jet INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2012, VOL. 58, NO. 2, PP. 177–192
Manuscript received July 16, 2011; revised April, 2012. DOI: 10.2478/v10177-012-0026-2

Near-Perfect Reconstruction Oversampled Nonuniform Cosine-Modulated Filter Banks Based on Frequency Warping and Subband Merging

Marek Parfieniuk and Alexander Petrovsky

***Abstract*—A novel method for designing near-perfect reconstruction oversampled nonuniform cosine-modulated filter banks is proposed, which combines frequency warping and subband merging, and thus offers more flexibility than known techniques. On the one hand, desirable frequency partitionings can be better approximated. On the other hand, at the price of only a small loss in partitioning accuracy, both warping strength and number of channels before merging can be adjusted so as to minimize the computational complexity of a system. In particular, the coefficient of the function behind warping can be constrained to be a negative integer power of two, so that multiplications related to allpass filtering can be replaced with more efficient binary shifts. The main idea is accompanied by some contributions to the theory of warped filter banks. Namely, group delay equalization is thoroughly investigated, and it is shown how to avoid significant aliasing by channel oversampling. Our research revolves around filter banks for perceptual processing of sound, which are required to approximate the psychoacoustic scales well and need not guarantee perfect reconstruction.**

***Keywords*—warped near-perfect reconstruction oversampled non-uniform cosine-modulated filter bank, allpass filter/transformation, subband/channel merging, frequency warping, critical bands, Bark scale.**

I. Introduction

NON-UNIFORM filter banks (FBs) are of considerable interest in digital signal processing, because they allow the frequency range to be partitioned into unequal subbands, which harmonizes with many real-world signal sources and phenomena, like human hearing. Over the years, a number of approaches have been proposed[1] for constructing non-uniform FBs, which can be broadly classified into the following categories: putting together independent filters with unequal bandwidths [1], arranging uniform FBs into an unbalanced tree of multiresolution signal decomposition [2], [3], mixing subbands of several uniform FBs which work in parallel [4], [5], warping uniform responses [6], [7], recombining outputs of a uniform FB [8] or simply adding them [9]–[11] in order to form wider subbands. Because each of these techniques has both advantages and drawbacks, all find applications, and there is still an interest in developing new subband decompositions which offer different design opportunities.

This work was supported by Bialystok University of Technology under the grant S/WI/4/08.

M. Parfieniuk and A. Petrovsky are with Department of Digital Media and Computer Graphics, Bialystok University of Technology, Wiejska 45A, 15-351 Bialystok, Poland, phone: +48 85 746 90 50 (e-mails: m.parfieniuk@pb.edu.pl; palex@bsuir.by).

[1]For the sake of brevity, only more recent works are cited in this introduction, which provide many more references.

Among the mentioned, warping can be considered most flexible, as obtainable partitionings are not constrained by a uniform grid. As this is especially advantageous in approximating the critical bands of hearing [12], [13], warped systems were used mainly in perceptual processing of sound signals [14]–[20]. Other areas of their applications are multiple description coding [21] and signal analysis [22].

In this paper, we propose to extend frequency warping with subband merging [9], [11], which results in even more design freedom. Namely, warping strength and number of channels before merging can be adjusted so as to reduce the computational complexity of a system without affecting accuracy of frequency partitioning. In particular, the allpass function behind warping can be constrained to use multiplications by negative integer powers of two, which can be implemented efficiently as hardwired binary shifts.

Unlike most authors, who study warped DFT FBs, we consider warped systems with cosine modulation [16]. Their practical advantage is that channel sequences are real-valued for inputs with the same property, which is natural and makes implementation of subband processing easier than in the case of complex values [23]. Nevertheless, some of the presented design insights apply also to warped DFT FBs. Similarly, even though our idea is introduced in the context of approximating the critical bands, it applies to other specifications for frequency partitioning.

The rest of the paper begins with introducing warped cosine-modulated FBs in Section II. In Section III, the idea of using both frequency warping and subband merging is presented, and our design methodology is described. Then, Section IV demonstrates practical examples of obtainable systems. In order to make the paper self-contained, some results of the theory of warped FBs, including authors' contributions, are elaborated in the appendices.

II. Warped Cosine-Modulated Filter Banks

A. Means of Nonuniform Frequency Partitioning

The transfer functions of the analysis and synthesis filters of an M-channel warped cosine-modulated FB can be described using the following expressions:

$$H_k(z) = a_k b_k U_k\big(A^{-1}(z)\big) + \overline{a}_k \overline{b}_k V_k\big(A^{-1}(z)\big), \quad (1a)$$

$$F_k(z) = \overline{a}_k b_k U_k\big(A^{-1}(z)\big) + a_k \overline{b}_k V_k\big(A^{-1}(z)\big), \quad (1b)$$

where $k = 0, \ldots, M-1$,

$$a_k = e^{j(-1)^k \frac{\pi}{4}}, \quad (2)$$

$$b_k = W_{2M}^{-(k+\frac{1}{2})(\frac{L-1}{2})}, \tag{3}$$

$$U_k(z) = P(zW_{2M}^{k+\frac{1}{2}}), \tag{4}$$

$$V_k(z) = P(zW_{2M}^{-(k+\frac{1}{2})}), \tag{5}$$

and W_{2M} denotes the $2M$th root of unity. The overbar indicates complex conjugate.

The real-coefficient lowpass prototype filter $P(z)$ of length L has a cutoff frequency $\omega_c = \pi/2M$ independent of warping strength. As explained in Appendix D-C, it can be designed using methods developed for uniform FBs, see e.g. [24].

The only difference between the above expressions and those for uniform cosine-modulated FBs, see [25], is the inverse allpass function $A^{-1}(z)$, instead of z, in (1). From an implementation point of view, this corresponds to replacing each of unit delays in a given uniform system with a stable and causal allpass filter $A(z)$ [26], [27]. This, so-called, allpass transformation is responsible for deforming the original magnitude responses in accordance with the phase response $\phi(\omega)$ of the filter, as explained in Fig. 1.

The technique resembles the frequency transformations for digital filters [28]. From a mathematical point of view, bilinear conformal mapping of the unit circle onto itself is established, so that the frequency scale undergoes warping[2], and subbands change both their widths and positions. It can be shown that warped subband decompositions are related to signal expansions in terms of Laguerre sequences [15].

An Rth-order allpass filter, whose IIR transfer function can always be factorized into first-order sections:

$$A(z) = \prod_{r=1}^{R} \left(\frac{\overline{\alpha}_r + z^{-1}}{1 + \alpha_r z^{-1}} \right), \tag{6}$$

has the phase response

$$\phi(\omega) = -R\omega + 2\sum_{r=1}^{R} \arctan \frac{|\alpha_r| \sin(\omega - \arg \alpha_r)}{|\alpha_r| \cos(\omega - \arg \alpha_r) - 1}, \tag{7}$$

which is a nonlinear function of frequency ω. The system is stable as long as all $|\alpha_r| \leq 1$.

Higher-order and complex-coefficient warpings offer great design flexibility, as they allow bandwidths to change non-monotonically and to narrow and widen alternately [7], [13], [29], [30]. However, they are not only computationally demanding but also increase both algorithmic delay and memory requirements of FBs, as shown in Appendix A. Thus, first-order real-coefficient allpass filters are much more practical, especially because monotonic increase/decrease in bandwidths they merely provide is sufficient to quite well approximate the psychoacoustic scales [12], [14]. The present paper focuses on first-order warpings and is aimed at proving that by extending them with subband merging, it is possible to greatly enhance design freedom without side-effects.

For clarity, we use α without a subscript to denote the coefficient of a first-order allpass function used to warp an FB. The greater absolute value of the coefficient, the stronger warping, as illustrated in Fig. 1. If $\alpha = 0$, then the allpass filter is simply the unit delay, and hence the corresponding transformation has no effect on the original FB. Negating α causes the reflection of the phase response, so that warping reverses in frequency.

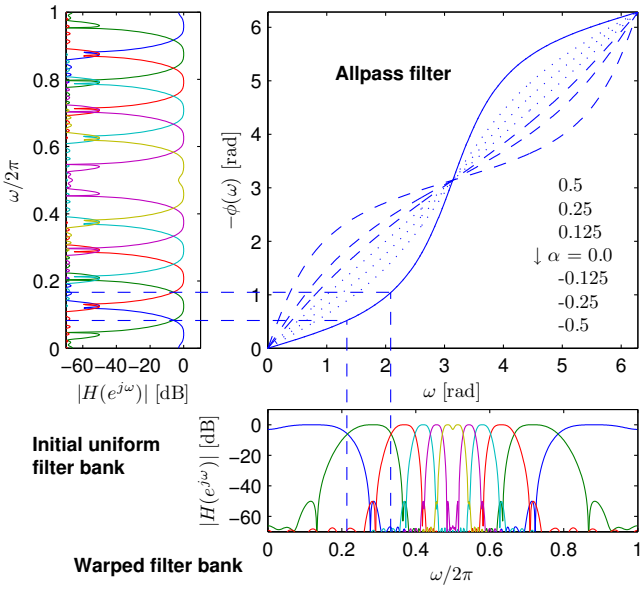

Fig. 1. Essence of warping.

B. Implementation Structure

Figure 2 (a) shows a scheme for implementing the systems of interest[3], which resembles the well-known polyphase structure for uniform modulated FBs [25], as prototype filtering is separated from modulation. The latter is represented using the $M \times 2M$ rectangular matrices:

$$[\mathbf{C}_\mathrm{A}]_{kn} = 2\cos\left[\frac{\pi}{M}\left(k+\tfrac{1}{2}\right)\left(n-\tfrac{L-1}{2}\right) + (-1)^k\tfrac{\pi}{4}\right] \tag{8a}$$

$$[\mathbf{C}_\mathrm{S}]_{kn} = 2\cos\left[\frac{\pi}{M}\left(k+\tfrac{1}{2}\right)\left(2M-1-n-\tfrac{L-1}{2}\right) - (-1)^k\tfrac{\pi}{4}\right], \tag{8b}$$

even though can be implemented using a fast algorithm.

However, there are notable differences compared to uniform FBs. Sampling rates of channel signals are changed outside the system, not at prototype filtering. The chain of $N = L-1$ delays related to the latter becomes an allpas chain. Reconstruction of the fullband signal is followed by a group delay equalization with a filter $C^N(z)$, which is equivalent to a series of N copies of a filter $C(z)$ that reverses the phase modification caused by a single allpass filter $A(z)$:

$$A(z)C(z) \approx z^{-D} \tag{9}$$

for some positive integer D.

In order to make the scheme of Fig. 2 (a) more general, we have supplemented it with delays for avoiding multiple frequency mapping, which are necessary only in the case of a higher-order warping, $R > 1$ in (6), as Appendix A explains.

[2]In some papers, see e.g. [22], the term "frequency warping" is used to refer to frequency mappings that have nothing to allpass filters.

[3]Subband merging is omitted from the picture for clarity.

