# Dimensionality Reduction of Room Acoustic Impulse Responses and Its Applications

Project-I (EE47002) report submitted to

Indian Institute of Technology Kharagpur

in partial fulfilment for the award of the degree of

Bachelor of Technology

in

Electrical Engineering

**by**

**Rudransh Gupta**

**(20EE38023)**

**Under the supervision of**

**Professor Anirban Mukherjee**

**Department of Electrical Engineering**

**Indian Institute of Technology Kharagpur**

**Autumn Semester, 2023-24**

**November 28, 2023**

**DEPARTMENT OF ELECTRICAL ENGINEERING**

**INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR**

**KHARAGPUR - 721302, INDIA**


## *CERTIFICATE*

This is to certify that the project report entitled "**Dimensionality Reduction of Room Acoustic Impulse Responses and Its Applications**" submitted by **Rudransh Gupta** (Roll No. 20EE38023) to Indian Institute of Technology Kharagpur towards partial fulfilment of requirements for the award of degree of Bachelor of Technology in Electrical Engineering is a record of bona fide work carried out by him under my supervision and guidance during Autumn Semester, 2023-24.

Date: November 28, 2023
Place: Kharagpur


Professor Anirban Mukherjee
Department of Electrical Engineering
Indian Institute of Technology Kharagpur
Kharagpur - 721302, India

# *Abstract*


Name of the student: **Rudransh Gupta** Roll No: **20EE38023**

Degree for which submitted: **Bachelor of Technology**

Department: **Department of Electrical Engineering**

Thesis title: **Dimensionality Reduction of Room Acoustic Impulse Responses and Its Applications**

Thesis supervisor: **Professor Anirban Mukherjee**

Month and year of thesis submission: **November 28, 2023**



The complexities of Room Acoustic Impulse Responses (RAIR), are pivotal in various acoustic signal processing applications like echo cancellation. RAIR characterizes how sound travels in an environment, considering both direct and indirect (reflected) sound paths from a source to a microphone. A notable challenge in real-world acoustic settings is the extensive nature of RAIRs, often comprising hundreds or thousands of elements, complicating their identification and management.

The focus of this study is to explore ways to simplify RAIRs by reducing their dimensionality, drawing inspiration from dynamic mode decomposition concepts. The goal is to develop a more manageable, lower-dimensional format for RAIRs, which would be simpler to identify and adjust for better acoustic results.

This study introduces a novel method for reducing the complexity of RAIRs and its application in acoustic system identification. Simulations demonstrate that this approach significantly improves system identification and tuning, providing an effective solution to a prevalent issue in acoustic signal processing.

# *Acknowledgements*

I would like to express my sincere gratitude to all those who have contributed to the completion of this work. Special thanks go to Professor Anirban Mukherjee for their invaluable guidance, support, and insightful feedback throughout the entire process.

I am also thankful for the Indian Institute of Technology Kharagpur for providing the necessary resources and environment for this research. I extend my appreciation to my colleagues and peers who have shared their expertise and perspectives, enriching the depth of this project.

I am grateful to my friends and family for their unwavering encouragement and understanding during the challenges and demands of this endeavor. Their support has been a constant source of motivation.

Lastly, I want to express my appreciation to the broader academic and research community whose work has laid the foundation for this study. The collaborative spirit and shared knowledge within this community have significantly contributed to the advancement of our understanding in this field.

Thank you all for being an integral part of this journey.

# Contents

# Chapter 1

# Introduction

## 1.1 Overview

Room Acoustic Impulse Responses (RAIRs) are critical in numerous applications within the field of acoustic signal processing. These applications range from system identification and acoustic echo cancellation to dereverberation, source separation, and beamforming. In realistic acoustic environments, RAIRs are often composed of a substantial number of coefficients, frequently reaching into the hundreds or thousands. This extensive nature makes RAIRs particularly challenging to identify and adjust, especially in environments with significant noise interference. To address this, there's a growing need to streamline RAIRs through a process known as RAIR dimensionality reduction. This process aims to condense RAIRs into more manageable, shorter forms for easier manipulation.

This study delves into the challenge of reducing RAIR dimensionality. We introduce a technique inspired by dynamic mode decomposition (DMD), a method originally formulated for analyzing fluid flow dynamics over time. DMD has since been adapted for various fields as an effective tool for dimensionality reduction. The core of DMD, akin to the eigensystem realization algorithm (ERA) used in control theory, lies in its ability to transform high-dimensional data into a concise, lower-dimensional format. This transformation is particularly effective as it retains the essential physical characteristics inherent in the data sequence. Leveraging this capability, we propose a novel approach to RAIR dimensionality reduction using DMD principles combined with singular value decomposition (SVD).

