

A Unified Approach to the Design of Interpolated and Frequency-Response-Masking FIR Filters

Wu-Sheng Lu

Dept. of Electrical and Computer Engineering
University of Victoria
Victoria, BC, Canada
Email: wslu@ece.uvic.ca

Takao Hinamoto

Hiroshima Institute of Technology
Hiroshima, 731-5193, Japan
Email: hinamoto@ieee.org

Abstract—We present a unified approach to minimax design of interpolated and frequency-response-masking FIR filters which are well-known classes of computationally efficient digital filter. The highly nonconvex minimax designs of these filters are carried out by jointly optimizing the subfilters involved using a convex-concave procedure (CCP). We explain why CCP is well-suited for the design problems at hand and present two design instances to demonstrate its performance.

I. INTRODUCTION

Interpolated FIR (IFIR) filters [1], [2] and frequency-response-masking (FRM) FIR filters [3] are well-known classes of computationally efficient digital filters because the total number of multipliers and adders required to implement these filters are considerably less than those required by their conventional counterparts [4]. There has been a great deal of work in the literature following the aforementioned original development, see for example [5]–[11] and the references therein. In this paper we propose a simple design technique that applies to both IFIR and FRM filters, where the subfilters involved are jointly optimized in the minimax sense. The core of the proposed unified design approach is the convex-concave procedure (CCP) [12]–[14] that allows the use of convex optimization to deal with nonconvex designs where the objective functions assume the form of difference of convex functions. We explain why the CCP is well-suited for the designs at hand and present two design examples to demonstrate its performance. Due to space limitation, our focus here is on the design of original IFIR and one-stage FRM filters. However, the design methodology developed here can readily be extended to cover a variety of IFIR and FRM variants.

II. PRELIMINARIES

A. Interpolated FIR Filters

IFIR filters are introduced in [1] and further investigated in [2]. As shown in Fig. 1, an IFIR filter is composed of a cascade of two FIR filters, whose transfer function assumes the form

$$H(z) = F(z^L)M(z) \quad (1)$$

where $L > 0$ is an integer that determines the degree of the filter's sparsity, hence its computational efficiency. Suppose the parent transfer function $F(z)$ represents a lowpass filter,

then $F(z^L)$ is a periodic filter whose widths of passband and transition bands are $1/L$ -th of those of $F(z)$. Consequently, by cascading $F(z^L)$ with a lowpass $M(z)$ (known as interpolator) that suppresses the undesired passbands of $F(z^L)$, a lowpass FIR filter with narrow passband and sharp transition band can be obtained with improved computational complexity relative to that of a conventional FIR with comparable approximation accuracy.

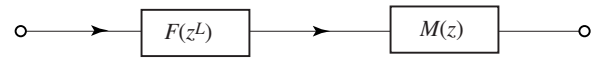


Fig. 1. An IFIR filter.

B. Frequency-Response-Masking Filters

Another class of FIR filters, known as FRM filters with very narrow transition bands is proposed in [3] and has since been a subject of study [5]–[10]. As illustrated in Fig. 2, a single

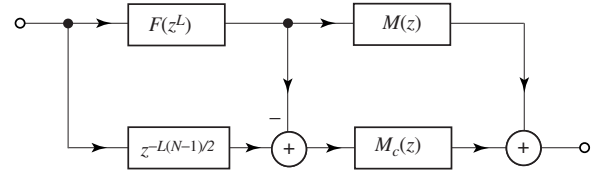


Fig. 2. A single-stage FRM filter.

stage FRM filter has a connected parallel structure where the periodic filter $F(z^L)$ and its delay-complementary periodic filters, $z^{L(N-1)/2} - F(z^L)$, are cascaded with masking filters $M(z)$ and $M_c(z)$, respectively, so as to produce an FIR filter with sharper transition bands. The transfer function of the FRM filter in Fig. 2 is given by

$$H_{FRM}(z) = F(z^L)M(z) + [z^{L(N-1)/2} - F(z^L)]M_c(z)$$

where N denotes the length of $F(z)$ and is assumed to be an odd integer throughout.

