MAXIMUM LIKELIHOOD ESTIMATION OF AR MODULATED SIGNALS

MOUNIR GHOGHO
National Polytechnics Institute of Toulouse, LEN7/GAPSE, France
email: ghogho@len7.enseeiht.fr

ABSTRACT

The desired signal is embedded in both multiplicative and additive noises. The multiplicative noise is modeled by a Gaussian AR process. Closed forms expressions are derived for the finite-sample Cramer-Rao bound and for the maximum likelihood estimator. A cyclic approach is used to initialize the maximum likelihood algorithm when the signal is a harmonic.

1 INTRODUCTION

Multiplicative noise causes serious difficulties in many signal processing problems for radar transmission and fading communication channels. White and colored multiplicative noises have been considered with increasing interest. Consider a deterministic signal c_t with a random time-varying amplitude and its corrupted noise version

$$\begin{aligned} x_t &= s_t c_t = (\mu + y_t) c_t \\ z_t &= x_t + \nu_t \qquad t = 0, ..., N-1 \end{aligned} \tag{1}$$

where μ is the mean of s_t, y_t is a zero-mean stationary Gaussian sequence and ν_t is a white Gaussian sequence with variance σ_ν^2. The modulated signal $\{c_t \neq 0,\, t = 0, ..., N-1\}$ is assumed to be deterministic and parametrized by a constant but unknown vector $\underline{\lambda} = (\lambda_1, ..., \lambda_m)^T$. We will focus in particular on the sinusoidal signal

$$c_t = \cos(\omega t + \phi) \tag{2}$$

where $f = \omega/2\pi$ and ϕ are the carrier frequency and the initial phase respectively.

The basic idea in this paper is to model the random amplitude $\{y_t\}$ by a Gaussian stationary AR(p) process with parameters σ_ϵ^2 and $\underline{a} = (a_1, ..., a_p)^T$. This assumption can be justified by the two following properties: 1) the spectrum of any real-valued stationary process can be approximated to any desired precision by a convenient AR spectrum; 2) zero-mean Gaussian processes are completely defined by their spectrum. This paper develops the maximum likelihood estimate (MLE) and the Cramer-Rao lower bound (CRB) for the model parameters. For c_t given by (2), the MLE initialization is provided by a cyclic approach which exploits the cyclostationarity properties of the signal.

2 CRAMER-RAO BOUNDS (CRB)

It is well known that the covariance matrix of any unbiased estimator cannot be lower than the inverse of the Fisher information matrix (FIM) known as the Cramer-Rao bound (CRB)

$$CRB(\underline{\theta}) = \mathbf{J}^{-1}_{\underline{\theta},\underline{\theta}} = E\left[\left(\frac{\partial L(\underline{\theta})}{\partial \underline{\theta}}\right)\left(\frac{\partial L(\underline{\theta})}{\partial \underline{\theta}}\right)^T\right]^{-1} \tag{3}$$

$L(\underline{\theta})$ being the log-likelihood function of the parameter vector describing the signal model. To satisfy the regularity conditions for the existence of the CRB, c_t is assumed differentiable with respect to $\underline{\lambda}$.

2.1 Multiplicative Noise Source

For a fixed realization $\underline{X} = (x_0, ..., x_{N-1})^T$ of length $N > p$, the log likelihood function, corresponding to $\underline{\theta} = (\sigma_\epsilon^2, \underline{a}^T, \underline{\lambda}^T)^T$, is given by

$$\begin{aligned} L(\underline{\theta}) &= -\frac{1}{2}\left[\ln\left((2\pi\sigma_\epsilon^2)^N \det \mathbf{R}_N\right) + \ln\det \mathbf{C}^2\right] \\ &\quad -\frac{1}{2}\left((\mathbf{C}^{-1}\underline{X} - \mu\underline{1}_N)^T \mathbf{R}_N^{-1} (\mathbf{C}^{-1}\underline{X} - \mu\underline{1}_N)\right)/\sigma_\epsilon^2 \end{aligned} \tag{4}$$

