DESIGN OF HIGHLY SELECTIVE QUASI-EQUIRIPPLE FIR LOWPASS FILTERS WITH APPROXIMATELY LINEAR PHASE AND VERY LOW GROUP DELAY

Thomas Kurbiel, Daniel Alfsmann and Heinz G. Göckler

Digital Signal Processing Group (DISPO), Ruhr-Universität Bochum, 44780 Bochum, Germany
phone: +49 234 32 22106, fax: +49 234 32 02106, email: {kurbiel, alfsmann, goeckler}@nt.rub.de, web: www.dsv.rub.de

ABSTRACT

In this contribution, the design of approximately linear-phase, low delay and quasi-equiripple FIR lowpass filters is resumed with particular emphasis on very low group delay. To this end, a novel unconstrained single-objective multi-criterion design algorithm is proposed. Results show that the mean group delay can be diminished down to roughly 20% of that of a corresponding linear-phase FIR filter. Moreover, the application of the proposed algorithm to the design of power complementary (square-root Nyquist) filters is demonstrated.

1. INTRODUCTION

Digital linear-phase (LP) FIR filters have been widely used in all fields of digital signal processing due to their robust stability, simplicity of implementation, and the absence of phase distortions. For some specific applications, however, the (normalised) group delay $\tau_g = (N-1)/2$ of standard LP FIR filters, where N represents the filter length, is unacceptably high. Very low group delay is a challenge, for instance, in case of signal processing in digitally implemented hearing aids, where the difference between the direct and the processed sound signals must not exceed a critical value [1]. Moreover, long delay is prohibitive, for instance, in two-way communication systems for intelligibility, and in feedback control systems for stability reasons.

A first idea of approaching the design of highly selective LD digital filters is to use standard recursive or non-recursive minimum-phase filters, respectively. However, the passband group delay distortion of these filter classes is frequently unacceptable – and group delay equalisation additionally counteracts to our design goal of very low group delay. Moreover, it is suspected that recursive minimum-phase filters (especially narrow-band ones) possess a prohibitively high minimum group delay being caused by the positive group delay contributions of the non-zero singularities of IIR filters within the *z*-plane unit circle [2]. Hence, in this investigation, the design of selective digital filters with very low group delay is focused on modified minimum-phase FIR filters, where the poles of the above IIR filters are replaced with zeros reducing, in contrast to the IIR counterpart, the passband group delay. Part of this group delay advantage is, however, lost again, since FIR filter orders are known to generally exceed those of corresponding IIR filters [2].

In the past, many attempts to the design of FIR lowpass filters have been made to approximate a desired magnitude response in conjunction with a linear or non-linear phase or, alternatively, with an approximately constant passband group delay response [3, 4, 5, 6]. In [7], a procedure for the design of FIR Nyquist filters with low group delay was proposed that is based on the Remez exchange algorithm. With these approaches a filter group delay can always be obtained that ranges below the group delay of a corresponding LP FIR filter. However, the minimisation of the minimum or mean value of the passband group delay was of no concern. In contrast, for instance, Lang [8] has shown that his algorithms for the constrained design of digital filters with arbitrary magnitude and phase responses have the potential to achieve a considerable reduction of group delay as compared to LP FIR filters, even for high order FIR filters.

The goal of this contribution is to present an alternative unconstrained approach to the design of highly selective FIR filters with *very low* and approximately constant group delay, and quasi-equiripple magnitude response. To this end, a novel multi-criterion objective function is proposed for optimisation allowing for individual weighting of the involved error functions. This multi-criterion objective function considers the deviation of the filter magnitude responses in passband and stopband and, in passband only, both the deviations of phase from linearity and of group delay from zero.

Subsequently, the multi-criterion objective function is introduced and discussed in section II. In section III, three illustrative design examples are presented and discussed, while in section IV some conclusions are drawn.

2. OPTIMISATION APPROACH

The design of FIR lowpass filters with approximately linear phase and very low group delay is first formulated as a multi-objective unconstrained optimisation problem. In practice, however, this problem is solved by combining all multiple objectives into one scalar objective function and applying standard unconstrained optimisation techniques. The independent parameters of the objective functions are the FIR filter coefficients.