C. Convex-Concave Procedure (CCP)

As well become clear in Sec. III, jointly optimizing all sub-filters involved in an IFIR or FRM filter is *not* a convex problem. The CCP is a heuristic method of convexifying non-convex problems. It is known that a function with continuous

second-order derivatives can be expressed as a difference of two convex functions [15] and, if the function in question has a bounded Hessian, such an expression can be explicitly constructed [12]. We are thus motivated to consider a nonconvex problem of the form

$$\text{minimize } f(\mathbf{x}) - g(\mathbf{x}) \quad (2a)$$

$$\text{subject to: } f_i(\mathbf{x}) \leq g_i(\mathbf{x}) \quad \text{for } i = 1, 2, \dots, m \quad (2b)$$

where all functions $f(\mathbf{x})$, $f_i(\mathbf{x})$, $g(\mathbf{x})$, and $g_i(\mathbf{x})$ are convex. A basic CCP algorithm for solving (2) includes two key steps: (i) In the k th iteration, convexify problem (2) by replacing $g(\mathbf{x})$ and $g_k(\mathbf{x})$ with their affine approximations

$$\hat{g}(\mathbf{x}, \mathbf{x}_k) = g(\mathbf{x}_k) + \nabla g(\mathbf{x}_k)^T (\mathbf{x} - \mathbf{x}_k) \quad (3a)$$

and

$$\hat{g}_i(\mathbf{x}, \mathbf{x}_k) = g_i(\mathbf{x}_k) + \nabla g_i(\mathbf{x}_k)^T (\mathbf{x} - \mathbf{x}_k) \quad (3b)$$

(ii) Solve the convex problem

$$\text{minimize } f(\mathbf{x}) - \hat{g}(\mathbf{x}, \mathbf{x}_k) \quad (4a)$$

$$\text{subject to } f_i(\mathbf{x}) \leq \hat{g}_i(\mathbf{x}, \mathbf{x}_k), \quad 1 \leq i \leq m \quad (4b)$$

Utilizing the convexity of all the functions involved in (2), it can be shown that the basic CCP outlined above is a descent algorithm [13] and the iterates $\{\mathbf{x}_k\}$ generated by solving (4) converge to a critical point of the original problem (2) [14]. Note that the CCP requires a feasible initial point $\mathbf{x}_0$ (that satisfies (2b)) to start the procedure.

III. MINIMAX DESIGN OF IFIR FILTERS

A. Problem Formulation

From (1), the frequency response of an IFIR filter is given by

$$H(e^{j\omega}) = F(e^{jL\omega})M(e^{j\omega}) \quad (5)$$

Assume both $F(z)$ and $M(z)$ are of linear phase response, the zero-phase frequency response of $H(z)$, denoted by $H_o(\omega)$, is given by

$$H_o(\omega) = [\mathbf{a}_f^T \mathbf{t}_f(L\omega)][\mathbf{a}_m^T \mathbf{t}_m(\omega)] \quad (6)$$

where $\mathbf{a}_f$ and $\mathbf{a}_m$ are coefficient vectors determined by the impulse responses of $F(z)$ and $M(z)$, respectively, and $\mathbf{t}_f(\omega)$ and $\mathbf{t}_m(\omega)$ are vectors with trigonometric components determined by the filter lengths and types (e.g. odd or even length and symmetrical or antisymmetrical filter coefficients).

Let $a_d(\omega)$ be the desired zero-phase response of the IFIR filter, the frequency-weighted minimax design of an IFIR filter amounts to finding $\mathbf{a}_f$ and $\mathbf{a}_m$ that solves the unconstrained problem

$$\text{minimize}_{\mathbf{a}_f, \mathbf{a}_m} \max_{\omega \in \Omega} w(\omega) |[\mathbf{a}_f^T \mathbf{t}_f(L\omega)][\mathbf{a}_m^T \mathbf{t}_m(\omega)] - a_d(\omega)| \quad (7)$$

where $w(\omega) > 0$ is a frequency-selective weight over $\omega \in \Omega$. Evidently, (7) is nonconvex with respect to design variables $\mathbf{a}_f$ and $\mathbf{a}_m$.

