Proceedings of Meetings on Acoustics
International Symposium on Musical Acoustics 2026 (ISMA 2026)
Helsinki, Finland, 15–17 June 2026

Spatial modal reverberation

Michele Ducceschi,[a] Craig Webb,[a] Matteo Gualeni[b] and Riccardo Russo[a]

[a]*University of Bologna, Bologna, Italy*; michele.ducceschi@unibo.it
[b]*Accademia del Suono, Milan, Italy*

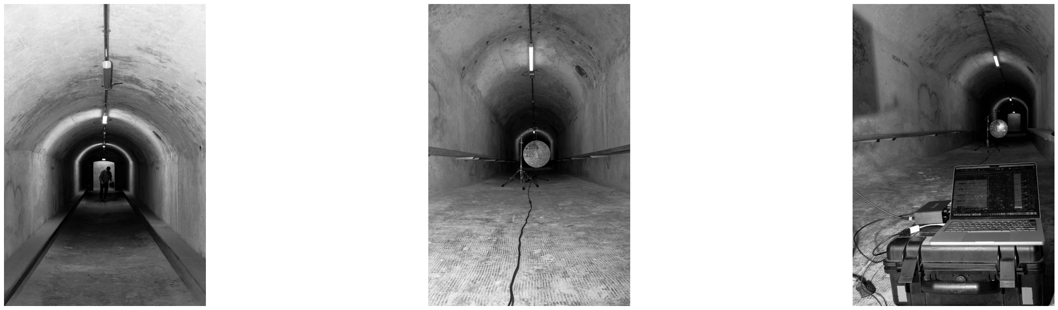
Spatial reverberation is a topic of growing interest in the audio industry. Many applications, such as virtual and augmented reality, gaming, and others, rely on the spatialisation of an acoustic field to achieve realistic audio rendering. Many established techniques exist for the purpose. Some, such as ambisonics, rely on a codification of the entire audio field. Others, such as Head-Related Transfer Function (HRTF) techniques, rely on an encoding of the receiver in the audio field. Dynamic rendering of a translating receiver is possible but usually comes at a cost, either through increased computational load (e.g., for HRTFs) or reduced perceptual rendering (e.g., for ambisonics). Here, we explore the possibility of rendering a moving receiver dynamically inside an acoustic space via a modal architecture. First, the poles of the acoustic space are retrieved for a number of source–receiver positions within the space. Then, the modal weights are computed at each location. Since the poles are invariant with location, the modal weights can be interpolated across locations to render dynamic reverberation at no additional cost, and without loss of quality.

1. INTRODUCTION

Artificial reverberation is a basic tool in music production, sound design, post-production, and a growing number of immersive applications such as gaming, virtual and augmented reality, and architectural auralisation.[17,18] Convolution reverberators reproduce the character of a measured room with very high fidelity but offer little parametric flexibility, and–more importantly for this paper–no direct mechanism for moving the listener inside the captured space without resampling a dense set of impulse responses (IRs). Algorithmic designs, from Schroeder allpass networks[1] to feedback delay networks (FDNs),[2,3] including their recent differentiable variants,[4–6] are efficient and tunable, but the mapping between internal coefficients and acoustic descriptors is indirect and the listener is generally a fixed point in an abstract topology. Wave-based solvers[7,8] achieve high physical accuracy but remain too costly for interactive use, particularly for large-volume spaces.

The middle path explored here is modal synthesis:[9,10] the impulse response is represented as a sum of damped sinusoids whose parameters—frequency, damping, and complex residue—carry direct acoustic meaning. In a recent contribution,[12] we showed that a sub-band frequency-domain implementation of the PolyMAX algorithm[14,15] is able to extract such modal parameters from measured IRs of large historical spaces in a robust and computationally tractable manner, supporting real-time synthesis of tens of thousands of oscillators on commodity hardware. In a complementary corpus-driven study,[13] the modal representation was further used to derive a parametric reverberator from a large dataset of measured rooms.

