DATARES AND PYRES: A ROOM DATASET AND A PYTHON LIBRARY FOR REVERBERATION ENHANCEMENT SYSTEM DEVELOPMENT, EVALUATION, AND SIMULATION

Gian Marco De Bortoli

Dept. of Information and Communications Engineering
Aalto University
Espoo, FI
gian.debortoli@aalto.fi

Karolina Prawda

Dept. of Physics, Engineering, and Technology
University of York
York, UK
karolina.prawda@york.ac.uk

Philip Coleman

L-Acoustics
67 Southwood Lane, Highgate, London N6 5EG, UK
phil.coleman@l-acoustics.com

Sebastian J. Schlecht

Multimedia Communications and Signal Processing
Friederich Alexander Universität
Erlangen-Nuremberg, DE
sebastian.schlecht@fau.de

ABSTRACT

Reverberation is crucial in the acoustical design of physical spaces, especially halls for live music performances. Reverberation Enhancement Systems (RESs) are active acoustic systems that can control the reverberation properties of physical spaces, allowing them to adapt to specific acoustical needs. The performance of RESs strongly depends on the properties of the physical room and the architecture of the Digital Signal Processor (DSP). However, room-impulse-response (RIR) measurements and the DSP code from previous studies on RESs have never been made open access, leading to non-reproducible results. In this study, we present DataRES and PyRES—a RIR dataset and a Python library to increase the reproducibility of studies on RESs. The dataset contains RIRs measured in RES research and development rooms and professional music venues. The library offers classes and functionality for the development, evaluation, and simulation of RESs. The implemented DSP architectures are made differentiable, allowing their components to be trained in a machine-learning-like pipeline. The replication of previous studies by the authors shows that PyRES can become a useful tool in future research on RESs.

1. INTRODUCTION

In the acoustical design of concert halls, reverberation is one of the main features [1, 2]. As different music genres can have different requirements regarding the reverberation properties of the hall, single-purpose halls can effectively host only a limited range of performances [3, 4, 5]. This limits their usage and leads to the need for multiple halls to host the diverse range of genres present in the contemporary music scene. In most cases, the construction of several single-purpose concert halls is not a feasible solution due to the high cost [2, 5].

Passive mechanical systems such as retractable curtains, suspended or rotating panels, reverberation chambers, and moving ceilings were introduced to control the geometry or the wall absorption of the physical space, thus dynamically altering its reverberation properties [1, 3, 5, 6]. Such solutions enable concert halls to host a broader range of music genres. Reverberation Enhancement Systems (RESs) are active systems, comprising microphones, a Digital Signal Processor (DSP), and loudspeakers, that provide an alternative to passive methods [1, 2, 7, 8]. Compared to their passive counterparts, RESs can be applied instantly, even during a performance. They allow a greater degree of reverberation variation at a lower cost and are easier to install and maintain while occupying less space [7, 9].

Although RESs have been present in scientific research and on the market for more than 60 years [7, 8], the related literature and research material are scarce compared to other topics in audio and acoustics. Initially, transducers and digital algorithms were not of sufficient quality to produce natural reverberation, which gave RESs a bad reputation [7]. On the one hand, it led to a progressive decline in their commercial use; on the other hand, the low margin of improvement discouraged the interest of the research community. The advances in loudspeaker and microphone design and artificial reverberation [7, 10, 11, 12] led to an increase in the use of RESs. However, the improvements in these systems were driven mainly by the industry, and thus not openly accessible.

Several commercial RESs have been available on the market for many years [5, 7, 8]. One common characteristic among them is the need for a physically room-informed tuning during the design and installation phases [13], which demonstrates a strong impact of the physical concert hall on the resulting reverberation enhancement. The number of degrees of freedom in RESs (e.g., number and position of transducers and physical space characteristics) is high [13], driving the need for systematic and generalized approaches in the design of new DSPs and perceptual studies of the produced reverberant field. To the best of our knowledge, there are no publicly available measurement datasets or open-source software in research on RESs, resulting in a lack of openly accessible resources. This, in turn, limits the development of RESs and makes most of the results in the related literature non-reproducible.

Research in RES is divided into two categories: signal processing studies and perceptual studies. Signal processing studies focus on the virtual room. The aim is to design DSP architec-

tures that provide better and deeper control over the parameters of the generated reverberation or increase the stability limit of the system [14, 15, 16, 17]. Perceptual studies focus on the physical room and its interaction with the virtual room, including phenomena such as the double-slope decay effect [18], the coloration in the reverb [19], and the geometric distribution of the reverberant field in the physical room [20].

In this work, we present two open-access resources: **DataRES**, a dataset of room impulse responses (RIRs) measured in rooms where RESs are installed, and **PyRES**, an open-source Python library for RES development, evaluation, and simulation. DataRES offers a diverse range of spaces, spanning from intimate listening rooms to expansive concert halls, with varying volumes, geometries, reverberation properties, and audio setups. DataRES is available online [21] under the CC BY-NC-SA 2.0 licence. PyRES interfaces with the dataset and provides functionalities for developing and simulating RESs. The DSP design is implemented with FLAMO [22], thus the DSP is defined as a chain of differential modules that can be trained with a machine-learning-like pipeline to meet a predefined target. PyRES also provides computations and visualizations of metrics used for the equalization and evaluation of RESs. PyRES is available online [23] under the MIT license. Together, DataRES and PyRES address the issue of low reproducibility of results and facilitate future research on RESs.

The remainder of the paper is organized as follows. Sec. 2 provides the theoretical background on the signal model and the system stability in RESs upon which DataRES and PyRES are designed. Sec. 3 describes the structure of DataRES and PyRES. Sec. 4 shows two applications of PyRES in system equalization and feedback attenuation. Sec. 5 summarizes and concludes the work.

