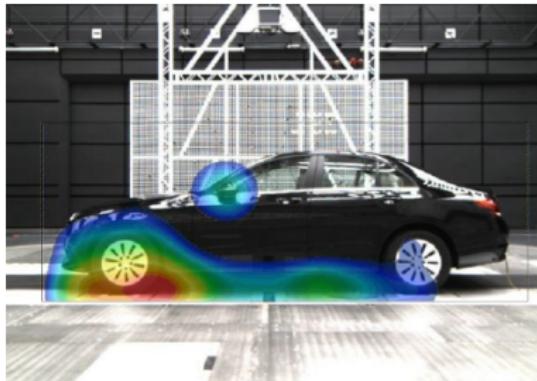


~ ECL Lecture - Microphone Array Techniques ~

Sound Source Localization - Beamforming



Simon Bouley - PhD, R & D Project Manager

simon.bouley@microdb.fr



www.microdb.vibratecgroup.com

MicrodB in a nutshell



Acoustic & Vibration
Measurement & Simulation

Since 1986



Sound source identification
Antennas & Software

Since 1994



Engineering service
technical assistance

Since 2005



MicrodB in a nutshell

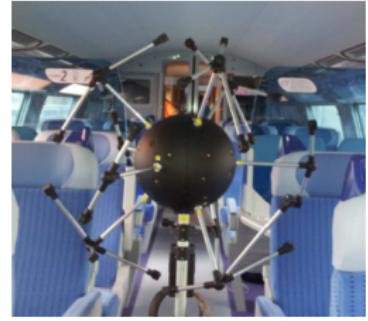
Industrial products for acoustic imaging



SoundCamera



SoundCamera Software



3DCam

MicrodB in a nutshell

From R & D to industrial tool in acoustic imaging



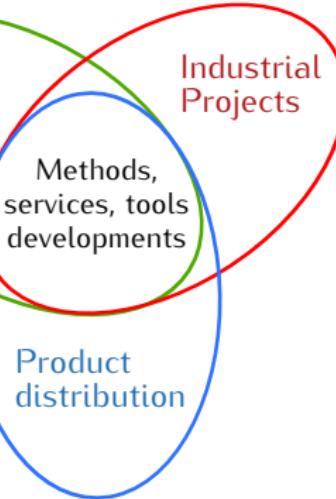
DGA



lva
laboratoire
vibrations
acoustique

LAUM

R & D
Projects



SIEMENS

NAVAL
GROUP

PSA
GROUPE



RENAULT

SAFRAN
AIRCRAFT ENGINES

AIRBUS

Objectives

This course is dedicated to Sound Source Localization and Quantification using array techniques :

Sound Source Localization (SSL) :

- Acoustic domain aiming at finding the most significative regions of noise emission in an industrial environnement
- The final goal is to reduce the noise emitted by a radiating object under study

Sound Source Quantification (SSQ) :

- Additional objective : define the absolute acoustic power of the sources identified in a localization step
- This aims at controlling that sub-elements comply to some acoustics specifications

How to achieve Sound Source Localization & Quantification ?

- Acoustic intensity technique : SSQ, standards
- Hemispheric power measurement : SSQ, standards
- Microphone array techniques : SSL & SSQ, no standard
 - Structure of many microphones in simultaneous recording
 - Signal processing techniques dedicated to draw *source maps*



Sound intensity probe



Hemispheric power measurement



Microphones array

Outline

1 Introduction

2 Industrial Needs

3 Overview of Sound Sources Localization Techniques

4 Signal Processing Prerequisites

5 Beamforming

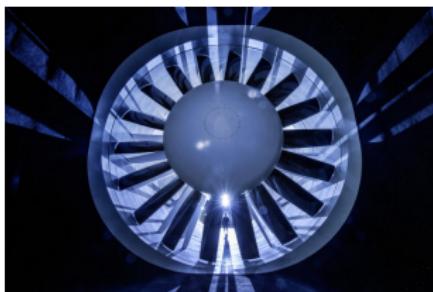
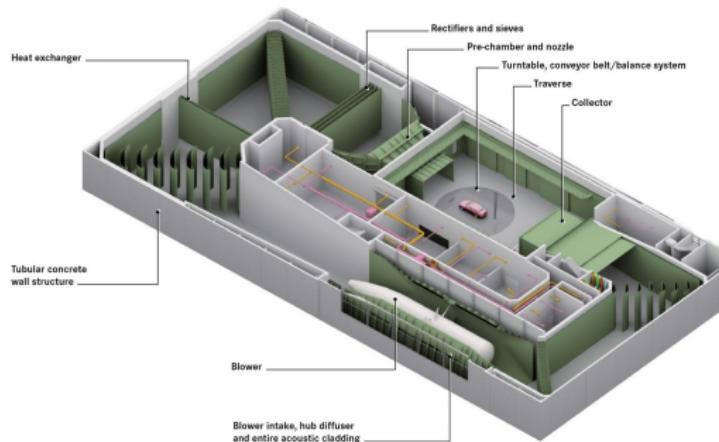
6 Conclusion

Industrial Needs

Automotive Wind Tunnel

Daimler Facility

- Study and reduce the noise of cars from interior & exterior points of view



Automotive Wind Tunnel

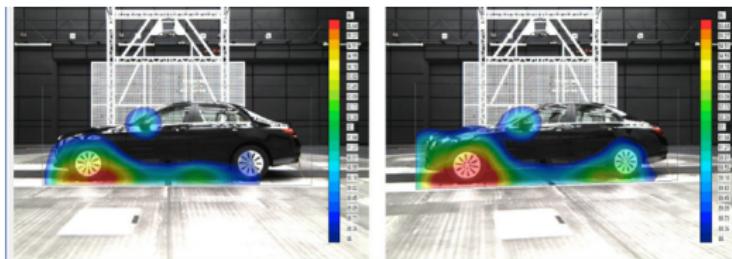
Array measurements and maps



Daimler aeroacoustic facility

Top and side arrays : ~ 360 channels

- Several configurations can be visually and quantitatively compared



Typical beamforming output maps

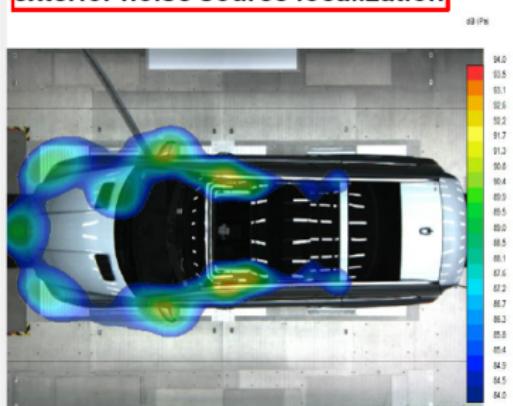
Automotive Wind Tunnel

Correlated maps

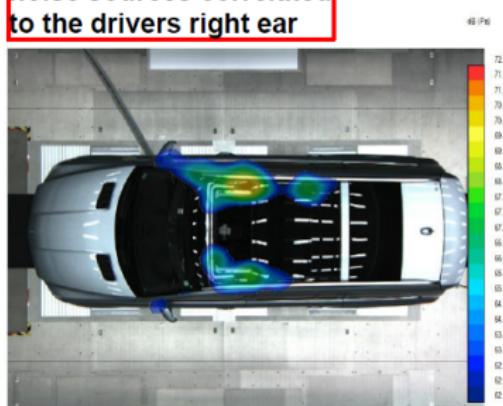
While arrays are located outside the vehicle, coherence treatments with binaural head microphones can provide information from the interior point of view



exterior noise source localization



noise sources correlated
to the drivers right ear



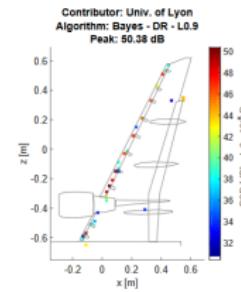
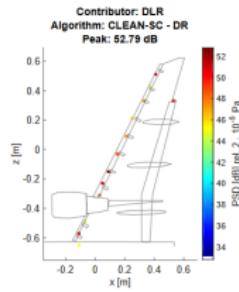
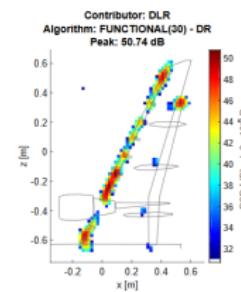
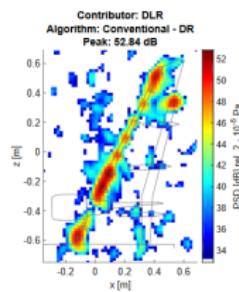
Airframe Noise

DLR Benchmark - Dornier-728 semispan model

Acoustic techniques comparison :



Wind-tunnel test section



PSD source maps at $f = 8496$ Hz

Bahr, C. J. et al., AIAA, 2017.

