UNIVERSITY OF MINNESOTA
ST. ANTHONY FALLS HYDRAULIC LABORATORY

Project Report No. 363

Time Series Analysis with Application to Velocity Fluctuation Data Measured in a Wind Tunnel Atmospheric Boundary Layer

by

Cesar Farell

and

Arun K.S. Iyengar

Prepared for

NATIONAL SCIENCE FOUNDATION
and
STRIPPER WELL OIL OVERCHARGE GRANT
DEPARTMENT OF ADMINISTRATION
STATE OF MINNESOTA

October 1994
Minneapolis, Minnesota

UNIVERSITY OF MINNESOTA
St. Anthony Falls Hydraulic Laboratory

Project Report No. 363

Time Series Analysis with Application to Velocity Fluctuation Data Measured in a Wind Tunnel Atmospheric Boundary Layer

by

Cesar Farell

and

Arun K.S. Iyengar

Prepared for

NATIONAL SCIENCE FOUNDATION
and
STRIPPER WELL OIL OVERCHARGE GRANT,
DEPARTMENT OF ADMINISTRATION
STATE OF MINNESOTA

October 1994

ABSTRACT

The techniques used to estimate the statistical parameters describing the turbulent velocity fluctuations in an urban atmospheric boundary layer simulated in a long test section wind tunnel are described. An extensive review of the theory underlying the selection of autocovariance and spectral density function estimators and of the equations needed for the digital evaluation of sample spectra using Fast Fourier Transforms is presented. A computer program for evaluation of sample spectra and autocovariances, developed on the basis of these equations and including the determination of confidence intervals for the spectral estimates, is given. The selection of series and subseries lengths and sampling intervals is illustrated through application of the program to autoregressive processes and to the atmospheric boundary layer velocity fluctuation data. The determination of integral length scales and the scales in the inertial subrange of the spectrum is discussed in particular for the velocity fluctuation data.

ACKNOWLEDGMENTS

The support of National Science Foundation under Grants NSF/ECE-8410532, NSF/ECE-6613554, and NSF/INT-8814085, and of the State of Minnesota under a Stripper Well Oil Overcharge Grant, is gratefully acknowledged.

TABLE OF CONTENTS

NOMENCLATURE

$X(t)$	=	stochastic process
$x_k(t)$	=	sample realization of $X(t)$
γ_{ij}	=	autocovariance of the stochastic process (delay $t_j - t_i$)
$\gamma_{XX}(u)$	=	autocovariance of a stationary process (with delay u)
$\Gamma_{XX}(f)$	=	spectrum of a stationary sample function
Δt	=	sampling rate
F	=	sampling frequency ($= 1/\Delta t$)
f_N	=	Nyquist folding frequency
$\Gamma_a(f)$	=	aliased spectrum (periodic with period F)
$c_{xx}(k)$	=	sample autocovariance function
$c_{XX}(k)$	=	biased autocovariance function estimator
$c'_{XX}(k)$	=	unbiased autocovariance function estimator
L	=	integral scale of the process
$w_B(l)$	=	Bartlett lag window
$w_T(l)$	=	Tukey lag window
$w_P(l)$	=	Parzen window
$w_{PP}(l)$	=	Parzen bias reducing window
$W(f)$	=	Fourier transform of w(u)
$C_{XX}(g)$	=	Fourier transform of $c_{xx}(u)$
$X_T(t)$	=	$X(t)$ for $0 \leq t \leq T$, 0 for $t \geq T$
$x_T(t)$	=	sample realization of $X_T(t)$
$a_T(f)$	=	$\frac{1}{T}\int_0^T x_T(t)\, e^{-j2\pi ft} dt$
$c_{xx}^{int}(u)$	=	sample autocovariance of $X(t)$ (integral reperesentation)
$c_{XX}^{int}(u)$	=	autocovariance function estimator (integral representation)
$C_{xx}(f)$	=	sample spectrum
$C_{XX}(f)$	=	spectrum estimator
$a_{T_a}(f)$	=	$\sum_{k=-\infty}^{\infty} a_T(f+Kf)$
M_f	=	no. of points in the frequency domain
M	=	no. of points in each subseries (time domain)
N	=	total no. of points
K	=	no. of subseries ($K = N/M$)
T	=	time record length of each subseries ($= M\Delta t$)

T_r	=	overall time record length (= $N\Delta t$)
Δf	=	frequency resolution
$\bar{U}$	=	longitudinal mean velocity
u'	=	longitudinal velocity fluctuation
σ	=	$\sqrt{\overline{(u')^2}}$
L_u^x	=	longitudinal turbulence integral length scale
$S(f)$	=	sample spectrum = $2\overline{C}_{xx}(f)$
PSD	=	power spectral density = $\frac{S(f)}{\sigma^2}$
$c_{xx_a}(n)$	=	sample autocovariance obtained from the aliased periodic realization $X_{T_a}(l)$
$C_{xx_a}(n)$	=	aliased periodic version of the sample spectrum
$c_{xx_{0a}}(n)$	=	sample circular autocovariance function
$C_{xx_{0a}}(n)$	=	sample spectrum of the time record extended by adding zeros

(overbar '$\bar{}$' denotes averaged quantities)

Common subscript notations :

'a' ≡ aliased function

'XX' ≡ estimator of the process

'xx' ≡ sample function estimate of the process

'T' ≡ finite time record

'0' ≡ indicates time record extended by zeros

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

The purpose of this work is two-fold: first, to present some measurements of velocity fluctuations in an urban atmospheric boundary layer simulated in a long test section wind tunnel, and second, and primarily, to analyze the techniques used to estimate the statistical parameters describing the fluctuations, specifically the autocovariance and spectral density functions. In satisfying the latter objective, the presentation is somewhat tutorial in nature, but this approach can be justified by noting that proper interpretation of experimental data for complex random phenomena requires a thorough understanding of the nature and limitations of the estimates of relevant statistical parameters and functions. The analysis is applicable to any random stationary process and is largely self-contained, although extensive use has been made of the classical book by Jenkins and Watts (1968) in its preparation.

Some basic concepts associated with acquisition of stochastic data are presented in Chapter 2, including a discussion of aliasing and of the effect of a finite record length on the ability to identify neighboring spectral peaks. The selection of autocovariance function estimators is discussed in Chapter 3, which begins with a review of the concepts of bias and mean square error. Some useful properties of autocovariance function estimators and an initial discussion of lag windows are also presented in this chapter. In Chapter 4, some basic results pertaining to the harmonic representation of stationary random processes are first reviewed. The digital evaluation and smoothing of spectral estimators, and the development of confidence intervals for the spectral estimates, are then discussed in detail, together with the criteria governing the selection of the sampling interval, the subseries length, and the total record length. This chapter includes all necessary equations for implementation of the numerical calculations in a computer.

First- and second-order autoregressive processes were used for verification of the programs developed for calculation of the autocovariance and spectral density functions. The known theoretical spectra for these processes were compared with the sample spectra obtained by applying the

numerical programs developed on realizations of these processes. Some of these comparisons are presented in Chapter 5. The effects of varying the length of the subseries used and the total record length have been illustrated using both the autoregressive processes and the experimental data presented in Chapter 6.

The velocity fluctuation data in Chapter 6 were obtained in an atmospheric boundary layer simulation carried out for urban terrain conditions in the St. Anthony Falls Laboratory boundary layer wind tunnel. The simulation technique was based on the use of quarter-elliptic, constant-wedge-angle vortex generators with height H = 1.20 m and a castellated barrier wall to produce the necessary initial momentum defect in the boundary layer, followed by a fetch of roughness elements representative of the urban terrain under consideration. The roughness selected consisted of a staggered arrangement of cubes with height k = 28 mm, placed 100 mm apart in rows transverse to the flow. The streamwise separation of the rows was also 100 mm. More details of the simulation can be found in Farell and Iyengar (1993). The St. Anthony Falls wind tunnel features a 1.70 m wide x 1.80 m high x 16 m long test section, with a ceiling height adjustable ± 0.30 m over most of its length, where wind velocities up to 45 m/s can be produced. A second test section, 2.4 m x 2.4 m x 18 m long, with velocities up to 20 m/s, is located in the return leg of the tunnel. The facility can be operated in both open and closed circuit modes. It has a 8-foot-diameter fan with manually variable pitch, powered by a variable speed 200 hp AC motor. The experiments reported in this work were conducted in the higher speed test section, with the tunnel operated in the closed circuit mode.

2. BASIC IDEAS

We begin by reviewing briefly some elementary concepts associated with random phenomena. Stochastic or random processes provide probabilistic descriptions of physical phenomena that develop in time according to well-defined laws of probability. Suppose that an experiment involving a stochastic process is performed and that the outcome of the experiment, represented, say, by a voltage as a function of time, is recorded. The records, $x_k(t)$, for

several experimental runs might appear as illustrated in Fig. 2.1, where the subscript k denotes a particular experimental outcome or realization. If the time of observation, t, for $x_k(t)$ extends from $-\infty$ to $+\infty$, the time series $x_k(t)$ is called a sample function. It should be emphasized that the sample function itself is not the random process. The stochastic process is the collection of all possible sample functions that the experiment might produce. The statistical properties of the stochastic process may be obtained by taking averages across the collection (known as an ensemble) of sample functions which make up the stochastic process.

In general, the statistical parameters which characterize a stochastic process may be computed from moments of the probability density functions associated with the process. In practice, only the low-order moments are used, particularly because estimates of high-order moments are not likely to converge. Two low-order moments of special significance are the mean and autocovariance functions. The mean value of the stochastic process at time t_i may be estimated from the relationship

$$\text{ave } \{x_k(t_i)\} \sim \frac{1}{N} \sum_{k=1}^{N} x_k(t_i) \tag{3}$$

where the subscript k again identifies a particular realization of the process. The autocovariance, γ_{ij}, is given by

$$\begin{aligned} \gamma_{ij} &= \text{Cov } \{x_k(t_i), x_k(t_j)\} \\ &= \text{ave } [x_k(t_i) - \text{ave } \{x_k(t_i)\}][x_k(t_j) - \text{ave } \{x_k(t_j)\}] \end{aligned} \tag{4}$$

In general, ave $[x_k(t_i)] \neq$ ave $[x_k(t_j)]$. Such a stochastic process is called nonstationary. If the ensemble averages are equal for all t_i and if the values of the autocovariance for a given delay time $(t_j - t_i)$ do not depend on t_i, the process is said to be weakly or covariance stationary. If the higher-

order statistical parameters are also time invariant then the process is is said to be strictly stationary. For both weakly stationary and strictly stationary processes, the mean value is a constant. The value of the autocovariance function depends only upon the time interval, $(t_j - t_i) = u$, and will be denoted by $\gamma_{XX}(u)$.

The experimental determination of the ensemble or probability average is generally impractical since a large number of experimental runs must be performed before significant statistical averages may be obtained. Fortunately, for most stationary random processes of interest, ensemble averages are equal to time averages along a single sample function. That is,

$$\text{ave } \{x_k(t_i)\} = \overline{x(t)}$$

$$\text{cov } \{x_k(t_i),\ x_k(t_i + u)\} = \overline{x(t)\ x(t+u)} \tag{5}$$

in which the bars represent time averages. This property of a process is known as ergodicity. If a random process is ergodic, all of the relevant statistical averages can be performed in time using only one sample function, rather than averaging across the ensemble of sample functions. Thus, the ergodic property results in considerable economy of experimental effort.

The principal emphasis herein will be focused upon digital time series, that is, the discrete representation of a continuous signal. The spectrum, $\Gamma_{XX}(f)$, will be defined for brevity in this section simply as the Fourier transform of the autocovariance function, as follows:

$$\Gamma_{XX}(f) = \int_{-\infty}^{\infty} \gamma_{XX}(u)\ e^{-j2\pi fu}\ du \tag{6}$$

where $j = \sqrt{-1}$, and f denotes frequency in Hz.

The definitions given by Eqs. (4) and (6) require an infinite length of continuous data, the sample function $x_k(t)$, for computation. However, in practice, one has a finite length of record digitized at equally spaced intervals in time. Let us briefly review what the principal effects of analyzing a finite, discrete record will be on the ability to estimate the spectrum, or equivalently, the autocovariance function.

Suppose that x(t) is a sample function and that the autocovariance function $\gamma_{XX}(u)$ and the spectrum $\Gamma_{XX}(f)$ are computed according to Eqs. (4) and (6), respectively. If x(t) is sampled at equally spaced points in time, only discrete values of the autocovariance function can be calculated, resulting in a discretized autocovariance function. With $\gamma_{XX}(u)$ effectively sampled at intervals of Δt, the expression for the autocovariance function at the sampling points $r\Delta t$, $r = 0, \pm 1, \pm 2, \ldots$ is

$$\gamma_{XX}(r\Delta t) = \int_{-\infty}^{\infty} \Gamma_{XX}(f)\, e^{j2\pi fr/F}\, df \tag{7}$$

where $F = \frac{1}{\Delta t}$. Following Cooley et al. (1967) we can express the integral in Eq. (7) as an infinite sum of integrals in the form

$$\gamma_{XX}(r\Delta t) = \sum_{k=-\infty}^{\infty} \int_{kF}^{(k+1)F} \Gamma_{XX}(f)\, e^{j2\pi fr/F}\, df \tag{8}$$

or, upon introducing the change of variables $f = f' + kF$, as

$$\gamma_{XX}(r\Delta t) = \sum_{k=-\infty}^{\infty} \int_{0}^{F} \Gamma_{XX}(f' + kF)\, e^{j2\pi (f'+kF)r/F}\, df'$$

But the exponential function in this equation is periodic, and therefore we get

$$\gamma_{XX}(r\Delta t) = \sum_{k=-\infty}^{\infty} \int_{0}^{F} \Gamma_{XX}(f + kF)\, e^{j2\pi fr/F}\, df$$

or

$$\gamma_{XX}(r\Delta t) = \int_{0}^{F} \left[\sum_{k=-\infty}^{\infty} \Gamma_{XX}(f + kF)\right] e^{j2\pi fr/F}\, df \tag{9}$$

The function in the brackets

$$\Gamma_a(f) = \sum_{k=-\infty}^{\infty} \Gamma_{XX}(f + kF) \tag{10}$$

is a periodic function of period F, formed by the superposition of the non-periodic function $\Gamma_{XX}(f)$ shifted by all multiples of the fundamental period. With this definition of $\Gamma_a(f)$ we have:

$$\gamma_{XX}(r\Delta t) = \int_{0}^{F} \Gamma_a(f)\, e^{j2\pi fr/F}\, df \tag{11}$$

This result shows that if the covariance function is known only at discrete points, then all that can be done about obtaining $\Gamma_{XX}(f)$ is to compute $\Gamma_a(f)$. Equation (11) shows furthermore that the discrete values of the autocovariance function, multiplied by 1/F, are the coefficients of the Fourier series expansion of the periodic function $\Gamma_a(f)$, and we have then the reciprocal relation:

$$\Gamma_a(f) = \frac{1}{F} \sum_{r=-\infty}^{\infty} \gamma_{XX}(r\Delta t)\, e^{-j2\pi fr/F} \qquad (12)$$

The superposition of frequencies expressed by Eq. (10) is called aliasing, since frequencies greater than $F/2 = 1/(2\Delta t)$ impersonate components with frequencies less than $1/(2\Delta t)$. The frequency $f_N = 1/(2\Delta t)$ is called the Nyquist or folding frequency. If the spectrum $\Gamma_{XX}(f)$ contains frequency components greater than f_N, then the spectrum computed from discrete data will be composed of contributions from $\Gamma_{XX}(f)$, $\Gamma_{XX}(f + 1/\Delta t)$, $\Gamma_{XX}(f - 1/\Delta t)$, $\Gamma_{XX}(f + 2/\Delta t)$, $\Gamma_{XX}(f - 2/\Delta t)$, and so on, as shown by Eq. (10). These frequencies are referred to as the aliases of the frequency f. The aliased spectrum yields a fallacious picture of the true spectrum, and it is impossible to recover $\Gamma_{XX}(f)$ from $\Gamma_a(f)$. Therefore, care must be taken to insure that the sampling frequency is high enough to avoid aliasing effects. The theoretical spectrum $\Gamma_{XX}(f)$ and the spectrum $\Gamma_a(f)$ obtained by sampling at time intervals larger and smaller than $1/(2f_{max})$ are depicted in Fig. 2.2 Aliasing effects may be avoided by choosing Δt such that $\Gamma_{XX}(f)$ is zero for $f \geq 1/(2\Delta t)$. In this case the function $\Gamma_a(f)$ is identically equal to $\Gamma_{XX}(f)$ in the interval $|f| \leq 1/(2\Delta t)$, as shown by Eq. (10), and no information is lost by using a discrete time series rather than a continuous function. Thus, if a process is bandwidth limited, analysis of the discrete (but still infinite) time series will yield the spectral density function provided that the sampling interval, Δt, is less than or equal to $1/(2f_{max})$, where f_{max} is the highest frequency present in the signal. In practice, the high frequencies are most conveniently eliminated by using an analog filter prior to digitizing the data. It should be noted that it is not possible to correct for aliasing after the data has been digitized by applying a low-pass digital filter to the data sequence. Digital filtering can eliminate components of the signal in the frequency range $|f| \leq 1/(2\Delta t)$, but it cannot separate components present in the original signal from those superimposed as a consequence of aliasing.

In the discussion of aliasing, the data sequence was discrete but infinite. The effects of a finite length of record can be examined most directly by considering an ensemble of continuous sample records of finite length for which the autocovariance function $\gamma_{XX}(u)$ can be determined up to some time lag, T. To compute the Fourier transform the autocovariance for the finite record must be defined on the interval $(-\infty,\infty)$. This can be accomplished by considering the function to be transformed as the product of the function $\gamma_{XX}(u)$ and a window w(u) defined as

$$w(u) = \begin{cases} 1, & 0 \leq |u| \leq T \\ 0, & |u| > T \end{cases} \tag{13}$$

Application of the convolution theorem to the transform of the product $\gamma_{XX}(u) \cdot w(u)$ yields

$$\Gamma_T(f) = \int_{-\infty}^{\infty} \Gamma_{XX}(g)\, W(f - g)\, dg \tag{14}$$

where the window function W(f) is the Fourier transform of w(u),

$$W(f) = \frac{T \sin(2\pi fT)}{\pi fT} \tag{15}$$

Equation (14) shows that the function computed from the finite autocovariance is a weighted or smoothed version of the true spectrum $\Gamma_{XX}(f)$. The constraint of a finite record length thus limits the ability to resolve adjacent frequencies. As might be expected the analysis of discrete, finite record lengths combines both aliasing and smoothing effects.

Since the quantity $\Gamma_T(f)$ evidently depends upon the true spectrum $\Gamma_{XX}(f)$ over some frequency band, the question arises immediately as to the resolution of the procedure, that is, the ability to identify spectral peaks. In the case of two sinusoidal signals of frequencies f_1 and f_2, for which the corresponding spectrum consists of two delta functions centered at these two frequencies, the smoothed spectrum given by Eq. (14) is proportional to the sum, $W(f-f_1) + W(f-f_2)$, of the window $W(f)$ centered first at f_1 and then at f_2.

As the spacing between f_1 and f_2 becomes smaller, the spurious components present in $\Gamma_T(f)$ at frequencies close to f_1 and f_2 merge together and a single intermediate peak may result. The problem is analogous to that in physical optics of the resolution of spectral lines from a light source using diffraction gratings. Rayleigh's criterion of resolution for the latter problem, in terms of the window $W(f)$, is that two close peaks are barely resolved (separated by a small dip in the curve) when the maximum of one peak is aligned with the first trough of the spectral window centered at the other peak. According to this criterion, the minimum record length (the window width in time space) necessary to resolve barely two peaks Δf apart is, for the rectangular window, $T \sim 0.7\ (1/\Delta f)$. For the Bartlett window, to be introduced later, the result is $T = 1/\Delta f$.

Figure 2.3 shows actual calculations using the window (13) for the combined sinusoidal signal. Both peaks are well resolved for a record length $T = 1/(\Delta f)$. The spectral estimates in Fig. 2.3 have been normalized through division by the product of the length of the record and the square of the amplitude of both signals. These results confirm those presented in Jenkins and Watts (1968). In general, the exact resolution of a digital technique will depend on both the shape of the window and the shape of the spectrum. These results imply also that for a record of length T no frequency resolution can be achieved in the range $0 \leq f \leq 1/T$. If examination of a signal of frequency f is required, the record length must be at least $1/f$.

3. SELECTION OF AN AUTOCOVARIANCE FUNCTION ESTIMATOR

In order to judge the relative merits of different estimators of either the spectrum or autocovariance, some criteria for comparison must be established. It should be clear that the estimate itself is a random variable, since a different value will be obtained for each realization of the random process. The bias of an estimator, $\hat{\theta}$, of a quantity θ is defined as

$$B = E[\hat{\theta}] - \theta \tag{16}$$

where E denotes the expected value or mean of the quantity in the brackets. If the bias is 0 the probability density function of the estimator $\hat{\theta}$ is centered at the true value of θ. The variance of the estimator is given by

$$\text{var}\ (\hat{\theta}) = E\{ [\hat{\theta} - E(\hat{\theta})]^2 \} \tag{17}$$

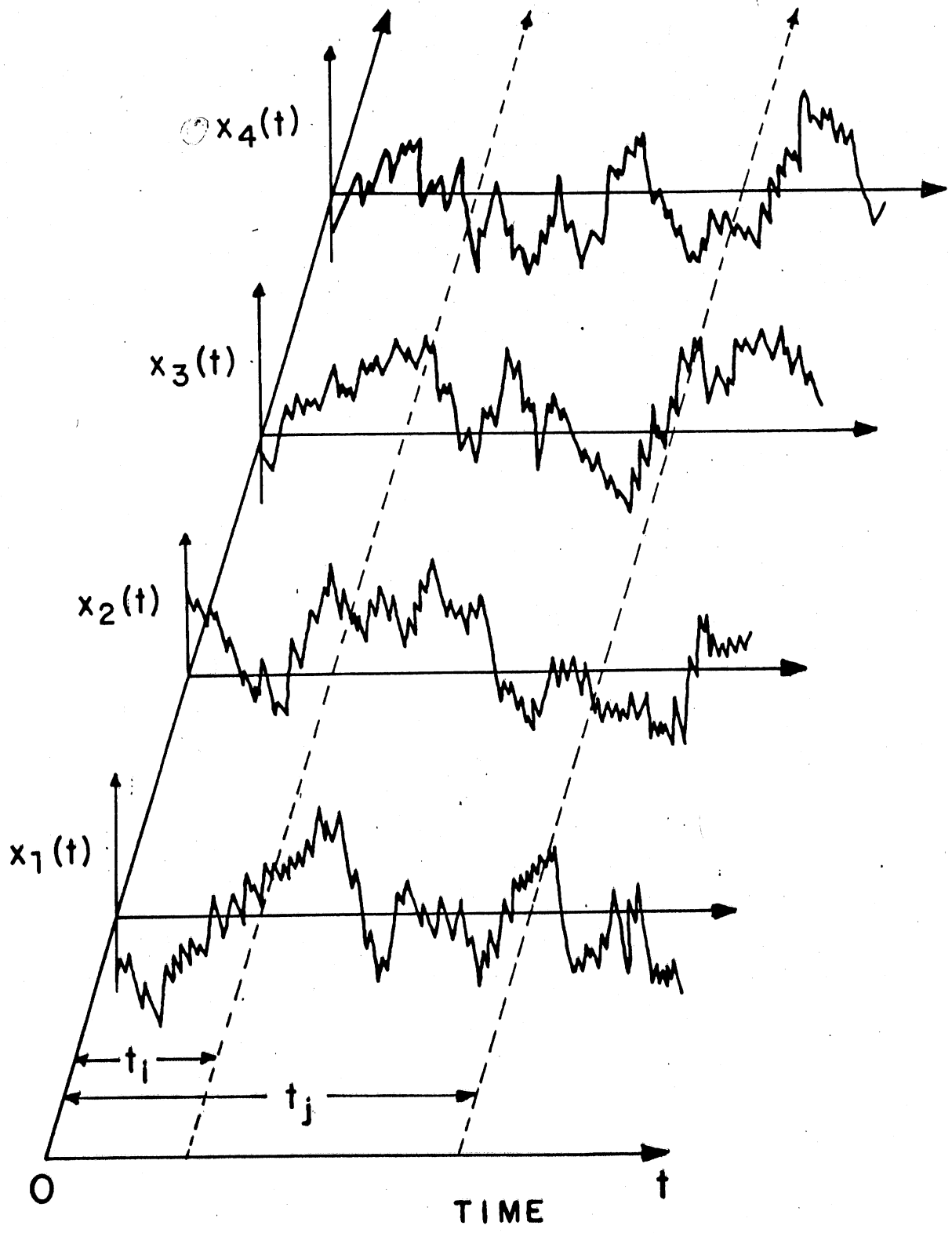


Fig. 2.1 Ensemble of sample functions

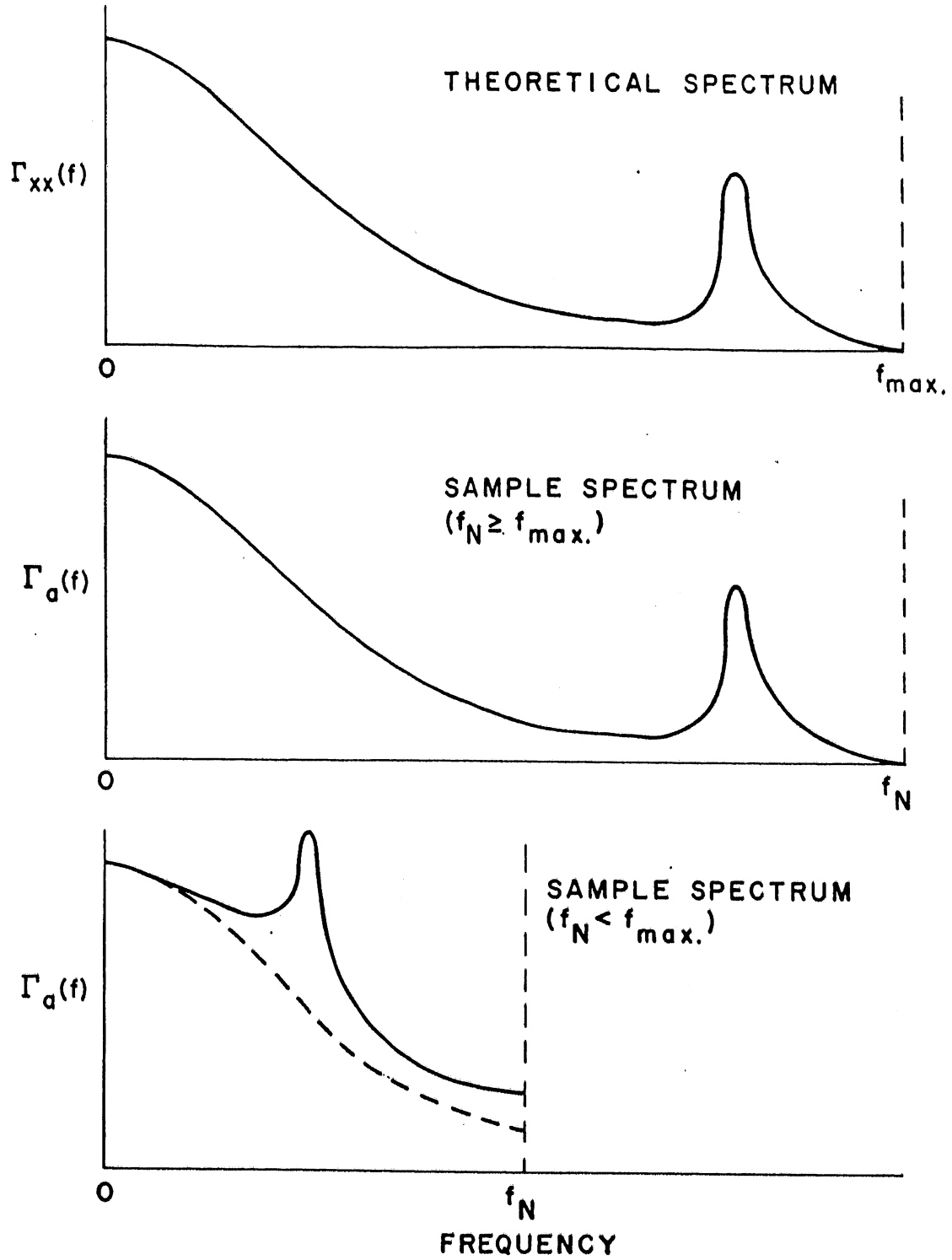
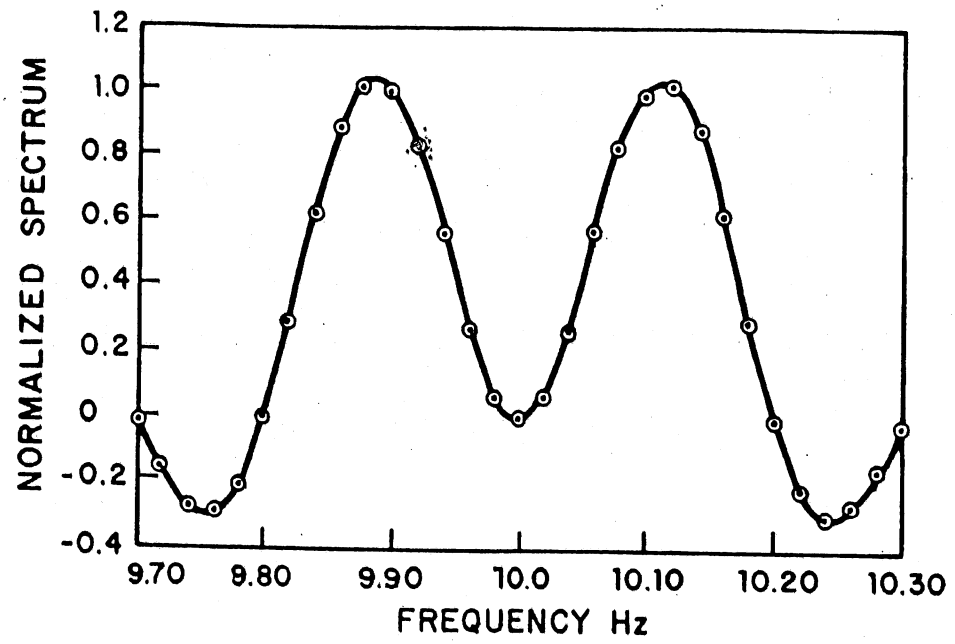
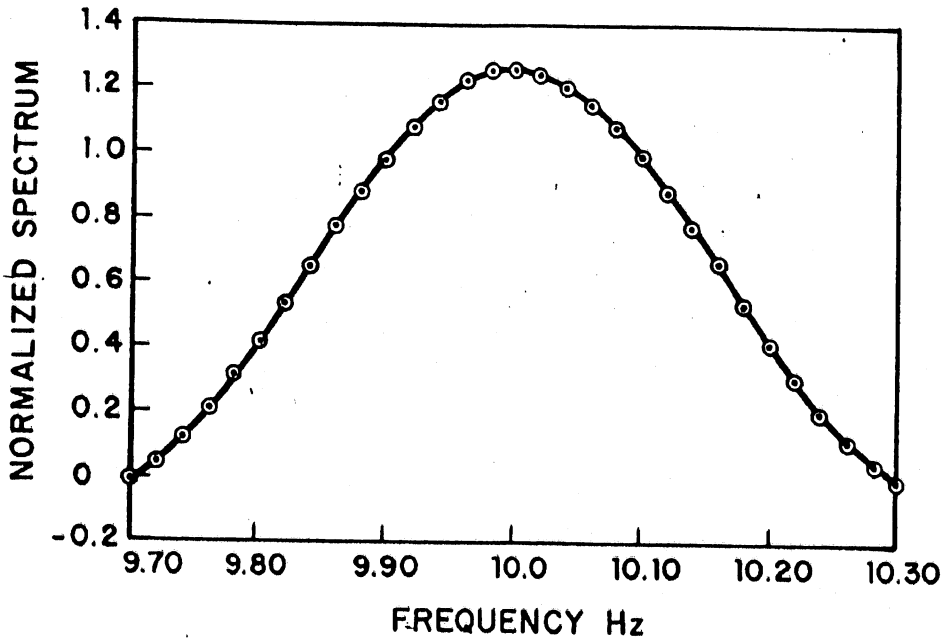


Fig. 2.2 Aliasing in spectrum



(a) $f_2 - f_1 = \frac{1}{T}$



(b) $f_2 - f_1 = \frac{1}{2T}$

Fig. 2.3 Resolution of peaks in the spectrum for different record lengths

and measure the spread of the probability density function of θ about its expected value. Generally, one should seek an estimator with both small bias and small variance. These two requirements, however, are not necessarily compatible, since reducing the variance in general increases the bias and vice versa. A compromise between bias and variance can be achieved by minimizing the mean square error of the estimator, given by

$$\text{mse} = E[(\hat{\theta} - \theta)^2] = \text{var}\ (\hat{\theta}) + B^2. \tag{18}$$

Although the mean square error criterion does not indicate how to choose an estimator, it does provide a means for comparison of the relative merits of different estimators.

In this and subsequent chapters, the stochastic process X(t) will be assumed covariance stationary, ergodic, and with zero mean value. If a given realization of this process, x(t), is sampled at intervals of Δt a sample autocovariance function may be defined

$$c_{xx}(k) = \begin{cases} \frac{1}{N} \sum_{i=1}^{N-|k|} x(i)x(i+|k|), & 0 \leq |k| \leq N-1 \\ 0, & |k| > N-1 \end{cases} \tag{19}$$

where $c_{xx}(k)$ is the sample autocovariance function at lag time $k\Delta T$, x(i) and $x(i + |k|)$ are the observations at times $i\Delta t$ and $(i + |k|)\Delta t$, respectively, and N is the number of observed data points in the discrete time series. The autocovariance function estimator which corresponds to the sample autocovariance function $c_{xx}(k)$ is

$$c_{XX}(k) = \begin{cases} \frac{1}{N} \sum_{i=1}^{N-|k|} X(i)X(i+|k|), & 0 \leq |k| \leq N-1 \\ 0, & |k| > N-1 \end{cases} \tag{20}$$

The estimator $c_{XX}(k)$ defined by Eq. (20) is a biased estimator. An unbiased estimator can be defined by

$$c'_{XX}(k) = \begin{cases} \frac{1}{N - |k|} \sum_{i=1}^{N-|k|} X(i)X(i+|k|), & 0 \leq |k| \leq N-1 \\ 0, & |k| > N-1 \end{cases} \tag{21}$$

where $c'_{XX}(k)$ is the estimator at lag time $k\Delta t$. More generally we will consider averaged biased and unbiased estimators, whose definition is discussed next.

According to a well-known result in probability theory a random variable defined as the average of K mutually independent random variables having equal variance σ^2 has variance equal to σ^2/K. On the basis of this result, Bartlett (1950) introduced the idea of subdividing the time series into K smaller subseries of M points each and computing the spectrum as the average of the sample spectra calculated for each subseries. Because, as will be seen in the next chapter, the variance of the sample spectra evaluated for each subseries is approximately constant if the number of points M in each subseries is not too small, the variance of the smoothed spectrum obtained by Bartlett's method is considerably reduced. The requirement that the subseries be stochastically independent is satisfied if the length $M\Delta t$ of each subseries is greater than twice the integral scale, L, of the process, defined by

$$L = \frac{1}{\gamma_{XX}(0)} \int_0^\infty \gamma_{XX}(u)\, du \tag{22}$$

(Tennekes and Lumley 1972, pp. 212-214).

We consider then the possibility of subdividing the time series into K smaller subseries and introduce the averaged biased and unbiased autocovariance function estimators defined respectively by

$$\bar{c}_{XX}(\ell) = \begin{cases} \frac{1}{K} \sum_{k=1}^{K} \frac{1}{M} \sum_{i=1}^{M-|\ell|} X_k(i) X_k(i+|\ell|), & 0 \le |\ell| \le M-1 \\ 0, & |\ell| > M-1 \end{cases} \tag{23}$$

and

$$\bar{c}'_{XX}(\ell) = \begin{cases} \frac{1}{K} \sum_{k=1}^{K} \frac{1}{M - |\ell|} \sum_{i=1}^{M-|\ell|} X_k(i) X_k(i+|\ell|), & 0 \le |\ell| \le M-1 \\ 0, & |\ell| > M-1 \end{cases} \tag{24}$$

The variance of either of these estimators affects the error in any sample estimate of the autocovariance function corresponding to a given realization of the stochastic process and may be quite large (Schaerf 1963, Jenkins and Watts 1968). It will now be shown, following Jenkins and Watts (1968), that the biased estimator has the smallest mean square error and is in this sense a better estimator than the unbiased estimator defined by Eq. (24).

The mean value of the estimator $\bar{c}_{XX}(\ell)$ is given by the expression

$$E[\bar{c}_{XX}(\ell)] = (1 - \frac{|\ell|}{M}) \gamma_{XX}(\ell) \tag{25}$$

and its bias is then

$$B(\ell) = - \frac{|\ell|}{M} \gamma_{XX}(\ell) \tag{26}$$

The contribution of the bias in the expected value of $c_{XX}(\ell)$ to the mean square error is, therefore, $(\frac{\ell}{M})^2 \gamma^2_{XX}(\ell)$. This term is small for processes having exponentially decreasing autocovariance functions and satisfying the condition that $M\Delta t$ is at least equal to 2L. The estimator $\bar{c}'_{XX}(\ell)$, on the other hand, is unbiased.

An expression for the covariance of the autocovariance function estimators will now be derived and the variance of the estimators will be obtained therefrom. The resulting expressions are exact for covariance stationary normal processes with zero mean; for processes which are not normal they hold with good approximation. A more detailed presentation of the derivation (in the case K = 1) and a discussion of the approximations involved can be found in Jenkins and Watts (1968, pp 412-418).

The covariance between the estimators $\bar{c}_{XX}(k)$ and $\bar{c}_{XX}(\ell)$ of the autocovariance function at lags $k\Delta t$ and $\ell\Delta t$, is given by

$$\text{cov}\,[\bar{c}_{XX}(k), \bar{c}_{XX}(\ell)] = E[\bar{c}_{XX}(k)\bar{c}_{XX}(\ell)] - [E\,\bar{c}_{XX}(k)]\, E[\bar{c}_{XX}(\ell)] \tag{27}$$

Introducing Eq. (23) for the averaged estimator and assuming for simplicity that both k and ℓ are positive, we obtain

$$\text{cov}\,[\bar{c}_{XX}(k)\bar{c}_{XX}(\ell)] = E\left\{\left[\frac{1}{K}\sum_{m=1}^{k}\frac{1}{M}\sum_{i=1}^{M-k} X_m(i)\; X_m(i+k)\right]\right.$$

$$\left.\times\left[\frac{1}{K}\sum_{n=1}^{K}\frac{1}{M}\sum_{j=1}^{M-\ell} X_n(j)\; X_n(j+\ell)\right]\right\}$$

$$-\left\{E\left[\frac{1}{K}\sum_{m=1}^{K}\frac{1}{M}\sum_{i=1}^{M-k} X_m(i)\; X_m(i+k)\right]\right.$$

$$\left.\times\; E\left[\frac{1}{K}\sum_{n=1}^{K}\frac{1}{M}\sum_{j=1}^{M-\ell} X_n(j)\; X_n(j+\ell)\right]\right\} \qquad (28)$$

By using the fourth product moment formula for normal random variables with zero means (see e.g., Parzen 1962, p. 93)

$$E[X(i)\; X(j)\; X(k)\; X(\ell)] = E[X(i)\; X(j)]\; E[X(k)\; X(\ell)]$$
$$+ E[X(i)\; X(k)]\; E[X(j)\; X(\ell)]$$
$$+ E[X(i)\; X(\ell)]\; E[X(j)\; X(k)]$$

and simplifying, one obtains from Eq. (28)

$$\text{cov}\,[\bar{c}_{XX}(\ell),\; \bar{c}_{XX}(k)] = \frac{1}{K^2}\sum_{m=1}^{K}\sum_{n=1}^{K}\frac{1}{M^2}\sum_{i=1}^{M-k}\sum_{j=1}^{M-\ell}$$

$$\left\{E[X_m(i)\; X_n(j)]\;\; E[X_m(i+k)\;\; X_n(j+\ell)]\right.$$

$$\left.+\; E[X_m(i)\;\; X_n(j+\ell)]\;\; E[X_m(i+k)\;\; X_n(j)]\right\} \qquad (29)$$

Since the subseries are assumed to be stochastically independent the correlation between data points separated by more than $M\Delta t$ is zero. The summations indicated in Eq. (29) can therefore be simplified, resulting in

$$\text{cov}\,[\bar{c}_{XX}(\ell),\ \bar{c}_{XX}(k)] = \frac{1}{K^2}\sum_{m=1}^{K}\frac{1}{M^2}\sum_{i=1}^{M-k}\sum_{j=1}^{M-\ell}\{E[X_m(i)\ X_m(j)]$$

$$\times\ E[X_m(i+k)\ X_m(j+\ell)] + E[X_m(i)\ X_m(j+\ell)]$$

$$\times\ E[X_m(i+k)\ X_m(j)]\} + \frac{1}{K^2}\sum_{m=1}^{K-1}\frac{1}{M^2}\sum_{i=1}^{M-k}\sum_{j=1}^{M-\ell}$$

$$\Big\{E[X_m(i)\ X_{m+1}(j)]\ E[X_m(i+k)\ X_{m+1}(j+\ell)]$$

$$+\ E[X_m(i)\ X_{m+1}(j+\ell)]\ E[X_m(i+k)\ X_{m+1}(j)]\Big\} \qquad (30)$$

Using the fact that the series is stationary, the final result for the covariance of the estimators is

$$\text{cov}[\bar{c}_{XX}(\ell),\ \bar{c}_{XX}(k)] = \frac{1}{KM^2}\sum_{i=1}^{M-k}\sum_{j=1}^{M-\ell}[\gamma_{XX}(j-i)\ \gamma_{XX}(j-i+\ell-k)$$

$$+\ \gamma_{XX}(j-i+\ell)\ \gamma_{XX}(j-i-k)] + \frac{(K-1)}{K^2M^2}\sum_{i=1}^{M-k}\sum_{j=1}^{M-\ell}$$

$$[\gamma_{XX}(M+j-i)\ \gamma_{XX}(M+j-i+\ell-k)$$

$$+\ \gamma_{XX}(M+j-i+\ell)\ \gamma_{XX}(M+j-i-k)] \qquad (31)$$

The variance of the autocovariance function estimator is obtained by setting $\ell = k$ in Eq. (31). Introducing a change of variables as outlined in Jenkins and Watts (1968) one obtains without difficulty for the variance of the estimator the expression

$$\text{var}\,[\bar{c}_{XX}(k)] = \frac{1}{KM^2}\sum_{r=-(M-1-k)}^{M-1-k}\Big\{(M-k-|r|)\ \gamma_{XX}^2(r) + \gamma_{XX}(r+k)\ \gamma_{XX}(r-k)\Big\}$$

$$+\ \frac{(K-1)}{K^2M^2}\sum_{r=-(M-1-k)}^{M-1-k}(M-k-|r|)\Big\{\gamma_{XX}^2(M+r)$$

$$+\ \gamma_{XX}(M+r+k)\ \ \gamma_{XX}(M+r-k)\Big\} \qquad (32)$$

The second term in Eq. (32) can be readily shown, at least qualitatively, to be much smaller than the first if M is not too small. Calculations for the variance of a first-order autoregressive process confirm this finding and the term will therefore be neglected.

An approximation due to Bartlett (1946) (see also Jenkins and Watts 1968, p. 178) valid only for sufficiently large M, and which consists essentially of cancelling the factor $(M-k-|r|)$ in the numerator with one of the factors M in the denominator and carrying out the summation from $-\infty$ to $+\infty$, can be used to simplify further the results in Eq. (32). We obtain thus

$$\mathrm{var}[\,\bar{c}_{XX}(k)\,] = \frac{1}{KM} \sum_{r=-\infty}^{\infty} [\gamma_{XX}^2(r) + \gamma_{XX}(r+k)\ \gamma_{XX}(r-k)\,] \qquad (33)$$

Using this simplified expression, one can immediately see that the variance of the autocovariance estimator depends on the total number of points in the time series and remains approximately constant regardless of the division into any number (K) of subseries, provided that the length of the subseries is not too small. Calculations of the variance of the autocovariance function estimator for first- and second-order autoregressive processes and for the pressure and velocity signals using Eq. (32) and different combinations of K and M confirm this fact. (The values of the autocovariance function were estimated from the longest series available.)

