CENTER FOR COMPUTER RESEARCH IN MUSIC AND ACOUSTICS
DEPARTMENT OF MUSIC, STANFORD UNIVERSITY
REPORT NO. STAN-M-116

Design Criteria for the Quadratically Interpolated FFT Method (III): Bias due to Amplitude and Frequency Modulation

October 13, 2004

Mototsugu Abe and Julius O. Smith III
E-Mail: {abemoto, jos}@ccrma.stanford.edu

Abstract

Quadratic peak interpolation in a zero-padded Fast Fourier Transform (FFT) has been widely used for sinusoidal parameter estimation in audio applications. In this series of papers, we investigate the estimation errors associated with various choices of analysis parameters, and provide precise criteria for designing the estimator. In this third paper, we focus on the bias caused by amplitude and frequency modulations of sinusoidal components. We investigate the first order modulations and derive a theory to predict the bias with a Gaussian window. We then extend it to other often-used windows, numerically confirm the results, and provide a criterion to determine the maximum window length as functions of prescribed maximum error bound and AM/FM rates of a signal. The results show, for example, if we are to bound the frequency bias below 1Hz using a Hann window for a sinusoid whose magnification rate is 100 times per second and chirp rate is 1kHz/s, we need to use a window whose length is less than 31ms.

1 Introduction

Among various approximate maximum likelihood (ML) sinusoidal parameter estimators [1, 2, 3, 4], quadratic interpolation of magnitude peaks in a zero-padded Fast Fourier Transform (FFT) [5] (referred to as the Quadratically Interpolated FFT, or QIFFT method) has been widely used due to its simplicity and accuracy. Although it works as an approximate ML estimator for well-separated sinusoids with a large zero-padding factor, its accuracy in practice is restricted by the choice of its design parameters, such as window type, FFT length and zero-padding factor. In this series of papers, we theoretically predict and numerically measure the estimation error bias associated with the choice of the design parameters, and precise design criteria for the estimator are defined. In the first paper [11], we investigate the bias due to quadratic interpolation and determine a criterion for choosing a zero-padding factor. In the second paper [12], we examine the bias due to interfering nearby signal components and detemine a criterion for a lower limit of window length.

In this final paper, we focus on the bias caused by amplitude and frequency modulations of sinusoidal components. A general requirement for the QIFFT, as well as most ML sinusoidal parameter estimators, is to choose window length enough short so that the sinusoidal components can be regarded as quasi-stationary. In practice, however, since audio signals are rarely completely time-invariant, the estimates are more or less biased by the AM and FM. Although accurate error bounds are important for designing the optimum estimator, only a few investigations on the bias with Gaussian windows can be found. Marques [6] proposed a method to estimate the AM and FM rates, and Peeters [7] derived mathematical formulas of the bias due to the AM and FM. Although their results are quite effective for a Gaussian window, since it is rarely used in practical applications due to its incompactness in both time and frequency domain, we need to determine a more general criterion for the bias with other often-used windows.

In this paper, we first derive mathematical formulas of the biases in frequency, amplitude and phase estimates with a Gaussian window. The results are essentially the same as the Peeters's formulas. We next extend them to other well-used windows, such as rectangular, Hann, Hamming and Blackman windows. We show approximate prediction formulas of the biases, and numerically confirm them. We finally provide a criterion to determine an upper limit of window length for given error bound and AM/FM rates.

2 QIFFT method

The Quadratically Interpolated FFT (QIFFT) method for estimating sinusoidal parameters from peaks in spectral magnitude data can be summarized as follows:

1. Calculate the amplitude and phase spectrum of audio data, by using a zero-padded FFT (points in Fig. 1).
2. Find the bin number of the maximum peak magnitude ($k_{\max}$).
3. Quadratically interpolate the log-amplitude of the peak using two neighboring samples (dotted line), and define $\hat{\Delta}$ as the distance in bins from $k_{\max}$ to the parabola peak.
4. Estimate the peak frequency in bins as $k_{\max} + \hat{\Delta}$, and define the estimated peak amplitude $\hat{A}$ as the interpolation (based on $\hat{\Delta}$) of the log amplitude samples at $k = k_{\max} - 1$, $k_{\max}$, and $k_{\max} + 1$.

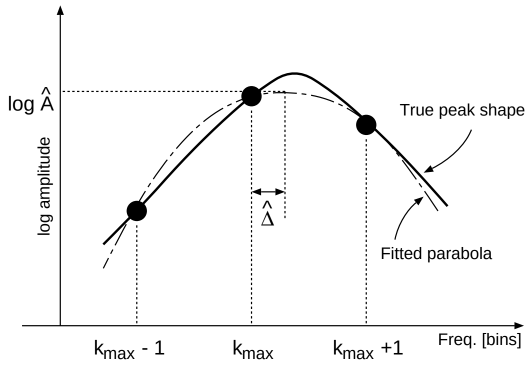


Figure 1: Quadratic interpolation of spectral peak

5. Estimate the phase, if needed, by interpolating[1] the phase spectrum based on the interpolated frequency estimate.
6. Subtract the peak from the FFT data for subsequent processing.
7. Repeat steps 2-6 above for each peak.

3 Preparation

For the sake of mathematical simplicity, we use continuous signal representation in the following discussion.

3.1 Signal

Let a sinusoid with the first order AM and FM be

$$x(t) = A_0 e^{\alpha_0 t} e^{j(\beta_0 t^2 + \omega_0 t + \phi_0)}, \tag{1}$$

where

$$\begin{aligned} A_0 &: \text{ instantaneous amplitude at } t = 0, \\ \omega_0 &: \text{ instantaneous frequency at } t = 0, \\ \phi_0 &: \text{ instantaneous phase at } t = 0, \\ \alpha_0 &: \text{ amplitude change rate}, \\ \beta_0 &: \text{ frequency change rate}. \end{aligned}$$

Without loss of generality, we assume the window center is placed at the time origin $t = 0$. A main concern of this paper is to clarify the dependency of the estimates ($\hat{\omega}_0$, $\hat{A}_0$ and $\hat{\phi}_0$) on the AM/FM rates (α_0 and β_0).

