15th Triennial World Congress, Barcelona, Spain

AN ADAPTIVE GRID POINT RPEM ALGORITHM FOR HARMONIC SIGNAL MODELING

Emad Abd-Elrady [1]

Department of Systems and Control, Information Technology, Uppsala University, P.O. Box 337, SE-751 05 Uppsala, Sweden
ea@SysCon.uu.se

Abstract: Periodic signals can be modeled as a real wave with unknown period in cascade with a piecewise linear function. A recursive Gauss-Newton prediction error method (RPEM) for joint estimation of the driving frequency and the parameters of the nonlinear output function parameterized in a number of *adaptively estimated* grid points is introduced. The Cramér-Rao bound (CRB) is derived for the suggested algorithm. Numerical examples indicate that the suggested algorithm gives better performance than using fixed grid point algorithms. *Copyright ©2002 IFAC*

Keywords: Frequency estimation, Identification algorithms, Nonlinear systems, Piecewise linear analysis, Prediction error methods.

1. INTRODUCTION

There is a quite substantial literature dealing with the problem of retrieving noisy sinusoidal signals, see for example (Nehorai and Porat, 1986) and (Händel and Tichavskỳ, 1994). In general, a periodic function with unknown fundamental frequency in cascade with a parameterized and unknown nonlinear function can be used as a signal model for an arbitrary periodic signal as shown in Fig. 1. This approach has two additional properties. First, it gives information on the underlying nonlinearity, in cases where the overtones are generated by nonlinear imperfections in the system. Second, prior information about the wave form can be used to increase the efficiency of the algorithm.

In (Wigren and Händel, 1996), (Abd-Elrady, 2000) and (Abd-Elrady, 2001*b*), the nonlinearity was chosen to be piecewise linear with the estimated parameters being the function values in a *fixed* set of grid points as shown in Fig. 2,

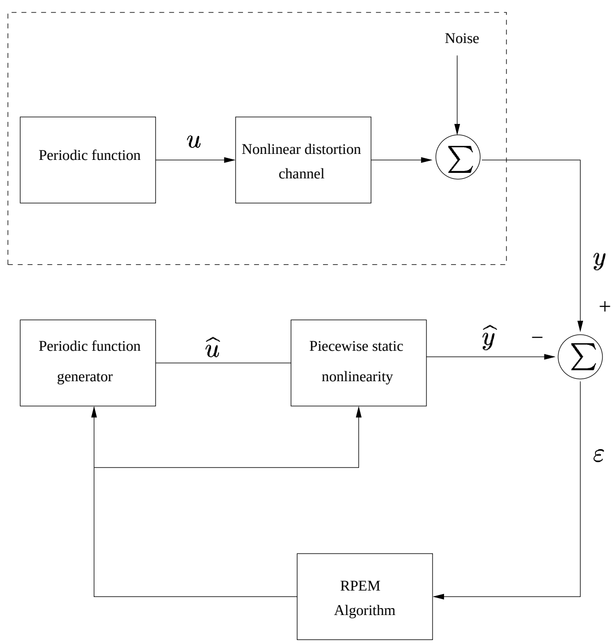

Fig. 1. The approach to harmonic signal modeling.

resulting in *fixed* grid point adaptation. In this paper, the RPEM algorithm introduced in (Abd-Elrady, 2000) is modified to enable the algorithm

[1] This work was supported by Swedish Research Council for Engineering Sciences under contract 98-654.

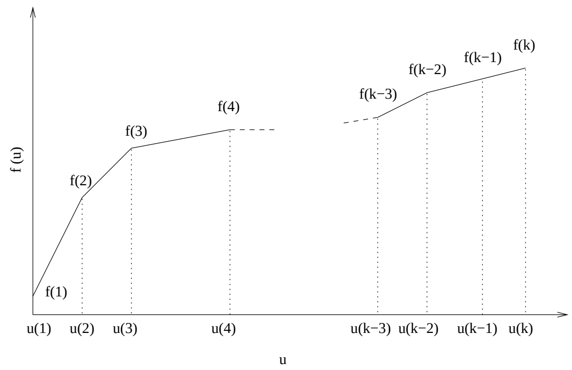

Fig. 2. Grids points, parameters and resulting piecewise linear model..

to estimate the grid points as well as the driving frequency and the parameters of the nonlinear output function, resulting in *adaptive* grid point adaptation. This expected to reduce modeling error since it gives the algorithm more freedom to choose the suitable grid points.

The contributions of this paper are as follows: The CRB is derived for the suggested method. Furthermore, the performance of the adaptive grid point algorithm is studied by numerical examples. The purpose is to investigate the performance as compared to fixed grid point algorithms.

2. THE SUGGESTED ALGORITHM

In order to define the parametric signal model, a periodic function being the input to the estimated static nonlinearity is needed. This function reflects any prior knowledge that is available. It could, for example, be chosen as a triangle wave in cases where the measured signal closely resembles that signal. Since the model is a cascade of two blocks as shown in Fig. 1, *the differential static gain* of the model will be a factor of two free parameters. It is nevertheless necessary for the algorithm to have information about where the static gain is located, or the criterion function may have an infinite number of minima. Hence, one of the parameters need to be fixed, cf. (Wigren, 1993). In (Wigren and Händel, 1996) and (Abd-Elrady, 2001*b*), this was done in the driving signal block. Here, however, the opposite situation will be investigated.

