RESIZING ROOMS IN CONVOLUTION, DELAY NETWORK, AND MODAL REVERBERATORS

Elliot K. Canfield-Dafilou and Jonathan S. Abel

Center for Computer Research in Music and Acoustics,
Stanford University, Stanford, CA 94305 USA
kermit|abel@ccrma.stanford.edu

ABSTRACT

In music recording and virtual reality applications, it is often desirable to control the perceived size of a synthesized acoustic space. Here, we demonstrate a physically informed method for enlarging and shrinking room size. A room size parameter is introduced to modify the time and frequency components of convolution, delay network, and modal artificial reverberation architectures to affect the listener's sense of the size of the acoustic space taking into account air and materials absorption.

1. INTRODUCTION

Computational methods for simulating reverberant environments are well developed [1], and find application in fields ranging from music recording to virtual reality and film audio production. Room acoustics is an approximately linear and time-invariant process, and there are several widely used methods for room acoustics simulation, including direct convolution with an impulse response [2], delay network-based methods [3], and modal reverberation [4].

In a number of scenarios, it is desirable to manipulate or control the perceived size of a given acoustic space. In a virtual reality or film setting, for instance, the size of the room might be changing over time, and it is preferable that the acoustics of the space change accordingly. In a music recording, performance, or composition environment, different sizes of acoustic space convey different musical impressions, and it is useful to have a palette of room size options associated with a given room response for artistic purposes. Larger spaces tend to be more reverberant and "darker" than smaller ones, but there does not seem to be a systematic way to manipulate the perceived size associated with a given room response. Rafii and Pardo [5] proposed finding relationships between subjective terms and reverb parameters, and Chourdakis and Reiss [6] proposed an adaptive reverberation algorithm based on learning parameters from user actions. In both cases, these methods can be used to modify a reverberation algorithm based on subjective characteristics rather than the physics of changing the dimensions of a room.

In this paper, we introduce a room size parameter for modifying the time and frequency content of an artificial reverberator. We demonstrate a physically informed method for changing the size of a room, taking into account the changes in geometry, absorbing surface area, and volume. We then show how to implement this room size parameter in convolution, delay network, and modal reverberation architectures.

This paper is organized as follows: section 2 introduces the acoustical concepts necessary for modifying room size. Section 3 discusses the implementation of the room resizing parameter in convolution, delay network, and modal reverberaters. Finally, section 4 offers some concluding remarks.

2. ON THE ACOUSTICS OF ROOM SIZE

The room response to a transient sound is often described as a sequence of events over time, a direct path followed by early reflections that give way to late-field reverberation, as seen in Fig. 1. The direct path carries with it information about the source direction, and arrives with a time delay and amplitude fixed according to the source-listener distance. The early reflections contain information about the geometry of the space, and can be simulated using details of the architecture of the space [7]. The late-field reverberation brings to the listener information about the volume of the space and materials present in the space through the frequency dependent rates of sound energy decay. Roughly speaking, the reverberation time is proportional to the ratio of the room volume to the room absorbing surface area [8].

If the room size were doubled, with everything else remaining the same, then the timing of the direct path and early reflections would be stretched by a factor of two. Similarly, if the room size were doubled, then its volume would increase by a factor of eight, while its absorbing area would increase by a factor of four, thereby doubling the reverberation time.

Reasoning along these lines was used by Spandöck in building scale models of proposed concert halls to test how they might sound when built [9]. Spandöck argued that a scale model of a concert hall made with the appropriate materials and filled with a dried gas would respond to a given high-frequency sound the way the larger actual space would respond to a low-frequency sound having the same relative wavelength. Spandöck describes using a magnetic tape deck to play back a sound into the scale model sped up by a factor of, say, eight, while simultaneously recording the response in the model. The recording was then played back, slowed by the same factor. In this way, the original pitch was restored, and the reverberation time increased to match that of the hypothesized full-scale hall.

As described in [10], this approach was independently discovered by Walter Murch while working as a sound editor for motion pictures in the late 1960s, and was used to make long-lasting reverberation. Spratt, et al. present a digital technique for implementing a real-time version of the method, using a loudspeaker and microphone in a physical room [10].

The technique is described as being mathematically equivalent to stretching the room impulse response in time, which has the effect of increasing the reverberation time, and stretching the reflection arrival times. Spratt, et al. argue that the method is similar to slowing the speed of sound or increasing the room size. However, doing either of these will not result in proper reverberation time as a function of frequency, as the relative absorption of sound by air and room materials will not be taken into account.

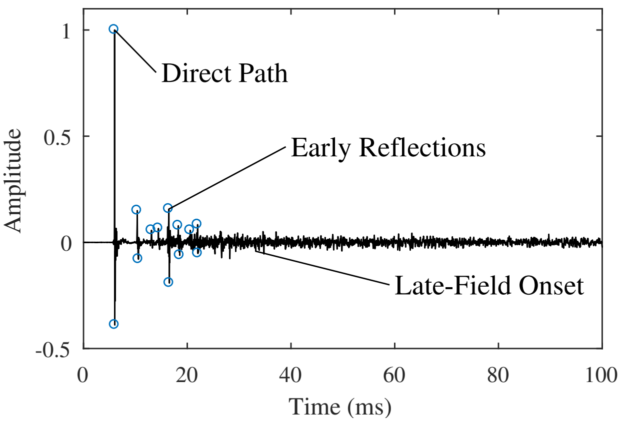

Figure 1: Example room impulse response showing the direct path, early reflections, and late-field reverberation onset.

