Correspondence

Comments on "Further Improvements in Digitizing Continuous-Time Filters"

Grigore Braileanu

***Abstract*—This correspondence points out that an algorithm proposed as a new digitizing method is, in fact, an extension of the linear interpolation design. Unlike previously published extensions, which incorporate the interpolation step into the s- to z-domain mapping step, this algorithm cascades the interpolator and the analog filter. Original design formulas derived for a rigorous comparison support the assertion that the above algorithm yields less accurate results than the other interpolation designs.**

***Index Terms*—Filter design and structures, sampling and interpolation, signal processing theory and methods.**

I. Introduction

A state-space algorithm, called "weighted sample" is presented in [1] as a new digitizing method. This algorithm implements a multiple-segment interpolator in a cascade defined in [1, Fig. 2], which is reproduced in Fig. 1 for convenience. Fig. 1 is a typical setup used, for example, by the step-invariance method (where S_C is a zero-order-hold element) and the linear interpolation method (where S_C is a two-point linear interpolator) [2]–[4]. The present correspondence points out that the above algorithm is, in fact, an immediate extension of the linear interpolation method from a two-point to an $(m+1)$-point interpolator S_C built with algebraic polynomials and cascaded with the analog filter $H(s)$. Next, it is shown that other digitizing options, better than the immediate extension, are available when the length m of the interpolation interval exceeds one. The most appealing of these options, which were previously proposed and analyzed in [5]–[7], are

i) the composition of the interpolation and filtering operators into just one operator by incorporating the interpolation step into the s- to z-domain mapping step (as opposed to the sequential approach [1], which cascades the interpolator and the filter);
ii) the causal and noncausal designs with optional additional poles at $z = 0$;
iii) the interpolation with trigonometric polynomials (as opposed to algebraic polynomials).

First of all, the algorithm proposed in [1] uses an algebraic polynomial interpolation to construct the output $x_C(t)$ of the m-segment interpolator along the current sampling period $kT \le t \le (k+1)T$ and applies this output sequentially to the input of the analog filter $H(s)$. Then, the output $y\{(k+1)T\}$ of the designed digital filter is derived from the response of the analog filter to $x_C(t)$ with an initial state at $t = kT$ equal to the state reached by the analog filter at the end of the previous computation done on $(k-1)T \le t \le kT$. Therefore, the resulting transfer function $H_D(z)$ is given by the impulse-response equivalent of the cascade $\{S_C, H(s)\}$, that is, the z transform of the sampled response $h_D(kT)$ of $\{S_C, H(s)\}$ to the input impulse $d(kT)$. For this reason, the "weighted sample" algorithm [1] will be referred to in the following as the *causal cascaded interpolation design* (CCID), where the causality ensues from the real-time type of calculation used in [1].

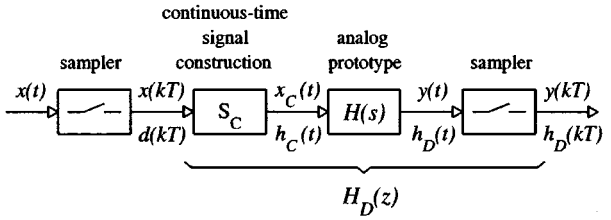

Fig. 1. Block diagram of the cascade-type interpolation design (CID).

The above mentioned alternative options i)–iii) stem from two results of the theory of interpolation at equally spaced points [8]. Specifically, the inherent oscillations of the interpolation error are usually exacerbated along the end segments of the interpolation interval, and at the same time, the interpolation with trigonometric polynomials enjoys much better numerical properties [8]; therefore, it should be preferred to the interpolation with algebraic polynomials [5]. Accordingly, trigonometric polynomials were used in [5]–[7] to calculate each new value $y\{(k+1)T\}$ from the response of the analog prototype $H(s)$ of order n_A determined along an *extended time segment* $(k-n_A+1)T \le t \le (k+1)T$ with a particular set of initial conditions. Specifically, along this time interval, the function $y(t)$, which is calculated at $t = (k-n_A+1)T, \ldots, kT$, must take on precisely the last n_A sampled values previously computed. Thus, this design, which is referred to as the *extended window design* (EWD), provides a natural match for the initial conditions of the analog and digital filters and incorporates the interpolation step into the s- to z-domain mapping step. Apparently, this makes the approximation $y(t)$ benefit from the increased accuracy of the central part of the m-segment interpolation interval that, usually, contains most of the signal energy [8]. This salient feature of the EWD is in contrast to the CCID, which uses only the right-end segment of the interpolation interval to calculate the current output of the digital equivalent. Finally, the problem created by the inaccurate end segments was addressed in [5]–[7] by the use of noncausal interpolators, which discard the end segments in the final stage of the design.

The main contribution of the present correspondence is a rigorous performance analysis of the existent digitizing methods based on interpolation with algebraic polynomials applied to causal, as well as noncausal, digital filter design. To this end, the paper addresses the fact that the CCID considered only causal algebraic interpolators, whereas the EWD has been presented in [5]–[7] under a very general and numerically stable algorithmic form that is more suitable for trigonometric interpolators. Thus, based on a unified approach, original closed-form design expressions that are suitable for the frequency domain analysis are derived below. As these expressions encompass both the causal and noncausal forms of design, the generic notations CID and EWD will be used hereafter. Finally, it is worth mentioning that the entire analysis is restricted to the interpolation with algebraic polynomials in order to i) provide the same basis for the performance evaluation of the CID and EWD and ii) point out that the latter yields more accurate digital equivalents to analog filters, even if algebraic polynomials are used instead of the trigonometric polynomials.

This correspondence is organized as follows. The general closed-form expressions of the CID and EWD transfer functions designed with interpolators that use algebraic polynomials are given in Section II. The

Manuscript received September 9, 1999; revised January 7, 2001. The associate editor coordinating the review of this paper and approving it for publication was Prof. Gregori Vazquez.

The author is with the Electrical Engineering Department, Gonzaga University, Spokane, WA 99258 USA.

Publisher Item Identifier S 1053-587X(01)07769-8.

performance analysis is done in Section III by analyzing the frequency responses of the analog filters and their digital equivalents. Concluding remarks are provided in Section IV.

