is called synchrosqueezing. Despite this technique obtains a poorer time-frequency localization than energy reassignment, it preserves the phase of the original transform and can be used for disentangling signal components as proposed in [36].

3 Self-tuning chirplet transform

The main idea of the self-tuning chirplet transform (STCT) is to fit the directional kernel of CT according to the TF accelerogram estimated using the STFT. This method is realized in a two-stage algorithm presented in Fig. 3. Firstly, the window parameter α is assumed to be equal to zero. Then, an accelerogram is computed using Eqs. (3), (6), or (7). The first stage is completed after estimating a dominant instantaneous CR as a function of time defined by:

$$r_x^g(t) = \hat{R}_x^g\big(t, \operatorname*{argmax}_{\omega}\big(E_x^g(t,\omega)\big)\big), \tag{15}$$

where $\hat{R}_x^g$ is a TF accelerogram (computed using one of the proposed estimators (3), (6) or (7)) which is a function of time t and of frequency ω. The operator argmax returns the frequency for which the maximum energy is reached. Finally, the STCT can be computed as:

$$y_x'^g(t,\omega) = \int_{\mathbb{R}} x(\tau)h(t-\tau)\mathrm{e}^{\mathrm{j}\frac{r_x^g(t)(t-\tau)^2}{2}}\mathrm{e}^{\mathrm{j}\omega(t-\tau)}\,\mathrm{d}\tau. \tag{16}$$

This new transform can also be used to compute second-order estimations of IF, SD and CR as for the STFT. A TF representation corresponding to the distribution of the spectral energy is provided by $|y_x'^g(t,\omega)|^2$ and its readability can be improved by a better TF localization using the time-frequency reassignment [24, 25].

3.1 Accuracy for a selected synthetic FM signal

Let us consider a synthetic simple model, which is an analytical FM signal:

$$x_{\sin}(t) = \exp\big(\mathrm{j}2\pi\sin(t)\big), \tag{17}$$

whose instantaneous CR is defined as the second-order derivative of its phase (expressed in rad/s^2):

$$R_{\sin}(t) = \frac{\mathrm{d}^2}{\mathrm{d}t^2}\left[\arg\big(x_{\sin}(t)\big)\right] = -2\pi\sin(t). \tag{18}$$

Its instantaneous angular frequency defined as the first-order derivative of the phase is:

$$\Omega_{\sin}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left[\arg\big(x_{\sin}(t)\big)\right] = 2\pi\cos(t), \tag{19}$$

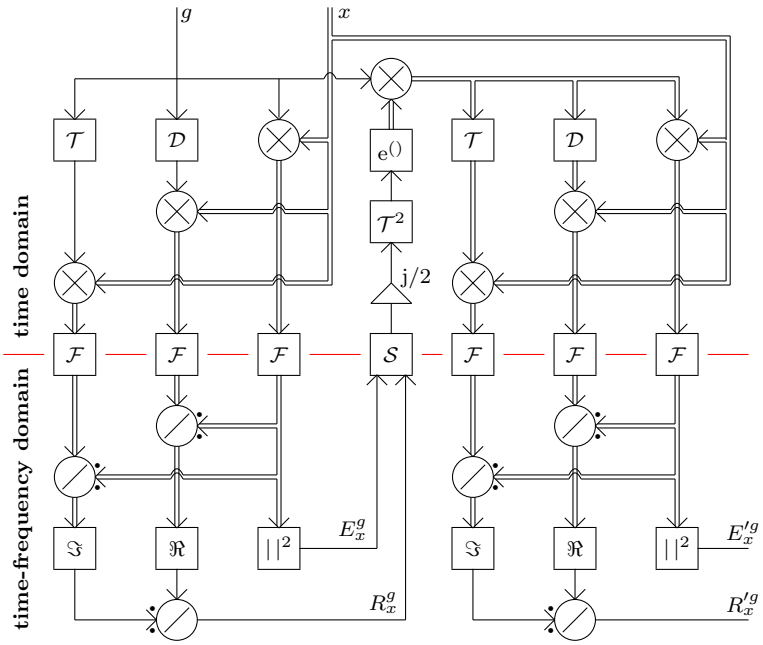

Figure 3: Block scheme of the proposed self-tuning chirplet spectrograph and accelerograph. In this variant, the CR estimator defined by Eq. (3) is applied. The legend for this diagram is presented in Fig. 1.

which is clearly seen in the presented spectrograms in Fig. 4 (as high energy density along this curve) and in accelerograms in Fig. 5 (directly, as value along this curve). The CR estimation for this signal is more precise in accelerograms computed through STCT in comparison to STFT. This is confirmed by the curves presented in Fig. 6 for a window width equal to 2.75 s which displays for the signal of Eq. (17), the relation between the energy localized by the transform where the CR estimation error is computed.

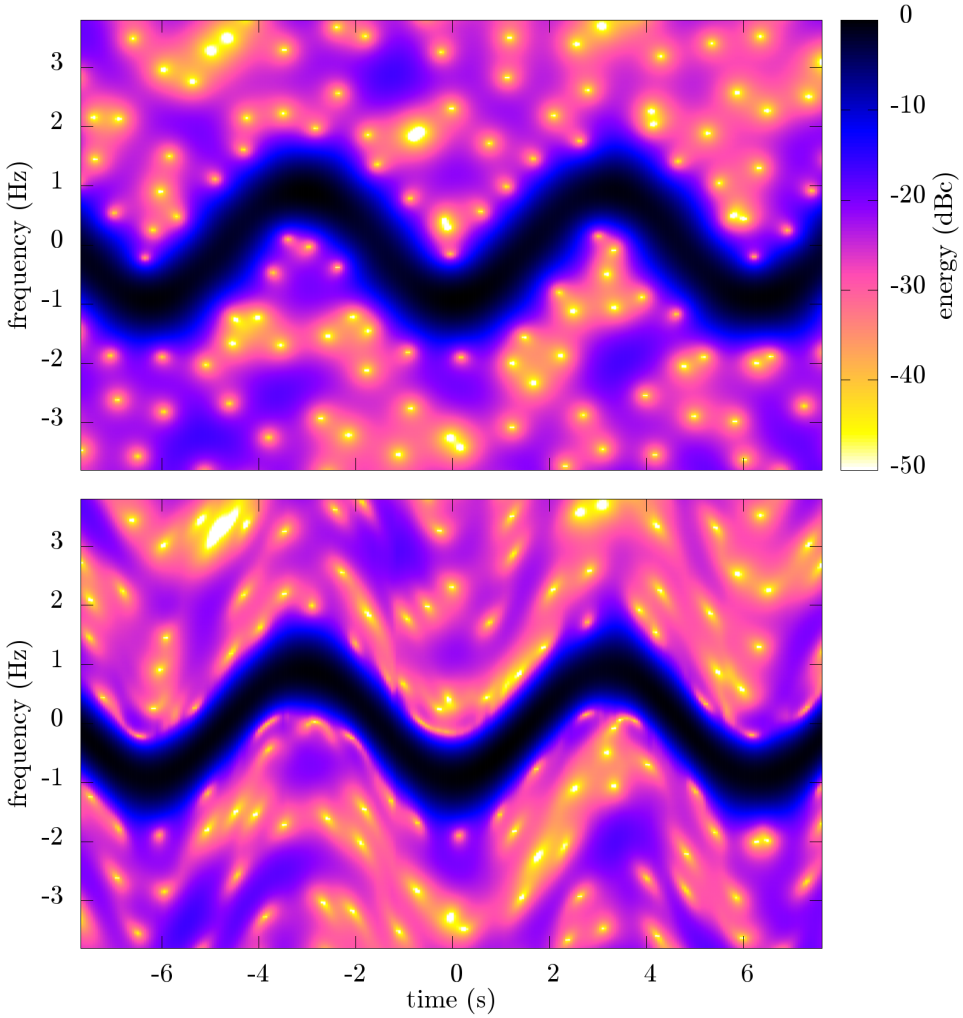

Figure 4: Classical energy spectrograms computed using STFT (upper) and STCT (lower) of signal defined by Eq. (17) in the presence of AWGN (SNR is 10 dB, sampling rate is 20 Sa/s).

