Non-stationary sinusoidal analysis

Sašo Muševič

MASTER THESIS UPF / 2009
Master in Sound and Music Computing

Master thesis supervisor:
Jordi Bonada
Department of Information and Communication Technologies
Universitat Pompeu Fabra, Barcelona

UNIVERSITAT POMPEU FABRA

Contents

1 Introduction **4**
1.1 Sinusoidal modeling 4
1.2 Spectral modeling synthesis 4

2 State of the art **6**
2.1 FFT based approaches 6
2.1.1 Quadratic Interpolated Fast Fourier Transform (QIFFT) . 6
2.1.2 Reassignment method (RM) 7
2.1.3 Derivative Analysis Method (DAM) 8
2.2 Quadratic Phase/Chirp transform (QPT) 9
2.3 Wigner-Ville transform (WVT) 11

3 Reassignment and derivative analysis method implementation **15**
3.1 General considerations 15
3.2 DAM 16
3.3 RM 17
3.4 Implementation specifics 18
3.4.1 Spectrum values at non FFT bin frequencies 18
3.4.2 Spectrum correction 18

4 Comparison of reassignment and derivative method analysis **21**
4.1 Frequency 21
4.1.1 DAM 21
4.1.2 RM 26
4.2 Amplitude Modulation 31
4.2.1 DAM 31
4.2.2 RM 36
4.3 Frequency Modulation 39
4.3.1 DAM 39
4.3.2 RM 41
4.4 Amplitude 44
4.4.1 DAM 44
4.4.2 RM 45
4.5 Theoretical extensions of DAM to 2nd order AM/FM estimation 46
4.6 Detailed study of DAM properties 48

5 Conclusion and future work **50**

Abstract

Signal analysis is a ground-level basis of many scientific applications. Since computers have become powerful enough to analyze/synthesize sound in real-time, field has received even more attention. Most of computer sound applications use Fourier analysis in one form or another. It is used for as simple cases as visualizing sound, extracting general sound features or providing basis for accurate sound parameterization techniques. Its popularity lies in the fact, that Fourier transform provides very human readable and sonically meaningful parameterization. With discovery of Fast Fourier Transform in 1965 and ideas like phase vocoding eventually led to development of Spectral Modeling Synthesis. This synthesis method essentially synthesizes sound from a spectrum, a word widely used to denote Fourier transform. Constructing a reasonably good spectrum from scratch can be a painstakingly long lasting task, so some kind of templates should be used. Such templates can be extracted from sounds of instruments. Such procedure can result in very naturally sounding synthetic sounds. Of course, a high quality analysis is a must, as the most interesting instruments tend to contain numerous details in form of subtle frequency and amplitude modulations, giving it its distinct character, recognized by producers and music lovers alike. To present day, sounds are analyzed by segmenting it into smaller pieces and applying analysis methods frame by frame. Such method is very practical, as it can be used in real-time manner, but forces one to make a tradeoff. Most analysis methods assume, that sound is completely stationary within one frame. In case that signal does not posses such property, smaller frames are taken. Clearly, frame cannot be made arbitrarily small for many different reasons, so a need of non-stationary signal analysis technique was obvious. When analyzing sound with intent to use it for spectral modeling synthesizer specifically, a specific form of non-stationary signal analysis is used: non-stationary sinusoidal analysis. In essence, such analysis should be able to detect fast frequency and amplitude changes of sinusoids, even if such changes occur inside a analysis frame, even if there is more sinusoids relatively close to each other in frequency and even in presence of noise.

Acknowledgments

Thanks

I'd like to thank my mentor Jordi Bonada, for finding even the tiniest bugs in my code and not giving up on me.

I'd like to thank following people that I've met in Barcelona in 2008-2009. You made my life there a joy, giving me inspiration and motivation for my academic work. Therefore I have to mention you all. Thanks again...

...to Leo, who I've never seen not being euphoric about life, each and every single day

...to Nate, one of the coolest and most creative guys ever

...to Robert, who taught me Scottish through endless jokes, which eventually turned against him

...to Juan, the biggest source of good music and positive energy

...to Mariona, the sweetest person I've ever met

Special thanks

Lat but not least, I'd like to give the biggest and most special thanks to my parents, who made my masters in Barcelona possible and knew better than me, how much I needed to do it.

1 Introduction

1

1.1 Sinusoidal modeling

Sinusoidal modeling is a very popular technique in computer sound and music based applications. Due to its flexibility and the fact that it is closely related to human perception of sound, makes it very attractive for practical applications, as model parameter manipulation results in sonically meaningful modification of resulting sound. Roots of ideas, that gave birth to sinusoidal model go back as far as mid sixties with digital phase vocoder [12] and gradually evolved to an idea of parametric model of a sound, which assumes that sound is composed of sinusoids and noise [26]. This assumption has survived years of challenges and had strong impact on research of sound and music synthesis. Further, Music Information Retrieval (MIR) community also profited greatly from its development, as many classification techniques rely greatly on sinusoidal analysis. Polyphonic transcription is mostly used in MIR related applications and greatly relies on sinusoidal analysis as well.

1.2 Spectral modeling synthesis

Area of spectral/sinusoidal modeling synthesizers has not yet reached it's full potential, as only a few commercial exist, namely Yamaha - Vocaloid, Virsynt - Tera,Camel Audio - Cameleon 5000, Synful - Orcehstra [20], to mention just a few. Considerable success was achieved in voice synthesis [8] but still, the area of spectral/sinusoidal model synthesis (SMS) is receiving considerable amount of academic attention, as many attractive, intriguing questions are still not answered [27]. Further it is estimated, that many of techniques are ready for practical use [27] and therefore its full potential is believed not to be exploited fully by the industry. It is important to mention, that an alternative to SMS - physical modeling (PM) has met similar problems [30], however next generation of sound synthesizers *could* in fact be hybrids between SMS and PM [30].
In a race for sound synthesizer market, the only two key measures are quality and efficiency. In SMS context, quality is easily measurable simply in quality of re-synthesized sound, compared to original, which often involves rather non-formal listening tests. Implicit quality, that is somehow hidden at first glance is quality of spectral model, eg: how many spurious data the model contains. In many of presented studies, quality of synthetic sounds (whose parameters are known) is often measured in error variance or maximum relative error or more complicated measures like energy of residual, noisiness of the residual and similar. Such measures are of course necessary in initial, research/prototype development phase. In more high level cases, appropriate for commercial purposes, signal parameters are rarely known, so simple mathematical measures often cannot be employed. Even when such measure would indeed exist, listening tests remain crucial, as certain (strictly mathematically) low error methods

might produce more audible artifacts as others. This makes quality assessment a bit tricky, as it is very much bound to human perception of sound. Luckily, sound compression research community succeeded to isolate perceptual properties, that are important for high quality sound impression and formulate them in technical terms [7].

When building sinusoidal model of a natural sound source, special care should be taken to frequency and amplitude modulations, as they seem to pose the biggest problems to all analysis methods, especially those incorporating sliding window method. Most analysis methods are still based on Fourier transform (FT), as it is by far the fastest and most direct analysis method. The downside is however a necessity to use sliding window approach, which dramatically decreases resolution, since in practive Gabor-Heisenberg inequality [13] makes a tradeoff between temporal and spectral resolutions unavoidable. In real-time applications, such trade off is unavoidable and is accepted in signal processing community nowadays. Nonetheless, every subtle change inside the windowed part of signal is averaged over the window length, so highly non-stationary signals produce bloated FTs, which is the main subject of present proposal.

Many types of everyday signals fall into category of *highly non-stationary*, most noticeably speech. Even for nearly stationary signals, methods that could detect even the slightest amount of non-stationarity are welcome, as **perfect** sound analysis (as far as humanly perceptible) in terms of sinusoidal model should be achieved as soon as possible. Such analysis quality would most probably set SMS right in the top of computer sound synthesis industry's interest. There has been endless efforts to overcome the inability of FT to detect amplitude modulation (AM) and frequency modulation (FM). These efforts have persisted through years, as perfect non-stationary spectral analysis proved to be a hard problem, especially in polyphonic case, which can be viewed as very high noise scenario. However it is very probable, that we will see a drastic change in near future, as linear AM/FM estimators have been around for years and only 1 step further, eg: generalization to 2nd order AM/FM estimators is most probably all that is needed. By appropriate window length selection, every imaginable modulation, let it be as non-linear as imaginable, could be very closely approximated by sequence of (possibly overlapping) 2nd order polynomials, so higher order AM/FM models most probably wouldn't bring much added value. Further, if mentioned methods could work well with highly noisy signals, they offer a good performance boost for polyphonic audio transcription, as instruments, overlapping in frequency possibly do not share the same AM/FM parameters, so this offers additional relevant parameters helpful with separation. It seems imperative for non-stationary audio analysis community to take the next step and explore generalization of existing methods to 2nd order FM/AM.

2 State of the art

Numerous methods for measuring have been developed in recent years. Below, categorization of selected methods is presented:

- FFT based approaches
  - Quadratic Interpolated Fast Fourier Transform
  - Reassignment Method
  - Derivative Analysis Method
- Quadratic Phase/Chirp transform
- Wigner-Ville transform

It is important to note, that all mentioned methods deal with 1st order, eg: linear AM/FM at best. This gives satisfactory results for many applications, but there is no reason not to continue research and try to reach higher order AM/FM estimations, as long as it doesn't imply *too much* computational overhead. By *too much* it is considered keeping computational complexity independent of signal properties as much as possible. In FFT based approaches that implies keeping number of required FFTs/frame independent of signal properties in order to keep frame processing time constant and predictable.

2.1 FFT based approaches

2.1.1 Quadratic Interpolated Fast Fourier Transform (QIFFT)

As all methods, QIFF had its stationary version first [16]. This initial version was used simply to acquire better peak location estimate in magnitude FFT, as it is essentially limited to bin number. Stationary sinusoid case was assumed at that point. First attempts to estimate linear AM/FM involved a second order polynomial interpolation of log-amplitude spectrum around the peak and second order polynomial interpolation of unwrapped phase around the corresponding peak in [24]. Because log-amplitude spectrum is interpolated and amplitude/phase are assumed to be quadratic functions of frequency, such interpolation is *exact*.
Assuming the same model, but estimating polynomial coefficients via minimization of *log spectral error*, similar approach was taken in [23]. Minimization of error involved iterative procedure, so high computational cost was involved.
A mathematically consistent, computationally efficient and generalized method was presented in [2]. In [2] and [23] a correction is made in order to decrease bias, that is introduced by the fact that spectral peak, picked from magnitude spectrum is shifted in presence of AM/FM modulation [2]. It's interesting to note, that computationally more expensive method, described in [23] does not suffers from this bias. The downside of all mentioned methods is, they assume use of Gaussian window (referred to as *direct method*) and require phase unwrapping. In [2] correction factors are introduced in order to adapt it to non-Gaussian

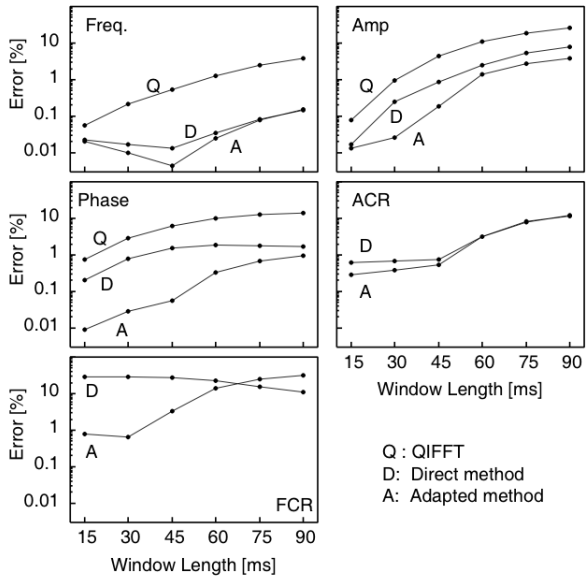


Figure 1: Figure shows RMS biases of QIFFT with Hann window. Taken from [2]

windows (referred to as *adapted method*). Quality is improved greatly by such corrections, yet method remains very sensitive to window shape [22], so use of alternative windows is difficult, as correction factors must be computed. In [2], tests were performed with 1000 sinusoids of random parameters, results are in figure 1. It is clear, that errors of all parameters rise with window length, reaching about 10% for amplitude, AM, FM at 90ms frame lengths, which is quite high.

2.1.2 Reassignment method (RM)

Idea of *reassignment* was first presented in [19], applied to spectrogram and was almost 15 years later generalized to any bilinear time-frequency or time-scale distribution in [4] and [3]. In case of Cohen class [9] time-frequency distributions (including STFT), reassignment methods has a nice visual explanation: relocate points in time-frequency plane from its current position to the *center of gravity*, located within support of smoothing kernel, centered at original position. To further clarify this idea, on which non-stationary analysis is based, figure 2 shows and illustration of transformation. One important property of this transform is, that it perfectly localizes linear chirp signals, which makes it interesting in our context. This fact is result of definition of gravity center and the fact that linear chirp ideally forms straight line in most time-frequency distribution plots, as shown in figure 2. Considering quadratic chirp signals we can intuitively (with help of figure 2) conclude, that it's time-frequency representa-

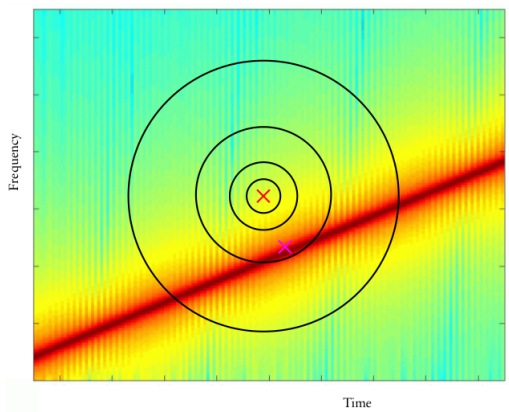


Figure 2: Reassignment: actual position (red cross) is reassigned to center of gravity of energy centre of area (violet cross), contained within kernel, whose shape and support is characterized with black circles (contours). From [14]

tion would form a parabola in time-frequency plot. An energy gravity center will be therefore always positioned below/above the parabola in concave/convex case respectively, hence 2nd and higher order chirps are **not** perfectly localized by reassignment method. It was however recently shown, that it is possible to extend reassignment method to estimate higher order modulations [29].
Use of reassignment method for non-stationanary sinusoid analysis was formalized in [25]. Once again it was discovered, that dealing with non-stationarity decreases flexibility concerning type of windows. Estimating AM/FM via reassignment requires computing FFT with use of original window, its 1st and 2nd derivative and time-ramped version of original window and its 1st derivative. FTs of window derivatives show significant side lobe amplification, therefore much stronger side lobe rejection is required. Squared Hanning window (equivalent to 3-term Nutall window with 30dB side lobe decay) and multiplication of Hanning and Hamming window is proposed and tested. However further study of low side lobe windows is encouraged, for reassignment method to achieve better results [25].