B. Convexification of (7) using CCP

By introducing an upper bound δ for the objective in (7) and treating it as an additional design variable, (7) becomes

$$\text{minimize } \delta \quad (8a)$$

$$\text{s.t. } [\mathbf{a}_f^T \mathbf{t}_f(L\omega)][\mathbf{a}_m^T \mathbf{t}_m(\omega)] \leq \delta_w + a_d(\omega) \quad (8b)$$

$$-[\mathbf{a}_f^T \mathbf{t}_f(L\omega)][\mathbf{a}_m^T \mathbf{t}_m(\omega)] \leq \delta_w - a_d(\omega) \quad (8c)$$

for $\omega \in \Omega$, where $\delta_w = \delta/w(\omega)$. Bearing CCP in mind, we add nonnegative term $q(\mathbf{x}, \omega) = \frac{1}{2}[(\mathbf{a}_f^T \mathbf{t}_f(L\omega))^2 + (\mathbf{a}_m^T \mathbf{t}_m(\omega))^2]$ to both sides of (8b) and (8c) to obtain an equivalent pair of constraints as

$$[\mathbf{a}_f^T \mathbf{t}_f(L\omega) + \mathbf{a}_m^T \mathbf{t}_m(\omega)]^2 \leq p(\mathbf{x}, \omega) + 2\delta_w + 2a_d(\omega) \quad (9a)$$

$$[\mathbf{a}_f^T \mathbf{t}_f(L\omega) - \mathbf{a}_m^T \mathbf{t}_m(\omega)]^2 \leq p(\mathbf{x}, \omega) + 2\delta_w - 2a_d(\omega) \quad (9b)$$

where $p(\mathbf{x}, \omega) = 2q(\mathbf{x}, \omega)$ with $\mathbf{x}^T = [\mathbf{a}_f^T \quad \mathbf{a}_m^T]$. Clearly, the functions on both sides of (9a) and (9b) are convex with respect to $\mathbf{x}$ and δ , therefore (9) is suitable for application of CCP. Note that the constraints in (9) can be expressed as

$$[(\mathbf{a}_f^T \mathbf{t}_f(L\omega)) + (-1)^i (\mathbf{a}_m^T \mathbf{t}_m(\omega))]^2 \leq \eta_i(\mathbf{x}, \mathbf{x}_k, \omega) \quad i = 0, 1 \quad (10)$$

where

$$\eta_i = p(\mathbf{x}_k, \omega) + \nabla p(\mathbf{x}_k, \omega)^T (\mathbf{x} - \mathbf{x}_k) + 2\delta_w + (-1)^i 2a_d(\omega)$$

hence the problem at hand in the k th iteration is to minimize upper bound δ subject to the constraints in (10). By defining

$$\mathbf{t}_i(\omega) = \begin{bmatrix} \mathbf{t}_f(L\omega) \\ (-1)^i \mathbf{t}_m(\omega) \end{bmatrix} \quad \text{for } i = 0, 1$$

the above problem can be cast as the SDP problem [16]

$$\text{minimize } \delta \quad (11a)$$

$$\text{subject to } \begin{bmatrix} \eta_i(\mathbf{x}, \mathbf{x}_k, \omega) & \mathbf{t}_i^T(\omega) \mathbf{x} \\ \mathbf{t}_i^T(\omega) \mathbf{x} & 1 \end{bmatrix} \succeq \mathbf{0} \quad (11b)$$

for $\omega \in \Omega_d, i = 0, 1$

where $\Omega_d = \{\omega_j, j = 1, 2, \dots, K\} \subset \Omega$ is a set of frequency grids that are sufficiently dense and placed uniformly over the frequency region of interest Ω . As mentioned earlier (see Sec. II.C), the iterates $\mathbf{x}_k$ generated from above design framework are guaranteed to converge.

C. Exact convexification of (8)

We now conclude this section with a remark on the exactness of convexifying the constraints in (8b) and (8c). By adding function $q(\mathbf{x}, \omega)$ to (8a) and (8b), (9a) and (9b) fit into (2b) nicely because all four functions on both sides of (9a) and (9b) are convex. Note that the Hessians of the terms on the left-hand sides of (9) are merely positive *semidefinite* (because there are $\mathbf{a}_f$ and $\mathbf{a}_m$ to make these terms zero for a given frequency ω). This means that adding anything less than $q(\mathbf{x}, \omega)$ to (8b) and (8c) will fail to produce equivalent constraints that fit into (2b). It is in this regard we refer the above process of convexifying (8b) and (8c) to as *exact*. We remark that achieving exact convexification is of importance because adding an unnecessarily “heavy” convex term to

convexity a constraint will lead to an unnecessarily loose convex constraint after its right-hand side is linearized, which will inevitably effect the design performance in a negative manner.