3.2 Typical AM/FM rates

Before detailed discussion, let us confirm typical values of α_0 and β_0 in actual audio. The root mean square (RMS) and maximum values in a female speech are shown in Table 1. For detailed discussion for obtaining these values, see appendix A.

[1] Quadratic interpolation is appropreate for sinusoids with FM.

Table 1: Typical values of the amplitude and frequency change rate and the fundamental frequency of a speech (from appendix A).

	RMS	MAX
α_0 [s^{-1}]	34	±100
β_0 [rad/s^2]	2300	±8000
ω_0 [rad/s]	1300	2500

Note that ω_0 and β_0 are values for the fundamental frequency, i,e, for the Nth partial, they are N times larger.

4 Bias with Gaussian Windows

Let a normalized Gaussian window be

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

where σ is the standard deviation ($\sqrt{1/e}$ width) of the Gaussian and $p \triangleq \frac{1}{2\sigma^2}$. The windowed Fourier transform of the signal is calculated as (see appendix B)

$$\begin{aligned} X(\omega) &= \int w(t)x(t)e^{-j\omega t}dt \\ &= Ae^{u(\omega)}e^{jv(\omega)}, \end{aligned} \tag{3}$$

where

$$u(\omega) = \frac{\alpha_0^2}{4p} - \frac{1}{4}\log(1+(\frac{\beta_0}{p})^2) - \frac{p}{4(p^2+\beta_0^2)}[\omega - \omega_0 - \frac{\alpha_0\beta_0}{p}]^2 \tag{4}$$

is the log-amplitude term and

$$v(\omega) = \phi_0 + \frac{\text{atan}(\beta_0/p)}{2} + \frac{\alpha_0^2}{4\beta_0} - \frac{\beta_0}{4(p^2+\beta_0^2)}[\omega - \omega_0 + \frac{p\alpha_0}{\beta_0}]^2 \tag{5}$$

is the phase term[2]. Applying the QIFFT method, we get the frequency estimate as[3]

$$\begin{aligned} \tilde{\omega}_0 &= \text{argmax}_{\omega}\ |X(\omega)| \\ &= \omega_0 + \frac{\alpha_0\beta_0}{p}, \end{aligned} \tag{6}$$

the amplitude estimate as

$$\begin{aligned} \tilde{A} &= |X(\tilde{\omega}_0)| \\ &= Ae^{\frac{\alpha_0^2}{4p} - \frac{1}{4}\log(1+(\frac{\beta_0}{p})^2)} \\ &\approx A\left(1 + \frac{\alpha_0^2}{4p} - \frac{\beta_0^2}{4p^2}\right), \end{aligned} \tag{7}$$

and the phase estimate as

$$\begin{aligned} \tilde{\phi} &= \text{ang}\,[X(\tilde{\omega}_0)] \\ &= \phi + \frac{\text{atan}(\beta_0/p)}{2} - \frac{\alpha_0^2\beta_0}{4p^2} \\ &\approx \phi + \frac{\beta_0}{2p}. \end{aligned} \tag{8}$$

Some characteristics of the biases can be predicted from these results:

[2]Although we show here the formulas for $\beta_0 \neq 0$, the following results are applicable to $\beta_0 = 0$ case.

[3]The quadratic interpolation for a Gaussian window is exact, so that the peak of the fitted parabola is the true peak. For the other windows, the estimates are biased by the interpolation [11].

Table 2: Equivalent standard deviations

Width	Rect	Hann	Hamm	Black
σ_0	0.288675	0.180756	0.200445	0.159485

Table 3: Theoretically predicted biases with Hann windows.

width	15ms	30ms	45ms	60ms
Freq. [%]	0.089	0.35	0.80	1.4
Amp. [%]	0.43	1.7	3.8	6.8
Phase [%]	0.53	2.1	4.3	6.7

For $\alpha_0 = 34[\mathrm{s}^{-1}]$, $\beta_0 = 2300[\mathrm{rad/s}^{-2}]$ and $\omega_0 = 1300[\mathrm{rad/s}]$. The frequency, amplitude and phase biases are normalized by ω_0, A_0 and π, respectively. Since the amplitude bias is not a monotonic distribution, the biases in the amplitude are the largest values below the α_0 and β_0.

1. The frequency estimate is biased by the multiple of AM and FM. This means the bias only appears when the sinusoid is modulated in both amplitude and frequency.
2. The amplitude estimate is biased by either AM or FM.
3. The phase estimate is mainly biased by FM. The bias only appears when the sinusoid is modulated in frequency.
4. Since $1/p$ is proportional to the square of the window length, use of a shorter window square-proportionally reduces the biases.

5 Bias with non-Gaussian windows

Due to its incompactness in both time and frequency domain, a Gaussian window is rarely used in practical applications. Since it is not always possible to induce analytic formulas for other windows, it is useful to provide a scheme to extend the above results to them. A natural approximation of a window with a Gaussian window may be setting the peak curvatures of the magnitude spectra to the same. This can be done analytically for some windows; at least numerically for any kind of windows. We show an analytical method for windows in the Blackman-Harris family here.

Let a window in the Blackman-Harris family be

$$w(t) = \sum_{l=-L}^{L} q_l e^{j2\pi t/T_w}, \tag{9}$$

where T_w denotes the window length, L is the order of the window, and q_ls are coefficients determine the window. Then the standard deviation of the equivalent Gaussian window can be calculated as (see appendix C)

$$\sigma = \sigma_0 T_w, \tag{10}$$

where

$$\sigma_0 \triangleq \sqrt{\frac{1}{12} + \frac{1}{q_0\pi^2}\sum_{l=1}^{L}\frac{(-1)^l q_l}{l^2}} \tag{11}$$

is a constants for each window. Some values of σ_0 are tabulated in Table 2.