The mean square error of the autocovariance function estimator $\bar{c}_{XX}(k)$ is the sum of the variance given by Eq. (32) and the square of the bias given by Eq. (26). The mean square error of the unbiased estimator, $\bar{c}'_{XX}(k)$, is equal to its variance, which can be expressed as

$$\mathrm{var}[\bar{c}'_{XX}(k)\,] = \frac{1}{K} \frac{1}{(M-k)^2} \sum_{r=-(M-1-k)}^{M-1-k} (M-k-|r|)$$

$$x[\gamma_{XX}^2(r) + \gamma_{XX}(r+k)\ \gamma_{XX}(r-k)\,] \qquad (34)$$

where the term analogous to the second term in Eq. (32) has been neglected. The mean square errors depend therefore on the unknown autocovariance function of the stochastic process. We can estimate their order of magnitude at a few points. At $k = 0$, both estimators have the same mean square error,

$$\text{mse}_{k=0} = \frac{2}{KM^2} \sum_{r=-(M-1)}^{M-1} (M - |r|)\ \gamma_{XX}^2(r) \qquad (35)$$

At $k = M - 1$, the mean square error of $\bar{c}_{XX}(k)$ is approximately $\frac{1}{KM^2}\gamma_{XX}^2(0)$, whereas the corresponding mean square error of $\bar{c}'_{XX}(k)$ is approximately $\frac{1}{K}\gamma_{XX}^2(0)$. Hence, the mean square error of the unbiased estimator of the autocovariance function, $\bar{c}'_{XX}(k)$, is M^2 times larger than that of the biased estimator $\bar{c}_{XX}(k)$ for a delay time $k\Delta t = (M - 1)\Delta t$. More generally, the biased estimator has always been found to have a smaller mean square error than the unbiased estimator, for all delays, except at zero delay where the mean square errors are equal. This relationship seems to be true for all processes, but a rigorous proof of this result is yet to be found. Schaerf (1963) has given proofs for various special cases.

We discuss now another property of the autocovariance function estimators, related to the selection of an estimator. Let a random variable Y be defined by the expression

$$Y = \lambda_1 X(t_1) + \lambda_2 X(t_2) + \ldots + \lambda_n X(t_n) \qquad (36)$$

where $\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_n$ are arbitrary coefficients and $X(t_1), X(t_2), \ldots, X(t_n)$ are the values of the process at times $t_1, t_2, \ldots, t_n$. Since the variance of Y cannot be negative the autocovariance matrix defined by

$$\begin{bmatrix} \gamma_{XX}(0) & \gamma_{XX}(t_2-t_1) & \ldots & \gamma_{XX}(t_n-t_1) \\ \gamma_{XX}(t_2-t_1) & \gamma_{XX}(0) & \ldots & \gamma_{XX}(t_n-t_2) \\ \gamma_{XX}(t_n-t_1) & & & \gamma_{XX}(0) \end{bmatrix}$$

is positive semidefinite, a property that is directly related to the existence of a nonnegative spectral density function for the process*.

Although the autocovariance matrix of a process as defined above is positive semidefinite, the corresponding matrix for a given autocovariance function estimator is not necessarily positive semidefinite. For example, the unbiased estimator $c'_{XX}(k)$ does not have this property (Schaerf 1963) and the spectral estimates obtained using this estimator may therefore exhibit physically impossible negative values for some frequencies. The positive semidefinite property of the covariance matrix appears to be a desirable property for the estimator. While the autocovariance matrix of the unbiased estimator is not positive semidefinite, that of the biased estimator is. In addition, the biased estimator has a smaller mean square error than the unbiased estimator. Hence, the biased estimator is preferred for estimation of the autocovariance function. It should be pointed out however, that the biased estimator still does not provide an entirely satisfactory estimate for the entire autocovariance function. Indeed, the corresponding sample estimator fails to damp out to zero at large lags, as is assumed to be the case for the true autocovariance function. This behavior is a consequence of the high correlation between adjacent autocovariance estimators, which results in spurious periodicities at large delays (Jenkins and Watts 1968, Fig. 5.13). In connection with this behavior, Parzen (1964) has shown that

---

* A theorem of Bochner (1937) states that the necessary and sufficient condition for a continuous function $f(x)$ to have the representation

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{jxu}\, d\phi(u)$$

where $\phi(u)$ is of bounded variation and $d\phi(u) \geq 0$, is that for any arbitrary finite set of points $x_1, x_2, \ldots, x_n$, and for arbitrary coefficients $a_1, a_2, \ldots, a_n$, the Hermitian form

$$\sum_{i,j=1}^{n} a_i \bar{a}_j\, f(x_i - x_j)$$

be positive semidefinite. A more general statement of the theorem is given in the reference cited.

$$\lim_{T\to\infty} E\left[\sum_{k=-(N-1)}^{N-1}\left\{\bar{c}_{XX}(k) - \gamma_{XX}(k)\right\}^2\right] = 2\pi\int_{-\pi}^{\pi}\Gamma_{XX}^2(\lambda)\,d\lambda \qquad (37)$$

That is, the mean square error of the autocovariance function estimator tends to a finite value different from zero as the length of record $T = N\Delta t \to \infty$. In other words, while the biased estimator is a consistent estimate of the autocovariance function at each lag, $k\Delta t$, in the sense that its expected value tends to the true autocovariance function as $T \to \infty$ [see Eq. (25)], it is not a consistent estimate of the entire autocovariance function, as Eq. (37) shows. Nevertheless, as will be discussed in the following chapter, the sample autocovariance estimate when properly transformed through the use of a suitable lag window as defined below does produce spectral estimates which provide a reasonable picture of the spectrum.

We close this chapter by introducing the so-called lag windows or covariance averaging kernels. The expected value of the biased autocovariance function estimator is $(1 - \frac{|\ell|}{M})\gamma_{XX}(\ell)$. The factor

$$w_B(\ell) = \begin{cases} 1 - \dfrac{|\ell|}{M}, & 0 \le |\ell| \le M \\ 0, & |\ell| > M \end{cases}$$

is known as the Bartlett lag window. There are other lag windows which are introduced actually to decrease the variance of the spectral estimators. Three windows most commonly used for this purpose are

1. The Tukey window,

$$w_T(\ell) = \begin{cases} \dfrac{1}{2}\left[1 + \cos\left(\dfrac{\pi\ell}{M}\right)\right], & 0 \le \dfrac{|\ell|}{M} \le 1 \\ 0, & \dfrac{|\ell|}{M} > 1 \end{cases}$$

2. The Parzen window,

$$w_p(\ell) = \begin{cases} 1 - 6\left(\dfrac{\ell}{M}\right)^2 + 6\left(\dfrac{|\ell|}{M}\right)^3, & 0 \le \dfrac{|\ell|}{M} \le \dfrac{1}{2} \\ 2\left[1 - \left(\dfrac{|\ell|}{M}\right)\right]^3, & \dfrac{1}{2} \le \dfrac{|\ell|}{M} \le 1 \\ 0, & \dfrac{|\ell|}{M} > 1 \end{cases}$$

3. The Parzen bias reducing window,

$$w_{pp}(\ell) = [1 + 6(\frac{|\ell|}{M})^2]\ w_p(\ell)$$

The choice of the parameter M, which can be less than or equal to the number of points in each subseries, will be discussed in the following chapter.

As discussed above, positive semidefiniteness of the autocovariance matrix for the product of the sample autocovariance function $c_{xx}(u)$, and the lag window, $w(u)$, is desirable so that no negative spectral estimates are obtained. Using the convolution theorem, the Fourier transform of this product is

$$\bar{C}_{xx}(f) = \int_{-\infty}^{\infty} C_{xx}(g)\ W(f-g)\ dg$$

where $C_{xx}(g)$ is the Fourier transform of $c_{xx}(u)$ and the spectral window $W(f)$ is the corresponding Fourier transform of $w(u)$. With the product $[c_{xx}(u) \cdot w(u)]$ calculated at discrete times, $u = \ell\Delta t$, we have [see Eq. (11)]

$$[c_{xx}(\ell)\ w(\ell)] = \int_0^{\frac{1}{\Delta t}} \bar{C}_{xx_a}(f)\ e^{j2\pi f\ell\Delta t}\ df$$

where

$$\bar{C}_{xx_a}(f) = \sum_{k=-\infty}^{\infty} \bar{C}_{xx}(f + \frac{k}{\Delta t})$$

For the Barlett and Parzen lag windows the corresponding spectral windows are nonnegative and therefore no negative spectral values can result, i.e., the autocovariance matrix is positive semidefinite. On the other hand, although the Tukey spectral window contains negative lobes, the magnitude of these lobes is rather small, and negative values of the sample spectrum are not likely to result. In fact, for sample spectra calculated using the Tukey window for first- and second-order autoregressive processes and for the pressure and velocity signals, no negative estimates have been obtained. It should be noted, however, that negative estimates can easily result when data with strong periodicities are analyzed using this window, as Fig. 2.3 illustrates.

4. ESTIMATION OF THE SPECTRAL DENSITY FUNCTION

4.1. Harmonic Representation of a Stationary Process. Let $x(t)$ denote a realization of a stationary process $X(t)$, $-\infty < t < \infty$, with zero mean. (More precisely, the requirement on the process $X(t)$ is that it be weakly stationary, that is, that its covariance function be continuous and depend only on the delay time). Representation of $x(t)$ either as a Fourier series or as a Fourier integral is not possible because $x(t)$ is not periodic nor does it satisfy the condition that $\int_{-\infty}^{\infty} |x(t)|dt$ be bounded. For the harmonic analysis of $X(t)$ we consider the random variable.

$$X_T(t) = \begin{cases} X(t) , & -T/2 \leq t \leq T/2 \\ 0 , & \text{otherwise} \end{cases} \tag{38}$$

The Fourier transform of a given realization, $x_T(t)$, of $X_T(t)$ is given by

$$z_T(f) = \int_{-\infty}^{\infty} x_T(t)\ e^{-j2\pi ft}\ dt$$

The difficulty to be resolved is that this integral diverges as $T \to \infty$. The device that is adopted depends on the fact that the random variable

$$Z_T(\mu,\lambda) = \int_{-T/2}^{T/2} X_T(t) \int_{\lambda}^{\mu} e^{-j2\pi ft}\ df\ dt$$

$$= \int_{-T/2}^{T/2} X_T(t)\ \frac{e^{-j2\pi t\mu} - e^{-j2\pi t\lambda}}{-j2\pi t}\ dt$$

whose realizations are determined by integrating the realizations $z_T(f)$ between λ and μ converges in the mean square sense to a limiting random variable $Z(\mu,\lambda)$ as $T \to \infty$. (Concise proofs of this and the following results can be found in Rosenblatt 1962). One can then show that

$$\lim_{\lambda \to -\infty} Z(\mu,\lambda) = Z(\mu) \tag{39}$$

exists as a limit in the mean square sense and that

$$Z(\mu,\lambda) = Z(\mu) - Z(\lambda)$$

The process $Z(\mu)$ satisfies the relations

$$E\{[Z(\mu) - Z(\lambda)][Z(\mu') - Z(\lambda')]^*\} = 0 \tag{40}$$

if the intervals (λ,μ), (λ',μ') are disjoint (the symbol * denotes the complex conjugate of a quantity) and

$$E[Z(\mu) - Z(\lambda)]^2 = F(\mu) - F(\lambda) \tag{41}$$

where the function $F(f)$ is the spectral distribution function of the process $X(t)$, that is

$$\gamma_{XX}(u) = \int_{-\infty}^{\infty} e^{j2\pi fu}\, dF(f) \tag{42}$$

[The derivative $dF/df = \Gamma_{XX}(f)$ is the spectral density function or spectrum of the process as defined by Eq. (6) .] In differential notation

$$E[dZ(\lambda)\, dZ^*(\mu)] = \delta_{\lambda,\mu}\, dF(\lambda) \tag{43}$$

where $\delta_{\lambda,\mu}$ is the Kronecker delta. Such processes $Z(\lambda)$ are called processes with orthogonal increments. In terms of $Z(\lambda)$ one can obtain a harmonic representation of the process $X(t)$ in the form of the random Fourier-Stieltjes integral,

$$X(t) = \int_{-\infty}^{\infty} e^{j2\pi ft}\, dZ(f) \tag{44}$$

The stochastic process $X(t)$ is thus represented as a sum of elementary harmonic components of continuously varying frequency. The amplitude of the contribution from an element of length df in the frequency domain

is given by the random variable $dZ(f)$. [The value of the increment $dZ(f)$ at any f depends on the particular realization of the process $X(t)$ considered.] Contributions from elements corresponding to different frequencies are stochastically orthogonal, according to Eq. (43). Consider the variance of the process $X(t)$,

$$\gamma_{XX}(0) = \int_{-\infty}^{\infty} dF(f)$$

[for example, $\gamma_{XX}(0) = \overline{u'^2}$ if $X(t)$ represents the velocity fluctuations, $u'(t)$, at a given point in a turbulent flow.] According to Eq. (43) the contribution to the variance between frequencies f and $f + df$ [represented by $dF(f)$ or by the area under the spectral density function $\Gamma_{XX}(f)$ between f and $f + df$] comes exclusively from the elementary harmonic component $e^{j2\pi ft}\, dZ(f)$; no frequencies other than f contribute to the ordinate at this frequency of the spectral denisty function, which can therefore be evaluated by analog filtering of the signal.

4.2. Alternative Definition of the Spectral Density Function. Given a realization $x(t)$ of a random process $X(t)$ the evaluation of the spectral density function $\Gamma_{XX}(f)$ can be efficiently carried out in digital form making use of the process $X_T(t)$ defined by Eq. (38) or alternatively by

$$X_T(t) = \begin{cases} X(t), & 0 \leq t \leq T \\ 0, & \text{otherwise} \end{cases} \tag{45}$$

which is somewhat more convenient for the calculations that follow. We define the Fourier transform

$$a_T(f) = \frac{1}{T}\int_{-\infty}^{\infty} x_T(t)\, e^{-j2\pi ft}\, dt = \frac{1}{T}\int_{0}^{T} x_T(t)\, e^{-j2\pi ft}\, dt \tag{46}$$

where $x_T(t)$ denotes a given realization of the process $X_T(t)$ defined by Eq. (45). Let $a_T^*(f)$ denote the complex conjugate of $a_T(f)$. We have

$$T\, a_T(f)\, a_T^*(f) = \int_0^T \int_0^T \frac{1}{T}\, x(t)\, x(t')\, e^{-j2\pi f(t-t')}\, dt\, dt'$$

Introducing the change of variables $u = t - t'$, $v = t'$ we get

$$T\, a_T(f)\, a_T^*(f) = \int_0^T \left[\frac{1}{T}\int_0^{T-u} x(v)\, x(v+u)\, dv\right] e^{-j2\pi fu}\, du$$

$$+ \int_{-T}^0 \left[\frac{1}{T}\int_{-u}^{T} x(v)\, x(v+u)\, dv\right] e^{-j2\pi fu}\, du$$

$$= \int_0^T \left[\frac{1}{T}\int_0^{T-u} x(v)\, x(v+u)\, dv\right] e^{-j2\pi fu}\, du$$

$$+ \int_{-T}^0 \left[\frac{1}{T}\int_0^{T+u} x(v')\, x(v'-u)\, dv'\right] e^{-j2\pi fu}\, du$$

$$= \int_0^T c_{xx}^{int}(u)\, e^{-j2\pi fu}\, du + \int_{-T}^0 c_{xx}^{int}(u)\, e^{-j2\pi fu}\, du$$

where $c_{xx}^{int}(u)$ is the sample autocovariance of the process $X(t)$ defined by

$$c_{xx}^{int}(u) = \begin{cases} \frac{1}{T}\int_0^{T-|u|} x(t)\, x(t+|u|)\, dt\,, & |u| \le T \\ 0\,, & |u| > T \end{cases} \tag{47}$$

The autocovariance function estimator corresponding to $c_{xx}^{int}(u)$ is

$$c_{XX}^{int}(u) = \frac{1}{T}\int_0^{T-|u|} X(t)\, X(t+|u|)\, dt\,, \qquad |u| \le T \tag{48}$$

and its mean value is given by

$$E\left[c_{XX}^{int}(u)\right] = \left(1 - \frac{|u|}{T}\right)\gamma_{XX}(u) \tag{49}$$

[Cf. Eqs. (20) and (25).] We have then

$$T\, a_T(f)\, a_T^*(f) = \int_{-T}^{T} c_{xx}^{int}(u)\, e^{-j2\pi fu}\, du$$

We define now the sample spectrum by

$$C_{xx}(f) = T\, a_T(f)\, a_T^*(f) \tag{50}$$

It follows that

$$C_{xx}(f) = \int_{-T}^{T} c_{xx}^{int}(u)\, e^{-j2\pi fu}\, du \tag{51}$$

Thus, the sample spectrum defined by Eq. (50) is the Fourier transform of the sample autocovariance function. The expected value of the sample spectrum estimator $C_{XX}(f)$ can be obtained from Eq. (51) as

$$E\left[C_{XX}(f)\right] = E\left[\int_{-T}^{T} c_{XX}^{int}(u)\, e^{-j2\pi fu}\, du\right]$$

$$= \int_{-T}^{T} \gamma_{XX}(u)\left[1 - \frac{|u|}{T}\right] e^{-j2\pi fu}\, du \tag{52}$$

Letting $T \to \infty$ one obtains

$$\lim_{T\to\infty} E\left[C_{XX}(f)\right] = \lim_{T\to\infty} \int_{-T}^{T} \gamma_{XX}(u)\left[1 - \frac{|u|}{T}\right] e^{-j2\pi fu}\, du$$

$$= \int_{\infty}^{\infty} \gamma_{XX}(u)\, e^{-j2\pi fu}\, du$$

$$= \Gamma_{XX}(f)$$

where $\Gamma_{XX}(f)$ is the spectral density function of the process [Cf. Eq. (42)]. In other words, $\Gamma_{XX}(f)$ can also be defined as

$$\Gamma_{XX}(f) = \lim_{T\to\infty} \{T\ E[a_T(f)\ a_T^*(f)] \tag{53}$$

This relationship is the basis for the method that is presented below for estimation of $\Gamma_{XX}(f)$. In this method, the fast Fourier transform is used (together with appropriate smoothing techniques) to calculate $a_T(f)$ at discrete frequencies. This procedure saves considerable time in comparison with the method in which the covariance function is evaluated first and the sprectrum then obtained by Fourier transforming the covariance function.

4.3. Digital Evaluation. If the process $X_T(t)$ is sampled at discrete equally-spaced points in time, expressions equivalent to those obtained in Chapter 2 for the covariance and spectrum can be obtained. The functions $a_T(f)$ and $(1/T)\ x_T(t)$ are a Fourier transform pair, satisfying Eq. (46) and the inverse relation*

$$\frac{1}{T}\, x_T(t) = \int_{-\infty}^{\infty} a_T(f)\ e^{j2\pi ft}\, df \tag{54}$$

If $x_T(t)$ is discretized at times $t = \ell\Delta t$ we obtain, as described in Chapter 2,

$$\frac{1}{T}\, x_T(\ell\Delta t) = \int_0^F a_{T_a}(f)\ e^{j2\pi f\ell/F}\, df \tag{55}$$

where $a_{T_a}(f) = \sum_{k=-\infty}^{\infty} a_T(f+kF)$ and $F = 1/\Delta t$. Since $a_{T_a}(f)$ is a periodic function (of period F) it can be represented by a Fourier series,

$$a_{T_a}(f) = \sum_{\ell=-\infty}^{\infty} a(\ell)\ e^{-j2\pi f\ell/F}$$

---

*At $t = 0$ and $t = T$ the integral on the right side of Eq. (54) represents respectively $1/(2T)\ x(0)$ and $1/(2T)\ x(T)$. To take this into account the definition of the process $X_T(t)$, Eq. (45), must be modified so that $X_T(0) = 1/2\ X(0)$, $X_T(T) = 1/2\ X(T)$.

where the coefficients, $a(\ell)$, are given by

$$a(\ell) = \frac{1}{F} \int_0^F a_{T_a}(f)\ e^{j2\pi f\ell/F}\ df$$

Using Eq. (55) we obtain

$$a(\ell) = \Delta t\ \frac{1}{T}\ x_T(\ell\Delta t)$$

and therefore

$$a_{T_a}(f) = \frac{\Delta t}{T} \sum_{\ell=-\infty}^{\infty} x_T(\ell\Delta t)\ e^{-j2\pi f\ell/F} \tag{56}$$

If we also discretize $a_{T_a}(f)$ at M_f frequencies $f = m\Delta f$ we have, with $M_f\Delta f = F = 1/\Delta t$ and $m = 0,1,\ldots,M_f-1$,

$$a_{T_a}(m\Delta f) = \Delta t \sum_{\ell=-\infty}^{\infty} \frac{1}{T}\ x_T(\ell\Delta t)\ e^{-j2\pi m\ell/M_f}$$

$$= \Delta t \sum_{k=-\infty}^{\infty} \sum_{\ell=kM_f}^{(k+1)M_f-1} \frac{1}{T}\ x_T(\ell\Delta t)\ e^{-j2\pi m\ell/M_f}$$

that is

$$a_{T_a}(m\Delta f) = \Delta t \sum_{\ell=0}^{M_f-1} \frac{1}{T}\ x_{T_a}(\ell\Delta t)\ e^{-j2\pi m\ell/M_f} \tag{57}$$

with

$$x_{T_a}(\ell\Delta t) = \sum_{k=-\infty}^{\infty} x_T[(\ell+kM_f)\Delta t]$$

The function $x_{T_a}(\ell\Delta t)$ is a discrete periodic function of time of period $M_f\Delta t$. If $M_f\Delta t > T$, then $x_{T_a}(\ell\Delta t)$ is exactly the same as $x_T(\ell\Delta t)$ within

the fundamental period, $0 \leq \ell \leq M_f$-*. The difference between the two functions in this case is that whereas $x_T(\ell\Delta t)$ is zero outside the fundamental period, $x_{T_a}(\ell\Delta t)$ is formed by the periodic extension (with period $M_f\Delta t$) of the values $x_T(\ell\Delta t)$, $0 \leq \ell \leq M_f-1$. If $M_f\Delta t < T$ the function $x_{T_a}(\ell\Delta t)$ is still a periodic function with period $M_f\Delta t$ but it must be computed within the fundamental period by aliasing the function $x_T(\ell\Delta t)$. As suggested by Cooley et al. (1967) it is possible to obtain $a_{T_a}(m\Delta f)$ in this case by carrying out this aliasing operation. This may be of interest in the case of the Fourier transform pair formed by the autocovariance $\gamma_{XX}(u)$ and spectrum $\Gamma_{XX}(f)$, where one may want to calculate the spectrum at a number of points less than the number M given by $M\Delta t = T$, to avoid obtaining a frequency resolution greater than necessary and thus eliminate unnecessary computing. The aliasing operation on the autocovariance must then be performed first. The reader is referred to the aforementioned paper for a more detailed discussion of this point.

The relationship between the functions $a_T(f)$ and $a_{T_a}(f)$ is the same as that between the spectrum $\Gamma_{XX}(f)$ and its aliased version $\Gamma_a(f)$, illustrated in Chapter 2. If the sampling interval is chosen such that $\Delta t \leq 1/(2f_{max})$, where f_{max} denotes the frequency above which $a_T(f)$ is negligible, then the aliasing in the frequency domain is negligible and we can write $a_T(f) = a_{T_a}(f)$, $-1/2\Delta t \leq f \leq 1/2\Delta t$. [The functions $a_T(f)$ and $a_{T_a}(f)$ are still different outside this interval.]

Equation (57) shows that the functions $(M_f\Delta t/T)\, x_{T_a}(\ell\Delta t)$, $\ell = 0,1,\ldots, M_f-1$, and $a_{T_a}(m\Delta f)$, $m = 0,.,\ldots, M_f-1$, with $M_f\Delta f = 1/\Delta t$, are a discrete finite Fourier transform pair corresponding to the Fourier integral transform pair $(1/T)\, x_T(t)$ and $a_T(f)$. We have

$$a_{T_a}(m\Delta f) = \frac{1}{M_f} \sum_{\ell=0}^{M_f-1} \frac{M_f\Delta t}{T}\, x_{T_a}(\ell\Delta t)\, e^{-j2\pi \frac{m\ell}{M_f}} \qquad (58)$$

and the inverse relation

---

*For $T = M_f\Delta t$ this is also the case except that

$$x_{T_a}(0) = x_{T_a}(M_f\Delta t) = \frac{x(0) + x(T)}{2}$$

(See footnote p. 27). In practice one can take $x_{T_a}(0)$ with negligible error.

$$\frac{M_f \Delta t}{T} x_{T_a}(\ell \Delta t) = \sum_{m=0}^{M_f-1} a_{T_a}(m\Delta f)\, e^{j2\pi \frac{m\ell}{M_f}} \tag{59}$$

In the following, for brevity, we will write simply $a_{T_a}(m)$ and $x_{T_a}(\ell)$ for, respectively, $a_{T_a}(m\Delta f)$ and $x_{T_a}(\ell\Delta t)$. We take in addition $T = M\Delta t$* [note the simplification in Eqs. (58) and (59)]. The sample spectrum defined by Eq. (50) can be expressed at $f = m\Delta f$ in terms of $a_{T_a}(m)$. We define

$$C_{xx_a}(m) = (M\Delta t)\, a_{T_a}(m)\, a^*_{T_a}(m) \tag{60}$$

$$= \frac{\Delta t}{M} \sum_{\ell=0}^{M-1} \sum_{\ell'=0}^{M-1} x_{T_a}(\ell)\, x_{T_a}(\ell')\, e^{-j2\pi m(\ell-\ell')/M}$$

that is, $C_{xx_a}(m)$ is the (periodic) sample spectrum obtained from the aliased function $a_{T_a}(m)$. Introducing the change of variables $n = \ell - \ell'$, $n' = \ell'$, we obtain

$$C_{xx_a}(m) = \Delta t \left[\sum_{n=0}^{M-1} \frac{1}{M} \sum_{n'=0}^{M-1-n} x_{T_a}(n')\, x_{T_a}(n'+n)\, e^{-j2\pi mn/M} \right.$$

$$\left. + \sum_{n=-(M-1)}^{-1} \frac{1}{M} \sum_{n'=-n}^{M-1} x_{T_a}(n')\, x_{T_a}(n'+n)\, e^{-j2\pi mn/M} \right]$$

or

$$C_{xx_a}(m) = \Delta t \left[\sum_{n=0}^{M-1} c_{xx_a}(n) \cdot e^{-j2\pi mn/M} + \sum_{n=-(M-1)}^{-1} c_{xx_a}(-n)\, e^{-j2\pi mn/M} \right] \tag{61}$$

---

*The number M_f of frequency points resulting from the choice $T = M\Delta t$ (where T is the record length and Δt the sampling interval) is the smallest possible to avoid aliasing in the time domain. The corresponding frequency spacing is $\Delta f = F/M = 1/M\Delta t = 1/T$. It can be shown using expressions derived later that at this spacing the correlation between neighboring (smoothed) spectral estimators is of the order of 0.2 for the Tukey and Bartlett windows and about 0.4 for the Parzen window (the correlation at double this spacing is practically zero for the Tukey window and less than 0.05 for the other two windows). Smaller frequency spacings (larger values of M) were found not to be necessary for the present applications. See, however, Jenkins and Watts (1968), pp. 260 and 283.

where

$$c_{xx_a}(n) = \begin{cases} \frac{1}{M} \sum_{n'=0}^{M-1-|n|} x_{T_a}(n')\, x_{T_a}(n'+|n|) , & 0 \leq |n| \leq M-1 \\ 0 , & |n| > M-1 \end{cases} \tag{62}$$

is the sample autocovariance function obtained from the aliased periodic realization $x_{T_a}(\ell)$. [For $M_f\Delta t \geq T$ the functions $x_{T_a}(\ell)$ and $x_T(\ell)$ coincide as explained for $0 \leq \ell \leq M_f-1$; the subscript a could therefore be dropped in Eq. (62)]. Equation (61) can be written in the equivalent forms [note that $c_{xx_a}(M) = 0$]

$$C_{xx_a}(m) = \Delta t \sum_{n=-M}^{M-1} c_{xx_a}(n)\, e^{-j2\pi mn/M} \tag{63}$$

or

$$C_{xx_a}(m) = \Delta t \sum_{n=0}^{M-1} [c_{xx_a}(n) + c_{xx_a}(M-n)]\, e^{-j2\pi mn/M} \tag{64}$$

Equation (63) corresponds to Eq. (51), which is its integral counterpart. Since we have chosen Δt so that there is no aliasing in the frequency domain, $C_{xx_a}(m) = C_{xx}(f)$ at $f = m\Delta f$, $m = 0,1,\ldots,M-1$, and

$$\lim_{T\to\infty} E[C_{XX_a}(m)] = \Gamma_{XX}(m\Delta f) \tag{65}$$

[See Eq. (53)]. Note that now the definition of the sample autocovariance function, Eq. (62), corresponds to the biased autocovariance function estimator defined in Chapter 3.

Remark. Applying the sampling theorem of Eqs. (58) and (59) to the Fourier integral transform pair $C_{xx}(f)$ and $c^{int}_{xx}(u)$ [see Eq. (51)] one obtains

$$C^{a}_{xx}(m\Delta f) = \frac{1}{M}\sum_{\ell=0}^{M-1} (M\Delta t)\; c^{int}_{xx_a}(\ell\Delta t)\; e^{-j2\pi m\ell/M} \qquad (66)$$

$$m = 0,1,\ldots,M-1$$

where $C^{a}_{xx}(m\Delta f)$ and $c^{int}_{xx_a}(\ell\Delta t)$ denote respectively the aliased periodic versions of the sample spectrum defined by Eq. (50) and of the sample autovariance defined by Eq. (47). If we assume as before that Δt is sufficiently small so that the aliasing in the frequency domain is negligible, then $C^{a}_{xx}(m\Delta f) = C_{xx}(f)$ at $f = m\Delta f$. The non-zero extent of the function $c^{int}_{xx}(u)$, on the other hand, is $2T$. Since we have chosen $M\Delta t = T$ (see footnote p. 30) we must carry out the aliasing operation in the time domain as explained to obtain $c^{int}_{xx_a}(\ell\Delta t)$ [see discussion following Eq. (57)]. Thus

$$c^{int}_{xx_a}(\ell\Delta t) = c^{int}_{xx}(\ell\Delta t) + c^{int}_{xx}\,(M-\ell)\Delta t \quad , \quad 0 \leq \ell \leq M-1 \qquad (67)$$

Equations (66) and (67) yield

$$C^{a}_{xx}(m\Delta f) = \Delta t \sum_{\ell=0}^{M-1} \{c^{int}_{xx}(\ell\Delta t) + c^{int}_{xx}[(M-\ell)\Delta t]\}\; e^{-j2\pi m\ell/M} \qquad (68)$$

which is the analog of Eq. (64). The corresponding analog of Eq. (63) follows immediately,

$$C^{a}_{xx}(m\Delta f) = \Delta t \sum_{\ell=-M}^{M-1} c^{int}_{xx}(\ell\Delta t)\; e^{-j2\pi m\ell/M} \qquad (69)$$

The difference between the results (63) and (64) on the one hand, and (68) and (69) on the other, is that while in Eqs. (68) and (69) the function $c^{int}_{xx}(\ell\Delta t)$ is defined by the integral (47) calculated at $u = \ell\Delta t$, in the first two equations $c_{xx_a}(n)$ is defined by the sum (62). The four equations represent exact results.[a] If the aliasing in the functions $a_{T_a}(m\Delta f)$ [and therefore $C_{xx_a}(m)$] and $C^{a}_{xx}(m\Delta f)$ is negligible, we have $C_{xx_a}(m) = C^{a}_{xx}(m\Delta f) = C_{xx}(f)$ at $f = m\Delta f$.

4.4. The Circular Autocovariance Function. Equation (64) is as Eq. (58) in the form of a finite Fourier transform, and the properties of this transform can then be used in all the calculations (see e.g., Cooley et al. 1967). The functions $C_{xx_a}(m)$, $m = 0,1,\ldots,M-1$, and $(M\Delta t)$ $[c_{xx_a}(n) + c_{xx_a}(M-n)$, $n = 0,1,\ldots,M-1]$, are a finite Fourier transform pair $(M\Delta t = T)$. The sum in the brackets is known as the sample circular autocovariance function and is denoted by $\tilde{c}_{xx_a}(n)$.

To obtain from the sample spectrum, $C_{xx_a}(m)$, the sample autocovariance function rather than the circular sample autocovariance function we define a modified process, $X_{0a}(\ell)$, by means of the sequence

$$x_{0_a}(\ell) = \begin{cases} x_{T_a}(\ell) , & \ell = 0,1,2,\ldots,M-1 \\ 0 , & \ell = M,M+1,\ldots,2M-1 \end{cases} \tag{70}$$

The equation corresponding to Eq. (58) is now

$$a_{0a}(m) = \frac{1}{2M} \sum_{\ell=0}^{2M-1} x_{0a}(\ell)\, e^{-j2\pi m\ell/2M} \tag{71}$$

The sample spectrum of the modified process is given by

$$C_{xx_{0a}}(m) = 2\,(M\Delta t)\, a_{0a}(m)\, a^{*}_{0a}(m) \tag{72}$$

and the equation corresponding to Eq. (64) is

$$C_{xx_{0a}}(m) = \Delta t \sum_{n=0}^{2M-1} \left[c_{xx_{0a}}(n) + c_{xx_{0a}}(2M-n) \right] e^{-j2\,mn/2M}$$

The (inverse) finite Fourier transform of $C_{xx_{0a}}(m)$ is the sample circular autocovariance function

$$\tilde{c}_{xx_{0a}}(n) = c_{xx_{0a}}(n) + c_{xx_{0a}}(2M-n)$$

multiplied by $2M\Delta t$. The sample periodic autocovariance function $c_{xx_{0a}}(n)$ is given by

$$c_{xx_{0a}}(n) = \frac{1}{2M} \sum_{k=0}^{2M-1-n} x_{0a}(k)\, x_{0a}(k+|n|)$$

But for $0 \leq n \leq M-1$ we have

$$2c_{xx_{0a}}(n) = c_{xx_a}(n)$$

and

$$c_{xx_{0a}}(2M-n) = 0$$

Therefore

$$(2M\Delta t)\, \tilde{c}_{xx_{0a}}(n) = (M\Delta t)\, c_{xx_a}(n)$$

that is, the inverse finite Fourier transform of the sample spectrum $C_{xx_{0a}}(m)$ defined by Eq. (72) yields, for $0 < |n| \leq M-1$, the values of the sample autocovariance function $c_{xx_a}(n)$ multiplied by $(M\Delta t)$.

4.5. Smoothing of Spectral Estimators. The sample spectrum, $C_{xx}(f)$, defined by Eq. (50) is a realization of the random variable that we have denoted by $C_{XX}(f)$ and whose expected value, given by Eq. (52), tends to $\Gamma_{XX}(f)$ as $T \to \infty$. An estimate of the probable difference between the sample spectrum and its expected value can be obtained in the form of confidence intervals, with the variance of the process $C_{XX}(f)$ giving an idea of the spread of the sample values $C_{xx}(f)$ about the mean. On the other hand the $E[C_{XX}(f)]$, given by Eq. (52), is different from $\Gamma_{XX}(f)$ for a finite record length T and the difference or bias must also be estimated.

It can be shown (Jenkins and Watts, 1968, pp. 412-418) that the covariance between the sample spectral estimators defined by Eq. (50), at frequencies f_1 and f_2, is given approximately by

$$\mathrm{Cov}[C_{XX}(f_1), C_{XX}(f_2)] \approx$$

$$\Gamma_{XX}^2(f) \left\{ \left[\frac{\sin \pi T(f_1+f_2)}{\pi T(f_1+f_2)} \right]^2 + \left[\frac{\sin \pi T(f_1-f_2)}{\pi T(f_1-f_2)} \right]^2 \right\} \qquad (73)$$

where $f_1 \leq f \leq f_2$. In the derivation of this result it is assumed that the spectral density function of the process is approximately constant in the range from f_1 to f_2. In addition, a contribution due to the fourth comulant of the process is assumed to be negligible. The latter assumption can be justified for a general class of non-normal processes; for a normal process the fourth cumulant is identically zero and Eq. (73) is therefore exact for normal white noise processes.* Equation (73) shows that for given $f_1 \neq f_2$ and T large the covariance between two spectral estimators is of order $1/T^2$ and that, to a good degree of approximation, spectral estimators at spacing greater than $1/T$ may be regarded as uncorrelated. [The covariance estimate (73) is actually identically zero when the frequencies $(f_1 \pm f_2)$ are multiplies of $1/T$]. On the other hand, for $f_1 = f_2$ and T large, we have

$$\mathrm{Var}[C_{XX}(f)] \approx \Gamma_{XX}^2(f) \tag{74}$$

For the expected value of the sample spectrum estimator, on the other hand, we have from Eq. (52)

$$E[C_{XX}(f)] = \Gamma_{XX}(f) - \int_{-T}^{T} \frac{|u|}{T} \gamma_{XX}(u)\, e^{-j2\pi fu}\, du \tag{75}$$

If T is large enough so that $\gamma_{XX}(u)$ is negligible for values of the delay times, u, of the order of T we have approximately

$$E[C_{XX}(f)] \approx \Gamma_{XX}(f) \tag{76}$$

Equation (74) shows clearly that the sample spectrum $C_{XX}(f)$ defined by Eq. (50) does not converge in any statistical sense to a limiting value as the record length, T, tends to infinity. Indeed, the variance of the estimator $C_{XX}(f)$ is a constant independent of T (the sample size) and sample spectra obtained from records of different length will therefore exhibit the same erratic behavior. If we define a coefficient of variability for the sample spectral estimator as the ratio of the root-mean-square value of the variance of the estimator to its expected value, we obtain for $C_{XX}(f)$

*An exact result valid for normal processes (without assuming $\Gamma_{XX}(f)$ constant between f_1 and f_2) is also given in the reference cited.

a coefficient of variability of 100%, independent of T. This renders $C_{XX}(f)$ useless for estimation purposes and shows that smoothing procedures, designed to reduce the variance of the estimator (at the expense as will be seen presently of increasing the bias), are needed if Eqs. (50) and (60) are to be used for spectral estimation.

A method to reduce the variance of the sample spectral estimator, due to Bartlett (1950), has already been introduced in Chapter 3.. In this method the time series is subdivided into K smaller stochastically independent time series and the sample spectrum is obtained by averaging over all the subseries. The variance of the averaged spectral estimator $\overline{C}_{XX}(f)$ is given by

$$\mathrm{Var}[\overline{C}_{XX}(f)] = \frac{\Gamma_{XX}^2(f)}{K}$$

The method is a particular case of a more general method due to Welch (1967) in which the average is taken over K overlapping subseries.

Let $M\Delta t$ be the length of each of the K subseries and $N\Delta t$ the total record length ($N = KM$). The averaged biased sample autocovariance function is given by [cf. Eq. (23)]

$$\overline{c}_{xx_a}(\ell) = \frac{1}{K} \sum_{k=1}^{K} \frac{1}{M} \sum_{i=(k-1)M}^{kM-1-|\ell|} x_{T_a}(i)\, x_{T_a}(i+|\ell|) \qquad 0 \leq |\ell| \leq M-1 \tag{77}$$

The results of the preceding chapters can be applied to each subseries without any change in notation. We have in particular

$$E[\overline{c}_{xx_a}(\ell)] = \left(1 - \frac{|\ell|}{M}\right) \gamma_{XX}(\ell\Delta t) \quad , \qquad \ell = 0,1,\ldots,M-1 \tag{78}$$

and

$$E[\overline{C}_{XX_a}(m)] = \Delta t \sum_{\ell=-(M-1)}^{M-1} \left(1 - \frac{|\ell|}{M}\right) \gamma_{XX}(\ell\Delta t)\, e^{-j2\pi m\ell/M} \quad , \tag{79}$$

$$m = 0,1,\ldots,M-1$$

In continuous time we have, instead of (47),

$$\bar{c}_{xx}^{int}(u) = \frac{1}{K} \sum_{k=1}^{K} \frac{1}{M} \int_{(k-1)M\Delta t}^{kM\Delta t-|u|} x(t)\, x(t+|u|)\, dt, \quad |u| \leq M\Delta t$$

and the expected values corresponding to those given by Eqs. (78) and (79) are given by Eqs. (49) and (52) with $T = M\Delta t$. If we assume $N\Delta t$ large enough so that when applying these equations to the entire record, that is with $K = 1$ and $M = N$, one can neglect $(|u|/T)\gamma_{XX}(u)$ in Eqs. (49) and (52), and $(|\ell|/N)\gamma_{XX}(\ell\Delta t)$ in Eqs. (78) and (79), one sees that Bartlett's procedure is essentially equivalent to multiplying the sample autocovariances $c_{xx}^{int}(u)$ and $c_{xx_a}(\ell)$, defined respectively by Eq. (47) with $T = N\Delta t$ and Eq. (62) with $M = N$, by the lag window

$$w_B(u) = \begin{cases} 1 - \dfrac{|u|}{M\Delta t} & |u| \leq M\Delta t \\ 0 & |u| > M\Delta t \end{cases} \tag{80}$$

in the continuous case, and by the lag window obtained by replacing u by $\ell\Delta t$ in (80), in the discrete case.*

Bartlett's smoothing procedure leads then naturally to the consideration of lag windows for the purpose of reducing the large variance of the sample spectrum estimator $C_{xx}(f)$. We consider in general functions $w(u)$ satisfying the conditions

$$w(0) = 1 \quad , \qquad w(u) = w(-u)$$

and (81)

$$w(u) = 0 \quad , \qquad |u| \geq T_0$$

and smoothed sample spectra defined by

---

*The difference between Eq. (62) with $M = N$ and Eq. (77), for a given lag $\ell\Delta t$, is that in (77) the summation contains $K(M-|\ell|)$ lagged products, whereas the sum in (62) contains all of these plus those formed with data points with the same lag but in different (consecutive) subseries, the total number being $N - |\ell| \sim N = kM$.

$$\bar{C}_{xx}(f) = \int_{-\infty}^{\infty} w(u)\, c_{xx}^{int}(u)\, e^{-j2\pi fu}\, du = \int_{-\infty}^{\infty} \bar{c}_{xx}(u)\, e^{-j2\pi fu}\, du \quad (82)$$

Applying the theorem of Eqs. (58) and (59) to the transform pair $\bar{C}_{xx}(f)$, $\bar{c}_{xx}(u)$, we obtain [Cf. Eqs. (68) and (69)]

$$\bar{C}_{xx}^{a}(m\Delta f) = \Delta t \sum_{\ell=-M}^{M-1} w(\ell\Delta t)\, c_{xx}^{int}(\ell\Delta t)\, e^{-j2\pi m\ell/M} \quad (83)$$

where $M\Delta t = T$ and Δt is chosen as before small enough so that the aliasing in the frequency domain is negligible and therefore $\bar{C}_{xx}^{a}(m\Delta f) = \bar{C}_{xx}(f)$ at $f = m\Delta f$. The window $w(\ell\Delta t)$ or simply $w(\ell)$ can also be applied on the autocovariance defined by the discrete sum (62). We have

$$\bar{C}_{xx_a}(m) = \Delta t \sum_{\ell=-M}^{M-1} w(\ell)\left[c_{xx_a}(\ell)\, e^{-j2\pi m\ell/M}\right] \quad (84)$$

Since both (63) and (69) yield $C_{xx}(f)$ at $f = m\Delta f$, we have also*

$$\bar{C}_{xx_a}(m\Delta f) = \bar{C}_{xx}^{a}(m\Delta f)\left[= \bar{C}_{xx}(f) \text{ at } f = m\Delta f\right] .$$

In the analysis below, therefore, we will use the continuous time expression (82) to discuss the properties of the smoothed sample spectrum, as was

---

*This can be readily shown by using a convolution result for the finite transformation defined by any one of the four equations (63), (69), (83) or (84). We have for example

$$\bar{C}_{xx_a}(m) = \frac{1}{2T} \sum_{n=-M}^{M-1} W(m-n)\, C_{xx_a}(n)$$

where

$$W(m-n) = \Delta t \sum_{k=-M}^{M-1} w(k)\, e^{-j\frac{2\pi}{M}(m-n)k}$$

and

$$C_{xx_a}(n) = \Delta t \sum_{\ell=-M}^{M-1} c_{xx_a}(\ell)\, e^{-j\frac{2\pi}{M}\ell n}$$

Eq. (51) already used at the beginning of this section in reviewing the properties of the sample spectrum $C_{xx}(f)$.

