Grouped Feedback Delay Networks for Modeling of Coupled Spaces

ORCHISAMA DAS, *AES Student Member* **AND JONATHAN S. ABEL,** *AES Fellow*
(orchi@ccrma.stanford.edu) (abel@ccrma.stanford.edu)

Center for Computer Research in Music and Acoustics, Stanford University, USA.

Delay Network reverberators are an efficient tool for synthesizing reverberation. We propose a novel architecture, called the *Grouped Feedback Delay Network* (GFDN) reverberator, with groups of delay lines sharing different target decay rates, and use it to simulate coupled room acoustics. Coupled spaces are common in apartments, concert halls and churches where two or more volumes with different reverberation characteristics are linked via an aperture. The difference in decay times (T_{60}s) of the coupled spaces leads to unique phenomena, such as multi-stage decay. Here, the GFDN is used to simulate coupled spaces with groups of delay line filters representing the T_{60}s of the coupled rooms. A parameterized, orthonormal mixing matrix is presented that provides control over the mixing times of the rooms and the amount of coupling between the rooms. As an example application, we measure a coupled bedroom and bathroom system separated by a door in an apartment, and use the GFDN to synthesize the late field for different openings of the door separating the two rooms, thereby varying coupling between the rooms.

0 INTRODUCTION

The study of reverberation, or how sound travels in an enclosed space, is complex and multi-disciplinary [1]. The history of creating artificial reverberation with digital signal processing techniques has been studied by Valimaki et al. in [2]. The Feedback Delay Network (FDN) is one such artificial reverberator. Feedback delay networks are efficient IIR structures for synthesizing room impulse responses (RIRs).

RIRs consist of a set of sparse early reflections which increase in density over time, building toward late reverberation where the impulse density is high and statistically Gaussian. Feedback delay networks are composed of delay lines in parallel, which are connected through a feedback matrix (or mixing matrix), which is unitary to conserve system energy [3]. Jot proposed adding decay filters to the delay lines to yield a desired frequency dependent T_{60} [4, 5]. Since then, FDNs have become one of the most popular structures for synthesizing reverberation due to the relative efficiency of the approach. Recent research on FDNs has focused on mixing matrix design to increase echo density [6], modal analysis [7, 8], time-varying FDNs [9], scattering FDNs [10], and reverberation time control by accurate design of the decay filters [11, 12].

While there has been research on designing FDNs to match a measured impulse response [12], the modeling of coupled spaces with delay network reverberators lacks investigation. Recently, the modeling of coupled rooms with scattering delay networks has been studied in [13]. Coupled spaces, where two volumes with different reverberation characteristics are linked by an aperture, are ubiquitous in the real world. They are found in concert halls, opera halls, and churches where columns, arches, domes, etc., divide the space into two or more subspaces with different absorption properties [14]. As a result of coupling, such a system exhibits non-exponential, multi-stage decay [15].

In this paper, we use a delay network architecture we recently proposed called the *Grouped Feedback Delay Network* (GFDN) [16] for modeling coupled spaces. In a GFDN, groups of delay lines have the same target T_{60} response associated with them, compared to traditional FDNs, in which all delay lines share the same decay characteristics. The interaction among the different delay line groups is controlled by a block mixing matrix. In a coupled room system, where two spaces have different decay times, two sets of delay lines are used in the GFDN to model the decay characteristics of the rooms. The diagonal sub-matrices of the block mixing matrix control the mixing in the individual rooms, whereas the off-diagonal sub-matrices control the coupling between the two rooms.

A multichannel delay-line feedback network to control direction-dependent energy decay, known as the Directional Feedback Delay Network (DFDN) was proposed in [17]. Similar to the GFDN, in a DFDN, $(L+1)^2$ delay lines

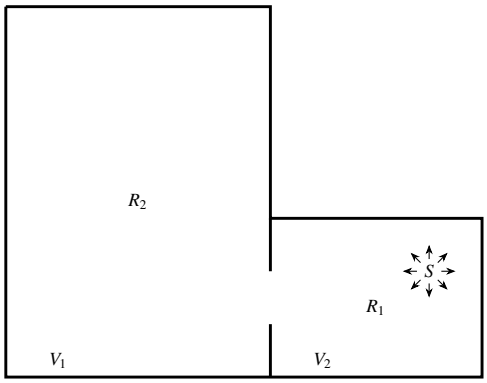


Fig. 1: Coupled rooms

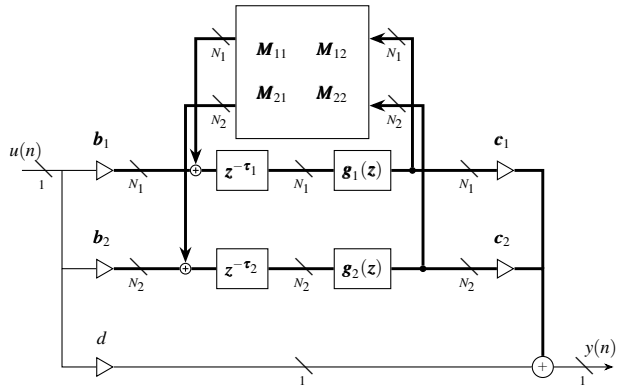


Fig. 2: GFDN block diagram.

are grouped to have the same length and a common gain, where L is the ambisonics order. The delay line groups are then fed into a linear transformation matrix, which controls the direction-dependent T_{60}. In [18], DFDNs were extended to include direction as well as frequency dependent energy decay. In contrast, the GFDN has been used to model coupled rooms, and rooms having walls and objects with different absorption characteristics [16].

In Section 1, we discuss coupled room acoustics and the phenomenon of multi-stage decay. In Section 2, we introduce the structure of the GFDN, and discuss methods for designing a parameterized, orthornormal mixing matrix which controls the coupling between two rooms with a *coupling coefficient* in Section 3. In Section 4, we use the GFDN to simulate the impulse response of a larger, more reverberant room coupled with a smaller, less reverberant room, such as box seating in an opera hall. The T_{60} response of the two rooms are modeled by simple first-order low shelf filters. We study the effect of coupling on the echo density profile of the GFDN [19]. In Section 5, we study measurements of a coupled bed-bath system (separated by a door) in an apartment. We fit T_{60} filters to individual room responses when the door is closed. We design FDNs to model the individual room impulse responses with the appropriate T_{60} filter and find an optimum mixing matrix that best matches the echo density profile of the measured impulse response. Finally, we couple the individual FDNs into a GFDN and evaluate the results for different values of the *coupling coefficient*, and compare them to measurements taken with different openings of the door separating the bedroom and bathroom (different amounts of coupling).

1 COUPLED ROOM ACOUSTICS

Two or more rooms can be coupled through an acoustically transparent aperture. The physics of sound propagation in coupled rooms was studied in [20]. A diffusion-equation model was applied to the study of acoustics in coupled-rooms in [21]. If the acoustic source is present in the smaller room with a shorter decay time, the sound will travel to the larger room and spill back into the first room. Such a configuration is shown in Fig. 1. The resulting impulse response will have a non-exponential decay. The first part of the decay has a steeper slope due to the short decay rate of the first room, whereas the latter part has a gentler slope representing the longer decay rate of the second room. Single-slope decays can occur when sound energy exchanges across coupling apertures between spaces is not significant or the coupled volume is not sufficiently separated, in which case the coupled spaces act as one. Similarly, multiple-slope decays, beyond double-slope decays can also occur [15].