The present work extends these ideas to the spatial setting. The starting observation is that the pole set of an acoustic space, as captured by Eq. (1), is a property of the geometry and of the boundary conditions; only the residues depend on the source and receiver positions through the mode shape evaluations. If the poles are extracted jointly across a set of source–receiver positions and identified as a *common modal basis*, then per-position residues can be fitted against this fixed basis and–crucially—interpolated across positions



(a) Longitudinal view. *(b) Source location.* *(c) Acquisition rig.*

***Figure 1:** Bunkervik tunnel during the November 2025 measurement campaign. The vaulted gallery is approximately 30 m long, with cross-section roughly uniform along the longitudinal axis. The Genelec studio monitor (visible behind the gong, panel (b)) was rotated through six cardinal orientations to approximate an omnidirectional source. The acquisition laptop and audio interface (panel (c)) were placed near one end of the tunnel.*

to render a continuously translating listener with minimal additional CPU cost. The approach is illustrated through a real measurement campaign in a strongly elongated, geometrically simple space, the Bunkervik tunnel in Brescia (Italy), and culminates in a real-time JUCE implementation.

The paper is organised as follows. Section 2 describes the venue and the November 2025 measurement campaign. Section 3 details the MATLAB-based offline pipeline, from sweep deconvolution and pre-processing to per-position modal fitting. Section 4 introduces the common-basis identification step and joint IIR/FIR fit, and reports validation results. Section 5 describes the real-time spatial reverberator and presents the plugin interface. Section 6 concludes.

2. THE BUNKERVIK TUNNEL AND MEASUREMENT CAMPAIGN

The Bunkervik (also known as “Il rifugio delle idee”) is a former wartime air-raid shelter located in via Federico Odorici 6/B in central Brescia, Italy. Built in the 1940s to protect civilians from Allied bombings, the tunnel is partially excavated below street level and exhibits a single, narrow, vaulted gallery whose total length is approximately 30 m. The walls are of unfinished masonry and exposed concrete; the cross-section is roughly semi-circular and approximately uniform along the longitudinal axis. Since 2016 the venue has been progressively reopened to the public as a contemporary-art and performance space;[22] it became the subject of the present work after one of the authors (M. Gualeni) recorded a series of drum performances inside the tunnel and developed an interest in capturing its acoustic signature in a faithful but musically usable form.

Two views of the gallery are reproduced in Fig. 1. The geometry is well approximated as one-dimensional: a listener walking along the central axis experiences a smooth, monotonic evolution of the reverberant signature as a function of the longitudinal coordinate. This makes the venue an ideal testbed for the proposed spatial-modal architecture, in which the listener position parameterises a single scalar coordinate x.

A. MEASUREMENT PROTOCOL

The acoustic survey was carried out in November 2025. Sine-sweep IRs were obtained using the exponential sine sweep (ESS) method.[19] A single Genelec studio monitor was used as the excitation source, placed at one end of the tunnel and oriented successively along six cardinal directions (UP, DOWN, FRONT, BACK, LEFT, RIGHT) to approximate an omnidirectional source by post-hoc averaging. Source gains and signal levels were held constant across all sweeps. The receiver was a Beyerdynamic free-field omnidirectional condenser microphone, mounted on a stand and positioned at three representative locations along the central axis of the tunnel, at approximate distances $x_1 \approx 8$ m, $x_2 \approx 10$ m, and $x_3 \approx 25$ m from the

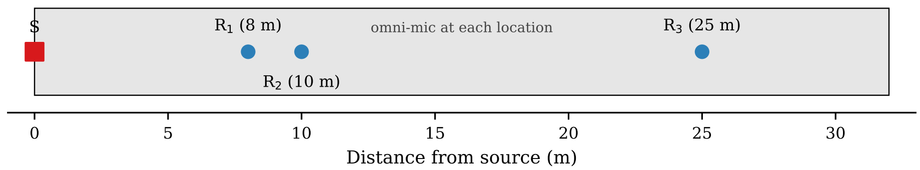


Figure 2: Schematic of the source–receiver geometry. A single Genelec source S at one end of the tunnel was rotated through six orientations; the omnidirectional receiver was placed in turn at three positions R_1, R_2, R_3 along the longitudinal axis.

source. The geometry is summarised schematically in Fig. 2. Audio was recorded at $f_s = 48$ kHz / 24 bit through a portable USB audio interface.

