G. Vairetti, E. D. Sena, M. Catrysse, S. H. Jensen, M. Moonen and T. V. Waterschoot,
"An Automatic Design Procedure for Low-Order IIR Parametric Equalizers,"
J. Audio Eng. Soc., vol. 66, no. 11, pp. 935–952, (2018 November.).
DOI: https://doi.org/10.17743/jaes.2018.0049

An Automatic Design Procedure for Low-Order IIR Parametric Equalizers

GIACOMO VAIRETTI,[1,2] *AES Student Member*, **ENZO DE SENA**[3], **MICHAEL CATRYSSE**[4],
(giacomo.vairetti@esat.kuleuven.be) (e.desena@surrey.ac.uk) (m.catrysse@televic.com)

SØREN HOLDT JENSEN[5], **MARC MOONEN,**[1] *AES Associate Member*, **AND**
(shj@es.aau.dk) (marc.moonen@esat.kuleuven.be)

TOON VAN WATERSCHOOT,[1,2] *AES Associate Member*
(toon.vanwaterschoot@esat.kuleuven.be)

[1]*KU Leuven, Dept. of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, Kasteelpark Arenberg 10, 3001 Leuven, Belgium*
[2]*KU Leuven, Dept. of Electrical Engineering (ESAT), ETC, e-Media Research Lab, Andreas Vesaliusstraat 13, 3000 Leuven, Belgium*
[3]*Institute of Sound Recording, University of Surrey, Guildford, Surrey, GU2 7XH, UK*
[4]*Televic N.V., Leo Bekaertlaan 1, 8870 Izegem, Belgium*
[5]*Dept. of Electronic Systems, Aalborg University, Fredrik Bajers Vej 7, 9220 Aalborg, Denmark*

Parametric equalization of an acoustic system aims to compensate for the deviations of its response from a desired target response using parametric digital filters. An optimization procedure is presented for the automatic design of a low-order equalizer using parametric infinite impulse response (IIR) filters, specifically second-order peaking filters and first-order shelving filters. The proposed procedure minimizes the sum of square errors (SSE) between the system and the target complex frequency responses, instead of the commonly used difference in magnitudes, and exploits a previously unexplored orthogonality property of one particular type of parametric filter. This brings a series of advantages over the state-of-the-art procedures, such as an improved mathematical tractability of the equalization problem, with the possibility of computing analytical expressions for the gradients, an improved initialization of the parameters, including the global gain of the equalizer, the incorporation of shelving filters in the optimization procedure, and a more accentuated focus on the equalization of the more perceptually relevant frequency peaks. Examples of loudspeaker and room equalization are provided, as well as a note about extending the procedure to multi-point equalization and transfer function modeling.

0 INTRODUCTION

Parametric equalization of an acoustic system aims to compensate for the deviations of its response from a target response using parametric digital filters. The general purpose is to improve the perceived audio quality by correcting for linear distortions introduced by the system [1–4]. Linear distortions, usually perceived as spectral coloration (i.e., timbre modifications) [5, 6], are related to changes in the magnitude and phase of the complex frequency response with respect to a target response. Even though phase distortions are perceivable in some conditions [7], their effect is usually small compared to large variations in the magnitude of the frequency response [8]. Consequently, a low-order equalizer should focus on correcting the magnitude response of the system, rather than its phase response.

Parametric equalizers using cascaded infinite impulse response (IIR) filter sections consisting of peaking and shelving filters are commonly used [9–12], especially when a low-order equalizer is required. Indeed, the possibility of adjusting gain, central frequency, and bandwidth of each section of the equalizer results in a greater flexibility and, if the values of the parameters are well-chosen, in a reduced number of equalizer parameters w.r.t., for instance, a graphic equalizer with fixed central frequencies and bandwidths, or a finite impulse response (FIR) filter. However, since manually adjusting the values of the control parameters, as often done, can be difficult or may lead to unsatisfactory results, the availability of automatic design procedures is beneficial.

For a parametric equalizer design procedure to be fully automatic, various relevant aspects should be considered such as the number of filter sections available, typically

fixed between 3 and 30 based on the application, and the structure of the filter sections, which can have different characteristics and be parametrized in different ways, especially in terms of the bandwidth parameter [9]. Other design choices pertain the definition of a target response, based on a prototype or defined by the user, and its 0-dB line, relative to which the global gain of the equalizer will be set, as well as preprocessing operations, such as smoothing of the system frequency response. Once all these aspects are determined, an automatic design procedure requires the definition of an optimization criterion (or cost function), typically in terms of a distance between the equalized system magnitude response and the target magnitude response, as well as the choice of an optimization algorithm for the estimation of the parameter values of the filter sections. The focus of this paper is on automatic parametric equalizer design procedures operating in a sequential way, optimizing one filter section at a time, starting with the one that reduces the cost function the most, i.e., in order of importance in the equalization [13, 14]. The idea is to select an initial filter section, to search for better parameter values by minimizing the cost function using an iterative optimization algorithm and then move to the initialization and optimization of the next filter section.

The choice of the cost function has a fundamental role in determining the final performance of the design procedure. The characteristics of the first- and second-order peaking and shelving filters used in minimum-phase low-order parametric equalizers are well suited for the equalization of the magnitude response and have only a small influence on the phase response. As a consequence, the cost function generally chosen uses the difference between the magnitudes of the equalized response and the target response, discarding the phase response. The procedure described by Ramos et al. [13] uses a cost function that is the average absolute difference between the equalized magnitude response and the target magnitude response, computed on a logarithmic scale. More recently, Behrends et al. [14] proposed a series of modifications to the aforementioned procedure, including the evaluation of the cost function on a linear scale. Such a choice is meant to favor the equalization of frequency peaks, which are known to be more audible than dips [15]. This is a desirable feature, especially for low-order equalizers, which also limits the selection of filters producing a sharp boost in the response that may cause clipping in the audio system.

In the proposed procedure, the focus on equalizing peaks is even more prominent. The cost function employed uses the sum of squared errors (SSE) between the equalized and the target complex frequency responses. Minimizing the SSE does not explicitly aim at maximizing the "flatness" of the equalized magnitude response, as for the procedures cited above, but rather at compensating for the deviations of the equalized response by putting more emphasis in the equalization of energetic frequency peaks over dips. Even though the use of the SSE may be a less intuitive way of defining the equalization problem, it brings some advantages over using the magnitude response error. Specifically, the SSE gives the possibility of computing analytical expressions for the gradients of the cost function w.r.t. the parameters of the filter sections, such that efficient line search optimization algorithms can be used, and of estimating the global gain of the equalizer (i.e., the 0-dB line). Moreover, if only the linear-in-the-gain structure of the parametric filters [9, 10] is used, the gain parameters can be estimated in closed form using least squares (LS), thus enabling the use of a grid search procedure for the initialization of the other filter parameters, as well as the inclusion of first-order shelving filters in the optimization procedure. It follows that most of the design aspects to be considered are based on the minimization of the cost function and not on arbitrary choices or assumptions regarding the magnitude response to be equalized, as in the procedures in [13] and [14] briefly described in Sec. 1.

The present paper is organized as follows: Sec. 1 gives an overview of the state-of-the-art procedures for automatic equalizer design using parametric IIR filters. Sec. 2 formalizes and discusses the equalization problem defined in terms of the SSE. In Sec. 3, linear-in-the-gain (LIG) parametric IIR filters are described and the closed-form expression for the gain parameter is derived. The proposed automatic procedure for parameter estimation of a low-order parametric equalizer is detailed in Sec. 4. In Sec. 5 results of the equalization of a loudspeaker response are evaluated using different error-based objective measures [4], as well as objective measures of perceived audio quality [6, 16, 17]. In Sec. 6 application to room response equalization is also considered. The modification to the proposed procedure for multi-point equalization and transfer function modeling is briefly discussed in Sec. 7. Sec. 8 concludes the paper.

0.1 Terminology

The following terms and conventions are defined and used throughout the paper. The term *system response* $H_0(k)$ indicates the frequency response to be equalized, which could be either a loudspeaker response, a room response, or a joint loudspeaker-room response. The radial frequency index k refers to the evaluation of the transfer function on the unit circle at the k^{th} radial frequency bin ω_k (k is short for $e^{j\omega_k/f_s}$, with f_s the sampling frequency). The *equalized response* $H_s(k)$ is defined as the system response filtered by the parametric equalizer having s filter sections. The term *parametric equalizer* refers to the cascade of S parametric filters, while the term *parametric filter* refers to either a *peaking* filter with filter order $m = 2$ or a *shelving* filter with filter order $m = 1$. A parametric filter has two possible implementation forms: a LIG form, typically used in the literature with a positive gain (in dB) to generate a *boost* in the filter response, and a nonlinear-in-the-gain (NLIG) form, typically used with a negative gain (in dB) to generate a *cut* in the filter response (see Sec. 3).

1 STATE-OF-THE-ART PROCEDURES

The purpose of parametric equalization is to compensate for the deviations of the system frequency response $H_0(k)$ from a user-defined target frequency response $T(k)$ using

a parametric equalizer of order M with overall response $F_M(k)$. In other words, the purpose is to filter $H_0(k)$ with the equalizer $F_M(k)$ in order to approximate the target response as closely as possible based on the following error:

$$E_M(k) = W(k)\big\{H_0(k) \cdot F_M(k) - T(k)\big\}. \tag{1}$$

with $W(k)$ a weighting function used to give more or less importance to the error at certain frequencies.

Different cost functions are possible. In the procedure proposed by Ramos et al. [13], the mean absolute error between the magnitudes in dB of the equalized response and the target response, computed on a logarithmic frequency scale, was chosen to account for the "double logarithmic behavior of the ear,"

$$\epsilon_M^{dB} = \frac{20}{N}\sum_k \Big| W(k)\Big\{\log_{10}|H_0(k)\cdot F_M(k)| - \log_{10}|T(k)|\Big\}\Big|, \tag{2}$$

with N the number of frequencies included in the frequency range of interest. The system magnitude response $|H_0(k)|$, as commonly done in low-order parametric equalization, is smoothed by a certain fractional-octave factor (usually $\frac{1}{8^{\text{th}}}$ or $\frac{1}{12^{\text{th}}}$) in order to remove narrow peaks and dips that are less audible [15] and to facilitate the search for the optimal parameter values. For each filter section, the procedure in [13] uses a heuristic algorithm to optimize the parameters. The procedure was extended in [18] to include second-order shelving and high-pass (HP) and low-pass (LP) filters in the equalizer design. The decision of including shelving filters has to be made by analyzing the error areas above and below the target magnitude response at the beginning and at the end of the frequency range of interest. A shelving (or HP/LP) filter is then included if the error area is larger than a predefined threshold, with the values of the filter parameters optimized using the same heuristic algorithm. Another extension proposed in [13] adds the possibility of reducing the order of the parametric equalizer by removing the peaking filters that are correcting for inaudible peaks and dips, according to psychoacoustic considerations [15].

In Behrends et al. [14] the higher perceptual relevance of spectral peaks is directly taken into account in the definition of the cost function by considering the error on a linear magnitude scale instead of a logarithmic scale, i.e.,

$$\epsilon_M^{\text{lin}} = \frac{1}{N}\sum_k \Big| W(k)\Big\{|H_0(k)\cdot F_M(k)| - |T(k)|\Big\}\Big|. \tag{3}$$

While the cost function used in Eq. (2) equally weights the error produced by deviations of the equalized magnitude response above and below the target response, the evaluation of the cost function on a linear scale as in Eq. (3) gives more importance to the portions of the equalized magnitude response that lie above the target, thus favoring the removal of frequency peaks, rather than the filling of the dips. In [14] Behrends et al. also suggest to employ a derivative-free algorithm, called the Rosenbrock method [19], which offers a gradient-like behavior and thus faster convergence.

A critical aspect of the procedures by Ramos et al. [13] and Behrends et al. [14] is the selection of the initial values of the parameters of each new parametric filter. The selection is done by computing the areas of the magnitude response above and below the target using either Eq. (2) or Eq. (3). The largest area becomes the one to be equalized, with the half-way point between the two zero-crossing points and the negation of its level (in dB) defining the central frequency and gain of the filter section, respectively, and the –3 dB points defining the bandwidth (or Q-value). This approach assumes that the system magnitude response is a combination of peaks and dips above and below the target magnitude response. The problem with such an assumption is that, in case of highly irregular system magnitude responses, the initial filter placement approach may provide initial values quite distant from a local minimizer. In this case, the reduction in the cost function provided by the initial filter may even be quite limited. Furthermore, the placement of the 0-dB line becomes an important aspect of the procedure for which a clear solution was not provided.

An example system magnitude response, similar to an example in [14], is given in Fig. 1 also showing the filter responses for the initial values computed with different procedures. Between 100 Hz and 16 kHz, there are seven error areas $A_1 - A_7$ above and below the predefined flat target magnitude response. The procedure by Ramos et al. [13] places the initial filter based on the largest error area computed according to Eq. (2), which is A_3 in the example; the parameters of the initial filter are chosen as described above; the irregularity of the system magnitude response makes the selection based on the half-way point of the area far from optimal, with the initial filter far from the optimal solution (also shown in the figure). The largest error area for the procedure in [14], computed according to Eq. (3), is instead A_2. As shown in the figure, using the same approach as in [13] leads to similar problems. A peak finding approach, as also suggested in [14], may provide a better initialization in this particular example, but it may not be effective in general and introduces the problem of defining the initial value for the bandwidth. The initial filter obtained with the proposed procedure is also shown in the figure. The initialization, which will be described in Sec. 4, is not based on the largest error area approach but on a grid search with optimal gain (in LS sense) computed w.r.t. the SSE. It can be seen that initial parameters are found quite close to the optimal ones.