2.1.3 Derivative Analysis Method (DAM)

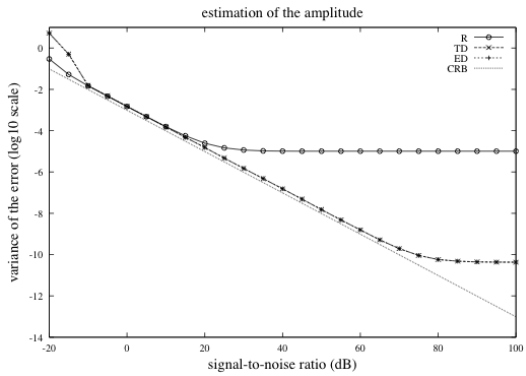
DAM is based on certain interesting properties of signals derivatives FTs . They greatly improve quality of analysis in terms of estimates of peak locations and heights in stationary case, as was first pointed out in [10] and [21]. Downside of this method is necessity to compute n^{th} derivative of signal. In practice, when dealing with sampled signals, using sample difference as approximation of derivative, high errors are associated with it [22]. Recently, this method was generalized for a non-stationary case [22] and signal derivatives are now com-

puted with high order derivation filter, which adds some computational cost, but is in the end very beneficial. An expected bias in FM estimation is observed in frequencies, close to Nyquist. Comparing this method to RM, it outperforms it in amplitude, frequency, phase and amplitude modulation, but not in frequency modulation estimation [22]. It is important to note, that using smaller window sizes, DAM outperforms RM in all estimations. To give a more exhaustive comparison of these state-of-the-art methods for non-stationary sinusoidal analysis, figures 3(a), 3(e), 3(c), 3(b), 3(d) show analysis errors for different estimators in respect to signal-to-noise ratio, as presented in [22]. Errors are compared to Cramer-Rao bounds as defined in [11]. Test signals consist of sinusoids of 99 different frequencies from 0 up to 0.75 of Nyquist frequency, other parameters are distributed inside reasonable bounds, but only non-zero AM/FM cases are shown in presented figures, as they are the most general ones. To estimate the quality of derivation filter, 2 versions of DAM were tested: derivation filter estimated derivative (ED) method and theoretic derivative (TD) method, where signal derivative is exact, taken directly from analytical definitions of signals. Such approach is taken, because derivation filter is assumed to be improved in further research, so TD method provides a lowest error performance, that ED method could achieve. It is very important to note, that RM and DAM method were proven to be theoretically equal [22], but perform differently in practice due to quite different approaches. In both methods it is required to compute a derivative: in RM case we compute the derivative of window and in DAM case we compute the derivative of the signal. In both cases this causes errors and they perform differently in practice.

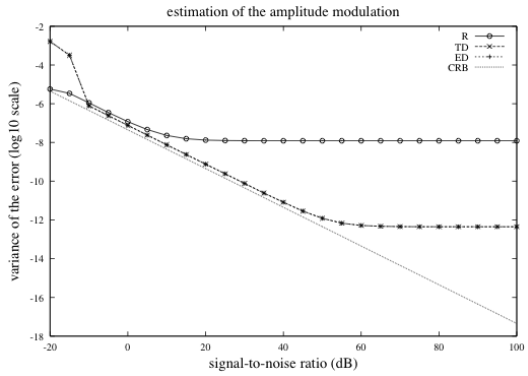
2.2 Quadratic Phase/Chirp transform (QPT)

Section 2.1 describes different FFT based approaches to detecting AM/FM modulation inside an analysis frame. Specifically, linear modulation was considered. However, another way of looking at the linear AM/FM estimation is, as a Maximum Likelihood (ML) problem [1]. Most of research in this field has been done by radar and seismic research community, therefore direct applications to sound analysis should be taken with care, as conditions are quite different. QPT can be, in its basic version, seen as generalized FT in a sense, that also gives estimates for linear FM chirps, eg: magnitude spectrum of QPT is 2D matrix, opposed to a vector in FT case. QPT magnitude plots usually give frequency on x-axis and FM rate on y-axis, as shown in figure 2.2, so to estimate FM and starting frequency, a peak of surface in 2Dt. Figure 2.2 was computed via *brute force* method, simply computing FFT for each chirp rate, which involves high computational complexity.

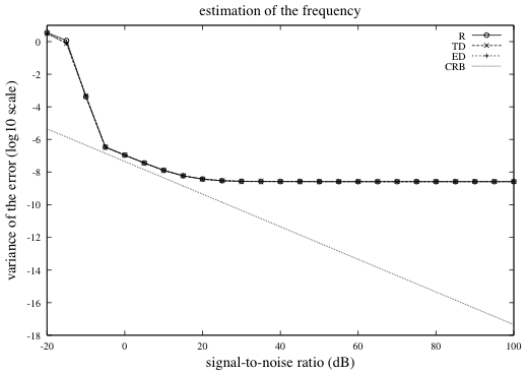
However, a fast algorithm, utilizing structures inherent to the transform, including optimizations similar to those in FFT algorithm [17], exists. Note, that AM case was not yet considered using QPT, but it is obvious that methods from section 2.1 could most likely be generalized to QPT. Intuitively, crossing chirps should pose no significant problem for QPT, as *up*-chirps and *down*-chirps should exhibit peaks on opposite sides of plane. However, completely symmetric



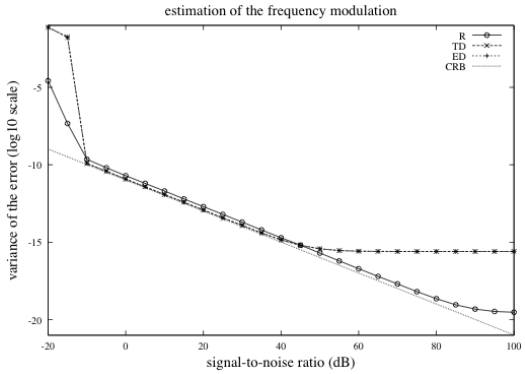
(a) Amplitude estimation error



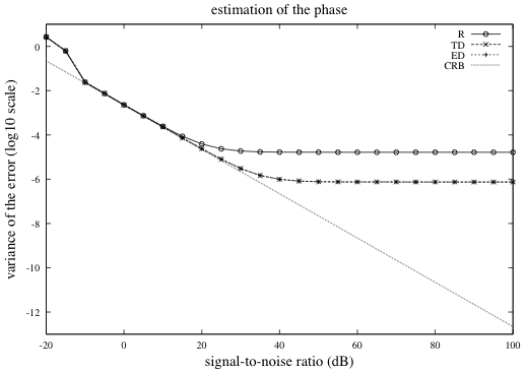
(b) Amplitude modulation estimation error



(c) Frequency estimation error



(d) Frequency modulation estimation error



(e) Phase estimation error

Figure 3: Comparison between reassignment (R) and 2 versions of DA (TD, ED) compared to Cramer-Rao bound (CRB), taken from [22]

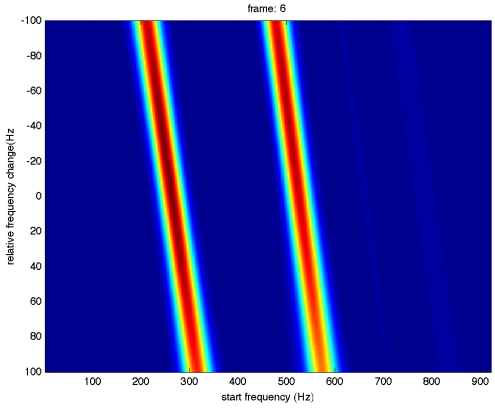
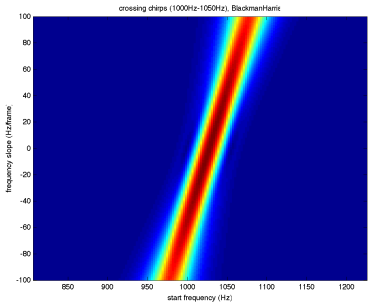
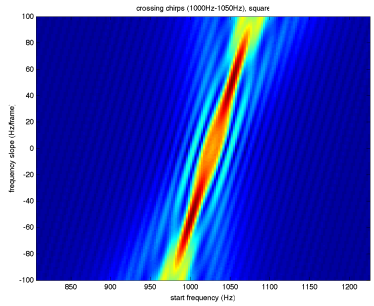


Figure 4: QPT plot of first 2 partials of violin glissando. Plot shows 2 peaks, that are approximately harmonic in frequency. The plot confirms, that analyzed frame actually covers a non-stationary part of sound, as peaks are not located at 0 FM value, but slightly in the negative FM rate region, suggesting decreasing frequency glissando. FM rate of second peak is approximately double as the FM rate of first peak, confirming the fact, that violin partials are nearly harmonic in frequency.

chirps (that is: frequency rate of first chirp is negative of the second and start frequency of first is the end frequency of the second) show very misleading plot, as peaks vanish completely, when frequency rate difference is small enough. Unfortunately, frequency change of 50Hz/frame is already small enough, to render QPT incapable of resolving crossing chirps, as shown in 5 (note, that 50Hz difference at higher frequencies represent a very low frequency change and can be expected in real world applications). Such shortcoming was expected, as QPT is theoretically not very different from FFT. Yet, it seems that battle is not completely lost, as QPT of non-smoothed (square windowed), shown in 5(b) indeed shows 2 peaks at expected locations. In spite of severe interference terms it's important to recognize, that interference terms form very prominent peaks and not so prominent *ridges*, which might just be enough to avoid detection of non-existant chirps. However, in more realistic example, using chirps that do not have the same absolute frequency rate, downside of this method seems to be unavoidable, as very significant, yet spurious peaks, whose location depends on window used, are produced in realistic case of 2 non-symmetric chirps, as demonstrated in figure 6. Unfortunately, QPT doesn't seem to be able to completely solve our (even) linear AM/FM problem.

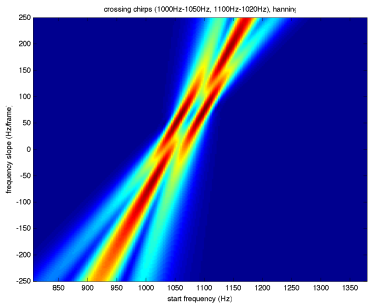


(a) Blackman-Harris window. Peaks at 50Hz/frame are smeared out.

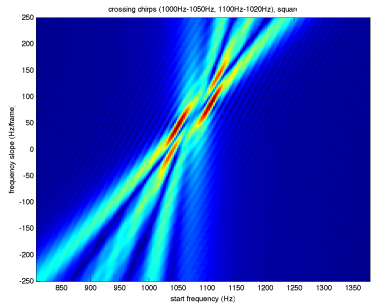


(b) Square window. Peaks at 50Hz are clearly visible, although severe side-lobe interference compared to (a) is present.

Figure 5: QPT plot of 2 crossing chirps using different windows. Start, end frequencies are 1000/1050Hz. Interestingly, QPT of non-smoothed (square window) frame is indeed able to recognize 2 chirps as opposed to Blackman-Harris windowed frame (note that use of Hanning window gives very similar results to Blackman-Harris)



(a) Hanning window. Four prominent peaks, instead of two.



(b) Square window. Four prominent peaks, instead of two.

Figure 6: QPT plot of 2 crossing chirps using different windows. Start/end frequencies for first/second partial are 1000/1050Hz, 1200/1020Hz respectively. In both cases, 4 instead of 2 prominent peaks were detected. One might want to find some kind of pattern in peak locations, but this seems unlikely to work, as peak locations depend significantly on window used (see (a), (b)).

2.3 Wigner-Ville transform (WVT)

Non-stationarity estimation methods, discussed in previous sections (2.1, 2.2) implicitly share a common property: a necessity to use sliding window approach. Another drawback, mentioned several times in subsections of section 2.1, is a bias in estimate of parameters, most importantly frequency and FM. This bias lies fundamentally in FT and windowing approach and it is not yet completely analyzed, in spite of numerous efforts (for some examples see [2], [22], [6]). These and many other facts have motivated research of alternative methods in non-stationarity estimates for sound and music.

A promising set of distributions are Cohen class time-frequency distributions [9]. One of its conveniences is broad area of distributions that it covers. In fact, it was proven in [4], that STFT belongs to Cohen class, which makes it specially attractive, as analysis methods that are proven to apply to this class are therefore directly applicable to STFT, as for example reassignment method [4] (reassignment was *invented* based on STFT [19] and was generalized to other time-frequency [4] like Cohen class later, but it could be the other way around as well). A very basic, but important Cohen class distribution is the Wigner-Ville [28] distribution, which is the simplest distribution in Cohen class. Other Cohen [9],[31] distributions could be viewed as smoothed WVT, which therefore possesses the highest time-frequency resolution. Unfortunately, its use in practice involving multicomponent signals (in present context, multi sinusoid/partial signal is assumed) is limited by cross-terms, resulting from interactions between components present in signal. Therefore, power of WVT is limited to mono-component signals. It is imperative to note, that most theoretical work on methods in 2.1 was done with single sinusoid case in mind as well, but it is often assumed, that effect of other partials can be neglected (possibly with use of correct window) and simple iterative procedure for dealing with multicomponent signal can be taken. In WVT however, number of cross-terms is N(N-1), if N is number of components. In time-frequency distribution, a cross-term is located at mean frequency and mean time, which is extremely inconvenient for harmonic signals, as cross-terms overlap with partials in WVT plot, as demonstrated in figure 7. Use of WVT has recently been used in estimating FM of sound samples [18]. Cross terms are eliminated by band pass filter bank, where each filter is centered at FFT peak, with cut-off frequency half way to neighboring peaks. Ideally we would wish, that result from such filtering would be a set of mono-component signals. However, in most general case, eg: a polyphonic audio recording, crossing chirp signals could still be contained signals after filtering and WVT of such signal type is shown in figure 8. As time-frequency planes can be considered images, pattern recognition techniques can be used to disambiguate cases like the one in figure 8. Indeed, it has been shown in [5], that Hough transform (HT) [15] efficiently extracts straight lines and suppresses modulated parts of time-frequency plane, which is exactly the procedure taken in [18]. However, accuracy of frequency, amplitude and AM/FM is not reported, as method in [18] is mainly used for improved partial tracking. So general parameter estimate quality comparison to other methods remains un-

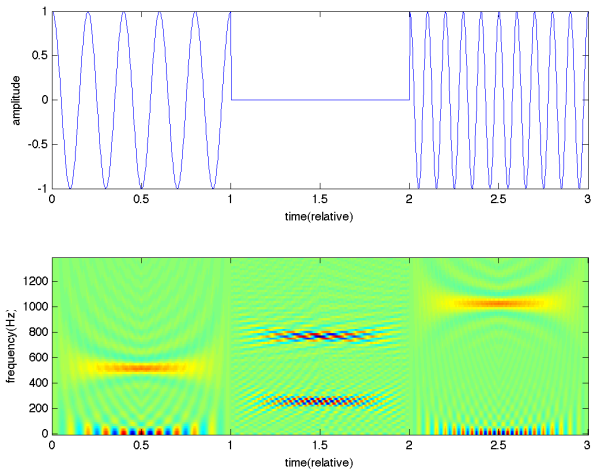


Figure 7: Upper plot: signal in time domain, bottom plot: WVT. Two stable sinusoids, 500/1000Hz,1 unit time span, separated with 1 unit time span. The middle region, from 1-2 time units should have no energy (upper plot), but severe cross-terms are observed in WVT (bottom plot). Cross-terms are however considerably more modulated then the terms, representing signal energy.

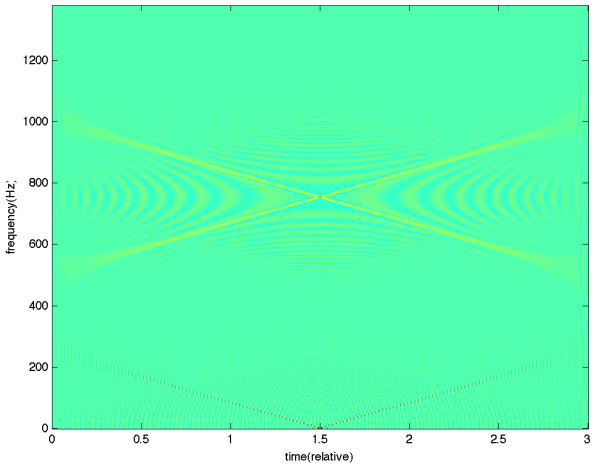


Figure 8: WVT of crossing linear chirps

done. However, it's main advantage to FFT based approaches (section 2.1) is robustness in presence of crossing chirps, so it provides strong tool for analysis of polyphonic signals, as this scenario is most likely to be observed in such cases. Note, that WVT seems to be the only well established analysis method so far, that can successfully separate crossing chirps. Still, cross influence of 2 chirps should be studied in more detail and amplitude and AM estimates should be evaluated. Importantly, WVT introduces great computational complexity and memory requirements.