II. Derivation of the CID and EWD Transfer Functions for Interpolation With Algebraic Polynomials

The following design of digital filters from analog prototypes is based on Fig. 1. For convenience, the continuous-time variable t will be normalized to the sampling period T; therefore, the sampling times will be designated by $t = k$, where k takes on integrals values. The continuous-time signals $x(t)$, $x_C(t)$, and $y(t)$ will be defined as in Fig. 1, and the input and output of the designed digital filter $H_D(z)$ will be denoted by $x_D(k)$ and $y_D(k)$, respectively. The orders of the analog and digital filters will be n_A and n_D. Moreover, the uppercase notations of functions of the variables s and z will, respectively, represent the Laplace and z-transforms of the corresponding signals denoted with lowercase letters.

A. General Time–Domain Invariant Synthesis

Given an analog filter $H(s)$, the traditional general time-domain invariant synthesis [4] generates a digital equivalent $H_D(z)$ such that equivalent inputs $x(t)$ and $x_D(k) = x(t)|_{t=k}$ yield equivalent outputs $y(t)$ and $y_D(k) = y(t)|_{t=k}$ along the entire time axis and under the *conditions of initial rest* [2, Sec. 3.5]. Thus, according to [4, Sec. 9-4], this time-domain equivalence implies the identity of the z-transforms $Y_D(z) \triangleq H_D(z)X_D(z)$ and $\mathcal{Z}\{Y(s)\}$. The latter expression uses the notation of [3, Sec. 2.4.1], where $\mathcal{Z}\{Y(s)\} \triangleq \mathcal{Z}\{[\mathcal{L}^{-1}[Y(s)]]|_{t=k}\}$ is the z-transform of the samples of $y(t)$. With this notation, the above identity gives the designed transfer function

$$H_D(z) = \frac{1}{X_D(z)}\mathcal{Z}\{Y(s)\}. \tag{1}$$

The crux of the design is the selection of the test signal $x(t)$ together with the method for calculating the corresponding response $y(t)$ of the analog prototype $H(s)$. Particular selections may produce the traditional impulse- and step-invariance methods, the linear interpolation method, or the two possible extensions of the latter—the CID and EWD—as shown below. At the same time, in order to satisfy the conditions of initial rest, the test signal applied to the interpolator S_C and the digital filter must be zero prior to some initial time t_0. Finally, the interpolation setup in Fig. 1 implies the identity $x_C(t) = x(t)$ for $t \geq t_0$. Therefore, in the context of time-invariant synthesis, the class of *admissible test signals* $x(t)$ corresponding to an $(m+1)$-point interpolation will be the set of polynomials—algebraic or trigonometric—whose zeros are m adjacent integers up to and including $t = t_0$. Restricting this correspondence only to algebraic polynomials, all the admissible test signals that satisfy the above requirements will be proportional to time-shifted variants of the generic polynomial

$$g(t) = \frac{1}{m!}t(t-1)(t-3)\ldots(t-m+1) \tag{2}$$

whose polynomial form leads to simple Laplace and z-transforms. For example

$$\begin{aligned} G(s) &\triangleq \mathcal{L}\{g(t)u(t)\} = \frac{P_m(s)}{m!s^{m+1}} \\ G_D(z) &\triangleq \mathcal{Z}\{[g(t)u(t)]|_{t=k}\} = \frac{z}{(z-1)^{m+1}} \end{aligned} \tag{3}$$

where $u(t)$ is the unit step, and $P_m(s)$ is a polynomial in s given by a straightforward calculation.

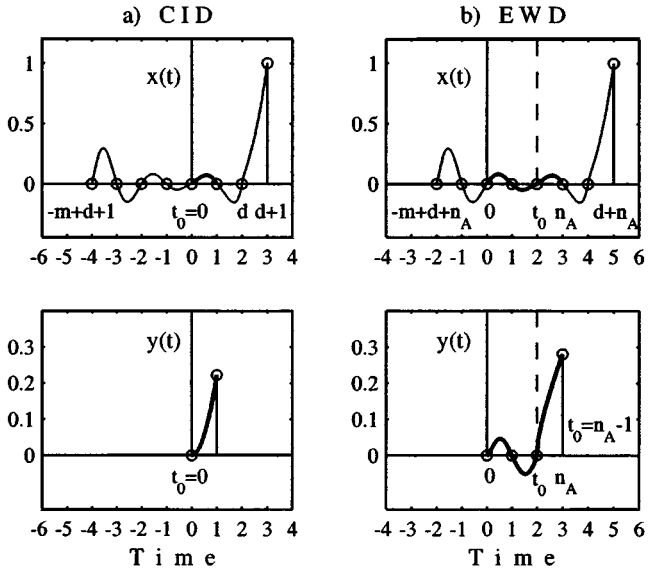

Fig. 2. Definition of the CID and EWD in terms of the test signals for a noncausal interpolation with $m = 7$, $d = 2$, $n_A = 3$ (the causal design is done with $d = 0$) (a) Current **CID** output $y(1)$ is computed from the *zero-state response* of the analog filter to the segment $x(t)$, $0 \leq t \leq 1$ of the $(m+1)$-point interpolation. (b) Current **EWD** output $y(n_A)$ is computed from the response of the analog filter to the segment $x(t)$, $0 \leq t \leq n_A$ with the auxiliary conditions $y(0) = y(1) = \cdots = y(n_A - 1) = 0$.

B. Definition of the CID and EWD Methods

Fig. 2(a) and (b) has the dual purpose of presenting the CID and EWD first and then defining the admissible test signals $x(t)$ necessary for the time-domain invariant synthesis based on the setup in Fig. 1.

In this section, Fig. 2 is used to present the CID and EWD recursive generation of the output; therefore, t_0 is the sampling time of the previously computed output, whereas $(t_0 + 1)$ is the current time. As the CID and EWD are variants of just one method—the interpolation design—they compute the next output sample in the same manner but with different initial conditions. Specifically, they compute the next output sample as the value at $t = t_0 + 1$ of the response of the analog filter $H(s)$ to the same signal $x_C(t)$ built by S_C from the last $(m+1)$ input samples on a time interval ending with $t = t_0 + d + 1$, that is, d sampling periods after the current time. Obviously, $d = 0$ corresponds to a causal design, whereas any positive integer d produces a digital filter whose causal implementation requires an additional output delay d. The *conceptual difference* between the CID and EWD stems from the way the initial conditions of the analog filter are defined. Thus, the CID takes the initial state of the filter $H(s)$ at t_0 equal to the final state during the calculation performed along the previous sampling period $t_0 - 1 \leq t \leq t_0$. Hence, only one segment of the interpolation interval is used (i.e., $t_0 \leq t \leq t_0 + 1$). By contrast, the EWD calculates each new value $y(t_0 + 1)$ from the response of the filter $H(s)$ determined along an extended time segment $t_0 - n_A + 1 \leq t \leq t_0 + 1$, whereas the function $y(t)$ takes on precisely the last n_A output values calculated at $t = t_0 - n_A + 1, \ldots, t_0 - 1, t_0$. The numerous options offered by the multiple-segment interpolation design are illustrated in Table I with reference to Fig. 3. At the same time, the main EWD features will be analyzed in Section III, together with the drawbacks of the CID, which are direct consequences of the single-segment sequential approach.

