value. Thus, we might expect that the true value is 0.81 (0.7/0.865). In both cases we expect that the true MSC is in about the 0.7 to 0.8 range, but our estimates varied from 0.4 to 0.7 because of misalignment when estimating the MSC.

IV. Conclusion

In conclusion, we see that even with a large number of FFT segments, estimates of the magnitude-squared coherence can be significantly biased downward, giving an erroneous indication of the value of the coherence. When the data are realigned and processed, estimates of the coherence are informative descriptors of the extent to which the ocean channel can be modeled by a linear time-invariant filter.

References

[1] G. M. Jenkins and D. G. Watts, *Spectral Analysis and Its Applications.* San Francisco: Holden-Day, 1968.
[2] L. H. Koopmans, *The Spectral Analysis of Time Series.* New York: Academic, 1974.
[3] G. C. Carter and C. H. Knapp, "Coherence and its estimation via the partitioned modified chirp-z transform," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-23, pp. 257–264, June 1975.
[4] G. C. Carter, C. H. Knapp, and A. H. Nuttall, "Estimation of the magnitude-squared coherence function via overlapped fast Fourier transform processing," *IEEE Trans. Audio Electroacoust.*, vol. AU-21, pp. 337–344, Aug. 1973.
[5] A. H. Nuttall and G. C. Carter, "Bias of the estimate of magnitude-squared coherence," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-24, pp. 582–583, Dec. 1976.
[6] J. S. Bendat and A. G. Piersol, *Random Data: Analysis and Measurement Procedures.* New York: Wiley, 1971. (New engineering text on coherence to appear in 1980.)
[7] W. G. Halvorsen and J. S. Bendat, "Noise source identification using coherent output power spectra," *J. Sound and Vibration*, pp. 15–24, Aug. 1975.
[8] D. R. Brillinger, *Time Series Data Analysis and Theory.* New York: Holt, Rinehart, and Winston, 1975.
[9] H. Akaike and Y. Yamanouchi, "On the statistical estimation of frequency response function," *Ann. Inst. Statist. Math.*, vol. 14, pp. 23–56, 1963.
[10] G. C. Carter, A. H. Nuttall, and P. G. Cable, "The smoothed coherence transform," *Proc. IEEE*, vol. 61, pp. 1497–1498, 1973.
[11] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-24, Aug. 1976.
[12] J. P. Kuhn, "Detection performance of the smoothed coherence transform (SCOT)," presented at 1978 Int. Conf. Acoust., Speech, Signal Processing.

A Weighted Overlap-Add Method of Short-Time Fourier Analysis/Synthesis

R. E. CROCHIERE

Abstract—**In this correspondence we present a new structure and a simplified interpretation of short-time Fourier synthesis using synthesis windows. We show that this approach can be interpreted as a modification of the overlap-add method where we inverse the Fourier transform and window by the synthesis window prior to overlap-adding. This simplified interpretation results in a more efficient structure for short-time synthesis when a synthesis window is desired. In addition, we show how this structure can be used for analysis/synthesis applications which require different analysis and synthesis rates, such as time compression or expansion.**

Manuscript received May 25, 1979; revised September 5, 1979.
The author is with Bell Laboratories, Murray Hill, NJ 07974.

I. Introduction

The concepts of short-time Fourier analysis and synthesis have been widely used for analyzing and modeling quasi-stationary (slowly time-varying) signals, such as speech. These concepts have evolved from two fundamental points of view.

One point of view is basically that of a filter-bank model. The input signal is filtered by a bank of bandpass filters which span the frequency range of interest. The outputs of the bandpass filters can then be used to define a short-time Fourier spectrum. This model is the basis for the phase vocoder developed by Flanagan and Golden [1]. In the synthesis procedure a signal can be reconstructed from its short-time Fourier spectra by summing the outputs of the bandpass filters. This method of analysis and synthesis is referred to as the filter-bank summation method. In practice, since the output signals from the filter bank are narrow-band signals, they can be sampled at a lower sampling rate than the initial input signal and interpolated back to a high sampling rate for synthesis.