3. MODAL REPRESENTATION AND OFFLINE PROCESSING

The acoustic transfer function from a fixed source to a receiver at position x in an LTI room admits a pole–residue decomposition[9,11]

$$H(j\omega; x) = \sum_{k=1}^{M} \left(\frac{r_k(x)}{j\omega - \lambda_k} + \frac{r_k^*(x)}{j\omega - \lambda_k^*} \right), \tag{1}$$

with M a large but finite mode count, $\lambda_k = -\sigma_k + j\sqrt{\Omega_k^2 - \sigma_k^2}$ the k-th complex pole (with damping σ_k and undamped angular frequency Ω_k), and $r_k(x) = a_k(x) + j\, b_k(x)$ the corresponding complex residue. The poles encode the geometry and boundary losses; the residues depend on the mode-shape evaluations at the source and receiver positions. The corresponding time-domain impulse response is

$$h(t; x) = \sum_{k=1}^{M} 2\, e^{-\sigma_k t} [a_k(x) \cos \omega_{d,k} t - b_k(x) \sin \omega_{d,k} t]\,, \tag{2}$$

where $\omega_{d,k} = \sqrt{\Omega_k^2 - \sigma_k^2}$. Equation (2) can be discretised by the unconditionally stable, low-dispersion second-order recursion described in,[12,16] allowing real-time synthesis of large modal banks.

A. SWEEP DECONVOLUTION AND PRE-PROCESSING

The recorded sweeps were converted into impulse responses through the two MATLAB scripts `IR_PreProcess.m` and `IR_Deconvolve.m`, illustrated below.

Pre-processing

For each receiver location, the six orientation-specific recordings (UP, DOWN, FRONT, BACK, LEFT, RIGHT) of identical sweeps are loaded by `IR_PreProcess.m`. A user-selected channel of each stereo file is extracted (no channel averaging is performed) and the six takes per orientation, when available, are time-aligned, trimmed to the common minimum length, and arithmetically averaged. The six orientation-averaged sweeps are then summed sample-wise to approximate an omnidirectional source recording. The script consistency-checks sample rates across the takes and saves both the per-orientation averages and the combined sweep to disk in MAT format.

Frequency-domain deconvolution

The combined recorded sweep y and the original excitation sweep x are then passed to `IR_Deconvolve.m`, which carries out a Wiener / Tikhonov-regularised frequency-domain deconvolution

$$H(\omega) = \frac{X^*(\omega)\,Y(\omega)}{|X(\omega)|^2 + \alpha}, \qquad \alpha = \alpha_0 \max_{\omega} |X(\omega)|^2, \quad \alpha_0 = 10^{-6}, \tag{3}$$

with X, Y the discrete Fourier transforms of x and y, respectively, and the small bias α ensuring numerical stability against the energy loss of X at the sweep extremes. The resulting time-domain IR $h(n) = \mathcal{F}^{-1}\{H\}$ is exported as a WAV file, ready for the modal-fitting stage. The procedure is repeated independently for each of the three receiver positions, yielding three impulse responses $h^{(i)}(n)$, $i = 1, 2, 3$.

B. PER-POSITION MODAL FITTING

Each of the three IRs was independently decomposed into pole–residue form following a sub-band peak-picking and least-squares fitting strategy described in[12] and adapted here. For each IR, an aligned signal and a set of band-wise reverberation times $T_{30,b}$, clarity C_{80}, and timbre ratios were first computed. Modal frequencies f_k were then extracted by linear-band peak-picking on the magnitude FFT, and per-mode dampings σ_k were assigned from the band-wise T_{30} via $\sigma_k = 3 \ln 10 / T_{60,b}$, with a global fallback for bands lacking valid decay estimates. Once poles λ_k were fixed, the corresponding complex residues $r_k = a_k + j b_k$ were fitted by banded linear least squares against the measured spectrum

$$\min_{a_k, b_k} \left\| H_{\mathrm{tgt}}(\omega) - \sum_k \left[\frac{r_k}{j\omega - \lambda_k} + \frac{r_k^*}{j\omega - \lambda_k^*} \right] \right\|_2^2 + \mu \|[\mathbf{a}; \mathbf{b}]\|_2^2, \tag{4}$$

within $N_b = 5$ overlapping sub-bands and a small Tikhonov regularisation $\mu = 10^{-10}$. Iterative residual peak-picking refinement and a final 256-tap FIR correction filter were applied to absorb any unmodelled spectral colouration in the early portion of the IR.