2.1 Multi-Objective Formulation

The frequency response of a real-valued digital FIR filter of length N is characterized by the scalar product [2]:

$$H(e^{j\Omega}) = \sum_{n=0}^{N-1} h_n e^{-jn\Omega} = \mathbf{h}^T \mathbf{c}, \tag{1}$$

where the N-dimensional coefficient vector

$$\mathbf{h}^T = [h_0, h_1, \cdots, h_{N-1}] \tag{2}$$

comprises the FIR filter parameters to be optimised, and vector $\mathbf{c}$ contains the corresponding exponential terms of (1).

For the design of a lowpass filter we define the desired filter magnitude response:

$$M(e^{j\Omega}) = \begin{cases} 1, & 0 \le \Omega \le \Omega_p \\ 0, & \Omega_s \le \Omega \le \pi, \end{cases} \tag{3}$$

where Ω_p and Ω_s represent the filter passband and stopband cut-off frequencies. Using (3), we set up *two* related error functions on a dense frequency grid to control the disjoint L_r-norms of the filter magnitude responses in passband and stopband during optimisation by using FFT techniques:

$$e_p(\mathbf{h}) = \left[\sum_{k=0}^{K_p}\left(\left|H\left(e^{j\Omega_k}\right)\right| - M\left(e^{j\Omega_k}\right)\right)^r\right]^{1/r}, \quad (4)$$

$$e_s(\mathbf{h}) = \left[\sum_{k=K_s}^{K/2}\left(\left|H\left(e^{j\Omega_k}\right)\right| - M\left(e^{j\Omega_k}\right)\right)^r\right]^{1/r}. \quad (5)$$

Hence, the frequency grid is defined equidistantly according to $\Omega_k = 2\pi k/K, \quad k = 0, 1, \ldots, K-1$, with K even. Moreover, index p refers to passband, where $\Omega_{K_p} \approx \Omega_p$, and index s denotes stopband with $\Omega_{K_s} \approx \Omega_s$. The exponent r, restricted to an even integer number, is set to $r = 2$ for least squares (L_2) optimisation, whilst $r = 20$ is chosen for quasi-equiripple ($L_{20} \approx L_\infty$) magnitude design [9].

Using the group delay frequency response of $H(e^{j\Omega})$ as given by [2]:

$$\tau_g(e^{j\Omega}, \mathbf{h}) = \frac{-\partial \arg(e^{j\Omega}, \mathbf{h})}{\partial \Omega} = \operatorname{Re}\left\{\frac{\mathbf{h}^T \mathbf{D}\mathbf{c}}{\mathbf{h}^T \mathbf{c}}\right\}, \quad (6)$$

where $\operatorname{Re}\{.\}$ denotes the real part of a complex-valued quantity, and $\mathbf{D} = \operatorname{diag}[0, 1, \cdots, N-1]$ represents an $N \times N$ diagonal matrix, a *third* objective function to be used for the minimisation of the maximum absolute value (L_∞-norm) of the passband group delay is readily formulated:

$$e_\tau(\mathbf{h}) = \max_{0 \le \Omega_k \le \Omega_p}\left|\tau_g(e^{j\Omega_k}, \mathbf{h})\right| = \left\|\tau_g(e^{j\Omega_k}, \mathbf{h})\right\|_\infty. \quad (7)$$

Obviously, in (7), the underlying passband group delay desired function is zero for all frequencies.

As an error function to control phase linearity during the optimisation process, we use the L_r-norm of the deviation of the phase curvature from zero. Hence, this *fourth* objective function is based on the first derivate of the group delay w.r.t. the normalised frequency Ω as follows:

$$\begin{aligned}
e_\phi(\mathbf{h}) &= \left[\sum_{k=0}^{K_p}\left|\frac{\partial \tau_g(e^{j\Omega}, \mathbf{h})}{\partial \Omega}\right|_{|\Omega=\Omega_k}^r\right]^{1/r} \\
&\sim \left[\sum_{k=1}^{K_p}\left|\tau_g\left(e^{j\Omega_k}\right) - \tau_g\left(e^{j\Omega_{k-1}}\right)\right|^r\right]^{1/r}, \quad (8)
\end{aligned}$$

where the frequency-continuous definition of $e_\phi(\mathbf{h})$ is approximated on the dense grid, as defined in (8). It should be noted that the grid density can be adapted to any value, as required. The (minimum) density we use throughout all designs is represented by $K = 1024$.

