M. Fontana, D. Fantini, G. Presti, M. Tiraboschi, and F. Avanzini,
"A Low-Dimensional Projection of Sound Absorption Coefficients for a Scattering Delay Network Reverberator,"
J. Audio Eng. Soc., vol. 74, no. 4, pp. 212–223 (2026 Apr.).
https://doi.org/10.17743/jaes.2022.0261.

A Low-Dimensional Projection of Sound Absorption Coefficients for a Scattering Delay Network Reverberator

MARCO FONTANA, (marco.fontana@unimi.it) **DAVIDE FANTINI,*** (davide.fantini@unimi.it) **GIORGIO PRESTI,** (giorgio.presti@unimi.it)

MARCO TIRABOSCHI, (marco.tiraboschi@unimi.it) **AND FEDERICO AVANZINI,** *AES Member* (federico.avanzini@unimi.it)

Laboratory of Music Informatics (LIM), Department of Computer Science, University of Milan, Milan, Italy

Scattering delay networks (SDNs), a class of artificial reverberators with physically interpretable parameters, provide an efficient synthesis of room acoustics accounting for wall absorption properties. This paper builds upon a previously proposed highly parametrized and real-time implementation of an SDN, exposing octave-band absorption coefficients for each room wall. The individual manipulation of these coefficients can be challenging due to their high dimensionality. A 2D parameter space (2PS) is proposed to facilitate the navigation of the absorption coefficients. The 2PS is obtained using principal component analysis as an initial dimensionality reduction of a dataset of absorption coefficients, followed by a relaxation procedure to create a seamless 2D representation of the coefficients. A twofold evaluation of the proposed 2PS was conducted: (a) the 2PS was compared to the original space and the raw principal component analysis in a reverb matching task for the tuning of SDN coefficients, and (b) a usability test with expert audio professionals supported the potential of the 2PS from a user standpoint.

Keywords: acoustic signal processing, artificial reverberation, dimensionality reduction, optimization, room acoustics

1 INTRODUCTION

Since the 1960s, digital artificial reverberation has gained popularity in various fields, including audio/music production, video games, and extended reality (XR). In XR, coherence with the visual stimulus and flexibility for source and receiver positions are essential. Consequently, physically based approaches for reverberation are preferred [1]. These approaches simulate sound wave propagation in a room given its geometry (i.e., shape and size) and acoustic properties, such as material sound absorption. Physically based models are computationally demanding; hence, reverberators based on delay networks may offer an efficient and controllable alternative. In particular, the structure of certain delay networks is tied to the physical properties of the simulated room, making them well-suited for XR applications. In this regard, scattering delay networks (SDNs) [2] approximate the acoustics of a shoebox room, with scattering nodes modeling the frequency-dependent absorption of each wall.

In a previous work [3], the present authors developed an SDN implementation for real-time room acoustics, which enables interactive control of parameters such as source and receiver positions, room sizes, and sound absorption for each wall. The latter is modeled by octave-band sound absorption coefficients [4], defined as the ratio between the absorbed and incident energies. For each wall of the SDN implementation, a filter is dynamically designed based on the absorption coefficients in the octave bands from 125 Hz to 16 kHz. This yields 48 absorption coefficients, given the eight octaves and the six shoebox walls. Consequently, exploring this high-dimensional parameter space to attain the desired effect may require users to endure strenuous and time-consuming tuning sessions. In general, this problem is typically encountered in audio effects controlled by numerous parameters. Accordingly, some researchers proposed to address this problem through dimensionality reduction techniques. The high-dimensional parameter space

*To whom correspondence should be addressed, email: davide.fantini@unimi.it.

is thus projected into a more manageable, low-dimensional representation, simplifying parameter manipulation while preserving essential characteristics. These techniques have mainly been applied to equalizers [5–9] and synthesizers [10–12], while they have yet to be explored for absorption coefficients.

In this regard, this paper proposes a 2D projection of a dataset comprising octave-band absorption coefficients measured for several materials. The projection consists of an initial dimensionality reduction performed with principal component analysis (PCA), followed by a custom relaxation algorithm based on Delaunay triangulation (see SEC. 3). This results in the proposed 2D parameter space (2PS), which provides a seamless transition between one material and its neighbors. Compared to established dimensionality reduction techniques, the proposed 2PS aims at creating a continuous and smooth 2D representation of a dataset of absorption coefficients yet allowing for a lossless reconstruction of the original points based on interpolation with barycentric coordinates. This feature is desirable when the dataset consists of measured real-world materials, whose coefficients are exactly reconstructed when the corresponding 2D point is selected, rather than being approximated.

The objective of the 2PS is to enable a streamlined parametrization of the SDN's absorption properties. Specifically, the 2PS simplifies the coefficients search through automated procedures and alleviates users from the individual manipulation of numerous parameters. This is supported by a twofold evaluation of the 2PS. First, an existing reverb matching method [13, 14] was adapted to tune the SDN's absorption coefficients based on target spatial room impulse responses (SRIRs; see SEC. 4). The 2PS yielded improved matching accuracy over the original space and the raw PCA technique. Secondly, a usability test with audio experts indicated that the 2PS outperforms the use of octave-band coefficients in the identification of a target reverberation.

2 BACKGROUND

2.1 Scattering Delay Networks

2.1.1 Theoretical Overview

SDNs [2] simulate the reverberation inside a shoebox room with explicit control of its physical properties, such as its geometry and absorption properties. This noteworthy feature is shared with physically based models, yet SDNs retain the computational efficiency of delay networks. The SDN structure is shown in Fig. 1, and it consists of interconnected nodes, which include the source, the receiver, and a scattering node placed on each wall where the first-order reflections occur. Accordingly, the direct path and the first-order reflections between the source and receiver are correctly rendered in time and space. On the other hand, the accuracy of higher-order reflections decreases as the reflection order increases.

In the scattering nodes, energy absorption is modeled with random-incidence absorption coefficients α [15]. The signals incoming to these nodes are scattered according to the scattering matrix $\mathbf{S}$ to obtain the outgoing signals. The wall absorption effect can be expressed as $\mathbf{S}^*\mathbf{Y}\mathbf{S} = (1-\alpha)\mathbf{Y}$, where $\mathbf{Y}$ is the Hermitian positive-definite matrix and $(\cdot)^*$ is the conjugate transpose. Given the reflectance coefficient $\rho = \sqrt{1-\alpha}$, $\mathbf{S}$ can be expressed as $\mathbf{S} = \rho\mathbf{A}$, where $\mathbf{A}$ is a lossless scattering matrix. This allows the energy loss at the SDN walls to match the corresponding physical room, resulting in an energy decay rate consistent with the Sabine [16] and Eyring [17] equations. To model frequency-dependent absorption properties, one can use the scattering matrix $\mathbf{S} = \mathbf{H}(z)\mathbf{A}$, where $\mathbf{H}(z)$ is a diagonal matrix composed of the filter $H(z)$. $H(z)$ is designed from the absorption coefficients, typically defined per octave band.

