

# Sampling plen-acoustic fields

---

Laurent DAUDET



Institut Langevin

Université Paris Diderot

[laurent.daudet@espci.fr](mailto:laurent.daudet@espci.fr)

Institut **Langevin**  
ONDES ET IMAGES

université  
**PARIS DIDEROT**  
PARIS 7



# Acknowledgements

---

- Joint work with  
**Rémi Mignot,**  
**Gilles Chardon,**  
Antoine Peillot,  
François Ollivier,  
Rémi Gribonval,  
Nancy Bertin,  
Albert Cohen....

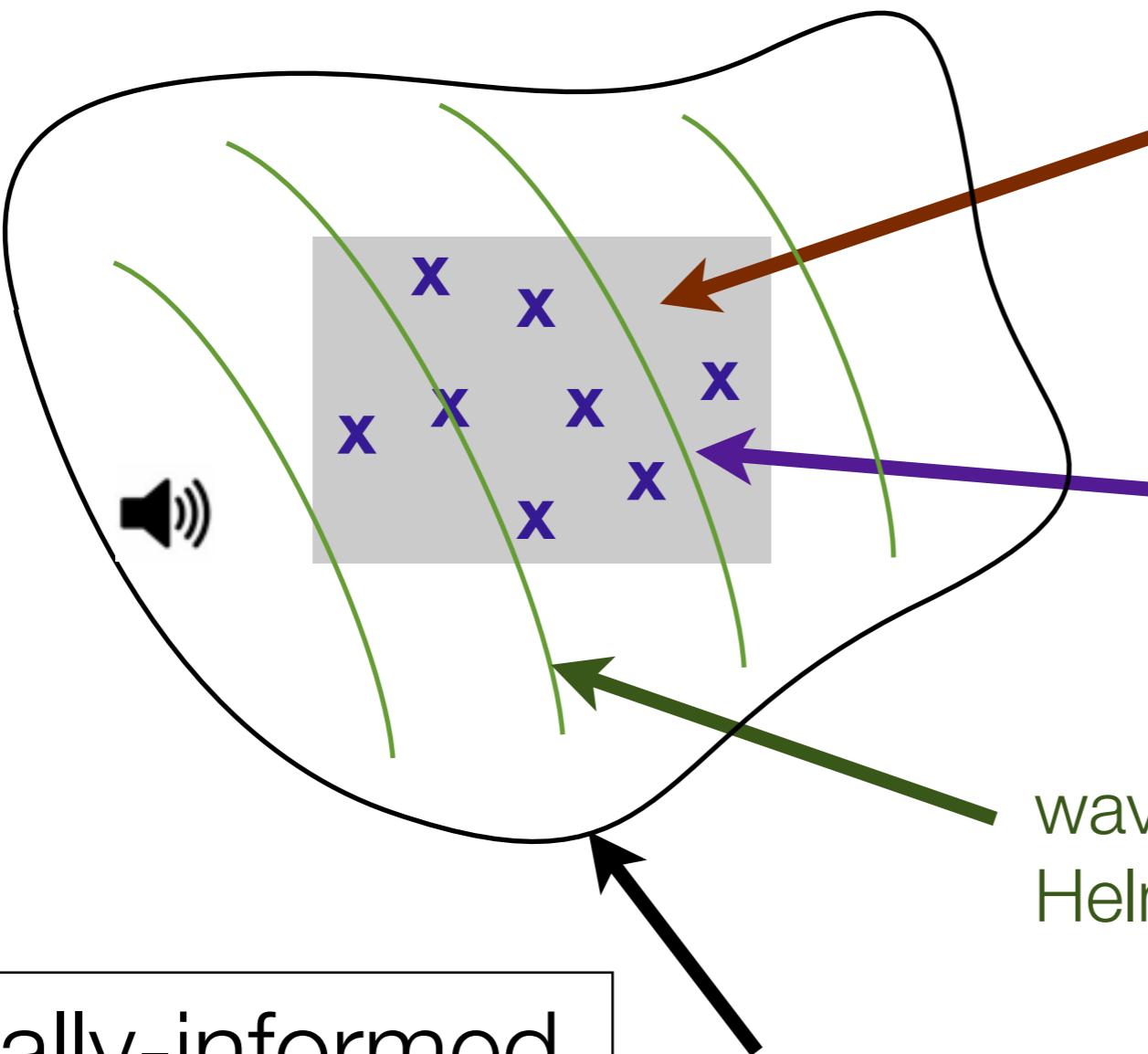


*in the framework of the ECHANGE project (2009-2012),  
funded by the French Agence Nationale de la Recherche*



<http://echange.inria.fr/>

# Acoustic fields and sampling : general framework



Physically-informed  
sampling-and-  
interpolation problem

We want to measure a set of Impulse Responses over an area of interest

Measurements are punctual (spatial samples)

waves satisfy the Helmholtz equation

boundary conditions are unknown

# Acoustic fields and sampling : general framework

---

## Outline of the presentation

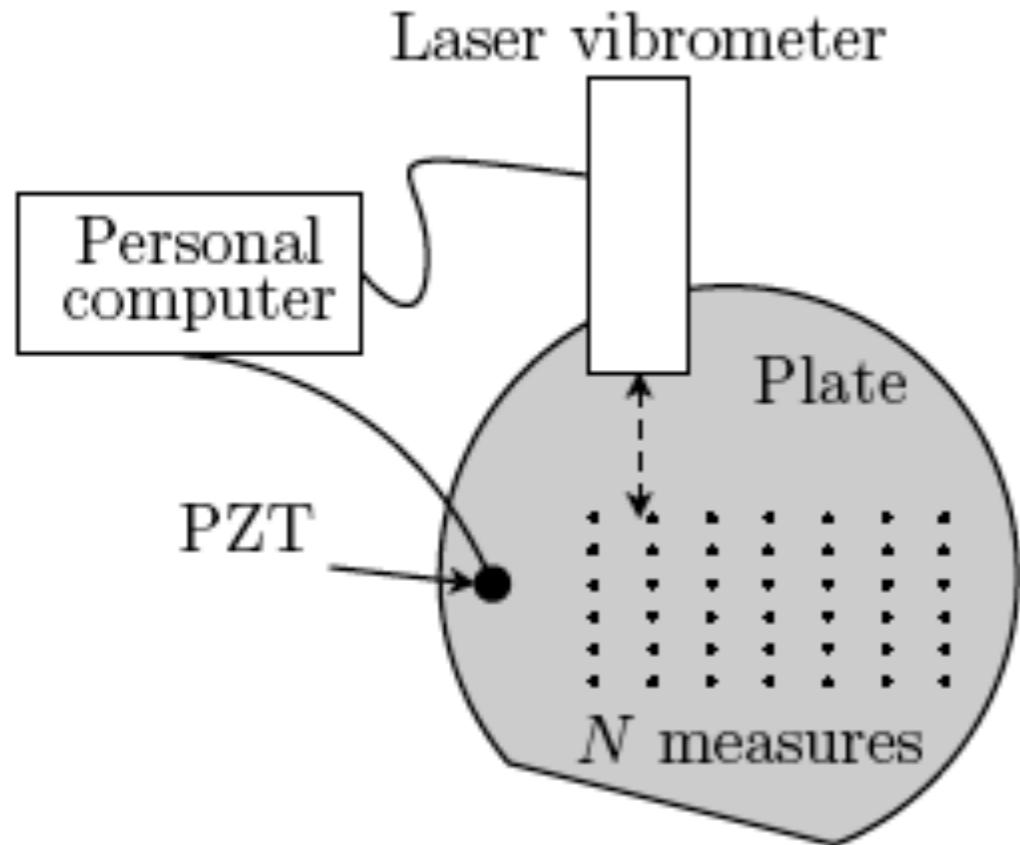
- the 2D
- the 3D

# Interpolating plates impulse responses

---

(with G. Chardon and A. Leblanc)

# Plate vibrations measurements



D-shaped plate  
known for chaotic  
behavior

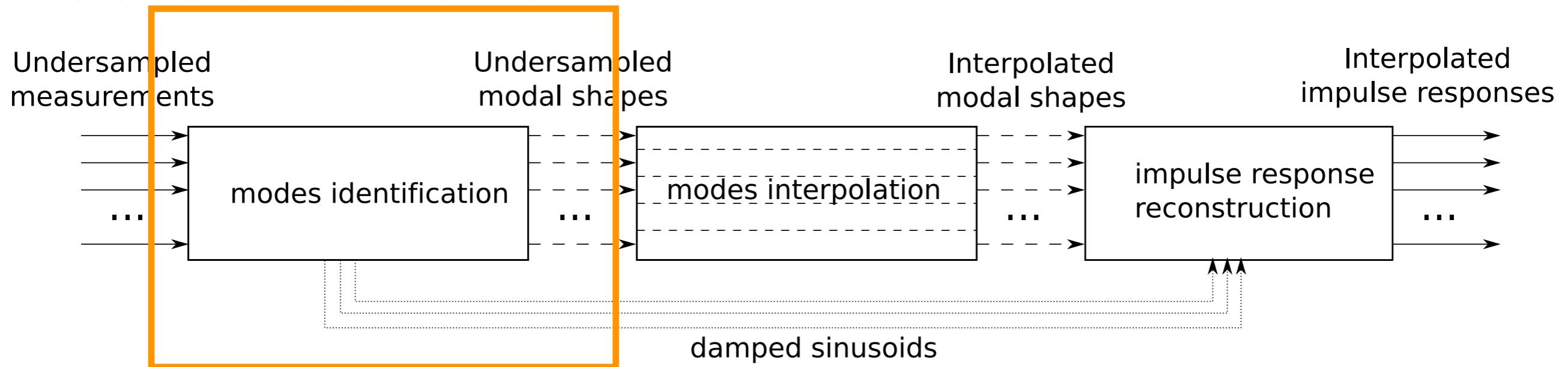
unknown dispersion relation

How many point measurements are necessary  
to recover the full vibration behavior ?

(fine sampling with laser vibrometer : 2 hrs)

# Recovering the impulse responses

## General framework



identify frequencies and damping jointly present in all measurements

Ex: SOMP

$$\operatorname{argmax} C_m(a, \omega, \phi) = \sum_k |\langle \hat{c}_{a,\omega,\phi}, u_{k,m} \rangle|^2$$

$$c_{a,\omega,\phi}(t) = e^{-at} \cos(\omega t + \phi)$$

$$u_{k,m+1} = u_{k,m} - P_m(u_{k,m})$$

# Plane-wave approximation

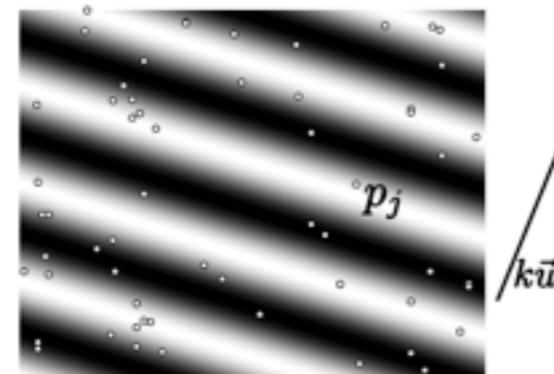
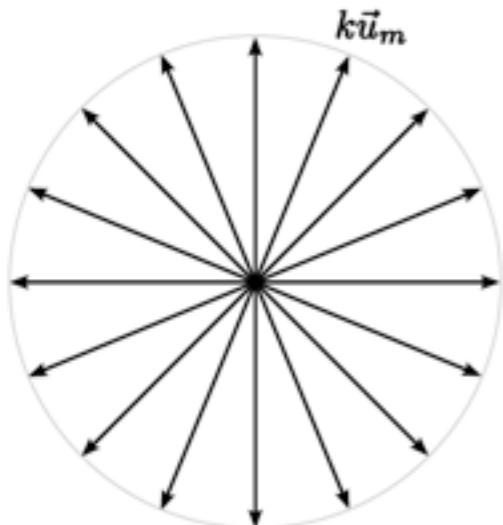
## Kirchhoff-Love model

$$\rho h \frac{\partial^2 w}{\partial t^2} + \alpha \frac{\partial w}{\partial t} + D \Delta^2 w = 0$$

Propagative modes

Evanescence modes

$$\rightarrow \mathbf{w}(x, y) \approx \left( \sum_n \alpha_n e^{i \vec{k}_n \cdot \vec{x}} + \beta_n e^{\vec{k}_n \cdot \vec{x}} \right) \mathbf{1}_{\mathcal{S}}(x, y)$$



$$||\vec{k}_n|| = k_0$$

## Global model

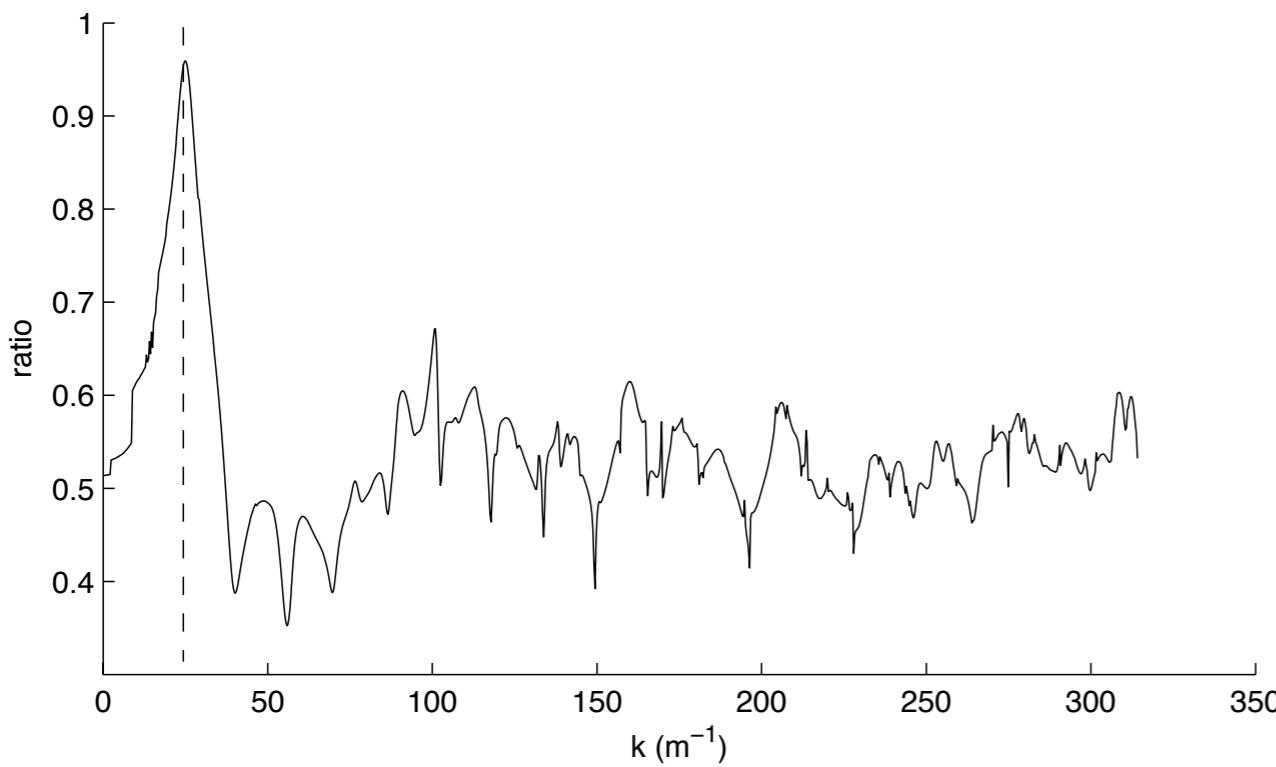
$$w(x, y, t) \simeq \sum_{n=0}^N A_n e^{-a_n t} \cos(\omega_n t + \phi_n) \left( \sum_{m=1}^M w_{nm}^- e^{i \vec{k}_{nm} \cdot \vec{x}} + \sum_{m=1}^M w_{nm}^+ e^{\vec{k}_{nm} \cdot \vec{x}} \right)$$

# Dispersion relation

---

$$\mathbf{w}(x, y) \approx \left( \sum_n \alpha_n e^{i\vec{k}_n \cdot \vec{x}} + \beta_n e^{\vec{k}_n \cdot \vec{x}} \right) \mathbf{1}_{\mathcal{S}}(x, y)$$
$$\|\vec{k}_n\| = k_0$$

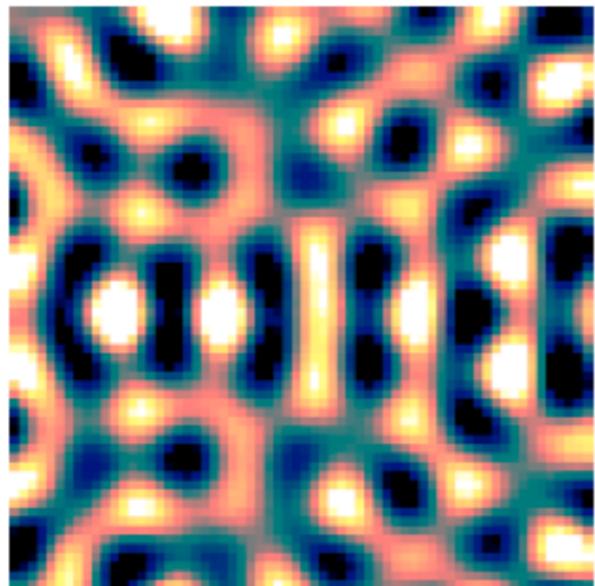
but  $k_0$  is unknown ...



Find best match ( $k_0$  s.t. measures are maximally projected on the set of plane waves on the circle of radius  $k_0$ )  
**“circle-MUSIC”**

# Recovering the modal shapes

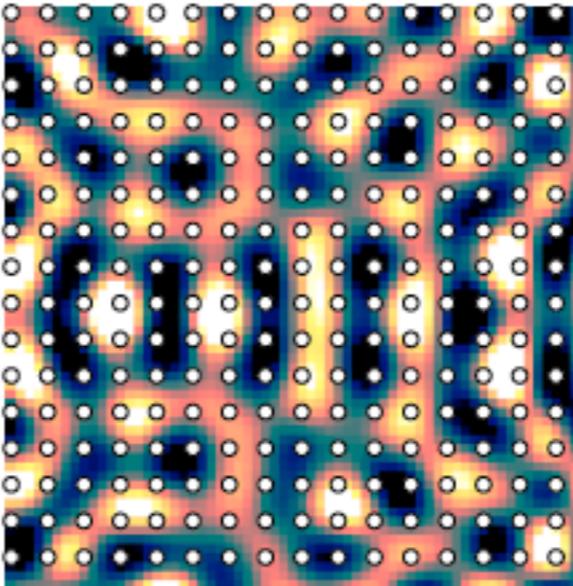
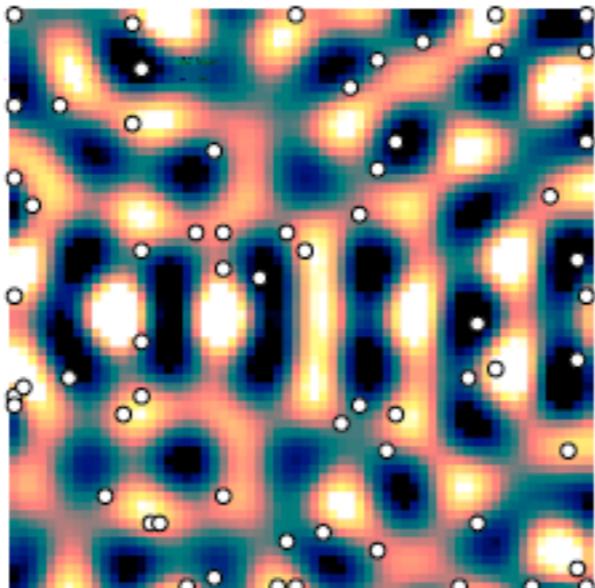
Fourier  
interp.



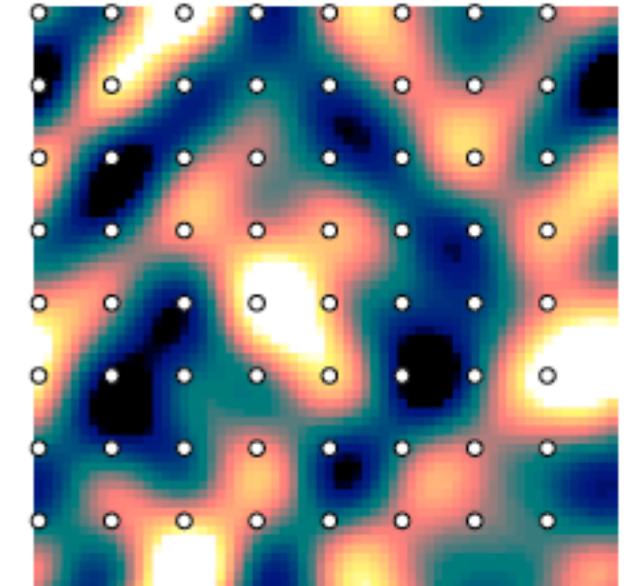
Reference

Random meas. pts

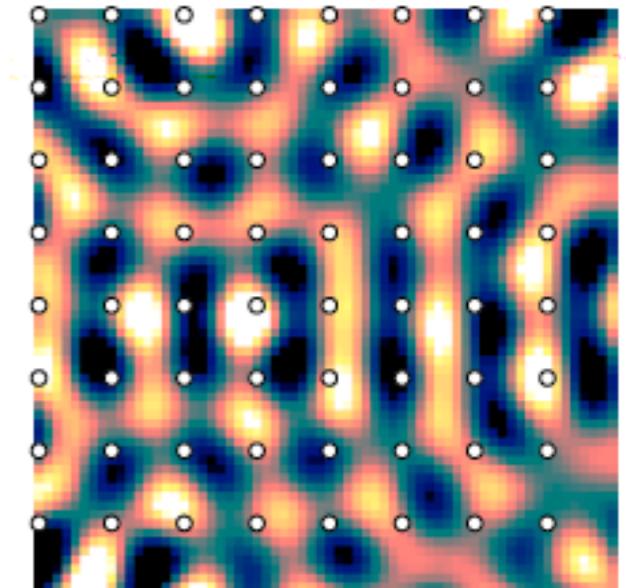
Structured  
sparsity  
interp.



