Signal Tracking Properties of A Class of Adaptive Notch Filters

Maciej Niedźwiecki and Piotr Kaczmarek *
Faculty of Electronics, Telecommunications and Computer Science
Department of Automatic Control, Gdańsk University of Technology
ul. Narutowicza 11/12, Gdańsk , Poland
maciekn@eti.pg.gda.pl, piokacz@proterians.net.pl

ABSTRACT

The signal tracking properties of two adaptive notch filtering algorithms are studied analytically using a linear filter approximation technique. Even though restricted to a single frequency case, the presented analysis provides valuable insights into the tracking mechanisms, including the speed/accuracy tradeoffs, the achievable performance bounds, and tracking limitations of the analyzed algorithms. Additionally, it allows one to formulate some useful rules of thumb for choosing design parameters.

1 Introduction

Consider the problem of elimination or extraction of a nonstationary sinusoidal signal $s(t)$ buried in noise

$$y(t) = s(t) + v(t) = \sum_{i=1}^{k} a_i(t) e^{j\sum_{s=1}^{t}\omega_i(s)} + v(t). \qquad (1)$$

where $v(t)$ is a complex white noise of variance σ_v^2. We will assume that $\mathrm{E}[v_R^2(t)] = \mathrm{E}[v_R^2(t)] = \sigma_v^2/2$, $\mathrm{E}[v_R(t)v_I(s)] = 0,\ \forall\ t,s$, where $v_R(t) = \mathrm{Re}[v(t)]$, $v_I(t) = \mathrm{Im}[v(t)]$, and that the (complex) amplitudes $a_i(t)$ and frequencies $\omega_i(t)$ in (1) are slowly time-varying.
The problem of elimination and extraction of complex sinusoidal signals (called cisoids) embedded in noise was considered by many authors - see e.g. [3], [4] and the references therein. In this paper we will compare signal tracking properties of two adaptive notch filters: the algorithm proposed in [1] and the multiple frequency tracker, described in [3]. Unlike the earlier studies [3], [4], which focused on the problem of frequency tracking, our primary interest will be in signal tracking characteristics of the compared filters.
The adaptive notch filter proposed in [1] combines the exponentially weighted least squares approach to amplitude tracking with gradient search approach to frequency tracking

$$\begin{aligned}
\widehat{f}_i(t) &= e^{j\widehat{\omega}_i(t)}\widehat{f}_i(t-1) \\
i &= 1,\ldots,k \\
\varepsilon(t) &= y(t) - \widehat{\mathbf{f}}^{\mathrm{T}}(t)\widehat{\boldsymbol{\alpha}}(t-1) \\
\mathbf{Q}(t) &= \frac{1}{\lambda}\left[\mathbf{Q}(t-1) - \frac{\mathbf{Q}(t-1)\widehat{\mathbf{f}}(t)\widehat{\mathbf{f}}^{\mathrm{H}}(t)\mathbf{Q}(t-1)}{\lambda + \widehat{\mathbf{f}}^{\mathrm{H}}(t)\mathbf{Q}(t-1)\widehat{\mathbf{f}}(t)}\right] \\
\mathbf{k}(t) &= \mathbf{Q}(t)\widehat{\mathbf{f}}(t) \\
\widehat{\boldsymbol{\alpha}}(t) &= \widehat{\boldsymbol{\alpha}}(t-1) + \mathbf{k}^*(t)\varepsilon(t) \\
g_i(t) &= \mathrm{Im}\{\varepsilon^*(t)\widehat{f}_i(t)\widehat{a}_i(t-1)\} \\
\widehat{\omega}_i(t+1) &= \widehat{\omega}_i(t) - \eta g_i(t) \\
i &= 1,\ldots,k \\
\widehat{s}(t) &= \sum_{i=1}^{k}\widehat{f}_i(t)\widehat{a}_i(t) \qquad (2)
\end{aligned}$$