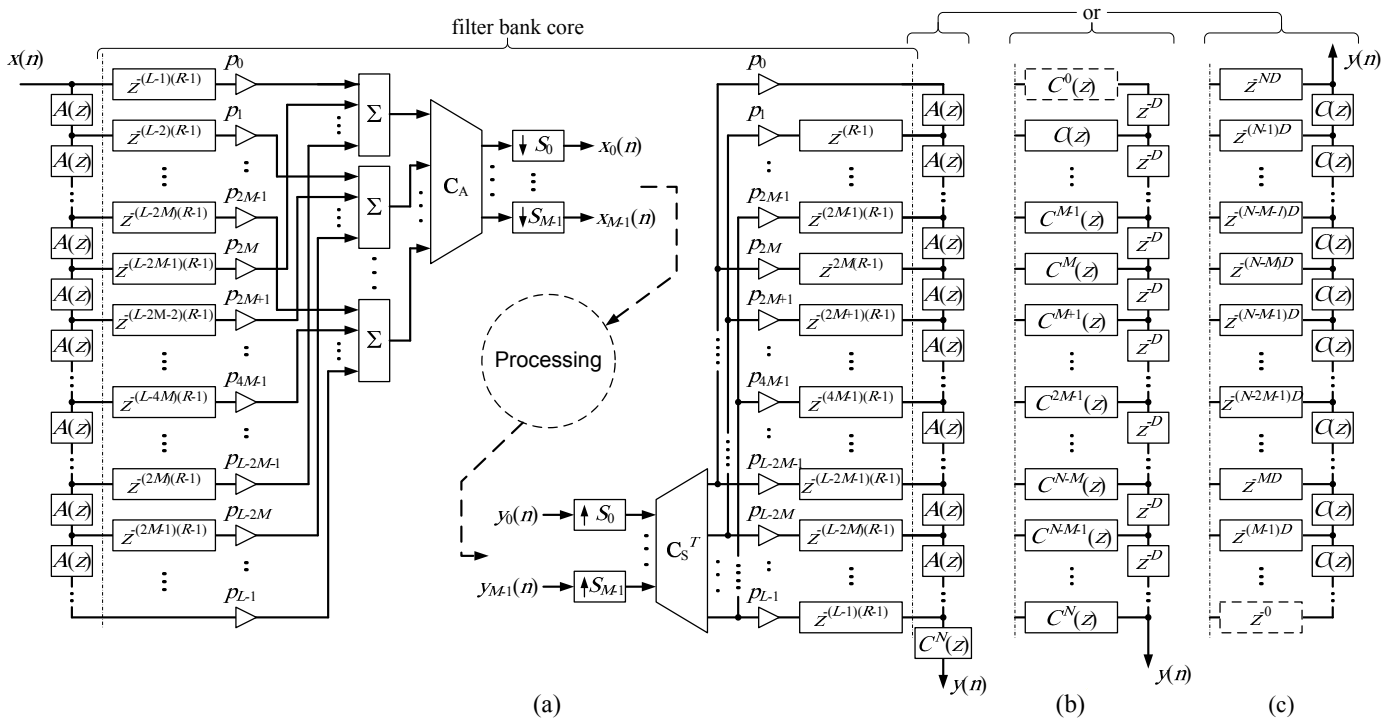

Fig. 2. Structure for warped cosine-modulated FBs and alternative schemes for reconstructing signals passed through an allpass filter chain.

More details on both subsampling and equalization are given in the following sections. Here, it is worth noticing that Figs. 2 (b) and (c) show alternative structures that combine equalization with reconstruction of a fullband signal [31], [32]. Appendix C thoroughly explains why it is better to use the synthesis part based on an allpass chain with equalization at its output, as shown in Fig. 2 (a).

Such a scheme is consistent, memory-efficient, and needs designing only one equalization filter. We exploit its great advantage that aliasing analysis and design of the filter bank core can be separated from optimizing group delay equalization. This can be considered an innovation over the known works, where all these issues are combined into a single, complicated optimization problem [7], [33]–[35]. In our approach, there are more design tasks, but they are simpler and require neither advanced solvers nor heavy computations.

C. Rules for Selecting Subsampling Ratios

In the systems under consideration, critical sampling without causing distortions is impossible. This is because they are nonfeasible-partition FBs [11], in which the cutoff frequencies of subband filters do not match any frequency partitioning given by the spectra replication that is related to reducing the sampling rate by an integer factor. As a result, aliasing caused by subsampling cannot be canceled out at synthesis.

Nevertheless, perfect reconstruction can be well approximated by using a selective prototype to attenuate minor aliasing terms and by preventing significant terms from arising at all. This requires channel subsampling ratios to be constrained in accordance with bandwidths and locations of subbands, which possibly result from merging narrower ones.

In [36], we have derived the following inequalities:

$$\left\lfloor \frac{n_k}{2f_{\mathrm{U}k}} f_{\mathrm{s}} \right\rfloor \geq S_k \geq \left\lceil \frac{n_k - 1}{2f_{\mathrm{L}k}} f_{\mathrm{s}} \right\rceil, \qquad 1 \leq n_k \leq \left\lfloor \frac{f_{\mathrm{U}k}}{f_{\mathrm{U}k} - f_{\mathrm{L}k}} \right\rfloor, \tag{10}$$

which assume that the signal of the kth channel is real-valued, occupies a frequency range from $f_{\mathrm{L}k}$ to $f_{\mathrm{U}k}$ and from $-f_{\mathrm{U}k}$ to $-f_{\mathrm{L}k}$, and is initially sampled at a frequency f_{s}, which then is reduced to $f_{\mathrm{s}k}$. The expressions determine ranges of the subsampling ratio $S_k = f_{\mathrm{s}}/f_{\mathrm{s}k}$ for which there is no significant irreversible aliasing. Of course, the greatest permissible value of S_k is of interest to us.

It is also noteworthy that both lower and upper frequencies, $f_{\mathrm{L}k}$ and $f_{\mathrm{U}k}$, respectively, must incorporate (at least partially) transition bandwidths after warping, which vary with frequency, as illustrated in Fig. 1.

The approach is described in more details in Appendix D. Even though it is necessary because of the properties of warped systems, it also follows the current trend of avoiding aliasing by oversampling in applications where severe modifications of subband signals cause aliasing cancellation to fail [5], [10]. As shown in Section IV, the technique performs well in practice. Signal distortion related to subband decomposition can easily be made insignificant compared to main processing effects, at the price of only slight redundancy within subband data.

D. Group Delay Equalization

After an allpass transformation, an FB becomes an IIR system, whose group delay varies with frequency. If the system was a near-perfect reconstruction FB, its core (the

central part of Fig. 2 (a) with channel processing omitted) still introduces marginal distortion, provided that subsampling has been modified as described in Section II-C. Thus, the existence of the core can be neglected [37], and we can assume that there is only the direct connection of two symmetric allpas chains. In each patch from input to output, there is a cascade of N allpass filters $A(z)$, whose phase response, $A^N(z)$, determines that of the warped FB [31].

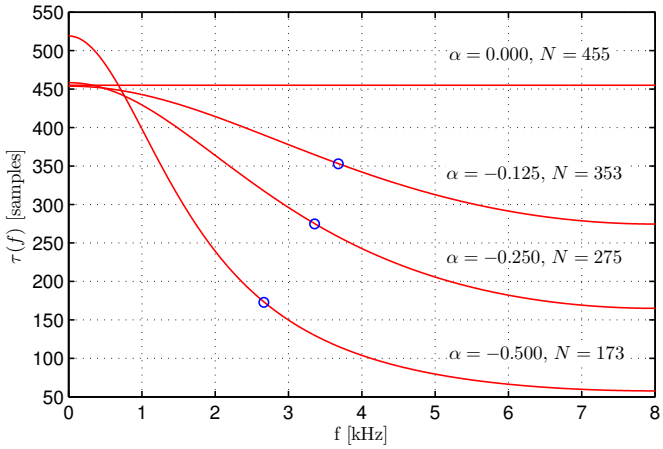

Fig. 3. Group delays of first-order allpass cascades.

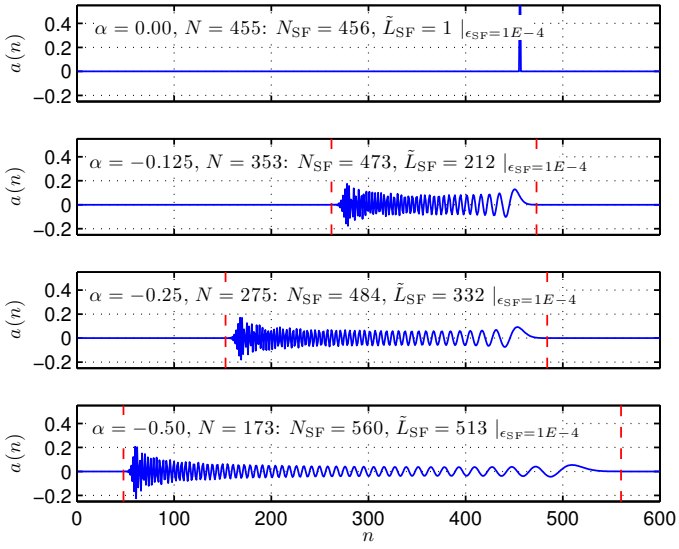

Fig. 4. Impulse responses of first-order allpass cascades.

Figure 3 shows the group delays of the cascades related to the FBs of Section IV. A circle indicates the group delay of an FB before the allpass transformation. Deviation from this value increases with intensifying warping, whereas changing the sign of α flips the curve vertically.

In Appendix B, several approaches to designing group delay equalizers are evaluated, which are mainly oriented to reversing the phase modification caused by a single allpass filter. If a cascade needs to be handled, such phase-correcting circuits can be connected in series, but using a single equivalent equalizer results in lower delay and makes error evaluation easier.

We propose to design such aggregate FIR equalizers by trimming and time-reversing the impulse response of the cascade [38]. Even though the response is infinite, it has a well determined significant fragment, as illustrated in Fig. 4. Thus, it can be assumed that only the samples with indexes from 0 to N_{SF} exist, which determine a finite-order polynomial almost equivalent to the original transfer function:

$$A^N(z) \approx \sum_{n=0}^{N_{\mathrm{SF}}} a(n) z^{-n}. \tag{11}$$

By reversing the order of the samples, a FIR filter is obtained that approximates the series connection of the inverse transfer function and delays:

$$C^N_{\mathrm{TRIM}}(z) = \sum_{n=0}^{N_{\mathrm{SF}}} a(N_{\mathrm{SF}} - n) z^{-n} \approx A^N(z^{-1}) z^{-N_{\mathrm{SF}}}. \tag{12}$$

Since $A^N(z^{-1}) = A^{-N}(z)$,

$$\begin{aligned} T_{\mathrm{TRIM}}(z) =& A^N(z) C^N_{\mathrm{TRIM}}(z) \\ \approx& A^N(z) A^{-N}(z) z^{-N_{\mathrm{SF}}} = z^{-N_{\mathrm{SF}}}. \end{aligned} \tag{13}$$

Obviously, the deviation of the product on the left-hand side from the pure delay depends on the residual after trimming.

The last significant sample can be determined as that after which all absolute values are less than some threshold. For brevity, careful optimization of the threshold subject to distortion constraints is left as a subject for a separate paper. For our purposes, it seems sufficient to prove that determining it roughly allows equalizer properties to be estimated accurately, and to show that $\epsilon_{\mathrm{SF}} = 10^{-4}$ suits our design examples. This is done empirically by evaluating four different allpass cascades such as magnitude distortions, group delay ripples, and equalizer order of $T_{\mathrm{TRIM}}(z)$ vary with ϵ_{SF}. Figure 5 shows this for the allpass cascades with $\alpha = -0.75$ and of lengths from 50 to 500. This is the strongest warping of interest to us, as its coefficient can be implemented with two binary shifts and one addition.

It can be observed that equalization accuracy, in terms of both the maximum magnitude ripple, $\max \Delta |T_{\mathrm{TRIM}}(e^{j\omega})$ and the peak-to-peak group delay, τ_{TRIMpp}, increases quickly with decreasing the threshold. On contrary, delay related to such equalization, determined as the average group delay $\tau_{\mathrm{TRIMavg}} \approx N_{\mathrm{SF}}$, is almost independent of threshold value. For slighter warpings, distortions and delay are lower for a given cascade length, but the shapes of curves are preserved. In particular, $\epsilon_{\mathrm{SF}} = 10^{-4}$ determines the knees in the distortion curves and ensures that group-delay ripples after equalization are not greater than several samples for low frequencies and quickly decrease with frequency to a fraction of sample. The corresponding magnitude distortion behaves similarly and does not exceed a few hundred dB. Thus, for a wide range of α and N, the proposed threshold value allows for accurately estimating the order and thus complexity of a trimming-based equalizer that introduces magnitude distortion comparable to that expected of the filter bank core and corrects group delay accurately. This is all what is necessary in our research, as will be shown in the rest of the paper.

