Fundamentals of Sound Source Localization

University of Pretoria – August 2018

Restricted © Siemens AG 2018

Realize innovation.

Fundamentals of Sound Source Localization
Agenda

- **Introduction to Sound Source Localization**
- Beamforming-based localization methods
- Far-field Deconvolution methods
- Map averaging methods

Restricted © Siemens AG 2018

Introduction to sound source localization
Overview of current techniques

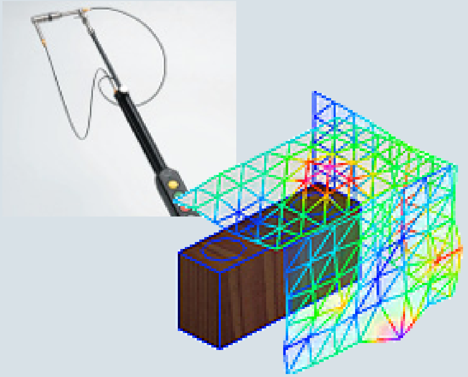
Simple → Advanced

Stethoscope

- ❑ Limited hardware

- ❑ No reportable data
- ❑ Sound pressure only
- ❑ Stationary sources only

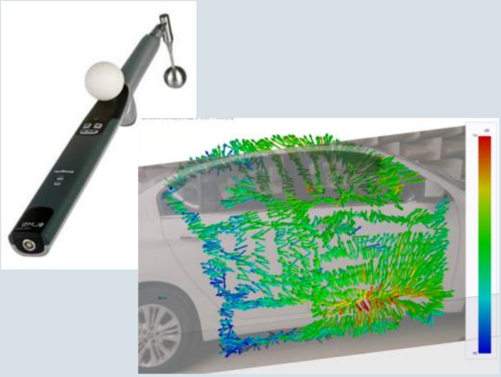
Sound Intensity



- ❑ ISO sound power & partial sound power

- ❑ Time-consuming
- ❑ Limited frequency range
- ❑ Stationary sources only

SoundBrush



- ❑ Fast results
- ❑ Limited hardware
- ❑ Directivity & sound power estimation

- ❑ Stationary sources only

Sound Camera

- ❑ Real-time results
- ❑ Stationary and transient sources

- ❑ Requires more advanced knowledge to operate

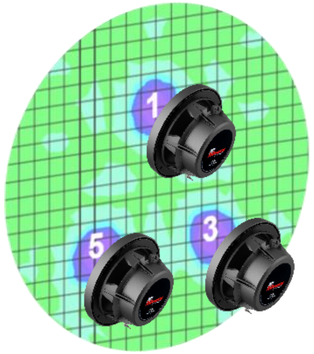
Restricted © Siemens AG 2018

Introduction to sound source localization
Why don’t we just measure sound pressure?

Measured
pressure field
(7 sources?)

Physical propagation

Back-propagation



Source
pressure field
(3 sources)

Sound pressure waves are deformed during physical propagation towards the array (interference effects, reflections, absorption, etc.)

Raw measured pressure maps do not show correct localization of sound sources

Need for specialized back-propagation techniques

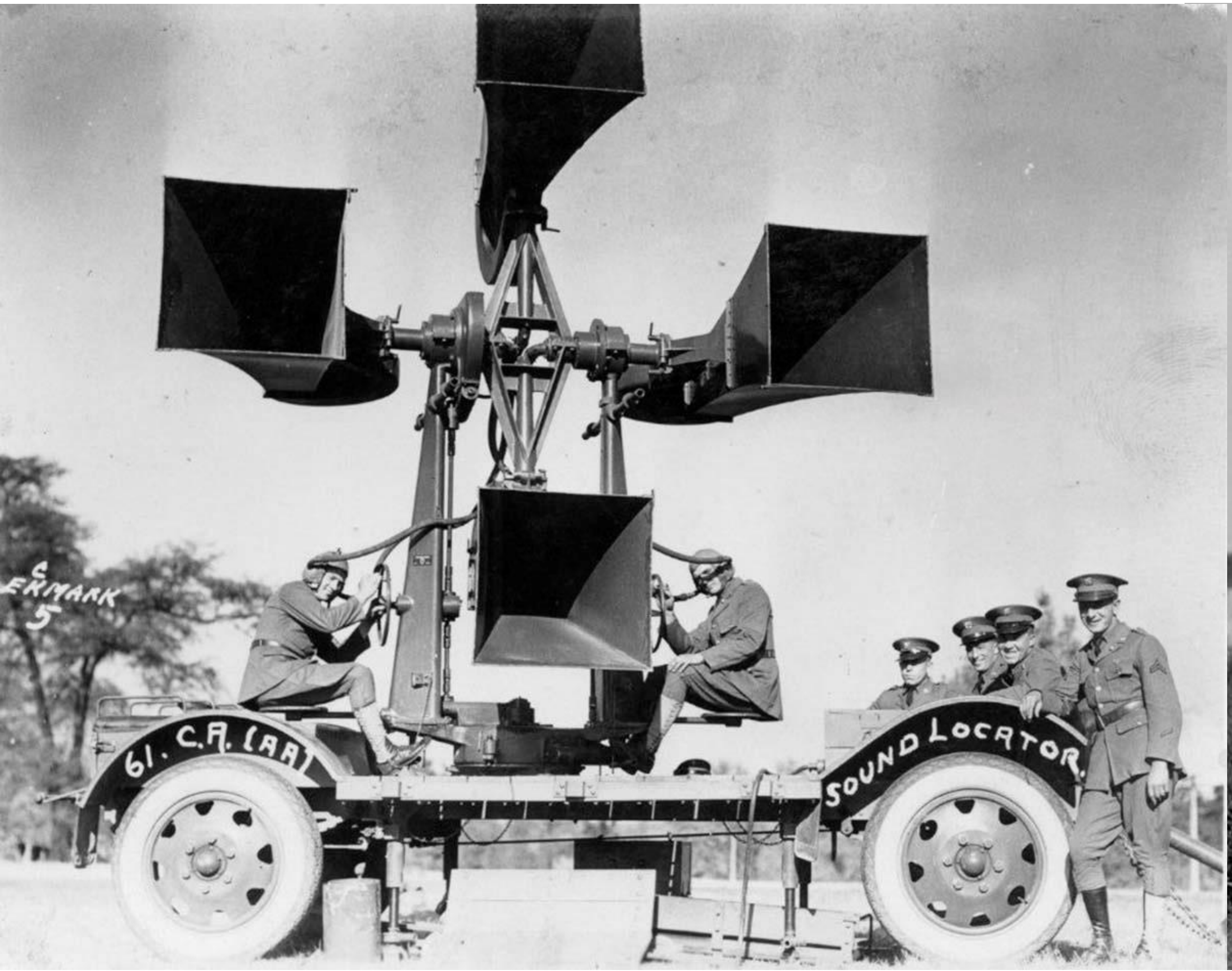
Wide variety of techniques available today: Beamforming, Holography, Deconvolution, etc.

No fit-for-all solution, every technique has its own boundary conditions, strengths and weaknesses

Microphone arrays are an expert tool

Restricted © Siemens AG 2018

Beamforming
A new method?



Restricted © Siemens AG 2018

Fundamentals of Sound Source Localization
Agenda

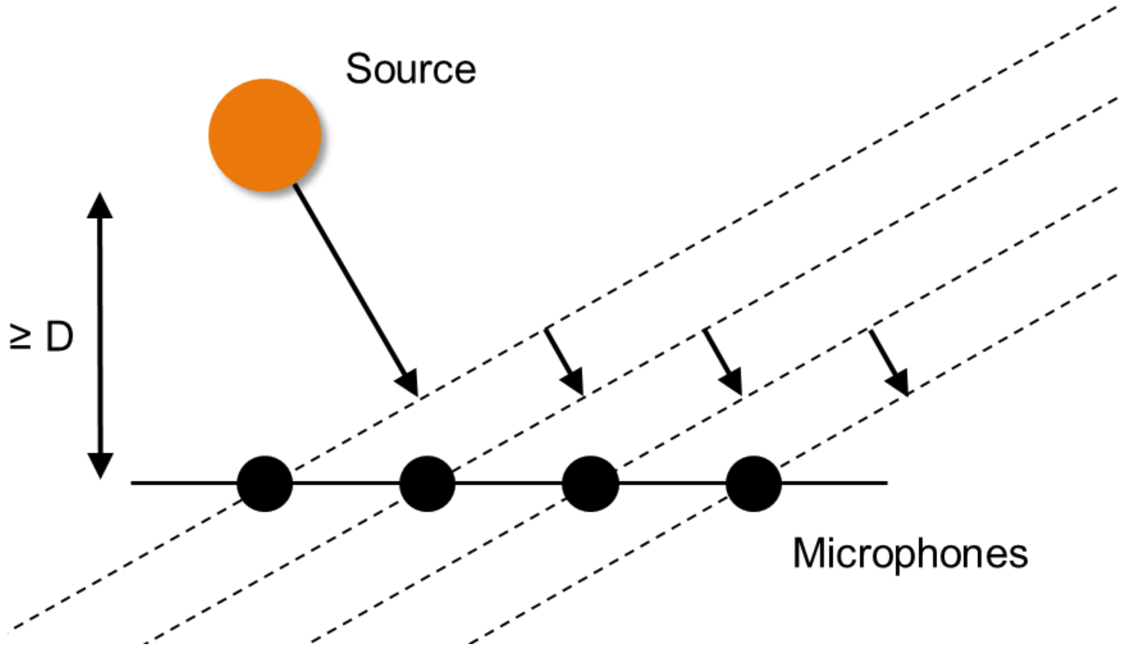
SIEMENS
Ingenuity for life

- Introduction to Sound Source Localization
- **Beamforming-based localization methods**
- Far-field Deconvolution methods
- Map averaging methods

Restricted © Siemens AG 2018

Beamforming
How does it work?

- Assume source emits **planar** pressure waves
- Assumption requires **far-field** measurement conditions
  → Distance to source ≥ array diameter D
- Microphone array consists of many microphones which record the sound pressure signal simultaneously
- Spatial distribution of sources causes a small time delay between the measured signals



Beamforming core principle:
Sound source origin can be extracted from time delay (or phase delay) information between microphones

Restricted © Siemens AG 2018

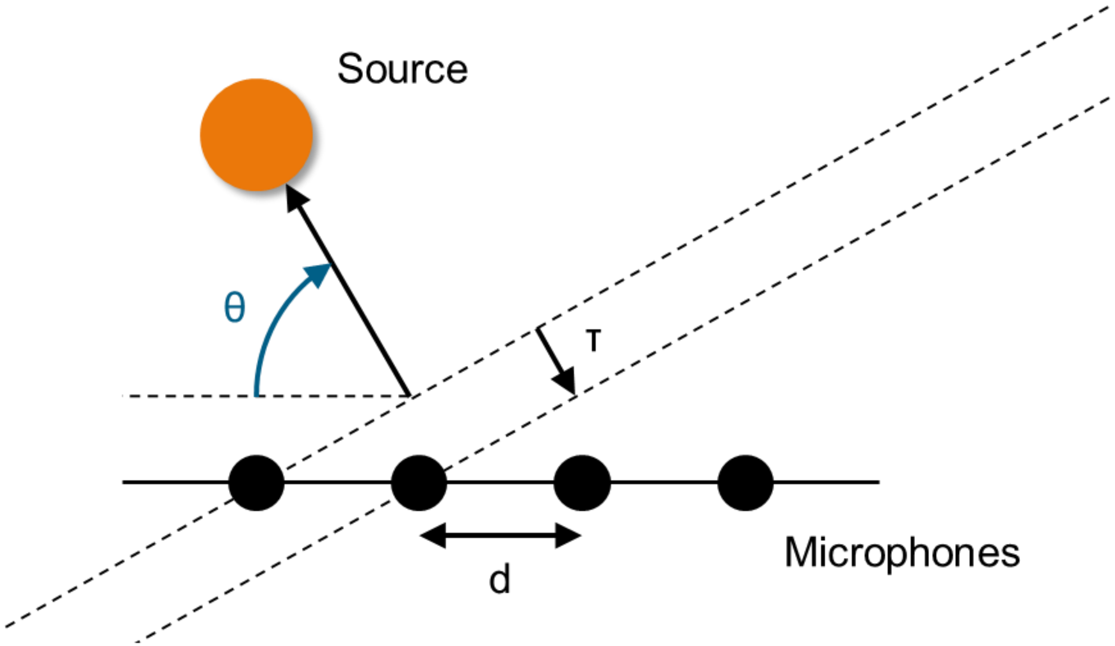
Beamforming
How does it work?