A comprehensive account of the properties discussed in the remaining of this chapter can be found in Jenkins and Watts (1968). The following paragraphs are therefore intended as brief summaries of results needed in the following chapter for discussion of the data. Let $W(f)$ be the Fourier transform of the lag windows $w(u)$, that is,*

$$W(f) = \int_{-\infty}^{\infty} w(u)\ e^{-j2\pi fu}\ du \tag{85}$$

Figure 4.1 shows the rectangular, Bartlett, Tukey and Parzen lag windows introduced in Chapter 3, and the corresponding spectral windows $W(f)$. The smoothed spectral estimator $\bar{C}_{xx}(f)$ is given by

$$\bar{C}_{xx}(f) = \int_{-\infty}^{\infty} C_{xx}(g)\ W(f-g)\ dg \tag{86}$$

Using Eqs. (73) and (86) it can be shown that the covariance between smoothed spectral estimators is given approximately by

$$\mathrm{Cov}[\bar{C}_{xx}(f_1), \bar{C}_{xx}(f_2)] \approx \frac{\Gamma_{xx}^2(f)}{T_r} \int_{-\infty}^{\infty} W(f_1-g)[W(f_2+g) + W(f_2-g)]\ dg\ , \quad f_1 \leq f \leq f_2 \tag{87}$$

provided that T_r is large enough so that the $(\sin \pi fT_r/\pi fT_r)^2$ terms in (73) behave as delta functions, and that $\Gamma_{xx}(f)$ is a smooth function over the width of the spectral window $W(f)$. Equation (87) shows that the covariance between smoothed spectral estimators at f_1 and f_2 is proportional to the

*Note that the theorem of Eqs. (58) and (59) applied to (85) yields

$$W_a(m) = \Delta t \sum_{\ell=0}^{M-1} w(\ell\Delta t)\ e^{-j2\pi\frac{m\ell}{M}}$$

which is the relationship used in the preceding footnote.

amount of overlap between the spectral windows centered at f_1 and f_2. The variance of the smoothed spectral estimator is given approximately by

$$\mathrm{Var}[\bar{C}_{XX}(f)] \approx \Gamma_{XX}^2(f)\ \frac{1}{T_r}\int_{-\infty}^{\infty} W^2(g)\ dg = \Gamma_{XX}^2(f)\ \frac{1}{T_r}\int_{-\infty}^{\infty} w^2(u)\ du \qquad (88)$$

[compare with Eq. (74)] and can be written in the form

$$\mathrm{Var}[\bar{C}_{XX}(f)] \approx \Gamma_{XX}^2(f)\ \frac{1}{b\,T_r} \qquad (89)$$

where b is the bandwidth of the window, defined as the width of a rectangular spectral window which gives the same variance. For the windows described in Chapter 3, $b = a/T$, with $a = 1.33$ for the Tukey window, $a = 1.5$ for the Bartlett window, and $a = 1.86$ for the Parzen window.

For the expected value of the smoothed spectral estimator one has

$$E[\bar{C}_{XX}(f)] = \int_{-T}^{T} \gamma_{XX}(u) \left[1 - \frac{|u|}{T_r}\right] w(u)\ e^{-j2\pi fu}\ du \qquad (90)$$

where $T = M\Delta t$. Assuming T_r large enough so that $\gamma_{XX}(u)$ is negligible for values of the delay time, u, of the order of T_r one obtains for the bias $B(f)$ the approximate expression

$$B(f) \sim \int_{-\infty}^{\infty} [w(u)-1]\ \gamma_{XX}(u)\ e^{-j2\pi fu}\ du \qquad (91)$$

Substituting for the lag windows $w(u)$ defined at the end of Chapter 3 (with ℓ/M replaced by $u/T = u/M\Delta t$), and neglecting higher order terms in u/T, one can obtain simple approximate expressions for the bias of the Bartlett, Tukey, and Parzen windows. The conclusions that can be drawn from the resulting expressions are as follows. First, the bias of the Bartlett window is larger in general (for the same truncation point T) than the bias of the Tukey and Parzen windows, except near a peak, where the Bartlett window produces small bias. For both the Tukey and Parzen windows, spectral peaks tend to be underestimated and spectral troughs overestimated: the narrower the peak or the trough, the larger the bias. Secondly, again for the same

truncation point T, the bias of the Tukey window is somewhat smaller than that of the Parzen window. If these windows are compared with equal bandwidth $b = a/T$, however, they produce almost identical smoothed spectra. Thirdly, and evidently, the bias is reduced for a given window as M increases and the width of the spectral window decreases (but the variance increases, see Eq. (89)).

The somewhat larger overall bias of the Bartlett smoothing procedure for given T is due to the relatively large side lobes of the corresponding spectral window, which produce a distortion of the spectrum at frequency f due to leakage from distant frequencies (see Eq. (86) and Fig. 4.1). Near a peak, however, the bias is small. (The bias given by the approximate expression referred to in the preceeding paragraph is actually zero at the peak.) Furthermore, as discussed in Chapter 3, the biased autocovariance function estimator resulting from application of Bartlett's procedure is to be preferred over the unbiased estimator for estimation of the autocovariance function. Under these conditions, and for simplicity, the Bartlett procedure will be adopted in Chapter 5 for spectral estimation. Actually, as illustrated in more detail in Jenkins and Watts (1968) with several examples, it is the choice of bandwidth and not the choice of window that is important in spectral analysis. Other windows could certainly be used if desired. To do this one could, as suggested by Enochson and Otnes (1968, p. 123), remove the bias associated with the Bartlett window by multiplying the sample autocovariance $\bar{c}_{xx}(u)$ given by Eq. (62) by $T/(T - |u|)$ (see footnote p. 37) and then introduce the window chosen. It is noted here in passing that, as a general rule, to achieve high fidelity (small bias) the bandwidth of the window must be of the same order as the width of the narrowest important detail of the spectrum.

To conclude this Section, we discuss the problem of establishing confidence intervals for the spectrum. It can be shown (Jenkins and Watts 1968, pp. 416-417) that the sample spectrum estimator $C_{xx}(f)$ is approximately distributed as a chi-squared random variable with 2 degrees of freedom. (This result is exact for normal processes). Using (74) and (76) we have

$$C_{xx}(f) \sim \frac{\Gamma_{xx}(f)}{2} \chi^2_2 \qquad (92)$$

where the mean and variance of the χ_2^2 random variable are respectively 2 and 4. the smoothed sample spectral estimator C_{xx_a} is on the other hand a weighted sum of random variables $C_{xx_a}(u)$ (see footnote p. 38), each distributed as a χ_2^2 random variable. One can then approximate the distribution of $\bar{C}_{xx}(f)$ by the random variable

$$\bar{C}_{xx}(f) \sim a_0 \chi_\nu^2 \tag{93}$$

with

$$\nu = \frac{2(E[\bar{C}_{xx}(f)])^2}{Var[\bar{C}_{xx}(f)]}$$

and

$$a_0 = \frac{E[\bar{C}_{xx}(f)]}{\nu}$$

(The expected value and variance of a χ_ν^2 distribution are respectively ν and 2ν.) Using (89) and neglecting the bias in (90) we obtain

$$\nu = 2bT_r = 2a\frac{T_r}{T}, \quad a_0 = \frac{\Gamma_{xx}(f)}{\nu} \tag{94}$$

since $\nu\bar{C}_{xx}(f)/\Gamma_{xx}(f)$ is distributed (approximately) as a χ_ν^2, reading off a table limits $x_\nu(\alpha/2)$ and $x_\nu(1-\alpha/2)$ such that

$$Pr\{\chi_\nu^2 \leq x_\nu(\alpha/2)\} = \alpha/2$$

and

$$Pr\{\chi_\nu^2 \leq x_\nu(1-\alpha/2)\} = 1 - \alpha/2$$

one has

$$Pr\left(x_\nu(\alpha/2) \leq \frac{\nu\bar{C}_{xx}(f)}{\Gamma_{xx}(f)} \leq x_\nu(1-\alpha/2)\right) = 1 - \alpha$$

and therefore the interval

$$\frac{\nu\bar{C}_{xx}(f)}{x_\nu(1-\alpha/2)}, \qquad \frac{\nu\bar{C}_{xx}(f)}{x_\nu(\alpha/2)}$$

is a $[(1-\alpha)100]\%$ confidence interval for $\Gamma_{xx}(f)$.

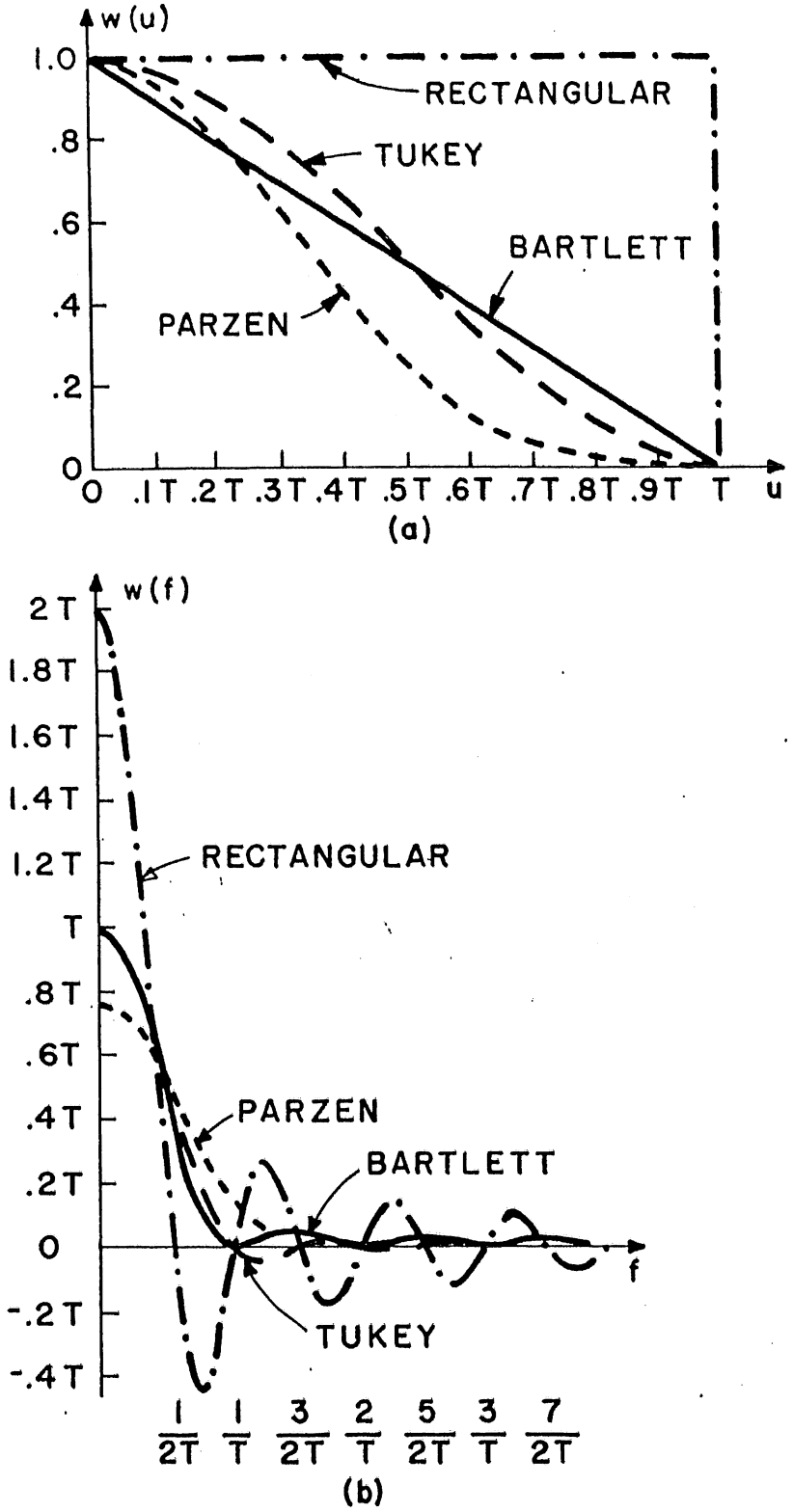


Fig. 4.1 Some common windows

5. AUTOREGRESSIVE PROCESSES

For purposes of comparison and discussion of results and for verification of the programs developed, first- and second-order autoregressive processes were used. These processes may be regarded as the outputs from first- and second-order linear systems (continuous or discrete) when the input is a purely random process (white noise signal). The autocovariance functions and spectra of autoregressive processes can be derived theoretically from the defining equations. Reference is made to Jenkins and Watts (1968) for the details of the derivations. Four types of spectra can be generated with second-order autoregressive processes. For the present application, the constants of the process were chosen so that the resulting spectra exhibited an intermediate peak.

To verify the validity of the technique used for calculation of the sample spectra, the exact spectra of the autoregressive processes were compared with the estimates obtained applying the spectral analysis programs developed to digitally generated realizations of the processes. The figures that follow show the effect, on the sample spectra of a second-order autoregressive process, of varying the number of points, M, in each subseries (subseries length $T = M\,\Delta t$), the number of subseries, K, and the total record length, $T_r = K\,T = K\,(M\Delta t)$. All sample spectra have been obtained using the same stochastic series (same data points). The theoretical spectrum is shown in each graph (continuous line). The sample spectral points are joined by straight lines.

The total number of points available in the series for data analysis was N = 12,288. Thus, for 128-point subseries, the maximum number of subseries that can be used is K = 96 (N = K x M = 96 x 128 = 12,288). Figures 5.1 and 5.2 show the sample spectra for M = 1024, K = 3, and M = 128, K = 96. Note the difference in the size of the confidence intervals plotted in each figure, and the difference in the bias of the spectral estimates. Figures 5.2 through 5.5 show the effect of varying K and M keeping the total number of points, N = K x M = 12,288, constant. Figures 5.5 through 5.8 show the effect of varying M while keeping the number of subseries, K, constant equal to 12.

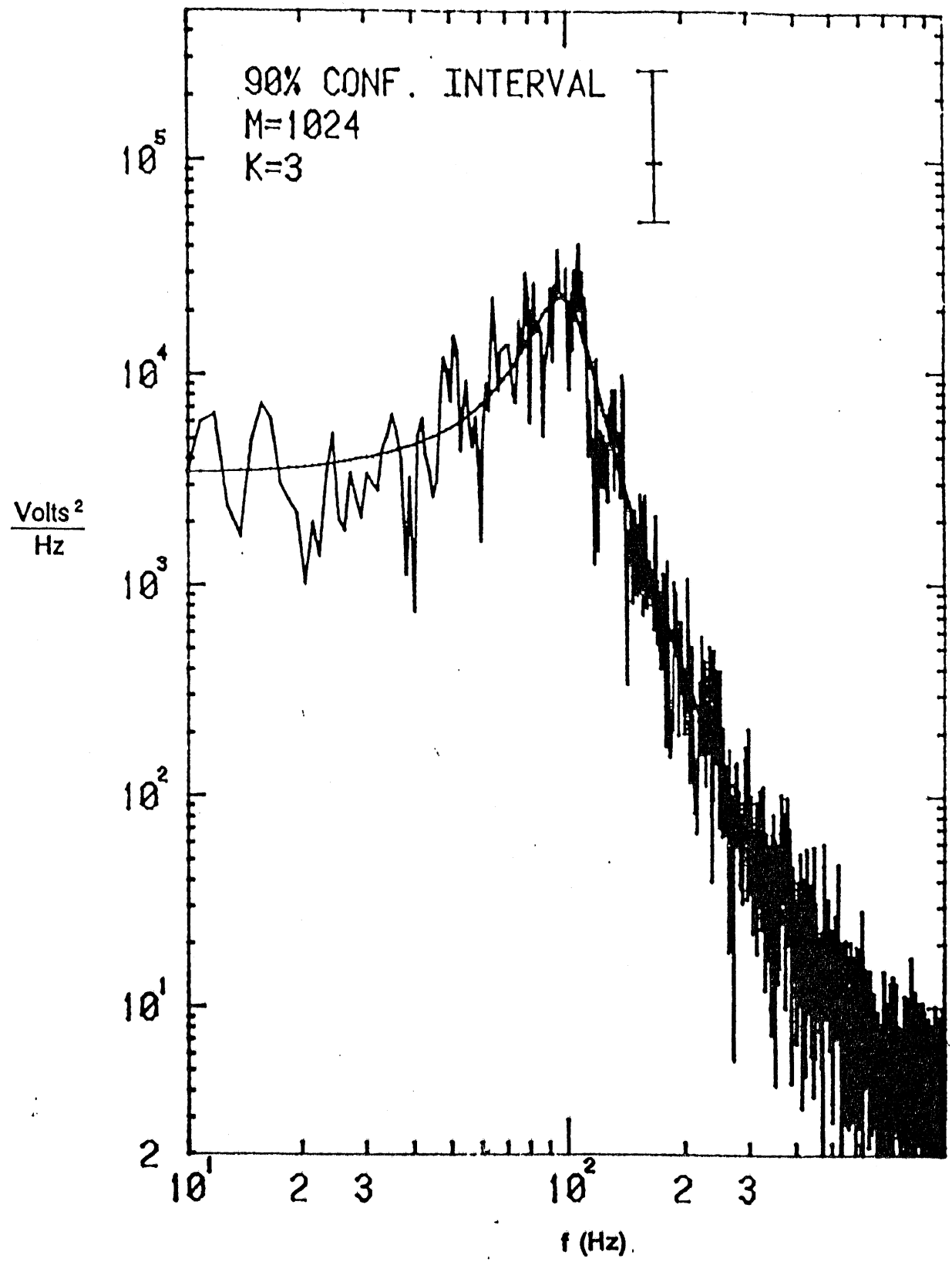


Fig. 5.1 Sample spectrum (M = 1024, K = 3, N = 3072)

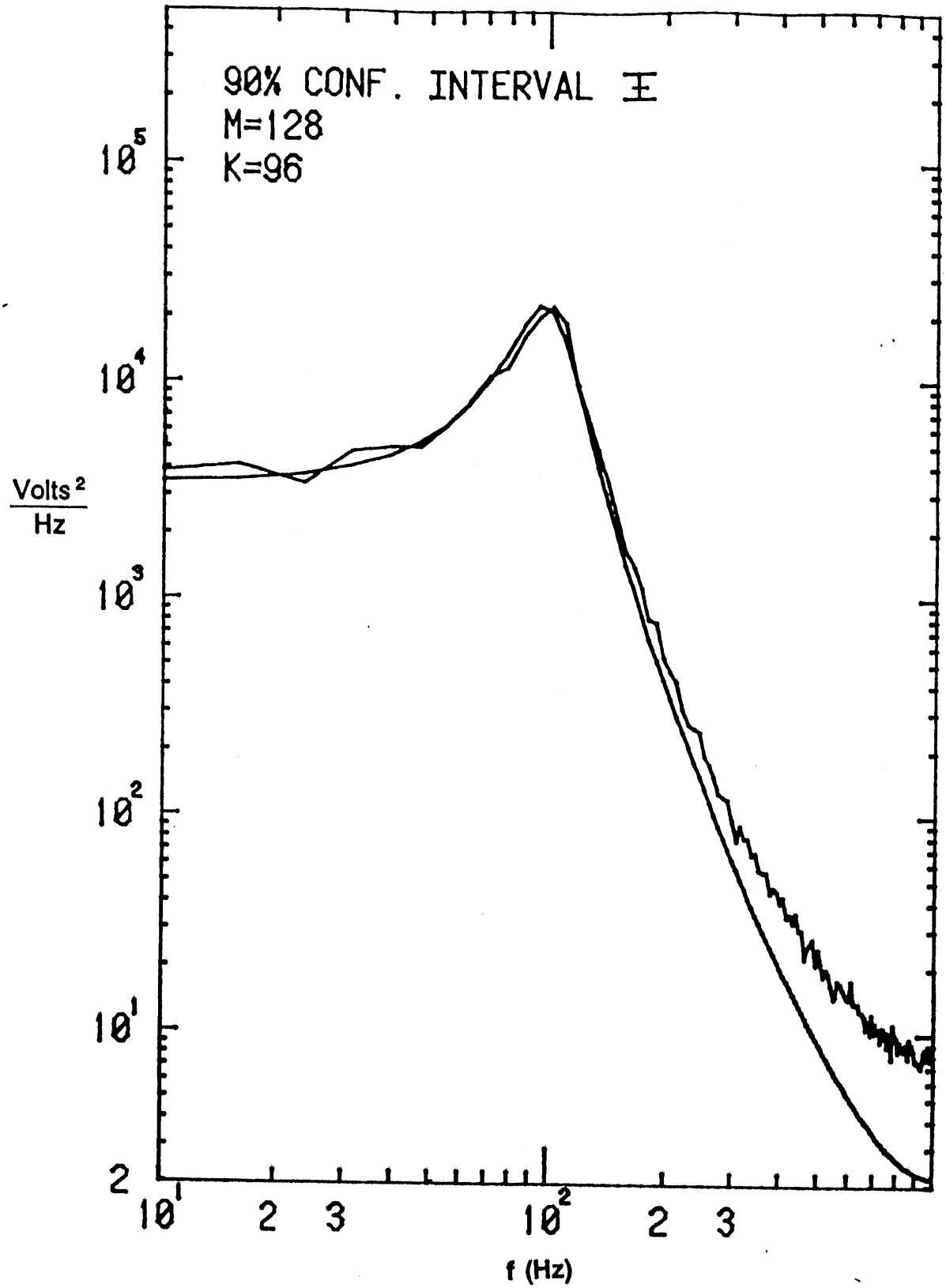


Fig. 5.2 Sample spectrum (M = 128, K = 96, N = 12,288)

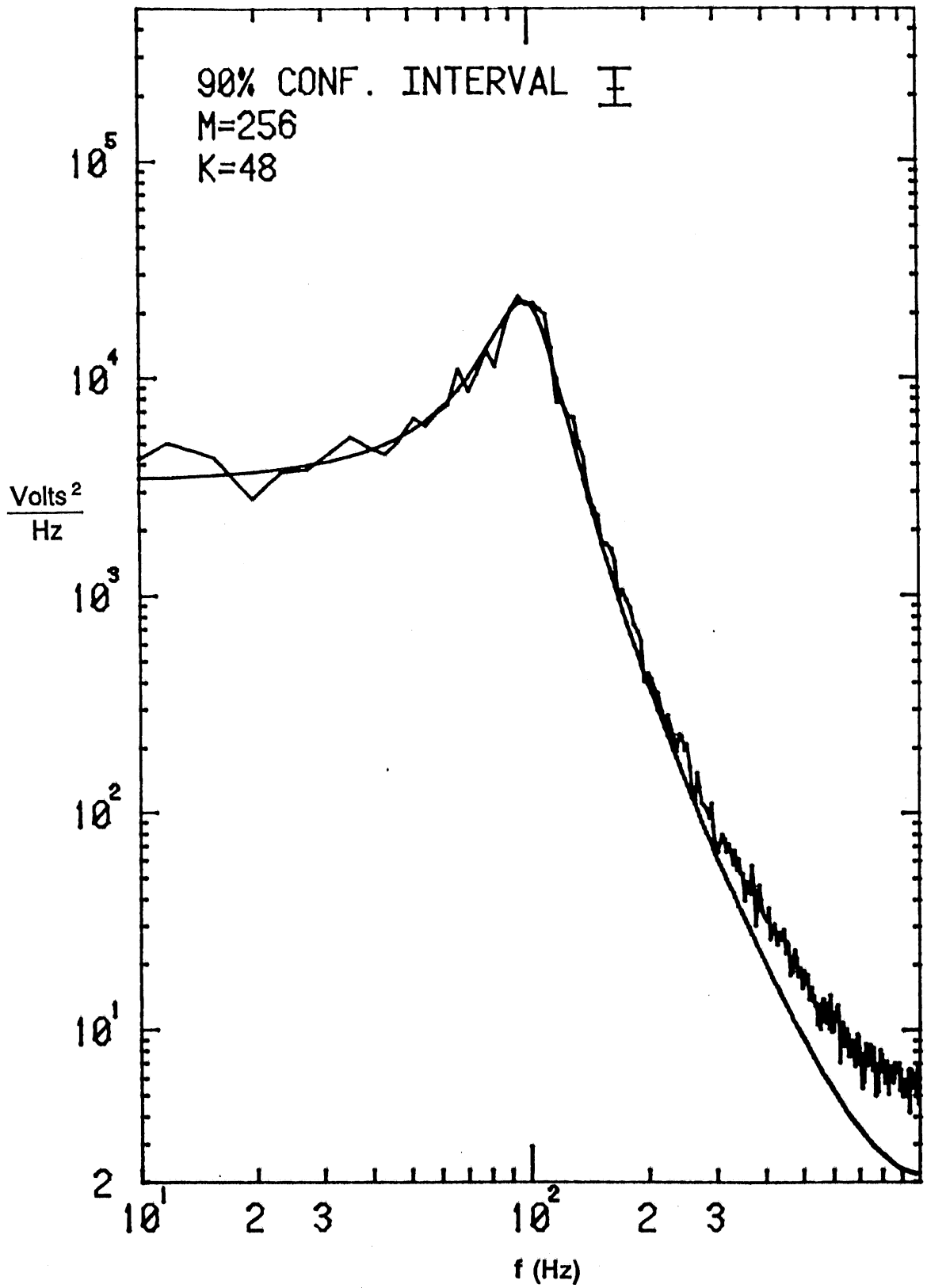


Fig. 5.3 Sample spectrum (M = 256, K = 48, N = 12,288)

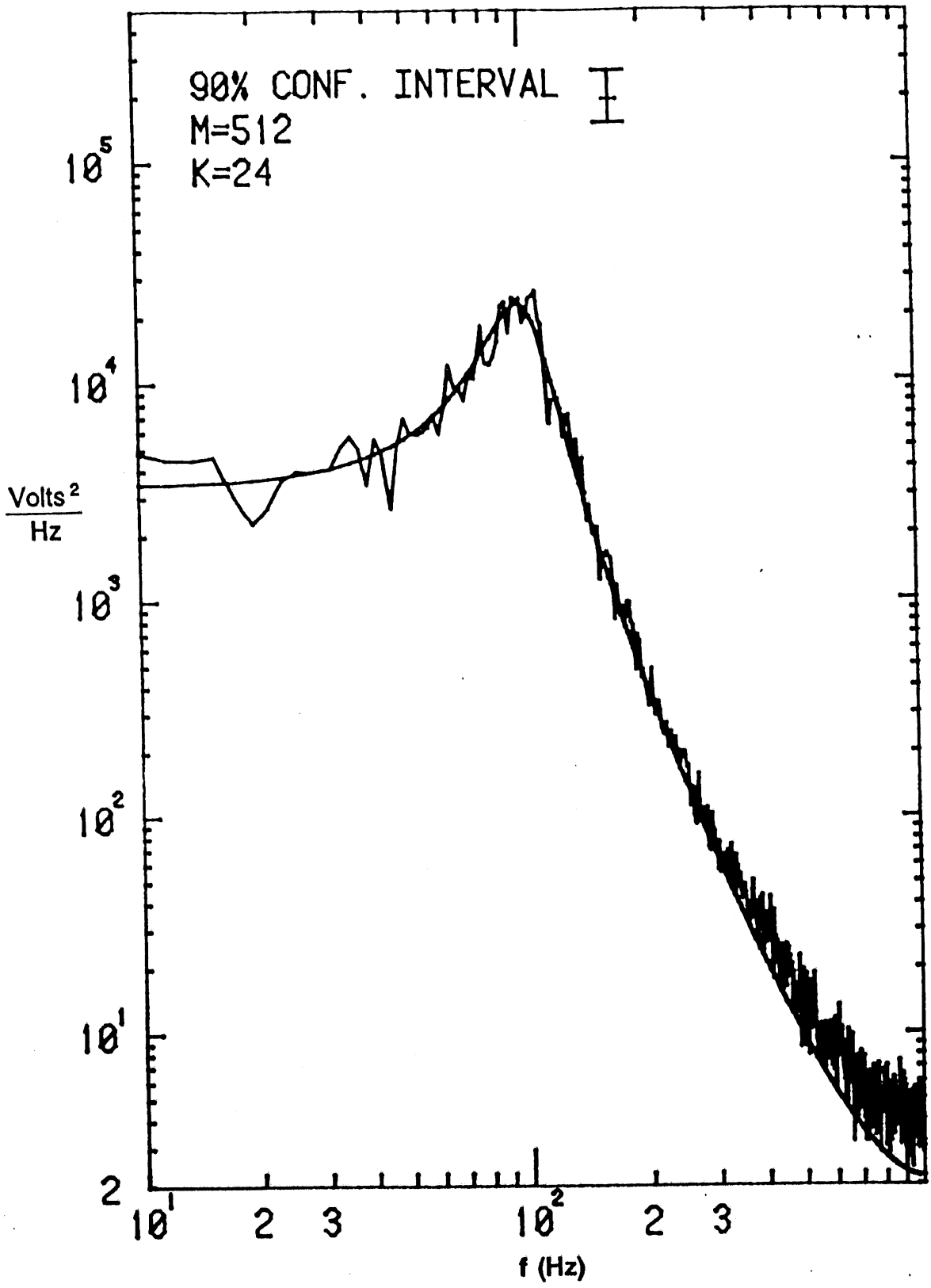


Fig. 5.4 Sample spectrum (M = 512, K = 24, N = 12,288)

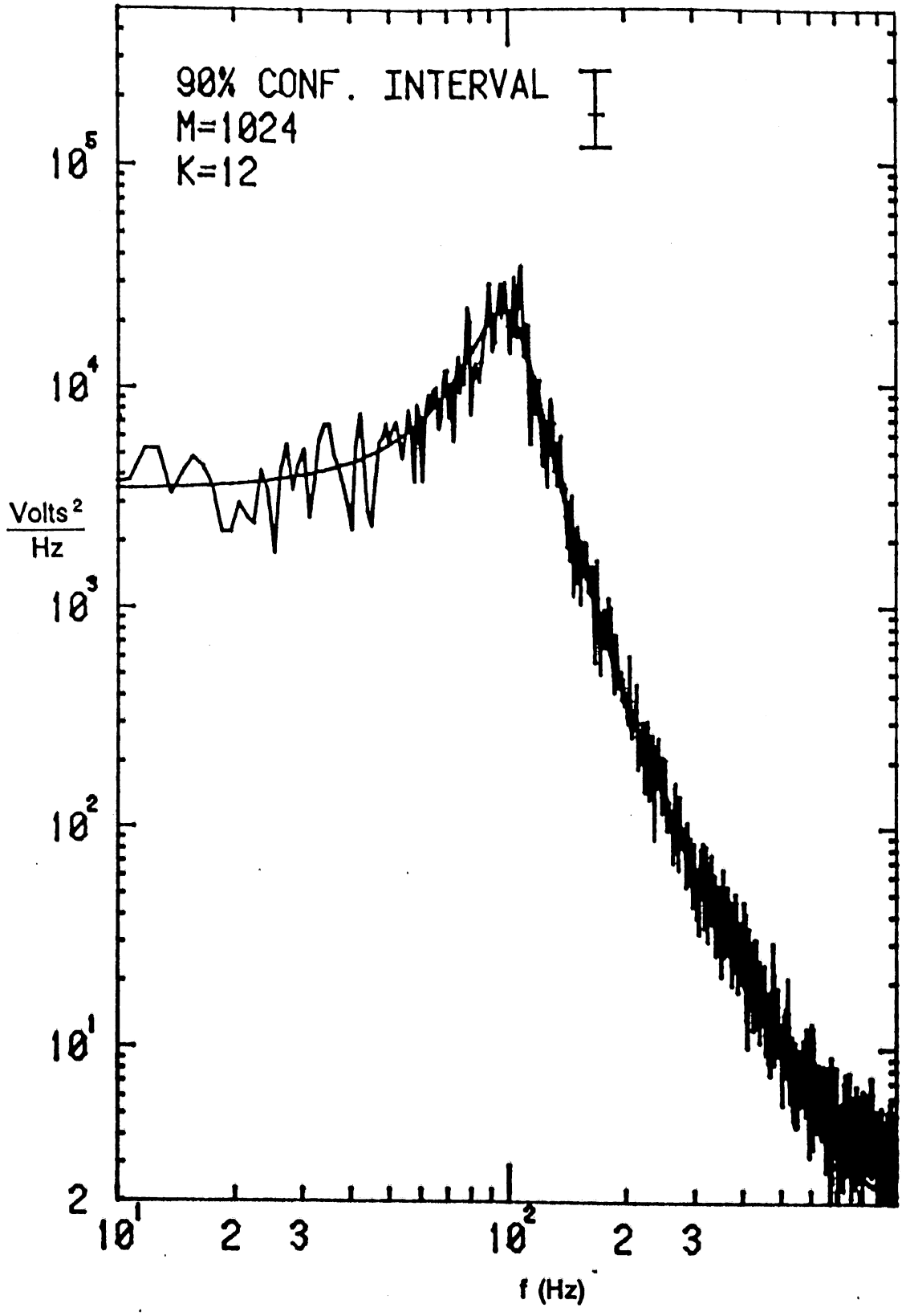


Fig. 5.5 Sample spectrum (M = 1024, K = 12, N = 12,288)

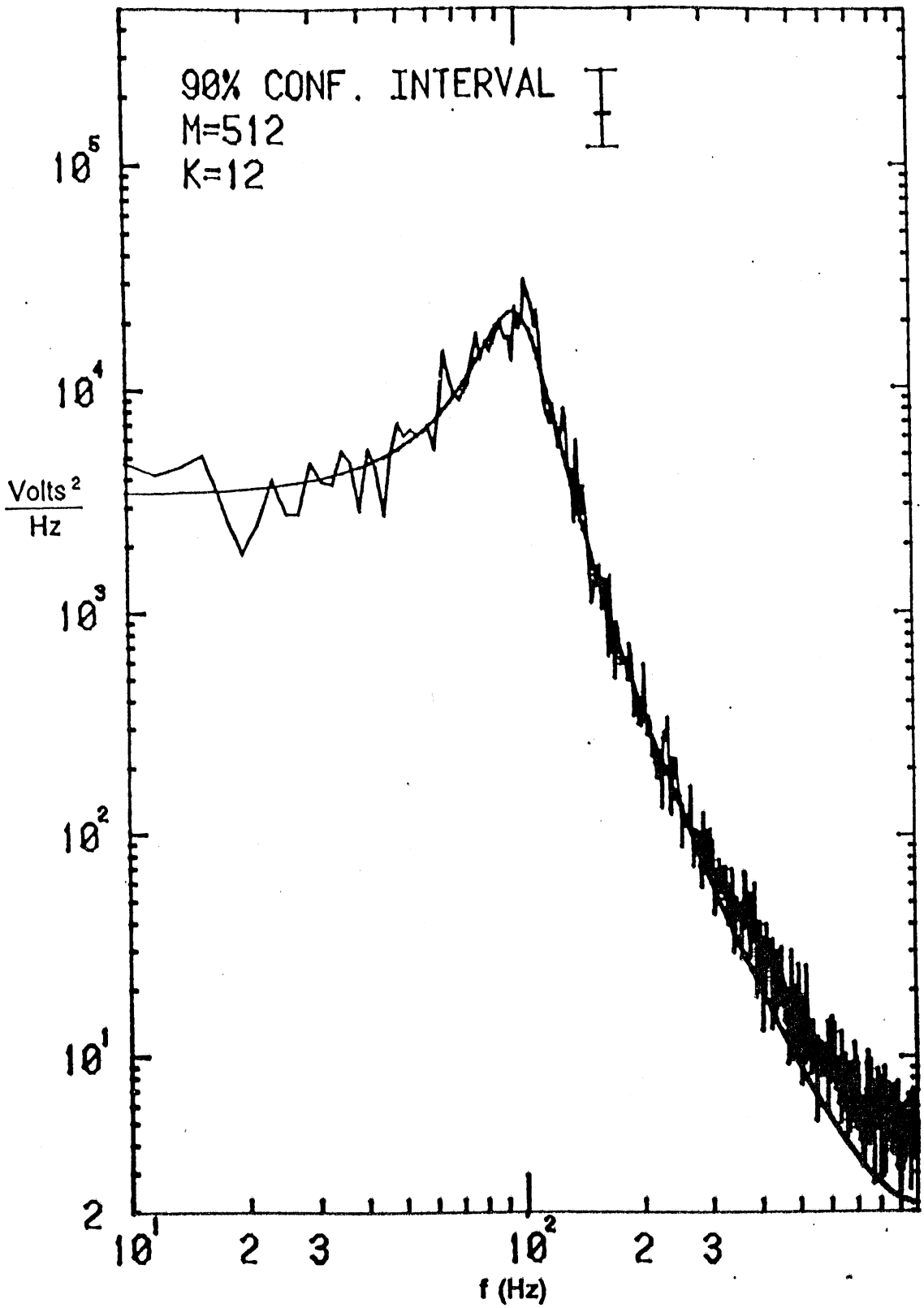


Fig. 5.6 Sample spectrum (M = 512, K = 12, N = 6144)

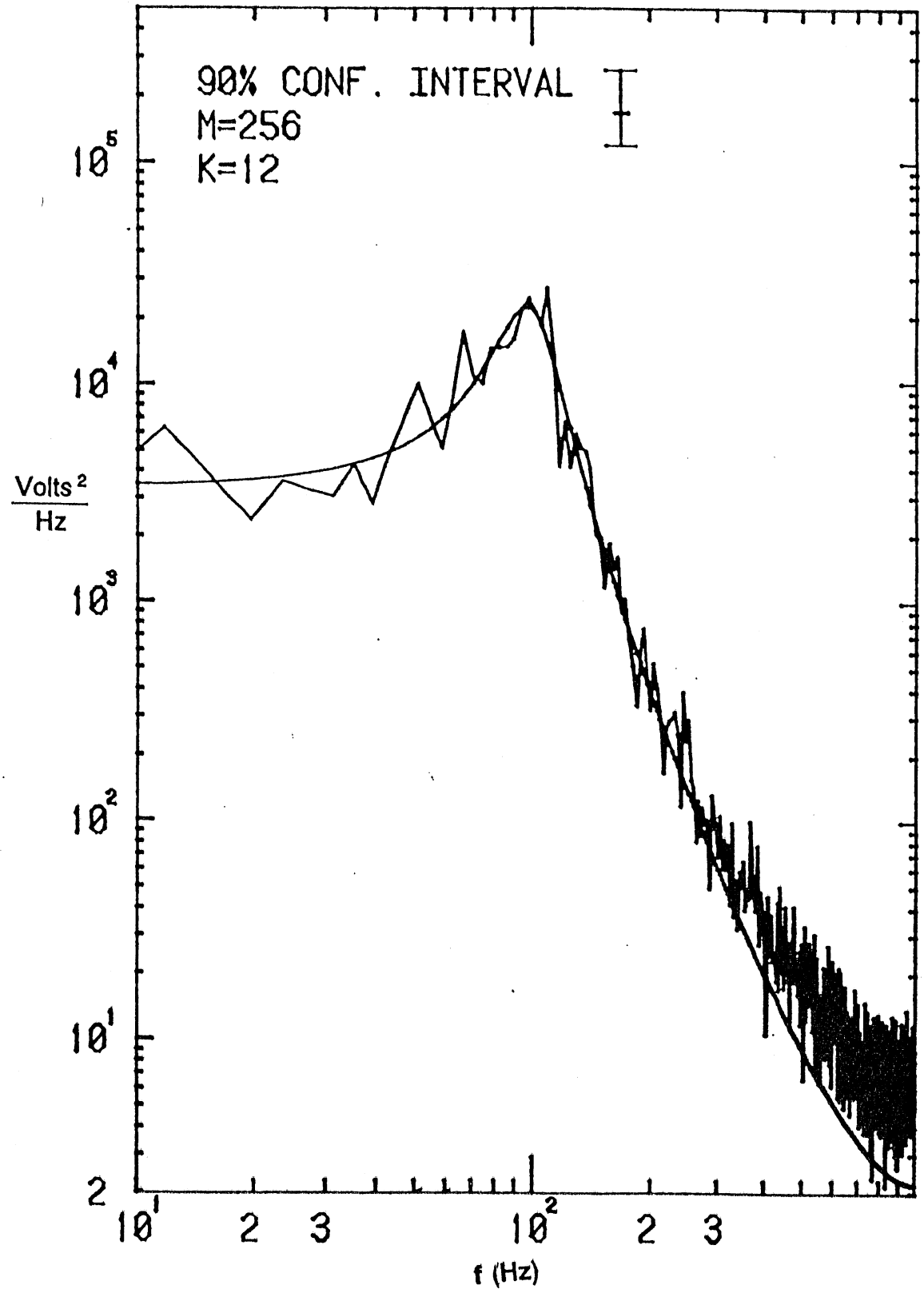


Fig. 5.7 Sample spectrum (M = 256, K = 12, N = 3072)

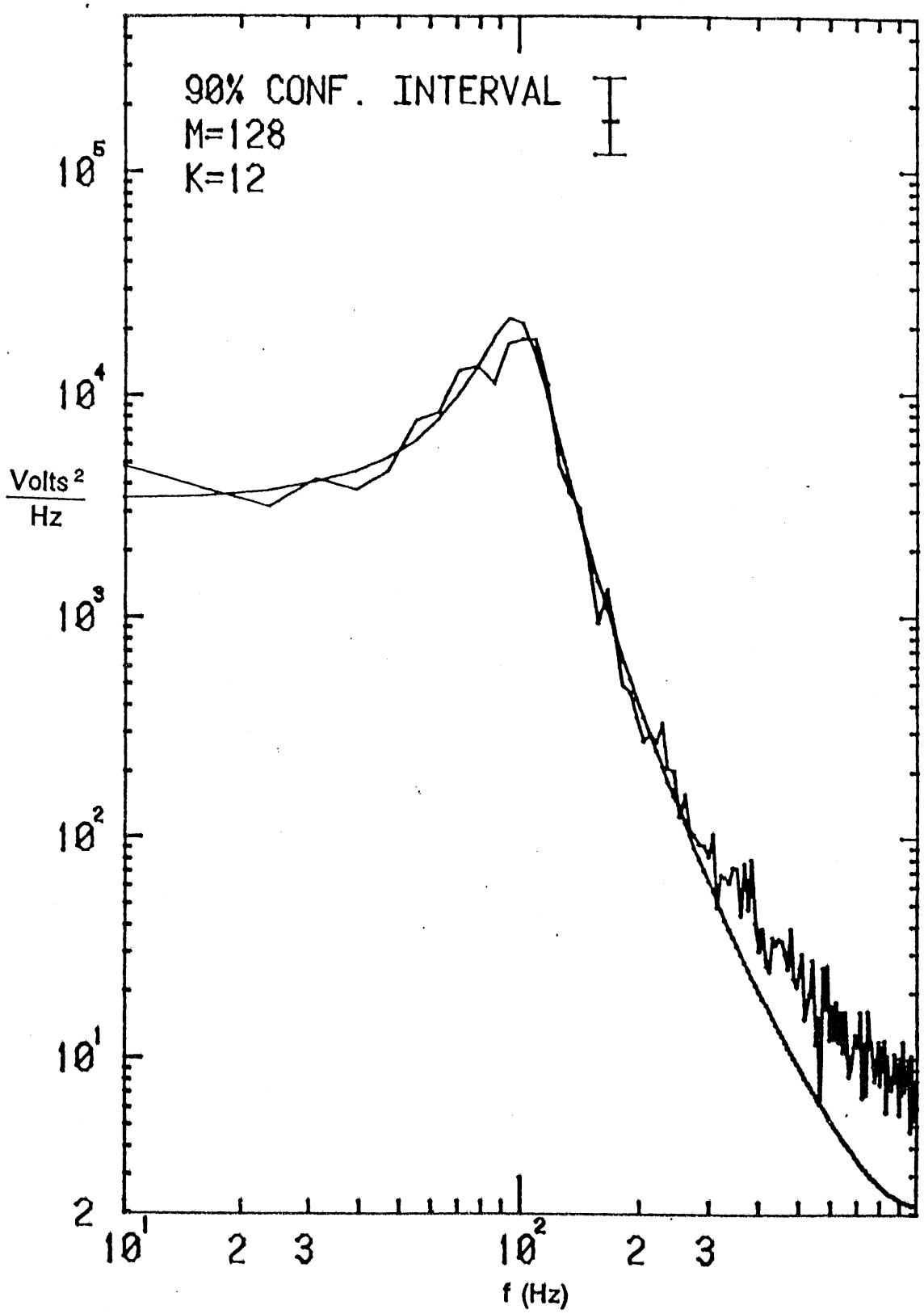


Fig. 5.8 Sample spectrum (M = 128, K = 12, N = 1536)

6. APPLICATION OF THE TECHNIQUES TO VELOCITY FLUCTUATION DATA

6.1. Introduction

In this chapter, a detailed discussion of velocity fluctuation data gathered in the SAFHL wind tunnel is presented to illustrate various aspects of spectral analysis and the application of the calculation programs introduced in the previous chapters. In the study of the structure of turbulence one must know how neighboring velocity fluctuations are related (both in time and space) and how the turbulent kinetic energy is distributed over various wavelengths. The autocorrelation and spectral density functions provide relevant fundamental information which can be used to characterize the various turbulence length scales and the exchange of energy between eddies of different wavelengths.

