THE CRAMÈR-RAO BOUND FOR THE ESTIMATION OF NOISY PHASE SIGNALS

A. M. Zoubir and A. Taleb

Australian Telecommunications Reasearch Institute
School of Electrical and Computer Engineering
Curtin University of Technology
GPO Box U1987, Perth WA 6845, Australia.
anisse@atri.curtin.edu.au, zoubir@ieee.org

ABSTRACT

This paper deals with noisy phase monocomponent signals in additive noise. This model is more appropriate for real world applications in particular for radar and communications. The problem is introduced and a maximum likelihood solution is proposed. Specifically, the Cramèr-Rao bound is explicitly derived and compared to the case of noise free phase.

1. INTRODUCTION

Non-stationary signals, i.e. signals whose spectral content varies with time, are very common in many signal processing applications. Examples include, but are not limited to, communications, sonar, radar, vibration, ultrasound imaging, speech and biological signals.

Many methods of analysis of non-stationary signals focus on nonparametric time-frequency distributions. However these suffer from the well known time-frequency resolution problem and an alternative to alleviate this problem lies in the use of parametric methods.

Monocomponent polynomial phase signals (pps), i.e. signals whose phase ϕ_t is modeled by a polynomial $\varphi_t = \sum_{i=0}^{n} a_i t^i$ and of constant amplitude A, have been first studied by Djurić and Kay [1]. Later, Peleg, Porat and Friedlander [2, 3] introduced the so-called polynomial phase transform as a tool for estimating the parameters of pps.

Most of the work cited above assumes that the signal phase fits perfectly to a polynomial of a certain degree. However, in a practical situation this is not true and a proper model should include a noise component on the phase, *i.e.*

$$\phi_t = \varphi_t + v_t \tag{1}$$

so that the received signal is

$$x_t = Ae^{j(\sum_{i=0}^{n} a_i t^i + v_t)} + n_t, \tag{2}$$

where v_t denote the noise on the phase, assumed to be real valued.

The motivation behind the noise component on the phase lies in the following

1. Errors due to an imperfect phase modeling by a certain basis function.
2. Phase noise in real signal sources, mainly oscillator phase noise, thermal noise, added amplifier noise prior or after frequency multiplication [4].
3. Noise due to the amplification of the received signal.

The noise in signal sources is highly relevant in both radar and communication systems. For instance in radar applications, spurious angle modulation of the transmitter will be transferred to the delayed returns from the ground clutter. As a consequence, when the Doppler frequency difference between target and clutter equals the offset frequency, the sidebands from the large clutter return may obscure the target signal.

This paper is concerned with an analysis of signals of the type given in Equation (2). Section 2 is dedicated to the mathematical signal model as well as to the main assumptions; identifiability issues will also be addressed. Section 3 discusses the maximum likelihood estimation while in Section 4 the Cramèr-Rao bounds are derived and compared to the noise free phase case. Finally, Section 5 concludes this paper.

2. SIGNAL MODEL AND ASSUMPTIONS

This section presents the signal model used throughout the paper. Notations and assumptions will be given.

2.1. Signal model

The observed signal model considered in this paper is

$$x_t = Ae^{j(\varphi_t + v_t)} + n_t, \tag{3}$$

where $t = 0, \ldots, T-1$ denotes the discrete time index. The phase function is modeled by,

$$\varphi_t = \sum_{i=0}^{n} a_i \xi_{ti}, \tag{4}$$

where the parameters $\{a_i,\ i = 0, \ldots, n\}$ are real valued, while $\{\xi_{ti},\ i = 0, \ldots, n\}$ is some arbitrary set of real valued basis functions. A usual example is the polynomial basis discussed in Section 1, where $\xi_{ti} = t^i$. We will denote by $\boldsymbol{\xi}$ the real $T \times (n+1)$ matrix whose entries are given by ξ_{ti}.

Unless specified, the noises v_t and n_t are assumed zero mean normal stationary and white. Furthermore, n_t is assumed to be complex circular.

This research is supported by the Australian Research Council through its Large Grant Scheme.

The estimation problem considered in this paper can be formulated as follows: Given a finite sample of x_t, estimate the amplitude of the signal A and its phase, *i.e.* the parameters $\{a_i,\ i = 0,\dots,n\}$. The noises variances can also be included into the estimation process since we do not assume their knowledge. However, unless in a detection context the latter are what is commonly known as nuisance parameters.

