Frequency Estimation Enhancement of Multiple Complex Exponentials

MARK LEE, DAVID CLARKE
Department of Engineering Science
University of Oxford
South Parks Road, Oxford, UK
UK

Abstract: - Frequency estimation of complex exponentials using the interpolated Fourier spectrum is analyzed. Data is categorized into two types: phase relationship between components – random/fixed phase difference; and availability of data – single-shot/continuous. When phase difference between exponentials is random the bias component, from interaction of secondary sidelobes, itself becomes random. This translation of bias to variance has considerable implications when choosing effective window functions in conjunction with averaging procedures to minimize mean square error. Optimal methods are identified for each type of data.

Key-Words: - Frequency estimation, Fourier, Interpolated spectrum, Complex exponentials, Window functions

1 Introduction

Frequency estimation of multiple complex exponentials continues to receive considerable attention in the signal processing literature. Non-parametric, Fourier based techniques have a long history and are commonly used in a wide range of applications.

Estimation of a single complex exponential is performed adequately by non-parametric techniques under the assumption of an integer number of cycles within the data set or a large number of periods. These problems are overcome by the use of interpolation of the spectral values surrounding the frequency of interest, and the optimal performance, given by the Cramér-Rao bound (CRB), is achieved. However, the presence of additional signals causes problems [1]. Sidelobes, from spectral leakage of secondary components, interfere with the location of the primary main-lobe; this is associated with a frequency estimation bias [3], the magnitude of which is a function of frequency separation. Judicious use of window functions reduces the influence of sidelobes at the expense of main-lobe width, resulting in the common bias-variance trade-off. However, resolution is severely affected.

Resolution is a source of advantage in parametric techniques [2]. Use of a pre-defined model achieves good estimation accuracy, whilst relaxation of the non-parametric technique assumptions gives superior distinction between two frequencies. However, computational burden is significantly higher for high-order model-based techniques.

A range of Fourier based procedures are available for many applications. The eventual choice of method depends upon such parameters as SNR, frequency separation and data length. This paper examines the performance of methods based on the discrete Fourier transform (DFT), shows how bias and variance of frequency estimation are linked and establishes the optimal method under a range of conditions with judicious use of averaging and windowing. It is shown how the optimal method depends upon signal phase properties and the availability of data.

2 Sources of estimation error

The time domain description of multiple complex exponentials is:

$$y(t) = \sum_{k=1}^{K} A_k e^{j(2\pi f_k t + \varphi_k)} + w(t), \quad 0 < t < t_{\mathrm{m}} \quad (1)$$

where A_k, f_k and φ_k denote amplitude, frequency and phase of the k^{th} signal, w(t) is complex white Gaussian additive noise of variance σ^2 and t_{m} is signal length.

2.1 Maximum likelihood estimation

It is well known that Fourier methods give the maximum likelihood frequency estimate [4] in the case of zero bias and zero leakage.

For K = 1, the first of these conditions is true when the DFT is used with a complex exponential. In the case of a single sinusoid the negative spectrum introduces a bias to the positive spectrum and frequency estimate, but use of complex exponentials allows significant mathematical simplification and zero bias [5]. If K > 1 bias is

introduced by secondary sidelobes; this case is dealt with in section 2.2.

The second condition becomes redundant, if K = 1, through the use of frequency domain interpolation of the spectral values around the frequency of interest. Many such techniques are available [6], [7] and relative performance difference between methods is often minimal. The mean square error then reduces to the estimation variance and is given by the Cramér-Rao bound [7]:

$$\sigma_f^2 = \frac{6 f_s^2}{(2\pi)^2 \rho N(N^2-1)}, \qquad (2)$$

where f_s is the sampling frequency, N the data length and ρ the SNR, A^2/σ^2. This lower bound is achieved for all SNR > SNR_c, the cut-off point at which it becomes impossible to reliably distinguish signal peaks from noise spikes in the spectrum [8].

However, deviation from the lower bound occurs when a non-rectangular window function is used on the time domain data. Typically, a Bartlett window will increase estimation standard deviation by 1.25 and a Parzen window by a factor 2. Fig. 1 shows the CRB is achieved using the Boxcar window for SNR > - 5 dB; use of other functions gives lower performance.

