Experimental Comparison of Fourier Transformation and Model-Fitting Methods of Spectral Analysis

William A. Gardner

Department of Electrical and Computer Engineering, University of California, Davis, Davis, California 95616

ABSTRACT

A summary of the results of an extensive comparative experimental study of Fourier transformation and model-fitting methods of spectral analysis of random time-series data is presented. It is illustrated that Fourier transformation methods can be an essential companion to model-fitting methods even for short data segments with underlying sharp spectral peaks. The best spectrum estimates can be obtained by taking advantage of the strengths of both types of methods. For example, it is shown that detection and estimation of the frequencies of spectral lines for short data segments can be best accomplished using certain parametric methods in conjunction with Fourier transformation methods to aid in model-order selection and identification of spurious peaks in the parametric spectrum estimate, and that estimation of amplitude and phase for sine-wave removal, given frequency estimates, and spectrum estimation after sine-wave removal can often be best accomplished with Fourier transformation methods alone.

I. INTRODUCTION

During the last twenty-five years there has been a tremendous amount of research work on parametric methods of spectral analysis based on fitting models to data. The primary driving force behind this work has been the promise of substantial improvements in spectral resolution performance over that obtainable with the older (sometimes called *classical*) nonparametric methods based on Fourier transformation of data. It has been found that the relative merits of the newer (sometimes called *modern*) methods depend strongly on the nature of the data being analyzed and the amount of advance knowledge about the underlying ideal spectrum being estimated. There have been two main classes of problems studied. The one of interest in this article concerns the estimation of an ideal spectrum of a theoretically infinite temporal series using a finite time segment of data. The other concerns the estimation of an ideal spectrum of a theoretically infinite spatial series using a set of finite space segments indexed by time. This latter problem area arises in applications involving arrays of sensors, such as geophones, hydrophones, radio antennas, and so on. The fact that the measurements of the finite-length spatial series can be averaged over time makes this spectral analysis problem significantly different from the former problem where there is only one finite-length time series and consequently no ensemble to average over.* Parametric methods of spectral analysis are usually superior to nonparametric methods when high-resolution spectral analysis of a time-indexed set of short spatial series is desired. The reason is that although only a short segment (in the lag parameter) of the autocorrelation can be measured, these measurements can be highly accurate. However, for the problem of spectral analysis of a single short time series, the relative merits of parametric and nonparametric methods are not at all clear cut and depend in a complicated way on what is known and what is unknown about the physical phenomenon giving rise to the data and therefore about the ideal spectrum to be estimated.

The primary objectives of this article are (1) to demonstrate by experiment that when the ideal spectrum to be estimated is truly unknown, the utility of parametric methods can be greatly enhanced when used jointly with methods based on Fourier transformation, and (2) to illustrate the relative strengths and weakness of the various methods and how to combine the strengths in a single hybrid method.

Because one of the most commonly referenced tutorial presentations on modern methods of spectral analysis is that of Kay and Marple [1], it was decided to adopt the same data set as that used in [1] to illustrate the various methods of spectral analysis described therein. This also enabled us to confirm that our computer programs were functioning properly. (They apparently were since they produced the same results as those presented in [1]). However, in order to demonstrate the effects of variability for each of the spectrum estimation methods studied, we needed more data than that presented in [1]. Therefore we also generated our own data set from a model described in Sec. II that is essentially the same as that in [1]. After it was confirmed that our programs produced the same results as those in [1] when operating on the data set given in [1], we then used our own data set for the experiments reported on here. This data set is published in [2].

The data model is especially appropriate for the purpose at hand because it does not give any particular method an unfair advantage by being particularly well matched to that method. That is, the model is not of the moving average (MA) type, the autoregressive (AR) type, or the ARMA type, nor is it

* There are also applications where an experiment can be repeated in such a way that the underlying ideal spectrum is the same for each repetition and, therefore, an ensemble of time segments of data can be obtained for averaging. We shall lump this type of problem together with the problem involving a time-indexed ensemble of space segments.

Received 2 September 1991; revised manuscript received 27 January 1992

International Journal of Imaging Systems and Technology, Vol. 4, 109–121 (1992)

CCC 0899–9457/92/020109–13$04.00

simply a sum of sine waves uniformly spaced in frequency. This seems to be a reasonable approach for studying the kind of spectral estimation problem where the ideal spectrum is truly unknown.

All methods studied are described within a unifying conceptual framework with unified notation in [2]. All but a few of the methods studied are also described more briefly in [1]. Extensive lists of references to prior work on these methods are given in [1] and [2]. Therefore, in order to minimize repetition, detailed descriptions of these methods are not given herein, and only the original sources are referenced. However, the readers' attention is drawn to the book by Marple [3], where computer programs for many of these methods are given.

II. EXPERIMENTAL RESULTS AND DISCUSSION

A summary and brief discussion of the results of an extensive experimental study are presented in this section. The methods compared include the periodogram with zero-padding, with and without data tapering, with and without frequency smoothing, and with and without sequential sine-wave removal and spectral-line reinsertion [2]; the minimum-leakage* (Capon) method [1, 2, 4]; the Yule–Walker [5, 6], Burg [7], and forward–backward least squares [8, 9] AR methods [1, 2]; the overdetermined-normal-equations method with biased and unbiased correlation estimates [2, 10]; and a singular-value-decomposition method [2, 11]. The relatively new signal-subspace methods are not included because they are not truly methods for spectral estimation. Rather they are primarily methods for estimating parameters of signals in noise.

The data was generated using a model that has the ideal spectrum and ideal autocorrelation shown in Fig. 1. Gardner [2] gives a table of 1024 time samples from this model, and graphs of data segments of lengths 64 and 256 from this data set are shown in Fig. 2. The data consists of three sine waves

* The most appealing interpretation of the Capon method is not as a model-fitting method, but rather as a data-adaptive Fourier analysis method in which leakage effects are directly minimized by linearly constrained least squares optimization of the set of bandpass filters used for Fourier analysis [2].

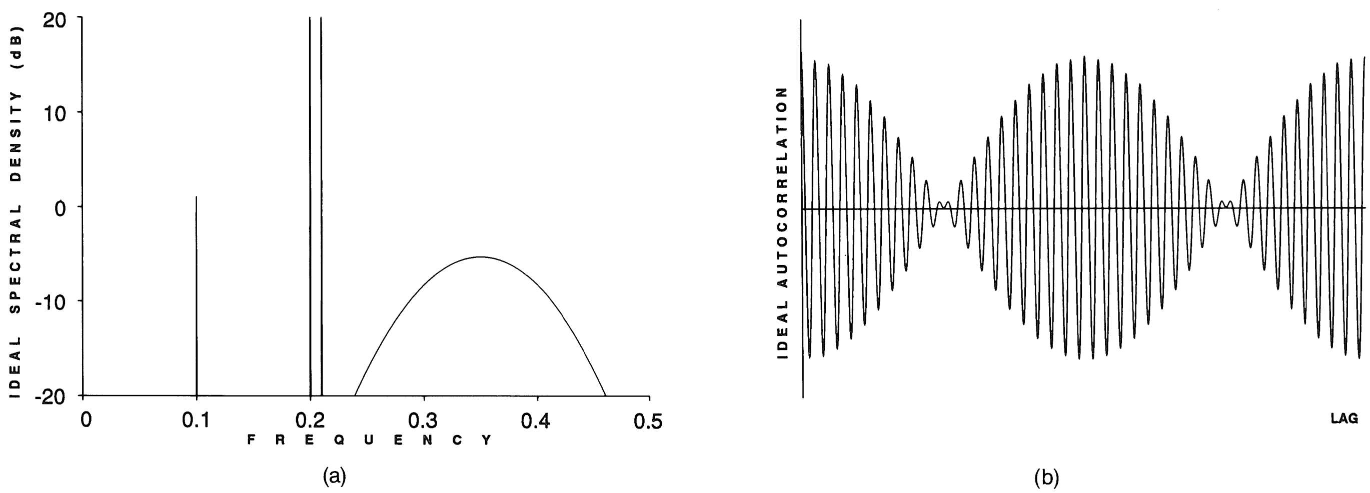


Figure 1. (a) Ideal spectrum for data used in experimental study (spectrum smoothed with rectangle window of width 1/256; highest peak is 21 dB). (b) Ideal autocorrelation for data used in experimental study (200 lag increments shown).

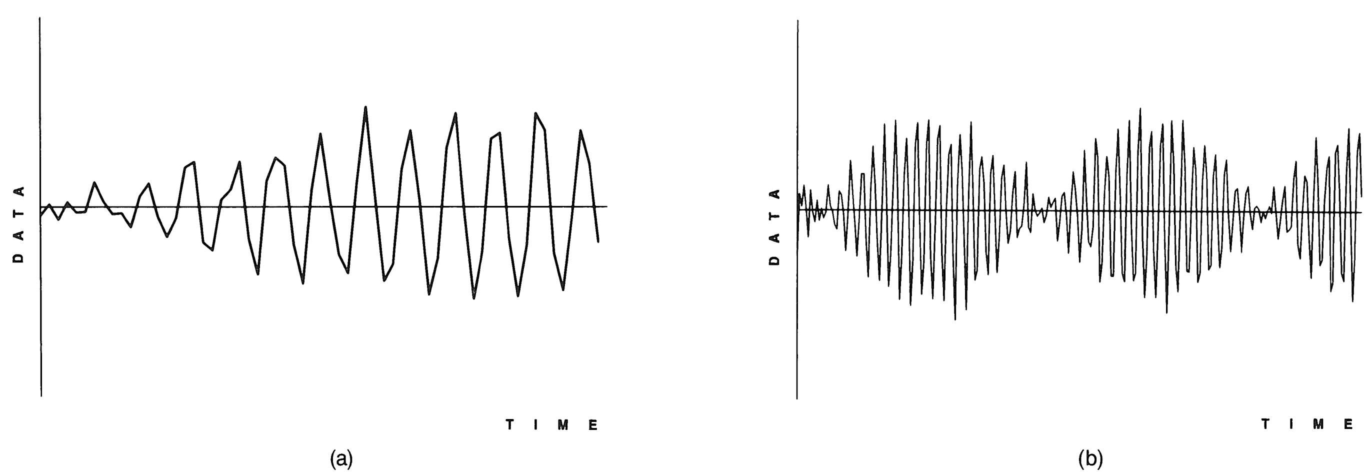


Figure 2. (a) Data segment of length 64 from Table 9-1 from Gardner [2] (starting point is 64). (b) Data segment of length 256 from Table 9-1 from Gardner [2] (starting point is 256).

in additive highly-colored Gaussian noise, with a band-limited Gaussian-shaped spectrum centered at $0.35/T_s$ Hz, where T_s is the time-sampling increment. The power of each of the two sine waves at frequencies $0.20/T_s$ Hz and $0.21/T_s$ Hz is -3 dB (relative to unity), and the phases are 106.2° and 41.5°, respectively. The power of the third sine wave at $0.10/T_s$ Hz is -23 dB and the phase is 32.6°. The power of the colored noise is -15 dB.