Other examples of automatic parametric equalizer design can be found in [20], where nonlinear optimization is used to find the parameters of a parametric equalizer starting from initial values selected using peak finding; in [21], where the gains of a parametric equalizer with fixed frequencies and bandwidths are estimated in closed form exploiting a self-similarity property of the peaking filters on a logarithmic scale; and in [22], where a gradient-based optimization of the parameters of an equalizer is proposed, which uses filters parametrized using the numerator and denominator coefficients of the transfer function and not a constrained form defined in terms of gain, frequency, and bandwidth, as the one used in this paper.

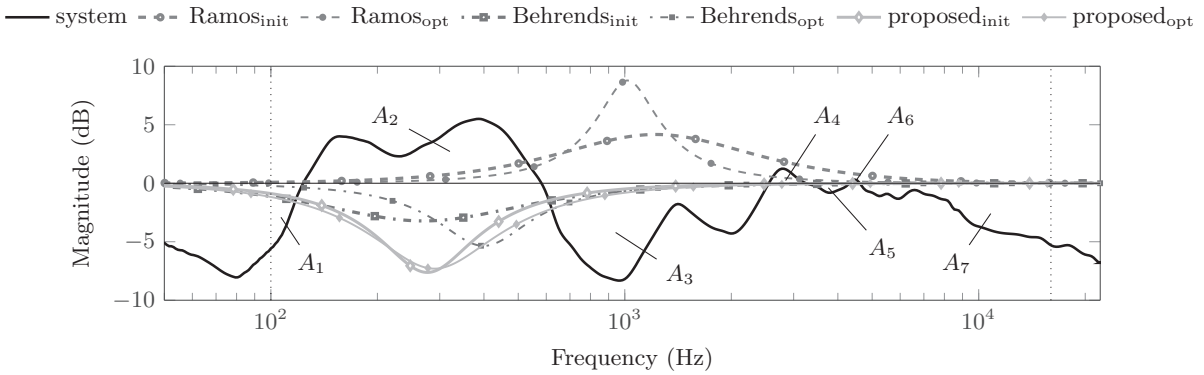

Fig. 1. Initialized (thick lines) and optimized (thin lines) responses of a single filter section using different procedures.

2 EQUALIZATION BASED ON THE SSE

In this paper a cost function is used based on the SSE between the frequency responses, i.e.,

$$\epsilon_M^{\text{SSE}} = \frac{1}{N}\sum_k \big(W(k)\big[H_0(k)\cdot F_M(k) - T(k)\big]\big)^2. \tag{4}$$

Such formulation, even though less intuitive than Eqs. (2) and (3), brings some advantages as will be detailed later on: (i) it provides an improved mathematical tractability of the equalization problem, with the possibility of computing analytical expressions for the gradients w.r.t. the filter parameters; (ii) when the parametric filter is in the LIG implementation form, it leads to a closed-form expression for the gain parameters (see Sec. 3), which simplifies the automatic design procedure; (iii) it provides a better way to initialize a parametric filter prior to optimization; (iv) it allows to include first-order shelving filters; (v) to estimate the global constant gain in closed-form; and (iv) it focuses on the equalization of the more perceptually relevant frequency peaks rather than the dips.

The parametric equalizer considered, comprising a cascade of minimum-phase parametric filters, has a minimum-phase response. An interesting property of a minimum-phase response is that its frequency response $H(\omega)$ is completely determined by its magnitude response. The phase $\phi_H(\omega)$ is, indeed, given by the inverse Hilbert transform $\mathcal{H}^{-1}\{\cdot\} = -\mathcal{H}\{\cdot\}$ of the natural logarithm of the magnitude [23, 24]:

$$\begin{aligned} &H(\omega) = |H(\omega)|e^{j\phi_H(\omega)}, \\ &\text{with } \phi_H(\omega) = -\mathcal{H}\{\ln|H(\omega)|\}. \end{aligned} \tag{5}$$

This is a consequence of the fact that the log frequency response is an *analytic signal* in the frequency domain

$$\ln H(\omega) = \ln|H(\omega)| + j\phi_H(\omega), \tag{6}$$

whose time-domain counterpart is the so-called *cepstrum* [23]. In the digital domain, the phase response of the minimum-phase frequency response $H(k)$ can be obtained as the imaginary part $\mathcal{I}$ of the DFT of the folded real periodic cepstrum $\hat{h}(n) = \text{IDFT}\{\ln|H(k)|\}$

$$\phi_H(k) = \mathcal{I}\{\text{DFT}\{\text{fold}\{\hat{h}(n)\}\}\} \tag{7}$$

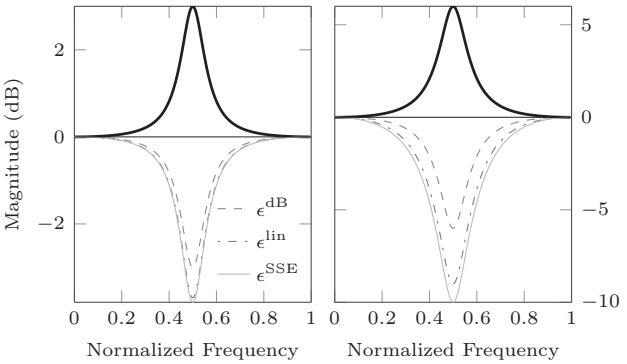

Fig. 2. Two peaking filters with gains $G = 3$ dB and $G = 6$ dB (thick lines) and the corresponding cut filter responses (thin lines) with gain optimized to give equal error using different cost functions.

where the DFT and IDFT operators indicate the discrete Fourier transform and its inverse, and the *fold* operation has the effect of folding the anti-causal part of $\hat{h}(n)$ onto its causal part. More details can be found in [25] or [26]. Thus, given the relation between the magnitude and the phase of a minimum-phase frequency response as given in Eq. (5), minimizing the cost function in Eq.(4), remarkably, still corresponds to a magnitude-only equalization.

The use of the SSE in Eq. (4) compared to the linear function in Eq. (3) puts more emphasis on the error generated by strong peaks, as described in more detail in Appendix A.2. Here an intuitive interpretation is given as follows. In Fig. 2, the boost magnitude response of two peaking filters with positive gains $G = 3$ dB and $G = 6$ dB is considered. A cut in the filter magnitude response, having the same central frequency and bandwidth, is obtained using a negative gain. The negative gain parameter is optimized such that the error w.r.t. the 0-dB line computed with the cost functions in Eqs. (2), (3), and (4) is equal to the one obtained for the boost response. For the cost function in Eq. (2), the cut filter response is obviously specular to the boost filter response on a logarithmic scale (the gain is $-G$), whereas for the cost functions in Eq. (3) and Eq. (4) it is not. This is the consequence of the fact that the evaluation of the error on a linear scale puts more weight on values above the 0-dB line. Whereas for the $G = 3$ dB gain case (left plot) the cost

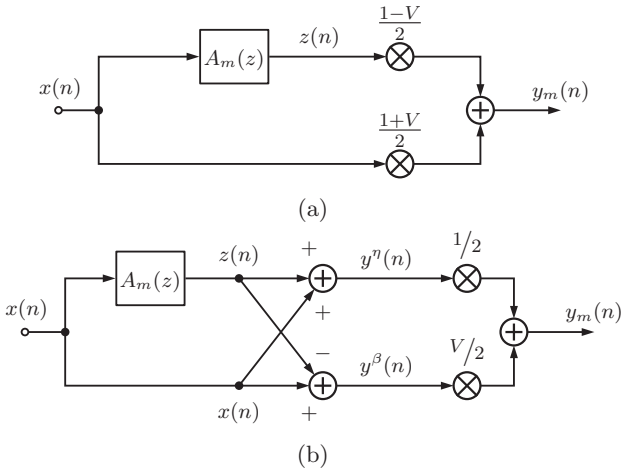

Fig. 3. The Regalia-Mitra parametric filter

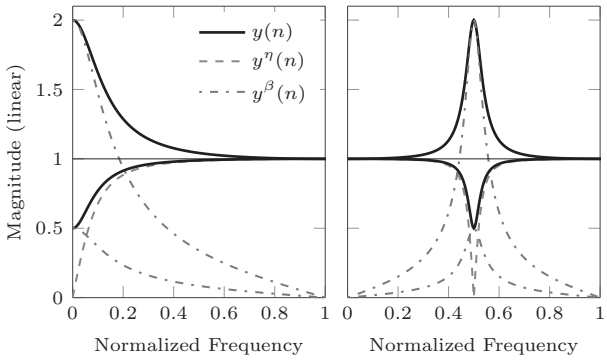

Fig. 4. Shelving and peaking filters in LIG form

functions in Eqs. (3) and (4) produce almost the same error; for higher gains (see right plot for $G = 6$ dB) the SSE gives more emphasis to errors above the 0-dB line.

3 LINEAR-IN-THE-GAIN PARAMETRIC FILTERS

Digital IIR filters used in parametric equalizers are first- and second-order IIR filters, with constraints on the filter magnitude response defined at the zero frequency, at the Nyquist frequency, and, for peaking filters, at the central frequency. Different parameterizations satisfying these constraints are possible, with different methods to compute the filter coefficients. However, even though the various parameterizations have different definitions for the bandwidth parameter, all parameterizations satisfying the same constraints are equivalent [9].

Among different possibilities, the structure of first- and second-order parametric filters originally proposed by Regalia and Mitra [10] is chosen here. This structure, shown in Fig. 3, comprises an all-pass (AP) filter $A_m(z)$ of order m and a feed-forward path. If the AP filter is independent from the gain parameter V, the parametric filter has a transfer function $F_m(z)$ which is linear in V,

$$F_m(z) = \frac{1}{2}[(1+V) + (1-V)A_m(z)] \tag{8}$$

$$= \frac{1}{2}[(1 + A_m(z)) + V(1 - A_m(z))], \tag{9}$$

where expression Eq. (9), corresponding to the equivalent filter structure in Fig. 3b, highlights this linear dependency [11, 12]. Given that for $V > 0$ the filter response is minimum-phase, whereas for $V < 0$ it is maximum-phase [10], only filters with positive linear gain will be considered.

Another characteristic of this filter structure, which is exploited in the proposed procedure, follows from the energy preservation property [27] of the AP filter: since the energy of the output signal of the AP filter is equal to the energy of its input signal, the signals $y^\eta(n) = x(n) + z(n)$, corresponding to a notch, and $y^\beta(n) = x(n) - z(n)$, corresponding to a resonance, are found to be orthogonal to each other. An intuitive proof is provided in Appendix A.3. It follows that, when the gain parameter V does not appear in the AP filter transfer function, the gain V is only acting on the resonant response $y^\beta(n)$, whereas the notch response $y^\eta(n)$ is not changed when V is modified. This can be seen in Fig. 4, showing the magnitude response of two shelving filters (left) and two peaking (right) filters in LIG form with gains $V = 2$ and $V = 0.5$, together with the corresponding notch and resonance responses. If should be noticed that the LIG filter structure is able to produce both a boost and a cut in the response, even though the cut response tends to have a reduced bandwidth [10], as discussed below.

3.1 First-Order Shelving Filters

A shelving filter is used whenever the lowest or highest portion of the system frequency response has to be enhanced or reduced. Shelving filters are described by a set of two parameters, namely the gain V and the transition frequency f_c, defined as the –3 dB notch bandwidth. By using the filter structure in Eq. (8) or Eq. (9), a first-order shelving filter at low frequency (LFs) or at high frequencies (HFs), respectively, by defining a first-order AP filter as

$$A_1^{\mathrm{LF}}(z) = \frac{a_{\mathrm{LF}} - z^{-1}}{1 - a_{\mathrm{LF}} z^{-1}}, \quad A_1^{\mathrm{HF}}(z) = \frac{a_{\mathrm{HF}} + z^{-1}}{1 + a_{\mathrm{HF}} z^{-1}}. \tag{10}$$

The LIG form is obtained by defining the parameter a in terms of the transition frequency f_c and the sampling frequency f_s as

$$a^b_{\mathrm{LF}} = \frac{1 - \tan(\pi f_c/f_s)}{1 + \tan(\pi f_c/f_s)}, \quad a^b_{\mathrm{HF}} = \frac{\tan(\pi f_c/f_s) - 1}{\tan(\pi f_c/f_s) + 1}. \tag{11}$$

As a consequence, the AP filter does not depend on the gain V. However, for $0 < V < 1$, when the filter represents a cut, the effective transition frequency of the filter response tends towards lower (or higher for the HF case) frequencies (see left plot of Fig. 4 or [10]). To obtain a cut response, for $0 < V < 1$, with response specular to the one obtained with the LIG form when V is replaced by $\frac{1}{V}$, the parameter a has to be modified to be dependent on the gain [12],

$$a^c_{\mathrm{LF}} = \frac{V - \tan(\pi f_c/f_s)}{V + \tan(\pi f_c/f_s)}, \quad a^c_{\mathrm{HF}} = \frac{\tan(\pi f_c/f_s) - V}{\tan(\pi f_c/f_s) + V} \tag{12}$$

which yields the NLIG form of a shelving filter.