IV. MINIMAX DESIGN OF FRM FILTERS

A. The Design Problem

The zero-phase frequency response of a single-level FRM filter is given by

$$H(\mathbf{x}, \omega) = [\mathbf{a}_f^T \mathbf{t}_f(L\omega)] [\mathbf{a}_m^T \mathbf{t}_m(\omega) - \mathbf{a}_c^T \mathbf{t}_c(\omega)] + \mathbf{a}_c^T \mathbf{t}_c(\omega) \quad (12)$$

where $\mathbf{a}_f$, $\mathbf{a}_m$, and $\mathbf{a}_c$ are coefficient vectors associated with filters $F(z^L)$, $M(z)$ and $M_c(z)$ in Fig. 2 and $\mathbf{t}_f(\omega)$, $\mathbf{t}_m(\omega)$, and $\mathbf{t}_c(\omega)$ are respective trigonometric vectors determined by the lengths and types of the FIR filters involved. Given a desired zero-phase frequency response $H_d(\omega)$, the frequency-weighted minimax design seeks to find $\mathbf{x} = [\mathbf{a}_f^T \ \mathbf{a}_m^T \ \mathbf{a}_c^T]^T$ by solving the problem

$$\text{minimize}_{\mathbf{x}} \max_{\omega \in \Omega} w(\omega) |H(\mathbf{x}, \omega) - H_d(\omega)| \quad (13)$$

where $w(\omega) > 0$ is a frequency-selective weight defined over a frequency region of interest, Ω . From (12), it is evident that (13) is a nonconvex problem with respect to design variable $\mathbf{x}$.

B. A CCP approach to solving (13)

By bounding the objective in (13) from above by δ and treating it as an additional design variable, we arrive at

$$\text{minimize } \delta \quad (14a)$$

$$\text{s.t. } H(\mathbf{x}, \omega) - \delta_w - H_d(\omega) \leq 0, \ \omega \in \Omega \quad (14b)$$

$$-H(\mathbf{x}, \omega) - \delta_w + H_d(\omega) \leq 0, \ \omega \in \Omega \quad (14c)$$

where $\delta_w = \delta/w(\omega)$. Defining

$$p(\mathbf{x}, \omega) = [\mathbf{a}_f^T \mathbf{t}_f(L\omega)]^2 + \frac{1}{2}[\mathbf{a}_m^T \mathbf{t}_m(\omega)]^2 + \frac{1}{2}[\mathbf{a}_c^T \mathbf{t}_c(\omega)]^2$$

and adding $p(\mathbf{x}, \omega)$ to both sides of (14b) and (14c), we obtain an equivalent pair of constraints as

$$u(\mathbf{x}, \omega) \leq p(\mathbf{x}, \omega) \quad (15a)$$

$$v(\mathbf{x}, \omega) \leq p(\mathbf{x}, \omega) \quad (15b)$$

where

$$u(\mathbf{x}, \omega) = p(\mathbf{x}, \omega) + H(\mathbf{x}, \omega) - \delta_w - H_d(\omega)$$

$$v(\mathbf{x}, \omega) = p(\mathbf{x}, \omega) - H(\mathbf{x}, \omega) - \delta_w + H_d(\omega)$$

