An FIR Notch Filter for Adaptive Filtering of a Sinusoid in Correlated Noise

Osman Kukrer and Aykut Hocanin

Department of Electrical and Electronics Engineering, Eastern Mediterranean University, Gazimagusa, Mersin 10, Turkey

A novel adaptive FIR filter for the estimation of a single-tone sinusoid corrupted by additive noise is described. The filter is based on an offline optimization procedure which, for a given notch frequency, computes the filter coefficients such that the frequency response is unity at that frequency and a weighted noise gain is minimized. A set of such coefficients is obtained for notch frequencies chosen at regular intervals in a given range. The filter coefficients corresponding to any frequency in the range are computed using an interpolation scheme. An adaptation algorithm is developed so that the filter tracks the sinusoid of unknown frequency. The algorithm first estimates the frequency of the sinusoid and then updates the filter coefficients using this estimate. An application of the algorithm to beamforming is included for angle-of-arrival estimation. Simulation results are presented for a sinusoid in correlated noise, and compared with those for the adaptive IIR notch filter.

1. INTRODUCTION

Estimation of sinusoidal signals and their frequencies from noisy measurements is important in many fields such as angle of arrival estimation, frequency-shift keying (FSK) demodulation, Doppler estimation of radar waveforms, biomedical engineering, sensor array processing, and cancellation of periodic interferences [1]. The observed signal has the following general form:

$$x(k) = a\cos(k\theta + \phi) + q(k). \qquad (1)$$

The problem in many applications is to recover the signal and/or its frequency θ, from the noisy observations $x(k)$. Various adaptive filtering algorithms have been introduced for solving such problems. The least-mean-square (LMS) algorithm [2] based on the FIR transversal filter has been widely used due to its simplicity and robustness. On the other hand, the performance of this algorithm deteriorates when the input signal is correlated [3]. Transform-domain techniques have been introduced to decorrelate the input signal and achieve faster convergence [3, 4]. Also, in certain applications, the filter length required for a satisfactory performance is large. The adaptive IIR filter, also known as the adaptive notch filter [5–7], has been introduced as an alternative to the LMS FIR filter. The IIR filter has the outstanding advantage of requiring considerably fewer coefficients compared

with its FIR counterpart. However, the performance of the IIR notch filter in correlated noise has not been studied well in the literature. In [8], an adaptive IIR notch filter for suppressing narrow-band interference is described, where the filter's bandwidth is adaptively controlled to maximize SNR.

In this paper, a notch filter based on the FIR structure is presented which has offline optimized magnitude responses [9] at all frequencies in the range $[0, \pi]$. The magnitude frequency response of the proposed filter is designed to minimize a criterion which depends on the noise suppression performance and takes into account the power spectrum of the noise. It is assumed that the noise power is concentrated in a certain frequency range which can be estimated. The filter coefficients are then adapted to track the input signal by using filter coefficients stored at preselected frequencies in the range $[0, \pi]$. In this way, an adaptive FIR notch filter with frequency responses optimized to reject correlated noise is obtained. This approach resembles that of [10] which employs online constraints for waveform estimation in the frequency domain. It is shown that the proposed filter provides performance gains compared with the adaptive IIR notch filter in terms of signal and frequency estimation, at frequencies outside of the noise band.

The proposed adaptive filter is suitable for adaptive line enhancer applications where the noise is correlated and the power spectra can be estimated (e.g., using periodogram

techniques). It can also be successfully employed in adaptive beamforming applications [2], where interference in certain directions can be effectively suppressed.

The paper is organized as follows. In Section 2, the optimization of the FIR notch filter is described. In Section 3, the adaptation algorithm for tracking the input signal is obtained and stability analysis is performed in Section 4. Section 5 describes an example application of the algorithm to beamforming. Section 6 presents the simulation results for the proposed method and for the adaptive IIR filter.

2. OFFLINE-OPTIMIZED FIR NOTCH FILTER

Consider a predictive FIR filter of length N,

$$\hat{x}_o(k + 1) = \sum_{n=0}^{N-1} h_{o,n} x_o(k - n), \qquad (2)$$

where $\hat{x}_o$ is the output of the filter and $h_{o,n}$, $n = 0, \ldots, N-1$, are the filter coefficients which will be optimized. In order that the filter can predict a sinusoidal signal $x_o(k) = A \cos(k\theta_0 + \phi)$ at frequency θ_0, the following equations must be satisfied:

$$\sum_{n=0}^{N-1} h_{o,n} \cos(n\theta_0) = \cos\theta_0,$$
$$\sum_{n=0}^{N-1} h_{o,n} \sin(n\theta_0) = -\sin\theta_0. \qquad (3)$$

Equation (3) can be written in matrix form as

$$\mathbf{A}\mathbf{h_o} = \mathbf{p}, \qquad (4)$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & \cos\theta_0 & \cos 2\theta_0 & \cdots & \cos\{(N-1)\theta_0\} \\ 0 & \sin\theta_0 & \sin 2\theta_0 & \cdots & \sin\{(N-1)\theta_0\} \end{bmatrix},$$
$$\mathbf{p} = \begin{bmatrix} \cos\theta_0 & -\sin\theta_0 \end{bmatrix}^T. \qquad (5)$$

This filter can be optimized by minimizing a cost function which depends on the noise suppression performance, subject to the constraints in (4). The frequency response of the filter in (2) is

$$H(\theta) = \sum_{n=0}^{N-1} h_{o,n} e^{-j(n+1)\theta}. \qquad (6)$$

In order to suppress the correlated noise frequency components and design the frequency response of the proposed filter, the following cost function is defined:

$$J_1 = \sum_{m=1}^{M} w_m |H(\theta_m)|^2, \qquad (7)$$

where θ_m, $m = 1, \ldots, M$, is the frequency range represented by M samples and w_m, $m = 1, \ldots, M$, are the weights that can be used to shape the frequency response. Minimization

of J_1 subject to the constraints in (4) can be achieved by using the method of Lagrange multipliers. Incorporating the constraint (4) in the cost function, we obtain

$$J_2 = \mathbf{h_o^T C h_o} + \boldsymbol{\lambda}^T(\mathbf{A h_o} - \mathbf{p}), \qquad (8)$$

where $\boldsymbol{\lambda} = [\lambda_1 \ \lambda_2]^T$ is the vector of Lagrange multipliers. The minimization of J_2 with respect to $\mathbf{h}_o$ results in the following equations:

$$\sum_{i=1}^{N} h_{o,i-1} c_{ij} = -\frac{1}{2}(\lambda_1 a_{1,j} + \lambda_2 a_{2,j}), \quad j = 1, \ldots, N, \qquad (9)$$

$$c_{ij} = \sum_{m=1}^{M} w_m \cos[(j - i)\theta_m], \quad i, j = 1, \ldots, N. \qquad (10)$$