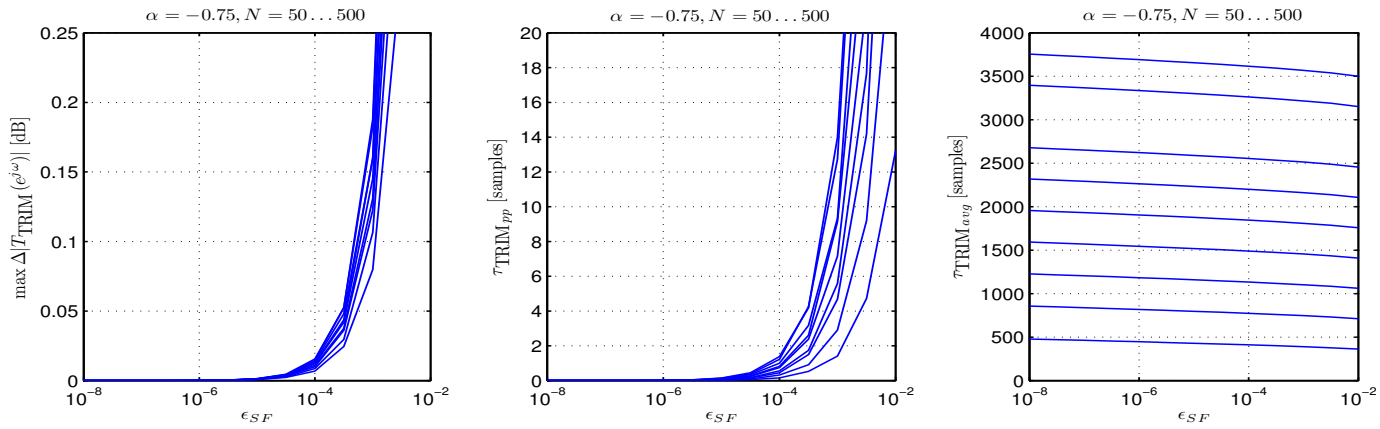

Fig. 5. Properties of equalizers designed using the trimming approach as functions of the threshold ϵ_{SF}.

Equalized signals are delayed by N_{SF} samples. As Fig. 6 shows, the delay grows linearly with adding stages to a cascade, at the rate determined by the warping coefficient. Thus, the problem of considerable delay arises for strong warpings and/or long cascades.

In such cases, equalization requires many computations, as significant parts are long. Some operations can by saved by omitting from the end of $C^N_{TRIM}(z)$ below-threshold samples that precede the significant ones, so that the filter is shortened to length $\tilde{L}_{\mathrm{SF}}$. As shown in Fig. 7, achievable savings depend on warping strength.

Dividing the number of significant samples by the cascade length results in the number of equalization coefficients that are associated with a single allpass filter. Fig. 8 clearly shows that dealing with whole cascade is much more efficient than stage-by-stage equalization.

If an application requires a warped FB to process signals without causing distortions, the delay is much more problematic than computational load, as it cannot be overcome by adding resources.

Finally, it should be noted that the trimming approach to design of equalization filters was independently considered by Löllmann, see e.g. [39], but in the context of DFT FBs with the troublesome synthesis structure of Fig. 2 (b) and without giving as general insights about the technique as our plots provide.

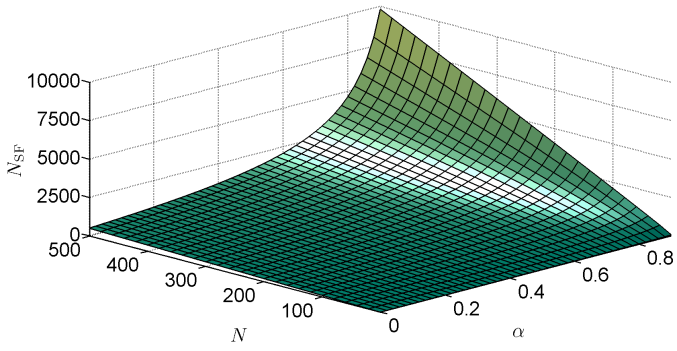

Fig. 6. Signal delay in allpass cascades equalized using the trimming approach ($\epsilon_{\mathrm{SF}} = 10^{-4}$).

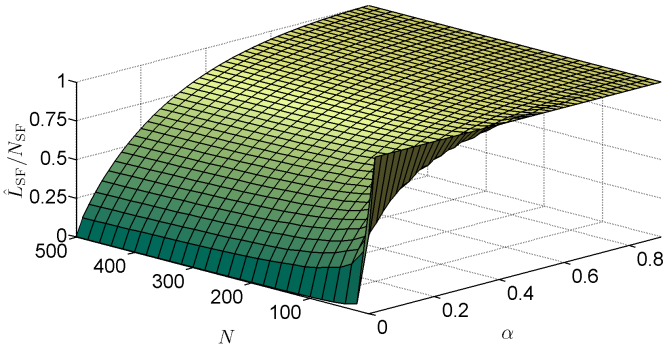

Fig. 7. Ratio between length of significant part and delay as a function of warping strength and cascade length ($\epsilon_{\mathrm{SF}} = 10^{-4}$).

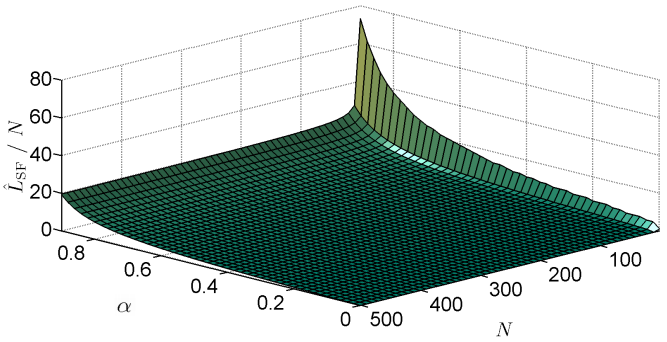

Fig. 8. Performance of the trimming approach ($\epsilon_{\mathrm{SF}} = 10^{-4}$). Please be aware that the viewpoint is different from Figs. 6 and 7.

E. Complexity Estimation

Warping significantly increases the multiplicative complexity of an FB. As channels are differently subsampled, compressors and expanders of sampling rate cannot be moved inside a system, in order to reduce the rate of both prototype filtering and modulation as it is typically seen in polyphase structures. Additional computational loads also appear, which are related to allpass filtering and to group delay equalization.

Therefore, for each sample of the fullband signal,

$$\mu_{\mathrm{A}} = \mu_{\mathrm{AP}} + \mu_{\mathrm{FIR}} + \mu_{\mathrm{MOD}} \tag{14}$$

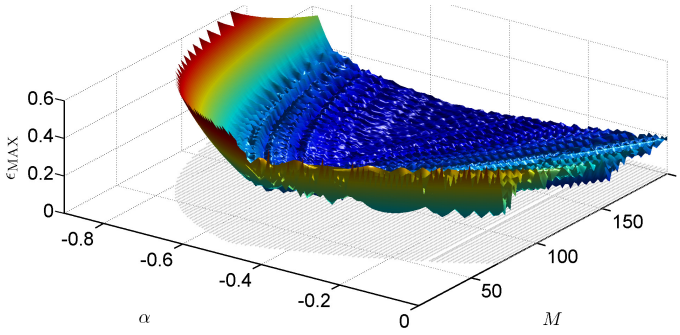

Fig. 9. Maximum error surface for $f_s = 16$ kHz.

and

$$\mu_S = \mu_A + \mu_{EQ} \tag{15}$$

multiplications are performed in analysis and synthesis parts of an FB, respectively.

The prototype filtering obviously requires $\mu_{FIR} = L$ multiplications. The prototype order also determines the complexity of the allpass chain, $\mu_{AP} = L - 1$, provided that first-order filters are used and each of them is implemented using one multiplication, as shown in [40].

The computational load of group delay equalization, μ_{EQ}, is more difficult to determine, because it depends also on warping strength and the structure selected for performing the task. Assuming that the approach of Section II-D is used, $\mu_{EQ} = \mu_{AP} + \tilde{L}_{SF}$.

Similarly, the cost of cosine modulation, μ_{MOD}, depends not only on the number of channels, but also on the algorithm used. For our purposes, the estimate $\mu_{MOD} = M \log_2 M$ is adequate, as explained in Section III-C.

It should be noted that the number of channels indirectly affects μ_{FIR} and μ_{AP}. Namely, subband merging requires the overlap among base subbands to be limited to only adjacent ones [11], so that for more channels, a more selective and thus longer prototype is necessary.

If the filter is designed using the Kaiser window function, its order necessary to have a stopband attenuation of δ for a transition bandwidth of Δf (expressed using normalized frequency) can be estimated as [24]

$$N \approx \frac{-20\log_{10}(\delta) - 7.95}{14.357\Delta f}. \tag{16}$$

For a given cut-off frequency, Δf must be selected to give the roll-off factor, ρ in [24], not greater than 1, in order to limit subband overlap as necessary.

In Sections III-C and IV, the above facts will be used to evaluate the complexities of particular FBs.

III. Enhancing Design Freedom by Subband Merging

A. Motivation

In the existing literature on FBs that mimic the human ear, see e.g. [2], [12], [14], the critical bands are approximated roughly. Mapping form the linear frequency axis to the Bark scale looks well only in terms of the ratio between bandwidth and center frequency of a subband. Obtainable frequency partitionings do not actually match the edges determined by Scharf [41]. The only exception we have noticed is [13].

Even though such an approach is considered satisfactory, the related design methods by itself cannot give better results. Thus our main idea was to develop a technique which allows a desirable frequency partitioning to be approximated accurately in terms of subband edges.

Additionally, the method should be more efficient than simply putting together independent modulated filters. Even though warped systems outperform such a solution, they are still much more computationally demanding than critically sampled FBs. Probably for this reason, multiplicative complexity of warped FBs was not thoroughly analyzed in papers by others, and there were no attempts to decrease it. Especially, to the best of our knowledge, nobody before us considered adjusting the warping coefficient value and number of channels so as to allow efficient implementation of the key blocks of a warped system. Infinite-precision multiplication-based implementation of the allpass chain was assumed, and warping coefficient was tuned carefully [12], [13], instead of trying to look for a way to employ extremely efficient multiplierless allpass filters.

We propose to address the problems of both accuracy and complexity by extending frequency warping with subband merging. The idea is to design a less or more warped FB with more channels than there are subbands in the desirable frequency partitioning, and then merge its subbands in order to approximate the reference ones. So far, such an approach has been applied only to uniform systems [9], [11], but without much success. In the case of warped ones, it seems much more beneficial, as shown in the following. Even though we limit our considerations to first-order warpings, the technique can be used in higher-order cases.

B. Approximating a Desirable Frequency Partitioning

In our approach, the fundamental question is how to assign channels of a warped FB to groups to be merged. We propose to simply group a sequence of subbands such that the lower frequency of the first of them and the upper frequency of the last are closest to the edges of a given reference subband. There is nothing to prevent a single subband from approximating a reference one by itself, but ambiguity arises when one subband is assigned to two groups. FBs for which the problem occurs are unable to well approximate the desirable partitioning, regardless of a way the ambiguity is solved, so they should be rejected from consideration.