Our method involves three key steps. Initially, we gather a diverse set of RAIRs, representing various potential source positions to a designated observation sensor, to form a training dataset. Following this, we apply SVD to the RAIR dataset, resulting in a semi-unitary signature matrix that characterizes the RAIRs. This signature matrix is then used to efficiently represent any new RAIR in a reduced dimension form, facilitating its application in real-time scenarios. Additionally, we demonstrate how this refined dimensionality reduction technique can be effectively applied to acoustic system identification, showcasing its practical utility in enhancing the accuracy and efficiency of processing RAIRs in complex acoustic environments.

## 1.2 Problem Formulation

Let's consider a scenario comprising $P$ distinct Room Acoustic Impulse Responses (RAIRs), each extending over a length $L$. These RAIRs encapsulate the responses emanating from $P$ discrete audio sources, such as loudspeakers, strategically positioned at different spots, all directing their output towards a singularly placed microphone. For the RAIR attributed to the $p$-th source, with $p$ spanning from 1 to $P$, it can be represented as a vertical vector thus:

$$h_p = \begin{bmatrix} h_{p,1} & h_{p,2} & \cdots & h_{p,L} \end{bmatrix}^T, \quad (1)$$

Here, the superscript $T$ symbolizes the vector's transposition.

Next, we define a matrix $H_{1:P-1}$ as follows:

$$H_{1:P-1} = \begin{bmatrix} h_1 & h_2 & \cdots & h_{P-1} \end{bmatrix}, \quad (2)$$

This results in an $L \times (P-1)$ matrix. When we apply Compact Singular Value Decomposition (SVD) to $H_{1:P-1}$, it decomposes into:

$$H_{1:P-1} = Q\Sigma V^T, \quad (3)$$

In this decomposition, $Q$ emerges as an $L \times R$ semi-unitary matrix, encompassing the left singular vectors of $H_{1:P-1}$. The variable $R$, representing the rank of $H_{1:P-1}$, adheres to the condition $R \leq \min\{L, P-1\}$. The matrix $\Sigma$, an $R \times R$ square diagonal matrix, holds the positive singular values along its diagonal. Moreover, $V$ represents a $(P-1) \times R$ semi-unitary matrix, containing the right singular vectors of $H_{1:P-1}$. Thus, the pseudo-inverse of $H_{1:P-1}$ is expressed as:

$$H^{\dagger}_{1:P-1} = V\Sigma^{-1}Q^T. \quad (4)$$

# Chapter 2

# Prior Art

## 2.1 Room Impulse Response Generator

Numerous techniques exist for the simulation of room acoustics. In our approach, we adopt the "Image Method", initially introduced by Allen and Berkley in 1979. Widely recognized and frequently utilized within the acoustic signal processing community, the Image Method offers a comprehensive solution for acoustic simulation. To facilitate its application, a MATLAB-compatible mex-function has been developed. This function, specifically designed for generating multichannel Room Impulse Responses (RIRs) through the Image Method, provides flexibility for the user. It allows for adjustments in various parameters such as the order of reflections, the dimensions of the room, and the directivity of the microphone.

### 2.1.1 Allen and Berkley's Image Method

The Image Model serves as an effective tool for simulating the reverberation effects in a room, tailored to specific locations of sound sources and microphones. Utilizing the principles of the Image Method, Allen and Berkley [1] have formulated an efficient approach for calculating a Finite Impulse Response (FIR). This FIR effectively models the acoustic channel existing between a source and a receiver within rectangular room environments.

### 2.1.2 Image Model

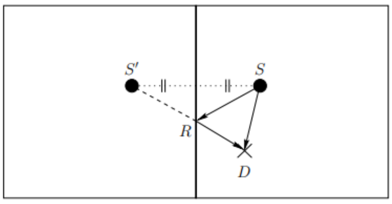
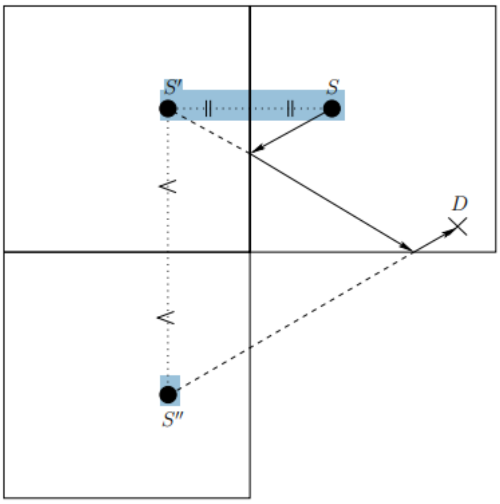



FIGURE 2.1: Path involving one reflection obtained using one image.

