Mean Absorption Coefficient Estimation from Impulse Responses

Deep Learning vs. Sabine

Corto Bastien[a] **Antoine Deleforge**[a] Cédric Foy[b]

[a]Inria Nancy – Grand Est
[b]Cerema, Ifsttar, UMRAE, Strasbourg

e-Forum Acusticum 2020
Dec 7-11, 2020

Antoine.Deleforge@inria.fr Cedric.Foy@cerema.fr

OUTLINE

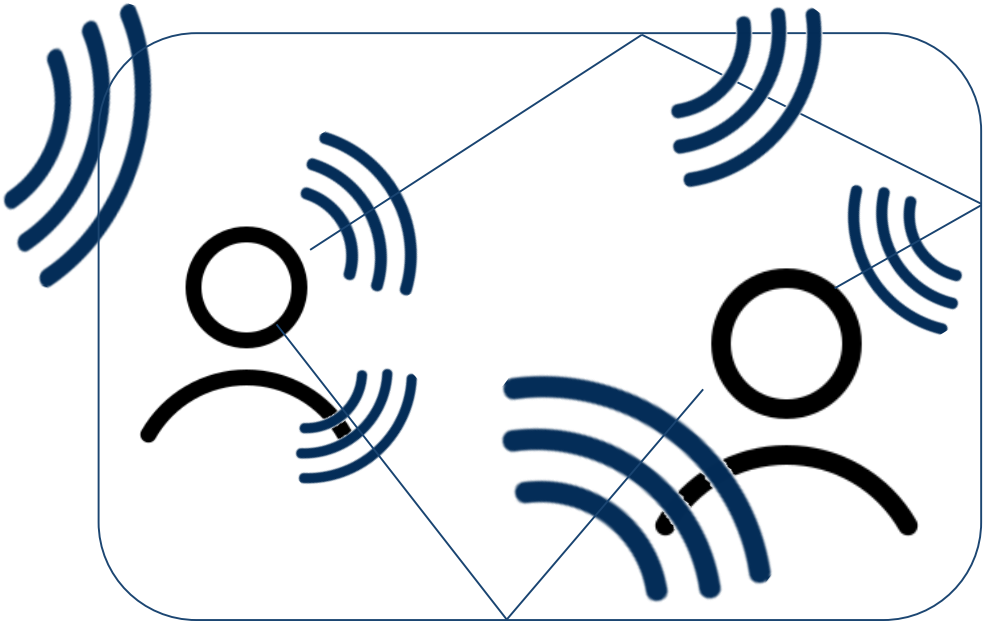
1. Introduction
2. Simulated Datasets
3. Neural Networks
4. Experiments & Results
5. Conclusion & Future Work

OUTLINE

1. **Introduction**
2. Simulated Datasets
3. Neural Networks
4. Experiments & Results
5. Conclusion & Future Work

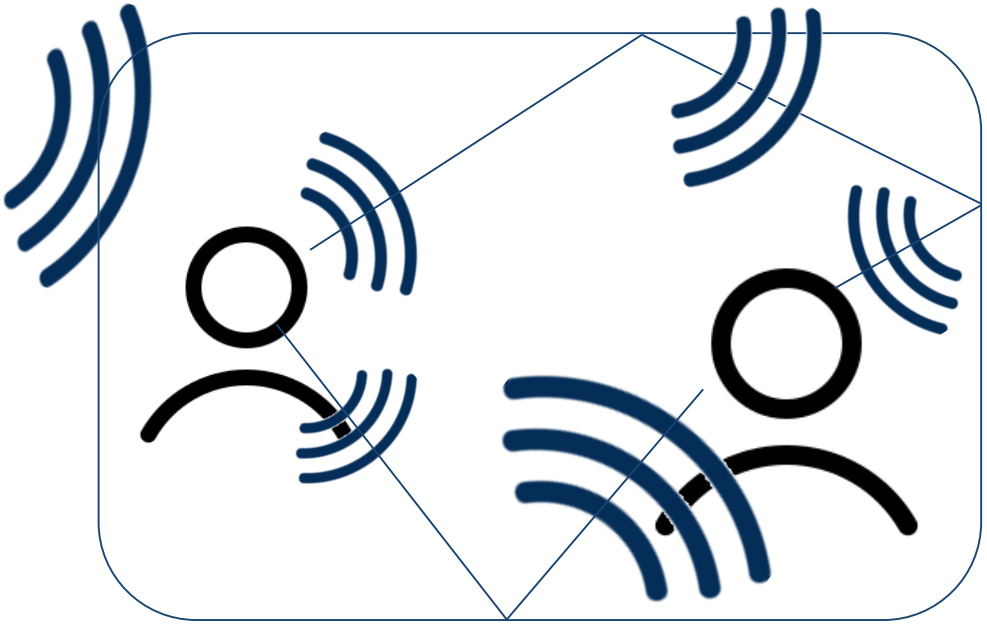
Noise disturbance

- Public health concern
- Affects speech intelligibility
- Reduces concentration



Noise disturbance

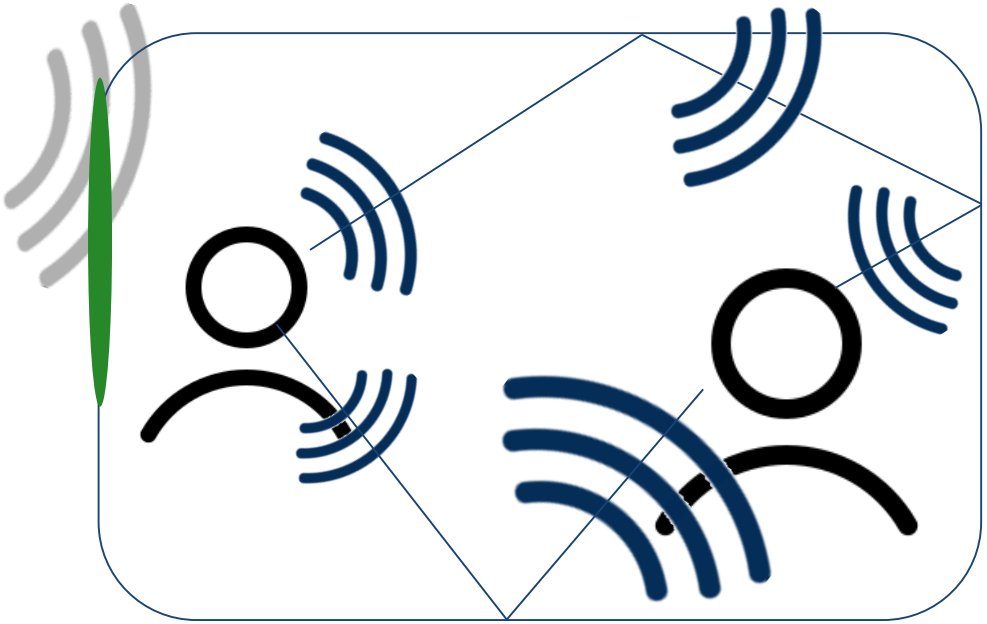
- Public health concern
- Affects speech intelligibility
- Reduces concentration



➡ Renovation of existing buildings

Noise disturbance

- Public health concern
- Affects speech intelligibility
- Reduces concentration

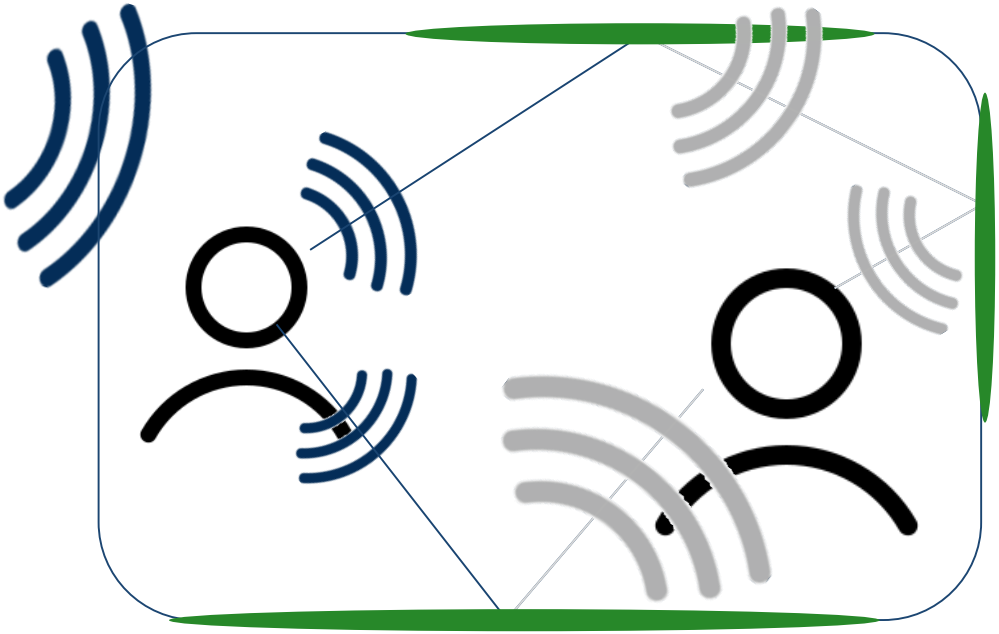


➡ Renovation of existing buildings

- Wall transmission

Noise disturbance

- Public health concern
- Affects speech intelligibility
- Reduces concentration

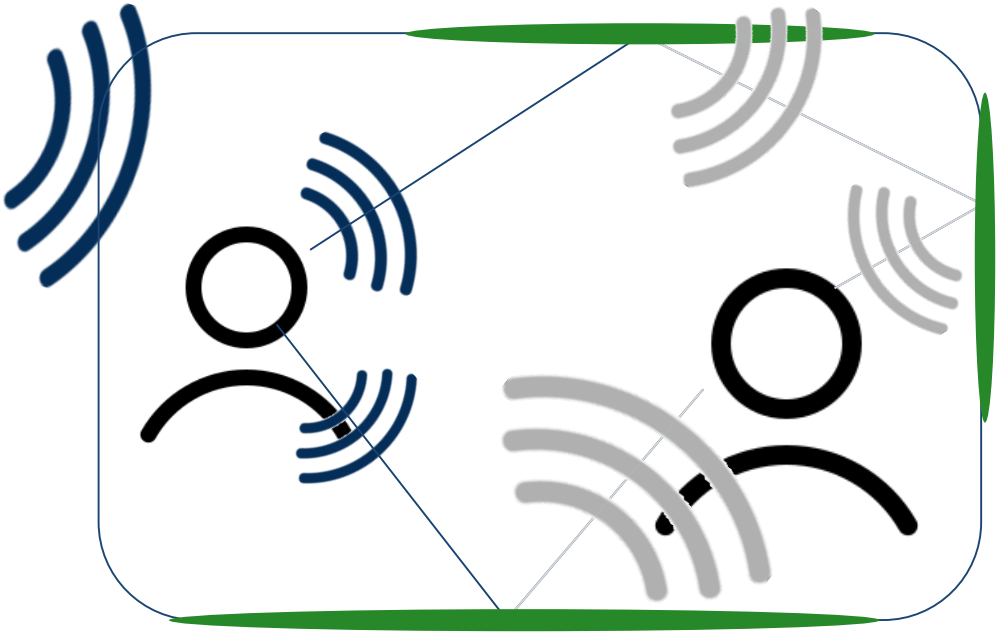


➡ Renovation of existing buildings

- Wall transmission
- <u>Wall absorption</u>

Noise disturbance

- Public health concern
- Affects speech intelligibility
- Reduces concentration



➡ Renovation of existing buildings
- Wall transmission
- Wall absorption

- First, a **diagnosis** is required:

Absorption coefficients estimation

Conventional approach

- Estimating the mean absorption coefficients in octave bands:

$$\bar{\alpha}(b) \overset{\text{def}}{=} \frac{1}{S_{\text{tot}}} \sum_{\text{surface } i} \alpha_i(b) S_i \qquad \left(b \in \{125, 250, 500, \ldots, 4000\} \text{ Hz} \right)$$

Conventional approach

- Estimating the mean absorption coefficients in octave bands:

$$\bar{\alpha}(b) \overset{\text{def}}{=} \frac{1}{S_{\text{tot}}} \sum_{\text{surface } i} \alpha_i(b) S_i \qquad \left(b \in \{125, 250, 500, \ldots, 4000\} \text{ Hz} \right)$$

Absorption coefficient in [0,1]

Conventional approach

- Estimating the mean absorption coefficients in octave bands:

$$\bar{\alpha}(b) \overset{\text{def}}{=} \frac{1}{\underbrace{S_{\text{tot}}}_{\sum_i S_i}} \sum_{\text{surface } i} \alpha_i(b) \underbrace{S_i}_{\text{Area in m}^2} \qquad \left(b \in \{125, 250, 500, \ldots, 4000\} \text{ Hz} \right)$$

Conventional approach

- Estimating the mean absorption coefficients in octave bands:

$$\bar{\alpha}(b) \overset{\text{def}}{=} \frac{1}{S_{\text{tot}}} \sum_{\text{surface } i} \alpha_i(b) S_i \qquad \left(b \in \{125, 250, 500, \ldots, 4000\} \text{ Hz} \right)$$

- Equations from reverberation theory:

$$\bar{\alpha}_{\text{Sabine}}(b) = 0.163 \frac{V}{S_{\text{tot}} R T_{60}(b)}, \qquad \bar{\alpha}_{\text{Eyring}}(b) = -\ln(1 - \bar{\alpha}_{\text{Sabine}}(b))$$

Conventional approach

- Estimating the mean absorption coefficients in octave bands:

$$\bar{\alpha}(b) \overset{\text{def}}{=} \frac{1}{S_{\text{tot}}} \sum_{\text{surface } i} \alpha_i(b) S_i \qquad \left(b \in \{125, 250, 500, \ldots, 4000\} \text{ Hz} \right)$$

- Equations from reverberation theory:

Volume in m³

$$\bar{\alpha}_{\text{Sabine}}(b) = 0.163 \frac{V}{S_{\text{tot}} RT_{60}(b)}, \qquad \bar{\alpha}_{\text{Eyring}}(b) = -\ln(1 - \bar{\alpha}_{\text{Sabine}}(b))$$

Conventional approach

- Estimating the mean absorption coefficients in octave bands:

$$\bar{\alpha}(b) \overset{\text{def}}{=} \frac{1}{S_{\text{tot}}} \sum_{\text{surface } i} \alpha_i(b) S_i \qquad \left(b \in \{125, 250, 500, \ldots, 4000\} \text{ Hz} \right)$$

- Equations from reverberation theory:

$$\bar{\alpha}_{\text{Sabine}}(b) = 0.163 \frac{V}{S_{\text{tot}} RT_{60}(b)}, \qquad \bar{\alpha}_{\text{Eyring}}(b) = -\ln(1 - \bar{\alpha}_{\text{Sabine}}(b))$$

