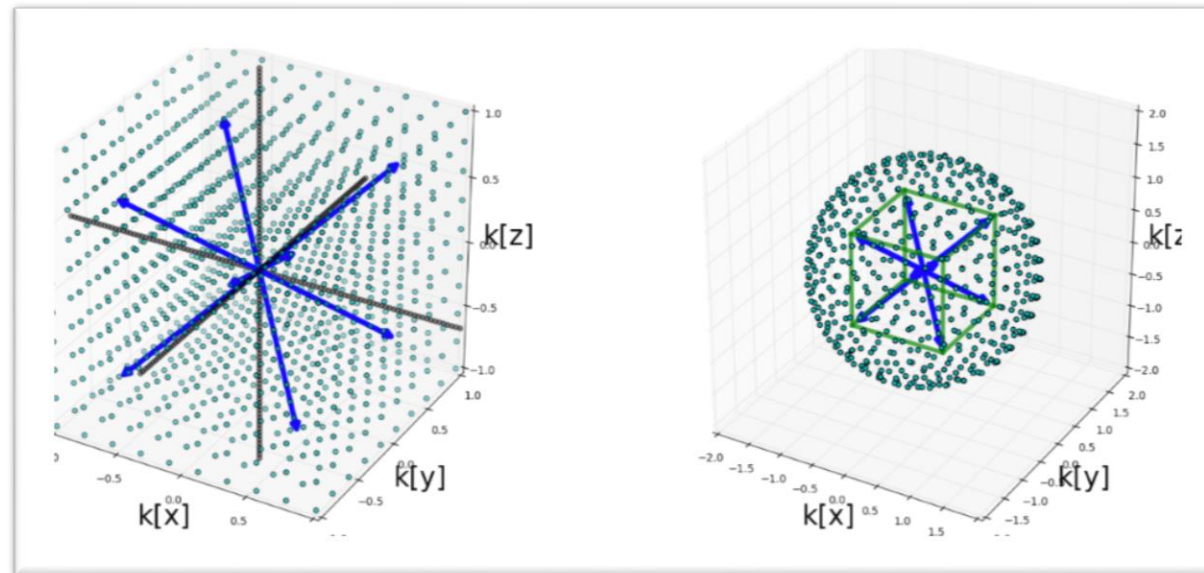


# JOINT ESTIMATION OF THE ROOM GEOMETRY AND MODES WITH COMPRESSED SENSING



Helena Peić Tukuljac, Thach Pham Vu,  
Hervé Lissek and Pierre Vandergheynst



April 19, 2018

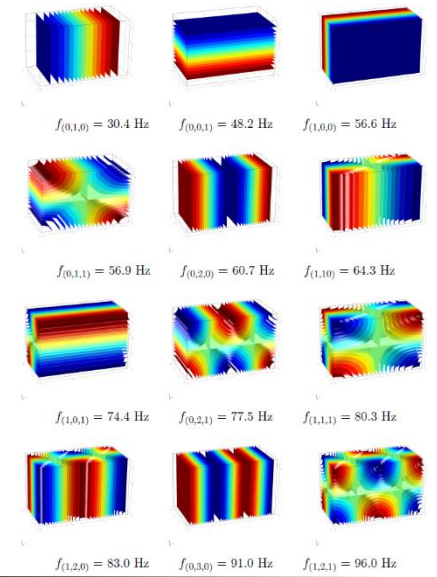
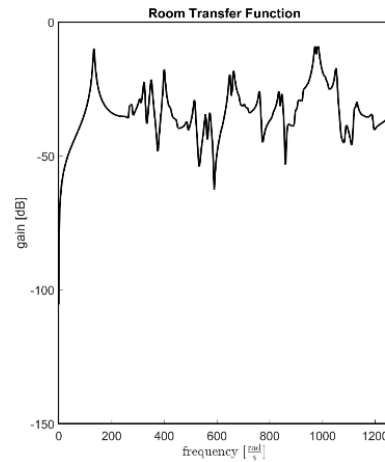
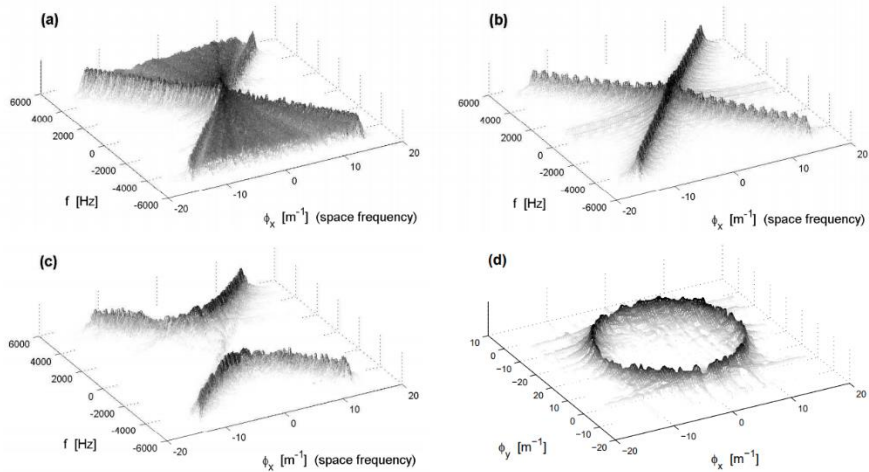


# Content Atlas

## Spatio-temporal sampling relations [1]

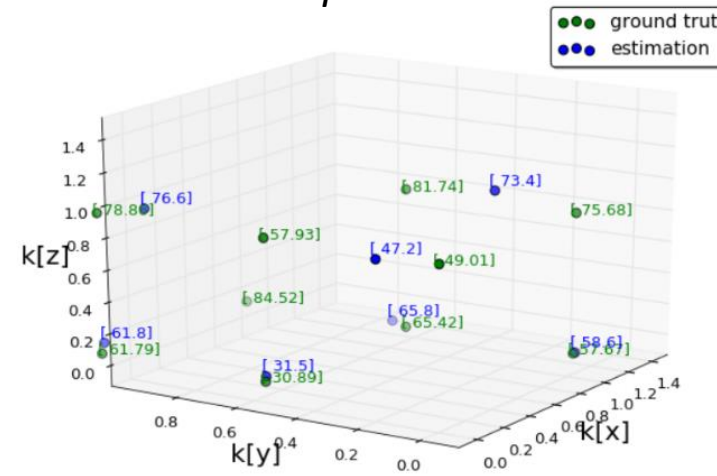
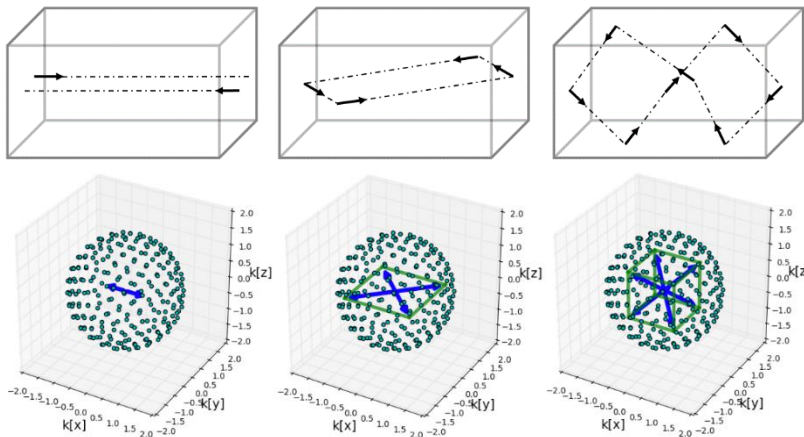
## Room transfer function [2]

## Room modes [3]



## Plane waves and wave vectors

## Acoustical k-space



Introduction

Contribution

# Spatio-Temporal Sampling Relations

Sampling steps:

$$\Delta t \quad \Delta x \quad \Delta y \quad \Delta z$$

Choosing  $\Delta t$  is easy: [1]

