APRIL 22 2026

Active acoustic enhancement systems: A review

Will J. Cassidy ; Gian Marco De Bortoli ; Karolina Prawda ; Philip Coleman ; Russell Mason ; Tapio Lokki ; Sebastian J. Schlecht ; Enzo De Sena

J. Acoust. Soc. Am. 159, 3533–3557 (2026)
https://doi.org/10.1121/10.0043585

JASA REVIEW-ARTICLE

Active acoustic enhancement systems: A review

Will J. Cassidy,[1,a] Gian Marco De Bortoli,[2] Karolina Prawda,[3] Philip Coleman,[4] Russell Mason,[1] Tapio Lokki,[2] Sebastian J. Schlecht,[5] and Enzo De Sena[1]

[1]*Institute of Sound Recording, University of Surrey, Guildford, GU2 7XH, United Kingdom*

[2]*Aalto Acoustics Lab, Department of Information and Communications Engineering, School of Electrical Engineering, Aalto University, FI-00076, Aalto, Finland*

[3]*AudioLab, School of Physics, Engineering and Technology, University of York, York, YO10 5DD, United Kingdom*

[4]*L-Acoustics, Highgate Studios, 53-79 Highgate Rd, London, NW5 1TL, United Kingdom*

[5]*Multimedia Communications and Signal Processing, University of Erlangen-Nuremberg, Cauerstr. 7, 91058 Erlangen, Germany*

ABSTRACT:

Active acoustic enhancement systems (AAESs) use microphones, loudspeakers and electronic processing to modify the reverberation of a space, offering flexible and cost-effective alternatives to passive variable acoustics. These systems can extend the reverberation time of a space and modify perceived characteristics such as wall distance, diffuseness and intimacy. In this article, the current literature is discussed, and common conditions of AAESs are demonstrated using simulations to help researchers to establish a comprehensive understanding of the field. A general model is first defined to approximate any AAES as a linear, time-invariant system of transfer functions. This is used to analyse the general stability condition, which is valuable for system tuning and prediction. The three main topologies of AAESs are presented, namely, in-line, regenerative and hybrid systems, describing the fundamental differences as well as the nuances of commercial implementations with a focus on signal processing techniques. Articles investigating AAESs have been summarised to allow readers to gauge the coverage of experimental research to date. The simulated contribution serves as an exploratory environment to compare AAES conditions, where code and audio examples are available online. Promising future trajectories are identified involving machine learning, artefact perception and expressive performance.

 https://doi.org/10.1121/10.0043585

(Received 28 October 2025; revised 23 February 2026; accepted 6 April 2026; published online 22 April 2026)

[Editor: James F. Lynch] Pages: 3533–3557

I. INTRODUCTION

Auditoria feature a variety of acoustic characteristics which may be well-suited to one type of event, but are unsuitable for another. Short early energy decay is necessary to achieve the clarity required for adequate speech communication, whereas musical performances can demand reverberant conditions. Baroque music was often written for small church performances with relatively short reverberation times (RTs), whereas the greater size of halls used during the Romantic period resulted in compositions being written for RTs of over two seconds (Beranek, 1992). As contemporary recording techniques were explored, popular music began to feature a wide range of reverberance. Drier concert venues were constructed where sound engineers could introduce additional reverberation via the front-of-house (FOH) loudspeakers if desired. This approach to increasing the reverberance of a space was unsuitable for most classical repertoire due to the lack of stage support required for musicians (Watanabe and Miyazaki, 2016), and to this day, orchestras still typically perform in purpose-built concert halls. Even for contemporary music, simply reverberating the FOH loudspeaker signals provides limited realism and listener envelopment (LEV). Hence, acoustically variable performance halls are universally valuable.

To increase the versatility of a space, passive variable acoustics can be used to mechanically vary the surface absorption and room volume (Beranek, 1992). Retractable curtains are common and can provide attenuation at mid and high frequencies, while rotating panels can expose either reflective, diffusive (scattering), or absorptive materials. These approaches typically lack low-frequency control, which is reportedly unsuitable for contemporary music (Hyon and Jeong, 2021), their installation can be expensive and pose structural challenges (Kaiser *et al.*, 2019), and switching configurations takes time and may be audible.

The concept of *active* variable acoustics aims to electronically alter the perceived characteristics of a space, which provides control beyond what is physically possible with passive solutions. An active acoustic enhancement system (AAES) functions by distributing microphones around a room to pick up acoustic energy radiated from natural sound sources. These signals are processed electronically, which may include filtering, modulation, and introducing artificial reflections, and are then routed to loudspeakers typically installed in the walls and ceiling of the environment. The ability to rapidly vary the electronic processing provides a significant advantage over passive variable acoustics by

[a]Email: w.cassidy@surrey.ac.uk

allowing presets to be prepared for different performances (Poletti, 2010), and future revisions to the system may be integrated as software updates, which improves installation longevity. Parameters that can be modified include the perceived room dimensions, amount of lateral energy, texture (how diffuse the room sounds), apparent source width (ASW), and RT, which can be extended far beyond that of mechanical means. AAESs are also beneficial in single-purpose settings such as rehearsal spaces and classrooms. Unsatisfactory acoustics may be improved by replacing missing reflections (Kok and Prinssen, 2003), reinforcing early energy for better speech intelligibility (Gade, 1997), and increasing the self-audibility of on-stage musicians (Ko *et al.*, 2012). Some active installations can increase RT independently of loudness, which is particularly useful for orchestral rehearsal spaces that require long RTs. Highly reverberant rooms can be loud, causing fatigue to performers from prolonged exposure, so using an AAES to extend the RT of a relatively dry rehearsal space can create a safer working environment (Hough *et al.*, 2025; Lokki *et al.*, 2000).

AAESs are usually categorised by three topologies based on the direct-to-reverberant ratio (DRR) of the signals received by the microphones (Poletti, 1998a). An *in-line* AAES relies on microphones placed near the stage to capture sources directly, often using microphone directivity to suppress room reverberation. Artificial reverberation can then be applied, either in the form of early reflections (Poletti, 2010) or full-length reverberation (Griesinger, 1991), which is reproduced by loudspeakers directly to the audience. This topology can be used to add supporting reflections to the existing space or to apply full-length reverberation to mask the original acoustics. *Regenerative* systems pick up sound sources using microphones placed around the room, and a controlled feedback loop is established with the loudspeakers. "Pure" regenerative AAESs do not use artificial reverberation and rely only on the natural propagation, reflection and diffusion of sound in the space to enhance the reverberation (de Koning, 1984; Parkin and Morgan, 1970). Due to the low DRR and relatively long initial time delay of the system inputs, regenerative AAESs are better suited to enhance the late reverberation, whereas in-line systems are more able to provide early energy control (Poletti, 2010). To operate with both of these principles, *hybrid* AAESs use near and far microphones with respect to the stage for increased flexibility, often utilising separate reverberators to process early and late energy (Poletti, 2006; Roskam *et al.*, 2022).

Despite providing considerably more flexibility than passive variable acoustics, AAESs encounter a range of challenges. As with any system of microphones and loudspeakers, feedback can be problematic. The power gain of an AAES must be kept below the instability threshold to prevent extreme system resonance. Assuming an AAES is stable, operating with high feedback gain may cause colouration, which can be perceived as annoying (Coleman *et al.*, 2024). One solution is to increase the channel count to reduce the gain of many isolated loops while maintaining the same power output (Griesinger, 1991). Some authors have suggested using around 100 channels, which introduces problems with cost, structural installation, and visible obtrusion (de Koning, 1984; Parkin and Morgan, 1970). Particularly in systems with low channel counts, time-varying reverberators can be utilised to increase the available gain before instability (GBI)—defined as the maximum gain applicable to the system feedback loop before instability is encountered—by around 3 to 6 dB (Griesinger, 1991; Miyazaki *et al.*, 2003; Schlecht and Habets, 2016). However, time-varying modulation can be prone to artefacts such as pitch shifting. Another solution is to increase the RT of the reverberator instead of the feedback gain, thus preventing colouration, but the resulting energy decay may exhibit multiple gradients, which can impact naturalness (Cassidy *et al.*, 2024).

Section II begins by introducing a signal model to describe a general AAES and its stability condition, which is referred to throughout this article. The most common AAES topologies are compared in Sec. II C with respect to the fundamental design choices, and Sec. II D introduces signal processing techniques used to provide greater enhancement headroom. These areas are then considered from the perspective of commercial implementations, exploring existing AAES installations in Sec. II E. The evaluation of active acoustics is detailed in Sec. III, which investigates how artefacts such as colouration and double-sloping can be quantified, as well as how perceptual experiments have been used to assess AAES performance. There is a distinct lack of comparison between AAES implementations in the literature, hence it is unclear what to expect when changing general parameters such as reverberator length, loop gain (μ) and transducer directivity. To demonstrate the behaviour of these systems under different parametric conditions, a set of simulations is analysed throughout Sec. IV. The purpose of this section is to provide a heuristic understanding of the differences between AAES implementations from an exploratory point of view. Section V summarises the common AAES design choices, and the most recent studies are discussed to underline current efforts and future directions in the field.

II. FUNDAMENTALS OF ACTIVE ACOUSTICS

This section provides an introduction to how a general AAES operates, using a signal model to identify its fundamental components and the electroacoustic stability condition. Variations in topological configurations are then derived from the signal model, and examples of commercial implementations are explored.

A. Signal model

To describe the architecture of different types of AAES, a general signal model is first presented from which they can be derived. Figure 1 shows a linear system consisting of S sound sources, R receivers, and a microphone array of size

https://doi.org/10.1121/10.0043585

M connected to a loudspeaker array of size L via an electronic processing unit. Due to popularity, the following model assumes a digital system, but analogue AAESs have been found in the past. Acoustic spaces are often approximated as linear time-invariant (LTI) systems because of the relative subtlety of temporal variation and nonlinearities in typical rooms. Approximating an AAES as an LTI system allows each quantity in Fig. 1 to be characterised as a matrix or vector of transfer functions (TFs), where each TF represents a discrete-time impulse response (IR) in the z-domain. This, therefore, assumes that each path in Fig. 1 can be sufficiently captured with a finite-length IR.

In Fig. 1, the sound sources (depicted as guitars) represent any natural acoustic source, such as an instrument or voice. Similarly, the receivers are depicted as pinnae, typically representing the sound received by members of the audience. The sound sources are presented as the input vector $\boldsymbol{u}(z) \in \mathbb{C}^{S\times 1}$, where $z = \rho\, e^{j\omega}$ is the complex frequency phasor on the z-plane with radius ρ and phase ω. The path of energy transfer from the sound sources to the receivers is characterised by the room transfer function (RTF) matrix $\boldsymbol{H}_{\mathrm{SR}}(z) \in \mathbb{C}^{R\times S}$, which includes the TFs of the transducers. The path from the sound sources to the microphones of the AAES is characterised by the RTF matrix $\boldsymbol{H}_{\mathrm{SM}}(z) \in \mathbb{C}^{M\times S}$, which similarly captures device TFs in addition to the room. The AAES microphone signals are then sent through a digital signal processing (DSP) unit $\boldsymbol{X}(z) \in \mathbb{C}^{L\times M}$, which is scaled by μ. The outputs of $\boldsymbol{X}(z)$ (denoted as the intermediate vector $\boldsymbol{x}(z) \in \mathbb{C}^{L\times 1}$) are routed to the AAES loudspeaker array, which forms paths to the receivers $\boldsymbol{H}_{\mathrm{LR}}(z) \in \mathbb{C}^{R\times L}$. A component of the AAES loudspeaker energy propagates back to the AAES microphones, defined as $\boldsymbol{H}_{\mathrm{LM}}(z) \in \mathbb{C}^{M\times L}$.

The source-to-receiver signals are presented as the output vector $\boldsymbol{v}(z) \in \mathbb{C}^{R\times 1}$, which is a superposition of the contributions from the room sources and the AAES loudspeakers,

$$\boldsymbol{v}(z) = \boldsymbol{H}_{\mathrm{SR}}(z)\boldsymbol{u}(z) + \boldsymbol{H}_{\mathrm{LR}}(z)\boldsymbol{x}(z). \tag{1}$$

The outputs of the DSP unit can be isolated as follows:

$$\boldsymbol{x}(z) = \mu\boldsymbol{X}(z)(\boldsymbol{H}_{\mathrm{SM}}(z)\boldsymbol{u}(z) + \boldsymbol{H}_{\mathrm{LM}}(z)\boldsymbol{x}(z)), \tag{2}$$

$$\boldsymbol{x}(z) = [\boldsymbol{I} - \mu\boldsymbol{X}(z)\boldsymbol{H}_{\mathrm{LM}}(z)]^{-1}\mu\boldsymbol{X}(z)\boldsymbol{H}_{\mathrm{SM}}(z)\boldsymbol{u}(z), \tag{3}$$

where $\boldsymbol{I}$ denotes the identity matrix of shape $L \times L$. Equation (1) can then be rewritten by substitution,

$$\boldsymbol{v}(z) = \boldsymbol{H}_{\mathrm{SR}}(z)\boldsymbol{u}(z) + \boldsymbol{H}_{\mathrm{LR}}(z) \times [\boldsymbol{I} - \mu\boldsymbol{X}(z)\boldsymbol{H}_{\mathrm{LM}}(z)]^{-1}\mu\boldsymbol{X}(z)\boldsymbol{H}_{\mathrm{SM}}(z)\boldsymbol{u}(z)\,. \tag{4}$$

In the practical implementation of the LTI model, each element of each quantity holds a TF evaluated for a finite set of points equally spaced around the unit circle of the z-plane, equivalent to the discrete Fourier transform (DFT) of the associated IRs. As such, Eq. (4) can be applied using $\rho = 1 \quad \omega = \omega_k : z = e^{j\omega_k} \forall k = 1, \ldots, K$, where K is the number of frequency bins (equal to the number of samples in its associated time-domain signal) and $\omega_k = 2\pi k/K$ is the uniformly-discretised normalised frequency variable. For direct multiplication to be possible, every matrix and vector of the model must have suitable dimensions. This can be achieved by zero-padding the time-domain signals to the same length so that their DFTs are also equal in length. The signal model can then be evaluated separately for each frequency bin k.

The signal model may be extended to represent time variation either end-to-end or for a subset of components. The former can be addressed by adding a time parameter to each quantity in Fig. 1, thereby defining the state of the TFs at each time step of the output signal. In practice, this is unlikely to be feasible due to high dimensionality, so a state-space representation of Eq. (4) may be constructed, which allows components to be varied at each time step of the simulation. In order for this time-domain model to be causal, it is assumed that there is at least one sample of group delay in either $\boldsymbol{X}[n]$ or $\boldsymbol{H}_{\mathrm{LM}}[n]$ [the time-domain representations of $\boldsymbol{X}(z)$ and $\boldsymbol{H}_{\mathrm{LM}}(z)$, respectively], which must be negated to compensate for the delay of the state-space representation. In practice, $\boldsymbol{H}_{\mathrm{LM}}[n]$ will exhibit a time delay due to the non-zero distance of any loudspeaker-microphone pair, and a non-zero sample delay will be introduced due to the analogue/digital conversion latency and buffer size used for the processing of $\mu\boldsymbol{X}[n]$.

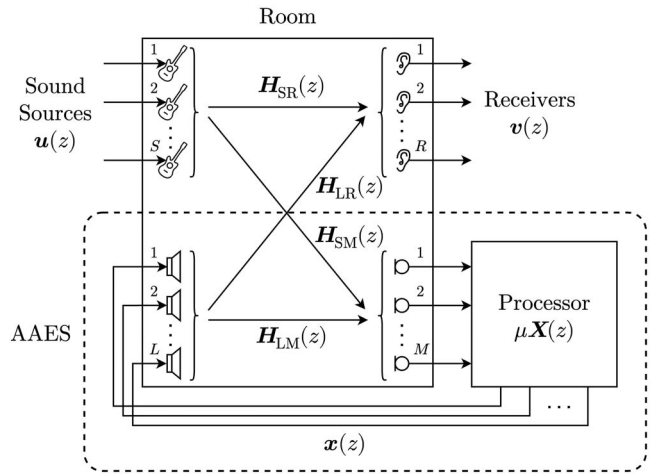

FIG. 1. Signal model of a general AAES, based on Poletti (2000, Fig. 11). The S sound sources and R receivers are depicted as guitars and pinnae, respectively. The active system is labelled "AAES" and comprises L loudspeakers, M microphones, and a processing unit. The solid arrows represent the signal flow with notations for all signal vectors (lower-case) and TF matrices (upper-case).

B. System stability

AAESs are prone to instability due to the inherent feedback component $\boldsymbol{H}_{\mathrm{LM}}(z)$. Instability is the condition that leads to an unbounded output for a given bounded input, or for a non-zero initial system state. This translates to the audio signals circulating in the feedback loop of the

system—labelled "AAES" in Fig. 1—being amplified at each loop iteration. This condition is potentially dangerous for listeners and equipment; it must be avoided.

An LTI system can be declared stable if all of its poles lie within the unit circle of the z-plane. In practice, evaluating the poles of an AAES is not feasible due to the complexity of the system, which grows with the IR length and the rank of the system. The rank corresponds to the number of unique loops in the feedback component of the system, which is at most equal to the minimum of the number of system loudspeakers and microphones (Poletti, 2011).

To discuss the stability analysis of an AAES, $X(z)$ will be assumed stable in this work. A single-channel system is first considered. To simplify the notation, $L = M = 1$ is assumed here, where the symbols $H_{\mathrm{LM}}(z)$ and $X(z)$ represent single TFs. Equation (4) becomes

$$\begin{aligned} \boldsymbol{v}(z) &= \boldsymbol{H}_{\mathrm{SR}}(z)\boldsymbol{u}(z) \\ &\quad + \boldsymbol{H}_{\mathrm{LR}}(z)\frac{\mu X(z)}{1-\mu X(z)H_{\mathrm{LM}}(z)}\boldsymbol{H}_{\mathrm{SM}}(z)\boldsymbol{u}(z)\,, \end{aligned} \tag{5}$$

and the stability of the AAES is governed by the closed-loop term,

$$\frac{1}{1-\mu X(z)H_{\mathrm{LM}}(z)}. \tag{6}$$

The Nyquist stability criterion (NSC) (Nyquist, 1932) is a procedure used to simplify the counting of poles outside the unit circle in a single-channel feedback system. It states that the stability of the closed-loop term can be inferred from the open-loop term $\mu X(z)H_{\mathrm{LM}}(z)$. The criterion considers whether the Nyquist plot of the open-loop term encircles the point $(+1, 0)$ on the complex plane. The closed-loop term is stable if the number of encirclements is zero. The Nyquist plot of the z-domain representation of the open-loop TF is obtained by mapping the unit circle of the z-plane to the complex plane through $\mu X(z)H_{\mathrm{LM}}(z)$. The NSC reduces to both the unconditional and conditional stability cases. Waterhouse (1965) revisited the NSC in the case of acoustic systems, affirming that only the unconditional stability is to be considered in this context. The revised version of the NSC states that when the open loop is real, its magnitude must be lower than unity. Therefore, a single-channel AAES is stable if

$$\begin{aligned} \mathfrak{Im}(\mu X(e^{j\omega})H_{\mathrm{LM}}(e^{j\omega}))) &\neq 0 \\ \mathfrak{Re}(\mu X(e^{j\omega})H_{\mathrm{LM}}(e^{j\omega}))) &\leq 1, \end{aligned} \tag{7}$$

where $\mathfrak{Im}(\cdot)$ and $\mathfrak{Re}(\cdot)$ represent the imaginary and real parts, respectively.