If a room is proportionally scaled, the echo pattern will be linearly stretched or squished. As a result, the echo density [11], or rate of reflections, will also be linearly scaled. However, to be physically accurate, one must also take the surface area and volume changes into account. Air absorption is nonlinear across frequency, and high frequencies will typically decay faster in a larger room than a small one. Additionally, simply enlarging or shrinking the room via the method described in [10] also proportionally scales all the room materials. For example, the pores in a carpet would be scaled, changing its contribution to the frequency response in the room. Here, we suggest a method for taking the air absorption and materials absorption into consideration when scaling the size of rooms.

3. SCALING ROOM SIZE

Perceived room size may be manipulated in the context of a number of artificial reverberation methods. The idea is to warp the time and frequency axes and adjust the decay times of a given reverberation impulse response according to a desired room size. In addition, the source loudness and radiation pattern may be adjusted according to the room size.

3.1. Reverberation Time

We first describe the change in reverberation time in response to a changing room size as a result of different relative contributions of materials absorption and air absorption.

As described in [8] and elsewhere, the decay over time of well-mixed acoustic energy in a room can be approximated by examining a room with volume V and having objects and surfaces with absorbing area A. The energy density $w(t)$ as a function of time t is assumed to be well mixed and independent of position within the room. After a period of time Δt, the total energy in the room, the product of the energy density and the volume, $Vw(t + \Delta t)$, will be that at time t minus what is lost due to interactions with absorbing surfaces and objects and air propagation,

$$Vw(t + \Delta t) = Vw(t) - Acgw(t)\Delta t - Vaw(t)\Delta t\,, \tag{1}$$

where the term $Acgw(t)\Delta t$ represents surface interaction absorption, and is proportional to the absorbing area A, sound speed c, a constant g, energy density $w(t)$, and time interval Δt, and the term $Vaw(t)\Delta t$ represents air absorption, and is proportional to the volume V, an absorption coefficient a, energy density $w(t)$, and time interval Δt. These absorption terms can be intuitively interpreted—the greater the time interval, the more energy that can be absorbed; the greater the energy density, the more energy that can "leave" the space during the time interval. Rearranging terms, and taking $\Delta t \to 0$, we have

$$\frac{w(t + \Delta t) - w(t)}{\Delta t} \to \frac{dw}{dt} = -\frac{1}{\tau}w(t), \tag{2}$$

and

$$w(t) = w_0 e^{-t/\tau}, \qquad t \geq 0\,, \tag{3}$$

with w_0 being the energy density at time $t = 0$, and τ being a time constant which increases with increasing volume, and decreases with increasing absorbing area,

$$\tau = \frac{V}{Acg + Va}\,. \tag{4}$$

In other words, the energy density in a well mixed room will decay exponentially, decreasing by a factor of $1/e$ every τ units of time. It is typical to measure reverberation time in terms of the time taken for energy to decrease 60 dB, T_{60}, in which case we have

$$T_{60} = \frac{\log_{10} 10^6}{\log_{10} e}\,\tau\,, \tag{5}$$

measured in units of seconds per 60 dB decay.

Energy density is also a function of frequency ω, $w(t, \omega)$, which was dropped from the discussion here for simplicity of presentation. It carries over to frequency-dependent materials and air absorption simply by making the absorbing area A and air absorption a frequency-dependent.

3.2. Adjusting Room Size

Consider a room described by a nominal length L_0. Now scale the room and all of its surfaces and objects to have a new characteristic length L. We want to understand how the decay time changes with a changing room size L. Using (4) and (5), and assuming that the room volume V is proportional to L^3 and the absorbing area A is proportional to L^2, the decay time of the resized room $T_{60}(L)$ is then

$$T_{60}(L) = \frac{L}{L_0\mu + L\alpha}\,, \tag{6}$$

where μ has been introduced to represent the materials absorption for the nominally sized room, α has been introduced to represent the materials absorption. Both μ and α are expressed in terms of 60 dB decay per unit time. Note also that

$$\alpha = \frac{\log_{10} 10^6}{\log_{10} e}\,a\,. \tag{7}$$

The decay time at the nominal room size,

$$T_0 = T_{60}(L_0)\,, \tag{8}$$

may be estimated from the room impulse response or otherwise modeled, and that the air absorption α is known, derived assuming a given temperature, pressure, and humidity, or tabulated [12, 13].

Accordingly, setting $L = L_0$ and solving (6) for the unknown materials absorption μ gives

$$\mu = \frac{1}{T_0} - \alpha\,. \tag{9}$$

Due to errors in estimating decay times from measured impulse responses, the reverberation time T_0 might exceed the air absorption-only reverberation time $1/\alpha$, and (9) would produce a negative value for μ. In these cases (or at such frequencies that this is true), a value of $\mu = 0$ is preferably used, and the reverberation time will not be affected by room size. If it is desired to have a changing reverberation time with room size, a small value for μ could be selected.