The pipeline yielded $M^{(1)} = 6031$, $M^{(2)} = 6324$, and $M^{(3)} = 16216$ stable modes for the three positions, respectively, with mid-band (500–1000 Hz) modal densities of 1.12, 1.37, and 2.80 modes/Hz. The growth in mode count and density with source–receiver distance reflects the increasingly diffuse character of the late-tail field at remote receivers, where the modal overlap regime is more thoroughly engaged. Per-band T_{60} profiles, mode-frequency histograms, and median residue magnitudes are summarised in Fig. 3.

4. COMMON MODAL BASIS AND JOINT IIR/FIR FIT

The independent fits of Section 3 produce three pole sets that are similar in distribution but not identical: small inevitable variations in peak-picking, banded clustering, and noise floor mean that the same physical mode is generally identified at slightly different frequencies and dampings across positions. Direct interpolation of residues between independent pole sets is therefore ill-defined—two residues to be interpolated must be attached to the same pole.

The remedy is the construction of a *common modal basis*, performed by the script `commonModalBasis.m` as follows. The pole sets $\{(f_k^{(i)}, \sigma_k^{(i)})\}$ from the three positions are concatenated and clustered along the frequency axis using a frequency-dependent tolerance: the clustering radius grows linearly from $\Delta f = 5 \times 10^{-3}$ Hz at 20 Hz to $\Delta f = 5$ Hz at 12 kHz, with hard clamping in $[0.05, 5]$ Hz. Within each cluster, candidate poles are merged into a single representative pole $(\hat{f}_k, \hat{\sigma}_k)$ obtained as a residue-magnitude-weighted mean of the cluster members. The procedure yields a unified pole set of size $\hat{M} = 6337$ over the band $[70, 12000]$ Hz.

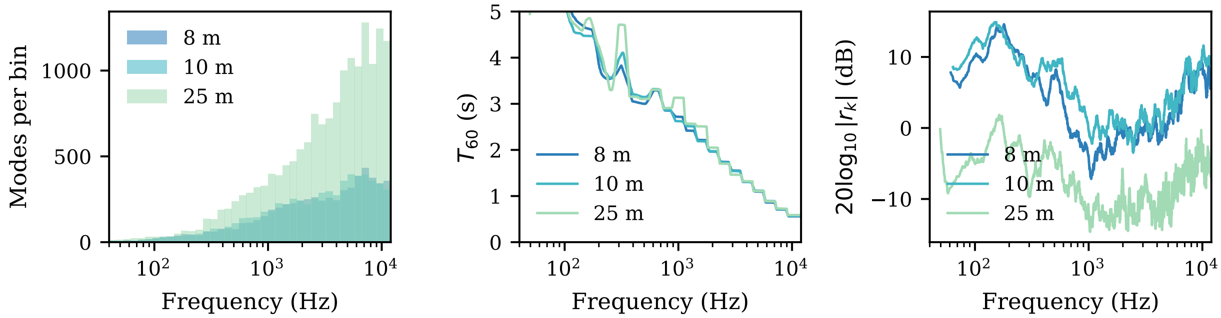


Figure 3: ***Per-position modal statistics extracted independently from the three Bunkervik IRs. Left: distribution of identified modal frequencies*** f_k***. Centre: per-mode reverberation time*** $T_{60} = 6.908/\sigma_k$ ***as a function of frequency. Right: smoothed magnitude*** $20\log_{10}|r_k|$ ***versus frequency. The decay profile is essentially position-invariant; the residue magnitudes drop systematically with source–receiver distance, as expected from a*** $1/r$***-type attenuation.***