As depicted in Figure 2.1, a sound source $S$ is placed adjacent to a rigid reflective wall. At the destination point $D$, two signals are received: one directly from $S$ and another reflected off the wall. The direct path's length is straightforwardly determined by the known positions of $S$ and $D$. Additionally, an imaginary mirror image of the source, denoted as $S'$, is considered behind the wall, equidistant from it as $S$. Due to the symmetrical arrangement, the triangle formed by $S$, the reflective spot $R$, and $S'$ is isosceles, making the sum of the lengths $SR + RD$ equal to the distance $S'D$. This symmetry aids in calculating the reflected path's length by measuring the distance from $S'$ to $D$. The single reflection is indicated by using only one image to compute this distance.




[h]

FIGURE 2.2: Path involving two reflections obtained using two images.

Figure 2.2 illustrates a scenario with two reflections, where the relevant path length is derived from $S''D$. In Figure 2.3, a path involving three reflections is considered, and its length is determined from $S'''D$. These concepts are extendable to three dimensions, incorporating reflections from both ceiling and floor surfaces.

In summary, the path lengths and corresponding delays of reflections are calculated by measuring the distances from the source's images to the destination. The reflection's strength depends on the path length and the number of reflections, which is inferred from the level of image sources used for the calculation.

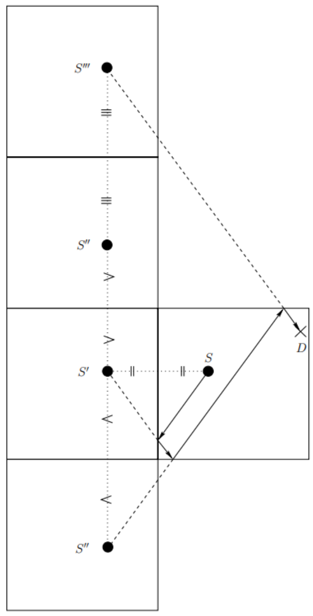



FIGURE 2.3: Path involving three reflections obtained using three images.

## 2.2 Image Method

In a rectangular room characterized by dimensions $L_x$, $L_y$, and $L_z$, consider a sound source at $\mathbf{r_s} = [x_s, y_s, z_s]$ and a microphone at $\mathbf{r} = [x, y, z]$, both relative to the room's origin at a corner. The position vectors of the source's images, relative to the microphone and created by reflections on the walls at $x = 0$, $y = 0$, and $z = 0$, are given by:

$$\mathbf{R_p} = [(1-2q)x_s - x, (1-2j)y_s - y, (1-2k)z_s - z], \quad (2.1)$$

where $p = (q, j, k)$ with $q, j, k \in \{0, 1\}$, resulting in eight unique combinations forming set $P$. The position of an image after $m$-th order reflections is found by adding $\mathbf{R_m}$ to $\mathbf{R_p}$, where:

$$\mathbf{R_m} = [2m_xL_x, 2m_yL_y, 2m_zL_z], \quad (2.2)$$

and $m = (m_x, m_y, m_z)$ with values ranging from $-N$ to $+N$. The order of reflection for an image at $\mathbf{R_p} + \mathbf{R_m} + \mathbf{r}$ is:

$$O_{p,m} = |2m_x - q| + |2m_y - j| + |2m_z - k|. \quad (2.3)$$

The distance $d$ between a source image and the microphone is:

$$d = \|\mathbf{R_p} + \mathbf{R_m}\|, \quad (2.4)$$

with the corresponding sound ray's time delay $\tau$ being:

$$\tau = \frac{d}{c} = \frac{\|\mathbf{R_p} + \mathbf{R_m}\|}{c}, \quad (2.5)$$

where $c$ is the speed of sound. The room's impulse response is given by:

$$h(\mathbf{r}, \mathbf{r_s}, t) = \sum_{p \in P} \sum_{m \in M} \frac{\beta_{x_1}^{|\mathrm{mx}-q|} \beta_{x_2}^{|\mathrm{mx}|} \beta_{y_1}^{|\mathrm{my}-j|} \beta_{y_2}^{|\mathrm{my}|} \beta_{z_1}^{|\mathrm{mz}-k|} \beta_{z_2}^{|\mathrm{mz}|} \delta(t-\tau)}{4\pi d}, \quad (2.6)$$

where $\beta_{v_1}, \beta_{v_2}$ are reflection coefficients of the six walls at $v = 0$ and $v = L_v$, respectively, and $M$ contains all $m$ values.

