The Smoothed Reassigned Spectrogram for Robust Energy Estimation

Erik M. Månsson
Acconeer AB
Scheelevägen 27
Lund, Sweden
erik@mansson.xyz

Maria Sandsten
Centre for Mathematical Sciences
Lund University
Lund, Sweden
maria.sandsten@matstat.lu.se

***Abstract*—The matched window reassigned spectrogram relocates all signal energy of an oscillating transient to the time- and frequency locations, resulting in a sharp peak in the time-frequency plane. However, previous research has shown that the method may result in split energy peaks for close components and in high noise levels, and the peak energy is then erroneously estimated. With use of novel knowledge on the statistics when subjected to noise, we propose a novel method, the smoothed reassigned spectrogram, for obtaining a stable and accurate measure of the signal energy from the peak value, with retained resolution properties. We also suggest a simple set of rules to enhance the reassigned spectrogram and speed up its calculation. Simulations are performed to verify the accuracy and an application example on radar data is shown.**

***Index Terms*—oscillating transient, time-frequency reassignment, reassignment vector statistics, smoothing kernels**

I. Introduction

Non-stationary stochastic processes and transient signals are all around us, from the changing temperature of our globe to the response from a radar system. The need for analyzing those kinds of signals arises practically everywhere, and is ever so challenging.

Usually, time-frequency (TF) representations serve as the common ground for further extraction of information from these signals. The research area of optimal TF kernels for different classes of signals and applications is extensive and well developed, where the suppression of cross-terms between TF components is of major interest. In recent years, most focus has been directed to instantaneous frequency estimation of linear and non-linear frequency modulated signals, which comprises all methods with aim to estimate time-varying frequencies. One such method is the *reassigned spectrogram* (RS), originally presented by Kodera et al. [1] and re-introduced by Auger and Flandrin [2], [3] with further developments in e.g. [4], [5].

The reassignment method exploits the phase information which typically is discarded in most kernel-based representations. Here, it is applied to increase the concentration of a single TF component by relocating mass to the local center of gravity. For multi-component signals, the reassignment improves the readability as cross-terms are reduced and the reassignment then squeezes the signal terms to be more localized. The method works well for long chirps as well as constant frequency signals.

For short multi-component oscillating transient signals we have previously invented a novel method, the matched window RS, for estimation of the instantaneous TF locations [6], [7]. In theory, all mass is perfectly localized in the TF plane. The estimates rely on a spectrogram window matching that of the envelope of the oscillating transient. It is therefore also suitable for parameter extraction of the envelope [8], [9]. The method has been applied for characterization of transient components within the echolocation beam of a dolphin [10], where the RS peak magnitude was used to detect components and then count them.

In the multi-component case, reassignment breaks down when the TF separation between signal components are too small – creating interference between them. Practically, knowledge of the derivatives of the reassignment vectors can be used to reduce the effect. Chassande-Mottin et al. have done some research on the topic, presented in [11] and a novel approach of using reassignment gradients to suppress interference of close components is presented in this submission.

For noisy signals, the estimates of the TF locations are often degraded, due to the reassignment vectors being sensitive to noise [8], [12]. One major problem is *peak splitting*, where the energy of the peak is split and spread over a larger area of the TF plane, resulting in erroneous RS peak magnitude estimates. A thorough statistical analysis of the reassignment vector sensitivity has been done by Chassande-Mottin et al. [13], [14] and more recently by Månsson in [12]. Based on this knowledge, an idea for a new method of resampling reassigned noisy spectrograms is proposed.

