FREQUENCY ESTIMATION AND DETECTION FOR SINUSOIDAL SIGNALS WITH ARBITRARY ENVELOPE: A NONLINEAR LEAST-SQUARES APPROACH

OLIVIER BESSON

ENSICA
Department of Avionics and Systems
31056 Toulouse, France
besson@ensica.fr

PETRE STOICA

Uppsala University
Systems and Control Group
75103 Uppsala, Sweden
ps@syscon.uu.se

ABSTRACT

In this paper, we consider the problem of estimating the frequency of a sinusoidal signal whose amplitude could be either constant or time-varying. We present a nonlinear least-squares (NLS) approach when the envelope is time-varying. We show that the NLS estimator can be efficiently implemented using a FFT. A statistical analysis shows that the NLS frequency estimator is nearly efficient. The problem of detecting amplitude time variations is next addressed. A statistical test is formulated, based on the statistics of the difference between two frequency estimates. The test is computationally efficient and yields as a by-product consistent frequency estimates under either hypothesis (i.e. constant or time-varying amplitude). Numerical examples are included to show the performance in terms of both estimation and detection.

1. INTRODUCTION

Estimating the parameters of sinusoidal signals with time-varying amplitude has been the topic of numerous studies in recent years (see e.g. [1]-[2] and references therein). This type of models proves to be relevant in many signal processing applications, such as precipitation and wind field velocity measurement via weather radar or lidar and vehicle speed determination by means of a Doppler radar (see [1]-[2] for details). For these and other applications the signal is well-described by the following equation:

$$y(t) = \alpha x(t) e^{i\omega_0 t} + e(t) \quad t = 0, 1, 2, \ldots \tag{1}$$

where α is a complex-valued amplitude, $x(t)$ is a real-valued time-varying envelope, ω_0 is the frequency, and $e(t)$ is a disturbance. In general, the parameter of interest is ω_0, the time-varying amplitude being considered as a nuisance, ***i.e.*** a multiplicative noise.

This paper is concerned with the application of the Nonlinear Least-Squares (NLS) approach to the model (1) with the following two goals:

Estimation. NLS frequency estimation has gained renewed interest in recent years [3] because of its computational simplicity (at least when a single frequency is sought, which is the case considered herein), its robustness to mismodelling the additive noise and its statistical accuracy whatever the hypothesis made on the noise (e.g. white or colored). Herein, we consider the application of the NLS approach to estimate ω_0 in (1).

Detection. In the above mentioned problems, the multiplicative model (1) has been proposed as an alternative to the usual model of a constant-amplitude sinusoidal signal in noise with the goal of obtaining a better representation of the physical phenomena giving rise to the observed signal. However, in many cases, the signal departs only slightly from the constant amplitude case, that is $x(t)$ in (1) is only slowly varying. Therefore a key issue, prior to or in conjunction with frequency estimation, is to decide whether the amplitude is time-varying or constant. In this paper, we present a conceptually simple and computationally efficient test for this detection problem.

2. NLS FREQUENCY ESTIMATION

In this section we present a NLS approach to estimate ω_0 (as well as α and $\{x(t)\}$). In this approach, the estimates are obtained as the minimizing arguments of the following criterion

$$\sum_{t=0}^{N-1} \left| y(t) - \alpha x(t) e^{i\omega t} \right|^2 \tag{2}$$

where N denotes the number of available data samples. It is well known that if $x(t)$ *were constant* (i.e. $x(t) \equiv 1$) then the estimate of ω_0 obtained by minimizing (2) would be given by the location of the highest peak of the data periodogram:

$$\hat{\omega}_0^{(0)} = \arg\max_{\omega} \frac{1}{N} \left| \sum_{t=0}^{N-1} y(t) e^{-i\omega t} \right|^2 \tag{3}$$

In the more general case in which $x(t)$ is *time-varying and unknown* we have the following:

Proposition 1 *The NLS estimate of ω_0 that minimizes (2) is given by*

$$\widehat{\omega}_0^{(1)} = \arg\max_{\omega} \frac{1}{N} \left| \sum_{t=0}^{N-1} y^2(t) e^{-i2\omega t} \right|^2 \tag{4}$$

Proof: see [4].

