GENERALIZED ADAPTIVE COMB FILTER WITH IMPROVED ACCURACY AND ROBUSTNESS PROPERTIES

Maciej Niedźwiecki and Michał Meller

Faculty of Electronics, Telecommunications and Computer Science
Department of Automatic Control, Gdańsk University of Technology
Narutowicza 11/12, Gdańsk , Poland
maciekn@eti.pg.gda.pl, michal.meller@eti.pg.gda.pl

ABSTRACT

Generalized adaptive comb filters can be used to identify/track parameters of quasi-periodically varying systems. In a special, signal case they reduce down to adaptive comb filters, applied to elimination or extraction of nonstationary multi-harmonic signals buried in noise. We propose a new algorithm which combines, in an adaptive way, results yielded by several, simultaneously working generalized adaptive comb filters. Due to its highly parallel estimation structure, the new algorithm is more accurate and more robust than the currently available algorithms.

1. INTRODUCTION

We will consider the problem of identification of quasi-periodically varying complex-valued systems, i.e., systems governed by

$$y(t) = \sum_{i=1}^{n} \theta_i(t)u(t-i+1) = \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta}(t) + v(t) \qquad (1)$$

where $t = \ldots,-1,0,1\ldots$ denotes the normalized discrete time, $y(t)$ denotes the system output, $\boldsymbol{\varphi}(t) = [u(t),\ldots,u(t-n+1)]^{\mathrm{T}}$ denotes regression vector, made up of the past input samples, $v(t)$ denotes measurement noise, and $\boldsymbol{\theta}(t) = [\theta_1(t),\ldots,\theta_n(t)]^{\mathrm{T}}$ is the vector of time-varying system coefficients, modeled as weighted sums of complex exponentials

$$\boldsymbol{\theta}(t) = \sum_{k=1}^{K} \boldsymbol{\alpha}_k(t)e^{j\sum_{i=1}^{t}\omega_k(i)}. \qquad (2)$$

All quantities in (1)–(2), except angular frequencies $\omega_1(t),\ldots,\omega_K(t)$, are complex-valued. Since the complex 'amplitudes' $\boldsymbol{\alpha}_k(t) = [a_{k1}(t),\ldots,a_{kn}(t)]^{\mathrm{T}}$, assumed to vary slowly with time, incorporate both magnitude and phase information, there is no explicit phase component in (2). Under certain circumstances (in the presence of several strong reflectors) the model (1)–(2) can be used to describe rapidly fading mobile radio channels [1], [2]. In this case $y(t)$ denotes the sampled baseband signal received by the mobile unit, $\{u(t)\}$ denotes the sequence of transmitted symbols, and $v(t)$ denotes channel noise.
We will assume that the frequencies $\omega_k(t)$ are harmonically related, namely

$$\omega_k(t) = l_k\omega_0(t), \;\; k = 1,\ldots,K \qquad (3)$$

This work was supported by the Foundation for the Polish Science and NCN.

where $\omega_0(t)$ denotes the slowly time-varying fundamental frequency and l_k are integer numbers. Such multiple frequencies, called harmonics, appear in the Fourier series expansions of periodic signals. For example, if parameter trajectory $\boldsymbol{\theta}(t)$ is periodic with period N, it admits the following Fourier representation

$$\boldsymbol{\theta}(t) = \sum_{k=0}^{N-1} \boldsymbol{\alpha}_k e^{jk\omega_0 t}, \quad \omega_0 = \frac{2\pi}{N}\,.$$

The notion of 'time-varying harmonics' can be regarded a natural extension of the Fourier analysis to quasi-periodically varying systems, such as (1)–(2). The choice of the multipliers $l_k, k = 1,\ldots,K$, depends on our prior knowledge of the system time variation. When all harmonics are expected to be present, one should set $l_k = k$. In the presence of odd harmonics only, the natural choice is $l_k = 2k-1$, etc.
In the special case where $n = 1$ and $\varphi(t) \equiv 1$, equations (1)–(3) describe a complex-valued multi-harmonic signal $s(t) = \theta(t)$ buried in noise

$$y(t) = s(t) + v(t), \quad s(t) = \sum_{k=1}^{K} a_k(t)e^{j\sum_{i=1}^{t}\omega_k(i)}. \qquad (4)$$

The problem of either elimination or extraction of multi-harmonic signals buried in noise can be solved using adaptive comb filters. For this reason the system identification/tracking algorithm described below can be considered a generalized comb filter.