SIEMENS
Ingenuity for life

- Time delay **τ** between microphone signals depends on:
  - Distance between microphones **d**
  - Direction of the source vs. the microphone array **θ**
  - Speed of sound in ambient conditions **c**

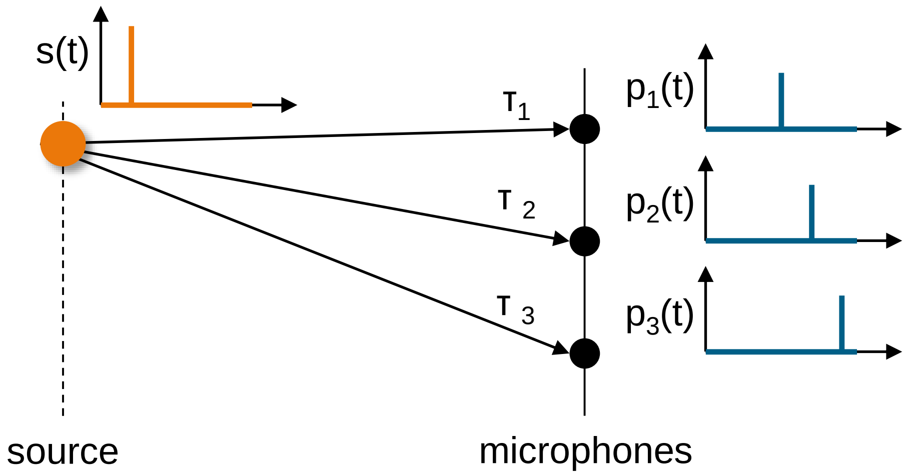
$$\tau = \frac{d \cdot cos(\theta)}{c}$$

- Beamforming is a ‘delay-and-sum’ method:
  - Apply reverse time delay for each potential source
  - Sum microphone signals
- Reverse time delay for true source directions results in in-phase summation of microphone signals
- **Summation amplitude ~ localization probability**



Restricted © Siemens AG 2018

Beamforming
Time-domain example

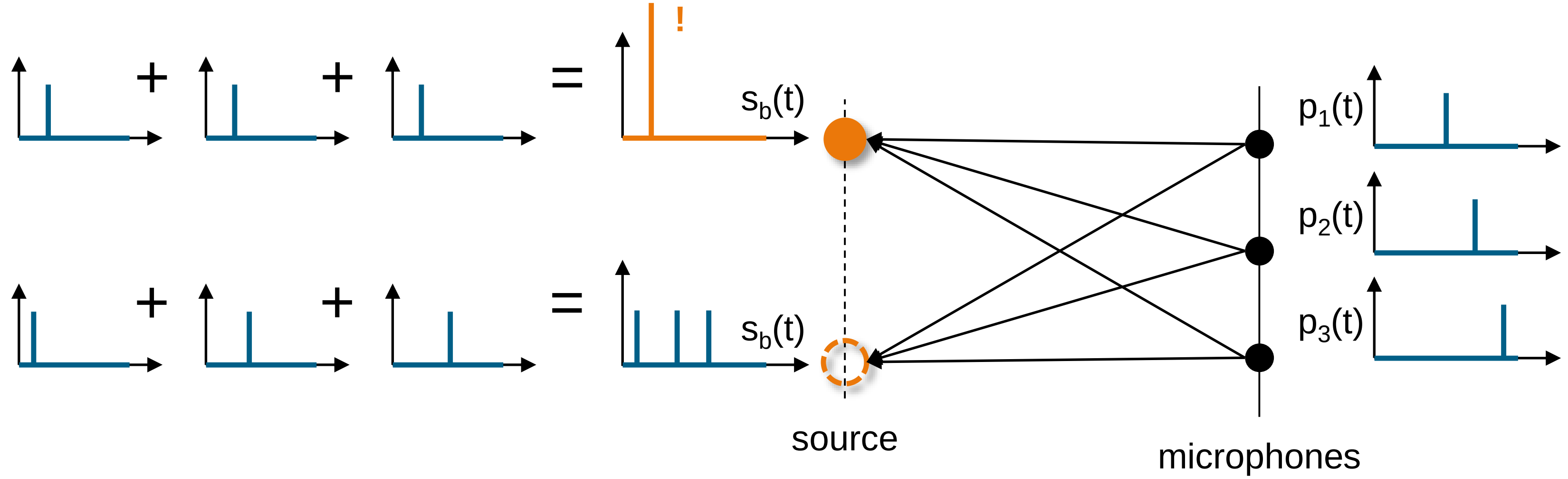


Propagation of source signal:

$$p_i(t) = s(t + \tau_i)$$

Back-propagation:

$$s_b(t) = \sum_i p_i(t - \tau_i)$$



In practice:

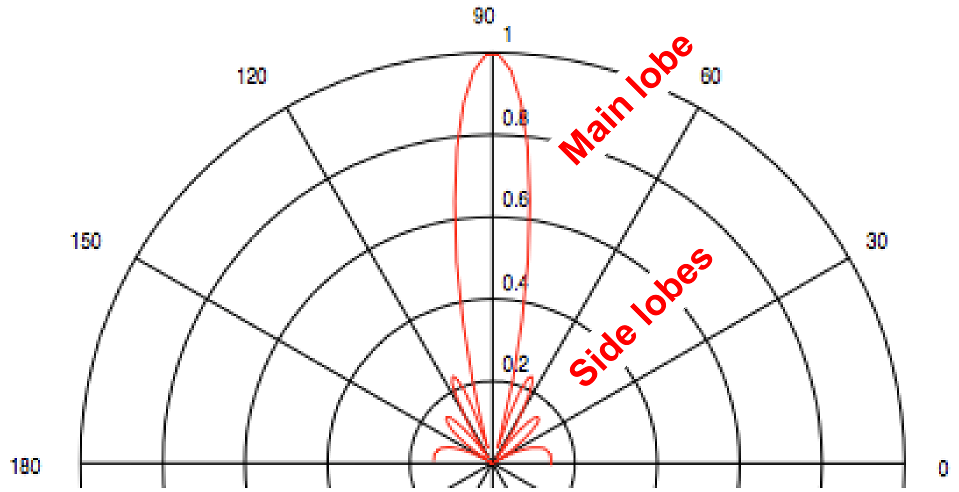
Generalized 'filter-and-sum' method implemented in frequency domain

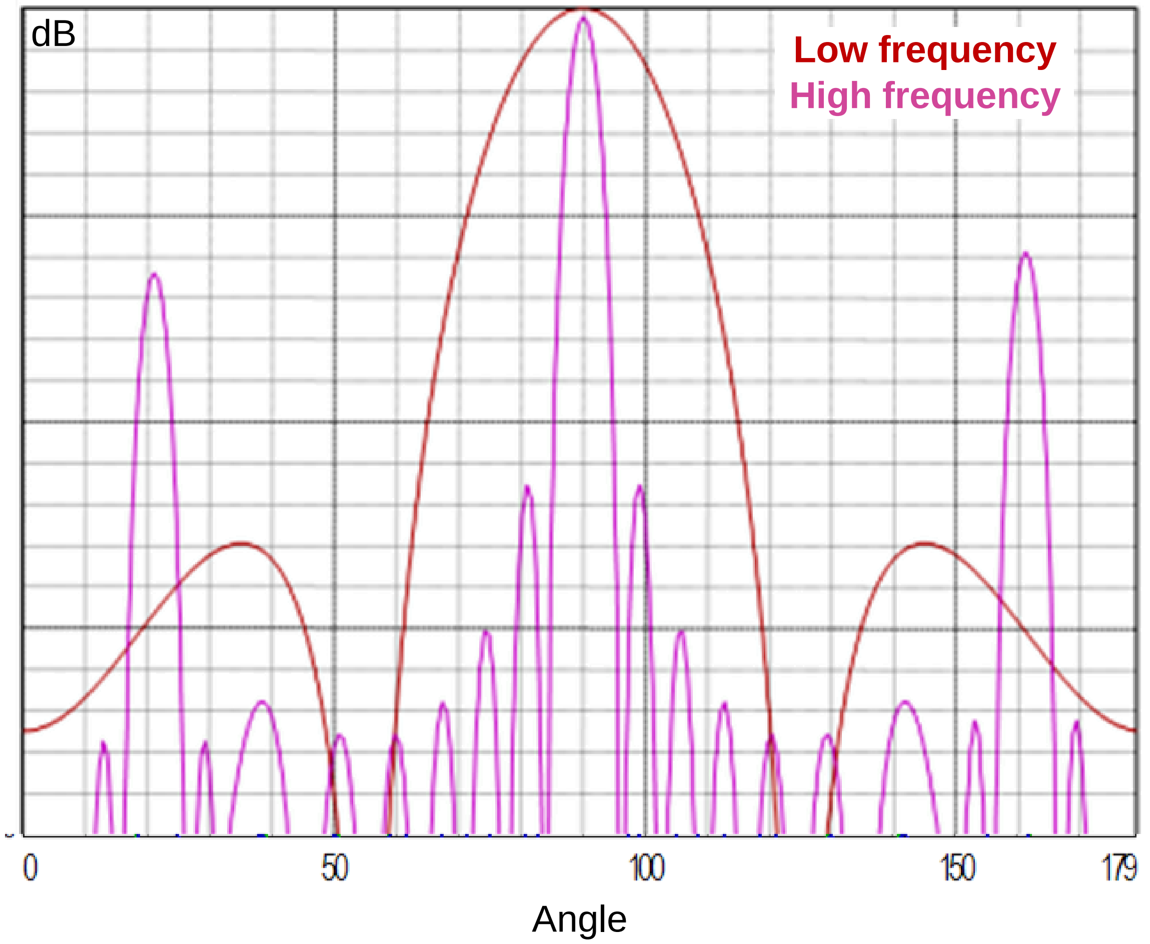
$$s(f) = \sum p_i(f) \cdot w(f) \cdot e^{-2\pi\tau j}$$

- **Source localization per frequency line possible**
- **Weighting function w(f) approximates physical source pressure level**

Restricted © Siemens AG 2018

Beamforming
What influences localization quality?