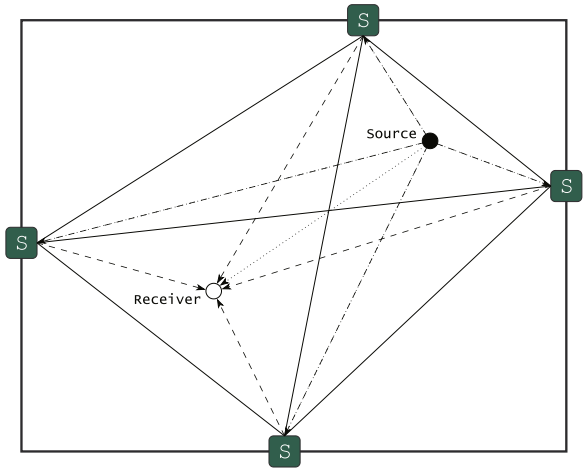


Fig. 1. Two-dimensional schema of the SDN structure adapted from [2, Fig. 4]. The blocks S correspond to the SDN wall nodes. Solid lines represent bidirectional delay lines connecting the wall nodes to each other. Dash lines represent the unidirectional delay line connecting the source to the wall nodes. Dashed lines represent the unidirectional delay lines connecting the wall nodes to the receiver. The dotted line represents the line-of-sight path. Figure reused from "A Highly Parametrized Scattering Delay Network Implementation for Interactive Room Auralization," © 2024 Marco Fontana et al., licensed under Creative Commons Attribution 4.0 International License.

2.1.2 Proposed Implementation

In a previous work [3], the present authors proposed an SDN implementation, which is publicly available both as a VST3 plugin and a standalone application.[1] This implementation provides a highly parametrized and real-time auralization of a 3D shoebox room. The parameters encompass source and receiver positions x_S and x_R, receiver orientation, shoebox dimensions, and absorption coefficients. The plugin supports mono, stereo, and up to fifth-order Ambisonics as output formats. Additionally, the binaural rendering toolbox [18] is employed for binaural output with custom head-related transfer functions.

The SDN models the frequency-dependent sound absorption of the shoebox walls. This is controlled by the absorption coefficients α_m^b, defined for each wall

[1] https://github.com/LIMUNIMI/Real-time-SDN.

$m \in \{1, \ldots, M\}$, with $M = 6$, and each octave band $b \in \{1, \ldots, 8\}$ between 125 Hz and 16 kHz. For each wall m, a third-order infinite impulse response filter is designed according to the coefficients α_m. This type of filter closely matches the coefficients of real-world materials while maintaining a low computational cost at runtime. Absorption coefficients α_m^b were processed before filter design. First, they were converted to reflectance coefficients ρ_m^b. To cover the entire spectrum, the values at the Nyquist frequency and at 0 Hz were extrapolated from the coefficients at the 16-kHz and 125-Hz bands, respectively. Then, linear interpolation was used to generate a denser set of points, facilitating filter design. Finally, the minimum phase frequency response was computed and the frequency points were weighted according to equivalent rectangular bandwidth [19]. Further details can be found in the present authors' previous work [3].

2.2 User Exploration of Audio Effects Parameters

The literature includes works to facilitate the tuning of audio effects with numerous parameters, which can prove challenging. This section provides an overview of such works. To the best of the authors' knowledge, no existing study is devoted to absorption coefficients used in reverberation effects. Therefore, the present focus is on similar approaches for equalizers and synthesizers in the following.

2.2.1 Equalization

The use of dimensionality reduction techniques for the projection of equalization parameters in a low-dimensional space is close to the use case of absorption coefficients. Mecklenburg and Loviscach [5] used self-organizing maps to derive a 2D space from the filter parameters of an equalizer. Neighbor interpolation allowed the continuous navigation of the 2D space through an interface that displayed descriptors such as "warm," "airy," and "boomy." In a similar work [6], the proposed 2D space was associated with four descriptors (tinny, warm, bright, dark) derived from user annotations of equalization curves. To obtain the 2D space, PCA was applied to 40-band weighting functions derived from the annotations, which quantify the influence of each band on the descriptor. Cartwright et al. [7] proposed a 2D space for exploring the mix of four audio tracks. This space was constructed by applying self-organizing maps to a corpus of equalizer parameters and a gain. The obtained mixes were then adjusted according to user ratings of the 2D points. In a subjective test, the proposed space facilitated the parameter exploration compared to the traditional mixing approach.

Stasis et al. [8] collected a dataset of equalization parameters tuned by users according to two descriptors: warm and bright. They proposed an equalization system capable of adapting to the input audio and based on a 2PS that was constructed by reducing the dataset dimensionality with a stacked autoencoder. In an equalization task, participants achieved improved separability of the descriptors in the 2D space compared to the original parameters. The approach was later expanded [9] with a clustering step, whereby equalization curves were grouped to identify salient subrepresentations within each descriptor. A user interface was designed to recommend changes according to the nearest cluster.

2.2.2 Synthesizers

The facilitation of parameter exploration has also interested synthesizers. Peachey et al. [10] trained a variational autoencoder to learn a 2D latent space from a dataset of synthesizer patches. The interpolation in the latent space allowed the generation of previously unseen patches. The method was evaluated by comparing the timbral characteristics of real and reconstructed patches. Esling et al. [11] used a variational autoencoder and normalizing flows to learn an organized latent space for audio rather than parameters. The proposed model allowed for synthesizer control while providing an invertible mapping to the space of the synthesizer's parameters. The model was later evaluated on larger sets of parameters [11]. Additionally, disentangling flows were used to learn the invertible mapping between the latent spaces for audio and parameters, and a semantic control for the synthesizer was incorporated. In a similar work [12], PCA was applied to the activation space of an early layer of a generative adversarial networks trained on synthesized sounds. This allowed the control of audio features by finding significant directions in the latent space. However, a quantitative assessment of the proposed method was not reported.

3 2D PARAMETER SPACE

3.1 Absorption Coefficients Dataset

The 2PS proposed in this paper is obtained by reducing the dimensionality of multiple datasets of absorption coefficients [20–24]. These datasets contain measurements for several real-world materials, including examples found in both nature and synthetic materials. The materials span from very reflective to very absorbent surfaces. For instance, the datasets include various types of glass, concrete, wood, and fabric, as well as snow, grass, and groups of people, among others. Each dataset entry is represented by $B = 6$ values in the range [0, 1], corresponding to the coefficients of the octave bands from 125 Hz to 4 kHz. In accordance with the ISO 354 standard [15], the 8-kHz and 16-kHz bands were disregarded. When working with the SDN, the coefficients for these two bands were replicated from a 4-kHz band. Further, 11 entries corresponding to flat absorption coefficients were included in the dataset. These entries are obtained by replicating each value in the set $\{0, 0.1, \ldots, 1\}$ for all octaves. A total of $N = 442$ entries were obtained after removing entries with identical absorption coefficients.

3.2 Initial Dimensionality Reduction

All the approaches discussed in SEC. 2.2 do not address the issue of perfectly reconstructing the original data points. Since the present approach is based on real-world measure-

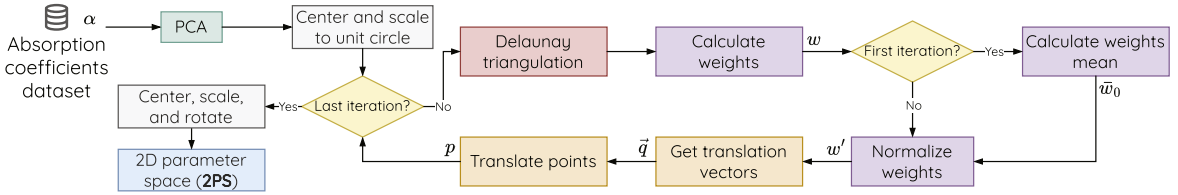


Fig. 2. Flowchart of the 2D projection of a dataset of octave-band absorption coefficients to obtain the proposed 2PS.

