Perceptual thresholds for realistic double-slope decay reverberation in large coupled spaces

Paul Luizard\textsuperscript{a} and Brian F. G. Katz

\textit{Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur, Centre National de la Recherche Scientifique, Boîte Postale 133, F91403 Orsay, France}

Catherine Guastavino

Centre for Interdisciplinary Research on Music Media and Technology, School of Information Studies, McGill University, Montreal, Canada

(Received 13 March 2013; revised 16 November 2013; accepted 12 November 2014)

Reverberation highly influences sound perception in enclosed spaces. The reverberation time (RT) metric, used to quantify reverberation in single volumes, is inappropriate for coupled spaces characterized by non-exponential double-slope energy decays. Previous research on reverberation perception of double-slope decays has been predominantly based on varying basic impulse response characteristics such as decay times corresponding to reverberation times of individual volumes presented as independent variables. Alternatively, several studies have employed geometrical room acoustic software simulations to generate collections of responses while varying architectural parameters such as coupling area and room volumes. To avoid issues related to geometrical acoustics simulations, such as position dependence and limitations of some software to properly simulate coupled volume behavior, this study examines perception of the variability of reverberation typical of a physical coupled volume system. Employing an established statistical model, the control parameter of coupling area aperture which acoustically connects the volumes serves as the independent variable. Two listening tests were conducted to determine perceptual thresholds using an ABX discrimination task. The range of tested values corresponded to physically realizable variations. Just noticeable differences (JNDs) were derived with an average JND of \( \approx 10\% \) variation of the coupling aperture. No significant differences were found between different musical excerpts. © 2015 Acoustical Society of America. [\texttt{http://dx.doi.org/10.1121/1.4904515}]

I. INTRODUCTION

Listeners’ experience in a concert venue is influenced by various factors\textsuperscript{1–3} such as reverberance, clarity, sound strength, and spatial impression. These factors can be related to physical and acoustic parameters of the room.\textsuperscript{4} However, small variations of physical parameters are not always perceived by listeners. In classical Sabinian single volume rooms, the reverberation time (RT) is defined as the time required for reflected sound to decay by 60 dB below the level of the direct sound. This acoustical parameter is appropriate for volumes in which reverberation profiles decay exponentially and discrimination thresholds, or just noticeable differences (JNDs), have been reported for reverberation in single volumes\textsuperscript{5–8} under various test conditions.

In coupled spaces, when an adjacent reverberant volume is connected to a less reverberant volume, energy exchanges between the two spaces can result in reverberation profiles with non-exponential decays, being the sum of several exponential decays. Uniquely describing this particular type of sound energy decay requires several acoustical parameters such as the decay rate of each slope and the initial level of each new slope. Figure 1 and Table I present an example of a double-slope decay and the parameters used for unique quantification.

In a real system, it is not possible to independently vary these parameters, as would be the case, for example, with electronic artificial reverberation units which can be independently adjusted. In a physically coupled system, variations of the physical coupling area between the volumes lead to changes in the decay profiles in terms of decay rate and level between slopes. As a design parameter, the coupling area offers the most common control mechanism in coupled volumes architectural applications after construction. In coupled volume concert halls, this is often represented by the configuration of an array of operable doors between the volumes, adjusted as a function of the desired acoustic conditions for a given performance.

As previous JND studies on reverberation have concentrated on single-slope decays, it is unclear what the perceptual limits are for complex variations of decay slopes as are commonly encountered in coupled volume systems. The present study aims at estimating perceptual thresholds for reverberation in coupled volumes as a function of the coupling aperture connecting two volumes and for different musical excerpts.

A. Overview of a coupled volume system

The energy decay profile in a coupled volume system, as show in Fig. 1, can be modeled as the superposition of a series of exponential decays. The parameters of these decays
can be approximated using the statistical approach developed by Cremer and Müller, which considers a closed system of two ideal (Sabinian) volumes exchanging sound energy. An acoustical power balance considering these cavities successively leads to a system of two differential equations. For a system with only two volumes, the solution of such a system in the transient excitation condition is a linear combination of exponentials with decay rates $\delta_i$ and initial energy levels $E_{ij}$, 

$$E_1(t) = E_{ii} \exp(-2\delta_{1i}t) + \frac{\delta_{1i} E_{i\Pi} \exp(-2\delta_{1i}t)}{\delta_1} \cdot (1)$$

$$\begin{align*}
\delta_{1\Pi} &= \frac{1}{2}(\delta_1 + \delta_2) \pm \sqrt{\frac{1}{4}(\delta_1 + \delta_2)^2 + k_1 k_2 \delta_1 \delta_2}; \\
\delta_1 &= \frac{cA_{11}}{8V_1}; \quad \delta_2 &= \frac{cA_{22}}{8V_2}; \\
k_1 &= \frac{S_c}{A_{11}}; \quad k_2 = \frac{S_c}{A_{22}},
\end{align*} (2)$$