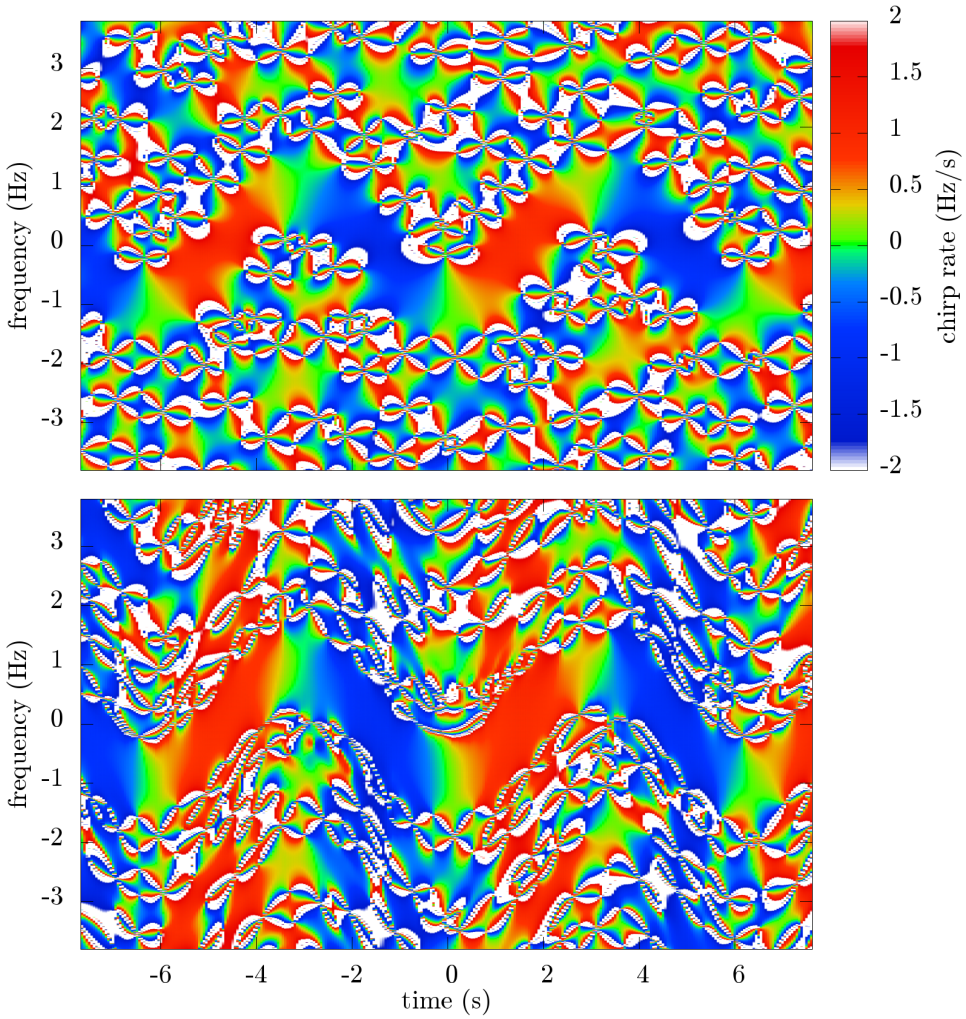

Figure 5: Time-frequency phase accelerograms computed using STFT (upper) and STCT (lower) of signal defined by Eq. (17) in the presence of AWGN (SNR is 10 dB, sampling rate is 20 Sa/s).

3.2 Application on real-world LFM sonar signal analysis

In this section, a real-world LFM signal recorded in shallow water of Baltic Sea is analyzed. This signal is initially produced by an active sonar developed in the Department of Marine Electronics Systems, Gdansk University of Technology [37, 38]. The analyzed signal is reflected from a real submarine during a military exercise close to the Polish coast. This is the reason, why the original parameters such as sampling rate, frequency band, pulse width, etc. which are confidential have been modified. The corresponding quadrature demodulated complex envelope is analyzed and, finally, its spectral energy is distributed around zero on the frequency axis in the presented spectrogram in Fig. 7. In this experiment, we estimate the energy distribution as well as the accelerogram used for the evaluation of statistical relations between these two representations (*cf.* Fig. 8). A

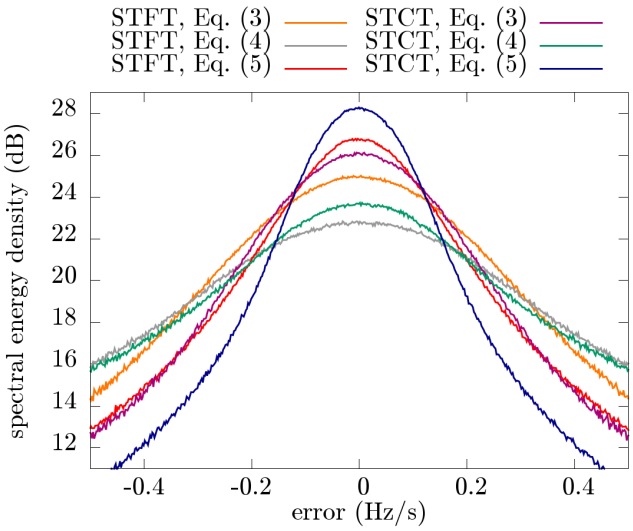

Figure 6: Relations between the error of estimated CR and the spectral energy density for different estimators for the window width equal to $T_{\text{win}} = 2.75$ s. The investigated signal is defined by Eq. (17) in the presence of AWGN (SNR is 10 dB, sampling rate is 20 Sa/s).

two-dimensional median filter with an empirically chosen kernel size was applied to the computed accelerogram in order to smooth the displayed results.

In Fig. 8, the curve which shows the relation between the observed energy and the estimated CR is presented. A significant local maximum in this curve is visible close to -10 kHz/s (inside a drawn circle), which is near to the original parameter of sounding LFM pulse signal. This value is assumed to be the window CR α in the next stage of the analysis, which involves the second time-frequency analysis. Now, our goal is to obtain a higher energy concentration in the spectrogram, which is necessary for a correct detection of submerged objects.

In Fig. 9, the comparison between curves which express for a given transform using a window parameterized by a fixed CR, the part of the corresponding energy density provided by its resulting TF representation. The highest percentage is reached by the STCT-based analysis for a CR equal to -10 kHz/s which is dominant for high energy density (bright red background). This confirms that for this CR, the energy concentration in the final spectrogram is the highest. In contrast, for window CRs equal to 10 and -30 kHz/s (black and blue lines), which are mismatched, the smearing of the energy is also clearly visible from this experiment.

