TABLE I
COMPARISON OF THE CC OF THE PROPOSED ALGORITHM WITH THOSE REQUIRED BY SOME WELL-KNOWN ALGORITHMS

Algorithm	N	mult/sample	add/sample
Rader-Brenner algorithm (FFT)	2048	9.50	38.75
Agarwal-Cooley nesting algorithm	2520	11.58	95.90
Split-nesting algorithm	2520	11.58	75.46
Proposed algorithm ($b = 4$)	2048	3.91	71.72

be one of these polynomial products. It can be computed in two steps as follows.

$$Y_{i_1}(z) = H_i(z)\, X_i(z) \qquad \text{mod } (z^{2^{b+2}} - 1) \tag{20}$$

$$Y_i(z) = Y_{i_1}(z) \qquad \text{mod } (z^{2^{b+1}} + 1). \tag{21}$$

The first step is a convolution of length 2^{b+2}, which is carried out by the FNT. The second one is a reduction requiring 2^{b+1} additions.

Computational Cost of the Proposed Algorithm: The CC associated with the algorithm concerns with the convolutions of length 2^{b+2} and the additions necessary for the various operations of reduction and reconstruction. The convolutions to be computed are those regarding the polynomial products (Remark 5) and those directly appearing in (2). Their number is $2^{t-b} - 3$ and they can be computed by the FNT. We remember that an FNT of length M requires $M \log_2 M$ additions. Consequently, each convolution requires $2^{b+3}(b+2)$ additions and 2^{b+2} multiplications. The overall CC due to the convolutions is

$$\begin{aligned} &2^{b+2}(2^{t-b} - 3) && \text{multiplications} \\ &(b+2)\, 2^{b+3}(2^{t-b} - 3) && \text{additions.} \end{aligned} \tag{22}$$

All the other operations (CRT, reduction, etc.) only imply additions; their number has been considered in detail in the previous remarks, i.e.,

$$\begin{aligned} \text{CRT (Remark 2):} \quad & 2^{t+2} - 2^{b+4} \\ \text{Reductions (Remark 4):} \quad & 2^t(4t - 4b - 9) + 2^{b+2} \\ \text{Reductions (Remark 5):} \quad & 2^{t+1} - 2^{b+3}. \end{aligned}$$

Therefore, the total CC amounts to

$$\begin{aligned} &2^{b+2}(2^{t-b} - 3) && \text{multiplications} \\ &2^t(4t + 4b + 13) - 2^{b+2}(6b + 17) && \text{additions} \end{aligned} \tag{23}$$

which corresponds to

$$\begin{aligned} &(4 - 3)\, 2^{b+2-t} && \text{mult/sample} \\ &4t + 4b + 13 - 2^{b+2-t}(6b + 17) && \text{add/sample.} \end{aligned} \tag{24}$$

This CC is much lower than that required by the methods already available in the technical literature, as can be seen from Table I.

It is interesting to note that the method holds up to a length $N \leqslant 2^{2b+3}$. With the usual values of b, N may be very large, and consequently it is not necessary to take into consideration the extension of the algorithm to longer convolutions.

APPENDIX
DERIVATION OF FORMULAS (13) AND (14)

For the sake of brevity, let us limit the derivation to the case $N_1 = N_2$. The circular convolution (11) is obtained by a polynomial transform of length N_1, root z_1^2 and mod $(z_1^{N_1} + 1)$

$$Y_{l_1}^{(c)}(z_1) = \frac{1}{N_1} \sum_{K=0}^{N_1-1} \bar{H}_K^{(c)}(z_1)\, \bar{X}_K^{(c)}(z_1)\, z_1^{-2Kl_1} \quad \text{mod } (z_1^{N_1} + 1);\ l_1 = 0, 1, \cdots, N_1 - 1 \tag{25}$$

where

$$\bar{H}_K^{(c)}(z_1) = \sum_{n_1=0}^{N_1-1} H_{n_1}(z_1)\, z_1^{2Kn_1}$$

$$\bar{X}_K^{(c)}(z_1) = \sum_{m_1=0}^{N_1-1} X_{m_1}(z_1)\, z_1^{2Km_1}. \tag{26}$$

