THE ROLE OF MODAL EXCITATION IN COLORLESS REVERBERATION

Janis Heldmann[1] *and Sebastian J. Schlecht*[1,2],

[1]Acoustics Lab, Dept. of Signal Processing and Acoustics
[2]Media Lab, Dept. of Media
Aalto University
Espoo, Finland
janis.heldmann@aalto.fi

ABSTRACT

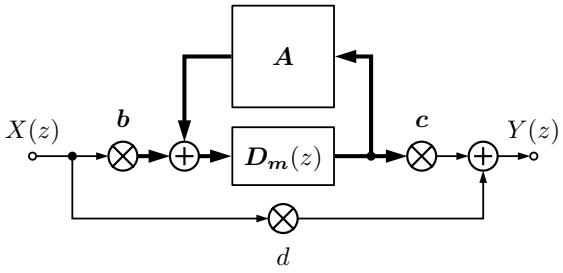
A perceptual study revealing a novel connection between modal properties of feedback delay networks (FDNs) and colorless reverberation is presented. The coloration of the reverberation tail is quantified by the modal excitation distribution derived from the modal decomposition of the FDN. A homogeneously decaying allpass FDN is designed to be colorless such that the corresponding narrow modal excitation distribution leads to a high perceived modal density. Synthetic modal excitation distributions are generated to match modal excitations of FDNs. Three listening tests were conducted to demonstrate the correlation between the modal excitation distribution and the perceived degree of coloration. A fourth test shows a significant reduction of coloration by the colorless FDN compared to other FDN designs. The novel connection of modal excitation, allpass FDNs, and perceived coloration presents a beneficial design criterion for colorless artificial reverberation.

1. INTRODUCTION

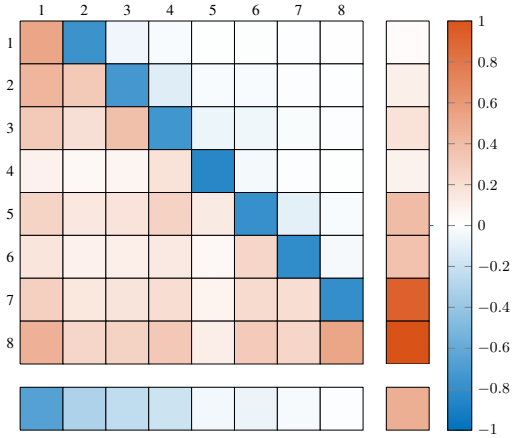
Two common goals in the design of artificial reverberation are to efficiently generate perceptually convincing room acoustics or aesthetic and musical sound effects [1]. Either approach is often derived from an ideal baseline system capable of producing a smooth and neutral reverberation tail—particularly sought after are reverberations that are colorless, i.e., spectrally flat and devoid of resonating peaks. Ideal colorless reverberation can be generated with exponentially decaying Gaussian noise [2], sparse velvet noise [3, 4] or by superposing a large number of modes [5].

Feedback delay networks (FDNs) are a computationally efficient structure for artificial reverberation with a mature theoretical foundation [6, 7, 8, 9], see Fig. 1a. In its early form, Schroeder and Logan [10] pursued colorless reverberation by concatenating delay-line-based allpass filters. Jot and Chaigne [7] introduced a two-stage design process, creating a lossless system and then introducing delay-proportional loss. The resulting decay is homogeneous, i.e., all modes decay according to a global frequency-dependent curve. Homogeneous decay prevents single modes from ringing past the overall decay and enforces a similar spectral smoothness over the duration of the reverberation tail.

To generate colorless reverberation with FDNs, additional recommendations on the actual choice of parameters, namely the feedback matrix and delay line lengths, were proposed, including co-



(a) *Block diagram of a feedback delay network (FDN) with a single-channel input and output and multi-channel recursive path. Thin and thick lines indicate single and multi-channel connections, respectively.*



(b) *The system matrix* $\boldsymbol{V} = \left[\begin{smallmatrix} \boldsymbol{A} & \boldsymbol{b} \\ \boldsymbol{c} & d \end{smallmatrix}\right]$ *of an allpass FDN with eight delay lines, i.e.,* $N = 8$*, and feedback gains* $\boldsymbol{A}$*, direct gain* d*, input and output gains* $\boldsymbol{b}$ *and* $\boldsymbol{c}$*, respectively. The color indicates the gain of each matrix entry.*

Figure 1: *Block diagram and parameters of an FDN.*

prime delays [10], non-degenerate delay distributions [11], sufficient number of modes [10, 7, 5], dense lossless feedback matrices [8, 12, 9], and diffusion filters [13, 14, 15]. Despite many such recommendations, large FDNs can have hundreds of tuning parameters. Practitioners commonly report that extensive tuning is required such that also machine-learning-based approaches were considered [16, 17].

This paper proposes a novel approach to designing colorless FDNs. To this end, we establish a connection between perceived coloration and the resonating modes of an FDN. We show that

the way modal excitation is distributed strongly affects coloration. In particular, coloration is low if the modes are excited equally. Further, we show that allpass filters can exhibit favorable modal excitation under certain conditions. Based on these relations, we propose a colorless FDN based on the recently developed allpass FDN design [18], see Fig. 1b.

We conducted four listening tests to demonstrate the impact of different modal distributions. We show that the proposed FDN has significantly higher ratings of colorlessness than other FDNs. It is emphasized that all studied FDN designs employ only the basic processing units.

The remaining paper is organized as follows: Section 2 introduces the background on FDNs, modal decomposition, and prior evaluation of coloration. Section 3 proposes an objective technique using modal excitation distributions. In Section 4, a perceptual evaluation is presented, and Section 5 presents the results and discussion of the findings. All examples in this paper can be reproduced by using the FDN Toolbox [19]. All generated sound examples are available online[1] [20].

2. BACKGROUND

2.1. Prior artificial reverberators

In this section, three of the most relevant structures for colorless artificial reverberation are reviewed.