where subscripts 1 and 2 correspond to the main room and the reverberation chamber, respectively. $E_1(t)$ is the energy density, $V_i$ are the volumes, $S_c$ is the coupling area between the volumes, $A_i$ are the equivalent absorption areas as per Sabine, and $c$ is the speed of sound. $\delta_{1\Pi}$ are the decay rates of the straight portions within the curved energy decays while $\delta_{1,2}$ are the natural sound energy decay rates of each independent room. The so-called coupling factors $k_{1,2}$ quantify the probability for energy density to enter the adjacent room. The initial energy levels $E_{ij}$ in Eq. (1) depend on the same architectural parameters.

Equations (1) and (2) show that varying $S_c$ leads to various decays $E_1(t)$. Figure 2 illustrates the interaction between the base acoustical parameters (decay times corresponding to each individual volume $DT_1$ and $DT_2$ and their intersection point $BP_L$) as a result of changes in the coupling area $S_c$, here represented relative to the total internal surface area $S_1$. It is clear from this illustration that for a given architecture (volumes and absorption characteristics fixed) a variation in the coupling aperture has a direct effect on both decay rates within each volume as well as the level difference between the two decays.

This theoretical model described above provides the basis of our understanding of coupled volume behavior, in the range of its associated approximations such as the limiting condition of lightly coupled volumes (small $k_i$). Beyond this condition, as the coupling area increases, and the separate volumes become more integrated, the simple energy exchange approximation will no longer apply. This model has been employed in various studies of coupled volumes and performance spaces.

One may note the obvious difficulties in applying acoustic metrics designed for single volume exponential decays to such a system. For example, if one examines the evolution of $BP_L$ which characterizes the bending point in the decay curve, for values where $BP_L > 10 \, dB$, the calculation of $BPL_{10}$, as defined in the ISO 3382-1 standard, will be directly affected by the contributions of the second slope. Similarly, previously defined parameters for characterizing double slope decays such as the “coupling coefficient” (Ref. 14), defined as $T_{30}/T_{15}$ [not to be confused with the coupling factor in Eq. (2)], will be highly dependent on $BP_L$, and whether it falls within the ranges of $BP_L > -5 \, dB$, $BP_L \in [-5, -35] \, dB$, or $BP_L < -35 \, dB$. In view of this, such inappropriate metrics should be avoided in coupled systems.

### B. Previous perceptual studies

Perceptual aspects of double-slope decay reverberation have previously been studied using several different types of stimuli. The room impulse response (RIR) synthesis process...
has been achieved by two basic means. The more typical approach is to directly apply a decay profile (generated as a sum of several linear decays) to frequency band filtered noise. An alternate method is to employ geometrical acoustics prediction software to simulate RIR of a given architectural configuration, implying the creation of a 3D modeling of the space.

Atal and Schroeder\textsuperscript{15} constructed curved energy decays with two different methods. First, by directly manipulating the early to late energy ratios, second by manipulating two independent exponential decays in order to create different curved sound decays. Participants were presented with RIR pairs, convolved with speech or music, containing exponential and non-exponential decay profiles, and were asked to judge which sound was more reverberant. Results showed a high correlation between subjective responses and the decay rate evaluated over the first 160 ms, or 15 dB as calculated from actual measured responses. This work provided the foundation for our current understanding of “running reverberance” as being dominated by early decay metrics.

Along a similar basis, Frissen \textit{et al.}\textsuperscript{16} used the combination of two independent exponential decays, convolved with a short syllable speech excerpt (350 ms). Participants had to compare the coupled response to a reference exponentially decaying sound (RT = 1.8 s). Results showed that the sounds were perceived to be different if the RT of the second slope was at least 1.5 times greater than the one corresponding to the exponential decay.

Picard\textsuperscript{17} also constructed double slope decays based on independent exponential decays, allowing for varying BP\textsubscript{L} for a variety of DT\textsubscript{1} and DT\textsubscript{2} conditions. Stimuli were RIR convolved with either short (0.24 s) or long (0.86 s) noise bursts, windowed with a 0.034 s fade-out. An ABX task was proposed to participants who compared single and double slope decays to obtain a perceptual equivalence point. Results showed that higher BP\textsubscript{L} were required when DT\textsubscript{1} and DT\textsubscript{2} were more similar.