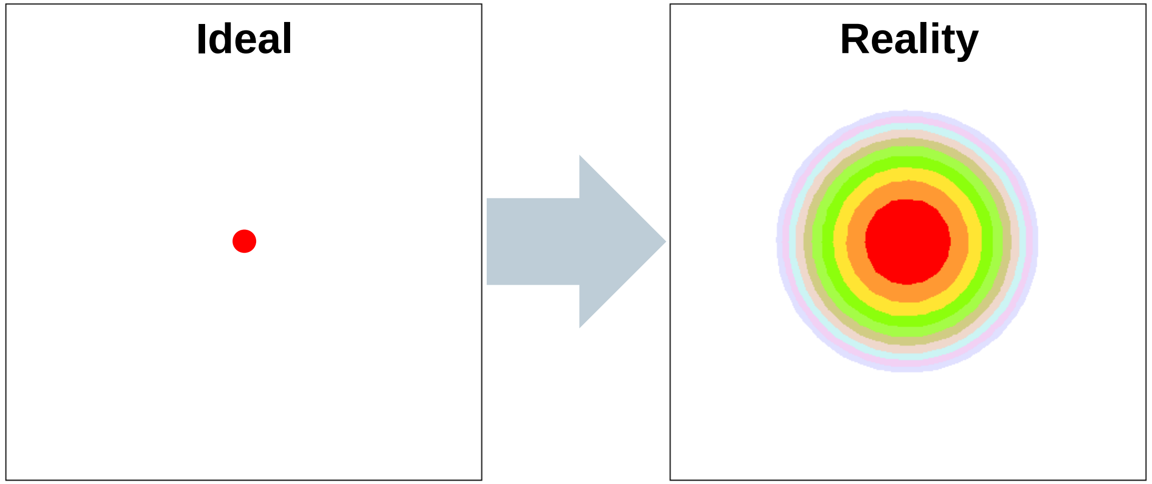


- Delay-and-sum processing has a sensitivity pattern around the main ‘beam’ direction which causes distortion

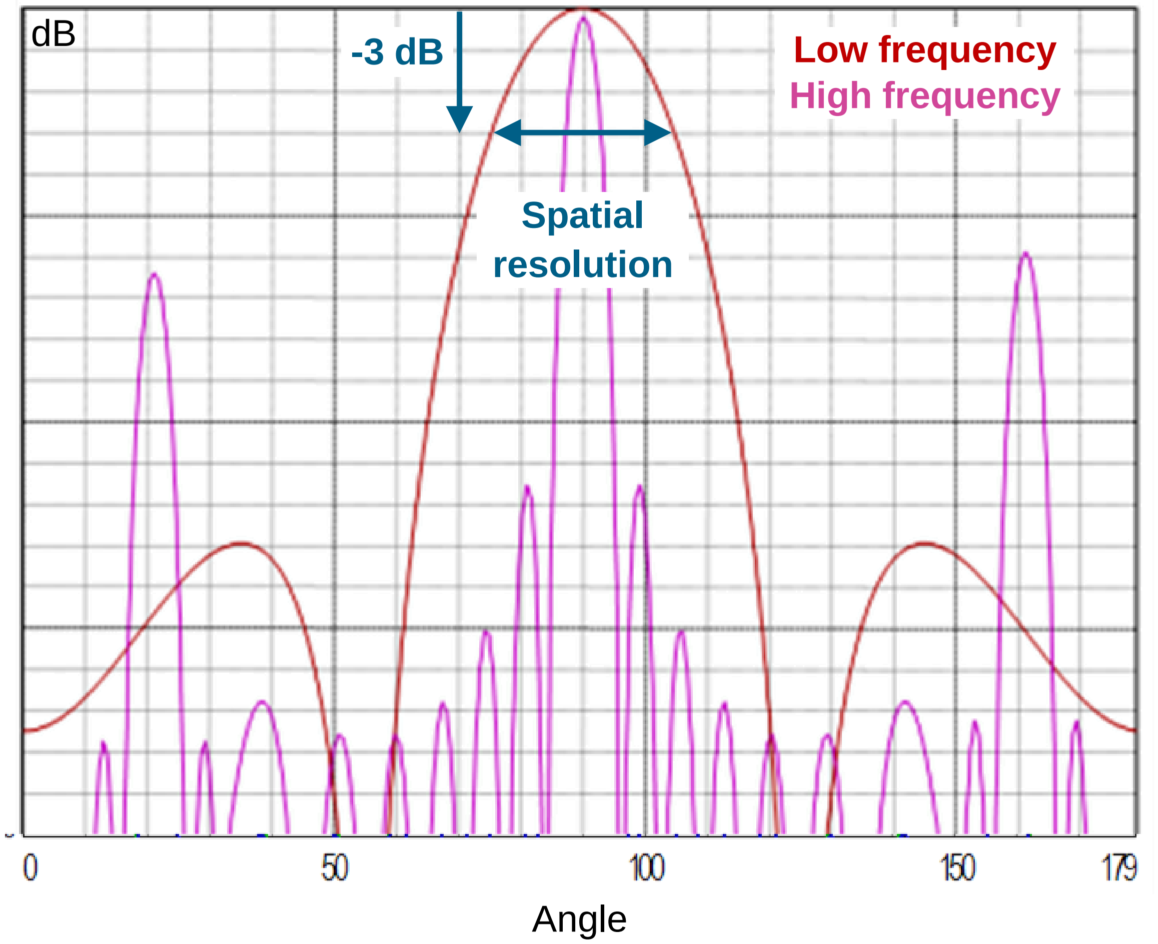
- Factors influencing the sensitivity pattern:
  - Localization technique
  - Analysis frequency
  - Array design

Restricted © Siemens AG 2018

Beamforming
Spatial resolution



Spatial resolution determines localization accuracy



- Beamforming identifies a source **region** with a size depending on the **spatial resolution:**

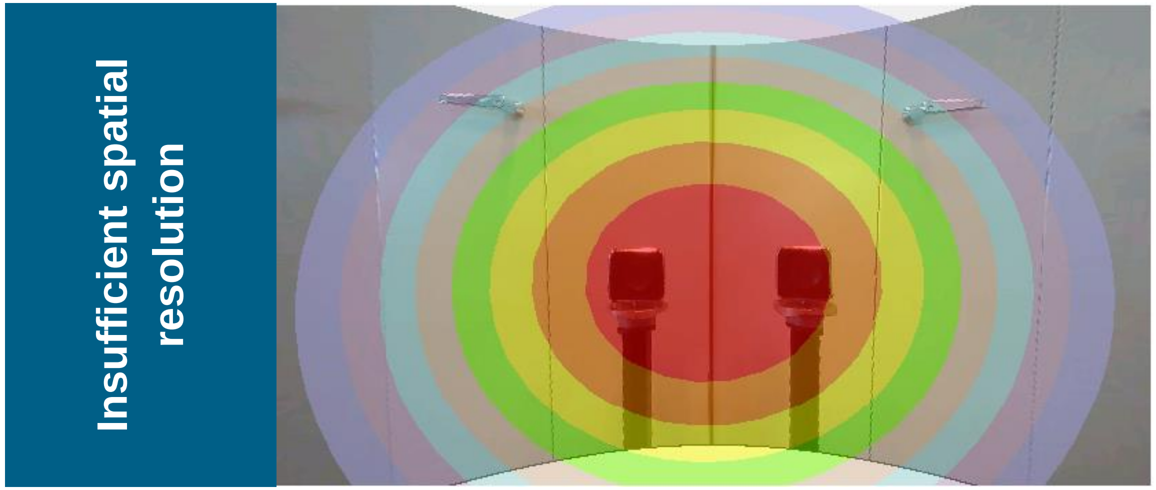
$$Spatial\ resolution \approx \frac{d}{D}\lambda$$

- d = array distance to source
- D = array diameter
- λ = analysis wavelength

Restricted © Siemens AG 2018

Beamforming

Spatial resolution – Practical implications

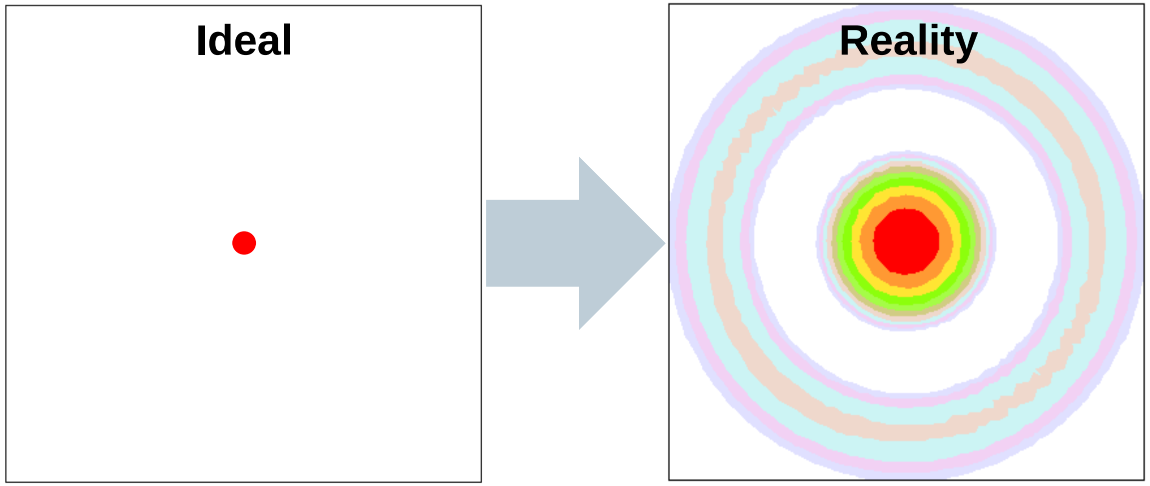


- Sufficient spatial resolution required for:
  - Precise localization of sound source origin
  - Separation of closely spaced sound sources
- Beamforming is very potent for high frequency sound source localization, but suffers in the low frequency range where the wavelengths are large

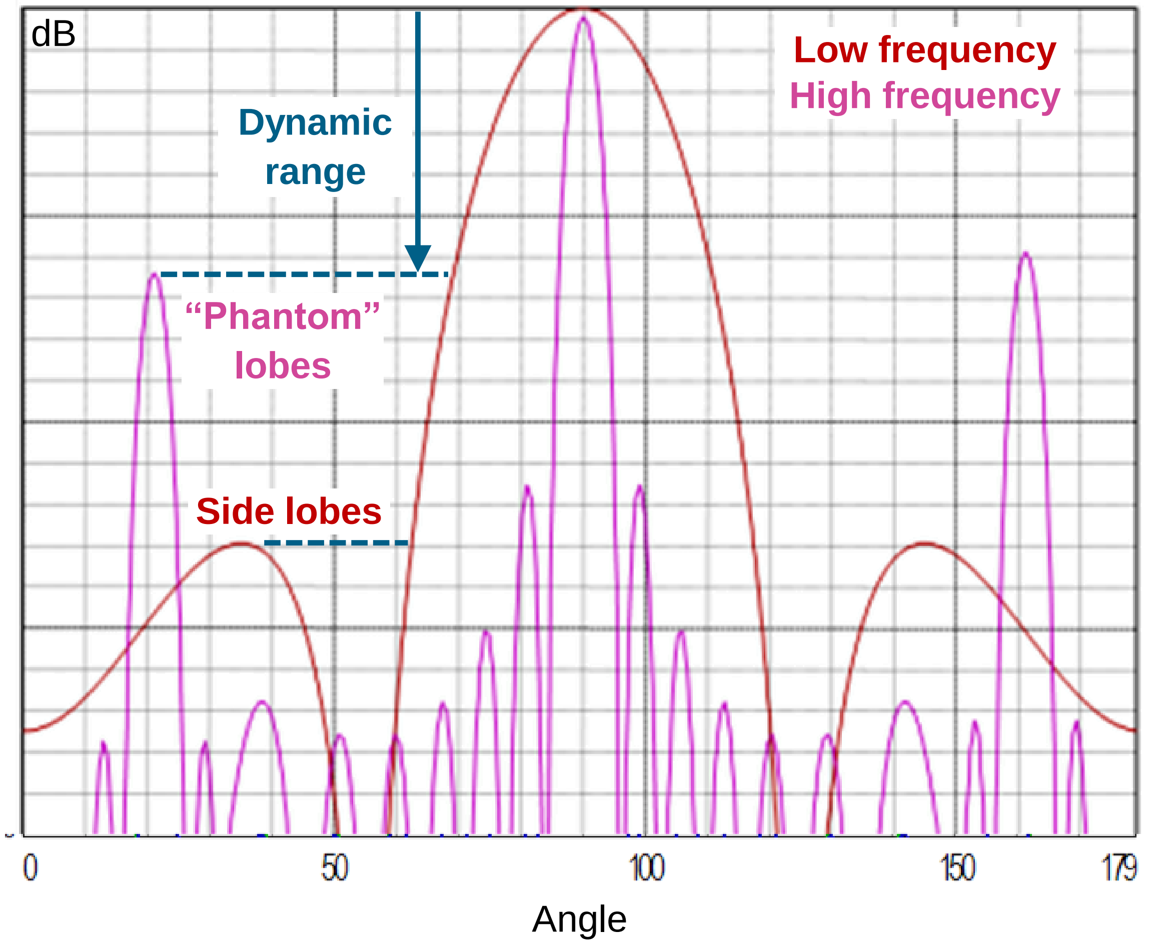
- Spatial resolution can be improved by:
  - Increasing the array **diameter**,
  - Moving **closer** to the source
  - Analyzing **higher frequency** contributions
- Other localization methods can improve the spatial resolution in certain conditions: Focalization, iNAH, Deconvolution, Bayesian Focusing, etc.