For example, the standard deviation of the equivalent Gaussian to the 30ms Hann window can be calculated as

$$\begin{aligned} \sigma &= 0.180756 \times 0.03\ [\mathrm{s}] = 0.005423\ [\mathrm{s}], \\ p &= 1/2\sigma^2 = 17004\ [\mathrm{s}^{-2}]. \end{aligned} \tag{12}$$

Theoretically predicted biases with the 30ms Hann window are shown in Figs. 2 (3D plot) and 4 (contour plot). The biases for the typical AM/FM rates (RMS values in Table 1) are shown in Table 3.

Table 4: Parameters for experiment

Parameter	Unit	Value
A_0		1.0
ω_0	[rad/s]	1300
ϕ_0	[rad]	random
α_0	[/s]	-100 to 100
β_0	[rad/s^2]	-8000 to 8000

6 Experiments

6.1 Conditions for numerical experiments

The signal to be tested is a complex AM/FMed sinusoid as

$$x(n) = A_0 e^{\alpha_0 T n} e^{j(\beta_0 T^2 n^2 + \omega_0 T n + \phi_0)}, \tag{13}$$

where $T = 1/44100$[s] is the sampling period and n is a discrete time index. The signal parameters are shown in Table 4. We use the FFT size $N = 16384$ and examine the rectangular, Hann, Hamming and Blackman windows whose sizes $M = 661, 1323, 1983$ and 2645, which correspond to 15, 30, 45 and 60ms respectively. The zero-padding factors are larger than 5 for all the windows. By changing α_0 and β_0, we estimate $\{\hat{A}_0, \hat{\omega}_0, \hat{\phi}_0\}$ using the QIFFT method, and evaluate the differences between the true values and the estimates. For each $\{\alpha_0, \beta_0\}$, we prepare 16 sets of sinusoids having different phases, and take the maximum errors out of them. The frequency, amplitude and phase biases are normalized by ω_0, A_0 and π, respectively.

6.2 Results

The results for the 30ms Hann window are shown in Figs. 3 (3D plot) and 5 (contour plot). The biases for the typical AM/FM rates for the other windows are shown in Table 5.

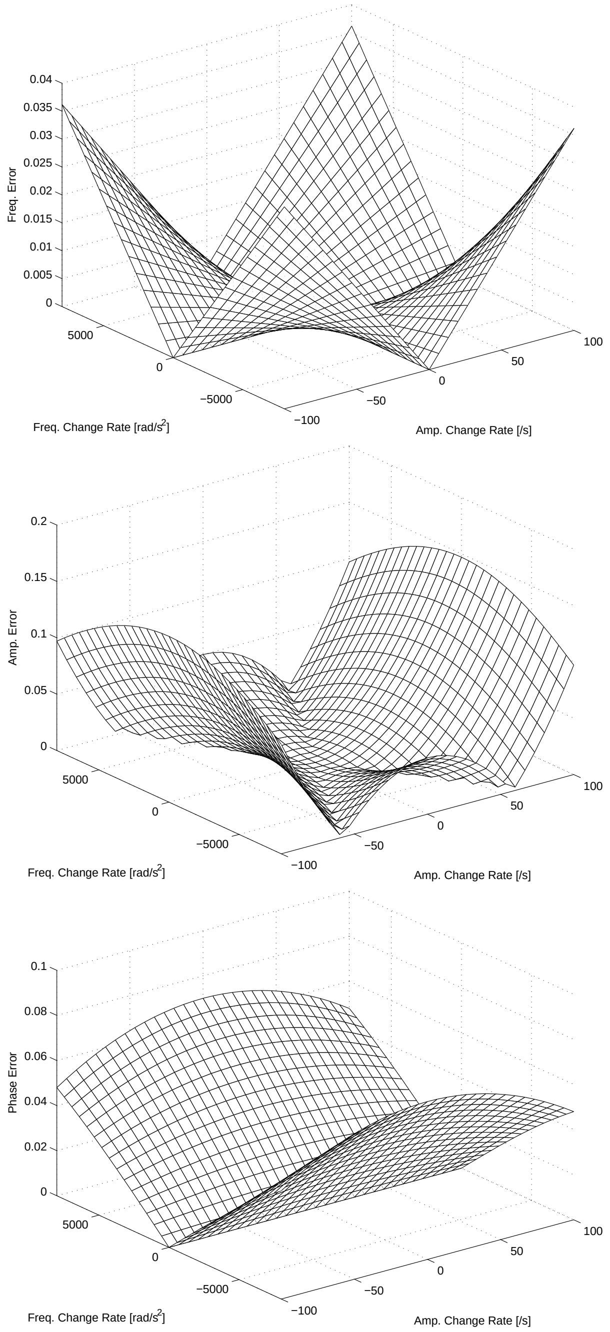
Comparing Figs. 3 and 2, we can confirm that the theoretical predictions qualitatively well approximate the experimental results. Although the theory slightly over-estimates the frequency bias, the amplitude and phase biases are quantitatively well fit to the theory as well.

Comparing Tables 3 and 5, we can confirm that 1) the biases are about square proportional to the window length as predicted in the theory, 2) the biases are inversely proportional to the main lobe width of the windows, i.e. Blackman < Hann < Hamming < rectangular, and 3) the 30ms windows are already too large to bound the amplitude and phase biases below 1%.

Table 5: Experimentally obtained biases

	Rectangular window			
width	15ms	30ms	45ms	60ms
Freq. [%]	0.090	0.37	0.87	1.9
Amp. [%]	1.1	4.4	10.0	18.3
Phase [%]	1.4	5.3	11.1	16.3
	Hann window			
width	15ms	30ms	45ms	60ms
Freq. [%]	0.062	0.25	0.57	1.0
Amp. [%]	0.42	1.7	3.9	7.0
Phase [%]	0.53	2.1	4.5	7.4
	Hamming window			
width	15ms	30ms	45ms	60ms
Freq. [%]	0.082	0.33	0.75	1.4
Amp. [%]	0.52	2.1	4.8	8.6
Phase [%]	0.65	2.6	5.4	8.4
	Blackman window			
width	15ms	30ms	45ms	60ms
Freq. [%]	0.056	0.22	0.51	0.92
Amp. [%]	0.33	1.3	3.0	5.4
Phase [%]	0.41	1.6	3.5	5.9

For $\alpha_0 = 34[\mathrm{s}^{-1}]$, $\beta_0 = 2300[\mathrm{rad/s}^{-2}]$ and $\omega_0 = 1300[\mathrm{rad/s}]$. Since the amplitude bias is not a monotonic distribution, the errors in the amplitude are the largest values below these α_0 and β_0.