ments, dimensionality reduction is treated solely as a way to project these data points onto a 2D plane, which then serves as the basis for interpolation. In this context, PCA was used as the initial dimensionality reduction step to map the dataset of B-dimensional absorption coefficients to a 2D space [see Fig. 3(a)], where B is the number of octave bands of the coefficients. PCA was chosen because it generates relatively evenly distributed 2D points without outliers or sparse clusters while also providing interpretable axes for the resulting 2D space (see SEC. 3.3). Alternative techniques for initial dimensionality reduction were contemplated but discarded. An autoencoder—with a two-neuron bottleneck layer—yielded results comparable to those obtained with PCA. Manifold learning techniques, including multidimensional scaling, Uniform Manifold Approximation and Projection, and t-distributed stochastic neighbor embedding, were also tested. However, they either yielded sparse clusters of points or, again, equivalent results to PCA. The 2D spaces obtained with the discarded techniques are provided in the supplementary materials.[2]

3.3 Relaxation Algorithm

The space obtained through PCA exhibits abrupt changes and discontinuities in the reconstructed absorption coefficients during exploration. To address this, the 2PS is introduced, derived from the initial PCA subspace using an iterative relaxation algorithm inspired by Voronoi relaxation [25]. This algorithm aims to produce a seamless 2D representation of the absorption coefficients by minimizing differences between neighboring points in the 2D space. At each iteration, the space is subdivided into triangles having the data points as vertices. The algorithm then analyzes the rate of change of the absorption coefficients between the vertices of each triangle. This rate of change is represented by the difference in the reconstructed absorption values between each point and its neighbors. The space's layout is then adjusted by expanding the area of triangles with higher rates of change and contracting those with lower rates.

3.3.1 Algorithm's Workflow

Fig. 2 shows the flowchart of the proposed 2D projection. Let α_n^b be the nth dataset's entry for the band $b \in \{1, \ldots, B\}$. As mentioned above, PCA is used to map the coefficients α_n into a 2D space of points p_n. As the distribution of the obtained 2D points p_n resembles a circular shape [see Fig. 3(a)], they are fit into the unit circle to constrain the parameters navigation within this region. This is achieved via a two-step process: a translation to center the points at coordinates (0, 0), followed by a scaling to map the convex hull of the points p_n onto the unit circle. Specifically, each point p_n is scaled as p_n/r_n, where r_n is the length of the segment that starts at the origin, intersects p_n, and ends on the boundary of the convex hull.

Then, the proposed relaxation algorithm is applied to move the 2D points p_n iteratively. First, a Delaunay triangulation is calculated to partition the convex hull into triangles having the points p_n as vertices. Then, the weight w_t, representing the coefficient's rate of change for each triangle t, is computed as follows. Let $d(c_t, p_{v_t})$ be the Euclidean distances between the centroid c_t of t and the points p_{v_t}, corresponding to the triangle's vertices v_t. The corresponding distance $d(C_t, \alpha_{v_t})$ in the original space is computed, where C_t is the centroid of the coefficients α_{v_t}, represented by the points p_{v_t} in the 2D space. The weight w_t is defined as:

$$w_t = \frac{1}{3} \sum_{n \in v_t} \frac{d(C_t, \alpha_n)}{d(c_t, p_n)}. \tag{1}$$

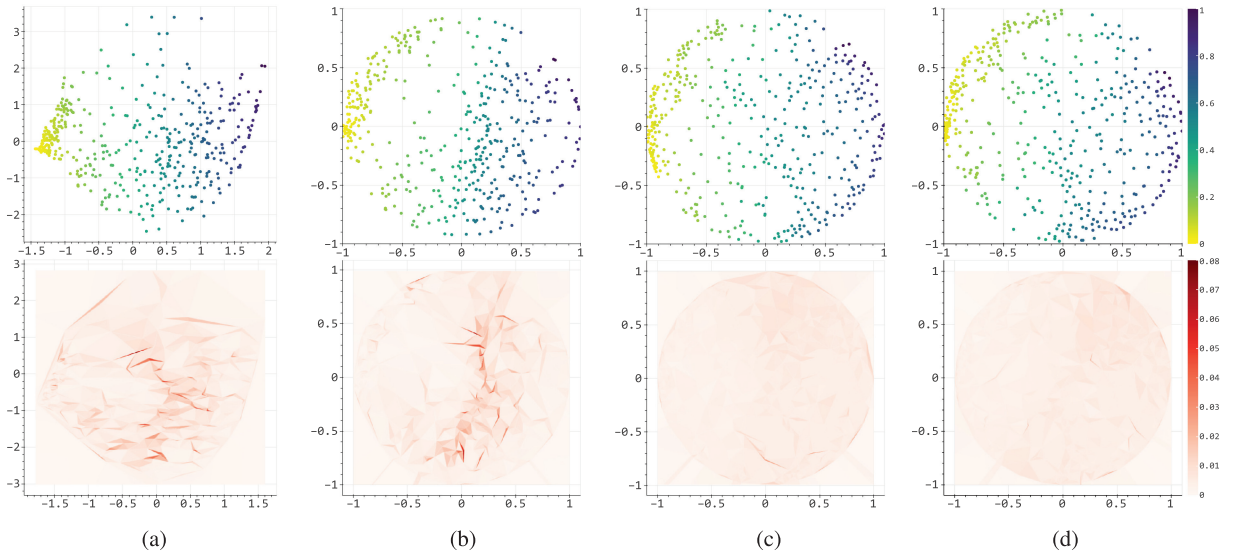
The smaller the distances in the 2D space compared to those in the original space, the higher the weight w_t, and vice versa. The weights are then normalized as $w'_t = (w_t - \bar{w}_0)/\bar{w}_0$, where $\bar{w}_0$ is the mean of weights w_t at the first iteration. For a given p_n, the translation vector $\vec{q}_n$ is so defined:

$$\vec{q}_n = \sigma \sum_{t \in T_n} - \overrightarrow{p_n c_t} \cdot w'_t, \tag{2}$$

where T_n is the set of triangles having p_n as vertex, $\overrightarrow{p_n c_t}$ is the vector from p_n to the centroid c_t, and σ controls the magnitude of the translation. The points p_n are translated by the vectors $\vec{q}_n$ to obtain the updated points for the next iteration. The larger the deviation between the distances $d(c_t, p_{v_t})$ and $d(C_t, \alpha_{v_t})$, the larger the translation of p_n given by $\vec{q}_n$. In particular, the contribution of a triangle $t \in T_n$ in translating p_n away from the centroid c_t is given by the weight w'_t, which is large for triangles with a small area compared to their B-dimensional counterpart.

This results in a progressive space relaxation, meaning that the difference of reconstructed absorption values between adjacent points is minimized as shown in Fig. 3. The top row of the figure shows the distribution of the 2D points at different steps of the relaxation algorithm. Furthermore, the bottom row of Fig. 3 shows how the points distribution affects the rate of change in absorption values when exploring the space. To generate these plots, the space was sampled on a 600 × 600 grid, the absorption coefficients

[2] https://doi.org/10.5281/zenodo.14389617.



(a) (b) (c) (d)

Fig. 3. On the top row, the proposed 2PS at different steps, where colors from yellow to blue are associated with increasing broadband absorption. On the bottom row, the evolution of the proposed 2PS's rate of change at different steps, where lighter colors represent a lower value than darker colors. (a) Raw 2D space generated by the PCA. (b) 2D space after the first iteration with points enclosed into the unit circle. (c) 2D space after 108,000 iterations of the relaxation algorithm. (d) Final 2PS in the last iteration (324,000).

were reconstructed at each point, and the average absolute difference was computed between the point and its neighbors using a kernel 3×3 grid. The figure shows that the 2D space becomes increasingly smooth as the iteration number grows.

3.3.2 Criterion Stop

The weights w_t were used as a criterion to stop the iterative algorithm. The iterations were considered as long as the mean of the weights w_t exhibited a decreasing trend. Then, the final iteration is the one with the minimum variance of the weights w_t. Specifically, the mean and variance weighted with the triangle areas were calculated. These values across the iterations are reported in the supplementary materials.[2] After the last iteration, the 2D points are again centered in the origin and scaled to the unit circle. In addition, they are rotated so that the point corresponding to the completely reflective material is situated in $(0, -1)$.

