

Highlights

An Efficient Equalization Filter Design Approach for the Accurate Sound Reproduction in a Room Environment with Wide-Band Content

Yongjie Zhuang, Guochenhao Song, Yangfan Liu

- Wide-band accurate sound reproduction in rooms
- An efficient Equalization filter design via convex optimization
- Order reduction based on room response characteristics

This manuscript was published in Applied Acoustics.

DOI: <https://doi.org/10.1016/j.apacoust.2024.109879>

This manuscript version is made available under the CC-BY-NC-ND 4.0 license <https://creativecommons.org/licenses/by-nc-nd/4.0/>

An Efficient Equalization Filter Design Approach for the Accurate Sound Reproduction in a Room Environment with Wide-Band Content

Yongjie Zhuang^{a,b}, Guochenhao Song^a, Yangfan Liu^{a,*}

^a*Ray W. Herrick Laboratories, Mechanical Engineering, Purdue University, 177 S. Russell St., West Lafayette, 47907, Indiana, USA*

^b*Electrical and Computer Engineering, Stony Brook University, 100 Nicolls Road, Stony Brook, 11794, New York, USA*

Abstract

In some sound reproduction applications, a desired sound field needs to be reproduced by remote loudspeakers in a room with sufficient accuracy over a wide frequency band. When low-frequency room modes exist and the low-frequency content is required to be reproduced accurately, an equalization filter with a long time-domain response may be necessary, resulting in a filter with a significantly high filter order, especially when the sampling rate is high for the high-frequency content. This imposes computational challenges on both filter design and implementation. In this article, a convex formulation is proposed for the wide-band sound reproduction filter design. To address issues due to the low-frequency room modes, a reduced-order technique is then proposed based on a sub-band filter structure where a low sampling rate is used to equalize the low-frequency room mode characteristics, and a high sampling rate is used for high-frequency features which usually has a short time span in response. The computational time can be reduced by several orders of magnitude and the numerical stability can be improved when the proposed order reduction method is used to eliminate redundant orders. Simulation and experimental results confirm the accuracy and efficiency advantages of the proposed method.

Keywords: Equalization filter design, convex optimization, optimal filter

*Corresponding author: Yangfan Liu
Email address: yangfan@purdue.edu (Yangfan Liu)

1. Introduction

In general, the purpose of room response equalization is to design equalization filters for loudspeakers to produce the desired sound at particular locations or regions in a given room environment. In typical room response equalization scenarios, the consideration of psychoacoustic effects simplifies the design process, as certain acoustic features may not significantly contribute to human perception of sound. For example, replacing low-frequency dominant poles with poles having smaller magnitudes reduces the extent of oscillations[1, 2] because human perception of sound is less sensitive to low-frequency content in general; using all-pole filters[3, 4, 5] to focus on modeling the spectral peaks instead of notches because the spectral peaks are more audible than the notches in human auditory systems[6], etc. However, the focus of this article centers on the precise reproduction of sound fields within a room environment, particularly for psychoacoustic subjective testing purposes. For an accurate assessment of human perception of sound, it is crucial to reproduce the sound field with accuracy, rather than solely reproducing the acoustic features previously deemed important.

One example of the aforementioned applications is the playback system for headphone-free psychoacoustic subjective tests[7, 8]. To ensure that the sounds being evaluated by subjects are not altered, it is crucial to accurately reproduce all sound signal features, including content over a broad frequency range by designing appropriate equalization filters. Those psychoacoustic experiments are usually performed in a room that is configured to mimic the real environment where the target sound is heard (e.g., an office, factory, or classroom environment depending on the purpose of the subjective test)[7, 8, 9, 10, 11, 12]. A remote hidden loudspeaker away from the subject instead of a headphone is used to minimize any potential influence of equipment appearance on the results of psychoacoustic tests. However, this approach presents challenges due to the low-frequency room modes present in large rooms. Consequently, the equalization filter impulse response must have a long time span. Furthermore, the required high sampling frequency to reproduce sound over the full audible frequency range also necessitates a large number of coefficients for the designed equalization filters, especially when using an FIR filter structure. Another challenge of designing such equalization

filters for subjective tests is that some response constraints may be needed in practical applications. For example, the loudspeaker should perform in its linear response range and the equalization filter should preserve the natural roll-off of loudspeakers at low and high frequencies[1, 13, 14]. Some subjective tests may also require the spectrum power at certain frequency bands to be limited or accurate within a certain small error range.

The frequency domain deconvolution (FDD) method, initially proposed by Kulp[15], is commonly used for high-order constrained filter design applications mentioned above. This method involves obtaining the reciprocal of the room responses in the frequency domain and then applying the fast inverse discrete Fourier transform, which is numerically efficient even for high-order cases. Constraining the filter responses in the frequency domain can also be convenient in the frequency domain. However, the FDD method may result in excessive gains and long impulse responses[1, 16]. Regularization parameters can be added to reduce the impulse responses[17, 18] which may sacrifice the sound reproduction accuracy in the low-frequency content. It is later demonstrated in Section 3 that the use of FDD may cause significant large errors in the low-frequency range.

The least-squares optimization method[1, 14, 19, 20, 21, 22, 23] is another viable approach for such accurate room equalization applications, as it typically yields an optimal equalization filter in terms of sound reproduction accuracy. However, the computational load in the filter design phase for least-squares optimization methods can be significantly high, especially when constraints are present. Therefore, there is a need for an efficient least-squares optimization method. Firstly, convex optimization can be employed to improve computational efficiency. Convex optimization methods have been widely used for efficient optimal filter design in various signal processing applications, e.g., beamformers design[24, 25], active noise control filter design[26, 27], source localization[28, 29, 30], hear-through filter design[31], etc. Secondly, multi-rate approaches can be used to reduce the number of design parameters. One sub-filter having a low sampling rate can cover the low frequencies[1, 32, 33, 34] which usually has a long time span. Thus, the total number of design parameters can be reduced.

In this article, a constrained equalization filter design approach using convex optimization is proposed to efficiently design a high order and high sampling rate equalization filter with filter responses constraints which can be directly used to design both a single full-band filter and sub-band filter structure. A reduced-order technique via sub-band methods is also proposed

to further reduce the computational loads by using the differences in time domain response lengths and sampling frequencies for different frequency bands. When multiple sub-filters are designed using the proposed method, they are not designed separately, instead, one convex optimization problem involving all sub-band filters is solved to achieve the optimal room equalization performance. It is demonstrated through simulation and experimental results that the proposed filter design method improves computational efficiency significantly and can achieve satisfactory equalization performance. For cases where the room is large and a full audio band is required to be reproduced, in the solving process of designing the optimal single full band filter, the traditional least-squares optimization method cannot converge within the practical computational time limit (on the order of several days), however, when the reduce order technique is applied, an equalization filter with good performance can be designed within half an hour. Even if computational power is not a concern, the reduced order technique can improve the numerical stability by eliminating the redundant orders while the traditional least-squares optimization method cannot converge to satisfactory filters due to numerical issues. Compared with the FDD method, the proposed method has better equalization performance, especially in the low-frequency range.

2. Theory

As mentioned earlier, if content in the full audio band is to be reproduced in a relatively large room, a high sampling rate needs to be used to design an optimal equalization filter, and a long impulse response is needed due to the low-frequency room modes, which, inevitably, leads to a large number of filter coefficients. If practical constraints, such as response magnitude constraints in particular bands, are further imposed, not only the computational efficiency in solving the optimal filter coefficients can be low, but numerical instability could also happen for such a large-scale problem. Convex optimization is known to have good computational efficiency, thus, in Section 2.1 and 2.2, the theory of proposed convex formulation is explained under the context of equalization filter design for both a single filter structure and a sub-band filter structure. In Section 2.3, details of how to apply the proposed sub-band-structure-based reduced-order technique to further improve the computational efficiency are described.

It is noted that the FIR filter structure is used. Although an IIR filter structure may reduce the filter parameters in real-time implementation, the

more complicated structure, especially the stability consideration, will make the filter design problem more challenging. If the IIR filter structure is preferred in such high-order applications, a common way is to design a high-order FIR filter using the proposed method in this article first, then fit a stable IIR filter from the pre-designed FIR filter [35].