2. GENERALIZED COMB FILTER

Suppose that

(A1) The sequence of regression vectors $\{\boldsymbol{\varphi}(t)\}$ is a wide sense stationary and ergodic process with a known positive definite correlation matrix $\mathrm{E}[\boldsymbol{\varphi}^*(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)] = \boldsymbol{\Phi} > 0$.

(A2) $\{v(t)\}$, independent of $\{\boldsymbol{\varphi}(t)\}$, is a sequence of zero-mean circular, independent and identically distributed random variables.

2.1 Unconstrained algorithm

Under assumptions (A1)–(A2), identification of the system (1)–(2) can be carried out using the multiple frequency version of the generalized adaptive notch filter (GANF) proposed in [3] (the signal-oriented version of this algorithm was analyzed in [4]). In addition to the instantaneous frequencies

$\omega_k(t)$, this algorithm tracks the frequency rates defined as $\delta_k(t) = \omega_k(t) - \omega_k(t-1)$:

$$
\begin{aligned}
\widehat{f}_k(t) &= e^{j[\widehat{\omega}_k(t-1)+\widehat{\delta}_k(t-1)]}\widehat{f}_k(t-1) \\
\widehat{\theta}(t|t-1) &= \sum_{k=1}^{K} \widehat{f}_k(t)\widehat{\alpha}_k(t-1) \\
\varepsilon(t) &= y(t) - \varphi^{\mathrm{T}}(t)\widehat{\theta}(t|t-1) \\
\widehat{\alpha}_k(t) &= \widehat{\alpha}_k(t-1) + \mu_k \Phi^{-1}\varphi^{\star}(t)\widehat{f}_k^{\star}(t)\varepsilon(t) \\
g_k(t) &= \mathrm{Im}\left[\frac{\varepsilon^{\star}(t)\varphi^{\mathrm{T}}(t)\widehat{f}_k(t)\widehat{\alpha}_k(t-1)}{\widehat{\alpha}_k^{\mathrm{H}}(t-1)\Phi\widehat{\alpha}_k(t-1)}\right] \\
\widehat{\delta}_k(t) &= \widehat{\delta}_k(t-1) - \eta_k g_k(t) \\
\widehat{\omega}_k(t) &= \widehat{\omega}_k(t-1) + \widehat{\delta}_k(t-1) - \gamma_k g_k(t) \\
\widehat{\theta}(t|t) &= \sum_{k=1}^{K} \widehat{\alpha}_k(t)\widehat{f}_k(t) \\
& k = 1,\ldots,K
\end{aligned}
\tag{5}
$$

where $\star$ denotes complex conjugation, H denotes Hermitian transpose (conjugate transpose) and $\mu_k > 0, \gamma_k > 0, \eta_k > 0$, $\eta_k \ll \gamma_k \ll \mu_k$, $k = 1,\ldots,K$, denote the user-dependent adaptation gains controlling the speed of frequency rate adaptation, frequency adaptation and amplitude adaptation, respectively. To reduce the number of design degrees of freedom, one can choose $\eta_1 = \ldots = \eta_K = \eta$, $\gamma_1 = \ldots = \gamma_K = \gamma$ and $\mu_1 = \ldots = \mu_K = \mu$.

The algorithm (5) is made up of K single-frequency GANF sub-algorithms [each taking care of one frequency component of $\theta(t)$], that work in parallel and are driven by the common prediction error $\varepsilon(t)$. In [5] it was shown that the number of frequency modes, as well as all initial conditions needed to smoothly start (start without initialization transients) the GANF algorithm, can be inferred from non-parametric DFT-based analysis of a short startup fragment of the input-output data. The tool that can be used for this purpose was termed generalized (system) periodogram, as in the signal case it reduces to the classical periodogram.

