ON THE BEHAVIOR OF DELAY NETWORK REVERBERATOR MODES

Orchisama Das, Elliot K. Canfield-Dafilou, Jonathan S. Abel

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

ABSTRACT

The mixing matrix of a Feedback Delay Network (FDN) reverberator is used to control the mixing time and echo density profile. In this work, we investigate the effect of the mixing matrix on the modes (poles) of the FDN with the goal of using this information to better design the various FDN parameters. We find the modal decomposition of delay network reverberators using a state space formulation, showing how modes of the system can be extracted by eigenvalue decomposition of the state transition matrix. These modes, and subsequently the FDN parameters, can be designed to mimic the modes in an actual room. We introduce a parameterized orthonormal mixing matrix which can be continuously varied from identity to Hadamard. We also study how continuously varying diffusion in the mixing matrix affects the damping and frequency of these modes. We observe that modes approach each other in damping and then deflect in frequency as the mixing matrix changes from identity to Hadamard. We also quantify the perceptual effect of increasing mixing by calculating the normalized echo density (NED) of the FDN impulse responses over time.

Index Terms— Feedback Delay Network, Artificial Reverberation, Modal Analysis, Normalized Echo Density, Mixing Matrix

1. INTRODUCTION

Artificial reverberation techniques aim to synthesize an impulse response using a combination of filters. An ideal artificial reverberator reproduces 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. In other words, the "echo-density" (number of echoes per second) should increase quadratically in time and is a good psychoacoustic measure of perceived reverberation. Another characteristic of an ideal reverberator is the frequency response—the number of modes should increase with frequency squared, which means high frequency modes are densely packed.

The first artificial reverberators were proposed by Schroeder [1] and used comb filters in parallel followed by allpass filters in series. While this architecture produces the desired increase in echo density over time, unnatural coloration in the impulse response persisted [2]. Gerzon [3] generalized delay network reverberators, suggesting the use of a unitary feedback matrix to mix the outputs of the delay lines into the inputs of each other. Jot and Chaigne later proposed the Feedback Delay Network (FDN) [4, 5] which matched the delay line lengths with shelf filters designed to yield a desired frequency dependent T_{60}. Since then, FDNs have gained popularity for creating efficient artificial reverberation. Some important contributions in FDN research include [6, 7] in which the authors propose a circulant feedback matrix for efficient implementation and maximum diffusion, and those by Schlecht [8, 9] which deal with time-varying FDNs and their practical implementation. Schlecht also studied the properties of mixing matrices that produce lossless FDNs [10]. In a recent paper [11], he investigated the modal decomposition of feedback delay networks using the Ehrlich-Aberth iteration for finding poles and also studied the statistical distribution of mode frequencies and amplitudes. In this paper, we explicitly derive the state transition matrix for the case of frequency dependent T_{60}s.

The design of an optimum mixing matrix to produce a desired perceptual effect is still somewhat ambiguous. In this paper, we first derive the FDN state transition matrix for scalar gains and first-order shelf filters (for frequency dependent decay rates), and show that the poles (modes) of the system are given by eigenvalues of the state transition matrix. Next, we use the concept of *homotopy* [12] to gradually alter the mixing matrix from a state of minimum diffusion to maximum diffusion and observe how the pole trajectories vary with mixing. We observe coupling between nearby modes, where they first approach each other and then deflect, similar to what was observed by Weinreich in piano strings [13]. We discuss different ways of designing delay line T_{60} filters to model rooms which have walls made of different materials. Then, we calculate the normalized echo density (NED) [14] of the impulse responses at different levels of mixing and compare how mixing affects the time at which the late reverb starts (mixing time). NED has been shown to be a good indicator of the psychoacoustic perception of reverb [15, 16]. The effect of mixing time on echo density, which is shown to be a polynomial function dependent on delay line lengths, has been studied in [17]. Using curve fitting and empirical estimation, we derive a parametric function relating mixing time to mean delay line length and mixing matrix. This relationship between mixing time and level of diffusion sheds some light on designing FDN mixing matrices to achieve a desired perceptual effect.