The measurements were performed in a neutral atmospheric boundary layer simulated for urban terrain conditions in the St. Anthony Falls Laboratory boundary layer wind tunnel. Brief descriptions of the simulation technique and the boundary layer tunnel have been given in Chapter 1. Reference is made to Farell and Iyengar (1993) for details of the simulation and the measurement techniques. The velocity signals were sampled and recorded using a LeCroy data acquisition system with a 12-bit digitizer. Direct on-line voltage spectra were also obtained using a Hewlett-Packard dynamic signal analyzer.

The spectra and autocorrelations were obtained by means of a FORTRAN program (referred to as the spectrum program from now on), developed on the basis of the concepts and equations presented in earlier chapters. A description of this program is given in the next section. Section 6.3 illustrates various aspects of time series analysis. Section 6.4 deals with the methods of computation of the integral length scale of the turbulence. The differences between direct voltage spectra and velocity spectra are examined in Appendix A. The DADiSP software has been extensively used as a graphics spreadsheet program in the various comparisons and computations.

6.2. Description of the Spectrum program

The spectrum program is written in FORTRAN language and can be run

on any PC with more than 540KB of available RAM memory. The program is based on the material presented in the earlier chapters and computes the power spectrum and the autocovariance of a regularly sampled time series. The power spectrum is computed from a fast Fourier transform (FFT) of the time data according to Eq. (72), Ch. 4. The program divides the overall time record into subseries (the number of subseries can be adjusted), with the overall mean removed, and the final spectrum is obtained as an average of the spectra of the subseries. The autocovariance is then obtained as the inverse FFT of the averaged power spectrum.

The program is written in an interactive mode. It reads from an user specified input data file and writes the spectra and the autocovariance into a specified output file[1]. The total number of data points, N, the sampling rate, Δt, and the number of points, M, in each subseries[2] should be specified by the user when prompted. A radix-2 algorithm is used in the computation of the FFT which requires the number of points in each subseries, M, to be an integral power of 2 (referred to as MLOG in the program). The value of MLOG should also be specified by the user.

The program has some restrictions when used on PC's due to the 640KB barrier of DOS. One can analyze a maximum total time record length of 32768 (32KB) data points and a maximum subseries record of 2048 points. The program outputs into a data file in a fashion easily readable by spreadsheet programs like Lotus 1-2-3[3] or DADiSP. The program requires for full display of the spectral estimates a software like DADiSP with a capability to handle arrays of virtually any size. If a time record of more than 32KB is to be analyzed, then the data could be divided into blocks of 32768 points and the program can then be used for each block.

Subroutine CIRDFT is used to calculate the spectrum and autocovariance according to section 4.4, Ch. 4. The spectrum of each subseries is obtained as the FFT of the subseries time data with M zeros

---

[1]Both filenames, including the extensions, should be at most 9 characters in length.

[2]The overall time record length, $T_r = N\Delta t$, is divided into K subseries of equal length T; $T = M\Delta t$, $N = KM$.

[3]Note: If Lotus 1-2-3 is used the filename should have an extension(for e.g. *.asc)

attached as shown in Eq. (70) (this is done so as to obtain the autocovariance directly from the inverse FFT of the spectrum). The frequency resolution of the resulting spectrum is half the frequency resolution $\Delta f = 1/(M\Delta t)$. It can be shown that, at even frequency points, the spectrum of the time record with zeros attached [as given by Eq. (72), Ch. 4] is exactly half the spectral estimate for the original time record. Spectral estimates at odd frequency points represent an interpolation between neighboring points and are not shown as output. The spectrum of each subseries is normalized using the variance, obtained by removing the overall mean, of that subseries. Hence, the final output is the averaged normalized spectrum. The program also gives the value of the integral length scale calculated directly from the value of the spectrum at zero frequency (see section 6.4). The subroutine contains a block that converts the voltage data gathered from the anemometer into velocities (see Appendix A). To this end, the program prompts the user to specify the fluid temperature, TE, during data acquisition.

Subroutine FFT implements the FFT algorithm and is used to compute the finite Fourier transform of each subseries. Finally, subroutine CONORM gives the confidence intervals for the averaged spectrum assuming that it follows a χ_ν^2 distribution (see section 4.5, Ch. 4). The chi-square distribution is approximated by a normal distribution since the number of degrees of freedom, ν, is usually large (> 30). A print out of the computer program used is given as Appendix B for further reference.

6.3. Time Series Analysis

6.3.1. Spectral density function. As will be discussed more fully later, an important task in the application of spectral analysis to turbulent flows, in particular in atmospheric boundary layer simulations, is the evaluation of the integral length scale from the value of the spectrum at zero frequency. To reduce the error in this evaluation, spectral estimates in the neighborhood of zero frequency are needed in addition to the zero frequency value. The number of low frequency spectral estimates available to this end depends on the frequency resolution of the data acquisition scheme. For the purpose of the present analysis, the data was sampled at intervals Δt = 2ms, 0.2ms, and 0.05ms, with filter cut off frequencies of 200hz, 1Khz, and 5Khz respectively. For the data sampled at Δt = 2ms, with the filter cutoff frequency at 200hz,

the maximum achievable frequency resolution is $\Delta f = 1/T = 0.244$hz, corresponding to the longest subseries record that can be analyzed with the spectrum program (M = 2048 points, T = 4.096s). Initially, the total time record length was fixed at N = 32768 points (T_r = 65.54s) and three subseries lengths (M = 512, T = 1.024s; M = 1024, T = 2.048s; M = 2048, T = 4.096s)[4] were tried. Figures 6.1 through 6.3 show log-log plots and linear-linear plots of the power spectra obtained for the three subseries lengths. On the log-log plots the points above 200hz have been truncated since these points are beyond the filter cutoff frequency, and only the points up to 50hz have been included on the linear plots for the sake of clarity. Also shown on the log-log plots are the 95% confidence intervals for the sample estimates $S(f)^{5,6} = 2\bar{C}_{xx}(f)$, based on the assumption that the quantity $\nu\bar{C}_{xx}(f)/\Gamma_{xx}(f)$ approximately follows a χ_ν^2 distribution (see Ch. 4). The 95% confidence intervals on the linear-linear plots are for the zero value of the spectrum and are based on the best estimate of $\Gamma_{xx}(0)$ (obtained with M = 2048, K = 48 as explained later).

These figures show clearly that the variance of the spectrum decreases as the number of subseries and hence the number of averages increase [Eq. (89), Ch. 4]. On the other hand, as discussed in Ch. 4, reducing the subseries size, M, results in an increase in bias. This increase in bias is in first approximation proportional to $d^2\Gamma_{xx}/df^2$ and, for the Bartlett window, inversely proportional to $T^2 = (M\Delta t)^2$, and is small if the spectrum line has small curvature. Spectral peaks where the curvature is not small, however, are underestimated when the subseries size is small (Jenkins and Watts, 1968, Ch. 6). By comparing the sample spectra shown in Figs. 6.1, 6.2, 6.3, and 6.4 [Figs. 6.4(a), 6.4(b) depict the "best" spectral estimate described in the

---

[4]If one assumes that the integral length scale is of the order of the distance from the floor (i.e., z = 20cm), the integral time scale obtained using the mean velocity at this point ($\bar{U}$ = 19.8m/s) is about 10ms which is much smaller (about 100 times) than the minimum subseries length (T = 1.024s) used in this analysis. The consecutive subseries can be assumed then to be uncorrelated (p. 13, Ch. 3).

[5]We use S(f) to denote twice the sample spectrum $\bar{C}_{xx}(f)$ defined in Chapter 4, so that $\int_{-\infty}^{\infty}\bar{C}_{xx}(f)df = \int_{0}^{\infty}S(f)df = \sigma^2$.

[6]Figures 6.1 through 6.13 and Fig. 6.18 show PSD $\equiv S(f)/\sigma^2$.

next paragraph] we see that for the three values of M selected there is no appreciable increase in bias with decreasing M over most of the frequency range. Although only faintly suggested by the data as presented, the low frequency peak is somewhat underestimated by the smaller subseries sizes (M = 1024 and M = 512).

The results also indicate that the spectral estimate obtained for M = 2048 (K = 16) was not averaged over a sufficient number of subseries to produce a reasonably small confidence interval for the spectral estimates [see in particular Fig. 6.3(b)]. In order to increase the number of averages without changing the frequency resolution, the total time record length was extended to N = 98304 points (T_r = 196.60s) and the spectral analysis was carried out with the same subseries length (M = 2048, T = 4.096s). As mentioned in Section 6.2, the data had to be split into 3 blocks and the spectrum program executed for each block. The mean of the overall time record (N = 98304), computed beforehand, was removed from each subseries. The resulting spectrum is an average over the three data blocks (i.e., K = 48). This spectrum, referred to herein as the "best" estimate, is depicted in Figs. 6.4(a) and 6.4(b) and shows clearly the reduction in variance resulting from averaging over the larger number of subseries.

A block-averaged spectral estimate was also obtained from the same three blocks of data using K = 96 subseries with M = 1024 points each. This estimate, shown in Figs. 6.5(a) and 6.5(b), is very close to the block averaged estimate with M = 2048 and K = 48, but it has of course half the variance and double the frequency resolution. All together, these results indicate that the spectral estimate obtained using 48 subseries records with M = 2048 points can be used as "best" estimate in the sense of satisfying the criteria of adequate frequency resolution at low frequencies and reasonably small variance. The same conclusion can be drawn from a similar analysis of the data at z = 60cm, shown in Figs. 6.6 through 6.10.

The spectral estimate obtained at z = 0.20m from the data gathered with the filter setting at 5Khz (Δt = 0.05ms), using K = 16 subseries of M = 2048 points (T = 0.1024s), is shown in Fig. 6.11(a). For smaller subseries lengths the record length is found to be less than twice the integral time scale (using the integral time scale value of 0.033s obtained later on) and hence consecutive subseries may not be statistically independent (see p.13, Ch. 3). For comparison purposes, spectral estimates were obtained anyway with

smaller subseries lengths M = 1024 (T = 0.0512s) and M = 512 (T = 0.0256s), and are shown respectively in Figs. 6.11(b) and 6.11(c). Linear plots of the spectral estimates of Figs 6.11(a), 6.11(b), and 6.11(c) (sampling interval Δt = 0.05ms) are shown for frequencies below 100hz in Fig. 6.11(d), together with the "best" spectral estimate (obtained as explained above from data sampled at Δt = 2ms intervals, using K = 48 subseries of M = 2048 points each, T = 4.096s). It is clear from the figure that, in addition to having a very poor frequency resolution, the spectral estimates obtained with the smaller subseries of 1024 and 512 points exhibit large errors at zero frequency, as expected.

The estimate obtained using 2048 point subseries also exhibits considerable error. In this case, the subseries length T = 0.1024s is only about 3 times the integral time scale of the velocity fluctuations. For this value of T, the frequency resolution is only about 10hz and the bias in the zero frequency estimate, inversely proportional to T^2 in first approximation, may be considerable. Due to the large variance of the spectral estimator at zero frequency, however, additional estimates would be needed for a precise determination of the bias.

To verify whether the short subseries length for M = 2048 and Δt = 0.05ms had any effect on the spectral estimates at intermediate frequencies within the inertial subrange, say, between 100hz and 500 hz [see Fig. 6.11(a)] sample spectra were obtained from N = 32768 data points sampled at Δt = 0.2ms intervals, with the filter cutoff frequency set at f_{cutoff} = 1Khz. These spectra are shown in Figs. 6.12(a) through 6.12(c) for M = 512, 1024, and 2048, respectively. (Note that for M = 512 points and Δt = 0.2ms, T = 0.1024s, the same length as for M = 2048 and Δt = 0.05ms.) The agreement with the sample spectrum of Fig. 6.11(a) at frequencies above, say, 50hz, is very good. Figure 6.12(d) shows a linear scale comparison, for frequencies below 100hz, of the spectra of Figs. 6.12(a) through 6.12(c) with the "best" estimate defined earlier (Δt = 2ms, K = 48, M = 2048, T = 4.096s). The low frequency values of the spectrum are underestimated by the sample spectra obtained with the Δt = 0.2ms sampling rate. The results in this and the preceding paragraph are verified by a similar analysis of the data at z = 60cm.

Final combined sample spectra over the range from 0 - 5khz, obtained by piecing together the spectral estimates calculated using the three

filter settings, are shown in Figs. 6.13(a) and 6.13(b) for z = 20cm and 60cm, respectively. The "best" estimate (f_{cutoff} =200hz, Δt = 2ms, N = 98304, M = 2048, K = 48) was used for the spectral values up to 200hz. To extend the spectrum to 1Khz, the sample spectrum for Δt = 0.2ms (f_{cutoff} = 1Khz), N = 32768, and M = 2048 (K = 16), was used. Beyond 1Khz the spectral estimates obtained for Δt = 0.05ms (f_{cutoff} = 5khz), N = 32768, and M = 2048 (K = 16) were employed. The need for division of the spectral calculation into frequency ranges depends on the reliability and frequency resolution required. For this particular application to wind velocity data, it is necessary to examine the low frequency range separately, using a sufficiently large sampling interval. The final combined sample spectra could have been obtained also by piecing together the spectral estimates for just the smallest and largest filter settings (f_{cutoff} = 200hz and 5khz). The preceding discussion depicts some of the difficulties involved in spectral analysis and emphasizes at the same time the need for careful experimental planning in the collection of data.

6.3.2. Autocovariance. Normalized sample autocovariances at z = 5cm, evaluated as the Fourier transforms of the sample spectra for N = 32768 and M = 512, M = 1024, and M = 2048, are shown respectively in Figs. 6.14(a) through 6.14(c). The results are nearly identical, within the limits of the truncation errors, for all the three cases, except at the higher delays approaching the subseries length T. Indeed, as shown in Ch. 3, the variance of the autocovariance function estimator is approximately independent of the subdivision of the time record into subseries, as long as the subseries length T is much greater than the integral time scale. Of course, the autocovariance function for longer subseries carry information about higher lags.

The autocovariance obtained as the inverse FFT of the sample spectrum should be identical to the autocovariance computed directly from its definition as an average of lagged products. The direct computation was carried out using the DADiSP software for the same N = 32768 data points of Figs. 6.14(a) through 6.14(c) and the result is shown in Fig. 6.15(a) up to a lag time of 2 seconds for comparison with Figs. 6.14. Because as shown by Eqs. (79) and (80), Ch. 3, the sample autocovariance obtained as the inverse FFT of the averaged sample spectrum is actually the autocovariance multiplied by a Bartlett lag window, in Fig. 6.15(b) the sample acvf estimates of Fig.

6.14(c) are shown divided by the Bartlett window. The results of Figs. 6.15(a) and 6.15(b) are practically identical within the truncation errors of the computations.

In Figs. 6.15 and 6.16, the sample autocovariances show the spurious periodicities at higher lags due to the high correlation of neighboring autocovariance estimators at large delays (Ch. 3). The biased estimator is a consistent estimator of the autocovariance at each lag in the sense that the expected value tends to the true value as $T \to \infty$. However, it is not a uniformly consistent estimator of the entire autocovariance function because its mean square error tends to a constant value different from zero as $T \to \infty$ [Eq. (37), Ch. 3]. The spurious periodicities decrease with increase in T as is evident from Fig. 6.16, which shows the autocovariance obtained as the FFT of the three block average referred to earlier as the "best" spectral estimate ($N = 98304$, $\Delta t = 2ms$, $M = 2048$, $K = 48$).

Finally, to stress the magnitude of possible errors due to an incorrect choice of data acquisition parameters, Fig. 6.17 shows a comparison of the sample autocovariance function obtained sampling at $\Delta t = 0.05ms$ ($f_{cutoff} = 5khz$), using $N = 32768$ and $M = 512$ ($T = M\Delta t = 0.0256s$), with the estimate obtained with $\Delta t = 2ms$ ($f_{cutoff} = 200hz$), $N = 98304$, and $M = 2048$ ($T = 4.096s$). The figure clearly shows the large error involved. It should be noted, however, that the error is exaggerated in this particular example because the subseries length $T = 0.0256s$ is less than the integral time scale of the process ($\approx 0.033s$).

6.4. Computation of Integral Length Scales

The calculation of integral length scales will be carried out by means of the following methods, which will be compared:

I. Using the zero value of the spectrum

II. Computing the area under the autocorrelation curve

III. Using the location of the spectral peak

6.4.1. Calculation of the integral scale using the zero value of the spectrum.

The integral length scale can be obtained using the value of the power spectrum at zero-frequency as

$$L_u{}^x = \frac{\bar{U}}{2} \frac{\Gamma_{xx}(0)}{\sigma^2} \tag{95}$$

where $\bar{U}$ = mean velocity[7]. In order to reduce the error in the determination of the zero frequency spectral value, the sample spectra obtained with the data sampled at Δt = 2ms intervals (f_{cutoff} = 200Hz), which provided the best frequency resolution at low frequencies, were used. The integral scale estimates obtained for N = 32768 and different subseries lengths using Eq. (95) above are listed as $(L_u{}^x)_1$ in Tables 6.1(a) and 6.1(b) for, respectively, z = 20cm and z = 60cm. One can see from the tables and more clearly from Figs. 6.1(b), 6.2(b), and 6.3(b), that the zero-frequency estimates for this number of data points possess large random errors.

Tables 6.2(a) and 6.2(b) (for z = 20cm and z = 60cm, respectively) show the integral length scale estimates evaluated from Eq. (95) using the block-averaged spectral estimates of Figs. 6.4(b) and 6.9(b) (Δt= 2ms, N = 98304, M = 2048, K = 48). The tables give also the length scales obtained using Eq. (95) for each of the three blocks of data used in the overall average. As in other tables, $(L_u{}^x)_1$ has been used as symbol for these estimates. Again, the zero-frequency estimates for N = 32768 differ considerably from block to block, in agreement with the variability of the low-frequency spectral estimates in Fig. 6.3(b). Tables 6.3(a) and 6.3(b) show the change in the $(L_u{}^x)_1$ values of Tables 6.2(a) and 6.2(b) when the same 98304 points are analyzed using M = 1024 and K = 32 for each data block. As Fig. 6.2(b) shows, the resulting confidence interval at zero frequency is somewhat smaller, but the length scale estimates still show a very large variability.

To reduce the error in the integral length scale estimates, an extrapolation to zero frequency of the sample spectral values in the neighborhood of zero can be used, taking advantage of the fact that for the Bartlett window the correlation between neighboring spectral estimates is about 0.2, while the correlation after two frequency spacings is negligible (see footnote p. 30, Ch. 4). Because the longitudinal velocity correlation is

---

[7]In terms of the sample spectrum S(f) defined on p. 56, we have

$$L_u{}^x = \frac{\bar{U}}{4} \frac{S(0)}{\sigma^2}$$

real and positive (no negative values have ever been measured), the longitudinal turbulence spectrum is symmetric and has an absolute maximum at zero wave number ($\kappa = 2\pi f/\bar{U}$). For small values of the frequency, therefore, the spectrum $\Gamma_{xx}(f)$ can be represented to third order by a parabola of the form Af^2+C. Such a parabola was fitted by least squares through the first few points of the sample spectra, selecting the number of points on the basis of curve trends and the mean square error of the fit. The values of integral length scales obtained by this extrapolation procedure are shown in Tables 6.1 through 6.3 using the symbol $(L_u^x)_2$. Figure 6.18 depicts for illustration the parabola fitted to the "best" spectral estimate.

Tables 6.4(a) and 6.4(b) present a summary of results organized by increasing number of subseries. Scale estimates obtained from sample spectra block-averaged over 192 subseries (N = 98304, M = 512) are included in addition to the estimates already shown in the preceding tables. The large variability of estimates obtained from spectral values at zero frequency is again evident. The integral scale estimates $(L_u^x)_2$ obtained from the parabola fit are in much better agreement. The last column in these tables shows the percentage difference between the parabola fit estimate for each K and the value of $(L_u^x)_2$ for K = 48, equal to 0.59m for z = 20cm and 0.77m for z = 60cm. These two values are taken herein as best estimates because the sample spectra for K = 48 and M = 2048 (T = MΔt = 4.096s) have better frequency resolution ($\Delta f = 1/T = 0.244$hz) and a reasonably small confidence interval in the neighborhood of zero frequency [see Figs. 6.4(b), 6.9(b)]. The large value of T minimizes also the bias, which for the Bartlett window is inversely proportional in first approximation to $T^2 = (M\Delta t)^2$.

Confidence intervals for the spectra corresponding to K = 64 and K = 96 are shown, respectively, in Figs. 6.1 and 6.5 for z = 20 cm, and 6.6 and 6.10 for z = 60 cm. The confidence intervals for the spectra corresponding to K = 192 (not shown) are smaller. The fact that for these three values of K the parabola fit estimates are within 1.7 and 2.6 percent, respectively, of the "best" estimates at the two heights shows that the reduction in variance associated with the larger K values compensates at least to some extent for the error due to the poorer frequency resolution, and that the bias contribution to the error is small. For a single block of data, N = 32768, the choice M = 512, K = 64 produced the smallest differences with the estimates for K = 48.

The results illustrated in Tables 6.4(a) and 6.4(b) show that it is possible to obtain quick and fairly accurate estimates of integral length scales with just a single block of data (N = 32768), if a parabola fit at low frequencies is used. The largest difference shown in the tables is 7.8 percent for K = 16, M = 2048 at z = 60cm, to be compared with differences of up to 24 percent for the $(L_u^x)_1$ values.

6.4.2. Calculation using the area under the autocovariance curve. The area under the autocovariance curve, normalized using the rms value of the fluctuations, $\sigma = \sqrt{\overline{(u')^2}}$, gives the integral time scale from which, using Taylor's hypothesis, the integral length scale can be calculated as $L_u^x = \bar{U}$(area). While the autocovariance function goes to zero as $T \to \infty$, sample autocovariances present spurious oscillations at large lags (due to the strong correlation of neighboring estimates), which fail to damp out as $T \to \infty$ (see p. 19, Ch. 3). It is then necessary to select a truncation point for calculation of the area, excluding the extraneous oscillations. Given the shape of the sample autocovariances shown in Figs. 6.14 and 6.15, the first zero crossing was taken to be the truncation point for the computation of the area. For the data of Fig. 6.14(c), corresponding to z = 5 cm, the truncation point was selected at $\tau \approx 0.16$s (the 80th point) and the corresponding area gave an integral scale of 38cm.

The selection of truncation point is particularly delicate in the case of correlations which actually exhibit negative values at intermediate lags. The validity of the procedure can be examined by verifying that the truncated oscillations fade away with increasing overall record length. As illustrated by Fig. 6.16, the sample autocovariances obtained using three data blocks and M = 2048 permit a satisfactory selection of truncation point. The integral length scale values obtained from the areas under these sample autocovariances were 0.60m for z = 20cm and 0.79m for z = 60cm.

While the truncation process is obviously arbitrary (Hinze, 1959), it should be noted that the behavior of the sample autocovariance at large lags is directly related to the noted variability of the sample spectral estimates at zero frequency and that, in fact, the use of different subseries lengths (with the same overall time record) in the calculation of the sample spectra is equivalent to truncating the autocovariance function at the lag

corresponding to the subseries length.

6.4.3. Computation of the Integral Length Scale Using the Location of the Spectral Peak. The location of the spectral peak of the non-dimensional spectrum $fS(f)/\sigma^2$ can be used to estimate the integral length scale. For exponentially decaying autocovariances, the relation between the peak and the integral length scale (the area under the normalized autocovariance $S(f)/\sigma^2$) is given by

$$L_u^x = (1/2\pi)(\bar{U}/f_{max})$$

The location of the peak of the non-dimensional spectrum($fS(f)/\sigma^2$) can best be determined from a linear-log plot of ($fS(f)/\sigma^2$) v/s wave number $f/\bar{U}$. The integral scale estimates obtained using this method for the two heights z = 20 cm and z = 60 cm are included in the comparison of results presented in Table 6.5.

A source of possible error in this method is the requirement of an exponential shape for the autocovariance. A second important source of error stems from the difficulty of determining accurately the position of the peak in sample non-dimensional spectra, in particular because the peak occurs normally at relatively small frequencies. However, the method affords quick initial estimates of integral length scales from spectral plots as given, for example, by hard-wired spectral analyzers.

A similar method sometimes used in atmospheric turbulence is to fit the von Karman spectral model to the sample spectrum and obtain L_u^x from the position of the peak (see e.g., Teunissen, 1980).

6.5. Conclusions

To obtain reliable spectral estimates and adequate frequency resolution over the range of frequencies of interest (0-5000 hz for the particular study used herein as example) it is necessary in many cases to gather data sets at different sampling rates and filter cutoff frequencies, each focusing on a specific frequency range. Integral scale estimates, in particular, require sufficiently large sampling rates to ensure that the subseries length is sufficient for adequate resolution at low frequencies. For complex turbulent flows, some trial and error using different record lengths and sampling intervals may be initially necessary to select the

appropriate values of these parameters to avoid aliasing and achieve an adequate trade-off between bias and variance errors.

The estimation of integral scales from the value of the sample spectrum at zero frequency is subject to significant errors. By fitting a parabola to the spectral estimates at low frequencies, however, it is possible to obtain quick and fairly accurate estimates of integral length scales from a reasonably small number of data points (32768 points were used in the examples herein). Depending on the sampling interval, a relatively small subseries length, say, M = 512 points, may produce the most accurate estimates. The curve fitting procedure is tantamount to smoothing out the spectrum near zero and a good frequency resolution is necessary to minimize possible bias errors.

Integral scales obtained from sample autocovariance areas depend on the selection of truncation point, which may be particularly delicate in the case of correlations which exhibit negative values at intermediate lags. The validity of the procedure can be examined by verifying that the truncated oscillations fade away with increasing overall record length. While the truncation process is arbitrary, the behavior of the sample autocovariance at large lags is directly related to the variability of the sample spectral estimates at zero frequency and, in fact, the use of different subseries lengths in the calculation of sample spectra is equivalent to truncating the autocovariance function at the lag corresponding to the subseries length.

BLOCK NO.	N	M	K	INTEGRAL LENGTH SCALES	
				$(L_u^x)_1$ (m)	$(L_u^x)_2$ (m)
I	32768	512	64	0.52	0.60
I	32768	1024	32	0.51	0.60
I	32768	2048	16	0.47	0.60

$\bar{U}$=19.63 m/s, u'=3.54 m/s.
Δt=2ms, $f_{cut\ off}$=200Hz, $\Delta f(M=2048)$=0.244Hz, $\Delta f(M=1024)$=0.488Hz, $\Delta f(M=512)$=0.976Hz

a) z = 20 cm

BLOCK NO.	N	M	K	INTEGRAL LENGTH SCALES	
				$(L_u^x)_1$ (m)	$(L_u^x)_2$ (m)
I	32768	512	64	0.86	0.79
I	32768	1024	32	0.86	0.80
I	32768	2048	16	0.59	0.83

$\bar{U}$=26.85 m/s, u'=2.80 m/s.
Δt=2ms, $f_{cut\ off}$=200Hz, $\Delta f(M=2048)$=0.244Hz, $\Delta f(M=1024)$=0.488Hz, $\Delta f(M=512)$=0.976Hz

b) z = 60 cm

Table 6.1. Variation of integral length scale with subseries size (Data block I)

BLOCK NO.	N	M	K	INTEGRAL LENGTH SCALES	
				$(L_u^x)_1$ (m)	$(L_u^x)_2$ (m)
I	32768	2048	16	0.47	0.60
II	32768	2048	16	0.57	-----
III	32768	2048	16	0.94	-----
Best Estimate	98304	2048	48	0.68	0.59

$\tilde{U}$=19.81 m/s, u'=3.51 m/s.
Δt=2ms, $f_{cut\ off}$=200Hz, Δf=0.244Hz

a) z = 20 cm

BLOCK NO.	N	M	K	INTEGRAL LENGTH SCALES	
				$(L_u^x)_1$ (m)	$(L_u^x)_2$ (m)
I	32768	2048	16	0.59	0.83
II	32768	2048	16	0.76	-----
III	32768	2048	16	0.87	-----
Average Spectrum	98304	2048	48	0.74	0.77

$\tilde{U}$=26.86 m/s, u'=2.76 m/s.
Δt=2ms, $f_{cut\ off}$=200Hz, Δf=0.244Hz

b) z = 60 cm

Table 6.2. Integral length scales from block averaged spectra (3 data blocks, N=98304, M = 2048, K = 48)

BLOCK NO.	N	M	K	INTEGRAL LENGTH SCALES	
				$(L_u^x)_1$ (m)	$(L_u^x)_2$ (m)
I	32768	1024	32	0.51	0.60
II	32768	1024	32	0.63	-----
III	32768	1024	32	0.91	-----
Best Estimate	98304	1024	96	0.69	0.58

$\bar{U}$=19.81 m/s, u'=3.51 m/s.
Δt=2ms, $f_{cut\ off}$=200Hz, Δf=0.488Hz

a) z = 20 cm

BLOCK NO.	N	M	K	INTEGRAL LENGTH SCALES	
				$(L_u^x)_1$ (m)	$(L_u^x)_2$ (m)
I	32768	1024	32	0.86	0.80
II	32768	1024	32	0.65	-----
III	32768	1024	32	0.91	-----
Average Spectrum	98304	1024	96	0.80	0.78

$\bar{U}$=26.86 m/s, u'=2.76 m/s.
Δt=2ms, $f_{cut\ off}$=200Hz, Δf=0.488Hz

b) z = 60 cm

Table 6.3. Integral length scales from block averaged spectra (3 data blocks, N=98304, M = 1024, K = 96)

K	M	N	INTEGRAL LENGTH SCALES		
			$(L_u^x)_1$ (m)	$(L_u^x)_2$ (m)	% (DIFF)
16	2048	32768	0.47	0.60	1.7
32	1024	32768	0.51	0.60	1.7
48	2048	98304	0.68	0.59	--
64	512	32768	0.52	0.60	1.7
96	1024	98304	0.69	0.58	1.7
192	512	98304	0.61	0.60	1.7

a) z = 20 cm

K	M	N	INTEGRAL LENGTH SCALES		
			$(L_u^x)_1$ (m)	$(L_u^x)_2$ (m)	% (DIFF)
16	2048	32768	0.59	0.83	7.8
32	1024	32768	0.86	0.80	3.9
48	2048	98304	0.74	0.77	---
64	512	32768	0.86	0.79	2.6
96	1024	98304	0.80	0.78	1.3
192	512	98304	0.77	0.75	2.6

b) z = 60 cm

Table 6.4. Comparison of integral length scales obtained using different subseries sizes and total record lengths with that from best estimate (N = 98304, M = 2048, K = 48)

z (cm)	INTEGRAL LENGTH SCALE (m)			
	FROM THE ZERO VALUE OF THE SPECTRUM	FROM PEAK LOCATION	FROM THE EXTRAPOLATED ZERO VALUE OF THE SPECTRUM	FROM AREA UNDER ACVF CURVE
20	0.62	0.80	0.59	0.60
60	0.74	0.78	0.77	0.79

Table 6.5. Integral length scales using different methods

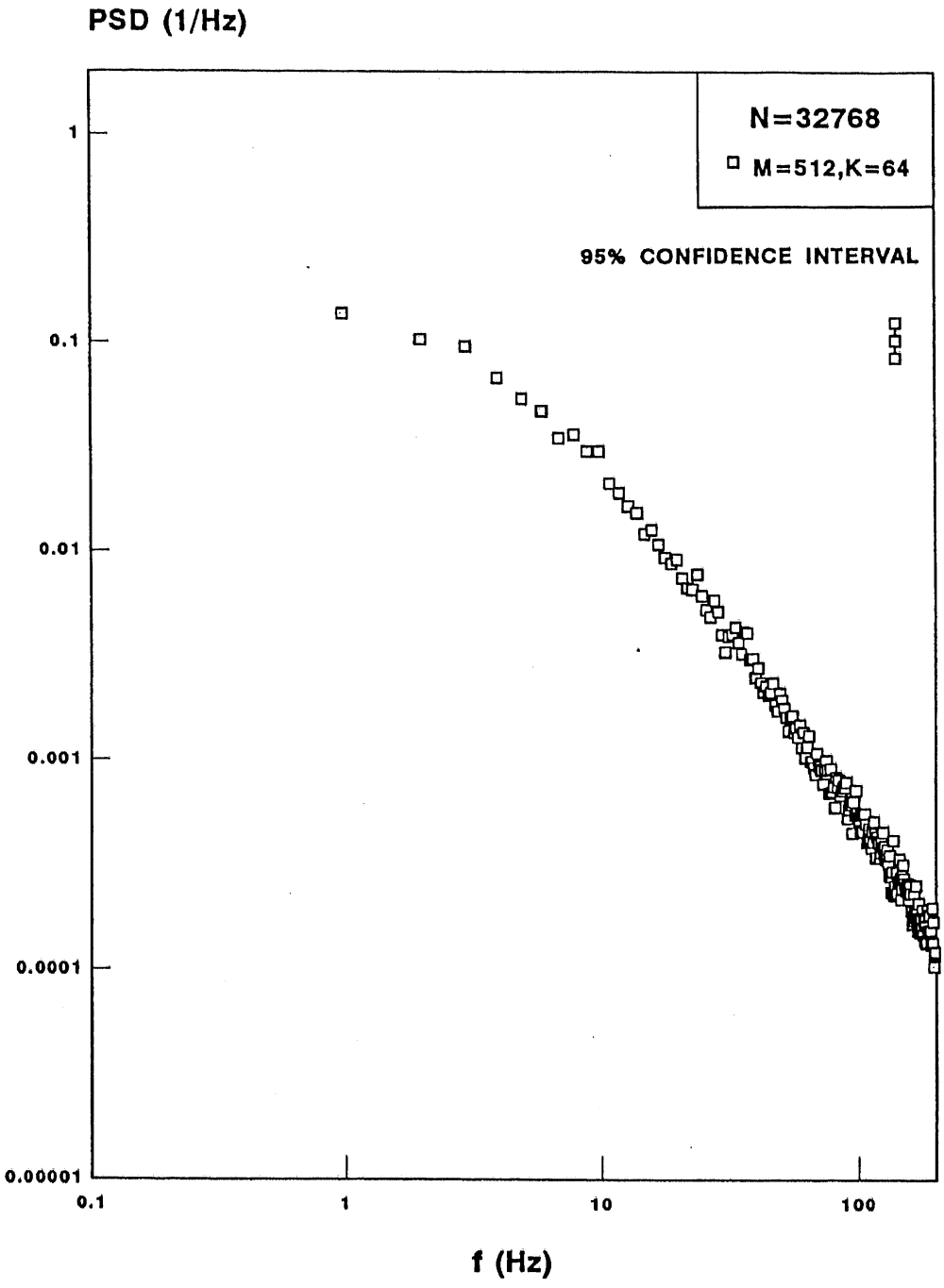


Fig. 6.1(a) Power spectral density at z = 20 cm

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz
N = 32768, M = 512, K = 64

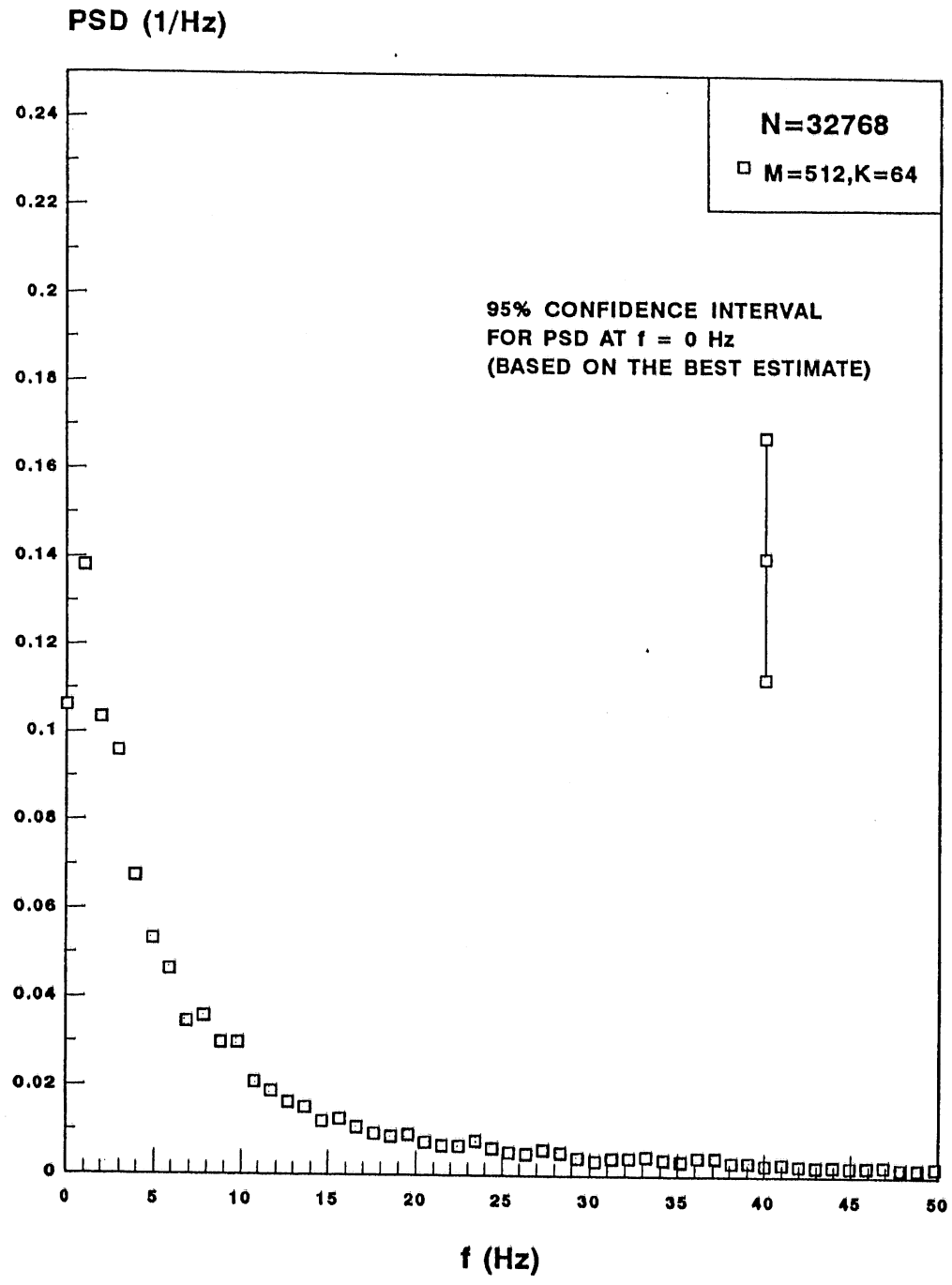


Fig. 6.1(b) Power spectral density at z = 20 cm

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz
N = 32768, M = 512, K = 64

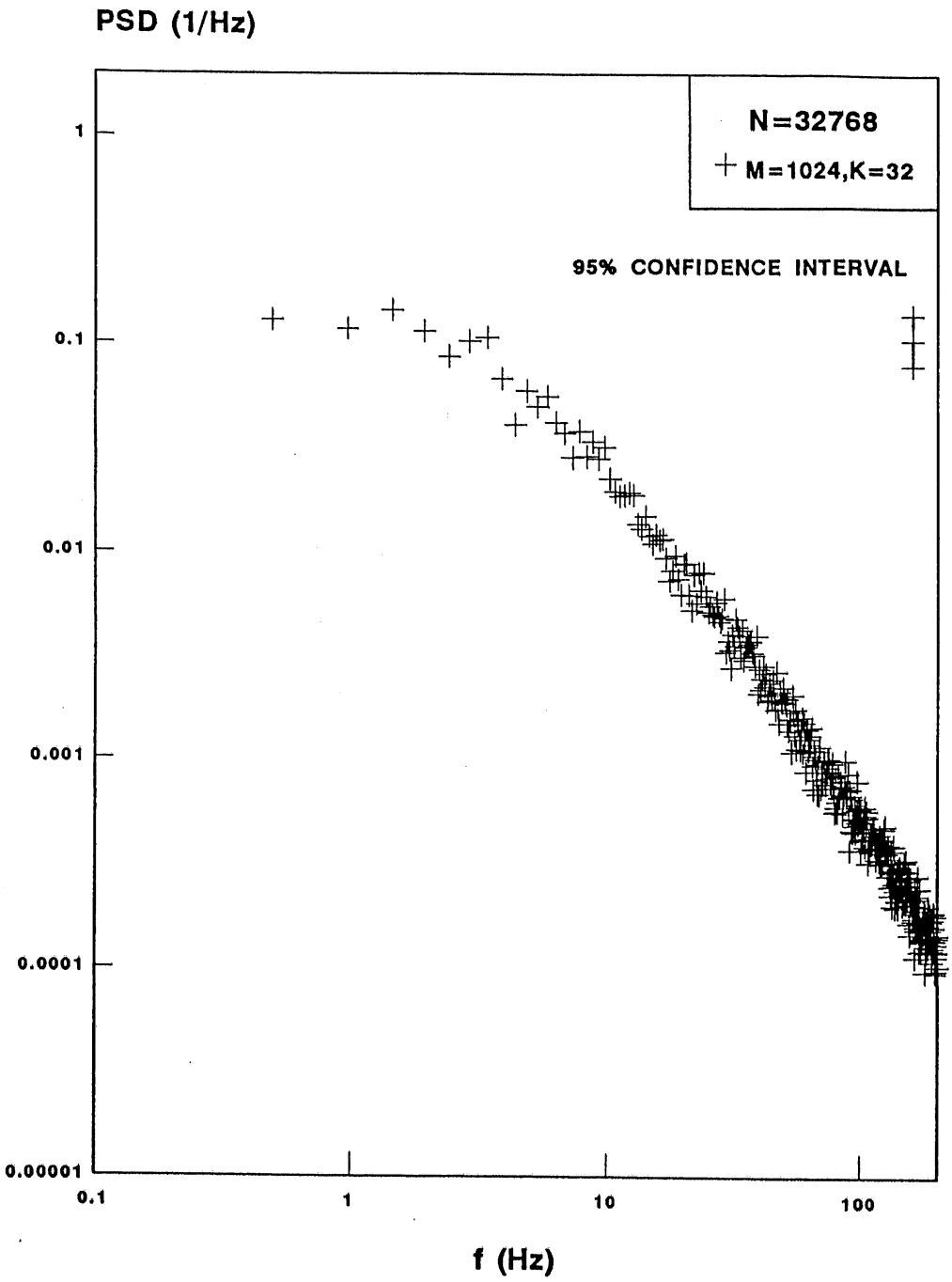


Fig. 6.2(a) Power spectral density at z = 20 cm

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz
N = 32768, M = 1024, K = 32