Fine grid



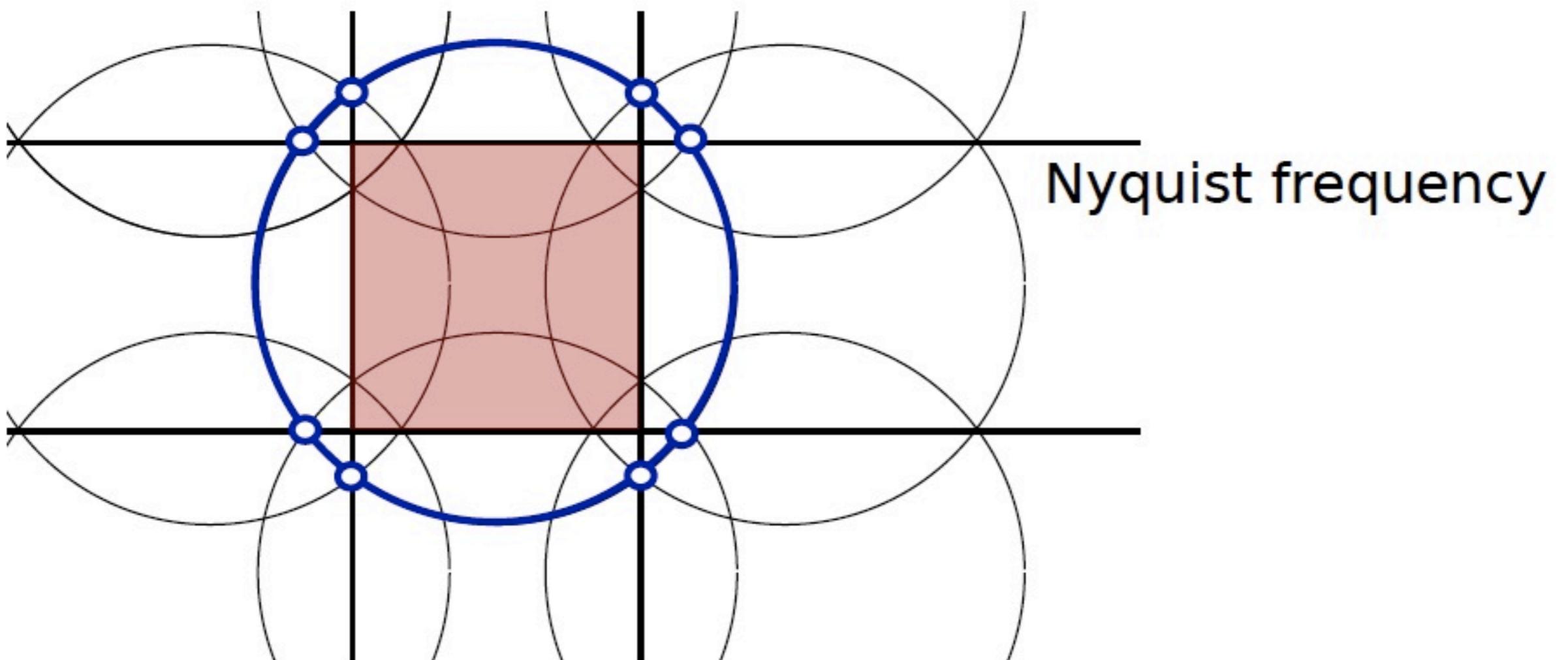
Coarse grid



**Structured sparsity  
disambiguates aliasing**

# Recovering the modal shapes

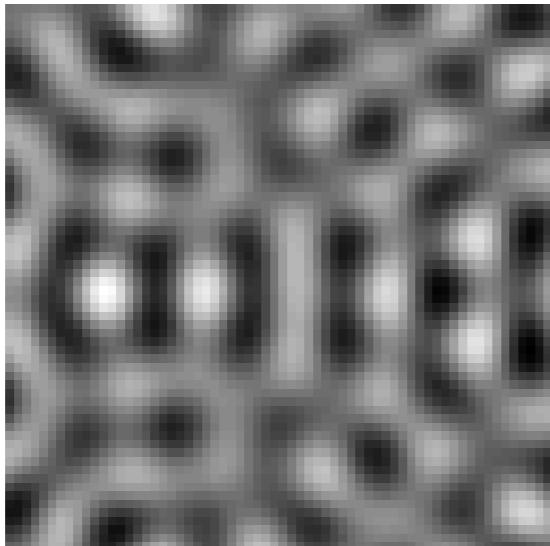
---



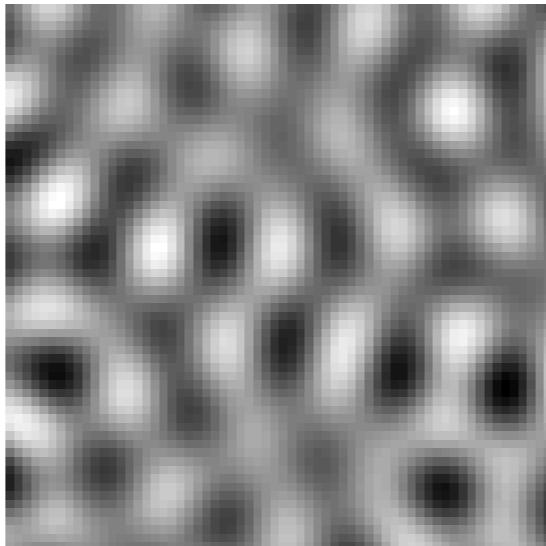
# Recovering the modal shapes

---

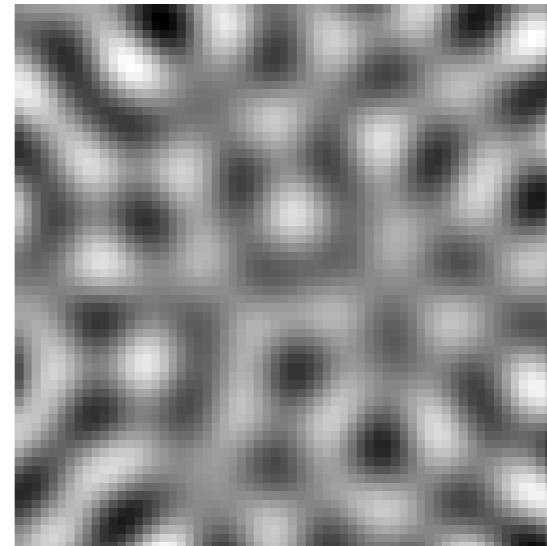
20453 Hz



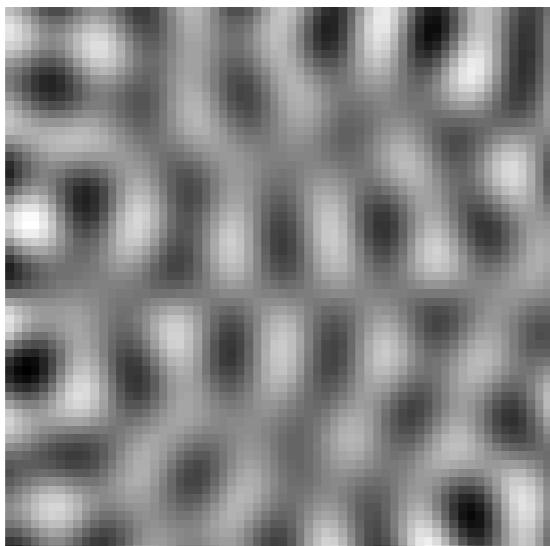
15688 Hz



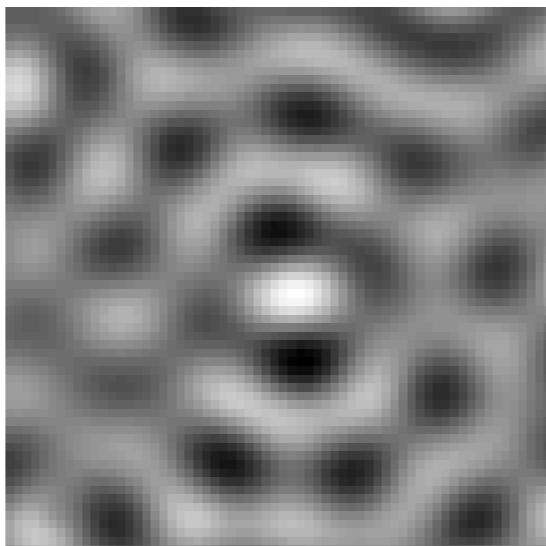
18951 Hz



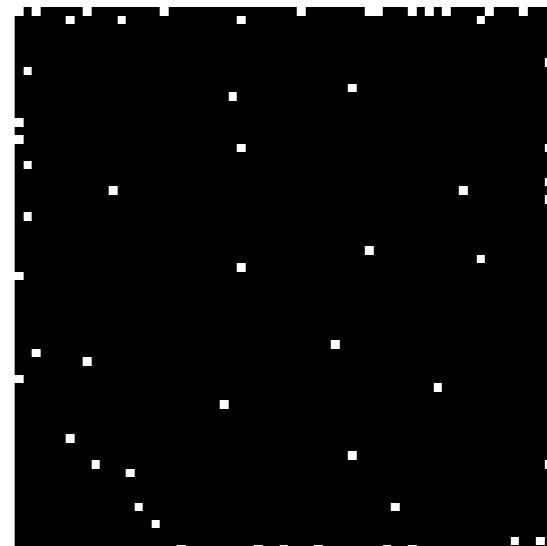
19467 Hz



13597 Hz

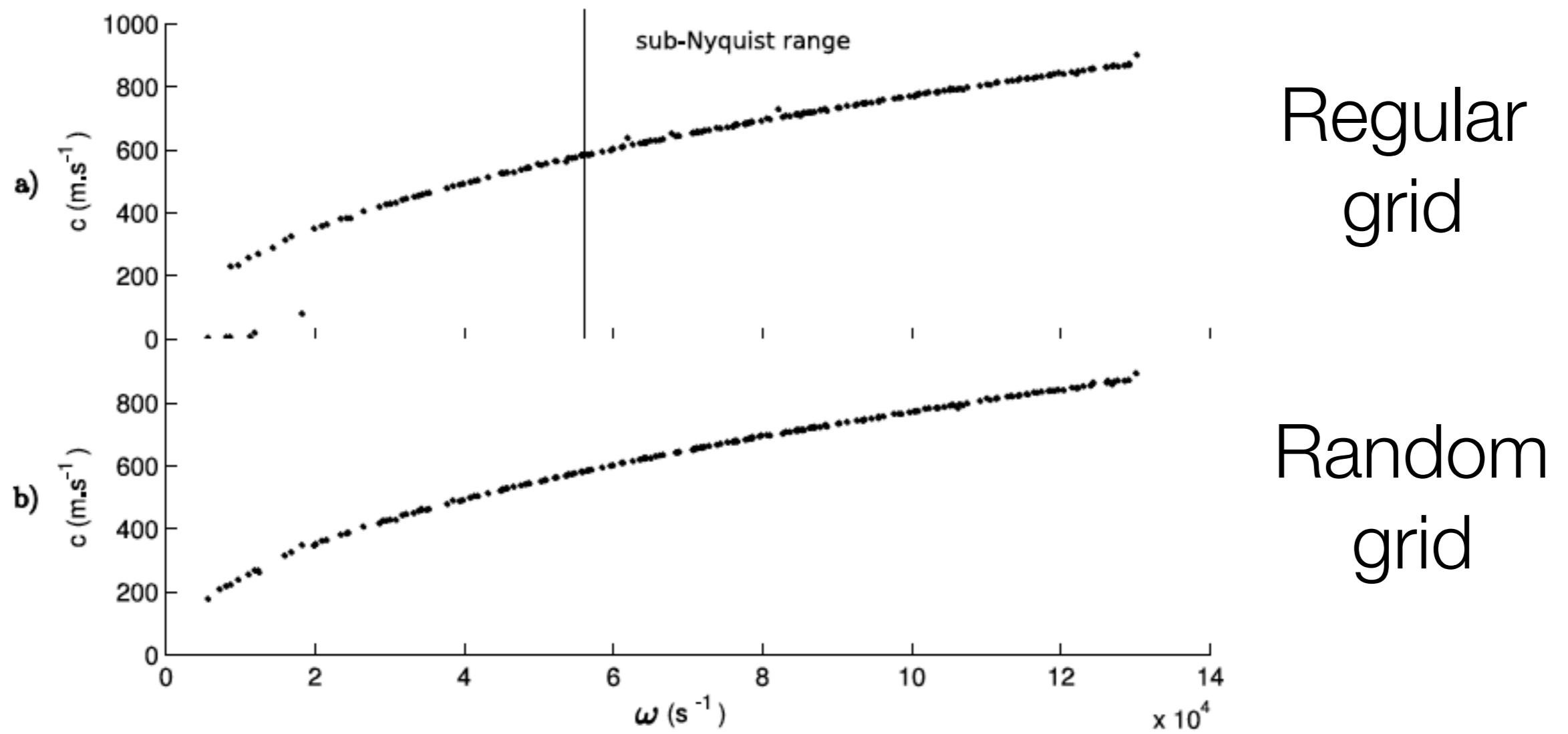


Sampling pattern

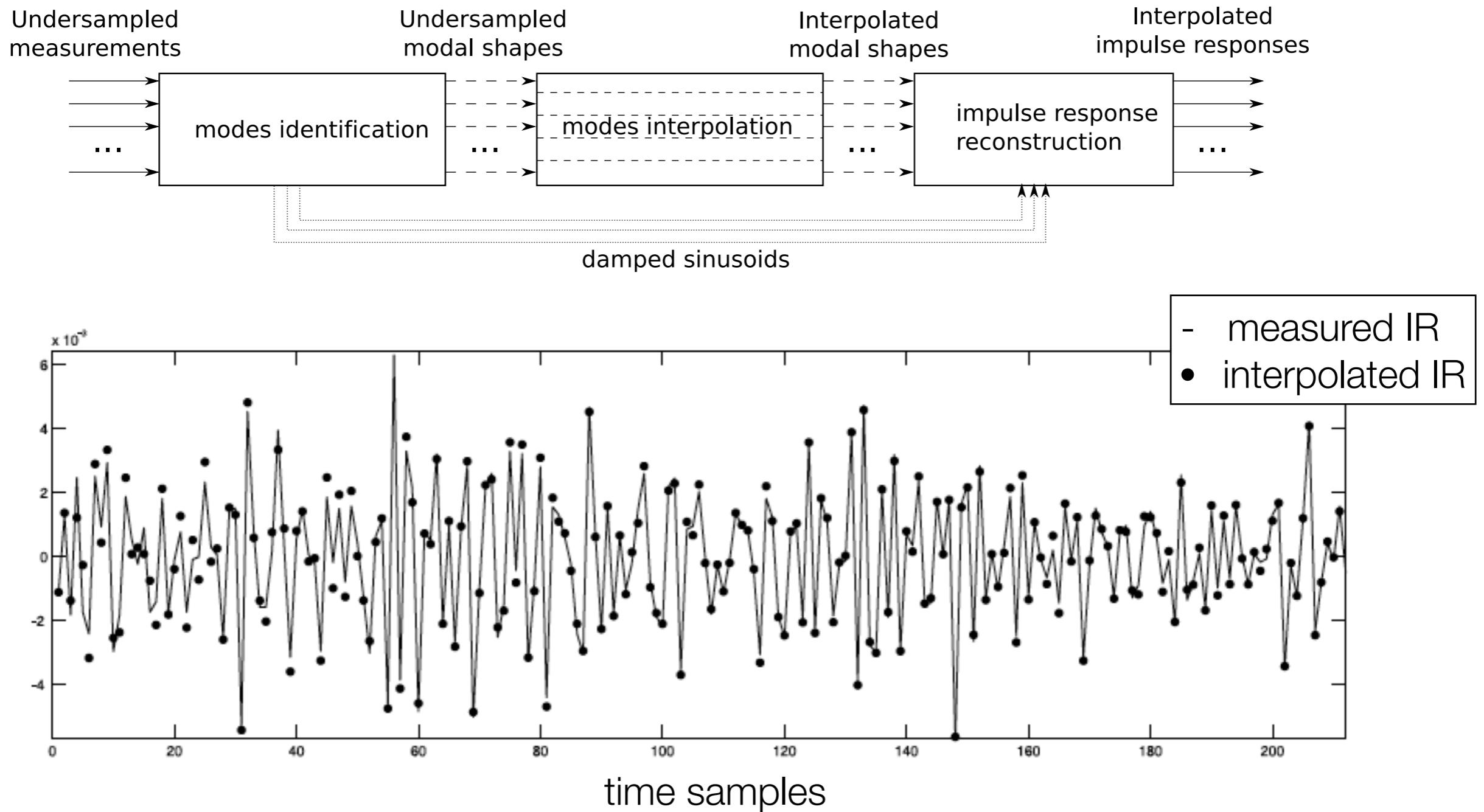


# Recovering the dispersion relation

---



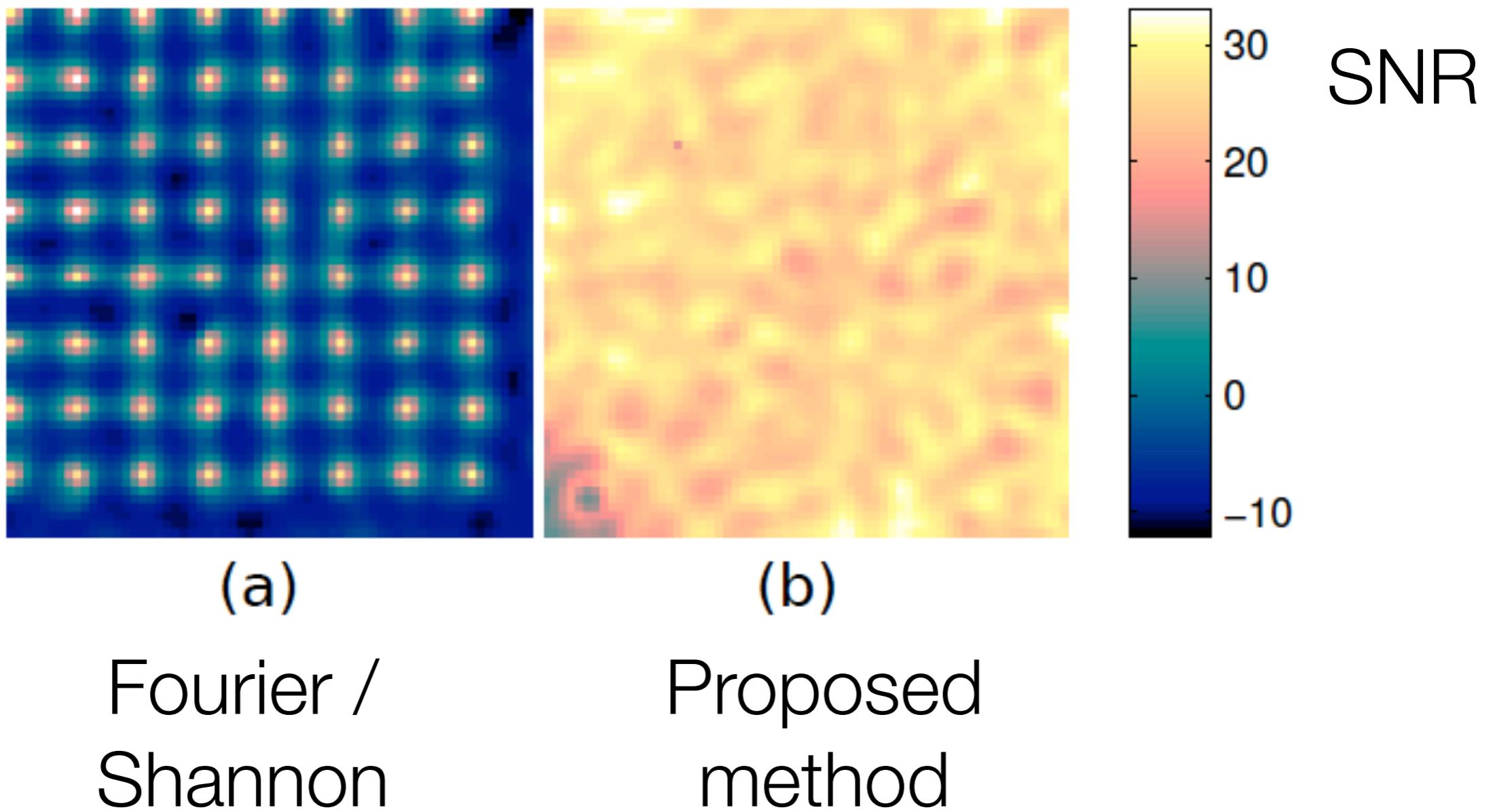
# Recovering the impulse responses



Can be used for calibration of tactile interfaces

# Recovering the impulse responses

---



# Recovering the impulse responses

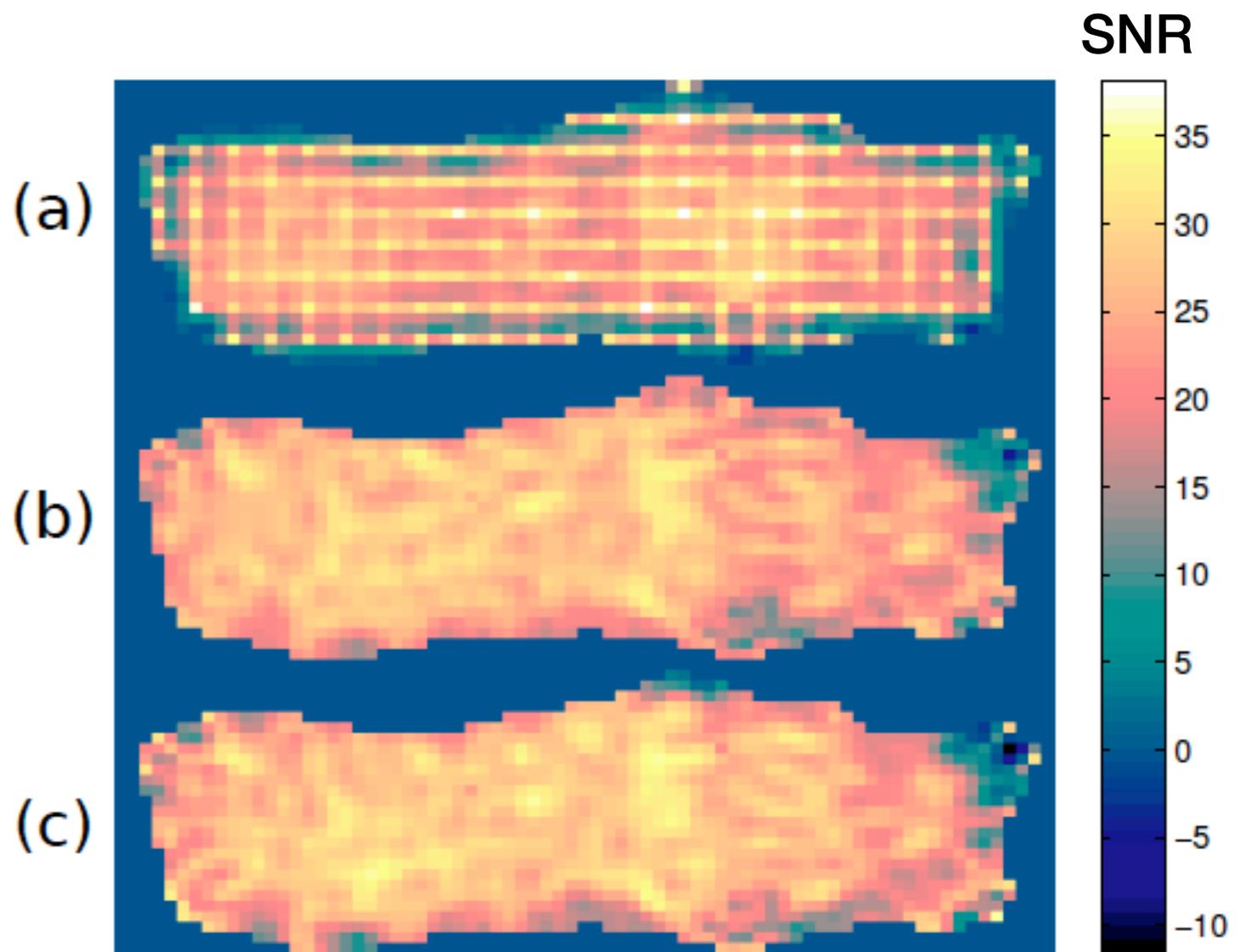
odd  
shape



Fourier interp

Proposed method  
with evanescent waves

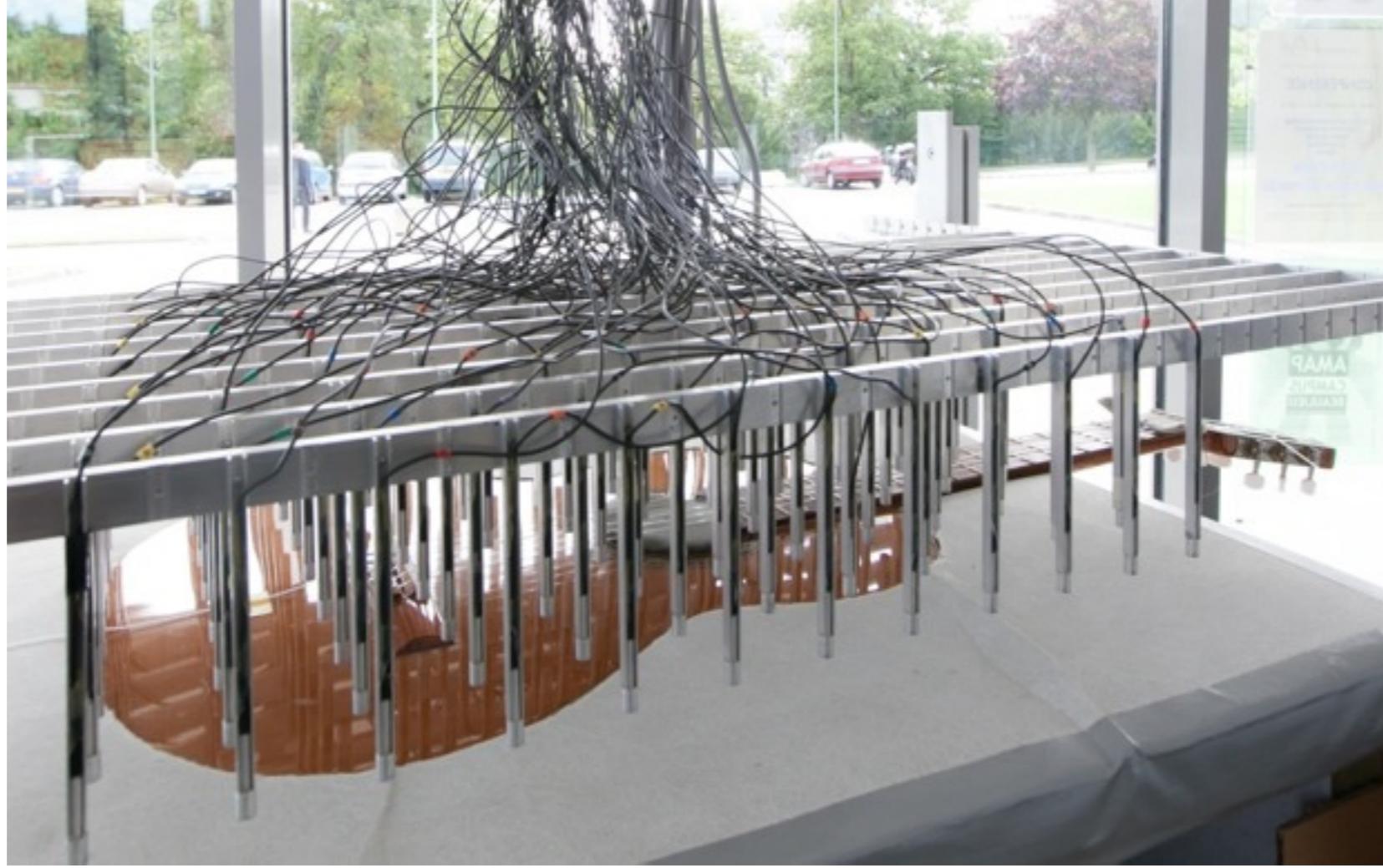
Proposed method  
without evanescent waves



# Compressive Nearfield Acoustic Holography

---

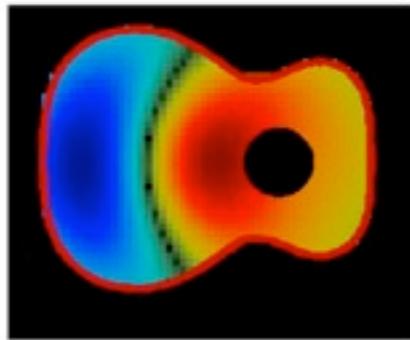
Extension to NAH with sub-Nyquist measurements  
(collab. F. Ollivier and R. Gribonval teams)



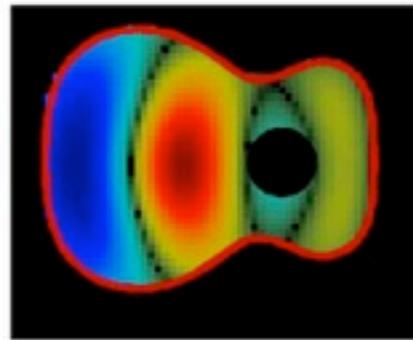
# Nearfield acoustical holography with a guitar

---

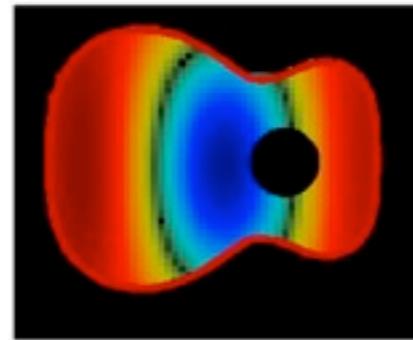
CS 120 micros - f = 275.1 Hz



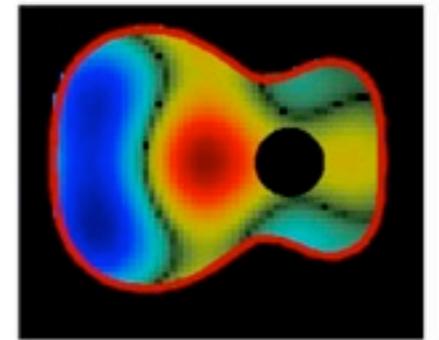
CS 120 micros - f = 310.4 Hz



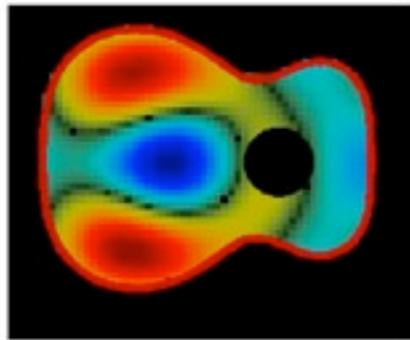
CS 120 micros - f = 339.0 Hz



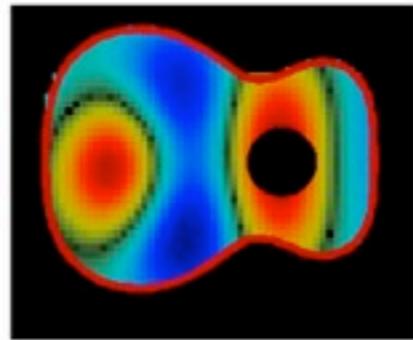
CS 120 micros - f = 396.3 Hz



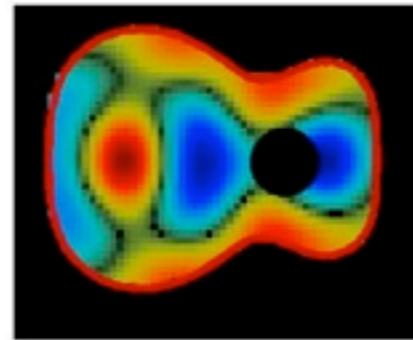
CS 120 micros - f = 470.2 Hz



