Non-parametric Envelope Estimation for the Matched Window Reassignment

Isabella Reinhold
Centre for Mathematical Sciences
Lund University
Lund, Sweden
isabella.reinhold@matstat.lu.se

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

***Abstract*—The reassigned spectrogram is a powerful tool for analysing non-stationary signals, and in an ideal setting it gives perfect time and frequency localisation. A method very well suited for oscillating transient signals is the matched window reassignment, which requires a matching window, i.e. the envelope of the transient, to be known or estimated beforehand. This paper proposes a novel method for estimating the envelope of noisy transients, using a non-parametric and computationally efficient approach. The estimated envelopes are used to calculate the matched window reassignment, obtaining estimates of the time-frequency centre of the transients. The reassignment using the estimated envelopes is shown to give good estimates of the time-frequency centres, and good localisation in time and frequency. The novel envelope estimation approach is illustrated on measured marine biosonar data.**

***Index Terms*—time-frequency analysis, oscillating transient, reassigned spectrogram**

I. Introduction

A common method for analysing non-stationary signals is the spectrogram, however a major drawback of this method is the trade-off between resolution in time and frequency. A long time-window can be used to get a reliable estimate of the frequency content, but the information on where in time that frequency content is will be poor. Conversely, a short time-window can be used to get better resolution in time, but with consequential poor frequency information. The reassignment method has been proposed as a solution to this problem [1]. The reassigned spectrogram can give perfect localisation in time and frequency and was first designed for linear chirps, but has in more recent years been adapted for oscillating transient signals, reassigning signal energy to the time-frequency centres (TF) of individual transients [2], [3].

Transients are common in ultrasonic and marine biosonar analysis, machine fault diagnosis, and biomedical signal processing [4]–[7]. The matched window reassignment (MWR) can identify the time centres and frequencies of individual transient components, but good performance relies on having good information about the transient envelope [3]. However, the signal envelope is usually not known and thus needs to be estimated. In previous studies, we have used parametric methods to estimate the envelope, assuming a Gaussian function and optimising the length [8], or by evaluating different functions and lengths [9]. These approaches either requires some prior information about the envelope or can be costly if many window functions or lengths are evaluated.

In this paper, we present a novel method for estimating the envelope of an oscillating transient from a noisy realisation. Our proposed method is non-parametric and requires no prior knowledge of the transient envelope. The estimated envelope is used to calculate the matched window reassignment, from which a TF centre estimate of the transient is obtained. The performance of the envelope estimation method is evaluated by examining the bias of the TF centres, and the Rényi entropy of the reassigned spectrograms. The estimation approach is automatic, and we evaluate the method for signals disturbed white and pink noise, respectively. The estimation approach is also evaluated on a measured marine biosonar signal.

Thanks to Josefin Starkhammar for providing beluga whale echolocation data.

II. The matched window reassignment

The reassigned spectrogram of a real-valued signal $x(t)$ and window function $h(t)$

$$RS_x^h(t,f) = \iint S_x^h(s,\xi)\delta\big(t-\hat{t}(s,\xi), f-\hat{f}(s,\xi)\big)ds d\xi, \quad (1)$$

where the integrals run from $-\infty$ to ∞ and $\delta(t,f)$ is the two-dimensional Dirac impulse, improves the readability of the spectrogram

$$S_x^h(t,f) = \big|F_x^h(t,f)\big|^2 = \left|\int x(s)h(s-t)e^{-i2\pi fs}ds\right|^2. \quad (2)$$

This is achieved by reassigning the signal energy of the spectrogram closer to the instantaneous frequencies of the signal components. The energy is moved according to the reassignment coordinates, which depend on the signal and window. While the reassignment method was originally invented for longer oscillating signals [1], the scaled reassigned spectrogram expands this theory to oscillating transients [2].

