ASYMPTOTICALLY EFFICIENT ESTIMATION OF SPECTRAL MOMENTS

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

Instituto de Telecomunicações and

Departamento de Engenharia Electrotécnica e de Computadores,

Instituto Superior Técnico, Lisboa, PORTUGAL,

Fax: 351.1.3523014 Email:d2287@beta.ist.utl.pt

ABSTRACT

This paper studies parametric estimation of spectral moments of a zero-mean complex Gaussian stationary process immersed in independent Gaussian noise. With the merit of the *maximum-likelihood* (ML) approach as motivation, this work exploits a Whittle's type objective function able to capture the relevant features of the *log-likelihood function,* while being much more manageable. The resulting estimates are strongly consistent and asymptotically efficient. As an example, application to Doppler weather radar data is considered.

EDICS category: 3.6.1 - Parameter Estimation: Single-Channel Time Series.

ABSTRACT

This paper studies parametric estimation of spectral moments of a zero-mean complex Gaussian stationary process immersed in independent Gaussian noise. With the merit of the *maximum-likelihood* (ML) approach as motivation, this work exploits a Whittle's type objective function able to capture the relevant features of the *log-likelihood function*, while being much more manageable. The resulting estimates are strongly consistent and asymptotically efficient. As an example, application to Doppler weather radar data is considered.

I. INTRODUCTION

In general, spectral estimation addresses the problem of *power spectral density* (PSD) inference from a finite *observation* of the underlying process. There are situations, however, in which the goal is the determination of PSD functionals, rather than the PSD itself; a relevant example is *spectral moments* (SM) estimation, the main theme of this paper. The motivation stems from our previous work on Doppler weather radar [1], [2], [3], where the goal is the determination of the three first SM, within the socalled *resolution volume* [4]. These moments are closely related to physical entities of interest: the zeroth moment (*mean power*) is related with the water content; the first moment (*mean frequency*) is related with the hydrometeors mean radial velocity; and the square root of the second centered moment (*spectral width*) is a measure of the hydrometeors velocity dispersion. Besides Doppler weather radar, SM estimation finds application on clear-air turbulence measurement; ultrasound imaging in medicine; and synthetic aperture radar, only to name a few.

Well known nonparametric SM estimation techniques are the *pulse pair* (PP) [5] and the periodogram based (PB) estimate [4]. For *bandlimited* processes, exact ML nonparametric estimates were derived in [2].

Concerning parametric approach, the optimality characteristics of the ML criterion (at least asymptotically) have fostered its study in several fields. However, determining the ML solution is often cumbersome (both in analytical and computational senses) [6]. As a consequence, the ML criterion has been put aside in what SM is concerned.

Our objective in this work is to come out with a parametric SM estimator that asymptotically exhibits ML properties having, however, a tolerable complexity. This is attained by replacing the log-likelihood function with a *close* Whittle's type objective function [7]. By hypothesis, the parametric model is known up to a multiplicative constant, a frequency shift, and a scale factor.

The main contributions presented in this work are the following: (a) a criterion yielding asymptotically efficient SM estimates; (b) the strong consistency of the proposed estimates; (c) an approximate expression for the estimate's bias; and (d) a procedure of *moderate* complexity providing the wanted SM estimates.

The paper is organized as follows: section II introduces notation, hypotheses about the underlying process, formulates the problem, and derives the approximate ML estimator; statistical characterization is also provided. Section III specializes the previous concepts to Doppler-spread data. Section IV develops an application using weather radar data.

II. PROBLEM FORMULATION AND APPROXIMATE ML ESTIMATOR

Let $\mathcal{X} = \{x(t),\, t \in \Re\}$ and $\mathcal{N} = \{n(t),\, t \in \Re\}$ be independent zero-mean complex Gaussian strictly stationary processes, with covariance functions (CF's) $R_x(\tau)$ and $R_n(\tau)$, respectively. The PSD associated to $R_x(\tau)$ is denoted by $S_x(f)$. We assume that $R_n(\tau)$ is known in advance (notice that, in many applications, $R_n(\tau)$ can be measured/estimated in with arbitrary precision). Since our approach is parametric, we write $R_x(\tau, \boldsymbol{\theta})$ and $S_x(f, \boldsymbol{\theta})$ with $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \Re^p$ and $p \geq 1$.

The problem of SM estimation is then stated as follows: given M consecutive samples $\mathbf{Y} = [Y_1, \ldots, Y_M]^T$, with $Y_i = y(iT_s) = x(iT_s) + n(iT_s)$, design estimators of the spectral moments (assumed to exist) given by

$$m_k(S_x) \equiv \int_{-\infty}^{\infty} f^k S_x(f, \boldsymbol{\theta})\, df \;\; = \;\; \frac{1}{(2\pi j)^k} R_x^{(k)}(0, \boldsymbol{\theta}), \qquad k = 1, 2, \ldots, \tag{1}$$

where ($j \equiv \sqrt{-1}$) and $R_x^{(k)}(0, \boldsymbol{\theta})$ stands for the k-th derivative of $R_x(0, \boldsymbol{\theta})$ with respect to τ, at $\tau = 0$. Besides m_k, we are also interested in the spectral width

$$\sigma(S_x) = \left[\frac{m_2}{m_0} - \left(\frac{m_1}{m_0} \right)^2 \right]^{1/2}. \tag{2}$$

Let $\widehat{\boldsymbol{\theta}}^{ml}$ be the ML estimate of parameter $\boldsymbol{\theta}$. The invariance of ML estimation [8] implies that $\widehat{m}_k^{ml} = m_k[S_x(f, \widehat{\boldsymbol{\theta}}^{ml})]$ and that $\widehat{\sigma}^{ml} = \sigma[S_x(f, \widehat{\boldsymbol{\theta}}^{ml})]$, i.e., ML estimation of SM reduces to ML estimation of parameter $\boldsymbol{\theta}$. Therefore, we will focus our attention on $\widehat{\boldsymbol{\theta}}^{ml}$.

