OPEN ACCESS Freely available online

G. M. De Bortoli, K. Prawda, and S. J. Schlecht,
"Active Acoustics With a Phase Cancelling Modal Reverberator,"
J. Audio Eng. Soc., vol. 72, no. 10, pp. 705–715 (2024 Oct.).
https://doi.org/10.17743/jaes.2022.0171.

Active Acoustics With a Phase Cancelling Modal Reverberator

GIAN MARCO DE BORTOLI,[1,*] **KAROLINA PRAWDA,**[1] **AND**
(gian.debortoli@aalto.fi) (karolina.prawda@aalto.fi)
SEBASTIAN J. SCHLECHT,[1,2] *AES Member*
(sebastian.schlecht@aalto.fi)

[1]*Acoustics Lab, Department of Information and Communications Engineering, Aalto University, Finland*
[2]*Media Lab, Department of Art and Media, Aalto University, Finland*

Active acoustics (AA) systems are used to electronically modify the acoustics of a room (e.g., in live music venues). AA systems have an inherent feedback component and can suffer from instability and coloration artifacts resulting from too high feedback gains. State-of-the-art methods can improve system stability and coloration, usually at the cost of complex implementations and long parameter-tuning sessions. They can also cause sound artifacts due to time-varying components, limiting the enhancement at low frequencies. This work proposes a time-invariant feedback attenuation method for low frequencies based on a modal reverberator. The attenuation is achieved through destructive acoustic interference, obtained via phase shifts between the input and output signals. The analyzed frequency range is 0–500 Hz, where the room transfer functions are considered highly invariant over time. The results show a gain-before-instability increase of more than 5 dB for a modal reverberator with high mode density in this frequency range. The improvement is also stable for low-magnitude changes in the room transfer functions over time. The proposed method provides a robust AA system with artificial reverberation for the low-frequency range and can be used alongside other established methods.

0 INTRODUCTION

Active acoustics (AA) systems are electroacoustic systems that modify the reverberant field in a closed space. They commonly consist of several microphones and loudspeakers, positioned inside a room, and a digital signal processor (DSP) controlling the time evolution of the signal circulating between the transducers. Fig. 1 illustrates a typical AA system with a sound source, audience, microphones, and loudspeakers in the physical room and a DSP unit as the virtual room. The microphones record the signals from the source, the DSP unit processes them, and the loudspeakers emit the processed signals into the physical room.

Due to the regenerative paths from the loudspeakers to the microphones, AA systems are susceptible to instability, which appears as ringing tones increasing in amplitude at each loop iteration up to the point of transducer saturation. As a consequence, stable AA systems that approach instability suffer from coloration artifacts. When time-variation is included in the DSP, coloration can be minimized, but modulation artifacts are introduced [1].

A single-channel AA system is stable if at all frequencies, the feedback loop transfer function (TF) is nonreal or lower than 0 dB in magnitude. The two conditions are called Nyquist's stability criterion [2, 3]. In multichannel systems, an Eigen-decomposition is applied per frequency to the matrix of loop TFs, and Nyquist's stability criterion is verified on each eigenvalue [4]. The maximum gain that can be achieved before the system becomes unstable is called gain-before-instability (GBI) [1].

Numerous DSPs described in the literature aim to enhance reverberation time and/or gain while maintaining the system stable. Usual techniques are broad-band and narrow-band channel equalization [5–8], spectrum decorrelation [1, 9, 10], external artificial reverberation [11–13], and acoustic-feedback cancellation [14–16]. In recent years, geometry-based and perception-based approaches have been suggested [17–19]. Since perception-based metrics are out of the scope of this work, these techniques are not discussed further. A detailed description of the state-of-the-art AA systems is given in SEC. 1.

*To whom correspondence should be addressed, email: gian.debortoli@aalto.fi. Last updated: July 5th, 2024.

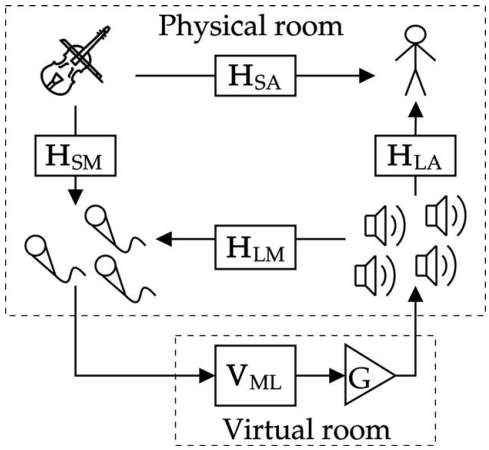


Fig. 1. Block diagram representing a typical AA system. The signal routing is depicted by solid arrows with annotations of involved TFs.

This study proposes a novel method to control the feedback in AA systems through modal synthesis of artificial reverberation and acoustic phase cancellation. This results in a low-complexity DSP implemented as a mixing matrix of modal reverberators. The impulse response (IR) of a modal reverberator is a superposition of characteristic resonances (e.g., room modes) [20, 21]. Each resonance is designed independently, allowing for the control of phase and magnitude responses of the reverb. Acoustic phase cancellation is obtained by destructive interference between sound waves contributing to the same sound field [22]. Tuning the phase of a frequency component for all the loop TFs can cause phase cancellation and suppress the feedback at that frequency.

The article is organized as follows. SEC. 1 describes the state-of-the-art AA techniques. SEC. 2 gives the background on multichannel system stability. SEC. 3 introduces the proposed method describing the mixing-matrix design. SEC. 4 explains the phase-response optimization that leads to phase cancellation. SEC. 5 presents validation results regarding stability improvement. SEC. 6 concludes the study.

1 STATE OF THE ART

In the following, the authors present an overview of the state-of-the-art AA techniques, together with their advantages and limitations. In an AA system, the system gain G is constrained by the GBI (cf. SEC. 2). If $G < GBI$, the system is stable, but G approaching the GBI results in audible artifacts. In time-invariant systems, the common problem is coloration, often referred to as "metallic" reverberation. This is because of the alternating peaks and valleys in the room TFs (RTFs): the feedback loop amplifies some frequencies more than others. For a typical concert-hall RTF, the magnitude response may fluctuate by up to even 10 dB with respect to the mean [9, 23, 24]. Time-varying techniques reduce coloration at the cost of introducing modulation artifacts, resulting in sound detuning and beating effects. To prevent artifacts, a G with a safe margin $\epsilon > 0$ dB from the GBI should be used:

$$G = GBI - \epsilon. \tag{1}$$

In time-invariant algorithms, [9] recommends $\epsilon \geq 2$ dB to avoid ringing effects, and a safer margin $\epsilon = 5$–7 dB is suggested by Svensson [25]. Time-varying algorithms do not have a typical ϵ value as it strongly depends on the technique.

System-stability improvement is a key feature of AA systems performance. A metric commonly used in the literature is the GBI increase or GBI improvement (GBIi) [1]. The GBIi is the GBI difference between a proposed DSP and a reference, with the system setup being equal. Typically, the reference is a unitary matrix $\mathbf{U}$, which is reduced to a unitary connection in a single-channel system. For the proposed method, the reference will be the GBI obtained without phase cancellation.

Channel equalization includes broad-band gain control and notch-filter–based methods. Broad-band gain control lowers the loudspeaker gains when a strong oscillation in the feedback loop is detected [6, 7]. It does not increase the GBI since the spectral shape of the loop TF is not altered. Notch-filter–based methods employ notch filters to suppress the most prominent feedback-loop peaks and can reach GBIi values up to 10 dB [5, 26]. Time-invariant implementations, however, require time-consuming manual tuning [8]. Signal-adaptive techniques can be complex because they must be able to distinguish between rising ringing tones and musical content [27].