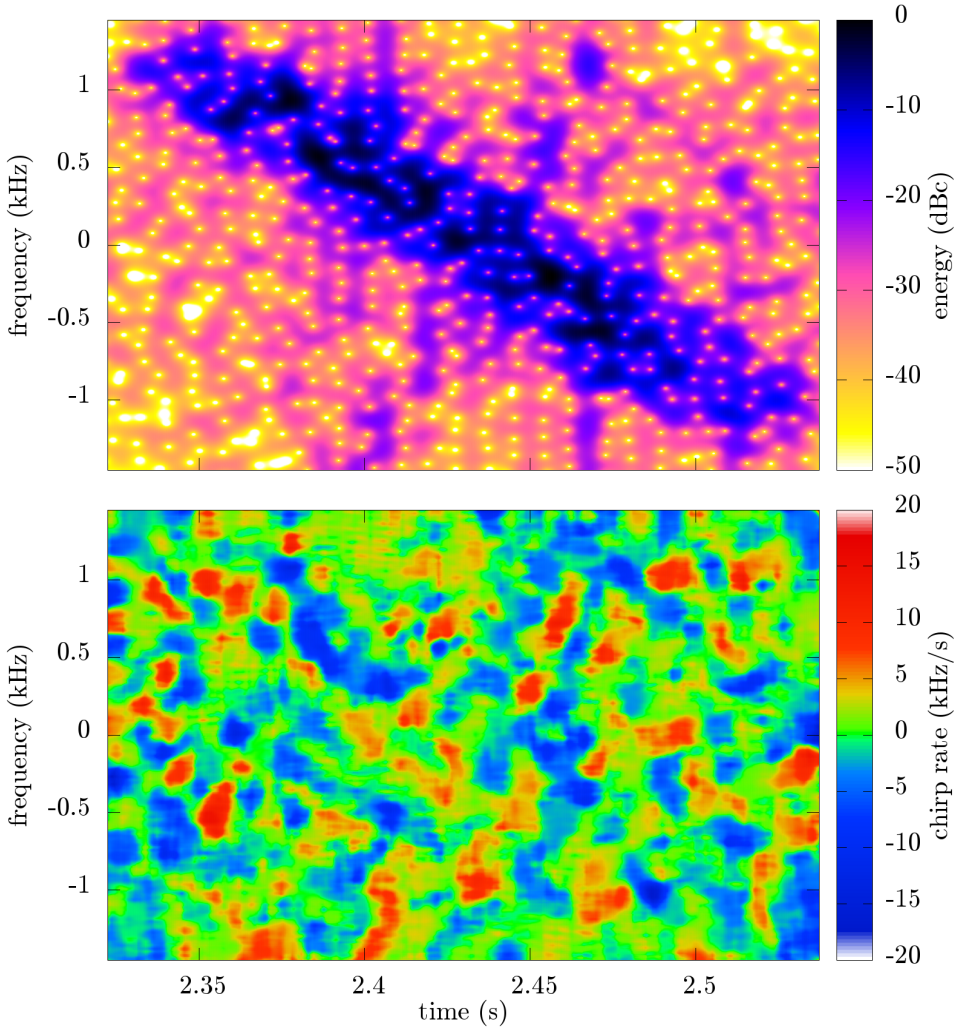

Figure 7: Classical STFT-based spectrogram (upper) and time-frequency phase accelerogram smoothed by a two-dimensional median filter (lower). The investigated signal is a reflection of a LFM sonar chirp from submarine recorded in shallow water of Baltic Sea during military exercise close to the Polish coast.

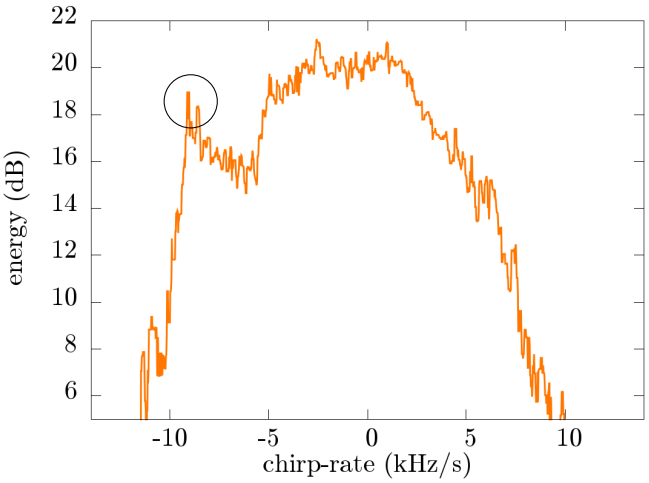

Figure 8: Statistical relations between estimated CR and spectral energy density. The investigated signal is a reflection of a LFM sonar chirp from submarine recorded in shallow water of Baltic Sea during military exercise close to the Polish coast.

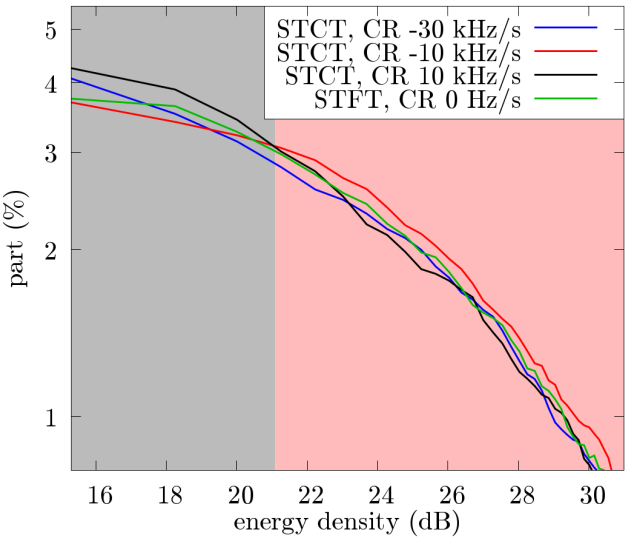

Figure 9: Statistical analysis of the energy density as a percentage share of spectrograms based on STFT and STCT. The investigated signal is a reflection of a LFM sonar chirp from submarine recorded in shallow water of Baltic Sea during military exercise close to the Polish coast.

4 Multi-window chirplet transform

Multi-window chirplet transform (MWCT) is proposed for the analysis of signals whose components are simultaneously characterized by a finite number of clearly different CRs. The procedure of this method begins from the creation of a vector of N representative CRs:

$$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_N), \tag{20}$$

where $\alpha_1 < \alpha_2 < ... < \alpha_N$. We assume that the range between α_1 and α_N covers the values which are associated with the considerable components of an analyzed signal. The accelerogram computed through the STFT can also be used to define this range (*e.g.* linearly spaced values in range $[\min(\hat{R}_x^g(t,\omega)), \max(\hat{R}_x^g(t,\omega))]$ can be chosen). In the remainder, all bold symbols (math font style) denote vectors, thus the corresponding TF functions (distributions and parameters) can now be considered as three-dimensional tensors. Then, CTs are computed for each CR present in vector $\boldsymbol{\alpha}$ using the chirped analysis windows defined by Eq. (2):

$$g_i(t,\omega) = h(t)\mathrm{e}^{\mathrm{j}\alpha_i \frac{t^2}{2}}\mathrm{e}^{\mathrm{j}\omega t}, \text{ for } i = 1, 2, ..., n. \tag{21}$$

Therefore, we can obtain a tensor made of these TF transforms denoted by:

$$\boldsymbol{y}_x^{\boldsymbol{g}}(t,\omega) = (y_x^{g_1}(t,\omega), y_x^{g_2}(t,\omega), ..., y_x^{g_N}(t,\omega)), \tag{22}$$

which are used to estimate the corresponding energy distributions:

$$\boldsymbol{E}_x^{\boldsymbol{g}}(t,\omega) = (E_x^{g_1}(t,\omega), E_x^{g_2}(t,\omega), ..., E_x^{g_N}(t,\omega)), \tag{23}$$

instantaneous frequencies:

$$\boldsymbol{\Omega}_x^{\boldsymbol{g}}(t,\omega) = (\hat{\Omega}_x^{g_1}(t,\omega), \hat{\Omega}_x^{g_2}(t,\omega), ..., \hat{\Omega}_x^{g_N}(t,\omega)), \tag{24}$$

spectral delays:

$$\boldsymbol{D}_x^{\boldsymbol{g}}(t,\omega) = (\hat{D}_x^{g_1}(t,\omega), \hat{D}_x^{g_2}(t,\omega), ..., \hat{D}_x^{g_N}(t,\omega)), \tag{25}$$

and, TF accelerograms:

$$\boldsymbol{R}_x^{\boldsymbol{g}}(t,\omega) = (\hat{R}_x^{g_1}(t,\omega), \hat{R}_x^{g_2}(t,\omega), ..., \hat{R}_x^{g_N}(t,\omega)) \tag{26}$$

both computed according to aforementioned Eqs. (3), (6), (7), (10), (11), and (14), where