Restricted © Siemens AG 2018

Beamforming
Dynamic range



Dynamic range determines localization confidence



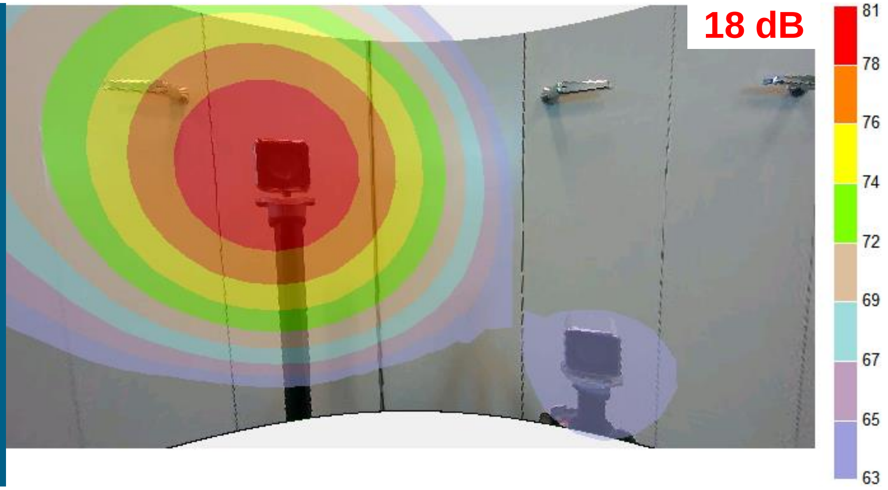
- Side lobes caused by redundancy in phase information between microphone pairs → microphone **layout**
- Phantom lobes caused by spatial undersampling of high frequency wavelengths → microphone **density**
- Dynamic range = zone between main lobe and highest side lobe which is **guaranteed distortion-free**

Restricted © Siemens AG 2018

Beamforming
Dynamic range – Practical implications

Low dynamic range

High dynamic range

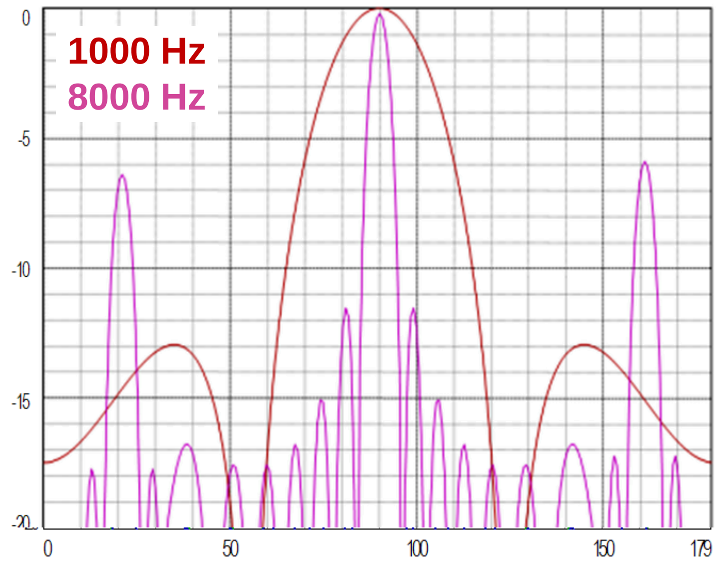


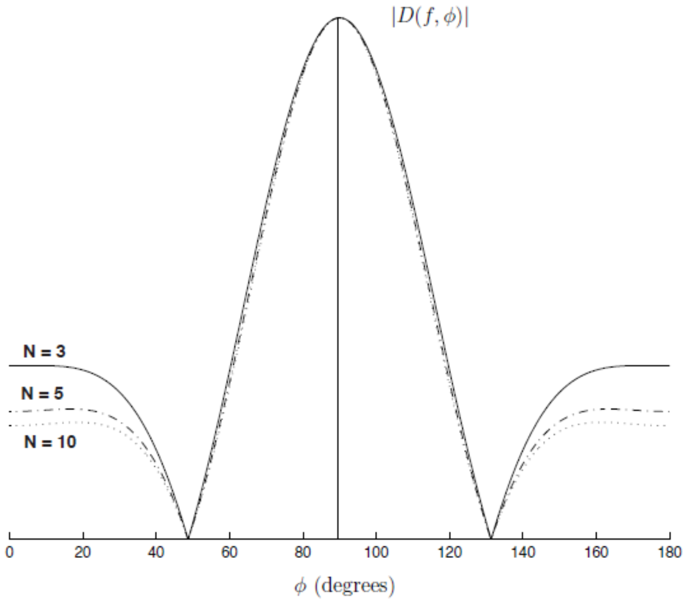
- Dynamic range applies **relative** to dominant source level
- Important aspect for sound source localization for **engineering** purposes:
  - Confidence in sources found within dynamic range
  - Ability to localize secondary (non-dominant) sources

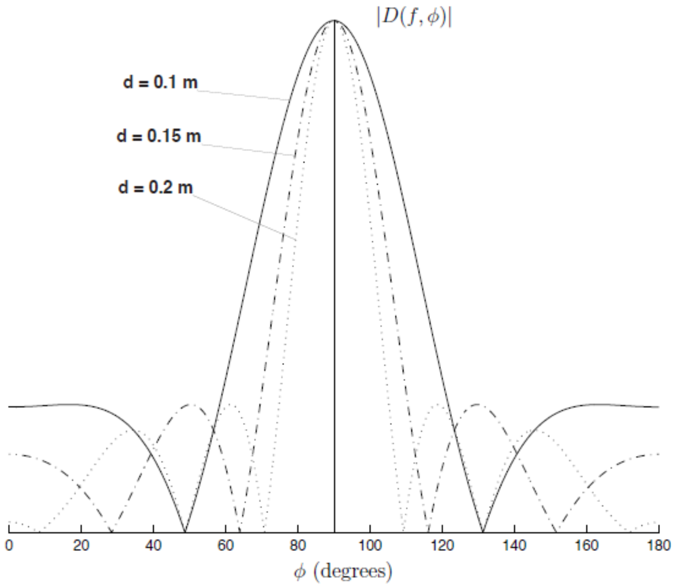
- Dominant factor in determining dynamic range is the **array design:**
  - Microphone layout
  - Array size
- Dynamic range is a **quality indicator** of a given array

Restricted © Siemens AG 2018

Beamforming
Influence of array design parameters







Increase frequency:
- Main lobe width smaller
- Phantom lobes appear at very high frequencies

Increase #mics *N*:
- Side lobe levels reduce
- Array size equal, so spacing *d* decreases

Increase array size:
- Main beam width reduces
- #mics *N* equal, so spacing *d* increases

Restricted © Siemens AG 2018

Focalization
A variant to Beamforming

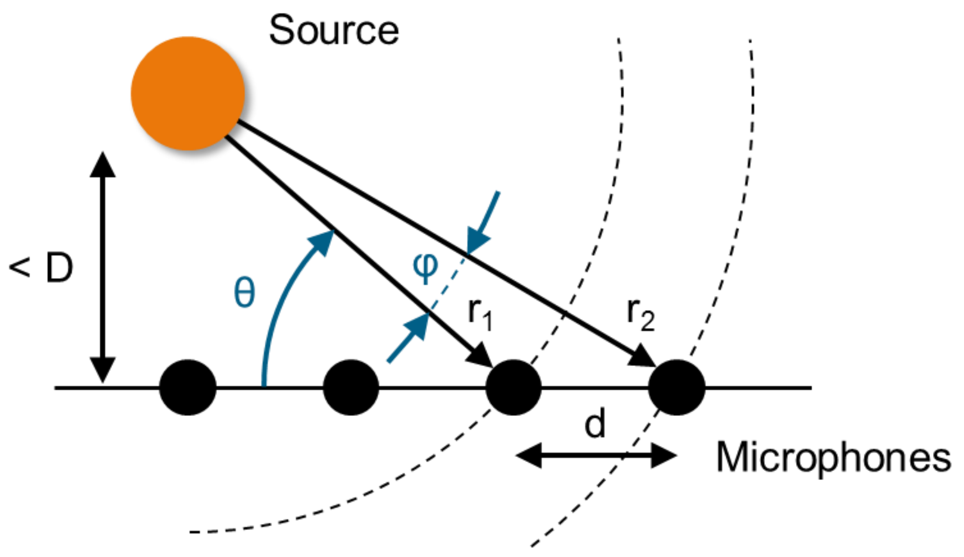
SIEMENS
Ingenuity for life

- Planar pressure wave assumption not valid in **near-field**
- When distance to source < array diameter D, pressure waves are better approximated as **spherical** waves
- Time delay between microphones can be expressed as:

$$\tau = \frac{r_2 - r_1}{c} = \frac{d \cdot \cos(\theta) \cdot (\cos\left(\frac{\varphi}{2}\right) + \sin(\frac{\varphi}{2}) \cdot \tan(90 - \frac{\varphi}{2}))}{c}$$

- At large source distances ($r_1 \rightarrow \infty$) angle φ becomes small ($\varphi \rightarrow 0$) and Focalization becomes more similar to classical Beamforming:

$$\left.\begin{array}{c} \sin\left(\frac{\varphi}{2}\right) \cdot \tan\left(90 - \frac{\varphi}{2}\right) \rightarrow 0 \\ \cos\left(\frac{\varphi}{2}\right) \rightarrow 1 \end{array}\right\} \tau \rightarrow \frac{d \cdot \cos(\theta)}{c}$$



Practical benefit:
Spatial resolution improves by factor ~2

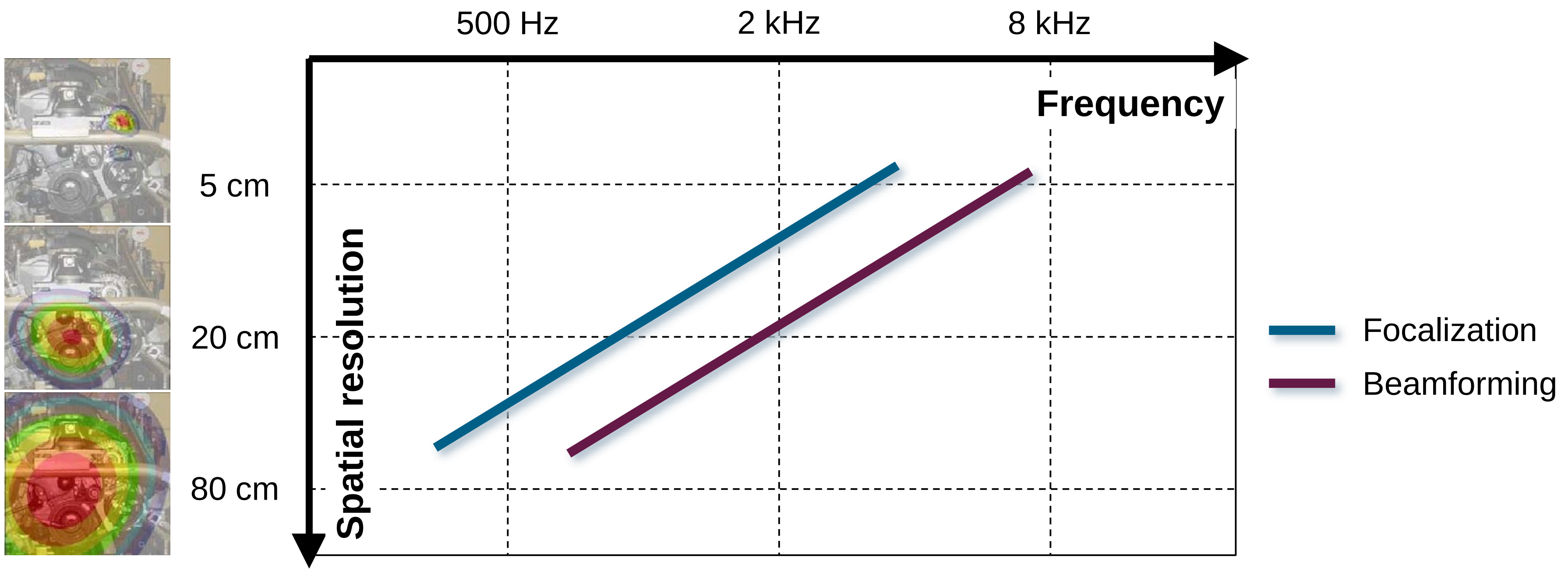
Restricted © Siemens AG 2018

Beamforming & Focalization
Summary

Beamforming spatial resolution ≈ λ (ideal)