For each method and each set of parameter values considered, three samples of spectrum estimates are shown superimposed. Thus resolution, leakage, and reliability properties are all reflected in these graphical results. Although a set of only three samples is insufficient for a quantitative study of variability, it is adequate for the qualitative study reported here, especially because the three samples of the phase of the oscillatory envelope of the data [which is shown in Fig. 2(b)] are distributed throughout the period of oscillation in the set of three samples. All methods are applied to two lengths of data segments: a short length, $N = 64$, and a long length, $N = 256$. The three statistical samples of data are obtained from Table 9-1 in [2] using staring points of 64, 128, and 192 for $N = 64$, and 256, 512, and 768 for $N = 256$. It can be seen from the segment of length 256 in Fig. 2(b) that the three segments of length 64 that were selected provide a uniform spread of starting phases for the strong beat-frequency envelope that modulates the data. None of the spectrum estimates produced by the various methods were normalized in any way. Thus the power-level estimation capability of each method is also illustrated in the graphical results.

For all the parametric methods considered that require use of a model-order-determining algorithm, the three methods FPE, AIC, and CAT were studied [1, 2, 12–14]. However, the performances of these methods were usually unacceptable. Trial-and-error experimentation revealed [by comparison with the ideal spectrum in Fig. 1(a)] that the order $M = 16$ was typically the best for $N = 64$, and $16 < M < 32$ (usually $M \cong 32$) was typically the best for $N = 256$. Consequently, results are presented for primarily $M = 16$ and $M = 32$. But a few samples of spectrum estimates obtained with estimated model orders are shown to illustrate their inferiority.

A. Fourier Transformation Methods. The first spectrum estimate considered is the periodogram without data tapering, time averaging, or frequency smoothing. This raw periodo-

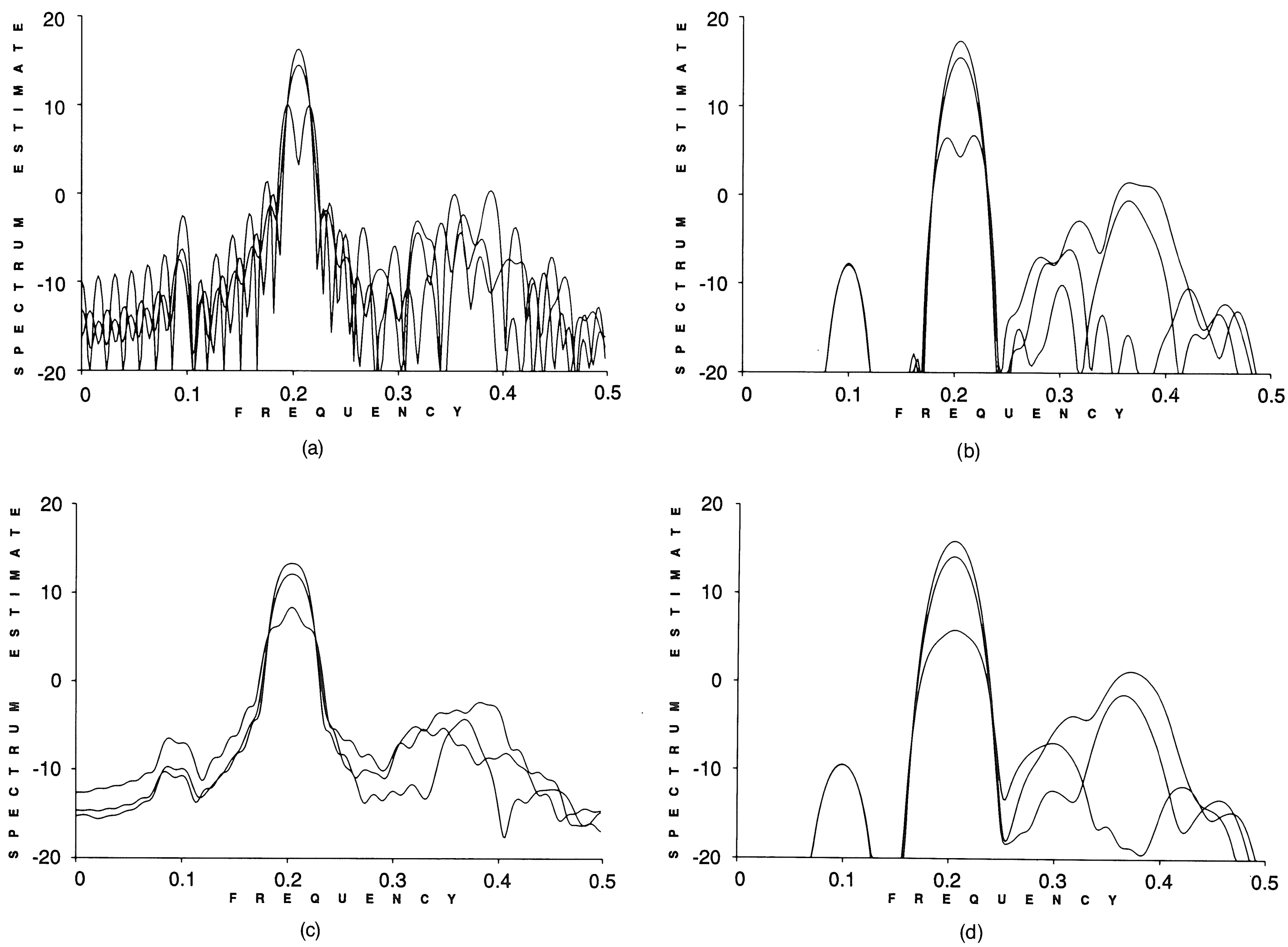


Figure 3. (a) Three periodograms for $N = 64$ data points each with $K = 8$ zero-padding factor and no data tapering. (b) Same as (a), but with raised-cosine data tapering. (c) Same as (a), but with frequency smoothing ($P = 2$). (d) Same as (b), but with frequency smoothing ($P = 2$).

gram was obtained using a DFT algorithm with $K=8$ zero-padding factor [i.e., $(K-1)N$ zeros are added so that the total number of points transformed is KN] and is shown in Figs. 3(a) and 4(a) for the two data-segment lengths of $N=64$ and $N=256$, respectively. It can be seen that the two closely spaced spectral lines are not reliably resolved for $N=64$, but are for $N=256$. These periodograms were then modified by use of a raised-cosine data-tapering window (with height of 2), and the results are shown in Figs. (3b) and (4b). All these periodograms were then frequency smoothed using smoothing parameter $P=2$ (PK DFT bins were averaged together), and the results are shown in Fig. 3(c), 3(d), 4(c), and 4(d). It can be seen that data tapering greatly reduces spectral leakage and that frequency smoothing improves reliability (although the amount of improvement is small for $P=2$), and that both techniques degrade resolution.

Since Figs. 3(a) and 3(b) suggested to us (pretending ignorance of the true model) that there were spectral lines in the vicinities of $f=0.1/T_s$ and $f=0.2/T_s$, the next method we studied removed sine waves from the data and then the preceding methods, corresponding to Figs. 3 and 4, were repeated, except that $P=4$ for $N=64$ and $P=8$ for $N=256$ were used instead of $P=2$ to obtain more reliability. The sine-wave removal was accomplished by using least squares estimates of the frequency (in the vicinity of the peaks in the raw periodogram), amplitude, and phase of each sine wave. These estimates are easily obtained *directly* from the DFT of the data [2]. Since a substantial spectral peak remained in the periodogram after subtraction from the data of a sine wave with frequency near $0.2/T_s$, the estimation procedure was repeated in this vicinity and a second sine wave was removed. Then the procedure was repeated in the vicinity of $f=0.1/T_s$, since a spectral peak was clearly evident in the periodogram with the leakage effects of the two peaks near $f=0.2/T_s$ largely removed. With the three estimated sine waves subtracted from the data, the periodogram was calculated for the residual time series. Finally, spectral lines with one DFT bin-width and with magnitude determined by the least squares estimates of the sine-wave amplitudes were added to the periodogram. The resultant spectrum estimates, which are shown in Fig. 5 for $N=64$ and Fig. 6 for $N=256$, exhibited greatly reduced spectral leakage as would be expected. Although the variability of the pair of spectral lines in the vicinity of $f=0.2/T_s$ is substantial for $N=64$, it is very small