Finally, to extend the range of applications of our optimisation approach, we introduce a *fifth* error function for the specific design of *power complementary* (square-root Nyquist) FIR filters potentially with low group delay, being used for data transmission and in standard quadrature mirror filter banks. To this end, we control the L_∞-norm of the filter magnitude response at the 3dB-point in the centre of the

transition band, $\Omega_t = (\Omega_p + \Omega_s)/2$, leading to:

$$e_t(\mathbf{h}) = \left|\left|H(e^{j\Omega_t})\right| - \frac{1}{\sqrt{2}}\right| = \left\||H(e^{j\Omega_t})| - \frac{1}{\sqrt{2}}\right\|_\infty. \quad (9)$$

Combining the five error functions (4), (5), (7), (8) and (9) to an objective *vector* function:

$$\mathbf{e}(\mathbf{h}) = \left[e_p(\mathbf{h}), e_s(\mathbf{h}), e_\tau(\mathbf{h}), e_\phi(\mathbf{h}), e_t(\mathbf{h})\right]^T, \quad (10)$$

our design problem is readily formulated as an unconstrained multi-objective least squares (L_2) optimization problem [10]:

$$\min_{\mathbf{h}\in\mathbb{R}^N}\|\mathbf{e}(\mathbf{h})\|_2^2 = \min_{\mathbf{h}\in\mathbb{R}^N}|\mathbf{e}(\mathbf{h})|^2 = \min_{\mathbf{h}\in\mathbb{R}^N}\mathbf{e}^T(\mathbf{h})\mathbf{e}(\mathbf{h}). \quad (11)$$

As a consequence of the non-linear nature of the objective vector function $\mathbf{e}(\mathbf{h})$ in dependence of the coefficient vector $\mathbf{h}$ to be optimised, the minimisation problem (11) is generally multi-modal and, hence, is expected to possess different (possibly many) local minima. Nevertheless, each optimum or sub-optimum solution of the minimisation process (11) shall represent a balanced compromise between the finally remaining error contributions of the five individual objective functions of (10) to the squared Euclidean norm of $\mathbf{e}(\mathbf{h})$. The improvement of an individual objective is always at the expense of some or all other errors of (10).

2.2 Single-Objective Formulation

The multi-objective $L_2/L_r/L_\infty$-norm based formulation of our optimisation problem given in section 2.1 is, in compliance with common practice, transformed to a single-objective form, which is generally better suitable for the application of standard optimisation procedures. To this end, the 5-dimensional error vector (10) is pre-multiplied by a 5-dimensional weighting vector according to:

$$e(\mathbf{h}) = \mathbf{w}^T\mathbf{e}(\mathbf{h}), \quad (12)$$

where $\mathbf{w}^T = [\alpha, \beta, \gamma, \delta, \eta]$, yielding the scalar objective function:

$$e(\mathbf{h}) = \alpha e_p(\mathbf{h}) + \beta e_s(\mathbf{h}) + \gamma e_\tau(\mathbf{h}) + \delta e_\phi(\mathbf{h}) + \eta e_t(\mathbf{h}). \quad (13)$$

Obviously, the individual weights of vector $\mathbf{w}$ have to be selected very carefully to obtain a balanced optimum result of the minimisation procedure. Needless to say that this requires a lot of procedural experience and sure instinct. Moreover, alternate specific selection of $\mathbf{w}$ allows for trading-off individual objectives against others.

The scalar (single-objective) form (13) of the error function is amenable to differentiation w.r.t. the coefficients $\mathbf{h}$ and, hence, to gradient-based optimisation techniques. As a result, for the optimisation of the vector $\mathbf{h}$ of N unknown coefficients subject to the minimisation of (13), we have used the gradient-based problem solver fminunc, which is offered by the MATLAB Optimization Toolbox. Due to the multi-modal nature of our non-linear optimisation problem, we are bound to start each optimisation from a suitable initial set $\mathbf{h}$. Details on finding these initial estimates will be discussed in conjunction with the presentation of the design examples in the subsequent section.

Table 1: Specification and performance of Ex. A versus ref.