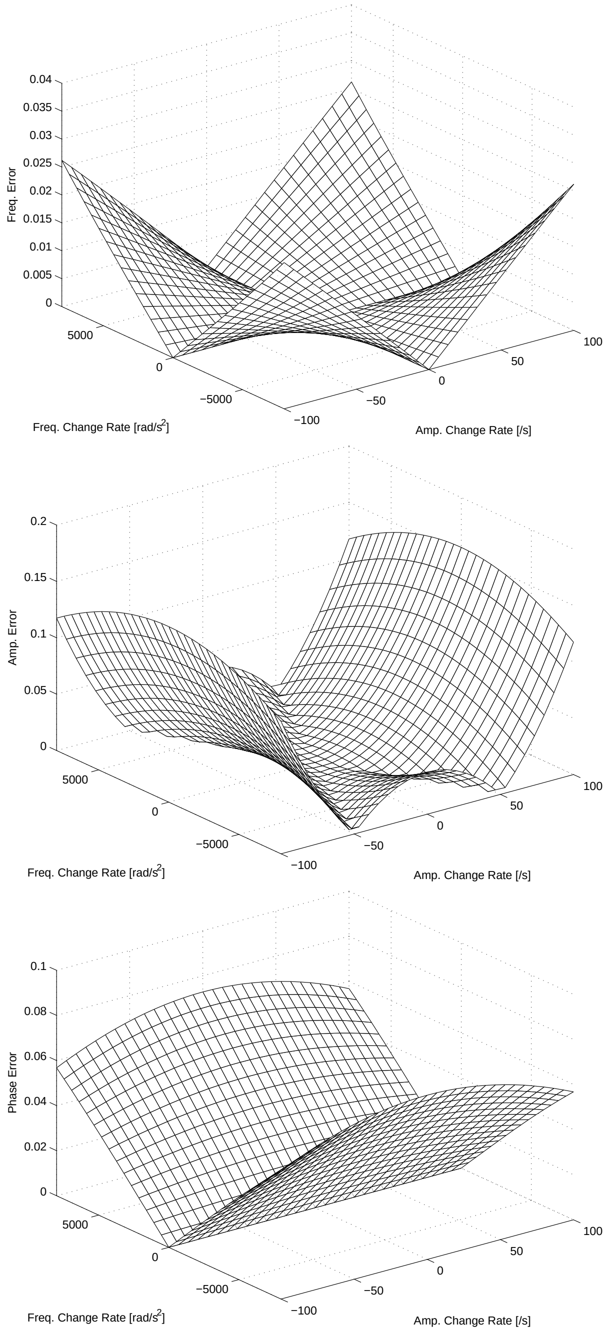


Figure 2: Theoretically predicted biases (3d plot): frequency (top), amplitude (middle), and phase (bottom).

Figure 3: Experimental results for a Hann window (3d plot): frequency (top), amplitude (middle), and phase (bottom).

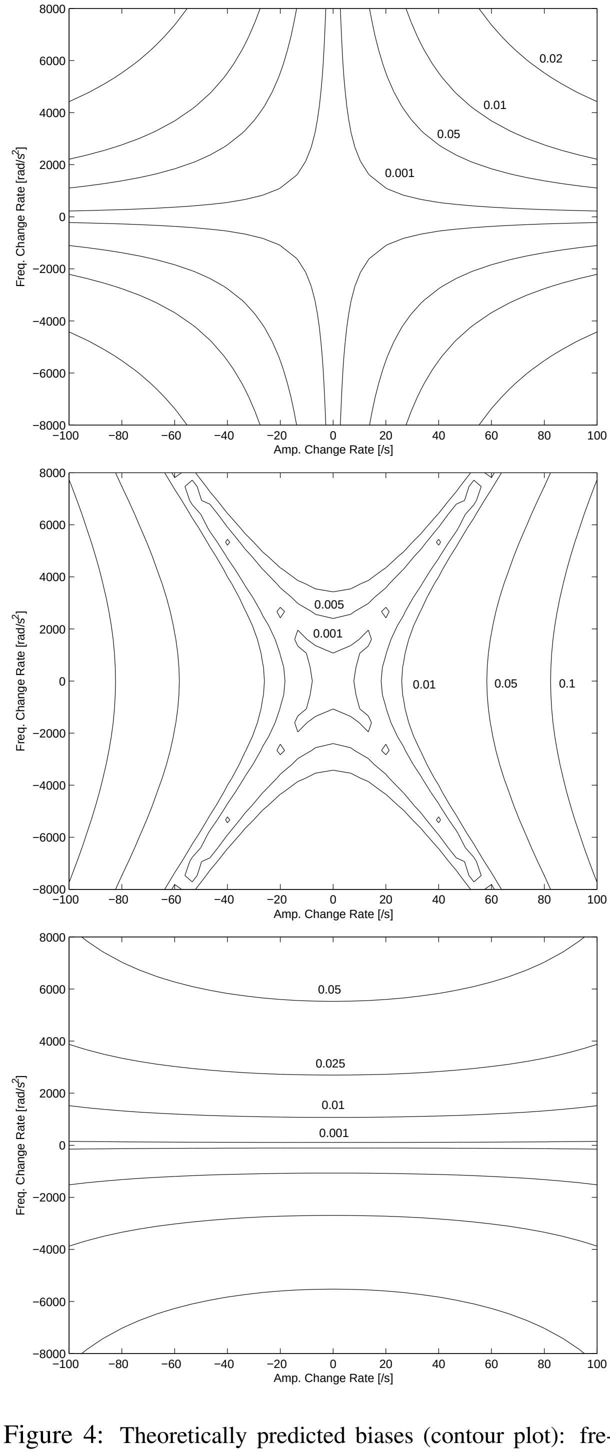


Figure 4: Theoretically predicted biases (contour plot): frequency (top), amplitude (middle), and phase (bottom).