Mathematically, the energy envelope of the Room Impulse Response (RIR) of a coupled system can be expressed as a weighted sum of two or more exponentials with different decay rates. For example, if there are two slopes with decay rates T_1 and T_2, then the energy envelope of the douple-slope decay maybe written as,

$$h_{env}(t) = \gamma_0 + \gamma_1 \exp\left(-\frac{t}{T_1}\right) + \gamma_2 \exp\left(-\frac{t}{T_2}\right). \quad (1)$$

Bayesian parameter estimation has been used to find multi-slope decay rates from measured RIRs in [15]. In this paper, we find the decay rates with constrained non-linear optimization, and update γs using weighted least squares (with more weight on the tail of the energy envelope).

Numerical evaluation of coupled room acoustics with various wave-based and geometrical room-acoustics software was done in [22]. The effect of changing coupling area on the room mode frequencies was investigated in [23]. Just-noticeable difference (JND) values for double-slope coupled-volume generated reverberation as a function of the coupling area was studied in [24]. Perceptual thresholds obtained from the study were 10% of variation of a given coupling area.

2 GROUPED FEEDBACK DELAY NETWORK ARCHITECTURE

A standard feedback delay network consists of N delay lines of length τ_i seconds $i = 1, 2, \ldots, N$, each with its associated decay filter, $g_i(z)$, connected through an $N \times N$ feedback matrix, $\boldsymbol{M}$. For a frequency dependent $T_{60}(z)$, the decay filter gains are related to the delay line length as

$$g_i(z) = 0.001 \exp\left(\frac{\tau_i}{T_{60}(z)}\right). \tag{2}$$

The same $T_{60}(z)$ is used to design the decay filters in all N delay lines. In the proposed grouped feedback delay network architecture, we use different $T_{60}(z)$ for each set of delay lines. In Fig. 2, a GFDN with two sets of delay lines are shown. For a total of N delay lines, N_1 delay lines have a decay response, $T_{60_1}(z)$, and N_2 delay lines have a decay response, $T_{60_2}(z)$, such that $N_1 + N_2 = N$. The two groups of decay filter gains, $g_1(z)$ and $g_2(z)$ are calculated according to the different $T_{60}(z)$s. The mixing matrix $\boldsymbol{M}$ is now an $N \times N$ block matrix made of the submatrices, $\boldsymbol{M}_{ij} \in \mathbb{R}^{N_i \times N_j}, i, j = 1, 2$. With $\boldsymbol{c}_i, \boldsymbol{b}_i, \boldsymbol{g}_i \in \mathbb{C}^{N_i \times 1}$ and $\boldsymbol{\tau}_i \in \mathbb{R}^{N_i \times 1}$, the transfer function of Fig. 2, $H(z)$, can be written as

$$\begin{aligned} H(z) &= \frac{Y(z)}{U(z)} \\ &= d + \begin{bmatrix}\boldsymbol{c}_1 & \boldsymbol{c}_2\end{bmatrix} \left(\begin{bmatrix} \boldsymbol{g}_1(z) z^{-\boldsymbol{\tau}_1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{g}_2(z) z^{-\boldsymbol{\tau}_2} \end{bmatrix} \right. \\ &\quad \left. \left(\boldsymbol{I} - \begin{bmatrix} \boldsymbol{g}_1(z) z^{-\boldsymbol{\tau}_1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{g}_2(z) z^{-\boldsymbol{\tau}_2} \end{bmatrix} \begin{bmatrix} \boldsymbol{M}_{11} \boldsymbol{M}_{12} \\ \boldsymbol{M}_{21} \boldsymbol{M}_{22} \end{bmatrix} \right)^{-1} \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{bmatrix} \right). \end{aligned} \tag{3}$$

3 MIXING MATRIX DESIGN

The mixing matrix determines the amount of coupling between various delay lines. This property controls the rate at which the echo density increases. A room with many objects and complex geometry will mix faster than an empty room with simple geometry. The mixing matrix can be designed to have a desired mixing time according to the method in [8], where the Kronecker product of a 2×2 rotation/reflection matrix (parameterized by an angle θ) with itself is taken $\log_2(N)$ times to give an $N \times N$ orthonormal matrix, $\boldsymbol{M}(\theta)$

$$\begin{aligned} \boldsymbol{R}(\theta) &= \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \\ \boldsymbol{M}_{N\times N}(\theta) &= \boldsymbol{R}(\theta) \otimes \boldsymbol{R}(\theta) \otimes \ldots \boldsymbol{R}(\theta). \end{aligned} \tag{4}$$

A well-diffused room with fast mixing time can be achieved by a scaled Hadamard mixing matrix ($\theta = \frac{\pi}{4}$). Similarly, a "room" with no mixing and no increase in echo density can be synthesized by an Identity mixing matrix ($\theta = 0$). The parameter θ can be chosen to give a desired mixing time, where $\theta = \frac{\pi}{4}$ yields the maximum amount of mixing and smaller positive values give less mixing.

In the GFDN, we can choose different, independent θ values for each delay line group (the diagonal submatrices $\boldsymbol{M}_{11}$ and $\boldsymbol{M}_{22}$). The off-diagonal submatrices ($\boldsymbol{M}_{12}$ and $\boldsymbol{M}_{21}$) then control how strongly coupled the groups are to each other. This gives us independent control over the intra- and inter-group mixing characteristics.

Let us consider two coupled rooms, R_1 and R_2. The diagonal submatrices that represent mixing in rooms R_1 and R_2 respectively can be characterized by two mixing angles, θ_1 and θ_2 depending on the occupancy of the rooms. The off-diagonal matrices represent the coupling between two rooms, and can be represented by matrices $\boldsymbol{R}_{12}, \boldsymbol{R}_{21}$, multiplied by a scalar, α, which represents the amount of coupling.

$$\boldsymbol{M} \propto \begin{bmatrix} \boldsymbol{M}(\theta_1) & \alpha \boldsymbol{R}_{12} \\ \alpha \boldsymbol{R}_{21} & \boldsymbol{M}(\theta_2) \end{bmatrix}. \tag{5}$$

This coupled mixing matrix is required to be orthonormal by design. Using this criteria, i.e., $\boldsymbol{M}^\top \boldsymbol{M} = \boldsymbol{I}$, we come up with the following constraints:

1. $\boldsymbol{R}_{12}$ and $\boldsymbol{R}_{21}$ need to be orthonormal.
2. $\boldsymbol{M}(\theta_1)^\top \boldsymbol{R}_{12} + \boldsymbol{R}_{21}^\top \boldsymbol{M}(\theta_2) = 0 \Rightarrow$
$$\boldsymbol{R}_{21} = -\boldsymbol{M}(\theta_2) \boldsymbol{R}_{12}^\top \boldsymbol{M}(\theta_1).$$
3. $\boldsymbol{M}$ needs to be scaled by $\frac{1}{\sqrt{1+\alpha^2}}$.

