PIPELINED IIR FILTER ARCHITECTURE USING POLE-RADIUS MINIMIZATION

Nigel Boston

Departments of Mathematics and Electrical and Computer Engineering, University of Wisconsin, Madison, WI 53706

Abstract. An extension of a polynomial consists of the polynomial plus higher power terms. Given a polynomial with real coefficients and an integer larger than its degree, a method is given that produces a finite list of extensions of degree this larger integer such that this list necessarily contains the extension whose largest root is as small as possible. This extension is called the pole radius minimizer. The pole radius minimizer is then found by the finite check of comparing the polynomials in the list. The method is applied to obtain filter transformations that are optimal as regards throughput, but also have considerable savings in hardware overhead compared with standard methods such as Scattered Lookahead and Minimum Order Augmentation. The table in section V gives an explicit comparison for various kinds of filters.

I. INTRODUCTION

The transfer function of an IIR filter is of the form $B(z)/A(z)$, where A and B are polynomials in $\mathbf{R}[z^{-1}]$. Such a filter can be transformed into a pipelined filter by adding additional poles (and cancelling zeros) so that the first few coefficients in the denominator are zero [1], [2]. Mathematically, this amounts to writing the transfer function as $(B(z)D(z))/(A(z)D(z))$, where the first few coefficients of $A(z)D(z)$ are zero. The idea is that the sequence of computations in a feedback loop determines a fundamental lower bound on the throughput of a VLSI architecture of an IIR filter. Zeroing out the first few coefficients results in algorithmic delays in the feedback loop that can be employed to pipeline the feedback loop and thereby improve the throughput. The additional terms in the numerator are implemented in a feedforward architecture and hence can be pipelined to an arbitrarily high level. In [3] it is noted that speedup is also accomplished by reducing the computational complexity of the inner feedback loops (removing the need for pipelining), for example by replacing multipliers by shifters and skeletal multipliers (multipliers consisting of shifters and a few adders).

This method, called the SPOT (Sum of Powers-Of-Two) transformation, is efficiently implemented by partitioning the feedback loops into three sections. The first "shifter" section consists of the innermost feedback loops, each coefficient being a power of 2 (thus requiring no pipelining). The second "skeletal-multiplier" section consists of shifters and adders, each coefficient is a short sum of powers of 2, and so it can be fully pipelined. The last "multiplier" section is composed of

Typeset by $\mathcal{AMS}$-TEX

multipliers. The outputs of the three sections are then added using the partially pipelined added tree. For more details, see [3].

Mathematically, this amounts to finding $D(z) \in \mathbf{R}[z^{-1}]$ such that the first few coefficients of $A(z)D(z)$ are powers of two or sums of a few such powers. We also ask that the roots of $D(z)$ all lie inside the unit circle, preferably as small as possible, so that the resulting filter is stable. In [3] a technique was given for doing this and resulted in better solutions as regards hardware overhead. In this paper, a new method for finding suitable, in fact best possible $D(z)$ is given. The results improve on the more ad hoc methods of [3] in that $D(z)$ of smaller degree are obtained and the coefficients in $A(z)D(z)$ are simpler powers of two. This further reduces the hardware overhead.

Mathematically, if we specify the first M nontrivial coefficients of $A(z)D(z)$ - say $A(z)D(z) = 1 + c_1z^{-1} + c_2z^{-2} + ... + c_Mz^{-M} + ...$, then since $A(z)$ is given, this determines the first M nontrivial coefficients of $D(z)$ - say $D(z) = 1 + d_1z^{-1} + d_2z^{-2} + ... + d_Mz^{-M} + ...$ The d_i's are explicit linear combinations of the c_i's. Let $f(z) = 1 + d_1z^{-1} + d_2z^{-2} + ... + d_Mz^{-M}$. We call a polynomial in $\mathbf{R}[z^{-1}]$ *stable* if all its roots lie inside the unit circle. If the truncation of $D(z)$ to degree M is f, then we call D a *stable extension* of f. Our problem then is to find a stable extension of our given f of as small degree L as possible. More precisely, we seek the extension of f of degree L that has largest root as small as possible (the pole radius minimizer).

Surprisingly a careful search of the mathematical literature turns up little work on this basic problem in analysis, although it can be considered as a generalization of the classical root-locus problem. A few papers [4], [5], [6], [7] addressing this same problem appeared in Russia around about 1940. For instance, [4] shows that if C is any closed star-shaped Jordan curve containing the origin in its interior, then there exists an extension of f all of whose roots lie on C. Thus stable extensions certainly always exist. These papers did not, however, quantify the degree of a minimal such extension, something we handle here.

We give a complete solution for small values of M and then give a general method that, given f and L, yields a finite list of extensions of f of degree L, one of which is necessarily the pole radius minimizer. This is considerably better than the usual pipelining overhead of $\deg(A)M$. For example, in a case with $M = 6$, for which [3] gave a $D(z)$ of degree 14, we obtain a $D(z)$ of degree 8. In many cases, we actually obtain a $D(z)$ of degree as small as possible, namely M, and give a constructive method that often yields this solution. In all cases our result yields the best possible transformation.

II. SMALL VALUES OF M

To begin, consider the case $M = 1$. Here, $f(z) = 1 + az^{-1}$ for some $a \in \mathbf{R}$.

Theorem 1: There exists a stable extension of degree L if and only if $\mid a \mid < L$.

Proof. Given such an a it has the stable extension $(1 + az^{-1}/L)^L$. Conversely, if it has a stable extension, then the roots of this extension, say $\alpha_1, ..., \alpha_L$, satisfy $\alpha_1 + ... + \alpha_L = -a$. Since by stability $\mid \alpha_i \mid < 1$ for all i, $\mid a \mid < L$ follows.