3.3.3 Inverse Mapping

The absorption coefficients for an arbitrary point p_n in the 2PS are obtained by inverse mapping. Specifically, the absorption coefficients associated with the vertices of the triangle where p_n lies are interpolated. The interpolation is performed using the barycentric coordinates of p_n as weights. This allows for a smooth transition when moving between adjacent triangles, as distant vertices have a negligible weight in the interpolation.

The choice of PCA resulted in axes exhibiting an interpretable meaning. Fig. 4 illustrates this meaning, showing the absorption coefficients derived from the inverse mapping of values sampled along the x-axis and y-axis of the 2PS. It is observed that the x-axis is related to average broadband absorption. Conversely, the y-axis is related to a frequency-dependent effect, varying from a low-pass to a high-pass behavior.

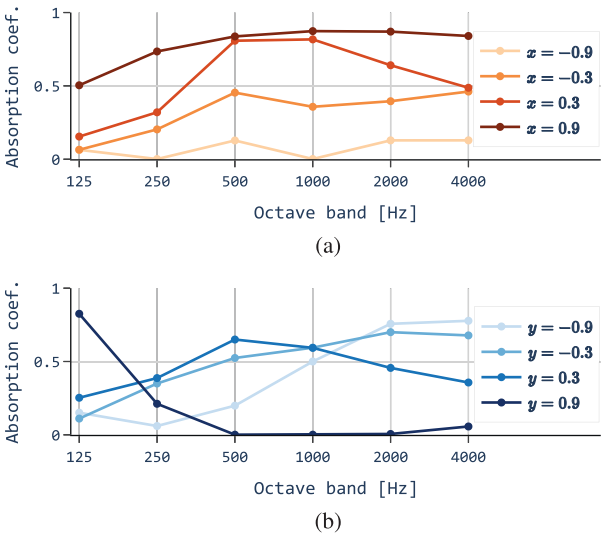


Fig. 4. Absorption coefficients obtained with inverse mapping from the 2PS. (a) The 2PS has been sampled at regular steps in the range $[-0.9, 0.9]$ of the x-axis, with $y = 0$ (orange lines), and (b) vice versa for the y-axis (blue lines).

4 AUTOMATIC TUNING OF ABSORPTION COEFFICIENTS

4.1 Reverb Matching Method

Reverb matching is the task of imitating the reverberation of a target environment. Several methods have been proposed in the literature to this end [26, 27]. This section builds upon a previous reverb matching method [13, 14]. This method was originally proposed to match the late reverberation of a target stereo RIR by tuning the parameters of a delay network. In this work, the method is adapted to match a target SRIR by tuning the SDN's absorption coefficients, with the other parameters assumed

as known. Specifically, the method navigates the 2PS to estimate the octave-band coefficients for each wall of the simulated shoebox. Given the assumptions made, the reverb matching presented here is not intended as a fully developed method for unsupervised scenarios but rather as a benchmark to assess the potential benefits of the 2PS over other approaches.

4.1.1 Method Workflow

This section describes the workflow of the proposed reverb matching method, which is shown in Fig. 5. Conversely, SEC. 4.1.2 provides details on the matching procedure. Let **h** be the target SRIR in B-format Ambisonic. It is assumed that **h** has been generated in a shoebox room, defined by the geometrical parameters Θ, including the room sizes; the positions x_S and x_R of the source and receiver, respectively; and the receiver orientation. The room wall $m \in \{1, \ldots, M\}$ is made of a material with absorption coefficients α_m^b, where $b \in \{1, \ldots, B\}$ is the octave band. Given **h** and Θ, the proposed method navigates the 2PS to estimate the coefficients α_m^b.

In the method workflow, seven monophonic audio signals are extracted from the SRIR **h**: the omnidirectional channel h_0 and six signals h_m, each related to a wall m. The matching procedure is first performed for h_0 to obtain the 2D parameters $\hat{p}_0$, which represent the overall absorption of the room, regardless of the directional effects. The signals h_m are employed to estimate the 2D parameters $\hat{p}_m$ of each wall m, with $\hat{p}_0$ used as a starting point. Specifically, each signal h_m represents the first-order reflection from wall m, allowing to disregard reflections originating from other walls. During the matching procedure, this first-order reflection serves as the reference for tuning the absorption coefficients, as it captures the spectral coloration associated with the corresponding wall. The reflection is extracted using a beamforming filter followed by a windowing operation, as described below.

4.1.1.1 Beamforming. Each signal h_m is obtained by isolating from **h** the sound reflections coming from the wall m. To this end, beamforming was performed, which is the spatial filtering in the spherical harmonics domain. For each wall m, the beamforming filter $\beta(\theta_m, \varphi_m)$ is applied to **h** to isolate the sound component at azimuth θ_m and elevation φ_m, which is the angle between the receiver x_R and the scattering node of the wall m. Beamforming was performed using the hyper-cardioid directivity pattern and implemented using the SPARTA suite [28].

4.1.1.2 Windowing. Two time windows are applied to each signal h_m. The window w_{in} removes any sound preceding the first-order reflection from the wall m:

$$w_{in}(t) = \begin{cases} 0 & t \le \gamma_m - l_{in} \\ f_{in}(t) & \gamma_m - l_{in} < t \le \gamma_m \\ 1 & t \le t_0 \end{cases}, \tag{3}$$

where γ_m is the instant when the first-order reflection from the wall m reaches the receiver, which is computed assuming a sound speed of 343 m/s. A cosine fade-in f_{in} of length $l_{in} = 1$ ms is applied before γ_m for a smooth transition. The earliest peak at time λ_m is identified in the windowed signal. This peak is searched among those greater than 10% of the maximum absolute amplitude of the signal. The function `find_peaks` from Scipy[3] was used to calculate the peaks. The second time window w_{out} consists of a cosine fade-out of 20 ms starting at λ_m. This window is intended to remove the reflections after the first one.

4.1.2 Matching Procedure

Let h_i with $i \in \{0, \ldots, M\}$ be the target signals to match. When $i = 0$, h_0 is the omnidirectional channel of the target SRIR **h**. When $i \in \{1, \ldots, M\}$, h_i is extracted from **h** with the beamforming filter $\beta(\theta_i, \varphi_i)$ and windowing operations described above, thus representing the isolated first-order reflection from wall i. Similarly, the corresponding signal $\hat{h}_i$ is extracted from the SRIR $\hat{\mathbf{h}}$ generated by the SDN given the known parameters Θ and the absorption coefficients $\hat{\alpha}_i$. The coefficients $\hat{\alpha}_i$ are reconstructed via inverse mapping from the 2D parameters $\hat{p}_i$ estimated at a given iteration. The matching procedure iteratively tunes the parameters $\hat{p}_i$ to minimize the difference between h_i and $\hat{h}_i$.

The iterative tuning of the absorption coefficients is based on a Bayesian optimization using a Gaussian process as a prior [29]. The objective is to minimize the loss function ℓ between H_i and $\hat{H}_i$, which represent the spectra of h_i and $\hat{h}_i$, respectively. The logarithmic spectral distance is used as the loss function:

$$\ell(H_i, \hat{H}_i) = \sqrt{\frac{1}{|F|} \sum_{f \in F} \left(20 \log_{10} \frac{|H_i(f)|}{|\hat{H}_i(f)|} \right)^2} \quad \text{[dB]}, \tag{4}$$

where F is the set of considered frequency bins that start from 20 Hz until the Nyquist frequency excluded.

