OPTIMAL DESIGN OF DIGITAL EQUIVALENTS TO ANALOG FILTERS

Grigore I. Braileanu

Gonzaga University
Spokane, WA 99258, USA

ABSTRACT

The proposed optimal algorithm for the digitizing of analog filters is based on two existing filter design methods: the extended window design (EWD) and the matched–pole (MP) frequency sampling design. The latter is closely related to the filter design with iterative weighted least squares (WLS). The optimization is performed with an original MP design that yields an equiripple digitizing error. Then, a drastic reduction of the digitizing error is achieved through the introduction of a fractional time shift that minimizes the magnitude of the equiripple error within a given frequency interval. The optimal parameters thus obtained can be used to generate the EWD equations, together with a variable fractional delay output, as described in an earlier paper. Finally, in contrast to the WLS procedure, which relies on a "good guess" of the weighting function, the MP optimization is straightforward.

Index Terms — IIR digital filters, interpolation, sampling rate conversion, variable fractional delay, approximation methods

1. INTRODUCTION

The design of infinite impulse response (IIR) digital filters is done either by digitizing a prototype analog filter or directly in the discrete domain with some computer–aided iterative procedure. A reason for using analog prototypes is that the design of analog filters is a well–established subject that includes not only closed form solutions and highly advanced analog approximation techniques, but also well–tested computer programs to carry out the designs. Also, many applications are defined in terms of analog transfer functions, but are implemented digitally for better accuracy and reliability. On the other hand, the computer–aided iterative IIR filter design is now an attractive alternative due to the tremendous computational power available to the designer. This type of design allows for a wider variety of user–defined frequency responses and, usually, relies on iterative weighted least squares (WLS) optimization techniques to minimize the error between the desired frequency response and the frequency response of the computer generated filter [1], [2]. Obviously, the WLS method can be used to digitize analog prototypes as well. Yet, the main shortcoming of this approach is the need for repeated trials with various weighting functions aiming at the minimization of the Chebyshev norm of the digitizing error. The present paper addresses this problem by proposing a new optimal algorithm for the approximation of analog filters by discrete filters.

First of all, Section 2 presents an original analysis of two existing interpolation techniques that provide a bridge between the analog prototype methods and the WLS approach: the extended window design (EWD) of IIR discrete equivalents of analog filters [3], and the so–called matched–pole (MP) frequency sampling design [4]. The EWD method is developed with an interpolation–based time domain algorithm, whereas the MP design is done with a frequency domain interpolation.

The main focus is on the EWD approach for both theoretical and practical reasons. First of all, it is argued that the time domain features of the EWD approach justify the search for some optimal value of a time-shift parameter that, when incorporated into the input interpolation process, will compensate for the input–output time delay due to the typical phase response of the analog prototype. An even more important feature of the EWD filters is the natural way of generating intersample output values that leads to the implementation of variable fractional delay (VFD) filters [5]. This procedure that has been developed in [6] is beyond the scope of this paper.

The EWD approach does not lend itself to a simple minimization of the digitizing error. Yet, since the EWD and MP filters are shown in Section 2 below to be identical, the optimization can be performed with the MP method. First, a fast MP algorithm is presented in Section 3 for the search of the design parameters that produce an equiripple digitizing error. Then, at the end of Section 3, another MP algorithm determines the value of the fractional time shift τ that minimizes the equiripple digitizing error within a predefined frequency interval. The actual transfer function of the digital filter is explicitly computed only once, at the very end of the optimization process. The convergence of the optimization algorithm, as well as the related problem of the best approximation in the sense of Chebyshev norm, is also discussed. The implications of the optimal time shift on the conventional WLS optimization techniques are mentioned in the concluding Section 4. Finally, it is shown that as soon as the optimal design parameters are selected, they can be used to implement variable fractional delay (VFD) filters with the EWD method.

2. PRELIMINARIES

Throughout the paper, the time is normalized to the sampling period, and so the folding frequency is $\omega_f = \pi$. The signals $x(t)$ and $y(t)$ are the input and output of the analog filter defined by

$$H_A(s)=\frac{Y(s)}{X(s)}=\frac{b(s)}{a(s)}=\frac{b(s)}{s^{n_A}+a_1 s^{n_A-1}+...+a_{n_A}}, \qquad (1)$$

where $b(s)$ and $a(s)$ are known polynomials. The digitizing problem consists in finding a transfer function,

$$H_D(z)=\frac{Y_D(z)}{X_D(z)}=\frac{f(z^{-1})}{g(z^{-1})}=\frac{f_0+f_1 z^{-1}+...+f_m z^{-m}}{1+g_1 z^{-1}+...+g_{n_A} z^{-n_A}} \qquad (2)$$

such that the frequency response $H_D(e^{j\omega})$ represents a "good approximation" of $H_A(j\omega)$. The input and output of the designed digital filter are denoted below by $x_D[k]$ and $y_D[k]$, respectively, while the sampled values of $x(t)$ and $y(t)$ are $x(k)$ and $y(k)$.

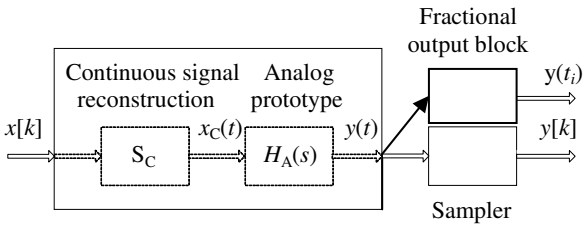

Fig. 1. Block diagram illustrating the extended window design (EWD), where S_C and $H_A(s)$ are *combined into just one block* with the same number of system modes as $H_A(s)$.