2.1. Control system description

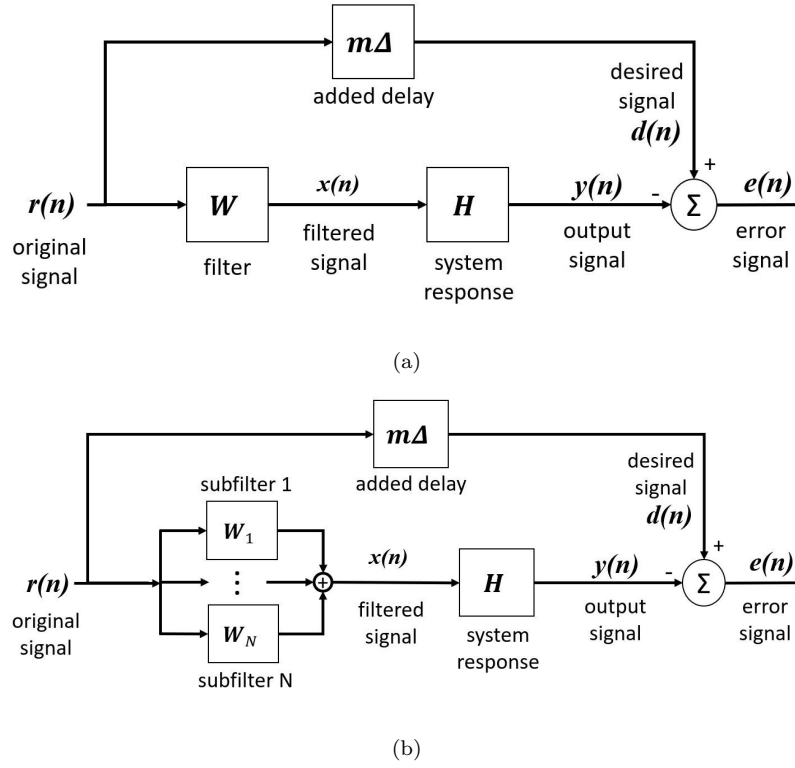


Figure 1: Block diagram for designing a constrained equalization filter (a) in a full band filter structure, (b) in a sub-filters structure.

The system block diagram of an equalization filter structure is shown in Fig. 1(a). In the diagram, $r(n)$ denotes the original signal that is to be reproduced at the desired location. W is the designed equalization filter. $x(n)$ is the output of the filter. H is the system response between the speaker and the listening location which includes both the loudspeaker's frequency response and the room response. $d(n)$ denotes the desired signal which is

obtained by delaying $r(n)$ by $m\Delta$, where Δ is the sampling interval of the original signal and m is an integer that is usually obtained through a few trials (for the applications shown in Section 3, the value of m ranges from 12k to 20k when the sampling rate is 48 kHz). This delay $m\Delta$ is added to ensure a satisfactory sound field reproduction performance if a causal equalization filter W is applied. $y(n)$ is the sound signal at a listening location which should be close to the delayed desired signal ($d(n)$). Thus, the power of error signals $e(n)$ is minimized in this design problem.

To ensure that the formulated design optimization problem is a convex problem, so that a global optimal equalization filter W can be obtained, an FIR filter is used as the equalization filter structure. However, the number of coefficients of this designed FIR filter, i.e., the design parameters, may be large when the required sampling rate is high and the impulse response is long. To reduce the computational load in the filter design process, a sub-band filter technique can be used which is shown in Fig. 1(b). The sub-filters 1 to N are N sub-filters that can be implemented in multi-rate systems to reduce the number of parameters in filter design as well as in the filter real-time implementation process. The meanings of other symbols in Fig. 1(b) are similar to those in Fig. 1(a). It is noted that, if the sub-band filters are to be implemented in real-time filtering, the system response H should also include the analysis filter bank and synthesis filter bank responses.

2.2. Convex H_2/H_∞ formulation

The formulations for both a single filter structure and a sub-band filter structure for the equalization filter design problem are described here. Firstly, the frequency responses of the equalization filter W in Fig. 1(a) can be expressed as

$$W(f_k) = \vec{F}(f_k, f_s, N_t)^T \vec{w}, \quad (1)$$

where T denotes the transpose operation, f_k denotes the k th frequency, $\vec{w}$ is the coefficients of the equalization filter, f_s denotes the sampling frequency of the equalization filter, N_t denotes the number of equalization filter coefficients, and

$$\vec{F}(f_k, f_s, N_t) = \begin{bmatrix} 1 & e^{-j2\pi f_k / f_s} & \dots & e^{-j2\pi f_k (N_t-1) / f_s} \end{bmatrix}^T. \quad (2)$$

For a sub-band filter structure (shown in Fig. 1(b)), the total frequency response of all sub-filters, $\tilde{W}(f_k)$, can be expressed as

$$\tilde{W}(f_k) = \sum_{i=1}^N W_i(f_k) = \sum_{i=1}^N \begin{bmatrix} 1 & e^{-j2\pi f_k / f_{s_i}} & \dots & e^{-j2\pi f_k (N_{t_i}-1) / f_{s_i}} \end{bmatrix} \vec{w}_i, \quad (3)$$

where $\vec{w}_i$ is the coefficients of the i th sub-filter, and f_{s_i} is the sampling frequency of i th sub-filter. It is noted that the Eq. (3) can be rearranged to

$$\tilde{W}(f_k) = \vec{\tilde{F}}(f_k)^T \vec{w}, \quad (4)$$

where

$$\vec{\tilde{F}}(f_k) = \begin{bmatrix} \vec{F}(f_k, f_{s_1}, N_{t_1}) \\ \vdots \\ \vec{F}(f_k, f_{s_N}, N_{t_N}) \end{bmatrix}, \quad \vec{w} = \begin{bmatrix} \vec{w}_1 \\ \vdots \\ \vec{w}_N \end{bmatrix}, \quad (5)$$

and N_{t_i} is the number of coefficients of the i th sub-filter. By comparing Eq. (1) and Eq. (4), it is obvious that the formulation to design the sub-filters (Fig. 1(b)) is equivalent to that for a single filter (Fig. 1(a)), if the modified Fourier transform kernel, $\vec{\tilde{F}}(f_k)$ is used. And designing a full-band filter is essentially a special case of this sub-filter approach. Then, the H_2/H_∞ formulation used in designing a single full-band equalization filter (Fig. 1(a)) can be applied directly to the design of equalization sub-filters (Fig. 1(b)). Thus, the rest of the descriptions in this subsection applies to both a single full-band filter structure and a sub-band filter structure.

The objective function $J(\vec{w})$ can be specified to minimize the total power of error signal $e(n)$ at all desired frequencies:

$$J(\vec{w}) = \sum_{k=1}^{N_f} J_k, \quad (6)$$

where N_f denotes the total number of frequencies considered in the desired equalization frequency band, and J_k is the expected power of error signal at the k -th frequency f_k , $e_f(f_k)$ (the subscript f is used to denote the frequency domain of the associated signal). For convenience, the frequencies can be equally spaced. However, non-uniform frequency resolution is also possible if needed. The N_f should be large enough compared with the filter order (i.e.,

the number of variables in the optimization problem) to prevent overfitting. J_k can be expressed as [36, 26]

$$\begin{aligned}
J_k &= \mathbb{E} (|e_f(f_k)|^2) \\
&= \mathbb{E} (|d_f(f_k) - y_f(f_k)|^2) \\
&= \mathbb{E} (|e^{-j2\pi f_k m \Delta} r_f(f_k) - H(f_k) \tilde{W}(f_k) r_f(f_k)|^2) \\
&= \left(1 - 2\Re \left(H(f_k) \tilde{W}(f_k) e^{+j2\pi f_k m \Delta} \right) + |H(f_k)|^2 |\tilde{W}(f_k)|^2 \right) S_{rr},
\end{aligned} \tag{7}$$

where S_{rr} is the auto spectral density function of the signal $r(n)$, $\Re$ denotes taking the real part of a complex-valued number. Minimizing the error signal power instead of directly inverting the room responses can inherently add higher weightings to frequencies having higher desired sound power. Equation (7) can be further reformulated as a standard convex quadratic function (See Appendix in Zhuang and Liu's previous work[26] for detail)

$$J_k = \vec{w}^T \mathbf{A}_J(f_k) \vec{w} + \vec{\mathbf{b}}_J^T(f_k) \vec{w} + S_{rr}(f_k), \tag{8}$$

where

$$\begin{aligned}
\mathbf{A}_J(f_k) &= |H(f_k)|^2 S_{rr}(f_k) \Re \left(\vec{F}(f_k)^* \vec{F}(f_k)^T \right), \\
\vec{\mathbf{b}}_J(f_k) &= -2S_{rr}(f_k) \Re \left(H(f_k) e^{+j2\pi f_k m \Delta} \vec{F}(f_k) \right),
\end{aligned}$$

and $*$ denotes the complex conjugate operation.

