SONIC – Self-optimizing narrowband interference canceler: comparison of two frequency tracking strategies

Maciej Niedźwiecki and Michał Meller

Abstract— **This paper presents a new approach to rejection of complex-valued sinusoidal disturbances acting at the output of a discrete-time linear stable plant with unknown and possibly time-varying dynamics. It is assumed that both the instantaneous frequency of the sinusoidal disturbance and its amplitude may be slowly varying with time and that the output signal is contaminated with wideband measurement noise. The proposed disturbance rejection algorithms are extensions of the algorithm proposed earlier for the constant-known-frequency case. The main purpose of this study is to compare cancellation efficiency under two approaches to frequency tracking: parametric (model-based) and nonparametric (based on decomposition of the signal space). We show that, rather surprisingly, a simple parametric scheme yields results that are of similar quality to those obtained using a high-resolution, computationally much more involved, nonparametric (ESPRIT) scheme.**

I. INTRODUCTION

We will consider the problem of reduction of a complex-valued narrowband disturbance at the output of a discrete-time system governed by (see Fig. 1)

$$y(t) = K_{\mathrm{p}}(q^{-1})u(t-1) + d(t) + v(t) \tag{1}$$

where $y(t)$ denotes the corrupted complex-valued system output, $t = \ldots, -1, 0, 1, \ldots$ denotes normalized time, $K_{\mathrm{p}}(q^{-1})$ denotes unknown transfer function of a linear single-input single-output stable plant, q^{-1} is the backward shift operator,

$$d(t) = a(t)e^{j\phi(t)}, \quad \phi(t) = \sum_{i=1}^{t} \omega(i) \tag{2}$$

is a complex-valued sinusoidal disturbance (cisoid) with slowly varying amplitude $a(t)$ and slowly varying frequency $\omega(t)$, $v(t)$ is a wideband measurement noise, and $u(t)$ denotes the input (control) signal. We will assume that the plant has no zeroes on the unit circle, i.e., $K_{\mathrm{p}}(e^{-j\omega}) \neq 0$, $\forall \omega \in (-\pi, \pi]$.

In certain applications, a reference sensor can be placed close to the source of noise/vibration, providing measurements $r(t)$ that are highly correlated with the unknown disturbance $d(t)$. In such cases, the effects of the disturbance on the output signal $y(t)$ can be minimized using adaptive feedforward compensation techniques. In this study, we will
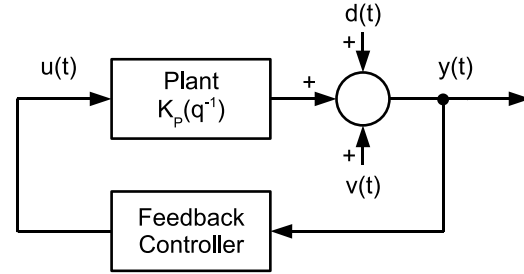
Fig. 1.   Block diagram of the disturbance rejection system.

assume that a reference signal is *not* available, i.e., one has access to the output signal only.

We will look for a feedback controller allowing for cancellation, or near-cancellation, of the sinusoidal disturbance, i.e., a controller generating the complex-valued feedback signal $u(t)$ that minimizes the system output in the mean-squared sense: $\mathrm{E}[\,|\,y(t)\,|^2\,] \longmapsto \min$.

II. THE KNOWN FREQUENCY CASE

Suppose that disturbance has the form

$$d(t) = a(t)e^{j\omega_0 t}$$

where $\omega_0 \in (-\pi, \pi]$ is a known frequency. In the case considered the problem of disturbance rejection is usually solved using the adaptive feedforward cancellation (AFC) technique [1] or the so-called internal model principle (IMP) [2]. For narrowband disturbances of known frequency both approaches are equivalent [3].