The following derivations of the CID and EWD transfer functions are based on the admissible test signals $x(t)$ that are necessary for the time-domain invariant synthesis applied to the setup in Fig. 1. To this end, according to Section II-A above, t_0 in Fig. 2 is viewed now as the time instant when the conditions of initial rest are applied. Apparently,

TABLE I
SIX VARIANTS OF THE DESIGN BASED ON $(m+1)$-POINT INTERPOLATORS ILLUSTRATED IN FIG. 3 FOR $m = 10$

	Name of design method	Type of interpolation polynomial	Domain of $y(t)$	n_A	n_D	Type of design
1	CID	algebraic	$-1 \le t \le 0$	10	19	Causal ($d = 0$)
2	A-EWD	algebraic	$-10 \le t \le 0$		10	
3	T-EWD	trigonometric	$-10 \le t \le 0$	10	10	
			$-8 \le t \le 0$	8		
4	CID	algebraic	$-3 \le t \le -2$	6	15	Noncausal ($d = 2$)
5	A-EWD	algebraic	$-8 \le t \le -2$		10	
6	T-EWD	trigonometric				

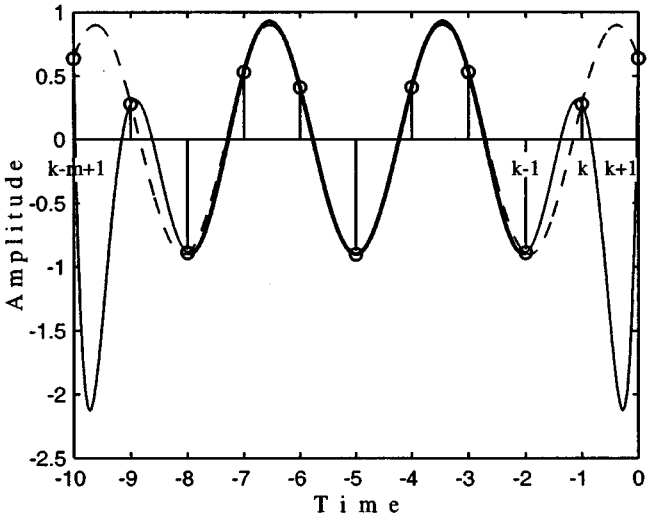

Fig. 3. Typical $(m+1)$-point interpolation of the input signal $(m = 10)$. *Dashed line*: $x(t)$—the original input. *Thin solid line*: $x_C(t)$—the entire interpolated signal. *Thick solid line*: the central part $-8 \le t \le -2$ of the interpolated signal considered to be "current input" by a noncausal design with $d = 2$. *Small circles*: the $(m+1)$ sampled values of the original input.

$t_0 = 0$ for the CID, and $t_0 = n_A - 1$ for the EWD, whereas the time origin shown in Fig. 2 is chosen such that all the signals considered in the remainder are zero along the negative side of the time axis.

C. Derivation of the CID Transfer Function

In the spirit of the general time-invariant synthesis, the CID transfer function $H_D(z) = H_{\mathrm{CID}}(z)$ is obtained by expressing (1) in terms of the signals $x(t)$ and $y(t)$ defined in Fig. 2(a). Accordingly, (1)–(3) yield the following expressions: $x(t) = g(t+m-d-1)u(t)$, $Y(s) = X(s)H(s)$, and

$$X(s) = \frac{P(s,m,d)}{m!s^{m+1}}, \quad X_D(z) = \frac{z^{m-d}}{(z-1)^{m+1}}$$
$$H_D(z) = H_{\mathrm{CID}}(z) = \frac{(z-1)^{m+1}}{z^{m-d}}\,\mathcal{Z}\left\{\frac{P(s,m,d)H(s)}{m!s^{m+1}}\right\}. \quad (4)$$

Moreover, the polynomial form of $x(t)$ makes the calculation of $X(s)$ straightforward, thus providing the coefficients of the polynomials $P(s,m,d)$. For the causal design (i.e., $d = 0$), the following relationships hold:

$$x(t) = \frac{1}{m!}t(t+1)(t+2)\ldots(t+m-1)u(t)$$
$$= (-1)^m g(-t)u(t)$$
$$P(s,m,0) = P_m(-s) \quad (5)$$

where the polynomials $P_m(s)$ have been defined in (3).

It is worth noting that the same expression (4) of $H_{\mathrm{CID}}(z)$, which was derived here as the generalized time-domain invariant equivalent of $H(s)$ alone, can be also derived as the *impulse-response invariant equivalent of the cascade* $\{S_C, H(s)\}$.

Finally, the CID characterization formulated in Proposition 1 below can be proven by considering (4), together with the total cancellation of the factors $(z-1)^{(m+1)}$ in $H_{\mathrm{CID}}(z)$.

Proposition 1: The causal implementation of the CID transfer function *increases the order of the digital filter* to $n_D = n_A + m - 1$, with $m-1$ poles at $z = 0$, and the nonzero poles $z_n = e^{s_n T}$, where the values s_n, $n = 1, 2, \ldots, n_A$ are the poles of the analog prototype.

D. Derivation of the A-EWD Transfer Function (Interpolation With Algebraic Polynomials)

While the CID is restricted to the interpolation with algebraic polynomials, two EWD forms corresponding to the interpolation with algebraic and trigonometric polynomials are available. They will be referred to as the A-EWD and T-EWD methods. In view of the facts that the T-EWD method was extensively analyzed in [7] and the CID was proposed in algebraic form, this correspondence is restricted to the interpolation with algebraic polynomials in order to provide the same basis for the performance evaluation of the CID and EWD. Thus, an expression similar to (4) is derived below for the particular A-EWD transfer functions designed with causal or noncausal algebraic interpolators.

Equation (1) is to be used here again for the particular forms of $x(t)$ and $y(t)$ corresponding to Fig. 2(b). Thus, $x(t) = g(t+m-L)u(t)$, where $L \triangleq n_A + d$. In addition, $X(s)$ and $XD(z)$ can be easily calculated as in (4) to give

$$X(s) = \frac{Q(s,m,L)}{m!s^{m+1}}, \quad L \triangleq n_A + d, \quad X_D(z) = \frac{z^{m+1-L}}{(z-1)^{m+1}} \quad (6)$$

where $Q(s,m,L)$ are polynomials in s and satisfy the following Proposition 2.

