

# Nonparametric Estimation of Mean Doppler and Spectral Width

José M. B. Dias, *Member, IEEE*, and José M. N. Leitão, *Member, IEEE*

**Abstract**—This paper proposes a new nonparametric method for estimation of spectral moments of a zero-mean Gaussian process immersed in additive white Gaussian noise. Although the technique is valid for any order moment, particular attention is given to the mean Doppler (first moment) and to the spectral width (square root of the centered second-spectral moment). By assuming that the power spectral density (PSD) of the underlying process is bandlimited, the maximum-likelihood estimates of its spectral moments are derived. A suboptimal estimate based on the sample covariance is also studied. Both methods are robust in the sense that they do not rely on any assumption concerning the PSD (besides being bandlimited). Under weak conditions, the set of estimates based on sample covariance is unbiased and strongly consistent. Compared with the classical pulse pair and the periodogram-based estimators, the proposed methods exhibit better statistical properties for asymmetric spectra and/or spectra with large spectral widths, while involving a computational burden of the same order.

**Index Terms**—Bandlimited processes, maximum-likelihood, nonparametric estimation, spectral moments.

## I. INTRODUCTION

THE GOAL of spectral estimation is to infer the power spectral density (PSD) from a finite observation of the underlying process. This subject has been extensively studied, yielding a large set of techniques, each suited for a particular situation. However, in many applications, the objective is the determination of PSD functionals rather than the PSD itself. This is the case of the spectral moments (SM's), particularly the mean power, the mean velocity, and the spectral width. Application areas are weather radar [1]–[8], ultrasound imaging in medicine [9]–[13], clear-air turbulence measurement [14], [15], synthetic aperture radar (SAR) [16]–[19], and electroencephalographic analysis [20]–[22] to name a few.

Concerning Doppler weather radars, the goal is the determination of the first three spectral moments. These moments are closely related to physical properties of the backscatters within the so-called resolution volume [5], [23]: the mean power (zeroth moment) is related to the reflectivity (which, in the case of rain, is an indicator of the water content), the mean frequency (first moment) is related to the mean radial velocity, and the spectral width (square root of the second-centered moment) is a measure of the velocity dispersion.

Ultrasound imaging is used in medicine for determination of soft-tissue acoustic parameters [24]. Of particular interest in the case of blood tissue are the three first-spectral moments [12], [25]. The mean power is related to the scatterers' (cells and other structural details) density, the mean frequency is related to the mean radial velocity, and the spectral width is a measure of the blood-flow turbulence.

In the atmosphere, turbulence can be thought of as random motion of a fluid (clear-air turbulence), thus causing variations in the refractive index [23]. These variations produce radar echoes that convey important information concerning atmospheric winds, turbulence, and other meteorological phenomena [14], [15]. Estimating the first three spectral moments from clear-air echoes finds important applications in airport hazard monitoring (for wind shear and wingtip vortices), detection of severe storms, and transport phenomena of various atmospheric pollutants.

Regarding SAR applications, the determination of the first spectral moment, the so-called Doppler centroid, is required for operations such as range-cell migration correction, azimuth compression, and image registration [19].

Very often, SM's must be estimated from a small number of discrete samples. Normally, underlying this restriction are either fast scanning rates or short-time observation windows. For example, in a typical weather radar, the number of estimates per complete aerial revolution can be as large as  $3 \times 360\,000$  [5], corresponding to 3 SM's, with a spatial resolution of  $100\text{ m} \times 1^\circ$ . This has to be done in real or near-real time, which means a few tens of seconds. Furthermore, in order to have, simultaneously, an acceptable azimuthal velocity (a few revolutions per minute) and to prevent the broadening of the lateral antenna pattern, the number of samples per estimate should not be large. Typical values are in the set {16, 32, 64, 128} (depending on the azimuthal velocity and on the lateral aerial resolution). A similar scenario can also be found in ultrasound imaging and clear-air turbulence measurement.

In summary, in many applications, SM's estimators should meet the requirement of being low in complexity (in a computational sense) yielding estimates of acceptable quality based on small sample sizes.

### A. Classical Estimators

As a statistical inference problem, the approach to SM estimation can be classified as parametric or nonparametric.

Nonparametric techniques are based either on the PSD or on the covariance function (CF) [26]. Accordingly, they are classified as spectral or covariance approaches, respectively. Seminal

Manuscript received August 13, 1997; revised January 11, 1999. This work was supported by the Portuguese PRAXIS XXI Program under Project 2/2.1.TIT/1580/95.

The authors are with the Instituto de Telecomunicações and Departamento de Engenharia, Electrotécnica e de Computadores, Instituto Superior Técnico, 1049-001 Lisboa, Portugal (e-mail: bioucas@lx.it.pt).

Publisher Item Identifier S 0196-2892(00)00016-4.

works [27]–[29] addressed a large set of combinations of spectral and covariance approaches, including continuous and discrete samples, real and complex processes, and white and colored noise.

In [2], various estimators of mean Doppler frequency and spectrum width are investigated. A technique therein called spectral processing uses the fast Fourier transform (FFT) as an efficient way, in a computational sense, to determine a spectral estimate. This method is, in essence, the one proposed in [28], in which the periodogram is replaced with a discrete version computed by means of the FFT. Since this estimator is based on the periodogram, it will be herein called the periodogram-based (PB) estimator. Related references on this topic are [30] and [31].

In the particular case of mean frequency, there exists a considerable number of methods besides those mentioned in the above paragraph. Namely, the Poly-Pulse-Pair family [32], which is based on the methodology introduced in [29], the periodogram maximization [4], the vector phase change [33], the scalar phase change [34], and the time derivative [35].

Despite the wide variety of available SM estimators, the pulse pair (PP) method introduced in [29], and the PB method, documented in [2], are the most widely used. This is a consequence of a good tradeoff between complexity (measured in number of floating-point operations) and quality of the estimates. Given a sample of size  $N$ , the complexities of the PP and PB algorithms are of the order of  $N$  and  $N \log N$ , respectively. The PP method has the additional advantage of allowing different sampling intervals, thus providing data to unfold the mean frequency when aliasing is present. On the other hand, the PB procedure allows us to edit the periodogram, which is important in some cases. Application and additional research of these techniques to the weather radar is carried out in [2], [5], and [36].

The PP and the PB statistical properties are well documented (see e.g., [2], [3], [29], [30], [36], [37]). Namely, it has been found that

1) *PP Method:*

- 1) The mean-frequency estimate is increasingly biased with the spectral skewness.
- 2) The spectral-width estimate is biased (its bias increases with the spectral width<sup>1</sup>).
- 3) The variances of both estimates are far from their Cramér–Rao bounds (CRB) for all ranges of spectral widths.

2) *PB Method:*

- 1) The mean-frequency estimator is biased due to the finite resolution associated with the FFT (this is a serious problem whenever the sample dimension is small).
- 2) The spectral-width estimator is biased due to the windowing effect associated with the FFT. This worsens as the spectral width increases.
- 3) The variance of both estimates at low spectral widths is close to the PP ones. However, at high spectral widths, the performance of the PB method is better.

<sup>1</sup> Work [36] proposes a modification of the spectral-width estimator introduced in [29], which is unbiased for Gaussian-shaped spectra.

Despite the reported shortcomings, the PP and the PB estimators are extensively used. As was mentioned previously, this results from the good tradeoff between complexity and performance.

Both PP and PB estimators are not optimal<sup>2</sup> in the sense that they were not derived from any optimality criterion, and they are not able to achieve uniformly optimal statistical properties. Concerning this matter, the maximum-likelihood (ML) criterion plays a prominent role given its optimal properties, at least in an asymptotic sense. In the field of ML spectral-moments estimation, various approaches have been proposed [2], [38]–[40]. Besides ML, maximum entropy [41] and risk-based [7] criteria have been suggested.

All the aforementioned criteria require parametrized PSD models. By constraining the search space, parametric approaches deliver estimates with lower uncertainty compared to nonparametric procedures. On the other hand, they generally have higher complexity than nonparametric methods.

Another important issue, concerning the parametric approach, is the choice of a model able to accurately fit the true PSD. In weather radar, the Gaussian spectral shape has been extensively used. Although having some experimental justification, the Gaussian-shape assumption is not without weaknesses. This is shown in [42], where a systematic and exhaustive measurement of spectral shape from precipitation echoes is reported. We quote from [42]: “. . . a Gaussian spectral shape agrees reasonably with a large fraction of the obtained spectra. However, in about a quarter of the cases, the deviation from the Gaussian shape is considerable, e.g., one or both edges may be too steep or too slight, the peak may be off-center, there may be more than one peak.” Thus, a more accurate spectral fitting should be looked for. This can be achieved, for example, with autoregressive moving average (ARMA) models of adequate dimension, of which the work reported in [15] is an example. However, the complexity inherent in the ARMA parameters estimation is unbearable in most applications.

As a conclusion to the above considerations, the classical nonparametric PP and PB estimators are characterized by having low complexity and tolerable, sometimes poor, performance. On the other hand, the parametric estimators rely strongly on spectral shape assumptions. In this sense, they are not robust. Moreover, their complexity is frequently incompatible with practical applications.

### B. Rationale of the Proposed Approach

The method herein presented assumes that the PSD of the underlying process is bandlimited. This hypothesis is meaningful, for example, in weather radar, ultrasound imaging, clear-air, and synthetic aperture radar (SAR) applications. In fact, in these cases, the PSD associated with each resolution volume is a weighted replica of the scatterers velocity distribution in the same volume [43], [44]. Since in each resolution volume the scatterers have a maximum and a minimum velocity, the PSD is bandlimited. The sampling theorem assures, therefore, that the CF can be exactly recovered from its discrete samples, as

<sup>2</sup> For Gaussian processes, if the samples can be split into independent pairs, which is an unrealistic assumption in most cases, the PP estimator is of maximum likelihood [29].

long as the sampling rate is equal to or greater than the Nyquist frequency. On the other hand, the SM's relate easily to the CF derivatives. In the sequel, the name bandlimited (BL) will be used to designate the proposed methodology.

The simple strategy we are proposing seems like it has not been fully explored in the literature. To our knowledge, only the work described in [10], in the field of ultrasound blood-velocity measurements, explores the bandlimited concept. However, their approach and methodology diverge from ours.

The paper organization is as follows. Section II introduces notation, formulates the SM estimation problem, and derives the ML estimates under the BL assumption. In Section III, the interpolation filter design is considered. A suboptimal estimator based on the sample covariance at different lags and its statistical characterization is presented in Section IV. Experimental results and comparisons with the PP and PB estimators are reported in Section V.