For the system TF of a multichannel AAES [see Eq. (4)], the stability is again governed by the closed-loop term,

$$[\boldsymbol{I} - \mu\boldsymbol{X}(z)\boldsymbol{H}_{\mathrm{LM}}(z)]^{-1}. \tag{8}$$

The multichannel approach is based on the multichannel version of the NSC, known as the generalised Nyquist stability criterion (GNSC) (MacFarlane and Postlethwaite, 1977). The GNSC states, again, that the stability of the system can be fully evaluated by analysing the open-loop TF matrix $\mu\boldsymbol{X}(z)\boldsymbol{H}_{\mathrm{LM}}(z)$. In the multichannel case, this means computing the eigenvalue decomposition of the open-loop matrix,

$$\mu\boldsymbol{X}(z)\boldsymbol{H}_{\mathrm{LM}}(z) = \boldsymbol{Q}(z)\boldsymbol{\Lambda}(z)\boldsymbol{Q}^{-1}(z), \tag{9}$$

where $\boldsymbol{Q}(z) \in \mathbb{C}^{L\times L}$ is a square matrix consisting of L eigenvectors across its columns, and $\boldsymbol{\Lambda}(z) \in \mathbb{C}^{L\times L}$ is a diagonal matrix of eigenvalues $\boldsymbol{\Lambda}(z) = \mathrm{diag}(\mu\lambda_1(z), \mu\lambda_2(z), \ldots, \mu\lambda_L(z))$. The open-loop matrix can also be derived by using the microphone input as the state, which leads to

$$\mu\boldsymbol{H}_{\mathrm{LM}}(z)\boldsymbol{X}(z) = \boldsymbol{Q}'(z)\boldsymbol{\Lambda}'(z)\boldsymbol{Q}'^{-1}(z) \tag{10}$$

having an alternative eigenvalue decomposition with eigenvectors $\boldsymbol{Q}'(z) \in \mathbb{C}^{M\times M}$ and eigenvalues $\boldsymbol{\Lambda}'(z) \in \mathbb{C}^{M\times M}$. The non-zero eigenvalues are shared between $\boldsymbol{\Lambda}(z)$ and $\boldsymbol{\Lambda}'(z)$. Thus, there are at most $P = \min(L, M)$ non-zero eigenvalues, equal to the rank of the system.

Considering the practical, discrete implementation of the LTI model, the eigenvalues are computed at the K frequency points. They are a discretisation of the continuous eigenfunctions, which are the characteristic polynomials of the open-loop TF matrix. Due to the nature of the eigenvalue decomposition, the eigenvectors are not unique, and therefore, the eigenvalues computed at different frequencies do not appear in the same order. Thus, two eigenvalues computed at adjacent frequencies with the same index p might not belong to the same eigenfunction (Ohsmann, 1990). By applying a continuity criterion to the computed eigenvalues at adjacent frequencies, it is possible to retrieve the exact continuous eigenfunction (MacFarlane and Postlethwaite, 1977; Schlecht and Weiss, 2024). The eigenfunctions are fully independent of one another, hence applying a continuity criterion after the eigenvalue decomposition provides a decomposition of the multichannel AAES into a superposition of P independent single-channel AAESs, i.e., eigenchannels (Poletti, 2010). Finally, the conditions from Eq. (7) can be applied to each continuous eigenfunction individually (Emami-Naeini and Kosut, 2012).

In practice, rooms exhibit stochastic behaviour above the Schroeder frequency f_{sch} (Schroeder and Kuttruff, 1962) and are sensitive to atmospheric changes (Prawda *et al.*, 2024; Prawda, 2025; Prawda *et al.*, 2023). Thus, the number of encirclements of $\mu X(z)H_{\mathrm{LM}}(z)$ around the point $(+1, 0)$ cannot be predicted with precision (Schroeder, 1964). Connor (1973) experimentally validated Eq. (7): the stability limit was correctly predicted in highly controlled environments, with an error of 0.5 dB, but changes in air temperature generated variance in the results. Therefore, the NSC is not a robust stability validation procedure for AAESs and, by extension, neither is the GNSC due to uncertainty in the data. Instead, all practical approaches aim to estimate the stability limit with empirical or statistical procedures.

https://doi.org/10.1121/10.0043585

In the early development of AAESs, the GBI was estimated heuristically. The first procedure to improve the system response was performed by iteratively increasing μ until instability, detecting the most resonant frequency, and suppressing said frequency with a notch filter (Boner and Boner, 1965, 1964). The first model attempting the prediction of the GBI was an empirical formula proposed by Schroeder (1964),

$$\mathrm{GBI}_{\mathrm{sch}} = -10\log_{10}\left(\log_{10}\left(\frac{T_{60}B}{22}\right)\right) - 3.8\ \mathrm{dB}, \quad (11)$$

where T_{60} is the RT of the room and B is the bandwidth of the system over which its frequency response is considered flat.

Ohsmann (1990) proposed an estimation method for the GBI in the multichannel case based on the GNSC. By substituting Eq. (9) into the closed-loop matrix, the following is obtained:

$$\begin{aligned}&\left[\boldsymbol{I} - \boldsymbol{Q}(z)\boldsymbol{\Lambda}(z)\boldsymbol{Q}^{-1}(z)\right]^{-1} = \boldsymbol{Q}\left(z\right)\left[\boldsymbol{I} - \boldsymbol{\Lambda}(z)\right]^{-1}\boldsymbol{Q}^{-1}(z),\\ &\left[\boldsymbol{I} - \boldsymbol{\Lambda}(z)\right]^{-1} = \mathrm{diag}\left(\frac{1}{1-\mu\lambda_1(z)},\ldots,\frac{1}{1-\mu\lambda_L(z)}\right).\end{aligned} \quad (12)$$

The GBI can be estimated through the P non-zero eigenvalues. Thus, the stable gain μ at the complex frequency $\hat{z}$ is $\mu : \mu < |\lambda_p(\hat{z})|^{-1}\ \forall\ p = 1,\ldots,P$. By applying this procedure at K discrete frequency values to obtain a collection of $P \times K$ eigenvalues, the GBI is estimated as

$$\mathrm{GBI} = \min_{k,p}\left(|\lambda_p(e^{j\omega_k})|^{-1}\right), \quad k = 1,\ldots,K, \quad p = 1,\ldots,P. \quad (13)$$

The constraint on the magnitude of the open-loop eigenvalues is stricter than the constraint on the real part given by Eq. (7). This stems from Eq. (12) and the formal stability condition of the geometric series. Further studies used a hybrid approach between Eqs. (7) and (13) (Poletti, 2000; Schlecht and Habets, 2016),

$$\begin{aligned}\mathrm{GBI} = \min_{k,p}&\left(\left\{\mathfrak{Re}(\lambda_p(e^{j\omega_k}))\right\}^{-1}\right),\\ &k = 1,\ldots,K, \quad p = 1,\ldots,P,\end{aligned} \quad (14)$$

which provides an estimation of the GBI closer to the true stability limit. However, this approach is sensitive to phase, and hence a conservative approach could prove more robust in practice.

A statistical approach to estimate the stability limit was proposed by Poletti (2000). All previous methods required direct measurements of either acoustical parameters of the physical space or the open-loop TF matrix, and provided an estimate of the GBI for a specific AAES installation. Instead, Poletti's model is a statistical theory of stability that is independent of the space of installation and can be applied prior to setup. The theory comprises a series of equations predicting the probability of system instability depending on the statistics of $\boldsymbol{X}(z)$, which are known at design. The theory is valid under the following assumptions:

(1) The analysis is restricted to frequencies higher than f_{sch}.
(2) The direct sound transmission from system loudspeakers to system microphones is negligible. Thus, $\boldsymbol{H}_{\mathrm{LR}}(z)$ is approximated as late reverberation only.
(3) The system is considered LTI.
(4) The cross-correlation between frequency samples of RTFs is zero. Due to this approximation, the accuracy of the analysis depends on the frequency resolution adopted (Poletti, 2000; Schroeder, 1962).

Poletti (2000) validated this approach with AAES simulations using synthetic RTFs following assumption (2). Assumption (2) may be dropped (Schroeder, 1962), but this has not been explored mathematically. Assumption (4) may be dropped at the cost of a more complex formulation (Poletti 2000).

C. System topologies

This section introduces and characterises in-line, regenerative and hybrid AAESs in terms of their design requirements, performance and commercial realisations.

1. In-line systems

Figure 2 describes a general in-line AAES as a version of the system in Fig. 1. Thus, the TF of the system can be derived from Eq. (4). The critical distances of the sources and loudspeakers are defined as the radii where the DRRs of $\boldsymbol{H}_{\mathrm{SM}}(z)$ and $\boldsymbol{H}_{\mathrm{LR}}(z)$ ($\mathrm{DRR}_{\mathrm{SM}}$ and $\mathrm{DRR}_{\mathrm{LR}}$) equal unity. The critical distance of the stage sources is labelled $D^{\mathrm{S}}_{\mathrm{crit}}$. In-line AAESs utilise microphones within $D^{\mathrm{S}}_{\mathrm{crit}}$, the radius that is influenced by the source and receiver directivity patterns and the room reverberation. In musical contexts, in-line AAESs often utilise directional microphones positioned over and pointing at the stage (Rumsey and Kok, 2014), and in speech contexts, microphones are often in close proximity to the speakers. As a result, only the stage sources are significantly enhanced, whereas sounds emanating from the audience remain subject to the original acoustics. The microphone signals are sent through $\mu\boldsymbol{X}(z)$, which may consist of an early reflections generator (Poletti, 2006), or full-length reverberation (Griesinger, 1991). Since in-line microphones capture a high DRR and have a relatively low time-of-flight from the stage sources, early energy can be manipulated effectively. This can be exploited to modify spatial characteristics such as localisation, ASW, and intimacy (Poletti, 2010), and improve speech intelligibility (Berkhout *et al.*, 1988; Prinssen and Holden, 1992). The reverberated microphone signals are reproduced by loudspeakers distributed throughout the space. By utilising directional microphones and loudspeakers, the loudspeaker-to-microphone feedback component $\boldsymbol{H}_{\mathrm{LM}}(z)$ is minimised to keep the risk of instability low (Poletti, 1996). This is represented in Fig. 2 by a dotted line, indicating that the system

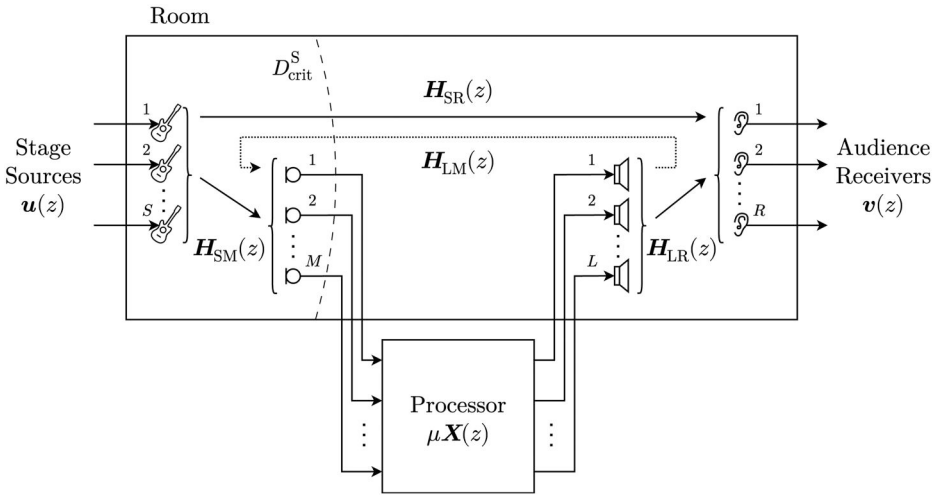

FIG. 2. A variation of Fig. 1 to represent a general in-line AAES. The critical distance $D^{\mathrm{S}}_{\mathrm{crit}}$ is illustrated for the stage sources. The feedback component, $\boldsymbol{H}_{\mathrm{LM}}(z)$, is minimised by design.

acts predominantly in a feed-forward manner, hence the name "in-line."

The strength of the source-to-receiver path, $\boldsymbol{H}_{\mathrm{SR}}(z)$, relative to the source-AAES-receiver path, $\mu\boldsymbol{H}_{\mathrm{LR}}(z)\boldsymbol{X}(z)\boldsymbol{H}_{\mathrm{SM}}(z)$, can affect characteristics of the resulting reverberation. It is possible that the stage sources significantly excite the acoustic space, but the AAES loudspeakers do not. For example, if the source were an orchestra, the signal at the listening position could be approximated as the summation of these terms,

$$\boldsymbol{v}(z) \approx \left\{\boldsymbol{H}_{\mathrm{SR}}(z) + \mu\boldsymbol{H}_{\mathrm{LR}}(z)\boldsymbol{X}(z)\boldsymbol{H}_{\mathrm{SM}}(z)\right\}\boldsymbol{u}(z), \tag{15}$$

where the high $\mathrm{DRR}_{\mathrm{SM}}$ would limit the contribution of $\boldsymbol{H}_{\mathrm{SM}}(z)$ to the overall reverberance. Depending on the value of $\mathrm{DRR}_{\mathrm{LR}}$, which may or may not be high, the summation can result in a superposition of the energy decay curves (EDCs) of the room $\boldsymbol{H}_{\mathrm{SR}}(z)$ and reverberation applied by $\boldsymbol{X}(z)$, potentially causing the double-slope effect (DSE; see Sec. III B). The feedback component is assumed negligible in this example, but it should be noted that this is rarely the case in practice (Vuichard and Meynial, 2000).

Alternatively, $\boldsymbol{H}_{\mathrm{SR}}(z)$ could be low in level relative to the loudspeakers. For example, a presenter could speak directly into a microphone connected to an input of the AAES. This would result in the reproduction system dominating the room excitation, so the signal at the listening position can be approximated as the artificially reverberated source without the room excitation due to the source,

$$\boldsymbol{v}(z) \approx \mu\boldsymbol{H}_{\mathrm{LR}}(z)\boldsymbol{X}(z)\boldsymbol{H}_{\mathrm{SM}}(z)\boldsymbol{u}(z). \tag{16}$$

In this case, $\mathrm{DRR}_{\mathrm{LR}}$ would have a prominent effect on the mix of active and passive reverberation in the audience, where a higher $\mathrm{DRR}_{\mathrm{LR}}$ would allow the artificial reverberation to dominate the energy decay. This is difficult to achieve in practice as the AAES loudspeakers would need to be in close proximity to the audience, such that $\mathrm{DRR}_{\mathrm{LR}}$ were high, and this would require many loudspeakers to avoid individual loudspeaker localisation and to ensure that the reverberation enhancement was spatially distributed. The concealment of such a system would also prove challenging. As such, in-line AAESs tend to exhibit some mixture of the active and passive reverberation. In the time domain, this is equivalent to the series convolution of the source with the reverberator IRs and the passive room impulse responses (RIRs), which can cause a convex-shaped EDC (Hak and Wenmaekers, 2013, 2015; Hughes, 2016; Vogel and de Vries, 1994).

2. Regenerative systems

Figure 3 shows the arrangement of a regenerative AAES, based on Fig. 1. The differences between this layout and that of an in-line system (Fig. 2) regard the acoustic positioning of the microphones and loudspeakers of the AAES. The microphones are positioned outside $D^{\mathrm{S}}_{\mathrm{crit}}$ to ensure the incident signals from the stage have a low DRR. These microphones are typically distributed around the walls and ceiling of the space, rather than over the stage. Another characteristic of a regenerative AAES is that the loudspeakers are positioned such that the microphones are also outside the critical distance of these loudspeakers, $D^{\mathrm{L}}_{\mathrm{crit}}$. As with the microphones, these are typically positioned around the walls and ceiling of the enclosure. The literature does not impose strict control over $\mathrm{DRR}_{\mathrm{LR}}$, but keeping the audience outside $D^{\mathrm{L}}_{\mathrm{crit}}$ may help to avoid loudspeaker localisation, and the loudness attenuation could be useful to control high sound pressure levels.

By placing AAES microphones in the acoustic proximity of the system loudspeakers, the feedback component $\boldsymbol{H}_{\mathrm{LM}}(z)$ is reinforced. This principle can be used to reintroduce the loudspeaker energy back into the system, known as regeneration. Due to the AAES microphone array's position, both the stage sources and the loudspeaker signals will be received by the microphones with low DRR. Since the direct component of the stage sources will be relatively low in level and will arrive significantly later than with stage microphones, regenerative systems do not provide as much early energy control as in-line systems (Poletti, 2010). Another disadvantage of using a purely regenerative AAES—that is, where $\boldsymbol{X}(z)$ does not introduce artificial reflections (de Koning, 1984)—is that to increase the RT of

https://doi.org/10.1121/10.0043585

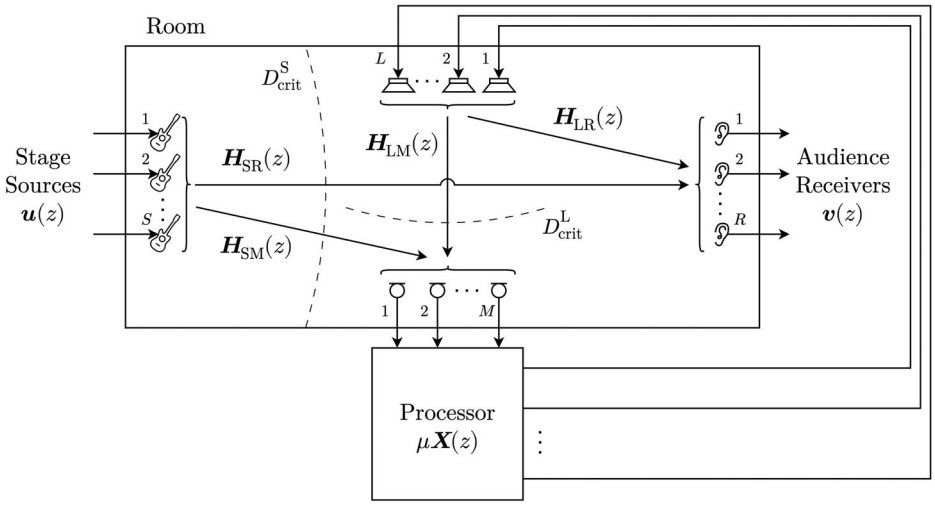

FIG. 3. A variation of Fig. 1 to represent a general regenerative AAES. The critical distances, $D^{\mathrm{S}}_{\mathrm{crit}}$ and $D^{\mathrm{L}}_{\mathrm{crit}}$, are illustrated for the stage sources and system loudspeakers, respectively.

a space, the system gain must be increased, which has been referred to as "longer is louder" (Bakker and Gillan, 2014). This imposes an enhancement limit due to the system's overall loudness, causing listener fatigue.