Define $\mathbf{R}_y(\boldsymbol{\theta}) \equiv E[\mathbf{Y}\mathbf{Y}^H] = \{R_y[(i-j)T_s, \boldsymbol{\theta}],\; i, j = 1, \ldots, M\}$ (where $(\cdot)^H$ stands for the Hermitian operator) with $R_y(\tau, \boldsymbol{\theta}) = R_x(\tau, \boldsymbol{\theta}) + R_n(\tau)$. Recalling that $\mathcal{X}$ and $\mathcal{N}$ are zero-mean complex Gaussian strictly stationary processes and assuming that $\mathbf{R}_y^{-1}(\boldsymbol{\theta})$ exists for $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, it follows that the probability density function of $\mathbf{Y}$ is

$$f_y(\mathbf{Y}=\mathbf{y}|\boldsymbol{\theta}) = \frac{1}{\pi^M |\mathbf{R}_y(\boldsymbol{\theta})|} \exp\left[-\mathbf{y}^H \mathbf{R}_y^{-1}(\boldsymbol{\theta})\mathbf{y}\right]. \tag{3}$$

Let $\boldsymbol{\theta}_0$ be the true parameter; its ML estimate is given by

$$\widehat{\boldsymbol{\theta}}^{ml} = \arg\max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \Lambda(\mathbf{Y}|\boldsymbol{\theta}) \quad \text{where} \quad \Lambda(\mathbf{Y}|\boldsymbol{\theta}) = \ln[f_y(\mathbf{Y}|\boldsymbol{\theta})]. \tag{4}$$

Finding $\widehat{\boldsymbol{\theta}}^{ml}$ demands explicit expressions for $|\mathbf{R}_y(\boldsymbol{\theta})|$ and $\mathbf{R}_y^{-1}(\boldsymbol{\theta})$. In most cases, these expressions are very difficult to obtain [6]. It is thus natural to apply iterative schemes to compute $\widehat{\boldsymbol{\theta}}^{ml}$; relevant examples are the estimation of structured covariance matrices [9], [10] and the estimation of ARMA parameters [11]; the expectation maximization algorithm have also been applied to iteratively solve ML maximization problems [12]. Regarding SM, the computational burden of these methods is unbearable in most practical applications.

Seeking for an estimator with ML (or nearly ML) features but with lighter complexity, we exploit an objective function $L(\mathbf{Y}|\boldsymbol{\theta})$ having the same structure of $\Lambda(\mathbf{Y}|\boldsymbol{\theta})$, but where $\mathbf{R}_y(\boldsymbol{\theta})$ is replaced with $\mathbf{R}_{cy}(\boldsymbol{\theta}) = \{R_{cy}[(i-j)T_s, \boldsymbol{\theta}],\, i,j = 1,\ldots,M\}$, and $R_{cy}(\tau,\boldsymbol{\theta}) = \sum_{k=+\infty}^{-\infty} R_y(\tau + kMT_s, \boldsymbol{\theta})$. Assuming that $R_y(iT_s,\boldsymbol{\theta})$ is negligible for $|i| \geq m$ and $M > 2m+1$, then $R_{cy}(\tau,\boldsymbol{\theta})$ is a periodic extension of $R_y(\tau,\boldsymbol{\theta})$. Observe that $\mathbf{R}_{cy}(\boldsymbol{\theta})$ is a right circulant matrix; with exception of the upper right and lower left corners, matrices $\mathbf{R}_{cy}(\boldsymbol{\theta})$ and $\mathbf{R}_y(\boldsymbol{\theta})$ are equal. Thus, if $M \gg 2m$, it is expectable that estimates based on $\mathbf{R}_{cy}(\boldsymbol{\theta})$ and on $\mathbf{R}_y(\boldsymbol{\theta})$ are *close*.

Consider the eigendecomposition $\mathbf{R}_{cy}(\boldsymbol{\theta}) = \mathbf{F}\boldsymbol{\Lambda}_y(\boldsymbol{\theta})\mathbf{F}^H$. Since $\mathbf{R}_{cy}(\boldsymbol{\theta})$ is right circulant, the eigenvector matrix is $\mathbf{F} = M^{-1/2}\{e^{-j\frac{2\pi}{M}ik},\, i,j = 0\ldots,M-1\}$, and the eigenvalue matrix $\boldsymbol{\Lambda}_y(\boldsymbol{\theta}) = \text{diag}\,[\lambda_0(\boldsymbol{\theta}),\ldots, \lambda_{M-1}(\boldsymbol{\theta})]$ is the discrete Fourier transform of the first row of $\mathbf{R}_{cy}(\boldsymbol{\theta})$ [8].