The aperiodic convolution (10) is obtained by a polynomial transform of length $2N_1$, root z_1 and mod $(z_1^{N_1} + 1)$

$$Y_{l_1}^{(a)}(z_1) = \frac{1}{2N_1} \sum_{K=0}^{2N_1-1} \bar{H}_K^{(a)}(z_1)\, \bar{X}_K^{(a)}(z_1)\, z_1^{-Kl_1} \quad \text{mod } (z_1^{N_1} + 1),\ l_1 = 0, 1, \cdots, 2N_1 - 1 \tag{27}$$

where

$$\bar{H}_K^{(a)}(z_1) = \sum_{n_1=0}^{N_1-1} H_{n_1}(z_1)\, z_1^{Kn_1}$$

$$X_K^{(a)}(z_1) = \sum_{m_1=0}^{N_1-1} X_{m_1}(z_1)\, z_1^{Km_1}. \tag{28}$$

By inserting (25) and (27) into formula (9), we have (13) and (14).

REFERENCES

[1] R. C. Agarwal and C. S. Burrus, "Fast convolution using Fermat number transform with application to digital filtering," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-22, pp. 87–97, Apr. 1974.

[2] H. J. Nussbaumer, *Fast Fourier and Convolution Algorithms*. Berlin, Germany: Springer-Verlag, 1981, sect. 8.4.

Spectral Leakage Suppression Properties of Linear and Quadratic Windowing

V. J. MATHEWS AND D. H. YOUN

***Abstract*—It is shown that the leakage suppression properties of segment averaging spectrum estimation methods using linear windows and equivalent quadratic windows are asymptotically the same, under the assumption that segments relatively far apart are uncorrelated. Thus,**

Manuscript received November 21, 1983; revised February 10, 1984, and May 14, 1984.

The authors are with the Department of Electrical and Computer Engineering, University of Iowa, Iowa City, IA 52242.

0096-3518/84/1000-1092$01.00 © 1984 IEEE

for large data lengths, one can effectively replace linear windowing by its equivalent quadratic windowing and hope to get similar leakage suppression at a substantially reduced computational cost. A simulation example that supports this conclusion is presented.

INTRODUCTION

The usefulness of linear windowing to reduce spectral leakage is documented in literature [1]-[4]. In spite of these efforts, linear windowing has been criticized in the past mainly because it is computationally expensive. The critics of linear windowing advocate quadratic modification of the estimated spectrum [5]-[7] (i.e., using a rectangular linear window and then convolving the estimated spectrum with a suitably selected smoothing function). Analytical expressions for the estimation mean, variance, and spectral resolution have been derived in [9] when combined linear and quadratic windowing are used in segment averaging methods of spectrum estimation. However, there have been conflicting opinions about the usefulness of quadratic windowing for leakage suppression [2], [6], [8]. It was shown in [9] that by choosing the quadratic windows properly, one can compensate for the bad sidelobe behavior of the linear window used. However, [2] points out that for periodograms, the leakage suppression brought about by the improved sidelobe structure is masked by estimation noise. In this paper, we will consider the effects of quadratic windowing on spectral leakage, when segment averaging methods of spectrum estimation are used. Under the reasonable constraint that segments relatively far apart are uncorrelated, we will show that one can avoid linear windowing (i. e., use only a rectangular linear window function) and obtain good leakage suppression, provided that the quadratic window used compensates for the bad sidelobe structure of the rectangular linear window function and the number of averaged segments is large enough. In the next section we will develop some theoretical aspects and back them up with a simulation example in Section III.