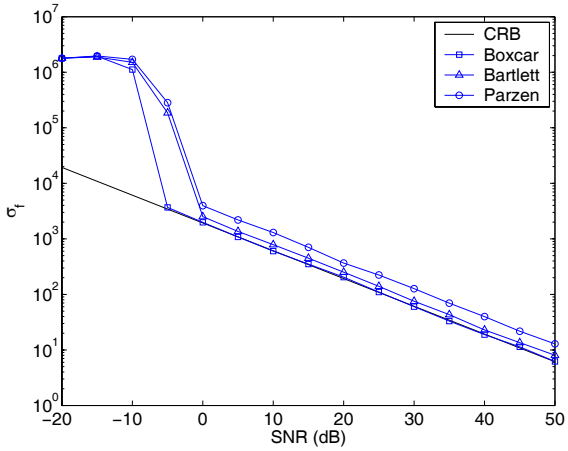

Fig. 1 – Deviation from the Cramér-Rao bound occurs with windowing.

2.2 Bias effect of additional complex signals

The spectrum of a time-limited complex exponential is given by the frequency shifted and scaled $\sin(f)/f$ function. Consider the situation when only one frequency is desired and the additional, spurious signals can be treated as unwanted narrowband noise. To establish the bias effect, of a spurious or secondary component of frequency f_2, on the frequency of interest, f_1, and how the effect varies with window function [9], it is useful to consider the simple cases of Boxcar and Bartlett functions, the spectra of which are:

$$Y_2^{box}(f) = \frac{A_2 \sin[\pi(f-f_2)\tau]}{\pi(f-f_2)}, \qquad (3)$$

and

$$Y_2^{bar}(f) = \frac{A_2 \sin^2[\pi(f-f_2)\tau/2]}{\pi^2(f-f_2)\ \tau/2}, \qquad (4)$$

where τ is the data length in the first case and half the data length in the second.

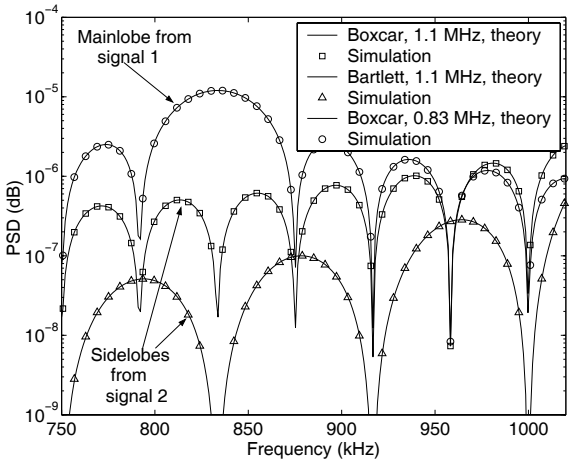

Fig. 2 – Bartlett sidelobes have lower magnitude and larger period.

The sidelobes of signal 2 introduce a bias to the spectral peak of signal 1; the magnitude of the frequency shift depends on the amplitude of the sidelobes at f_1, found from eqns. 3 & 4.

This leads to the familiar trade-off between bias and variance. A noisy signal, with well separated frequencies, will yield lower errors with a rectangular window. However, closely spaced exponentials, with high SNR, require the use of more complex time domain functions. Fig. 2 illustrates the impact of spurious sidelobes on the main-lobe with $f_2/f_1 = 1.3$, $A_2/A_1 = 1$ and $t = 20\mu s$. The impact of additional signals is reduced when the data is windowed with a Bartlett function. The combined spectrum is not shown for reasons of clarity and although the resultant bias is small compared to sidelobe period the error is significant for many applications.

3 Minimizing estimation error

In order to optimize the frequency estimation procedure it is useful to measure the estimation error over a large number of estimations, *N*. A suitable measure is the root mean square error:

$$RMSE = \sqrt{B^2 + \sigma^2}, \tag{5}$$

where the bias, B, is

$$B = \frac{1}{N}\sum_{n=1}^{N} f_{\mathrm{p}}^{n} - f, \tag{6}$$

and the variance is given by

$$\sigma^2 = \frac{1}{n-1}\sum_{n=1}^{N} (f_{\mathrm{p}}^{n} - f)^2, \tag{7}$$

where f_{p}^{n} is the n^{th} frequency estimate.