Once the common basis $\{(\hat{f}_k, \hat{\sigma}_k)\}$ is fixed, per-position residues and a 256-tap FIR correction filter are re-fitted by a joint, alternating least-squares procedure. Writing $H^{(i)}_{\mathrm{FIR}}(\omega)$ for the frequency response of the position-i FIR and stacking the $\hat{M}$ basis functions into a matrix $R(\omega)$, the forward model

$$H^{(i)}_{\mathrm{tgt}}(\omega) \approx H^{(i)}_{\mathrm{FIR}}(\omega) \cdot R(\omega) \begin{bmatrix} \mathbf{a}^{(i)} \\ \mathbf{b}^{(i)} \end{bmatrix} \tag{5}$$

is alternately solved for $\mathbf{h}^{(i)}_{\mathrm{FIR}}$ (time-domain ridge least squares with identity prior to discourage comb-filtering, and an explicit constraint $h_{\mathrm{FIR}}(0) = 1$) and for the residue vector $[\mathbf{a}^{(i)}; \mathbf{b}^{(i)}]$ (banded frequency-domain least squares with the FIR scaling absorbed into the design matrix). Damped updates with relaxation factors $\alpha_{\mathrm{FIR}} = 0.5$ and $\alpha_{\mathrm{res}} = 0.4$ ensure stable convergence over $K = 6$ outer iterations, and an accept-if-better safeguard prevents any iteration from increasing the time-domain nRMSE.

The output of this stage is, for each measurement position i, a tuple $(\mathbf{a}^{(i)}, \mathbf{b}^{(i)}, \mathbf{h}^{(i)}_{\mathrm{FIR}})$ of residues and FIR taps attached to a common, position-invariant pole set $\{(\hat{f}_k, \hat{\sigma}_k)\}_{k=1}^{\hat{M}}$. This is the structure that enables interpolation, as described in Section 5.

The quality of the joint fit is summarised in Table 1 and illustrated in Fig. 4. Final time-domain normalised RMSE values lie below 3.5% at all three positions, with the FIR stage reducing the IIR-only error by a factor of 2–3. Mode counts and mid-band densities ramp distinctly from the near-source position x_1 to the far position x_3, with a near-tripling of $\rho_{500\text{–}1\mathrm{k}}$ at the far end—consistent with a more diffuse, modal-overlap-dominated field at x_3, where directional source effects have averaged out, and with the systematic drop in residue magnitudes visible in the right panel of Fig. 3.

Table 1: ***Time-domain normalised RMSE for the three measurement positions, for the IIR-only fit (residues against the common basis with*** $H_{\mathrm{FIR}} = \delta$***) and after the alternating IIR/FIR joint optimisation. Mode counts and mid-band modal densities are also reported.***

Position	Mode count M	$\rho_{500\text{–}1\mathrm{k}}$ (modes/Hz)	nRMSE (IIR only)	nRMSE (joint)
$x_1 = 8$ m	6 031	1.12	3.96×10^{-2}	1.60×10^{-2}
$x_2 = 10$ m	6 324	1.37	8.68×10^{-2}	3.47×10^{-2}
$x_3 = 25$ m	16 216	2.80	3.95×10^{-2}	1.30×10^{-2}