The optimization procedure assumes that ℓ follows a multivariate Gaussian distribution. The Matérn kernel [30, SEC. 4.2] is used as the covariance function between the current parameters and the new candidate ones. At each iteration, an acquisition function A selects the next set of candidate parameters $\hat{p}_i$ from the 2PS. Bayesian optimization was implemented using Scikit-Optimize,[4] with the option `gp_hedge` selected for the acquisition function A. The iteration limit of the matching procedure was set to 150, as larger values yielded negligible improvements.

4.1.3 Refinement of Reverberation Time

In a preliminary evaluation of the reverb matching method, it was observed that the absorption coefficients were estimated with reasonable accuracy; however, this did not always translate into a good match of the reverberation characteristics. In particular, notable mismatches were found between the reverberation times T_{60} of the target and the estimated SRIRs. To address this, a refinement step was introduced to slightly adjust the estimated absorption coefficients with a broadband offset. Since T_{60} is inversely

[3] https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.find_peaks.html.

[4] https://doi.org/10.5281/zenodo.1157319.

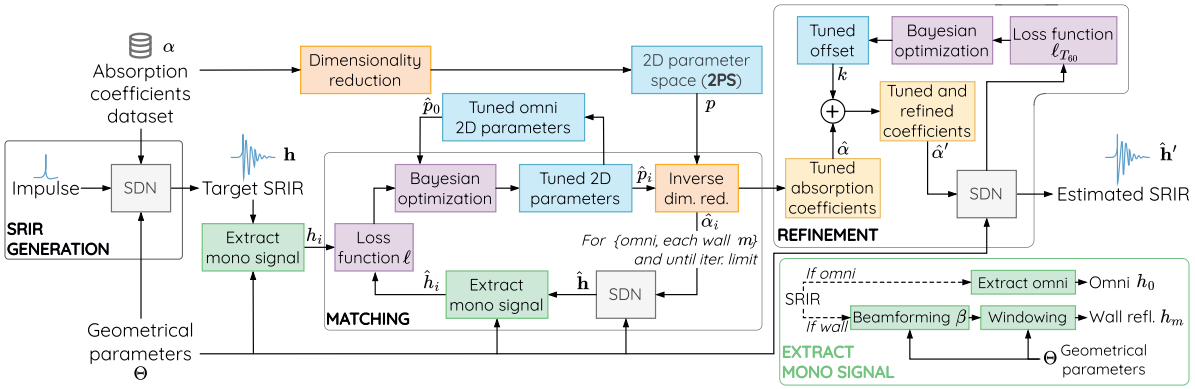


Fig. 5. Workflow of the reverb matching method for tuning the SDN's absorption coefficients to match a target SRIR.

related to the overall sound absorption in the room [16, 17], this broadband offset effectively shifts the global energy decay rate. The rationale behind this refinement is to decouple the estimation of the spectral shape from the broadband energy absorption. The refinement step seeks the optimal offset k that, when added to the estimated coefficients $\hat{\alpha}$, minimizes the mismatch in T_{60} between the target and estimated SRIRs. Thus, the offset k is intended to compensate for potential broadband biases while preserving the spectral profile estimated in the initial matching.

The refinement uses the same matching procedure described in SEC. 4.1.2, with some key differences. Specifically, Bayesian optimization is employed to find the value of k that minimizes T_{60}. At each iteration, the refined coefficients $\hat{\alpha}'$ are obtained from the previously tuned coefficients $\hat{\alpha}$ as $\hat{\alpha}' = \hat{\alpha} + k$. The refined coefficients $\hat{\alpha}'$ and the parameters Θ are used to set the SDN, which generates the SRIR $\hat{\mathbf{h}}'$. The objective of the optimization is to minimize the loss function $\ell_{T_{60}}$, defined as the mean absolute difference in T_{60} between $\hat{\mathbf{h}}'$ and the target SRIR $\mathbf{h}$:

$$\ell_{T_{60}}(\mathbf{h}, \hat{\mathbf{h}}') = \frac{1}{C} \sum_{c=0}^{C} \left| T_{60}(\mathbf{h}[c]) - T_{60}(\hat{\mathbf{h}}'[c]) \right| \text{[ms]}, \qquad (5)$$

where the reverberation time T_{60} is calculated for each channel c of the SRIRs. As the refinement step only targets reverberation time, it omits beamforming, windowing, and dimensionality reduction. The number of optimization iterations was limited to 150. The offset k was searched within the range $[-0.2, +0.2]$, and the resulting coefficients were clipped to the valid range [0, 1].

4.2 EVALUATION

4.2.1 Evaluation Procedure

4.2.1.1 Target Spatial Room Impulse Responses. The reverb matching method was evaluated on 16 target SRIRs in fourth-order Ambisonics. The target SRIRs were generated by the SDN whose parameters were defined as follows. For each wall, the absorption coefficients were sampled from the materials in the dataset described in SEC. 3.1. The geometrical parameters Θ were randomly selected. The room sizes were sampled from a Gaussian distribution with a mean of 7 m and standard deviation of 5 m. The normalized source and receiver positions were sampled from a uniform distribution between 0.01 and 0.99. The receiver orientation was set to 0, while the line of sight and air absorption parameters were disabled. The line of sight parameter was enabled only for the refinement step.

4.2.1.2 Conditions. The reverb matching performances were evaluated with four conditions. The condition *2PS* corresponds to the coefficients search in the proposed 2PS using the reverb matching method. For the condition *OctBand*, the reverb matching method directly tunes the octave-band absorption coefficients; thus, no dimensionality reduction is involved. In the condition *PCA*, the method operates in the 2D space generated by PCA without relaxation [see Fig. 3(a)]. To prevent a biased estimation, the absorption coefficients used to generate the target SRIRs were removed from the 2D spaces of the conditions *2PS* and *PCA*. The refinement step was performed for all the conditions *2PS*, *OctBand*, and *PCA*. Finally, the condition *Baseline* is a custom naive approach used as a control condition. For each wall m, the RMS is calculated for each octave b of the target magnitude spectrum H_m. The RMS values are linearly scaled, clipped within the range [0, 1], and inverted to obtain the estimated baseline absorption coefficients.

4.2.1.3 Metrics. The reverb matching performances are quantified by the absolute difference between the target absorption coefficients α_m^b and the estimated ones $\hat{\alpha}_m^b$. For each SRIR, the overall error is calculated by averaging the differences in each band b and wall m. Then, acoustic features on the SRIRs were measured, including reverberation time T_{60}, early decay time *EDT*, clarity index C_{80}, and center time T_S. For each SRIR channel, the absolute difference between the acoustic features of the target SRIR $\mathbf{h}$ and the corresponding estimated SRIR $\hat{\mathbf{h}}$. The overall feature error was calculated by averaging the differences of each channel.

4.2.2 Results

Table 1 shows the error metrics obtained for each condition. *2PS* exhibits the lower error for the matching of absorption coefficients. As expected, *Baseline* condition has the highest error, followed by the *OctBand* and *PCA*. A Friedman test showed a significant difference between

Table 1. Reverb matching performances for different conditions, quantified as the MAE between the metrics computed for the target and estimated SRIRs.