## II. PROBLEM FORMULATION

The BL method computes SM estimates from discrete samples. It is assumed, however, that there is an underlying process defined on  $\mathfrak{R}$ , from which discrete samples are taken. This is, in fact, a typical situation in all the applications referred to in the introduction.

Let  $\mathcal{X} = \{x(t), t \in \mathfrak{R}\}$  and  $\mathcal{N} = \{n(t), t \in \mathfrak{R}\}$  be independent, stationary, zero-mean, normal, complex processes, with covariance functions  $R_x(\tau) \equiv E[x(t+\tau)x^*(t)]$  and  $R_n(\tau) \equiv E[n(t+\tau)n^*(t)]$ , respectively. Also, define  $y(t) \equiv x(t) + n(t)$ .

Wide-sense stationarity and strict-sense stationarity are equivalent in real normal processes. This is not the case with complex processes. A zero-mean, wide-sense, stationary complex, normal process  $\mathcal{X}$  is strict-sense stationary if and only if  $E[x(t+\tau)x(t)]$  is a function only of  $\tau$  [26]. We further assume that the processes  $\mathcal{X}$  and  $\mathcal{N}$  are circular, which means that  $E[x(t+\tau)x(t)] = 0$  and  $E[n(t+\tau)n(t)] = 0$  for all  $t, \tau \in \mathfrak{R}$  [45].

It should be stressed that the circular and normal assumptions are, to a great extent, physically meaningful in many applications such as weather radar, ultrasound imaging, clear-air turbulence measurement, and SAR [46]–[51]. Nevertheless, the Gaussian hypothesis could be discharged, since it plays no role in the BL approach (except for its statistical characterization).

Define  $S_x(f)$  and  $S_n(f)$  as the PSD's of processes  $\mathcal{X}$  and  $\mathcal{N}$ , respectively. The SM estimation problem is stated as follows. Given the  $N$ -dimensional sample vector  $\mathbf{Y} \equiv [Y_1, \dots, Y_N]^T$  with  $Y_i \equiv y(iT_s)$ , find estimators for the entities below (assumed to exist).

- 1) the  $k$ th spectral moment, defined as

$$m_k(S_x) \equiv \int_{-\infty}^{\infty} f^k S_x(f) df \quad k = 0, 1, \dots \quad (1)$$

- 2) the normalized  $k$ th spectral moment defined as

$$\mu_k(S_x) \equiv \frac{\int_{-\infty}^{\infty} f^k S_x(f) df}{\int_{-\infty}^{\infty} S_x(f) df} \quad k = 1, 2, \dots \quad (2)$$

- 3) the spectral width defined as

$$\sigma(S_x) \equiv [\mu_2(S_x) - \mu_1^2(S_x)]^{1/2}. \quad (3)$$

Estimators of  $m_k(S_x)$ ,  $\mu_k(S_x)$ , and  $\sigma(S_x)$  will be denoted by  $\hat{m}_k(\mathbf{Y})$ ,  $\hat{\mu}_k(\mathbf{Y})$ , and  $\hat{\sigma}(\mathbf{Y})$ , respectively.

It will be assumed that the noise CF  $R_n(\tau)$  is known. This is not a severe restriction since, in most applications, this function can be estimated with arbitrary precision in the signal absence.

### A. Bandlimited Solution

Assume that process  $\mathcal{X}$  has absolutely continuous spectral distribution function  $W_x$ . Then  $S_x = W'_x$ , where  $(\cdot)'$  denotes the derivative. Further assume that  $S_x$  is of energy type and bandlimited, i.e.,

$$\begin{aligned} \mathbf{C}_1 & \int_{-\infty}^{\infty} S_x^2(f) df < \infty \\ \mathbf{C}_2 & S_x(f) = 0, f \notin I_S \equiv [-f_m, f_m] \end{aligned}$$

where

$$f_m \leq f_s/2 \quad \text{with} \quad f_s = 1/T_s.$$

Assumptions  $\mathbf{C}_1$  and  $\mathbf{C}_2$  and the Schwarz' inequality assure that for any  $k = 0, 1, \dots$

$$\left| \int_{-\infty}^{\infty} f^k S_x(f) df \right|^2 \leq \int_{-f_m}^{f_m} |f|^{2k} df \int_{-\infty}^{\infty} S_x^2(f) df < \infty \quad (4)$$

that is, the SM's of  $S_x$  do exist for any  $k = 0, 1, \dots$ .

Inequality (4) for  $k = 0$  implies that  $S_x$  is absolutely integrable. This condition assures that

$$R_x(\tau) = \int_{-\infty}^{\infty} S_x(f) e^{j2\pi f\tau} df \quad (5)$$

exists and is continuous and nonnegative definite on  $\mathfrak{R}$  (Bochner's Theorem [26, Theorem 5.1]). The above properties of  $R_x(\tau)$ , together with inequality (4), assure that  $R_x^{(k)}(\tau)$  [the  $k$ th derivative of  $R_x(\tau)$ ] with  $k = 0, 1, \dots$ , has the following properties [26, Theorem 5.2]:

$$\begin{aligned} \mathbf{P}_1 & R_x^{(k)}(\tau) \text{ exists for } \tau \in \mathfrak{R} \\ \mathbf{P}_2 & R_x^{(k)}(\tau) = (2\pi j)^k \int_{-\infty}^{\infty} f^k S_x(f) e^{j2\pi f\tau} df \\ \mathbf{P}_3 & R_x^{(k)}(\tau) \text{ is uniformly continuous on } \mathfrak{R}. \end{aligned}$$

Property  $\mathbf{P}_2$  allows us to write the SM's as

$$m_k(S_x) = \frac{R_x^{(k)}(0)}{(2\pi j)^k}, \quad k = 0, 1, \dots \quad (6)$$

Given the assumptions  $\mathbf{C}_1$  and  $\mathbf{C}_2$ , a slight modification of the Witteraker–Shannon–Kotel'nikov sampling theorem (see e.g., [52]) leads to [48]

$$R_x^{(k)}(\tau) = \sum_{i=-\infty}^{\infty} R_x(iT_s) h^{(k)}(\tau - iT_s), \quad k = 1, \dots \quad (7)$$

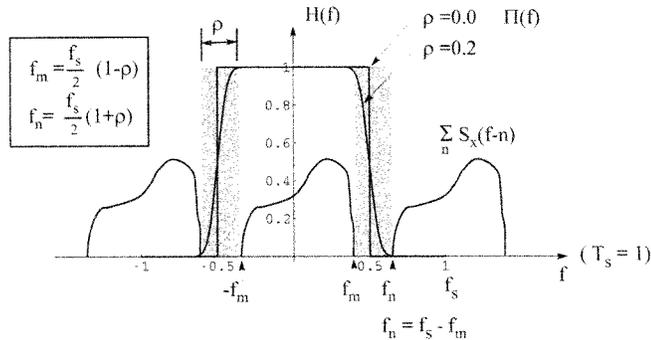


Fig. 1. Interpolation filter.

where  $h(\tau)$  is a well-behaved function whose Fourier transform fulfills

$$H(f) = \begin{cases} T_s & |f| \leq f_m \\ 0 & |f| \geq f_n \equiv f_s - f_m. \end{cases} \quad (8)$$

Fig. 1 shows two interpolation filters satisfying (8). If  $H$  was chosen to be the ideal lowpass filter with cutoff frequency  $f_s/2$  ( $H(f) = T_s \Pi(f/f_s)$ ), then (7) would be exactly the Whittaker–Shannon–Kotelnikov sampling theorem. The problem with the ideal lowpass filter is that its time response tends to zero with  $1/\tau$  as  $|\tau| \rightarrow \infty$ . This low vanishing rate is not tolerable, mainly for small sample sizes. However, this problem can be lightened by choosing a filter  $H$  with a smooth variation in the interval  $[f_m, f_n]$ . This issue will be addressed in Section III.

Placing (7) into (6) leads to

$$m_k(S_x) = \frac{1}{(2\pi j)^k} \sum_{i=-\infty}^{\infty} r_i h_i^k, \quad k = 0, 1, \dots \quad (9)$$

with  $r_i \equiv R_x(iT_s)$  and  $h_i^k \equiv h^{(k)}(-iT_s)$  for  $i \in \mathbf{Z} \equiv \{\dots, -1, 0, 1, \dots\}$ . Expression (9) is the basis of our approach to SM estimation. The following sections are mostly devoted to the design of suitable coefficients  $h_i^k$  and to the estimation of sequence  $\{r_i, i \in \mathbf{Z}\}$ .

### B. Maximum-Likelihood Estimate

As expressed in (9), estimates  $\hat{m}_k(\mathbf{Y})$  are somehow functions of the infinite sequence  $\{\hat{r}_i, i \in \mathbf{Z}\}$ . However, since our approach is nonparametric, given a sample vector  $\mathbf{Y}$  of size  $N$ , only the sequence  $\{\hat{r}_i, i = -N + 1, \dots, N - 1\}$  can be estimated from  $\mathbf{Y}$ . Thus, the infinite sum (9) cannot be computed and should be replaced by a truncated version. This raises no serious problems if the truncation error is small. Assume that for a given positive number  $\varepsilon_k$ , there exists an integer  $M_k$  such that

$$\frac{1}{(2\pi)^k} \sum_{|i| > M_k} |h_i^k| < \varepsilon_k \quad k = 0, 1, \dots \quad (10)$$

Under these conditions, the magnitude of the error between  $m_k$  and sum (9), truncated to  $|i| \leq M_k$ , is smaller than  $r_0 \varepsilon_k$  (notice that the nonnegative nature of  $\{r_i, i \in \mathbf{Z}\}$  implies that  $|r_i/r_0| \leq 1$ ). Since  $\mu_k = m_k/r_0$ , the magnitude of the error between  $\mu_k$  and the truncated version is smaller than  $\varepsilon_k$ .

In what follows, we assume that  $\varepsilon_k$  is negligible compared to the estimation error and denote the  $k$ th SM as the sum (9) truncated to  $M_k$ . The choice of the interpolation function  $h(\tau)$  and the size  $M_k(\varepsilon_k)$  of the truncated sum (9) will be addressed in Section III.