are convex with respect to $\mathbf{x}$ and δ . This is because

$$\begin{aligned} p(\mathbf{x}, \omega) + H(\mathbf{x}, \omega) &= \\ \frac{1}{2}[\mathbf{a}_f^T \ \mathbf{a}_m^T] \mathbf{M}_0 [\mathbf{a}_f^T \ \mathbf{a}_m^T]^T &+ \frac{1}{2}[\mathbf{a}_f^T \ \mathbf{a}_c^T] \mathbf{N}_1 [\mathbf{a}_f^T \ \mathbf{a}_c^T]^T + \mathbf{a}_c^T \mathbf{t}_c \\ p(\mathbf{x}, \omega) - H(\mathbf{x}, \omega) &= \\ \frac{1}{2}[\mathbf{a}_f^T \ \mathbf{a}_m^T] \mathbf{M}_1 [\mathbf{a}_f^T \ \mathbf{a}_m^T]^T &+ \frac{1}{2}[\mathbf{a}_f^T \ \mathbf{a}_c^T] \mathbf{N}_0 [\mathbf{a}_f^T \ \mathbf{a}_c^T]^T - \mathbf{a}_c^T \mathbf{t}_c \end{aligned}$$

where $\mathbf{M}_i = \mathbf{m}_i \mathbf{m}_i^T$ and $\mathbf{N}_i = \mathbf{n}_i \mathbf{n}_i^T$ for $i = 0, 1$ with

$$\mathbf{m}_i = \begin{bmatrix} \mathbf{t}_f(L\omega) \\ (-1)^i \mathbf{t}_m(\omega) \end{bmatrix} \text{ and } \mathbf{n}_i = \begin{bmatrix} \mathbf{t}_f(L\omega) \\ (-1)^i \mathbf{t}_c(\omega) \end{bmatrix}$$

Note that $p(\mathbf{x}, \omega)$ itself is also convex w.r.t. $\mathbf{x}$, hence the constraints in (15) fit nicely into CCP which in turn lead (15) to a pair of convex constraints by linearizing $p(\mathbf{x}, \omega)$ in a vicinity of the k th iterate $\mathbf{x}_k$ to $\tilde{p}(\mathbf{x}, \mathbf{x}_k, \omega)$ which is given by

$$\tilde{p}(\mathbf{x}, \mathbf{x}_k, \omega) = p(\mathbf{x}_k, \omega) + \nabla^T p(\mathbf{x}_k, \omega)(\mathbf{x} - \mathbf{x}_k)$$

where

$$\begin{aligned} \nabla^T p(\mathbf{x}_k, \omega) &= \\ [2(\mathbf{a}_f^T \mathbf{t}_f(L\omega)) \mathbf{t}_f^T(L\omega) \ (\mathbf{a}_m^T \mathbf{t}_m(\omega)) \mathbf{t}_m^T(\omega) \ (\mathbf{a}_c^T \mathbf{t}_c(\omega)) \mathbf{t}_c^T(\omega)] \end{aligned}$$

The convexity of $p(\mathbf{x}, \omega)$ ensures that $p(\mathbf{x}, \omega) \geq \tilde{p}(\mathbf{x}, \mathbf{x}_k, \omega)$, hence the design problem at hand can be relaxed to a sequence of convex subproblems of the form

$$\text{minimize } \delta \quad (16a)$$

$$\text{s.t. } u(\mathbf{x}, \omega_i) \leq \tilde{p}(\mathbf{x}, \mathbf{x}_k, \omega_i), \ 1 \leq i \leq K \quad (16b)$$

$$v(\mathbf{x}, \omega_i) \leq \tilde{p}(\mathbf{x}, \mathbf{x}_k, \omega_i), \ 1 \leq i \leq K \quad (16c)$$

where $\{\omega_i, i = 1, 2, \dots, K\} \subset \Omega$ is a set of K frequency grids that are uniformly placed over Ω . Since both $u(\mathbf{x}, \omega)$ and $v(\mathbf{x}, \omega)$ are convex quadratic and $\tilde{p}(\mathbf{x}, \mathbf{x}_k, \omega_i)$ is linear w.r.t. $\mathbf{x}$, (16) can readily be cast as an SDP problem [16]. Finally, we remark that with an argument similar to that in Sec. III.C, the above process of convexifying the constraints in (14b) and (14c) can be shown to be exact.