$\sigma_\epsilon^2 \mathbf{R}_N$ denotes the covariance matrix of $(y_0, ..., y_{N-1})^T$, $\underline{1}_N$ is the $(N \times 1)$ vector whose elements are 1 and $\mathbf{C} = diag\{c_0, ..., c_{N-1}\}$. The inverse of $\mathbf{R}_N$ is given by the well-known Gohberg-Semencul formula

$$\mathbf{R}_N^{-1} = \mathbf{D}_1\mathbf{D}_1^T - \mathbf{D}_2\mathbf{D}_2^T \tag{5}$$

where $\mathbf{D}_1$ and $\mathbf{D}_2$ are lower triangular Toeplitz matrices whose (i,j)-th elements are

$$d_1(i,j) = \begin{cases} 1, & i = j \\ a_{i-j}, & i > j \\ 0, & i < j \end{cases} ;\; d_2(i,j) = \begin{cases} a_{N-i+j}, & i \geq j \\ 0, & i < j \end{cases}$$

Using (5), it has been shown for the pure AR case that the quadratic form in the observations can be expressed as a quadratic form in the autoregressive parameters [2]. Applying this result to the vector $(\mathbf{C}^{-1}\underline{X} - \mu\underline{1}_N)$ yields to

$$(\mathbf{C}^{-1}\underline{X} - \mu\underline{1}_N)^T \mathbf{R}_N^{-1} (\mathbf{C}^{-1}\underline{X} - \mu\underline{1}_N) = \sum_{i,j=0}^{p} a_j Q_{ij} a_j \tag{6}$$

where $a_0 = 1$ and for $i + j \leq N-1$

$$\begin{aligned} Q_{ij} &= \sum_{t=0}^{N-1-i-j} (x_{t+i}c_{t+i}^{-1} - \mu)(x_{t+j}c_{t+j}^{-1} - \mu) \\ &= -Q_{N-1-i,N-1-j} \end{aligned} \tag{7}$$

Let $\mathbf{G}_k,\ \mathbf{H}_{k,l},\ k,l=1,...,m$ be the $(p+1)\times(p+1)$ symmetric matrices whose elements are given by

$$g_k(i,j) = \sum_{t=0}^{N-1-i-j} \frac{\partial \ln|c_{t+i}c_{t+j}|}{\partial\lambda_k}, \quad i,j=0,...,p \quad (8)$$

$$h_{k,l}(i,j) = \sum_{t=0}^{N-1-i-j} \frac{\partial \ln|c_{t+i}|}{\partial\lambda_k}\frac{\partial \ln|c_{t+j}|}{\partial\lambda_l} \quad (9)$$

$\widetilde{\mathbf{G}}_k$ and $\widetilde{\mathbf{R}}_{p+1}$ are the $p\times(p+1)$ matrices formed by the p last rows of the matrices $\mathbf{G}^k$ and $\mathbf{R}_{p+1}$.

Proposition 1. *The exact FIM $\mathbf{J}_{\underline{\theta},\underline{\theta}}$ is given as follows: the components $\mathbf{J}_{\mu,\mu}$, $\mathbf{J}_{\mu,\sigma_\epsilon^2}$, $\mathbf{J}_{\mu,\underline{a}}$, $\mathbf{J}_{\sigma_\epsilon^2,\sigma_\epsilon^2}$, $\mathbf{J}_{\sigma_\epsilon^2,\underline{a}}$ and $\mathbf{J}_{\underline{a},\underline{a}}$ are the same as in the pure AR case, i.e. without modulation; the other elements of $\mathbf{J}_{\underline{\theta},\underline{\theta}}$ are:*

$$\begin{aligned}
\mathbf{J}_{\mu,\lambda_k} &= \frac{1}{2}\frac{\mu}{\sigma_\epsilon^2}\left(1,\underline{a}^T\right)^T \ \mathbf{G}_k\ \left(1,\underline{a}^T\right)\\
\mathbf{J}_{\sigma_\epsilon^2,\lambda_k} &= \frac{1}{2\sigma_\epsilon^2} g_k(0,0)\\
\mathbf{J}_{\underline{a},\lambda_k} &= -\left(\widetilde{\mathbf{R}}_{p+1}\cdot *\widetilde{\mathbf{G}}_k\right)\left(1,\underline{a}^T\right)^T\\
\mathbf{J}_{\lambda_k,\lambda_l} &= \mu^2\left(1,\underline{a}^T\right)^T\ \mathbf{H}_{k,l}\ \left(1,\underline{a}^T\right)+2\,h_{k,l}(0,0)
\end{aligned}$$

where $k,l=1,...,m$ and *the operator* $(\cdot *)$ *denotes the element by element matrix multiplication. The matrix $\mathbf{J}_{\underline{\theta},\underline{\theta}}$ is completed by symmetry.*