Another option would be to redefine the parameter a in order to obtain a single expression that provides specular responses for a boost with gain V and a cut with gain $\frac{1}{V}$ [1, 9, 28]. However, the resulting filter structure of the *proportional* shelving filter is nonlinear in the gain parameter.

Finally, it should be noticed, also from the left plot of Fig. 4, that the notch response $y^{\eta}(n)$ of the LF shelving filter corresponds to a first-order HP filter (i.e., when $V = 0$). The same is true also for the notch response of the HF shelving filter, which corresponds to a first-order LP filter.

3.2 Second-Order Peaking Filters

Peaking filters are used to compensate for peaks or dips in the system magnitude response. As for first-order shelving filters, second-order peaking filters can be implemented with the filter structure in Eq. (8) by defining a second-order AP filter as

$$A_2(z) = \frac{a + d(1+a)z^{-1} + z^{-2}}{1 + d(1+a)z^{-1} + az^{-2}}, \tag{13}$$

with $d = -\cos(2\pi f_0/fs)$, where f_0 is the central frequency of the peaking filter. The LIG form is obtained by defining the bandwidth parameter a as

$$a^b = -\frac{\tan(\pi f_b/f_s) - 1}{\tan(\pi f_b/f_s) + 1}, \tag{14}$$

with f_b defined as the –3 dB notch bandwidth obtained for $V = 0$ [9, 10]. Similar to first-order shelving filters, peaking filters do not show a specular response when replacing V by $\frac{1}{V}$ (see right plot of Fig. 4 or [10]). In order to obtain symmetric boost and cut responses, either the NLIG form [12] for $0 < V < 1$, with

$$a^c = -\frac{\tan(\pi f_b/f_s) - V}{\tan(\pi f_b/f_s) + V}, \tag{15}$$

or the proportional filters in [1, 9, 28] could be used. In both cases, the linear dependency w.r.t the gain parameter is lost. Only the LIG form is used in the proposed automatic equalization procedure. It is possible in any case to convert the parameters of a filter, either shelving or peaking, from the LIG form to the NLIG or the proportional form.

3.3 LS Solution for the Gain Parameter

The advantage of the LIG form is that the linearity and orthogonality properties described above enable a closed-form solution for the estimation problem of the gain parameter. When the equalizer is made of only one parametric filter, the cost function in Eq. (4) can be written as

$$\epsilon_m^{\mathrm{SSE}} = \frac{1}{N}\sum_k \Big(W(k)\Big\{\frac{1}{2}H_0(k)[F_m^{\eta}(k) + V\,F_m^{\beta}(k)] - T(k)\Big\}\Big)^2, \tag{16}$$

where $F_m^{\eta}(k) = 1 + A_m(k)$ and $F_m^{\beta}(k) = 1 - A_m(k)$, respectively, and $k = 1, \ldots, N$. The minimization of the cost function is performed by setting to zero the first-order partial derivative of $\epsilon_M^{\mathrm{SSE}}$ w.r.t. V. The LS solution is obtained by

$$\hat{V} = \frac{\sum_k |W(k)|^2\, F_m^{\beta}(k) H_0^*(k) T(k)}{\sum_k |W(k)|^2 |H_0(k)|^2 |F_m^{\beta}(k)|^2} \tag{17}$$

with $\{\,\cdot\,\}^*$ indicating complex conjugation, which is independent from $F_m^{\eta}(k)$ because of the orthogonality between $F_m^{\eta}(k)$ and $F_m^{\beta}(k)$ (see details in Appendix A.3 and A.4). This feature will be also used in the parameter initialization as described in Sec. 4. Indeed, if the equalizer is designed one parametric filter at a time, the optimal value $\hat{V}_s$ of the gain parameter of the s^{th} filter section, is obtained by substituting the system frequency response $H_0(z)$ in Eq. (17) with the equalized response $H_{s-1}(z)$.

4 PROPOSED DESIGN PROCEDURE

The aim of the proposed procedure is to design a parametric equalizer of order M as a cascade of S filter sections, each consisting of a parametric filter of order $m_s = 1$ (shelving) or $m_s = 2$ (peaking) having frequency response $F_{m_s}(k)$ defined as in Eqs. (8–9), i.e.,

$$F_M(k) = C\prod_{s=1}^{S} F_{m_s}(k), \text{ with } M = \sum_{s=1}^{S} m_s, \tag{18}$$

where s indicates the filter section index and C a global gain. The parameter values of the s^{th} filter section are optimized so as to minimize the cost function $\mathcal{F}(a_s, d_s, V_s)$, defined as

$$\epsilon_{s,m_s}^{\mathrm{SSE}} = \frac{1}{N}\sum_k \Big(W(k)\Big\{H_{s-1}(k)F_{m_s}(k) - T(k)\Big\}\Big)^2, \tag{19}$$

with H_{s-1} the system response filtered by the equalizer comprising the previous $s-1$ filter sections.

The proposed design procedure consists of the steps depicted in Fig. 5 and detailed in the rest of the section. A preliminary step is to define a target response $T(k)$ and a minimum-phase preprocessed version of the system response $H_0(k)$. Optionally, the value of the global gain C can be estimated in closed-form using LS. The design of each new filter section can be divided into two stages. The first stage provides initial parameter values by means of a grid search, in which the optimal gain parameter for predefined discrete values of the central frequency and bandwidth is estimated as described above. The second stage consists of a line search optimization, which is intended to iteratively refine the initial parameter values and reach a local minimum of the cost function.

4.1 Spectral Preprocessing

The spectral preprocessing of the system frequency response follows the steps outlined in [25]: first, the system magnitude response $|H_0(k)|$ is smoothed according to the Bark frequency scale, in order to approximate the critical bands of the ear, using a moving-average (MA) filter with bandwidth increasing with frequency. Apart for frequencies below 500 Hz, at which the smoothing is performed

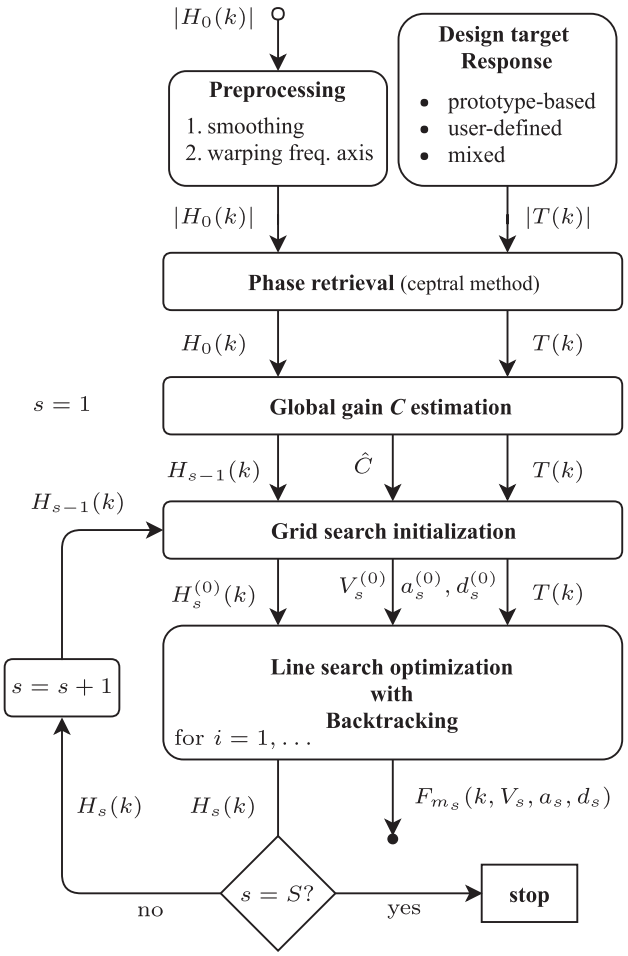

Fig. 5. Schematics of the proposed design procedure.

over a fixed 100 Hz interval, the bandwidth of the filter is set to an interval equal to 20% of the frequency. The amount of smoothing can then be controlled by the length of the window of the MA filter; either fractional critical bandwidth smoothing or fractional-octave smoothing can be easily used instead.

The second (and optional) step of the spectral preprocessing in [25] is to warp the frequency axis in order to approximate the Bark frequency scale, i.e., to allocate a higher resolution to the LFs. An alternative, also adopted in this paper, is to resample the frequency axis from linear to logarithmic, by defining a logarithmically spaced axis, e.g., with $\frac{1}{48^{\text{th}}}$-octave resolution as in [13] and thus evaluating the magnitude response at those frequency points (e.g., using Horner's method [23], after the phase retrieval step explained below). Yet another way of favoring the equalization of a given frequency range, which can be used in conjunction with the strategies above, is to tune the weighting function $W(k)$ in Eq. (19) accordingly.

Finally, the cost function in Eq. (19) requires the minimum-phase response $H_0(k)$ to be retrieved from the preprocessed system magnitude response. A common solution, also suggested in [25], to create a minimum-phase frequency response is by means of the cepstral method [23, 25], where the smoothed (and/or warped) magnitude response is used to retrieve the corresponding phase response, as given in Eqs. (5–7). Notice that, in order to avoid time-aliasing given by deep notches that can remain in the magnitude response after smoothing (e.g., towards 0 Hz), it is advisable to increase the FFT size to a high power of two and to clip the response as suggested in [26].

4.2 Target Response

Although the choice of the target response is arbitrary, it should be made cautiously. If the target response is too distant from the system frequency response, the equalization will be more difficult to be realized. For instance, if the lower cut-off frequency of the target response is below the lower cut-off frequency of the loudspeaker, the equalizer would contain a parametric filter with positive gain, which would move the loudspeaker driver outside its working range.

There is no complete agreement on the optimal target response for loudspeaker/room response equalization, and no single target for all sound reproduction purposes and all listeners can be defined [29]. It is out of the scope of this paper to discuss the characteristics of an optimal, according to some criterion, target response for different sound reproduction systems and situations. Here only a brief overview of different approaches and guidelines is given. The target response can be defined in its magnitude and then its phase can be retrieved with the cepstral method.

4.2.1 Prototype-Based

A prototype target magnitude response can be defined as, e.g., a band-pass filter transfer function or the magnitude response of a different loudspeaker. In this case particular attention should be given to matching the cut-off frequencies of the system magnitude response and of the prototype target response, in order to avoid overloading of the loudspeaker driver. Another option is to use a strongly smoothed version of the system magnitude response, such as the one-octave smoothed response [20] or smoothing based on power averaged sound pressure [30], which eliminates peaks and dips while preserving the coarse spectral envelope of the system response.

4.2.2 User-Defined

A target magnitude response can be obtained as an interpolation of a set of points defined w.r.t. the system magnitude response [14]. In this way it is easy to match the cut-off frequencies of the system magnitude response and to determine any desired characteristic of the response in the pass-band.

4.2.3 Mixed Strategies

A combination of the two approaches can be used. For instance, the target magnitude response may be obtained by smoothing the system magnitude response in the LFs and in the HFs, whereas the response in the middle range may be defined by the user, e.g., a flat response or a boost at LFs.

4.3 Optimal Global Gain

Another aspect to consider is the optimization of the global gain C of the parametric equalizer, or, equivalently, the setting of the 0-dB line. Indeed, this has an influence on the characteristics of the filters selected by the design procedure. Centering a loudspeaker response around 0 dB would most likely avoid the selection of wide-band filters. However, in case of a room response, it is more difficult to determine the level at which the response should be centered, so that wide-band filters, with possibly high gains, are more likely to be selected, especially if the target response is not chosen carefully.

As described in Sec. 1, the placement of the 0-dB line is a critical aspect in the procedures proposed in [13] and [14]; the requirement for the system magnitude response to be centered around the 0-dB line of the target response in order to create error areas to be equalized is somewhat arbitrary. A possibility would be to place the 0-dB line by visual inspection or as the mean of the magnitude response of the system within a frequency range of interest (e.g., mid frequencies). This solution is not guaranteed to be an optimal one.

The use of the cost function based on the SSE, instead, allows the estimation of a global gain using LS, similarly to the estimation of the linear gain described in Sec. 3.3; by replacing the parametric equalizer $F_M(k)$ in Eq. (4) by a constant C, an estimate for the global gain $\hat{C}$ is given as

$$\hat{C} = \frac{\sum_k |W(k)|^2 \, H_0^*(k) T(k)}{\sum_k |W(k)|^2 |H_0(k)|^2} \tag{20}$$

This global gain C can be regarded as a scaling factor that centers the system response around the 0-dB line that minimizes the cost function in Eq. (4). Since the SSE puts more emphasis on the peaks (see Sec. 2), the system magnitude response will tend to have dips that are more prominent than the peaks w.r.t. the target response. This may not be desirable, as the design procedure may favor the boost of spectral dips rather than the cut of spectral peaks. If desired, this may be avoided by adding an offset of a few dB to the global gain in order to restore the emphasis on the equalization of peaks over dips.