Substituting for μ in (6) gives T_0, the decay time as a function of room size L. In the case that μ is given by (9) and not modified, a little algebra gives an expression for $T_{60}(L)$ in terms of the nominal decay time T_0 and the decay time if the only absorption of sound energy were due to air $T_{\text{air}} = 1/\alpha$,

$$T_{60}(L) = \frac{L \cdot T_0 T_{\text{air}}}{L_0 \cdot T_{\text{air}} + (L - L_0) \cdot T_0}\,. \tag{10}$$

As an example of a changing reverberation time as a function of room size, consider the reverberation time of a church with a 10-meter nominal size, shown as a line with markers in Fig. 2. Also shown are the reverberation times of hypothesized churches that are 2, 4, 8, and 16 times as large, and 2, and 4 times as small. For reference, the reverberation time associated with air absorption only, α for 50% humidity and 25° C is shown in Fig. 3. Generally speaking, a doubling of the room size doubles the reverberation time. However, for large rooms and high frequencies (where the air absorption and materials absorption are somewhat comparable), a doubling of the room size increases the reverberation time by a good bit less than the factor of two seen at low frequencies or for small rooms. Note that this is the case for high frequencies in Fig. 2.

The effect of a finite air absorption may be exaggerated or suppressed by reducing or increasing—or even replacing—the air absorption characteristic shown in Fig. 3. The idea is to have different frequency bands express different reverberation times, scaling with room size. In doing so, when solving (9) for the materials absorption $\mu(\omega)$, any frequencies ω producing values less than zero should be set to zero. That is,

$$\mu(\omega) = \max\left(0,\, \frac{1}{T_{60}(L_0, \omega)} - \alpha(\omega)\right)\,. \tag{11}$$

As an example of a changing room size with a modified air absorption characteristic, Fig. 4 shows the reverberation times of Fig. 3 with a wacky air absorption.

3.3. Implementation in Common Reverberation Architectures

As described in [1], there are many commonly used reverberation algorithms. Here, we show how to resize rooms using three common methods: direct convolution, feedback delay network, and modal reverberators. The room impulse response is stretched in time and its decay rate as a function of frequency is modified to properly account for the changing relative importance of air absorption and materials absorption.

When using a convolutional reverberator (see Fig. 5), the room impulse response is resampled in time according to a room size

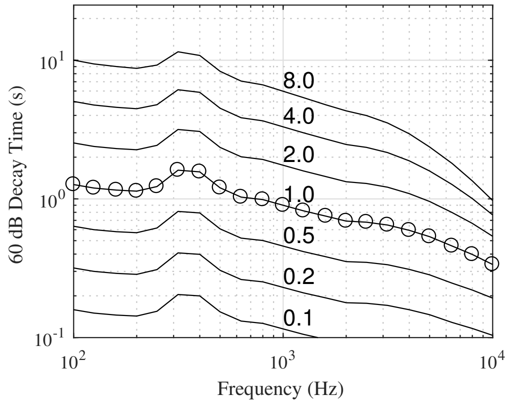

Figure 2: Example reverberation time of a small church as a function of room size taking air absorption into consideration. The markers show the measured reverberation times and the traces show the decay times when the nominal length of the church is scaled.

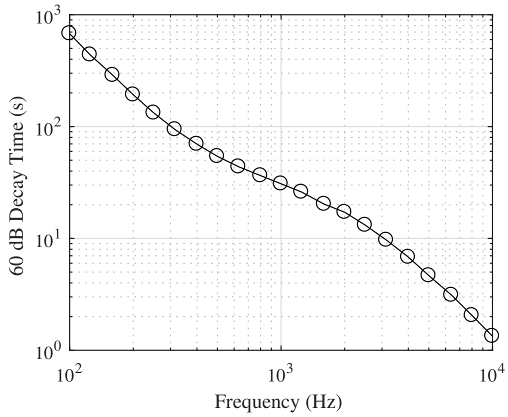

Figure 3: Reverberation time of a room with perfectly reflecting walls filled with STP air at 50% humidity.

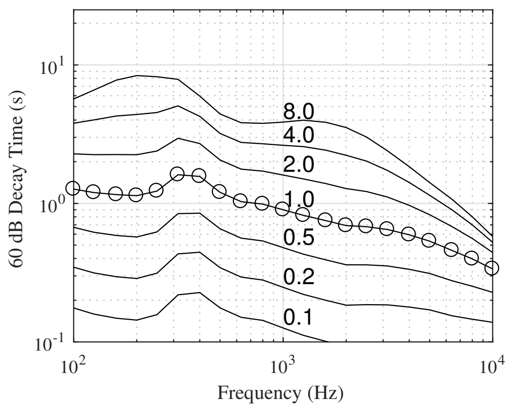

Figure 4: Example reverberation time of a small church as a function of room size, with a strange air absorption characteristic.

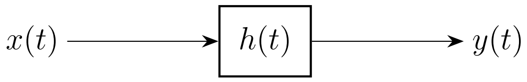

Figure 5: A convolutional reverberator showing an input signal $x(t)$, convolved with a room impulse response $h(t)$ to produce a reverberated output $y(t)$.

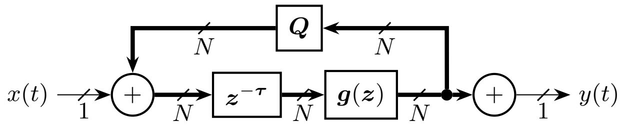

Figure 6: A feedback delay network reverberator, including a set of N delay lines z^{-T_n}, filters $g_n(z)$, $n = 1, 2, \ldots, N$, and an orthonormal mixing matrix $\boldsymbol{Q}$.