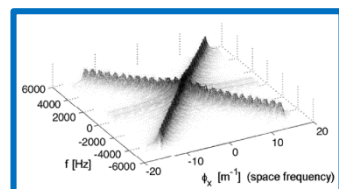
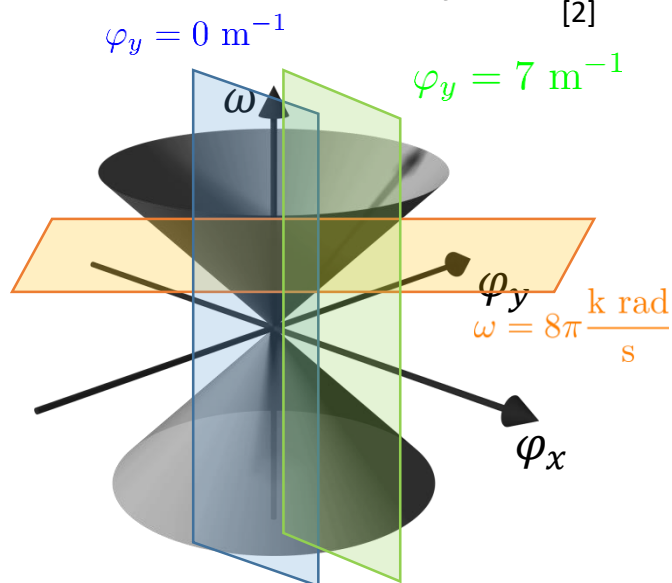
Sampling frequencies:

$$\omega = \frac{2\pi}{\Delta t} \quad \varphi_x = \frac{2\pi}{\Delta x} \quad \varphi_y = \frac{2\pi}{\Delta y} \quad \varphi_z = \frac{2\pi}{\Delta z}$$

$$\omega \geq 2\omega_{\text{cutoff}}$$

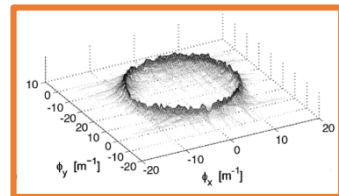
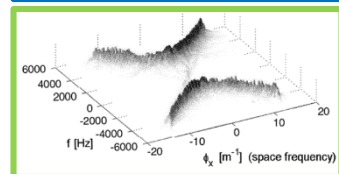
Two points of view on spatial sampling:

View #1: Plenacoustic function [2]



[3]

Connected plane wave sparsity and compressed sensing in 2014



$$\varphi_x^2 + \varphi_y^2 \leq \frac{\omega^2}{c^2}$$

$$\varphi_x^2 + \varphi_y^2 + \varphi_z^2 \leq \frac{\omega^2}{c^2}$$

View #2: Courant–Friedrichs–Lewy condition [4]

If a wave is moving across a discrete spatial grid and we want to compute its amplitude at discrete time steps of equal duration, then this duration must be less than the time for the wave to travel to adjacent grid points

$$\Delta t \sum_{i=1}^n \frac{u_{x_i}}{\Delta x_i} \leq C_{\text{max}}$$

$$\Leftrightarrow \frac{c\Delta t}{\Delta x_i} \leq \frac{1}{\sqrt{2}}$$

$$\Leftrightarrow \frac{c\Delta t}{\Delta x_i} \leq \frac{1}{\sqrt{3}}$$

Can we introduce a parametric solution and some assumptions that will reduce the complexity of the problem?

[1] H. Nyquist, "Certain topics in telegraph transmission theory," Transactions of the American Institute of Electrical Engineers, vol. 47, no. 2, pp. 617–644, April 1928.

[2] T. Ajdler, L. Sbaiz, and M. Vetterli, "The plenacoustic function and its sampling," IEEE Transactions on Signal Processing, vol. 54, no. 10, pp. 3790–3804, Oct 2006.

[3] R. Mignot, G. Chardon and L. Daudet, "Low Frequency Interpolation of Room Impulse Responses Using Compressed Sensing," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 22, no. 1, pp. 205–216, Jan. 2014

[4] Lewy, H., Friedrichs, K., and Courant, R.. "Über die partiellen Differenzgleichungen der mathematischen Physik." Mathematische Annalen 100 (1928): 32–74.

# Wave equation and its parametrized solution

$$\Delta p(t, \mathbf{X}) - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} p(t, \mathbf{X}) = 0$$

$$p(t, \mathbf{X}) = \sum_{q \in \mathbb{I}} A_q \Phi_q(\mathbf{X}) g_q(t)$$

Spatial dependency:

$$\Phi_q(\mathbf{X}) \approx \sum_{r=1}^R a_{q,r} e^{j \mathbf{k}_{q,r} \cdot \mathbf{X}}$$

Temporal dependency:

$$g_q(t) = e^{j k_q c t} \quad k_q = \frac{\omega_q - j \xi_q}{c}$$

$$p(t, \mathbf{X}) = \sum_{q,r} \alpha_{q,r} e^{j(k_q c t + \mathbf{k}_{q,r} \cdot \mathbf{X})}$$

The goal: Fast and efficient **parameter learning**  
from a *small* number of microphone measurements

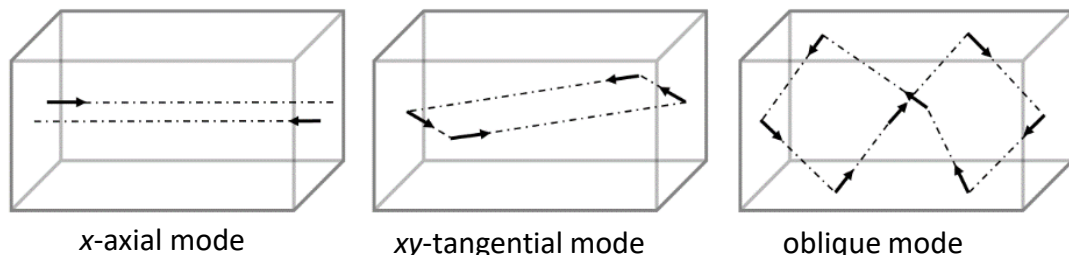
$$p(t, \mathbf{X}) = \sum_{q,r} \alpha_{q,r} e^{j((\omega_q - j \xi_q)t + \mathbf{k}_{q,r} \cdot \mathbf{X})}$$

# A set of assumptions for a well-posed problem

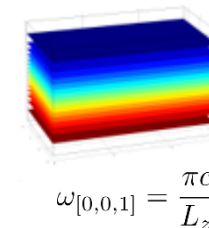
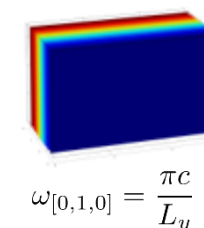
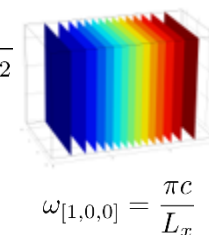
What natural **assumptions** can we introduce to reduce the complexity of the mathematical model at a low cost of approximation losses?