The main design task is to determine such a warping coefficient value and number of channels, α and M, that subbands of the corresponding FB, after merging, accurately match a desirable frequency partitioning. This requires approximation accuracy to be measured, for which purpose we use maximum (MAX) and root-mean-square (RMS) error metrics based on the bandwidth-weighted deviation of the upper frequency from the reference one:

$$\epsilon_{MAX} = \max_{k=0\ldots\tilde{M}-2} \frac{|f_{Uk} - \hat{f}_{Uk}|}{|\hat{f}_{Uk} - \hat{f}_{Lk}|}, \tag{17}$$

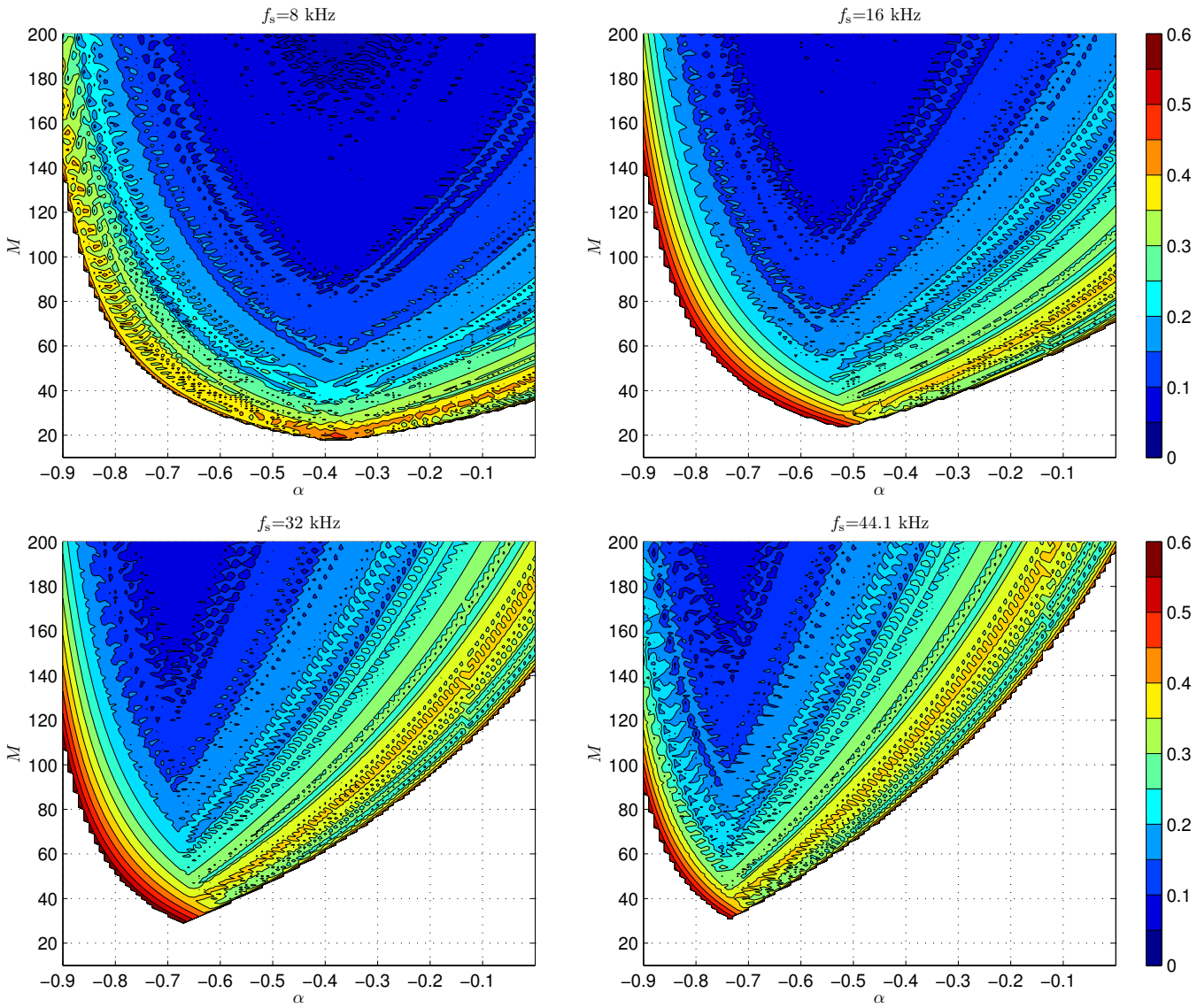

Fig. 10. Contour plots of ϵ_{MAX} for different sampling frequencies.

and

$$\epsilon_{\mathrm{RMS}} = \sqrt{\frac{1}{M-1}\sum_{k=0}^{\tilde{M}-2}\left(\frac{f_{\mathrm{U}k}-\hat{f}_{\mathrm{U}k}}{\hat{f}_{\mathrm{U}k}-\hat{f}_{\mathrm{L}k}}\right)^2}, \tag{18}$$

respectively, where $\tilde{M} \leq M$ denotes the number of channels after merging. The lower and upper edges of the kth subband obtained from merging are denoted by $f_{\mathrm{L}k}$ and $f_{\mathrm{U}k}$, respectively, whereas the hats indicate the edges of the corresponding reference band.

Such measures are much more restrictive than those of most of previous papers, as mentioned in Section III-A.

Having specified the measures, we can see how they depend on α and M. As an example let us consider approximating the critical band of hearing [41], the edges of the 22 first of which can be found in the second column of Table I. Scanning a wide range of parameter values: $-0.9 \leq \alpha \leq 0$ and $20 \leq M \leq 200$, results in error surfaces like that of Fig. 9. The experiment has been repeated for four sampling frequencies that are used in sound processing: 8, 16, 32, and 44.1 kHz, i.e. for different numbers of critical bands in the frequency range. The obtained surfaces of the maximum error are visualized in Fig. 10 using contour plots. They are wedge-like bounded because there are ambiguities in subband merging for the omitted pairs of parameter values.

Evidently, using both merging and warping needs less channels to achieve the same accuracy as only merging at $\alpha = 0$. For a fixed number of channels, intensifying warping improves accuracy, but only to some degree: too much warping worsens the approximation. If the number is small, employing warping is necessary to make approximation possible. For a fixed warping, accuracy can always be improved by adding more channels. The greater sampling frequency, the more channels are necessary to achieve desirable accuracy. From another point of view, the larger difference between the narrowest and widest reference bandwidths, which is caused by widening the frequency range, the more warping outperforms merging.

Only the general trends are characterized above. As the errors oscillate, there are many local minimums, separate or forming grooves. It is clear that the same accuracy can be achieved using different combinations of warping strength and number of channels, which correspond to points irregularly distributed over the parameter plane.

Fig. 11 shows, for 8 and 16 kHz sampling rates, the intersections of both RMS and MAX error surfaces with the planes determined by the warping coefficient values for which allpass filters can be implemented without multipliers, i.e. 0, $-1/2$, $-1/4$, and $-1/8$. The observations already made are confirmed, but the plots also prove that the errors are highly correlated, so our attention can be focused only on the second of them. Evidently, the critical bands can be approximated as accurately as necessary even with a constrained value of the warping coefficient. Because local minimums exist, the minimum necessary number of channels is much lower than that resulting from the general trend of error convergence.

Thus, the only reliable optimization technique is to carefully evaluate all potentially interesting points of the set, especially its front and subsets determined by particular values of α or M that are advantageous from an implementation point of view. By comparing FBs that correspond to such points, with respect of group delay, frequency selectivity and complexity, the system can be determined that best balances response quality with implementation requirements.

It seems impossible to analytically describe and solve such a design task. However, preparing data and plots can easily be automatized and takes no more than several minutes with a PC computer. Then, human intuition is necessary for selecting the best compromise.

C. Minimizing Complexity

In order to find the system of the minimum possible computational load for a given accuracy of frequency partitioning, an error plot must be analyzed together with that of multiplicative complexity. Fig. 12 shows the latter plot for both analysis FB and full subband analysis-synthesis system with group delay equalization. Computational loads were estimated as described in Section II-E, with omitting the term related to the allpass-chain if a multiplierless implementation is possible. This is why there are grooves in the generally smooth surfaces. There are marked the parameter combinations which give accuracies of 0.15 (dotted line) and 0.3 (solid line) at minimum complexity (number of channels, in fact) for the warping strengths of our primary interest.

The experiment confirms our intuition that for a given partitioning precision, the complexity of a system can be minimized by making a compromise between warping strength and the number of channels, i.e. by trading off the computational load related to the allpass transformation for that of the modulation. The former is predominant when frequency partitioning is based on warping (the number of channels to be merged is comparable with that of the desirable subbands), whereas the latter plays a key role in merging-based designs (many channels compensate for slight warping). Pure merging is not the best solution, even though it simplifies design and implementation, as there are no the side-effects of the allpass transformation.

If phase equalization can be neglected, it is advantageous to minimize the number of channels at the cost of the strongest warping. If group delay has to be equalized, a compromise between warping and merging is necessary to minimize the computational load. In the latter case, obtainable savings in multiplications are lower but still evident. Precise evaluations can be found in Section IV.

It is noteworthy that the parametric equalization circuits of Appendix B-B can be implemented without multiplications if the allpass coefficient is a power of two. On the other hand, efficient algorithms for cosine modulation of complexity $O(M \log_2 M)$, see e.g. [25], usually can be constructed only for certain numbers of channels (a power of two or a divider of prototype length). As our approach allows for constraining the number without deteriorating frequency partitioning much, it facilitates employing fast algorithms.

However, such modulation schemes seem not very advantageous for the systems under consideration. In the next section it is shown that avoiding aliasing by channel oversampling, as we propose, results in only slight redundancy within subband data. On average, only 2–3 subband values are generated for each sample of a fullband signal. The values are sums of small subsets of outputs of the modulation stage, and subset cardinalities are comparable or even lower than $\log_2 M$. As fast implementations with intricate data flows generally do not support skipping operations, most of outputs would be computed without a need. Thus, we have concluded that it is most practical to straightforwardly compute inner-products of rows of the matrices (8) and sample vectors. It can be easily implemented as a simple MAC-based computational scheme, in which it is easy to identify and skip operations related to an unnecessary output. Moreover half of $2M$ multiplications related to an inner-product can be saved by exploiting (anti)symmetries of matrix rows. For such an implementation, the estimate $\mu_{\mathrm{MOD}} = M \log_2 M$ seems even inflated.

IV. Design Examples

The results of the previous sections have been used to design four FBs that approximate the critical bands at 16 kHz sampling rate. Their magnitude responses before and after merging are represented in Fig. 13 by dashed and solid lines, respectively. The vertical lines represent the bounds of the critical bands. The figure also shows the magnitudes, $|T(f)|$, and group delays, $\tau(f)$, of the distortion transfer functions of the subband analysis-synthesis systems with group delay equalization[4]. Design parameters, prototype orders and error values accompany the plots, whereas the details of merging and subsampling are given in Table I.

The prototypes have been designed using the Kaiser window approach [24]. In order to limit the algorithmic delay, L was set to $4M$. For such a constraint, limiting overlap between subbands before merging as required in Section II-E results in stopband attenuation about 60 dB.

[4]Unequalized group delays differ marginally from those in Fig. 3.

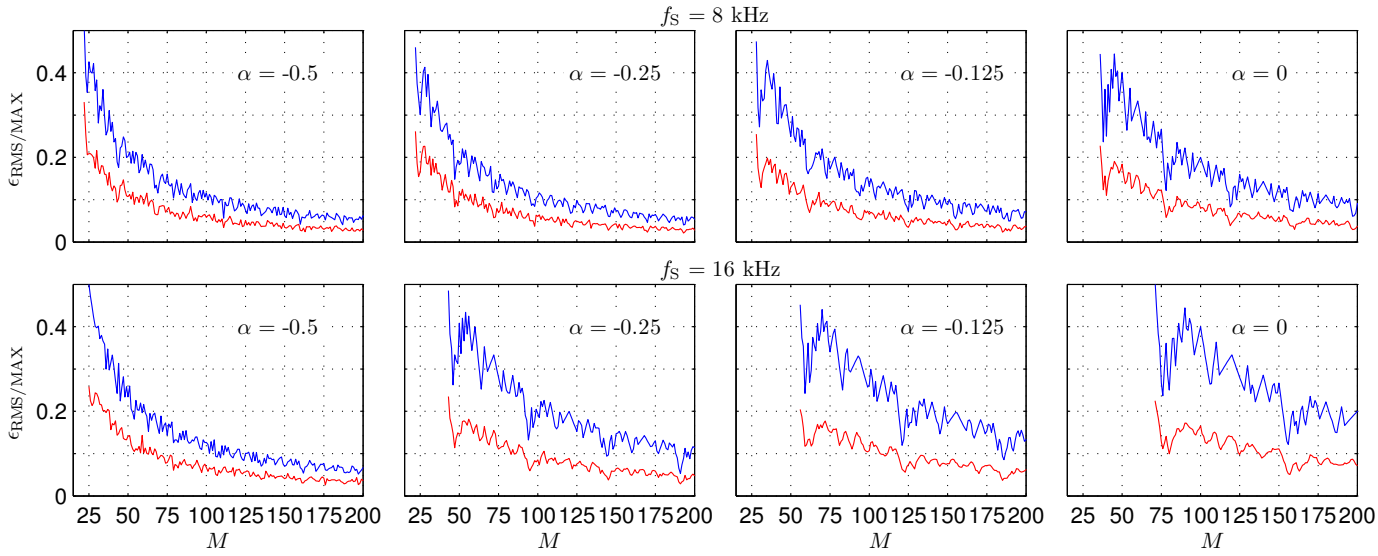

Fig. 11. Dependencies of RMS and maximum errors on the number of channels.