The low DRR of the AAES loudspeaker signals results in natural decorrelation via the transfer path $\boldsymbol{H}_{\mathrm{LM}}(z)$, which is important to decrease the risk of colouration via feedback, allowing more energy to be regenerated before instability. Vuichard and Meynial (2000) explained that when the distance between a microphone and its corresponding loudspeaker is on the order of $D^{\mathrm{L}}_{\mathrm{crit}}$, the signal may be considered sufficiently decorrelated. They also found that closing this distance to $0.2D^{\mathrm{L}}_{\mathrm{crit}}$ approximately maintained the decorrelation, provided that the transducer directivities were exploited to attenuate the direct path. The natural decorrelation provided by $\boldsymbol{H}_{\mathrm{LM}}(z)$ is crucial when the system is purely regenerative; the microphones of the AAES do not strictly need to be in the reverberant field of the loudspeakers if $\boldsymbol{X}(z)$ provides sufficient decorrelation. Placing microphones in the reverberant field in small or highly absorptive rooms can prove challenging, so applying artificial reverberation may be beneficial in these cases.

Similar to in-line AAESs, the perceived result can change given different types of input sources, but with an additional interaction between μ and the degree of regeneration. In the case where $\boldsymbol{u}(z)$ represents an orchestra, both $\boldsymbol{H}_{\mathrm{SR}}(z)$ and $\boldsymbol{H}_{\mathrm{SM}}(z)$ would be high in level, so when setting μ such that significant energy were reintroduced via feedback, the signal at the listening position would be represented by Eq. (4) in full. However, if μ were lowered below a threshold of significant regeneration but the AAES were still audible to the listener, then the receiver vector could be approximated according to Eq. (15), where the AAES would act as an in-line system. Regenerative AAESs can therefore also be prone to the DSE.

3. Hybrid systems

The third type of AAES is known as the hybrid system, which combines the early reflection control of an in-line system with late regeneration (Kaiser *et al.*, 2019; Poletti, 2010, 2011). The layout of a hybrid system, shown in Fig. 4, is visually similar to that of the regenerative system (Fig. 3). The difference is that a subset of the microphones is positioned within $D^{\mathrm{S}}_{\mathrm{crit}}$. In Fig. 4, there are M_1 microphones within $D^{\mathrm{S}}_{\mathrm{crit}}$, and M_2 microphones outside such that $M_1 + M_2 = M$, where M is the total number of microphones in the general model (Fig. 1). All M microphones are routed through $\mu\boldsymbol{X}(z)$, but the processing is typically separate for the early and late signals. The processing may still be defined using a single matrix, and hence Eq. (4) holds.

Hybrid AAES implementations utilise the direct and reverberant fields to control early and late energy effectively (Poletti, 2010). Typically, the M_1 microphones close to the stage area are connected to an in-line early reflections generator, and M_2 microphones in the audience area are part of a regenerative system (Poletti, 2010). This allows early reflections to be controlled with a sufficiently high DRR, and the regenerative system can be used to enhance late reverberation. Separating in-line and regenerative components also allows the early-to-late ratio to be varied, enabling the balance between ASW and LEV to be tuned (Poletti, 2010).

If $\boldsymbol{u}(z)$ were a vocalist singing directly into a microphone of a hybrid AAES, then $\boldsymbol{H}_{\mathrm{SR}}(z)$ would be insignificant. In this case, given significant regeneration due to μ, the signals at the listening position could be approximated as

$$\boldsymbol{v}(z) \approx \boldsymbol{H}_{\mathrm{LR}}(z)[\boldsymbol{I} - \mu\boldsymbol{X}(z)\boldsymbol{H}_{\mathrm{LM}}(z)]^{-1}\mu\boldsymbol{X}(z)\boldsymbol{H}_{\mathrm{SM}}(z)\boldsymbol{u}(z), \tag{17}$$

that is, Eq. (4) without the source-to-receiver component. Alternatively, if regeneration were insignificant due to μ, then the system could be approximated by Eq. (16).

D. Signal processing

Given an AAES setup, the DSP unit $\boldsymbol{X}(z)$ is responsible for controlling how the signal in the feedback loop evolves in time and frequency. A variety of architectures are

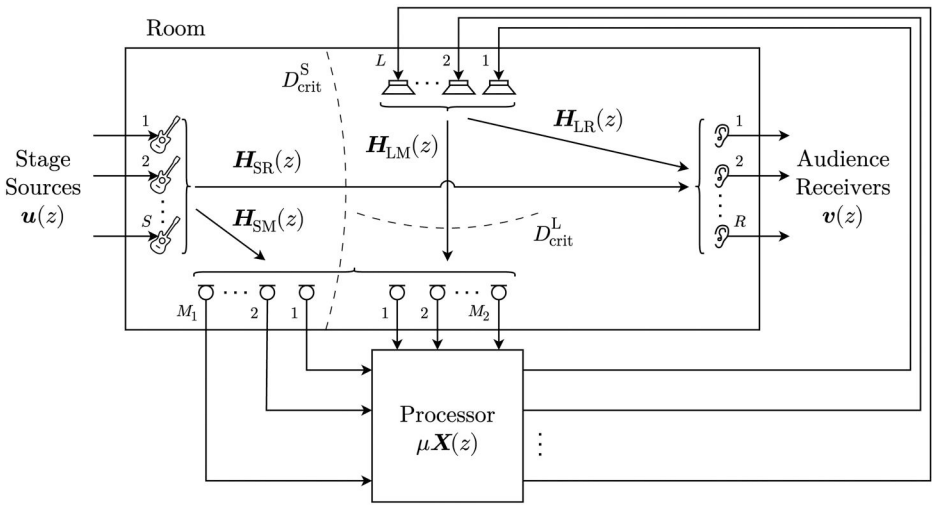

FIG. 4. A variation of Fig. 1 to represent a general hybrid AAES. The critical distances, $D^{\mathrm{S}}_{\mathrm{crit}}$ and $D^{\mathrm{L}}_{\mathrm{crit}}$, are illustrated for the stage sources and system loudspeakers, respectively.

available that share the criteria of stability and flexibility, offering versatile acoustic enhancement while avoiding colouration. Several DSP techniques have been adopted by AAESs, including those aimed at increasing the GBI, which leads to greater RT extension before colouration, and others that extend the RT independently of μ.

Techniques that aim to increase the GBI focus on providing effective decorrelation in the open-loop TF matrix $\mu\boldsymbol{X}(z)\boldsymbol{H}_{\mathrm{LM}}(z)$. As described in Secs. II B and III A, the open-loop TF matrix is frequency dependent, thus a fixed μ may cause certain frequencies to resonate strongly, but not others. Statistically, an RTF between two points outside of their relative critical distance in a typical auditorium has a Rayleigh-distributed magnitude that presents an average peak-to-mean ratio of 5 dB, with extreme values up to 15 dB (Schroeder, 1962, 1964). Providing $\boldsymbol{X}(z)$ is a diagonal matrix, the system open-loop TF matrix retains those statistics (Poletti, 2000). The average peak-to-mean ratio of the RTFs limits the overall enhanced reverberation gain. This is due to some frequencies [i.e., ω_k such that $|\mu\boldsymbol{X}(e^{j\omega_k})\boldsymbol{H}_{\mathrm{LM}}(e^{j\omega_k})|$ has a maximum] being close to their stability limit, and some others [i.e., ω_k such that $|\mu\boldsymbol{X}(e^{j\omega_k})\boldsymbol{H}_{\mathrm{LM}}(e^{j\omega_k})|$ has a minimum] being far from enhancement instead.

Increasing the amount of inter-channel mixing, i.e., utilising the non-diagonal elements of $\boldsymbol{X}(z)$, can decrease the peak-to-mean ratio by smoothing $|\mu\boldsymbol{X}(e^{j\omega_k})\boldsymbol{H}_{\mathrm{LM}}(e^{j\omega_k})|$ toward its mean magnitude value (Griesinger, 1991; Poletti, 2004). Inter-channel mixing works well for regenerative and hybrid systems with many transducers, where a larger number of channels leads to a higher smoothing of $|\mu\boldsymbol{X}(e^{j\omega_k})\boldsymbol{H}_{\mathrm{LM}}(e^{j\omega_k})|$. By employing an optimisation algorithm, the entries of $\boldsymbol{X}(z)$ can be tuned to achieve a high degree of spectral flatness in the feedback loop (De Bortoli *et al.*, 2024a).

Employing time variation in $\boldsymbol{X}(z)$ is another technique that affects the signal regeneration. Decorrelation of successive feedback loop iterations can be achieved by modulating the open-loop TFs over time (Svensson, 1994). This has the effect of smoothing $|\mu\boldsymbol{X}(e^{j\omega_k})\boldsymbol{H}_{\mathrm{LM}}(e^{j\omega_k})|$ over time (in different ways depending on what is varied), which reduces the accumulation of energy at the open-loop resonances. Its main disadvantage is the introduction of modulation artefacts such as beating, pitch shifting, and amplitude modulation, which degrade the naturalness of the enhanced reverberation (Meynial and Vuichard, 1999; Svensson, 1994). Hence, a compromise must be found between modulation subtlety and the effectiveness of resonance suppression.

Common time-varying filters in AAESs, namely, phase modulation, delay modulation, and frequency shifting, were initially designed for single-channel systems without external reverberation (Svensson, 1995). Amongst them, frequency shifting proved superior, theoretically providing a perfect smoothing of $|\mu\boldsymbol{X}(e^{j\omega_k})\boldsymbol{H}_{\mathrm{LM}}(e^{j\omega_k})|$ (Nielsen and Svensson, 1999; Poletti, 2004; Schroeder, 1964), and in practice achieving values close to those predicted. Consequently, more complex techniques have proven effective for multichannel AAESs, such as modulating channel routing (Miyazaki *et al.*, 2003). Schlecht and Habets (2015, 2016) further developed frequency shifting into a multichannel design that maintained the same degree of open-loop smoothing but produced less perceptible artefacts. It should be noted that the strength of such artefacts can depend on the enhancement context, for example, where the sensitivity to pitch shifting can be lower for speech communication than for music (Woszczyk, 2011). In general, the improvement from temporal variation decreases with the number of channels (Poletti, 2011).

Feedback attenuation is a technique that reduces signal regeneration in the feedback loop by cancelling the loudspeaker contribution from the signals received by the microphones. For this reason, the only technique described in this section requires knowledge of the physical space. Ideal feedback attenuation can theoretically lead to infinite GBI. In practice, this approach requires inverting the feedback TF matrix, which disregards the time variation of the physical room when precomputed and becomes increasingly computationally expensive for higher channel counts when performed in real-time. The approach proposed by Abel *et al.* (2018)

https://doi.org/10.1121/10.0043585

utilises the precomputed inverse TF matrix, but only the part with a low degree of temporal variation. De Bortoli *et al.* (2024b) suggested the use of acoustic phase cancellation, limiting the analysis to the low frequency range. In-line systems inherently minimise the feedback component through transducer positioning; thus, this approach mostly benefits regenerative and hybrid systems.

The relation between the RT extension and μ can be disentangled by defining $\boldsymbol{X}(z)$ as a full-length reverberator, whereby increasing the RT of $\boldsymbol{X}(z)$ can extend the resulting RT of the space while keeping μ constant (Griesinger, 1991; Poletti, 1995). In addition, a full-length reverberator grants control over the reflection pattern and the absorption characteristics of the enhanced reverberant field (Poletti, 1993, 1996, 2006). Introducing such a reverberator comes with the risk of causing the DSE; if the RT is greater than approximately two times that of the passive space, the shape of the EDC can sound unnatural (Cassidy *et al.*, 2024). It is also possible for colouration to be caused by the increased peak-to-mean ratio in the open-loop TF matrix due to the resonances of the reverberator coinciding with the resonances of the physical room (Poletti, 2000). The degradation of the open-loop TF statistics may be tackled by employing a reverberation unit with allpass properties (Poletti, 1995).

Some solutions in the literature combine full-length reverberation, channel mixing, and time variation, providing high GBI while allowing independent control over the RT extension. Miyazaki *et al.* (2003) and Schlecht and Habets (2015) merged channel mixing and time variation into one design. Griesinger (1991), Lokki and Hiipakka (2001), and Schlecht and Habets (2015) proposed different implementations of time-varying reverberators. Poletti (2006) divided the processing path into two parallel routes: one for early reflections, employing channel mixing; and one for late reverberation, employing an allpass reverberator.

E. Commercial implementations

Table I presents the commercial AAESs discussed in this section.

1. In-line systems

The System for Improved Acoustic Performance (SIAP) is an in-line AAES developed by Prinssen and Holden (1992). This system was constructed using multiple modules, each consisting of two microphones and several reverberation processors. Typically, around 24 microphones were positioned roughly 10 m above the front of the stage area, the signals of which were processed using both early reflections and late reverberation generators. The early and late reverberated signals were reproduced by ten foldback loudspeakers near the stage area, and approximately 36 loudspeakers distributed around the space to ensure an even spread of medial and lateral energy. Each loudspeaker-microphone TF was equalised in third-octave bands to reduce the risk of feedback. The SIAP has been used specifically for speech reinforcement tasks where the passive conditions are poor. In the speech setting, early reflections are added to improve speech intelligibility for the audience and to improve audibility on stage.

The concept of wave field synthesis was used to motivate an in-line AAES called the acoustic control system (ACS) (Berkhout *et al.*, 1993; Berkhout *et al.*, 1997). Wave field synthesis is based on Huygens's principle, which states that any point on a wavefront can be considered as the incidence of a separate point source (Berkhout *et al.*, 1993). In the ACS, approximately 20 microphones were used to capture the wave field of stage sources, modify it, and reproduce it using a loudspeaker array (Poletti, 2011). This AAES features multiple modules, including "speech," "theatre," and "reverberation hall," which can be used independently or in combination (Berkhout *et al.*, 1988).

In 1975, a US patent was filed by Jaffe (1977) for the Electronic Sound Enhancing System, which was later known as the Electronic Reflected Energy System (ERES) (Jaffe and Scarbrough, 1992). This in-line system used six microphones in the stage area to capture sources, which were summed into one signal and reproduced by four loudspeakers at the front of the stage and four in the audience area. The active system introduced time delays to the two groups of loudspeakers, so that the stage-end loudspeakers produced early reflections with a short delay, while the audience-end loudspeakers were delayed further to enhance late reverberation. The audience loudspeakers were low-pass filtered to enhance the energy only in the low frequencies. This system was based on time delays and did not introduce additional artificial reverberation.

In 1985, another US patent was filed for an in-line AAES invention. This was the Reverberation-on-Demand System (RODS) by Barnett (1987), which applied dynamic gating to five microphones spaced around the auditorium (Griesinger, 1991). The microphone signals were split into eight weighted frequency bands and were then analysed to determine the derivative of the mean input level. If the mean were constant or increasing, the input signals would be fed into multiple delay lines, and if the mean were decreasing, the output of the delay lines would be sent to three sets of loudspeakers (Griesinger, 1991). This architecture was designed to avoid feedback and, as a result, would enhance stop-chords but not continuous music.

The Lexicon Acoustic Reinforcement and Enhancement System (LARES) was developed by Griesinger (1991) as an in-line AAES utilising Lexicon reverberation units with time-varying characteristics (see Sec. II D). The system used a small number of microphones near the stage sources and a large number of loudspeakers distributed around the audience. The LARES was first installed at the Elgin Theatre in Toronto, using two Lexicon 480 processors to reverberate two microphone signals via 16 reverberators, producing 120 loudspeaker outputs. The loudspeakers were mounted in the ceiling across the audience area, with each reverberator output channel feeding a bank of loudspeakers. These were interleaved with other banks so that adjacent loudspeakers were uncorrelated, increasing perceived diffusion. The

TABLE I. A comparison of AAES installations. Artificial reverberation is defined as applying additional reflection density in $\boldsymbol{X}(z)$. The year of each AAES is stated as the primary publication year unless otherwise informed.

System	Year	Topology	Artificial reverberation	Time-varying	Primary publication
AR	1964	Regenerative	False	False	Parkin and Morgan, 1970
ERES	1977	In-line	False	False	Jaffe, 1977
MCR	1984	Regenerative	False	False	de Koning, 1984
RODS	1987	In-line	True	False	Barnett, 1987
AFC	1990	Hybrid	True	True	Kawakami and Shimizu, 1990
LARES	1991	In-line	True	True	Griesinger, 1991
SIAP	1992	In-line	True	Optional	Prinssen and Holden, 1992
ACS	1993	In-line	True	False	Berkhout *et al.*, 1993
VRAS	1996	Hybrid	True	False	Poletti, 2006
Carmen	2004	Regenerative	False	False	Schmich and Vian, 2004
Constellation	2006	Hybrid	True	False	Meyer Sound, 2024
Vivace	2008	In-line/Hybrid	True	True	Müller-BBM Acoustic Solutions GmbH, 2023
Amadeus	2019	Hybrid	True	Optional	Kaiser *et al.*, 2019
CarmenCita	2020	Regenerative	True	False	Jagla and Chervin, 2020
Ambiance	2022	Hybrid	True	False	Roskam *et al.*, 2022

LARES has a number of presets similar to the ACS, such as "speech" and "church," offering solutions ranging from speech reinforcement to high RT extension (Gade, 1997).

The Vivace system is a commercial AAES by Müller-BBM Acoustic Solutions GmbH (2023). This system uses low-latency convolution as the basis of the artificial reverberation, which also includes a time-varying component (Poletti, 2011; Walter and Melchior, 2008). Vivace has mostly been referred to as an in-line system in the literature (Kaiser, 2009; Ko, 2015; Poletti, 2010, 2011; Walter and Melchior, 2008) however according to Müller-BBM Acoustic Solutions GmbH (2023), the AAES uses "ambience" microphones around the room in addition to stage area microphones, suggesting it may be similar to a hybrid system which is supported by Hyon and Jeong (2021) and Möller (2022). A large number of loudspeakers are distributed along the walls and ceiling of the space, and the AAES also utilises FOH line arrays.

2. Regenerative systems

Parkin and Morgan (1970) developed the assisted resonance (AR) system, which aimed to modify the energy decay of discrete room modes with a large number of tuned electronic resonators. This was achieved by placing microphones inside Helmholtz resonators, whose signals were amplified and connected to loudspeakers in quarter-wavelength tubes or tuned boxes, depending on the target frequency. Each resonator was tuned for a narrow frequency band (quality ≈ 30) and placed at the corresponding pressure antinode in the room. The AR system used 172 channels to extend frequencies from 58 to 700 Hz. Parkin and Morgan stated that the RT of an empty auditorium can be extended by around 1.8 times in the 125 Hz octave band.