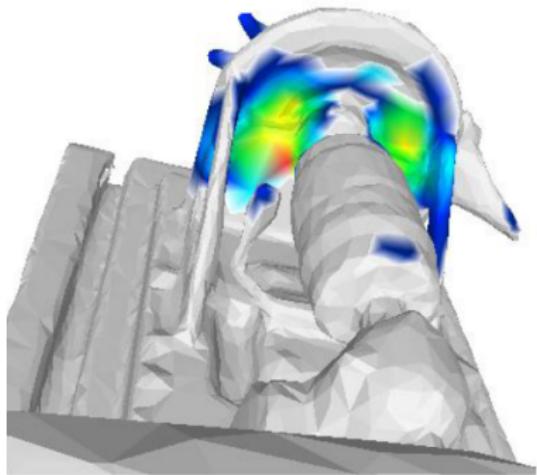
Three-Dimensional Inverse method

Renault Electric motor

Objectives : absolute acoustic sound power per component



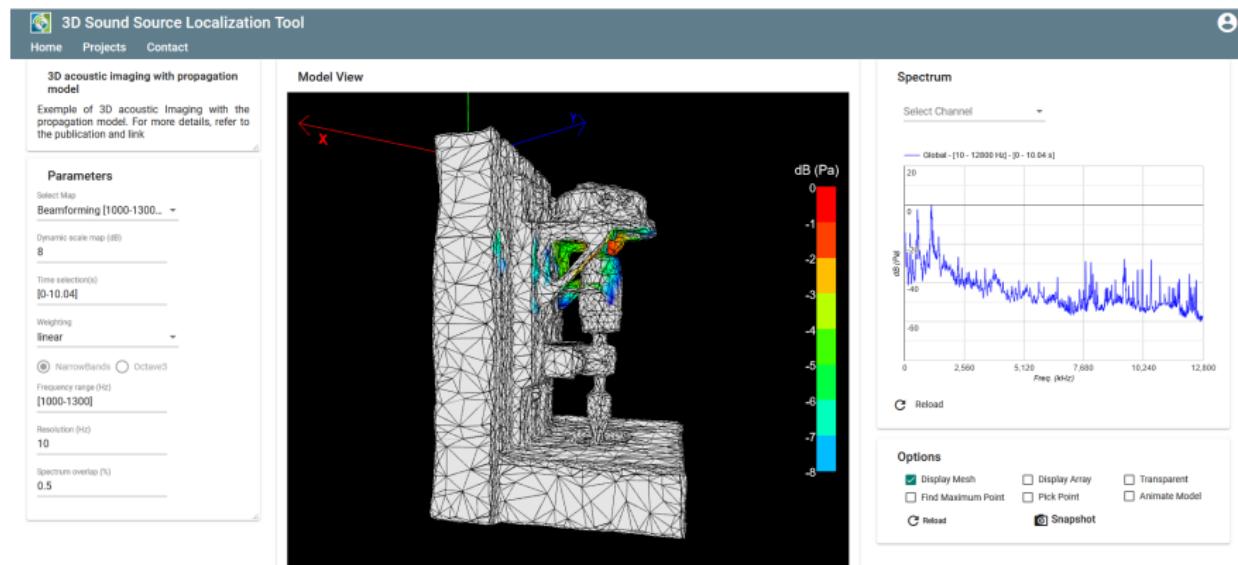
Electric motor set-up



PSD source maps at $f = 1$ kHz

Three-Dimensional Inverse method

Renault Electric motor



powered by MicrodB  and Rstone 

MicrodB 3D sound source localization web browser viewer

<https://saas.microdb.fr/lug/>

Underwater Acoustics

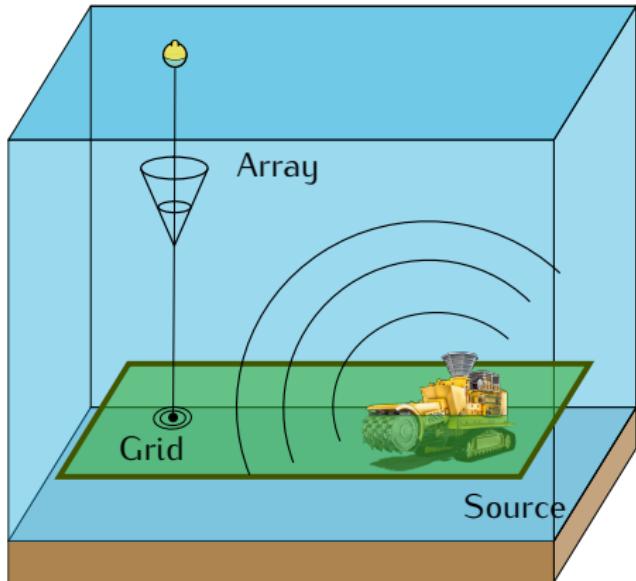
ABYSOUND Project

Abysound Project partners :

Naval-Group, IFREMER,
LMA, GIPSA-Lab, LSIS-Lab,
OSEAN, SEMANTIC TS, MicrodB

Objectives :

- ⇒ Acoustic monitoring of deep sea mining zones
- ⇒ Control and reduce their environmental impacts on biodiversity.



MicrodB and GIPSA-Lab :

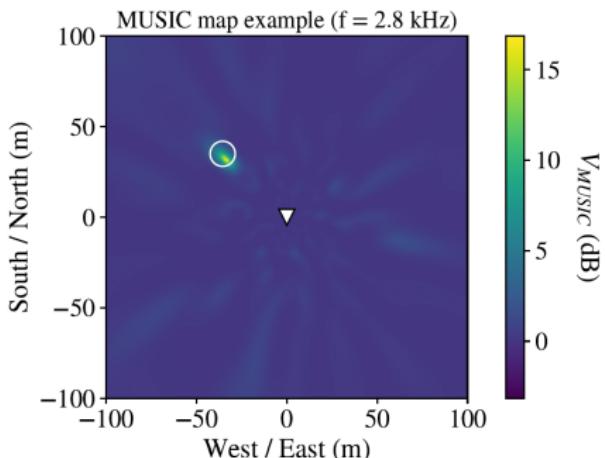
- Localization and quantification of acoustic sources

Underwater Acoustics

ABYSOUND Project



Hydrophone array

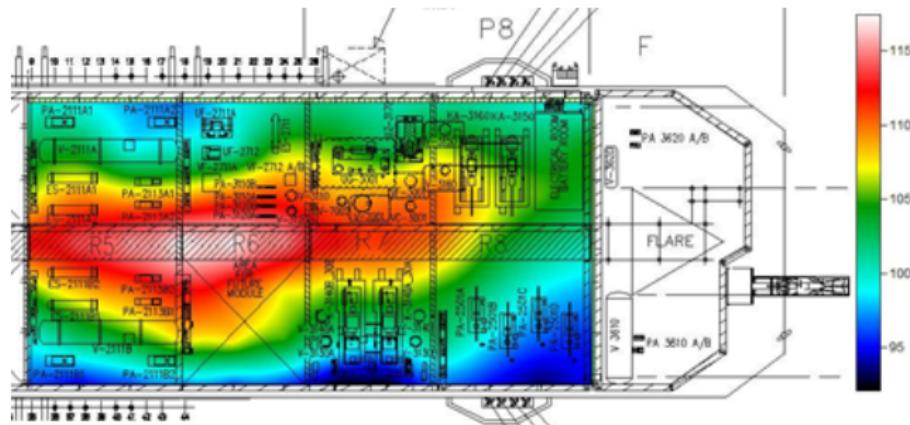


Localization map

Baron et al., IEEE JOE, submitted.

Overview of Sound Sources Localization Techniques

Acoustic measurement

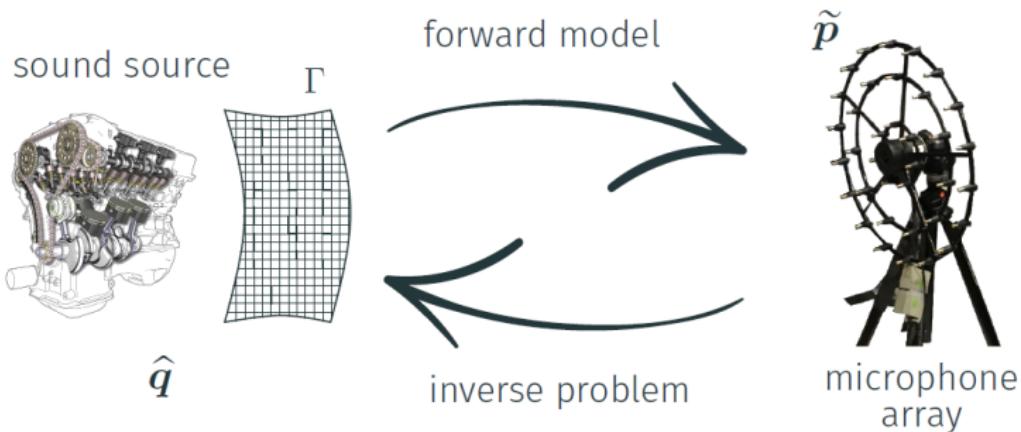


Acoustic map of a factory

The acoustic field is measured with a microphone moved around the radiating device.

- allows to measure directivity of a device
 - does not provide sound localization and quantification on the device

Acoustic Imaging



All localization methods require :

- a microphone array
- a source source grid on the radiating device
- a forward model
- a back-propagation method

Classical forward model

Monopole radiation

$$G(\mathbf{r}) = A \frac{\exp(i2\pi k\mathbf{r})}{4\pi r}$$

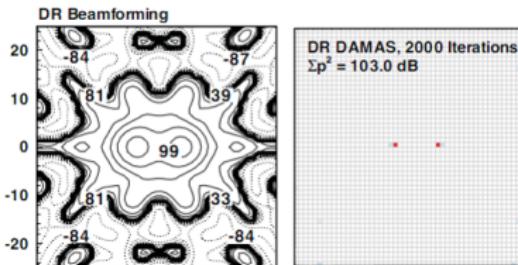
Acoustic Imaging

What distinguishes these methods ?

- the inverse problem technique resolution (beamforming, inverse method, iterative technique)
- rarely, the forward model

The performances depend on the method choice :

