Perception of bass with some musical instruments in concert halls

H. Tahvanainen, J. Pätynen and T. Lokki
Department of Media Technology, Aalto University, P.O. Box 15500, 00076 Aalto, Finland
henna.tahvanainen@aalto.fi

The perception of many orchestral instruments with their fundamental frequencies in the low frequency region may suffer from a lack of the bass in concert halls. In particular, the perception may be affected by the seat-dip attenuation which occurs typically between 80-1000 Hz due to the direct sound propagating across the audience area at near-grazing angles. This paper presents studies on the perception of bass with some musical instruments in four existing concert halls. In a paired comparison listening test, the assessors compared the amount of bass and overall clarity (articulation) in the concert halls with three musical excerpts. The excerpts are obtained by convolving anechoic recordings of cello, double bass, and tuba with room impulse responses of four European concert halls measured with a calibrated loudspeaker orchestra. The results show that perceived amount of bass is clearly stronger in the shoebox-shaped halls with wideband seat-dip attenuation than in the vineyard halls with narrowband seat-dip attenuation. Clarity, on the other hand, depends on the musical context.

1 Introduction

Sufficient bass in concert halls is considered important for spatial impression and acoustic warmth [1]. Bass sound in concert halls has been addressed in many occasions [2, 3, 4, 5], but these studies have received critique of being limited to the 125 Hz octave band and above [6, 7]. Thus, the results do not cover the lowest frequencies of many orchestral instruments, such as cello or double bass, that extend to the 63 Hz and 31.5 Hz octave bands. The 125 Hz octave band can also ignore the seat-dip attenuation, in particular if the attenuation bandwidth is narrow [8].

The seat-dip effect is the excess attenuation of low frequencies in addition to the geometrical spreading. The excess attenuation due to the sound propagating at near-grazing angles across the concert hall stalls [9, 10] and it is the result of a destructive interference between the direct sound and reflections from the seat tops and the floor between the seats [11]. The main excess attenuation lies between 80-300 Hz with a magnitude up to -30 dB and it can span up to 1 kHz. The attenuation reaches its maximum within the first 20 ms after the direct sound. Importantly, the attenuation recovers to varying degree over time depending what kind of non-grazing reflections are provided by the concert hall geometry. Both ceiling reflections [12] and lateral reflections [8] have been proposed to correct the seat-dip attenuation.

With such strong attenuation in the early arriving sound, the perception of bass is considered to suffer [2], especially with instruments whose fundamental frequencies fall in the range of seat-dip attenuation. Bradley [12] has estimated that the attenuation of the double bass due to the seat-dip effect can be 6 dB in addition to the geometrical spreading. Indeed, the double bass and the cello, and their articulation is particular, are often regarded weak in many concert halls. The seat-dip effect was in fact first noticed in the New York Philharmonic Hall because it lacked bass and the instruments sounded dull [9, 10]. Then the weakness of the cellos and the double basses were particularly commented on [13].

There is very little research on the topic of the perception of the seat-dip effect in real concert halls, and how the effect influences the perception of the musical instruments, especially their timbre. The threshold of audibility for the seat-dip attenuation has been obtained in a simulated hall [14]. It has also been suggested that more bass could be perceived when the low frequency content is attenuated in the direct sound due to the seat-dip effect, but retained in the reflected sound [14, 15].

This paper studies how cello, double bass, and tuba playing together are perceived in concert halls in terms of amount of bass and clarity. The seat-dip effect is also connected to the perception of bass. The method is a paired comparison listening test using loudspeaker orchestra measurements in real concert halls convolved with anechoic instrument recordings. Three different musical excerpts with the lowest played notes at 62 Hz were selected. The excerpts turn out to have an influence on the perception of clarity. The amount of the bass, on the other hand, depends very clearly on the concert hall.

2 The studied concert halls

This study includes two shoebox and two vineyard halls: Berlin Konzerthaus (BK), Vienna Musikverein (VM), Berlin Philharmonie (BP), and Helsinki Music Centre (MT). Their properties are listed in Table 1. The reverberation time (*RT*) is averaged over the 63 Hz - 2 kHz octave bands while clarity (C_{80}) and strength (*G*) are both averaged over the 63 - 500 Hz octave bands. The main attenuation frequency of the seat-dip effect (SDE) depends on the height of the seat back rest which corresponds to a quarter of the wavelength of the interfering sound [11, 12]. The width of the attenuation depends on the raking of the floor and the seat type (seats with an open underpass vs. closed seats or seats blocked by stepwise raking floor) [16]. The properties of the seats and the floor are very similar within the chosen shoebox halls and the vineyard halls.