2. REVERBERATION ENHANCEMENT SYSTEM

RESs are multichannel audio systems consisting of microphones and loudspeakers distributed throughout a space, along with a DSP that controls the time and frequency evolution of the audio signals flowing between the transducers [7, 8]. Figure 1 shows a block diagram representation of a typical RES. The physical room is the concert hall that hosts n_{S} stage sources, n_{M} microphones, n_{L} loudspeakers, and n_{A} audience positions. The virtual room is a DSP that implements a virtual space, enhancing the sound field in the physical room.

The RIRs in the physical room are as follows: $\mathbf{H}_{\mathrm{SA}}[t] \in \mathbb{R}^{n_{\mathrm{A}} \times n_{\mathrm{S}}}$ between the stage sources and the audience positions; $\mathbf{H}_{\mathrm{SM}}[t] \in \mathbb{R}^{n_{\mathrm{M}} \times n_{\mathrm{S}}}$ between the stage sources and the microphones; $\mathbf{H}_{\mathrm{LM}}[t] \in \mathbb{R}^{n_{\mathrm{M}} \times n_{\mathrm{L}}}$ between the loudspeakers and the microphones; $\mathbf{H}_{\mathrm{LA}}[t] \in \mathbb{R}^{n_{\mathrm{A}} \times n_{\mathrm{L}}}$ between the loudspeakers and the audience positions; t is the discrete time variable. The feedforward path between the microphones and the loudspeakers contains the virtual room, which is a matrix of impulse responses (IRs) $\mathbf{V}_{\mathrm{ML}}[t] \in \mathbb{R}^{n_{\mathrm{L}} \times n_{\mathrm{M}}}$, and the system gain $G \in \mathbb{R}^{+}$.

By applying the z-Transform to all the RIR matrices, the complete RES transfer function (TF) $\mathbf{H}_{\mathrm{RES}}(z)$ is defined in the complex domain as

$$\begin{aligned} \mathbf{H}_{\mathrm{RES}}(z) &= \mathbf{H}_{\mathrm{SA}}(z) + \\ &+ \mathbf{H}_{\mathrm{LA}}(z)\big(\mathbf{I} - \mathrm{G}\mathbf{V}_{\mathrm{ML}}(z)\mathbf{H}_{\mathrm{LM}}(z)\big)^{-1}\mathrm{G}\mathbf{V}_{\mathrm{ML}}(z)\mathbf{H}_{\mathrm{SM}}(z), \end{aligned} \tag{1}$$

where z is the complex variable. $\mathbf{H}_{\mathrm{RES}}(z)$ is the summation of $\mathbf{H}_{\mathrm{SA}}(z)$, the physical-room acoustical path, and the electro-acoustic path. $\mathbf{H}_{\mathrm{LM}}(z)$ is the feedback component of the system which makes Eq. (1) recursive. Consequently, RESs are prone to instability [7, 8] caused by a too high gain in the feedback loop, which results in the energy of the looping signal increasing at each loop iteration. This is dangerous for both people and equipment and must be avoided. Assuming that the system gain G has a magnitude of one, the feedback loop gain is a function of frequency that depends on the interaction between $\mathbf{H}_{\mathrm{LM}}$ and $\mathbf{V}_{\mathrm{ML}}$. Thus, the upper gain limit ensuring stability is different for each frequency. The lowest value of the upper gain limit across frequencies is defined as the Gain Before Instability (GBI) [8]. To ensure system stability, $G <$ GBI.

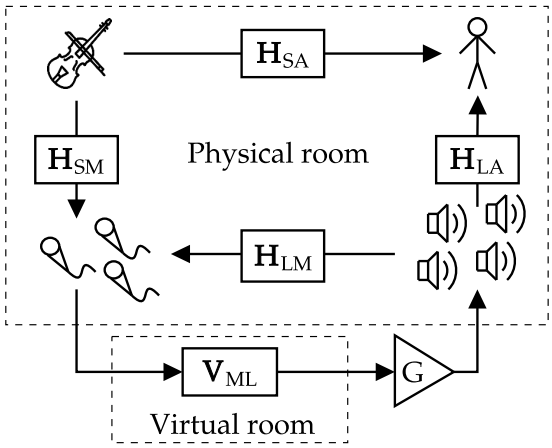


Figure 1: Block diagram of the signal flow in a typical RES. The signal routing is depicted by solid arrows with annotations of the involved TFs.

In the literature, there is no direct method for computing the GBI; however, three estimation methods are identified: statistical, deterministic, and empirical. The former two consider the open-loop TF matrix $\mathbf{O}_{\mathrm{MM}}(z) \in \mathbb{C}^{n_{\mathrm{M}} \times n_{\mathrm{M}}}$

$$\mathbf{O}_{\mathrm{MM}}(z) = \mathrm{G}\mathbf{H}_{\mathrm{LM}}(z)\mathbf{V}_{\mathrm{ML}}(z). \tag{2}$$

The statistical approach analyses the statistical distribution of either the real part or the magnitude of $\mathbf{O}_{\mathrm{MM}}(z)$ and computes the risk of instability as a variable of G. The GBI is then estimated by taking the value of G with 50% risk of instability [24].

The deterministic approach applies eigen-decomposition to $\mathbf{O}_{\mathrm{MM}}(z)$:

$$\mathbf{O}_{\mathrm{MM}}(z) = \mathbf{Q}(z)\Lambda(z)\mathbf{Q}^{-1}(z), \tag{3}$$

where $\Lambda(z) \in \mathbb{C}^{n_{\mathrm{M}} \times n_{\mathrm{M}}}$ is a complex diagonal matrix containing the open-loop eigenvalues and $\mathbf{Q}(z) \in \mathbb{C}^{n_{\mathrm{M}} \times n_{\mathrm{M}}}$ is a complex square matrix containing the open-loop eigenvectors along its columns. At each frequency, the upper gain limit is estimated to be the eigenvalue having either the maximum magnitude or the maximum real part [16, 24, 25].