In order to fix the static gain in an amplitude subinterval I_o (cf. (Abd-Elrady, 2000)) *contained in exactly one* of the subintervals of the nonlinear block, the driving input signal $\widehat{u}(t,\boldsymbol{\theta_l})$ is modeled as

$$\widehat{u}(t,\boldsymbol{\theta_l}) = X\Lambda(\omega t), \quad \boldsymbol{\theta_l} = (X \ \omega)^T \tag{1}$$

where t denotes discrete time, $\omega \in [0,\pi]$ denotes the unknown normalized angular frequency, $\omega = 2\pi f/f_s$ where f is the frequency, and f_s is the sampling frequency. X is a (possibly time varying) parameter recursively estimated to allow the linear block of the model to adapt its static gain so that the data in I_o can be explained. The fact that $\Lambda(.)$ is periodic now means

C1) $\Lambda(\omega(t + \frac{2k\pi}{\omega})) = \Lambda(\omega t), \qquad k \in Z.$

Then let one period of $\Lambda(\omega t)$ be divided into L disjoint intervals $I_j,\ j = 1,\cdots,L$, and assume

C2) $\Lambda(\omega t)$ is a monotone function of ωt on each interval $I_j,\ j = 1,\cdots,L$.

Remark 1: **C2)** is introduced to avoid restrictions that would reduce the generality of the approach. This can be explained as follows. Assume that one static nonlinearity is used and $\Lambda(\omega t) = \sin(wt)$ then the model output $f_1(\theta_1,\sin(wt))$ is obtained. If the unknown parameter vector θ_1 of the nonlinear block is fixed, $f_1(\theta_1,\sin(w(\pi/w - t))) = f_1(\theta_1,\sin(wt))$ holds for all t. This means that the model signal in half of the time intervals of length π/w is given by the signal in the remaining time intervals.

A piecewise linear model is used for the parameterization of the nonlinearity, cf. (Wigren, 1993). Choosing I_o to be contained in the first interval I_1, the grid points are defined as

$$\boldsymbol{g_j} = \begin{cases} (u^1_{-k_1^-} u^1_{-k_1^-+1} \cdots u_{o^-} u_{o^+} \cdots u^1_{k_1^+-1} u^1_{k_1^+}), j = 1 \\ (u^j_{-k_j^-} u^j_{-k_j^-+1} \cdots u^j_{-1} u^j_1 \cdots u^j_{k_j^+-1} u^j_{k_j^+}), j \neq 1. \end{cases} \tag{2}$$

Then with $f_j(\boldsymbol{\theta_j},\boldsymbol{g_j},\widehat{u}(t,\boldsymbol{\theta_l}))$ denoting the nonlinearity to be used in I_j, The parameters $\boldsymbol{\theta_j}$ are chosen as the values of $f_j(\boldsymbol{\theta_j},\boldsymbol{g_j},\widehat{u}(t,\boldsymbol{\theta_l}))$ in the grid points, i.e.

$$\boldsymbol{\theta_j} = (f^j_{-k_j^-} \cdots f^j_{-1} f^j_1 \cdots f^j_{k_j^+})^T, \ j = 1,\cdots,L.$$

$$f_j(\boldsymbol{\theta_j},\boldsymbol{g_j},u^j_i) = \begin{cases} K_o\widehat{u}(t,\boldsymbol{\theta_l}) + f_o, \quad \widehat{u} \in I_o, \ j = 1 \\ f^j_i, \ \forall j, \ i = -k_j^-,..,-1,1,..,k_j^+. \end{cases} \tag{3}$$

Here K_o is the user chosen static gain constant defined as

$$\frac{\partial f_1(f_o,\boldsymbol{\theta_1},\boldsymbol{g_1},\widehat{u}(t,\boldsymbol{\theta_l}))}{\partial\widehat{u}} = K_o, \ \widehat{u}(t,\boldsymbol{\theta_l}) \in I_o. \tag{4}$$

Thus the model output becomes

$$\widehat{y}(t,\boldsymbol{\theta}) = f_j(\boldsymbol{\theta_j},\boldsymbol{g_j},\widehat{u}(t,\boldsymbol{\theta_l})), \ \widehat{u} \in I_j, \ j = 1,\cdots,L.$$
$$\boldsymbol{\theta} = (\boldsymbol{\theta_l}^T \ \boldsymbol{\theta_n}^T \ \boldsymbol{\theta_g}^T)^T, \ \boldsymbol{\theta_n} = (f_o \ \tilde{\boldsymbol{\theta}}_{\boldsymbol{n}}^T)^T,$$
$$\tilde{\boldsymbol{\theta}}_{\boldsymbol{n}} = (\boldsymbol{\theta_1}^T \cdots \boldsymbol{\theta_L}^T)^T, \ \boldsymbol{\theta_g} = (\boldsymbol{g_1}^T \cdots \boldsymbol{g_L}^T)^T. \tag{5}$$

Remark 2: $\widehat{u}(t,\boldsymbol{\theta_l}) \in I_j$ means that the phase ωt is such that I_j is in effect, cf. (Abd-Elrady, 2000).

A piecewise linear function of $\widehat{u}(t,\boldsymbol{\theta_l})$ can now be constructed from the linear segments with end points in (u^j_{i-1}, f^j_{i-1}) and (u^j_i, f^j_i). A RPEM then follows by a minimization of

$$V(\boldsymbol{\theta}) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} E[\varepsilon^2(t,\boldsymbol{\theta})] \tag{6}$$

where $E[.]$ denotes expectation. Here, $\varepsilon(t,\boldsymbol{\theta}) = y(t) - \widehat{y}(t,\boldsymbol{\theta})$ denotes the prediction error and $y(t)$ is the measured signal to be modeled.

