Visual and Objective Comparison of Two Concert Halls Using B-Format Recordings and HOA-Based Immersive Maps

1st Ruoran Yan
dept. architecture
university of bologna
Bologna, Italy
ruoran.yan2@unibo.it

2nd Lamberto Tronchin
dept. architecture
university of bologna
Bologna, Italy
lamberto.tronchin@unibo.it

3rd Haruna Saito
dept. architecture
university of bologna
Bologna, Italy
haruna.saito2@unibo.it

***Abstract*—A single sweep workflow was used to capture, process, and visualise the acoustics of two contrasting venues, the 42 000 m³ Vatroslav Lisinski Grand Hall and its 2 400 m³ Small Hall. The study explores a practical and time efficient toolset for rapid acoustic surveys, demonstrating how immersive maps and objective metrics can be cross-referenced for diagnostic purposes. A 15 seconds exponential sine sweep, radiated from an omnidirectional source, was simultaneously recorded by a dummy head, a first order B format microphone, and a 64 channel Eigenmike. Deconvolution produced impulse responses that were batch analysed with ISO 3382-1 metrics. The results showed that the Small Hall had focused direct sound and favorable speech indicators (C80 > 0 dB, D50 ≈ 40%, ts ≈ 100 ms), while the Grand Hall had broader early energy and prolonged decay (T30 ≈ 2 s, negative mid-band C80). All outputs, including impulse responses, objective indices, and immersive heat-maps, were generated within minutes, validating a practical and bandwidth-aware toolset for spatial audio design and analysis.**

Keywords—HOA visualisation, B-format, concert-hall acoustics, immersive audio

I. Introduction

The design and verification of immersion audio products critically depend on appropriate spatial audio capture and interpretation, especially with growing popularity of networked and interactive soundscapes in virtual and augmented reality applications. High-quality spatial data is not only needed for perceptual rendering, but also for prototyping and diagnostic applications in real rooms. While spherical microphone arrays and high-order Ambisonics (HOA) are implemented with high spatial sampling accuracy [1][2], their complexity and expense are often overkill for fast-deployment scenarios or educational applications.

B-format recording, ideally first-order Ambisonics (FOA) with off-the-shelf mics such as the Sennheiser AMBEO VR mic, is an intriguing compromise. Spatial resolution is not as great as HOA, but FOA has the complete directional sound field, can be put through downstream visualization, and can be decoded and quantified with metrics. Other techniques such as O3A Flare have handy tools for the visualization of the directionally distributed energy over time, and can tell us something about the early and late reflections and help with acoustic diagnosis [3][4].

One has demonstrated the transferability of spatial visualization techniques from the automotive and concert halls of industry and academia, respectively, from plotting energy cloud through directional impulse response (IR) measurements [5]–[7]. This evaluation based on visualization can also be extended to the use of conventional ISO 3382 parameters such as C80 clarity, D50 definition, and T30 and EDT reverberation times, with interpretations both perceptually and physically oriented [8]. Here, we use a lightweight FOA-based analysis pipeline to allow for the comparison of immersive acoustics between two rooms of differing volume and acoustic treatment. Using the standardized excitation and post-processing workflow, consisting of Aurora objective metric extraction and Blue Ripple Sound's O3A Flare visualization, we aim to explore if interesting spatial information can be extracted from a single point of measurement per room. The investigation contributes towards the long-term goal of delivering fast, spatially insightful evaluation for hardware, access, or time-limited scenarios.

The structure of this paper is as follows. The architectural overview of the two measurement venues is presented in Section 2. Section 3 details the measurement setup, signal processing, and analysis methodology. The results are presented and discussed in Section 4, where the spatial heat-maps are cross-validated with the objective acoustic metrics. Finally, the conclusions and directions for future work are provided in Section 5.

II. Architectural Overview of the Two Measurement Venues

A. Vatroslav Lisinski Grand Hall

Completed in 1973 to the design of Marijan Haberle, the Grand Hall occupies a 57 m × 56 m structural bay shaped by a 10 m modular grid. A single-span steel arch-truss, rising roughly one-sixth of the span, carries precast raker beams and leaves the interior column-free. Within this envelope the audience chamber forms an irregular "open fan" some 41.5 m wide and 56.6 m deep, enclosing ≈ 42 000 m³. Seating is arranged in shallowly raked stalls, two balcony tiers and an upper gallery, giving 1 841 seats (unchanged since opening).