2.1.1. Schroeder series allpass

The first approach to achieve colorless artificial reverberation is to assure a flat frequency response by the allpass criterion of the z-domain transfer function $H(z)$, i.e.,

$$|H(z)| = 1, \qquad \text{for } |z| = 1. \tag{1}$$

Schroeder and Logan proposed an allpass reverberator [10] using a series of feedforward-feedback delay filters. The transfer function of a single Schroeder allpass is

$$H_{\text{Schroeder}}(z) = \frac{g + z^{-m}}{1 + g z^{-m}}, \tag{2}$$

where g is the real-valued feedforward and feedback gain and m is the delay in samples.

Nonetheless, the Schroeder reverberator exhibits strong coloration, which led Schroeder to formulate further criteria to achieve colorless artificial reverberation [10]. These included overlapping normal modes over the whole frequency range, equal reverberation times for each mode, sufficient echo density, absence of periodicity in time domain, and no periodic or comb-like frequency responses. It was also noted that *too many* series filters lead to non-exponential decay envelopes and should be avoided [10]. In this study, we nonetheless included a chain of eight Schroeder allpasses for illustrative purpose.

2.1.2. Homogeneous FDN

The FDN is a combination of a feedback matrix with input, output and direct gains and can be stated in the so-called delay state space form

$$H(z) = \boldsymbol{c}^\top [\boldsymbol{D}_{\boldsymbol{m}}(z) - \boldsymbol{A}]^{-1} \boldsymbol{b} + d \tag{3}$$

with $N \times N$ feedback matrix $\boldsymbol{A}$, $N \times 1$ input gain vector $\boldsymbol{b}$, $1 \times N$ output gain $\boldsymbol{c}$, scalar direct gain d, and $N \times N$ delay matrix $\boldsymbol{D}_{\boldsymbol{m}}(z) = \text{diag}\big([z^{-m_1}, z^{-m_2}, \ldots, z^{-m_N}]\big)$. Here, only the case of single-input single-output (SISO) FDNs is considered. The delay lengths are defined by the vector $\boldsymbol{m} = [m_1, \ldots, m_N]$ in samples and total system order is $\mathfrak{N} = \sum_{i=1}^{N} m_i$. The feedforward and feedback gains can be collected in the system matrix

$$\boldsymbol{V} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{b} \\ \boldsymbol{c} & d \end{bmatrix}. \tag{4}$$

The feedback matrix is the product of the unilossless matrix $\boldsymbol{U}$, e.g., an orthogonal matrix [9], and the diagonal gain matrix $\boldsymbol{\Gamma}$

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Gamma}. \tag{5}$$

As $\boldsymbol{U}$ is unilossless, the reverberation time T_{60} of all modes is only dependent on $\boldsymbol{\Gamma}$, which in turn derives from the attenuation-per-sample γ by $\boldsymbol{\Gamma} = \text{diag}(\gamma^{\boldsymbol{m}})$. For a dense feedback matrix $\boldsymbol{U}$, i.e., without zero entries, the FDN exhibits a dense distribution of overlapping modes and generates an exponentially decaying noise response.

Common feedback matrices are the random-orthogonal matrix as shown in Fig. 2b, the Hadamard matrix shown in Fig. 2a [21], and the Householder matrix shown in Fig. 2c [13]. Additionally, the Schroeder series allpass can be found as a special case of the FDN [18], see Fig. 2d. Concrete parameters for these examples are given in Section 3.1.

2.1.3. Uniallpass FDN

The uniallpass FDN was recently proposed in [18] as a general class of FDN system matrices $\boldsymbol{V}$ which are allpass for arbitrary delay lengths $\boldsymbol{m}$. Further, the allpass FDN can also be homogeneous, i.e., satisfy (5). For a given feedback matrix $\boldsymbol{A}$, the input, output, and direct gains $\boldsymbol{b}, \boldsymbol{c}, d$ are selected such that the system fulfills the allpass criterion. The allpass FDN generalizes other allpass topologies such as the Schroeder series allpass and Gardener's nested allpass [22]. The system matrix of a homogeneous uniallpass FDN is shown in Fig. 1b.

2.2. Modal decomposition

The impulse response $h(n)$ of any linear and time-invariant filter can be decomposed into finitely many modes $h_i(n)$, i.e., decaying complex exponentials,