In practical applications, constraints may be imposed on the amplitude of the equalization filter response at various frequencies:

$$|\tilde{W}(f_k)|^2 \leq \tilde{C}(f_k), \tag{9}$$

or sub-filters' responses:

$$|W_i(f_k)|^2 \leq C_i(f_k), \tag{10}$$

are sometimes needed to ensure loudspeakers operate in their linear response regions, or, if a sub-band filter structure is used in real-time implementation, to ensure each sub-band filter has proper roll-off response characteristics in their transition bands. The $\tilde{C}(f_k)$ and $C_i(f_k)$ are nonnegative constants specified according to the required magnitude constraints. Equations (9) and (10) can be reformulated as standard convex quadratic inequalities:

$$\vec{w}^T \Re \left(\vec{F}(f_k)^* \vec{F}(f_k)^T \right) \vec{w} - \tilde{C}(f_k) \leq 0 \tag{11}$$

$$\vec{w}_i^T \Re \left(\vec{F}(f_k, f_{s_i}, N_{t_i})^* \vec{F}(f_k, f_{s_i}, N_{t_i})^T \right) \vec{w}_i - C_i(f_k) \leq 0 \quad (12)$$

Sometimes, a constraint on the allowed maximum error can be applied to enforce the sound field reproduction accuracy being within prescribed bounds at particular frequencies:

$$J_k \leq B(f_k), \quad (13)$$

where $B(f_k)$ is the prescribed square error bound. Equation (13) can also be reformulated into a standard convex quadratic function as [36, 26]

$$\vec{w}^T \mathbf{A}_J(f_k) \vec{w} + \vec{\mathbf{b}}_J^T(f_k) \vec{w} + S_{rr}(f_k) - B(f_k) \leq 0, \quad (14)$$

where $\mathbf{A}_J(f_k)$ and $\vec{\mathbf{b}}$ are defined in Eq. (8).

By using Eq. (6) as the objective function, Eq. (11) (12) (14) as the constraints, the convex optimization for designing the equalization filter (for either a single full band filter or a sub-band filter structure) using the H_2/H_∞ formulation can be written as

$$\begin{aligned} \min_{\vec{w}} . \quad & \vec{w}^T \sum_{k=1}^{N_f} \mathbf{A}_J(f_k) \vec{w} + \sum_{k=1}^{N_f} \vec{\mathbf{b}}_J^T(f_k) \vec{w} + \sum_{k=1}^{N_f} S_{rr}(f_k) \\ \text{s.t.} \quad & \vec{w}^T \Re \left(\vec{F}(f_k) \vec{F}(f_k)^H \right) \vec{w} - \tilde{C}(f_k) \leq 0, \quad \text{for required } f_k \\ & \vec{w}_i^T \Re \left(\vec{F}(f_k, f_{s_i}, N_{t_i}) \vec{F}(f_k, f_{s_i}, N_{t_i})^H \right) \vec{w}_i - C_i(f_k) \leq 0, \quad \text{for required } f_k \text{ and } i \\ & \vec{w}^T \mathbf{A}_J(f_k) \vec{w} + \vec{\mathbf{b}}_J^T(f_k) \vec{w} + S_{rr}(f_k) - B(f_k) \leq 0, \quad \text{for required } f_k. \end{aligned} \quad (15)$$

This convex optimization problem can be reformulated to a cone programming problem which is one type of convex optimization problem and it can be solved efficiently by suitable convex optimization toolboxes [37, 38] or by primal-dual interior-point methods that are designed for cone programming problems [39, 40, 41, 42]. Demonstration of efficiently solving a similar optimization can be referred to Zhuang and Liu's previous work [26] in which the computational time is improved by several orders.

2.3. A reduced-order technique

Although some efficient algorithms, such as primal-dual interior-point methods designed for convex programming, can be applied to efficiently solve the optimization Eq. (15), the computational load can still be significant if the designed equalization filter has a large number of coefficients due to high

sampling frequency and long impulse response time. When sound reproduction is performed in a large room, the impulse response length is long mainly due to low-frequency modes in the room response, a sub-band filter-structure-based reduced-order technique is proposed in this section to further reduce the number of filter parameters to be calculated in the optimal filter design process.

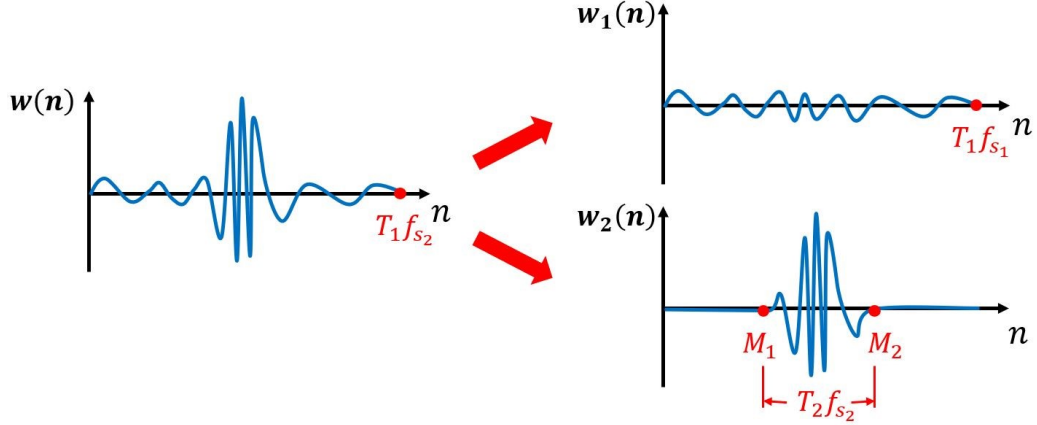


Figure 2: An illustration of dividing an equalization filter into two sub-filters with different impulse response lengths and sampling frequencies.

In large room sound reproduction applications, although the desired equalization filter has a long impulse response and high sampling frequency, the higher frequency information is usually concentrated in a short time span. This implies that, if the equalization filter to be designed includes several sub-band filters, only sub-filters at low frequencies (which are to equalize the room response features that are dominated by low-frequency room modes) have long impulse response times, while the sub-filters in high-frequency sub-bands may only have a short impulse response time length. Order reduction can then be achieved by using a low sampling rate for low-frequency sub-filters that have long response times, while a high sampling rate for high-frequency sub-filters with short response time length. An illustration of this is shown in Fig. 2, where $w(n)$ represents the desired full band equalization filter whose length is T_1 and a high sampling frequency, f_{s_2} , has to be used to effectively reproduce the high-frequency content; $w_1(n)$ represents a sub-filter that has a longer length T_1 but with a lower sampling frequency f_{s_1} , and $w_2(n)$ represents a sub-filter that has a shorter length T_2 and higher

sampling frequency f_{s_2} . It is also noted that if the loudspeaker is at a distant location from the listening location, the high-frequency sub-filter $w_2(n)$ may exhibit a pure time delay behavior due to the large added delay (i.e., $m\Delta$ in Fig. 1), which, in Fig. 2, corresponds to the zero response in impulse response before the time M_1 . The parameter M_1 can be estimated to be the difference of added delay ($m\Delta$ in Fig. 1) and the time for sound propagating from the loudspeaker to the listening location. Once the time delays M_1 is estimated, there is no need to have undetermined filter coefficients in $w_2(n)$ before M_1 , which is an additional order reduction. Due to the existence of low-frequency room modes in a relatively large room, $T_1 \gg T_2$ and $f_{s_1} \ll f_{s_2}$, this then suggests that the required number of design parameters for a sub-band structure, $T_1 f_{s_1} + T_2 f_{s_2}$, is much smaller than that for a full band filter structure, $T_1 f_{s_2}$.

The implementation of this reduced-order technique is convenient if the convex formulation in the previous subsection is used. For example, when implementing the reduced-order technique for the case in Fig. 2, the only formulation that needs modification is the Fourier transform vector $\vec{F}(f_k, f_{s_2}, T_1 f_{s_2})$ associated with the sub-filter $w_2(n)$. More specifically, the Fourier transform vector $\vec{F}(f_k, f_{s_2}, T_1 f_{s_2})$ should be modified to a reduced-order Fourier transform vector $\vec{F}_r(f_k, f_{s_2}, M_1, M_2)$:

$$\vec{F}_r(f_k, f_{s_2}, M_1, M_2) = \begin{bmatrix} e^{\frac{-j2\pi f_k M_1}{f_{s_2}}} & \dots & e^{\frac{-j2\pi f_k M_2}{f_{s_2}}} \end{bmatrix}^T, \quad (16)$$

which is mathematically equivalent to assuming the impulse response for the sub-filter $w_2(n)$ before time M_1/f_{s_2} is zero. Thus, the designed parameters can be significantly reduced to reduce the computational load and avoid potential numerical instability issues for large-scale optimization problems.

In practice, the approximate choice of f_{s_1} can be obtained by checking the room impulse responses to see above what frequency the response tends to be concentrated in a small time interval.