The rest of the paper is organized as follows: in §2 we setup the FDN state space equations and derive the state transition matrix for frequency independent and frequency dependent decay rates. We show that the modes of the FDN are eigenvalues of the state transition matrix. In §3, we discuss how to smoothly vary the mixing matrix from zero to maximum diffusion. In §4, we see how mixing affects the mode trajectories and discuss how different decay filters can be designed. We also generate FDN impulse responses for different mixing matrices and show how their NEDs evolve with time. Finally, we come up with a parametric equation relating the mixing time to the mean delay line length and the mixing matrix. We conclude the paper in §5 and delineate scope for future work.

2. FEEDBACK DELAY NETWORK MODAL DECOMPOSITION

Let us consider a feedback delay network with N delay lines of length $\tau_1, \tau_2, \ldots, \tau_N$ samples each as shown in Fig. 1. The mixing matrix denoted by $\boldsymbol{M}$ is typically orthonormal and of the order $(N \times N)$, the 60 dB decay time of each delay line (T_{60}) is controlled by the filter $g_i(z)$. The input and output at time n is given by $u(n)$ and $y(n)$ respectively. The vectors $\mathbf{b}$ and $\mathbf{c}$ denote the input and output gains and d is the direct path gain.

The FDN can be represented as a state space system, with the state vector $\mathbf{x}$ defined by the values stored in each memory element.

$$\mathbf{x} = [x_1, x_2, \ldots, x_{\tau_1}, x_{\tau_1+1}, \ldots, x_{\sum_i \tau_i}]^T \tag{1}$$

With a state transition matrix $\boldsymbol{A}$, the state space equations can be written as

$$\begin{aligned} \mathbf{x}(n) &= \boldsymbol{A}\mathbf{x}(n-1) + \mathbf{b}u(n) \\ y(n) &= \mathbf{c}^T\mathbf{x}(n) + du(n) \\ Y(z) &= \left[\mathbf{c}^T(z\boldsymbol{I} - \boldsymbol{A})^{-1}\mathbf{b} + d\right] U(z) \end{aligned} \tag{2}$$

2.1. State Transition Matrix

It is obvious from (2) that the modes of the FDN are the eigenvalues of the matrix $\boldsymbol{A}$. In this section, we will derive $\boldsymbol{A}$ for two cases—first the case of frequency independent T_{60}s, in which case the delay line filter is a scalar gain, and second the more commonly used frequency dependent T_{60} case, in which $g_i(z)$ is typically given by a low-shelf filter.

Since $\boldsymbol{A}$ is a large sparse matrix of the order $(\sum_{i=1}^N \tau_i \times \sum_{i=1}^N \tau_i)$, we will use block matrix notation, and denote $\mathbf{x}$ as a stacked vector $[\tilde{\mathbf{x}}_1^T, \ldots, \tilde{\mathbf{x}}_N^T]$ and $\boldsymbol{A}$ as a block matrix with $N \times N$ sub-matrices.

$$\begin{bmatrix} \tilde{\mathbf{x}}_1(n) \\ \tilde{\mathbf{x}}_2(n) \\ \vdots \\ \tilde{\mathbf{x}}_N(n) \end{bmatrix} = \begin{bmatrix} \tilde{\boldsymbol{A}}_{11} & \tilde{\boldsymbol{A}}_{12} & \ldots & \tilde{\boldsymbol{A}}_{1N} \\ \tilde{\boldsymbol{A}}_{21} & \tilde{\boldsymbol{A}}_{22} & \ldots & \tilde{\boldsymbol{A}}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\boldsymbol{A}}_{N1} & \tilde{\boldsymbol{A}}_{2N} & \ldots & \tilde{\boldsymbol{A}}_{NN} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{x}}_1(n-1) \\ \tilde{\mathbf{x}}_2(n-1) \\ \vdots \\ \tilde{\mathbf{x}}_N(n-1) \end{bmatrix} \tag{3}$$

2.1.1. Frequency independent decay times

In this case, $g_i(z) = g_i$ for $i = 1 \ldots N$. Each sub-vector of the state vector $\mathbf{x}$ can be written as

$$\tilde{\mathbf{x}}_i^T = [x_{\sum_{j=1}^{i-1} \tau_j + 1}, \ldots, x_{\sum_{j=1}^{i} \tau_j}]^T \tag{4}$$