$$h(n) = \sum_{i=1}^{\mathfrak{N}} h_i(n), \tag{6}$$

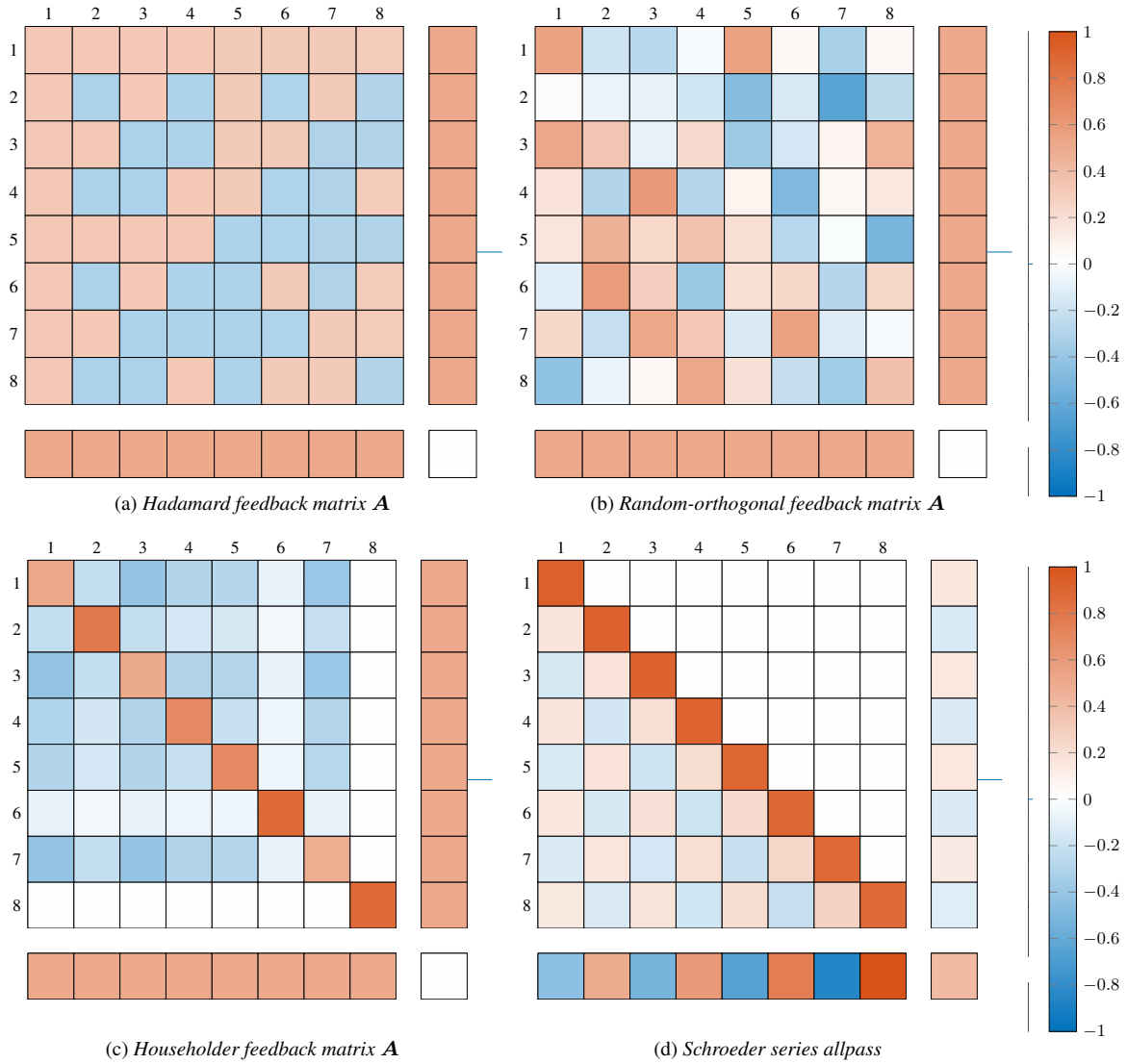
where $\mathfrak{N}$ is the number of modes and n denotes the discrete time. The modal decomposition procedure for FDNs as proposed in [23] is briefly explained in the following.

The transfer function of the FDN (3) can be expressed as a rational fraction in the form of

$$H(z) = \frac{q_{\boldsymbol{m},\boldsymbol{A},\boldsymbol{b},\boldsymbol{c},d}(z)}{p_{\boldsymbol{m},\boldsymbol{A}}(z)} \tag{7}$$

with polynomials $p_{\boldsymbol{m},\boldsymbol{A}}(z)$ and $q_{\boldsymbol{m},\boldsymbol{A},\boldsymbol{b},\boldsymbol{c},d}$. Based on partial fraction expansion, the transfer function in (7) can be expressed as

[1] https://www.sebastianjiroschlecht.com/publication/OnColorlessReverb



(a) *Hadamard feedback matrix* $\boldsymbol{A}$

(b) *Random-orthogonal feedback matrix* $\boldsymbol{A}$

(c) *Householder feedback matrix* $\boldsymbol{A}$

(d) *Schroeder series allpass*

Figure 2: *System matrices* $\boldsymbol{V} = \left[\begin{smallmatrix} \boldsymbol{A} & \boldsymbol{b} \\ \boldsymbol{c} & d \end{smallmatrix}\right]$ *of four different FDN with eight delay lines, i.e.,* $N = 8$*. For Fig. 2a-2c only the matrix is strongly defined such that all values of input and output gains,* $\boldsymbol{b}$ *and* $\boldsymbol{c}$ *are set to* 0.5 *and the direct gain is* $d = 0$*.*

a superposition of complex one-pole resonators with pole λ_i and residue ρ_i, i.e.,

$$H(z) = d + \sum_{i=1}^{\mathfrak{N}} \frac{\rho_i}{1 - \lambda_i z^{-1}} \tag{8}$$

and in the time domain

$$h_i(n) = |\rho_i||\lambda_i^n| e^{\imath(n\angle\lambda_i + \angle\rho_i)}. \tag{9}$$

Each mode is then characterized by the frequency $\angle\lambda_i$, decay $|\lambda_i|$, amplitude $|\rho_i|$ and phase $\angle\rho_i$.

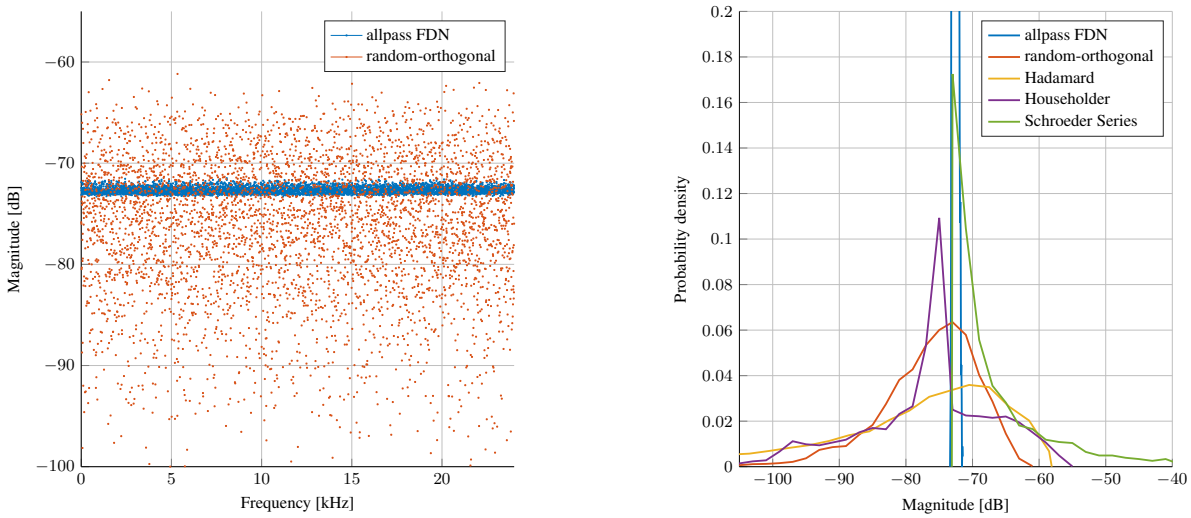
2.3. Modal properties

The quantitative as well as the qualitative aspects of the set of modes used to generate artificial reverberation are important to achieve a natural-sounding reverb.