where $\widehat{\boldsymbol{\alpha}}(t) = [\widehat{a}_1(t),\ldots,\widehat{a}_k(t)]^{\mathrm{T}}$ and $\widehat{\mathbf{f}}(t) = [\widehat{f}_1(t),\ldots,\widehat{f}_k(t)]^{\mathrm{T}}$.
In the above algorithm λ $(0 < \lambda < 1)$, usually set close to one, denotes the so-called forgetting constant, which controls the rate of amplitude adaptation, and $\eta > 0$, usually set close to zero, denotes the stepsize coefficient, which controls the rate of frequency adaptation.
The initial conditions for (2) should be set to $\widehat{\boldsymbol{\alpha}}(0) = \mathbf{0}$ and $\mathbf{Q}(0) = c\mathbf{I}_k$, where $\mathbf{I}_k$ denotes the $k \times k$ identity matrix and c is a large positive constant - this is a standard initialization procedure for all RLS-type recursive estimation algorithms [5].
The multiple frequency tracker, which bears some resemblance to (2), will be described in Section 2.
It is worth noticing that the algorithm (2) is a special (signal) case of a generalized adaptive notch filtering algorithm proposed in [2] for the purpose of identification/tracking of quasi-periodically varying complex-valued dynamic systems.

2 Tracking analysis

Before we start analyzing tracking properties of the algorithm (2), we will convert it into a more convenient form by applying the linear time-varying transformation: $\widehat{\boldsymbol{\beta}}(t) = \widehat{\mathbf{F}}(t)\ \widehat{\boldsymbol{\alpha}}(t)$, $\mathbf{l}(t) = \widehat{\mathbf{F}}^*(t)\ \mathbf{k}(t)$ and $\mathbf{P}(t) =$

* This work was supported by KBN under Grant 4 T11A 01225.

$\widehat{\mathbf{F}}^*(t)\mathbf{Q}(t)\widehat{\mathbf{F}}(t)$, where $\widehat{\mathbf{F}}(t) = \mathrm{diag}\{\widehat{f}_1(t),\ldots,\widehat{f}_k(t)\}$. Using this transformation and setting $\widehat{\mathbf{A}}(t) = \mathrm{diag}\{e^{j\widehat{\omega}_1(t)},\ldots,e^{j\widehat{\omega}_k(t)}\}$ one can rewrite (2) in the following form

$$\begin{aligned}
\varepsilon(t) &= y(t) - \mathbf{1}_k^{\mathrm{T}}\widehat{\mathbf{A}}(t)\widehat{\boldsymbol{\beta}}(t-1)\\
\mathbf{P}(t) &= \frac{1}{\lambda}\,\widehat{\mathbf{A}}^*(t)\left[\mathbf{P}(t-1)\right.\\
&- \left.\frac{\mathbf{P}(t-1)\mathbf{1}_k\mathbf{1}_k^{\mathrm{T}}\mathbf{P}(t-1)}{\lambda+\mathbf{1}_k^{\mathrm{T}}\mathbf{P}(t-1)\mathbf{1}_k}\right]\widehat{\mathbf{A}}(t)\\
\mathbf{l}(t) &= \mathbf{P}(t)\mathbf{1}_k\\
\widehat{\boldsymbol{\beta}}(t) &= \widehat{\mathbf{A}}(t)\widehat{\boldsymbol{\beta}}(t-1) + \mathbf{l}^*(t)\varepsilon(t)\\
g_i(t) &= \mathrm{Im}\{\varepsilon^*(t)e^{j\widehat{\omega}_i(t)}\widehat{\beta}_i(t-1)\}\\
\widehat{\omega}_i(t+1) &= \widehat{\omega}_i(t) - \eta g_i(t)\\
i &= 1,\ldots,k\\
\widehat{s}(t) &= \sum_{i=1}^{k}\widehat{\beta}_i(t)
\end{aligned}\tag{3}$$

where $\widehat{\boldsymbol{\beta}}(t) = [\widehat{\beta}_1(t),\ldots,\widehat{\beta}_k(t)]^{\mathrm{T}}$, $\widehat{\beta}_i(t) = \widehat{f}_i(t)\widehat{a}_i(t)$, $i = 1,\ldots,k$ and $\mathbf{1}_k = [\underbrace{1,\ldots,1}_{k}]^{\mathrm{T}}$.

It should be stressed, that the algorithms (2) and (3) are strictly input-output equivalent, i.e. when started with the same initial conditions ($\boldsymbol{\beta}(0) = \boldsymbol{\alpha}(0), \mathbf{P}(0) = \mathbf{Q}(0)$) they yield identical signal estimates $\widehat{s}(t)$.