For discrete simulations, Peterson proposed a modification to this method, replacing each impulse in (2.6) with the response of a Hanning-windowed ideal low-pass filter, enhancing accuracy in time-of-arrival simulations even at lower sampling frequencies. The reverberation time $RT60$ is calculated using the Sabin-Franklin formula:

$$RT60 = \frac{24\ln(10)V}{c\sum_{i=1}^{6} S_i(1-\beta_i^2)}, \quad (2.7)$$

with $V$ as the room volume, and $S_i$, $\beta_i$ as the surface area and reflection coefficient of the $i$-th wall, respectively.

## 2.3 Implementation

The Image Method, as previously detailed, has been implemented through a MATLAB mex-function, developed using C++. This implementation yields a Dynamic-Link Library (DLL) compatible with MATLAB, functioning as a regular MATLAB function. Notably, the C++ version exhibits enhanced performance compared to its MATLAB counterpart.

The function for generating room impulse responses, *rir_generator*, is defined as:

```
function [h, beta_hat] = rir_generator(c, fs, r, s, L, beta, nsample,
                                       mtype, order, dim, orientation,
                                       hp_filter);
```

Input parameters include:

- $c$: Sound velocity in meters per second.
- $fs$: Sampling frequency in Hertz.
- $r$: An $M \times 3$ matrix detailing the receiver(s) coordinates in meters.
- $s$: A $1 \times 3$ vector for the source coordinates in meters.
- $L$: A $1 \times 3$ vector defining the room dimensions in meters.

- *beta*: A $1 \times 6$ vector for reflection coefficients $[\beta_{x1}\beta_{x2}\beta_{y1}\beta_{y2}\beta_{z1}\beta_{z2}]$ or the reverberation time (RT60) in seconds.

Optional input parameters, along with their default values, are:

- *nsample*: Number of samples to be calculated, defaulting to RT60fs.
- *mtype*: Microphone type, options include 'omnidirectional', 'subcardioid', etc., with 'omnidirectional' as the default.
- *order*: Maximum reflection order, default is -1.
- *dim*: Room dimension (2 or 3), with 3 as the default.
- *orientation*: Microphone direction, specified in azimuth and elevation angles in radians, defaulting to [0 0].
- *hp_filter*: High-pass filter usage, with 'true' as the default setting.

Output parameters are:

- *h*: An $M \times nsample$ matrix containing the calculated room impulse responses.
- *beta_hat*: Reflection coefficient returned if reverberation time is input.

The '*rir_generator*' function supports multi-channel calculations, enabling simultaneous RIR computation for multiple receiver positions.

### 2.3.1 Example

**Algorithm 1** MATLAB Code Implementation

| | |
|---|---|
| $c = 340$ | ▷ Sound velocity (m/s) |
| $fs = 16000$ | ▷ Sample frequency (samples/s) |
| $r = [2, 1.5, 2]$ | ▷ Receiver position [x y z] (m) |
| $s = [2, 3.5, 2]$ | ▷ Source position [x y z] (m) |
| $L = [5, 4, 6]$ | ▷ Room dimensions [x y z] (m) |
| $beta = 0.4$ | ▷ Reverberation time (s) |
| $n = 4096$ | ▷ Number of samples |
| $h = \text{rir\_generator}(c, fs, r, s, L, beta, n)$ | ▷ Calculate impulse response |

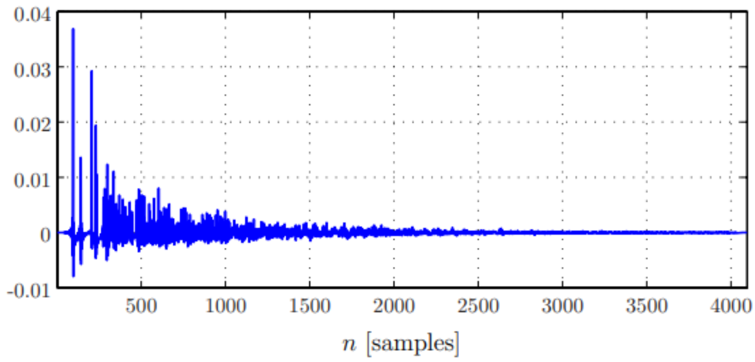



FIGURE 2.4: Output of Example

# Chapter 3

# Dimensionality Reduction

## 3.1 Mathematical Illustration

Consider a collection $\{h_1, h_2, \ldots, h_P\}$, with $P \leq L$, serving as the basis for deriving the semi-orthogonal Room Acoustic Impulse Response (RAIR) matrix $Q$ of dimensions $L \times R$. We extract the leading $R'$ columns from $Q$, corresponding to the $R'$ highest singular values obtained from $H_{1:P-1}$. Here, $R' \leq R \leq L$, to construct the RAIR signature matrix $U$ of size $L \times R'$. This matrix $U$ is then employed for the reduction of RAIR dimensionality.