II. The matched window reassigned spectrogram

The short-time Fourier transform (STFT) of an oscillating transient signal on the form

$$x(t) = a(t - t_0) \cdot \exp(-i2\pi\nu_0 t) \tag{1}$$

is defined as

$$F^h(t,\nu) = \int x(s)\, h^*(s-t) \exp(-i2\pi\nu s)\, \mathrm{d}s \tag{2}$$

using a unit energy window function h. The resulting spectrogram is then defined as

$$S^h(t,\nu) = \left|F^h(t,\nu)\right|^2. \tag{3}$$

The reassigned time and frequency $(\hat{t}^h, \hat{\nu}^h)$ for a specific point (t,ν) in the spectrogram can be calculated from

$$\hat{t}^h(t,\nu) = t + c_t\tilde{t}^h,\ \tilde{t}^h = \mathrm{Re}\left\{\frac{F^{\mathcal{T}h}(t,\nu)}{F^h(t,\nu)}\right\} \tag{4}$$

$$\hat{\nu}^h(t,\nu) = \nu - c_\nu\tilde{\nu}^h,\ \tilde{\nu}^h = \frac{1}{2\pi}\,\mathrm{Im}\left\{\frac{F^{\mathcal{D}h}(t,\nu)}{F^h(t,\nu)}\right\} \tag{5}$$

where $F^{\mathcal{T}h}(t,\nu)$ and $F^{\mathcal{D}h}(t,\nu)$ are the STFTs of the signal $x(t)$ using $t\cdot h(t)$ and $dh(t)/dt$ as window functions respectively [2]. The reassigned spectrogram (RS) can then be expressed as

$$RS(t,\nu) = \iint S(t',\nu')\delta(t-\hat{t}(t',\nu'),\nu-\hat{\nu}(t',\nu'))\,\mathrm{d}t'\,\mathrm{d}\nu' \tag{6}$$

with

$$\iint f(t,\nu)\delta(t-\hat{t},\nu-\hat{\nu})dtd\nu = f(\hat{t},\hat{\nu}) \tag{7}$$

and where originally, $c_t = c_\nu = 1$ [1], [2]. However, for oscillating transients with an envelope matching the window, i.e., $h(t) = a(-t)$, $c_t = c_\nu = 2$ is the optimal choice [6], [7]. This matched window RS gives perfect localization of the total signal energy E to the signal time- and frequency locations (t_0,ν_0) of the noise free signal, i.e.,

$$RS(t,\nu) = E\cdot\delta(t-t_0,\nu-\nu_0). \tag{8}$$

Although the matched window RS works with any envelope/window, in this paper we constrain it to be Gaussian:

$$a(t) = \frac{1}{\pi^{1/4}\sqrt{\lambda}}\exp\left(-\frac{t^2}{2\lambda^2}\right), \tag{9}$$

and the corresponding matched window $h(t) = a(-t)$. We work with the assumption that a deterministic signal $x(t)$ is subjected to additive analytic (and circular) white Gaussian noise with variance σ_n^2, forming a noisy signal $y(t)$. The local signal to noise ratio (SNR) is defined as $\rho(t,\nu) = S^h(t,\nu)/2\sigma_n^2$ and the peak SNR is defined as $\rho_0 = \max_{t,\nu}\rho(t,\nu)$.

An example of a spectrogram containing two close components with time- and frequency locations $(75, 0.075)$ and $(100, 0.1)$, both with $\lambda = 10$ and peak SNR $\rho_0 = 15$ dB, is shown to the left in figure 1. The corresponding matched window RS is shown to the right. The spectrogram is not able to resolve into two separate components as clearly as the RS. However, the peak energy estimates of the RS are degraded as the peaks are split and spread out due to the noise.

Previous findings have shown that for $\rho(t,\nu) \gtrsim 10\,\mathrm{dB}$, the reassignment vector distribution is approximately Gaussian with an unbiased expectation value [8], [12]. Furthermore, for $\rho_0 \gtrsim 10\,\mathrm{dB}$ the RS distribution for a Gaussian signal with a scale λ may be approximated as a Gaussian with scale $\lambda/(0.6\cdot\sqrt{\rho_0})$, see section 4.5 in [12].