The second point of view is basically that of a block-by-block analysis in time. The input signal is time windowed into overlapping finite duration time segments [8]. Each segment is then Fourier transformed to give a short-time Fourier spectrum. Schafer and Rabiner showed how this analysis technique can be efficiently implemented using the FFT algorithm [2]. Allen [3] carefully discussed a method of synthesizing a signal from its short-time Fourier spectra by inverse transforming each sample of the short-time spectra to recover the short-time segments of the signal in time. These overlapped signal segments are then appropriately summed (overlapped and added) to reproduce the time signal. This method is referred to as the overlap-add synthesis method.

More recently, Allen and Rabiner compared the effects of modifications of the short-time spectrum in the filter-bank summation and overlap-add methods. Portnoff [5], [9] has fully developed the use of a synthesis window and has developed general expressions for short-time analysis/synthesis using distinct analysis and synthesis windows [6], [9]. He has proposed an implementation of short-time synthesis based on the FFT algorithm in which inverse transformed short-time Fourier spectra are appropriately interpolated by the synthesis filter to obtain the desired reconstructed output. In this paper we show that Portnoff's windowed synthesis procedure can be greatly simplified and that it can be implemented in the form of a weighted overlap-add procedure of short-time synthesis. In this technique the short-time spectra are inverse transformed to produce the short-time signal segments. These short-time signal segments are weighted by the synthesis window. They are then overlapped and added in a manner similar to the overlap-add synthesis procedure. The synthesis window has the same effect in the synthesis procedure of the filter-bank summation method.

Thus, two principle methods of implementation (or interpretation) of short-time Fourier analysis/synthesis can be readily defined (various modifications of these two basic methods, of course, can be generated). They are the filter-bank summation method with decimation (reduced sampling rate) and interpolation in each channel, and the weighted overlap-add method in which the input of the discrete Fourier transform is windowed by the analysis window and the output of the inverse transform is windowed by the synthesis window in the manner to be described in this paper. Both methods are completely general in terms of implementating arbitrary analysis and synthesis windows (filters) and they are mathematically identical and interchangeable (assuming the analysis and synthesis windows are appropriately defined). The mathematical basis is similar to that of Portnoff [6] and, depending on the specific synthesis windows defined, the methods of filter-bank sum and overlap-add analyzed by Allen [3] and Allen and Rabiner [4] result.

II. Review of the Mathematical Framework for Short-Time Analysis/Synthesis

As stated above, the mathematical framework considered here is similar to that derived by Portnoff—only the interpretation of the DFT synthesis procedure is changed. Therefore, the notation that we will use is similar to Portnoff's, with small changes for convenience, and we refer to [6] for details of the derivations.

A. Analysis

The discrete short-time Fourier transform of a signal $x(m)$, sampled at equispaced frequencies every R samples in time m is commonly defined in the form

$$X_k(sR) = \sum_{m=-\infty}^{\infty} h(sR - m)\, x(m)\, W_M^{-mk} \tag{1}$$

where

$$W_M = e^{j2\pi/M}. \tag{2}$$

M is the number of frequency samples, k is the discrete frequency index, $h(m)$ is the analysis window, and s denotes the time index of the short-time transform at the decimated sampling rate (decimated by the integer factor R).

By a change of variables $r = m - sR$, (1) can be modified to the form

$$\begin{aligned} X_k(sR) &= \sum_{r=-\infty}^{\infty} h(-r)\, x(r + sR)\, W_M^{-(sR+r)k} \\ &= W_M^{-sRk}\, \tilde{X}_k(sR) \end{aligned} \tag{3}$$

where $X_k(sR)$ is the short-time transform referenced to a fixed time origin $m = s = 0$ and $\tilde{X}_k(sR)$ is the short-time transform referenced to a linearly increasing (sliding) time reference $m = sR$, which corresponds to the origin of the sliding analysis window. As seen by (3), these two short-time transforms are related by a linear phase component and their magnitudes are identical. The short-time transform $\tilde{X}_k(sR)$ can be expressed in the DFT form

$$\tilde{X}_k(sR) = \sum_{m=0}^{M-1} \tilde{x}_m(sR)\, W_M^{-mk} \tag{4}$$

where $\tilde{x}_m(sR)$ is the time-aliased signal

$$\tilde{x}_m(sR) = \sum_{l=-\infty}^{\infty} x(sR + lM + m)\, h(-lM - m). \tag{5}$$

The interpretation of this time-aliased signal will become more clear in the next section.