The scaled reassignment coordinates are calculated according to

$$\hat{t}(t,f) = t + c_t \Re\left(\frac{F_x^{th}(t,f)}{F_x^h(t,f)}\right),$$
$$\hat{f}(t,f) = f - c_f \frac{1}{2\pi} \Im\left(\frac{F_x^{dh/dt}(t,f)}{F_x^h(t,f)}\right), \tag{3}$$

where c_t and c_f are the scaling constants, $\Re$ and $\Im$ represent the real and imaginary parts and F_x^h, F_x^{th} and $F_x^{dh/dt}$ are the short-time Fourier transforms (STFTs) that use $h(t)$, $t \cdot h(t)$ and $dh(t)/dt$ as window functions.

We have previously presented the matched window reassignment (MWR) [3], which for $c_t = c_f = 2$ gives perfect localisation at the TF centre of a general oscillating transient

$$x(t) = a(t - t_I)\cos(2\pi f_I t), \tag{4}$$

where $a(t)$ is the envelope function, t_I and f_I are the time and frequency shifts, if the window function is $h(t) = a(-t)$. This result holds for any time and frequency shifts due to the linearity of the Fourier transform and thus the reassignment coordinates [10].

III. Non-parametric envelope estimation

Estimating $a(t)$ from the real-valued $x(t)$ in Eq. (4) is trivial, given that the signal is measured long enough and is adequately sampled. It is a more difficult problem if the signal has added noise. In this paper, we will consider signals

$$y(t) = x(t) + \epsilon(t), \tag{5}$$

where $\epsilon(t)$ is zero-mean noise uncorrelated with $x(t)$. The noise can be either white or pink ($1/f$). An example of such a signal with white noise is shown in the top plot of Figure 1, the transient $x(t)$ is shown in the bottom plot.

A. Estimation approach

The method we propose is designed to be used automatically on noisy signals with one transient component and consists of four steps:

1) Denoise the signal
2) Calculate the envelope of the denoised signal using the Hilbert transform
3) Cut and smooth the shape of the envelope
4) Estimate the mass centre point and adjust envelope length so that point is centred lengthwise

In step 1 we suggest to denoise the signal. This helps create consistent results for the following steps but is not always necessary especially if the signal has high SNR. The denoising can be done by decomposing the signal using the discrete wavelet transform, which is computationally efficient. In our analysis we used the Daubechies db4 wavelet [11], a four-level decomposition, and the Stein's Unbiased Risk Estimate threshold [12]. Figure 1 shows an illustration of this step.

Step 2 calculates the envelope of the denoised signal. This is done by using the Hilbert transform of the real-valued denoised signal to get the corresponding analytic signal. The envelope is then obtained by the absolute value of the analytic signal. The Hilbert transform does require that the centre frequency of the transient is not too close to zero, but as this paper deals with oscillating transients we will assume that the centre frequency is sufficiently high, otherwise it is likely that the transient would not be oscillating. Figure 1 shows an example of the result of this step, it can see seen that the envelope is not very smooth at this stage.

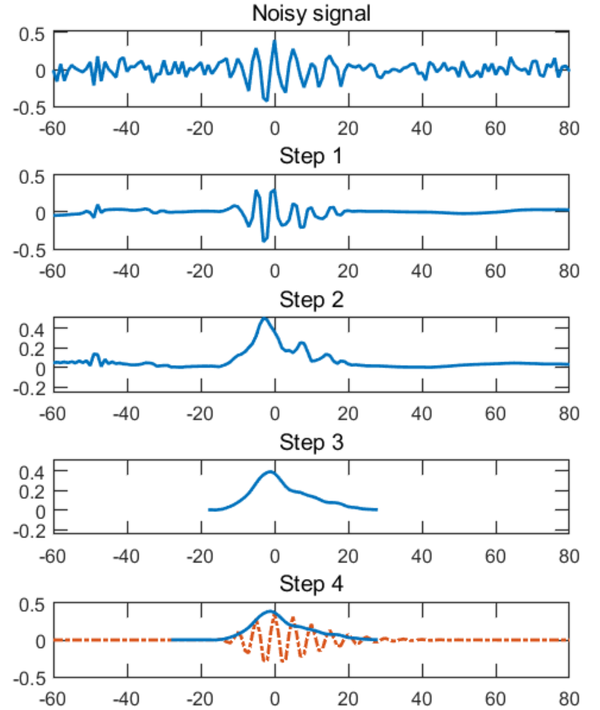

Fig. 1. The steps of the envelope estimation demonstrated on a transient with white noise (SNR = -2 dB) and denoised using discrete wavelet transform. The lowest plot shows the final estimate superimposed on the transient without noise.