Table 1: The properties of the measured unoccupied concert halls. S stands for shoebox, V for vineyard. Reverberation time (*RT*) is averaged over 63-2000 Hz octave bands and clarity (C_{80}) and strength (*G*) over 63-500 Hz.

Abbr.	Name	N	Hall, Floor, Seat type	RT (s)	C_{80} (dB)	G (dB)	SDE (Hz)
BK	Berlin Konzerthaus	1600	S, flat open	2.3	-3.1	4.4	182
VM	Vienna Musikverein	1700	S, flat open	2.8	-3.8	4.6	177
BP	Berlin Philharmonie	2200	V, raked closed	2.1	-1.8	1.3	96
MT	Helsinki Music Centre	1700	V, raked blocked	2.4	-2.1	2.6	125

Figure 1 shows the plans of concert halls with the positions of the measurement loudspeakers (1-6) and the receiver location R at 19 m from the stage. The loudspeakers correspond to the locations of the cello section (1-3), double bass section (4-6), and tuba (6) in typical symphony orchestra.

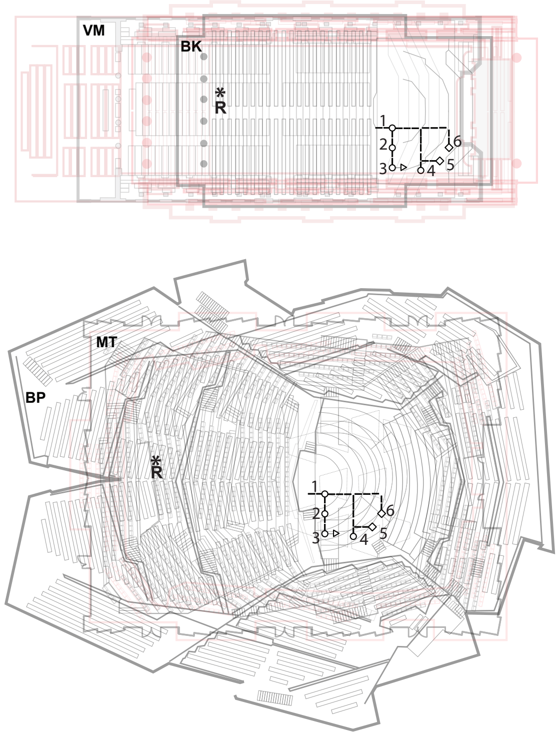

Figure 1: The plans of the studied concert halls and the loudspeaker orchestra measurement setup with the numbers sources indicated.

Figure 2 shows the cumulative time-frequency development of the impulse response in each concert hall averaged over the loudspeaker sources in 10 ms time window increments, starting from 20 ms (see [8] for more details on the method). The 20 ms curve is drawn with a thick line. The second highest curve show the frequency response at 200 ms, and the highest curve the overall frequency response. The fundamental frequencies of the notes played in the selected musical excerpts as well as the second and third partials of the double bass and tuba (dashed line) and the cello (solid line) are plotted on the frequency axis for comparison.

It is noticable that the seat-dip attenuation at the 20 ms curve is more narrowband in the vineyard halls than in the shoebox halls. In addition, the main attenuation frequency is lower in the vineyard halls than in the shoebox halls due to floor raking and less high seat back rests. Furthermore, the difference between the 20 ms and the full response curve is greater in the shoebox halls than in the vineyard halls.

In the vineyard halls the fundamentals of the double bass and the tuba (62 Hz - 124 Hz) lie at the main attenuation frequencies whilst the higher partials are well beyond the seat-dip attenuation. In the shoebox halls, the fundamental frequencies of the double bass and tuba are below the seat-dip attenuation, while all the other partials and the cello fundamentals are within the attenuation band.