2.1. The Extended Window Design of IIR Digital Filters

The basic idea of this design is illustrated in Fig. 1, where the continuous signal reconstruction block S_C is an $(m+1)$–point interpolator. Fig. 2 shows the recursive generation of the output $y[k]$ from the response $y(t)$ of the analog prototype $H_A(s)$ to the (virtual) signal $x_C(t)$ that interpolates the last $(m+1)$ input samples $x[k]$ up to the current time $t=k$. Yet, the seemingly natural implementation that would process the two internal blocks S_C and $H_A(s)$ as individual blocks is unacceptable. Indeed, this would increase the order of the final digital filter $H_D(z)$ to $n_D=n_A+m$, and significantly alter the frequency response $H_D(e^{j\omega})$ because of the cascade interconnection. By contrast, the EWD approach [3], [4], is characterized by a particular choice of the auxiliary conditions needed to calculate the response $y(t)$ of $H_A(s)$. These conditions, represented in Fig. 2 by stars, are the last n_A output samples grouped in the vector $\mathbf{y}_{aux} = [y(k–n_A),..., y(k–1)]$, for $\tau = 0$, or $\mathbf{y}_{aux} = [y(k–\tau–n_A),..., y(k–\tau–1)]$, for $0 < \tau < 1$. It is worth noting that the fractional time-shift parameter, τ, can be easily introduced in the design through the interpolator, as shown below in (4) and (4'). In either case (i.e., $\tau = 0$, or $0 < \tau < 1$), this design provides a natural match for the initial conditions of the analog and digital filters, and so it incorporates the interpolation step into the s- to z-domain mapping step. Thus, the two internal blocks S_C and $H_A(s)$ in Fig. 1 are now processed together, rather than cascaded. Consequently, it was proven in [3] and [4] that the denominators $a(s)$ and $g(z^{-1})$ in (1) and (2) are related by the expressions

$$a(s) = \prod_{n=1}^{n_A} (s - p_n), \qquad g(z^{-1}) = \prod_{n=1}^{n_A} (1 - e^{p_n} z^{-1}). \qquad (3)$$

The reason that this method is referred to as the extended window design is that, at each current time k, *the interpolation time window* extends to the left of the current time k along a segment of length $m \geq n_A$. Nevertheless, it is worth noting that the actual interpolation is transparent to the designer. Specifically, the virtual continuous signal $x_C(t)$ is built sequentially, one sampling period at a time, from the last $(m+1)$ input samples. The design parameter m is chosen to be larger than n_A by two to five units, according to a trade-off between the IIR filter complexity and the digitizing accuracy. In this paper, m will be always an odd number for the sake of a more concise presentation. Thus, a new integer parameter, M, can be introduced as $M = (m+1)/2$. In addition, a distinction is made between the vector $\mathbf{w}$ of the *frequency nodes, which cancel the digitizing error*, $\{w_n \mid 0 < w_n < \pi, n = 1, \ldots, M\}$, and the much larger vector $\boldsymbol{\omega}$ of frequencies $\{\omega \mid 0 < \omega < \pi\}$ that are used for the numerical computation of the digitizing error.

The design starts with the continuous–time reconstruction of the input sequence through the interpolation expression

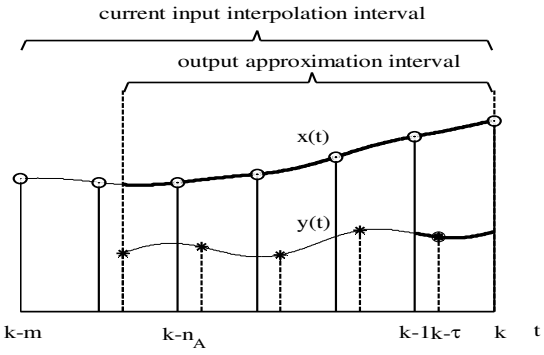

Fig.2. Interpolation intervals for $m = 6$, $n_A = 4$, and $\tau = 0.7$. The thick line of $y(t)$, $k–1 < t < k$, yields an optional interpolated output.

$$x_C(t) = \sum_{\substack{n=-M \\ n \neq 0}}^{M} c_n e^{j w_n (t-\tau)}, \quad k - m \leq t \leq k, \qquad (4)$$

which includes an optional time shift τ. *The parameter τ may be chosen to be zero,* or *optimized* according to the procedure proposed in Section 3.3 below. The expression $x_C(t)$ is uniquely computed from the $(m+1)$ input samples $x(\ell)$, $\ell = k–m, \ldots, k$, by solving $(m+1)$ linear algebraic equations for the coefficients c_n,

$$\sum_{\substack{n=-M \\ n \neq 0}}^{M} c_n e^{j w_n (\ell-\tau)} = x(\ell), \quad \ell = k - m, \ldots, k. \qquad (4')$$

Finally, an efficient numerical method [3] expresses the response $y(t)$ of the analog filter to $x_C(t)$ in terms of matrix invariants that provide intersample values [6] but, also, relate the input vector $[x(k–m), ..., x(k)]$ to the vector $[y(k–\tau–n_A),..., y(k–\tau–1)]$ of auxiliary conditions. The result is the EWD transfer function $H_{EWD}(z) = H_D(z)$. Both the numerical generation of the EWD filter [3] and the closed–form solution [4] provide the following:

Theorem: The real–time implementation of the EWD transfer function (2), possibly designed with an optional time shift τ, is of order $m \geq n_A$, and has n_A poles, z_n, related to the poles p_n of the analog prototype by $z_n = e^{p_n}$, $n = 1, \ldots, n_A$, while the remaining $(m–n_A)$ poles are placed at $z = 0$. The current discrete output corresponds to the relationship $y_{EWD}[k] = y(k - \tau)$.