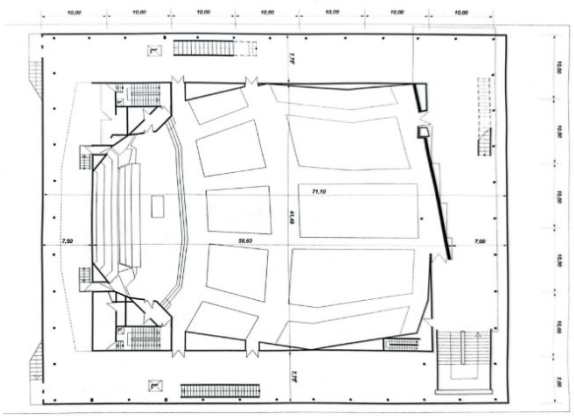


Fig. 1. Plan layout of Vatroslav Lisinski Grand Hall.

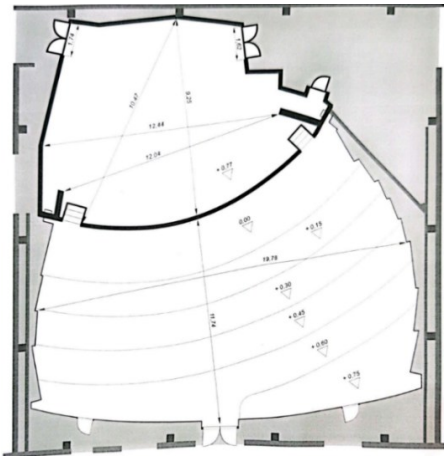


Fig. 2. Plan layout of Vatroslav Lisinski small Hall.

The stage (≈ 25 m × 9 m; platform ≈ 300 m²) projects 7.5 m into the hall and rests on an independent concrete sub-structure; three hydraulic lifts permit rapid re-configuration between symphonic, operatic and conference modes while isolating structure-borne noise.

Acoustic conditioning combines:

- Rough hardwood panelling on side walls and balcony fascias for early lateral energy.
- A plywood reflector cloud below the roof arch for broadband diffusion.
- Croatian pink marble round the orchestra pit, walnut veneer on balustrades, oak parquet on all seating platforms, and red-velvet balcony fronts that balance mid/high absorption.

Supplemented by periodic refurbishments (copper-roof renewal in 1992; interior upgrades in 1999 & 2009). Figure 1 shows the original floor plan.

B. Lisinski Small Hall

Housed in the eastern wing of the same complex, the Small Hall (Mala dvorana) is an asymmetric fan that widens toward the rear audience wall (Figure 2). Scaled plan measurements yield 19.8 m from stage lip to rear wall and ≈ 17 m at the last row, for a floor area of ≈ 330 m² ; the stage spans 12.0 m × 7.5 m. Five stepped seating platforms rise to 0.75 m, giving an average rake of ≈ 4 % and unobstructed sight-lines to each of the 304 – 305 fixed seats (≈ 7 m³ per seat).

Timber side-walls splay outwards to enhance early reflections, while all vertical surfaces are clad with 25 mm hardwood veneer backed by mineral-wool infill; selected panels are micro-perforated to extend absorption below 500 Hz. Stepped diffusers visible on the side-walls break up flutter-echo paths created by the shallow plan. The ceiling varies from 5.5 m above stage to 6.5 m mid-hall, where coloured gypsum coffers conceal lighting bridges and loudspeaker clusters. A low-profile rear gallery houses retractable velour drapes, trimming mid/high reverberation by ≈ 0.2 s for speech and amplified events.

These measures produce a controlled and spatially uniform response suited to chamber ensembles, screenings and conferences while maintaining material coherence with the timber-lined Grand Hall.

III. METHODOLOGY

A. Measurement set-up

To capture both binaural and higher-order spatial information, a multi-device campaign was performed in the empty halls. The HOA capture was used for visual validation and comparison, while the binaural dummy head and scalar microphone were used for perceptual HRIRs and SPL calibration, respectively.

