OPTIMIZATION OF WEIGHTING FACTORS IN THE PEAK MATCHED MULTIPLE WINDOW METHOD

Maria Hansson

Dept. of Applied Electronics, Signal Processing Group
Lund University, Sweden
mh@tde.lth.se

ABSTRACT

Peak matched multiple windows are found as the Karhunen-Loève basis functions of a predefined peaked spectrum. With a penalty function, an optimization procedure can be constrained with resulting control of sidelobes. Weighting factors, included in the averaging of the periodograms, can be designed to fulfill certain constraints. Desirable properties are low variance and small bias. This paper presents a discussion of minimization of variance at the peak compared with the optimization that also include the neighbourhood of the peak.

1. INTRODUCTION

Spectrum estimation methods can be divided into two groups, parametric and non-parametric. The usual argument for using parametric methods, e.g., AR-estimates, is that the frequency resolution is much higher than for the non-parametric. The algorithms for computing AR-estimates are, however, often sensitive to model errors and computationally demanding. With the exploitation of the FFT-algorithm, the non-parametric methods and especially the windowed periodogram, are less complicated when it comes to calculations.

The advantage of the multiple window methods is reduced variance without decreased resolution or use of more data compared to the use of single window methods, [1,2,3]. Multiple window methods are thereby especially useful for non-stationary data.

2. PROBLEM FORMULATION

The power density spectrum, $S_x(f)$, of the zero mean complex valued stationary random process, $x(n)$, is given. The spectrum is assumed to have a peak located at $f = 0$.

With use of the N samples $\mathbf{x} = [x(0) \ \ldots \ x(N-1)]^T$, where the superscript T denotes the vector transposition, the spectrum should be estimated by

$$\hat{S}_x(f) = \sum_{i=0}^{K-1} \alpha_i \hat{S}_i(f) \tag{1}$$

where

$$\hat{S}_i(f) = | \sum_{n=0}^{N-1} x(n) h_i(n) e^{-j2\pi f n} |^2 \tag{2}$$

is the estimated subspectrum number i. Equation (2) is a periodogram obtained by using the data window $\mathbf{h}_i = [h_i(0) \ \ldots \ h_i(N-1)]^T$. A weighted sum of K subspectra provides the estimate in Eq. (1).

2.1. Window Estimation

The windows $\mathbf{h}_i$, $i = 0 \ \ldots \ K-1$, $K > 1$, will be designed to give a low variance and small bias estimate of $S_x(f)$. Reduction of the variance is achieved with uncorrelated subspectrum $\hat{S}_i(f)$ at the peak. Bias is reduced by preventing leakage from distant frequencies. To prevent leakage from regions outside a predetermined interval of width B, the Fourier transforms $H_i(f)$ of $\mathbf{h}_i$, $i = 0 \ \ldots \ K-1$ have to be band-limited to the interval $(-B/2, B/2)$. The mainlobe of $H_i(f)$ should be inside this band and the sidelobes should be as low as possible.

The power of the output signal within the frequency interval $(-B/2, B/2)$ is

$$P_B = \sum_{i=0}^{K-1} \alpha_i \int_{-B/2}^{B/2} |H_i(f)|^2 S_x(f) df = \sum_{i=0}^{K-1} \alpha_i \mathbf{h}_i^T \mathbf{R}_B \mathbf{h}_i. \tag{3}$$

The $(N \times N)$ Toeplitz covariance matrix $\mathbf{R}_B$ has the elements $r_B(l) = r_x(l) * B \ \mathrm{sinc}(Bl)$, $0 \leq |l| \leq N-1$, where $r_x(l)$ is the covariance function of $x(n)$, $\mathrm{sinc}(u) = \sin(\pi u)/\pi u$ and $*$ denotes the convolution operator. The covariance function $r_x(l)$ is found from a known peaked spectrum $S_x(f)$,

$$S_x(f) = e^{\frac{-2C|f|}{10B\log_{10}(e)}}, \quad |f| \leq 1/2. \tag{4}$$

The optimization is performed subject to the constraint

$$P_Z = \sum_{i=0}^{K-1} \alpha_i \int_{-1/2}^{1/2} |H_i(f)|^2 S_z(f) df = \sum_{i=0}^{K-1} \alpha_i \mathbf{h}_i^T \mathbf{R}_Z \mathbf{h}_i = 1 \tag{5}$$

where $S_z(f)$ has the corresponding Toeplitz covariance matrix $\mathbf{R}_Z$.