An entirely new approach to narrowband disturbance suppression, called SONIC (self-optimizing narrowband interference canceler), was described in [4] (for real-valued signals) and [5] (for complex-valued signals). The complex-valued SONIC algorithm can be summarized as follows:

$$z(t) = e^{j\omega_0}\left[(1 - c_\mu)z(t-1) - \frac{c_\mu}{\widehat{\mu}(t-1)}y(t-1)\right]$$

$$r(t) = \rho r(t-1) + |z(t)|^2$$

$$\widehat{\mu}(t) = \widehat{\mu}(t-1) - \frac{z^*(t)y(t)}{r(t)}$$

$$\widehat{d}(t+1|t) = e^{j\omega_0}[\,\widehat{d}(t|t-1) + \widehat{\mu}(t)y(t)\,]$$

$$u(t) = -\frac{\widehat{d}(t+1|t)}{k_{\mathrm{n}}}. \tag{3}$$

where $\widehat{d}(t+1|t)$ denotes the one-step-ahead prediction of $d(t)$, μ denotes a *complex-valued* adaptation gain, c_μ

is a real-valued constant (e.g. 0.01), k_{n} denotes the assumed (nominal) complex-valued plant gain at the frequency ω_0, $z(t)$ is the local estimate of the sensitivity derivative $\partial y(t; \mu)/\partial \mu$, $*$ denotes complex conjugation, and ρ ($0 < \rho < 1$) is the user-dependent forgetting constant that determines the effective width of the local analysis interval, equal to $1/(1 - \rho)$ samples. Computational burden associated with (3) is equal to 33 real multiply/add operations and 3 real division operations per time update.

SONIC is an adaptive feedback disturbance rejection scheme that combines the coefficient fixing technique, used to "robustify" self-tuning minimum-variance regulators, with automatic adaptation gain tuning. It consists of two loops: the inner control loop [the last two recursions of (3)] that predicts and cancels the disturbance, and the outer, self-optimization loop [the first three recursions of (3)], that automatically adjusts adaptation gain so as to optimize the overall system performance, and – even more importantly – to compensate modeling error caused by adopting in (3) the nominal gain k_{n} in lieu of the true (usually unknown or uncertain) plant gain $k_{\mathrm{p}} = K_{\mathrm{p}}(e^{-j\omega_0})$. SONIC can be easily extended to systems with multiharmonic disturbances and to systems with an extra transport delay. Because of the lack of space none of these extensions will be discussed in this paper.

Both theoretical analysis and simulation experiments confirm very good tracking properties of the proposed scheme. In particular, when the amplitude of the disturbance evolves according to the random-walk model, one can show that, under Gaussian assumptions, SONIC converges locally to the optimal controller.

III. THE UNKNOWN FREQUENCY CASE

When the instantaneous frequency $\omega(t)$ is time-varying but known, the only change that should be introduced into the SONIC algorithm is replacement of ω_0 with $\omega(t)$. This suggests the following extension of (3)

$$
\begin{aligned}
z(t) &= e^{j\widehat{\omega}(t|t-1)}\Big[(1 - c_\mu)z(t-1) \\
&\qquad - \frac{c_\mu}{\widehat{\mu}(t-1)} y(t-1) \Big] \\
r(t) &= \rho r(t-1) + |z(t)|^2 \\
\widehat{\mu}(t) &= \widehat{\mu}(t-1) - \frac{z^*(t)y(t)}{r(t)} \\
\widehat{d}(t+1|t) &= e^{j\widehat{\omega}(t|t-1)}[\widehat{d}(t|t-1) + \widehat{\mu}(t)y(t)] \\
\widehat{\omega}(t+1|t) &= g[\mathcal{Y}(t)] \\
u(t) &= -\frac{\widehat{d}(t+1|t)}{k_{\mathrm{n}}}.
\end{aligned} \tag{4}
$$

where $\widehat{\omega}(t+1|t)$ denotes the one-step-ahead frequency predictor based on the measurements available at instant t: $\mathcal{Y}(t) = \{y(i), i \leq t\}$. Following [6], such a frequency-adaptive algorithm will be further referred to as extended SONIC, or xSONIC. We will consider two methods of frequency estimation: the parametric, model-based approach,

and the nonparametric approach, based on decomposition of the signal space.

