FROM THE MULTIPLE FREQUENCY TRACKER TO THE MULTIPLE FREQUENCY SMOOTHER

Maciej Niedźwiecki

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

ABSTRACT

The problem of extraction/elimination of nonstationary sinusoidal signals from noisy measurements is considered. This problem is usually solved using adaptive notch filtering (ANF) algorithms. It is shown that the accuracy of frequency estimates can be significantly increased if the results obtained from ANF are backward-time filtered by an appropriately designed lowpass filter. The resulting adaptive notch smoothing (ANS) algorithm can be employed to perform many off-line signal processing tasks, such as elimination of sinusoidal interference from a prerecorded signal. In the single sinusoid case, we show that when the unknown signal frequency drifts according to the random-walk model, the optimally tuned ANS algorithm is, under Gaussian assumptions, statistically efficient, i.e. , it attains the Cramér-Rao type lower smoothing bound, which limits accuracy of *any* frequency estimation scheme.

Index Terms— frequency estimation, adaptive filters

1. PROBLEM STATEMENT

Consider the problem of extraction or elimination of nonstationary complex sinusoidal signals (cisoids) from noisy measurements $y(t)$

$$y(t) = \sum_{l=1}^{k} s_i(t) + v(t) = \mathbf{1}_k^{\mathrm{T}}\mathbf{s}(t) + v(t)$$

where $\mathbf{s}(t) = [s_1(t), \dots, s_k(t)]^{\mathrm{T}}$, $\mathbf{1}_k = [1, \dots, 1]^{\mathrm{T}}$ and

$$s_i(t) = a_i(t) e^{\left[j \sum_{l=1}^{t} \omega_i(l)\right]}, \quad i = 1, \dots, k\ . \tag{1}$$

We will assume that the complex-valued amplitudes $a_i(t)$ and real-valued instantaneous frequencies $\omega_i(t) \in [-\pi, \pi]$ are slowly varying quantities, and that the measurement noise $v(t)$ is circular white.

Due to its large practical relevance, the problem of retrieving noisy sinusoidal signals has attracted a great deal of interest in the literature — see, e.g., [1] and references therein. Recursive estimation of signal parameters — its amplitudes and frequencies — can be achieved in many different ways using the so-called adaptive notch filtering (ANF) algorithms. In this paper, we will focus our attention on a particular form of ANF known as the multiple frequency tracker (MFT), the algorithm proposed and analyzed in [2]. Let $\widehat{\mathbf{s}}(t) = [\widehat{s}_1(t), \dots, \widehat{s}_k(t)]^{\mathrm{T}}$ and $\widehat{\mathbf{\Omega}}(t) = \mathrm{diag}\{e^{j\widehat{\omega}_1(t)}, \dots, e^{j\widehat{\omega}_k(t)}\}$, where $\widehat{s}_i(t)$ and $\widehat{\omega}_i(t)$ denote the current signal and frequency estimates, respectively. The MFT algorithm can be summarized as follows

$$\begin{aligned}
\widehat{\mathbf{s}}(t) &= \widehat{\Omega}(t-1)\widehat{\mathbf{s}}(t-1) + \mathbf{P}^{-1}(t)\mathbf{1}_k\varepsilon(t) \\
\varepsilon(t) &= y(t) - \mathbf{1}_k^{\mathrm{T}}\widehat{\mathbf{\Omega}}(t-1)\widehat{\mathbf{s}}(t-1) \\
\mathbf{P}(t) &= \mathbf{1}_k\mathbf{1}_k^{\mathrm{T}} + \lambda\widehat{\mathbf{\Omega}}(t)\mathbf{P}(t-1)\widehat{\mathbf{\Omega}}^*(t-1) \\
\widehat{\omega}_i(t) &= \widehat{\omega}_i(t-1) + (1-\rho)\ \mathrm{Arg}\left[\frac{\widehat{s}_i(t)}{\widehat{s}_i(t-1)e^{j\widehat{\omega}_i(t-1)}}\right] \\
& i = 1, \dots, k
\end{aligned} \tag{2}$$

It can be controlled by means of adjusting two user-dependent coefficients: the forgetting constant λ, $0 < \lambda < 1$, which determines the speed of amplitude tracking, and another forgetting factor ρ, $0 < \rho < 1$, which determines the speed of frequency tracking.