Replacing $\mathbf{R}_y(\boldsymbol{\theta})$ with $\mathbf{R}_{cy}(\boldsymbol{\theta})$ in $\Lambda(\mathbf{Y}|\boldsymbol{\theta})$, the objective function $L(\mathbf{Y}|\boldsymbol{\theta})$ is equivalent to

$$\begin{aligned} L(\mathbf{Y}|\boldsymbol{\theta}) &= \text{const} - M^{-1}\left[\ln|\mathbf{R}_{cy}(\boldsymbol{\theta})| + \mathbf{Y}^H \mathbf{R}_{cy}^{-1}(\boldsymbol{\theta})\mathbf{Y}\right] && (5)\\ &= \text{const} - M^{-1}\left[\ln|\boldsymbol{\Lambda}_y(\boldsymbol{\theta})| + \left\{\mathbf{F}^H\mathbf{Y}\right\}^H \boldsymbol{\Lambda}_y^{-1}(\boldsymbol{\theta})\left\{\mathbf{F}^H\mathbf{Y}\right\}\right] && (6)\\ &= \text{const} - M^{-1}\left(\sum_{i=0}^{M-1}\ln\lambda_i(\boldsymbol{\theta}) + \sum_{i=0}^{M-1}\frac{\tilde{\lambda}_i}{\lambda_i(\boldsymbol{\theta})}\right), && (7)\end{aligned}$$

where $\tilde{\lambda}_i$ is the periodogram of sequence $\{Y_i\}_{i=1}^{i=M}$ computed at $F_i = i/M$, for $i = 0,\ldots,M-1$, and $\lambda_i(\boldsymbol{\theta}) = \lambda(F_i,\boldsymbol{\theta})$, with (we assume that $M > 2m+1$)

$$\lambda(F,\boldsymbol{\theta}) = \sum_{k=0}^{M-1} R_{cy}(kT_s,\boldsymbol{\theta})e^{-j2\pi Fk} = \sum_{k=-\infty}^{\infty} R_y(kT_s,\boldsymbol{\theta})e^{-j2\pi Fk}. \tag{8}$$

The last sum in (8) is, by definition, the *discrete time Fourier transform* (DTFT) of $R_y(kT_s, \boldsymbol{\theta})$.

The estimation criterion (4), with $\Lambda(\mathbf{Y}|\boldsymbol{\theta})$ replaced by $L(\mathbf{Y}|\boldsymbol{\theta})$, leads to an estimate herein denoted by $\widehat{\boldsymbol{\theta}}$ and termed *approximate maximum likelihood* (AML). It corresponds to a generalization of Whittle's ML approximation method [7], and is significantly more manageable than the exact log-likelihood function while behaving (asymptotically) in the same way. Related results are found in [6]. This work proposes the application of the concepts just stated to SM estimation; this makes sense if the AML estimate $\widehat{\boldsymbol{\theta}}$ is a *good* substitute for $\widehat{\boldsymbol{\theta}}^{ml}$. Results presented ahead provide insight into this matter.

A. Asymptotic Properties

Result 1: Let $\{Y_t\}$ be a strictly stationary zero-mean complex Gaussian sequence, with covariance $E[Y_{t+k}Y_t^*] = R_y(kT_s, \boldsymbol{\theta}_0)$ such that $\boldsymbol{\theta}_0$ belongs to the closed compact set $\boldsymbol{\Theta}$ contained in an open set $S \subset \Re^p$, and $\sum_{k=-\infty}^{\infty} |k||R_y(kT_s, \boldsymbol{\theta})|^2 < \infty$ for $\boldsymbol{\theta} \in S$. The partial derivatives $\partial\lambda^{-1}(F, \boldsymbol{\theta})/\partial\theta_i$ are continuous for all components $\theta_i \in \boldsymbol{\theta}$, with $(F, \boldsymbol{\theta}) \in [-1/2, 1/2] \times S$. If $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in S$ and $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_2$, then $\lambda(F, \boldsymbol{\theta}_1) \neq \lambda(F, \boldsymbol{\theta}_2)$ for almost all F. Under these conditions, the AML estimate $\widehat{\boldsymbol{\theta}}(M)$ is strongly consistent: $\widehat{\boldsymbol{\theta}}(M) \to \boldsymbol{\theta}_0$, with probability one (w.p.1), as $M \to \infty$.

Result 2: Assume the hypotheses of Result 1 and that $\partial^2\lambda(F, \boldsymbol{\theta})/(\partial\theta_i\partial\theta_j)$ is continuous for all components θ_i of $\boldsymbol{\theta}$, with $(F, \boldsymbol{\theta}) \in [-1/2, 1/2] \times S$. Hence, the random vector $\boldsymbol{\xi}(M) = \sqrt{M}(\widehat{\boldsymbol{\theta}}(M) - \boldsymbol{\theta}_0)$ is asymptotically zero-mean Gaussian with covariance matrix $\boldsymbol{\Gamma}^{-1}(\boldsymbol{\theta}_0)$, where

$$\boldsymbol{\Gamma}(\boldsymbol{\theta}) = -\int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{\nabla\lambda(F, \boldsymbol{\theta})\nabla^T\lambda(F, \boldsymbol{\theta})}{\lambda^2(F, \boldsymbol{\theta})}\, dF \tag{9}$$

is the *normalized asymptotic information matrix*.

B. Bias

Result 3: Assume the hypotheses of Result 2 and that $\sum_{k=-\infty}^{\infty} |k||R_y(kT_s, \boldsymbol{\theta})| < \infty$ for $\boldsymbol{\theta} \in S$. Under these conditions, the estimator bias verifies

$$M\{\Gamma_{ij}(\boldsymbol{\theta}_0) - M^{-1}\varepsilon_{ij}(\boldsymbol{\theta}_0)\}E[(\hat{\boldsymbol{\theta}}(M) - \boldsymbol{\theta}_0)] = \{\varepsilon_i(\boldsymbol{\theta}_0)\} + O(M), \tag{10}$$

with

$$\varepsilon_{i_1 \ldots i_k}(\boldsymbol{\theta}) \quad = \quad \int_{-1/2}^{1/2} \beta(F, \boldsymbol{\theta}) \frac{\partial^k \lambda^{-1}(F, \boldsymbol{\theta})}{\partial\theta_{i_1} \ldots \partial\theta_{i_k}}\, dF, \tag{11}$$

$$\beta(F,\boldsymbol{\theta}) \;\;=\;\; \sum_{k=-\infty}^{\infty} |k| R_y(kT_s,\boldsymbol{\theta}) e^{-j2\pi Fk}, \tag{12}$$

where $O(M)$ is proportional to M^{-1} and $\boldsymbol{\Gamma}$ is the asymptotic information matrix given by (9).