As is straightforward to check, the algorithm (3) is almost identical with the algorithm known as multiple frequency tracker (MFT), proposed by Tichavský and Händel in their seminal paper [3] (see equations (9) - (11) in [3]). It turns out that the only difference lies in the frequency update mechanism, which in the case of MFT has the form (in our notation)

$$\begin{aligned}
g_i(t) &= \mathrm{Arg}\left[\frac{\widehat{\beta}_i(t)}{\widehat{\beta}_i(t-1)e^{j\widehat{\omega}_i(t)}}\right]\\
\widehat{\omega}_i(t+1) &= \widehat{\omega}_i(t) - \eta g_i(t)\\
i &= 1,\ldots,k
\end{aligned}$$

We will analyze (3) using the approximating linear filter (ALF) technique, introduced in [3]. Approximating linear filters characterize the relation between the sequences of estimation errors and the sequences of measurement noise $v(t)$ and of the one-step changes of the true frequency $\omega(t+1) - \omega(t)$, provided that the analyzed algorithms operate in a neighborhood of their equilibrium state.

Similarly as in [3], we will consider the single frequency case ($k = 1$) and steady state tracking conditions. Note that for $k = 1$, the scalar (1×1) counterpart of the matrix $\mathbf{P}(t)$, denoted by $\mathrm{p}(t)$, tends to a constant steady state value $\mathrm{p}(\infty) = \lim_{t\mapsto\infty}\mathrm{p}(t) = 1-\lambda = \mu$. Hence, in the case considered, one can rewrite (3) in a much simpler form

$$\begin{aligned}
\varepsilon(t) &= y(t) - e^{j\widehat{\omega}(t)}\widehat{\beta}(t-1)\\
\widehat{\beta}(t) &= e^{j\widehat{\omega}(t)}\widehat{\beta}(t-1) + \mu\varepsilon(t)\\
g(t) &= \mathrm{Im}\{\varepsilon^*(t)e^{j\widehat{\omega}(t)}\widehat{\beta}(t-1)\}\\
\widehat{\omega}(t+1) &= \widehat{\omega}(t) - \eta g(t)\\
\widehat{s}(t) &= \widehat{\beta}(t)
\end{aligned}\tag{4}$$

For MFT the analogous equations are identical with (4), except that

$$g(t) = \mathrm{Arg}\left[\frac{\widehat{\beta}(t)}{\widehat{\beta}(t-1)e^{j\widehat{\omega}(t)}}\right]\tag{5}$$

Denote by $\Delta\widehat{\beta}(t) = \widehat{\beta}(t) - \beta(t) = \widehat{s}(t) - s(t)$ the signal estimation error and let

$$\begin{aligned}
\Delta\widehat{\phi}(t) &= \beta^*(t)\Delta\widehat{\beta}(t) = \Delta\widehat{\phi}_R(t) + j\Delta\widehat{\phi}_I(t)\\
e(t) &= \beta^*(t)v(t) = e_R(t) + je_I(t)\\
w(t+1) &= \omega(t+1) - \omega(t)
\end{aligned}\tag{6}$$

Using the technique proposed in [3], the following result can be proved

Proposition 1

Assume that the sequences $\{e(t)\}$ and $\{w(t)\}$ are uniformly small so that one can neglect higher than first-order moments of their elements. Then the algorithm (4) applied to signal

$$y(t) = \beta(t) + v(t), \quad \beta(t) = e^{j\omega(t)}\beta(t-1)\tag{7}$$

can be approximately described by the following linear filtering equations

$$\begin{aligned}
\Delta\widehat{\phi}_R(t) &= F(q^{-1})e_R(t)\\
\Delta\widehat{\phi}_I(t) &= G_1(q^{-1})e_I(t) + G_2(q^{-1})w(t)
\end{aligned}\tag{8}$$

where q^{-1} denotes the backward shift operator ($q^{-1}x(t) = x(t-1)$) and

$$\begin{aligned}
F(q^{-1}) &= \frac{1-\lambda}{1-\lambda q^{-1}}\\
G_1(q^{-1}) &= \frac{1-\lambda+(\lambda-\delta)q^{-1}}{1-(\lambda+\delta)q^{-1}+\lambda q^{-2}}\\
G_2(q^{-1}) &= -\frac{b^2\lambda}{1-(\lambda+\delta)q^{-1}+\lambda q^{-2}}
\end{aligned}\tag{9}$$

with $b = |\beta(t)|$ and $\delta = 1 - \eta b^2$.