Proposition 2: If the EWD output interval $0 \le t \le n_A$ is centered within the input interval $-(m-L) \le t \le L$ as in Fig. 2(b) (i.e., always in the case of causal interpolation with $m = n_A$), then the following relations hold:

$$Q(s,m,L) = P(-s,\ m,\ d), \quad L \triangleq A + d = m - d$$
$$m = n_A + 2d. \quad (7)$$

Proof: For $L = m - d$, the particular form of the generic polynomial $g(t)$ defined in (2) generates the CID and EWD test signals $x_{\mathrm{CID}}(t) = g(t+L-1)u(t)$ and $x_{\mathrm{EWD}}(t) = g(t+m-L)u(t) = (-1)^m g(-t+L-1)u(t)$.Then, based on (4), (6), and $\mathcal{L}\{t^k u(t)\} = k!/s^{k+1}$, a straightforward algebraic manipulation completes the proof

$$X_{\mathrm{CID}}(s) = \mathcal{L}\{g(t+L-1)u(t)\} = \frac{P(s,m,d)}{m!s^{m+1}}$$
$$X_{\mathrm{EWD}}(s) = (-1)^m\mathcal{L}\{g(-t+L-1)u(t)\} = \frac{P(-s,m,d)}{m!s^{m+1}}$$
$$= \frac{Q(s,m,L)}{m!s^{m+1}}. \quad (8)$$

□

Next, let $H(s)$ be a transfer function with the denominator $a_0\, s^{n_A} + \cdots + a_{n_A}$ and poles s_n. According to Fig. 2(b), the general time-invariant synthesis applied to the EWD imposes the initial conditions $y(t) = 0, t = 0, 1, \ldots, n_A - 1$. Therefore, the expression of $Y(s)$ in (1) exhibits now a zero-state component $Y_{zs}(s)$ as well as a zero-input component $Y_{zi}(s)$ and can be written as $Y(s) = Y_{zs}(s) + Y_{zi}(s) = X(s)H(s) + Y_{zi}(s)$. In fact, $y_{zi}(t)$ is a linear combination of the modes of $H(s)$, and its coefficients are to be calculated as to satisfy the above initial conditions. The design formula (1) is implemented by using (6) and applying the operator $\mathcal{Z}$ to both components of $Y(s)$.This gives the following design equations, which are also illustrated in the Appendix:

$$\begin{aligned}\mathcal{Z}\{X(s)H(s)\} &= \mathcal{Z}\left\{\frac{Q(s,m,L)H(s)}{m!s^{m+1}}\right\}\\ &\triangleq \mathcal{Z}\left\{\frac{b_0 s^{n_A+m-1} + \cdots + b_{n_A+m-1}}{s^{m+1}(a_0 s^{n_A} + \cdots + a_{n_A})}\right\}\\ &\triangleq \frac{z(\beta_0 z^{n_A+m} + \cdots + \beta_{n_A+m})}{(z-1)^{m+1}(\alpha_0 z^{n_A} + \cdots + \alpha_{n_A})}\end{aligned} \tag{9a}$$

$$\begin{aligned}H_D(z) &= H_{\mathrm{EWD}}(z)\\ &= \frac{(z-1)^{m+1}}{z^{m+1-L}}\left[\mathcal{Z}\{X(s)H(s)\} + \mathcal{Z}\{Y_{zi}(s)\}\right]\\ &= \frac{(z-1)^{m+1}}{z^{m+1-L}}\\ &\quad\cdot\left[\frac{z(\beta_0 z^{n_A+m} + \cdots + \beta_{n_A+m})}{(z-1)^{m+1}(\alpha_0 z^{n_A} + \cdots + \alpha_{n_A})}\right.\\ &\quad\left. + \frac{z(c_1 z^{n_A-1} + \cdots + c_{n_A})}{\alpha_0 z^{n_A} + \cdots + \alpha_{n_A}}\right]\end{aligned} \tag{9b}$$

$$\begin{aligned}(z-1)^{m+1} &\triangleq z^{m+1} + d_1 z^m + \cdots + d_{m+1}\\ c_1 &= -\beta_0\\ \left\{c_k\right. &= \left.-\beta_{k-1} - \sum_{n=1}^{k-1} c_{k-n} d_n, k = 2, \ldots, n_A\right\}.\end{aligned} \tag{9c}$$

Apparently, the expression $\mathcal{Z}\{Y_{zi}(s)\}$ is a ratio of two polynomials of same degree n_A; its poles are $z_n = e^{s_n T}$, and the numerator coefficients c_n are chosen such that $\lim_{Z\to\infty}\{z^{n_A-1}\mathcal{Z}\{Y(s)\}\} = 0$, thus satisfying the initial conditions. This is equivalent to the recursive equations above, which yield the coefficients c_k such that the n_A leading terms of the numerator of $H_{\mathrm{EWD}}(z)$ in (9) are canceled. At the same time, the factors $(z-1)^{m+1}$ are canceled altogether.

It is worth noting the excellent accuracy of the numerical calculation of $\mathcal{Z}\{X(s)H(s)\}$, which stems from the fact that, according to a good design practice, $H(s)$ is usually already defined in factored form. Thus, knowing the poles of $H(s)$, the method of residues offers closed-form expressions for the accurate computation of all partial fraction expansion coefficients, including those of the multiple pole at $s = 0$. This approach gives $H_D(z)$ directly in a form suitable for finite-word arithmetic implementation, corresponding to cascade or parallel realizations. At the same time, the obvious pole-zero cancellation involving the factors $(z-1)^{m+1}$ is dealt with during this easy-to-program algebraic process. The algebraic analysis of (9) yields the following.

Proposition 3: The causal implementation of the EWD transfer function is of order $n_D = m \geq n_A + d$, with n_A poles related to the poles s_n of the analog prototype by the expression $z_n = e^{s_n T}, n = 1, 2, \ldots, n_A$, whereas the remaining poles, if any, are placed at $z = 0$. In particular, the causal design with $d = 0$ and $m = n_A$ generates a *digital filter of exactly the same order as the analog prototype*.

III. Performance Evaluation

The CID and A-EWD will be evaluated in terms of the digitizing error expressed in decibels $E_D(\omega) = 20\log_{10}(|H(j\omega) - H_D(e^{j\omega})|)$, $0 \leq \omega \leq \omega_f = \pi$, where $H_D(e^{j\omega}) \triangleq H_D(z)\big|_{z=e^{j\omega}}$ is the frequency response of the designed digital filter, and ω_f is the folding frequency (i.e., half the sampling frequency).

