Beamforming Techniques and Acoustic Imaging Applications

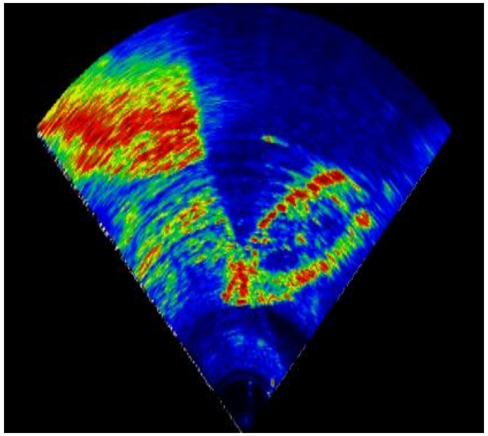
May 2023

Danilo Greco

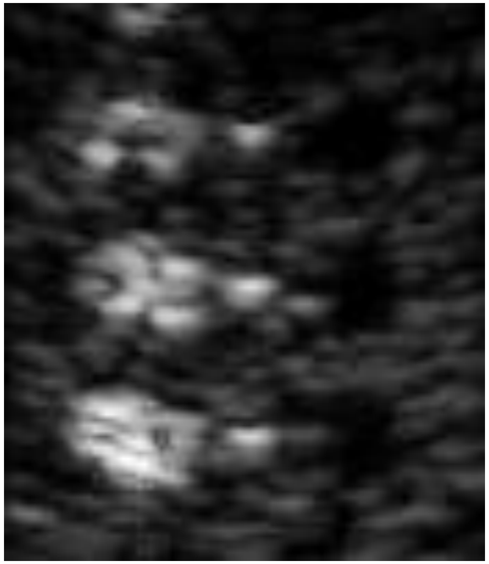
Presentation Overview

- Introduction on Acoustic Imaging
- The Beamforming
- A synthesis method for a very large bandwidth acoustic imaging system
- Results.

What is acoustic imaging?



- Acoustic imaging is an imaging modality which exploits acoustic wave propagation in a medium to recover the acoustic intensity distribution in a given spatial region.
- Well-known and widespread acoustic imaging applications are e.g. sonar and echography.
- Acoustic imaging in air till now has received comparatively little attention.
- In automated surveillance acoustic imaging can provide great advantages, used both alone or jointly with other imaging modalities (Video, infrared etc.).



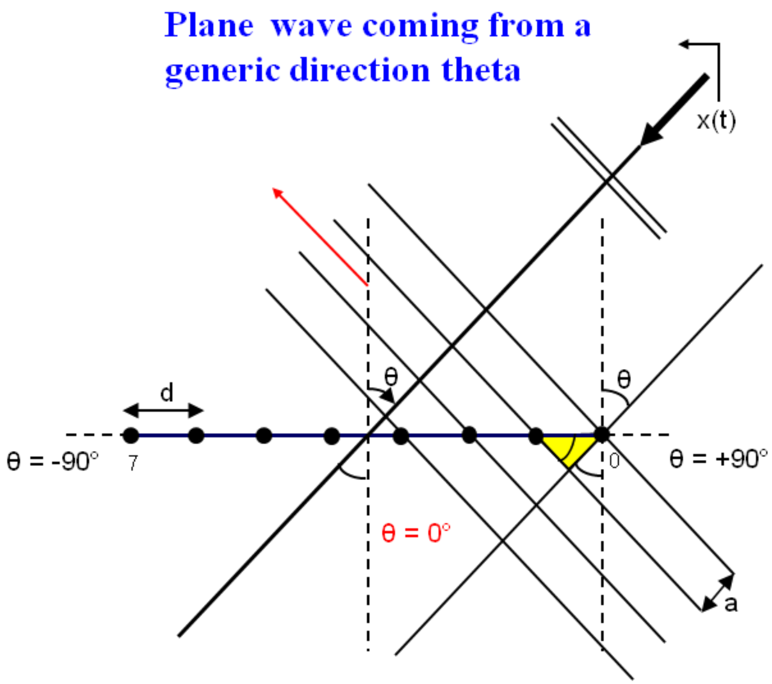
Acoustic imaging in air: assumptions

- Passive acoustic imaging: the imaging device does not emit sounds but acquires environmental sounds.
- The acoustic waves propagating in free space are spatially sampled by a set of omnidirectional microphones (microphone array) placed on a plane.
- Far field condition: the acoustic sources are sufficiently distant from the array; the acoustic waves can be approximated as plane waves.
- In far-field conditions the position of the acoustic source in respect to the array can be described by two angles (e.g. azimuth and elevation in spherical coordinates) while the distance from the array is not considered

Acoustic imaging: key point

- In an acoustic image each pixel corresponds to the acoustic intensity of the source whose location is described by a particular couple of angles.
- Each microphone receives a combination of plane waves coming from different directions. How to separate the different waves impinging on the array?
- The relative delay of the plane wave at each microphone position depends on wave direction of arrival (DOA). As a consequence we can sum coeherently the signal contribution related to one particular DOA, while attenuating all the other contributions.

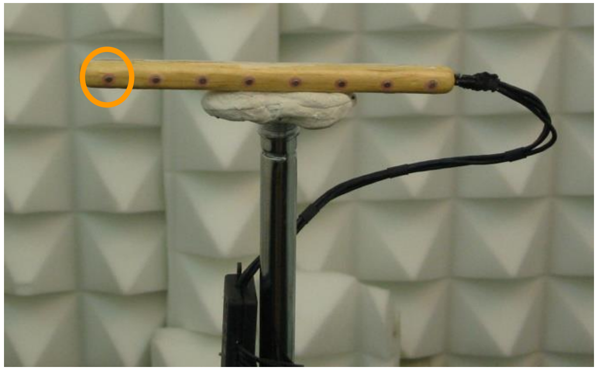
A simpler case: linear array



$$x_n(t,\theta) = x\left(t - d_n \sin(\theta)/c\right)$$

Fourier transform

$$\tilde{x}_n(f,\theta) = e^{-j2\pi f \cdot d_n \cdot \frac{\sin(\theta)}{c}}\,\tilde{x}(f,\theta)$$



d_n: position of the n-th microphone
θ: angle of arrival of the plane wave
$d_n \cdot \sin(\theta)/c$: delay of the wave front sampled by the n-th microphone

The Beamforming Concept

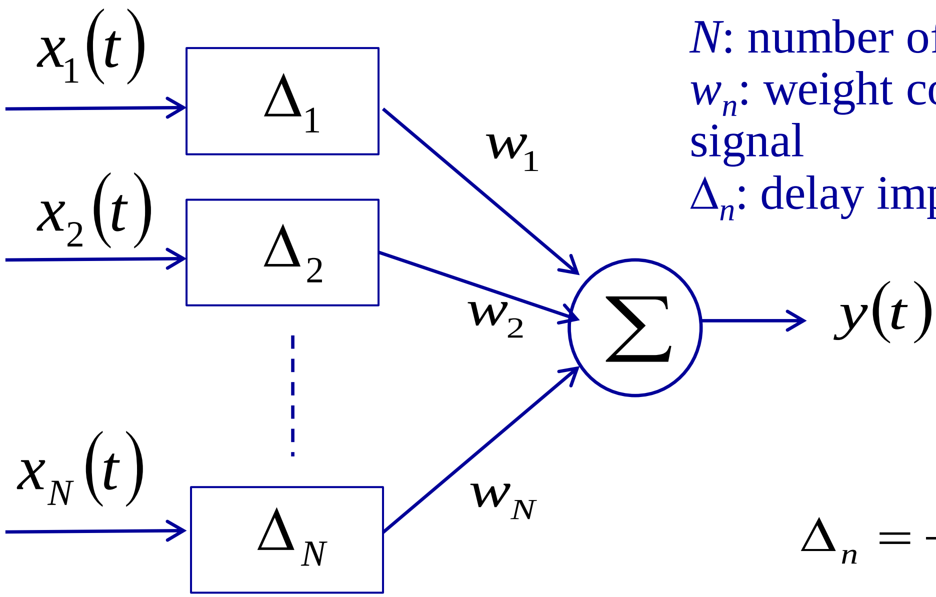
-Beamforming is a spatial filter that linearly combines the temporal signals spatially sampled by an array

-There are two major categories of beamforming algorithms: data-independent and data-dependent

-In imaging applications, data-independent beamforming is used and successfully implemented

-Beamforming arranges the acquired signals in such a way as to amplify the signal coming from a fixed direction (steering direction) and to reduce all the signals coming from any other direction.