Outline of proof:

Using the approximation $e^{j\Delta\widehat{\omega}(t)} \cong 1 + j\Delta\widehat{\omega}(t)$, where $\Delta\widehat{\omega}(t) = \widehat{\omega}(t) - \omega(t)$, and neglecting all terms of order higher than one in $\Delta\widehat{\omega}(t)$ and $\Delta\widehat{\beta}(t-1)$, the error equations for (4) can be written down in the form

$$\begin{aligned}
\Delta\widehat{\phi}_R(t) &\cong \lambda\Delta\widehat{\phi}_R(t-1) + \mu e_R(t)\\
\Delta\widehat{\phi}_I(t) &\cong \lambda\Delta\widehat{\phi}_I(t-1) + \lambda b^2\Delta\widehat{\omega}(t) + \mu e_I(t)\\
\Delta\widehat{\omega}(t+1) &\cong \delta\Delta\widehat{\omega}(t) - \eta\Delta\widehat{\phi}_I(t-1)\\
&+ \eta e_I(t) - w(t+1)
\end{aligned}\tag{10}$$

All approximations hold for sufficiently high signal-to-noise ratio (SNR) and for sufficiently low rate of frequency changes compared with 1/SNR.
Solving equations (10) for $\Delta\widehat{\phi}_R(t)$ and $\Delta\widehat{\phi}_I(t)$ one arrives at (9). The complete proof can be found in [6]. ■

It is easy to check that for any λ and δ from the interval (0,1) the poles of all transfer functions in (9) lie inside the unit circle in the complex plane. Hence, under the constraint mentioned above, the approximating linear filter associated with (4) is stable.
Following many earlier tracking studies, we will assume that the frequency $\omega(t)$ evolves according to the random walk model, i.e. that the frequency increments $w(t)$ form a zero-mean white noise sequence with variance σ_w^2, independent of $v(t)$. It is straightforward to check that $e(t)$, similarly as $v(t)$, is a complex-valued white noise obeying $\sigma_e^2 = \mathrm{E}[|e(t)|^2] = b^2\sigma_v^2$, $\mathrm{E}[e_R^2(t)] = \mathrm{E}[e_I^2(t)] = \sigma_e^2/2$, $\mathrm{E}[e_R(t)e_I(s)] = 0$, $\forall t, s$. Hence, using standard results from the linear filtering theory, one arrives at

$$\mathrm{E}[(\Delta\widehat{\phi}_R(t))^2] = I[F(z)]\ \mathrm{E}[e_R^2(t)] \tag{11}$$

$$\mathrm{E}[(\Delta\widehat{\phi}_I(t))^2] = I[G_1(z)]\ \mathrm{E}[e_I^2(t)] + I[G_2(z)]\ \mathrm{E}[w^2(t)]$$

where

$$I[X(z)] = \frac{1}{2\pi j}\oint X(z)X(z^{-1})\frac{dz}{z}$$

is an integral evaluated along the unit circle in the z-plane, and $X(z)$ denotes any stable proper rational transfer function.
By means of residue calculus one obtains

$$I[F(z)] = \frac{1-\lambda}{1+\lambda} \cong \frac{\mu}{2} \tag{12}$$

$$I[G_1(z)] = \frac{1+\delta-\lambda-3\lambda\delta+2\lambda^2}{(1-\lambda)(1+2\lambda+\delta)} \cong \frac{\gamma}{2\mu} + \frac{\mu}{2}$$

$$I[G_2(z)] = \frac{b^4\lambda^2(1+\lambda)}{(1-\lambda)(1-\delta)(1+2\lambda+\delta)} \cong \frac{b^4}{2\mu\gamma}$$

where $\gamma = 1-\delta = b^2\eta$ and all approximations hold for sufficiently small values of μ and γ.
For a constant-modulus signal it holds that $|\Delta\widehat{\phi}(t)|^2 = (\Delta\widehat{\phi}_R(t))^2 + (\Delta\widehat{\phi}_I(t))^2 = b^2|\Delta\widehat{\beta}(t)|^2$. Therefore, after combining (11) with (6), (7) and (12), one arrives at the following expression for the steady state mean-squared signal estimation error

$$\mathrm{E}[|\widehat{\beta}(t)-\beta(t)|^2] \cong \left[\frac{\gamma}{4\mu}+\frac{\mu}{2}\right]\sigma_v^2 + \frac{b^2}{2\mu\gamma}\ \sigma_w^2 \tag{13}$$