The denoising step should make sure that the areas outside the transient are close to zero, however it cannot be expected that these areas are exactly zero or without smaller areas of higher amplitudes, see step 2 in Figure 1. Therefore, step 3 is needed. The envelope is at this stage of equal length with the signal, and the transient can be located by finding the envelope point that has the highest value. Seen from that maximum, we expect the envelope to taper down on both sides and after reaching values close to zero, we would not like the amplitude to increase again. Because of this a tolerance can be set of what is considered close enough to zero, and when, seen from the maximum, the envelope first goes below this tolerance the envelope is cut. This should be done on both sides of the maximum. The smoothing of the cut envelope can be accomplished by a short moving average or low-pass filter. Figure 1 shows the result of step 3 on the example signal, it can be seen that only slight smoothing is used, this is to preserve the general shape of the transient envelope. In this paper the length of the moving average filter is 15% of the length of the envelope after the cut in step 3.

The tolerance in step 3 can be set by considering the median value of the signal envelope. If the signal is relatively long and the transient relatively short, the median value should be an estimation of the general "close to zero" level outside the transient envelope. However, if the signal is relatively short compared to the transient, the median is not necessary a good estimate. Then a good tolerance can instead be a fraction of the maximum amplitude already calculated in step 3. We recommend calculating both these potential tolerances and choosing the smallest of them as the final tolerance.

The envelope now looks good enough, however, step 4 significantly improves the TF centre estimate given by the reassigned spectrogram that uses the estimated envelope. First, numerically calculate the mass centre of the envelope from step 3, that point should then be centred lengthwise in the envelope. This is best done by adding samples to the window, either in the beginning or end, depending on if the mass centre point was larger or smaller than half the envelope length. To avoid sudden changes in the amplitude of the window, it is preferable to add samples with the same value as its neighbour, i.e. the value of the first or last sample of the envelope. This will of course slightly alter the mass centre, but this change seems to not significantly affect performance. In Figure 1 samples are added to the beginning of the envelope.

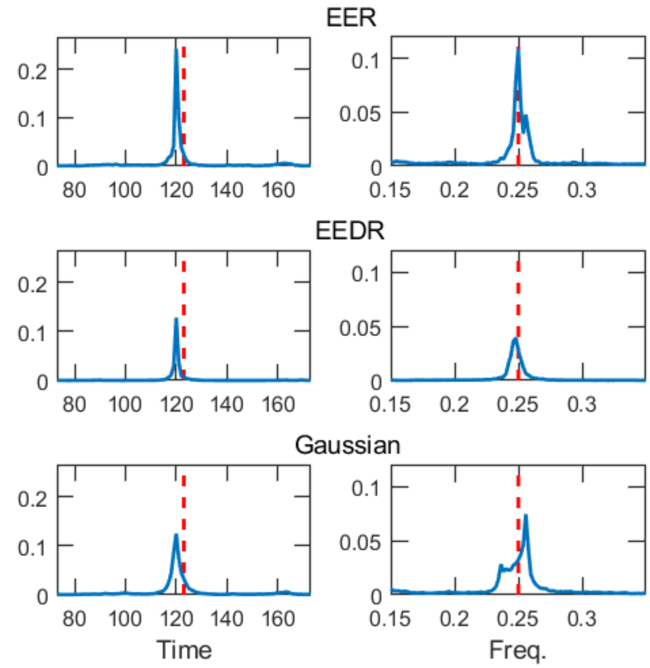

Fig. 2. Time and frequency marginals of the reassigned spectrograms of a transient Eq. (8) with white noise (SNR = 0 dB), for EER, EEDR and Gaussian $p = 8$. The vertical, dashed red lines show the true time and frequency centres.

