Abstract—An iterative second-order cone programming (SOCP) approach is proposed in this paper. The original nonconvex design problem is first relaxed into an SOCP problem, which can provide a lower bound on the optimal value of the original problem. For reducing the gap between the original and the convex problem, an iterative procedure is developed. The initial point of the iterative procedure can be chosen as the solution obtained from the relaxed SOCP problem. Unlike other iterative approaches, the convergence of the proposed iterative procedure is definitely guaranteed. Design examples demonstrate the effectiveness of the proposed method.

I. INTRODUCTION

Large numbers of methods [1]-[10] have been proposed to design IIR digital filters. Two major difficulties need to be resolved in design procedures: 1) Due to the existence of the denominator, either linearization or nonlinear optimization techniques have to be deployed; 2) Stability constraints must be incorporated in the design procedure if the phase response is also concerned. In [1], a linear programming (LP) approach was presented. In order to tackle the first difficulty, the denominator part of the frequency response error is neglected. This makes the design result not to be a truly minimax solution. Some other methods [2]-[6] employ iterative procedures to solve this problem. At each iteration, the denominator obtained from the preceding iteration is treated as a part of the weighting function. This strategy has been applied to design IIR digital filters under different optimization criteria. Unfortunately, so far the convergences of these iterative procedures are not strictly guaranteed. On the other hand, the performance of iterative procedures is greatly dependent on the selection of initial points. Generally speaking, stability criteria can be categorized into two groups, i.e., time-domain stability criteria [7] and frequency-domain stability criteria [1], [6], [8]. Compared with time-domain stability criteria, frequency-domain stability criteria are much more tractable. However, most of stability criteria proposed in frequency domain are only sufficient conditions for stability. This means that some stable filters can be excluded from the admissible solutions.

Recently, convex optimization techniques [11] have been widely used in digital filter designs [1]-[6], [9], [10], [12]. When the design problem is cast to a convex form, the global optimum can be attained, if it exists. In an SOCP problem [11], [13], a linear function is minimized with the linear and second-order cone (SOC) constraints. SOCP is a general extension of linear and quadratic programs (QP). Thus, SOCP can be used to formulate more complicated design problems and handle more general types of quadratic constraints.

This paper is organized as follows. The minimax IIR digital filter design problem is first formulated in Section II. Then, the original nonconvex problem is relaxed into a convex form and the iterative procedure is developed. In Section III, two examples are presented to illustrate the effectiveness of the proposed method. Finally, conclusions are made in Section IV.

II. MINIMAX DESIGN OF IIR DIGITAL FILTER

A. Problem Formulation

The transfer function of an IIR digital filter is defined as

\[ H(z) = \frac{P(z)}{Q(z)} = \frac{\sum_{n=0}^{N} a_n z^{-n}}{1 + \sum_{n=1}^{M} b_n z^{-n}} = p^T \phi(z) \]

where \( p = [p_0, p_1, \ldots, p_N]^T \), \( a_n \) and \( q = [1, q_1, \ldots, q_M]^T \), and \( \phi(z) = [1, z^{-1}, \ldots, z^{-N}] \). The design problem of IIR digital filters in the minimax sense can be expressed as

\[ \min_{x} \max_{\omega \in \Omega} W(\omega) |H(e^{j\omega}) - D(e^{j\omega})| \]

where \( x = [p^T, q^T] \), \( D(e^{j\omega}) \) is the ideal frequency response, \( W(\omega) \geq 0 \) is a given nonnegative weighting function, and \( \Omega \) is the union of frequency bands of interest over \([0, \pi]\). By introducing an auxiliary variable \( \delta \), (2) can be further cast as

\[ \min_{x} \delta \]

\[ \text{s.t.} \quad q_0 = 1 \]

\[ W(\omega) \left[ P(e^{j\omega}) - D(e^{j\omega})Q(e^{j\omega}) \right] \leq \delta |Q(e^{j\omega})|^2, \quad \omega \in \Omega \]  \hspace{1cm} (3b)

where \( e(\omega) = W(\omega) \left[ \begin{array}{c} \text{Re}\{\phi(e^{j\omega})\} \\ \text{Im}\{\phi(e^{j\omega})\} \\ -\text{Re}\{D(e^{j\omega})\phi(e^{j\omega})\} \\ -\text{Im}\{D(e^{j\omega})\phi(e^{j\omega})\} \end{array} \right] \) (4)