Reverberation time in sec

Conventional approach

- Estimating the mean absorption coefficients in octave bands:

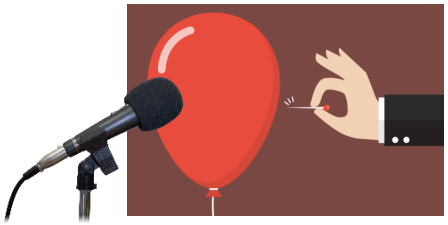
$$\bar{\alpha}(b) \overset{\text{def}}{=} \frac{1}{S_{\text{tot}}} \sum_{\text{surface } i} \alpha_i(b) S_i \qquad \left(b \in \{125, 250, 500, \ldots, 4000\} \text{ Hz} \right)$$

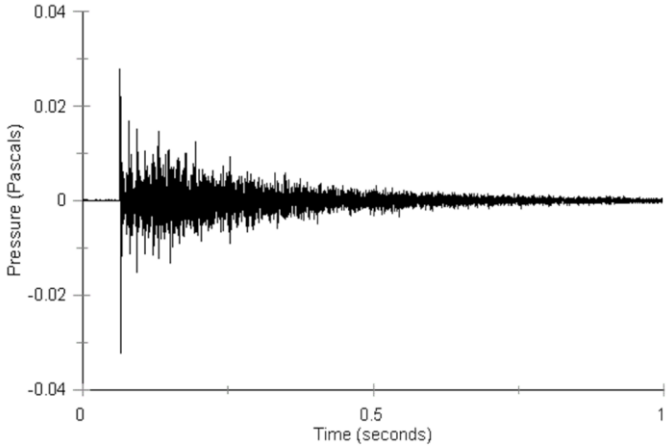
- Equations from reverberation theory:

$$\bar{\alpha}_{\text{Sabine}}(b) = 0.163 \frac{V}{S_{\text{tot}} RT_{60}(b)}, \qquad \bar{\alpha}_{\text{Eyring}}(b) = -\ln(1 - \bar{\alpha}_{\text{Sabine}}(b))$$

Reverberation time in sec

➡ Obtained by Schroeder integration on a ***Room Impulse Response*** **(RIR)**





Conventional approach

- Estimating the mean absorption coefficients in octave bands:

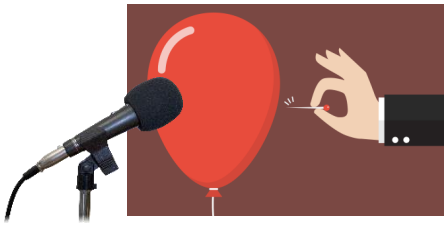
$$\bar{\alpha}(b) \overset{\text{def}}{=} \frac{1}{S_{\text{tot}}} \sum_{\text{surface } i} \alpha_i(b) S_i \qquad \left(b \in \{125, 250, 500, \ldots, 4000\} \text{ Hz} \right)$$

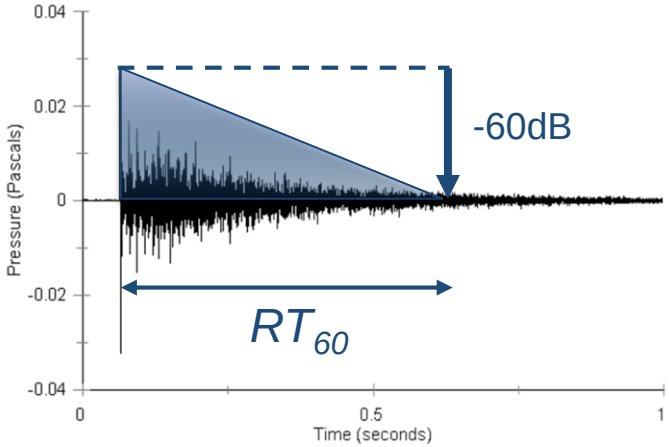
- Equations from reverberation theory:

$$\bar{\alpha}_{\text{Sabine}}(b) = 0.163 \frac{V}{S_{\text{tot}} RT_{60}(b)}, \qquad \bar{\alpha}_{\text{Eyring}}(b) = -\ln(1 - \bar{\alpha}_{\text{Sabine}}(b))$$

Reverberation time in sec

➡ Obtained by Schroeder integration on a ***Room Impulse Response*** **(RIR)**





Conventional approach

- Estimating the mean absorption coefficients in octave bands:

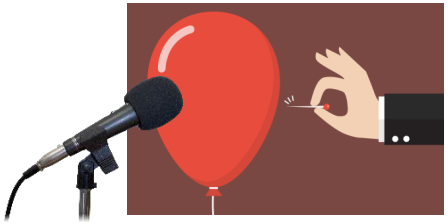
$$\bar{\alpha}(b) \overset{\text{def}}{=} \frac{1}{S_{\text{tot}}} \sum_{\text{surface } i} \alpha_i(b) S_i \qquad \left(b \in \{125, 250, 500, \ldots, 4000\} \text{ Hz} \right)$$

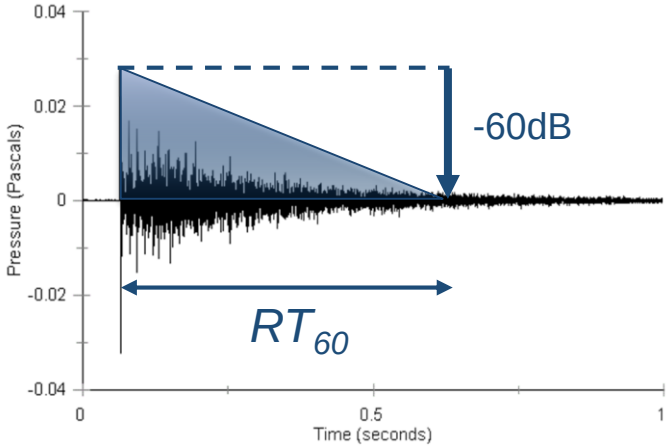
- Equations from reverberation theory:

$$\bar{\alpha}_{\text{Sabine}}(b) = 0.163 \frac{V}{S_{\text{tot}} RT_{60}(b)}, \qquad \bar{\alpha}_{\text{Eyring}}(b) = -\ln(1 - \bar{\alpha}_{\text{Sabine}}(b))$$

Reverberation time in sec

➡ Obtained by Schroeder integration on a ***Room Impulse Response*** **(RIR)**

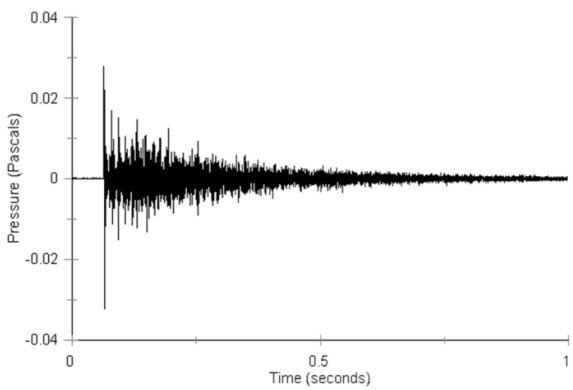


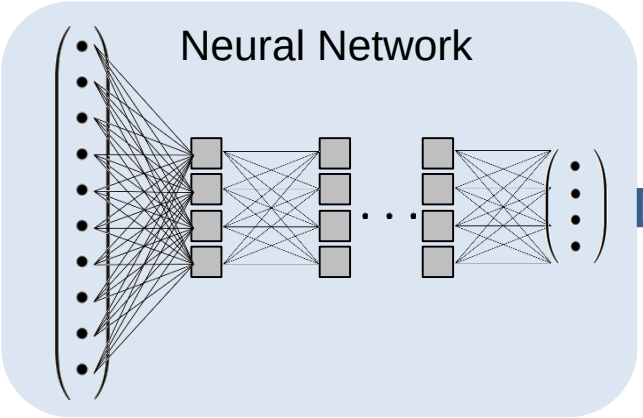


Limits:

- ✕ Assumes a diffuse sound field
- ✕ Only valid in ***homo-genous conditions***
- ✕ Requires V and S_{tot}

Proposed approach

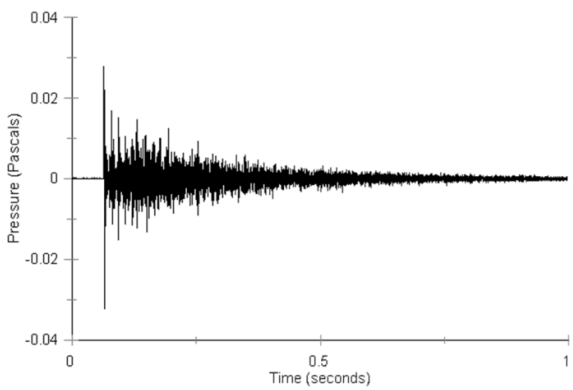


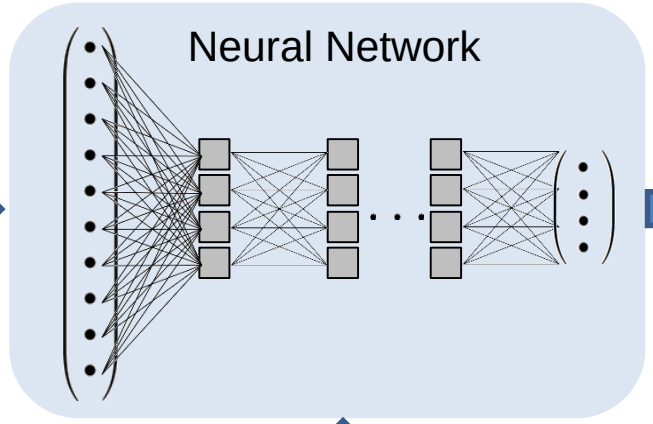


Mean absorption coefficients

$$\bar{\alpha}_{\mathrm{NN}}(125\mathrm{Hz})$$
$$\bar{\alpha}_{\mathrm{NN}}(250\mathrm{Hz})$$
$$\vdots$$
$$\bar{\alpha}_{\mathrm{NN}}(4000\mathrm{Hz})$$

Proposed approach





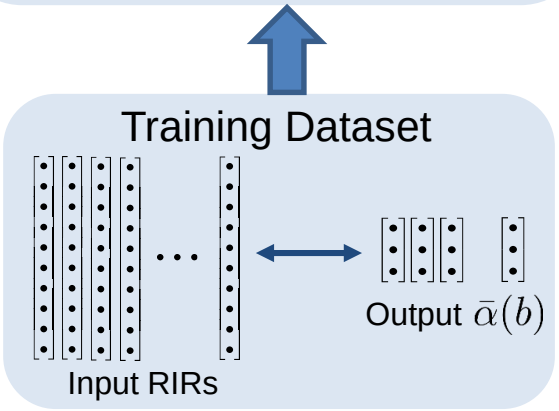
Mean absorption coefficients

$\bar{\alpha}_{\mathrm{NN}}(125\mathrm{Hz})$

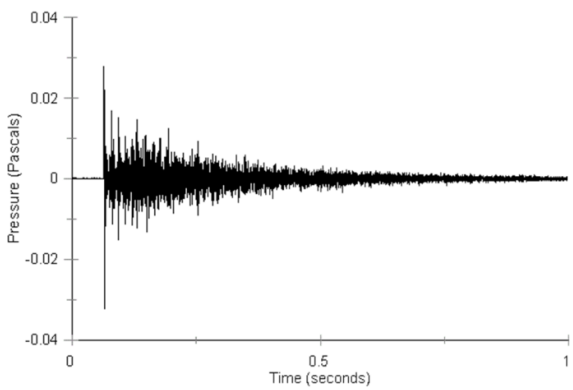
$\bar{\alpha}_{\mathrm{NN}}(250\mathrm{Hz})$

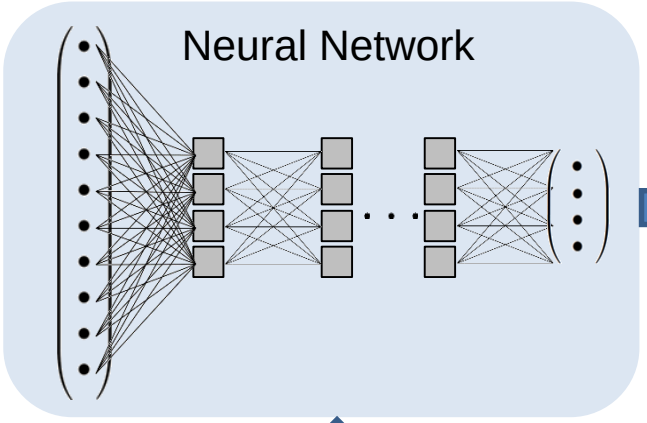
$\vdots$

$\bar{\alpha}_{\mathrm{NN}}(4000\mathrm{Hz})$



Proposed approach





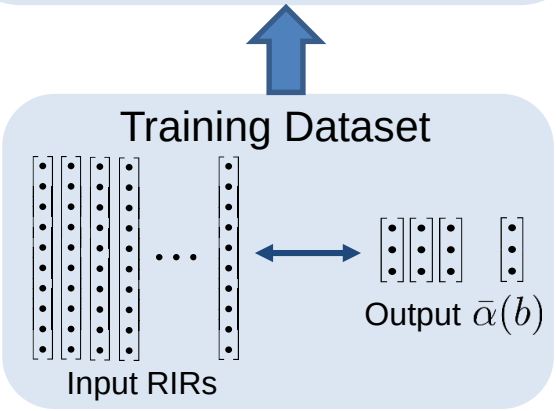
Mean absorption coefficients

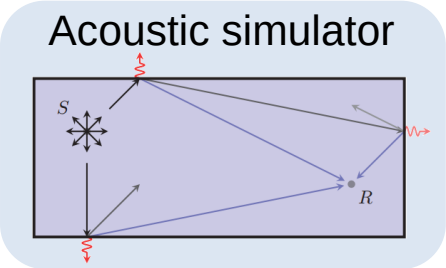
$\bar{\alpha}_{\mathrm{NN}}(125\mathrm{Hz})$