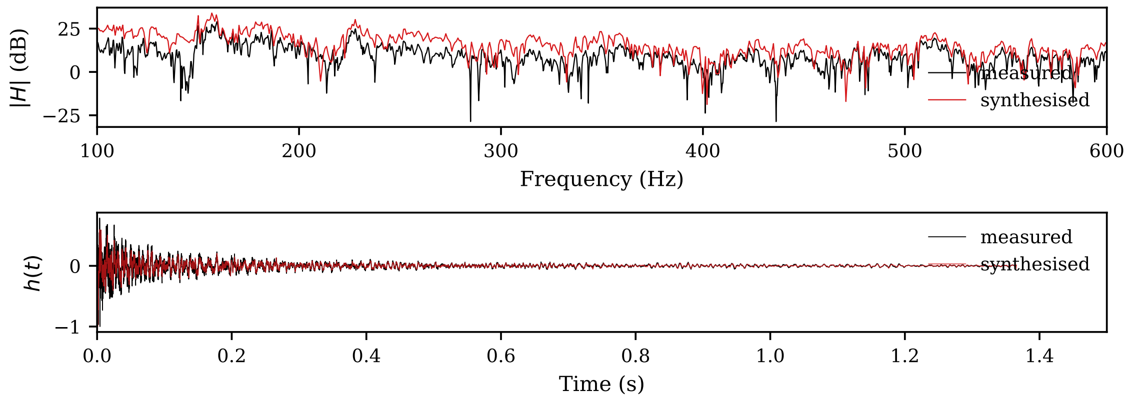


Figure 4: Frequency- and time-domain comparison of the measured (black) and synthesised (red) impulse response at the $x_1 = 8$ ***m position, after the common-basis joint IIR/FIR fit. Top: magnitude responses on*** $[100, 600]$ ***Hz. Bottom: time-domain waveforms over the first*** 1.5 ***s. The synthesised IR captures both the discrete narrow modes below*** ~ 300 ***Hz and the smoother broadband envelope, and tracks the measured decay through several reverberation times.***

5. PLUGIN ARCHITECTURE AND INTERFACE

A. TIME-VARYING MODAL ARCHITECTURE

For an arbitrary listener position $x \in [x_1, x_3]$, position-dependent residues and FIR taps are obtained by piecewise linear interpolation of the three measured triplets:

$$\mathbf{a}(x) = \mathrm{lerp}(\mathbf{a}^{(i)}, \mathbf{a}^{(i+1)}; x), \qquad \mathbf{b}(x) = \mathrm{lerp}(\mathbf{b}^{(i)}, \mathbf{b}^{(i+1)}; x), \tag{6}$$

with i chosen so that $x \in [x_i, x_{i+1}]$, and analogously for the FIR. *The pole set is unchanged*. The resulting position-dependent acoustic transfer function is $H(z; x) = H_{\mathrm{FIR}}(z; x) \cdot \mathrm{Delay}(z; x) \cdot \mathrm{ModalBank}(z; x)$, where the propagation delay $\tau(x) = |x - x_S|/c$, with $c \approx 343$ m/s, is implemented as a fractional-delay circular buffer with linear interpolation. Each modal recursion takes the form

$$q_k[n] = G_{1,k}\, q_k[n-1] + G_{2,k}\, q_k[n-2] + P_{0,k}(x)\,(f[n] + f[n-1]) + P_{\mathrm{d},k}(x)\,(f[n] - f[n-1]), \tag{7}$$

where $G_{1,k}, G_{2,k}$ depend only on the (position-invariant) poles and are precomputed; position dependence enters only through $P_{0,k}(x), P_{\mathrm{d},k}(x)$, derived from the interpolated residues following.[12] Block-rate parameter updates are smoothed by one-pole low-passes ($f_c = 1$ Hz) to suppress zipper artefacts; modal states, FIR delay line, and circular buffer are preserved across JUCE `processBlock()` boundaries. The size of the common basis, $\hat{M} = 6337$ modes, is well within the regime that the time-stepping scheme of[12] sustains in real time on commodity hardware: preliminary benchmarks indicate steady-state CPU usage around 25% of a core for a single instance at $f_s = 48$ kHz with 256-sample blocks and modulation active on an Apple M2.

B. PLUGIN INTERFACE

The architecture above is exposed to the user through the single-pane interface of the companion JUCE plugin, *Bunkervik Spatial Reverb*, reproduced in Fig. 5 together with a section-by-section walkthrough of its five DSP-affecting blocks.

Aura controls. Two large knobs and a slider control a high-frequency *modal warping* of the position-invariant pole frequencies. CUTOFF (default 5 kHz) and MAX FREQ (default 18 kHz) define a piecewise-linear map $\hat{f}_k \mapsto \hat{f}'_k$ leaving modes below the cutoff untouched and stretching $[f_c, f_s/2]$ onto $[f_c, f_{\max}]$; BOOST scales the slope. Tonal control in the brilliance band is thereby obtained without altering modal density or the spatial residue structure. An aura change triggers a one-shot recomputation of $G_{1,k}, G_{2,k}$ on a worker thread, with audio-thread crossfading.