Delay-and-sum beamforming



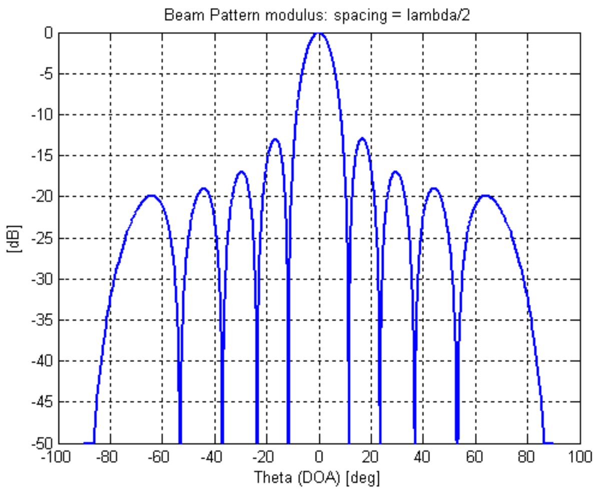
N: number of microphones
w_n: weight coefficient mutiplied by the n-th signal
Δ_n: delay imposed at the n-th signal

$$\Delta_n = -d_n \cdot \frac{\sin(\theta_0)}{c}$$

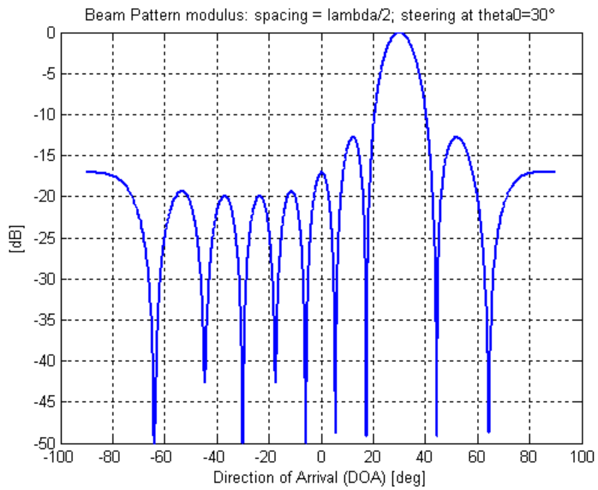
<u>Beam Pattern</u>: beamforming response to a plane wave, function of the angle of arrival theta and the frequency.

$$BP(\theta, f) = \sum_{n=0}^{N-1} w_n e^{-j2\pi f \cdot d_n \cdot \frac{\sin(\theta)-\sin(\theta_0)}{c}}$$

A typical beam pattern shape



Steering at 0° deg



Steering at 30° deg

$$\lambda = \frac{v}{f}$$

10 microphones

f = 5 kHz

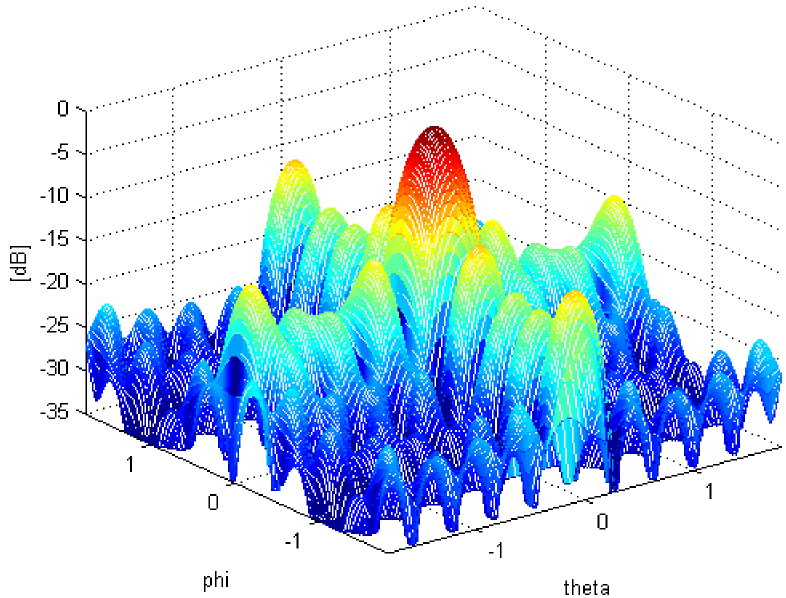
Total array aperture D = 30 cm

d = wavelength/2

No weighting

Adding one dimension: 2D array, 2 angles of arrival

$$BP(\theta, \varphi, f) = \sum_{n=0}^{N-1} w_n e^{-j2\pi f \cdot \left[x_n \cdot \frac{\sin(\theta) - \sin(\theta_0)}{c} + y_n \cdot \frac{\sin(\varphi) - \sin(\varphi_0)}{c} \right]}$$



Two angle of arrival: theta and phi.

Two steering angles: theta0 and phi0.

Two microphone coordinates x_n and y_n.

How the beam pattern shape impacts on the performance of the imaging system?

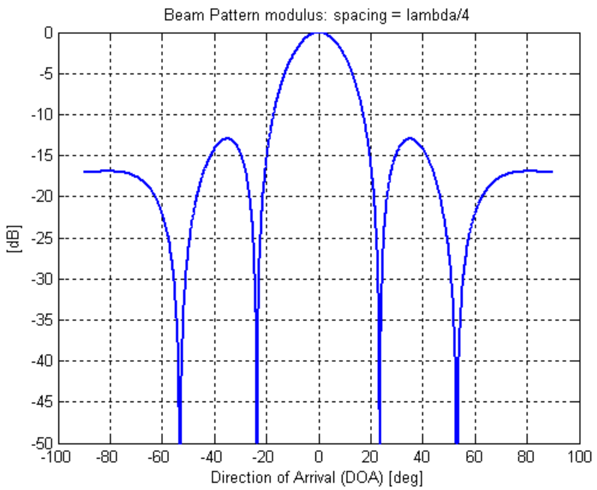
Two empirical rules:

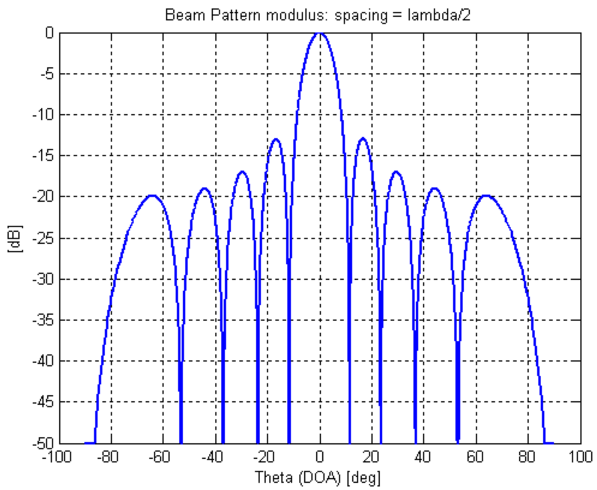
1. **The main lobe width determines the image resolution**
2. **The side lobe eight determines the image SNR**

The image quality improves as the main lobe shrinks and the side lobes decrease

The beam pattern can be interpreted as a point spread function between the acoustic sources in the real space and the image: a single point source in the real space will result in an image reproducing the beam pattern shape.

How to improve the beam pattern shape? (1)





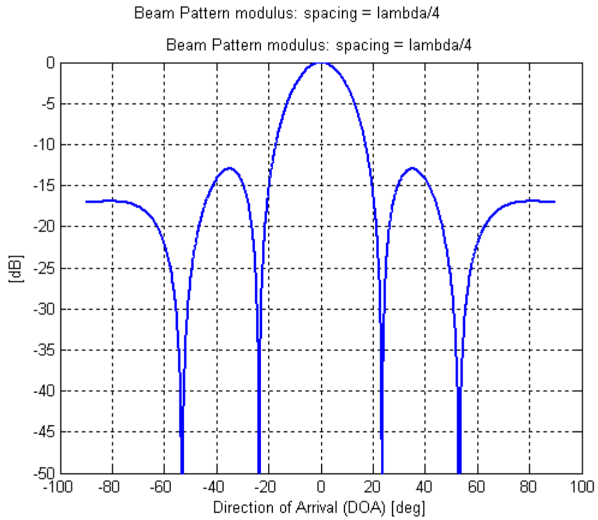
Spacing at lambda/4
Aperture 15 cm

Spacing at lambda/2
Aperture 30 cm

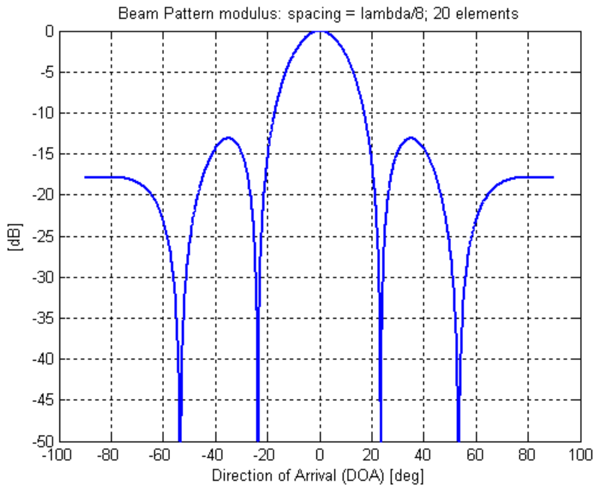
•Increasing the overall array aperture (size) makes the Beam Pattern shrink.

•Practical constraints on the maximum array aperture given by the application.

How to improve the beam pattern shape? (2)



10 microphones
Aperture 15 cm



20 microphones
Aperture 15 cm

•Increasing the number of microphones causes a decrease in the side lobe level.

•Cost and computational constraint on the number of available microphones.

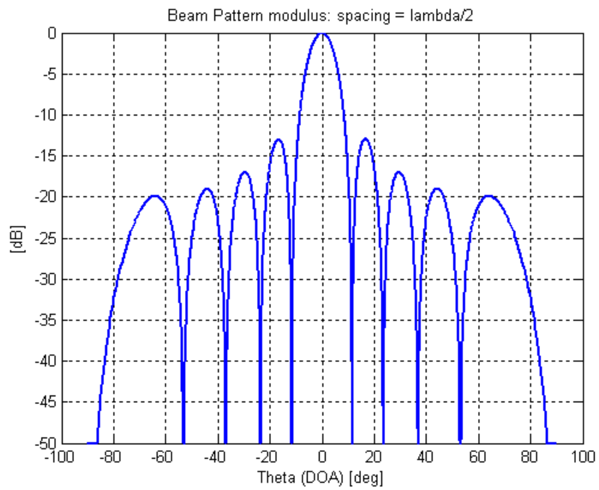
How to improve the beam pattern shape? (3)

•Acting on the weight coefficients allows to decrease the side lobe level.