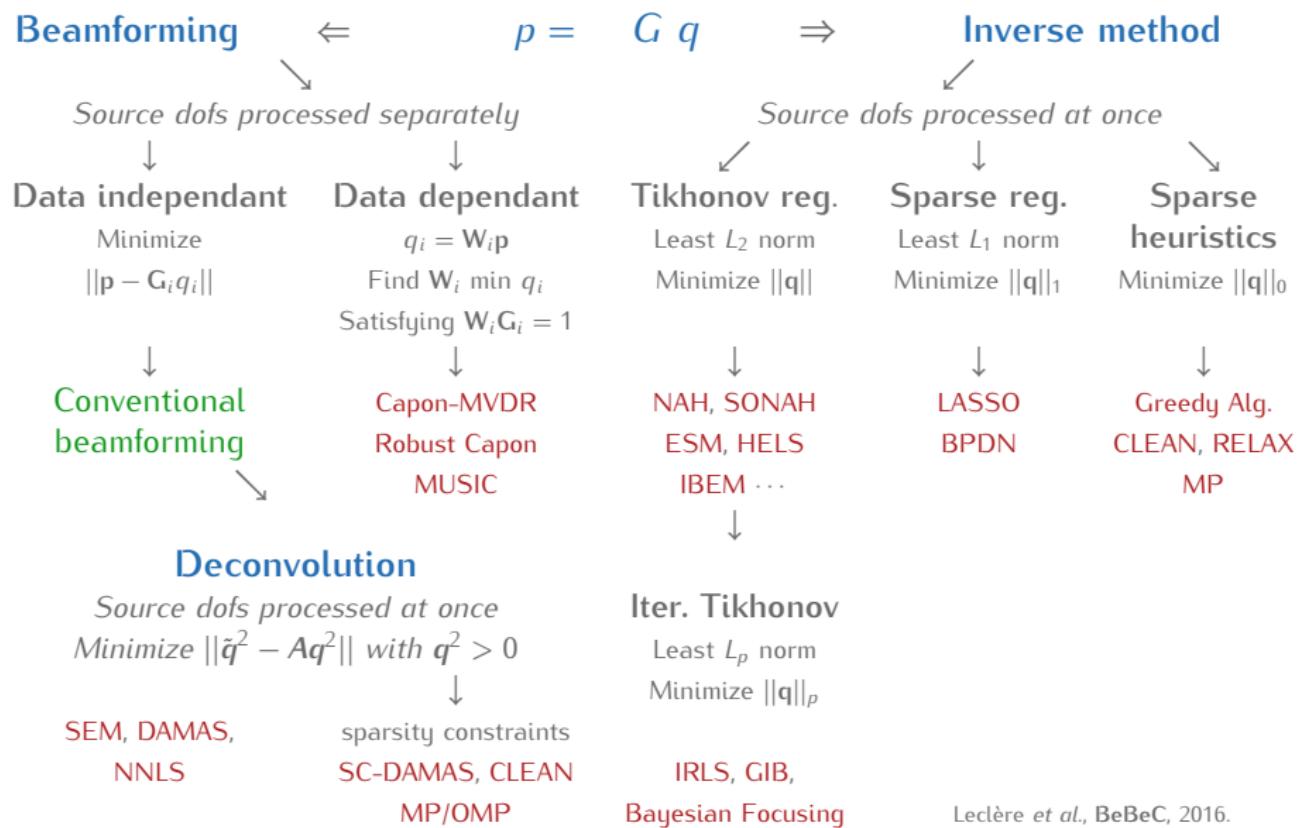
- Computation time (from one second to several hours for one frequency)
- Spatial resolution
- Source quantification accuracy
- Model bias sensitivity



Beamforming is the simplest, the most widespread and one of the most robust method for source localization.

Brooks & Humphrey, AIAA, 2006.

Tentative Summary



Beamforming

Advantages :

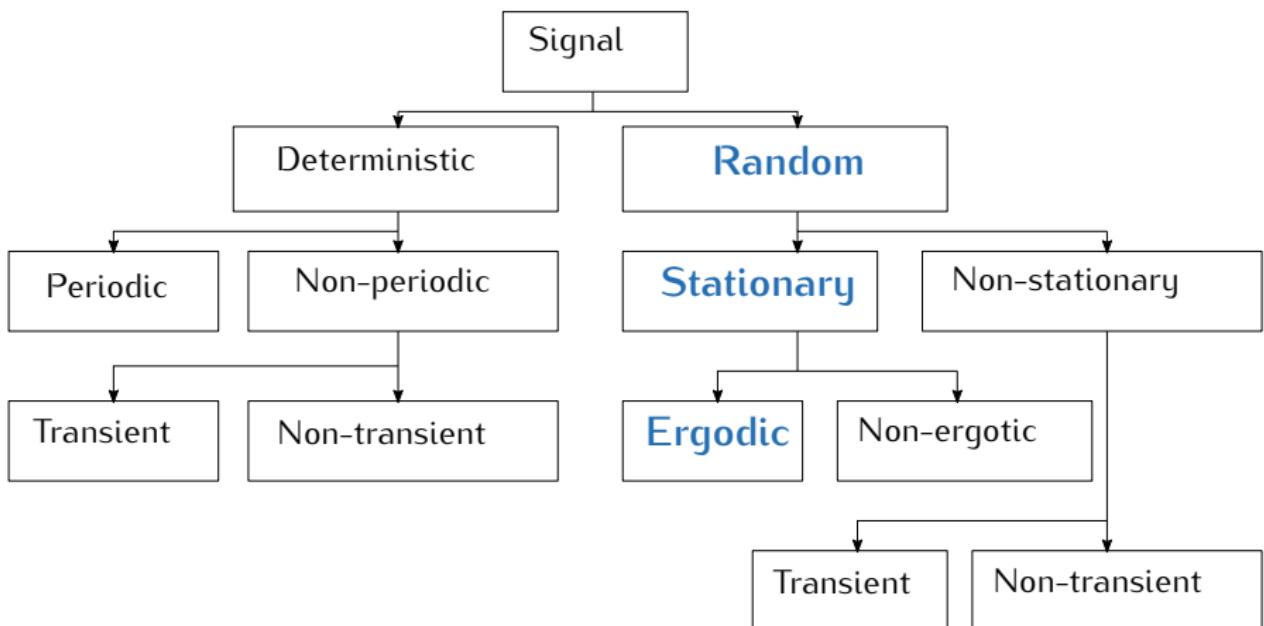
- Versatility : for any geometry of both the array and the calculation grid
- Time or frequency processing
- Robustness
- Implementation simplicity

Drawbacks :

- Non-quantitative
- Inaccurate resolution in low frequencies

Signal Processing Prerequisites

Signal Classification



Bendat & Piersol, Random Data, 2010.

Basic definitions

A random signal represents one of the many possible observations of a random physical phenomenon that might have occurred.

⇒ A time record of a random process cannot be analytically described.

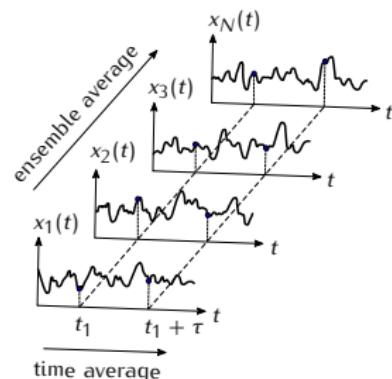
An **ensemble** of time records can be studied through its statistical properties :

- mean value at time t : $\mu_x(t)$
- autocorrelation at two time $t + \tau$: $R_{xx}(t, t + \tau)$

If $\mu_x(t) = \mu_x$ and $R_{xx}(t, t + \tau) = R_{xx}(\tau)$, the random process is *weakly stationary*.

A **stationary** random process is **ergodic** if the **ensemble** and the **time** averages produce the same results.

One can draw statistics over time from one signal instead of repeating N times the random process.



Spectral Density Functions

Fundamental tools in signal processing are the spectral density functions :

- Power Spectral Density : $S_{xx}(\omega) \in \mathbb{R}$
- Cross Power Spectral Density : $S_{xy}(\omega) \in \mathbb{C}$

Wiener–Khintchin Theorem

$$S_{xx}(\omega) = \text{FT}\{R_{xx}(\tau)\}, \quad S_{xy}(\omega) = \text{FT}\{R_{xy}(\tau)\}, \quad \text{for ergodic stationary signals}$$

- $|S_{xy}(\omega)|$: coherent energy spread over signals $x(t)$ and $y(t)$ at frequency ω
- $\varphi(S_{xy}(\omega))$: phase relationship between the coherent part of signals $x(t)$ and $y(t)$ at frequency ω

The cross power spectral density (cross-spectrum) reveals information on the acoustic field acquired **coherently** by distinct microphones.

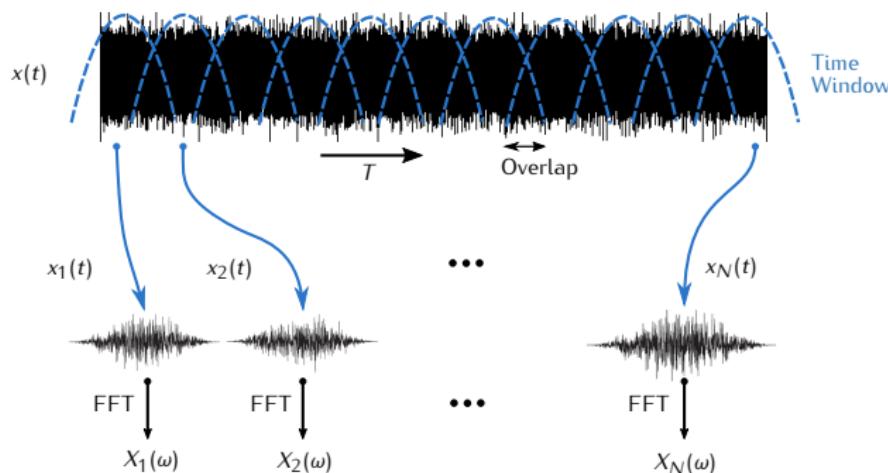
$S_{xy}(\omega)$ depends on many factors : source and microphone positions, source strength...

(C)PSD Estimation

Welch's Method

As signals are time finite and noisy, only a PSD estimation is obtainable. A usual method is the Welch process to compute a periodogram :