3 Listening test

The production of the listening test material consisted of three stages: recording of the instruments in an anechoic room, measuring the room impulses responses of the concert halls with the loudspeaker orchestra, and convolving these

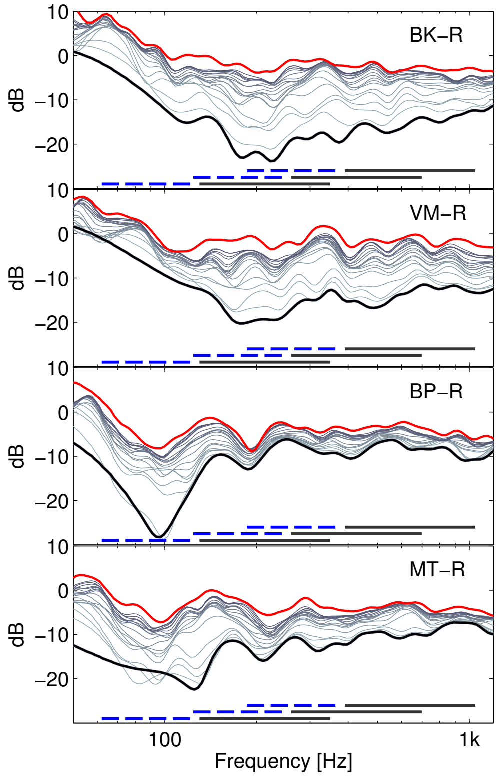

Figure 2: Time-frequency development of the concert halls with one-third octave smoothing applied. The tuba and double bass registers is marked with a dashed line and the cello range with the solid line underneath the curves. The three levels of the lines indicate the fundamental, second, and third partial.

two signals for the final listening via a spatial audio system. Since the low frequencies are at the reproduction limit of the equipment, they were compared in the convolution and the final listening stage.

3.1 Material

The musical instruments were recorded individually in an anechoic room with several musical pieces played by professional musicians [17]. Only one cello and one double bass were recorded and the sections were created artificially by copying these recordings with varied parameters [18]. The room is fully anechoic above 125 Hz, but the decay time of the instruments is typically higher than that of the room. Out of the anechoic recordings, the musical excerpts selected for this study were from the second movement of Bruckner's 8th symphony, and the first movement of Beethoven's 7th symphony. Only the parts for double bass, cello, and tuba were included. The excerpts are listed below and their scores are shown in Figure 3.

- M1: Bruckner, bars 25-32, duration 10.8 s, 7 cellos (sources 1-3) and 5 double basses (sources 4-6)
- M2: Bruckner, 37-40, 5.4 s, 7 cellos (1-3) and 5 double basses (4-6), tuba (6)
- M3: Beethoven, 29-30, 6.5 s, 7 cellos (1-3) and 4 double basses (4-5)

The spatial room impulse responses in the concert halls were measured using a calibrated loudspeaker orchestra

Figure 3: The musical scores of the excerpts used.

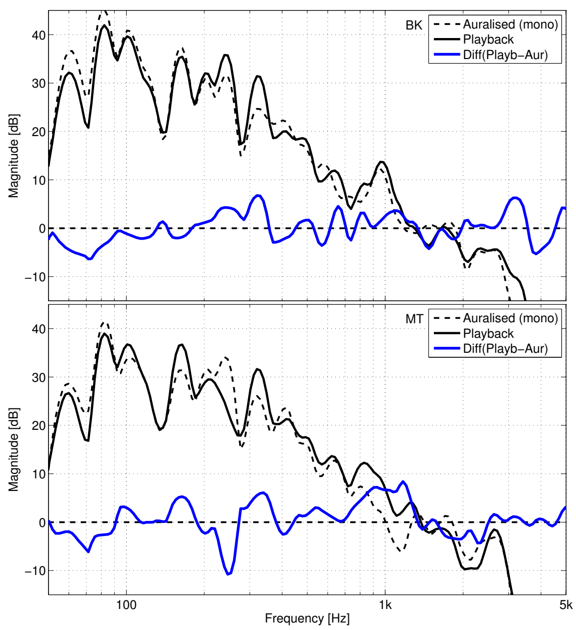

Figure 4: Reproduction of low frequencies with M2 in BK and MT. The dashed black curve to the auralised M2, the thick black curve refers to the playback of M2 in the listening room, and the thick blue curve is the difference between these two.

that simulates a symphony orchestra in terms of locations on the stage and the approximate instrument directivity [19]. The concert halls were measured unoccupied, and the receiver location remained constant in each hall for direct comparison between the halls. In order to better approximate the directivity of the cello at high frequencies, an auxiliary loudspeaker was placed on the floor in the source channel 3.