The sub-matrices, $\tilde{\boldsymbol{A}}_{ij}$ are of the order $(\tau_i \times \tau_j)$ for $i, j = 1 \ldots N$. The diagonal blocks can be written as

$$\tilde{\boldsymbol{A}}_{ii} = \begin{bmatrix} 0 & \ldots & 0 & g_i M_{ii} \\ 1 & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & 1 & 0 \end{bmatrix}, \tag{5}$$

where M_{ii} is the i, ith element of $\boldsymbol{M}$. The off-diagonal blocks have zeros everywhere, except the $(1, \tau_j)$-th element which is given by

$$\tilde{\boldsymbol{A}}_{ij}(1, \tau_j) = g_j M_{ji} \tag{6}$$

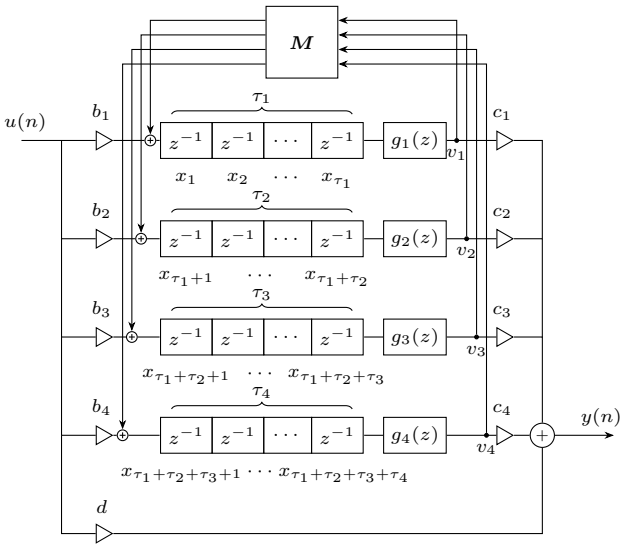


Figure 1: State space FDN block diagram.

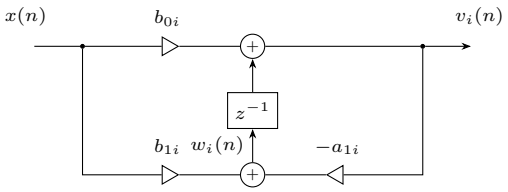


Figure 2: First order shelf filter block diagram.

2.1.2. Frequency dependent decay times

In this case, the delay-line filter is a first-order low shelf filter of the form

$$g_i(z) = \frac{b_{0i} + b_{1i} z^{-1}}{1 - a_{1i} z^{-1}} \tag{7}$$

This ensures that the lower frequency modes have a longer decay than the higher frequency modes, which is consistent with what is observed in actual rooms. We design a shelf filter by setting the filter coefficients according to desired gains at DC and Nyquist. The direct form II diagram of the shelf filter is given in Fig. 2. The state vector now has an additional state w_i per sub-vector.

$$\tilde{\mathbf{x}}_i^T = [x_{\sum_{j=1}^{i-1} \tau_j + 1}, \ldots, x_{\sum_{j=1}^{i} \tau_j}, w_i]^T \tag{8}$$

The sub-matrices, $\tilde{\boldsymbol{A}}_{ij}$ are of the order $(\tau_i + 1 \times \tau_j + 1)$ for $i, j = 1 \ldots N$. The diagonal blocks can be written as

$$\tilde{\boldsymbol{A}}_{ii} = \begin{bmatrix} 0 & \ldots & 0 & b_{0i} M_{ii} & M_{ii} \\ 1 & \ldots & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \ldots & 1 & b_{1i} - a_{1i} b_{01} & -a_{1i} \end{bmatrix} \tag{9}$$

The off-diagonal blocks have zeros everywhere, except the $(1, \tau_j)$

and $(1, \tau_j + 1)$-th elements which are given by

$$\begin{aligned} \tilde{\boldsymbol{A}}_{ij}(1, \tau_j) &= b_{0j} M_{ji}, \\ \tilde{\boldsymbol{A}}_{ij}(1, \tau_j + 1) &= M_{ji} \end{aligned} \tag{10}$$

