A Complex Adaptive Notch Filter

Article *in* IEEE Signal Processing Letters · December 2010

DOI: 10.1109/LSP.2010.2075925 · Source: IEEE Xplore

P.A. Regalia
The Catholic University of America

A Complex Adaptive Notch Filter

Phillip A. Regalia, *Fellow, IEEE*

***Abstract*—A complex adaptive notch filter is developed, for tracking single-sided (a.k.a. analytic or complex) tones immersed in background noise. A complex all-pass based realization is pursued which inherits useful properties from its real counterpart: independent tuning of the notch frequency and attenuation bandwidth, easy realization of the complementary band-pass filter, unbiased frequency estimation, and faster convergence and tracking than a gradient descent algorithm.**

***Index Terms*—Adaptive filters, chirp, frequency estimation.**

I. Introduction

Adaptive notch filtering aims to estimate unknown frequencies of periodic components buried in noise, and/or retrieve such periodic components, finding thus wide application in communication systems and in periodic noise suppression. Many designs have been advanced over the years [1]–[16], exploiting constrained poles and zeros either through direct coefficient scaling (e.g., [1], [2], [4], [12], [13]) or through all-pass decompositions (e.g., [3], [5]–[7], [9], [10], [15]); the latter feature independent tuning of the notch frequency and attenuation bandwidth, and offer ease of realizing the complementary band-pass filter to retrieve the periodic component.

While most the these designs focus on real coefficient filters, complex notch filters also find applications, as cogently surveyed in [8], particularly in communication systems using quadrature modulation. The intent of this note is to develop an all-pass based complex notch filter—as contrasted to the coefficient scaling approach of [8]—for tracking single sided (or analytic) periodic components. The proposed design mimics an earlier real-coefficient prototype [7] and thus inherits its attractive features: independent tuning of the notch frequency and attenuation bandwidth, unbiased frequency estimation, and a wider basin of attraction than a gradient-based algorithm.

Section II develops the complex notch filter, while Section III develops the adaptation algorithm and its convergence properties. Simulation results confirming the design are presented in Section IV, while comparisons with a gradient descent approach are developed in Section V, to establish the superior basin of attraction of the proposed design. Concluding remarks are synthesized in Section VI.

Manuscript received July 12, 2010; revised August 30, 2010; accepted September 03, 2010. Date of publication September 13, 2010; date of current version September 23, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jen-Tzung Chien.

The author is with the Department of Electrical Engineering and Computer Science, Catholic University of America, Washington, DC 20064 USA (e-mail: regalia@cua.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2010.2075925

II. Filter Realization

Consider a first-order all-pass transfer function $A(z)$:

$$A(z) = \frac{z^{-1} - \alpha}{1 - \alpha z^{-1}},$$

where $0 \leq \alpha < 1$ is a real coefficient. The phase shift of $A(e^{j\omega})$ is zero at $\omega = 0$, reaching $-\pi$ radians at $\omega = \pi$. If the output of this filter is subtracted from its input, a zero (or notch) is obtained at $\omega = 0$. By applying a frequency translation, the notch frequency at $\omega = 0$ can be moved to an arbitrary position. In particular, consider the transformation $z \to e^{-j\theta} z$ applied to the all-pass function $A(z)$, which generates a first-order complex all-pass transfer function denoted $C(z)$:

$$\begin{aligned} C(z) &= A(z)|_{z \to e^{-j\theta} z} \\ &= \frac{e^{j\theta} z^{-1} - \alpha}{1 - \alpha e^{j\theta} z^{-1}}. \end{aligned}$$

Two complementary transfer functions may then be defined through

$$\begin{aligned} G(z) &= \tfrac{1}{2}[1 - C(z)]; \\ H(z) &= \tfrac{1}{2}[1 + C(z)]. \end{aligned}$$

These are readily shown (e.g., [17]) to satisfy the relations

$$\begin{aligned} G(e^{j\omega}) + H(e^{j\omega}) &= 1 \\ |G(e^{j\omega})|^2 + |H(e^{j\omega})|^2 &= 1, \qquad \text{for all } \omega. \end{aligned}$$