2.2. Note on the identifiability of the model

Before a complete study of an estimation procedure, one must always check for the identifiability of the model. To show the identifiability, it is necessary to address the uniqueness of the stochastic decomposition of the observed signal. In other words, does another decomposition of x_t as (3) exist for a different set of parameters? This ensures a one-to-one mapping from the parameters space to the distributions set. Let such other decomposition of x_t,

$$x_t = Ae^{j(\varphi_t+v_t)} + n_t = A'e^{j(\varphi'_t+v'_t)} + n'_t \tag{5}$$

where $n'_t,\ v'_t$ have the same distribution as v_t and n_t respectively but different variances.

Applying the expectation operator on both sides of (5) leads to:

$$E[x_t] = Ae^{j\varphi_t}e^{-\sigma_v^2} = A'e^{j\varphi'_t}e^{-\sigma_{v'}^2} \tag{6}$$

which implies that $e^{j\varphi_t} = e^{j\varphi'_t}$ and thus $\varphi_t = \varphi'_t \mod 2\pi$. This is the only uniqueness result which can be obtained concerning the phase. For the uniqueness of the signal phase parameters it is essential to force the parameters to lie within a certain range which depends on the basis functions used to model the phase. For instance, for a polynomial basis, it is then well known, that to guarantee the uniqueness of the phase parameters these must be restricted in the range $|a_i| \le \pi/i!,\ i = 0,1,\dots,n$. In the remainder of the paper, Ω will denote the range of parameters which guarantees such uniqueness.

Now we turn to the identifiability of the signal amplitude A. It is clear that if no phase noise is present, $\sigma_v^2 = \sigma_{v'}^2 = 0$, then $A = A'$. When $\sigma_v^2 \ne 0$, we resort to the second order moment of x_t,

$$\begin{aligned} E[(x_t - E[x_t])^2] &= A^2 e^{j2\phi_t}\{e^{-4\sigma_v^2} - e^{-2\sigma_v^2}\} \\ &= A'^2 e^{j2\phi'_t}\{e^{-4\sigma_{v'}^2} - e^{-2\sigma_{v'}^2}\} \end{aligned} \tag{7}$$

Combining this equation with (6) and after some algebraic manipulations one obtains $\sigma_v^2 = \sigma_{v'}^2$. Using the latter in (6) leads to $A = A'$. One can also deduce immediately that $\sigma_n^2 = \sigma_{n'}^2$ which concludes the identifiability issue.

3. THE MAXIMUM LIKELIHOOD ESTIMATOR

Consider the model of equation (3) and collect the observed data into $X^{(T)} = \{x_0, x_1, \dots, x_{T-1}\}$. Given the parameters vector $\theta = (A, a_0, \dots, a_n, \sigma_n^2, \sigma_v^2)^\dagger \in \Theta = \mathbb{R}^+ \times \Omega \times \mathbb{R}^{+2}$, the likelihood function writes as $\ell(X^{(T)}|\theta) = \prod_{t=0}^{T-1} p(x_t|\theta)$ since the noises v_t and n_t are assumed white. Moreover, since v_t and n_t are independent,

$$p(x_t|\theta) = \int_{\mathbb{R}} p_n\Big(x_t - A\exp(j\phi_t)\Big)p_v\Big(\phi_t - \sum_{i=0}^{n} a_i\xi_{ti}\Big)d\phi_t$$

where p_v and p_n denote respectively the probability density functions of the phase noise v_t and the additive noise n_t,

$$p_n(z) = \frac{1}{\pi\sigma_n^2}\exp(-\frac{|z|^2}{\sigma_n^2}), \quad z \in \mathbb{C} \tag{8}$$

$$p_v(u) = \frac{1}{\sqrt{2\pi}\sigma_v}\exp(-\frac{u^2}{2\sigma_v^2}), \quad u \in \mathbb{R} \tag{9}$$

The maximum likelihood estimator writes then as:

$$\hat{\theta} = (\hat{A}, \hat{a}_0, \dots, \hat{a}_n, \hat{\sigma}_n^2, \hat{\sigma}_v^2)^\dagger = \underset{\theta\in\Theta}{\operatorname{argmax}}\ \ell(X^{(T)}|\theta) \tag{10}$$

This estimator holds when the noise variances, σ_v^2 and σ_n^2, are unknown. If these are known, then they are simply replaced into the likelihood expression and the latter is optimized with respect to the remaining parameters. For instance, in some radar applications σ_n^2 can be assumed known by estimating the noise level when no pulse is present.