Proofs of results 1 and 3 are carried out in [3]; result 2 is proved in [6, chapter II].

III. APPLICATION TO DOPPLER-SPREAD TARGETS

We are interested in the SM of a process with CF

$$R_x(\tau,\boldsymbol{\theta}) = \theta_0\, r_x(\theta_2 \tau) \exp(j2\pi\theta_1\tau), \tag{13}$$

where $r_x(0) = 1$, $r_x(\tau)$ is *a priori* known, $\theta_0 > 0$, $\theta_2 > 0$. The CF (13) is typical of *frequency-spread* targets as it is the case in all the applications mentioned in the introduction.

For the CF (13), the objective function (7) becomes

$$L(\mathbf{Y}|\boldsymbol{\theta}) = \text{const} - M^{-1}\left(\sum_{i=0}^{M-1} \ln \lambda_0(F_i) + \sum_{i=0}^{M-1} \frac{\tilde{\lambda}_i}{\lambda_0(F_i - \theta_1 T_s)}\right), \tag{14}$$

where $\lambda_0(F)$ is the DTFT of $R_y[kT_s,(\theta_0, 0, \theta_2)]$. To maximize (14) we propose the Newton-type iteration

$$\widehat{\boldsymbol{\theta}}^{k+1} = \widehat{\boldsymbol{\theta}}^{k} + \boldsymbol{\Gamma}^{-1}(\widehat{\boldsymbol{\theta}}^{k}) \nabla L(\mathbf{Y}|\widehat{\boldsymbol{\theta}}^{k}). \tag{15}$$

Compared with the Newton-Raphson iterative scheme, expression (15) uses matrix $-\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}}^{k})$ instead of $\mathbf{H}(\mathbf{Y}|\widehat{\boldsymbol{\theta}}^{k})$, the Hessian of $L(\mathbf{Y}|\widehat{\boldsymbol{\theta}}^{k})$. An informal justification for (15) is the following: under regularity conditions as those of Result 2, $\mathbf{H}(\mathbf{Y}|\widehat{\boldsymbol{\theta}}) = -\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}}) + \boldsymbol{\Delta}$, where $E\,\|\boldsymbol{\Delta}\|^2 = O(M)$ [3]. Hence, for $\|\widehat{\boldsymbol{\theta}}^{k} - \widehat{\boldsymbol{\theta}}\| \simeq 0$, it follows that $\mathbf{H}(\mathbf{Y}|\widehat{\boldsymbol{\theta}}^{k}) \simeq -\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}}^{k}) + \boldsymbol{\Delta}'$, with $E\,\|\boldsymbol{\Delta}'\|^2 = O(M)$. It should be noted, however, that $\mathbf{H}(\mathbf{Y}|\widehat{\boldsymbol{\theta}}^{k})$ may largely diverge from $\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}}^{k})$, as greater displacements $\|\widehat{\boldsymbol{\theta}}^{k} - \widehat{\boldsymbol{\theta}}\|$ are considered. Nevertheless, the iterative scheme (15) has shown to be much more robust than the pure Newton-Raphson: the reason is that, even for moderate displacements of $\|\widehat{\boldsymbol{\theta}}^{k} - \widehat{\boldsymbol{\theta}}\|$, $L(\mathbf{Y}|\widehat{\boldsymbol{\theta}}^{k})$ exhibits large deviations from the parabolic shape; it often happens that $\mathbf{H}(\mathbf{Y}|\widehat{\boldsymbol{\theta}}^{k})$ has very small (negative), or even positive eigenvalues, whereas $\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}}^{k})$ keeps close to $\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})$. Another advantage of (15) is that $\boldsymbol{\Gamma}^{-1}(\widehat{\boldsymbol{\theta}}^{k})$ can be computed off-line.

The choice of the starting point $\widehat{\boldsymbol{\theta}}^{0}$ is crucial; notice that (14) depends on the periodogram $\tilde{\lambda}_i$, upon which the PB estimate is built:

$$\widehat{m}_k^{pb} = \frac{1}{M} \sum_{i=-M/2}^{M/2-1} \left(\frac{if_s}{M}\right)^k [\tilde{\lambda}_i - N_i], \tag{16}$$

where M is assumed even and N_i is the DTFT of $R_n(\tau)$. Furthermore, by applying (1) to (13), we obtain

$$\theta_0 = m_0 \quad (17)$$

$$\theta_2 = 2\pi\sigma_x(S_x)\left[-r_r^{(2)}(0) - \left(r_i^{(1)}(0)\right)^2\right]^{-1/2} \quad (18)$$

$$\theta_1 = \frac{m_1(S_x)}{m_0(S_x)} - \left(\frac{\theta_2}{2\pi}\right) r_i^{(1)}(0), \quad (19)$$

where $r_i(\tau)$ and $r_r(\tau)$ are the real and imaginary parts of $r_x(\tau)$, respectively. Therefore, concerning the starting point $\widehat{\boldsymbol{\theta}}^0$ of the Newton-type algorithm, we proceed as follows: (a) compute $(\widehat{m}_0^{pb}, \widehat{m}_1^{pb}, \widehat{m}_2^{pb})$ according to (16); (b) use equations (17), (18), and (19) to obtain $\widehat{\boldsymbol{\theta}}^0 = (\widehat{\theta}_1^0, \widehat{\theta}_2^0, \widehat{\theta}_3^0)$.