The Multichannel Amplification of Reverberation (MCR) system is an example of a regenerative AAES whose channels operated as isolated loops (de Koning, 1984). The MCR utilised a similar number of microphones to loudspeakers with an almost one-to-one mapping. For example, at the Phillips POC Congress Centre in Eindhoven, 90 microphones were mounted in the ceiling, which were routed to 110 loudspeakers distributed around the walls, ceiling, and stage area. The principle of this system was to form many closed microphone-loudspeaker loops, whose gain could be electronically increased to artificially extend the room's RT. Each closed loop exhibited unique resonances associated with the room modes at the transducer positions, so de Koning found that increasing the channel count reduced the overall variance of the system's frequency response. He stated that, to minimise colouration of the system, each channel must be responsible for no more than 0.8% of the reverberant energy. Strøm *et al.* (1986) implemented a 40-channel MCR system in a multi-purpose concert hall. The system sound was reported as satisfactory for symphonic music by musicians and the audience, but this was not confirmed by a formal listening test.

A virtual wall concept was used to create the Carmen system, whereby sound energy incident on a wall could be captured, altered, and re-emitted at a similar position to alter the perceived characteristics of the wall (Schmich and Vian, 2004). The Carmen system operated using 20–30 "cells," each consisting of a coincident microphone-loudspeaker pair and an electronic processing unit which did not introduce artificial reverberation. Since the cells at the boundaries faced inward, the system was regenerative by nature. These cells could then be used to artificially reduce the effective absorption of the walls, and delays could be introduced between the sound incidence and re-emission to increase the perceived distance of the walls. Reportedly, this system could naturally double the RT of a room with 24 cells (Schmich and Vian, 2004).

As an extension of the Carmen system, Jagla and Chervin (2020) created CarmenCita by adding artificial reverberation. CarmenCita used 16 Carmen-like cells

distributed around the walls and ceiling of the audience area, with an artificial reverberator inserted into the microphone-to-loudspeaker path of each cell. A global controller allowed the gains of each loudspeaker channel to be changed to adjust the spatial contribution of the AAES, for example, to increase lateral energy.

3. Hybrid systems

In 1990, the active field control (AFC) system was developed by Yamaha, which was one of the first AAESs to address the processing of early and late reverberation separately (Kawakami and Shimizu, 1990; Miyazaki *et al.*, 2003; Watanabe and Miyazaki, 2016). This system used an array of four to eight microphones above the stage area, the signals of which were reverberated using an early reflections generator and were reproduced laterally via loudspeakers on the walls near the stage (Miyazaki *et al.*, 2003). Another four to eight microphones were placed outside the critical distance of the stage sources, the signals of which were sent through a time-varying reverberator (described in more detail in Sec. II D) and were reproduced by loudspeakers on the walls and ceiling of the audience area as part of a regenerative subsystem. More recently, studies have been conducted with the AFC to compare methods of early reflections control (Hashimoto *et al.*, 2020), as well as to demonstrate the range of applications of AAESs (Hashimoto *et al.*, 2023).

Poletti (1998a) also mentioned the concept of a hybrid AAES that used both in-line and regenerative components to provide greater enhancement flexibility. Poletti then introduced the Variable Room Acoustics System (VRAS) (Poletti, 2006), a hybrid AAES which is now the basis of Meyer Sound's Constellation (Meyer Sound, 2024). The VRAS originally used approximately eight microphones above the stage and 16 microphones distributed in the ceiling above the audience, with 64 loudspeakers positioned in the ceiling, walls and under the balcony. Similar to the AFC system, a separate reverberator was used for the in-line early reflections processing and the late reverberation processing, which allowed the ratio of early-to-late reverberation to be changed in order to prioritise source clarity or LEV (Poletti, 2010). The main parameters of Constellation include the length of the VRAS reverberator, the strength of the system via μ (which also affects the regenerative coupling), and low- and high-frequency equalisation (EQ) (Ellison and Schwenke, 2025).

In 2019, Kaiser *et al.* (2019) provided an objective evaluation of the AAES by Amadeus Acoustics, which is a hybrid system that uses a regenerative component and artificial reverberation. They described two case studies: a multipurpose room with 14 microphones and 109 loudspeakers, and a concert hall using 32 microphones and 64 loudspeakers. The system allowed the user to select presets aiming to emulate different types of space. The Amadeus AAES is, in most cases, time-invariant; however, "optional microcrossfade phase modulation" is available for special installations, such as in museums (Werner, 2025).

Ambiance is another hybrid AAES, which was developed by L-Acoustics (Roskam *et al.*, 2022). Typically, 8–32 microphones are distributed above the stage and audience areas, the signals of which are processed independently by a gain-delay matrix and a reverberator. The signals are typically reproduced by 32–64 loudspeakers around the walls and ceiling of the audience area. Ambiance uses a relatively low number of channels compared to what has been seen in the past for AAESs with a regenerative component.

4. Signal processing

The SIAP system addressed the correction of insufficient passive room echo density, i.e., the number of reflections per second. This AAES was tuned *in situ* to generate the reflections missing from the acoustic space, which contributed to both the early and late reverberation echo density (Kok and Prinssen, 2003). The first 200–300 ms of the artificial energy decay could be modified in terms of level, delay and frequency-magnitude response, and the late reverberation could be changed regarding echo density (Prinssen and Holden, 1992). With a small number of input signals, the reverberator produced a relatively large number of decorrelated outputs, each routed to an individual loudspeaker.

The reverberator used in the RODS was a 72-tap digital delay line with each tap spaced by approximately 30 ms, resulting in around two seconds of sparse artificial reflections (Jaffe and Scarbrough, 1992). Each tap could also be filtered and adjusted in level, allowing different reflection patterns and absorption characteristics to be modelled. In contrast to the sparse delay lines of the RODS, the Lexicon processors of the LARES resulted in an exponential increase in echo density over time, which resembled natural acoustics (Griesinger, 1991). The Lexicon 480 L owner's manual states that the early processing consists of diffuse clusters of reflections with variable density (Lexicon, 2024).

In the CarmenCita system (Jagla and Chervin, 2020), the microphone signals were sent through a reverberator based on the work of Jot and Chaigne (1991), which described a range of reverberation algorithms, particularly the feedback delay network (FDN). At each loop iteration, the delayed signals from the previous iteration are further delayed and mixed, resulting in an exponential increase in echo density, as observed in LARES. The CarmenCita system also introduced EQ within the FDN. By filtering each channel in the feedback loop, the decay of energy over time became frequency-dependent where some frequencies decayed more rapidly than others.

In some cases, convolutional reverberators have been used to apply finite impulse response (FIR) filters instead of infinite impulse response (IIR) filters. An advantage of FIR convolution is that the IRs may be measured from real spaces, which is the case in the Vivace system (Müller-BBM Acoustic Solutions GmbH, 2023). Potential downsides

include increased computational complexity and non-trivial interpolation between reverberation settings.

The reverberators used in the AFC system were based on multi-tap fractional delay lines (Kawakami and Shimizu, 1990; Miyazaki *et al.*, 2003). After mixing eight microphone signals into four channels, each channel was written to a delay line (usually a circular buffer). By positioning several taps (read pointers) along each delay line, the channel signals can be read at a range of time delays, simulating artificial reflections. Due to the processing constraints in the 1990s, only around 30–100 delay taps were feasible per channel in an AFC system (Kawakami and Shimizu, 1990). The reverberated channels would then be routed to loudspeakers via a 4×32 gain-delay matrix (Miyazaki *et al.*, 2003).

The concept of geometric active acoustics has recently been promoted by Werner *et al.* (2023), whereby artificial reflections were generated according to an acoustic model of the installation space. Directional microphones were utilised to determine the angles of sound incidence, from which a vector model generated a set of early reflections, which were distributed to selected loudspeakers. At this point, DSP could be introduced to manipulate spatial and spectral attributes. The regenerative component of the AAES would then reverberate the generated reflections, and further algorithmic reverberation could be applied for more control.

The VRAS uses orthogonal matrices for both the early and late reverberation processing (Poletti, 2006). An orthogonal matrix is a real-valued square matrix, $\boldsymbol{A}$, such that $\boldsymbol{A}^T = \boldsymbol{A}^{-1}$ (Jot and Chaigne, 1991). Orthogonal matrices preserve total output power for constant input power and are regarded as multichannel all-pass filters that help define the stability condition of feedback systems (Poletti, 2010, 2006). The type of orthogonal matrix can have a significant effect on the resulting reverberation (Välimäki *et al.*, 2012).

The AFC system uses two applications of time-varying processing (Watanabe and Miyazaki, 2016). Within the reverberator, the read position of each fractional delay tap is modulated over a distinct time range and at a distinct rate. This shifts the system's eigenfrequencies over time, reducing the probability that a given frequency will be reinforced by feedback. Additionally, a so-called electric microphone rotator (EMR) modulates the microphone-to-loudspeaker mapping over time. It can be represented as a modulated weighting of different elements of the DSP TF matrix [$\mu\boldsymbol{X}(z)$ in Fig. 1], which modifies the magnitudes of the eigenvalues over time.

Using a time-varying reverberator allows fewer transducers to be used while maintaining acceptable colouration, and higher μ can be applied before instability (Griesinger, 1991). Griesinger reported that the LARES system provided at least 6 dB of extra stability headroom compared to time-invariant systems, and the SIAP system achieved a 3–6 dB GBI increase. However, Poletti demonstrated that this benefit diminishes for higher channel counts ($N \geq 16$) due to the inherent reduced risk of instability (Poletti, 2010, 2004).

III. EVALUATION OF ACTIVE ACOUSTICS

It is common for AAES evaluation to be conducted using room acoustic metrics such as T_{30}, early decay time (EDT) and sound strength (ISO 3382-1, ISO, 2009), as well as visual representations of RIRs. While these are useful for demonstrating enhancement performance, such as RT extension and speech intelligibility, the listener's impression should also be evaluated, especially since room acoustic metrics do not usually account for electronic artefacts that may be introduced by an AAES.

A. Colouration

Colouration is the perceptual phenomenon associated with distorting an audio signal from its natural frequency response (Rubak and Johansen, 2003). This distortion can occur acoustically, such as comb filtering caused by reflective surfaces near a listener; electronically, including the frequency-dependent impedances of an analogue audio circuit; and electroacoustically, for example, due to the frequency response of a mobile phone loudspeaker. Two prevalent types of colouration exist, namely, spectral and temporal colouration (Rubak and Johansen, 2003). The former refers to distorting a signal's frequency-magnitude response so that its timbre is noticeably different from what was expected. Most audio systems involve some form of spectral colouration, ranging from subtle differences among studio microphones to the strong band-limiting of a megaphone. This colouration is often perceived as unpleasant, but in some cases it is desired, such as in vintage audio equipment.

Temporal colouration refers to the distortion of the transient response of an audio system (Rubak and Johansen, 2003). An example is the build-up of diffuse reflections in reverberation, which can be perceived as time smearing, reducing speech intelligibility and signal clarity. This smearing can also be perceived as pleasant, contributing to the "luscious" quality of late reverberation. Another form of temporal colouration is known as ringing, where feedback systems with low damping can cause resonant oscillations. This is generally unwanted in audio systems because of the inharmonicity of resonances, which can become distracting in music and fatiguing for listeners, especially at high frequencies.

Regenerative and hybrid AAESs are particularly prone to ringing. When a sound is emitted inside an AAES, the spectral and temporal colouration of the passive room is first imposed on the signal via $\boldsymbol{H}_{\mathrm{SM}}(z)$. The frequency responses of the system microphones are also characterised by $\boldsymbol{H}_{\mathrm{SM}}(z)$, and the loudspeakers are similarly represented by $\boldsymbol{H}_{\mathrm{LR}}(z)$. The passive room and transducers also contribute to the colouration applied by the loudspeaker-to-microphone feedback path, $\boldsymbol{H}_{\mathrm{LM}}(z)$. The statistical properties of this TF matrix have a profound effect on the overall system colouration, due to the closed loop established by Eq. (8). When an audio system applies spectral colouration in a closed loop, an input signal is repeatedly subjected to the colouration

https://doi.org/10.1121/10.0043585

over time, which increases the variance of the frequency-magnitude response if the loop is time-invariant. This causes certain frequencies to decay more slowly than the mean, which results in ringing tones. This effect is exaggerated at low damping (high μ) because the decay rate of a given frequency decreases nonlinearly as its damping approaches zero. If the decay times of some frequencies are significantly longer than that of the mean energy, it is possible that the ringing tones will become the only remaining audible components of the signal.

Some studies have considered the statistical distribution of system frequency-magnitude responses for colouration analysis (Coleman *et al.*, 2024; Meynial and Vuichard, 1999; Mohlin and Höstmad, 2022; Poletti, 1998b). By using an ideal model of a diffuse room known as a Rayleigh model, a colouration target can be formed based on its RIR frequency density (Kuttruff, 2000). The standard deviation, σ, of the reverberant-field frequency-magnitude response describes how far certain frequencies stray away from the mean magnitude. The coefficient of variation can be found by dividing σ by the mean magnitude, m, which in the Rayleigh case is

$$\frac{\sigma}{m} = \sqrt{\frac{4}{\pi} - 1} \approx 0.523. \tag{18}$$

The coefficient of variation of any RIR frequency-magnitude response can be compared to this value to suggest how far from ideal its frequency-magnitude variability is. However, some assumptions of this measurement must be taken into account by truncating the RIR to ensure the diffuse field hypothesis holds, and limiting the frequency range between f_{sch} and the cutoff frequency of the measurement transducers (Meynial and Vuichard, 1999). It may be more appropriate to replace the mean with a smoothed moving average of the magnitude spectrum, since coarse variations in the magnitude response are expected in natural-sounding rooms. Hence, the coefficient of variation can be found by computing the standard deviation of the ratio of the raw and logarithmically-smoothed magnitudes (Meynial and Vuichard, 1999). Coleman *et al.* (2024) explored using this method to suggest how perceptually annoying the colouration of an AAES was. According to the linear regression of the listener ratings against the coefficients of variation, a latter value of 0.523 (the Rayleigh target) corresponded to colouration ratings between "inaudible" and "audible but not annoying," whereas coefficients of around 0.68 aligned with "annoying" colouration.

Some approaches to measuring sound colouration in rooms have been based on the cepstrum coefficients of an RIR, i.e., the inverse Fourier transform of the log-magnitude spectrum (Rindel, 2016). Recently, Neeten *et al.* (2025) proposed a cepstrum-based measure for AAESs that involves computing the ratio of the geometric and arithmetic means of the liftered (filtered in the quefrency domain) cepstrum coefficients. This was shown to increase GBI when used to inform closed-loop EQ, but the measure has not yet been validated perceptually.

Perceptual tests have been run to determine the objective factors affecting the colouration of electroacoustic systems. Early work on colouration included an investigation by Toole and Olive (1986) to determine the amplitude thresholds at which resonances became unnoticeable. Subjects listened to the summation of a dry programme item with a band-limited and time-delayed version, such that a certain frequency resonated. The subjects were asked to control the amplitude of the resonance until it "disappeared," which was repeated for frequencies between 130 Hz and 10 kHz, and for quality factors between 1 and 50. It was found that the amplitude of the resonance had more of an effect on the perception threshold than its duration. The authors commented that a high frequency resolution was required for objective measurements to identify resonances that were perceptually significant. For certain programme items, some "plainly visible" resonances were reportedly inaudible, which suggests the objective evaluation of temporal colouration might not be trivial.

Haeussler and van de Par (2015) investigated the perceptual effects of a convolution of two RIRs in terms of the resultant colouration. Subjects were asked to vary the colouration of a test room to match that of a stimulus, which consisted of a piano recording convolved with pairs of RIRs. The convolution of the first and second RIRs used by Haeussler and van de Par resulted in a slightly convex energy decay, and they found that the convolution worsened the perceived colouration of the piano, which supported objective results whereby the standard deviation of the log-magnitude spectrum increased by a factor of $\sqrt{2}$.

B. The double-slope effect

Sound energy in a room can be expected to decay exponentially (known as Sabine decay) if it has uniformly distributed absorption (Mortessagne *et al.*, 1993; Nilsson, 2004), as well as sufficient reflective randomisation (Joyce, 1978; Kanev, 2011). However, the acoustic energy decay of a space can consist of multiple exponential curves known as multi-exponential decay, or the DSE (Bradley and Wang, 2005). This can occur passively in coupled volumes where multiple spaces are connected via an aperture (Balint and Kaiser, 2018; Pu *et al.*, 2011; Xiang *et al.*, 2009), and in AAESs where the artificial reverberation acts as a secondary room.

In-line systems aim to reduce the significance of the feedback component $\boldsymbol{H}_{\mathrm{LM}}(z)$. This minimises the interaction between the active system and the passive room, hence a large difference may exist between the passive reverberation of $\boldsymbol{H}_{\mathrm{SR}}(z)$ and any artificial reverberation of $\boldsymbol{X}(z)$. The point at which the DSE can be seen with in-line AAESs is influenced by the ratio of RTs between the natural and artificial reverberation. Griesinger (1991) stated that a double slope decay may be observed when the enhancement critical distance[1] is greater than the passive critical distance of a source, $D^{\mathrm{S}}_{\mathrm{crit}}$. In a regenerative or hybrid system, the

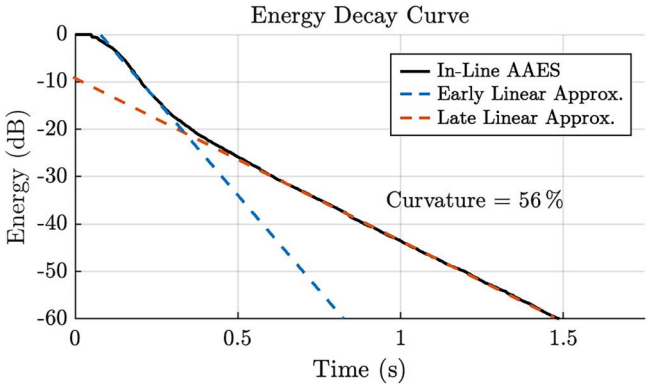

FIG. 5. Curvature estimation of a double-sloped EDC from a simulated 16×4 in-line AAES. The reverberator used an RT twice that of the passive room, and $\mu = -16$ dB.

relationship between the room and the reverberator is more complex due to regenerative coupling.

The degree of double sloping exhibited by an AAES can be determined by considering the RT of the system IR before and after the turning point of the energy decay (Cassidy *et al.*, 2024). The curvature is defined as the percentage increase in the late energy decay gradient with respect to the early decay, $|m_l/m_e - 1| \times 100\,\%$, where m_e and m_l are the early and late gradients, respectively, calculated using two linear approximations of the EDC, as demonstrated in Fig. 5.

Jagla and Chervin (2020) considered the active-passive coupling of a regenerative AAES that utilised artificial reverberation. They used the CarmenCita system, comparing an increase in master gain for a dry room versus a concert hall (mid-frequency RTs of 0.1 and 1.1 s, respectively). The RT of the reverberator was set to approximately 1.7 s, and the master gain was increased. It was found that in the dry room, the DSE was caused and the level of the secondary slope increased with the master gain, similar to an in-line system. In the concert hall, no significant double sloping was caused since the active and passive RTs were similar, and the master gain determined the *gradient* of the EDC.