The objective of this paper is, as stated in the introduction, to estimate the grid points recursively in addition to the estimation of the fundamental frequency and the parameters of the nonlinear output function. This will give the algorithm more freedom to choose the grid points and achieve a better performance. The negative gradient of $\widehat{y}(t,\boldsymbol{\theta})$ is needed in the formulation of the recursive algorithm, see (Söderström and Stoica, 1989). It is given by (for $\widehat{u}(t,\boldsymbol{\theta_l}) \in I_j, j = 1,\cdots,L$)

$$\boldsymbol{\psi}(t,\boldsymbol{\theta}) =$$
$$\left(\frac{\partial f_j(.)}{\partial \widehat{u}}\boldsymbol{\psi_l}(t) \;\; \frac{\partial f_j(.)}{\partial \widehat{f}_o} \;\; 0..0 \;\; \frac{\partial f_j(.)}{\partial \widehat{\boldsymbol{\theta}}_{\boldsymbol{j}}} \;\; 0..0 \;\; \frac{\partial f_j(.)}{\partial \widehat{\boldsymbol{g}}_{\boldsymbol{j}}} \;\; 0..0 \right)^T \tag{7}$$

where

$$\boldsymbol{\psi_l}(t) = \left(\Lambda(\phi) \mid_{\phi=\omega t} \;\; Xt\frac{d\Lambda(\phi)}{d\phi} \mid_{\phi=\omega t} \right)^T \tag{8}$$

Thus the RPEM algorithm becomes

$$\varepsilon(t) = y(t) - \widehat{y}(t)$$
$$\lambda(t) = \lambda_o\lambda(t-1) + 1 - \lambda_o$$
$$S(t) = \boldsymbol{\psi}^T(t)\boldsymbol{P}(t-1)\boldsymbol{\psi}(t) + \lambda(t)$$
$$\boldsymbol{P}(t) = (\boldsymbol{P}(t-1) - \boldsymbol{P}(t-1)\boldsymbol{\psi}(t)S^{-1}(t)\boldsymbol{\psi}^T(t)\boldsymbol{P}(t-1))/\lambda(t)$$

$$\begin{pmatrix} \widehat{\boldsymbol{\theta}}_{\boldsymbol{l}}(t) \\ \widehat{\boldsymbol{\theta}}_{\boldsymbol{n}}(t) \\ \widehat{\boldsymbol{\theta}}_{\boldsymbol{g}}(t) \end{pmatrix} = \left[\begin{pmatrix} \widehat{\boldsymbol{\theta}}_{\boldsymbol{l}}(t-1) \\ \widehat{\boldsymbol{\theta}}_{\boldsymbol{n}}(t-1) \\ \widehat{\boldsymbol{\theta}}_{\boldsymbol{g}}(t-1) \end{pmatrix} + \boldsymbol{P}(t)\boldsymbol{\psi}(t)\varepsilon(t) \right]_{D_M} \tag{9}$$

$$\phi = \widehat{\omega}(t)(t+1)$$
$$\widehat{u}(t+1) = \widehat{X}(t)\Lambda(\phi)$$
$$\boldsymbol{\psi_l}(t+1) = \left(\Lambda(\phi) \;\; \widehat{X}(t)(t+1) \;\; \frac{d\Lambda(\phi)}{d\phi} \right)^T$$

when $\widehat{u}(t+1) \in I_1$

when $\widehat{u}(t+1) \in [\widehat{u}^1_{-1}, u_o-]$

$$\widehat{y}(t+1) = \frac{\widehat{f}^1_{-1}u_o- - (K_o u_o- + \widehat{f}_o)\widehat{u}^1_{-1}}{u_o- - \widehat{u}^1_{-1}} + \frac{K_o u_o- + \widehat{f}_o - \widehat{f}^1_{-1}}{u_o- - \widehat{u}^1_{-1}}\,\widehat{u}$$
$$\frac{\partial f_1(.)}{\partial \widehat{u}^1_{-1}} = \frac{(\widehat{f}^1_{-1} - K_o u_o- - \widehat{f}_o)u_o-}{(u_o- - \widehat{u}^1_{-1})^2} + \frac{K_o u_o- + \widehat{f}_o - \widehat{f}^1_{-1}}{(u_o- - \widehat{u}^1_{-1})^2}\,\widehat{u}$$
$$\frac{\partial f_1(.)}{\partial \widehat{u}} = \frac{K_o u_o- + \widehat{f}_o - \widehat{f}^1_{-1}}{u_o- - \widehat{u}^1_{-1}}$$
$$\frac{\partial f_1(.)}{\partial \widehat{f}^1_{-1}} = \frac{u_o- - \widehat{u}}{u_o- - \widehat{u}^1_{-1}}, \qquad \frac{\partial f_1(.)}{\partial \widehat{f}_o} = \frac{\widehat{u} - \widehat{u}^1_{-1}}{u_o- - \widehat{u}^1_{-1}}$$
$$\frac{\partial f_1(.)}{\partial \widehat{u}^1_l} = 0, \quad l \neq -1, \qquad \frac{\partial f_1(.)}{\partial \widehat{f}^1_l} = 0, \quad l \neq -1, 0$$

end

when $\widehat{u}(t+1) \in [u_o-, u_o+]$

$$\widehat{y}(t+1) = K_o\widehat{u} + \widehat{f}_o$$
$$\frac{\partial f_1(.)}{\partial \widehat{u}} = K_o, \qquad \frac{\partial f_1(.)}{\partial \widehat{f}_o} = 1$$
$$\frac{\partial f_1(.)}{\partial \widehat{f}^1_l} = 0, \quad l \neq 0, \qquad \frac{\partial f_1(.)}{\partial \widehat{u}^1_l} = 0, \quad \forall\, l$$