Let $\boldsymbol{R}_{12} = \boldsymbol{M}(\theta_1)^{\frac{1}{2}} \boldsymbol{M}(\theta_2)^{\frac{1}{2}} = \boldsymbol{M}(\frac{\theta_1}{2}) \boldsymbol{M}(\frac{\theta_2}{2})$. Then, $\boldsymbol{R}_{21} = -\boldsymbol{M}(\frac{\theta_2}{2}) \boldsymbol{M}(\frac{\theta_1}{2})$. Now, the orthonormal mixing matrix is

$$\boldsymbol{M} = \frac{1}{\sqrt{1+\alpha^2}} \begin{bmatrix} \boldsymbol{M}(\theta_1) & \alpha \boldsymbol{M}(\frac{\theta_1}{2}) \boldsymbol{M}(\frac{\theta_2}{2}) \\ -\alpha \boldsymbol{M}(\frac{\theta_2}{2}) \boldsymbol{M}(\frac{\theta_1}{2}) & \boldsymbol{M}(\theta_2) \end{bmatrix}. \tag{6}$$

Let $1/\sqrt{1+\alpha^2} = \cos\phi$ and $\alpha/\sqrt{1+\alpha^2} = \sin\phi$, then our mixing matrix is characterized by a *coupling angle*, $\phi \in [0, \frac{\pi}{4}]$ radians. When $\phi = 0$, we get minimum coupling (diagonal $\boldsymbol{M}$), and when $\phi = \frac{\pi}{4}$, we get maximum coupling between the two rooms. The final parameterized coupled mixing matrix is

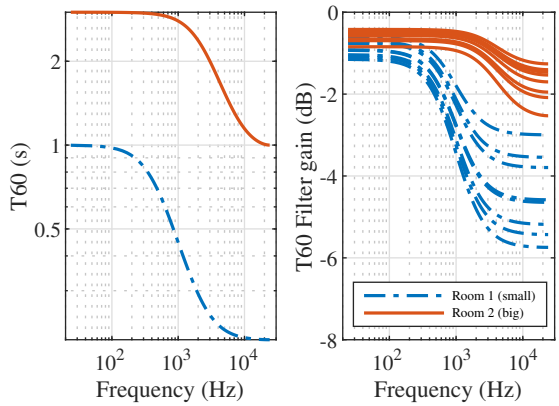
$$\boldsymbol{M}(\theta_1, \theta_2, \phi) = \begin{bmatrix} \boldsymbol{M}(\theta_1) \cos\phi & \boldsymbol{M}(\frac{\theta_1}{2}) \boldsymbol{M}(\frac{\theta_2}{2}) \sin\phi \\ -\boldsymbol{M}(\frac{\theta_2}{2}) \boldsymbol{M}(\frac{\theta_1}{2}) \sin\phi & \boldsymbol{M}(\theta_2) \cos\phi \end{bmatrix}. \tag{7}$$

In equation (7), the block mixing matrix is scaled by a 2×2 rotation matrix. We can also scale it by a Householder reflection matrix instead,

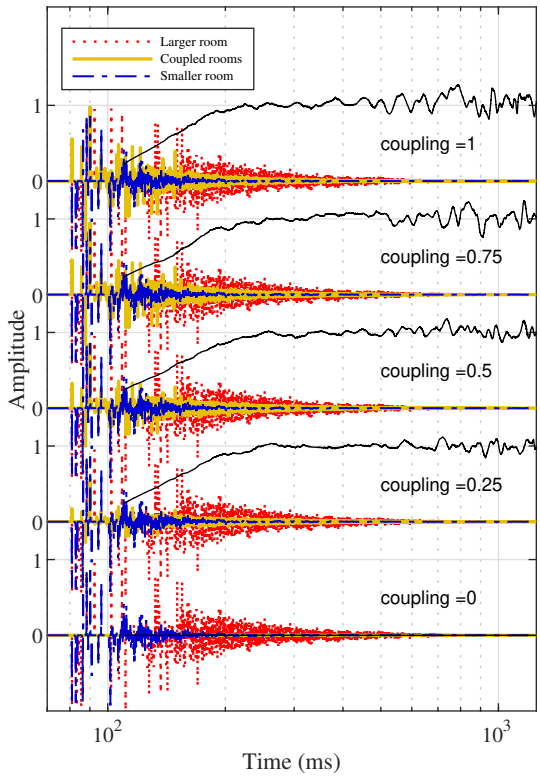
$$\begin{aligned} \boldsymbol{H} &= \boldsymbol{I} - 2\boldsymbol{u}\boldsymbol{u}^\top \\ \boldsymbol{u} &= \begin{bmatrix}\cos\phi & -\sin\phi\end{bmatrix}^\top \\ \boldsymbol{H} &= \begin{bmatrix} -\cos 2\phi & \sin 2\phi \\ \sin 2\phi & \cos 2\phi \end{bmatrix} \end{aligned} \tag{8}$$

Such matrices are typically used to model scattering junctions [25]. In case of coupled rooms, it describes the exchange of energy between the rooms.

The suggested technique can be applied recursively to model any arbitrary number of coupled rooms having multi-slope decay with many delay line sub-groups. Each sub-group will have an associated mixing matrix with an appropriate mixing structure that will constitute the diagonal block of the overall block mixing matrix. Each off-diagonal sub-matrix will have an appropriate scaling to reflect space-to-space coupling. This overall block mixing matrix can be made orthonormal by choosing an orthonormal matrix of coupling coefficients (proof in Appendix). An arbitrary coupling matrix with desired room-to-room



(a) Desired T_{60} response (left). Delay line T_{60} filters (right).



(b) Impulse responses for different coupling coefficients. Normalized Echo Density (NED) in black.

Fig. 3: Coupled Rooms modeled with GFDN

coupling coefficients can be selected, and converted to its closest orthonormal form with the sign-agnostic Procrustes method suggested in [26].

4 SYNTHESIZED EXAMPLE

To demonstrate an example[1], we designed a 16 delay line GFDN, with 8 delay lines each representing the smaller and larger room. The source is placed in the less reverberant room, R_1, and listener is placed in the more

[1]All sound examples are available at `https://ccrma.stanford.edu/~orchi/FDN/GFDN/coupled_rooms.html`

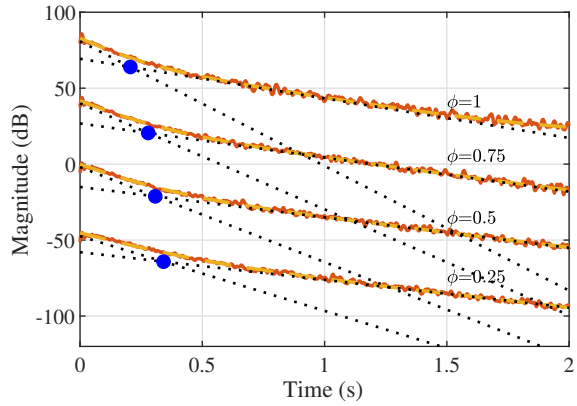


Fig. 4: Two-stage decay in coupled GFDN impulse response for varying values of ϕ. Red line indicates energy envelope, yellow line is the curve fit and the black dotted lines are the 2-stage decay fits. The blue dot is the *turning point*.

reverberant room, R_2. These source and listener locations are determined by the $\boldsymbol{b_1}, \boldsymbol{b_2}$ and $\boldsymbol{c_1}, \boldsymbol{c_2}$ coefficients respectively. The T_{60} filters of the two rooms, shown in Fig. 3a, are first order low shelf filters parameterized by the DC and Nyquist gains and transition frequency. The smaller room, R_1, (in blue) has a shorter decay time, with $T_{60}(0) = 1$ s, $T_{60}(\infty) = 0.2$ s and $f_T = 1$ kHz. The larger room, R_2 (in red) has $T_{60}(0) = 3$ s, $T_{60}(\infty) = 1$ s and $f_T = 4$ kHz. The decay filters, $\boldsymbol{g}_1(z), \boldsymbol{g}_2(z)$ calculated according to (2) are shown in Fig. 3a.