Next suppose $M = 2$. Here, $f(z) = 1 + az^{-1} + bz^{-2}$ for some $a, b \in \mathbf{R}$. Reasoning as above, such a polynomial with a stable extension must satisfy $\mid a \mid < L$ and $\mid b \mid < L(L-1)/2$. The converse is, however, not true. There is a certain region $V_L \subseteq \mathbf{R}^2$ in which the polynomial $f(z)$ has a stable extension of degree L. Clearly, $V_2 \subseteq V_3 \subseteq V_4 \subseteq ...$ What more can we say?

First, V_L is symmetric about the b-axis since $f(z)$ has a stable extension of degree L if and only if $f(-z)$ does. We can compute the first few V_L explicitly. Note that V_2 is simply the region in which f is itself stable.

Theorem 2: V_2 is the interior of the region in $\mathbf{R}^2$ bounded by the three lines $b = 1, b = a - 1, b = -a - 1$.

Moreover, the boundary curves of V_2 correspond to the following f

$1 + az^{-1} + z^{-2}$ $(-1 \leq a \leq 1)$

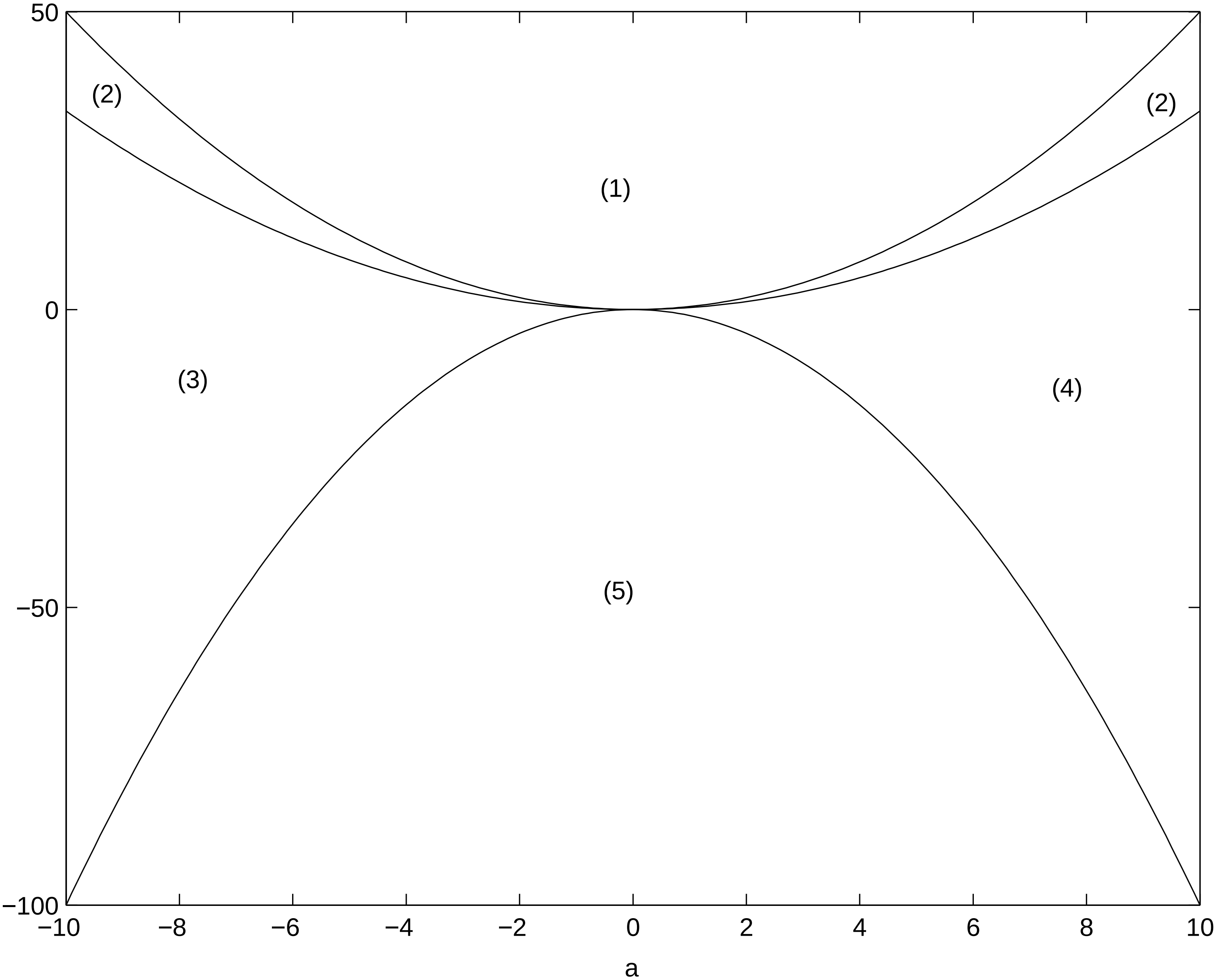
$1 + az^{-1} + (-a - 1)z^{-2}$ $(-2 \leq 0 \leq 0)$

$1 + az^{-1} + (a - 1)z^{-2}$ $(0 \leq a \leq 2)$

Note that this is traditionally called the *stability triangle* in the DSP literature. Next consider the case of V_3. The pole radius of a polynomial will denote the maximum of the absolute values of its roots (so that a stable polynomial is precisely one with pole radius 1). We can prove the following very useful theorem by the method described below.