Spectral decorrelation techniques aim at smoothing the loop magnitude response by averaging the RTF's peaks and valleys through either delay modulation, phase modulation, or frequency shift of successive iterations of the feedback loop [1, 27]. Delay modulation has a simple implementation and can reach a GBIi of 4–8 dB, but it does not offer improvement below 500–700 Hz. Phase modulation provides a lower GBIi than delay modulation, but it is more consistent at low frequencies—at the expense of frequency and amplitude modulation artifacts. Frequency shift provides the best stability improvement in this category. In a single-channel system, it can completely smooth the feedback-loop magnitude response and reach a GBIi of 10–12 dB. But it is susceptible to modulation artifacts at high gains, requiring $\epsilon = 6$ dB [1, 9, 27] and the improvement decreases with increasing number of channels [28]. A time-varying mixing matrix is a suitable choice for multichannel systems. With a minimum of 16 channels, it reaches a GBIi of 10–12 dB with better perceptual results than frequency shift [10].

External-reverberation–based algorithms aim to control the reverberation time, sometimes at the cost of a negative GBIi. With proper design using all-pass reverberators and a unitary mixing matrix, the GBIi can be zero [11]. Feedback delay networks are a suitable choice for time-varying reverberators. This technique can achieve between 2 and 7 dB in GBIi already with a low number of channels and a low feedback-delay-network order [12, 29, 13].

Acoustic-feedback cancellation predicts the feedback signals from the loudspeakers to the microphones and subtracts them from the microphone signals [15]. Ideally, it can completely suppress the feedback signals, granting an unlimited GBIi. Adaptive algorithms usually require complex implementation and high computational costs. Thus, adaptive cancellation is challenging for real-time, multiple-input multiple-output systems with many channels [27].

2 PROBLEM FORMULATION

Fig. 1 depicts a multichannel AA system and the interaction between its elements: The *physical room* is the closed space in which the transducers are placed, and the DSP component is the *virtual room*. In the physical room, there is one natural sound source, $n_{\mathrm{S}} = 1$, a grid of n_{A} space-sampling points representing a live audience, n_{M} microphones, and n_{L} loudspeakers.

The physical room has RTFs from the natural source to the audience points, $\mathbf{H}_{\mathrm{SA}}(e^{j\omega_k}) \in \mathbb{C}^{n_{\mathrm{A}}\times n_{\mathrm{S}}}$; from the natural source to the microphones, $\mathbf{H}_{\mathrm{SM}}(e^{j\omega_k}) \in \mathbb{C}^{n_{\mathrm{M}}\times n_{\mathrm{S}}}$; from the loudspeakers to the microphones, responsible for the feedback, $\mathbf{H}_{\mathrm{LM}}(e^{j\omega_k}) \in \mathbb{C}^{n_{\mathrm{M}}\times n_{\mathrm{L}}}$; and from the loudspeakers to the audience points, $\mathbf{H}_{\mathrm{LA}}(e^{j\omega_k}) \in \mathbb{C}^{n_{\mathrm{A}}\times n_{\mathrm{L}}}$. ω_k is the discrete frequency variable expressed in radian per second, j is the imaginary unit, and e is Euler's number. The virtual room includes a mixing matrix containing the feedforward TFs from microphones to loudspeakers—$\mathbf{V}_{\mathrm{ML}}(e^{j\omega_k}) \in \mathbb{C}^{n_{\mathrm{L}}\times n_{\mathrm{M}}}$—and a real scalar gain G.

AA-systems stability is determined by the product of the feedforward TFs and feedback RTFs, which results in the matrix of the loop TFs:

$$\mathbf{F}_{\mathrm{MM}}(e^{j\omega_k}) = \mathrm{G}\,\mathbf{H}_{\mathrm{LM}}(e^{j\omega_k})\mathbf{V}_{\mathrm{ML}}(e^{j\omega_k}), \tag{2}$$

containing all the TFs from any microphone to any microphone.

All elements of $\mathbf{F}_{\mathrm{MM}}(e^{j\omega_k})$ are typically nonzero, meaning all the channels are coupled. Applying Eigendecomposition to $\mathbf{F}_{\mathrm{MM}}(e^{j\omega_k})$ at each frequency ω_k provides a new formulation of the full system as a superposition of decoupled single-channel systems, that is,

$$\mathbf{F}_{\mathrm{MM}}(e^{j\omega_k}) = \mathrm{G}\,\mathbf{Q}(e^{j\omega_k})\Lambda(e^{j\omega_k})\mathbf{Q}^{-1}(e^{j\omega_k}), \tag{3}$$

where $\mathbf{Q}(e^{j\omega_k})$ are the eigenvectors and $\Lambda(e^{j\omega_k})$ is a diagonal matrix containing the system's eigenvalues $\{\lambda_i(e^{j\omega_k})\}$, for $1 \le i \le n_{\mathrm{M}}$. Each eigenvalue fully describes the behavior of the corresponding decoupled single-channel system as a function of frequency.

The overall system is stable if all the single-channel systems are stable. Recalling Nyquist's stability criterion and using the conservative magnitude approach [10, 30], a single-channel system is stable if its eigenvalue magnitude is less than 0 dB at all frequencies.

The maximum eigenvalue of the full system at frequency ω_k is

$$\lambda_{\max}(e^{j\omega_k}) = \max_{1\le i\le n_{\mathrm{M}}} \left\{ \left| \lambda_i(e^{j\omega_k}) \right| \right\}, \tag{4}$$

and the maximum eigenvalue over frequency is

$$\hat{\lambda}_{\max} = \max_{\omega_k} \left\{ \lambda_{\max}(e^{j\omega_k}) \right\}. \tag{5}$$

$\hat{\lambda}_{\max}$ sets the system's GBI value (i.e., the maximum stable G) as

$$\mathrm{GBI} = \frac{1}{\hat{\lambda}_{\max}}. \tag{6}$$

It is important to underline that even though the matrix product is not commutative, the eigenvalues are independent of the order of the factors in Eq. (2).

In a concert-hall scenario, the audio quality of the enhanced reverberant field is of primary importance. Thus, the artificial reverberation of commercial AA systems must blend with the natural acoustics of the room in which they are deployed. In this work, the authors propose modal synthesis to tailor the reverberation as required from the physical room and phase cancellation to raise the GBI and obtain higher amplification.

3 PROPOSED METHOD

To maximize the GBIi, the proposed method minimizes the magnitude of the system eigenvalues without changing the reverb energy. This is achieved by controlling the phase of the feedforward TFs and causing phase cancellation in the loop TFs. Phase cancellation pairs well with modal synthesis for virtual reverberation since the phase content of each characteristic resonance can be controlled individually.

The proposed method requires prior knowledge of the $\mathbf{H}_{\mathrm{LM}}$. Furthermore, the frequency band is restricted to 0–500 Hz, where RTFs are more stable toward small changes in transducer positions, atmospheric changes, or room occupancy, and thus, they are assumed to be time-invariant [31, 32].

3.1 Modal Synthesis

The building block of the virtual room is the resonance filter. The TF of a real second-order resonance filter can be expressed, through partial fraction expansion, as a sum of two complex first-order resonators

$$H(z) = \frac{b(z)}{a(z)} = \frac{r}{1 - pz^{-1}} + \frac{\overline{r}}{1 - \overline{p}z^{-1}}, \tag{7}$$

where z is a complex number; a and b are the denominator and the numerator, respectively, of the rational TF $H(z)$; r is the residue and p is the pole of the complex resonator; and the overline indicates the complex conjugate operator. The residue r contains information on the initial phase θ and the amplitude α of the mode, while the pole p is characterized by the resonance frequency ω_{r} and the decay rate β:

$$r = \alpha e^{j\theta} \qquad p = \beta e^{j\psi}, \tag{8}$$

and $\psi = \omega_r/f_{\mathrm{s}}$ with $0 \le \psi \le \pi$ being the angle of the pole on the complex plane, defined as the resonance frequency ω_r in radian per second normalized over the time-sampling rate f_{s} in hertz.