$\bar{\alpha}_{\mathrm{NN}}(250\mathrm{Hz})$

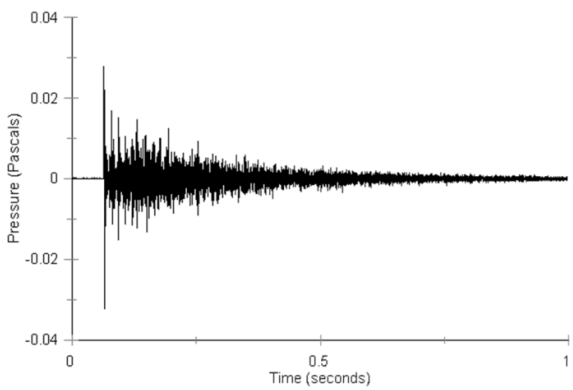
$\vdots$

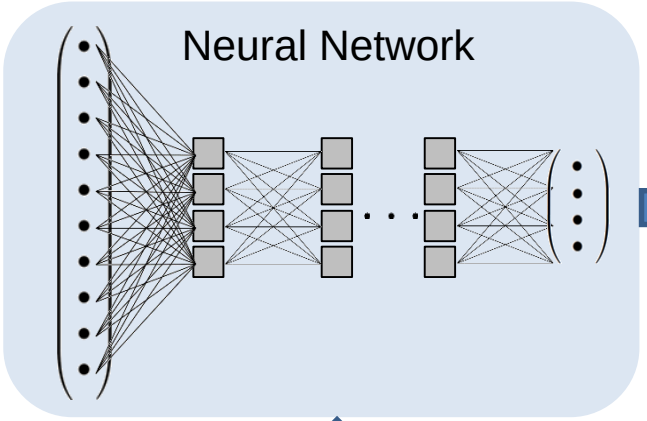
$\bar{\alpha}_{\mathrm{NN}}(4000\mathrm{Hz})$





Proposed approach





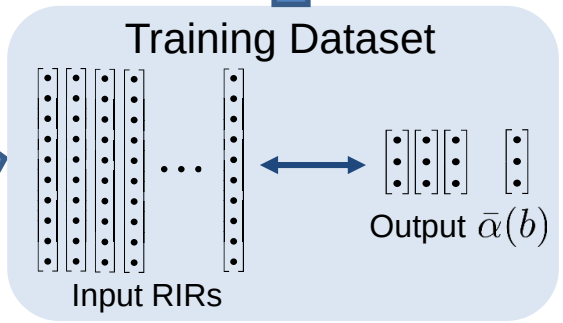
Mean absorption coefficients

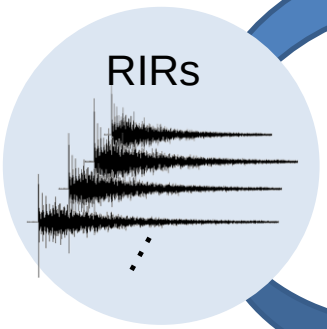
$\bar{\alpha}_{\mathrm{NN}}(125\mathrm{Hz})$

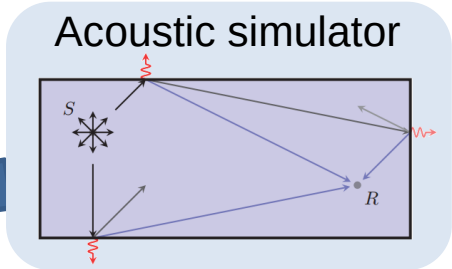
$\bar{\alpha}_{\mathrm{NN}}(250\mathrm{Hz})$

$\vdots$

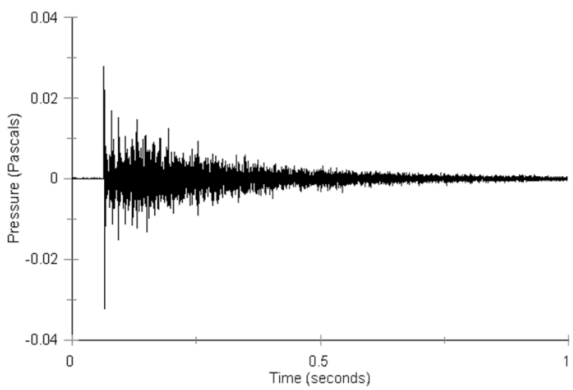
$\bar{\alpha}_{\mathrm{NN}}(4000\mathrm{Hz})$

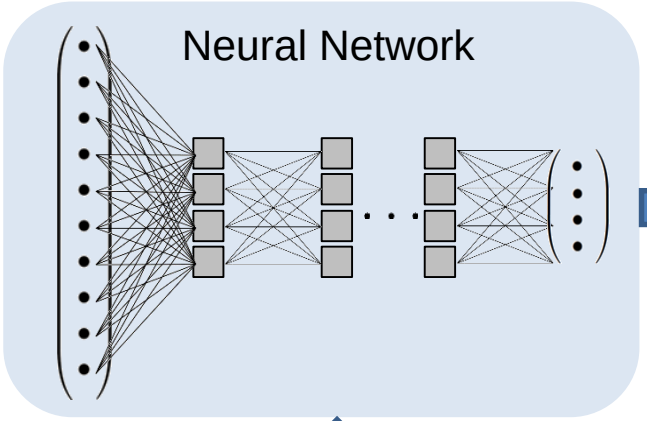






Proposed approach





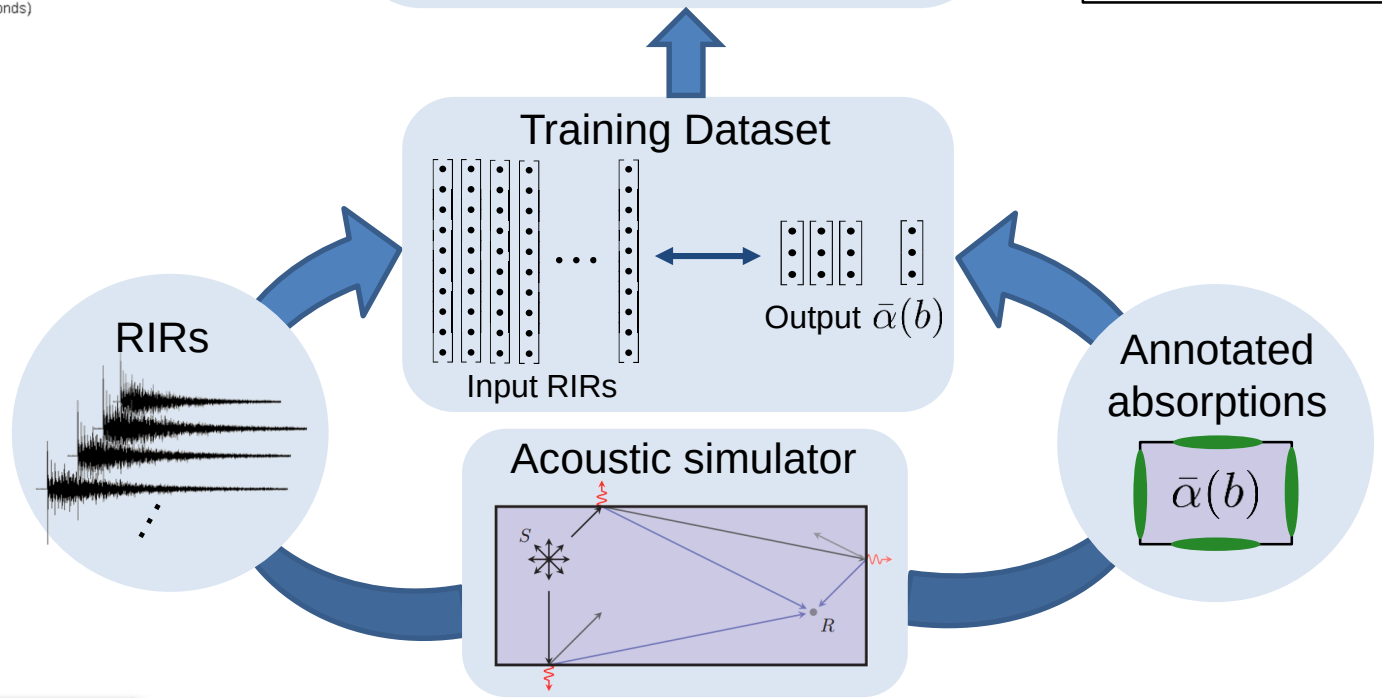
Mean absorption coefficients

$\bar{\alpha}_{\mathrm{NN}}(125\mathrm{Hz})$

$\bar{\alpha}_{\mathrm{NN}}(250\mathrm{Hz})$

$\vdots$

$\bar{\alpha}_{\mathrm{NN}}(4000\mathrm{Hz})$



OUTLINE

1. **Introduction**
2. Simulated Datasets
3. Neural Networks
4. Experiments & Results
5. Conclusion & Future Work

OUTLINE

1. Introduction
2. **Simulated Datasets**
3. Neural Networks
4. Experiments & Results
5. Conclusion & Future Work

RIR training set simulation

Tradeoffs needed to be considered

RIR training set simulation

Tradeoffs needed to be considered

Realism vs. Computational complexity	

RIR training set simulation

Tradeoffs needed to be considered

Realism vs. Computational complexity

Roomsim *[Schimmel et al. 2007]* (shoebox)

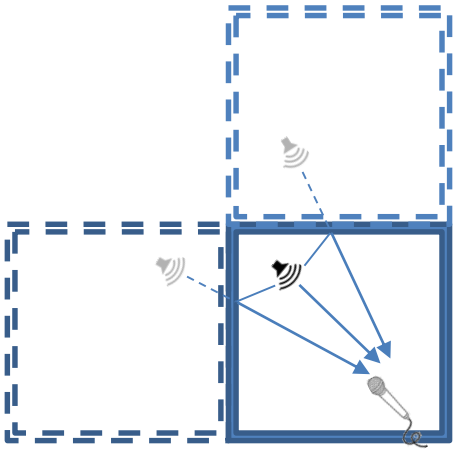
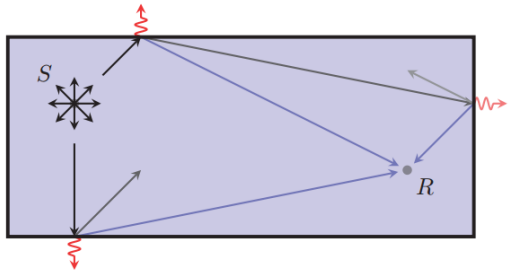


Image source method

+



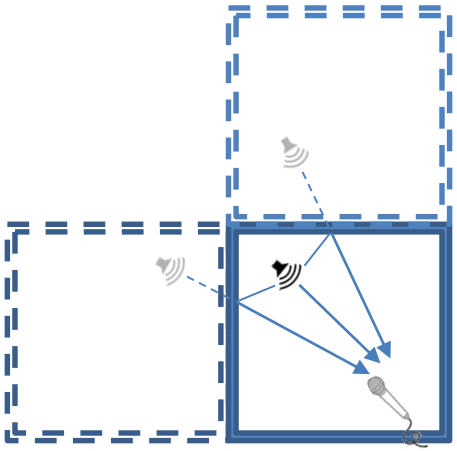
Diffuse rain algorithm

RIR training set simulation

Tradeoffs needed to be considered

Realism vs. Computational complexity

Roomsim *[Schimmel et al. 2007]* (shoebox)



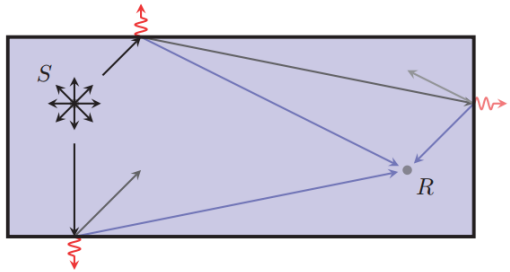


Image source method + Diffuse rain algorithm

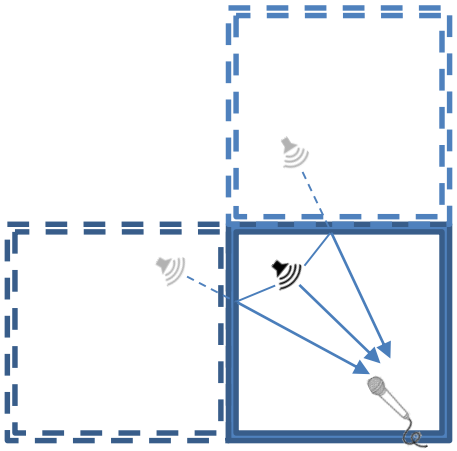
Echo order: 50

Unique, random diffusion coefficient

RIR training set simulation

Tradeoffs needed to be considered

Realism vs. Computational complexity	Diversity vs. Model capactiy
Roomsim *[Schimmel et al. 2007]* (shoebox)	



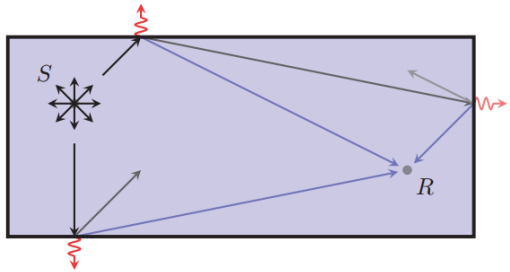


Image source method + Diffuse rain algorithm

Echo order: 50

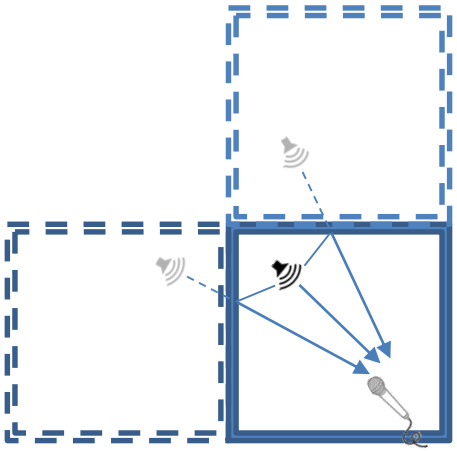
Unique, random diffusion coefficient

RIR training set simulation

Tradeoffs needed to be considered

Realism vs. Computational complexity	Diversity vs. Model capactiy

Roomsim *[Schimmel et al. 2007]* (shoebox)	• Focus on housing or office-type buildings:
	  • Heights in [2.5, 4] meters
	  • Length/Width in [1.5, 10] meters



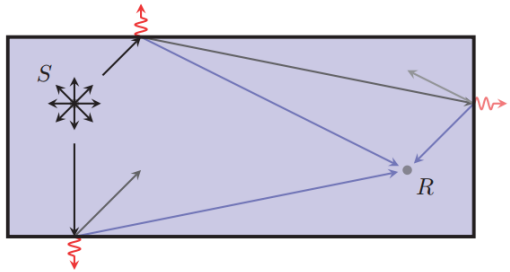
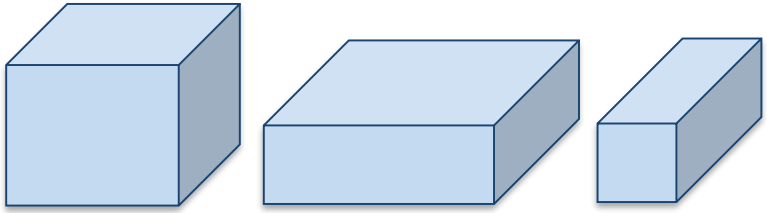