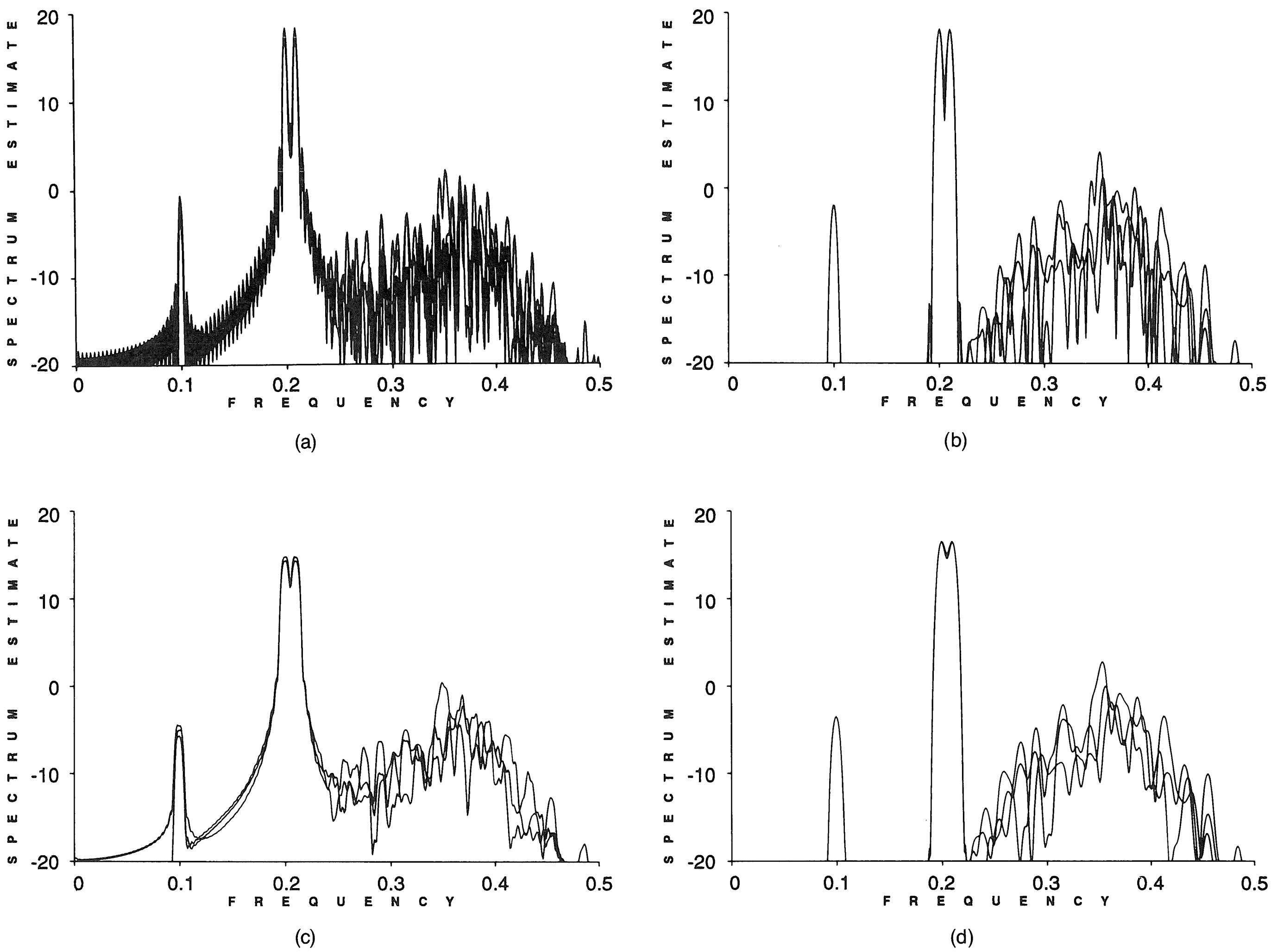


Figure 4. (a) Three periodograms for $N=256$ data points each with $K=8$ zero-padding factor and no data tapering. (b) Same as (a), but with raised-cosine data tapering. (c) Same as (a), but with frequency smoothing ($P=2$). (d) Same as (b), but with frequency smoothing ($P=2$).

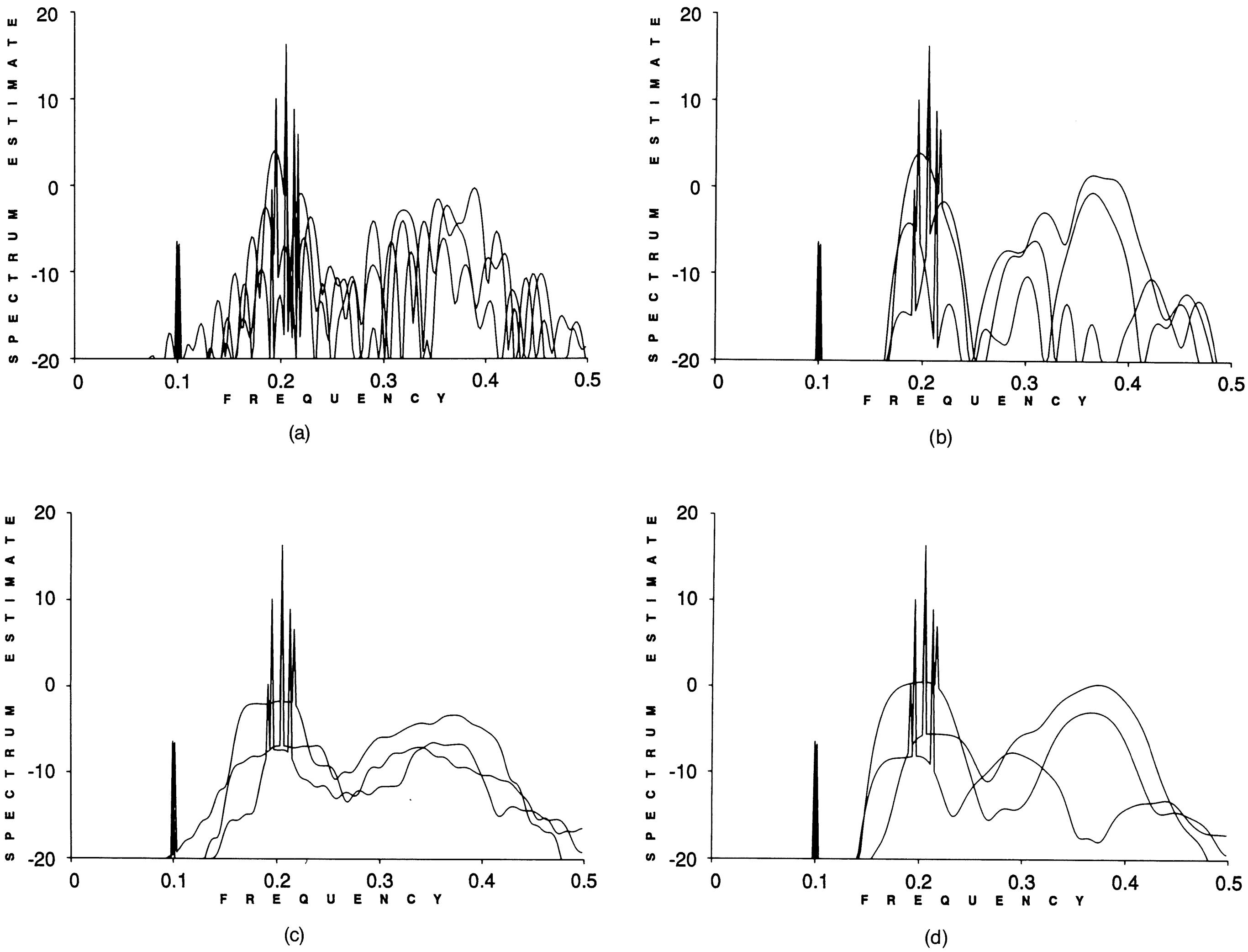


Figure 5. (a) Three periodograms with sine-wave removal and spectral-line reinsertion for $N = 64$ data points each, and with $K = 8$ zero-padding factor and no data tapering. (b) Same as (a), but with raised-cosine data tapering. (c) Same as (a), but with frequency smoothing ($P = 4$). (d) Same as (b), but with frequency smoothing ($P = 4$).

for $N = 256$. Also, the variability of the single weak spectral line at $f = 0.1/T_s$ is quite small, even for $N = 64$.

B. Minimum-Leakage Method. The next spectrum estimates considered are those provided by the minimum-leakage (ML) method [1, 2, 4]. Four versions of this method were studied. These include the original method by Capon and a modified method that normalizes the power estimate by a data-adaptive (rather than fixed) resolution bandwidth to obtain a power spectral density estimate [2, 15]. Each of these methods has two versions corresponding to use of either the covariance-type data-correlation matrix (which prevents the adaptive leakage-minimizing filter from running off the ends of the finite segment of data) or the autocorrelation-type data correlation matrix (which allows the filter to run completely off both ends of the data) [2]. The best results were obtained by the unmodified method using the covariance-type data-correlation matrix. Results for this method only are presented. These results are shown in Fig. 7. For $N = 64$, it can be seen that both resolution and reliability performance for $M = 24$ is definitely better than that provided by the periodogram both with and without sine-wave removal. However, this superior performance for $N = 64$ is obtained with the unfair advantage of knowing that the filter order $M = 24$ yields the best results. If the ideal spectrum were not known in advance, then this choice for M would not necessarily have been made. The ML spectrum estimates for $M < 20$ were too smooth and those for $M > 24$ exhibited spurious peaks and excess variability. This is illustrated in Fig. 14. For $N = 256$, the best ML estimate, which is shown in Fig. 7(d), does not approximate the ideal spectrum as accurately as does the Fourier transformation method (with sine-wave removal) shown in Figs. 6(c) and 6(d).