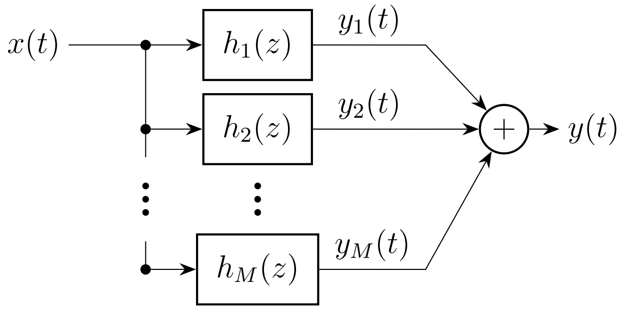

Figure 7: A modal reverberator having a parallel set of M mode filters $h_m(z)$, each characterized by a mode frequency ω_m, mode decay time τ_m, and mode amplitude γ_m.

control, and its decay rate as a function of frequency is modified. Depending on whether the room is being made larger or smaller, a high-frequency reverberant room response may be synthesized to extend the reverberation to frequencies which are warped into the audio band. A second method is described where an existing impulse response is resynthesized from its room-size-modified echo density.

For a reverberator implemented using a network of delay lines (see Fig. 6), the delay times are stretched according to the room size control, and the feedback filters are warped and scaled according to the new decay times and delay lengths. A second method is also described where the delay line lengths are not adjusted but the filters and mixing matrix are modified to account for the room resizing.

In a modal reverberator (see Fig. 7), the room size control modifies the mode frequencies and dampings. Additionally, high-frequency or low-frequency modes may need to be synthesized.

3.3.1. Convolution Reverberator

In the case of a convolutional reverberator [2, 14], the given or nominal room impulse response, $h_0(t)$, associated with a nominal room size L_0, may be resampled according to the new room size L to produce an adjusted impulse response $h_L(t)$,

$$h_L(t) = h_0\left(t \cdot \frac{L_0}{L}\right) . \tag{12}$$

As seen in Fig. 8, this adjusted impulse response may then be used to process an input signal $x(t)$ to produce a reverberated output $y(t)$ associated with the room of size L.

In the case that the room size L is smaller than the nominal room size L_0, the resampling will shorten the impulse response, thereby increasing its bandwidth. Preferably, the resampling would include the step of low-pass filtering so as to avoid aliasing if the increased bandwidth exceeds the Nyquist limit.

If L is larger than L_0, then the resampled (i.e., interpolated) impulse response will be longer than the original impulse response, and have decreased bandwidth. In this case, the adjusted impulse response may be extended to the Nyquist limit by first estimating reverberation characteristics such as decay times, equalization, echo density, and the like for that band. For instance, the decay times may be assumed to decrease in a manner typical of air absorption with increasing frequency above the original bandwidth. A trend could be fit to the decay characteristic of the nominal impulse response, and extended in frequency. Similarly, the equalization could be extrapolated to higher frequencies by noting the trend near the nominal band edge.

The mechanism of increasing the reverberation time by multiplying the reverberation impulse response by a growing exponential (as used in a number of commercially available convolutional reverberators) will generate unwanted artifacts, including a bloom in energy at the end of the impulse response [15, 16]. Resampling the impulse response as described above generally avoids this difficulty, though extending the impulse response to below the noise floor also would be of benefit. It should be noted that such a mechanism for lengthening reverberation time, even when applied to a properly extended room response, is not preferred, as the timing of temporal features, such as significant early reflections, are not appropriately modified.

As described above, the reverberation time of a room with with a modified size is roughly scaled by the relative change in size. It is affected by the different relative absorptions of air and materials, with materials absorption accounting for a greater portion of the decay in smaller rooms. The resampling of the impulse response described above has the effect of simultaneously stretching the reverberation time and compressing the associated frequency axis,

$$\tilde{T}_{60}(L, \omega) = \frac{L}{L_0} T_0\left(\omega \cdot \frac{L}{L_0}\right) , \tag{13}$$

where $\tilde{T}_{60}(L, \omega)$ is the frequency-dependent reverberation time of the stretched impulse response $h_L(t)$, and $T_0(\omega)$ is that of the given impulse response $h(t)$. For example, if a room impulse response were stretched by a factor of two, the reverberation time at 500 Hz would be twice that of the original impulse response at 1000 Hz. As a result, when the given reverberation time $T_0(\omega)$ is not relatively constant with frequency, the reverberation time produced by resampling $h(t)$ will differ from the desired one given by (10), and it is preferable to modify the reverberation time of the stretched impulse response accordingly.

As shown in Fig. 9, this may be accomplished by splitting the resampled room impulse response $h_L(t)$ into a set of frequency bands (for instance, half-octave-wide bands or ERB bands). Each band is then windowed with a growing or shrinking exponential function to give it the desired reverberation time. Then the windowed bands are summed to form a room response having the appropriate amplitude envelope as a function of frequency. This process could also be applied to the given impulse response $h(t)$

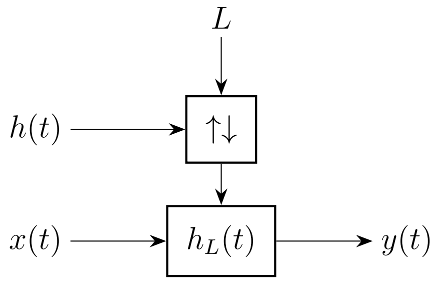

Figure 8: A convolutional reverberator showing a process operating on the impulse response $h(t)$ so that it is time-stretched (resampled) according to a room size control.