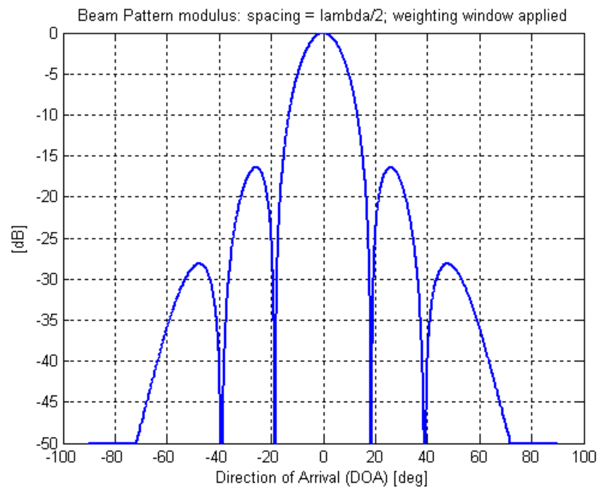
$$B(\theta, f) = \sum_{n=0}^{N-1} w_n e^{-j2\pi f \cdot dn \cdot \frac{\sin(\theta)-\sin(\theta_0)}{c}}$$

•The drawback is a main lobe enlargement.

•Weight coefficients can be set to predefined values (apodization windows) or found through the minimization of a suitable cost function.



Uniform weighting

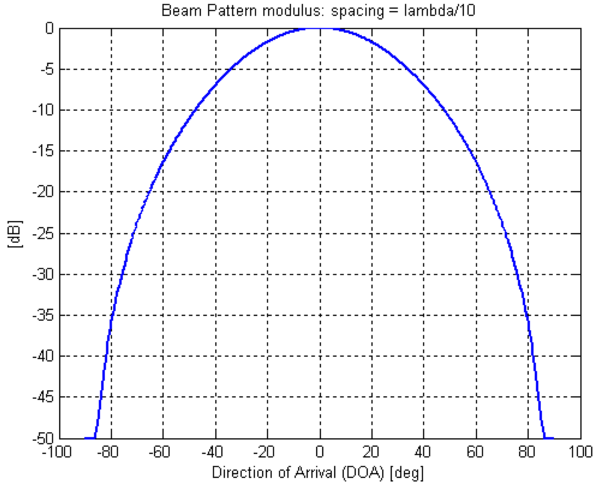


Optimized weighting

What happens for the lowest frequencies?

The wavelength may become equal or even larger than the array aperture D.

Example:
Frequency 1020 Hz
Array aperture = 30 cm
D/lambda = 1.11
Interelement spacing = lambda/10

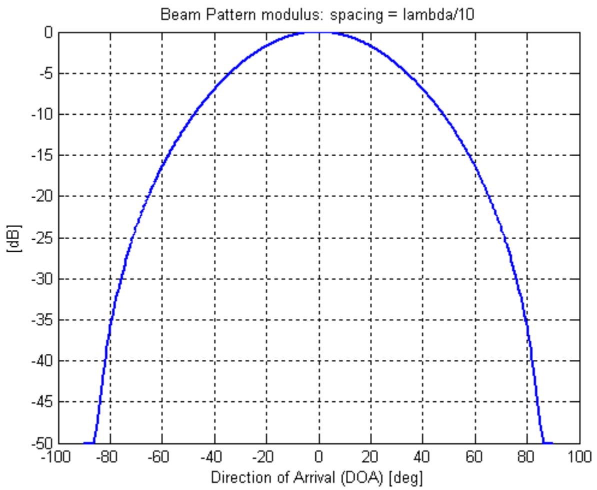


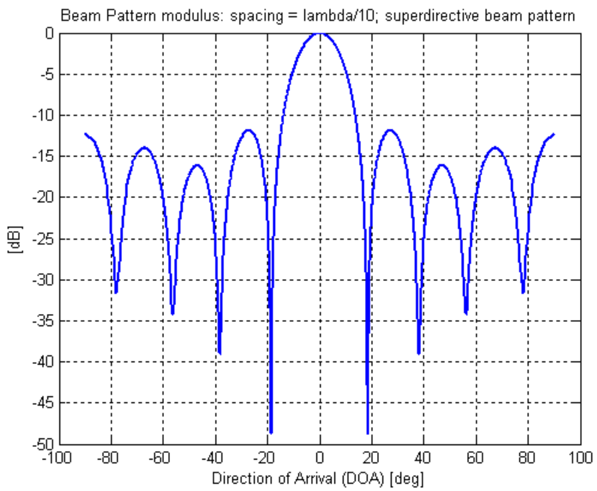
•With uniform weights the main lobe broadens unacceptably.

•Conventional weighting windows are not useful (further main lobe increase)

Superdirectivity theory

When D<= lambda it is possible to shrink the main lobe employing specific weighting windows with alternate signs.





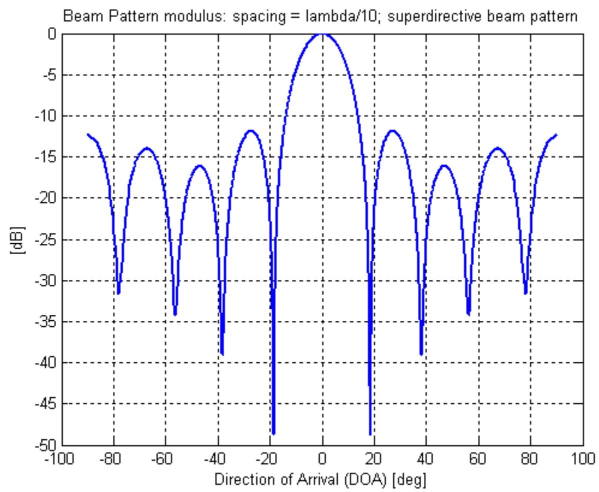
Uniform weighting

Superdirective weighting

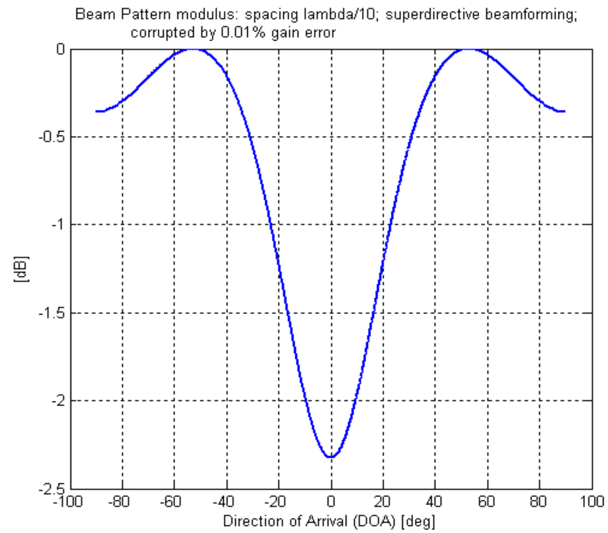
Superdirectivity theory: sensitivity to errors

Superdirective beamforming is often extremely sensitive to deviations from the nominal values in the array characteristics (es. microphone gain and phase, microphone position).

A small deviation in the nominal values can completely degrade the BP



1% deviation from the nominal microphone gains



In superdirective design is mandatory to set constraints on the BP <u>robustness</u>

What happens at the highest frequencies?

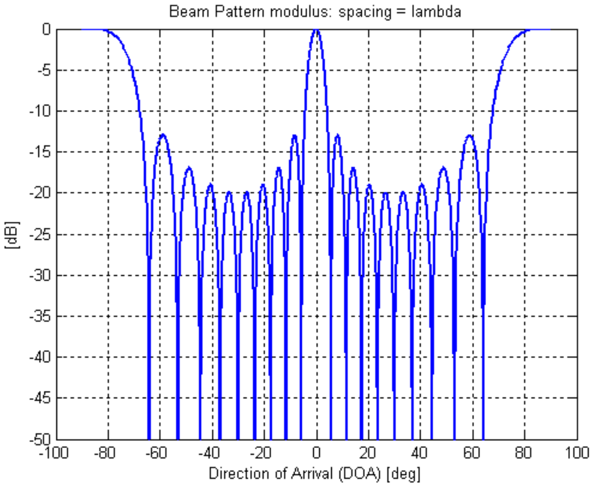
The interelement spacing may become higher than lambda/2.

Nyquist theorem is not respected.

Grating lobes (replicas of the main lobe) appear in the BP.

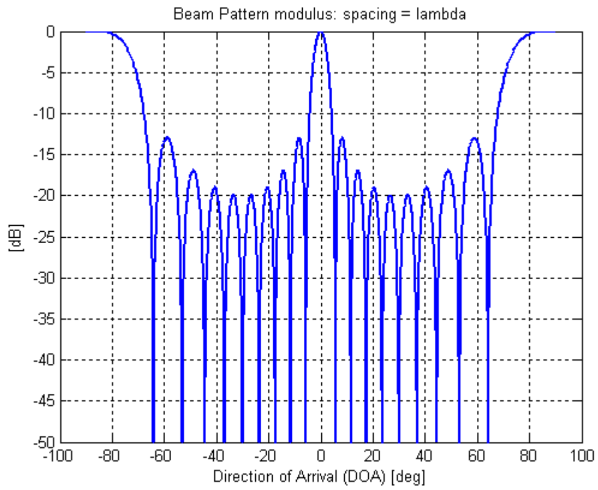
Grating lobes can not be lowered by weighting windows.