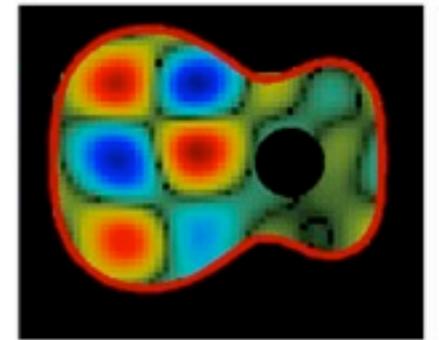
CS 120 micros - f = 506.4 Hz



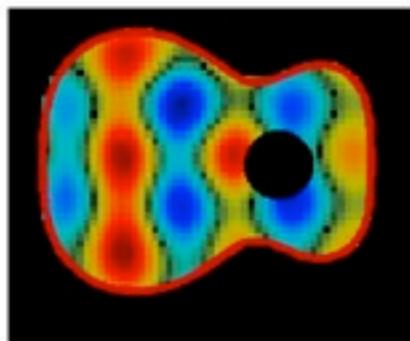
CS 120 micros - f = 647.1 Hz



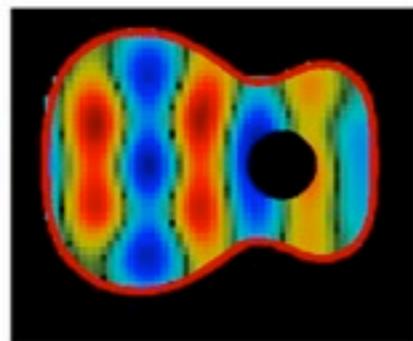
CS 120 micros - f = 707.1 Hz



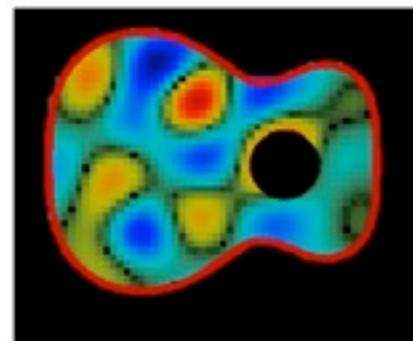
CS 120 micros - f = 802.0 Hz



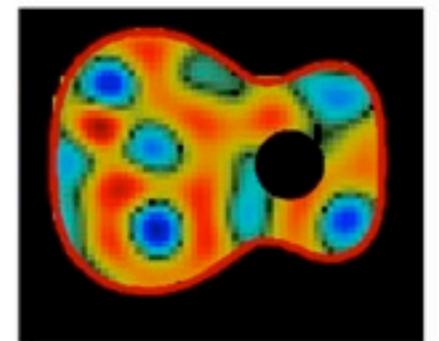
CS 120 micros - f = 861.2 Hz



CS 120 micros - f = 1008.0 Hz



CS 120 micros - f = 1237.4 Hz



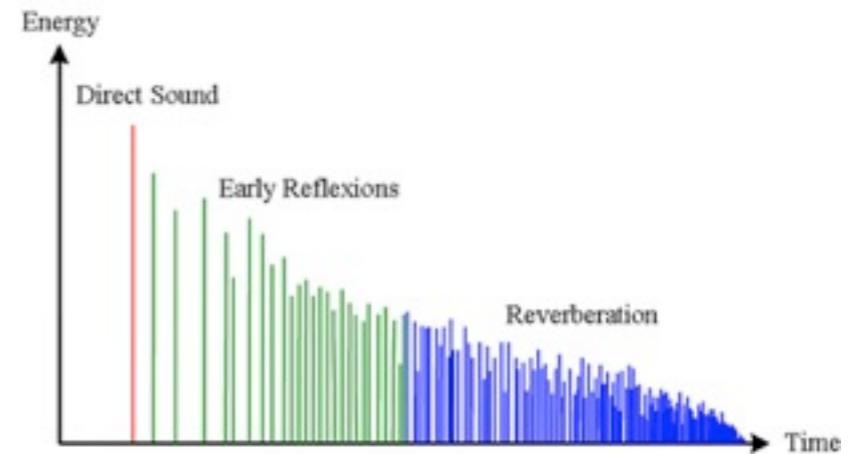
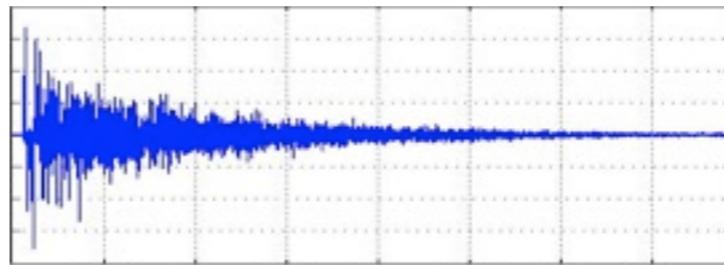
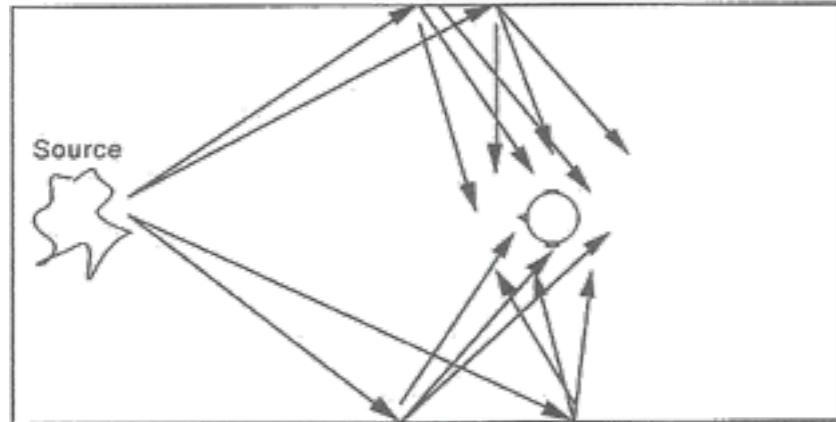
# Compressively sampling the plenacoustic function

---

(with R. Mignot)

# Compressively sampling the plenacoustic function

- The acoustics of a room is characterized by the full set of impulse responses between sources and receivers.



- (abstract) notion of “plenacoustic function” (Ajler / Vetterli, 2002)  
 $p = f(S, R, t, \text{room characteristics})$

“How many microphones do we need to place in the room in order to completely reconstruct the sound field at any position in the room?”

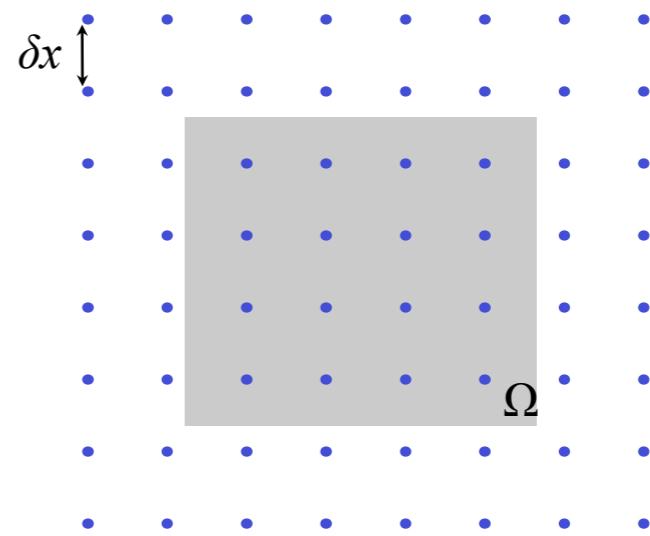
# Compressively sampling the plenacoustic function

- uniform sampling [Ajdler/Vetterli, IEEE TSP 2006]

- Sampling

- with a cutoff frequency  $f_c$ , time sampling is done at  $F_s > 2 f_c$
    - in space, the maximal sampling step  $\delta_x$  is

$$\delta_x < \frac{c_0}{2f_c}$$



- Sampling on a large 3D domain might require a *very* large number of measurements
  - $f_c = 17 \text{ kHz} \rightarrow 10^6 \text{ measurements / m}^3$  !

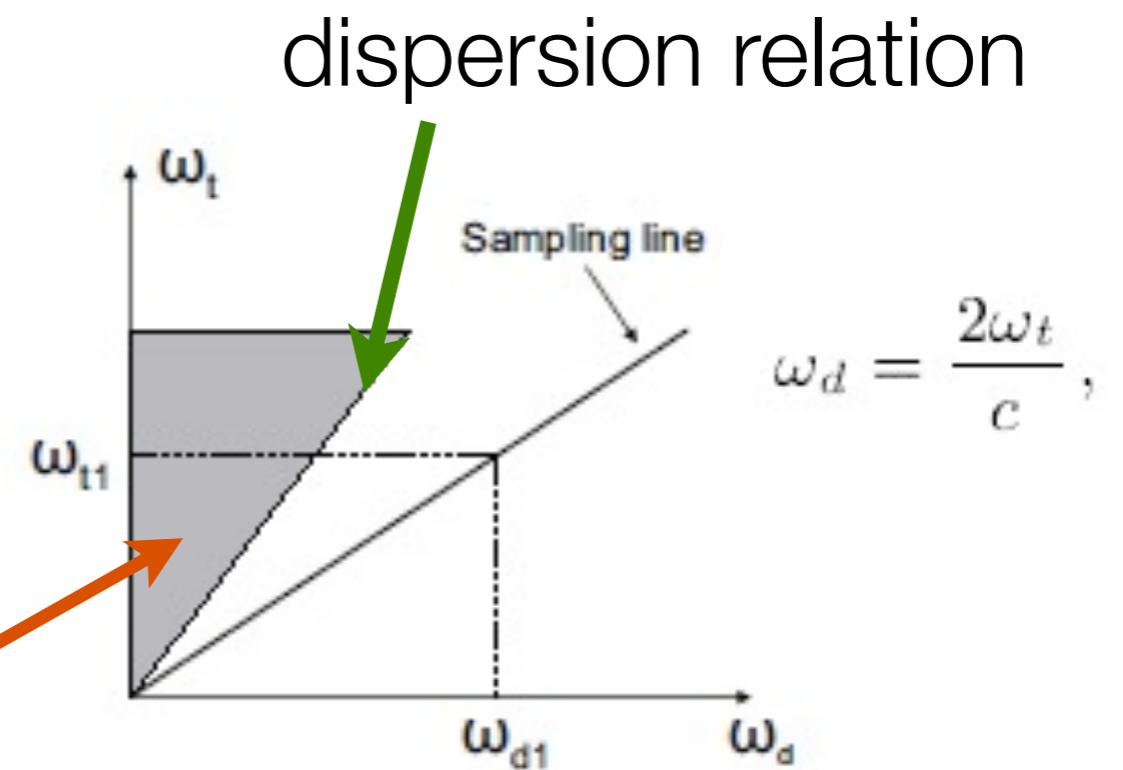
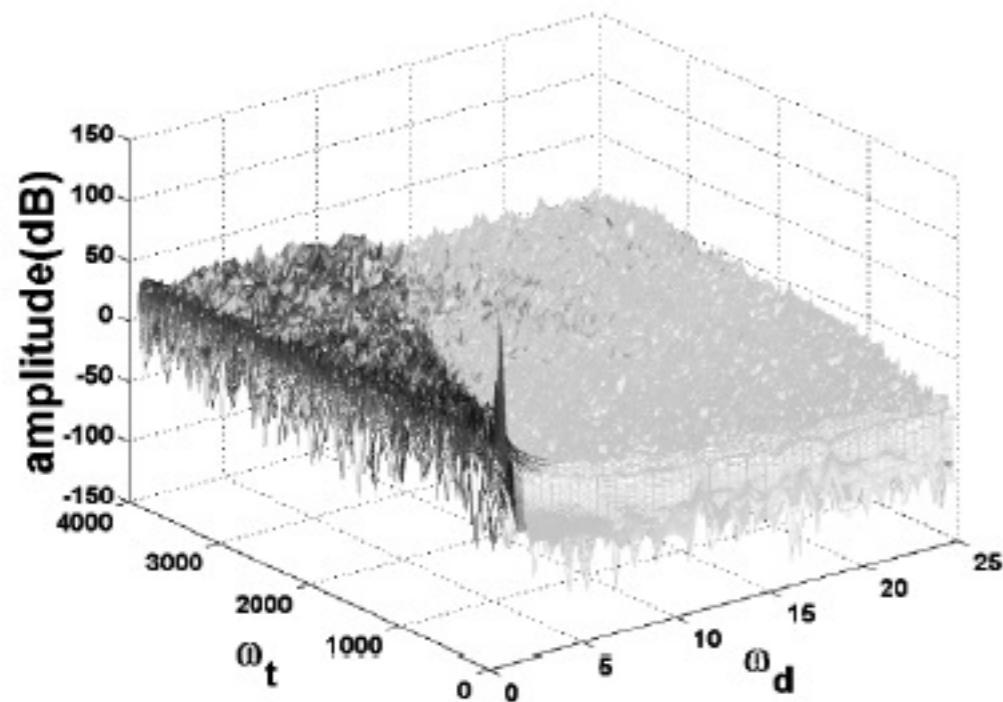
Use physics ! We are not sampling any random signal in  $[0, F_c]$  but solutions of the **wave equation**.