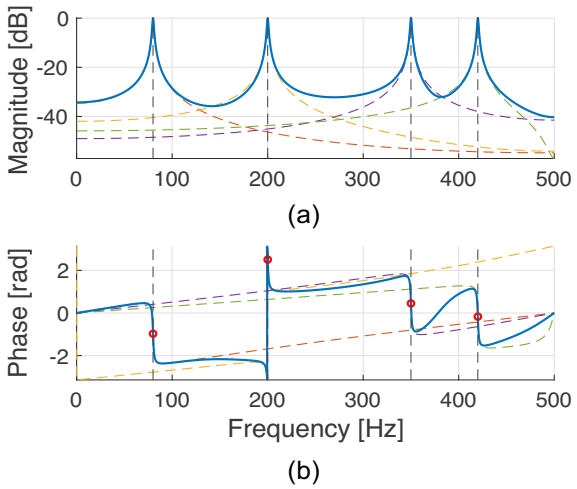


Fig. 2. Magnitude response (a) and phase response (b) of a modal reverberator obtained by superposing four different resonance filters.

The virtual modal reverberator is obtained as a parallel structure of n_R resonance filters [20], one for each mode,

$$V(z) = \sum_{q=1}^{n_\mathrm{R}} H_q(z), \tag{9}$$

where $H_q(z)$ is the TF of the qth resonance filter [see Eq. (7)]. An example of the magnitude and phase responses of a modal reverberator containing four modes is depicted in Fig. 2.

The optimization on the phase θ is performed only for the positive frequencies [i.e., only on the first addendum of Eq. (7), $\tilde{H}(z) = \frac{r}{1-pz^{-1}}$]. The negative frequencies are then reconstructed with the following property: $h(t) = 2\,\mathrm{Re}\{\tilde{h}(t)\}$, where $h(t)$ is the IR of $H(z)$ and $\tilde{h}(t)$ is the IR of $\tilde{H}(z)$, obtaining a real-valued signal.

3.2 Mixing Matrix

The mixing matrix shapes the spectra of the loudspeaker input signals from the microphone output signals. Each element *(l,m)* of the mixing matrix, which is the feedforward TF from the mth microphone to the lth loudspeaker, is a modal reverberator:

$$V_{\mathrm{l,m}}(z) = \sum_{q=1}^{n_\mathrm{R}} \tilde{H}_q^{l,m}(z), \tag{10}$$

where

$$\tilde{H}_q^{l,m}(z) = \frac{r_q^{l,m}}{1 - p_q^{l,m} z^{-1}} \qquad r_q^{l,m} = \alpha_q^{l,m} e^{j\theta_q^{l,m}} \qquad p_q^{l,m} = \beta_q^{l,m} e^{j\psi_q^{l,m}}. \tag{11}$$

To maximize phase cancellation and to study only its contribution to the system stability, the elements of the mixing matrix have the same magnitude peaks, decay rates, and resonance frequencies:

$$\alpha_q^{l,m} = \alpha_q \quad \forall\,(l,m) \qquad \beta_q^{l,m} = \beta_q \quad \forall\,(l,m) \qquad \psi_q^{l,m} = \psi_q \quad \forall\,(l,m), \tag{12}$$

but their phases $\theta_q^{l,m}$ are different. The numerator and denominator of $\tilde{H}_q^{l,m}(z)$ are separated:

$$\tilde{H}_q^{l,m}(z) = \frac{\alpha_q e^{j\theta_q^{l,m}}}{1 - p_q z^{-1}} = e^{j\theta_q^{l,m}} \frac{\alpha_q}{1 - p_q z^{-1}}. \tag{13}$$

A different phase is then assigned to each transducer. Thus, the numerator is further divided into two complex exponentials:

$$e^{j\theta_q^{l,m}} = e^{j(\theta_q^m + \theta_q^l)} = e^{j\theta_q^m} e^{j\theta_q^l}. \tag{14}$$

This design reduces the number of degrees of freedom of the optimization process. The dimensionality of the problem scales linearly with the number of transducers instead of quadratically—as in the case of independent phases for each element of $\mathbf{V}_\mathrm{ML}$—which is convenient in case of a high number of channels regarding optimization and computational speed.

The resonance filter described in Eqs. (13) and (14) comprises a series of three filters. One is a complex first-order resonance filter with zero phase at the resonance. The other two are complex exponentials only applying a phase shift of θ_q^m and θ_q^l, respectively. From now on, the resonance filter will be called the magnitude filter, and the complex exponentials will be called phase filters.

Once resonance frequencies, decay rates, and phase values are set for each mode in each element of the mixing matrix, the DSP is normalized. Because α_q affects the norm of the matrix $\mathbf{V}_\mathrm{ML}$, it is set to 0 dB, and the mixing matrix is scaled by its inverse Frobenius norm:

$$\left(\mathbf{V}_\mathrm{ML}(e^{j\omega_k})\right)_\mathrm{normalized} = \frac{\mathbf{V}_\mathrm{ML}(e^{j\omega_k})}{\|\mathbf{V}_\mathrm{ML}\|_\mathrm{F}}. \tag{15}$$

After Eq. (15), the Frobenius norm of the mixing matrix is equal to one. Recalling *Parseval's theorem*, $\|\mathbf{V}_\mathrm{ML}\|_\mathrm{F}$ is computed in the frequency domain as

$$\|\mathbf{V}_\mathrm{ML}\|_\mathrm{F} = \sqrt{\sum_{l=1}^{n_\mathrm{L}} \sum_{m=1}^{n_\mathrm{M}} \frac{1}{2\pi} \sum_{k=1}^{N_\mathrm{b}} \left| V_{\mathrm{l,m}}(e^{j\omega_k}) \right|^2}, \tag{16}$$

where N_b is the number of frequency bins. It is possible to equalize the modal reverb by controlling α_q of each resonance filter individually. The normalization process remains unchanged. However, equalization is outside the scope of this work as it would affect system stability.

The last step in the mixing matrix design process is to reconstruct the negative frequencies of the impulse response to generate real loudspeaker signals:

$$\mathbf{V}_\mathrm{ML}[n] = 2\,\mathrm{Re}\left\{\mathrm{DFT}^{-1}\left\{\left(\mathbf{V}_\mathrm{ML}(e^{j\omega_k})\right)_\mathrm{normalized}\right\}\right\}, \tag{17}$$

where $\mathbf{V}_\mathrm{ML}[n]$ is the time-domain formulation of the mixing matrix and DFT^{-1} is the inverse discrete Fourier transform.

A block diagram of the full DSP architecture for a system setup of two microphones, two loudspeakers, and two modes for the modal reverb is depicted in Fig. 3. The filters are separated into three stages:

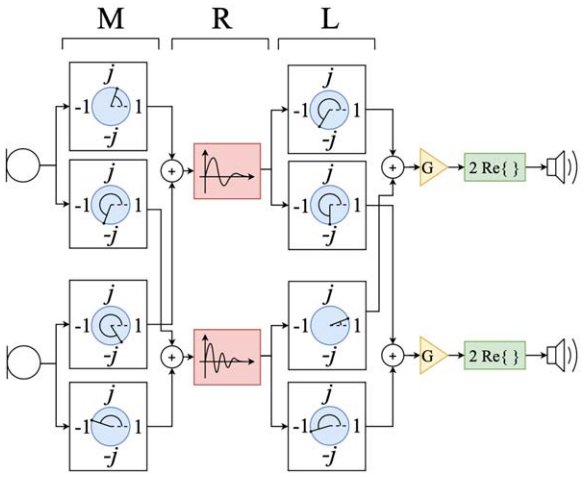


Fig. 3. Block diagram representing the full processing path from microphones to loudspeakers for a system comprising two microphones, two loudspeakers, and a modal reverberator containing two modes.