The covariance matrix $\mathbf{R}_Z = \mathbf{R}_G$ corresponds to a penalty frequency function $S_g(f) = G$, $B/2 <| f |\leq 1/2$, and $S_g(f) = 1$, $-B/2 \leq f \leq B/2$, which is used to reduce the leakage from the sidelobes.

Equation (5) is a constraint that bounds the output power from the filterbank. The window design problem is formulated as

$$\max_{\mathbf{h}_i, \alpha_i} P_B \text{ subject to } P_Z = 1. \tag{6}$$

The solution with respect to $\mathbf{h}_i$ is given by the generalized eigenvalue problem

$$\mathbf{R}_B\mathbf{q}_i = \lambda_i \mathbf{R}_Z \mathbf{q}_i, \quad i = 0 \dots N-1, \tag{7}$$

where the eigenvalues are ordered in decreasing magnitude, $\lambda_0 \geq \lambda_1 \geq \dots \geq \lambda_{N-1}$. The eigenvectors corresponding to the K largest eigenvalues are used as windows, $\mathbf{h}_i = \mathbf{q}_i = [q_i(0)\ q_i(1)\ \dots\ q_i(N-1)]^T$, $i = 0 \dots K-1$. The solution of Eq. (7) with $G = 0$ dB ($\mathbf{R}_Z = \mathbf{I}$) gives windows which are the Karhunen-Loève eigenvectors of $\mathbf{R}_B$. These windows are called the Peak Matched Multiple Windows (PM MW). With $G = 30$ dB the windows have suppressed sidelobes by 30 dB outside the optimization interval. The windows are named PM_30 MW.

2.2. Optimization of Weighting Factors

The criterion for optimization of the weighting factors $\boldsymbol{\alpha} = [\alpha_0\ \alpha_1\ \dots\ \alpha_{K-1}]^T$ is minimization of

$$\xi = \sum_{n=-M}^{M} \text{Variance } \hat{S}_x(f_n). \tag{8}$$

with $f_n = \frac{n}{2N}$ and

$$\text{Variance } \hat{S}_x(f_n) = \sum_{j=0}^{K-1}\sum_{i=0}^{K-1} \alpha_i \alpha_j \text{cov}[\hat{S}_i(f_n)\hat{S}_j(f_n)]. \tag{9}$$

The parameter M defines the interval around the peak to be included in the criterion. Defining $\mathbf{x}^T\mathbf{\Phi}(f_n)\mathbf{h}_i = A_i$, the covariance can be expressed as

$$\begin{aligned}\text{cov}[\hat{S}_i(f_n)\hat{S}_j(f_n)] &= \text{cov}[A_i^H A_i A_j^H A_j] \\ &= E[A_i A_i^H A_j^H A_j] - E[A_i A_i^H]E[A_j^H A_j] \\ &= E[A_i A_j^H]E[A_i^H A_j] + E[A_i A_j]E[A_i^H A_j^H] \\ = |\mathbf{h}_i^T \mathbf{\Phi}^H(f_n)\mathbf{R}_X\mathbf{\Phi}(f_n)\mathbf{h}_j|^2 &+ |\mathbf{h}_i^T \mathbf{\Phi}(f_n)\mathbf{R}_X\mathbf{\Phi}(f_n)\mathbf{h}_j|^2, \end{aligned} \tag{10}$$

where $\mathbf{\Phi}(f_n) = \text{diag}[1\ e^{-j2\pi f_n}\ \dots\ e^{-j2\pi(N-1)f_n}]$ and $E[\mathbf{x}\mathbf{x}^T] = \mathbf{R}_X$. The windows are the eigenvectors from the solution of Eq. (7) and they are the approximate eigenvectors of $\mathbf{R}_X$ for sufficiently large N.

The window $\mathbf{h}_0$ filters the spectrum at the frequency $f_n = 0$ and the resulting value is λ_0. This value is squared and multiplied by α_0^2. The window $\mathbf{h}_i$ has its center-frequency at $f = \pm\frac{i}{2N}$ and filters the spectrum with the resulting value λ_i which is squared and multiplied by α_i^2. The sum of the result from all window functions are the variance of the estimate $\hat{S}_x(0)$.