In spite of its simplicity, the gradient frequency tracking mechanism adopted in (5) has very good statistical properties – as shown in [3] for the single frequency case ($K = 1$), when the instantaneous frequency drifts according to the integrated random walk model (quasi-linear variation), the optimally tuned GANF algorithm (5) is statistically efficient, i.e., under Gaussian assumptions it reaches the Cramér-Rao-type lower frequency and frequency rate tracking bounds.

When applied to identification of the system (1)–(3) the GANF algorithm (5) has two serious drawbacks.

First, it does not take into consideration the harmonic structure (3), i.e., the estimated frequencies are regarded as mutually unrelated quantities, while the true harmonics vary in a coordinated way. Hence, even though such an unconstrained multiple-frequency generalized adaptive notch filter can be used to identify the multi-harmonic system/signal, its tracking characteristics will be generally inferior to those offered by solutions that incorporate the harmonic constraints (the analysis carried out for stationary multi-harmonic signals by Nehorai and Porat [6], shows that significant improvements in the Cramér-Rao bounds can be achieved if the harmonic structure is taken into account in the estimation process).

Second, the algorithm (5) is not robust to incorrect frequency matching. While the strong frequency components, i.e., those characterized by large values of the signal-to-noise ratio $\mathrm{SNR}(t) = \| \alpha_k(t) \|^2 / \sigma_v^2$ are usually tracked successfully, the weak ones may be difficult to follow – even if the initial frequency assignment is correct, the sub-algorithms tracking such weak components may, after some time, lock onto the neighboring, stronger components, corresponding to higher or lower frequencies. Moreover, when the system/signal is nonstationary, the 'strength' of different harmonic components may also vary with time, which further complicates the picture. Note that in the parallel estimation structure underlying (5), the neglected harmonic components become a part of the prediction error $\varepsilon(t)$. Hence, since all sub-algorithms are driven by the same error signal, a failure of even one of them adversely affects performance of the entire structure. In extreme cases such a failure may even cause the filter divergence.

2.2 Constrained algorithm

Note that, according to (3), the estimates of the k-th harmonic $\omega_k(t)$ and its rate of change $\delta_k(t)$ can be used to obtain the estimates of all other harmonics using the relationships

$$
\widehat{\omega}_{i|k}(t) = \frac{\widehat{\omega}_k(t)l_i}{l_k}\ ,\ \widehat{\delta}_{i|k}(t) = \frac{\widehat{\delta}_k(t)l_i}{l_k}\ ,\ i = 1,\ldots,K,\ i \neq k. \tag{6}
$$

This simple observation is the cornerstone of the constrained tracking algorithm summarized below:

$$
\begin{aligned}
\widehat{f}_{i|k}(t) &= e^{j[\widehat{\omega}_{i|k}(t-1)+\widehat{\delta}_{i|k}(t-1)]}\widehat{f}_{i|k}(t-1) \\
\widehat{\theta}_k(t|t-1) &= \sum_{i=1}^{K} \widehat{f}_{i|k}(t)\widehat{\alpha}_{i|k}(t-1) \\
\varepsilon_k(t) &= y(t) - \varphi^{\mathrm{T}}(t)\widehat{\theta}_k(t|t-1) \\
\widehat{\alpha}_{i|k}(t) &= \widehat{\alpha}_{i|k}(t-1) + \mu_k \Phi^{-1}\varphi^{\star}(t)\widehat{f}_{i|k}^{\star}(t)\varepsilon(t) \\
g_k(t) &= \mathrm{Im}\left[\frac{\varepsilon_k^{\star}(t)\varphi^{\mathrm{T}}(t)\widehat{f}_{k|k}(t)\widehat{\alpha}_{k|k}(t-1)}{\widehat{\alpha}_{k|k}^{\mathrm{H}}(t-1)\Phi\widehat{\alpha}_{k|k}(t-1)}\right] \\
\widehat{\delta}_{k|k}(t) &= \widehat{\delta}_{k|k}(t-1) - \eta_k g_k(t) \\
\widehat{\omega}_{k|k}(t) &= \widehat{\omega}_{k|k}(t-1) + \widehat{\delta}_{k|k}(t-1) - \gamma_k g_k(t) \\
\widehat{\delta}_{i|k}(t) &= \frac{\widehat{\delta}_{k|k}(t)l_i}{l_k}\ ,\ \ i \neq k \\
\widehat{\omega}_{i|k}(t) &= \frac{\widehat{\omega}_{k|k}(t)l_i}{l_k}\ ,\ \ i \neq k \\
\widehat{\theta}_k(t|t) &= \sum_{i=1}^{K} \widehat{\alpha}_{i|k}(t)\widehat{f}_{i|k}(t) \\
& k = 1,\ldots,K \\
& i = 1,\ldots,K
\end{aligned}
\tag{7}
$$