Consider a new data vector, namely

$$\{X^{(T)}, \Phi^{(T)}\} = \{x_0, x_1, \dots, x_{T-1}, \phi_0, \phi_1, \dots, \phi_{T-1}\}$$

We refer to this vector as the *complete data* while $\Phi^{(T)}$ is what is commonly known as *missing or unobserved data*. This data vector contains both the signal samples as well as the noisy phase samples, though the latter are actually unknown. The likelihood associated with the complete data is seen to be

$$\ell^c(X^{(T)}, \Phi^{(T)}|\theta) = \prod_{t=0}^{T-1} p_n\Big(x_t - A\exp(j\phi_t)\Big)p_v\Big(\phi_t - \sum_{i=0}^{n} a_i\xi_{ti}\Big) \tag{11}$$

The complete data likelihood can then be easily seen to belong to the exponential family. Furthermore we have:

$$\ell(X^{(T)}|\theta) = \int \ell^c(X^{(T)}, \Phi^{(T)}|\theta)\ d\phi_0 d\phi_1 \cdots d\phi_{T-1} \tag{12}$$

It is well known that the maximum likelihood estimates of the '*natural*' parameters in an exponential family are easily computed by the use of sufficient statistics. The natural form of the likelihood combined with the EM algorithm is used in [5] to derive a practical estimator of the parameters θ.

The score functions for this estimation problem are given by, $\dfrac{\partial L(X^{(T)}|\theta)}{\partial\theta_i}$ for $i = 1,\dots,n+3$, where $L(X^{(T)}|\theta)$ is the log-likelihood: $L(X^{(T)}|\theta) = \log \ell(X^{(T)}, \theta)$. For each θ_i one can see that

$$\frac{\partial L(X^{(T)}, \theta)}{\partial\theta_i} = E\left[\sum_{t=0}^{T-1} \frac{\partial \log p(x_t, \phi_t|\theta)}{\partial\theta_i}\,\middle|\, X^{(T)}, \theta\right] \tag{13}$$

Hence the score of θ is the conditional expectation of the complete likelihood score given the observed data. Now observe that,

$$\begin{aligned} \log p(x_t, \phi_t|\theta) = &-\frac{1}{\sigma_n^2}|x_t - A\exp(j\phi_t)|^2 - \log\sigma_n^2 \\ &- \frac{1}{2\sigma_v^2}\Big(\phi_t - \sum_{i=0}^{n} a_i\xi_{ti}\Big)^2 - \frac{1}{2}\log\sigma_v^2 \end{aligned} \tag{14}$$

From which it is straightforward to derive the score of the complete data likelihood and by the use of Eq. (13) the score of the likelihood,

$$\frac{\partial L(X^{(T)},\theta)}{\partial a_k} = \frac{1}{\sigma_v^2}\sum_{t=0}^{T-1}\xi_{tk}E[v_t|x_t,\theta], \quad k=0,\ldots,n \tag{15}$$

$$\frac{\partial L(X^{(T)},\theta)}{\partial A} = \frac{-1}{\sigma_n^2}\sum_{t=0}^{T-1}E[n_t^* e^{j\phi_t} + n_t e^{-j\phi_t}|x_t,\theta], \tag{16}$$

$$\frac{\partial L(X^{(T)},\theta)}{\partial \sigma_n} = \frac{2T}{\sigma_n^3}\left\{\frac{1}{T}\sum_{t=0}^{T-1}E[|n_t|^2|x_t,\theta] \ -\sigma_n^2\right\}, \tag{17}$$

$$\frac{\partial L(X^{(T)},\theta)}{\partial \sigma_v} = \frac{T}{\sigma_v^3}\left\{\frac{1}{T}\sum_{t=0}^{T-1}E\left[v_t^2|x_t,\theta\right] \ -\sigma_v^2\right\}, \tag{18}$$

4. CRAMÈR-RAO BOUNDS

The ij-th entry of the Fisher information matrix $I(\theta)$ is given by:

$$J_{ij}(\theta) = -E\left[\frac{\partial^2 L(X^{(T)},\theta)}{\partial\theta_i\partial\theta_j}\right] \tag{19}$$