The empirical approach analyzes Eq. (1). Since the room TFs (RTFs) not involved in the feedback loop do not affect the stability of the system, the analysis is often restricted to the closed-loop TF matrix $\mathbf{C}_{\mathrm{ML}}(z) \in \mathbb{C}^{n_{\mathrm{L}} \times n_{\mathrm{M}}}$

$$\mathbf{C}_{\mathrm{ML}}(z) = \big(\mathbf{I} - \mathrm{G}\mathbf{V}_{\mathrm{ML}}(z)\mathbf{H}_{\mathrm{LM}}(z)\big)^{-1}\mathrm{G}\mathbf{V}_{\mathrm{ML}}(z). \tag{4}$$

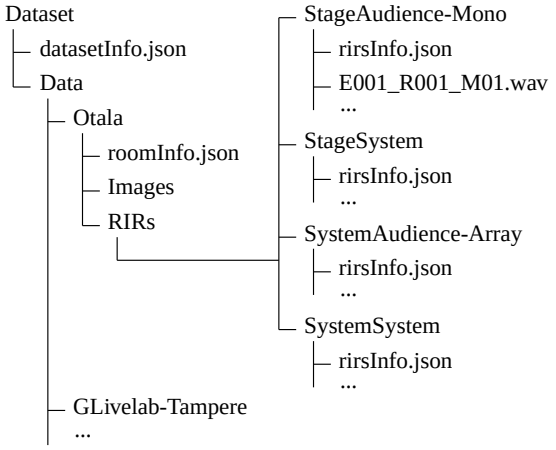


Figure 2: DataRES file structure.

The empirical approach then computes Eq. (4) for multiple values of G [26, 27]. Spectral analysis is then applied to $\mathbf{C}_{\mathrm{ML}}(z)$ to detect which frequencies are the most likely to cause system instability. Due to the inherent time-varying behavior of the RTFs of the physical room [28, 29, 30], all the estimations described tend to diverge from the true GBI.

When the system is stable, some frequencies will still decay faster than others because $\mathbf{O}_{\mathrm{MM}}(z)$ is a function of frequency. Therefore, when G is too close to instability, strong coloration artifacts arise in the form of long-ringing tones at frequencies for which the upper gain limit is close to the GBI [17, 31]. This further limits the gain value at which the system can operate.

3. DATARES AND PYRES

In this section, we present DataRES and PyRES. Their structure and functionality are described using the signal model explained in Sec. 2.

3.1. DataRES

DataRES comprises RIRs measured in rooms that host a RES. The rooms are used for research and development on RESs, either in the academic or industrial environment, or for hosting live performance events. The spaces differ in volume, geometry, and reverberation properties, and the multichannel audio setups differ in the number and position of transducers. Thus, the dataset serves as a selection of different conditions suitable for testing new virtual-room designs and evaluating perceptual metrics.

The structure of the dataset is shown in Fig. 2. The data is organized into folders for each room, e.g., *GLivelab-Tampere*, *Otala*, each including an *Images* and a *RIRs* subfolder. All room folders contain the same type of information, except when confidentiality issues prevent it.

The *Images* folder contains images of the room in PNG, JPEG, or PDF format. The image content ranges from floor plans and audio setup schemes to room photos. They provide information on the geometry of the physical space and the multichannel audio system. The photos are in 360-degree view in case the measurements are taken with microphone arrays or binaural heads, thus allowing for the use of the dataset in virtual reality applications.

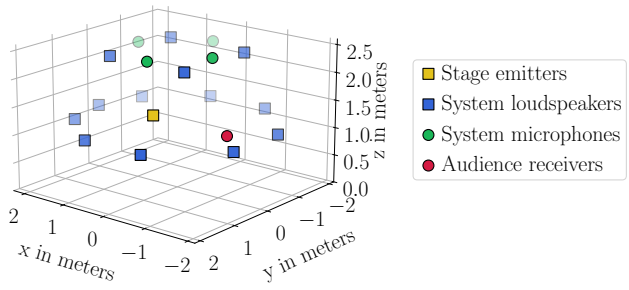
The *RIRs* folder contains the measured RIRs grouped according to Fig. 1: *StageAudience* contains $\mathbf{H}_{\mathrm{SA}}$, *StageSystem* contains $\mathbf{H}_{\mathrm{SM}}$, *SystemAudience* contains $\mathbf{H}_{\mathrm{LA}}$, and *SystemSystem* contains $\mathbf{H}_{\mathrm{LM}}$. The labels *-Mono*, *-Array*, and *-Binaural* in the folder names specify the type of microphone used for the measurements. The RIRs are stored in WAV format. Each RIR measurement file contains the RIR of one emitter-receiver pair, where emitter stands for stage source in $\mathbf{H}_{\mathrm{SA}}$ and $\mathbf{H}_{\mathrm{SM}}$ and for system loudspeaker in $\mathbf{H}_{\mathrm{LA}}$ and $\mathbf{H}_{\mathrm{LM}}$, and receiver stands for audience position in $\mathbf{H}_{\mathrm{SA}}$ and $\mathbf{H}_{\mathrm{LA}}$ and for system microphone in $\mathbf{H}_{\mathrm{SM}}$ and $\mathbf{H}_{\mathrm{LM}}$. The filename pattern of each RIR measurement file is *Exxx_Ryyy_Mzz.wav*, where *xxx* is the number of the emitter, *yyy* is the number of the receiver, and *zz* is a number ranging from one to the number of measurements that were taken for the *xxx-yyy* pair. The measurements between system loudspeakers and system microphones were taken using the audio setup equipment, hence directionality and polar patterns were given by the setup. In RESs, the system microphones are always omnidirectional or cardioid mono-channel microphones. Thus, $\mathbf{H}_{\mathrm{SM}}$ and $\mathbf{H}_{\mathrm{LM}}$ are mono-channel audio files, and the label *-Mono* in the folder name is omitted. When recording the audience positions, instead, multiple types of receivers were considered: omnidirectional mono-channel microphones to simplify the use of the dataset for simulations; microphone arrays for spatial analysis of the reverberant field; binaural heads for fast binauralization of the RES IR and rendering in virtual reality applications.