3 Reassignment and derivative analysis method implementation

3.1 General considerations

Reassignment and derivative analysis deserve a special place in non-stationary sinusoidal analysis. Currently the most accurate and the only methods, able to detect linear AM/FM in a computationally very low manner, making them attractive for real-time applications. Signals under investigation is considered non-stationary, frequency and amplitude are assumed to change linearly inside one analysis frame. This way, instantaneous phase, amplitude and corresponding signal can be written as functions of time in the following way

$$\varphi(t) = \phi_0 + \omega_0 t + \frac{\psi_0}{2} t^2 \tag{1}$$

$$a(t) = \lambda_0 + \mu_0 t \tag{2}$$

$$s(t) = \exp(a(t) + i\varphi(t)) \tag{3}$$

Definition of Short Time Fourier Transform (STFT), a function S_w of time t and frequency ω:

$$S_w(t, \omega) = \int_{-\infty}^{\infty} s(\tau) w(\tau - t) exp(-i\omega(\tau - t)) d\tau \tag{4}$$

,where $w(t)$ is a windowing function. Present definition is slightly different from usual: the time reference slides with the window, creating a phase shift of $-\omega t$, which is of no major concern.
Each partial is represented as Fourier transform of window function, centered at ω_0 multiplied by $s_0 = \exp(\lambda + i\phi)$. This results can be derived in the following

way. [22] with:

$$\mathfrak{F}\{s(t)w(t)\} = \exp(\lambda + i\phi)\mathfrak{F}\{\mu_0 t + i\left(\omega_0 t + \frac{\psi_0}{2}t^2\right)\} \tag{5}$$

$$= s_0 \int_{-\infty}^{+\infty} w(t) \exp\left(\mu_0 t + i\left((\omega_0 - \omega)t + \frac{\psi_0}{2}t^2\right)\right) dt \tag{6}$$

$$\Gamma(\omega, \mu_0, \psi_0) = \int_{-\infty}^{+\infty} w(t) \exp\left(\mu_0 t + i\left((\omega_0 - \omega)t + \frac{\psi_0}{2}t^2\right)\right) dt \Rightarrow \tag{7}$$

$$\mathfrak{F}\{s(t)w(t)\} = \exp(\lambda + i\phi)\Gamma(\omega_0 - \omega, \mu_0, \psi_0) = X(\omega) \tag{8}$$

Generally, all parameters apart from phase and amplitude can be incorporated in Γ. By substituting $\omega_\Delta = \omega_0 - \omega$, we can rewrite the Γ function as following:

$$\Gamma(\omega_\Delta, \mu_0, \psi_0) = \int_{-\infty}^{+\infty} w(t) \exp\left(\mu_0 t + i\left(\omega_\Delta t + \frac{\psi_0}{2}t^2\right)\right) dt \tag{9}$$

This integral only has a simple analytical solution in the case of Gauss window [2]. For more common cosine based window functions such integrals pose quite a challenge. We can look at Γ as a kind of 'modified' spectrum of window. Obviously, this 'modification' depends on AM/FM, whereas the frequency of sinusoid only defines the central location in spectrum. That is, we should be able to observe an image of $\Gamma(\omega_\Delta, \mu_0, \psi_0)$, centered at ω_0 in magnitude spectrum. Thus, if AM/FM values can somehow be estimated from the signal, than it is possible to use this transformed window extract amplitude and phase values (λ, ϕ).

3.2 DAM

DAM was only recently (2008) generalized for use in non-stationary conditions in [22]. The main idea of DAM is to consider signal derivatives and its spectrums. First and second derivative of signal defined by (3) are:

$$s'(t) = (\mu_0 + i(\omega_0 + \psi t))s(t) \tag{10}$$

$$s''(t) = ((\mu_0^2 - \omega_0^2 - 2\omega_0\psi_0 t - \psi_0^2 t^2) + i(\psi_0 + 2\mu_0\omega_0 t))s(t) \tag{11}$$

Considering STFT of that and with assumptions from [22]:

- at spectral peaks, contributions of all terms, that depend on t are negligible
- even/odd functions have real/imaginary spectrums respectively

Complete theoretical derivation is out of scope of this work and can be found in [22]. We can conclude, that for FFT bin of frequency ω_k closest to the peak of the partial the following equation will give a very good estimate to the frequency of the partial:

$$\hat{\omega}_0 = \Im\left(\frac{S'_w(\omega_k)}{S_w(\omega_k)}\right) \tag{12}$$

One could argue, that spectral peak FFT bin frequency is a good approximation of partial frequency. Unfortunately, it was shown in [6], that spectral peak is shifted in frequency, when frequency and/or amplitude modulation is present. Therefore, above equation should provide a way to bypass effects of AM/FM at least to some extent. Even in stationary cases, equation 12 helps to improve frequency resolution, as it gives frequency estimate more accurate than FFT resolution allows [10]. It is trivial to express AM estimate:

$$\hat{\mu}_0 = \Re\left(\frac{S'_w(\hat{\omega}_0)}{S_w(\hat{\omega}_0)}\right) \tag{13}$$

Using second derivative and same assumptions, frequency modulation can be estimated by:

$$\hat{\psi}_0 = \Im\left(\frac{S''_w(\hat{\omega}_0)}{S_w(\hat{\omega}_0)}\right) - \hat{\mu}_0\hat{\omega}_0 \tag{14}$$

We assume, that our estimate of frequency partial $\hat{\omega}_0$ is very close to the actual frequency of the partial and since we used that estimate to calculate AM and FM, the the spectrum S_w can be 'corrected' with $\Gamma(0, \hat{\mu}_0, \hat{\psi}_0)$ (eg: $\omega \approx 0$). Now, we can finally estimate amplitude and phase:

$$\hat{\lambda}_0 = \left|\frac{S_w(\hat{\omega}_0)}{\Gamma(0, \hat{\mu}_0, \hat{\psi}_0)}\right| \tag{15}$$

$$\hat{\phi}_0 = \angle\left(\frac{S_w(\hat{\omega}_0)}{\Gamma(0, \hat{\mu}_0, \hat{\psi}_0)}\right) \tag{16}$$

3.3 RM

Reassignment was generalized by Auger and Flandrin in [4]. We can write any spectrum as product of amplitude and phase functions of frequency:

$$S_w(t,\omega) = \exp(\lambda(t,\omega) + i\phi(t,\omega)) \Rightarrow \tag{17}$$

$$\log(S_w(t,\omega)) = \lambda(t,\omega) + i\phi(t,\omega) \Rightarrow \tag{18}$$

$$\Im(\log(S_w(t,\omega))) = \phi(t,\omega), \Re(\log(S_w(t,\omega))) = \lambda(t,\omega) \tag{19}$$

Intuitively, computing first order time derivative of phase results in frequency, second order time derivative of phase results in linear frequency modulation and time derivative of amplitude results in linear amplitude modulation:

$$\hat{\omega}(t,\omega) = \frac{\partial}{\partial t}\phi(t,\omega) = \Im\left(\frac{\partial}{\partial t}S_w(t,\omega)\right) = \omega - \Im\left(\frac{S_{w'}(t,\omega)}{S_w(t,\omega)}\right) \tag{20}$$

$$\hat{\mu}(t,\omega) = \frac{\partial}{\partial t}\lambda(t,\omega) = \Re\left(\frac{\partial}{\partial t}S_w(t,\omega)\right) = -\Re\left(\frac{S_{w'}(t,\omega)}{S_w(t,\omega)}\right) \tag{21}$$

$$\hat{\psi}(t,\omega) = \frac{\partial^2}{\partial t^2}\phi(t,\omega) = \Im\left(\frac{S_{w''}(t,\omega)}{S_w(t,\omega)}\right) - \Im\left(\left(\frac{S_{w'}(t,\omega)}{S_w(t,\omega)}\right)^2\right) \tag{22}$$

Complete mathematical derivations are out of scope of this work and can be found in [14]. In practice, discrete spectral peak frequency ω_m is used:

$$\hat{\omega}_0 = \hat{\omega}(t, \omega_m), \hat{\mu}_0 = \hat{\mu}(t, \omega_m), \hat{\psi}_0 = \hat{\psi}(t, \omega_m) \tag{23}$$

However, slight improvement (inspired by DAM method) is to use estimated frequency $\hat{\omega}_0$ when computing AM,FM estimates:

$$\hat{\omega}_0 = \hat{\omega}(t, \omega_m), \hat{\mu}_0 = \hat{\mu}(t, \hat{\omega}_0), \hat{\psi}_0 = \hat{\psi}(t, \hat{\omega}_0) \tag{24}$$

Again, estimate of frequency $\hat{\omega}_0$ is very close to the actual frequency and it was used to calculate AM and FM. Therefore we use the same spectrum S_w 'correction' as in DAM (eg: $\omega \approx 0$):

$$\hat{\lambda}_0 = \left| \frac{S_w(\hat{\omega}_0)}{\Gamma(0, \hat{\mu}_0, \hat{\psi}_0)} \right| \tag{25}$$

$$\hat{\phi}_0 = \angle \left(\frac{S_w(\hat{\omega}_0)}{\Gamma(0, \hat{\mu}_0, \hat{\psi}_0)} \right) \tag{26}$$

It's important to realize, that such minor improvement might affect AM and FM estimates significantly in cases, when FFT bins are large, eg: for small windows. In such cases the frequency estimate can differ significantly from nearest FFT bin frequencies, thus causing significant differences in AM/FM estimates.
Hainsworth has shown in [14] that in practice, RM exhibits an estimate bias. It is not present in formulation of RM, but is introduced in practice when discrete formulation is used. Bias depends on window function used and the distance of actual partial frequency to nearest FFT bin. It could be decreased by technique, described in [14] section 3.4. however, an original version of RM was considered in this study.

3.4 Implementation specifics

3.4.1 Spectrum values at non FFT bin frequencies

First step of both algorithms is estimating frequency of the partial. This estimates are then used to calculate AM/FM and eventually amplitude and phase estimates. But such procedure requires calculations of various spectrum values at non FFT bin frequencies and are thus unknown at that point. It is possible to interpolate spectrum values and thus acquire a more or less accurate estimate of spectrum at arbitrary frequency. However, acquired frequency estimate $\hat{\omega}_0$ can be used to compute DFT at that exact frequency for some extra computational cost with means of DFT:

$$S_w(\hat{\omega}_0, t) = \sum_{n=0}^{N-1} s[n] w[n - tF_S] \exp(-i\hat{\omega}_0 \frac{n}{F_S}) \tag{27}$$

,where N is length window s[n], w[n] are discrete variants of signal and window function respectively, and F_S is sampling frequency. Such approach was used in both DAM and RM implementations used for present study.

3.4.2 Spectrum correction

In subsection 3.1, we defined effect of non-stationarity defined by equation (8). For Gauss window, defined as $w(t) = \sqrt{\frac{p}{\pi}} e^{-pt^2}$ analytical solution of equation (8) would yield [2]:

$$\Gamma(\omega, \mu_0, \psi_0) = \exp(u(\omega, \mu_0, \psi_0) + iv(\omega, \mu_0, \psi_0)), \tag{28}$$

$$u(\omega, \mu_0, \psi_0) = \frac{\mu_0^2}{4p} - \frac{1}{4}\log\left[1 + \left(\frac{\psi_0}{p}\right)^2\right] - \frac{p}{4(p^2 + \psi_0^2)}\left(\omega - \omega_0 + \frac{\psi_0\mu_0}{p}\right)^2, \tag{29}$$

$$v(\omega, \mu_0, \psi_0) = \frac{\mu_0^2}{4\psi_0} + \frac{1}{2}\arctan\left(\frac{\psi_0}{p}\right) - \frac{p}{4(p^2 + \psi_0^2)}\left(\omega - \omega_0 + \frac{p\mu_0}{\psi_0}\right)^2 \tag{30}$$

A maximum in amplitude spectrum would than be located at:

$$\hat{\omega}_0 = argmax|(\lambda_0 + i\phi)\Gamma(\omega, \mu_0, \psi_0)| = \omega_0 + \frac{\mu_0\psi_0}{p} \tag{31}$$

Analytically expressed bias of frequency estimator is very valuable, as it offers potential to reduce the bias. For more common cosine based window functions such integrals are more complicated. A cosine based window function is defined by:

$$w(t) = \sum_{k=0}^{K} a_k \cos(2\pi kt) \tag{32}$$

Such windows are only defined for $t = -\frac{1}{2}...+\frac{1}{2}$ and are 0 outside that region. In practice window can be stretched in time domain as much as needed, therefore integral in (8) can be written as:

$$\Gamma(\omega, \mu_0, \psi_0) = \int_{-\frac{T}{2}}^{+\frac{T}{2}} w(t)\exp\left(\mu_0 t + j\left(\omega t + \frac{\psi_0}{2}t^2\right)\right)dt \tag{33}$$

,where T is length of window. In such case, it is a difficult task to compute an integral by hand, but *Wofram research Mathematica* gives the following result:

$$\Gamma(\omega,\mu_0,\psi_0) =$$
$$\sum_{k=0}^{K}\frac{1}{\sqrt{\psi_0}}\left(\frac{1}{4}+\frac{i}{4}\right)a_k e^{\frac{i(2ik\pi+T(\mu_0+i\omega))^2}{2T^2\psi_0}}\sqrt{\pi}$$
$$\left(\mathrm{Erf}\left[\frac{\left(\frac{1}{4}-\frac{i}{4}\right)(4k\pi+T(-2i\mu_0+T\psi_0+2\omega))}{T\sqrt{\psi_0}}\right]-\mathrm{Erf}\left[-\frac{\left(\frac{1}{4}-\frac{i}{4}\right)(-4k\pi+T(2i\mu_0+T\psi_0-2\omega))}{T\sqrt{\psi_0}}\right]\right.$$
$$+e^{\frac{4k\pi(\mu_0+i\omega)}{T\psi_0}}\left(\mathrm{Erf}\left[\frac{\left(\frac{1}{4}-\frac{i}{4}\right)(4k\pi+T(2i\mu_0+T\psi_0-2\omega))}{T\sqrt{\psi_0}}\right]-\mathrm{Erf}\left[-\frac{\left(\frac{1}{4}-\frac{i}{4}\right)(-4k\pi+T(-2i\mu_0+T\psi_0+2\omega))}{T\sqrt{\psi_0}}\right]\right. \tag{34}$$

,where Erf is complex error function. In case of $\psi_0 = 0$ above term results in an undefined value (multiplication of 0 and infinity) and we have to compute the integral again, with assumption $\psi_0 = 0$:

$$\Gamma(\omega,\mu_0,0) = \sum_{k=0}^{K}\frac{a_k e^{-\frac{1}{2}T(\mu_0+i\omega)}T^2\left(e^{T(\mu_0+i\omega)}-1\right)(\mu_0+i\omega)\cos(k\pi)}{4k^2\pi^2+T^2(\mu_0+i\omega)^2} \tag{35}$$

Again, above term is not defined when $k = 0, \omega = 0, \mu_0 = 0$ therefore we require another integral for this case:

$$\Gamma(0,0,0) = a_0 T \tag{36}$$

With above equations amplitude and phase corrections can be calculated very precisely. Unfortunately, it is quite difficult to analytically derive an equation that would show, how much a spectrum peak is shifted away from actual frequency, as was done for Gauss window, see equation (31).
RM requires first and second order derivatives of the window function, which makes common windows like Hanning, Hamming rather unsuitable, as their derivatives exhibit poor side lobe rejection ratios [25]. Proposed squared Hanning window implies calculation of integral [8] with $cos^2(x)$ terms. This yields similar result as equations (34), (35), (36), but it contains a bit more terms and will be, for the sake of readiness, omitted from the text.
In practice parameter estimates never reach exact 0, but small values will cause numerical errors or force algorithm to multiply zero and infinity (in case of MATLAB). Some reasonable small thresholds for parameters should be used in order to avoid such situations. In present implementation, the following thresholds were used:

$$\psi_T = 0.5, \mu_T = 0.01, w_T = 0.01 \tag{37}$$

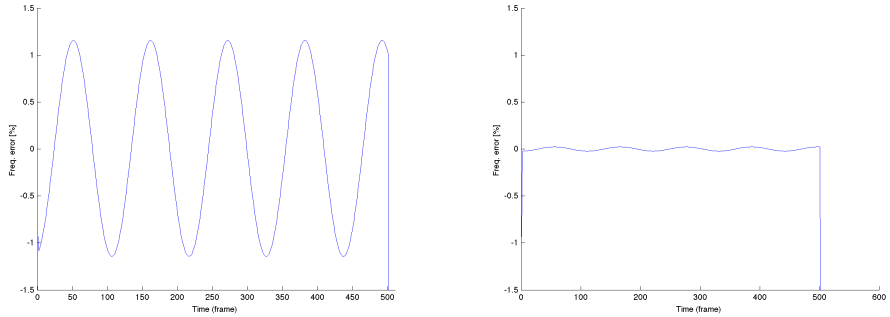
For all DAM analysis Hanning window was used and in all RM analysis squared Hanning window was used.