- $\mathbf{M}(e^{j\omega_k}) \in \mathbb{C}^{n_\mathbf{R} \times n_\mathbf{M}}$ is the matrix containing all the phase filters preceding the magnitude filters.
- $\mathbf{R}(e^{j\omega_k}) \in \mathbb{C}^{n_\mathbf{R} \times n_\mathbf{R}}$ is the diagonal matrix containing all the magnitude filters.
- $\mathbf{L}(e^{j\omega_k}) \in \mathbb{C}^{n_\mathbf{L} \times n_\mathbf{R}}$ is the matrix containing all the phase filters following the magnitude filters.

Neglecting the real-part operator, the mixing matrix can be rewritten as follows:

$$\mathbf{V}_{\mathrm{ML}}(e^{j\omega_k}) = \mathbf{L}(e^{j\omega_k})\mathbf{R}(e^{j\omega_k})\mathbf{M}(e^{j\omega_k}), \tag{18}$$

and the loop-TF matrix becomes

$$\mathbf{F}_{\mathrm{MM}}(e^{j\omega_k}) = \mathrm{G}\,\mathbf{H}_{\mathrm{LM}}(e^{j\omega_k})\mathbf{L}(e^{j\omega_k})\mathbf{R}(e^{j\omega_k})\mathbf{M}(e^{j\omega_k}). \tag{19}$$

4 PHASE OPTIMIZATION

According to Eq. (6), one needs to minimize $\hat{\lambda}_{\max}$ to maximize the GBI. Obtaining sufficient phase cancellation requires multiple channels and a high frequency resolution. Those conditions, in turn, increase the size of $\mathbf{F}_{\mathrm{MM}}$ and force heavier usage of the Eigen-decomposition, which may lead to a surge in the computational cost of the phase optimization, regardless of the modal-reverberator design.

For a full-rank matrix $\mathbf{A} \in \mathbb{C}^{n \times m}$

$$\mathrm{rank}\big(\mathbf{A}\big) = \min\{n, m\}. \tag{20}$$

Since for any two matrices $\mathbf{A}$ and $\mathbf{B}$,

$$\mathrm{rank}(\mathbf{AB}) = \min\{\mathrm{rank}(\mathbf{A}), \mathrm{rank}(\mathbf{B})\}, \tag{21}$$

the rank of the mixing matrix is

$$\mathrm{rank}\big(\mathbf{V}_{\mathrm{ML}}\big) = \min\{n_\mathrm{L}, n_\mathrm{M}, n_\mathrm{R}\}, \tag{22}$$

and the number of eigenvalues in the system is

$$\mathrm{rank}\big(\mathbf{F}_{\mathrm{MM}}\big) = \min\{n_\mathrm{L}, n_\mathrm{M}, n_\mathrm{R}\}. \tag{23}$$

However, the eigenvalues of a matrix product—as specified in SEC. 2—are independent of the order of the factors. The loop-TF matrix $\mathbf{T}_{\mathrm{RR}}(e^{j\omega_k}) \in \mathbb{C}^{n_\mathrm{R} \times n_\mathrm{R}}$ can be computed as

$$\mathbf{T}_{\mathrm{RR}}(e^{j\omega_k}) = \mathrm{G}\,\mathbf{M}(e^{j\omega_k})\mathbf{H}_{\mathrm{LM}}(e^{j\omega_k})\mathbf{L}(e^{j\omega_k})\mathbf{R}(e^{j\omega_k}), \tag{24}$$

retaining the eigenvalues of $\mathbf{F}_{\mathrm{MM}}$.

Considering a multiple-mode modal reverb, each magnitude filter has a high magnitude around its resonance and a low magnitude elsewhere. When two modes are far enough from each other, their reciprocal influence is low. In such a case, $\lambda_{\max}(e^{j\omega_k})$ can be approximated by only the mode with the highest magnitude in (ω_k). This work considers a reverberation time $T_{60} = 1.4$ s for all modes, with a resonance bandwidth of approximately 1.4 Hz.

A single-mode reverberator, $\mathrm{rank}\big(\mathbf{R}(e^{j\omega_k})\big) = 1$, is first considered, such that $\mathrm{rank}\big(\mathbf{T}_{\mathrm{RR}}(e^{j\omega_k})\big) = 1$. Thus, there is only one eigenvalue. Since $n_\mathrm{R} = 1$, $\mathbf{M}(e^{j\omega_k})$ is a row-vector, $\mathbf{L}(e^{j\omega_k})$ is a column-vector, and $\mathbf{R}(e^{j\omega_k})$ is a single complex element. Due to the order of the terms in Eq. (24), $\mathbf{T}_{\mathrm{RR}}(e^{j\omega_k}) = T_{\mathrm{RR}}(e^{j\omega_k})$ is a 1-by-1 matrix. From Eq. (3), $\Lambda(e^{j\omega_k}) = T_{\mathrm{RR}}(e^{j\omega_k})$ and the magnitude of the eigenvalue is $\big|\lambda(e^{j\omega_k})\big| = \big|T_{\mathrm{RR}}(e^{j\omega_k})\big|$. The computation of Eq. (24) is fast, and it directly provides the eigenvalue that is needed to minimize, hence bypassing the Eigen-decomposition.

In the proposed method, the system is a superposition of single-mode reverberators, which can be studied independently. Due to the magnitude fluctuations in the RTFs and the nonzero bandwidth of the magnitude filters, obtaining phase cancellation at ψ_q might lead to a lower magnitude value at ψ_q than at the neighboring frequencies. Therefore, the authors study the magnitude of resonance filters not only at the exact values of ψ_q but also in a surrounding frequency range $[\omega_1, \omega_2]_q$ chosen as

$$\begin{aligned} \big|T_{\mathrm{RR}}(e^{j\omega_1})\big| = \big|T_{\mathrm{RR}}(e^{j\omega_2})\big| = \max_\omega\Big\{\big|T_{\mathrm{RR}}(e^{j\omega_k})\big|\Big\} - \delta \\ \text{with} \quad \omega_1 < \psi_q \wedge \omega_2 > \psi_q, \end{aligned} \tag{25}$$

where δ is the difference between the maximum and minimum magnitude values of $\lambda_{\max}(e^{j\omega_k})$, from Eq. (4), when substituting the DSP with a unitary matrix.

The optimal set of phase shifts of the phase filters in $\mathbf{M}(e^{j\omega_k})$ and $\mathbf{L}(e^{j\omega_k})$ for the qth mode is

$$\begin{aligned} &(\theta_q^l, \theta_q^m)^\star = \arg\min_{\theta_q^l, \theta_q^m}\left(\max_{\omega_k}\Big\{\big|T_{\mathrm{RR}}(e^{j\omega_k})\big|\Big\}\right) \\ &\text{where} \quad \omega_k \in [\omega_1, \omega_2]_q. \end{aligned} \tag{26}$$

The optimization algorithm is a Quasi-Newton method receiving a set of randomized phase shifts as initial conditions. Because the algorithm is likely to find a local minimum, the choice of the initial condition affects the optimization results. The optimization process success and speed are also affected by the cardinality of the set (θ_q^l, θ_q^m), which grows with the number of transducers, increasing the chance of incurring a local minimum.

5 EVALUATION

In the following, the authors present the evaluation of the proposed method in a real room. They assess the per-

Table 1. Characteristics for the measured room. Dimensions correspond to width, length, and height, respectively, and the T_{60} values are reported for octave frequency bands with center frequencies f_c.

Dimensions (m)	$4.5 \times 6.1 \times 3.8$					
Volume (m^3)	104.3					
f_c (Hz)	125	250	500	1,000	2,000	4,000
T_{60} (s)	0.22	0.20	0.21	0.22	0.20	0.18

formance of their algorithm in terms of GBIi, resolution of ringing tones and instability, and spatial energy distribution.

5.1 Physical-Room Characteristics

In the evaluation, the authors considered both measured and simulated feedback IRs. In simulations, the IRs were obtained using the image-source method [33]. Table 1 shows the dimensions and the T_{60} values of the room used for the IR modeling.