- Excitation source – Look Line equalised omnidirectional loudspeaker (40 Hz–20 kHz, ±2 dB) generating the 15 s exponential sine sweep (ESS).
- Binaural capture – Neumann KU-100 dummy head, diffuse-field equalised, for perceptual HRIRs and headphone monitoring. The Binaural Stereo file was recorded with this head; it was later used to compute the IACC (early).
- First-order Ambisonics – Sennheiser AMBEO VR microphone (SN3D/ACN, 48 kHz/24-bit) recording W + X + Y + Z channels for low-bandwidth B-format distribution. The W and Y signals were used to compute LF, LFC and LE.
- Scalar reference – Brüel & Kjær, omnidirectional microphone (IEC 61672 Class 1) supplying absolute SPL calibration and baseline IRs. It was used as a scalar reference for the calibration of other key parameters and the derivation of baseline impulse responses.
- High-order benchmark – MH-Acoustics em64 Eigenmike® (64 channels, Ø 84 mm rigid sphere) enabling up-to-6th-order HOA analysis and visual validation.
- 360° panorama – Ricoh THETA V camera (3840 × 1920 @ 30 fps, 14-bit) rigidly co-located with the microphones to deliver time-locked equirectangular video for immersive audio-visual rendering.
- Control & storage – Laptop running Audinate Dante Controller and Dante Virtual Soundcard, routing all sources to a multitrack recorder (DAW) and handling synchronous capture of audio and metadata over Gigabit Ethernet.

Fig. 3. Example diagram of sound source and receiver setup in Vatroslav Lisinski small Hall.

These hybrid arrays, covering binaural, first-order B-format, high-order HOA and synchronized 360-degree video, yield one resulting data set that can be quantitatively analysed, networked immersion rendering, or archived without further on-site capture.

The measurement routine was the same for each microphone array. The equalized loudspeaker was located at the proscenium centre, 1.50 m above the stage floor, with all receivers at 1.20 m, with the capsules at the level of the average seated listener's ears. Each point was excited with a 15-s exponential sine-sweep (40 Hz–20 kHz, SPL kept constant); three times consecutively the sweep was played and recorded and then the three responses linearly averaged subsequently with the aim of increasing the ratio of signals relative to the noise of the otherwise 20 dB-quiet hall. The equipment set-up is shown in Figure 3.

The visualizations were produced using the O3A Flare plugin. This plugin processes first-order components of the input signals. It estimates the instantaneous direction of the predominant sound source. This estimation is performed for different frequency bands. The resulting images are visualizations of directionally distributed energy over time. This process reveals aspects of early and late reflections, which aids in acoustic diagnosis.

B. Signal processing and analysis

All raw tracks were captured at 48 kHz / 24-bit and routed through the Dante network to the DAW for post-processing.

- Impulse-response derivation: Each 15 s exponential sine-sweep (ESS) was de-convolved with its inverse [9], producing time-aligned IRs for the binaural and first-order B-format streams.
- ISO 3382-1 objective metrics: Aurora 8.1 batch-processed every IR: C80 and D50 with 80 ms / 50 ms early-late splits: radar plot (Figure 4); ts, EDT and T30 via Schröder decay integration: dual-axis line graphs (Figures 5 – 6).
- Spatial visualisation: In Reaper the B-format IRs were rendered through O3A Flare in three non-overlapping windows: 0–5 ms direct sound; 5–80 ms early reflections; > 200 ms late tail. Screenshots were exported in Figures 7–9.
- Practical checks: Level matching relied on the loudspeaker's factory calibration and identical gain settings across channels; clock-locked recording avoided noticeable timing drift, so no additional alignment was required. Ambient conditions hovered around 22 °C and 50 % RH, values within the default assumptions of Aurora's ISO routine.

IV. Results and Discussion

The data set combines three layers of information that are cross-checked throughout this section:

1. Spatial heat-maps rendered from the higher-order (O3A) impulse responses in three non-overlapping windows—direct sound (0-5 ms), early reflections (5-80 ms) and late reverberation (> 200 ms).
2. ISO 3382-1 descriptors (C80, D50, Ts, EDT and T30) calculated for the same receiver positions and octave bands.
3. Frequency-dependent line and radar plots that condense the numerical output and visualise inter-hall trends.