Image source method + Diffuse rain algorithm

Echo order: 50

Unique, random diffusion coefficient

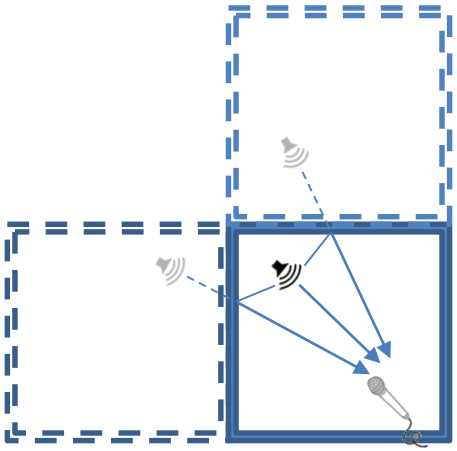


RIR training set simulation

Tradeoffs needed to be considered

Realism vs. Computational complexity

Roomsim *[Schimmel et al. 2007]* (shoebox)



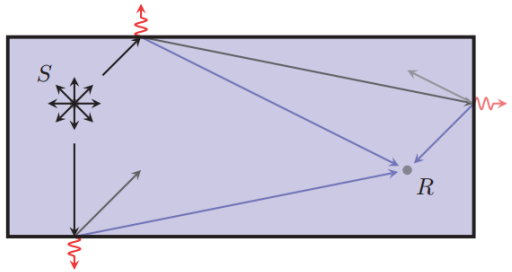


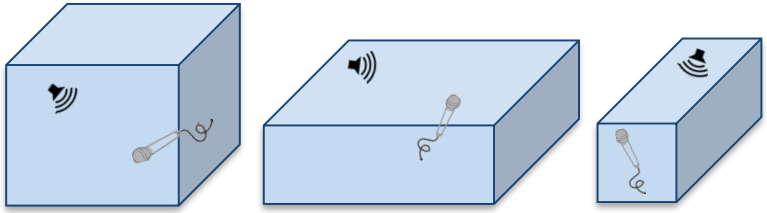
Image source method + Diffuse rain algorithm

Echo order: 50

Unique, random diffusion coefficient

Diversity vs. Model capactiy

- Focus on housing or office-type buildings:
  - Heights in [2.5, 4] meters
  - Length/Width in [1.5, 10] meters



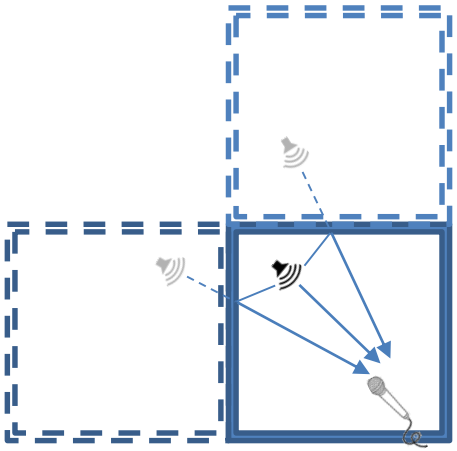
- Omnidirectional source and receiver placed uniformly at random in the room

RIR training set simulation

Tradeoffs needed to be considered

Realism vs. Computational complexity

Roomsim *[Schimmel et al. 2007]* (shoebox)



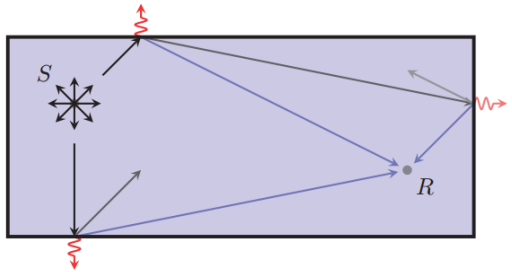


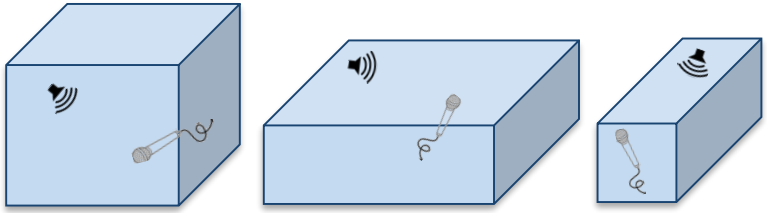
Image source method + Diffuse rain algorithm

Echo order: 50

Unique, random diffusion coefficient

Diversity vs. Model capactiy

- Focus on housing or office-type buildings:
  - Heights in [2.5, 4] meters
  - Length/Width in [1.5, 10] meters



- Omnidirectional source and receiver placed uniformly at random in the room

- Training set: 15,000 RIRs
- Development set : 5,000 RIRs

Room acoustics sampling

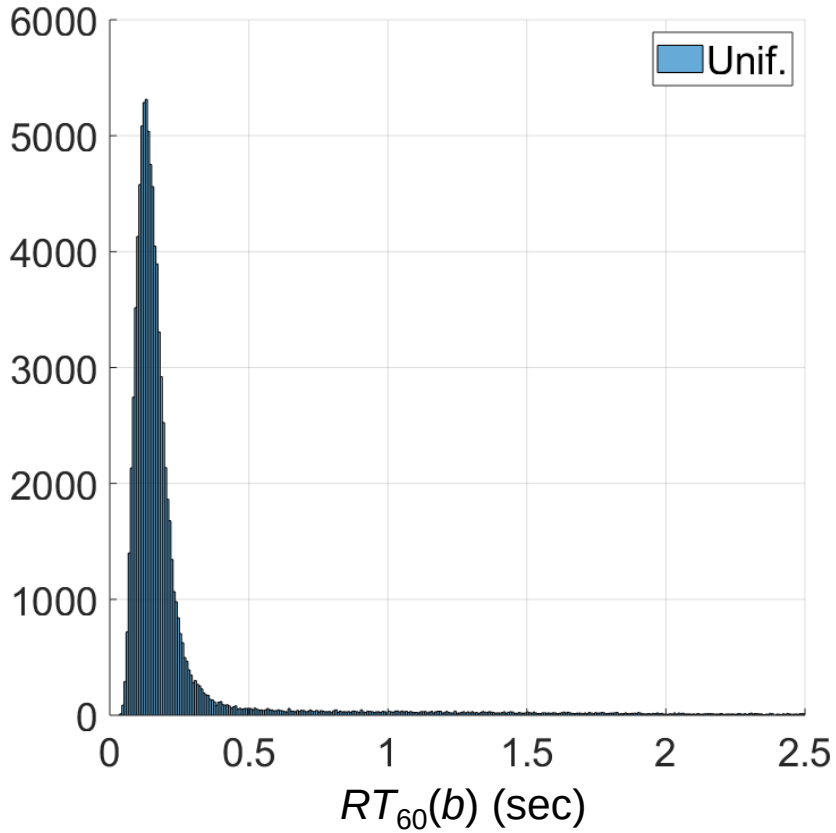
1) Uniformly sampled (**Unif.**) dataset

Absorption coefficients of 6 surfaces in 6 octave bands drawn uniformly at random in [0, 1]

Room acoustics sampling

1) Uniformly sampled (**Unif.**) dataset

Absorption coefficients of 6 surfaces in 6 octave bands drawn uniformly at random in [0, 1]



Room acoustics sampling

2) Reflectivity-Biased (**RB**) dataset

For each surface type (wall, ceiling, floor), toss a coin:

Room acoustics sampling

2) Reflectivity-Biased (**RB**) dataset

For each surface type (wall, ceiling, floor), toss a coin:

- On heads: frequency-independent absorption coefficient in [0,0.12] (*reflective*)

Room acoustics sampling

2) Reflectivity-Biased (**RB**) dataset

For each surface type (wall, ceiling, floor), toss a coin:

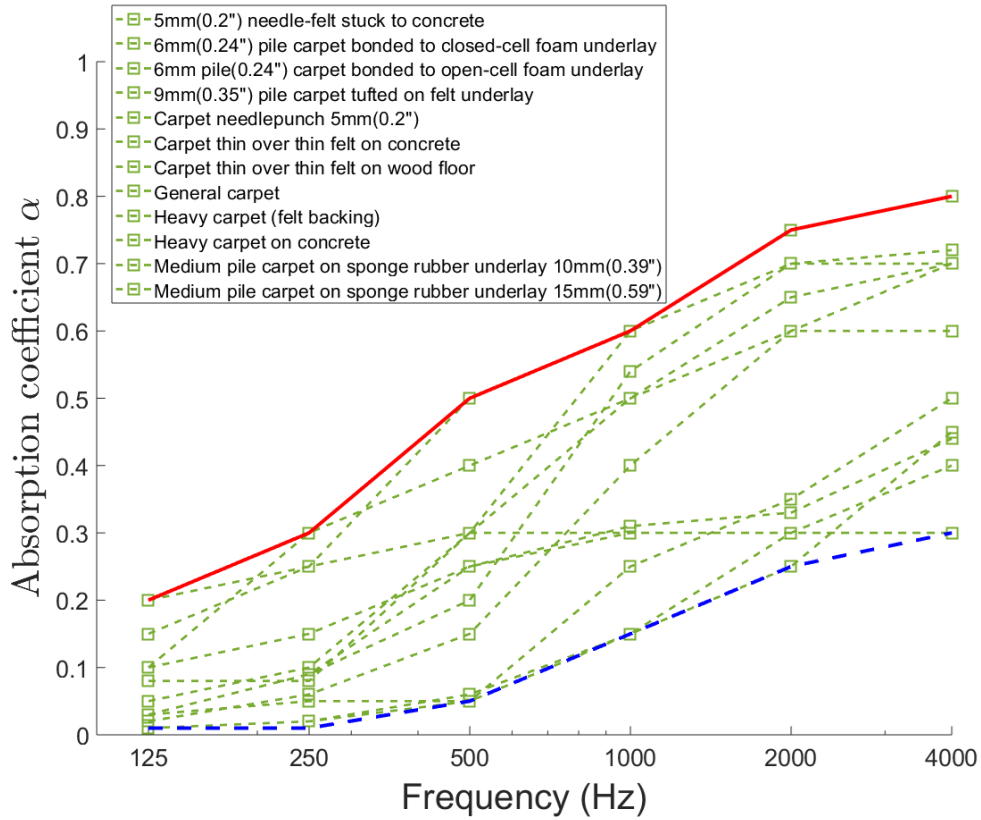
- On heads: frequency-independent absorption coefficient in [0,0.12] (*reflective*)
- On tails: random absorption profile inside a realistic range (*treated*)

Room acoustics sampling

2) Reflectivity-Biased (**RB**) dataset

For each surface type (wall, ceiling, floor), toss a coin:

- On heads: frequency-independent absorption coefficient in [0,0.12] (*reflective*)
- On tails: random absorption profile inside a realistic range (*treated*)

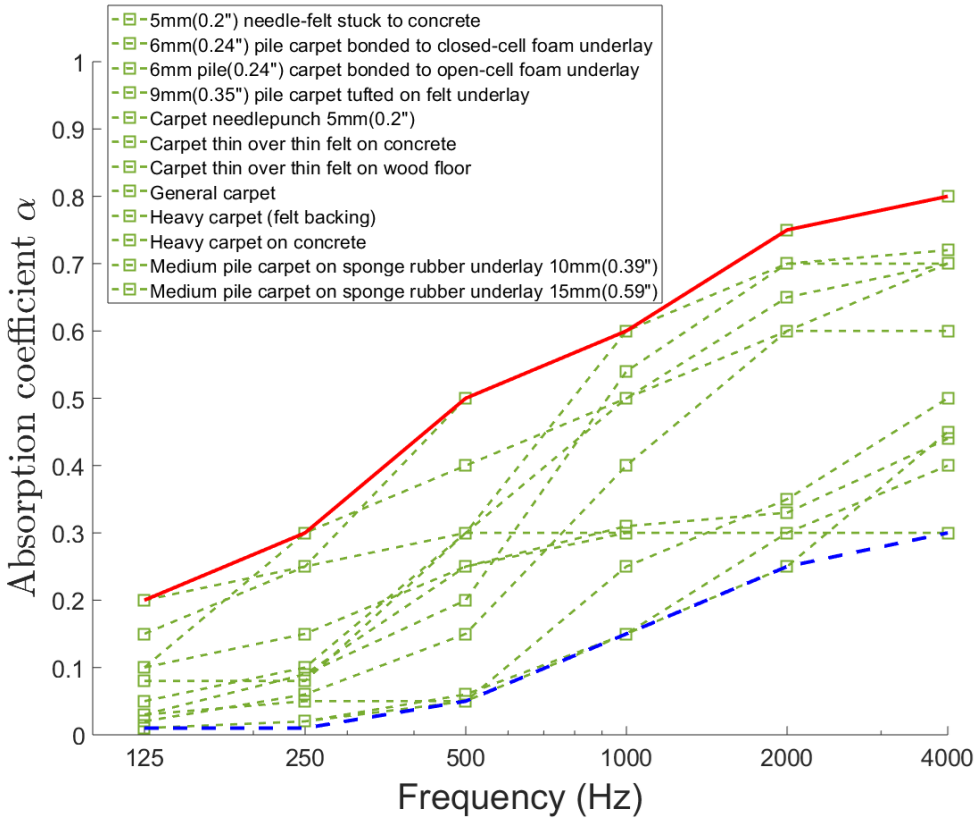


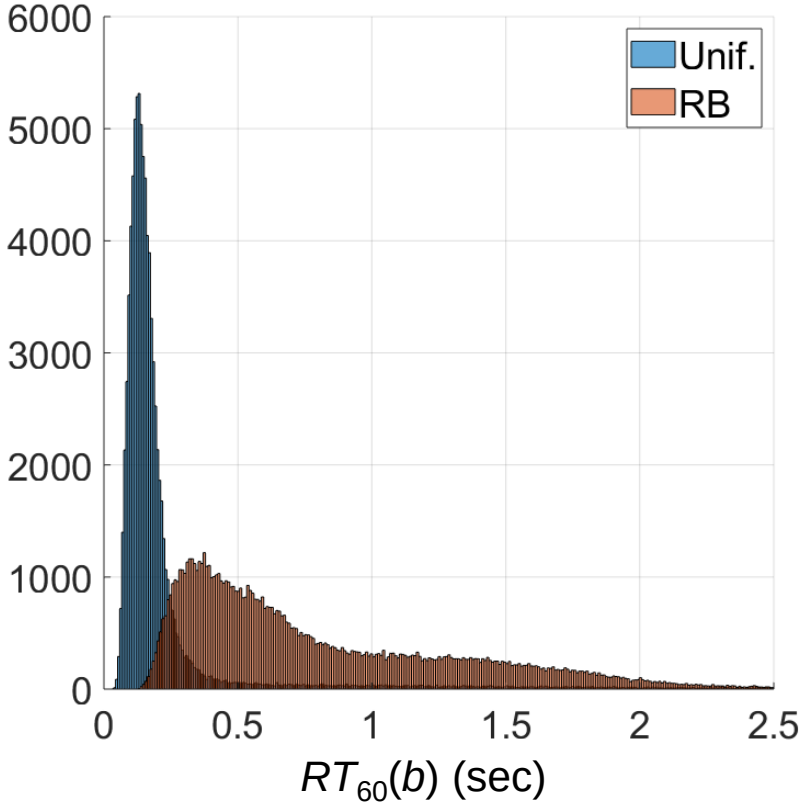
Room acoustics sampling

2) Reflectivity-Biased (**RB**) dataset

For each surface type (wall, ceiling, floor), toss a coin:

- On heads: frequency-independent absorption coefficient in [0,0.12] (*reflective*)
- On tails: random absorption profile inside a realistic range (*treated*)





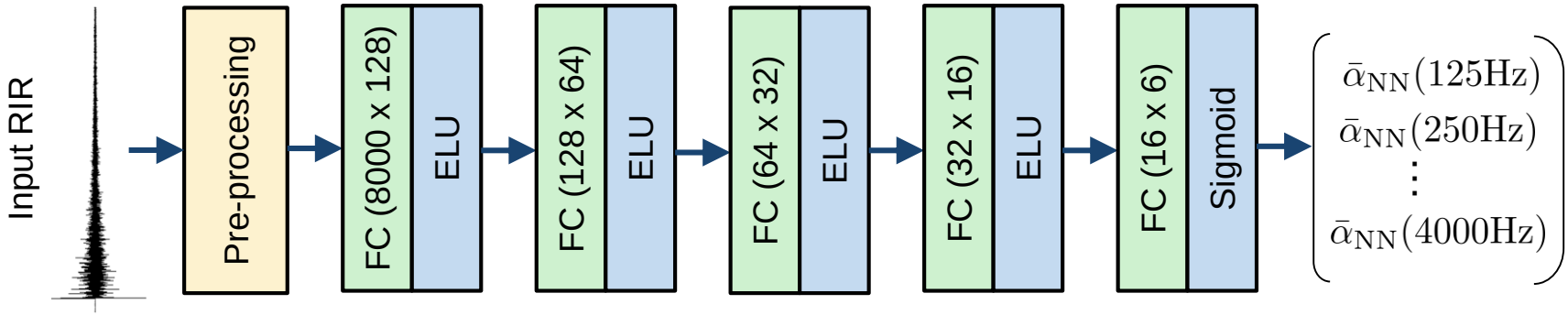
OUTLINE

1. Introduction
2. **Simulated Datasets**
3. Neural Networks
4. Experiments & Results
5. Conclusion & Future Work

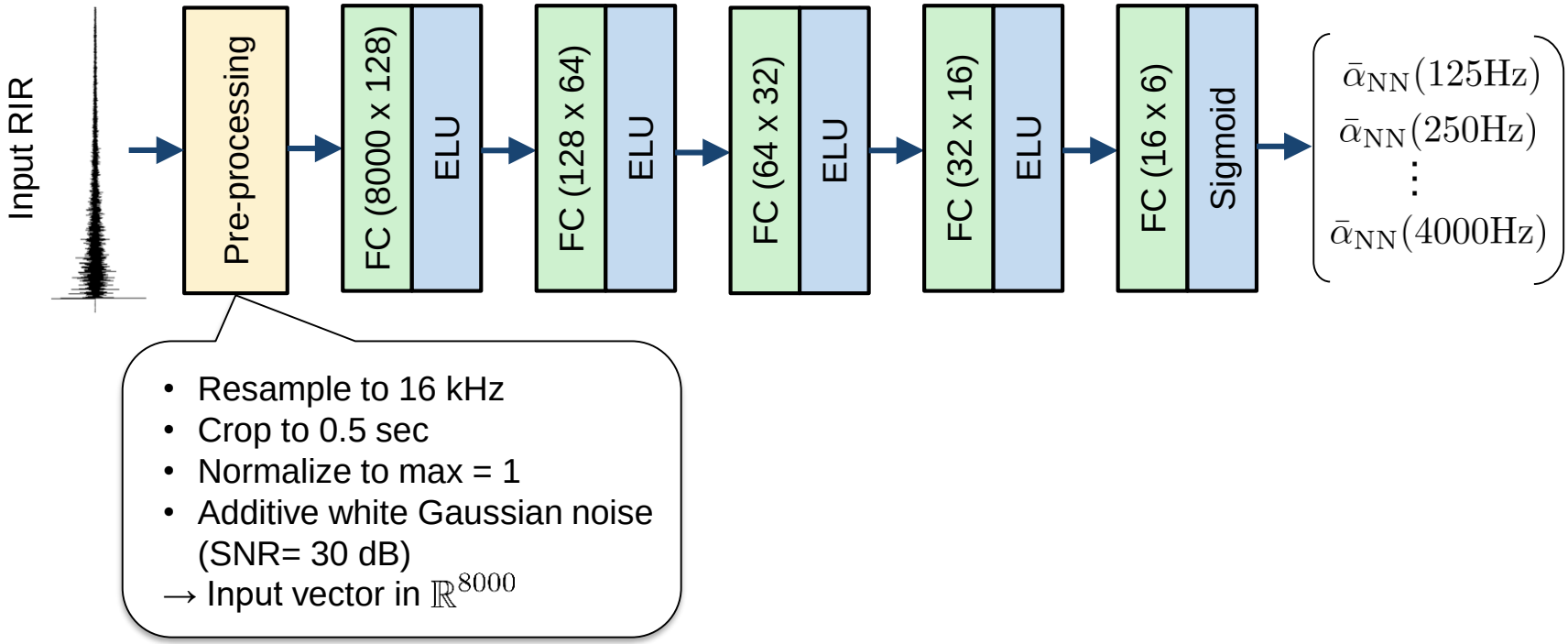
OUTLINE

1. Introduction
2. Simulated Datasets
3. **Neural Networks**
4. Experiments & Results
5. Conclusion & Future Work

1) Multilayer Perceptron (MLP)



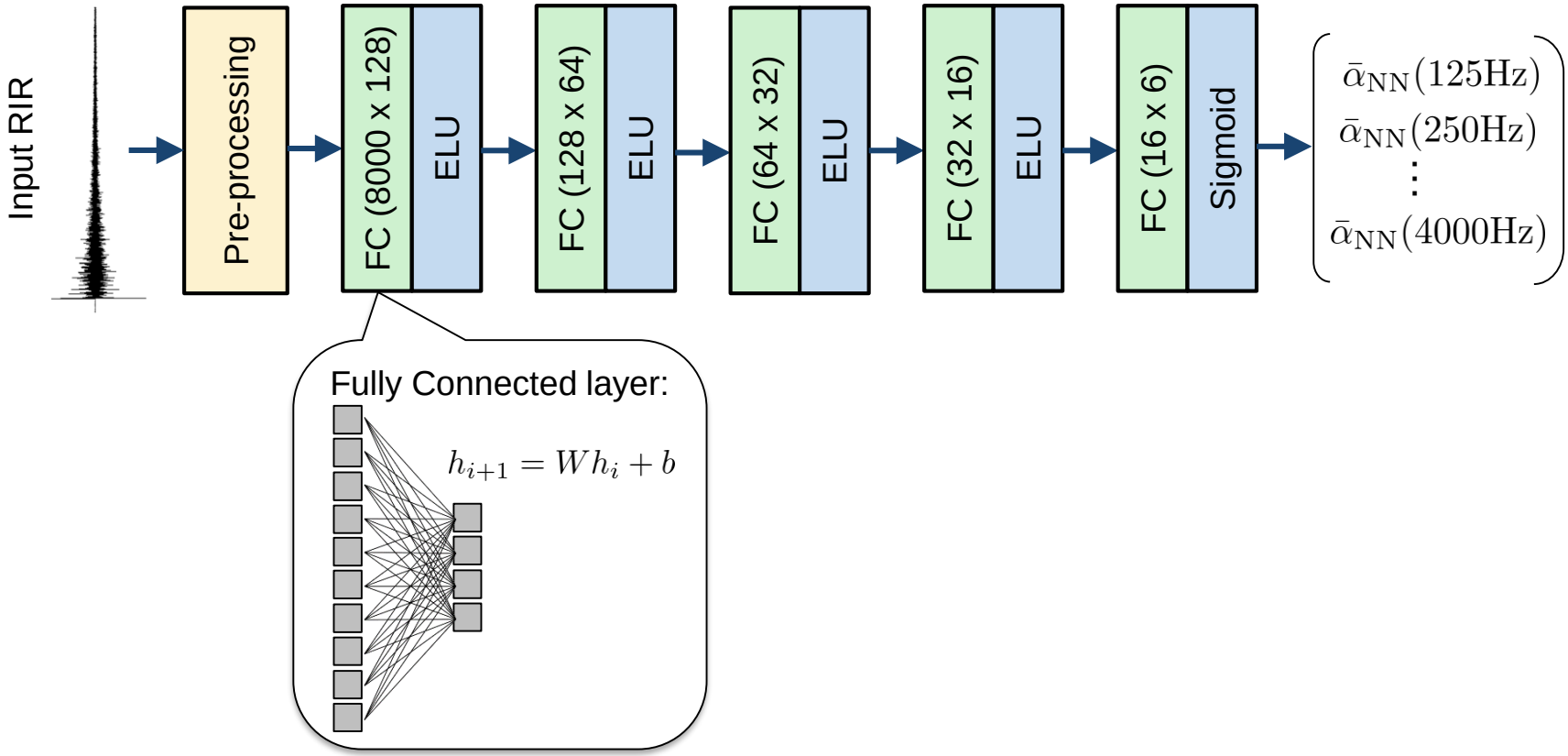
1) Multilayer Perceptron (MLP)



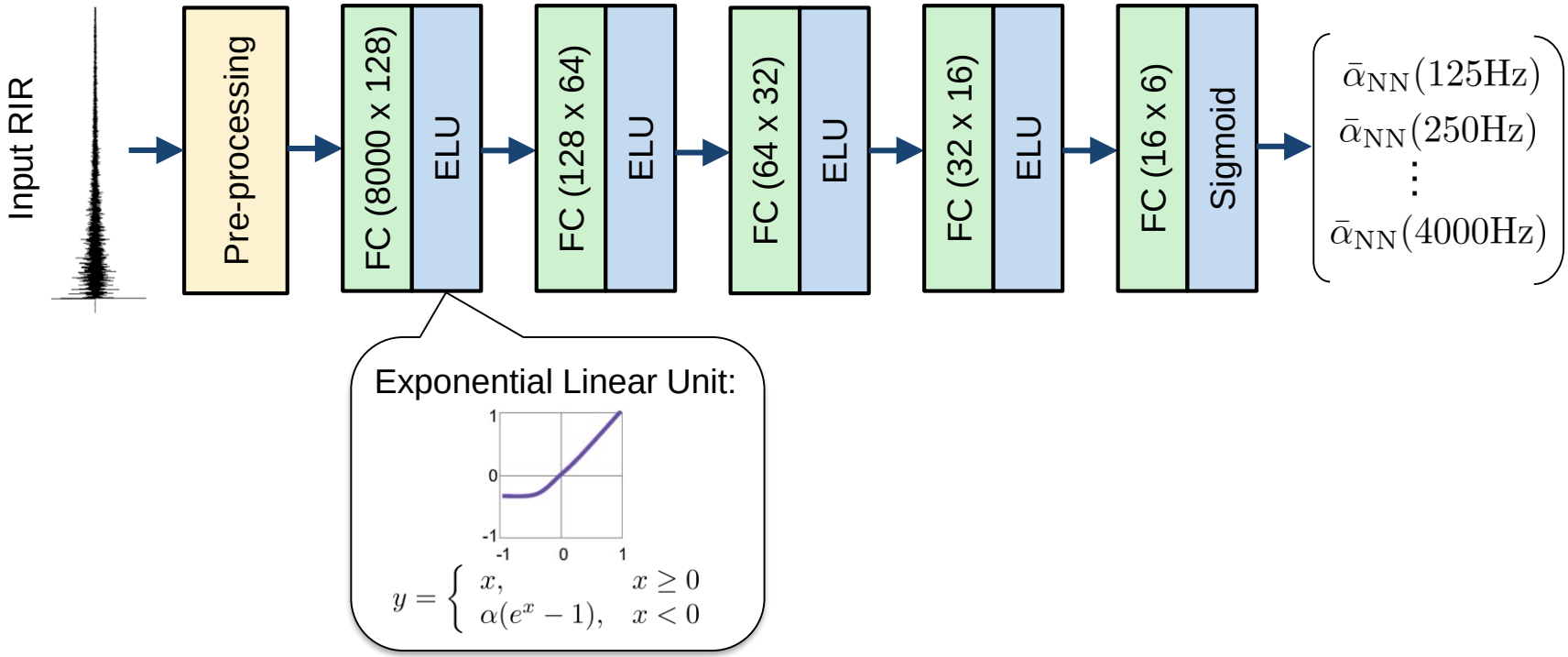
- Resample to 16 kHz
- Crop to 0.5 sec
- Normalize to max = 1
- Additive white Gaussian noise (SNR= 30 dB)

→ Input vector in $\mathbb{R}^{8000}$

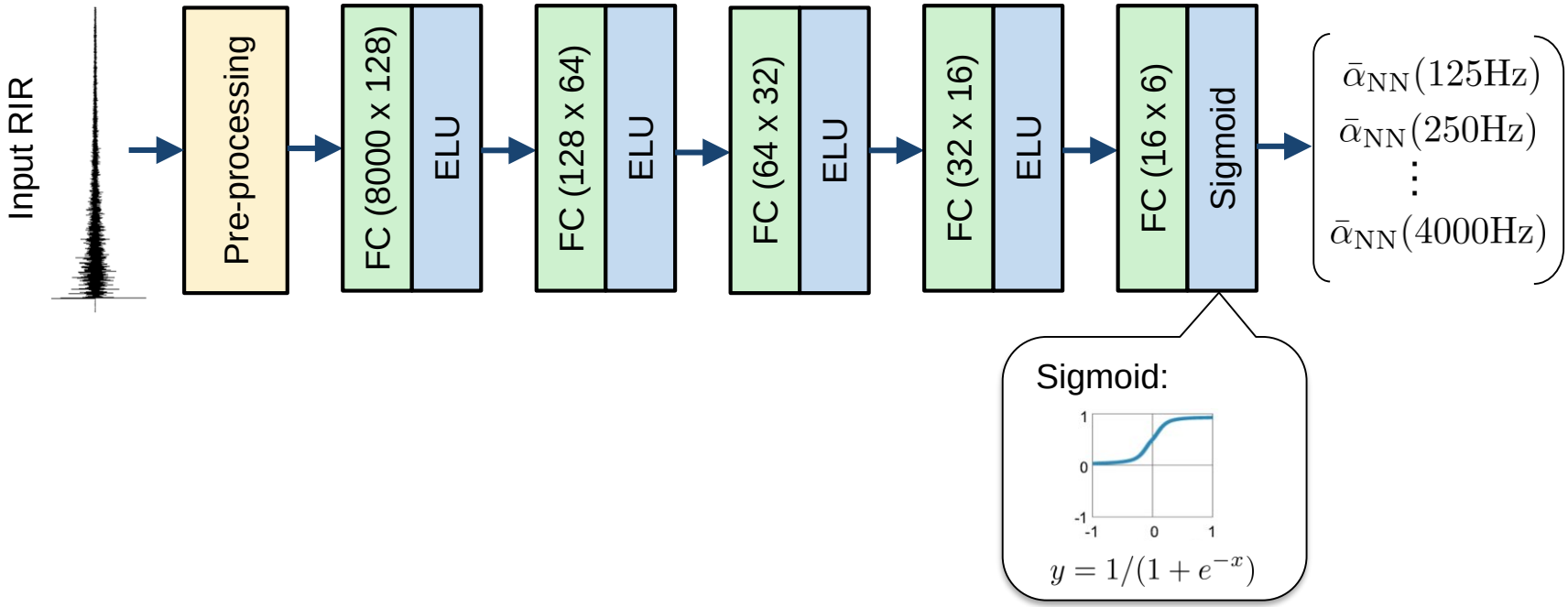
1) Multilayer Perceptron (MLP)