2.2. The importance of the EWD approach

The interest in the EWD method is due to three reasons: (i) available closed–form solutions [4] can be used to build accurate benchmark examples, (ii) the time–domain form of the design reveals ways toward the reduction of the digitizing error, and (iii) the EWD interpolation algorithms provide a simple and accurate computation of interpolated output values [3], [6]. For example, the time–domain features of the EWD approach, illustrated in Fig. 2, justify the search for some optimal value, τ_{opt}, of a time-shift parameter, τ. Indeed, the need for this optimization stems from the fact that both the conventional and optimal digitizing methods suffer from limitations due to the restrictive nature of the discrete representations of continuous systems. Specifically, the difference equations of IIR filters are traditionally designed as if the input and output filter sequences were in strict correspondence with the sampled values of their continuous counterparts. This design does not reflect the fact that the continuous output incorporates a time delay caused by the typical phase response of the analog prototype.

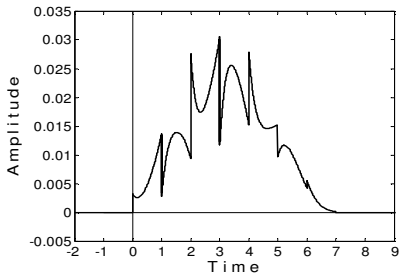

Fig.3. The function $h_\tau(t)$ with compact support, $t \in [0,n_A]$, for n_A=7.

The delay affects the accuracy of the current input and output components of the discrete equations in the following ways: (i) the last few input samples have little effect on the current output sample, (ii) the input and output sequences are not well correlated at sampling times, and (iii) the output samples also depend on input samples that occurred prior to the input samples in the current equation. Thus, it is apparent that, in the context of the EWD method, there should be an optimal time shift that minimizes the digitizing error.

While the EWD filters are essentially designed in the time domain, the digitizing error is defined in the frequency domain and is controlled by the M parameters $\{w_n \mid n = 1, \ldots, M\}$ and the time shift τ. The optimization problem is addressed in Section 3.

2.3. The Matched–Pole Frequency Sampling Design

The EWD optimization proposed in Section 3 below is based on the identity between the EWD filters and the filters designed with the so–called matched–pole (MP) method [4]. The latter is related to the frequency sampling method for IIR filter design [2, pp. 218–221] that is used to find a discrete transfer function (2) whose frequency response interpolates a given complex expression $H_A(j\omega)$ at equally spaced frequency points. This method consists in solving the equations

$$\frac{f\left(e^{-j\omega_n}\right)}{g\left(e^{-j\omega_n}\right)} = H_A(j\omega_n), \qquad n = 1,\ldots,(n_A+m), \tag{5}$$

for the (n_A+m) unknown coefficients in (2). The reason given in [2] for the inaccurate results is the poor estimation of the denominator $g(z^{-1})$, which may even produce an unstable filter.

By contrast, the digitizing of analog filters can be done with a predetermined denominator $g(z^{-1})$. Indeed, since both $H_A(s)$ and $H_D(z)$ defined in (1) and (2) are analytic rational functions, the properties of analytic functions impose some rather powerful constraints on their behavior within the respective regions of convergence. One consequence of these constraints is the relationship $z_n = e^{p_n}$ mentioned in the above Theorem. In general, given an analog system, the Laplace transforms of all the system modes and signals have the poles related to the poles of their z-transform counterparts by the s–to–z domain mapping $z = e^s$. Accordingly, the digitizing of filters by frequency sampling can be reduced to the computation of the numerator $f(z^{-1})$ in (2), while the denominator $g(z^{-1})$ is fixed and defined by (3). The design parameters n_A, m, and τ are chosen to be the same as those of the EWD filter, while (5) is replaced by $m+1 = 2M$ equations,

$$f\left(e^{-jw_n}\right) = e^{-j\tau w_n} H_A(jw_n)\, g\left(e^{-jw_n}\right), \quad n = \pm 1, \ldots, \pm M, \tag{6}$$

where the only unknowns are the $(m+1)$ coefficients of $f(z^{-1})$.

With the poles of $g(z^{-1})$ rigorously matched to the poles of $a(s)$, the IIR filter design has now become an FIR frequency–sampling problem that is referred to as the *matched-pole* (MP) *frequency sampling design*.

The conceptual difference between the two types of interpolation represented by (5) and (6) can be better understood by analyzing the properties of the function

$$H_\tau(s) \triangleq e^{-\tau s} H_A(s)\, g\left(e^{-s}\right) = e^{-\tau s} b(s) \prod_{n=1}^{n_A} \frac{1-e^{-(s-p_n)}}{s-p_n} \tag{7}$$

which is directly related to the right-hand side of (6).

Property 1: $H_\tau(s)$ is an entire function of the complex variable s, since $b(s)$ is a polynomial and the pole and zero of each one of the n_A factors in (7) cancel out.

Property 2: Let $h_\tau(t)$ be the inverse Laplace transform of $H_\tau(s)$. Then, $h_\tau(t)$ is a function with compact support of length n_A, as illustrated in Fig. 3.

Proof: We introduce the following notations:

$$H_0(s) = b(s) H_R(s), \quad H_R(s) = \prod_{n=1}^{n_A} R_n(s), \quad R_n(s) = \frac{1-e^{-(s-p_n)}}{s-p_n},$$

and the inverse Laplace transforms $h_R(t) = L^{-1}\{H_R(s)\}$ and

$$r_n(t) = L^{-1}\left\{\frac{1-e^{-(s-p_n)}}{s-p_n}\right\} = e^{p_n t}\left[u(t)-u(t-1)\right], \tag{8}$$

where $u(t)$ is the unit step. The function $r_n(t)$ has the compact support [0,1], Then, $h_R(t)$ is the result of the repeated convolution

$$h_R(t) = \left(r_1 * r_2 * \cdots * r_{n_A}\right)(t) \tag{9}$$

of n_A functions with support [0,1] and so, according to a convolution theorem, it has the compact support [0, n_A]. Finally, $h_0(t)$ and $h_\tau(t)$ have also a compact support of length n_A, since $h_0(t)$ is a linear combination of derivatives of $h_R(t)$, and $h_\tau(t) = h_0(t-\tau)$.