**Assumption #1:** simple (rectangular) room geometries  $L_x \times L_y \times L_z$

Axial modes imply room shape



$$\omega_n = \pi c \sqrt{\left(\frac{n_x}{L_x}\right)^2 + \left(\frac{n_y}{L_y}\right)^2 + \left(\frac{n_z}{L_z}\right)^2}$$



**Assumption #2:** lightly damped rooms

$$\mathbf{k}_q = [k_x \ k_y \ k_z]^T$$

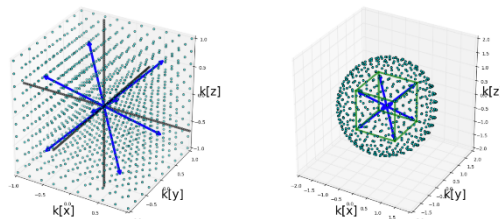
$$k_q = \frac{\omega_q - j\xi_q}{c}$$

$$\xi \ll \omega \quad |\mathbf{k}| \approx \left| \frac{\omega}{c} \right| = \sqrt{k_x^2 + k_y^2 + k_z^2} \quad \text{Spherical Search space}$$

Separability of spatial and temporal parameter estimation:

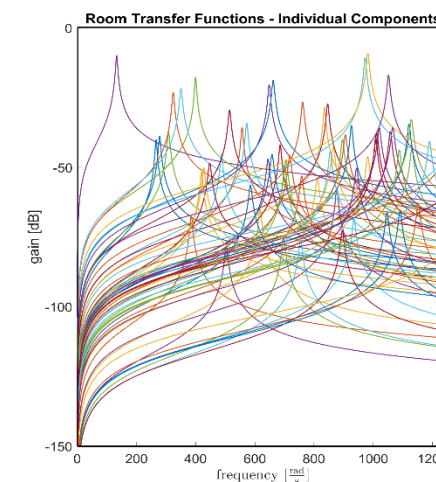
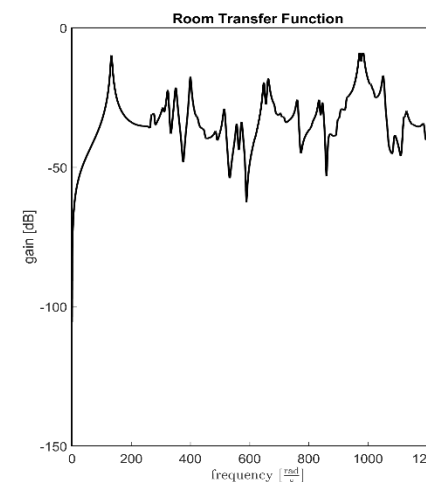
1. *temporal*: estimate  $\omega$  and build a ball  $r = \frac{\omega}{c}$

2. *spatial*: estimate the direction of  $\mathbf{k}$

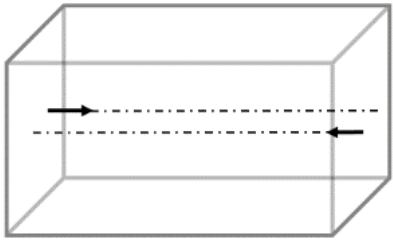
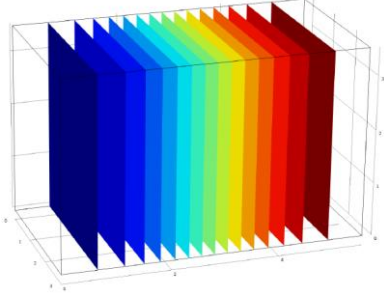
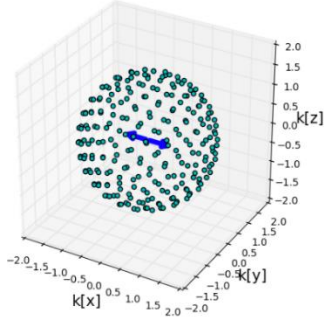
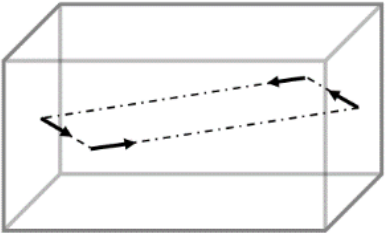
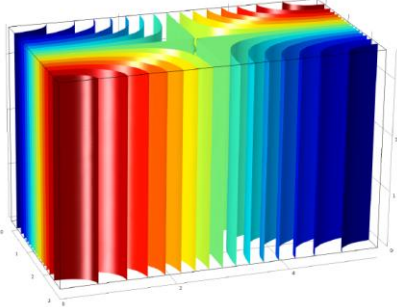
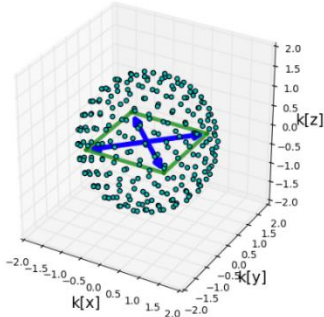
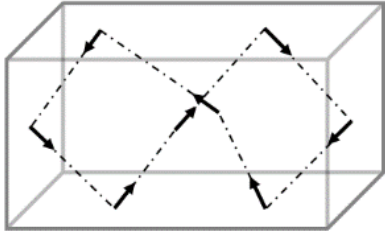
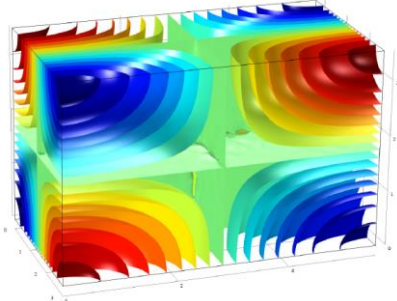
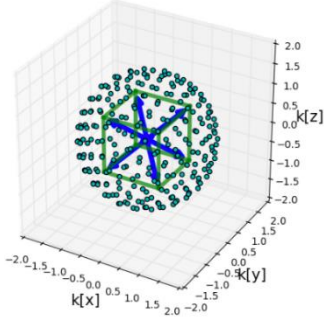