IV. Evaluation

The estimated transient envelope, $\hat{a}(t)$, is used to calculate the MWR, by having $h(t) = \hat{a}(-t)$ in Eq. (1) and its associated calculations. The calculation of the MWR also requires a decision on what signal to use, the original noisy one or the denoised signal obtained in step 1 of the estimation approach, both options will be evaluated in this paper. The first option, we will call the estimated envelope reassignment (*EER*) and calculates the spectrogram and reassignment with the original noisy signal. The second option calculates the spectrogram and reassignment with the denoised signal, we call this the estimated envelope denoised reassignment (*EEDR*). The performance of these two options, which use the estimated envelope, will be compared to calculating the spectrogram and reassignment with a *Gaussian* window, a common window choice if information is limited, and the noisy signal.

The signal envelope used for this evaluation is asymmetrical with a heavier tail, since this is a common appearance of measured transients [13]–[15]. Such an asymmetrical envelope can be modelled by a skewed distribution function, in this paper we use the Gumbel distribution

$$a(t) = \frac{1}{p} e^{-\frac{t}{p} - e^{-\frac{t}{p}}}, \quad t \in \mathbb{R}, \tag{6}$$

where p is the scaling parameter. This means that the envelope will be skew in time and symmetric in frequency, which in turn gives a time mass centre, t_0, that is not equal to the time shift, t_I, since the mass centre is

$$t_0 = \int |a(t - t_I)|^2 (t - t_I) dt, \tag{7}$$

and $a(t - t_I)$ is not symmetric. The frequency centre can be calculated in an analogue way, but since the frequency distribution is symmetric, it holds that $f_0 = f_I$.

To evaluate the EER, EEDR and Gaussian approaches, we use a set of simulated oscillating transient signals with noise

$$y(t) = a(t - t_I) \cos(2\pi f_0 t) + \epsilon(t), \, t = 0, 1, \ldots 249, \tag{8}$$

where the envelope $a(t)$ is the Gumbel distribution in Eq. (6). The sampling rate is 1 Hz, time shifts $t_I \in \mathcal{U}(50, 170)$, giving a similar but not equal range for the time centre t_0, frequency centres are $f_0 \in \mathcal{U}(0.1, 0.4)$ and scaling $p \in \mathcal{U}(4, 14)$. The spectrogram is calculated with 256 positive frequency bins, making the number of time and frequency bins approximately equal. Two types of noise are evaluated, white noise and pink noise, 1000 simulations are used for each noise type and a range of SNR. The SNR range $(-4, 10)$ dB for the white noise and $(-2, 12)$ dB for the pink noise, where

$$\text{SNR} = 10 \log_{10}(\text{signal energy/noise variance}) \tag{9}$$

and the signal energy is calculated as the L^2-norm of $x(t)$.

The performance of the EER, EEDR and Gaussian is evaluated on two aspects. The first is the bias of the TF centre estimate $(\hat{t}_0, \hat{f}_0)$. According to previous studies [16], the estimated TF centres are obtained from the location of the highest peak in the reassigned spectrogram using the respective windows/envelopes. The time and frequency marginals of the reassigned spectrograms in Figure 2 illustrates how this can be done. All the marginals have clear peaks close to the true TF centres. The bias is calculated as the radial distance,