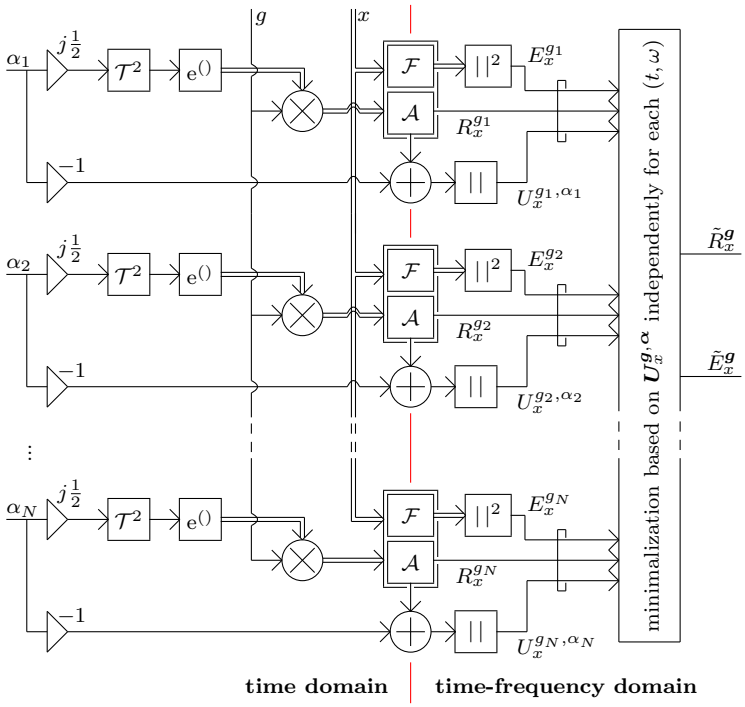

Figure 10: Block scheme of multi-window chirplet spectrograph and accelerograph. The legend for this diagram is presented in Fig. 1.

$\boldsymbol{g}$ denotes the vector of the used windows. Subsequently, a new measure of unmatching $\boldsymbol{U}_x^{\boldsymbol{g}}(t,\omega)$ can be defined based on the vectors $\boldsymbol{\alpha}$ and $\boldsymbol{R}_x$ as follows:

$$\boldsymbol{U}_x^{\boldsymbol{g},\boldsymbol{\alpha}}(t,\omega) = |\boldsymbol{R}_x^{\boldsymbol{g}}(t,\omega) - \boldsymbol{\alpha}| = (|\hat{R}_x^{g_1}(t,\omega) - \alpha_1|, |\hat{R}_x^{g_2}(t,\omega) - \alpha_2|, ..., |\hat{R}_x^{g_N}(t,\omega) - \alpha_N|). \quad (27)$$

Each element of $\boldsymbol{U}_x^{\boldsymbol{g},\boldsymbol{\alpha}}$ is a TF distribution defined as the absolute value of the difference between each expected CR α_i and the TF accelerogram computed using its corresponding chirped analysis window. Finally, a new energy distribution can be derived from $\boldsymbol{E}_x^{\boldsymbol{g}}$ (combination from its elements) according to the following formula:

$$\tilde{E}_x^{\boldsymbol{g}}(t,\omega) = E_x^{g_k}(t,\omega), \text{ with } k = \underset{i}{\operatorname{argmin}} |\hat{R}_x^{g_i}(t,\omega) - \alpha_i| \text{ and } i = 1,2,...,N. \quad (28)$$

The same approach as for the energy can be applied to derive other parameters, such as IF Ω_x, SD D_x, and CR R_x, respectively:

$$\tilde{\Omega}_x^{\boldsymbol{g}}(t,\omega) = \hat{\Omega}_x^{g_k}(t,\omega), \text{ with } k = \underset{i}{\operatorname{argmin}} |\hat{R}_x^{g_i}(t,\omega) - \alpha_i| \text{ and } i = 1,2,...,N, \quad (29)$$

$$\tilde{D}_x^{\boldsymbol{g}}(t,\omega) = \hat{D}_x^{g_k}(t,\omega), \text{ with } k = \underset{i}{\operatorname{argmin}} |\hat{R}_x^{g_i}(t,\omega) - \alpha_i| \text{ and } i = 1,2,...,N, \quad (30)$$

as well as

$$\tilde{R}_x^{\boldsymbol{g}}(t,\omega) = \hat{R}_x^{g_k}(t,\omega), \text{ with } k = \underset{i}{\operatorname{argmin}} |\hat{R}_x^{g_i}(t,\omega) - \alpha_i| \text{ and } i = 1,2,...,N. \quad (31)$$

Moreover, a modified time-frequency reassignment method can also be applied such as for the STFT through Eq. (13), using these new estimated IFs and SDs. A diagram of this procedure in presented in Fig. 10 for the energy and CR estimation.

4.1 LFM signal detection in AWGN

In this section we consider the detection of LFM pulses in additive white Gaussian noise (AWGN) channel for signal-to-noise ratio (SNR) levels which are equal to 0 and 10 dB (2 cases). This testing signal consists of two pulses which occur simultaneously. Their CRs are equal to, respectively, $r_1 = 0.5$ kHz/s and $r_2 = 1$ kHz/s and their frequencies are $f_1 = 200$ Hz and $f_2 = 400$ Hz. Moreover, each pulse is weighted by a Tukey window [29] whose width is equal to 0.3 s.
This signal definition can be expressed by the following formula:

$$x_{\text{2LFM}}(t) = w_{\text{Tukey}}(t)\exp(\text{j}2\pi f_1 t + \text{j}\pi r_1 t^2) + w_{\text{Tukey}}(t)\exp(\text{j}2\pi f_2 t + \text{j}\pi r_2 t^2) + b(t), \quad (32)$$

where $w_{\text{Tukey}}(t)$ is the envelope of a Tukey window and b(t) is a white Gaussian noise. The detection is based on the statistical relations between the resulting spectral energy and the accelerograms corresponding to local maximums in the energy-CR profiles. In Fig. 11, detection results for a series of LFM chirp signals in AWGN channel are presented. First of all, we can observe that using a wide window (length of 1 s) to compute a single STFT leads to poor results (see the red curves). This situation is similar when a mismatched chirplet transform is used (not presented here). Some improvements can be observed for the use of a short window (length of 0.1 s) represented by the blue curves. For this variant, we can observe a satisfying detection for a SNR equal to 10 dB (see the cyan circles). However, for a low SNR equal to 0 dB, we miss one of the components. Finally, the use of the multi-window chirplet transform (MWCT) allows excellent detection even for SNR equal to 0 dB (see the gray circles). For this experiment, $N = 151$ uniformly spaced CR values in range $[0, 1500]$ Hz (resp. $[-1500, 0]$ Hz) have been used to compute the MWCT.

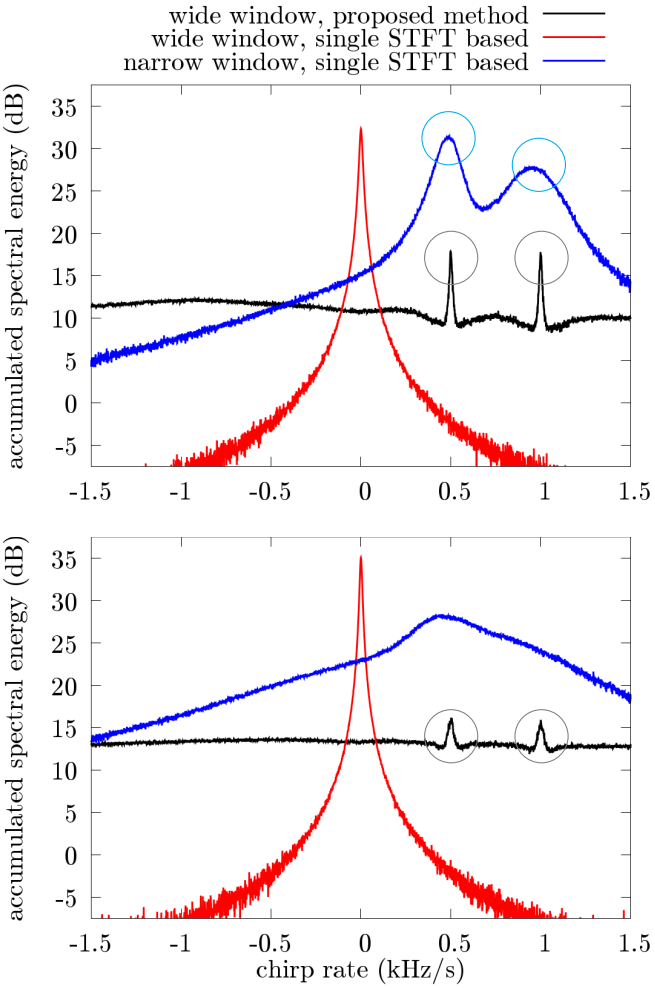

Figure 11: Averaged relation between CR and spectral energy of two-component LFM signal in the presence of AWGN, where SNR is equal to 0 (lower) and 10 dB (upper). The proposed multi-window method and the method based on a single STFT are compared. Successful detections are indicated by circles.