**Assumption #3:** low frequency domain analysis

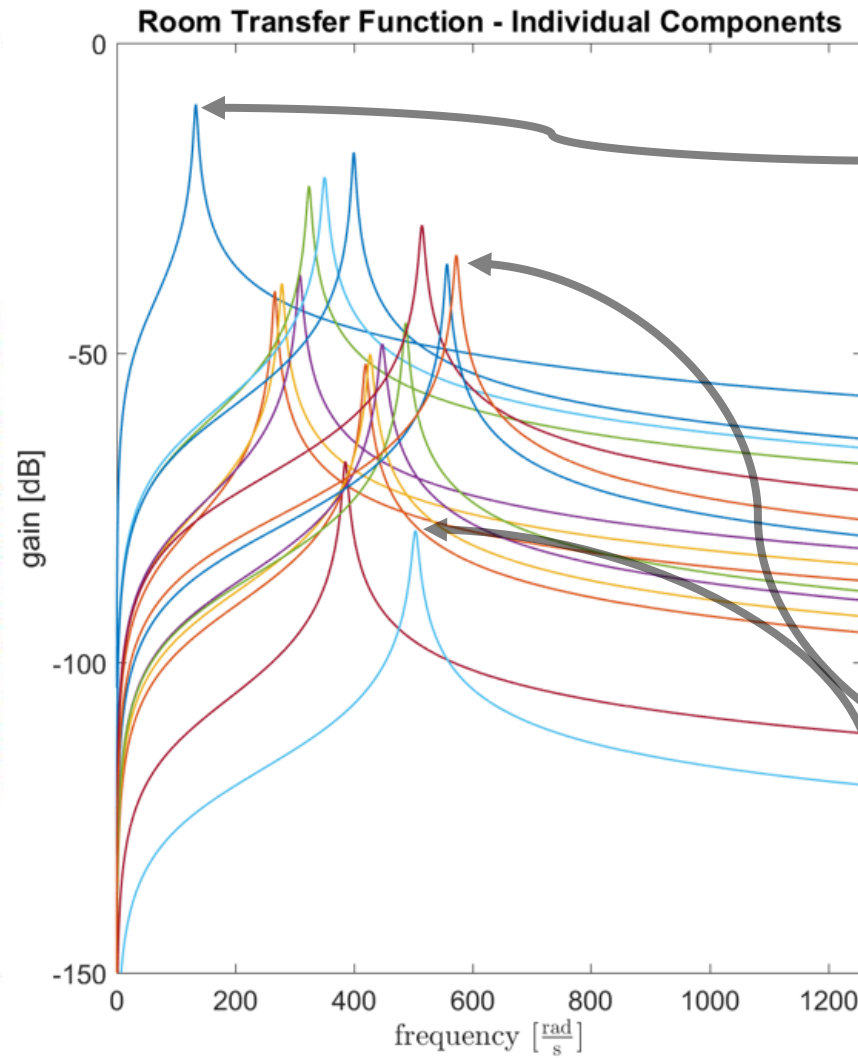
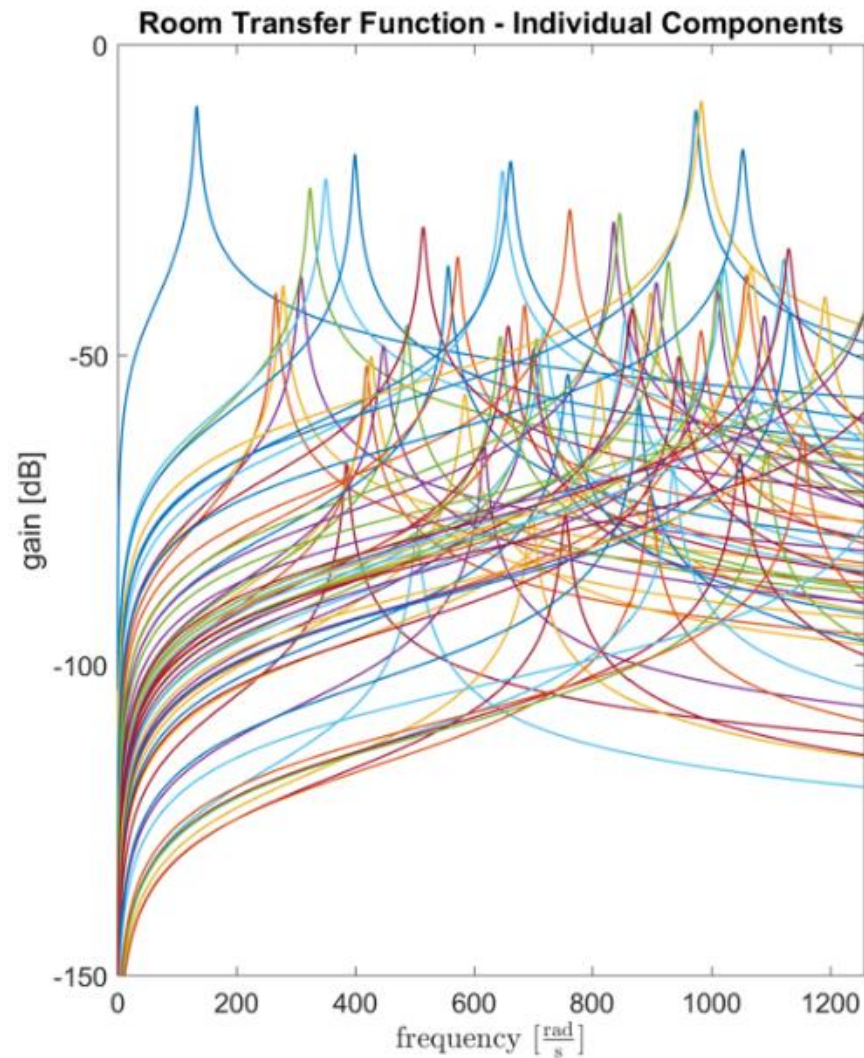
Room mode sparsity



# Three points of view on room modes

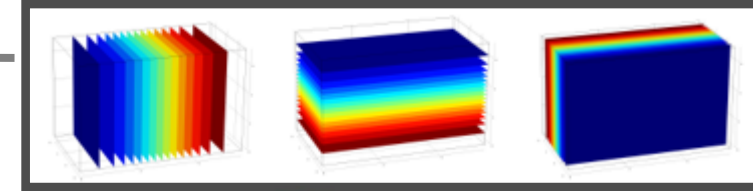
	Plane waves	Pressure isosurfaces	Wave vectors in search space $r = \frac{\omega}{c}$
x-axial mode			
xy-tangential mode			
oblique mode			

# Focusing on the low frequencies and room mode assignment

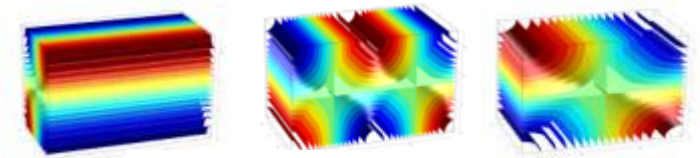


*Sound pressure isosurfaces:*

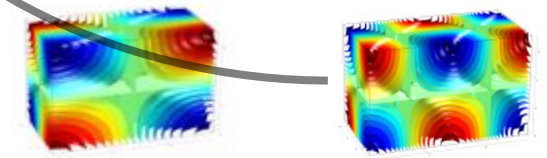
Axial modes



Tangential modes



Oblique modes

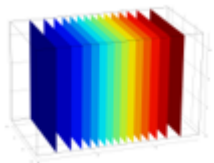


# ReSEMBLE algorithm

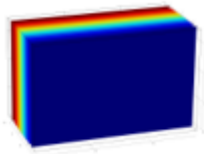
How to choose the partitioning frequency  $f_p$  ( $\omega_p$ ) ?

Based on the approximate room size:  
Such that we capture all the basic axial modes

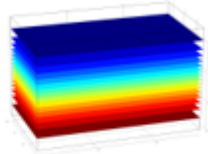
$$\omega_c = \frac{\pi c}{\tilde{L}_{\min}}$$



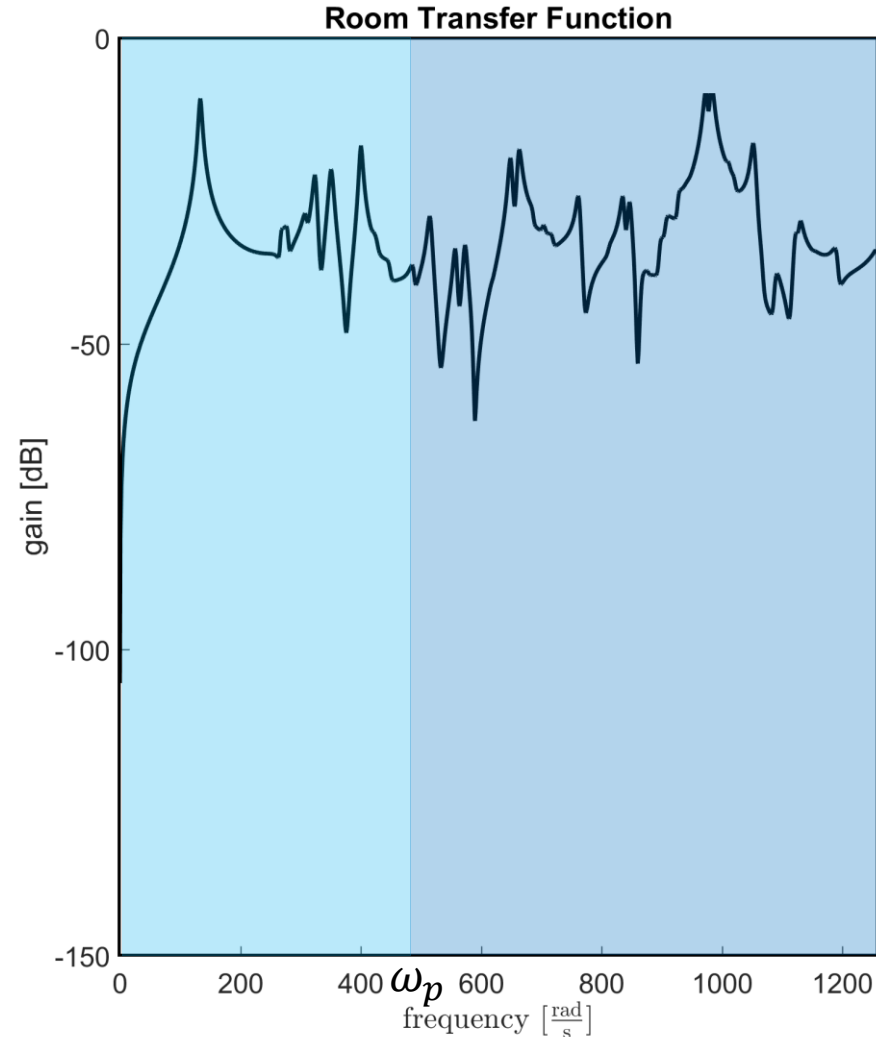
$$\omega_{[1,0,0]} = \frac{\pi c}{L_x}$$



$$\omega_{[0,1,0]} = \frac{\pi c}{L_y}$$



$$\omega_{[0,0,1]} = \frac{\pi c}{L_z}$$



procedure RESEMBLE( $\mathbf{R}, \mathbf{X}$ )