IV. EXAMPLE: DOPPLER WEATHER RADAR DATA

In this section we take $r_x(\tau) = \exp(-2\pi^2\tau^2)$ and $R_n(iT_s) = N_0\delta(i)$. This is typical of weather radar data [13]. Hence, it follows that

$$\lambda_0(F) = \frac{\theta_0}{\sqrt{2\pi}(\theta_2 T_s)} \sum_{k=-\infty}^{\infty} \exp\left(-\frac{1}{2}\left[\frac{1}{(\theta_2 T_s)}(F-k)\right]^2\right) + N_0. \quad (20)$$

In all the results presented the sample size is $M = 128$. For each value of $\boldsymbol{\theta}$, the iterative scheme (15) was applied over one hundred independent data sets, using the stop rule $|(\theta_0^k - \theta_0^{k-1})/\theta_0^k| < 10^{-4}$, $|(\theta_1^k - \theta_1^k)| < 10^{-4}T_s^{-1}$, and $|(\theta_2^k - \theta_2^{k-1})/\theta_2^k| < 10^{-4}$. The Newton-type scheme converged on all runs (2100), thus confirming the robustness above stated. The mean number of iterations, although function of $\boldsymbol{\theta}_0$, was never greater than five, leading to $O(M)\log M$ mean number of floating point operations per estimate.

Solid lines in Fig. 1 and Fig. 3 plot the normalized asymptotic CRB (standard deviation) of parameters $\theta_1 T_s$ and $\theta_2 T_s$ as function of $\theta_2 T_s$. The curves are parameterized by the SNR. The points incide the circles represent sample standard deviations obtained by running the proposed procedure. Within the uncertainty associated with the sample standard deviation (roughly 0.1%), these estimates exhibit a standard deviation very close to the asymptotic CRB.

Fig. 2 plots the ratio between the sample variance and the normalized asymptotic CRB (VAR$_0$) of $\theta_1 T_s$, resulting from Monte Carlo simulation (100 runs per point), for five different estimators. The AR(i) estimator is implemented by applying the SM definition to the spectrum $\widehat{S}_x$ estimated by assuming an *autoregressive* process of order i. The first thing to note is that the PP and AR(1) estimators have the same performance (roughly, 1.5 times the CRB); this is shown in [3]. The AR(3) and PB estimators seriously degrade their performance for $\theta_2 T_s > 0.175$. It is interesting to note that the AR(3) estimator, compared with the AR(1), yields higher variance for $\theta_2 T_s > 0.175$. The justification is that the spectrum associated to the AR(1) estimator is symmetric around $\theta_1 T_s$, this not being true for the AR(3). Therefore, the AR(1) conveys more *a priori* information about the underlying spectral shape.

Concerning bias, Result 3 led us to the conclusion that $\widehat{\theta}_1$ is unbiased whereas $\widehat{\theta}_0$ and $\widehat{\theta}_2$ are biased. Fig. 4 plots the relative bias of $\widehat{\theta}_2$, given by (10), for different SNR's. The points inside the circles show the relative sample bias obtained by Monte Carlo simulations. The accordance between the theoretical curve and the simulation results is evident.

Fig. 5 plots the relative sample bias of $\widehat{\theta}_2 T_s$ for SNR $= 10dB$, drawn from the Monte Carlo simulation, showing that the AML estimator is clearly the best. The PB estimator exhibits two kinds of bias: (a) due to the mean value of the periodogram, which is the convolution of the true spectrum with a Bartlet window (this bias, which tends asymptotically to zero, affects only the estimates at low spectral widths, and can be minimized by convolving the periodogram with an appropriate window); (b) due to aliasing and independent of the sample size. The bias pointed in (b), which can only be avoided by choosing a correct value of T_s, affects also the AR(i) estimators regardless of i. On the contrary, the AML presents no bias at high SNR, despite the referred aliasing.

V. Conclusions

This paper addressed the estimation of spectral moments of Gaussian processes immersed in Gaussian noise. An estimator exploiting a Wittle's type objective function, herein termed *approximate maximum likelihood* (AML), was introduced. Results stating its strong consistency and asymptotical efficiency were presented. A relevant feature of the AML estimator is that it depends on the periodogram, upon which the well known periodogram based (PB) estimator is built. This fact was exploited, yielding a Newton-type iterative algorithm initialized with the PB estimate. Furthermore, the Hessian matrix, needed in the Newton-Raphson method, was replaced with the asymptotic Fisher information (which can be computed off-line), leading to a highly robust scheme. As a practical example, a Gaussian shaped spectrum, typical of weather radar data, was considered. For a sample size $M = 128$, SNR $< 10dB$, and spectral width $\theta_2 T_s \in [0.05, 0.3]$, the proposed estimators is, nearly, unbiased and efficient. Comparisons with the periodogram based, the pulse pair, and the autoregressive based estimators are, in all cases, favorable to the AML.

References

[1] J.M.B. Dias and J.M.N. Leitão , ''Maximum likelihood estimation of spectral moments at low signal to noise ratios'', *in Proceedings of ICASSP'93*, vol. IV, pp. 149--152, Minneapolis, April 1993.

