Pergamon

PII: S0005-1098(96)00145-8

Automatica, Vol. 33, No. 2, pp. 209–215, 1997

0005-1098/97 $17.00 + 0.00

Brief Paper

Unbiased Estimation of a Sinusoid in Colored Noise via Adapted Notch Filters*

S. BITTANTI,† M. CAMPI‡ and S. M. SAVARESI†

Key Words—Frequency estimation; bias, sinusoids; notch filters; ARMA models.

Abstract—Standard notch models for the estimation of a sinusoid in noise are characterized by two poles on the unit circle and two zeros aligned with the poles. These models provide biased estimates of the unknown frequency, the bias increasing with the zero–pole distance. On the other hand, keeping the zeros far from the poles has the beneficial effect of making the identification algorithm less sensitive to the adopted initialization. In this paper we remove the pole–zero alignment constraint. The extra degree of freedom thus obtained is used to alleviate the traditional trade-off between unbiasedness of the estimates and robustness against poor initializations. The proposed technique applies to the case of both white and colored additive noise. © 1997 Elsevier Science Ltd. All rights reserved.

1. *Introduction*

The problem addressed in this paper is the identification of the frequency of a sinusoid corrupted by noise. A typical way of posing this problem is to assume that the observed data are given by

$$y(t) = s(t) + n(t), \quad t \in \mathbb{Z}, \tag{1}$$

where $s(t) = A\sin(\Omega_0 t + \phi)$ is a sinusoid with unknown frequency, amplitude and phase, and $n(\cdot)$ is a stationary and ergodic stochastic process with $E[n(t)] = 0$, $\mathrm{var}\,[n(t)] = \sigma^2 > 0$. The question of estimating the unknown frequency Ω_0 when the noise statistics are known has been extensively studied in the past four decades. In this framework, reference is often made to the following notch model:

$$G(z^{-1}) = \frac{1 - 2\rho\cos(\Omega)\,z^{-1} + \rho^2 z^{-2}}{1 - 2\cos(\Omega)z^{-1} + z^{-2}}, \quad \rho < 1, \tag{2}$$

where z^{-1} is the unit delay operator. This model is characterized by two poles on the unit circle, at frequencies $\pm\Omega$, and two zeros at $\rho e^{\pm j\Omega}$.

For the detection of the hidden sinusoid frequency, one considers the output of the model (2) fed by white noise as a potential mimic of the true signal, and sets up the frequency estimation issue as the problem of minimizing the loss function

$$J(\Omega) = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} E[\varepsilon(t,\Omega)^2], \tag{3}$$

where

$$\varepsilon(t,\Omega) = y(t) - \hat{y}(t,\Omega), \quad \hat{y}(t,\Omega) = [1 - G^{-1}(z^{-1})]y(t).$$

(here $\hat{y}(t,\Omega)$ is the one-step-ahead prediction of the output of the notch model with white input, and $\varepsilon(t,\Omega)$ is the associated prediction error).

However, this prediction-error identification technique has a major drawback in that it provides biased estimates of the frequency. This well-known feature is due to the fact that the minimum of the loss function (2) coincides with Ω_0 only when $\Omega_0 = \frac{1}{2}\pi$; otherwise, the minimum is shifted towards $\frac{1}{2}\pi$, and this results in a biased estimate of the frequency. To compensate for biasedness, a classical simple remedy is to let $\rho \to 1$ (indeed, the parameter ρ is usually called the 'de-biasing' parameter). In this way, the shape of the loss function becomes more peaked, and the bias is reduced. However, as $\rho \to 1$, the optimization problem associated with the search for the sinusoid frequency becomes more and more intractable, and accurate initialization becomes of major importance. Furthermore, in practice, one has to deal with a *finite* number of data, so that (3) must be replaced by its sample counterpart. In this case, a peculiar phenomenon enters the game: as $\rho \to 1$, the loss function is infested with numerous ripples, resulting in a bunch of local minima (see Remark 2 in Section 2).

To overcome these major drawbacks, in this paper we propose to drop the constraint that the zeros of the notch model are aligned with the poles. By suitably exploiting the so obtained extra degrees of freedom, we shall be able to design a new estimation technique that provides unbiased estimates without oversharpening effects on the loss function, and a modest ripple effect. With this objective in mind, we replace the model (2) with

$$G_f(z^{-1}) = \frac{C(z^{-1}, f(\Omega), \rho)}{D(z^{-1}, \Omega)} = \frac{1 - 2\rho f(\Omega)z^{-1} + \rho^2 z^{-2}}{1 - 2\cos(\Omega)z^{-1} + z^{-2}}. \tag{4}$$

Then the main design tool is the function $f(\Omega)$. As we shall see, its suitable shaping (adapted to the noise statistics) turns out to be much more effective and suitable than the naive remedy of pushing ρ close to 1. The main advantage provided by this extra degree of freedom is that one can keep the zeros of the model far from the unit circle, thus obtaining an enlargement of the attraction domain in the loss function (in this way, the initialization issue becomes less dramatic), and a smoothing of the sample loss function, thus alleviating the false minima effect.

The problem of estimating the frequency of harmonic signals embedded in noise is a classical problem in the fields of signal processing, system identification and control system design. Different techniques have been introduced and applied (for a detailed review see e.g. the classical paper of Kay and Marple (1981) or the more recent work of Händel

* Received 25 April 1995; revised 6 December 1995; revised 19 March 1996; received in final form 23 May 1996. A preliminary version of this paper was presented at the 5th IFAC Symposium on Adaptive Systems in Control and Signal Processing, which was held in Budapest, Hungary during June 1995. The Published Proceedings of this IFAC meeting may be ordered from the Customer Support Departments of the Elsevier Science Regional Sales Offices (see p. ii of *Automatica*) or from Elsevier Science Ltd, The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, U.K. This paper was recommended for publication in revised form by Associate editor Bo Wahlberg under the direction of Editor Torsten Söderström. Corresponding author Professor Sergio Bittanti. Tel. +39 2 2399 3543; Fax +39 2 2399 3587; E-mail bittanti@elet.polimi.it.