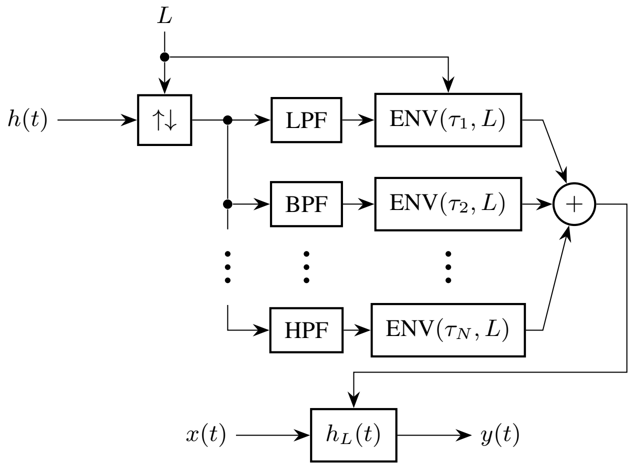

Figure 9: A convolutional reverberator in which a room size parameter modifies the resampling amount as well as modifying the frequency-dependent decay rates by $e^{-t/\boldsymbol{\tau}} \rightarrow e^{-(tL_0)/(\boldsymbol{\tau}L)}$.

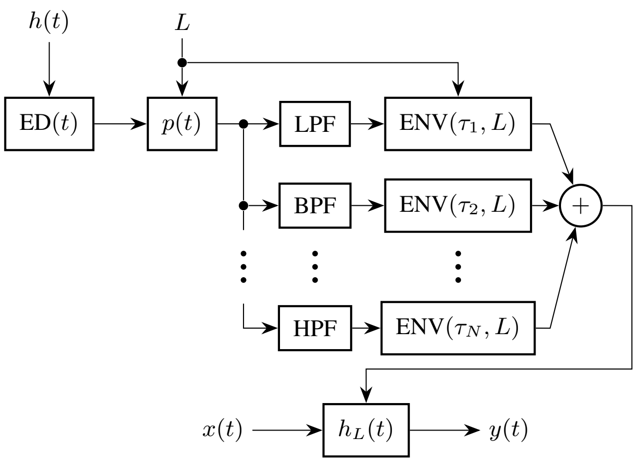

Figure 10: A convolutional reverberator in which a pulse sequence is synthesized from the echo density estimated from a desired room impulse response, split into frequency bands, and the bands windowed and summed to form an impulse response used in a convolutional reverberator. The timing of the pulses and duration of the band envelopes are adjusted according to room size.

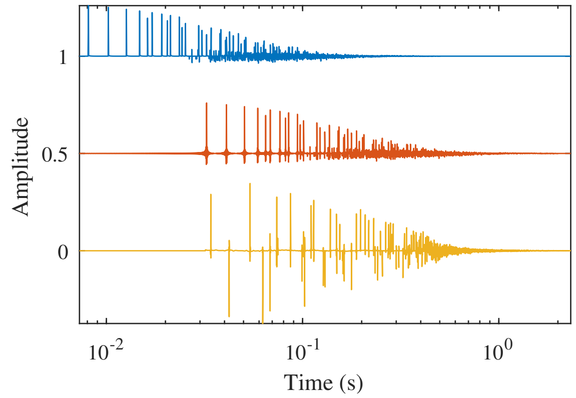

Figure 11: Time domain plots of the early reflections of impulse responses for use with a convolution reverberator showing the original IR (top), the IR stretched by a factor of 4 through resampling (middle), and resynthesized from its echo density, stretched by a factor of 4 (bottom). As a result of plotting these IRs on a logarithmic time axis, the stretched IRs appear shifted by an amount $\log_{10} 4$. Note that the ideal sinc interpolation used for the middle example filters each pulse, while the bottom example shows how resynthesizing the stretched impulse response from a statistical model does not preserve the exact echo sequence but does not have the same filtering as a result of the resampling.

before resampling, with the band windowing anticipating the reverberation time changes produced by the resampling.

It should be pointed out that while Spratt and Abel [10] describe resampling the room impulse response as similar to changing the sound speed or resizing the room, this is only true if everything about the room is resized, including materials absorption features. Here, we desire to scale the room size without modifying the materials or air properties, and it is thus preferred to correct the reverberation time produced by resampling as described above.

Finally, we note that the method described in [17] to synthesize impulse responses from balloon pop recordings may be adapted to synthesize room impulse responses at different room sizes. The process is shown in Fig. 10. Echo density is measured along the given impulse response $h(t)$, and the impulse response root energy over time (e.g., an amplitude envelope) in a set of frequency bands is estimated. A statistically independent, but perceptually identical, nominal impulse response $h_L(t)$ is then synthesized by randomly generating a set of full-bandwidth pulses, $p(t)$, according to the measured echo density, NED. (Note that in cases where the reverberation becomes quickly dense, white Gaussian noise may be used in place of the statistical pulse sequence.) This pulse sequence is then split into a set of frequency bands, and the estimated amplitude envelopes are imprinted on the pulse sequence bands before being summed to form the nominal impulse response.

To generate impulse responses of different room sizes, the same process is used, with the pulse times being scaled by the room size or with the echo density used to generate the pulse times being scaled by the inverse room size. This pulse sequence is processed as above, but with the band root energy envelopes resampled according to the room size ratio L/L_0, and preferably the envelopes modified to bring the band reverberation times in line with the desired $T_{60}(L, \omega)$ described by (10) or (9) and (6). Fig. 11

shows an impulse response resized to be twice as large through resampling and by generating a statistically similar, but stretched, pulse sequence from normalized echo density.

3.3.2. Feedback Delay Network Reverberator

Artificial reverberators are often implemented as networks of delay lines with filtering, mixing, and feedback. One such reverberator structure is the feedback delay network (FDN) [3]. The FDN reverberator employs a tapped delay line to generate the direct path and early reflections. A set of delay lines with output filtering and feedback through a unitary mixing matrix is used to produce the late-field reverberation.