4 Comparison of reassignment and derivative method analysis

4.1 Frequency

4.1.1 DAM

In stationary case, DAM frequency estimate accuracy without zero padding reaches around 1% in lower and about 3% in high frequencies. Zero padding significantly improves estimate, demonstrated in figures 9(a), 9(b). Although error exhibits stable oscillation, more extensive tests have shown, that the error is not biased significantly, as figures 9(a), 9(b) suggest. In fact, most of frequency error is produced by variance, rather than bias. Further, it is evident that error is correlated with exact location of window with respect to sinusoid. It is important to note, that amplitude of error oscillations can be decreased with zero-padding or increased window length, but frequency of error oscillation remains constant even when window size is changed.
In non-linear AM/FM case frequency estimates remain very accurate, even for

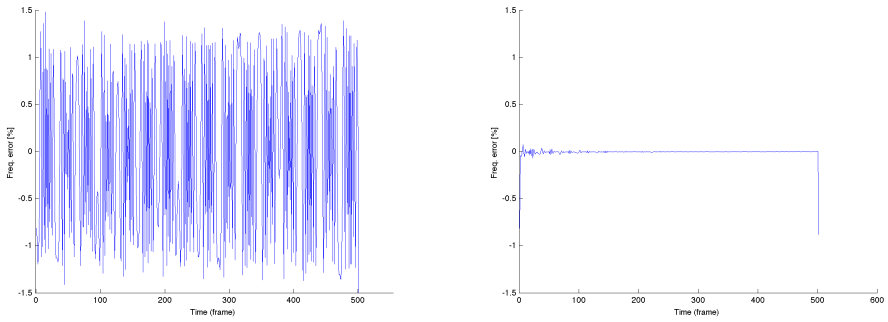


(a) No zero padding

(b) Zero padding factor: 2

Figure 9: DAM: Relative frequency estimate error for stable sinusoid of 1000Hz, 177 samples window (4 x wavelength), 1000ms duration, different zero paddings

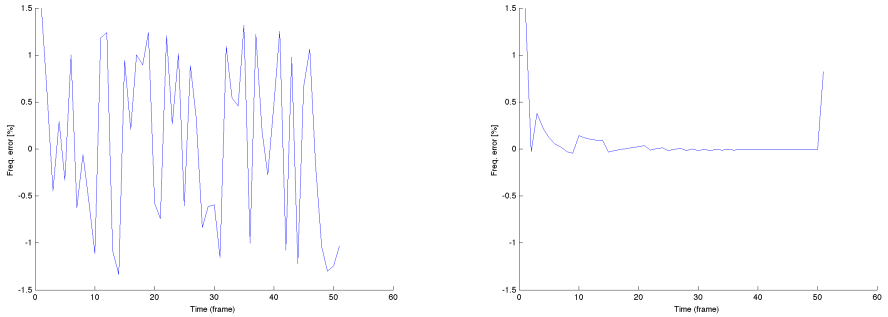
extreme modulations without zero padding. Examples are shown in figures 10(a), 10(b) and 11(a), 11(b). In figures 11(a), 11(b) signal is shortened to 100ms while maintaing the same absolute frequency change, non-linear AM is added and FM is also non-linear. Clearly, frequency estimate is still confined in 1.5% interval without zero-padding, thus we can conclude that variance of error does not change significantly. However, higher accuracy test using zero-padding reveals increased bias (figure 11(b)).
There are at least 2 reasons, why such bias occurs. One of them is phenomena, observed in [6] and analytically derived in [2], although only for Gauss window. Spectral peak is not located exactly at middle frame frequency, when AM/FM sinusoid is analyzed, so even very accurate peak detection technique



(a) No zero padding

(b) Zero padding factor: 2

Figure 10: DAM: Relative frequency estimate error for linear FM chirp, 1000-5000Hz, 177 samples window (4 x wavelength of minimal frequency), 1000ms duration, different zero paddings



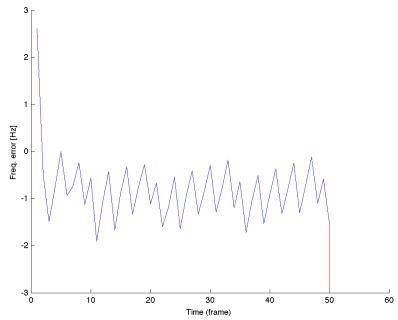
(a) No zero padding

(b) Zero padding factor: 2

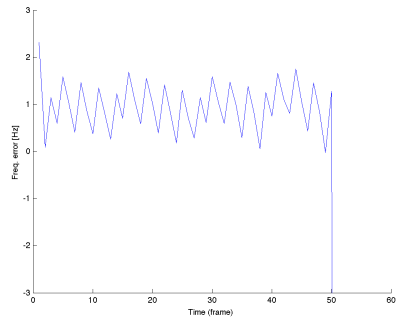
Figure 11: DAM: Relative frequency estimate error for quadratic FM, 1000-5000Hz, quadratic AM, log amplitude 0-5, 177 samples window (4 x wavelength of minimal frequency), 100ms duration, different zero paddings

would not estimate frequency correctly. Since DAM uses imaginary parts of signal and signal derivative spectrum for frequency estimate, it is not straight forward to say, that estimate will be biased in some way (eg: positive negative), if (positive/negative) spectral shift is introduced. For this purpose, absolute frequency errors will be studied. From figures 12(a), 12(b), 12(d), 12(c) we can conclude, that positive FM causes negative bias and vice versa, whereas positive AM causes positive bias and vice versa.

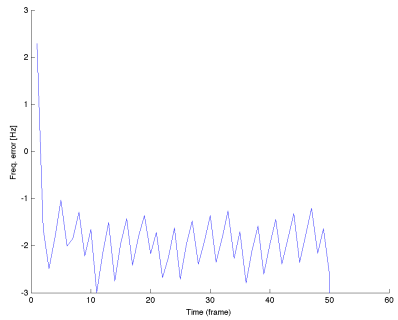
Further, errors depend solely on AM/FM and not on frequency, which makes DAM method less suitable for low frequencies, however signal derivative approximations forces it to make larger errors in higher frequency range as well, as described in [22]. Important note is, that frequency estimate bias seems to be



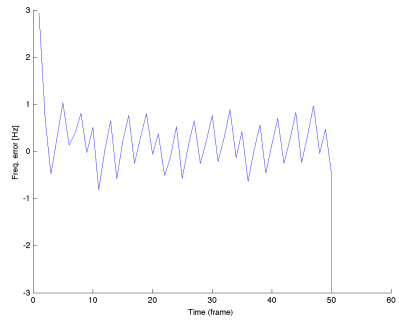
(a) Positive FM, no AM



(b) Negative FM, no AM



(c) Positive FM, negative AM



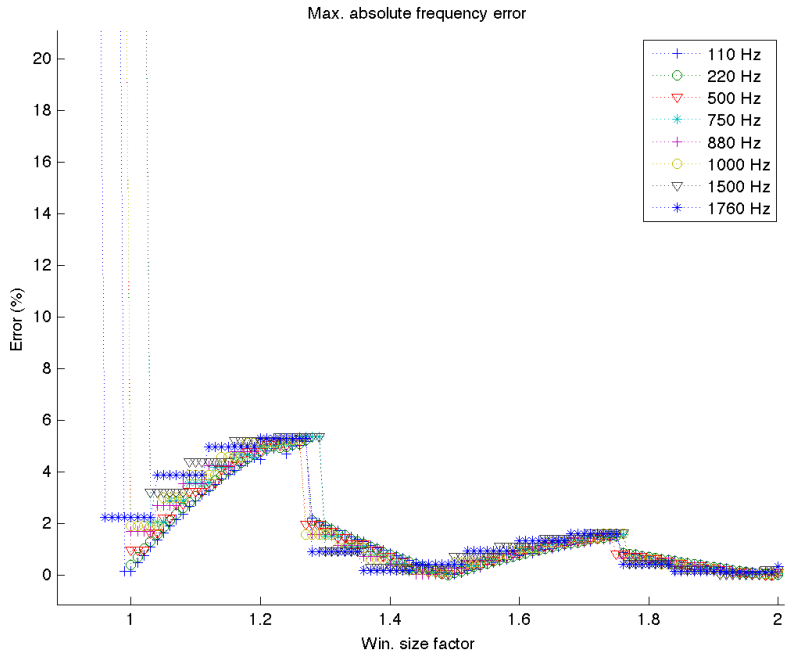
(d) Positive FM, positive AM

Figure 12: DAM: Absolute frequency estimate error for linear FM, frequency range 1000-5000Hz, linear AM 5, 177 samples window (4 x wavelength of minimal frequency), 100ms duration, zero padding factor 2

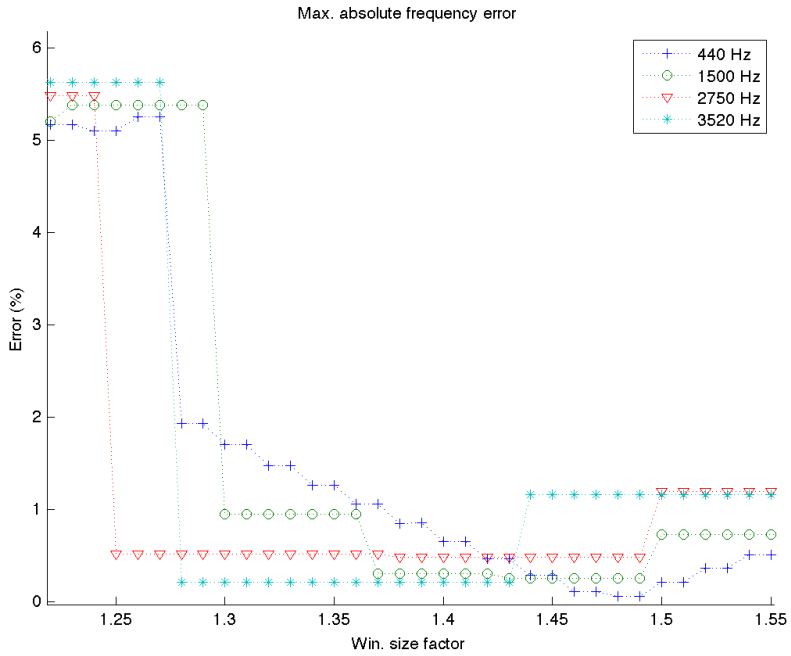
dependent on AM and FM in manner, suggesting linear independence of both effects. For certain combinations of AM/FM, eg: positive AM and positive FM the bias is decreased significantly, see figure 12(d). This differs from analytical solution derived in [2], defined by equation (31) where clearly, estimate bias is not a linear function of AM and FM. However, above result depends on window function and cannot be generalized. Unfortunately, such calculations cannot easily be done for more cosine type window functions, as it was shown in subsection 3.4.2.

Generally attractive property of any spectral analysis technique is high accuracy in as wide band conditions as possible. Frequency accuracy of DAM method was tested for window sizes ranging from narrow to wide band conditions. Figure 13(a) shows maximum relative error, that DAM will make for different window sizes, all in wide band window size range.

Very interestingly, window sizes at around 1.3-1.5 of period length seem to be a good trade off between window size and accuracy for nearly all frequen-



(a) Frequency error in wide band conditions for stable sinusoids of different frequencies.



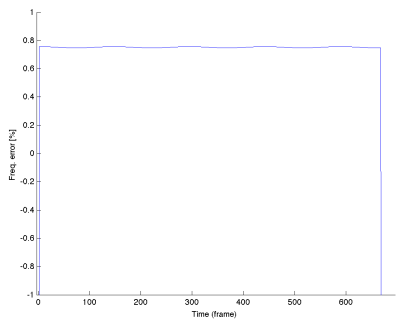
(b) Maximum relative error in wide band conditions for stable sinusoids of different frequencies, zoom in.

Figure 13: DAM: Frequency errors in wide band conditions.

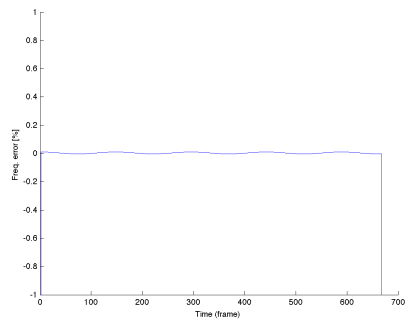
cies. It can be observed (figure 13(b)), that for higher frequency sinusoids a slightly shorter window sizes of 1.3 fundamental period are already acceptable. Approaching very high frequencies (close to Nyquist) such investigation of frequency accuracy in wide band becomes more and more difficult, as window sizes shrink down to several samples and transition from wide band to narrow band happens within few samples.

4.1.2 RM

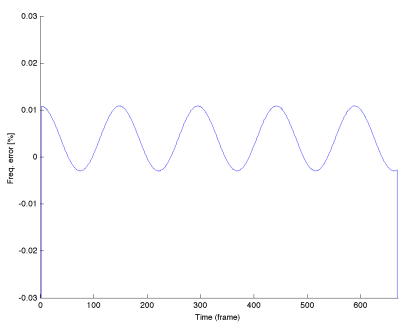
RM frequency estimate interestingly exhibits an inherent bias. Zero-padding however, reduces this bias dramatically, as is clearly seen on figures 14(a), 14(b) and 14(c). This phenomena were observed by Hainsworth and described in [14]. A steady oscillation of error exists and can be observed when zooming in significantly. There is no doubt that this oscillation is caused by the same phenomena as the one noticed using DAM, shown in figures 9(a) and 9(b). It seems that this is intrinsic property of Fourier transform, but further investigation is out of scope of this work.