The transfer function $G(z)$ now has a zero (or notch) at $\omega = \theta$, with a 3 dB bandwidth Ω between the frequencies $\omega_\pm$ for which $C(e^{j\omega_\pm}) = \pm j$, leading to the formula

$$\Omega = \frac{\pi}{2} - 2\tan^{-1}(\alpha).$$

A flow graph of the complex notch filter realization appears in Fig. 1; the input sequence is designated $u(n)$, while the output is $e(n)$. The state-space description of the system is

$$\begin{aligned} x(n{+}1) &= e^{j\theta(n)} \alpha x(n) + e^{j\theta(n)} \sqrt{1 - \alpha^2}\, u(n) \\ e(n) &= -\frac{\sqrt{1 - \alpha^2}}{2} x(n) + \frac{1 + \alpha}{2} u(n). \end{aligned}$$

The complementary band-pass filter is provided through $H(z)$, obtained by adding (rather than subtracting) the all-pass filter output and input, and offering recovery of a complex tone.

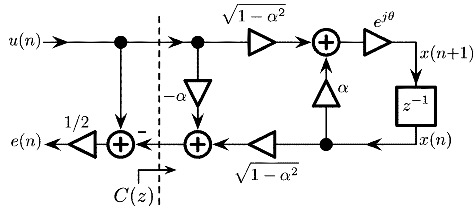

Fig. 1. Flow graph of complex notch filter.

III. Adaptation Algorithm

Suppose the input is a complex single-sided tone plus background noise:

$$u(n) = Ae^{j(\omega_0 n+\phi)} + b(n).$$

Here A is a scale factor,[1] ω_0 is the unknown frequency, ϕ is a random phase uniformly distributed over $[0, 2\pi)$, and $b(n)$ is a white complex circular Gaussian noise process, of variance

$$\sigma^2 = E|b(n)|^2 = E\{(\mathrm{Re}[b(n)])^2\} + E\{(\mathrm{Im}[b(n)])^2\}.$$

Consider the following adaptation algorithm:

$$\theta(n{+}1) = \theta(n) + \mu\,\mathrm{Im}[e(n)x^*(n)]. \tag{1}$$

Here $\mu > 0$ is a small adaptation step size. For sufficiently slow adaptation, the evolution of the adaptation algorithm is weakly linked to an ordinary differential equation [18]–[20] of the form

$$\frac{d\theta}{dt} = E\{\mathrm{Im}[e(n)x^*(n)]|\theta\}$$

where, as notationally emphasized, the right-hand side expectation is to be evaluated for a fixed θ, with the expression so obtained interpreted as a function of θ which drives the differential $d\theta/dt$.

Let $G(z)$ and $F(z)$ be the transfer functions linking the input $u(n)$ to the notch output $e(n)$ and the filtered regressor $x(n)$, respectively; these are given as

$$G(z) = \frac{1+\alpha}{2}\,\frac{1-e^{j\theta}z^{-1}}{1-\alpha e^{j\theta}z^{-1}};$$
$$F(z) = \sqrt{1-\alpha^2}\,\frac{e^{j\theta}z^{-1}}{1-\alpha e^{j\theta}z^{-1}}.$$

The expectation $E\{e(n)\,x^*(n)\}$ may then be expressed as the inner product

$$E\{e(n)x^*(n)\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathcal{S}_u(\omega)G(e^{j\omega})F^*(e^{j\omega})d\omega$$

in which $\mathcal{S}_u(\omega)$ is the input power spectral density:

$$\mathcal{S}_u(\omega) = 2\pi A^2\delta(\omega-\omega_0) + \sigma^2.$$

This gives for the expectation

$$E\{e(n)x^*(n)\} = A^2G(e^{j\omega_0})F^*(e^{j\omega_0})$$
$$+\frac{\sigma^2}{2\pi}\int_{-\pi}^{\pi} G(e^{j\omega})F^*(e^{j\omega})d\omega$$

in which

$$G(e^{j\omega})F^*(e^{j\omega}) = \frac{(1+\alpha)\sqrt{1-\alpha^2}}{2}\,\frac{e^{j(\omega_0-\theta)}-1}{|1-\alpha e^{j(\omega_0-\theta)}|^2}.$$