A. Parametric approach

We will examine two types of frequency updates

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

$$
\widehat{\omega}(t+1|t) = \widehat{\omega}(t|t-1) - \eta\mathrm{Im}\left[\frac{y^*(t)e^{j\widehat{\omega}(t|t-1)}}{\widehat{d}^*(t|t-1)} \right] \tag{6}
$$

borrowed from the adaptive notch filtering (ANF) algorithms described in [7] and [8], respectively. In both cases η ($0 < \eta \ll 1$) denotes a small adaptation gain that controls the rate of frequency adaptation. We will show that even though in the case where the nominal gain coincides with the true plant gain ($k_{\mathrm{n}} = k_{\mathrm{p}}$) both frequency estimation rules yield comparable tracking results, only the first one (5) is robust to modeling errors.

To get insight into the way modeling errors affect the algorithm (4), we will analyze its inner loop under the assumption that the adaptation gain μ is fixed and not updated:

$$
\begin{aligned}
\widehat{d}(t+1|t) &= e^{j\widehat{\omega}(t|t-1)}[\widehat{d}(t|t-1) + \mu y(t)] \\
\widehat{\omega}(t+1|t) &= g[\mathcal{Y}(t)] \\
u(t) &= -\frac{\widehat{d}(t+1|t)}{k_{\mathrm{n}}}.
\end{aligned} \tag{7}
$$

To derive analytical results we will assume that disturbance is governed by

$$
\begin{aligned}
d(t) &= e^{j\omega(t-1)}d(t-1) \\
\omega(t) &= \omega(t-1) + w(t)
\end{aligned} \tag{8}
$$

where $w(t)$ denotes an uniformly small perturbation. According to (8) $d(t)$ is a constant-modulus cisoid with unknown magnitude $a = |d(t)|$ and slowly varying frequency $\omega(t)$. Our approach will be based on averaging. Consider a local analysis window $T = [t_1, t_2]$, covering $n = t_2 - t_1 + 1$ consecutive time instants ($n \gg 2\pi/\omega(t), \forall t \in T$). If the transfer function of the plant $K_{\mathrm{p}}(e^{-j\omega})$ is a smooth function of ω, and if the instantaneous frequency of the disturbance changes sufficiently slowly with time, the true response of the plant to the narrowband excitation $u(t)$ can be approximated by

$$
K_{\mathrm{p}}(q^{-1})u(t-1) \cong k_T u(t-1), \ t \in T
$$

where $k_T = \sum_{t \in T} K_{\mathrm{p}}(e^{-j\omega(t)})/n$ denotes the average plant gain over the interval T. Using this approximation, one can express plant output in the form

$$
y(t) \cong d(t) - \beta\widehat{d}(t|t-1) + v(t) , \ t \in T \tag{9}
$$

where $\beta = k_T/k_{\mathrm{n}}$ — the ratio of the average plant gain to the nominal (assumed) gain — denotes the local modeling error. In our local analysis, β will be regarded as a time-invariant quantity.

Denote the cancellation error by $\Delta\widehat{d}(t) = d(t) - \beta\widehat{d}(t|t-1)$, and the one-step-ahead frequency prediction error by $\Delta\widehat{\omega}(t) = \omega(t) - \widehat{\omega}(t|t-1)$. To establish the dependence of $\Delta\widehat{d}(t)$ and $\Delta\widehat{\omega}(t)$ on $v(t)$ and $w(t)$, we will employ the approximating linear filter (ALF) technique, proposed by Tichavský and Händel [7], for the purpose of analyzing adaptive notch filters. Using this approach, one arrives at the following approximations valid for the algorithm (7)+(5) (see Appendix I):

$$\Delta\widehat{x}(t) = (1 - \mu\beta)\Delta\widehat{x}(t-1) + ja^2\Delta\widehat{\omega}(t-1)$$
$$- \mu\beta z(t-1)$$
$$\Delta\widehat{\omega}(t+1) = \Delta\widehat{\omega}(t) - \frac{\eta}{a^2}\operatorname{Im}[\mu\beta\Delta\widehat{x}(t)] - \frac{\eta}{a^2}\operatorname{Im}[\mu\beta z(t)]$$
$$+ w(t+1) \qquad (10)$$

where $\Delta\widehat{x}(t) = \Delta\widehat{d}(t)d^*(t)$ and $z(t) = v(t)d^*(t)$.