Using geometrical acoustics prediction software, Bradley and Wang\textsuperscript{10} proposed to examine perceptual differences related to architectural parameters. RIRs were synthesized with \textsc{odeon},\textsuperscript{18} varying room volumes, absorption coefficients, and coupling area, which were then convolved to provide binaural renderings of running music. The disparity between the theoretical coupled system response and the simulated RIRs was noted. Subjective ratings of reverberation and clarity were made for various configurations. Higher BP\textsubscript{L} and greater decay ratio (DT\textsubscript{2}/DT\textsubscript{1}) lead to a greater degree of perceived reverberation, or reverberance. These results corroborate those of Atal \textit{et al.}\textsuperscript{15} linking general reverberance to the early portion of the decay, which would be more prominent with greater DT\textsubscript{2} and higher BP\textsubscript{L}.

Similarly, Ermann\textsuperscript{19} generated sound stimuli by convolving music and organ stop-chord recordings with RIRs calculated with \textsc{catt-acoustic},\textsuperscript{20} for two room models to investigate both perception and preference of coupled reverberation. Participants (college students and attendees at an audio conference) were provided explanations on the theory of coupled volume acoustics. They were presented with pairs of stimuli and asked if the sounds were the same or different, if different then which condition had a more pronounced double-slope decay, and which they preferred. Results did not show a clear tendency but subjects were more likely to differentiate the stimuli when the late parts of the decays were more varied from one another. Full length and truncated examples were tested, highlighting no reduction in detection ability from only running reverberation conditions. Comparisons between listener position versus coupling configuration showed higher subjective sensitivity to position, indicating other factors were in play. Finally, the difficulty of the task was highlighted by a non-negligible percentage of respondents who could not correctly perform the reference task (false positive detection between two identical stimuli).

In summary, perception of double-slope reverberation has been studied with considerations based on the impulse response itself and associated metrics. Certain previous studies\textsuperscript{15–17} manipulated various slope values to build multi-slope impulse responses, which are not necessarily linked to a plausible coupled volume system. Variations in the total energy of the decay were noted as a potential perceptual cue. Others\textsuperscript{10,19} considered the simulation of example coupled volume systems using ray-tracing software and adjusted coupling area as one possible variable. Poor correspondence between the predicted response and the theoretical response, as well as positional variations and other early reflection effects of using a precise geometry, were noted as potentially problematic.

C. The present study

This study uses the average dimensions and absorption characteristics\textsuperscript{21} of some existing coupled volume concert halls to define a space with realistic characteristics. Using a simple statistical model of sound energy decay enables total control of the synthesis parameters to construct impulse responses, i.e., apply decay curves on filtered noise, as was performed by Seraphim\textsuperscript{3} in the fundamental experiments on difference limen of reverberation times in single volume rooms, as referenced in the ISO-3382 standard.\textsuperscript{13}

In addition, the fact that no spatial dependency is involved in the present model of energy decay allows for considering the results at any location of sound source, receiver, and coupling area within the main room of the coupled volume system.

II. EXPERIMENTS

The aim of these perceptual experiments is to evaluate the JND for coupled reverberation as a function of the variation of the acoustical coupling area as a natural control parameter. A range of physically valid coupling area values was used to generate the room impulse responses and the results are expressed in terms of the size of the coupling area and various acoustical parameters.

Short duration stimuli were used for the discrimination task, with the assumption that participants would therefore base their judgment primarily on the stop-chord reverberation, heard after the direct sound is emitted when only reflected sound propagate throughout the space, as opposed...
to the running reverberation heard while an instrument is playing. This study was separated into two distinct experiments, each having a different reference sound stimulus based on the extreme values of coupling area in order to double-check the validity of results. In other words, each experiment can be seen as estimating how sensitive people are to the variation of coupling area from a given initial—or reference—value. Thus, the experiments are based on two different initial coupling areas and the consistency of JNDs over these two conditions is investigated. While the first experiment was 2 h long, the second one was only 1 h long as fewer stimuli were presented (stated times included pauses to ensure participant concentration).

A. Stimuli

The stimuli presented to participants consisted of impulse responses convolved with three short musical excerpts: bongos from Bang & Olufsen, Music For Archimedes, a symphony orchestra, and a solo soprano singer recorded at Aalto university. The impulse responses were constructed by applying the decay profile to normally distributed random number sequence used in previous studies. Following previous work by Moorer, this noise was first equalized in third octave bands from 63 Hz to 16 kHz to obtain the spectrum presented in Fig. 3. A collection of curved temporal decays, presented in Fig. 4, were applied to this noise to obtain the series of test impulse responses.