Property 3: The Fourier transform of $h_\tau(t)$ is $H_\tau(j\omega)$, which is simply obtained from (7) by replacing s with $j\omega$.

It follows that the fundamental difference between the direct interpolation (5) and the FIR-like interpolation (6) stems from the dual form of the Nyquist-Shannon sampling theorem [1], where ω is viewed as a "time variable" whereas t becomes a "frequency variable." Thus, the above derivations show that the conditions of the sampling theorem are far from being satisfied by the method described in [2], where $h_A(t) = L^{-1}\{H_A(s)\}$ extends over an excessively wide time segment, but are fully satisfied by the MP design since $h_\tau(t)$ is strictly zero outside an interval of length n_A.

2.4 Identity of the EWD and MP Frequency Sampling Design

Since (4') has a unique solution, each sampled exponential from the set of linearly independent functions $\{e^{jw_n t},\ n = \pm 1, \ldots, \pm M,\ w_{-n} = -w_n\}$ is exactly reconstructed by the EWD interpolator S_C *with the optional time shift* τ. Then, for each $x(t) = e^{jw_n t}$, the exact sampled response of the filter $H_A(s)$ to $x_C(t)$ (see Fig. 1) is

$$y_{EWD}[k] = e^{-j\tau w_n} H_A(jw_n)\, e^{jw_n k}. \tag{10}$$

At the same time, the response of the MP digital filter to a discrete time sequence, $e^{jw_n k}$, is $y_{MP}[k] = H_{MP}(e^{jw_n})\, e^{jw_n k}$ where, according to (6), $H_{MP}(e^{jw_n}) = e^{-j\tau w_n} H_A(jw_n)$. Thus, the last two expressions, together with (10), yield the $y_{MP}[k] = y_{EWD}[k]$. Finally, $H_{MP}(z) \equiv H_{EWD}(z)$, since $H_{MP}(z)$ and $H_{EWD}(z)$ have the

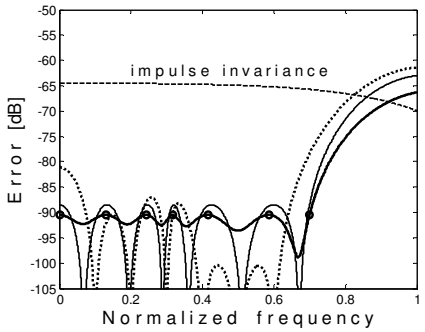

Fig. 4. Digitizing errors, $|E(\omega)|_{\mathrm{dB}}$, for the lowpass analog filter

$$H_{\mathrm{A}}(s)=\frac{0.0033(s^2+12.26)(s^2+4)(s^2+2.69)}{(s+0.768)(s^2+0.23s+0.967)(s^2+0.735s+0.86)(s^2+1.27s+0.686)}.$$

Impulse invariant design (top dashed line); EWD design with τ=0 and equally spaced nodes (dotted line), equiripple EWD with τ =0 (thin solid line), and best Chebyshev norm approximation (thick solid line). The frequency is normalized to π.

same denominator, $g(z^{-1})$, by design, and the (m+1) coefficients of the MP numerator, $f(z^{-1})$, are uniquely determined by (6).

3. EQUIRIPPLE OPTIMIZATION OF EWD FILTERS

3.1. Computation of the Digitizing Error

The optimization is done with a fast algorithm based on the MP interpolation equations (6) and consists in finding the (m+1) real coefficients of $f(z^{-1})$ that minimize the Chebyshev norm,

$$\left\|E(\omega)\right\| = \max_{0<\omega<\omega_{\max}} \left|E(\omega)\right|, \qquad (11)$$

of the digitizing error,

$$E(\omega) = \frac{f\left(e^{-j\omega}\right)}{g\left(e^{-j\omega}\right)} - e^{-j\tau\omega} H_{\mathrm{A}}(j\omega), \qquad (12)$$

where $\omega_{\max}$ is predetermined, $g(z^{-1})$ is given by (3), and τ is to be obtained in the final stage of the optimization in Section 3.3 below.

Fig. 4 shows typical plots of $|E(\omega)|_{\mathrm{dB}} = 20 \log_{10} |E(\omega)|$, corresponding to $\tau = 0$, $m = 11$, and $\omega_{\max} = 0.7\ \pi$. First, a near-optimal EWD filter is obtained by computing the frequency nodes that yield an equiripple magnitude $|E(\omega)|_{\mathrm{dB}}$ within the passband and transition bands of the analog prototype $H_{\mathrm{A}}(s)$, like that one exemplified in Fig. 4 with solid line. In general, the error of the best Chebyshev norm approximation discussed in Section 3.3 does not need to cancel, as shown in Fig. 4 by the thick solid line.

The symmetry properties of the transfer functions and $E(\omega)$ allow for the derivations below to be done only for the M positive nodes, whereas the actual nodes are w_n, $n = \pm1, \ldots, \pm M$, $w_{-n} = -w_n$. Similarly to the Remez exchange algorithm [7], instead of trying to reduce the error at each iteration, Algorithm 3 in Section 3.2 will adjust the reference $\{w_n \mid n = 1, \ldots, M\}$ such that it converges to a point set $\{w_n^*\}$ that yields an equiripple $|E(\omega)|_{\mathrm{dB}}$ on $\omega \in [0,\omega_{\max}]$.