JSON files provide information about the stored data. *datasetInfo.json* contains high-level information about each room in the dataset: use of the room, volume, reverberation time in octave bands, number of transducers in the audio setup, temperature, relative humidity, level of room occupancy, and additional occlusion of one or more system microphones at recording time, a bool value indicating if microphone arrays or binaural heads were used for the audience positions, and the directory to the room folder. Each room folder contains a *roomInfo.json* file that holds low-level information about the room: image directory and filenames; measurement procedure (i.e., which method was used and further details depending on the method), number and positioning of stage emitters, audience receivers, and system transducers, RIRs encoding information (i.e., time-sampling rate, length of the audio files in samples, and number of bits per sample), and directories. Finally, in each *RIRs* subfolder, *rirsInfo.json* file contains the measurement filenames.

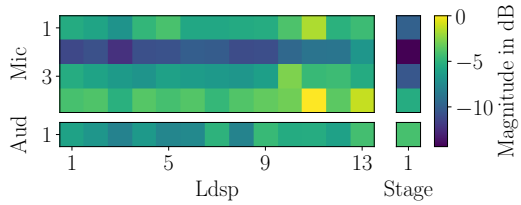
Currently, DataRES includes measurements from four different rooms: two research and development rooms and two live music clubs. Table 1 shows the current list of rooms and their attributes. *Otala* is a listening room at Aalto University, Espoo, Finland. *ImmersiveLab* is a space used for research and development on RESs. *GLivelab-Helsinki* and *GLivelab-Tampere* are two live

Table 1: List of rooms contained in the dataset. For each room, volume (V) in m^3, average reverberation time (RT) in s, and number of system loudspeakers and microphones are reported.

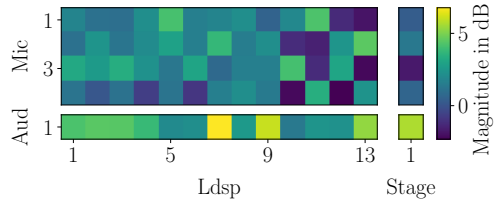
Room	$V\,[m^3]$	$RT\,[s]$	n_{L}	n_{M}
Otala	96	0.27	13	4
ImmersiveLab	800	0.86	32	16
GLivelab-Helsinki	700	0.51	42	6
GLivelab-Tampere	2600	0.44	41	8



(a) Scatter plot of the system audio setup.

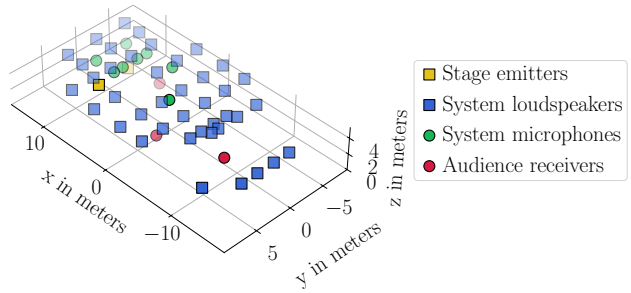


(b) Matrix containing the energy of the RIR for each emitter-receiver pair.

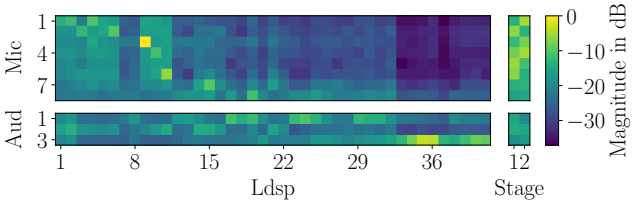


(c) Matrix containing the direct-to-reverberant ratio of the RIR for each emitter-receiver pair.

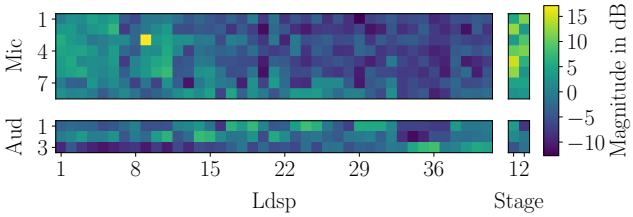
Figure 3: Research and development room Otala, Aalto University, Espoo, Finland.



(a) Scatter plot of the system audio setup.



(b) Matrix containing the energy of the RIR for each emitter-receiver pair.



(c) Matrix containing the direct-to-reverberant ratio of the RIR for each emitter-receiver pair.

Figure 4: Live music club GLivelab, Tampere, Finland.

music clubs in Helsinki and Tampere, respectively, that primarily host musical performances in genres such as rock, jazz, and blues. The rooms are always recorded under default conditions: the room is empty, the temperature and humidity follow the air conditioning schedule, and the transducers are positioned as during system operation. To facilitate studies on DSP robustness against RIR time variation, the dataset also provides measurements of some rooms under non-default conditions. If the RIRs are measured after one or more conditions were altered, a label is added to the name of the respective folder: *-Tn* for changes in the atmospheric conditions, *-Pn* for changes in the positions of the system transducers, *-Cn* for changes in the room occupancy, and *-Ln* means that one or more system microphones were occluded during the measurement. In each label, *n* is an integer iterator that indicates how many times the room was recorded for the same type of condition alteration.

Figures 3 and 4 illustrate the data from *Otala* and *GLivelab-Tampere*, respectively. The figures are produced with PyRES by accessing the information in the dataset. Figures 3a and 4a show schemes of transducer setup, with emitters marked with squares (stage sources in green and system loudspeakers in blue), and receivers indicated with circles (system microphones in red and audience positions in yellow). The markers' level of opacity decreases with the growing distance from the observer. Figures 3b and 4b show matrices of the RIR energy for each emitter-receiver pair. Figures 3c and 4c show matrices of the RIR direct-to-reverberant ratios of each emitter-receiver pair.