Remark 1 *Note that if c_t is close to 0 for some values of t then the CRB of $\underline{\lambda}$ is close to 0 since $\mathbf{J}_{\underline{\lambda},\underline{\lambda}}$ tends to ∞. Indeed, if some elements of $\{c_t\}$ are 0, then $\underline{\lambda}$ can be exactly determined from the zero locations.*

2.2 Multiplicative and Additive Noise Sources

The parameter vector describing the signal z_t in (1) is $\underline{\theta}=(\sigma_\epsilon^2,\underline{a}^T,\underline{\lambda}^T,\sigma_\nu^2)^T$. The FIM elements can be written as [3]

$$\mathbf{J}_{\theta_k,\theta_l} = \frac{\partial \underline{M}^T}{\partial\theta_k}\mathbf{R}_z^{-1}\frac{\partial \underline{M}}{\partial\theta_l} + \frac{1}{2}\ tr\left\{\mathbf{R}_z^{-1}\frac{\partial \mathbf{R}_z}{\partial\theta_k}\mathbf{R}_z^{-1}\frac{\partial \mathbf{R}_z}{\partial\theta_l}\right\} \quad (10)$$

where $\underline{M}^T$ and $\mathbf{R}_z$ is the mean vector and the covariance matrix of $\underline{Z}=(z_o,...,z_{N-1})^T$. In (10), tr denotes the trace operator. The processes x_t and ν_t are independent. Thus $\underline{M}^T$ and $\mathbf{R}_z$ are given by

$$\underline{M} = \mu\,(c_o,...,c_{N-1})^T \quad (11)$$

$$\mathbf{R}_z = \sigma_\epsilon^2\mathbf{C}\mathbf{R}_N\mathbf{C}+\sigma_\nu^2\mathbf{I}_N \quad (12)$$

where $\mathbf{I}_N$ is the N-dimensional identity matrix. The partial derivatives of $\underline{M}$ and $\mathbf{R}_z$ with respect to $\underline{\theta}$ are

$$\frac{\partial\underline{M}}{\partial\mu} = \frac{1}{\mu}\underline{M};\ \frac{\partial\underline{M}}{\partial\lambda_k}=\mu\frac{\partial\mathbf{C}}{\partial\lambda_k}\underline{1}_N \quad (13)$$

$$\frac{\partial\mathbf{R}_z}{\partial\sigma_\epsilon^2} = \mathbf{C}\mathbf{R}_N\mathbf{C};\ \frac{\partial\mathbf{R}_z}{\partial\sigma_\nu^2}=\mathbf{I}_N \quad (14)$$

$$\begin{aligned}
\frac{\partial\mathbf{R}_z}{\partial a_k} &= \sigma_\epsilon^2\mathbf{C}\frac{\partial\mathbf{R}_N}{\partial a_k}\mathbf{C}\\
&= -\sigma_\epsilon^2\mathbf{C}\mathbf{R}_N\frac{\partial\mathbf{R}_N^{-1}}{\partial a_k}\mathbf{R}_N\mathbf{C} \quad (15)
\end{aligned}$$

$$\frac{\partial\mathbf{R}_z}{\partial\lambda_k} = 2\sigma_\epsilon^2\mathbf{C}\mathbf{R}_N\frac{\partial\mathbf{C}}{\partial\lambda_k} \quad (16)$$

The other derivatives are identically zero. Using (5), explicit expressions of the derivatives of $\mathbf{R}_N^{-1}$ with respect to a_k are derived in [3]

$$\frac{\partial\mathbf{R}_N^{-1}}{\partial a_k} = \mathbf{Z}_k\mathbf{D}_1^T+\mathbf{D}_1\mathbf{Z}_k^T-\mathbf{Z}_{N-k}\mathbf{D}_2^T-\mathbf{D}_2\mathbf{Z}_{N-k}^T \quad (17)$$

where $\mathbf{Z}_k$ is the down shift matrix

$$z_k(i,j)=\begin{cases}1, & i-j=k\\ 0, & \text{otherwise}\end{cases} \quad (18)$$

Using the previous formula, a closed-form expression for $\mathbf{J}_{\underline{\theta},\underline{\theta}}$ is obtained.