Separate the measurements with  $f_p$ :  $\mathbf{R} = \mathbf{R}^l + \mathbf{R}^h$ .

for  $i_l \in \{1, \dots, N_l\}$  do

step 1: estimate  $(\omega_{i_l}, \xi_{i_l})$  from  $\mathbf{R}_{i_l}^l$

step 2: estimate  $\mathbf{k}_{i_l}$  from  $\mathbf{r}_{i_l}^l$

step 3: compute new residual  $\mathbf{R}_{i_l+1}^l$

end for

Recover the room size  $\tilde{L}_x, \tilde{L}_y, \tilde{L}_z$  from basic axial room modes and form the regular wave vector grid.

for  $i_h \in \{N_l + 1, \dots, N\}$  do

step 1: get  $\omega_{i_h}$  and  $\mathbf{k}_{i_h}$  from the wave vector grid

step 2: estimate  $\xi_{i_h}$  from  $\mathbf{R}_{i_h}^h$

step 3: compute new residual  $\mathbf{R}_{i_h+1}^h$

end for

Estimate the expansion coefficients  $\{\alpha\}_{n=1, v=1}^{N, V}$  using least square approach.

end procedure



# ReSEMBLE algorithm

Temporal dictionary:  $\theta[i] = e^{\xi_n[i]t} e^{j\omega_n[i]t}$   
 Spatio-temporal dictionary:  $\Sigma[:, i] = e^{\xi_{i_l}t} e^{j\omega_{i_l}t} e^{\mathbf{X} \cdot \mathbf{k}[i]}$

procedure RESEMBLE( $\mathbf{R}, \mathbf{X}$ )

Separate the measurements with  $f_p$ :  $\mathbf{R} = \mathbf{R}^l + \mathbf{R}^h$ .

```

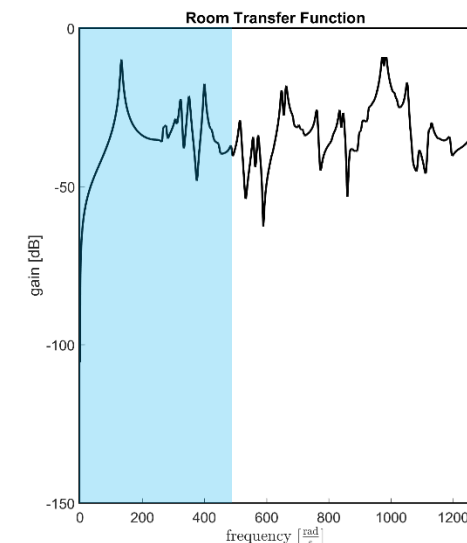
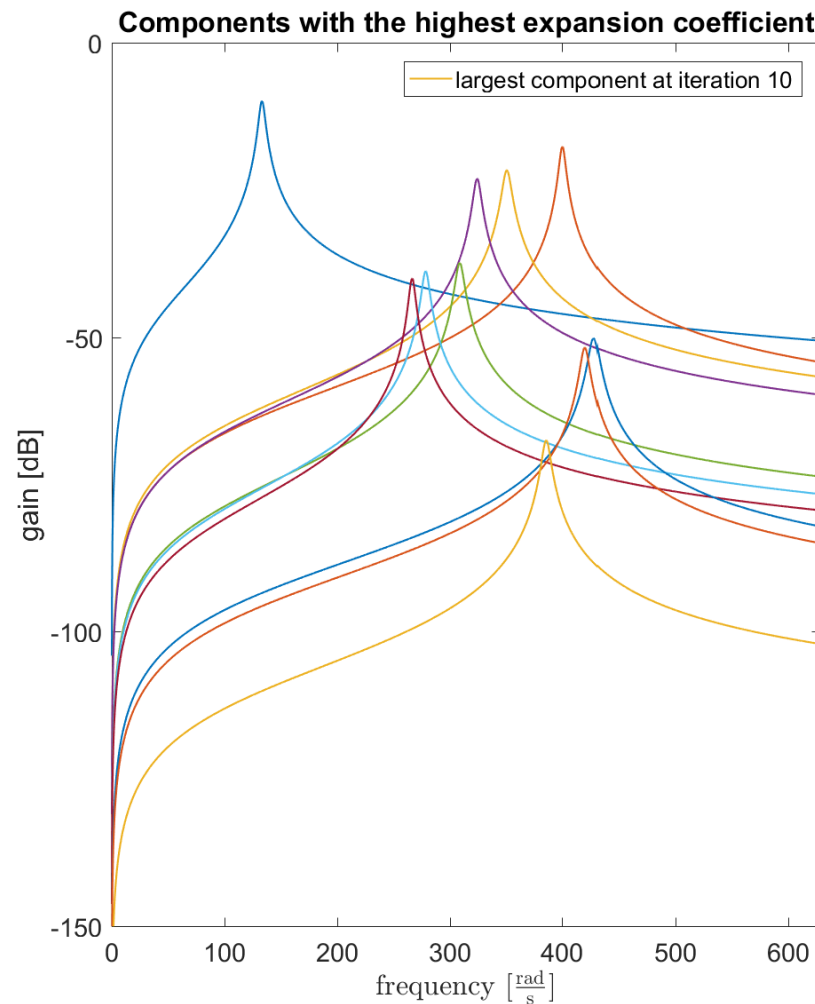
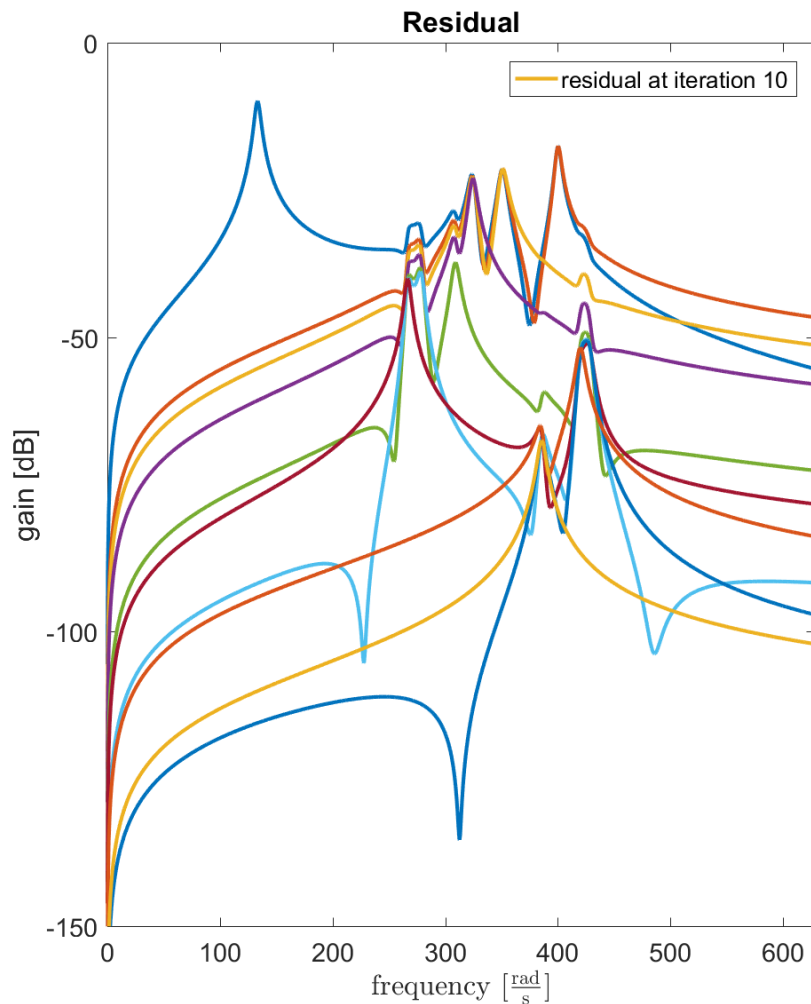
for  $i_l \in \{1, \dots, N_l\}$  do
    step 1: estimate  $(\omega_{i_l}, \xi_{i_l})$  from  $\mathbf{R}_{i_l}^l$ 
    step 2: estimate  $\mathbf{k}_{i_l}$  from  $\mathbf{r}_{i_l}^l$ 
    step 3: compute new residual  $\mathbf{R}_{i_l+1}^l$ 
end for
    
```