For $f_n \neq 0$, the process is repeated for the modulated spectrum $\mathbf{\Phi}^H(f_n)\mathbf{R}_X\mathbf{\Phi}(f_n)$. For $i = 0$ the window $\mathbf{h}_0$ filters the modulated spectrum $S_x(f - f_n)$. The resulting value is λ_n. The windows $\mathbf{h}_i$, $i \neq 0$, filter the modulated spectrum at two different frequencies $\pm\frac{i}{2N}$. Since the spectrum is no longer symmetric around the zero frequency, the resulting powers from the two filters are not equal. The result will be $\frac{1}{2}\lambda_{|n+i|}$ and $\frac{1}{2}\lambda_{|n-i|}$ where the factor $\frac{1}{2}$ is a consequence of the window function which has half its power at each side of the zero frequency. The variance of the estimate $\hat{S}_x(f_n)$ is given by the squares of the the two λ multiplied with α_i^2.

The variance in Eq. (9) simplifies to

$$\text{Variance } \hat{S}_x(f_n) = \sum_{i=0}^{K-1} \alpha_i^2 (\frac{1}{2}\lambda_{|n-i|} + \frac{1}{2}\lambda_{|n+i|})^2. \tag{11}$$

The second term of Eq. (10) is large only for frequencies inside the interval $|\ f\ | < B/2$ or in an equal interval around the Nyquist frequency, and is excluded.

Minimization of Eq. (11) without any constraint will result in a zero valued parameter vector $\boldsymbol{\alpha}$. A suitable constraint is to demand the bias of the peak to be zero which simplifies to

$$\begin{aligned}\text{Bias } \hat{S}_x(0) &= \sum_{i=0}^{K-1} \alpha_i \mathbf{h}_i^T \mathbf{R}_X \mathbf{h}_i - S_x(0) \\ &= \boldsymbol{\alpha}^T \lambda_0 - S_x(0)\end{aligned} \tag{12}$$

where $\boldsymbol{\lambda_0} = [\lambda_0\ \dots\ \lambda_{K-1}]^T$.

The minimization of Eq. (11) is performed subject to the constraint of unbiased peak, Eq. (12). The Lagrangian is

$$\mathcal{L} = \boldsymbol{\alpha}^T\mathbf{\Sigma}\boldsymbol{\alpha} + \gamma(S_x(0) - \boldsymbol{\alpha}^T\boldsymbol{\lambda_0}) \tag{13}$$

where $\mathbf{\Sigma} = \sum_{n=-M}^{M} \text{diag}[\lambda_{|n|}^2\ \ (\frac{1}{2}\lambda_{|n-1|} + \frac{1}{2}\lambda_{|n+1|})^2\ \dots\ (\frac{1}{2}\lambda_{|n-K+1|} + \frac{1}{2}\lambda_{|n+K-1|})^2]$.

The gradient of $\mathcal{L}$ with respect to $\boldsymbol{\alpha}$ is

$$\frac{d\mathcal{L}}{d\boldsymbol{\alpha}} = 2\mathbf{\Sigma}\boldsymbol{\alpha} - \gamma\boldsymbol{\lambda_0} \tag{14}$$

and a solution is found when $\frac{d\mathcal{L}}{d\boldsymbol{\alpha}} = 0$. The solution is a minimum since $\mathbf{\Sigma}$ is positive definite and it is given by

$$\boldsymbol{\alpha} = \frac{S_x(0)\mathbf{\Sigma}^{-1}\boldsymbol{\lambda_0}}{\boldsymbol{\lambda_0}^T\mathbf{\Sigma}^{-1}\boldsymbol{\lambda_0}} \tag{15}$$

where M is a parameter.

3. EVALUATION

To investigate the weighting factors of Eq. (15), a test spectrum is used, defined as

$$S_t(f) = S_x(f - 0.25) + S_x(f + 0.25). \tag{16}$$

With the use of this spectrum instead of $S_x(f)$, an evaluation where the second term of Eq. (10) influences the result at the peak is avoided. Small crosscorrelation terms of $S_x(f - 0.25)$ and $S_x(f + 0.25)$ will, however, be present.

The number of windows is chosen to $K = 8$ and the window length is set to $N = 128$. The optimization interval is defined to $B = 0.08$ and $C = 20$.