4.4 Grid Search Initialization and Constraints

The initialization of the parameters of each new parametric filter in the cascade, as well as the selection of either a peaking or a shelving filter, is performed in an automatic way by means of a grid search using a discrete set of possible frequency and bandwidth values. A pole grid is defined, similarly to [31], where the radius and angle of complex poles determine respectively the bandwidth f_b and central frequency f_0 of the peaking filters. The radius of the real poles defines the transition frequencies f_c of LF (positive real poles) and HF (negative real poles) shelving filters. The gain for the filters built using each pole p in the grid is defined by LS estimation as described in Sec. 3.3, and the parameters of the filter that reduces the SSE the most are selected as initial parameter values of the current filter section. The gain can be limited based on hardware specifications, by defining a minimum (e.g., $V_{\min} = 0.25$) and a maximum value (e.g., $V_{\max} = 4$). Note that, being the system response minimum-phase, the gain V will always be positive [10]. The rest of this section discusses a strategy for constructing the pole grid, of which an example is given in Fig. 7 and a way to impose and check constraints on the bandwidth in terms of the commonly used Q-factor.

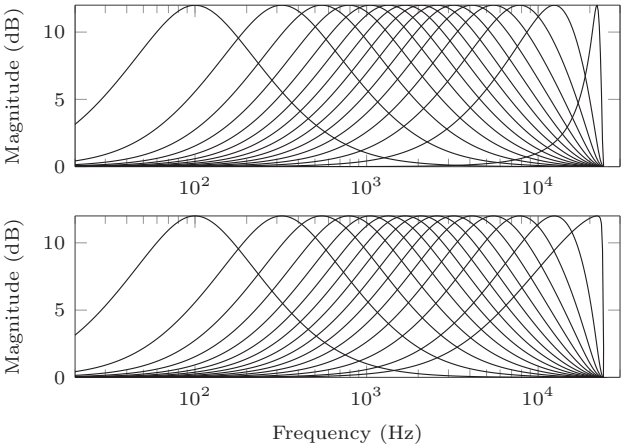

Fig. 6. Magnitude response of constant-Q (top) and constant relative bandwidth (bottom) peaking filters.

Given the critical-band smoothing and the logarithmic resolution of the frequency axis, the angle $\sigma = 2\pi f_0/f_s$ of the complex poles, which define peaking filters, can be discretized according to a logarithmic or a Bark-scale distribution with minimum and maximum angles defined, for instance, by the frequency limits of the equalization. The radius $\rho = \sqrt{a}$ of the complex poles $p = \rho e^{j\sigma}$ can be defined between a lower and an upper limit determined by the constraints imposed on the gain and bandwidth parameters for the different values of σ. It is common to define constraints in terms of the Q-factor, which provides an indication of the filter bandwidth relative to its central frequency [12]. The parameter a can be converted into the corresponding Q-factor in closed form, but the two cases of $V > 1$ and $V < 1$ must be addressed separately. Filters in the LIG form defined in terms of the parameters a and d (see Sec. 3.2), can be converted in the corresponding LIG *boost* form and NLIG *cut* forms defined in terms of Q and the auxiliary variable $K = \tan(\pi f_0/f_s)$ as in [12], respectively with

$$Q^b = \frac{\sin(2\pi f_0/f_s)}{2\tan(\pi f_b/f_s)} = \frac{\sin(\sigma)}{2\frac{1-a^b}{1+a^b}} \text{if } V > 1\,, \tag{21}$$

$$Q^c = \frac{\sin(2\pi f_0/f_s)}{2V\tan(\pi f_b/f_s)} = \frac{\sin(\sigma)}{2V\frac{1-a^b}{1+a^b}} \text{if } V < 1\,. \tag{22}$$

The Q-factor can be limited as well in order to avoid filters too narrow-band (e.g., $Q_{\max} = 10$) or too wide-band (e.g., $Q_{\min} = 0.5$).

However, for given fixed values of Q and V, the bandwidth (in octaves) of a peaking filter is not actually constant, but it reduces for increasing frequencies so that the filter response on a logarithmic scale becomes asymmetric when f_0 approaches $\frac{f_s}{2}$ (top plot of Fig. 6).

In order to keep the relative bandwidth approximately constant over the whole frequency range (bottom plot of Fig. 6), the radius ρ of the complex poles is set to decrease exponentially with increasing angle σ, according to $\rho = R^{\frac{\sigma}{\pi}}$, with R the value of the radius defined at the Nyquist frequency [32]. The value for R can be computed to match the response of a filter defined in terms of a given Q [12] at a given angular frequency σ_q. The parameter a_q is computed from Eqs. (21–22) as

$$a_q^b = \frac{2Q^b - \sin(\sigma_q)}{2Q^b + \sin(\sigma_q)} \quad \text{if } V > 1, \tag{23}$$

$$a_q^c = \frac{2VQ^c - \sin(\sigma_q)}{2VQ^c + \sin(\sigma_q)} \quad \text{if } V < 1, \tag{24}$$

from which the corresponding $R = a_q^{\frac{\pi}{2\sigma_q}}$ is obtained. The limits for R are computed the same way inserting the constraints in Eqs. (23–24). The minimum and maximum radius at the Nyquist frequency for $V > 1$ ($R^b_{\min}$, $R^b_{\max}$) are computed from Eq. (23) for $Q = Q^b_{\min}$ and $Q = Q^b_{\max}$, whereas for $V < 1$, $R^c_{\min}$ and $R^c_{\max}$ are computed from Eq. (24) for $Q = Q^c_{\min}$ and $Q = Q^c_{\max}$, with $V = V_{\min}$. This results in two partially overlapping allowed areas of the unit disc, one valid when $V > 1$ and the other when $V < 1$, where generally $R^c_{\min} < R^b_{\min}$ and $R^c_{\max} < R^b_{\max}$.

In general, the bandwidth constraints for filter with $V > 1$ (Q^b) and filters with $V < 1$ (Q^c) can be chosen to be different, with the limitation dictated by the requirement of having a positive value for a (and thus ρ real). From Eq. (24) with $\sigma_q = \frac{\pi}{2}$, it is required that $Q^c_{\min} > \frac{1}{2V_{\min}}$, thus trading-off between sharp cut filters with high gain and broader cut filters with limited gain. Also, it is required from Eq. (23) that $Q^b_{\min} > 0.5$ (which is anyway quite wide, approximately 2.5 octaves). Notice that for very large bandwidths, the filter responses tend to skew towards the Nyquist frequency but less dramatically than for the filters with fixed Q (see Fig. 6). A unique allowed area could be found by setting $Q^c_{\min} = \frac{Q^b_{\min}}{V_{\min}}$ and $Q^c_{\max} = \frac{Q^b_{\max}}{V_{\min}}$, but this would lead to filters with cut responses ($V < 1$) much narrower compared to boost responses ($V > 1$).

Regarding the values for R between $R^c_{\min}$ and $R^b_{\max}$, it is suggested in [31] to set the desired number of radii (for each angle) and distribute them logarithmically in order to increase density towards the unit circle (obtaining the so-called Bark-exp grid [31]) and thus to increase the resolution of narrow peaking filters. If the allowed areas do not coincide, the complex poles with smaller radius are valid only for $\hat{V} < 1$ (i.e., cut responses), whereas they would produce too wide boost responses for $\hat{V} > 1$. On the other hand, complex poles very close to the unit circle, valid for $\hat{V} > 1$, would produce too narrow cut responses for $\hat{V} < 1$. It is then necessary to check the constraints after the estimation of the optimal gains $\hat{V}$ and select the initial filter as the one that minimizes the cost function within the constraints. This can be done by checking that the parameter $a_s = \rho_s^2$ of the selected complex pole $p_s = \rho_s e^{j\sigma_s}$ satisfies $a^b_{\min} \le a_s \le a^b_{\max}$ or $a^c_{\min} \le a_s \le a^c_{\max}$, where $a^b_{\min}$ and $a^b_{\max}$ are computed from Eq. (23) for $Q = Q^b_{\min}$ and $Q = Q^b_{\max}$, and $a^c_{\min}$ and $a^c_{\max}$ from Eq. (24) for $Q = Q^c_{\min}$ and $Q = Q^c_{\max}$, with $V = V_{\min}$, where σ_q is replaced by σ_s.

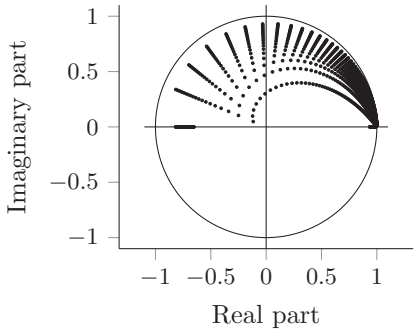

Fig. 7. A Bark-exp pole grid for the grid-search.

Finally, the radius of the real poles, determining the transition frequency f_c of the shelving filters, may be set arbitrarily within the range of equalization. The effective transition frequency corresponding to ρ can be easily computed from Eq. (11) and Eq. (12), for $V > 1$ and $V < 1$, respectively. An upper and a lower limit for the radius of real poles can be imposed using Eq. (11) for $V > 1$ and using Eq. (12) with $V = V_{\min}$ for $V < 1$. It is also possible to include first-order HP/LP filters in the grid search by forcing the gain of the shelving filters to zero, effectively using only their notch responses as mentioned in Sec. 3.

An example Bark-exp pole grid is shown in Fig. 7 with $Q^c_{\min} = Q^b_{\min} = 0.75$ and $Q^c_{\max} = Q^b_{\max} = 10$, and with $V_{\max} = \frac{1}{V_{\min}} = 4$, where a_q^b and a_q^c in Eq. (23) and Eq. (24) are evaluated at $\sigma_q = \frac{\pi}{4}$, giving a good balance between narrow and wide band filters. The central frequencies f_0 are distributed between 100 Hz and 21 kHz, with poles having 75 possible angles and 20 possible radii. The cut-off frequencies f_c of the candidate shelving and high/low-pass filters are linearly distributed between 100 Hz and 1 kHz, and between 18 kHz and 21 kHz.

4.5 Line Search

Once the pole $p_s = \rho_s e^{j\sigma_s}$ corresponding to the optimal parametric filter in the grid search is selected, the parameters $d_s^{(0)} = -\cos(\sigma_s)$ and $a_s^{(0)} = \rho_s^2$ are used as the initial conditions of a line search optimization [33], meant to refine their value and reduce the cost function $\mathcal{F}(a_s, d_s, V_s)$ further. In the optimization, $\sigma_s^{(0)}$ is used instead of $d_s^{(0)}$ to take into account its cosinusoidal nature, important in the computation of the search direction. The cost function in Eq. (19), indeed, allows the computation of the gradients w.r.t. the filter parameters, thus enabling the use of gradient-based algorithms, such as steepest descent (SD), quasi-Newton or Gauss-Newton (GN) algorithms, which guarantee fast convergence to a local minimum, provided that the initial values are chosen properly. The assumption that the initial filter parameters obtained with the grid search are sufficiently close to a local minimum is reasonable, as long as the density of the poles in the grid is sufficiently high. The same assumption is required also for the derivative-free algorithms in [13] and in [14], in order to guarantee convergence in a relatively small number of iterations, with the

exception that in those cases the initial filter parameters are obtained by an indirect minimization of the cost function, without verifying whether the initial values provide a good starting point for the equalization.

The parameter vector, initialized as $\boldsymbol{\theta}_s^{(0)} = [a_s^{(0)}, \sigma_s^{(0)}]^T$ for a complex pole (peaking filter), or $\boldsymbol{\theta}_s^{(0)} = a_s^{(0)}$ for a real pole (shelving filter), is updated at each iteration $i = 0, 1, 2, \ldots$ as

$$\boldsymbol{\theta}_s^{(i+1)} = \boldsymbol{\theta}_s^{(i)} + \mu^{(i)} \boldsymbol{p}^{(i)}, \tag{25}$$

where $\mu^{(i)}$ indicates the step size, and $\boldsymbol{p}^{(i)}$ the search direction along which the step is taken in order to reduce the cost function in Eq. (19), such that

$$\mathcal{F}(\boldsymbol{\theta}_s^{(i)} + \mu^{(i)} \boldsymbol{p}^{(i)}, V_s^{(i)}) < \mathcal{F}(\boldsymbol{\theta}_s^{(i)}, V_s^{(i)}), \tag{26}$$

where $V_s^{(0)}$ is the gain estimated in the grid search, which is updated by LS estimation at each evaluation of the cost function. In other words, the search direction $\boldsymbol{p}^{(i)}$ has to be a descent direction, i.e., $\boldsymbol{p}^{(i)T} \nabla \mathcal{F}_s^{(i)} < 0$ with $\nabla \mathcal{F}_s^{(i)} = \nabla \mathcal{F}(\boldsymbol{\theta}_s^{(i)}, V_s^{(i)})$ the gradient of the cost function (i.e., the vector of its first-order partial derivatives) w.r.t. the parameters in $\boldsymbol{\theta}_s^{(i)}$,

$$\nabla \mathcal{F}_s^{(i)} = \frac{\partial \mathcal{F}_s^{(i)}}{\partial \boldsymbol{\theta}_s^{(i)}} = \left[\frac{\partial \mathcal{F}_s^{(i)}}{\partial a_s^{(i)}}, \frac{\partial \mathcal{F}_s^{(i)}}{\partial \sigma_s^{(i)}} \right]^T, \tag{27}$$

with $\{ \cdot \}^T$ indicating the vector transpose. The analytic expressions for the gradient are given in Appendix A.5.