Example
Spacing = lambda
10 microphones.
Aperture = 30 cm
Frequency = 10200 Hz

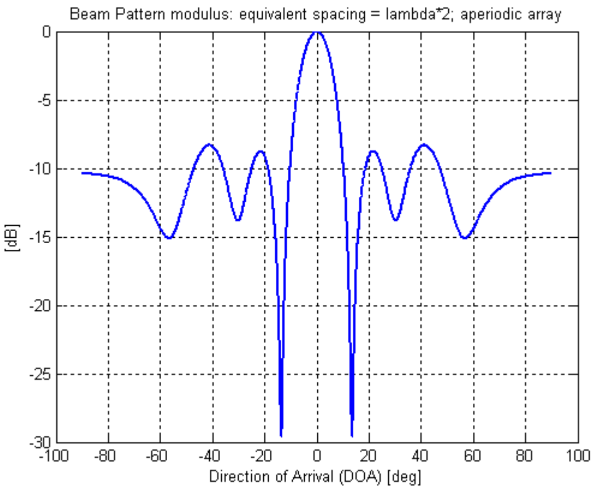


Aperiodic arrays

Breaking the periodicity of the microphones positions is possible to eliminate grating lobes while mantaining the same aperture and the same number of microphones.



Periodic displacement



Aperiodic displacement

To obtain the best result the positions need to be optimized

Broadband beamforming

•A typical audio signal covers a wide frequency band (more than 5 kHz)

•The beamforming shape depends on the frequency of the received signal.

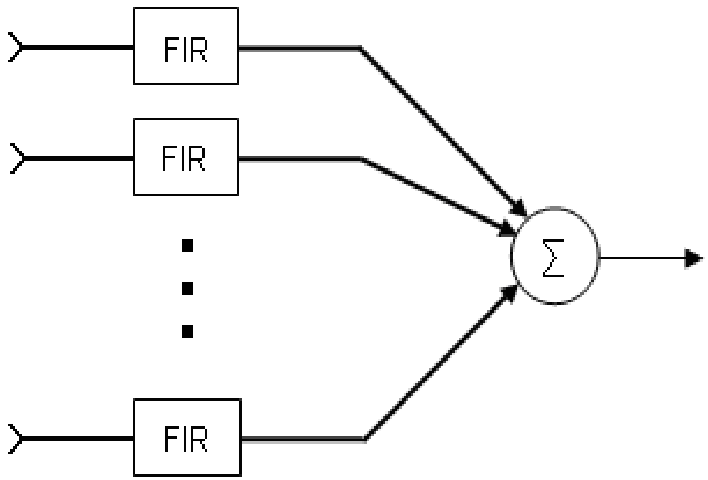
•In particular the weighting coefficients used to model the beamforming can be very different depending on the frequency component of the signal processed.

•Frequency–variant weight coefficients are needed.

•The delay-and-sum beamforming structure is not sufficient.

•A different beamforming structure has to be adopted.

Filter-and-Sum beamforming



The frequency variant weight coefficient are given by the frequency responses of the FIR filters.

$$BP(\theta, f) = \sum_{n=0}^{N-1} e^{-j2\pi f \cdot d_n \cdot \frac{\sin(\theta)-\sin(\theta_0)}{c}} \cdot H_n(f)$$

$H_n(f)$: frequency response of the n-th filter
K: filters order
$w_{n,k}$: k-th coefficient of the n-th filter
T_c: sampling period

$$H_n(f) = \sum_{k=0}^{K-1} w_{n,k} \cdot e^{-j2\pi f \cdot kT_c}$$

Design of very large bandwidth acoustic imaging system

A case study (1)

Design requirements

Frequency band of interest: 300 Hz – 5 kHz

Steering range: about 105° on horizontal and vertical axes .

Image Resolution and SNR: as much as possible.

Design constraints.

32 Microphones.

2D array of maximum aperture 0.8 m x 0.8 m.

Very large bandwidth (more than 4 octaves) and array with limited spatial aperture and reduced number of sensors.

A case study (2)

Working conditions	Problems	Solutions
D<λ at the lowest frequencies	Very poor spatial directivity	Superdirectivity
Superdirective array	High sensibility to sensor mismatching	Robust synthesis method
d>λ/2 at the highest frequencies	Spatial undersampling with *grating lobes*	Optimized Aperiodic array

Beamforming synthesis: cost function

$$J(\mathbf{w},\mathbf{d})=\int_{\theta_{0\min}}^{\theta_{0\max}}\int_{\varphi_{0\min}}^{\varphi_{0\max}}\int_{\theta_{\min}}^{\theta_{\max}}\int_{\varphi_{\min}}^{\varphi_{\max}}\int_{f_{\min}}^{f_{\max}}\left|B(\mathbf{w},\mathbf{d},\theta_0,\varphi_0,\theta_0,\varphi_0,f)-1\right|^2+C\left|B(\mathbf{w},\mathbf{d},\theta,\varphi,\theta_0,\varphi_0,f)\right|^2 df\,d\theta\,d\varphi\,d\theta_0\,d\varphi_0$$

Adherence term

Energy term

w: filter coefficients vector

d: microphone positions vector

Joint optimization of weights and positions: accounts for superdirectivity and aperiodicity

$$B(\mathbf{w},\mathbf{d},\theta_0,\varphi_0\theta,\varphi,f)=\sum_{n=0}^{N-1}\sum_{k=0}^{K-1}w_{n,k}e^{-j2\pi f\cdot\left[x_n\cdot\frac{\sin(\theta)-\sin(\theta_0)}{c}+y_n\cdot\frac{\sin(\varphi)-\sin(\varphi_0)}{c}+kT_c\right]}$$

How to achieve robustness?

Optimization of the mean performance i.e. the multiple integrals of the cost function over the sensors' phase and gain considered as random variables.

$$A_n = a_n \cdot e^{-\gamma_n}$$

$$J^{tot}(\mathbf{w},\mathbf{d}) = \int_{A_0} \cdots \int_{A_{N-1}} J(\mathbf{w},\mathbf{d},A_0,\ldots,A_{N-1}) \cdot f_A(A_0) \cdots f_A(A_{N-1}) \; dA_0 \cdots dA_{N-1}$$

Robust Cost function

Cost function for a BP perturbed by a specific phase and gain error realization.

PDF of the random variable A_n

S. Doclo and M. Moonen, "Design of Broadband Beamformers Robust Against Gain and Phase Errors in the Microphone Array Characteristics," *IEEE Trans. Signal Proc.*, vol. 51, pp. 2511-2526, October 2003.

Computational issues (1)

- Large number of variables to be optimized:
E.g. 32 microphone, FIR filters of 70-th order : 2336 variables.

- Heavy cost function computation: lots of integrals to be calculated.

- Cost function is not convex in respect to the microphone positions.

- In this form the cost function minimization is unfeasible.

- Need to:
    - Simplify the cost function evaluation
    - Devise a smart global minimum search strategy

Computational issues (2)

$$u = \sin(\theta) - \sin(\theta_0)$$
$$v = \sin(\varphi) - \sin(\varphi_0)$$

Substituting u and v into the BP expression we obtain:

$$B(\mathbf{w}, \mathbf{d}, u, v, f) = \sum_{n=0}^{N-1} \sum_{k=0}^{K-1} w_{n,k} e^{-j2\pi f \cdot \left[x_n \cdot \frac{u}{c} + y_n \cdot \frac{v}{c} + kT_c \right]}$$

A new cost function can be defined as follows:

$$J(\mathbf{w}, \mathbf{d}) = \int_{u_{\min}}^{u_{\max}} \int_{v_{\min}}^{v_{\max}} \int_{f_{\min}}^{f_{\max}} \left| B(\mathbf{w}, \mathbf{d}, u, v, f) - 1 \right|^2 + C \left| B(\mathbf{w}, \mathbf{d}, u, v, f) \right|^2 df du dv$$

The new cost function is a good approximation of the original one, allowing to reduce the number of integrals.

$$u_{\max} = \sin(\theta_{\max}) - \sin(\theta_{0\min})$$
$$u_{\min} = \sin(\theta_{\min}) - \sin(\theta_{0\max})$$
$$v_{\max} = \sin(\varphi_{\max}) - \sin(\varphi_{0\min})$$
$$v_{\min} = \sin(\varphi_{\min}) - \sin(\varphi_{0\max})$$

Computational issues (3)

The vector **w** can be extracted from the multiple integrals in the robust cost function obaining a quadratic form in **w.**

$$J^{tot}\left(\mathbf{w},\mathbf{d}\right)=\mathbf{w}\cdot\mathbf{M}(\mathbf{d})\cdot\mathbf{w}^{T}-2\mathbf{w}\cdot\mathbf{r}^{T}(\mathbf{d})+s$$

Under opportune hypotheses also the integrals on the variables A_n can be extracted from the matrix **M** and the vector **r** and calculated in closed form obtaining:

$$J^{tot}\left(\mathbf{w},\mathbf{d}\right)=\mathbf{w}\cdot\mathbf{A}\otimes\tilde{\mathbf{M}}(\mathbf{d})\cdot\mathbf{w}^{T}-2\mathbf{w}\cdot\mathbf{a}\otimes\tilde{\mathbf{r}}^{T}(\mathbf{d})+s$$

For each element of $\tilde{\mathbf{M}}$ and $\tilde{\mathbf{r}}$ the integral on the frequency can be calculated in closed form, moreover it can be demonstrated that the two integrals in *u* and *v* can be converted into a single integral.