The impulse responses of the coupled GFDN as a function of linearly spaced coupling angles (normalized by $\frac{\pi}{4}$) are shown in Fig. 3b. As expected, when $\phi = 0$, the rooms are decoupled and the GFDN gives zero output. Increasing ϕ increases diffusion between the two rooms, giving denser reverb. The normalized echo density (NED) [19], which is a perceptual measure of reverberation that helps quantify when early decay switches to late reverb, is plotted in black. The NED plots show that denser reverberation is achieved more quickly as ϕ increases. The decay profile of the smaller room dominates as ϕ increases.

It can be noted from the figure, as well as the sound examples, that the early reflections are too sparsely distributed to be an accurate representation of the RIR of the rooms we are trying to model. While all RIRs are composed of a direct path and early reflections followed by a dense Gaussian distributed late field, FDNs are typically used to synthesize the late field only, whereas the early reflections are synthesized by geometric methods or convolution. Fade-in control of the late field synthesized by FDNs, to the early part has been studied in [27]. Some guidelines for choosing delay line lengths and attenuating filters to capture the early response have been proposed in [28].

Two-stage decay plots for various amounts of coupling are shown in Fig. 4. Both decay rates, T_1 and T_2, decrease as the coupling angle increases, and the resulting RIR decays faster. This can be corroborated by listening to the sound examples. The *turning point* [21], which is the time

at which the two slopes intersect, shifts to the left as coupling angle increases.

5 FITTING TO MEASUREMENTS

Measurements of a coupled bathroom and bedroom in the author's apartment were taken with an omnidirectional condensor microphone (*Beyerdynamic MM1*) and loudspeaker (*PreSonus Eris E5*) with a *Behringer UM2* interface. The volume of the bedroom is 28.77 m^3, and that of the bathroom is 9.36 m^3. Although the bathroom is smaller in size, it is more reverberant because of tiled walls. The loudspeaker was kept 40.6 cm from the door separating the two rooms, facing the bedroom, at a height of 38 cm. The microphone was kept in the bedroom at a distance of 94 cm from the speaker, and at a height of 66 cm. A 10 s long allpass sine sweep [29] was played from $20 - 20000$ Hz at a sampling rate of 48 kHz. The bathroom door was initially closed, and then opened in intervals. RIR measurements were taken for different openings of the door aperture. Two more sets of measurements were done - one with the speaker in the bedroom and microphone in the bathroom, and another with both the speaker and the microphone in the bathroom. The axes of the speaker and microphone were not aligned.

5.1 Designing FDN to match individual room response

We designed FDNs to match the measured RIRs of the bedroom and bathroom individually (when the door is closed). The spectrograms of the measured RIRs are shown in Fig. 5a, 5b. The FDN has 16 delay lines with mutually prime lengths, ranging from 5 ms to 10 ms. The T_{60} response in one-fifth octave bands was calculated from the measured RIRs by band-pass filtering with a second order Butterworth filter, and fitting straight lines to the energy envelope. For the bedroom, T_{60} values below 200 Hz, and for the bathroom, T_{60} values below 50 Hz were discarded because of the high noise floor (poor signal to noise ratio) in lower frequencies. To get the frequency response, the remaining estimated T_{60} values were interpolated with a cubic spline to a linear frequency axis of 2048 bins, ranging from 0 to $\frac{f_s}{2}$ Hz, where f_s is the sampling rate in Hz. From 200 Hz to DC, the response was made to roll-off at 10 ms/octave.

We converted the interpolated T_{60} frequency response to a minimum-phase impulse response, and used frequency-warped Prony's method [30] to find the T_{60} filter coefficients. The warping factor was -0.85 and the Prony filter order was 12. The measured, interpolated, Prony fit and warped Prony fit T_{60}s for the two rooms is shown in Fig. 6.

To find the optimum mixing matrix, we synthesized the RIR for a grid of mixing angle values, $\theta \in [0, \frac{\pi}{4}]$. For each mixing angle in the search grid, we found the error between the NEDs of the late fields of the synthesized and measured RIR. The mixing angle that produced the least error was selected as the optimum. This is because the NED is a good perceptual measure of the diffused field [31]. The mixing matrix was designed according to (4). The values of the optimum mixing angle for the bedroom and bathroom are $0.8\frac{\pi}{4}$ and $0.6\frac{\pi}{4}$ respectively. The resulting synthesized RIRs of the two rooms are shown in Fig. 5c, 5d.

5.2 Designing GFDN to match coupled room response

Once we have tuned the FDN parameters for the individual rooms, we couple them with a 32 delay line GFDN. The first set of 16 delay lines with a shorter T_{60} response (Fig. 6a) model the bedroom, and the second set of 16 delay lines with the longer T_{60} response (Fig. 6b), model the bathroom. The mixing angles calculated in Section 5.1 are used, with $\theta_1 = \frac{0.8\pi}{4}$ and $\theta_2 = \frac{0.6\pi}{4}$.

