Sinusoidal Modeling

summer 2006 lecture on analysis,
modeling and transformation of audio signals

Axel Röbel
Institute of communication science TU-Berlin
IRCAM Analysis/Synthesis Team

25th August 2006

Contents

1 Sinusoids plus noise sound modeling

- In the previous lectures we have been using a generic representation of sound in terms of the Fourier spectrum.
- Most of the algorithms so far did not make use of a representation of the sound in terms of an explicit signal model.
- A signal model was implicitly used for example in the phase vocoder time stretching algorithm [Röb06c, section 3] and for fundamental frequency estimation [Röb06e].
- Higher level of sound representation try to distinguish the perceptually different components:
  - **sinusoids** and **noise**.
- In the following we will see how we may represent a sound signal by means of the sinusoids plus noise signal model. An introduction can be found in [Ser97], sound transformation applications are explained in [ABLS02]
- open source software library for sinusoidal modeling and transformation can be found at `http://www.cerlsoundgroup.org/Loris/` and `http://clam.iua.upf.edu/`

1.1 Sinusoids

Why sinusoids ?

- real world excitation signals (source filter model) are often periodic such that they can be represented by means of a superposition of harmonically related sinusoids.
- free oscillation of physical systems can generally be characterized by means of a superposition of modes, where each mode contributes a sinusoid with characteristic frequency to the output signal.
- if modes are not too dense the related sound will be perceived as rather clean.

Each sinusoidal component is identified by its index k and each individual component has time varying amplitude $a_k(n)$ and time varying phase $\phi_k(n)$. A single sinusoidal component can be represented as

$$P_k(n) = a_k(n)\cos(\phi_k(n)). \tag{1}$$

or in complex notation

$$P_k(n) = a_k(n)e^{\phi_k(n)}. \tag{2}$$

For a time continuous sinusoids the frequency is the time derivative of the phase. It is convenient to define the frequency of the discrete time sinusoid as the phase difference of subsequent samples.

$$\omega_k(n) = \frac{\phi_k(n+1) - \phi_k(n-1)}{2} \tag{3}$$

Without any further constraints each sound signal could be interpreted as a sinusoid if we would set $a_0(n) = s(n)$ and $\phi_0(n) = 0$.

The idea however is that the sinusoidal components are perceived as individual entities. As a vague constraint for sinusoidal components it is required that the amplitude $a_k(n)$ and the derivative of the unwrapped phase with respect to time of the related continuous time phase $\frac{\partial \phi_k(t)}{\partial t}$ is sufficiently small such that the perceived quality is close to a stationary sinusoid.

The complete set of sinusoidal components of a signal $s(n)$ are represented by means of the superposition

$$s(n) = \sum_k P_k(n) = \sum_k a_k(n)\cos(\phi_k(n)). \tag{4}$$

1.2 Noise

Having detected all sinusoidal components with parameters $a_k(n)$ and $\phi_k(n)$ we may subtract them from the signal. The remaining signal is called the residual.

The residual combines signal noise and modeling error.

Noise/sinusoid classification

- For a sinusoid plus noise model a classification procedure is required that distinguishes sinusoidal and noise peaks of the signal spectrum.
- Common techniques are based on amplitude level and smoothness of the amplitude and frequency trajectory. In that case sinusoids forming amplitude or frequency trajectories that are not sufficiently smooth are removed from the set of sinusoids. For harmonic sounds the sinusoidal selection is simplified because the frequency positions where sinusoids are expected are confined to the integer multiples of the fundamental frequency.
- Their exist few algorithms that allow to distinguish between spectral peaks representing sinusoids and noise. Common techniques are based on features that are derived

from the form of the phase and amplitude spectrum [RZR04, HMW01, Rod97, Tho82].

2 Overview over the sinusoidal analysis/synthesis model

pre processing : the sinusoidal analysis is performed on the STFT of the signal. The STFT parameters window size, DFT size and frame offset have to be chosen such that the interesting sinusoids are resolved [Röb06b].

peak detection : each STFT frame is analyzed to find the spectral peaks (section **3**).

sinusoidal parameter estimation : for each peak that has been selected the sinusoidal parameters are estimated (section **4**).

sinusoidal peak continuation : for synthesis of the sinusoids a complete trajectory of amplitude, frequency, and phase is required. The STFT provides values only at a grid given by the hop size of the analysis. The values in between the frames have to be interpolated and, therefore, peaks in consecutive frames have to be matched (connected) to be able to create complete trajectories.

residual creation : if a residual signal is desired, the sinusoidal parameters for all sinusoids have to be interpolated form frame rate to sample rate and the sinusoids have to be synthesized and subtracted from the signal.

noise model : A dedicated noise model can be fitted to the residual spectrum. Common choice is based on a source filter model [Röb06d], using a spectral envelope of the residual and excitation using white noise.

3 Peak detection

- It is a fundamental property of a sinusoid that it will create a prominent local peak in the spectrum,
- a spectral peak is a local maximum of the magnitude spectrum,
- for each spectral frame the spectral peaks are determined by means of searching these local maxima,
- amplitude thresholds or other classification schemes may be used to prevent the need to process a large number of peaks that later are qualified as noise,

4 Parameter estimation

Having selected the candidate peaks one needs to determine the parameters of the related sinusoids.

The minimum set of parameters comprises: amplitude and frequency. In many cases phase is estimated as well. Proper phase estimation is essential to be able to subtract the sinusoid from the sound.

4.1 stationary sinusoids

Remember: DFT spectrum of stationary sinusoid

$$s(n) = e^{j(\Omega n+\phi)} \tag{5}$$

using analysis window $v(n)$ is given by the window spectrum $V(\omega)$ moved to the location of the sinusoid frequency [Röb06a, section 4.1]

$$X(w) = (e^{j((m+\frac{M-1}{2})\Omega+\phi)}) \cdot (e^{-j\frac{M-1}{2}\omega}) \cdot |V(\omega - \Omega)|. \tag{6}$$

Due to linearity of DFT transformation the result for sinusoidal amplitude $a(n) = A$ multiplies the result by A.

Parameter estimate for a stationary sinusoid in noise:

frequency : frequency location of the maximum of the peak ω_0.

amplitude : amplitude value at location ω_0 of the spectrum divided by the maximum of the spectrum of the analysis window. From FT of the analysis window we find

$$\max(|V(\omega)|) = \sum_{n=0} M - 1v(n) \tag{7}$$

phase : estimated from the phase spectrum at position ω_0. Attention, remove the phase trend first !!!