3.2. PyRES

PyRES is a Python library that implements the signal flow depicted in Fig. 1 in a modular way, thus granting independent control over the physical room, the virtual room, and the interaction between the two. PyTorch [32], FLAMO [22], and pyfar [33] Python libraries and packages are used as the back-end. PyRES serves as a platform for the development and simulation of RESs.

PyRES is composed of three main classes: `PhRoom`, `VrRoom`, and `RES`. `PhRoom` is a base class used to implement the physical room. Its subclass `PhRoom_dataset` interfaces with the dataset, loads the information of the requested room, and stores the RIRs in four separate variables: `h_SA`, `h_SM`, `h_LM`, and `h_LA`.

`VrRoom` is a base class used to implement the virtual room. It receives the number of system microphones and loudspeakers, defining its input-output connections. Its subclasses are DSP architectures defined through FLAMO [22] as chains of DSP modules. Currently, PyRES includes five DSP architectures:

- a truncated unitary random mixing matrix
- a random finite-impulse-response (FIR) filter mixing matrix [31]
- a multichannel phase-cancelling modal reverberator [17]
- a feedback delay network (FDN)
- a multichannel unitary reverberator [34]

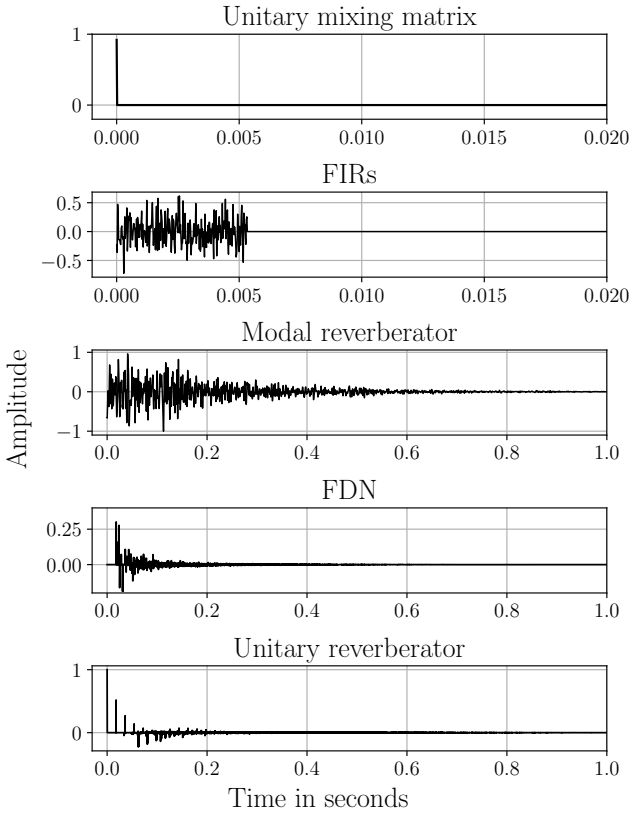


Figure 5: Virtual-room designs in PyRES.

Figure 5 IRs of these DSP architectures. The unitary mixing matrix and the FIRs have short IRs and are suitable for system equalization (see Sec. 4.1). The modal reverberator, the FDN, and the unitary reverberator have long IRs suitable for reverberant field control (see Sec 4.2).

Since the DSP is implemented using the FLAMO library [22], each processing module of the architecture is differentiable, and thus the parameters that define its behavior are trainable in a machine-learning pipeline. Figure 6 shows how PyRES handles the training of the DSPs. `PhRoom` and `VrRoom` are used to define the model, i.e., open or closed loop. The loss function is defined according to the optimization target (see Sec. 4). The training dataset and optimization algorithm are handled by FLAMO.

`RES` receives an instance from each, `PhRoom` and `VrRoom`, as input arguments and combines them in a RES. It allows control and simulation of the system by implementing the full signal flow shown in Fig. 1. The most important methods of the `RES` class are:

- `open_loop()`, which returns the open-loop from Eq. (2) as a chain of FLAMO modules
- `open_loop_eigenvalues()`, which returns the complex eigenvalues of the open-loop TF matrix as in Eq. (3)
- `estimate_GBI()`, which estimates the GBI based on either the real part or the magnitude of the open-loop eigenvalues
- `closed_loop()`, which returns the closed-loop from Eq. (4) as a chain of FLAMO modules
- `system_simulation()`, which returns the complete RES IRs and TFs (see Eq. (1))

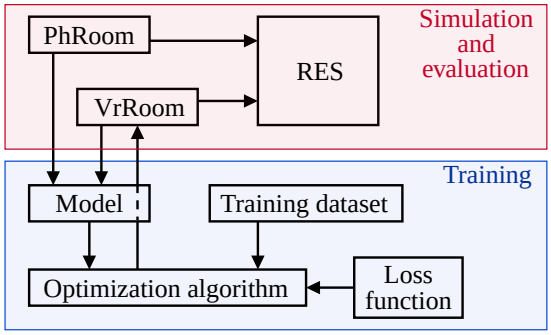


Figure 6: PyRES pipelines for the training of the DSP and the simulation and evaluation of the RES.

Figure 7 depicts the output of the `system_simulation()` method for three physical rooms from the dataset and all DSP architectures from PyRES. For each combination, Fig. 7 shows a box plot of the magnitude distribution of the open-loop eigenvalue set (top) and the IR spectrogram at an audience position (bottom) for the corresponding RES with $G = \text{GBI} - 2$ dB. In the box plot, the mean is represented as a thick black line. The left and right whiskers represent the 25th and 75th percentiles, respectively. The dot at -2 dB in each boxplot indicates the maximum value in the distribution. Outliers are not shown because they do not add information about system stability or spectral flatness.

For each RES, the virtual room affects the variance of the eigenvalue magnitude distribution, while the physical room affects the difference between the mean and the maximum value. The spectrograms show that the physical room determines which frequencies are affected by coloration, while the reverberation time of the response is affected by both physical and virtual rooms. In the last row of Fig. 7, the frequency range is 0–500 Hz because the phase-canceling modal reverberator is designed to work in this interval [17].