II. SOME THEORETICAL ASPECTS

Let $x(k)$; $k = 0, 1, \cdots, p-1$, be a stationary time series. The $2M$-point estimate of $G_x(f)$, the spectrum of $x(k)$, using the weighted, overlapped, segment averaging (WOSA) algorithm [10], [11], is computed as

$$\hat{G}_x^{(l)}(f) = \frac{1}{N} \cdot \sum_{i=0}^{N-1} \hat{G}_{x,i}(f) \tag{1}$$

where

$$\hat{G}_{x,i}(f) = \frac{1}{r} \cdot \frac{1}{2M} \cdot |F\{x(iR+n)w_l(n)\}|^2 \tag{2}$$

is the spectrum of the ith weighted segment of $x(k)$, $F\{\cdot\}$ denotes the $2M$-point discrete Fourier transform of $\{\cdot\}$, R is the number of samples between adjacent segments, N is the total number of segments, $w_l(n)$ is a linear window of length $L, L \leqslant M$, and

$$r = \frac{1}{2M} \sum_{n=0}^{L-1} w_l^2(n). \tag{3}$$

Let

$$G_l(f) = \frac{1}{2M} \cdot |F\{w_l(n)\}|^2 \tag{4}$$

be the spectrum of $w_l(n)$ and let $W_q(f)$ be a quadratic window such that

$$\frac{1}{2M} \cdot \sum_{f=0}^{2M-1} W_q(f) = 1. \tag{5}$$

Now, convolving the estimate in (1) with the quadratic window function $W_q(f)$ yields a smoothed spectrum estimate, $\hat{G}_x^{(q)}(f)$. That is,

$$\hat{G}_x^{(q)}(f) = \hat{G}_x^{(l)}(f) \circledast W_q(f) \tag{6}$$

where $\circledast$ denotes complex convolution. We will restrict our attention to those $W_q(f)$ such that the effective window function $W_e(f) = G_l(f) \circledast W_q(f) \geqslant 0$ for all f. This will ensure that the expected value of $\hat{G}_x^{(q)}(f)$ is nonnegative.

It is easy to show that

$$E\{\hat{G}_x^{(l)}(f)\} = G_x(f) \circledast G_l(f) \tag{7}$$

and

$$E\{\hat{G}_x^{(q)}(f)\} = G_x(f) \circledast G_l(f) \circledast W_q(f) = G_x(f) \circledast W_e(f) \tag{8}$$

where $E\{\cdot\}$ denotes the statistical expectation of $\{\cdot\}$. Let

$$\hat{G}_{x,i}(f) = E\{\hat{G}_x^{(l)}(f)\} + Q_i(f) = G_x(f) \circledast G_l(f) + Q_i(f) \tag{9}$$

where $Q_i(f)$ denotes the estimation error for the spectrum estimate of the ith segment. Then, $\hat{G}_x^{(l)}(f)$ may be written as

$$\hat{G}_x^{(l)}(f) = G_x(f) \circledast G_l(f) + Q(f) \tag{10}$$

where

$$Q(f) = \frac{1}{N} \sum_{i=0}^{N-1} Q_i(f). \tag{11}$$

We will assume that the ith and jth segments are uncorrelated if $|i-j| > S$ where S is smaller than N. Assuming stationary Gaussian time series, one can show [2], [9] that, for $i = 0, 1, \cdots, N-1$,

$$V(f_1, f_2, 0) = E\{Q_i(f_1)\, Q_i(f_2)\}$$
$$\cong \frac{1}{(2M)^2} \sum_{\mu=0}^{2M-1} \sum_{\nu=0}^{2M-1} \cdot \frac{G_x(\mu) G_x(\nu) W_l(f_1-\mu) W_l^*(f_1-\nu) W_l^*(f_2-\mu) W_l(f_2-\nu)}{r^2 (2M)^2} \tag{12}$$

where * denotes complex conjugates and

$$W_l(f) = F\{w_l(n)\}. \tag{13}$$

Let

$$V(f_1, f_2, \tau) = E\{Q_i(f_1) Q_{i-\tau}(f_2)\} \tag{14a}$$

and

$$\rho(f_1, f_2, \tau) = V(f_1, f_2, \tau)/V(f_1, f_2, 0). \tag{14b}$$

Then, direct computation will yield

$$E\{Q(f_1)\, Q(f_2)\}$$
$$= \frac{V(f_1, f_2, 0)}{N} \sum_{\tau=-S}^{S} \frac{(N-|\tau|)}{N} \rho(f_1, f_2, \tau) \tag{15a}$$

$$\cong \frac{V(f_1, f_2, 0)}{N} \sum_{\tau=-S}^{S} \rho(f_1, f_2, \tau) \tag{15b}$$

if $S \ll N$. Since both $V(f_1, f_2, 0)$ and $\rho(f_1, f_2, \tau)$ are not functions of N and $S \neq N$, we have that, for fixed S, $E\{Q(f_1)Q(f_2)\}$ is inversely proportional to N, the total number of segments used.