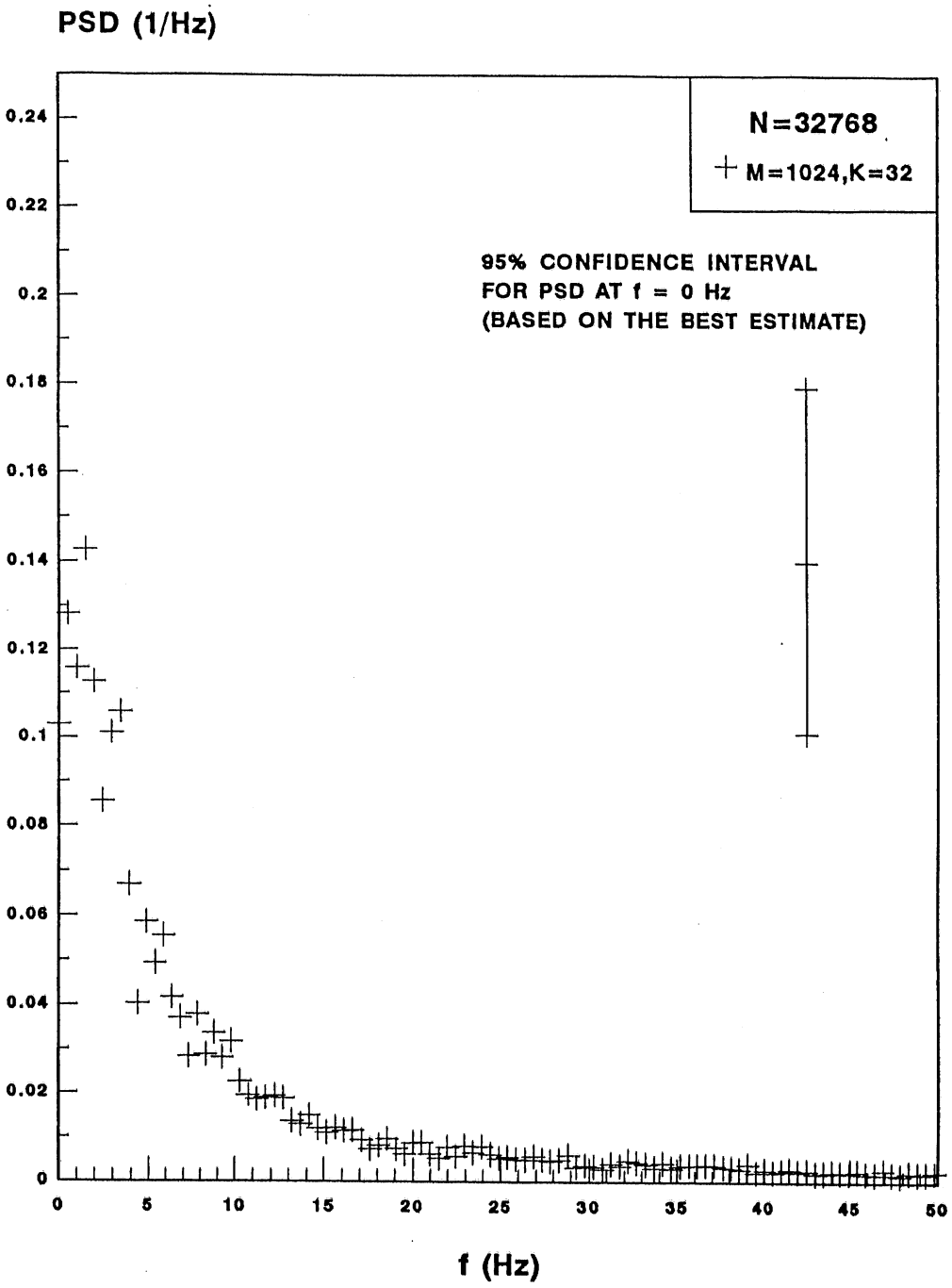


Fig. 6.2(b) Power spectral density at z = 20 cm

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz
N = 32768, M = 1024, K = 32

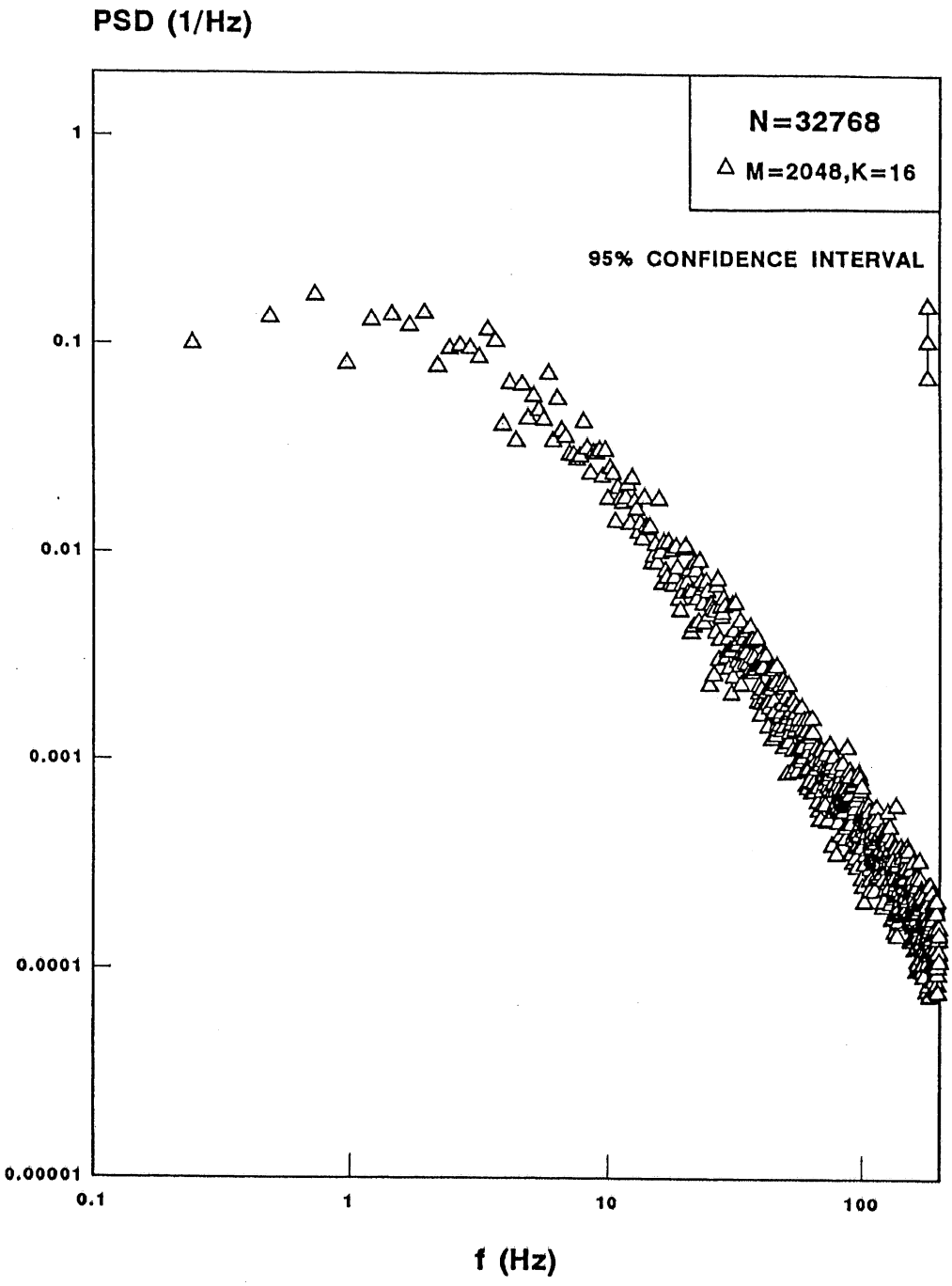


Fig. 6.3(a) Power spectral density at z = 20 cm

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz
N = 32768, M = 2048, K = 16

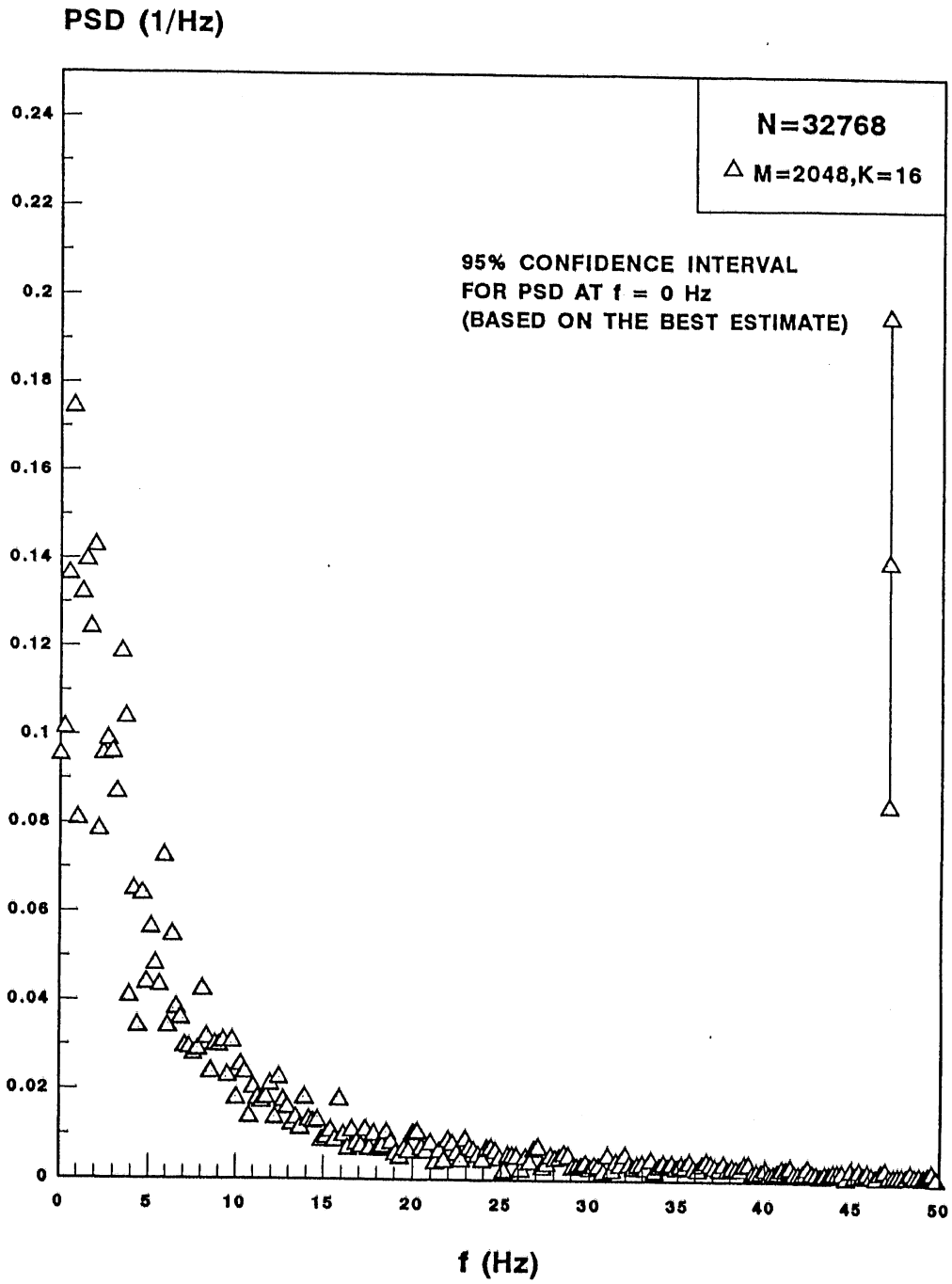


Fig. 6.3(b) Power spectral density at z = 20 cm

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz
N = 32768, M = 2048, K = 16

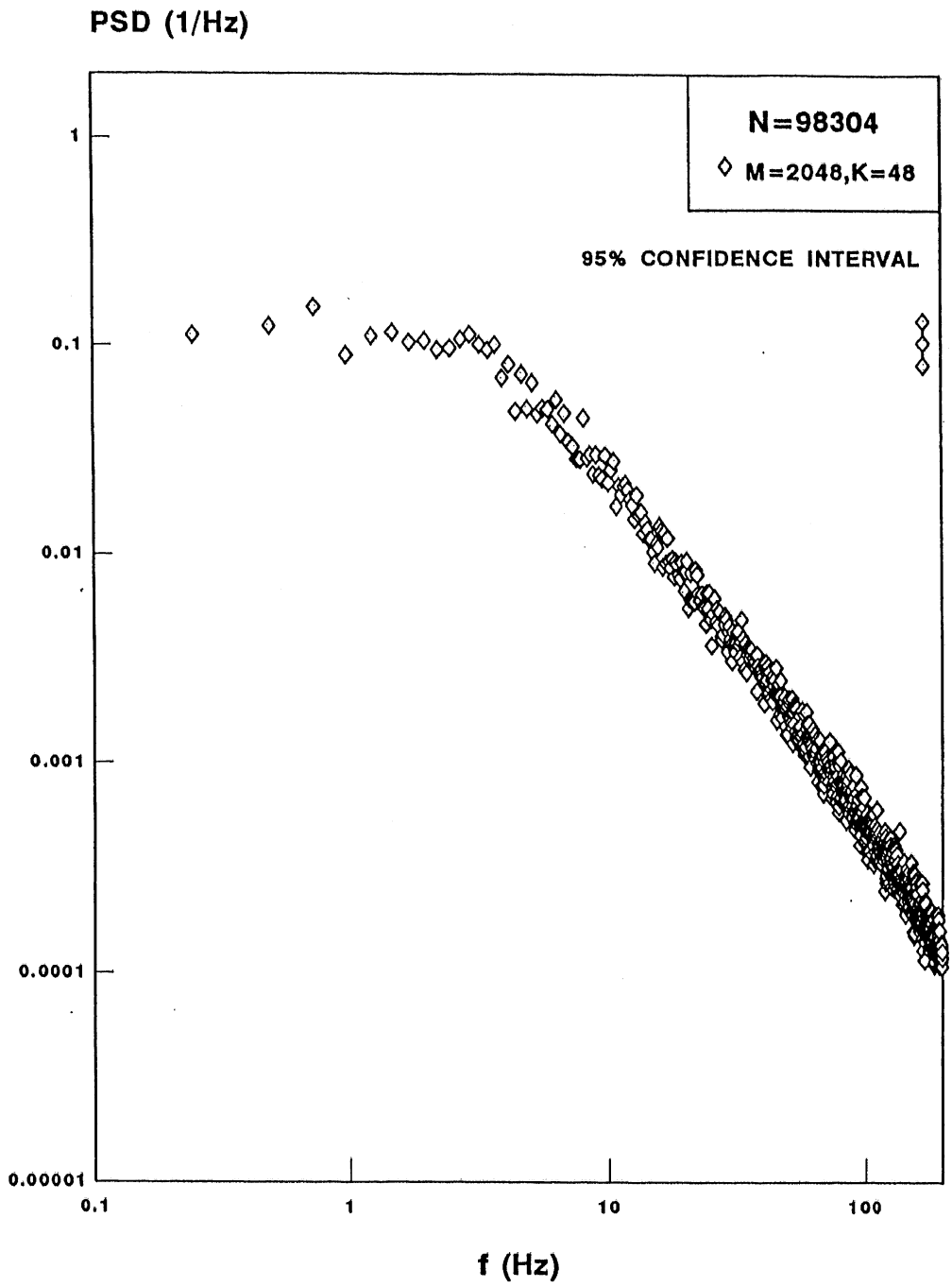


Fig. 6.4(a) Block averaged power spectral density at z = 20 cm

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz
3 data blocks, N = 98304, M = 2048, K = 48

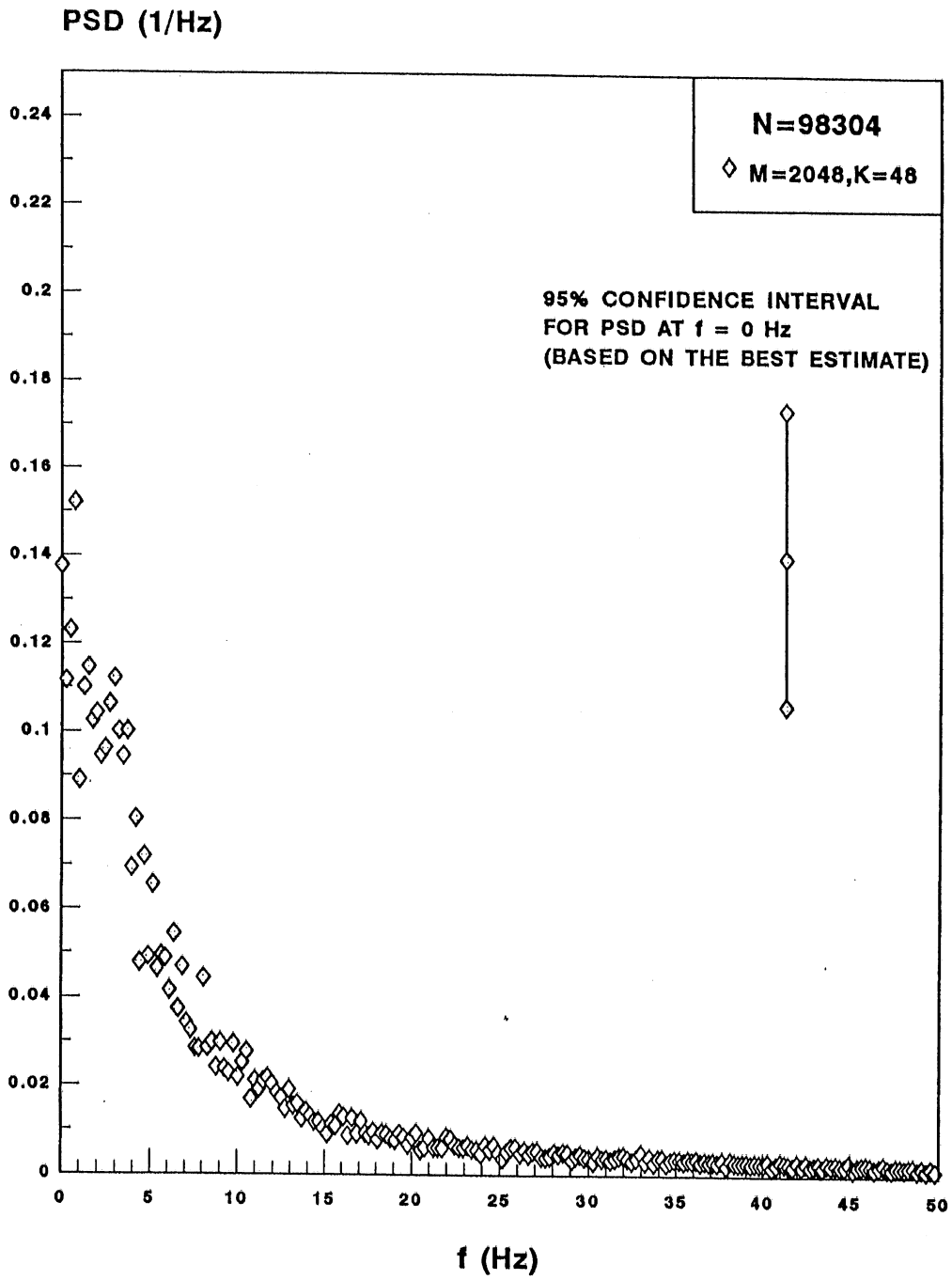


Fig. 6.4(b) Block averaged power spectral density at z = 20 cm

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz

3 data blocks, N = 98304, M = 2048, K = 48

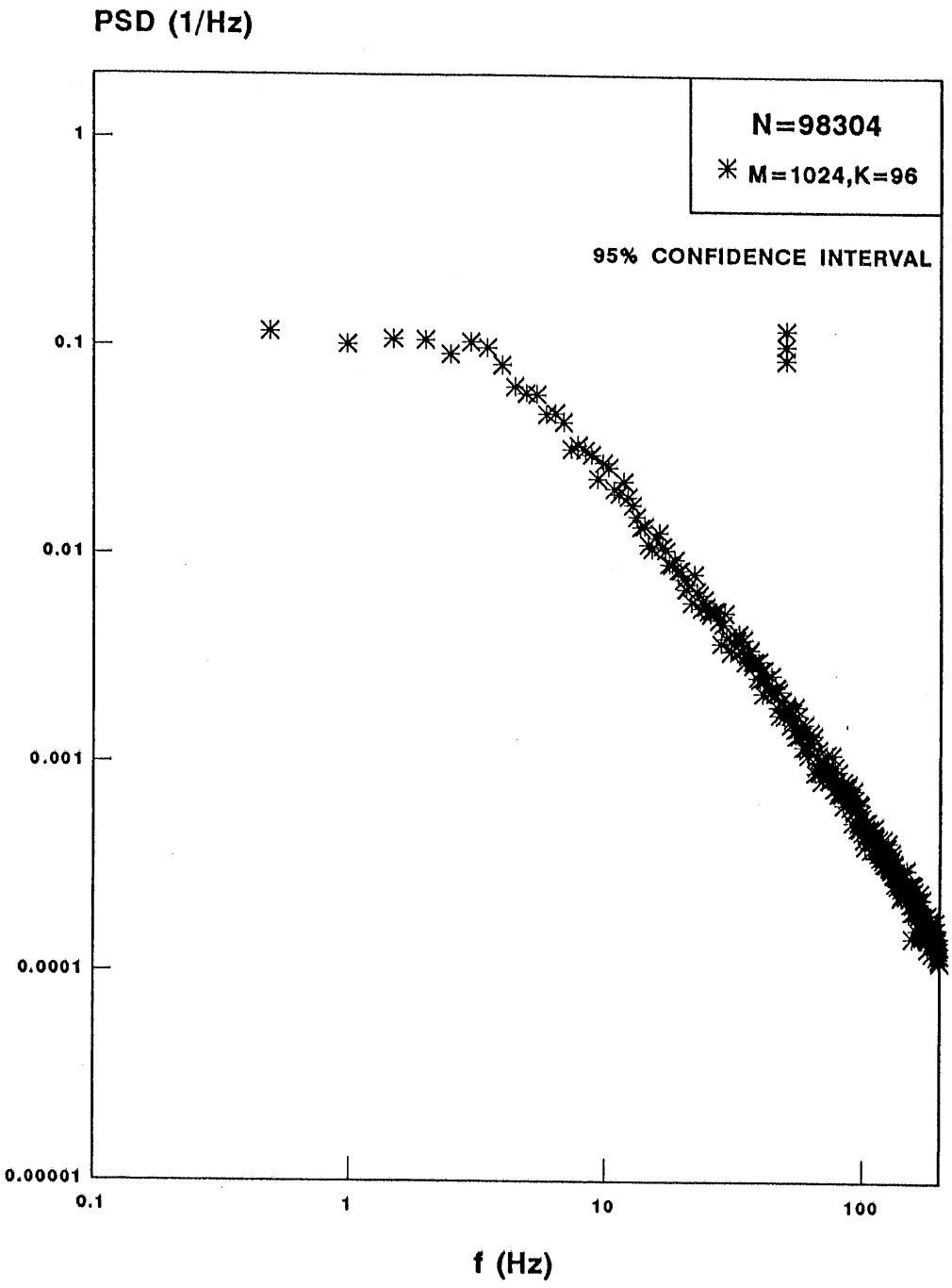


Fig. 6.5(a) Block averaged power spectral density at z = 20 cm

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz
3 data blocks, N = 98304, M = 1024, K = 96

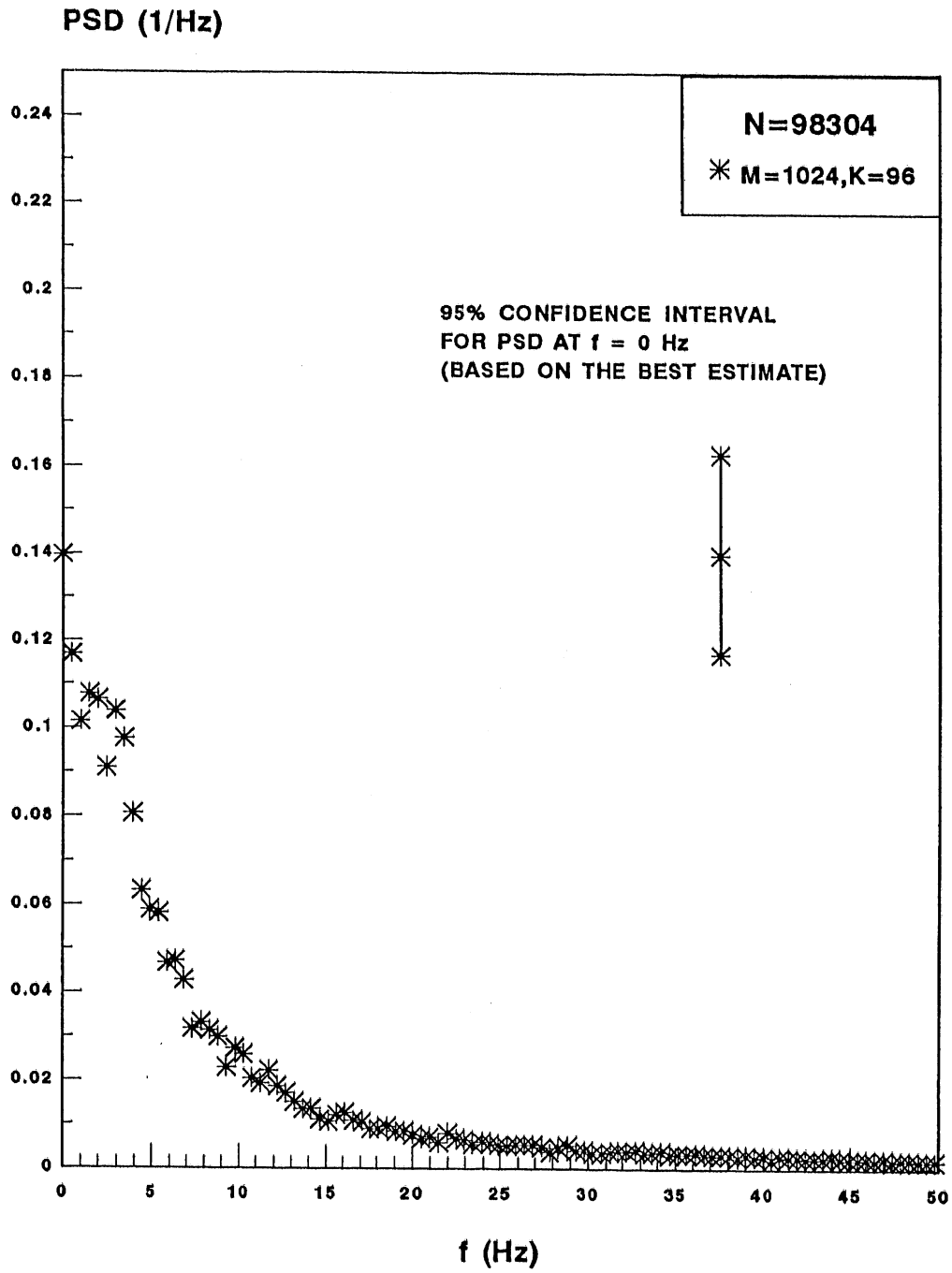


Fig. 6.5(b) Block averaged power spectral density at z = 20 cm

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz

3 data blocks, N = 98304, M = 1024, K = 96

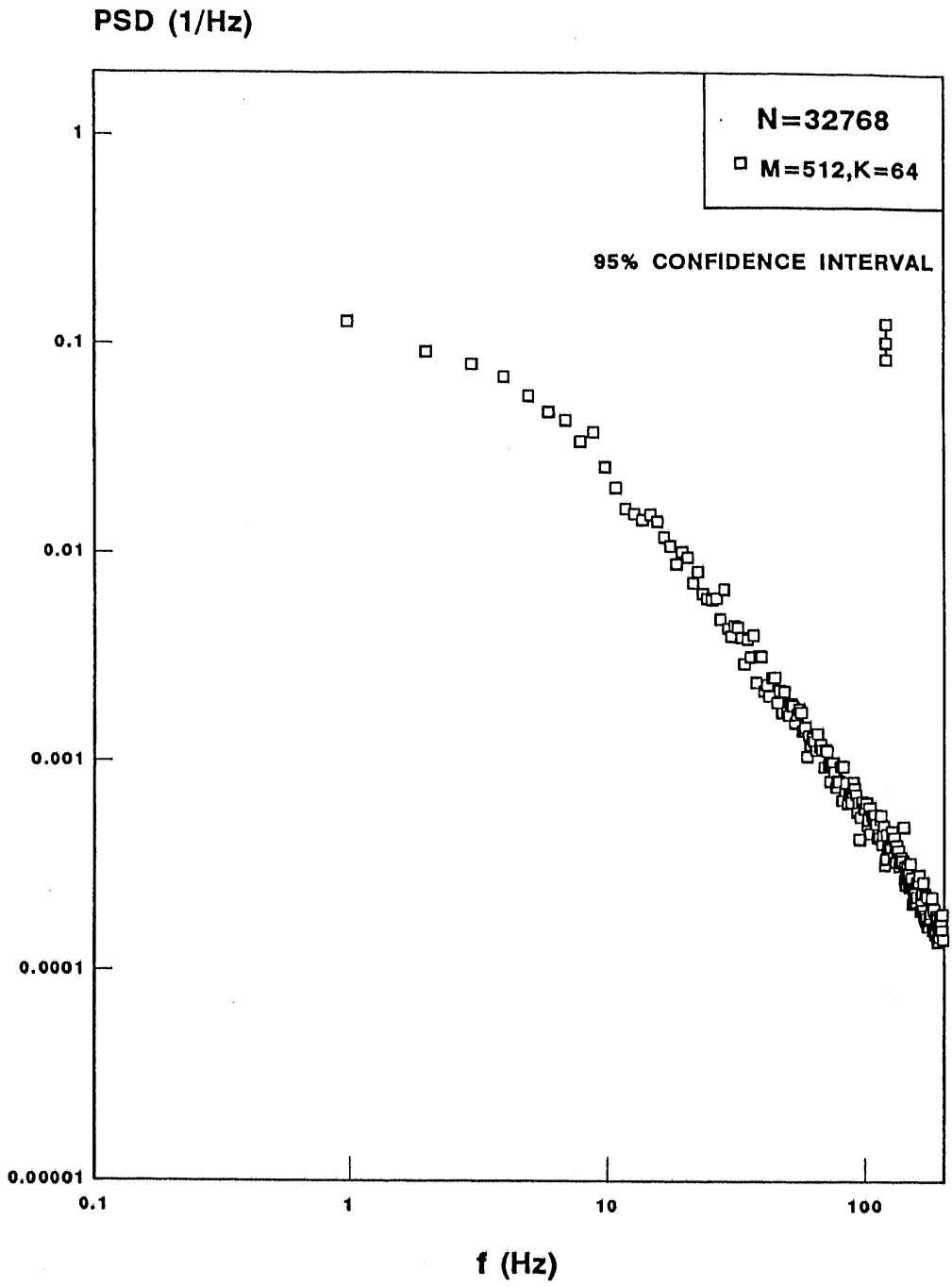


Fig. 6.6 Power spectral density at z = 60 cm

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz
1 data block, N = 32768, M = 512, K = 64

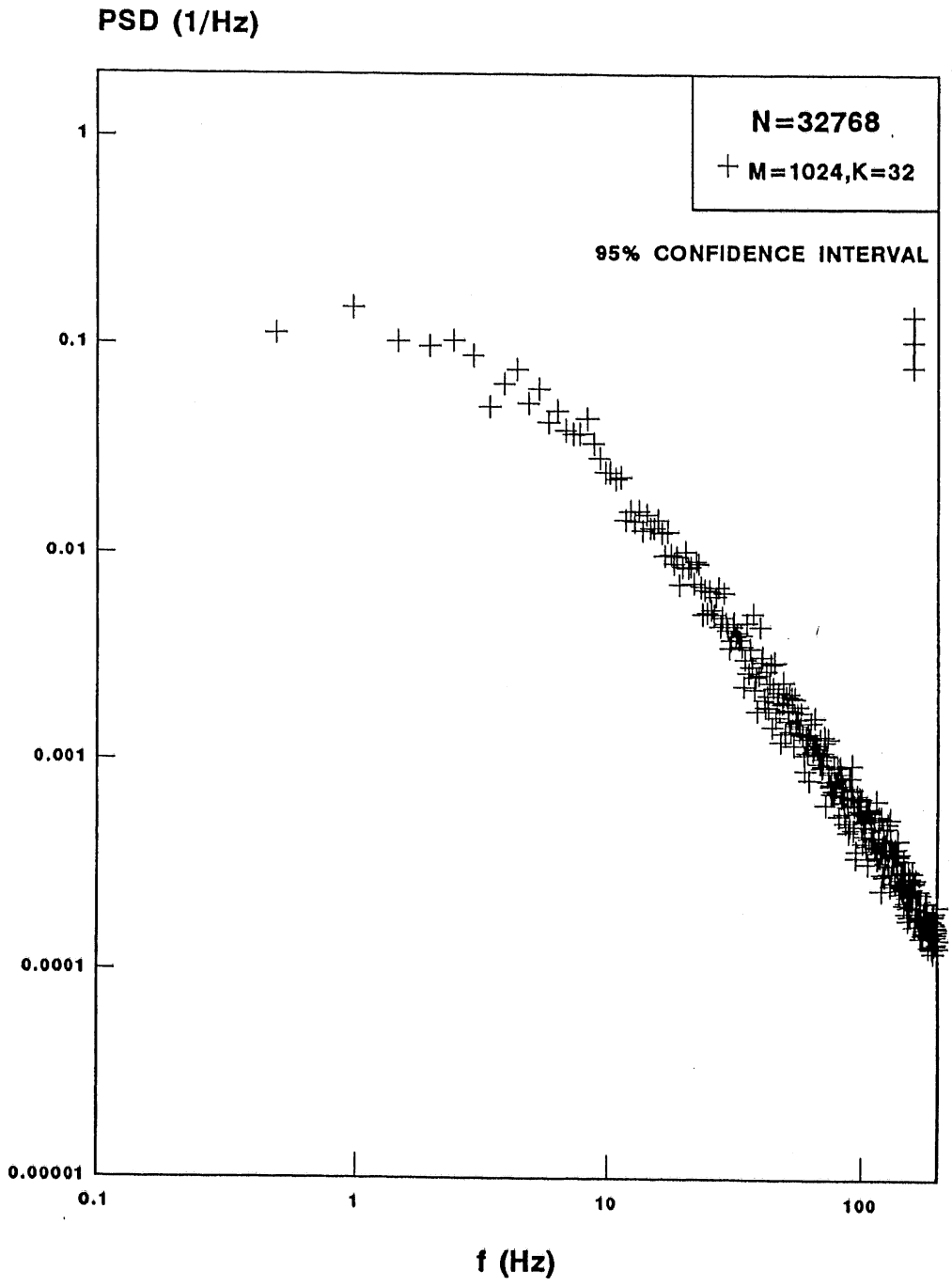


Fig. 6.7 Power spectral density at z = 60 cm

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz
1 data block, N = 32768, M = 1024, K = 32

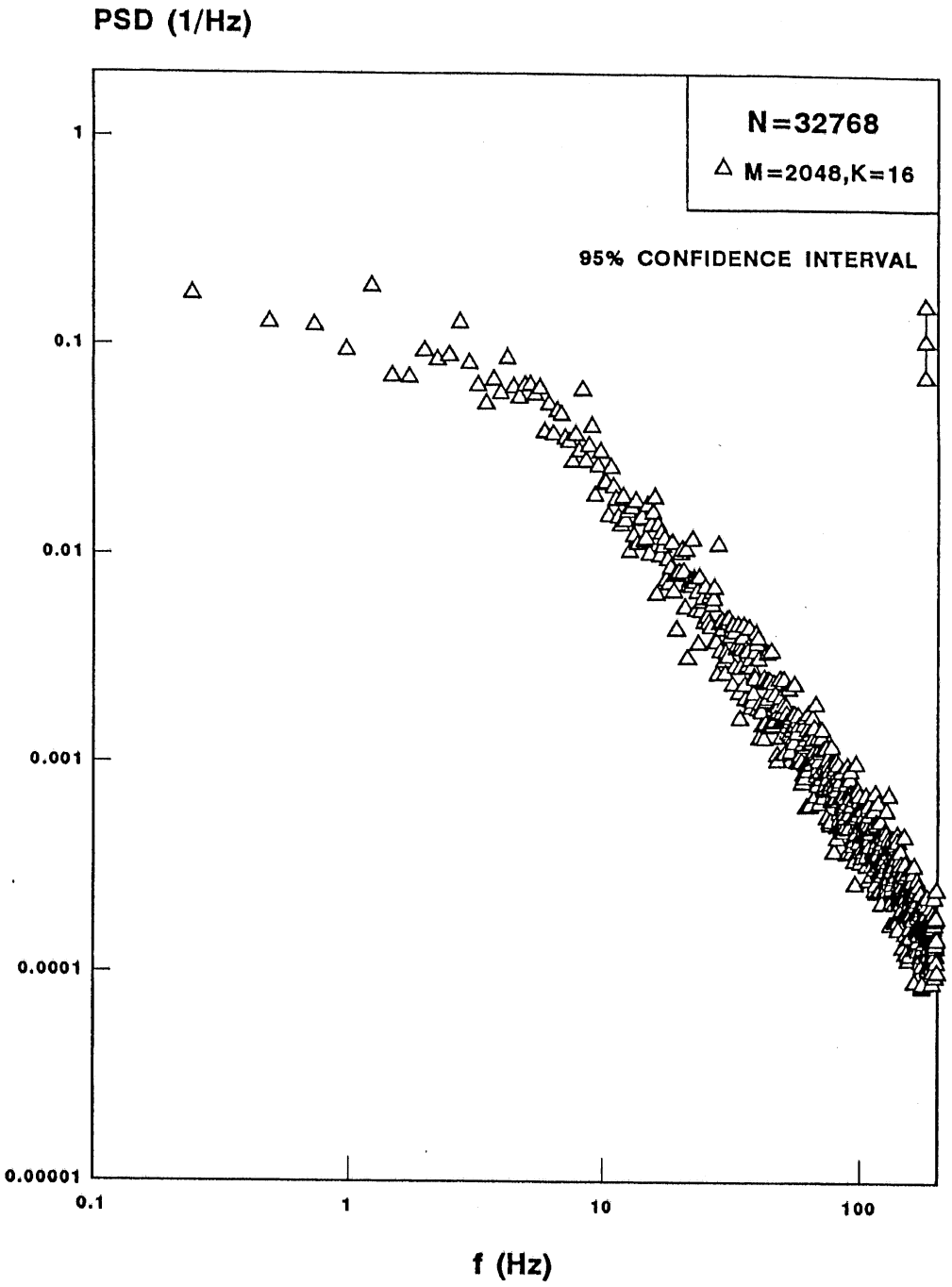


Fig. 6.8 Power spectral density at z = 60 cm

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz
1 data block, N = 32768, M = 2048, K = 16

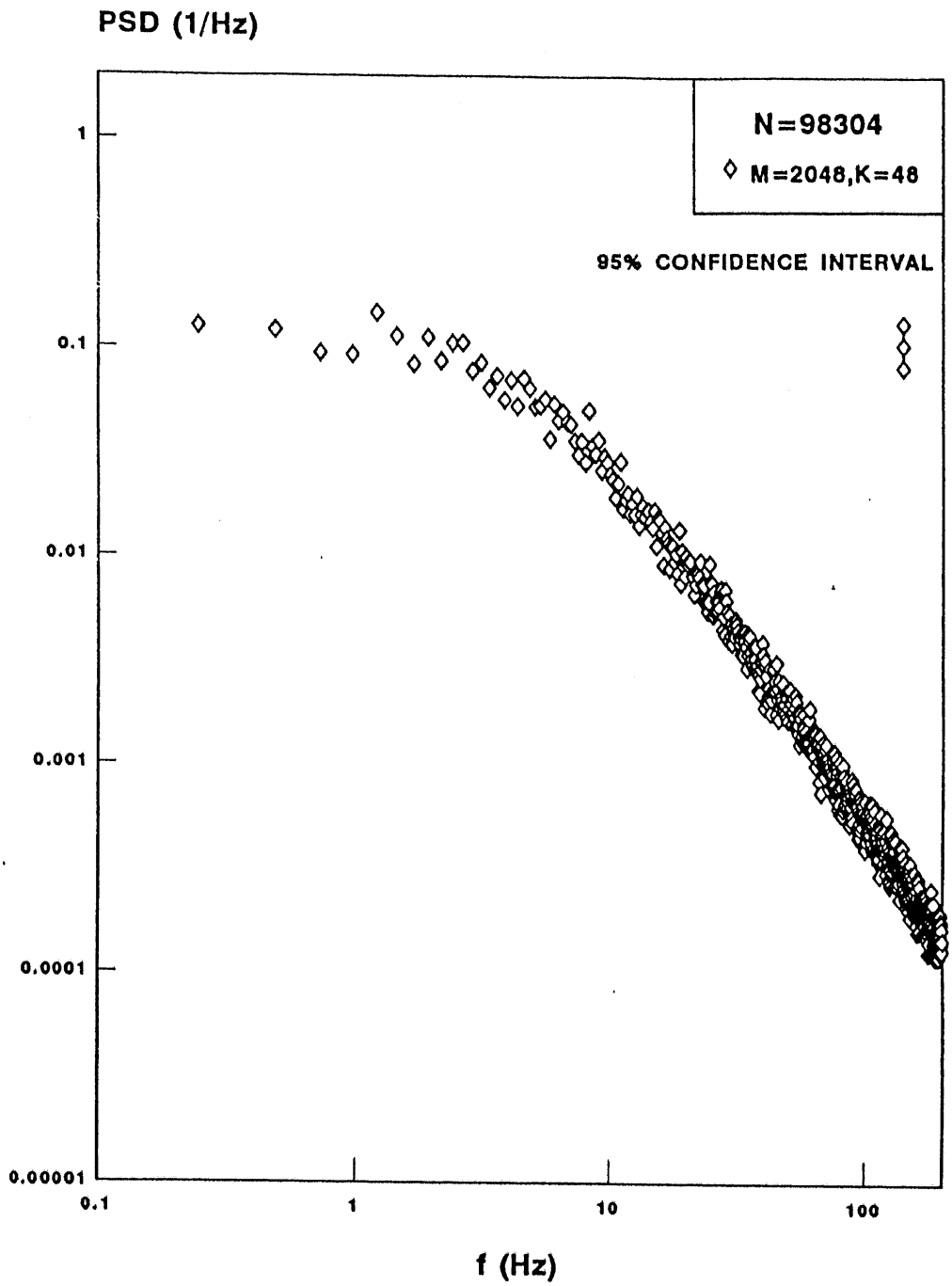


Fig. 6.9(a) Block averaged power spectral density at z = 60 cm

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz
3 data blocks, N = 98304, M = 2048, K = 48