end

when $\widehat{u}(t+1) \in [u_o+, \widehat{u}^1_1]$

$$\widehat{y}(t+1) = \frac{(K_o u_o+ + \widehat{f}_o)\widehat{u}^1_1 - \widehat{f}^1_1 u_o+}{\widehat{u}^1_1 - u_o+} + \frac{\widehat{f}^1_1 - K_o u_o+ - \widehat{f}_o}{\widehat{u}^1_1 - u_o+}\,\widehat{u}$$
$$\frac{\partial f_1(.)}{\partial \widehat{u}^1_1} = \frac{(\widehat{f}^1_1 - K_o u_o+ - \widehat{f}_o)u_o+}{(\widehat{u}^1_1 - u_o+)^2} - \frac{\widehat{f}^1_1 - K_o u_o+ - \widehat{f}_o}{(\widehat{u}^1_1 - u_o+)^2}\,\widehat{u}$$
$$\frac{\partial f_1(.)}{\partial \widehat{u}} = \frac{\widehat{f}^1_1 - K_o u_o+ - \widehat{f}_o}{\widehat{u}^1_1 - u_o+}$$
$$\frac{\partial f_1(.)}{\partial \widehat{f}_o} = \frac{\widehat{u}^1_1 - \widehat{u}}{\widehat{u}^1_1 - u_o+}, \qquad \frac{\partial f_1(.)}{\partial \widehat{f}^1_1} = \frac{\widehat{u} - u_o+}{\widehat{u}^1_1 - u_o+}$$
$$\frac{\partial f_1(.)}{\partial \widehat{u}^1_l} = 0, \quad l \neq 1, \qquad \frac{\partial f_1(.)}{\partial \widehat{f}^1_l} = 0, \quad l \neq 0, 1$$

end

when $\widehat{u}(t+1) \in I_j, \qquad j = 1,\cdots,L$

when $\widehat{u}(t+1) \in [\widehat{u}^j_i, \widehat{u}^j_{i+1}]$

$$\forall\, i = \begin{cases} -k^-_1, .., -2, 1, .., k^+_1 - 1, & j = 1 \\ -k^-_j, \cdots, k^+_j - 1, & j = 2, \cdots, L. \end{cases}$$

$$\widehat{y}(t+1) = \frac{\widehat{f}^j_i \widehat{u}^j_{i+1} - \widehat{f}^j_{i+1}\widehat{u}^j_i}{\widehat{u}^j_{i+1} - \widehat{u}^j_i} + \frac{\widehat{f}^j_{i+1} - \widehat{f}^j_i}{\widehat{u}^j_{i+1} - \widehat{u}^j_i}\,\widehat{u}$$
$$\frac{\partial f_j(.)}{\partial \widehat{u}^j_i} = \frac{(\widehat{f}^j_i - \widehat{f}^j_{i+1})\widehat{u}^j_{i+1}}{(\widehat{u}^j_{i+1} - \widehat{u}^j_i)^2} + \frac{\widehat{f}^j_{i+1} - \widehat{f}^j_i}{(\widehat{u}^j_{i+1} - \widehat{u}^j_i)^2}\,\widehat{u}$$
$$\frac{\partial f_j(.)}{\partial \widehat{u}^j_{i+1}} = \frac{(\widehat{f}^j_{i+1} - \widehat{f}^j_i)\widehat{u}^j_i}{(\widehat{u}^j_{i+1} - \widehat{u}^j_i)^2} - \frac{\widehat{f}^j_{i+1} - \widehat{f}^j_i}{(\widehat{u}^j_{i+1} - \widehat{u}^j_i)^2}\,\widehat{u}$$
$$\frac{\partial f_j(.)}{\partial \widehat{u}} = \frac{\widehat{f}^j_{i+1} - \widehat{f}^j_i}{\widehat{u}^j_{i+1} - \widehat{u}^j_i}$$
$$\frac{\partial f_j(.)}{\partial \widehat{f}^j_i} = \frac{\widehat{u}^j_{i+1} - \widehat{u}}{\widehat{u}^j_{i+1} - \widehat{u}^j_i}, \qquad \frac{\partial f_j(.)}{\partial \widehat{f}^j_{i+1}} = \frac{\widehat{u} - \widehat{u}^j_i}{\widehat{u}^j_{i+1} - \widehat{u}^j_i}$$
$$\frac{\partial f_j(.)}{\partial \widehat{u}^j_l} = 0, \; l \neq i, i+1, \qquad \frac{\partial f_j(.)}{\partial \widehat{f}^j_l} = 0, \; l \neq i, i+1$$

end

$$\frac{\partial f_j(.)}{\partial \widehat{\boldsymbol{\theta}}_{\boldsymbol{j}}} = \left(\frac{\partial f_j(.)}{\partial \widehat{f}^j_{-k^-_j}} \cdots \frac{\partial f_j(.)}{\partial \widehat{f}^j_{-1}} \;\; \frac{\partial f_j(.)}{\partial \widehat{f}^j_1} \cdots \frac{\partial f_j(.)}{\partial \widehat{f}^j_{k^+_j}} \right)$$
$$\frac{\partial f_j(.)}{\partial \widehat{\boldsymbol{g}}_{\boldsymbol{j}}} = \left(\frac{\partial f_j(.)}{\partial \widehat{u}^j_{-k^-_j}} \cdots \frac{\partial f_j(.)}{\partial \widehat{u}^j_{-1}} \;\; \frac{\partial f_j(.)}{\partial \widehat{u}^j_1} \cdots \frac{\partial f_j(.)}{\partial \widehat{u}^j_{k^+_j}} \right)$$
$$\boldsymbol{\psi}^T(t+1) =$$
$$\left(\frac{\partial f_j(.)}{\partial \widehat{u}}\boldsymbol{\psi_l}^T(t+1) \;\; \frac{\partial f_j(.)}{\partial \widehat{f}_o} 0..0 \frac{\partial f_j(.)}{\partial \widehat{\boldsymbol{\theta}}_{\boldsymbol{j}}} 0..0 \frac{\partial f_j(.)}{\partial \widehat{\boldsymbol{g}}_{\boldsymbol{j}}} 0..0 \right)$$

end

Here D_M indicates that the algorithms described in (Ljung and Söderström, 1983) are used to keep the predictor in the model set.