C. Speech intelligibility

Speech intelligibility is imperative in classrooms, lecture theatres and cultural centres, and is even prioritised over sound quality in some public address systems (Svensson, 1994). It is also common for multi-purpose halls to be built for the minimum-energy context of speech communication, with the intention of using an AAES to provide additional reverberation for musical and theatrical performances (Poletti, 2011; Woszczyk, 2011). However, some passive spaces are unsuitable for speech, such as large lecture theatres with insufficient acoustic treatment, or rooms which are too dry to provide the necessary strength. In the former case, an AAES can be used for speech reinforcement rather than for RT extension.

The intelligibility of speech is mostly affected by the direct and "pseudodirect" energy of an RIR (0–20 ms) (Berkhout *et al.*, 1993), so applying artificial reflections within this timeframe requires inputs to the AAES to have a short initial time delay gap. Conveniently, personal microphones can be worn in most speech-related enhancement contexts, which allows short delays to be achieved. Similarly, it can be beneficial to place loudspeakers close to the audience to reduce the time-of-flight delay of the reproduction (Berkhout *et al.*, 1988).

The speech transmission index (STI) and rapid STI (RASTI) have been used to evaluate speech intelligibility in AAESs (Błasiński *et al.*, 2023; Haeussler and van de Par, 2019), where values greater than 0.6 are generally considered acceptable (Kaiser *et al.*, 2016; Pastusiak *et al.*, 2025). These metrics are based on modulation TFs, which describe the loss of detail from a source to a receiver for a given modulation frequency and octave band (IEC 60268-16; IEC, 2020). Clarity of speech (C_{50}) and definition (D_{50}) have also been used for evaluating electroacoustic systems in rooms (Haeussler and van de Par, 2019; Hatziantoniou *et al.*, 2013), which are based on the first 50 ms of the RIR (ISO 3382-1; ISO, 2009). To evaluate speech intelligibility via a listening test, the percentage of correctly understood words can be used to determine the speech reception threshold in different environments (Błasiński *et al.*, 2023; Haeussler and van de Par, 2019).

In applications where an RT extension is desired, AAESs have been shown to provide greater speech intelligibility than in equivalently reverberant passive spaces. Berkhout *et al.* (1988) investigated using the ACS to extend late reverberation and found that RASTI scores of around 0.5–0.6 could be achieved when extending the RT from 1.2 s up to around 4.5 s. This was most effective when the late energy was 20 dB below the early energy, indicating that the DSE contributed to this observation by preserving the early energy. Błasiński *et al.* (2023) found that extending the RT of a space beyond 2.5 s with an AAES resulted in higher speech intelligibility (both STI and listener ratings) than passive rooms of the same RT, and stated that the DSE was present. Utilising energy decay curvature when extending the RT has also been suggested to preserve musical clarity (C_{80}) (Hopper, 2012).

AAES enhancement of speech intelligibility for fixed RTs appears to be more challenging. Gade (1997) investigated using the speech preset of the LARES to provide a small number of strong early reflections for speech reinforcement. While the sound strength increased from 0.0 dB to 1.7 dB, the RASTI decreased from 0.62 to 0.57, although this is still borderline acceptable. Hatziantoniou *et al.* (2013) explored the effects of a distributed loudspeaker system in a classroom with poor speech intelligibility. Comparing single-loudspeaker reproduction to six loudspeakers, the STI remained low (around 0.4), but slight increases in C_{50} and D_{50} were observed for positions towards the rear of the classroom.

D. High-level perceptual attributes

Ko *et al.* (2012) conducted a listening test with 20 professional musicians, evaluating three conditions of an

https://doi.org/10.1121/10.0043585

IR-based AAES installed at McGill University, Canada. The first condition was the passive room only, the second used a moderate level of the AAES, and the third was the same as the second, but with a late reverberation gain of +3 dB. The musicians were asked to complete a questionnaire after performing as part of a string quartet in each of the conditions, evaluating 11 attributes including ease of hearing others, amount of reverberation and height sensation. Once the results were collected, Ko *et al.* used principal component analysis (PCA) to extract groups of scored attributes that were influenced in similar ways, revealing three main components: spatial energy, support and tonal quality. They also found that attributes related to support and clarity explained 68% of the variance in the musicians' responses, suggesting these qualities of a space have a profound impact on performers.

A similar investigation into acoustic factors affecting performers was conducted by Rosseel and van Waterschoot (2023). This study asked four vocalists to interact with a virtual rehearsal space (essentially an in-line AAES), and a total of 24 attributes were elicited from them through a rapid sensory analysis technique known as flash profiling (Zacharov, 2018). The vocalists were then asked to rate four room models using their chosen descriptors. From their ratings, the attributes were analysed using agglomerative hierarchical clustering, a method that groups parameters based on their influence, similar in concept to PCA. The clusters were identified as reverberance, directness and spectral content. These underlying components are similar to those found by Ko *et al.* (2012), where spectral content can be linked to tonal quality, reverberance to spatial energy and directness to support.

Gade (1997) explored the objective and subjective effects of the LARES installed in a concert hall in Vejle, Denmark. Two sessions were conducted in which musicians performed for an audience under different acoustic conditions, and both the musicians and the audience rated the reverberation and degree of realism on a questionnaire. The musicians were additionally asked to rate the support of the space, and the audience was asked to rate timbre. The acoustic conditions were varied by selecting different presets of the LARES (e.g., "Large Hall" and "Chamber"), of which some objective measurements were presented by Gade, but it is unclear as to whether these settings varied in other respects, such as lateral energy fraction. The musicians' ratings revealed a strong positive correlation between the support ratings and the reverberance of the spaces. It was also found that an increase in RT of around 50% relative to passive RT was the limit of realism for the musicians, suggesting it was similar to what would be possible with passive acoustic enhancement. The audience rated the LARES system presets as more realistic on average than the musicians, and the point at which the reverberance became "excessive" was around twice the passive RT, which was 50% higher than the musicians'.

The listening experiment by Ko *et al.* (2012) was later repeated by Chon *et al.* (2015) with some changes. Instead of using live musicians, recordings of string quartets were made using a head-and-torso simulator under the same acoustic conditions as before. These recordings were then played to 12 audio engineers in random order, rather than in increasing RT as Ko *et al.* did. Chon *et al.* asked listeners questions about naturalness, source distance, room size, clarity, loudness, and preference. No significant differences were found in naturalness ratings across conditions, but perceived size increased with stimulus RT, suggesting that listeners perceived the enhanced acoustic conditions with similar naturalness to the passive space, despite their noticeable differences. Chon *et al.* also found that listeners preferred the second condition (moderate enhancement) overall, whereas Ko *et al.* found that the moderate and extended AAES conditions were rated equally in terms of preference by performers.

Lokki *et al.* (2007) used multiple time-varying AAESs in a room to vary the RT across different dimensions, thereby changing the shape of the perceived space. This concept was validated by asking ten listeners to make sounds in the centre of a room (clapping, talking, singing, etc.), after which they would draw the shape and size of the perceived space. Different artificial room shapes were created by varying the RTs of four independent AAESs positioned on the walls. When long RTs were applied to each wall, the listeners drew large rectangles and squares that aligned with the virtual aspect ratio of the RTs. When the system was turned off, multiple participants appeared to perceive the space as smaller than the physical room, which Lokki *et al.* suggested was due to the surprisingly dry nature of the passive space. An open wall condition seemed to have caused more confusion than the other conditions, but multiple subjects positioned the front wall closer to the listener than the back, as expected.

IV. SIMULATIONS

The literature reviewed previously has revealed numerous approaches to AAESs and reverberators. In this section, simulations are presented to illustrate the influence of the reverberator parameters, EQ, passive room characteristics and transducer directivities on AAES performance. The AAESToolbox[2] provides a MATLAB environment to generate AAES simulations based on the signal model from Sec. II A and to visualise the results. A single shoebox room was defined as a base simulation, and parameters were varied individually to highlight their effects. The simulations in this section were generated under the following assumptions: the passive room and transducers are LTI; the passive RIR becomes a decaying Gaussian random process after the perceptual mixing time (t_{mp}); and all transducers have a flat frequency response. Further information on approximations is provided throughout this section.

A. Simulation workflow

The passive room TF matrices $\boldsymbol{H}_{\mathrm{SM}}(z)$, $\boldsymbol{H}_{\mathrm{SR}}(z)$, $\boldsymbol{H}_{\mathrm{LM}}(z)$ and $\boldsymbol{H}_{\mathrm{LR}}(z)$ were simulated in the time domain using

AKtools (Brinkmann and Weinzierl, 2017), approximating the passive room as an LTI system. A hybrid approach was used to combine the image-source method (ISM) with stochastic reverberation, simulated according to Borß and Martin (2009). The ISM was applied up to $2t_{mp}$, using $t_{mp} = 20V/S + 12$ where V denotes the room volume and S denotes its surface area (Lindau *et al.*, 2012). Since the room size was consistent across simulations, $2t_{mp} \approx 120$ ms in all cases. The stochastic reverberation was faded in linearly from the first reflection of the ISM to $2t_{mp}$. Since $f_{sch} < 30$ Hz in all simulations, which is below the reproduction capability of loudspeakers typically deployed for active acoustics, modelling low-frequency modal behaviour was not considered necessary. However, in smaller rooms, more accurate simulations could be applied, such as using a wave-based solution for low frequencies. Absorption coefficients were specified in octave bands for each surface, and the microphones and loudspeakers used ideal cardioid and omnidirectional polar patterns with flat frequency responses. To compensate for the lack of diffusion modelling, the transducer positions were initially randomised within a 10 cm cube centred on their initial positions, but fixed in all simulations.

Once the passive RIRs were simulated, the DSP unit [$\boldsymbol{X}(z)$ in Fig. 1] was generated for multiple time-invariant cases using the FDN toolbox (FDNTB) (Schlecht, 2020) and saved as IRs. For each AAES condition, the system IRs were loaded, and the DFT of the single-channel receiver RIR was calculated according to Eq. (4) using $\boldsymbol{u}(z) = 1$ and $z = e^{j\omega_k}$. The loop gain μ was defined relative to the GBI [estimated using Eq. (14)] unless stated otherwise. For the time-varying reverberator cases, this was applied using a time-domain implementation of the AAES simulation so as to define the modulation state at each time step. This still yields an LTI approximation of the AAES, but eigenvalue modulation during the simulation captures the effects of time variation. All IRs were written into 32-bit 48 kHz waveform files.

The base AAES simulation was defined for comparison throughout this section, and the controlled variables in each simulation used the base configuration unless stated otherwise. The base room was a shoebox with dimensions $35 \times 18 \times 8$ m (length, width, and height, respectively). The absorption coefficients were chosen by selecting one material for the four walls, one for the floor and one for the ceiling using the data from Vorländer (2008). Each of the seven octave bands was then multiplied by scale factors to approximately match the RT of each band in a reference theatre measurement. The resulting absorption coefficients are displayed in Table II, as well as the RTs of each octave band.

A 16×16 regenerative AAES was defined around the walls and ceiling of the audience area, shown in Fig. 6, with mannequins representing the source and receiver in the stage and audience areas, respectively. The transducer positions were based on typical audience-focused installations, similar to those used in previous work (Cassidy *et al.*, 2024), in which the coverage did not include the stage area. The source and receiver were both omnidirectional, positioned at a height of 1.2 m in the centre of the stage and audience regions, 17.5 m apart. All loudspeakers were cardioid, whereas the six overhead microphones were omnidirectional, and the wall microphones were cardioid. Blue circles indicate the front of each transducer in Fig. 6. Full coordinates, directivities and rotations can be found in the AAESToolbox.[2]

TABLE II. Absorption coefficients and T_{30} values of the base simulation.

	Octave Band Centres (Hz)						
	125	250	500	1 k	2 k	4 k	8 k
Walls	0.46	0.75	0.55	0.49	0.65	0.60	0.76
Floor	0.20	0.25	0.28	0.36	0.33	0.46	0.55
Ceiling	0.70	0.55	0.42	0.43	0.70	0.57	0.65
T_{30} (s)	0.82	0.74	0.87	0.83	0.66	0.62	0.53

The DSP unit in the base AAES was simply an identity matrix $\boldsymbol{X}(z) = \boldsymbol{I}_{16}$, meaning that no artificial reverberation was applied, and each microphone was routed directly to its neighbouring loudspeaker. The μ was set to -5 dB to allow RT extension to be observed with minimal temporal colouration. All channels were unequalised in the base condition.

B. Reverberator

This section explores the influence of the loop gain μ, the ratio of the artificial RT relative to the passive room, and time-varying modulation.

1. Loop gain (μ)

Using the base AAES condition described in Sec. IV A, RIRs from the receiver position were simulated for $\mu = -\infty$ dB (system off), -4 and -2 dB. Spectrograms of the simulated RIRs are shown in Fig. 7, where the extension of power across time due to regeneration can be observed as μ increases, indicated by the increasing wideband T_{30} (dashed). The RT extension observed in (b) versus (a) is similar to decreasing the absorption coefficients of the passive room. However, as μ is increased to -2 dB in (c), colouration becomes apparent (see Sec. III A) as an oscillation emerges around 800 Hz. The issue with such a condition is twofold: the resonance may be perceived as unpleasant ringing, and the mean RT extension across frequency becomes limited as μ increases past this point. This is because if the mean energy is far from instability relative to a narrow band, then due to the nonlinear extension of resonances towards instability, the mean will not exhibit much extension before the band becomes unstable.

2. RT ratio

RT ratio refers to the ratio of T_{30} values between the AAES reverberator and the passive room, where a ratio of 1 indicates matched RTs. To demonstrate the effect of RT ratio on the energy decay, a simulation was performed using the base condition but with an LTI FDN inserted as $\boldsymbol{X}(z)$.

https://doi.org/10.1121/10.0043585

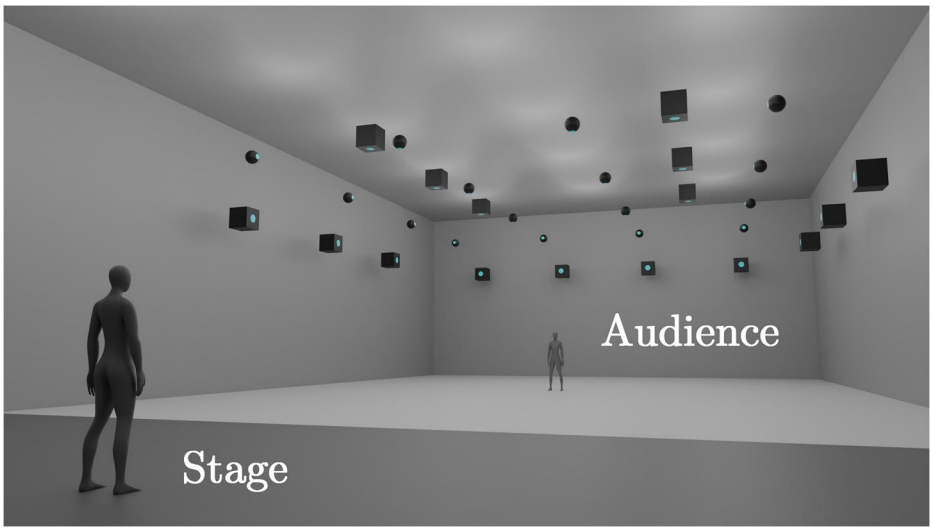

FIG. 6. The 16×16 regenerative simulation base condition, depicting the source position (left mannequin), receiver position (right mannequin), microphones (small spheres), and loudspeakers (large cubes). Blue dots indicate transducer orientation.

This was generated using the FDNTB (Schlecht, 2020), where the RT at DC was set to the product of the passive room RT at 1 kHz (0.868 s) and the RT ratio. The RT at the Nyquist frequency (24 kHz) was set to half of the RT at DC. This produced a square matrix of uncorrelated, frequency-dependent IRs with uniform weighting such that each microphone was reverberated and routed to each loudspeaker. Figure 8 presents the EDCs of the simulations with RT ratios varying from 0.5 to 2.0, shown alongside the passive room (dashed). As the RT ratio increases, the energy decay exhibits the DSE more distinctly (see Sec. III B). Note that the secondary gradients pivot about a common point, known as the turning point (Atalay *et al.*, 2022), which is largely determined by the level of the AAES at the receiver relative to the passive room. Each condition used the same absolute loop gain, $\mu = -68$ dBFS (decibels full-scale, i.e., relative to digital clipping), to preserve the turning point. If μ were to increase, this turning point would translate towards the origin.

It has been suggested that the DSE begins to occur when the RT ratio exceeds unity (Poletti, 2006), and the naturalness of reverberation can be negatively impacted beyond a ratio of 2 (Cassidy *et al.*, 2024). In Fig. 8, RT extension is visible in all cases, but the Sabine decay is mostly preserved for ratios ≤ 1.

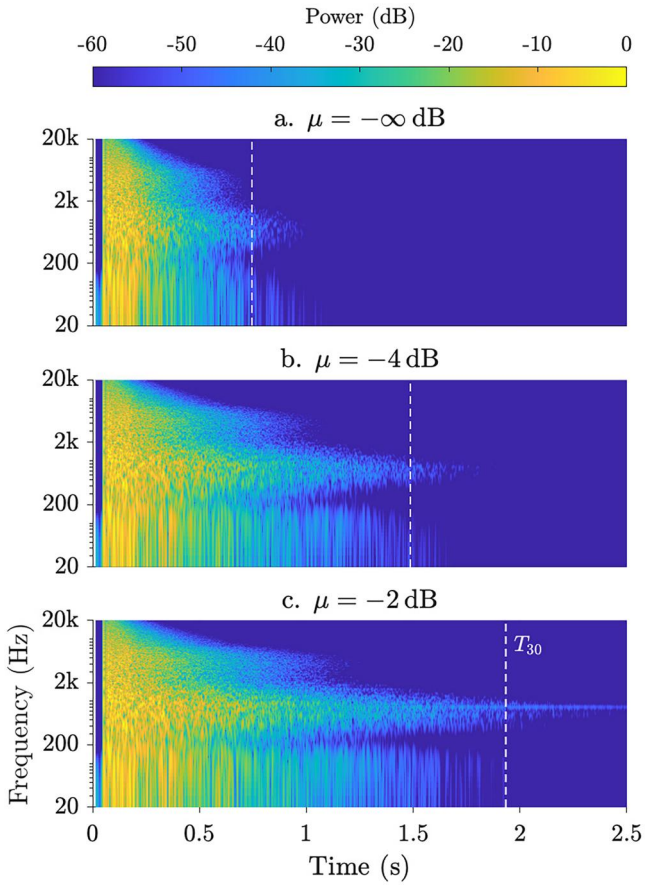

FIG. 7. Power spectrograms of the 16×16 pure regenerative AAES with increasing loop gain μ. Wideband T_{30} values are indicated by dashed lines.