The imaginary part thus becomes

$$\begin{aligned}&\mathrm{Im}(E\{e(n)x^*(n)\})\\ &= \frac{A^2}{2}\,\frac{(1+\alpha)\sqrt{1-\alpha^2}}{|1-\alpha e^{j(\omega_0-\theta)}|^2}\sin(\omega_0-\theta)\\ &\quad+\frac{\sigma^2(1+\alpha)\sqrt{1-\alpha^2}}{2\pi}\underbrace{\int_{-\pi}^{\pi}\frac{\sin(\omega-\theta)}{|1-\alpha e^{j(\omega-\theta)}|^2}d\omega}_{=0}\end{aligned} \tag{2}$$

in which the noise-induced term vanishes since it is the integral over one period of a function odd about $\omega = \theta$. The associated differential equation thus becomes

$$\frac{d\theta}{dt} = \frac{A^2}{2}\,\frac{(1+\alpha)\sqrt{1-\alpha^2}}{|1-\alpha e^{j(\omega_0-\theta)}|^2}\sin(\omega_0-\theta). \tag{3}$$

Convergence of θ to ω_0 in this differential equation is shown by choosing as a Lypunov function

$$L(t) = [\omega_0-\theta(t)]^2,$$

to obtain

$$\begin{aligned}\frac{dL(t)}{dt} &= \frac{dL}{d\theta}\,\frac{d\theta}{dt}\\ &= -\frac{A^2(1+\alpha)\sqrt{1-\alpha^2}}{|1-\alpha e^{j(\omega_0-\theta)}|^2}(\omega_0-\theta)\sin(\omega_0-\theta) < 0,\\ &\text{for } \theta \neq \omega_0\end{aligned}$$

assuming $\omega_0-\theta$ is restricted to the principal value range $-\pi \le \omega_0-\theta \le \pi$. This shows that $L(t)$ is monotonically decreasing, so that $\theta(t) \to \omega_0$, as desired. If $|\omega_0-\theta| > \pi$, then θ converges to $\omega_0+2\pi k$ for an appropriate integer k; since θ intervenes in the filter computations only through the factor $e^{j\theta(n)}$, a modulo-2π ambiguity in θ proves innocuous.

Obtaining step size bounds for the actual update algorithm (1) is more delicate, since it is at best a stochastic approximation to the differential equation (3). A rough though usable bound may be obtained, however, if we examine a simplified approximation. In particular, consider the mean evolution of (1), setting $\overline{\theta}(n) = E[\theta(n)]$ for notational simplicity:

$$\begin{aligned}\overline{\theta}(n{+}1) &= \overline{\theta}(n) + \mu E\{\mathrm{Im}[e(n)x^*(n)]\}\\ &= \overline{\theta}(n) + \mu\frac{A^2}{2}\,\frac{(1+\alpha)\sqrt{1-\alpha^2}}{|1-\alpha e^{j(\omega_0-\overline{\theta})}|^2}\sin(\omega_0-\overline{\theta}).\end{aligned} \tag{4}$$

Here we have replaced a right-hand side expectation with the expression from (2), obtained under the assumption that $e(n)$ and $x(n)$ had attained stationarity. This yields a crude approximation, since time variations in $\theta(n)$ will induce nonstationarity in

[1]We may take A real without loss of generality: If A were complex, its phase angle could be absorbed into ϕ, leaving behind a real-valued scale factor.

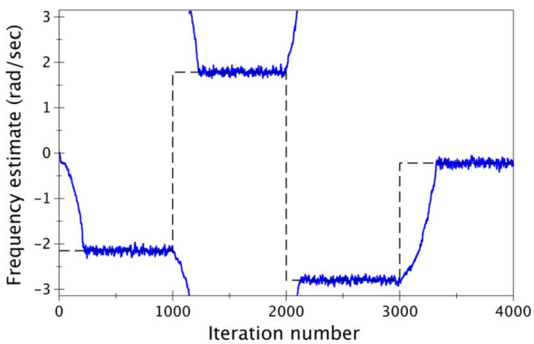

Fig. 2. Simulation using a frequency hop experiment.

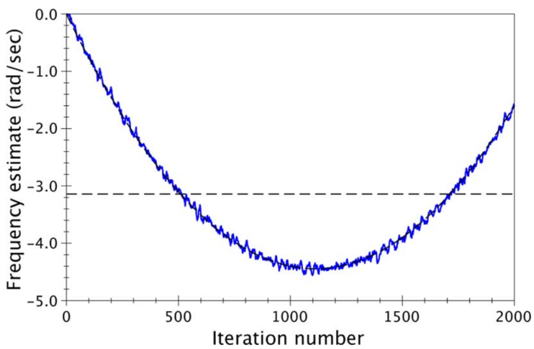

Fig. 3. Simulation using a quadratic chirp.