In (9), $a_{1,j}$ and $a_{2,j}$ are the elements of $\mathbf{A}$. Equation (9) can be put into matrix form as

$$\mathbf{C h_o} = -\frac{1}{2}\mathbf{A^T}\boldsymbol{\lambda}. \qquad (11)$$

In order to solve for the multipliers, we substitute (11) in (4). Then the vector $\boldsymbol{\lambda}$ can be solved as

$$\boldsymbol{\lambda} = -2(\mathbf{A C^{-1} A^T})^{-1}\mathbf{p}. \qquad (12)$$

Finally, the filter coefficient vector $\mathbf{h}_o$ can be obtained from (11) as

$$\mathbf{h_o} = \mathbf{C^{-1} A^T}(\mathbf{A C^{-1} A^T})^{-1}\mathbf{p}. \qquad (13)$$

3. THE ADAPTIVE FILTER

Consider an adaptive predictive FIR filter of the form

$$\hat{x}(k + 1) = \sum_{n=0}^{N-1} h_n(k) x(k - n) = \mathbf{h^T}(k)\mathbf{x}_N(k), \qquad (14)$$

where $\hat{x}(k + 1)$ is the output of the filter, $\mathbf{h}(k) = [h_0(k) \ \cdots \ h_{N-1}(k)]^T$ is the vector of the filter coefficients at time step k, and

$$\mathbf{x}_N(k) = [x(k)x(k-1)\cdots x(k-N+1)]^T \qquad (15)$$

is the observed data vector, where $x(k) = x_o(k) + q(k)$. It is assumed that $q(k)$ is a zero-mean Gaussian random process. The adaptive FIR filter will be designed such that it predicts a sinusoid x_o of any frequency θ_s by utilizing the optimum filter coefficients computed at a selected frequency θ_0, where θ_s is assumed to lie in a certain neighborhood of θ_0. The prediction error is defined as

$$e_p(k + 1) = x(k + 1) - \hat{x}(k + 1). \qquad (16)$$

With the filter coefficient vector fixed at $\mathbf{h}_o$, an error transfer function can be defined as

$$H_{e,o}(z) = \frac{E_p(z)}{X(z)} = 1 - \sum_{n=0}^{N-1} h_{o,n} z^{-(n+1)}. \qquad (17)$$

Note that (17) is a notch filter tuned at the frequency θ_0. Therefore, the first pair of the zeros $(z_{0,1}, z_{0,2})$ of the polynomial in (17) corresponds to the frequency θ_0. Hence $H_{e,o}(z)$ can be written as

$$H_{e,o}(z) = (1 - 2\cos(\theta_0)z^{-1} + z^{-2})\prod_{l=3}^{N}(1 - z_{0,l}z^{-1}), \quad (18)$$

where $\{z_{0,l},\ l = 3,\ldots,N\}$ are the remaining zeros of the polynomial. The proposed adaptive filter is based on the following parameterized error transfer function:

$$H_e(z, \alpha) = (1 - \alpha z^{-1} + z^{-2})H_n(z), \quad (19)$$

where $H_n(z)$ denotes the product term in (18), which has constant coefficients of the powers of z. Now, $H_e(z, \alpha)$ can be written in the form of $H_{e,o}(z)$ in (17) as

$$H_e(z, \alpha) = 1 - \sum_{n=0}^{N-1} h_n(\alpha)z^{-(n+1)}. \quad (20)$$

The α-dependent filter coefficients can be obtained by expanding $H_e(z, \alpha)$ in (19) as a power series. It is obvious that the coefficients will be linear in α, as follows:

$$h_n(\alpha) = a_n + \alpha b_n, \quad n = 0,\ldots,N-1. \quad (21)$$

The filter will be optimal with respect to the cost function in (7) when $\alpha = \alpha_0 = 2\cos(\theta_0)$. However, it can be argued that the noise power will not change significantly when α is in the vicinity of α_0. Note that (20) is a notch filter tuned at the frequency $\theta = \cos^{-1}(\alpha/2)$. Filtered prediction error is defined as follows:

$$\widetilde{e}_p(k+1) = \sum_{n=0}^{N-1} h_{e,n}e_p(k-n), \quad (22)$$

where $\mathbf{h_e}$ may be chosen equal to $\mathbf{h_o}$. This error may be assumed to be equal to the difference between the original signal $x_o(k+1)$ and the prediction, where θ_s is the original signal frequency,

$$\widetilde{e}_p(k+1) \cong x_o(k+1) - \hat{x}(k+1) = x_o(k+1) - \mathbf{h}^T(\alpha)\mathbf{x}_N(k). \quad (23)$$

Now, if the following assumption is made:

$$x_o(k+1) \cong \mathbf{h}^T(\alpha_s)\mathbf{x}_N(k), \quad (24)$$

where $\alpha_s = 2\cos\theta_s$, then (23) can be written as

$$\widetilde{e}_p(k+1) = [\mathbf{h}(\alpha_s) - \mathbf{h}(\alpha)]^T\mathbf{x}_N(k) = (\Delta\mathbf{h})^T\mathbf{x}_N(k). \quad (25)$$

The correction in the coefficient vector can be approximated by

$$\Delta\mathbf{h} \cong \left.\frac{\partial\mathbf{h}}{\partial\alpha}\right|_{\alpha_s}\Delta\alpha \cong [-2\sin(\theta_s)\Delta\hat{\theta}_s]\mathbf{b}, \quad (26)$$

where $\mathbf{b} = [b_0 \cdots b_{N-1}]^T$, and $\Delta\hat{\theta}_s$ is the correction in the estimated frequency of the signal. Substituting (26) in (25)

and solving for $\Delta\hat{\theta}_s$, the updating scheme for $\hat{\theta}_s$ is obtained as

$$\hat{\theta}_s(k+1) = \hat{\theta}_s(k) - \mu\frac{\widetilde{e}_p(k+1)}{[2\sin(\hat{\theta}_s(k))](\mathbf{b}^T\mathbf{x}_N(k))}, \quad (27)$$

where μ is a suitably chosen stepsize. The coefficient vector is then updated using

$$\mathbf{h}(k+1) = \mathbf{h}[\alpha(k+1)] = \mathbf{a} + \mathbf{b}\alpha(k+1), \quad (28)$$

where $\alpha(k+1) = 2\cos(\hat{\theta}_s(k+1))$. Note that the term $\mathbf{b}^T\mathbf{x}_N(k)$ in (27) may become arbitrarily small since it is a linear combination of noisy sinusoids. In such case, the correction in the frequency cannot be solved from (25) and (26). Higher-order terms may be required in the series expansion in (26) for the solution of $\Delta\hat{\theta}_s$. However, in such case, $\widetilde{e}_p$ will become nonlinear in $\Delta\hat{\theta}_s$. This is avoided by equating the correction $\Delta\hat{\theta}_s$ to zero in such a singular case, where μ is set to zero whenever

$$|\mathbf{b}^T\mathbf{x}_N(k)| < \varepsilon. \quad (29)$$

Here, ε is a threshold for successive updates. When ε is too small, it may lead to instability. On the other hand, a large value of ε may result in decreased tracking performance. Note that ε must be chosen less than the maximum amplitude of $\mathbf{b}^T\mathbf{x}_N(k)$. A reasonably accurate initial value of the frequency estimate can be obtained by a periodogram with a relatively short length FFT.