This parameter estimate is assigned to the center of the analysis window.

It has been shown that the procedure above implements a maximum likelihood estimate *MLE* of the sinusoidal parameters.

***MLE*:** parameter values that create observed signal with maximum probability.

4.2 DFT interpolation

The MLE procedure above fails for the DFT spectrum because the maximizer of the spectrum is always confined to the bin positions.

Bin positions do not align with sinusoidal frequencies.

Solutions

- zero padding : increase analysis frame by means of adding zeros after windowing. Zero padding decreases the frequency distance between bins. Processing time scales with DFT size N according to $N\log(N)$. zero padding is rather costly.
- Quadratic interpolation of the DFT spectrum : *QIFFT*
  - select maximum bin and the two direct neighbors.
  - select perform a second order (quadratic) interpolation of log-amplitude spectrum and unwrapped phase spectrum.
  - apply the parameter estimation procedure to the quadratically interpolated peak spectrum.

In the real world applications both solutions are mixed. According to Taylor series approximation the error of the quadratic interpolation will become smaller with a smaller the distance of the supporting points to the maximum.

5 Estimator performance evaluation

The quantitative evaluation is usually performed by means of parameter estimation from single sinusoids in noise.

The estimation error is shown as a function of the SNR.

Two error contributions, bias and variance, are distinguished. Denote P and unknown parameter to be estimated and $\hat{P}$ the estimate that an estimator F will produce. Then we can define the bias as:

$$B_F = E(\hat{P}) - P, \tag{8}$$

where $E()$ denotes the expected value, generally the sample mean. The bias is the average or systematic error of the estimator.

The variance is then defined as

$$\sigma_F^2 = E((\hat{P} - E(\hat{P}))^2). \tag{9}$$

It tells us about the variation of the estimate around its average value.

The mean squared error MSE can now decomposed into bias and variance:

$$\mathrm{MSE}(F) = E((\hat{P} - P)^2) = \frac{1}{L}\sum_{n=0}^{L}(\hat{P}(n) - E(\hat{P}) + E(\hat{P}) - P)^2 \quad (10)$$

$$= \frac{1}{L}\sum_{n=0}^{L}(\hat{P}(n) - E(\hat{P}) + B_F)^2 \quad (11)$$

$$= \frac{1}{L}\sum_{n=0}^{L}((\hat{P}(n) - E(\hat{P}))^2 + 2(\hat{P}(n) - E(\hat{P}))B_F + B_F^2) \quad (12)$$

$$= \sigma_n^2 + B_F^2 + 2B_F((\frac{1}{L}\sum_{n=0}^{L}\hat{P}(n)) - E(\hat{P})) \quad (13)$$

$$= \sigma_F^2 + B_F^2 + 2B_F(E(\hat{P}) - E(\hat{P})) \quad (14)$$

$$= \sigma_F^2 + B_F^2 \quad (15)$$

This tells us that the mean squared error can be decomposed into the squared bias and the variance. The squared bias as the average error is the indicator for systematic errors. The variance is the indicator for noise sensitivity.

5.1 Cramer Rao bound

The Cramer-Rao theorem provides a lower bound for the variance of an unbiased estimator. An unbiased estimator is an estimator for that $B_F = 0$. If we denote the Cramer Rao bound of the estimation of parameter λ as $\mathrm{CRB}(\hat{\lambda})$ and if $\sigma_{\hat{\lambda}}$ is the variance of an estimator that provides estimates for variable λ then this variance is bounded by the Cramer-Rao bound

$$\sigma_F^2 \geq \mathrm{CRB}(\hat{\lambda}) \tag{16}$$

The Cramer Rao bound is a function of the Fisher information of the probability distribution of the data x given the parameter λ $P(x|\lambda)$ [Kay88].

The Cramer Rao bounds for sinusoidal parameter estimation for the case of a single stationary complex exponential of length N and amplitude A in stationary complex white Gaussian noise with variance σ_z

$$s(n) = Ae^{j(wn+\phi)} + z(n) \tag{17}$$

are [RB98]:

$$\text{Amplitude:} \quad \mathrm{CRB}(\hat{A}) = \frac{\sigma_z}{N} \tag{18}$$

$$\text{Frequency:} \quad \mathrm{CRB}(\hat{w}) = \frac{6\sigma_z^2}{A^2 N^3} \tag{19}$$

$$\text{Phase:} \quad \mathrm{CRB}(\hat{\phi}) = \frac{\sigma_z^2}{2NA^2} \tag{20}$$

(21)

The bounds decrease with increasing observation length and with decreasing noise level.