One can now easily show that [2], [9] the variance of the spectrum estimate $\hat{G}_x^{(q)}(f)$ is given by

$$\operatorname{var}\{\hat{G}_x^{(q)}(f)\} = \frac{1}{(2M)^2} \sum_{\mu=0}^{2M-1} \sum_{\nu=0}^{2M-1} \cdot E\{Q(\mu)Q(\nu)\} W_q(f-\mu) W_q^*(f-\nu). \quad (16)$$

From (15a)–(16) we can see that var $\{G_x^{(q)}(f)\}$ is also inversely proportional to N, when S is kept constant. The estimation noise, whose mean squared value is given by (16), is also correlated [2].

Let $W_q(f)$ be such that $W_e(f) = G_l(f) \circledast W_q(f)$ when convolved with $G_x(f)$ produces smaller spectral leakage than when $G_l(f)$ is convolved with $G_x(f)$ [i.e., $W_e(f)$ has a better sidelobe structure than $G_l(f)$]. Then from (6), (8), and (10),

$$\hat{G}_x^{(q)}(f) = G_x(f) \circledast W_e(f) + Q(f) \circledast W_q(f). \quad (17)$$

The first term of the right-hand side of (17) shows smaller spectral leakage than $G_x(f) \circledast G_l(f)$. However, this reduction in spectral leakage is masked by the second term, $Q(f) \circledast W_q(f)$ which has an rms value of the order of var$^{1/2}$ $\{\hat{G}_x^{(q)}(f)\}$ and, as discussed before, is inversely proportional to $N^{1/2}$.

The above discussion implies that for a fixed S, as the number of segments increases, the masking of the leakage suppression effect by the correlated estimation noise decreases, thereby making it possible for one to apply only a proper quadratic window and obtain good leakage suppression properties when the observation time is very large.

Remark: The discussions in this section imply that a linear window with spectrum $G_l(f)$ and a quadratic window $W_q(f)$ such that

$$G_l(f) = W_q(f) \circledast G_R(f), \quad (18)$$

where $G_R(f)$ is the spectrum of the rectangular window, have asymptotically equivalent leakage suppression properties, under the condition that segments more than S apart are uncorrelated. However, it should be understood that by increasing the overlap between adjacent segments, we cannot obtain the same result. There is a definite limit to the variance reduction possible with increased overlap [12].

III. A Simulation Example

The true spectrum of the signal used in this example, in decibels, is plotted as curve 1 in all the figures. The signal was obtained using three independent zero-mean, white, Gaussian sequences of variances 1, 0.001, and 0.0001, respectively, sampled at 2000 Hz. The first two sequences were processed through 10th-order Butterworth bandpass filters with passbands 200–400 Hz and 600–800 Hz, respectively, and were added together along with the third sequence. Signals similar to this have been used in [2].

In this example, we will compare spectrum estimates of this signal, obtained using 1) the WOSA algorithm using a 64 point Hamming window, 128 point fast Fourier transforms (FFT), and 50 percent overlap and 2) quadratic modification of the WOSA estimate using a 64 point rectangular window, 128 point FFT's, and 50 percent overlap. The quadratic window $W_q(f)$ was such that the corresponding effective window was the same as the spectrum of the Hamming window used in the WOSA method.

Figs. 1(a), (b) and 2(a), (b) display the spectrum estimates obtained using 500 and 20 000 data points, respectively. In Figs. 1(a) and 2(a), curves marked 2 are the estimates in decibels obtained using the linear Hamming window. In Figs. 1(b) and 2(b), we have plotted the results obtained using the rectangular window (curves marked 2) and using quadratic modification of these estimates (curves marked 3).