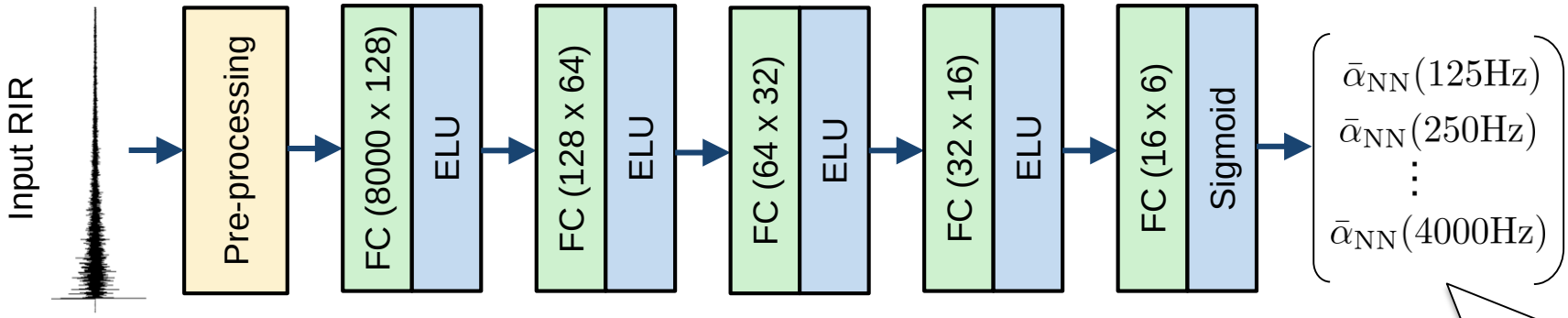
1) Multilayer Perceptron (MLP)



1) Multilayer Perceptron (MLP)

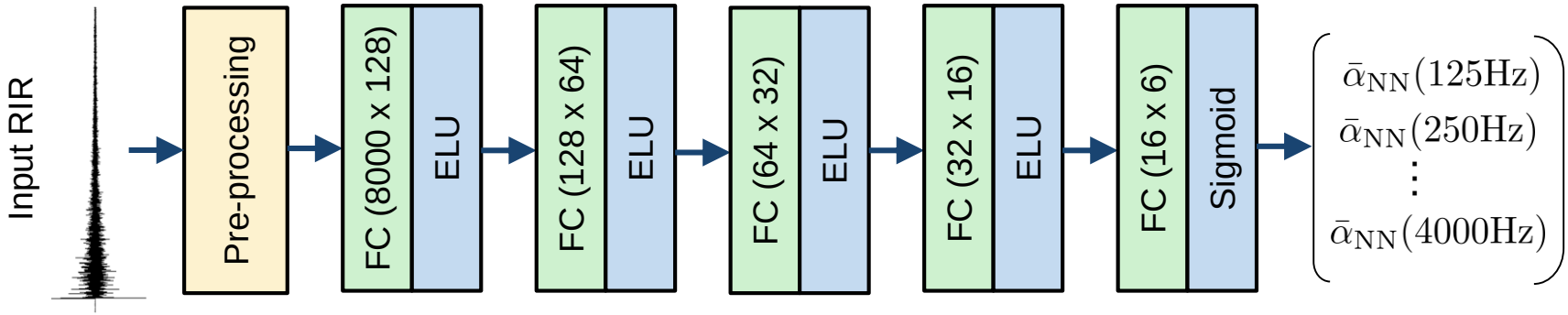


1) Multilayer Perceptron (MLP)

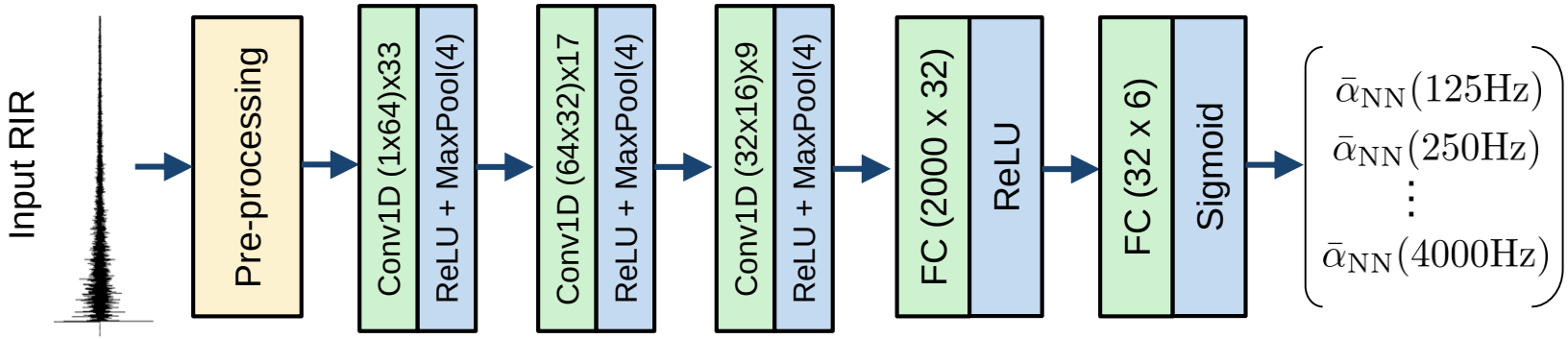


- Output vector in $[0,1]^6$
- Loss Function = Mean Squared Error
- Optimal parameters on dev. set over 200 epochs

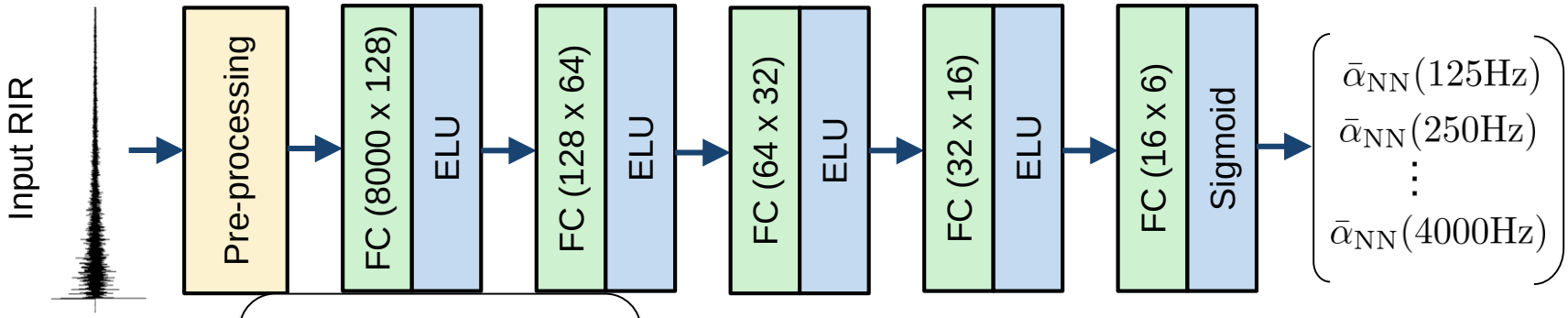
1) Multilayer Perceptron (MLP)



2) Convolutional Neural Newtork (CNN)

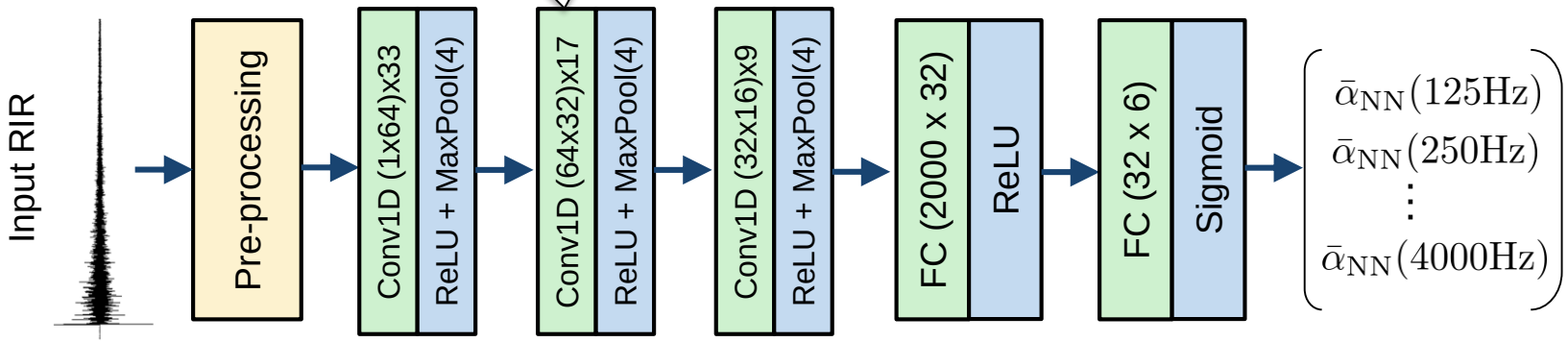


1) Multilayer Perceptron (MLP)

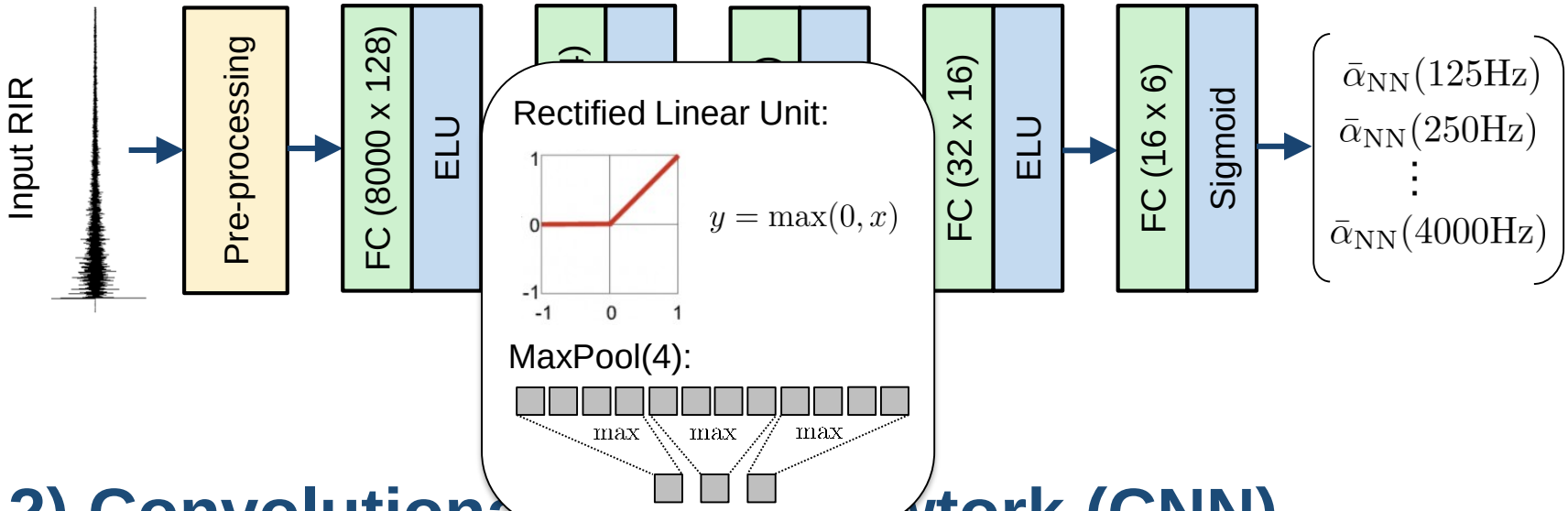


1D convolutional layer:
- 64 input channels
- 32 output channels
- Kernel size: 17

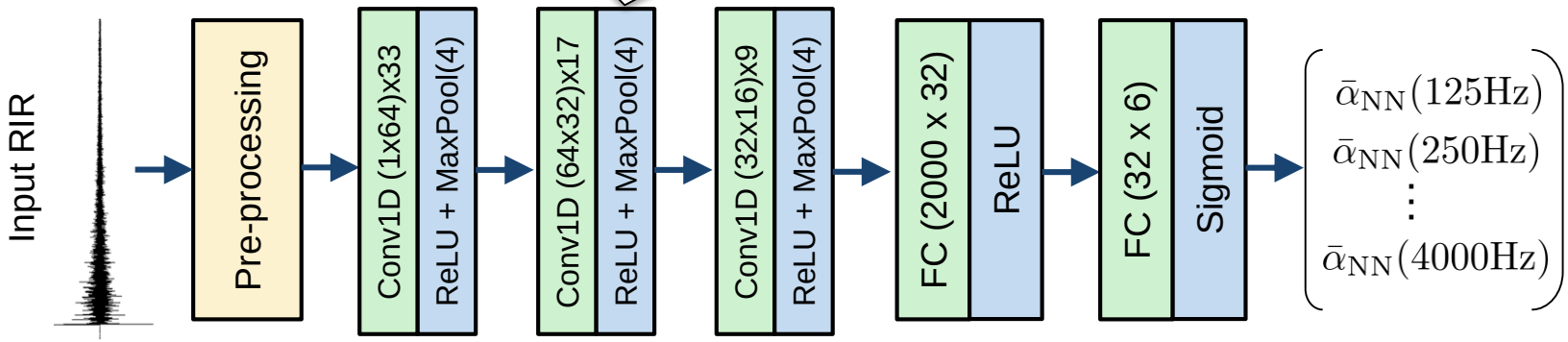
2) Convolutional Neural Newtork (CNN)