For a given test RAIR $h'$ not in $\{h_1, h_2, \ldots, h_P\}$ and with length $L$, its condensed representation $h''$ of length $R'$ is calculated as:

$$h'' = U^\top h'. \tag{3.1}$$

The error vector $e'$ representing the difference between $h'$ and its reconstruction $Uh''$ is given by:

$$e' = h' - Uh''. \tag{3.2}$$

Minimizing the squared Euclidean norm $\|e'\|_2^2$ with respect to $h''$ naturally results in $h'' = U^\top h'$. Substituting this into the error expression yields:

$$\|e'\|_2^2 = h'^\top h' - h'^\top U U^\top h'. \tag{3.3}$$

Consequently, the normalized misalignment, which is expected to diminish as $P$ grows, is formulated as:

$$\frac{\|h' - Uh''\|_2^2}{\|h'\|_2^2} = 1 - \frac{h'^\top UU^\top h'}{h'^\top h'}. \tag{3.4}$$

### 3.1.1 Pseudo-Code

**Algorithm 2** RAIR Dimensionality Reduction and Testing

```
H ← zeros(n, P)
for j = 1, ..., P do
    source_pos = source_positions(j, :)
    H(:, j) = rir_generator(c, fs, microphone_pos, source_pos, L, beta, n)
end for

[U, S, V] = svd(H, 'econ')
R = findRankForEnergyThreshold(singular_values, 0.95)
U_reduced = U(:, 1 : R)                                          ▷ Dimensionality reduction

test_source_pos = [x, y, z]
h_prime = rir_generator(c, fs, microphone_pos, test_source_pos, L, beta, n)
h_prime = h_prime(:)
h_prime_low_dim = U_reduced' * h_prime                           ▷ Low-dimensional representation
h_prime_reconstructed = U_reduced * h_prime_low_dim              ▷ Reconstruction

misalignment_numerator = norm(h_prime − h_prime_reconstructed, 2)^2
misalignment_denominator = norm(h_prime, 2)^2
misalignment = (misalignment_numerator/misalignment_denominator)
misalignments = 10 * log_10(misalignments                        ▷ Misalignment in dB
```

### 3.1.2 Rank For Energy Threshold

The MATLAB function `findRankForEnergyThreshold` is designed to determine the minimum rank $R$ that is required to capture a specified fraction of the energy in a system. This energy is quantified by the square of the system's singular values. If the user does not specify an energy threshold, the function defaults to capturing 95% of the total energy. The function calculates the total energy by summing the squares of all the singular values. It then incrementally adds the square of each singular value to a running total of captured energy. When the captured energy constitutes the desired fraction of the total energy, as dictated by the energy threshold, the function returns the rank $R$ at which this condition is satisfied.

**Algorithm 3** Find Minimum Rank Based on Energy Threshold

```
function findRankForEnergyThreshold(singular_values, energy_threshold)
    if nargin < 2 then
        energy_threshold = 0.95                    ▷ Default energy threshold
    end if
    total_energy = sum(singular_values.^2)
    energy_captured = 0
    R = 0
    for i = 1 to length(singular_values) do
        energy_captured = energy_captured + singular_values(i)^2
        if energy_captured/total_energy ≥ energy_threshold then
            R = i
            break
        end if
    end for
    return R
end function
```

# Chapter 4

# Applications in System Identification

## 4.1 Mathematical Illustration

Consider a linear single-input, single-output (SISO) acoustic system within an environment afflicted by noise. The system's output at any given time $k$ is represented by the equation:

$$d(k) = \mathbf{h}^\top \mathbf{x}(k) + n(k) = y(k) + n(k), \tag{4.1}$$

where $\mathbf{x}(k)$ is a column vector of length $L$ signifying the input signal and is defined as:

$$\mathbf{x}(k) = \begin{bmatrix} x(k) \\ x(k-1) \\ \vdots \\ x(k-L+1) \end{bmatrix},$$

with $x(k)$ being the zero-mean input signal. The vector $\mathbf{h}$ symbolizes the system's unknown Room Acoustic Impulse Response (RAIR), and $n(k)$ is the zero-mean noise which is uncorrelated with $x(k)$. Under these conditions, the objective is to estimate the vector $\mathbf{h}$ based on the known inputs $\mathbf{x}(k)$ and output $d(k)$.