† Politecnico di Milano, Dipartimento di Elettronica e Informazione, Piazza Leonardo da Vinci, 32 20133, Milano, Italy.

‡ Universitá degli studi di Brescia, Via Branze, 38, 25121, Brescia, Italy.

(1993)): among them, the use of notch models is comparatively new, and has turned out to be handy and effective.

The idea of using adaptive notch models for estimating the frequency of sinusoids in noise probably goes back to Friedlander and Smith (1984) and Rao and Kung (1984); in the filters they used, no constraints on the positions of poles and zeros were imposed. Numerical instability in the minimization of the associated loss function was the main drawback of their approach.

The one-parameter-per-sinusoid constrained notch model was first proposed by Nehorai (1985) (and independently by Ng, 1987). The notch model with poles constrained to lie on the unit circle, and zeros on the same radial position as poles, turned out to be simpler and more efficient than previous schemes: recursive adaptation formulas were easy to calculate and intrinsically stable except for frequencies close to 0 or π. Successive work (Stoica and Nehorai, 1988; Rao and Peng, 1988; Dragošević and Stanković, 1989; Chicaro and Ng, 1990; Dragošević, 1993) provided detailed analyses of its features from different points of view. In Rao and Peng (1988) and Händel and Nehorai (1994) an analysis of the tracking characteristics of the constrained notch filter was proposed. In Stoica and Nehorai (1988) and Dragošević (1993) a detailed analysis of asymptotic properties of the algorithm was provided, with an explicit computation of the asymptotical bias and variance of the estimate for the case $\rho \to 1$. In Hush *et al.* (1986) and Kwan and Martin (1989) two different constrained notch filter structures, providing no bias in the case of white noise, were proposed, while the general case of colored noise was not considered. In Chambers (1990) an alternate method (based upon all-pass filters) for the design of such filters was presented, while in Nishimura (1993) these white-noise-unbiasing filters were used within an innovative line-enhancer structure. The general problem of the unbiased estimation of a sinusoid embedded in colored noise has been considered in Quinn and Fernandes (1991), where a computationally expensive procedure based on an ARMA-like filter is proposed, suitable for high-precision off-line estimation. The procedure proposed there differs from the others mentioned above mainly in that it does not make use of the classical biquadratic constrained notch filter. In the present paper (which generalizes the results presented in Bittanti *et al.*, 1995) we present a general method (suitable both for white and colored noise) to obtain unbiased frequency estimates. It is worth noting that in the white noise case (Section 2.1) the modified filter coincides with the filter first proposed in Hush *et al.* (1986), even though it is obtained following a different design rationale.

The paper organization is as follows. In Section 2 we consider the frequency-domain interpretation of prediction-error identification methods (which represent a key tool for the development of our method), and a min–max procedure for the selection of the optimal zeros position is given. The procedure holds in general for white and colored noise; however, for clarity of exposition we shall treat the cases of white noise and colored noise in two different subsections (Sections 2.1 and 2.2 respectively). A maximum-likelihood-type algorithm for the estimation of the frequency of the sinusoid is introduced in Section 3, while the performance of our method are tested in Section 4 by some illustrative simulations. Some conclusive remarks end the paper in Section 5.

2. *Unbiased frequency estimation via notch filters*

Assuming that the numerator $C(z^{-1}, f(\Omega), \rho)$ of the model (4) is a Hurwitz polynomial, the prediction error associated with the model $G_f(z^{-1})$ is given by the expression

$$\varepsilon(t, \Omega) = \frac{D(z^{-1}, \Omega)}{C(z^{-1}, f(\Omega), \rho)} y(t).$$

Since $y(\cdot)$ is quasi-stationary, so is $\varepsilon(\cdot, \Omega)$. From identification theory in a quasi-stationary context (see e.g. Ljung, 1987), the loss function (3) can then be written as

$$J(\Omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{D(e^{j\omega}, \Omega)}{C(e^{j\omega}, f(\Omega), \rho)} \right|^2 S_y(\omega) \, d\omega, \tag{5}$$

where $S_y(\omega)$ is the power spectrum of $y(\cdot)$. Denoting the power spectrum of the sinusoid $s(\cdot)$ by

$$S_s(\omega) = A^2[0.25\delta(\omega + \Omega_0) + 0.25\delta(\omega - \Omega_0)]$$

and the spectrum of the noise $n(\cdot)$ by $S_n(\omega)$, and substituting into (5), one obtains $J(\Omega) = J_1(\Omega) + J_2(\Omega)$, with

$$J_1(\Omega) = \frac{\frac{1}{2}A^2}{2\pi} \left| \frac{D(e^{j\Omega_0}, \Omega)}{C(e^{j\Omega_0}, f(\Omega, \rho)} \right|^2, \tag{6}$$

$$J_2(\Omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{D(e^{j\omega}, \Omega)}{C(e^{j\omega}, f(\Omega), \rho)} \right|^2 S_n(\omega) \, d\omega. \tag{7}$$

From (6) and (4), the explicit form of $J_1(\Omega)$ can be easily computed as

$$J_1(\Omega) = \frac{A^2}{4\pi} \times \frac{4(\cos\Omega_0 - \cos\Omega)^2}{[\cos 2\Omega_0 - 2\rho f(\Omega)\cos\Omega_0 + \rho^2]^2 + [\sin 2\Omega_0 - 2\rho f(\Omega)\sin\Omega_0]^2}, \tag{8}$$

while $J_2(\Omega)$ depends on the spectral characteristics of $n(\cdot)$. In connection with (7) and (8), it is worth noticing that;

(i) the minimum of $J_1(\Omega)$ *is located at* $\Omega = \Omega_0$, *independently of the particular shape of the function* $f(\cdot)$;

(ii) $J_2(\Omega)$ *does not depend on* Ω_0.