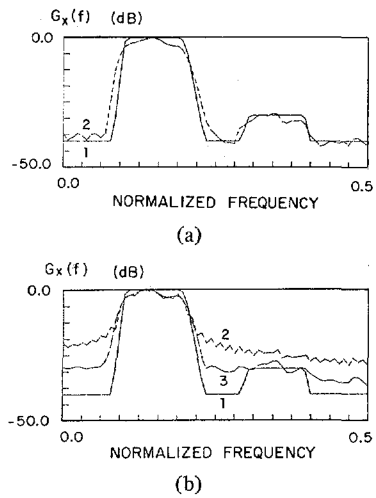

Fig. 1. Spectrum estimates using 500 data points. (a) 1) True spectrum and 2) using a linear Hamming window. (b) 1) True spectrum, 2) using a linear rectangular window, and 3) after quadratic windowing.

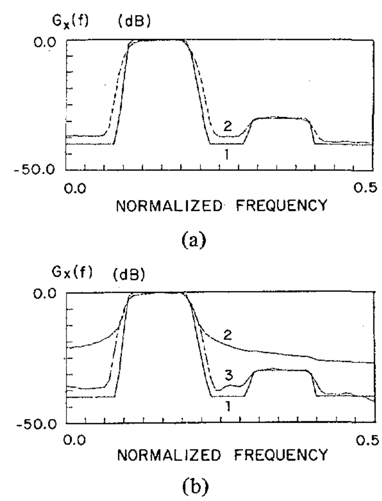

Fig. 2. Spectrum estimates using 20 000 data points. (a) 1) True spectrum and 2) using a linear Hamming window. (b) 1) True spectrum, 2) using a linear rectangular window, and 3) after quadratic windowing.

We can observe from Figs. 1 and 2 that as the observation time increases (thereby increasing the number of segments), the estimates using the quadratic window became more and more comparable to those obtained using linear window only. In fact, plot 2 in Fig. 2(a) and plot 3 in Fig. 2(b) are virtually identical.

Remarks:

1) As discussed before, one of the major advantages of quadratic windowing is the reduced number of multiplications involved in the computations. It has been shown in [9] that for the same effective window function, one can use larger nonoverlapped segments and no linear windowing (thereby reducing the number of segments and computations) without sacrificing the stability of the estimates. Table I demonstrates the approximate savings in the number of multiplications for different segment lengths and overlaps. In Table I, we have neglected the computations for quadratic windowing operations. Also, the percentage savings is approximately independent of the data length.

TABLE I
APPROXIMATE PERCENTAGE REDUCTION IN MULTIPLICATIONS OVER 64 POINT LINEAR WINDOWING WITH 50 PERCENT OVERLAP FOR DIFFERENT INITIAL SEGMENT LENGTHS AND OVERLAP WHEN ONLY QUADRATIC WINDOWING IS USED

Case	Segment Length	Percent Overlap	Percent Savings in Multiplications
1	64	50	10
2	64	0	55
3	128	0	50
4	256	0	45

To obtain all the leakage suppression capability of the sidelobe structure of the effective window function using only quadratic windowing, the number of averaged segments N should be such that the estimation noise in (17) is of the order of the sidelobe levels of the effective window. However, a smaller number of segments will work if the lowest level of the spectrum being estimated is much larger than that of the sidelobe levels. To decide whether one can use quadratic windowing alone and obtain good leakage suppression, one should first make a decision on an acceptable estimation noise level based on prior knowledge of the dynamic range of the spectra and see if the data length available is long enough to yield smaller estimation variances than this level. Exact expressions for estimation variance may be found in [9].

IV. CONCLUSIONS

We have shown that linear windowing and its equivalent quadratic windowing have asymptotically similar leakage suppression properties, under the assumption that data segments relatively far apart are uncorrelated. Thus, when the total number of segments tends to be large, one can effectively replace linear windowing with equivalent quadratic windowing and hope to get almost the same leakage suppression at substantially reduced computational costs. However, when the number of segments tends to be small, one still needs linear windowing for effective leakage suppression and this supports the conclusions of [2]. The example presented supports the above conclusions.