3. Time variation

Using a time-varying reverberator can provide effective suppression of temporal colouration (Griesinger, 1991;

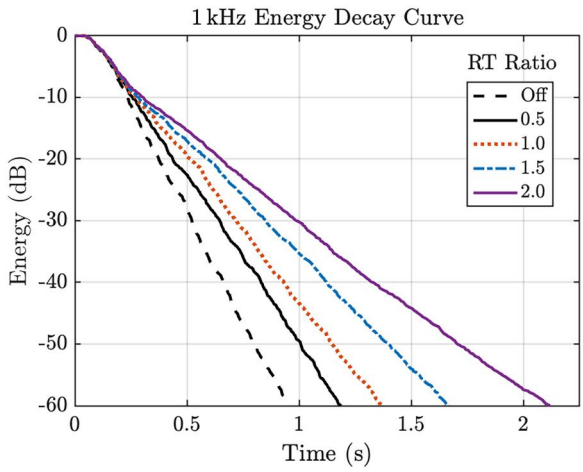

FIG. 8. EDCs of the regenerative AAES receiver RIRs using an FDN reverberator with increasing RT ratio for a fixed absolute loop gain. Each EDC represents the 1 kHz octave band.

Miyazaki *et al.*, 2003), which can lead to a greater average energy extension. However, high degrees of amplitude and pitch modulation can prove detrimental to the listening experience, so time variation should be used in moderation (Poletti, 2000). To demonstrate the benefits of time-varying processing, reverberators were injected into a 16×4 in-line AAES configuration, similar to the layout in Fig. 6, but using only the four frontal microphones translated towards the stage. Given the nature of an in-line system, μ is typically low to avoid feedback. In this simulation, μ is initially set to -3 dB to cause severe colouration for illustrative purposes. The FDNTB (Schlecht, 2020) was used to produce time-invariant and time-varying reverberators with an RT ratio of 2, and frequency dependence as described in Sec. IV B 2. For the time-varying condition, the following parameters were set: `modulationFrequency` $= 0.05$ Hz; `modulationAmplitude` $= 0.3$; `spread` $= 50$ (for more information, see Schlecht, 2020).

The spectrograms of the time-invariant and time-varying conditions are compared in Fig. 9. The time-varying condition used the same *absolute* μ as the time-invariant condition (-31.3 dBFS) to ensure the change in colouration was due to the time variation alone. Figure 9(b) shows the significant resonance mitigation as a result of the time variation, and Fig. 9(c) applies the same *relative* μ to the time-varying condition as the time-invariant condition to demonstrate the effective reverberation extension.

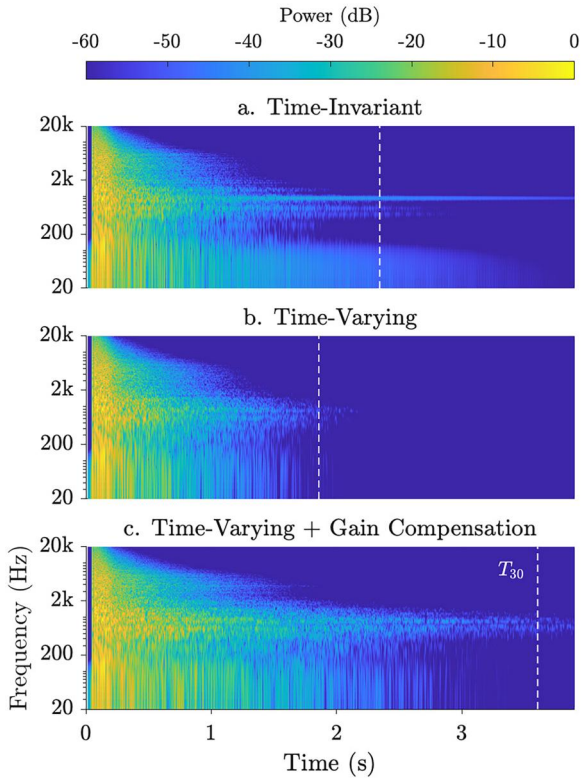

FIG. 9. Power spectrograms of the 16×4 in-line AAES using an FDN reverberator with time variation disabled (a) and enabled (b). Part (c) applies loop gain compensation to (b) to match the extension observed in (a). Wideband T_{30} values are indicated by dashed lines.

When comparing Figs. 9(a)–9(c), two observations can be made. First, the temporal colouration appears low after time-varying processing, with the horizontal streaks in Fig. 9(a) more spread across frequency in Fig. 9(c). Second, the mean energy is greater in Fig. 9(c), indicated by the increase in wideband T_{30} (dashed), despite the system operating at the same μ as (a). This can be explained by the resonances in Fig. 9(a), where the eigenvalue magnitudes are far from the mean, limiting how close to instability the mean can get. In Fig. 9(c), the eigenvalue maxima are suppressed over time, allowing the mean to be brought closer to instability.

Although this example of time variation seems ideal, there may be perceptual artefacts that are not represented by spectrograms. For example, the system response to a sinusoidal input may cause audible frequency modulation, amplitude modulation, or frequency spreading (i.e., the signal getting noisier over time). The degree of time variation must be tuned accordingly.

C. EQ

EQ can be a valuable tool in AAESs for reducing sources of colouration. The frequency responses of microphones and loudspeakers used in an AAES contribute to the colouration perceived by the listener, especially when regeneration is present (see Sec. III A). By equalising the transducers independently from the room they are situated in, for example, by analysing them in an anechoic chamber, electronic colouration can be reduced while preserving the frequency-dependent impedances of the passive room. When considering colouration due to electroacoustic feedback, however, the frequency response characteristics of the passive room have a significantly greater influence than the transducers. Thus, EQ may be used to attenuate specific problematic frequencies in the feedback loop, or the open loop may be inverse-filtered to increase the available GBI across all frequencies. The latter results in a corrective whitening of the magnitude response of $\mu \boldsymbol{X}(z)\boldsymbol{H}_{\mathrm{LM}}(z)$ before applying creative EQ to achieve more realistic tonal characteristics and frequency-dependent energy decay.

One approach to the inverse filter problem equalises every microphone-to-loudspeaker routing [i.e., each element of $\boldsymbol{X}(z)\boldsymbol{H}_{\mathrm{LM}}(z)$] independently. This is useful for AAESs utilising time-varying processing based on routing modulation (such as the EMR used in the AFC system; Watanabe and Miyazaki, 2016), since any permutation of inter-channel mixing would remain as frequency-independent as possible. Third-octave EQ resolution was common in early systems (Prinssen and Holden, 1992), but higher-order filters are more computationally feasible today (De Bortoli *et al.*, 2024a). Another way to apply EQ is to analyse a specific AAES condition by performing the eigenvalue decomposition of the open-loop TF matrix as described in Sec. II A. Then, a global EQ curve can be applied to every element of $\boldsymbol{X}(z)$ such that the GBI is minimised across frequency (De Bortoli *et al.*, 2024a).

https://doi.org/10.1121/10.0043585

To demonstrate the global filtering method, the regenerative base condition [$\boldsymbol{X}(z) = \boldsymbol{I}_{16}$] was used with $\mu = -2$ dB to induce temporal colouration. Two second-order bi-quadratic IIR filters were cascaded to form $H_{\mathrm{EQ}}(z)$, which was applied to every channel such that $\boldsymbol{X}(z) = H_{\mathrm{EQ}}(z)\boldsymbol{I}$. The coefficients of $H_{\mathrm{EQ}}(z)$ were tuned to attenuate the most significant resonances at the receiver position. Figures 10(a) and 10(b) show spectrograms of the receiver RIRs before and after applying $H_{\mathrm{EQ}}(z)$ [magnitude response in Fig. 10(e)], respectively, both using the same absolute μ (−34.2 dBFS). Figure 10(c) shows the equalised condition when matching the relative μ of Fig. 10(a). This process has a similar effect to applying time variation (see Sec. IV B 3), whereby reducing resonances provides greater headroom for enhancement. The wideband T_{30} (dashed) slightly increases from Figs. 10(a)–10(c), and using a higher-order filter would improve this effect.

Figures 10(d) and 10(f) present the open-loop eigenvalues for the unequalised and equalised conditions, respectively, normalised to 0 dBFS. The magnitudes of the eigenvalues are presented as dots, and the phases of the 35 largest-magnitude eigenvalues are displayed as semi-transparent circles. Before EQ, the largest-magnitude eigenvalues are clustered around 800 Hz in Fig. 10(d), whereas after EQ, the maxima become more evenly distributed across frequency. Similarities can be drawn between the eigenvalue magnitudes in Figs. 10(d) and 10(f) and the resonant characteristics of the spectrograms in Figs. 10(a) and 10(c), where eigenvalues near a magnitude of 0 dBFS and 0 rad phase in the former correspond to significant resonances in the latter. Interestingly, the longest-lasting resonance in Fig. 10(c)—around 800 Hz—does *not* correspond to the greatest-magnitude eigenvalue in Fig. 10(f), but rather an eigenvalue with near-zero phase (marked with an asterisk). This emphasises the necessity to consider the phase during stability analysis, which reiterates why more accurate GBI estimation can be achieved by analysing the real part of the open-loop eigenvalues rather than the magnitude (see Sec. II B).

D. Passive acoustics

The reverberation of an AAES is influenced by passive acoustics, especially in a pure regenerative system, i.e., when using a gain-delay matrix instead of a reverberator. As a room approaches anechoic conditions, the reverberance of $\boldsymbol{H}_{\mathrm{SR}}(z)$, $\boldsymbol{H}_{\mathrm{LR}}(z)$, $\boldsymbol{H}_{\mathrm{SM}}(z)$, and $\boldsymbol{H}_{\mathrm{LM}}(z)$ decreases, and it becomes increasingly difficult to place microphones outside of $D^{\mathrm{S}}_{\mathrm{crit}}$ and $D^{\mathrm{L}}_{\mathrm{crit}}$, thus limiting the available natural decorrelation. Installing an AAES with its microphones within $D^{\mathrm{L}}_{\mathrm{crit}}$ means, by definition, that a maximum of 50% of the energy incident at the microphones due to the loudspeakers will be reverberant, which can limit the available RT extension.

To demonstrate this, the base room was compared to a more reverberant room formed by halving each absorption coefficient in Table II. Two reverberation conditions were applied to each room. The first was a pure regenerative condition, where the processor was defined as a Hadamard matrix to provide a fully dense routing between all transducers, but with orthogonality for stability control (see Sec. II E 4). The second utilised artificial reverberation, which was formed using the previous condition, but where the processor was an FDN with fully dense routing. The FDN parameters were consistent with those used in Sec. IV B 2, where the RT ratio was 2.0 for each passive room condition. Both active conditions used the base loop gain of $\mu = -5$ dB.

The EDCs are presented in Fig. 11 for the system off (dashed), a pure regenerative AAES (solid), and the regenerative condition with artificial reverberation (dotted). The pure regenerative condition in Fig. 11(b) (solid) shows a

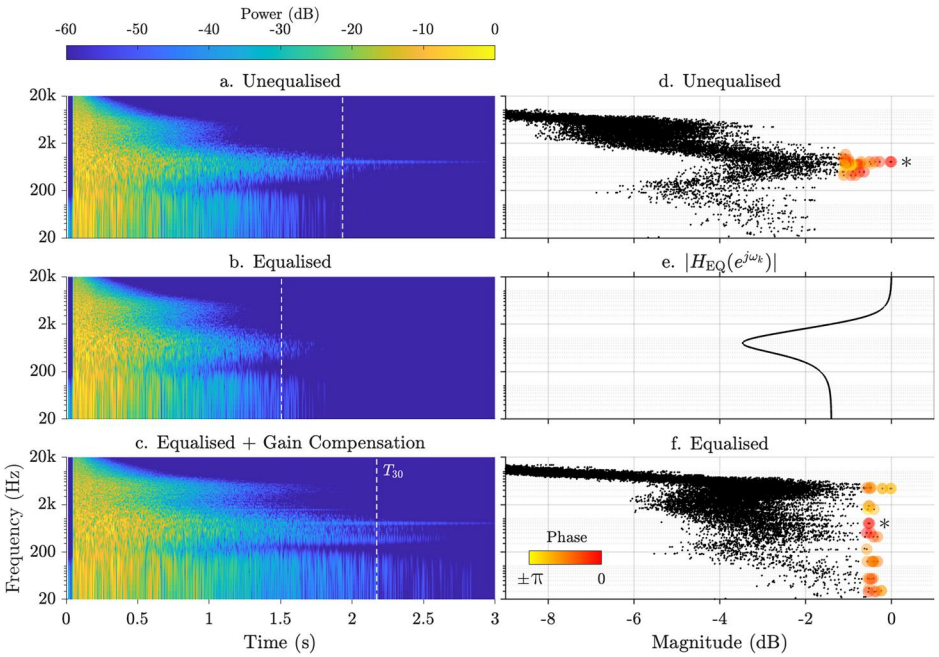

FIG. 10. The effects of closed-loop EQ on the pure regenerative AAES condition (no artificial reverberation). Parts (a), (b), and (c) show power spectrograms of the receiver RIR from the pre-EQ, post-EQ and post-EQ with gain compensation conditions, respectively, where (a) and (c) use $\mu = -2$ dB. Parts (d) and (f) show the pre- and post-EQ eigenvalue magnitudes $|\lambda_p(e^{j\omega_k})|$ of the open-loop TF $\mu\boldsymbol{X}(z)\boldsymbol{H}_{\mathrm{LM}}(z)$ where $p = 1, \ldots, 16$, and $\omega_k = 2\pi k/K$ for $k = 1, \ldots, K$ where K is the number of frequency bins. Phases of the 35 maxima are identified as semi-transparent circles, and the eigenvalues corresponding to the longest resonances are marked as asterisks in (d) and (f). Part (e) displays the magnitude of the two cascaded second-order IIR filters, $|H_{\mathrm{EQ}}(e^{j\omega_k})|$. All axes are aligned between adjacent plots. Wideband T_{30} values are indicated by dashed lines.

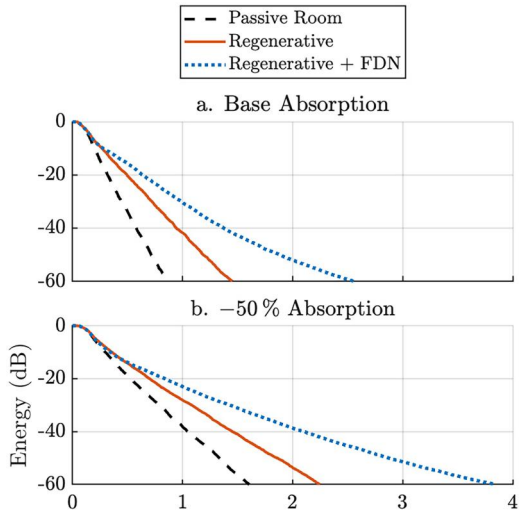

FIG. 11. EDCs of passive rooms (dashed), 16 × 16 pure regenerative systems (solid), and 16 × 16 regenerative systems using FDN reverberators (dotted) for the absorption conditions (a) and (b).

similar RT extension in the more reverberant space as in the base room, while the regenerative with FDN conditions demonstrate how significant RT extensions can be produced using the same μ values as the pure regeneration. The conditions applying the FDN show a slight curvature due to the interaction between the reverberator and the regeneration, which is not observed in the pure regenerative conditions. Therefore, if Sabine decay is desired, a regenerative AAES with little or no artificial reverberation may be preferable at the expense of limited RT extension. Alternatively, more channels could be introduced.

It is also important to note that the RT can sometimes be influenced by the emergence of strong resonances, causing colouration. To illustrate this, Fig. 12 shows the colouration values of the base room and the reduced-absorption space for increasing μ using the measure by Meynial and Vuichard (1999). The base room is seen to exhibit a greater rate of increase in colouration as μ increases. Therefore, further colourless RT extension can be achieved in more reverberant spaces with increasing μ as a result of the improved closed-loop decorrelation provided by the more reverberant space.

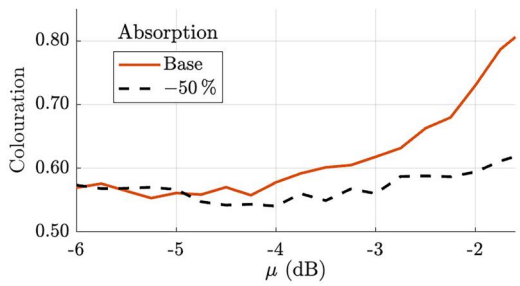

FIG. 12. Colouration with increasing loop gain μ relative to the GBI for the two absorption conditions using the pure regenerative AAES. The colouration measure follows Meynial and Vuichard (1999).

E. Microphone and loudspeaker layout

Increasing the number of channels in a regenerative system can reduce the severity of colouration by lowering the gain required by each closed loop to achieve the same power output (Griesinger, 1991). At the limit of stability, only the most prominent eigenfrequencies will oscillate with low damping factors as the rest of the magnitude response decays (Poletti, 2010). Thus, if each closed loop operates further below its limit of stability, the energy will decay more uniformly across frequency. Therefore, maximising the channel count can result in a smoother overall frequency-magnitude response, allowing greater RT extensions without colouration.

A higher number of channels can also reduce the variability of the system resonance, since room modes vary in strength across the room. The resonant characteristics of a closed loop in an AAES are influenced by the position and routing of its transducers, which may consist of independent microphone-loudspeaker loops (de Koning, 1984), or there could be a more complex mapping between multiple microphones and loudspeakers (Miyazaki *et al.*, 2003; Poletti, 2006). By increasing the number of these channel loops with different transducer positions, more variations of resonances contribute to a more uniform energy decay across frequency (Poletti, 2010).

It is common to use small-diaphragm condenser microphones in active acoustics due to the typically flat frequency-magnitude response, high sensitivity, and low visual obtrusion. Utilising directional microphones can also prove useful at the expense of colouration for off-axis incidence, where minimising feedback is crucial, i.e., for in-line systems. Typical directivities include cardioid (Griesinger, 1991), supercardioid (Prinssen and Holden, 1992) and hypercardioid (Poletti, 2006) patterns which can effectively increase D^{S}_{crit} by rejecting reflected energy. Directional microphones can also increase the GBI of an AAES by rejecting direct loudspeaker energy, reducing the risk of colouration and instability due to increased decorrelation. Pressure-operated (omnidirectional) microphones are commonly found in regenerative systems, but directional microphones can still be valuable when microphones are in close proximity to associated loudspeakers, such as in virtual wall systems (Jagla and Chervin, 2020; Schmich and Vian, 2004).