3.1 Introduction of random relative phase

The phase coherency between spectrum 1 and sidelobes 2 is found from the ratio of frequency separation, $\Delta f_{21} = f_2 - f_1$, and the sidelobe period, $2/\tau$:

$$\theta_1 = \pi\tau\,\Delta f_{21}. \tag{8}$$

Note that in Fig. 2 the sidelobes are out of phase with the main-lobe, hence the slopes are maximized and the effect on main-lobe position is severe. However, eqns. 3 & 4 contain no phase information, so the effect of relative phase between the signals, $\theta_2 = \varphi_2 - \varphi_1$, must be included in the combined expression:

$$\begin{aligned} Y(f) &= \frac{\sin[\pi(f - f_1)\tau]}{\pi(f - f_1)} \cdots \\ &+ \cos(\theta_1 + \theta_2) \times \frac{\sin[\pi(f - f_2)\tau]}{\pi(f - f_2)}, \end{aligned} \tag{9}$$

which is valid for frequency values surrounding the main-lobe and when spurious frequencies are higher. It is useful to interpret this equation. Spectral phase, θ_1, determines the potential magnitude of the combined spectrum due to the interaction of spurious sidelobes and relative phase, θ_2, determines the actual magnitude, which varies as cos (θ_2).

Fig. 3 shows the impact of spurious sidelobes on the frequency of interest. Frequency separation and relative phase is varied under zero noise conditions. From this graph the importance of phase is obvious. As phase moves through one period, the bias moves through its full range. In fact, return to eqn. 1 and impose a uniform distribution on φ_k, such that $\Delta\varphi = \varphi_2 - \varphi_1$ is uniformly distributed: $0 < \Delta\varphi < 2\pi$, and the bias at a given frequency separation becomes random.

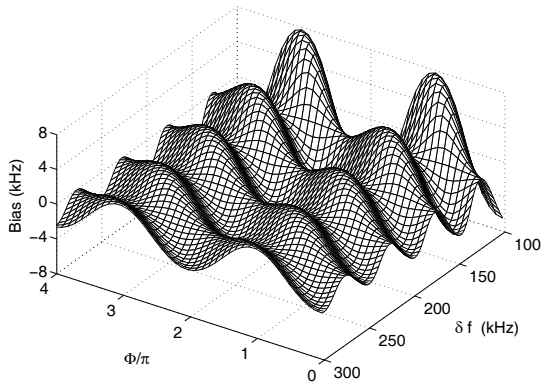

Fig. 3 – Bias changes periodically with Δf and θ, with bias being higher at small frequency separations.

Further, impose random signal frequencies such that each frequency deviates through a range larger than the sidelobe period, $2/\tau$, then the zero noise bias can be translated into standard deviation.

3.2 Theory linking bias with variance

From the definition of variance, eqn. 7, it is possible to predict random-phase frequency estimation standard deviation arising from constant-phase bias. The bias values represent an envelope of possible estimation errors over many samples, hence the introduction of random phase will change the estimation error according to cos (θ_2). The standard deviation is found by multiplying the bias envelope by the root mean square value of $\cos^2(\theta_2)$:

$$\begin{aligned} &\sqrt{\frac{1}{\pi}\int_0^{\pi} \cos^2(\theta)\, d\theta} \\ &= \left(\frac{1}{\pi}\left[\frac{\theta}{2} - \frac{1}{4}\sin(2\theta)\right]_0^{\pi}\right)^{1/2} = \frac{1}{\sqrt{2}}. \end{aligned} \tag{10}$$

Fig. 4 shows the bias-induced variance is linked to zero-phase bias as expected.

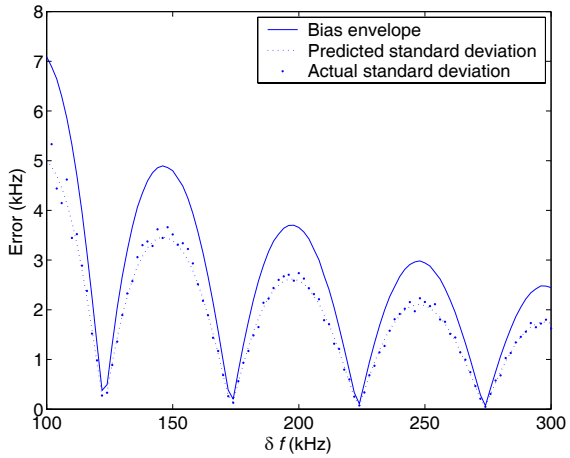

Fig. 4 – Theoretical values of random phase variance are confirmed with simulation.