Prior research on modal density focused on the number of modes $\mathfrak{N}$ and the frequency distribution of the modes $\angle\lambda_i$. Karjalainen and Järveläinen studied uniform frequency distributions over the Bark scale [5]. For equal modal excitation, the minimum number of modes distributed over the frequency range of 80 Hz-10 kHz was found to be at least 3000 modes[2], while 10000 modes were used as a reference sample. Room acoustics modeling can be readily based on synthetically distributed modes [24].

For FDNs, it was found that the modal frequencies are commonly distributed uniformly across the frequency spectrum [23]. As such, no coloration is expected due to clusters of modes. Specific relations between modal phases can lead to prominent constructive and destructive interference, such as in the early part of

[2]In this paper, modes are defined as one-pole complex resonators as opposed to second-order sections like in [5]. All reported numbers are adjusted to this definition.



(a) *Modal decomposition for allpass FDN and random-orthogonal FDN.*

(b) *Modal excitation distributions of various FDN designs.*

Figure 3: *The pdf-normalized histograms of modal excitation distributions in (b) are derived from modal decomposition of FDNs, exemplary shown in (a) for the allpass FDN and the homogeneous FDN with random-orthogonal mixing matrix.*

the impulse response. The late response is typically unaffected on account of the phases drifting apart over time.

Especially important for the remaining paper is the absolute value of the residue that is equivalent with the excitation of each mode $|\rho_i|$. For homogeneous FDNs satisfying (5), all system modes decay equally, i.e., $|\lambda_i| = \gamma$. While typically, the decay of the modes is fully determined by the target reverberation time T_{60}, the modal excitation is largely unconstrained by design.

The impact of modal excitation envelopes on the perception of coloration in artificial reverberation has not yet been the target of investigation. To evaluate its role, a suitable subjective test design is needed.

3. ANALYSIS OF MODAL EXCITATION

Uniallpass and homogeneous FDNs have fulfilled the criteria, set forth by Schroeder and Karjalainen, of producing a sufficient number of modes and equal decay rates across frequency. And yet, informal listening tests have shown that coloration can still be perceived, especially in later reverberation. Therefore, another unconsidered aspect of the quantity or quality of modes has an effect on the coloration. In the following, we investigate the modal excitation distribution.

3.1. Specification of analyzed FDNs

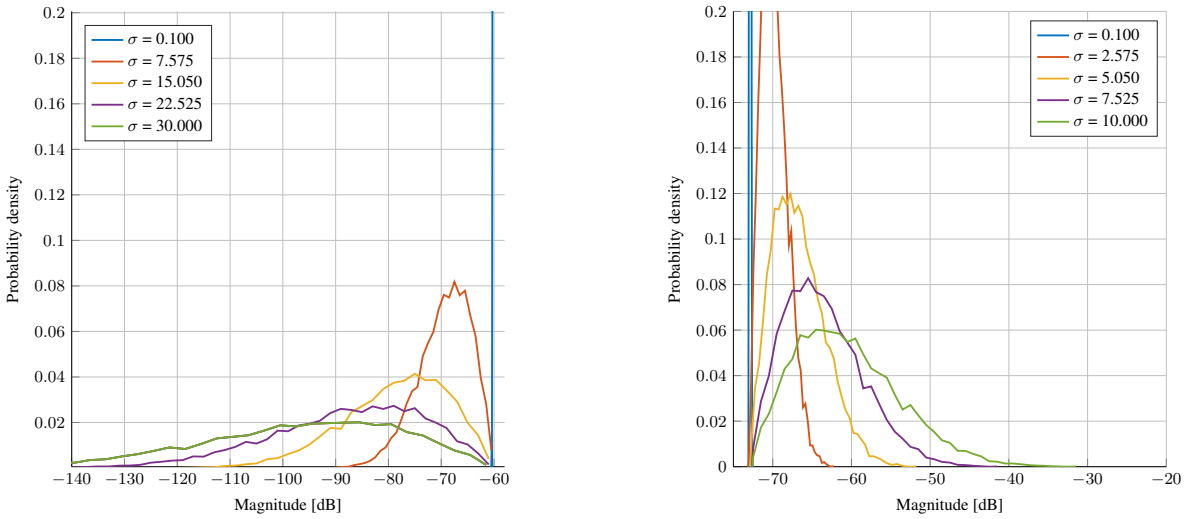
In this paper, we study five FDNs. All FDNs share the same size $N = 8$ and delays $\boldsymbol{m} = [809, 877, 937, 1049, 1151, 1249, 1373, 1499]$. The delays are prime numbers logarithmically distributed between 800 and 1500 samples. The prime delays avoid double poles for the Schroeder series allpass, which would cause a non-exponential decay [10]. The number of modes is $\mathfrak{N} = 8944$. All FDNs are homogeneous with an attenuation-per-sample $\gamma = 0.9999$, equal to a T_{60} of 1.44 s at a sample rate of 48 kHz.

We included three conventional FDNs with random-orthogonal, Hadamard and Householder feedback matrices with scaled input and output gains, $b_i = 0.5$ and $c_i = 0.5$ for all i, respectively, and direct gain $d = 0$. Further, two allpass FDNs are included. The Schroeder series allpass with N units in series, where the feedforward/-back gain in (2) is set to homogenize the decay, i.e., $g_i = \gamma^{m_i}$. The proposed allpass FDN is a solution to the completion problem in [18][3]. The completion problem determines $\boldsymbol{b}$, $\boldsymbol{c}$ and d for a given feedback matrix $\boldsymbol{A}$, such that the system matrix $\boldsymbol{V}$ becomes allpass. All FDN system matrices $\boldsymbol{V}$ are depicted in Fig. 1b and 2. The matrix multiplication in the Hadamard, Householder and allpass series FDNs can be implemented more efficiently than the proposed allpass and random-orthogonal FDNs.