Similarly, as in the majority of available ANF algorithms, MFT is a causal estimation scheme, which means that the instantaneous frequency estimates are obtained in terms of the current and past measurements only. While in all real-time applications causality is an obvious requirement, in many off-line processing tasks, such as elimination of a sinusoidal interference from a prerecorded signal, estimation can be based on both past and "future" measurements. When appropriately designed, such noncausal estimators, which incorporate smoothing, yield smaller estimation errors than their causal counterparts. We will show how the multiple frequency tracker can be turned into a statistically efficient multiple frequency smoother.

2. ADAPTIVE NOTCH SMOOTHING ALGORITHM — THE SINGLE FREQUENCY CASE

The proposed smoothing procedure is two-pass. During the first, forward-time pass, the MFT algorithm, described above,

This work was supported by MNiSW under Grant N514 011 31/3091.

is used to obtain preliminary (biased) frequency estimates. The second, backward-time pass, is needed to perform filtering of the sequence of frequency estimates obtained in the first stage of processing. The transfer function of the smoothing (lowpass) filter depends on the frequency tracking characteristics of MFT and will be determined analytically.

2.1. Filtering

For a single noisy cisoid ($k = 1$) the steady-state version of the MFT algorithm (2) can be written in the form

$$\begin{aligned}\widehat{s}(t) &= e^{j\widehat{\omega}(t-1)}\widehat{s}(t-1) + \mu\varepsilon(t)\\ \varepsilon(t) &= y(t) - e^{j\widehat{\omega}(t-1)}\widehat{s}(t-1)\\ \widehat{\omega}(t) &= \widehat{\omega}(t-1) + \gamma\ \mathrm{Arg}\left[\frac{\widehat{s}(t)}{\widehat{s}(t-1)e^{j\widehat{\omega}(t-1)}}\right]\end{aligned} \tag{3}$$

where $\mu = 1 - \lambda$ and $\gamma = 1 - \rho$ denote the corresponding (small) adaptation gains.

Using the approximating linear filter technique, Tichavský and Händel [2] established the following steady-state relationship between the frequency estimation error $\Delta\widehat{\omega}(t) = \widehat{\omega}(t) - \omega(t)$, the one-step frequency changes $w(t) = \omega(t) - \omega(t-1)$, and the scaled measurement noise $z(t) = \mathrm{Im}[v(t)/s(t)]$:

$$\Delta\widehat{\omega}(t) \cong \frac{F(q^{-1}) - 1}{1 - q^{-1}}\ w(t) + (1 - q^{-1})F(q^{-1})z(t) \tag{4}$$

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

where q^{-1} denotes the backward shift operator.

Following [2], to obtain more specific analytical results, we will assume that:

(A1) The measurement noise $\{v(t)\}$ is a zero-mean circular white sequence of complex random variables with variance σ_v^2.

(A2) $\{s(t)\}$ is a constant modulus signal, i.e. $|s(t)| = |a|$, $\forall t$.

(A3) The process $\{w(t)\}$, independent of $\{v(t)\}$, is a zero-mean white noise with variance σ_w^2 (random-walk frequency drift).

Note that under (A1) and (A2) $\{z(t)\}$ is a real-valued white noise with variance $\sigma_z^2 = \sigma_v^2/(2|a|^2)$. Under (A1)—(A3) the mean-squared frequency estimation error can be expressed in explicit form as

$$\begin{aligned}\mathrm{E}\{\,[\widehat{\omega}(t) - \omega(t)]^2\} &\cong \frac{\sigma_v^2}{|a|^2}\ \frac{(1-\lambda)(1-\rho)}{1 + 3\lambda + \rho(1-\lambda)}\\ &+ \sigma_w^2\ \frac{\rho^2(1-\lambda)^2 + 2\lambda(\rho + \lambda - 2\rho\lambda)}{[1 + 3\lambda + \rho(1-\lambda)](1-\lambda)(1-\rho)}\end{aligned} \tag{5}$$