The characteristics of these decays depend on architectural parameters of the rooms, such as the volumes, the equivalent absorption areas, and the size of coupling area. The architectural parameters used to define the coupled volume system are provided in Table II. These values produced independent reverberation times which were slightly exaggerated as compared to typical constructions in order to obtain two distinct slopes within the curved decays. While absorption and room morphology were held constant, the variable element between the synthesized RIR was the size of coupling area, set to 1%, 2%, 3%, 4%, 5%, 9%, and 10% of the total inner surface of the main room ($S_1 = 4560 \text{ m}^2$), which represents an interval from $S_c = 45.6$ to 456 m$^2$. This represents the realistic situation of a coupled volume concert hall, wherein the coupling area is changed through variations in opening of an array of doors. This range of values covers coupling areas which can be found in actual coupled volume concert halls (see Table III). The non-regular steps between 1% and 10% were chosen considering the non-linear relationship between the coupling area variation and the corresponding acoustical parameters (see Figs. 2 and 5).

These acoustical parameters correspond to the time-energy decay curves presented in Fig. 4, which define the envelopes of the impulse responses. $DT_i$ are calculated directly from the decay rate $\delta_i$ of each slope, corresponding to the argument of the exponential in each term in the first line of Eq. (1). These decay rates, combined with the initial energy density of each term, allow for obtaining two exponential decays corresponding to each slope which intersect at one point. The so-called bending point is estimated as the closest point on the decay curve to this intersection point.

To make sure that the only differences in loudness are due to the differences between energy decays, the sounds are

<table>
<thead>
<tr>
<th>Location</th>
<th>$S_1$ (m$^2$)</th>
<th>Max $S_c$ (m$^2$)</th>
<th>Max $S_c$ %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lucerne, Switzerland</td>
<td>5050</td>
<td>910</td>
<td>18.0</td>
</tr>
<tr>
<td>Lahti, Finland</td>
<td>4410</td>
<td>615</td>
<td>14.8</td>
</tr>
<tr>
<td>Philadelphia, USA</td>
<td>5700</td>
<td>555</td>
<td>9.7</td>
</tr>
<tr>
<td>Dallas, USA</td>
<td>5750</td>
<td>521</td>
<td>9.1</td>
</tr>
<tr>
<td>Birmingham, UK</td>
<td>5990</td>
<td>195</td>
<td>3.3</td>
</tr>
</tbody>
</table>

FIG. 3. Spectrum of the noise used to generate impulse responses.

FIG. 4. Temporal energy decays from Cremer–Müller statistical theory for different aperture sizes expressed in percent of total inner surface of the main room. Lower decays (darker) correspond to small coupling areas while higher decays (lighter) stand for large coupling areas. Dots show the bending point for each curve.

TABLE II. Input parameters to Cremer’s theory implementation. RT stands for the natural reverberation time of each independent room.

<table>
<thead>
<tr>
<th></th>
<th>$V_i$ (m$^3$)</th>
<th>$S_i$ (m$^2$)</th>
<th>$a_i$</th>
<th>RT,</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main room ($i = 1$)</td>
<td>19 000</td>
<td>4560</td>
<td>0.55</td>
<td>1.22</td>
</tr>
<tr>
<td>Reverberation chamber ($i = 2$)</td>
<td>6050</td>
<td>2040</td>
<td>0.10</td>
<td>4.77</td>
</tr>
</tbody>
</table>

TABLE III. Surface specifications of various existing coupled volume concert halls, from Beranek (Ref. 21). $S_i$: inner surface of the main room; $S_c$: coupling area expressed in terms of surface unit and in % of $S_1$. 

<table>
<thead>
<tr>
<th>Location</th>
<th>$S_1$ (m$^2$)</th>
<th>Max $S_c$ (m$^2$)</th>
<th>Max $S_c$ %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lucerne, Switzerland</td>
<td>5050</td>
<td>910</td>
<td>18.0</td>
</tr>
<tr>
<td>Lahti, Finland</td>
<td>4410</td>
<td>615</td>
<td>14.8</td>
</tr>
<tr>
<td>Philadelphia, USA</td>
<td>5700</td>
<td>555</td>
<td>9.7</td>
</tr>
<tr>
<td>Dallas, USA</td>
<td>5750</td>
<td>521</td>
<td>9.1</td>
</tr>
<tr>
<td>Birmingham, UK</td>
<td>5990</td>
<td>195</td>
<td>3.3</td>
</tr>
</tbody>
</table>
normalized to 90% of their maximum value after convolution.

B. Participants

A total of 21 participants completed the first experiment and 19 new participants completed the second one. They were 16 males and 24 females, aged 18 to 51, the mean age being 27. All of them had more than ten years of practice on a musical instrument or in sound engineering. They were tested for normal hearing using a standard audiometric test from 250 Hz to 8 kHz. Participants received $20 (Canadian) for their participation.