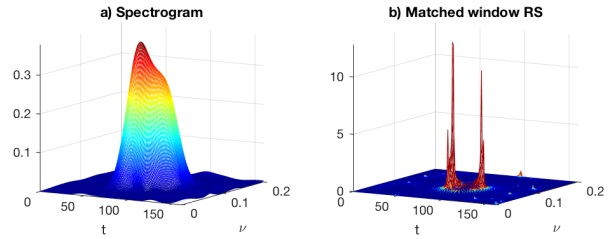

Fig. 1. a) Spectrogram of two close transient oscillating components with peak SNR, $\rho_0 = 15$ dB. b) The corresponding matched window RS.

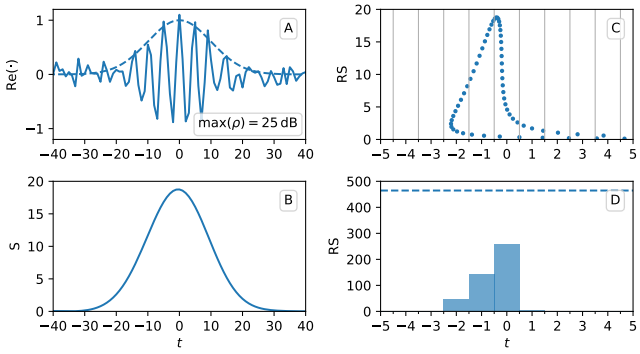

Fig. 2. An example of the matched window RS of a noisy Gaussian pulse. Plot A shows the real part of the signal, B shows the cross-section of the spectrogram, and C and D shows the raw and binned RS respectively. For comparison, the dashed line in D shows the total energy in the spectrogram.

III. The Smoothed Reassigned Spectrogram

In current literature, the reassigned spectrogram of a signal is often represented in one of two ways. In the first way, here referred to as the *raw* representation, the individual reassigned points of the spectrogram are shown. In the second way, here referred to as the *binned* representation, the reassigned spectrogram is naively resampled by assigning (binning) each point to the closest point in the *discrete-time* and *discrete-frequency* spectrogram. It is reasonable to assume that if the data is to be further processed, binning is preferred.

Both representations are shown in figure 2 for a single matched Gaussian pulse with $\lambda = 10$ and $\rho_0 = 25\,\mathrm{dB}$. To make it easier to illustrate, we only look at a cross-section of the spectrogram which is taken where the frequency $\nu_0 = 0.17$ matches that of the signal. Note that the two right plots (C and D) showing the RS representations are zoomed in by a factor of 8 in the time domain. Here we try to demonstate a typical example of the RS spreading and deteriorating when a signal is subjected to noise.

In this case, naive binning of the RS resulted in significant deterioration. This could perhaps be compensated with wider bins. But then, the localization of the peak would be subject to a larger quantization error. Also, no matter the bin width, if the signal center align between two bins, the energy is often split between the two. Thus, the binned RS is simply not robust to noisy signals.

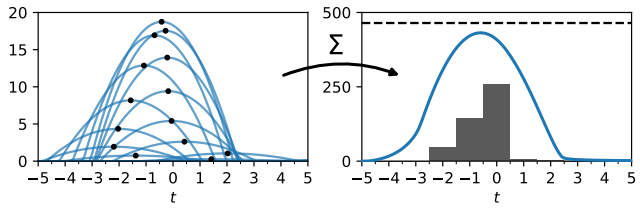

Fig. 3. An example of placing kernels on each reassigned point. The resulting SRS is shown in the right plot. For comparison, in the right plot, the bars show the binned RS, and the dashed lines shows the total energy.

A. Definition

We now propose a new method, the *smoothed reassigned spectrogram*, or SRS for short. The purpose is to be able to obtain a robust measure of the signal energy/power in its peak with retained resolution properties. The basic idea is to place a kernel on each reassigned point with some scale relating to that of the RS distribution. Essentially this means smoothing the reassigned points before resampling to obtain a well behaved version of the RS. This is conceptually similar to kernel density estimation or kernel smoothing.