4.2 TF localization performance for LFM signal in AWGN

In this section, we analyze the degrading effect of AWGN in terms of TF localization when the reassignment method is used. We consider a LFM chirp sampled at 1000 Sa/s whose CR is constant and equal to $r_1 = 1$ kHz/s with an initial frequency equal to 0 Hz. This signal has a constant amplitude equal to 1. At the beginning, we estimate its energy distribution in the TF domain based on single- or multi-window STFT and we look for the energy maximums at each instant. Their arguments as a function of time denoted by $\omega_{\max}(t)$ are Fourier frequencies which should be located close to a signal attractor in the TF domain if a noise density is relatively weak. The attractor is defined by the following

linear relation: $f = r_1 t$. Then, in the obtained points $\big(t, \omega_{\max}(t)\big)$, we estimate IF and SD using, respectively, Eqs. (10) and (11). Finally, the performance is expressed by the root mean squared error (RMSE) expressed in hertz, as follows:

$$\xi = \frac{1}{2\pi}\sqrt{\frac{1}{M}\sum_{m=0}^{M}\Big(\hat{\Omega}\big(t_m, \omega_{\max}(t_m)\big) - 2\pi r_1 \hat{\Gamma}\big(t_m, \omega_{\max}(t_m)\big)\Big)^2}, \tag{33}$$

where $m = 1, 2, ..., M$ is an sampling index. In Fig. 12, we present results of the estimation of ξ comparing the proposed MWCT method to the single STFT-based approach for 6 different windows whose widths are equalto 50 ms, 100 ms, 175 ms, 300 ms, 425 ms, and 550 ms. These curves are functions of the SNR. We analyze this signal in the time range between $T_{\min} = -0.5$ s and $T_{\max} = 0.5$ s using $N = 321$ CRs non-uniformly distributed in the range between -5.6713 kHz/s and 5.6713 kHz/s. This set is defined as follows: $\{r_1, r_2, ..., r_N\}$, where

$$r_n = r_{\mathrm{u}} \tan\left(\pi \frac{n - 81}{180}\right), \tag{34}$$

and $r_{\mathrm{u}} = 1$ Hz/s.

Looking at Fig. 12, the curves for STFT are shaped by the Heisenberg-Gabor Uncertainty Principle. From these curves, we can point the best (quasi-optimal) window width which is 100 ms and which has the smallest RMSE in every SNR point. The windows which are narrower or wider are characterized by higher RMSE values. This is the classic situation in the TF analysis. For each of these curves, we observe 2 ranges: anormal (type II error) and normal (type I error, where straight dashed line is observed), respectively, for low and high SNRs. However, an additional issue influences on the results for the proposed MWCT method. Above all, in this case, we observe the significat improvment for any window and for both estimators. The Heisenberg-Gabor Uncertainty Principle still plays an important role, however we can notice that the adaptation works well. In general, the curves for the proposed MWCT method can be divided into 4 parts:

- for low SNR and for all widths – anormal range – method does not work at all,
- for moderate SNR and for all widths – normal range – the adaptation works well, however its impact is limited by the inaccuracy of chirp-rate estimation,
- for high SNR and for wide widths – transitional range – the adaptation works still better and its impact increases comparatively to other effects,
- for very high SNR and for wide widths – "super" range – the adaptation works very well and its impact is constant comparatively to other effects.

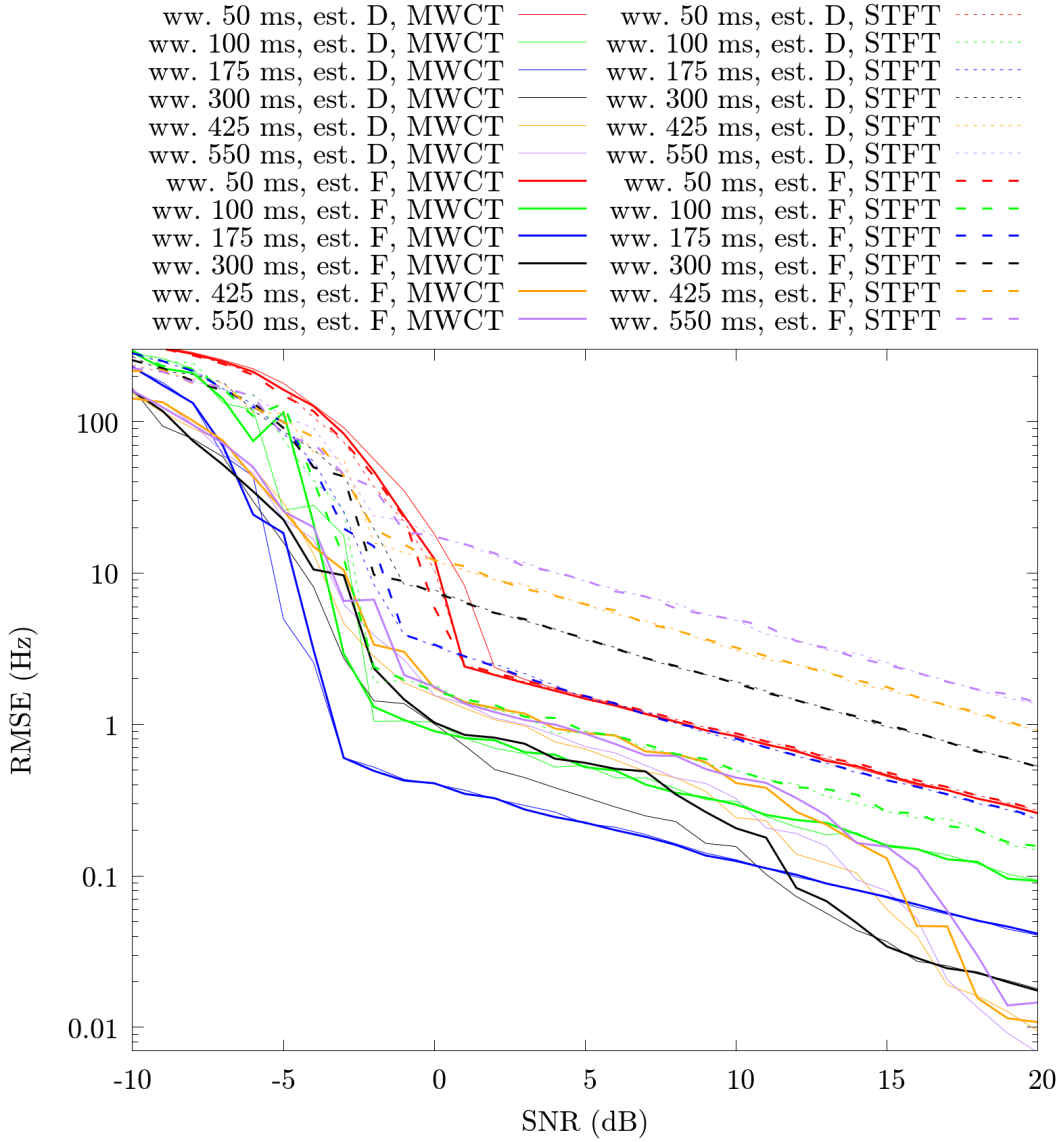

Figure 12: Comparison of the proposed multi-window method (MWCT) to the single STFT-based approach (STFT) for 2 CR estimators ("D" and "F") and for 6 different windows, whose widths are equal to: 50, 100, 175, 300, 425, and 550 ms, as a function of SNR. The infinite LFM chirp, whose CR is equal to 1 kHz/s, is analyzed.