For convenience in the later discussion, the magnitude of weighted frequency response error in (2), i.e., \( |W(\omega)H(e^{j\omega}) - D(e^{j\omega})| \), is replaced by \( |W(\omega)|H(e^{j\omega}) - D(e^{j\omega})|^2 \) in (3b).
Obviously, these two design problems are equivalent to each other.

B. Convex Relaxation

Note that only the squared magnitude of the denominator is required on the right hand side of the inequality constraint (3b). Actually, our design method is based on this fact.

It is known that [9]
\[
Q(z)Q(z^{-1}) = \left(1 + \sum_{m=0}^{M} q_m z^{-m}\right) \left(1 + \sum_{m=0}^{M} q_m z^{m}\right)
\]

where
\[
d_m = \sum_{m=0}^{M} q_m d_m^{m-m}
\]

From (6), some important properties can be derived: 1) \(d_0 = ||q||^2; 2) d_M = q_M^T 3) ||d_m|| \leq ||q||^2 = d_0\). Here, \(||\cdot||\) denotes the Euclidean norm of a vector. On the unit circle, (5) can be cast as a standard SOC constraint [13]. Thus, (3) is reformulated as

\[
Q(e^{i\omega}) = d_0 + \sum_{m=1}^{M} 2 d_m \cos(m\omega) = d^T s(\omega)
\]

where \(d = [d_0, d_1, \ldots, d_M]^T\), and \(s(\omega) = [1, 2\cos(\omega), \ldots, 2\cos(M\omega)]^T\). Using (7), the inequality constraint (3b) can be transformed to a hyperbolic constraint, which can be further cast as a standard SOC constraint [13]. Thus, (3) is formulated as

\[
\min \delta
\]

s.t. \(q_0 - 1 = 0\)
\[
\text{Re}\{f(e^{i\omega})\} \leq \delta d^T s(\omega), \ \omega \in \Omega,
\]
\[
\|Q(e^{i\omega})\|_2 = \|f^{T}(\omega)q\|_2 = d^T s(\omega), \ \omega \in [0, \pi]
\]

where
\[
f(\omega) = \left[\text{Re}\{\varphi(e^{i\omega})\}, \quad \text{Im}\{\varphi(e^{i\omega})\}\right]
\]

Note that the hyperbolic constraint (8b) needs to be satisfied over \(\Omega\), while (8c) must be satisfied at any \(\omega \in [0, \pi]\). Due to the existence of the quadratic equality constraint (8c), the problem (8) is still nonconvex. However, we can relax it into a convex problem by replacing (8c) with another hyperbolic constraint \(Q(e^{i\omega_0})^2 \leq d^T s(\omega_0)\). We then obtain a lower bound on the optimal value of (8) by solving the following convex problem

\[
\min \delta
\]

s.t. \(q_0 - 1 = 0\)
\[
d_m - q_m = 0
\]
\[
-\delta + d_m \leq 0, \ m = 1, 2, \ldots, M
\]
\[
-\delta - d_m \leq 0, \ m = 1, 2, \ldots, M
\]
\[
\|f^{T}(\omega)q\|_2 \leq d^T s(\omega), \ \omega \in [0, \pi]
\]

For the SOC problem (16), the variables are \(\delta\), \(\Delta x(\omega)\) (or \(v^{(k)}\)) and \(u^{(k)}\), and \(d^{(k)}\).