The analogous approximations for the algorithm (7)+(6) take the form (see Appendix II):

$$\Delta\widehat{x}(t) = (1 - \mu\beta)\Delta\widehat{x}(t-1) + ja^2\Delta\widehat{\omega}(t-1)$$
$$- \mu\beta z(t-1)$$
$$\Delta\widehat{\omega}(t+1) = \Delta\widehat{\omega}(t) - \frac{\eta}{a^2}\operatorname{Im}[\beta\Delta\widehat{x}(t)] - \frac{\eta}{a^2}\operatorname{Im}[\beta z(t)]$$
$$+ w(t+1) \qquad (11)$$

ALF equations (10) and (11) can be used to study tracking performance of the corresponding algorithms.

In the absence of modeling errors ($\beta = 1$) both algorithms perform similarly. Actually, suppose that $\{w(t)\}$ is a white noise sequence, i.e., that the instantaneous frequency evolves according to the random-walk model (which is a widely used benchmark in tracking analysis). Then, after elementary but tedious ALF-based calculations, one can establish analytical expressions for the steady-state mean-squared frequency estimation error and a steady-state mean-squared cancellation error, respectively – see [7] and [8] for more details. Using these expressions, one can show that, when optimally tuned, both algorithms yield the same mean-squared errors, and that in the Gaussian case they both are statistically efficient frequency trackers, reaching the lower frequency tracking bound established in [7] (called the posterior Cramér-Rao bound).

In the presence of modeling errors ($\beta \neq 1$) situation is different. According to the ALF equations (10), tracking properties of the algorithm (5)+(7) can be quantified in terms of $\mu\beta$ and η. Since modeling error β does not interfere with η, the value of η can be safely fixed at a constant level, based on our prior knowledge of the expected rate of frequency variation. Moreover, when the automatically adjusted adaptation gain μ is allowed to take complex values, modeling error can be compensated by feedback – as shown in [5], $\widehat{\mu}(t)$ converges in mean to such value $\mu_0 \in \mathbb{C}$ which guarantees that $\mu_0\beta = \mu_{\mathrm{opt}} > 0$, where μ_{opt} denotes the optimal value of μ in the absence of modeling error.

Note that the analogous error compensation is not possible when the algorithm (6)+(7) is used. According to the ALF equations (11), tracking properties of this algorithm depend on the values of $\mu\beta$ and $\eta\beta$. Hence, for a fixed value of $\eta > 0$, modeling error (which is unknown and possibly time-varying) will affect tracking performance of the algorithm no matter how μ is chosen.

The discussion carried out above shows clearly only the algorithm (7)+(5) is robust to modeling errors.

B. Nonparametric approach

Following [9], one can track the instantaneous frequency of the disturbance using an adaptive, e.g. sliding window, version of the ESPRIT (Estimation of Signal Parameters via Rotational Invariance Techniques) algorithm. ESPRIT is a high-resolution subspace-based method used in spectral analysis. The essence of the sliding window ESPRIT lies in estimating the frequency-dependent signal subspace rotation matrix $\mathbf{\Psi}(t)$ (the so-called spectral matrix), based on the results of singular value decomposition (SVD) of the $N \times M$ data matrix $\mathbf{X}(t)$, made up of $K = N+M-1$ recent samples of the analyzed signal $s(t)$:

$$\mathbf{X}(t) = \begin{bmatrix} s(t) & s(t-1) & \cdots & s(t-M+1) \\ s(t-1) & s(t-2) & \cdots & s(t-M) \\ \vdots & \vdots & & \vdots \\ s(t-N+1) & s(t-M) & \cdots & s(t-N-M+2) \end{bmatrix}$$

When the reference signal is not available, to guarantee sufficient spectral richness of $s(t)$ one can analyze the mixture (the reconstructed unattenuated output signal)

$$s(t) = \widehat{d}(t|t-1) + y(t)$$

Rather than presenting mathematical details of the ESPRIT approach, we will cite – after [10] – the corresponding Matlab code for detection of P complex exponentials

```
[L, S, U] = svd(X);
Us = U(:, 1 : P);
U1 = Us(1 : (M − 1), :);
U2 = Us(2 : M, :);
[LL, SS, UU] = svd([U1U2]);
UU12 = UU(1 : P, (P + 1) : (2 ∗ P));
UU22 = UU((P + 1) : (2 ∗ P), (P + 1) : (2 ∗ P));
Psi = −UU12 ∗ inv(UU22);
phi = eig(Psi);
omegahat = angle(diag(phi));
```

In the procedure listed above, which should be executed every time instant t, `X` stands for $\mathbf{X}(t)$, `Psi` stands for $\mathbf{\Psi}(t)$ and `omegahat` denotes $\widehat{\omega}(t+1|t)$.