To model the RAIR $\mathbf{h}$, a real-valued linear filter $\mathbf{g}$ of dimensions $L \times 1$ is used to estimate $d(k)$ as:

$$\hat{y}(k) = \mathbf{g}^\top \mathbf{x}(k), \tag{4.2}$$

where $\hat{y}(k)$ signifies the estimated output. If we express $\mathbf{g}$ as $\mathbf{U}\mathbf{g}'$, where $\mathbf{g}'$ is a filter of reduced length $R'$, substitution into equation (4.2) yields:

$$\hat{y}(k) = \mathbf{g}'^{\top}\mathbf{x}'(k), \tag{4.3}$$

with $\mathbf{x}'(k) = \mathbf{U}^{\top}\mathbf{x}(k)$ being the transformed input signal vector of reduced length $R'$. Thus, the focus shifts to estimating the shorter filter $\mathbf{g}'$ rather than the full-length filter $\mathbf{g}$.

The discrepancy between the actual output $d(k)$ and the estimated $\hat{y}(k)$ is quantified by the error term:

$$\epsilon(k) = d(k) - \hat{y}(k). \tag{4.4}$$

The Mean Squared Error (MSE) criterion for this scenario, represented by $J(\mathbf{g}')$, is subsequently formulated as:

$$J(\mathbf{g}') = E[\epsilon^2(k)] = \sigma_d^2 - 2\mathbf{g}'^{\top}\boldsymbol{\gamma}_{\mathbf{x}'d} + \mathbf{g}'^{\top}\boldsymbol{\Phi}_{\mathbf{x}'}\mathbf{g}', \tag{4.5}$$

where $\sigma_d^2$ is the variance of $d(k)$, and $\boldsymbol{\gamma}_{\mathbf{x}'d}$ and $\boldsymbol{\Phi}_{\mathbf{x}'}$ are the cross-correlation vector and the covariance matrix, respectively. The minimization of $J(\mathbf{g}')$ yields the Wiener solution:

$$\mathbf{g}'_W = \boldsymbol{\Phi}_{\mathbf{x}'}^{-1}\boldsymbol{\gamma}_{\mathbf{x}'d}. \tag{4.6}$$

The optimal solution for the Wiener filter in terms of identifying the RAIR $\mathbf{h}$ is therefore given by:

$$\mathbf{g}_W = \mathbf{U}\mathbf{g}'_W = \mathbf{U}\boldsymbol{\Phi}_{\mathbf{x}'}^{-1}\boldsymbol{\gamma}_{\mathbf{x}'d} = \mathbf{U}(\mathbf{U}^{\top}\boldsymbol{\Phi}_{\mathbf{x}}\mathbf{U})^{-1}\mathbf{U}^{\top}\boldsymbol{\gamma}_{\mathbf{xd}}, \tag{4.7}$$

where $\boldsymbol{\Phi}_{\mathbf{x}}$ and $\boldsymbol{\gamma}_{\mathbf{xd}}$ are the covariance matrix and the cross-correlation vector for the untransformed signals.

### 4.1.1 Pseudo-Code

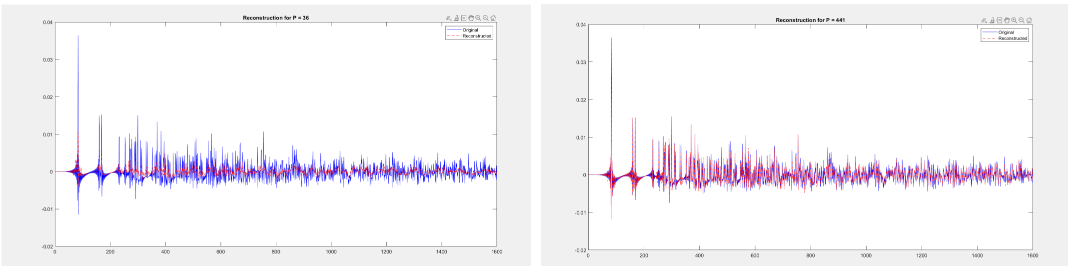
**Algorithm 4** Excitation Signal Generation and System Identification

```
alpha = 0.9                                                          ▷ AR(1) coefficient
ar_coeffs = [1, −alpha]
excitation_signal = filter(1, ar_coeffs, randn(n_obs, 1))            ▷ Colored noise generation

h_prime = rir_generator(c, fs, microphone_pos, test_source_pos, L, beta, n)

d = conv(excitation_signal, h_prime, ′same′)                         ▷ Convolution with test RAIR
signal_power = bandpower(d)                                          ▷ Signal power calculation
noise_power = signal_power/(10^(SNR_target/10))                      ▷ Noise power for desired SNR
noise = sqrt(noise_power/2) ∗ randn(size(d))                         ▷ White Gaussian noise
d_with_noise = d + noise                                             ▷ Signal with added noise

X_toep = toeplitz(excitation_signal(n_obs : −1 : 1), excitation_signal(n_obs : −1 : n_obs − n_rir + 1))′   ▷ Input signal vector
Rdx_conventional = X_toep ∗ d_with_noise                             ▷ Cross-correlation
Rxx_conventional = X_toep ∗ X_toep′                                  ▷ Auto-correlation
gw_conventional = Rxx_conventional\Rdx_conventional                  ▷ Wiener filter estimate
h_estimated_conventional = X_toep′ ∗ gw_conventional                 ▷ Estimated RAIR
e_conventional = d_with_noise − h_estimated_conventional             ▷ Error calculation

misalignment_numerator_conventional = norm(e_conventional, 2)^2
misalignment_denominator_conventional = norm(d_with_noise, 2)^2
misalignments_conventional = 10∗log_10(misalignment_numerator_conventional/misalignment
▷ Misalignment in dB
```