Remark 2 *The method for the FIM computation in subsection 2.2 is also valid in the case where there is no additive noise ($\sigma_\nu^2=0$). However, proposition 1 gives a more compact form which is computationally much less demanding. Indeed, it only needs matrices of size $(p+1)\times(p+1)$ rather than $N\times N$.*

3 MAXIMUM LIKELIHOOD ESTIMATOR

Consider first the multiplicative noise source. For a fixed realization $\underline{X}=(x_0,...,x_{n-1})^T$, the approximate likelihood function is given by the conditional PDF of $(x_p,....,x_{n-1})^T$ given the initial values $x_0,...,x_{p-1}$

$$\begin{aligned}
L(\underline{\theta})\approx l(\underline{\theta}) &= \frac{1}{(2\pi\sigma_\epsilon^2)^{\frac{N-p}{2}}\left|c_p...c_{N-1}\right|}\exp\left\{-\frac{1}{2\sigma_\epsilon^2}\right.\\
&\left.\sum_{t=p}^{N-1}\left(\frac{x_t}{c_t}-\mu+\sum_{k=1}^{p}a_k\left(\frac{x_{t-k}}{c_{t-k}}-\mu\right)\right)^2\right\}
\end{aligned} \quad (19)$$

The (approximate) MLE $\widehat{\underline{\theta}}_{ML}$ of $\underline{\theta}$ is the one which maximizes $l(\underline{\theta})$ over a subset $\mathbf{\Theta}=\Theta_\lambda\times\Theta_a\times\Theta_{\sigma_\epsilon^2}$ of the multidimensional Euclidean space $\mathbb{R}^{p+m+1}$. Setting the partial derivative of $l(\underline{\theta})$ with respect to σ_ϵ^2 to 0 yields

$$\widehat{\sigma}_{\epsilon,ML}^2=\frac{1}{N-p}\sum_{t=p}^{N-1}\left(\frac{x_t}{c_t}-\mu+\sum_{k=1}^{p}a_k\left(\frac{x_{t-k}}{c_{t-k}}-\mu\right)\right)^2 \quad (20)$$

Substituting $\widehat{\sigma}_{\epsilon,ML}^2$ into (19) and dropping constant terms, we need now to maximize

$$\begin{aligned}
&J_1(\mu,\underline{a},\underline{\lambda}) = -\ln|c_p...c_{N-1}|-\tfrac{1}{2}(N-p)\\
&\ln\left[\frac{1}{N-p}\sum_{t=p}^{N-1}\left(\frac{x_t}{c_t}-\mu+\sum_{k=1}^{p}a_k\left(\frac{x_{t-k}}{c_{t-k}}-\mu\right)\right)^2\right]
\end{aligned} \quad (21)$$

3.1 Zero-Mean Case: $\mu=0$

$J_1(\underline{a},\underline{\lambda})\stackrel{\triangle}{=}J_1(0,\underline{a},\underline{\lambda})$ is quadratic with respect to $\underline{a}$. Thus, $\widehat{\underline{a}}_{ML}$ is given by the following orthogonal projection

$$\begin{aligned}
\widehat{\underline{a}}_{ML} &= \mathbf{P}_{\mathbf{X}./\mathbf{C}}(\underline{X}./\underline{C})\\
&= -\left[(\mathbf{X}./\mathbf{C})^T\ (\mathbf{X}./\mathbf{C})\right]^{-1}(\mathbf{X}./\mathbf{C})^T(\underline{X}./\underline{C})
\end{aligned} \quad (22)$$