The minimum value of the frequency tracking error is achieved for:

$$\mu = \mu_{\mathrm{opt}} = \frac{-u + \sqrt{u^2 + 4u}}{2}\ , \quad \gamma = \gamma_{\mathrm{opt}} = \frac{\mu_{\mathrm{opt}}}{2 - \mu_{\mathrm{opt}}}$$

where

$$u = \kappa + \sqrt{\kappa^2 + 8\kappa} \tag{6}$$

and $\kappa = |a|^2\sigma_w^2/\sigma_v^2$ is a scalar coefficient that can be regarded as a measure of signal nonstationarity. As shown in [2], under Gaussian assumptions, the minimum mean-squared tracking error obtained for the optimal settings

$$\begin{aligned}&\mathrm{E}\{\,[\widehat{\omega}(t) - \omega(t)]^2|\mu_{\mathrm{opt}}, \gamma_{\mathrm{opt}}\}\\ &\qquad\cong (-1 + \sqrt{1 + 4u^{-1}})\ \sigma_w^2\end{aligned} \tag{7}$$

attains its lower limit known as the posterior Cramér-Rao bound [3]. This means that in the case considered the optimally tuned MFT algorithm (3) is a *statistically efficient* procedure for tracking randomly drifting frequency.

2.2. Smoothing

To obtain a smoothed estimate of $\omega(t)$, further denoted by $\widetilde{\omega}(t)$, we will pass the estimates yielded by the "pilot" MFT algorithm (3) through an appropriately designed noncausal filter $G(q^{-1}) = \ldots + g_{-1}q^{-1} + g_0 + g_1q^1 + \ldots$

$$\widetilde{\omega}(t) = G(q^{-1})\widehat{\omega}(t)\ . \tag{8}$$

The filter $G(q^{-1})$ will be designed so as to minimize the mean-squared frequency matching error $\mathrm{E}[(\Delta\widetilde{\omega}(t))^2]$, where $\Delta\widetilde{\omega}(t) = \widetilde{\omega}(t) - \omega(t)$. Combining (4) and (8), one arrives at

$$\Delta\widetilde{\omega}(t) \cong \frac{X(q^{-1}) - 1}{1 - q^{-1}}\ w(t) + (1 - q^{-1})X(q^{-1})z(t) \tag{9}$$

where

$$X(q^{-1}) = F(q^{-1})G(q^{-1}) \tag{10}$$

Under (A1)—(A3), it holds that

$$\mathrm{E}\{\,[\Delta\widetilde{\omega}(t)]^2\} \cong \frac{1}{2\pi}\int_{-\pi}^{\pi} f\left[X(e^{-j\xi})\right]d\xi \tag{11}$$

$$f[X] = c_1(X-1)(X^*-1)H_1H_1^* + c_2XX^*H_2H_2^*$$

where $c_1 = \sigma_w^2$, $c_2 = \sigma_z^2$, $H_1(e^{-j\xi}) = 1/(1-\xi^{-1})$, $H_2(e^{-j\xi}) = 1 - \xi^{-1}$. To avoid confusion with ω, the standard Fourier-domain angular frequency variable is denoted here by ξ.

Minimization of (11) is pretty straightforward — the problem can be solved by minimizing $f\left[X(e^{-j\xi})\right]$ for every value of $\xi \in [-\pi, \pi]$. Setting $\partial f/\partial X^*|_{X=X_{\mathrm{opt}}} = 0$, one obtains $X_{\mathrm{opt}} = c_1H_1H_1^*/(c_1H_1H_1^* + c_2H_2H_2^*)$ or equivalently

$$X_{\mathrm{opt}}(q^{-1}) = \frac{c}{c + (1-q^{-1})^2(1-q)^2}\ . \tag{12}$$

where $c = c_1/c_2 = 2\sigma_w^2|a|^2/\sigma_v^2 = 2\kappa$.

It turns out that $X_{\rm opt}(q^{-1})$, given by (12), can be expressed in the form $F(q^{-1})F(q)$. To determine coefficients $\lambda_o = 1-\mu_o$ and $\rho_o = 1-\gamma_o$, which allow for such factorization, one must solve the following two algebraic equations

$$\mu_o^2\gamma_o^2 = 2(1-\mu_o)\kappa \ , \quad \mu_o = (2-\mu_o)\gamma_o \qquad (13)$$