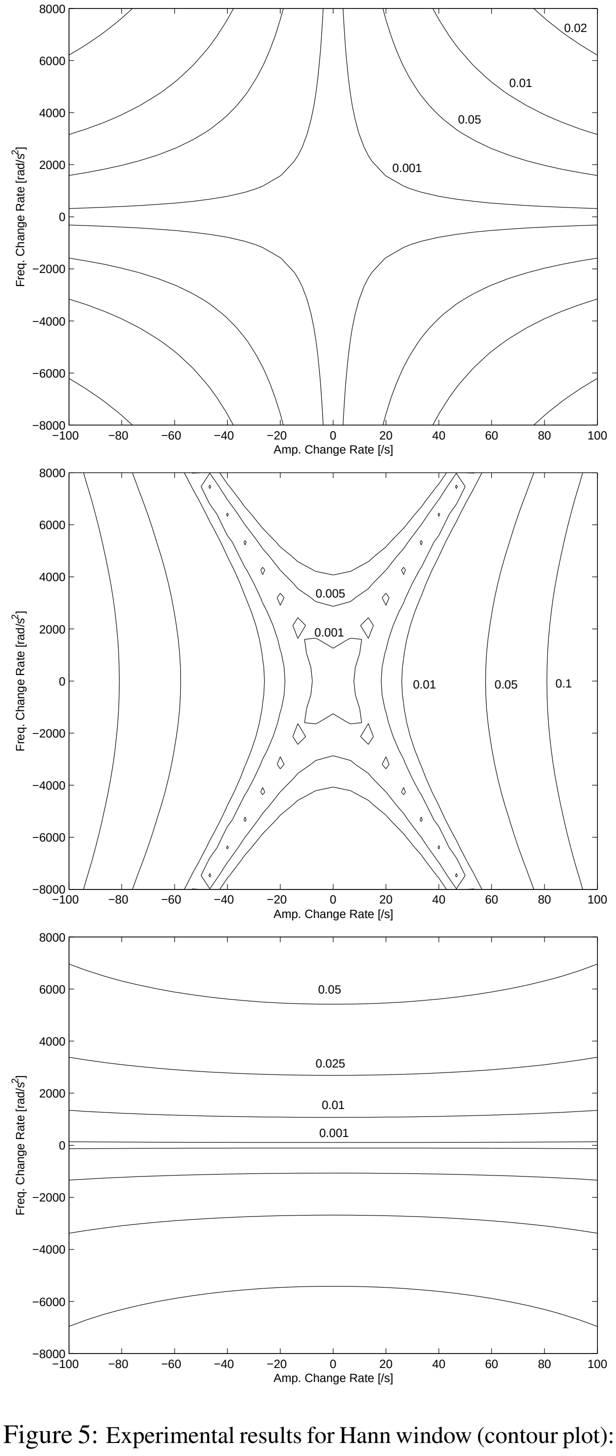


Figure 5: Experimental results for Hann window (contour plot): frequency (top), amplitude (middle), and phase (bottom).

7 Criterion on maximum window length

It may be convenient to formulate a criterion on the window length to bound the maximum biases below a prescribed amount for given AM/FM rates.

Let B_{f} be the maximum frequency bias in Hz, α_{m} be the maximum AM rate and β_{m} be the maximum FM rate. Then from Eqs.(6) and (10), we get

$$2\pi B_{\mathrm{f}} = 2\alpha_{\mathrm{m}}\beta_{\mathrm{m}}\sigma_0^2 T_w^2. \tag{14}$$

This leads

$$T_w \leq \sqrt{\frac{\pi B_{\mathrm{f}}}{\alpha_{\mathrm{m}}\beta_{\mathrm{m}}\sigma_0^2}}. \tag{15}$$

Similar bounds can be derived for the amplitude and phase biases. Let B_{a} be the amplitude bias which is the ratio to the true amplitude and B_{p} be the phase bias in radians. Then we get

$$T_w \leq \mathrm{Min}\left(\sqrt{\frac{2B_{\mathrm{a}}}{\sigma_0^2\alpha_{\mathrm{m}}^2}}, \left(\frac{B_{\mathrm{a}}}{\beta_{\mathrm{m}}^2\sigma_0^4}\right)^{\frac{1}{4}} \right), \tag{16}$$

$$T_w \leq \sqrt{\frac{B_{\mathrm{p}}}{\beta_{\mathrm{m}}\sigma_0^2}}. \tag{17}$$

For example, if we are to bound the maximum frequency bias of a Hann window below $B_{\mathrm{f}} = 1.16$Hz (10 cents of 200Hz), for $\alpha_{\mathrm{m}} = 68[\mathrm{s}^{-1}]$ and $\beta_{\mathrm{m}} = 4600[\mathrm{rad/s}^{-2}]$ (twice the RMSs), we have to set the length to

$$T_w \leq 18.9[\mathrm{ms}]. \tag{18}$$

8 Discussion and Summary

In this paper, we investigated the estimation bias due to AM and FM of sinusoidal components in the QIFFT method. We derived a theory to predict the biases in amplitude, frequency and phase estimates, for both Gaussian and non-Gaussian windows, and numerically confirmed it. The results show that the theory well fit to the experimentally obtained biases. We also provided a criterion for choosing window length which bounds the bias below a prescribed value.

Some important results in this paper are 1) if we use a 30ms Hann window, the biases in a natural speech are around a few percent, and 2) the biases are square proportional to the window length. This amount of the biases may cause audible errors, and become square-proportionally worse for a longer window. Another point to be mentioned is that the typical frequency change rate used in our experiments is for the fundamental frequency component in a speech. Since the rate is k times larger for the kth partial, it raises the phase bias linearly and the amplitude bias square-proportionally.

If we wish to predict the frequency bias for non-Gaussian windows more precisely, we may have to consider an additional compensation scheme. A simple solution is to adjust p to a more adequate value. From our informal experiment, $p \approx 24000$ seems to be well fit to the experimental result for the 30ms Hann window.