Remark 3: It is considered here that some precautions are taken to prevent grid miss ordering during the estimation process. This is stated as

C3) Grid ordering is included in the definition of the model set.

3. THE CRAMÉR-RAO BOUND

In this section the CRB of the proposed parameterization is calculated. Introduce the following condition:

C4) The linear block and the static nonlinearity of the system are contained in the model set.

$$
\boldsymbol{I}(t)=\left(\begin{array}{cccccccccc}
I_{X,X} & I_{X,w} & I_{X,f_o} & 0\cdots 0 & I_{X,f_i^j} & I_{X,f_{i+1}^j} & 0\cdots 0 & I_{X,u_i^j} & I_{X,u_{i+1}^j} & 0\cdots 0\\
I_{w,X} & I_{w,w} & I_{w,f_o} & 0\cdots 0 & I_{w,f_i^j} & I_{w,f_{i+1}^j} & 0\cdots 0 & I_{w,u_i^j} & I_{w,u_{i+1}^j} & 0\cdots 0\\
I_{f_o,X} & I_{f_o,w} & I_{f_o,f_o} & 0\cdots 0 & I_{f_o,f_i^j} & I_{f_o,f_{i+1}^j} & 0\cdots 0 & I_{f_o,u_i^j} & I_{f_o,u_{i+1}^j} & 0\cdots 0\\
0 & 0 & 0 & & 0 & 0 & & 0 & 0 & \\
\vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & \\
0 & 0 & 0 & & 0 & 0 & & 0 & 0 & \\
I_{f_i^j,X} & I_{f_i^j,w} & I_{f_i^j,f_o} & 0\cdots 0 & I_{f_i^j,f_i^j} & I_{f_i^j,f_{i+1}^j} & 0\cdots 0 & I_{f_i^j,u_i^j} & I_{f_i^j,u_{i+1}^j} & 0\cdots 0\\
I_{f_{i+1}^j,X} & I_{f_{i+1}^j,w} & I_{f_{i+1}^j,f_o} & 0\cdots 0 & I_{f_{i+1}^j,f_i^j} & I_{f_{i+1}^j,f_{i+1}^j} & 0\cdots 0 & I_{f_{i+1}^j,u_i^j} & I_{f_{i+1}^j,u_{i+1}^j} & 0\cdots 0\\
0 & 0 & 0 & & 0 & 0 & & 0 & 0 & \\
\vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & \\
0 & 0 & 0 & & 0 & 0 & & 0 & 0 & \\
I_{u_i^j,X} & I_{u_i^j,w} & I_{u_i^j,f_o} & 0\cdots 0 & I_{u_i^j,f_i^j} & I_{u_i^j,f_{i+1}^j} & 0\cdots 0 & I_{u_i^j,u_i^j} & I_{u_i^j,u_{i+1}^j} & 0\cdots 0\\
I_{u_{i+1}^j,X} & I_{u_{i+1}^j,w} & I_{u_{i+1}^j,f_o} & 0\cdots 0 & I_{u_{i+1}^j,f_i^j} & I_{u_{i+1}^j,f_{i+1}^j} & 0\cdots 0 & I_{u_{i+1}^j,u_i^j} & I_{u_{i+1}^j,u_{i+1}^j} & 0\cdots 0\\
0 & 0 & 0 & & 0 & 0 & & 0 & 0 & \\
\vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & \\
0 & 0 & 0 & & 0 & 0 & & 0 & 0 &
\end{array}\right) \tag{10}
$$

Then there are vectors $\boldsymbol{\theta^o}$ such that the output of the static nonlinearity of the system is described by

$$
y(t)=f_j(\boldsymbol{\theta_n^o},\boldsymbol{\theta_g^o},\widehat{u}(t,\boldsymbol{\theta_l^o}))+w(t),\ \widehat{u}\in I_j. \tag{11}
$$

where *w(t)* is the disturbance which satisfies the following condition,

C5) $E[w(t)w(s)]=\sigma^2\delta_{t,s}$ and $w(t)$ is zero mean Gaussian.

Furthermore, introduce the following condition,

C6) $N>N_o$ such that there exist a time instant $t<N_o$ where $\widehat{u}\in I_j$ and $\widehat{u}\in[u_i^j,u_{i+1}^j]$ $\forall i,j\in\{i=-k_j^-,\dots,k_j^+-1,\ j=1,\dots,L\}$.

***Remark 4*:** **C6)** means that there is signal energy in each subinterval of the model, cf. (Wigren and Händel, 1996).