Similar to (5), the new algorithm is composed of K sub-algorithms, which track different harmonics. However, unlike (5), the k-th sub-algorithm estimates on its own *all* signal components[1], and works out its *own* prediction error $\varepsilon_k(t)$.

[1]To avoid confusion, the corresponding estimates are double indexed – $\widehat{\alpha}_{i|k}(t)$, $\widehat{\delta}_{i|k}(t)$ and $\widehat{\omega}_{i|k}(t)$ denote the estimates of $\alpha_i(t)$, $\delta_i(t)$ and $\omega_i(t)$, respectively, yielded by the k-th sub-algorithm.

We will show that such an estimation redundancy makes the constrained algorithm more robust than the unconstrained one. Note that only the k-th harmonic is directly estimated by the k-th sub-algorithm – the remaining frequency estimates are obtained indirectly via (6).
Since each of the sub-algorithms generates its own estimate of the parameter vector, some sort of information fusion is needed to arrive at the final estimate aggregating, in a statistically meaningful way, all partial estimates $\widehat{\boldsymbol{\theta}}_k(t|t)$, $k = 1,\ldots,K$. Our fusion formula will take a form of a convex combination

$$\widehat{\boldsymbol{\theta}}(t|t) = \sum_{k=1}^{K} \mu_k(t)\widehat{\boldsymbol{\theta}}_k(t|t) \tag{8}$$

where the weights $\mu_k(t) \geq 0, k = 1,\ldots,K$, $\sum_{k=1}^{K} \mu_k(t) = 1$, further referred to as credibility coefficients, are evaluated in terms of the locally observed prediction errors yielded by the corresponding sub-algorithms. The details of this construction will be given in the next section.

3. INFORMATION FUSION

In this section we will derive the information fusion formula. We will start from a brief description of the classical Bayesian estimation principle. Then we will present a non-standard Bayesian solution based on the so-called prequential approach.

3.1 Classical Bayesian approach

Consider K hypothetical, mutually exclusive statistical models $\mathscr{M}_k, k = 1,\ldots,K$, of the data-generating mechanism (system or signal) which might have produced the observed data $\Omega(t)$. Suppose that, based on each model, one works out an optimal, in the mean-squared sense, one-step-ahead prediction of the signal $y(t)$

$$\mathscr{M}_k \rightarrow \ \widehat{y}_k(t+1|t) = \boldsymbol{\varphi}^{\mathrm{T}}(t+1)\widehat{\boldsymbol{\theta}}_k(t+1|t) \tag{9}$$

where $\widehat{\boldsymbol{\theta}}_k(t+1|t) = \mathrm{E}[\boldsymbol{\theta}(t+1)|\Omega(t),\mathscr{M}_k]$ denotes the one-step-ahead predictor of $\boldsymbol{\theta}(t+1)$.
It is well-known [7] that the optimal Bayesian predictor, minimizing the quadratic loss function