These parts are clearly visible for the window whose width is equal to 300 ms (black curve; section points are localized on the SNR axis near to: -1, 9, and 15 dB). We can also observe that the adaptation does not work for very narrow windows, however its positive impact is clear and significant for the optimal window in the case of the classical STFT. This gives the practical meaning of the proposed method.

In Figs. 13 and 14, estimated TF energy distributions using a single STFT and the proposed MWCT method are presented. We analyze a AM FM test signal which is defined

as follows:

$$x(t) = \frac{1}{t_{\text{sec}}}(t - t_0)\exp(\text{j}\pi r_0 t^2), t > -0.5, \tag{35}$$

where $t_0 = -0.5$ s, $r = 1$ kHz/s, and $t_{\text{sec}} = 1$ s. The signal is analyzed in the presence of moderate AWGN. We present energy distributions before (Fig. 13) and after (Fig. 14) TF reassignment. In our subjective opinion, the proposed MWCT method does not clearly improve the readability, however it can be used to find true TF localization of local ridge maximums (attractor).

4.3 Analysis of the vibrato performed by a flute

We have conducted the time-frequency analysis of the vibrato performed by a flute. Its results are presented in Figs. 15 and 16 including an energy distribution and phase

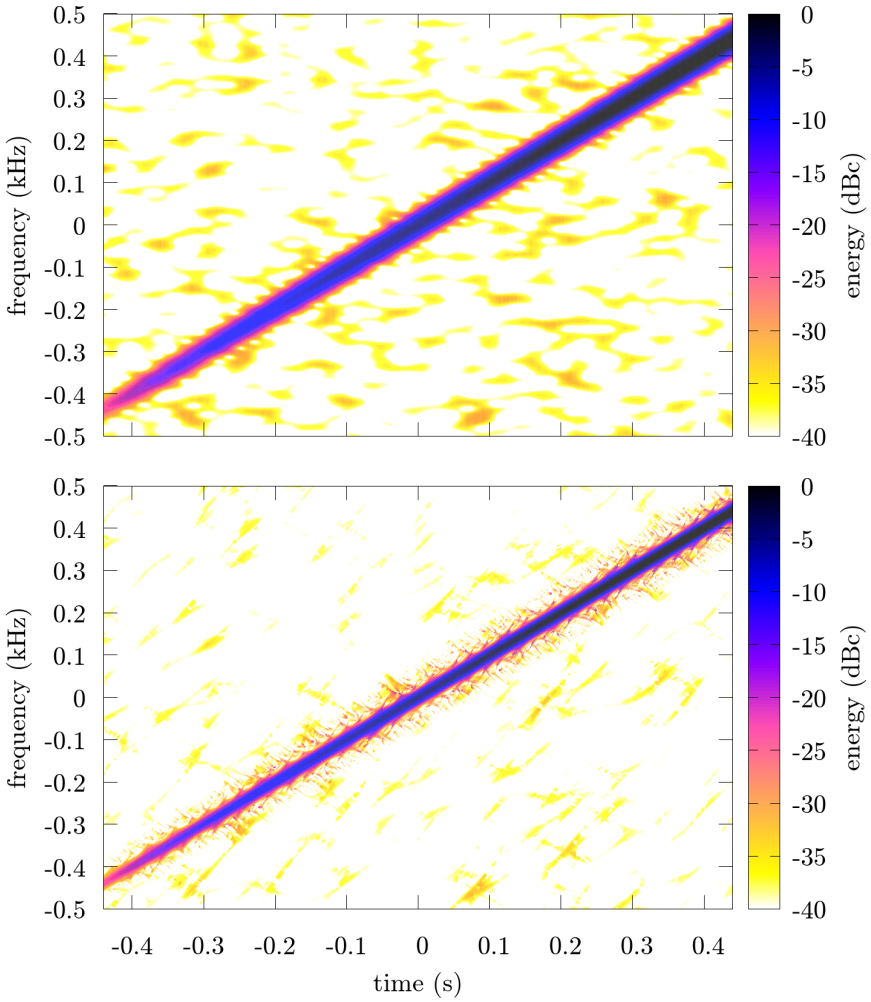

Figure 13: Classical single STFT-based (upper) and proposed multi-window (lower) spectrograms of AM FM test signal in moderate AWGN (SNR = 10 dB).

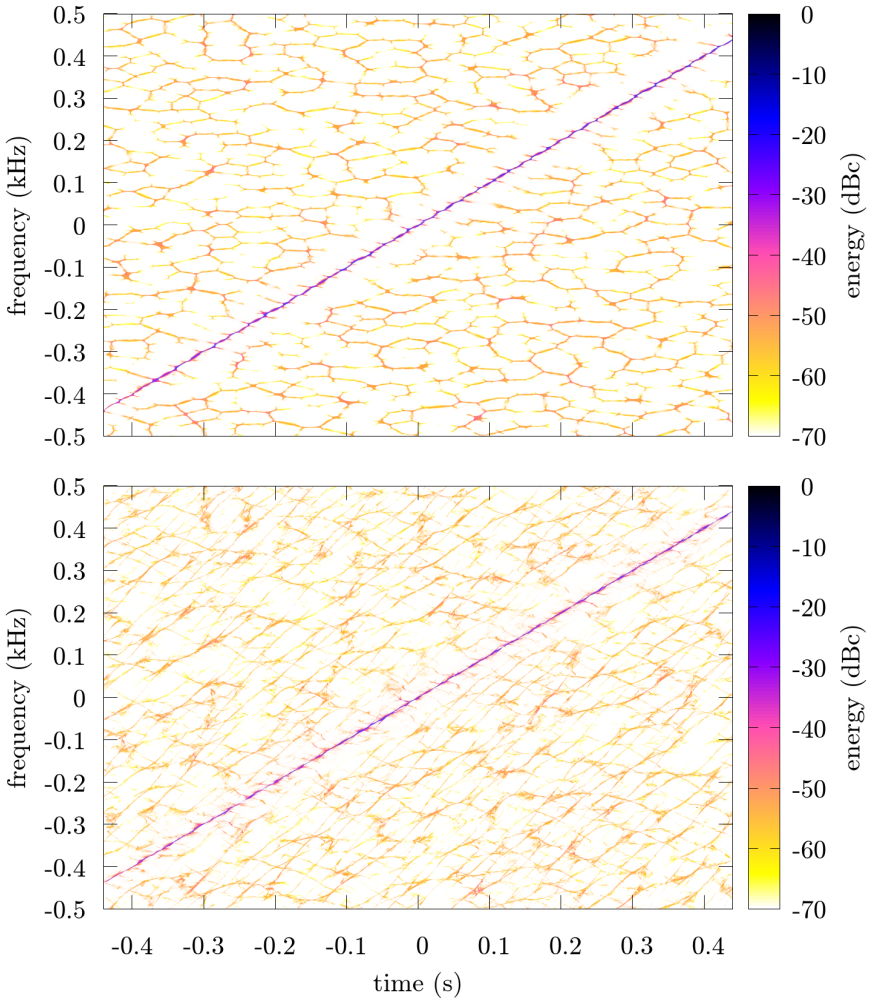

Figure 14: Reassigned single STFT-based (upper) and proposed multi-window (lower) spectrograms of AM FM test signal in moderate AWGN (SNR = 10 dB).

accelerograms using a single STFT as well as using the proposed multi-window method. The energy distribution is estimated by a single STFT. The window width is equal to 125 ms. The selected analyzed signal contains components which are characterized by positive, neutral, and negative values of the CR. They occur alternately (vibrato). We can observe that the introduced multi-window method allows to estimate the CR for an audio signal.