Spatial controls. The SRC$\to$30 M strip is a position widget mapping $x \in [0, 30]$ m onto the modal-bank input of (6), with the source S at the left edge. MIC POS exposes the same parameter as a knob in metres for DAW automation. MIC ROT biases the residue interpolation between the broadside-averaged stack and a partial reconstruction restricted to the FRONT/BACK orientations of Section 2. DAMPING is a global multiplier on σ_k, contracting or extending all per-mode decays uniformly.

Mix controls. GAIN is a wet-bus trim, BLEND a dry/wet mix with optional equal-power crossfade. The default places BLEND at unity for use on a send.

Modulation. Two LFOs are routed through a $2 \times N$ depth matrix, each row of which selects a TARGET from a drop-down (MIC POSITION, AURA BOOST, MIC ROT, DAMPING, GAIN, BLEND, NONE) and reports two depths in $[-99, +99]$. The two LFO units carry a rate knob (in Hz, or beat-synchronised divisions $1/4, 1/8, \ldots$ via the toggle pill) and a waveform selector (sine, triangle, square). The factory routing assigns LFO $1 \to$ MIC POSITION at 0.04 Hz with depth 50% (a ~ 25 s cyclic listener walk) and LFO $2 \to$ AURA BOOST at 0.05 Hz with depth 25% (slow timbral breathing).

Output. Stereo L/R meters monitor the wet output post-mix; the BYPASS button switches the plugin to a sample-accurate dry path for A/B comparison.

Figure 5: The **Bunkervik Spatial Reverb** ***plugin interface, with a section-by-section walkthrough. The single-pane layout exposes the architecture of Section 5.A through five visual blocks. The information jewel at the top renders an arc-meter visualisation of the*** **CUTOFF/MAX FREQ** ***pair. The modulation block is what makes the spatial-modal architecture musically alive: a slow modulation of*** **MIC POSITION** ***drives the time-varying interpolation*** (6) ***continuously and audibly translates the listener up and down the tunnel, while the precomputed pole structure ensures that modal frequencies and decays remain perfectly stable; CPU cost is independent of modulation depth and rate. By contrast, a convolution-based equivalent would require crossfading between many static IRs.***

6. CONCLUDING REMARKS

A spatial extension of the modal-reverberation framework has been described and demonstrated through a real measurement and plugin development case study. The central idea is the construction of a *common modal basis*—a single set of position-invariant poles—against which per-position residues and FIR correction filters are fitted. Because the poles encode the geometry and the boundary losses of the room, while only the residues depend on source–receiver positioning through the mode-shape evaluations, this decomposition makes residues directly interpolable across measurement positions. The pole set and the modal recursion coefficients $G_{1,k}, G_{2,k}$ derived from it can therefore be precomputed once, while a moving listener is rendered at run time simply by linearly interpolating $4\hat{M}$ scalars (residues and FIR taps) and passing them to a fixed bank of second-order resonators. This has minimal additional computational cost compared to a static modal reverberator as the update is not required at every time-step, in testing every 8th step produced high quality results with no artefacts.

The case study—the Bunkervik tunnel in Brescia, an essentially one-dimensional vaulted gallery—was deliberately chosen to expose the mechanism in its simplest form: a single longitudinal coordinate parameterising the listener and three measurement positions sufficing to span the perceptually relevant variation. Time-domain nRMSE figures below 3.5% at all three positions, and a working JUCE implementation with stable, audibly artefact-free listener-walk modulation, support the practical viability of the approach. Future work will address the extension of the common-basis identification to two- and three-dimensional source–receiver grids, the joint identification of source-position dependence (currently fixed), and a formal perceptual evaluation against alternative spatial-reverberation methods such as ambisonics, HRTF-based renderers, and parametric FDN approaches.[20,21]

ACKNOWLEDGEMENTS

Funded by the European Research Council (ERC) under EU Horizon 2020, Grant No. 950084 NEMUS. The authors thank the Comune di Brescia and Associazione Culturale Palazzo Monti for access to the Bunkervik venue.