where the operator $(./)$ denotes the element by element matrix division. $\underline{C}=(c_p,...,c_{N-1})^T$, $\mathbf{C}$ and $\mathbf{X}$ are the $(N-p)\times p$ Toeplitz matrices whose (i,j)-th elements are given by $c_{p+i-j-1}$ and $x_{p+i-j-1}$ respectively. If the parameter vector $\underline{\lambda}$ is known, then the MLE of σ_ϵ^2 and $\underline{a}$ are given by (20) and (22). If $\underline{\lambda}$ is unknown, its MLE is obtained by the argument of the supremum of

$$\begin{aligned}
&J(\underline{\lambda}) = -\ln|c_p...c_{N-1}|-\tfrac{1}{2}(N-p)\\
&\ln\left[\tfrac{1}{N-p}(\underline{X}./\underline{C})^T\left(\mathbf{I}-(\mathbf{X}./\mathbf{C})\mathbf{P}_{\mathbf{X}./\mathbf{C}}\right)(\underline{X}./\underline{C})\right]
\end{aligned} \quad (23)$$

where $\mathbf{I}$ is the $(N-p)$-dimensional identity matrix. In other words, the maximization of $L(\underline{\theta})$ with respect to $\underline{\theta}$ is equivalent to the maximization of $J(\underline{\lambda})$ with respect to $\underline{\lambda}$ only. The MLE of σ_ϵ^2 and $\underline{a}$ are obtained by replacing $\underline{\lambda}$ by $\widehat{\underline{\lambda}}_{ML}(=\arg\max\, J(\underline{\lambda}))$ in (20) and (22) respectively.

3.2 Non Zero-Mean Case: $\mu \neq 0$

Setting the partial derivative of $J_1(\mu, \underline{a}, \underline{\lambda})$ with respect to μ to 0 yields

$$\widehat{\mu}_{ML} = \frac{1}{(N-p)(1+\sum_{k=1}^{p} a_k)} \sum_{t=p}^{N-1} \left(\frac{x_t}{c_t} + \sum_{k=1}^{p} a_k \frac{x_{t-k}}{c_{t-k}} \right) \tag{24}$$

Substituting $\widehat{\mu}_{ML}$ into (21), the maximization criterion becomes then

$$\begin{aligned} J_2(\underline{a}, \underline{\lambda}) = - \ln |c_p ... c_{N-1}| - \tfrac{1}{2}(N-p) \\ \ln \left[\tfrac{1}{N-p} \sum_{t=p}^{N-1} \left(\tfrac{x_t}{c_t} - \mu_0 + \sum_{k=1}^{p} a_k \left(\tfrac{x_{t-k}}{c_{t-k}} - \mu_k \right) \right)^2 \right] \end{aligned} \tag{25}$$

where

$$\mu_k = \frac{1}{N-p} \sum_{t=p}^{N-1} \frac{x_{t-k}}{c_{t-k}}, \qquad k = 0, ..., p \tag{26}$$

The maximization criterion is quadratic with respect to $\underline{a}$ as in the zero-mean case. Thus,

$$\begin{aligned} \widehat{\underline{a}}_{ML} &= \mathbf{P}_{(\mathbf{X}./\mathbf{C}-\mathbf{1}\boldsymbol{\Lambda})} (\underline{X}./\underline{C} - \mu_0 \underline{1}) \\ &= - \left[(\mathbf{X}./\mathbf{C} - \mathbf{1}\boldsymbol{\Lambda})^T \, (\mathbf{X}./\mathbf{C} - \mathbf{1}\boldsymbol{\Lambda}) \right]^{-1} \\ &\quad (\mathbf{X}./\mathbf{C} - \mathbf{1}\boldsymbol{\Lambda})^T (\underline{X}./\underline{C} - \mu_0 \underline{1}) \end{aligned} \tag{27}$$

where $\boldsymbol{\Lambda} = diag\{\mu_1, ..., \mu_p\}$ and $\underline{1}$ is the $(N-p)$ vector and $\mathbf{1}$ is the $(N-p) \times p$ matrix whose elements equal 1. Hence, the maximization of $L(\underline{\theta})$ over the whole parameter vector $\underline{\theta}$ is equivalent to the maximization of