C. Procedure

Stimuli were presented over headphones (Sennheiser HD650) in an acoustically treated laboratory using a Matlab graphical user interface on a Mac Pro computer connected to an audio interface (Grace Designs m904) at a sampling rate of 44.1 kHz with a bit-depth of 16 bits.

The experiments employed a standard method of constant stimuli paradigm, with a two-interval, forced-choice task. The ABX form was chosen with a standard sound corresponding to the largest coupling area set to 10% of the inner surface of the main room for the first experiment. The standard sound for the second one was the smallest coupling area set to 1% of the inner surface of the main room. At each trial, two different sounds A and B were presented to the participant. A or B were randomly the standard sound or one of the six other sounds. The third presented sound, called X, was pseudo-randomly A or B. The participants could listen to the three sounds as many times as necessary. The task was to choose whether sound X was sound A or sound B. A practice session of ten trials (without feedback) allowed participants to become familiar with the task and the sounds of the real test. As the participants had different habits and tolerance thresholds in terms of sound intensity regarding critical listening situation, they were told to set the sound level as they thought it was loud enough to hear subtle differences, within a possible range of 4 dB, at 88.2 dB sound pressure level ref. 20 \(\mu\)Pa, in order to keep the results comparable. This level adjustment was made during the practice session and would remain unchanged during the test.

The first experiment was organized in three blocks corresponding to the three anechoic sounds. In each block, the six pairs of sounds (standard RIR + 1 out of the six others) were presented ten times in random order for a total of 60 trials per block so that the duration of the first experiment was

---

**TABLE IV.** Double slope decay parameters for various coupling area sizes. Those parameters are defined in Table I.

<table>
<thead>
<tr>
<th>(S_c/S_i) (%)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>(S_c (m^2))</td>
<td>45.6</td>
<td>91.2</td>
<td>136.8</td>
<td>182.4</td>
<td>228.0</td>
<td>410.4</td>
<td>456.0</td>
</tr>
<tr>
<td>(DT_1 ,(s))</td>
<td>1.22</td>
<td>1.20</td>
<td>1.18</td>
<td>1.16</td>
<td>1.13</td>
<td>1.00</td>
<td>0.96</td>
</tr>
<tr>
<td>(DT_2 ,(s))</td>
<td>4.03</td>
<td>3.49</td>
<td>3.11</td>
<td>2.84</td>
<td>2.63</td>
<td>2.21</td>
<td>2.17</td>
</tr>
<tr>
<td>(BP_t ,(s))</td>
<td>0.835</td>
<td>0.675</td>
<td>0.571</td>
<td>0.486</td>
<td>0.411</td>
<td>0.180</td>
<td>0.140</td>
</tr>
<tr>
<td>(BP_L ,(dB))</td>
<td>-37.0</td>
<td>-29.7</td>
<td>-25.1</td>
<td>-21.5</td>
<td>-18.3</td>
<td>-8.5</td>
<td>-6.8</td>
</tr>
</tbody>
</table>
about 2 h (including an audiogram) with pauses. The second experiment was slightly different and focused on one block corresponding to the anechoic recording of orchestra, composed of 90 trials with the same six pairs of sound stimuli for a total duration of 1 h. The task can be seen as looking for a difference between the sound corresponding to a standard coupling area and the sounds corresponding to different coupling areas, even if the participants were not told that the reference was present at each trial.

The study was presented as dealing with reverberation in coupled spaces, without any further explanation of the coupled volume acoustic specificity. Thus, participants passed the test with their own previous knowledge of acoustics and critical listening. Participants filled out a questionnaire at the end of the test. They were asked to rate the perceived difficulty of the task and to describe in their own words the strategy they used to discriminate the sounds. They were also asked about how often they attended live concerts and in which venues.

III. RESULTS

Results are expressed in terms of proportion of correct answers as a function of the various sizes of coupling area or the corresponding analysis parameters. To obtain psychometric functions, the data were fitted with cumulative Gaussians $\Psi(x)$ free to vary in position and slope by changing parameters of Eq. (3), using the software package PSIGNIFIT. An example can be seen in Fig. 6.

$$\Psi(x) = \gamma + (1 - \gamma - \lambda) F(x, \alpha, \beta),$$
$$F(x, \alpha, \beta) = \frac{1}{\beta \sqrt{2\pi}} \int_{-\infty}^{x} \exp \left( \frac{z - x}{2\beta^2} \right) dz.$$  \hspace{1cm} (3)

Since the ABX form was used, the theoretical minimum performance is the ratio 0.5 which corresponds to random answers or chance, when discrimination is impossible due to the strong similarity of stimuli. The maximum performance lies within the interval $[1 - \lambda; 1]$, where $\lambda$ depends on the participant's results, with a maximum value of 0.06, as advised by Wichmann and Hill.