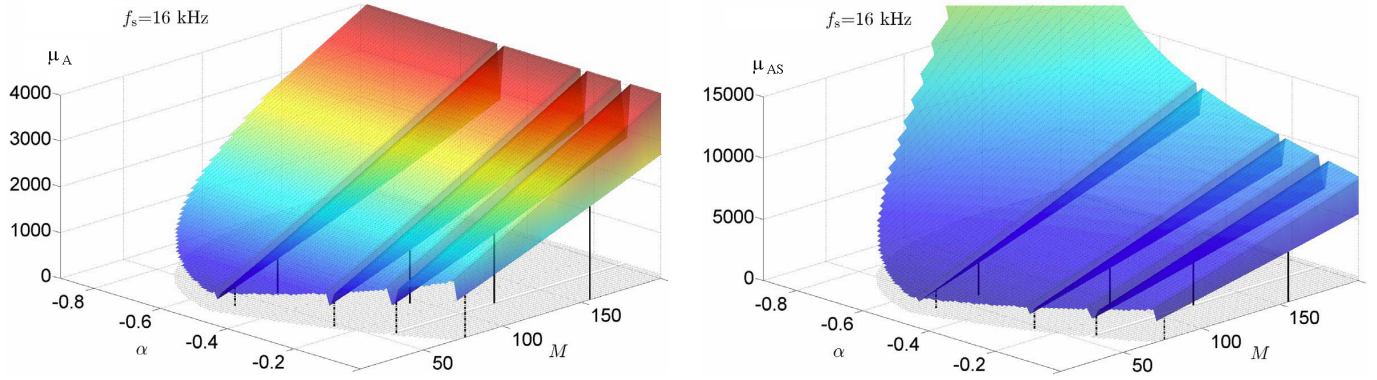

Fig. 12. Complexity surfaces of an analysis FB and of the corresponding equalized analysis-synthesis system ($\mu_{AS} = \mu_A + \mu_S$).

The FBs correspond to different tradeoffs between warping strength and number of channels. Their parameters were selected by identifying the first local minimums in the maximum error plots in Fig. 11. Except the most warped one, the FBs are characterized by similar approximation accuracies, group delays, and distortion levels. The main difference between systems is how transition bandwidth depends on the subband number. The less warping, the less it increases with frequency. In the system with $\alpha = -1/8$, both narrow and wide subbands are characterized by similar transition bandwidths, whereas in the FB without warping, there is no difference at all. The latter case can be considered unnatural in some applications.

Warping allows the overlap between subbands to be made proportional to their bandwidths, so that prototype length can be shortened without affecting stopband attenuation. This also affects the total oversampling ratio, which is abbreviated as TOR and computed as the sum of the reciprocals of channel subsampling ratios, determined as described in Section II-C. Widening transition bandwidth increases the risk of aliasing, so that some margins must be added to subband edges before applying (10), which then results in lower subsampling ratios.

Only slight redundancy within subband data is necessary to avoid aliasing by channel oversampling. For all systems, the total oversampling ratio about 2 is sufficient for keeping aliasing errors comparable with magnitude distortions.

For each FB, Table II shows multiplicative complexity terms related to its different parts in accordance with Section II-E. Allpass chains are assumed to be implemented without multiplications, but operation counts for multiplier-based versions are given in brackets for comparison. The savings depend on the ratio between the prototype length and the number of channels, and on whether group delay is equalized. They are 20–30 % in our case and should be similar for other practical designs.

The systems based on both warping and merging are more efficient than those using only merging. The difference depends on whether group delay is equalized or not, but even in the former case, it is considerable, i.e. 13–39%. Thus, from the point of view of multiplicative complexity, it is best to minimize the number of channels at the price of maximizing warping.

If, however, group delay is an issue, systems with no or with only slight warping are better choices. From the plots, it is evident that phase distortions caused by warping can be corrected accurately, but this comes at the price of delaying signal much. Reducing the delay requires not only finding the

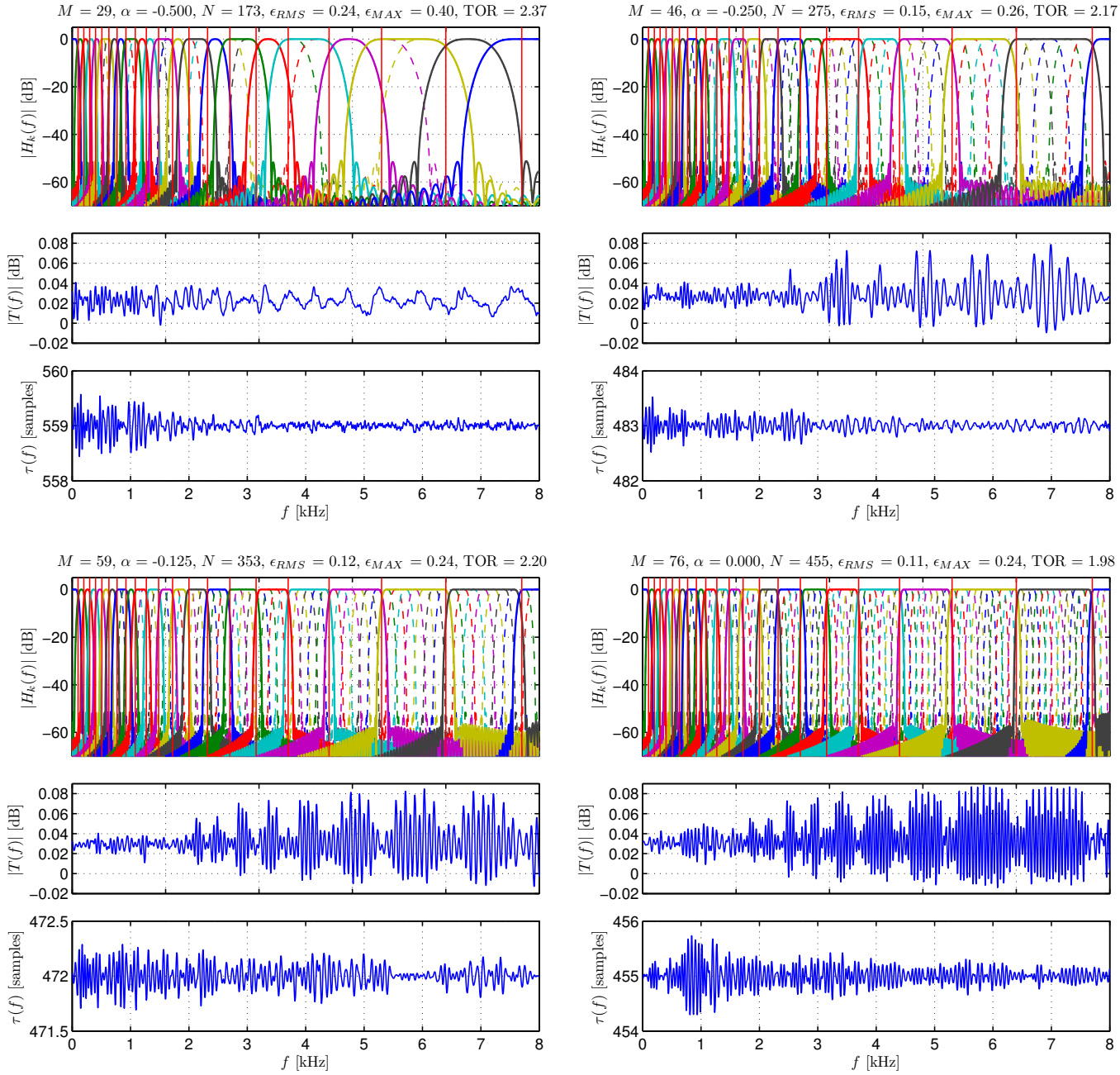

Fig. 13. Magnitude responses of FBs based on warping and merging.

tradeoff between warping and merging, but primarily needs sacrificing both filter selectivity and accuracy of frequency partitioning. The designed FBs are characterized by delays of 29–35 ms at stopband attenuations about 60 dB, which is quite a good result and seems impossible to be improved without relaxing the requirements on the latter quantity.

The issue is related not to the design technique but to the fundamental limitation of signal analysis, which states that improving time resolution requires sacrificing frequency resolution and vice versa. Compared to other FB classes, our approach offers great flexibility in trading-off one of these quantities for another.

V. Concluding Remarks

Combining frequency warping and subband merging results in a very flexible technique for designing nonuniform FBs. In addition to making good approximation of peculiar frequency partitionings possible, the mix allows developers to explore different tradeoffs among complexity, selectivity, delay and distortions, that is among those FB properties that are essential from a practical point of view. Neither warping nor merging has itself such a potential, whereas design methods that use efficient critically sampled FBs with aliasing cancellation cannot give frequency partitioning that does not match a uniform grid. Even though oversampling and merely approximating perfect reconstruction are disadvantages of the presented systems,

TABLE I
THE CRITICAL BANDS [41] AND PARAMETERS OF FBS APPROXIMATING THEM AT $f_s = 16$ KHZ

Critical band		System parameters | Merged subbands / Subsampling ratio S_k			
No.	Freq. range kHz	$\alpha = -0.5$ $M = 29$	$\alpha = -0.25$ $M = 46$	$\alpha = -0.125$ $M = 59$	$\alpha = 0$ $M = 76$
1	0.00 – 0.10	0 – 0 / 57	0 – 0 / 51	0 – 0 / 50	0 – 0 / 50
2	0.10 – 0.20	1 – 1 / 34	1 – 1 / 30	1 – 1 / 30	1 – 1 / 30
3	0.20 – 0.30	2 – 2 / 24	2 – 2 / 21	2 – 2 / 21	2 – 2 / 21
4	0.30 – 0.40	3 – 3 / 37	3 – 3 / 33	3 – 3 / 33	3 – 3 / 33
5	0.40 – 0.51	4 – 4 / 30	4 – 4 / 27	4 – 4 / 27	4 – 4 / 27
6	0.51 – 0.63	5 – 6 / 22	5 – 5 / 35	5 – 5 / 34	5 – 5 / 35
7	0.63 – 0.77	7 – 7 / 28	6 – 6 / 30	6 – 6 / 30	6 – 6 / 30
8	0.77 – 0.92	8 – 8 / 25	7 – 8 / 15	7 – 8 / 15	7 – 8 / 24
9	0.92 – 1.08	9 – 10 / 20	9 – 9 / 28	9 – 9 / 28	9 – 9 / 36
10	1.08 – 1.27	11 – 11 / 30	10 – 11 / 17	10 – 11 / 17	10 – 11 / 24
11	1.27 – 1.48	12 – 12 / 27	12 – 13 / 20	12 – 13 / 20	12 – 13 / 20
12	1.48 – 1.72	13 – 14 / 13	14 – 15 / 17	14 – 15 / 17	14 – 15 / 23
13	1.72 – 2.00	15 – 15 / 15	16 – 17 / 15	16 – 17 / 15	16 – 18 / 15
14	2.00 – 2.32	16 – 17 / 9	18 – 19 / 13	18 – 20 / 13	19 – 21 / 13
15	2.32 – 2.70	18 – 18 / 11	20 – 22 / 11	21 – 23 / 11	22 – 25 / 11
16	2.70 – 3.15	19 – 20 / 7	23 – 24 / 15	24 – 27 / 7	26 – 29 / 12
17	3.15 – 3.70	21 – 21 / 8	25 – 28 / 6	28 – 31 / 8	30 – 34 / 8
18	3.70 – 4.40	22 – 23 / 5	29 – 31 / 7	32 – 36 / 7	35 – 41 / 7
19	4.40 – 5.30	24 – 24 / 4	32 – 35 / 4	37 – 42 / 4	42 – 49 / 6
20	5.30 – 6.40	25 – 26 / 2	36 – 39 / 2	43 – 49 / 2	50 – 60 / 2
21	6.40 – 7.70	27 – 27 / 4	40 – 44 / 4	50 – 56 / 4	61 – 72 / 4
22	7.70 – 8.00[5]	28 – 28 / 6	45 – 45 / 18	57 – 58 / 18	73 – 75 / 21

TABLE II
COMPLEXITY OF DESIGNED FBS

Filter bank		Complexity terms				Total complexity	
α	M	μ_{AP}	μ_{FIR}	μ_{MOD}	μ_{EQ}	μ_A	μ_{AS}
-0.5	29	0 (173)	174	141	513	315 (488)	1143 (1489)
-0.25	46	0 (275)	276	254	332	530 (805)	1392 (1942)
-0.125	59	0 (353)	354	347	212	701 (1054)	1614 (2320)
0	76	0	456	475	0	931	1862

such FBs are useful in sound enhancement and multiple description coding.