The invariance principle of ML estimation [53] states that if  $g(\boldsymbol{\theta}): \boldsymbol{\Theta} \rightarrow \boldsymbol{\Phi}$  is a mapping of  $\boldsymbol{\Theta}$  onto  $\boldsymbol{\Phi}$ , with  $\boldsymbol{\Theta} \subset \mathbb{R}^{n_1}$ ,  $\boldsymbol{\Phi} \subset \mathbb{R}^{n_2}$ , and  $n_1 \leq n_2$ , and  $\hat{\boldsymbol{\theta}}^{ml}$  is the ML estimate of  $\boldsymbol{\theta}$ , then  $g(\hat{\boldsymbol{\theta}}^{ml})$  is the ML estimate of  $g(\boldsymbol{\theta})$ . By applying this principle to a truncated version of (9), we obtain ML estimates of the SM's given by

$$\hat{m}_k^{ml} = \frac{1}{(2\pi j)^k} \sum_{i=-M_k}^{M_k} \hat{r}_i^{ml} h_i^k, \quad k = 0, 1, \dots \quad (11)$$

$$\hat{\mu}_k^{ml} = \frac{(2\pi j)^{-k}}{\hat{r}_0^{ml}} \sum_{i=-M_k}^{M_k} \hat{r}_i^{ml} h_i^k, \quad k = 1, 2, \dots \quad (12)$$

The same concept applies to the ML spectral width estimator. Hence

$$\hat{\sigma}^{ml} = [\hat{\mu}_2^{ml} - (\hat{\mu}_1^{ml})^2]^{1/2}. \quad (13)$$

Expressions (11)–(13) are exact ML estimates (under the bandlimited and negligible  $\varepsilon_k$  assumptions). They are simple functions of ML sequences  $\{\hat{r}_i^{ml}, i = -M_k, \dots, M_k\}$ . However, to our knowledge, they have not been derived and used in the literature.

## III. INTERPOLATION FILTER

A desirable characteristic of coefficients  $h_i^k$  is that their vanishing rate be as large as possible so that, given  $\varepsilon_k$ , inequality (10) is satisfied for the lowest possible value of  $M_k$ .

In accordance with the considerations of Section II-A, in case of  $f_s = 2f_m$  (sampling rate equal to Nyquist rate), there is no alternative to the ideal low-pass filter  $H(f) = T_s \Pi(f/f_s)$ . However, if the sampling rate is greater than the Nyquist rate, then the vanishing rate of coefficients  $h_i^k$  can be significantly increased. This is achieved by choosing an interpolation filter  $H(f)$  with a finite roll-off rate within the interval  $[f_m, f_n]$  ( $H(f)$  is, by assumption, even). Fig. 1 illustrates this concept (for simplicity,  $f_s = 1$ ). The gray rectangles demarcate the frequency intervals for which  $H$  is not defined. The length of these intervals is characterized by the roll-off factor  $\rho = (f_n - f_m)/f_s$  with  $0 \leq \rho < 1$ .

The behavior of  $h^{(k)}(\tau)$ , and consequently of  $h_i^k$ , depends on the continuity properties of  $H_k(f) = (2\pi j)^k H(f)$  and of its successive derivatives. If  $H_k(f)$  and all its derivatives up to order  $n$  exist and are of bounded variation in  $\mathbb{R}$ , then its inverse Fourier transform  $h^{(k)}(\tau)$  tends to zero at least as fast as  $1/\tau^{n+1}$  when  $|\tau| \rightarrow \infty$ . This result is a minor variation of [54, Theorem 1, p. 95].

In this work, we have chosen  $H(f)$  to be the raised cosine filter with impulse response [55 p. 179]

$$h(\tau) = T_s \frac{\sin(\pi f_s \tau)}{(\pi f_s \tau)} \frac{\cos(\pi f_s \rho \tau)}{1 - (2f_s \rho \tau)^2}$$

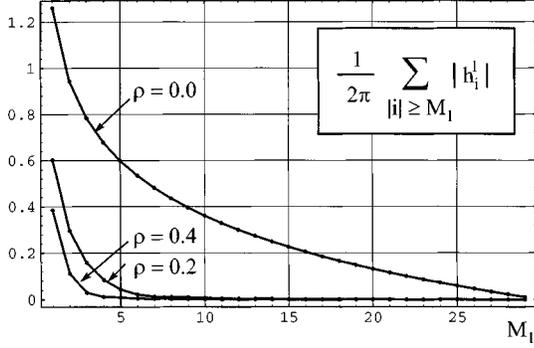


Fig. 2. Upper bound of truncation error associated with interpolation coefficients  $h_i^1$ . Parameter  $\rho$  is the roll-off factor of the raised-cosine filter.

where  $0 \leq \rho < 1$  is the roll-off factor. If  $\rho > 0$ , coefficients  $h_i^k$  tend to zero with  $1/i^3$  as  $|i| \rightarrow \infty$ , as opposed to  $1/i$  for  $\rho = 0$ .

Fig. 2 illustrates the behavior of  $((2\pi)^{-1} \sum_{|i| \geq M_k} |h_i^1|)$  for different values of the roll-off factor  $\rho$ . For example, if  $\rho = 0.2$ , which demands a sampling rate 1.25 above the Nyquist rate, an upper bound  $\varepsilon_1 = 0.01$  is achieved with just  $M_1 = 7$ . In the case of  $h_i^2$ , the same value of  $M_1 = 7$  leads to  $\varepsilon_2 < 0.0025$ .

#### IV. SAMPLE COVARIANCE-BASED ESTIMATE

Under weak regularity conditions, estimates  $\hat{m}_k^{ml}$  and  $\hat{\mu}_k^{ml}$ , for  $k = 0, 1, \dots$ , and  $\hat{\sigma}^{ml}$  given by (11)–(13), respectively, are best asymptotic normal estimates [44]. Despite the goodness of ML estimates, analytic solutions to the correspondent maximization problem are available only in a few simple cases. Numerical solutions are, very often, difficult to program and are computationally intensive (see e.g., [56]–[58]).

Taking into account (9) or ML estimators (11)–(13), it seems reasonable to exploit the properties of

$$\hat{m}_k = \frac{1}{(2\pi j)^k} \sum_{i=-M}^M \hat{r}_i h_i^k, \quad k = 0, 1, \dots \quad (14)$$

$$\hat{\mu}_k = \frac{(2\pi j)^{-k}}{\hat{r}_0} \sum_{i=-M}^M \hat{r}_i h_i^k, \quad k = 1, 2, \dots \quad (15)$$

$$\hat{\sigma} = [\hat{\mu}_2 - (\hat{\mu}_1)^2]^{1/2} \quad (16)$$

where,  $\{\hat{r}_i, i = -M, \dots, M\}$  is the sequence of unbiased sample covariance estimates. Assume that  $M_k = M$  for  $k = 1, 2, \dots$ . Given the sample vector  $\mathbf{Y}$  of size  $N > M$ , the unbiased sample covariances are given by

$$\hat{r}_i \equiv \begin{cases} \frac{1}{N-|i|} \sum_{n=1}^{N-|i|} y_{n+i} y_n^* - R_n(iT_s) & 0 \leq i \leq M \\ \hat{r}_i = \hat{r}_{-i}^* & -M \leq i \leq 0 \end{cases} \quad (17)$$

$$= \text{tr}\{\mathbf{J}_i \mathbf{Y} \mathbf{Y}^H\} - R_n(iT_s) \quad |i| \leq M \quad (18)$$

where  $(\cdot)^H \equiv ((\cdot)^*)^T$ ,  $\text{tr}\{\cdot\}$  denotes the trace operator, and  $\mathbf{J}_i$  is a null matrix, except for the  $i$ th diagonal, where its entries are

set to  $(N - |i|)^{-1}$ . The even nature of  $h(\tau)$  and the Hermitian property  $\hat{r}_{-i} = \hat{r}_i^*$  allow writing

$$\hat{m}_k = \frac{2j}{(2\pi j)^k} \left[ \sum_{i=1}^M \text{Im}[\hat{r}_i] h_i^k \right] \quad k = 1, 3, \dots \quad (19)$$

$$\hat{m}_k = \frac{2}{(2\pi j)^k} \left[ \hat{r}_0 h_0^k + \sum_{i=1}^M \text{Re}[\hat{r}_i] h_i^k \right] \quad k = 0, 2, \dots \quad (20)$$

meaning that even SM's depend only on  $\text{Re}[\hat{r}_i]$  and odd SM's depend only on  $\text{Im}[\hat{r}_i]$ .

Since the number  $M$  of sample covariances needed to determine the SM is moderate (typically smaller than 10), estimates  $\{\hat{r}_i, i = 0, \dots, M\}$  can be computed directly with approximately  $MN$  complex floating-point operations. Alternatively, they can be computed by inverting the periodogram, which has a complexity of the order of  $N \ln N$ .

The two sets of estimators (11)–(13) and (14)–(16) will be termed ML method and BL method, respectively.

#### A. Statistical Characterization

The statistical characterization of  $\hat{m}_k$  and  $\hat{\mu}_k$  for  $k = 0, 1, \dots$ , and  $\hat{\sigma}$  given by (14)–(16), respectively, was carried out in [44]. Herein, pertinent statistical properties of those estimators are stated without proof.

The PSD of sequence  $\{R_y(iT_s), -\infty < i < \infty\}$ , denoted by  $\lambda_y(F)$ , plays a central role in some of the following results. By definition

$$\lambda_y(F) \equiv \sum_{i=-\infty}^{\infty} R_y(iT_s) e^{-j2\pi F i}, \quad F \in \left[-\frac{1}{2}, \frac{1}{2}\right]. \quad (21)$$

Since  $R_y(iT_s) = R_x(iT_s) + R_n(iT_s)$ , then  $\lambda_y(F) = \lambda_x(F) + \lambda_n(F)$ .

In order to relate  $\lambda_y(F)$  with the underlying continuous echo and noise spectra  $S_x$  and  $S_n$ , define

$$S(F) = f_s \sum_{i=-\infty}^{\infty} S_x[(F-i)f_s] + S_n[(F-i)f_s] \quad (22)$$

is assumed to exist. By applying the Poisson formula to (22), after some manipulation, one concludes that  $S(F) = \lambda_y(F)$  for  $F \in \mathfrak{R}$ . Noting that  $S_x(f)$  has limited support  $[-f_m, f_m]$  with  $f_m \leq f_s/2$ , it results that

$$\lambda_y(F) = f_s S_x(Ff_s) + \lambda_n(F), \quad F \in \left[-\frac{1}{2}, \frac{1}{2}\right] \quad (23)$$

with

$$\lambda_n(F) = f_s \sum_{i=-\infty}^{\infty} S_n[(F-i)f_s] = \sum_{i=-\infty}^{\infty} R_n(iT_s) e^{-j2\pi F i}.$$

Very often, the noise CF verifies  $R_n(iT_s) = 0$  for  $|i| \geq 1$ , which defines a white noise sequence. In this case

$$\lambda_y(F) = f_s S_x(Ff_s) + R_n(0), \quad F \in \left[-\frac{1}{2}, \frac{1}{2}\right]. \quad (24)$$

In what follows, it is convenient to define

$$T_{ij} \equiv \frac{N}{(2\pi)^{i+j}} \text{tr}\{\mathbf{R}_y \mathbf{A}_i \mathbf{R}_y \mathbf{A}_j\}, \quad i, j = 0, 1, \dots \quad (25)$$

where  $\mathbf{R}_y \equiv E[\mathbf{Y}\mathbf{Y}^H]$

$$\mathbf{A}_k = \sum_{i=-M}^M a_i^k \mathbf{J}_i, \quad (26)$$

and

$$a_i^k = (-j)^k h_i^k. \quad (27)$$

The statistical characterization next presented is sectioned into three parts corresponding to zeroth SM, normalized SM's, and spectral width.