A. Outlier detection

Among the 21 participants in the first experiment, three obtained random-like distributed results which means that they did not succeed in hearing the differences. Five to eight of them, depending on the original sound (orchestra, soprano, or bongos), had results which lead to poor values of goodness-of-fit. The latter was estimated using statistical deviance, expressed as Eq. (4), or log-likelihood ratio, which quantifies the difference between the psychometric function and the experimental points. Finally, 9 to 12 psychometric functions per original sound have been used so that the mean value of statistical deviance is inferior to 4. In the second experiment, seven participants had satisfactory results with respect to goodness-of-fit, among the 19 initial participants. Two of them were dismissed because of the poor goodness-of-fit value and the others obtained random-like answers suggesting that the task was too demanding for them. While the exclusion of subjects is a regrettable situation, and far from ideal, the difficulty in accomplishing the task should be emphasized, as similar problems with subject reliability were also identified in previous studies under similar conditions.

$$D = 2 \sum_{i=1}^{K} \left( n_i \log_{10} \left( \frac{y_i}{p_i} \right) + n_i (1 - y_i) \log_{10} \left( \frac{1 - y_i}{1 - p_i} \right) \right),$$ \hspace{1cm} (4)

where $y_i$ stands for the experimental data and $p_i$ refers to the predicted model of psychometric function.

B. JND estimation

The discrimination threshold (or JND) is determined from the slope of the psychometric function. It is defined, as per convention, as the difference between the abscissa values that correspond to the second and third quarters of the performance scores, which are approximately 75% [point of subjective equality (PSE)] and 87.5% (see Fig. 6). As widely used in psychophysics, Weber’s law can be written as follows:

$$\frac{\text{JND}}{\text{PSE}} = \frac{\text{JND}_{rel}}{1},$$ \hspace{1cm} (5)

where JND is the value of absolute JND, PSE is the corresponding point of subjective equality, and JND$_{rel}$ is the relative value of JND, expressed in percent and assumed to be a constant quantity.

The first experiment, with the reference sound stimulus corresponding to the largest coupling area, lead to similar JND values of about 30 m$^2$, estimated with the three original recordings, as presented in Fig. 7. While these results are of the same order of magnitude as the first experiment, results from the second experiment, which used a reference sound stimulus corresponding to the smallest coupling area, lead to much smaller JND estimations, below 10 m$^2$. The same
difference between experiments 1 and 2 is observed for the PSE values. Writing the results as the ratio of absolute JND and PSE, as in Eq. (5), allows for considering the relative JND, which is independent from the size of coupling area.

An average relative JND of 10% of the coupling area is found in the first experiment (see Fig. 7). It means that for a given size of coupling area the audience will perceive a change in reverberation if the coupling area is modified by at least 10% of its current size.

The average value of JND in the first experiment of the current study condition is around 30 m², depending on the type of musical stimuli. The corresponding mean PSE value is approximately 300 m². This is a relatively high value considering coupling areas in real coupled volume concert halls and would be smaller if the reference sound is chosen as corresponding to a smaller coupling area, as shown with the comparison of the first and second experiments in Fig. 7. However, the PSE itself should not influence the relative JND value since the absolute JND is directly proportional to the PSE, according to Eq. (5) derived from Weber’s law. Again, the second experiment shows that the relative JND, expressed in percent, maintains a similar value while the PSE is strongly reduced.

Furthermore, the deviance, or log-likelihood ratio [see Eq. (4)], as estimated by the PSIGNIFIT package and its average value is inferior to 5 for every type of sound, as presented in Fig. 7. This quantity can be seen as an estimation of judgment reliability, which characterizes the dispersion of experimental data. The threshold value of 5 has been chosen to be close to the one used in a previous study, in order to keep the same order of requirement for judgment reliability. This quantity was used to detect outliers, as explained in Sec. III A.

JNDs have also been calculated for the decay curve descriptors. Figure 8 shows the relative JND estimated for each acoustical parameter. Mean JND for DT₁ varies from 4% to 7% depending on the stimulus which is in accordance with previously estimated values for single slope sound decay by Seraphim. In large volumes as well as in small rooms, slightly higher values are also found, around 10%, with noise, speech, or Western music stimuli. Another study has found even higher reverberation threshold values above

---

**FIG. 7.** Results of the discrimination experiment in terms of JND, PSE, and deviance for the aperture parameter, presented as boxplots. For each sound, large rectangles represent lower and upper quartiles, the line in the middle is the median, and extreme small lines show the smallest and largest observation. Snd₁, bongos; Snd₂, soprano singer; and Snd₃, orchestra.