Observe that the derived formula includes terms proportional to the adaptation gains $\mu = 1-\lambda$ and $\gamma = 1-\delta$, and terms inversely proportional to μ and γ. This stays in agreement with the well-known fact in adaptive filtering: the adaptation gains should be chosen so as to compromise between the tracking speed of an adaptive filter (which increases with growing μ and γ) and its noise rejection capability (which decreases with growing μ and γ) [5].
Denote by μ_β and γ_β the values of μ and γ that minimize the mean-squared signal estimation error. Straightforward calculations yield

$$\mu_\beta = \sqrt[4]{2\xi}, \quad \gamma_\beta = \sqrt{2\xi}$$

$$\mathrm{E}[|\widehat{\beta}(t)-\beta(t)|^2|\mu_\beta,\gamma_\beta] \cong \sqrt[4]{2\xi}\ \sigma_v^2 \tag{14}$$

where

$$\xi = \frac{b^2\sigma_w^2}{\sigma_v^2} \tag{15}$$

Note that the optimal values of design parameters and the best achievable performance are functions of a scalar coefficient ξ - the product of the signal-to-noise ratio b^2/σ_v^2 and the variance of frequency changes σ_w^2. The coefficient ξ can be regarded a measure of signal nonstationarity and plays an important role in analysis of tracking capabilities of the algorithm (4).
Our first remark will concern the problem of choice of design variables μ and γ (or equivalently λ and δ). First of all, recall that γ, equal to $b^2\eta$, is a function of a signal power $b^2 = |\beta(t)|^2$. Therefore, unless $|\beta(t)|$ is constant (which we have been assuming so far) and known *a priori*, the user does not have full control over the adaptation gain γ. This obvious drawback can be eliminated by replacing the correction term $g(t)$ in (4) with the normalized correction term

$$\bar{g}(t) = \frac{g(t)}{\widehat{b}^2(t)} = \mathrm{Im}\left[\frac{\varepsilon^*(t)e^{j\widehat{\omega}(t)}\widehat{\beta}(t-1)}{\widehat{b}^2(t)}\right] \tag{16}$$

where $\widehat{b}^2(t)$ denotes a local estimate of $b^2 = |\beta(t)|^2$, for example $\widehat{b}^2(t) = \lambda_o\widehat{b}^2(t-1) + (1-\lambda_o)|\widehat{\beta}(t)|^2$, where $0 \le \lambda_o < 1$ is the local averaging coefficient (e.g. $\lambda_o = 0.9$). Careful analysis shows that such modification does not change equations of the approximating linear filter associated with (4), provided that δ is redefined as $\delta = 1-\eta$. In this case γ is equal to η, i.e. it is an entirely user-dependent quantity. The modifications described above can be easily extended to the multiple frequency case.
Even though our optimization study was not based on realistic assumptions (the random walk model of signal frequency variation can be criticized as rather naive), its results, summarized in (14), have some practical relevance as they suggest useful tuning rules. Observe that $\gamma_\beta = \mu_\beta^2$. Therefore, to make tuning easier it may be worthwhile to set $\gamma = \mu^2$. The problem is then reduced to selection of a single design parameter μ. Optimization of μ can be performed either sequentially (e.g. by setting $\mu(t) = \mu[\widehat{\xi}(t)]$, where $\widehat{\xi}(t)$ is a continuously updated local estimate of the rate of signal nonstationarity), or using parallel estimation approach (which combines in a rational way the results yielded by a bank of adaptive filters with different settings - see e.g. [5]). Because of the lack of space such fully adaptive procedures will not be discussed here.

3 Comparison with MFT

While the frequency tracking properties of the multiple frequency tracker (5) are known (see [3]), the results of its signal tracking analysis are new and seem to be presented for the first time.

Proposition 2

Assume that all conditions of ALF analysis are fulfilled. Then the MFT algorithm applied to the signal (7) can be approximately described by equations (8) with

$$F(q^{-1}) = \frac{1-\lambda}{1-\lambda q^{-1}} \tag{17}$$

$$G_1(q^{-1}) = \frac{(1-\lambda)(1-\rho q^{-1})}{1-(2\lambda+\rho-\rho\lambda)q^{-1}+\lambda q^{-2}}$$

$$G_2(q^{-1}) = -\frac{b^2\lambda}{1-(2\lambda+\rho-\rho\lambda)q^{-1}+\lambda q^{-2}}$$

where $\rho = 1-\eta$.