Note that the results in this paper are also applicable to other sinusoidal parameter estimation methods based on spectral peak detection. Even if we use the true ML estimator for a stable sinusoid [1], the bias due to AM/FM still exists as well.

As a method to reduce the AM/FM bias, the multi-resolution STFT [8] in which shorter windows are used for higher frequency partials can be effectively used. Another method is to estimate AM and FM rates explicitly, by which we may correct the bias as well as being able to utilize the estimated rates for further spectral modeling. Some methods have been proposed for this purpose, such as the frequency-varying sinusoidal modeling [6], the dynamic channel vocoder model [9], and the linear chirp modeling [10]. We also propose a method to estimate α_0 and β_0 by extending the QIFFT method in our related paper [14].

References

[1] D. C. Rife and R. R. Boorstyn: "Single-Tone Parameter Estimation from Discrete-Time Observations," IEEE Trans. Info. Theory, 20, 5, 591/598 (1974).

[2] D. J. Thomson: "Spectrum Estimation and Harmonic Analysis," Proc. of the IEEE, 70, 9, 1055/1096 (1982).

[3] B. G. Quinn: "Estimation of Frequency, Amplitude, and Phase from the DFT of a Time Series," IEEE Trans. Signal Processing, 45, 3, 814/817 (1997).

[4] M. D. Macleod: "Fast Nearly ML Estimation of the Parameters of Real or Complex Single Tones or Resolved Multiple Tones," IEEE Trans. Signal Processing, 46, 1, 141/148 (1998).

[5] J. O. Smith III and X. Serra: "PARSHL: A program for the analysis/synthesis of inharmonic sounds based on a sinusoidal representation," in Proc. ICMC'87, available at http://www-ccrma.stanford.edu/~jos/parshl.

[6] J. S. Marques and L. B. Almeida: "Frequency-Varying Sinusoidal Modeling of Speech," IEEE Trans. Acoust., Speech, Signal Processing, 37, 5, 763/765 (1989).

[7] G. Peeters and X. Rodet: "SINOLA: A New Analysis/Synthesis Method using Spectrum Peak Shape Distortion, Phase and Reassigned Spectrum," Proc. ICMC'99 (1999).

[8] S. N. Levine and J. O. Smith III: "A Switched Parametric & Transform Audio Coder," Proc. IEEE ICASSP'99, 985/988 (1999).

[9] A. S. Master: "Sinusoidal Modeling Parameter Estimation Via a Dynamic Channel Vocoder Model," Proc. IEEE ICASSP'02, 1857/1860 (2002).

[10] A. S. Master and Y. Liu: "Nonstationary Sinusoidal Modeling with Efficient Estimation of Linear Frequency Chirp Parameters," Proc. IEEE ICASSP'03, 656/659 (2003).

[11] M. Abe and J. O. Smith III: "Design Criteria for the Quadratically Interpolated FFT Method (I): Bias due to Interpolation," Technical Report STAN-M-114, Dept. of Music, Stanford University, August, (2004).

[12] M. Abe and J. O. Smith III: "Design Criteria for the Quadratically Interpolated FFT Method (II): Bias due to Interfering Components," Technical Report STAN-M-115, Dept. of Music, Stanford University, August, (2004).

[13] M. Abe and J. O. Smith III: "Correcting Bias in a Sinusoidal Parameter Estimator based on Quadratic Interpolation of FFT Magnitude Peaks," Technical Report STAN-M-117, Dept. of Music, Stanford University, August, (2004).

[14] M. Abe and J. O. Smith III: "AM/FM Rate Estimation and Bias Correction for Time-Varying Sinusoidal Modeling," Technical Report STAN-M-118, Dept. of Music, Stanford University, August, (2004).

Appendix

A Example of AM/FM rates

In this appendix, we confirm AM and FM rates in an actual sound. The sample to be tested is a female speech of about 7 second long. The waveform and spectrogram are shown in Fig.6(a) and (b).

To estimate the amplitude change rate, we first get the amplitude envelope $a(n)$ by calculating the moving average of the root-mean-square (RMS) of the waveform, as shown in 6(c). We used 20ms Hann window for the moving average. We may roughly estimate $\alpha(n)$ as

$$\alpha(n) = \frac{a(n+1) - a(n-1)}{2Ta(n)}, \tag{19}$$

which is sometimes referred as a relative derivative or a log-derivative of $a(n)$. T denotes the sampling period of the amplitude envelope.

The result is shown in Fig.6(d). Since the relative derivatives become infinity at each onset, which should not be modeled by sinusoidal components, we may ignore the large spikes. Therefore, we may roughly say that the maximum amplitude change rate is around $\pm 100[\mathrm{s}^{-1}]$ and the RMS is about $34[\mathrm{s}^{-1}]$.

To estimate the frequency change rate, we extract the fundamental frequency $f(n)$ of the signal by simply picking the peaks of the lowest frequency component in the spectrogram. By evaluating flatness of the spectrum around the peaks, we discriminate consonants and extract vowel intervals. The result is shown in Fig.6(e).

We can roughly estimate the frequency change rate β as

$$\beta(n) = 2\pi \frac{f(n+1) - f(n)}{2T}, \tag{20}$$

where T denotes the sampling period of the STFT.

The result is shown in Fig.6(f). We may roughly say the maximum frequency change rate is about ± 8000 [rad/s^2], and the RMS is about 2300 [rad/s^2].