The equipment used for the measurements included two Genelec 1030A loudspeakers, two Genelec 1029A loudspeakers, four G.R.A.S. 0.5-in free-field microphones of type 46AF, a Presonus Digimax FS eight-channel microphone preamplifier, a G.R.A.S. power module of type 12AG, and a Motu 16A audio interface. To increase the GBI, [25] suggests directional microphones, positioned close to the sound source, and directional loudspeakers, positioned close to the audience. Due to the small volume of the room selected for this work, the authors did not follow such guidelines.

The measurement signal was a 5-s–long exponential sine sweep. The signal was played once for each loudspeaker and recorded simultaneously by all microphones. The procedure was repeated for each loudspeaker, obtaining a set of 16 recordings, which were then deconvolved to obtain the IRs [34]. Measured IRs achieved an average signal-to-noise ratio of 50 dB [35], where the noise power was computed as the scaled median of samples' energy to achieve a more stable estimation [36].

Since the method is used in the 0–500 Hz frequency region, the IRs were downsampled from 44.1 to 1 kHz. Such a narrow frequency interval allowed for an extremely high frequency resolution: $N_b = 2^{13}$ frequency bins used for the DFT computation resulted in frequency spacing of 0.061 Hz between adjacent bins.

Fig. 4 shows the placement of the transducers, the natural source, and one audience position in the room. Each transducer has been assigned a number from 1 to 4, and their coordinates in space are shown in an $[x, y, z]$ format. All the IRs between transducers were measured while all the IRs related to natural source and audience were simulated. Fig. 5(a) displays the RTF magnitude of the loudspeaker 1 and microphones 1 to 4. Fig. 5(b) presents the RTF magnitude of the natural source and microphones 1 to 4.

5.2 Gain-Before-Instability Improvement

In this section, the authors study the effect of phase optimization on the system GBI. The loop iterations between

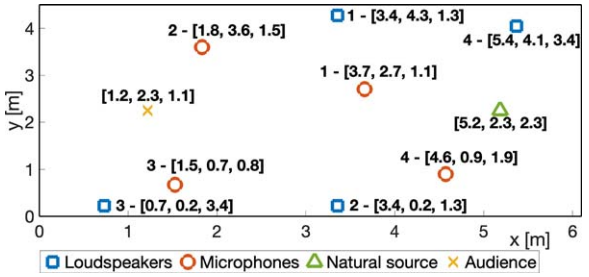
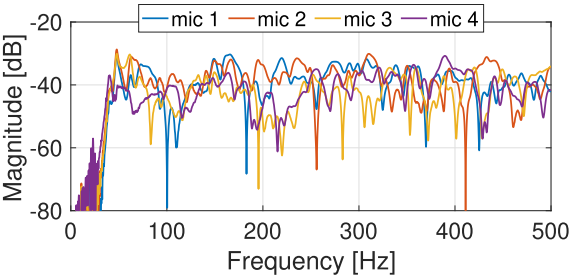
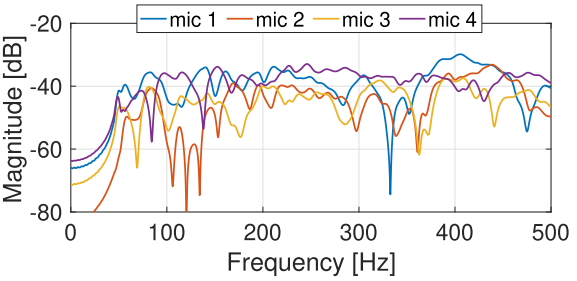


Fig. 4. Top view of the analyzed physical room showing the placement of transducers, natural source, and audience position: blue squares for the loudspeakers, red circles for the microphones, green triangle for the natural source, and yellow cross for the audience.



(a) RTFs magnitude from loudspeaker 1 to all microphones.



(b) RTFs magnitude from natural source to all microphones.

Fig. 5. RTFs magnitude responses measured in the physical room considering the placement depicted in Fig. 4. (a) RTFs magnitude from loudspeaker 1 to all microphones. (b) RTFs magnitude from natural source to all microphones.

the transducers will give coloration to the reverb due to the irregular shape of the RTF magnitude, enhancing frequencies coinciding with the modes and attenuating frequencies corresponding to the nodes. Attenuated frequencies are inherently more stable than enhanced ones. To disentangle the algorithm's performance from the positioning of the resonances along the frequency axis, results are provided both in terms of single-mode and multiple-mode attenuation.

The authors analyzed 200 modes linearly spaced in frequency between 50 and 450 Hz, with a frequency spacing of 2 Hz. At each mode, the phase filters were assigned randomized phase shifts as initial conditions for the Quasi-Newton method. To account for the optimization result variability due to the incurrence of local minima, the process was run 15 times for each mode.

First, a reverberator with only one mode was considered. The GBI was computed according to Eqs. (26), (24), and

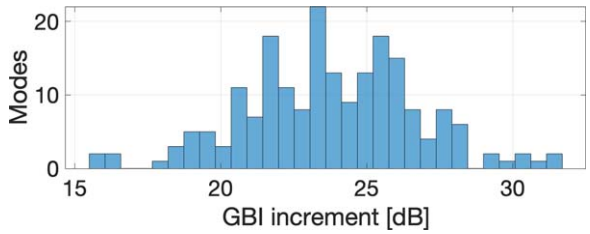


Fig. 6. Distribution of GBIi obtained through optimization of 200 single-mode reverberators with different frequencies linearly spaced from 50 to 450 Hz.

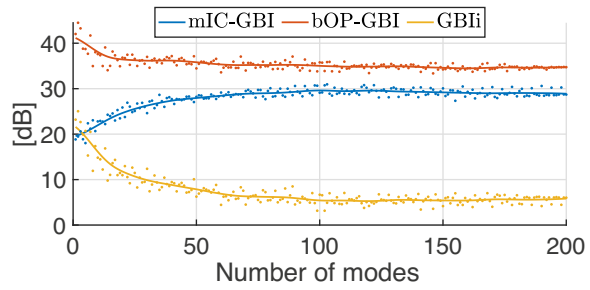


Fig. 7. GBI obtained with nonoptimized (blue-middle) and optimized mixing matrix (red-top) and GBIi (yellow-bottom). The dots are the real computed values; the solid lines are obtained by smoothing the dotted data.

(6), with both the initial conditions and the optimized phase shifts at each optimization run. Out of all the trials, two sets of phase values were kept: the one that provided median GBI between the initial-condition sets (mIC) and the one that resulted in the best GBI between the optimized-phase sets (bOP). Two mixing matrices were then composed for each mode using mIC and bOP in the phase filters, respectively. Both mixing matrices were normalized according to Eq. (15), and the corresponding GBI values were computed according to Eqs. (2)–(6). The GBIi was obtained by subtracting the mIC-GBI from the bOP-GBI. Fig. 6 shows the distribution of the GBIi for all the analyzed modes. Most GBIi values lie between 20 and 28 dB, with some outliers below 20 dB and around 30 dB.

Then, the authors studied how bOP-GBI, mIC-GBI, and GBIi change with multiple modes in the mixing matrix. Multiple subsets of the previously analyzed modes were considered. Each subset contained between 2 and 200 modes linearly spaced in frequency between 50 and 450 Hz. bOP-GBI, mIC-GBI, and GBIi were obtained in the same manner as in the single-mode case.

The results in Fig. 7 show the GBI and GBIi values obtained from each subset of modes as red-top dots for the bOP-GBI, blue-middle dots for the mIC-GBI, and yellow-bottom dots for the GBIi. The solid lines are obtained by smoothing the corresponding scattered data with a 30-datapoints–long Hanning window.

Fig. 7 indicates that the number of modes used in the reverberator affects the GBI. bOP-GBI decreases at first but stabilizes for more than 20 modes. Conversely, mIC-GBI significantly increases with the growing number of modes and saturates for more than 70 modes. Those tendencies have a strong effect on GBIi, which initially declines and converges around 5.5 dB for more than 100 modes.