When $f(\Omega) = \cos\Omega$ (the standard choice leading to the classical notch model (2)), an inspection of the expressions (7) reveals that the function $J_2(\Omega)$ is in general non-constant. As a consequence, the minimum of $J(\Omega)$ is pushed away from that of $J_1(\Omega)$ (which corresponds to the correct frequency $\Omega = \Omega_0$) by an amount that depends on the gradient of $J_2(\Omega)$ at $\Omega = \Omega_0$, and on the signal-to-noise ratio. In light of this, *we propose to select the function* $f(\cdot)$ *in the notch filter* (4) *so as to render* $J_2(\Omega)$ *constant and to obtain an unbiased estimate of the frequency* Ω_0. In view of the expression (7), this objective calls for a suitable adaptation of the location of the notch-model zeros to the noise spectrum. Thus we arrive at a constrained optimization problem over $J_2(\Omega)$ with respect to $f(\cdot)$. In what follows, we emphasize the dependence of $J_2(\Omega)$ by $f(\cdot)$ by writing $J_2(\Omega, f(\cdot))$.

A general procedure to put the above idea into practice can be sketched as follows.

Procedure for the selection of $f(\cdot)$.

1. Find $f_m(\cdot)$ as the function such that

$$\bar{J}_2(\Omega, f_m(\cdot)) \le \bar{J}_2(\Omega, b)$$

$$\forall\Omega \in [0, 2\pi] \quad \text{and} \quad \forall b \in \left(-\frac{\sqrt{1+\rho^2}}{\rho\sqrt{2}}, \frac{\sqrt{1+\rho^2}}{\rho\sqrt{2}} \right)$$

(the bound on b forces $C(z^{-1}, \Omega)$ to be Hurwitz).

2. Find $M = \max_\Omega \bar{J}_2(\Omega, f_m(\Omega))$.
3. Find $f_F(\Omega)$ such that $\bar{J}_2(\Omega, f_F(\Omega)) = M \ \forall\Omega \in [0, 2\pi]$.

The function $f_m(\cdot)$ will be referred to as the *extremal* function. Correspondingly, the pair $(\Omega, f_m(\Omega))$ defines the 'bottom valley' of $J_2(\cdot, \cdot)$. Obviously, M is the ordinate of the 'highest peak' of such a bottom valley. Finally, $f_F(\cdot)$ defines the *max–min isocline* of $J_2(\cdot, \cdot)$. Note in passing that, in view of the definition of M as a max–min of $J_2(\cdot, \cdot)$, the function $f_F(\cdot)$ always exists. Obviously, one can define all isoclines of $J_2(\cdot, \cdot)$ by considering all the functions $f(\cdot)$ such that $J_2(\Omega, f(\Omega)) = c$, $c \ge M$; such an $f(\cdot)$ will be called a *flattening function*. The function $f_m(\cdot)$ defined above has the distinctive feature of being the flattening function with lowest altitude for J_2.

We shall now give more concreteness to the procedure outlined above by first focusing on a white-noise disturbance $n(\cdot)$, and then elaborating further the case of a colored disturbance.