The search direction generally has the form

$$\boldsymbol{p}^{(i)} = -\{\boldsymbol{B}^{(i)}\}^{-1} \nabla \mathcal{F}_s^{(i)}, \tag{28}$$

where $\boldsymbol{B}^{(i)}$ is a symmetric and nonsingular matrix, whose form differentiates the different methods. When $\boldsymbol{B}^{(i)}$ is an identity matrix, $\boldsymbol{p}^{(i)}$ is the SD and Eq. (28) corresponds to the SD method. When $\boldsymbol{B}^{(i)}$ is the exact Hessian $\nabla^2 \mathcal{F}_s^{(i)}$ (i.e., the matrix of second-order partial derivatives) Eq. (28) corresponds to the Newton method. The Hessian can be approximated at each iteration without the need for computing the second-order partial derivatives, leading to quasi-Newton methods, such as the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm. The GN method, instead, computes the search direction by expressing the derivatives of $\mathcal{F}_s^{(i)}$ in terms of the Jacobians $\nabla \boldsymbol{e}_s^{(i)}$, as

$$\begin{aligned}
&\boldsymbol{p}^{(i)} = -\left(\nabla \boldsymbol{e}^{(i)H} \nabla \boldsymbol{e}^{(i)} \right)^{-1} \nabla \boldsymbol{e}^{(i)H} \boldsymbol{e}^{(i)}, \quad \text{with} \\
&\nabla \boldsymbol{e}^{(i)} = \frac{\partial \boldsymbol{e}^{(i)}}{\partial \boldsymbol{\theta}_s^{(i)}} = \left[\frac{\partial \boldsymbol{e}^{(i)}}{\partial a_s^{(i)}}, \frac{\partial \boldsymbol{e}^{(i)}}{\partial \sigma_s^{(i)}} \right]^T, \\
&\boldsymbol{e}^{(i)} = [e(1, \boldsymbol{\theta}_s^{(i)}, V_s^{(i)}), \ldots, e(N, \boldsymbol{\theta}_s^{(i)}, V_s^{(i)})]^T, \\
&e(k, \boldsymbol{\theta}_s^{(i)}, V_s^{(i)}) = W(k) \left\{ \frac{1}{2} H_{s-1}(k) F_{m_s}^{(i)}(k) - T(k) \right\}, \\
&F_{m_s}^{(i)}(k) = F_{m_s}(k, \boldsymbol{\theta}_s^{(i)}, V_s^{(i)})
\end{aligned} \tag{29}$$

with $\{ \cdot \}^H$ indicating Hermitian transpose, where the Jacobians are obtained as an intermediate step in the calculation of the gradients (see Appendix A.5). The GN method approximates the Hessian with $\nabla \boldsymbol{e}^{(i)H} \nabla \boldsymbol{e}^{(i)}$, thus having convergence rate similar to the Newton method, i.e., faster than the SD method.

The convergence rate of line search algorithms also depends on the choice of the step size $\mu^{(i)}$. In order to select a value of $\mu^{(i)}$ that achieves a significant reduction of $\mathcal{F}_s^{(i)}$ without the need to optimize for $\mu^{(i)}$, backtracking with Armijo's sufficient decrease condition [33] is used. The backtracking strategy consists in starting with a large step size $\mu^{(i)} < 1$ ($\mu^{(i)} = 1$ for Newton and quasi-Newton methods) and iteratively reducing it by means of a contraction factor $\kappa \in (0, 1)$, such that $\mu^{(i)} \leftarrow \kappa \mu^{(i)}$. At each repetition of the backtracking, a sufficient decrease condition is evaluated to ensure that the algorithm gives reasonable descent along $\boldsymbol{p}^{(i)}$. The condition in Eq. (26) is, however, not sufficient to ensure convergence to a local minimum. A different condition is then required such as the commonly used Armijo's sufficient decrease condition

$$\mathcal{F}(\boldsymbol{\theta}_s^{(i)} + \mu^{(i)} \boldsymbol{p}^{(i)}, V_s^{(i)}) \leq \gamma \mu^{(i)} \boldsymbol{p}^{(i)T} \nabla \mathcal{F}(\boldsymbol{\theta}_s^{(i)}, V_s^{(i)}) \tag{30}$$

with $\gamma \in (0, 1)$, which states that a decrease in $\mathcal{F}_s^{(i)}$ is sufficient if proportional to both $\mu^{(i)}$ and $\boldsymbol{p}^{(i)T} \nabla \mathcal{F}_s^{(i)}$. A final value for $\mu^{(i)}$ is obtained when Armijo's condition is fulfilled or when it becomes smaller than a predefined value $\mu_{\min}$. Also, the parameters in $\boldsymbol{\theta}_s^{(i)} + \mu^{(i)} \boldsymbol{p}^{(i)}$ should be checked to ensure that $a_s^{(i)}$ and $\sigma_s^{(i)}$ still satisfy the constraints described in the previous section. Stability is guaranteed by $a_{\max} < 1$.

The line search for the current stage terminates when $\boldsymbol{p}^{(i)T} \nabla \mathcal{F}_s^{(i)} \leq \tau$, with τ a specified tolerance or when a maximum number of iterations I is reached. It should be mentioned that it is possible to include a closed-form expression of V in terms of a_s and d_s in the filter transfer function $F_{m_s}(k)$ in Eq. (19), at the expense of more complicated analytic expressions for the gradients. Another alternative is to include the gain V in the vector of parameters $\boldsymbol{\theta}_i$ and perform the line-search without updating the gain parameter between two iterations. However, experimental results showed that the speed of convergence and the final result of these two alternatives are comparable to the results of the line-search algorithm described above.

5 LOUDSPEAKER EQUALIZATION EXAMPLE

In this section an example of parametric equalization of a loudspeaker response is presented. The aim is to show the performance of the proposed procedure described above, in comparison to the state-of-the-art procedures presented in Sec. 1. In an attempt to keep the comparison as fair as possible, the same target response, the same range of equalization 100 Hz–21 kHz, and the same pre-processing (logarithmic frequency axis, Bark-scale smoothing, etc.) is used for the three procedures considered. The target response is built to match the pass-band characteristics of the loudspeaker response, using second-order high-pass and low-pass Butterworth filters with cut-off frequency of 250 Hz and 22 kHz, respectively. The loudspeaker response is scaled so that the 0-dB line of the target response corresponds to the response mean value between 400 Hz and 6 kHz, which satisfies the requirement of the state-of-the-art procedures of having peaks and dips to be equalized (see Fig. 8). The same termination conditions are used for all procedures; the

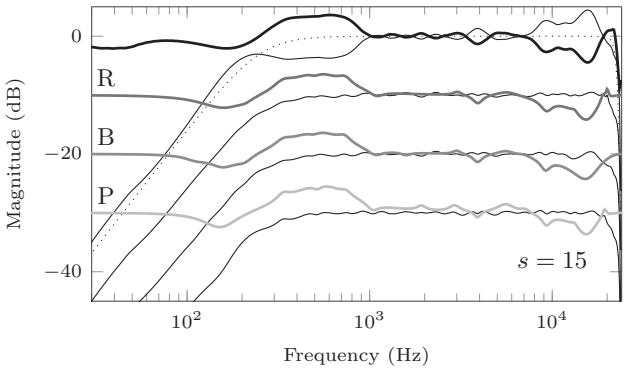

Fig. 8. Loudspeaker equalization. Top: the unequalized response (solid) with the target response (dotted) and the ideal high-order FIR equalizer (thick); From top to bottom (10 dB offset): the equalized response (solid) and the corresponding equalizer (thick) using procedures R, B, and P.

Table 1. Error-based objective measures.

measure	procedure	$s = 0$	$s = 5$	$s = 10$	$s = 15$
n_i	R	0	232	552	824
	B	0	260	593	885
	P	0	29	53	70
SFM_s	R	0.922	**0.991**	**0.996**	**0.998**
	B	0.922	0.986	0.990	0.993
	P	0.922	0.983	0.991	0.995
SDM_s	R	0.435	**0.100**	**0.045**	**0.030**
	B	0.435	0.118	0.078	0.051
	P	0.375	0.101	0.064	0.035

algorithm moves to the next filter section whenever either a maximum number of iterations (e.g., $I = 100$) is reached, or the step size gets smaller than a given value (e.g. $\mu_{\min} = 10^{-4}$), or the reduction in the cost function in a number of previous iterations (e.g., 10) is less than a predefined tolerance value (e.g., $\tau = 10^{-8}$). The Rosenbrock method [19] is applied for both the state-of-the-art procedures, using a step expansion factor $\alpha = 1.5$ and a step contraction factor $\zeta = 0.75$, starting from an initial variation of 0.5% of the value of the initial filter parameters (see [14]). In the procedure by Ramos et al. (**R**) [13], the Q-factor of the filter is initialized based on the bandwidth of the selected error area, while in the one by Behrends et al. (**B**) [14] it is set to $Q_0 = 2$.

The Bark-exp grid used in the proposed procedure (**P**) is the one in Fig. 7. In the example, the GN algorithm is used in the line search, which provides very similar results as SD in a much smaller number of iterations. The initial step size is set to $\mu^{(i)} = 0.9$, the contraction factor for the backtracking to $\kappa = 0.8$, and the Armijo's condition constant to $\gamma = 0.05$. The global gain C is estimated as explained in Sec. 4.

The error produced by the different procedures with increasing number of filter sections s is shown in Fig. 9. As expected, the proposed procedure (P) performs best in minimizing the normalized SSE (NSSE), i.e., the error in Eq. (19) normalized w.r.t. the error in Eq. (4) computed without equalizer ($F_M(k) = 1$) and converted to decibels; the procedure by Ramos et al. (R), with cost function as in Eq. (2), outperforms the other procedures in minimizing the logarithmic error, whereas the procedure by Behrends et al. (B) fails to minimize the linear cost function in Eq. (3) more than the other procedures (at least in this example). Procedure P is the one that, for all cost functions considered, is able to achieve the largest error reduction in the first two stages. Also, the error for procedure P exhibits a staircase-like behavior, which is due to the vicinity of the initial parameter values to a local minimum and the subsequent small improvement given by the line search. In general, the different procedures for an increasing number of stages are not too different from each other in terms of equalization performance, all capable of attaining the target response to a certain degree, as can be seen in Fig. 8 for $s = 15$. A difference is found in the total number of iterations (n_i), with procedure P using the GN algorithm having an order of magnitude less than the other procedures, including the backtracking (see Table 1), due to the efficiency of both the initialization and the GN algorithm. However, the grid search and each iteration of the line search are computationally more demanding than the iterations of the Rosenbrock algorithm, eventually obtaining similar execution times for the different procedures.

Apart from the performance evaluation based on the different cost functions themselves, other measures are considered, namely the spectral flatness measure (SFM) and the spectral distance measure (SDM) described in [4]. The SFM is the ratio between the geometric mean and the

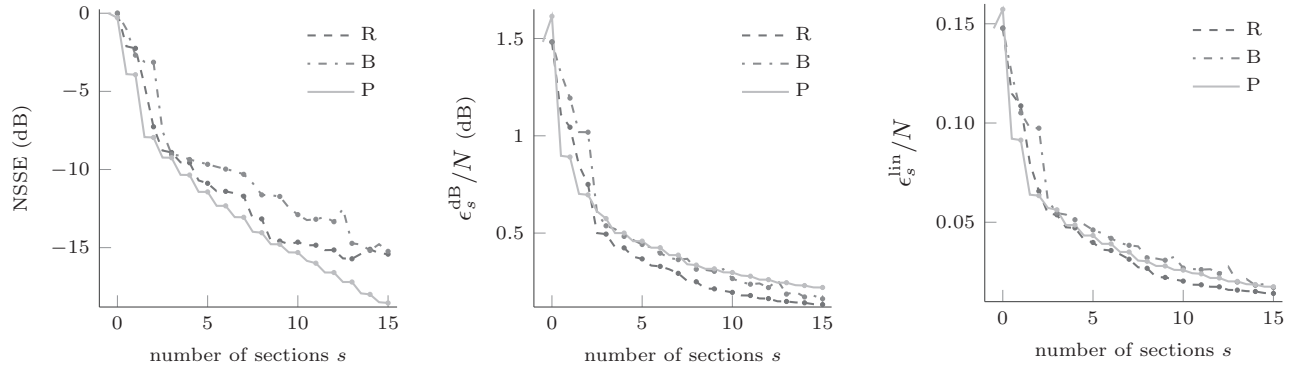

Fig. 9. The error produced by the different procedures at each stage according to the different cost functions.

arithmetic mean of the power spectrum (on a linear frequency scale). The target response not necessarily being flat in the range of equalization, the measure is computed using the power spectrum of the equalized system response divided by the target response $\tilde{H}_s(k) = \frac{H_s(k)}{T(k)}$

$$\mathrm{SFM}_s = N\frac{\sqrt[n]{\prod_k |\tilde{H}_s(k)|^2}}{\sum_k |\tilde{H}_s(k)|^2}, \tag{31}$$

so that the ideal high-order FIR equalizer defined as $D(k) = \frac{T(k)}{H_0(k)}$ has SFM=1. The SDM is also based on the power spectrum of the responses, and it is given by

$$\mathrm{SDM}_s = \sqrt{\sum_k \left| \frac{|\bar{H}_s(k)|^2 - |\bar{T}(k)|^2}{N} \right|^2}, \tag{32}$$

where in this case $\bar{H}_s(k)$ and $\bar{T}(k)$ are the loudspeaker and target responses resampled on a logarithmic frequency scale with $\frac{1}{5}$ octave resolution [4]. Results for these two measures are shown in Table 1 for the different procedures using equalizers with 5, 10, and 15 parametric filters. For both measures, the greatest improvement is achieved with 5 filters only, with smaller improvements for increasing number of filters. It is interesting to notice that, even though not specifically designed to maximize the SFM of the loudspeaker response as for procedures B and R, the proposed procedure (P) achieves a good level of flatness. Regarding the SDM, procedure P achieves a performance close to that of procedure R. Also notice that the optimization of the global gain (at $s = 0$) already contributes to a reduction of the SDM.