3. Results

3.1. Simulation results

3.1.1. Simulation setup and simulated responses

To investigate the proposed filter design approach. Rooms of two different sizes (denoted as “small room” and “large room”) are simulated using

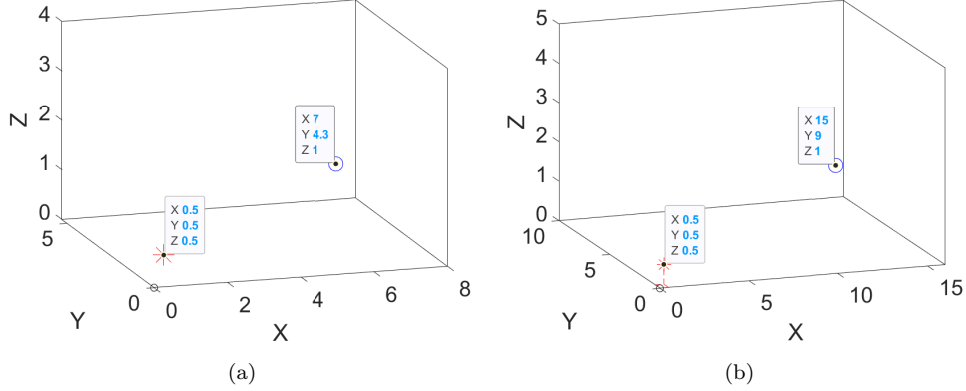


Figure 3: An illustration of setups for simulating (a) the small room and (b) the large room. The stars represent the source and the circles represent the listening location. (Unit: meters)

the image sources method. The dimensions of the simulated small room are 8 meters in length, 5.3 meters in width, and 4 meters in height. And the dimensions of the simulated large room are 16 meters in length, 10 meters in width, and 5 meters in height. In both rooms, the frequency-dependent reflection coefficients of each surface are obtained by assuming the floor is carpeted on concrete and all other surfaces are acoustic tiles on rigid surfaces[43]. The speaker is assumed to be a monopole positioned at a corner 0.5 meters away from the floor and the other two surfaces. The listening location is assumed to be 1 meter in height and 1 meter away from both surfaces. The illustration of setups for the two simulated rooms is shown in Fig. 3. The white noise is used as the desired sound clip to show the equalization performance when equal weighting is applied to all frequencies.

The impulse responses and the frequency responses of the paths from the speaker input to the listening location in two simulated rooms are shown in Fig. 4. A point-wise smoothing was implemented with a moving Hann window which covers a one-third octave band centered on the frequency [8, 44]. From Fig. 4(a)(c), the high-frequency content is concentrated in a narrow time-domain region while the low-frequency content spreads out over a long time span. This is one of the key motivations in the proposed reduced-order technique using sub-band design. In the frequency response, “True responses” demonstrates the actual simulated results while the “Used responses” demonstrates the responses used to design filters. The only dif-

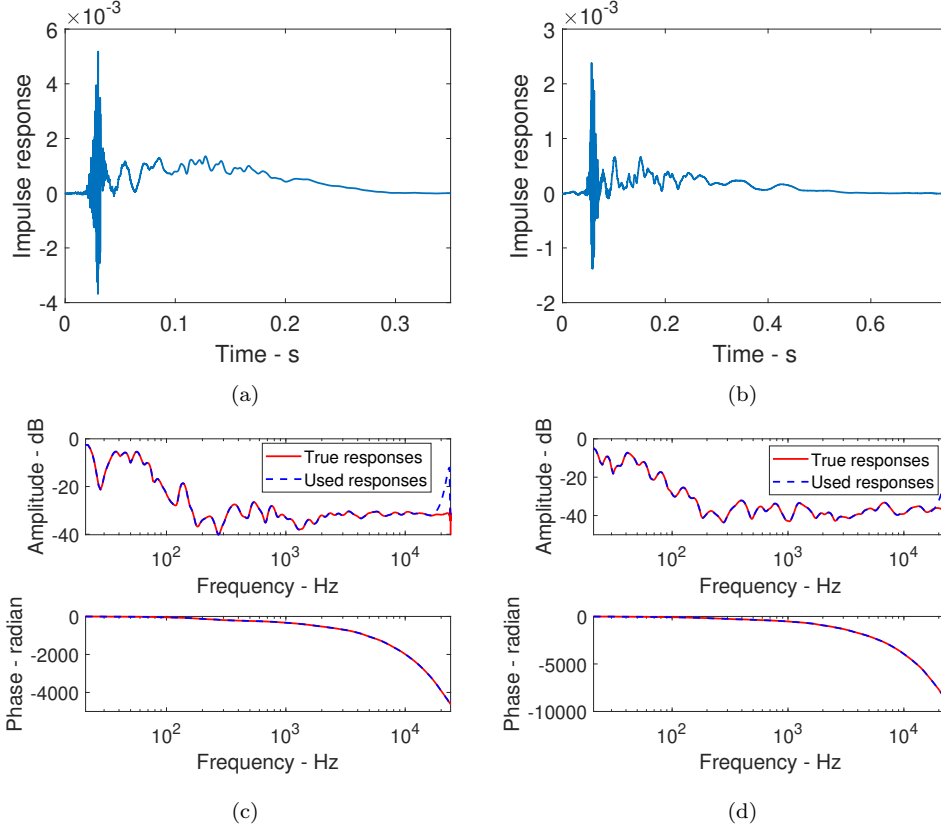


Figure 4: Impulse responses and frequency responses of the room response (the path from the speaker input to the listening location). (a): impulse responses in the small room case; (b): impulse responses in the large room case; (c): frequency responses in the small room case; (d): frequency responses in the large room case.

ference between the “True responses” and the “Used responses” is that the “Used responses” have higher responses towards the cut-off frequency to give a better roll-off in the designed equalization filter.

3.1.2. Comparison of room equalization performance in simulated room cases

To design the equalization filters, the sampling rate for the full-band filter is 48 kHz, and the sampling rates for the two sub-filters are 2.4 kHz and 48 kHz for all simulated room cases. The test signal is white noise. The delay added is 0.3125 s (i.e., $m = 15000$) in the small room case and 0.2625 s (i.e., $m = 12600$) in the large room case for both methods. The high-sampling-rate sub-filter W_2 starts at 0.25 s (i.e., M_1 in Fig. 2 is 12000) for the small

Table 1: Comparison of the number of design parameters using the traditional least-squares optimization method and the proposed method in small room case.

	Traditional Method	Proposed Method	
		sub-filter 1	sub-filter 2
Time length (s)	0.5	0.5	0.05
Sampling rate (kHz)	48	2.4	48
Filter order	24000	1200	2400

Table 2: Comparison of the number of design parameters using the traditional least-squares optimization method and the proposed method in large room case.

	Traditional Method	Proposed Method	
		sub-filter 1	sub-filter 2
Time length (s)	0.9	0.9	0.06
Sampling rate (kHz)	48	2.4	48
Filter order	43200	2160	2880

room case and starts at 0.16 s ($M_1 = 7680$). Frequency response magnitudes are constrained to prevent excessively large amplification at very low and high frequencies by using the response magnitude constraints Eqs. (11) (12). The filter orders using the proposed reduced-order technique and traditional least-squares optimization methods are compared in Table 1 and Table 2. From the tables, the total number of filter orders is reduced from 24000 to $1200+2400 = 3600$ (i.e., 85% reduction or around 6.5 times smaller) in the small room case and from 43200 to $2160+2880 = 5040$ (i.e., 88.3% reduction or around 8.5 times smaller) in the large room case. This significant reduction in the number of design parameters can reduce the filter design solution time by several orders of magnitude. The required computational time is plotted in Fig. 5. Compared with 25 hours of computational time using the traditional method using filter order 24000, the proposed method requires only 10 minutes because the filter order can be reduced to 3600. In Fig. 6, it demonstrates that when a low-order traditional method is used (either order of 3600 or 7200), the equalization performance is poor compared with the proposed method (order of 3600). Even if the computational time is not a concern, directly using the high-order traditional method (order of 24000) is not reliable because numerical issues may happen when excessive redundant information is present [27]. In the case of using an order of 24000, Fig. 6

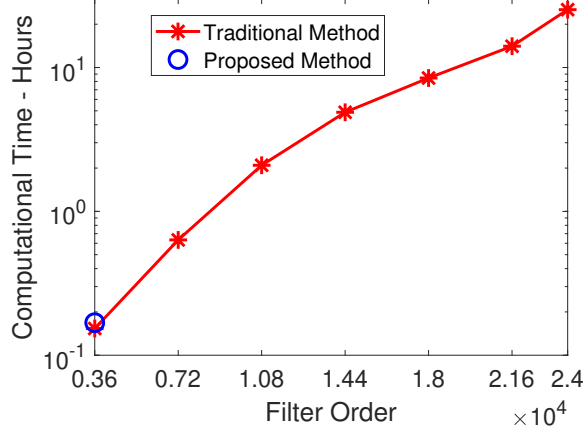


Figure 5: A comparison of required computational time for the traditional and proposed methods applied to the small room case.

shows that the traditional method cannot converge to satisfactory results in practice. Thus, the proposed order reduction method can improve the numerical stability.