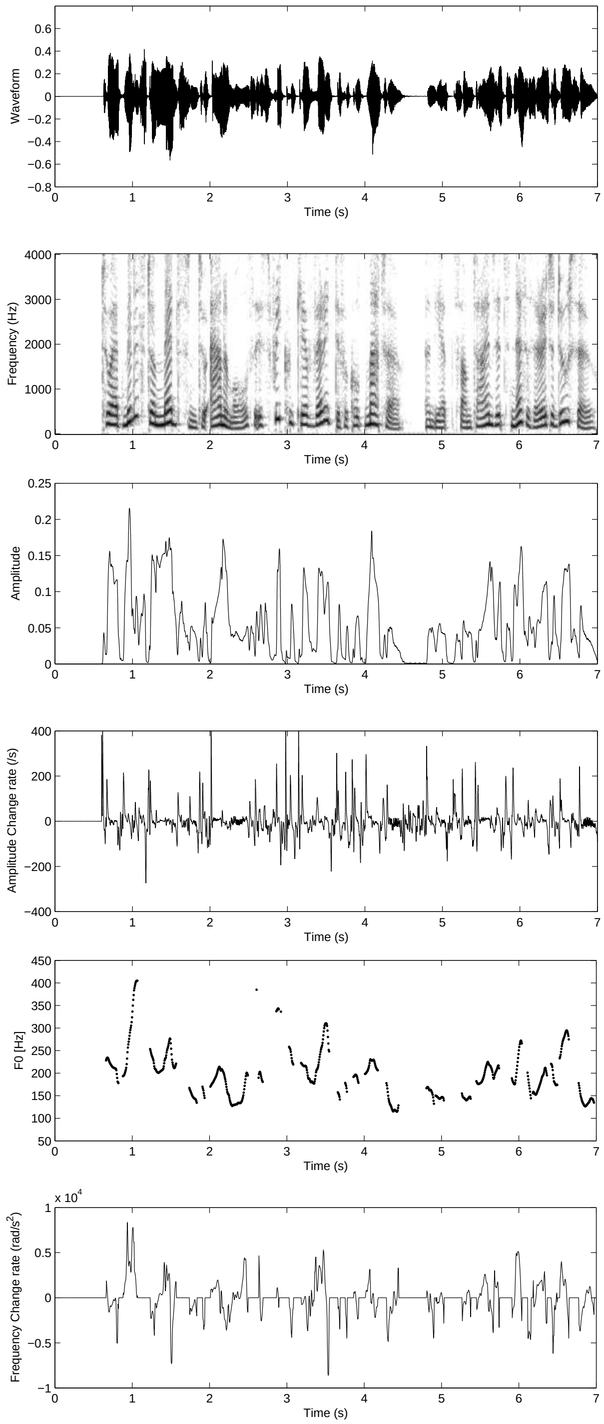


Figure 6: Female speech: (a) waveform, (b) spectrogram, (c) amplitude envelope (20ms Hann windowed average), (d) amplitude change rate, (e) estimated fundamental frequency (42ms Hann windowed DFT, zero pad factor = 8, QIFFT), (f) frequency change rate of the fundamental frequency.

B Gaussian windowed Fourier transform of an AM/FM sinusoid

Let us calculate the windowed Fourier transform of the AM/FM sinusoid shown in Eq.(1).

$$
\begin{aligned}
X(\omega) &= \int w(t)x(t)e^{-j\omega t}dt \\
&= \int e^{-pt^2} A e^{\alpha t} e^{j(\beta t^2+\omega_0 t+\phi)} e^{-j\omega t}dt \\
&= Ae^{j\phi} \int e^{-(p-j\beta)t^2+[\alpha-j(\omega-\omega_0)]t}dt \\
&= Ae^{j\phi} \int e^{-(p-j\beta)[t^2-\frac{[\alpha-j(\omega-\omega_0)]}{(p-j\beta)}t]}dt \\
&= Ae^{j\phi} e^{\frac{[\alpha-j(\omega-\omega_0)]^2}{4(p-j\beta)}} \int e^{-(p-j\beta)[t-\frac{[\alpha-j(\omega-\omega_0)]}{2(p-j\beta)}]}dt \\
&= \sqrt{\pi}Ae^{j\phi} e^{\frac{[\alpha-j(\omega-\omega_0)]^2}{4(p-j\beta)}} \sqrt{\frac{1}{p-j\beta}} \ .
\end{aligned}
\tag{21}
$$

Denoting

$$
\begin{aligned}
\xi^2 &= \frac{1}{p^2+\beta^2}, \\
\omega' &= (\omega-\omega_0), \\
\theta &= \mathrm{atan}(\beta/p),
\end{aligned}
\tag{22}
$$

and using

$$
\frac{1}{p-j\beta} = \frac{1}{p^2+\beta^2}(p+j\beta) = \xi^2(p+j\beta) = e^{\log\xi+j\theta},
\tag{23}
$$

we get

$$
\begin{aligned}
X(\omega) &= \sqrt{\pi}Ae^{j\phi}\exp\left(\frac{1}{2}(\log\xi+j\theta)\right)\exp\left(\frac{1}{4}\xi^2(p+j\beta)(\alpha-j\omega')^2\right) \\
&= \sqrt{\pi}A\exp\left(j\phi+\frac{1}{2}(\log\xi+j\theta)+\frac{1}{4}\xi^2(p+j\beta)(\alpha^2-\omega'^2-j2\alpha\omega')\right) \\
&= \sqrt{\pi}A\exp\left(j\phi+\frac{1}{2}(\log\xi+j\theta)+\frac{1}{4}\xi^2(p\alpha^2+j\beta\alpha^2-p\omega'^2-j\beta\omega'^2-j2p\alpha\omega'+2\alpha\beta\omega')\right).
\end{aligned}
\tag{24}
$$

For $\beta = 0$, we get

$$
\begin{aligned}
X(\omega) &= \sqrt{\pi}A\exp\left(j\phi+\frac{1}{2}\log\xi+\frac{1}{4}\xi^2(p\alpha^2-p\omega'^2-j2p\alpha\omega')\right) \\
&= \sqrt{\frac{\pi}{p}}A\exp\left(\frac{\alpha^2}{4p}-\frac{1}{4p}(\omega-\omega_0)^2\right)\exp\left(j\left(\phi-\frac{\alpha}{2p}(\omega-\omega_0)\right)\right).
\end{aligned}
\tag{25}
$$