Finally, the concert halls were auralised by convolving the anechoic recordings with the spatial room impulses responses [20]. The auralisations were played back using a 24-channel spatial sound reproduction system in an acoustically treated room. The system complies with the ITU-R BS.1116-1 recommendation for subjective audio evaluation systems. The system consisted of twenty Genelec 8020B and four Genelec 1029A calibrated loudspeakers. The loudspeakers were positioned on five levels of elevation: at 0°(ear level) [azimuth angles 0°, ±22.5°, ±45°, ±67.5°, ±90°, ±135°, 180°], 30° [azimuth angles 0°,±45°, ±135°], 45° [azimuth angles ±90°], 90° (on top of the listening position) and at 35° [azimuth angles ±40°, ±150°]. The nominal loudspeaker distance was 1.5 m from the listening position. The background noise level (A-weighted, slow) was measured to be 29 dB. The level of playback (A-weighted, slow) in the listening room was 50-65 dB depending on the musical excerpt.

3.2 Reproduction of the lowest frequencies

Because of operating at the lower limit of the audio equipment, the reproduction of the low frequencies in the listening room deserves attention. According to the manufacturer specifications, the loudspeakers used in the reproduction have their cut-off frequency at 66 Hz and below that the roll-off is 24 dB/octave. The lowest fundamental note in the musical excerpts is played by the double bass at B_1=61.7 Hz.

The reproduction is compared between the auralised excerpts and the playback excerpts. The auralised signal is the convolution between the anechoic recordings and the six loudspeaker orchestra sources summed into a mono signal and it is the input to the spatial reproduction system. The output is the playback excerpt measured with a G.R.A.S measuring microphone in the sweet spot of the listening room. An example of the frequency responses of these two cases and their difference is plotted in Figure 4 for BK and MT with musical excerpt M2. These halls were perceived with the most and the least amount of bass. The playback system is able to reproduce the low frequencies until 50 Hz, and the maximum difference occurs at around 72 Hz for the very low frequencies. In general, below 200 Hz, the difference between the auralised excerpt and the listened excerpt is very similar between halls.

3.3 Listening test

The listening test was a paired comparison test composed of two parts. In the first part, the assessors were asked to compare the amount of bass, and in the second part the overall clarity of the excerpts. Altogether, there were 54 comparison pairs per part (6 hall pairs, 1 receiver position, 3 musical excerpts, 3 repeats per pair). The order of both the pairs and the parts was randomised. The assessors could listen to the sample pairs as many times as they wanted. At the end, they were also asked for comments about the test. 7 male and 2 female assessors aged 21-43 participated in the listening test. All assessors had background in acoustics and none of them reported a hearing impairment. The test took on average 35 mins (Part I - 15 min, Part II - 20 min).

4 Results

The boxplots in Figure 5 show the results of the two parts of the listening test grouped by the musical excerpt.

The results of comparing the amount of bass in different halls are very clear; BK was considered to have the most bass in almost all cases and with all musical excerpts by all assessors. The differences in ranking between the shoebox halls and the vineyard halls are significant (Kruskal-Wallis test, $H(1) = 70.35, p < 0.001$), so that shoebox halls in general have a higher perceived level in bass than the vineyard halls. Figure 6 combines the results of the ranks

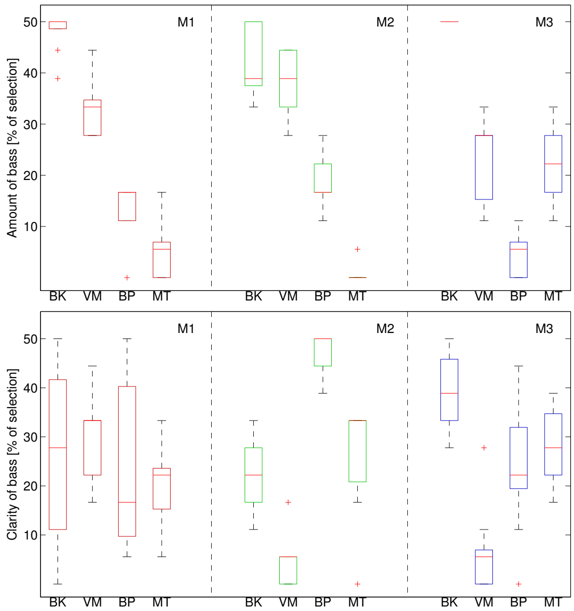

Figure 5: Analysis of the paired comparison test. The y-axis indicates the percentage of cases when an assessor has chosen the concert hall against any other concert hall in the comparison test with respect to amount of bass and clarity. The maximum for a concert hall is 50 %, which means it has been chosen in all comparison pairs.