From a subjective point of view, it is commonly accepted that a flat (in the pass-band) frequency response is perceived as more natural, and that deviations from this are perceived as spectral coloration. Perceptual objective measures based on these assumptions are used here for speech and music stimuli. These are the average log-spectral difference measure (LSDM) [16] and a measure based on a linear distortion auditory model [17], referred to here as perceptual linear distortion measure (PLDM).

The LSDM is the square difference between the logarithm of the magnitude responses of a clean speech segment convolved with the target response $S_T(k)$ (reference) and the same speech segment convolved with the equalized loudspeaker response $S_{H_s}(k)$,

$$\mathrm{LSDM}_s = \sqrt{\frac{1}{N}\sum_k \left[\log(S_{H_s}(k)) - \log(S_T(k))\right]^2} \tag{33}$$

The average LSDM is then computed for those segments (in this case of 25 ms with 15 ms overlap) where active speech is detected. The PLDM is a measure of the perceived subjective naturalness of speech or music w.r.t. linear distortions, represented by spectral ripples and tilts in the magnitude response (see [17] for detailed information).

Results obtained using a male voice speech signal [34] and a music signal consisting of the instrumental introduction of a rock song [35] (having a wide-band spectrum) are shown in Table 2. For both measures considered, a strong improvement for the low-order equalizers ($s = 5$) is shown, with procedure R slightly better than procedure P, except for the LSDM using the music signal. For the LSDM, the use of higher-order equalizers improves the scores only slightly, whereas a more consistent improvement is still visible for the PLDM with $s = 10$. Notice, however, that small differences in the score values will most likely not be perceived as a difference in sound quality.

Table 2. Perception-based objective measures.

measure	procedure	$s = 0$	$s = 5$	$s = 10$	$s = 15$
$\mathrm{LSDM}_{\mathrm{speech}}$	R	0.350	**0.144**	**0.118**	**0.099**
	B	0.350	0.158	0.143	0.135
	P	0.350	0.160	0.142	0.118
$\mathrm{LSDM}_{\mathrm{music}}$	R	0.308	0.165	0.143	0.127
	B	0.308	0.171	0.163	0.159
	P	0.296	**0.146**	**0.136**	**0.110**
$\mathrm{PLDM}_{\mathrm{speech}}$	R	0.785	**0.349**	**0.222**	0.178
	B	0.785	0.366	0.258	**0.174**
	P	0.785	0.399	0.259	0.217
$\mathrm{PLDM}_{\mathrm{music}}$	R	0.960	**0.380**	**0.267**	**0.224**
	B	0.960	0.414	0.314	0.257
	P	0.960	0.405	0.278	0.238

6 ROOM EQUALIZATION EXAMPLE

The proposed procedure can be applied to the equalization of the combined loudspeaker/room response without major modifications. Differently from loudspeaker equalization, the purpose of room transfer function (RTF) equalization is not only to obtain a more balanced response w.r.t. a target response but also to compensate (as much as possible) for strong resonances at LFs. Thus, the smoothing should be less prominent with fractional-octave smoothing (e.g., $\frac{1}{6^{\mathrm{th}}}$) preferred over Bark-scale smoothing, which has constant resolution below 500 Hz. Based on the amount of smoothing, which determines the level of detail in the spectral envelope of the response, the number of parametric filters required to attain the target response with a certain accuracy may vary.

The definition of the target response is a critical issue. RTFs have a more irregular frequency structure than loudspeaker responses, which cannot be easily recognized as deviations from a flat response. Moreover, spectral complexity combined with a less aggressive smoothing result in less smooth error surfaces produced by the cost functions, presenting a large number of local minima. In order to obtain a target response that produces peaks and dips in the RTF to be equalized, as required by the state-of-the-art procedures, one could use a strongly smoothed (e.g., one-octave resolution) version of the response as the target, which however may not provide a desired response. These procedures can be still used in general but are more likely to incur into problems.

The proposed procedure is instead less sensitive to the selection of the target response and will always start optimizing a new filter section from an initial point reasonably close to a useful local minimum, provided that the density

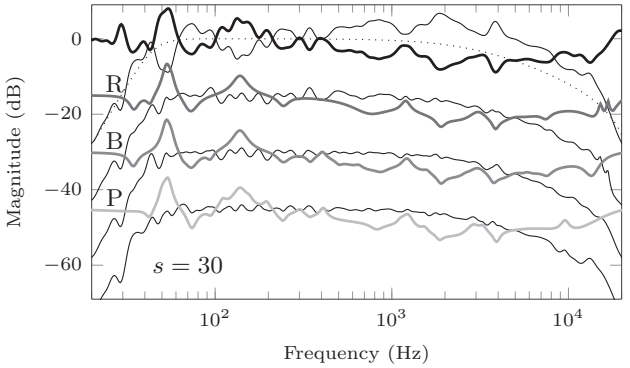

Fig. 10. Room equalization. Top: the unequalized response (solid) with the target response (dotted) and the ideal high-order FIR equalizer (thick); From top to bottom (15 dB offset): the equalized response (solid) and the corresponding equalizer (thick) using procedures R, B, and P.

of the poles in the grid is proportional to the amount of smoothing applied. In this example, the number of possible angles was increased to 300. This, however, makes the grid search at each stage more computationally demanding. If efficiency is an issue, an option is to start from a dense grid and effectively use only a subset of the grid points, different for every filter section (e.g., by taking one every $n = 4$ angles and shift over one angle for the following $n - 1$ filter sections). Simulation results show that this solution leads to very similar results to those obtained with the solution based on the full grid.

In the example shown here, the magnitude of an RTF measured in the parliament hall of the Provinciehuis Oost-Vlaanderen in Ghent, Belgium, having a reverberation time of 1.5 s, has been smoothed with $\frac{1}{6}$ octave resolution and the equalization is evaluated in the range 30 Hz–18 kHz. The target response used is a combination of a forth-order high-pass Butterworth filter with cut-off frequency of 45 Hz and a first-order low-pass Butterworth filter with cut-off frequency of 3 kHz, which produces a slight roll-off at higher frequencies (see Fig. 10). Table 3 provides results of the different error functions and other error measures produced by the different procedures. Given the more complicated error function surfaces, the GN algorithm in the proposed procedure (P) now needs more iterations than in the previous example but still two or three times fewer than the other two procedures. As in the loudspeaker example, all procedures are able to achieve equalization with a comparable accuracy (see also Fig. 10 for $s = 30$), with the strongest improvement obtained in the first 10 stages. The proposed procedure achieves better results in terms of NSSE and of SDM and slightly worse performances in the error measures related to the flatness of the equalized response. Also in this example, procedure B is slightly outperformed by the other procedures in its own cost function but has better performance in terms of the SFM. It should also be noted that procedure R does not achieve any improvement with the inclusion of the last 10 filter sections ($s = 30$). In this particular case, the Rosenbrock method gets stuck into a local minimum and is unable to correct for the largest error area, which is then selected at each new filter initialization thus failing to produce any further performance improvement.

Table 3. Error-based objective measures (RTF).

measure	procedure	$s = 0$	$s = 10$	$s = 20$	$s = 30$
n_i	R	0	642	1150	1584
	B	0	668	1353	2026
	P	0	231	595	792
NSSE$_s$	R	0	**–12.6**	–14.6	–14.6
	B	0	–9.4	–13.9	–16.5
	P	0	–11.9	**–15.7**	**–18.3**
$\epsilon_s^{\mathrm{dB}}/N$ Eq. (2)	R	3.91	**0.80**	**0.54**	0.54
	B	3.91	0.96	0.57	**0.42**
	P	3.91	0.94	0.65	0.51
$\epsilon_s^{\mathrm{lin}}/N$ Eq. (3)	R	0.402	**0.077**	0.051	0.051
	B	0.402	0.082	0.052	**0.037**
	P	0.402	**0.077**	**0.048**	**0.037**
SFM$_s$ Eq. (31)	R	0.915	0.966	0.979	0.978
	B	0.915	**0.986**	**0.992**	**0.995**
	P	0.915	0.960	0.966	0.979
SDM$_s$ Eq. (32)	R	1.146	0.176	0.106	0.104
	B	1.146	0.186	0.106	0.066
	P	1.146	**0.146**	**0.078**	**0.049**

7 NOTE ON MULTI-POINT EQUALIZATION AND TRANSFER FUNCTION MODELING

When the aim is to improve the response of a loudspeaker at multiple listening angles or the room response at multiple positions, a single equalizer can be designed based on a prototype response that contains the common acoustic features of the multiple responses. Averaging and smoothing the magnitude responses was proven to offer an effective solution [36], which also increases robustness to spatial variations. The proposed procedure can then be extended to multi-point equalization by including an averaging operation in the prepocessing step.

A very similar procedure to the one described for automatic design of a parametric equalizer can be applied to the problem of transfer function modeling. This idea can be useful, for instance, to model the ideal FIR equalizer $D(k) = \frac{T(k)}{H_0(k)}$ using a low order parametric filter, the response of a graphic equalizer [21], or more generally any minimum-phase transfer function. For the modeling problem, the cost function becomes

$$\epsilon_M^{\mathrm{SSE}} = \frac{1}{N}\sum_k \big(W(k)\big[D(k) - F_M(k)\big]\big)^2. \tag{34}$$

Also in this case, LIG parametric filters can be used with the possibility of computing the gradients w.r.t. the other filter parameters.

8 CONCLUSION

An automatic procedure for the design of a low-order parametric equalizer has been proposed, which uses a series of second-order peaking filters and first-order shelving filters. The proposed procedure minimizes the SSE in Eq. (4) between the complex-valued minimum-phase

system and target responses, instead of the commonly used difference in the magnitude responses, bringing some advantages, such as an improved mathematical tractability of the equalization problem, with the possibility of computing analytical expressions for the gradients w.r.t. the filter parameters and a closed-form solution for the estimation of the gain parameters, an improved parameter initialization, the inclusion of shelving filters in the optimization procedure, and a more accentuated focus on the equalization of frequency peaks over dips. Examples of loudspeaker and room response equalization have shown that effective equalization using a small number of parametric filters can be achieved. The proposed procedure can be extended to multi-point equalization, by means of a prototype average response, and to transfer function modeling.

9 ACKNOWLEDGMENT

This research work was carried out at the ESAT Laboratory of KU Leuven, in the frame of (i) the FP7-PEOPLE Marie Curie Initial Training Network "Dereverberation and Reverberation of Audio, Music, and Speech (DREAMS)," funded by the European Commission under Grant Agreement no. 316969, (ii) KU Leuven Research Council CoE PFV/10/002 (OPTEC), (iii) KU Leuven Impulsfonds IMP/14/037, (iv) KU Leuven Internal Funds C2-16-00449 and VES/16/03, and (v) was supported by a Postdoctoral Fellowship (F+/14/045) of the KU Leuven Research Fund. The authors would like to thank Televic N.V. for the use of their equipment and their premises, especially Vincent Soubry and Frederik Naessens for the useful collaboration in the early stage of this work. The authors would also like to thank Martin Møller (Bang & Olufsen) for providing the loudspeaker responses and Rainer Huber (Uni Oldenburg) for the fruitful discussions and the code for the PLDM. The scientific responsibility of this work is assumed by its authors.

10 REFERENCES

[1] V. Välimäki and J. D. Reiss, "All about Audio Equalization: Solutions and Frontiers," *Appl. Sciences*, vol. 6, no. 5, p. 129 (2016). https://doi.org/10.3390/app6050129

[2] L. D. Fielder, "Analysis of Traditional and Reverberation-Reducing Methods of Room Equalization," *J. Audio Eng. Soc.*, vol. 51, pp. 3–26 (2003 Jan./Feb.).

[3] M. Karjalainen, T. Paatero, J. N. Mourjopoulos, and P. D. Hatziantoniou, "About Room Response Equalization and Dereverberation," *Proceedings 2005 IEEE Workshop Appls. Signal Process. Audio Acoust. (WASPAA '05)* (2005), pp. 183–186. https://doi.org/10.1109/aspaa.2005.1540200

[4] M. Karjalainen, E. Piirilä, A. Järvinen, and J. Huopaniemi, "Comparison of Loudspeaker Equalization Methods Based on DSP Techniques," *J. Audio Eng. Soc.*, vol. 47, pp. 14–31 (1999 Jan./Feb.).