[2] J.M.B.Dias and J. M.N. Leitão , ''Nonparametric estimation of spectral moments in weather radar'', *in Advanced Weather Radar Systems - COST'75*, Brussels, 1994. Kluwer Academic Publishers.

[3] J.B. Dias, *Estimation of Spectral Moments: Conceptual and Computational Aspects*, PhD thesis, Instituto Superior Tecnico, Lisbon, 1994, (In Portuguese).

[4] D. S. Zrnic, ''Estimation of spectral moments for weather echos'', *IEEE Trans.Geosci. Electron.*, vol. GE-17, pp. 113--128, Oct. 1979.

[5] K. S. Miller and M. M. Rochewarger, ''A covariance approach to spectral moment estimation'', *IEEE Trans. Infom. Theory*, vol. IT-18, pp. 588--596, Sep. 1972.

[6] K. Dzhparidze, *Parameter Estimation and Hypothesis Test in Spectral Analysis of Stationary Time Series*, Spriger-Verlag, New York, 1985.

[7] P. Whittle, ''Estimation and information in stationary time series'', *Arkiv Matematick*, vol. B-2, pp. 423--434, 1953.

[8] L. L. Scharf, *Statistical Signal Processing. Detection, Estimation and Time Series Analysis*, Addison-Wesley, New York, 1991.

[9] J. P. Burg and D. G. Luenberger, ''Estimation of structured covariance matrices'', *Proc. IEEE*, vol. 70, pp. 963--974, 1982.

[10] D. B. Williams and D. H. Johnson, ''Robust estimation of structured covariance matrices'', *IEEE Trans. Signal Processing*, vol. 41, pp. 2891-2906, 1993.

[11] R. Kumaresan and L. L. Scharf, ''An algorithm for pole-zero modeling and spectral analysis'', *IEEE Trans. Signal Processing*, vol. 34, pp. 637--640, June 1986.

[12] M.I. Miller and D.L. Snyder, ''The role of likelihood and entropy in incomplete data problems: application to estimating point processes intensity and toeplitz constrained covariances'', *IEEE Trans. Signal Processing*, vol. 75, pp. 892--907, July 1987.

[13] 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.

List of figure captions

Figure 1: The solid lines plot the normalized asymptotic CRB (standard deviation) of the first spectral moment (mean frequency). The points inside the circles show the sample standard deviation (100 Monte Carlo runs per point) of the AML estimator.

Figure 2: Ratio between the sample variance (100 Monte Carlo runs per point) and the normalized asymptotic CRB (VAR_0) of the first spectral moment (mean frequency), for five different estimators.

Figure 3: The solid lines plot the normalized asymptotic CRB (standard deviation) of the second centered spectral moment (spectral width). The points inside the circles show the sample standard deviation (100 Monte Carlo runs per point) of the AML estimator.

Figure4: The solid lines plot the relative bias of the second centered spectral moment (spectral width) according to Result 3. The points inside the circles represent the sample bias (100 Monte Carlo runs per point) of the AML estimator for SNR = 10dB.

Figure5: Relative sample bias (100 Monte Carlo runs per point) of the second centered spectral moment (spectral width), for five estimators.

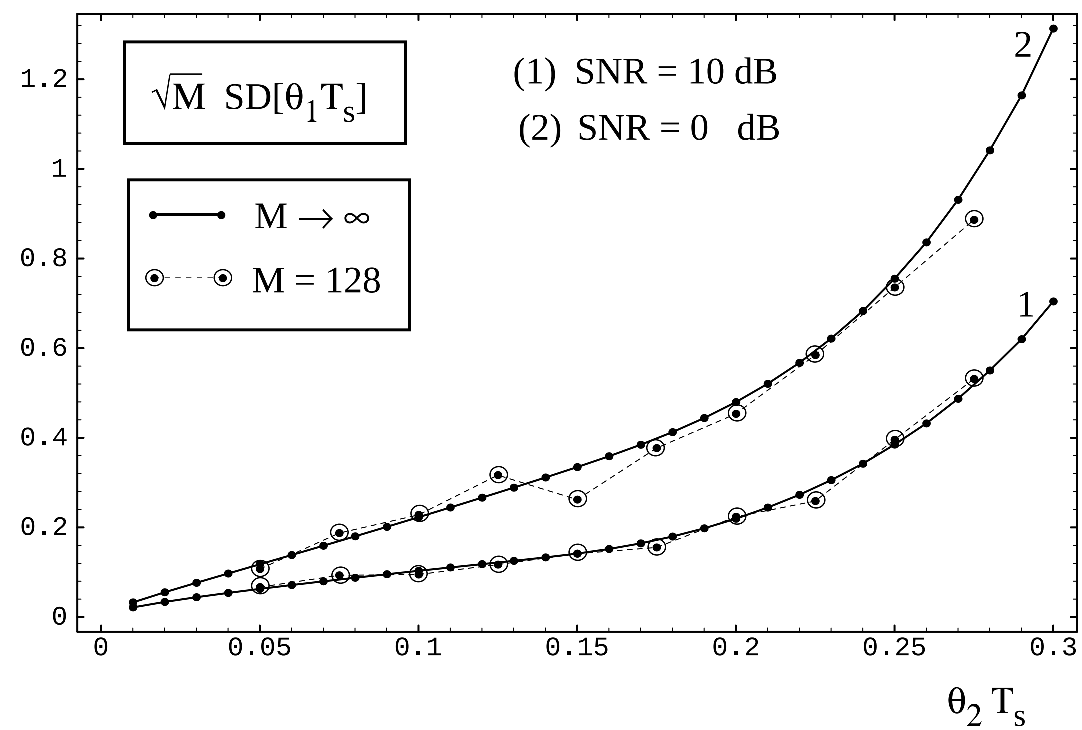

Fig. 1. The solid lines plot the normalized asymptotic CRB (standard deviation) of the first spectral moment (mean frequency). The points inside the circles show the sample standard deviation (100 Monte Carlo runs per point) of the AML estimator.