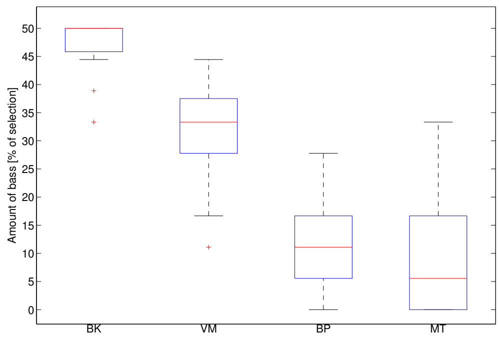

Figure 6: Combined results of all musical excerpts for the amount of bass. The maximum for a concert hall is 50 %, which means it has been chosen in all comparison pairs.

for the amount of bass with all musical excerpts. It shows clearly that BK was considered to have the most bass.

As for clarity, all assessors found the comparison of clarity more difficult than that of the amount of bass, and it took on average 5 minutes longer to complete this part. The hall that was assessed the clearest depends on the musical excerpt, and the variation between the assessors is greater than in assessing the amount of bass.

With M1, the hall rankings in clarity are very equal with VM having the highest mean, although the mean ranks are not significantly different ($H(3) = 3.38, p < 0.34$) . It may be that assessing clarity from a tremolo of low-register strings was not an easy task. With M2, the highest clarity was perceived in BP, followed by MT. The mean ranks between the vineyard halls and shoebox halls is significant ($H(1) = 16.67, p < 0.001$) With M3, BK was perceived both the clearest and the strongest in bass. The mean rank of VM is significantly worse than BK or MT ($H(3) = 20.65, p < 0.001$).

Finally, the correlation between the main attenuation frequency of the seat-dip effect and the perceived amount of bass is $\rho = 0.78$. For comparison, the correlation between the attenuation frequency and clarity is $\rho = -0.22$.

5 Discussion

The shoebox halls are in general characterised with a full overall sound and strong bass, and this can be explained by looking at the time-frequency responses: the level of bass is more than 10 dB higher in the shoebox halls than in the vineyard halls (see Figure 2). The fact that the main attenuation frequency of the seat-dip effect is higher in the shoebox halls than in the vineyard halls seem to help retain a high level of bass below the main attenuation frequency. Furthermore, the wideband attenuation caused by the seat-dip effect does not seem to hamper the perception of the amount of bass in the shoebox halls.

In general, the level of reverberation is considered favourable for the perceived of amount of bass [12]. Other factors that may have influence these results are the seat upholstery and the absorption of low frequencies by the walls and the ceiling [1]. Furthermore, in this listening test, two assessors reported that the apparent width was confused with the amount of bass.

Whereas the results for the amount of bass are straightforward, the same does not apply in the case of clarity. The standard clarity values C_{80} do not seem to explain the results alone. Clarity itself is not a clearly defined concept; it can have various interpretations between listeners. It involves the articulation of the instruments, their attack and separability, and many properties of the concert hall such as reverberation time. In addition, the presence of bass may influence it. In this study, 8 out of 9 assessors reported that they had difficulties in comparing the clarity between a hall with a blurred bass and a hall with no bass.

The reverberation time in VM is considerably higher than in the other halls, and this may have adversely influenced the perception of clarity. In particular in M2, the tuba fused to the string instruments due to the long reverberation. The reason why VM performs a bit better in M1 than in the others is that the tremolo parts benefit from reverberation resulting in a good blending. In this case, clarity or articulation may not be considered important.

The research here was conducted in unoccupied halls whereas concerts in reality have an audience. However, the audience does not affect the seat-dip attenuation below 800 Hz [10]. In addition, the low frequency absorption coefficients for empty and occupied seats do not seem to differ considerably [1]. Thus, the results obtained with unoccupied halls can be considered valid.

6 Conclusions

The amount of bass and clarity in four concert halls was compared in a listening test with three musical excerpts containing double bass and cello sections, and a tuba. The perceived amount of bass is clearly stronger in the shoebox-shaped concert halls with wideband seat-dip attenuation than in the vineyard halls with narrowband seat-dip attenuation. The perception of clarity depends on the musical excerpt (M1-M3).

All in all, it can be said that if these three instrument groups are playing a tremolo (M1), then the amount of bass is important. If the instruments have a clear melody or phrase (M2), then articulation or clarity is important. If the instruments are accompanying other instruments (M3), then the amount of bass is important, and clarity seems to influence the quality of bass.