To obtain (4) we assume that $\sum_{t=0}^{N-1} x^2(t) = 1$ and that $\omega_0 \leq \pi/2$; otherwise, the signal model in (1) is ambiguous. Note that (4) can be obtained by dividing by two the peak location of the FFT (with possibly zero-padding) of $y^2(t)$. Interestingly enough, the NLS frequency estimator (4) has been used, on more or less heuristical grounds, in the communication literature where it is referred to as "squaring loop" (see [5, chapter 6]).

We now analyze the performance of the NLS frequency estimator. The following proposition gives the asymptotic (*i.e.* large sample) variance of the estimate (4).

Proposition 2 *Under the assumption that $x(t)$ in (1) is a Gaussian stationary process and that $e(t)$ is a white circular Gaussian noise with zero mean and variance σ_e^2, the asymptotic variance of $N^{3/2}\left(\widehat{\omega}_0^{(1)} - \omega_0\right)$ is given by:*

$$\lim_{N\to\infty} N^3 \mathcal{E}\left(\widehat{\omega}_0^{(1)} - \omega_0\right)^2 = \frac{6}{SNR_1}\left[1 + \frac{1}{2} SNR_1^{-1}\right] \tag{5}$$

where $SNR_1 = \mathcal{E}\left[|\alpha|^2 x^2(t)\right]/\sigma_e^2$.

Proof: see [4]

Remark 1 *We note that the Cramér-Rao Bound for the problem at hand is given in the high SNR case and under the assumption that $x(t)$ obeys a finitely-dimensional parametric model by [6]:*

$$CRB = \frac{1}{N^3}\frac{6}{SNR_1} \tag{6}$$

Hence the variance of the NLS frequency estimator can be quite close to the CRB, at least for a sufficiently high Signal to Noise Ratio. Observe that, in contrast to (6), the variance expression (5) for the NLS frequency estimate does not rely on a specific model for the envelope.

3. NLS-BASED DETECTION

As indicated in the introduction, it is of major interest to decide between the following two hypotheses, using a data sample of length N :

$$\begin{aligned} \mathbf{H}_0 \quad &: \quad y(t) = Ae^{i\omega_0 t} + e(t) \\ \mathbf{H}_1 \quad &: \quad y(t) = \alpha x(t) e^{i\omega_0 t} + e(t) \end{aligned} \tag{7}$$

Deciding between $\mathbf{H}_0$ and $\mathbf{H}_1$ in (7) is a binary hypothesis problem, which could be solved in an optimal manner e.g. by using a Likelihood Ratio Test (LRT) [7]. However, a LRT would require further assumptions on the structure of $x(t)$. Such assumptions would seldomly be met in practice; anyway, the LRT would be optimal only in a certain class of envelopes. Finally, the computational burden associated with the LRT is generally high. A test different from LRT was derived in [8] but it is still rather involved computationally. In contrast, we are interested in deriving a test which 1) should be computationally simple 2) should provide, as a by-product, accurate frequency estimates and 3) should be robust to mismodelling the envelope. Our detection scheme relies on the following observation. Under $\mathbf{H}_0$, both $\widehat{\omega}_0^{(0)}$ and $\widehat{\omega}_0^{(1)}$ provide consistent (and efficient or nearly efficient) estimates of the frequency. In contrast, under $\mathbf{H}_1$, only $\widehat{\omega}_0^{(1)}$ will have this property. In fact, under $\mathbf{H}_1$, $\widehat{\omega}_0^{(0)}$ is likely to be asymptotically biased (depending on the type of variation of $x(t)$) and will thus be quite different from $\widehat{\omega}_0^{(1)}$. Hence, the difference $\left(\widehat{\omega}_0^{(1)} - \widehat{\omega}_0^{(0)}\right)$ can serve as a good indicator of whether the amplitude is constant or time-varying.

We now formalize the previous idea. The following proposition is needed to derive the test.

Proposition 3 *Under $\mathbf{H}_0$, the vector $N^{3/2}\widehat{\omega} \overset{\text{def}}{=} N^{3/2}\begin{pmatrix} \widehat{\omega}_0^{(0)} - \omega_0 \\ \widehat{\omega}_0^{(1)} - \omega_0 \end{pmatrix}$ is asymptotically Gaussian distributed with zero-mean and covariance matrix given by*

$$\lim_{N\to\infty} N^3 E\left\{\widehat{\omega}\widehat{\omega}^T\right\} = \frac{6}{SNR_0}\begin{pmatrix} 1 & 1 \\ 1 & 1 + \frac{1}{2}SNR_0^{-1} \end{pmatrix} \tag{8}$$

where $SNR_0 = |A|^2/\sigma_e^2$.