$e(n)$ and $x(n)$. For sufficiently small μ however, corresponding to slow adaptation, such an approximation may be tolerable.

Now, for $\overline{\theta}$ sufficiently close to ω_0, we may use the further approximation

$$\frac{\sin(\omega_0 - \overline{\theta})}{|1 - \alpha e^{j(\omega_0 - \overline{\theta})}|^2} \approx \frac{\omega_0 - \overline{\theta}}{(1-\alpha)^2}, \qquad \text{for small } |\omega_0 \text{ - } \overline{\theta}|$$

to simplify (4) to

$$\omega_0 - \overline{\theta}(n{+}1) = \left(1 - \mu\frac{A^2}{2}\left(\frac{1+\alpha}{1-\alpha}\right)^{3/2}\right)(\omega_0 - \overline{\theta}(n)).$$

This recursion will induce $|\omega_0 - \overline{\theta}(n{+}1)| < |\omega_0 - \overline{\theta}(n)|$ provided the step-size μ lies in the range

$$0 < \mu < \frac{1}{A^2}\left(\frac{1-\alpha}{1+\alpha}\right)^{3/2}. \tag{5}$$

When instead $\overline{\theta}$ is displaced from ω_0, we in fact have the strict inequality

$$\left|\frac{\sin(\omega_0 - \overline{\theta})}{|1 - \alpha e^{j(\omega_0 - \overline{\theta})}|^2}\right| < \left|\frac{\omega_0 - \overline{\theta}}{(1-\alpha)^2}\right|, \qquad \overline{\theta} \neq \omega_0$$

so that the upper bound from (5) still ensures monotonic convergence in mean, but becomes conservative. We note finally that in practice, this upper bound should be scaled back due to the actual algorithm (1) having filtered background noise and nonstationary components that this simplified mean analysis neglects. Further insights into the relevant advantages and drawbacks of "slow adaptation" analysis techniques may be found in [21].

IV. Simulation Results

Fig. 2 shows the frequency estimates $\theta(n)$ (reduced modulo-2π) obtained from a single run of the adaptive complex notch filter using the proposed algorithm (1) in a frequency hop experiment, using unit signal power $(A = 1)$ and a signal-to-noise ratio of 0 dB. The periodic component of the input signal abruptly changes its frequency ω_0 every 1000 iterations, giving a piecewise stationary signal. The bandwidth parameter was fixed to $\alpha = 0.7$, using a step size of $\mu = 0.02$ [roughly one-third the upper bound from (5)]. The adaptive filter is clearly able to distinguish positive from negative frequencies, as required in analytic signal processing.

Fig. 3 shows the result of tracking a quadratically varying frequency, in which $u(n) = e^{j\phi(n)} + b(n)$ using

$$\phi(n) = \phi_2 n^2 + \phi_3 n^3$$

with $\phi_2 = -0.004$ and $\phi_3 = 1.2 \times 10^{-6}$. The signal to noise ratio is again 0 dB, using a step size $\mu = 0.008$ and now a bandwidth parameter $\alpha = 0.9$, to demonstrate the tracking abilities of the proposed design. We remark that the input signal is nonstationary in this case, which significantly complicates attempts to bound the range of usable step size values. For this example, nonetheless, respectable tracking performance is observed using a step size value which here is about two-thirds the upper bound from (5) that was obtained under stationarity assumptions.

V. Comparison With Gradient Descent

The proposed adaptation algorithm is not a gradient descent procedure applied to the output power $E|e(n)|^2$. Such an approach is pursued in [8], albeit using a different parametrization which results in biased frequency estimates [8, eq. (17)] since, as with the real-coefficient case (e.g., [9, Ch.10]), the noise gain varies with the center frequency. The realization proposed in this note does not suffer this limitation, although we show that, for the same filter bandwidth, the proposed algorithm (1) has a superior basin of attraction.

To this end, we note that the bandwidth of the notch filter $G(z)$ does not change with the notch frequency (as controlled by θ). As such, the presence of white background noise contributes a constant (dependent on α) to the cost function $E|e(n)|^2$; the minimum of this cost function thus occurs when the notch frequency θ aligns with the signal frequency ω_0.