Minimization strategy (1)

•For a fixed microphone displacement the global minimum of the robust cost function can be calculated in closed form.

•On the contrary the presence of local minima in respect to the microphone position prevents from using gradient-like iterative methods.

Hybrid strategy analytic and stochastic based on Simulated Annealing algorithm.

M. Crocco and A. Trucco, “A Synthesis Method for Robust Frequency-Invariant Very Large Bandwidth Beamforming,” *18th Europ. Signal Processing Conf. (EUSIPCO-2010)* , Aalborg (Denmark), pp. 2096-2100, October 2003.

Simulated annealing

•Iterative procedure aimed at minimizing an energy function f(y).

•At each iteration, a random perturbation is induced in the current state y_i.

•If the new configuration, y*, causes the value of the energy function to decrease, then it is accepted.

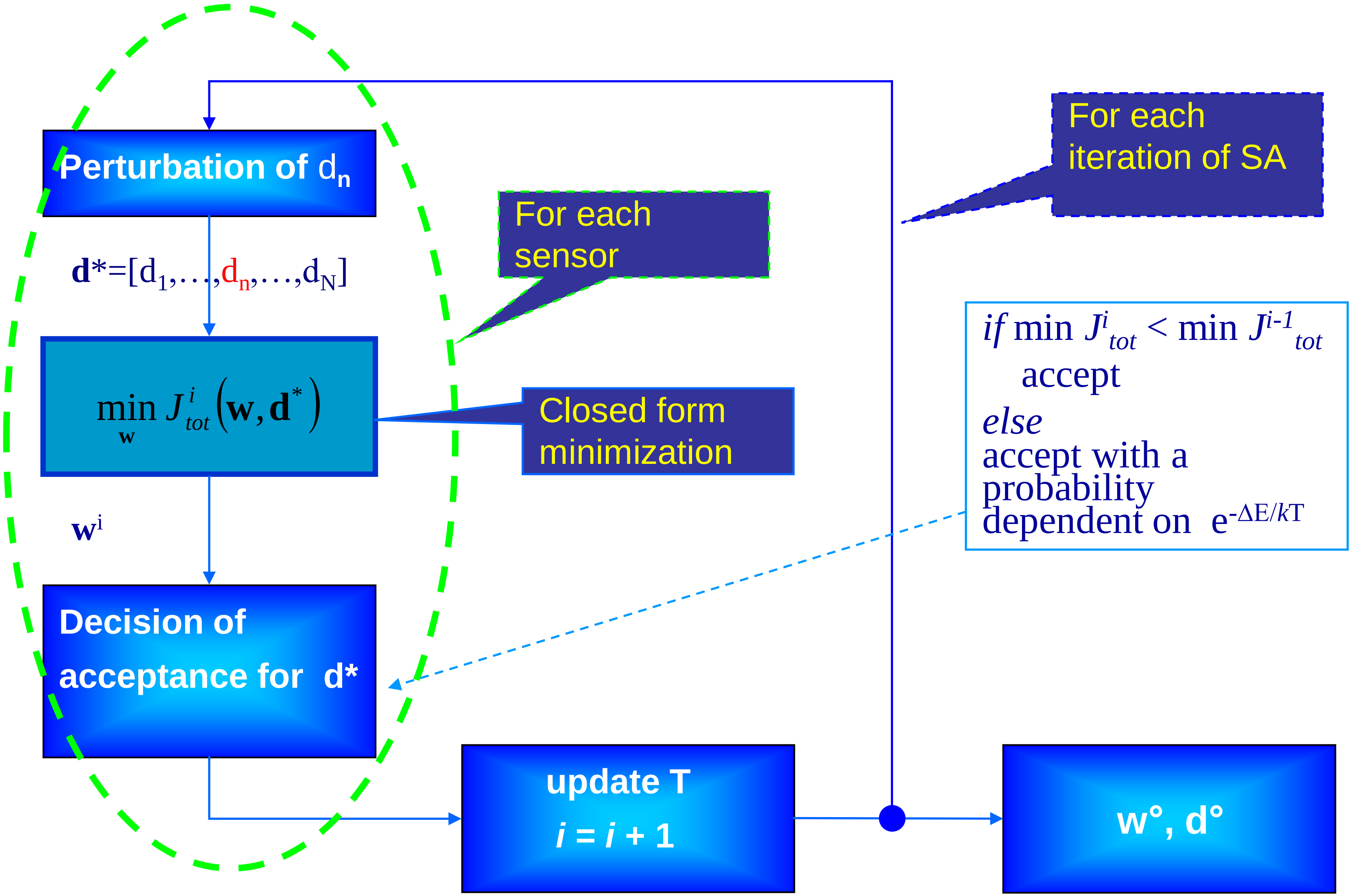
•If y *causes the value of the energy function to increase, it is accepted with a probability dependent on the system temperature, in accordance with the Boltzmann distribution.

•The temperature is a parameter that is gradually lowered, following the reciprocal of the logarithm of the number of iterations.

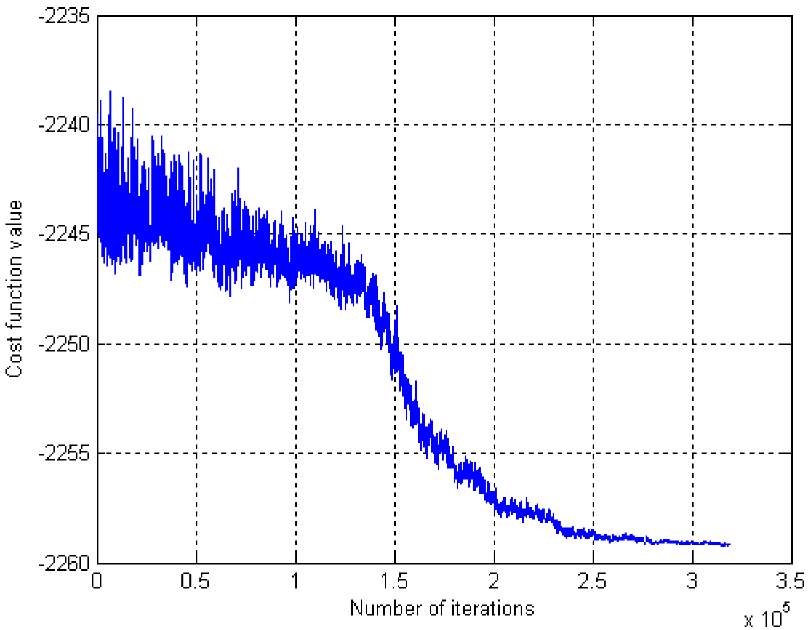
•The higher the temperature, the higher the probability of accepting a perturbation causing a cost increase and of escaping, in this way, from unsatisfactory local minima.

S. Kirkpatrick, C.D. Gellat, Jr, and M. P. Vecchi, "Optimization by simulated annealing," *Sci.*, vol. 220, no. 4598, pp. 671-680, 1983.