To this end, the interpolation of $H_\tau(j\omega)$ defined by (6) and (7) is done with the Lagrange formula, thus calculating $E(\omega)$ directly, without the coefficients of the complex polynomial $f(z^{-1})$:

$$f\left(z^{-1}\right) \triangleq \sum_{\substack{\ell=-M \\ \ell\neq 0}}^{M} H_\tau(jw_\ell) \prod_{\substack{n=-M \\ n\neq\ell \\ n\neq 0}}^{M} \frac{z_n^{-1}-z^{-1}}{z_n^{-1}-z_\ell^{-1}}, \qquad (13)$$

where $z_n^{-1}=e^{-jw_n}$. First of all, the interpolation of $H_\tau(j\omega)$ with $f\left(e^{-j\omega}\right)$ and the interpolation of $e^{-j\frac{m}{2}\omega}H_\tau(j\omega)$ with $e^{-j\frac{m}{2}\omega}f\left(e^{-j\omega}\right)$ are equivalent. Then, since m is odd, the interpolation function

$$e^{j\frac{m}{2}\omega}f\left(e^{-j\omega}\right)=f_0e^{j\frac{m}{2}\omega}+\ldots+f_{M-1}e^{j0.5\omega}+f_Me^{-j0.5\omega}+\ldots+f_me^{-j\frac{m}{2}\omega}$$

can be expressed in terms of a linear combination of the linearly independent trigonometric functions $\{\cos(n\text{-}0.5)\omega,\ \sin(n\text{-}0.5)\omega,\ n=1, \ldots, M\}$ with period 4π. The trigonometric form of (13),

$$e^{j\frac{m}{2}\omega}f\left(e^{-j\omega}\right)=\sum_{\substack{\ell=-M \\ \ell\neq 0}}^{M} e^{j\frac{m}{2}w_\ell}H_\tau(jw_\ell)\prod_{\substack{n=-M \\ n\neq\ell \\ n\neq 0}}^{M}\frac{\sin\frac{\omega-w_n}{2}}{\sin\frac{w_\ell-w_n}{2}}, \qquad (14)$$

can be expressed in terms of the M positive frequency nodes alone. First, by using the symmetry properties of $H_\tau(j\omega)$ and grouping some terms, the following trigonometric polynomials are defined:

$$P_\ell(\omega) = \prod_{\substack{n=1 \\ n\neq\ell}}^{M}\frac{\cos\omega-\cos w_n}{\cos w_\ell-\cos w_n}. \qquad (15)$$

Then, $f(e^{-j\omega})$ becomes

$$f(e^{-j\omega})=e^{-j\frac{m}{2}\omega}\cos 0.5\omega\sum_{\ell=1}^{M}\frac{\mathrm{Re}\left\{e^{j\frac{m}{2}w_\ell}H_\tau(jw_\ell)\right\}}{\cos 0.5w_\ell}P_\ell(\omega)$$

$$+je^{-j\frac{m}{2}\omega}\sin 0.5\omega\sum_{\ell=1}^{M}\frac{\mathrm{Im}\left\{e^{j\frac{m}{2}w_\ell}H_\tau(jw_\ell)\right\}}{\sin 0.5w_\ell}P_\ell(\omega). \qquad (16)$$

Finally, the expression (16) replaces $f(e^{-j\omega})$ in (12) to yield $E(\omega)$ for given sets of frequencies, $\omega \in [0, \omega_{\max}]$, and frequency nodes, $\{w_n \mid n=1,\ldots,M\}$, grouped in the vectors $\boldsymbol{\omega}$ and $\mathbf{w}$, respectively. In this form, the numerical processing of $E(\omega)$ can be efficiently done with vector and matrix invariants, as shown in the next section.

3.2. Computation of the Equiripple Digitizing Error

The smoothness of the function $H_\tau(j\omega)$, together with the properties discussed in Section 2.3, suggests a smooth magnitude $|E(\omega)|_{\mathrm{dB}}$ as well. In fact, $|E(\omega)|_{\mathrm{dB}}$ is characterized by (M+1) lobes on the imposed interval $\omega \in [0,\omega_{\max}]$.

The equiripple magnitude $|E(\omega)|_{\mathrm{dB}}$ is obtained with the iterative Algorithm 3, which calls Algorithms 1 and 2. With m and $\omega_{\max}$ fixed, Algorithm 3 is called repeatedly for each new value of the time shift τ, according to the optimization process presented in Section 3.3 below. The algorithms are described in pseudo-code.

Algorithm 1. Determine: E_ℓ — the maxima of $|E(\omega)|_{\mathrm{dB}}$, v_ℓ — the abscissa of E_ℓ, and the lengths Δ_ℓ, $\ell = 0,1,\ldots,M$, of the error lobes on the segments $[0,w_1]$, $\{[w_\ell,w_{\ell+1}],\ \ell = 1,\ldots,M\text{–}1\}$, and $[w_M, \omega_{\max}]$

- compute:
  $E_0 = \max\{\mathrm{Edb}(\boldsymbol{\omega})\}$ on $\omega \in [0, w_1]$, v_0, $\Delta_0 = w_1$
- FOR $\ell = 1,\ldots,M\text{–}1$, compute:
  $E_\ell = \max\{\mathrm{Edb}(\boldsymbol{\omega})\}$ on $\omega \in [w_{\ell+1},w_\ell]$, v_ℓ, $\Delta_\ell = w_{\ell+1} - w_\ell$
  END FOR
- compute:
  $E_M = \max\{\mathrm{Edb}(\boldsymbol{\omega})\}$ on $\omega \in [w_M, \omega_{\max}]$, v_M, $\Delta_M = \omega_{\max} - w_M$
  $E_{\inf} = \min\{E_0, \ldots, E_M\}$, $E_{\sup} = \max\{E_0, \ldots, E_M\}$

Moreover, a detailed analysis of the expression $|E(\omega)|_{\text{dB}}$ calculated in terms of (15) and (16) justifies the following:

Proposition: The maximum value E_ℓ of any given lobe of $|E(\omega)|_{\text{dB}}$ decreases when the length $\Delta_\ell = w_{\ell+1} - w_\ell$ decreases and, conversely, E_ℓ increases when Δ_ℓ increases.