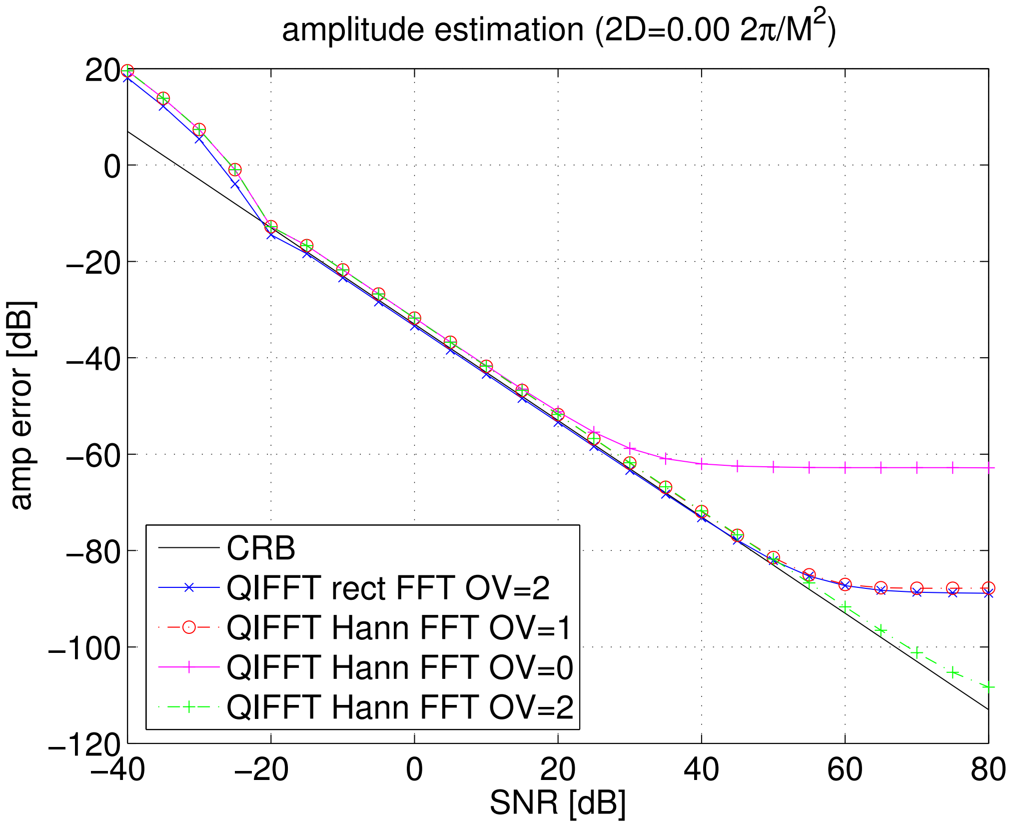

Figure 1: Estimation error and Cramer Rao bound for estimation of sinusoidal amplitude using QIFFT with different zero padding and different analysis windows. (window length M=1000, FFT size $N = 2^{\mathrm{nextpow2}(M)+OV}$ ¡).

As first example consider an experiment that evaluates different zero padding factors and different analysis windows for the estimation of the amplitude. The axis of the CRB graphs display the SNR as x-axis such that moving to the right will decrease the noise variance. On the y-axis the MSE of the error of the estimator is displayed. The error curves can be divided into three regions.

middle section:

- the error follows the CRB (the error is dominated by the variance)
- curves are close to the CRB (estimator is rather efficient)

left section:

- section the estimator variance increases stronger than the CRB,
- threshold effects (noise peaks are selected)

right section:

- with decreasing noise the variance part of the MSE will fall below the bias
- estimator errors saturate at a fixed level given by the estimator bias,

Conclusion

The present curves show clearly that the bias decreases with the zero padding factor (interpolation errors become smaller). Moreover the rectangular window has larger bias then the Hanning window because the mainlobe of the rectangular window is narrower and less well approximated by a quadratic function. Note, however, that the rectangular window is closer to the CRB in the middle section. This shows that the down weighting that the other windows apply to the border regions of the data decreases estimator efficiency.

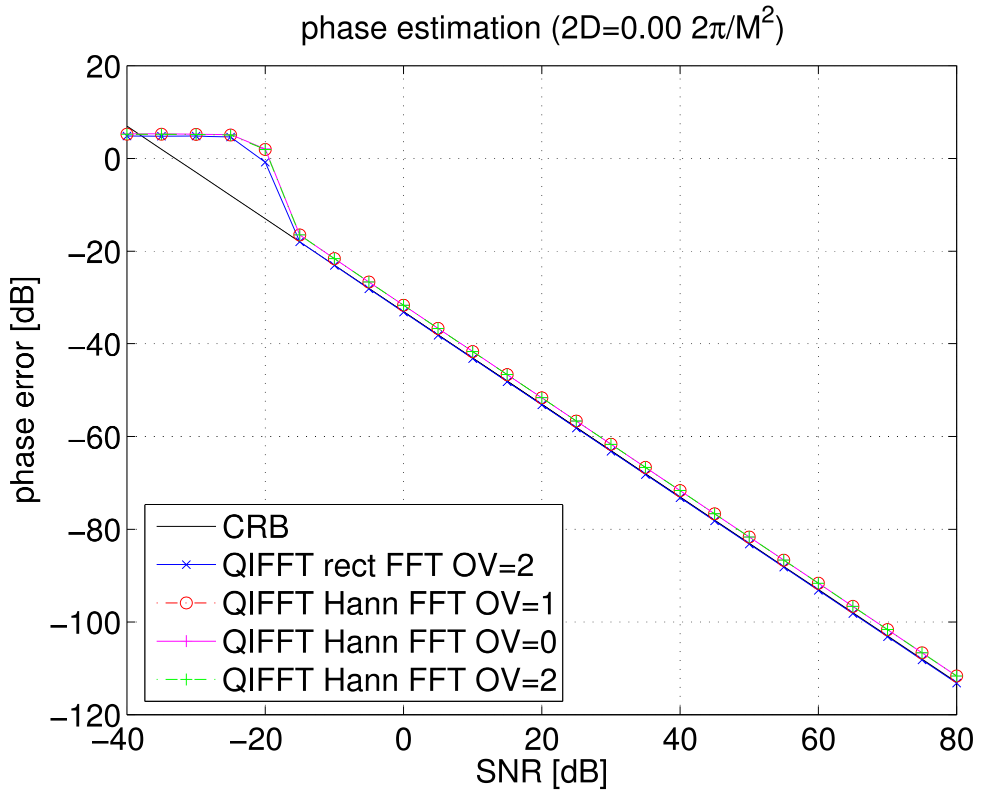

Figure 2: Estimation error and Cramer Rao bound for estimation of sinusoidal phase using QIFFT with different zero padding and different analysis windows. (window length M=1000, FFT size $N = 2^{\mathrm{nextpow2}(M)+OV}$).

The phase estimation error does not show any bias. Because the phase is constant within the peak a small error of the frequency estimator will not change the phase estimate. The threshold effects show a maximum error. This is due to the use of the 2π phase range which cannot create errors larger than $\pm\pi$.