V. DESIGN EXAMPLES

A. Example 1

The CCP-based algorithm proposed in Sec. III was applied to design a lowpass IFIR filter with normalized passband edge $\omega_p = 0.15\pi$, stopband edge $\omega_a = 0.2\pi$, sparsity factor $L = 4$, and orders of $F(z)$ and $M(z)$ being 31 and 16, respectively. The frequency weight $w(\omega)$ was set to $w(\omega) \equiv 1$ for ω in the passband and $w(\omega) \equiv 2$ for ω in the stopband. A total of $K = 1400$ frequency grids were uniformly placed in $[0, \omega_p] \cup [\omega_a, \pi]$. It took the algorithm 54 iterations to converge to an IFIR filter whose amplitude response is depicted in Fig. 3. The largest passband ripple and smallest stopband attenuation were found to be 1.9750×10^{-3} and 60.1411 dB, respectively. For comparison, probably the best results found in the literature were achieved by the method in [2], the reader is referred to Example 10.29 of text [4] which describes a design of IFIR filter based on [2] with the same specifications and practically the same design performance as that reported above.

B. Example 2

The CCP-based algorithm proposed in Sec. IV was applied to design a lowpass FRM filter with the same design specifications as in the first example in [3] and [10] with the lengths of $F(z)$, $M(z)$, and $M_c(z)$ being 45, 41, and 33, respectively, $L = 9$, $\omega_p = 0.6\pi$ and $\omega_a = 0.61\pi$. A trivial weight $w(\omega) \equiv 1$ was utilized. With $K = 1100$, it took the algorithm 60 iterations to converge to an FRM filter whose amplitude response is shown in Fig. 4. The largest passband ripple and smallest stopband attenuation were found to be 0.0668 and 42.44 dB, respectively, which are favorably compared with

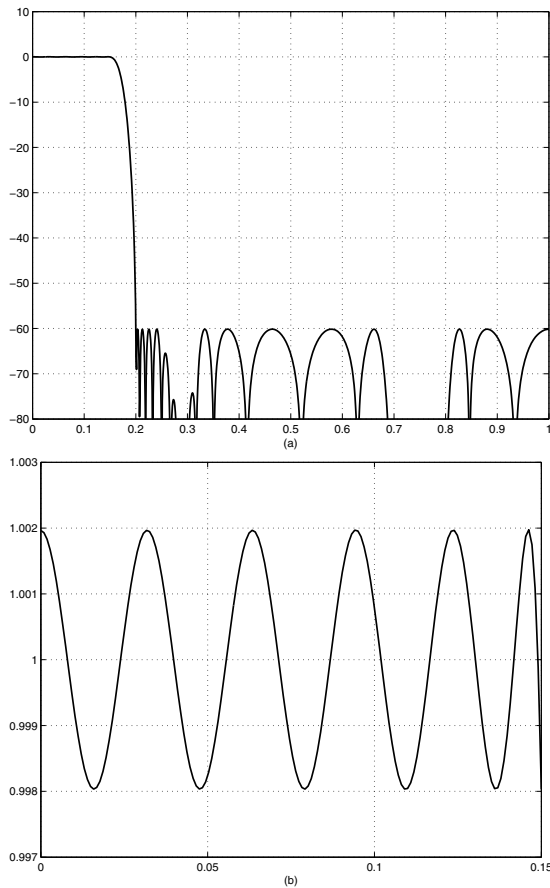


Fig. 3. Amplitude response of the IFIR filter in Example 1 over (a) the entire baseband and (b) the passband.

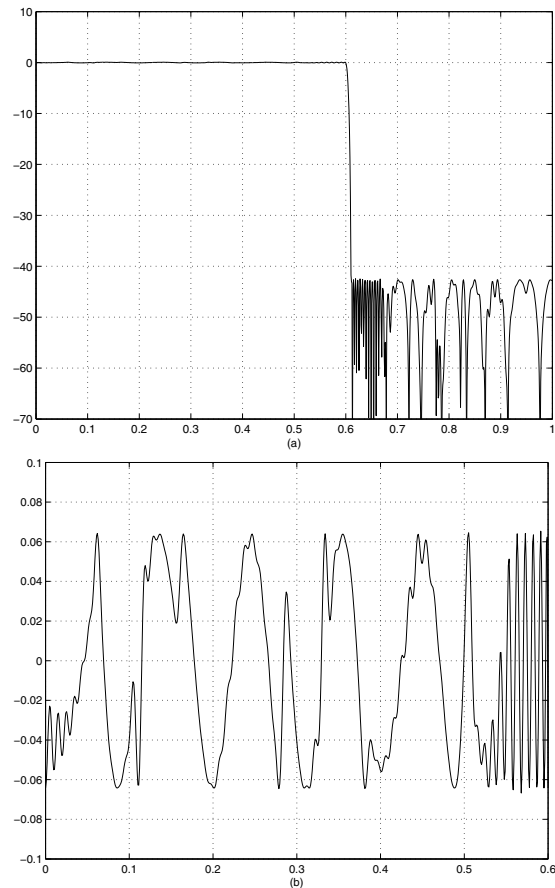


Fig. 4. Amplitude response of the FRM filter in Example 2 over (a) the entire baseband and (b) the passband.