REFERENCES

1 M. R. Schroeder, "Natural sounding artificial reverberation," *J. Audio Eng. Soc.*, vol. 10, no. 3, pp. 219–223, 1962.

2 J.-M. Jot and A. Chaigne, "Digital delay networks for designing artificial reverberators," in *Proc. 90th Conv. Audio Eng. Soc.*, Paris, 1991.

3 V. Välimäki, 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.

4 G. Dal Santo, K. Prawda, S. J. Schlecht, and V. Välimäki, "Differentiable feedback delay network for colorless reverberation," in *Proc. DAFx-23*, Copenhagen, 2023.

5 G. Dal Santo, B. Alary, K. Prawda, S. J. Schlecht, and V. Välimäki, "RIR2FDN: An improved room impulse response analysis and synthesis," in *Proc. DAFx-24*, Guildford, 2024.

6 A. I. Mezza, R. Giampiccolo, E. De Sena, and A. Bernardini, "Data-driven room acoustic modeling via differentiable feedback delay networks with learnable delay lines," *EURASIP J. Audio Speech Music Process.*, vol. 2024, art. 51, 2024.

7 B. Hamilton and S. Bilbao, "FDTD methods for 3-D room acoustics simulation with high-order accuracy in space and time," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 25, no. 11, pp. 2112–2124, 2017.

8 S. Bilbao, "Modeling of complex geometries and boundary conditions in finite difference / finite volume time domain room acoustics simulation," *IEEE Trans. Audio Speech Lang. Process.*, vol. 21, no. 7, pp. 1524–1533, 2013.

9 J. S. Abel, S. Coffin, and K. Spratt, "A modal architecture for artificial reverberation with application to room acoustics modeling," in *Proc. 137th AES Conv.*, 2014.

10 S. J. Schlecht and E. A. P. Habets, "Modal decomposition of feedback delay networks," *IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5340–5351, 2019.

11 H. Kuttruff, *Room Acoustics*, 6th ed., CRC Press, 2017.

12 M. Ducceschi, C. J. Webb, and G. Fratoni, "Efficient synthesis of large room impulse responses in the modal domain," in *Proc. I3DA 2025*, Bologna, 2025.

[13] M. Ducceschi, L. Gabrielli, R. Simionato, and R. Russo "A corpus-driven parametric modal reverberator," in *Proc. DAFx-26*, Cambridge, MA, 2026.
[14] P. Guillaume, P. Verboven, S. Vanlanduit, H. Van der Auweraer, and B. Peeters, "A poly-reference implementation of the least-squares complex frequency-domain estimator," in *Proc. IMAC*, vol. 21, 2003.
[15] B. Peeters, H. Van der Auweraer, P. Guillaume, and J. Leuridan, "The PolyMAX frequency-domain method: A new standard for modal parameter estimation?" *Shock Vib.*, vol. 11, no. 3–4, pp. 395–409, 2004.
[16] M. van Walstijn, V. Chatziioannou, and A. Bhanuprakash, "Implicit and explicit schemes for energy-stable simulation of string vibrations with collisions: refinement, analysis, and comparison," *J. Sound Vib.*, vol. 569, 117968, 2024.
[17] L. Savioja and N. Xiang, "Simulation-based auralization of room acoustics," *Acoustics Today*, vol. 16, no. 4, pp. 49–57, 2020.
[18] M. Vörländer, *Auralization*, Springer, 2020.
[19] A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine technique," in *Proc. 108th AES Conv.*, Paris, 2000.
[20] V. Välimäki, K. Prawda, and S. J. Schlecht, "Two-stage attenuation filter for artificial reverberation," *IEEE Signal Process. Lett.*, vol. 31, pp. 391–395, 2024.
[21] S. J. Schlecht, J. Fagerström, and V. Välimäki, "Decorrelation in feedback delay networks," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 31, pp. 3478–3487, 2023.
[22] Comune di Brescia, "Bunkervik – Il rifugio delle idee," `https://www.comune.brescia.it/news/bunkervik-0`, last accessed April 2026.