Recover the room size  $\tilde{L}_x, \tilde{L}_y, \tilde{L}_z$  from basic axial room modes and form the regular wave vector grid.

```

for  $i_h \in \{N_l + 1, \dots, N\}$  do
    step 1: get  $\omega_{i_h}$  and  $\mathbf{k}_{i_h}$  from the wave vector grid
    step 2: estimate  $\xi_{i_h}$  from  $\mathbf{R}_{i_h}^h$ 
    step 3: compute new residual  $\mathbf{R}_{i_h+1}^h$ 
end for
    
```

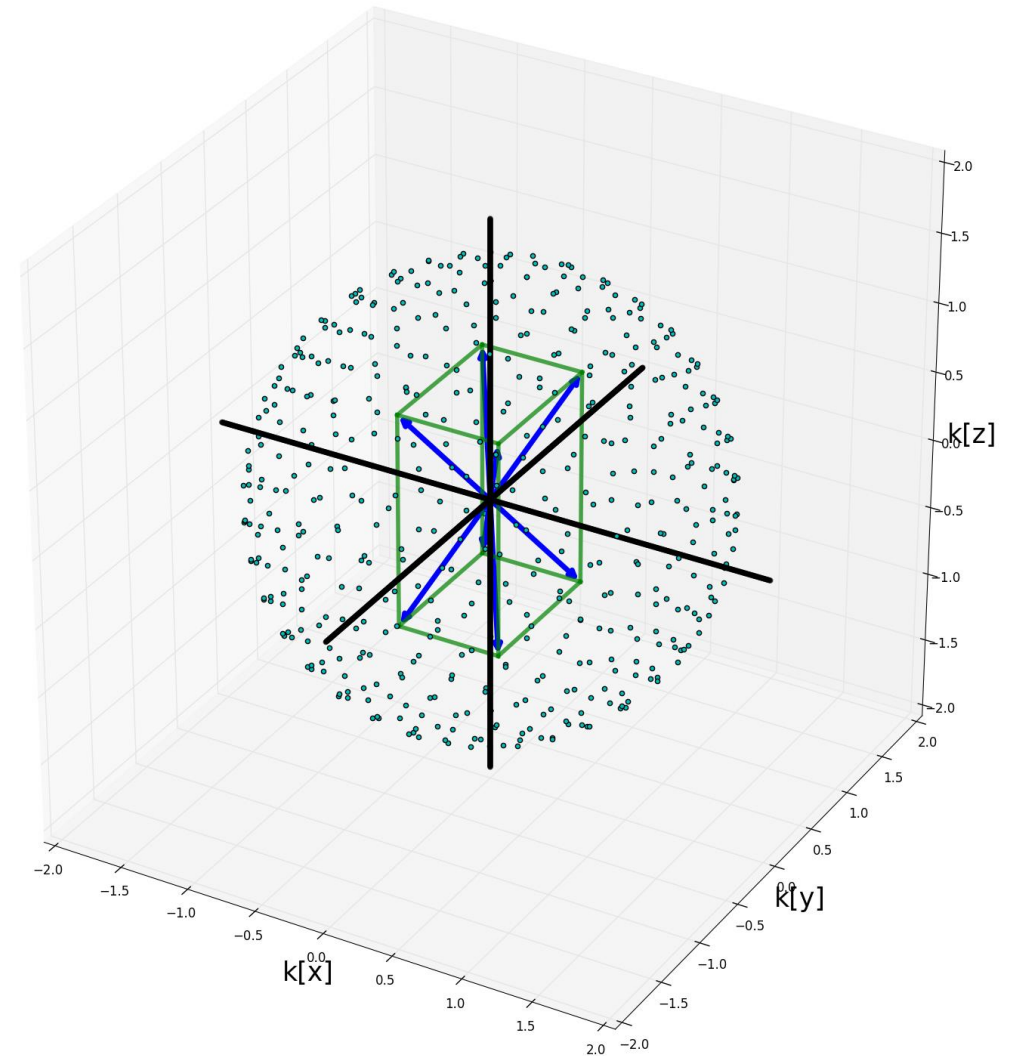
Estimate the expansion coefficients  $\{\alpha\}_{n=1, v=1}^{N, V}$  using least square approach.  
 end procedure



# Structured group sparsity

Separability of spatial and temporal parameter estimation:

1. *temporal*: estimate  $\omega$  and build a ball  $r = \frac{\omega}{c}$
2. *spatial*: estimate the direction of  $\mathbf{k}$



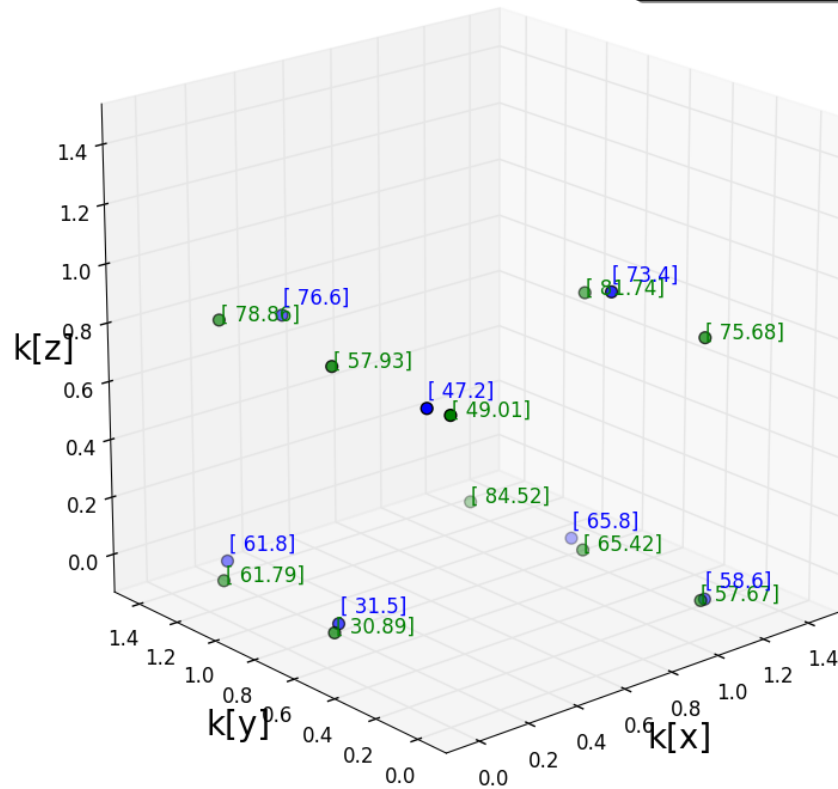
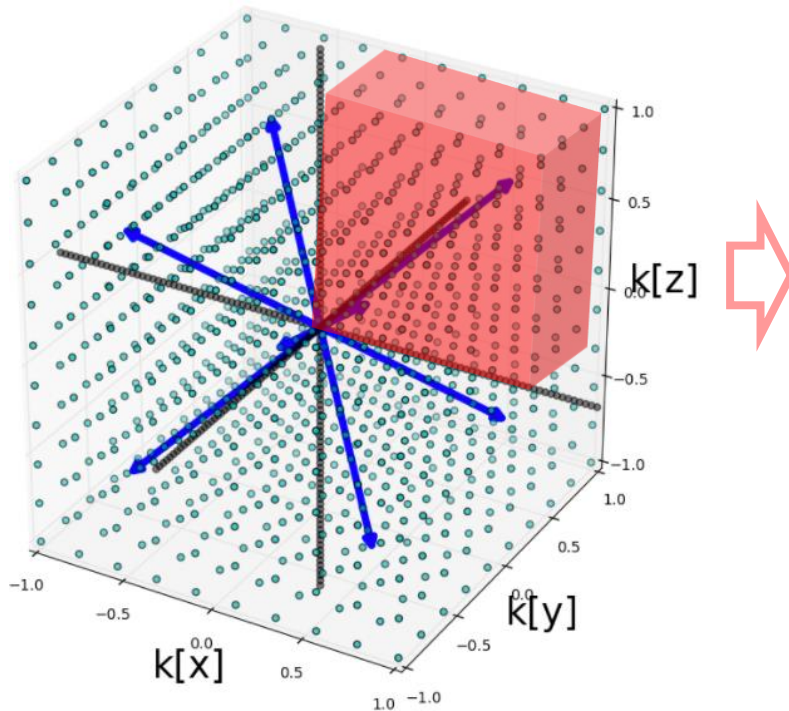
# k-space estimation results