Focalization spatial resolution ≈ ½ λ (ideal)

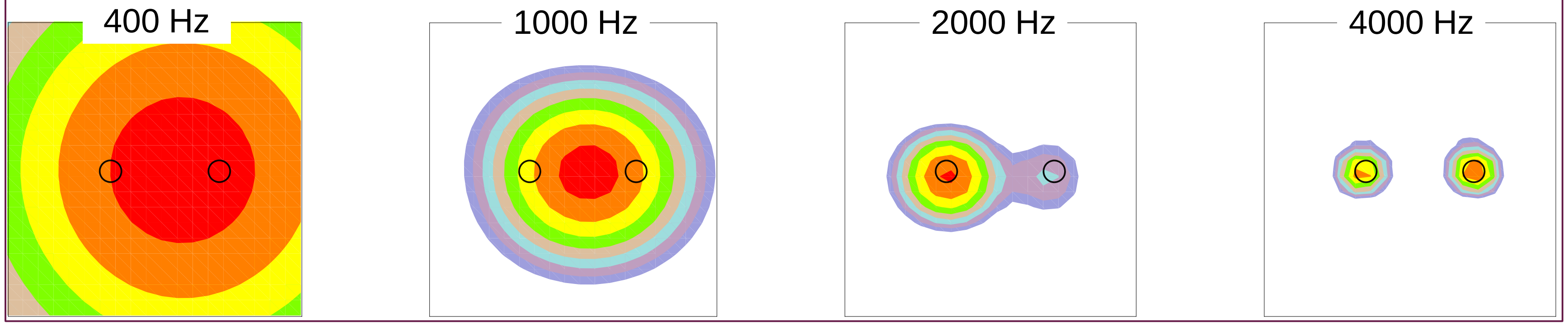
- Many factors have been seen to influence real-world localization performance
- Best results obtained at **high analysis frequencies** and **short measurement distances**



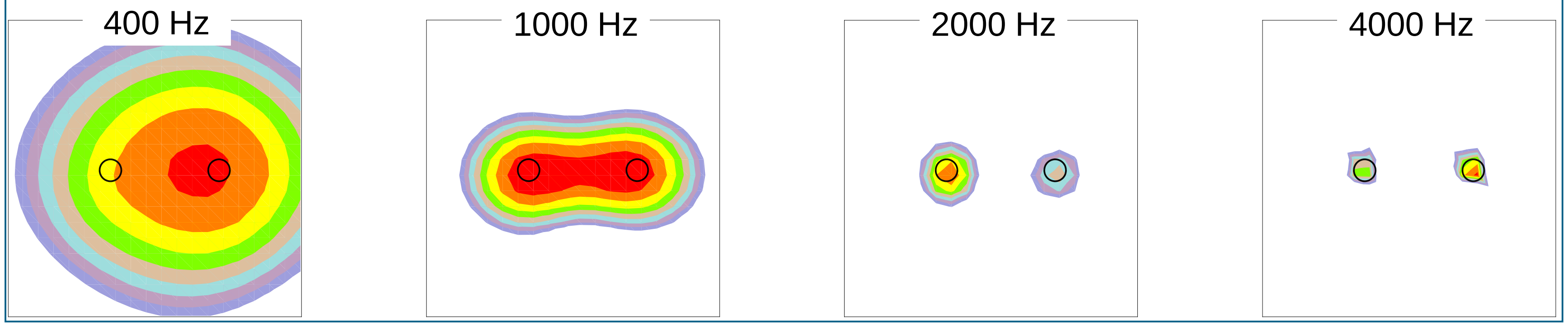
Restricted © Siemens AG 2018

Beamforming & Focalization
Synergy

Beamforming @ 40 cm



Focalization @ 7 cm



Best results obtained with combination of Beamforming and Focalization

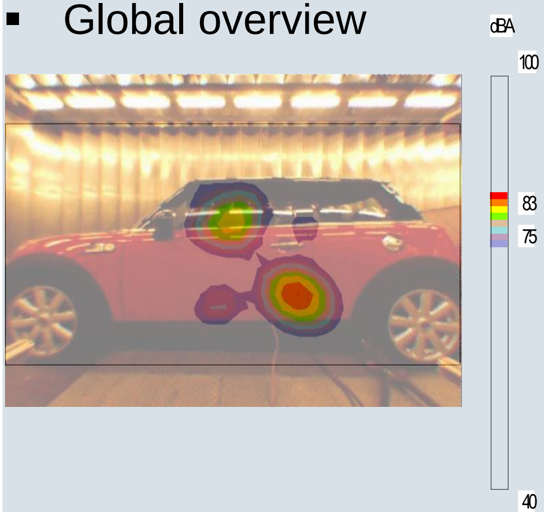
Restricted © Siemens AG 2018

Beamforming & Focalization
Synergy – Practical example

Where does wind noise leak into this vehicle?

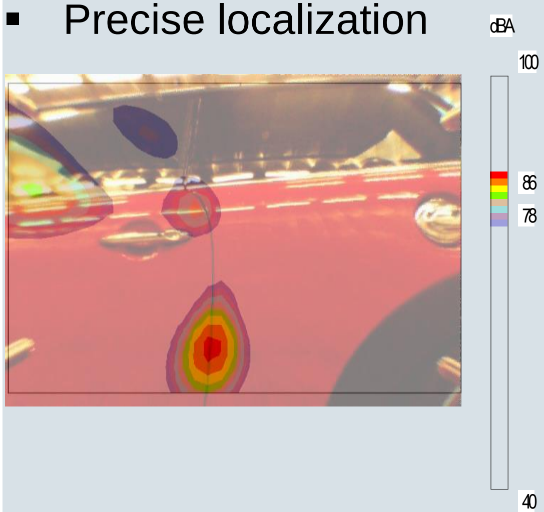
1. Beamforming

- Far-field technique
- Low resolution
- Global overview



2. Focalization

- Near-field technique
- High resolution
- Precise localization



Restricted © Siemens AG 2018

Beamforming & Focalization

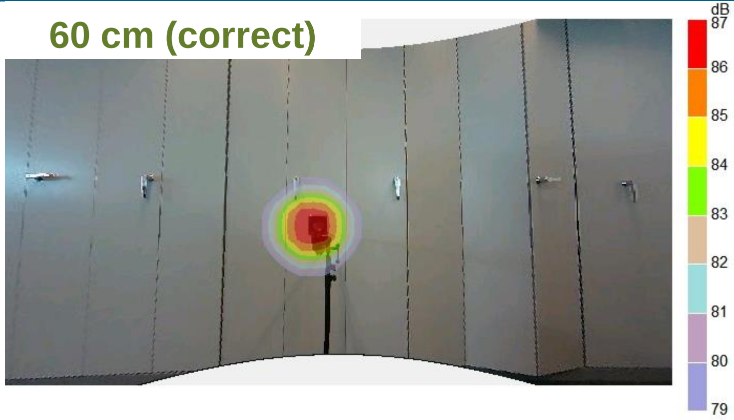
Measurement distance vs. calculation distance

Far-field: time delay dominated by distance between microphones → minor influence by incorrect distance

30 cm (too close)

Grid step: 0.014 / 0.014 - Grid size: 55 / 31 - distance :0.300 - [2828 - 3563 Hz]

60 cm (correct)



Grid step: 0.029 / 0.029 - Grid size: 55 / 31 - distance :0.600 - [2828 - 3563 Hz]

100 cm (too far)

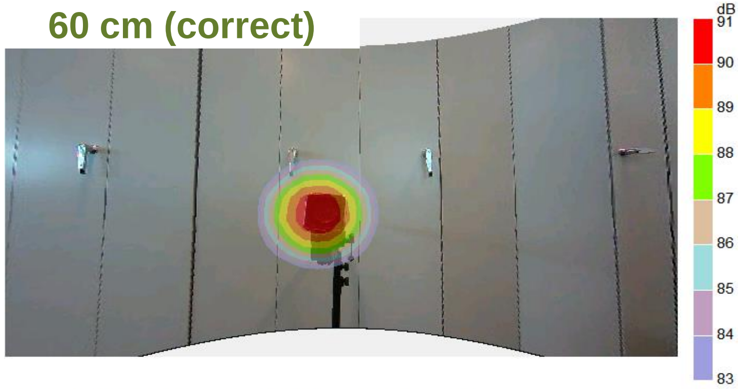
Grid step: 0.048 / 0.048 - Grid size: 55 / 31 - distance :1.000 - [2828 - 3563 Hz]

Near-field: spherical waves sensitive to propagation distance → major influence by incorrect distance

30 cm (too close)

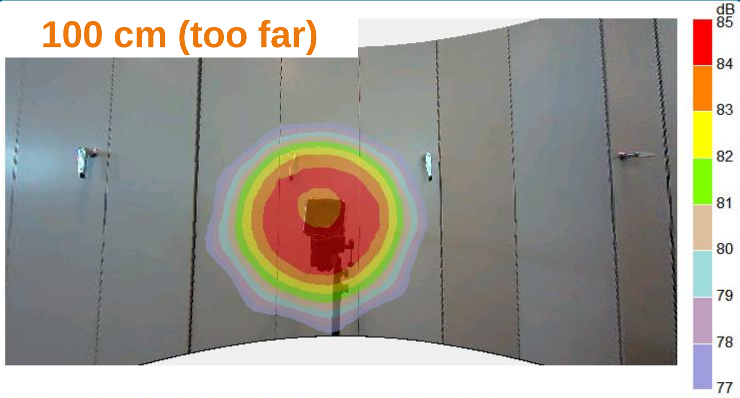
Grid step: 0.005 / 0.005 - Grid size: 55 / 31 - distance :0.100 - [2828 - 3563 Hz]

60 cm (correct)



Grid step: 0.014 / 0.014 - Grid size: 55 / 31 - distance :0.300 - [2828 - 3563 Hz]

100 cm (too far)



Grid step: 0.048 / 0.048 - Grid size: 55 / 31 - distance :1.000 - [2828 - 3563 Hz]

Restricted © Siemens AG 2018

Fundamentals of Sound Source Localization
Agenda

- Introduction to Sound Source Localization
- Beamforming-based localization methods
- **Far-field Deconvolution methods**
- Map averaging methods

Restricted © Siemens AG 2018

Far-field Deconvolution
What is it?

SIEMENS
Ingenuity for life

What?

- **Iterative & quantitative** source localization method
- Increases spatial resolution vs. Beamforming-based methods by **factor ~4 to 5**, and enables **objective source ranking** via Sound Power estimations

Why?

- Beamforming underperforms in cases with **multiple sources** in **close proximity** at **large distances**
- Alternative methods not always applicable in those scenario's (e.g. near-field Holography)

How?

- Assume discrete **monopole source distribution** with **uncorrelated** spherical sound radiation
- Optimal monopole distribution to match measured Beamforming result found via **iterative optimization**

When?

- **Far-field**, distance to source > array diameter
- Localization only: **low to mid** frequency range
- Quantification: **low to high** frequency range
- **Uncorrelated** sources only (e.g. aero-acoustics)

Restricted © Siemens AG 2018

Far-field Deconvolution

How does it work?