C. Yule–Walker, Burg, and Forward–Backward Least Squares AR Methods. The least squares autoregressive spectrum estimates provided by the methods of Yule–Walker (YW), or maximum entropy, Burg, and forward–backward (FB) linear prediction [1, 2, 5–9] are presented in Figs. 8, 9, and 10 for $N = 64$ with $M = 16$ and for $N = 256$ with $M = 16$ and $M = 32$. It can be seen that the YW method is consistent-

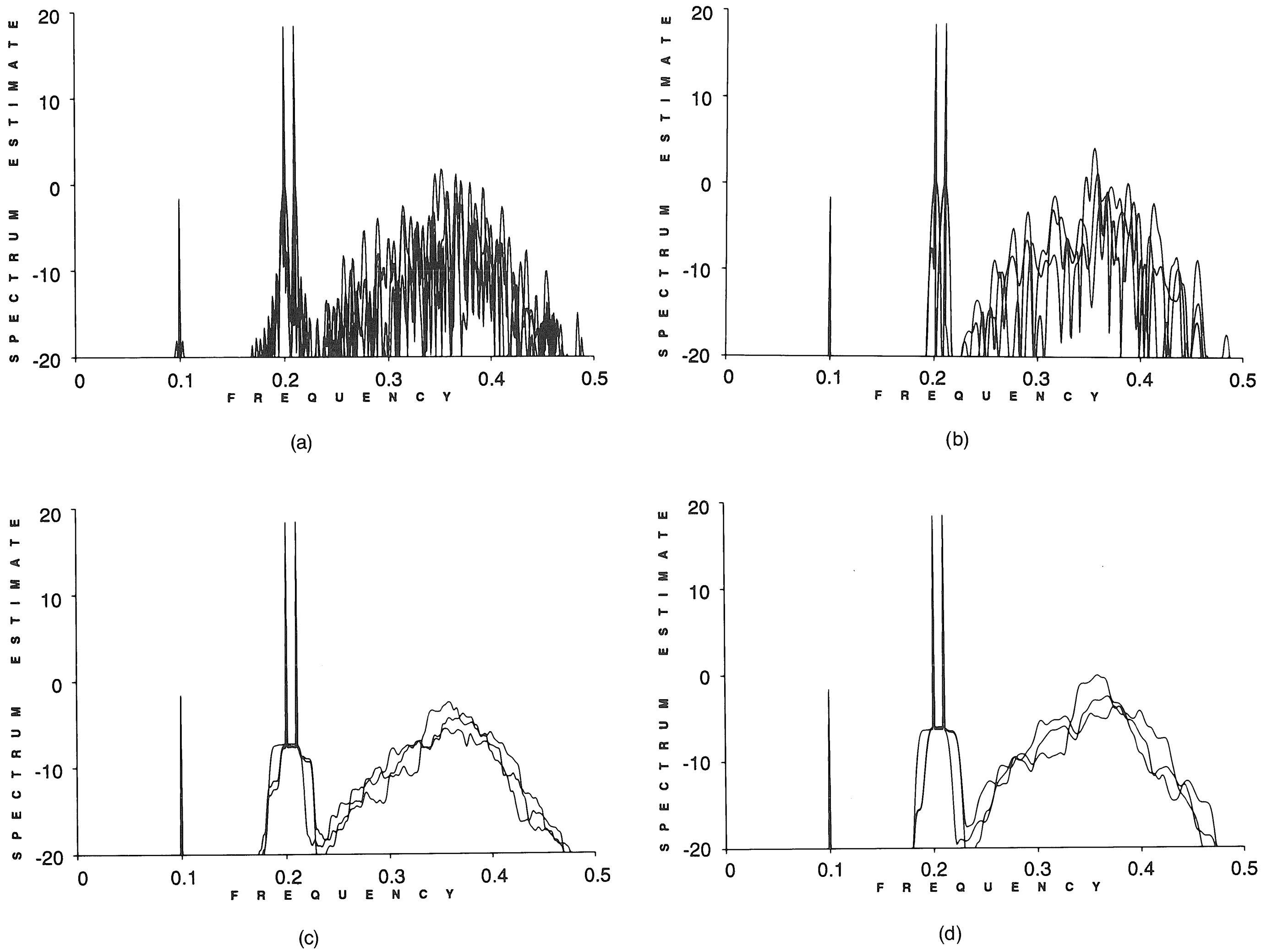


Figure 6. (a) Three periodograms with sine-wave removal and spectral-line reinsertion for $N = 256$ data points each, and with $K = 8$ zero-padding factor and no data tapering. (b) Same as (a), but with raised-cosine data tapering. (c) Same as (a), but with frequency smoothing ($P = 8$). (d) Same as (b), but with frequency smoothing ($P = 8$).

ly inferior to the ML method shown in Fig. 7,* except in the accuracy of the overall level of the spectrum estimate. For $N = 64$, it barely resolves the peak at $f = 0.1/T_s$ and does not resolve the pair of peaks near $f = 0.2/T_s$, nor does it produce as accurate an estimate of the continuous part of the spectrum centered at $f = 0.35/T_s$. For $N = 256$, both resolution and variability are poorer than they are for both the ML method [Fig. 7(d)] and the Fourier transformation method [Figs. 6(c) and 6(d)]. The Burg method is also consistently inferior to the ML method (except for the accuracy of the overall level and the relative heights of the weak and strong spectral lines) because of its higher variability. However, the Burg method is superior to the Fourier transformation method [Figs. 6(c) and 6(d)] for $N = 64$. The FB method clearly does the best job of resolving the three spectral lines for $N = 64$, and it is clearly the best in terms of low variability in the locations of the spectral lines. However, the Burg method is superior in terms of the accuracy of estimating the relative heights of the weak and strong spectral lines. For $N = 256$, the same conclusions apply in comparing the Burg and FB methods. However, these two methods are inferior to the Fourier transformation method [and the ML method in Fig. 7(d) except for the accuracy of the overall level] for $N = 256$. Comparison of Figs. 9(b) and 9(c) and 10(b) and 10(c) with Figs 6(c) and 6(d) reveals that the Fourier transformation methods provides considerably more accurate estimates of the ideal spectrum for the relatively long data segments ($N = 256$).

The superiority of the Burg and FB methods relative to the Fourier transformation method for $N = 64$ must be tempered by the fact that these two parametric methods were given the unfair advantage of having the best model orders specified. When this advantage is removed by using the estimated orders produced by the model-order-determining methods AIC, FPE, and CAT, both parametric methods become inferior to

* This result does not support the commonly held belief that in general the YW method provides higher resolution than the ML method [1]. But previous studies focused on the version of the ML method that uses the autocorrelation-type data-correlation matrix, which yields lower resolution than the version that uses the covariance-type correlation matrix.

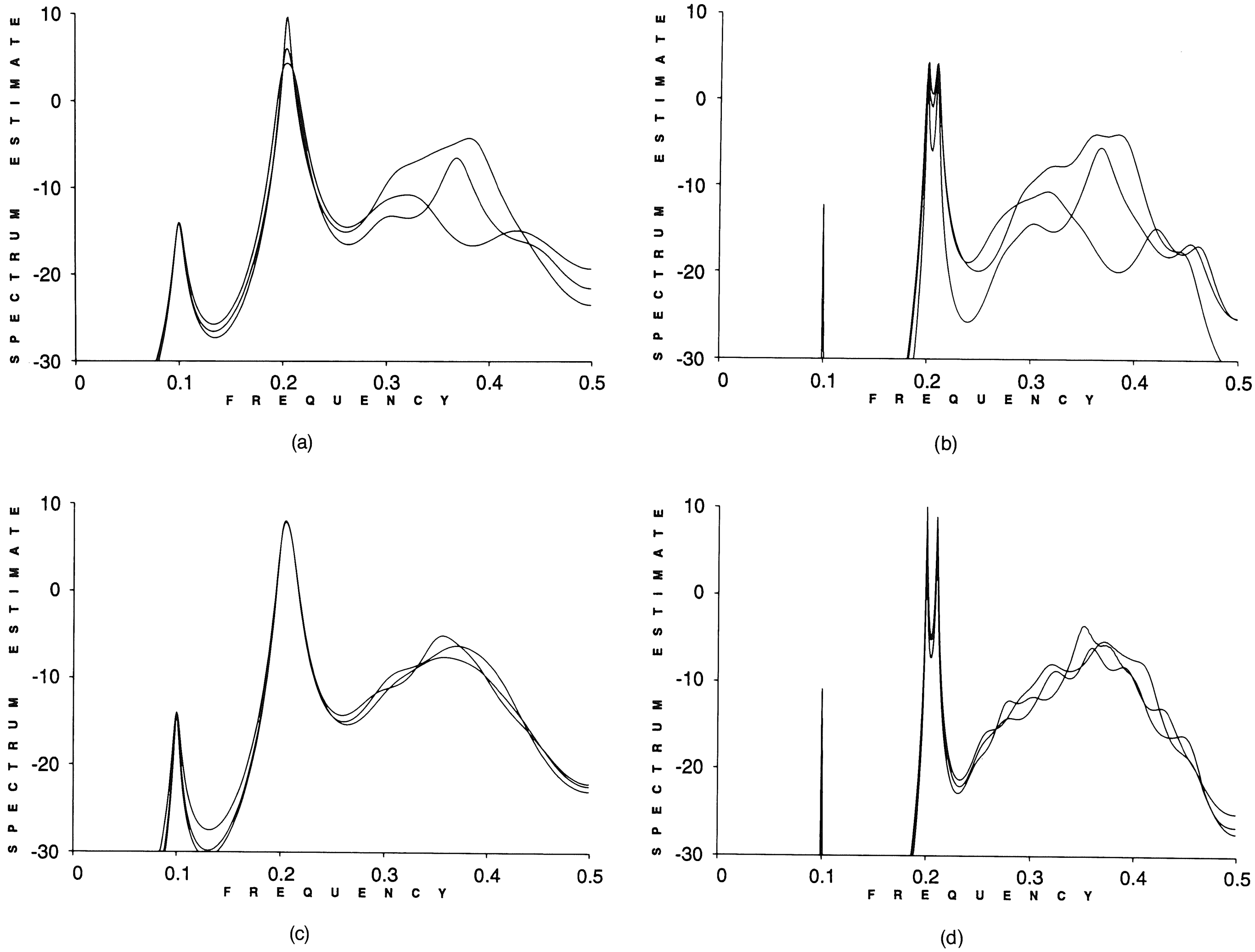


Figure 7. (a) Three minimum-leakage spectrum estimates with $N = 64$, $M = 16$. (b) Same as (a), but with $M = 24$. (c) Same as (a), but with $N = 256$, $M = 16$. (d) Same as (a), but with $N = 256$, $M = 32$.

the Fourier transformation method because the estimated order is too high for $N = 64$, resulting in spurious peaks and high variability.