With the thresholding of the second term, the update equation for $\hat{\theta}_s(k)$ can be written as

$$\hat{\theta}_s(k+1) = \hat{\theta}_s(k) - \mu(k)\frac{\widetilde{e}_p(k+1)}{[2\sin(\hat{\theta}_s(k))](\mathbf{b}^T\mathbf{x}_N(k))}, \quad (30)$$

where

$$\mu(k) = \begin{cases} \mu & |\mathbf{b}^T\mathbf{x}_N(k)| > \varepsilon, \\ 0 & \text{otherwise.} \end{cases} \quad (31)$$

In the implementation of the proposed adaptive filter, the frequency range $[0, \pi]$ is divided into $L = 18$ intervals. The optimum filter coefficient vector is calculated at the centre frequency $\theta_0(l)$, $l = 1,\ldots,L$, of each interval. Figure 1 shows the frequency response of the optimized filter tuned at the frequency $\pi/4$. The vectors $(\mathbf{a}, \mathbf{b})$ are then calculated offline (only once) by (18) using symbolic computation and then stored in processor memory. Given the frequency estimate $\hat{\theta}_s(k)$ at time k, the filter coefficient vector is then calculated using $\mathbf{h}(k) = \mathbf{a}(l) + \mathbf{b}(l)\alpha(k)$, where l is the index of the interval to which $\hat{\theta}_s(k)$ belongs. Note that the larger L is, the smaller the deviation of the cost function value will be, at any frequency $\hat{\theta}_s(k)$ in an interval l, from the optimal value for $\theta_0(l)$. Increasing L will not complicate the design of the filter. However, with a larger L, the variance of the frequency estimate will decrease at the increased cost of the time taken to search for the interval l to which $\hat{\theta}_s(k)$ belongs. Table 1 shows the variation of the variance of the frequency estimation, where the frequency to be estimated is located at the center of each interval. It can be observed that for large intervals (small L), there is a large variability in the variance as the
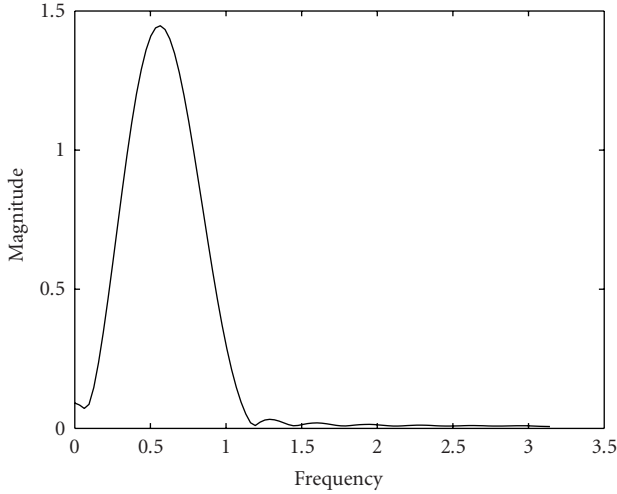
FIGURE 1: Magnitude response of optimized notch filter at $\theta_0 = \pi/4$.

TABLE 1: Effect of the number of intervals L on the variation of estimated frequency variance in an interval.

L	$\theta_s(\times 180/\pi)$	$\sigma_{\hat{\theta}_s}^2 (\times 10^{-5})$
	36	6.9772
10	45	5.7712
	54	21.644
	40	10.763
18	45	5.3937
	50	14.722
	42	4.4113
30	45	5.2202
	48	11.385

frequency to be estimated approaches the end points of each interval. For $L = 10$, at the center of the interval ($\theta_s = 45°$), $\sigma_{\hat{\theta}_s}^2 = 5.7712 \times 10^{-5}$. At the end of the corresponding interval ($\theta_s = 54°$), it reaches a value of $\sigma_{\hat{\theta}_s}^2 = 21.644 \times 10^{-5}$. However, for $L = 30$, at the center $\sigma_{\hat{\theta}_s}^2 = 5.2202 \times 10^{-5}$ and at the end ($\theta_s = 48°$), the variance reaches to $\sigma_{\hat{\theta}_s}^2 = 11.385 \times 10^{-5}$.

Further, it should be noted that increasing the filter length N improves the performance of the proposed algorithm. The larger N is, the sharper the notch in the frequency response at the signal frequency. While increasing N, L should also be increased to minimize the variability in estimation variance.

The computational complexity of the proposed filter is comparable with that of the standard LMS (requires approximately $2N$ multiplications and $2N$ additions per sample), but is much lower than the transform-domain LMS. The algorithm requires approximately $3N$ multiplications and $3N$ additions per sample, where N is the filter length. The offline optimization is done once and requires a single application of the FFT where the length is approximately $2N$. The IIR-ANF has a low complexity of approximately 10 multiplications and 10 additions.

A summary of the algorithm is given below.

Offline

(1) Select $\theta_0(l)$, $l = 1, \ldots, L$, uniformly distributed in $[0, \pi]$.
(2) For $l = 1, \ldots, L$, select weights $w_m(l)$, $m = 1, \ldots, M$ such that (7) is minimized for $\theta_0(l)$.
(3) For $l = 1, \ldots, L$, compute $\mathbf{h}_o$ using (13). Then, compute the vectors $(\mathbf{a}(l), \mathbf{b}(l))$ using the following procedure:

  (a) find the zeros of the polynomial $H_{e,o}(z)$ in (17);
  (b) using (18) find the coefficients of $H_n(z)$ in (19) (symbolic computation is used);
  (c) find $h_n(\alpha)$, then $a_n(l)$ and $b_n(l)$.

Online

(1) Given the frequency estimate $\hat{\theta}_s(k)$ at time step k, find l such that

$$\theta_o(l) - \frac{\pi}{2L} < \hat{\theta}_s(k) \leq \theta_o(l) + \frac{\pi}{2L}. \tag{32}$$

(2) Compute $\mathbf{h}[\hat{\theta}_s(k)]$ using (28) ($k + 1$ replaced by k).
(3) Compute the signal prediction $\hat{x}(k + 1)$ using (14).
(4) Compute the error using (16).
(5) Update the frequency estimate using (27) if (29) is not satisfied. Otherwise, $\hat{\theta}_s(k + 1) = \hat{\theta}_s(k)$.