Since a linear phase shift in the DFT frequency domain corresponds to a circular or modulo rotation in the time domain, $X_k(sR)$ can also be defined as the DFT of the circularly rotated version of $\tilde{x}_m(sR)$, i.e.,

$$X_k(sR) = \sum_{m=0}^{M-1} x_m(sR)\, W_M^{-mk} \tag{6}$$

where

$$x_m(sR) = \tilde{x}_{((m-sR))_M}(sR) \tag{7}$$

and where $((n))_M$ denotes n reduced modulo M. Equations (1)–(7) define the basic mathematical framework for short-time analysis.

B. Synthesis

The synthesis of a signal $\hat{x}(n)$, from its discrete short-time transform $\hat{X}_k(sR')$, sampled every R' samples in time n, can be obtained from the following generalized synthesis formula [6, (3.16)]

$$\hat{x}(n) = \sum_{s=-\infty}^{\infty} f(n - sR')\, \frac{1}{M} \sum_{k=0}^{M-1} \hat{X}_k(sR')\, W_M^{nk} \tag{8}$$

where $f(n)$ defines the synthesis window. The conditions on the windows (filters) $f(n)$ and $h(n)$, in order to achieve an exact resynthesis, are derived in [6]. By defining $\hat{x}_n(sR')$ as the inverse DFT of $\hat{X}_k(sR')$, i.e.,

$$\hat{x}_n(sR') = \frac{1}{M} \sum_{k=0}^{M-1} \hat{X}_k(sR')\, W_M^{nk}, \tag{9}$$

(8) becomes

$$\hat{x}(n) = \sum_{s=-\infty}^{\infty} f(n - sR')\, \hat{x}_n(sR'). \tag{10}$$

In a manner similar to that in the analysis stage, a short-time transform $\hat{\tilde{X}}_k(sR')$ can be defined as

$$\hat{\tilde{X}}_k(sR') = W_m^{sR'k}\, \hat{X}_k(sR') \tag{11}$$

and it corresponds to the short-time transform referenced to the linearly increasing (sliding) time frame $n = sR'$. The inverse transform of $\hat{\tilde{X}}_k(sR')$ can be defined as $\hat{\tilde{x}}_n(sR')$, and from (11) and the relation between linear phase shift in the DFT frequency domain [used to derive (7)] it can be shown that

$$\hat{\tilde{x}}_n(sR') = \hat{x}_{((n+sR'))_M}(sR'), \tag{12}$$

or equivalently

$$\hat{x}_n(sR') = \hat{\tilde{x}}_{((n-sR'))_M}(sR'). \tag{13}$$

Applying (13) to (10) gives

$$\hat{x}(n) = \sum_{s=-\infty}^{\infty} f(n - sR')\, \hat{\tilde{x}}_{((n-sR'))_M}(sR'). \tag{14}$$

Thus, it is seen that at the time reference n replaced by $n + s_0R'$ the s_0 term in (14) contributes a component $f(n)\, \hat{\tilde{x}}_n(s_0R')$ to the time shifted signal $\hat{x}(n + s_0R')$. It should be noted that each term in the sum of (14) is itself a sequence in n, with each term concentrated in a different region of the n axis. This form will be useful in the next section to define the weighted overlap-add synthesis procedure.

