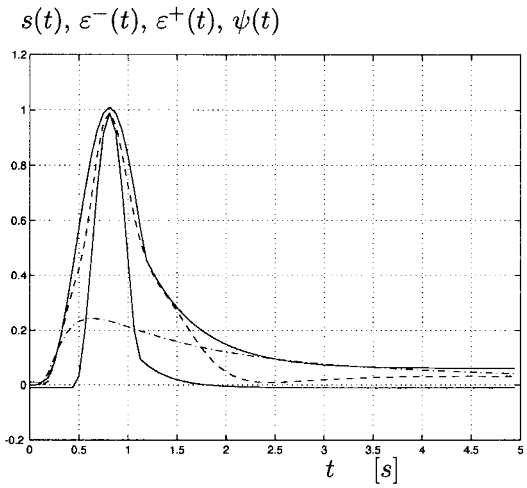

Fig. 3. Input signal s (dash-dotted line), specified output envelope ε (solid line), and resulting noiseless output signal ψ (dashed line).

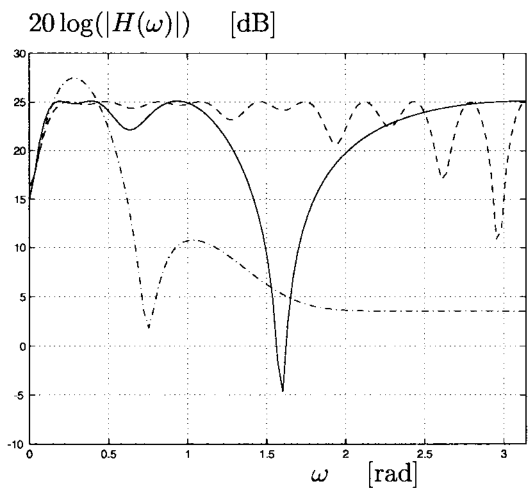

Fig. 4. Resulting frequency response of equalizer for an L_∞-design (solid line) and an L_2-design (dashed line).

normalized frequency of $1/T$ as input and produces an output that lies within the envelope given by the DSX-3 pulse template [1], [3].

Fig. 3 shows the coaxial cable impulse response s, the output envelope ε, and the resulting noiseless output signal ψ. The solution corresponds to an L_∞-design with $N = 8$ coefficients obtained with the semi-infinite simplex procedure. Fig. 4 shows the resulting frequency response of the equalizer designed when using both L_∞ and L_2 design techniques. Fig. 4 also shows a comparison with the common FIR filter $(a = 0, N = 20)$, where the order N has been chosen to match the L_∞-norm corresponding to the L_∞-optimal Laguerre filter of order 8.

V. Conclusions

The envelope-constrained (EC) IIR filter design problem has been formulated as a special case of a general frequency domain L_∞ optimization problem. The optimization problem is cast as a semi-infinite linear program that can be solved by using numerically efficient simplex extension algorithms. An orthonormal Laguerre network is used as an example of a recursive filter offering a low-order alternative to the conventional FIR filter. A numerical example concerning the equalization of a digital transmission channel is included to demonstrate the efficiency of the design method.

References

[1] CCITT, "Physical/electrical characteristic of hierarchical digital interfaces," G.703, Fascicle III, 1991.

[2] Z. Zang, B. Vo, A. Cantoni, and K. L. Teo, "Applications of discrete-time Laguerre networks to envelope constrained filter design," in *Proc. ICASSP*, May 1996, vol. 3, pp. 1363–1366.

[3] Bell Commun., "DSX-3 isolated pulse template and equations," Tech. Ref. TR-TSY-000499, issue 2, pp. 9–17, Dec. 1988.

[4] J. W. Lechleider, "A new interpolation theorem with application to pulse transmission," *IEEE Trans. Commun.*, vol. 39, pp. 1438–1444, 1991.

[5] E. J. Anderson and P. Nash, *Linear Programming in Infinite-Dimensional Spaces*. New York: Wiley, 1987.

[6] D. Burnside and T. W. Parks, "Optimal design of FIR filters with the complex Chebyshev error criteria," *IEEE Trans. Signal Processing*, vol. 43, pp. 605–616, Mar. 1995.

[7] M. A. Masnadi-Shirazi and N. Ahmed, "Optimum Laguerre networks for a class of discrete-time systems," *IEEE Trans. Signal Processing*, vol. 39, pp. 2104–2108, Sept. 1991.