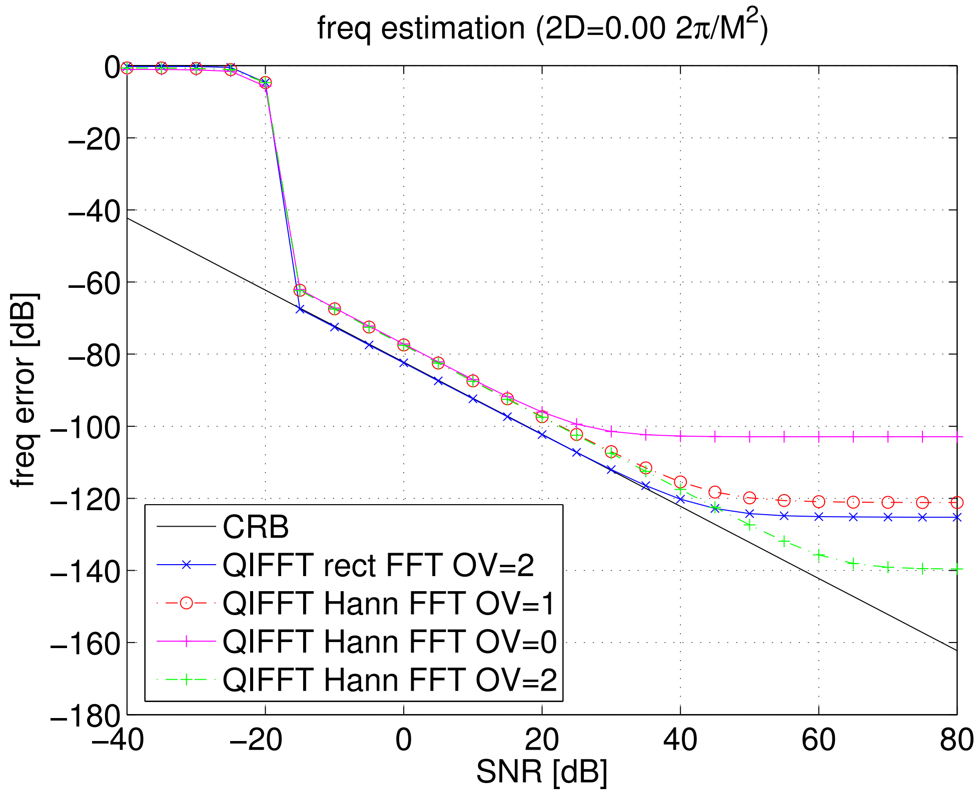

Figure 3: Estimation error and Cramer Rao bound for estimation of sinusoidal frequency using QIFFT with different zero padding and different analysis windows. (window length M=1000, FFT size $N = 2^{\mathrm{nextpow2}(M)+OV}$).

The frequency estimation error is similar to the amplitude estimation error with bias for high SNR and threshold for low SNR. The main difference is that the frequency error shows largest distance between the CRB and the estimator MSE. This due to the fact that the frequency estimation is the central part of the algorithm. Phase and amplitude use the frequency to determine their estimates. For amplitude and for frequency however, the final estimate does not change strongly with the frequency position such that they are less influenced by noise. Due to the flat top of the peak however, the frequency estimate is influenced much more by the noise such that it shows the largest sensitivity to noise. Note that the sensitivity stronger for Hanning windows which have a mainlobe with a larger plateau which is easily affected by noise;

6 non stationary sinusoids

Real world signals are never stationary.

Non-stationary sinusoids have been studied either

- with linear AM/FM

$$s(n) = (A + a(n - n_0))e^{i(\phi+\Omega(n-n_0)+D(n-n_0)^2)}, \tag{22}$$

- or with linear FM and exponential AM

$$s(n) = Ae^{a(n-n_0)}e^{i(\phi+\Omega(n-n_0)+D(n-n_0)^2)}. \tag{23}$$

To understand the impact of the time varying parameters a mathematical study of the spectral peak and its local maximum as a function of the parameters and the analysis window is required.

- For the complete linear model there exist only approximate solutions if the analysis window is Gaussian [Pee01].

- For the exponential amplitude model and a Gaussian window a complete mathematical solution is possible [AS05].

We reproduce the results for the exponential amplitude evolution and a Gaussian analysis window

$$w(n) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{n^2}{2\sigma^2}} = \sqrt{\frac{p}{\pi}} e^{-pn^2}, \tag{24}$$

with the shortcut notation $p = \frac{1}{2\sigma^2}$. Following [AS05] the FT spectrum is

$$X(\omega) = \sum_{n=-\infty}^{\infty} w(n) s(\omega) e^{-j\omega n} = e^{u(\omega)+jv(\omega)}. \tag{25}$$

The log amplitude spectrum $u(\omega)$ is given by

$$u(\omega) = \log(A) + \frac{a^2}{4p} - \frac{1}{4}\log(1 + (\frac{D}{p})^2) - \frac{p}{4(p^2+D^2)}[\omega - \Omega - \frac{aD}{p}]^2, \tag{26}$$

and the phase spectrum $v(\omega)$ is given by

$$v(\omega) = \phi + \frac{a^2}{4D} + \frac{1}{2}\text{atan}(\frac{D}{p}) - \frac{D}{4(p^2 + D^2)}[\omega - \Omega + \frac{pa}{D}]^2 \tag{27}$$

Slightly different results are obtained by means of second order Taylor approximation of the FT spectrum of eq. (22).

Note, that the amplitude and the phase spectrum of the exponential AM linear FM chirp are exactly quadratic functions such that the QIFFT method can be used to estimate all parameters from the three central bins of the main lobe of the peak.

6.1 Bias in th QIFFT method

Apply QIFFT method to the log amplitude an phase spectra to understand the bias.

Frequency:

- local maximum is at

$$\hat{\Omega} = \max_{\omega} u(\omega) = \Omega + \frac{aD}{p} \tag{28}$$

- the amplitude at that position is

$$u(\hat{\Omega}) = \log(\hat{A}) = e^{\log(A) + \frac{a^2}{4p} - \frac{1}{4}\log(1+(\frac{D}{p})^2)}, \tag{29}$$

- and the phase estimate is

$$v(\hat{\Omega}) = \hat{\phi} = \phi - \frac{a^2 D}{4p^2} + \frac{1}{2}\text{atan}(\frac{D}{p}). \tag{30}$$

In the general case these estimates do not match the correct values.

- The frequency estimate is biased if frequency slope D and the log amplitude slope a are present,
- the amplitude estimate $\hat{A}$ is biased if frequency slope D or log amplitude slope a are present,
- the phase estimate $\hat{\phi}$ is biased whenever the frequency slope is not zero.

Note, that amplitude and phase bias may significantly increase the residual energy if no bias correction scheme is applied.

This is especially true for vibrato signals.

6.2 slope estimation

The advantage of the analytic results is that the bias can simply be corrected as soon as log amplitude slope and frequency slope are estimated.

The first and second order derivatives with respect to ω of the log amplitude spectrum $u(\omega)$ and the phase spectrum $v(\omega)$ at the position of the local maximum are

$$v'(\hat{\Omega}) = -\frac{a}{2p}, \quad (31)$$

$$u''(\hat{\Omega}) = -\frac{p}{2(p^2+D^2)}, \quad (32)$$

$$v''(\hat{\Omega}) = -\frac{D}{2(p^2+D^2)}. \quad (33)$$

From these equations [AS05] have derived an estimate for a and D as follows

$$\hat{a} = -2pv'(\hat{\Omega}) \quad (34)$$

$$\hat{D} = p\frac{v''(\hat{\Omega})}{u''(\hat{\Omega})}. \tag{35}$$

These estimates may be used to correct the estimates above.