# Compressively sampling the plenacoustic function

S fixed, for a given room  $f(x, t)$

Let us sample  $f$  along the x-axis



The useful information

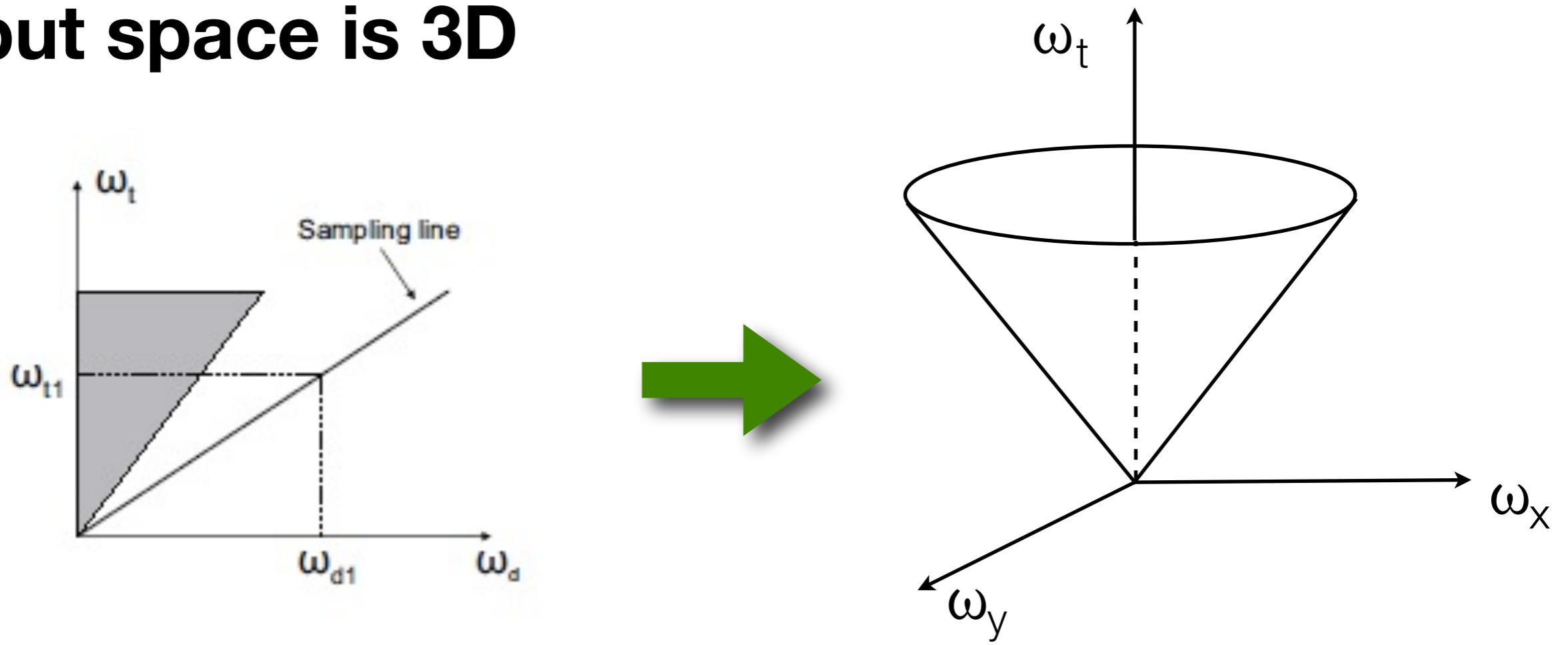
lives here

(Ajdler / Vetterli, 2002)

leads to 1:2 savings in sampling along a line

# Compressively sampling the plenacoustic function

**but space is 3D**



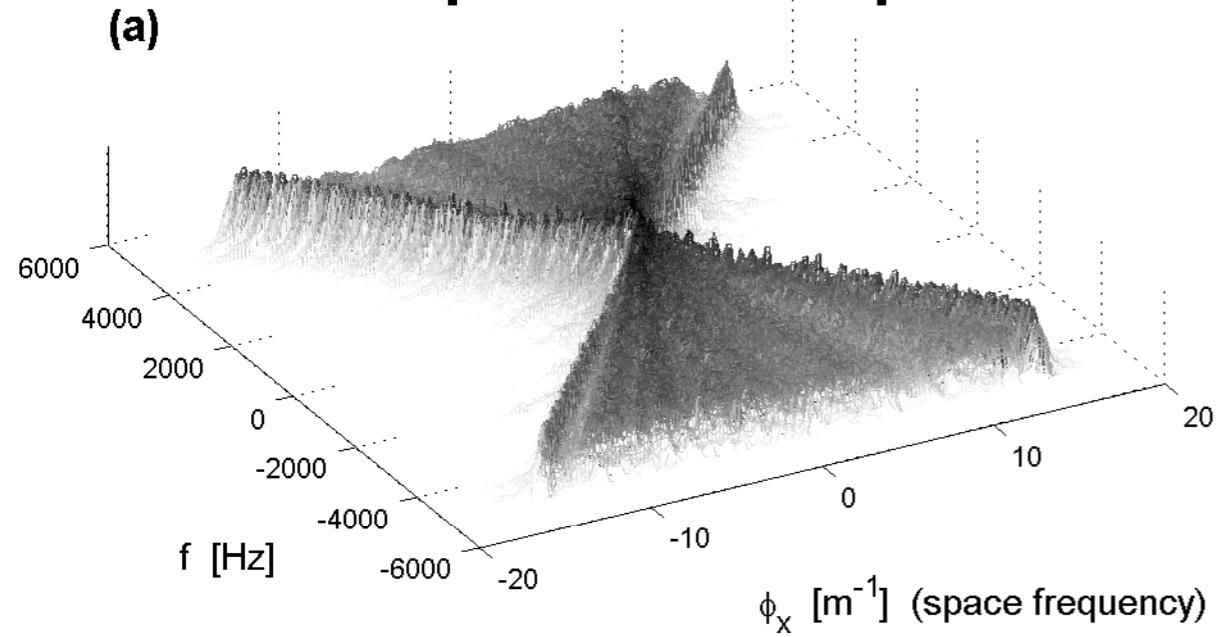
In the 4-D space (3-D + t), the useful information *ideally* lives **on** a cone along the temporal frequency axis

We should “only” have to sample a 3-D surface in a 4-D space

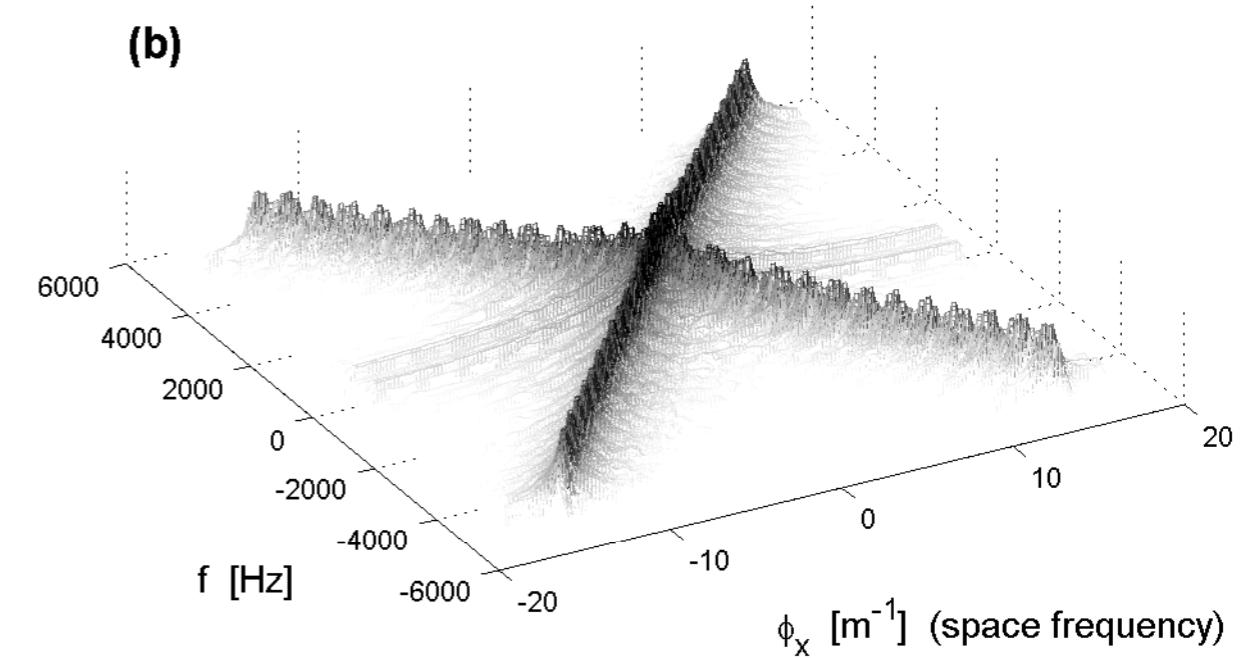
# Compressively sampling the plenacoustic function

Numerical simulation in 2D

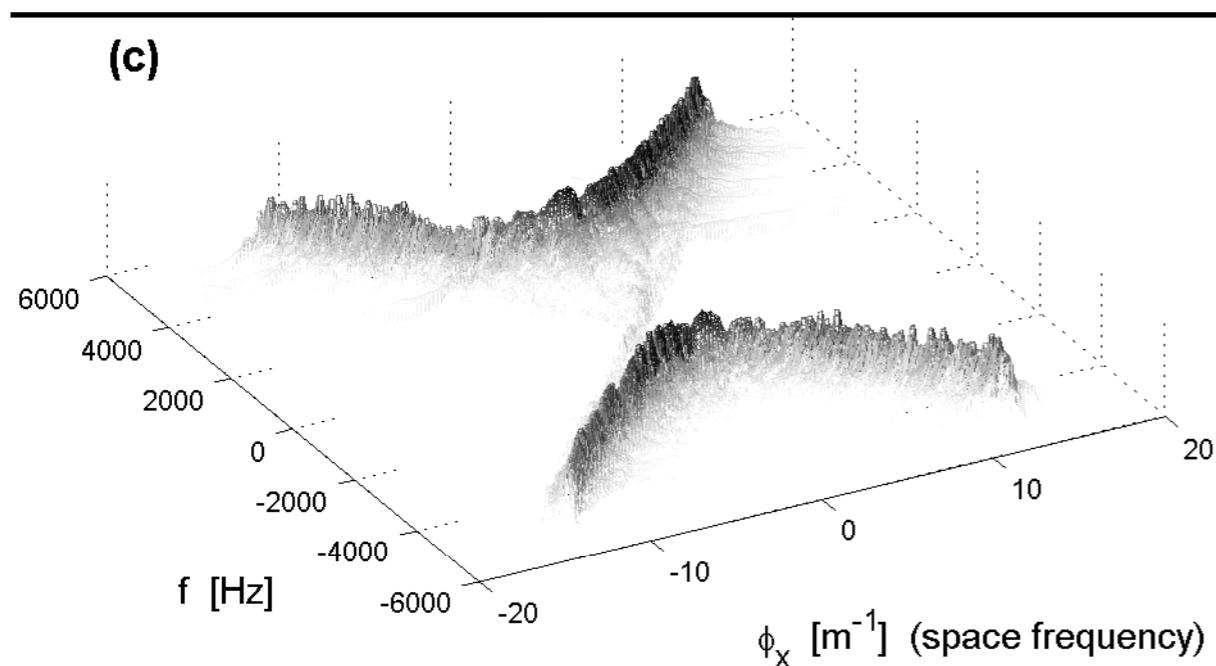
**Sampled in 1D in space**



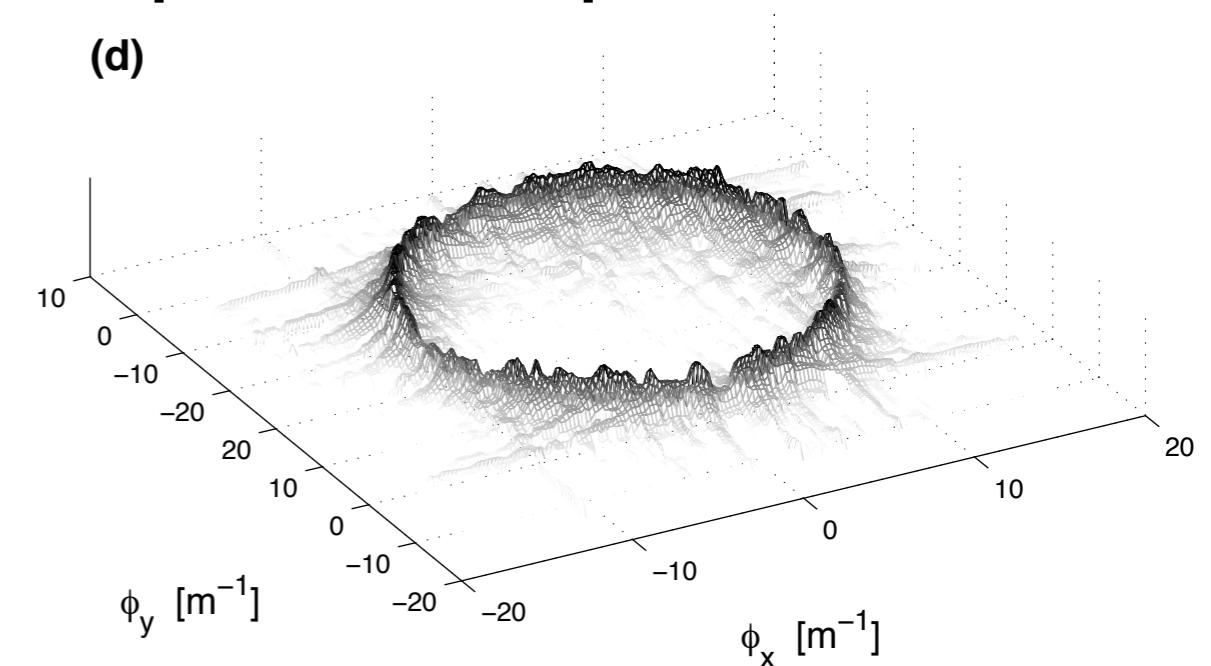
(b)



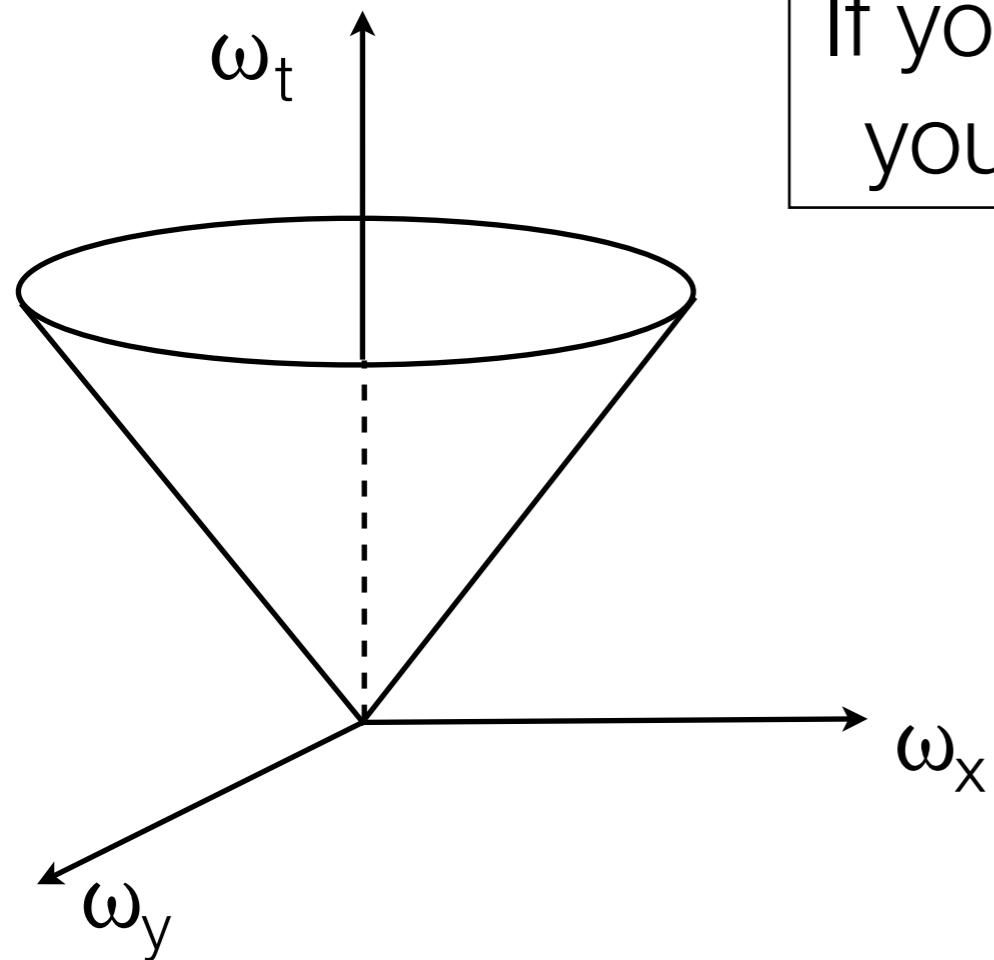
**Sampled in 2D in space**



(d)



# Compressively sampling the plenacoustic function



If you sample in  $n$ -D a wave in  $n$ -D  
you get one dimension «for free»

... however, we cannot use this  
directly as we don't sample in  
the 4D Fourier space, but this  
provides sparsity

Same gain as with integral theorems but  
with more freedom in the sampling scheme  
(and therefore more control on the stability of numerical schemes)

# Compressively sampling the plenacoustic function

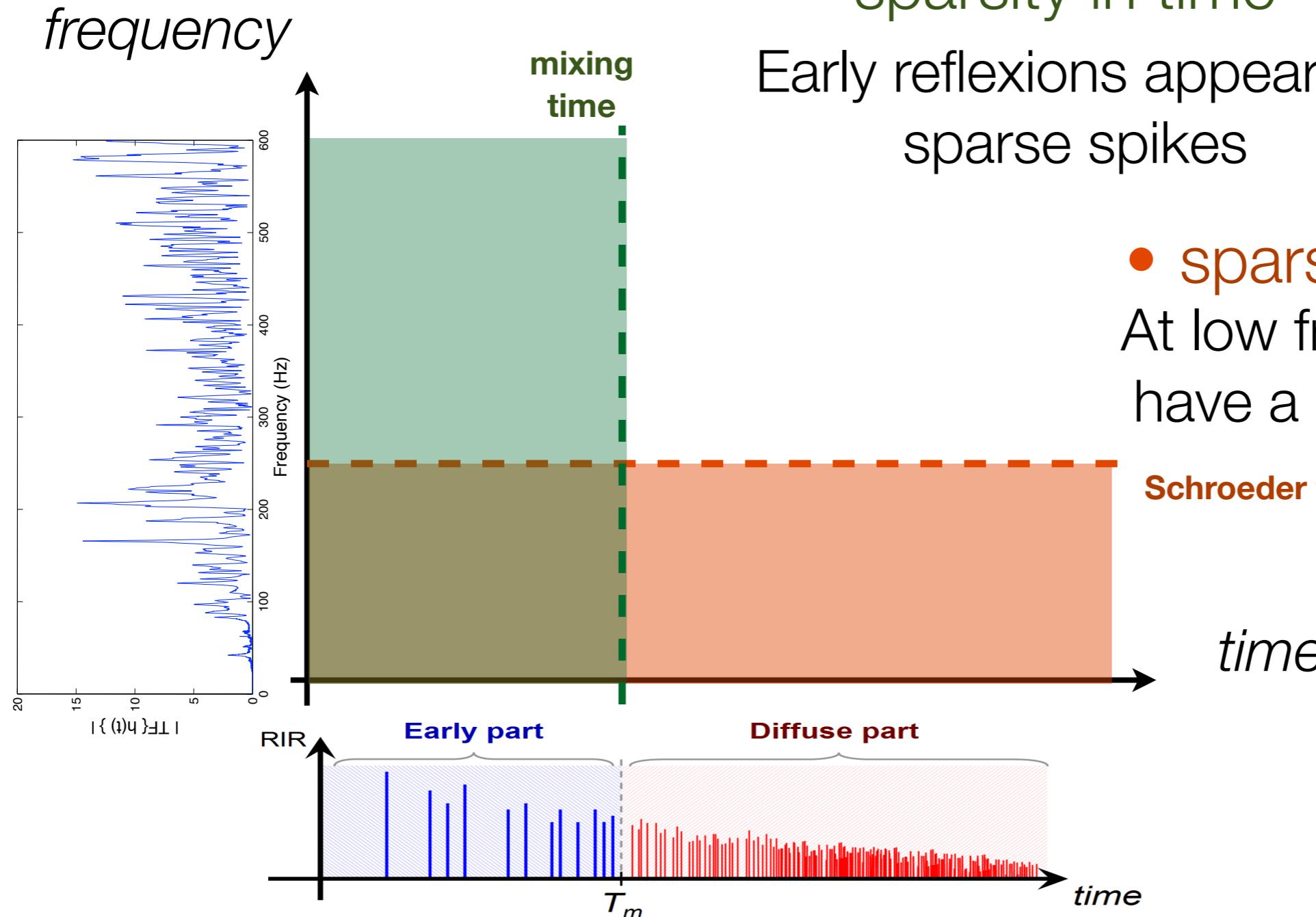
In practice, 2 forms of (approximate) sparsity