First of all, a byproduct of Sections II-C and D is the derivation of $H_C(j\omega)$ defined as the Fourier transform of the unit impulse response of S_C, which is shown in Fig. 1. Thus, assuming $H(s) = 1$ and using the fact that the rational terms in z found in (4) and (9) implement a summations of weighted time shifts, the following expressions are obtained:

$$\begin{aligned}H_C(j\omega) &= e^{j\tau\omega}(1 - e^{-j\omega})^{m+1}\frac{p(j\omega)}{m!(j\omega)^{m+1}}\\ &\begin{cases}p(j\omega) = P(j\omega, m, d), & \tau = d+1, \text{ for CID}\\ p(j\omega) = Q(j\omega, m, L), & \tau = L, \text{ for A-EWD.}\end{cases}\end{aligned} \tag{10}$$

In particular, for $m = 1$ and $d = 0$, either expression of $H_C(j\omega)$ will give the well-known frequency response of the linear interpolator (see, e.g., either [2, Sec. 8.2] or [4, Sec. 8.2]). Typical plots for $20\log_{10}(|H_C(j\omega)|)$ and the corresponding interpolation error $E_I(\omega) = 20\log_{10}(|1 - H_C(j\omega)|)$ are shown in Figs. 4 and 5. In addition, by applying the relationships (7) of Proposition 2 to (10), an important result is obtained in the form of Proposition 4.

Proposition 4: If the EWD output interval $0 \leq t \leq n_A$ is centered within the input interval $-(m - L) \leq t \leq L$, then $H_{C_{\mathrm{CID}}}(j\omega) = H^*_{C_{\mathrm{EWD}}}(j\omega)$, where $*$ denotes the conjugate. As a consequence, the magnitude responses $|H_C(j\omega)|$, as well as the deviations from the unity gain $|1 - H_C(j\omega)|$, are identical for the CID and A-EWD interpolators when $m = n_A + 2d$ and $L = m - d$. Note that this always holds for causal interpolation (i.e., $d = 0$), which corresponds to an A-EWD computed with $m = L = n_A$.

The strong distortional effect of the CID, evidenced by the plots of $H_C(j\omega)$ shown in Figs. 4 and 5, is a direct consequence of (4) and (10). Specifically, for any digital filter, $H_D(e^{jw})$ is a periodic function of ω with period $\omega_s = 2\pi$ and, when obtained with the CID, it can be expressed as

$$H_D(e^{j\omega}) = H_C(j\omega)H(j\omega), \quad -\pi < \omega < \pi \tag{11}$$

under the assumption of negligible aliasing. Equation (11) clearly shows that any interpolation device that is cascaded with the analog prototype will irreversibly distort the frequency response. This multiplicative distortion appears strictly additive in the decibel plots shown in Figs. 6–8 (except for the interval $0.9\omega_f \leq \omega \leq \omega_f$ affected by some aliasing). Indeed, the distorted magnitude response of a CID filter designed with $m = 10$, represented by the thin solid line in Fig. 6, is the sum of the dashed line in Fig. 6 and the thin solid line in Fig. 4, that is, $|H_D(e^{j\omega})|_{\mathrm{dB}} = |H(j\omega)|_{\mathrm{dB}} + |H_C(j\omega)|_{\mathrm{dB}}$. Likewise, in Fig. 7, the corresponding digitizing error $E_D(\omega) = 20\log_{10}(|H(j\omega) - H_D(e^{j\omega})|)$ represented by the thin solid line coincides with the dash-dotted line, which is the sum of the dashed line in Fig. 6 and the thin solid line in Fig. 5, that is, $E_D(\omega) = |H(j\omega) - H_C(j\omega)H(j\omega)|_{\mathrm{dB}} = |H(j\omega)|_{\mathrm{dB}} + |1 - H_C(j\omega)|_{\mathrm{dB}}$.

By contrast, the EWD example of Tables A-II and A-III in the Appendix is shown in Figs. 6–8 to be dramatically more accurate and free of the distortional effect of (11), in spite of Proposition 4. For example, the large deviations of the plots in Fig. 5 in the region $0.7\omega_f \leq \omega \leq \omega_f$

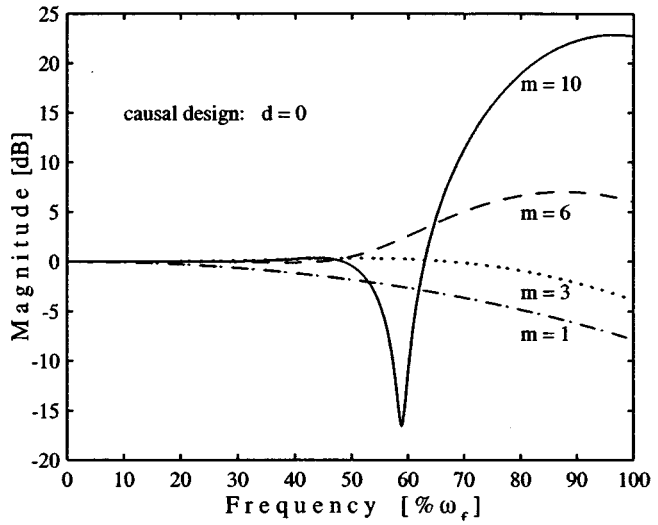

Fig. 4. Magnitude of $H_C(j\omega)$ for the interpolator S_C in Fig. 1 built with polynomials of degrees $m = 1, 3, 6$, and 10.

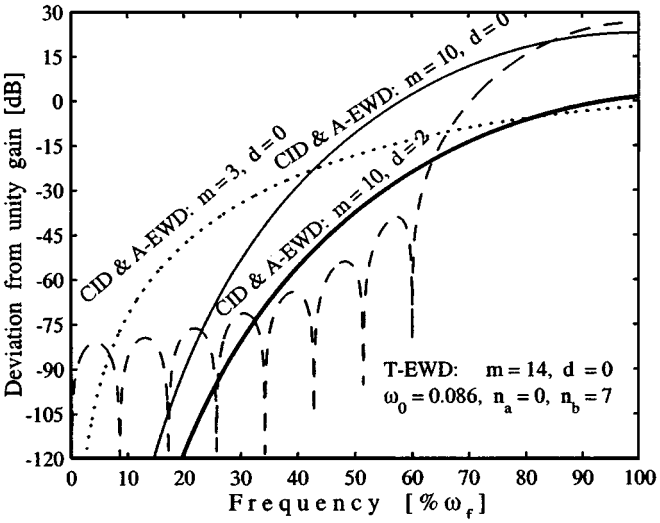

Fig. 5. Deviation of $H_C(j\omega)$ from the (ideal) unity gain for a few algebraic and trigonometric interpolators: $E_I(\omega) = 20\log_{10}(|1 - H_C(j\omega)|)$. This figure supports the use of the noncausal A-EWD (e.g., with $d = 2$) and the more computationally expensive T-EWD (the T-EWD parameters ω_0, n_a, and n_b were defined in Algorithm 1 of [7]).