Performance	Design	Ref.
Passband Ripple	0.5dB	
Stopband Atten.	50dB	
Filter Length N	25	20
Passband min. τ_g	5.9	9.5
Passband mean τ_g	6.0	9.5
Passband max. τ_g	6.1	9.5

(a) Magnitude responses

(b) Group delay

(c) Zero plot (design)

(d) Zero plot (reference)

Figure 1: Design of Ex. A (solid) versus reference (dashed)

3. DESIGN EXAMPLES

Subsequently, we present and investigate three different design results of an FIR lowpass filter of length $N = 25$ obtained with the single-objective multi-criterion optimisation procedure as described in section 2.2. The passband and stopband cut-off frequencies are fixed to $\Omega_p = 0.4\pi$ and $\Omega_s = 0.6\pi$, respectively. To obtain quasi-equiripple magnitude and phase responses, the exponent r of (4), (5) and (8) is set to $r = 20$. In addition, the impact of the weighting vector $\mathbf{w}^T$ of (12) and (13) on the frequency responses is studied. Each design is compared with a LP or minimum-phase (MP) FIR reference filter with the same magnitude response.

3.1 Example A: Approximately Linear Phase

For Ex. A, the weights are chosen as follows: $\mathbf{w}^T = [1, 15, 0.01, 3, 0]$. As a result, we expect a high stopband attenuation, forced by weight $\beta = 15$ of (5), and a highly linear phase controlled by weight $\delta = 3$ of (8).

The design results of Ex. A are presented in Tab. 1 and depicted in Fig. 1, respectively, along with a LP FIR reference filter that meets the same magnitude specifications. Despite the fact that our design requires a higher filter length (excess order of 5), its group delay ranges completely below the constant value of $\tau_g^{ref} = 9.5$ of the LP reference filter. Especially in passband, the approximately constant group delay amounts to roughly 60% of that of the reference filter only.

As to be seen from Figs. 1(a),(c), the zeros effectuating the stopband of the transfer function $H(z)$ of our design are located slightly off (i.e. inside) the z-plane unit circle. Hence, all zeros within the unit circle (there are no more *group delay neutral* zeros on the periphery of the unit circle) contribute to the overall group delay by at least a small negative amount [2]. Obviously, to maintain the specified stopband rejection, extra zeros inside but close to the unit circle are needed, which explains the excess filter order of our design compared to that of the reference filter.

Finally, it is somewhat unexpected that the mean group delay of $\tau_g^{mean} = 6.0$ of our design is that low despite the very low weight of $\gamma = 0.01$ put on the group delay error function (7). Hence, we suspect that this error function has a high overall impact on the scalar objective function (13). Furthermore, according to our experience gained from many designs with the above fixed weight vector, it should be noted that an increase of the filter length predominantly increases the stopband attenuation of our design, whereas all other filter properties remain essentially unchanged.
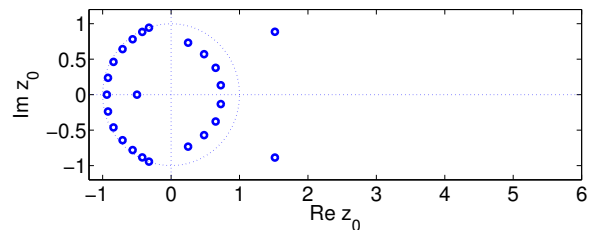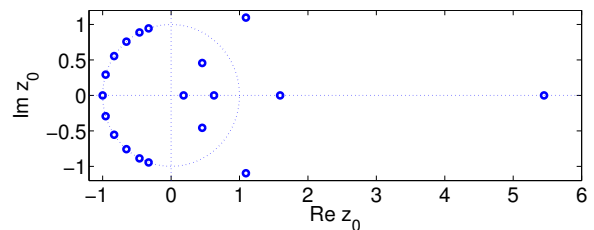
3.2 Example B: Very Low Group Delay