4. APPLICATIONS

In this section, two of the authors' previous studies are revisited using PyRES. These studies belong to the category of signal processing and system stability.

4.1. System equalization

The optimization of a mixing matrix of FIR filters has been proposed to obtain spectral flatness in the feedback loop [31]. With the same approach, RES equalization can be achieved by optimizing the FIR filters towards a desired equalization curve. In FLAMO, the coefficients of an FIR filter are defined as a set of learnable parameters for PyTorch. In PyRES, the mixing matrix of FIR filters is an instance of the multi-input multi-output `Filter` class from FLAMO [22], containing one FIR filter for each microphone-loudspeaker pair. Thus, the training pipeline presented in Fig. 6 can be used to find the optimal set of coefficients that flatten the magnitude distribution of the open-loop eigenvalues. This leads to a distribution of energy across frequencies and channels that follows the target of optimization. The coefficients of the FIR filters were initialized by drawing from the standard normal distribution.

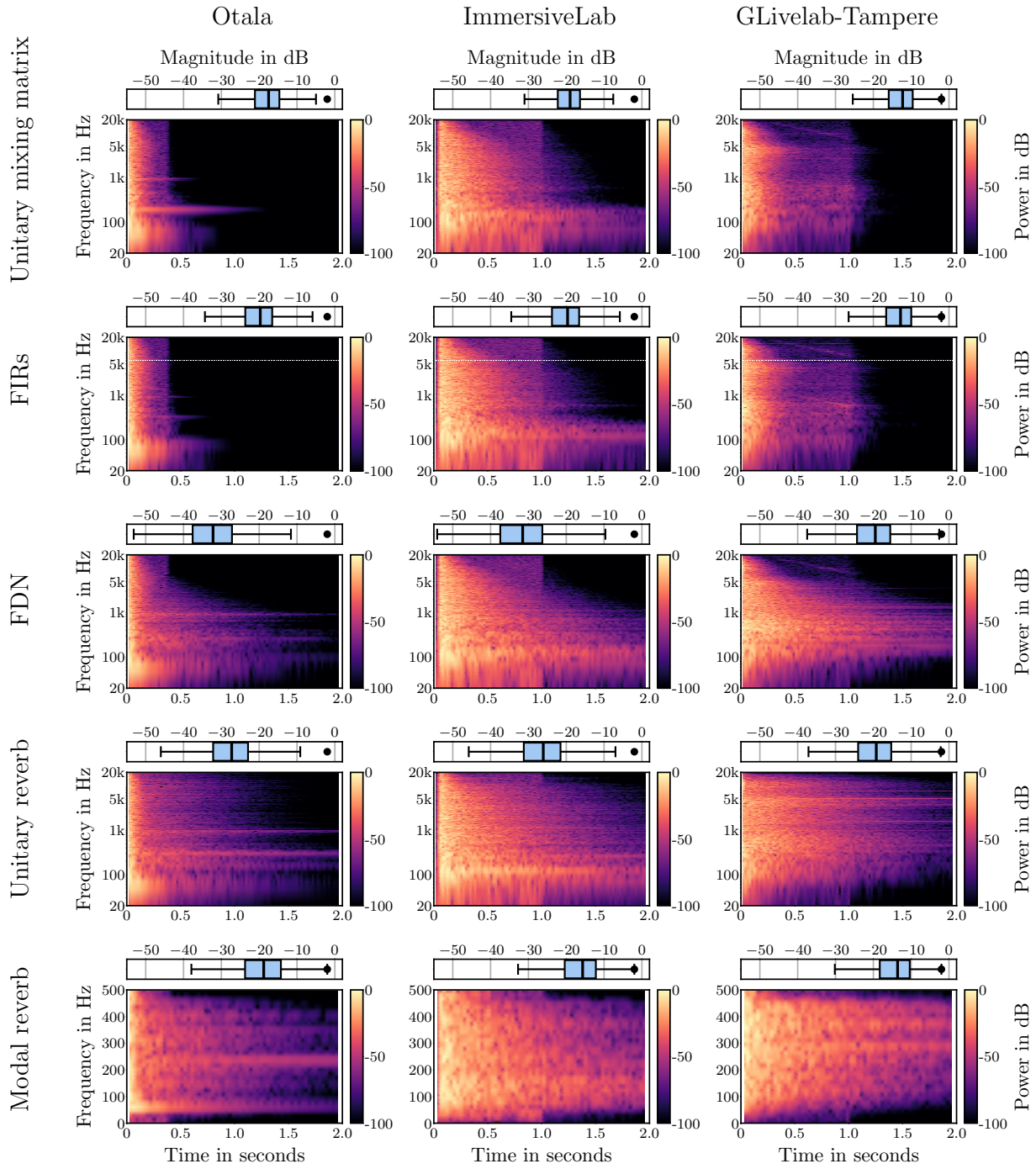


Figure 7: Eigenvalue magnitude distribution (top) and spectrogram of a RES IR (bottom) for each physical room-virtual room pair.

The equalization curve used as the optimization target was designed as a flat response in the range 20–10k Hz, followed by the mean magnitude of the eigenvalues per frequency bin, smoothed with a moving average. The loss function was the mean squared error between the equalization curve and the magnitude of the open-loop eigenvalues. Before the optimization, the system gain was set to $G = \mathrm{GBI} - 1$ to show strong coloration, and was maintained after optimization.

The left side of Fig. 8 shows the comparison of the magnitude distribution of the open-loop eigenvalues in the range 20–10k Hz before (Init) and after (Opt) optimization. The right side of Fig. 8 shows the comparison of the spectrogram of the complete system IR (see Eq. (1)) at an audience position before (Init) and after (Opt) optimization. The magnitude of the eigenvalues has converged to the target response, thus redistributing the energy in the open-loop evenly across frequencies. This leads to a suppression of the ringing tones at $\approx$ 300 Hz and $\approx$ 900 Hz and a reduction in the coloration below 100 Hz.