# Chapter 5

# Simulation and Results

In the subsequent analysis, we examine the efficacy of RAIRs' compact representations obtained via Dimensionality Reduction techniques. The focus is on assessing the system identification accuracy with varying numbers of $P$, representing uniformly spaced potential sound sources within a rectangular field. The term $P$ quantifies these sources, and the uniform spacing along the horizontal and vertical axes is indicated by $\delta_x$ and $\delta_y$, respectively. The size of each Room Acoustic Impulse Response (RAIR) is set to $L = 1600$. Additionally, the effectiveness of the suggested RAIR dimensionality reduction technique in system identification will be evaluated.




(a) P = 36 (b) P = 441

FIGURE 5.1: Original Signal and it's Reconstruction

### 5.0.1 For Varying Values of P

In this simulation, we consider the value of $P$ from 9 to 676. For each value of P we have a corresponding value of $\delta_x$ and $\delta_y$. We first generate the $P$ RAIRs and then

Table 5.1: Distribution of Potential Sound Sources

| $P$ (Number of Sources) | $\delta_x$ (Spacing on x-axis) | $\delta_y$ (Spacing on y-axis) |
|---|---|---|
| 9 | 0.2 m | 0.2 m |
| 36 | 0.1 m | 0.1 m |
| 121 | 0.05 m | 0.05 m |
| 289 | 0.03 m | 0.03 m |
| 441 | 0.025 m | 0.025 m |
| 676 | 0.02 m | 0.02 m |

generate the test RAIR and we calculate the normalized misalignment between the test RAIR and it's reconstruction according to equation (3.4).

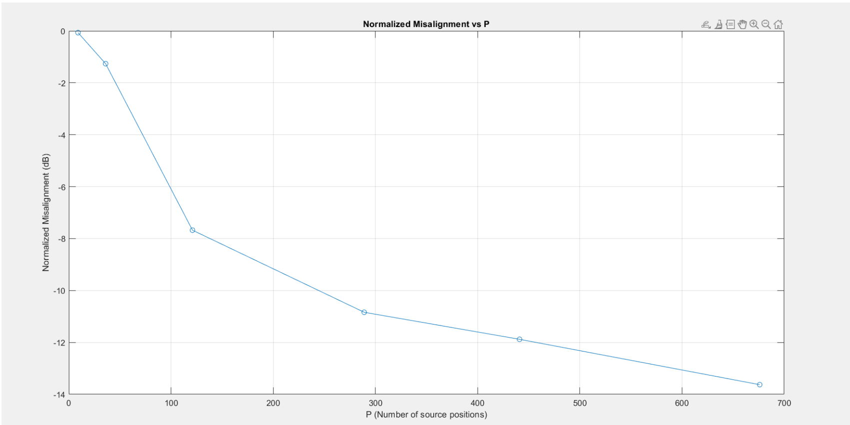



Figure 5.2: Normalized Misalignment vs P

The graph clearly demonstrates a trend where the normalized misalignment diminishes with increasing values of $P$. Particularly at larger $P$ values, the normalized misalignment attains substantially negative values. This trend underscores the efficiency of the dimensionality reduction technique applied to RAIRs. Essentially, a more negative value of normalized misalignment correlates with a reduced discrepancy between the original signal and its reconstructed version using the lower-dimensional representation.

### 5.0.2 System Identification

Employing an autoregressive (AR) model, we engineer a colored excitation signal. This process involves channeling white Gaussian noise through a primary order AR mechanism, encapsulated by the equation $\frac{1}{1-\alpha z^{-1}}$, with the parameter $\alpha$ chosen as 0.9. To fabricate the observable signal $d(k)$, we initially convolve the colored excitation signal with the RAIR, subsequently introducing white Gaussian noise to calibrate the SNR.