Although all three order-determining methods, FPE, AIC, and CAT, produced comparable order estimates for each spectrum estimation method, these estimates were too low for the YW method ($5 < M < 11$ for $N = 64$ and $12 < N < 22$ for $N = 256$) and too high for the Burg method ($M > 48$ for $N = 64$ and $M > 84$ for $N = 256$) and for the FB method, except for $N = 256$ ($M \cong 24$ for $N = 64$, $M \cong 32$ for $N = 256$). Unfortunately, in practice, where there is only one data segment to analyze, there is apparently no way to determine from these parametric methods alone which peaks are correct and which are spurious. Sample spectrum estimates with estimated orders are shown in Figs. 11–14 to illustrate their inferiority.

D. Overdetermined-Normal-Equations AR Method. The spectrum estimates provided by the overdetermined-normal-equations (ODNE) AR method [2, 10] are shown in Figs. 15 and 16 for the parameter Q, which specifies the number of normal equations, given by $Q = 48$. The cases included are $N = 64$ with $M = 16$, and $N = 256$ with $M = 16$ and $M = 32$. The results shown in Fig. 15 were obtained using the biased correlation estimates from the autocorrelation method of least squares, whereas the results shown in Fig. 16 were obtained using the corresponding unbiased estimates [1, 2]. As expected, the method that uses the unbiased estimates provides better resolution but the increase in variability is surprisingly small. For $N = 64$, the overall performance of the unbiased ODNE method is slightly inferior to that of the Burg method and is strongly inferior to that of the ML and FB methods. For $N = 256$, the performance of the unbiased ODNE method is comparable to that of the Burg and FB methods and is therefore inferior to that of both the ML method [Fig. 7(d)] and the Fourier transformation method [Figs. 6(c) and 6(d)]. Also, the accuracy of the overall level of the spectrum estimate provided by both ODNE methods is poor.

E. Singular-Value-Decomposition Method. The spectrum estimates provided by the singular-value-decomposition (SVD) method [2, 11] are shown in Fig. 17. Since the data is not from an AR model and does not consist of simply strong sine waves in a white-noise background, the eigenvalues do

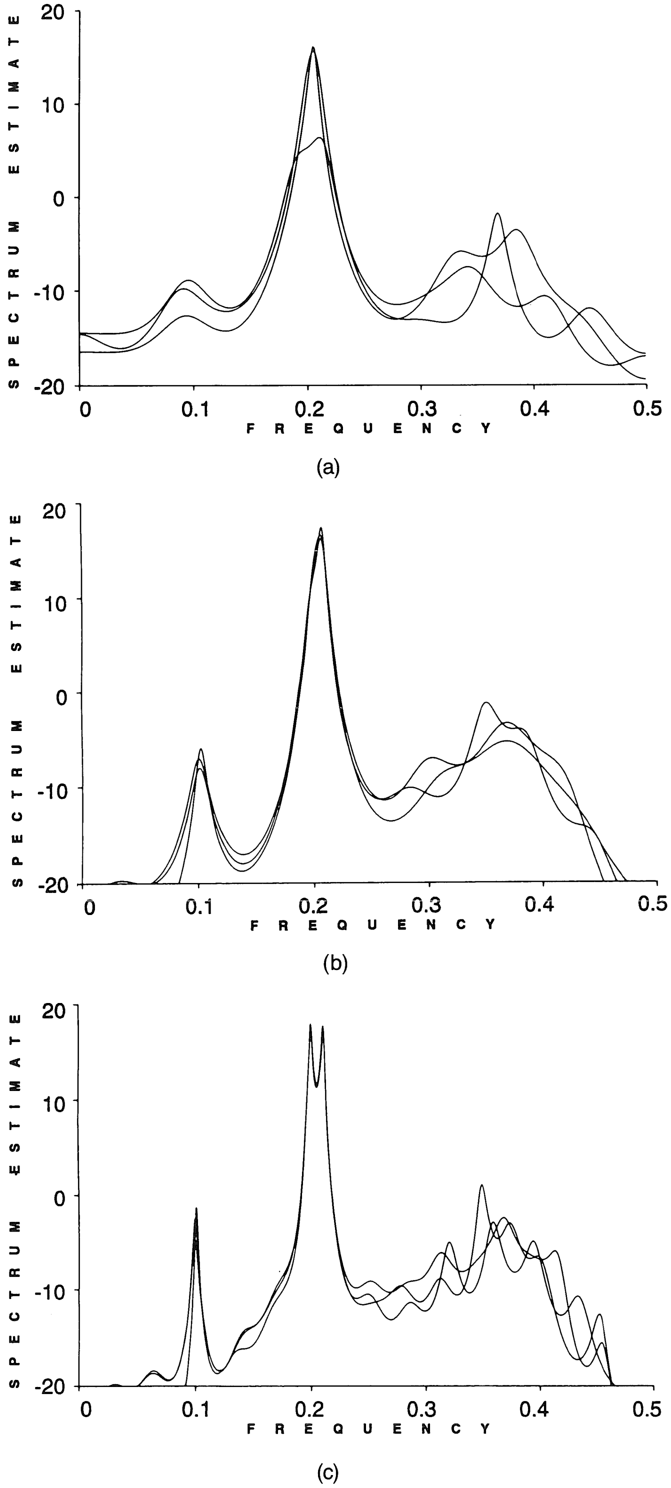


Figure 8. (a) Three Yule–Walker AR spectrum estimates with $N = 64$, $M = 16$. (b) Same as (a), but with $N = 256$, $M = 16$. (c) Same as (a), but with $N = 256$, $M = 32$.

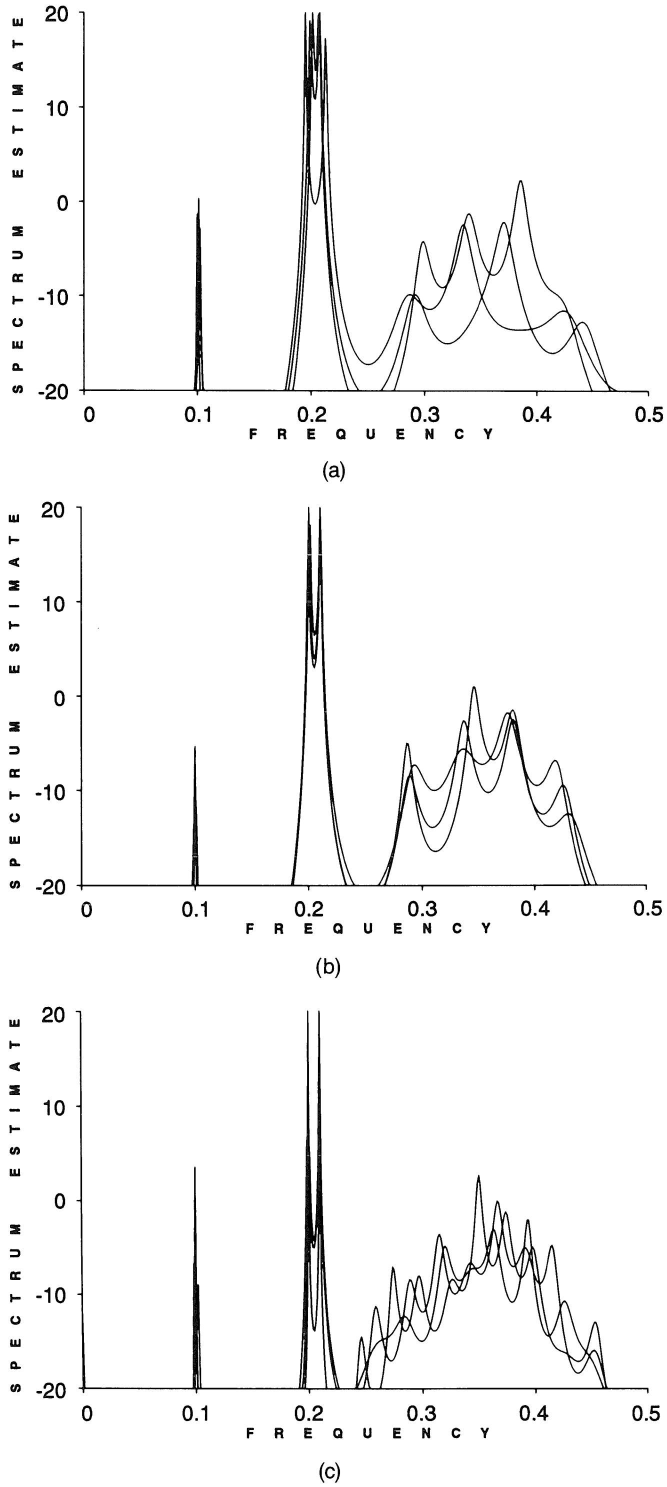
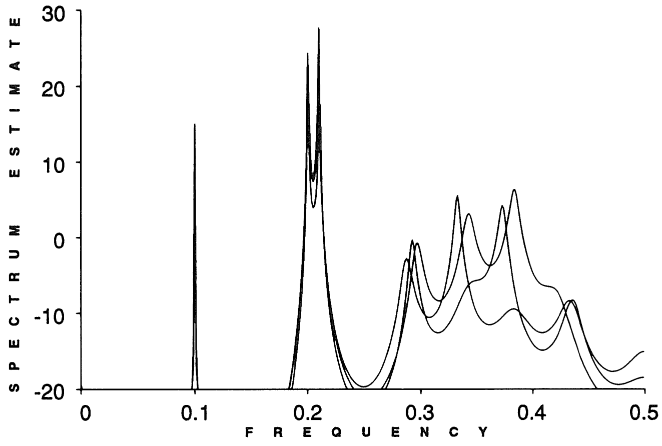
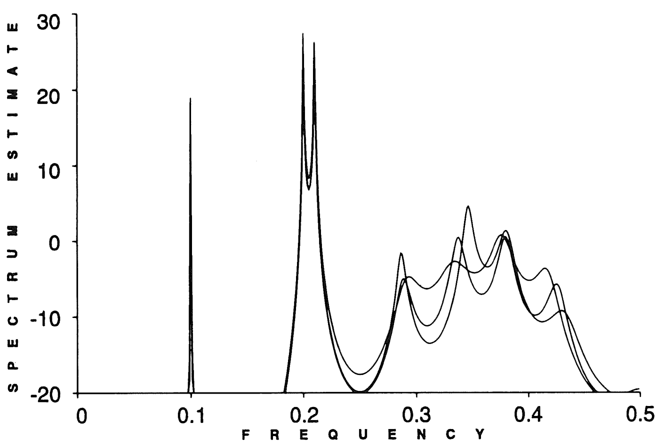