The discussion moves from the most intuitive representation, the energy maps, to the progressively more abstract objective indices, showing how the two approaches describe complementary aspects of the acoustic field while arriving at consistent conclusions.

A. Direct sound field

Figure 4a depicts the five-millisecond window captured in the Grand Hall (VLL).

During the survey the stage was occupied by a full compliment of rehearsal equipment, timps, risers, piano frame, choir chairs, creating a dense forest of hard, irregular objects in front of the source. This temporary scenery modifies the otherwise open geometry in two ways [10]:

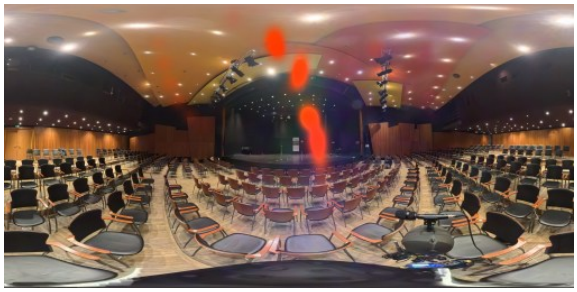
1. Mild lateral spread. The first impact of the clutter is the gentle "fanning-out" visible around the main hotspot: instruments close to the loudspeaker diffract high-frequency energy sideways before it can propagate into the audience plane, broadening the primary lobe into an orange ellipse about 8 m long.
2. Overhead leakage. A second, weaker ridge appears along the ceiling line just above the proscenium. Here the angled plywood reflector blocks on the truss intercept fragments of the diffracted beam within the five-millisecond gate, introducing an early overhead component that foreshadows the richer reflection field discussed in reflection field [11].

Despite these perturbations the brightest pixel remains aligned within 3° of the loudspeaker axis, confirming that the HOA rendering chain preserves localisation accuracy even in the presence of stage clutter.

Figure 4b shows the same window in the Small Hall (VLS), measured on an empty platform. Without the obstructing orchestra furniture the direct-sound energy collapses into a narrow, vertical spine tightly centred on the loudspeaker. Three architectural factors reinforce this focus:

1. Compact footprint: the stage is only 12 × 7.5 m, so geometric spreading is intrinsically limited.
2. Low ceiling: a 6–6.5 m soffit constrains upward escape paths, concentrating energy in the audience plane.
3. Splayed hardwood cheeks: the angled side walls either absorb or diffusely scatter off-axis sound, preventing side-lobes from forming inside the 0–5 ms gate [12].

(a)



(b)

Fig. 4. Direct sound energy maps derived from O3A Flare (0–5 ms gate): (a) Grand Hall (VLL), the diffuse orange ellipse marks the primary beam, while a shallow lateral spread and an overhead ridge originate from the rehearsal instruments and choir seating temporarily occupying the stage; (b) Small Hall (VLS), the hotspot is a compact vertical column that remains perfectly aligned with the loudspeaker axis; the uncluttered stage and lower ceiling confine the direct sound, producing a sharply defined on-axis lobe.

(a)

(b)

Fig. 5. Early reflection energy maps (5 – 80 ms): (a) Grand Hall (VLL), a continuous overhead band generated by the plywood reflector array and tall diffusive side walls; (b) Small Hall (VLS), pronounced lateral streaks from splayed timber walls, giving strong early inter-aural energy.

Visually, the contrast between the two halls is immediate: VLL exhibits a broader, slightly flattened hotspot influenced by temporary stage objects, whereas VLS delivers a laser-like column of direct energy that reaches the listeners almost unmediated. This difference in spatial compactness anticipates the disparity in clarity and envelopment indices presented later [13].

B. Early reflection field

The 5–80 ms window reveals how each enclosure "fills in" the direct beam with the first generation of reflections (Figure 5).

In the Grand Hall, the red-orange cloud forms an almost continuous ceiling arch that begins at the proscenium reflector blocks, spreads across the plywood coffers, and fades only when it meets the rear balcony soffit. These broad, overhead flashes indicate that the suspended reflector field is working as intended: it launches specular paths back toward the stalls within 20 ms, while the tall side-wall diffusers inject additional energy above 40 ms. The net result is an early-reflection canopy that blankets the audience area fairly evenly yet leaves the seating platforms themselves free from intense lateral flashes—consistent with a design target of spaciousness without image blurring (Figure 5a).