The designed equalization filter using the traditional FDD method and the proposed method is compared in Fig. 7. The designed two sub-filters using the proposed method are also shown in Fig. 7(a)(b). Although the designed equalization filters are similar using the two approaches over a wide range of frequencies, the proposed method has a significantly better equalization performance at the low-frequency range which is shown in Fig. 8. In Fig. 8(c)(d), the frequency responses after combining the designed equalization filters with the room responses are demonstrated. Ideally, it should be 0 dB across the audible frequency range. The equalization performance using the FDD has larger variations compared with the proposed method because the proposed method gives the optimal equalization filters for the same equalization filter length. The simulated reproduced sound power spectral density (PSD) is also shown Fig. 9.

3.2. Experimental results

In this section, the proposed equalization filter design approach is applied to a sound playback system for a psychoacoustic subjective test in an actual large room, which requires the sound clip being tested to be reproduced accurately for a wide frequency band at the listening location[8].

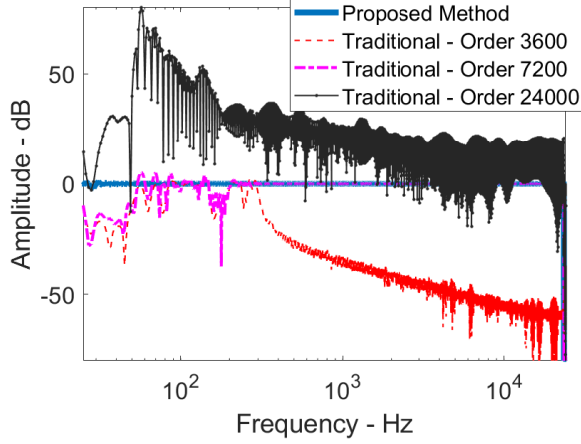


Figure 6: A comparison of equalization performance for the traditional and proposed methods applied to the small room case.

3.2.1. Experimental setup and measured responses

Since the objective of this particular psychoacoustic test is to evaluate office noise, an office mock-up was set up to create a natural testing environment for subjects. The desired sound to be reproduced is the background noise measured in an actual office environment[8]. The prefiltered sound clip is played by a hidden distant speaker and the desired sound should be reproduced at the listening location. Hidden loudspeakers are used to ensure that the sound evaluation is not influenced by the subjects' perceived cue that sound is artificially generated by an audio system. The dimensions of the room are 13.1 meters in length, 8.5 meters in width, and 6.7 meters in height. Figure 10(a) shows the layout of the experimental setup including the listening location and the speaker location, and Figure 10(b) shows a picture of the listening location. Because there is sound scattering and absorption from multiple objects, such as the partitions and tables, the sound field reproduction is more complicated compared with that in an empty room if an accurate sound field reproduction performance is required. Since the listening location is fixed in the room and it is not likely to have a large variation in subjects' head position and orientation, the requirement of reproducing the sound at a particular location is considered to be sufficient for this application. Also, no artificial head or torso is involved in measurement related to the sound reproduction system, since the binaural effect will occur naturally due to head and torso scattering when a subject is sitting at the

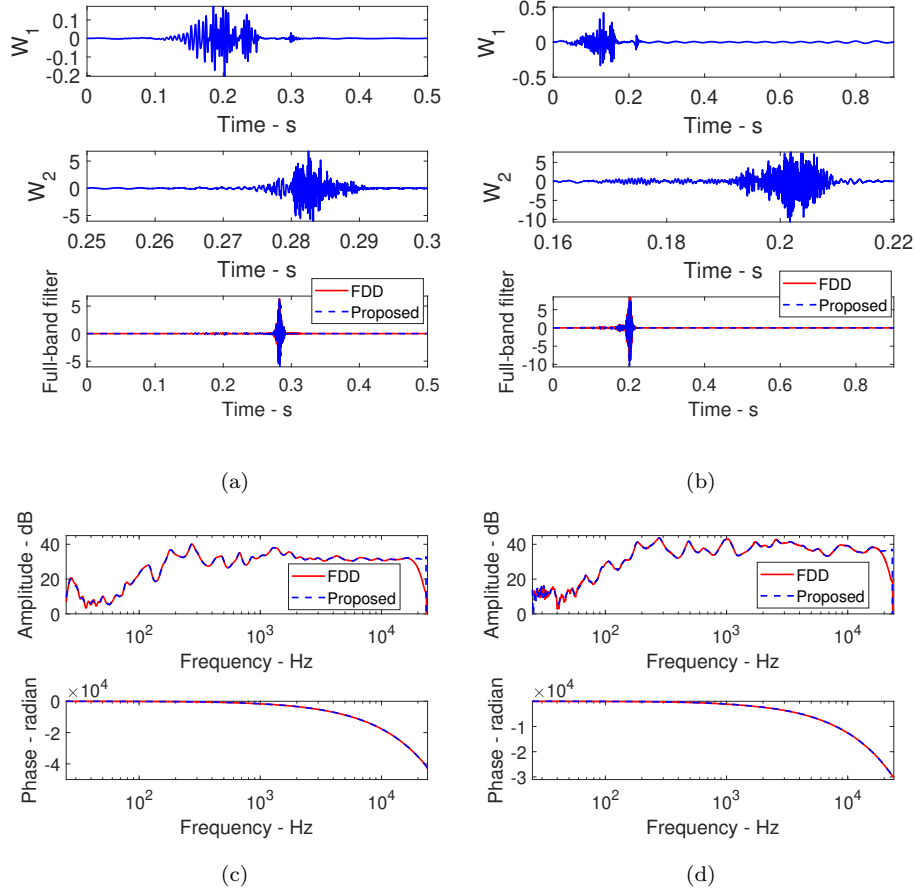


Figure 7: The comparison of the designed equalization filter using the frequency-domain deconvolution (denoted as “FDD” in the plot) and the proposed method in time and frequency domain. (a): time domain for the small room case; (b): time domain for the large room case; (c): frequency domain for the small room case; (d): frequency domain for the large room case.

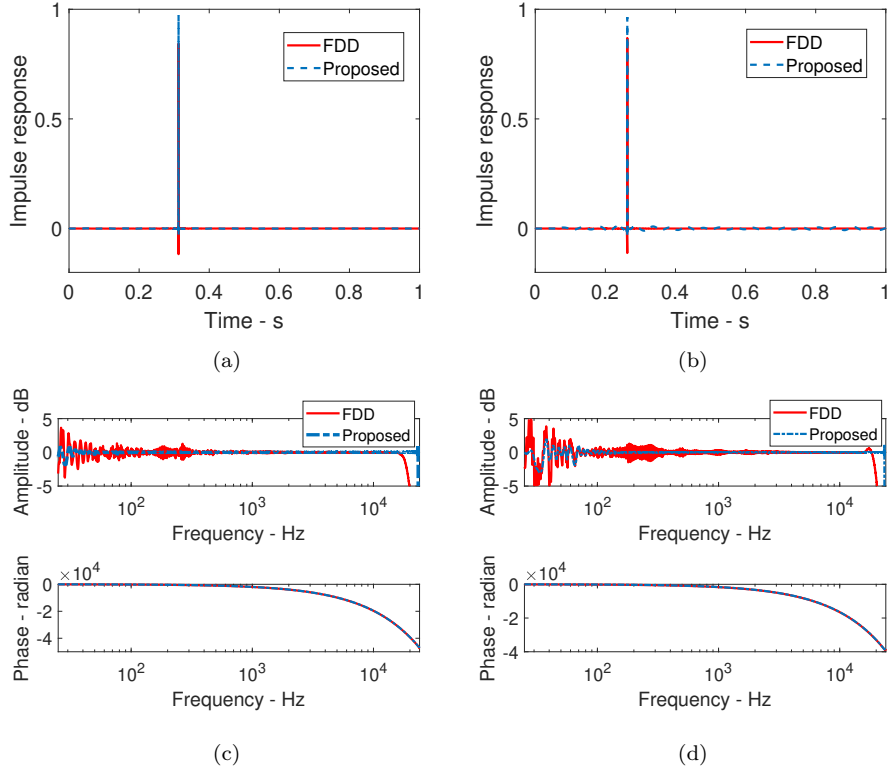


Figure 8: The comparison of the convolution result of impulse responses and the multiplication of frequency responses of the designed equalization filters and room responses using the frequency-domain deconvolution (denoted as “FDD” in the plot) and the proposed method. (a): impulse responses in the small room case; (b): impulse responses in the large room case; (c): frequency responses in the small room case; (d): frequency responses in the large room case.

listening location in the reproduced sound field.