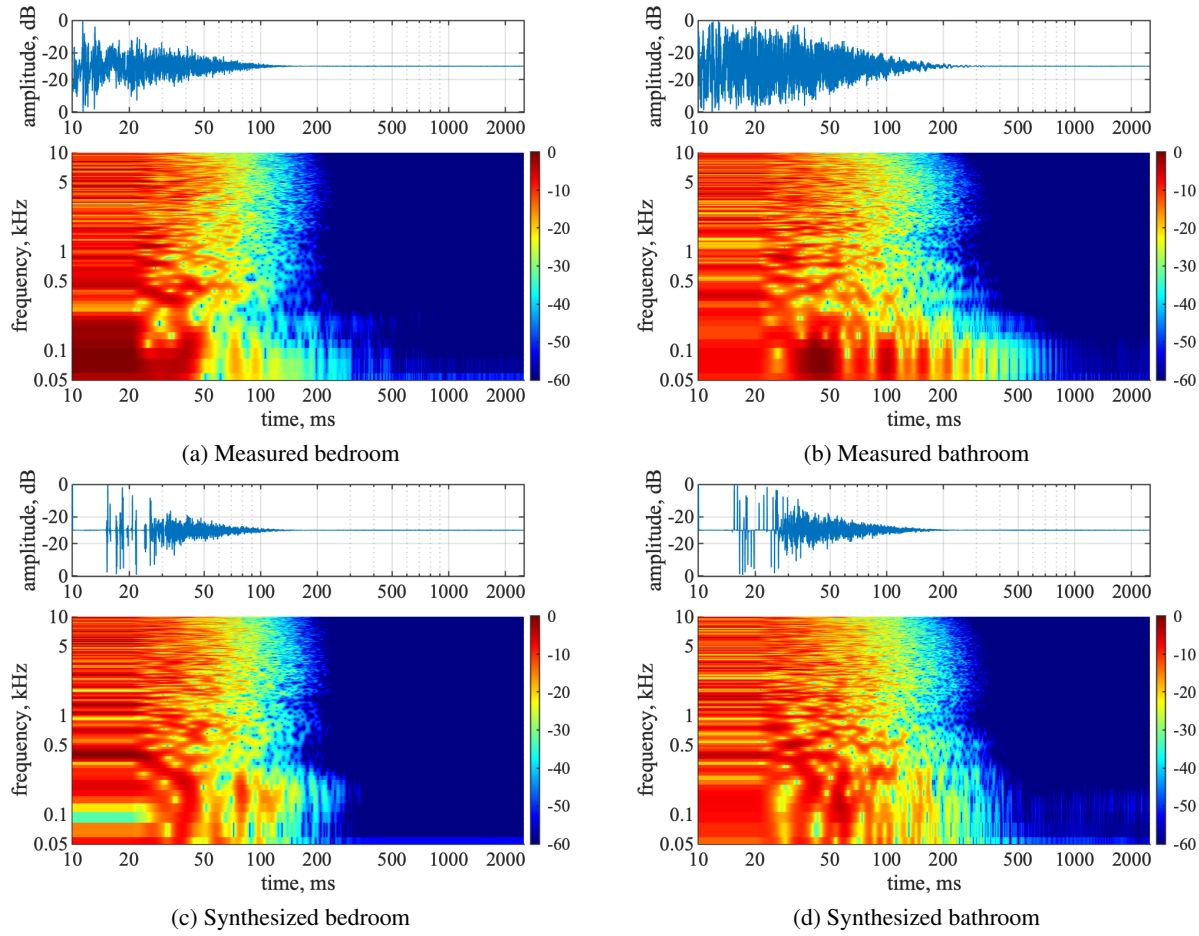
The coefficients $\boldsymbol{b_1}, \boldsymbol{b_2}, \boldsymbol{c_1}, \boldsymbol{c_2}$ from (3) determine the speaker and microphone locations. These scalar driving gains are set to either zero or selected randomly based on which room the speaker and mic are located in. In the first scenario, both the speaker and microphone are virtually placed in the bedroom by setting $\boldsymbol{b_2}, \boldsymbol{c_2} = \boldsymbol{0}$. For progressively increasing areas of the door aperture, the optimum coupling angle, ϕ, is found by a linear grid search after comparing the half octave T_{60}s of the measured and synthesized RIRs, and choosing the ϕ which minimizes their squared error. The mixing matrix is calculated according to (7).

The T_{60}s of the measured RIRs for various areas of door aperture are shown in Fig. 7a. The T_{60}s of the RIRs synthesized with the GFDN with the optimum values of the normalized coupling angle, $\frac{4\phi}{\pi}$, are shown in Fig, 7b. Points below 200 Hz are discarded because of the previously mentioned high noise floor in low frequencies. In Fig. 7a, the T_{60} measured in the bedroom increases as the door opens (or, as there is more coupling between the rooms). This is expected, since the bathroom has a longer decay time than the bedroom. According to Sabine's theory [32], increase in volume accompanied by a decrease in the absorbing area leads to an increase in decay time. This behavior is mimicked by the GFDN output as coupling angle increases.

In the second scenario, the speaker is placed in the bedroom ($\boldsymbol{b_2} = \boldsymbol{0}$) and the microphone is placed in the bathroom ($\boldsymbol{c_1} = \boldsymbol{0}$). Similar T_{60} plots are shown in Fig. 8 after optimizing the coupling angle, ϕ. The behavior observed in Fig. 8a is opposite to that in Fig. 7a, i.e, the decay time decreases with more coupling. This is because the energy from the more reverberant bathroom dissipates more quickly into the bedroom as the aperture size increases. There is no output from the GFDN when $\phi = 0$ because the rooms are decoupled, whereas complete decoupling is not possible in reality even with a closed door, since it transmits some acoustic energy. Sound examples of the measured and synthesized late field for different amounts of coupling are available[2].

Two stage decay observed in the measured and synthesized RIRs for the second scenario, when the speaker is in the bedroom and the microphone is in the bathroom, for

[2]See footnote 1



(a) Measured bedroom

(b) Measured bathroom

(c) Synthesized bedroom

(d) Synthesized bathroom

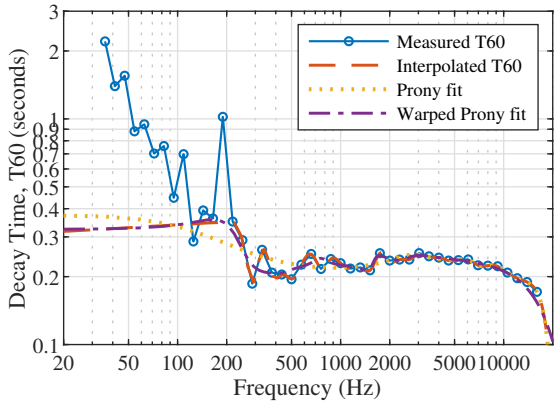
Fig. 5: Measured and synthesized RIRs (and their spectrograms) of the bedroom and bathroom when the door between them is closed (decoupled). The RIRs are plotted as $\frac{\tanh \beta h}{\tanh \beta}$ with $\beta = 3$.

various amounts of coupling, is shown in Fig. 9. The associated values of the two T_{60}s from (1) are also shown in the plots. When the coupling angle is very small, as in Fig. 9b, only single-slope decay is observed. Although the individual values of the two decay times of the measured and synthesized RIRs are not identical, some common behaviors are exhibited. In both the measured and synthesized RIRs, T_{60_1} and T_{60_2} decrease with the increase in coupling. Similar behavior was observed in Fig. 4. As the door opens (or equivalently, the coupling angle in the GFDN increases), more sound travels to the bedroom and decays faster, resulting in a shorter T_{60}. The T_{60}s associated with the measured RIRs decay more rapidly than the ones associated with the synthesized RIRs.

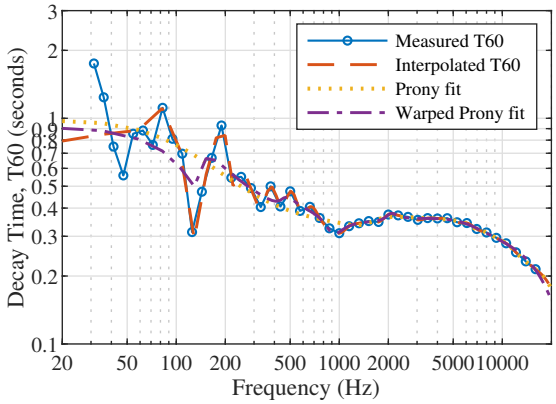
6 CONCLUSION

A novel architecture of the Feedback Delay Network, called the *Grouped Feedback Delay Network*, has been discussed and applied to simulate reverberation in coupled spaces. The GFDN has different T_{60} filters in different groups of delay lines, and a block mixing matrix that controls the inter- and intra-group mixing. The GFDN has been used to simulate coupled rooms, which are commonly found in concert halls, apartments and places of religious worship. Groups of delay lines in the GFDN model multiple rooms with different T_{60} responses. An orthonormal mixing matrix is designed with parameters that control the mixing time in individual rooms, as well as the amount of coupling between the rooms. An example has been demonstrated that studies the effect of coupling on the echo density profile and two-stage decay of the GFDN impulse response. Measurements of a coupled bed-bath system have been taken for varying aperture sizes, and strategies have been discussed for modeling it with the GFDN.