APPENDIX A
MULTIPLE FREQUENCY MAPPING IN HIGHER-ORDER WARPINGS

Direct substitution of delays in a uniform FB for an Rth-order allpass filter gives a result different from what is expected of flexible warping. Namely, each of the resulting frequency responses is composed from R images of the original response, as demonstrated in Fig. 14 (a).

The problem arises because the phase response (7) is always monotone decreasing, which is a consequence of the equivalence (6) between a higher-order allpass filter and a cascade of first-order ones. Thus Rth-order allpass transformation corresponds to R mappings of the frequency interval $-\pi \le \omega \le \pi$ onto its R times shorter fragments. From another point of view, such multiple mapping can be perceived as a single mapping of the extended interval $-R\pi \le \omega \le R\pi$ onto $-\pi \le \omega \le \pi$.

The issue can be avoided by reducing to ω the term $R\omega$ in (7), so that the resulting phase shift

$$\hat{\phi}(\omega) = \phi(\omega) + (R-1)\omega \tag{19}$$

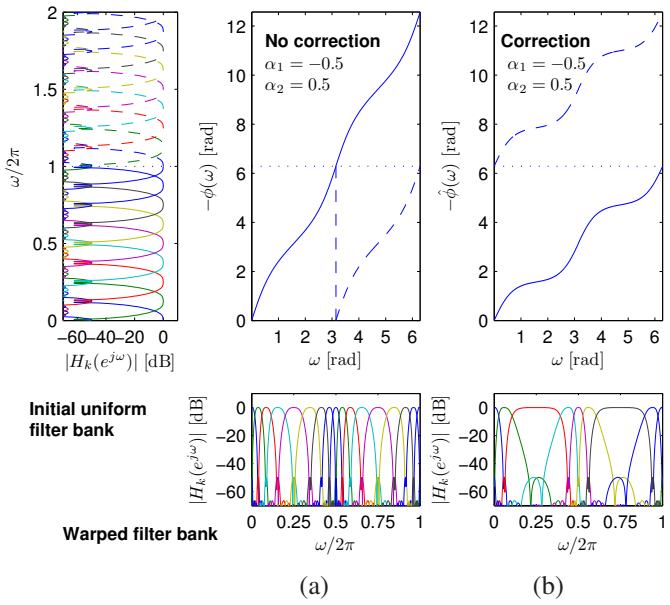

Fig. 14. Second-order warping ($\alpha_1 = -0.5$ and $\alpha_2 = 0.5$) (a) without and (b) with the correction of multiple frequency mapping.

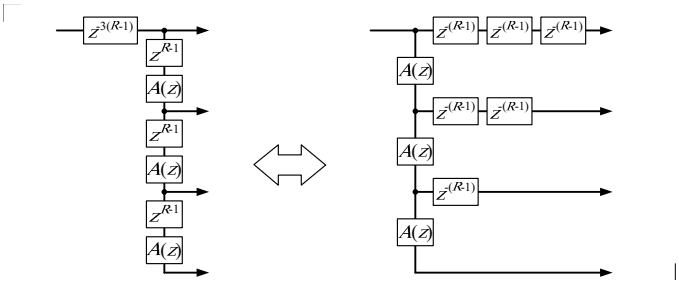

Fig. 15. Causal correction of multiple frequency mapping.

is limited to $\langle -\pi, \pi \rangle$ [29], as illustrated in Fig. 14. Such a modification requires connecting a filter in series with the reciprocals of unit delays, which results in

$$\hat{A}(z) = z^{+(R-1)} A(z). \tag{20}$$

Even though such a solution seems impractical due to the use of the noncausal transfer function, it can easily be applied to allpass chains [29], as explained in Fig. 15. The idea is to precede the chain of the non-causally corrected allpass filters with appropriately selected additional delays. Moving the delays from the input to the outputs, we can annihilate noncausal blocks and obtain the equivalent causal structure.

Such correction usually significantly increases memory requirements, and, even more importantly, increases by $N(R-1)$ samples the total delay introduced by a warped FB with the chain of N allpass filters.

APPENDIX B
ALTERNATIVE METHODS FOR GROUP DELAY EQUALIZATION

A. Noncausal Filtering

Obviously, phase distortion caused by an allpass filter $A(z)$ can be counterbalanced by the inverse transfer function

$A^{-1}(z)$. The latter is unstable but can be decomposed into two stable filters, one of which is noncausal [42]. This makes an implementation possible, as time-reversing can be simulated by buffering signals, even infinite-length ones, like sound [42]. However, such systems are difficult to design and implement, especially in cases of memory or delay limits.

Time-reversing and trimming sequences have also been used to develop warped wavelet FBs in [15]. Even though the related theory is nice and innovative, its practical use is limited due to high computational load and memory requirements. Moreover, the related time-frequency representations of signals are redundant, similarly to our FBs.

B. Polynomial and Binomial Equalization Filters

In [37], it has been shown that the phase modifications caused by a first-order allpass filter can be compensated for by passing its output signal through the FIR transfer function

$$C_{\mathrm{POLY}}(z) = (1+\alpha z^{-1}) \sum_{n=0}^{D-1} (-\alpha)^n z^{-(D-1)+n}. \tag{21}$$

The polynomial is constructed so that the product

$$A(z)C_{\mathrm{POLY}}(z) = T_{\mathrm{POLY}}(z) = z^{-D} - (-\alpha)^D \tag{22}$$

converges to a pure delay with increasing D, whereas the term $(-\alpha)^D$ decreases to zero, provided that $|\alpha| < 1$.

Obviously, the greater magnitude of the allpass coefficient, the higher order of the polynomial is required to keep distortions within acceptable limits. As

$$\left|T_{\mathrm{POLY}}(e^{j\omega})\right| = \sqrt{1 - 2(-\alpha)^D \cos D\omega + \alpha^{2D}}, \tag{23}$$

$$\arg\left\{T_{\mathrm{POLY}}(e^{j\omega})\right\} = \arctan \tfrac{-\sin D\omega}{\cos D\omega - (-\alpha)^D}, \tag{24}$$

both magnitude distortion caused by such equalization and phase error remaining after it exhibit equiripple behavior, with extrema at integer multiples of $\omega = \frac{\pi}{D}$. The group delay of an equalized allpass filter

$$\tau_{\mathrm{POLY}}(\omega) = D \tfrac{(-\alpha)^D \cos(D\omega) - 1}{2(-\alpha)^D \cos(D\omega) - \alpha^{2D} - 1}, \tag{25}$$

has the peak-to-peak ripple

$$\left|\tau_{\mathrm{POLY}}(0) - \tau_{\mathrm{POLY}}(\tfrac{\pi}{D})\right| = \left|-2D \tfrac{(-\alpha)^D}{\alpha^{2D}-1}\right|. \tag{26}$$

The latter expression allows us to determine D which gives equalization accuracy not worse than that achieved using the trimming approach. Scan over α and N results in the plot of Fig. 16, which can be compared to Fig. 8. Clearly, it is much better to equalize many allpass filters at once, in respect of both complexity and delay, which are about four times lower for the trimming technique.

In [31], we have shown that computations and memory can be saved by restricting D to be a power of 2, and by factorizing (21), which requires $D-1$ multiplications, in the following way:

$$C_{\mathrm{BI}}(z) = (1+\alpha z^{-1})(z^{-1} - \alpha) \prod_{n=1}^{\log_2 D - 1} (z^{-2^n} + \alpha^{2^n}). \tag{27}$$

Although still

$$A(z)C_{\mathrm{BI}}(z) = T_{\mathrm{BI}}(z) = z^{-D} - \alpha^D, \tag{28}$$

each of $\log_2 D$ binomials that replace the polynomial requires only one multiplication. Moreover, the number of necessary unit delays decreases in the same way.

The results in Fig. 16 can easily be transformed into those in Fig. 17, which shows how many binomials are necessary to equalize group delay as accurately as using an trimming-based filter. Even though the binomial-based equalization is most efficient, the problems with magnitude distortion and enormous delay still exist.

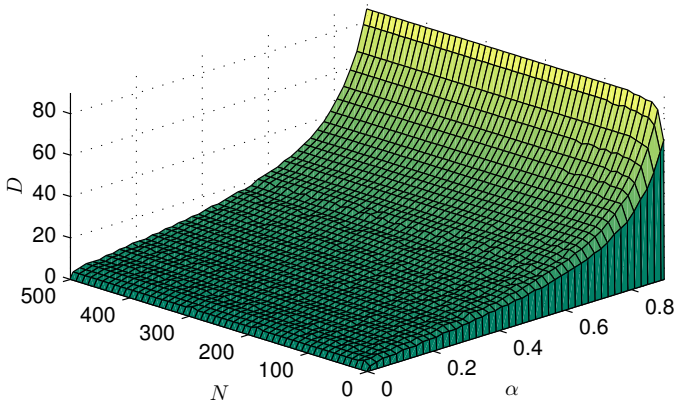

Fig. 16. Performance of the polynomial-based equalization for cascades of allpass filters.

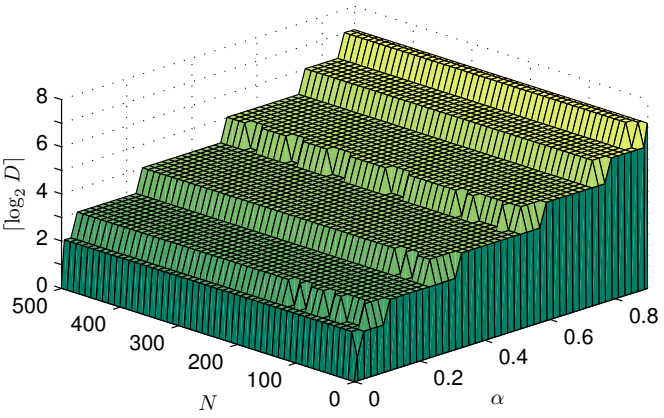

Fig. 17. Performance of the binomial-based equalization for cascades of allpass filters.