REFERENCES

[1] F. J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proc. IEEE*, vol. 66, pp. 51–83, Jan. 1978.
[2] C. V. Schooneveld and D. J. Frijling, "Spectral analysis: On the usefulness of linear tapering for leakage suppression," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 323–329, Apr. 1981.
[3] D. R. Brillinger, "The key role of tapering in spectrum estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 1075–1076, Oct. 1981.
[4] A. H. Nuttall, "Some windows with very good sidelobe behavior," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 84–91, Feb. 1981.
[5] C. K. Yuen, "On the smoothed periodogram method for spectrum estimation," *Signal Processing*, vol. 1, pp. 83–86, 1979.
[6] ——, "Quadratic windowing in the segment averaging method for power spectrum computation," *Technometrics*, vol. 20, pp. 195–200, May 1978.
[7] E. A. Sloan, "Comparison of linearly and quadratically modified spectral estimates of Gaussian signals," *IEEE Trans. Audio Electroacoust.*, vol. AU-17, pp. 133–137, June 1969.
[8] C. Bringham, M. D. Godfrey, and J. W. Tukey, "Modern techniques of power spectrum estimation," *IEEE Trans. Audio Electroacoust.*, vol. AU-15, pp. 56–66, June 1967.
[9] A. H. Nuttall and G. C. Carter, "Spectrum estimation using combined time and lag weighting," *Proc. IEEE*, vol. 70, pp. 1115–1125, Sept. 1982.
[10] P. D. Welch, "The use of FFT for estimation of power spectra: A method based on time averaging of short modified periodograms," *IEEE Trans. Audio Electroacoust.*, vol. AU-15, pp. 70–73, June 1967.
[11] A. H. Nuttall, "Spectral estimation by means of overlapped fast Fourier transform processing," NUSC Tech. Rep. 4169, 13, Oct. 1971.
[12] G. C. Carter, C. H. Knapp, and A. H. Nuttall, "Estimation of magnitude-squared coherence function via overlapped fast Fourier transform processing," *IEEE Trans. Audio Electroacoust.*, vol. AU-21, pp. 337–344, Aug. 1973.

Pole-Zero Analysis of Voiced Speech Using Group Delay Characteristics

NAOKI MIKAMI AND RYOJI OHBA

***Abstract*—The group delay characteristics of the vocal tract are derived from the approximate spectral envelope defined by peaks at the harmonics of the pitch frequency. Formant and anti-formant frequencies are extracted from the group delay by a simple peak-picking method.**

I. INTRODUCTION

Considerable studies have been reported to extend linear prediction to a speech production system which contains both poles and zeros. One of them, which was proposed by Yegnanarayana [1], is based on the fact that a spectrum can be decomposed into an AR (all-pole) part and an MA (all-zero) part by using the negative derivative of the phase spectrum (NDPS), that is, the group delay characteristics. In his method, the group delay characteristics of the vocal tract are calculated from the cepstrum, eliminating effects of the glottal source by a short-pass lifter (filter in the quefrency domain). However, there are two defects in this method. First, the analysis is very sensitive to the lifter length [1]. Second, as in the case of general cepstrum analysis, the resulting estimate is seriously affected by the low-level part in the spectral fine structure [2].

In this correspondence, the authors propose an improved pole-zero analysis method which reduces these defects by using an approximate spectral envelope of the voiced speech instead of the spectrum itself to calculate the cepstrum.

II. PROPERTY OF GROUP DELAY

Suppose that the vocal tract may be represented by a cascade of M resonators and N anti-resonators; its frequency response is

$$H(\omega) = K \frac{\prod_{n=1}^{N} (b_n - j\omega)(b_n^* - j\omega)}{\prod_{m=1}^{M} (a_m - j\omega)(a_m^* - j\omega)} \quad (1)$$

Manuscript received October 5, 1982; revised May 9, 1983, September 26, 1983, and March 5, 1984.

The authors are with the Department of Applied Physics, Faculty of Engineering, Hokkaido University, Sapporo 060, Japan.

0096-3518/84/1000-1095$01.00 © 1984 IEEE