Why the position of points deviates the most over z-axis?  
 This is why the x-, y- and xy-modes are precise and z-, xz-, xy- and xyz-modes are far off from the ground truth

$$|k| \approx \left| \frac{\omega}{c} \right|$$

$$k_{iq} = \frac{\omega_q - j\xi_q}{c} \quad \xi \ll \omega$$

- undamped case
- estimation



```

procedure RESEMBLE(R, X)
  Separate the measurements with  $f_p$ :  $\mathbf{R} = \mathbf{R}^l + \mathbf{R}^h$ .
  for  $i_l \in \{1, \dots, N_l\}$  do
    step 1: estimate  $(\omega_{i_l}, \xi_{i_l})$  from  $\mathbf{R}_{i_l}^l$ 
    step 2: estimate  $\mathbf{k}_{i_l}$  from  $\mathbf{r}_{i_l}^l$ 
    step 3: compute new residual  $\mathbf{R}_{i_l+1}^l$ 
  end for
  
```

```

  Recover the room size  $\tilde{L}_x, \tilde{L}_y, \tilde{L}_z$  from basic axial room modes and form the regular wave vector grid.
  for  $i_h \in \{N_l + 1, \dots, N\}$  do
    step 1: get  $\omega_{i_h}$  and  $\mathbf{k}_{i_h}$  from the wave vector grid
    step 2: estimate  $\xi_{i_h}$  from  $\mathbf{R}_{i_h}^h$ 
    step 3: compute new residual  $\mathbf{R}_{i_h+1}^h$ 
  end for
  Estimate the expansion coefficients  $\{\alpha\}_{n=1, v=1}^{N, V}$  using least square approach.
end procedure
  
```

$$\tilde{L}_x, \tilde{L}_y, \tilde{L}_z$$

# ReSEMBLE algorithm

Imply the wave vectors from the reconstructed room shape:

$$(n_x, n_y, n_z) \in \mathcal{N}_0^3 \setminus (0, 0, 0)$$

$$k_x = n_x \frac{\pi}{\tilde{L}_x} \quad k_y = n_y \frac{\pi}{\tilde{L}_y} \quad k_z = n_z \frac{\pi}{\tilde{L}_z}$$

After applying the *high* part of the algorithm, the Pearson correlation coefficient that the approximation is good (e.g. 82% for only 19-microphone setting and  $f_c = 200\text{Hz}$ ), but it should be further improved once the deviation of the wave vectors is efficiently characterized

procedure RESEMBLE( $\mathbf{R}, \mathbf{X}$ )

Separate the measurements with  $f_p$ :  $\mathbf{R} = \mathbf{R}^l + \mathbf{R}^h$ .

for  $i_l \in \{1, \dots, N_l\}$  do

step 1: estimate  $(\omega_{i_l}, \xi_{i_l})$  from  $\mathbf{R}_{i_l}^l$

step 2: estimate  $\mathbf{k}_{i_l}$  from  $\mathbf{r}_{i_l}^l$

step 3: compute new residual  $\mathbf{R}_{i_l+1}^l$

end for

Recover the room size  $\tilde{L}_x, \tilde{L}_y, \tilde{L}_z$  from basic axial room modes and form the regular wave vector grid.

for  $i_h \in \{N_l + 1, \dots, N\}$  do

step 1: get  $\omega_{i_h}$  and  $\mathbf{k}_{i_h}$  from the wave vector grid

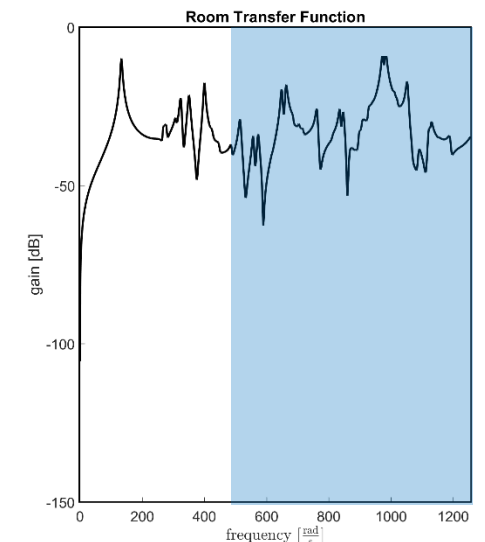
step 2: estimate  $\xi_{i_h}$  from  $\mathbf{R}_{i_h}^h$

step 3: compute new residual  $\mathbf{R}_{i_h+1}^h$

end for

Estimate the expansion coefficients  $\{\alpha\}_{n=1, v=1}^{N, V}$  using least square approach.

end procedure



# In the spirit of *open research* and *acoustic data augmentation*

[1]

7 commits | 1 branch | 0 releases | 1 contributor

Search master | New pull request | Create new file | Upload files | Find file | Close or download

Files Add files via upload | Latest commit 23 Oct '18 28 days ago

- manuscript Add files via upload 28 days ago
- matlab\_code Add files via upload 29 days ago
- room\_time\_images Add files via upload a month ago
- README.mxd Update README.mxd 29 days ago

README

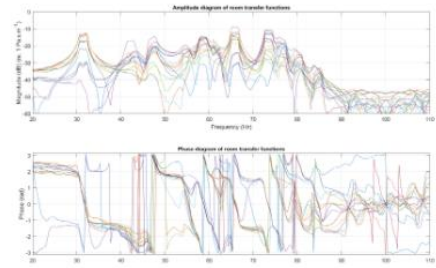
## JOINT ESTIMATION OF ROOM GEOMETRY AND MODES

Here you can find a framework for retrieving room shape and modes from microphone measurements. We use a combination of the following techniques:

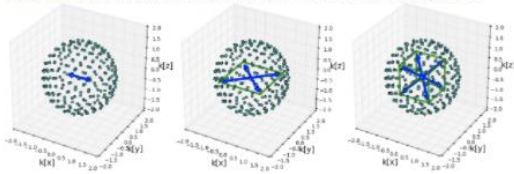
1. curve fitting based on rational fraction polynomial method
2. sampling on the sphere close to uniform
3. matching pursuit
4. group sparsity
5. compressed sensing

The solution is limited to rectangular rooms that is lightly damped.

The measurements and the room details are available here: <https://zenodo.org/record/1169161#.WqFicQyW-W>  
 We have used the **Rational Fraction Polynomial Method** by Cristian Gutierrez Acuna  
<https://ch.mathworks.com/matlabcentral/fileexchange/31805-rational-fraction-polynomial-method>  
 focused=5049557&tab=function for curve fitting of our room transfer functions.



and Suite of functions to perform uniform sampling of a sphere by Anton Semchko  
<https://ch.mathworks.com/matlabcentral/fileexchange/37004-suite-of-functions-to-perform-uniform-sampling-of-a-sphere>  
 s.sid=prof.contriblink in order to discover the type of the room modes that corresponds to our resonant frequencies:



Here we provide the matlab version of the code and python version is available upon request: [helena.peictukuljac@epfl.ch](mailto:helena.peictukuljac@epfl.ch).