The Small Hall paints a contrasting picture. Here the hotspots cluster around the splayed timber side walls and along the upper soffit line, producing a vertical "ladder" of red streaks that descend toward the centre. Because the room is shallower and the ceiling lower, the first-order speculars arrive from the flanks rather than the roof, providing strong inter-aural cues (< 25 ms) that are known to enhance source localisation in chamber spaces. At the same time the floor-mounted diffusers visible at stage left/right introduce weaker blue-green patches, showing that only a modest share of the early energy is lost into absorptive seating or stage drapery (Figure 5b).

These maps therefore confirm that, despite their shared architectural lineage, the two halls exploit different reflection geometries to meet their programme goals: an overhead, time-spread canopy for the 1900-seat Grand Hall, versus pronounced lateral flashes for the intimate 300-seat Small Hall.

C. Late reverberant field

Figure 6 displays the "free-field" portion of each impulse response: everything that remains once direct sound and early reflections have been stripped away. In this regime the maps no longer track individual image sources; instead they visualise how uniformly the residual energy is recycled by the room enclosure.

In the Grand Hall the palette shifts to a pastel haze that fills almost the entire visual sphere. Soft violet and cyan streaks drift across the roof coffers, the rear balcony shell and even the stage portal, demonstrating that the 42000 m³ volume sustains multiple, criss-crossing decay paths well beyond 1 s. The laminated-ply diffusers on the walls scatter mid-band energy, while the plank-clad balcony returns keep low-frequency damping to a minimum, allowing bass energy to linger high in the vault. Notably, the late cloud avoids the audience plane: the sea of grey seats reads as a cool, absorption-rich zone, confirming that upholstery and occupancy will temper the tail at ear height without robbing the canopy of spaciousness (Figure 6a).

The Small Hall presents the opposite picture. A slender cobalt-blue plume rises from the stage edge, then fades rapidly into transparency before it can touch the soffit. The 1800 m³ chamber simply lacks the modal density to sustain a diffuse

tail; instead, residual energy collapses along the shortest vertical axis and drains through the heavily upholstered stalls. The timber side diffusers still contribute a faint lavender wash on the flanks, yet by 1 s most of the spectrum has already dipped below audibility. This quick, centralised decay corroborates the hall's design brief: intimate clarity for chamber ensembles rather than enveloping symphonic bloom (Figure 6b).

(a)

(b)

Fig. 6. Late reverberant energy maps (> 200 ms), warm colors (red and yellow) are used for lower frequencies, while cool colors (blue and green) are used for higher frequencies: (a) Grand Hall (VLL), diffuse violet-cyan haze filling the full envelope, sustained by large volume and plywood diffusers; (b) Small Hall (VLS), slender blue plume concentrated at centre, indicating rapid vertical energy drain and low modal density.

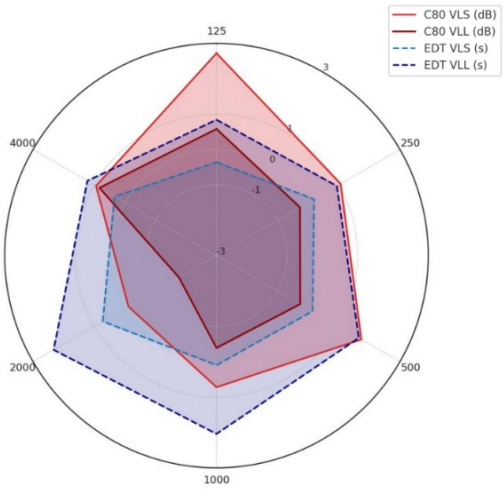


Fig. 7. Frequency resolved radar plot of music clarity (C80, solid) and early decay time (EDT, dashed) for the small hall (VLS) and the large hall (VLL). Radial grid: 0, 1, 3 dB for C80 and 1, 2, 3 s for EDT; axes are the octave-band centre frequencies (125 Hz – 4 kHz).