III. Weighted Overlap-Add Implementation of Short-Time Analysis/Synthesis

With the above mathematical framework, we can now illustrate the weighted overlap-add implementation. To perform the analysis we first select the appropriate section of the input signal and window it by the analysis window. We will assume

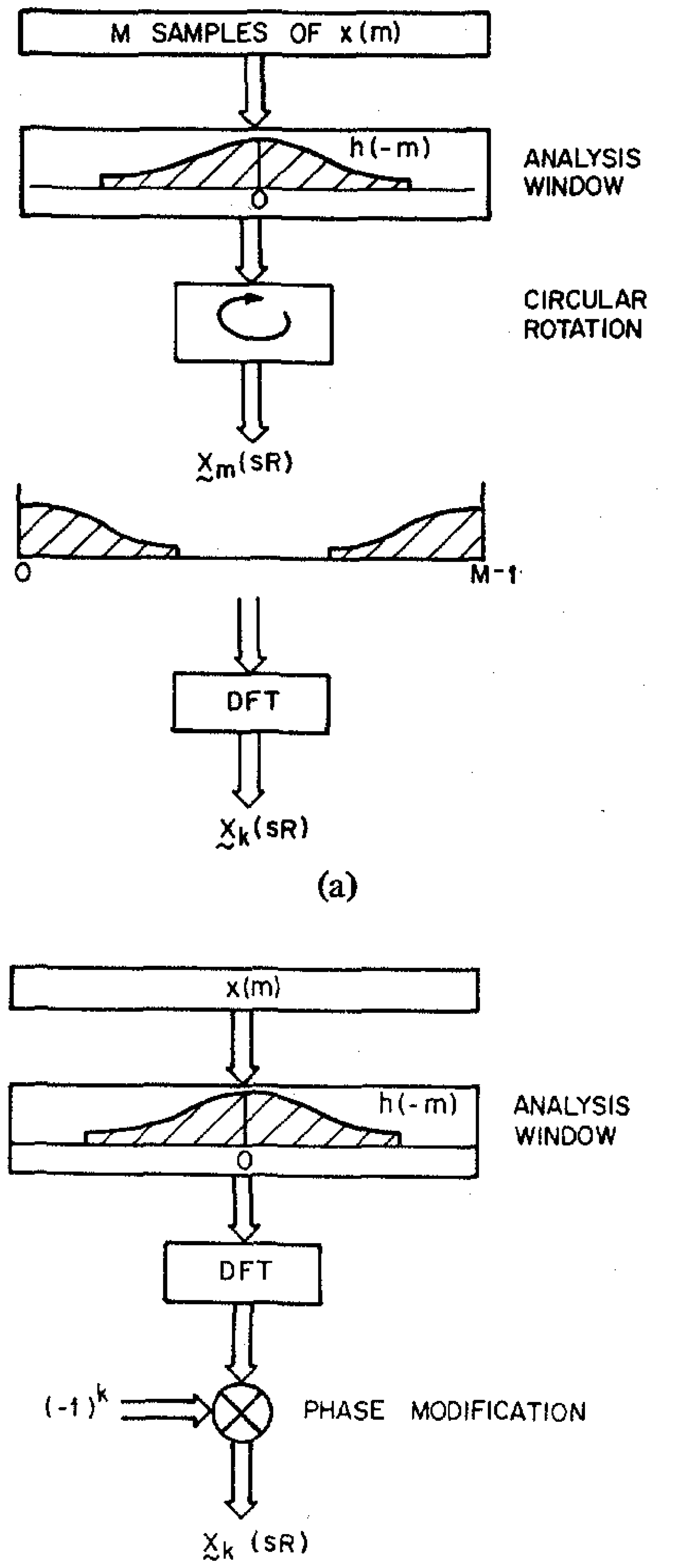

Fig. 1. (a) Generation of the short-time signal $\underset{\sim}{x}_m(sR)$ by windowing and circulary shifting by $M/2$ samples. (b) An equivalent method based on a phase modification of the DFT.

that the duration of $h(m)$ is less than M where M is the transfrom size (assumed to be even). We select a block of M samples $(x(m + sR - M/2), m = 0, 1, 2, \cdots, M - 1)$ and window the block by $h(-m)$, as seen in Fig. 1(a), such that $h(0)$ aligns with the $m = M/2$ sample in the block. Assuming that $h(m)$ is a symmetric zero-phase (odd number of taps) FIR window and that we wish to align the center sample of the window to the first sample of the transform, the windowed signal can be circularly rotated by $M/2$ samples as seen in Fig. 1(a). The resulting signal corresponds to the short-time signal $\underset{\sim}{x}_m(sR)$ according to (5), and its DFT corresponds to $\underset{\sim}{X}_k(sR)$. Since a circular shift by $M/2$ samples corresponds to a phase shift of $\exp(-j2\pi k(M/2)/M) = (-1)^k$, an alternate form of this implementation is shown in Fig. 1(b).