3. HOMOTOPY

Now that we have derived the state transition matrix of the FDN, hence its modes, we wish to observe how these modes change as we modify the mixing matrix $\boldsymbol{M}$. For this, we must alter the mixing matrix smoothly from a state of no mixing to maximum mixing. A slow and continuous change from one function to another is known as *homotopy* [12]. In [8], the authors create a feedback matrix evolution equivalent to a recursive update using linear modulation functions. In this paper, we parameterize $\boldsymbol{M}$ as a function of θ. Starting with the (2×2) rotation matrix, $\boldsymbol{R}(\theta)$ which is orthonormal, we can generate an $(N \times N)$ orthonormal mixing matrix $\boldsymbol{M}(\theta)$ by taking the Kronecker product of $\boldsymbol{R}(\theta)$ with itself $\log_2 N$ times. It is to be noted that the Kronecker product of two orthonormal matrices is also orthonormal [18].

$$\begin{aligned} \boldsymbol{R}(\theta) &= \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \\ \boldsymbol{M}_{N\times N}(\theta) &= \boldsymbol{R}(\theta) \otimes \boldsymbol{R}(\theta) \otimes \ldots \boldsymbol{R}(\theta) \end{aligned} \tag{11}$$

Starting with $\theta = 0$ gives us $\boldsymbol{M}(0) = \boldsymbol{I}_{N\times N}$, i.e, N parallel delay lines with no mixing among them. By incrementally increasing θ by ϵ radians, new mixing matrices $\boldsymbol{M}(\theta + \epsilon)$ are generated with increased mixing among delay lines, until $\theta = \frac{\pi}{4}$, when the mixing matrix becomes Hadamard and depicts the case of maximum mixing. From now on, we will define θ as the mixing angle, and parametrize the mixing matrix as $\boldsymbol{M}(\theta)$.

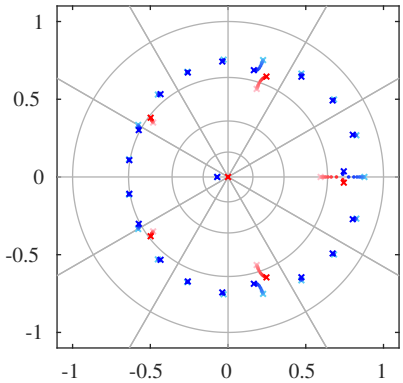
4. RESULTS AND DISCUSSION

4.1. Mode Trajectories

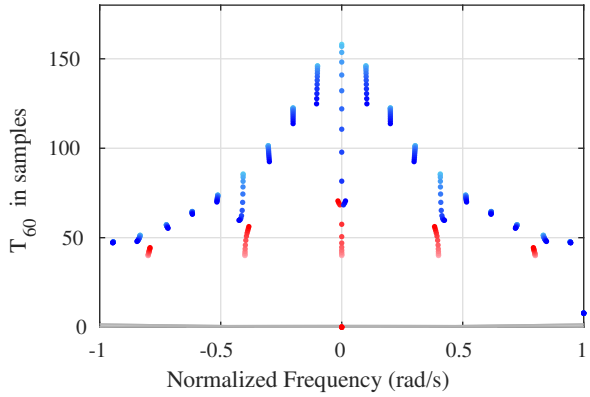
In Fig. 3 we show the modes in the complex plane and the frequency dependent T_{60} for a FDN with two delay lines, with $\tau_1 = 5$ (in red) and $\tau_2 = 19$ (in blue); $g_1(z)$ is a constant with $T_{60} = 40$ samples, $g_2(z)$ is a low-shelf filter with $T_{60_{DC}} = 150$ samples and $T_{60_{Nyquist}} = 50$ samples. The color intensity increases from light to dark as θ goes from 0 to $\frac{\pi}{4}$. We have squared the radii of the modes in Fig. 3a to make the pole movement easier to see.