IV. COMPUTER SIMULATIONS

Two simulation experiments were arranged to compare tracking/cancellation properties of the parametric scheme (4)+(7), and the nonparametric, ESPRIT-based scheme. The simulated discrete-time plant $K_{\mathrm{p}}(q^{-1}) = 0.0952/(1 - 0.9048q^{-1})$ was adopted from [9] and corresponds to a continuous-time plant with transfer function
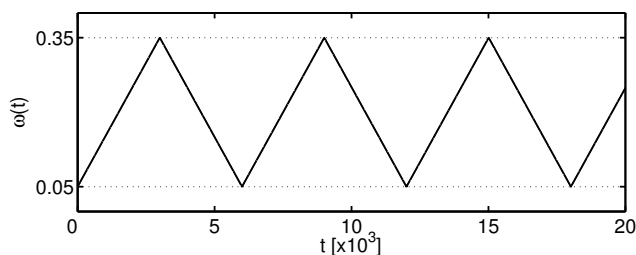
Fig. 2. Example of instantaneous frequency changes (for the sweep rate equal to $\Delta_\omega = 10^{-4}$ [rad/Sa]).

$K_{\mathrm{p}}(s) = 1/(1 + 0.01s)$, sampled at the rate of 1 kHz. The nominal gain was set to $k_{\mathrm{n}} = 1$. To avoid erratic behavior of the cancellation algorithm, the maximum allowable value for $|\widehat{\mu}(t)|$ was set to $\mu_{\max} = 0.1$.

In the first experiment disturbance was a chirped signal, governed by (8), and the instantaneous frequency was changing in a sawtooth manner between 0.05 and 0.35 – see Fig. 1. In each of the linear growth/decay periods the evolution of $\omega(t)$ was governed by $\omega(t) = \omega(t-1) \pm \Delta_\omega$, where Δ_ω denotes the frequency sweep rate.

Tracking/cancellation performance of the compared algorithms depends on their design parameters η, K and M, on the sweep rate Δ_ω, and on the signal-to-noise ratio SNR=a^2/σ_v^2. Therefore, not to compare "apples with oranges", we have optimized both schemes for each of the considered sweep rates (10^{-5}, $3 \cdot 10^{-5}$, $6 \cdot 10^{-5}$, 10^{-4}, $3 \cdot 10^{-4}$, $6 \cdot 10^{-4}$ and 10^{-3} [rad/Sa]) and signal-to-noise ratios (20, 40 [dB]), i.e., we selected such value of η (for the parametric scheme), and such values of K and $M < K$ (for the nonparametric scheme) which yield the smallest mean-squared cancellation errors. The values of η were chosen from the interval [0.01, 0.18], the values of K were selected from the set $\{4, 6, 8, 12, 16, 24, 32, 48, 64\}$ and the values of M – from the set $\{2, 4, 6, 8, 12, 16, 24, 32\}$.

For SNR=40 dB the optimal values of η ranged from 0.12 (for the smallest sweep rate) to 0.18 (for the largest sweep rate), the optimal values of K ranged from 8 (for the smallest sweep rate) to 4 (for the largest sweep rate), and the optimal value of K was always equal to 2. For SNR=20 dB the optimal values of η ranged from 0.06 to 0.14, and the optimal values of (K, M) ranged from (64,32) to (8,2).

The results, gathered in Figs. 3 and 4, clearly indicate that under a wide range of operating conditions the parametric scheme and nonparametric scheme perform similarly. Rather surprisingly, the computationally involved ESPRIT algorithm, which has a reputation for being the highest-resolution frequency estimation tool, does not outperform a simple gradient algorithm (5).