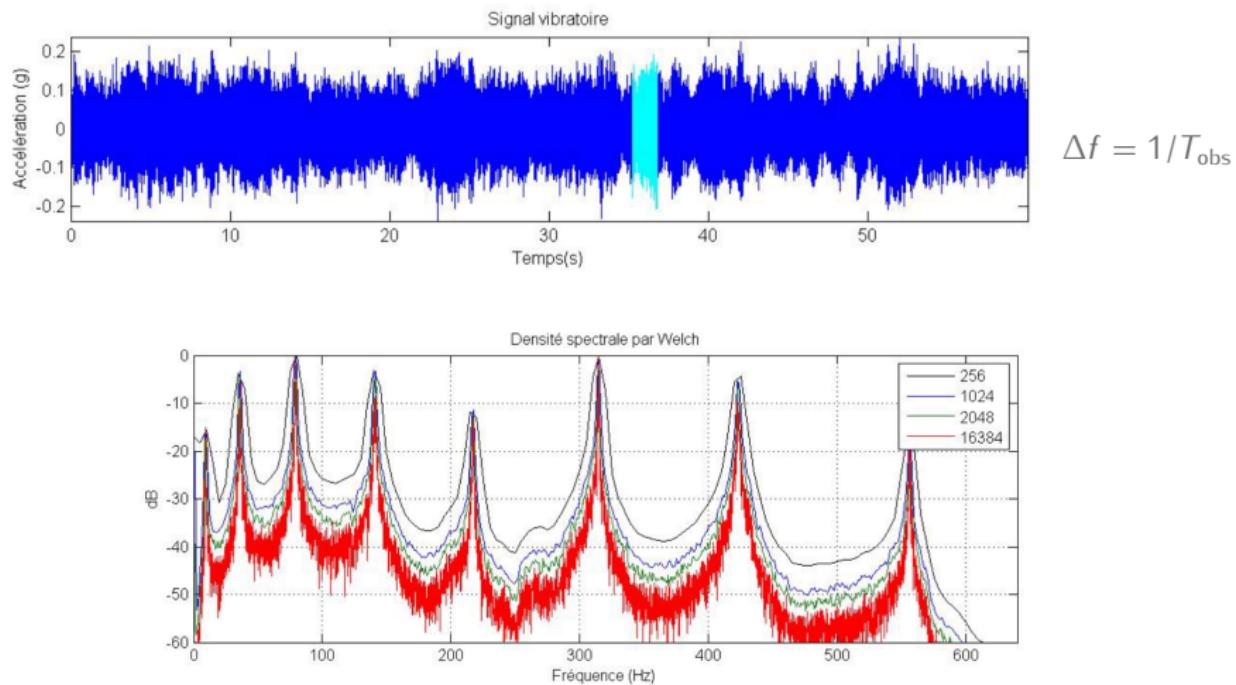
$$\hat{S}_{xx}(\omega) = \langle X(\omega)X^H(\omega) \rangle = \langle |X(\omega)|^2 \rangle, \quad \hat{S}_{xy}(\omega) = \langle X(\omega)Y^H(\omega) \rangle.$$



$$\hat{S}_{xx}(\omega) = \frac{1}{N} \sum_{i=1}^N |X_i(\omega)|^2.$$

(C)PSD Estimation

Time block size influence

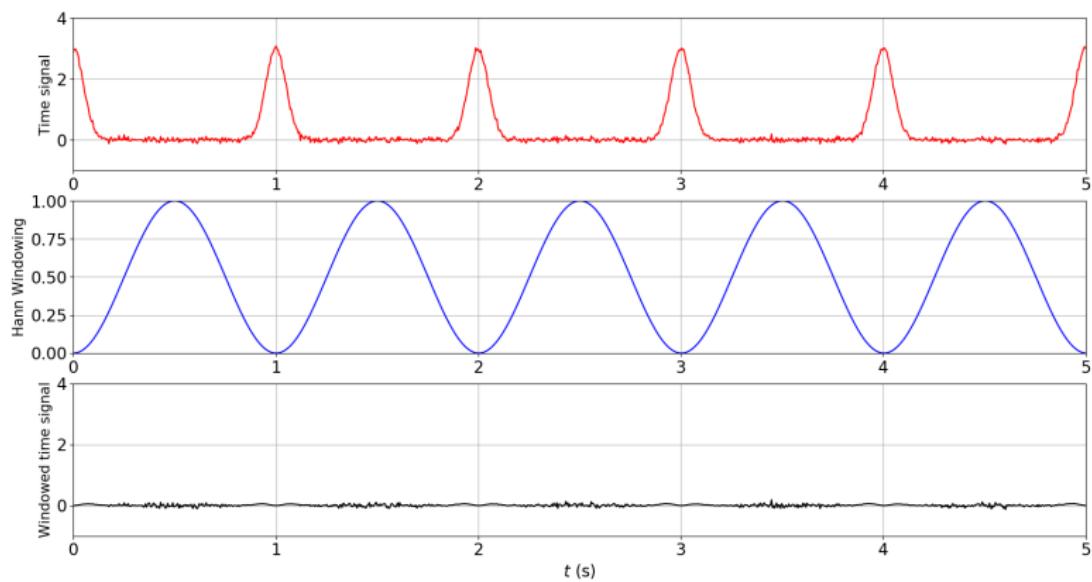


A well-adapted time block size reduces the PSD variance without too much over-estimation.

Ollivier, Analyse de signaux aléatoires, 2014.

(C)PSD Estimation

Overlap influence

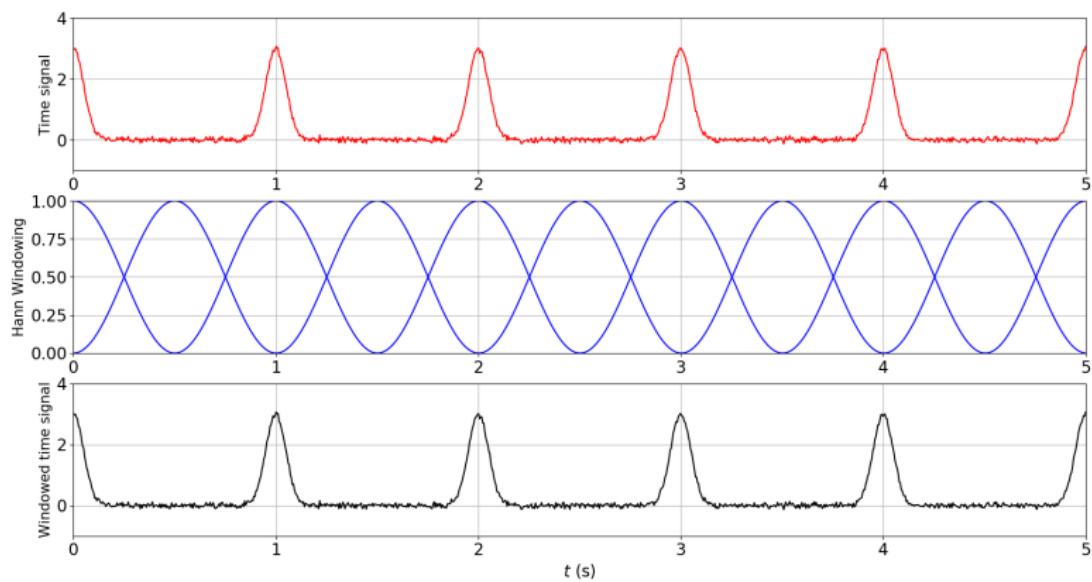


Time signal, Hann time windowing, Windowed time signal

Necessity to overlap the snapshots to avoid information loss.

(C)PSD Estimation

Overlap influence



Time signal, Hann time windowing, Windowed time signal

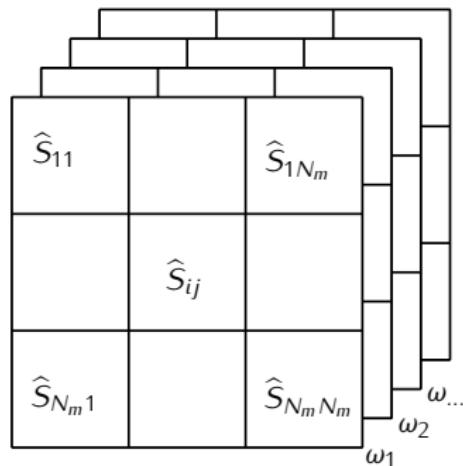
With a well-designed overlap (here 50 %), all the events of the time signal are considered.

(C)PSD Estimation

For the cross power spectral density :

$$\hat{S}_{xy}(\omega) = \frac{1}{N} \sum_{i=1}^N X_i(\omega) Y_i^H(\omega).$$

All the information used in most array techniques is carried by a **Cross Spectral Matrix** (CSM), noted S_{pp} , which contains the measured data.



Array processing's goal :

⇒ recover the sources positions from S_{pp} and the geometrical configuration.

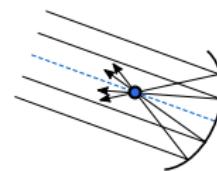
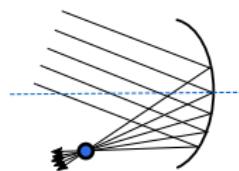
Beamforming

Beamforming

Basic Principle



Acoustic mirror - Royal Air Force, Denge, UK



As for radio waves, parabolic antennas were dedicated to hear environmental noise. The device is either steered in a certain direction or the microphone is moved to the focus point. A **beam is formed** in this direction.

Time Formulation

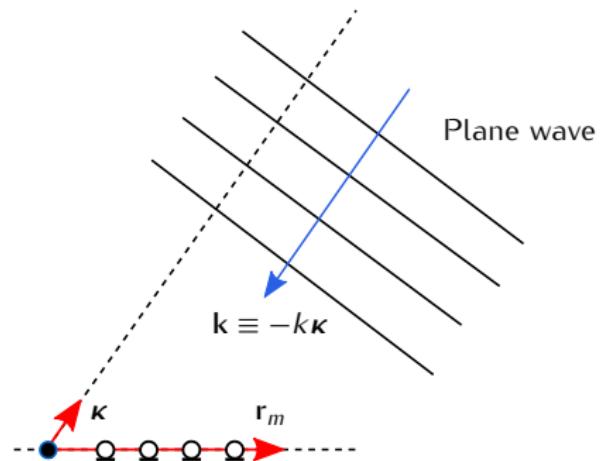
Delay and Sum Beamformer

$$p_m(t) = p_0(t - \Delta_m(\kappa_0))$$

For a focus angle κ :

- 1 Compensate the **delays** measured by the microphones due to the focus direction angle κ
- 2 **Sum** the delayed microphone signals

$$b(\kappa, t) = \frac{1}{N_m} \sum_{m=1}^{N_m} p_m(t + \Delta_m(\kappa)),$$



$$b(\kappa, t) = \frac{1}{N_m} \sum_{m=1}^{N_m} p_0 \left(t - \frac{(\kappa_0 - \kappa) \cdot r_m}{c_0} \right)$$

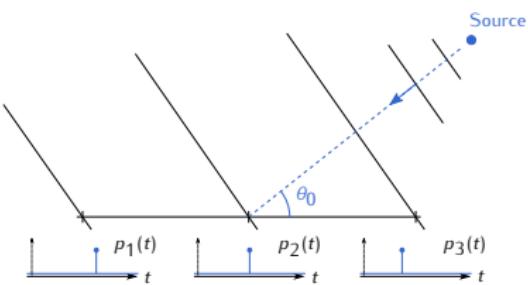
$$\Delta_m(\kappa) = \frac{\kappa \cdot r_m}{c_0}$$

The beamformer $b(\kappa, t)$ seeks the direction where its value is the highest, when the delayed pressures are summed up *coherently* ($\kappa = \kappa_0$).