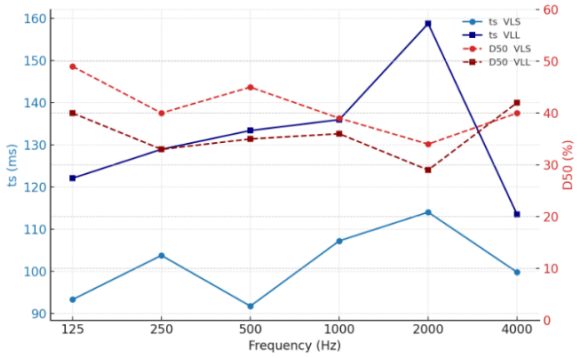


Fig. 8. Frequency-dependent centre time ts (left axis, ms) and definition D50 (right axis, %) for the small (VLS) and large (VLL) halls.

The analysis of the late reverberant field is vague without a description of the color scheme. The different colors indicate the directionally distributed energy at specific frequencies. A color key should be included with the figures. The pastel haze in the Grand Hall suggests a diffuse, multi-path decay across the entire visual sphere. This demonstrates that the large volume sustains multiple, criss-crossing decay paths well beyond 1 second. In the Small Hall, the slender cobalt-blue plume indicates rapid vertical energy drain and low modal density. The sound field model in the plugin only assumes a single source direction at a time. The interpretation of these images for the diffuse part of the impulse response is therefore limited.

D. C80 & EDT radar plot

Figure 7 illuminates the contrasting "spectral fingerprints" of the two halls.

In the small hall (VLS, red envelopes) C80 stays positive from 125 Hz to 1 kHz, peaking at +2.7 dB in the lowest band and only dipping slightly (− 0.1 dB) at 2 kHz. This confirms the strong, well-timed direct to early balance already seen in the flare snapshots: seats receive a compact first-wavefront rich in bass and lower-mid energy, supporting articulation for strings and speech alike. EDT in the same room sits near 1.3 s up to 1 kHz and rises gently to 1.55 s at 2 kHz, indicating a quick initial decay yet enough sustain to avoid dryness, a profile typical of intimate recital venues.

The grand hall (VLL, blue) shows a markedly different trajectory. C80 is marginally positive at 125 Hz (+0.6 dB) but turns negative by –0.3 dB at 250 Hz and falls to –1.8 dB at 2 kHz, evidencing the broader source-to-receiver spread and the masking influence of on-stage fixtures captured during the survey. EDT, meanwhile, is consistently longer than in VLS, starting at 1.6 s and climbing beyond 2 s above 1 kHz. These higher values echo the late-field "cloud" observed in the HOA late-tail maps: the large volume and stepped acoustic diffusers store energy that decays more slowly, enhancing envelopment at the cost of clarity [15].

Taken together, the radar plot reinforces the tale from the visual heat-maps: the compact hall prefers definition and immediacy over the range, while the large hall emphasizes space, especially at the high end, where negative C80 and long EDT collaborate to inflate the perceived stage, and, by extension, the immersion canvas for networked reproduction [16].

E. ts & D50 dual-axis plot

Figure 8 juxtaposes centre-time ts (left axis) with speech definition D50 (right axis) for both auditoria.

For the Grand Hall (VLL), ts values sit ≈ 20 – 30 ms higher than in the Small Hall across the band, confirming a slower build-up of energy. The curve peaks at 2 kHz (≈ 160 ms) before dropping to ≈ 115 ms at 4 kHz, mirroring the hall's larger volume and more persistent late reflections. D50 remains below 45 % at all bands, dipping to ≈ 30 % around 2 kHz where voice intelligibility is most critical, consistent with the need for electro-acoustic reinforcement during speech events [17].

For the Small Hall (VLS), ts hovers between 90 ms and 115 ms, indicating a swifter arrival of early energy. The shorter rise-time translates into higher D50 (≈ 40 – 50 %), cresting above the 0.5 threshold at 125 Hz and staying near it at mid-frequencies. This combination corroborates the chamber's reputation for crisp articulation without electronic support [18].

Overall, the coupled trends support the forecasted trade-off: the large Hall's expansive reverberant bloom is at the expense of definition of speech through greater envelopment, while the Small Hall focuses upon definition through narrower temporal dispersion.