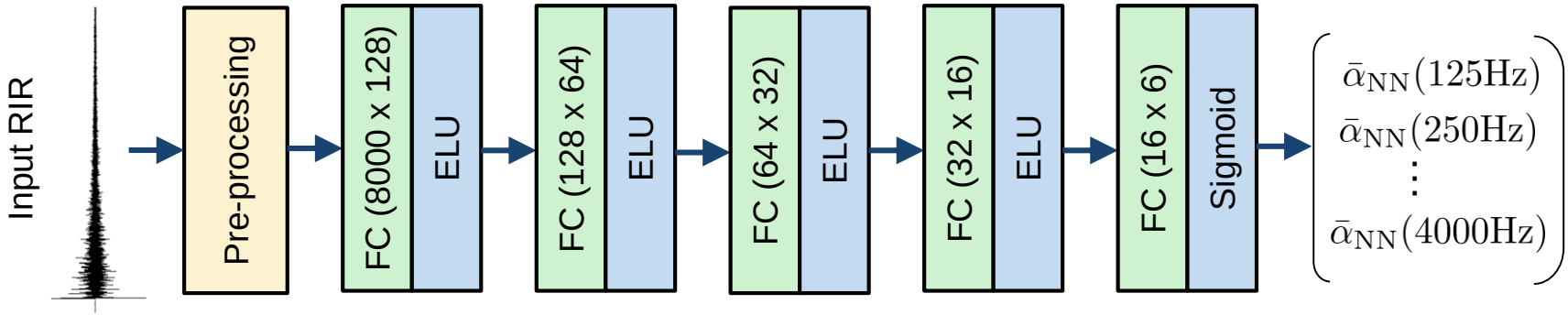
1) Multilayer Perceptron (MLP)



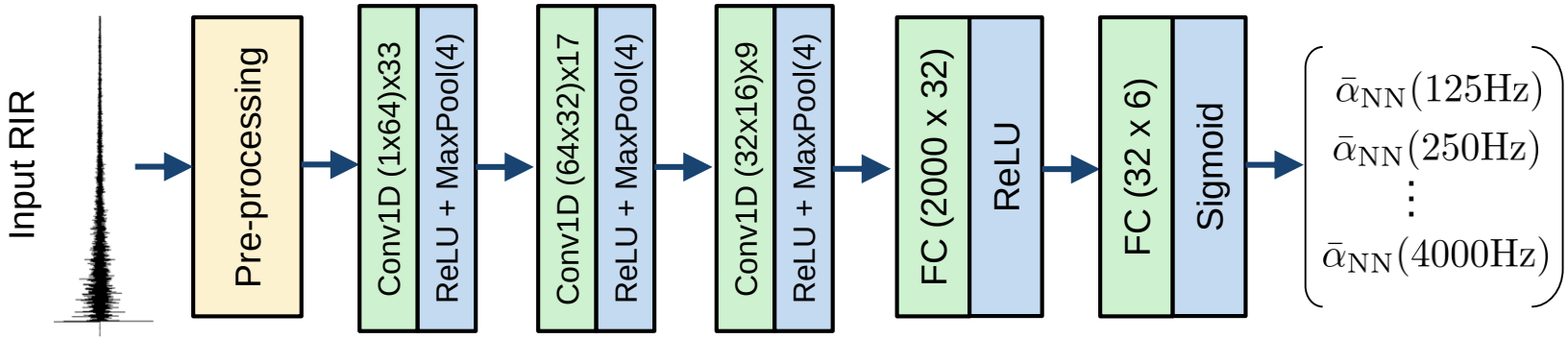
2) Convolutional Neural Newtork (CNN)



1) Multilayer Perceptron (MLP)



2) Convolutional Neural Newtork (CNN)



OUTLINE

1. Introduction
2. Simulated Datasets
3. **Neural Networks**
4. Experiments & Results
5. Conclusion & Future Work

OUTLINE

1. Introduction
2. Simulated Datasets
3. **Neural Networks**
4. **Experiments & Results**
5. Conclusion & Future Work

4. Experiments & Results

Test set: Same simulator, real material profiles, 5 room geometries, 500 RIRs

4. Experiments & Results

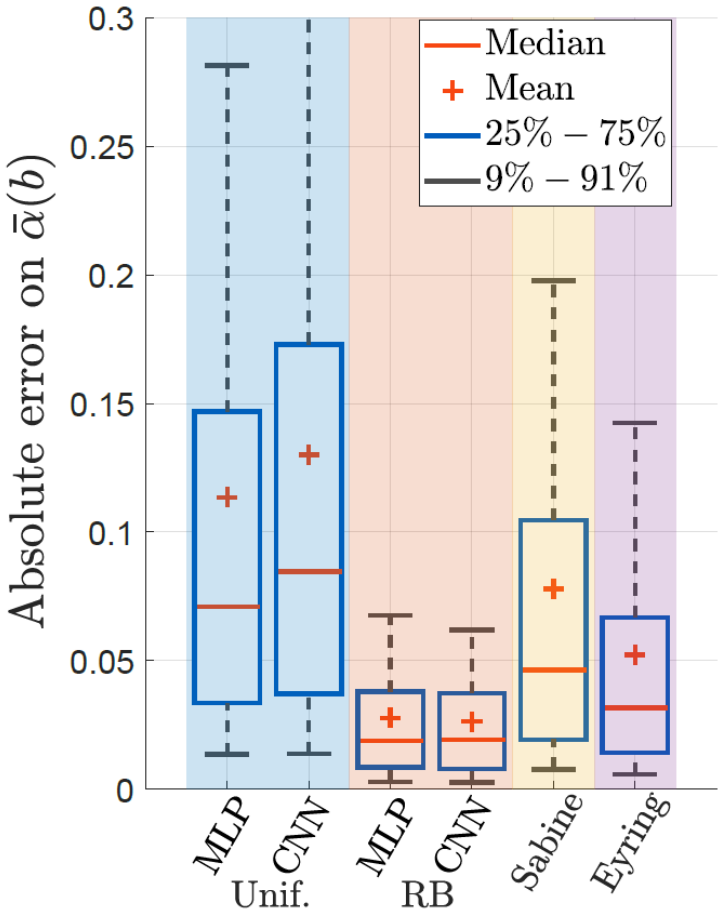
Test set: Same simulator, real material profiles, 5 room geometries, 500 RIRs

- Comparing the two training sets (**Unif.**, **RB**) and the two neural networks (**MLP, CNN**) against Sabine and Eyring's laws (given true S_{tot} and V)

4. Experiments & Results

Test set: Same simulator, real material profiles, 5 room geometries, 500 RIRs

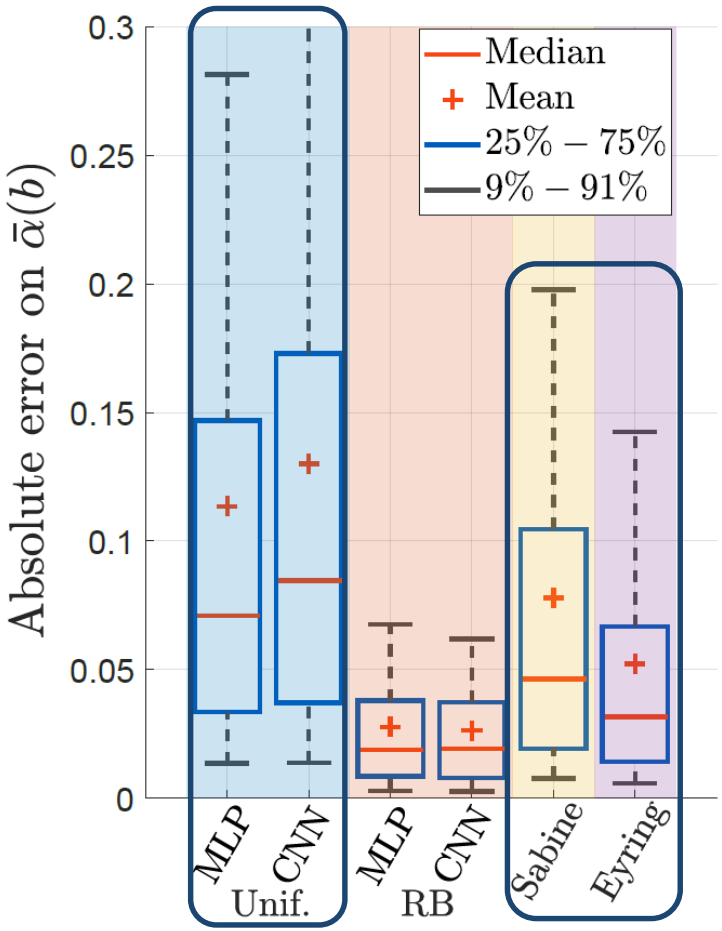
- Comparing the two training sets (**Unif.**, **RB**) and the two neural networks (**MLP, CNN**) against Sabine and Eyring's laws (given true S_{tot} and V)



4. Experiments & Results

Test set: Same simulator, real material profiles, 5 room geometries, 500 RIRs

- Comparing the two training sets (**Unif.**, **RB**) and the two neural networks (**MLP, CNN**) against Sabine and Eyring's laws (given true S_{tot} and V)

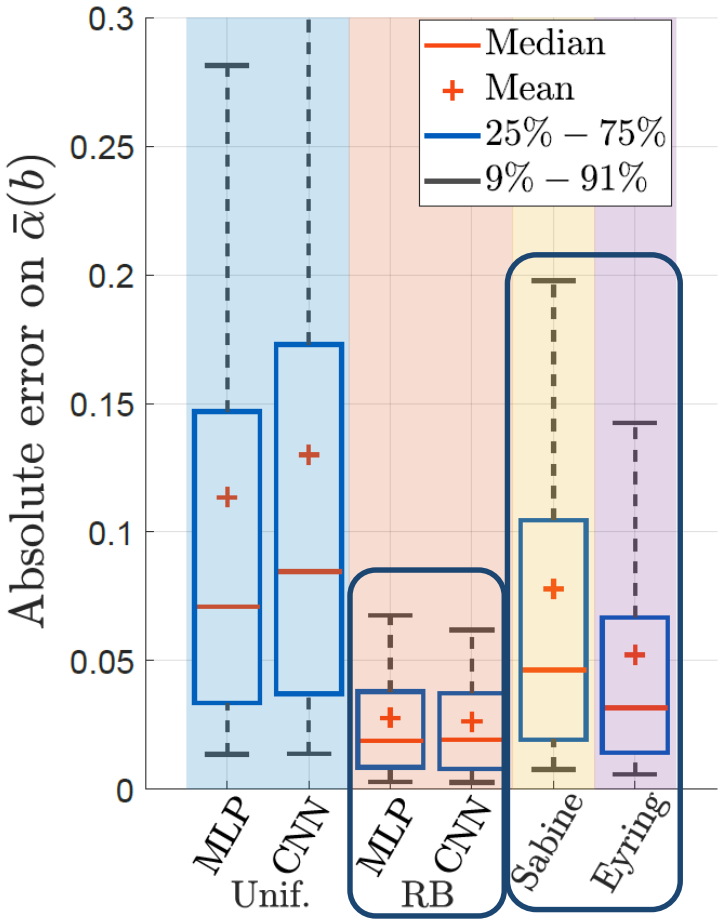


➡ Training on uniformly sampled acoustics fails to outperform reverberation theory

4. Experiments & Results

Test set: Same simulator, real material profiles, 5 room geometries, 500 RIRs

- Comparing the two training sets (**Unif.**, **RB**) and the two neural networks (**MLP, CNN**) against Sabine and Eyring's laws (given true S_{tot} and V)

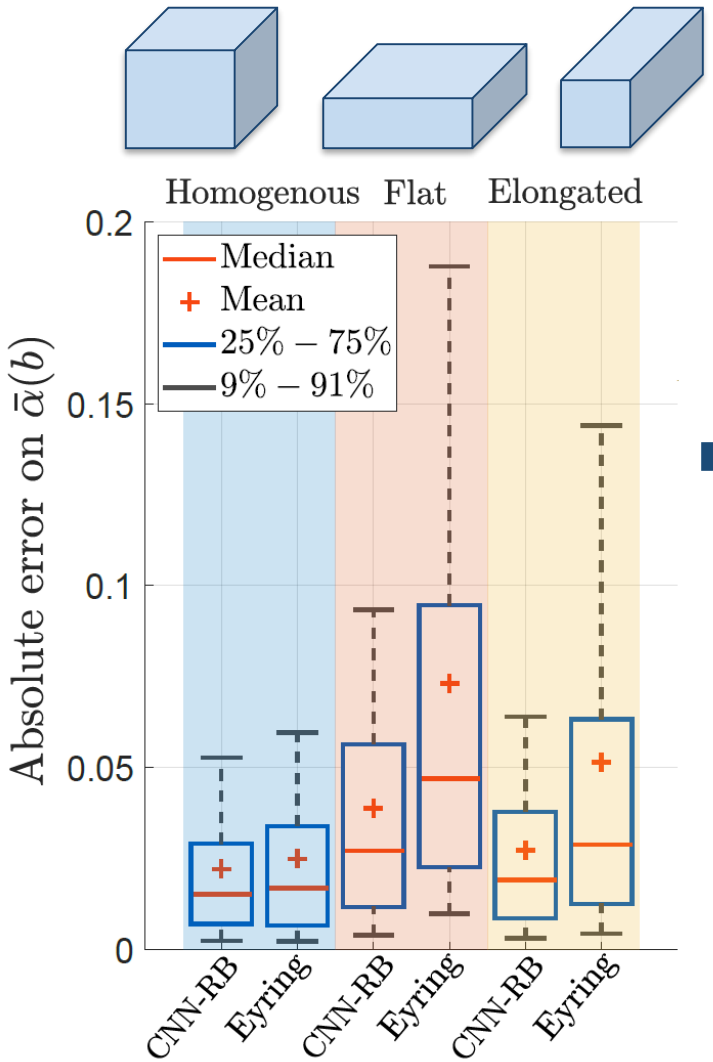


➡ Training on uniformly sampled acoustics fails to outperform reverberation theory

➡ Training on the reflectivity-biased set significantly outperforms both baselines