The upper part of Fig. 2 shows the overall implementation of the short-time analysis based on (3)–(7) and on the interpretation in Fig. 1(b). The input signal is buffered with an M sample buffer. The contents of the buffer are copied every R samples and windowed by the analysis window $h(-m)$. This windowed segment of speech is transformed to give the short-time Fourier transform in the sliding time reference as described above and in Fig. 1. It is then multiplied by the phase factor $(-1)^k W_M^{-sRk}$ to convert it to the fixed time reference.

From (8)–(14), a similar inverse synthesis structure can be derived, as shown in the lower part of Fig. 2. The short-time transform $\hat{X}_k(sR')$ is first multiplied by $W_M^{sR'k}$, according to (11) to convert it from a fixed time reference to the linearly

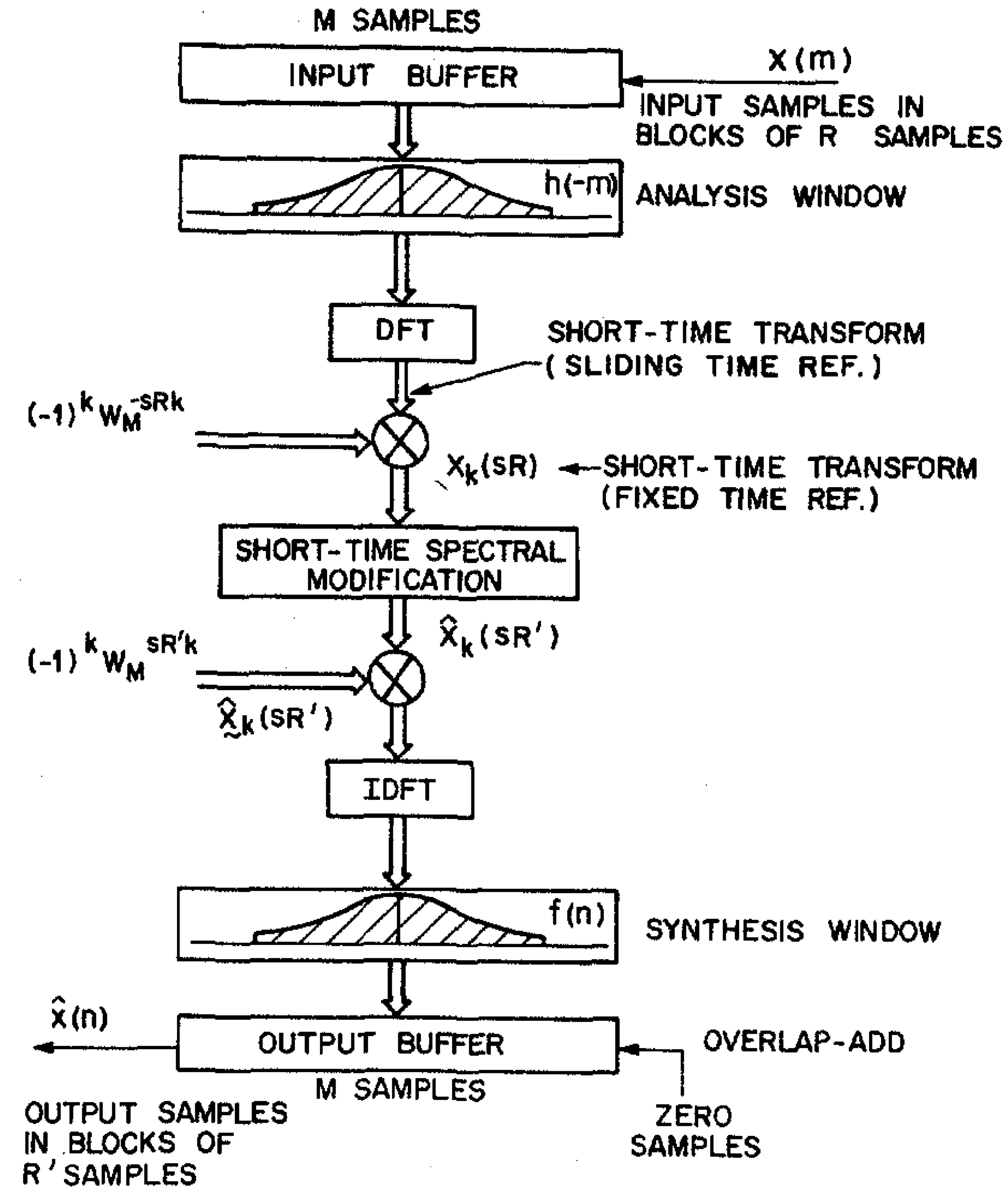

Fig. 2. A weighted overlap-add procedure for short-time Fourier analysis/synthesis using linear phase modification in the frequency domain.

increasing time referenced short-time transform $\hat{\underset{\sim}{X}}_k(sR')$. It is then multiplied by $(-1)^k$ to account for the circular rotation in the time domain by $M/2$ samples (i.e., the inverse operation to that shown in Fig. 1). The resulting phase-shifted signal is then inverse transformed and windowed by the synthesis window. These inverse-transformed windowed short-time signals are then summed into the output buffer in an overlap-add manner according to (14). At the synthesis (block) time s_0, the s_0 term in (14) is summed into the output buffer. The resulting structure is seen in the lower part of Fig. 2.