3.2. FDN modal distribution

We analyze the modal distribution of the five different FDN designs. First, the FDN is decomposed into its modes; see (8) and (9) [23]. As all FDNs are homogeneous, the modal decays are equal, and as we are primarily concerned with the late tail, we do not study further the modal phase. In Fig. 3a, the modal excitation $|\rho_i|$ is plotted against the modal frequency for the random-orthogonal and the proposed allpass FDN. The modal excitations are gathered and displayed in a histogram, see Fig. 3b. The histogram is normalized as a probability density function estimate, where the number of observations in the bin is divided by the total number of observations multiplied by the width of the bin. It can be observed that the modal excitation distributions differ in shape and spread. The distributions show varying slopes away from a central peak density. The random-orthogonal FDN has a wide distribution in comparison to the steep and narrow distribution of the

[3]The homogeneous allpass FDN is fully determined by the similarity matrix $\boldsymbol{X}$, which satisfies [18, Eq. (72)]. Here, $\boldsymbol{X} = \mathrm{diag}([1.00, 1.2542, 1.5285, 1.9522, 2.5489, 3.5589, 4.7309, 7.0943])$.



(a) *Flipped Rayleigh distribution with increasing* σ.

(b) *Rayleigh distribution with increasing* σ.

Figure 4: *Probability density functions of (a) flipped Rayleigh distributions with input* $x < 0$*, scaling factor* $0.1 \leq \sigma \leq 30$ *and gain shift* $\mu = -60\,\mathrm{dB}$*, and (b) Rayleigh distributions with input* $x < 0$*, scaling factor* $0.1 \leq \sigma \leq 10$ *and gain shift* $\mu = -73\,\mathrm{dB}$.

proposed allpass FDN. The Householder FDN has a steep peak at $-75\,\mathrm{dB}$ combined with a moderately flat floor of higher and lower modal excitation. The series allpass has a steep maximum and a continuously decaying density towards high modal excitation.

3.3. Synthetic modal distribution

To study the modal excitation more systematically, synthetic distributions with simple parameters are generated. By inverse transform sampling, a set of modes can be drawn from an arbitrary probability distribution. While there are many possible distributions, we found that the Rayleigh distribution [25] shows similar characteristics as the FDNs described above. The Rayleigh distribution is given by

$$f(x;\sigma;\mu) = \frac{x-\mu}{\sigma^2} e^{-(x-\mu)^2/(2\sigma^2)}, \quad x \geq 0, \tag{10}$$

with the Rayleigh scaling factor σ and the gain shift μ of all modal excitation.

While scaling factors with $0 < \sigma < 1$ generate steep distributions similar to those of the proposed allpass FDN, increasing the scaling factor σ flattens the distribution with low densities towards stronger excitation. For $x < 0$, the Rayleigh distribution can be flipped and therefore produces low densities for low modal excitation. A variety of Rayleigh distributions depicted as probability density functions can be seen in Fig. 4. Approximate scaling factors are $\sigma = 0.3$ for the proposed allpass FDN, while $\sigma = 10$ and $\sigma = 16$ are flipped scaling factors for random-orthogonal and Hadamard feedback matrices, respectively. The modal phases $\angle\rho_i$ are randomized uniformly between 0 and 2π. The modal frequencies $\angle\lambda_i$ are complex conjugated pairs for a real-valued time-domain response and are distributed between 0 and 2π. For better control, the modal frequencies are distributed with low discrepancies from an equidistribution, i.e., $\angle\lambda_i = 2\pi(i+u)/\mathfrak{N}$ with $u \sim \mathcal{U}(-\zeta,\zeta,)$, where ζ indicates the extent of the discrepancy and $\mathcal{U}(x,y,)$ denotes a uniform distribution between x and y. For all the synthetic distributions, the discrepancy is $\zeta = 0.5$.

3.4. Modal distribution and allpass property

For allpass filters, we show that the modal excitation is determined by the modal frequency distribution. Allpass filters are characterized by reciprocal location of poles and zeros, i.e.,

$$H(z) = \prod_{i=1}^{\mathfrak{N}} \frac{1-\overline{\lambda_i}^{-1} z^{-1}}{1-\lambda_i z^{-1}}, \tag{11}$$

such that for each pole λ_i, there is a zero at $1/\overline{\lambda_i}$, where $\bar{\cdot}$ denotes the complex conjugate. The residues are then computed as

$$\rho_j = (1-|\lambda_j|^{-2}) \prod_{i=1, i\neq j}^{\mathfrak{N}} \frac{1-\overline{\lambda_i}^{-1}\lambda_j^{-1}}{1-\lambda_i\lambda_j^{-1}}. \tag{12}$$

Here, we limit our analysis to systems with only single poles, such that for homogeneous decay, i.e., $|\lambda_i| = \gamma$, the modal excitation $|\rho_i|$ is entirely determined by the modal frequencies $\angle\lambda_i$. In the following, we establish a heuristic relation between the distribution of the modal frequency and the modal excitation.

The dominating factor in (12), is $1/(1-\lambda_i\lambda_j^{-1})$ for poles i and j being in close proximity, i.e., the closer λ_i and λ_j are, the larger is $|\rho_j|$. For equidistributed modal frequencies, i.e., $\angle\lambda_i = i2\pi/\mathfrak{N}$, all modal excitation are equal, i.e., $|\rho_i| = |\rho_j|$ for all i and j. For low-discrepancy distributions, the larger the discrepancy ζ, the wider distributed is the modal excitation, see Fig. 5. For uniformly distributed modal frequencies, i.e., $\angle\lambda_i \sim \mathcal{U}(-\pi,\pi,)$, the distribution appears widest. By comparing with Fig. 3b, the proposed allpass FDN has a low discrepancy ζ, while the series allpass has high ζ.

Please note, the synthetic modal distributions are not allpass in general. However, narrow distributions with small σ and low discrepancy ζ tend to generate flat magnitude responses. Therefore, the allpass property is not strictly necessary for narrow modal excitation in FDNs, albeit it is a supportive design constraint.