(a) No zero padding



(b) Zero padding factor: 2

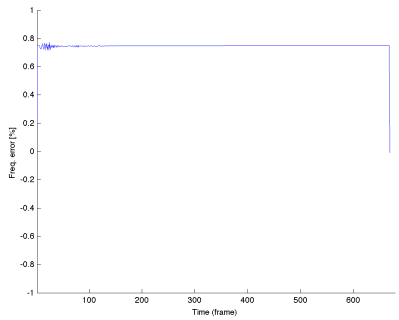


(c) Zero padding factor: 2, zoom in

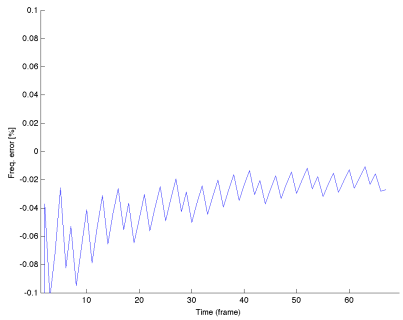
Figure 14: RM: Relative frequency estimate error for stable sinusoid of 1000Hz, no AM, 177 samples window (4 x wavelength), 1000ms duration, different zero paddings

Figures 15(a),15(b),15(c) show analysis of chirp signal. Plots confirm, that estimate is biased and suggest, that estimate bias depends on frequency and on exact frequency position between adjacent FFT bins. This causes frequency error to exhibit abrupt jumps, as frequency passes middle of two adjacent FFT bins. Hence we can conclude, that for specific frequency estimate and zero padding, a unique frequency bias exist. This may lead to 'correction' function,

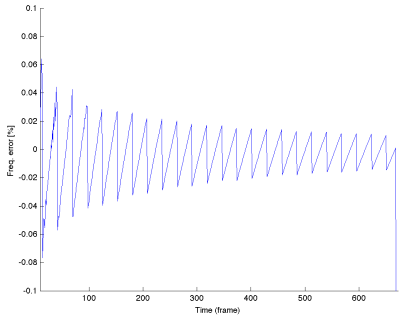
which might be able to completely remove bias. Indeed, such correction was already mentioned in subsection 3.3 and was proposed by Hainsworth in [14].



(a) No zero padding



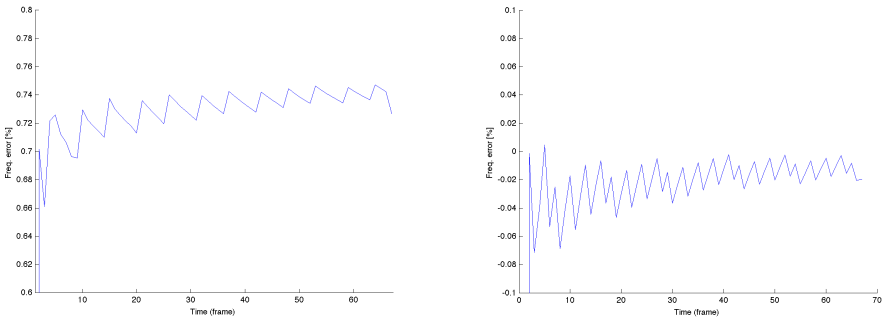
(b) Zero padding factor: 2



(c) Zero padding factor: 2, zoom in

Figure 15: RM: Relative frequency estimate error for linear FM chirp, 1000-5000Hz, no AM, 177 samples window (4 x wavelength), 1000ms duration, different zero paddings

Figures 16(a), 16(b) show analysis of the same signal shortened to 100ms while absolute frequency difference is kept the same, linear AM is added. Please note, that both figures have the same span in y-coordinate, but 16(a) is centered at 0.7, whereas 16(b) is centered at 0. Comparing these 2 plots we can conclude, that zero padding decreases bias, but does not (significantly) decrease amplitude of error oscillations. It is further evident, that increased FM and/or AM adds additional bias to frequency estimate.
Following plots of absolute frequency error are exactly the same as in DAM case and further reveal effect of AM/FM on frequency estimate. As expected, frequency bias behaves very similarly as in case of DAM (positive FM causes negative frequency bias and vice versa, positive AM causes positive frequency bias and vice versa), but it seems like AM has a bit weaker effect in RM case.
In wide band conditions, RM behaves similarly to DAM, see figure 18. First

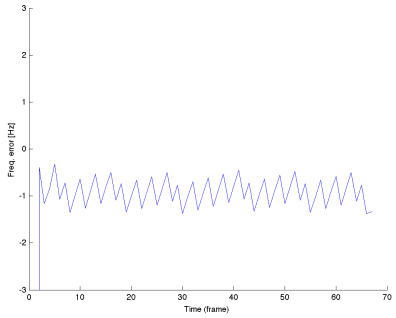


(a) No zero padding

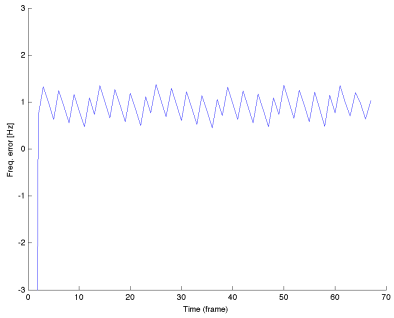
(b) Zero padding factor: 2

Figure 16: RM: Relative frequency estimate error for linear FM chirp, 1000-5000Hz, linear AM, log amplitude 0-5,177 samples window (4 x wavelength), 100ms duration, different zero paddings.

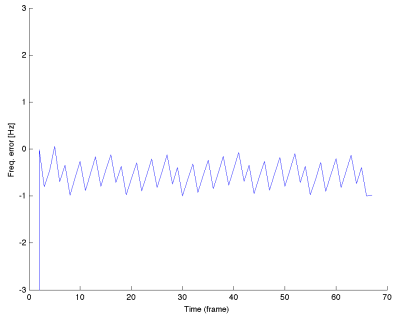
significant minimum for all frequencies is located around factor 1.5, which is slightly higher than in DAM case. Most probably, it depends on window function. However, it seems like there is always a significant minimum below window size double the length of period, which makes both methods good candidates for wide band analysis, as far as frequency estimate is considered.



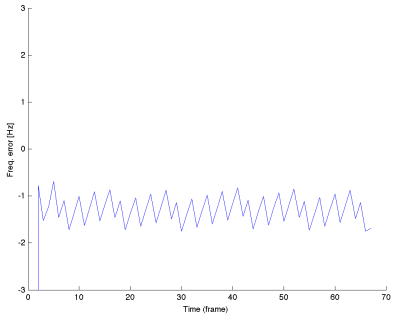
(a) Positive FM, no AM



(b) Negative FM, no AM



(c) Positive FM, positive AM



(d) Positive FM, negative AM

Figure 17: RM: Absolute frequency estimate error for linear FM, frequency range 1000-5000Hz, linear AM of 5, 177 samples window (4 x wavelength of minimal frequency), 100ms duration, zero padding factor 2

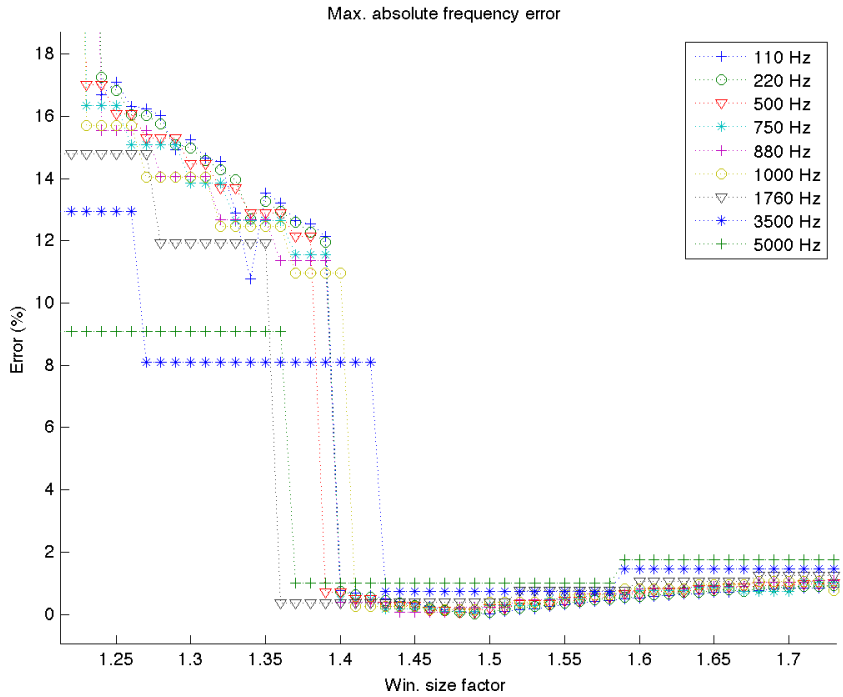


Figure 18: RM: Maximum relative error in narrow band conditions for stable sinusoids of different frequencies..

4.2 Amplitude Modulation

4.2.1 DAM

Amplitude modulation estimate exhibits static error oscillations, much like frequency estimate, as seen on figure 4.2.1. Amplitude of this error oscillation does change very little with actual AM value (see figure 4.2.1), eg: the error oscillation amplitude is nearly constant, forcing DAM to make larger relative errors for small AM values, at least for sinusoids of approximately 1000Hz. Comparing this plot to DAM frequency estimate plots for same conditions 9(a), 9(b) it seems straight forward to claim, that AM error is caused by frequency error, because frequency is directly used to calculate AM estimate and because of visual resemblence. In controlled conditions it is possible to replace frequency estimate with actual frequency and thus test exactly how much AM estimate depends on frequency estimate. Doing so reveals, that frequency estimate has actually very weak influence on AM estimate. Specifically, AM estimate error is reduced for less than 0.1%, which does not represent significant improvement.

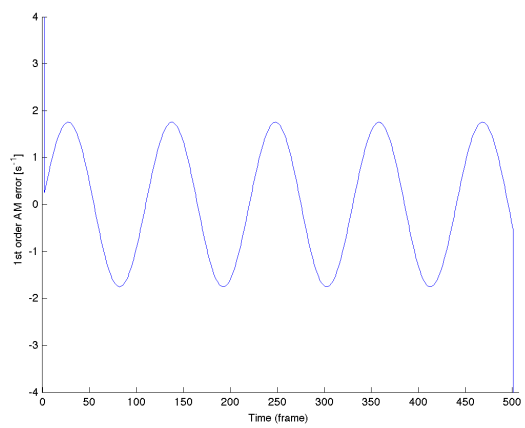


Figure 19: DAM: AM estimate error for stable sinusoid, 1000Hz, AM rate: 50, 177 samples window (4 x wavelength)

To further investigate effect of frequency on AM estimate, a more detailed look inside mechanics of DAM is needed. In figure 21 we see a plot of AM estimate (eg: the value of $\Re\left(\frac{S'_w}{S_w}(\omega)\right)$) with respect to frequency. Frequency estimate is very close to actual, but value of $\Re\left(\frac{S'_w}{S_w}(\omega)\right)$ at both frequencies is far from correct AM value. Note, that despite this plot was take from a random frame, it is representative and we can conclude, that frequency estimate error has a minor effect on AM estimate. However, FFT frequency bins can be very far apart for smaller windows, thus using frequency estimate instead of nearest FFT bin frequency seems to be a good procedure. Another interesting relationship is

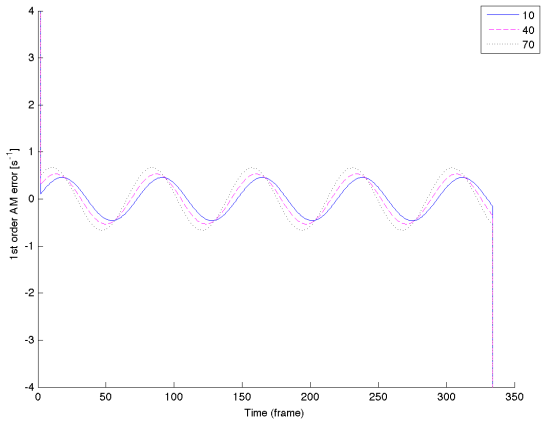


Figure 20: DAM: AM estimates for different AM rates, sinusoid 1000Hz, 177 sample window size (4 x wavelength), no FM, no zero padding

AM vs fundamental frequency. Figure 22 shows such graph for whole frequency band. Note, that for frequencies above 14500Hz errors increase drastically, as it is predicted for DAM. Similarly than in frequency estimation case, the only way to drastically improve AM accuracy is to enlarge window size. Figures 23(a) and 23(b) show maximum relative error that DAM makes for different frequencies, window sizes and 2 different AM rates. Please note, that in both figures some of frequencies may not enter 15% maximum error range are therefore not shown. We can see that relative errors can reach very high values, when AM value is lowered (figure 23(b)), confirming observation from beginning of this section: DAM make an absolute error in estimate, that is weakly correlated with actual AM value. This makes relative errors very high for small AM values. From comparing figures 23(a) and 23(b) it is further evident, that window size should be choose very carefully for low AM values (eg: window size should be exact multiple of half the wavelength), to avoid big relative AM errors. Interestingly, if window size is exactly multiple of half the wavelength, than relative AM error is roughly the same at both AM rates.

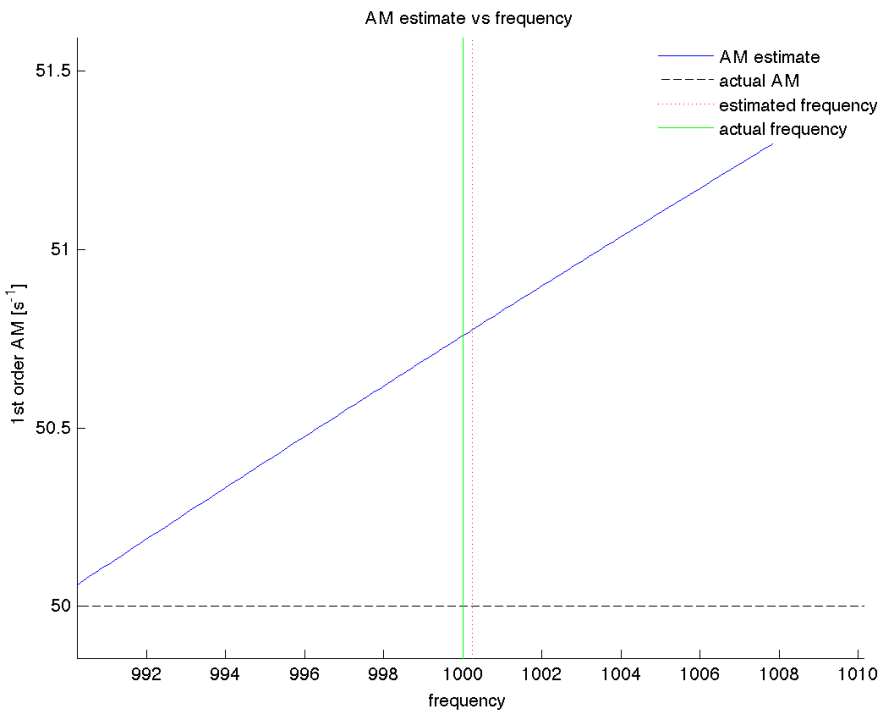


Figure 21: DAM: $\Re\left(\frac{S'_w}{S_w}(\omega)\right)$ for frequencies around estimated and actual frequency. Sinusoid of 1000Hz, no AM/FM, window size 177 (4 x fundamental period), zero padding factor 2