[5] A. Gabrielsson, B. Lindström, and O. Till, "Loudspeaker Frequency Response and Perceived Sound Quality," *J. Acoust. Soc. Am.*, vol. 90, no. 2, pp. 707–719 (1991). https://doi.org/10.1121/1.401941

[6] B. C. Moore and C.-T. Tan, "Perceived Naturalness of Spectrally Distorted Speech and Music," *J. Acoust. Soc. Am.*, vol. 114, no. 1, pp. 408–419 (2003). https://doi.org/10.1121/1.1577552

[7] L. G. Johansen and P. Rubak, "The Excess Phase in Loudspeaker/Room Transfer Functions: Can it Be Ignored in Equalization Tasks?" presented at the *100th Convention of the Audio Engineering Society* (1996 May), convention paper 4181.

[8] R. Plomp and H. Steeneken, "Effect of Phase on the Timbre of Complex Tones," *J. Acoust. Soc. Am.*, vol. 46, no. 2B, pp. 409–421 (1969). https://doi.org/10.1121/1.2002342

[9] R. Bristow-Johnson, "The Equivalence of Various Methods of Computing Biquad Coefficients for Audio Parametric Equalizers," presented at the *97th Convention of the Audio Engineering Society* (1994 Nov.), convention paper 3906.

[10] P. Regalia and S. Mitra, "Tunable Digital Frequency Response Equalization Filters," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 35, no. 1, pp. 118–120 (1987). https://doi.org/10.1109/tassp.1987.1165037

[11] U. Zölzer and T. Boltze, "Parametric Digital Filter Structures," presented at the *99th Convention of the Audio Engineering Society* (1995 Oct.), convention paper 4099.

[12] U. Zölzer, *Digital Audio Signal Processing* (John Wiley & Sons, 2008). http://doi.org/10.1002/9780470680018

[13] G. Ramos and J. J. Lopez, "Filter Design Method for Loudspeaker Equalization Based on IIR Parametric Filters," *J. Audio Eng. Soc.*, vol. 54, pp. 1162–1178 (2006 Dec.).

[14] H. Behrends, A. von dem Knesebeck, W. Bradinal, P. Neumann, and U. Zolzer, "Automatic Equalization Using Parametric IIR Filters," *J. Audio Eng. Soc.*, vol. 59, pp. 102–109 (2011 Mar.).

[15] F. E. Toole and S. E. Olive, "The Modification of Timbre by Resonances: Perception and Measurement," *J. Audio Eng. Soc.*, vol. 36, no. 3, pp. 122–142 (1988 Mar.).

[16] A. Gray and J. Markel, "Distance Measures for Speech Processing," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 24, no. 5, pp. 380–391 (1976). http://doi.org/10.1109/TASSP.1976.1162849

[17] B. C. Moore and C.-T. Tan, "Development and Validation of a Method for Predicting the Perceived Naturalness of Sounds Subjected to Spectral Distortion," *J. Audio Eng. Soc.*, vol. 52, no. 9, pp. 900–914 (2004 Sep.).

[18] G. Ramos and P. Tomas, "Improvements on Automatic Parametric Equalization and Cross-Over Alignment of Audio Systems," presented at the *126th Convention of the Audio Engineering Society* (2009 May), convention paper 7759.

[19] H. Rosenbrock, "An Automatic Method for Finding the Greatest or Least Value of a Function," *Comput. J.*, vol. 3, no. 3, pp. 175–184 (1960). https://doi.org/10.1093/comjnl/3.3.175

[20] T. Corbach, A. von dem Knesebeck, K. Dempwolf, M. Holters, P. Sorowka, and U. Zölzer, "Automated Equalization for Room Resonance Suppression," *Proceedings 12th Int. Conf. Digital Audio Effects (DAFx09)* (2009).

[21] J. S. Abel and D. P. Berners, "Filter Design Using Second-Order Peaking and Shelving Sections," *Proceedings 30th Int. Comput. Music Conf. (ICMC)*, Miami, USA (2004).

[22] Z. Chen, Y. Liu, G. Geng, and F. Yin, "Optimal Design of Digital Audio Parametric Equalizer," *J. Inform. Comput. Sci.*, vol. 11, no. 1, pp. 57–66 (2014). https://doi.org/10.12733/jics20102563

[23] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing (2nd Ed.)* (Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1999).

[24] M. Hayes, J. Lim, and A. Oppenheim, "Signal Reconstruction from Phase or Magnitude," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 28, no. 6, pp. 672–680 (1980). https://doi.org/10.21236/ada078745

[25] J. O. Smith, *Spectral Pre-Processing for Audio Digital Filter Design* (Ann Arbor, MI: Michigan Publishing, University of Michigan Library, 1983).

[26] J. O. Smith, *Introduction to Digital Filters with Audio Applications,* online book, 2007 edition, available: http://ccrma.stanford.edu/∼jos/filters/, accessed Oct. 2016.

[27] P. A. Regalia, S. K. Mitra, and P. Vaidyanathan, "The Digital All-Pass Filter: A Versatile Signal Processing Building Block," *Proceedings IEEE*, vol. 76, no. 1, pp. 19–37 (1988). https://doi.org/10.1109/5.3286

[28] J.-M. Jot, "Proportional Parametric Equalizers-Application to Digital Reverberation and Environmental Audio Processing," presented at the *139th Convention of the Audio Engineering Society* (2015 Oct.), convention paper 9358.

[29] F. Toole, "The Measurement and Calibration of Sound Reproducing Systems," *J. Audio Eng. Soc.*, vol. 63, pp. 512–541 (2015 Jul./Aug.). https://doi.org/10.17743/jaes.2015.0064

[30] J. Abildgaard Pedersen and K. Thomsen, "Fully Automatic Loudspeaker-Room Adaptation—The Room-Perfect System," presented at the *AES 32nd International Conference: DSP for Loudspeakers* (2007 Sep.), conference paper 6.

[31] G. Vairetti, E. De Sena, M. Catrysse, S. H. Jensen, M. Moonen, and T. van Waterschoot, "A Scalable Algorithm for Physically Motivated and Sparse Approximation of Room Impulse Responses with Orthonormal Basis Functions," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 25, no. 7, pp. 1547–1561 (2017). https://doi.org/10.1109/taslp.2017.2700940

[32] M. Karjalainen and T. Paatero, "Equalization of Loudspeaker and Room Responses Using Kautz Filters: Direct Least Squares Design," *EURASIP J. Adv. Signal Process.*, vol. 2007, no. 1, pp. 185–185 (2007). https://doi.org/10.1155/2007/60949

[33] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. (Springer, 2006), ch. 3, "Line Search Methods," pp. 30–65. https://doi.org/10.1007/978-0-387-40065-5

[34] G. Waters, "Sound Quality Assessment Material—Recordings for Subjective Tests: Users Handbook for the EBU–SQAM Compact Disk," European Broadcasting Union (EBU), Tech. Rep. (1988).

[35] Radiohead, "Burn the Witch," in *A Moon Shaped Pool*. XL Records, 2016, (12 s–20 s).

[36] S. Cecchi, L. Palestini, P. Peretti, L. Romoli, F. Piazza, and A. Carini, "Evaluation of a Multipoint Equalization System Based on Impulse Response Prototype Extraction," *J. Audio Eng. Soc.*, vol. 59, pp. 110–123, (2011 Mar.).

APPENDIX A

A.1 SSE Minimum-Phase Cost Function

Here the SSE cost function in Eq. (4) for the minimum-phase equalization problem is analyzed using the relation by which the frequency response of a minimum-phase transfer function $H(k)$ can be written as

$$H(k) = |H(k)|e^{-j\mathcal{H}\{\ln|H(k)|\}} \tag{A.1}$$

For simplicity, the weighting matrix in Eq. (4) is set to $W(k) = 1$ and the notation is simplified. The cost function in Eq. (4) can be elaborated in terms of magnitude and phase of the frequency responses involved, as shown in Eq. (A.2), where the Euler's rule and the linear property of the Hilbert transform have been used ($\{\cdot\}^*$ indicates complex conjugation). It can be noticed that the optimal equalizer, for which $E^2(k) = 0$, is defined as $F(k) = \frac{T(k)}{H(k)} = \frac{|T(k)|}{|H(k)|}e^{j(\phi_T(k)-\phi_H(k))}$.

This cost function has a quadratic form, which assumes large values whenever the power of the equalized magnitude response is significantly larger than the power of the target response, and whenever the difference between their magnitude responses on a natural logarithmic scale is large (i.e., when the value of cos is far from one).

$$\begin{aligned}
E^2(k) &= (H(k)F(k) - T(k))^*(H(k)F(k) - T(k)) \\
&= (F^*(k)H^*(k) - T^*(k))(H(k)F(k) - T(k)) \\
&= |F(k)H(k)|^2 + |T(k)|^2 - |F(k)H(k)T(k)| \\
&\quad \times \left(e^{-j(\phi_F(k)+\phi_H(k)-\phi_T(k))} - e^{j(\phi_F(k)+\phi_H(k)-\phi_T(k))}\right) \\
&= |F(k)H(k)|^2 + |T(k)|^2 \\
&\quad - 2|F(k)H(k)T(k)|\cos(\phi_F(k) \\
&\quad + \phi_H(k) - \phi_T(k)) \\
&= |F(k)H(k)|^2 + |T(k)|^2 \\
&\quad - 2|F(k)H(k)T(k)|\cos(-j\mathcal{H}\{\ln|F(k)|\} \\
&\quad - j\mathcal{H}\{\ln|H(k)| + j\mathcal{H}\{\ln|T(k)|\}) \\
&= |F(k)H(k)|^2 + |T(k)|^2 \\
&\quad - 2|F(k)H(k)T(k)|\cos(-j\mathcal{H}\{\ln|F(k)| \\
&\quad + \ln|H(k)| - \ln|T(k)|\}) \\
&= |\tilde{H}(k)|^2 + |T(k)|^2 \\
&\quad - 2|\tilde{H}(k)T(k)|\cos(-j\mathcal{H}\{\ln|\tilde{H}(k)| \\
&\quad - \ln|T(k)|\})
\end{aligned} \tag{A.2}$$

To simplify the analysis even further, a zero-phase, flat target response ($T(k) = 1$) is considered, for which the cost

function assumes the form

$$E^2(k) = |\tilde{H}(k)|^2 + 1 - 2|\tilde{H}(k)|\cos(-j\mathcal{H}\{\ln|\tilde{H}(k)|\}) \tag{A.3}$$

In this case, it is easy to see that the error is larger when the equalized magnitude response of $\tilde{H}(k)$ has values larger than one, with the error increasing more than linearly for increasing magnitude, which explains the focus on the equalization of strong peaks.

A.2 The Orthogonality Property of the Regalia-Mitra Parametric Filters

A brief explanation of the property introduced in Sec. 3 is provided here. Define a vector $\boldsymbol{x}$ containing N samples of the input signal $x(n)$ and the vector $\boldsymbol{z}$ containing N samples of the output signal $z(n)$ of the all-pass filter $A_m(z)$. A known property of an all-pass filter is the preservation of the energy, such that the energy of the input signal is equal to the energy of the output signal

$$\sum_{n=-\infty}^{\infty}|z(n)|^2 = \sum_{n=-\infty}^{\infty}|x(n)|^2 \tag{A.4}$$

or, in terms of vector inner products, $\langle \boldsymbol{z}, \boldsymbol{z}\rangle = \langle \boldsymbol{x}, \boldsymbol{x}\rangle$. With reference to Fig. 3b, the output of the filter $y_m(n)$ is formed from the weighted summation of two signals, $y^{\eta}(n) = x(n) + z(n)$ and $y^{\beta}(n) = x(n) - z(n)$. The orthogonality of these two signals can be assessed from their inner product,

$$\langle \boldsymbol{y}^n, \boldsymbol{y}^r\rangle = \langle \boldsymbol{x}+\boldsymbol{z}, \boldsymbol{x}-\boldsymbol{z}\rangle \tag{A.5}$$

$$= \langle \boldsymbol{x}, \boldsymbol{x}\rangle - \langle \boldsymbol{x}, \boldsymbol{z}\rangle + \langle \boldsymbol{z}, \boldsymbol{x}\rangle - \langle \boldsymbol{z}, \boldsymbol{z}\rangle = 0, \tag{A.6}$$

which follows from the equality stated above, $\langle \boldsymbol{z}, \boldsymbol{z}\rangle = \langle \boldsymbol{x}, \boldsymbol{x}\rangle$, and from $\langle \boldsymbol{x}, \boldsymbol{z}\rangle = \langle \boldsymbol{z}, \boldsymbol{x}\rangle$.