$$\mathrm{E}\left\{|y(t+1) - \widehat{y}(t+1|t)|^2\right\}$$

takes the form (assuming that one of the models is a true system/signal description)

$$\widehat{y}(t+1|t) = \sum_{k=1}^{K} \mu_k(t)\widehat{y}_k(t+1|t) \tag{10}$$

where the weights $\mu_k(t)$ are given by

$$\mu_k(t) = P(\mathscr{M}_k|\Omega(t)) = \frac{p(\Omega(t)|\mathscr{M}_k)\pi(\mathscr{M}_k)}{\sum_{k=1}^{K} p(\Omega(t)|\mathscr{M}_k)\pi(\mathscr{M}_k)} \tag{11}$$

i.e., they are equal to posterior probabilities of the respective models. By $\pi(\mathscr{M}_k)$ we denote the prior probability of the model $\mathscr{M}_k$.
Note that the predictor (10) can be rewritten in the form

$$\widehat{y}(t+1|t) = \boldsymbol{\varphi}^{\mathrm{T}}(t+1)\widehat{\boldsymbol{\theta}}(t+1|t)$$

where

$$\widehat{\boldsymbol{\theta}}(t+1|t) = \sum_{k=1}^{K} \mu_k(t)\widehat{\boldsymbol{\theta}}_k(t+1|t) \tag{12}$$

denotes the combined estimate of $\boldsymbol{\theta}(t+1)$.
Even though intuitively appealing, the classical Bayesian solution has very limited practical usefulness as it applies to a narrow class of optimal predictors inferred from the underlying data-generating rules. As a result, it can't be used to combine the results yielded by *ad hoc* predictors, such as those described in the previous section.

3.2 Prequential approach

The prequential (predictive + sequential) approach, introduced by Dawid [8], is a general framework for assessing and comparing the predictive performance of forecasting systems. According to the prequential principle, the assessment of the quality of a forecasting system, given a sequence of observed outcomes, should depend only on the forecasts it in fact delivered for that sequence.
To guarantee adaptivity of the fusion rule, at each time instant t our prequential analysis will be restricted to a local decision window $T(t) = [t-M+1,\ldots,t]$, covering the last M input/output measurements $\Omega_T(t) = \{y(i),\boldsymbol{\varphi}(i), i \in T(t)\}$.
We will assume, in addition to (A2), that the measurement noise is distributed according to the complex generalized Gaussian law [9]

$$v \sim \mathscr{CGN}(\psi,\beta)$$
$$p_V(v;\psi,\beta) = \frac{\beta}{2\psi\Gamma(1/\beta)} \exp\left\{-\left(\frac{|v|}{\psi}\right)^{\beta}\right\} \tag{13}$$

where $\Gamma(\cdot)$ is the Gamma function, $\psi > 0$ is the unknown scale parameter and $\beta > 0$ is the known shape parameter.
We note that (12) has a very flexible form which adapts to a large family of symmetric distributions from super-Gaussian to sub-Gaussian, including specific densities such as Laplacian ($\beta = 1$) and Gaussian ($\beta = 2$).
According to the rules of prequential analysis, in order to assess a forecasting system, all parameters [such as $\boldsymbol{\theta}(t)$] and hyperparameters (such as ψ) that are not known at the instant t, should be replaced by their current estimates, i.e., estimates based on $\Omega(t)$, or should be eliminated, e.g. integrated out [10]. For the system (1) this is equivalent to adopting (for assessment purposes only) the following hypothetical model of system parameter variation

$$\mathscr{H}_k(t): \ \ \boldsymbol{\theta}(i) = \widehat{\boldsymbol{\theta}}_k(i|i-1), \ \ i \in T(t)$$