The GFDN is an efficient way of rendering reverberation in coupled rooms for virtual reality applications. It can also be used to synthesize the late field in augmented reality and architectural acoustics applications, when measurements are available. Future work includes designing a listening test to evaluate the results of the measured and synthesized late-fields. Similar perceptual studies have been performed in [33] to evaluate the naturalness of synthesized binaural RIRs. Moreover, frequency-dependent coupling to model diffraction through the aperture connecting the two rooms can be explored.



(a) Bedroom



(b) Bathroom

Fig. 6: T_{60} responses (measured and fit) of the two individual rooms when door is closed.

7 ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their constructive feedback, which has helped to improve the quality of the manuscript. We would also like to thank Elliot K. Canfield-Dafilou for his contributions to the previously published DAFx paper on this topic.

APPENDIX

If there are N rooms coupled with each other, with coupling coefficients $\alpha_{ij},\ i,j \in 1,\ldots,N$, and each room has its own mixing angle $\theta_i\ i \in 1,\ldots,N$, then we have a block mixing matrix whose diagonals are the individual room mixing matrices without any coupling, and the off-diagonal elements represent room-to-room coupling. Let us define the block mixing matrix,

$$\boldsymbol{F}(\boldsymbol{M}_1,\ldots\boldsymbol{M}_N) = \begin{cases} \boldsymbol{M}_i,\ i=j \\ \alpha_{ij}\boldsymbol{M}_i^{\frac{1}{2}}\boldsymbol{M}_j^{\frac{1}{2}},\ i\neq j \end{cases} \tag{1}$$

Here, $\boldsymbol{M}_i = \boldsymbol{M}(\theta_i)$ and $\boldsymbol{M}_i^{\frac{1}{2}} = \boldsymbol{M}(\frac{\theta_i}{2})$, are both individually orthonormal. To preserve energy in the system, we want, $\boldsymbol{F}^\top\boldsymbol{F} = \boldsymbol{I}$.

$$(\boldsymbol{F}^\top\boldsymbol{F})_{ij} = \begin{cases} \boldsymbol{M}_i^\top\boldsymbol{M}_i + \sum_{k=1,k\neq i}^{N}\alpha_{ik}^2,\ i=j \\ \boldsymbol{M}_i^{\frac{\top}{2}}\boldsymbol{M}_j^{\frac{1}{2}}(\alpha_{ij}+\alpha_{ji}+\sum_{k=1,k\neq i,j}^{N}\alpha_{ki}\alpha_{kj}),\ i\neq j \end{cases} \tag{2}$$

From these equations, we get the following conditions for $\boldsymbol{F}$ to be orthonormal,

1. $\boldsymbol{F}$ needs to be scaled by $\frac{1}{\sqrt{\prod_{i=1}^{N}(1+\sum_{k=1,k\neq i}^{N}\alpha_{ik}^2)}}$.
2. $\alpha_{ij}+\alpha_{ji}+\sum_{k=1,k\neq i,j}^{N}\alpha_{ki}\alpha_{kj} = 0,\ i\neq j$

Let $\boldsymbol{\Phi} \in \mathbb{R}^{N\times N}$ be the matrix of coupling coefficients,

$$\boldsymbol{\Phi} = \begin{bmatrix} 1 & \alpha_{12} & \cdots & \alpha_{1N} \\ \alpha_{21} & 1 & \cdots & \alpha_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{N1} & \alpha_{N2} & \cdots & 1 \end{bmatrix} \tag{3}$$

$$(\boldsymbol{\Phi}^\top\boldsymbol{\Phi})_{ij} = \begin{cases} 1+\sum_{k=1,k\neq i}^{N}\alpha_{ik}^2,\ i=j \\ \alpha_{ij}+\alpha_{ji}+\sum_{k=1,k\neq i,j}^{N}\alpha_{ki}\alpha_{kj},\ i\neq j \end{cases}$$

It is evident that for $\boldsymbol{F}$ to be orthonormal, the matrix of coupling coefficients $\boldsymbol{\Phi}$ needs to be orthonormal. To do so, we have to scale $\boldsymbol{\Phi}$ such that,

$$\tilde{\boldsymbol{\Phi}} = \boldsymbol{\Phi}/\sqrt{\prod_{i=1}^{N}(\sum_{j=1}^{N}\phi_{ij}^2)}.$$

For meaningful coupling, $\boldsymbol{\Phi}$ should also have antisymmetric off-diagonal elements, i.e., $\alpha_{ij} = -\alpha_{ji}$, since the coupling coefficient between rooms 1 and 2 should be the same as that between rooms 2 and 1.

A.1 REFERENCES

[1] B. A. Blesser, "An interdisciplinary synthesis of reverberation viewpoints," *J. Audio Eng. Soc.*, vol. 49, no. 10, pp. 867–903 (2001).

[2] V. Valimaki, J. D. Parker, L. Savioja, J. O. Smith, J. S. Abel, "Fifty years of artificial reverberation," *IEEE Trans. on Audio, Speech, and Language Process.*, vol. 20, no. 5, pp. 1421–1448 (2012).

[3] M. A. Gerzon, "Unitary (energy-preserving) multichannel networks with feedback," *Electron. Lett.*, vol. 12, no. 11, pp. 278–279 (1976).

[4] J. M. Jot, A. Chaigne, "Digital delay networks for designing artificial reverberators," presented at the *Audio Eng. Soc. Conv. 90* (1991).

[5] J. M. Jot, "An analysis/synthesis approach to real-time artificial reverberation," presented at the *IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 2, pp. 221–224 (1992).

[6] S. J. Schlecht, E. A. P. Habets, "Dense Reverberation with Delay Feedback Matrices," presented at the *IEEE Workshop Appl. Signal Process. Audio Acoust.*, pp. 150–154 (2019).

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

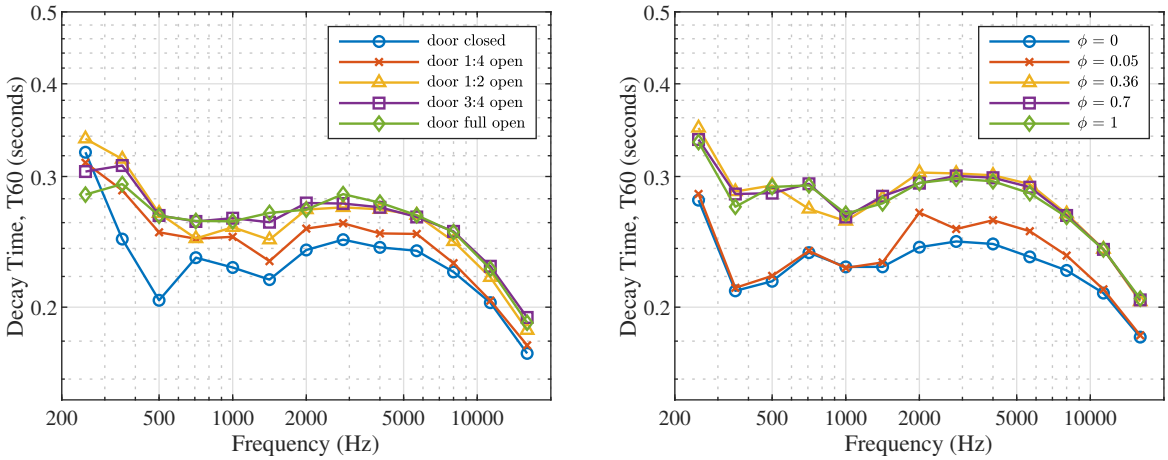


Fig. 7: Left - measured RIR T_{60}, Right - GFDN RIR T_{60} when both speaker and microphone are in the bedroom.