**FIG. 8.** Results of the discrimination experiment in terms of relative JND for the acoustic parameters, presented as boxplots. Snd₁, bongos; Snd₂, soprano singer; and Snd₃, orchestra.
TABLE V. Differences between double slope decay parameters for various coupling areas and the reference areas of 10% (experiment 1) and 1% (experiment 2) of the audience room surface.

<table>
<thead>
<tr>
<th></th>
<th>Standard $S_r/S_1 = 10%$</th>
<th></th>
<th>Standard $S_r/S_1 = 1%$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_r/S_1 (%)$</td>
<td>10</td>
<td>20</td>
<td>30</td>
</tr>
<tr>
<td>$\Delta(S_r/S_1)(%)$</td>
<td>9</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>$\Delta(S_r)(m^2)$</td>
<td>410.4</td>
<td>364.8</td>
<td>319.2</td>
</tr>
<tr>
<td>$\Delta(DT_1)(s)$</td>
<td>0.26</td>
<td>0.24</td>
<td>0.23</td>
</tr>
<tr>
<td>$\Delta(DT_2)(s)$</td>
<td>1.86</td>
<td>1.33</td>
<td>0.95</td>
</tr>
<tr>
<td>$\Delta(BP_1)(s)$</td>
<td>0.694</td>
<td>0.535</td>
<td>0.431</td>
</tr>
<tr>
<td>$\Delta(BP_2)(dB)$</td>
<td>30.3</td>
<td>22.9</td>
<td>18.3</td>
</tr>
</tbody>
</table>

25% in single rooms, using Chinese music motifs as stimuli, the latter being responsible for the greater JND values according to the authors.

Furthermore, the JND for DT$_2$ is much greater and varies from 20% to 40% in this experiment. The same results are found for the other parameters, the time and level of bending point. This similarity appears as an expected result since these parameters are all strongly linked to the late decay time through the characteristics of the second slope of energy decay.

Table V summarizes the actual differences between each parameter and the reference one, the latter being relative to the coupling areas of 10% and 1% of the total primary (audience) room surface, respectively, used in the first and second experiments. It shows objective differences between the sounds which participants had to compare. Although the progression is not linear and differences between sounds are not the same in both cases, results are still consistent, which confirms the JND value.

An one-way analysis of variance was performed on the means of JND obtained in the first experiment with the different types of sounds. The differences did not reach statistical significance since the probability of the test is too high ($0.3 < F < 2; 0.1 < p < 0.8$) in all cases. This is consistent with previous studies showing that the discrimination of stopped chord reverberant stimuli is not strongly affected by the nature of the sound stimulus.27

The second experiment, which used the orchestra sample only and whose reference sound was the one based on a coupling area of 1% of the total primary room surface, instead of 10% as in the first experiment, lead to similar JND, the mean value being 13% of the reference coupling area. Figures 7 and 8 show the results of JND, PSE, and deviance for the second experiment in comparison with the results from the first experiment. Relative JND values are consistent in both experiment. One can notice in Fig. 7 that the absolute JND and PSE values are much lower in the second experiment while the relative JND remains similar to the ones in the first experiment. This suggests that Eq. (5) is validated for two extreme coupling areas, i.e., the relative JND is constant over a wide range of coupling areas among those generally used in concert halls (see Table III).

C. Post-questionnaire analysis

Participants were asked to rate the perceived difficulty of the task, on a scale ranging from 0 (very easy) to 10 (very difficult). The means over all participants in the first part of experiment were 6 for the orchestra stimuli, 7 for the bongos, and 8 for the soprano singer. The orchestra as the only proposed sound in the second experiment received a subjective difficulty rating from 2 to 8 with average of 5 by the retained participants, while the average rating was 6 among the 19 participants. So, even if no significant difference was observed for the JNDs between the different musical excerpts, the soprano was judged to be harder to discriminate than the orchestra or the bongos, possibly because of its narrower frequency bandwidth and its temporal softness, without strong attacks. The bongos instead presented strong temporal peaks because of the percussive nature of the instrument and the orchestra sample had a very rich and large frequency spectrum which might make it easier to discriminate when only a slight difference existed between two stimuli. Although the sound engineering population would be expected to be more adept at judging different degrees of reverberation, participants having such a background interestingly rated this test as more difficult than other participants.

Participants reported using various strategies in order to compare the different sounds presented in a trial. Twelve mentioned focusing on the temporal aspects—attacks and mostly decay of sound—while nine participants paid more attention to the frequency content of the various sounds, e.g., comparing the difference of decay rate of high vs low frequencies or listening to the spectral content. Finally, eight participants stated clearly paying attention to both temporal and spectral aspects of the sound stimuli. The remaining participants did not explicitly mention any comparison strategy.