'induced' by the k-th prediction algorithm.
Based on the same principle the unknown scale parameter ψ will be treated as a nuisance parameter with assigned noninformative (improper) prior distribution

$$\pi(\psi|\mathscr{H}_k(t)) = \pi(\psi) \propto \frac{1}{\psi} \tag{14}$$

the form of which stems from the Jeffrey's rule [7].
Finally, we will assume that the hypotheses $\mathscr{H}_k(t)$ are equiprobable

$$\pi(\mathscr{H}_k(t)) = \frac{1}{K}, \ \ k = 1,\ldots,K \tag{15}$$

The 'optimal' prequential rule, allowing one to combine different forecasts, takes the same form as the classical Bayesian rule (10). The only thing that changes is the way of computing the credibility coefficients $\mu_k(t)$, which now are given by

$$\mu_k(t) = P(\mathscr{H}_k(t)|\Omega_T(t)) \propto p(\Omega_T(t)|\mathscr{H}_k(t))$$
$$= \int_0^\infty p(\Omega_T(t)|\mathscr{H}_k(t),\psi)\pi(\psi|\mathscr{H}_k(t))\pi(\mathscr{H}_k(t))d\psi \quad (16)$$

The prequential likelihood function can be expressed in the form

$$p(\Omega_T(t)|\mathscr{H}_k(t),\psi) = \prod_{i\in T(t)} p_V(\varepsilon_k(i))$$
$$= \left[\frac{\beta}{2\psi\Gamma(1/\beta)}\right]^M \exp\left\{-\frac{\sum_{i\in T(t)}|\varepsilon_k(i)|^\beta}{\psi^\beta}\right\} \quad (17)$$

where

$$\varepsilon_k(i) = y(i) - \widehat{y}_k(i|i-1) = y(i) - \boldsymbol{\varphi}^{\mathrm{T}}(i)\widehat{\boldsymbol{\theta}}_k(i|i-1)$$

denotes the prediction error yielded by the k-th predictor. Combining (16) with (14), (15) and (17), after elementary calculations, one arrives at

$$\mu_k(t) \propto \left[\sum_{i\in T(t)} |\varepsilon_k(i)|^\beta\right]^{-M/\beta}. \quad (18)$$

We note that if the one-step-ahead prediction errors are replaced with the so-called matching errors (deleted residuals), the expression (18) is identical with that derived in [11] for the purpose of combining results yielded by several signal smoothers (the smoothed estimates depend on both past and 'future' signal values).
For wide decision windows, i.e., for large values of M, even small differences in the prediction error statistics result in large differences in the values of the corresponding credibility coefficients. Consequently, the major contribution to $\widehat{y}(t+1|t)$ in (12) is due to the predictor that was 'recently the best'. In such a case the weighted estimation rule (12) *de facto* reduces itself to $\widehat{y}(t+1|t) = \widehat{y}_{k^\star(t)}(t+1|t)$ where $k^\star(t) = \arg\max_{1\le k\le K}\mu_k(t)$. When $\beta = 2$, maximization of $\mu_k(t)$ is equivalent to minimization of $\sum_{i\in T(t)}|\varepsilon_k(i)|^2$. This can be regarded as the time-localized version of the Rissanen's predictive least squares principle [12].

4. EXPERIMENTAL AND SIMULATION RESULTS

Because of the lack of space, only the results obtained for the signal-oriented version of the proposed algorithm will be reported here.
In our first experiment we analyzed a real-world acoustic signal - the sound of a motorcycle engine noise, sampled at a frequency 1.1 kHz. The four seconds long recording includes the periods of acceleration (twice), gear shift and braking. The spectrogram of the signal, which consists of $K = 12$ harmonics, is shown in Fig. 1. Fig. 2 shows the time plot of this signal, along with the *a posteriori* estimation errors $\varepsilon_p(t) = s(t) - \widehat{s}(t|t)$ yielded by the unconstrained algorithm (5) and by the constrained algorithm (7)–(8). Note that the constrained algorithm performs considerably better than the