$$\begin{aligned} J(\underline{\lambda}) = - \ln |c_p ... c_{N-1}| - \tfrac{1}{2}(N-p) \ln \left[\tfrac{1}{N-p} (\underline{X}./\underline{C} \right. \\ \left. - \mu_0 \underline{1})^T \left(\mathbf{I} - (\mathbf{X}./\mathbf{C} - \mathbf{1}\boldsymbol{\Lambda}) \mathbf{P}_{(\mathbf{X}./\mathbf{C}-\mathbf{1}\boldsymbol{\Lambda})} \right) (\underline{X}./\underline{C} - \mu_0 \underline{1}) \right] \end{aligned} \tag{28}$$

with respect to $\underline{\lambda}$ only.

3.3 Influence of Additive Noise onto the MLE

The signal x_t is now contaminated by a Gaussian additive white noise. Let $SNR(t)$ define the instantaneous signal-to-noise ratio

$$SNR(t) = 10 \log \frac{c_t^2 E\{s_t^2\}}{\sigma_\nu^2} \tag{29}$$

Two approaches are possible to estimate the model parameters. First, the exact likelihood function can be maximized with respect to all the parameters since the noisy signal vector $\underline{Z} = (z_0, ..., z_{N-1})^T$ is Gaussian with mean and covariance function given by (11) and (12). This method is computationally rather expensive. An alternative approach ignores the additive noise and uses the MLE in subsections (3.1) and (3.2). Section 6 simulations show that the later approach provides an accurate estimate of the sinusoidal parameters, especially when the mean value of SNR: $\overline{SNR} = \frac{E\{s_t^2\}}{2\sigma_\epsilon^2}$ is not too low.

4 INITIALIZATION: CYCLIC APPROACH

This section focuses on the special case of harmonic signals (2). Since c_t is periodic in this case, the k-th moments of x_t are periodically time-varying. Cyclic cumulants with their consistent estimates have been proposed in [4]. AR processes belong to the class of mixing processes. Hence, these estimates are consistent and asymptotically normal. Using these statistics, consistent estimators of ω and ϕ have been proposed in [7].

4.1 Non Zero-Mean Case: $\mu \neq 0$

The mean of z_t is periodically time-varying $m_{1z}(t) = E\{z_t\} = \mu \cos(\omega t + \phi)$. The cyclic mean is given by [4]

$$\begin{aligned} c_{1z}(\alpha) &\triangleq \lim_{N\to\infty} \tfrac{1}{N} \sum_{t=0}^{N-1} m_{1z}(t) e^{-j\alpha t} \\ &= \mu \left[e^{i\phi} \delta(\alpha - \omega) + e^{-i\phi} \delta(\alpha + \omega) \right] \end{aligned} \tag{30}$$

where $\delta(.)$ is the Kronecker delta function. Replacing $c_{1z}(\alpha)$ by its consistent estimate

$$\widehat{c}_{1z}(\alpha) = \tfrac{1}{N} \sum_{t=0}^{N-1} z_t e^{-j\alpha t} \tag{31}$$

yields to

$$\widehat{\omega}^1 = \arg\max_{\alpha > 0} |\widehat{c}_{1z}(\alpha)| \, ; \; \widehat{\phi}^1 = \arg \left[\widehat{c}_{1z}(\widehat{\omega}^1) \right] \tag{32}$$

4.2 Zero-Mean Case: $\mu = 0$

If $\mu = 0$, the cyclic mean contains no information about the sinusoidal parameters. The instantaneous variance of z_t is given by

$$m_{2z}(t; 0) = m_{2y}(0) \cos^2(\omega t + \phi) + \sigma_\nu^2 \tag{33}$$

The cyclic variance of the zero-mean cyclostationary signal z_t is given by:

$$\begin{aligned} c_{2z}(\alpha; 0) := \lim_{N\to\infty} \tfrac{1}{N} \sum_{t=0}^{N-1} m_{2z}(t; 0) e^{-j\alpha t} = \\ \tfrac{1}{2} m_{2y}(0) \left[1 + e^{i2\phi} \delta(\alpha - 2\omega) + e^{-i2\phi} \delta(\alpha + 2\omega) \right] \end{aligned} \tag{34}$$

where $m_{2y}(0)$ is the variance of the stationary signal y_t. The cyclic variance can be estimated consistently from a single realization [4]

$$\widehat{c}_{2z}(\alpha; 0) = \tfrac{1}{N} \sum_{t=0}^{N-1} z_t^2 e^{-j\alpha t} \tag{35}$$