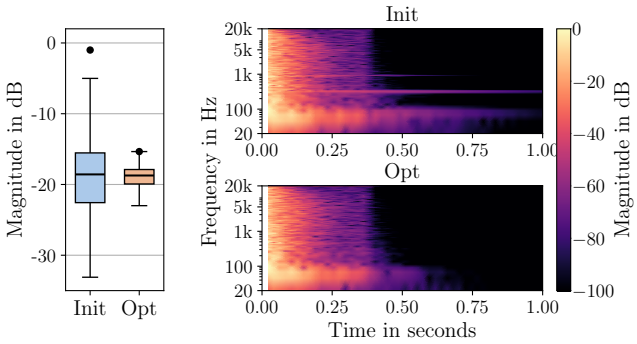


Figure 8: Magnitude distributions of the open-loop eigenvalues (left) for the RES in Otala with FIR filters at initialization stage (blue) and after optimization (orange) and spectrograms of the complete RES IR at an audience position (right) for the RES at initialization stage (top) and after optimization (bottom).

4.2. Feedback attenuation

Feedback attenuation improves the stability of RESs. By lowering the contribution of the system loudspeakers to the sound energy captured by the system microphones, the amount of energy regenerated by the feedback loop decreases, leading to an increase in GBI. PyRES includes the implementation of a phase-canceling modal reverberator, which controls the synthetic reverberation in a RES while achieving feedback attenuation [17].

Phase cancellation is produced at the resonance frequencies of the modal reverberator to reduce the energy in the feedback loop without altering the energy of the reverberator IRs. In PyRES, this is achieved by optimizing the phases of the resonance filters [35] within the DSP architecture using the training pipeline depicted in Fig. 6. For this example, the reverberator had 120 resonance filters, all with a reverberation time of 1 s. The phases were initialized by drawing from a uniform distribution in the range 0–2π rad. The loss function consisted of two components: the magnitude of the open-loop eigenvalues at the resonances of the modal reverberator and the mean squared error between the peaks in the magnitude responses of the modal reverberator before and after optimization. Thus, the aim was to minimize the open-loop eigenvalues and preserve the energy in the reverberator IRs. Before optimization, the system gain was set to $G = \mathrm{GBI}_{init} = -19.62$ dB and was maitained after optimization. The increase in GBI was $\mathrm{GBI}_{opt} - \mathrm{GBI}_{init} = 16.75$ dB, showing an improvement of 11.25 dB from the original results [17], which is attributed to a more sophisticated training procedure and loss function, implemented in PyRES.

Figure 9 presents a similar comparison as in Fig. 8 for the phase-canceling modal reverberator. The box plots show a decrease in the variance of the eigenvalue magnitude distribution. The top spectrogram shows strong coloration due to a too high gain G. The optimized system in the bottom spectrogram is far from instability, and thus the reverberation time profile is flatter and closer to the desired value of 1 s.

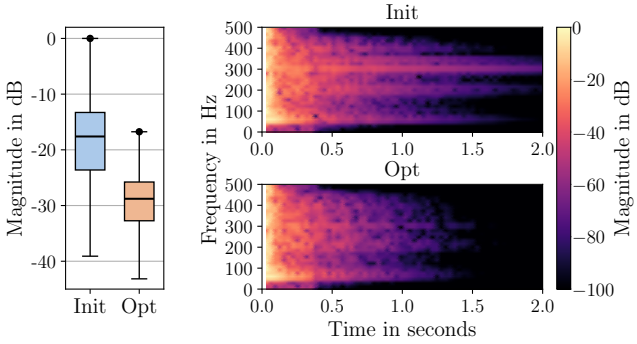


Figure 9: Magnitude distributions of the open-loop eigenvalues (left) for the RES in Otala with a phase-canceling modal reverberator at initialization stage (blue) and after optimization and system gain increment (orange) and spectrograms of the complete RES IR at an audience position (right) for the RES at initialization stage (top) and after optimization and system gain increment (bottom).

5. CONCLUSIONS

In this work, we presented DataRES and PyRES. DataRES is a dataset of RIR measurements from different rooms, providing multiple conditions for RES studies. PyRES is a Python library that interfaces with the dataset and allows the development and optimization of virtual-room designs, the simulation and control of RESs, and the evaluation of the enhanced reverberant field. Together, they promote the transparency and reproducibility of future studies on RESs.

The application of PyRES to the authors' previous work has demonstrated that the dataset and library can be utilized not only to replicate results from previous studies but also to generate new research on RESs. A comparison of several RES implementations has shown that the physical room and virtual room play distinct roles in defining the system eigenvalues and the enhanced reverberant field.

6. ACKNOWLEDGMENTS

The work of the first author was funded by the Academy of Finland, project No. 357391.

7. REFERENCES

[1] H. Kuttruff, *Room Acoustics*, CRC Press, Boca Raton, Florida, USA, sixth edition, 2017.

[2] M. Barron, *Auditorium Acoustics and Architectural Design*, Spon Press, Abingdon, Oxford, UK, sixth edition, 2010.

[3] P. W. Barnett, "A review of reverberation enhancement systems," in *6th Audio Eng. Soc. Conf.: Sound Reinforcement*, 1988.

[4] A. C. Gade, *Springer Handbook of Acoustics*, chapter Acoustics in Halls for Speech and Music, pp. 317–366, Springer New York, New York, NY, USA, 2014.

[5] J. Hyon and D. Jeong, "Variable acoustics in performance venues - a review," *J. Acoust. Soc. Korea*, vol. 40, no. 6, pp. 626–648, 2021.

[6] K. Prawda, S. J. Schlecht, and V. Välimäki, "Calibrating the Sabine and Eyring formulas," *J. Acoust. Soc. Am.*, vol. 152, no. 2, pp. 1158–1169, 2022.

[7] M. A. Poletti, "Active acoustic systems for the control of room acoustics," *Building Acoustics*, vol. 18, no. 3-4, pp. 237–258, 2011.