The mean update term of a gradient descent algorithm is proportional to $-\partial E|e(n)|^2/\partial\theta$, as given by

$$\begin{aligned} -\frac{\partial E|e(n)|^2}{\partial\theta} &= \beta\,\mathrm{Im}\left(\frac{e^{j(\theta-\omega_0)}}{|1-\alpha e^{j(\theta-\omega_0)}|^2}\,\frac{1-e^{j(\omega_0-\theta)}}{1-\alpha e^{j(\omega_0-\theta)}}\right) \\ &\triangleq f_{\mathrm{grad}}(\omega_0 - \theta) \end{aligned}$$

where the scale factor β may be absorbed into the step size of a gradient descent algorithm. The mean driving term $E\{\mathrm{Im}[e(n)\; x^*(n)]\}$ from (2) is also a function of $(\omega_0 - \theta)$; denoting this by $f_{\mathrm{prop}}(\omega_0 - \theta)$, a calculation will show that, subject to the normalization constraint $f'_{\mathrm{grad}}(0) = f'_{\mathrm{prop}}(0)$,

$$\frac{f_{\mathrm{grad}}(\omega_0-\theta)}{f_{\mathrm{prop}}(\omega_0-\theta)} = \frac{(1-\alpha)[1+\alpha-2\alpha\cos(\omega_0-\theta)]}{|1-\alpha e^{j(\omega_0-\theta)}|^2} \le 1 \tag{6}$$

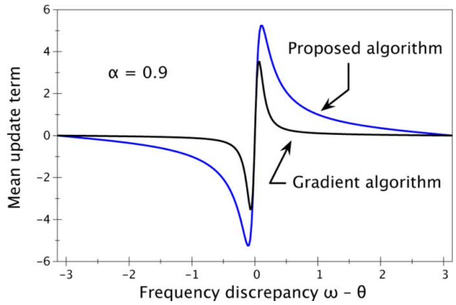

Fig. 4. Comparing the mean driving terms of the proposed algorithm and a gradient descent algorithm, when normalized for the same local convergence properties.

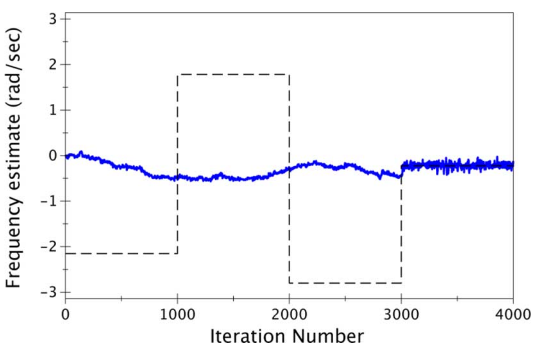

Fig. 5. Frequency hop experiment using a gradient descent algorithm instead of the proposed algorithm.

for all discrepancies $\omega_0 - \theta$, and for any $0 \leq \alpha < 1$. Thus the gradient descent algorithm will always have a weaker driving term, when normalized for the same local convergence properties $\left[f'_{\mathrm{grad}}(0) = f'_{\mathrm{prop}}(0)\right]$. Fig. 4 illustrates this inferiority, by plotting the mean update terms versus the frequency discrepancy $\omega_0 - \theta$ for the two schemes, for a particular value of α. The two curves are normalized for the same slope at the zero crossing, to ensure the same local convergence properties. Thus the proposed scheme exhibits faster tracking and convergence than a gradient descent procedure.

As an example, Fig. 5 plots the frequency estimates using a gradient descent algorithm, for the same bandwidth parameter $\alpha = 0.7$ and the same input signal used in Fig. 2, along with the same initialization $\theta(0) = 0$. Only the final frequency in the hop sequence is identified in Fig. 5, as it happens to fall within a basin of attraction of the initialization point.[2] The remaining frequencies simply lie outside the basin of attraction of the gradient algorithm, thus illustrating a clear performance defect compared to the proposed algorithm (1). Similar behavior is observed with the gradient descent algorithm of [8], with the further drawback of bias in the frequency estimate.

[2]Increasing the step size can increase the basin of attraction, at the expense of increasing the estimation variance at a convergent point; the step size has adjusted to offer comparable variance to Fig. 2 for the identified frequency.