To be able to apply the bias correction scheme to non Gaussian windows a linear scaling of the correction factors has been proposed in [AS05]. Scaling factors have been optimized using signals with slight or medium modulation.

6.3 Alternative approach

For eq. (22) the bias disappears completely whenever $D = 0$.

The frequency slope estimator that has been derived in [Pee01] for eq. (22) is the same as the frequency slope estimator for the exponential AM signal.

Experimentally one obtains could frequency slope estimation with this estimator even for non Gaussian windows.

in this case the effective p of the non Gaussian window is simply the std deviation of the window itself.

- estimation of the frequency slope using the method shown above
- demodulation of the signal related to the spectral peak by means of multiplication with a complex exponential chirp with frequency slope $-D$.

$$s_d(n) = e^{-Dn^2}, \tag{36}$$

- application of the QIFFT method.

- approximate demodulation can be obtained by means of convolution of the spectral peak to be analyzed and the main lobe of the deconvolution signal $s_d(n)$.

6.4 Experimental investigation of the bias correction effect

Experimental investigation of the estimation errors for different methods using the signal model in eq. (22) and randomly selected signal parameters can be used to compare estimator performance.

Range of randomly selected signal parameters (uniform distribution):

- frequency: Ω selected from $[0.1, 0.3]\pi$),
- phase ϕ selected from $[-\pi, \pi]$,
- amplitude slope a selected from $[-1, 1]A/M$,
- frequency slope D selected from $[-2, 2]2\pi/M^2$ (frequency changes within a window by not more then $\frac{2\pi}{M}$).

Note that for real world signals with vibrato frequency slope increases linearly with partial number.

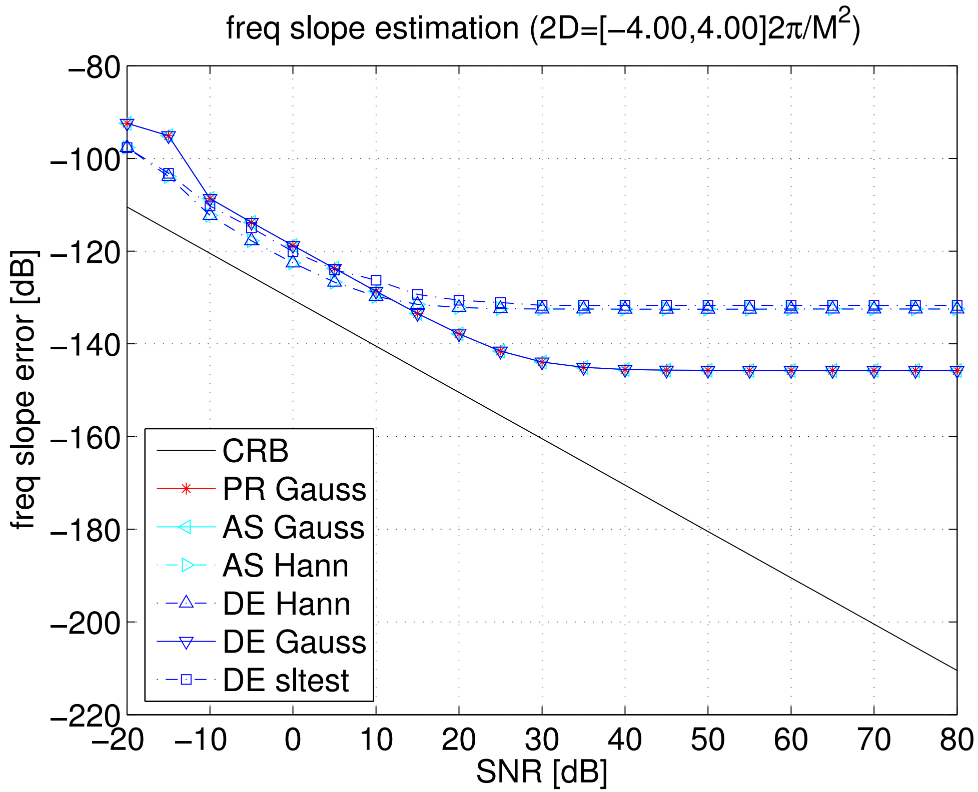

Figure 4: Estimation error and Cramer Rao bound for estimation of frequency slope for using different analysis windows and different estimation procedures . (window length M=1001, FFT size $N = 4096$).

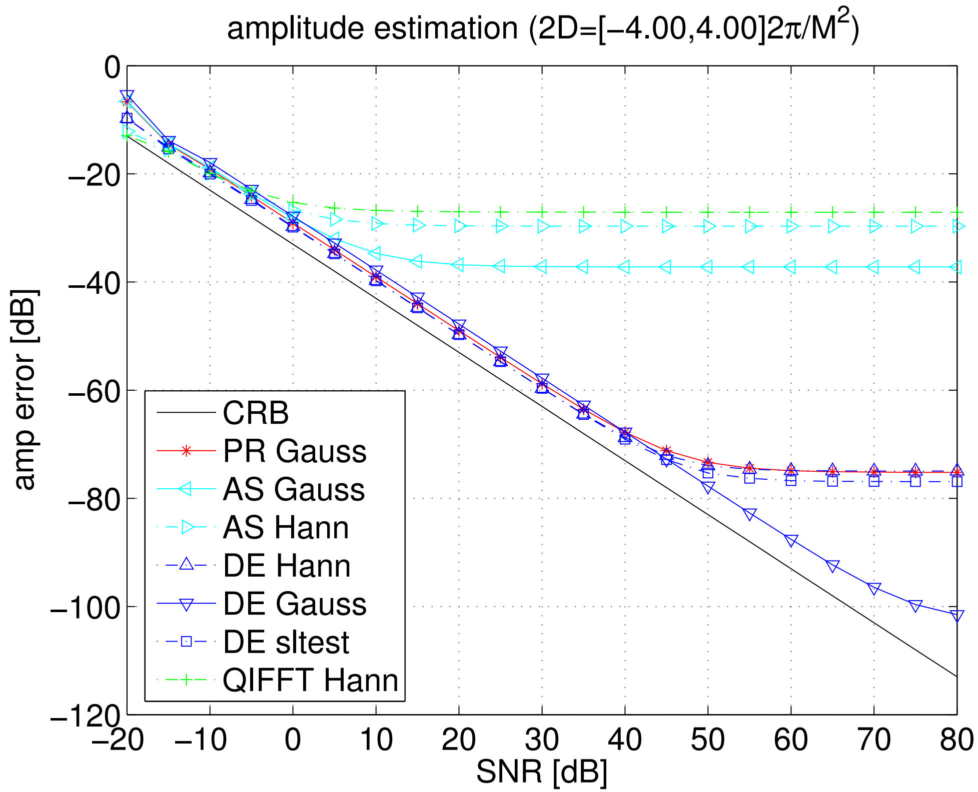

Figure 5: Estimation error and Cramer Rao bound for estimation of amplitude for using different analysis windows and different estimation procedures . (window length M=1001, FFT size $N = 4096$).