Three cases are studied where different values of M are used. In case 1 the minimization is made just at the peak frequency ($M = 0$) which gives $\alpha_i \propto 1/\lambda_i$. Case 2 uses $M = 3$ which minimizes the variance in the interval $|f| < 0.01$. In case 3, which minimizes in the interval $|f| < 0.04$ ($M = 10$), the resulting α_i is proportional to λ_i. The three cases are depicted in Figure 1.

These weighting factors are used to calculate the estimated spectrum $\hat{S}_t(f_n) = \sum_{i=0}^{K-1} \alpha_i\ \mathbf{h}_i^T\mathbf{\Phi}^H(f_n)\mathbf{R}_T\mathbf{\Phi}(f_n)\mathbf{h}_i$,

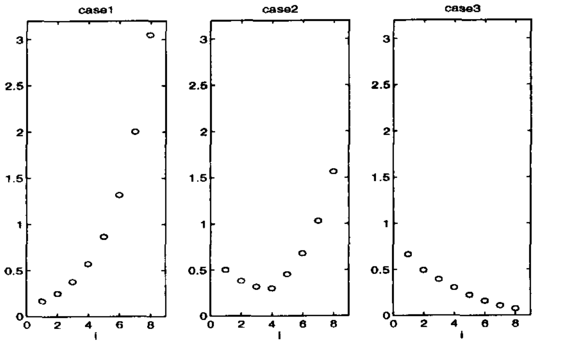

Figure 1: The weighting factors, $\boldsymbol{\alpha}$, resulting from the minimization of variance in different frequency intervals, case 1 $M = 0$, case 2 $M = 3$, case 3 $M = 10$.

where $\mathbf{R}_T$ is the corresponding covariance function of the testspectrum $S_t(f)$. The estimated spectra for the three cases and the two different types of windows are shown in Figure 2. The dotted line is the true peaked spectrum $S_t(f)$ and the dashed line is the estimated spectrum when the PM MW are used. It is seen that the bias at the peak is zero which was the constraint in the optimization of the variance. This is not the case for the PM_30 MW (solid line). The PM MW are better approximations of the Karhunen-Loève basis functions to $\mathbf{R}_X$. The result is compared to the Thomson multiple window method. The Thomson windows are derived from Eq. (7) with $r_x(l) = r_z(l) = \delta(l)$ for B=0.08, $\alpha_i = 1/8$, $i = 0 \ldots 7$ (Thomson8 MW) and B=0.04, $\alpha_i = 1/4$, $i = 0 \ldots 3$ (Thomson4 MW). The bias of the Thomson method is large at the peak and the resulting estimate is smoothed the spectrum.

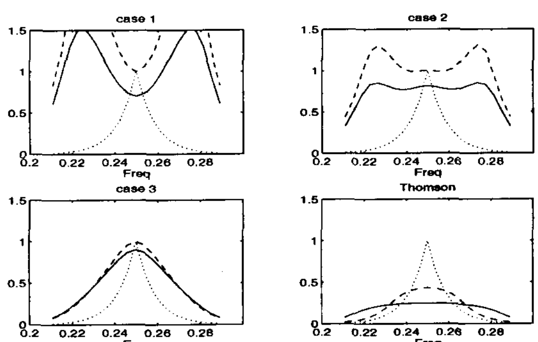

Figure 2: The estimated spectra for the three cases and for the Thomson method, case 1 $M = 0$, case 2 $M = 3$, case 3 $M = 10$ (solid line=PM_30 MW, dashed line=PM MW, dotted line=$S_t(f)$). Thomson method (solid line=Thomson8 MW, dashed line=Thomson4 MW).

The variance is calculated with use of Eqs. (9,10), where $S_x(f_n)$ and $\mathbf{R}_X$ are exchanged for $S_t(f_n)$ and $\mathbf{R}_T$. To obtain a comparable measure for different frequencies, the variance is normalized with the squared expected value of the estimate. The variance of the PM MW (dashed line) is lower than for the PM_30 MW (solid line). This is compensated with the better sidelobe suppression of the PM_30 MW. For $M = 0$, the variance at the peak is small (1/8 for the PM MW). This value is the optimal variance for $K = 8$ windows. Outside the peak, however, the variance is increased. In case 2, the variance is small in the optimization interval (marked with dotted lines) and increased outside. In case 3, the weighting factors cooperate to minimize the variance in the total interval $|f| < 0.04$. The conclusion is that the weighting factors of case 3 is preferrable. The variance of the Thomson8 MW method is smaller at the peak but larger outside than the peak match technique.