Indeed, this property, which is obvious on the factors $\mathrm{P}_\ell(\omega)$ defined by (15), propagates throughout the terms of (16).

The iterative search for an equiripple $|E(\omega)|_{\text{dB}}$ is done below by adjusting the lengths Δ_ℓ according to the above Proposition. The numerical examples illustrated in Figs. 4-6 are done with $n_{\text{A}} = 7$, $m = 11$, a vector $\boldsymbol{\omega}$ of length $L = 256$, and a vector $\mathbf{w}$ of length $M = (m+1)/2 = 6$. In the algorithms below, the element–by–element multiplication and division of vectors are represented by $\circ$ and $\circ/$, respectively. The vector multiplication is represented by *.

Algorithm 2: Compute $E(\boldsymbol{\omega})$ and $|E(\omega)|_{\text{dB}}$ for a given time shift τ and vector $\mathbf{w}$ (use the vectors $\mathbf{h}$, $\mathbf{r}$, $\mathbf{i}$, $\mathbf{c}$, and $\mathbf{g}$ from Algorithm 3)

- given the vector $\mathbf{w}$, compute the following:

  $\mathbf{f_w} = e^{j\frac{m}{2}\mathbf{w}} \circ H_\tau(j\mathbf{w}), \quad \mathbf{r_w} = \mathrm{Re}\{\mathbf{f_w}\}, \quad \mathbf{i_w} = \mathrm{Im}\{\mathbf{f_w}\},$

  and the matrix P: $L \times M$ defined by the vector $\mathbf{c}$ and (15)
- compute $\mathrm{E}(\boldsymbol{\omega})$:

  $\mathbf{f} = \mathbf{r} \circ \{\mathrm{P} * [\mathbf{r_w} \circ/ \cos(0.5\mathbf{w})]\} + j\ \mathbf{i} \circ \{\mathrm{P} * [\mathbf{i_w} \circ/ \sin(0.5\mathbf{w})]\}$

  $\mathrm{E}(\boldsymbol{\omega}) = \mathbf{f} \circ/ \mathbf{g} - \mathbf{h}, \quad \mathrm{Edb}(\boldsymbol{\omega}) = 20\log_{10}|E(\boldsymbol{\omega})|$

Algorithm 3: Determine the vector $\mathbf{w}$ that provides an equiripple magnitude $|E(\omega)|_{\text{dB}}$ on $0 < \omega \le \omega_{\max}$ for a given time shift τ

- **Initialization** (done once and for all for)
  - build the vector $\boldsymbol{\omega}$ with L frequency points equally spaced on $0 < \omega \le \omega_{\max} < \pi$ (usually, $L \le 256$)
  - compute the vectors $\mathbf{h}$, $\mathbf{r}$, $\mathbf{i}$, $\mathbf{c}$, and $\mathbf{g}$:

    $\mathbf{h} = e^{-j\tau\boldsymbol{\omega}} \circ H_{\text{A}}(j\boldsymbol{\omega}), \mathbf{r} = e^{-j\frac{m}{2}\boldsymbol{\omega}} \circ \cos 0.5\boldsymbol{\omega}, \ \mathbf{i} = e^{-j\frac{m}{2}\boldsymbol{\omega}} \circ \sin 0.5\boldsymbol{\omega},$

    $\mathbf{c} = \cos\boldsymbol{\omega}, \quad \mathbf{g} = g\left(e^{-j\boldsymbol{\omega}}\right)$

  - build the vector $\mathbf{w}$ with M frequency nodes equally spaced on $0 < \omega < \omega_{\max} < \pi$
  - call **Algorithm 2** to compute $E(\boldsymbol{\omega})$
  - call **Algorithm 1** to compute $\{E_0, \ldots, E_M\}$, E_{inf} and E_{sup}
- $\delta = 1, \quad K = 0$
- WHILE $E_{\text{sup}} - E_{\text{inf}} > 0.5$ dB (*threshold set by the designer*)

  $\delta = \max\{\delta/2, 2^{-20}\}$

  WHILE $K < 4$
  - sort the set $\{E_0, \ldots, E_M\}$ and divide it evenly into two sets: the set {Low} with the lowest maxima, and the set {High} with the highest maxima
  - IF the lobes in the two sets interchange from the previous run, THEN $K=K+1$, ELSE, $K=0$, END IF
  - IF M is even, THEN $\Delta=\Delta_{M/2}$, ELSE $\Delta=0$, END IF

    increase the lengths Δ_ℓ^{Low} in the set {Low}, and decrease the lengths $\Delta_\ell^{\text{High}}$ in the set {High}:

    $\Delta_\ell^{\text{Low}} \leftarrow (1+\delta)\,\Delta_\ell^{\text{Low}}, \quad \Delta_\ell^{\text{High}} \leftarrow (1-\delta)\,\Delta_\ell^{\text{High}}$
  - $\mathrm{S} = \mathrm{SUM}\{\Delta_\ell^{\text{Low}}\} + \mathrm{SUM}\{\Delta_\ell^{\text{High}}\} + \Delta$

    $\Delta_\ell^{\text{Low}} \leftarrow \omega_{\max}\,\Delta_\ell^{\text{Low}}/\mathrm{S}, \quad \Delta_\ell^{\text{High}} \leftarrow \omega_{\max}\,\Delta_\ell^{\text{High}}/\mathrm{S}$

    IF M is even, THEN $\Delta_{M/2} \leftarrow \omega_{\max}\,\Delta_{M/2}/\mathrm{S}$, END IF
  - update the vector $\mathbf{w}$ for the new lengths Δ_ℓ
  - call **Algorithm 2** to compute $E(\boldsymbol{\omega})$
  - call **Algorithm 1** to get $\{E_0, \ldots, E_M\}$, E_{inf} and E_{sup}

  END WHILE

  END WHILE

The $(m+1)$ coefficients of $f(z^{-1})$ can now be determined from (6).