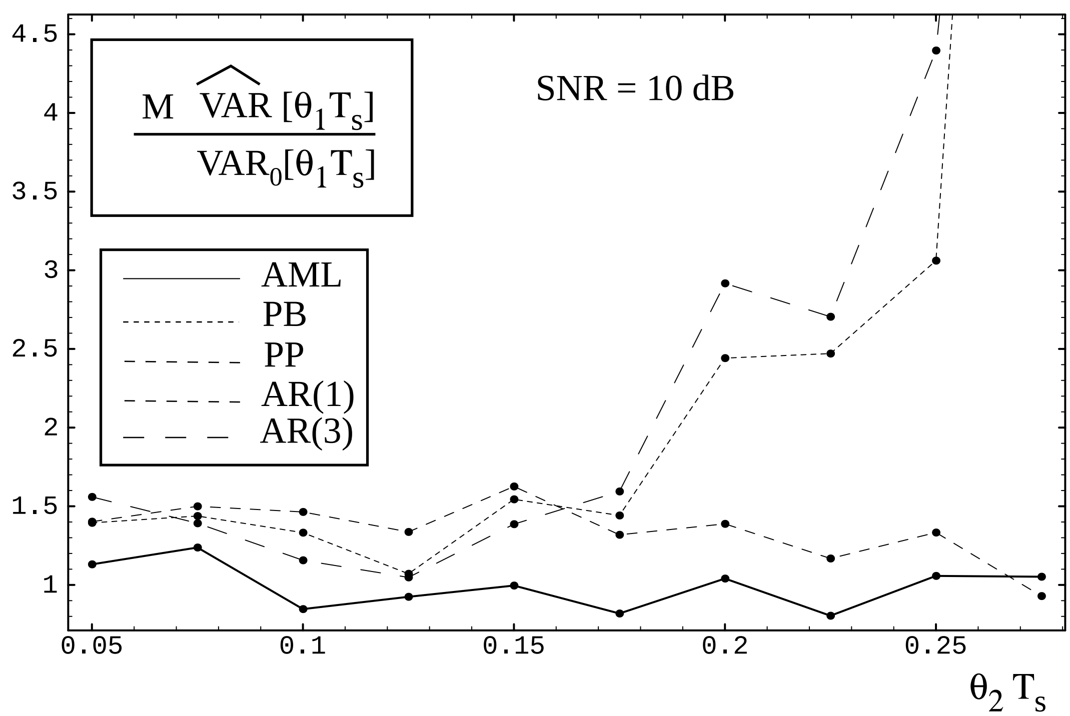

Fig. 2. Ratio between the sample variance (100 Monte Carlo runs per point) and the normalized asymptotic CRB (VAR_0) of the first spectral moment (mean frequency), for five different estimators.

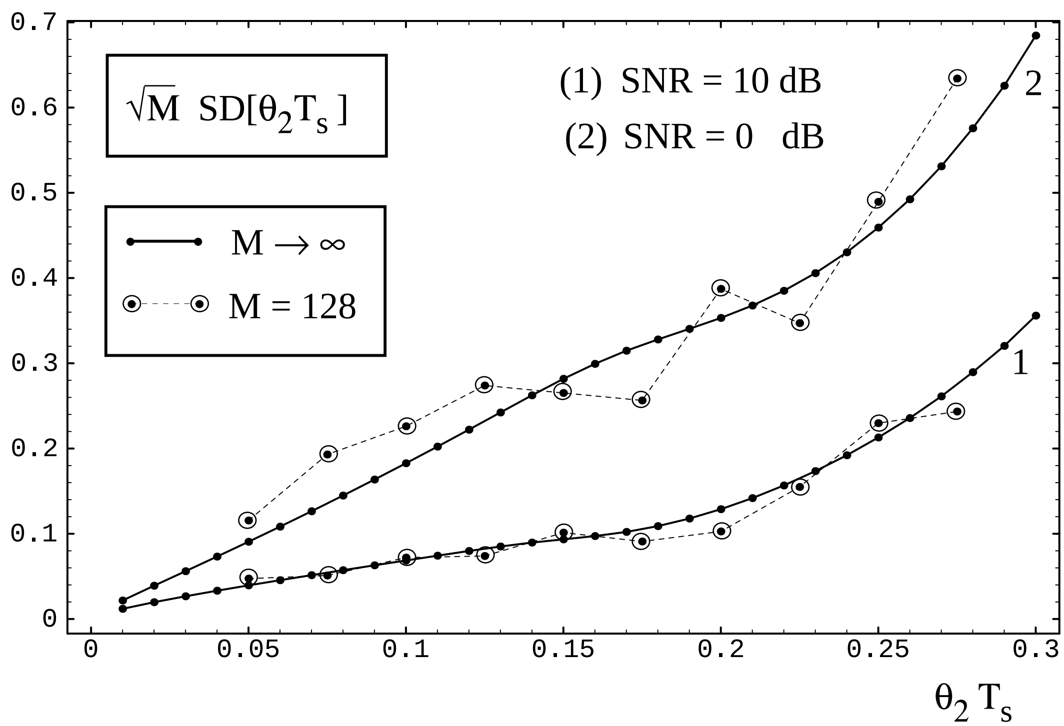

Fig. 3. The solid lines plot the normalized asymptotic CRB (standard deviation) of the second centered spectral moment (spectral width). The points inside the circles show the sample standard deviation (100 Monte Carlo runs per point) of the AML estimator.