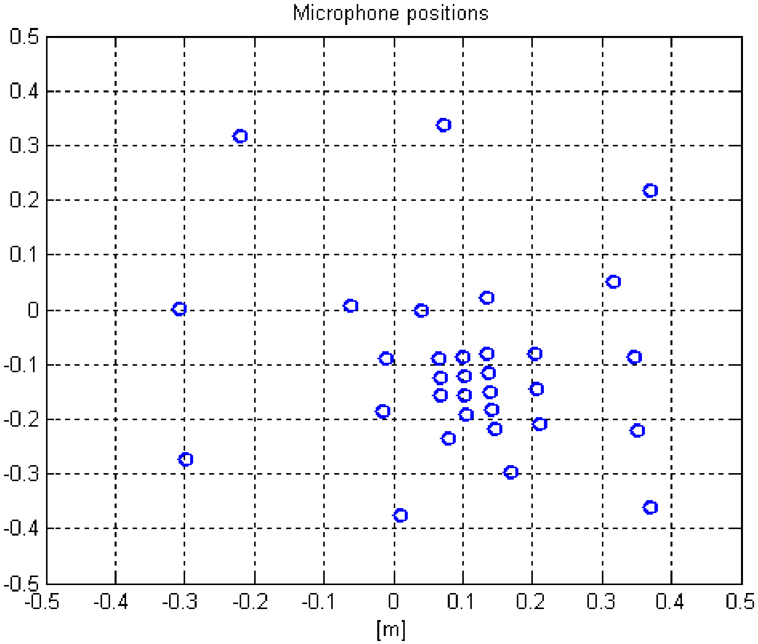
Minimization scheme



Results

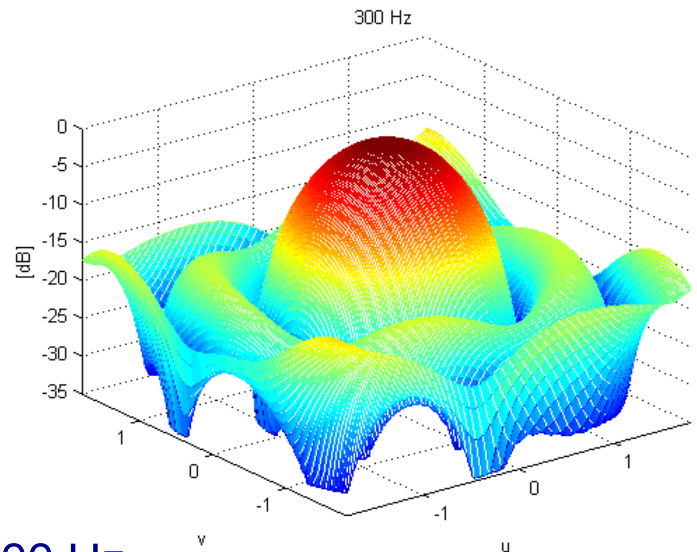


Cost function vs number of iterations

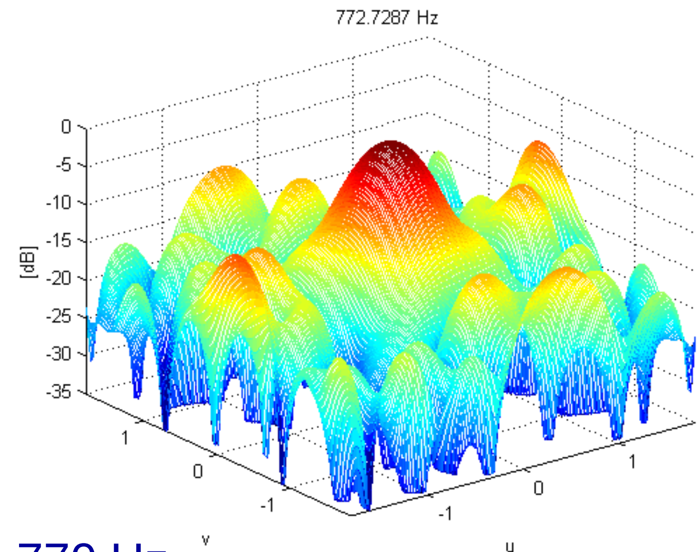


Final microphone displacement

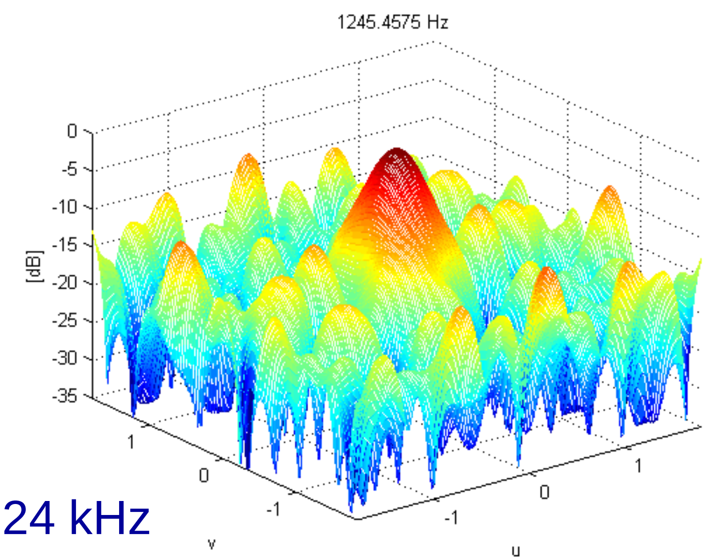
Results: BP modulus (1)



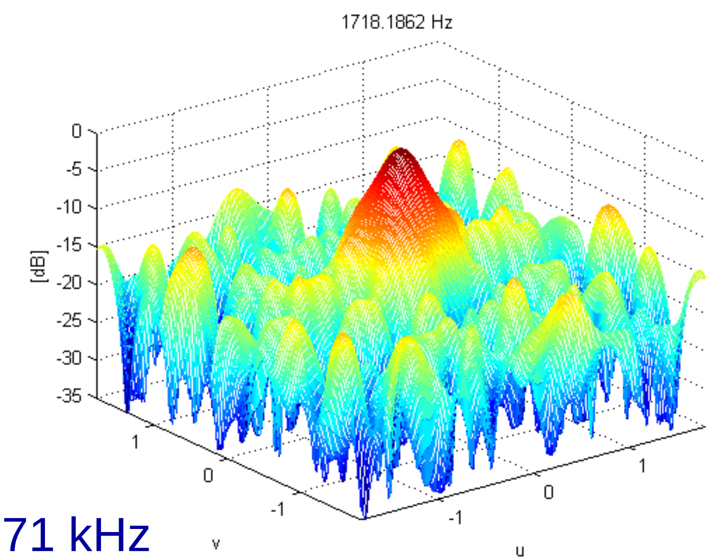
300 Hz



770 Hz

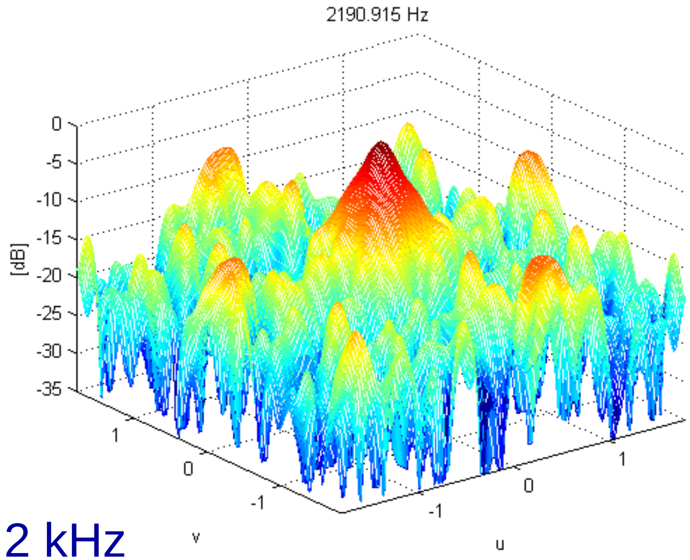


1.24 kHz

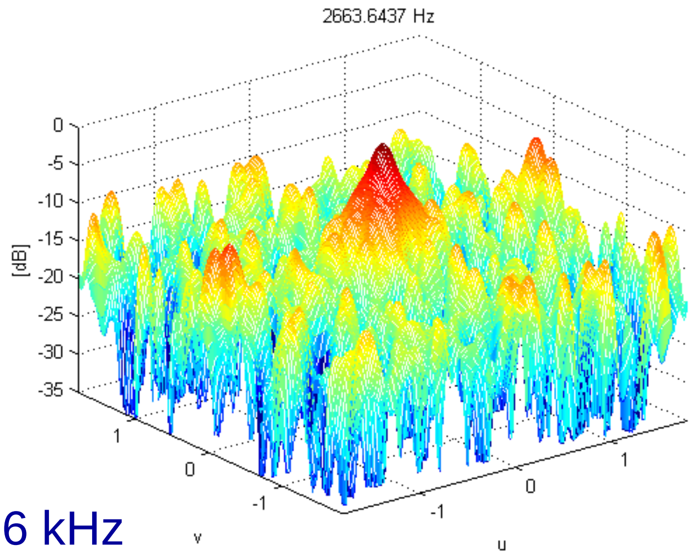


1.71 kHz

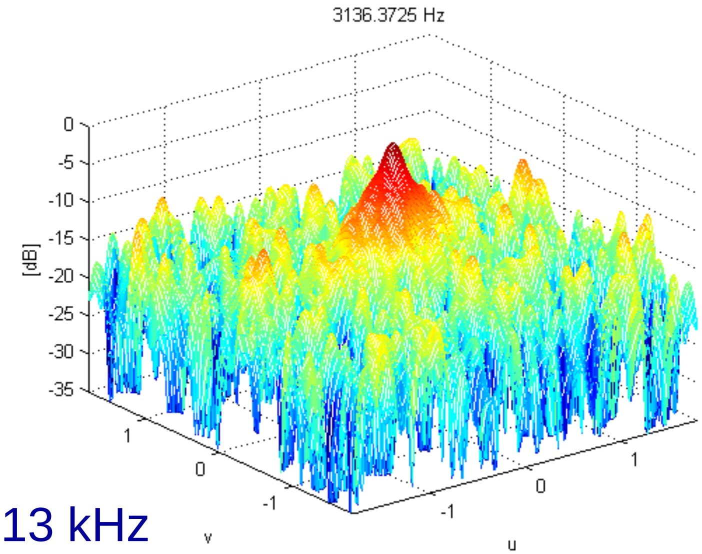
Results: BP modulus (2)