1) *Zeroth Spectral Moment:* The estimate  $\hat{m}_0 = \hat{r}_0$  exhibits the following statistical properties.

- 1) It is unbiased and has variance given by

$$V(\hat{m}_0) \equiv E\{(\hat{m}_0 - m_0)(\hat{m}_0 - m_0)^*\} = \frac{T_{00}}{N} \quad (28)$$

with  $T_{00}$  introduced in (25).

- 2) It is consistent with probability one (w.p.1) and also in the mean-square sense, verifying

$$\lim_{N \rightarrow \infty} NV(\hat{m}_0) = \int_1^{\frac{1}{2}} \lambda_y^2(F) dF \quad (29)$$

where  $\int_1 \equiv \int_{-1/2}^{1/2}$ .

- 3) The sequence  $\sqrt{N}(\hat{m}_0 - m_0)$  is asymptotically normal with zero-mean and variance given by (29).

2) *Normalized Spectral Moments:* The estimates  $\hat{\mu}_k$  with  $k = 1, 2, \dots$  exhibit the following statistical properties.

- 1) They are asymptotically unbiased

$$E[\hat{\mu}_k] = \mu_k - \frac{1}{Nm_0^2} [T_{0k} - \mu_k T_{00}] + O(N^{-2}) \quad (30)$$

where symbol  $O(N^{-2})$  means that  $N^2|O(N^{-2})| < \alpha$  for some  $\alpha > 0$ , and every  $N \geq N_0$ .

- 2) They are consistent w.p.1 and in the mean-square sense as well

$$E[(\hat{\mu}_k - \mu_k)^2] = \frac{1}{Nm_0^2} [T_{kk} - 2\mu_k T_{k0} + \mu_k^2 T_{00}] + O(N^{-2}). \quad (31)$$

- 3) The sequence  $\sqrt{N}(\hat{\mu}_k - \mu_k)$  is asymptotically normal with zero-mean and variance given by

$$\frac{(2\pi)^{-2k}}{m_0^2} \int_1 \lambda_y^2(F) |\overline{H}_k(F) - (2\pi)^k \mu_k \overline{H}_0(F)|^2 dF \quad (32)$$

where  $\overline{H}_k(F)$  is the discrete Fourier series (DFS) of sequence  $\{a_{-i}^k, |i| \leq M\}$  given by

$$\overline{H}_k(F) \equiv \sum_{i=-M}^M a_{-i}^k e^{-j(2\pi F)i}. \quad (33)$$

Function  $\overline{H}_k(F)$ , defined in (33), is closely related to the interpolation function  $H_k(f)$  introduced in Section III. Suppose

for a while that  $M = \infty$ . Recalling that  $a_i^k = (-j)^k h_i^k$  and that  $h_i^k = h^{(k)}(-iT_p)$ , expression  $\overline{H}_k(F)$  is then given by

$$\overline{H}_k(F) = (-j)^k \sum_{i=-\infty}^{\infty} h^{(k)}(iT_s) e^{-j(2\pi F)i}. \quad (34)$$

Applying the Poisson formula to (34) yields

$$\overline{H}_k(F) = \frac{(-j)^k}{T_s} \sum_{i=-\infty}^{\infty} H_k[(F-i)f_s]. \quad (35)$$

Filter  $H_k(f) = (j2\pi f)^k H(f)$  has support  $[-f_n, f_n]$ . Since  $f_n = (1+\rho)f_s/2$ , the neighbor replicas  $H_k[(F-i)f_s]$  and  $H_k[(F-i-1)f_s]$  overlap if  $\rho > 0$ . Nevertheless, in the interval  $[-f_m, f_m]$ , with  $f_m = (1-\rho)f_s/2$ , there is no overlapping. This implies that

$$\overline{H}_k(F) = (2\pi F f_s)^k \quad |F| \leq \frac{1}{2}(1-\rho) \quad (36)$$

since  $H(f) = T_s$  for  $f \in [-f_n, f_n]$ .

For  $M$  finite, filter  $\overline{H}_k(F)$  is obtained by noting that its coefficients  $\{a_{-i}^k, |i| \leq M\}$  are given by the product  $\{a_{-i}^k \text{rect}(i/(2M))\}_{i = \dots, -1, 0, 1, \dots}$ . Thus,  $\overline{H}_k(F)$  is the period 1 convolution of (35) with the DFS of  $\text{rect}(i/(2M))$ . Considering that  $M$  was chosen in such a way that  $\sum_{|i|>M} |h_i^k|$  is small, this convolution does not change substantially (35). Under these conditions, we conclude that

$$\lim_{N \rightarrow \infty} NV(\hat{\mu}_k) \simeq \frac{1}{m_0^2} \int_1 \lambda_y^2(F) [(Ff_s)^k - \mu_k]^2 dF. \quad (37)$$

3) *Spectral Width:* The estimate  $\hat{\sigma}$  exhibits the following statistical properties.

- 1) It is asymptotically unbiased with

$$E[\hat{\sigma}] = \sigma - \frac{1}{2\sigma Nm_0^2} \left[ (T_{02} - \mu_2 T_{00}) + (T_{11} - 2\mu_1 T_{10} + \mu_1^2 T_{00}) + \frac{1}{2\sigma^2} (T_{22} - 2\mu_2 T_{20} + \mu_2^2 T_{00}) \right] + O(N^{-2}). \quad (38)$$

- 2) It is consistent w.p.1 and also in the mean-square sense with

$$E[(\hat{\sigma} - \sigma)^2] = \frac{1}{4\sigma^2 Nm_0^2} [T_{22} - 4\mu_1 T_{12} - 2aT_{02} + 4\mu_1^2 T_{11} + 4\mu_1 a T_{01} + a^2 T_{00}] + O(N^{-2}) \quad (39)$$

where  $a = \sigma^2 - \mu_1^2$ .

- 3) The sequence  $\sqrt{N}(\hat{\sigma} - \sigma)$  is asymptotically normal with zero-mean and variance given by

$$\frac{(2\pi)^{-4}}{4m_0^2 \sigma^2} \int_1 \lambda_y^2(F) |\overline{H}_2(F) - 2(2\pi)\mu_1 \overline{H}_1(F) - (2\pi)^2 a \overline{H}_0(F)|^2 dF. \quad (40)$$

Again, using the fact that  $\overline{H}_k(F) \simeq (2\pi F f_s)^k$  for  $|F| \leq (1/2)(1 - \rho)$ , the variance (40) is approximately given by

$$\lim_{N \rightarrow \infty} NV(\hat{\mu}_k) \simeq \frac{1}{4m_0^2 \sigma^2} \int_1 \lambda_y^2(F) [(F f_s) - \mu_1]^2 - \sigma^2]^2 dF. \quad (41)$$

The sample covariances  $\{\hat{r}_k, k = 0, \dots, k = p\}$  of a Gaussian AR( $p$ ) process are asymptotically efficient [53, Theorem 5.8]. Since any process can be approximated by an AR( $\infty$ ) process, it is reasonable to admit that the sample covariances  $\{\hat{r}_k, k = 0, \dots, k = M\}$  are asymptotically efficient, as well as the SM's estimators herein proposed, which are based on the sample covariances.

Asymptotic variances of  $\hat{\mu}_1$  given by (37), with  $k = 1$ , and of  $\hat{\sigma}$  given by (41), are exactly the same as presented in [2] and [30] for the PB method. Despite this asymptotic similarity, the BL estimators of  $\mu_1$  and  $\sigma$  perform better than the corresponding PB ones for small sample sizes. This will become clear in the next section.

## V. PERFORMANCE EVALUATION

This section presents theoretical and simulation results concerning the mean frequency  $\mu_1$  and the spectral width  $\sigma$ , aiming at the comparison of the BL method (BLM), the ML method (MLM), the PB method (PBM), and the PP method (PPM). The following statistics are evaluated:

- 1) BLM biases  $E[\hat{\mu}_1 - \mu_1]$  and  $E[\hat{\sigma} - \sigma]$  given by (30) and (38), respectively;
- 2) BLM mean square errors (MSE's)  $E[(\hat{\mu}_1 - \mu_1)^2]$  and  $E[(\hat{\sigma} - \sigma)^2]$  given by (31) and (39), respectively;
- 3) *Cramér-Rao Bounds* (CRB's) of parameters  $\mu_1$  and  $\sigma$  with respect to the covariance sequence  $\{r_i, i = 0, \dots, M\}$ ;
- 4) sample mean biases  $T[\hat{\mu}_1 - \mu_1]$  and  $T[\hat{\sigma} - \sigma]$  of the BLM, the MLM, the PPM, and the PBM ( $T[x] \equiv (1/N) \sum_{i=1}^N x_i$  denotes the sample mean of  $N$  independent samples of random variable  $x$ );
- 5) BL estimates of  $\mu_1$  and  $\sigma$  computed according to (15) and (16), respectively, whereas ML estimates of these parameters are given by (12) and (13), respectively;
- 6) Sample MSE's  $T[(\hat{\mu}_1 - \mu_1)^2]$  and  $T[(\hat{\sigma} - \sigma)^2]$  of the BLM, the MLM, the PPM, and the PBM.

The following remarks are in order.