In its simplest form, assuming a fixed kernel, the SRS may be expressed as

$$SRS(t,\nu) \triangleq \iint RS(t',\nu') \cdot K(t-t',\nu-\nu')\,\mathrm{d}t'\,\mathrm{d}\nu' \,. \quad (10)$$

Simply put, the SRS becomes a convolution between the kernel and the RS. However, in practice for the discrete case, we let the scale of individual kernels depend on the local SNR, but more on that later. Two choices of kernels were considered - a Gaussian G and a parabola P:

$$K^G(t,\nu) = \exp\left[-\frac{1}{2}\left(\frac{t^2}{\sigma_{K,t}^2}+\frac{\nu^2}{\sigma_{K,\nu}^2}\right)\right] \quad (11)$$

$$K^P(t,\nu) = \max\left[0,\, 1-\frac{1}{2}\left(\frac{t^2}{\sigma_{K,t}^2}+\frac{\nu^2}{\sigma_{K,\nu}^2}\right)\right] \quad (12)$$

where $\sigma_{K,t}$ and $\sigma_{K,\nu}$ are scaling parameters. Other choices are possible, but we choose to limit our work to these two kernels. Consider a Gaussian centered in (t_0,ν_0), where the noise-free RS is perfectly localized in the center point with the total signal energy E. That is, $RS(t,\nu) = E\cdot\delta(t-t_0,\nu-\nu_0)$. The SRS should retain that $SRS(t_0,\nu_0) = E$, which corresponds to the mass in $RS(t_0,\nu_0)$ for the noise-free case. Given this requirement, $K(0,0) = 1$. Figure 3 shows a cross-section example of placing such kernels on the reassigned points of the discrete spectrogram.

B. Kernel selection

Both the defined kernels – the Gaussian (11) and the parabola (12) – fulfill the requirement stated in the previous section. Also, the scaling factors $\sigma_{K,t}$ and $\sigma_{K,\nu}$ can be chosen such that the kernels become circularly symmetric in the normalized variables as shown in section 3.2 in [12].

We continue to approach this problem by studying the marginal distribution in time. For the matched window RS case, extension to both the time and frequency domain is then trivial since the distributions are circularly symmetric in normalized variables [12], [13]. With the approximations in chapter 4 in [12], we let the marginal distribution of the RS with unit energy and scale (std. dev.) be

$$RS(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right) \quad (13)$$

We define a Gaussian kernel K^G scaled with a factor a relative to the RS as

$$K^G(x) = \exp\left(-x^2/2a^2\right) \quad (14)$$

Here, the SRS is the convolution between the RS and kernel:

$$SRS(x) = \frac{1}{\sqrt{1+a^{-2}}}\cdot\exp\left(-\frac{x^2}{2\,(1+a^2)}\right) \quad (15)$$

$$= E_0^{SRS}\cdot\exp\left(-\frac{x^2}{2\,\sigma_{SRS}^2}\right) \quad (16)$$

As the kernel grows wider, i.e., as $a\to\infty$, then $E_0^{SRS}\to 1$ and asymptotically, $\sigma_{SRS} = \sqrt{1+a^2}\to a$. In other words, the energy is recovered, and the SRS width becomes that of the kernel. If we want to recover $95\,\%$ of the energy, we may pick $a\approx 3$. This gives an SRS scale of $\sqrt{1+a^2}\approx 3$. To clarify, if we use Gaussian kernels ~ 3 times wider than the estimated RS distribution scale, we expect to recover $95\,\%$ of the signal energy in the SRS peak.

A desirable kernel property is finite support, which makes the SRS practically feasible to compute. The parabola is often used in kernel smoothing since it is optimal in the sense that it minimizes spread [15]. In the context of the marginal distribution, we define the parabolic kernel as

$$K^P(x) = \begin{cases} 1-x^2/2a^2 & \text{for } x^2 < 2a^2 \\ 0 & \text{otherwise} \end{cases} \quad (17)$$

Note that it matches the first two terms of the Taylor series of the Gaussian kernel. One drawback compared to the Gaussian kernel is that the resulting SRS distribution will not be Gaussian. However, we can still calculate the expected energy recovery, E_0^{SRS}, numerically:

$$SRS(0) = \int_{-\sqrt{2}a}^{\sqrt{2}a}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)\cdot\left[1-\frac{x^2}{2a^2}\right]\mathrm{d}x \quad (18)$$

Recovering $95\,\%$ of the energy again gives a scale $a\approx 3$.