- sparsity in time

Early reflexions appear as sparse spikes

- sparsity in frequency

At low frequencies, rooms have a “modal” response



# Compressively sampling the plenacoustic function

---

The goal of this study is

to take advantage of this prior information on the signals

*directly at the acquisition stage*

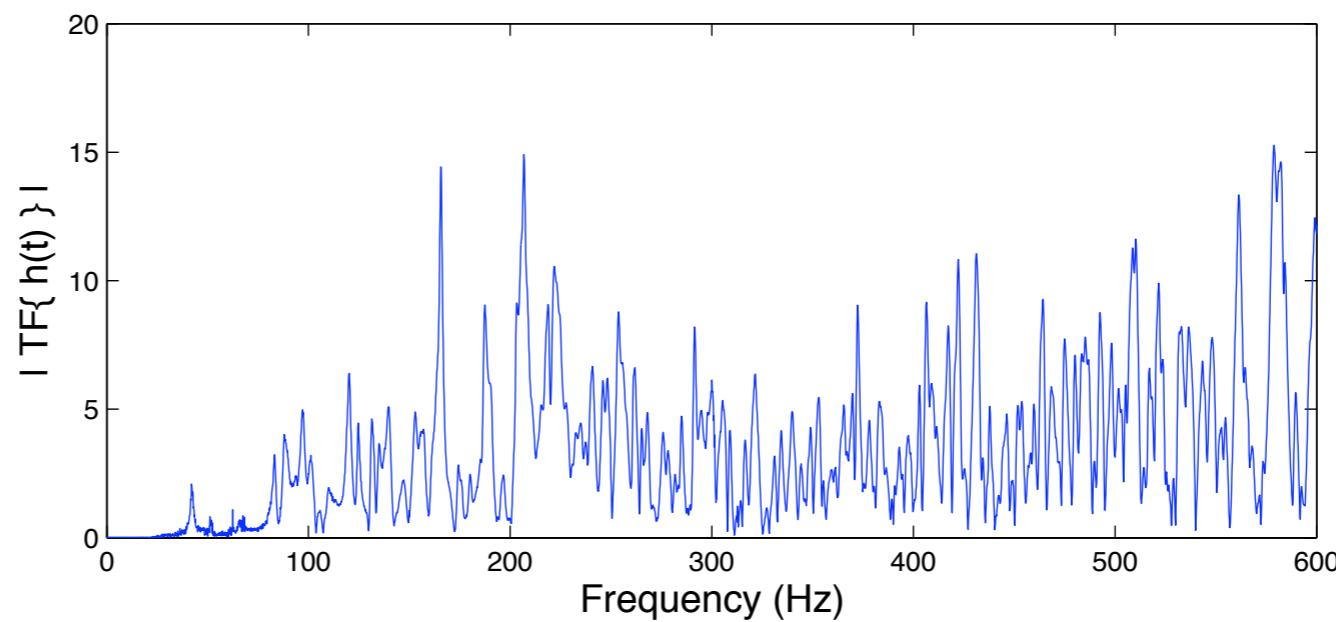
in order **to reduce the number of measurements**,

using the sampling paradigm of compressed sensing

# Compressively sampling the plenacoustic function

- sparsity in frequency

At low frequencies, rooms have a “modal” response



$$\tilde{p}(\vec{X}, t) = \sum_{q,r} \tilde{\alpha}_{q,r} e^{j(k_q c_0 t + \vec{k}_{q,r} \cdot \vec{X})} \quad \|\vec{k}_{q,r}\| = k_q$$

structured dictionary of plane waves, sparse in the number of modes

2-step algorithm: find frequencies (joint sparsity across measurements)  
then modes

# Compressively sampling the plenacoustic function

---

3 different sparse models

- sparsity in frequency
  - structured plane wave model (CS1)  
at a given frequency, not sparse in the k domain
  - unstructured plane wave model (CS2)  
sparse in frequency and in the k domain
- sparsity in time  
image source model

# Sparsity in frequency (method CS1)

---

- Wave equation

$$c_0^2 \Delta p(\vec{X}, t) - \partial_t^2 p(\vec{X}, t) = 0$$

- Modal theory for closed rooms (with ideally rigid walls)

$$p(\vec{X}, t) = \sum_{q \in \mathbb{Z}^\star} A_q \phi_q(\vec{X}) e^{j k_q c_0 t}$$

with walls attenuation  $k_q \in \mathbb{C}$

# Sparsity in frequency (method CS1)

---

## Plane wave approximation

each spatial mode  $\Phi_q$  (with wavenumber  $k_q$ ) is approximated by R plane waves (whose vectors all have the same amplitude  $|k_q|$ )

$$\phi_q(\vec{X}) \approx \sum_{r=1}^R a_{q,r} e^{j\vec{k}_{q,r}\vec{X}} \quad \|\vec{k}_{q,r}\|_2 = |k_q|$$

Within a finite frequency range  $[0, \omega_c]$  containing Q real modes (2Q complex), the plenacoustic function can be approximated as a sum of 2QR damped harmonic plane waves

$$e^{j(k_q c_0 t + \vec{k}_{q,r} \vec{X})}$$

with coefficients  $\alpha_{q,r} = A_q a_{q,r}$

# Sparsity in frequency (method CS1)

---

## Algorithm CS1

- measure Impulse Response at  $M$  points in volume  $\Omega$
- estimate shared modes frequencies using a joint sparsity (across  $M$  measurements) model for damped sinusoids (ex. MUSIC, SOMP, ...)
- for a given mode
  - construct corresponding dictionary of plane waves (sampling of the sphere)
  - get coefficients by least-square projection
- interpolate plenacoustic function at every position of  $\Omega$

$$\tilde{p}(\vec{X}, t) = \sum_q \tilde{\alpha}_{q,r} e^{j(k_q c_0 t + \vec{k}_{q,r} \cdot \vec{X})}$$

## Sparsity in frequency (method CS2)

---

- Assume a rectangular room  
space variables get de-coupled

$$p(\vec{X}, t) = \sum_{q \in \mathbb{Z}^*} A_q F_{xq}(x) F_{yq}(y) F_{zq}(z) e^{j k_q c_0 t}$$

one mode : 8 plane waves       $\vec{k} = [\pm k_x, \pm k_y, \pm k_z]$

- formulation as a standard sparse optimization problem
- has to be further simplified for processing
- greedy approach

# Sparsity in time with the image source model

- sparsity in time

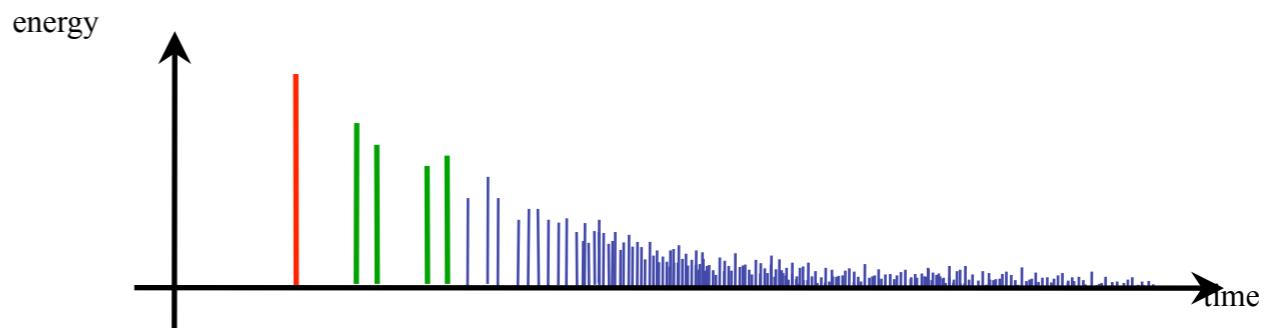
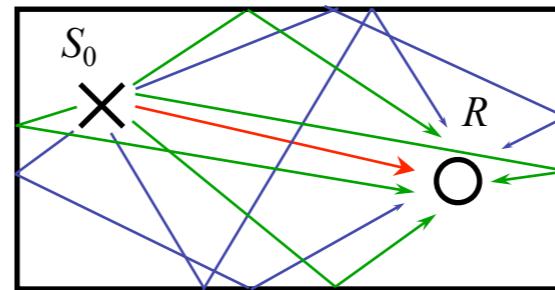
In an enclosed space:

direct path,

1<sup>st</sup>-order reflections,

and multiple reflections.

 sources  
 receiver

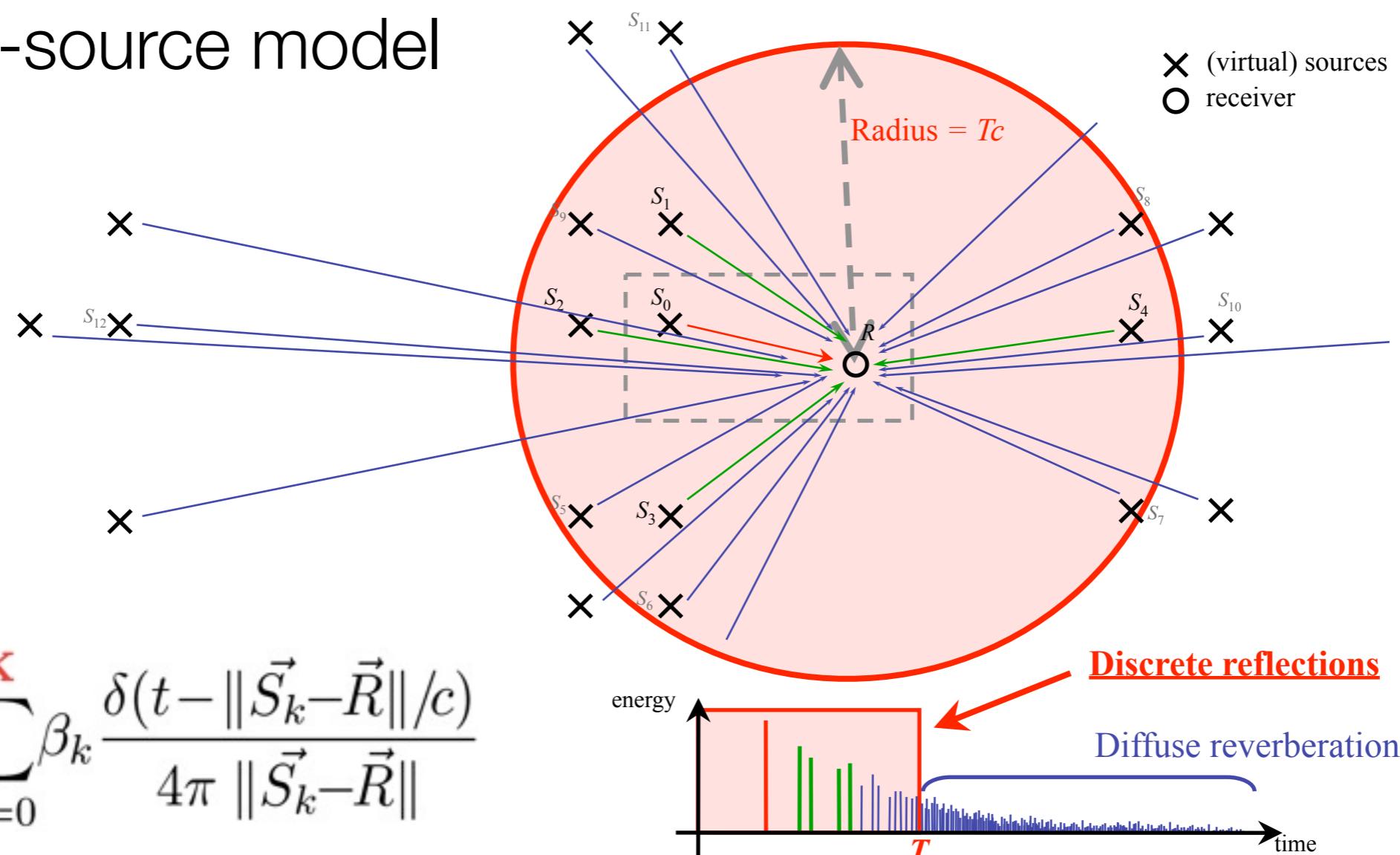


$$p(t, \vec{R}) =$$

# Sparsity in time with the image source model

- sparsity in time

image-source model

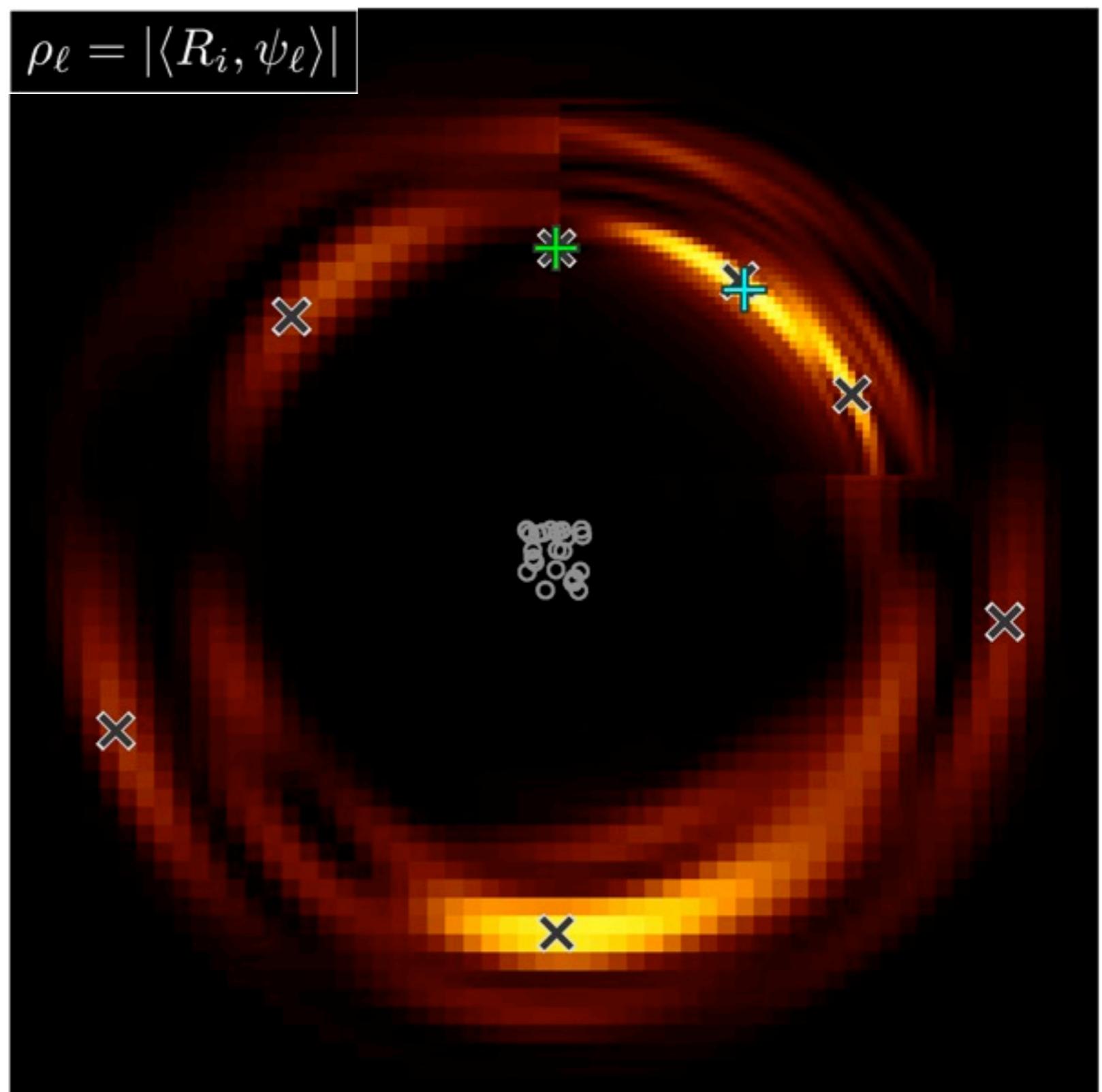


unstructured dictionary of monopole sources,  
sparse in the number of virtual sources

Toy example:

Simulation of 7 synthetic sources in a  
2D free field.

- microphones
- ✗ synthetic sources
- + estimated sources

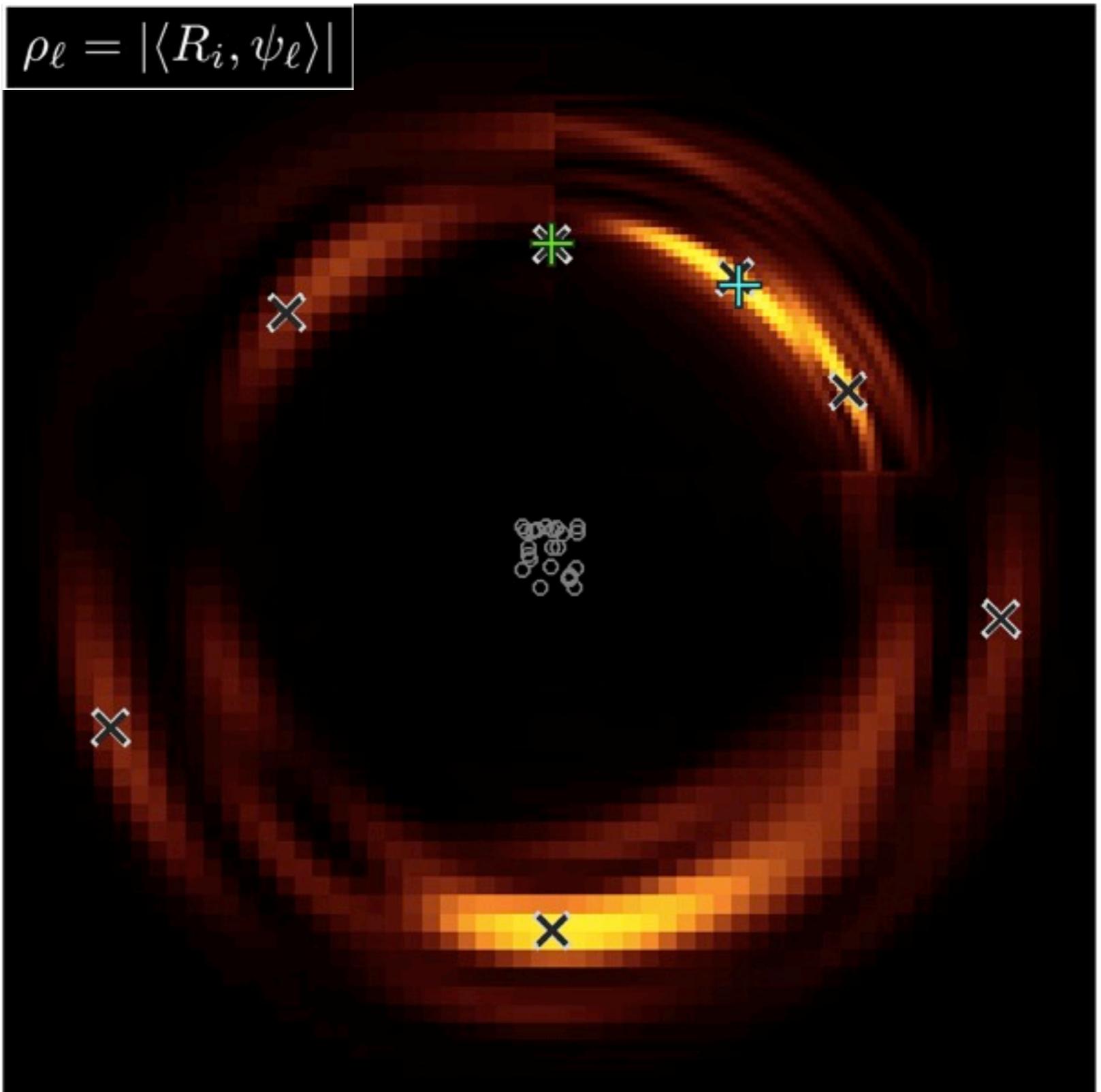


Toy example:

$$\rho_\ell = |\langle R_i, \psi_\ell \rangle|$$

Simulation of 7 synthetic sources in a  
2D free field.

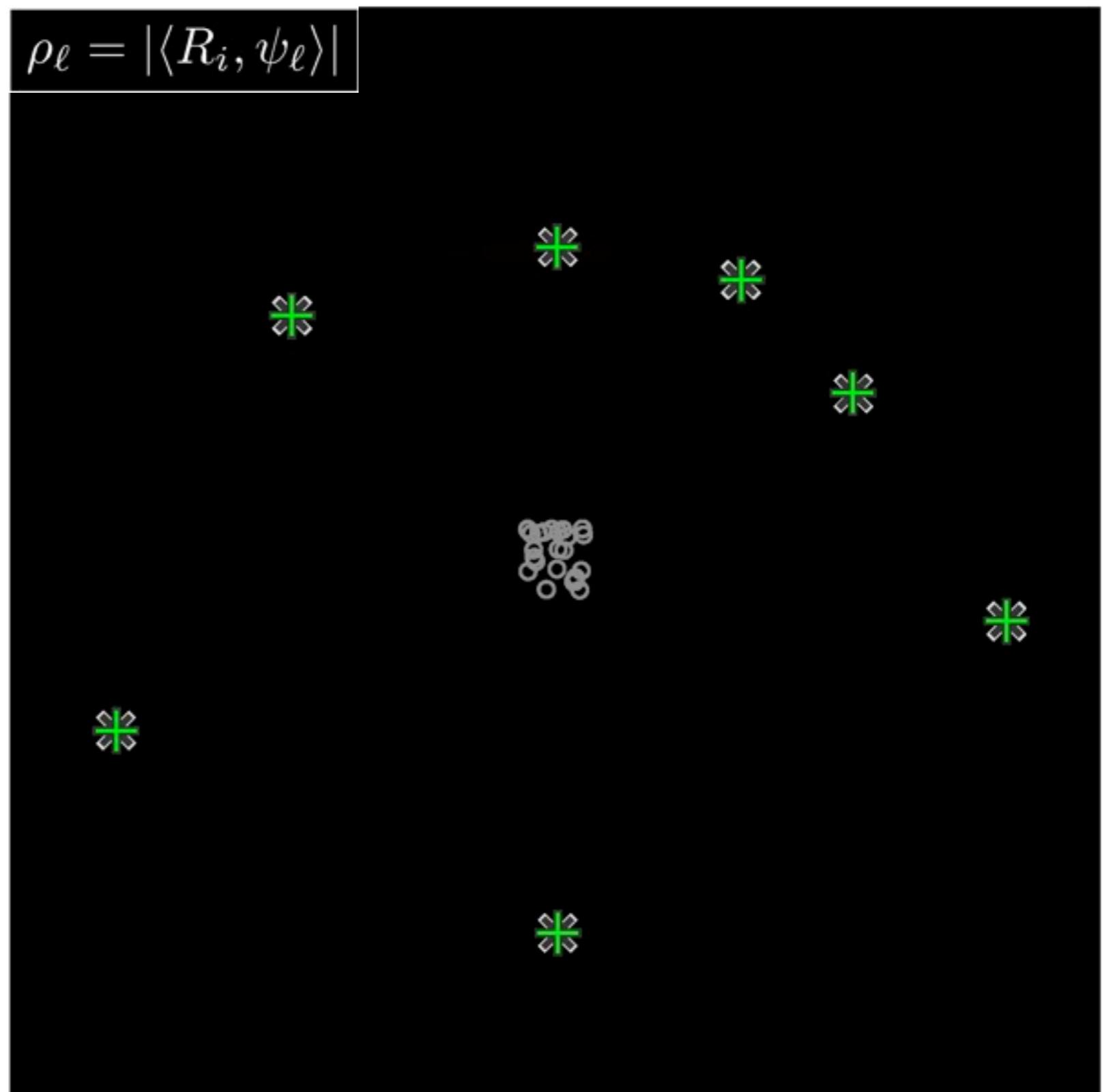
- microphones
- ✗ synthetic sources
- + estimated sources