- 1) The random vector  $\mathbf{Y}$  was generated according to  $\mathbf{Y} = \mathbf{R}_y^{1/2} \mathbf{W}$ , where  $\mathbf{W}$  is a zero-mean  $N$ -dimensional, normal, circular-complex random vector of covariance matrix  $\mathbf{I}$  (identity), and  $\mathbf{R}_y^{1/2} = (\mathbf{R}_y^{1/2})^H$  is the square root of covariance matrix  $\mathbf{R}_y = \mathbf{R}_y^{1/2} \mathbf{R}_y^{1/2}$ , which exists since  $\mathbf{R}_y$  is Hermitian and nonnegative definite
- 2) The SM's are functions of the covariance sequence  $\{r_i, i = 1, \dots, M\}$ . On the other hand, the probability density function of the random vector  $\mathbf{Y}$  of size  $N$  depends on  $\{r_i, i = 1, \dots, N - 1\}$ . Since the present approach is nonparametric (i.e., no assumption is made concerning the structure of the

covariance function (besides being bandlimited), the CRB should be determined with respect to the whole set of parameters  $\{r_i, i = 1, \dots, N - 1\}$ . Define  $\boldsymbol{\theta} \equiv (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ , where  $\boldsymbol{\theta}_1 = (r_0, \text{Re}(r_1), \dots, \text{Re}(r_{N-1}))$  and  $\boldsymbol{\theta}_2 = (\text{Im}(r_1), \dots, \text{Im}(r_{N-1}))$ . The variance of any unbiased estimate of  $\mathbf{g}(\boldsymbol{\theta})$ , such as  $\hat{\mathbf{g}}(\mathbf{Y})$  subject to some regularity conditions, verifies [26]

$$\begin{aligned} \mathbf{V}(\boldsymbol{\theta}) &\equiv E\{(\hat{\mathbf{g}} - \mathbf{g})(\hat{\mathbf{g}} - \mathbf{g})^H\} \\ &\geq \mathbf{G}(\boldsymbol{\theta}) \mathbf{J}^{-1}(\boldsymbol{\theta}) \mathbf{G}^H(\boldsymbol{\theta}) \end{aligned} \quad (42)$$

where  $\mathbf{G}(\boldsymbol{\theta})$  is the Jacobian matrix of  $\mathbf{g}(\boldsymbol{\theta})$ ,  $\mathbf{J}$  is the information matrix, and  $\mathbf{G} \mathbf{J}^{-1} \mathbf{G}^H$  is the CRB.

From (12), for  $k = 1$  and (13), and also taking into account (19) and (20), one gets the Jacobians of  $\mu_1$  and  $\sigma$ , respectively, as

$$\frac{d\mu_1}{d\boldsymbol{\theta}} = \frac{1}{\pi r_0} \left[ -\pi \mu_1, \underbrace{0, \dots, 0}_{N-1}, h_1^1, \dots, h_M^1, \underbrace{0, \dots, 0}_{N-1-M} \right] \quad (43)$$

$$\frac{d\sigma}{d\boldsymbol{\theta}} = \frac{1}{2\sigma} \left[ \frac{d\mu_2}{d\boldsymbol{\theta}} - 2\mu_1 \frac{d\mu_1}{d\boldsymbol{\theta}} \right] \quad (44)$$

where

$$\begin{aligned} \frac{d\mu_2}{d\boldsymbol{\theta}} &= -\frac{1}{2\pi^2 r_0} \left[ \mu_2 2\pi^2 + h_0^2, h_1^2, \dots, h_M^2, \right. \\ &\quad \left. \underbrace{0, \dots, 0}_{N-1-M}, \underbrace{0, \dots, 0}_{N-1} \right]. \end{aligned} \quad (45)$$

For normal circular-complex random vectors, the element  $J_{kl}$  of the information matrix is given by [26]

$$J_{kl} = \text{tr} \left\{ \mathbf{R}_y^{-1} \mathbf{R}_y^{(k)} \mathbf{R}_y^{-1} \mathbf{R}_y^{(l)} \right\}. \quad (46)$$

Since  $\mathbf{R}_y$  is Hermitian and Toeplitz, its derivatives with respect to  $\theta_k$  are very easy to determine. To compute the whole matrix  $\mathbf{J}$ , we adopted the method proposed in [48] and [59], with complexity  $O(N^2 \ln N)$ . Notice that computing  $\mathbf{J}$  according to (46) (one inverse, three matrix products, and one trace) would have the unbearable complexity of  $O(N^4)$ .

- 3) The ML estimates  $\hat{\mu}^{ml}$  and  $\hat{\sigma}^{ml}$  depend both on the ML sequence  $\{\hat{r}_0^{ml}, \hat{r}_1^{ml}, \dots, \hat{r}_M^{ml}\}$ . This is a well-known instance of estimation of structured covariance matrix, particularly of ML estimation of Toeplitz covariance matrices [56]–[58], [60]–[62].

In this work, we adopted the methodology proposed in [56] with slight adaptations. Basically, the iterative expectation-maximization scheme [63], [64] was used to generate a sequence of covariance matrices  $\{\hat{\mathbf{R}}_y^t\}$ , for  $t = 0, 1, \dots$ , with increasing likelihood

- 4) The CRB refers to unbiased estimates. However, the methods herein considered are biased, mainly in what

spectral width is concerned. Nevertheless, the CRB is plotted jointly with the MSE's of the estimates. The main reason for this is that in many situations, the BLM is almost unbiased, being therefore comparable with the CRB.

The results next presented are grouped under the following major topics: a) symmetric spectra; b) asymmetric spectra; c) sample size; and d) SNR.

Gaussian and linear combination of rectangular spectral shapes are considered. The underlying reasons for studying linear combinations of rectangular forms are the following.

- According to (37) and (41), the performance of the BLM depends above all on how the spectral mass is clustered in the neighborhood of the respective SM. It does not depend, therefore, on small spectrum details. Linear combination of rectangular forms is a simple way of placing spectral mass.
- As stated and documented in the introduction, the PPM performance degrades as the degree of skewness and/or spectral width increases. It is therefore important to compare the performances of the BL and the PP methods for spectra having those features. The rectangular spectrum is the form that exhibits the largest spectral width in the smaller frequency interval. In addition, by using linear combination of rectangular forms, it is straightforward to build highly skewed spectra.

The Gaussian spectral shape, considered in all works on SM's estimation, is also included in this study, mainly for reference purposes.

In all the results presented, it is assumed that noise is white, i.e.,  $R_n(iT_s) = 0$  for  $|i| \geq 1$ , thus leading to a PSD associated with  $\{R_y(iT_s)\}$  given by (24).

In all simulations next presented, the sample mean  $T$  is obtained from 100 Monte Carlo simulations per point.

#### A. Symmetric Spectra

Figs. 3 and 4 display statistics relative to the Gaussian-shaped PSD

$$\lambda_y(F) = \frac{1}{\sqrt{2\pi}(\sigma T_s)} \exp\left(-\frac{F^2}{2(\sigma T_s)^2}\right) + R_n(0) \quad (47)$$

$$F \in \left[-\frac{1}{2}, \frac{1}{2}\right]$$

which, assuming a negligible aliasing (i.e.,  $\sigma T_s < 0.25$ ), is obtained from the underlying continuous PSD

$$S_x(f) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(f^2/2\sigma^2)}. \quad (48)$$

Spectrum (48) has power  $m_0 = 1$ , mean frequency  $\mu_1 = 0$ , and spectral width  $\sigma$ . The sample size is set to  $N = 32$  and the SNR = 20 dB ( $R_n(0) = 0.01$ ).

Fig. 3—*Mean Frequency*: In this figure, the abscissa is the scaled spectral width  $\sigma T_s$ , which takes values in the interval  $[0.02, 0.25]$  ( $\sigma = 0.02$  corresponds nearly to a sinusoid, whereas  $\sigma > 0.25$  would lead to aliasing). The ordinate is the root mean square error (RMSE) of  $\hat{\mu}_1$  scaled by  $\sqrt{N}T_s$ .

We call attention to the closeness between the theoretical BLM MSE and the CRB. For the present choice of spectrum parameters, the BLM is almost efficient. This fact is confirmed

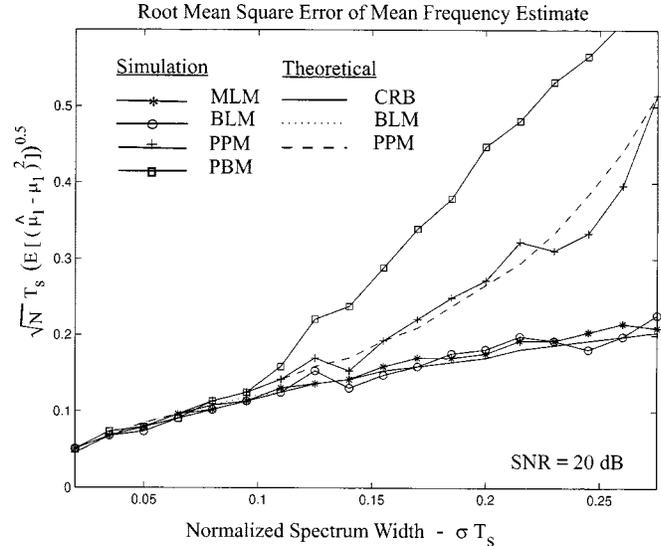


Fig. 3. RMSE of mean-frequency estimate  $\hat{\mu}_1$  for the Gaussian-shaped spectrum (47),  $N = 32$ , and SNR = 20 dB. The abscissa is the spectral width scaled by  $T_s$ , and the ordinate is the RMSE of  $\hat{\mu}_1$  scaled by  $\sqrt{N}T_s$ . The sample MSE is obtained from 100 Monte Carlo simulations per point. The CRB-0.5 of  $\mu_1$  and the standard deviation of the PPM are also plotted.

by the BLM simulation results. The behavior of the MLM can be taken as equal to that of the BLM (within the limits of sample variability). Concerning the PPM and the PBM, it is clear they perform worse when  $\sigma > 0.1$ .

The theoretical PPM curve plotted in Fig. 3 is the variance of  $\hat{\mu}_1$  given in [2]. We stress the agreement between the variance and the sample MSE of PPM. This is a consequence of having unbiased mean frequency PPM estimates when the spectrum exhibits symmetry about the mean frequency [29]. However, this is not so when the spectrum is asymmetric.

Fig. 4—*Spectral Width*: In this figure, the abscissa is the same as in Fig. 3, and the ordinate is the RMSE of  $\hat{\sigma}$  scaled by  $\sqrt{N}T_s$ .

The theoretical BLM MSE presented in Fig. 4 is greater than the CRB for  $\sigma \leq 0.1$ . However, for  $\sigma > 0.1$ , these two statistics are nearly equal. The MLM performance is very close to the CRB for all values of  $\sigma$  considered. The PPM and the PBM behaviors are similar to that of BLM for  $\sigma < 0.1$ . Yet, for  $\sigma > 0.1$ , the performance of the former two degrades severely.