The playback system that produced the filtered sound clip consists of a LynxONE sound card, a Furman SP20AB amplifier, and a loudspeaker (ALTEC N1201-8A). The frequency response function from the computer output signal to sound pressure at the listening location was measured by collecting the input and output signals and computing the cross-spectral density function between input and output signals and the power spectral density function of the input signals using Welch’s method. A point-wise smoothing was then implemented with a moving Hann window which covers a one-third octave band centered on the frequency [8, 44]. The measured

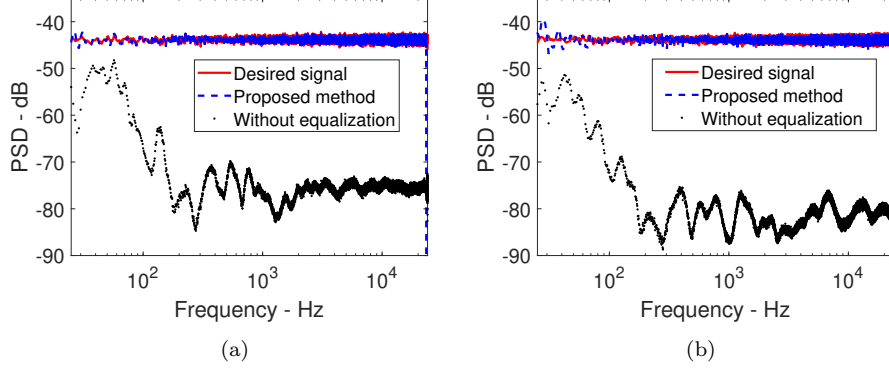


Figure 9: Comparison of power spectral density (PSD) of reproduced signal and desired signal at listening location for simulated (a) small room and (b) large room cases.

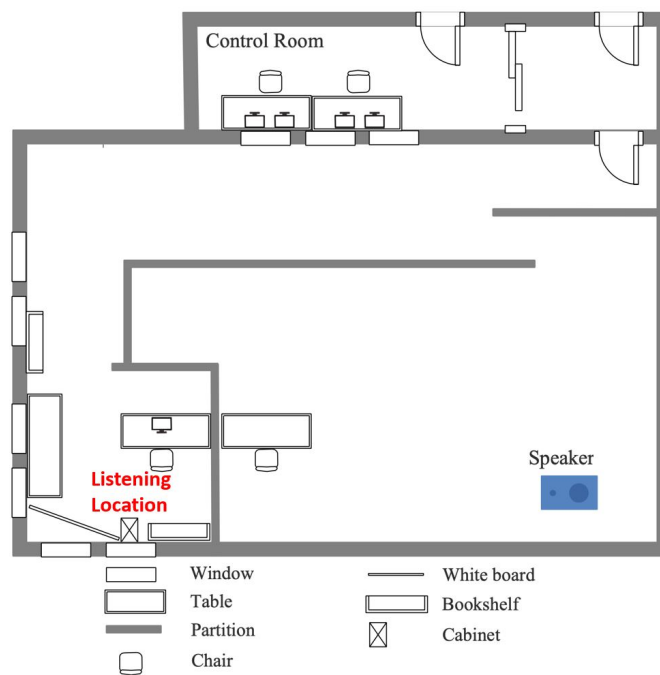
impulse response can be obtained by doing an inverse Fourier transform of the computed frequency response. The measured impulse response and the frequency response of the paths from the speaker input to the listening location in this room setup are shown in Fig. 11. From Fig. 11(a), it is clear that the high-frequency content is concentrated in a narrow time span while the low-frequency content spreads over a wide time span. Similar to the simulation case, the modified response (“Used responses”) is used in the design phase instead of the “True responses” to get a better roll-off towards the cut-off frequency.

3.2.2. Comparison of room equalization performance for the actual room case

Table 3: Comparison of design parameters using different methods.

	Direct Method	Proposed Method	
		sub-filter 1	sub-filter 2
Time length (s)	0.8	0.8	0.0628
Sampling rate (kHz)	48	2.4	48
Filter order	38400	1920	3000

To design the equalization filters, the sampling rate for the full-band filter is 48 kHz, and the sampling rates for the two sub-filters are 2.4 kHz and 48 kHz. The desired sound clip was an ambient broadband noise measured in an office space. The delay added is 0.4167 s (i.e., $m = 20000$) for both methods. The high-sampling-rate sub-filter W_2 starts at 0.33 s (i.e., M_1 in Fig. 2



(a)



(b)

Figure 10: (a) Layout of the experimental setup, (b) picture of the listening location.

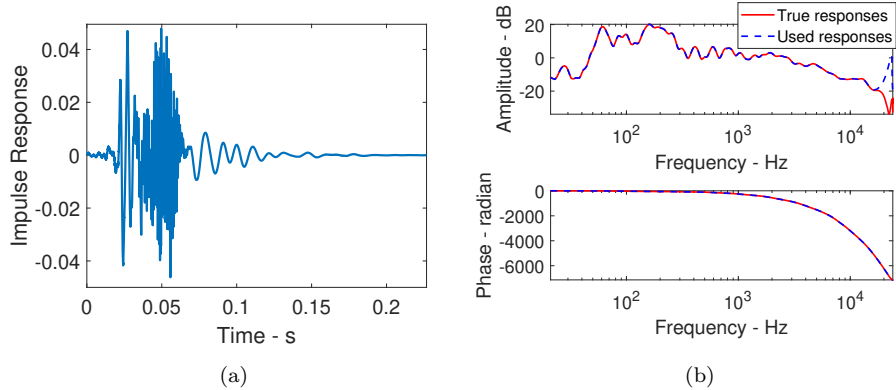


Figure 11: (a) Impulse responses and (b) frequency responses of the room response (the path from the speaker input to the listening location).

is 16000). Frequency response magnitudes are constrained to prevent excessively large amplification at very low and high frequencies. The filter orders using the proposed reduced-order technique and traditional least-squares optimization methods are compared in Table 3. From the tables, the total number of filter orders is reduced from 38400 to $1920+3000 = 4920$ (i.e., 87.2% reduction or around 8 times smaller). Similar to the simulated room cases, this significant reduction in the number of design parameters can reduce the filter design solution time by several orders of magnitude from days to below half an hour.

The designed equalization filter using the traditional FDD method and the proposed method is compared in Fig. 12. For this complicated room environment, the difference in the designed equalization filters in the frequency domain is more obvious compared with the simulated room cases in Fig. 7. In Fig. 13, the impulse responses and frequency responses after combining the designed equalization filters with the room responses are demonstrated. The complicated room environment setup further increases the required filter length which makes the FDD method perform worse. Also, the complicated real-life noise in office setup has large power level variations across the audible frequencies. Compared with the FDD method, the use of the proposed method inherently adds higher weightings to frequencies having higher desired sound power. The proposed method has very small variations around 0 dB across the audible frequency range (within 2 dB variations). However, the FDD method may have variations up to 20 dB which is not suitable for this psychoacoustic testing.

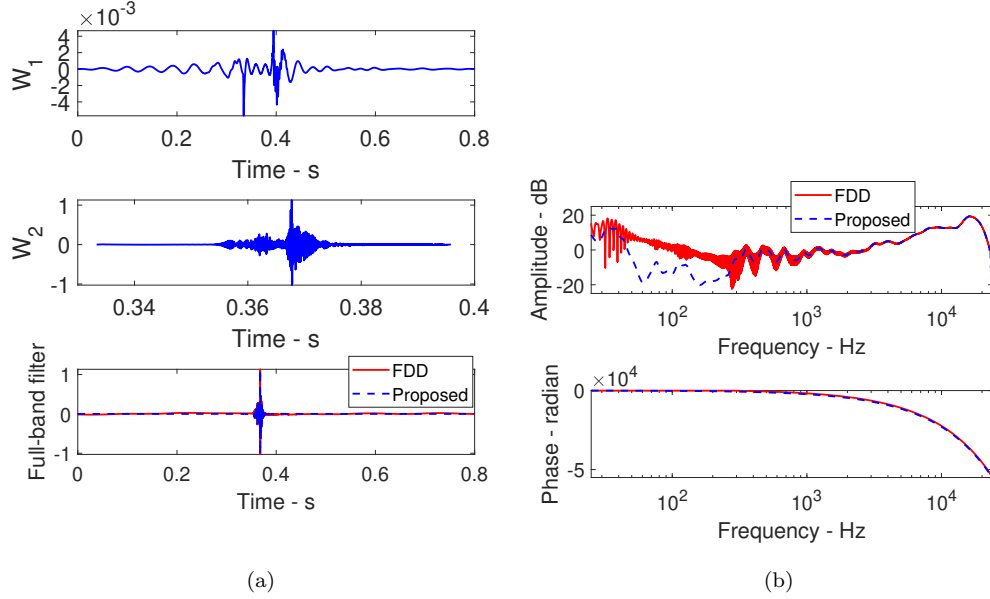


Figure 12: The comparison of the designed equalization filter using the frequency-domain deconvolution (denoted as “FDD” in the plot) and the proposed method in (a) time and (b) frequency domain.