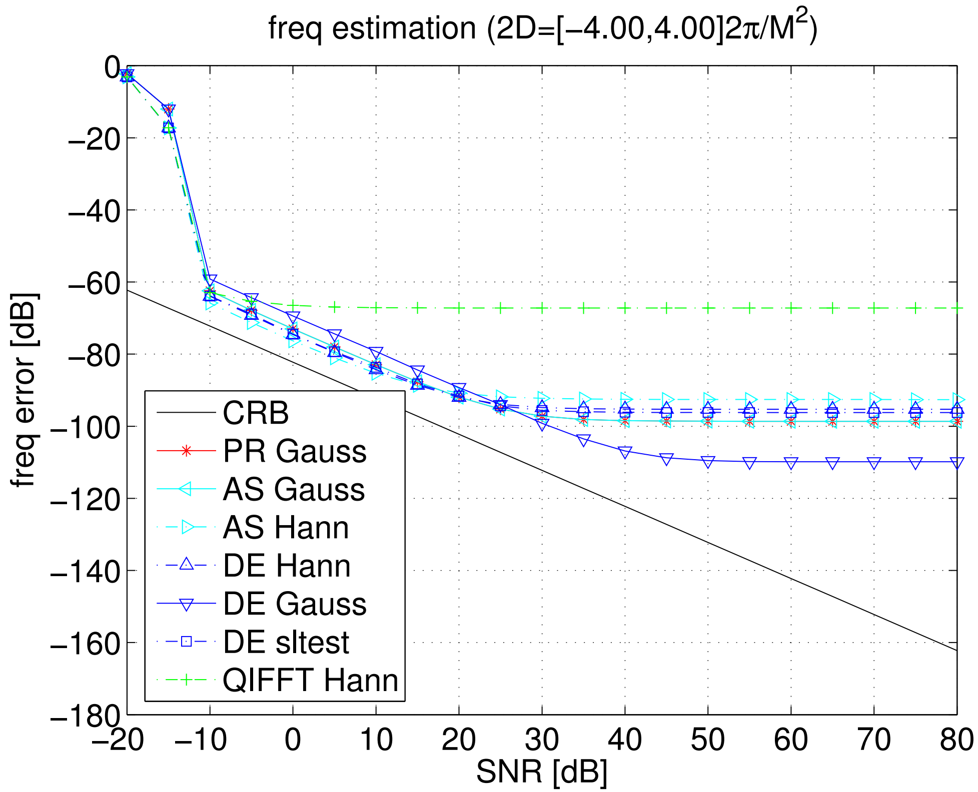

Figure 6: Estimation error and Cramer Rao bound for estimation of frequency for using different analysis windows and different estimation procedures . (window length M=1001, FFT size $N = 4096$).

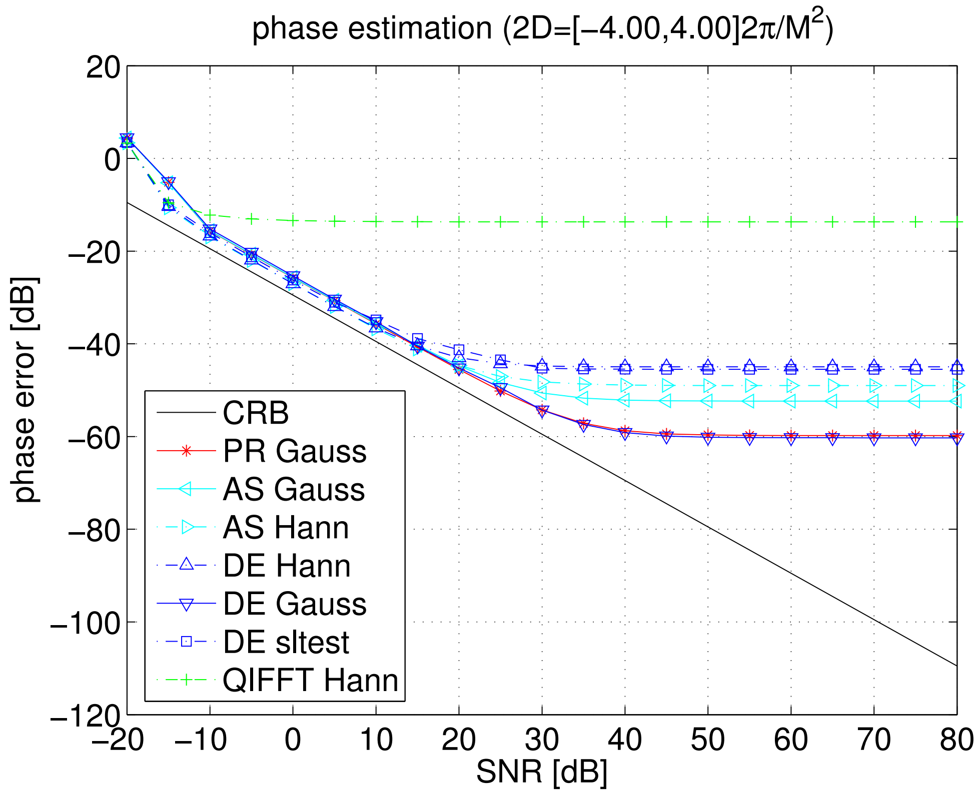

Figure 7: Estimation error and Cramer Rao bound for estimation of phase for using different analysis windows and different estimation procedures . (window length M=1001, FFT size $N = 4096$).

7 Sinusoidal continuation problem

After the estimation of the sinusoidal parameters for the spectral peaks in the individual frames these peaks have to be connected to form sinusoidal trajectories.

There have been proposed many algorithms to find proper peak connections. Because different situations (vibrato, polyphony, noise level) require different approaches no algorithm is best for all situations. The original algorithm has been proposed in [MQ86]. It is based on the simple idea to connect each peak in the previous frame to the peak in the next frame that is closest in frequency.