Finally, the results may change when the entire orchestra is included in the auralisation. In particular, the higher partials of the double bass and the cello can be complimented by the violins. Furthermore, there is some evidence that the entire orchestra may mask the perceived level of bass [7]. Therefore, for future studies the inclusion of the the entire orchestra is seen necessary for comparison.

Acknowledgments

The research leading to these results has received funding from the Academy of Finland, project no. [257099].

7 References

[1] L. L. Beranek, *Music, Acoustics, and Architecture*, 2nd ed. New York, USA: Springer- Verlag, 2004.

[2] J. S. Bradley and G. Soulodre, "Factors influencing the perception of bass," *Journal of the Acoustical Society of America*, vol. 101, p. 3135, 1997.

[3] M. Barron, "Bass sound in concert auditoria," *Journal of the Acoustical Society of America*, vol. 97, no. 2, pp. 1088–1098, 1995.

[4] G. A. Soulodre and J. S. Bradley, "Subjective evaluation of new room acoustic measures," *Journal of the Acoustical Society of America*, vol. 98, no. 1, pp. 294–301, 1995.

[5] L. Beranek, "The sound strength parameter G and its importance in evaluating and planning the acoustics of halls for music," *Journal of the Acoustical Society of America*, vol. 129, no. 5, pp. 3020–30 260, 2011.

[6] L. Kirkegaard and T. Gulsrud, "In search of a new paradigm: How do our parameters and measurement techniques constrain approaches to concert hall design?" *Acoustics Today*, pp. 7–14, January 2011.

[7] N. Nishihara and T. Hidaka, "Loudness perception of low tones undergoing partial masking by higher tones in orchestral music in concert halls," *Journal of the Acoustical Society of America*, vol. 132, no. 2, pp. 799–803, 2012.

[8] J. Pätynen, S. Tervo, and T. Lokki, "Analysis of concert hall acoustics via visualisations of time-frequency and spatiotemporal responses," *Journal of the Acoustical Society of America*, vol. 133, no. 2, pp. 842–857, 2013.

[9] T. Schultz and B. Watters, "Propagation of sound across audience seating," *Journal of the Acoustical Society of America*, vol. 36, no. 5, pp. 885–896, 1964.

[10] G. Sessler and J. West, "Sound transmission over theatre seats," *Journal of the Acoustical Society of America*, vol. 36, no. 9, pp. 1725–1732, 1964.

[11] K. Ishida, "Investigation of the fundamental mechanism of the seat-dip effect – Using measurements on a parallel barrier scale model," *Journal of the Acoustical Society of Japan (E)*, vol. 16, no. 2, pp. 105–114, 1995.

[12] J. Bradley, "Some further investigations of the seat dip effect," *Journal of the Acoustical Society of America*, vol. 90, no. 1, pp. 324–333, 1991.

[13] L. L. Beranek, F. R. Johnson, T. Schultz, and B. Watters, "Acoustics of Philharmonic Hall, New York, during its first season," *Journal of the Acoustical Society of America*, vol. 36, no. 9, pp. 1247–1262, 1964.

[14] W. Davies, T. Cox, and Y. Lam, "Subjective perception of seat dip attenuation," *Acta Acustica united with Acustica*, vol. 82, pp. 784–792, 1996.

[15] A. Walther, P. Robinson, and O. Santala, "Effect of spectral overlap on the echo threshold for single reflection conditions," *Journal of the Acoustical Society of America*, vol. 134, no. 2, pp. EL158–EL164, 2013.

[16] W. Davies and T. Cox, "Reducing seat dip attenuation," *Journal of the Acoustical Society of America*, vol. 108, no. 5, pp. 443–448, 2000.

[17] J. Pätynen, V. Pulkki, and T. Lokki, "Anechoic recording system for symphony orchestra," *Acta Acustica united with Acustica*, vol. 94, no. 6, pp. 856–865, December 2008.

[18] J. Pätynen, S. Tervo, and T. Lokki, "Simulation of the violin section sound based on the analysis of orchestra performance," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2011)*, New Paltz, NY, USA, 2011, pp. 173–176.

[19] J. Pätynen, "A virtual loudspeaker orchestra for studies on concert hall acoustics," PhD Dissertation, Aalto University School of Science, Espoo, Finland, 2011.

[20] S. Tervo, J. Pätynen, and T. Lokki, "Spatial decomposition method for room impulse responses," *Journal of Audio Engineering Society*, vol. 61, no. 1, pp. 1–13, 2013.