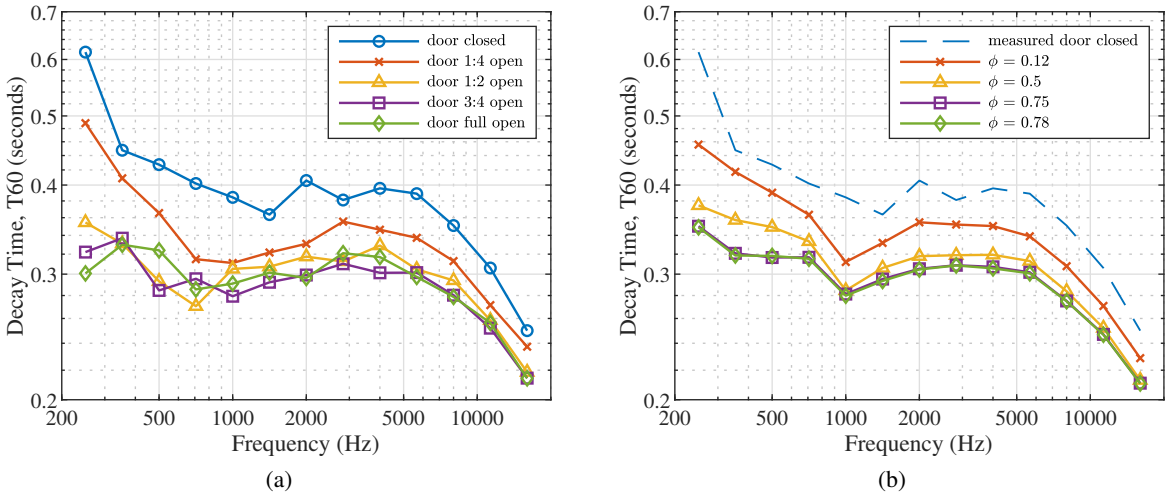


Fig. 8: Left - measured RIR T_{60}, Right - GFDN RIR T_{60} when the speaker is in the bedroom (less reverberant) and microphone is in the bathroom (more reverberant).

[8] O. Das, E. K. Canfield-Dafilou, J. S. Abel, "On The Behavior of Delay Network Reverberator Modes," presented at the *IEEE Workshop Appl. Signal Process. Audio Acoust.*, pp. 50–54 (2019).

[9] S. J. Schlecht, E. A. P. Habets, "Time-varying feedback matrices in feedback delay networks and their application in artificial reverberation," *J. Acoust. Soc. Amer.*, vol. 138, no. 3, pp. 1389–1398 (2015).

[10] S. J. Schlecht, E. A. Habets, "Scattering in feedback delay networks," *IEEE/ACM Trans. on Audio, Speech, and Language Process.*, vol. 28, pp. 1915–1924 (2020).

[11] S. J. Schlecht, E. A. P. Habets, "Accurate reverberation time control in feedback delay networks," *Int. Conf. Digit. Audio Effects*, pp. 337–344 (2017).

[12] K. Prawda, V. Välimäki, S. J. Schlecht, "Improved Reverberation Time Control for Feedback Delay Networks," presented at the *Int. Conf. Digit. Audio Effects* (2019).

[13] T. B. Atalay, *Scattering Delay networks with Aperture Size control for Simulating Coupled Room Acoustics*, Master's thesis, Middle East Technical University (2019).

[14] J. S. Anderson, M. Bratos-Anderson, "Acoustic coupling effects in St Paul's cathedral, London," *J. Sound Vib.*, vol. 236, no. 2, pp. 209–225 (2000).

[15] N. Xiang, P. Goggans, T. Jasa, P. Robinson, "Bayesian characterization of multiple-slope sound energy decays in coupled-volume systems," *J. Acoust. Soc. Amer.*, vol. 129, no. 2, pp. 741–752 (2011).

[16] O. Das, J. S. Abel, E. K. Canfield-Dafilou, "Delay Network Architectures for Room and Coupled Space Modeling," presented at the *Int. Conf. Digit. Audio Effects*, pp. 234–241 (2020).

[17] B. Alary, A. Politis, S. J. Schlecht, V. Välimäki, "Directional feedback delay network," *J. Audio Eng. Soc.*, vol. 67, no. 10, pp. 752–762 (2019).

[18] B. Alary, A. Politis, "Frequency-dependent directional feedback delay network," presented at the *IEEE Int.*