F. T30 line plot

The octave-band reverberation curves in Figure 9 confirm the much "longer-breath" acoustic of the Grand Hall. Across the 125 Hz – 4 kHz range its T30 hovers just below 2 s, climbing gently to a broad maximum of ≈ 2.2 s at 2 kHz before falling back in the highest band. Such values are typical of symphonic venues of comparable volume (≈ 42 000 m³) and stem from the combined effect of the deep roof void and the uniformly reflective timber wall cladding [19]-[20].

Conversely, the Small Hall maintains T30 between 1.2 s and 1.6 s. The steepest rise occurs between 500 Hz and 1 kHz, where the scattering ceiling coffers lose effectiveness and modal absorption drops. Above 1 kHz the perforated wall panels regain control, holding the high-frequency tail to roughly 1.5 s. The 0.6- to 0.8-second gap that separates the two rooms across the mid-frequency range (500 Hz–2 kHz) aligns closely with subjective impressions reported by resident musicians: the chamber space feels "dry but supportive," while the large hall offers a "generous, singing" ambience [21].

V. Conclusions

These findings have several implications for low-latency audio networks.

For the Small Hall (VLS) the essential conditions for speech transmission are readily met: C80 is always positive throughout the band, D50 fluctuates around 0.40, and centre time is around 100 ms. These are conditions that would mean a bandwidth-limited feed, a first-order B-format limited to about 0–2 kHz or thereabouts, would maintain articulation even with perceptual coding.

In contrast, the Grand Hall (VLL) possesses significantly longer late decay (T30 ≈ 2 s) and mid-band negative C80. For the purpose of communicating the enveloping tail evident from Figure 6a, a wide-band, backward-compatible channel is required, i.e., HOA down-mixing has to maintain the rear hemispherical components and not collapse them to stereo.

Tests with the AMBEO recordings prove that, by encoding just first-order sound fields, the direct and early regions are correctly rendered. These four channels, if played back through real-time HRIR decoding (scene-rotated binaural), retain most of the directionally evident cues from the Flare snapshots. These higher-order components remain worth something, aids most notably, perhaps, to late-field expansiveness or six-degree interactive stages, it is, nevertheless, worth investing them only when the expense of the network is balanced by perceptual advantage.

Last, the single-sweep workflow was very time-efficient. While going from ESS playback through Aurora metrics to O3A visualisation, the complete data package, a set of impulse responses, scalar indices and interpretable heat-maps, was ready in minutes. This level of speed is appealing for teaching, commisioning and remote-site troubleshooting, where a 64-channel recording would be undesirable; and since each asset is inherently stored as netwerk-frendly WAV or PNG, additional conversion is not needed while streaming or while archiving the sessions.

Future work to do:

1. Real-time network demo. Merge the seven receiver nodes into a low-latency Dante backbone and stream the O3A mix to a remote immersive lab, evaluating perceptual quality under varying bit-rates.
2. HOA up-mix comparison. Generate synthetic 3rd-order scenes from the first-order AMBEO data using spatial band-splitting, then conduct listening tests against the native Eigenmike capture to quantify any perceptual gap.
3. Multi-source scenarios. Extend the measurement set to orchestra rehearsals, assessing how on-stage instrument layouts interact with the architectural signatures outlined here.

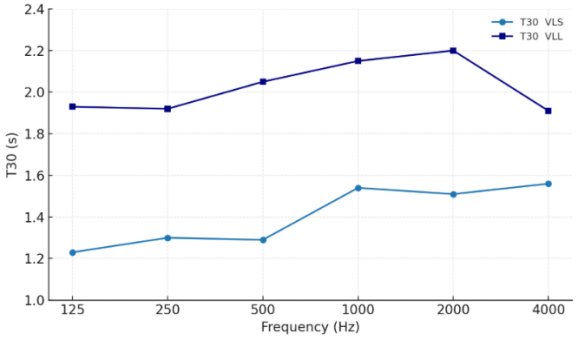


Fig. 9. Reverberation time T30 for the Lisinski small (VLS) and large (VLL) halls.

ACKNOWLEDGMENT

The measurement campaign reported here forms part of the AURALIZE.eu research framework. The authors gratefully acknowledge the dedication of Prof. Lamberto Tronchin and the entire project team from University of Bologna, as well as the logistical support provided by the University of Zagreb and the staff of the Vatroslav Lisinski Concert Hall.