In [32], we have noticed a similar technique that uses the cascade of allpass filters

$$C_{\mathrm{AP}}(z) = \tfrac{z^{-1}-\alpha}{1-\alpha z^{-1}} \prod_{n=1}^{\log_2 D - 1} \tfrac{z^{-2^n} + \alpha^{2^n}}{1 - \alpha^{2^n} z^{-2^n}}. \tag{29}$$

Such equalization results in the following allpass function

$$A(z)C_{\mathrm{AP}}(z) = T_{\mathrm{AP}}(z) = \tfrac{z^{-D} - \alpha^D}{1 - \alpha^D z^{-D}}, \tag{30}$$

and thus causes no magnitude distortions. However, via simple trigonometric manipulations, one can prove that

$$\arg\left\{T_{\mathrm{AP}}(e^{j\omega})\right\} = \arg\left\{T_{\mathrm{BI}}(e^{j\omega})\right\} - \arctan \tfrac{\alpha^D \sin D\omega}{1 - \alpha^D \cos D\omega}, \tag{31}$$

and

$$\arg\left\{T_{\mathrm{AP}}(e^{j\omega})\right\} + D\omega = 2(\arg\left\{T_{\mathrm{BI}}(e^{j\omega})\right\} + D\omega). \quad (32)$$

So, phase ripples remaining after allpass-based equalization have the same shape as for the binomial approach, but are twice as large in amplitude. Another drawback is related to doubling the number of delays.

An advantage of analytically-described equalization is that it can be used in tunable warped FBs, in which frequency partitioning is changed on-the-fly by modifying the warping coefficient [16], [43].

C. General-Purpose Equalization Filters

Over the years, a fair number of general-purpose methods for group delay equalization have been developed (see e.g. [44] and references therein), but their applicability to warped systems is limited. Firstly, the high-dimensionality of the optimization problem of coefficient synthesis causes such methods to fail (in terms of convergence) if a longer cascade of allpass filters has to be equalized, or to give, after heavy computations, results not better than the algorithms described above. Secondly, obtainable IIR solutions often suffer from instability. Finally, some methods are aimed at equalization limited to only a range of frequencies.

APPENDIX C
ALTERNATIVE SYNTHESIS STRUCTURES FOR WARPED FILTER BANKS WITH GROUP DELAY EQUALIZATION

In [37], the structure of Fig. 2 (b) has been proposed in the context of warped DFT FBs. The wideband signal passed thorough an allpass chain (possibly followed by an FB core that has, or well approximates, the perfect reconstruction property, an thus is transparent for the rest of a system) is restored using a delay chain, whose inputs are preceded by cascades of an equalization filter $C(z)$. Both cascade lengths and delay amounts are selected so that the transfer function

$$T_n(z) = A^n(z)C^n(z)z^{-D(N-n)} \approx z^{-DN}, \quad (33)$$

where $n = 0, \ldots, N$, characterizes the nth path from the input to the output, assuming that (9) is satisfied.

The solution is not very useful in practice, especially because of its enormous computational complexity, as $C(z)$ is computed $\frac{N}{2}(N+1)$ times for each output sample. Another problem is that paths differ in the number of allpass filters, and thus in phase distortion to be compensated for. Thus each patch needs an individually designed equalizer of different length, which is especially troublesome if one wants to replace a cascade of independent filters, $C^n(z)$, with a more efficient single equalizer, as done in [39]. Additionally, it is difficult to evaluate phase distortion at the output.

In order to reduce the complexity of the scheme, we have introduced its transposition of Fig. 2 (c) [31]. The idea is to interchange compensation filters with delays, so that the former form a chain, and the latter are cascaded at chain inputs. Although the system is equivalent to the previous one, as (33) still describes the relationship between the input and output, $C(z)$ is computed only N times. The gain is incomparably higher than that obtained in [39] by aggregating equalization filters. Unfortunately, the memory requirements are huge, because of the presence of $\frac{N}{2}(N+1)D$ unit delays in the branches. Moreover, there is still a need to design equalization filters separately, one-by-one, because in spite of their theoretical equivalence, they form a hierarchy. In addition to equalizing an allpass cascade, a filter complements equalization of a longer cascade, which makes error evaluation difficult.

After considering the aforementioned schemes, we concluded in [31] that computational complexity, memory requirements, and design ease are reconciled in the structure of Fig. 2 (a), which separates fullband signal reconstruction from phase correction. The former is based on the allpass chain symmetric to that in the analysis FB, whose output is postfiltered in order to equalize group delay.

Although the presence of the chain increases the complexity, evident advantages are obtained in exchange. There are no cascades of delays in the structure, and each path from the input to the output is characterized by the same transfer function

$$T_n(z) = A^N(z)C^N(z) \approx z^{-DN}, \quad (34)$$

for $n = 0, \ldots, N$. Thus, instead of $L = N + 1$ cascades of from 0 to N allpass filters, only the last of them is of interest, which makes both evaluation of the output distortion and optimal design of an equalizer much easier then in the approaches mentioned above. Additionally, the scheme we recommend is easier to implement owing to its increased coherence and modularity.

A. Critically Sampled Subband Decomposition Systems with Allpass Filter Chains

It should be noted that several systems have been developed over the years, in which perfect reconstruction can be achieved, at least theoretically, even though they are based on allpass filter chains, and their subbands are critically sampled [21], [45]. However, these solutions have clear limitations, and there are conceptual differences between them and warped FBs in our sense. Namely, the systems can be considered degenerated because their FIR polyphase matrices are constrained to be paraunitary or orthogonal, respectively. Thus, in the first case, a system is warped only partially, whereas in the second case, where the DFT or DCT matrices can be used, the analysis filters obtained have low selectivity. Moreover, the synthesis structures, which are very complicated and similar in both cases, work differently from FBs, because modifications of a single channel are spread over all frequencies, so that the paradigm of subband processing breaks down.

Therefore, we have good reasons to content ourselves with only mentioning such systems.

APPENDIX D
AVOIDING ALIASING BY OVERSAMPLING

A. Problem Nature

Our discussion is limited to the simplest way of changing the sampling rate, that is to using decimators with integer ratios

at outputs of analysis FBs and expanders at inputs of synthesis FBs. This is because other approaches, like modulating a subband in order to shift it inside a feasible position [11], seriously affect the coherence of a system without significantly decreasing its oversampling ratio.

In warped cosine modulated FBs, subband widths and positions are virtually unrestricted. Especially, subbands are not integer-positioned [46], i.e. their lower frequencies are not integer multiples of the bandwidths. The opposite obviously occurs in uniform FBs as well as in nonuniform ones built upon them. It implies feasibility of frequency partitionings [11], and thus reconciles maximal decimation with perfect reconstruction via aliasing cancellation.

In cosine-modulated FBs, channel signals are usually real-valued and thus have spectra symmetric about the Nyquist frequency. The asymmetry occurs when the allpass filter and/or prototype have complex coefficients, and for complex-valued input signals. As such situations are uncommon, we give up investigating them.

Because two-part spectral structure of subbands is inherited by aliasing terms, subsampling can cause both parts of the same term to overlap. In FBs with warping and/or merging, the related error cannot usually be canceled using another subband, because the latter is differently subsampled, so that spectral replicas are not positioned as necessary for mutual compensation of distortions. Moreover, warping affects transition bandwidths of a channel filter unequally, so that a spectral replica is not necessarily equivalent to its reflected version.

Formal consideration of these issues leads to huge equations and hard optimization problems, and thus its practical usefulness seems little. Thus, unlike others [6], [7], [10], [33], [39], we do not analyze the distortion transfer function of a nonuniform FB [47] and give up the idea of approaching perfect reconstruction by employing aliasing cancellation. Instead, we have deduced that similar or even better results can be obtained by analyzing channels independently, one after another, using bandpass sampling theory.

We are interested in minimizing aliasing errors by carefully selecting subsampling ratios for channels, so that selectivity of the prototype completes the work. The idea is to keep channels somewhat oversampled, which prevents severe overlapping between subband spectrum and its replicas which arise from subsampling. As a result, channel signals contain no significant aliasing terms which must be canceled during subband synthesis. Only minor aliasing terms exist, whose level is determined by the stopband attenuation of the prototype filter, like in uniform near-perfect reconstruction FBs [11], [25], which use partial aliasing cancellation.

B. Constraints for Subsampling Ratios

Permissible subsampling ratios can be derived by identifying kth channel signal with a bandpass signal. This allows applying the fundamental principle of uniform sampling of such signals:

$$2\frac{f_{\mathrm{U}k}}{n_k} \le f_{\mathrm{s}k} \le 2\frac{f_{\mathrm{L}k}}{n_k-1}, \qquad 1 \le n_k \le \left\lfloor \frac{f_{\mathrm{U}k}}{f_{\mathrm{U}k}-f_{\mathrm{L}k}} \right\rfloor, \quad (35)$$

which can be found e.g. in [46]. The symbols have already been defined in Section II-C.

The inequalities determine acceptable sampling frequencies, but can easily be converted into the expression (10) for the corresponding subsampling ratios.

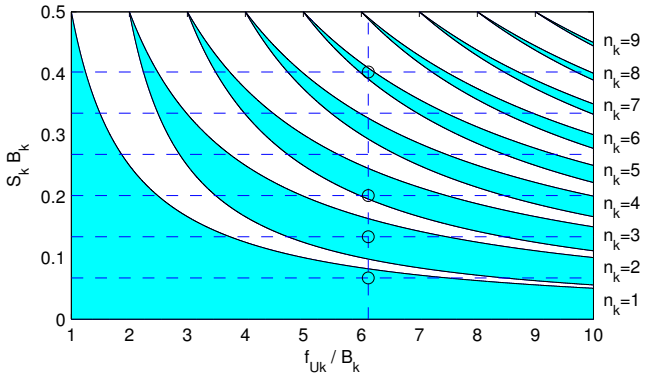

Fig. 18. Relations between subband positioning and permissible subsampling ratios in a cosine-modulated FB.

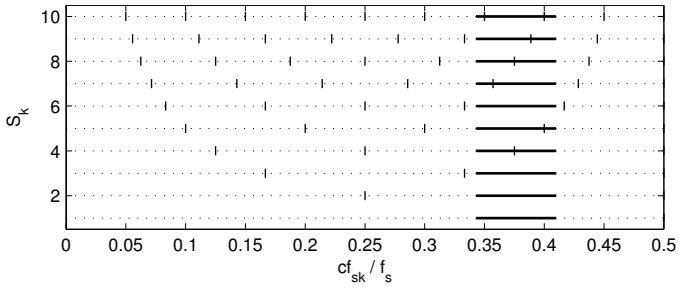

Fig. 19. Locations of a hypothetical subband with respect of multiples of lowered sampling frequencies.

Without the roundings, the inequalities (10) can be illustrated graphically with hyperboles, as shown in Fig. 18, where $B_k = f_{\mathrm{U}k} - f_{\mathrm{L}k}$. The shaded regions between the curves represent the cases in which subband positioning and subsampling ratio are related in such a way that there is no severe aliasing. The remaining area corresponds to unacceptable relations, and thus to distortion existence.

The diagram is equivalent to that of [46, Fig. 4] but allows direct determination of permissible subsampling ratios S_k, instead of sampling frequencies f_{sk}. Namely, for a given subband, the relation between its positioning and bandwidth can be represented with a vertical line. In turn, a horizontal line can be drawn to illustrate the bandwidth widening caused by subsampling. If the lines intersect inside a shaded area, sampling frequency is reduced without severe aliasing. Otherwise, subsampling causes unrecoverable errors.

In Fig. 18, the lines for a hypothetical kth subband with $f_{\mathrm{U}k} = 0.41$ and $B_k = 0.067$ are drawn. The cases of permissible subsampling are indicated with circles. Fig. 19 shows the positioning of the same subband (horizontal bold lines) against a background of multiples of the lowered sampling frequency (vertical strokes) for different subsampling ratios (a single level of the plot). It is obvious that (10) prevents such a multiply from occurring inside the subband.

Obviously, the rules we have developed can be applied to FBs in which subband merging alters frequency partitioning.

In such cases, a subsampling ratio for a channel after merging must be calculated based on the total bandwidth of the merged subbands as well as on the lower and upper frequencies of the first and last of them, respectively. It also should be taken into account that transition bands exist in real-world magnitude responses and their widths are affected by warping.