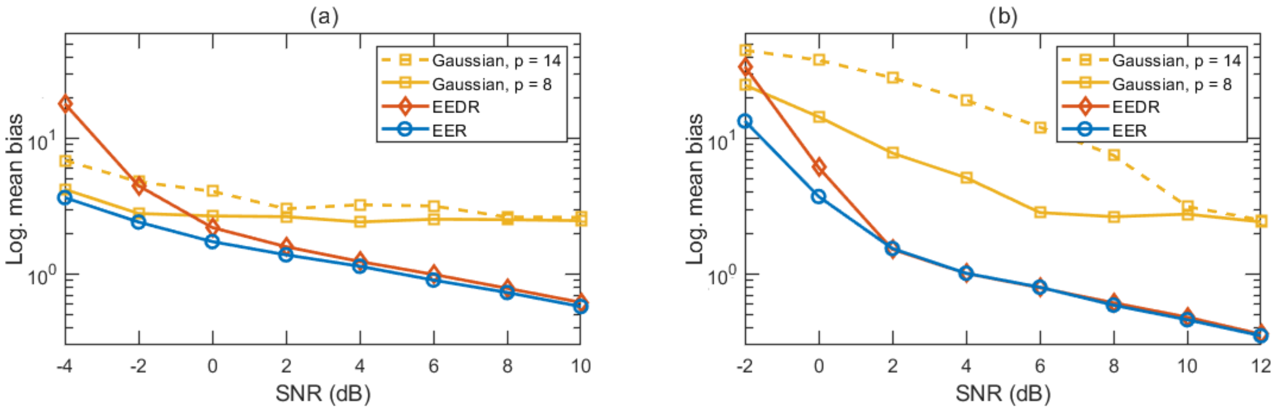

Fig. 3. Mean bias in estimates of the TF mass centres obtained from the reassigned spectrograms with the different windows for the transient signals Eq. (8) with different SNR and (a) white noise, (b) pink noise. Logarithmic scale is used on the y-axis.

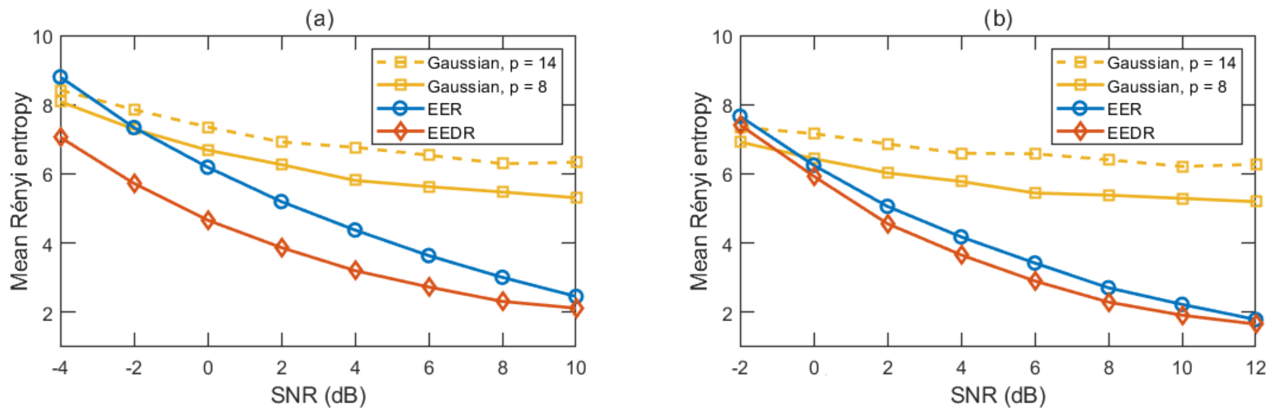

Fig. 4. The mean Rényi entropy of the reassigned spectrograms with the different windows for the transient signals Eq. (8) with different SNR and (a) white noise, (b) pink noise.

$r = \sqrt{(t_0 - \hat{t}_0)^2 + (f_0 - \hat{f}_0)^2}$, where the true t_0 is given by Eq. (7) and f_0 by Eq. (8). The second aspect is how localised the signal energy is after reassignment, which is measured using the Rényi entropy, where a lower entropy is better [17].

The EER and EEDR are non-parametric methods but using the Gaussian window is not. The length of the Gaussian window needs to be defined, fine tuning of the scaling parameter is possible but requires additional computations and evaluation [8]. Therefore, the two fixed lengths of the Gaussian window are used, the first is optimal for the median length of the signal envelope, $p = 8$, and the other is a long window, $p = 14$, that will always fit the entire transient inside the window.

V. Results and Discussion

The mean biases for the estimates of the TF mass centres for the simulations with white noise and pink noise are shown in Figure 3(a)–(b) respectively. The EER outperforms the Gaussian windows for all evaluated SNR and both noise types, there is however only small differences for the lowest SNR. This is because the envelope estimation is harder for low SNR, therefore the estimated envelopes will on average only be slightly better matches compared to the Gaussian window, especially for $p = 8$. As the SNR increases, the mean biases from the EER and EEDR steadily decrease, the same is not in general seen for the Gaussian windows as the mean biases will reach a plateau. Both window lengths reach the same plateau, suggesting that this is the bias we will always get by estimating the mass centre of the transient with a heavy tail, using the symmetric Gaussian function.