For our problem, simple computations show that

$$E\left[\frac{\partial^2 L(X^{(T)},\theta)}{\partial a_i\partial a_j}\right] = -\frac{1}{\sigma_v^4}\sum_{t=0}^{T-1}E[E[v_t|x_t]^2]\xi_{ti}\xi_{tj} \tag{20}$$

$$E\left[\frac{\partial^2 L(X^{(T)},\theta)}{\partial A^2}\right] = -\frac{2}{\sigma_n^4}\sum_{t=0}^{T-1}E[E[\Re\{n_t e^{-j\phi_t}\}|x_t]^2] \tag{21}$$

$$E\left[\frac{\partial^2 L(X^{(T)},\theta)}{\partial a_i\partial A}\right] = \frac{-2A}{\sigma_n^4}\sum_{t=0}^{T-1}\xi_{ti}E\{\Im\{E[n_t^* e^{j\phi_t}|x_t]^2\}\} \tag{22}$$

We will only consider the Cramèr-Rao bound for the estimation of A and the phase parameters and we will assume that the nuisance parameters σ_v^2 and σ_n^2 are known. To make some developments to equations (20-22) we need the following lemma

Lemma 1 *Let $x_t = Ae^{j\phi_t} + n_t$, then for any stationary process q_t for which $E[q_t|x_t]$ exists, $E[q_t|x_t]$ is a* stationary process*. As a corollary the moments of $E[q_t|x_t]$ are independent of t.*

Proof: First notice that Ae^{jv_t} is stationary since v_t was assumed stationary. Furthermore since n_t is stationary Gaussian and circular, so is $n_t e^{-j\varphi_t}$ since $e^{-j\varphi_t}$ is a rotation in the complex plane. Hence $Ae^{jv_t} + n_t e^{-j\varphi_t} = x_t e^{-j\varphi_t}$ is a stationary process. Now, for any random process q_t, $E[q_t|x_t] = E[q_t|x_t e^{-j\varphi_t}]$ since multiplying by $e^{-j\varphi_t}$ represent a rotation in the complex plane which is one-to-one. Thus, if q_t is stationary and since $x_t e^{-j\varphi_t}$ is stationary then $E[q_t|x_t]$ is stationary. ■

By noticing that v_t, $n_t e^{-j\phi_t}$ are stationary processes, and by the help of Lemma 1, one deduces that the conditional moments in (20-22) are stationary processes. Therefore their moments can be factored out of the sum. Now observe that, by an integration by parts, we have

$$E[v_t|x_t] = j\frac{A\sigma_v^2}{\sigma_n^2}\{x_t^* E[e^{j\phi_t}|x_t] - x_t E[e^{-j\phi_t}|x_t]\} \tag{23}$$

Moreover, by noticing that $x_t^* E[e^{j\phi_t}|x_t]$ and $x_t E[e^{-j\phi_t}|x_t]$ have the same distribution, the previous equations become

$$E\left[\frac{\partial^2 L(X^{(T)},\theta)}{\partial a_i\partial a_j}\right] = \frac{4A^2}{\sigma_n^4}E[E[\Im\{x_t^* e^{j\phi_t}\}|x_t]^2]\sum_{t=0}^{T-1}\xi_{ti}\xi_{tj} \tag{24}$$

$$E\left[\frac{\partial^2 L(X^{(T)},\theta)}{\partial A^2}\right] = \frac{4TE[E[\Re\{n_t e^{-j\phi_t}\}|x_t]^2]}{\sigma_n^4} \tag{25}$$

$$E\left[\frac{\partial^2 L(X^{(T)},\theta)}{\partial a_i\partial A}\right] = 0 \tag{26}$$

We will denote $J(\sigma_v)$ the following partitioned Fisher information matrix

$$\begin{bmatrix} E\left[\left(\frac{\partial L(X^{(T)},\theta)}{\partial A}\right)^2\right] & \mathbf{0} \\ \mathbf{0} & E\left[\frac{\partial L(X^{(T)},\theta)}{\partial \boldsymbol{a}}\frac{\partial L(X^{(T)},\theta)}{\partial \boldsymbol{a}}^\dagger\right] \end{bmatrix} \tag{27}$$

which can be rewritten as $J(\sigma_v) = E[E[\boldsymbol{k}|x_t]E[\boldsymbol{k}|x_t]^\dagger]$ with

$$\boldsymbol{k} = \begin{bmatrix} n_t e^{-j\phi_t} + n_t^* e^{j\phi_t}\mathbf{1}^\dagger \\ j\frac{A}{\sigma_n^2}\{x_t^* e^{j\phi_t} - x_t e^{-j\phi_t}\}\boldsymbol{\xi}^\dagger \end{bmatrix} \tag{28}$$

where $\mathbf{1}$ denote the $T\times 1$ vector with all entries equal to one, notice that $\mathbf{1}^\dagger\mathbf{1} = T$. Using classical inequalities we have

$$J(\sigma_v) \le E[\boldsymbol{k}\boldsymbol{k}^\dagger] = \begin{bmatrix} \frac{2}{\sigma_n^2}\mathbf{1}^\dagger\mathbf{1} & \mathbf{0} \\ \mathbf{0} & \frac{2A^2}{\sigma_n^2}\boldsymbol{\xi}^\dagger\boldsymbol{\xi} \end{bmatrix} = J(0) \tag{29}$$