Consider a FDN with N delay lines $z^{-\tau_n}, n = 1, 2, \ldots, N$ having delays τ_n and feedback filtering $g_n(z)$. The feedback filters are typically designed so that they produce similar dB attenuation per unit delay-time according to a desired decay time as a function of frequency [3]. The unitary matrix $\boldsymbol{Q}$ represents state mixing, and controls the rate of echo density increase. An identity mixing matrix $\boldsymbol{Q} = \boldsymbol{I}$ feeds each delay line to itself with no mixing between delay lines and produces a constant echo density. A Hadamard mixing matrix $\boldsymbol{Q} = \boldsymbol{H}$ generates significant mixing between delay lines, producing a rapidly increasing echo density.

To change the room size to L from a nominal L_0, the delay line lengths can be changed proportionately, as seen in Fig. 12,

$$\tau_n(L) = \frac{L}{L_0}\tau_n(L_0), \quad n = 1, 2, \ldots N\,. \tag{14}$$

Interpolated delay lines can be used to implement the desired early reflection delay times, but allpass filters are suggested to implement any fractional portion of delays used in the feedback loop so as to prevent unwanted magnitude filtering that would affect the resulting decay time.

The feedback filters $g_n(z)$ need not be modified, as the increased (or decreased) delay line lengths will result in proportionally longer (or shorter) decay times as the filters are, in effect, being applied less (or more) often. However, if desired, the feedback filters $g_n(z)$ can be modified so as to properly account for the effect of air absorption on the decay time. Additionally, note that by changing the feedback delay line lengths τ_n, the mixing matrix $\boldsymbol{Q}$ need not be modified in response to a changing room size, as the room mixing time will simply scale with the delay line lengths.

It might be the case that it is desired to leave the feedback delay lines fixed, independent of room size. In such scenarios, it is possible to change the apparent size of the room by adjusting the reverberation time and echo density profile (e.g., mixing time) by (i) modifying the feedback filters $g_n(z) \rightarrow (g_n(z))^{1/L}$, and (ii) modifying the mixing matrix $\boldsymbol{Q}$ so as to slow the state mixing, and therefore the rate of echo density increase, for larger rooms, and speed state mixing for smaller rooms as seen in Fig. 13. Fig. 14 shows the impulse response of a FDN resized by modifying the delay line lengths compared to modifying the mixing matrix and decay filters.

3.3.3. Modal Reverberator

As presented in [4], the modal reverberator implements reverberation as a parallel sum of resonant filters $h_m(t)$, each representing a room resonance or mode, and each characterized by a mode frequency ω_m, mode decay rate σ_m, and mode amplitude γ_m,

$$h(t) = \sum_{m=1}^{M} h_m(t), \tag{15}$$

where,

$$h_m(t) = \gamma_m e^{j\omega_m t - \sigma_m t}\,. \tag{16}$$

A number of options are described for implementing such filters in [4], including biquad structures, phasor filters, and heterodyning-modulation architectures.

To implement a changing room size in a modal reverberator, the mode parameters are adjusted accordingly. The mode frequen-

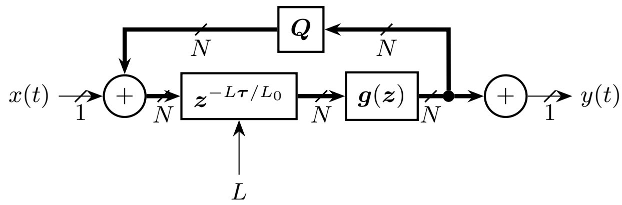

Figure 12: A delay network reverberator in which delay lengths are adjusted according to a room size control.

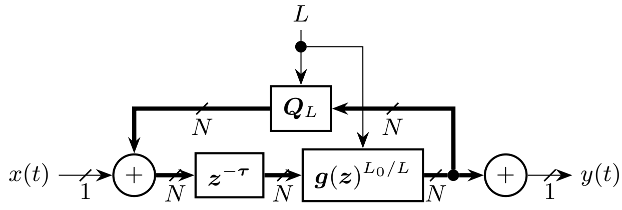

Figure 13: A delay network reverberator in which the feedback filters and mixing matrix are adjusted according to a room size control.

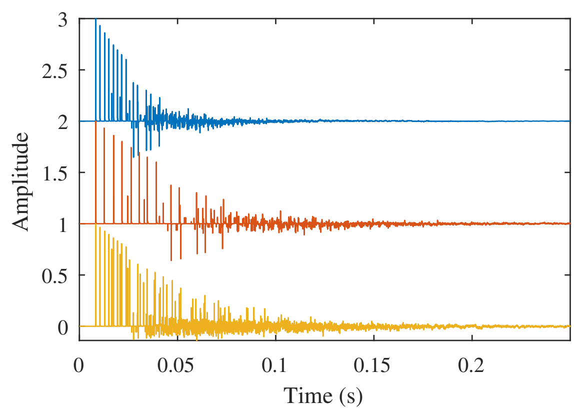

Figure 14: Time domain plots of the impulse response from a FDN showing the original IR (top), the IR stretched by a factor of 2 by modifying the delay lines (middle), and stretched by a factor of 2 by modifying the mixing matrix and decay filters (bottom). Note how the method that modifies the delay line lengths preserves the reflections exactly, just scaled by the room size parameter while modifying the mixing matrix and decay rates changes the echo pattern.