It should be noted that subsampling in nonuniform FBs was already independently studied in [48], where, however, neither such explicit expressions as (10) nor an accessible graphical approach have been given.

C. A Remark on Prototype Filter Design

The following two facts prove that there is no point in developing dedicated algorithms which evaluate warped responses during prototype synthesis.

Firstly, an allpass transformation of an FB in which channel signals are not subsampled causes some deformation of the magnitude response of the distortion transfer function only along the frequency axis. More precisely, amplitude distortions that already exist, which depend on the used prototype, are only shifted and stretched or compressed in frequency. Their presence and level themselves are unconnected with the warping. In particular, if a nonsubsampled perfect reconstruction FB is warped, only phase distortions characterize the analysis-synthesis system obtained.

Secondly, in our approach, prototype filters are only expected to have high stopband attenuation and possibly a narrow transition band (if high total subsampling ratio is desirable), which is usually required of prototypes for uniform FBs that approximate perfect reconstruction [24], [25]. Thus, the same design methods can be successfully used for warped FBs.

REFERENCES

[1] *Recommendation ITU-R BS.1387-1 — Method for Objective Measurements of Perceived Audio Quality*. International Telecommunications Union, 1999, ITU-R.

[2] A. Karmakar, A. Kumar, and R. K. Patney, "Design of optimal wavelet packet trees based on auditory perception criterion," *IEEE Signal Process. Lett.*, vol. 14, no. 4, pp. 240–243, Apr. 2007.

[3] A. Pandharipande and S. Dasgupta, "On biorthogonal nonuniform filter banks and tree structures," *IEEE Trans. Circuits Syst. I*, vol. 49, no. 10, pp. 1457–1467, Oct. 2002.

[4] W. Zhong, G. Shi, X. Xie, and X. Chen, "Design of linear-phase nonuniform filter banks with partial cosine modulation," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3390–3395, Jun. 2010.

[5] Z. Cvetković and J. D. Johnston, "Nonuniform oversampled filter banks for audio signal processing," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 5, pp. 393–399, Sep. 2003.

[6] E. Galijašević and J. Kliewer, "Design of allpass-based non-uniform oversampled DFT filter banks," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)*, vol. 2, Orlando, FL, 13–17 May 2002, pp. 1181–1184.

[7] H. W. Löllmann and P. Vary, "Least-squares design of DFT filter-banks based on allpass transformation of higher order," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2393–2398, Apr. 2010.

[8] X. Xie, S. Chan, and T. Yuk, "Design of linear-phase recombination nonuniform filter banks," *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2809–2814, Jul. 2006.

[9] J. Ogale and S. Ashok, "Cosine modulated non-uniform filter banks," *Journal of Signal and Information Processing*, vol. 2, pp. 178–183, 2011.

[10] Y. Deng, V. Mathews, and B. Farhang-Boroujeny, "Low-delay nonuniform pseudo-QMF banks with application to speech enhancement," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 2110–2121, May 2007.

[11] J. Li, T. Q. Nguyen, and S. Tantaratana, "A simple design method for near-perfect-reconstruction nonuniform filter banks," *IEEE Trans. Signal Process.*, vol. 45, no. 8, pp. 2105–2109, Aug. 1997.

[12] J. O. Smith III and J. S. Abel, "Bark and ERB bilinear transforms," *IEEE Trans. Speech Audio Process.*, vol. 7, no. 6, pp. 697–708, Nov. 1999.

[13] P. Ghosh and S. Narayanan, "Bark frequency transform using an arbitrary order allpass filter," *IEEE Signal Process. Lett.*, vol. 17, no. 6, pp. 543–546, Jun. 2010.

[14] T. Gülzow, A. Engelsberg, and U. Heute, "Comparison of a discrete wavelet transformation and nonuniform polyphase filterbank applied to spectral-subtraction speech enhancement," *Signal Process.*, vol. 64, no. 1, pp. 5–19, Jan. 1998.

[15] G. Evangelista and S. Cavaliere, "Discrete frequency warped wavelets: Theory and applications," *IEEE Trans. Signal Process.*, vol. 46, no. 4, pp. 874–885, Apr. 1998.

[16] A. Petrovsky, M. Parfieniuk, and K. Bielawski, "Psychoacoustically motivated nonuniform cosine modulated polyphase filter bank," in *Proc. 2nd Int. Workshop on Spectral Methods and Multirate Signal Process. (SMMSP)*, Toulouse, France, 7–8 Sep. 2002, pp. 95–101.

[17] X. Zhang, L. Huang, and G. Evangelista, "Warped filter banks used in noisy speech recognition," in *Proc. 4th Int. Conf. Innovative Computing, Information and Control (ICICIC)*, 7–9 Dec. 2009.

[18] M. Livshitz, M. Parfieniuk, and A. Petrovsky, "Wideband CELP coder with multiband excitation and multilevel vector quantization based on reconfigurable codebook," *Digital Signal Process. (KBWP, Moscow, Russia)*, no. 2, pp. 20–35, 2005, in Russian.

[19] A. Borowicz, M. Parfieniuk, and A. Petrovsky, "An application of the warped discrete Fourier transform in the perceptual speech enhancement," *Speech Comm.*, vol. 48, pp. 1024–1036, 2006.

[20] H. W. Löllmann and P. Vary, "Uniform and warped low delay filter-banks for speech enhancement," *Speech Communication*, vol. 49, no. 207, pp. 574–587, 2007.

[21] B. Shankar and A. Makur, "Allpass delay chain-based IIR PR filterbank and its application to multiple description subband coding," *IEEE Trans. Signal Process.*, vol. 50, no. 4, pp. 814–823, Apr. 2002.

[22] S. Caporale, L. De Marchi, and N. Speciale, "Frequency warping biorthogonal frames," *IEEE Trans. Signal Process.*, vol. 59, no. 6, pp. 2575–2584, Jun. 2011.

[23] W. H. Chin and B. Farhang-Boroujeny, "Subband adaptive filtering with real-valued subband signals for acoustic echo cancellation," *IEE Proc.–Vis. Image Signal Process.*, vol. 148, no. 4, pp. 283–288, 2001.

[24] A. Datar, A. Jain, and P. Sharma, "Design of Kaiser window based optimized prototype filter for cosine modulated filter banks," *Signal Processing*, vol. 90, pp. 1742–1749, 2010.

[25] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[26] P. Vary, "Digital filter banks with unequal resolution," in *Proc. EUSIPCO Short Communication Digest*, Lausanne, Switzerland, Sep. 1980, pp. 41–42.

[27] G. Doblinger, "An efficient algorithm for uniform and nonuniform digital filter banks," in *Proc. of Int. Symp. on Circuits and Systems (ISCAS)*, vol. 1, Singapore, Jun. 1991, pp. 646–649.

[28] A. G. Constantinides, "Spectral transformations for digital filters," *IEE Proc.*, vol. 117, no. 8, pp. 1585–1590, 1970.

[29] M. Kappellan, B. Strauss, and P. Vary, "Flexible nonuniform filterbanks using allpass transformation of multiple order," in *Proc. 8th European Signal Process. Conf. (EUSIPCO)*, vol. 3, Trieste, Italy, 7–13 Sep. 1996, pp. 1745–1748.

[30] T. von Schroeter, "Frequency warping with arbitrary allpass maps," *IEEE Signal Process. Lett.*, vol. 6, no. 5, pp. 116–118, May 1999.

[31] M. Parfieniuk and A. Petrovsky, "Reduced complexity synthesis part of non-uniform near-perfect-reconstruction DFT filter bank based on allpass transformation," in *Proc. 16th European Conf. on Circuits Theory and Design (ECCTD)*, vol. III, Cracow, Poland, 1–4 Sep. 2003, pp. III-5–III-8.

[32] H. W. Löllmann and P. Vary, "Parametric phase equalizers for warped filter-banks," in *Proc. 14th European Signal Processing Conf. (EUSIPCO)*, Florence, Italy, 4–8 Sep. 2006.

[33] J. M. de Haan, I. Claesson, and H. Gustafsson, "Least squares design of nonuniform filter banks with evaluation in speech enhancement," in *Proc. Int. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, vol. VI, Hong Kong, China, Apr. 2003, pp. 109–112.

[34] B. Vo and S. Nordholm, "Non-uniform DFT filter bank design with semi-definite programming," in *Proc. Int. Symp. Signal Process. Information Technology (ISSPIT)*, Darmstadt, Germany, 2003, pp. 42–45.

[35] G. Kumar, P. Reddy, and B. Charles, "Design of non-uniform filter banks: Quadratic optimization with linear constraints," *ARPN J. Engineering Applied Sciences*, vol. 2, no. 3, pp. 24–27, Jun. 2007.

[36] M. Parfieniuk and A. Petrovsky, "Simple rule of selection of subsampling ratios for warped filter banks," in *Proc. VIII Int. Conf. "Modern Communication Systems"*, Naroch, Belarus, 29 Sep.–3 Oct. 2003, pp. 130–134, haapec. Issue of Trans. Belarus. Eng. Acad., No. 1(15)/3.

[37] E. Galijašević and J. Kliewer, "Non-uniform near-perfect-reconstruction oversampled DFT filter banks based on allpass transforms," in *Proc. 9th IEEE DSP Workshop (DSP 2000)*, Hunt, TX, 15–18 Oct. 2000.

[38] A. Piotrowski and M. Parfieniuk, *Digital Filter Banks: Analysis, Synthesis, and Implementation for Multimedia Systems*. Bialystok, Poland: Wydawnictwo Politechniki Bialostockiej, 2006, in Polish.

[39] H. W. Löllmann and P. Vary, "Improved design of oversampled allpass transformed DFT filter-banks with near-perfect reconstruction," in *Proc. 15th European Signal Process. Conf. (EUSIPCO)*, Poznan, Poland, 3–7 Sep. 2007, pp. 50–54.

[40] S. K. Mitra and K. Hirano, "Digital all-pass networks," *IEEE Trans. Circuits Syst. I*, vol. 21, no. 5, pp. 688–700, Sep. 1974.

[41] B. Scharf, "Critical bands," in *Foundations of Modern Auditory Theory*, J. Tobias, Ed. New York, NY: Academic Press, 1970, pp. 159–202.

[42] C. D. Creusere and S. K. Mitra, "Efficient audio coding using perfect reconstruction noncausal IIR filter banks," *IEEE Trans. Speech Audio Process.*, vol. 4, no. 2, pp. 115–123, Mar. 1996.

[43] G. Evangelista and S. Cavaliere, "Time-varying frequency warping: Results and experiments," in *Proc. 2nd COST G-6 Workshop on Digital Audio Effects (DAFx)*, Trondheim, Norway, 9–11 Dec. 1999, pp. 13–16.

[44] M. Quélhas and A. Petraglia, "Optimum design of group delay equalizers," *Digital Signal Processing*, vol. 21, pp. 1–12, 2011.

[45] C. Feldbauer and G. Kubin, "Critically sampled frequency-warped perfect reconstruction filter bank," in *Proc. European Conf. on Circuit Theory and Design (ECCTD)*, vol. III, Cracow, Poland, 1–4 Sep. 2003, pp. III–109–III–112.

[46] R. G. Vaughan, N. L. Scott, and D. R. White, "The theory of bandpass sampling," *IEEE Trans. Signal Process.*, vol. 39, no. 9, pp. 1973–1984, Sep. 1991.

[47] P. Q. Hoang and P. P. Vaidyanathan, "Non-uniform multirate filter banks: Theory and design," in *Proc. IEEE Int. Symp. Circuits Systems (ISCAS)*, Portland, OR, 8–11 May 1989, pp. 371–374.

[48] J. D. Griesbach, M. R. Lightner, and D. M. Etter, "Subband adaptive filtering decimation constraints for oversampled nonuniform filterbanks," *IEEE Trans. Circuits Syst. II*, vol. 49, no. 10, pp. 677–681, Oct. 2002.