4. STABILITY ANALYSIS

The updating equation for the estimated frequency in (27) is a nonlinear stochastic discrete-time equation. An exact stability analysis is only possible by using Lyapunov's direct method which is analytically intractable for this system. Therefore, an approximate stability analysis is performed when $\hat{\theta}_s$ is assumed to be close to the original signal frequency θ_s. The perturbation in $\hat{\theta}_s$ is defined as

$$\delta\hat{\theta}_s(k) = \hat{\theta}_s(k) - \theta_s. \tag{33}$$

The corresponding perturbation in the parameter α is

$$\delta\alpha(k) = \alpha(k) - \alpha_s \cong -2\sin(\theta_s)\delta\hat{\theta}_s(k). \tag{34}$$

In order to simplify the analysis, it will be assumed that there is no filtering on the prediction error. The prediction error in (16) can be expressed as

$$e_p(k + 1) = x_o(k + 1) + q(k + 1) - \hat{x}(k + 1). \tag{35}$$

Now, (28) can be used to write

$$\mathbf{h}(k) = \mathbf{h}[\alpha(k)] = \mathbf{a} + \mathbf{b}\alpha(k)$$
$$= \mathbf{h}(\alpha_s) + \mathbf{b}\delta\alpha(k). \tag{36}$$

The predicted signal in (35) can be written as

$$\hat{x}(k+1) = \mathbf{h}^T(k)\mathbf{x}_N(k) = \mathbf{h}^T(\alpha_s)\mathbf{x}_N(k) + \mathbf{b}^T\mathbf{x}_N(k)\delta\alpha. \tag{37}$$

The input vector in (37) can be written as

$$\mathbf{x}_N(k) = \mathbf{x}_{o,N}(k) + \mathbf{q}(k), \tag{38}$$

where $\mathbf{q}(k) = [q(k) \; q(k-1) \; \cdots \; q(k-N+1)]^T$, leading to

$$\mathbf{h}^T(\alpha_s)\mathbf{x}_N(k) = \mathbf{h}^T(\alpha_s)\mathbf{x}_{o,N}(k) + \mathbf{h}^T(\alpha_s)\mathbf{q}(k)$$
$$= x_o(k+1) + \hat{q}(k+1). \tag{39}$$

Substituting (39) and (37) in (35),

$$e_p(k+1) = q(k+1) - \hat{q}(k+1) - \mathbf{b}^T\mathbf{x}_N(k)\delta\alpha. \tag{40}$$

Substituting (33), (34), and (40) in (27), we obtain

$$\delta\hat{\theta}_s(k+1) = \left[1 - \mu(k)\frac{\sin(\theta_s)}{\sin(\hat{\theta}_s(k))}\right]\delta\hat{\theta}_s(k)$$
$$- \mu(k)\frac{q(k+1) - \hat{q}(k+1)}{[2\sin(\hat{\theta}_s(k))](\mathbf{b}^T\mathbf{x}_N(k))}. \tag{41}$$

Equation (41) is a nonlinear stochastic equation in discrete-time. Linearization of this equation around $\delta\hat{\theta}_s(k) = 0$ gives

$$\delta\hat{\theta}_s(k+1) = (1 - \mu(k))\delta\hat{\theta}_s(k) - \mu(k)\frac{q_d(k+1)}{2\sin(\theta_s)d_1(k)}. \tag{42}$$

In (42), $q_d(k+1) = q(k+1) - \hat{q}(k+1)$ and $d_1(k) = \mathbf{b}^T\mathbf{x}_N(k)$.

Taking the expectation of (42), the time-dependent part of the second term becomes

$$E\left\{\mu(k)\frac{q_d(k+1)}{d_1(k)}\right\} = \mu p_t E\left\{\frac{q_d(k+1)}{d_1(k)} \mid |d_1(k)| > \varepsilon\right\}, \tag{43}$$

where $p_t = P\{|d_1(k)| > \varepsilon\}$. In Appendix A, it is shown that this term is negligible. Taking expectation, (42) becomes

$$E\{\delta\hat{\theta}_s(k+1)\} = (1 - \bar{\mu})E\{\delta\hat{\theta}_s(k)\}, \tag{44}$$

where $\bar{\mu} = \mu \cdot P\{|d_1(k)| > \varepsilon\}$. The first-order discrete-time equation in (44) is stable if $0 < \bar{\mu} < 2$, in which case $\lim_{k\to\infty} E\{\delta\hat{\theta}_s(k)\} = 0$, implying that the frequency estimates are unbiased. Using (42), it is also possible to show that the frequency estimate converges to its true value in the mean, square sense, and the variance of the frequency estimate, which is obtained for white noise, is given as

$$\sigma_{\hat{\theta}_s}^2 \cong \frac{\mu^2 p_t \sigma_n^2(1+G_n)}{2\varepsilon A\theta_s \sin^2(\theta_s)(1-\lambda^P)}\sqrt{1 - \frac{\varepsilon^2}{A^2}}, \tag{45}$$

where $G_n = \|\mathbf{h}(\theta_s)\|^2$, $\lambda = (1 - 2\bar{\mu} + \mu^2 p_t)$, $P = \pi/\theta_s$, and A is the amplitude of $\mathbf{b}^T\mathbf{x}_N$. The derivation of (45) is outlined in Appendix B.
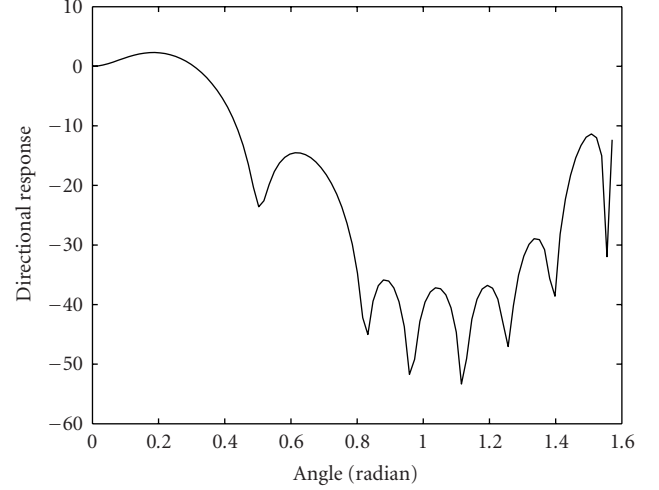

FIGURE 2: Directional response of the beamformer ($\theta_0 = 0°$).

5. APPLICATION TO BEAMFORMING

The proposed method is well suited to be applied in beamforming applications with angle-of-arrival estimation. Consider an array of N sensors with real gains and an incident signal $x_0(k)$. The output of the nth sensor is then

$$x_n(k) = x_0(k)e^{-jn\theta}, \tag{46}$$

where θ is the angle of arrival (AOA) of the signal. The beamformer output can be written as

$$y(k) = x_0(k)\sum_{n=0}^{N-1} h_n e^{-jn\theta}. \tag{47}$$

The directional response of the beamformer is

$$H(\theta) = \sum_{n=0}^{N-1} h_n e^{-jn\theta}. \tag{48}$$

Following the procedure given in Section 2, the gain of the beamformer at a selected angle of arrival can be made unity, while the weights can be chosen to shape the response. Figure 2 shows the response where interference arising in the directions from 0.8 to 1.4 radian is suppressed by approximately -40 dB. Note that since the gains of the beamformer are real, the response is symmetrical about $0°$.