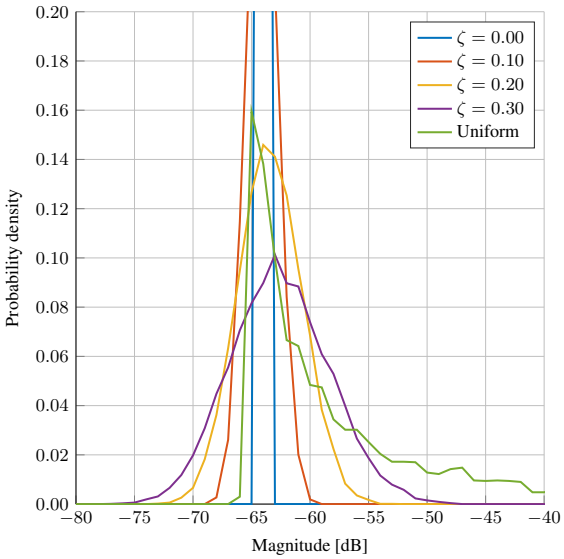


Figure 5: *Probability density function of the modal excitation distribution of allpass filters with* $\mathfrak{N} = 10000$. *A uniform modal frequency distribution is compared to low-discrepancy distributions with discrepancy* ζ.

4. PERCEPTUAL EVALUATION OF MODAL EXCITATION DISTRIBUTION

We conducted a listening test to evaluate the correlation between the modal excitation distribution and the perceived coloration in the impulse response.

The test is based on a modified MUSHRA and includes a training and evaluation phase [26]. In the training phase, the participant is familiarized with all sound samples and can adjust the overall loudness. The loudness shall further be unchanged. The test is conducted at home with the participants' choice of headphones[4].The test was deployed to the participants via MUSHRAM, a set of MATLAB routines provided by [27].

The reference sample for all tests is exponentially decaying Gaussian white noise as it is an ideal reverberation tail [2]. In the evaluation phase, the reference signal is compared to five sound samples evaluated with a slider between 0 and 100. The slider is labeled with 0 - *certainly colored*, 25 - *rather colored*, 50 - *fairly colored / colorless*, 75 - *rather colorless*, and 100 - *certainly colorless*. The test items also include a hidden reference and the test design requires at least one sample to be rated as *certainly colorless* to assure the listener can identify the reference sample.

Each individual sample was RMS normalized, and subsequently, all samples were normalized by the highest peak of all test samples in all four tests. In informal tests, differences in the early part of the impulse responses were observed, such as a slow build-up of energy in the impulse response generated from the allpass FDNs. Therefore, all impulse responses were convolved with a pink noise burst of 300 ms to smoothen time-domain differences.

The reverberation time of all test items is set to $1.44\,\mathrm{s}$. In particular, all tested FDNs are homogeneous. The synthetic modal synthesis samples have a low-discrepancy modal frequency distribution with $\zeta = 0.5$.

4.1. Test 1: Equal modal excitation with varying number of modes

The first test evaluates the number of required modes to generate colorless artificial reverberation when all modes are excited equally. A set of modes is generated and superposed as in (6). The tested number of modes $\mathfrak{N}$ are $1000, 3000, 6000, 10000, 20000$. All modal excitation are equal at a value of μ= $-60\,\mathrm{dB}$.

4.2. Test 2: Synthetic modal excitation distribution sampled from Rayleigh distribution

For the second test, a set of $\mathfrak{N} = 20000$ modes is generated with a varying distribution of the modal excitation. The excitation follows a Rayleigh distribution with an increasing scaling parameter σ, see Fig. 4b. The σ are linearly distributed between 0.1 and 10, concretely as $[0.1, 2.6, 5.1, 7.5, 10]$. In prior informal listening tests, $\sigma = 0.1$ with $\mathfrak{N} = 20000$ was perceived to be similar to white noise and similar to the top sample in Test 1. A scaling parameter of $\sigma = 10$ was found to clearly exhibit metallic artifacts. The gain shift of the distributed modes is set to $\mu = -73\,\mathrm{dB}$, similar to the modal excitation distributions seen in Fig. 3b.

4.3. Test 3: Synthetic modal excitation distribution sampled from flipped Rayleigh distribution

Some modal excitation distributions such as Hadamard and random-orthogonal feedback matrices show a steep slope towards higher modal excitation and a slower decreasing slope towards lower modal excitation, as visible in Fig. 3b, see also Sec. 3.3. The resulting modal excitation distributions are depicted in Fig. 4a. Informal listening tests showed that the σ exhibiting clearly metallic coloration of the sound sample exceeds the prior test by a factor of three. The linearly distributed σ parameters are $[0.1, 7.6, 15.1, 22.5, 30]$ with $\mu = -60\,\mathrm{dB}$.

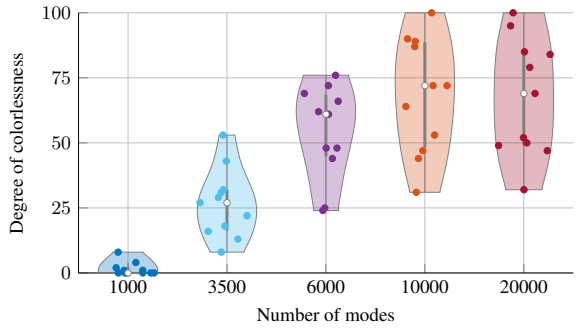
4.4. Test 4: Evaluation of FDN artificial reverberation

The last test evaluates the coloration of the FDNs as described in Sec. 3.1. The impulse responses have been generated by standard time-domain recursion. The selection is comprised of typical choices for computational efficient artificial reverberation prominent in literature and the proposed allpass FDN that is designed to achieve a colorless reverb.