[1]

February 8, 2018

Dataset Open Access

## Room Impulse Response measurements of a rectangular room


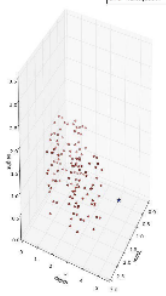
Pham Vu Thach; Peic Tukuljac Helena

This archive contains the data for a Gitlab hosted project ([https://github.com/epfl-lts2/joint\\_estimation\\_of\\_room\\_geometry\\_and\\_modes](https://github.com/epfl-lts2/joint_estimation_of_room_geometry_and_modes)). This archive allows users to extract the Room impulse responses (RIRs) measurements for a real rectangular room. A total of 132 measurements are included. Furthermore, users have the option to perform several post processing steps on the RIRs such as filtering, downsampling and truncation in time domain. Please refer to the guidelines pdf for more information.

Preview

Page: 1 of 2 Automatic Zoom

pass filtering, downsampling and truncation in time domain.

The room used for measurements (top) and the positions of the microphones and speaker for the measurements (right)

About the measurements :

- Measurements are done in 12 groups, each group contains 11 microphones placed at different positions. This gives a total of 132 microphone measurements.
- The original data file is `TF_data.mat` which contains 132 transfer functions corresponding to

[2]

Indexed in

OpenAIRE

Publication date:  
February 8, 2018

DOI:  
DOI 10.5281/zenodo.1169161

Keyword(s):

Room Impulse Response Compressive sensing ICASSP  
rectangular room room shape

Meeting:

ICASSP 2018, Calgary, Alberta, Canada, 15–20 April 2018

License (for files):

Creative Commons Attribution 4.0

Versions

Version 1.0 10.5281/zenodo.1169161 Feb 8, 2018

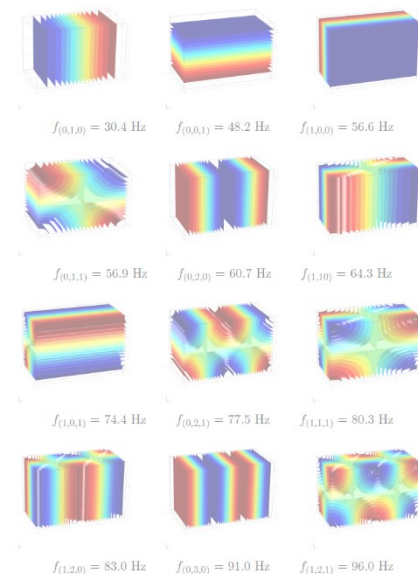
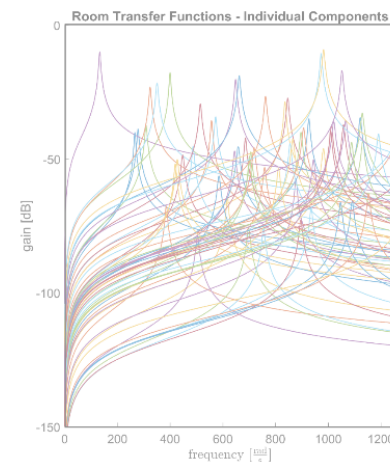
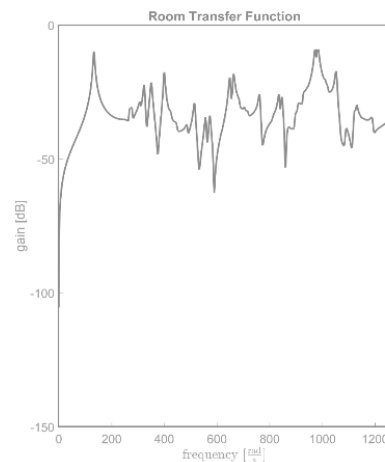
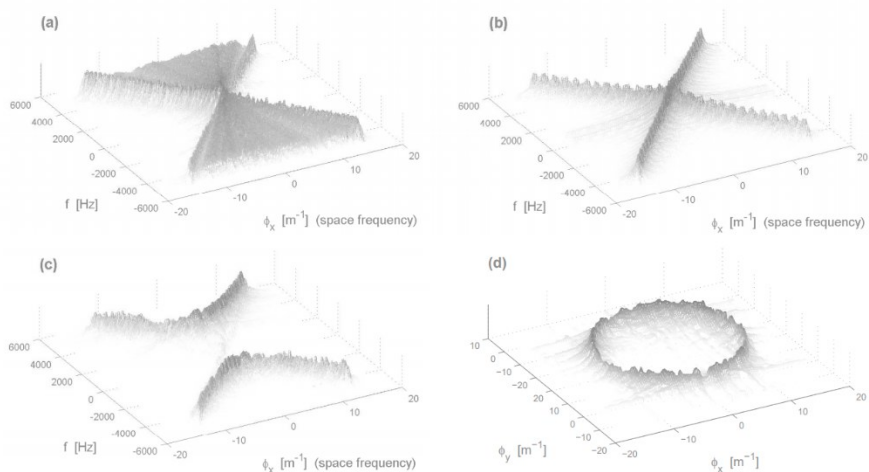
Cite all versions? You can cite all versions by using the DOI 10.5281/zenodo.1169160. This DOI represents all versions, and will always resolve to the latest one. [Read more.](#)

# Content Atlas

## Spatio-temporal sampling relations [1]

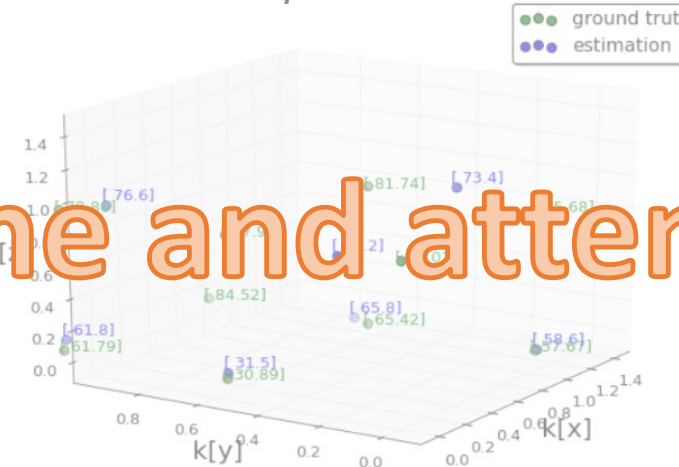
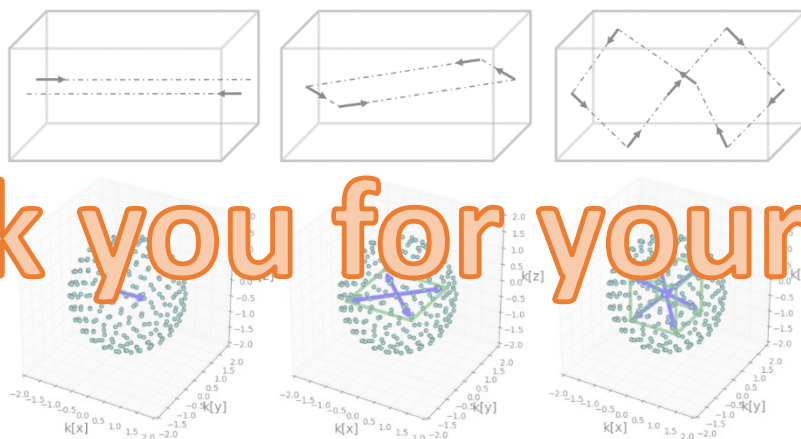
## Room transfer function [2]

## Room modes [3]



## Plane waves and wave vectors

## Acoustical k-space



Thank you for your time and attention!

Introduction

Contribution