Some advantages of ESPRIT (faster and more accurate response to changes) are revealed by our second experiment, where frequency changed in a step manner. The results, shown in Fig. 5, were obtained for SNR=30 dB and for design parameters which guarantee that the compared algorithms yield the same mean-squared cancellation errors in the constant-frequency case ($\eta = 0.05$, $K = 32$, $M = 16$).
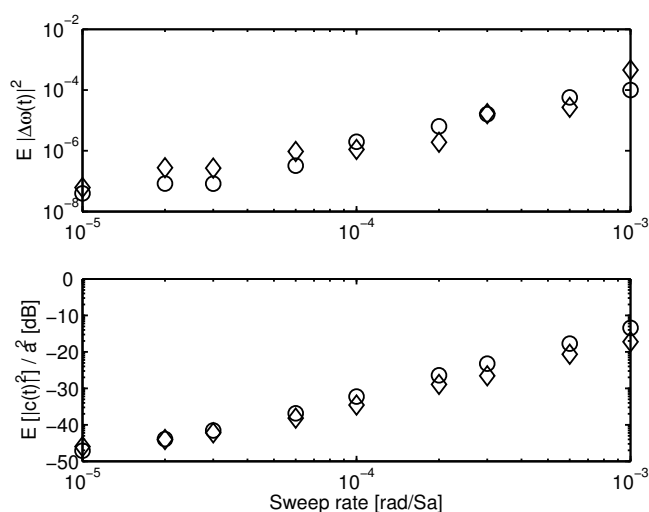

Fig. 3. Comparison of the mean-squared frequency tracking errors (upper plot) and normalized mean-squared cancellation errors (lower plot), obtained for the xSONIC algorithm equipped with the parametric frequency estimator (circles) and nonparametric (ESPRIT) frequency estimator (diamonds). All results were obtained for SNR=40 dB.
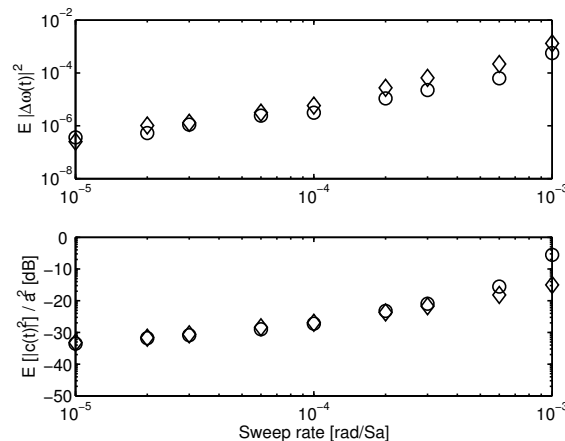

Fig. 4. Comparison of the mean-squared frequency tracking errors (upper plot) and normalized mean-squared cancellation errors (lower plot), obtained for the xSONIC algorithm equipped with the parametric frequency estimator (circles) and nonparametric (ESPRIT) frequency estimator (diamonds). All results were obtained for SNR=20 dB.

V. CONCLUSION

We have examined how selection of the frequency tracking algorithm (parametric or high-resolution nonparametric) affects cancellation efficiency of an active narrowband noise controller. We have demonstrated that, at least in the single frequency case, a simple parametric tracker yields similar results as a computationally more demanding ESPRIT-based tracker.

REFERENCES

[1] S.M. Kuo and D.R. Morgan, *Active Noise Control Systems: Algorithms and DSP Implementations*. New York: Wiley, 1996.
[2] S.J. Elliott and P.A. Nelson, "Active noise control," *IEEE Signal Process. Mag.*, vol. 10, pp. 12–35, Oct. 1993.
[3] M. Bodson, "Rejection of periodic disturbances of known and time-varying frequency," *Int. J. Adapt. Contr. and Signal Process.* vol. 19, pp. 67–88, 2005.
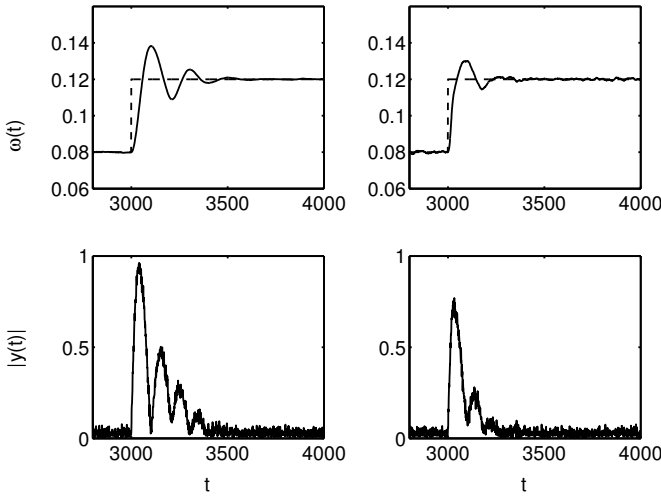
Fig. 5. Transient responses to abrupt frequency change, obtained for the xSONIC algorithm equipped with the parametric frequency estimator (left plots) and nonparametric (ESPRIT) frequency estimator (right plots).