Proof: see [4].

A simple consequence of Proposition 3 is:

Corollary 1 *Under $\mathbf{H}_0$, we have*

$$N^{3/2}\left(\widehat{\omega}_0^{(1)} - \widehat{\omega}_0^{(0)}\right) \overset{\text{as}}{\sim} \mathcal{N}(0, \quad 3 \times SNR_0^{-2}) \tag{9}$$

where $\overset{\text{as}}{\sim}$ means asymptotically distributed.

Hence, under $\mathbf{H}_0$, $\mathcal{T} \overset{\text{def}}{=} N^3\left(\widehat{\omega}_0^{(1)} - \widehat{\omega}_0^{(0)}\right)^2 SNR_0^2/3$ is asymptotically $\mathcal{X}^2(1)$. This property can be used to obtain a statistical test. This test will have a Probability of False Alarm as defined by

$$P_{FA} = \Pr\left\{\mathcal{T} > \gamma \mid \mathbf{H}_0\right\} \tag{10}$$

where the threshold γ is used to control the value of P_{FA}. To summarize, the detection scheme proposed is as follows.

Outline of the detection scheme

Step 1 From a $\mathcal{X}^2(1)$ table, obtain the threshold γ for a given P_{FA}.

Step 2 Compute the frequency estimates $\widehat{\omega}_0^{(0)}$ and $\widehat{\omega}_0^{(1)}$, as given by (3) and (4). Obtain the NLS estimates of A and σ_e^2 as

$$\widehat{A}^{(0)} = \frac{1}{N} \sum_{t=0}^{N-1} y(t) e^{-i\widehat{\omega}_0^{(0)} t} \qquad (11)$$

$$\widehat{\sigma}_e^{2(0)} = \frac{1}{N} \sum_{t=0}^{N-1} \left| y(t) - \widehat{A}^{(0)} e^{i\widehat{\omega}_0^{(0)} t} \right|^2 \qquad (12)$$

Step 3 Compute an estimate $\widehat{\mathcal{T}}$ of $\mathcal{T}$:

$$\widehat{\mathcal{T}} = \frac{N^3 \left| \widehat{A}^{(0)} \right|^4}{3\widehat{\sigma}_e^{4(0)}} \left(\widehat{\omega}_0^{(1)} - \widehat{\omega}_0^{(0)} \right)^2 \qquad (13)$$

Step 4 If $\widehat{\mathcal{T}} < \gamma$, accept $\mathbf{H}_0$. Otherwise, accept $\mathbf{H}_1$.

Note that, once $\mathbf{H}_0$ (resp. $\mathbf{H}_1$) is accepted, a corresponding frequency estimate is already available: $\widehat{\omega}_0^{(0)}$ (resp. $\widehat{\omega}_0^{(1)}$).

Remark 2 *It should be noted that under* $\mathbf{H}_1$, $\widehat{\sigma}_e^{2(0)}$ *as given by (12) could be quite large (since* $\widehat{\omega}_0^{(0)}$ *is poor). This in turn decreases* $\widehat{\mathcal{T}}$ *and hence the chance to detect* $\mathbf{H}_1$*. In order to remedy this problem, we propose to use the following estimator of* σ_e^2*:*

$$\widehat{\sigma}_e^{2(1)} = \frac{1}{N} \left(\sum_{t=0}^{N-1} |y(t)|^2 - \left| \sum_{t=0}^{N-1} y^2(t) e^{-i2\widehat{\omega}_0^{(1)} t} \right| \right) \qquad (14)$$

The estimator (14), which corresponds to the NLS estimate of σ_e^2 *under* $\mathbf{H}_1$*, is a consistent estimate of* σ_e^2 *under* $\mathbf{H}_0$ *as well. Comparing (12) and (14) and noting that (4) is already available, an alternative way to estimate* $|A|$ *is given by:*

$$\left| \widehat{A}^{(1)} \right|^2 = \frac{1}{N} \left| \sum_{t=0}^{N-1} y^2(t) e^{-i2\widehat{\omega}_0^{(1)} t} \right| \qquad (15)$$

4. NUMERICAL EXAMPLES

In this section, we illustrate the performance of the proposed estimation and detection methods. We begin with the NLS frequency estimator. Let us consider a signal as generated by equation (1) where: $\alpha = 1$, $\omega_0 = 2\pi \times 0.18$, and $e(t)$ is a white noise with zero mean and variance σ_e^2. The envelope $x(t)$ is a real-valued second-order autoregression with variance σ_x^2 and poles at $\rho e^{\pm i2\pi f}$. The signal to noise ratio is chosen as $SNR_1 = \sigma_x^2/\sigma_e^2 = 20dB$. We ran 1000 Monte-Carlo simulations to estimate the mean squared errors (MSE) of the frequency estimate proposed here. The so-obtained MSE values, along with the CRB are shown in Figures 1 and 2 for varying N and f, respectively.