Theorem 3: For given $a, b \in \mathbf{R}$, but varying $c \in \mathbf{R}$, the polynomial $1 + az^{-1} + bz^{-2} + cz^{-3}$ with smallest pole radius has $c \in \{ab, (b/a)^3, (ab)/2 - a^3/8, (a(9b - 2a^2) \pm 2(a^2 - 3b)^{3/2})/27\}$.

Moreover, there is a simple look-up criterion. If $b > a^2/2$, then $c = (ab)/2 - a^3/8$ (region 1 in the figure below). If $a^2/3 < b < a^2/2$, then $c = (b/a)^3$ (region 2 below). If $-a^2 < b < a^2/3$, then $c = (a(9b - 2a^2) \pm 2(a^2 - 3b)^{3/2})/27$, with the plus sign if $a < 0$ (region 3 below) and the minus sign if $a > 0$ (region 4). If $b < -a^2$, then $c = ab$ (region 5 below).



The point is that now in order to find a pole radius minimizer we have to check at most five candidate polynomials. The finer analysis simplifies further to a look-up. These five candidate polynomials arise respectively from imposing a relation on the roots α, β, γ of the cubic, namely $\alpha = -\beta, \alpha\beta = \gamma^2, \alpha + \beta = \gamma, \alpha = \beta$ respectively (the last one giving two possible values of c). This then yields the following V_3.

Corollary 1: V_3 is the interior of the region in $\mathbf{R}^2$ bounded by the five lines $b = -1, b = 2a - 3, b = -2a - 3, b = a, b = -a$ and the parabola $b = (1/4)a^2 + 1$.

Moreover, the boundary curves correspond to the following extensions of f

$1 + az^{-1} - az^{-2} - z^{-3}$ $(-3 \le a \le -2)$

$1 + az^{-1} + ((1/4)a^2 + 1)z^{-2} + (a/2)z^{-3}$ $(-2 \le a \le 2)$

$1 + az^{-1} + az^{-2} + z^{-3}$ $(2 \le a \le 3)$

$1 + az^{-1} + (-2a - 3)z^{-2} + (a + 2)z^{-3}$ $(-3 \le a \le -1)$

$1 + az^{-1} - z^{-2} - az^{-3}$ $(-1 \le a \le 1)$

$1 + az^{-1} + (2a - 3)z^{-2} + (a - 2)z^{-3}$ $(1 \le a \le 3)$

The case of V_4 is actually simpler than that of V_3. We can prove the following very useful theorem.

Theorem 4: For given $a, b \in \mathbf{R}$, but varying $c, d \in \mathbf{R}$, the polynomial $1 + az^{-1} + bz^{-2} + cz^{-3} + dz^{-4}$ with smallest pole radius is one of the following three:

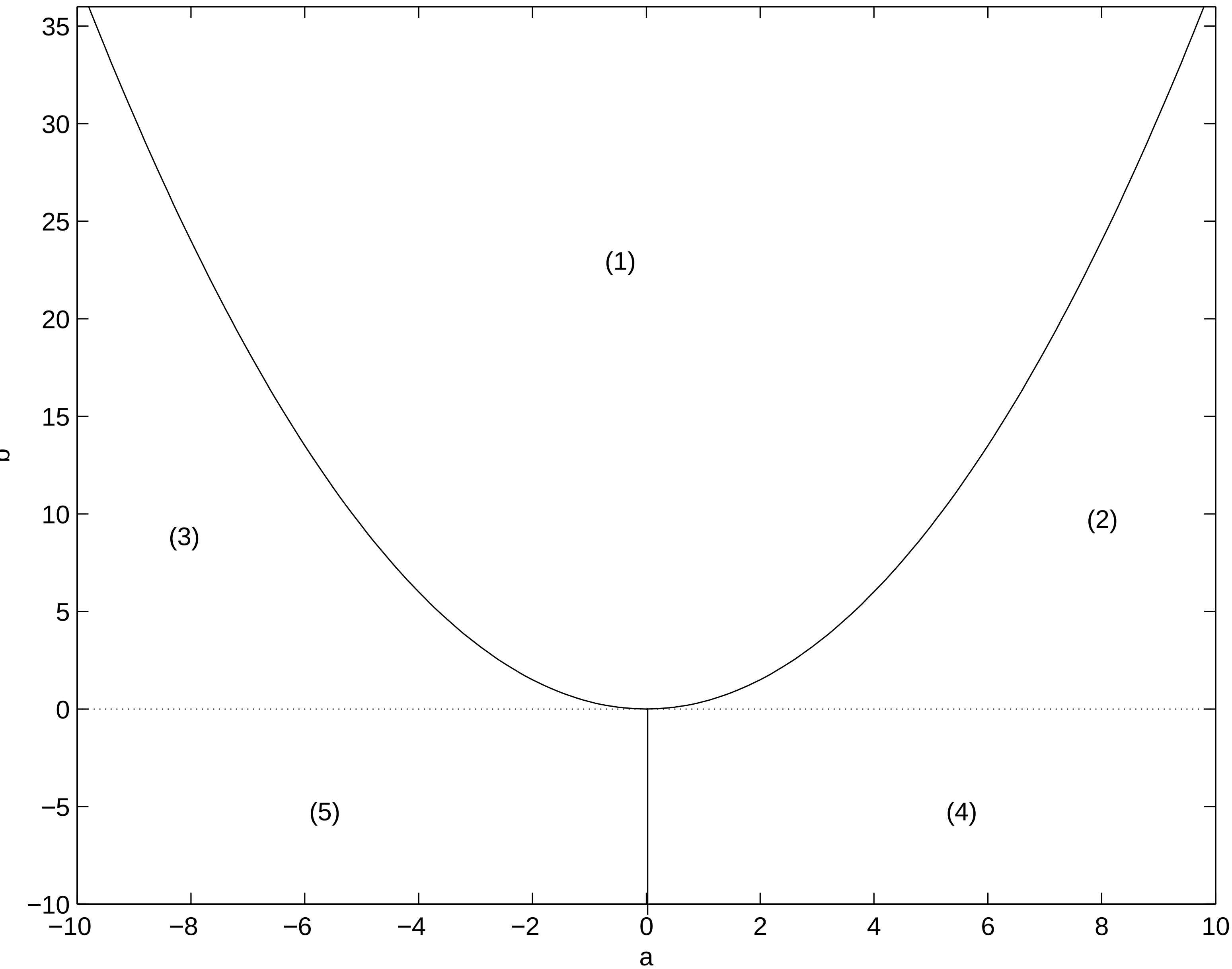
(i) $1 + az^{-1} + bz^{-2} + (ab/2 - a^3/8)z^{-3} + ((a^2 - 4b)^2/64)z^{-4}$