$J(0)$ is the Fisher information matrix when no noise is present in the phase, see for instance [2] for the polynomial basis function case. Equation (29) leads to a quite expected result: *There is a loss of information related to the presence of noise in the phase.* The effect of this would be to raise the Cramèr-Rao bound for the amplitude and phase parameters

$$J(\sigma_v)^{-1} \ge J(0)^{-1} \tag{30}$$

Equality will hold when $E[e^{j\phi_t}|x_t] \overset{\text{a.e.}}{=} e^{j\phi_t}$. By examining the joint density function of x_t and ϕ_t one deduces that for $\sigma_n^2 > 0$, ϕ_t must be certain and equal to φ_t. Hence, equality holds only for $\sigma_v^2 = 0$.

One may notice that for the case of a noisy phase, the decoupling between the estimation of the phase parameters and the signal amplitude is preserved.

It is useful to define a relative Fisher information matrix as $\tilde{J}(\sigma_v) = J(0)^{-1}J(\sigma_v)$ which preserves the symmetry and the non-negative definiteness of $J(\sigma_v)$. The choice of this representation is motivated by the fact that we compare the achievable accuracy in estimating the signal parameters when no noise is present with the one obtained when a noise of variance σ_v^2 is present in the phase. This matrix can be easily seen to be independent from the choice of the basis functions modeling the phase, in fact

$$\tilde{J}(\sigma_v) = \frac{2}{\sigma_n^2}\times \begin{bmatrix} E[E[\Re\{n_t e^{-j\phi_t}\}|x_t]^2] & \mathbf{0} \\ \mathbf{0} & E[E[\Im\{x_t^* e^{j\phi_t}\}]^2]I_n \end{bmatrix} \tag{31}$$

For small values of σ_v^2, it is possible to use a Taylor expansion of matrix elements around $\sigma_v^2 = 0$,

$$\frac{2}{\sigma_n^2} E[E[\Re\{n_t e^{-j\phi_t}\}|x_t]^2] = 1 - \sigma_v^2 + O(\sigma_v^4) \qquad (32)$$

$$\frac{2}{\sigma_n^2} E[E[\Im\{x_t^* e^{j\phi_t}\}]^2] = 1 - (1 + \frac{2A^2}{\sigma_n^2})\sigma_v^2 + O(\sigma_v^4) \qquad (33)$$

which leads to the Taylor expansion of $\tilde{J}(\sigma_v)$,

$$J(\sigma_v) = I_{n+1} - \sigma_v^2 \begin{bmatrix} 1 & 0 \\ 0 & (1 + \frac{2A^2}{\sigma_n^2}) I_n \end{bmatrix} + O(\sigma_v^4) \qquad (34)$$

This expression shows, quite surprisingly, that the relative loss of information on the phase parameters due to σ_v^2, is proportional to the SNR as defined by A^2/σ_n^2. For large values of the SNR the relative Fisher information matrix decreases quickly when σ_v^2 increases.

Figure 1 shows the value of $\frac{-1}{2\sigma_n^2} E[E[x_t^* e^{j\phi_t} - x_t e^{-j\phi_t}|x_t]^2]$ for different values of the SNR. This corresponds to the relative Fisher information matrix for the phase parameters. This decrease does not mean that when the SNR is low we have better estimates. It only says that the higher the SNR is, the higher is our loss of precision when there is some noise on the phase compared to the case of noise free phase. For instance, if no additive noise is present, *i.e.* SNR=∞, then the Fisher information matrix $J(0)$ tends toward infinity and a small amount of phase noise makes $I(\sigma_v)$ finite which translates into a curve equal to 1 at $\sigma_v^2 = 0$ and 0 when $\sigma_v^2 > 0$.