[8] Z. Zang, A. Cantoni, and K. L. Teo, "Continuous-time envelope-constrained filter design via Laguerre filters and $\mathcal{H}_\infty$ optimization methods," in *Proc. ICASSP*, May 1997, vol. 1, pp. 55–58.

[9] X. Chen and T. W. Parks, "Design of FIR filters in the complex domain," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 144–153, Feb. 1987.

[10] D. G. Luenberger, *Linear and Nonlinear Programming*. Reading, MA: Addison-Wesley, 1984.

Comments on "A Weighted Least-Squares Method for the Design of Stable 1-D and 2-D IIR Digital Filters"

Phillip A. Regalia

***Abstract*—A recent iterative procedure by Lu *et al.* for designing IIR filters is recognized as the Sanathanan–Koerner or Steiglitz–McBride iteration, which was proposed originally in 1963 and 1965, respectively, for system identification purposes. We re-examine some claims and issues related to Lu *et al.*'s paper[1] in view of various properties that have been deduced for the Steiglitz–McBride iteration in the identification literature. In particular, the Steiglitz–McBride procedure is known not to minimize a least-squares distance measure, contrary to Lu *et al.*'s claim. It can, however, provide good models and remain stable during successive iterations under specific conditions which are reviewed within.**

I. Introduction

In a recent paper [1], Lu *et al.* proposed an iterative least-squares procedure for designing IIR filters. In the one-dimensional (1-D) case, the iterative procedure seeks nth-order polynomials $D_k(z)$ and $N_k(z)$, with $D_k(z)$ monic, which minimize

$$J_k = \int_0^\pi \frac{W(\omega)}{|D_{k-1}(e^{j\omega})|^2} \left|F_d(\omega)D_k(e^{j\omega}) - N_k(e^{j\omega})\right|^2 d\omega \qquad (1)$$

[1] W. S. Lu *et al.*, *IEEE Trans. Signal Processing*, vol. 46, pp. 1–10, Jan. 1998.

Manuscript received March 30, 1998; revised June 19, 1998. The associate editor coordinating the review of this paper and approving it for publication was Dr. Ali H. Sayed.

The author is with the Département Signal et Image, Institut National des Télécommunications, Evry, France (e-mail: regalia@int-evry.fr).

Publisher Item Identifier S 1053-587X(99)04658-9.

in which

- $F_d(\omega)$ desired frequency response of the filter to be designed;
- $W(\omega)$ weighting function;
- $D_{k-1}(z)$ monic polynomial determined from the procedure at iteration $k-1$.

From an initial "denominator" polynomial $D_0(z)$, minimizing successively $J_1, J_2, J_3, \cdots$ gives rise to an iterative procedure in which convergence, if it occurs, takes the form of $D_k(z)$ approaching a limit polynomial $D_\infty(z)$ as k increases. At any convergent point, yielding, thus, $D_k(z) = D_{k-1}(z) = D_\infty(z)$, the criterion $J_k = J_\infty$ appears as

$$J_\infty = \int_0^\pi W(\omega)\left|F_d(\omega) - \frac{N_\infty(e^{j\omega})}{D_\infty(e^{j\omega})}\right|^2 d\omega \tag{2}$$

which is a weighted L_2-norm of the error function $F_d - N/D$.

This iterative procedure is precisely that proposed in 1963 by Sanathanan and Koerner [2] using frequency response measurements and in 1965 by Steiglitz and McBride [3], using time domain measurements, both in a system identification context. In that context, $F_d(\omega)$ represents the frequency response of an unknown system, and $W(\omega)$ is the power spectral density function of the input process driving the unknown system; this function is constant in Sanathanan and Koerner. Various properties of the Steiglitz–McBride iteration have been deduced over the years, and the intent of this correspondence is re-examine some of the issues raised by Lu *et al.* in view of some known results on this procedure.

II. On the Desired Response

In order to draw more closely on the similarities with system identification, we assume that $F_d(\omega)$ admits a convergent Fourier series expansion

$$F_d(\omega) = \sum_{k=-\infty}^{+\infty} f_k\, e^{-jk\omega}, \qquad \text{with} \sum_{k=-\infty}^{+\infty} |f_k|^2 < \infty.$$

We assume, moreover, that the coefficients $\{f_k\}$ in this series expansion are real.