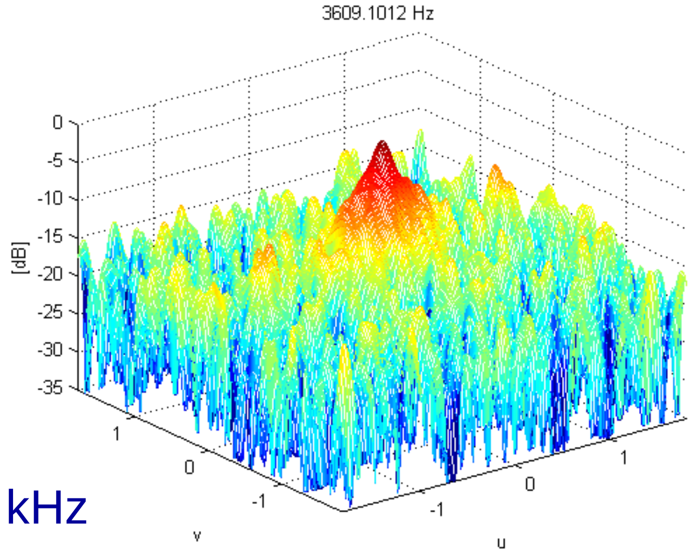
2.2 kHz



2.66 kHz

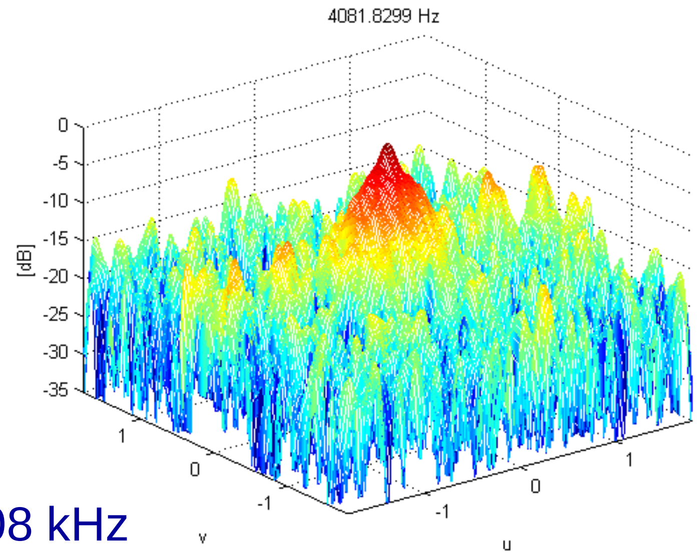


3.13 kHz

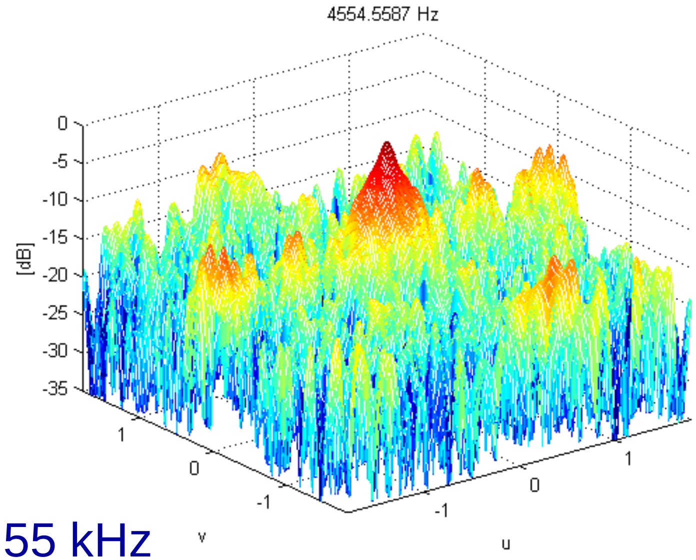


3.6 kHz

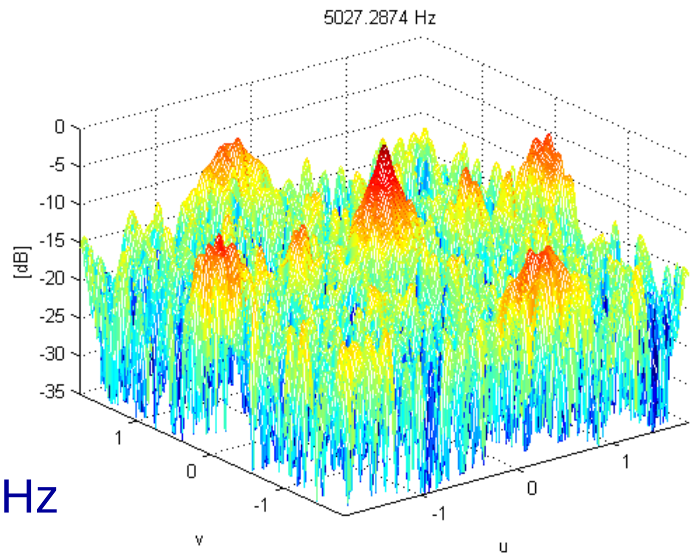
Results: BP modulus (3)



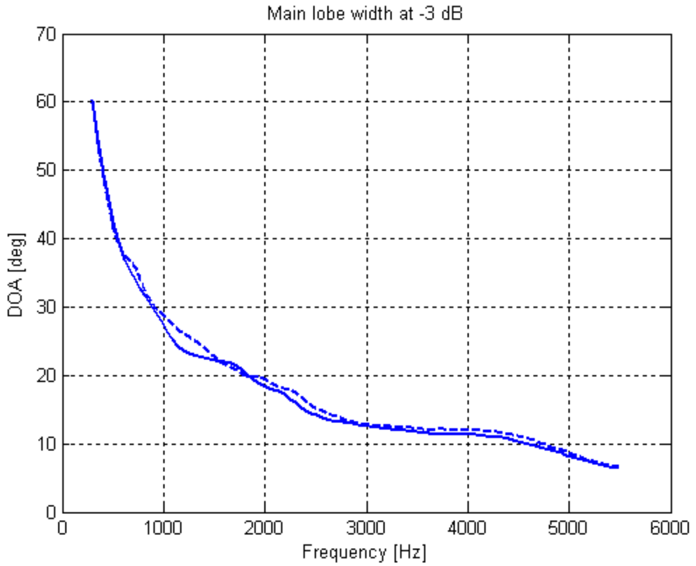
4.08 kHz



4.55 kHz



5 kHz



Performance evaluation (1)

Directivity [dB]

- SNR increase yielded by the array in respect to the single sensor toward isotropic noise.

$$S_a(f) = \frac{\left|BP(\theta_0, \varphi_0, f)\right|^2}{\frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} BP\left|(\theta, \varphi, f)\right|^2 \sin(\theta)\, d\theta d\varphi}$$

White Noise Gain (WNG) [dB]

- SNR increase yielded by the array in respect to the single sensor towards sensor self-noise

$$S_s(f) = \frac{\left|BP(\theta_0, \varphi_{0,} f)\right|^2}{\sum_{n=0}^{N-1}\left|H_n(f)\right|^2}$$

- <u>It measures the system robustness towards imperfections in the array characteristics</u>

Performance evaluation (2)

Expected Beam Power Pattern (EBPP)

- Expectations of the squared modulus of the perturbed beam pattern.

$$B_e^2(\theta,\varphi,f)=E\left\{\left|B_r(\theta,\varphi,f)\right|^2\right\}=\int_{A_0}\cdots\int_{A_{N-1}}\left|B_r(\theta,\varphi,f)\right|^2$$

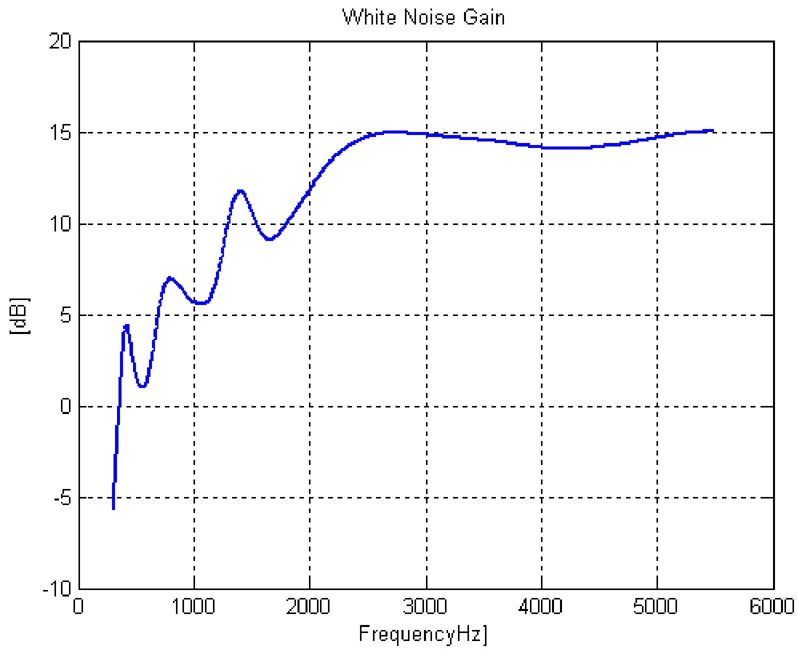
$$\cdot f_{A_0}(A_0)\cdots f_{A_{N-1}}(A_{N-1})\,dA_0\cdots dA_{N-1}$$

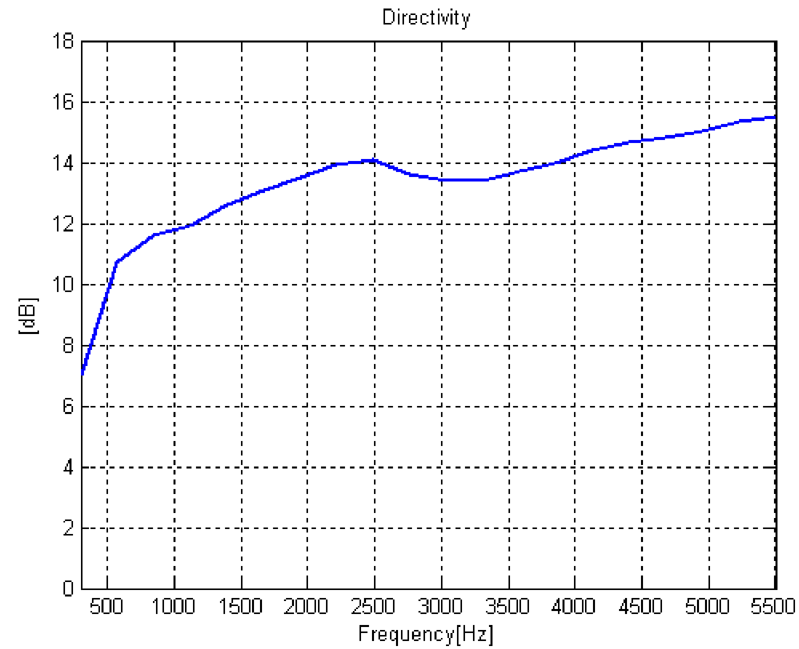
- It allows evaluating the impact on the BP of a given variance in the array characteristics (gain and phase of the microphone responses).
- Under the assumption of small variances can be approximated as:

$$B_e^2(\theta,\varphi,f)=\left|B(\theta,\varphi,f)\right|^2+\frac{1}{S_s(f)}\left(\sigma_g^2+\sigma_\psi^2\right)$$

- The second term sets a threshold below which the EBPP cannot decrease.