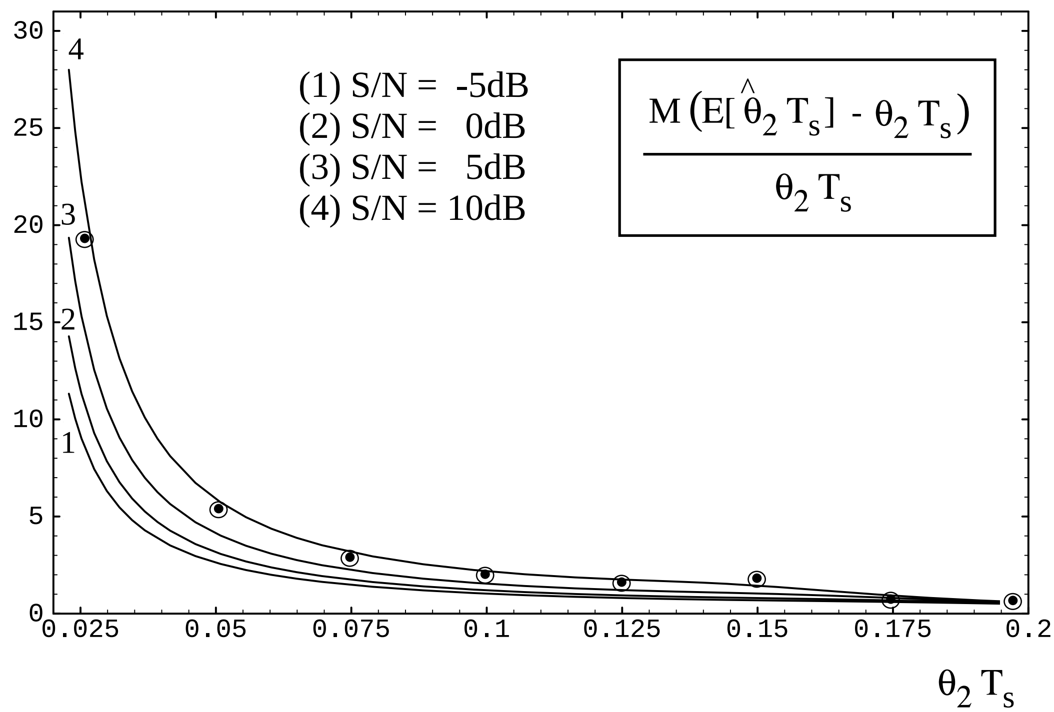

Fig. 4. The solid lines plot the relative bias of the second centered spectral moment (spectral width) according to Result 3. The points inside the circles represent the sample bias (100 Monte Carlo runs per point) of the AML estimator for SNR = 10dB.

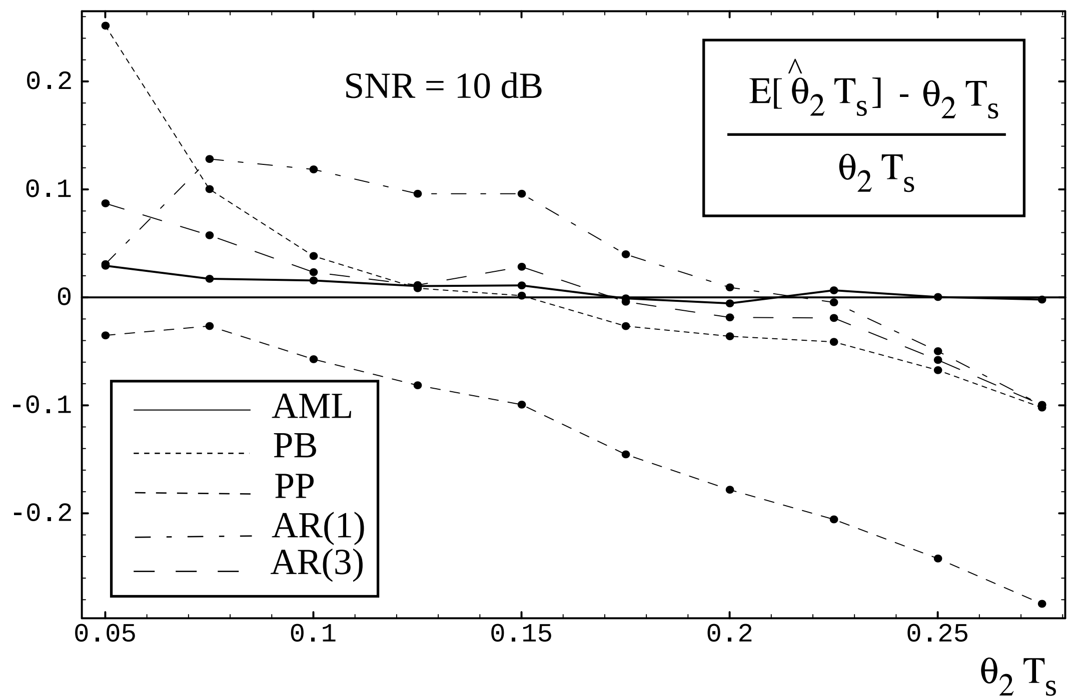

Fig. 5. Relative sample bias (100 Monte Carlo runs per point) of the second centered spectral moment (spectral width), for five estimators.