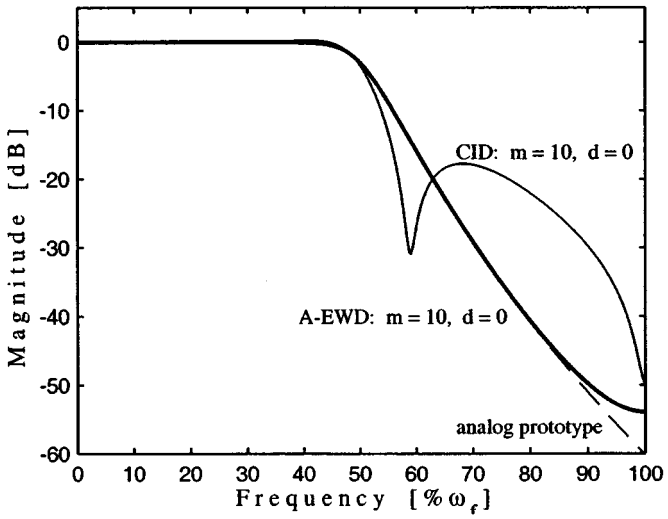

Fig. 6. Magnitude responses of digital equivalents of a wideband lowpass Butterworth filter of passband [0, 1] and order 10 ($\omega_f = 2$). *Thin solid line*: CID filter of order $n_D = 19$. *Thick solid line*: A-EWD filter of order $n_D = 10$. *Illustration of the distortional effect of the CID method*: Both filters are designed with the same interpolator of degree $m = 10$ in Figs. 4 and 5, yet only the CID log-magnitude plot in decibels is additively distorted with the large deviations of the plot in Fig. 5 in the region $0.7\omega_f \leq \omega \leq \omega_f$.

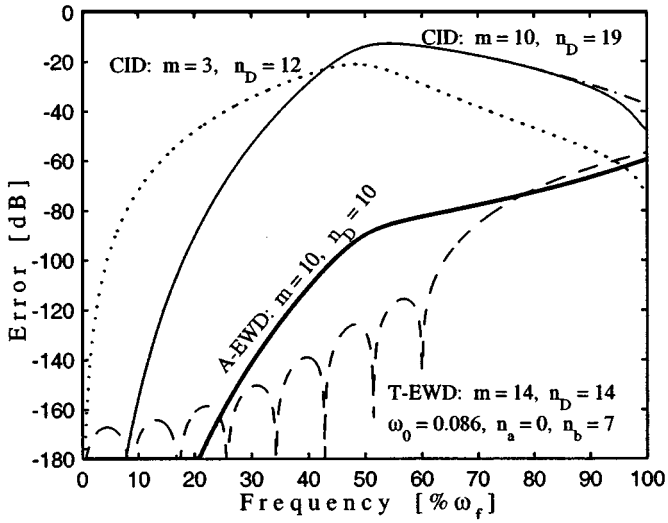

Fig. 7. Magnitude of digitizing errors $E_D(\omega) = 20\log_{10}(|H(j\omega) - HD(e^{j\omega})|)$ corresponding to various causal digital filter implementations of the wideband lowpass filter in Fig. 6. *Illustration of the distortional effect of the CID method*: The error $E_D(\omega)$ of the CID filter designed with $m = 10$, which is represented by the *thin solid line*, coincides with the expression $20\log_{10}(|H(j\omega) - H_C(j\omega)H(j\omega)|)$, which is represented by the dash-dotted line.

do not yield additive alterations in the EWD plots shown in Figs. 6 and 7. Indeed, while Proposition 4 makes both the CID and A-EWD have the same magnitude of the corresponding expression (11), the latter does not represent the frequency response of the A-EWD filter due to the additional terms containing the coefficients c_k in (9). This somewhat surprising property of the EWD interpolators is due to what may be called a symbiotic composition of the interpolation and analog-filtering operators. In addition, considering the causal design (i.e., $d = 0$), it is obvious that although EWD requires $m \geq n_A + d$, it suffices to choose $m = n_A$ for the order of $H_D(z)$ to remain exactly $n_D = n_A$ (i.e., equal to the number of imposed initial conditions). This salient feature of the above symbiosis, which is exemplified in Fig. 7 with the A-EWD digitizing error, indeed suggests a "blending" of the interpolator S_C and the analog prototype $H(s)$, as opposed to the CID where the cascaded S_C increases the order of $H_D(z)$ to $n_D = n_A + m - 1$.

The above properties, which are typical of the CID and EWD, can be given a simple qualitative explanation by using Fig. 3 and Table I, where a large value of m was chosen ($m = 10$) for the sake of a clear graphical illustration. The key to the entire discussion is the property of the interpolation at equally spaced points, which yields a relatively accurate central interval but produces large errors along the end segments [8]. Thus, the poor behavior of CID filters can be explained by the fact that the current $y(k+1)$ is calculated from the response $y(t)$ to the highly inaccurate end segment of the interpolation (i.e., $-1 \leq t \leq 0$ in Fig. 3). At the same time, the EWD calculates each value $y(k+1)$ from the response $y(t)$ to the entire interpolated input $x_C(t)$ on $-10 \leq t \leq 0$. Hence, most of the energy of $x_C(t)$ corresponds to the accurate central part ($-8 \leq t \leq -2$) whereas, usually, the recent values of the input have little effect on the output sample $y(k+1)$. This reasoning also explains why noncausal EWD designs