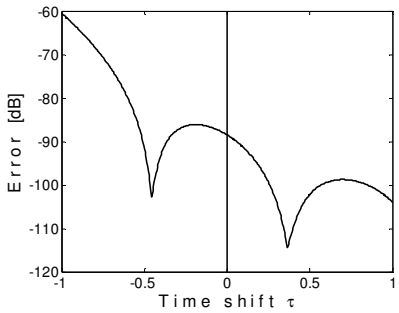

Fig. 5. The Chebyshev norm $\|E(\omega)\|$ plotted as a function of the time shift τ for the EWD designs illustrated in Fig. 4 (τ_{opt}=0.365).

Based on the properties stated in the Proposition above, an analysis of Algorithm 3 shows that the inner WHILE loop converges to a limit cycle for each new value of the parameter δ. When a limit cycle is reached, the lobes in the sets {Low} and {High} interchange at each new node adjustment done with the same δ. The inner loop is terminated when this interchange occurs three times. As δ is halved on each new run of the outer loop, the difference $(E_{\text{sup}}–E_{\text{inf}})$ corresponding to each new limit cycle decreases, and the algorithm ends when a threshold (e.g., 0.5 dB) is reached. The equiripple design illustrated in Fig. 4 took 124 iterations in 0.17 seconds to complete on an Intel Pentium computer (3 GHz, 1 GB of RAM).

In the next section we argue that, given $H_{\text{A}}(s)$, m, and $\omega_{\max}$, the equiripple minimization is optimum and unique for an optimal τ.

3.3. The Optimization Algorithm

The vector $\mathbf{w}$ that provides an equiripple magnitude $|E(\omega)|$ for a given time shift τ and frequency range $0 < \omega \le \omega_{\max} < \pi$ is obtained with Algorithms 1– 3. As mentioned in Section 2.2, an optimal time shift, τ_{opt}, is needed due to the inherent time delay between the input and the output of the analog prototype. Thus, the interpolation can be substantially improved by repeatedly running the above algorithms for values of τ within the interval [0,1] and plotting $\|E(\omega)\|$ as a function of τ, as illustrated in Fig. 5. Typically, this plot is very smooth and has a single minimum in the interval [0,1]. Accordingly, a divide–and–conquer procedure can be applied by running the above algorithms for, say, 16 values of τ $(0 \le \tau < 1)$, and then, if necessary, repeating the procedure with smaller increments of τ around the previous minimum.

It is worth noting that the focus on the EWD filters restricts the approximation to an MP interpolation problem which, in general, is not the same as the problem of the best possible Chebyshev approximation in the complex plane [8]. While the Remez Alternation Theorem [7] implies that the best Chebyshev approximation in the space of *real functions* is the same as the best *interpolation* that minimizes the Chebyshev norm (11), this is not generally true of the space of complex functions. Yet, the analyticity properties of the EWD problem lead to a simple algorithm for the Chebyshev optimization $[-\omega_{\max},\omega_{\max}]$. The only purpose of this algorithm is to let the designer test the following two conjectures.

Conjecture 1: For $\tau \ne \tau_{\text{opt}}$, the equiripple EWD error differs only slightly from the Chebyshev error, as illustrated in Fig. 4.

Conjecture 2: If the equiripple optimization is done with the optimal time shift, τ_{opt}, then $H_{\text{EWD}}(e^{j\omega})$ coincides with the best Chebyshev approximation of $e^{-j\tau_{\text{opt}}\omega} H_{\text{A}}(j\omega)$, as shown in Fig. 6.

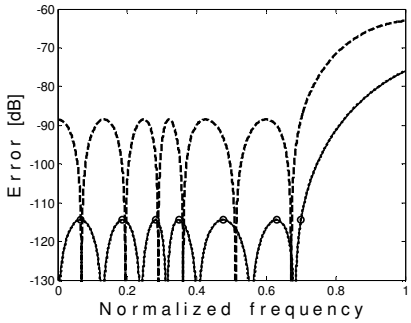

Fig.6. Digitizing errors, $|E(\omega)|_{\mathrm{dB}}$, for the analog filter defined by $H_{\mathrm{A}}(s)$ in Fig. 4: equiripple EWD design with $\tau = 0$ (top dotted line), optimal equiripple EWD for $\tau_{\mathrm{opt}} = 0.365$ (solid line), and best Chebyshev norm approximation for $\tau_{\mathrm{opt}} = 0.365$ (thick dashed line); actually, the latter two plots coincide.

The best Chebyshev approximation defined in Section 3.1 in terms of (11) and (12) is now given a specific formulation based on the fact that $g(z^{-1})$ is fixed and expressed by (3). Moreover, $g(\mathrm{e}^{-j\omega})$ does not cancel for any real value of ω since the system (1) is assumed to be stable. Now, let $\varphi_k(\omega) = \mathrm{e}^{-jk\omega}/g(\mathrm{e}^{-j\omega})$, $k = 0,\ldots,m$, be the basis functions of the complex space Φ of functions

$$\varphi(\omega) \triangleq H_{\mathrm{D}}\left(e^{j\omega}\right) = \sum_{k=0}^{m} f_k\, \varphi_k(\omega) \qquad (17)$$