(ii) $1 + az^{-1} + bz^{-2} + ((-a^3 \pm a^2\sqrt{a^2 - 8b} + 4ab)/8)z^{-3} + ((-a^4 \pm a^3\sqrt{a^2 - 8b} + 4a^2b + 8b^2)/32)z^{-4}$

(iii) $1 + az^{-1} + bz^{-2} + ((-9a^3 \pm \sqrt{3}(3a^2 - 8b)^{3/2} + 36ab)/72)z^{-3} + ((-27a^4 \pm 3\sqrt{3}a(3a^2 - 8b)^{3/2} + 108a^2b - 72b^2)/864)z^{-4}$

where the $\pm$ are chosen to be the same sign in the coefficients of z^{-3} and z^{-4}.

Moreover, there is a simple look-up criterion. If $8b > 3a^2$, then (i) is best (region 1 in the figure below). If this fails, and $b > 0$, then (ii) is best, choosing $+$ if $a > 0$ (region 2 below) and $-$ if $a < 0$ (region 3 below). If $b < 0$, then (iii) is best, choosing $+$ if $a > 0$ (region 4 below) and $-$ if $a < 0$ (region 5 below).



These polynomials arise from imposing respectively that the four roots be of the form $(\alpha,\alpha,\beta,\beta),(\alpha,\alpha,-\alpha,\beta),(\alpha,\alpha,\alpha,\beta)$. Note that if $a^2 < 8b/3$, then neither of the last two polynomials has real coefficients and in this case the pole radius minimizer is necessarily the first polynomial.

Corollary 2: V_4 is the interior of the region in $\mathbf{R}^2$ bounded by the four lines $b = a-2, b = -a-2, b = 3a-6, b = -3a-6$ and the parabola $b = (1/4)a^2+2$.

Moreover, the boundary curves correspond to the following extensions of f

$1 + az^{-1} + ((1/4)a^2+2)z^{-2} + az^{-3} + z^{-4}$ $(-4 \le a \le 4)$

$1 + az^{-1} + (-3a-6)z^{-2} + (3a+8)z^{-3} + (-a-3)z^{-4}$ $(-4 \le a \le -2)$

$1 + az^{-1} + (-a-2)z^{-2} - az^{-3} + (a+1)z^{-4}$ $(-2 \le a \le 0)$

$1 + az^{-1} + (a-2)z^{-2} - az^{-3} + (-a+1)z^{-4}$ $(0 \le a \le 2)$

$1 + az^{-1} + (3a-6)z^{-2} + (3a-8)z^{-3} + (a-3)z^{-4}$ $(2 \le a \le 4)$

In general, we have the following. Here a *variety* is any subset of $\mathbf{R}^L$ given by a finite number of polynomial equations. The Hilbert basis theorem implies that every variety is a union of finitely many irreducible subvarieties, hence reducing each problem we face to a finite check as in the given examples.

Theorem 5: For given $a_1,...,a_M \in \mathbf{R}$ and given $L > M$, there is a computable variety $V(M,L)$ of dimension M such that the extension $1 + a_1z^{-1} + ... + a_Lz^{-L}$ of degree L of $1 + a_1z^{-1} + ... + a_Mz^{-M}$ of minimal pole radius has $(a_1,...,a_M) \in V(M,L)$.

For instance, Theorem 3 gives that $V(2,3)$ is the union of the irreducible subvarieties $V_1 = \{(a,b,c) : c = ab\}, V_2 = \{(a,b,c) : a^3c = b^3\}, V_3 = \{(a,b,c) : 8c + a^3 = 4ab\}, V_4 = \{(a,b,c) : (27c - a(9b-2a^2))^2 = 4(a^2-3b)^3\}$.

Our problem amounts to identifying this variety. In the case that $M = 1$, it is easy to see that the pole radius minimizer corresponds to imposing the condition that all the roots be equal. In other cases it is harder but there is always just a short list of finite cases to check.

A related problem is to describe the set $\{a_1,...,a_L) \in \mathbf{R}^L : 1+a_1z^{-1}+...+a_Lz^{-L}$ is stable $\}$. This region is called a *stability domain* and is studied in the theory of robust control, where ellipsoidal approximations to this non-convex set [8] and exact descriptions for small L (e.g. $L = 3$ in [9]) are found.