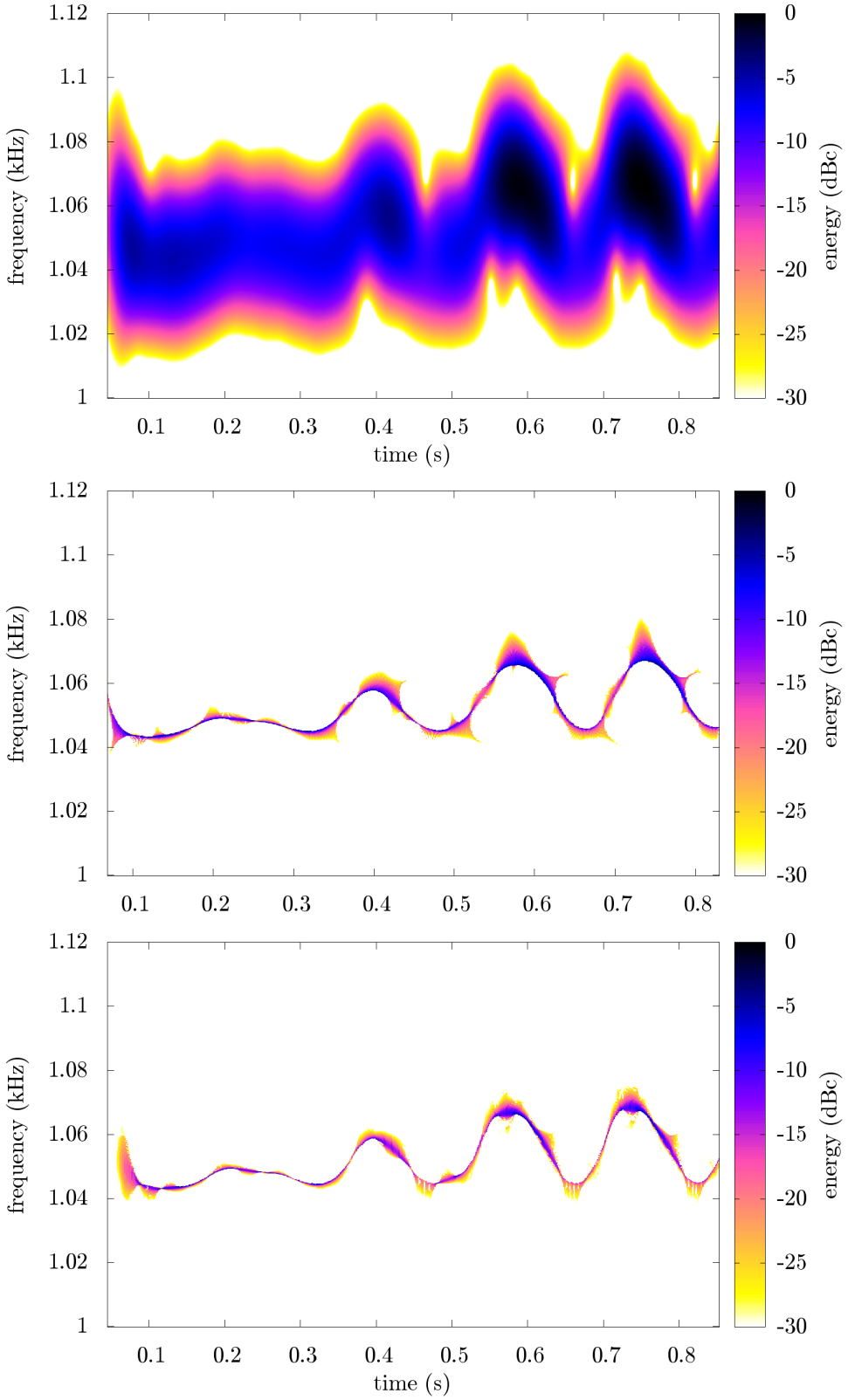

Figure 15: Single STFT-based spectrogram (top), single STFT-based reassigned spectrogram (middle) and multi-window reassigned energy distribution (bottom) of a vibrato performed by a flute. A slight improvement of the TF localization using the multi-window reassignment method is visible in comparison with the single STFT-based version.

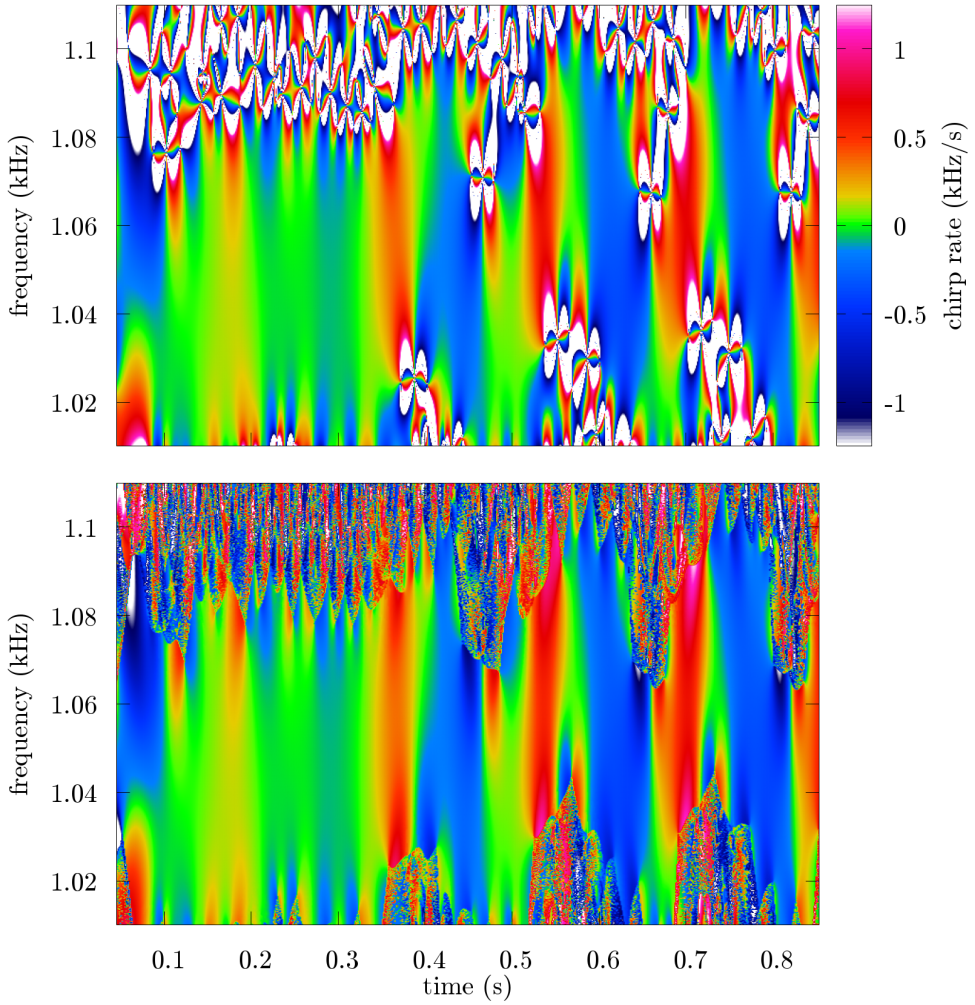

Figure 16: Time-frequency phase accelerograms of the vibrato performed by a flute STFT-based (upper) and estimated using multi-window method (lower).

5 Conclusion

We have proposed to use estimated CRs in the time-frequency domain (TF accelerogram) in order to locally adapt the analysis window used to compute improved TF representations. In the presented experiments, the tapering Blackman-Harris window was modulated and used as a directional kernel. We have considered this approach as a special case of an adaptive chirplet transform. We have defined three CR estimators for this purpose. We have shown that the introduced approach can improve the performance of the TF attractor localization through TF reassignment as well as it significantly increases the possibility of LFM chirp detection whose CR is *a priori* unknown. However, we do not notice any significant improvement in the TF representation readability. This problem

remains to be further investigated.

We have introduced two adaptive CT algorithms which can be efficiently implemented by FFT – self-tuning and multi-window. We have provided an implementation of these methods as the open source ccROJ project [30], which is C++ framework for the time-frequency analysis. Our proposed implementations use an elegant mathematical formulation of several TF signal parameters estimators through specific analysis windows. This results in a negligible increase of the computation run-time complexity as expressed in Table 5 in terms of "Big O" notation. We intend to further develop this project. Then, we have shown the practical significance of the proposed approach in applications such as signal detection in the time-frequency domain and by using CR versus spectral energy dependence.