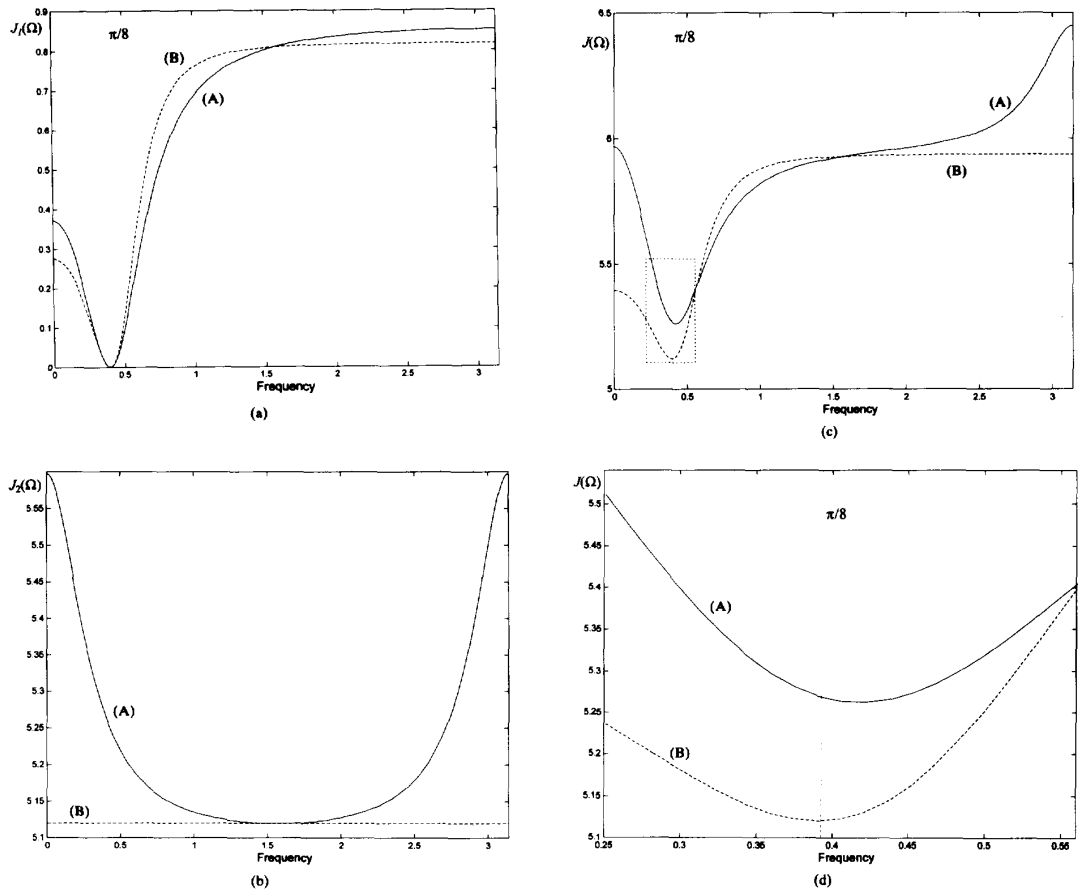

Fig. 1. (a) Shapes of J_1 in the case of $\Omega_0 = \frac{1}{8}\pi$ ($A = 1$, $\rho = 0.75$): (A) Radial zeros; (B): unbiasing zeros. (b) Shapes of J_2 ($n \approx \mathrm{WN}(0, 4)$, $\rho = 0.75$); (A) radial zeros; (B), unbiasing zeros. (c) Shapes of $J = J_1 + J_2(A = 1$, $n \approx \mathrm{WN}(0, 4)$, $\rho = 0.75$): (A) radial zeros; (B) unbiasing zeros. (d) A zoom over (c).

2.1. *Unbiased estimation in white noise.* If $n(\cdot)$ is an uncorrelated sequence having zero mean and variance σ^2, i.e. $n(\cdot) \approx \mathrm{WN}(0, \sigma^2)$, then $S_n(\omega)$ is a constant, and the expression for J_2 can be explicitly evaluated. Using the *Rugizka algorithm* (see e.g. Åström, 1970) leads to

$$J_2(\Omega, f(\cdot)) = \frac{\sigma^2}{2\pi} \times \frac{2[1 + 2\cos^2\Omega + 2\rho^2\cos^2\Omega - \rho^4 - 8\rho\cos(\Omega)f(\Omega) + 4\rho^2 f(\Omega)^2]}{1 + \rho^2 - \rho^4 - \rho^6 - 4\rho^2 f(\Omega)^2 + 4\rho^4 f(\Omega)^2}.$$

A cumbersome but simple inspection shows that, for each $\Omega \in [0, 2\pi]$, the function $J_2(\Omega, b)$ has a unique minimum with respect to b, given by

$$f_m(\Omega) = \frac{(1 + \rho^2)\cos\Omega}{2\rho}. \tag{9}$$

Furthermore, it turns out that the 'bottom valley' defined by the pair $(\Omega, f_m(\Omega))$ has the property of being horizontal, namely

$$J_2(\Omega, f_m(\Omega)) = \frac{\sigma^2}{2\pi}\frac{2}{1 + \rho^2} = M \quad \forall\Omega \in [0, 2\pi]. \tag{10}$$

In other words, the extremal function $f_m(\cdot)$ is in fact the max–min flattening function:

$$f_F(\Omega) = f_m(\Omega) \quad \forall\Omega \in [0, 2\pi].$$

Remark 1 (*Unbiased property*). To compare the classical choice $f(\Omega) = \cos\Omega$ (radial positioning of the zeros with respect to the poles) with the choice (9), consider the case when the sinusoidal signal is $s(t) = \sin(\pi/8t)$ and the disturbance is $n \approx \mathrm{WN}(0, 4)$, when $\rho = 0.75$ is used. The functions J_1 and J_2 are shown in Figs 1(a) and (b) respectively, while $J = J_1 + J_2$ can be seen in Figs 1(c) and (d) respectively. With the radial choice, the function J_2 has the effect of displacing the minimum of J from its correct position Ω_0 (associated with the J_1 minimum) towards $\frac{1}{2}\pi$. Accordingly, the estimate is asymptotically unbiased only if $\Omega_0 = \frac{1}{2}\pi$. In contrast, with the optimized choice of the zeros, the function J_2 turns out to be flat, so that the minimum point of J coincides with that of J_1. This means that the estimate of Ω_0 is unbiased, regardless of the value of Ω_0.

Remark 2 (*Sample loss function and false minima*). In the practical implementation of the optimization problem associated with the minimization of the statistical loss function $J(\Omega)$ given by (3), one must resort to a sample version of $J(\Omega)$:

$$J_N(\Omega) = \frac{1}{N}\sum_{t=0}^{N} \varepsilon(t, \Omega)^2. \tag{11}$$

When using $J_N(\Omega)$, in place of $J(\Omega)$, a new problem arises, namely that of false minima. In this respect, it is important to point out the typical behavior of the sample loss function for different choices of ρ. Typical diagrams of $J_N(\Omega)$ for a number of values of ρ are given in Fig. 2 (the curves depicted

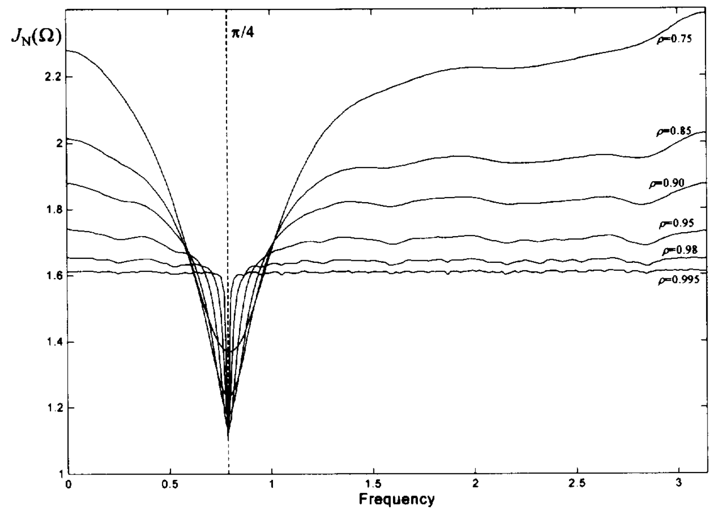

Fig. 2. The shape of $J_N(\Omega)$ when $\Omega_0 = \frac{1}{4}\pi$, $A = 1$, $\sigma^2 = 2$ and $N = 1000$.

are those of an unbiasing-zeros model: with respect to the false-minima problem, the curves of a radial-zeros model exhibit the same qualitative behavior). The important problem of false minima in the sample loss function is a further reason for keeping the zeros of the notch filter far away from the unit circle. As can be seen, ripple phenomena affect the sample loss function more dramatically when ρ is close to 1.