The overall structure of Fig. 2 is seen to be a modification of the overlap-add structure. Input samples $x(n)$ are shifted into the input buffer in blocks of R samples and samples of the output signal $\hat{x}(n)$ are shifted out of the output buffer in blocks of R' samples, with zero valued samples filling in the rightmost R' samples of the output buffer. The entire structure is implemented in a block-by-block manner with the operations between the two buffers being performed once per block.

Note that for generality we have allowed R and R' to be different. This can be useful in applications such as time expansion or compression where the input and output rates may be different [6].

If the operation of the phase modification in Fig. 2 is commutable with the short-time spectral modification, then it is possible to replace the phase modifications $(-1)^k W_M^{-sRk}$ and $(-1)^k W_M^{sR'k}$ with a single phase modification of the form $W_M^{s(R'-R)k}$. Furthermore, if $R' = R$, the phase modifications can be eliminated entirely and the spectral modifications can be made directly on the short-time transform in the sliding time reference (again *only* if the phase modification and spectral modification operations are commutable).

By using the fact that multiplication by a linear phase shift in the DFT frequency domain corresponds to a circular rotation in the time domain, the alternate analysis structure of Fig.1(a) results. Similary, an alternate synthesis structure can be defined in which $\hat{X}_k(sR')$ is first inverse transformed and then circularly rotated by $((n + sR' + M/2))_M$ samples accord-