Split $F_d(\omega)$ into its "causal" and "strictly anticausal" parts as

$$F_+(\omega) = f_0 + f_1 e^{-j\omega} + f_2 e^{-j2\omega} + \cdots \quad \text{(causal)}$$
$$F_-(\omega) = f_{-1}\, e^{j\omega} + f_{-2}\, e^{j2\omega} + \cdots \quad \text{(strictly anticausal)}.$$

Most applications in system identification assume a causal system $F_d(\omega)$, meaning $F_-(\omega) = 0$. This constraint is not immediately absorbed into the formulation of [1], which may have its dangers. For example, a rational model $N(z)/D(z)$ lending some of its approximation power to the anticausal part $F_-(\omega)$ will necessarily have poles outside the unit circle; if such a rational function is misread as causal, it becomes unstable. It is therefore desirable to avoid target functions $F_d(\omega)$ having nontrivial anticausal parts.

Some of the results quoted subsequently are simpler when the weighting function $W(\omega)$ is chosen as a constant, although remarks concerning more general weighting functions will be accommodated where appropriate. As far as the influence of a nonzero anticausal part $F_-(\omega)$ is concerned, we shall treat it here briefly for the case of a constant weighting function $W(\omega)$.

Consider first the least-squares criterion (2) when $W(\omega)$ is constant. If we restrict the search space to stable polynomials for $D(z)$, it is straightforward to show that the error surface corresponding to (2) is theoretically insensitive to the anticausal part $F_-(\omega)$. Numerically, however, a nonzero $F_-(\omega)$ may introduce complications, which favors setting $F_-(\omega)$ to zero outright.

For the Steiglitz–McBride iteration (1), on the other hand, the influence of a nonzero anticausal part $F_-(\omega)$ is not so benign. Let δ be any constant for which the causal function

$$R(\omega) \triangleq \delta - F_+(\omega)F_-(-\omega) \tag{3}$$

is positive real. [If $F_+(\omega)$ and $F_-(\omega)$ are both bounded functions, then so is the real part of $F_+(\omega)F_-(-\omega)$; we may then take $\delta \geq \sup_\omega \operatorname{Re} F_+(\omega)F_-(-\omega)$]. We may then show that the stationary points of the iterative procedure

$$\hat{J}_k = \int_{-\pi}^{\pi} \left|\frac{F_+(\omega)D_k(e^{j\omega}) - N_k(e^{j\omega})}{D_{k-1}(e^{j\omega})}\right|^2 d\omega + \int_{-\pi}^{\pi} R(\omega)\left|\frac{D_k(e^{j\omega})}{D_{k-1}(e^{j\omega})}\right|^2 d\omega$$

coincide with the stationary points of the original procedure (1) if $W(\omega)$ is constant. A proof may be developed using the methods of [5] and is omitted for brevity. The key point is that a nonzero anticausal part $F_-(\omega)$ has an influence akin to an output disturbance in an identification context; the second term in the above expression can bias the final solution, except when its kernel $R(\omega)$ reduces to a constant [4], [5, Sec. 8.4], corresponding here to $F_-(\omega) = 0$ by way of (3). Note that if we use a dense frequency grid as in [1], (I)FFT techniques can be used to zero out the anticausal part of $F_d(\omega)$ (to within the frequency resolution of the dense grid), and therefore, we usually assume $F_-(\omega) = 0$ in the discussions to follow.

III. On the Quality of a Converged Model

Lu *et al.* claim that any convergent point of the procedure (1) will yield a local minimum of the L_2-norm criterion from (2). This claim is known to be incorrect [6], [5, Th. 7.2 versus Th. 8.5]; in particular, the Steiglitz–McBride procedure is not an optimization algorithm in any correct sense of the term. It is true that in many cases studied (see [7] and [8] for concrete examples), the convergent point(s) of the Steiglitz–McBride procedure appear(s) acceptably close to a *global* minimum of the L_2-norm criterion from (2).

One of the few formal results supporting the good quality models furnished by the Steiglitz–McBride method can be found in [9] for the case in which the weighting function $W(\omega)$ is constant.