Then the following theorem holds

***Theorem* 1.** Under the Conditions **C1)-C6)**, the CRB for $(\boldsymbol{\theta_l}^T\ \boldsymbol{\theta_n}^T\ \boldsymbol{\theta_g}^T)^T$ is given by

$$
CRB(\boldsymbol{\theta})=\sigma^2\left(\sum_{t=1}^{N}\boldsymbol{I}(t)\right)^{-1} \tag{12}
$$

where $\boldsymbol{I}(t)$ given by (10) for $\widehat{u}(t,\boldsymbol{\theta_l})\in[u_i^j,u_{i+1}^j]\in I_j,\ i=-k_j^-,\dots,k_j^+-1,\ j=1,\dots,L$ and

$$
\begin{aligned}
I_{X,X}&=\left[\frac{\partial f_j(.)}{\partial u}\right]^2\ \Lambda^2(\phi)\\
I_{w,w}&=\left[\frac{\partial f_j(.)}{\partial u}\right]^2\ X^2t^2\left[\frac{d\Lambda(\phi)}{d\phi}\right]^2\\
I_{f_o,f_o}&=\left[\frac{\partial f_j(.)}{\partial f_o}\right]^2\\
I_{X,w}&=\left[\frac{\partial f_j(.)}{\partial u}\right]^2\ Xt\ \Lambda(\phi)\frac{d\Lambda(\phi)}{d\phi}\\
I_{X,f_o}&=\frac{\partial f_j(.)}{\partial u}\frac{\partial f_j(.)}{\partial f_o}\Lambda(\phi)
\end{aligned}
$$

$$
\begin{aligned}
I_{w,f_o}&=\frac{\partial f_j(.)}{\partial u}\frac{\partial f_j(.)}{\partial f_o}Xt\frac{d\Lambda(\phi)}{d\phi}\\
I_{X,f_i^j}&=\frac{\partial f_j(.)}{\partial u}\frac{\partial f_j(.)}{\partial f_i^j}\Lambda(\phi)\\
I_{X,f_{i+1}^j}&=\frac{\partial f_j(.)}{\partial u}\frac{\partial f_j(.)}{\partial f_{i+1}^j}\Lambda(\phi)\\
I_{X,u_i^j}&=\frac{\partial f_j(.)}{\partial u}\frac{\partial f_j(.)}{\partial u_i^j}\Lambda(\phi)\\
I_{X,u_{i+1}^j}&=\frac{\partial f_j(.)}{\partial u}\frac{\partial f_j(.)}{\partial u_{i+1}^j}\Lambda(\phi)\\
I_{w,f_i^j}&=\frac{\partial f_j(.)}{\partial u}\frac{\partial f_j(.)}{\partial f_i^j}Xt\frac{d\Lambda(\phi)}{d\phi}\\
I_{w,f_{i+1}^j}&=\frac{\partial f_j(.)}{\partial u}\frac{\partial f_j(.)}{\partial f_{i+1}^j}Xt\frac{d\Lambda(\phi)}{d\phi}\\
I_{w,u_i^j}&=\frac{\partial f_j(.)}{\partial u}\frac{\partial f_j(.)}{\partial u_i^j}Xt\frac{d\Lambda(\phi)}{d\phi}\\
I_{w,u_{i+1}^j}&=\frac{\partial f_j(.)}{\partial u}\frac{\partial f_j(.)}{\partial u_{i+1}^j}Xt\frac{d\Lambda(\phi)}{d\phi}\\
I_{f_o,f_i^j}&=\frac{\partial f_j(.)}{\partial f_o}\frac{\partial f_j(.)}{\partial f_i^j}\\
I_{f_o,f_{i+1}^j}&=\frac{\partial f_j(.)}{\partial f_o}\frac{\partial f_j(.)}{\partial f_{i+1}^j}\\
I_{f_o,u_i^j}&=\frac{\partial f_j(.)}{\partial f_o}\frac{\partial f_j(.)}{\partial u_i^j}\\
I_{f_o,u_{i+1}^j}&=\frac{\partial f_j(.)}{\partial f_o}\frac{\partial f_j(.)}{\partial u_{i+1}^j}\\
I_{f_i^j,f_i^j}&=\left[\frac{\partial f_j(.)}{\partial f_i^j}\right]^2\\
I_{f_{i+1}^j,f_{i+1}^j}&=\left[\frac{\partial f_j(.)}{\partial f_{i+1}^j}\right]^2\\
I_{f_i^j,f_{i+1}^j}&=\frac{\partial f_j(.)}{\partial f_i^j}\frac{\partial f_j(.)}{\partial f_{i+1}^j}\\
I_{u_i^j,u_i^j}&=\left[\frac{\partial f_j(.)}{\partial u_i^j}\right]^2
\end{aligned}
$$

$$
\begin{aligned}
I_{u_{i+1}^j, u_{i+1}^j} &= \left[\frac{\partial f_j(.)}{\partial u_{i+1}^j}\right]^2 \\
I_{u_i^j, u_{i+1}^j} &= \frac{\partial f_j(.)}{\partial u_i^j}\frac{\partial f_j(.)}{\partial u_{i+1}^j} \\
I_{f_i^j, u_i^j} &= \frac{\partial f_j(.)}{\partial f_i^j}\frac{\partial f_j(.)}{\partial u_i^j} \\
I_{f_i^j, u_{i+1}^j} &= \frac{\partial f_j(.)}{\partial f_i^j}\frac{\partial f_j(.)}{\partial u_{i+1}^j} \\
I_{f_{i+1}^j, u_i^j} &= \frac{\partial f_j(.)}{\partial f_{i+1}^j}\frac{\partial f_j(.)}{\partial u_i^j} \\
I_{f_{i+1}^j, u_{i+1}^j} &= \frac{\partial f_j(.)}{\partial f_{i+1}^j}\frac{\partial f_j(.)}{\partial u_{i+1}^j} \\
\phi &= \omega t
\end{aligned} \tag{13}
$$

***Proof*:** See Appendix A. □

4. NUMERICAL EXAMPLES

The adaptive grid point algorithm was studied by numerical examples in (Abd-Elrady, 2001*a*) to investigate its local convergence and the ability to track both the damped amplitude and the frequency variations. Also, the following examples were performed.