MAE	α	T_{60} [ms] (%)	*EDT* [ms] (%)	C_{80} [dB] (%)	T_S [ms] (%)
Baseline	0.255 ± 0.212	24.3(20.3) ± 31.2	1.5(5.1) ± 2.3	4.6(19.7) ± 4.4	1.3(5.0) ± 1.7
OctBand	0.169 ± 0.190	3.8(3.9) ± 8.1	0.1(0.8) ± 0.2	**0.5**(**1.3**) ± 0.8	0.1(0.6) ± 0.1
PCA	0.122 ± 0.125	3.2(4.0) ± 4.7	0.2(0.9) ± 0.2	0.9(3.1) ± 1.5	0.2(0.9) ± 0.3
2PS	**0.102** ± 0.106	**2.6**(**3.2**) ± 4.1	**0.1**(**0.7**) ± 0.2	0.6(1.3) ± 1.0	**0.1**(**0.5**) ± 0.1

Note. Metrics include the absorption coefficients α and acoustic features, for which the percentage error in relation to the feature measured on the target SRIR is also reported in parentheses. Bold text indicates the best metric value for each column. MAE = mean absolute error.

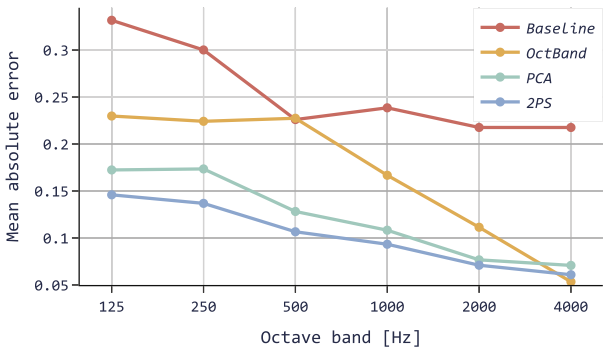
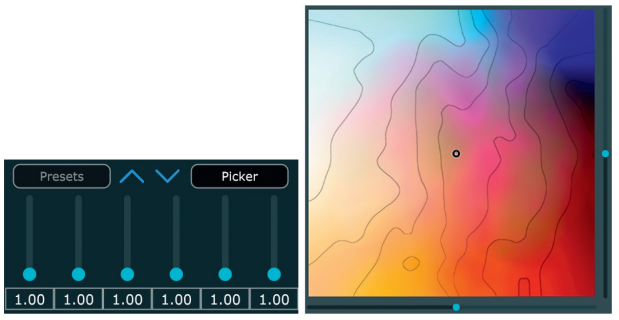


Fig. 6. Mean absolute error for each octave band between the target absorption coefficients and those estimated with the four conditions of reverb matching.



(a) (b)

Fig. 7. The two interfaces compared in the usability test. (a) *OctBand*; (b) *2PS*.

the conditions ($p < 0.001$). A post hoc Nemenyi test indicated that all differences between the conditions were significant, except for the conditions *PCA* and *2PS*. Similar statistical tests were conducted for the acoustics parameters and found that *2PS* significantly outperformed all the other conditions for T_s. For T_{60}, *2PS* significantly outperformed *OctBand* but not *PCA*, while for C_{80}, *2PS* outperformed *PCA*, but no significant difference was found with *OctBand*. Regarding the acoustic feature, Table 1 also shows the percentage error in relation to the feature measured on the target SRIR. The *2PS* and *PCA* conditions achieved better performances in this regard. *2PS* achieved the lowest error for the reverberation time T_{60} with a percentage error of 3.2%.

Fig. 6 shows the mean absolute error between the target and estimated absorption coefficients for each band *b*. A decreasing trend was observed for the error as the band increased. *2PS* exhibits the lowest error for each band except the last one, despite the improvement over *PCA* getting smaller at higher bands.

The superior performance of *2PS* over *PCA* is likely attributable to the "relaxed" properties of the parameter space. While PCA yields an irregular points distribution [see Fig. 3(a)], the proposed relaxation algorithm rearranges the points to obtain a more uniform distribution without abrupt transitions yet expands information-dense areas. This results in a smoother loss landscape for the Bayesian optimizer, thereby mitigating the convergence to local minima and facilitating precise fine tuning.

5 USABILITY TEST

5.1 Experimental Design

5.1.1 Task and Setup

A usability test was conducted to subjectively evaluate the proposed 2PS for absorption coefficients. The technical setup included a laptop connected to a 27-in display, a pair of Sennheiser HD 650 headphones, and a MOTU UltraLite mk3 Hybrid audio interface. The participant was presented with two tracks in the DAW Reaper. The first track included the SDN plugin parametrized with target octave-band absorption coefficients. All SDN walls shared the same coefficients. Furthermore, the SDN's geometrical parameters and the stereo output mode were fixed for all target reverberations. Participants were free to listen to the three audio samples included in the track (speech, snares, and instrumental music). The second track was a replica of the first, except for the SDN absorption coefficients, which were initialized to 1 to obtain a dry output. During the test, participants were instructed to manually tune the coefficients of the second track to achieve a reverberation effect as similar as possible to the first track. The change in a coefficient for a wall directly affected all the other walls.

Participants tuned the SDN absorption coefficients with the two interfaces shown in Fig. 7: *OctBand* and *2PS*. *OctBand* is equipped with six sliders, one for each octave band of the absorption coefficients. *2PS* features a colored square representing the proposed 2PS. By moving the cursor within the square, the user can navigate the absorption coefficients reconstructed through inverse mapping from the 2PS. The reconstructed absorption coefficients were visible when us-

Table 2. Results of the usability test in terms of the mean and standard deviation of the absolute error between the target and manually tuned absorption coefficients, trial duration, and SUS scores.

	Coefficients Error	Duration (s)	SUS Score
OctBand	0.108 ± 0.111	266.0 ± 58.6	67.29 ± 17.1
2PS	**0.074** ± 0.076	**236.5** ± 72.2	**74.79** ± 17.5

Note. Bold text indicates the best metric value for each column.

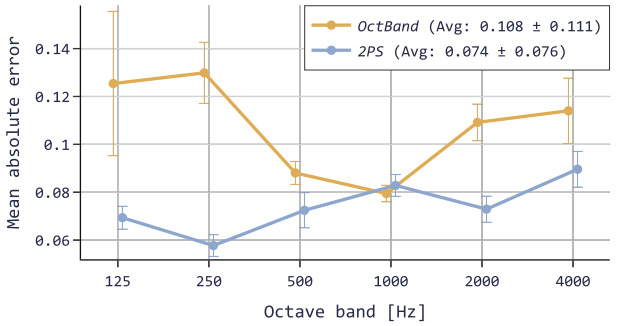


Fig. 8. Mean absolute error for each octave between the target and manually tuned absorption coefficients in the usability test. The whiskers represent the error variance.

ing the *2PS* interface. The square's colors were obtained by mapping the reconstructed absorption coefficients for each 2D point to an RGB triplet. The red component was associated with the mean of the first two octaves, the third and fourth octaves were associated with green, and the fifth and sixth octaves were associated with blue. The colors were then smoothed, and contours were added to enclose similar broadband absorption values.

5.1.2 Protocol and Metrics

Before the test, participants were allowed to familiarize themselves with the two interfaces for a couple of minutes each. The usability test comprised six trials, each with a distinct set of target absorption coefficients. The participants were asked to complete each trial within five minutes. Participants alternatively used one interface for the first half of trials and the other interface for the second half. The Latin square design was adopted to distribute the six sets of target coefficients across participants, such that an equal distribution of interface-set pairs every 12 participants was obtained.

Objective and subjective metrics were collected. Following each half of trials, participants completed the system usability scale (SUS; in Italian [31]) for the used interface. Upon completion of the test, participants responded to a final questionnaire with questions requiring them to select one of the two interfaces or neither (see Table 3). The task performances were objectively measured by the duration of each trial and the absolute error between the target coefficients and those manually set by the participants.