Consistent estimators of ω and ϕ are obtained by

$$\widehat{\omega}^1 = \frac{1}{2} \arg\max_{\alpha > 0} |\widehat{c}_{2z}(2\alpha; 0)| \, ; \; \widehat{\phi}^1 = \frac{1}{2} \arg \left[\widehat{c}_{2z}(2\widehat{\omega}^1; 0) \right] \tag{36}$$

The AR parameters can also be estimated to avoid local maxima in the MLE procedure. Indeed, even if $J(\underline{\lambda})$ is independent of $\underline{a}$, its estimate will allow the "goodness" of $\widehat{\underline{\lambda}}_{ML}$ to be checked via relations (22) and (27). The AR parameters estimates can be obtained either by using a special ARMA$(2p, 2p)$ representation of z_t [1][5] or by direct modeling of the demodulated signal [7].

5 COMPUTATIONAL ASPECT

The orthogonal projection formulas (22) and (27) can be computed without matrix inversion by using the efficient Levinson-Durbin algorithm. Because of the derivatives of $J(\underline{\lambda})$ with respect to $\underline{\lambda}$ are nonlinear, a numerical optimization technique has to be used. Local gradient methods, e.g. BFGS quasi-Newton optimization, can be used successfully if the deterministic signal c_t is slowly time-varying or contains only slow oscillations in the studied time interval. On the other hand, if c_t has fast oscillations, the maximum achieved by conventional gradient type algorithms may be a local rather than a global maximum even after a quite good initialization. For instance, consider the harmonic case with a relative high frequency. Simulations have shown that $J(\underline{\lambda})$ has many stationary points even in a small neighborhood of the initial point. To overcome this problem, we use the simulated annealing (SA) algorithm [6] which belongs to the class of stochastic optimization techniques. This algorithm is based on an analogy to the state evolution of material when it is cooled slowly. The initial estimator $\widehat{\omega}^1$ and $\widehat{\phi}^1$ of section 4 are asymptotically normal with variances $var(\widehat{\omega}^1)$ and $var(\widehat{\phi}^1)$ which decay as N^{-3} and N^{-1} respectively [7]. Thus, the search of the global supremum is restricted to the subspace $\mathbf{\Omega} = \mathbf{E} \cap \Theta_\lambda$ where $\mathbf{E}$ is given by

$$\mathbf{E} = \left\{ \underline{\lambda} \Big/ \left|\lambda_1 - \widehat{\omega}^1\right| \le \beta\sqrt{var\left(\widehat{\omega}^1\right)} \text{ and } \left|\lambda_2 - \widehat{\phi}^1\right| \le \beta\sqrt{var\left(\widehat{\phi}^1\right)} \right\} \quad (37)$$

β is a positive scalar which controls the area of $\mathbf{E}$. For instance, $\beta = 3$ asymptotically ensures that the true parameter vector belongs to $\mathbf{\Omega}$ with probability 0.99. When N is not large enough, β must to be taken greater than 3. In our simulations, β is set to 10. The annealing schedule is then as follows. The sequence of decreasing temperatures obeys the exponential law $T_k = T_0 \exp(-k^{1/2})$, $k = 1, ..., K$, where $T_0 = 200$ and K is the number of required temperatures. At each temperature T_k, the new candidate $\underline{\lambda}^k$ is obtained by perturbing the current parameter estimates $\underline{\lambda}^{k-1}$ by a random vector $\underline{\eta} = (\eta_1, \eta_2)^T$ generated by the bi-dimensional distribution

$$\varphi_k(\underline{\eta}) = \prod_{m=1}^{2} \frac{\xi}{1+\eta_m/T_k} \qquad \xi > 0 \quad (38)$$

Using Boltzman distribution, the new candidate is accepted with probability p:

$$p = \min\left\{1, \exp(\Delta/T_k\right\} \quad (39)$$

where $\Delta = J(\underline{\lambda}^k) - J(\underline{\lambda}^{k-1})$. That means that the new candidate is always accepted if $\Delta > 0$ and it is accepted stochastically with probability $\exp(\Delta/T)$ if $\Delta < 0$. Thus, at high temperatures, candidate with lower score can be accepted. That enables the algorithm to escape from local maxima.