Introduce the infinite Hankel matrix

$$\Gamma_F = \begin{bmatrix} f_1 & f_2 & f_3 & \cdots \\ f_2 & f_3 & f_4 & \cdots \\ f_3 & f_4 & f_5 & \cdots \\ \vdots & \vdots & \vdots & \iddots \end{bmatrix}$$

and set its singular values [called the Hankel singular values of $F_d(\omega)$] in decreasing order: $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \cdots$. Let $F(z) = N_\infty(z)/D_\infty(z)$ be the transfer function obtained at any stationary (but not necessarily convergent) point of the Steiglitz–McBride procedure. In [9], it is shown that whenever $F(z)$ is stable, the resulting L_2 error norm can be bounded as

$$\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |F_d(\omega) - F(e^{j\omega})|^2\, d\omega\right)^{1/2} \leq \sigma_{n+1}. \tag{4}$$

An adjustment to this bound to handle the case in which $F_-(\omega) \neq 0$ can likewise be found in [9] but is omitted here for brevity.

This bound is relevant because the Hankel singular values have a direct bearing on how well a given function can be approximated in rational form. A celebrated result of Adamjan *et al.* [10] asserts that σ_{n+1} is precisely the distance, in Hankel norm, between $F_d(\omega)$ and the set of rational functions of degree not exceeding n. These same Hankel singular values also provided loose, although quite usable, bounds on the attainable approximation error with other criteria, such as the L_2 and L_∞ norms (see, e.g., [5, Sec. 4.7]). Since the global minimum of (2) also satisfies the bound (4) when $W(\omega)$ is constant (although the local minima need not), it follows that any stationary point of the Steiglitz–McBride iteration will lie near a global minimum of the L_2 criterion whenever σ_{n+1} is sufficiently small. In particular, if estimates of the Hankel singular values of $F_d(\omega)$ can be obtained, the bound (4) can be used to estimate an order n, leading to a good approximation in the filter design.

When $W(\omega)$ is nonconstant, a similar proximity between a convergent point of the Steiglitz–McBride procedure and a global minimum of (2) may still be observed [11]; see [8] for further analysis of this phenomenon.

IV. On Convergence of the Procedure

Although Lu *et al.* attempt a convergence inference by appealing (incorrectly) to a minimization argument, the question of whether the Steiglitz–McBride algorithm will always converge has never been satisfactorily answered, except for the particular case in which $F_d(\omega)$ is already a rational function of degree n or less [4]. The major obstacle is that as mentioned earlier, the Steiglitz–McBride algorithm is not an optimization routine. In particular, no known distance measure has been uncovered whose minima the Steiglitz–McBride algorithm can claim to seek.

The question of whether a stationary point always exists *can* be given a qualified affirmative answer, at least for the case where the weighting function $W(\omega)$ is constant. As in [12], we say that $F_d(\omega)$ is "n-fold bounded-input–bounded-output (BIBO) stable" if

$$\sum_{k=1}^{\infty} k^n |f_{k+n}| < \infty.$$

This is more stringent than BIBO stability but less stringent than exponential stability. Whenever $F_d(\omega)$ is n-fold BIBO stable, the Steiglitz–McBride iteration admits a stationary point for which the resulting model $N_\infty(z)/D_\infty(z)$ is asymptotically stable [12]. (The n-fold BIBO stability is sufficient for this result, but its necessity remains open.) This result, however, does not claim that the set of stationary points will always include an attractor point for the iteration. An extension of this result to nonconstant weighting functions $W(\omega)$ remains open.

V. On the Stability of Successive Iterates

Lu *et al.* raise the important question of whether each step of the Steiglitz–McBride iteration will yield $D_k(z)$ as a stable polynomial. For the case of a constant weighting function, $D_k(z)$ is stable whenever $D_{k-1}(z)$ is stable [5, Th. 8.1], [12]. For nonconstant weighting functions $W(\omega)$, by contrast, the polynomial $D_k(z)$ can become unstable [4]. Having $F_-(\omega)$ nonzero may also induce risks, as noted earlier.

In this light, the additional inequality constraints absorbed by Lu *et al.* for ensuring stability are of definite interest. We may remark, however, that the constraint in question [1, eq. (20) or (21)] is sufficient, but not necessary, for stability of $D_k(z)$. Absorbing such a constraint is thus tantamount to restricting the search space, which may conceivably give poor results in some cases, particularly since [1, eq. (20) or (21)] becomes more stringent as n increases.