The simulated sound reproduction performance is shown in Fig. 14 and the experimentally measured sound reproduction performance is shown in Fig. 15(a). In Fig. 14, compared with the reproduced signal without using an equalization filter, the reproduced signal after using the equalization filter designed by the proposed method is much closer to the desired signal. To further confirm the sound field reproduction performance, the sound pressure level (SPL) of the desired signal and the reproduced signal measured in the experiment are compared in Fig. 15(a). In the experiment, the reproduced signal after using the equalization filter designed by the proposed method is still close to the desired signal. Thus, the sound field reproduction performance of using the proposed equalization filter design method is experimentally validated. To show the spatial variations of the reproduced sound signal using the proposed method, the measured SPL at three locations 10 cm to the left of, to the right of, and behind the listening location (i.e., the location where measurement was conducted) are shown in Fig. 15(c)-(d). Given that human subjects are seated in front of a table, 10 cm can be considered a relatively large distance for the purpose of this study, as

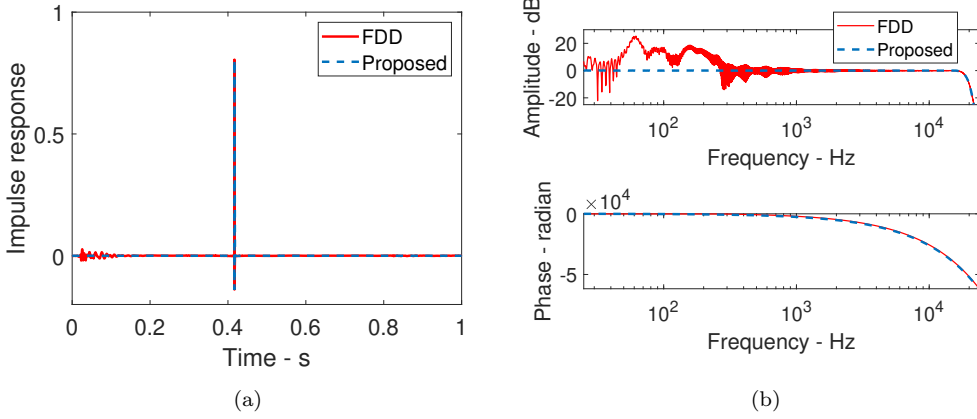


Figure 13: The comparison of (a) the convolution result of impulse responses and (b) the multiplication of frequency responses of the designed equalization filters and room responses using the frequency-domain deconvolution (denoted as “FDD” in the plot) and the proposed method.

head movement is typically limited. In Fig. 15(c)-(d), the equalization performance is still satisfactory for a wide range of frequencies except for some spikes and notches due to the change of low-frequency room modes which is typical for a least-squares optimization-based local sound field reproduction method.

4. Conclusion

In this article, a convex formulation for designing a constrained equalization filter for local sound field reproduction using H_2/H_∞ formulation is proposed. This formulated optimization problem can be solved by some computationally efficient optimization algorithms. Either the design of a full band equalization filter or the design of multiple sub-filters can be formulated in a similar way by using this method. Multiple sub-band filters can be designed simultaneously by solving one convex optimization problem such that the total sound field reproduction performance can be optimal. Constraints on the frequency responses of each sub-band filter and the total equalization filter effect can be specified in a convenient way. A reduced-order technique by using the sub-band filters is also proposed to further reduce the number of filter design parameters. The order reduction performance of this sub-band structure is particularly effective for sound reproduction by remote loudspeakers in a relatively large room. In this order reduction method, filters

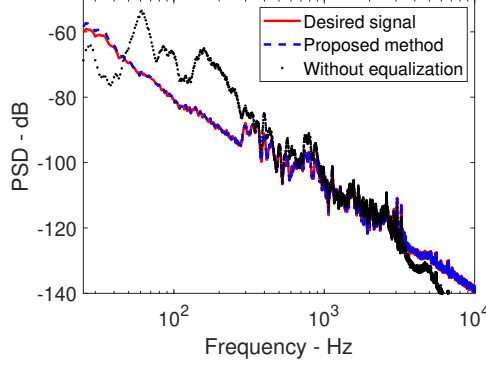


Figure 14: Comparison of simulated power spectral density (PSD) of reproduced signal and desired signal at the listening location.

for low-frequency sub-bands, which are to equalize room response features dominated by low-frequency room modes, can be designed using a low sampling rate, while the high-frequency sub-band filters usually only have short impulse response time length and can be designed using a high sampling frequency.

The proposed method was investigated in simulated room cases and an experiment with a playback system for psychoacoustic subjective tests. Results show that, by using the proposed method, the desired sound signal can be reproduced at the required location successfully. Compared with the traditional least-squares optimization methods, the use of a reduced-order technique can result in a satisfactory sound field reproduction performance while reducing the number of required design parameters significantly. This reduction in the number of design parameters can reduce the computational time from the order of days to within half an hour. The elimination of redundant orders using the proposed reduced-order technique can also improve numerical stability. Compared with the traditional frequency domain deconvolution methods, the proposed method has a better sound reproduction accuracy at the low-frequency range. The sound reproduction performance is also measured at locations 10 cm to the left of, to the right of, and behind the listening location (measuring location) to demonstrate the spatial robustness of the proposed method.

With the recent trend of studying psychoacoustic effects in office or factory environment[7, 8, 9, 10, 11], the proposed method can be applied to efficiently design equalization filters for these studies. Precise reproduction

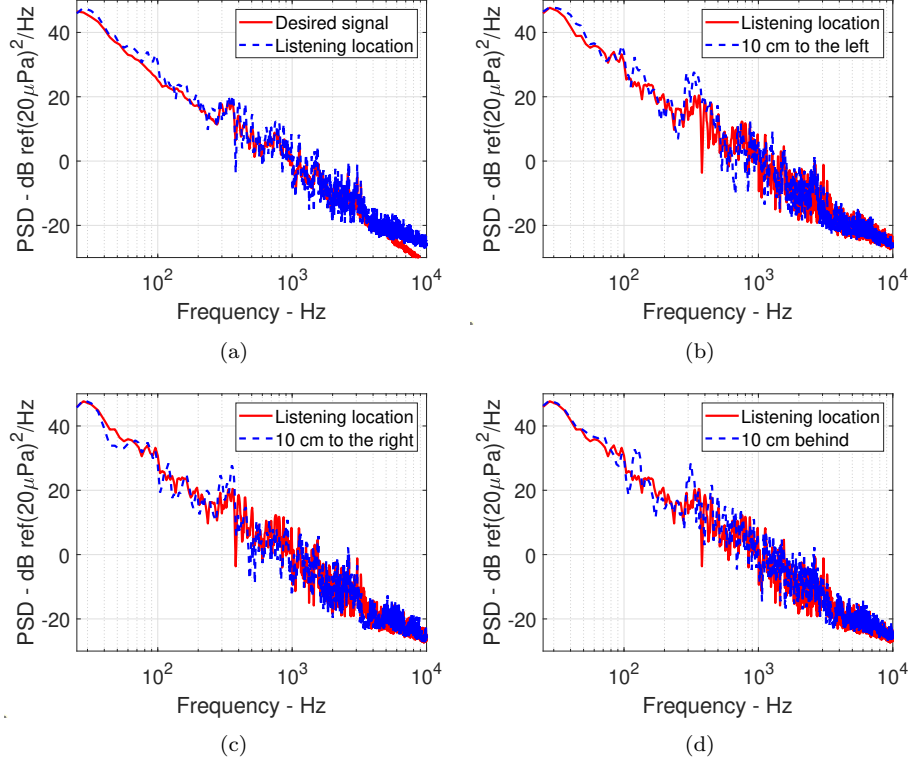


Figure 15: Comparison of experimentally measured sound pressure level (SPL) at (a) listening location and location 10 cm to the (b) left of, (c) right of, and (d) behind the listening location. The subfigures (b) – (c) are used to demonstrate the spatial variations after equalization

can reduce the bias in the sound perception study and the improvements in computational efficiency can shorten the research cycle. The proposed method can also be applied to obtain better acoustic environments such as in virtual conference rooms, etc. In the future, the automatic choice of filter lengths, the number of subbands, and the sampling frequency for each band based on the impulse responses of rooms can be investigated.

Acknowledgments

The authors thank Beijing Ancsonic Technology Co., Ltd. for providing financial support.

Author Contributions

- Yongjie Zhuang: conceptualization, methodology, software, validation, formal analysis, investigation, writing—original draft preparation, visualization
- Guochenhao Song: data curation, validation
- Yangfan Liu: conceptualization, resources, writing—review and editing, supervision, funding acquisition, project administration.