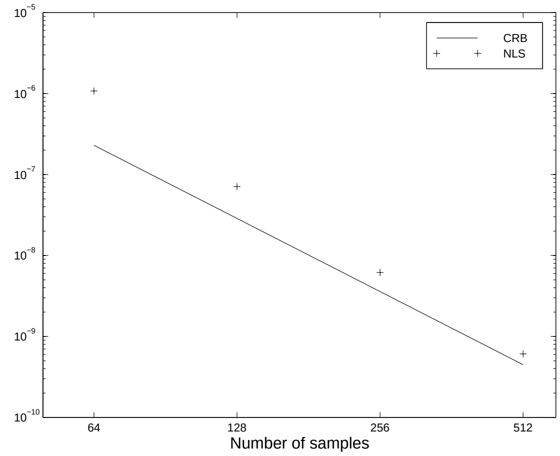

Figure 1: CRB and MSE of the frequency estimator versus the number of samples. $\rho = 0.95$ and $f = 0.01$.

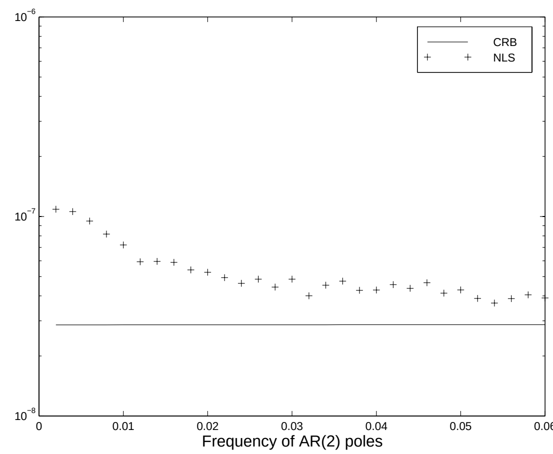

Figure 2: CRB and MSE of the frequency estimator versus the $AR(2)$ pole frequency. $\rho = 0.95$ and $N = 128$.

Despite the fact that our simple semi-parametric estimator does not use any information about the envelope, its performance is seen to get quite close to the parametric CRB.

The performance of the proposed detection test is now illustrated. It will be measured by P_{FA} and the Probability of Miss defined by

$$P_M = \Pr \{ \mathcal{T} < \gamma \mid \mathbf{H}_1 \} \qquad (16)$$

Two variants of the detection scheme were tested, which correspond to different estimates of A and σ_e^2:

Variant 1: A is estimated as in (11) and σ_e^2 by (12).

Variant 2: A is estimated as in (15) and σ_e^2 by (14).

We begin with illustrating the false alarm performances. To this end, 5000 Monte-Carlo simulations were run in each case considered. The signal was generated according to (7) where $A = e^{j\varphi}$ and $\varphi \in [0, 2\pi]$ is fixed in each simulation set. The Signal to Noise Ratio is defined as before: $SNR_0 = |A|^2 / \sigma_e^2$. The frequency is given by $\omega_0 = 2\pi \times f_0 = 2\pi \times 0.18$. *The theoretical P_{FA} is 0.01.* Tables 1 and 2 show the empirical P_{FA}'s obtained from Monte-Carlo simulations, as a function of N and SNR_0.

These tables indicate that the test performs as predicted by the theory, at least when a sufficient number of samples is available (typically $N > 64$). Additionally, the performance remains stable over a wide range of SNR_0's. Finally, we note that P_{FA} is only slightly increased when variant 2 is used.

Next, we illustrate the performance of the proposed test in terms of the miss probability. We consider the same example as before for the time-varying amplitude case. We ran 5000 Monte-Carlo simulations to estimate P_M for $P_{FA} = 0.01$. Figures 3 and 4 display the empirical P_M as a function of N and f. Some remarks on these figures are in order:

- A considerable improvement is achieved when using variant 2 compared to variant 1, in all cases tested.
- As the bandwidth of the envelope increases, P_M decreases, which is a logical behavior. The result of the test essentially depends on the ratio between the envelope bandwidth and the center frequency f_0. Also, the fact that P_M may not be negligible for small values of f (even though it is quite small for variant 2) is not a real problem. Indeed, when the envelope varies very slowly, the conventional constant amplitude frequency estimators are quite accurate.