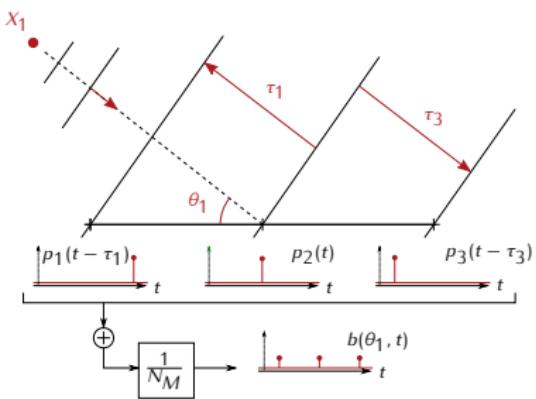
Time Formulation

Delay and Sum Beamformer

For a ground truth angle θ_0 :



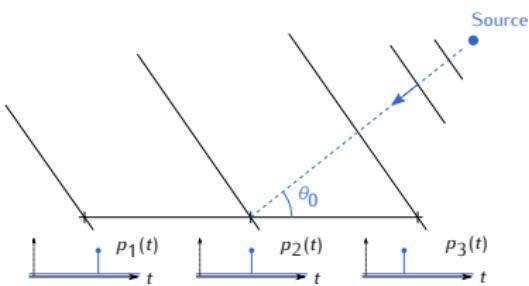
Test a focus angle θ_1 :



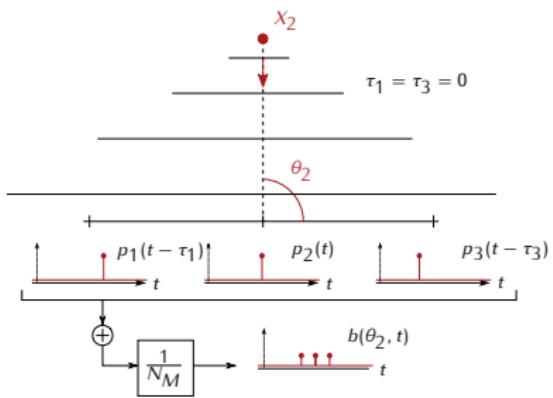
Time Formulation

Delay and Sum Beamformer

For a ground truth angle θ_0 :



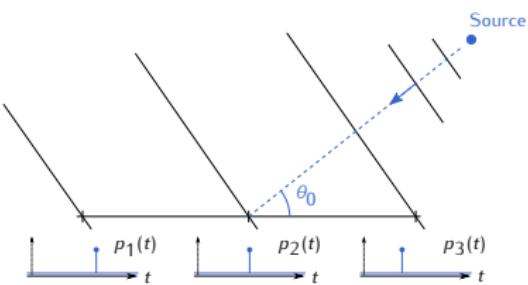
Test a focus angle θ_2 :



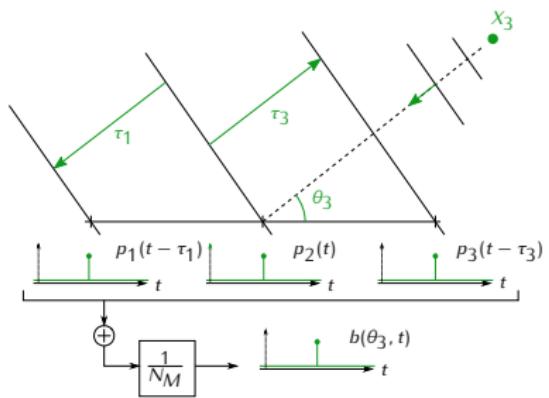
Time Formulation

Delay and Sum Beamformer

For a ground truth angle θ_0 :



Test a focus angle θ_3 :



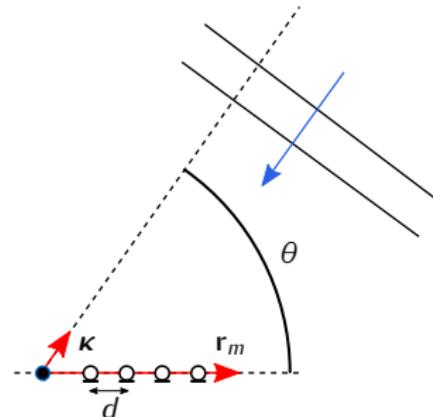
Time Formulation

Basic propagation model

Far-field / Near-field expressions

Far-Field (Fraunhofer region)

$$\Delta_m = \frac{md}{c_0} \cos \theta_m$$

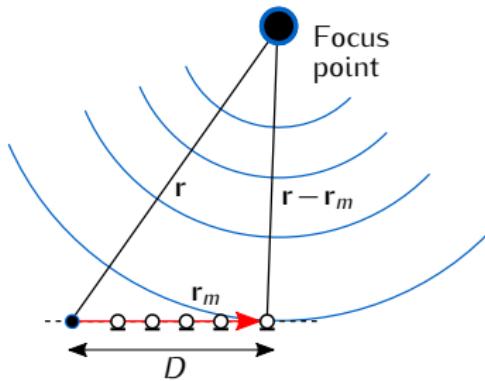


Near-Field (Fresnel region)

$$\Delta_m = \frac{|r| - r_m(r)}{c_0}$$

Fresnel region :

$$r < \frac{D^2}{\lambda}$$



Conventional Beamforming

Frequency Formulation

Source Location

$\mathbf{P} = [P_1(f), P_i(f), \dots, P_{N_m}(f)]$ is the vector of complex pressure values measured by microphones, indexed i at frequency f for a given snapshot.

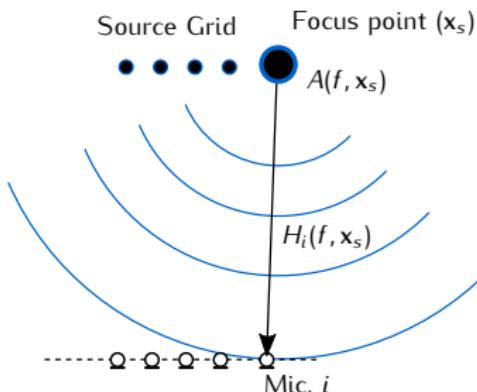
Suppose a single source with complex amplitude $A(f)$ for this snapshot and a known transfer function $H_i(f, x_s)$.

$$\text{Direct propagation : } \mathbf{P} = A\mathbf{H}$$

For a fixed frequency f , we seek the source point x_s so that $\mathbf{P}(f)$ and $\mathbf{H}(f, x_s)$ are proportional, i.e. co-linear, done by computing the normalized cross-product between the two (complex) vectors :

$$\gamma^2(x_s) = \frac{|\mathbf{H}^H(x_s)\mathbf{P}|^2}{||\mathbf{H}(x_s)||^2||\mathbf{P}||^2}, \quad \text{coherence indicator.}$$

Elias, VKI, 1997.



Conventional Beamforming

Source Location

Coherence Indicator for random sources

For random sources, the pressure is averaged on the N snapshots to drop incoherent noise out.

$$\langle |\mathbf{H}^H \mathbf{P}|^2 \rangle = \left\langle \sum_{i,j=1}^{N_m} H_i^H P_i P_j^H H_j \right\rangle = \sum_{i,j=1}^{N_m} H_i^H \langle P_i P_j^H \rangle H_j = \sum_{i,j=1}^{N_m} H_i^H \hat{\mathbf{S}}_{ij} H_j = \mathbf{H}^H \hat{\mathbf{S}}_{pp} \mathbf{H}$$

$$\langle \|\mathbf{P}\|^2 \rangle = \left\langle \sum_{i=1}^{N_m} |P_i|^2 \right\rangle = \sum_{i=1}^{N_m} \langle |P_i|^2 \rangle = \sum_{i=1}^{N_m} \hat{S}_{ii}^2 = \text{Tr}(\hat{\mathbf{S}}_{pp}), \quad \text{Tr : matrix trace}$$

The localization function for random sources reads :

$$\boxed{\gamma^2(\mathbf{x}_s) = \frac{\mathbf{H}^H(\mathbf{x}_s) \hat{\mathbf{S}}_{pp} \mathbf{H}(\mathbf{x}_s)}{\|\mathbf{H}(\mathbf{x}_s)\|^2 \text{Tr}(\hat{\mathbf{S}}_{pp})}, \quad 0 \leq \gamma^2 \leq 1.}$$

Conventional Beamforming

Source Amplitude

The objective of finding the acoustic source location and its amplitude is a simple minimization problem :

For a given measurement vector \mathbf{P} and a known library of transfer vectors \mathbf{H} , find A which minimize the quadratic error :

$$\varepsilon(A) = \|\mathbf{P} - A\mathbf{H}\|^2.$$

Solution :

$$\varepsilon(A) = (\mathbf{P} - A\mathbf{H})^H(\mathbf{P} - A\mathbf{H}) = \mathbf{P}^H\mathbf{P} - A\mathbf{P}^H\mathbf{H} - A^H\mathbf{H}^H\mathbf{P} + A^H A \mathbf{H}^H \mathbf{H}.$$

We are free to use A and A^H as independent variables. For a fixed value of \mathbf{x}_s :

$$\frac{\partial \varepsilon(\hat{A})}{\partial A^H} = 0 \Leftrightarrow -\mathbf{H}^H\mathbf{P} + \hat{A}\mathbf{H}^H\mathbf{H} = 0,$$

$$\hat{A}(\mathbf{x}_s) = \frac{\mathbf{H}^H(\mathbf{x}_s)\mathbf{P}}{\|\mathbf{H}(\mathbf{x}_s)\|^2}$$

Conventional Beamforming

Source Amplitude

With this amplitude, the value of the residue ε is :

$$\varepsilon(A) = \|\mathbf{P}\|^2 - \frac{|\mathbf{H}^H(\mathbf{x}_s)\mathbf{P}|^2}{\|\mathbf{H}(\mathbf{x}_s)\|^2}$$

The source location is then where ε is minimum, or for which

$$\gamma^2(\mathbf{x}_s) = 1 - \frac{\varepsilon}{\|\mathbf{P}\|^2}, \quad \text{is maximum.}$$

For random stationnary sources, the mean squared amplitude reads :

$$\sigma^2(\mathbf{x}_s) = < |\hat{A}(\mathbf{x}_s)|^2 > = \frac{\mathbf{H}^H(\mathbf{x}_s)\hat{S}_{pp}\mathbf{H}(\mathbf{x}_s)}{\|\mathbf{H}(\mathbf{x}_s)\|^4}$$

Beamforming is basically a vector-matrix-vector product between :

- The measured data cross spectral matrix \hat{S}_{pp}
- The model transfer vector \mathbf{H}

The principle of Delay and Sum is recovered in $\mathbf{H}^H(\mathbf{x}_s)\mathbf{P}$ through the conjugate operator (**time delays**) and the vector–vector product (**sum**).