To obtain closed form expressions for the Cramèr-Rao bound, one needs to invert the matrix $\boldsymbol{\xi}^\dagger \boldsymbol{\xi}$ which depends on the choice of the basis functions. This task can be simple if, for instance, one uses orthogonal basis functions. By orthogonal, it is meant that $\sum_{t=0}^{T-1} \xi_{ti}\xi_{tj} = \delta_{ij} \sum_{t=0}^{T-1} \xi_{ti}^2, \ \forall i, j$ and, obviously, $\boldsymbol{\xi}^\dagger \boldsymbol{\xi}$ is diagonal. This has the advantage of providing decoupled estimates of the phase parameters as well as simplifying estimation procedures which are based on the method of least squares.

For non-orthogonal basis functions the task of inverting this matrix, even numerically, can be very complex. For example, with the polynomial basis functions the matrix $\boldsymbol{\xi}^\dagger \boldsymbol{\xi}$ is ill-conditioned thus preventing the use of numerical procedures. Peleg and Porat [2] propose the use of an asymptotic (large T) expression for its inverse.

5. CONCLUSION

In this paper we presented a more general model for constant amplitude monocomponent signals. The novelty lies in the introduction of a noise component in the phase which is highly relevant in many radar and communications applications. We showed the identifiability of the signal model for the case of Gaussian phase noise with any variance. For other noise distributions identifiability is not guaranteed, it depends on the shape of the characteristic function of the noise. We derived explicit formulas for the Cramér-Rao lower bound on the variance of the estimated signal parameters for a general basis function specification. These bounds are compared to the noise free case and a quantitative *loss of information* is derived. It is shown that for a fixed phase noise variance, the loss of information increases when the SNR increases. This can be justified by the fact that at very low SNR, estimation errors are due mainly to the additive noise.

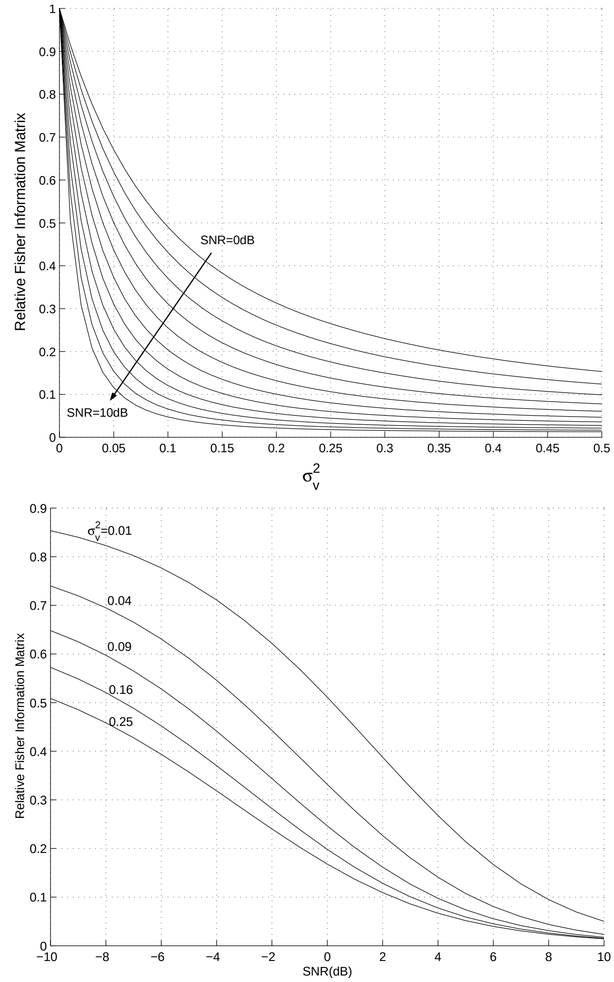

Fig. 1. Decrease of the relative Fisher information matrix for the phase parameters when the SNR increases

6. REFERENCES

[1] P. M. Djurić and S. M. Kay, "Parameter Estimation of Chirp Signals," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, no. 12, pp. 2118–2126, December 1990.

[2] S. Peleg and B. Porat, "The Cramér-Rao Lower Bound for Signals with Constant Amplitude and Polynomial Phase," *IEEE Transactions on Signal Processing*, vol. 39, pp. 749–752, 1991.

[3] S. Peleg and B. Friedlander, "The Discrete Polynomial-Phase Transform," *IEEE Transactions on Signal Processing*, vol. 43, pp. 1901–1914, 1995.

[4] W.P. Robins, *Phase Noise In Signal Sources*, IEE Telecommunications series 9. Peter Peregrinus Ltd., London, UK., 1984.

[5] A. Taleb and A.M. Zoubir, "Modelling and Analysis of Nonstationary Noisy Phase Signals," *IEEE Transactions on Signal Processing*, submitted.