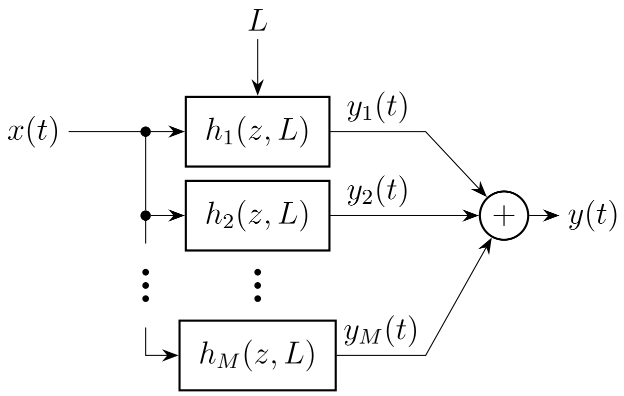

Figure 15: A modal reverberator having mode frequencies, decay times, and amplitudes modified according to a room size control.

cies would be changed in inverse proportion to the varying room size,

$$\omega_m(L) = \frac{L_0}{L}\omega_m(L_0), \quad m = 1, 2, \dots M\,, \tag{17}$$

as seen in Fig. 15. One way to understand this is to consider a closed path among a set of reflecting surfaces that creates a resonance. If the path length were twice as long, the associated travel time would be twice as long, and the frequency reduced to half its original value.

The mode decay rates would be modified according to the scaled decay times at the new mode frequencies as described above in (6),

$$T_{60}(L, \omega_m(L)) = \frac{L}{L_0\mu(\omega_m(L)) + L\alpha(\omega_m(L))}\,, \tag{18}$$

where the decay times $T_{60}(L, \omega_m(L))$ can be found by interpolation if they are not directly available. The decay rates $\sigma_m(L)$ at room size L are then

$$\sigma_m(L) = \frac{\ln 1000}{T_{60}(L, \omega_m(L))}\,. \tag{19}$$

If the room size L is made smaller than the nominal room size L_0, then the mode frequencies will be increased. Those modes with frequencies that become larger than the Nyquist limit can be eliminated, for instance, not computed or their amplitudes reduced to zero.

If the room size L is made larger than the nominal room size L_0, then the mode frequencies will be decreased. Those modes with frequencies that become smaller than the audio band lower limit, or the lower limit of what can be reproduced with the target sound reproduction system, can be eliminated. As in the case of manipulating a convolution impulse response for changing room size, an increase in room size may significantly reduce the bandwidth of the modal reverberator response, and additional bandwidth would be preferably created. This may be done by synthesizing additional high-frequency modes, for example by statistically generating additional new high-frequency modes by extrapolating the density of mode frequencies and the decay rates from the known lower-frequency modes.

As an alternative to eliminating and synthesizing modes to accommodate a changing room size, the mode frequencies ω_m can be warped within the audio band to generate new frequencies ν_m according to a first-order allpass characteristic,

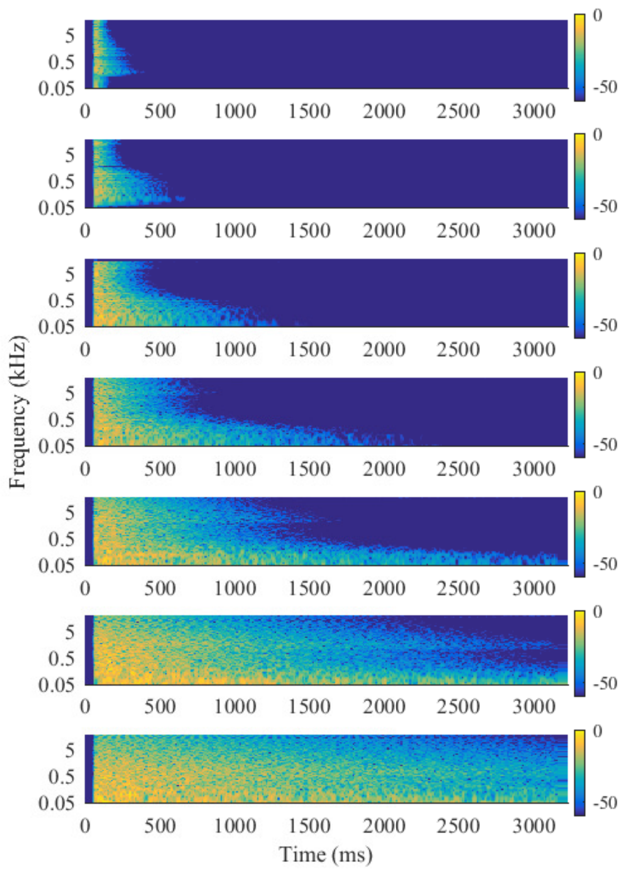

Figure 16: Spectrograms of a modal impulse response resized by factors of $1/4, 1/2, 1, 2, 4, 8$, and 16.

$$e^{-j\nu_m} = \frac{\rho + e^{-j\omega_m}}{1 + \rho e^{-j\omega_m}}\,, \tag{20}$$

that is,

$$\nu_m = j \ln\left\{\frac{\rho + e^{-j\omega_m}}{1 + \rho e^{-j\omega_m}}\right\}. \tag{21}$$

Here, the allpass parameter ρ is chosen according to the room size ratio L/L_0, and a little algebra gives

$$\rho = \frac{L - L_0}{L + L_0}\,. \tag{22}$$

Doing so will scale the low frequencies according to the desired linear characteristic

$$\nu_m(L) \approx \frac{L_0}{L}\omega_m(L_0), \quad |\omega_m| \ll 1\,, \tag{23}$$

with the high frequencies being warped to map the band edge ω onto the band edge ν.