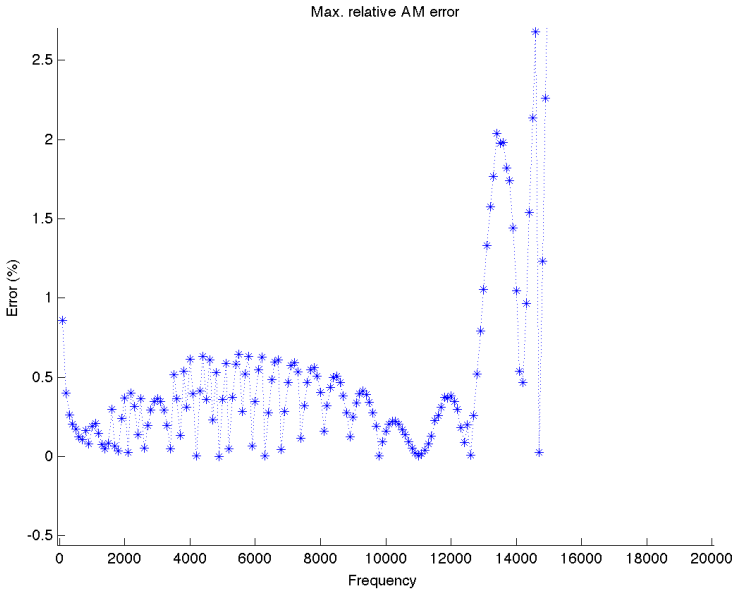
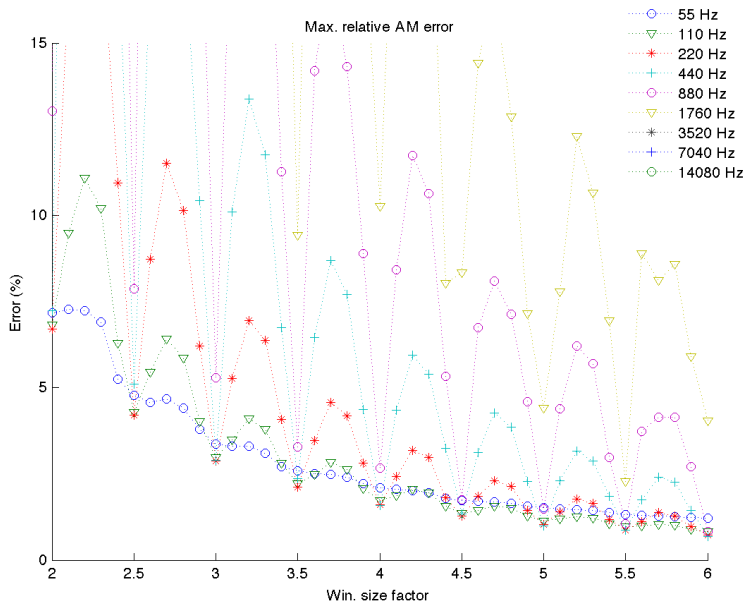
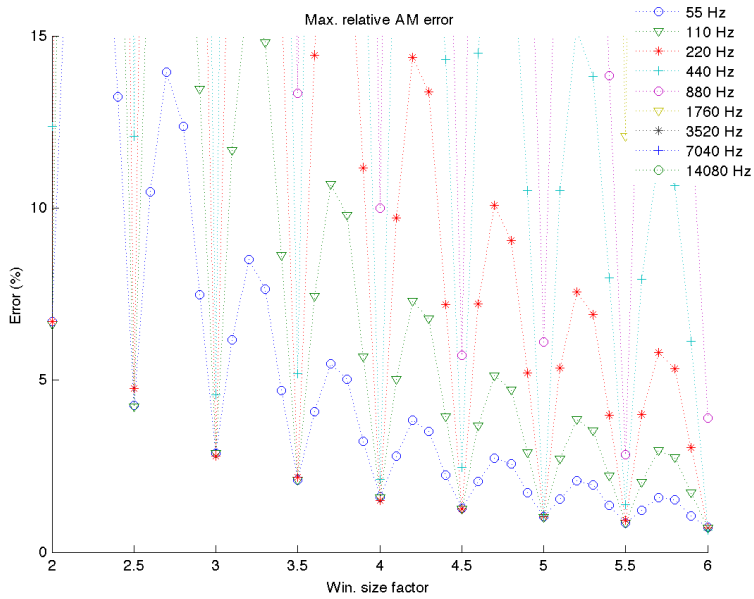


Figure 22: DAM: maximum AM estimate error in respect to frequency, AM rate: 50, window size 4 x wavelength for all frequencies.



(a) AM rate: 50 s^{-1}



(b) AM rate: 50 s^{-1}

Figure 23: DAM: maximum AM estimate error in respect to frequency and window size, different AM rates.

4.2.2 RM

AM estimate of RM behaves pretty similarly as its frequency estimate. Bias and oscillation around average are present, see figure 4.2.2. Since AM estimate depends directly on frequency estimate, we can proceed similarly as in DAM case and substitute frequency estimate with actual frequency, when computing AM estimate. The difference in estimates is almost negligible, less than 0.1% and is comparable to analogous difference in DAM case. Figure 25 shows AM/frequency estimate dependence from a random frame. Again, even a perfect frequency estimate would not yield correct AM. Obviously, whole curve is place much too high to come close to correct value even at actual frequency.

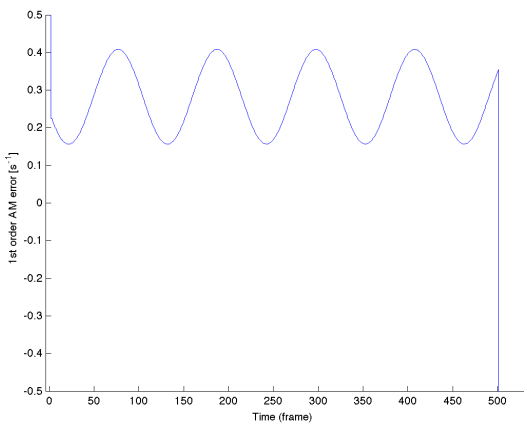


Figure 24: RM: AM estimate error for stable sinusoid, 1000Hz, AM rate 50, 177 samples window (4 x wavelength)

In contrast to DAM, absolute AM bias and variance are correlated with absolute AM value. In figures 26(a) and 26(b) we can see, how bias and variance increase as AM increases. Substituting estimate frequency for actual does not decrease neither bias nor variance significantly. Again, the only effective parameter that decreases error is window size. Figures 27(a) and 27(a) show maximum relative AM error in respect to frequency and window size, for 2 different AM rates. Comparing these plots we can again see, that careful choice of window size of exact multiple of half the wavelength will force RM to make very comparable relative AM error at high and low AM rates. This phenomena is not so explicit for higher frequencies, as can be seen in figure 27(b).

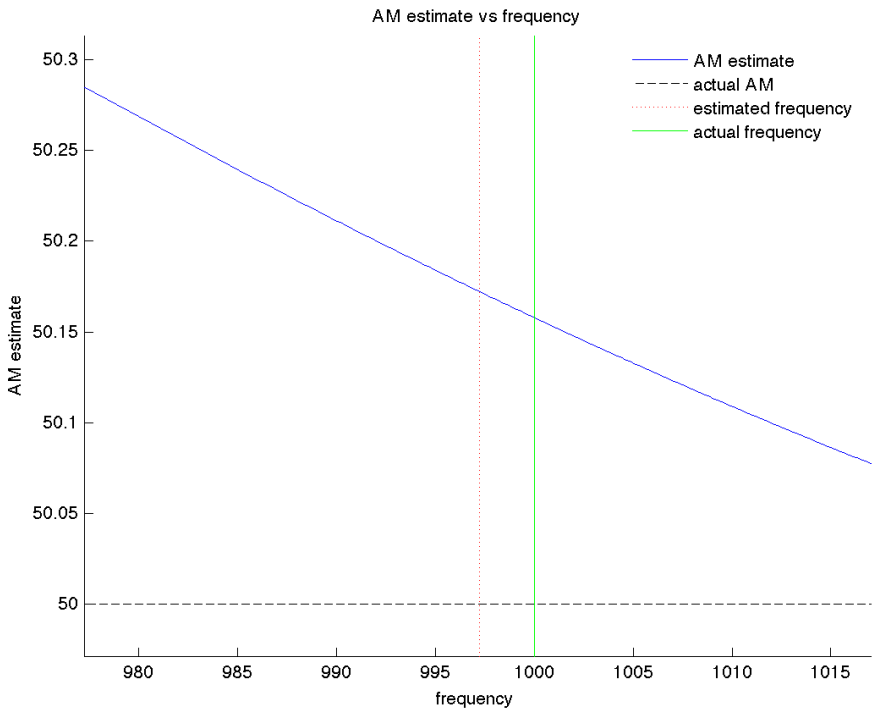
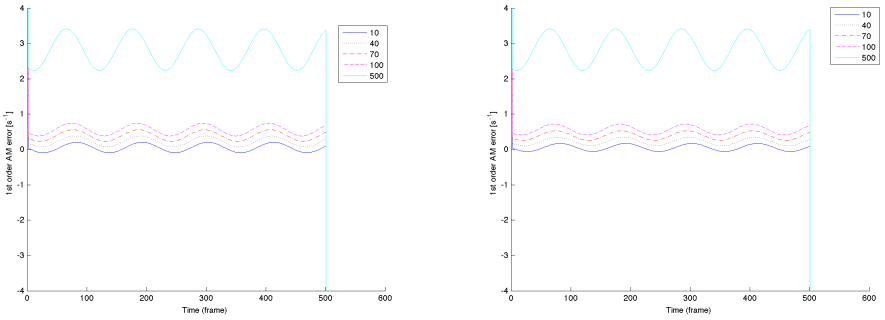


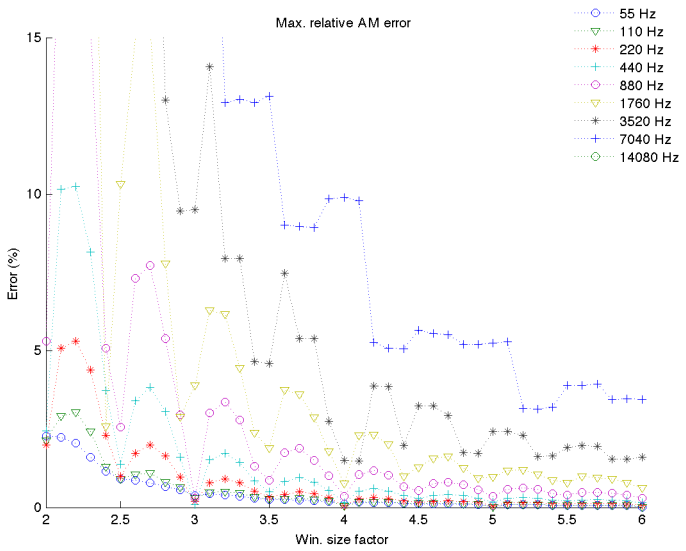
Figure 25: RM: $-\Re\left(\frac{S_{W'}}{S_W}(\omega)\right)$ for frequencies around estimated and actual frequency. Sinusoid of 1000Hz, no AM/FM, window size 177 (4 x fundamental period), no zero padding



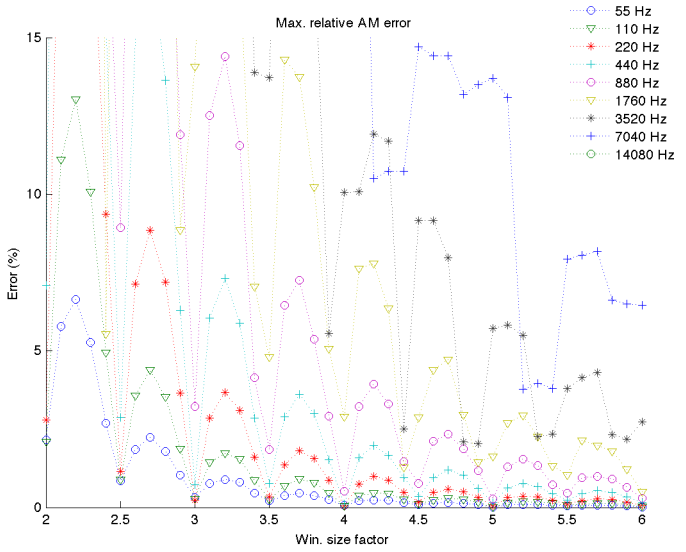
(a) AM error for different AM rates, estimated frequency is used

(b) AM error for different AM rates, actually frequency is used

Figure 26: RM: AM estimates for different AM rates, 177 sample window size (4 x wavelength), no FM, no zero padding



(a) AM rate: 50 s^{-1}



(b) AM rate: 10 s^{-1}

Figure 27: RM: maximum AM estimate error in respect to frequency and window size, different AM rates.

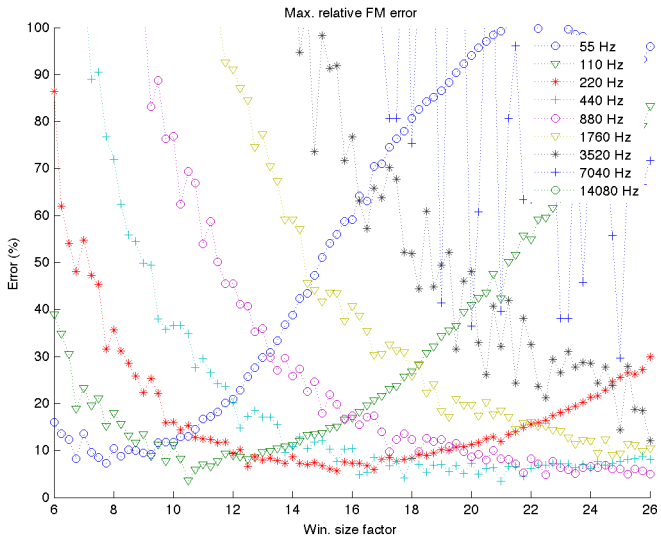
4.3 Frequency Modulation

4.3.1 DAM

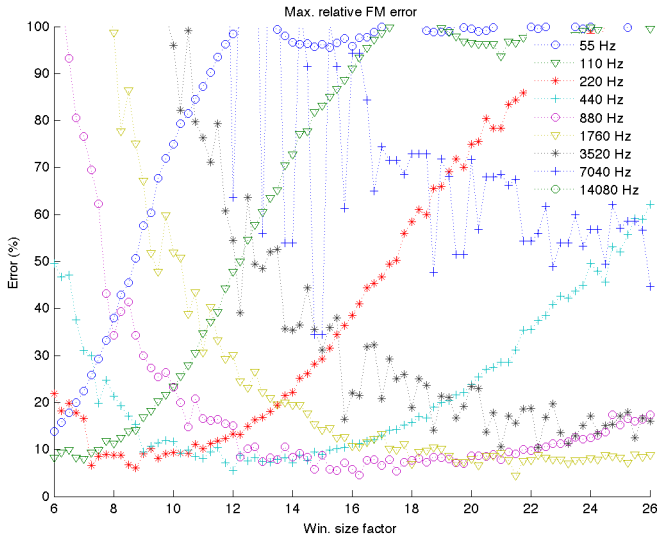
In section 3.2 the following equations for frequency, AM and FM estimates of DAM were given respectively:

$$\hat{\omega} = \Im\left(\frac{S'_w(\omega_k)}{S_w(\omega_k)}\right), \hat{\mu} = \Re\left(\frac{S'_w(\hat{\omega})}{S_w(\hat{\omega})}\right), \hat{\psi} = \Im\left(\frac{S''_w(\hat{\omega})}{S_w(\hat{\omega})}\right) - 2\hat{\mu}\hat{\omega}$$

,where ω_k is frequency of a FFT bin, nearest to spectrum peak, S_w is STFT of signal and S'_w, S''_w are STFTs of first and second order signal derivative and. Frequency estimate depends solely on frequency of selected FFT bin, whereas AM estimate depends on frequency estimate. Lastly, FM estimate should depend significantly more on frequency estimate as both terms $\Im\left(\frac{S''_w(\hat{\omega})}{S_w(\hat{\omega})}\right)$ and $2\hat{\mu}\hat{\omega}$ depend on frequency estimate, the second term being a product of frequency estimate and frequency dependent estimate, eg: second order dependance. Thus, an error in frequency estimate seems to have a big effect on FM, at least in theory. Test have shown, that in fact, frequency estimate has a bigger effect on FM estimate than on AM estimate. However, similar to the case of AM estimate, even if frequency and AM estimates are substituted with actual values, FM estimate accuracy doesn't improve significantly and for sure cannot be used to get perfect FM estimate.
Further, FM estimates seem to be very inaccurate: maximum relative errors are above 10% for all frequencies, considering window sizes smaller than 6 x wavelength. For bigger windows however, errors drop under 10% relative error. Figures 28(a) and 28(b) show, that for different frequencies, different relative window sizes should be used to achieve good FM estimate. From plots it's clearly seen, how too large and too small relative window size causes FM accuracy to drop. It is clear from the plots that accurate FM estimation of several partials, even in monophonic case requires DAM to be repeated for each partial with different window size, which makes computational cost of such algorithm highly dependable on number of partials. Anyhow, FM accuracy seems to be quite unsatisfactory for most practical uses and was thus not studied with AM present, as it would result in even poorer accuracy.