2.2. *The colored-noise case.* In this subsection we address the problem of finding a flattening function in the general colored-noise case. Differently from the white-noise case, an explicit expression for the flattening function cannot be worked out in general, and a numerical approach must be used.

In order to discuss the necessary modifications with respect to the white-noise case, we now run through the main steps of our procedure again:

(*i*) *Computing $J_2(\Omega, b)$.* Since

$$J_2(\Omega, b) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{D(e^{j\omega}, \Omega)}{C(e^{j\omega}, b, \rho)}\right|^2 S_n(\omega)\,d\omega,$$

function $J_2(\Omega, b)$ depends on $S_n(\omega)$, i.e. on the characteristics of the noise in which the sinusoid is embedded.

In order to compute $J_2(\Omega, b)$, there are two possibilities.

(a) The noise is modeled as the output of an ARMA model, fed by white noise:

$$n(t) = \frac{E(z^{-1})}{F(z^{-1})}e(t), \quad e(\cdot) \approx \mathrm{WN}\,(0, \sigma^2),$$

where, in a practical problem, the polynomials $E(z^{-1})$ and $F(z^{-1})$ are estimated from data. Once such a description of the noise is obtained, $J_2(\Omega, b)$ can be computed as

$$J_2(\Omega, b) = \frac{\sigma^2}{2\pi}\int_{-\pi}^{\pi}\left|\frac{D(e^{j\omega}, \Omega)E(e^{j\omega})}{C(e^{j\omega}, b, \rho)F(e^{j\omega})}\right|^2 d\omega.$$

by means of the Rugizka algorithm, or numerically.

(b) From the sequence $\{n(t)\}_{t=1}^{N}$, $J_2(\Omega, b)$ can be evaluated numerically for various values (Ω, b) (within the range of interest) by filtering $n(\cdot)$ with $D(z^{-1}, \Omega)/C(z^{-1}, \Omega, b, \rho)$ and by computing the variance of the filtered signal.

(*ii*) *Searching for $f_F(\cdot)$.* Once $J_2(\Omega, b)$ has been obtained (numerically, or as an explicit function), the procedure described in Section 2.1 can be applied to work out the max-min flattening function $f_F(\cdot)$. In most cases the procedure leads to a numerical evaluation of $f_F(\cdot)$ over a finite number of points. Then one has to resort to an interpolation procedure (e.g. polynomial interpolation) in order to obtain an explicit form for $f_F(\cdot)$. This is illustrated by means of the following example.

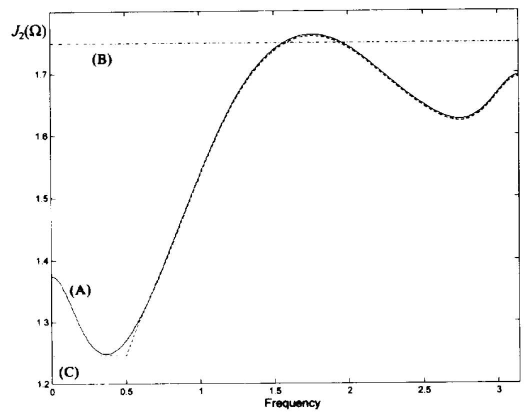

Fig. 3. Shape of $J_2(\Omega)$: (A) radial zeros; (B), global-flattening zeros; (C) local-flattening zeros.

Example 1. Assume that $n(\cdot)$ is generated by the ARMA model

$$n(t) = \frac{E(z^{-1})}{F(z^{-1})}e(t), \tag{12a}$$

where

$$E(z^{-1}) = z^{-2} - 1.2\cos(\tfrac{1}{2}\pi)z^{-1} + 0.36, \tag{12b}$$

$$F(z^{-1}) = z^{-2} - 0.4\cos(\tfrac{1}{3}\pi)z^{-1} + 0.04, \tag{12c}$$

and $e(\cdot)$ is a zero-mean white noise with variance $\sigma^2 = 1$.

If one resorts to the standard radial-zeros notch model, the shape of J_2 is that shown in Fig. 3 (line A). It is then apparent that the radial notch filter leads to unbiased estimates only if the true frequency Ω_0 equals one of the three stationary points of J_2.

With our approach, one has a further degree of freedom, namely the parameter b appearing in the polynomial $C(z^{-1}, b, \rho)$, by means of which the position of the notch-model zeros can be modified. In Fig. 4 a three-dimensional plot of $J_2(\Omega, b)$ is depicted. For the sake of clarity, we have set $b = \cos\Omega + \Delta b$ and represented J_2 as a function of Δb (it has been truncated for values of J_2 higher

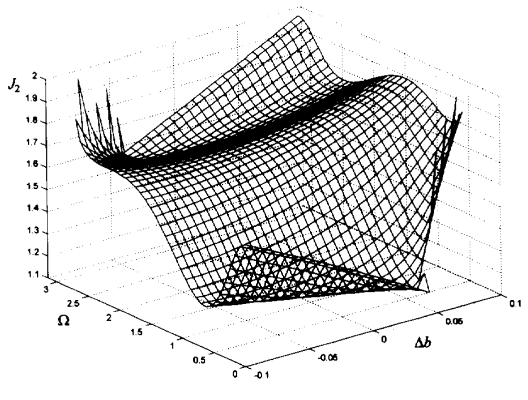

Fig. 4. Three-dimensional plot of the function $J_2(\Omega \cos \Omega + \Delta b)$.

than 2). The contour plot of this surface is given in Fig. 5. Line B is the max–min flattening function. The function $J_2(\Omega, f_F(\Omega))$ thus obtained is plotted in Fig. 3 as line B.