III. THE MAIN METHOD

We sketch the proof of the theorem by illustrating it in a straightforward case that brings in all the aspects of the problem. Suppose $M = 3$ and $L = 4$. In other words, we are given $1 + az^{-1} + bz^{-2} + cz^{-3}$ $(a,b,c \in \mathbf{R})$ and we seek $d \in \mathbf{R}$ such that $1 + az^{-1} + bz^{-2} + cz^{-3} + dz^{-4}$ has pole radius as small as possible.

Suppose the roots of the quartic are r,s,t,u. Our hypotheses impose

$$r + s + t + u = -a$$

$$rs + rt + ru + st + su + tu = b$$

$$rst + rsu + rtu + stu = -c$$

and we wish to minimize $\max(|r|,|s|,|t|,|u|)$. This suggests Lagrange multipliers, except for the problem that $\max(|r|,|s|,|t|,|u|)$ is not differentiable. The way around this is to break up the problem into cases where the Lagrange multiplier method is legitimate (i.e. local differentiability) and cases where things degenerate.

The Lagrange multiplier method works in cases where an optimal solution has its largest roots being a pair of nonreal complex conjugates, say r and s. Then we seek to minimize (the differentiable) rs (which is locally the square of the pole radius) subject to the three equalities above. Differentiating

$$rs+\lambda(r+s+t+u+a)+\mu(rs+rt+ru+st+su+tu-b)+\nu(rst+rsu+rtu+stu+c)$$

with respect to r, s, t, u respectively gives:

$$s+\lambda+\mu(s+t+u)+\nu(st+su+tu)=0$$

$$r+\lambda+\mu(r+t+u)+\nu(rt+ru+tu)=0$$

$$\lambda+\mu(r+s+u)+\nu(rs+ru+su)=0$$

$$\lambda+\mu(r+s+t)+\nu(rs+rt+st)=0$$

These together with the three equalities above yield 7 equations in 7 unknowns (treating a, b, c as given). There is no nicely expressed solution (unlike e.g. the case $c = ab/2 - a^3/8$ of theorem 3 above, which comes from Lagrange multipliers, corresponding simply to $r + s = t$), but for given a, b, c, NSolve on the software package Mathematica yields all solutions.

For example, taking $a = -0.8883, b = 1.469, c = 1.516$, Mathematica gives 24 solutions, leading to just 2 possible real values of $rstu$, namely 0.297564 and -0.549381. These give pole radii 1.566 and 1.547 respectively. Comparing these with the pole radii coming out of the degenerate cases below, we see that $d = -0.549381$ gives the pole radius minimizer. A simulation using successive approximation confirms this.

The other possibility is that an optimal solution occurs for a degenerate case, where two roots not in a complex conjugate pair have equal absolute value. So we might have here (i) $r = s$, (ii) $r = -s$ (but this actually never gives a pole radius minimizer - it is always beaten by one of the others), or (iii) r and s a complex conjugate pair with $rs = t^2$.

From (i), our equations simplify to

$$2r+t+u=-a, r^2+2rt+2ru+tu=b, r^2t+r^2u+2rtu=-c$$

Mathematica will explicitly solve these, giving $rstu = r^2tu$ as a complicated algebraic expression in a, b, c involving cube roots. We can in practice either use NSolve with numerical values of a, b, c or use the following simple trick. If $f(z) = z^4+az^3+bz^2+cz+d$ has a repeated root r, then r is a root of $f'(z)$ too. We know f' since we are given a, b, c. For each of the 3 roots of this f', we plug them back into f to find d. We then find the pole radius of each of these 3 polynomials f.

From (iii), letting $r = te^{i\theta}, s = te^{-i\theta}$ and introducing new variable $v = 1 + 2\cos(\theta)$, our equations simplify to

$$tv+u=-a, t^2v+tvu=b, t^3+t^2vu=-c$$

Again, Mathematica gives explicit solutions, involving roots of a certain quintic with coefficients expressions in a, b, c. The numerical method works better.

Continuing our example above, case (i) yields just one real solution, namely 0.297564 (which also comes out of the Lagrange multiplier method since $r = s$ gives a solution to the four Lagrange multiplier equations). As noted earlier, this fails to be the pole radius minimizer. Case (iii) yields only one solution with $rstu = t^3u$ real, namely -146.396 (note that $v = 1.14844$, yielding a real θ). This gives pole radius 3.572.

In conclusion, we obtain a short list of candidates for pole radius minimizer. Comparing these we have found the pole radius minimizer occurs for $d = -0.549381$ and is 1.547.

By these means we find that V_L is always bowl-shaped, consisting of a top parabola and a bottom made up of L straight lines. This result allows one to find precisely the shortest stable extension of any given degree $M = 2$ polynomial (by finding the first V_L containing the corresponding point (k, l) in $\mathbf{R}^2$). We could attempt to continue this into higher dimensions, but since values of M are in practice around about 6, this is not beneficial.

IV. PRACTICAL EXAMPLES

Example 1:

Let $A(z) = 1 - 3.6951z^{-1} + 9.4066z^{-2} - 15.9195z^{-3} + 21.0099z^{-4} - 21.0668z^{-5} + 16.7477z^{-6} - 10.1227z^{-7} + 4.6058z^{-8} - 1.4077z^{-9} + 0.2436z^{-10}$. This is the denominator of the transfer function for a tenth order low-pass elliptic filter with stop band edge at 0.4 sampling frequency, passband ripple of 0.5, and stopband attenuation of 40dB. Let $M = 6$. We seek a polynomial $D(z) = 1 + d_1z^{-1} + ... + d_Lz^{-L}$, with L as small as possible, such that the first M coefficients $c_1, ..., c_M$ of $A(z)D(z)$ are small powers of 2 (or possibly short sums of these). The authors give an example with $L = 14$ and $c_1 = 0, c_2 = 0, c_3 = 2, c_4 = -0.25, c_5 = -1, c_6 = 2$. Using 2 skeletal multipliers, they find a $D(z)$ of degree 8 with $c_1 = 0, c_2 = 2, c_3 = 2, c_4 = 2, c_5 = 2.75, c_6 = 2.875$. We obtain a $D(z)$ of degree 8 using no skeletal multipliers, thus reducing the number of adders required to implement.