(a) FM rate: 1 octave/s



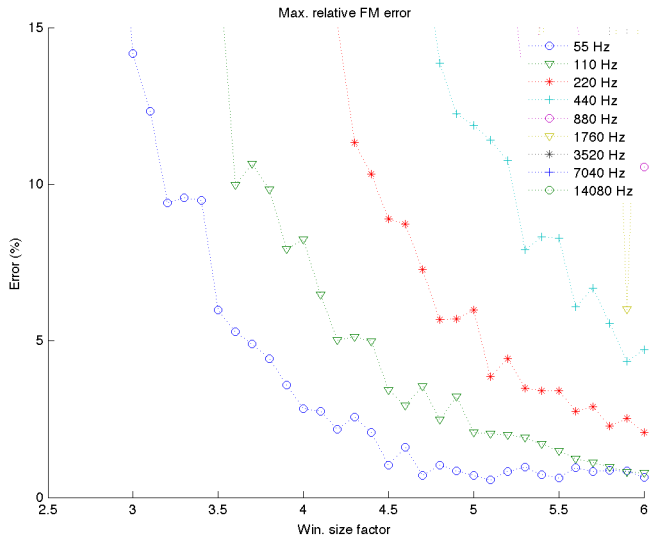
(b) FM rate: 2 octave/s

Figure 28: DAM: maximum relative FM errors for different frequencies, window sizes and FM rates, no AM, 0-100% relative error range

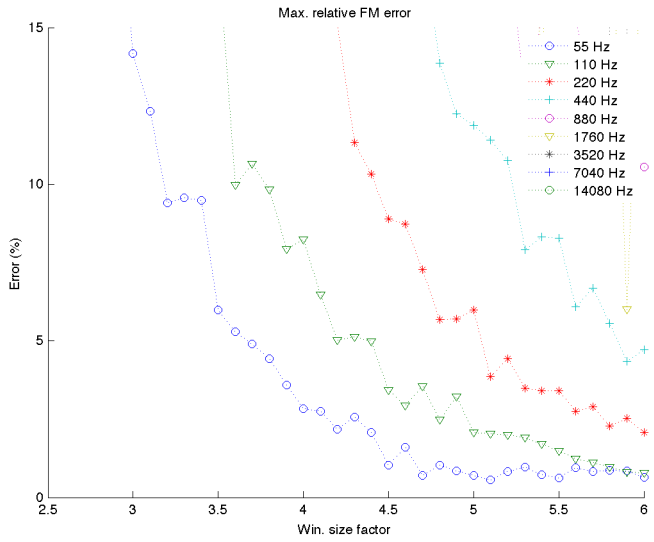
4.3.2 RM

In section 3.3 the following equations for frequency, AM and FM estimates of RM were given respectively: $\hat{\omega} = \omega_k - \Im\left(\frac{S_{w'}(\omega_k)}{S_w(\omega_k)}\right), \hat{\mu} = -\Re\left(\frac{S_{w'}(\hat{\omega})}{S_w(\hat{\omega})}\right), \hat{\psi} = \Im\left(\frac{S_{w''}(\hat{\omega})}{S_w(\hat{\omega})}\right) - \Im\left(\left(\frac{S_{w'}(\hat{\omega})}{S_w(\hat{\omega})}\right)^2\right)$,where ω_k is frequency of a FFT bin, nearest to spectrum peak and $S_{w'}, S_{w''}$ STFT with first, second order window derivative of the window used respectively. Again, frequency estimate depends only on select FFT bin frequency, AM estimate relies on frequency estimate and both terms of FM estimate, $\Im\left(\frac{S_{w''}(\hat{\omega})}{S_w(\hat{\omega})}\right)$ and $\Im\left(\left(\frac{S_{w'}(\hat{\omega})}{S_w(\hat{\omega})}\right)^2\right)$ depend on frequency estimate, the second term even exhibits second order dependance. Note however, that FM estimate does not use AM estimate, as it is the case in DAM. Similarly as in DAM FM estimation case and both (RM and DAM) AM estimation cases, substituting frequency estimation with actual frequency, to compute FM estimation, does not improve results significantly.

However, RM estimates FM fairly well, compared to DAM even for nearly wide band conditions. Figures 29(a) and 29(a) show maximum error RM makes for window sizes from 2.5 to 6 times wavelength for different FM rates. Figures 30(a) and 30(b) show maximum error for large windows, from 6-26 times wavelength for different FM rates. Unfortunately, best FM estimate accuracy at different frequencies is scattered without any apparent pattern, making it difficult to design a simple algorithm, that would estimate FM equally good for all frequencies. Minimum maximum errors reach around 1% for lower frequencies and go below 0.1% for frequencies higher than 700Hz, assuming the right window size is choosen.

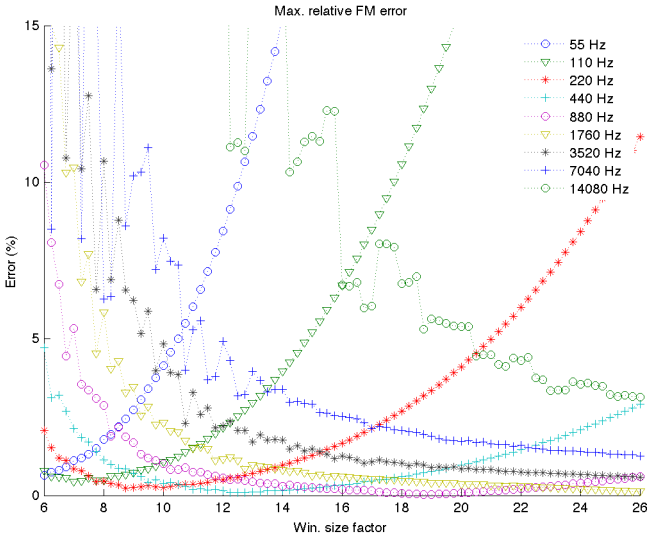


(a) FM rate: 1 octave/s

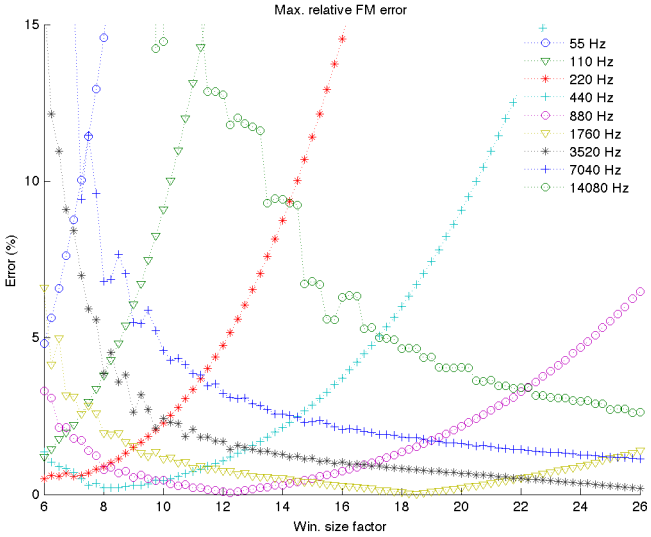


(b) FM rate: 2 octave/s

Figure 29: RM: maximum relative FM errors for different frequencies, window sizes (from almost wide to narrow band) and FM rates, no AM, 0-15% relative error range



(a) FM rate: 1 octave/s



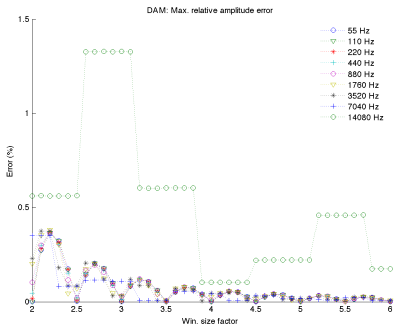
(b) FM rate: 2 octave/s

Figure 30: RM: maximum relative FM errors for different frequencies, window sizes and FM rates, no AM, 0-15% relative error range

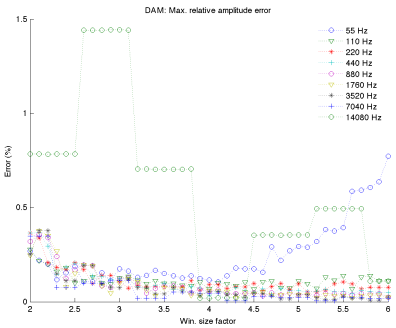
4.4 Amplitude

4.4.1 DAM

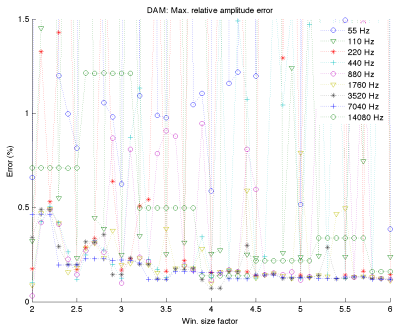
Amplitude estimate seems to be one of the most accurate ones of all. Using correction, that follows from equation (8) DAM makes an excellent estimate of amplitude. However, it seems that AM has a big impact on accuracy. Figures 31(a), 31(b), 31(c), 31(d) show combinations with and without AM/FM. Obviously, AM has a huge impact, as accuracy for most cases exceeds 1.5%, while no AM case exhibit relative errors well below 1%. Such effect was observed already in section 4.1.1 with frequency estimate.



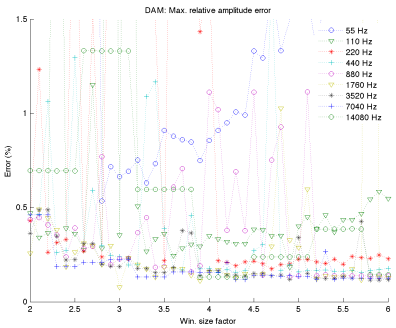
(a) FM rate: 0 octave/s, AM rate: 0 /s



(b) FM rate: 2 octave/s, AM rate: 0 /s



(c) FM rate: 0 octave/s, AM rate: 50 /s

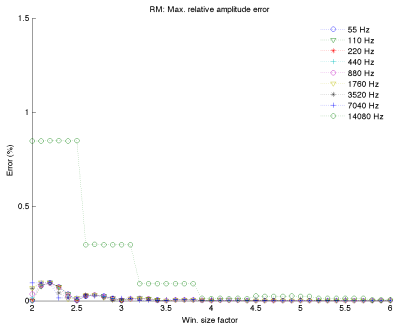


(d) FM rate: 2 octave/s, AM rate: 50 /s

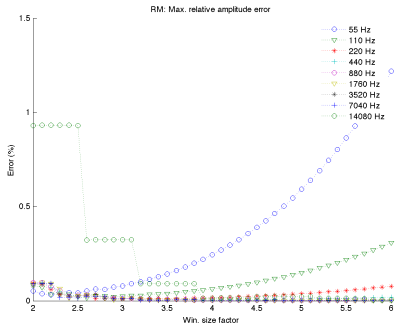
Figure 31: DAM: maximum relative amplitude error for different AM/FM, different window lengths, zero padding factor 2, relative error range 1.5%

4.4.2 RM

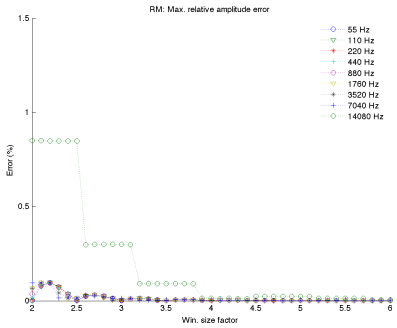
RM method estimates amplitude very accurately even with strong AM/FM present. Presented plots 32(a), 32(b), 32(c), 32(d) show accuracy, well below 1% for all frequencies at certain window size. AM has again a strong effect, but it does not cause estimates to exceed 1% even for windows 6 x wavelength. Similar phenomena, that AM has lower impact on estimate was already mentioned in section 4.1, where frequency estimate was studied.



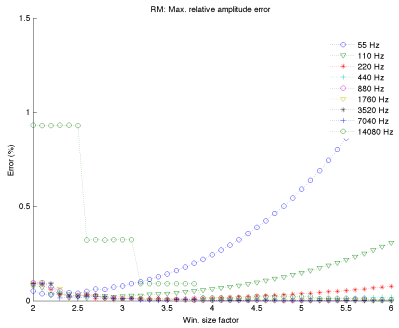
(a) FM rate: 0 octave/s, AM rate: 0 /s



(b) FM rate: 2 octave/s, AM rate: 0 /s



(c) FM rate: 0 octave/s, AM rate: 50 /s



(d) FM rate: 2 octave/s, AM rate: 50 /s

Figure 32: RM: maximum relative amplitude error for different AM/FM, different window lengths, zero padding factor 2, relative error range 1%

4.5 Theoretical extensions of DAM to 2nd order AM/FM estimation

As mentioned in introduction a robust second order modulation estimation is a very attractive option. To account for second order AM/FM, model of signal now changes slightly:

$$\varphi(t) = \phi_0 + \omega_0 t + \frac{\psi_0}{2}t^2 + \frac{\gamma_0}{3}t^3 \tag{38}$$

$$a(t) = \lambda_0 + \mu_0 t + \frac{\tau_0}{2}t^2 \tag{39}$$

$$s(t) = \exp(a(t) + i\varphi(t)) \tag{40}$$

Computing first three derivatives of the signal yields following equalities:

$$\frac{s'(t)}{s(t)} = \mu_0 + \tau_0 t + i(\omega_0 + \psi_0 t + \gamma_0 t^2) \tag{41}$$

$$\frac{s''(t)}{s(t)} = (\tau_0 + i\,(2t\gamma_0 + \psi_0)) + \left(\mu_0 + t\tau_0 + i\left(t^2\gamma_0 + t\psi_0 + \omega_0\right)\right)^2 \tag{42}$$

$$\begin{aligned}\frac{s'''(t)}{s(t)} &= i\gamma_0 + 3\,(\tau_0 + i\,(2t\gamma_0 + \psi_0))\left(\mu_0 + t\tau_0 + i\left(t^2\gamma_0 + t\psi_0 + \omega_0\right)\right) + \\ &\quad \left(\mu_0 + t\tau_0 + i\left(t^2\gamma_0 + t\psi_0 + \omega_0\right)\right)^3\end{aligned} \tag{43}$$

Without going further into details following the idea presented in [22] and taking the same assumptions as in section 3.2 frequency and AM estimate do not change:

$$\hat{\omega}_0 = \Im\left(\frac{S'_w(\omega_k)}{S_w(\omega_k)}\right) \tag{44}$$

$$\hat{\mu}_0 = \Re\left(\frac{S'_w(\hat{\omega}_0)}{S_w(\hat{\omega}_0)}\right) \tag{45}$$

$$\tag{46}$$

However, considering second order signal derivative, FM estimate equation

changes.