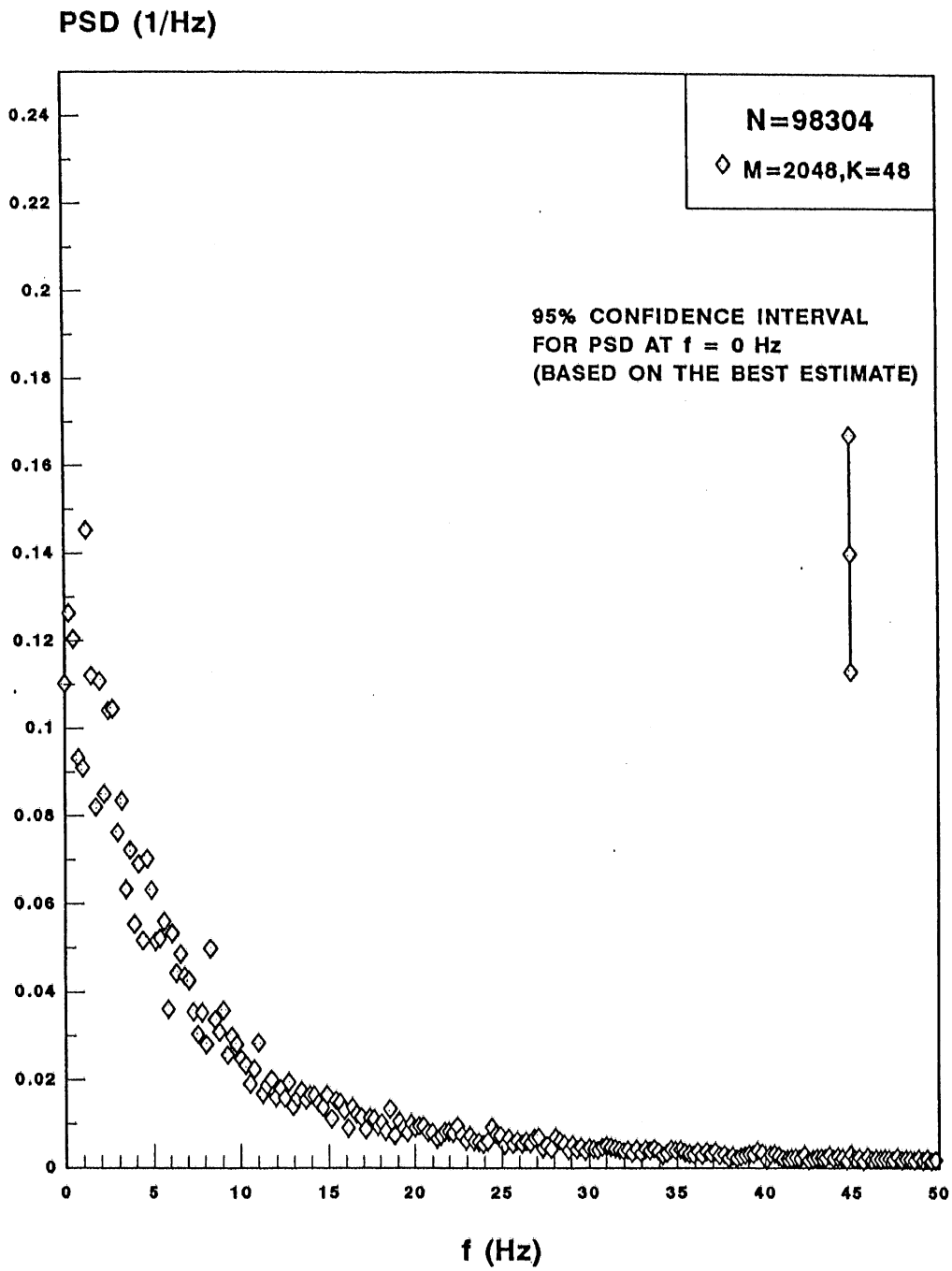


Fig. 6.9(b) Block averaged power spectral density at z = 60 cm

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz

3 data blocks, N = 98304, M = 2048, K = 48

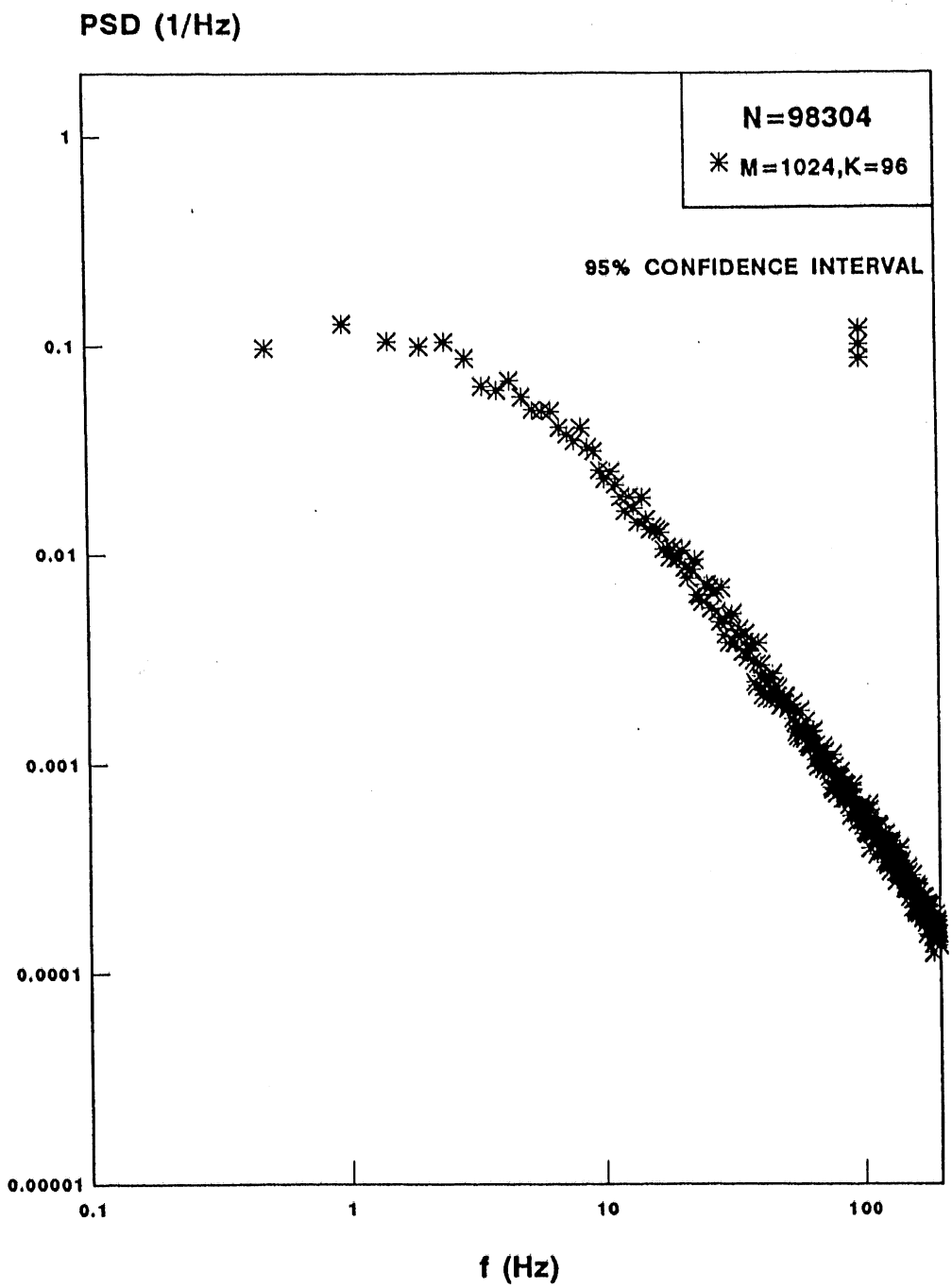


Fig. 6.10(a) Block averaged power spectral density at z = 60 cm

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz

3 data blocks, N = 98304, M = 1024, K = 96

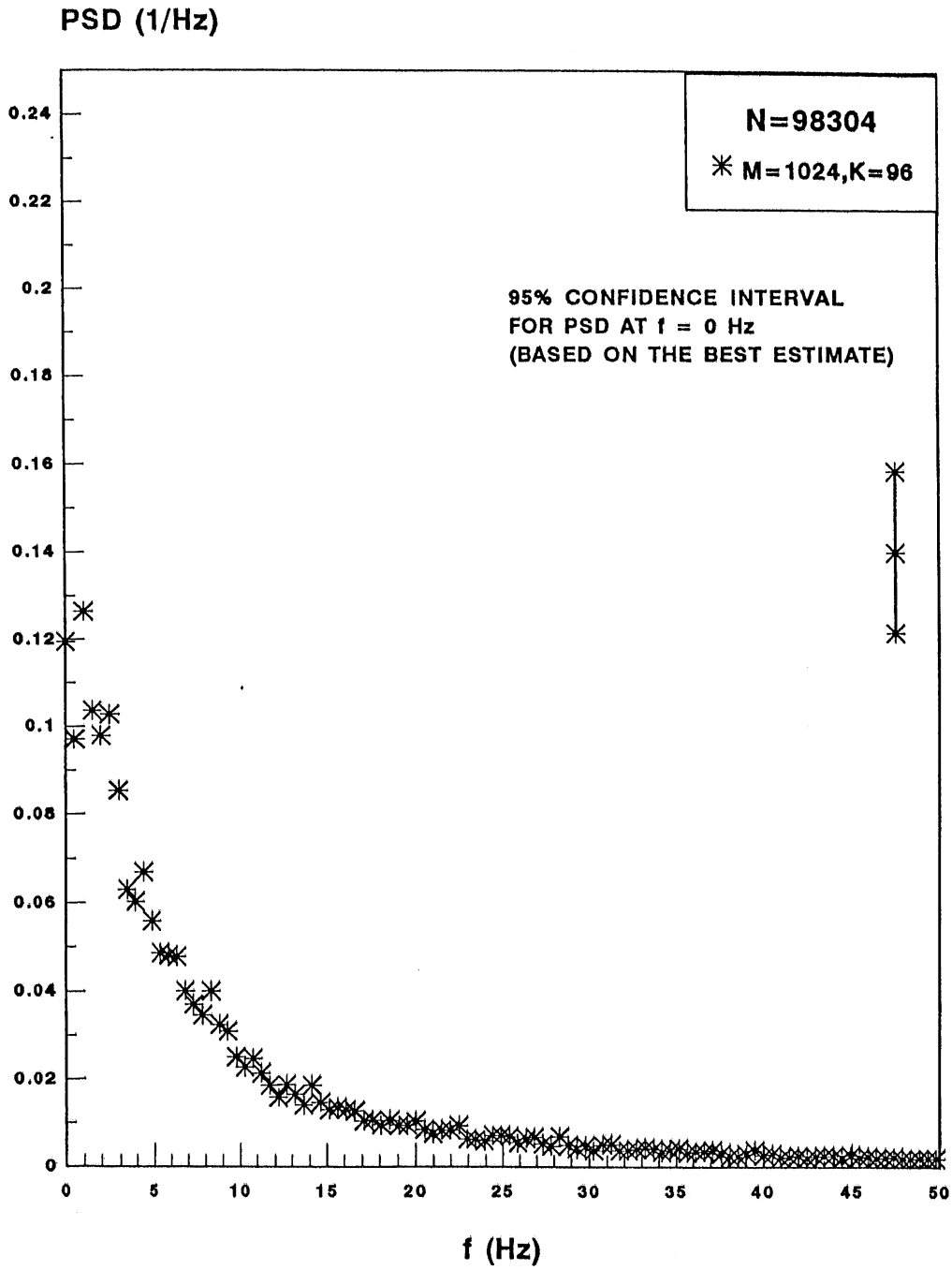


Fig. 6.10(b) Block averaged power spectral density at z = 60 cm

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz
3 data blocks, N = 98304, M = 1024, K = 96