(derivation is elementary but tedious). After extracting γ_o from the second equation and substituting it in the first equation, one obtains

$$\mu_o^4 - 2(1-\mu_o)(2-\mu_o)^2\kappa = 0 \ . \qquad (14)$$

Note that the substitution

$$u = \frac{\mu_o^2}{1-\mu_o} \qquad (15)$$

turns the fourth-order equation (14) into the second-order equation $u^2 - 2\kappa u - 8\kappa = 0$, which, given that $u > 0$, is equivalent to (6). Furthermore, solving (15) for $\mu_o > 0$ one obtains expression which is identical with $\mu_{\rm opt}$. The same holds for γ_o, i.e. $\gamma_o = \gamma_{\rm opt}$. In this way one arrives at

$$X_{\rm opt}(q^{-1}) = F_{\rm opt}(q^{-1})F_{\rm opt}(q) \qquad (16)$$

where $F_{\rm opt}(q^{-1}) = F(q^{-1})|\lambda = \lambda_{\rm opt}, \rho = \rho_{\rm opt}$.

Finally, given that the forward-time MFT algorithm is optimally tuned, one obtains [cf. (10)]

$$G_{\rm opt}(q^{-1}) = \frac{X_{\rm opt}(q^{-1})}{F_{\rm opt}(q^{-1})} = F_{\rm opt}(q) \qquad (17)$$

Since the filter $G_{\rm opt}(q^{-1})$ is *anticausal*, the smoothed estimate $\widetilde{\omega}(t)$ can be obtained by means of backward-time filtering of the estimates yielded by the MFT algorithm (3). Such backward-time filtering can be performed recursively using the following equations

$$\begin{aligned}\widetilde{\omega}(t) &= (2\lambda + \rho - \rho\lambda)\widetilde{\omega}(t+1) - \lambda\widetilde{\omega}(t+2) \\ &\quad + (1-\rho)(1-\lambda)\widehat{\omega}(t) \\ t &= N-2, \ldots, 1\end{aligned} \qquad (18)$$

where N denotes the number of available (prerecorded) data samples and $\widetilde{\omega}(N) = \widehat{\omega}(N), \widetilde{\omega}(N-1) = \widehat{\omega}(N-1)$. Note that the optimal gains $\lambda_{\rm opt}$ and $\rho_{\rm opt}$, usually not known *a priori*, were replaced with λ and ρ, respectively — the gains used in the tracking algorithm. Making such a choice is equivalent to adopting $G(q^{-1}) = F(q)$.

Once the smoothed frequency trajectory is available, one can use it to improve estimation results by running — as a follow up to (3) — another algorithm that incorporates the smoothed frequency estimates

$$\begin{aligned}\widetilde{\varepsilon}(t) &= y(t) - e^{j\widetilde{\omega}(t-1)}\widetilde{s}(t-1) \\ \widetilde{s}(t) &= e^{j\widetilde{\omega}(t-1)}\widetilde{s}(t-1) + \mu\widetilde{\varepsilon}(t) \ .\end{aligned} \qquad (19)$$

Note that the "frequency guided" filter (19), originally proposed in [4] for the purpose of fixed-lag smoothing, does not estimate the instantaneous frequency on its own — it relies on frequency estimates obtained from the smoothing filter (18).

3. STATISTICAL EFFICIENCY

When the estimated frequency is time-invariant and the measurement noise is circular Gaussian, the lower Cramér-Rao frequency estimation bound (CRB) is proportional to $1/N^3$ [5]. When the instantaneous frequency is a stochastic variable, the classical Cramér-Rao inequality does not apply. A bound that is similar to the CRB, and can be applied to signals with randomly drifting frequency, was derived by Tichavský [3], based on a more general result due to van Trees [6]. This bound was called in [3] the posterior Cramér-Rao bound (PCRB). Let $\widehat{\boldsymbol{\omega}} = [\widehat{\omega}(1), \ldots, \widehat{\omega}(N)]^{\rm T}$ be any estimator (possibly biased) of the vector of instantaneous frequencies $\boldsymbol{\omega} = [\omega(1), \ldots, \omega(N)]^{\rm T}$, based on the entire observation history $\mathcal{Y}(N) = \{y(1), \ldots, y(N)\}$. Denote by $\mathbf{F}_N$ the the so-called posterior Fisher information matrix for the problem at hand