To demonstrate the effect of microphone directivity in an AAES, the base simulation was modified into an 8×8 pure regenerative system, depicted in Fig. 13(a). The system microphones, shown as small spheres, were positioned immediately below the loudspeakers, facing downwards. The routing of microphones to loudspeakers was such that each microphone only fed its nearest neighbour, resulting in eight single-channel loops. All loudspeakers were modelled with cardioid directivity, and the microphones were switched between omnidirectional and cardioid patterns. This simulation used the more absorptive set of absorption coefficients from Sec. IV D to increase the contrast in DRR.

https://doi.org/10.1121/10.0043585

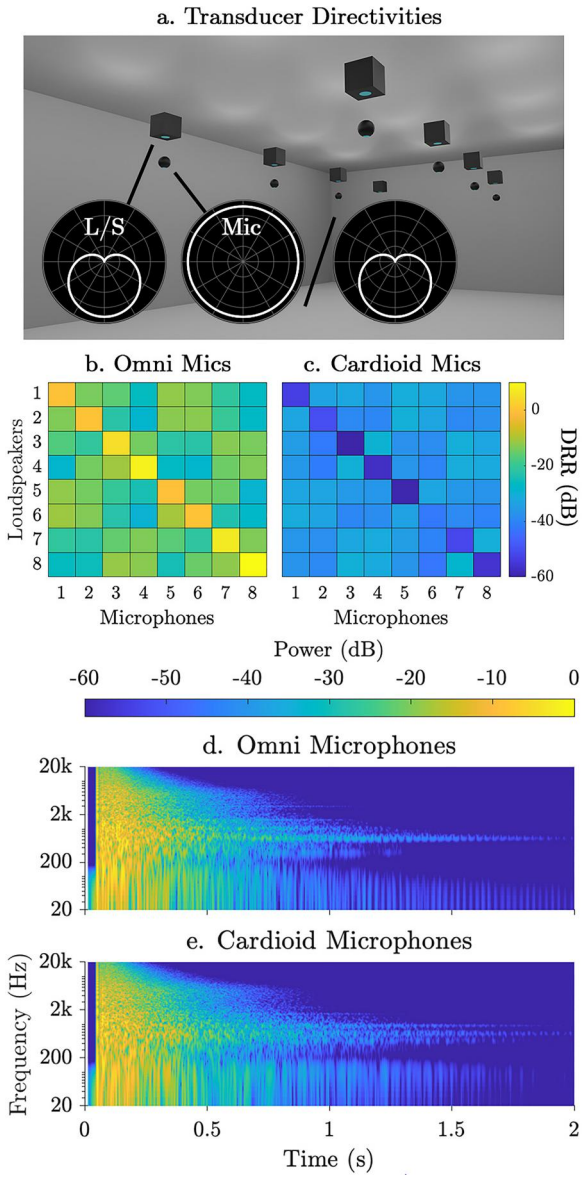

FIG. 13. Simulation configuration diagram (a), DRR heatmaps (b) and (c), and spectrograms (d) and (e) for AAES simulations using omnidirectional versus cardioid microphones. The diagonal elements of (b) and (c) represent the routing of each microphone to its neighbouring loudspeaker.

Figure 13(b) and 13(c) present the DRRs of each loudspeaker-to-microphone path for the cardioid and omnidirectional conditions, respectively.

The diagonal elements exhibit a maximum difference of around 70 dB between the conditions, indicating a high level of rejection of the neighbouring loudspeakers in the cardioid condition. When the microphones are omnidirectional, the DRR of non-diagonal elements is approximately 20 dB greater than the cardioid condition, since the off-axis cardioid incidence also provides some direct signal attenuation. The spectrograms of the receiver RIRs for each condition are shown in Fig. 13(d) and 13(e) using $\mu = -2$ dB. Some of the visible resonances in Fig. 13(d), such as 30 and 600 Hz and 3 kHz, become more spread over frequency when the microphones are switched to cardioid Fig. 13(e). This is because when the DRR of a loudspeaker-to-microphone path decreases, the peak of its frequency-magnitude response relative to the mean also decreases [Vuichard and Meynial, 2000; Fig. 1(c)], which makes the formation of large peaks and valleys during feedback more likely.

V. SUMMARY AND FUTURE PERSPECTIVES

In this article, prior sections have identified the distinct structures of AAES types, commercial implementations, and the artefacts of such systems from a perceptual perspective. Simulations show the effects of varying common AAES parameters. To conclude, the common design choices are justified, organised by high-level topology based on Secs. I–IV. The current research trajectories in active acoustics and motivation for future work are also discussed.

A. Summary of design choices

The previous simulations have demonstrated the complexity of the design space for AAESs, being influenced by the system topology, microphone directivity, reverberation processing and passive room acoustics. Three main AAES topologies characterise the main capabilities and decisions for AAES designers, especially regarding their capacity to affect the early energy and the inherent electroacoustic coupling in the design.

In-line AAESs typically use a small number of microphones near the stage area of an auditorium to capture stage sources with a high DRR. These signals reach the microphones after a relatively short time of flight and are then reflected and reproduced around the walls and ceiling of the audience area. Due to the direct and low-latency properties of the microphone signals, early reflections can be manipulated to modify features such as ASW, intimacy and speech intelligibility. Directional microphones are commonly utilised, pointing towards the stage to avoid feedback by reducing the strength of $\boldsymbol{H}_{\mathrm{LM}}(z)$. Some in-line reverberators utilise time-varying processing to further reduce the risk of instability and colouration, but artefacts may be noticeable, like pitch shifting. Full-length artificial reverberation has been used in some in-line instances, but they can be prone to the DSE if the reverberator RT is greater than that of the passive room.

Regenerative AAESs use more microphones placed outside $D^{\mathrm{S}}_{\mathrm{crit}}$ to ensure the system input is reverberant. The microphones are also placed outside $D^{\mathrm{L}}_{\mathrm{crit}}$ to ensure signals in the feedback loop are sufficiently decorrelated, which is crucial for colourless regeneration when artificial reverberation is not being used. Due to the low DRR of the input signals and the longer time-of-flight from the stage to the microphones, regenerative systems are better suited to enhancing late reverberation, where features such as lateral energy fraction and LEV can be changed. The DSP unit $\boldsymbol{X}(z)$ in a regenerative system may be as simple as a gain-delay matrix, which can extend the RT of the room while avoiding the DSE, but with limited effect. Artificial reverberation can be introduced to extend RT while keeping μ constant, and to provide further control over perceptual

attributes. To encourage significant regeneration, μ must be relatively high, which can cause colouration and ringing. Hence, EQ and time-varying processing can be useful for increasing the GBI, allowing a greater RT extension before colouration. Orthogonal feedback matrices have also been used to clearly define the stability condition for regenerative AAESs.

Hybrid topologies aim to provide greater flexibility by featuring directional stage microphones that feed an in-line subsystem and microphones distributed around the concert hall that populate a regenerative subsystem. This allows both early and late energy to be controlled through separate processing while utilising common system components, such as loudspeakers. Table I presents an overview of AAES installations detailing their year of publication, topology, and reverberation type.

B. Discussion of future perspectives

Recent studies in active acoustics have focused on improving the perceptual understanding of these systems, exploring stability control and optimisation, and their performative and compositional applications. The influence of the DSE on listeners of an AAES has been investigated by Cassidy *et al.* (2024) and Błasiński *et al.* (2024). Cassidy *et al.* found that naturalness decreased significantly when the reverberator RT exceeded twice the passive room RT, resulting in pronounced EDC curvature. This detrimental effect was alleviated in more reverberant spaces and when using lower μ. Błasiński *et al.* found that the EDT did not accurately model the subjective reverberance ratings for double-sloped stimuli, as was suggested by the literature, whereas T_{20} proved to be a more useful predictor. Other perceived factors of AAESs have been studied, including colouration (Coleman *et al.*, 2024), where μ values between –5.4 and –8.5 dB represented a threshold of "slightly annoying" resonance in their eight-channel regenerative system. Aydin *et al.* (2025) found that the reported immersion increased with active enhancement compared to anechoic reproduction. Work on the objective evaluation of AAESs has also been conducted recently, with a study by Möller and Błasiński (2023) approaching the analysis with traditional metrics from ISO (2009), such as RT and clarity, and the change in RT between AAES conditions has been used for comparison. Werner *et al.* (2023) proposed using the spatial decomposition method to better evaluate the directional component of AAES reverberation, allowing the directions of arrival of an ambisonics receiver to be visualised within specific time regions under different enhancement conditions.

Developments by Green (2025) have explored designing passive acoustics specifically for the installation of an AAES, which differs from the often retrospective approach. This involved using curved ceiling reflectors to reduce the strength of the ceiling reflection, thereby allowing the AAES to introduce a higher virtual reflection and reducing the space's overall reverberance to control the resulting sound strength of the system.

Given the complexity of factors affecting the operation and perception of AAESs, system optimisation has been a particular topic of interest in recent years (Dal Santo *et al.*, 2025; De Bortoli *et al.*, 2024b,a; Zhang *et al.*, 2024). De Bortoli *et al.* (2024b) developed a method of resonance reduction based on a modal reverberator whereby phase shifts were introduced to destructively interfere via electroacoustic feedback, resulting in over 5 dB of increased headroom. A differentiable framework was also proposed by De Bortoli *et al.* (2024a) to optimise FIR-based EQ applied to an AAES reverberator using loss functions based on the eigenvalues of the feedback loop. A similar study by Zhang *et al.* (2024) approached electroacoustic resonance suppression with artificial neural networks, but by considering the closed-loop system in the frequency domain through a recursive training process. This resulted in a model capable of providing real-time feedback suppression. The use of deep learning in active acoustics is promising but has seen little development to date. However, with the emergence of open-source differentiable reverberators and frameworks, such exploration has become particularly accessible (Dal Santo *et al.*, 2023; Dal Santo *et al.*, 2025; De Bortoli *et al.*, 2025a; Mezza *et al.*, 2024). To aid the training of deep-learning-based AAES models and to increase the reproducibility of results in AAES studies, a dataset of measured RIRs from AAES installations is available (De Bortoli *et al.*, 2025b).

The relevance of achieving a natural-sounding acoustic response with AAESs has been questioned multiple times in the literature (Möller and Błasiński, 2023; Poletti, 2010), reframing active acoustics as a creative tool rather than merely a corrective one. Poletti (2010) discussed the reluctance to merge technology with art, stemming from the desire to preserve the art form's original setting. AAESs offer the potential for a new state-of-the-art in the entertainment industry, but the barriers of tradition may need to be broken. Given the relatively recent application of active acoustics to contemporary music venues, the topic of integrating object-based spatial audio with AAESs has been raised, such as combining L-ISA with Ambiance (Roskam *et al.*, 2022), and Dolby Atmos with Constellation (Ellison and Schwenke, 2025). A number of authors have also begun to explore the performative possibilities of AAESs (Blome *et al.*, 2025; Hashimoto *et al.*, 2023; Ohgi *et al.*, 2023; Roß *et al.*, 2023; Tampu, 2024), for example, by automating system parameters to transport listeners through different sonic environments and by exploiting artefacts as a compositional aid. Tampu (2024) demonstrated how stable resonances and RT extension could be used during a saxophone performance, relying on the performer's acoustic interaction with the system to develop movement and immersion. Audience enhancement is another area of interest that increases excitement and participation during sing-along sections and applause (Roskam *et al.*, 2022).

ACKNOWLEDGEMENTS

The first author was funded by the University of Surrey Doctoral College and L-Acoustics, and the second author by the Research Council of Finland, Project No. 357391.

https://doi.org/10.1121/10.0043585

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

DATA AVAILABILITY

The simulation code is openly available at the AAESToolbox GitHub repository.

[1]Griesinger (1991) defines the "enhancement critical distance" of an in-line system as the radius around an omnidirectional source at which the energies of the direct sound and the electroacoustic response are equal.
[2]https://github.com/wjcassidy/AAESToolbox/tree/review-simulations.

Abel, J. S., Callery, E. F., and Canfield-Dafilou, E. K. (**2018**). "A feedback canceling reverberator," in *Proceedings of the 21st International Conference on Digital Audio Effects (DAFx-18)*, pp. 100–106.

Atalay, T. B., Sü Gül, Z., De Sena, E., Cvetković, Z., and Hacıhabiboğlu, H. (**2022**). "Scattering delay network simulator of coupled volume acoustics," IEEE/ACM Trans. Audio. Speech Lang. Process. **30**, 582–593.

Aydin, A., Zhang, K. Y.-Y., Kelly, J., King, R., and Woszczyk, W. (**2025**). "A study on reverberation in a virtual acoustic setting using the Lexicon 960L," in *Proceedings of the 158th Convention of the Audio Engineering Society*, Warsaw, Poland (May 22–24).

Bakker, R., and Gillan, S. (**2014**). "The history of active acoustic enhancement systems," in *Proceedings of the Reproduced Sound 2014*, Birmingham, UK (October 14).

Balint, J., and Kaiser, F. (**2018**). "Multi-exponential decay curves in auditoriums," in *Proceedings of the Institute of Acoustics*, pp. 418–426.

Barnett, P. W. (**1987**). "Acoustic systems," U.S. patent US4649564A.

Beranek, L. L. (**1992**). "Concert hall acoustics," J. Acoust. Soc. Am. **92**(1), 1–39.

Berkhout, A. J., De Vries, D., Hemingway, J. R., and Griffioen, A. (**1988**). "Experience with the Acoustical Control System ACS," in *Proceedings of the 6th International Conference of the Audio Engineering Society: Sound Reinforcement*, Nashville, TN (May 5–8), pp. 534–555.

Berkhout, A. J., De Vries, D., and Sonke, J. J. (**1997**). "Array technology for acoustic wave field analysis in enclosures," J. Acoust. Soc. Am. **102**(5), 2757–2770.

Berkhout, A. J., De Vries, D., and Vogel, P. (**1993**). "Acoustic control by wave field synthesis," J. Acoust. Soc. Am. **93**(5), 2764–2778.

Błasiński, Ł., Felcyn, J., and Kociński, J. (**2024**). "Perception of reverberation length in rooms with reverberation enhancement systems," J. Acoust. Soc. Am. **156**(1), 450–462.

Błasiński, Ł., Pastusiak, A., Kociński, J., and Buszkiewicz, M. (**2023**). "Objective measurements and subjective assessment of the speech intelligibility in rooms with an active acoustic enhancement systems," in *Proceedings of INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, pp. 3480–3489.

Blome, M., Mommertz, E., and Hoffbauer, E. (**2025**). "A concert hall without artistic limits–The Bergson approach," in *Proceedings of the Institute of Acoustics*.

Boner, C. P., and Boner, C. R. (**1964**). "A procedure for controlling room-ring modes and feedback modes in sound systems with narrow-band filters," J. Audio Eng. Soc. **13**(4), 297–299.

Boner, C. P., and Boner, C. R. (**1965**). "Minimizing feedback in sound systems and room-ring modes with passive networks," J. Acoust. Soc. Am. **37**(1), 131–135.

Borß, C., and Martin, R. (**2009**). "An improved parametric model for perception-based design of virtual acoustics," in *Proceedings of the 35th International Conference of the Audio Engineering Society*, London, UK (February 11–13).

Bradley, D. T., and Wang, L. M. (**2005**). "The effects of simple coupled volume geometry on the objective and subjective results from nonexponential decay," J. Acoust. Soc. Am. **118**(3), 1480–1490.

Brinkmann, F., and Weinzierl, S. (**2017**). "AKtools—An open software toolbox for signal acquisition, processing, and inspection in acoustics," in *Proceedings of the 142nd Convention of the Audio Engineering Society*, Berlin, Germany (May 20–23).

Cassidy, W. J., Coleman, P., Mason, R., and De Sena, E. (**2024**). "Naturalness of double-slope decay in generalised active acoustic enhancement systems," in *Proceedings of the 27th International Conference on Digital Audio Effects*, pp. 262–269.

Chon, S. H., Ko, D., and Kim, S. (**2015**). "Listeners' response to string quartet performances recorded in virtual acoustics," in *Proceedings of Sound Reinforcement Engineering and Technology*, Montral, Canada (July 15–17).

Coleman, P., Epain, N., Venkatesh, S., and Roskam, F. (**2024**). "Exploring perceptual annoyance and colouration assessment in active acoustic environments," in *Proceedings of the Audio Engineering Society International Conference on Acoustics & Sound Reinforcement*, Le Mans, France, Paper No. 3.

Connor, W. K. (**1973**). "Experimental investigation of sound system-room feedback," J. Audio Eng. Soc. **21**(1), 27–32.

Dal Santo, G., De Bortoli, G. M., Prawda, K., Schlecht, S. J., and Välimäki, V. (**2025**). "FLAMO: An open-source library for frequency-domain differentiable audio processing," in *Proceedings of the ICASSP 2025*, Hyderabad, India, pp. 1–5.

Dal Santo, G., Prawda, K., Schlecht, S. J., and Välimäki, V. (**2023**). "Differentiable feedback delay network for colorless reverberation," in *Proceedings of the 26th International Conference on Digital Audio Effects*, Copenhagen, Denmark (September 4–7).

De Bortoli, G. M., Dal Santo, G., Prawda, K., Lokki, T., Välimäki, V., and Schlecht, S. J. (**2024a**). "Differentiable active acoustics—Optimizing stability via gradient descent," in *Proceedings of the 27th International Conference on Digital Audio Effects*, Guildford, UK (September 3–7), pp. 254–261.

De Bortoli, G. M., Prawda, K., Coleman, P., and Schlecht, S. J. (**2025a**). "DataRES and PyRES: A room dataset and a Python library for reverberation enhancement system development, evaluation, and simulation," in *Proceedings of the 28th International Conference on Digital Audio Effects*, Ancona, Italy (September 2–5), pp. 251–258.

De Bortoli, G. M., Prawda, K., Coleman, P., and Schlecht, S. J. (**2025b**). "DataRES," Zenodo version. 2.0.1 [dataset], 10.5281/zenodo.15737243.

De Bortoli, G. M., Prawda, K., and Schlecht, S. J. (**2024b**). "Active acoustics with a phase cancelling modal reverberator," J. Audio Eng. Soc. **72**(10), 705–715.

de Koning, S. H. (**1984**). "The MCR system—Multiple-channel amplification of reverberation," Philips Tech Rev. **41**(1), 12–23.

Ellison, S., and Schwenke, R. (**2025**). "Integrating spatial audio with active acoustics through the lens of egocentric and allocentric frames of reference," in *Proceedings of the Institute of Acoustics*.

Emami-Naeini, A., and Kosut, R. L. (**2012**). "The generalized Nyquist criterion and robustness margins with applications," in *Proceedings of the CDC 2012—51st IEEE Conference on Decision and Control (CDC)*, Maui, HI (December 10–13), pp. 226–231.

Gade, A. C. (**1997**). "Evaluation of a reverberation enhancement system installed in a small multi purpose hall," Acta Acust. united Ac. **83**, 522–529.

Green, E. (**2025**). "Bechstein Hall: Augmenting spatial perception using active acoustics," in *Proceedings of the Institute of Acoustics*.

Griesinger, D. (**1991**). "Improving room acoustics through time-variant synthetic reverberation," in *Proceedings of the 90th Convention of the Audio Engineering Society*, Paris, France (February 19–22), Paper No. 3014.

Haeussler, A., and van de Par, S. (**2015**). "Spectral and perceptual properties of a transfer chain of two rooms," in *Proceedings of the 10th European Congress and Exposition on Noise Control Engineering (EuroNoise)*, pp. 2077–2082.