5.2 Results

The test was conducted by 12 experts with at least two years of experience in the field of audio and music production (mean: 10.6 ± 9.4 years). Table 2 shows the test results in terms of the error in the absorption coefficients tuning, trial duration, and SUS score. The proposed *2PS* interface outperformed *OctBand* across all these metrics. Fig. 8 shows the mean error in the coefficients tuning for each octave. *2PS* improves the error for the octaves at the extremities, where *OctBand* yielded higher errors. *2PS* also exhibits a lower variance of the error at these octaves, particularly at 125 Hz. Conversely, *OctBand* and *2PS* exhibited similar errors for the middle octaves, where *OctBand* yielded a lower error compared to the other octaves. This could be attributed to the higher sensitivity of the human ear to these frequencies. A Mann-Whitney U test indicated a significant difference between the two interfaces for the distribution of errors across participants and octaves ($p < 0.001$)—null hypothesis of normality was rejected for both distributions ($p < 0.001$).

Regarding trial duration, participants were on average about 30 s faster in conducting the task when using *2PS* compared to *OctBand*. Further, *2PS* yielded a SUS score greater than 7.5 compared to *OctBand*. Nevertheless, no significant difference was found for duration and SUS score between the two interfaces according to Mann-Whitney U and Wilcoxon signed-rank tests, respectively.

Table 3 reports the results of the final questionnaire regarding the participants preferences for the two interfaces. The majority of participants indicated that *2PS* facilitated the exploration of the reverberation effects. Conversely, most of the participants indicated that *OctBand* allowed them to achieve better results. Interestingly, this subjective feeling is in contrast with the objective performances of the task, wherein *2PS* exhibited a lower error for the coefficients tuning. This is probably caused by a series of compounding factors. Given the prior experience of the users selected for this study, it is reasonable to expect them to identify the more familiar *OctBand* as the easier to use and better performing interface. This is a meaningful insight into user behavior, as interface familiarity can bias subjective assessments of task performance. Additionally, the design of the *2PS* interface is intentionally more suited for a free-form exploration of the existing possibilities rather than a precision-based task. This supports the perceived ease of use and performance of *OctBand*.

The comments in the open section of the questionnaire reinforce these hypotheses, as the subjects indicated that the *2PS* interface fosters creativity in parameter exploration and may be more intuitive and user-friendly for beginner users after proper instructions. Conversely, they deemed that expert users, such as themselves, may feel more confident with *OctBand*, which allows for finer tuning of absorption coefficients.

In addition, participants suggested some improvements for *2PS*, such as including labels to describe the two axes. In a further question, all participants indicated that they would include both *OctBand* and *2PS* in their own interface. This was the original intention for integrating *2PS* in this SDN implementation, where the desired reverb effect is

Table 3. Results of the final questionnaire in which participants answered selecting one of the two interfaces or neither.

Question	*OctBand* (%)	*2PS* (%)	Indifferent (%)
Which mode of interaction provided the easiest way to explore the variety of possible reverberations?	25.0	66.7	8.3
Which mode of interaction enabled you to achieve better results?	58.3	25.0	16.7
Which mode of interaction allowed you to focus more on the sound?	25.0	25.0	50.0
Which mode of interaction did you find easiest to accomplish the required task?	58.3	41.7	0.0
In general, which mode of interaction did you prefer?	33.3	50.0	16.7

searched in the 2PS, followed by the fine-tuning with the sliders. Finally, most of the participants indicated an overall preference of the *2PS* interface.

6 CONCLUSION

This paper addressed the challenge of individually tuning octave-band absorption coefficients by proposing a 2PS to facilitate their exploration. The 2PS was constructed by applying a custom dimensionality reduction method to a dataset of absorption coefficients. The proposed 2PS was designed and evaluated for the parameters of an SDN. The 2PS has the potential to improve the conventional octave-band representation in both the machine-driven and user-driven identification of target coefficients. This is supported by the findings of a twofold evaluation encompassing (a) an automated reverb matching approach and (b) a usability test with expert participants. The use of 2PS led to improved performance in both evaluations, although some results remain mixed. These findings highlight the need for further refinements, particularly concerning the reverb matching procedure.

The proposed approach is potentially adaptable to streamline the navigation of other audio effects exhibiting numerous parameters. A well-suited potential use case is represented by equalizers. In future works, the reverb matching method will be evaluated on SRIRs not generated by the SDN. Further, an extended subjective evaluation will be conducted to assess user interaction with the 2PS on different tasks.

7 ACKNOWLEDGMENT

This work is part of SONICOM, a project that has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 101017743.

8 REFERENCES

[1] V. Valimaki, J. D. Parker, L. Savioja, J. O. Smith, and J. S. Abel, "Fifty Years of Artificial Reverberation," *IEEE Trans. Audio Speech Lang. Process.*, vol. 20, no. 5, pp. 1421–1448 (2012 Jul.). https://doi.org/10.1109/TASL.2012.2189567.

[2] E. De Sena, H. Hacıhabiboğlu, Z. Cvetković, and J. O. Smith, "Efficient Synthesis of Room Acoustics via Scattering Delay Networks," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 23, no. 9, pp. 1478–1492 (2015 Sep.). https://doi.org/10.1109/TASLP.2015.2438547.

[3] M. Fontana, G. Presti, D. Fantini, F. Avanzini, and A. Reyes-Lecuona, "A Highly Parametrized Scattering Delay Network Implementation for Interactive Room Auralization," in Proc. *27th Intl. Conf. Digital Audio Effects (DAFx)*, pp. 286–293 (Guildford, UK) (2024 Sep.).

[4] J. Huopaniemi, L. Savioja, and M. Karjalainen, "Modeling of Reflections and Air Absorption in Acoustical Spaces a Digital Filter Design Approach," in Proc. *Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)* (New Paltz, NY) (1997 Oct.). https://doi.org/10.1109/ASPAA.1997.625594.

[5] S. Mecklenburg and J. Loviscach, "subjEQt: Controlling an Equalizer Through Subjective Terms," in *CHI'06 Extended Abstracts on Human Factors in Computing Systems*, pp. 1109–1114 (ACM, New York, NY, 2006). https://doi.org/10.1145/1125451.1125661.

[6] A. T. Sabin and B. Pardo, "2DEQ: An Intuitive Audio Equalizer," in Proc. *7th ACM Conf. Creativity and Cognition*, pp. 435–436 (Berkeley, CA) (2009 Oct.). https://doi.org/10.1145/1640233.1640339.

[7] M. Cartwright, B. Pardo, and J. Reiss, "MIXPLORATION: Rethinking the Audio Mixer Interface," in Proc. *19th Intl. Conf. Intelligent User Interfaces*, pp. 365–370 (Haifa, Israel) (2014 Feb.). https://doi.org/10.1145/2557500.2557530.

[8] S. Stasis, R. Stables, and J. Hockman, "Semantically Controlled Adaptive Equalisation in Reduced Dimensionality Parameter Space," *Appl. Sci.*, vol. 6, no. 4, paper 116 (2016 Apr.). https://doi.org/10.3390/app6040116.

[9] S. Stasis, J. Hockman, and R. Stables, "Navigating Descriptive Sub-Representations of Musical Timbre," in Proc. *Intl. Conf. New Interfaces for Musical Expression*, pp. 56–61 (Copenhagen, Denmark) (2017 May). https://doi.org/10.5281/zenodo.1176171.

[10] M. Peachey, S. Oore, and J. Malloch, "Creating Latent Representations of Synthesizer Patches Using Variational Autoencoders," in Proc. *4th Intl. Symp. Internet of Sounds*, pp. 1–7 (Pisa, Italy) (2023 Oct.). https://doi.org/10.1109/IEEECONF59510.2023.10335466.