The EEDR performs significantly worse that the EER for low SNR. This is due to the denoising sometimes affecting the oscillations of transient, which is suggested by step 1 in Figure 1. If the denoised signal has a damaged transient and is used for the reassignment calculations, this will affect especially the frequency centre estimate but to a lesser degree also the time centre estimate. This effect is dependent on the denoising method, a better tailored denoising method might mitigate this result, but a more careless choice will also worsen the effect. For higher SNR it is easier to separate the transient from the noise and the EEDR performs as well as the EER.

The mean Rényi entropies are shown in Figure 4, a low Rényi entropy indicates a clean TF representation with localised energy. The EEDR has the lowest entropy, followed by the EER, except for the signals with pink noise and SNR $= -2$ dB. The low entropy of the EEDR is because the signal is denoised and does not indicate higher localisation around

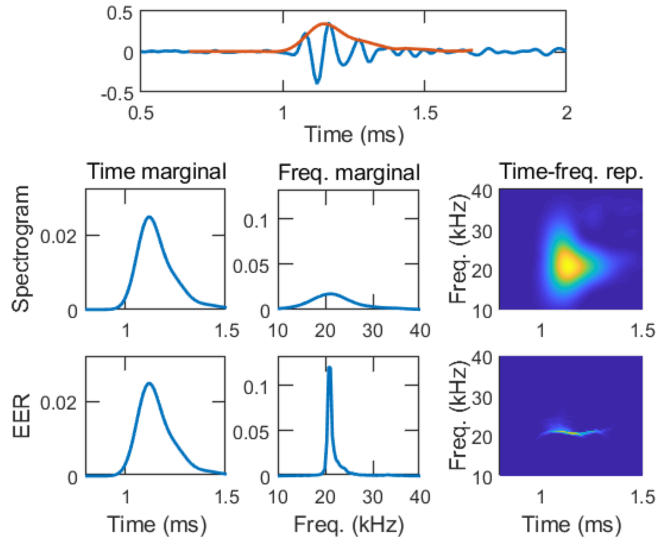

Fig. 5. Example of the proposed EER on an echolocation signal from a beluga whale. The top plot shows the signal and the estimated envelope, the two bottom rows show the time and frequency marginals as well as the TF representations of the spectrogram and reassigned spectrogram with the estimated envelope.

the TF centre of the transient. There are only small differences between the EEDR and EER, and they converge when the SNR increases. All methods have decreasing Rényi entropy when the SNR increases, but the decrease is much faster for the EER and EEDR, suggesting that using the estimated envelopes to get a matched window is preferred.

VI. Marine biosonar example

It is common to model the echolocation clicks of dolphins as Gaussian enveloped oscillating signals, however it has lately been shown that other envelopes might be better suited [15], [18]. The EER is here applied to a transient echolocation signal from a beluga whale (*Delphinapterus leucas*), the signal is sampled with 1 MHz and recorded by one of 47 simultaneously sampling hydrophones as described in [19]. The echolocation signal is shown in the top plot of Figure 5, superimposed is the non-parametric and automatically estimated envelope. In Figure 5, shows both the spectrogram and the reassigned spectrogram calculated with the estimated envelope and their time and frequency marginals.

VII. Conclusion

In this paper we propose a novel method for estimating the envelope of a noisy oscillating transient. This envelope is then used as a matched window when calculating the reassigned spectrogram of that transient. The estimation method is automatic, computationally efficient and requires no prior knowledge of the transient envelope.

The results show that reassignment calculations with the estimated envelopes gives good estimates of and localisation at the TF centres of simulated transients with white noise (from SNR = -4 dB) or pink noise (from SNR = 0 dB). The estimation method requires a denoising step, it is evaluated if the denoised signal should also be used for the reassignment calculations (EEDR) or not (EER). The results show that the EER gives the most consistent results, and outperforms the results obtained when assuming a Gaussian envelope.