It is not so unusual to have a priori information on the approximate location of the unknown frequency, say $\Omega_0 \in [\Omega_1, \Omega_2]$. If so, it can be of interest to find an isocline of J_2 in the region $[\Omega_1, \Omega_2]$ with lower ordinate than the max–min value M of J_2. In such a way, one obtains a twofold benefit: first, the fact that the max–min isocline over $[\Omega_1, \Omega_2]$ is lower than the global max–min isocline reduces the effect of noise in the optimization problem; second, the search for a flattening function over a restricted frequency interval is easier and faster.

Example 1 (Continued). With reference to the above example, suppose that we have the a priori information that Ω_0 belongs to some interval, for instance that $\Omega_0 \in [0, 0.5]$. Then it is apparent from Fig. 4 that there are infinitely many isoclines crossing the frequency region of interest. Among them, one can select an isocline at a lower ordinate than the one corresponding to $f_F(\cdot)$. For instance, by selecting line C in Fig. 5, one obtains the partial isocline displayed in Fig. 3 again, as line C.

3. *Recursive algorithm for the frequency estimation*

In this section we present a simple recursive identification algorithm for the estimation of Ω_0 by means of the unbiasing notch model proposed in this paper. We present the general case when $n(\cdot)$ is a colored noise, and the sample loss function is given by (11).

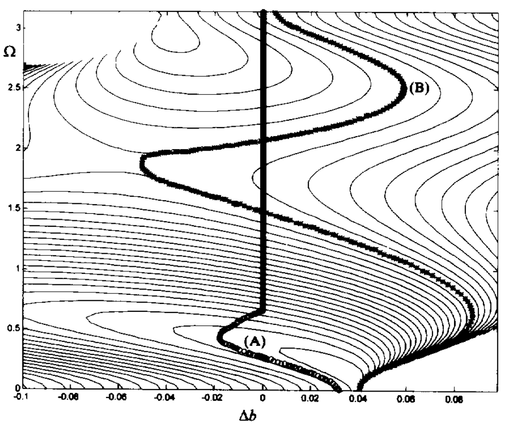

Fig. 5. Contour plot of $J_2(\Omega, \cos \Omega + \Delta b)$.

Denoting the derivative of $\varepsilon(t, \Omega)$ with respect to Ω by $\psi(t, \Omega)$, the basic maximum-likelihood recursion (see e.g. Söderström and Stoica, 1989) can be written as follows:

- $$\hat{\Omega}(t) = \hat{\Omega}(t-1) - \frac{\varepsilon(t, \hat{\Omega}(t-1))\psi(t, \hat{\Omega}(t-1))}{S(t)}, \tag{13a}$$

where $S(t) = \sum_{i=1}^{t} [\psi(i, \hat{\Omega}(i-1))]^2$ can be computed recursively as

- $$S(t) = S(t-1) + [\psi(t, \hat{\Omega}(t-1))]^2. \tag{13b}$$

$\varepsilon(t, \hat{\Omega}(t-1))$ can be approximately computed starting from the prediction error equation

$$C(z^{-1}, \hat{\Omega}(t-1))\varepsilon(t, \hat{\Omega}(t-1)) = D(z^{-1}, \hat{\Omega}(t-1))y(t),$$

which in this specific case leads to

- $$\begin{aligned}\varepsilon(t, \hat{\Omega}(t-1)) = y(t) &- 2\cos[\hat{\Omega}(t-1)]\,v(t-1) \\ &+ y(t-2) + 2\rho f(\hat{\Omega}(t-1))\varepsilon(t-1, \hat{\Omega}(t-2)) \\ &- \rho^2 \varepsilon(t-2, \hat{\Omega}(t-3)).\end{aligned} \tag{13c}$$

The approximate gradient $\psi(t, \hat{\Omega}(t-1))$ is obtained by differentiating both sides of the prediction error equation

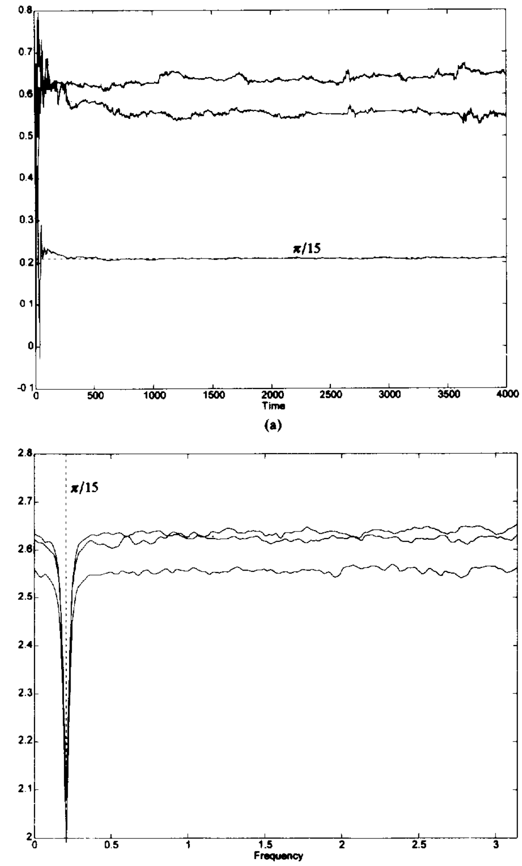

Fig. 6. (a) Three runs of simulation for the estimation of $\Omega_0 = \frac{1}{15}\pi$, using a radial-zeros model with $\rho = 0.98$. (b) Shape of the sample loss function J_N for each noise realization, using a radial-zeros model with $\rho = 0.98$.