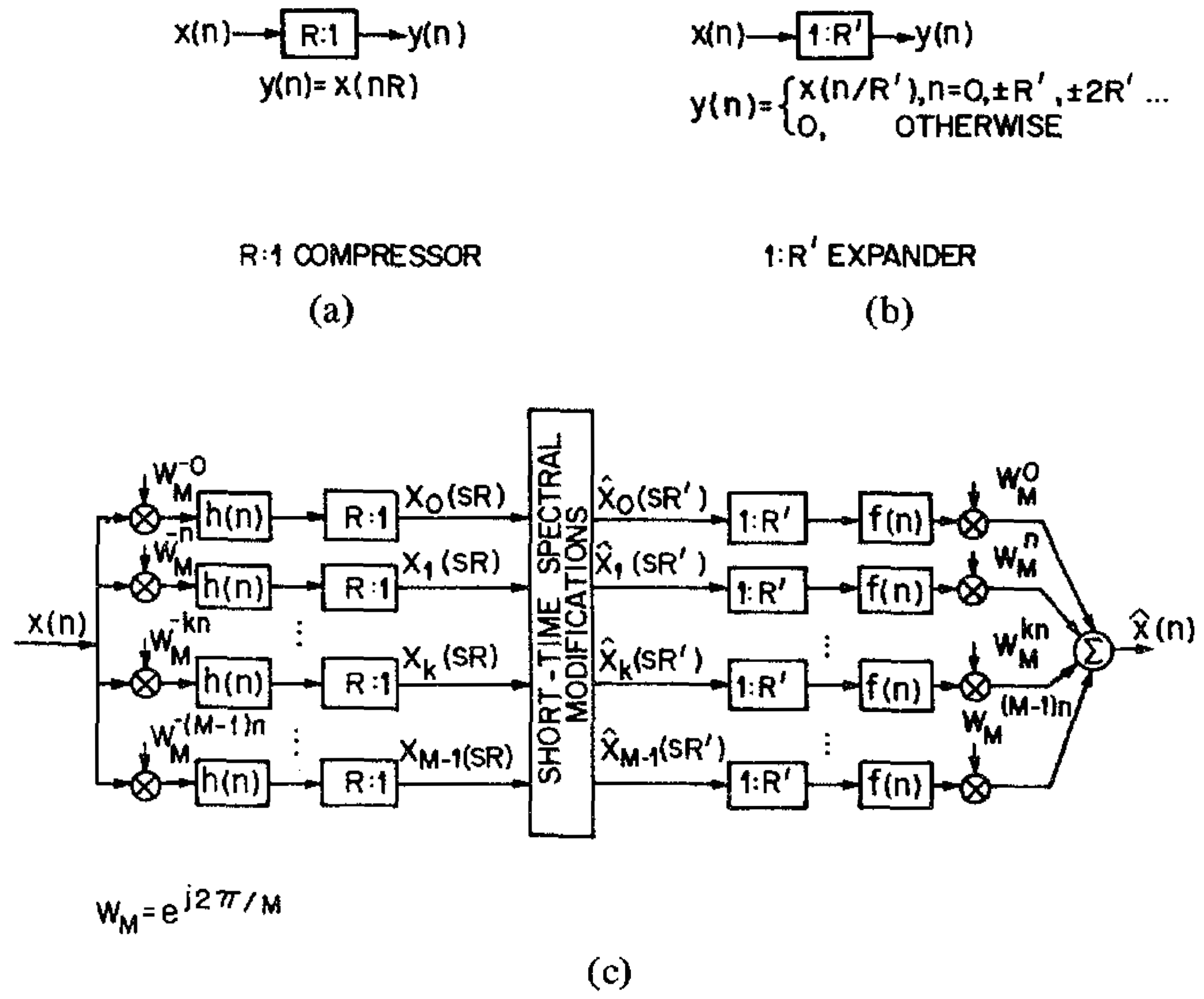

Fig. 3. The filter-bank summation interpretation of short-time Fourier analysis/synthesis.

ing to (12) to give $\hat{x}_n(sR')$ (multiplies by $W_M^{s(R'-R)k}$ can also be implemented with a circular shift). This signal is then weighted by $f(n)$ and overlap-added into the output buffer in the same manner as in Fig. 2.

The advantages of the implementation of Fig. 2 of the short-time synthesis over the structure proposed by Portnoff [5], [6] is that it requires substantially less internal storage of data and it is more easily programmed. Only M samples of partial sums of products in the overlap-add output buffer need be stored, as opposed to requiring the storage of many inverse transformed short-time segments. The number of actual arithmetic operations, however, is the same as that of Portnoff's (assuming the alternate method of circular rotation is used in place of the phase modification).

Another advantage of the implementation of Fig. 2 is that it clearly shows the relationship of the filter-bank synthesis method to that of the overlap-add synthesis method. For this reason we refer to this method as a weighted overlap-add technique. If an M point rectangular synthesis window is used, then the synthesis method is identical to the overlap-add technique proposed by Allen [3].