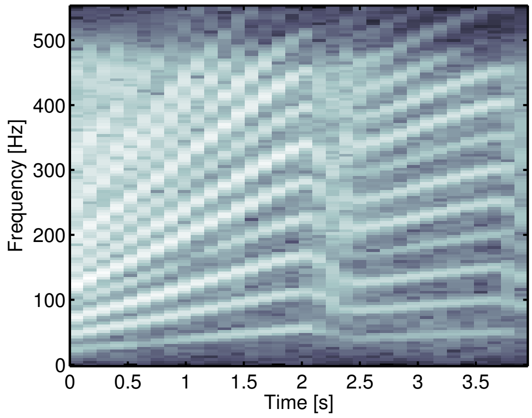

Figure 1: Spectrogram of the engine-induced acoustic noise used in the real-world experiment.

unconstrained one. The same adaptation gains were used in all sub-algorithms. The adaptation gain μ was set to 0.05. The remaining two gains were chosen in agreement with the rules of thumb proposed in [4] : $\gamma = \mu^2/2$, $\eta = \mu^3/8$. The width M of the local decision window was equal to 25. The complex-valued version of the signal was obtained using the Hilbert transform.
The second experiment involved the artificially generated multi-harmonic signal governed by (4), consisting of 4 harmonic components embedded in Gaussian or Laplacian noise with standard deviation $\sigma_v = 0.05$. The frequencies and amplitudes of this signal were varying with time according to

$$\omega_0(t) = \pi\left[0.04 + 0.02\sin\frac{2\pi t}{5000}\right], \quad \omega_k(t) = k\omega_0(t)$$
$$a_k(t) = \frac{1}{\sqrt{k}}\left[0.1 + 0.05\sin\left(\frac{2\pi t}{5000\sqrt{k}} + \phi_k\right)\right]$$
$$\phi_k \sim \mathscr{U}[0,2\pi), \quad k = 1,\ldots,4$$

where $\mathscr{U}[0,2\pi)$ denotes uniform distribution over $[0,2\pi)$. The width of the decision window and the values of the adaptation gains were the same as in the first experiment.
Table 1 summarizes estimation results obtained for 250 Monte Carlo experiments - each simulation run corresponded to a different realization of noise and different initial phase shifts ϕ_k, $k = 1,\ldots,4$. The length of each test signal was equal to 5000. First, for each simulation run, the mean-squared values of the *a posteriori* errors $\varepsilon_p(t)$ were computed by means of time averaging (to eliminate the initial transient effects, the results obtained for the first 250 signal samples were discarded). Then, the mean and median estimation scores were evaluated as ensemble means and ensemble medians, respectively.
Note a dramatic improvement (by 3 to 6 orders of magnitude) provided by the constrained algorithm. Large differences between the mean scores and median scores, observed for all sub-algorihms, reveal one of the sources of this improvement – effective 'elimination' of sub-algorithms which lose track of the harmonics they were initially locked on (which usually results in an attempt to track a whole bunch of nonexistent harmonics). Another important factor is the coordinated search enforced by the comb structure of each sub-algorithm.

Table 1: Mean and median estimation scores obtained for the nonstationary signal with four harmonics, embedded in Gaussian ($\mathscr{N}$) or Laplacian ($\mathscr{L}$) noise: U – unconstrained algorithm; S_1, S_2, S_3, S_4 – four sub-algorithms of the constrained algorithm; C – constrained algorithm.

mean scores

D	U	S_1	S_2	S_3	S_4	C
$\mathscr{N}$	1.27 E−005	3.10 E−007	1.63 E−006	2.26 E−006	2.33 E−005	2.53 E−011
$\mathscr{L}$	1.36 E−005	8.47 E−008	1.04 E−006	3.06 E−006	3.37 E−005	2.65 E−011

median scores

D	U	S_1	S_2	S_3	S_4	C
$\mathscr{N}$	9.03 E−006	2.74 E−011	3.17 E−011	3.52 E−011	2.65 E−011	2.41 E−011
$\mathscr{L}$	1.03 E−005	2.95 E−011	3.21 E−011	3.64 E−011	2.85 E−011	2.50 E−011