The iteration procedure will stop when the following condition is satisfied:
\[
\text{Re}\{e^{(\delta^{(k)}, \omega)}\} \leq \epsilon
\]

where \(\epsilon\) is a prescribed small positive tolerance. When \(k\) is large enough, \(\text{Re}\{e^{(\delta^{(k)}, \omega)}\} \approx \text{Re}\{e^{(\delta^{(k)}, \omega)}\} \approx 0\). Coupled with (14) and (15), it follows that \(|u^{(k)}| \approx 0\), which means there is no significant change on \(q^{(k)}\). Then the design problem can be equivalently stated as: Given a denominator \(q^{(k)}\), find an optimal numerator in the following iterations. The optimal point can be easily obtained, since \(|W(\omega)|H(e^{i\omega} - D(e^{i\omega}))|^{2}\) with
a fixed denominator \( q^{(b)} \) is essentially a quadratic function of \( p^{(b)} \).

D. Stability Constraint

It is well-known that at the \( k \)th iteration step the positive-realness based stability constraint [7] can be formulated as a linear inequality constraint:

\[
\text{Re}\{\phi_k^T(e^{(\omega)})\} \cdot q^{(i)} \geq \mu, \quad \omega \in [0, \pi]
\] (20)

where \( \mu \) is a small and positive number.

Another less conservative stability constraint is based on the Rouche’s theorem [8], which is stated as: If all zeros of \( Q^{(k)}(z) \) lie inside the circle \( C = \{ z : |z| < \rho, \rho \leq 1 \} \) and \( |Q^{(k)}(z) - Q^{(k-1)}(z)| < |Q^{(k-1)}(z)| \), all zeros of \( Q^{(k)}(z) \) will lie inside \( C \). In order to incorporate this constraint into (16), the initial denominator must be chosen with all zeros inside \( C \). Then the Rouche’s theorem based stability constraint can be written as

\[
\|f^{(i)}(\omega)u^{(i)}\| \leq |Q^{(k)}(e^{(\omega)})|, \quad \omega \in [0, \pi]
\] (21)

which is an SOC constraint.

Some remarks are made here:

1) Although only the lower bound on the optimal value of the original nonconvex problem (8) can be obtained by solving (10), in practice, the corresponding \( x \) and \( d \) can be chosen as the initial point, i.e., \( x^{(0)} \) and \( d^{(0)} \), for the iterative procedure.

2) In practice, the constraints (16c) and (16d) can be removed from (16). Large numbers of simulations show that their contributions to restricting \( d^{(i)} \) are negligible.

3) In order to guarantee the convergence, we incorporate (15) into the iterative procedure. However, it should be noted that this constraint is only a sufficient condition for the convergence. Therefore, the proposed method cannot be guaranteed to achieve the globally optimum.

4) In order to increase the convergence speed, the constraint (15) can be modified as

\[
-(d^{(i)} - d^{(i-1)}) + 2\alpha q^{(i)}u^{(i)} \geq \lambda^{(i-1)}
\] (22)

where \( \lambda^{(i-1)} \) is a positive number, whose value depends on the previous iteration. In our designs, \( \lambda^{(i-1)} \) is set as

\[
\lambda^{(i-1)} = \begin{cases} 0.5 \cdot e^{(i-1)} - \alpha \|u^{(i-1)}\| & \text{if } \|u^{(i-1)}\| \geq 10^{-4} \\ 0 & \text{otherwise} \end{cases}
\] (23)

where \( e^{(i)} \) denotes \( e(d^{(k-1)}, q^{(k-1)}) \).

III. SIMULATIONS

In this section, two examples are presented to demonstrate the effectiveness of the proposed method. We use the ScDuMi [14] MATLAB toolbox to solve the SOCP problem (16). Besides the peak errors and the \( L_2 \) errors of magnitude (MAG) and group delay (GD) in passbands and stopbands, the maximum (weighted) magnitude of the frequency response error is also adopted as a measurement:

\[
e_{\text{MAX}} = \max_{\omega \in \Omega} W(\omega) |H(e^{(i)}(\omega)) - D(e^{(i)}(\omega))|
\] (24)

The weighting function is always set to 1 over passbands and stopbands, and 0 over other transition bands. All hyperbolic constraints in (16) and stability constraint (20) or (21) are implemented on a set of dense grid points.