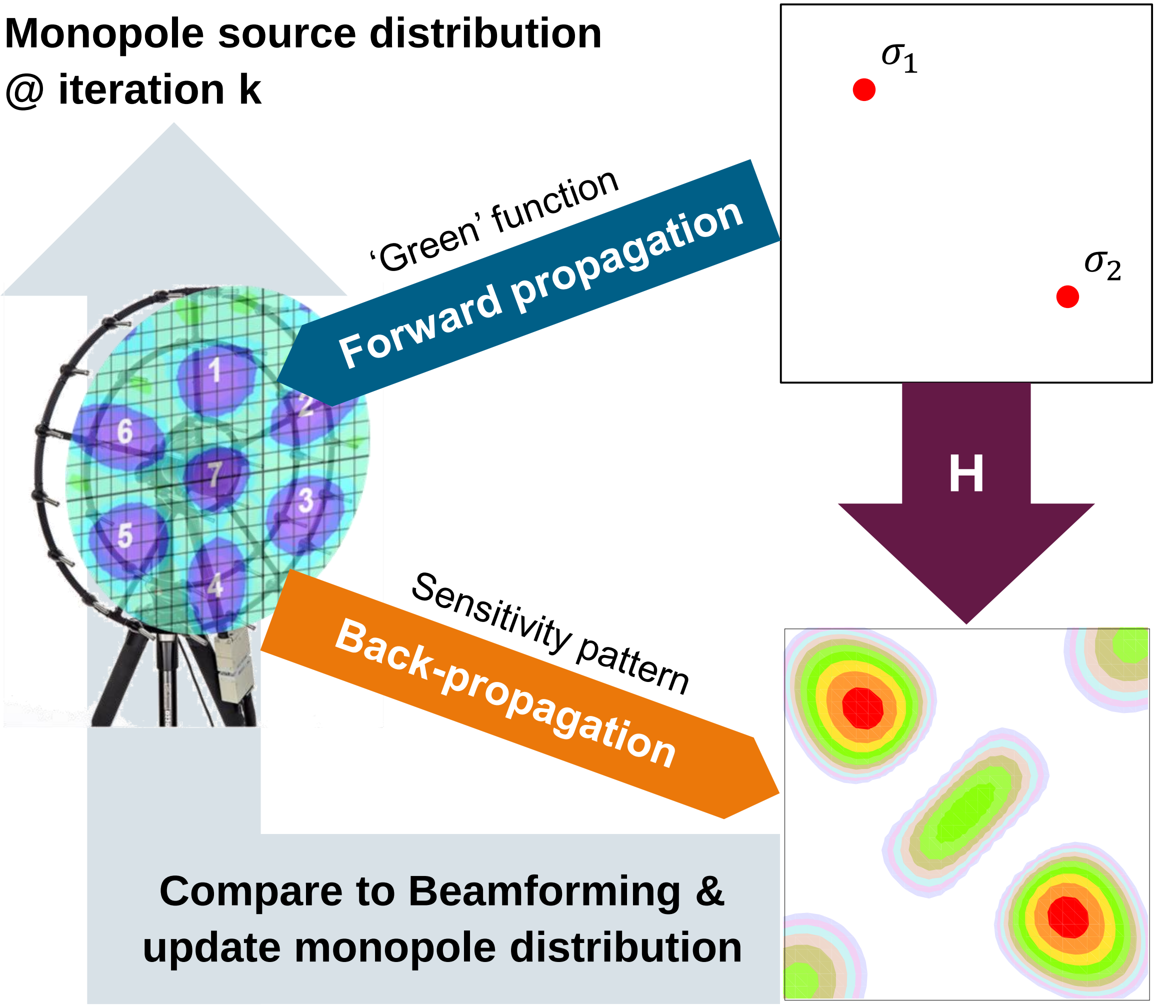
- Find the monopole distribution for which the **predicted** Beamforming result matches the **measured** result

- Mathematically expressed as the minimization of the following cost function (**CIRA** method):

$$\min_{\sigma,\sigma_i \geq 0} C(\sigma) = \sum_i \left(\sum_j (H_{ij} \cdot \sigma_j) - B_i \right)^2$$

- σ = distribution of monopole sources $\sigma_{1..J}$
- H_{ij} = transformation matrix from monopole @ point j to Beamforming result @ point i
- B_i = measured Beamforming result @ point i

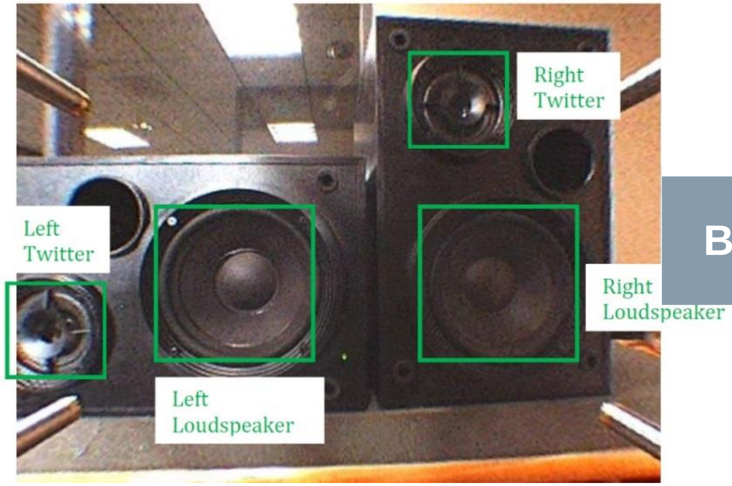
Monopole source model dramatically improves spatial resolution and allows Sound Power estimation!



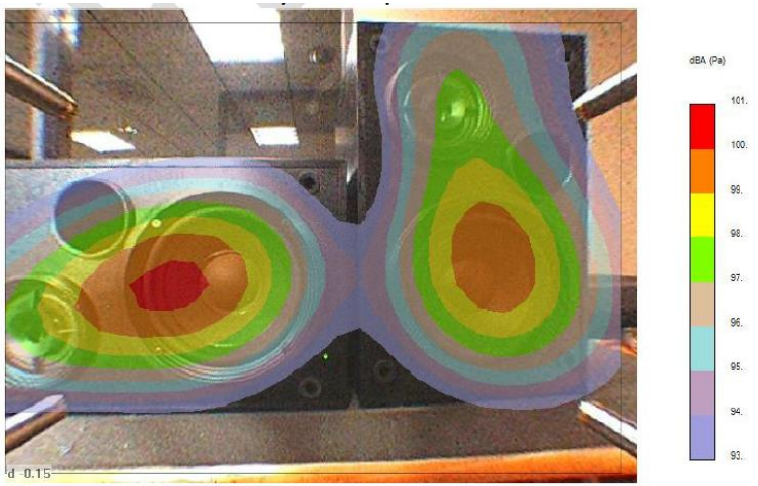
Restricted © Siemens AG 2018

Far-field Deconvolution

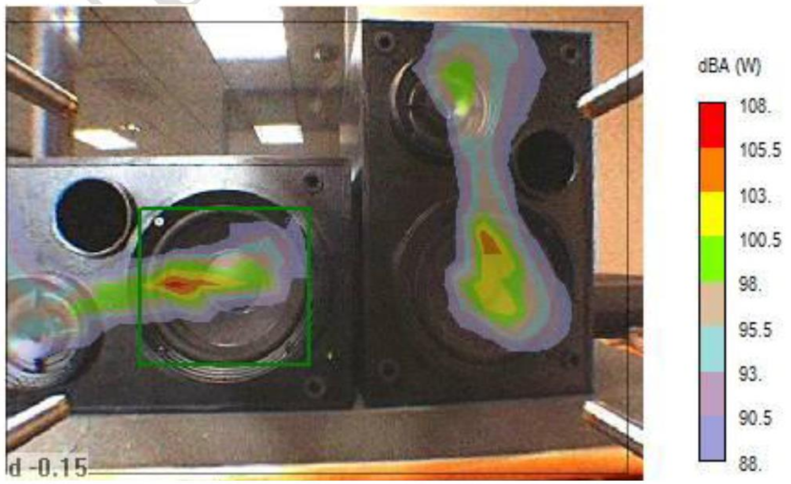
CIRA – Practical examples



Beamforming (1 – 2 kHz)



CIRA (1 – 2 kHz)

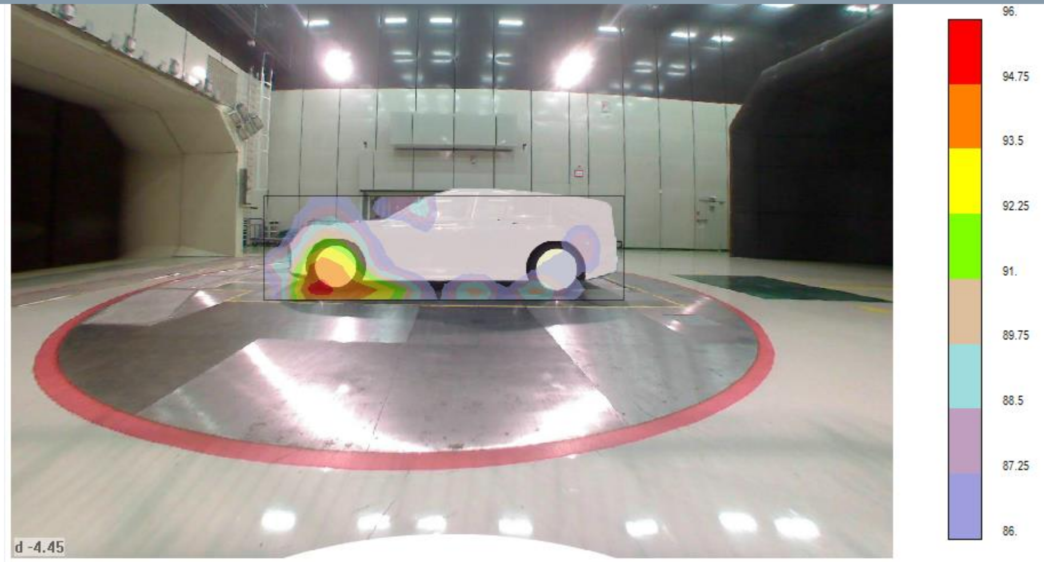


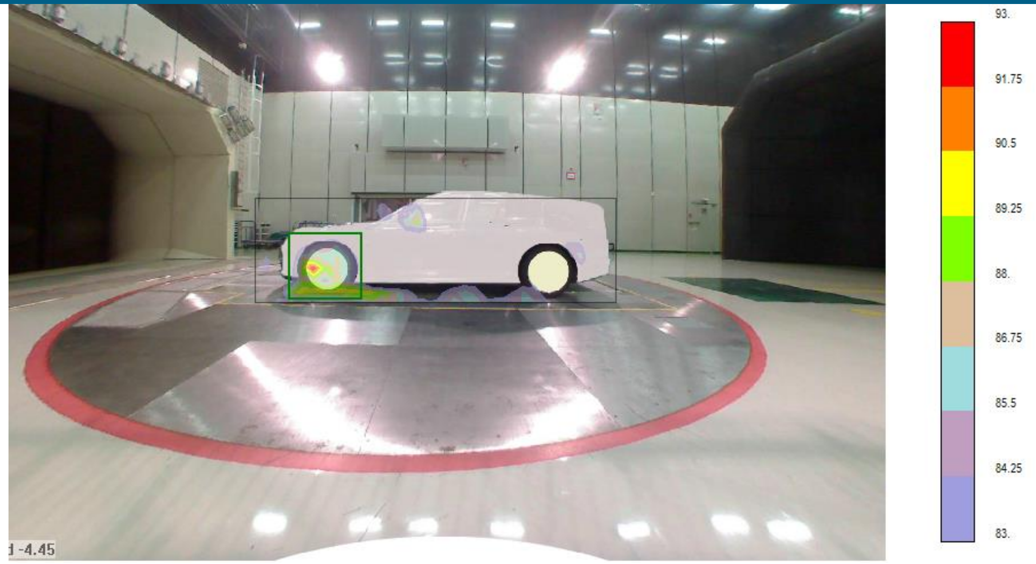
- Dramatic improvement of **spatial resolution**
- Dramatic improvement of **dynamic range** (8 dB → 20 dB)
- Quantified results in **[dBW]**