[11] P. Esling, N. Masuda, A. Bardet, R. Despres, and A. Chemla-Romeu-Santos, "Flow Synthesizer: Universal Audio Synthesizer Control With Normalizing Flows," *Appl. Sci.*, vol. 10, no. 1, paper 302 (2020 Jan.). https://doi.org/10.3390/app10010302.

[12] K. Tahiroğlu, M. Kastemaa, and O. Koli, "GanSpaceSynth: A Hybrid Generative Adversarial Network Architecture for Organising the Latent Space Using a Dimensionality Reduction for Real-Time Audio Synthesis," in Proc. *Conf. AI Music Creativity* (Online) (2021 Jul.). https://doi.org/10.5281/zenodo.5137902.

[13] R. Bona, D. Fantini, G. Presti, et al., "Automatic Parameters Tuning of Late Reverberation Algorithms for Audio Augmented Reality," in Proc. *17th Intl. Audio Mostly Conf.*, pp. 36–43 (St. Pölten, Austria) (2022 Sep.). https://doi.org/10.1145/3561212.3561236.

[14] D. Fantini, G. Presti, M. Geronazzo, et al., "Co-Immersion in Audio Augmented Virtuality: The Case Study of a Static and Approximated Late Reverberation Algorithm," *IEEE Trans. Visualiz. Comput. Graphics*, vol. 29, no. 11, pp. 4472–4482 (2023 Nov.). https://doi.org/10.1109/TVCG.2023.3320213.

[15] ISO, "Acoustics — Measurement of Sound Absorption in a Reverberation Room," *Standard 354:2003* (2003 May).

[16] W. C. Sabine, *Collected Papers on Acoustics* (Harvard University Press, Cambridge, MA, 1922).

[17] C. F. Eyring, "Reverberation Time in 'Dead' Rooms," *J. Acoust. Soc. Am.*, vol. 1, no. 2A, pp. 217–241 (1930 Jan.). https://doi.org/10.1121/1.1915175.

[18] D. González-Toledo, L. Molina-Tanco, M. Cuevas-Rodriguez, P. Majdak, and A. Reyes-Lecuona, "The Binaural Rendering Toolbox. A Virtual Laboratory for Reproducible Research in Psychoacoustics," in Proc. *Forum Acusticum*, pp. 331–337 (Turin, Italy) (2023 Sep.). https://doi.org/10.61782/fa.2023.1042.

[19] J. Huopaniemi, *Virtual Acoustics and 3-D Sound in Multimedia Signal Processing*, Ph.D. thesis, Helsinki University of Technology, Espoo, Finland (1999 Nov.).

[20] M. Vorländer, *Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms, and Acoustic Virtual Reality* (Springer, Berlin, Germany, 2007). https://doi.org/10.1007/978-3-540-48830-9.

[21] Odeon A/S, "Global Material Library," https://odeon.dk/download/materials/Material.Li8 (accessed 12 Dec. 2024).

[22] Acoustic Traffic LLC, "Absorption Coefficients," https://www.acoustic.ua/st/web_absorption_data_eng.pdf (accessed 12 Dec. 2024).

[23] M. Egan, *Architectural Acoustics* (J. Ross Publishing, Plantation, FL, 2007).

[24] JCW Acoustic Supplies, "Absorption Coefficients of Common Building Materials and Finishes," https://www.acoustic-supplies.com/absorption-coefficient-chart/ (accessed 12 Dec. 2024).

[25] S. Lloyd, "Least Squares Quantization in PCM," *IEEE Trans. Inform. Theory*, vol. 28, no. 2, pp. 129–137 (1982 Mar.). https://doi.org/10.1109/TIT.1982.1056489.

[26] S. J. Schlecht and E. A. Habets, "Accurate Reverberation Time Control in Feedback Delay Networks," Proc. *Intl. Conf. Digital Audio Effects (DAFx)*, pp. 337–344 (Edinburgh, UK) (2017 Sep.).

[27] K. Prawda, V. Välimäki, and S. Schlecht, "Improved Reverberation Time Control for Feedback Delay Networks," in Proc. *Intl. Conf. Digital Audio Effects (DAFx)*, paper 46 (Birmingham, UK) (2019 Sep.).

[28] L. McCormack and A. Politis, "SPARTA & COMPASS: Real-Time Implementations of Linear and Parametric Spatial Audio Reproduction and Processing Methods," in Proc. *AES Intl. Conf. Immersive and Interactive Audio* (2019 Mar.), e-Brief 111.

[29] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian Optimization of Machine Learning Algorithms," in P. Bartlett, F. Pereira, C. J. Burges, L. Bottou, and K. Q. Weinberger (eds.), *Advances in Neural Information Processing Systems 25*, pp. 2951–2959 (Neural Information Processing Systems, La Jolla, CA, 2012).

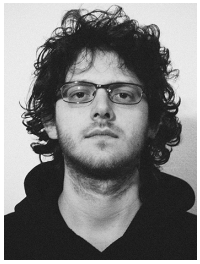
[30] C. K. Williams and C. E. Rasmussen, *Gaussian Processes for Machine Learning*, vol. 2 (MIT Press, Cambridge, MA, 2006).

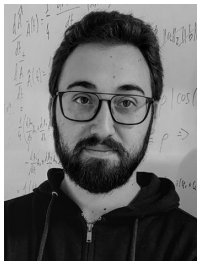
[31] S. Borsci, S. Federici, and M. Lauriola, "On the Dimensionality of the System Usability Scale: A Test of Alternative Measurement Models," *Cogn. Process.*, vol. 10, no. 3, pp. 193–197 (2009 Aug.). https://doi.org/10.1007/s10339-009-0268-9.

THE AUTHORS


Marco Fontana


Davide Fantini


Giorgio Presti


Marco Tiraboschi


Federico Avanzini

Marco Fontana is a research fellow in the Department of Computer Science at the University of Milan. He received his M.Sc. degree in computer science from the University of Milan in 2023. He is involved in the SONICOM project, funded under the EU's Horizon2020 Programme, which leverages artificial intelligence to design immersive audio technologies. His research interests include spatial audio, extended reality, and digital reverberation.

•

Davide Fantini is a research fellow in the Department of Computer Science at the University of Milan. He received his M.Sc. and Ph.D. degrees in computer science from the University of Milan in 2019 and 2024, respectively. He is involved in the SONICOM project, funded under the EU's Horizon2020 Programme, which leverages artificial intelligence to design immersive audio technologies. His research interests include head-related transfer functions, binaural rendering, artificial reverberation, and spatial hearing.

•

Giorgio Presti is an adjunct professor at the Department of Computer Science at the University of Milan. He carries out research in the field of sound and music computing, specialized in the creation of new digital tools for music signal analysis and production and aimed at exploring the natural phenomena from which the music experience emerges. He is also a sound designer, author of interactive multimedia installations, and electroacoustic performances.

•

Marco Tiraboschi is a research fellow in the Department of Computer Science at the University of Milan. He earned his M.Sc. and Ph.D. in computer science from the University of Milan in 2020 and 2024, respectively. He is currently involved in the S-TWIN project, funded by the PRIN program, which investigates novel applications of the Digital Twin paradigm for patients with cochlear implants. His research interests span physically motivated acoustic analysis, mathematical models of sound, and the intersections of music and computational neuroscience.

•

Federico Avanzini is a full professor in the Department of Computer Science at the University of Milan. His main research interests concern algorithms for sound synthesis and processing and 3D sound rendering, with applications to such domains as assistive technologies, virtual musical instruments, digital cultural heritage, and digital learning. He has been principal investigator and scientific responsible of EU, national, and industry-funded projects. He is currently serving as conference coordinator in the permanent Steering Committee of the Sound & Music Computing Conference and Summer School and president of the Italian Music Informatics Association (AIMI).