5. REFERENCES

[1] G. Zhou. *Random amplitude and polynomial phase modeling of nonstationary processes using higher-order and cyclic statistics*. PhD thesis, University of Virginia, Charlottesville, January 1995.

[2] O. Besson and P. Stoica. Sinusoidal signals with random amplitude: least-squares estimators and their statistical analysis. *IEEE Trans. Sig. Proc.*, 43(11):2733–2744, Nov. 1995.

[3] P. Stoica and A. Nehorai. Statistical analysis of two nonlinear least-squares estimators of sine-wave parameters in the colored noise case. *Circ. Sys. Sig. Proc.*, 8(1):3–15, 1989.

[4] O. Besson and P. Stoica. Nonlinear least-squares frequency estimation and detection for sinusoidal signals with arbitrary envelope. Tech. Rep. 97 093R, Uppsala University, Aug. 1997.

[5] J. Proakis. *Digital Communications*. McGraw-Hill, New York, 3rd edition, 1995.

[6] J.M. Francos and B. Friedlander. Bounds for estimation of multicomponent signals with random amplitude and deterministic phase. *IEEE Trans. Sig. Proc.*, 43(5):1161–1172, May 1995.

[7] L.L. Scharf. *Statistical Signal Processing: Detection, Estimation and Time Series Analysis*. Addison Wesley, 1991.

[8] G. Zhou. A statistical test to discern random from constant amplitude harmonics. In *Proc. ICASSP-97*, pp 3685–3688.

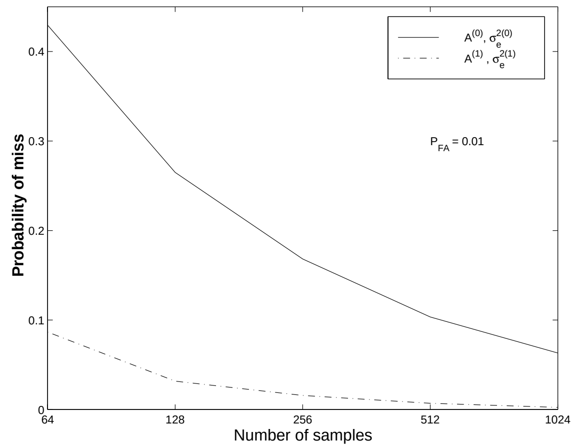

Figure 3: Probability of miss versus N. $f_0 = 0.18$, $\rho = 0.95$, $f = 0.01$ and $SNR_1 = 20dB$.

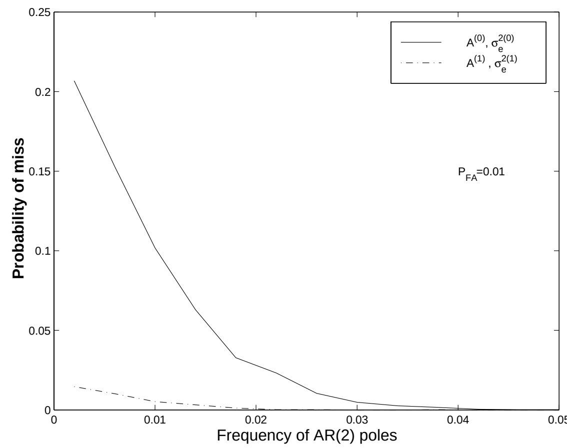

Figure 4: Probability of miss versus f. $\rho = 0.95$, $f_0 = 0.18$, $N = 512$ and $SNR_1 = 20dB$.

N	**Variant 1**	**Variant 2**
64	0.012	0.0168
128	0.0106	0.013
256	0.0104	0.0128
512	0.0092	0.0108
1024	0.0118	0.012

Table 1: Empirical P_{FA} versus N. $SNR_0 = 10dB$.

$SNR_0(dB)$	**Variant 1**	**Variant 2**
0	0.017	0.0234
5	0.0128	0.0136
10	0.0094	0.0122
15	0.0122	0.0152
20	0.0088	0.011
25	0.0088	0.0112

Table 2: Empirical P_{FA} versus SNR_0. $N = 128$.