[8] P. U. Svensson, *On Reverberation Enhancement in Auditoria*, Ph.D. thesis, Chalmers University of Technology, 1994.

[9] R. W. Schwenke and J. R. Duty, "Electro-acoustic architecture: Is it green?," *J. Acoust. Soc. Am.*, vol. 127, no. 3, pp. 1724, 2010.

[10] V. Välimäki, J. Parker, L. Savioja, J. O. Smith, , and J. Abel, "Fifty years of artificial reverberation," *IEEE Trans. Audio, Speech Lang. Process.*, vol. 20, no. 5, pp. 1421–1448, 2012.

[11] V. Välimäki, J. Parker, L. Savioja, J. O. Smith, , and J. Abel, "More than 50 years of artificial reverberation," in *60th Audio Eng. Soc. Int. Conf.: DREAMS*, 2016.

[12] T. Watanabe, D. Hashimoto, H. Miyazaki, and R. Bakker, "Comparison of effectiveness of acoustic enhancement systems—comparison of in-line," in *Audio Eng. Soc. Conv. 144*, 2018.

[13] M. Nagata, "Active sound field control systems in auditoriums - expectations and precautions," *J. Acoust. Soc. Jpn. (E)*, vol. 12, no. 6, pp. 283–290, 1991.

[14] J. Abel, E. F. Callery, and E. K. Canfield-Dafilou, "A feedback canceling reverberator," in *Proc. Int. Conf. Digital Audio Effects*, 2018.

[15] S. J. Schlecht and E. A. P. Habets, "Reverberation enhancement systems with time-varying mixing matrices," in *59th Audio Eng. Soc. Int. Conf.: Sound Reinforcement Engineering and Technology*, 2015.

[16] S. J. Schlecht and E. A. P. Habets, "The stability of multichannel sound systems with time-varying mixing matrices," *J. Acoust. Soc. Am.*, vol. 140, no. 1, pp. 601–609, 2016.

[17] G. De Bortoli, K. Prawda, and S. J. Schlecht, "Active acoustics with a phase cancelling modal reverberator," *J. Audio Eng. Soc.*, vol. 72, no. 10, pp. 705–715, 2024.

[18] W. J. Cassidy, P. Coleman, R. Mason, and E. De Sena, "Naturalness of double-slope decay in generalised active acoustic enhancement systems," in *Proc. Int. Conf. Digital Audio Effects*. University of Surrey, 2024, pp. 262–269.

[19] P. Coleman, N. Epain, S. Venkatesh, and F. Roskam, "Exploring perceptual annoyance and colouration assessment in active acoustic environments," in *Audio Eng. Soc. Conf.: Acoustics & Sound Reinforcement*, 2024.

[20] V. Werner, S. Neeten, and F. Kaiser, "Evaluation of a geometric approach to active acoustics," *Proc. Inst. Acoust*, vol. 45, no. 2, 2023.

[21] G. De Bortoli, K. Prawda, P. Coleman, and S. J. Schlecht, "DataRES: Dataset for Research on Reverberation Enhancement Systems (2.0.0) [data set]. Zenodo," Accessed April 06, 2025, https://doi.org/10.5281/zenodo.15737243.

[22] G. Dal Santo, G. De Bortoli, K. Prawda, S. J. Schlecht, and V. Välimäki, "FLAMO: An open-source library for frequency-domain differentiable audio processing," in *Proc. ICASSP*, 2025, pp. 1–5.

[23] G. De Bortoli, "PyRES GitHub Repository," Accessed April 06, 2025, https://github.com/GianMarcoDeBortoli/PyRES.git.

[24] M. A. Poletti, "The stability of single and multichannel sound systems," *Acta Acust. un. Acust.*, vol. 86, no. 1, pp. 163–178, 2000.

[25] M. Ohsmann, "Analyse von mehrkanalanlagen," *Acta Acust. un. Acust.*, vol. 70, no. 4, pp. 233–246, 1990.

[26] M. R. Schroeder, "Improvement of acoustic-feedback stability by frequency shifting," *J. Acoust. Soc. Am.*, vol. 36, no. 9, pp. 1718–1724, 1964.

[27] P. U. Svensson, "Computer simulations of periodically time-varying filters for acoustic feedback control," *J. Audio Eng. Soc.*, vol. 43, no. 9, pp. 667–677, 1995.

[28] W. K. Connor, "Experimental investigation of sound-system-room feedback," *J. Audio Eng. Soc.*, vol. 21, no. 1, pp. 27–32, 1973.

[29] K. Prawda, S. J. Schlecht, and V. Välimäki, "Short-time Coherence Between Repeated Room Impulse Response Measurements," *J. Acoust. Soc. Am.*, vol. 156, no. 2, pp. 1017—-1028, 2024.

[30] K. Prawda, "Sensitivity of room impulse responses in changing acoustic environment," in *Proc. ICASSP*, Hyderabad, India, 2025.

[31] G. De Bortoli, G. Dal Santo, K. Prawda, T. Lokki, V. Välimäki, and S. J. Schlecht, "Differentiable active acoustics: Optimizing stability via gradient descent," in *Proc. Int. Conf. Digital Audio Effects*, 2024, pp. 254–261.

[32] A. Paszke, S. Gross, F. Massa, et al., "Pytorch: An imperative style, high-performance deep learning library," *Adv. Neural Inf. Process. Syst.*, vol. 32, 2019.

[33] "Pyfar, a python package for acoustics research," Available at http://pyfar.org, Accessed April 01, 2025.

[34] M. A. Poletti, "A unitary reverberator for reduced colouration in assisted reverberation systems," in *Proc. INTER-NOISE and NOISE-CON Cong. and Conf.*, 1995, vol. 3, pp. 1223–1232.

[35] J. S. Abel, S. Coffin, and K. Spratt, "A modal architecture for artificial reverberation with application to room acoustics modeling," in *Audio Eng. Soc. Conv. 137*, Los Angeles, USA, Oct. 2014.