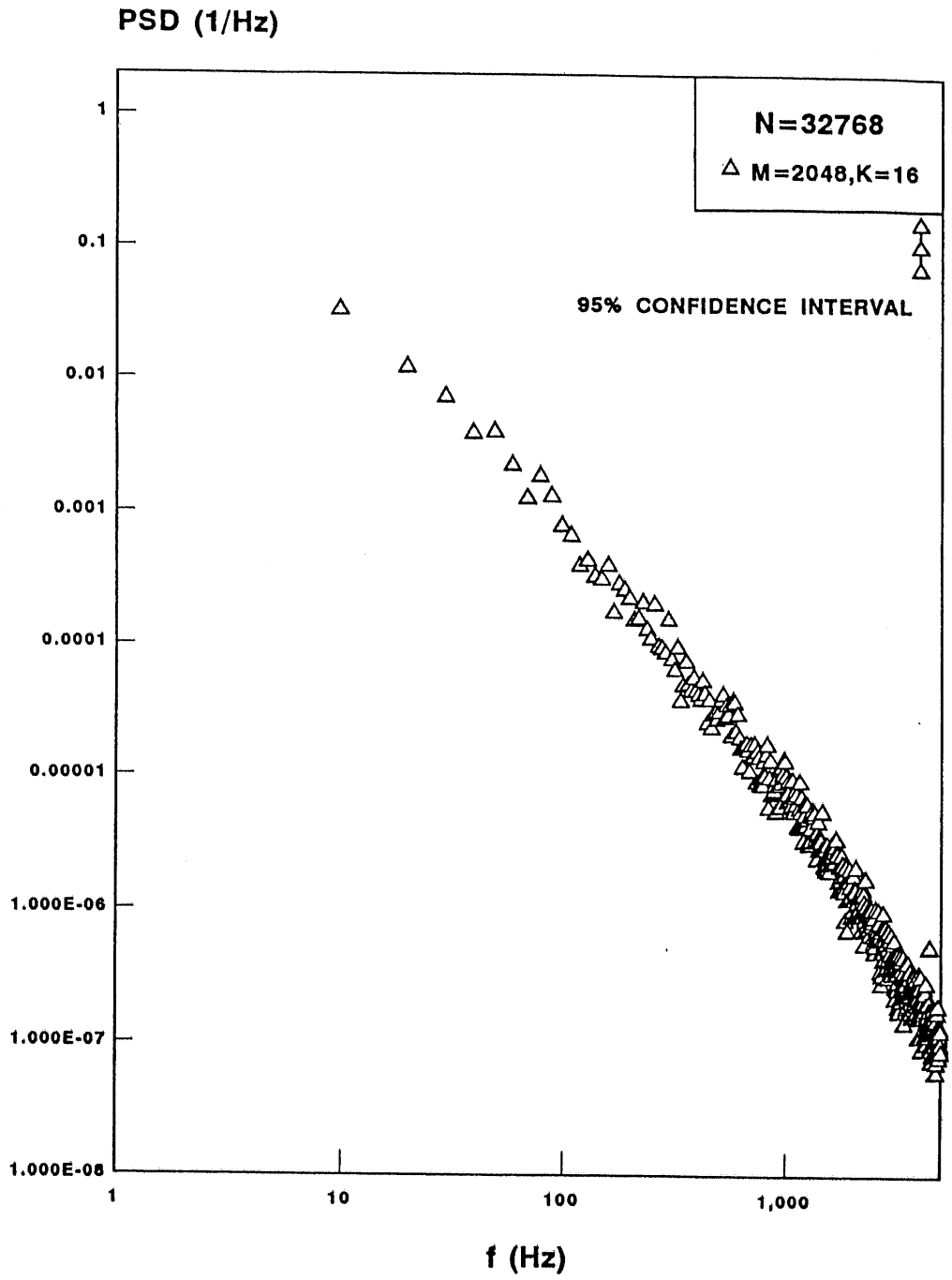


Fig. 6.11(a) Power spectral density at z = 20 cm

Δt = 0.05 ms, $f_{cut\ off}$ = 5 kHz
1 data block, N = 32768, M = 2048, K = 16

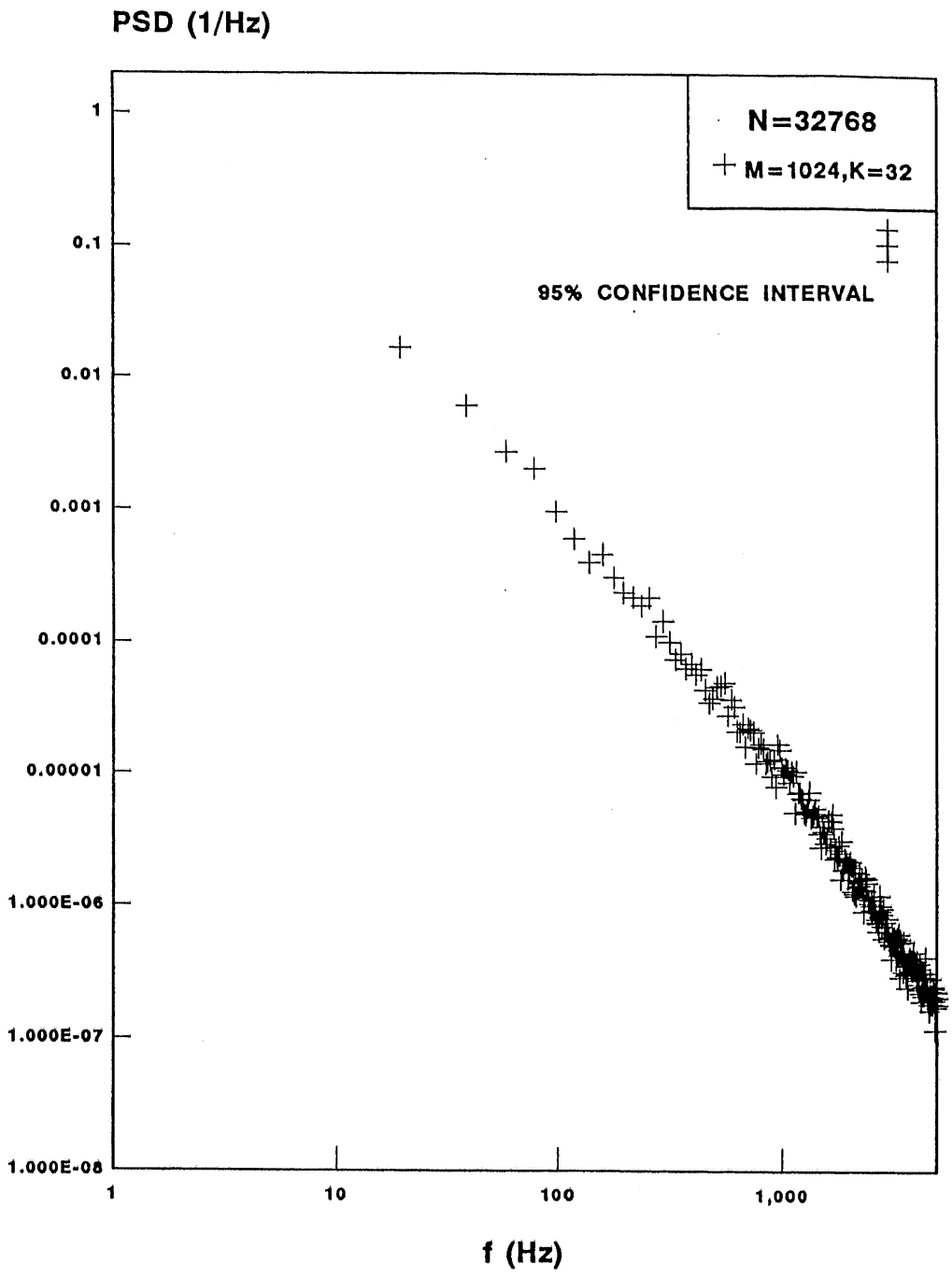


Fig. 6.11(b). Power spectral density at z = 20 cm

Δt = 0.05 ms, $f_{cut\ off}$ = 5 kHz
1 data block, N = 32768, M = 1024, K = 32

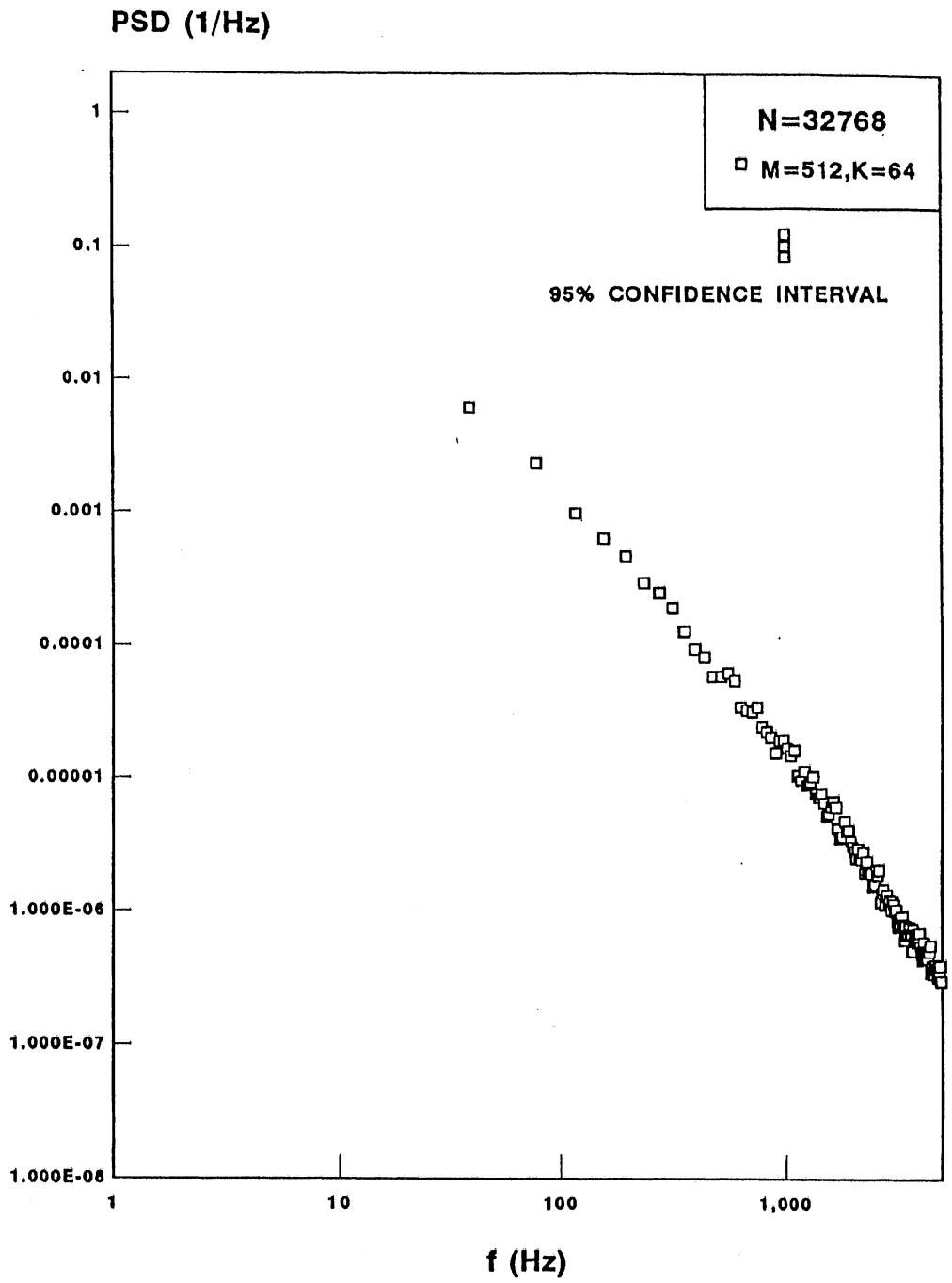


Fig. 6.11(c) Power spectral density at z = 20 cm

Δt = 0.05 ms, $f_{cut\ off}$ = 5 kHz
1 data block, N = 32768, M = 512, K = 64

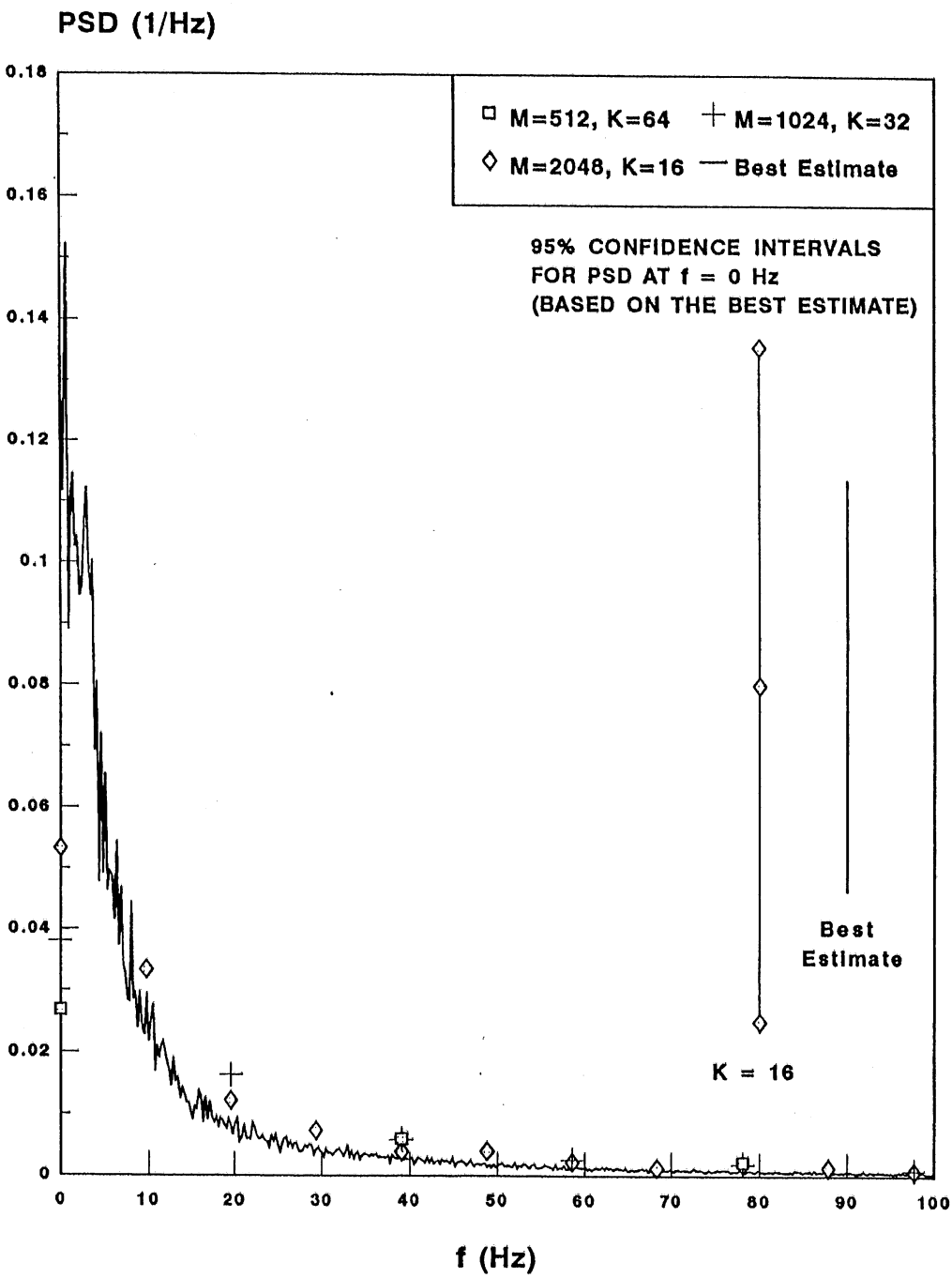


Fig. 6.11(d) Comparison with best estimate of Power spectral density

z =20cm, Δt = 0.05 ms, $f_{cut\ off}$ = 5 kHz , N = 32768

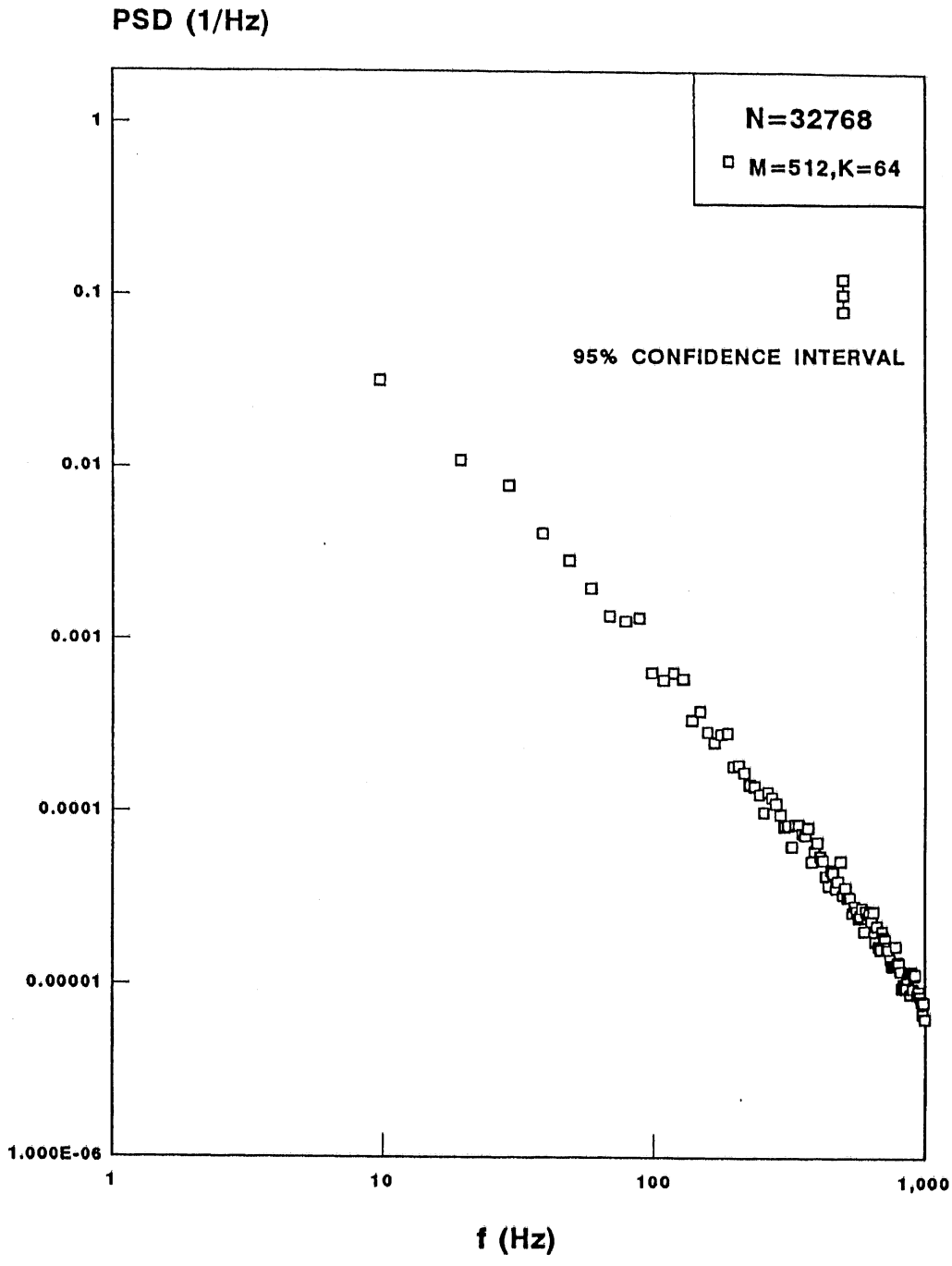


Fig. 6.12(a) Power spectral density at z = 20 cm

Δt = 0.2 ms, $f_{cut\ off}$ = 1 kHz
1 data block, N = 32768, M = 512, K = 64

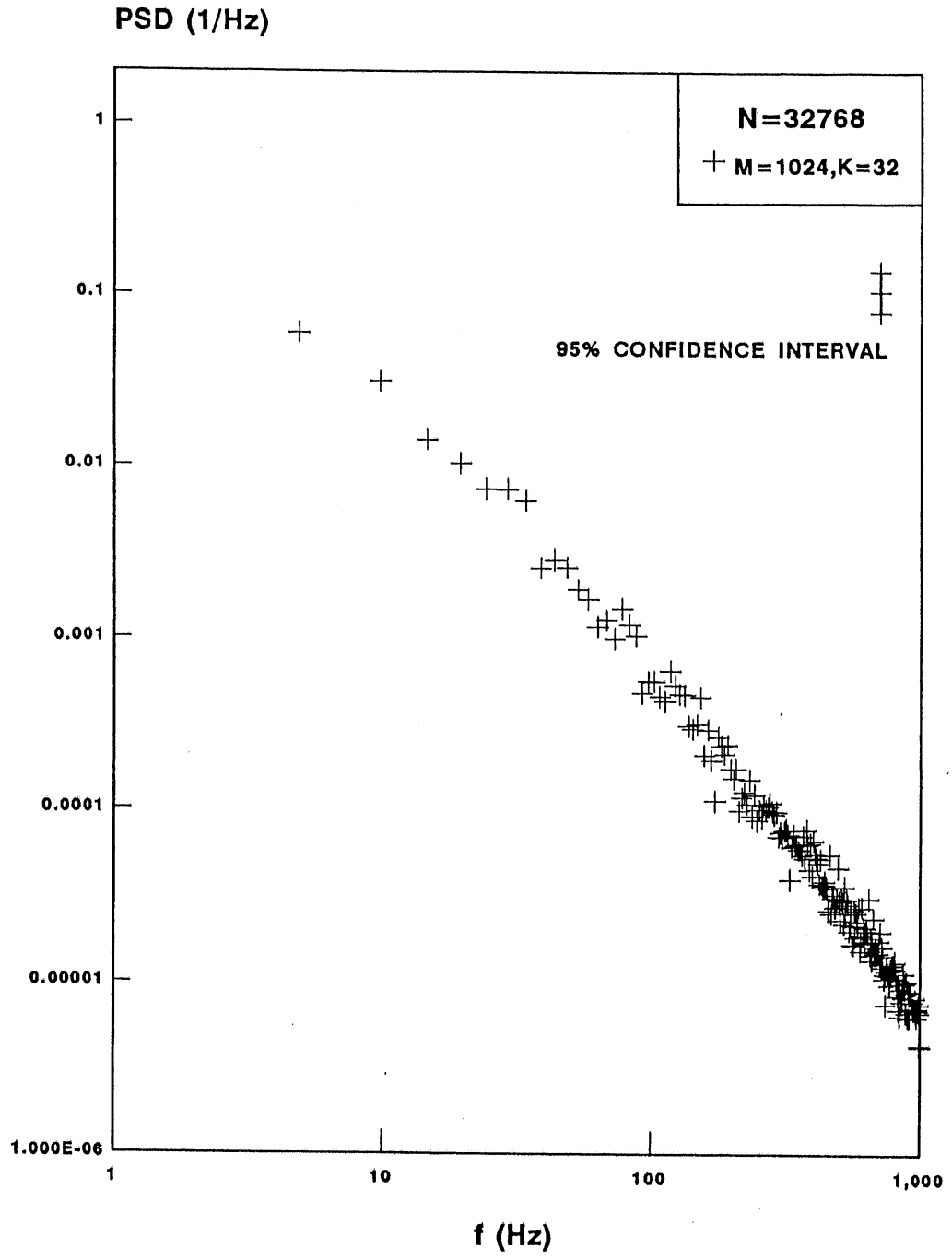


Fig. 6.12(b) Power spectral density at z = 20 cm

Δt = 0.2 ms, $f_{cut\ off}$ = 1 kHz
1 data block, N = 32768, M =1024, K = 32

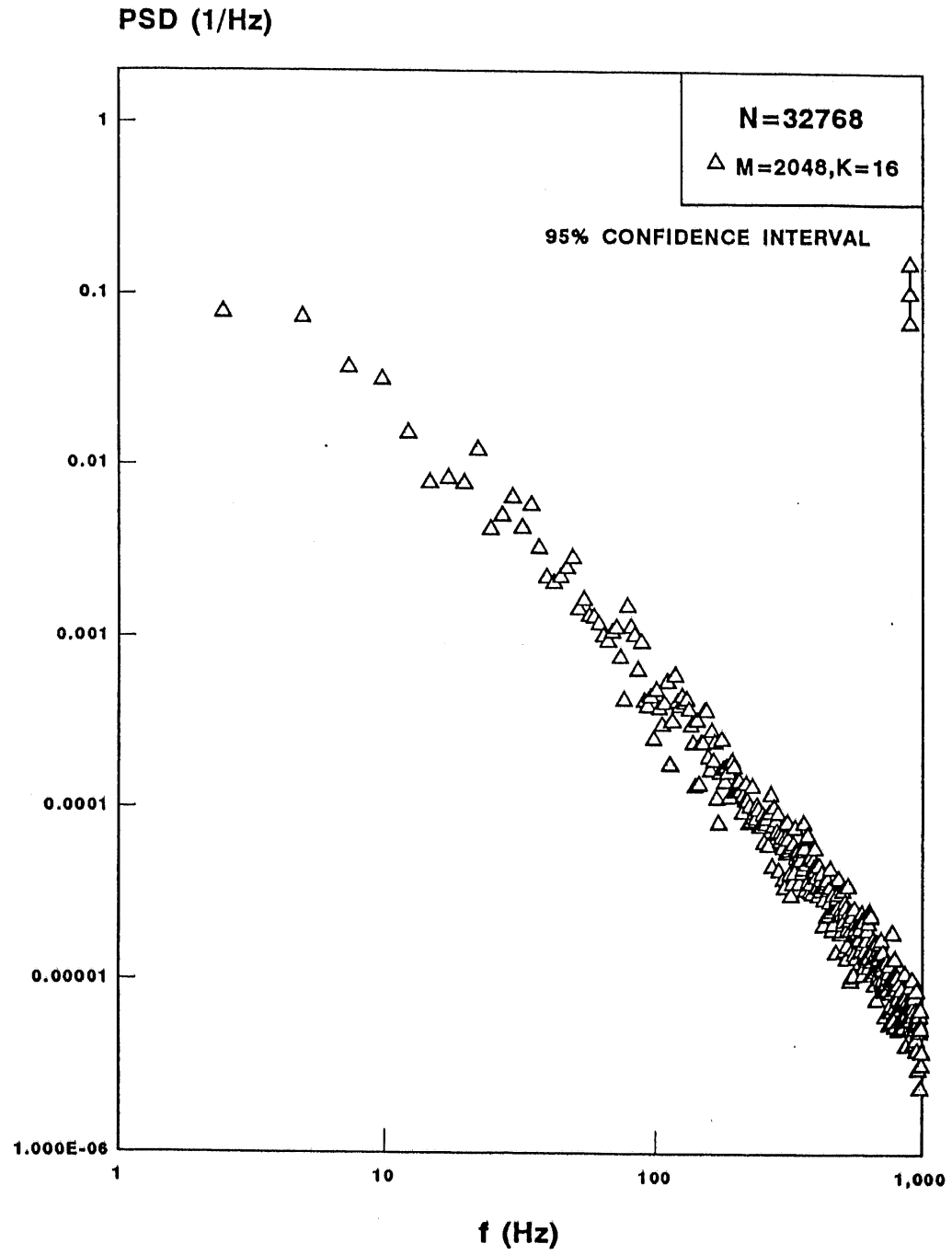


Fig. 6.12(c) Power spectral density at z = 20 cm

Δt = 0.2 ms, $f_{cut\ off}$ = 1 kHz
1 data block, N = 32768, M = 2048, K = 16

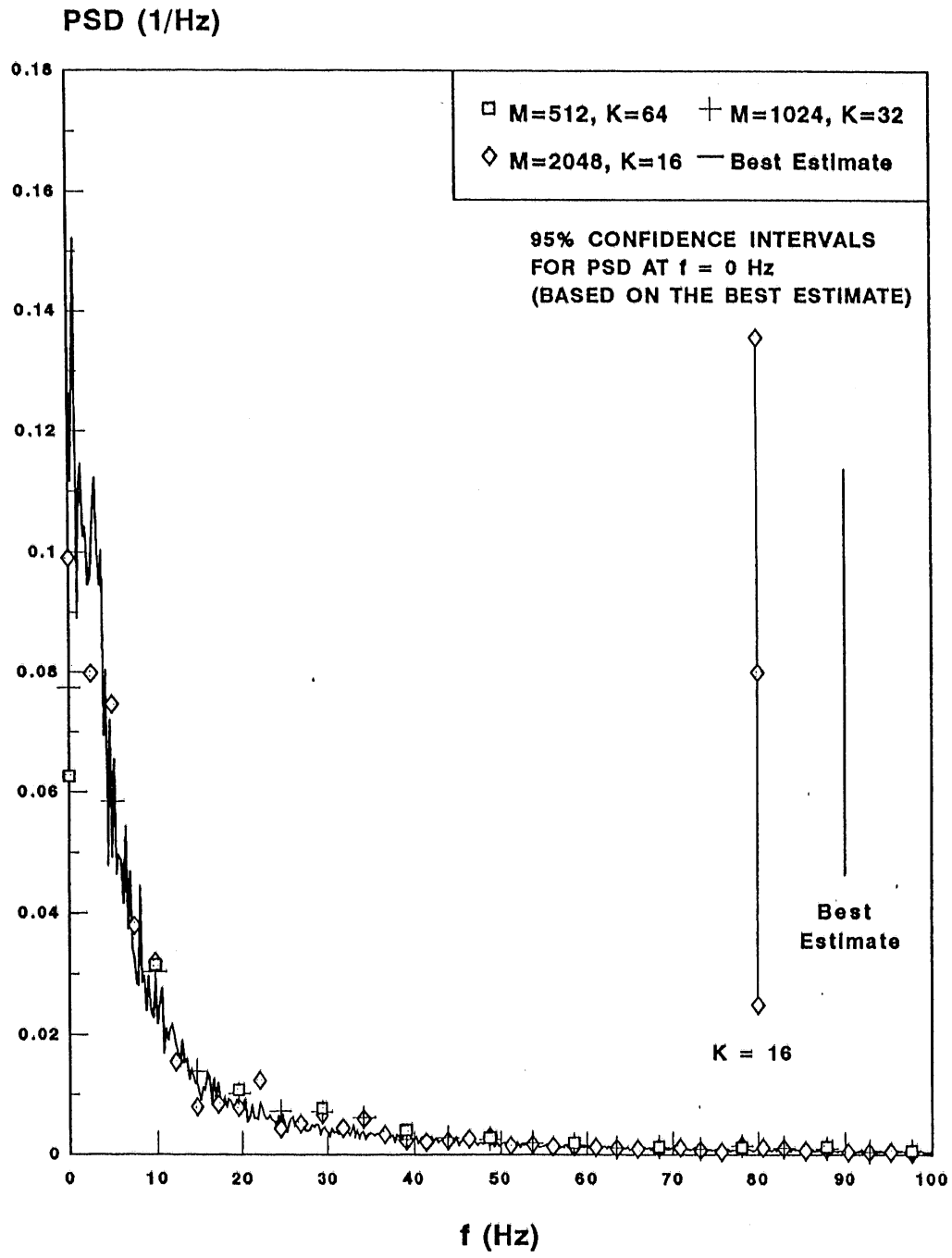


Fig. 6.12(d) Comparison with best estimate of Power spectral density

$z = 20$ cm, $\Delta t = 0.2$ ms, $f_{cut\ off} = 1$ kHz , $N = 32768$

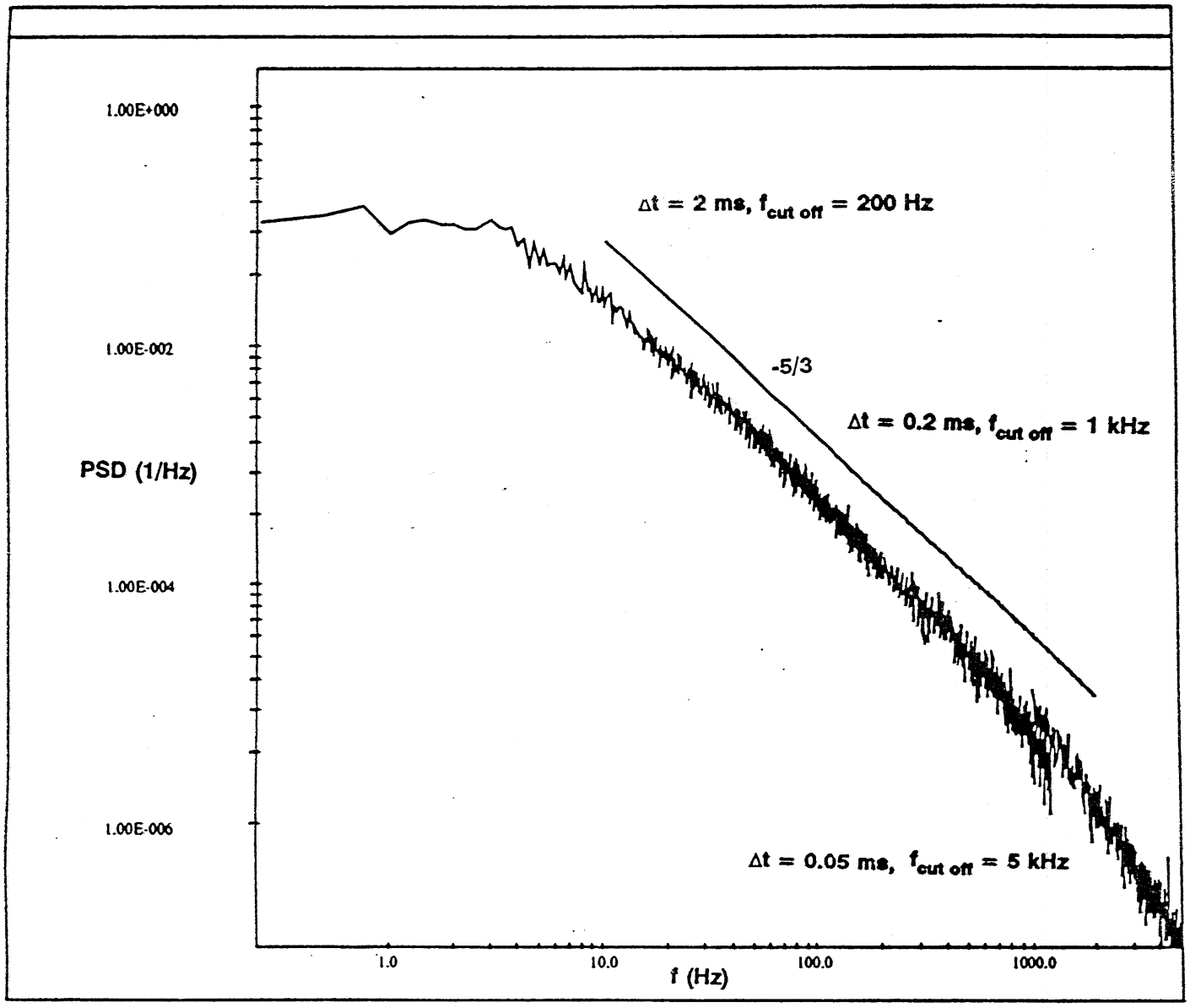


Fig. 6.13(a) Power spectral density from 0 - 5 kHz at z = 20 cm

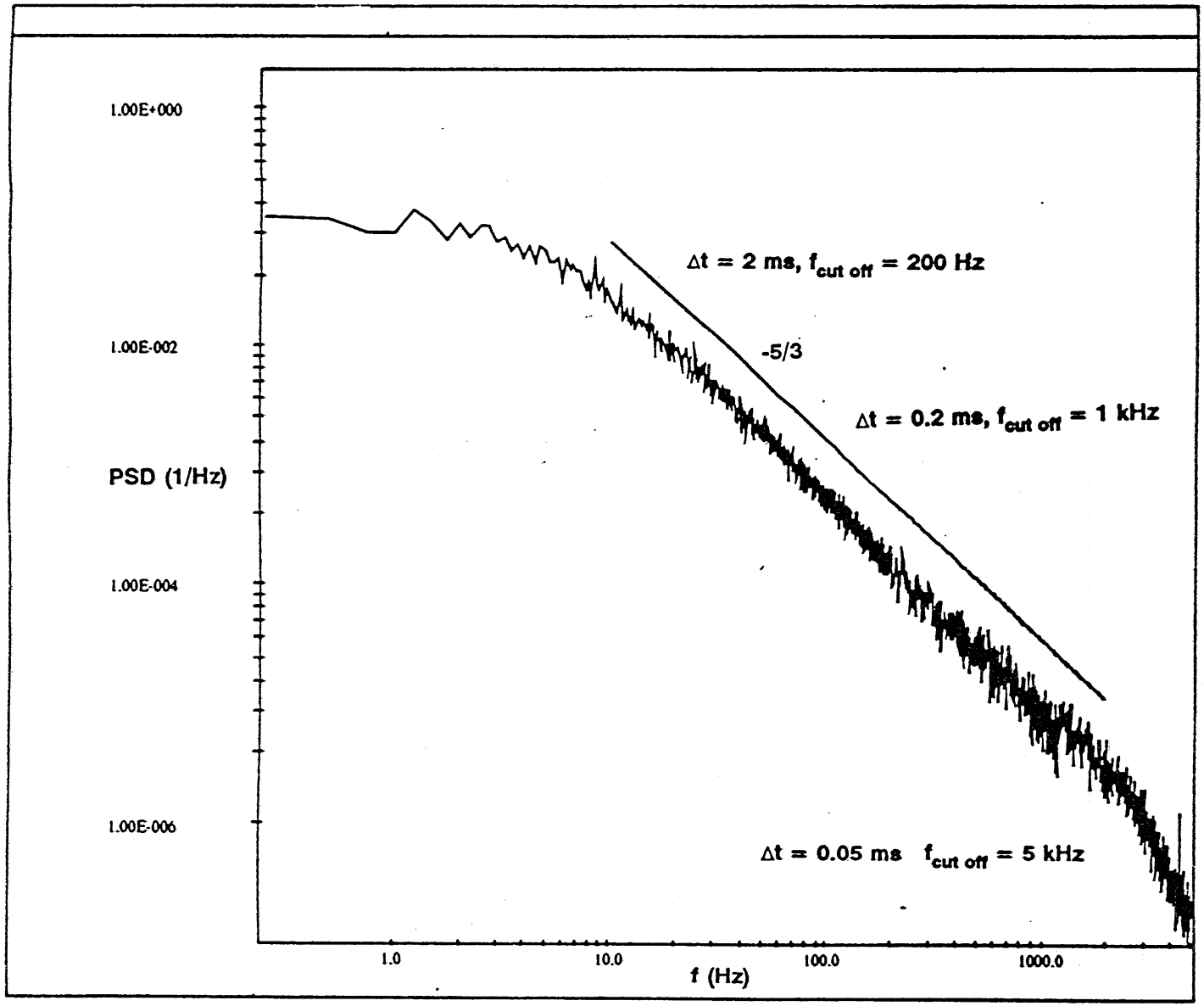


Fig. 6.13(b) Power spectral density from 0 - 5 kHz at z = 60 cm

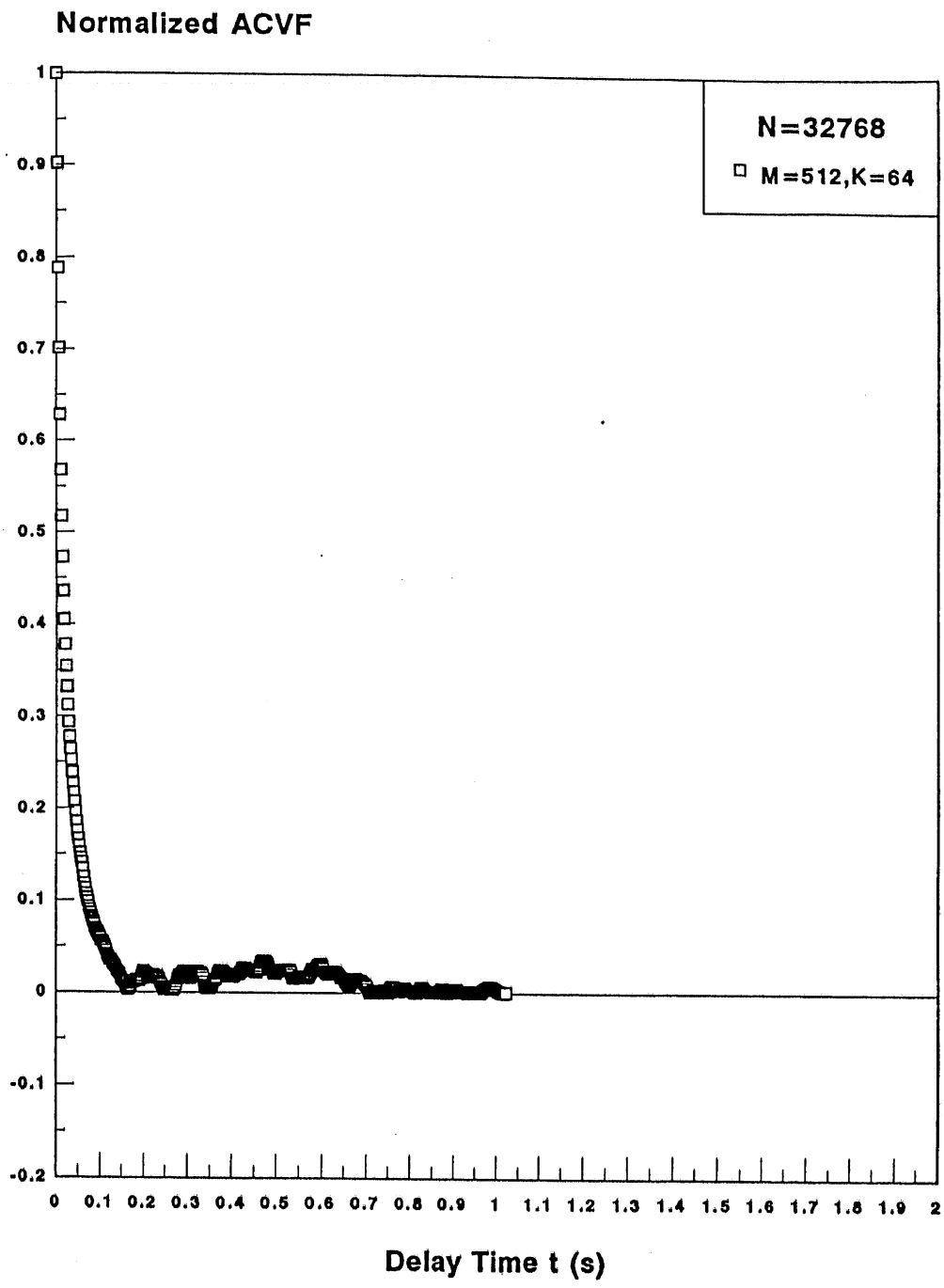


Fig. 6.14(a) Normalized autocovariance at z = 5 cm

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz
N = 32768, M = 512, K = 64

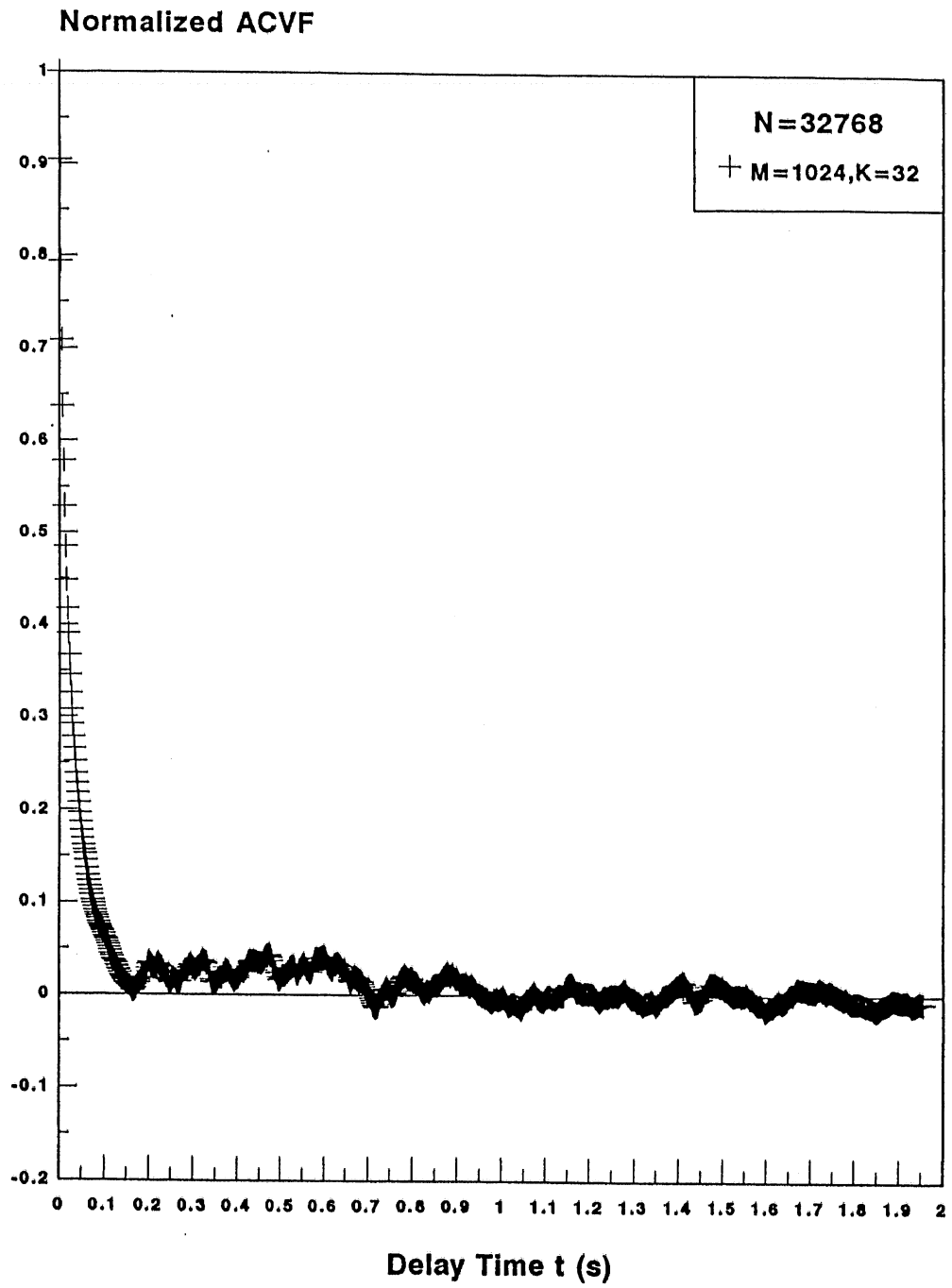


Fig. 6.14(b) Normalized autocovariance at z = 5 cm

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz
N = 32768, M = 1024, K = 32

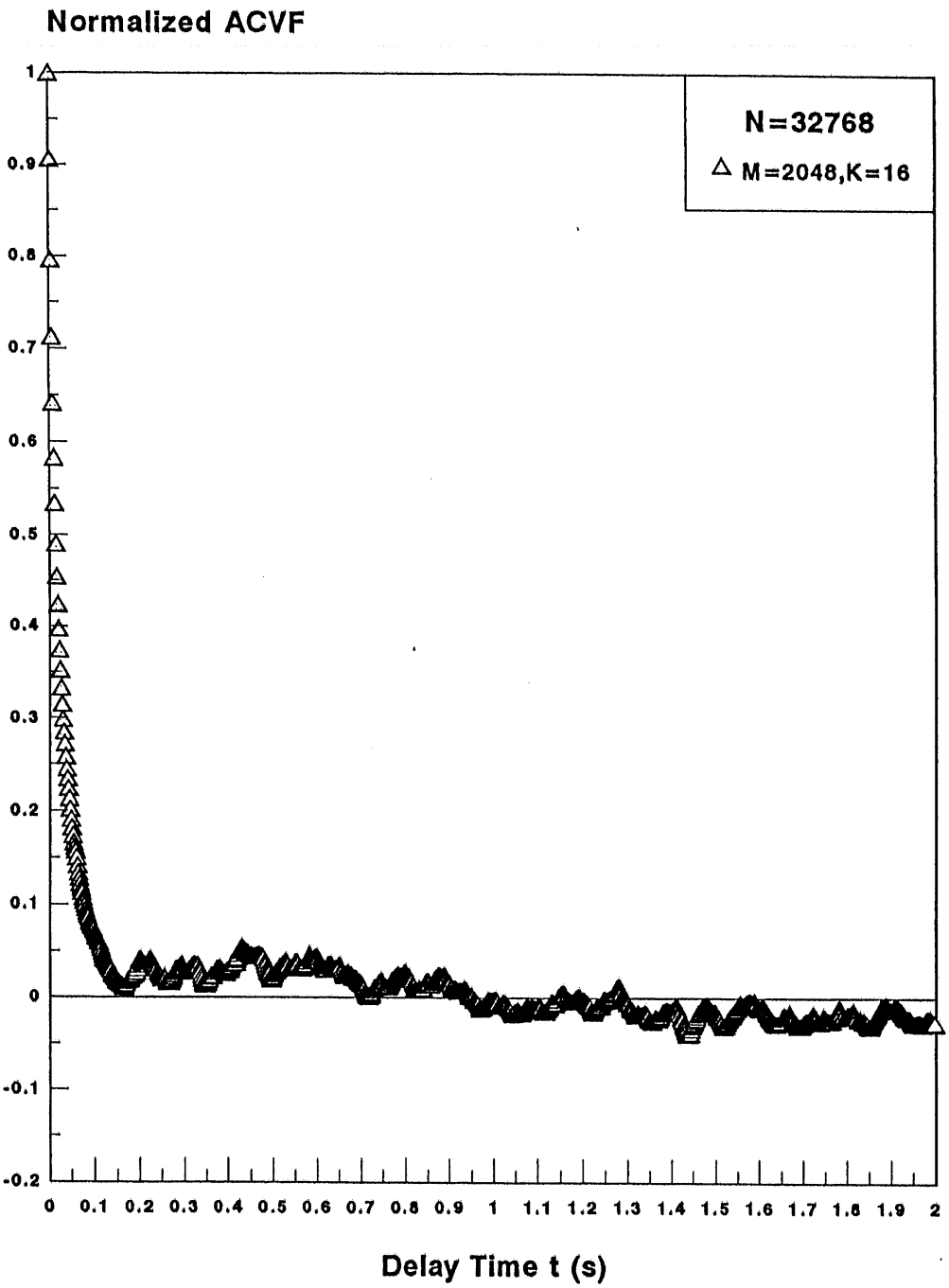


Fig. 6.14(c) Normalized autocovariance at z = 5 cm

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz
N = 32768, M = 2048, K = 16

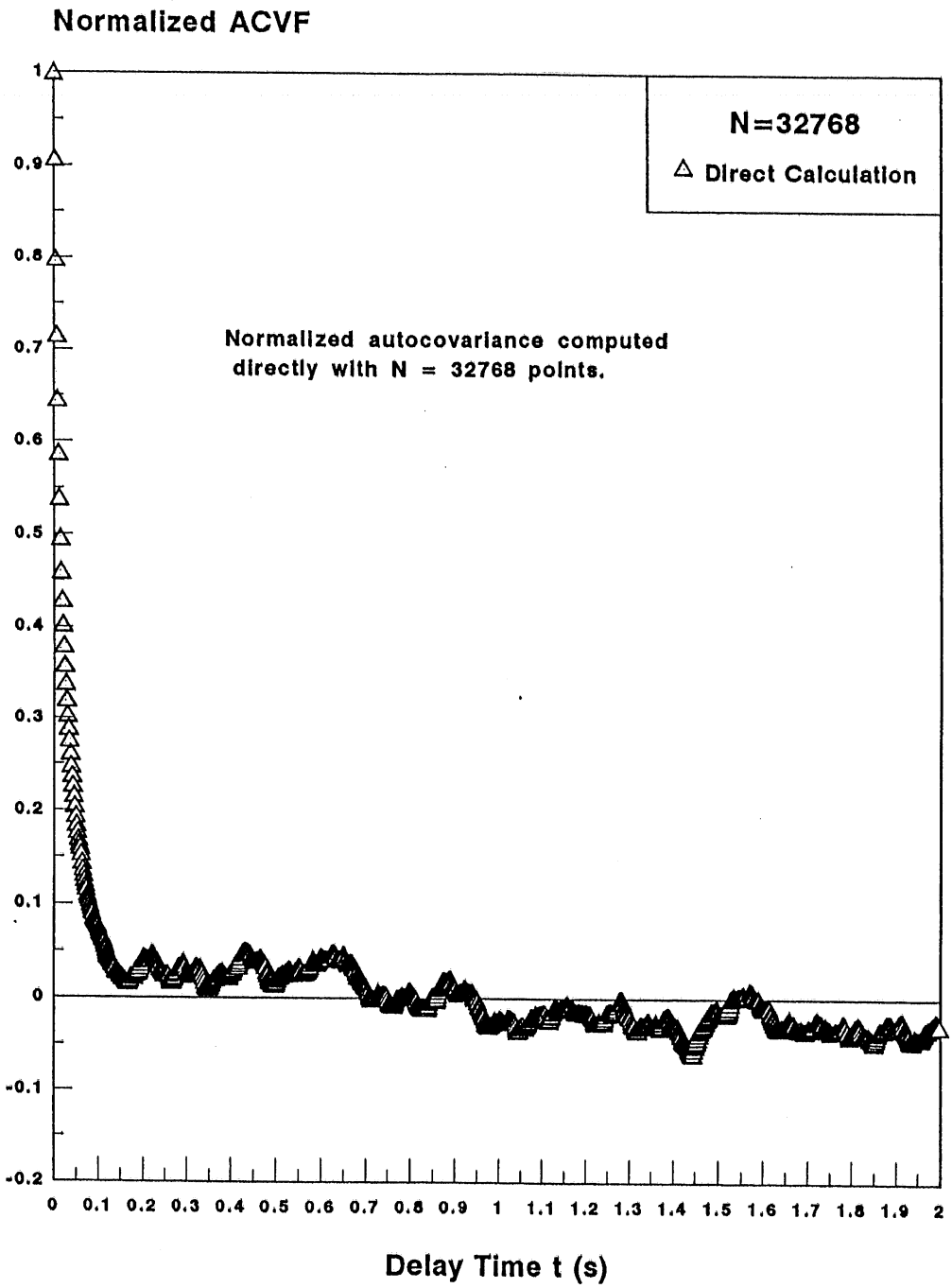


Fig. 6.15(a) Normalized autocovariance at z = 5 cm
Direct calculation

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz
N = 32768

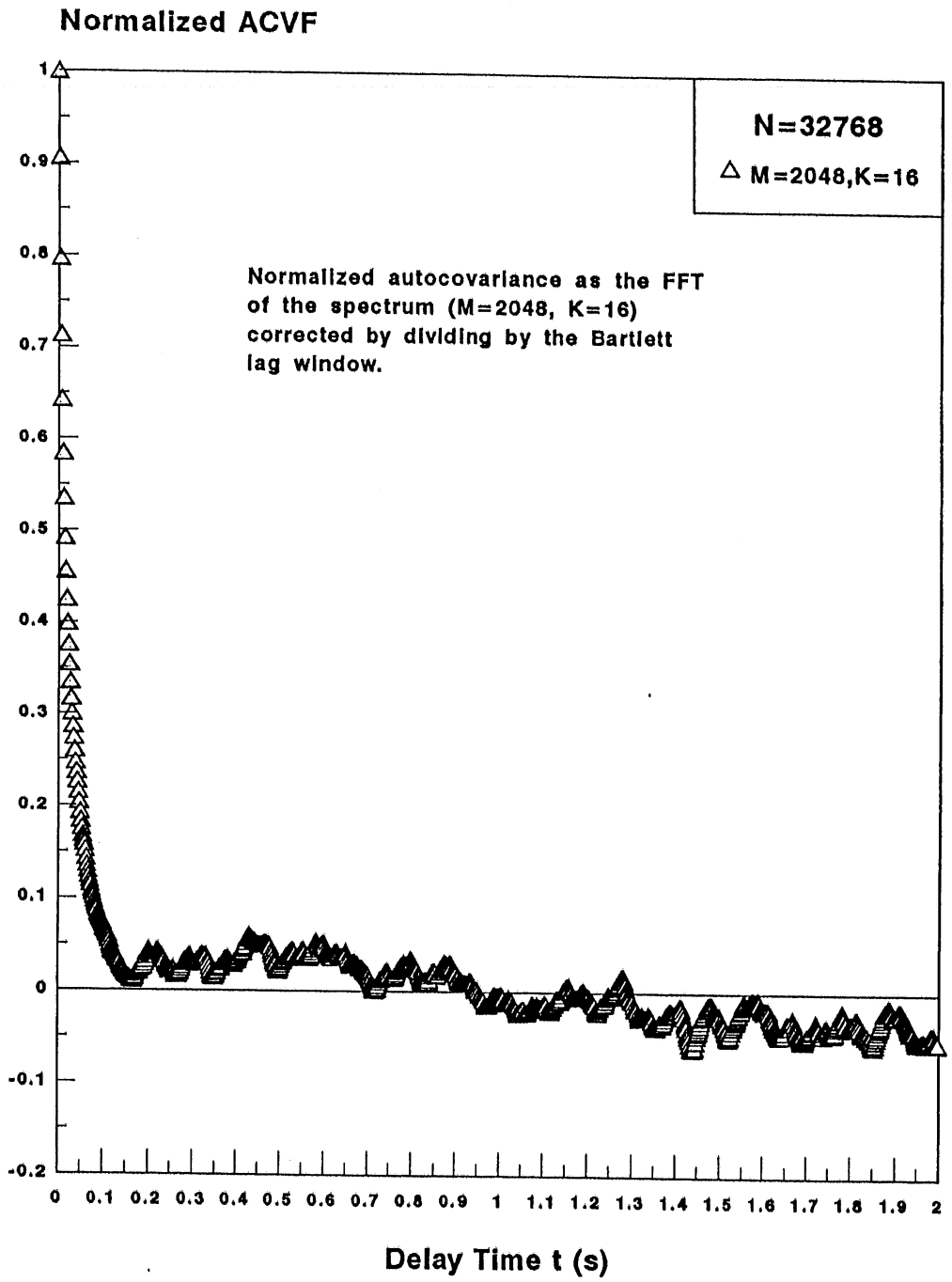


Fig. 6.15(b) Normalized autocovariance of Fig. 6.14(c) corrected by dividing by the Bartlett window

Δt = 2 ms, $f_{cut\ off}$ = 200 Hz, N = 32768, M =2048

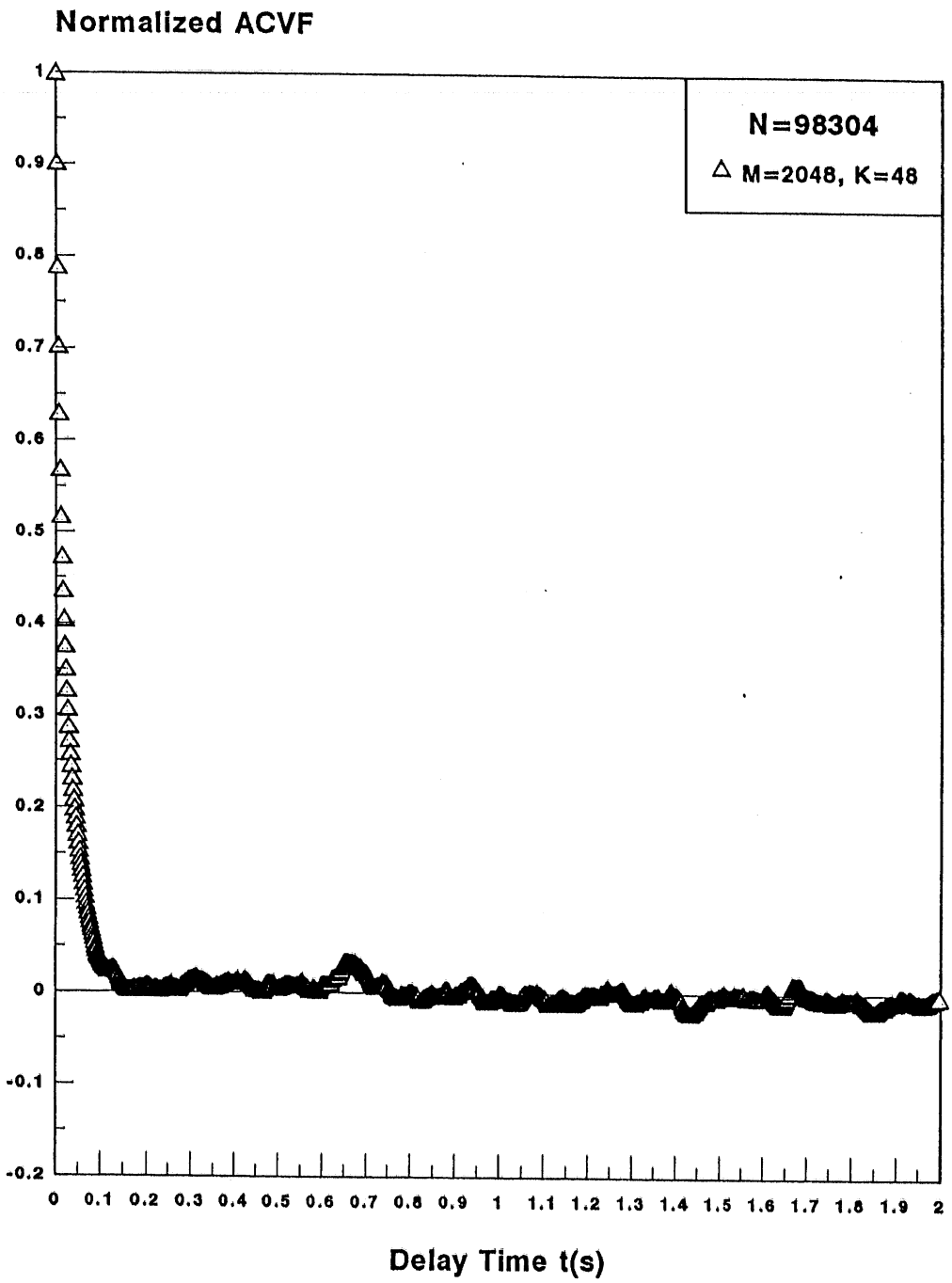


Fig. 6.16 Normalized autocovariance from best estimate of spectrum
z = 5 cm, Δt = 2 ms, $f_{cut\ off}$ = 200 Hz
N = 98304, M = 2048, K = 48

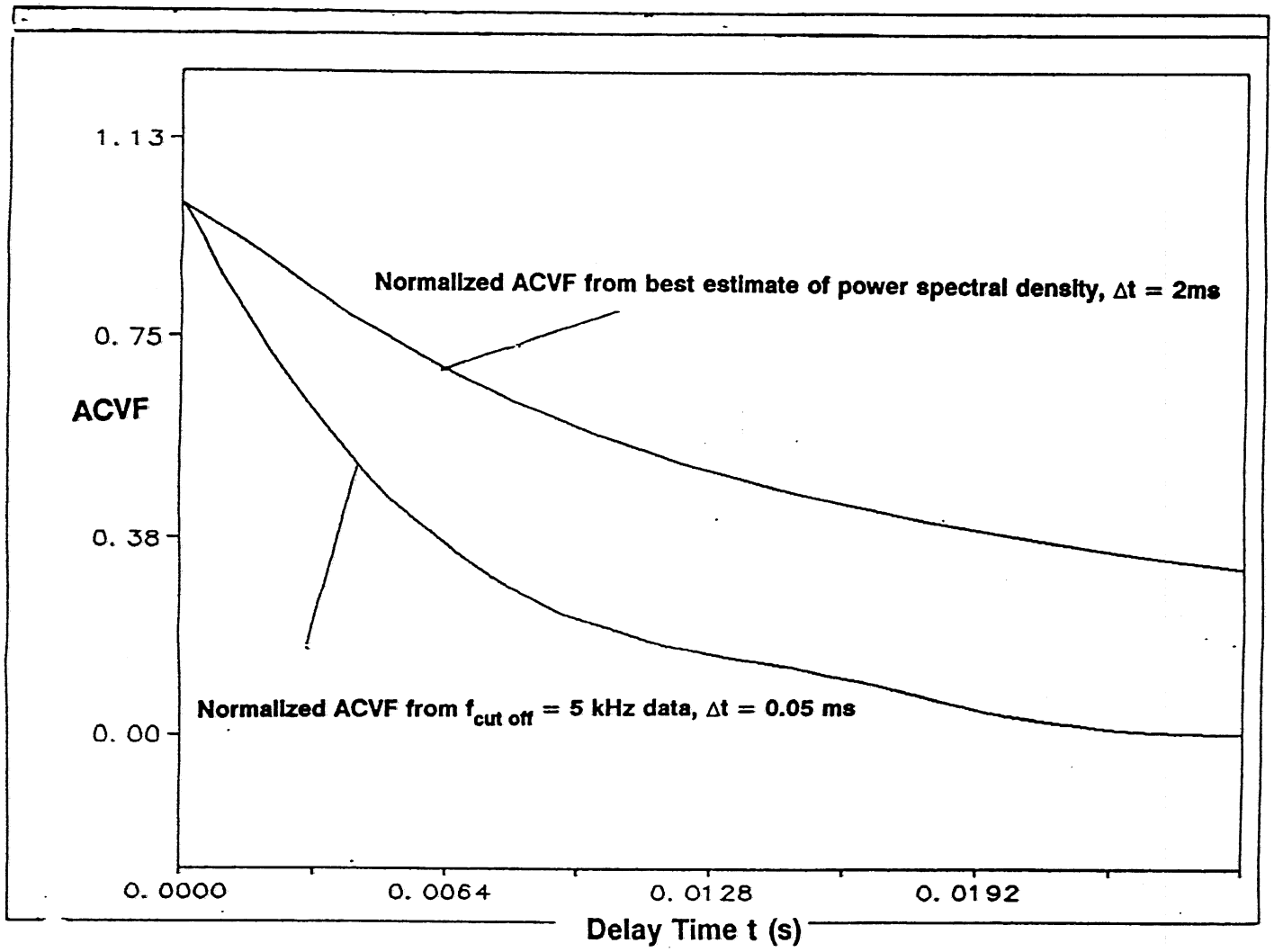


Fig. 6.17 Comparison of normalized autocovariance obtained from the 5khz data with that from best estimate of power spectrum

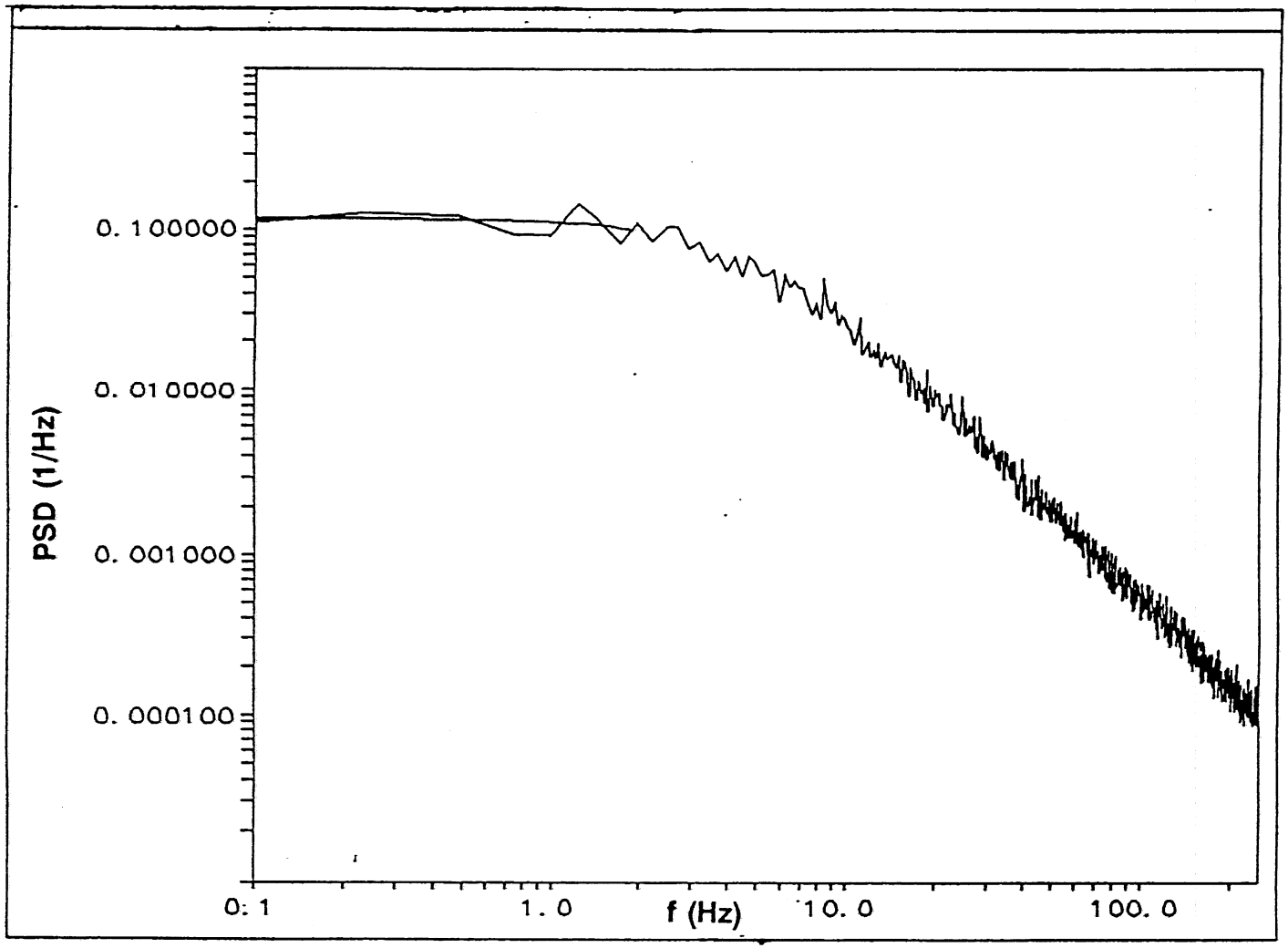


Fig. 6.18 Parabolic fit to first few points of the power spectral density
z = 60 cm, N = 98304, M =2048, K = 48, Δt = 2 ms

7. CHARACTERISTICS OF DYNAMIC SIGNAL ANALYZER

7.1. Introduction

The Hewlett Packard dynamic signal analyzer (HP35660A) is a self-contained instrument designed for digital analysis of experimental data, in particular for the determination of single-channel linear and power spectral densities and two-channel coherence and cross-spectral density functions. A brief discussion of the principles of operation of this instrument is presented in this appendix, in the light of the material in this report.

The instrument uses M = 1024 points for calculation of the sample spectra. In the terminology of this report, the subseries length used for calculation of the spectra is T = 1024 Δt, where the sampling interval Δt can be varied depending on the frequency content of the signal. To avoid aliasing in the frequency domain, frequencies above $1/(2\Delta t)$ must be filtered out. This is done using both analog and digital filtering, as explained subsequently. The filtering operation has a ramp response characteristic and actually suppresses signal components with frequencies above $1/(2.56\Delta t)$. The number of points in the frequency domain, M_f, is selected equal to M = 1024 (Ch. 4, p. 30). Considering only nonnegative frequencies, the number of frequency points is 512. To take into account the response charateristic of the anti-alias filter, the analyzer displays only 512/(2.56/2) = 400 points and discards the rest. Let f_{disp} be the maximum displayed frequency and $f_{max} = 1/(2\Delta t)$ the maximum frequency in the sample. The following relations then apply:

- frequency resolution $\Delta f = f_{disp}/400$ hz
- subseries length $T = 1024(\Delta t) = 1/\Delta f = 400/f_{disp}$ (seconds)
- $\Delta t = (400/f_{disp})/1024 = 1/(2.56\ f_{disp})$ (seconds)
- $f_{max} = 1/(2\Delta t) = f_{disp}\ (512/400)$ hz.

The sampling interval Δt has a minimum value Δt_{min} dictated by a front anti-alias filter through which the signal is made to pass before being digitally sampled. The instrument's maximum display frequency in the single channel mode is 102.4Khz. (When both channels of the instrument are used the maximum display frequency is one half this value, 51.2Khz.) For the maximum display frequency of 102.4Khz, we have Δf_{max} = 102,400/400 hz = 256hz, T = $M\Delta t_{min} = 1/\Delta f_{max}$ = 3.90625ms, Δt_{min} = 3.90625/1024 ms $\approx 3.815 \times 10^{-3}$ms, and $f_{max} = 1/(2\Delta t_{min})$ = 131.072Khz. The front anti-alias filter eliminates all

components with frequencies larger than this value of f_{max}.

From the instrument's anti-alias filter the signal passes through an A/D converter which always digitizes the signal at Δt_{min}, regardless of the actual value of Δt selected. The digitized signal then passes through a digital filter which ensures protection from aliasing for any frequency span of interest (that is, for any selected value of $\Delta t > \Delta t_{min}$). The relationships between Δt and the parameters f_{disp}, f_{max}, Δf, and T have been given above. The output from the digital filter goes into a buffer. On the signal in the buffer we have the choice of applying several data windows: Hanning, uniform, flat top, force or exponential. A discussion of some of the mathematics associated with the use of these data windows and the calculation of linear and power spectral densities is presented in the next section. The H.P. dynamic signal analyzer uses a radix-2 FFT algorithm to perform a discrete fast Fourier transform on the windowed signal from the buffer.

The transformed signal, now in the frequency domain, is displayed depending on the user's choice. The H.P. analyzer allows for averaging up to 99,999 sample spectra, for a total record length equal to $T = 1024(\Delta t)$ times the number of averages. Confidence intervals can then be obtained as explained in Chapter 4, p. 42.

7.2. Measurement of Power Spectral Densities: Basic Relations

The Hanning, uniform, flat top, force and exponential windows in the H.P. analyzer are *data windows*, applied directly to the data in the time domain. Their Fourier transforms are often referred to as *frequency windows*. On the other hand, the windows introduced in Chapters 3 and 4 are applied on the sample autocovariance (*lag windows*) and the sample spectrum (*spectral windows*)[1]. The effects of data windows on spectral estimates are briefly examined in this section, together with basic definitions and equations relating to linear and power spectral densities.

Time or data windows are used in the H.P. analyzer to compensate for the so-called *leakage effect*, which can best be understood by observing that for a purely periodic signal the truncated time trace $x_T(t)$ defined by

---

[1] Note that the definition of $x_T(t)$ in Chapter 4, Eq. (45), implies the use of a uniform data window.

Eq. (45), Chapter 4, is no longer periodic. The analyzer performs an FFT on the windowed time input to obtain the linear spectrum $a_T(f)$. For the uniform data window, the linear spectrum is defined by Eq. (46), Chapter 4. For the Hanning, flat top, force, and exponential windows, the linear spectrum is also defined by Eq. (46) but with the time signal multiplied by the selected window inserted in place of $x_T(t)$.

The H.P. dynamic signal analyzer computes the sample power spectrum from the relation

$$S_{xx}(f) = a_T(f) . a_T^*(f) \tag{96}$$

where $a_T^*(f)$ is the conjugate of $a_T(f)$. The sample power spectral density, denoted herein as PSD(f) for short, is obtained as the sample spectrum normalized with respect to a bandwidth of 1hz. According to the definitions in the dynamic analyzer manual

$$\text{PSD} = \frac{\text{sample power spectrum}}{\text{BW}} \tag{97}$$

where BW = bandwidth = $K_W(\Delta f)$, K_W is a window factor, and $\Delta f=(1/T)$. The window factor is necessary to correct for the effect that the data window has on the amplitude (and therefore the power) of the signal. For the uniform window, $K_W = 1$ and Eq. (97) is equivalent to Eq. (50), Ch. 4. For the Hanning, flat top, force and exponential windows, the factors K_W can be calculated numerically from the analytical expressions for the windows, as explained in Appendix 7.1. Such calculation was carried out as verification exercise for the Hanning and flat top windows and resulted in K_W values, respectively, of 1.5 and 3.82.

From Eqs. (96) and (97) we have

$$\frac{\text{PSD}(f)}{[a_T(f) . a_T^*(f)]} = \frac{T}{K_W} = \frac{\text{PSD}(f)}{|A(f)|^2} \tag{98}$$

where PSD(f) denotes the power spectral density obtained with the analyzer and $|A(f)|^2 = a_T(f).a_T^*(f)$. If the uniform window is chosen, the averaging of successive sample power spectra obtained using the H.P. analyzer averaging capability introduces the Bartlett spectral and lag windows, and the analyzer produces the same result as the methodology of Chapter 4, with the total record length T_r equal to the record length T times the number of averages.

7.2.1 Verification of Eq. (98). The verification was carried out using both sine waves of known frequency and amplitude and random noise as input signals. The sample linear spectrum $a_T(f)$ was obtained as the FFT of the windowed time input and the product $a_T(f).a_T^*(f) = |A(f)|^2$ was calculated at each frequency from this linear spectrum. Simultaneously, the sample power spectrum PSD(f) was obtained using the power spectral density function of the analyzer. Values of T/K_W were then calculated using Eq. (98). The procedure was carried out for the Hanning, flat-top and uniform windows, using the instrument's maximum display frequency of 102.4Khz, corresponding to T = 3.90625ms and Δf = 256hz. For the sine wave inputs, Table 7.1 gives the values of T/K_W and K_W calculated at the frequency f of the input sinusoidal signals. For the random noise inputs, the values of T/K_W and K_W in Table 7.2 were calculated for an arbitrarily chosen frequency of 10.24Khz.