For Ex. B, the weights are chosen as follows: $\mathbf{w}^T = [1, 15, 0.01, 0, 0]$. Here, we expect a lowpass filter with
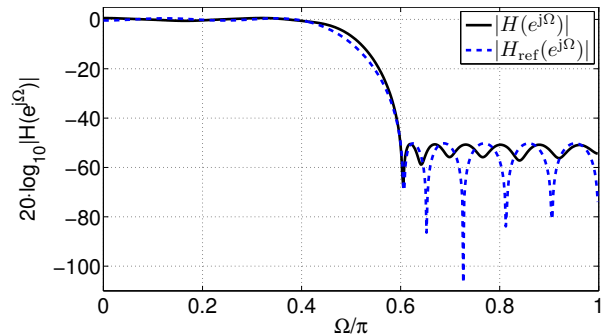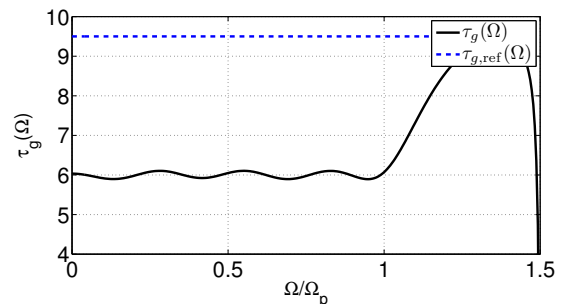
high stopband attenuation, forced by weight $\beta = 15$ of (5), and very low group delay in the passband controlled by weight $\gamma = 0.01$ of (7).

The design results are presented in Tab. 2 and Fig. 2, respectively, along with the corresponding LP FIR reference filter. The very small magnitude deviation in passband is deemed the main reason for the excess filter length of the LP reference by 9, although the stopband zeros of our design are again located inside the z-plane unit circle; cf. Fig. 2(a). Comparing group delay, we have obtained $\tau_g^{mean} = 0.14 \cdot \tau_g^{ref}$, however, at the expense of the loss of phase linearity!
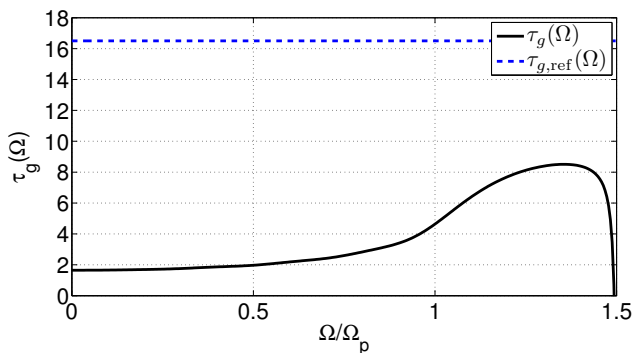
Finally we report that, like in this example, very low group delay is only achievable in connection with very small

Table 2: Specification and performance of Ex. B versus ref.

Performance	Design	Ref.
Passband Ripple	0.01dB	
Stopband Atten.	52dB	
Filter Length N	25	34
Passband min. τ_g	1.7	16.5
Passband mean τ_g	2.3	16.5
Passband max. τ_g	4.7	16.5
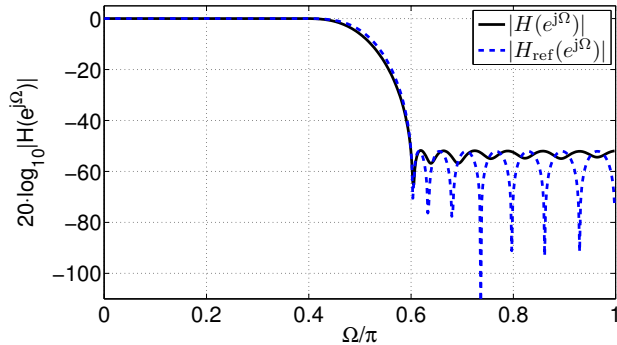
(a) Magnitude responses

(b) Group delay

Figure 2: Design of Ex. B (solid) versus reference (dashed)

magnitude ripple in passband, i.e. reduction of group delay concurrently decreases passband magnitude deviation.

3.3 Example C: Square-Root Nyquist Filter

For Ex. C, the weights are chosen as follows: $\mathbf{w}^T = [1, 1, 0.1, 0, 1]$. Hence, we expect a nearly power complementary FIR square-root Nyquist filter as a result of weight $\eta = 1$ of (9) with low group delay controlled by weight $\gamma = 0.1$ of (7). To support magnitude symmetry w.r.t. $\Omega = \pi$, the same weights $\alpha = \beta = 1$ are put on (4) and (5).