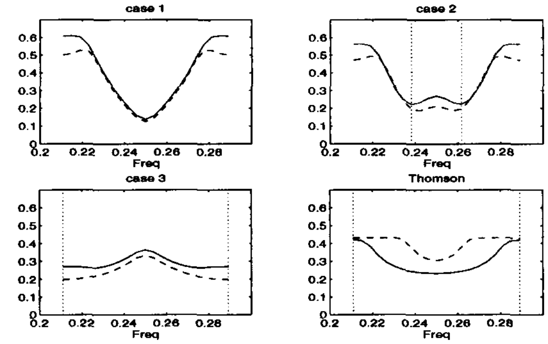

Figure 3: The estimated normalized variance for the three cases and for the Thomson method, case 1 $M = 0$, case 2 $M = 3$, case 3 $M = 10$ (solid line=PM_30 MW, dashed line=PM MW). Thomson method (solid line=Thomson8 MW, dashed line=Thomson4 MW).

3.1. Normalization for unbiased white noise spectrum

When the windows and the weighting factors optimized in Section 2 are used in the estimator of Eq. (1), the property of bias are considered for a peaked spectrum. The bias for smooth spectra are not studied, even if it is an important property for a spectrum estimator since smooth parts often are included in a spectrum. Therefore, an unbiased estimate for the spectrum of white noise is desirable. This property is attained if

$$\begin{aligned} \text{Bias } \hat{S}_w(f) &= \sum_{i=0}^{K-1} \alpha_i \mathbf{h}_i^T \mathbf{\Phi}^H(f) \mathbf{R}_W \mathbf{\Phi}(f) \mathbf{h}_i - S_w(f) \\ &= \sum_{i=0}^{K-1} \sigma_w^2 \alpha_i \mathbf{h}_i^T \mathbf{h}_i - S_w(f) = 0 \qquad (17) \end{aligned}$$

where $\mathbf{R}_W = \sigma_w^2 \mathbf{I}$ is the covariance matrix for the white noise process. The spectrum $S_w(f) = \sigma_w^2$ for all frequencies. For the PM MW, $\mathbf{h}_i^T \mathbf{h}_i = 1$, and Eq. (17) simplifies to $\sum_{i=0}^{K-1} \alpha_i = 1$ as a constraint for unbiased white noise spectrum. The PM_30 MW fulfill $\mathbf{h}_i^T \mathbf{R}_G \mathbf{h}_i = 1$, but $\mathbf{h}_i^T \mathbf{h}_i \neq 1$ in the general case. To achieve an unbiased spectrum the windows are normalized, $\frac{\mathbf{h}_i}{\sqrt{\mathbf{h}_i^T \mathbf{h}_i}}$. Unbiased white noise spectrum is now given as $\boldsymbol{\alpha}_n = \frac{\boldsymbol{\alpha}}{\boldsymbol{\alpha}^T \mathbf{1}}$ where $\mathbf{1}$ is a column vector with ones in all positions.

The variance for the peaked spectrum does not alter for the PM MW as the spectrum is just scaled with a factor. For the PM_30 MW, the normalization of windows will give a small change in variance. The average squared bias has decreased. The constraint of unbiased peak is, however, no longer valid.

For white noise spectrum the variance (0.219) of the PM MW method, case 3, is small as the windows are orthogonal. The PM_30 MW, case 3, do not have this property which will give a slightly higher variance (0.220). A comparison between the methods shows that the Thomson8 MW, as expected, are preferable (0.125), but that the PM MW and PM_30 MW methods are comparable to the Thomson4 MW (0.25).

4. NUMERICAL EXAMPLES

Simulated ARMA-process data are used to test the performance of the proposed methods with the normalized weighting factors of case 3 in Section 2. The simulation examples aim at illustrating two different types of spectrum dynamics; spectrum with peaks and notches and low-frequency dominant spectrum with large dynamics. The alteration in normalized mean square error, bias and variance with changing frequency for the different spectrum examples will be illustrated.

Example 1

The ARMA spectrum with two poles, $p_{1,2} = 0.95e^{\pm j2\pi 0.1}$ and two zeros $z_{1,2} = 0.95e^{\pm j2\pi 0.3}$ is studied. The results are depicted in Figure 4 with PM _30 MW as the solid line and PM MW as the dashed line. The results from the Thomson8 MW (dash-dotted line) and the single Hanning window (dotted line) are also shown.