### Toy example:

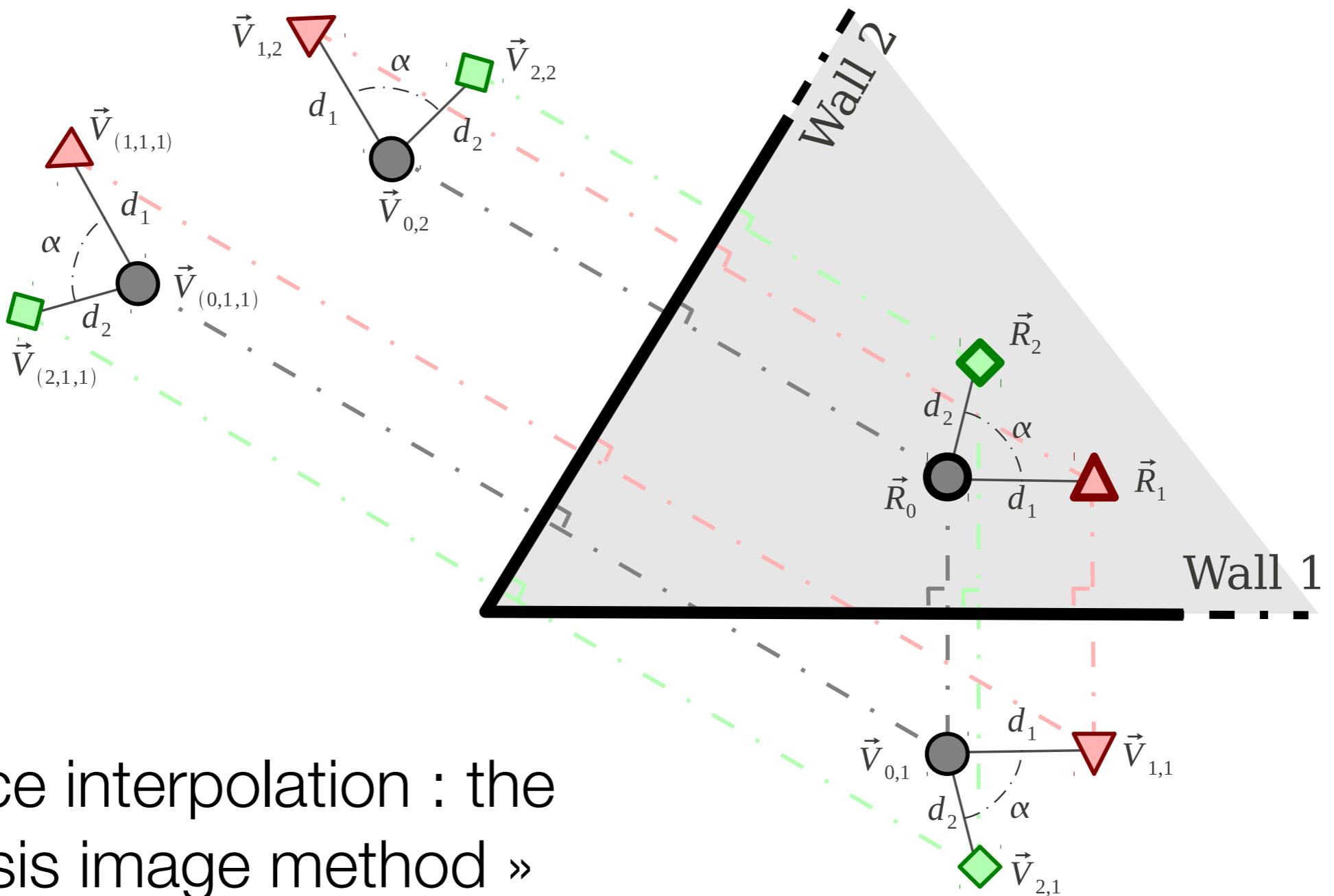
Simulation of 7 synthetic sources in a  
2D free field.

- microphones
- ✗ synthetic sources
- + estimated sources



## Multi-resolution MP algorithm

# Sparsity in time with the image source model

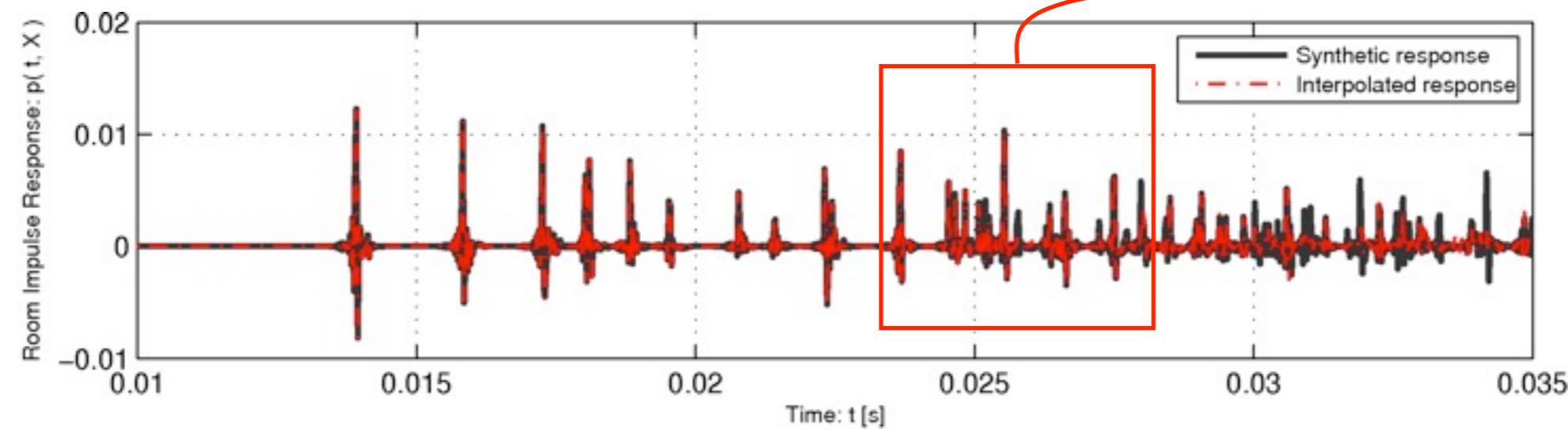
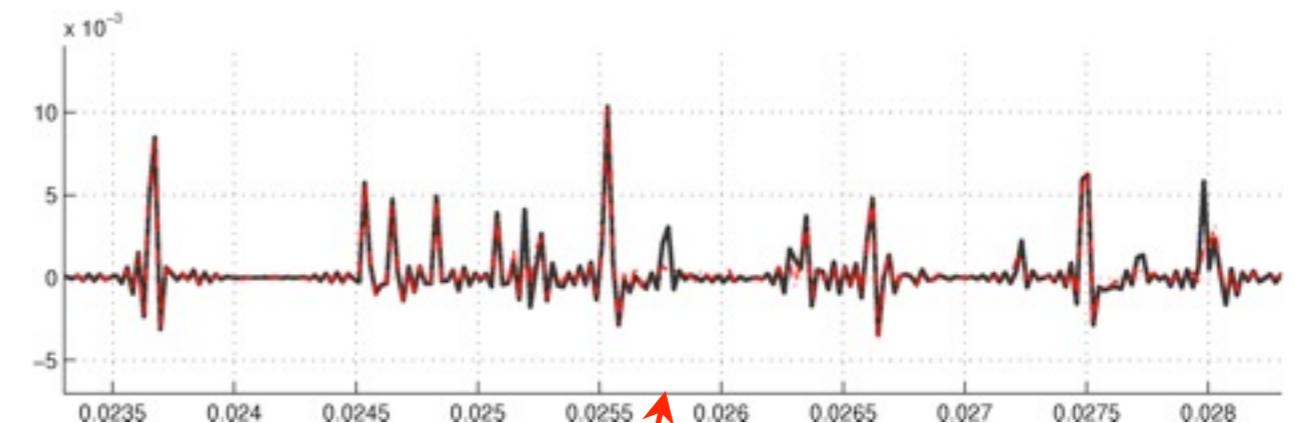
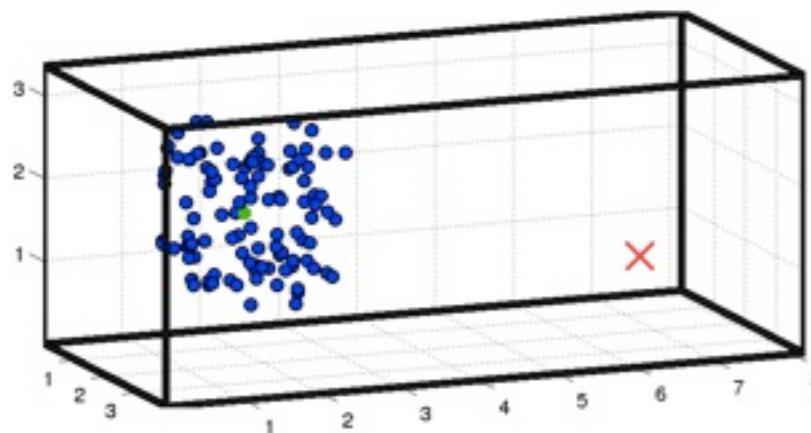


Source interpolation : the  
« basis image method »

# Sparsity in time with the image source model

## ■ Results on **synthetic** Room Impulse Responses:

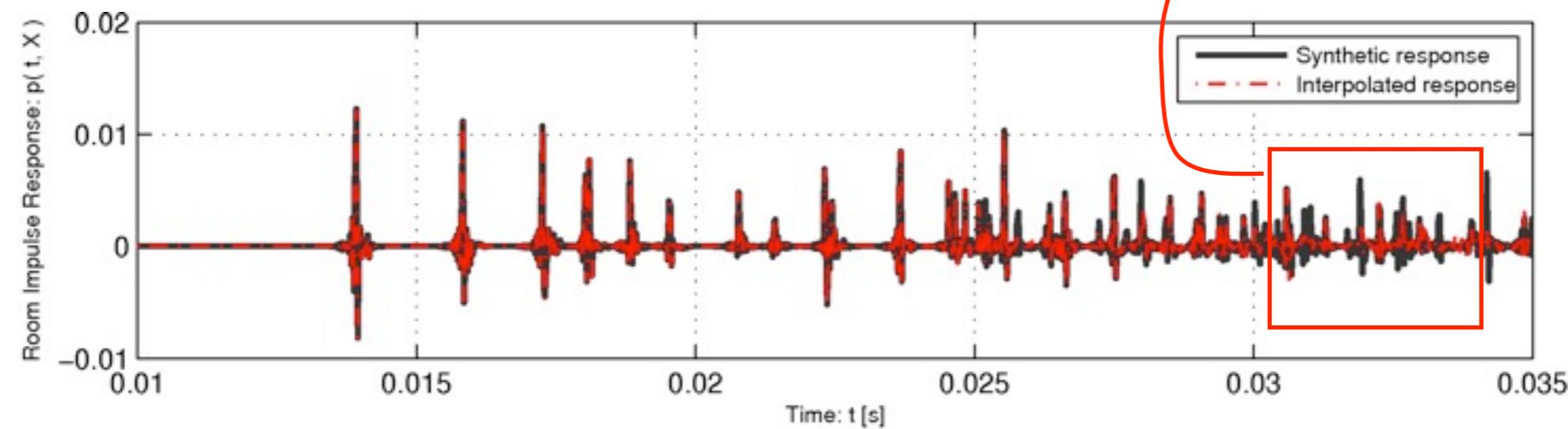
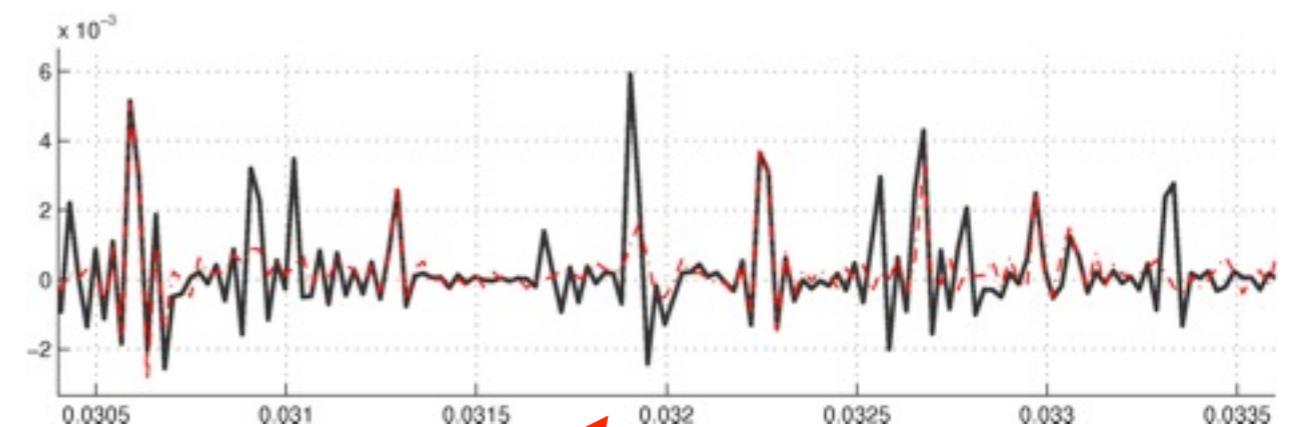
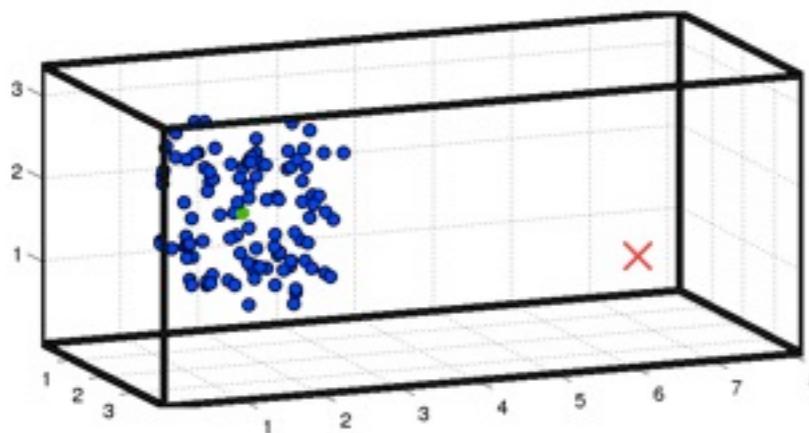
- Sampling, Analysis with  $M = 120$  microphones (●) and 1 fixed source (✗)
- Simulation using [Allen,JASA,1976], for a rectangular room ( $3.8 \times 8.15 \times 3.6$ )m
- Interpolation on another position (●)
- $F_s = 44100$  Hz,  $f_c = 20000$  Hz.



# Sparsity in time with the image source model

## ■ Results on **synthetic** Room Impulse Responses:

- Sampling, Analysis with  $M = 120$  microphones (●) and 1 fixed source (✗)
- Simulation using [Allen,JASA,1976], for a rectangular room ( $3.8 \times 8.15 \times 3.6$ )m
- Interpolation on another position (●)
- $F_s = 44100$  Hz,  $f_c = 20000$  Hz.



# Microphone array experiment

---



Room size  
 $L_x = 3.88 \text{ m}$   
 $L_y = 8.15 \text{ m}$   
 $L_z = 3.3 \text{ m}$

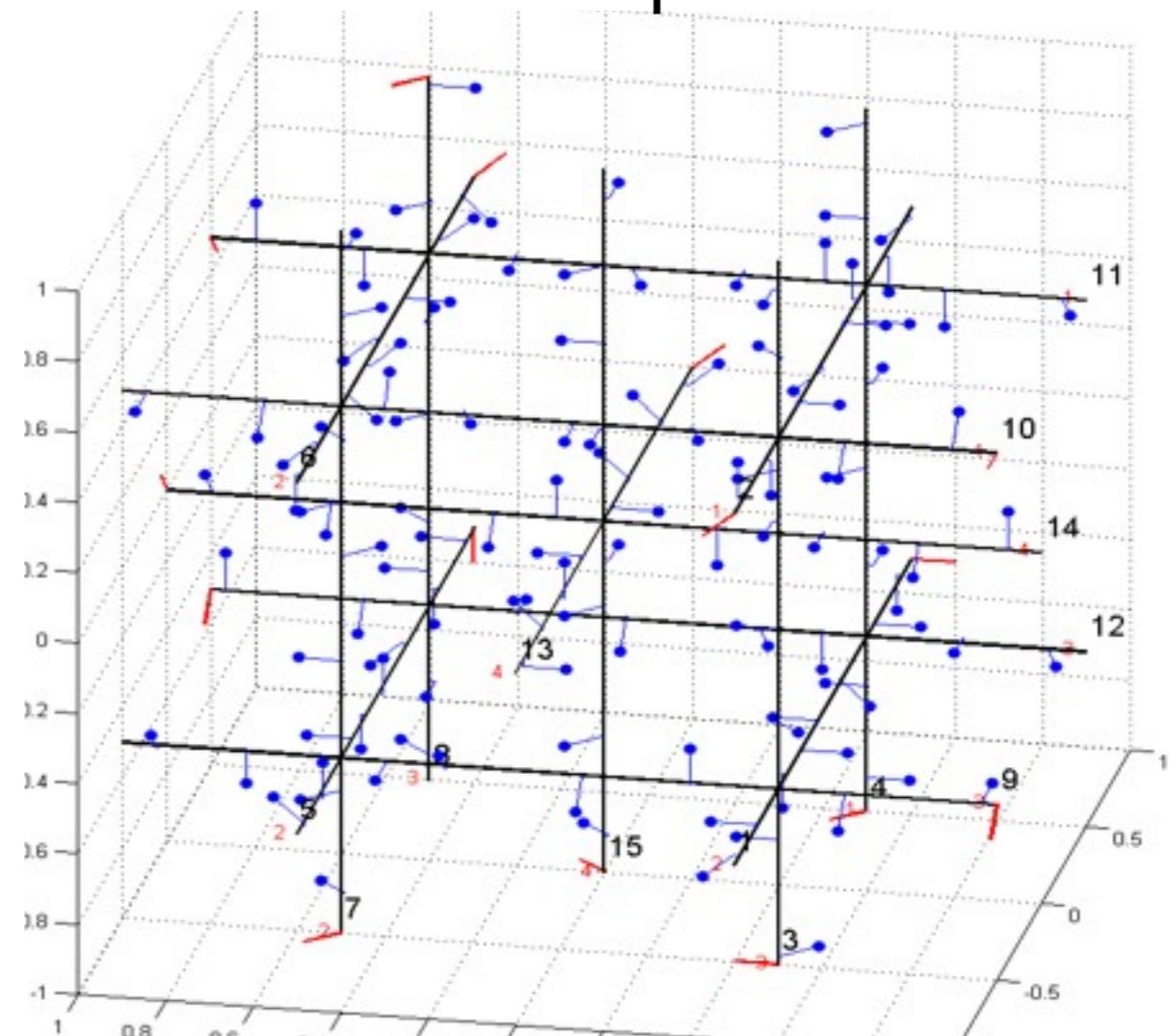
# Compressively sampling the plenacoustic function

## experimental setup



IJLRA, team MPIA

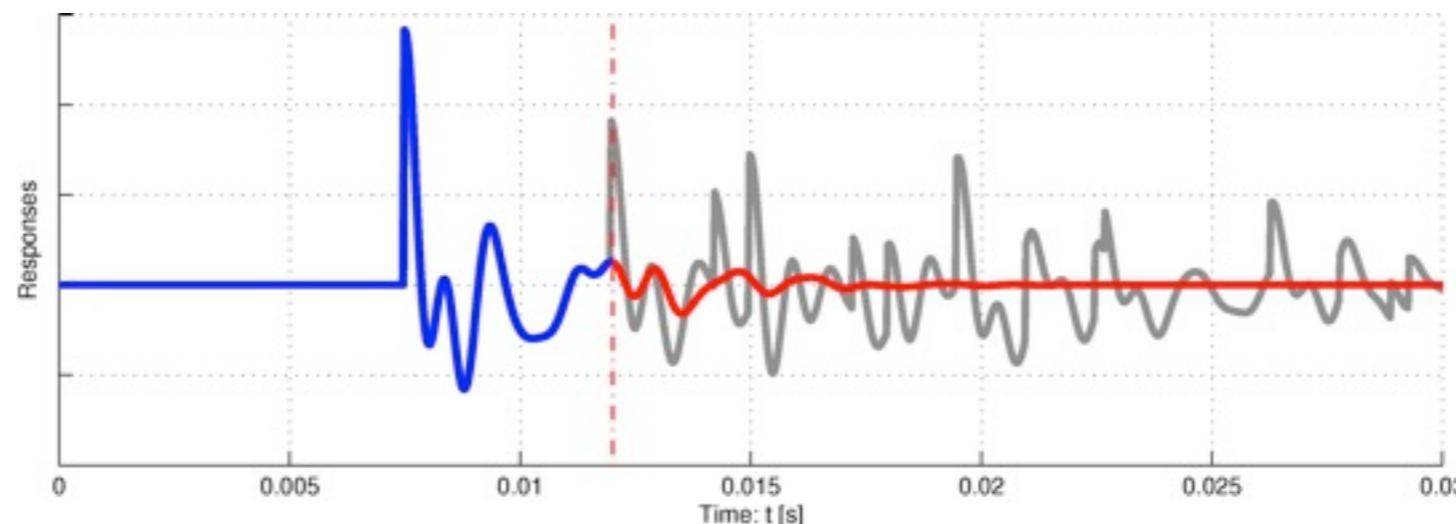
Irregular sampling of  
a  $2^*2^*2$ m cube  
120 microphones



# Calibration issues

## ■ Microphone array experiment

- First calibration:
  - Estimation of the microphone positions  $\mathbf{R}_m$  using [Ono,Waspaa,2009].
- **Measures:**  $H_{\text{measured}} = \mathbf{H}_{\text{room}} \mathbf{G}$ , where  $\mathbf{G} = H_{\text{excitation}} \ H_{\text{loudspeaker}} \ H_{\text{microphone}}$ 
  - In practice, we need to estimate  $\mathbf{G}$ .
  - Problem: without anechoic measures, we need to make “in situ” estimations.