The theoretical PPM curve plotted in Fig. 4 is the variance of  $\hat{\sigma}$  given in [2]. The slight increase between the variance and the sample MSE of the PPM is due to the nonnegligible bias of this estimator for  $\sigma > 0.15$ .

In conclusion, for SNR = 20 dB,  $N = 32$ ,  $\mu = 0$ ,  $\sigma T_s \in [0.02, 0.25]$ , and Gaussian-shaped spectra, we have the following.

- 1) The BLM and MLM produce estimates of  $\hat{\mu}_1$  and  $\hat{\sigma}$  that are practically unbiased.
- 2) The BLM and MLM produce estimates of  $\hat{\mu}_1$  that are practically efficient.
- 3) The BLM produce estimates of  $\hat{\sigma}$  that are practically efficient for  $\sigma > 0.1$ .
- 4) The MLM delivers estimates of  $\hat{\sigma}$  that are practically efficient.

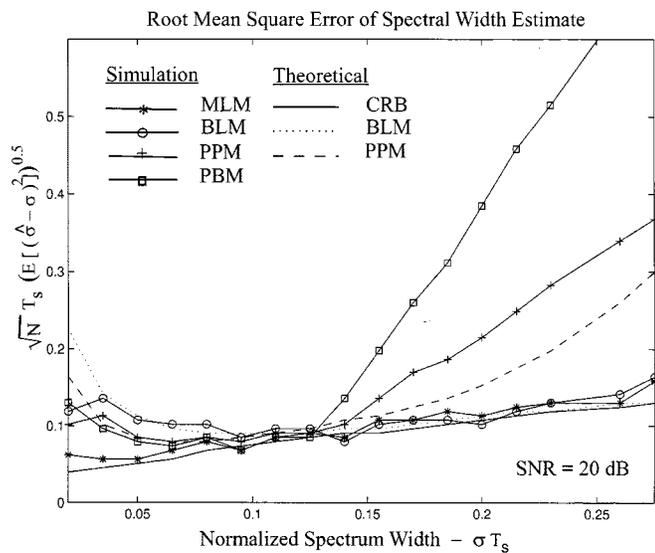


Fig. 4. RMSE of spectrum-width estimate  $\hat{\mu}_1$ , for Gaussian-shaped spectrum (47). Axes and curves have the same meaning as in Fig. 3.

- 5) The BLM performs similarly to the PPM and the PBM for low spectral widths. For high spectral widths, the BLM generates better estimates than the PPM and much better estimates than the PBM.

**B. Asymmetric Spectra**

Results presented in Figs. 6–9 correspond to the linear combination of rectangular-shaped spectra

$$\lambda_y(F) = \frac{8f_s}{4 + \gamma} \underbrace{\{\Pi(2Ff_s) + \gamma\Pi[8(Ff_s - 0.3125)]\}}_{f_s S_x(f = Ff_s)} + R_n(0). \tag{49}$$

Fig. 5 shows  $S_x(f)$ . For  $\gamma = 0$ ,  $S_x$  is symmetric with respect to the origin, having SM's  $m_0 = 1$ ,  $\mu_1 = 0$ , and  $\sigma = 0.144$ . For  $\gamma = \infty$ ,  $S_x$  is symmetric with respect to  $F = 0.3125$ , having SM's  $m_0 = 1$ ,  $\mu_1 = 0.3125$ , and  $\sigma = 0.036$ . Varying  $\gamma$  in the interval  $(0, \infty)$  yields spectra with different degrees of asymmetry.

Studying the behavior of SM estimators in the case of asymmetric spectra is of interest, because this situation is found in many applications (e.g., about 25% of the weather radar spectra are asymmetric [42]). On the other hand, it has been reported that the PPM's performance degrades as the degree of spectral asymmetry increases [2], [29].

From the results presented in Figs. 6–9, we conclude the following.

- 1) The BLM and the MLM produce estimates of  $\hat{\mu}_1$  and  $\hat{\sigma}$  that are practically unbiased.
- 2) The bias  $E[\hat{\sigma} - \sigma]$  presented in Fig. 8 follows roughly the sample bias  $T[\hat{\sigma} - \sigma]$ .
- 3) The BLM and the MLM produce estimates of  $\hat{\mu}_1$  and  $\hat{\sigma}$  that are practically efficient.
- 4) The BLM and the MLM have performance somewhat superior to that of PPM and much superior to that of PBM for the class of asymmetric spectra considered.

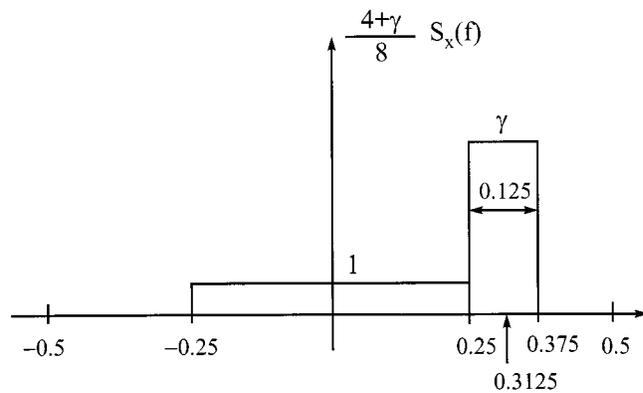


Fig. 5. Asymmetric spectrum (49). Different degrees of asymmetry are obtained by varying parameter  $\gamma$ . Maximum skewness is obtained, approximately, for  $\gamma = 5$ .

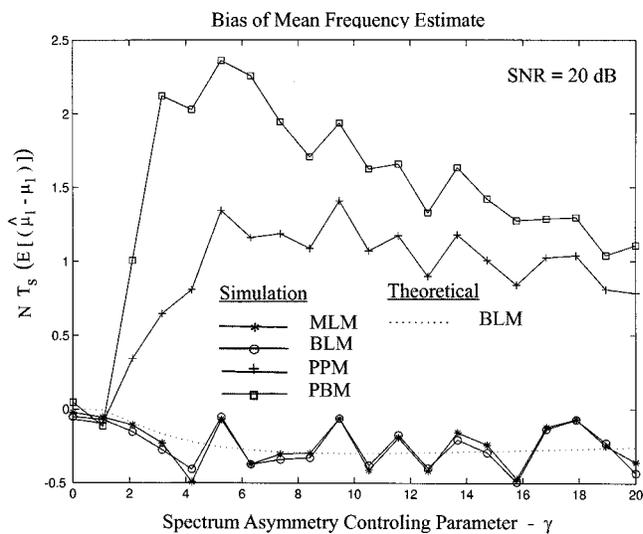


Fig. 6. Bias of mean-frequency estimate  $\hat{\mu}_1$  for the spectrum of Fig. 5,  $N = 32$ , and SNR = 20 dB. Parameter  $\gamma$  controls the degree of skewness. The ordinate is the bias of  $\hat{\mu}_1$  scaled by  $NT_s$ . The sample bias is obtained from 100 Monte Carlo simulations per point.

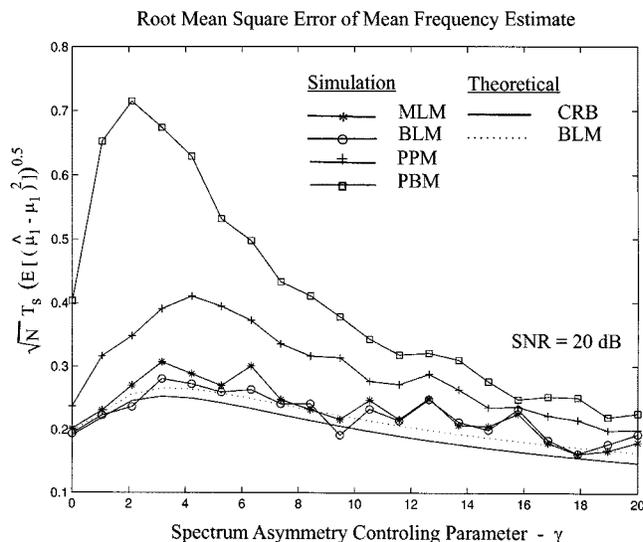


Fig. 7. RMSE of the mean-frequency estimate  $\mu_1$  for the asymmetric spectrum of Fig. 5,  $N = 32$ , and SNR = 20 dB. The abscissa is the same as in Fig. 6.

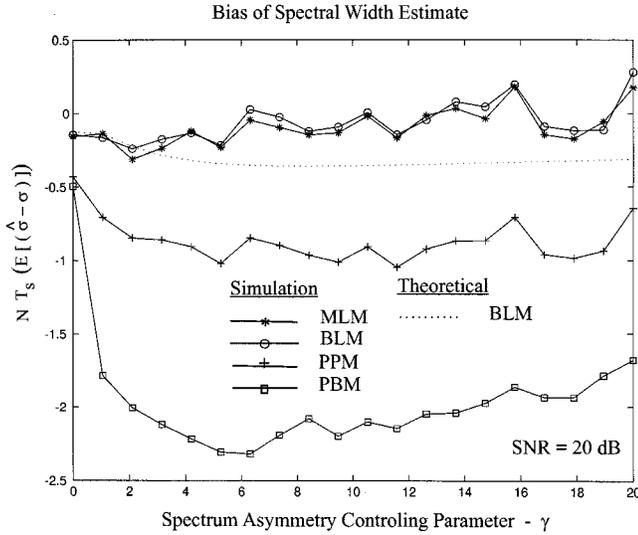


Fig. 8. Bias of the spectrum-width estimate  $\hat{\sigma}$  for the spectrum of Fig. 5,  $N = 32$ , and SNR = 20 dB. Axes and curves have the same meaning as in Fig. 6 and Fig. 7.

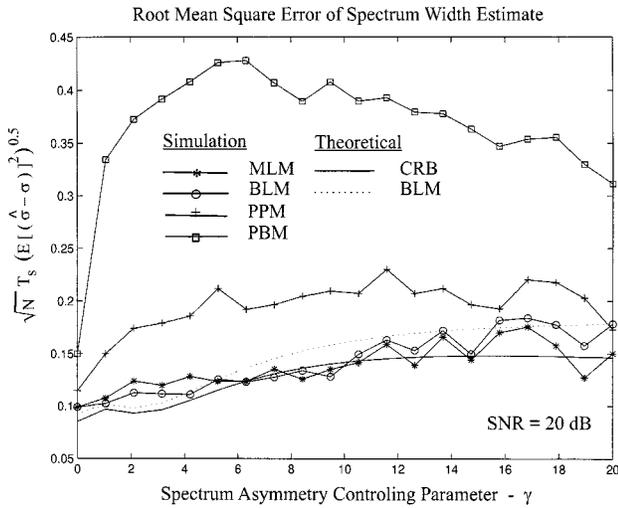


Fig. 9. RMSE of the spectrum-width estimate  $\hat{\sigma}$  for the spectrum of Fig. 5,  $N = 32$ , and SNR = 20 dB. Axes and curves have the same meaning as in Fig. 7.