Conflicts of Interest

The authors declare no conflict of interest. The funder had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

- [1] S. Cecchi, A. Carini, S. Spors, Room response equalization—a review, *Applied Sciences* 8 (1) (2017) 16.
- [2] A. Maamar, I. Kale, A. Krukowski, B. Daoud, Partial equalization of non-minimum-phase impulse responses, *EURASIP Journal on Advances in Signal Processing* 2006 (2006) 1–8.
- [3] J. Durbin, The fitting of time-series models, *Revue de l’Institut International de Statistique* (1960) 233–244.
- [4] J. Mourjopoulos, M. Paraskevas, Pole and zero modeling of room transfer functions, *Journal of Sound and Vibration* 146 (2) (1991) 281–302.
- [5] A. Carini, S. Cecchi, F. Piazza, I. Omicciolo, G. L. Sicuranza, Multiple position room response equalization in frequency domain, *IEEE transactions on audio, speech, and language processing* 20 (1) (2011) 122–135.

- [6] F. E. Toole, S. E. Olive, The modification of timbre by resonances: Perception and measurement, *Journal of the Audio Engineering Society* 36 (3) (1988) 122–142.
- [7] G. Song, P. Davies, Y. Liu, Development of metric models to predict annoyance due to tonal office noise, *Science and Technology for the Built Environment* 28 (8) (2022) 1054–1068.
- [8] G. Song, Annoyance thresholds of tones in noise as related to building services equipment (Dec 2020). doi:10.25394/PGS.13360649.v1.
- [9] N. Broner, Determination of the Relationship Between Low-frequency HVAC Noise and Comfort in Occupied Spaces: Psycho-acoustic Phase, Vipac Engineers & Scientists Limited, 2004.
- [10] J. Lee, J. M. Francis, L. M. Wang, How tonality and loudness of noise relate to annoyance and task performance, *Noise Control Engineering Journal* 65 (2) (2017) 71–82.
- [11] E. Bowden, L. Wang, Relating human productivity and annoyance to indoor noise criteria systems: a low frequency analysis, *ASHRAE Transactions* 111 (1) (2005) 684–692.
- [12] G. Vettori, L. Di Leonardo, S. Secchi, A. Astolfi, L. Bigozzi, Primary school children’s verbal working memory performances in classrooms with different acoustic conditions, *Cognitive Development* 64 (2022) 101256.
- [13] T. Paatero, M. Karjalainen, Kautz filters and generalized frequency resolution: Theory and audio applications, *Journal of the Audio Engineering Society* 51 (1/2) (2003) 27–44.
- [14] M. Karjalainen, T. Paatero, Equalization of loudspeaker and room responses using kautz filters: Direct least squares design, *EURASIP Journal on Advances in Signal Processing* 2007 (2006) 1–13.
- [15] B. D. Kulp, Digital equalization using fourier transform techniques, in: *Audio Engineering Society Convention 85*, Audio Engineering Society, 1988.

- [16] O. Kirkeby, P. A. Nelson, Digital filter design for inversion problems in sound reproduction, *Journal of the Audio Engineering Society* 47 (7/8) (1999) 583–595.
- [17] O. Kirkeby, P. A. Nelson, H. Hamada, F. Orduna-Bustamante, Fast deconvolution of multichannel systems using regularization, *IEEE Transactions on speech and audio processing* 6 (2) (1998) 189–194.
- [18] O. Kirkeby, P. Rubak, A. Farina, Analysis of ill-conditioning of multi-channel deconvolution problems, in: *Proceedings of the 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. WASPAA'99* (Cat. No. 99TH8452), IEEE, 1999, pp. 155–158.
- [19] P. A. Nelson, F. Orduna-Bustamante, H. Hamada, Multi-channel signal processing techniques in the reproduction of sound, in: *Audio Engineering Society Conference: UK 7th Conference: Digital Signal Processing (DSP)*, Audio Engineering Society, 1992.
- [20] O. Kirkeby, P. A. Nelson, F. Orduna-Bustamante, H. Hamada, Local sound field reproduction using digital signal processing, *The Journal of the Acoustical Society of America* 100 (3) (1996) 1584–1593.
- [21] P. M. Clarkson, J. Mourjopoulos, J. Hammond, Spectral, phase, and transient equalization for audio systems, *Journal of the Audio Engineering Society* 33 (3) (1985) 127–132.
- [22] T. Mei, A. Mertins, M. Kallinger, Room impulse response reshaping/shortening based on least mean squares optimization with infinity norm constraint, in: *2009 16th International Conference on Digital Signal Processing*, IEEE, 2009, pp. 1–6.
- [23] M. Kolundžija, C. Faller, M. Vetterli, Multi-channel low-frequency room equalization using perceptually motivated constrained optimization, in: *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Ieee, 2012, pp. 533–536.
- [24] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, B. Ottersten, Convex optimization-based beamforming, *IEEE Signal Processing Magazine* 27 (3) (2010) 62–75.

- [25] S. E. Nai, W. Ser, Z. L. Yu, H. Chen, Beampattern synthesis for linear and planar arrays with antenna selection by convex optimization, *IEEE Transactions on Antennas and Propagation* 58 (12) (2010) 3923–3930.
- [26] Y. Zhuang, Y. Liu, Constrained optimal filter design for multi-channel active noise control via convex optimization, *The Journal of the Acoustical Society of America* 150 (4) (2021) 2888–2899.
- [27] Y. Zhuang, Y. Liu, A numerically stable constrained optimal filter design method for multichannel active noise control using dual conic formulation, *The Journal of the Acoustical Society of America* 152 (4) (2022) 2169–2182.
- [28] T. Shi, Y. Liu, J. S. Bolton, Spatially sparse sound source localization in an under-determined system by using a hybrid compressive sensing method, *The Journal of the Acoustical Society of America* 146 (2) (2019) 1219–1229.
- [29] C. Meng, Z. Ding, S. Dasgupta, A semidefinite programming approach to source localization in wireless sensor networks, *IEEE signal processing letters* 15 (2008) 253–256.
- [30] X. Qu, L. Xie, An efficient convex constrained weighted least squares source localization algorithm based on tdoa measurements, *Signal Processing* 119 (2016) 142–152.
- [31] Y. Zhuang, Y. Liu, A constrained optimal hear-through filter design approach for earphones, in: *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, Vol. 263, Institute of Noise Control Engineering, 2021, pp. 1329–1337.
- [32] S. Bharitkar, C. Kyriakakis, Multirate signal processing for multiple listener low frequency room acoustic equalization, in: *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers*, 2004., Vol. 1, IEEE, 2004, pp. 263–267.
- [33] P. Rubak, L. Johansen, Listening test results from a new digital loudspeaker/room correction systems, in: *Audio Engineering Society Convention* 110, Audio Engineering Society, 2001.

- [34] G. Long, Y. Wang, T. C. Lim, Optimal parametric design of delayless subband active noise control system based on genetic algorithm optimization, *Journal of Vibration and Control* 28 (15-16) (2022) 1950–1961.
- [35] Y. Zhuang, Y. Liu, A stable iir filter design approach for high-order active noise control applications, in: *Acoustics*, Vol. 5, MDPI, 2023, pp. 746–758.
- [36] J. Cheer, S. J. Elliott, Multichannel control systems for the attenuation of interior road noise in vehicles, *Mechanical Systems and Signal Processing* 60 (2015) 753–769.
- [37] M. Grant, S. Boyd, CVX: Matlab software for disciplined convex programming, version 2.1 (Mar. 2014).
- [38] M. Grant, S. Boyd, Graph implementations for nonsmooth convex programs, in: V. Blondel, S. Boyd, H. Kimura (Eds.), *Recent Advances in Learning and Control, Lecture Notes in Control and Information Sciences*, Springer-Verlag Limited, 2008, pp. 95–110.
- [39] S. Boyd, S. P. Boyd, L. Vandenberghe, *Convex optimization*, Cambridge university press, 2004.
- [40] J. F. Sturm, Implementation of interior point methods for mixed semidefinite and second order cone optimization problems, *Optimization methods and software* 17 (6) (2002) 1105–1154.
- [41] J. F. Sturm, Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones, *Optimization methods and software* 11 (1-4) (1999) 625–653.
- [42] R. H. Tütüncü, K.-C. Toh, M. J. Todd, Solving semidefinite-quadratic-linear programs using SDPT3, *Mathematical programming* 95 (2) (2003) 189–217.
- [43] L. E. Kinsler, A. R. Frey, A. B. Crippens, J. V. Sanders, *12. architectural acoustics* (1999).
- [44] P. D. Hatziantoniou, J. N. Mourjopoulos, Generalized fractional-octave smoothing of audio and acoustic responses, *Journal of the Audio Engineering Society* 48 (4) (2000) 259–280.