For $\beta \neq 0$, we get

$$
\begin{aligned}
X(\omega) &= \sqrt{\pi}A\exp\left(j\phi+\frac{1}{2}(\log\xi+j\theta)+\frac{1}{4}\xi^2(p\alpha^2+j\beta\alpha^2-p[\omega'^2-\frac{2\alpha\beta}{p}\omega']-j\beta[\omega'^2+\frac{2p\alpha}{\beta}\omega'])\right) \\
&= \sqrt{\pi}A\exp\left(j\phi+\frac{1}{2}(\log\xi+j\theta)+\frac{1}{4}\xi^2(p\alpha^2+j\beta\alpha^2-p[\omega'-\frac{\alpha\beta}{p}]^2+\frac{\alpha^2\beta^2}{p}-j\beta[\omega'+\frac{p\alpha}{\beta}]^2+j\frac{p^2\alpha^2}{\beta})\right) \\
&= \sqrt{\pi}A\exp\left(j\phi+\frac{1}{2}(\log\xi+j\theta)+\frac{1}{4}\xi^2(\frac{\alpha^2}{p}(p^2+\beta^2)+j\frac{\alpha^2}{\beta}(\beta^2+p^2)-p[\omega'-\frac{\alpha\beta}{p}]^2-j\beta[\omega'+\frac{p\alpha}{\beta}]^2)\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \sqrt{\pi}A\exp\left(j\phi + \frac{1}{2}(\log\xi + j\theta) + \frac{\alpha^2}{4p} + j\frac{\alpha^2}{4\beta} + \frac{1}{4}\xi^2(-p[\omega' - \frac{\alpha\beta}{p}]^2 - j\beta[\omega' + \frac{p\alpha}{\beta}]^2)\right) \\
&= \sqrt{\pi}A\exp\left(\frac{1}{2}\log\xi + \frac{\alpha^2}{4p} - \frac{p}{4}\xi^2[\omega' - \frac{\alpha\beta}{p}]^2\right)\exp\left(j(\phi + \frac{\theta}{2} + \frac{\alpha^2}{4\beta} - \frac{\beta}{4}\xi^2[\omega' + \frac{p\alpha}{\beta}]^2)\right) \\
&= \sqrt{\pi}A\exp\left(\frac{\alpha^2}{4p} - \frac{1}{4}\log(p^2+\beta^2) - \frac{p}{4(p^2+\beta^2)}[\omega - \omega_0 - \frac{\alpha\beta}{p}]^2\right) \\
&\qquad \exp\left(j(\phi + \frac{\mathrm{atan}(\beta/p)}{2} + \frac{\alpha^2}{4\beta} - \frac{\beta}{4(p^2+\beta^2)}[\omega - \omega_0 + \frac{p\alpha}{\beta}]^2)\right).
\end{aligned} \tag{26}
$$

C Fitting Gaussian to a window in Blackman-Harris Family

There exist several methods to approximate a window in the Blackman-Harris family by a Gaussian. Here, we approximate it by fitting the curvature of the peaks of the spectra.

The Fourier transform of the Gaussian window in Eq.(2) can be written as

$$
W_g(\omega) = e^{-\frac{1}{2}\sigma^2\omega^2}. \tag{27}
$$

Note that the peak value is normalize to 1. The second derivative at the peak is

$$
W_g''(0) = -\sigma^2. \tag{28}
$$

A window in the Blackman-Harris family can be written as

$$
w_b(t) = \sum_{l=-L}^{L} q_l e^{j2\pi t/T_w} = \sum_{l=-L}^{L} q_l e^{j\pi t/\gamma}, \tag{29}
$$

where T_w denotes the window length and $\gamma = T_w/2$. Actual values of q_ls for some often-used windows are shown in Table 6.

The Fourier transform of the window in Eq.(29) can be calculated as

$$
W_b(\omega) = \frac{1}{q_0}\sum_{l=-L}^{L}(-1)^l q_l \frac{\sin(\gamma\omega)}{\gamma\omega - l\pi}. \tag{30}
$$

Note that the peak value is normalize to 1. The first derivative is

$$
W_b'(\omega) = \frac{1}{q_0}\sum_{l=-L}^{L}(-1)^l q_l \frac{(\gamma\omega - l\pi)\gamma\cos(\gamma\omega) - \gamma\sin(\gamma\omega)}{(\gamma\omega - l\pi)^2}. \tag{31}
$$

The second derivative is

$$
W_b''(\omega) = \frac{1}{q_0}\sum_{l=-L}^{L}(-1)^l q_l \frac{\gamma^2[2 - (\gamma\omega - l\pi)^2]\sin(\gamma\omega) - 2\gamma^2(\gamma\omega - l\pi)\cos(\gamma\omega)}{(\gamma\omega - l\pi)^3}. \tag{32}
$$

Table 6: Coefficients of windows in Blackman-Harris family.

Window	L	q_0	$q_{\pm 1}$	$q_{\pm 2}$
Rect	0	1.0	—	—
Hann	1	0.5	0.25	—
Hamming	1	0.54	0.23	—
Blackman	2	0.42	0.25	0.04

For $\omega = 0$, by using $\sin(x) \approx x - x^3/6$ and $\cos(x) \approx 1 - x^2/2$, we get

$$W_b''(0) = -\gamma^2 \left[\frac{1}{3} + \frac{4}{q_0 \pi^2} \sum_{l=1}^{L} (-1)^l q_l \frac{1}{l^2} \right] = -T_w^2 \left[\frac{1}{12} + \frac{1}{q_0 \pi^2} \sum_{l=1}^{L} \frac{(-1)^l q_l}{l^2} \right]. \tag{33}$$

Note that the positive and negative terms are combined and the range of the summation has been changed. From Eq.(28) and Eq.(33), we get the B-H window equivalent Gaussian as

$$\sigma = T_w \sqrt{\frac{1}{12} + \frac{1}{q_0 \pi^2} \sum_{l=1}^{L} \frac{(-1)^l q_l}{l^2}}. \tag{34}$$

For example, a Hann-window-equivalent Gaussian can be calculated as

$$\sigma = T_w \sqrt{\frac{1}{12} - \frac{1}{2\pi^2}} \approx 0.180756028\ T_w. \tag{35}$$

The values for the other types of windows are shown in Table 2.