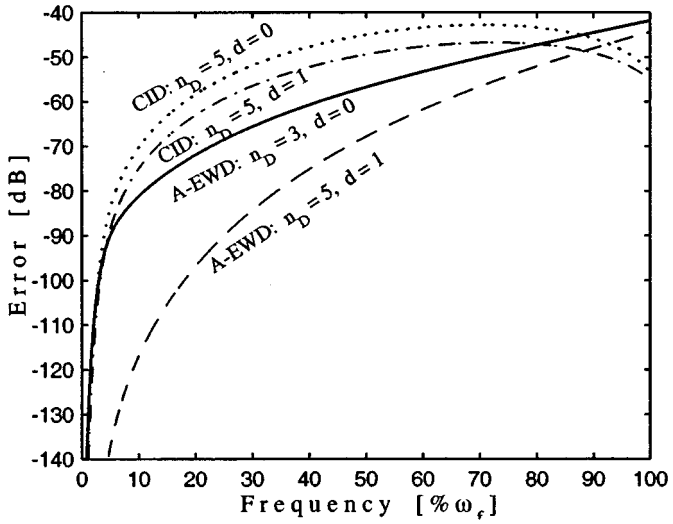

Fig. 8. Magnitude of digitizing errors for $H(s) = (20s + 2)/[(s + 1)(s^2 + 2s + 2)]$, $\omega_f = 40$. The same interpolator ($m = 3$) produces the CID filters of order $n_D = 5$ and the causal A-EWD of order $n_D = 3$. Note that the noncausal design with $d = 1$ makes the A-EWD error decrease more substantially than the CID error.

TABLE II
COEFFICIENTS OF THE A-EWD EQUATIONS (9) FOR THE FILTER IN FIG. 8

	Coefficients of the Intermediate Numerator		Coefficients of the Causal A-EWD Digital Filter		
			Denominator		Numerator
z^6	β_0	0.			
z^5	β_1	0.0025520215			
z^4	β_2	-0.0129490536			
z^3	β_3	0.0365145364	α_0	1.	0.010238537
z^2	β_4	0.0549197859	α_1	-2.767696115	0.0815736775
z	β_5	-0.0693503881	α_2	2.558638882	-0.08286628
1	β_6	-0.0108254181	α_3	-0.790081283	-0.0080844505

like 5 and 6 in Table I increase the accuracy even more dramatically than their CID counterparts (see Figs. 5 and 8). Moreover, the second design 3 of Table I,which has been done with $n_A = 8$ and $m = 10$, reduces the digitizing error by discarding the erroneous left-end interpolation segments (e.g., see the T-EWD plot in Fig. 7, with $n_A = 10$ and $m = 14$). Finally, the analysis of the better approach—the T-EWD, which is based on frequency selective interpolators [6], [7]—is beyond the scope of this paper.

IV. CONCLUSIONS

A unified approach was developed for the interpolation based design of digital filters. Original closed- form design formulas that were suitable for a rigorous frequency domain analysis, were derived for two conceptually opposed options. Specifically, (4) and (9) were derived for the CID and EWD, respectively, and were given general forms that encompass both causal and noncausal designs. Moreover, it was shown that the algorithm proposed in [1] as a new digitizing method is a cascade-type extension of the linear interpolation design, and in fact, it addresses only the *causal* CID. Conceptual arguments, as well as a quantitative analysis, supported the conclusion that this extension is less accurate than the other existent extensions of the linear interpolation design.

First of all, in the case of the causal design ($d = 0$), (4) yields the same result as the state-space algorithm of [1]. In particular, the design done with $m = 1$, $d = 0$, and (4) gives the transfer function obtained with the linear interpolation method [2, Sec. 8.2 and 8.4] or [3, Sec. 4.4.2]. This can be illustrated with the transfer function $F(s) = 4/(s + 4)$ used in [1] to prove that the proposed state-space algorithm was not the same as the zero-order hold method. Indeed, the transfer function $F_1(z) = (0.1418z + 0.1278)/(z - 0.7304)$ (i.e., [1, Eq. (45)]) obtained for the sampling period $T = \pi/40$ is precisely $H_D(z)$ given by the linear interpolation method, as well as (4) above with $m = 1$, $d = 0$, and appropriate time-domain denormalization.

Second, the argument was made in this correspondence that the EWD, which was previously developed in [5]–[7] and briefly analyzed in Sections II and III, circumvents the drawbacks of the cascade approach. Indeed, in all the examples shown in Figs. 6–8, the A-EWD exhibits better performance, even in the case of narrowband lowpass filters (see the example in Fig. 8, which illustrates the performance of a few digital equivalents of the narrowband lowpass filter of [1, Eq. (46)]). Moreover, unlike the CID equivalents, whose orders exceed the order n_A of the analog filter by $(m - 1)$, EWD designs done with $m = n_A$ preserve the order n_A.

Third, the analysis of Section III underlined the importance of the noncausal design. Indeed, it is usually enough to choose $d = 1$, or $d = 2$, to dramatically reduce the digitizing error. The inherent additional d-step delay (which is necessary for the real-time implementations of the resulting filters) is fully acceptable in most applications, with the possible exception of those cases when the digital filters are placed in feedback loops.

Finally, while the main focus was on the A-EWD in order to provide the same basis for the comparison with the CID, it is worth mentioning that the advantages of frequency-selective interpolators [6], [7], which are typical of the T-EWD, make the T-EWD the method of choice. Nevertheless, the A-EWD is a convenient alternative for the design of causal low-order lowpass filters in environments characterized by significant aliasing.

APPENDIX
NUMERICAL EXAMPLES

The implementation of the A-EWD equations (9) is illustrated here for the transfer function used in Fig. 8, $H(s) = (20s + 2)/[(s+1)(s^2+2s+2)]$. A causal three-segment polynomial interpolator of degree $m = 3$ yields both the CID filter of order $n_D = 5$ and the A-EWD filter of order $n_D = n_A = 3$ designed with $L = 3$ and $d = 0$. The denormalization of the design equations required by the actual sampling period $T = \pi/40$ can be incorporated into the polynomials $P(s, m, d)$ and $Q(s, m, L)$ of (4), (6), and (9), which, consistently with (7), become $Q(s, 3, 3) = P(-s, 3, 0) = (2T^2s^2 - 6Ts + 6)/T^3$. The first step of the EWD design generates the coefficients $\beta_0, \ldots, \beta_6$ and $\alpha_0, \ldots, \alpha_3$ of the intermediate transfer function $\mathcal{Z}\{X(s)H(s)\}$ in (9), that is, the z-transform of the inverse Laplace transform of $Q(s, m, L)H(s)/(m!s^{m+1})$. The coefficients are shown in Table II. The recursive equations that give the coefficients c_k of (9) are $c_1 = -\beta_0, c_2 = -\beta_1 + 4c_1$, and $c_3 = -\beta_2 + 4c_2 - 6c_1$. According to Section II-D, these equations calculate the coefficients c_k such that the n_A leading terms of the numerator of $H_{\mathrm{EWD}}(z)$ in (9) are canceled, thus satisfying the EWD initial conditions. It can be proven that the relation $\beta_0 = 0$ always holds; therefore, the coefficients c_k become $c_1 = 0$, $c_2 = -0.002\,552\,021\,5$, and $c_3 = 0.002\,740\,967\,6$. Moreover, there will be total cancellation of the factors $(z - 1)^{m+1}$ in (9), and $z^{m+1-L} = z$ since the minimum-order causal A-EWD

TABLE III
BUTTERWORTH ANALOG FILTER OF ORDER 10 IN FIGS. 6 AND 7

Factors of the Denominator		
s^2	s	1
1.	0.3128689	1.
1.	0.907981	1.
1.	1.414214	1.
1.	1.782013	1.
1.	1.975377	1.