A.3 Gain LS Estimation

The first-order partial derivative of the cost function in Eq. (19) w.r.t. the gain parameter V is given by Eq. (A.7).

$$\begin{aligned}
\frac{\partial \epsilon_m^{\mathrm{SSE}}}{\partial V} &= \frac{1}{N}\sum_k \Big(\frac{1}{2}W(k)H_s(k)F_m^{\beta}(k)\Big)^* \\
&\quad \times \Big(W(k)\big[H_s(k)\cdot F_m(k) - T(k)\big]\Big) \\
&= \frac{1}{N}\sum_k \Big(\frac{1}{2}F_m^{\beta *}(k)H_s^*(k)W^*(k)\Big)\Big(\frac{1}{2}W(k)H_s(k)F_m^{\eta}\Big) \\
&\quad + \frac{1}{N}\sum_k \Big(\frac{1}{2}F_m^{\beta *}(k)H_s^*(k)W^*(k)\Big)\Big(\frac{V}{2}W(k)H_s(k)F_m^{\beta}\Big) \\
&\quad - \frac{1}{N}\sum_k \Big(\frac{1}{2}F_m^{\beta *}(k)H_s^*(k)W^*(k)\Big)\Big(W(k)T(k)\Big).
\end{aligned} \tag{A.7}$$

Since orthogonality in the time domain (see Appendix A.3) implies orthogonality also in the frequency domain, the first summation in the equation equals zero ($F_m^{\eta}(k)$ and $F_m^{\beta}(k)$ are orthogonal). By setting $\frac{\partial \epsilon_m^{\mathrm{SSE}}}{\partial V} = 0$, the following equation is obtained

$$\begin{aligned}
&V\sum_k \Big(\frac{1}{2}F_m^{\beta *}(k)H_s^*(k)W^*(k)\Big)\Big(\frac{1}{2}W(k)H_s(k)F_m^{\beta}\Big) \\
&= \sum_k \Big(\frac{1}{2}F_m^{\beta *}(k)H_s^*(k)W^*(k)\Big)\Big(W(k)T(k)\Big).
\end{aligned} \tag{A.8}$$

which provides an estimate for the gain as

$$\hat{V} = \frac{\sum_k \big[F_m^{\beta *}(k)H_s^*(k)W^*(k)\big]\big[W(k)T(k)\big]}{\sum_k \big[F_m^{\beta *}(k)H_s^*(k)W^*(k)\big]\big[W(k)H_s(k)F_m^{\beta}(k)\big]} \tag{A.9}$$

which is equivalent to the expression in Eq. (17).

A.4 Gradients and Jacobians Expressions

Based on the method chosen to perform the line search at each stage, the search direction $\boldsymbol{p}_i$ requires the computation of the gradients $\nabla\mathcal{F}_s^{(i)} = \frac{\partial \mathcal{F}_s^{(i)}}{\partial \boldsymbol{\theta}_s^{(i)}}$ (in the SD and quasi-Newton methods) or of the Jacobians $\nabla \boldsymbol{e}_s^{(i)} = \frac{\partial \boldsymbol{e}_s^{(i)}}{\partial \boldsymbol{\theta}_s^{(i)}}$ (in the GN method), where[1]

$$\mathcal{F}_s^{(i)} = \frac{1}{N}\sum_k e(k, \boldsymbol{\theta}_s^{(i)}, V_s^{(i)})^2 = \frac{1}{N}\boldsymbol{e}_s^{(i)H}\boldsymbol{e}_s^{(i)} \tag{A.10}$$

The gradient of the cost function can be written as

$$\nabla\mathcal{F}_s^{(i)} = \frac{2}{N}\sum_k \frac{\partial e_s^{(i)}(k)}{\partial \boldsymbol{\theta}_s^{(i)}} e_s^{(i)}(k) = \frac{2}{N}\nabla \boldsymbol{e}_s^{(i)H}\boldsymbol{e}_s^{(i)}, \tag{A.11}$$

where the Jacobian is given by, for $k = 1, \ldots, N$,

$$\frac{\partial e_s^{(i)}(k)}{\partial \boldsymbol{\theta}_s^{(i)}} = \frac{1}{2}W(k)H_{s-1}(k)\frac{\partial F_{m_s}(k, \boldsymbol{\theta}_s^{(i)}, V_s^{(i)})}{\partial \boldsymbol{\theta}_s^{(i)}}, \tag{A.12}$$

so that the partial derivatives $\frac{\partial F_{m_s}^{(i)}(k)}{\partial \boldsymbol{\theta}_s^{(i)}}$ for peaking and shelving filters are required. In order to use the Newton method, the exact Hessian $\nabla^2\mathcal{F}_s^{(i)} = \frac{\partial^2 \mathcal{F}_s^{(i)}}{\partial \boldsymbol{\theta}_s^{(i)2}}$ should be computed. Analytic expressions for the second-order partial derivatives can be obtained, but the advantages of using the Newton method are outweighed by a higher complexity.

Peaking Filters

The frequency response of peaking filters in the LIG form can be written, substituting Eq. (13) in Eq. (9), as

$$F_2(k) = \frac{(1+a)(1+2dk+k^2) + V(1-a)(1-k^2)}{2(1+d(1+a)k+ak^2)}, \tag{A.13}$$

and its first-order partial derivatives w.r.t. the parameters a and $\sigma = \cos^{-1}(-d)$ as

$$\frac{\partial F_2(k)}{\partial a} = \frac{(1-V)(1-k^2)(1+2dk+k^2)}{2(1+d(1+a)k+ak^2)^2} \tag{A.14}$$

[1]For convenience here k stands for $e^{-j\frac{\omega_k}{f_s}}$, so that the transfer functions are evaluated at $z = e^{j\frac{\omega_k}{f_s}}$, and z^{-1} can be substituted by k. $\{\cdot\}^H$ indicates Hermitian transpose.

$$\frac{\partial F_2(k)}{\partial \sigma} = \frac{\sin(\sigma)(1-V)(1-k^2)(1-a^2)k}{2(1+d(1+a)k+ak^2)^2}. \quad \text{(A.15)}$$

Shelving and HP/LP Filters

The frequency response of low frequency (LF) and high frequency (HF) shelving filters in the LIG form can be written, substituting Eq. (10) in Eq. (9), respectively as

$$F_1^{\mathrm{LF}}(k) = \frac{(1+a)(1-k)+V(1-a)(1+k)}{2(1-ak)} \quad \text{(A.16)}$$

$$F_1^{\mathrm{HF}}(k) = \frac{(1+a)(1+k)+V(1-a)(1-k)}{2(1+ak)}, \quad \text{(A.17)}$$

and its partial derivatives w.r.t. the parameter a as

$$\frac{\partial F_1^{\mathrm{LF}}(k)}{\partial a} = \frac{(1-V)(1-k^2)}{2(1-ak)^2} \quad \text{(A.18)}$$

$$\frac{\partial F_1^{\mathrm{HF}}(k)}{\partial a} = \frac{(1-V)(1-k^2)}{2(1+ak)^2}. \quad \text{(A.19)}$$

The frequency response for HP and LP filters is obtained by setting $V = 0$ in Eqs. (A.16) and (A.17), respectively, and their partial derivatives w.r.t. a by setting $V = 0$ in Eqs. (A.18) and (A.19).

THE AUTHORS


Giacomo Vairetti


Enzo De Sena


Michael Catrysse


Søren Holdt Jensen


Marc Moonen


Toon van Waterschoot

Giacomo Vairetti received the B.Sc. in 2010 and the M.Sc. (cum laude) in 2012, both in computer engineering at Politecnico di Milano (Italy). In 2018, he received the Ph.D. in electrical engineering at KU Leuven (Belgium), where he was a Marie Curie Fellow. He was a visiting student at the Signal Processing and Acoustics Dept. of Aalto University (Finland) in 2012 and at the Electronic Systems Dept. of Aalborg University (Denmark) in 2014. His research interests are in signal processing and system identification applied to room acoustic modeling, sound synthesis, and audio reproduction.

•

Enzo De Sena received the B.Sc. in 2007 and M.Sc. (cum laude) in 2009, both from the Universita degli Studi di Napoli "Federico II" (Italy) in telecommunication engineering. In 2013, he received the Ph.D. degree in electronic engineering from Kings College London (UK), where he was also a Teaching Fellow from 2012 to 2013. Between 2013 and 2016 he was a Postdoctoral Research Fellow at the Katholieke Universiteit Leuven (Belgium). Since September 2016 he is a Lecturer in audio at the Institute of Sound Recording at the University of Surrey (UK). He held visiting positions at Stanford University (USA), Aalborg University (Denmark), and Imperial College London (UK). He is a former Marie Curie Fellow. His current research interests include room acoustics modeling, surround sound, microphone beam forming, and binaural modeling. For more information see www.desena.org.

•

Michael Catrysse received the M.Sc. and Ph.D. degrees in electronic engineering from Katholieke Universiteit Leuven, Leuven, Belgium, in 1998 and 2004, respectively. He joined Televic, Izegem, Belgium, in 2013 as the Director of Technology and Innovation. Before joining Televic, he was an R&D Manager and Plant Manager with Picanol Group, Ieper, Belgium. Prior to this he held positions as a Researcher (FMTC, Belgium), an Innovation Consultant, a Project Manager (Centexbel, Belgium), and Innovation Manager (Domo). He has been involved in several nationally and European-funded research projects, both as a Participant and a Project Coordinator. He has (co)authored more than 30 publications. Since 2017 he is the owner of CoEnCo Micro Breweries, Oostkamp, Belgium.

•

Søren Holdt Jensen received the M.Sc. degree in electrical engineering from Aalborg University (AAU), Denmark, in 1988, and the Ph.D. degree (in signal processing) from the Technical University of Denmark (DTU) in 1995. He is Full Professor in signal processing at Aalborg University. Before joining the Electronic Systems Dept. (AAU), he was with the Telecommunications Laboratory of Telecom Denmark, Ltd., Copenhagen; the Electronics Institute of Technical University of Denmark; the Scientific Computing Group of Danish Computing Center for Research and Education (UNI C), Lyngby; the Electrical Engineering Dept. of KU Leuven, Belgium; and the Center for PersonKommunikation (CPK) of AAU. His current research interest are in statistical signal processing, numerical algorithms, optimization engineering, machine learning, and digital processing of acoustic, audio, communication, image, multimedia, speech, and video signals. He is co-author of the textbook *Software-Defined GPS and Galileo Receiver—A Single-Frequency Approach*, Birkhäuser, Boston, USA, also translated to Chinese: National Defence Industry Press, China. Prof. Jensen has been Associate Editor for the *IEEE Transactions on Signal Processing, IEEE/ACM Transactions on Audio, Speech and Language Processing, Elsevier Signal Processing,* and *EURASIP Journal on Advances in Signal Processing*. He is a recipient of an individual European Community Marie Curie (HCM: Human Capital and Mobility) Fellowship, former Chairman of the IEEE Denmark Section and the IEEE Denmark Sections Signal Processing Chapter (founder and first chaiman). He is member of the Danish Academy of Technical Sciences (ATV) and has been member of the Danish Council for Independent Research (2011–2016) appointed by Danish Ministers of Science.

•

Marc Moonen is a Full Professor at the Electrical Engineering Dept. of KU Leuven, where he is heading a research team working in the area of numerical algorithms and signal processing for digital communications, wireless communications, DSL, and audio signal processing. He received the 1994 KU Leuven Research Council Award, the 1997 Alcatel Bell (Belgium) Award (with Piet Vandaele), the 2004 Alcatel Bell (Belgium) Award (with Raphael Cendrillon), and was a 1997 Laureate of the Belgium Royal Academy of Science. He received journal best paper awards from the IEEE Transactions on Signal Processing (with Geert Leus and with Daniele Giacobello) and from Elsevier Signal Processing (with Simon Doclo). He was chairman of the IEEE Benelux Signal Processing Chapter (1998–2002), a member of the IEEE Signal Processing Society Technical Committee on Signal Processing for Communications, and President of EURASIP (European Association for Signal Processing, 2007–2008 and 2011–2012). He has served as Editor-in-Chief for the *EURASIP Journal on Applied Signal Processing* (2003–2005), Area Editor for Feature Articles in *IEEE Signal Processing Magazine* (2012–2014), and has been a member of the editorial board of *IEEE Transactions on Circuits and Systems II, IEEE Signal Processing Magazine, Integration-the VLSI Journal, EURASIP Journal on Wireless Communications and Networking,* and *Signal Processing*. He is currently a member of the editorial board of *EURASIP Journal on Advances in Signal Processing*.

•

Toon van Waterschoot received the M.Sc. (2001) and Ph.D. (2009) degrees in electrical engineering, both from KU Leuven, Belgium, where he is currently a tenure-track Assistant Professor. He has previously held teaching and research positions with the Antwerp Maritime Academy, the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT), and the Research Foundation – Flanders (FWO) in Belgium, with Delft University of Technology in The Netherlands, and with the University of Lugano in Switzerland. His research interests are in signal processing, machine learning, and numerical optimization, applied to acoustic signal enhancement, acoustic modeling, audio analysis, and audio reproduction. He has been serving as an Associate Editor for the *Journal of the Audio Engineering Society* (AES) and for the *EURASIP Journal on Audio, Music, and Speech Processing*, and as a Guest Editor for *Elsevier Signal Processing*. He is a Member of the Board of Directors of the European Association for Signal Processing (EURASIP) and a Member of the IEEE Audio and Acoustic Signal Processing Technical Committee (AASP-TC). He was the General Chair of the 60th AES International Conference in Leuven, Belgium (2016), and has been serving on the Organizing Committee of the European Conference on Computational Optimization (EUCCO 2016) and the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2017). He is a member of EURASIP, IEEE, ASA, and AES.