The design results are presented in Tab. 3 and Fig. 3, respectively, along with a MP FIR reference filter that meets the same magnitude specifications [11]. Power complementarity is assessed by the logarithmic distortion function [11]

$$a_{\text{dist}}(e^{j\Omega}) = 20 \cdot \log_{10} \sum_{k=0}^{K/2} \left[\left| H\left(e^{j\Omega_k}\right) \right|^2 + \left| H\left(e^{j(\Omega_k - \pi)}\right) \right|^2 \right],$$

as depicted in Fig. 3(c). The reasonably close similarity of both designs of this example confirms that our approach also

Table 3: Specification and performance of Ex. C versus ref.

Performance	Design	Ref.
Passband Ripple	0.01dB	
Stopband Atten.	43dB	
Filter Length N	25	28
FB-Distortion max.	0.45dB	0.2dB
Passband min. τ_g	1.4	1.5
Passband mean τ_g	1.9	2.0
Passband max. τ_g	3.5	3.7

has the potential to design standard quadrature mirror filters (SQMF) with common quality. Moreover, our design approach overcomes the restriction to even filter lengths in contrast to the approach by Herrmann & Schüssler [11].

3.4 Numerical Properties of the Design Algorithm

Due to the non-linear nature of the single-objective error function $e(\mathbf{h})$, as defined by (12) and (13), the minimisation problem (11) is multi-modal and, hence, there may exist many local minima, where each (sub-)minimum obtained highly depends on the choice of the initial coefficient vector $\mathbf{h}^0$. To overcome this deficiency, we adopt a heuristic approach: We perform each filter design M times, always starting from a different, randomly selected initial coefficient vector $\mathbf{h}_m^0$, $m = 1, \dots, M$, where only the best (sub-)optimum solution is retained ($M = 30 \dots 60$). In future, this *multistart* procedure will be combined with more intelligent strategies of genetic algorithms.

Experience with our design algorithm has shown that, too, the choice of the weight vector $\mathbf{w}^T$ has a high impact on the quality of the final optimum solution. On the one hand, there are weight ratios that make the algorithm diverge since, most probably, no (sub-)optimum solution exists for this very weight vector. This is, for instance, the case, if the group delay requirement is too tight by imposing a too high weight on (7). On the other hand, minor changes of some individual weights may lead to completely different final solutions, or even to divergence, respectively. All these issues call for further detailed investigations in the future.

4. CONCLUSION

In this contribution, an effective and flexible approach to the design of highly selective quasi-equiripple FIR filters with very low group delay and/or approximately linear phase has been proposed. This method is based on the solution of a multi-objective unconstrained optimisation problem allowing for individual weighting of the involved error functions. To this end, three L_{20}-norm and up to two L_{∞}-norm error functions are combined to a weighted least squares (L_2) optimisation approach. The impact of both the individual objective functions and the associated weights have been investigated. Furthermore, we have presented a method to solve the multi-objective optimisation problem by using a MATLAB routine, and provided three examples. Finally, we have discussed the properties of the design algorithm with reference to the examples.

Future investigations will be devoted to the design of narrow-band FIR lowpass filters with very low group delay, and their application in uniform, complex-modulated low delay filter banks. Furthermore, we will compare the poten-