Conventional Beamforming

Summary

To do it properly, the SSL and SSQ processes should be done in two steps :

- Localize the source : find the maximum position x_{\max} of the coherence indicator map :

$$\gamma^2(x_s) = \frac{\mathbf{H}^H(x_s) \hat{\mathbf{S}}_{pp} \mathbf{H}(x_s)}{\|\mathbf{H}(x_s)\|^2 \text{Tr}(\hat{\mathbf{S}}_{pp})}, \quad 0 \leq \gamma^2 \leq 1.$$

- Estimate the source amplitude in x_{\max} :

$$\sigma^2(x_s) = \frac{\mathbf{H}^H(x_s) \hat{\mathbf{S}}_{pp} \mathbf{H}(x_s)}{\|\mathbf{H}(x_s)\|^4}.$$

In practice, software and acoustic teams often prefer to look at a single map : $\sigma^2(x_s)$.

Linear array performances for far-field acoustic wave

2D Configuration

Here a linear array includes N microphones over a length L .

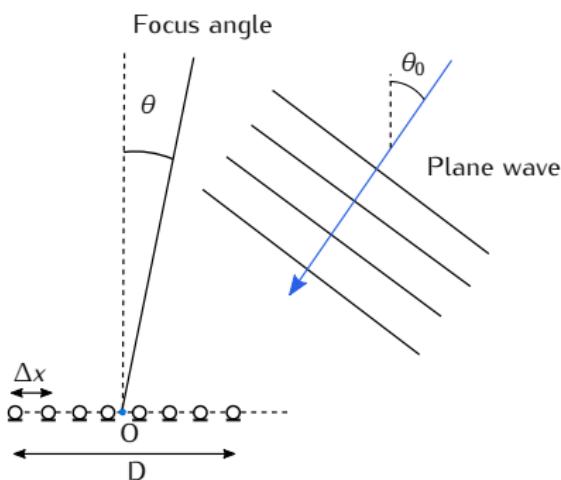
The pressure field due to a plane wave, measured by the microphone n is :

$$p_n = \exp(i k x_n \sin \theta_0).$$

$$H_n(\theta) = \exp(i k x_n \sin \theta).$$

$$Y^2(\theta_0, \theta) = \frac{|\mathbf{H}^H(\theta)\mathbf{P}(\theta_0)|^2}{\|\mathbf{H}(\theta)\|^2 \|\mathbf{P}(\theta_0)\|^2} = \left(\frac{\sin z}{N \sin(z/N)} \right)^2,$$

$$\text{with } z = \pi \frac{L}{\lambda} \frac{N}{N-1} (\sin \theta - \sin \theta_0).$$



Jupyter Notebook :

[Link](#)

Linear array performances for far-field acoustic wave

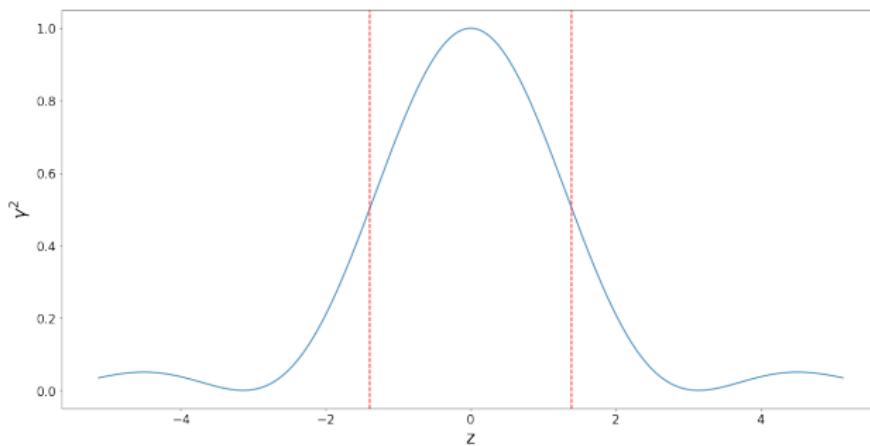
Resolution

Parameters

$L = 1 \text{ m}$

$N = 10 \text{ microphones}$

$f = 500 \text{ Hz}$



Resolution = ability to discriminate two close sources
= overall width of the main lobe 3 dB (50 %) under the maximum.

Linear array performances for far-field acoustic wave

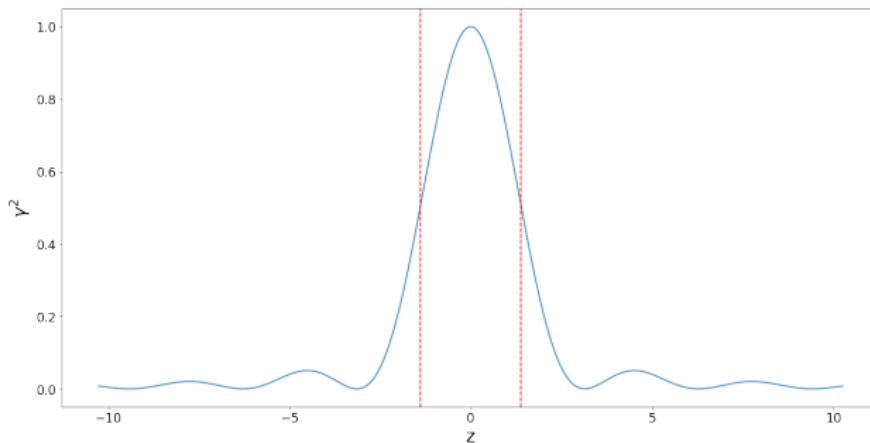
Resolution

Parameters

$L = 1 \text{ m}$

$N = 10 \text{ microphones}$

$f = 1000 \text{ Hz}$



$$\Delta\theta_{3\text{dB}} \simeq 0.9 \frac{\lambda}{L \cos \theta_0}$$

The resolution is inversely proportional of the frequency

Linear array performances for far-field acoustic wave

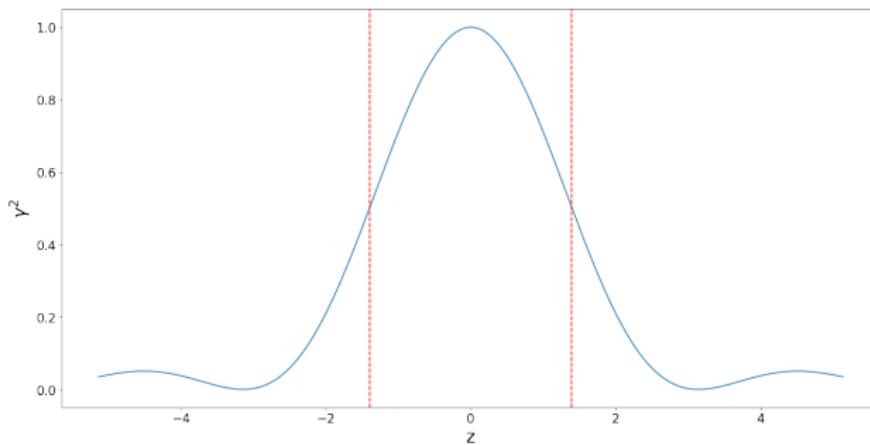
Resolution

Parameters

$L = 1 \text{ m}$

$N = 10 \text{ microphones}$

$f = 500 \text{ Hz}$



Resolution = ability to discriminate two close sources
= overall width of the main lobe 3 dB (50 %) under the maximum.

Linear array performances for far-field acoustic wave

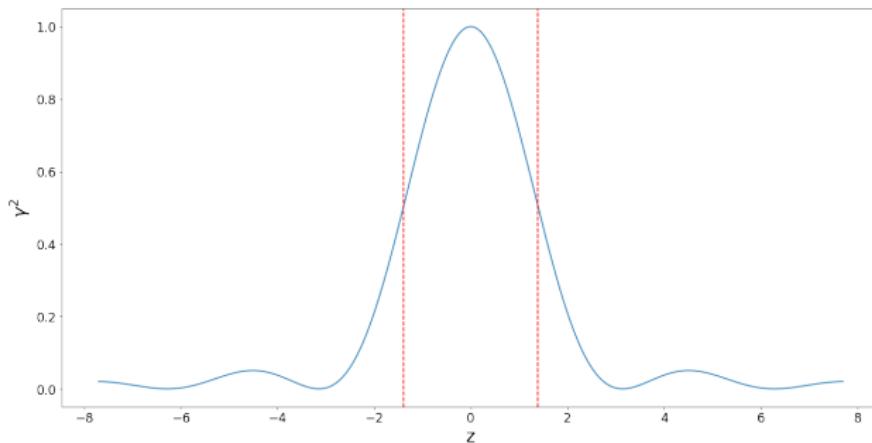
Resolution

Parameters

$L = 1.5 \text{ m}$

$N = 10 \text{ microphones}$

$f = 500 \text{ Hz}$



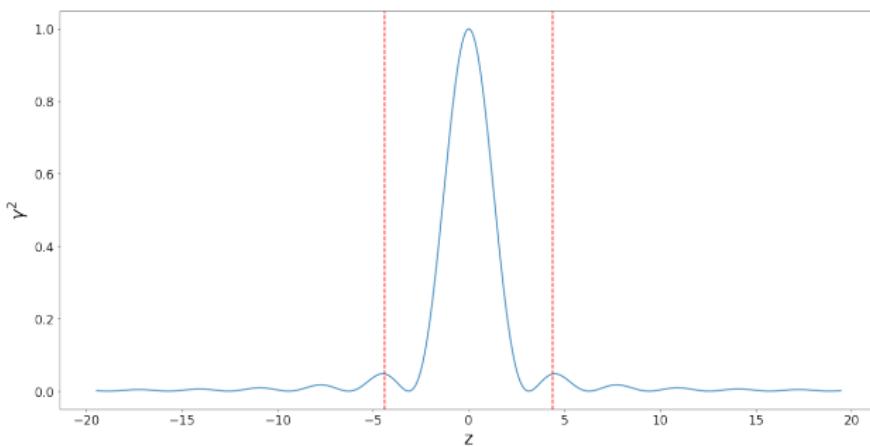
$$\Delta\theta_{3\text{dB}} \simeq 0.9 \frac{\lambda}{L \cos \theta_0}$$

Wide array improves the spatial resolution

Linear array performances for far-field acoustic wave Dynamics

$\gamma^2(z)$ has side lobes, with a maximum around $z = \frac{3\pi}{2}; \frac{5\pi}{2}; \dots$

If a second source of small amplitude is present, it may be hidden by the side lobes of the first one.