The authors attribute the change in GBIi over the number of modes to the following factors:

- A system with multiple modes performs as well as its worst-performing mode.
- When multiple modes are present, if they are close to each other in frequency, their magnitudes overlap. The assumption that system stability in a frequency bin is only approximated by considering the mode with the highest magnitude is invalid. Some GBIi obtained by optimizing the modes individually is lost due to interaction between the modes. The higher the density of the modes along the frequency axis, the more prominent this effect becomes.
- Due to normalization, the higher the number of modes considered, the lower the energy in each of them. For this reason, the system designed with mIC becomes more stable with an increasing number of modes.

In a real-world scenario, the number of modes must be high enough to generate a perceptually natural modal reverberator in the considered frequency band. Although recommendations for the number of modes exist for broad-band and narrow-band signals [21], they are not readily applicable to this method. Therefore, it was found empirically through informal listening tests that $n_R \geq 140$ is sufficient for the frequency band 50–450 Hz.

The GBI data from Fig. 7 was compared with Poletti's Unitary Allpass Reverberator [11] (UAR)—as an alternative method for time-invariant artificial-reverberation–based AA systems; a mixing matrix with exponentially decaying white Gaussian noise (WGN) with $T_{60} = 1.4$ s in each element; and a random unitary matrix (RUM) as reference. The authors considered the bOP-GBI and mIC-GBI values with $140 \leq n_R \leq 200$ for the proposed method. To consider the result variance due to the stochastic components of the UAR, WGN, and RUM, the authors generated 15 instances of each of them. They also band-passed the resulting signals—for the UAR and WGN—between 50 and 450 Hz and normalized the matrices according to Eq. (15). The GBI was then computed for all 15 instances with the same procedure as applied to the proposed method.

Fig. 8 shows the GBI comparison for the analyzed techniques. The results of mIC-GBI are similar to the UAR and WGN. bOP-GBI achieves a higher GBI for all considered mode numbers and a GBI distribution narrower than the UAR and WGN. In the UAR IRs, most of the energy is concentrated in a direct-path component, and the energy contribution from early reflections and late reverberation is low. Regarding the direct-to-reverberant ratio, the UAR can be considered a middle point between the WGN and RUM. The proposed mixing matrix has IRs with an exponential decay and no strong direct component, similar to the WGN.

Finally, the authors studied the effect of RTF time variation on GBI and GBIi. They used three sets of measured

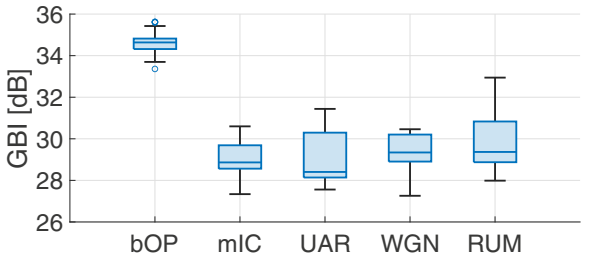
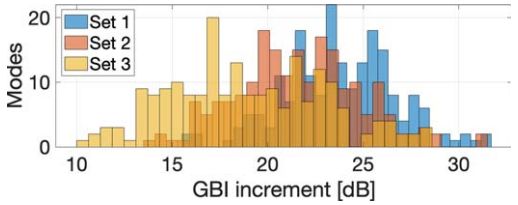
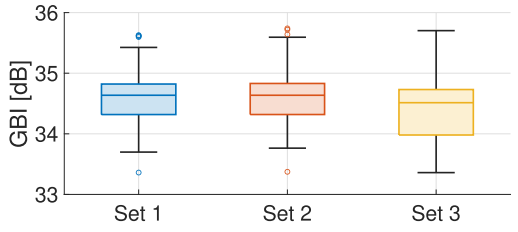


Fig. 8. GBI comparison between the presented method and three alternative methods: UAR, WGN, and RUM.

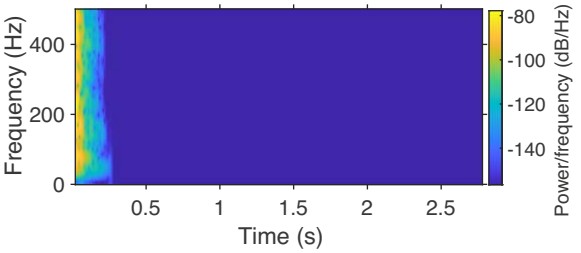


(a) Single-mode reverberator.

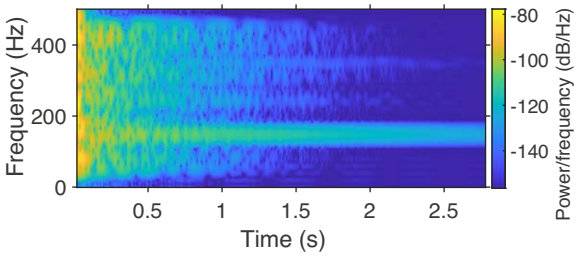


(b) Multi-mode reverberator.

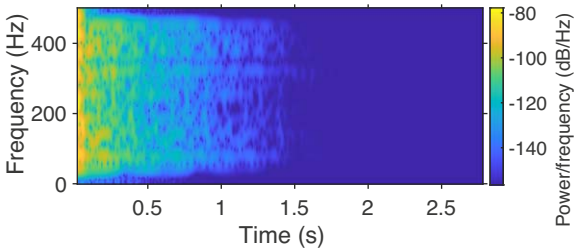
Fig. 9. Time evolution of (a) GBIi distribution and (b) GBI distribution when the RTF time variance is considered.



(a) No modal reverberation.



(b) Non-optimized modal reverberation.



(c) Optimized modal reverberation.

Fig. 10. Spectrograms of the IR from the natural source to the audience position with no modal reverberation (a), nonoptimized modal reverberation (b), and optimized modal reverberation (c). The DSPs have the same amplifier G and the target T_{60} is 1.4 s.

IRs for this purpose: *Set 1* is described in SEC. 5.1; *Sets 2* and *3* were captured in the same room, using the same equipment and procedure, but with 30-min breaks between the last measurement of one set and the first measurement of the following one. Room occupancy and transducer positioning were the same for all sets. The authors used the same sets of phases, bOP and mIC, to repeat computations from Eqs. (2)–(6) for *Sets 2* and *3*.

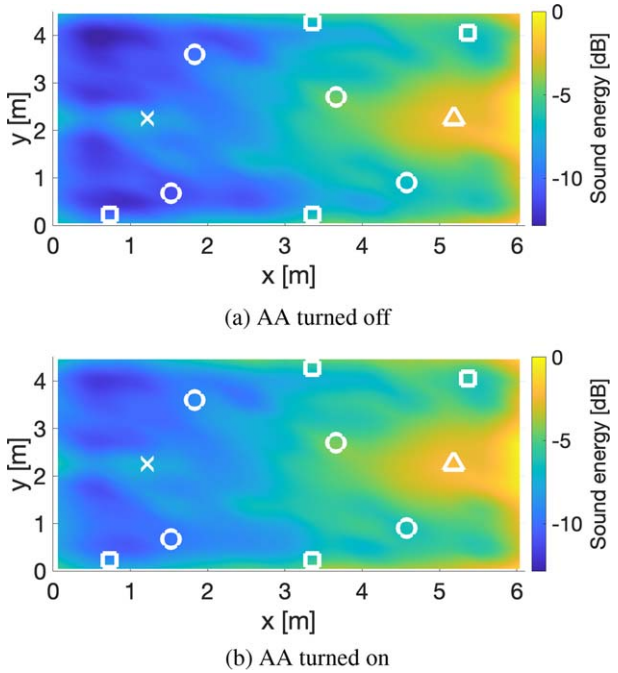
Fig. 9(a) shows the GBIi distribution for all the 200 analyzed modes from *Sets 1*, *2*, and *3* (blue-right, red-middle, and yellow-left bars, respectively). All three distributions have a Gaussian shape, although the distributions become wider from *Set 1* to *Set 3*. The mean GBIi decreases for the successive sets. However, Fig. 9(a) shows that the system is robust toward small time variance because most of the modes have a GBIi >12.5 dB for each measurement set. Fig. 9(b) shows the bOP-GBI range of values variation from *Set 1* to *Sets 2* and *3*. There is no significant difference in GBI distribution between *Sets 1* and *2*. *Set 3* has a wider range of values and a lower median value, but it presents a high degree of overlap with the other two sets.