$$\Im\left(\frac{s''(t)}{s(t)}\right) = \psi_0 + 2\gamma_0 t + 2(\mu_0\omega_0 + \mu_0\psi_0 t + \mu_0\gamma_0 t^2 + \tau_0\omega_0 t + \tau_0\psi_0 t^2 + \tau_0\gamma_0 t^3) \Rightarrow \tag{47}$$

$$\Im\left(\frac{S_w''(\omega_k)}{S_w(\omega_k)}\right) = \psi_0 + 2\mu_0\omega_0 \Rightarrow \tag{48}$$

$$\psi_0 = \Im\left(\frac{S_w''(\omega_k)}{S_w(\omega_k)}\right) - 2\mu_0\omega_0 \tag{49}$$

$$\Re\left(\frac{S_w''(\omega_k)}{S_w(\omega_k)}\right) = \mu_0^2 - \omega_0^2 + \tau_0 \Rightarrow \tag{50}$$

$$\tau_0 = \Re\left(\frac{S_w''(\omega_k)}{S_w(\omega_k)}\right) - \mu_0^2 + \omega_0^2 \tag{51}$$

Therefore, 2nd order AM is easy to calculate, but equation for second order FM is somehow not yet straightforward, so we consider 3rd signal derivative:

$$\Im\left(\frac{S_w'''(\omega_k)}{S_w(\omega_k)}\right) = 2\gamma_0 - \omega_0^3 + 3\mu_0\psi_0 + 3\mu_0^2\omega_0 + 3\tau_0\omega_0 \Rightarrow \tag{52}$$

$$\gamma_0 = \frac{1}{2}\left(\omega_0^3 - 3\mu_0\psi_0 - 3\mu_0^2\omega_0 - 3\tau_0\omega_0 + \Im\left(\frac{S_w'''(\omega_k)}{S_w(\omega_k)}\right)\right) \tag{53}$$

Presented steps gradually lead to estimations of all parameters, as parameters estimated in the beginning are used in later equations, so implementing an algorithm that follows these equations is straight forward. Finally, spectrum correction is much more difficult in this case, as several definite integrals, leading from equation (8), but generalized for second order AM/FM cannot be expressed as any known function. Therefore, accurate amplitude and phase estimators are not easy to implement. However, considering the errors, that DAM makes for linear AM/FM it is no surprise, that second order AM/FM estimates, as defined above give enormous errors rendering it useless in practice. Tests have indeed shown, that second order AM/FM estimators using DAM exhibit very high errors and can be regarded as unusable in practice. It would be however reasonable to consider using same technique with RM, since theoretical equivalence of DAM and RM was proven in [22].
Another attractive option is to head another way directly from equation (41), by computing derivative of both sides:

$$\frac{s''(t)s(t) - s'(t)^2}{s(t)^2} = \tau_0 + i(\psi_0 + \gamma_0 t) \text{ or equivalently} \tag{54}$$

$$s''(t)s(t) - s'(t)^2 = (\tau_0 + i(\psi_0 + \gamma_0 t))\, s(t)^2 \text{ leading to} \tag{55}$$

$$s_1 = (\tau_0 + i(\psi_0 + \gamma_0 t))\, s_2 \tag{56}$$

Now, we could proceed similarly as in [22], with substituted signals. All assumptions stated in [22] now hold for substituted signals. In order to acquire

substitutes we have to calculate products $s_2''(t)s(t), s_2'(t)^2, s_2(t)^2$. In case, when signal is multi-component (contains more than 1 sinusoid), squared signal and signal/signal derivative products double amplitudes and frequencies of all partials and produce artifacts at sums of frequencies, eg:

$$s(t) = \sum_{k=1}^{n} \exp(a_k(t) + i\varphi_k(t)) \Rightarrow \tag{57}$$

$$s^2(t) = \sum_{k=1}^{n}\sum_{j=1}^{n} \exp(a_k(t) + a_j(t) + i(\varphi_k(t) + \varphi_j(t))) \tag{58}$$

$$= \sum_{k=1}^{n} \exp(2a_k(t) + 2i\varphi_k(t)) + \sum_{k=1}^{n}\sum_{j=1 j\neq k}^{n} \exp(a_k(t) + a_j(t) + i(\varphi_k(t) + \varphi_j(t))) \tag{59}$$

Clearly, frequency and amplitude of every partial is doubled and n(n-1) artifacts are added. Frequencies and amplitudes of artifacts are sums of all possible pairs of partials. For harmonic or nearly harmonic timbres this is highly undesirable, as overlapping partials pose a difficult problem. It is interesting to see, that similar cross-terms are observed using Wigner-Ville distributions. Solutions, employed in those cases can therefore be used in our case, like for example [18]. A filterbank of bandpass filters, centered at peaks with cutoff frequencies halfway between peaks efficiently separates partials and allows us to avoid cross terms using mentioned method. Unfortunately, tests have shown that even for single sinusoid case, above procedure gives very high error for second order modulation estimates as well and is very much unusable in practice.

4.6 Detailed study of DAM properties

In order to further explore, why DAM makes such considerable errors, we start with equation (41) and compute STFT of both sides:

$$S_w'(\omega, t) = \mathfrak{F}\{(\mu_0 + \tau_0 t + i(\omega_0 + \psi_0 t + \gamma_0 t^2))\} * S_w(\omega, t) \tag{60}$$

$$= (\mu_0\delta(\omega) + i\tau_0\delta'(\omega) + i(\omega_0\delta(\omega) + i\psi_0\delta'(\omega) - \gamma_0\delta''(\omega))) * S_w(\omega, t) \tag{61}$$

$$= (\mu_0\delta(\omega) - \psi_0\delta'(\omega) + i(\omega_0\delta(\omega) + \tau_0\delta'(\omega) - \gamma_0\delta''(\omega))) * S_w(\omega, t) \tag{62}$$

From definition of Dirac delta function we know that $(\delta^{(n)} * f)(x) = (-1)^n f^{(n)}(x)$, thus we can write above equation in the following, more intuitive form:

$$S_w'(\omega, t) = \mu_0 S_w(\omega, t) + \psi_0 S_w^{(1)}(\omega, t) + i(\omega_0 S_w(\omega, t) - \tau_0 S_w^{(1)}(\omega, t) - \gamma_0 S_w^{(2)}(\omega, t)) \tag{63}$$

,where $S_w^{(n)}(\omega)$ denotes n-th derivative of spectrum of windowed portion of signal with respect to frequency, which brings us to reassignment like term:

$$S'_w(\omega,t) = \mu_0 S_w(\omega,t) - i\psi_0 S_{tw}(\omega,t) + i(\omega_0 S_w(\omega,t) + i\tau_0 S_{tw}(\omega,t) + \gamma_0 S_{t^2w}(\omega,t)) \tag{64}$$

$$= \mu_0 S_w(\omega,t) - \tau_0 S_{tw}(\omega,t) + i(\omega_0 S_w(\omega,t) - \psi_0 S_{tw}(\omega,t) + \gamma_0 S_{t^2w}(\omega,t)) \tag{65}$$

,where S_{tw}, S_{t^2w} are STFTs obtained by using time ramped window and time squared ramped window respectively, much like those in reassignment.
We can see from (65) that spectrum of signal derivative is a sum of components, that are directly computable from original spectrum and parameter estimates. It might be informative to plot magnitude spectrum while varying the parameters, as such plots might reveal, how the errors are generated.

5 Conclusion and future work

Presented comparison of DAM and RM differs significantly from the one, done by Marchand in [22]. RM seems to perform better in all conditions, which is in contrast with [22]. However, a possible reason for this could be the slight improvement of RM, described in section 3.3. It's reasonable to claim, that his slight improvement makes RM perform much better for small window lengths, which was the main advantage of DAM, reported by Marchand in [22]. However, the amount of accuracy gained was not measured. Such comparison is straightforward and most of current implementation can be reused, so it seems like a reasonable next step.

Depending on purpose, more or less accurate AM/FM is desired. From results presented in section 4 it is clear, that for accuracies below 1% for all frequency range RM should be reran at different window sizes, increasing computational cost significantly. Of course, a non real-time analysis can still benefit from such analysis. Designing such an algorithm requires exact relationship between frequency and best window size. Plots of such dependency were presented in section 4, but more exact mathematical definition would be needed. Further, the method of removing bias in RM case as described by Hainsworth in [14] should be generalized to squared Hanning window and implemented.

Frequency and AM estimates exhibit very predictable errors in both DAM and RM case. As mentioned in section 4, there exists very high probability, that this errors is closely related to phase of partial under study and since it is present in DAM and RM cases, it is very likely that is inherent to discrete Fourier transform. Removing this fluctuations would greatly reduce errors in both DAM and RM cases, thus it is worth to put some effort into such research.

As already mentioned in section 4.5 RM could be extended to estimate second order AM/FM. This would most probably imply using second derivative of window. Currently preferred squared Hanning window might not provide sufficiently low side lobe rejection ratio, possibly forcing to design new kind of window function. This probably leads to compromise between side lobe rejection ratios and main lobe widths of original, first and second derivative. New type of window will also require recalculation of integrals for spectrum correction, however they might be simpler than the ones for 'raised cosine' type windows. The fact, that frequency peak in magnitude spectrum shifts in frequency in presence of AM/FM is also one of the problems to solve. To achieve this, correction function for other than Gauss windows should be analyzed to give analytical expression for this estimate biases. Looking at correction functions for cosine based windows in subsection 3.4.2, this looks like a hard problem. However, despite fairly large number of terms containing complex error function, there is many useful assumptions that one can make, that will significantly reduce the complexity. The fact that bias expression cannot be very complicated is supported by plots of such functions (computed numerically) by Betser et al in [6]. In small frequency error range, bias looks very much linear, therefore even linear approximation would suffice.

References

[1] T.J. Abatzoglou. Fast maximnurm likelihood joint estimation of frequency and frequency rate. *Aerospace and Electronic Systems, IEEE Transactions on*, AES-22(6):708–715, 1986.

[2] M. Abe and J.O. Smith. AM/FM rate estimation for time-varying sinusoidal modeling. In *Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference on*, volume 3, pages iii/201–iii/204 Vol. 3, 2005.

[3] F. Auger and P. Flandrin. Generalization of the reassignment method to all bilinear time-frequency and time-scale representations. In *Acoustics, Speech, and Signal Processing, 1994. ICASSP-94., 1994 IEEE International Conference on*, volume iv, pages IV/317–IV/320 vol.4, 1994.

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

[5] S. Barbarossa. Analysis of multicomponent LFM signals by a combined Wigner-Houghtransform. *Signal Processing, IEEE Transactions on*, 43(6):1511–1515, 1995.

[6] M. Betser, P. Collen, G. Richard, and B. David. Estimation of frequency for AM/FM models using the phase vocoder framework. *Signal Processing, IEEE Transactions on*, 56(2):505–517, 2008.

[7] Karlheinz Brandenburg, Jürgen Herre, James D Johnston, Yannick Mahieux, and Ernst F Schroeder. ASPEC-Adaptive spectral entropy coding of high quality music signals. In *Aud. Eng. Soc., 1991 (AES 91), Proceedings of 90th Convention*, volume 90, page 3011, February 1991.

[8] O. Celma, A. Loscos, J. Ortolà, X. Serra, and J Bonada. Singing voice synthesis combining excitation plus resonance and sinusoidal plus residual models. In *International Computer Music Conference, 2001. Proceedings (ICMC '01)*, 2001.

[9] Leon Cohen. *Time Frequency Analysis: Theory and Applications*. Prentice Hall PTR, facsimile edition, December 1994.

[10] Myriam Desainte-Catherine and Sylvain Marchand. High precision fourier analysis of sounds using signal derivatives. *Audio Engineering Society, 1953 Journal of*, 1998.

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

[12] J. L. Flanagan and R. M. Golden. Phase vocoder. *Bell Syst. Tech. J.,vol.45,pp. 1493-1509*, 1966.

[13] D. Gabor. Theory of communication. *Journal IEE*, 93(26):120, November 1946.

[14] Stephen Webley Hainsworth, Stephen W Hainsworth, and Stephen W Hainsworth. *Techniques for the Automated Analysis of Musical Audio*. PhD thesis, Signal Processing Group Signal Processing Group Engineering Department Cambridge University, 2004.

[15] P.V. Hough. Methods and means to recognize complex patterns. United States Patent Office, U.S. Patent 3.069.654, 1962.

[16] Julius O. Smith III and Xavier Serra. PARSHL: an Analysis/Synthesis program for Non-Harmonic sounds based on a sinusoidal representation. In *International Computer Music Conference, 1987. Proceedings (ICMC '87)*, Stanford, California 94305, 1987. Department of Music, Stanford University.

[17] M. Z. Ikram, K. Abed-Meraim, and Y. Hua. Fast quadratic phase transform for estimating the parameters of multicomponent chirp signals. *Digital Signal Processing*, 7(2):127–135, 1997.

[18] Corey Kereliuk and Philippe Depalle. Improve hidden markov model partial tracking through time-frequency analysis. In *Digital Audio Effects, 2008. Proceedings of the 11th Int. Conference on (DAFx-08)*, Espoo Finland, 2008.

[19] K. Kodera, R. Gendrin, and C. Villedary. Analysis of time-varying signals with small BT values. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 26(1):64–76, 1978.

[20] E. Lindemann. Music synthesis with reconstructive phrase modeling. *Signal Processing Magazine, IEEE*, 24(2):80–91, 2007.

[21] Sylvain Marchand. Improving spectral analysis precision with an enhanced phase vocoder using signal derivatives. *Digital Audio Effects Workshop, 1998. Proceedings of (DAFx '98)*, pages 114—118, 1998.

[22] Sylvain Marchand and Philippe Depalle. Generalization of the derivative analysis method to Non-Stationary sinusoidal modeling. *Digital Audio Effects, 2008. Proceedings of the 11th Int. Conference on (DAFx-08)*, 2008.

[23] L.S. Marques and L.B. Almeida. Frequency-varying sinusoidal modeling of speech. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 37(5):763–765, 1989.

[24] Geoffroy Peeters and Xavier Rodet. SINOLA: a new Analysis/Synthesis method using spectrum peak shape distortion, phase and reassigned spectrum. *International Computer Music Conference, 1999. Proceedings of (ICMC '99)*, pages 153—156, 1999.

[25] Axel Röbel. Estimating partial frequency and frequency slope using reassignment operators. *International Computer Music Conference, 2002. Proceedings (ICMC '02)*, pages 122—125, 2002.

[26] X. Serra. *A System for Sound Analysis/Transformation/Synthesis based on a Deterministic plus Stochastic Decomposition.* PhD thesis, Stanford University, 1989.

[27] X Serra. State of the art and future directions in musical sound synthesis. In *Multimedia Signal Processing, 2007. MMSP 2007. IEEE 9th Workshop on*, 2007.

[28] J. Ville. Théorie et applications de la notion de signal analytique. *Cables et Transmission*, 2A:61–74, 1948.

[29] Mark Sandler Wen Xau. Notes on model-based non-stationary sinusoid estimation methods using derivative. In *Digital Audio Effects, 2009. Proceedings of the 12th Int. Conference on (DAFx-09)*, 2009.

[30] Gerhard Widmer, Davide Rocchesso, Vesa Välimäki, Cumhur Erkut, Fabien Gouyon, Daniel Pressnitzer, Henri Penttinen, Pietro Polotti, and Gualtiero Volpe. Sound and music computing: Research trends and some key issues. *Journal of New Music Research*, 36(3):169, 2007.

[31] E Wigner. On the quantum correction for thermodynamic equilibrium. *Physical Review*, 40(5):749, 1932.