Figure 9. (a) Three Burg AR spectrum estimates with $N = 64$, $M = 16$ (highest peak is 33 dB). (b) Same as (a), but with $N = 256$, $M = 16$ (highest peak = 34 dB). (c) Same as (a), but with $N = 256$, $M = 32$ (highest peak = 35 dB).

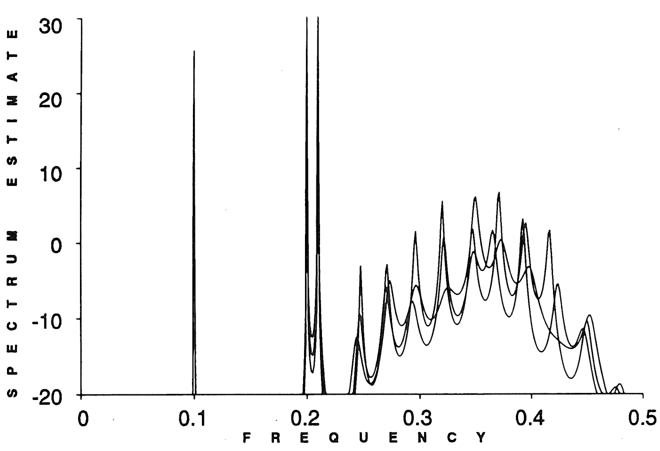
not partition into one set of relatively large values corresponding to spectral features of interest and a remaining set of negligible values or a set of small values corresponding to a flat spectrum. As a consequence, the procedure of using a test on the eigenvalue ratio ρ (sum of M largest eigenvalues divided by sum of all Q eigenvalues) to determine a best rank M [2] cannot be expected to perform well, particularly since there will be no obvious way to set the threshold in such a test. Therefore M was just chosen to be $M = 16$ and $M = 32$, and the corresponding ratio ρ was calculated as an aside. For



(a)



(b)



(c)

Figure 10. (a) Three forward–backward linear-prediction AR spectrum estimates with $N=64$, $M=16$. (b) Same as (a), but with $N=256$, $M=16$. (c) Same as (a), but with $N=256$, $M=32$ (highest peak is 39 dB).

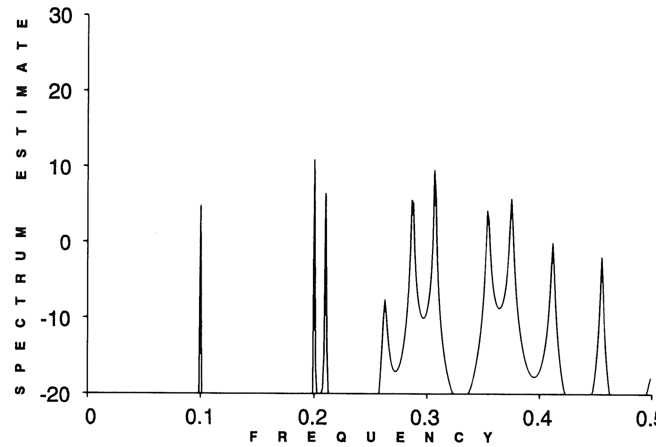
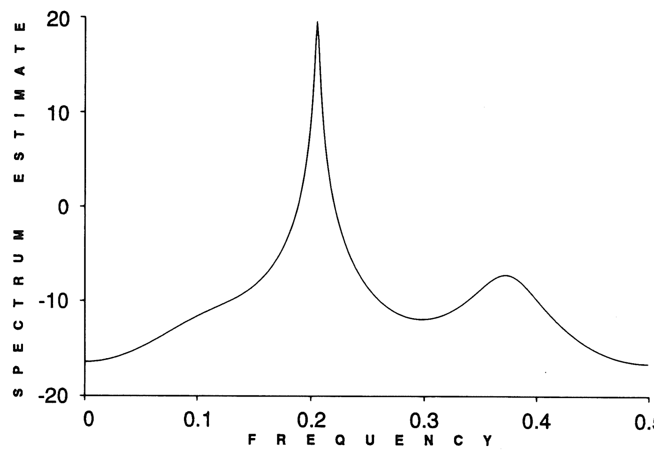
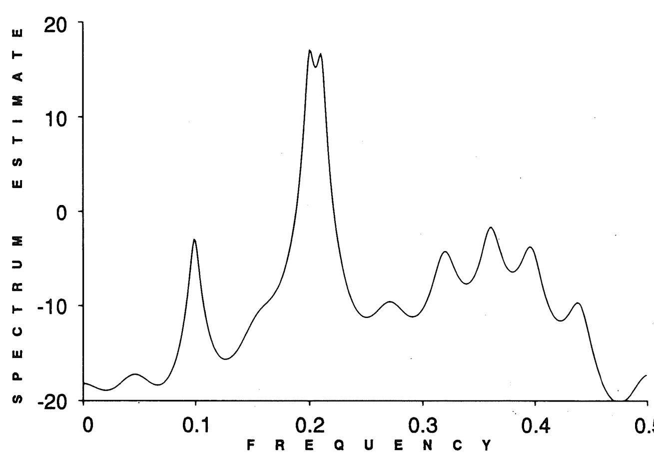


Figure 11. Sample of poor spectrum estimate obtained with the forward–backward linear-prediction method with too high an order ($M=24$) for $N=64$.

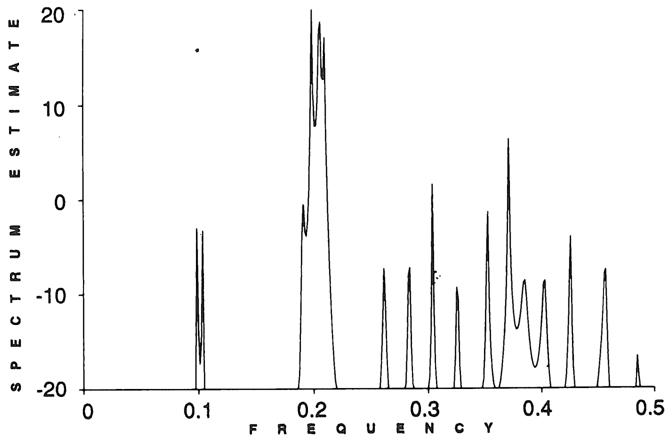


(a)

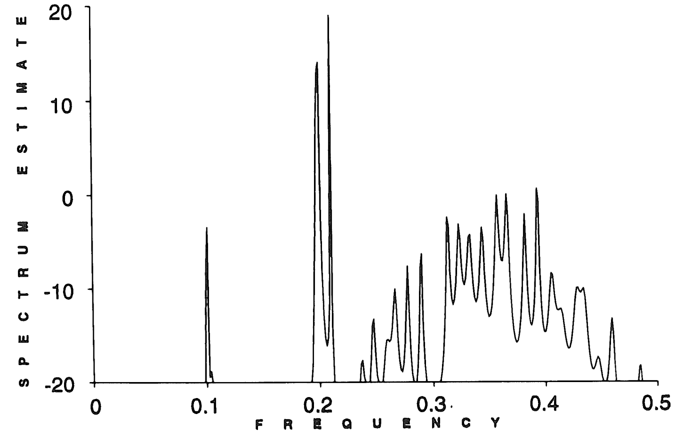


(b)

Figure 12. (a) Sample of poor spectrum estimate obtained with the Yule–Walker AR method with too low an order ($M=7$) for $N=64$. (b) Same as (a), but with $M=32$, $N=256$.