The output of the beamformer can be written in general as

$$y(k) = \sum_{n=0}^{N-1} h_n(k)x_n(k) = \mathbf{h}^T(k) \cdot \mathbf{x_N}(k), \tag{49}$$

where

$$\mathbf{x_N}(k) = \begin{bmatrix} x_0(k) & x_1(k) & \cdots & x_{N-1}(k) \end{bmatrix}^T$$
$$= \mathbf{x}_N^{(s)}(k) + \mathbf{x}_N^{(i)}(k),$$
$$\mathbf{x}_N^{(s)}(k) = x_0^{(s)}(k) \cdot \begin{bmatrix} 1 & e^{-j\theta_0} & \cdots & e^{-j(N-1)\theta_0} \end{bmatrix}^T, \tag{50}$$
$$\mathbf{x}_N^{(i)}(k) = x_0^{(i)}(k) \cdot \begin{bmatrix} 1 & e^{-j\theta_i} & \cdots & e^{-j(N-1)\theta_i} \end{bmatrix}^T.$$

TABLE 2: Bias of the frequency estimates.

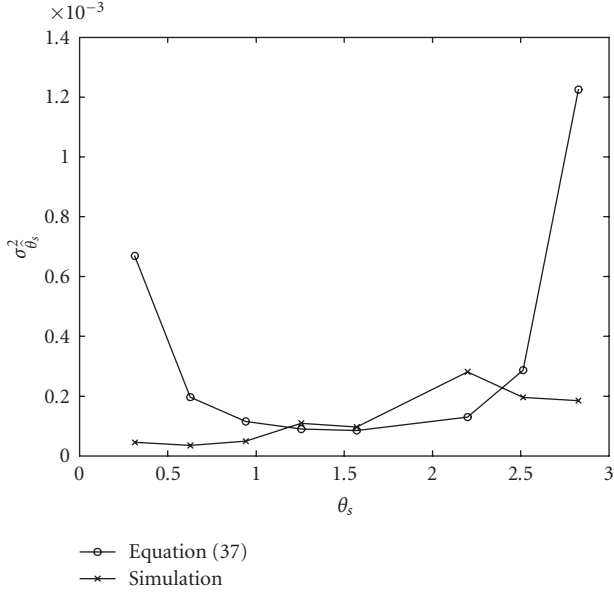| θ_s	0.3142	0.6283	0.9425	1.5708	2.5133	2.8274
$\hat{\theta}_s$ (AWGN)	0.3126	0.6290	0.9420	1.5591	2.5172	2.8278
$\hat{\theta}_s$ (CGN)	0.3146	0.6287	0.9431	1.5677	2.5160	2.8280

FIGURE 3: Approximate theoretical and computed variance of frequency estimates for AWGN.

In (50), θ_0 and θ_i are the angles-of-arrival of the main signal and the interference, respectively. In (49) it is assumed that the AOA of the main signal is not known and the gains $h_n(k)$ are adapted to estimate this angle. For this adaptation, the error signal is

$$e(k) = y(k) - \mathbf{h}^T(k) \cdot \mathbf{x_N}(k) \qquad (51)$$

and is complex in general. Therefore, the update equation for the AOA estimate should be written as

$$\hat{\theta}_s(k+1) = \hat{\theta}_s(k) - \mu \frac{\mathrm{Re}\,\{e(k)\}}{[2\sin(\hat{\theta}_s(k))]\,\mathrm{Re}\,\{\mathbf{b}^T\mathbf{x}_N(k)\}}. \qquad (52)$$

6.   SIMULATION RESULTS

For frequency estimation of a noisy sinusoid, the parameters used in the simulations are $N = 16$, $a = 1$, $L = 18$, $\mu = 0.01$. Table 2 shows the estimates of selected frequencies in the range $[0, \pi]$ in AWGN ($\sigma_q^2 = 0.25$) and in correlated Gaussian noise (CGN) ($\sigma_q^2 = 0.30$), averaged over 30 000 samples. The estimates are generally unbiased with a maximum absolute error of 0.7%.

Figure 3 shows the theoretical and computed variance of estimated frequency in AWGN. There is good agreement
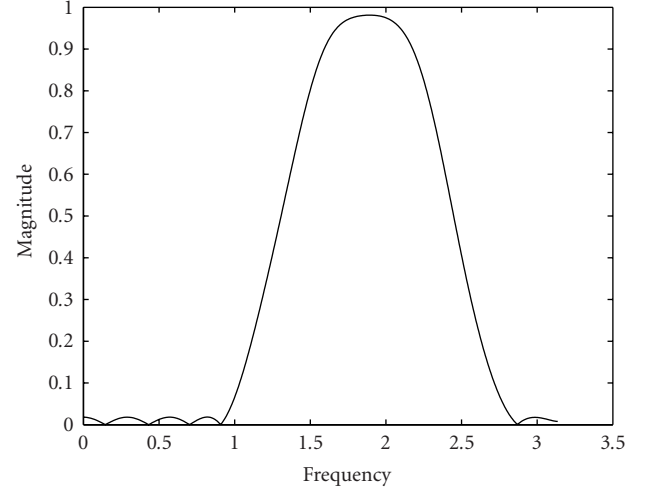

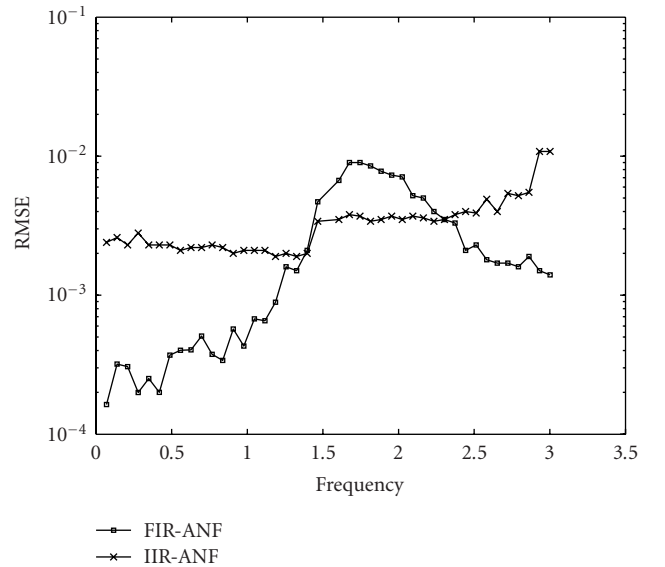FIGURE 4: Magnitude response of the noise filter.