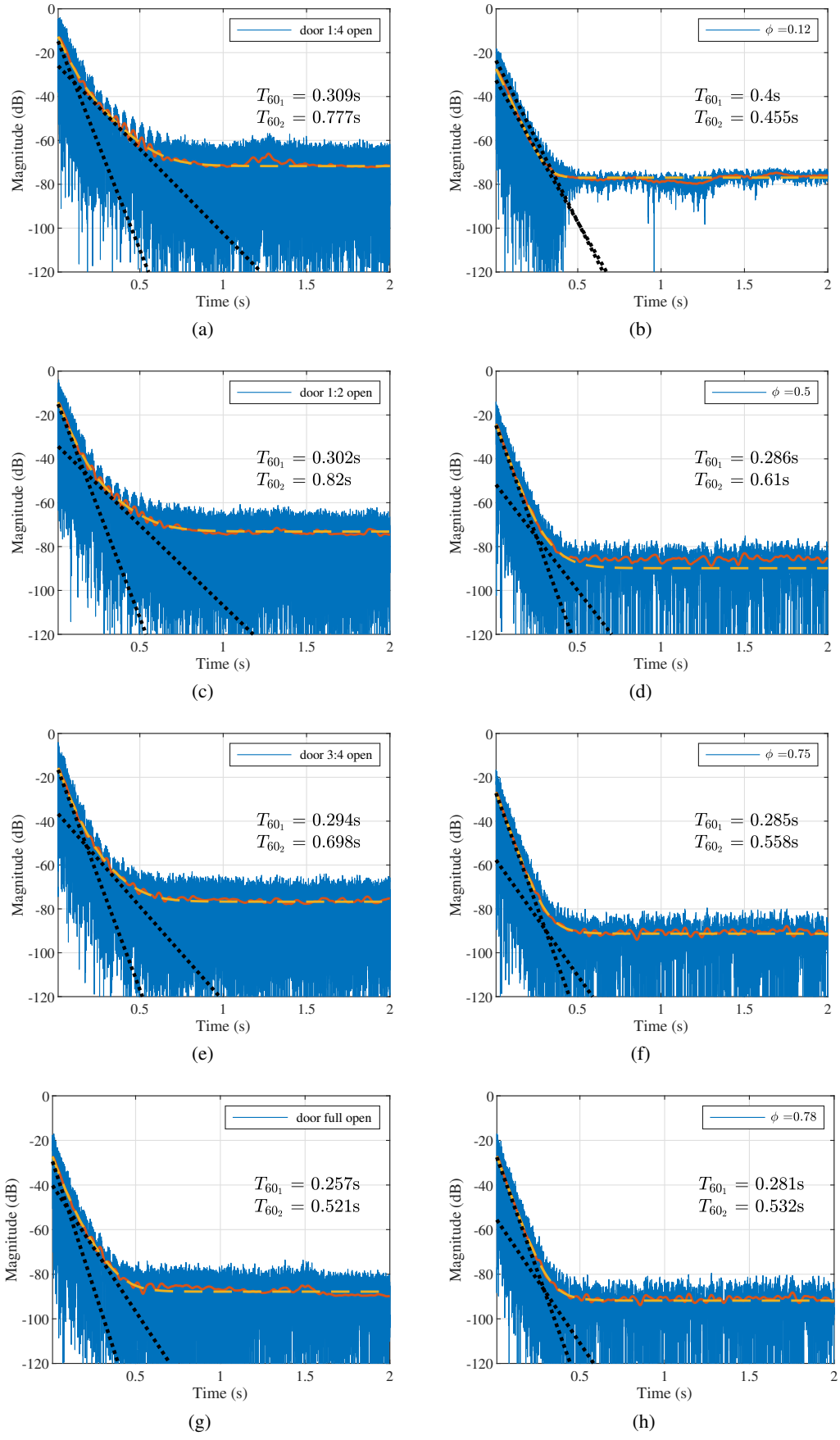


Fig. 9: Two-stage decay observed when the speaker is in the bedroom and mic is in the bathroom, for various amounts of coupling. Right - Measured, Left - GFDN Synthesized. Blue - RIR, Red - energy envelope, Yellow - fitted curve, Black - 2-stage decay fit

Conf. on Acoust., Speech, Signal Process. (ICASSP), pp. 176–180 (2020).

[19] J. S. Abel, P. Huang, "A simple, robust measure of reverberation echo density," presented at the *Audio Eng. Soc. Conv. 121* (2006).

[20] C. F. Eyring, "Reverberation time measurements in coupled rooms," *J. Acoust. Soc. Amer.*, vol. 3, no. 2A, pp. 181–206 (1931).

[21] N. Xiang, Y. Jing, A. C. Bockman, "Investigation of acoustically coupled enclosures using a diffusion-equation model," *J. Acoust. Soc. Amer.*, vol. 126, no. 3, pp. 1187–1198 (2009).

[22] A. Weber, B. Katz, "Numerical simulation round robin of a coupled volume case: Preliminary results," presented at the *Int. Cong. on Acoustics* (2019).

[23] M. Meissner, "Computer modelling of coupled spaces: variations of eigenmodes frequency due to a change in coupling area," *Archives of Acoustics*, vol. 34, no. 2, pp. 157–168 (2009).

[24] P. Luizard, B. F. Katz, C. Guastavino, "Perceptual thresholds for realistic double-slope decay reverberation in large coupled spaces," *J. Acoust. Soc. Amer.*, vol. 137, no. 1, pp. 75–84 (2015).

[25] E. De Sena, H. Hacıhabiboğlu, Z. Cvetković, J. O. Smith, "Efficient synthesis of room acoustics via scattering delay networks," *IEEE/ACM Trans. on Audio, Speech and Language Process. (TASLP)*, vol. 23, no. 9, pp. 1478–1492 (2015).

[26] S. J. Schlecht, E. A. P. Habets, "Sign-agnostic matrix design for spatial artificial reverberation with feedback delay networks," presented at the *Audio Eng. Soc. Int. Conf. Spatial Reproduction-Aesthetics Sci.* (2018).

[27] N. Meyer-Kahlen, S. J. Schlecht, T. Lokki, "Fade-In Control for Feedback Delay Networks," presented at the *Int. Conf. on Digit. Audio Effects* (2020).

[28] J. Stautner, M. Puckette, "Designing multi-channel reverberators," *Comp. Music Journal*, vol. 6, no. 1, pp. 52–65 (1982).

[29] E. K. Canfield-Dafilou, J. S. Abel, "An allpass chirp for constant signal-to-noise ratio impulse response measurement," presented at the *Audio Eng. Soc. Conv. 144* (2018).

[30] A. Härmä, M. Karjalainen, L. Savioja, V. Välimäki, U. K. Laine, J. Huopaniemi, "Frequency-warped signal processing for audio applications," *J. Audio Eng. Soc.*, vol. 48, no. 11, pp. 1011–1031 (2000).

[31] P. Huang, J. S. Abel, H. Terasawa, J. Berger, "Reverberation echo density psychoacoustics," presented at the *Audio Eng. Soc. Conv. 125* (2008).

[32] W. C. Sabine, *Collected papers on acoustics* (Peninsula Publishing, Los Alto, CA) (1993).

[33] N. Xiang, U. Trivedi, B. Xie, "Artificial enveloping reverberation for binaural auralization using reciprocal maximum-length sequences," *J. Acoust. Soc. Amer.*, vol. 145, no. 4, pp. 2691–2702 (2019).

NOMENCLATURE

FDN = Feedback Delay Network

GFDN = Grouped Feedback Delay Network
RIR = Room Impulse Response
NED = Normalized Echo Density

T_{60} = 60 dB decay time, calculated from the decay time, or time constant, T, using the relation, $T_{60} = T \ln(1000)$.

THE AUTHORS

Orchisama Das

Jonathan Abel

Orchisama Das received the B. Eng. degree in Instrumentation and Electronics Engineering from Jadavpur University, India, in 2016. She is currently a Ph.D. candidate at the Center for Computer Research in Music and Acoustics at Stanford University. Her research interests are room acoustics modeling, sound analysis-synthesis and statistical signal processing. She has been a teaching assistant for various signal processing courses at CCRMA. In 2015, Orchisama worked at the University of Calgary funded by a Mitacs Globalink fellowship. She has interned at Tesla Motors and Facebook Reality Labs. In 2019, she was a visiting researcher in the Sound Analysis-Synthesis team at IRCAM.

●

Jonathan Abel is a consulting professor at the Center for Computer Research in Music and Acoustics (CCRMA) in the Music Department at Stanford University where his research interests include audio and music applications of signal and array processing, parameter estimation, and acoustics. From 1999 to 2007, Abel was a co-founder and chief technology officer of the Grammy Award-winning Universal Audio, Inc. He was a researcher at NASA/Ames Research Center, exploring topics in room acoustics and spatial hearing on a grant through the San Jose State University Foundation. Abel was also chief scientist of Crystal River Engineering, Inc., where he developed their positional audio technology, and a lecturer in the Department of Electrical Engineering at Yale University. As an industry consultant, Abel has worked with Apple, FDNY, LSI Logic, NRL, SAIC and Sennheiser, on projects in professional audio, GPS, medical imaging, passive sonar and fire department resource allocation. He holds Ph.D. and M.S. degrees from Stanford University, and an S.B. from MIT, all in electrical engineering. Abel is a Fellow of the Audio Engineering Society.