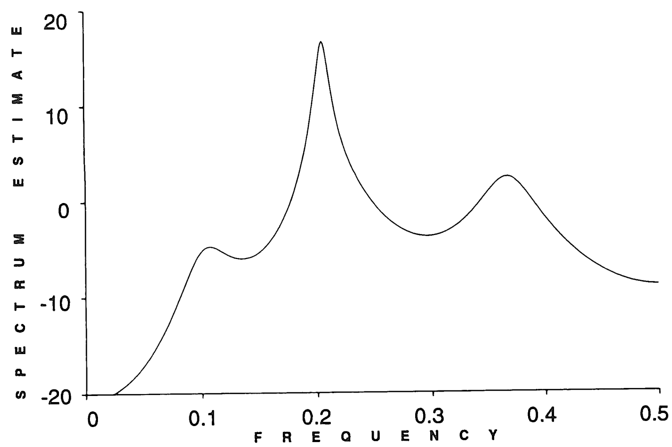


(a)

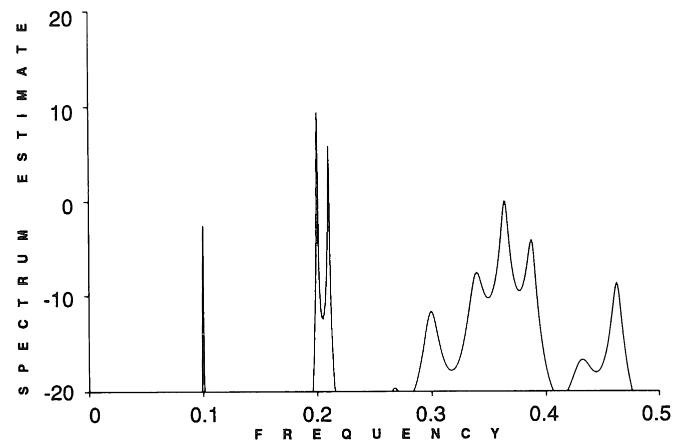


(b)

Figure 13. (a) Sample of poor spectrum estimate obtained with the Burg AR method with too high an order ($M = 48$) for $N = 64$. (b) Same as (a), but with $M = 90$, $N = 256$.



(a)



(b)

Figure 14. (a) Sample of poor spectrum estimate obtained with the minimum-leakage method with too low an order ($M = 12$) for $N = 64$. (b) Same as (a), but with too high an order ($M = 32$).

$M = 16$, $\rho \cong 0.96$ for $N = 64$ and $\rho \cong 0.97$ for $N = 256$. For $M = 32$, $\rho \cong 0.99$ for $N = 64$ and $N = 256$. It can be seen from Fig. 17 that for both $N = 64$ and $N = 256$ the performance is inferior to that of all other parametric methods as well as the ML method and the Fourier transformation methods [Figs. 5(c), 5(d), 6(c) and 6(d)]. As with all parametric methods studied, experimentation with the order M showed that the best results were obtained with $M \cong 16$ for $N = 64$ and with $M \cong 32$ for $N = 256$.

III. CONCLUSIONS

The results and discussion in Sec. II illustrate that parametric methods have a crucial weakness, and this is model-order selection. When knowledge of the ideal spectrum cannot be used to guide model-order selection, the spectrum estimates produced by parametric methods with short data segments are likely to yield either inadequate resolution (order too low) or spurious peaks and/or high variability (order too high), and without further experimentation the user who does not have knowledge of the ideal spectrum cannot know if either of these situations has occurred. If the application is such that processing time and complexity are not constrained, then the best approach would be to use a combination of direct Fourier transformation (with various amounts of data tapering and spectral smoothing, and possibly sine-wave removal) and model fitting with various orders, and possibly even various models, in an experimentation mode, until confidence is gained that none of the sharp peaks retained are likely to be spurious, and that smooth portions and deep valleys are revealed as accurately as possible. To illustrate the quality of spectrum estimates obtainable with such hybrid methods, the FB method was used to detect spectral lines and estimate their frequencies, and the DFT (with $K = 8$ zero-padding factor) was then used to estimate the amplitudes and phases of the three corresponding sine waves; after subtraction of these estimated sine waves from the data, the 32-point smoothed periodogram ($P = 4$ and $K = 8$) was computed. The results are shown in Fig. 18. This is clearly the most accurate of all spectrum estimates considered for short data segments ($N = 64$).

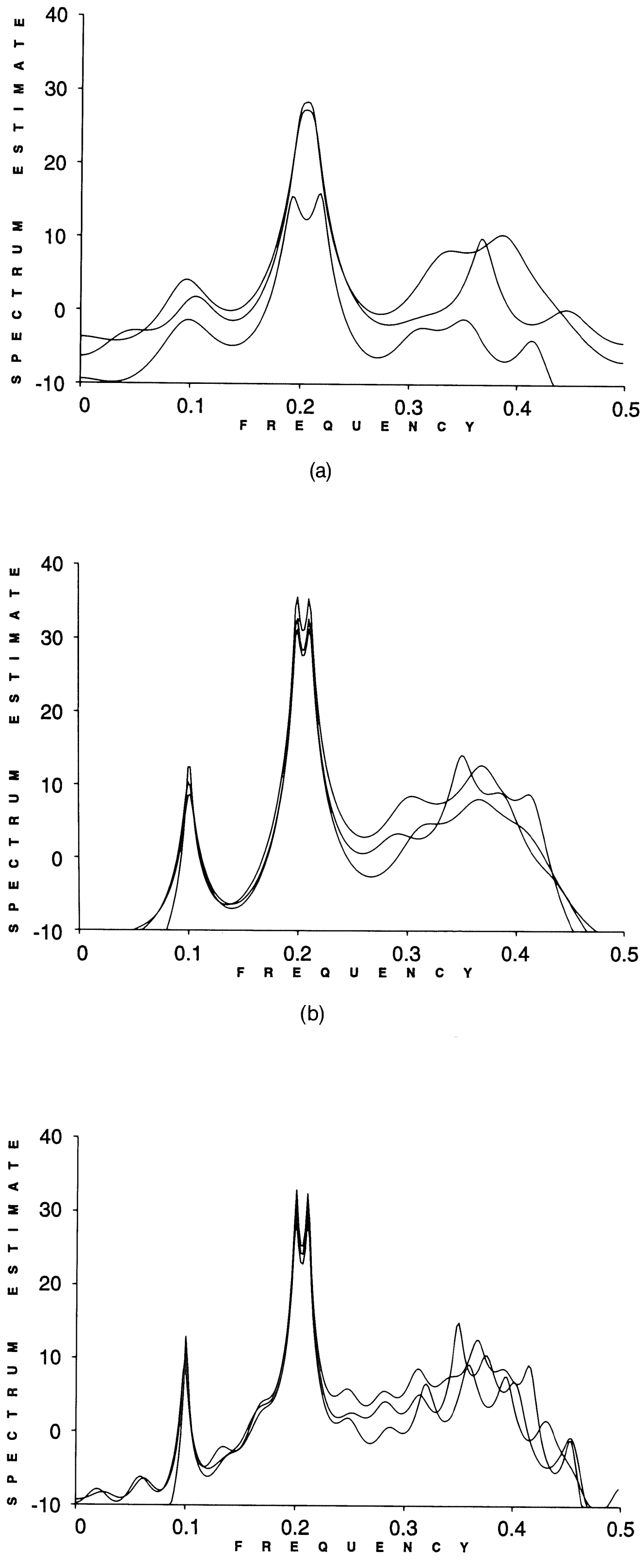


Figure 15. (a) Three overdetermined-normal-equations AR spectrum estimates obtained using biased autocorrelation estimates and $Q = 48$ normal equations with $N = 64$, $M = 16$. (b) Same as (a), but with $N = 256$, $M = 16$. (c) Same as (a), but with $N = 256$, $M = 32$.