First, we solve for $d_1, ..., d_M$ in terms of $c_1, ..., c_M$. This gives $d_1 = 3.6591 + c_1, d_2 = 4.2472 + 3.6951c_1 + c_2, d_3 = -3.1451 + 4.2472c_1 + 3.6951c_2 + c_3, ...$ For any given choice of $c_1, ..., c_M$ (which then fixes $d_1, ..., d_M$), we have a case of our problem, going from degree 6 to degree 8. Since each choice is handled in a fraction of a second, letting $c_1, ..., c_M$ run through the set $\Sigma := \{0, \pm 1, \pm 2, \pm 0.5, \pm 4, \pm 0.25\}$ is feasible. (Note that computations are actually made to 15 decimal places, with just 4 reported at each stage).

It turns out that our optimal degree 8 polynomial corresponds to $c_1 = -0.5, c_2 = 2, c_3 = 1, c_4 = 2, c_5 = 1, c_6 = 4$ and has 2 pairs of equal roots. Equating its first 6 coefficients to the $d_1, d_2, ..., d_6$ already computed above yields 6 equations in the 6 unknowns, numerically solved by Mathematica to give 15 solutions. We compute the coefficients of z^{-7} and z^{-8} for each, yielding only 3 distinct cases in which these coefficients are real. Computing the pole radii of these three octics yields the (provably best) pole radius minimizer, namely:

$$d_0 = 1.0000, d_1 = 3.1951, d_2 = 4.3996, d_3 = 3.1216, d_4 = 2.0033,$$
$$d_5 = 3.0170, d_6 = 4.1240, d_7 = 2.9186, d_8 = 0.8965.$$

Here the pole radius is 0.9929. Note that trying for a degree 7 extension yields (according to the method of section 2, finding roots of $f'(z)$) no optimal extensions with real d_7. So our degree 8 extension is the smallest degree solution. If a smaller pole radius is desired, then one can use higher degree polynomials, e.g. $L = 10$, (and again this method produces the optimal such). The trade-off between degree and pole radius will depend on practical matters, but this method gives the best pole radius for any chosen L.

In terms of hardware overhead, the method of [3] uses 10 MAC (adder-multiplier pairs) and 11 A (adders). There are 6 extra adders because of the skeletal multipliers. Our method uses 10 MAC but only 5 A. This compares very favorably with the 30 MAC for the Scattered Lookahead method [1] and 42 MAC for the Minimum Order Augmentation (MOA) method [2].

Example 2:

Consider a sixth order low pass Butterworth filter with 0.3 cutoff frequency and $M = 6$. MATLAB gives the denominator of its transfer function as $A(z) = 1-2.3797z^{-1}+2.9104z^{-2}-2.0551z^{-3}+0.8779z^{-4}-0.2099z^{-5}+0.0218z^{-6}$. A new feature here is that we can obtain a suitable $D(z)$ of degree 6 (so best possible!) in a completely straightforward way. Namely, as usual, we compute that if $A(z)D(z) = 1+c_1z^{-1}+...+c_6z^{-6}+...$, then $D(z) = 1+d_1z^{-1}+...+d_6z^{-6}+...$ with $d_1, ..., d_6$ simple linear combinations of $c_1, ..., c_6$. Explicitly $D(z) = 1+(2.3797+c_1)z^{-1}+(2.7526+2.3797c_1+c_2)z^{-2}+(1.6795+2.7526c_1+2.3797c_2+c_3)z^{-3}+(-0.0017+1.6795c_1+2.7526c_2+2.3797c_3+c_4)z^{-4}+(-1.1146-0.0017c_1+1.6795c_2+2.7526c_3+2.3797c_4+c_5)z^{-5}+(-1.1346-1.1146c_1-0.0017c_2+1.6795c_3+2.7526c_4+2.3797c_5+c_6)z^{-6}+....$ Letting $c_1, ..., c_6$ run through Σ, we test each degree 6 polynomial for stability and obtain some examples! The one with smallest pole radius (namely 0.6894) has $c_1 = -2, c_2 = 2, c_3 = -1, c_4 = 0.25, c_5 = 0, c_6 = 0$. Here,

$$d_0 = 1.0000, d_1 = 0.3797, d_2 = -0.0068, d_3 = -0.0662,$$
$$d_4 = 0.0147, d_5 = 0.0902, d_6 = 0.0998.$$

Another notable example is that of $c_1 = -1, c_2 = 0.5, c_3 = c_4 = c_5 = c_6 = 0$, which has considerable pipelining but still yields a stable degree 6 polynomial with largest pole radius 0.8116.

In all these cases, the hardware overhead is 6 MAC plus 5 A. The method of [3] needs 14 MAC plus 11 A and the pole radius, 0.7177, is not as good, whereas Scattered Lookahead and MOA each take 18 MAC (pole radii 0.8085 and 0.7807 respectively).

V. A GENERAL METHOD

Despite the speed and guaranteed optimality of the method, the idea of testing many choices of $c_1, ..., c_M$ is unappealing to some. A heuristic for obtaining a typically good (but not guaranteed to be best) choice is as follows.