6 SIMULATION RESULTS

This section illustrates the performance of the ML algorithm by Monte Carlo experiments. For instance, consider the problem of Doppler frequency estimation [1]. In this case, the received signal can be modeled by (1) and (2) with μ=0. Let $f = 0.18$, $\phi = 0.2\pi$, y_t is the low pass AR(2) process whose poles are located at $0.9e^{\pm j2\pi 0.08}$ and σ_ϵ^2=1. Figure 1 displays the MLE mean square errors (MSE) and the CRB for f versus the number of samples in the case of multiplicative noise source. Figure 2 shows the influence of additive noise (SNR=10dB) onto the MLE performance. 500 Monte Carlo runs were carried out in each case. The results depicted in figures 1 and 2 show the efficiency of the ML algorithm.

7 CONCLUSION

The maximum likelihood estimator and the Cramer-Rao bound are computed for the parameter vector of the multiplicative model . The maximization of the likelihood function over the whole parameter vector is shown equivalent to the maximization of a criterion with respect to the modulated signal parameters only. When this signal is a sinusoid, the conventional gradient type algorithms often converge to a local optimum. The simulated annealing algorithm has been used successfully to overcome this problem.

References

[1] O. Besson and F. Castanié, "On estimating the frequency of a sinusoid in autoregressive multiplicative noise," Signal processing, vol.30, no. 1, January 1993.

[2] G. E. P. Box and G. M. Jenkins, Time series analysis: Forecasting and control, Englewood, NJ: Prentice Hall, 1971.

[3] J. Francos and B. Friedlander, "Bounds for estimation of multicomponent signals with random amplitude and deterministic phase," *IEEE Trans. ASSP,* vol. 43, pp. 1161-1172, May 1995

[4] M. A. Gardner, "Exploitation of spectral redundancy in cyclostationary signals," *IEEE signal processing mag.,* pp. 14-36, Apr. 1991.

[5] M. Ghogho and B. Garel, "On AR modulated harmonics: CRB and Parameter estimation ," *in Proc. ICASSP-96,* (Atlanta), May 1996.

[6] S. Kirkpatrick and al., "Optimization by simulated annealing," Sci., vol. 220, pp. 671-680, 1983.

[7] G. Zhou and G. Giannakis, "Harmonics in Gaussian multiplicative and additive noise: Performance Analysis of Cyclic Estimators," *IEEE Trans. ASSP*, vol. 43, pp. 1445-1460, June 1995.

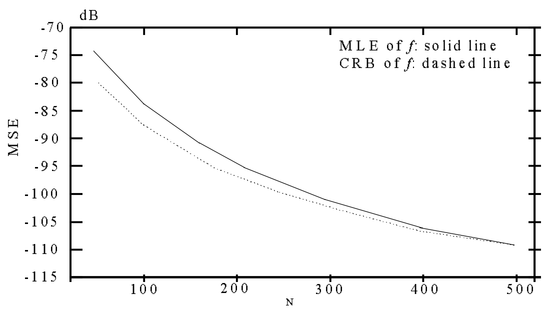

Fig. 1. Multiplicative noise source

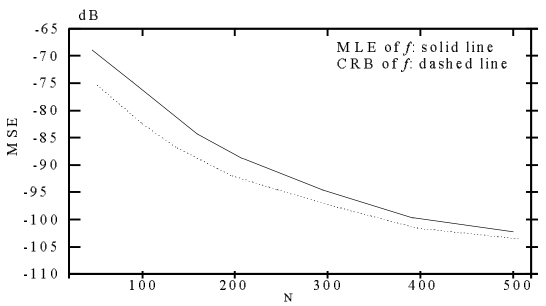

Fig. 2. Additive and multiplicative noise sources