The results in Tables 7.1 and 7.2 confirm the validity of Eqs. (97) and (98). The values K_W = 1.0, 1.5, and 3.82 are in perfect agreement with the values calculated from the analytical expressions for the windows. For the sine wave inputs, the results in Table 7.1 show that the power $|A(f)|^2 = a_T(f).a_T^*(f)$ is approximately constant independent of the window used, while the power spectral density PSD(f) decreases with increasing window bandwidth. The opposite occurs for the random noise inputs, as can be seen in Table 7.2. This different behavior is due to the fact that for the sine wave inputs the power is concentrated in a narrow band which, for the instrument settings used to construct Tables 7.1 and 7.2, is much smaller than the bandwidths $K_W(\Delta f)$ of the windows examined (ideally, the spectrum would be a Dirac delta function).

For additional verification, a different procedure was followed using a much smaller frequency resolution (see Appendix 7.2). An input sine wave with nominal frequency f_n = 1Khz was analyzed using f_{disp} = 3.2Khz, $\Delta f = f_{disp}/400$ = 8hz, $T = 1/\Delta f$ = 125ms. The marker (on the analyzer display options) was used every 8hz to get the values of the linear spectrum $a_T(f)$. These values were squared and added up (this sum is denoted by Σ^2 in Appendix 7.2) and then divided by the window bandwidth, $BW = K_W(\Delta f)$. As the results in the appendix show, the same number is obtained for the different windows. This number, multiplied by Δf, is the mean square voltage of the signal.

7.3. Comparison of the Dynamic Analyzer Spectrum with that Obtained Using the Spectrum Program

A comparison for a given signal of the sample spectrum given by the dynamic signal analyzer in the AC coupled mode (using a uniform data window) with the spectral estimate obtained using the spectrum program is shown in Fig. 7.1. Since the dynamic analyzer gives voltage spectra, the comparison is made with the voltage spectra obtained with the program. The analyzer uses M = 1024 points and the selected number of averages (the number of subseries) was K = 50. The spectrum program was used with M = 1024 and N = 32768 (K = 32). The spectral estimates are nearly identical except at low frequencies due to the high pass filter of the dynamic analyzer. Spectral values obtained with the analyzer in the AC coupled mode are not accurate below a frequency of 1hz due to the characteristics of the high pass filter. Spectral estimates at frequencies lower than 1hz (as low as 0.488mhz for single channel measurements) can be obtained using the analyzer in a DC coupled mode. Even in the DC coupled mode the spectral estimates at zero frequency and the first few frequency points are not accurate due to characteristics of the analyzer's input amplifiers. Hence, very long time records may be necessary if accurate estimates at the low frequencies are needed, with consequently small frequency spans.

Appendix 7.1

A data window modifies the input signal and hence a factor has to be applied on the resulting spectra to correct for the effect of the window. The factor can be calculated numerically from the analytical expression for the window. If input data $x(r\Delta t)$ are modified by a data window $d(r\Delta t)$ before calculation of the sample spectrum, the resulting spectral estimates must be divided by the factor

$$K_W = \frac{1}{M} \sum_{r=0}^{M-1} (d(r\Delta t))^2$$

to correct for the loss of amplitude caused by the data taper (Newland 1975, pp. 252–256). A straightforward FORTRAN program was written in order to compute this factor for the Hanning window defined by the relation

$$d(r\Delta t) = 2\left[\sin\left(\frac{\pi r \Delta t}{T}\right)\right]^2$$

The computation yielded a value of 1.5 which agrees very well with the results in Section 7.2.1. A similar routine for the flat-top window using the analytical expression for this window provided by Hewlett Packard resulted in a window factor of 3.82.

Appendix 7.2

A fixed sine signal of amplitude 5 volts and frequency 1Khz was used as the source signal for this verification procedure. The display window on the signal analyzer was set to a frequency span of 3.2Khz (f_{disp} = 3.2Khz) which resulted in a frequency resolution of Δf = 8hz. The values of the linear spectrum resulting from the application of each of the three windows, Hanning, uniform and flat top, were recorded every 8hz using the analyzer's marker function. The quantity $\Sigma^2/(K_W \Delta f)$ was calculated as shown below.

Uniform window (K_W =1):

$$\Sigma = [(5.0144)^2]^{\frac{1}{2}} = 5.0144 \text{ Volts}$$

$$\frac{\Sigma^2}{\Delta f . K_W} = \frac{(5.0144)^2}{(8.0)(1.0)} = 3.143.$$

Hanning window (K_W = 1.5):

$$\Sigma = [(2.508)^2+(5.01)^2+(2.509)^2]^{\frac{1}{2}} = (37.685)^{\frac{1}{2}} = 6.1388 \text{ Volts}$$

$$\frac{\Sigma^2}{\Delta f . K_W} = \frac{(6.1388)^2}{(8.0)(1.5)} = 3.140.$$

Flattop window (K_W = 3.82):

$$\Sigma = [(0.089)^2+(1.016)^2+(3.28)^2+(4.86)^2+(5.01)^2+ (4.862)^2+(3.282)^2+(1.017)^2+(0.089)^2]^{\frac{1}{2}} = (95.95)^{\frac{1}{2}} = 9.7955 \text{ Volts}$$

$$\frac{\Sigma^2}{\Delta f . K_W} = \frac{(9.7955)^2}{(8.0)(3.82)} = 3.140.$$

SIGNAL LEVEL AND FREQUENCY	WINDOW TYPE	A(f) (mvrms)	PSD(f) V^2_{rms}/hz (x 10^9)	$\|A(f)\|^2$ V^2_{rms} (x 10^6)	(T/K_W) (msec)	K_W
10 mv f = 10.24kHz	Uniform	7.0955	196.67	50.346	3.906	1.0
	Hanning	7.0880	130.87	50.251	2.604	1.5
	Flattop	7.0859	51.35	50.210	1.023	3.82
30 mv f = 10.24kHz	Uniform	20.073	1574.0	402.9	3.9	1.0
	Hanning	20.054	1047.3	402.2	2.6	1.5
	Flattop	20.049	411.11	402.0	1.023	3.82

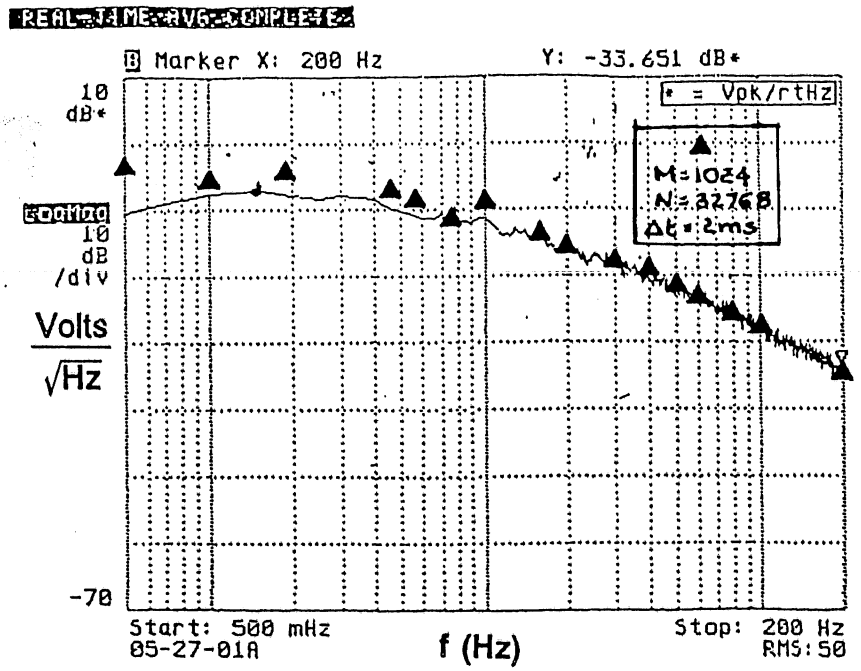
Table 7.1. Sine wave input, frequency = 10.24 kHz

H.P. Dynamic analyzer settings :
Δf = 256hz (measurements at full span= 102.4Khz), T=3.906ms, No. of averages = 50

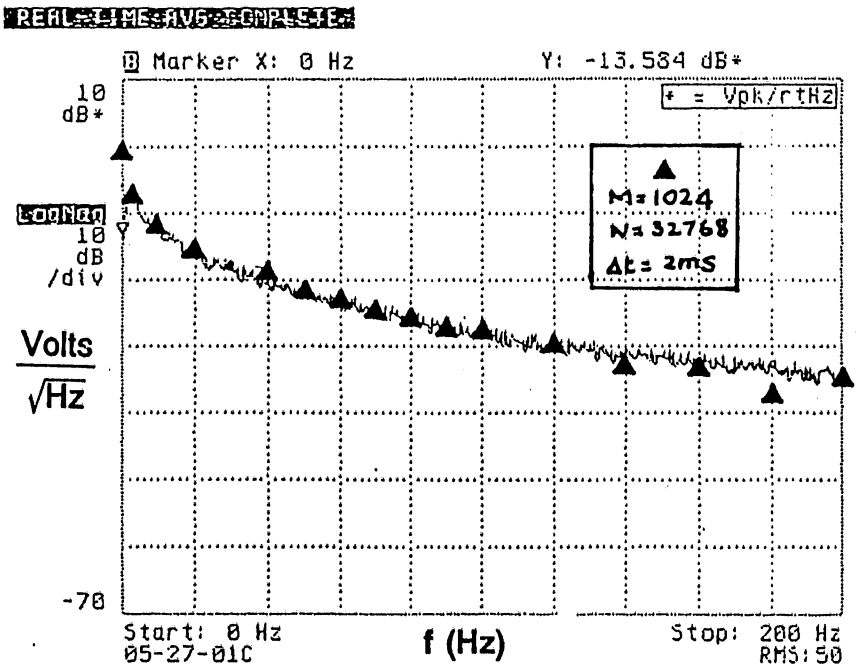
SIGNAL LEVEL (V)	WINDOW TYPE	A(f) (mvrms)	PSD(f) V^2_{rms}/hz (x 10^9)	$\|A(f)\|^2$ V^2_{rms} (x 10^3)	(T/K_W) (msec)	K_W
5	Uniform	89.407	31.23	7.9937	3.906	1.0
	Hanning	108.59	30.71	11.792	2.604	1.5
	Flattop	176.78	31.96	31.252	1.023	3.82
2	Uniform	83.49	27.23	6.9706	3.9063	1.0
	Hanning	107.63	30.17	11.583	2.604	1.5
	Flattop	185.66	35.25	34.468	1.023	3.82
0.5	Uniform	9.49	351.5E-03	89.97 E-03	3.906	1.0
	Hanning	11.68	355.3E-03	136.4 E-03	2.604	1.5
	Flattop	16.94	293.5E-03	287.0 E-03	1.022	3.82

Table 7.2. Results with a random noise as the input signal

H.P. Dynamic analyzer settings:
Δf = 256hz (measurements at full span= 102.4Khz), T=3.906ms; all measurements were done at a frequency of 10.24Khz (arbitrarily chosen), No. of averages = 50



a) Log-Log plot



b) Log-Linear plot

Fig.7.1 Comparison of H.P. dynamic analyzer spectrum with estimate from the program

APPENDIX A VELOCITY SPECTRA V/S VOLTAGE SPECTRA

The equation relating the bridge voltage output E of the constant temperature anemometer to the velocity V is

$$\frac{E^2}{(T_s-T_e)} = a\,(\rho V)^{\frac{1}{2}} + b \qquad (A.1)$$

where T_s= sensor operating temperature, T_e= fluid temperature, ρ = fluid density at T_e, and the constants a and b depend on the probe used. For the measurements in this report, a=0.08972 and b=0.085202. Fig. A.1 shows the variation of E with V and the regression curve used to get a and b. Due to the nonlinearity of the E–V relationship the velocity spectra are in general not the same as the voltage spectra. The slope dE/dV is largest at low velocities and can be obtained from Eq. (A.1) as

$$\frac{dE}{dV} = \frac{a^2 \rho\,(T_s-T_e)^2}{4E\,[E^2-b(T_s-T_e)]} \qquad (A.2)$$

In the present analysis a=0.08972, b=0.085202, ρ=1.180, T_e=299K, and T_s=523K. Substituting these values in Eq. (A.2) we get

$$\frac{dE}{dV} = \frac{119.15}{E\,[E^2-19.09]} \qquad (A.3)$$

For the measurements at height z=5cm the slope dE/dV varied from about 0.176 volts/(m/s) for V = 9.5 volts to 0.124 Volts/(m/s) for V = 10.5 volts.

Corresponding velocity and voltage sample spectra at z = 5cm are shown in Figs. A.2 and A.3. (The data analysis was carried out using N = 32768 points, M = 1024 points, and Δt=2ms.) A 10-point moving average of the difference between the two spectra is shown in Fig. A.4. As expected the voltage spectrum does not exactly match the velocity spectrum, particularly at the larger frequencies where the contributions from the nonlinearity of the E–V relationship are larger.

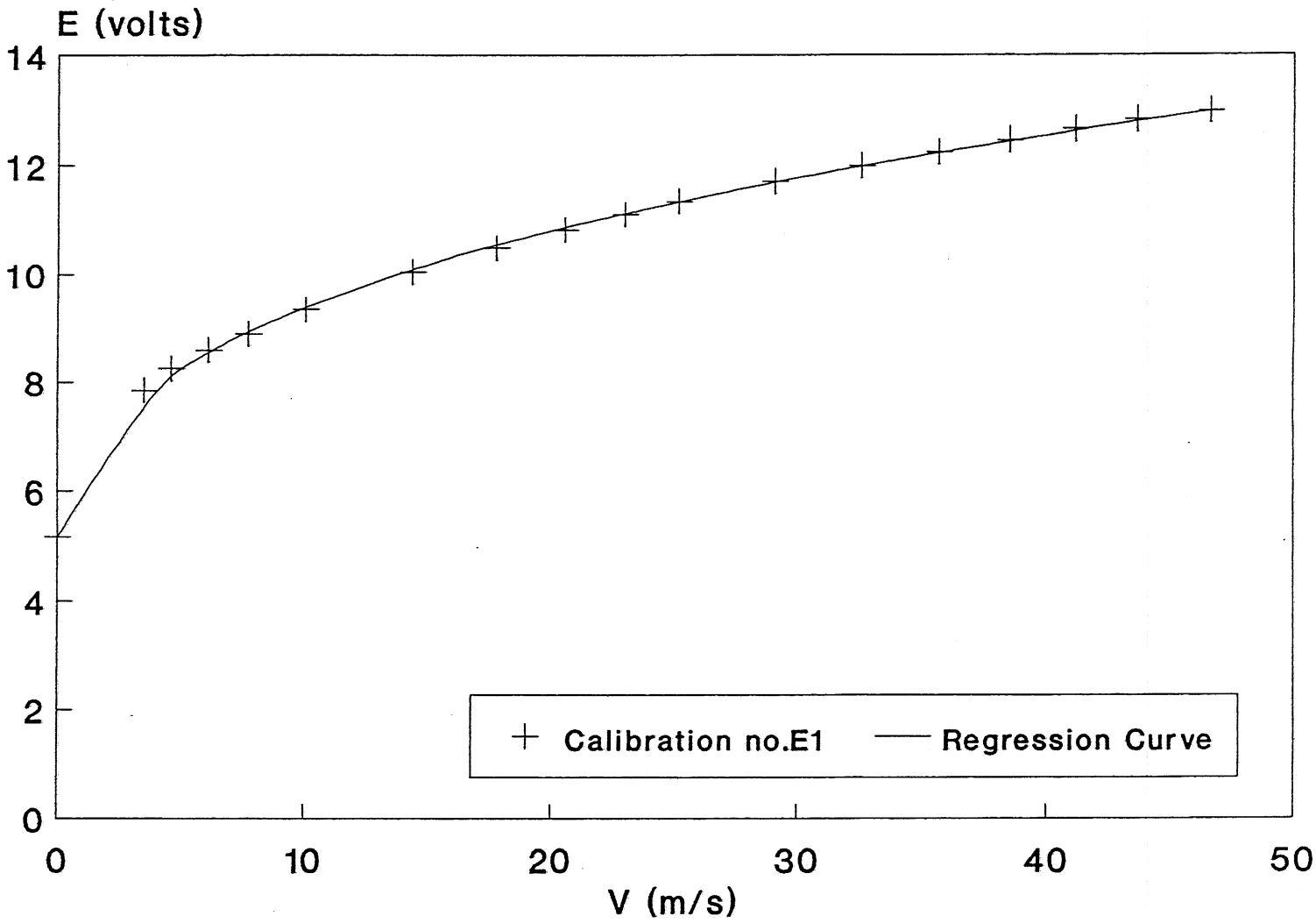


Fig. A.1 Calibration of the hot film probe
Probe: TSI 1210-60

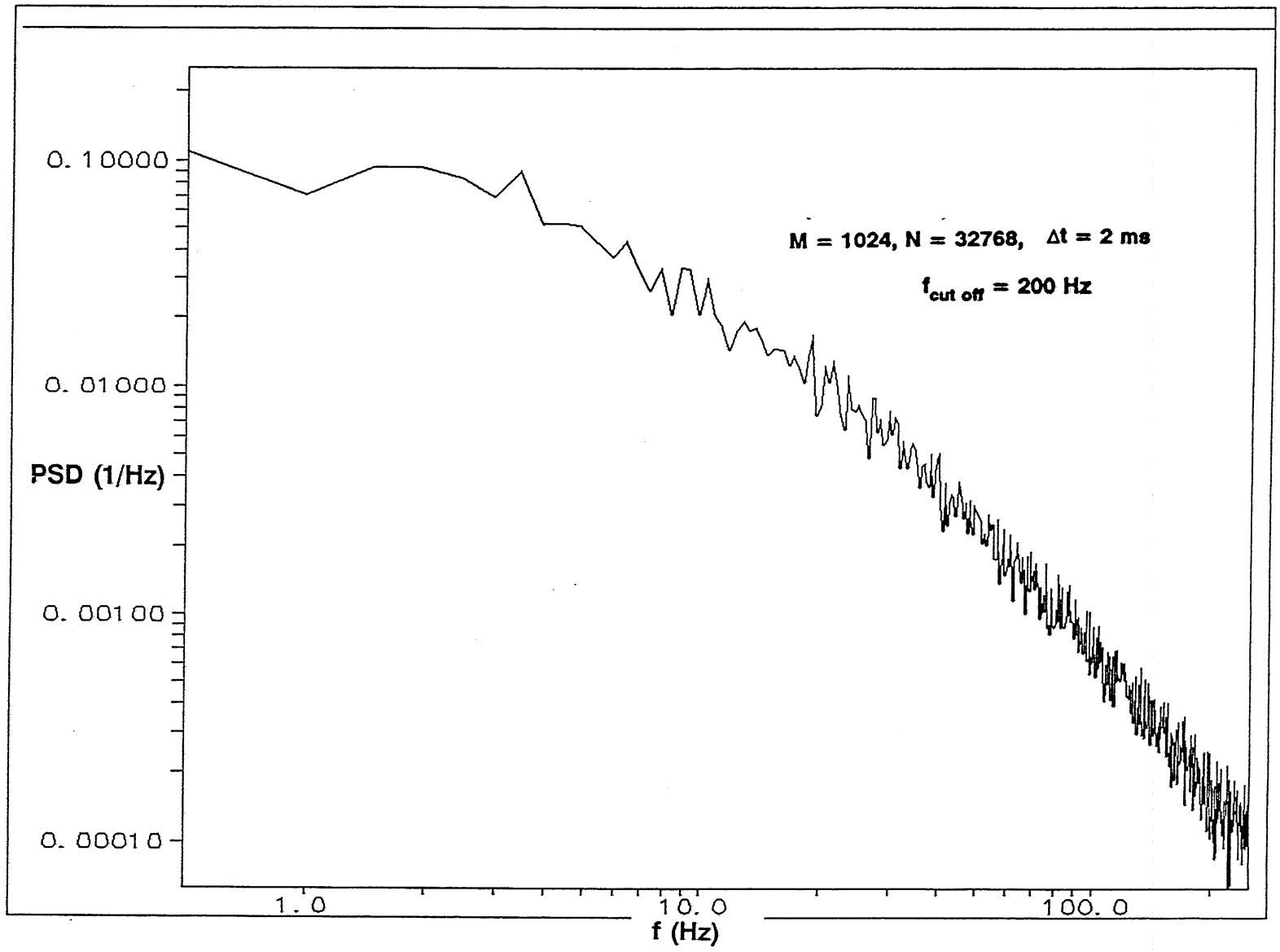


Fig. A.2 Power spectral density of velocity fluctuations

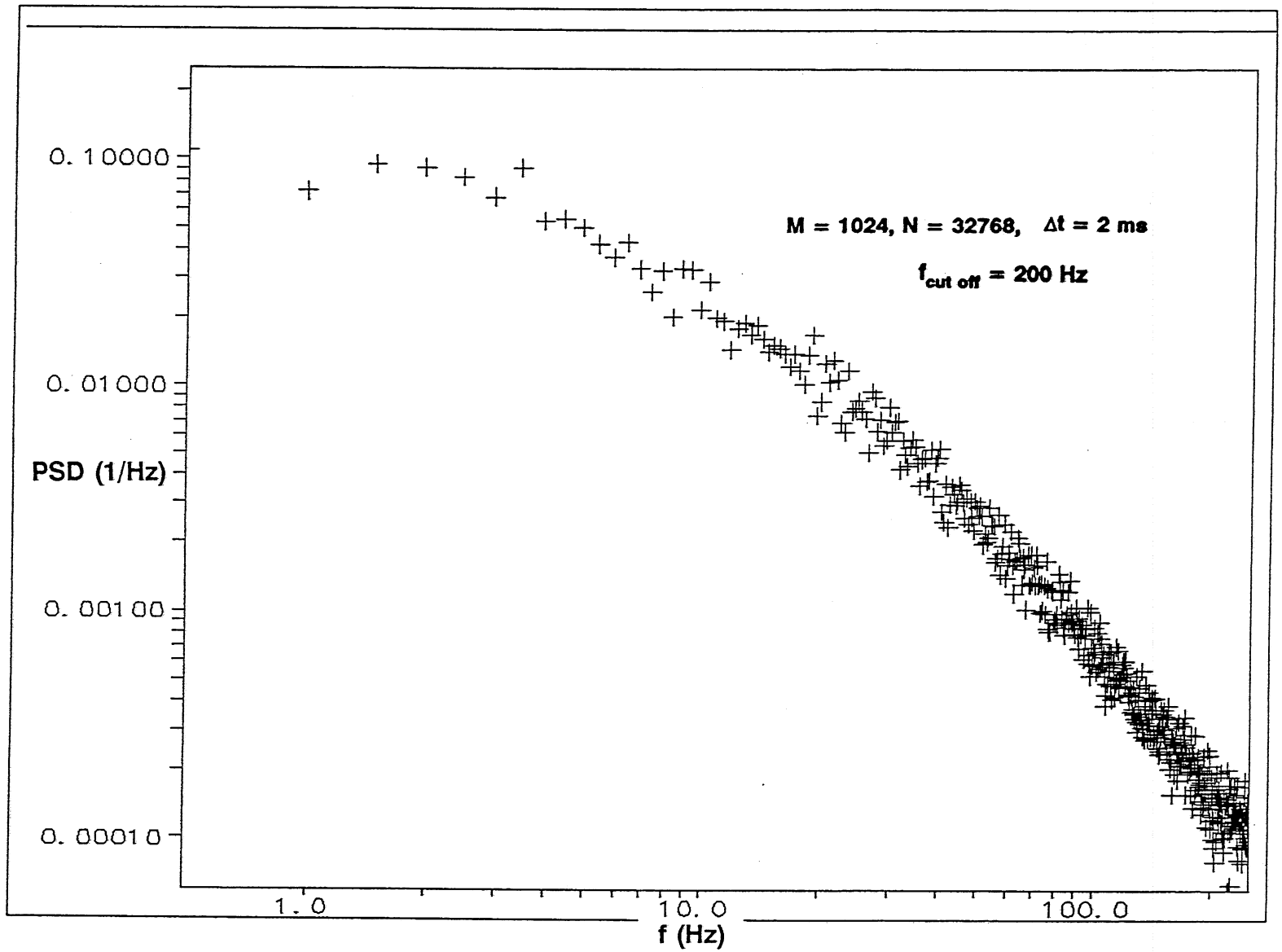


Fig. A.3 Power spectral density of voltage fluctuations

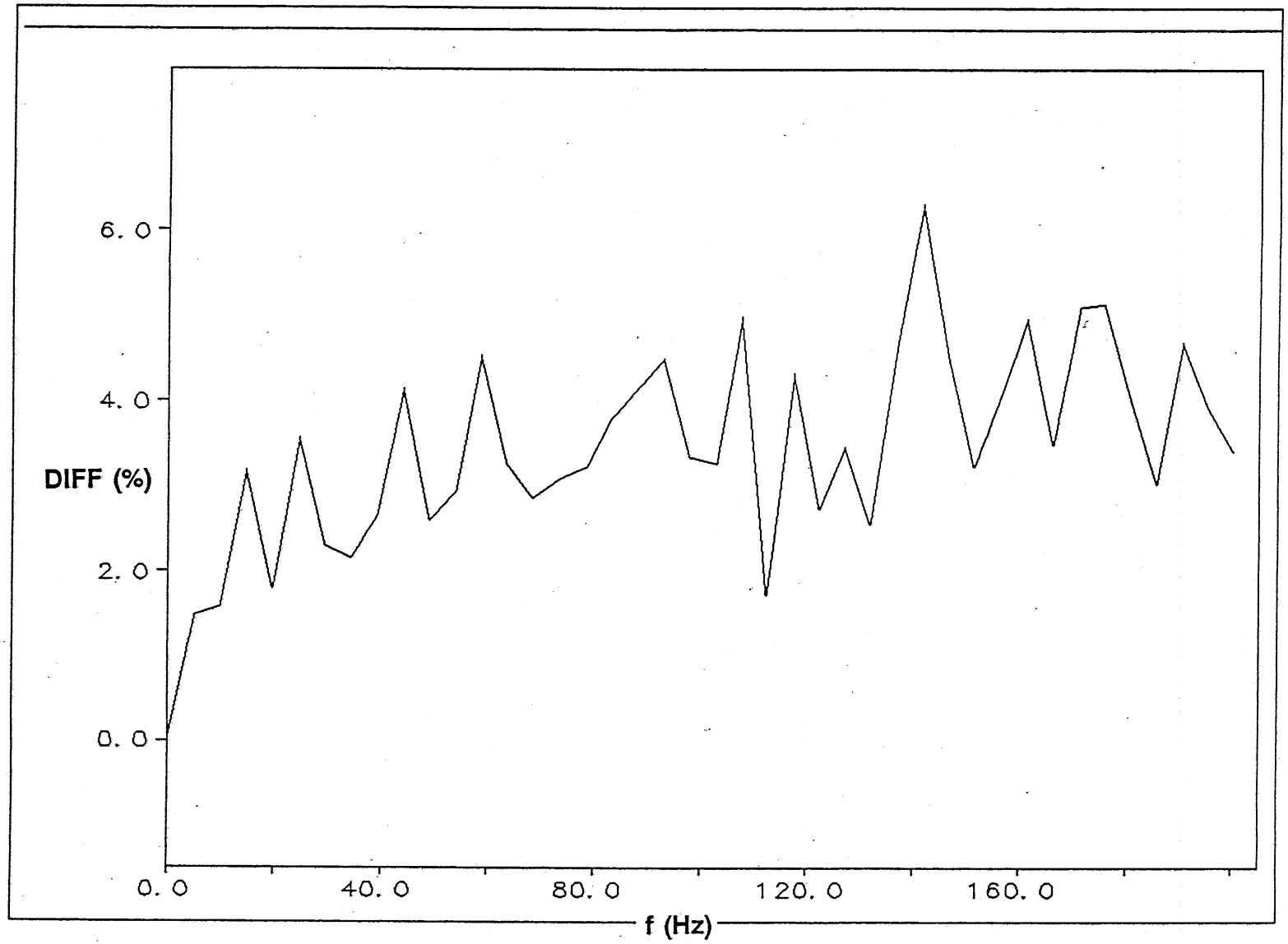


Fig. A.4 Ten-point moving average of the difference between the power spectral densities of voltage and velocity fluctuations

APPENDIX B SPECTRUM PROGRAM

```
C*************************************************************
C PROGRAM SPEVEL.FOR IS TO OBTAIN THE POWER SPECTRUM AND
C THE AUTOCOVARIANCE FUNCTION. THE ANALYSIS IS DONE
C BY ADDING THE SAME NUMBER OF ZEROS IN EACH SUBSERIES.
C CONFIDENCE INTERVAL FOR THE SPECTRUM IS CALCULATED
C USING THE APPROXIMATE METHOD. THE PROGRAM TAKES VOLTAGE DATA
C AS ITS INPUT. CONVERSION TO VELOCITY IS CARRIED OUT IN
C THE CIRDFT SUBROUTINE. THE CALIBRATION COEFFICIENTS HAVE
C TO BE CHANGED ACCORDING TO THE MOST RECENT CALIBRATION.
C*************************************************************
C
      CHARACTER*1 ALT
      CHARACTER*9 INP,OUT
C
      COMMON/CONST/ N,M,MLOG,K,TE
      COMMON/SMALL/ DT,DF
      COMMON/VARBL/ X(1:33000),SPE(0:4096),AUT(0:4096)
      COMMON/CONFI/ ALPHA,BETA,RALPHA,ALOGL,ALOGU,ALOWER,AUPPER
C
C
 199  WRITE(6,1)
    1 FORMAT(1x ,'ENTER INPUT FILENAME(< 9 CHARACTERS) ==>')
      READ(5,2) INP
    2 FORMAT(A9)
      OPEN (UNIT=8,FILE=INP,STATUS='OLD')
      WRITE(6,3)
    3 FORMAT(1x,'ENTER OUTPUT FILENAME(< 9 CHARACTERS) ==>')
      READ(5,2)OUT
      OPEN (UNIT=9,FILE=OUT,STATUS='unknown')
C
C INPUT QUANTITIES   (M * K = N)

      WRITE(6,4)
    4 FORMAT(1x,'ENTER SAMPLING RATE (DELTA T IN SECS) ==>')
      READ(5,5) DT
    5 FORMAT(F11.4)
      WRITE (6,6)
    6 FORMAT(1x,'ENTER TOTAL NO OF SAMPLES (N) ==>')
      READ(5,7)N
    7 FORMAT(I5)
      WRITE(6,8)
    8 FORMAT(1x,'ENTER NO SAMPLES IN SUBSERIES (M <= 2048) ==>')
      READ(5,9)M
    9 FORMAT(I4)
      WRITE(6,10)
   10 FORMAT(1x,'ENTER MLOG = LOG(M) ==>')
      READ(5,11)MLOG
   11 FORMAT(I2)
      K=N/M
C
C TE IS THE AVERAGE FLUID TEMPERATURE DURING DATA ACQUISITION
      WRITE(6,15)
   15 FORMAT(1x,'GIVE THE VALUE OF TEMP TE IN K ==>')
      READ(5,*)TE
C
C INPUT CONFIDENCE INTERVAL PARAMETER ALPHA
C
      ALPHA=0.05
C
C CALCULATE FMAX & DELTA F
C
      FMAX=1./2./DT
      DF=1./DT/FLOAT(M)
      WRITE(6,50) DF
   50 FORMAT(///,15X,'FREQUENCY RESOLUTION = ',F10.4,//)
C
C GET DATA
```

```
      DO 100 JH=1,N
      READ(8,*) X(JH)
  100 CONTINUE
C
C CALL SUBROUTINES THAT CALCULATE SPECTRUM, AUTOCOVARIANCE
C AND CONFIDENCE INTERVAL
C
      CALL CIRDFT
      CALL CONORM
C
C
      RALPHA=100*(1-ALPHA)
      WRITE(6,668) RALPHA
  668 FORMAT(/,F6.2,' PERCENT CONFIDENCE INTERVAL')
      WRITE(6,669) ALOWER,AUPPER
  669 FORMAT(/,2X,'LOWER =',E13.5,/,2X,'UPPER =',E13.5)
C
      WRITE(6,670)
  670 FORMAT(/,1x,'COMPUTATION FINISHED. DO YOU WISH ANOTHER RUN?(Y/N)')
      READ(5,671) ALT
  671 FORMAT(A1)
      IF(ALT .EQ. 'Y' .OR. ALT .EQ. 'y') GOTO 199
      STOP
      END
C
C
C
      SUBROUTINE CIRDFT
C*****************************************************************
C  THIS SUBROUTINE IS TO CALCULATE  SPECTRUM AND
C  THE AUTOCOVARIANCE FUNCTION BASED ON SECTION 4.4
C  IN CHAPTER 4
C*****************************************************************
      COMMON/CONST/ N,M,MLOG,K,TE
      COMMON/SMALL/ DT,DF
      COMMON/VARBL/ X(1:33000),SPE(0:4096),AUT(0:4096)
      COMMON/CONFI/ ALPHA,BETA,RALPHA,ALOGL,ALOGU,ALOWER,AUPPER
C     COMMON/STAT/ XVAR,XMEAN
      COMPLEX A,ACON,AUTC
      DIMENSION SUM(0:4096),A(4096),ACON(4096)
      DIMENSION AUTC(4096)
C
C
C  CALCULATION OF VELOCITY FROM VOLTAGE DATA ,X,
C  AND THE COMPUTATION OF VARIANCE, MEAN OF THE
C  OVERALL TIME RECORD
C
      XMEAN=0.0
      XVAR=0.0
      R=FLOAT(N)
C
      RHO=(1.192*294.6/TE)
      DO 45 JV=1,N
      E2=(X(JV)**2)/(523-TE)
      E2=((E2-0.085202)/0.08972)**2
   45 X(JV)=E2/RHO
C
      DO 50 JV=1,N
      XMEAN=XMEAN+(X(JV)/R)
   50 CONTINUE
      DO 51 JV=1,N
      XVAR=XVAR+(X(JV)-XMEAN)**2
   51 CONTINUE
      XVAR=XVAR/R
      WRITE(6,75)XVAR,XMEAN
   75 FORMAT(//,2x,'VARIANCE=',E13.5,5X,'MEAN VALUE=',E13.5)
C
C   DEFINE PARAMETERS
C
      MT2=M*2
```

```
      M2M1=MT2-1
      M2LOG=MLOG+1
C
      DO 100 IK=0,M2M1
  100 SUM(IK)=0.
C
C
      DO 300 II=1,K
      IM=II-1
C     REMOVAL OF THE OVERALL MEAN
C
      DO 240 IL=1,M
  240 A(IL)=X(IL+IM*M)-XMEAN
C     APPENDING M ZEROES TO THE SUBSERIES. (TO GET THE
C     AUTOCOVARIANCE DIRECTLY)
      DO 250 IQ=M+1,MT2
  250 A(IQ)=0.
C
C     CALCULATION OF VARIANCE `VARC',AND MEAN UARM
C     FOR EACH SUBSERIES
      VARC=0.0
      UARM=0.0
      DO 255 I=1,M
      UARM=UARM+(X(I+IM*M)/FLOAT(M))
  255 CONTINUE
      DO 257 I=1,M
      VARC=VARC+((X(I+IM*M)-XMEAN)**2)
  257 CONTINUE
      VARC=VARC/FLOAT(M)
C
C     CALL THE FAST FOURIER TRANSFORM ROUTINE
      CALL FFT(A,M2LOG,MT2)
C
      DO 260 IS=1,MT2
      ACON(IS)=CONJG(A(IS))
C
C  FINDING THE SPECTRUM BASED ON Eq. (72), Ch. 4
C  THEN, CHANGING SUBSCRIPT BACK TO THE ACTUAL CASE(START FROM ZERO)
C  ALSO SPECTRUM IS NORMALIZED BY THE VARIANCE OF THAT SUBSERIES
C
      SPE(IS-1)=(FLOAT(M)*DT*A(IS)*ACON(IS)*4.)/VARC
  260 SUM(IS-1)=SUM(IS-1)+SPE(IS-1)
      WRITE(6,*)VARC,UARM
  300 CONTINUE
C
C  SMOOTHING BY AVERAGING
C
      DO 350 IA=0,M2M1
  350 SPE(IA)=SUM(IA)/FLOAT(K)
C
C  CALCULATION OF INTEGRAL LENGTH SCALE FROM THE VALUE OF SPECTRUM
C  AT ZERO FREQUENCY
C
      XLUU=(XMEAN*SPE(0))/2.
C
C
C  OBTAIN NORMAZLIZED CIRCULAR AUTOCOVARIANCE FUNCTION BY
C  INVERSE FOURIER TRANSFORM (Section 4.4, Ch. 4)
C
      DO 360 IG=1,MT2
  360 AUTC(IG)=SPE(IG-1)/DT/FLOAT(MT2)/2.
C
      CALL FFT(AUTC,M2LOG,MT2)
C
      DO 370 IH=0,M2M1
  370 AUT(IH)=AUTC(IH+1)
C
C  OBTAIN THE AUTOCOVARIANCE FUNCTION
C
      DO 380 IZ=0,M2M1
```

```
  380 AUT(IZ)=AUT(IZ)*2.*FLOAT(MT2)
C
C-----------------------------------------------------------------------
C    WRITE THE SPECTRUM AND AUTOCOVARIANCE TO THE OUTPUT FILE IN A FORMAT
C    THAT CAN BE READ EASILY BY ANY GRAPHICS PACKAGE.
C-----------------------------------------------------------------------
      WRITE(9,501)XLUU
  501 FORMAT(2X,'THE INTEGRAL LENGTH SCALE OF THE PROCESS IS',5X,F7.3)
      DO 503 I=0,M-1,2
      FREQ=FLOAT(I)*DF/2
      TIMS=FLOAT(I)*DT/2
      SPE(I)=2*SPE(I)
      WRITE(9,502)FREQ,SPE(I),TIMS,AUT(I)
  502 FORMAT(F10.2,E13.5,E13.4,E13.5)
  503 CONTINUE
      WRITE(6,504)XLUU
  504 FORMAT(/,2X,'INTEGRAL LENGTH SCALE =',3X,F7.3)
C
      RETURN
      END
C*************************************************************
C  SUBROUTINE TO CALCULATE FFT
C*************************************************************
      SUBROUTINE FFT(A,MLOG,NB)
      COMPLEX A(NB),U,W,T
C DIVIDE ALL ELEMENTS BY NB
      DO 10 J1=1,NB
   10 A(J1)=A(J1)/FLOAT(NB)
C REORDER SEQUENCE ACCORDING TO FIG. 12.8
      NBD2=NB/2
      NBM1=NB-1
      J=1
      DO 40 L=1,NBM1
      IF(L .GE. J) GOTO 20
      T=A(J)
      A(J)=A(L)
      A(L)=T
   20 K=NBD2
   30 IF(K .GE. J) GOTO 40
      J=J-K
      K=K/2
      GOTO 30
   40 J=J+K

C CALCULATE FFT

      PI=4.0*ATAN(1.0)
      DO 60 M1=1,MLOG
      U=(1.0,0.0)
      ME=2**M1
      K=ME/2
      AK=FLOAT(K)
      W=CMPLX(COS(PI/AK),-SIN(PI/AK))
      DO 60 J2=1,K
      DO 50 L2=J2,NB,ME
      LPK=L2+K
      T=A(LPK)*U
      A(LPK)=A(L2)-T
   50 A(L2)=A(L2)+T
   60 U=U*W
      RETURN
      END
C
C
      SUBROUTINE CONORM
C*************************************************************
C THIS SUBROUTINE IS TO OBTAIN THE CONFIDENCE INTERVAL
C OF THE SAMPLE SPECTRUM ESTIMATOR ASSUMING THAT THE
C ESTIMATOR IS AS A CHI-SQUARE DISTRIBUTION. IT IS
C DONE BY USING A NORMAL DISTRIBUTION APPROXIMATION AS
```

```
C THE DEGREE OF FREEDOM IS LARGE (GREATER THAN 30).
C*******************************************************
      COMMON/CONST/ N,M,MLOG,K,TE
      COMMON/SMALL/ DT,DF
      COMMON/VARBL/ X(1:33000),SPE(0:4096),AUT(0:4096)
      COMMON/CONFI/ ALPHA,BETA,RALPHA,ALOGL,ALOGU,ALOWER,AUPPER
      DOUBLE PRECISION FNORM,ASQ,DNORM,ANORM
C
      BETA=(1.-ALPHA)/2.
C
C  CALCULATE THE DEGREE OF FREEDOM `NU'
C
      NU=3*K
      ANU=FLOAT(NU)
C
      PAR=2./9./ANU
C
      FNORM=0.
      DNORM=1.D-2
      ANORM=-DNORM/2.
      PI=4.0*ATAN(1.0)
      DEN=SQRT(2.*PI)
C
   50 ANORM=ANORM+DNORM
      ASQ=ANORM**2
      VAR=EXP(ASQ/(-2.))
      FNORM=FNORM+VAR/DEN*DNORM
C
      DUM3=ABS(FNORM-BETA)
      IF(DUM3 .LT. 1E-3) GOTO 60
      IF(ANORM .GT. 4.) RETURN
      GOTO 50
C
C PRINT OUT LOWER AND UPPER XP IN NORMAL DISTRIBUTION
C
   60 XP=ANORM
      XM=-XP
C     WRITE(9,70) XM,XP
C  70 FORMAT(1H1,////,1X,'NORMAL DISTRIBUTION',//,2X,'XLOWER = ',E13.5,
C    +          //,2X,'XUPPER = ',E13.5)
C
C GET AND PRINT OUT LOWER AND UPPER XP IN CHI-SQUARE DISTRIBUTION
C
      XPNEW=ANU*(1.-PAR+XP*SQRT(PAR))**3
      XMNEW=ANU*(1.-PAR+XM*SQRT(PAR))**3
      ALOWER=ANU/XPNEW
      AUPPER=ANU/XMNEW
      RALPHA=100.*(1.-ALPHA)
      ALOGL=ALOG10(ALOWER)
      ALOGU=ALOG10(AUPPER)

      RETURN
      END
C=======================================================================
```

REFERENCES

Bartlett, M.S., "On the Theoretical Specification and Sampling Properties of Autocorrelated Time Series", Journal of the Royal Statistical Society, London, Ser. B., Vol. 8, pp. 27-41, 1946.

Bartlett, M.S., "Periodogram Analysis and Continuous Spectra," Biometrika, Vol. 37, Parts 1 and 2, pp. 1 -16, 1950.

Bochner, S., "Lectures on Fourier Analysis," Edwards Brothers, Ann Arbor, Michigan, 1937.

Cooley, J.W., Lewis, P.A.W., and Welch, P.D., "Application of the Fast Fourier Transform to Computation of Fourier Integrals, Fourier Series, and Convolution Integrals," IEEE transactions on Audio and Electracoustics, Vol. AU-15, pp. 79-84, 1967.

Enochson, L.D., and Otnes, R.K., "Programming and Analysis for Digital Time Series Data," The Shock and Vibration Information Center, Naval Research Laboratory, Washington, D.C., Vol. SVM-3, 1968.

Farell, C., and Iyengar, A. K., "Experiments on the Simulation of a Neutral Atmospheric Boundary Layer," 7th U.S. Natl. Conf. Wind Eng., Los Angeles, 1993.

Hinze, J.O., "Turbulence," Mcgraw-Hill, 1959.

Jenkins, G.M., and Watts, D.G., "Spectral Analysis and Applications," Holden-Day, San Francisco, 1968.

Newland, D.E., "An Introduction to Random Vibrations and Spectral Analysis," Longman Group Limited, London, 1975.

Parzen, E., "Stochastic Processes," Holden-Day, 1962.

Parzen, E., "On Statistical Spectral Analysis," Proceedings of Symposia in Applied Mathematics, American Mathematical Society, Vol. 16, 1964.

Rosenblatt, M., "Random Processes," Oxford University Press, 1962.

Schaerf, M.C., "Estimation of the Covariance and Autoregressive Structure of a Stationary Time Series," Ph.D. Thesis, Stanford University, 1963.

Tennekes, H., and Lumley, J.L., "A First Course in Turbulence," M.I.T. Press, 1972.

Teunissen, H.W., "Structure of Mean Winds and Turbulence in the Planetary Boundary Layer over Rural Terrain," Boundary-Layer Meteorology, 19, pp. 187-221, 1980.

Welch, P.D., "The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging over Short, Modified Periodograms," IEEE transactions on Audio and Electroacoustics, Vol. AU-15, pp. 70-73, 1967.