5.3 System Stability

To visualize the improvement in system stability, the authors computed the IR from the natural source to the audience position with a nonoptimized DSP and an optimized DSP. For both, they considered $n_R = 148$, which led to GBIi $= 5.54$ dB, and a feedforward gain G $=$ GBI$_{mIC} - 1$ dB, where GBI$_{mIC}$ is the GBI value obtained with 148 modes and mIC. The target T_{60} was 1.4 s.

Fig. 10 shows the spectrograms of the IR in three system configurations: no AA [Fig. 10(a)], nonoptimized modal reverberation [Fig. 10(b)], and optimized modal reverberation [Fig. 10(c)]. Both AA settings significantly extend the reverberation compared to the quickly decaying physical IR.

The chosen value of G brings the nonoptimized system close to instability, leading to a strong ringing tone at 180 Hz and a longer-than-intended overall reverberation. The optimized system, instead, is far from instability, and no ringing tone is visible. The optimized system also shows a constant decay rate in frequency, meaning that all modes have similar reverberation times, which is close to the target



(a) AA turned off

(b) AA turned on

Fig. 11. Sound-field-energy space distribution along the plane parallel to the ground. (a) AA turned off. (b) AA turned on.

value. Sound samples of the displayed spectrograms and more audio examples are available online.[1]

5.4 Spatial Reverberation Distribution

In this section, the authors study the spatial distribution of the sound energy obtained with the optimized AA system from SEC. 5.3. They considered a plane parallel to the ground at $z = 1.1$ m. They simulated the IRs from natural sources and loudspeakers to $n_\mathrm{A} = 143$ evenly distributed observation points on the plane. Then, the authors computed the IR from the natural source to each observation point with the AA system both turned off and on. They retrieved the sound-energy matrix $\mathbf{E} \in \mathbb{R}^{n_\mathrm{A}}$ for the n_A observation points as

$$\mathbf{E} = \frac{1}{2\pi} \sum_{k=1}^{N_\mathrm{b}} \left| \mathbf{H}_\mathrm{SA}(e^{j\omega_k}) + \mathbf{H}_\mathrm{AA}(e^{j\omega_k}) \right|^2, \tag{27}$$

where $\mathbf{H}_\mathrm{AA}(e^{j\omega_k})$ is the TF matrix from the natural source to the observation points considering only the active contribution of the AA system:

$$\mathbf{H}_\mathrm{AA}(e^{j\omega_k}) = \mathbf{H}_\mathrm{LA}(e^{j\omega_k}) \frac{\mathbf{G}\mathbf{V}_\mathrm{ML}(e^{j\omega_k})}{\mathbf{I} - \mathbf{F}_\mathrm{MM}(e^{j\omega_k})} \mathbf{H}_\mathrm{SM}(e^{j\omega_k}), \tag{28}$$

$\mathbf{I} \in \mathbb{R}^{n_\mathrm{M} \times n_\mathrm{M}}$ being an identity matrix.

Fig. 11 shows a continuous distribution of sound field energy along the plane obtained by the cubic interpolation of the $n_\mathrm{A} = 143$ observation-points energy values. Transducers and natural source are depicted in white using the same symbols as in Fig. 4. The top pane displays the distribution when the AA system is off (i.e., $\mathbf{H}_\mathrm{AA} = \mathbf{0}$). The bottom pane shows the distribution when the AA system is on. Fig. 11 reveals that the AA system preserves the natural sound field energy distribution and does not create areas with unexpectedly high or low sound energy. A statistical analysis of the sound-energy enhancement given by the AA system reveals a mean energy increment of 0.63 dB with a standard deviation of 0.42 dB.

[1] gianmarcodebortoli.github.io/AA-modalReverberator/.

6 CONCLUSION

The present work proposed a novel method for feedback control in multiple-input multiple-output AA systems with external reverberation. The DSP is a mixing matrix whose elements are modal artificial reverberators. Feedback attenuation is obtained through phase cancellation, which is generated by optimizing the phase values for each mode in the mixing matrix.

Results for single-mode reverberators show a GBIi between the optimized and nonoptimized phase shifts ranging from 20 to 28 dB. Results for multiple-mode reverberators show a GBIi >5 dB. Time variation in the RTFs marginally affects the results, because the GBIi decreases but remains positive and higher than 10 dB. The artificial reverberation obtained in the room is well distributed in space, and the phase optimization of the modal reverberator successfully removed the ringing tones.

7 REFERENCES

[1] P. U. Svensson, *On Reverberation Enhancement in Auditoria*, Ph.D. thesis, Chalmers University of Technology, Göteborg, Sweden (1994 Nov.).

[2] H. Nyquist, "Regeneration Theory," *Bell Syst. Tech. J.*, vol. 11, no. 1, pp. 126–147 (1932 Jan.). https://doi.org/10.1002/j.1538-7305.1932.tb02344.x.

[3] R. V. Waterhouse, "Theory of Howlback in Reverberant Rooms," *J. Acoust. Soc. Am.*, vol. 37, no. 5, pp. 921–923 (1965 May). https://doi.org/10.1121/1.1909480.

[4] M. Ohsmann, "Analyse von Mehrkanalanlagen," *Acta Acust. un. Acust.*, vol. 70, no. 4, pp. 233–246 (1990 Apr.).

[5] C. P. Boner and C. R. Boner, "Minimizing Feedback in Sound Systems and Room-Ring Modes With Passive Networks," *J. Acoust. Soc. Am.*, vol. 37, no. 1, pp. 131–135 (1965 Jan.). https://doi.org/10.1121/1.1909290.

[6] E. T. Jr. Patronis, "Electronic Detection of Acoustic Feedback and Automatic Sound System Gain Control," presented at the *57th Convention of the Audio Engineering Society* (1977 May), paper 1235.

[7] E. T. Jr. Patronis, "Acoustic Feedback Detector and Automatic Gain Control," US Patent 4,079,199 (1978 Mar.).

[8] C. P. Boner and C. R. Boner, "A Procedure for Controlling Room-Ring Modes and Feedback Modes in Sound Systems With Narrow-Band Filters," presented at the *16th Convention of the Audio Engineering Society*, (1964 Oct.), paper 324.

[9] M. R. Schroeder, "Improvement of Acoustic-Feedback Stability by Frequency Shifting," *J. Acoust.*

Soc. Am., vol. 36, no. 9, pp. 1718–1724 (1964 Sep.). https://doi.org/10.1121/1.1919270.

[10] S. J. Schlecht and E. A. P. Habets, "The Stability of Multichannel Sound Systems With Time-Varying Mixing Matrices," *J. Acoust. Soc. Am.*, vol. 140, no. 1, pp. 601–609 (2016 Jul.). https://doi.org/10.1121/1.4955285.

[11] M. Poletti, "A Unitary Reverberator for Reduced Colouration in Assisted Reverberation Systems," in *Proceedings of the INTER-NOISE and NOISE-CON Congress and Conference*, pp. 1223–1232 (Newport Beach, CA) (1995 Jul.).

[12] T. Lokki and J. Hiipakka, "A Time-Variant Reverberation Algorithm for Reverberation Enhancement Systems," in *Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFx)*, pp. 28–32 (Limerick, Ireland) (2001 Dec.).

[13] S. J. Schlecht and E. A. P. Habets, "Reverberation Enhancement Systems With Time-Varying Mixing Matrices," in *Proceedings of the 59th AES International Conference: Sound Reinforcement Engineering and Technology* (2015 Jul.), paper 3-2.