All participants reported attending concerts regularly, with an average number of four concerts per month with a majority of live acoustic instruments in a variety of reverberant environments. The venues mentioned included large concert halls, theaters, churches, and to a lesser extent jazz clubs and other small venues.

IV. DISCUSSION

It should be noted that all participants had a minimum of 10 years of musical training and their results may not be generalized to less experienced listeners. In addition, the proposed sounds were monophonic signals with no specific direction of sound incidence and the synthesized impulse responses did not provide distinct early reflections with lower temporal density than purely reverberated sound.
Thus, on the one hand, stimuli might be less ecologically valid than what would have been generated with impulse responses having more distinct direct sound and early reflections, synthesized in stereophonic or even binaural signals, but on the other hand they did not restrict the study to a specific room geometry and configuration (e.g., the general room shape, the source and receiver location relative to the open coupling area from which the late reverberation is introduced in the main room), as per classical tests on reverberation perception. In that sense, the results are not specific to a given geometrical configuration and are more likely to apply generally, as stated in the Introduction. At each trial of the listening test, the task was to identify the differences of reverberation between two similar sound stimuli, synthesized with impulse responses made of reverberated sound. Hence, their relative differences were more important than their absolute sound quality—which was criticized by two participants only—for the specific purpose of the study.

Furthermore, the results expressed in terms of acoustical parameters should be considered with caution since comparing JND for DT$_{1}$ and DT$_{2}$ in Fig. 8 shows that the values are much lower for DT$_{1}$ and one could conclude that listeners are more sensitive to the early part of decay as compared to the late part represented by DT$_{2}$. However, this interpretation may be biased since those JND values are estimated out of the same series of RIRs and the difference of obtained values is exclusively due to the early and late slope variation within the initial collection of energy decays presented in Figs. 4 and 5. The same explanation can justify the values obtained for the bending point time and level. As mentioned in the Introduction, acoustical parameters are dependent on the coupling area: the four acoustical parameters and the architectural parameter (size of coupling area) are coupled and cannot be considered separately.

Studying the influence of these parameters independently would require a different protocol for stimuli synthesis. The different slopes and relative level should be handled in order to sculpt the decay curve as desired as was made by Frissen et al.,$^{16}$ e.g., maintaining the decay rate of each slope while varying the initial level of the second one. This procedure would lead to a JND estimation for the second slope initial level (directly related to BP$_{t}$), linked to the late reverberation, at given decay rate of each slope. However, such a response would be totally artificial in the context of coupled volumes because the slopes and their initial levels would vary together in a real room. In the same manner, the decay rate of the second slope could be varied while maintaining its initial level and the decay rate of the first slope. Hence various procedures are possible but do not fit with the present study’s intention to focus on realistic relationships between the acoustical parameters through their interdependency driven by the coupling area.

Finally, further studies could include replicating similar listening tests with variations of individual RTs in each room by means of changing the relative room volumes or inner absorption before varying the coupling area, as can be made in actual concert halls with movable ceiling and absorptive banners. This would change the slope ratio when the rooms are coupled and could affect perceptual thresholds. Since the late reverberation is more audible when the slope ratio is high, i.e., the secondary chamber volume is much more reverberant than the primary audience room, then smaller changes in size of the coupling area might lead to audible differences, making the JND lower than the one found here, where typical overall absorption values$^{21}$ were used in each room.

V. CONCLUSION

This perceptual study determined JND values for double-slope coupled-volume generated reverberation as a function of an architectural parameter, namely, the coupling area. Perceptual thresholds obtained were 10% of variation of a given coupling area. The stimuli used for this study were generated using short anechoic sounds of soprano, bongos, and symphony orchestra convolved with various room impulse responses whose differences were only due to the coupling area variation, according to a simple statistical model of coupled volume sound energy decay. The present JND value is scarcely comparable with previous work where JND values were determined relative to single-slope reverberation only. However, results from the present study are consistent with those obtained in previous perceptive studies$^{10,16,17,19}$ on coupled reverberation in the sense that differences are heard for large variations of slopes and bending point location in the time-energy space, those variations depending physically on the size of coupling area.

ACKNOWLEDGMENTS

The authors would like to acknowledge the Center for Interdisciplinary Research in Music Media and Technology (CIRMMT) for providing the test room and funding the participants, Dr. Ilja Frissen for the useful discussions on data analysis, and Vincent Rogerone for running participants in the second experiment. This research was supported in part by a McGill University William Dawson Award to C. Guastavino.