References

[1] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Trans. on Signal Processing*, vol. 43, pp. 1068–1089, May 1995.

[2] 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, January 2015.

[3] M. Sandsten, J. Brynolfsson, and I. Reinhold, "The matched window reassignment," in *26th European Signal Processing Conference, EUSIPCO 2018*. European Association for Signal Processing (EURASIP), 2018, pp. 2340–2344.

[4] C. Wei, W. W. L. Au, D. R. Ketten, Z. Song, and Y. Zhang, "Biosonar signal propagation in the harbor porpoise's (Phocoena phocoena) head: The role of various structures in the formation of the vertical beam." *J. Acoust. Soc. Am.*, vol. 141, no. 6, p. 4179, 2017.

[5] J. Starkhammar, I. Reinhold, P. Moore, D. Houser, and M. Sandsten, "Detailed analysis of two detected overlaying transient components within the echolocation beam of a bottlenose dolphin (Tursiops truncatus)," *J. Acoust. Soc. Am.*, vol. 145, no. 4, pp. 2138–2148, 2019.

[6] J. Pons-Llinares, M. Riera-Guasp, J. A. Antonino-Daviu, and T. G. Habetler, "Pursuing optimal electric machines transient diagnosis: The adaptive slope transform." *Mechanical Systems and Signal Processing*, vol. 80, pp. 553 – 569, 2016.

[7] Q. He and X. Ding, "Sparse representation based on local time-frequency template matching for bearing transient fault feature extraction." *Journal of Sound and Vibration*, vol. 370, pp. 424 – 443, 2016.

[8] J. Brynolfsson and M. Sandsten, "Parameter estimation of oscillating Gaussian functions using the scaled reassigned spectrogram," *Signal Processing*, vol. 150, pp. 20–32, 9 2018.

[9] J. Brynolfsson, I. Reinhold, and M. Sandsten, "A time-frequency-shift invariant parameter estimator for oscillating transient functions using the matched window reassignment," *Signal Processing*, vol. 183, p. 107913, 2021.

[10] E. Chassande-Mottin, F. Auger, and P. Flandrin, "Reassignment," in *Time-Frequency Analysis*, F. Hlawatsch and F. Auger, Eds. London: U.K.: ISTE, 2008, ch. 9, pp. 249–277.

[11] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Comm. Pure Appl. Math*, vol. 41, no. 7, pp. 909–996, 1988.

[12] C. M. Stein, "Estimation of the mean of a multivariate normal distribution," *The Annals of Statistics*, vol. 9, no. 6, pp. 1135–1151, 1981.

[13] A. Venkitaraman and C. S. Seelamantula, "Temporal envelope fit of transient audio signals," *IEEE Signal Processing Letters*, vol. 20, no. 12, pp. 1191–1194, 2013.

[14] R. F. Lyon, A. G. Katsiamis, and E. M. Drakakis, "History and future of auditory filter models," in *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, 2010, pp. 3809–3812.

[15] J. Brynolfsson, I. Reinhold, J. Starkhammar, and M. Sandsten, "The matched reassignment applied to echolocation data," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2019, pp. 8236–8240.

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

[17] R. G. Baraniuk, P. Flandrin, A. J. E. M. Janssen, and O. J. J. Michel, "Measuring time-frequency information content using Rényi entropies," *IEEE Trans. on Information Theory*, vol. 47, no. 4, pp. 1391–1409, May 2001.

[18] C. Capus, Y. Pailhas, K. Brown, D. Lane, P. Moore, and D. Houser, "Bio-inspired wideband sonar signals based on observations of the bottlenose dolphin (Tursiops truncatus)," *J. Acoust. Soc. Am.*, vol. 121, no. 1, pp. 594–604, 2007.

[19] J. Starkhammar, M. Amundin, J. Nilsson, T. Jansson, S. Kuczaj, M. Almqvist, and H. Persson, "47-channel burst-mode recording hydrophone system enabling measurements of the dynamic echolocation behavior of free-swimming dolphins," *J. Acoust. Soc. Am.*, vol. 126, no. 3, pp. 959–962, 2009.