those achieved in [3] (0.0896 and 40.96dB), which has been a benchmark for performance comparison, and those reported in [10] (0.0674 and 42.25dB) which are among the best in the literature.

VI. CONCLUSION

We have proposed a CCP-based unified approach to the design of minimax IFIR and FRM filters, which is conceptually simple and produces designs that are comparable with, sometimes even better than, the best designs available from the literature.

REFERENCES

- [1] Y. Neuvo, C. Y. Dong, and S. K. Mitra, "Interpolated finite impulse response filters," *IEEE Trans. ASSP*, vol. 32, pp. 563–570, June 1984.
- [2] T. Saramäki, Y. Neuvo, and S. K. Mitra, "Design of computationally efficient interpolated FIR filters," *IEEE Trans. CAS*, vol. 35, pp. 70–88, Jan. 1988.
- [3] Y. C. Lim, "Frequency-response masking approach for the synthesis of sharp linear phase digital filters," *IEEE Trans. CAS*, vol. 33, pp. 357–364, April 1986.
- [4] S. K. Mitra, *Digital Signal Processing — A Computer-Based Approach*, 3rd ed., McGraw Hill, 2006.
- [5] Y. C. Lim and Y. Lian, "The optimum design of one- and two-dimensional FIR filters using the frequency-response masking technique," *IEEE Trans. CAS II*, vol. 40, pp. 88–95, Feb. 1993.
- [6] Y. C. Lim and Y. Lian, "Frequency-response masking approach for digital filter design: Complexity reduction via masking filter factorization," *IEEE Trans. CAS II*, vol. 41, pp. 518–525, Aug. 1994.
- [7] T. Saramäki, Y. C. Lim, and R. Yang, "The synthesis of half-band filter using frequency-response masking technique," *IEEE Trans. CAS II*, vol. 42, pp. 58–60, Jan. 1995.
- [8] T. Saramäki and H. Johansson, "Optimization of FIR filters using frequency response masking approach," in *Proc. 2001 ISCAS*, vol. II, pp. 177–180, May 2001.
- [9] L. C. R. de Barcellos, S. L. Netto, and P. S. R. Diniz, "Optimization of FRM filters using the WLS-Chebyshev approach," *Circuits, Systems and Signal Processing*, vol. 22, no. 2, pp. 99–113, March 2003.
- [10] W.-S. Lu and T. Hinamoto, "Optimal design of frequency-response-masking filters using semidefinite programming," *IEEE Trans. CAS I*, vol. 50, pp. 557–568, April 2003.
- [11] Y. Wei and D. Liu, "Improved design of frequency-response masking filters using band-edge shaping filter with non-periodical frequency response," *IEEE Trans. Signal Processing*, vol. 61, no. 13, pp. 3269–3278, July 2013.
- [12] A. L. Yuille and A. Rangarajan, "The concave-convex procedure," *Neural Computation*, vol. 15, no. 4, pp. 915–936, 2003.
- [13] T. Lipp and S. Boyd, "Variations and extensions of the convex-concave procedure," *Research Report*, Stanford University, Aug. 2014.
- [14] G. R. Lanckreijdt and B. K. Sriperumbudur, "On the convergence of the concave-convex procedure," in *Advances in Neural Information Processing Systems*, pp. 1759–1767, 2009.
- [15] P. Hartman, "On functions representable as a difference of convex functions," *Pacific J. of Math.*, vol. 9, no. 3, pp. 707–713, 1959.
- [16] A. Antoniou and W.-S. Lu, *Practical Optimization: Algorithms and Engineering Applications*, Springer, 2007.