The difference reported in item 2 is due to the low sample-size dimension used in the simulations [for  $N = 32$ , terms of order  $O(N^{-2})$  are still not negligible].

### C. Sample Size

Figs. 10 and 11 concern the MSE's of  $\hat{\mu}_1$  and of  $\hat{\sigma}$ , respectively, as functions of the sample size  $N$ . Data were generated according to the spectrum (49), taking  $\gamma = 5$  (this value of  $\gamma$  corresponds, approximately, to the maximum of spectral skewness). The MLM was not considered in this simulation, because the BLM is practically efficient (see curves of CRB and BLM in Figs. 10 and 11). With respect to the PPM and PBM, the BLM exhibits superior performance. This behavior is due to the large

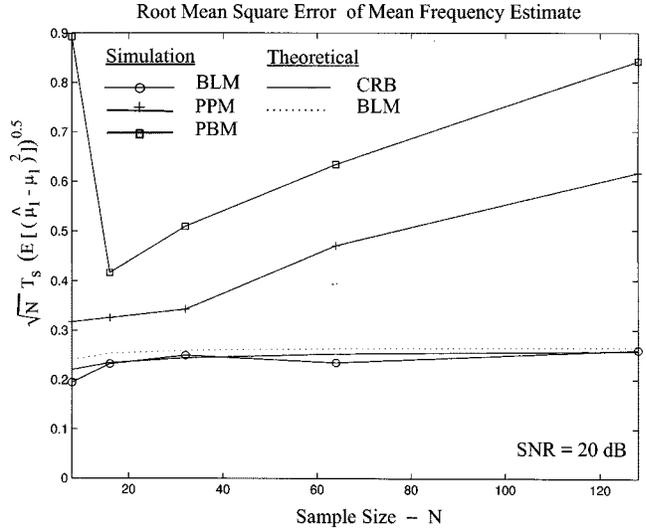


Fig. 10. RMSE of the mean-frequency estimate  $\hat{\mu}_1$  for the spectrum of Fig. 5,  $\gamma = 5$ , SNR = 20 dB, and  $N$  variable. Coordinate and curves have the same meaning as in Fig. 3.

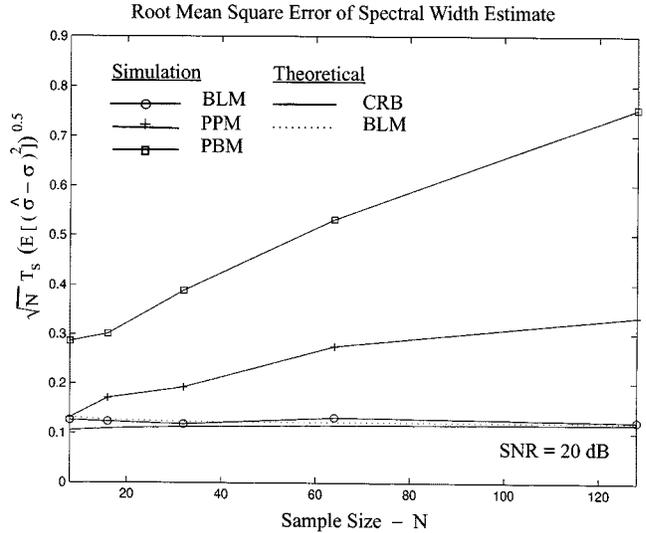


Fig. 11. RMSE of the spectrum-width estimate  $\hat{\sigma}$  for the spectrum of Fig. 5,  $\gamma = 5$ , SNR = 20 dB, and  $N$  variable. Ordinates and curves have the same meaning as in Fig. 3.

bias that PPM and PBM have (independently of  $N$ ), for asymmetric spectra.

### D. Signal to Noise Ratio (SNR)

With the purpose of studying the dependence of the BLM on the SNR, we have performed simulations using data generated according to the spectrum (49), taking  $\gamma = 5$  and SNR  $\in [5, 25]$  dB. For SNR  $\geq 10$  dB, the BLM MSE is close to the CRB, smaller than the PPM one, and much smaller than the PBM one. For  $5 \text{ dB} \leq \text{SNR} < 10 \text{ dB}$ , the MSE (of  $\mu_1$  and  $\sigma$ ) depart from the CRB, being smaller than the MSE of PBM, however. Still considering the interval  $5 \text{ dB} \leq \text{SNR} < 10 \text{ dB}$ , and in comparison to the PPM, the BLM exhibits better performance in the mean frequency estimation and comparable performance in the spectral-width estimation.

## VI. CONCLUSIONS

This paper introduces a novel nonparametric method for the estimation of SM's of stationary zero-mean normal circular-complex processes immersed in additive white Gaussian noise. The approach applies to discrete samples.

By assuming that the PSD is bandlimited (a hypothesis valid in applications such as weather radar, ultrasound imaging, clear-air turbulence measurement, SAR, and electroencephalographic analysis), it was shown that the SM's can be obtained by linear combination of covariance-function samples taken at the sampling instants with coefficients obtained from adequate interpolating functions.

The number of terms of the linear combinations referred to previously is infinite. However, in a practical implementation, they must be somehow truncated. In order to minimize the truncation error, a raised cosine-type interpolating filter was adopted. This family of functions assures a vanishing rate proportional to  $k^{-3}$  ( $k$  denotes the  $k$ th multiple of the sampling period). In what concerns the SM's  $\mu_1$  and  $\mu_2$ , it was shown that a roll-off factor of only 0.2 in the raised-cosine filter and seven terms in the linear combinations lead to a truncation error smaller than 0.01 and 0.0025, respectively.

By applying the invariance principle of the maximum-likelihood estimation to the truncated linear combinations, (assuming negligible truncation errors in comparison with the estimation standard deviation) maximum-likelihood estimators of the SM's were obtained. This procedure was termed maximum-likelihood method (MLM).

An estimator with the same formal structure as the MLM, but with the maximum-likelihood covariance sequence replaced by the unbiased sample-covariance sequence, was also studied. This procedure was named bandlimited method (BLM).

The BLM exhibits the following statistical properties.

- 1) Nonnormalized SM estimators are unbiased, mean square consistent, consistent with probability one, and asymptotically normal.
- 2) Normalized SM estimators are asymptotically unbiased (with bias proportional to  $N^{-1}$ ), mean square consistent, consistent with probability one, and asymptotically normal.
- 3) The spectral-width estimator is asymptotically unbiased (with bias proportional to  $N^{-1}$ ), mean square consistent, consistent with probability one, and asymptotically normal.

The following conclusions concerning the mean frequency  $\mu_1$  and the spectral width  $\sigma$  are based on simulation results.

- 1) The BLM and the MLM performances are much better than the ones of the PPM and the PBM for asymmetric spectra.
- 2) The BLM and MLM performances are much better than the ones of the PPM and the PBM, for spectra with medium or large spectral widths.
- 3) The BLM performance is comparable to the ones of the PPM and the PBM for spectra with small spectral widths.
- 4) The BLM and the MLM MSE's are close to the  $\text{CRB}(\mu_1)$ .

5) The BLM MSE is close to the  $\text{CRB}(\sigma)$  for  $\text{SNR} \geq 10$  dB.

6) The BLM MSE is close to the  $\text{CRB}(\sigma)$  for  $\sigma > 0.1$ .

The results presented correspond to Gaussian and rectangular-shaped spectra. However, it was verified (by simulation) that the drawn conclusions remain valid for other spectral shapes (e.g., triangular and rational). The justification for this fact is due to the behavior of terms  $T_{ij}$  defined in (25), on which all considered statistics depend. They tend asymptotically to integrals [e.g., (37) and (40)], having little sensitivity to the spectral detail.

## REFERENCES

- [1] J. Dias and J. Leitão, "Nonparametric estimation of mean velocity and spectral width in weather radar," in *IEEE Int. Geoscience and Remote Sensing Symp.—IGARSS'95*, Florence, July 1995, pp. 2121–2125.
- [2] D. S. Zrnic, "Estimation of spectral moments for weather echos," *IEEE Trans. Geosci. Electron.*, vol. GE-17, pp. 113–128, Oct. 1979.
- [3] —, "Spectrum width estimates for weather echos," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-15, pp. 613–619, Sept. 1979.
- [4] P. Mahapatra and D. Zrnic, "Practical algorithms for mean velocity estimation in pulse Doppler weather radars using a small number of samples," *IEEE Trans. Geosci. Remote Sensing*, vol. GE-21, pp. 491–501, Oct. 1983.
- [5] R. J. Doviak and D. S. Zrnic, "The Doppler weather radar," in *Aspects of Modern Radar*, E. Brookner, Ed. Norwood, MA: Artech, 1988.
- [6] —, *Doppler Radar and Weather Observations*. New York: Academic, 1993.
- [7] E. Chornobody, "Optimal mean velocity estimation for Doppler weather radars," *IEEE Trans. Geosci. Remote Sensing*, vol. 31, pp. 575–586, May 1993.
- [8] J. Dias and J. Leitão, "Asymptotically efficient estimation of spectral moments," *IEEE Trans. Signal Processing*, vol. 43, pp. 2222–2225, Sept., 1995.
- [9] B. Angelsen, "Instantaneous frequency, mean frequency, and variance of mean frequency estimators for ultrasonic blood velocity Doppler signals," *IEEE Trans. Biomed. Eng.*, vol. BME-28, pp. 733–741, Nov. 1981.
- [10] B. Angelsen and K. Kristoffersen, "Discrete time estimation of the mean Doppler frequency in ultrasonic blood velocity measurements," *IEEE Trans. Bio. Eng.*, vol. 30, pp. 207–214, Apr. 1983.
- [11] W. Barder, J. Eberhard, and S. Karr, "A new time domain technique for velocity measurements using Doppler ultrasound," *IEEE Trans. Biomed. Eng.*, vol. BME-32, pp. 213–229, Mar. 1985.
- [12] K. Kristoffersen, "Time-domain of the center frequency and the spread of Doppler spectra in diagnostic ultrasound," *IEEE Trans. Ultrason. Ferroelect. Freq. Contr.*, vol. 35, pp. 685–700, Nov. 1988.
- [13] J. David, S. Jones, and D. Giddens, "Modern spectral analysis techniques for blood flow velocity and spectral measurements with pulsed Doppler ultrasound," *IEEE Trans. Biomed. Eng.*, vol. BME-38, pp. 589–597, June 1991.
- [14] K. S. Gage and B. B. Balsley, "MST radar studies of wind and turbulence in the middle atmosphere," *J. Atmos. Terrest. Phys.*, vol. 46, pp. 793–753, 1984.
- [15] R. D. Palmer, J. R. Cruz, and D. Zrnic, "Enhanced autoregressive moving average spectral estimation applied to measurement of Doppler spectral width," *IEEE Trans. Geosci. Remote Sensing*, vol. 29, pp. 358–368, May 1991.
- [16] E. Li, D. Curlander, and C. Wu, "Doppler estimation for spaceborne synthetic aperture radars," *IEEE Trans. Geosci. Remote Sensing*, vol. GE-23, pp. 576–585, Jan. 1985.
- [17] S. Madsen, "Estimating the Doppler centroid of SAR data," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 25, pp. 134–141, Mar. 1989.
- [18] R. Bamler and H. Runge, "PRF-ambiguity resolution by wavelength diversity," *IEEE Trans. Geosci. Remote Sensing*, vol. 29, pp. 997–1003, Nov. 1991.
- [19] F. Wong and I. Cumming, "A combined SAR Doppler centroid estimation scheme based upon signal phase," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 696–707, May 1996.
- [20] D. Michael and J. Houchin, "Automatic EEG analysis: A segmentation procedure based on the autocorrelation function," *Electroencephalogr. Clin. Neurophysiol.*, vol. 46, pp. 232–235, 1979.