This algorithm may create unreasonable jumps. An improved strategy compares amplitude and frequency difference for the candidates to connect and connects only peaks that do not exceed a minimum variation for both parameters. Unconnected peaks belong to dying partials. Peaks without any connections may represent a new born sinusoid [ABLS02].

The variation thresholds can be adapted to favor smoothness of amplitude and frequency trajectories.

Recent algorithms try to incorporate a trajectory model into the peak continuation algorithm [LMR04].

8 Parameter interpolation

For synthesis of the sinusoid from the estimated parameters an interpolation from the analysis frame rate to the sample rate has to be obtained.

The problem has been solved in [MQ86].

Given are frame parameters of frame at position n_i, $[A(n_i), \phi(n_i), \omega(n_i)]$ and the following frame at position n_{i+1}, $[A(n_{i+1}), \phi(n_{i+1}), \omega(n_{i+1})]$

Use lowest order that uniquely determines an interpolating polynomial.

Amplitude interpolation:

- 2 points given $\longrightarrow$ linear interpolation

$$A(n) = \frac{A(n_i)(n_{i+1} - n) + A(n_{i+1})(n - n_i)}{n_{i+1} - n_i} \tag{37}$$

Phase and frequency are not independent, phase interpolation has to be consistent with the frequencies at the frame boundaries.

- 4 values are given phase at left and right frame boundary as well as frequency at frame boundaries.
- lowest polynomial order is 3,
- third order phase polynomial:

$$\phi(n) = qn^3 + rn^2 + sn + t \tag{38}$$

- second order frequency polynomial:

$$\omega(n) = 3qn^2 + 2rn + s \tag{39}$$

- coordinate system located at frame n_i argument is time difference $n_d = n_{i+1} - n_i$
- phase and frequency given at left frame boundary yields

$$\phi(0) = t = \hat{\phi}(n_i) \tag{40}$$

$$\omega(0) = s = \hat{\omega}(n_i) \tag{41}$$

- frequency constraints, phase at right boundary is known only up to an integer multiple of 2π

$$\phi(n_d) = qn_d^3 + rn_d^2 + \hat{\omega}(n_i)n_d + \hat{\phi}(n_i) = \hat{\phi}(n_{i+1}) + 2\pi M \quad (42)$$

$$\omega(n_d) = 3qn_d^2 + 2rn_d + \hat{\omega}(n_i) = \hat{\omega}(n_{i+1}) \quad (43)$$

- 3 unknowns and 2 equations, solving for q and r we get a solution depending on M

$$r = \frac{3}{n_d^2}(\hat{\phi}(n_{i+1}) - (\hat{\omega}(n_i)n_d + \hat{\phi}(n_i)) + 2\pi M) \quad (44)$$

$$-\frac{1}{n_d}(\hat{\omega}(n_{i+1}) - \hat{\omega}(n_i)) \quad (45)$$

$$q = -\frac{2}{n_d^3}(\hat{\phi}(n_{i+1}) - (\hat{\omega}(n_i)n_d + \hat{\phi}(n_i)) + 2\pi M) \quad (46)$$

$$+\frac{1}{n_d^2}(\hat{\omega}(n_{i+1}) - \hat{\omega}(n_i)) \quad (47)$$

- select M we require minimum curvature of the frequency trajectory, curvature is proportional to q, so we select M that minimizes

$$\mathrm{MIN} = q(M)^2 = (\frac{-2\Delta_\phi - 4\pi M + n_d(\hat{\omega}(n_{i+1}) - \hat{\omega}(n_i))}{n_d^3})^2 \tag{48}$$

where $\Delta_\phi = \hat{\phi}(n_{i+1}) - (\hat{\omega}(n_i)n_d + \hat{\phi}(n_i))$

- setting the derivative with respect to M to zero we get

$$0 = n_d(\hat{\omega}(n_{i+1}) - \hat{\omega}(n_i)) - 2\Delta_\phi - 4\pi M \tag{49}$$

- and solving for M yields

$$\hat{M} = \frac{1}{2\pi}(\frac{n_d}{2}(\hat{\omega}(n_{i+1}) - \hat{\omega}(n_i)) - \hat{\phi}(n_{i+1}) + \hat{\omega}(n_i)n_d + \hat{\phi}(n_i)) \tag{50}$$

- the M selected has to be integer so we select $M = \mathrm{round}(\hat{M})$.

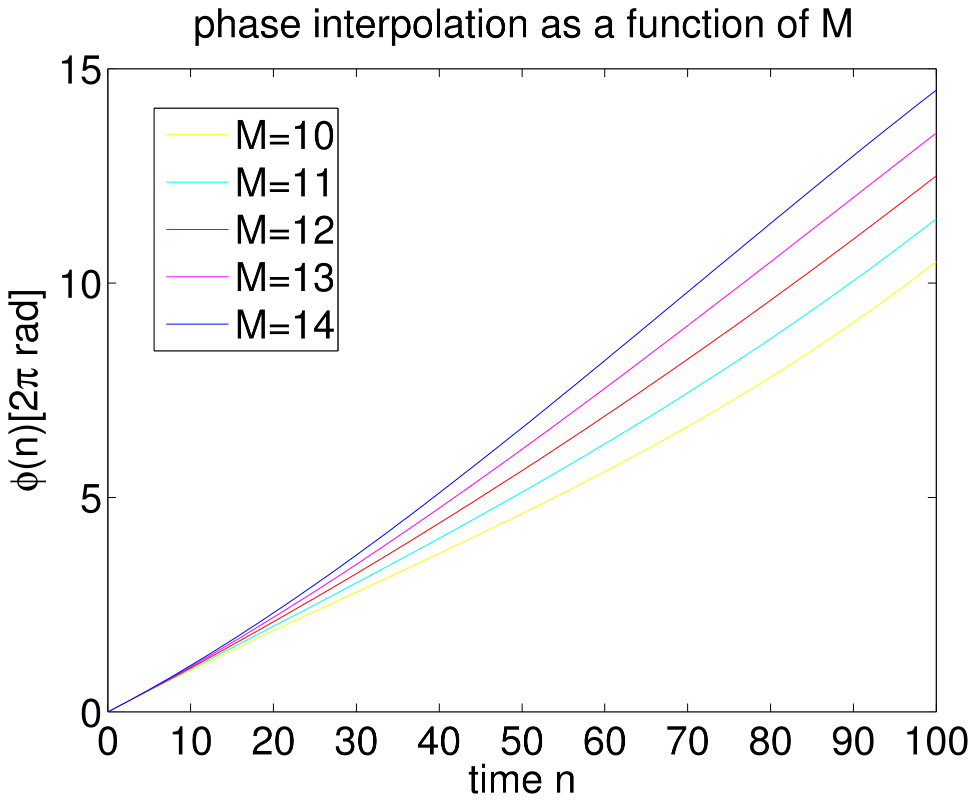

Figure 8: phase interpolation for varying M.