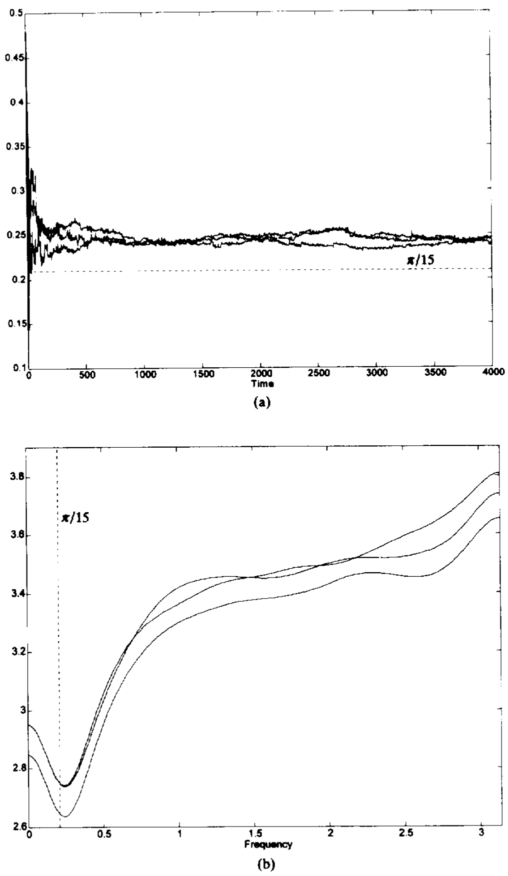

Fig. 7. (a) Three runs of simulation for the estimation of $\Omega_0 = \frac{1}{15}\pi$, using a radial-zeros model with $\rho = 0.75$. (b) Shape of the sample loss function J_N for each noise realization, using a radial-zeros model with $\rho = 0.75$.

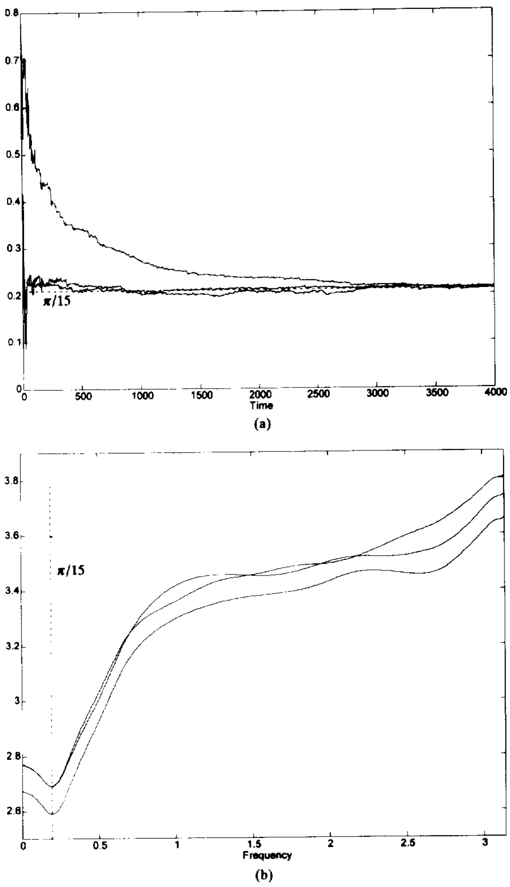

Fig. 8. (a) Three runs of simulation for the estimation of $\Omega_0 = \frac{1}{15}\pi$, using a local-flattening function with $\rho = 0.75$. (b) Shape of the sample loss function J_N, for each noise realization, using a local-flattening function with $\rho = 0.75$.

with respect to Ω, thus obtaining

- $\psi(t, \hat{\Omega}(t-1))$

$$= 2\sin[\hat{\Omega}(t-1)]y(t-1) + 2\rho \left.\frac{\partial f}{\partial \Omega}\right|_{\Omega=\hat{\Omega}(t-1)} \varepsilon(t-1)$$

$$+ 2\rho f(\hat{\Omega}(t-1))\psi(t-1) - \rho^2 \psi(t-2). \quad \text{(13d)}$$

Notice that (13c, d) are recursive equations, whose dynamics are determined by the polynomial $C(z^{-1}, \hat{\Omega}(t-1))$.

4. *Discussion via simulation examples*

We now present an illustrative simulation example, with the purpose of comparing the performance of the new technique with that achievable with the radial-zeros model. The main conclusion is that, when resorting to the radial-zeros model, one has to compromise between the amount of asymptotic bias in the estimate and the size of the attractiveness region of the optimal estimate. In contrast, the proposed technique guarantees asymptotic unbiasedness without deteriorating the degree of attractiveness.

Example 1 (Continued). Consider the signal $y(t) = s(t) + \sqrt{2}n(t)$, where $s(t) = \sin(\pi/15t)$ and $n(t)$ is the colored noise (12). First, the estimation of the frequency $\Omega_0 = \frac{1}{5}\pi$ has been performed with the radial positioning of the zeros with the choice $\rho = 0.98$ and initial condition $\hat{\Omega}(0) = \frac{1}{8}\pi$. The results of three simulation runs performed with the radial-zeros version ($f_F(\Omega) = 2\cos\Omega$) of the recursive algorithm described in Section 3, using 4000 data snapshots, are depicted in Fig. 6(a). In two cases the estimation algorithm ends in a local minimum far away from the correct frequency. This can be explained by means of Fig. 6(b), where the sample loss function (11) based on the whole set of 4000 data is represented. As can be seen, notwithstanding the large number of data, there are quite a lot of false minima in J_N. This can be explained according to the discussion of the ripple effect in Remark 2.