Table 1: Run-time complexity in units of time for the proposed methods. T denotes the number of time instants where a FFT of length F is computed to obtain a TF representation (STFT or CT). Our new proposed multi-window method is computed through $n \ll F$ distinct CTs (in this paper N has the largest value equal to 321).

Subroutine	Arithmetic operations in units of time
FFT	$\mathcal{O}(F\log(F))$
STFT or CT	$\mathcal{O}(TF\log(F))$
IF, SD or CR Eq. (3) estimation	$3\mathcal{O}(TF\log(F))$
CR Eq. (6) or (7) estimation	$5\mathcal{O}(TF\log(F))$
STCT (compute a CR and a CT)	$8\mathcal{O}(TF\log(F))$ or $12\mathcal{O}(TF\log(F))$
MWCT (compute N CRs and a CT)	$4\mathcal{O}(NTF\log(F))$ or $6\mathcal{O}(NTF\log(F))$

References

[1] P. Flandrin, Time-Frequency/Time-Scale Analysis, Wavelet Analysis and Its Applications, Elsevier Science, 1998.

[2] B. Boashash, Time-frequency signal analysis and processing: a comprehensive reference, Academic Press, 2015.

[3] M. Kepesi, L. Weruaga, Adaptive chirp-based time-frequency analysis of speech signals, Speech Communication 48 (5) (2006) 474–492.

[4] L. Weruaga, M. Kepesi, The fan-chirp transform for non-stationary harmonic signals, Signal Processing 87 (6) (2007) 1504–1522.

[5] M. Bartkowiak, Application of the fan-chirp transform to hybrid sinusoidal+noise modeling of polyphonic audio, in: 2008 16th European Signal Processing Conference, 2008, pp. 1–5.

[6] L. Angrisani, M. D'Arco, A measurement method based on an improved version of the chirplet transform for instantaneous frequency estimation, in: Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference. Rediscovering Measurement in the Age of Informatics, Vol. 2, 2001, pp. 1123–1129.

[7] N. A. Khan, B. Boashash, Instantaneous frequency estimation of multicomponent nonstationary signals using multiview time-frequency distributions based on the adaptive fractional spectrogram, IEEE Signal Processing Letters 20 (2) (2013) 157–160.

[8] B. Boashash, S. Ouelha, An improved design of high-resolution quadratic time-frequency distributions for the analysis of nonstationary multicomponent signals using directional compact kernels, IEEE Transactions on Signal Processing 65 (10) (2017) 2701–2713.

[9] M. Brajovic, V. Popovic-Bugarin, I. Djurovic, S. Djukanovic, Post-processing of time-frequency representations in instantaneous frequency estimation based on ant colony optimization, Signal Processing 138 (2017) 195–210.

[10] M. A. Awal, S. Ouelha, S. Dong, B. Boashash, A robust high-resolution time–frequency representation based on the local optimization of the short-time fractional fourier transform, Digital Signal Processing 70 (Supplement C) (2017) 125 – 144.

[11] S. Mann, S. Haykin, The chirplet transform: physical considerations, IEEE Transactions on Signal Processing 43 (11) (1995) 2745–2761.

[12] K. Czarnecki, D. Fouan, Y. Achaoui, S. Mensah, Fast bubble dynamics and sizing, Journal of Sound and Vibration 356 (2015) 48–60.

[13] K. Czarnecki, The instantaneous frequency rate spectrogram, Mechanical Systems and Signal Processing 66–67 (2016) 361–373.

[14] D. Fourer, F. Auger, K. Czarnecki, S. Meignen, P. Flandrin, Chirp rate and instantaneous frequency estimation: application to recursive vertical synchrosqueezing, IEEE Signal Processing Letters 24 (11) (2017) 1724–1728.

[15] S. G. Mallat, Z. Zhang, Matching pursuits with time-frequency dictionaries, IEEE Transactions on Signal Processing 41 (12) (1993) 3397–3415.

[16] P. L. Shui, Z. Bao, H. T. Su, Nonparametric detection of FM signals using time-frequency ridge energy, IEEE Transactions on Signal Processing 56 (5) (2008) 1749–1760.

[17] S. Chikkerur, A. N. Cartwright, V. Govindaraju, Fingerprint enhancement using STFT analysis, Pattern Recognition 40 (1) (2007) 198–211.

[18] P. M. Djuric, S. M. Kay, Parameter estimation of chirp signals, IEEE Transactions on Acoustics, Speech, and Signal Processing 38 (12) (Dec 1990) 2118–2126.

[19] S. Barbarossa, Analysis of multicomponent LFM signals by a combined wigner-hough transform, IEEE Transactions on Signal Processing 43 (6) (1995) 1511–1515.

[20] J. C. O'Neill, P. Flandrin, Chirp hunting, in: Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, 1998, pp. 425–428.

[21] X.-G. Xia, Discrete chirp-Fourier transform and its application to chirp rate estimation, IEEE Transactions on Signal Processing 48 (11) (2000) 3122–3133.

[22] Y. Pantazis, O. Rosec, Y. Stylianou, Chirp rate estimation of speech based on a time-varying quasi-harmonic model, in: 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, 2009, pp. 3985–3988.

[23] I. Djurovic, Viterbi algorithm for chirp-rate and instantaneous frequency estimation, Signal Processing 91 (5) (2011) 1308–1314.

[24] K. Kodera, C. de Villedary, R. Gendrin, A new method for the numerical analysis of non-stationary signals, Physics of the Earth and Planetary Interiors 12 (1976) 142–150.

[25] F. Auger, P. Flandrin, Improving the readability of time-frequency and time-scale representations by the reassignment method, IEEE Transactions on Signal Processing 43 (5) (1995) 1068–1089.

[26] G. K. Nilsen, Recursive time-frequency reassignment, IEEE Transactions on Signal Processing 57 (8) (2009) 3283–3287.

[27] S. C. Pei, S. G. Huang, STFT with adaptive window width based on the chirp rate, IEEE Transactions on Signal Processing 60 (8) (2012) 4065–4080.

[28] N. A. Khan, M. Sandsten, Time–frequency image enhancement based on interference suppression in wigner–ville distribution, Signal Processing 127 (2016) 80–85.

[29] F. J. Harris, On the use of windows for harmonic analysis with the discrete Fourier transform, Proceedings of the IEEE 66 (1978) 51–83.

[30] F. Auger, K. Czarnecki, D. Fourer, ccROJ – time-frequency C++ framework, https://github.com/dsp-box/ccROJ, SoftwareX (submitted).

[31] D. Nelson, Cross-spectral methods for processing speech, The Journal of the Acoustical Society of America 110 (2001) 2575–2592.

[32] S. A. Fulop, K. Fitz, Algorithms for computing the time-corrected instantaneous frequency (reassigned) spectrogram, with applications, The Journal of the Acoustical Society of America 119 (2006) 360–371.

[33] R. Tao, Y.-L. Li, Y. Wang, Short-time fractional Fourier transform and its applications, IEEE Transactions on Signal Processing 58 (5) (2010) 2568–2580.

[34] D. Fourer, F. Auger, P. Flandrin, Recursive versions of the levenberg-marquardt reassigned spectrogram and of the synchrosqueezed STFT, in: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2016, pp. 4880–4884.

[35] F. Auger, P. Flandrin, Y. T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, H. T. Wu, Time-frequency reassignment and synchrosqueezing: An overview, IEEE Signal Processing Magazine 30 (6) (2013) 32–41.

[36] D. Fourer, J. Harmouche, J. Schmitt, T. Oberlin, S. Meignen, F. Auger, P. Flandrin, The ASTRES toolbox for mode extraction of non-stationary multicomponent signals, in: Proc. EUSIPCO, Kos island, Greece, 2017, pp. 1170–1174.

[37] J. Marszal, Implementation of contemporary technologies in the modernization of naval sonars, Hydroacoustics 16 (2013) 167–180.

[38] J. Marszal, Digital signal processing applied to the modernization of Polish Navy sonars, Polish Maritime Research 21 (2) (2014) 65–75.