***Example* 1: Comparison with the fixed grid point algorithm.**

In order to compare the performance of the adaptive grid point algorithm with the fixed grid point algorithm (cf. (Abd-Elrady, 2000)), 100 Monte Carlo simulations were performed with different noise realizations. The data were generated according to: the driving wave was given by $u(t, \boldsymbol{\theta_l}) = \sin \omega t$ where $\omega = 2\pi \times 0.05$. The static nonlinearity was chosen as

$$
f(u) = \begin{cases} (5/3)u^2 + 0.15 & u \geq 0.3 \\ u & -0.3 \leq u < 0.3 \\ -(5/3)u^2 - 0.15 & u < -0.3 \end{cases} \tag{14}
$$

The algorithms were initialized with $\lambda(0) = 0.95$, $\lambda_o = 0.99$, $\boldsymbol{P}(0) = 0.01I$, $K_o = 1$, $X = 1$, $f_o = 0$ and $\omega(0) = 2\pi \times 0.02$. Further, two static nonlinearities ($L = 2$) were used, where $\widehat{u}(t, \boldsymbol{\theta_l}) \in I_1$ for positive slopes and $\widehat{u}(t, \boldsymbol{\theta_l}) \in I_2$ for negative slopes, respectively. The nonlinearities were initialized as straight lines with unity slope in the following grid points

$$
\begin{aligned}
\boldsymbol{g_1} &= (-2, , -1, -0.3, -0.1, 0.1, 0.3, 1, 2), \\
\boldsymbol{g_2} &= (-2, , -1, -0.3, 0.3, 1, 2).
\end{aligned} \tag{15}
$$

The mean square error (MSE) for the two algorithms for different signal to noise ratios (SNR) was calculated. The results are plotted in Fig. 3 which shows that the adaptive grid point algorithm gives lower MSE than the fixed grid point algorithm for moderate and high SNR. This results indicates that the modeling error is lower for the adaptive grid point algorithm.

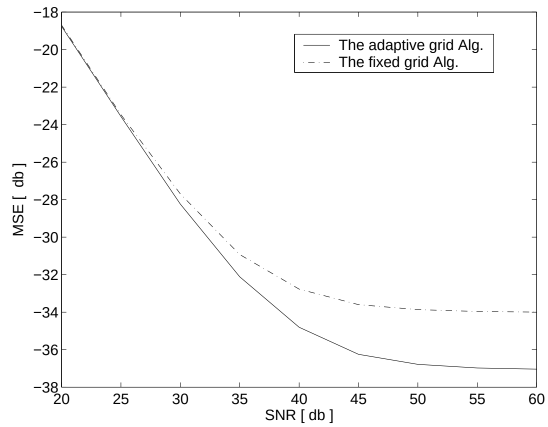

Fig. 3. The MSE results.

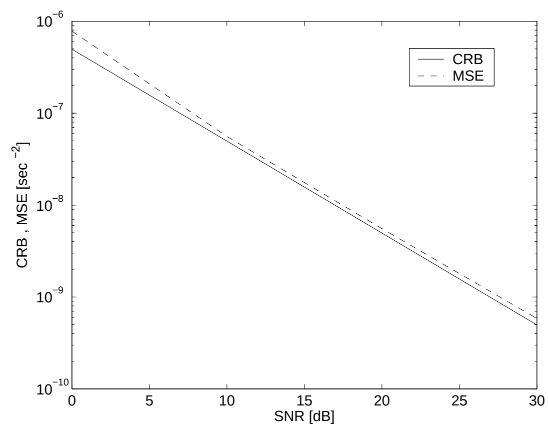

Fig. 4. Statistical results.

***Example* 2: Performance of the adaptive grid point algorithm as compared to the Cramér-Rao bound (CRB).**

In order to compare the performance of the adaptive grid point algorithm with the derived CRB for the fundamental frequency estimation, 100 Monte Carlo simulations were performed with different noise realizations. The data were generated as in Example 1 with a static nonlinearity given by

$$
\begin{aligned}
&\boldsymbol{g_1} = (-1, -0.3, -0.15, 0.15, 0.3, 1), \\
&\boldsymbol{g_2} = (-1, -0.3, 0.3, 1), \\
&\boldsymbol{\theta_1} = (-0.8, -0.3, 0.3, 0.8),\ \widehat{u}(t, \boldsymbol{\theta_l}) \in I_1 \\
&\boldsymbol{\theta_2} = (-0.8, -0.5, 0.5, 0.8),\ \widehat{u}(t, \boldsymbol{\theta_l}) \in I_2.
\end{aligned} \tag{16}
$$

Also, the algorithm was initialized as in Example 1 except that $\boldsymbol{P} = 0.0001I$. The statistics is based on excluding simulations that did not satisfy a margin of 5 standard deviations (as predicted by the CRB) from the true fundamental frequency. Both the CRB for the fundamental frequency estimate and the MSE value were evaluated for different SNR. The statistical results are plotted in Fig. 4 which shows that the adaptive grid point algorithm gives good results, in particular for moderate values of the SNR.

5. CONCLUSIONS

A recursive harmonic signal estimation scheme has been presented. It estimates the grid points in addition to the fundamental frequency and the parameters of the static nonlinearity. Local convergence of the suggested algorithm was investigated by numerical examples and the CRB was calculated for this algorithm. Monte Carlo experiments show that the suggested algorithm gives significantly better results than using fixed grid point algorithms.

ACKNOWLEDGMENT

The author would like to thank Professor Torbjörn Wigren and Professor Torsten Söderström for many valuable comments on this work.