FIGURE 5: Computed RMSE of the FIR-ANF and the IIR-ANF.

except at the edges of the frequency range. This is due to the terms in the denominator in (45) which become zero at the edges. The RMSE performance of the FIR notch filter (FIR-ANF) is compared with that of the IIR notch filter (IIR-ANF, constrained poles and zeros [6]). The noise is correlated Gaussian with variance $\sigma_q^2 = 0.30$ and is obtained by filtering white noise using a filter having the magnitude response shown in Figure 4. In order that a fair comparison is made, the stepsize of IIR-ANF is adjusted to have the same convergence rate as the FIR-ANF. Alternatively, the RMSE for the two methods could have been fixed to observe the improvement in the convergence rate. Figure 5 shows the computed RMSE values over the complete frequency range. It is observed that the RMSE of the FIR-ANF is less than that of
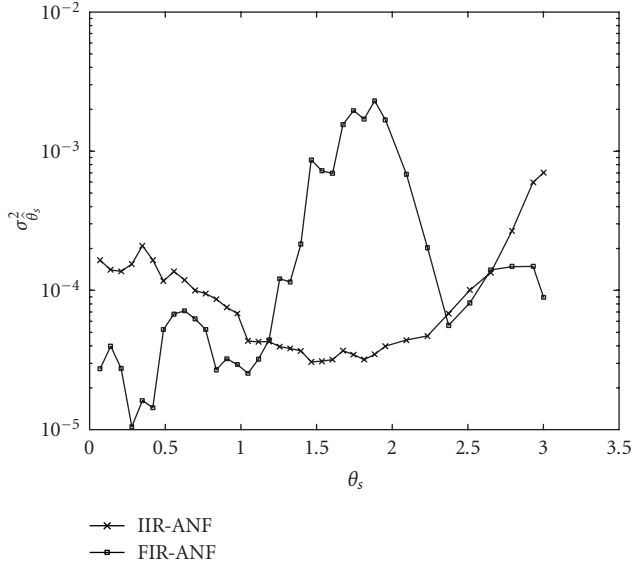
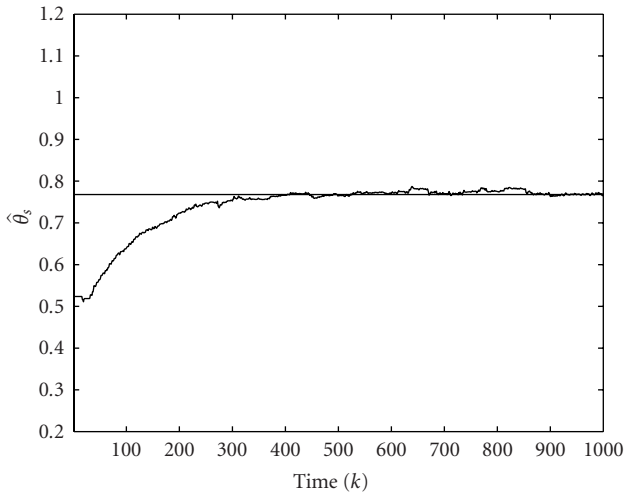FIGURE 6: Computed estimated frequency variances of FIR-ANF and IIR-ANF.

FIGURE 8: Convergence of the estimated frequency from an initial value of 1.92 rad. to 2.80 rad. (FIR-ANF).

FIGURE 7: Convergence of the estimated frequency from an initial value of $\pi/6$ to $\pi/4$ (FIR-ANF).

exponential, that is, $(1 - \bar{\mu})^k$, which is the solution of (44). It is also important to note that the initial frequency estimate does not have to be in very close vicinity of the true one. Figure 8 shows the convergence of the estimated frequency when the initial value is at the center of the correlated noise band (1.92 rad.), which poses the biggest challenge for the algorithm (it still converges to the true frequency which is 2.80 rad.). However, as the initial frequency is far from the steady state value for this case, the linear model in (42) is not valid any more and the response of the error is not exponential. It should be noted here that, whatever parameters are chosen, IIR-ANF does not converge under the same conditions. Figure 9 shows the responses of the frequency estimates with FIR and IIR-ANFs for the case where the frequency variances are equated.

A beamforming application is also simulated where the actual angle-of–arrival of the signal is $4°$ and the initial estimate is $0°$. Figure 10 shows the convergence of the estimated AOA to the true one. The frequencies of the signal and the interference are 0.087 radian and 0.87 radian, respectively. Sensor inputs are assumed to be corrupted by zero-mean Gaussian noise with uniform directional density. The signal-to-interference ratio (SINR) of the beamformer input is SINR = 4.15. After convergence, the signal-to-interference ratio is calculated as SINR = 84.9, with an increase by a factor of 20.45.

7. CONCLUSIONS

A new adaptive notch FIR filter is introduced. This filter has the novel feature that its frequency responses can be optimized in an offline manner. The proposed filter is considerably more flexible in shaping the frequency response, and thereby rejecting noise in selected frequency ranges. Unlike the IIR filter, the adaptive FIR filter is always stable for

IIR-ANF over almost the entire frequency range, except in $1.5 < \theta_s < 2.3$ rad. (around the center of the noise band).