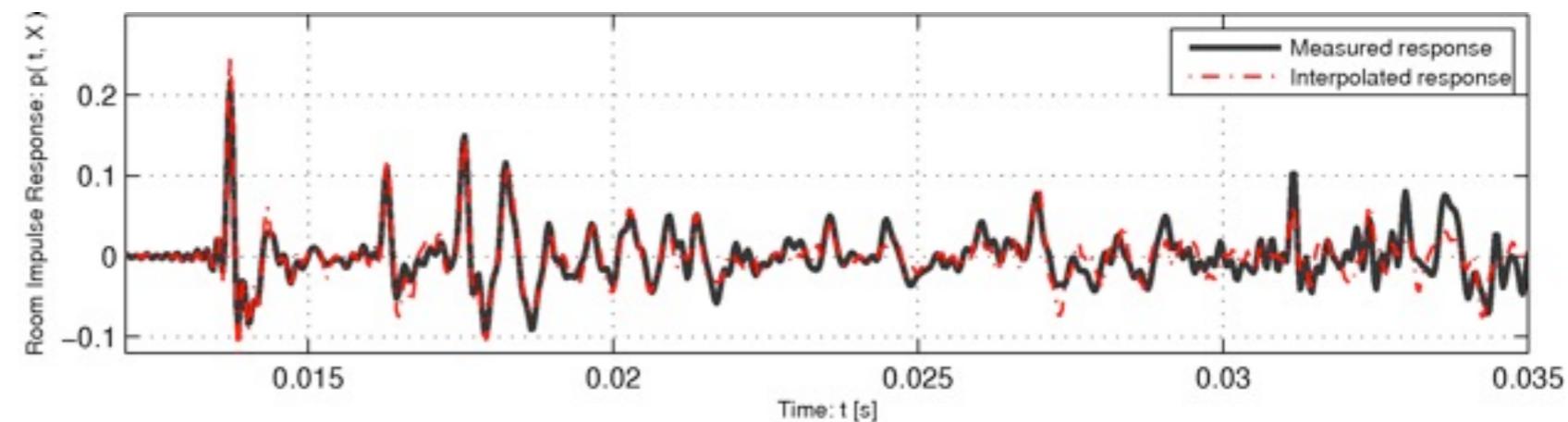
- Microphone responses have to be estimated for each microphone !

# Compressively sampling the plenacoustic function

## Microphone array experiment

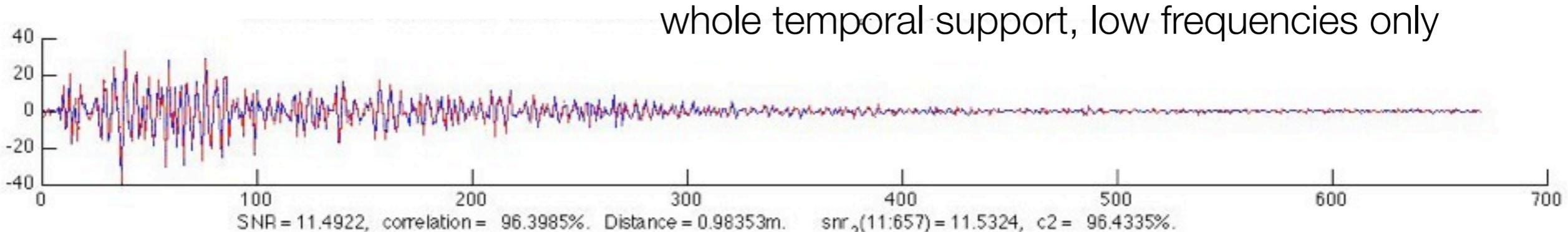
Analysis with 119 microphones, interpolation of the 120<sup>th</sup> response.

- sparsity in time



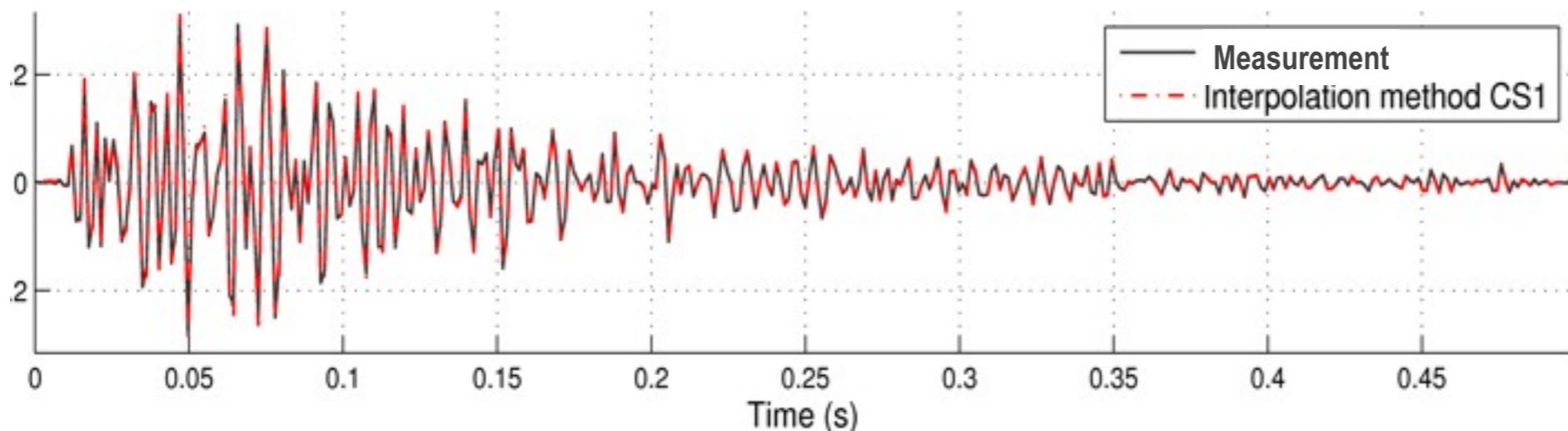
full band, beginning of impulse responses only

- sparsity in frequency

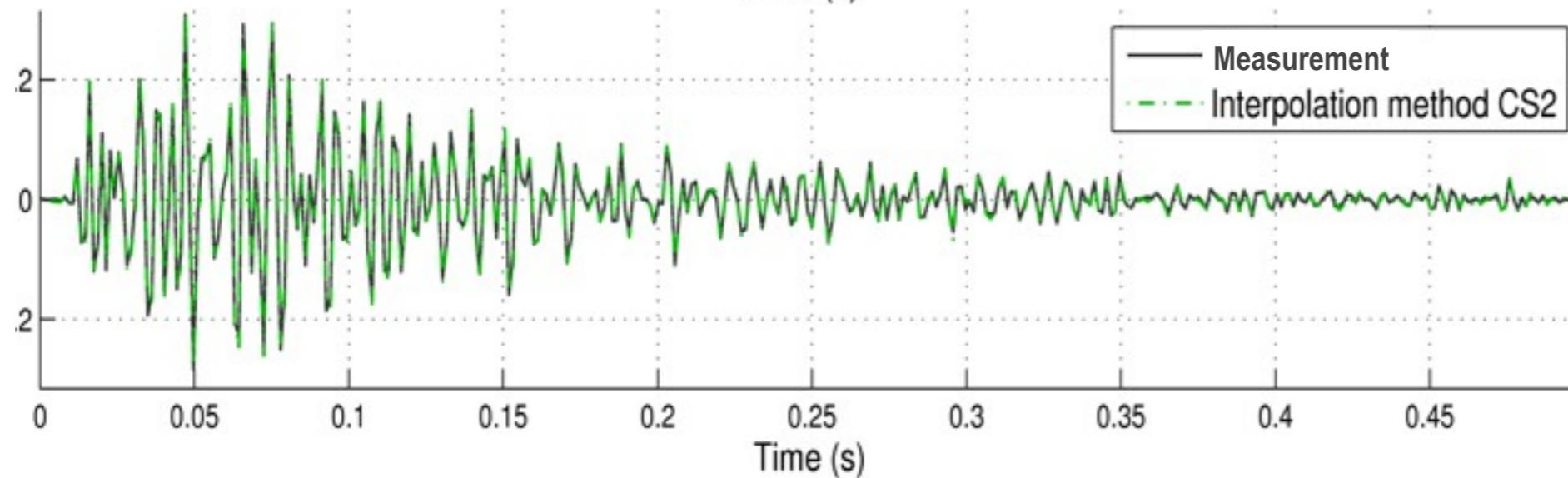


# Microphone array experiment

- sparsity in frequency
- Interpolation at a point near the centre of the volume  
(distance to the center = 32 cm)  
 $F_s = 750 \text{ Hz}$ ,  $f_c = 300 \text{ Hz}$ ,  $K = 216$  modes,  $M = 119$



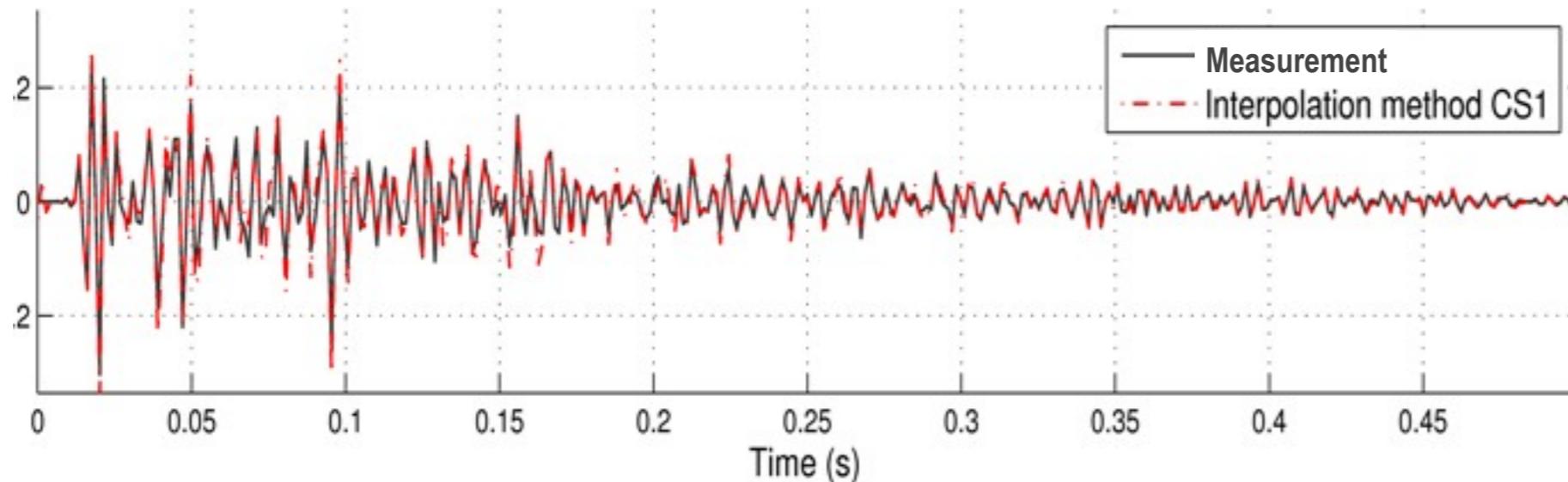
SNR = 21.8 dB  
Corr = 99.7 %



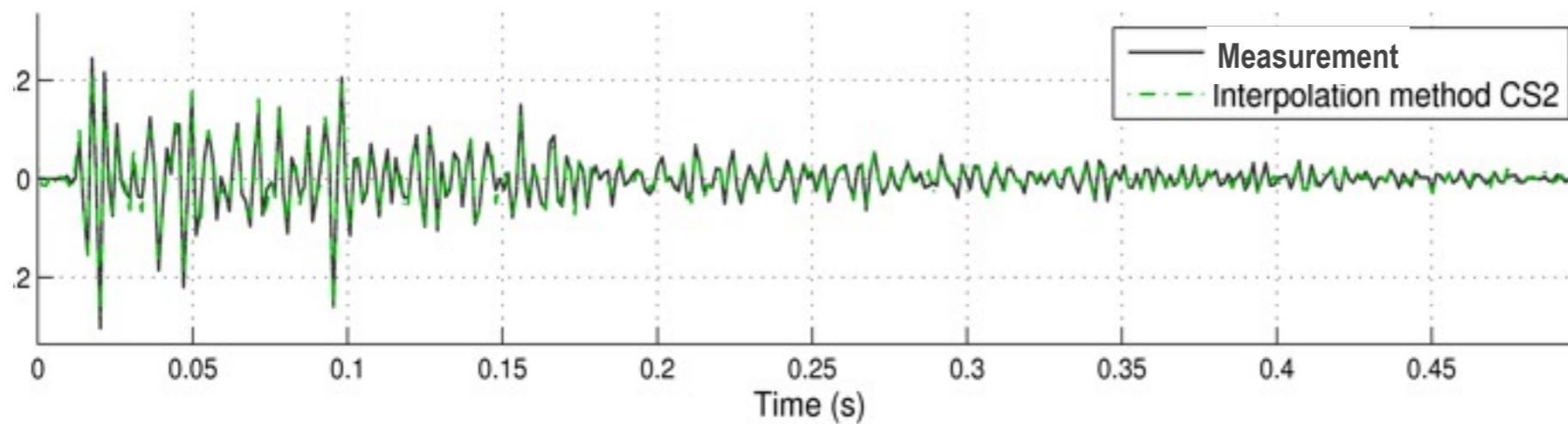
SNR = 19.1 dB  
Corr = 99.4 %

# Microphone array experiment

- sparsity in frequency
- Interpolation at a point further away from the center  
(distance to the center = 90 cm)  
 $F_s = 750 \text{ Hz}$ ,  $f_c = 300 \text{ Hz}$ ,  $K = 216$  modes,  $M = 119$



SNR = 9.7 dB  
Corr = 95.8 %

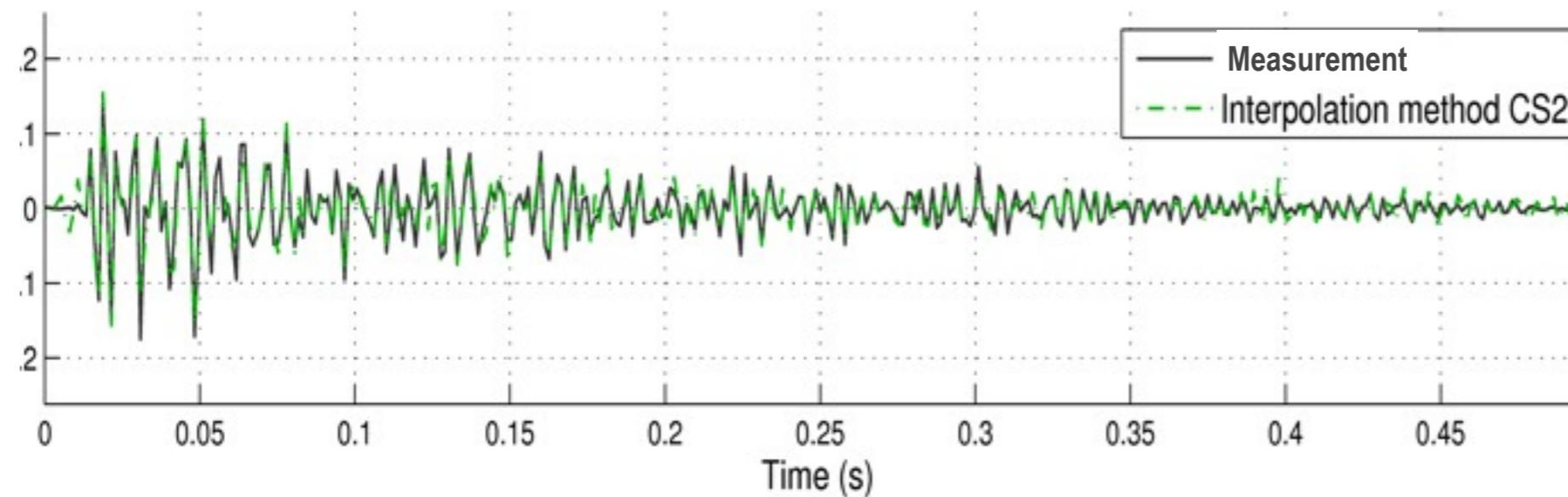
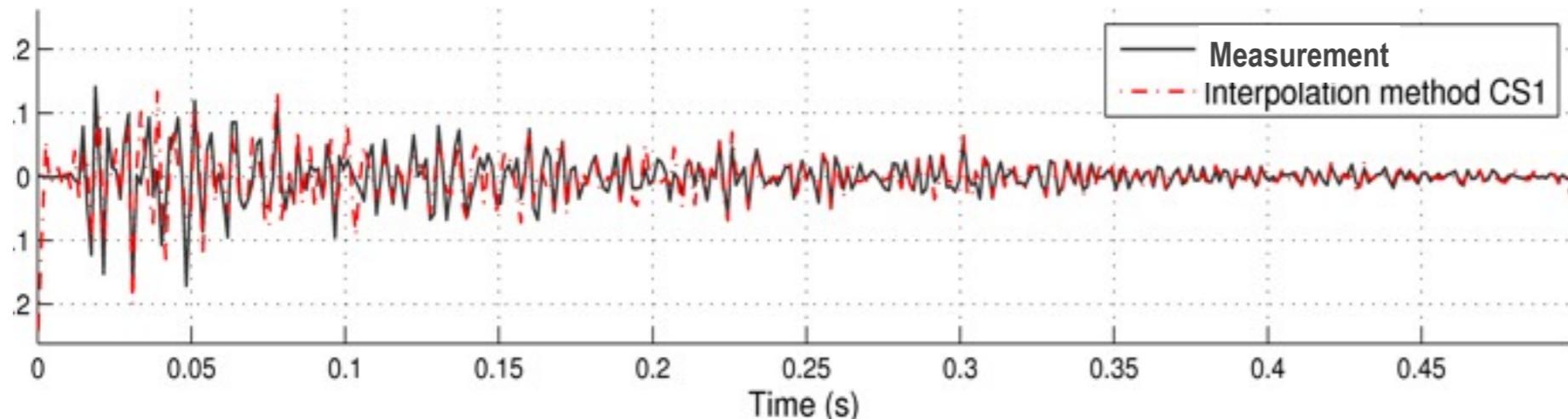


SNR = 12.3 dB  
Corr = 97.0 %

# Microphone array experiment

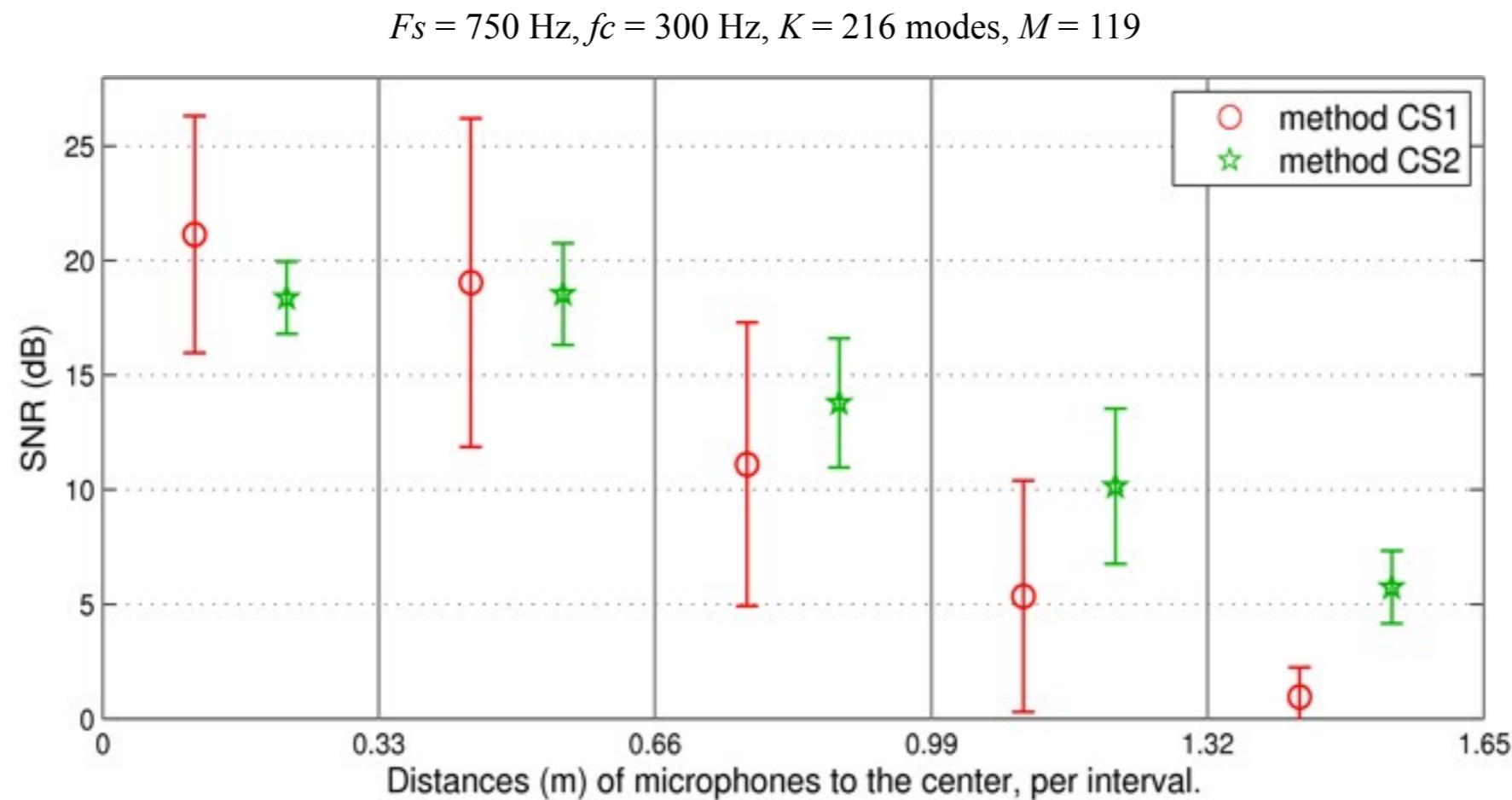
- sparsity in frequency

- Interpolation at a point at the edge of the volume  
(distance to the center = 134 cm)  
 $F_s = 750 \text{ Hz}$ ,  $f_c = 300 \text{ Hz}$ ,  $K = 216$  modes,  $M = 119$