C. Practical considerations

So far we studied the method in continuous time and frequency, but practically, the SRS makes most sense in the discrete-time and discrete-frequency case, where it actually will be applied. We previously assumed a known RS distribution scale, giving a fixed scale kernel K, but in practice it has to be estimated in some way. In the matched Gaussian case, the approximate scale is given by the peak SNR, ρ_0 [12]. Since all points are ideally reassigned to the signal center (peak),

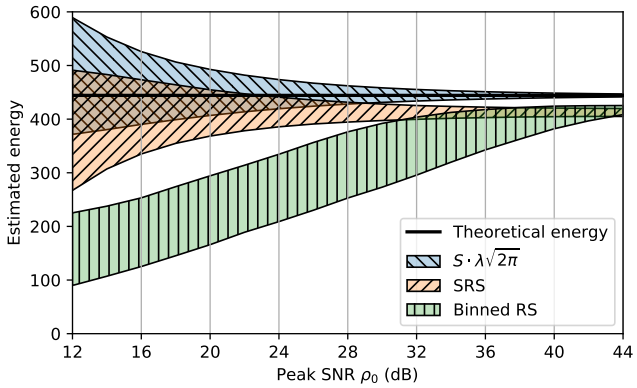

Fig. 4. The energy estimated from the peak the spectrogram, binned RS, and SRS. The span between the 25th and 75th percentile is shown. The thick line shows the theoretical noise-free signal energy.

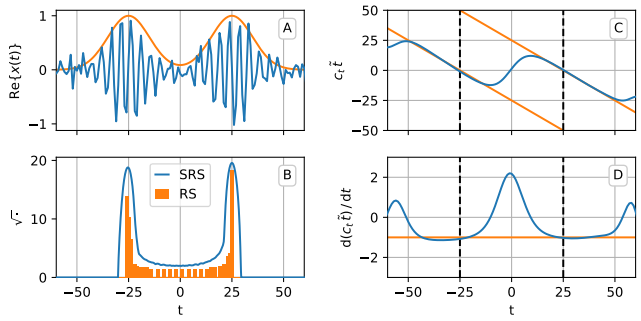

Fig. 5. An example of a signal with multiple close components - two Gaussians close in time. Plot A shows the signal and its envelope, B shows the RS and SRS, C shows the actual and ideal reassignment vector in time, and D shows the derivative of the reassignment vector in time.

we suggest using the local SNR at the point where all energy is reassigned, to estimate the distribution scale. That is, we obtain the kernel scale from $\rho(\hat{t}, \hat{\nu})$.

Now, we suggest a simple set of three basic criteria which a reassigned point must meet. These criteria have hard thresholds – either a point is included or discarded. This drastically speeds up the average computation time for the SRS.

Points with an SNR $\rho \lesssim 5\,\mathrm{dB}$ have a reassignment vector with significantly biased expectation value, see section 3.3.1 in [12]. Therefore, we suggest that such points are discarded. Now, if the SNR where the point is reassigned to $(\hat{t}, \hat{\nu})$ is low, it is likely that the point was not properly reassigned. As such, we propose discarding points reassigned to a location with an SNR below some threshold at least as high as the previous. This is useful since it sets a limit to the kernel width.

Points with large reassignment vectors have higher variance and contribute to the heavy tails of the RS distribution, see chapter 4 in [12]. These "far out" points can be discarded to some extent. Consider the marginal distribution of the spectrogram, which is a Gaussian, and let λ be its standard deviation. If we want to recover $\sim 95\,\%$ of its energy, this corresponds to reassigning all points from within $\pm 2\lambda$. Therefore, we suggest discarded points which are reassigned further than $\sim 2\lambda$.