Figure 6 shows the variances of the estimated signal frequency for the two methods. It is again observed that the variance of FIR-ANF is better than that of IIR-ANF over the frequency range, except in $1.2 < \theta_s < 2.4$ rad. As the signal frequency approaches the noise band, the variance of FIR-ANF rapidly increases above that of the IIR-ANF. This is due to the inefficiency of the FIR filter to suppress noise components which are very near the signal frequency. Figure 7 shows the convergence of the estimated frequency from an initial value of $\theta_s(0) = \pi/6$ to the actual value $\theta_s = \pi/4$ with the FIR-ANF. It is observed that the convergence is almost
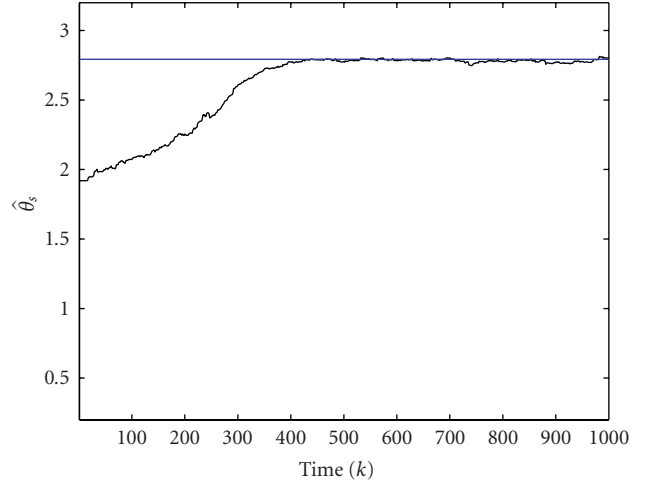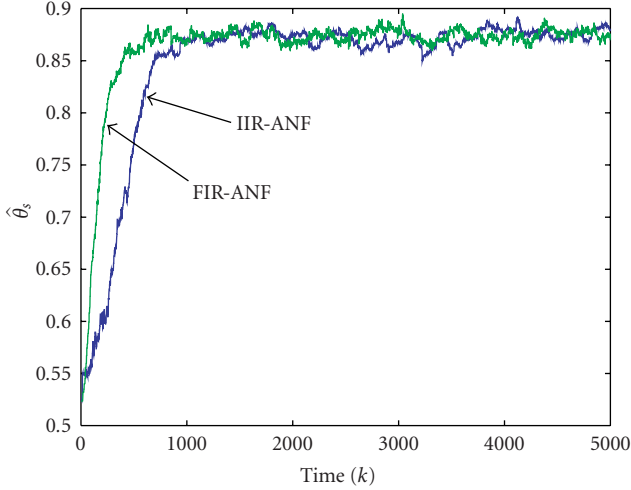
FIGURE 9: Convergence of the estimated frequency from an initial value of $\hat{\theta}_s(0) = 0.52$ rad. to $\theta_s = 0.87$ rad. (FIR-ANF and IIR-ANF).
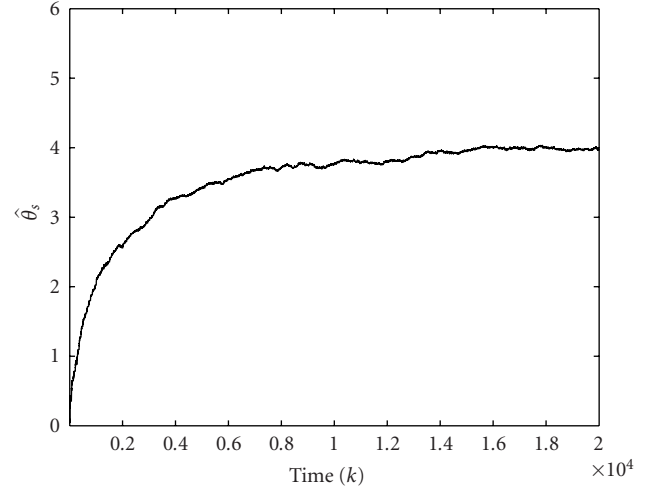

FIGURE 10: Convergence of the estimated AOA from an initial value of $0°$ to $4°(\mu = 0.0003, \varepsilon = 6\sin\hat{\theta}_s)$.

suitable choice of stepsizes. The algorithm can effectively be applied to beamforming problems with AOA estimation, whereas the IIR counterpart is inapplicable. Simulation results indicate that except for the frequency range around the peak noise power, the FIR-ANF is superior in estimating the sinusoid and its frequency in CGN, compared with the IIR notch filter.

APPENDICES

A.   EXPECTED VALUE OF THE SECOND TERM IN (41)

In [11], an approximation for the expected value of a function of two random variables is given as

$$E\{g(X, Y)\} \cong g_0 + \frac{1}{2}\left(\frac{\partial^2 g}{\partial x^2}\sigma_X^2 + \frac{\partial^2 g}{\partial y^2}\sigma_Y^2\right) + \frac{\partial^2 g}{\partial x \partial y}\rho\sigma_X\sigma_Y,$$
(A.1)

where $g_0 = g(\mu_X, \mu_Y)$ and the derivatives are evaluated at (μ_X, μ_Y). For $g(x, y) = x/y$ (A.1), gives

$$E\{g(X, Y)\} \cong \frac{\mu_X}{\mu_Y} - \frac{\rho\sigma_X\sigma_Y}{\mu_Y^2} + \frac{\mu_X\sigma_Y^2}{\mu_Y^3},$$
(A.2)

which can be applied to the expected value in (43). Letting $X = q_d(k + 1)$ and defining Y as a random variable taking values which satisfy the threshold. The fact that $\mu_X = E\{q_d(k + 1)\} = 0$ gives

$$\eta = E\left\{\frac{q_d(k + 1)}{d_1(k)} \mid |d_1(k)| > \varepsilon\right\} \cong -\frac{\rho\sigma_X\sigma_Y}{\mu_Y^2},$$
(A.3)

where $\mu_Y = E\{d_1(k) \mid |d_1(k)| > \varepsilon\}$.

Similar to (39), $d_1(k)$ can be written as

$$d_1(k) = \mathbf{b}^T\mathbf{x}_N(k) = \mathbf{b}^T\mathbf{x}_{o,N}(k) + \mathbf{b}^T\mathbf{q}(k)$$
$$= s(k) + q_2(k),$$
(A.4)

where $s(k)$ is a sinusoid. Let $s(k) = A\sin((k + 1)\theta_s)$, where A is determined by the vector $\mathbf{b}$. Therefore, $\mu_Y \cong s(k)$ whenever $k \in I_\in = \{k : |d_1(k)| > \varepsilon\}$ and

$$\eta \cong -\frac{\sigma_{XY}}{A^2\sin^2((k + 1)\theta_s)} \quad k \in I_\in.$$
(A.5)

The time-average of the expected value is written as

$$\bar{\eta} \cong -\frac{\rho\sigma_X\sigma_Y}{A^2}\frac{\theta_s}{\pi}\sum_{k \in I_\in}\frac{1}{\sin^2((k + 1)\theta_s)}$$
$$\cong -\frac{\sigma_{XY}\theta_s}{A^2\pi}\frac{1}{\theta_s}\int_{\sin^{-1}\varepsilon/A}^{\pi - \sin^{-1}\varepsilon/A}\frac{1}{\sin^2\theta}d\theta$$
$$= -\frac{2\sigma_{XY}}{A\pi\varepsilon}\sqrt{1 - \frac{\varepsilon^2}{A^2}} = -\frac{2\sigma_{XY}}{A^2\pi\tau}\sqrt{1 - \tau^2},$$
(A.6)

where $\tau = \varepsilon/A$ (is of the order of 0.1). The amplitude A can be estimated if the filter coefficients are approximated by those of an LMS adaptive line enhancer [12]:

$$h_n = \frac{2a^*}{N}\cos((n + 1)\theta_s), \quad n = 0, \ldots, N - 1,$$
(A.7)

where a^* is approximately equal to one for large N. The derivatives with respect to α are obtained as

$$b_n = \frac{\delta h_n}{\delta\alpha} = \frac{a^*(n + 1)}{N\sin\theta_s}\sin((n + 1)\theta_s), \quad n = 0, \ldots, N - 1.$$
(A.8)

Evaluation of $\mathbf{b}^T\mathbf{x}_{o,N}(k)$ shows that it contains three sinusoidal functions of the time k, with amplitudes which are of the orders of $1/N$, 1, and N. Hence, the amplitude A can be approximately obtained as

$$A = \frac{a^*(N + 1)}{4\sin\theta_s}a.$$
(A.9)