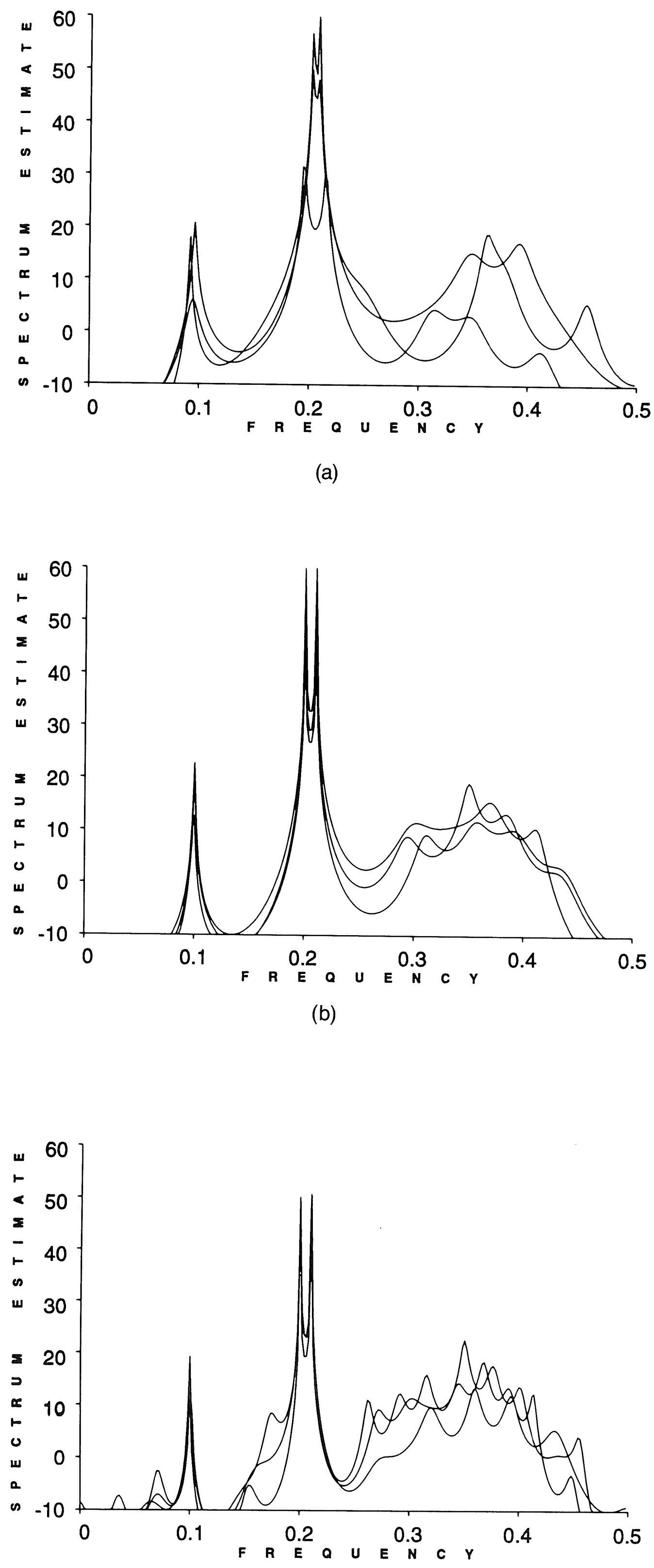


Figure 16. (a) Three over-determined-normal-equation AR spectrum estimates obtained using unbiased autocorrelation estimates and $Q = 48$ normal equations with $N = 64$, $M = 16$. (b) Same as (a), but with $N = 256$, $M = 16$ (highest peak is 64 dB). (c) Same as (a), but with $N = 256$, $M = 32$.

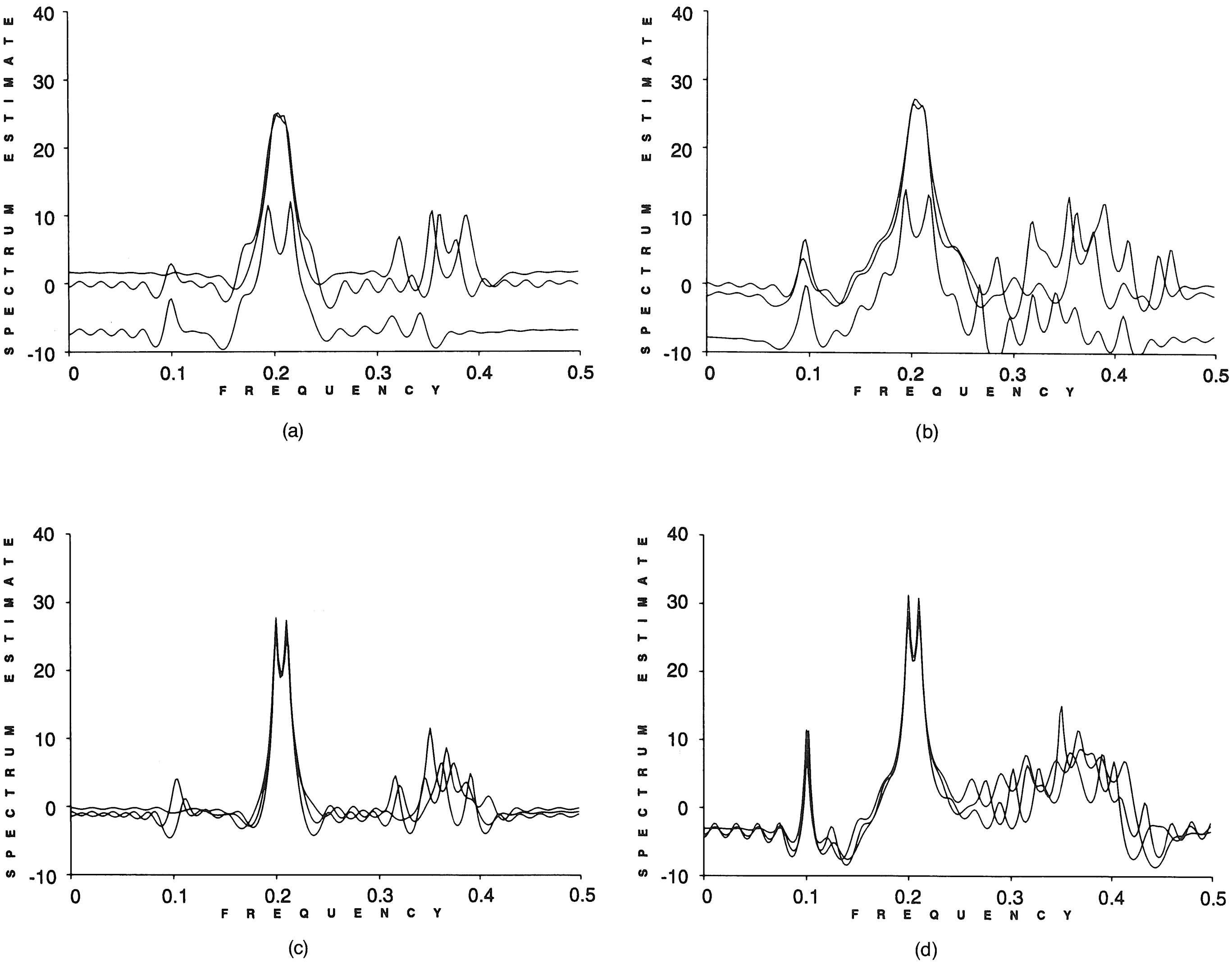


Figure 17. Three singular-value-decomposition spectrum estimates obtained using $Q = 48$, $N = 64$, $M = 16$. (b) Same as (a), but with $M = 32$. (*c*) Same as (a), but with $N = 256$, $M = 16$. (d) Same as (a), but with $N = 256$, $M = 32$.

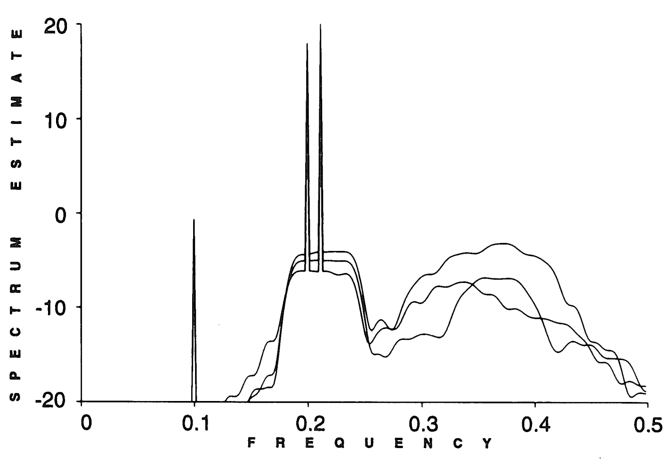


Figure 18. Three spectrum estimates obtained using the hybrid method described in Sec. III.

ACKNOWLEDGMENT

The author expresses his gratitude to Messrs. B. G. Agee, W. A. Brown, C. K. Chen, J. H. Reed, and R. S. Roberts for their assistance in this experiment, which was conducted in 1986.

REFERENCES

1. S. M. Kay and S. L. Marple, Jr. "Spectrum analysis—A modern perspective," *Proc. IEEE* **69**, 1380–1419 (1981). Also in *Modern Spectrum Analysis, II*, edited by S. B. Kesler (IEEE Press, NY, 1986).
2. W. A. Gardner, *Statistical Spectral Analysis: A Nonprobabilistic Theory* (Prentice-Hall, Englewood Cliffs, NJ, 1987).
3. S. L. Marple, *Digital Spectral Analysis with Applications* (Prentice-Hall, Englewood Cliffs, NJ, 1987).
4. J. Capon, "High-resolution frequency–wave number spectrum analysis," *Proc. IEEE* **57**, 1408–1418 (1969).
5. G. U. Yule, "Why do we sometimes get nonsense correlations

between time-series?—A study in sampling and the nature of time-series," *J. R. Stat. Soc.* **89**, 1–69 (1926).

6. G. Walker, "On periodicity in series of related terms," *Proc. R. Soc.* **131**, 518–532 (1931).
7. J. P. Burg, "A new analysis technique for time-series data," *NATO Advanced Study Institute on Signal Processing with Emphasis on Underwater Acoustics*. (Enschede, The Netherlands, August 12–23, 1968).
8. T. J. Ulrych and R. W. Clayton, "Time-series modeling and maximum entropy," *Physics Earth Planet. Inter.* **12,** 188–200 (1976).
9. A. H. Nuttall, "Spectral analysis of a univariate process with bad data points, via maximum entropy and linear predictive techniques," Tech. Report Tr-5303, Naval Underwater Systems Center, New London, CT, March 26, 1976 (unpublished).
10. J. A. Cadzow, "Spectral estimation: An overdetermined rational model equation approach," *Proc IEEE* **70**, 907–939 (1982).
11. J. A. Cadzow, B. Baseghi, and T. Hsu, "Singular-value decomposition approach to time-series modeling," *IEE Proc.. F* **130**, 202–209 (1983).
12. H. Akaike, "Power spectrum estimation through autoregression model fitting," *Ann. Inst. Stat. Math.* **21**, 407–419 (1969).
13. H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control* **AC-19**, 716–723 (1974).
14. E. Parzen, "Some recent advances in time-series modeling," *IEEE Trans. Auto. Control* **AC-19**, 723–730 (1974).
15. M. A. Lagunas-Hernandez and A. Gasull-Llampallas, "An improved maximum likelihood method for power spectral density estimation," *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-32**, 170–173 (1984).