Haeussler, A., and van de Par, S. (**2019**). "Crispness, speech intelligibility, and coloration of reverberant recordings played back in another reverberant room (room-in-room)," J. Acoust. Soc. Am. **145**(2), 931–944.

Hak, C. C. J. M., and Wenmaekers, R. H. C. (**2013**). "Room in room acoustics: Using convolutions to find the impact of a listening room on recording acoustics," in *Proceedings of the International Symposium on Room Acoustics*, Toronto, Canada (June 9–11), pp. 1–9.

Hak, C. C. J. M., and Wenmaekers, R. H. C. (**2015**). "Room in room acoustics: The influence of the direct/diffuse sound field ratio in a listening room on played back recorded acoustics," *Proceedings of the 10th European Congress and Exposition on Noise Control Engineering (EuroNoise)*, pp. 1–5.

Hashimoto, D., Miyazaki, H., Bakker, R., and Kim, S. (**2023**). "Sound environment control using an immersive audio system—Advancing sound experience for creators and listeners using active field control," in *Proceedings of the International Conference on Spatial & Immersive Audio, Audio Engineering Society*, Huddersfield, UK (August 23–25), Paper No. 45.
Hashimoto, D., Watanabe, T., and Miyazaki, H. (**2020**). "Comparison of methods to control early reflections for acoustic enhancement systems," in *Proceedings of the 149th Convention of the Audio Engineering Society*, October 27–30, Paper No. 10422.
Hatziantoniou, P. D., Tatlas, N.-A., and Potirakis, S. M. (**2013**). "Optimizing teaching room acoustics: Investigating the exclusive use of a distributed electroacoustic installation to improve the speech intelligibility," in *Proceedings of the 134th Convention of the Audio Engineering Society*, Rome, Italy (May 4–7), Paper No. 8871.
Hopper, H. A. (**2012**). "Reverberation enhancement for small rooms," Ph.D. thesis, University of Southampton, Southampton, UK.
Hough, C., Boulter, N., and Tait, S. (**2025**). "The use of electroacoustic enhancement systems in the design of orchestral rehearsal rooms," J. Acoust. Soc. Am. **157**(4), A338.
Hughes, R. J. (**2016**). "The room-in-room effect and its influence on perceived room size in spatial audio reproduction," in *Proceedings of the 141st Convention of the Audio Engineering Society*, Los Angeles, CA, Paper No. 9621.
Hyon, J, and Jeong, D. (**2021**). "Variable acoustics in performance venues—A review," J. Acoust. Soc. Korea **40**(6), 626–648.
IEC (**2020**). IEC 6026816, "Sound system equipment—Part 16, Objective rating of speech intelligibility by speech transmission index" (IEC, Geneva, Switzerland).
ISO (**2009**). ISO 3382-1, "Acoustics—Measurement of room acoustic parameters—Part 1: Performance spaces" (ISO, Geneva, Switzerland).
Jaffe, J. C. (**1977**). "Electronic sound enhancing system," U.S. patent US4061876A.
Jaffe, J. C., and Scarbrough, P. H. (**1992**). "Electronic architecture: Toward a better understanding of theory and application," in *Proceedings of the 93rd Convention of the Audio Engineering Society*, San Francisco, CA, Paper No. 3382.
Jagla, J., and Chervin, P. (**2020**). "The effect of electronic reverberation in a hybrid reverberation enhancement system: CarmenCita," in *Proceedings of Forum Acusticum*, Lyon, France (December 7–11), pp. 1223–1227.
Jot, J.-M., and Chaigne, A. (**1991**). "Digital delay networks for designing artificial reverberators," in *Proceedings of the 90th Convention of the Audio Engineering Society*, Paris, France (February 19–22), Paper No. 3030.
Joyce, W. B. (**1978**). "Exact effect of surface roughness on the reverberation time of a uniformly absorbing spherical enclosure," J. Acoust. Soc. Am. **64**(5), 1429–1436.
Kaiser, F. (**2009**). "Acoustic enhancement systems," Ph.D. thesis, Graz University of Technology, Graz, Austria.
Kaiser, F., Frischmann, C., Werner, V., and Rohde, T. (**2019**). "Room acoustic evaluation of active acoustics systems—Results from measurements," in *Proceedings of the International Symposium on Room Acoustics*, Amsterdam, The Netherlands (September 15–17), pp. 177–185.
Kaiser, F., Werner, V., and Rohde, T. (**2016**). "Active acoustics, speech enhancement and noise masking in multipurpose venues," in *Proceedings of the 29th International Convention of the Tonmeistertagung*, Cologne, Germany (November 17–20).
Kanev, N. (**2011**). "Sound decay in a rectangular room with specular and diffuse reflecting surfaces," in *Proceedings of Acusticum 2011*, pp. 1935–1940.
Kawakami, F., and Shimizu, Y. (**1990**). "Active field control in auditoria," Appl. Acoust. **31**(1–3), 47–75.
Ko, D. (**2015**). "Virtual acoustics for musicians: Exploring the influence of an electronic acoustic enhancement system on music performance," Ph.D. thesis, McGill University, Montreal, Canada.
Ko, D., Woszczyk, W., and Chon, S. H. (**2012**). "Evaluation of a new active acoustics system in performances of five string quartets," in *Proceedings of the 132nd Convention of the Audio Engineering Society*, Paper No. 8603.
Kok, B., and Prinssen, W. (**2003**). "Acoustical enhancement systems: Design criteria and evaluation of room acoustical parameters based on in situ measurements," J. Acoust. Soc. Am. **114**, 2342.
Kuttruff, H. (**2000**). *Room Acoustics*, 4th ed. (Spon Press, London, UK).
Lexicon (**2024**). "Lexicon 480L owner's manual," https://lexiconpro.com/en/product_documents/480l_manual_rev23pdf (Last viewed July 16, 2024).
Lindau, A, Kosanke, L, and Weinzierl, S. (**2012**). "Perceptual evaluation of model- and signal-based predictors of the mixing time in binaural room impulse responses," J. Audio Eng. Soc. **60**(11), 887–898.
Lokki, T., and Hiipakka, J. (**2001**). "A time-variant reverberation algorithm for reverberation enhancement systems," in *Proceedings of the 4th International Conference on Digital Audio Effects*, pp. 28–32.
Lokki, T., Kajastila, R., and Takala, T. (**2007**). "Virtual acoustic spaces with multiple reverberation enhancement systems," in *Proceedings 30th International Conference of the Audio Engineering Society*.
Lokki, T., Nummela, J., and Lahti, T. (**2000**). "An electro-acoustic enhancement system for rehearsal rooms," in *Proceedings of the European Acoustics Association Symposium on Architectural Acoustics*, Madrid, Spain (October 14–20), Paper AAQ06.
MacFarlane, A. G. J., and Postlethwaite, I. (**1977**). "The generalized Nyquist Stability Criterion and Multivariable Root Loci," Int. J. Control **25**(1), 81–127.
Meyer Sound (**2024**). "Constellation," https://meyersound.com/product/constellation/ (Last viewed July 26, 2024).
Meynial, X., and Vuichard, O. (**1999**). "Objective measure of sound colouration in rooms," Acta Acust. united Ac. **85**(1), 101–107.
Mezza, A. I., Giampiccolo, R., De Sena, E., and Bernardini, A. (**2024**). "Data-driven room acoustic modeling via differentiable feedback delay networks with learnable delay lines," J. Audio. Speech. Music Proc. **2024**(1), 51.
Miyazaki, H., Watanabe, T., Kishinaga, S., and Kawakami, F. (**2003**). "Active field control (AFC): reverberation enhancement system using acoustical feedback control," in *Proceedings of the 115th Convention of the Audio Engineering Society*, New York (October 10–13), Paper No. 5861.
Mohlin, P., and Höstmad, P. (**2022**). "Blind estimation of sound coloration in rooms using chi-square distributions of damping constants," J. Acoust. Soc. Am. **152**(1), 456–469.
Möller, H. (**2022**). "Room acoustic design for electronic enhancement systems," in *Proceedings of the Baltic-Nordic Acoustics Meeting*, Aalborg, Denmark (May 9–11), pp. 261–267.
Möller, H., and Błasiński, Ł. (**2023**). "Room acoustic measurements in halls with electro-acoustic enhancement systems," in *Proceedings of the Institute of Acoustics*.
Mortessagne, F., Legrand, O., and Sornette, D. (**1993**). "Role of the absorption distribution and generalization of exponential reverberation law in chaotic rooms," J. Acoust. Soc. Am. **94**(1), 154–161.
Müller-BBM Acoustic Solutions GmbH. (**2023**). "Vivace—Our acoustic solutions for perfect sound," https://vivace.mbbm-aso.com/vivace/ (Last viewed December 14, 2023).
Neeten, S., Frischmann, C., Werner, V., and Kaiser, F. (**2025**). "Objective assessment of coloration in active acoustic systems: real-room measurements with time-invariant filter equalization," in *Proceedings of the Institute of Acoustics*.
Nielsen, J. L., and Svensson, U. P. (**1999**). "Performance of some linear time-varying systems in control of acoustic feedback," J. Acoust. Soc. Am. **106**(1), 240–254.
Nilsson, E. (**2004**). "Decay processes in rooms with non-diffuse sound fields. Part I: Ceiling treatment with absorbing material," Build. Acoust. **11**(1), 39–60.
Nyquist, H. (**1932**). "Regeneration theory," Bell Syst. Tech. J. **11**(1), 126–147.
Ohgi, H., Miyazaki, H., Kim, S., and Uhm, S. (**2023**). "Transcending boundaries: Unleashing musical expression through immersive sound image and reverberation control system," in *Proceedings of the 155th Convention of the Audio Engineering Society*.
Ohsmann, M. (**1990**). "Analyse von mehrkanalanlagen" ("Analysis of multi-channel systems"), Acta Acust. united Ac. **70**(4), 233–246.
Parkin, P. H., and Morgan, K. (**1970**). "Assisted resonance in The Royal Festival Hall, London: 1965–1969," J. Acoust. Soc. Am. **48**(5A), 1025–1035.
Pastusiak, A., Błasiński, Ł., and Kociński, J. (**2025**). "Listening effort in reverberant rooms: A comparative study of subjective perception and objective acoustic metrics," Arch. Acoust. **50**, 321–329.
Poletti, M. A. (**1993**). "On controlling the apparent absorption and volume in assisted reverberation systems," Acta Acust. united Ac. **78**(2), 61–73.
Poletti, M. A. (**1995**). "A unitary reverberator for reduced colouration in assisted reverberation systems," in *Proceedings of INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, pp. 1223–1232.

Poletti, M. A. (**1996**). "An assisted reverberation system for controlling apparent room absorption and volume," in *Proceedings of the 101st Convention of the Audio Engineering Society*, Paper No. 4365.
Poletti, M. A. (**1998a**). "The analysis of a general assisted reverberation system," Acta Acust. united Ac. **84**, 766–775.
Poletti, M. A. (**1998b**). "The measurement of coloration in electroacoustic enhancement systems," J. Acoust. Soc. Am. **103**(5), 3026–3027.
Poletti, M. A. (**2000**). "The stability of single and multichannel sound systems," Acta Acust. united Ac. **86**(1), 163–178.
Poletti, M. A. (**2004**). "The stability of multichannel sound systems with frequency shifting," J. Acoust. Soc. Am. **116**(2), 853–871.
Poletti, M. A. (**2006**). "The control of early and late energy using the variable room acoustics system," in *Proceedings of Acoustics*, Christchurch, New Zealand (November 20–22), pp. 215–218.
Poletti, M. A. (**2010**). "Active acoustic systems for the control of room acoustics," in *Proceedings of the International Symposium on Room Acoustics*, Melbourne, Australia (August 29–31).
Poletti, M. A. (**2011**). "Active acoustic systems for the control of room acoustics," Build. Acoust. **18**(3-4), 237–258.
Prawda, K. (**2025**). "Sensitivity of room impulse responses in changing acoustic environment," in *Proceedings of ICASSP 2025*, Hyderabad, India (April 6–11), pp. 1–5.
Prawda, K., Schlecht, S. J., and Välimäki, V. (**2023**). "Time variance in measured room impulse responses," in *Proceedings of Forum Acusticum*, Turin, Italy (September 10–15), pp. 1639–1646.
Prawda, K., Schlecht, S. J., and Välimäki, V. (**2024**). "Short-time coherence between repeated room impulse response measurements," J. Acoust. Soc. Am. **156**(2), 1017–1028.
Prinssen, W., and Holden, M. (**1992**). "System for improved acoustic performance," in *Proceedings of the Institute of Acoustics*, pp. 93–101.
Pu, H., Qiu, X., and Wang, J. (**2011**). "Different sound decay patterns and energy feedback in coupled volumes," J. Acoust. Soc. Am. **129**(4), 1972–1980.
Rindel, J. H. (**2016**). "Detection of colouration in rooms by use of cepstrum technique," in *Proceedings of the German Annual Conference on Acoustics (DAGA)*, Aachen, Germany (March 14–17).
Roskam, F., Laval, J., Corteel, E., and Willsallen, S. (**2022**). "ABBA voyage: Object-based mixing meets acoustic enhancement," in *Proceedings of the Meeting of the Institute of Acoustics Reproduced Sound*, Bristol, UK (November 17).
Rosseel, H., and van Waterschoot, T. (**2023**). "Interactive perceptual evaluation of auralized acoustics in rehearsal spaces for vocal music," in *Proceedings of Forum Acusticum*, Turin, Italy (September 10–15), pp. 4123–4130.
Roß, B., Werner, V., Frischmann, C., and Neeten, S. (**2023**). "A comparison of musical performances in different acoustic environments created with an active acoustics system," in *Proceedings of the Institute of Acoustics*.
Rubak, P., and Johansen, L. G. (**2003**). "Coloration in natural and artificial room impulse responses," in *Proceedings of the 23rd International Conference of the Audio Engineering Society*, Copenhagen, Denmark (May 23–25), pp. 171–189.
Rumsey, F., and Kok, B. (**2014**). "Acoustic enhancement systems," J. Audio Eng. Soc. **62**(6), 449–453.
Schlecht, S. J. (**2020**). "FDNTB: The feedback delay network toolbox," in *Proceedings of the 23rd International Conference on Digital Audio Effects*, Vienna, Austria (September 9–11), pp. 211–218.
Schlecht, S. J., and Habets, E. A. P. (**2015**). "Reverberation enhancement systems with time-varying mixing matrices," in *Proceedings of the Sound Reinforcement Engineering and Technology Conference*, Montreal, Canada (July 15–17).
Schlecht, S. J., and Habets, E. A. P. (**2016**). "The stability of multichannel sound systems with time-varying mixing matrices," J. Acoust. Soc. Am. **140**(1), 601–609.
Schlecht, S. J., and Weiss, S. (**2024**). "Reconstructing analytic dinosaurs: Polynomial eigenvalue decomposition for eigenvalues with unmajorised ground truth," in *Proceedings of the 32nd European Signal Processing Conference (EUSIPCO)*, pp. 1287–1291.
Schmich, I., and Vian, J.-P. (**2004**). "CARMEN: A physical approach for room acoustic enhancement system," *in Proceedings of the Joint Congress CFA/DAGA*, Strasbourg, France (March 22–25).
Schroeder, M. R. (**1962**). "Frequency-correlation functions of frequency responses in rooms," J. Acoust. Soc. Am. **34**(12), 1819–1823.
Schroeder, M. R. (**1964**). "Improvement of acoustic-feedback stability by frequency shifting," J. Acoust. Soc. Am. **36**(9), 1718–1724.
Schroeder, M. R., and Kuttruff, K. H. (**1962**). "On frequency response curves in rooms. Comparison of experimental, theoretical, and Monte Carlo results for the average frequency spacing between maxima," J. Acoust. Soc. Am. **34**(1), 76–80.
Strøm, S., Krokstad, A., Sørsdal, S., and Stensby, S. (**1986**). "Design of room acoustics and a MCR reverberation system for Bjergsted concert hall in Stavanger," Appl. Acoust. **19**(6), 465–475.
Svensson, P. U. (**1994**). "On reverberation enhancement in auditoria," Technical Report F94-01 (Chalmers University of Technology, Göteborg).
Svensson, P. U. (**1995**). "Computer simulations of periodically time-varying filters for acoustic feedback control," J. Audio Eng. Soc. **43**(9), 667–677.
Tampu, E. M. (**2024**). "Active acoustics: A compositional and performative approach to regenerative systems," Master's thesis, Aalto University, Aalto, Finland.
Toole, F. E., and Olive, S. E. (**1986**). "The perception of sound coloration due to resonances in loudspeakers and other audio components," in *Proceedings of the 81st Convention of the Audio Engineering Society*, Los Angeles, CA (November 12–16), Paper No. 2406.
Välimäki, V., Parker, J. D., Savioja, L., Smith, J. O., and Abel, J. S. (**2012**). "Fifty years of artificial reverberation," IEEE Trans. Audio. Speech Lang. Process. **20**(5), 1421–1448.
Vogel, P., and de Vries, D. (**1994**). "Electroacoustic system response in a hall: A convolution of impulse sequences," J. Audio Eng. Soc. **42**(9), 684–690.
Vorländer, M. (**2008**). "Back matter," in *Auralization*, 1st ed. (Springer, Berlin, Heidelberg, Germany), pp. 303–335.
Vuichard, O., and Meynial, X. (**2000**). "On microphone positioning in electroacoustic reverberation enhancement systems," Acta Acust. united Acust. **86**(5), 853–859.
Walter, F., and Melchior, F. (**2008**). "On the measurement of electro acoustic enhanced sound fields," in *Proceedings of the 124th Convention of the Audio Engineering Society*, Amsterdam, The Netherlands (May 17–20), Paper No. 7468.
Watanabe, T., and Miyazaki, H. (**2016**). "Electronic shell—Improvement of room acoustics without orchestra shell utilizing active field control," in *Proceedings of the 140th Convention of the Audio Engineering Society*, Paris, France, pp. 3291–3291.
Waterhouse, R. V. (**1965**). "Theory of howlback in reverberant rooms," J. Acoust. Soc. Am. **37**(5), 921–923.
Werner, V. (**2025**). (private communication).
Werner, V., Neeten, S., and Kaiser, F. (**2023**). "Evaluation of a geometric approach to active acoustics," in *Proceedings of the Institute of Acoustics*.
Woszczyk, W. (**2011**). "Active acoustics in concert halls—A new approach," Arch. Acoust. **36**(2), 379–393.
Xiang, N., Jing, Y., and Bockman, A. C. (**2009**). "Investigation of acoustically coupled enclosures using a diffusion-equation model," J. Acoust. Soc. Am. **126**(3), 1187–1198.
Zacharov, N. (**2018**). *Sensory Evaluation of Sound*, 1st ed. (CRC Press, Boca Raton, FL).
Zhang, H., Zhang, Y., Yu, M., and Yu, D. (**2024**). "Advancing acoustic howling suppression through recursive training of neural networks," in *Proceedings of the ICASSP 2024*, Seoul, Korea (April 14–19), pp. 711–715.