We illustrate this by sticking with the case $M = 6$. Suppose $A(z) = 1 + a_1z^{-1} + a_2z^{-2} + a_3z^{-3} + a_4z^{-4} + a_5z^{-5} + a_6z^{-6} + ...$ Solving for the coefficients of $D(z)$, we get $D(z) = 1 + (-a_1 + c_1)z^{-1} + (a_1^2 - a_2 - a_1c_1 + c_2)z^{-2} + (-a_1^3 + 2a_1a_2 - a_3 + a_1^2c_1 - a_2c_1 - a_1c_2 + c_3)z^{-3} + (a_1^4 - 3a_1^2a_2 + a_2^2 + 2a_1a_3 - a_4 - a_1^3c_1 + 2a_1a_2c_1 - a_3c_1 + a_1^2c_2 - a_2c_2 - a_1c_3 + c_4)z^{-4} + (-a_1^5 + 4a_1^3a_2 - 3a_1a_2^2 - 3a_1^2a_3 + 2a_2a_3 + 2a_1a_4 - a_5 + a_1^4c_1 - 3a_1^2a_2c_1 + a_2^2c_1 + 2a_1a_3c_1 - a_4c_1 - a_1^3c_2 + 2a_1a_2c_2 - a_3c_2 + a_1^2c_3 - a_2c_3 - a_1c_4 + c_5)z^{-5} + (a_1^6 - 5a_1^4a_2 + 6a_1^2a_2^2 - a_2^3 + 4a_1^3a_3 - 6a_1a_2a_3 + a_3^2 - 3a_1^2a_4 + 2a_2a_4 + 2a_1a_5 - a_6 - a_1^5c_1 + 4a_1^3a_2c_1 - 3a_1a_2^2c_1 - 3a_1^2a_3c_1 + 2a_2a_3c_1 + 2a_1a_4c_1 - a_5c_1 + a_1^4c_2 - 3a_1^2a_2c_2 + a_2^2c_2 + 2a_1a_3c_2 - a_4c_2 - a_1^3c_3 + 2a_1a_2c_3 - a_3c_3 + a_1^2c_4 - a_2c_4 - a_1c_5 + c_6)z^{-6} + ...$

First, we should test whether we can choose $c_1, ..., c_6$ so that the above degree 6 truncation is stable. Naturally, one might seek to make the coefficients as small as possible, i.e. choose c_1 to be $\pm$ the power of 2 closest to a_1 and c_2 to be $\pm$ the power of 2 closest to $-a_1^2 + a_2 + a_1c_1$ and so on. This can be automated so that MATLAB computes the transfer function, then finds successively $c_1, ..., c_6$, and then tests the resulting degree 6 polynomial for stability. This takes a fraction of a second.

Sometimes this works spectacularly. For instance, for the filter of Example 3 above, the method above gives $c_1 = -2, c_2 = 2, c_3 = -1, c_4 = 0.25, c_5 = -0.125, c_6 = 0.25$ and a largest pole radius of 0.7259. This is not far from the best choice we obtained using search, where c_5 and c_6 are taken to be 0 instead and the largest pole radius is 0.6894.

Unfortunately there are also cases where this rounding idea does not give values of $c_1, ..., c_6$ for which $D(z)$ is stable, but such values do exist. For example, the tenth order highpass Butterworth filter with cutoff frequency 0.3 in [3] has a stable degree 6 $D(z)$ coming from $c_1 = -2, c_2 = 2, c_3 = -0.5, c_4 = 0, c_5 = -1, c_6 = 2$, whereas the rounding idea yields $c_1 = -4, c_2 = 8, c_3 = -8, c_4 = 0.5, c_5 = 8, c_6 = -8$ and an unstable $D(z)$.

A general proposed method is as follows. For any given filter, we use MATLAB to compute $a_1, ..., a_6$. We then (1) compute the largest absolute value of a root of the degree 6 truncation of $D(z)$ (given above) for the particular choice of $c_1, ..., c_6$ given by the rounding idea. This is almost instantaneous and often works, i.e. yields a stable $D(z)$. If it produces an unstable $D(z)$, then we (2) try each of the 11^6 choices of $c_1, ..., c_6 \in \Sigma$. In many cases, we obtain instances where this largest absolute value is < 1, which leads to an optimal solution with $L = M = 6$. If this fails, we can either enlarge Σ or increase the attempted degree L, as explained below.

Here is a step-by-step elaboration of the proposed algorithm for general M.

Precomputation: Dividing $A(z)D(z) = 1 + c_1z^{-1} + ... + c_Mz^{-M} + ...$ by $A(z) = 1 + a_1z^{-1} + ... + a_Mz^{-M} + ...$ yields a symbolic expansion for $D(z)$ (as done for $M = 6$ at the start of this section).

Step 1: For the given filter, compute the first M coefficients of the numerator $A(z)$ of its transfer function.

Step 2: Since the ith coefficient of the symbolic expansion of $D(z)$ involves a single c_i, recursively set c_i to be the power of 2 that minimizes that coefficient of $D(z)$.

Step 3: Computes the roots of the $1 + d_1z^{-1} + ... + d_Mz^{-M}$ so produced. If it is stable, finish.

Step 4: If it is unstable, then for each of the 11^M choices of $c_1, ..., c_M \in \Sigma$, compute $1 + d_1z^{-1} + ... + d_Mz^{-M}$ and test for stability. This can be done best by starting with choices close to that produced in step 2. Stop when a stable example is found.