[4] M. Niedźwiecki and M. Meller, "Self-optimizing adaptive vibration controller," *IEEE Trans. on Automatic Control*, vol. 54, pp. 2087–2099, Sept. 2009.
[5] M. Niedźwiecki and M. Meller, "A new approach to active noise and vibration control - Part I: The known frequency case," *IEEE Trans. on Signal Processing*, vol. 57, pp. 3373–3386, Sept. 2009.
[6] M. Niedźwiecki and M. Meller, "A new approach to active noise and vibration control - Part II: The unknown frequency case," *IEEE Trans. on Signal Processing*, vol. 57, pp. 3387–3398, Sept. 2009.
[7] P. Tichavský and P. Händel, "Two algorithms for adaptive retrieval of slowly time-varying multiple cisoids in noise," *IEEE Trans. on Signal Processing*, vol. 43, pp. 1116–1127, May 1995.
[8] M. Niedźwiecki and P. Kaczmarek, "Tracking analysis of a generalized adaptive notch filter," *IEEE Trans. on Signal Processing*, vol. 54, pp. 304–314, Jan. 2006.
[9] X. Guo and M. Bodson, "Adaptive rejection of multiple sinusoids of unknown frequency," in *Proc. European Control Conference*, Kos, Greece, 2007, pp. 121–128.
[10] D.G. Manolakis, V.K. Ingle and S.M. Kogon, *Statistical and adaptive signal processing: spectral estimation, signal modeling, adaptive filtering, and array processing*. Boston: McGraw-Hill, 2000.

APPENDIX I [derivation of (10)]

According to [7], when carrying out the ALF analysis of (8), one should examine dependence of $\Delta\widehat{d}(t)$ and $\Delta\widehat{\omega}(t)$ on $v(t)$ and $w(t)$, neglecting higher than first-order terms of all quantities listed above, including all cross-terms. To arrive at the recursive formula for $\Delta\widehat{d}(t)$, note that

$$\Delta\widehat{d}(t) = d(t) - \beta\widehat{d}(t\,|\,t-1) = e^{j\omega(t-1)}d(t-1)$$
$$- \beta e^{j\omega(t-1)}e^{-j\Delta\widehat{\omega}(t-1)}[\widehat{d}(t-1\,|\,t-2)$$
$$+ \mu\Delta\widehat{d}(t-1) + \mu v(t-1)]. \qquad (12)$$

When tracking is satisfactory, i.e., when $\Delta\widehat{\omega}(t)$ is small, one obtains $e^{-j\Delta\widehat{\omega}(t-1)} \cong 1 - j\Delta\widehat{\omega}(t-1)$ which, after substitution in (12) and neglecting all higher-order terms, leads to the following approximation

$$\Delta\widehat{d}(t) = e^{j\omega(t-1)}d(t-1)$$
$$- \beta e^{j\omega(t-1)}[1 - j\Delta\widehat{\omega}(t-1)]\widehat{d}(t-1\,|\,t-2)$$
$$- \beta\mu e^{j\omega(t-1)}\Delta\widehat{d}(t-1) - \beta\mu e^{j\omega(t-1)}v(t-1).$$

Furthermore, since $\widehat{d}(t-1\,|\,t-2) = [d(t-1) - \Delta\widehat{d}(t-1)]/\beta$, after substitution, regrouping, and neglecting the term proportional to $\Delta\widehat{\omega}(t-1)\Delta\widehat{d}(t-1)$, one obtains

$$\Delta\widehat{d}(t) = e^{j\omega(t-1)}[1 - \beta\mu]\Delta\widehat{d}(t-1) - \beta\mu e^{j\omega(t-1)}v(t-1)]$$
$$+ je^{j\omega(t-1)}d(t-1)\Delta\widehat{\omega}(t-1). \qquad (13)$$

Finally, after multiplying both sides of (13) by $d^*(t) = e^{-j\omega(t-1)}d^*(t-1)$, one arrives at the first recursion of (10).