[14] P. Gil-Cacho, T. van Waterschoot, M. Moonen, and S. H. Jensen, "Regularized Adaptive Notch Filters for Acoustic Howling Suppression," in *Proceedings of the 17th European Signal Processing Conference*, pp. 2574–2578 (Glasgow, Scotland) (2009 Aug.).

[15] J. S. Abel, E. F. Callery, and E. K. Canfield-Dafilou, "A Feedback Canceling Reverberator," in *Proceedings of the 21st International Conference on Digital Audio Effects (DAFx)*, pp. 100–106 (Aveiro, Portugal) (2018 Sep.).

[16] T. van Waterschoot and M. Moonen, "Adaptive Feedback Cancellation for Audio Applications," *Signal Process.*, vol. 89, no. 11, pp. 2185–2201 (2009 Nov.). https://doi.org/10.1016/j.sigpro.2009.04.036.

[17] P. Coleman, N. Epain, S. Venkatesh, and F. Roskam, "Exploring Perceptual Annoyance and Colouration Assessment in Active Acoustic Environments," in *Proceedings of the AES International Conference on Acoustics & Sound Reinforcement* (2024 Jan.), paper 3.

[18] F. Kaiser, C. Frischmann, V. Werner, and T. Rohde, "Room Acoustics Evaluation of Active Acoustics Systems – Results from Measurements," in *Proceedings of the International Symposium on Room Acoustics (ISRA)*, pp. 177–185 (Amsterdam, The Netherlands) (2019 Sep.).

[19] V. Werner, S. Neeten, and F. Kaiser, "Evaluation Of A Geometric Approach To Active Acoustics," in *Proceedings of the Institute of Acoustics*, vol. 45 (Pt. 2), paper 24 (Athens, Greece) (2023 Sep.). https://doi.org/10.25144/16001.

[20] J. S. Abel, S. Coffin, and K. Spratt, "A Modal Architecture for Artificial Reverberation With Application to Room Acoustics Modeling," presented at the *137th Convention of the Audio Engineering Society* (2014 Oct.), paper 9208.

[21] M. Karjalainen and H. Järveläinen, "More About This Reverberation Science: Perceptually Good Late Reverberation," presented at the *111th Convention of the Audio Engineering Society* (2001 Nov.), paper 5415.

[22] J. Zhang, Y. Wen, Y. Zhang, Q. Zhou, and Z. Chu, "A Brief Review of Independent Sound Zone Technology," *Noise Vibr. Worldw.*, vol. 54, no. 7-8, pp. 414–422 (2023 Jul.). https://doi.org/10.1177/09574565231179950.

[23] T. Lokki, J. Nummelo, and T. Lahti, "An Electro-Acoustic Enhancement System for Rehearsal Room," in *Proceedings of the EAA Symposium on Architectural Acoustics*, paper AAQ06 (Madrid, Spain) (2000 Oct.).

[24] J. L. Nielsen and U. P. Svensson, "Performance of Some Linear Time-Varying Systems in Control of Acoustic Feedback," *J. Acoust. Soc. Am.*, vol. 106, no. 1, pp. 240–254 (1999 Jul.). https://doi.org/10.1121/1.427053.

[25] P. U. Svensson, "Influence of Electroacoustic Parameters on the Performance of Reverberation Enhancement Systems," *J. Acoust. Soc. Am.*, vol. 94, no. 1, pp. 162–171 (1993 Jul.). https://doi.org/10.1121/1.407094.

[26] C. P. Boner and C. R. Boner, "Behavior of Sound System Response Immediately Below Feedback," *J. Audio Eng. Soc.*, vol. 14, no. 3, pp. 200–203 (1966 Jul.).

[27] T. van Waterschoot and M. Moonen, "Fifty Years of Acoustic Feedback Control: State of the Art and Future Challenges," *Proceedings of the IEEE*, vol. 99, no. 2, pp. 288–327 (2011 Feb.). https://doi.org/10.1109/JPROC.2010.2090998.

[28] M. A. Poletti, "The Stability of Multichannel Sound Systems With Frequency Shifting," *J. Acoust. Soc. Am.*, vol. 116, no. 2, pp. 853–871 (2004 Aug.). https://doi.org/10.1121/1.1763972.

[29] T. Lokki, J. Pätynen, T. Peltonen, and O. Salmensaari, "A Rehearsal Hall With Virtual Acoustics for Symphony Orchestras," presented at the *126th Convention of the Audio Engineering Society* (2009 May), paper 7695.

[30] M. A. Poletti, "The Stability of Single and Multichannel Sound Systems," *Acta Acust. un. Acust.*, vol. 86, no. 1, pp. 163–178 (2000 Jan.).

[31] P. Dietrich, *Uncertainties in Acoustical Transfer Functions: Modeling, Measurement and Derivation of Parameters for Airborne and Structure-Borne Sound* (Logos Verlag, Berlin, Germany, 2013).

[32] W. K. Connor, "Experimental Investigation of Sound-System-Room Feedback," *J. Audio Eng. Soc.*, vol. 21, no. 1, pp. 27–32 (1973 Feb.).

[33] A. Politis, "shoebox-roomsim," https://github.com/polarch/shoebox-roomsim (accessed May 5, 2023).

[34] A. Farina, "Simultaneous Measurement of Impulse Response and Distortion With a Swept-Sine Technique," presented at the *108th Convention of the Audio Engineering Society* (2000 Feb.), paper 5093.

[35] A. B. Carlson and P. B. Crilly, *Communication Systems*, 5th ed. (McGraw-Hill, New York, NY, 2010).

[36] K. Prawda, S. J. Schlecht, and V. Välimäki, "Robust Selection of Clean Swept-Sine Measurements in Non-Stationary Noise," *J. Acoust. Soc. Am.*, vol. 151, no. 3, pp. 2117–2126 (2022 Mar.). https://doi.org/10.1121/10.0009915.

THE AUTHORS


Gian Marco De Bortoli


Karolina Prawda


Sebastian J. Schlecht

Gian Marco De Bortoli is a doctoral candidate at Aalto University, Espoo, Finland. He received a B.Eng. degree in Mechanical Engineering from University of Padua, Italy, in 2020, followed by an M.Sc. degree in Music and Acoustic Engineering from Politecnico di Milano, Italy, in 2023. He has been part of the Aalto Acoustics Lab since 2022. His research interests include active acoustics, differentiable digital signal processors, and artificial reverberation.

•

Karolina Prawda is a Lecturer (Assistant Professor) in Audio Engineering for Creative Technologies at the AudioLab of the School of Physics, Engineering and Technology at the University of York. In 2017, she received an M.Sc. degree in Acoustic Engineering from AGH University of Krakow, Poland, and in 2022, she was awarded a D.Sc. (Tech.) degree in Acoustics and Audio Signal Processing from Aalto University, Finland. In 2022–2024, she was a postdoctoral researcher at the Department of Information and Communications Engineering of Aalto University. Her research interests include acoustic metrology, sound propagation in inhomogeneous media, and artificial reverberation. Dr. Prawda is a member of the Audio Engineering Society and Acoustical Society of Finland. She is currently the vice-Chair of the AES Technical Committee on Audio Algorithms.

•

Sebastian J. Schlecht is an Associate Professor of Signal Processing at the Chair of Multimedia Communications and Signal Processing at the University of Erlangen-Nuremberg, Germany. He earned a Diploma in Applied Mathematics from the University of Trier, Germany, in 2010, followed by an M.Sc. in Digital Music Processing at Queen Mary University of London, UK, in 2011. He completed his Doctorate at the International Audio Laboratories Erlangen, Germany, in 2017. From 2012 to 2019, he also served as an external research and development consultant at Fraunhofer IIS, Erlangen, Germany. Between 2019 and 2024, he held the position of Professor of Practice for Sound in Virtual Reality at Aalto University, Finland. He has received multiple best paper awards, including those from *JAES*, WASPAA, DAFx (2018, 2021, 2022, and 2023), and AES AVAR.