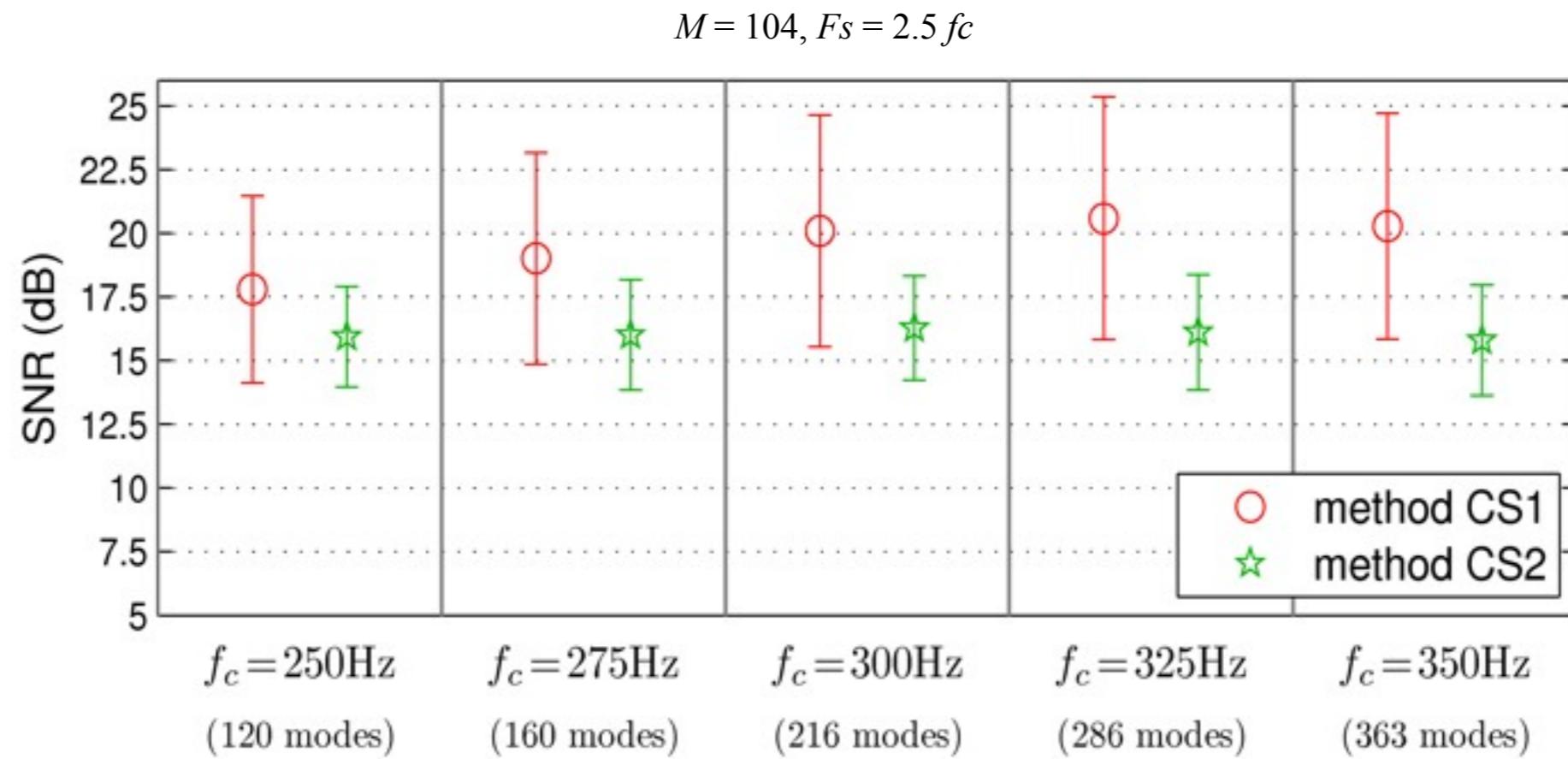
# Microphone array experiment

- sparsity in frequency
  - Dependency on the distance to the center



# Microphone array experiment

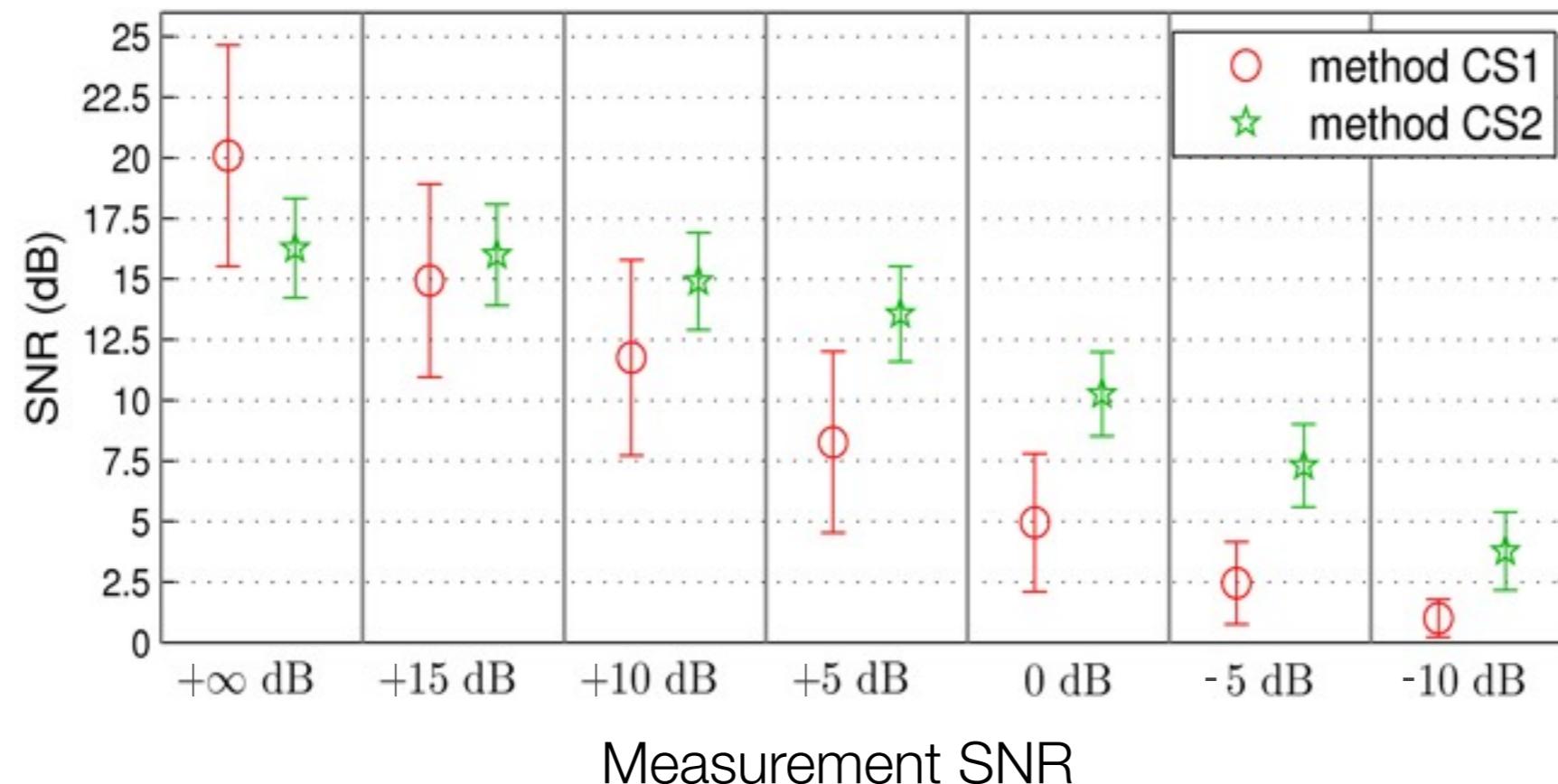
- sparsity in frequency
  - Dependency on the frequency cutoff



# Microphone array experiment

- sparsity in frequency
- Robustness to additional noise

$F_s = 750$  Hz,  $f_c = 300$  Hz,  $K = 216$  modes,  $M = 104$

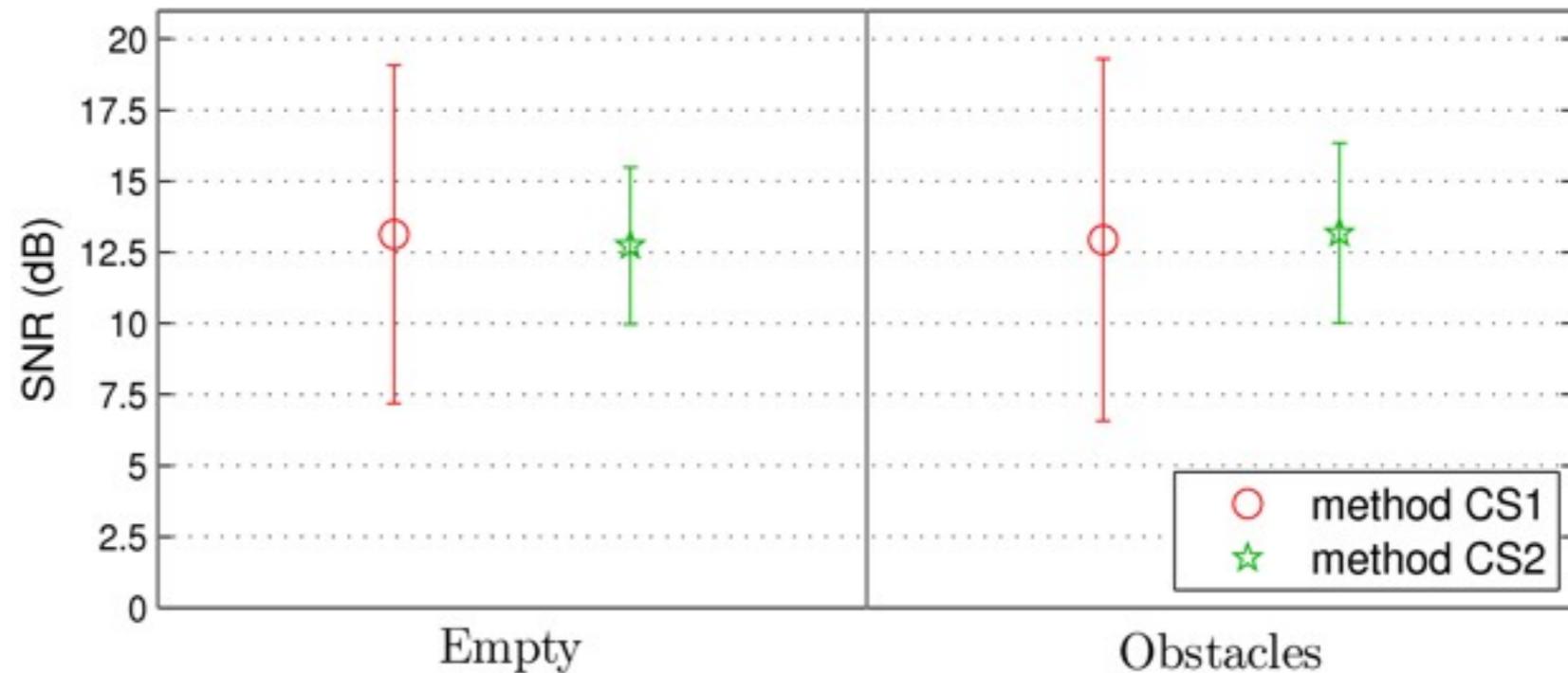


# Microphone array experiment

---

- Comparison between an **empty** room and a room with **obstacles** : open door and window, chair, wood panel (remember that CS2 assumes a perfect rectangular room)

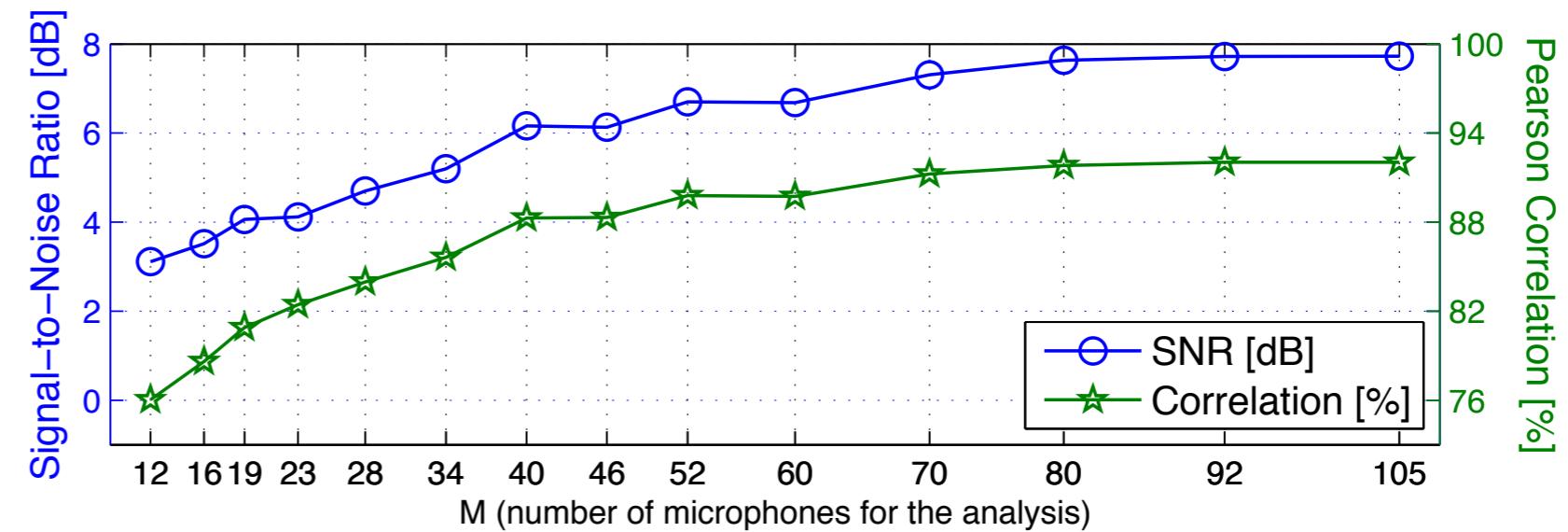
$F_s = 750 \text{ Hz}, f_c = 300 \text{ Hz}, K = 216 \text{ modes}, M = 104$



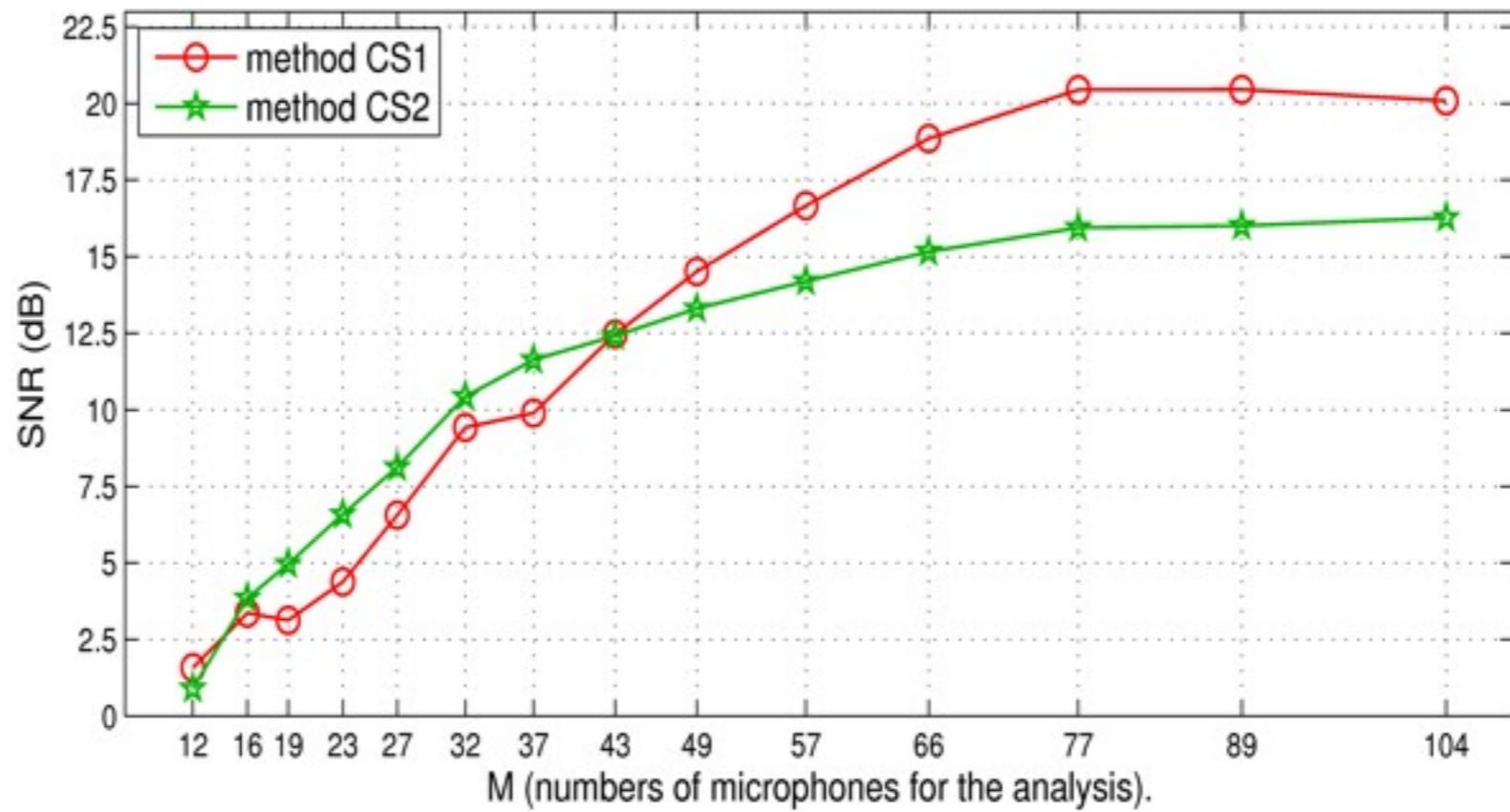
# Compressively sampling the plenacoustic function

- How many microphones ?

- sparsity  
in time

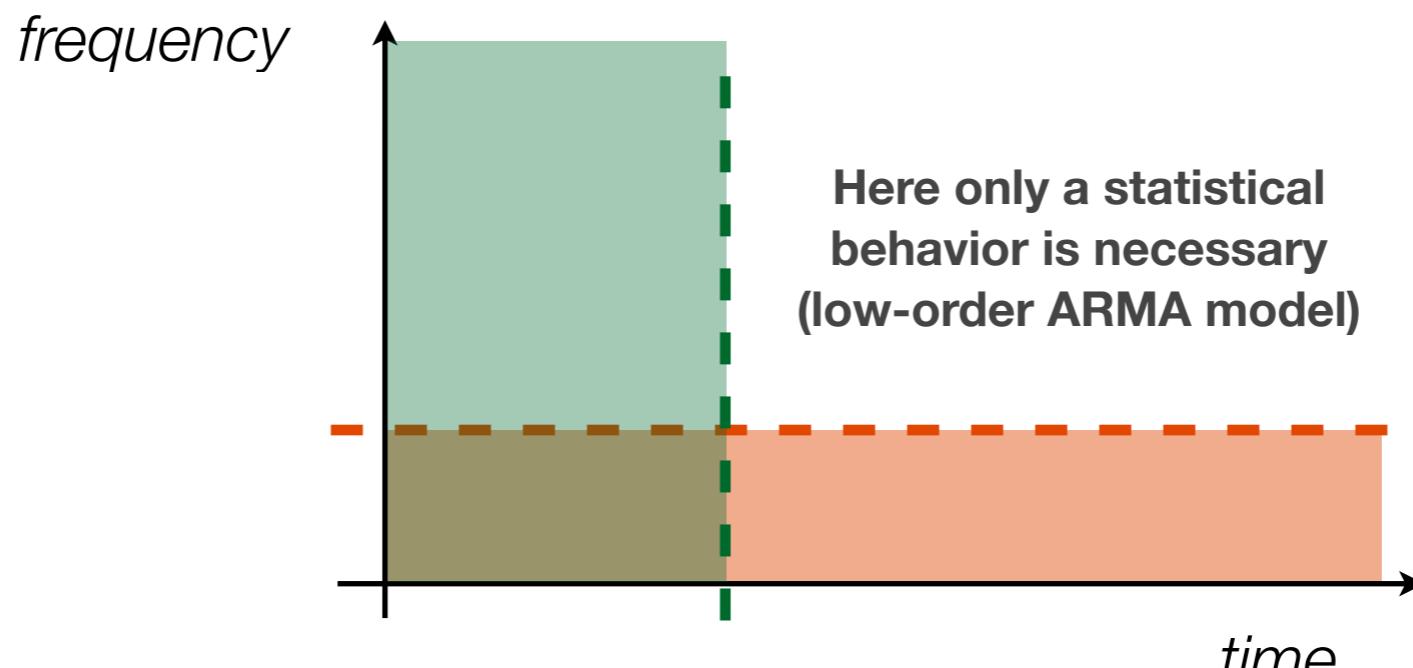


- sparsity in frequency



# Compressively sampling the plenacoustic function

- Yet to be done : fusion of the 2 approaches



- "How many microphones do we need to place in the room in order to completely reconstruct the sound field at any position in the room?"  
→ maybe not a crazy number !  
Sparsity makes it realistic to experimentally sample the plenacoustic function in a whole 3D volume.
- Applications : virtual acoustics, source localization, dereverberation ...
- Many open questions : calibration, algorithms for large data size, structured sparsity, dictionary design, optimal sampling distribution, ...

# Conclusion

---

# Conclusion

---

- Sparse wavefield models can be useful to spatially interpolate impulse responses of acoustic systems
- A model must be constructed according to the physics of the problem.
- 2D vibration of plates
  - unknown dispersion relation
  - evanescent waves
- 3D room acoustics
  - sparsity assumptions valid only at low freq. or early reflexions
  - hybrid models
- Still many open questions !

# References

---

- Mignot R., Daudet L. and Ollivier F., *Interpolation of room impulse responses in 3d using compressed sensing*, Proceedings of Acoustics'2012, Nantes (2012).
- Chardon G. and Daudet L., *Narrowband source localization in an unknown reverberant environment using wavefield sparse decomposition*, Proceedings of International Conference on Acoustics Speech and Signal Processing (ICASSP'2012), Kyoto (2012).
- Mignot R., Daudet L. and Ollivier F., *Compressed sensing for acoustic response reconstruction: interpolation of the early part*, Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, Mohonk, New-York (2011).
- Mignot R. and Daudet L., *Compressively sampling the plenacoustic function*, Proceedings of the SPIE XIVth Conference on Wavelets and Sparsity, San Diego, vol.8138-08, invited paper (2011).
- Chardon G., Leblanc A. and Daudet L., *Plate impulse response spatial interpolation with sub-Nyquist sampling*, Journal of Sound and Vibration, Vol 330 (23), pp. 5678-5689, (Nov. 2011).
- And other journal papers to come (please contact me)