- [21] B. Saltzberg, W. Buton, and N. Burch, "Moments of the power spectral density from samples of the autocorrelation function," *Electroenceph. Clin. Neurophysiol.*, vol. 61, pp. 89–93, 1985.
- [22] H. Strube, "Spectral moments from the ACF or directly from the signal," *Signal Processing*, vol. 9, pp. 51–55, Aug., 1985.
- [23] A. Ishimaru, *Wave Propagation and Scattering in Random Media*. New York: Academic, 1978, vol. 1.
- [24] A. Macovski, *Medical Imaging Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1983.
- [25] K. Hutchinson, "Axial variation of Doppler ultrasound spectral shape in the *in vivo* post-stenotic shear layer," *J. Vascular Invest.*, vol. 1, pp. 15–25, Mar., 1995.
- [26] K. S. Miller, *Complex Stochastic Processes. An Introduction to Theory and Applications*. Reading, MA: Addison-Wesley, 1974.
- [27] K. Miller and M. Roehwarger, "On estimating spectral moments in the presence of colored noise," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 303–309, May 1970.
- [28] —, "Estimation of spectral moments of time series," *Biometrika*, vol. 57, pp. 513–517, 1970.
- [29] —, "A covariance approach to spectral moment estimation," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 588–596, Sept. 1972.
- [30] T. Berger and H. H. Groginsky, "Estimation of spectral moments of pulse trains," in *Proc. Int. Conf. Information Theory*, Tel-Aviv, Israel, 1973, pp. 1522–1553.
- [31] D. Sirmans and B. Bumgarner, "Numerical comparison of five mean frequency estimators," *J. Appl. Meteorol.*, vol. 14, pp. 991–1003, Sept., 1975.
- [32] R. Strauch, R. Kropfli, W. Sweezy, W. Moninger, and R. Lee, "Improved Doppler velocity estimates by the poly-pulse-pair method," in *Amer. Meteorol. Soc., Proc. 18th Conf. Radar Meteorology*, Atlanta, GA, 1978.
- [33] G. Hyde and K. Perry, "Doppler phase difference integrator," Lincoln Lab., Mass. Inst. Technol., Cambridge, MA, 1958.
- [34] D. Sirmans, "Real time estimates of mean velocity by averaging quantized phase displacement of Doppler radar echoes," 1973 SWIECO Record Tech. Papers, ser. IEEE Catalog No. 73CHO 719-5 SWIECO, 1973.
- [35] P. Bello, "Some techniques for the instantaneous measurement of multipath and Doppler spread," *IEEE Trans. Commun. Technol.*, vol. CT-13, pp. 285–292, 1965.
- [36] D. S. Zmic, "Spectral moments estimate from correlated pulse pairs," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-13, pp. 344–354, July 1977.
- [37] J. Dias and J. Leitão, "Estimation of spectral moments in Doppler weather radar," in *Proc. URSI Int. Symp. Signals, Systems and Electronics—ISSSE'92*, Paris, France, 1992, pp. 492–495.
- [38] P. Whittle, "Estimation and information in stationary time series," *Arkiv Matematik*, vol. B-2, no. 23, pp. 423–434, 1953.
- [39] M. J. Levin, "Power spectrum parameter estimation," *IEEE Trans. Inform. Theory*, vol. IT-11, pp. 100–107, Nov. 1965.
- [40] J. Dias and J. Leitão, "Maximum likelihood estimation of spectral moments at low signal to noise ratios," in *IEEE Int. Conf. Acoustics, Speech, and Signal Processing—ICASSP'93*, vol. 4, Minneapolis, MN, Apr. 1993, pp. 149–152.
- [41] J. Klostermeyer, "Maximum entropy estimation of Doppler shift and spectral width of VHF radar signals," *Radio Sci.*, vol. 24, pp. 47–63, 1989.
- [42] L. H. Janssen and G. A. Van der Spek, "The shape of Doppler spectra from precipitation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-21, pp. 208–219, Mar. 1985.
- [43] R. R. Rogers, "Radar measurement of velocities of meteorological scatterers," *J. Atmos. Sci.*, vol. 20, pp. 170–174, Mar., 1963.
- [44] J. Dias, "Estimation of Spectral Moments: Conceptual and Computational Aspects," Ph.D. thesis (in Portuguese), Inst. Superior Técnico, Lisbon, Portugal, 1994.
- [45] B. Picinbono, "On circularity," *IEEE Trans. Signal Processing*, vol. 42, pp. 3473–3482, Dec. 1994.
- [46] P. R. Wallace, "Interpretation of the fluctuation echo from randomly distributed scatterers. Part 2," *Can. J. Phys.*, vol. 31, pp. 995–1009, 1953.
- [47] R. S. Berkowitz, *Modern Radar, Analysis, Evaluation and System Design*. New York: Wiley, 1965.
- [48] J. Dias, "Aspects of a meteorological radar system," Ph.D. dissertation (in Portuguese), Dep. Eng. Elect. Comput., Instituto Superior Técnico, Lisbon, Portugal, 1991.
- [49] P. Kelly, H. Derin, and K. Hartt, "Adaptive segmentation of speckled images using a hierarchical random field model," *IEEE Trans. Acoust. Speech, Signal Process.*, vol. 36, pp. 1628–1641, Oct. 1988.
- [50] J. Goodman, "Statistical properties of laser speckle patterns," in *Laser Speckle and Related Phenomena*. ser. Topics in Applied Physics, J. C. Dainty, Ed. Heidelberg: Springer-Verlag, 1975, vol. 9, pp. 9–75.
- [51] —, "Some fundamental properties of speckle," *J. Opt. Soc. Amer.*, vol. 66, pp. 1145–1150, Nov. 1976.
- [52] A. Zayed, *Advances in Shannon's Sampling Theory*. Boca Raton, FL: CRC, 1993.
- [53] B. Porat, *Digital Processing of Random Signals*. Englewood Cliffs, NJ: Prentice-Hall, 1994.
- [54] A. Papoulis, *Signal Analysis*. New York: McGraw-Hill, 1977.
- [55] J. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.
- [56] M. Miller and D. Snyder, "The role of likelihood and entropy in incomplete data problems: Application to estimating point processes intensity and Toeplitz constrained covariances," *Proc. IEEE*, vol. 75, pp. 892–907, July 1987.
- [57] J. Burg and D. Luenberger, "Estimation of structured covariance matrices," *Proc. IEEE*, vol. 70, pp. 963–974, Dec. 1982.
- [58] D. Williams and D. Johnson, "Robust estimation of structured covariance matrices," *IEEE Trans. Signal Processing*, vol. 41, pp. 2891–2906, Sept. 1993.
- [59] J. Dias and J. Leitão, "A method for computing the information matrix of stationary Gaussian processes," in *Proceedings of EUSIPCO'96*, vol. I, Trieste, Sept. 1996, pp. 240–243.
- [60] T. Chonavel and P. Loubaton, "A maximum likelihood estimation algorithm for autocovariance matrices of band-limited processes," in *Proc. EUSIPCO—1994*, M. Holt, C. Cowan, and W. Sandham, Eds., Edinburgh, Scotland, 1994, pp. 498–501.
- [61] D. Fuhrmann and M. Miller, "On the existence of positive-definite maximum likelihood estimates of structured covariance matrices," *IEEE Trans. Inform. Theory*, vol. 34, pp. 722–729, July, 1988.
- [62] B. Veen and L. Scharf, "Estimation of structured covariance matrices and multiple window spectrum analysis," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1467–1471, Aug., 1990.
- [63] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood estimation from incomplete data via the EM algorithm," *J. R. Stat. Soc. B*, vol. 39, pp. 1–38, 1977.
- [64] A. Little and D. Rubin, *Statistical Analysis with Missing Data*. New York: Wiley, 1987.



**José M. B. Dias** (S'87–M'95) received the E.E., M.Sc., and Ph.D. degrees in electrical and computer engineering, all from Instituto Superior Técnico (IST), Technical University of Lisbon, Portugal, in 1985, 1991, and 1995, respectively.

He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, IST. He is also a Researcher with the Communication Theory and Pattern Recognition Group of the Institute of Telecommunications, Lisbon, Portugal. His research interests include remote sensing, signal

and image processing, and telecommunications.



**José M. N. Leitão** (M'95) received the E.E. and Ph.D. degrees in electrical engineering in 1970 and 1983, respectively, both from Instituto Superior Técnico (IST), Technical University of Lisbon, Portugal. He received the Agregado degree in electrical and computer engineering, also from IST, in 1992.

He was with the Laboratory of Physiology, Instituto Gulbenkian de Ciência, in Oeiras, Portugal, from 1970 to 1972. After spending three years at the University of Tübingen, Germany, he joined the faculty of IST in 1976, where he is currently a full Professor

with the Department of Electrical and Computer Engineering. He is also the coordinator of the Communication Theory and Pattern Recognition Group of the Institute of Telecommunications, Lisbon, Portugal. His main research interests are communication and information theory, pattern recognition, and signal and image processing.