(a) Magnitude responses
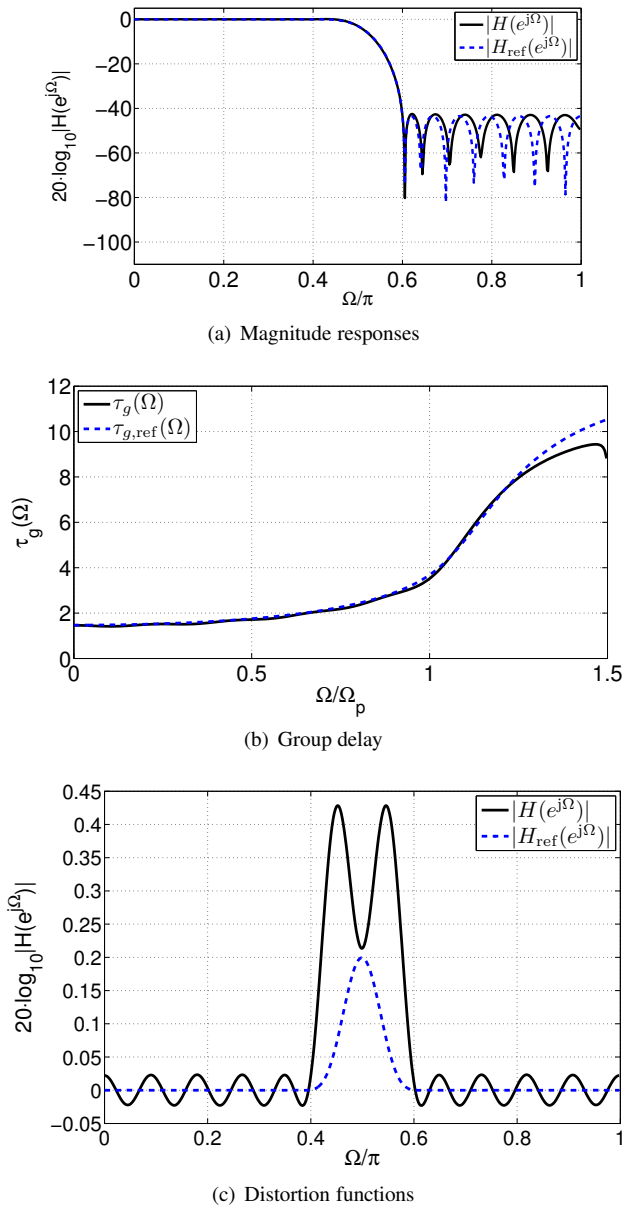

(b) Group delay

(c) Distortion functions

Figure 3: Design of Ex. C (solid) versus reference (dashed)

tial of our unconstrained design approach with that of constrained design algorithms such as those proposed in [8, 12] and, possibly, transform our procedure into a constrained design problem. Finally, we will fathom the assertion of Lang in [8] that IIR filters have the potential of the lowest group delay, possibly ranging below that of FIR filters with very low delay, as presented here.

REFERENCES

[1] M. A. Stone and B. C. J. Moore, "Tolerable hearing aid delays. III. effects on speech production and perception of across-frequency variation in delay," *Ear & Hearing*, vol. 24, no. 2, pp. 175–183, April 2003.

[2] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Prentice-Hall, London, 1989.

[3] T. F. Liau, M. A. Razzak, and L. G. Cuthbert, "Phase constraints on FIR digital filters," *Electronics Letters*, vol. 17, no. 24, pp. 910–911, November 1981.

[4] H. Baher, "FIR digital filters with simultaneous conditions on amplitude and delay," *Electronics letters*, vol. 18, no. 7, pp. 296–297, April 1982.

[5] E. Sharestani and L. G. Cuthbert, "FIR digital filters designed to a group-delay characteristic," *Electronics letters*, vol. 21, no. 12, pp. 542–544, June 1985.

[6] M. Akazawa and M. Ikehara, "Simultaneous approximation of magnitude and group delay in FIR digital filters," *Proc. 12th EUSIPCO 2004, Vienna, Austria*, pp. 117–120, 2004.

[7] X. Zhang and T. Yoshikawa, "Design of FIR Nyquist filters with low group delay," *IEEE Trans. Signal Processing*, vol. 47, no. 5, pp. 1454–1458, May 1999.

[8] M. Lang, "Algorithms for the constrained design of digital filters with arbitrary magnitude and phase responses," Ph.D. dissertation, Vienna Univ. of Technology, Austria, June 1999.

[9] C. S. Burrus and J. A. Barreto, "Least *p*-power error design of FIR filters," *Proc. IEEE Int. Symposium on Circuits and Systems, San Diego, USA*, vol. ISCAS-92, pp. 545–548, May 1992.

[10] J. J. E. Dennis and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Englewood Cliffs, Prentice-Hall, 1983.

[11] H. G. Göckler and A. Groth, *Multiratensysteme*. Schlembach Verlag, Wilburgstetten, Germany, 2004.

[12] K. Tsui, S. Chan, and K. Yeung, "Design of FIR digital filters with prescribed flatness and peak error constraints using second-order cone programming," *IEEE Trans. Circuits Systems II*, vol. 52, pp. 601–605, 2005.