The first example is taken from [4]. The desired frequency response in [4] is defined as a smooth function. This smooth-transition-band method serves to reduce the Gibb’s ripple but at a possible cost of design resource at some unimportant frequency band(s). Therefore, the desired frequency response in this design is defined as

\[
D(e^{(\omega)}) = \begin{cases} e^{-j5\omega} & 0 \leq \omega \leq 0.5\pi \\ 0 & 0.55\pi \leq \omega < \pi \end{cases}
\]

The filter orders are set to \( M = N = 18 \). Like [4], we adopt (20) as the stability constraint used at each iteration, and the initial point is obtained by solving (10). Other parameters are selected as \( \alpha = 0.9, \varepsilon = 10^{-5}, \) and \( \mu = 10^{-6} \). It takes 29 iterations to stop. The maximal pole radius of the designed filter is 0.970. The design results are shown in Fig. 1. The changes of \( \log_{10}|e(d^{(i)}, q^{(i)})| \) are shown in Fig. 2. The plot shows nearly linear convergence on \( \log_{10}|e(d^{(i)}, q^{(i)})| \) especially when \( k < 20 \). All measurements are given in Table I. For comparison, we also design the lowpass filter using the method proposed in [4] under the same set of specifications except the initial point chosen as \( p^{(0)} = q^{(0)} = 0 \) for \( 0 \leq n \leq N \) and \( 1 \leq m \leq M \).

The second example is to design a highpass filter with the following ideal frequency response

\[
D(e^{(\omega)}) = \begin{cases} e^{-j4\omega} & 0.75\pi \leq \omega < \pi \\ 0 & 0 \leq \omega \leq 0.7\pi \end{cases}
\]

The filter orders are set as \( M = N = 15 \). Other parameters are chosen as the same in Example 1. It takes 27 iterations to find the final solution. Its maximal pole radius is 0.910. The design results are shown in Fig. 3. The changes of \( \log_{10}|e(d^{(i)}, q^{(i)})| \) are also shown in Fig. 4. All measurements are summarized in Table II. For comparison, we also utilize the method proposed in [4] to design the highpass IIR digital filter under the same specifications. The corresponding measurements are also given in Table II. It is shown that the proposed method can achieve better performances.

IV. CONCLUSIONS

In this paper, we have presented an iterative method for designing IIR digital filters under the minimax sense. Since the design problem is essentially nonconvex, first it is relaxed into a convex problem in an SOCP form. A lower bound on the optimal value of the original problem can be obtained by solving the convex design problem. In order to reduce the gap between these two design problems, an iterative procedure is then applied. In our designs, the initial point of the iterative procedure is obtained by solving the relaxed convex problem. At each iteration, an SOCP problem is solved. Stability, some constraint needs to be incorporated to control the maximal pole radius. Unlike some other approaches in which the design problem is formulated as LP or QP forms, stability constraints in quadratic forms can also be incorporated.

Compared with other iterative design methods, the obvious advantage of the proposed method is that the convergence can be strictly guaranteed, which has been shown by the analyses in Section II and demonstrated by examples. When \( \lambda^{(i)} \) in
(23) is chosen properly, the convergence speed can be improved, which has been demonstrated by the simulation results. By using the well-developed SeDuMi, the design problem can be efficiently solved.

REFERENCES


TABLE I. MEASUREMENTS OF DESIGN RESULTS IN EXAMPLE 1

<table>
<thead>
<tr>
<th>Method</th>
<th>Passband (MAG (dB) (Peak / $L_2$))</th>
<th>Stopband (MAG (dB) (Peak / $L_2$))</th>
<th>GD ($L_2$)</th>
<th>$\epsilon_{MM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed</td>
<td>-33.47/ -40.00</td>
<td>-33.20/ -38.39</td>
<td>2.54/ 0.20</td>
<td>0.0213</td>
</tr>
</tbody>
</table>