The bias is smallest for the single Hanning window, but the results for the PM_30 MW are better than for the Thomson8 MW for all frequencies. Both the peak matched methods are comparable in estimating the peak. The notch is, however, estimated with more success with use of the PM_30 MW as the leakage properties are better. The variance result of the PM_30 MW is about the same as the Thomson8 MW but taking the better bias property into account, the PM_30 MW is the method to use for spectrum with peaks and notches as the normalized mean square error for the PM_30 MW method gives the smallest value for almost all frequencies.

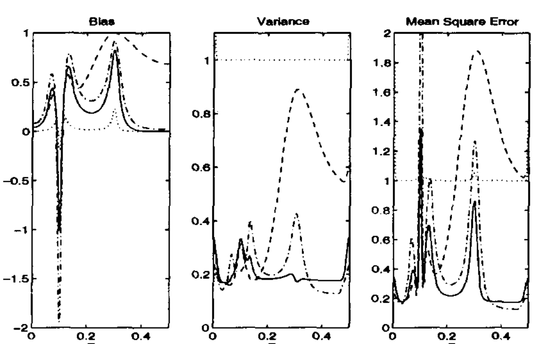

Figure 4: The normalized bias; variance; mean square error, (solid line=PM_30 MW, dashed line=PM MW, dash-dotted line=Thomson8 MW, dotted line=Hanning window).

Example 2

In the following example a random process is generated by filtering white noise through a third-order lowpass Butterworth-filter with cutoff frequency $f = 0.1$. Figure 5 indicates that the proposed methods give reliable results also for flat spectra, although the Thomson8 MW method shows smaller variations in the passband than the peak matched technique. The single Hanning window is able to estimate spectrum variations of more than 60 dB for the third order process. An advantage for the Hanning window is the decay of the sidelobes, (18 dB/octave), which contributes to the ability to track a steep slope. The multiple window methods all have very little decay of the sidelobes. The first prolate function window of the Thomson8 MW can be used as a single window to estimate large dynamics as the sidelobes are suppressed 140 dB. The following windows limit the capability to about -40 dB. For the PM MW and PM_30 MW methods, all the windows have about the same sidelobe suppression. The PM_30 MW reach down to -50 dB which is lower than the Thomson8 MW.

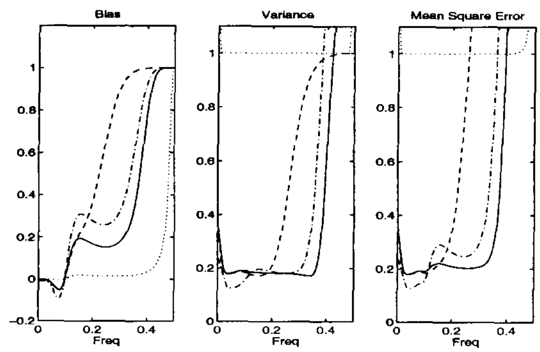

Figure 5: The normalized bias; variance; mean square error, (solid line=PM_30 MW, dashed line=PM MW, dash-dotted line=Thomson8 MW, dotted line=Hanning window).

5. CONCLUSIONS

Windowed periodograms are combined with weighting factors to the multiple window spectrum estimate. Minimization of variance at the peak with respect to the weighting factors give an estimate with optimal variance at the peak but unacceptable result in the neighbourhood.

One solution is to minimize the variance at the peak as well as in the neighbourhood. These weighting factors possess good qualities as the resulting spectrum estimate has low variance and small bias in the predefined frequency range around the peak.

REFERENCES

[1] M. Hansson, T. Gänsler and G. Salomonsson, "A Multiple Window Method for Estimation of a Peaked Spectrum",*Proc. of the ICASSP-95*, pp 1617-1620, May 1995.

[2] D.J. Thomson, "Spectrum Estimation and Harmonic Analysis",*Proc. of the IEEE*, vol. 70, pp 1055-1096, Sept 1982.

[3] C. T. Mullis and L. L. Scharf, "Quadratic Estimators of the Power Spectrum", In S. Haykin, editor,*Advances in Spectrum Analysis and Array Processing*, chapter 1, pp 1-57, Prentice Hall, 1991.