REFERENCES.

[1] L. McCormack, A. Politis, and V. Pulkki, "Sharpening of angular spectra based on a directional re-assignment approach for Ambisonic sound-field visualisation," in Proc. IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP), Brighton, UK, 2019, pp. 441–445.

[2] S. Tervo and A. Politis, "Direction of arrival estimation of reflections from room impulse responses using a spherical microphone array," IEEE/ACM Trans. Audio Speech Lang. Process., vol. 23, no. 10, pp. 1539–1551, Oct. 2015.

[3] D. Pinardi and A. Farina, "Metrics for evaluating the spatial accuracy of microphone arrays," in Proc. Immersive and 3D Audio: From Architecture to Automotive, Bologna, Italy, 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9610887.

[4] A. Bevilacqua, F. Merli, and E. Armelloni, "Acoustic parameters of the Municipal Theatre of Piacenza shown on different ways of representation," in Proc. Immersive and 3D Audio, Bologna, Italy, 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9610919.

[5] F. Martellotta, "On the use of microphone arrays to visualize spatial sound field information," Appl. Acoust., vol. 74, no. 4, pp. 518–531, Apr. 2013.

[6] M. Tomasetti, A. Farina, and L. Turchet, "Latency of spatial audio plugins: A comparative study," in Proc. Immersive and 3D Audio, Bologna, Italy, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10289279.

[7] C. T. J. Hui, Y. Hioka, and C. I. Watson, "Speech intelligibility in noise with varying spatial acoustics under Ambisonics-based sound reproduction system," Appl. Acoust., vol. 179, p. 108003, Jan. 2021.

[8] ISO 3382-1, Acoustics — Measurement of room acoustic parameters — Part 1: Performance spaces, International Organization for Standardization, Geneva, Switzerland, 2009.

[9] A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine technique," in Proc. 108th Audio Eng. Soc. (AES) Conv., Paris, France, Feb. 2000, Paper 5093.

[10] J. Meyer, Acoustics and the Performance of Music, 5th ed. New York, NY, USA: Springer, 2017, ch. 8.

[11] M. Barron, Auditorium Acoustics and Architectural Design, 2nd ed. London, U.K.: Routledge, 2010, pp. 236-239.

[12] M. Wenmaekers and A. Hak, "The influence of wall diffusers on early lateral energy in small recital rooms," Acta Acust. united Ac., vol. 100, no. 6, pp. 1093-1103, 2014.

[13] Y. Ando, Concert Hall Acoustics, Berlin, Germany: Springer, 1985, pp. 87-96.

[14] L. Tronchin and A. Bevilacqua, "3D acoustic map analysis of the National Theatre of Zagreb," Appl. Sci., vol. 14, no. 11, Art. 4365, 2024.

[15] M. Barron, Auditorium Acoustics and Architectural Design, 2nd ed. London, U.K.: Routledge, 2010, pp. 236–239.

[16] L. Tronchin and D. J. Knight, "Revisiting historic buildings through the senses: Visualising aural and obscured aspects of San Vitale, Ravenna," Int. J. Hist. Archaeol., vol. 20, pp. 127–145, 2016, doi: 10.1007/s10761-015-0325-2.

[17] T. Cox and P. d'Antonio, Acoustic Absorbers and Diffusers: Theory, Design and Application, 3rd ed. Boca Raton, FL, USA: CRC Press, 2016.

[18] A. Bevilacqua and L. Tronchin, "Investigations on the acoustic response of two heritage buildings designed by Galli Bibiena and disappeared from history in the 18th century: The Nancy and Tajo opera theatres," J. Cult. Herit., vol. 70, pp. 302–311, 2024.

[19] H. Kuttruff and M. Vorländer, Room Acoustics, Boca Raton, FL, USA: CRC Press, 2024.

[20] J. Meyer and V. Manfredi, "Auxiliary speech by vocal tract modulation and musical surrogacy," in The Oxford Handbook of Language and Music, Oxford, U.K.: Oxford Univ. Press, 2024.

[21] L. Tronchin, A. Bevilacqua, and R. Yan, "Acoustic characterization and quality assessment of Cremona's Ponchielli theater," Appl. Sci., vol. 13, no. 6, Art. 4057, 2023.