Obviously, in order to alleviate this effect, one possibility is to decrease the value of ρ. Thus the above simulations with the same data have been repeated with $\rho = 0.75$, leading to the diagrams of Fig. 7(a), which correspond to the loss functions of Fig. 7(b). It is clear that the loss functions are now smooth, so the problem of false minima is alleviated. However, the global minimum is not centered at the correct value $\Omega_0 = \frac{1}{15}\pi$ of the frequency (Fig. 7b), so the estimates are biased (Fig. 7a).

Turning now to our estimation approach, the performance achieved with the unbiasing choice of the zeros using the local-flattening function described in Section 2 (see Fig. 3, line C) and $\rho = 0.75$ has been tested again with the 4000 snapshots used in the preceding trials. In Figs 8(a, b) the diagrams of the estimates obtained and of the associated sample loss functions are given.

5. *Conclusions*

In this paper the problem of extracting a spectral line from noisy signals has been considered. The main contribution is the development of a technique for the *unbiased* estimation of the spectral line frequency in *colored noise*. The main limitation of the proposed method is the assumption that the spectral density of the background noise is a priori available. Although in some applications such an assumption can be taken for granted, it is of interest to probe further the case when both the frequency and the noise statistics are unknown. Along this line, a possibility would be to resort to a sequence of alternate steps, where a spectral line estimation iteration is followed by a noise statistics estimation step, up to convergence. This should be the subject of further research.

Acknowledgements—The work reported in this paper was supported by the European *HCM* project 'System Modelling and Identification', the Italian MURST project 'Model Identification, System Control, Signal Processing' and the CNR Center for System Theory of Milano. The authors are grateful to Dr Fabrizio Lorito for stimulating exchanges of ideas on the subject of sinusoids estimation.

References

Åström, K. J. (1970). *Introduction to Stochastic Control Theory.* Academic Press, New York.

Bittanti, S., M. Campi and S. Savaresi (1995). Unbiased estimation of a sinusoid in noise via notch filters. In *Proc. IFAC Symposium on Adaptive Control and Signal Processing,* Budapest, pp. 165–170.

Chambers, J. A. (1990). Frequency tracking using constrained adaptive notch filters synthesised from allpass sections. *IEE Proc.,* Pt F, **137,** 475–481.

Chicharo, J. F. and T. S. Ng (1990). Gradient-based adaptive IIR notch filtering for frequency estimation. *IEEE Trans. Acoust., Speech, Sig. Process.,* **ASSP-38,** 769–777.

Dragošević, M. V. (1993). Stationary points of the recursive generalized least squares algorithm for adaptive notch filtering. *IEEE Trans. Sig. Process.,* **SP-41,** 1672–1675.

Dragošević, M. V. and S. S. Stanković (1989). A generalized least squares method for frequency estimation. *IEEE Trans. Acoust., Speech, Sig. Process.,* **ASSP-37,** 805–819.

Friedlander, B. and J. O. Smith (1984). Analysis and performance evaluation of an adaptive notch filter. *IEEE Trans. Inf. Theory,* **IT-30,** 283–295.

Händel, P. (1993). Estimation methods for narrow-band signals. PhD Thesis, Uppsala.

Händel, P. and A. Nehorai (1994). Tracking analysis of an adaptive notch filter with constrained poles and zeros. *IEEE Trans. Sig. Process.,* **SP-42,** 281–291.

Hush, D. R., N. Ahmed, R. David and S. Stearns (1986). An adaptive IIR structure for sinusoidal enhancement, frequency estimation, and detection. *IEEE Trans. Acoust., Speech, Sig. Process.,* **ASSP-34,** 1380–1390.

Kay, S. M. and S. L. Marple, Jr (1981). Spectrum analysis—a modern perspective. *Proc. IEEE,* **69,** 1380–1419.

Kwan, T. and K. Martin (1989). Adaptive detection and enhancement of multiple sinusoids using a cascade IIR filter. *IEEE Trans. Circuits Syst.* **CAS-36,** 937–947.

Ljung, L. (1987). *Systems Identification: Theory for the User.* Prentice-Hall, Englewood Cliffs, NJ.

Nehorai, A. (1985). A minimal parameter adaptive notch filter with constrained poles and zeros. *IEEE Trans. Acoust, Speech, Sig. Process.,* **ASSP-33,** 983–996.

Ng, T. S. (1987). Some aspects of an adaptive digital notch filter with constrained poles and zeros. *IEEE Trans. Acoust., Speech, Sig. Process.,* **ASSP-35,** 158–161.

Nishimura, S. (1993). Adaptive detection and enhancement of closely spaced sinusoids using multirate techniques. In *Proc. IEEE International Conf. on Acoustics, Speech and Signal Processing,* Vol. III, pp. 400–403.

Quinn, B. G. and J. M. Fernandes (1991). A fast efficient technique for the estimation of frequency. *Biometrika,* **78,** 489–497.

Rao, D. V. B. and S. Y. Kung (1984). Adaptive notch filtering for the retrieval of sinusoids in noise. *IEEE Trans. Acoust., Speech, Sig. Process.,* **ASSP-32,** 791–802.

Rao, B. D. and R. Peng (1988). Tracking characteristics of the constrained IIR adaptive notch filter. *IEEE Trans. Acoust., Speech, Sig. Process.,* **ASSP-36,** 1466–1479.

Söderström, T. and P. Stoica (1989). *System Identification.* Prentice-Hall, Englewood Cliffs, NJ.

Stoica, P. and A. Nehorai (1988). Performance analysis of an adaptive notch filter with constrained poles and zeros. *IEEE Trans. Acoust., Speech, Sig. Process.,* **ASSP-36,** 911–919.