IV. Evaluation

The performance of the estimated signal energy/power of the SRS is evaluated in a simulation study. We use the same signal as before – a Gaussian envelope with $\lambda = 10$. Again, for simplicity, we only look at the cross-section of the spectrogram and its reassignment in the time domain. All rules proposed in the previous section were applied, and a parabolic kernel with a relative scale of $a = 3$ was used. With this kernel, the expected energy recovery is $\sim 94\,\%$. Also, since we ignore points reassigned further than 2λ, that in itself limits the energy recovery by $\sim 4\,\%$.

In this simulation study the analyzed signal was realized $N = 5000$ times. From each realization, the peak value was saved. The spectrogram (S) is the baseline which the binned RS and SRS are compared against. To estimate the total energy $\hat{E}$ from the spectrogram peak,

$$\hat{E} = \max_t[S(t)] \cdot \sqrt{2\pi} \cdot \lambda. \tag{19}$$

For the binned RS and SRS, the estimated energy $\hat{E}$ was obtained directly from the peak.

For every peak SNR swept, the 25th and 75th percentiles was calculated, which is shown in figure 4. The binned RS clearly suffers from significant deterioration, worsening as more noise is added. For the lower SNRs, the RMSE of the SRS is only slightly worse than the spectrogram baseline. At higher SNRs, the RMSE goes to some lower limit caused by the bias. Overall, SRS behavior is very similar to that of the spectrogram. This is great since it means that using the SRS, we can reassign the spectrogram for increased resolution, while retaining information of signal energy/power.

V. Discarding Interference in Multi-Component Signals

Reassignment breaks down between signal components which are too close. An example of this is shown in figure 5, where two Gaussians are placed fairly close to each other in time. It can be seen that there is some interference between the components.

We study plot C, where the ideal reassignment $(t_0 - t)$ (orange lines) in time for the two components are shown. The actual reassignment $(c_t\tilde{t})$ (blue line) shows a transition of the reassigning points from one center to the other.

Since ideally, $c_t\tilde{t} = t_0 - t$, then $\mathrm{d}\big(c_t\tilde{t}\big) \,/\, \mathrm{d}t = (c_t\tilde{t})' = -1$, shown in plot D. We then confirm that when the reassignment is well behaved, $(c_t\tilde{t})' \approx -1$. Therefore, we suggest discriminating points which have $(c_t\tilde{t})' \not\approx -1$. Exactly how to set the limits has not been thoroughly studied. For the results later shown, the limits used are $-1.5 < (c_t\tilde{t})' < -0.5$. Expanding this to the time-frequency plane is simple – there we would put a constraint on the reassignment vector gradient instead of just in the derivate in the time dimension.

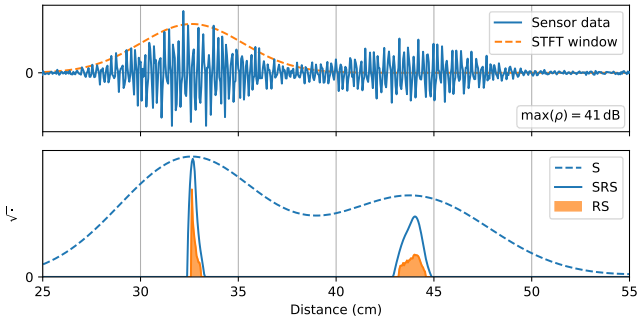

Fig. 6. An example of processing data from the Acconeer A1 sensor. The top plot shows the data and the manually matched window. The bottom plot shows the slice of the spectrogram (S), binned RS, and SRS, matching the frequency of the signal. The spectrogram has been scaled up for a better view.

VI. Application on radar data

The Acconeer A1 radar sensor is a tiny pulsed coherent 60 GHz radar with very low power consumption, ideal for IoT devices. It sends out pulses with a near Gaussian envelope which reflections are sampled with an effective sampling period of ~ 3 ps, corresponding to a ~ 0.5 mm depth resolution.