6. REFERENCES

Abd-Elrady, E. (2000). Convergence of the RPEM as applied to harmonic signal modeling. Technical Report 2000-027. Information Technology, Uppsala University, Uppsala, Sweden.

Abd-Elrady, E. (2001*a*). An adaptive grid point algorithm for harmonic signal modeling. Technical Report 2001-018. Information Technology, Uppsala University, Uppsala, Sweden.

Abd-Elrady, E. (2001*b*). Study of a nonlinear recursive method for harmonic signal modeling. In: *The Twentieth IASTED International Conference on Modeling, Identification and Control.* Innsbruck, Austria.

Händel, P. and P. Tichavskỳ (1994). Adaptive estimation for periodic signal enhancement and tracking. *Int. J. Adaptive Control, Signal Process., 8, pp. 447-456.*

Ljung, L. and T. Söderström (1983). *Theory and Practice of Recursive Identification.* M.I.T. Press, Cambridge, MA.

Nehorai, A. and B. Porat (1986). Adaptive comb filtering for harmonic signal enhancement. *IEEE Trans. Acoustics, Speech, Signal Processing* **ASSP-34**, 1124–1138.

Söderström, T. and P. Stoica (1989). *System Identification.* Hemel Hempstead, UK: Prentice-Hall.

Wigren, T. (1993). Recursive prediction error identification using the nonlinear Wiener model. *Automatica* **29**, 1011–1025.

Wigren, T. and P. Händel (1996). Harmonic signal modeling using adaptive nonlinear function estimation. In: *IEEE International Conference on Acoustics, Speech and Signal Processing.* Atlanta, GA.

APPENDIX A: PROOF OF THEOREM 1

The log-likelihood function is given by

$$l(\boldsymbol{\theta}) = \kappa - \frac{1}{2\sigma^2} \sum_{t=1}^{N} (y(t) - \widehat{y}(t, \boldsymbol{\theta}))^2 \tag{17}$$

where κ is a constant. Let,

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_l}} & \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_n}} & \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_g}} \end{pmatrix} \tag{18}$$

where

$$\begin{aligned}
\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_l}} &= \begin{pmatrix} \frac{\partial l(\boldsymbol{\theta})}{\partial X} & \frac{\partial l(\boldsymbol{\theta})}{\partial \omega} \end{pmatrix} \\
\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_n}} &= \begin{pmatrix} \frac{\partial l(\boldsymbol{\theta})}{\partial f_o} & \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_1}} & \cdots & \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_L}} \end{pmatrix} \\
\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_j}} &= \begin{pmatrix} \frac{\partial l(\boldsymbol{\theta})}{\partial f^j_{-k_j^-}} & \cdots & \frac{\partial l(\boldsymbol{\theta})}{\partial f^j_{-1}} & \frac{\partial l(\boldsymbol{\theta})}{\partial f^j_{1}} & \cdots & \frac{\partial l(\boldsymbol{\theta})}{\partial f^j_{k_j^+}} \end{pmatrix} \\
\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_g}} &= \begin{pmatrix} \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{g_1}} & \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{g_2}} & \cdots & \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{g_L}} \end{pmatrix} \\
\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{g_j}} &= \begin{pmatrix} \frac{\partial l(\boldsymbol{\theta})}{\partial u^j_{-k_j^-}} & \cdots & \frac{\partial l(\boldsymbol{\theta})}{\partial u^j_{-1}} & \frac{\partial l(\boldsymbol{\theta})}{\partial u^j_{1}} & \cdots & \frac{\partial l(\boldsymbol{\theta})}{\partial u^j_{k_j^+}} \end{pmatrix}
\end{aligned} \tag{19}$$

Then, the Fisher informatiom matrix (Söderström and Stoica, 1989) can be written as

$$\begin{aligned}
\boldsymbol{J} &= -E \, \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}^T \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \\
&= -E \begin{pmatrix}
\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_l}}^T \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_l}} & \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_l}}^T \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_n}} & \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_l}}^T \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_g}} \\
\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_n}}^T \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_l}} & \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_n}}^T \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_n}} & \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_n}}^T \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_g}} \\
\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_g}}^T \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_l}} & \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_g}}^T \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_n}} & \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_g}}^T \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_g}}
\end{pmatrix}
\end{aligned} \tag{20}$$

In order to calculate $\boldsymbol{J}$, note that For $\widehat{u}(t, \boldsymbol{\theta_l}) \in [u_i^j, u_{i+1}^j] \in I_j$, $i = -k_j^-, \ldots, k_j^+ - 1$, $j = 1, \ldots, L$, it holds that

$$\begin{aligned}
\frac{\partial \widehat{y}(t, \boldsymbol{\theta})}{\partial X} &= \frac{\partial f_j(.)}{\partial u} \, \Lambda(\phi) \\
\frac{\partial \widehat{y}(t, \boldsymbol{\theta})}{\partial w} &= \frac{\partial f_j(.)}{\partial u} \, Xt \frac{d\Lambda(\phi)}{d\phi}
\end{aligned} \tag{21}$$

Thus using the signal model **C4)** and **C5)**, the blocks of (20) can be evaluated as in (Abd-Elrady, 2001*a*). Introduce the notation in (13) and use the facts that $\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta_m} \partial \boldsymbol{\theta_n}} = 0$ and $\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{g_m} \partial \boldsymbol{g_n}} = 0$ for $m \neq n$ and $\boldsymbol{J} = 1/\sigma^2 \sum_{t=1}^{N} \boldsymbol{I}(t)$. Then (12) directly follows from **C6)**, see (Abd-Elrady, 2001*a*).