The bias-induced variance remains a function of frequency separation and relative phase, but the magnitude of this effect is smaller with non-rectangular functions. Fig. 5 compares Boxcar, Bartlett and Parzen windows using theory and simulation. The optimal choice of function depends on frequency separation. Closely spaced frequencies give lower bias-induced variance with a rectangular window whereas estimation of widely spaced frequencies is improved by the Parzen function. The periodic nature of error with Δf arises from secondary sidelobes. The Boxcar sidelobes have double the frequency of Bartlett sidelobes, and the Parzen is shown to have a broader main-lobe.

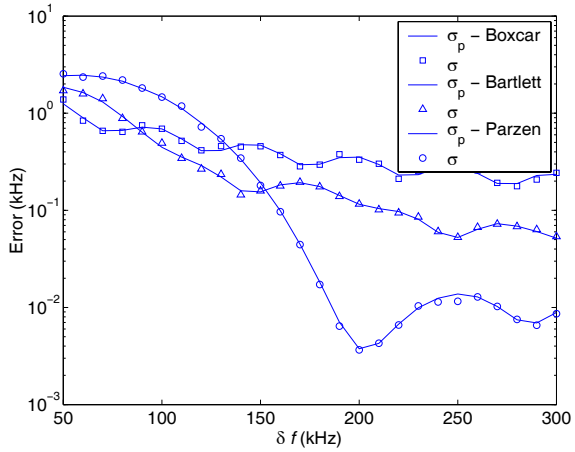

Fig. 5 – Choice of window function depends on frequency separation.

3.3 Improving RMSE through averaging

All frequency estimation applications can be broadly categorized as belonging to one of two groups: 1) limited data length available requiring a one-shot approach, and 2) continuous data stream enabling averaging.

Bias and variance reduction must be performed by choosing the optimal estimation procedure. If the application falls within category 1 then for an unbiased estimate, the variance performance must lie close to the CRB.

However, for applications falling within category 2, averaging can be used to reduce the estimation variance according to eqn. 11.

$$RMSE = \sqrt{B^2 + \frac{\sigma^2}{n}}, \qquad (11)$$

where n is the number of frequency estimates used. Variance tends to zero hence the RMSE is limited to a floor set by the bias. This occurs when the relative frequency and phase of complex exponentials is fixed and bias does not vary between estimates.

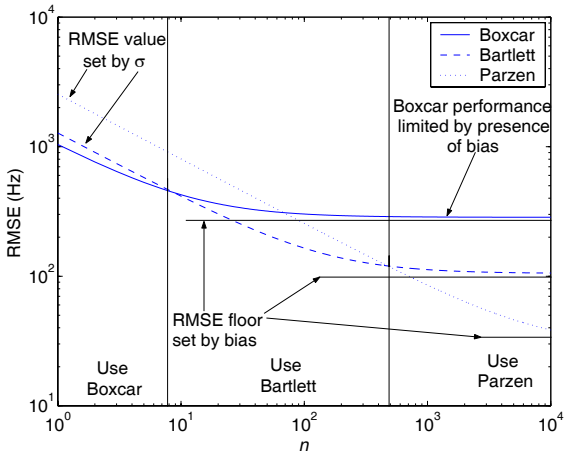

Fig. 6 - Zero relative phase estimation.

Fig. 6 illustrates that averaging reduces variance but not bias. Here, f_2 / f_1 is uniformly distributed between 1.2 and 1.4, $A_2 / A_1 = 10$ and $t = 20\mu s$. The bias sets a lower bound on the RMSE, $RMSE \geq B$, and is approached as n increases:

$$\lim_{n \to \infty} RMSE = B, \qquad (12)$$

but, if $\sigma > B$, $RMSE$ approaches σ as n approaches zero:

$$\lim_{n \to 0} RMSE = \sigma. \qquad (13)$$

The choice of window function then simply reduces to determining the urgency of estimation and working with the trade-off between *RMSE* and *n*.