An alternate approach is to reparametrize $N_k(z)$ and $D_k(z)$ such that stability becomes inherent to the parametrization. One possibility may be found in [5, Sec. 8.7], using normalized lattice parameters. Instead of working with the direct-form parameters, with their well-recognized numerical drawbacks, each step determines reflection coefficients that , by virtue of the procedure, are inherently bounded by one, thus ensuring stablility. Moreover, the stationary points coincide with those obtained from the direct-form version. The sequence of iterates, however, is generally different from that obtained with the direct-form version, and like that version, convergence remains difficult to prove.

VI. Concluding Remarks

The design examples from [1] illustrate that the Steiglitz–McBride algorithm has clear utility in digital filter design. Indeed, it is already used for this purpose in the Signal Processing Toolbox of Matlab [13]. A clearer understanding of some of the method's properties, as outlined in this correspondence, may be beneficial toward further improvements to the method developed in [1].

We should also note that the iterative quadratic maximum likelihood (IQML) technique for direction-of-arrival estimation is known to be equivalent to the Steiglitz–McBride iteration [14]. Some recent results on IQML, concerning (lack of) consistency [15] and conditions for local convergence with expressions for asymptotic variances [16], are thus of direct relevance to the further study of this family of iterative methods.

Acknowledgment

The author would like to thank Prof. P. Stoica for useful remarks on an earlier draft of this note.

References

[1] W.-S. Lu, S.-C. Pei, and C.-C. Tseng, "A weighted least-squares method for the design of stable 1-D and 2-D IIR digital filters," *IEEE Trans. Signal Processing*, vol. 46, pp. 1–10, Jan. 1998.

[2] C. K. Sanathanan and J. Koerner, "Transfer function synthesis as a ratio of two complex polynomials," *IEEE Trans. Automat. Contr.*, vol. AC-8, pp. 56–58, Jan. 1963.

[3] K. E. Steiglitz and L. E. McBride, "A technique for the identification of linear systems," *IEEE Trans. Automat. Contr.*, vol. AC-10, pp. 461–464, Oct. 1965.

[4] P. Stoica and T. Söderström, "The Steiglitz–McBride identification algorithm revisited—Convergence analysis and accuracy aspects," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 712–717, June 1981.

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

[6] T. Söderström and P. Stoica, "On some system identification techniques for adaptive filtering," *IEEE Trans. Circuits Syst.*, vol. 35, pp. 457–461, Apr. 1988.

[7] H. Fan and M. Nayeri, "On reduced-order identification: Revisiting 'On some system identification techniques for adaptive filtering'," *IEEE Trans. Circuits Syst.*, vol. 37, pp. 1144–1151, Sept. 1990.

[8] H. Fan and M. Doroslovački, "On 'global convergence' of Steiglitz–McBride adaptive algorithm," *IEEE Trans. Circuits Syst. II*, vol. 40, pp. 73–87, Feb. 1993.

[9] P. A. Regalia and M. Mboup, "Undermodeled adaptive filtering: An *a priori* error bound for the Steiglitz–McBride method," *IEEE Trans. Circuits Syst. II*, vol. 43, pp. 105–116, Feb. 1996.

[10] V. M. Admjan, D. Z. Arov, and M. G. Kreĭn, "Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur–Takagi problem," *Math. USSR Sbornik*, vol. 15, pp. 31–73, 1971.

[11] H. Fan and K. Jenkins, "A new adaptive IIR filter," *IEEE Trans. Circuits Syst.*, vol. CAS-33, pp. 939–947, 1986.

[12] P. A. Regalia, M. Mboup, and M. Ashari, "On the existence of stationary points for the Steiglitz–McBride algorithm," *IEEE Trans. Automat. Contr.*, vol. 42, pp. 1592–1596, Nov. 1997.

[13] *Matlab Signal Processing Toolbox User's Guide*. Natick, MA: The Mathworks, 1996.

[14] J. H McClellan and D. Lee, "Exact equivalence of the Steiglitz–McBride iteration and IQML," *IEEE Trans. Signal Processing*, vol. 39, pp. 509–512, Feb. 1991.

[15] P. Stoica, J. Li, and T. Söderström, "On the inconsistency of IQML," *Signal Process.*, vol. 56, pp. 185–190, Jan. 1997.

[16] J. Li, P. Stoica, and A.-S. Liu, "Comparative study of IQML and MODE direction-of-arrival estimators," *IEEE Trans. Signal Processing*, vol. 46, pp. 149–160, Jan. 1998.