To derive the second recursion, note that

$$g(t) = \text{Arg}\left[\frac{\widehat{d}(t+1\,|\,t)e^{-j\widehat{\omega}(t\,|\,t-1)}}{\widehat{d}(t\,|\,t-1)}\right]$$
$$= \text{Im}\left\{\log\left[1 + \frac{\mu y(t)}{\widehat{d}(t\,|\,t-1)}\right]\right\} \cong \text{Im}\left[\frac{\mu y(t)}{\widehat{d}(t\,|\,t-1)}\right].$$

Assuming that the cancellation error $\Delta\widehat{d}(t)$ is small, which means that $d(t) \cong \beta\widehat{d}(t\,|\,t-1)$, one arrives at

$$g(t) \cong \text{Im}\left[\frac{\mu\beta y(t)}{d(t)}\right] = \frac{1}{a^2}\text{Im}[\mu\beta\Delta\widehat{d}(t)d^*(t)]$$
$$+ \frac{1}{a^2}\text{Im}[\mu\beta v(t)d^*(t)]. \qquad (14)$$

Combining (5) with (14), one obtains

$$\widehat{\omega}(t+1\,|\,t) = \widehat{\omega}(t\,|\,t-1) + \frac{\eta}{a^2}\text{Im}[\mu\beta\Delta\widehat{d}(t)d^*(t)]$$
$$+ \frac{\eta}{a^2}\text{Im}[\mu\beta v(t)d^*(t)].$$

Finally, subtracting this equation (sidewise) from $\omega(t+1) = \omega(t) + w(t+1)$, one arrives at the second recursion of (10).

APPENDIX II [derivation of (11)]

Derivation of the first recursion of (11) coincides with that given in Appendix I. To obtain the second approximation, note that $d(t) \cong \beta\widehat{d}(t\,|\,t-1)$ [which holds true for small $\Delta\widehat{d}(t)$] entails $|\widehat{d}(t\,|\,t-1)|^2 \cong |d(t)/\beta|^2 = a^2/|\beta|^2$, leading to

$$\frac{y^*(t)e^{j\widehat{\omega}(t\,|\,t-1)}}{\widehat{d}^*(t\,|\,t-1)} = \frac{y^*(t)e^{j\widehat{\omega}(t\,|\,t-1)}\widehat{d}(t\,|\,t-1)}{|\widehat{d}(t\,|\,t-1)|^2}$$
$$\cong \frac{y^*(t)e^{j\widehat{\omega}(t\,|\,t-1)}\widehat{d}(t\,|\,t-1)|\beta|^2}{a^2}$$

Furthermore

$$y^*(t)e^{j\widehat{\omega}(t\,|\,t-1)}\widehat{d}(t\,|\,t-1)|\beta|^2$$
$$= [\Delta\widehat{d}^*(t) + v^*(t)]\frac{d(t) - \Delta\widehat{d}(t)}{\beta}$$
$$\cong \frac{\Delta\widehat{d}^*(t)d(t)}{\beta} + \frac{v^*(t)d(t)}{\beta} = \Delta\widehat{x}^*(t)/\beta + z^*(t)/\beta$$

and hence

$$\widehat{\omega}(t+1\,|\,t) \cong \widehat{\omega}(t\,|\,t-1) + \frac{\eta}{a^2}\text{Im}[\beta\Delta\widehat{x}(t)] + \frac{\eta}{a^2}\text{Im}[\beta z(t)]$$

Subtracting this equation, sidewise, from $\omega(t+1) = \omega(t) + w(t+1)$, one obtains the second recursion of (11).