Restricted © Siemens AG 2018

Far-field Deconvolution
CIRA – Practical examples

Beamforming	CIRA





- Aero-acoustic sources in wind tunnel @ 4.45 meters
- Beamforming provides high-confidence but low-accuracy initial guess on source positions

- Confirms Beamforming result and improves spatial resolution
- Sources quantified → allows objective comparison

Restricted © Siemens AG 2018

Far-field Deconvolution
CIRA vs. Clean-SC

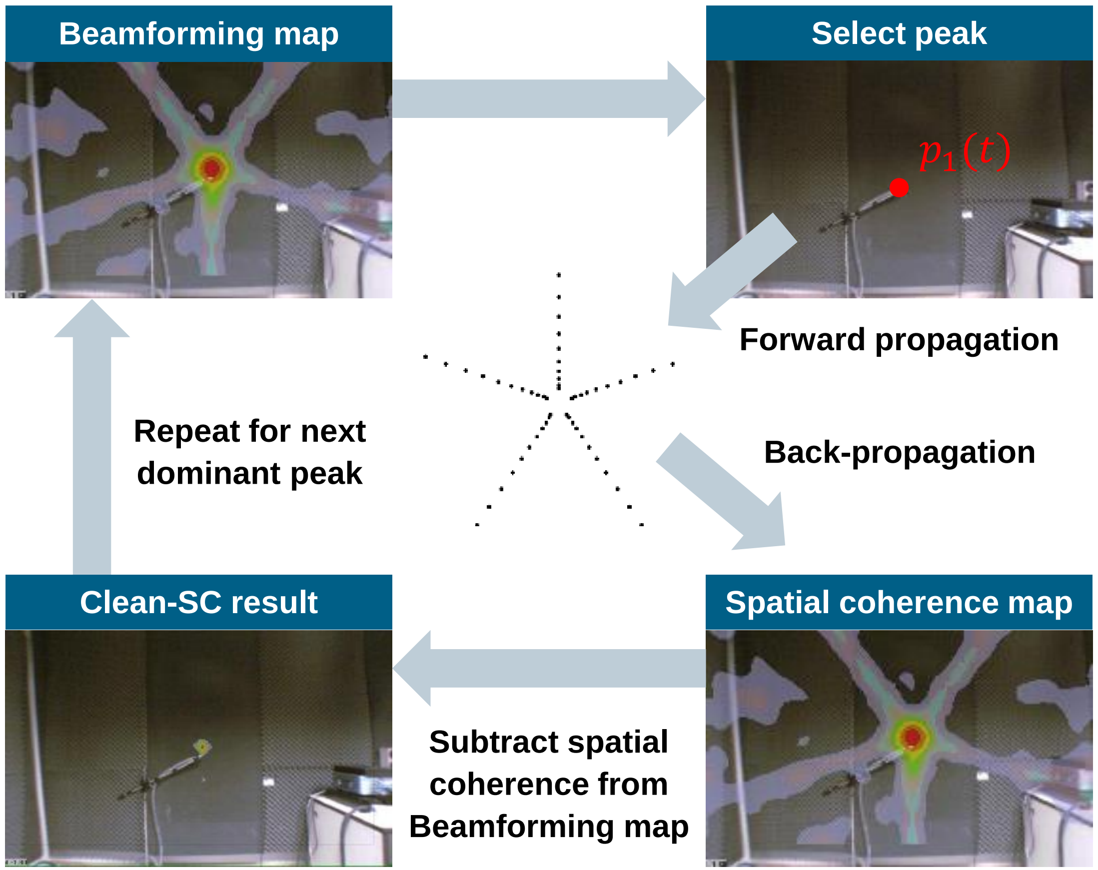
Iterative Least-squares

- All monopoles are processed in the same loop
- Finds optimal monopole distribution through iterative optimization of cost function (previous slides)
- **CIRA** in Testlab: improved convergence criteria

Iterative Cleaning

- Each monopole is processed in a separate loop
- Finds optimal monopole position for dominant source in current loop, then removes coherent components
- **Clean-SC** in Testlab: based on spatial coherence

Clean-SC procedure



Restricted © Siemens AG 2018

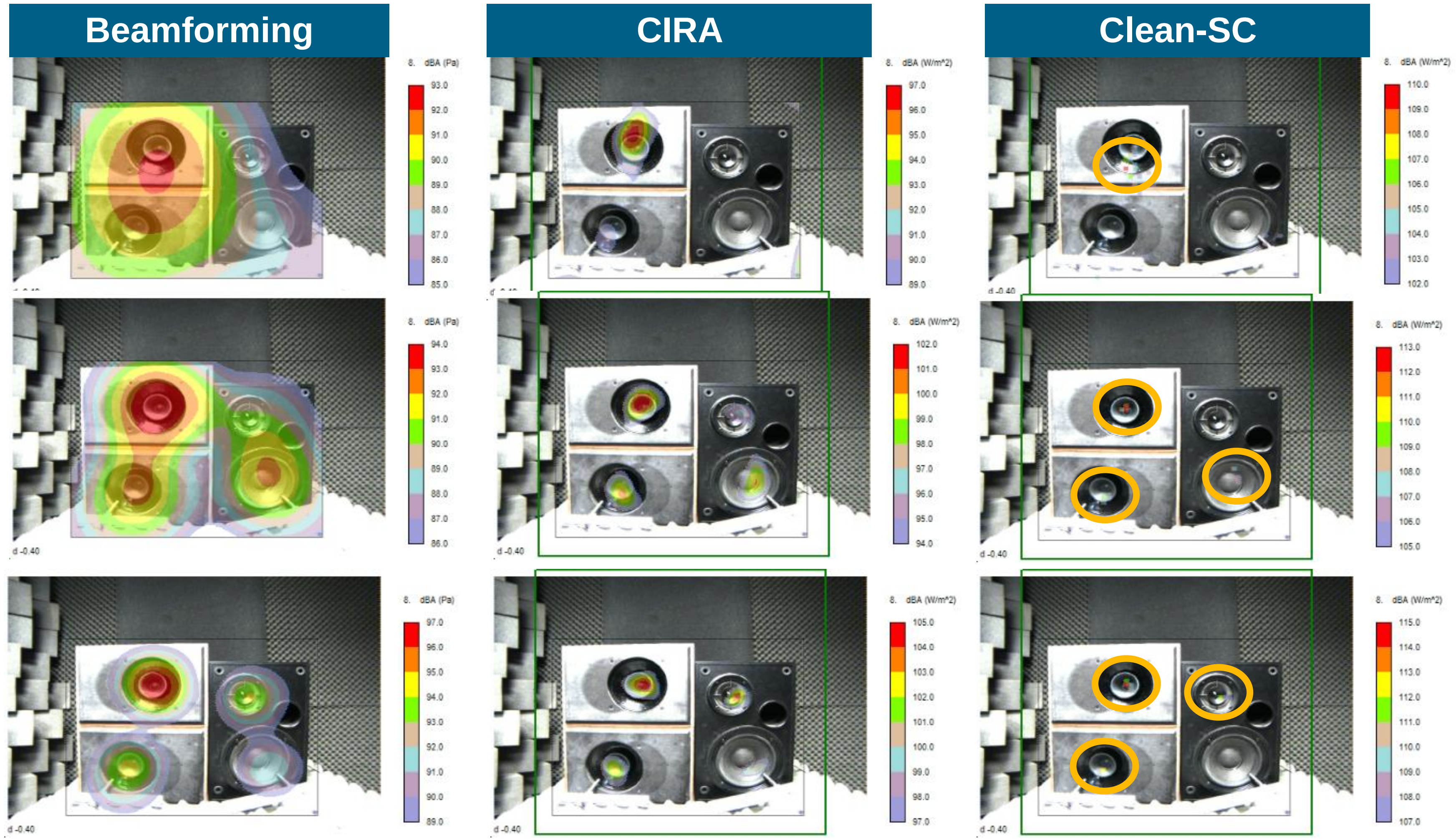
Far-field Deconvolution

CIRA vs. Clean-SC – Practical example

900 – 1120 Hz

1420 – 1780 Hz

2250 – 2820 Hz



Far-field Deconvolution

CIRA vs. Clean-SC – Summary

SIEMENS
Ingenuity for life

CIRA	Clean-SC
Advantages vs. Beamforming ▪ Improves dynamic range ▪ Improves spatial resolution (reconstruction with monopoles) ▪ Improves source separation capability ▪ Quantitative results in Sound Power	**Advantages vs. Beamforming** ▪ Strongly improves dynamic range ▪ Improves the spatial resolution (reconstruction with monopoles) ▪ Quantitative results in Sound Power ▪ Very fast calculation time for Deconvolution method
Disadvantages ▪ Only valid for uncorrelated sources, sensitive to correlation between sources ▪ Calculation time of multi-variable optimization	**Disadvantages** ▪ Only valid for uncorrelated sources, **very** sensitive to correlation between sources ▪ Does not improve source separation capability

Restricted © Siemens AG 2018

Fundamentals of Sound Source Localization
Agenda

- Introduction to Sound Source Localization
- Beamforming-based localization methods
- Far-field Deconvolution methods
- **Map averaging methods**

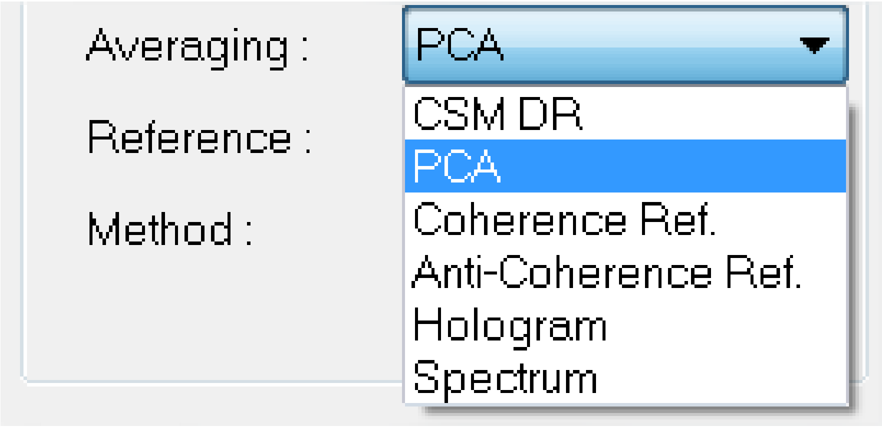
Restricted © Siemens AG 2018

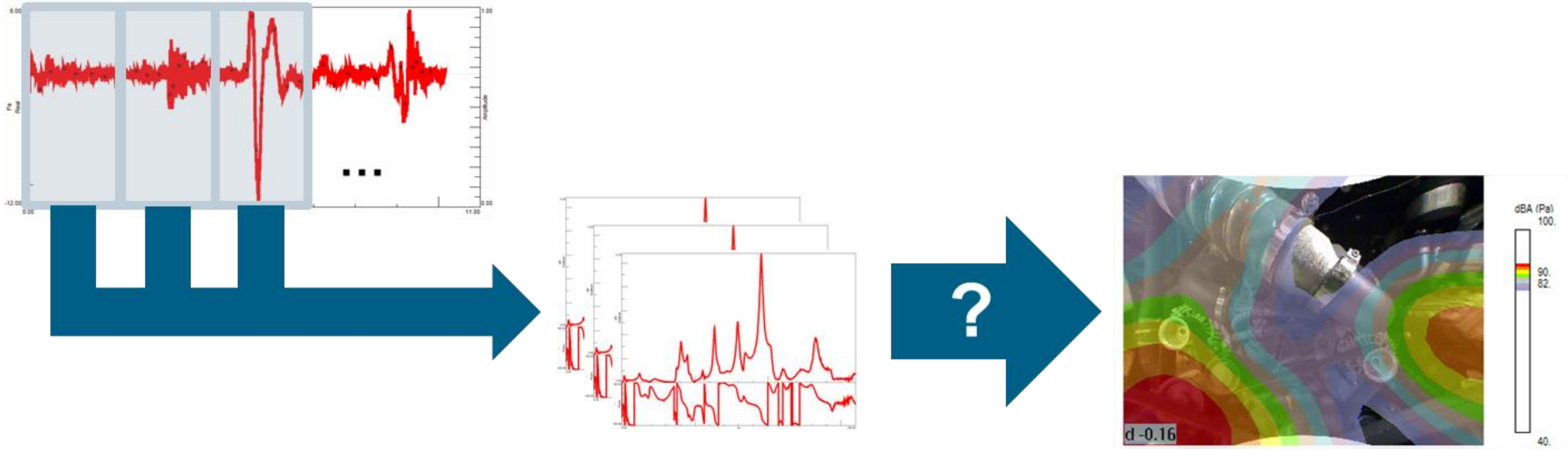
Map averaging methods
How do we handle time-domain averaging?