Results: WNG and directivity

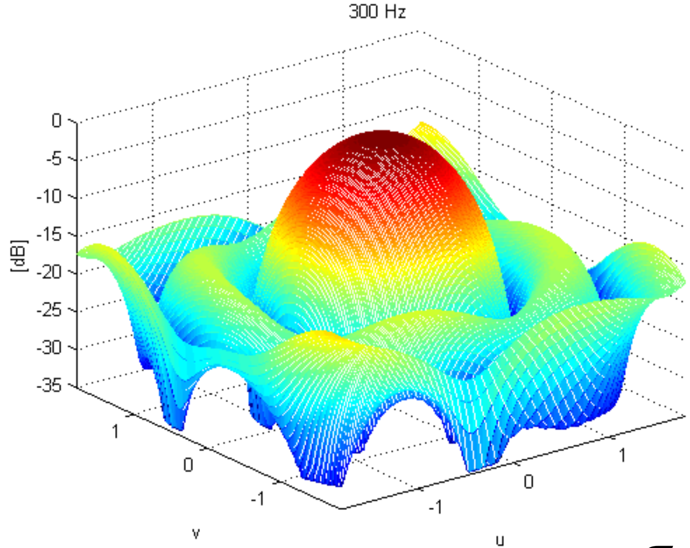




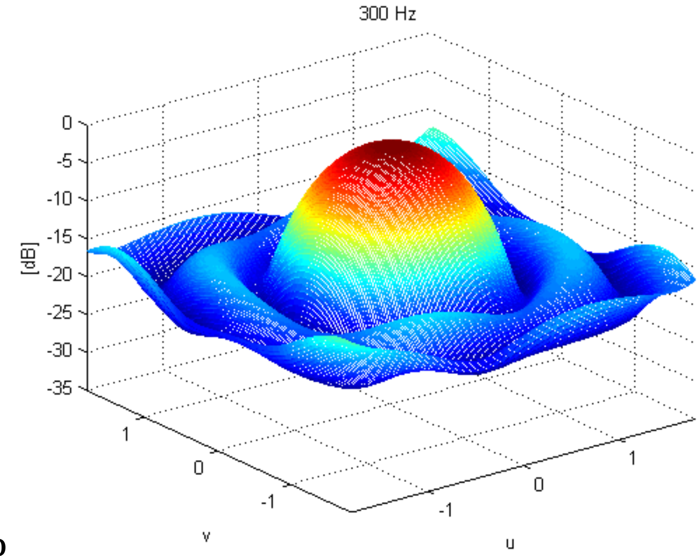
White noise gain vs. frequency

Directivity vs. frequency

Results: BP vs EBPP



300 Hz

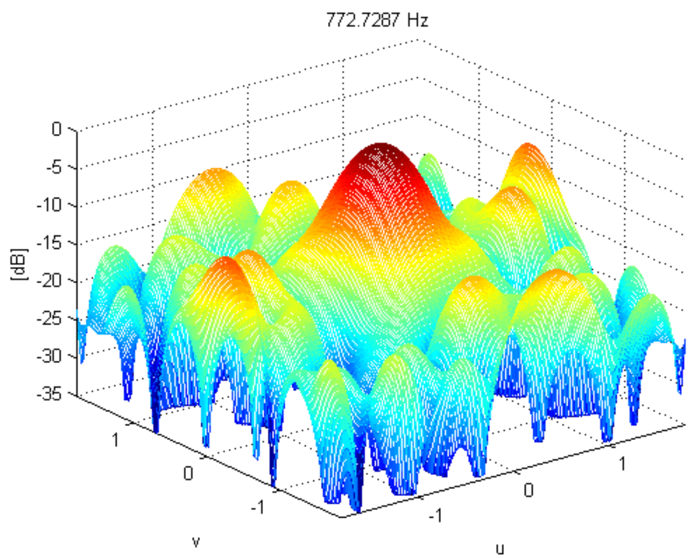


BP modulus

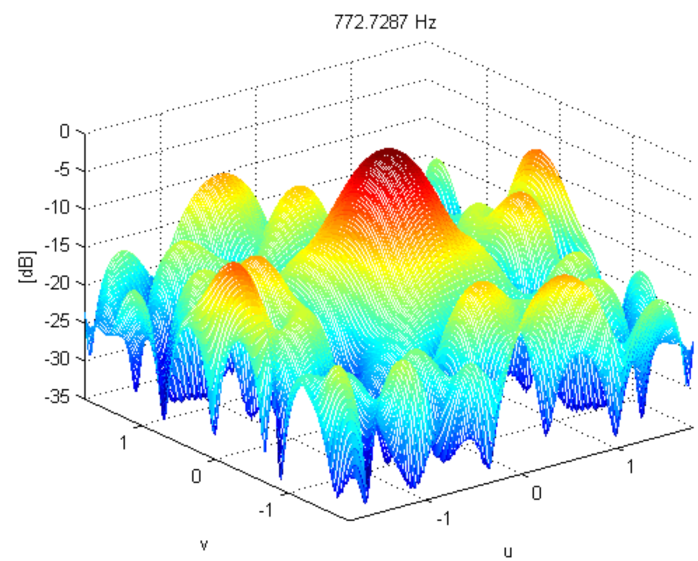
$$\sigma_g = 0.03 = 3\%$$

$$\sigma_\psi = 0.035 rad \cong 2^\circ$$

EBPP

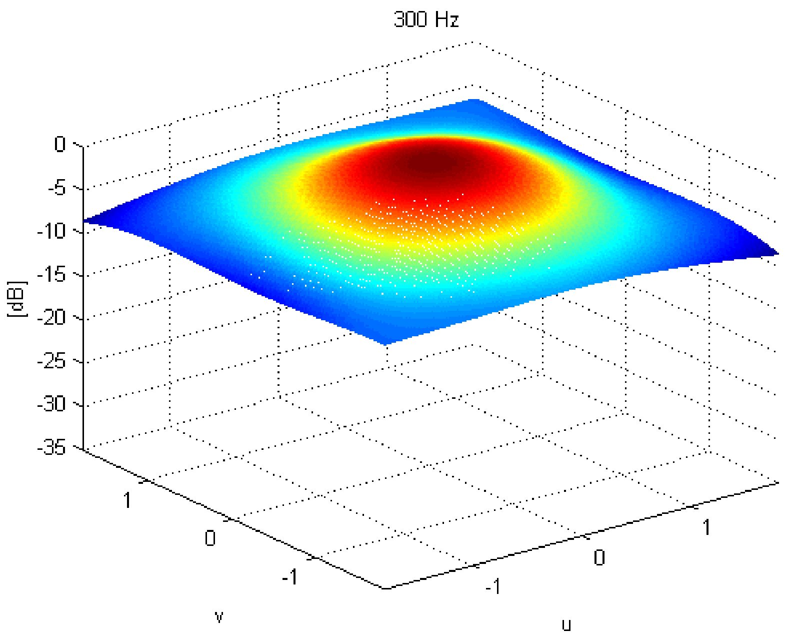


770 Hz

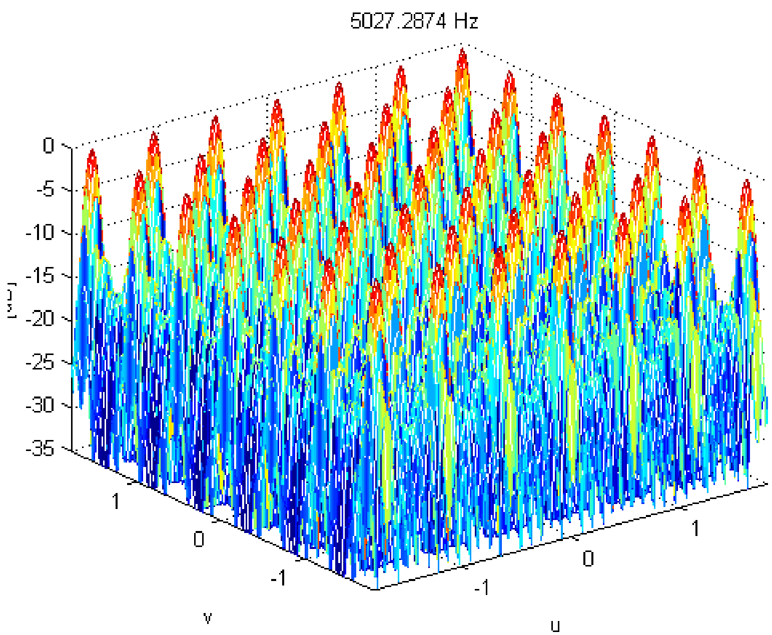


Are superdirectivity and aperiodic array mandatory?

Same array aperture and number of microphones.



300 Hz: uniform weghting
(no superdirectivity)



5 kHz: periodic array

Conclusions

•A method for the design of an acoustic imaging system able to encompass a very large frequency bandwidth has been proposed.

•The method deals with both robust superdirective beamforming and aperiodic arrays allowing to obtain satysfying beam patterns with limited array aperture and number of microphones.

•Further research will be devoted to tune the trade-off between resolution and SNR of the imaging system by controlling the side lobe level and main lobe width.

Additional constraints on the microphone positions will be inserted in the design method in order to take into account the physical supporting structure of the array.