IV. Relationship to the Modified Filter-Bank Summation Method

An alternative interpretation of short-time analysis/synthesis is that of the modified filter-bank summation approach shown in Fig. 3. Since this structure is derived from the same mathematical framework as the weighted overlap-add structure, it is clear that they are equivalent structures. The values $X_k(sR)$ in Fig. 3 are identical to those in Fig. 2. Both structures can implement arbitrary analysis and/or synthesis windows. In the case where $R = R' = 1$ and $f(n) = \delta(n)$ (a unit pulse, $\delta(n) = 1$ for $n = 0$, and $\delta(n) = 0$ for $n \neq 0$), it can be seen that the modified filter-bank summation method becomes identical to the filter-bank summation approach discussed by Allen and Rabiner [4].

It is also interesting to note that if we take the transpose of either of the structures of Fig. 2, or Fig. 3, the structures remain the same; however, the roles of the analysis and synthesis windows and R and R' are interchanged. This is consistent with the concepts of transposition of linear time varying systems discussed by Claasen and Mecklenbrauker [7].

V. Conclusion

In this paper we have proposed a weighted overlap-add implementation of short-time analysis/synthesis. The advantages of this scheme are that it is more efficient, in terms of storage, and easier to program than the method used by Portnoff. It also clearly illustrates the relationship between the weighted overlap-add and the filter-bank summation methods of short-time analysis/synthesis using arbitrary analysis and synthesis windows. We have also attempted to point out the differences between the sliding and the fixed time references which are often confused in these methods. The basic mathematical framework is similar to that derived by Portnoff, including effects of short-time modification—only the interpretation of the structure is different.

Acknowledgment

The author would like to thank J. B. Allen and M. R. Portnoff for comments in the development of this paper.

References

[1] J. L. Flanagan and R. Golden, "Phase vocoder," *Bell Syst. Tech. J.*, vol. 45, pp. 1493–1509, Nov. 1966.

[2] R. W. Schafer and L. R. Rabiner, "Design and simulation of a speech analysis-synthesis system based on short-time Fourier analysis-synthesis system based on short-time Fourier analysis," *IEEE Trans. Audio Electroacoust.*, vol. AU-21, pp. 165–174, June 1973.

[3] J. B. Allen, "Short-term spectral analysis synthesis, and modification by discrete Fourier transform," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-25, pp. 235–238, June 1977.

[4] J. B. Allen and L. R. Rabiner, "A unified approach to short-time Fourier analysis and synthesis," *Proc. IEEE*, vol. 65, pp. 1558–1564, Nov. 1977.

[5] M. R. Portnoff, "Implementation of the digital phase vocoder using the fast Fourier transform," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-24, pp. 243–248, June 1976.

[6] M. R. Portnoff, "Time scale modification of speech based on short-time Fourier analysis," Ph.D. dissertation, Massachusetts Inst. Tech., Cambridge, Apr. 1978.

[7] T. A. C. M. Claasen and W. F. G. Mecklenbrauker, "On the transportation of linear time-varying discrete-time networks and its application to multirate digital systems," *Philips J. Res.*, vol. 23, pp. 78–102, 1978.

[8] D. Gabor, "Theory of communication," *J. IEE*, no. 93, pp. 429–457, Nov. 1946.

[9] M. R. Portnoff, "Time-frequency representation of digital signals and systems based on short-time Fourier analysis," *IEEE Trans. Acoust., Speech, Signal Processing*, this issue, pp. 55–69.

Generating Covariance Sequences and the Calculation of Quantization and Rounding Error Variances in Digital Filters

JEAN-PIERRE DUGRÉ, ALOYSIUS A. L. BEEX, AND LOUIS L. SCHARF

Abstract—**A linear algorithm is given for the generation of covariance sequences for rational digital filters using numerator and denominator coefficients directly. There is no need to solve a Lyapunov equation or**

Manuscript received December 22, 1978; revised June 8, 1979 and July 23, 1979. This work was partially supported by the Office of Naval Research, Statistics and Probability Branch, Arlington, VA, under Contract N00014-75-C-0518.

The authors are with the Department of Electrical Engineering, Colorado State University, Fort Collins, CO 80523.

0096-3518/80/0200-0102$00.75 © 1980 IEEE