The modes are in complex conjugate pairs. With no mixing, they are distributed uniformly in frequency. As mixing increases, the modes in the delay lines that are close in frequency approach each other in damping rapidly, and then deflect in frequency. This coupling behavior is more accentuated in the lower frequencies, with the modes at DC approaching each other and bifurcating. The modes at higher frequencies show less movement, and the mode at Nyquist remains stationary. In fact, once fully mixed, the shorter delay line that started with a constant T_{60} shows low-pass behavior.

The coupling among these modes is similar to what was observed by Weinreich in piano strings that are tuned with a slight deviation in frequency [13]. Due to this mistuning, the strings have slightly different angular frequencies and are coupled due to the dynamic motion of the bridge. The frequencies of the normal modes of the strings as a function of the deviation in angular frequency follow the same trajectory as our delay line modes when mixing increases.



(a) Poles in the complex plane.



(b) Frequency dependent T_{60}.

Figure 3: Modes of FDN for two delay line case with varying amounts of mixing. Red indicates a delay line with five taps and a scalar gain. Blue indicates a delay line with nineteen taps with a shelf filter. Lighter colors show the low mixing case and darker colors maximum mixing.

4.2. Designing different decay filters

Typically in an FDN reverberator, the decay filters are designed such that all delay lines independently produce the same T_{60} frequency response. However, it is likely that the physical configuration of a room would require multiple, concurrent T_{60} responses. One might use one set of filters to model air absorption while another set of filters can be used to model the absorption due to the materials in the room. For example, a church might have a pair of parallel walls with glass windows, and another set of walls covered with drapes. The absorption coefficient of glass decreases with increase in frequency, so its T_{60} response can be modeled by a low-shelf filter, while the T_{60} response of the drapery is better modeled with low-shelf resonant filter [19]. It would be realistic to assign part of the delay line filters to model the glass and the others to model the drapery. In [20], the T_{60} filters were designed such that all walls mimic the frequency-dependent absorption of cotton carpet. The mixing matrix could be adjusted to emulate the occupancy of the church—adding furniture would increase interaction between the modes, which is equivalent to increasing the mixing among the delay lines. Similarly, FDNs could have different filters in the delay lines with high diffusion to mimic the reverberant characteristics of coupled rooms [21]. Designing mixing matrices for coupled rooms

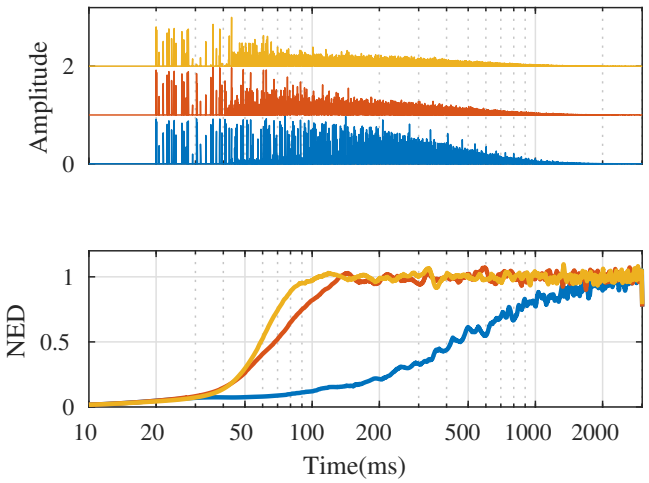


Figure 4: Impulse Responses and NED profiles for mixing matrices with θ set to $\frac{\pi}{40}$, $\frac{\pi}{8}$, and $\frac{\pi}{4}$, corresponding to 10%, 50%, and 100% mixing respectively (bottom to top).

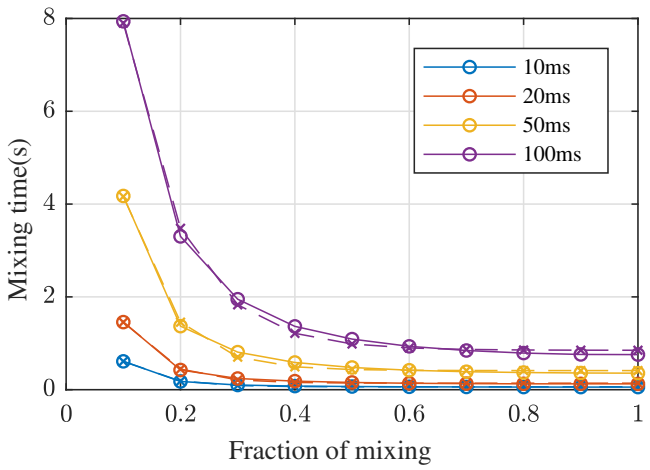


Figure 5: Mixing time as a function of mixing angle θ (divided by $\frac{\pi}{4}$) for average delay line lengths $\bar{\tau}$ of 10, 20, 50, and 100 ms (bottom to top). Dashed lines are fitted curves using the parametric equation.