TABLE IV
A-EWD DIGITAL FILTER OF ORDER 10 DESIGNED FROM THE FILTER IN TABLE III

Zeros of the A-EWD Filter (factor: 0.226329246e-05)		Poles of the A-EWD Filter	
Real Part	Imaginary Part	Real Part	Imaginary Part
-6.73605469			
-1.63016176			
-0.58451141		0.0151248676	±0.78198973
-0.26054507		0.0835006202	±0.482943981
-9.44924093	±76.7326155	0.146224024	±0.295078168
-0.0631632673	±0.044451201	0.186582771	±0.16139036
0.0172102122	±0.023798308	0.205572464	±0.051556184

implies $L = m = 3$. Thus, the transfer function of the digital filter will have the numerator and the denominator of same degree $n_D = m = 3$. The coefficients of the numerator are shown in Table II, together with the coefficients $\alpha_0, \ldots, \alpha_3$ of the denominator.

Next, the transfer function of the analog Butterworth filter of order 10 used in the examples shown in Figs. 6 and 7 is considered. The numerator is 1, and the denominator is given in factored form in Table III. The design equations (9) yield a minimum-order A-EWD digital equivalent of this filter, which is represented in Table IV in pole-zero form.

REFERENCES

[1] C. Wan and A. M. Schneider, "Further improvements in digitizing continuous-time filters," *IEEE Trans. Signal Processing*, vol. 45, pp. 533–542, Mar. 1997.

[2] A. V. Oppenheim, A. S. Willsky, and I. T. Young, *Signals and Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1983.

[3] G. F. Franklin, J. D. Powell, and M. L. Workman, *Digital Control of Dynamic Systems*. New York: Addison-Wesley, 1990.

[4] R. E. Ziemer, W. H. Tranter, and D. R. Fannin, *Signals and Systems: Continuous and Discrete*. New York: Macmillan, 1993.

[5] G. Braileanu, "Matrix operators for numerically stable representation of stiff, linear dynamic systems," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 974–980, Aug. 1990.

[6] G. Braileanu, "A new approach to the digitizing of continuous-time systems with application to interpolation and digital filter design," in *Proc. Third Int. Conf. Adv. Commun. Contr. Syst.*, Victoria, BC, Canada, Oct. 1991, pp. 645–656.

[7] G. Braileanu, "Extended-window interpolation applied to digital filter design," *IEEE Trans. Signal Processing*, vol. 44, pp. 457–472, Mar. 1996.

[8] M. J. D. Powell, *Approximation Theory and Methods*. London, U.K.: Cambridge Univ. Press, 1981.

Reply to Comments on "Further Improvements in Digitizing Continuous-Time Filters"

Alan M. Schneider and Chunxi Wan

Dr. Braileanu compares his A-EWD design technique with our WS technique (which he calls CID) and cites the relative errors of the two techniques as a function of frequency. The error plot he presents in his Fig. 8, which we have redrawn as Fig. 1, illustrates that the two techniques exhibit similar small errors in the filter passband but differ in the filter stopband. The magnitude of the errors for both techniques is on the order of $10^{(-2)}$ to $10^{(-7)}$. In many practical applications, the energy content in the stopband of a system contributes an insignificant level to the time signal passing through the filter. Small errors in insignificant contributions represent third- and fourth-order effects and are of little practical importance. In reality, analog filters built with real components have a tolerance on the order of one part in 100, and the spectral errors resulting from component tolerance spread far outweigh the simulation differences reported by Dr. Braileanu. Further, comparing the size of low-level errors in the frequency-domain description of the different filter design methods begs the real question: What effect do the errors cause in the fidelity of the time-domain signals? In reality, the time-domain response of different filter implementations is remarkably tolerant of small changes in the filter's frequency response. This tolerance is the result of the temporal averaging inherent in the convolution process.

Fig. 1 is a replot of two curves of Dr. Braileanu's Fig. 8: the CID $nd = 5, d = 0$ curve (our WS design) and the A-EWD $nd = 3, d = 0$ curve. The only change is the use of the logarithmic scale for frequency for the purpose of placing the passband into broader perspective. The bandwidth of the prototype analog filter for this plot is 5 rad/s. The sampling rate is 80 rad/s, which is 16 times the bandwidth. This plot shows that the error of either filter in the passband ranges from $10^{(-3)}$ to $10^{(-7)}$. The error in the stopband ranges from $10^{(-2)}$ to $10^{(-3)}$.

The tenth-order Butterworth filter Dr. Braileanu uses in his comparison (Fig. 6) was designed with a small ratio of sampling frequency to bandwidth (4:1 to be precise), or to put it another way, the sampling rate is twice the Nyquist limit. In many signal processing applications, it is indeed desirable to operate close to the Nyquist limit, and this is possible because the signal is preconditioned while in its analog form prior to being sampled. The examples we cite represent signals that are oversampled (sampling rate to bandwidth) by factors of 10 to 20 since they are signals in a control system that have not been preconditioned by an analog prefilter. Again, the significant region over which the filters should be compared is the passband. The oversampled passband represents a small part of the total spectrum of the input signal; in this region, both techniques exhibit very small errors.

The realm of application of causal and noncausal filter differs. Causal filters are used in real-time applications. The fourth paragraph of IV. Conclusions agrees. It says "the additional d-step delay (in noncausal filters) is fully acceptable in most applications, possibly with the exception of those cases when the filter is placed in feedback

Manuscript received April 24, 2001; revised June 6, 2001. The associate editor coordinating the review of this paper and approving it for publication was Prof. Gregori Vasquez.

A. M. Schneider is with the Department of Mechanical and Aerospace Engineering, University of California at San Diego, La Jolla, CA 92093-0411 USA.

C. Wan is with the Beijing Institute of Technology, Beijing, China.

Publisher Item Identifier S 1053-587X(01)08242-3.