The covariance σ_{XY} can be obtained as

$$\sigma_{XY} = E\{q_d(k+1) \cdot d_1(k)\} = \mathbf{b}^T(\mathbf{r}_q - \mathbf{Q}\mathbf{h}(\alpha_s)), \quad \text{(A.10)}$$

where $\mathbf{Q}$ is the autocorrelation matrix of the noise sequence and $\mathbf{r}_q = E\{q(k+1) \cdot \mathbf{q}(k)\}$. It is difficult to derive a general result from (A.10) for any given correlated noise sequence. To gain an insight as to the order of this term, we consider white noise. In this case, $\mathbf{r}_q = \mathbf{0}$ and $\mathbf{Q} = \text{diag}[\sigma_q^2, \ldots, \sigma_q^2]$ resulting in

$$\sigma_{XY} = \sigma_q^2 \mathbf{b}^T \mathbf{h}(\alpha_s), \quad \text{(A.11)}$$

which can be calculated using (A.7) and (A.8) as

$$\sigma_{XY} = \frac{\sigma_q^2 (\mathrm{a}^*)^2}{4N^2 \sin^2 \theta_s}$$
$$\times \{(2N + \csc^2 \theta_s) \sin(2N\theta_s) - 2N \cot \theta_s \cos(2N\theta_s)\}. \quad \text{(A.12)}$$

Combining all the above results, it can be easily shown that the second term in (43) (except for μ) is of the order of

$$\frac{8\sigma_q^2}{\pi \tau N^3 a^2 \sin^2 \theta_s}. \quad \text{(A.13)}$$

Note that the above expression is also approximately equal to the bias in the frequency estimate. Hence, for $0 < \theta_s < \pi$ and for sufficiently large N the bias is negligible.

B. VARIANCE OF THE FREQUENCY ESTIMATE

Variance of the frequency estimate is given by

$$v(k) = E\left([\hat{\theta}_s(k) - \theta_s]^2\right) = E\left([\delta\hat{\theta}_s(k)]^2\right). \quad \text{(B.1)}$$

If the simplifying assumption is made that $\delta\hat{\theta}_s(k)$, $q(k+1)$, and $\hat{q}(k+1)$ are uncorrelated, the following can be obtained from (42):

$$v(k+1) = (1 - 2\bar{\mu} + \mu^2 p_t)v(k) + \frac{\mu^2 p_t \sigma_n^2(1+G_n)}{4\sin^2 \theta_s}D(k), \quad \text{(B.2)}$$

where $D(k) = E\{1/d_1^2(k) \mid |d_1(k)| > \varepsilon\}$.

The steady state solution of the difference equation in (B.2) can be written as

$$\lim_{k\to\infty} v(k) = \lim_{k\to\infty} \frac{\mu^2 p_t \sigma_n^2(1+G_n)}{4A^2 \sin^2 \theta_s} \sum_{\substack{i=1 \\ i\in I_\in}}^{k} \frac{\lambda^{k-i}}{\sin^2[(i+1)\theta_s]}, \quad \text{(B.3)}$$

where $\lambda = (1 - 2\bar{\mu} + \mu^2 p_t)$.

The summation term in (B.3) may be approximately evaluated as

$$\sum_{\substack{i=1 \\ i\in I_\in}}^{k} \frac{\lambda^{k-i}}{\sin^2[(i+1)\theta_s]} \simeq \left(\frac{1-\lambda^k}{1-\lambda^P}\right)\frac{2A}{\varepsilon\theta_s}\sqrt{1 - \frac{\varepsilon^2}{A^2}}, \quad \text{(B.4)}$$

where $P = \pi/\theta_s$. In the limit, as k goes to infinity. (B.3) and (B.4) lead to (43).

REFERENCES

[1] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1993.

[2] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, Englewood Cliffs, NJ, USA, 2002.

[3] F. Beaufays, "Transform-domain adaptive filters: an analytical approach," *IEEE Transactions on Signal Processing*, vol. 43, no. 2, pp. 422–431, 1995.

[4] L. S. Resende, J. M. T. Romano, and M. G. Bellanger, "Split wiener filtering with application in adaptive systems," *IEEE Transactions on Signal Processing*, vol. 52, no. 3, pp. 636–644, 2004.

[5] A. Nehorai, "A minimal parameter adaptive notch filter with constrained poles and zeros," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 4, pp. 983–996, 1985.

[6] P. Stoica and A. Nehorai, "Performance analysis of an adaptive notch filter with constrained poles and zeros," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 6, pp. 911–919, 1988.

[7] G. Li, "A stable and efficient adaptive notch filter for direct frequency estimation," *IEEE Transactions on Signal Processing*, vol. 45, no. 8, pp. 2001–2009, 1997.

[8] A. Mvuma, S. Nishimura, and T. Hinamoto, "Adaptive IIR notch filter with controlled bandwidth for narrow-band interference suppression in DS CDMA system," in *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS '03)*, vol. 4, pp. IV-361–IV-364, Bangkok, Thailand, May 2003.

[9] A. Hocanin and O. Kukrer, "Estimation of the frequency and waveform of a single-tone sinusoid using an offline-optimized adaptive filter," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05)*, vol. 4, pp. 349–352, Philadelphia, Pa, USA, March 2005.

[10] B. Rafaely and S. J. Elliot, "A computationally efficient frequency-domain LMS algorithm with constraints on the adaptive filter," *IEEE Transactions on Signal Processing*, vol. 48, no. 6, pp. 1649–1655, 2000.

[11] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, NewYork, NY, USA, 1991.

[12] J. T. Rickard and J. R. Zeidler, "Second-order output statistics of the adaptive line enhancer," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 27, no. 1, pp. 31–39, 1979.

Osman Kukrer was born in 1956 in Larnaca, Cyprus. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the Middle East Technical University (METU), Ankara, Turkey, in 1979, 1982, and 1987, respectively. From 1979 to 1985, he was a Research Assistant in the Department of Electrical and Electronics Engineering, METU. From 1985 to 1986 he was with the department of Electrical and Electronics Engineering, Brunel University, London, UK. He is currently a Professor in the Department of Electrical and Electronic Engineering, Eastern Mediterranean University, Gazimagusa, North Cyprus. His research interests include power electronics, control systems, and signal processing.

Aykut Hocanin was born in Paphos, Cyprus, in 1970. He received the B.S. degree in electrical and computer engineering from Rice University, Houston, Texas, USA, in 1992, and the M.E. degree from Texas A&M University, College Station, Texas, USA, in 1993. He received the Ph.D. degree in electrical and electronics engineering from Boğaziçi University, Istanbul, Turkey, in 2000. He joined the faculty of Eastern Mediterranean University, Gazimağusa, North Cyprus, in 2000, where he is currently an Assistant Professor and Vice Chair in the Department of Electrical and Electronics Engineering. His current research interests include receiver design for wireless systems, multiuser techniques for CDMA, detection and estimation theory.