Step 5: In the event that none of the above yield a stable $D(z)$, we can enlarge Σ to include more powers of 2 ($\{\pm 8, \pm 0.125\}$) or try the next highest L. Enlarging Σ is more computationally expensive but the filter designed will have fewer MAC, whereas a higher L will give a quicker design solution but lead to a filter with more MAC. There is a trade-off and the implementer's priorities will decide between the two.

For the 9 examples in Table IV of [3], (1) works for 2 of them, (2) for a further 2, and (3) for the remaining 5. Below, we summarize our results in a table similar to Table IV of [3]. Each case is for $M = 6$. The largest pole radii are given in brackets.

Filter Specifications	**This Method**	**SPOT**	**Lookahead**	**MOA**
Low Elliptic, $N = 10$ $W_p = 0.4, R_p = 40, R_s = 0.5$	10 MAC + 5 A (0.9930)	10 MAC + 11A (0.97)	30 MAC (0.9981)	42 MAC (0.97)
Low Butterworth, $N = 6$ $W_p = 0.3$	6 MAC + 5A (0.6894)	14 MAC + 11 A (0.7177)	18 MAC (0.8085)	18 MAC (0.7807)
High Elliptic, $N = 6$ $W_p = 0.4, R_s = 40, R_p = 0.5$	6 MAC + 5 A (0.7865)	6 MAC +11 A (0.6727)	18 MAC (0.9698)	14 MAC (0.97)
High Butterworth, $N = 10$ $W_p = 0.3$	6 MAC + 5 A (0.9430)	10 MAC + 11 A (0.8882)	30 MAC (0.8805)	18 MAC (0.8998)
High ChebyshevII, $N = 8$ $W_p = 0.4, R_s = 40$	6 MAC + 5 A (0.7729)	6 MAC + 11 A (0.6691)	24 MAC (0.8921)	14 MAC (0.8938)

Note that in some cases our pole radii are worse than those of the others. That is because we have sought to minimize the degree of the stable extension. If instead we allow the degree to go as high as the others, i.e. accept the same number of adder-multiplier pairs, then we can get extremely small pole radii.

VI. REMARKS ON THE ALGORITHM

We are computing the degree 6 truncation of $D(z) = (1+c_1z^{-1}+c_2z^{-2}+...)/(1+a_1z^{-1}+a_2z^{-2}+...)$. In the rounding idea of the last section, in effect c_i is chosen close to a_i to try to make the roots of $D(z)$ small. If we set $\epsilon_i = c_i - a_i$, then $D(z) = 1+(\epsilon_1z^{-1}+\epsilon_2z^{-2}+...)/(1+a_1z^{-1}+a_2z^{-2}+...) = 1+\epsilon'_1z^{-1}+\epsilon'_2z^{-2}+...$, where $\epsilon'_1 = \epsilon_1, \epsilon'_2 = \epsilon_2 - a_1\epsilon_1, \epsilon'_3 = \epsilon_3 - a_1\epsilon_2 + (a_1^2 - a_2)\epsilon_1, ...$ We wish to have the polynomial $z^6 + \epsilon'_1z^5 + ... + \epsilon'_6$ have all its roots be small, in particular at least be inside the unit circle. Bounds for the largest root in terms of the coefficients are well-known - see [10], especially Chapter VII. These lead to a variety of situations in which the rounding method gives an immediate result.

ACKNOWLEDGEMENTS. The author thanks Naresh Shanbhag for very helpful discussions and input on this topic. He was partially supported by NSF grants DMS 99-70184 and EIA-0088063.

REFERENCES

1. K.K.Parhi and D.G.Messerschmitt, "Pipeline interleaving and parallelism in recursive digital filters," *IEEE Trans. Acoust. Speech and Signal Proc.* vol. 37, 1989, pp. 1099–1117.

2. Y.Lim and B.Liu, "Pipelined recursive filter with minimum order augmentation," *IEEE Trans. Signal Processing,* vol. 40, 1992, pp. 1643–1651.

3. M.Sharma and N.R.Shanbhag, "Architecture driven filter transformations," presented at the International Symposium on Circuits and Systems, Geneva, Switzerland, May, 2000.

4. N.G.Tschebotaröw, "Über die Fortsetzbarke-it von Polynomen auf geschlossene Kurven," *C.R. (Doklady) Acad. Sci. URSS (N.S.),* vol. 32, 1941, pp. 3–6

5. L.Gavrilov, "On K-prolongeable polynomials," *C.R. (Doklady) Acad. Sci. URSS (N.S.),* vol. 37, 1942, pp. 246–249.

6. L.Gawrilov, N. Tschebotaröw, "Über F-Polynome. VI. K-Polynome mit verschobenem Zentrum," *Bull. Soc. Phys.-Math. Kazan (3),* vol. 12, 1940, pp. 183–195.

7. N.G.Tschebotareff, "On continuable polynomials. I. General statement of the problem." in *Memorial volume dedicated to D.A.Grave [Sbornik posvjascenii pamjati D.A.Grave],* Moscow: publisher unknown, 1940, pp. 268–282.

8. D.Henrion, D.Peaucelle, D.Arzelier, and M.Sebek, "Ellipsoidal approximation of the stability domain of a polynomial," in *Proceedings of the European Control Conference ECC,* Porto, Portugal, 2001, pp. 384–389.

9. J.Ackermann, "Parameter space design of robust control systems,"*IEEE Trans. Automatic Control,* vol. 25, no. 6, 1980, pp. 1058–1072.

10. M.Marden, *The geometry of the zeros of a polynomial in a complex variable,* AMS Math Surveys, No. 3, New York, 1949, pp. 1–183.