The top plot in figure 6 shows data where two objects has been placed in front of the sensor, closely enough for the reflected pulses to interfere with each other. There are nonlinear effects and a slight correlation between the data points. The figure also shows the by hand matched Gaussian window ($\lambda = 27$ mm). Other methods of matching the window exists, such as minimizing the Rényi entropy [6], [8]. In the bottom figure, the extreme smoothing of the usual spectrogram (S) is clearly seen as well as the split energy of the standard binned RS, especially for the weaker peak around 44 cm. However, with use of the SRS, the energy levels of both peaks are recovered.

VII. Conclusions

In this paper, we brought up the subject on how reassigned spectrograms are represented/resampled. A big shortcoming with naively resampling (binning) the reassigned spectrogram is that it deteriorates when subjected to noise. To mitigate this, we proposed a novel method – the *smoothed reassigned spectrogram*. By utilizing knowledge about the reassigned spectrogram distribution, applying kernels to the reassigned points, a more well behaved distribution was obtained. Additionally, a novel idea of using reassignment gradients to suppress interference of close components is presented.

References

[1] K. Kodera, C. D. Villedary, and R. Gendrin, "A new method for the numerical analysis of non-stationary signals," *Physics of the Earth and Planetary Interiors*, vol. 12, no. 2, pp. 142 – 150, 1976.

[2] F. Auger and P. 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, 05 1995.

[3] P. Flandrin, *Explorations in Time-frequency Analysis*. Cambridge University Press, 2018.

[4] F. Auger, E. Chassande-Mottin, and P. Flandrin, "On phase-magnitude relationships in the short-time Fourier transform," *IEEE Signal Processing Letters*, vol. 19, no. 5, pp. 267–270, 2012.

[5] D. Fourer, F. Auger, and 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)*. IEEE, 2016, pp. 4880–4884.

[6] M. Hansson-Sandsten and J. Brynolfsson, "The scaled reassigned spectrogram with perfect localization for estimation of Gaussian functions," *IEEE Signal Processing Letters*, vol. 22, no. 1, pp. 100–104, 2015.

[7] M. Sandsten, J. Brynolfsson, and I. Reinhold, "The matched window reassignment," in *EUSIPCO*, Rome, Italy, 2018.

[8] J. Brynolfsson and M. Sandsten, "Parameter estimation of oscillating Gaussian functions using the scaled reassigned spectrogram," *Signal Processing*, vol. 150, pp. 20–32, 2018.

[9] J. Brynolfsson, I. Reinhold, J. Starkhammar, and M. Sandsten, "The matched reassignment applied to echolocation data," in *Proceedings of the ICASSP*. Brighton, U.K.: IEEE, 2019.

[10] I. Reinhold, M. Sandsten, and J. Starkhammar, "Objective detection and time-frequency localization of components within transient signals," *J. Acoust. Soc. Am. (JASA)*, vol. 143, no. 4, pp. 2368–2378, 2018.

[11] E. Chassande-Mottin, F. Auger, and P. Flandrin, "Reassignment," *Time-Frequency Analysis: Concepts and Methods*, pp. 249–277, 2008.

[12] E. M. Månsson, "On the statistics and practical application of the reassignment method for Gabor spectrograms," Mathematical Statistics, Centre for Mathematical Sciences, Lund University, Lund (Sweden), Master thesis, June 2019.

[13] E. Chassande-Mottin, P. Flandrin, and F. Auger, "Statistique des vecteurs de reallocation du spectrogramme," Laboratories de Physique, ENS-Lyon (URA 1325 CNRS), Lyon (France), Rapport interne, 1 1996.

[14] ——, "On the statistics of spectrogram reassignment vectors," *Multidimensional Systems and Signal Processing*, vol. 9, no. 4, pp. 355–362, 10 1998.

[15] V. A. Epanechnikov, "Non-parametric estimation of a multivariate probability density," *Theory of Probability & Its Applications*, vol. 14, no. 1, pp. 153–158, 1969.