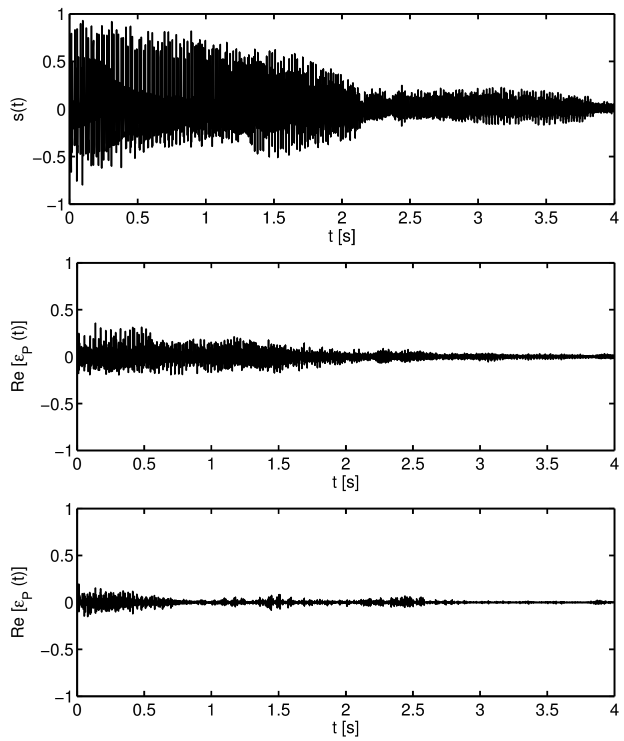

Figure 2: Time plots of the original motorcycle signal (top figure) and the real parts of the *a posteriori* errors yielded by the unconstrained algorithm (middle figure) and constrained algorithm (bottom figure).

5. CONCLUSION

We have proposed a new generalized adaptive comb filtering (GACF) algorithm with a doubly-parallel estimation structure. The final parameter/signal estimates are evaluated as weighted combinations of the estimates provided by several GACF sub-algorithms working in parallel and locked on different harmonics. The corresponding weights are computed using the non-standard Bayesian approach based on the Dawid's concept of prequential probability. The new algorithm outperforms, both in terms of robustness and estimation accuracy, the previously proposed solution.

REFERENCES

[1] G.B. Giannakis and C. Tepedelenlioğlu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Proc. IEEE*, vol. 86, pp. 1969–1986, 1998.

[2] J. Bakkoury, D. Roviras, M. Ghogho and F. Castanie, "Adaptive MLSE receiver over rapidly fading channels," *Signal Processing*, vol. 80, pp. 1347–1360, 2000.

[3] M. Niedźwiecki and M. Meller, "Identification of quasi-periodically varying systems with quasi-linear frequency changes," Proc. 18th IFAC World Congress, Milano, Italy, pp. 9070-9078, 2011.

[4] M. Niedźwiecki and M. Meller, "New algorihms for adaptive notch smoothing," *IEEE Trans. Signal Process.*, vol. 59, pp. 2024-2037, 2011.

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

[6] A. Nehorai and B. Porat, "Adaptive comb filtering for harmonic signal enhancement," *IEEE Trans. Acoust. Speech, Signal Process.*, vol. 34, pp. 1124-1138, 1986.

[7] G. Box and G. Tiao, *Bayesian Inference in Statistical Analysis*. New York: Addison-Wesley, 1973.

[8] A.P. Dawid, "Present position and potential developments: some personal views. Statistical theory. The prequential approach," (with discussion), J. Roy. Stat. Soc., vol. A 147 pp. 278-292, 1984.

[9] M. Novey, T. Adali and A. Roy, "A complex generalized Gaussian distribution – characterization, generation and estimation," *IEEE Trans. Signal Process.*, vol. 58, pp. 1427-1433, 2010.

[10] A.P. Dawid and V.G. Vovk, "Prequential probability: principles and properties," Bernoulli, vol. 5 pp. 125-162, 1999.

[11] M. Niedźwiecki, "Easy recipes for cooperative smoothing," *Automatica*, vol. 46, pp. 716–720, 2010.

[12] J. Rissanen, "A predictive least-squares principle," IMA J. Math. Control Info., vol. 3, pp. 211-222, 1986.