4. Experiments & Results

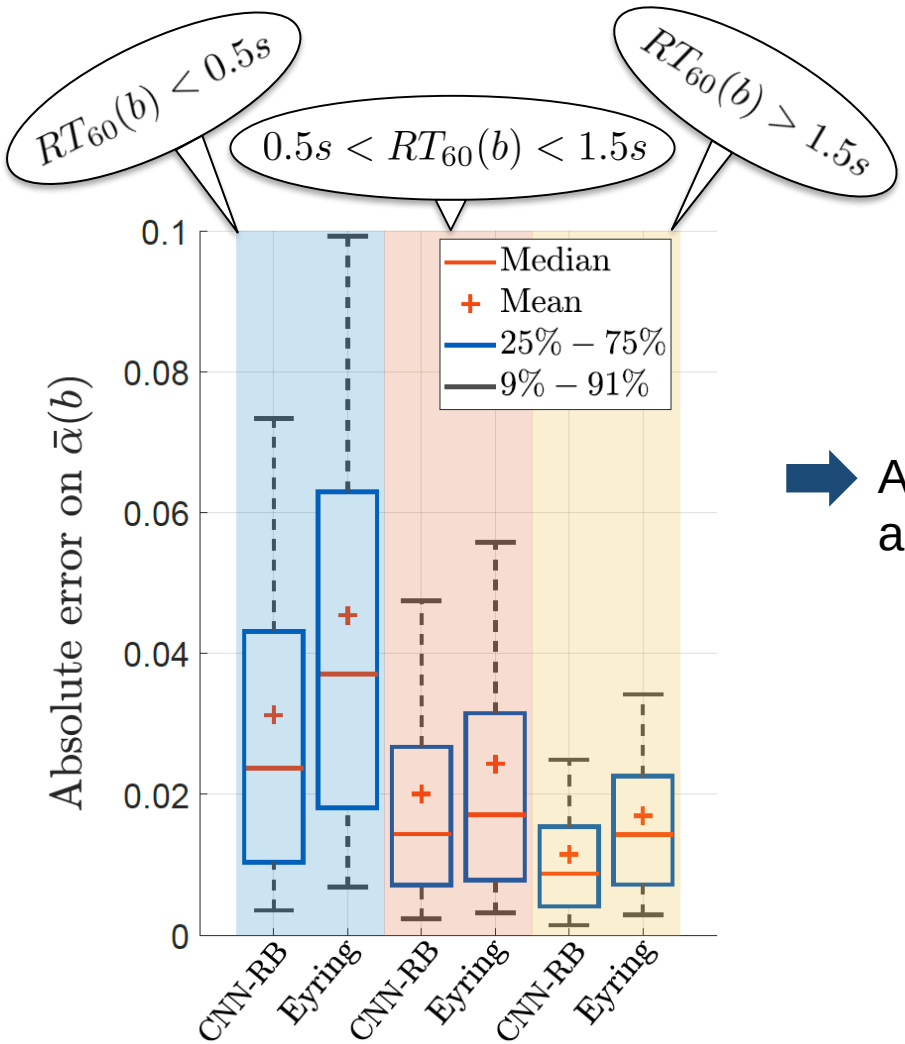
- Influence of room geometry:



➡ Advantage of the learning-based approach clearer under non-homogenous geometries

4. Experiments & Results

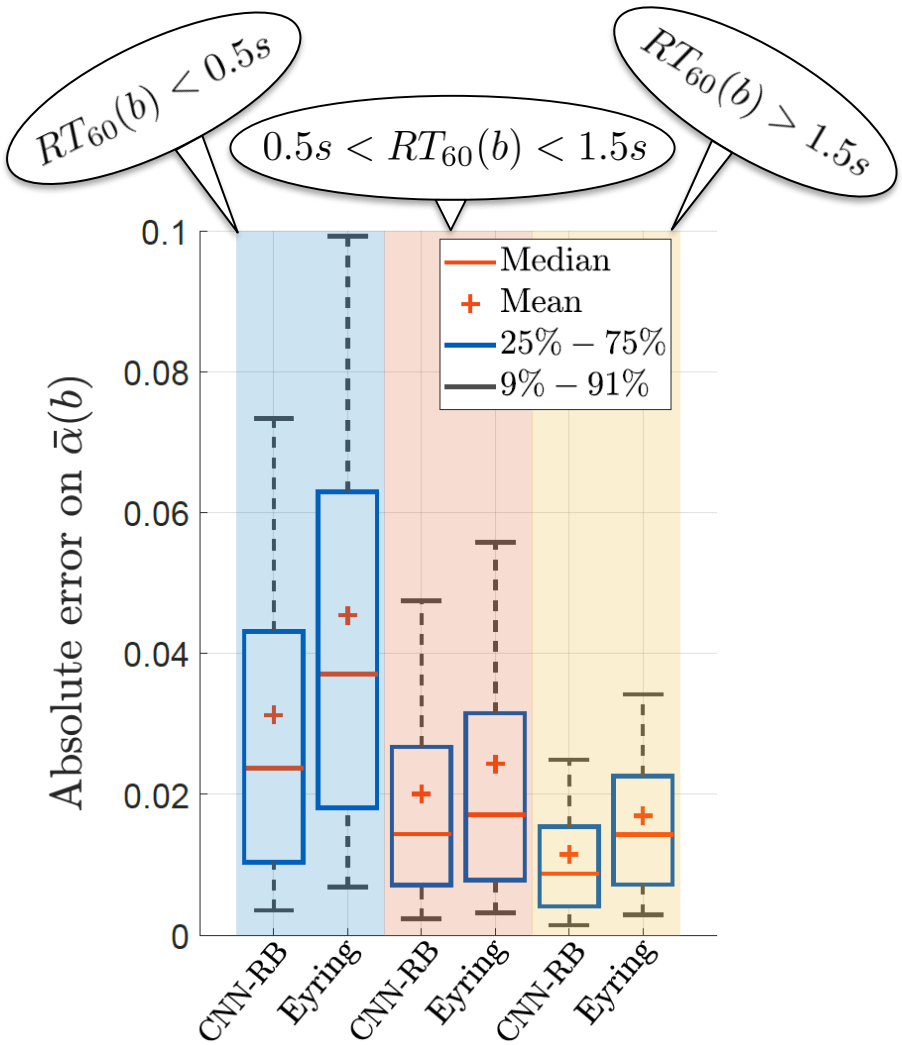
- Influence of reverberation time:



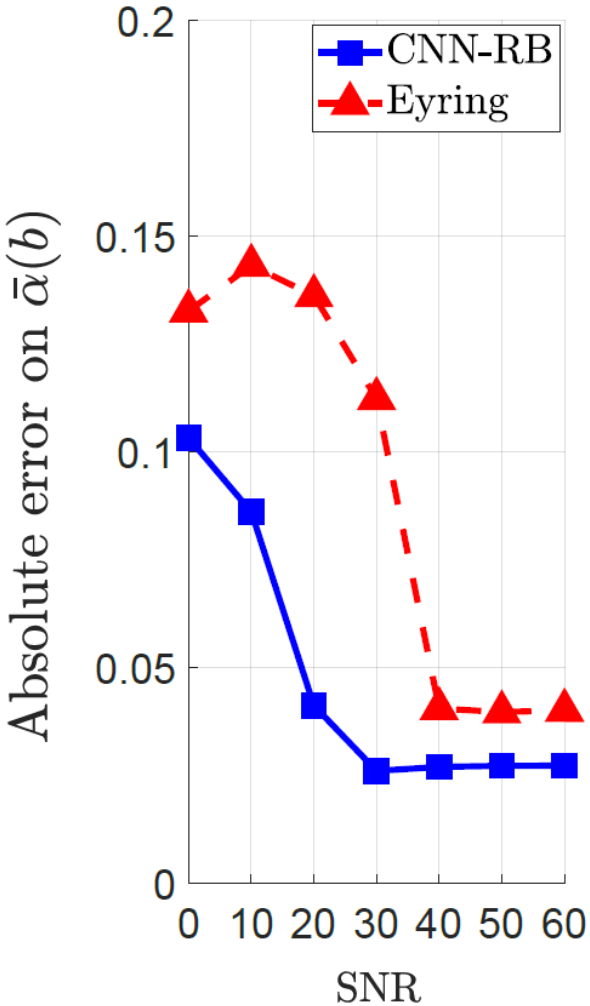
➡ A higher reverberation time helps both approaches

4. Experiments & Results

- Influence of reverberation time:



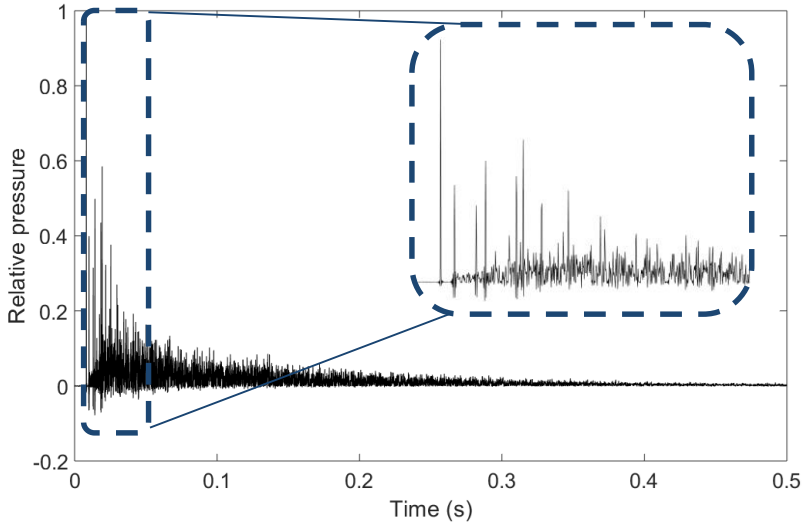
- Influence of noise:



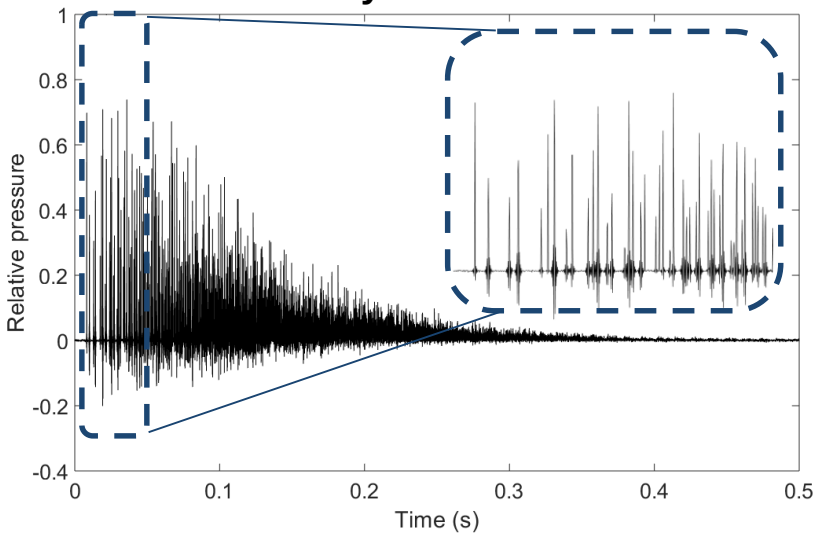
4. Experiments & Results

- Tests on another simulator: **pyroomacoustics** *[Scheibler et al. 2018]*

Roomsim RIR



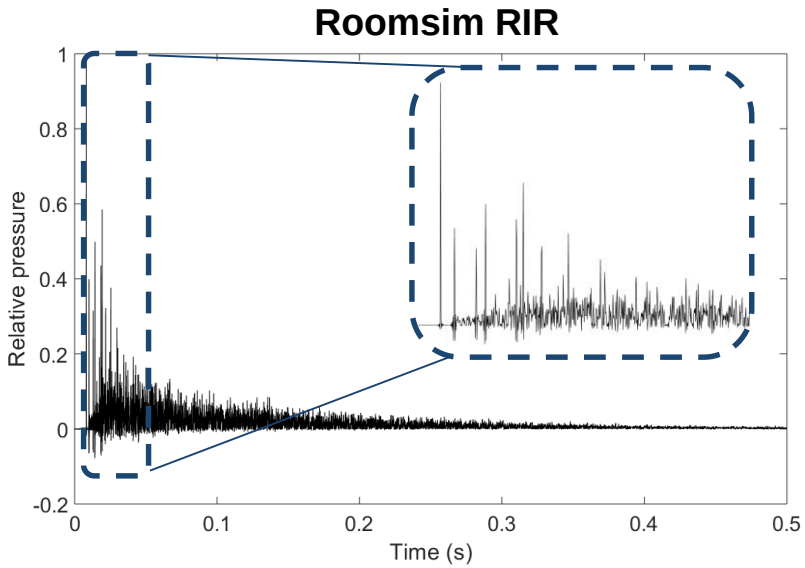
Pyroom. RIR

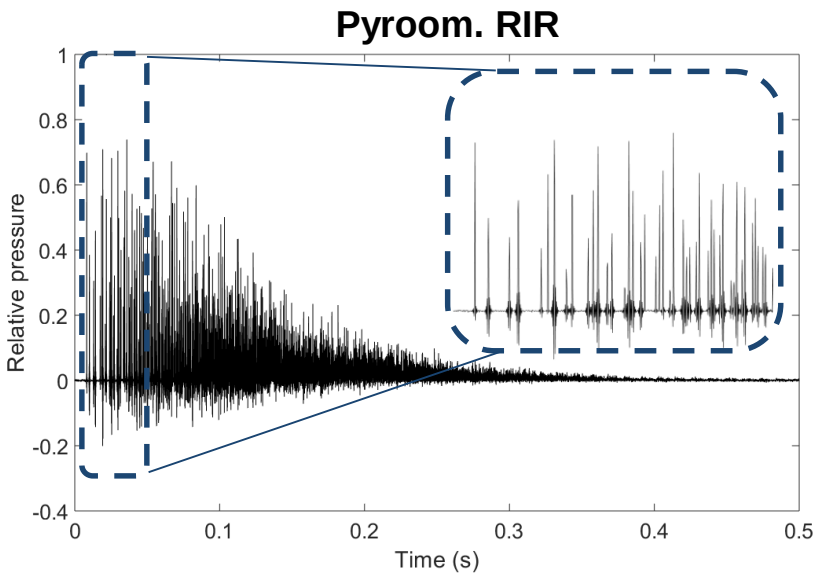


- No diffusion
- Frequency-independent $\bar{\alpha}(b)$
- Sinc kernel

4. Experiments & Results

- Tests on another simulator: **pyroomacoustics** *[Scheibler et al. 2018]*





- No diffusion
- Frequency-independent $\bar{\alpha}(b)$
- Sinc kernel

Mean errors on $\bar{\alpha}(b)$:

	Roomsim	**Pyroom.**
CNN-RB	0.026	**0.085**
MLP-RB	0.024	0.207
Eyring	0.043	0.093

➡ Encouraging robustness to the specific training simulator

OUTLINE

1. Introduction
2. Simulated Datasets
3. Neural Networks
4. **Experiments & Results**
5. Conclusion & Future Work

OUTLINE

1. Introduction
2. Simulated Datasets
3. Neural Networks
4. Experiments & Results
5. **Conclusion & Future Work**

Conclusion

- Simulation-based learning is a promising approach for the acoustic diagnosis of rooms from RIRs
- By exploiting the fine temporal structure of RIRs, the networks outperform reverberation theory in non-homogenous conditions
- Careful acoustic sampling in the training set is crucial

Conclusion

- Simulation-based learning is a promising approach for the acoustic diagnosis of rooms from RIRs
- By exploiting the fine temporal structure of RIRs, the networks outperform reverberation theory in non-homogenous conditions
- Careful acoustic sampling in the training set is crucial

Future work

- Generalization to real measured RIRs
  - Domain adaptation techniques
- Estimating the absorption profiles of individual surfaces
  - Geometry-informed models

THANK YOU