with bleed between them has been previously studied in [22].

4.3. Echo Density and Mixing Time

In Fig. 4, we show how the normalized echo density profile (calculated with a 50 ms Hanning window) changes with change in mixing angle θ for a FDN with 16 delay lines, with delay lengths uniformly distributed between 10–20 ms at a sampling frequency of 48 kHz with $T_{60_{DC}} = 4$ s and $T_{60_{Nyquist}} = 2$ s. As expected, as the amount of mixing increases, the echo density becomes Gaussian more quickly. Audio examples of the impulse responses are available at `https://ccrma.stanford.edu/~orchi/FDN/IR.html`

A FDN artificial reverberator has many parameters that need to be tuned, such as mixing matrix, number of delay lines, their lengths etc. to produce a desired perceptual effect. The tuning of these parameters is still considered somewhat an art. The mixing time of a feedback delay network is defined as the time in which the echo density reaches a threshold, $\mathcal{T} = 0.9$ in this case [15]. In [17], the authors come up with a closed form solution for choosing a mean delay length, given a desired mixing time. We wish to relate mixing time with the mixing angle. Figure 5 shows the mixing time against the mixing angle for the same FDN for 4 mean delay line lengths, $\bar{\tau}$ of 10, 20, 50, and 100 ms. For each mean delay line length, we perform Monte Carlo simulations, such that each of the 16 delay line lengths are random and uniformly distributed on the interval

$$\tau_i \sim \mathcal{U}\left[\frac{\bar{\tau}_i}{\phi}, \bar{\tau}_i\phi\right], \tag{12}$$

where ϕ is the golden ratio, and plot the average mixing time. An exponential relationship is observed between mixing time and mixing angle θ. For $\theta = 0$, there is no mixing and the echo density never becomes Gaussian, which means the mixing time is theoretically infinite. For a given mean delay line length, it is possible to fit a curve to this plot using non-linear least squares and estimate what value of θ would produce the desired mixing time. The parametric equation for t_{mix} in seconds, given θ and a mean delay line length of $\bar{\tau}$ in ms is given in (13).

$$\begin{aligned} t_{mix} &= \alpha \exp\left(-\beta\theta\right) + \gamma \\ \alpha &= 0.397\bar{\tau} - 20.02\gamma \\ \beta &= 20.246 - 0.09\bar{\tau} \\ \gamma &= 0.0075\bar{\tau} \end{aligned} \tag{13}$$

Combined with Schlecht's method for predicting mean delay line length [17], our method for picking a mixing matrix $\boldsymbol{M}(\theta)$ to achieve a desired mixing time will aid in the design of FDN reverberators. These results are also directly applicable to the FDN resizing algorithm described in [23].

5. CONCLUSION

We have studied the modal decomposition of a Feedback Delay Networks. While the modal decomposition for a FDN with a scalar gain is well known [6], we explicitly derived the modal decomposition with a shelf filter in the delay lines. We changed the mixing matrix smoothly from minimum mixing to maximum mixing and observed coupling among nearby modes from different delay lines. We used a 2 delay line example with a small number of modes for visualizing modal behavior. We also studied how mixing affects the echo density profile and mixing time of the FDN impulse response, and came up with a parametric equation to choose a mixing matrix, given a mean delay line length and a desired mixing time.

We could not do modal decompositions of FDNs with longer delay lines that are commonly used in practice, because of the limited efficiency of MATLAB's `eig` function when dealing with large matrices. Nor could we find a good numeric method that could efficiently and accurately give us all the eigenvalues of a large, sparse, non-symmetric matrix. It would be interesting to use the numerical method described in [11] to find the eigenvalues of large state transition matrices with many more states in a future work. We could then generalize the observed coupled modal behavior in this paper for practical FDNs with thousands of modes.