Dynamics = Difference between main lobe and the higher side lobe amplitudes

Linear array performances for far-field acoustic wave

Spatial aliasing

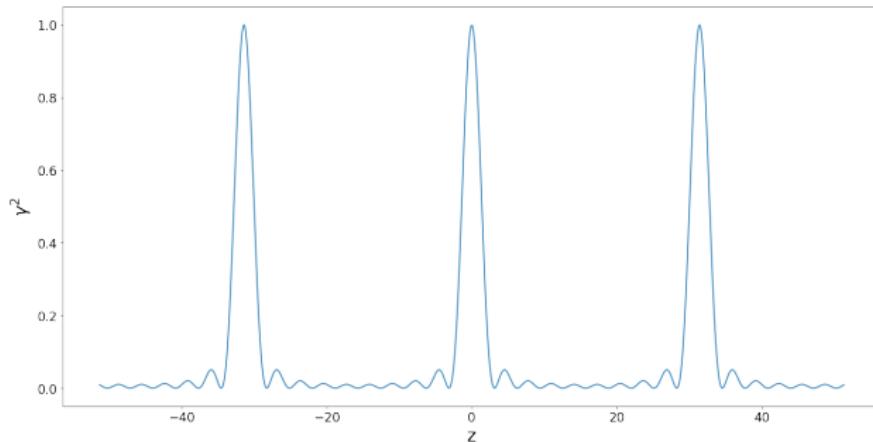
Due to its periodicity, $\gamma^2(z)$ is equal to unity for $\sin \theta = \sin \theta_0 + m \frac{\lambda}{\Delta x}$.

Parameters

$$L = 1 \text{ m}$$

$$N = 10 \text{ microphones}$$

$$f = 5000 \text{ Hz}$$



Ambiguous localisation, the same pressure is measured by all microphones

Linear array performances for far-field acoustic wave

Spatial aliasing

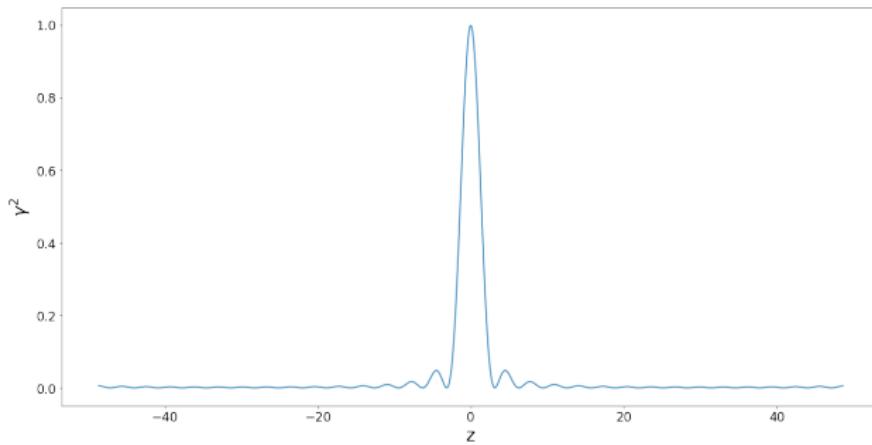
Due to its periodicity, $\gamma^2(z)$ is equal to unity for $\sin \theta = \sin \theta_0 + m \frac{\lambda}{\Delta x}$.

Parameters

$$L = 1 \text{ m}$$

$$N = 20 \text{ microphones}$$

$$f = 5000 \text{ Hz}$$



$$|\sin \theta_0 + m \frac{\lambda}{\Delta x}| > 1, \quad \Delta x < \frac{\lambda}{2}.$$

Shannon theorem for spatial sampling

2D Beamforming

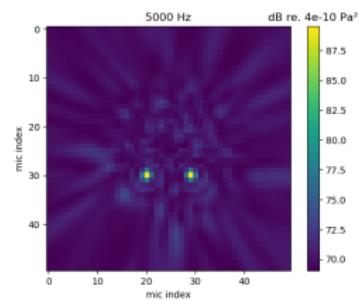
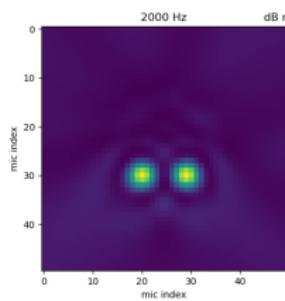
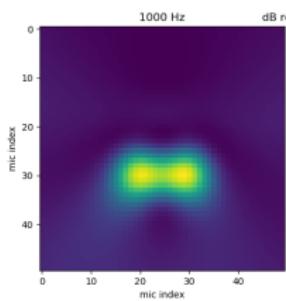
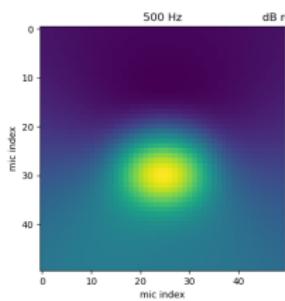
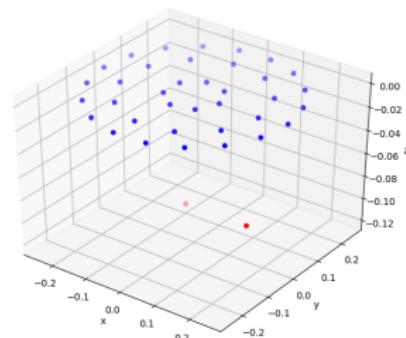
Resolution & Dynamics performances

Circle-Plane array

D = 25 cm

N = 36 microphones

2 sources of same amplitude



2D Beamforming

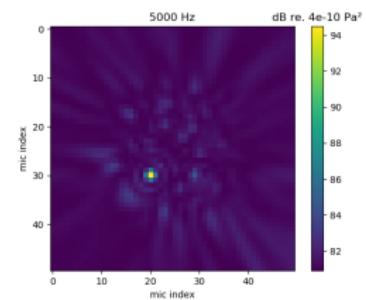
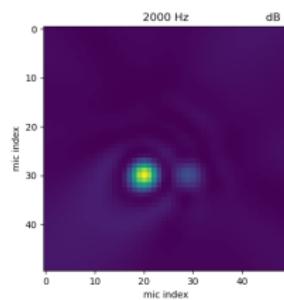
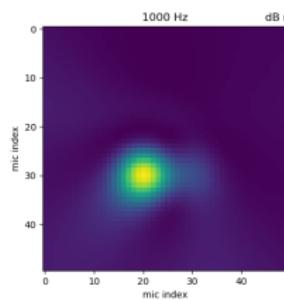
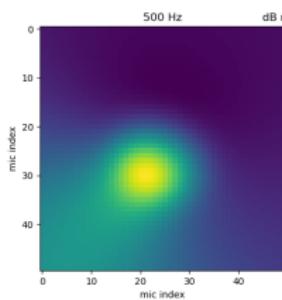
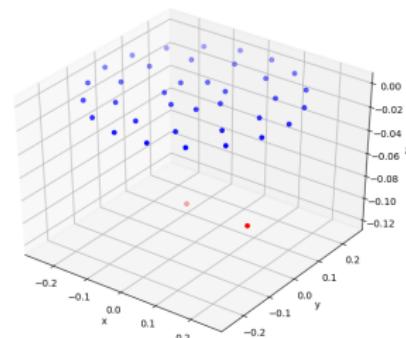
Resolution & Dynamics performances