Note that if it is desired to retain the original reverberation equalization, the mode amplitudes can be adjusted with room size to account for the changing equalization resulting from a changing modal density. Where the modal density is increased, the mode energy (the square of the mode magnitude) is proportionally increased. Fig. 16 shows spectrograms of the impulse response corresponding to a modal reverberator resized by various scale factors.

Finally, the circumstance in which only aspects of the room were made larger or smaller—say only a pair of walls being moved further apart—can be accommodated by having certain modes be unaffected or only modestly affected. Similarly, in the delay network reverberator structures above, only certain delay lines could be affected or others only modestly affected by a changing room size. This would be similar to changing the shape of the room.

3.4. Changing room size in real time

It may be desirable to modify the size of the room in real time. All three of the models presented here may experience undesirable pitch gliding artifacts if one were to modify the filter parameters in real time. Instead, it would be better to run multiple reverberators in parallel and cross-fade between them. In some situations, it may be beneficial to stretch the decay rates without modifying the modal frequencies to make the transitions across room size more smooth even though this is less physically accurate.

4. CONCLUSION

Here we have shown how a room size parameter can be introduced to scale the size of a virtual room in convolution, delay network, and modal reverberation algorithms. If a room is resized, the modal frequencies will be proportionally raised or lowered because of the scaling of the geometry of the space. Because resizing the room changes the surface area and volume, we must adapt the frequency dependent delay rates to account for these changes. Furthermore, we must also adapt the filtering to account for the fact that the material properties should remain unchanged. We do this by decoupling the modal frequencies and decay rates. There are clear trade offs in the complexity and sound of our various solutions, but these methods allow one to take an existing reverberant characteristic and stretch or shrink the size of the room with a physically informed method.

5. REFERENCES

[1] Vesa Välimäki, Julian D. Parker, Lauri Savioja, Julius O. Smith, and Jonathan S. Abel, "Fifty years of artificial reverberation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 5, pp. 1421–48, 2012.

[2] Guillermo Garcia, "Optimal filter partition for efficient convolution with short input/output delay," in *Proceedings of the 113th Audio Engineering Society Convention*, 2002.

[3] Jean-Marc Jot and Antoine Chaigne, "Digital delay networks for designing artificial reverberators," in *Proceedings of the 90th Audio Engineering Society Convention*, 1991.

[4] Jonathan S. Abel, Sean Coffin, and Kyle Spratt, "A modal architecture for artificial reverberation with application to room acoustics modeling," in *Proceedings of the 137th Audio Engineering Society Convention*, 2014.

[5] Bryan Pardo Zafar Rafii, "Learning to control a reverberator using subjective perceptual descriptors," in *Proceedings of the 10th International Society for Music Information Retrieval Conference*, 2009.

[6] Emmanouil Theofanis Chourdakis and Joshua D. Reiss, "Automatic control of a digital reverberation effect using hybrid models," in *Proceedings of the 60th International Audio Engineering Society Conference: DREAMS (Dereverberation and Reverberation of Audio, Music, and Speech)*, 2016.

[7] Jeffrey Borish, "Extension of the image model to arbitrary polyhedra," *Journal of the Acoustical Society of America*, vol. 75, no. 6, pp. 1827–36, 1984.

[8] Wallace Clement Sabine, *Collected papers on acoustics*, Peninsula Publishing, Los Alto, CA, 1993.

[9] Friedrich Spandöck, "Die Vorausbestimmung der Akustik eines Raumes mit hilfe von Modellversuchen," in *Proceedings of the 5th International Conference on Acoustics*, 1965, vol. 2, p. 313.

[10] Kyle Spratt and Jonathan S. Abel, "All natural room enhancement," in *Proceedings of the International Computer Music Conference*, 2009, pp. 231–4.

[11] Jonathan S. Abel and Patty Huang, "A simple, robust measure of reverberation echo density," in *Proceedings of the 121st Audio Engineering Society Convention*, 2006.

[12] American National Standards Institute, Committee S1, Acoustics, *Method for Calculation of the Absorption of Sound by the Atmosphere, ANSI S1.26-2009*, American National Standards Institute,, New York, NY, Sept. 1995.

[13] International Organization for Standardization, Committee ISO/TC 43, Acoustics, Sub-Committee SC 1, Noise, Acoustics, *Attenuation of sound during propagation outdoors-Part 1: Calculation of the absorption of sound by the atmosphere, ISO9613-1*, International Organization for Standardization, Geneva, Switzerland, 1993.

[14] William G. Gardner, "Efficient convolution without input-output delay," *Journal of the Audio Engineering Society*, vol. 43, no. 3, pp. 127–136, 1995.

[15] Jonathan S. Abel and Nicholas J. Bryan, "Methods for extending room impulse responses beyond their noise floor," in *Proceedings of the 129th Audio Engineering Society Convention*, 2010.

[16] Elliot K. Canfield-Dafilou and Jonathan S. Abel, "On restoring prematurely truncated sine sweep room impulse response measurements," in *Proceedings of the 20th International Conference on Digital Audio Effects*, 2017.

[17] Jonathan S. Abel, Nicholas J. Bryan, Patty P. Huang, Miriam Kolar, and Bissera V. Pentcheva, "Estimating room impulse responses from recorded balloon pops," in *Proceedings of the 129th Audio Engineering Society Convention*, 2010.