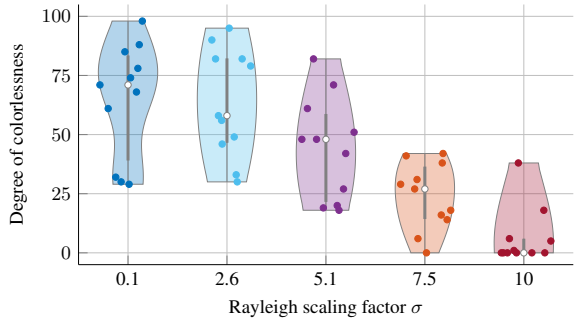
5. RESULTS AND EVALUATION

In total, 11 experienced listeners obtained the MUSHRAM MATLAB files, took the listening test remotely, and submitted their results via an online cloud service. The results of the listening test are shown by violin plots with median values, 25 percent and 75 percent quartiles, and individual ratings. The results are discussed

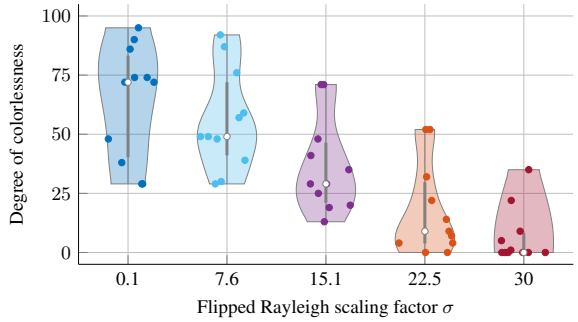
[4]Due to Covid-19 restrictions, it was not possible to conduct the listening tests in a controlled lab environment.



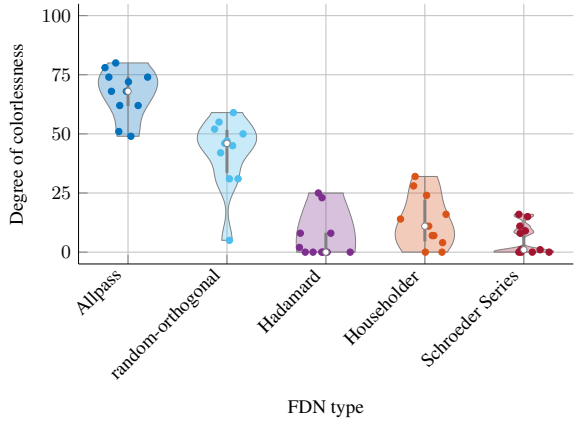
(a) *Test 1 with increasing number of modes* $\mathfrak{N}$ *between 1000 and 20000.*



(b) *Test 3 with synthetic Rayleigh distribution and* $\mathfrak{N} = 20000$.



(c) *Test 3 with synthetic flipped Rayleigh distribution and* $\mathfrak{N} = 20000$.



(d) *Test 4 for five FDNs as specified in Sec. 3.1 and* $\mathfrak{N} = 8944$.

Figure 6: *Violin plots of all four test results. Reverberation time is* T_{60} =1.44 s *and the number of modes is given by* $\mathfrak{N}$.

in the following subsections. As the listening test has been done remotely and not in a controlled environment, some problems could have been caused by the use of different headphones, and background noises in the participants' environment.

5.1. Results 1: Equal modal excitation with varying number of modes

The violin plot in Fig. 6a shows the evaluation of test samples with increasing mode number. Similarly high ratings were given to samples with 10000 and 20000 modes with medians 72 and 69, which indicates that only some participants rated them as colorless as decaying white noise. The sample with 1000 modes was identified as *certainly colored* throughout. As all participants identified the reference signal correctly as *certainly colorless*, the evaluated reference has been excluded from the violin plot. The results of the first listening test approve the correlation between increasing number of modes and perceived colorlessness, as stated by Karjalainen in [5]. $\mathfrak{N} = 3000$ was rated as *rather colored*, although the same number was rated colorless when distributed on a Bark scale [5]. Also, a full colorlessness was not achieved.

5.2. Results 2: Synthetic modal excitation distribution sampled from Rayleigh distribution

The result of the second test is shown in Fig. 6b. In accordance with the results of Test 1, the sample with $\sigma = 0.1$ achieved the highest median rating of 71, followed by the second test sample with $\sigma = 2.6$ and median 58. With further increasing σ, the samples were rated as *fairly colored / colorless* for $\sigma = 5.1$, *rather colored* for $\sigma = 7.5$ and *certainly colored* for $\sigma = 10$. Even though none of the samples achieved a median rating equal to the reference, the spread of the Rayleigh distribution clearly has an effect on the perceived coloration. Therefore, small number of modes located at higher modal excitation can cause perceived coloration in the test samples.

5.3. Results 3: Synthetic modal excitation distribution sampled from flipped Rayleigh distribution

The flipped Rayleigh distributions were identified as *rather colorless* for the most compact modal distribution of $\sigma = 0.1$ and *certainly colored* for the most spread distribution of $\sigma = 30$ in Fig. 6c. The intermediate samples with $\sigma = 7.6, 15.1$ and 22.5 were evaluated with a decreasing degree of colorlessness. Fig. 4a shows that the slope towards higher energy decreases less significantly than for the Rayleigh distributions with $x > 0$. Therefore, the number of modes in the high modal excitation region determines the coloration, rather than those with low modal excitation.

5.4. Results 4: Evaluation of FDN artificial reverberation

The ratings of the FDN impulse responses are shown in Fig. 6d. This result coincides with the prior test samples with a similarly narrow modal excitation distribution. The result for the proposed allpass FDN is a median of 68 and was therefore perceived as *rather colorless*. The random-orthogonal FDN was rated as *fairly colored / colorless*, while the remaining FDNs were perceived as *certainly colored*. Again, an evaluation as *certainly colorless* was not achieved, although it was found that the correlation between perceived coloration and modal excitation distribution envelope is applicable to FDN impulse responses as well.

6. CONCLUSIONS

In this work, a novel connection between modes-perception and allpass FDNs was proposed to improve colorless reverberation. A new correlation between the modal excitation distribution and perceived coloration in artificial reverberation was subjectively evaluated by a number of listening tests. It was found that coloration increases for flat distributions and decreases for narrow distributions. The proposed coloration measure is consistent between synthetic and FDN modal excitation distributions. We further showed that allpass filters have narrow modal excitation distribution when the modal frequencies have a low discrepancy distribution. From these results, a colorless FDN which is allpass and homogeneous was proposed, and a listening test demonstrated that it is less colored than other FDN designs. However, full colorlessness could not be achieved, which requires further investigation.