$$\mathbf{F}_N = \frac{1}{\sigma_w^2}\left[\mathbf{G}_N + 2\kappa\mathbf{H}_N\right]$$

$$\mathbf{G}_N = \begin{bmatrix} 1 & -1 & 0 & \ldots & 0 & 0 \\ -1 & 2 & -1 & \ldots & 0 & 0 \\ 0 & -1 & 2 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 2 & -1 \\ 0 & 0 & 0 & \ldots & -1 & 1 \end{bmatrix}$$

$$\mathbf{H}_N = \begin{bmatrix} N & N-1 & N-2 & \ldots & 2 & 1 \\ N-1 & N-1 & N-2 & \ldots & 2 & 1 \\ N-2 & N-2 & N-2 & \ldots & 2 & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 2 & 2 & 2 & \ldots & 2 & 1 \\ 1 & 1 & 1 & \ldots & 1 & 1 \end{bmatrix}$$

Finally, denote by $[\mathbf{F}^{-1}]_{ij}$ the (i,j)th element of $\mathbf{F}^{-1}$.

According to Tichavský [3], under (A1)—(A3) and under Gaussian assumptions imposed on $\{v(t)\}$ and $\{w(t)\}$, the asymptotic (steady-state) value of the PCRB for any causal estimator of $\omega(t)$ can be obtained from

$$\begin{aligned}\text{LTB} &= \lim_{t\mapsto\infty} \inf_{\widehat{\omega}(\cdot)} \mathrm{E}\{\,[\widehat{\omega}(t) - \omega(t)]^2\} \\ &= \lim_{n\mapsto\infty} \left[\mathbf{F}_{2n+1}^{-1}\right]_{2n+1,2n+1} \ .\end{aligned} \qquad (20)$$

This bound, further referred to as the lower tracking bound (LTB), can be computed analytically and is identical with the right-hand side of (7).

The analogous quantity, limiting the steady-state performance of *any* frequency estimation scheme (including all noncausal estimators), will be called the lower smoothing bound (LSB). It is given by

$$\begin{aligned}\text{LSB} &= \lim_{t\mapsto\infty} \inf_{\widetilde{\omega}(\cdot)} \mathrm{E}\{\,[\widetilde{\omega}(t) - \omega(t)]^2\} \\ &= \lim_{n\mapsto\infty} \left[\mathbf{F}_{2n+1}^{-1}\right]_{n+1,n+1} \ .\end{aligned} \qquad (21)$$

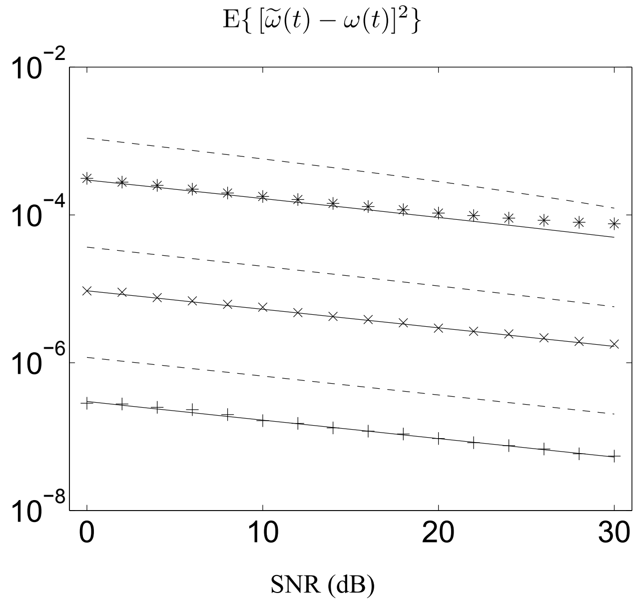

Fig. 1. Comparison of the theoretical values of the lower smoothing bound (solid lines) and lower tracking bound (broken lines) with experimental results obtained for the signal with randomly drifting frequency for three different speeds of frequency variation: $\sigma_w = 0.01$ ($*$), $\sigma_w = 0.001$ ($\times$), $\sigma_w = 0.0001$ ($+$), and for 16 different SNR values.

The lower smoothing bound can be evaluated numerically as $\left[\mathbf{F}_{2n+1}^{-1}\right]_{n+1,n+1}$ provided that n is sufficiently large.