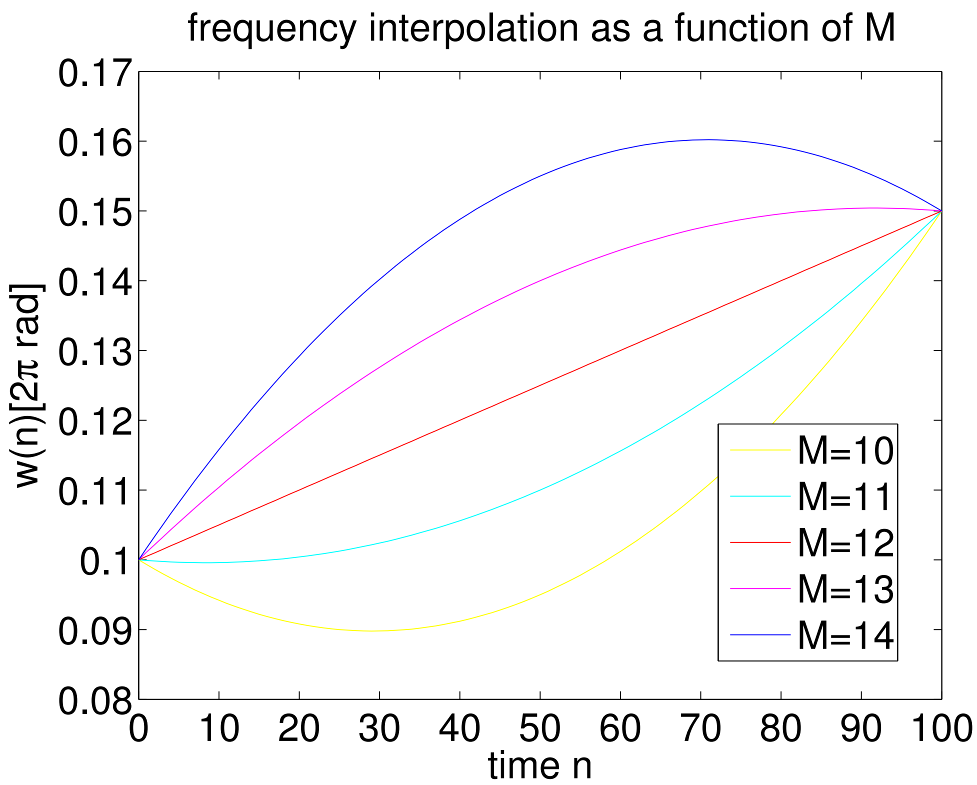

Figure 9: frequency interpolation for varying M. Limiting values are $\hat{\omega}(n_i)) = 0.1$ and $\hat{\omega}(n_{i+1})) = 0.15$ (normalized frequency).

References

[ABLS02] X. Amatriain, J. Bonada, A. Loscos, and X. Serra. Spectral processing. In U. Zölzer, editor, *Digital Audiuo Effects*, chapter 10, pages 373–438. John Wiley & Sons, 2002. 3, 43

[AS05] M. Abe and J. O. Smith. AM/FM rate estimation for time-varying sinusoidal modeling. In *Proc. Int. Conf. on Acoustics, Speech and Signal Processing*, pages 201–204 (Vol. III), 2005. 30, 34, 35

[HMW01] S.W. Hainsworth, M.D. Macleod, and P.J. Wolfe. Analysis of reassigned spectrograms for musical transcription. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 23–26, 2001. 8

[Kay88] S. Kay. *Modern Spectral Estimation*. Prentice Hall, 1988. 20

[LMR04] M. Lagrange, S. Marchand, and J-B. Rault. Using linear prediction to enhance the tracking of partials. In *Proc. Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, 2004. 44

[MQ86] R. J. McAulay and T. F. Quatieri. Speech analysis-synthesis based on a sinusoidal representation. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 34(4):744–754, 1986. 43, 45

[Pee01] G. Peeters. *Modèles et modification du signal sonore adapté à ses charactéristiques locales*. PhD thesis, Univertsité Paris 6, 2001. french only, http://recherche.ircam.fr/equipes/analyse-synthese/peeters/ARTICLES/Peeters_2001_PhDThesisv1.1.pdf. 30, 36

[RB98] B. Ristic and B. Boashash. Comments on ”The Cramer-Rao lower bounds for signals with constant amplitude and polynomial phase”. *IEEE Transactions on Signal Processing*, 46(6):1708–1709, 1998. 21

[Röb06a] A. Röbel. Analysis, modelling and transformation of audio signals - Part I: Fundamentals of discrete fourier analysis. lecture slides, 2006. AMT : Part I. 13

[Röb06b] A. Röbel. Analysis, modelling and transformation of audio signals - Part II: Analysis/resynthesis with the short time fourier transform. lecture slides, 2006. AMT : Part II. 9

[Röb06c] A. Röbel. Analysis, modelling and transformation of audio signals - Part III: signal modifications using the STFT. lecture slides, 2006. AMT : Part III. 3

[Röb06d] A. Röbel. Analysis, modelling and transformation of audio signals - Part IV: Source filter modeling and spectral envelope estimation. lecture slides, 2006. AMT : Part IV. 10

[Röb06e] A. Röbel. Analysis, modelling and transformation of audio signals - Part V: Fundamental frequency estimation. lecture slides, 2006. AMT : Part V. 3

[Rod97] X. Rodet. Musical sound signal analysis/synthesis: Sinusoidal+residual and elementary waveform models. In *Proc IEEE Time-Frequency and Time-Scale Workshop 97, (TFTS'97)*, page ??, 1997. 8

[RZR04] A. Röbel, M. Zivanovic, and X. Rodet. Signal decomposition by means of classification of spectral peaks. In *Proc. Int. Computer Music Conference (ICMC)*, pages 446–449, 2004. 8

[Ser97] X. Serra. *Musical signal processing*, chapter Musical Sound Modeling with Sinusoids and Noise, pages 91–122. Studies on New Music Research. Swets & Zeitlinger B. V., 1997. 3

[Tho82] D. J. Thomson. Spectrum estimation and harmonic analysis. *Proceedings of the IEEE*, 70(9):1055–1096, 1982. 8