Sections 3.1 and 3.2 developed theory behind translation of bias to variance. When relative phase difference is itself a random variable, eqn. 11 changes to:

$$(RMSE)^2 = (B^2 + \sigma^2)/\sqrt{n}, \qquad (14)$$

hence RMSE, and not just variance, can be reduced to an arbitrarily small value.

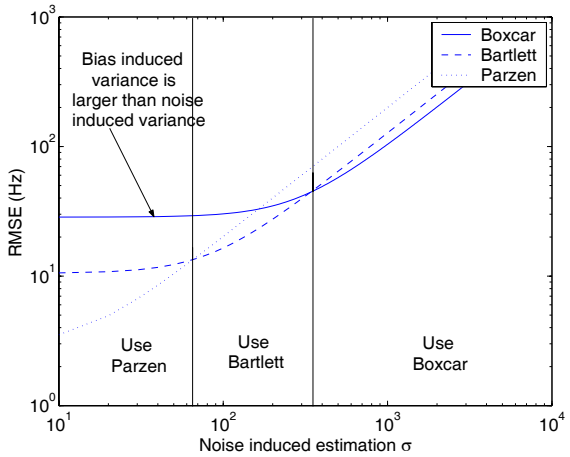

Fig. 7 – The criteria for window function choice is changed when relative phase is random.

Fig. 7 shows RMSE variation with noise induced variance when the parameter values are the same as for Fig. 6. At high SNR the impact of bias-induced variance is minimal and a Boxcar function offers optimal performance. At low SNR, bias-induced variance dominates and a Parzen function is preferred.

4 Conclusion

Many frequency estimation problems involve a single frequency corrupted by Gaussian noise and spurious frequencies. The choice of estimation procedure is a function of data parameters, particularly the rate of variation of relative phase between complex exponentials. With fixed phase, frequency estimation performance quickly reaches a lower bound set by the sidelobe-induced bias as averaging is performed. If the application enables further data collection, it becomes optimal to use a powerful bias-reduction window function, such as Parzen, to reduce RMSE. Averaging compensates for the increase in noise dependent variance.

When components have random phase, bias is translated to variance, hence averaging continues to reduce RMSE. The choice of window function then reduces to an estimate of the SNR. When data quality is high, noise-induced variance is small, bias-induced variance dominates and a suitable window function is Parzen. However, if SNR is poor, narrow main-lobe width is required and a Boxcar function is optimal.

This work has emphasized the importance of understanding the available data set. Optimal methods have been suggested for data sets classified by 1) feasibility of averaging and 2) nature of phase relationship.

References:

[1] Manolakis, D.G. *et al*, *Statistical and Adaptive Signal Processing*, McGraw-Hill, 2000.

[2] Jenkins, G.M. *et al*, *Spectral Analysis and its Applications,* Holden-Day, 1968.

[3] Proakis, J.G. *et al*, *Advanced Digital Signal Processing*, Macmillan, 1992.

[4] Rife, D.C. *et al*, Single-tone parameter estimation from discrete-time observations, *IEEE Transactions on Information Theory*, Vol.20, No.5, 1974, pp. 591-598.

[5] Sharma, P.L. *et al*, Auto-regressive spectral estimation of noisy sinusoids, *Electrical Engineering Department, University of Akron, Ohio,* 1982, pp. 1028-1041.

[6] McIntyre, M.C. *et al*, A new fine-frequency estimation algorithm based on a parabolic regression, *Defense Research Establishment, Atlantic, Nova Scotia*, 1992, pp. 541-544.

[7] Schoukens, J., *et al*, The interpolated fast Fourier transform: A comparative study, *IEEE Transactions on Instrumentation and Measurement*, Vol.41, No.2, 1992, pp. 226-232.

[8] Aboutanios, E., A modified dichotomous search frequency estimator, *IEEE Signal Processing Letters,* Vol.11, No.2, 2004, pp. 186-188.

[9] So, H.C. *et al*, Comparison of various periodograms for single tone detection and frequency estimation, *IEEE International Symposium on Circuits and Systems,* 1997, pp. 2529-2532.

[10] Ofelli, C., The influence of windowing on the accuracy of multifrequency signal parameter estimation, *IEEE Transactions on Instrumentation and Measurement,*Vol.42, No.2, 1992, pp. 256-261.