Since, in the case of smoothing, the amount of data used for estimation is effectively doubled compared to tracking, it holds that LSB $<$ LTB, i.e., the limiting accuracy of non-causal estimators exceeds accuracy of their causal counterparts. In the randomly drifting frequency scenario, the achievable reduction rate approaches 4 for small values of κ.

Numerical tests clearly demonstrate that the proposed ANS algorithm is statistically efficient. Figure 1 shows a comparison of the theoretical values of the lower smoothing bound with experimental results obtained for the constant modulus signal, subject to random-walk frequency drift, for three different speeds of frequency variation σ_w (0.01, 0.001 and 0.0001) and 16 different SNR values, ranging from 0 dB to 30 dB. The mean-squared frequency estimation errors were evaluated (for the optimally tuned ANS algorithm) by means of joint time and ensemble averaging. First, for each realization of the measurement noise sequence and each realization of the frequency trajectory, the mean-squared errors were computed from 200 iterations of the ANS filter (after the algorithm has reached its steady state). The obtained results were next averaged over 200 realizations of $\{v(t)\}$ and 200 realizations of $\{w(t)\}$. Note the good agreement between the theoretical curves and the results of computer simulations.

4. MULTIPLE FREQUENCY SMOOTHER

Extension of the smoothing scheme, described in Section 2 to the multiple frequencies case is pretty straightforward. The two additional steps that should be taken after computing the frequency estimates $\widehat{\omega}_i(t), i = 1, \dots, k$ using the "pilot" MFT algorithm (2), can be summarized as follows:

smoothing filter:

$$\begin{aligned} \widetilde{\omega}_i(t) &= F(q)\widehat{\omega}_i(t) \\ i &= 1, \dots, k \end{aligned} \tag{22}$$

frequency-guided filter:

$$\begin{aligned} \widetilde{\mathbf{s}}(t) &= \widetilde{\mathbf{\Omega}}(t-1)\widetilde{\mathbf{s}}(t-1) + \mathbf{Q}^{-1}(t)\mathbf{1}_k\varepsilon(t) \\ \varepsilon(t) &= y(t) - \mathbf{1}_k^{\mathrm{T}}\widetilde{\mathbf{\Omega}}(t-1)\widetilde{\mathbf{s}}(t-1) \\ \mathbf{Q}(t) &= \mathbf{1}_k\mathbf{1}_k^{\mathrm{T}} + \lambda\widetilde{\mathbf{\Omega}}(t)\mathbf{Q}(t-1)\widetilde{\mathbf{\Omega}}^*(t-1) \end{aligned} \tag{23}$$

where $\widetilde{\mathbf{\Omega}}(t) = \mathrm{diag}\{e^{j\widetilde{\omega}_1(t)}, \dots, e^{j\widetilde{\omega}_k(t)}\}$.

The efficient initialization procedure, which can be used to identify the number of frequency modes k and to determine the initial frequency estimates, needed to start (2), was presented in [7].

5. REFERENCES

[1] P. A. Regalia, *Adaptive IIR Filtering in Signal Processing and Control*, New York: Marcel Dekker, 1995.

[2] P. Tichavský and P. Händel, "Two algorithms for adaptive retrieval of slowly time-varying multiple cisoids in noise", *IEEE Trans. Signal Process.*, vol. 43, pp. 1116–1127, 1995.

[3] P. Tichavský, "Posterior Cramér-Rao bound for adaptive harmonic retrieval", *IEEE Trans. Signal Process.*, vol. 43, pp. 1299–1302, 1995.

[4] M. Niedźwiecki and A. Sobociński, "A simple way of increasing estimation accuracy of generalized adaptive notch filters", *IEEE Signal Process. Lett.*, vol. 14, pp. 217–220, 2007.

[5] D.C. Rife and R.R. Boorstyn, "Single-tone parameter estimation from discrete-time observations", *IEEE Trans. Inf. Theory*, vol. 20, pp. 591–598, 1974.

[6] H.L. van Trees, *Detection, Estimation and Modulation Theory*, New York: Wiley, 1968.

[7] M. Niedźwiecki and P. Kaczmarek, "Identification of quasi-periodically varying systems using the combined nonparametric/parametric approach", *IEEE Trans. Signal Process.*, vol. 53, pp. 4588–4598, 2005.