Although allpass FDNs support the design of colorless reverberation, it is neither necessary nor sufficient, see Schroeder series allpass. Instead, non-allpass FDNs can be optimized to have a narrow modal excitation distribution. Further, the current study is limited to spectral flatness and does not solve temporal roughness.

7. ACKNOWLEDGMENTS

We like to thank Thomas McKenzie, Nils Meyer-Kahlen and Michael McCrea for their helpful suggestions and many clarifications.

8. REFERENCES

[1] Vesa Välimäki, Julian D Parker, Lauri Savioja, Julius O Smith III, and Jonathan S Abel, "Fifty years of artificial reverberation," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 20, no. 5, pp. 1421 – 1448, 2012.

[2] James Anderson Moorer, "About this reverberation business," *Comput. Music J.*, vol. 3, no. 2, pp. 13 – 17, 1979.

[3] Hanna Järveläinen and Matti Karjalainen, "Reverberation Modeling Using Velvet Noise," in *Proc. Audio Eng. Soc. Conf.*, 2007, pp. 1 – 9.

[4] Vesa Välimäki and Karolina Prawda, "Late-Reverberation Synthesis using Interleaved Velvet-Noise Sequences," *IEEE/ACM Trans. Audio, Speech, Language Process.*, pp. 1149 – 1160, 2021.

[5] Matti Karjalainen and Hanna Järveläinen, "More about this reverberation science: Perceptually good late reverberation," in *Proc. Audio Eng. Soc. Conv.*, 2001, pp. 1 – 8.

[6] Michael A Gerzon, "Synthetic stereo reverberation: Part One," *Studio Sound*, vol. 13, pp. 632 – 635, 1971.

[7] Jean Marc Jot and Antoine Chaigne, "Digital delay networks for designing artificial reverberators," in *Proc. Audio Eng. Soc. Conv.*, Paris, France, 1991, pp. 1 – 12.

[8] Davide Rocchesso and Julius O Smith III, "Circulant and elliptic feedback delay networks for artificial reverberation," *IEEE Trans. on Speech and Audio Process.*, vol. 5, no. 1, pp. 51 – 63, 1997.

[9] Sebastian J Schlecht and Emanuël A P Habets, "On Lossless Feedback Delay Networks," *IEEE Trans. Signal Process.*, vol. 65, no. 6, pp. 1554 – 1564, 2016.

[10] Manfred R Schroeder and B F Logan, ""Colorless" artificial reverberation," *IRE Transactions on Audio*, vol. AU-9, no. 6, pp. 209 – 214, 1961.

[11] Sebastian J Schlecht and Emanuël A P Habets, "Feedback Delay Networks: Echo Density and Mixing Time," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 25, no. 2, pp. 374 – 383, 2016.

[12] Davide Rocchesso, "Maximally diffusive yet efficient feedback delay networks for artificial reverberation," *IEEE Signal Process. Lett.*, vol. 4, no. 9, pp. 252 – 255, 1997.

[13] Riitta Väänänen, Vesa Välimäki, Jyri Huopaniemi, and Matti Karjalainen, "Efficient and Parametric Reverberator for Room Acoustics Modeling," in *Proc. Int. Comput. Music Conf.*, Thessaloniki, Greece, 1997, pp. 200 – 203.

[14] Luke Dahl and Jean Marc Jot, "A Reverberator based on Absorbent All-pass Filters," in *Proc. Int. Conf. Digital Audio Effects (DAFx)*, Verona, Italy, 2000, pp. 1 – 6.

[15] Sebastian J Schlecht and Emanuël A P Habets, "Scattering in Feedback Delay Networks," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 28, pp. 1915 – 1924, 2020.

[16] Michael Chemistruck, Kyle Marcolini, and Will Pirkle, "Generating Matrix Coefficients for Feedback Delay Networks Using Genetic Algorithm," in *Proc. Audio Eng. Soc. Conv.*, 2012, pp. 1 – 6.

[17] Justin Shen and Ramani Duraiswami, "Data-driven feedback delay network construction for real-time virtual room acoustics," in *AM '20: Proc. Int. Conf. Audio Mostly*, 2020, pp. 46–52.

[18] Sebastian J Schlecht, "Allpass Feedback Delay Networks," *IEEE Trans. Signal Process.*, vol. 69, pp. 1028–1038, 2021.

[19] Sebastian J Schlecht, "FDNTB: The Feedback Delay Network Toolbox," in *Proc. Int. Conf. Digital Audio Effects (DAFx)*, Vienna, Austria, 2020, pp. 211–218.

[20] Nils Werner, Stefan Balke, Fabian-Rober Stöter, Meinard Müller, and Bernd Edler, "trackswitch.js: A Versatile Web-Based Audio Player for Presenting Scientifc Results," in *3rd web audio conference*, London, UK, 2017, pp. 1–6.

[21] Julius O Smith III, *Physical Audio Signal Processing*, W3K Publishing, 2010.

[22] William G Gardner, "A real-time multichannel room simulator," *J. Acoust. Soc. Am.*, vol. 92, no. 4, pp. 1 – 23, 1992.

[23] Sebastian J Schlecht and Emanuël A P Habets, "Modal Decomposition of Feedback Delay Networks," *IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5340–5351, 2019.

[24] Jonathan S Abel, Sean Coffin, and Kyle Spratt, "A Modal Architecture for Artificial Reverberation with Application to Room Acoustics Modeling," in *Eur. Signal Process. Conf. (EUSIPCO)*, 2014, pp. 1 – 10.

[25] Athanasios Papoulis and S Unnikrishna Pillai, *Probability, random variables, and stochastic processes*, McGraw-Hill Companies, 2002.

[26] ITU-R, "BS.1534 : Method for the subjective assessment of intermediate quality levels of coding systems," Tech. Rep., 2015.

[27] Emmanuel Vincent, Maria Jafari, and Mark Plumbley, "Preliminary guidelines for subjective evalutation of audio source separation algorithms," in *UK ICA Research Network Workshop*, Southampton, United Kingdom, 2006, pp. 1–4.