Operating at a sampling rate of 16000 Hz, we leverage the RAIR set $\{h_1, h_2, \ldots, h_P\}$ as the foundational training dataset. Following this, we construct the semi-orthogonal signature matrix $U$ for RAIR, spanning $L \times R'$ dimensions, in alignment with the protocol delineated in equation (3). The ensuing phase involves pinpointing the test RAIR via a Wiener filter, crafted as per equation (4.7). The precision of system identification is quantified through the normalized misalignment metric between $h'$ and $g_W$, in harmony with the framework set forth in equation (3.4).

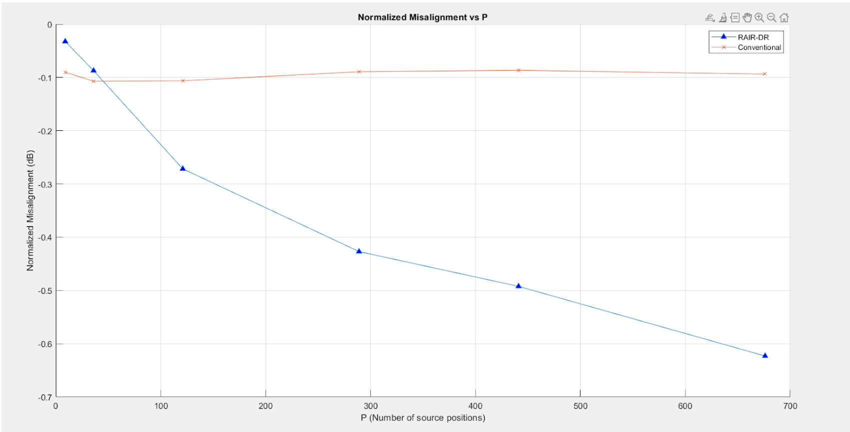



FIGURE 5.3: Normalized misalignment for identification of acoustic system with RAIR-DR method and its conventional method without dimensionality reduction across varying values of P

We assess the performance of system identification across various counts of $P$ by setting the Signal-to-Noise Ratio (SNR) at 15 dB. The observation period is established at 4000 samples, and $R'$ is designated as $P-1$. Figure 5.3 illustrates the relationship between the normalized misalignment and varying values of $P$. From this figure, it's evident that the RAIR-Dimensionality Reduction (RAIR-DR) technique surpasses the conventional method, particularly when the number of potential sources, represented by $P$, is not exceedingly low. These observations solidly indicate that the RAIR-DR approach, with an optimally chosen value of $P$, more effectively identifies acoustic systems.

# Chapter 6

# Future Plan

The outcomes of our study reveal that employing RAIR dimensionality reduction significantly enhances system identification, especially when dealing with a high count of sound sources.

Subsequent investigations will focus on the RAIR-DR method's efficacy under varying Signal-to-Noise Ratio (SNR) conditions, in comparison to traditional methods. Additionally, we plan to scrutinize the performance of the RAIR-DR method under different observation lengths (measured in samples).

An analysis will also be conducted to explore the influence of varying the Rank $R$ on the performance of the RAIR-DR method against the conventional approach.

The comprehensive results from these experimental scenarios are expected to demonstrate the superiority of the RAIR-DR methodology in advancing acoustic system identification, particularly over conventional techniques that do not incorporate dimensionality reduction.

# Bibliography


Allen, J. and Berkley, D. (1979). "image method for efficiently simulating small room acoustics". *J. Acoustical Soc. Amer.*, 65.

G. Huang, J. Benesty, I. C. and Chen, J. (2023). Kronecker product multichannel linear filtering for adaptive weighted prediction error-based speech dereverberation. *IEEE/ACM Trans. Audio, Speech, Lang. Process*, 30.

Gongping Huang, Jacob Benesty, J. C. (2023). Dimensionality reduction of room acoustic impulse responses and applications to system identification. *IEEE SIGNAL PROCESSING LETTERS*, 30.

Habets, D. E. A. (2010). Room impulse response generator.

J. Benesty, J. Chen, Y. H. and Dmochowski, J. (2007). On microphone-array beamforming from a mimo acoustic signal processing perspective. *IEEE Trans. Audio, Speech, Lang. Process.*, 15.

W. Yang, G. Huang, J. C. J. B. I. C. and Kellermann, W. (2021). Robust dereverberation with kronecker product-based multichannel linear prediction. *IEEE Signal Process Letters*, 28.