VI. Concluding Remarks

The proposed design is a complex-valued counterpart to an earlier real-coefficient scheme [7], and successfully inherits its advantages. Although we have focused on the single tone case for brevity, extensions to multiple tones may be pursued in manners analogous to those developed in the real case [4], [9], [10], [15].

References

[1] A. Nehorai, "A minimal parameter adaptive notch filter with constrained poles and zeros," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-33, no. 8, pp. 983–996, Aug. 1985.

[2] J. M. Travassos-Romano and M. G. Bellanger, "Fast least-squares adaptive notch filtering," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 7, pp. 1536–1540, Sep. 1988.

[3] N. I. Choi, C.-H. Choi, and S. U. Lee, "Adaptive line enhancement using an IIR lattice notch filter," *IEEE Trans. Acoust., Speech, Signal Process,*, vol. 37, no. 4, pp. 585–589, Apr. 1989.

[4] T. Kwan and K. Martin, "Adaptive detection and enhancement of multiple sinusoids using a cascade IIR filter," *IEEE Trans. Circuits Syst,*, vol. 36, no. 7, pp. 937–947, Jul. 1989.

[5] J. A. Chambers and A. G. Constantinides, "Frequency tracking using constrained adaptive notch filters synthesized from allpass sections," *Proc. Inst. Elect. Eng. F*, vol. 137, pp. 475–481, Dec. 1990.

[6] A. Zalnierniunas, "Adaptive notch wave digital filters," in *Int. Symp. Circuits Syst.*, New Orleans, LA, May 1990, pp. 3518–3521.

[7] P. A. Regalia, "An improved lattice-based adaptive IIR notch filter," *IEEE Trans. Signal Process.*, vol. 39, pp. 2124–2128, Sep. 1991.

[8] S.-C. Pei and C.-C. Tseng, "Complex adaptive IIR notch filter algorithm and its applications," *IEEE Trans. Circuits Syst. II*, vol. 41, no. 2, pp. 158–163, Feb. 1994.

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

[10] V. DeBrunner and S. Torre, "Multiple fully adaptive notch filter design based on allpass sections," *IEEE Trans. Signal Process.*, vol. 48, pp. 550–552, Feb. 2000.

[11] Y. C. Lim, Y. X. Zou, and N. Zheng, "A piloted adaptive notch filter," *IEEE Trans. Signal Process.*, vol. 53, pp. 1310–1323, Apr. 2005.

[12] Y. Xiao, L. Ma, K. Khorasani, and A. Ikuta, "Statistical performance of the memoryless nonlinear gradient algorithm for the constrained adaptive IIR notch filter," *IEEE Trans. Circuits Syst. I*, vol. 52, no. 8, pp. 1691–1702, Aug. 2005.

[13] J. Zhou, "Simplified adaptive algorithm for constrained notch filters with guaranteed stability," *Proc. Inst. Elect. Eng., Vis., Image Signal Process.*, vol. 153, no. 5, pp. 574–580, 2006.

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

[15] J. E. Cousseau, S. Werner, and P. D. Donate, "Factorized all-pass based IIR adaptive notch filters," *IEEE Trans. Signal Process.*, vol. 55, pp. 5225–5236, Nov. 2007.

[16] M. Niedzwiecki, "Generalized adaptive notch smoothing revisited," *IEEE Trans. Signal Process.*, vol. 58, pp. 1565–1576, Mar. 2010.

[17] P. P. Vaidyanathan, P. A. Regalia, and S. K. Mitra, "Design of doubly complementary IIR digital filters using a single complex all-pass filter, with multirate applications," *IEEE Trans. Circuits Syst.*, vol. 34, no. 4, pp. 378–389, Apr. 1987.

[18] L. Ljung, "Analysis of recursive stochastic algorithms," *IEEE Trans. Automat. Contr.*, vol. 22, no. 8, pp. 551–575, Aug. 1977.

[19] A. Benveniste, M. Metivier, and P. Priouret*, Algorithmes Adaptatifs et Approximations Stochastiques*. Paris, France: Masson, 1987.

[20] H. J. Kushner*, Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems*. Boston, MA: Birkhäuser, 1990.

[21] A. H. Sayed*, Fundamentals of Adaptive Filtering*. New York: Wiley, 2003.