Outline of proof:

The first two error equations of (10) remain valid for the multiple frequency tracker. The third equation has the form

$$\Delta\widehat{\omega}(t+1) = \rho\Delta\widehat{\omega}(t) + \frac{\eta(1-q^{-1})}{b^2}\Delta\widehat{\phi}_I(t) - w(t+1)$$

Transfer functions, given by (17), can be easily obtained by solving ALF equations with respect to $\Delta\widehat{\phi}_R(t)$ and $\Delta\widehat{\phi}_I(t)$. The complete proof can be found in [6]. ■

One can check that the substitution $\delta = \lambda + \rho - \rho\lambda$ (or equivalently $\gamma = \eta\mu$) converts transfer functions (9), derived in the previous subsection for the algorithm (4), into the transfer functions (17), characterizing local behavior of the MFT algorithm. It is therefore clear that for a single noisy cisoid both algorithms have essentially the same signal tracking properties. Interestingly, a similar conclusion was reached in [4], where the *frequency tracking* properties of MFT were compared with the analogous properties of yet another three adaptive notch filtering algorithms.

4 Computer simulations

Figure 1 shows comparison of the theoretical variance of the signal estimation error, given by (14), and the mean square errors obtained via numerical simulation. The generated signal consisted of a single cisoid ($k = 1$) with constant amplitude $b = 1$, embedded in white complex Gaussian noise with variance $\sigma_v^2 = 0.2$ (SNR=7 dB). The evolution of the instantaneous frequency $\omega(t)$ was modeled as a random walk process with the variance of frequency increments set to $\sigma_w^2 = 10^{-7}$ and with the starting value set to $\omega(0) = \pi/2$.

According to (14), to optimize signal tracking one should set μ to $\mu_\beta = \sqrt[4]{2\xi} \cong 0.032$ and set γ to $\gamma_\beta = \sqrt{2\xi} = 0.001$ (i.e. set η to $\eta_\beta = \gamma_\beta/b^2 = \gamma_\beta$). The analysis was carried around the optimal point (μ_β, γ_β). In the first experiment γ was set to its optimal value γ_β and μ was changed around μ_β. In the second experiment μ was set to μ_β and γ was changed around γ_β. Both plots shown in Figure 1 were obtained by double averaging. First, the mean-squared signal estimation errors were computed for different pairs (μ, γ) and for a given frequency trajectory from 10000 iterations of the algorithm (after the algorithm has reached its steady state). The obtained results were next averaged over 50 realizations of $\{w(t)\}$, i.e. over 50 different frequency trajectories. Note good agreement between theoretical curves and the results of computer simulations.

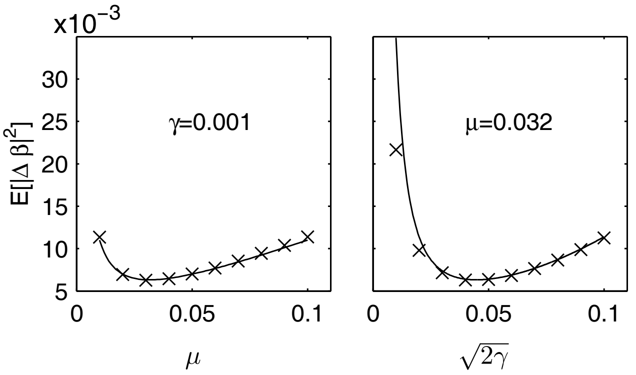

Figure 1: Variance of the signal estimation error for a single noisy cisoid with a random walk frequency drift. The theoretical results (solid lines) are compared with simulation results ($\times$) for different values of μ and γ.

References

[1] Niedźwiecki, M. and P. Kaczmarek (2004). Adaptive notch filters based on combined parametric and nonparametric approach. *Proc. IFAC Workshop on Adaptation and Learning in Control and Signal Processing*, Yokohama, Japan, 439–444.

[2] Niedźwiecki, M. and P. Kaczmarek (2004). Identification of quasi-periodically varying systems using the combined nonparametric/parametric approach. *IEEE Trans. on Signal Processing* (accepted)

[3] Tichavský P. and P. Händel (1995). Two algorithms for adaptive retrieval of slowly time-varying multiple cisoids in noise. *IEEE Trans. on Signal Processing*, **vol. 43**, 1116–1127.

[4] Tichavský P. and A. Nehorai (1997). Comparative study of four adaptive frequency trackers. *IEEE Trans. on Signal Processing*, **vol. 45**, 1473–1484.

[5] Niedźwiecki, M. (2000). *Identification of Time-varying Processes.* Wiley. New York.

[6] Niedźwiecki, M. and P. Kaczmarek (2005). Tracking analysis of a generalized adaptive notch filter. *IEEE Trans. on Signal Processing* (accepted)