that approximate $H(j\omega) = \mathrm{e}^{-j\tau\omega} H_{\mathrm{A}}(j\omega)$. The *necessary* conditions of [8, Sect. 3.2] set for the coefficients $\{f_k \,|k = 0, \ldots, m\}$ of the *unique* polynomial $f(z^{-1})$ that minimizes the Chebyshev norm (11) of the error (12) are met for $\omega \in [-\omega_{\max}, \omega_{\max}]$: (i) the space Φ satisfies the Haar condition [8] since it possesses the property that every function in Φ which is not identically zero vanishes at no more than m points, (ii) the maxima of the equiripple $|E(\omega)|$ have *the same value* at r points, and (iii) $m+2 \le r \le 2m+3$. Considering the $2M$ zeros of $|E(\omega)|$, it can be shown that, in general, $r = m+2$ with one maximum at $\omega = 0$, whereas $r = m + 3$ occurs only when τ is in a narrow neighborhood of τ_{opt} . This always amounts to $(M+1)$ maxima within the interval $\omega \in [0, \omega_{\max}]$, as seen in Figs. 4 and 6 (computed with $m = 11$ and $M = 6$).

The Lawson algorithm for the best Chebyshev norm approximation [9], outlined in Algorithm 4 below, can be used to test the above conjectures. Typical plots are given in Figs. 4 and 6.

Algorithm 4. Iterative reweighted least squares error computation.

- **Initialization** (done once and for all)
  - build the vector $\boldsymbol{\omega}$ with L frequency points, ω_i, equally spaced on $-\omega_{\max} \le \omega \le \omega_{\max}$ (usually, $L = 512$)
  - initialize the weight vector: $\boldsymbol{\eta} = [1, \ldots, 1]$ of length L
- FOR $\ell = 1,\ldots, L_0$ (usually $L_0 < 50$)
  - Compute the filter coefficients $\{f_k \,|k = 0, \ldots, m\}$ as the LS solution of the overdetermined system of equations

$$\eta_i \sum_{k=0}^{m} f_k\, \varphi_k(\omega_i) = \eta_i\, e^{-j\tau\omega_i} H_{\mathrm{A}}(j\omega_i), \quad i=1,\ldots, L,$$

  - compute $|E(\omega_i)|$, $i = 1, \ldots, L$, where $E(\omega)$ is defined by (12)
  - update the weight vector $\boldsymbol{\eta}$ with Lawson's formula:

$$\eta_i \leftarrow \frac{\eta_i\, |E(\omega_i)|}{\sum_{i=0}^{2L+1} \eta_i\, |E(\omega_i)|}, \quad i=1,\ldots, L.$$

END FOR

It is worth noting that, although the Lawson algorithm is known to converge very slowly, the analyticity properties of the present approximation problem make the error $E(\omega)$ become practically steady after just a few iteration steps. Finally, since it is proven that, under the above conditions (i)–(iii), the Lawson algorithm converges to the unique best Chebyshev norm approximation [9], the above test done with Algorithm 4 also supports the assertion that the method proposed in this paper yields the unique optimal IIR filter for the parameters m, $\omega_{\max}$, and τ_{opt} .

4. CONCLUSIONS

The paper has developed an original algorithm for the optimal digitizing of analog filters with equiripple error, and proposed a drastic reduction of the digitizing error by choosing $e^{-\tau s} H_{\mathrm{A}}(s)$ as target of the optimization, rather than $H_{\mathrm{A}}(s)$. This modification of the analog transfer function amounts only to a slight increase of the inherent group delay of $H_{\mathrm{A}}(s)$ by τ ($\tau < 1$).

The proposed method is meant to be used only for the design of the optimal digital filter. Then, the EWD algorithm [6] would provide the equations for intersample output, as well as VFD implementations when necessary. At the same time, the optimal τ can be also used by the WLS digitizing of analog filters, according to the previously proven equivalence between the EWD and WLS designs [4]. In addition, Figs. 4 and 6 suggest that the existing methods that determine the best Chebyshev approximation in the space of complex functions benefit of the time shift optimization as well. Yet, all these methods, including the WLS design, require computationally expensive iterative procedures that may or may not converge, and so they are not suitable for the optimization of the time shift τ presented in Section 3.3 above. By contrast, the proposed optimization is completely automated: it needs only the initial parameters, m and $\omega_{\max}$, which are kept unchanged for each new time shift trial: m is chosen one to five units greater than n_{A}, while the interval $[0,\omega_{\max}]$ must cover the passband and transition bands of the analog filter.

5. REFERENCES

[1] S. K. Mitra, *Digital Signal Processing. A Computer–Based Approach,* McGraw–Hill, New York, NY, 2006.

[2] T. W. Parks and C. S. Burrus, *Digital Filter Design*, John Wiley, New York, 1987.

[3] G. I. Braileanu, "Extended–Window Interpolation Applied to Digital Filter Design," *IEEE Trans. Signal Process.*, vol. 44, pp. 457–472, 1996.

[4] G. I. Braileanu, "Equivalence between the Extended Window Design of IIR Filters and Least Squares Frequency Domain Designs," *Proc. 2003 Int. Conf. on Acoustics, Speech, and Signal Process.* (*ICASSP* 2003), Hong Kong, vol. VI, pp. 21–24, 2003.

[5] T. I. Laakso, V. Välimäki, M. Karjalainen, and U.K. Laine, "Splitting the Unit Delay," *IEEE Signal Processing Magazine*, vol. 13, pp. 30–60, 1996.

[6] G. I. Braileanu, "Digital Filters with Implicit Interpolated Output," *IEEE Trans. Signal Process.*, vol.45, pp.2551–2560, 1997.

[7] M. J. D. Powell, *Approximation Theory and Methods*, Cambridge Univ. Press, 1981.

[8] G. Meinardus, *Approximation of Functions. Theory and Numerical Methods*, Springer–Verlag, New York, 1967.

[9] J. R. Rice and K. H. Usow," The Lawson Algorithm and Extensions," *Math. Comp.*, vol. 22, pp. 118-127, 1968.