Circle-Plane array

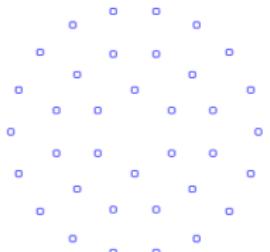
D = 25 cm

N = 36 microphones

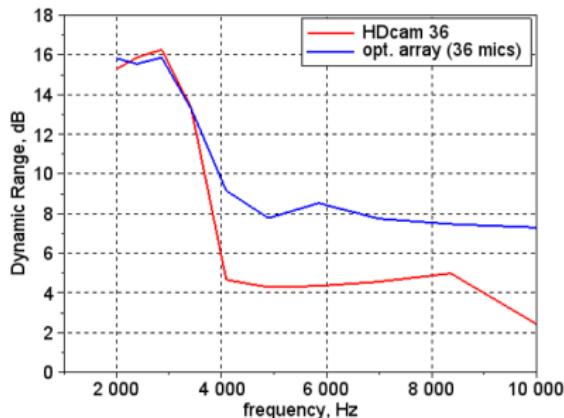
2 sources of unequal amplitudes



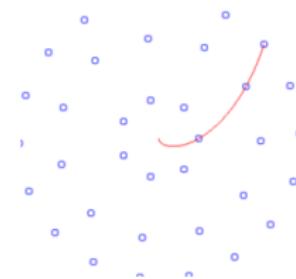
Array Optimization



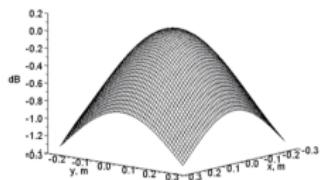
HDcam 36



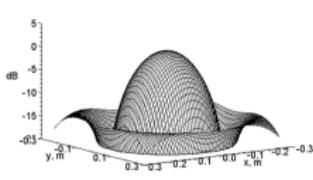
4 dB gain



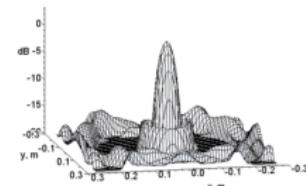
Optimal Array



< 2 kHz



3 kHz



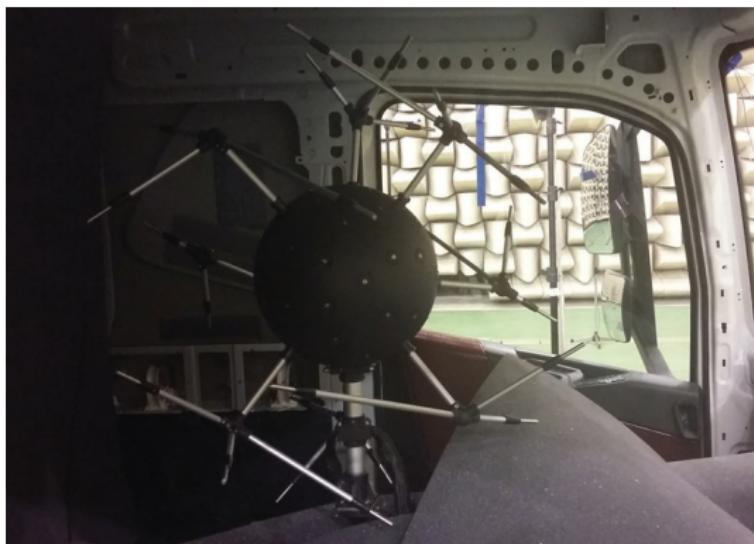
8 kHz

Dynamic optimization = optimal distribution to reduce maximum side lobe level

3D Beamforming

Internal acoustic localization

A 2D plane array cannot distinguish if the acoustic waves are propagating **in front of or behind** it.



Spherical 3Dcam inside truck cabin

⇒ For internal acoustic localization, a 3D array is required to recover the complete acoustic field

3D Beamforming

Application

Internal acoustic source localization in a car cabin, in a wind tunnel.



Identify noise sources



Test modifications



*Evaluate influence
of modifications*

Enhancement of beamforming method

Deconvolution

Beamforming

Advantages	Drawbacks
Robustness Short computation time Precise resolution for high frequencies	No quantitative if more than one source Large resolution for low frequencies Presence of side lobes Decorrelated sources hypothesis

Deconvolution methods

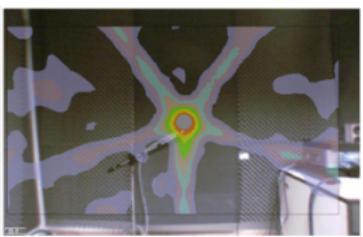
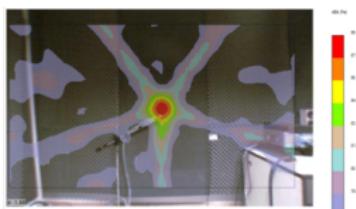
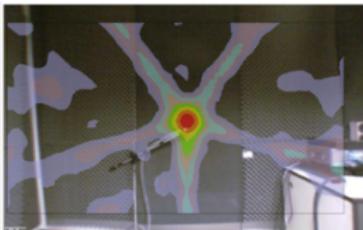
⇒ Remove the array response (**PSF**) of the beamforming maps

(Methods : DAMAS, CLEAN, CLEAN-SC, CIRA...)

Point Spread Function (PSF) = Response of an imaging system to a point source,
system's impulse response

Deconvolution

Main principle



Principle

The side lobes correlated to the maximum sources are iteratively subtracted from the initial map.

How to take into account physical transfer functions ?

In a reverberant environment, the acoustic pressure measured by microphones is modified by **diffraction** / **reflection** / **scattering** phenomena.

Beamformer

$$\sigma^2(x_s) = \frac{\mathbf{H}^H(x_s) \widehat{S}_{pp} \mathbf{H}(x_s)}{||\mathbf{H}(x_s)||^4}$$

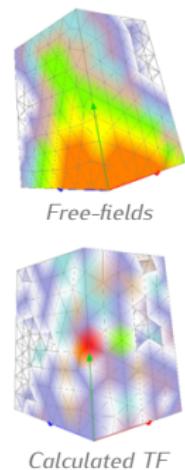
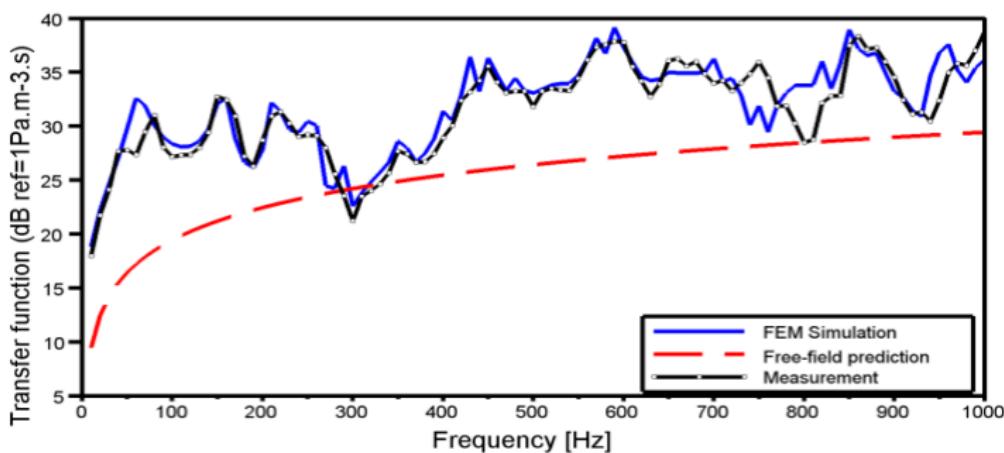
To consider a true environment where the free-fields hypothesis is no longer appropriate, the transfer function **H** must be changed.



How to take into account physical transfer functions ?

To estimate the acoustic transfert functions, three methods :

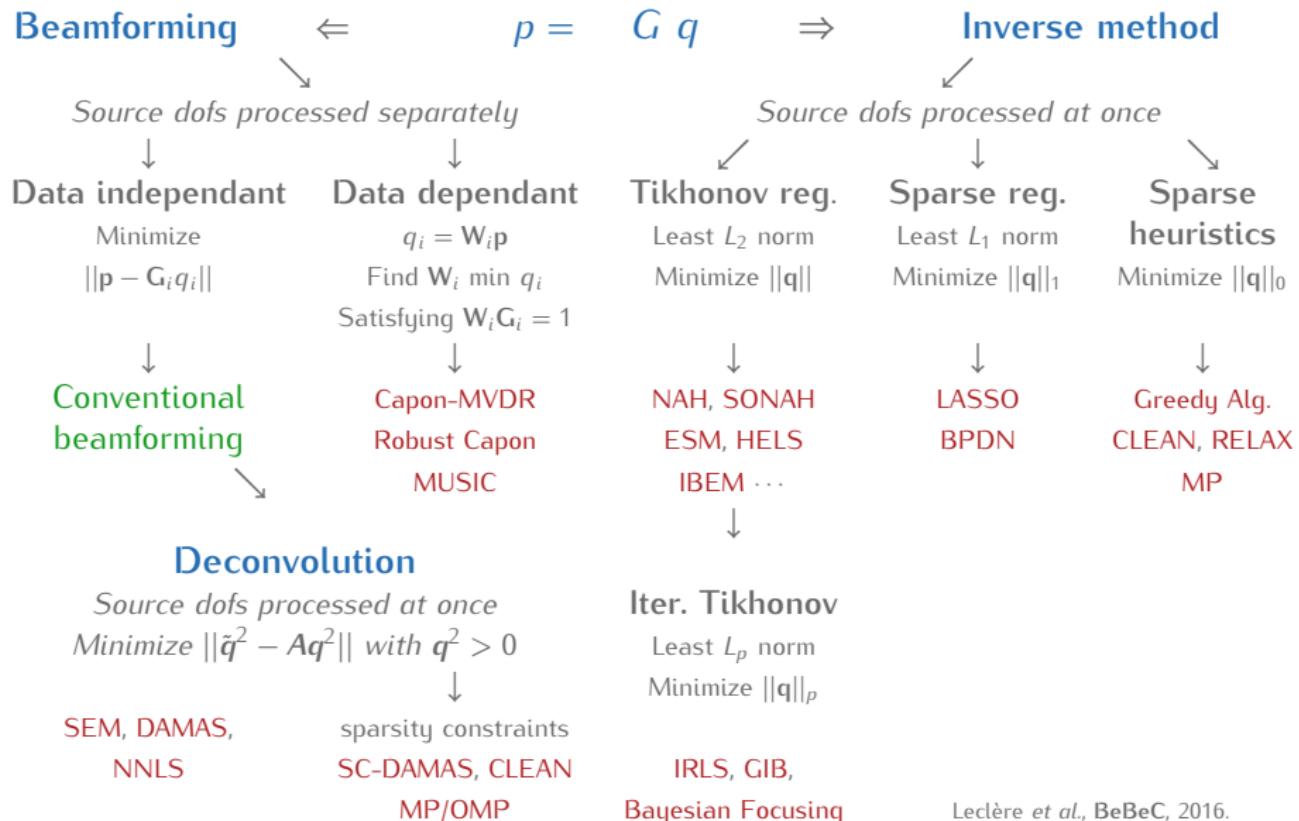
- Measurement : needs to place a calibrated source on all the surface, defined as a grid (at least hundreds of points)
- Analytical model : only for simplest case : reflecting plane, sphere, cylindrical or annular duct...
- Numerical simulation : mostly FEM, still costly and time-consuming



Conclusion

Conclusion

A leafy tree of methods...



Conclusion

Beamforming is the simplest, the most widespread and one of the most robust method for source localization.

It represents an excellent first tool to deal with acoustic imaging issues.

BUT

- No quantitative if more than one source
- Presence of side lobes
- Uncorrelated sources hypothesis (inexact directivity reconstruction)
- Frequency limitations :
 - High frequency : aliasing due to inter-microphone step (Shannon theorem)
 - Low frequency : resolution (depend on the array size)

All those limitations can be filled by advanced methods, often dedicated to specific experimental configurations.

Any questions ?

For specific questions / remarks / internship

✉ simon.bouley@microdb.fr



www.microdb.vibratecgroup.com