6. REFERENCES

[1] M. R. Schroeder, "Natural sounding artificial reverberation," *Journal of the Audio Engineering Society*, vol. 10, no. 3, pp. 219–223, 1962.

[2] J. A. Moorer, "About this reverberation business," *Computer music journal*, pp. 13–28, 1979.

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

[4] J.-M. Jot and A. Chaigne, "Digital delay networks for designing artificial reverberators," in *Audio Engineering Society Convention 90*. Audio Engineering Society, 1991.

[5] J.-M. Jot, "An analysis/synthesis approach to real-time artificial reverberation," in *[Proceedings] ICASSP-92: 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2. IEEE, 1992, pp. 221–224.

[6] D. Rocchesso and J. O. Smith, "Circulant and elliptic feedback delay networks for artificial reverberation," *IEEE Transactions on Speech and Audio Processing*, vol. 5, no. 1, pp. 51–63, 1997.

[7] D. Rocchesso, "Maximally diffusive yet efficient feedback delay networks for artificial reverberation," *IEEE Signal Processing Letters*, vol. 4, no. 9, pp. 252–255, 1997.

[8] S. J. Schlecht and E. A. Habets, "Time-varying feedback matrices in feedback delay networks and their application in artificial reverberation," *The Journal of the Acoustical Society of America*, vol. 138, no. 3, pp. 1389–1398, 2015.

[9] ——, "Practical considerations of time-varying feedback delay networks," in *Audio Engineering Society Convention 138*. Audio Engineering Society, 2015.

[10] ——, "On lossless feedback delay networks," *IEEE Transactions on Signal Processing*, vol. 65, no. 6, pp. 1554–1564, 2017.

[11] ——, "Modal decomposition of feedback delay networks," *arXiv preprint arXiv:1901.08865*, 2019.

[12] J.-H. He, "Homotopy perturbation technique," *Computer methods in applied mechanics and engineering*, vol. 178, no. 3-4, pp. 257–262, 1999.

[13] G. Weinreich, "Coupled piano strings," *The Journal of the Acoustical Society of America*, vol. 62, no. 6, pp. 1474–1484, 1977.

[14] J. S. Abel and P. Huang, "A simple, robust measure of reverberation echo density," in *Audio Engineering Society Convention 121*. Audio Engineering Society, 2006.

[15] P. Huang, J. S. Abel, H. Terasawa, and J. Berger, "Reverberation echo density psychoacoustics," in *Audio Engineering Society Convention 125*. Audio Engineering Society, 2008.

[16] A. Lindau, L. Kosanke, and S. Weinzierl, "Perceptual evaluation of model-and signal-based predictors of the mixing time in binaural room impulse responses," *Journal of the Audio Engineering Society*, vol. 60, no. 11, pp. 887–898, 2012.

[17] S. J. Schlecht and E. A. Habets, "Feedback delay networks: Echo density and mixing time," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 2, pp. 374–383, 2017.

[18] P. A. Regalia and M. K. Sanjit, "Kronecker products, unitary matrices and signal processing applications," *SIAM review*, vol. 31, no. 4, pp. 586–613, 1989.

[19] J. S. Abel, D. P. Berners, and J. Herrera, "Signal processing techniques for digital audio effects," *Course Notes for Stanford Music 424*, 2004.

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

[21] J. S. Abel and M. J. Wilson, "Luciverb: Iterated convolution for the impatient," in *Audio Engineering Society Convention 133*. Audio Engineering Society, 2012.

[22] S. J. Schlecht and E. A. Habets, "Sign-agnostic matrix design for spatial artificial reverberation with feedback delay networks," in *AES International Conference on Spatial Reproduction-Aesthetics and Science*. Audio Engineering Society, 2018.

[23] E. K. Canfield-Dafilou and J. S. Abel, "Resizing rooms in convolution, delay network, and modal reverberators," in *Proceedings of the 21st International Conference on Digital Audio Effects*, 2018.