SIEMENS
Ingenuity for life

- One time block per microphone is sufficient to calculate one hologram result

- Different strategies exist for combining holograms from multiple time blocks, depending on the source type:
  - Stationary vs. transient
  - Correlated vs. uncorrelated





Restricted © Siemens AG 2018

Map averaging methods
Spectrum averaging

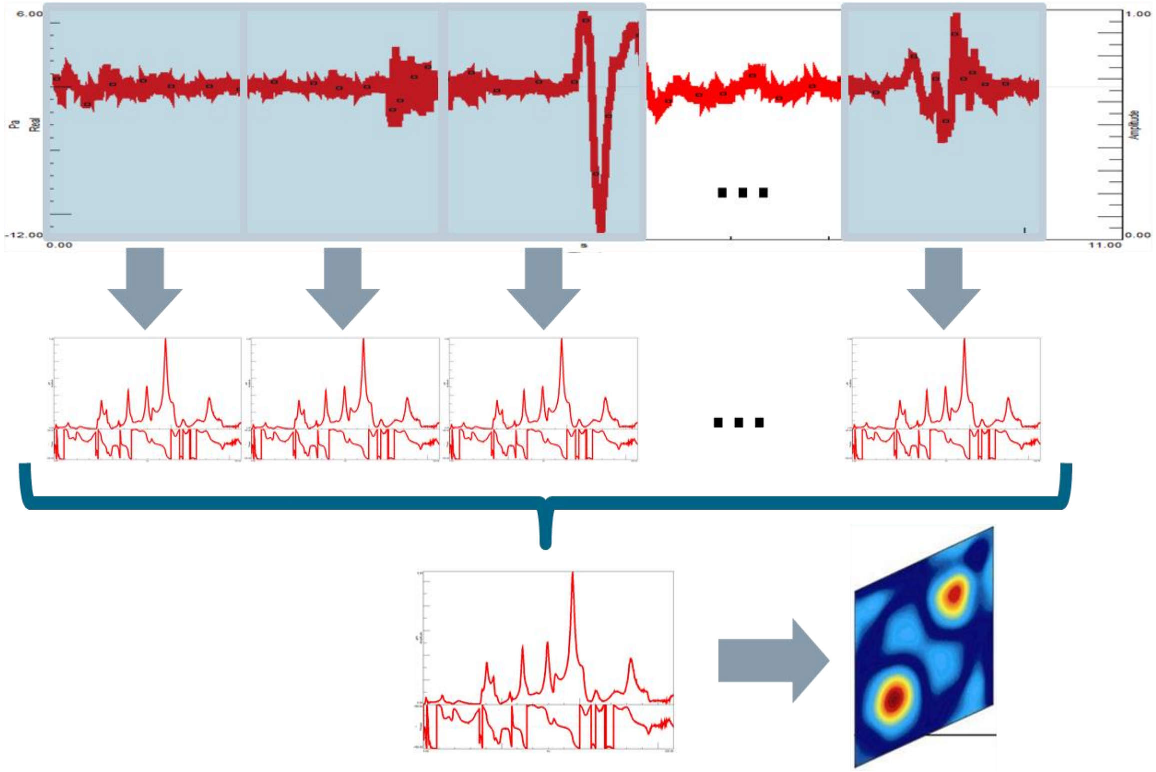
Process

1) Calculate averaged spectrum over all time blocks
2) Calculate hologram from averaged spectrum

Usage

- Extremely fast calculation time → animations!
- Requires phase reference signal
- Requires stationary and correlated sources to avoid information loss during averaging

Spectrum averaging

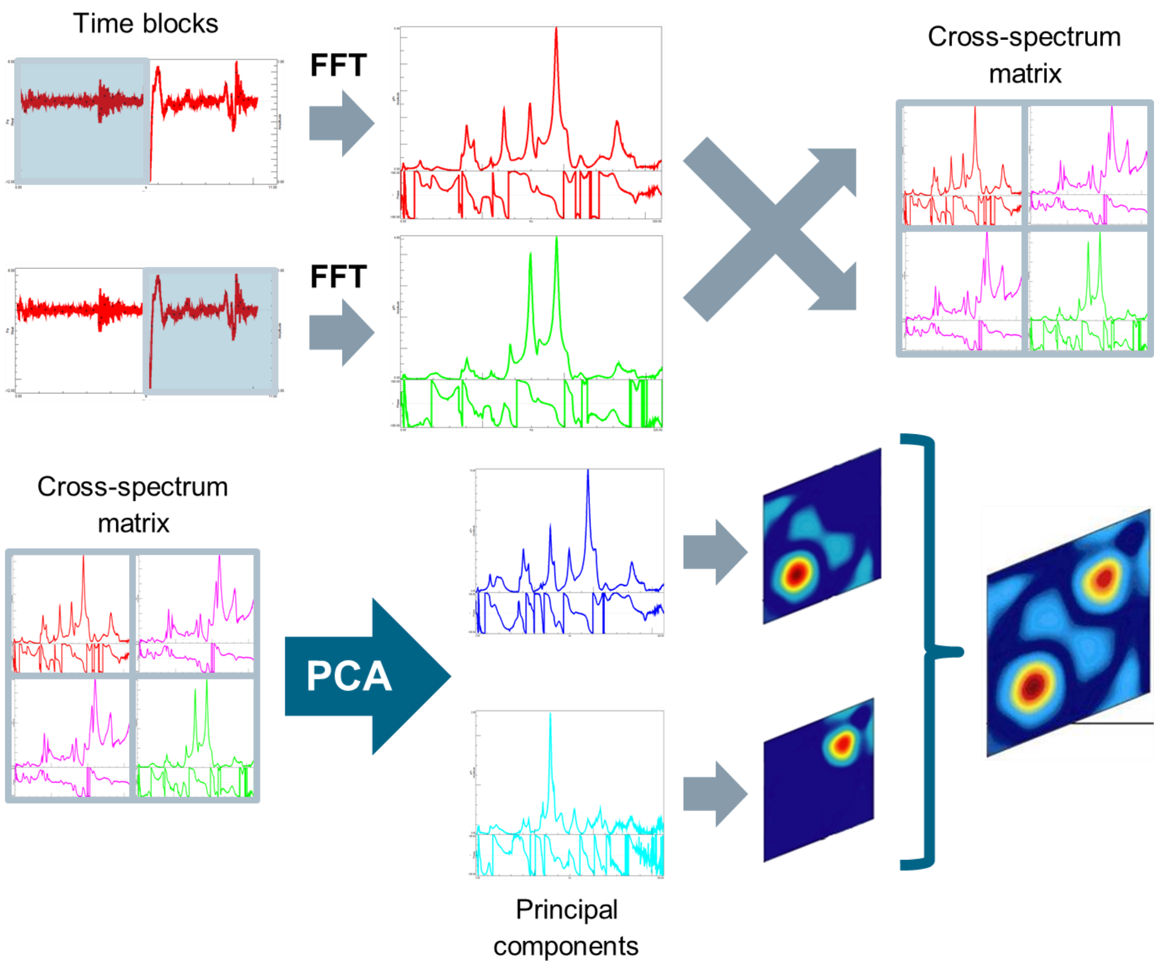


Map averaging methods
Principal Components Averaging

Process
1) Calculate cross-spectrum matrix between time blocks
2) Apply PCA to CSM → identifies independent sources
3) Calculate spatially averaged hologram over all principal components

Usage
▪ Does not eliminate uncorrelated sources and does not require phase reference signal → general purpose
▪ Only suitable for stationary sources

Principal Components Averaging



Map averaging methods
Hologram averaging

SIEMENS
Ingenuity for life

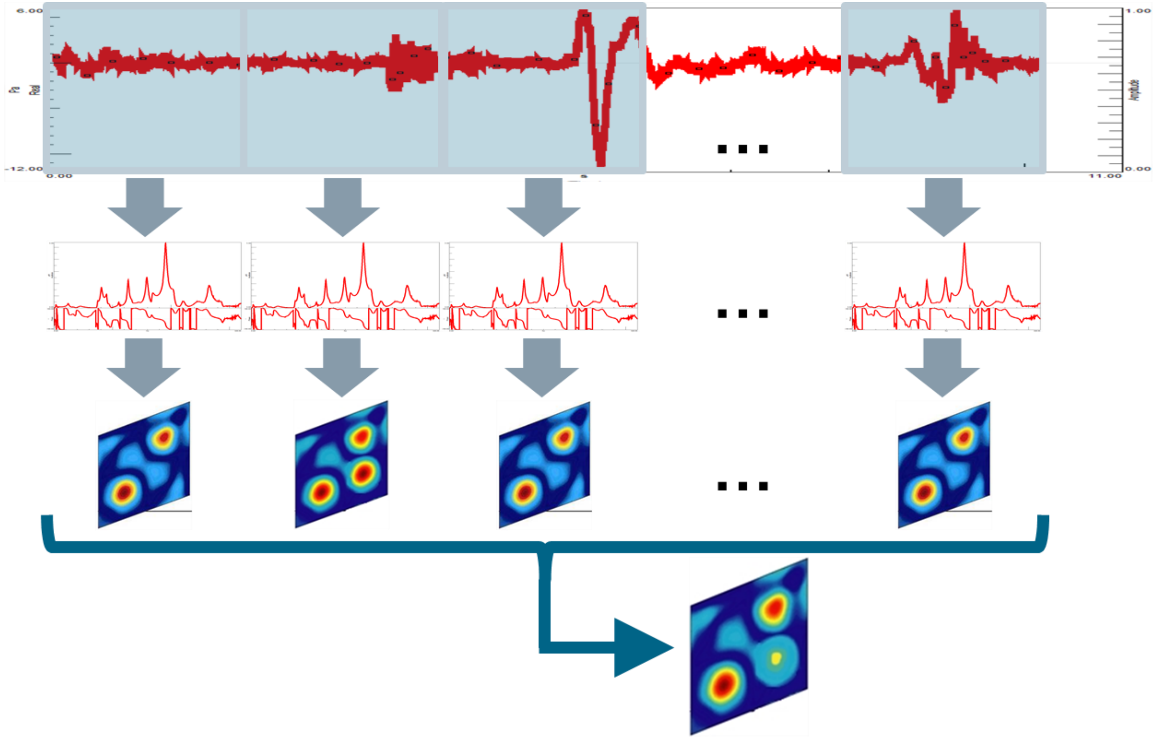
Process

1) Calculate one spectrum per time block
2) Calculate one hologram per spectrum
3) Calculate spatially averaged hologram over all time blocks

Usage

- Averaging process does not completely eliminate transient components
- Does not require phase reference signal
- High computational effort

Hologram averaging



Restricted © Siemens AG 2018

Map averaging methods

Cross-Spectrum Matrix Diagonal Removed averaging

Process

1) Remove diagonal components from CSM in the classical Beamforming expression

2) Apply Hologram averaging to the result

Usage

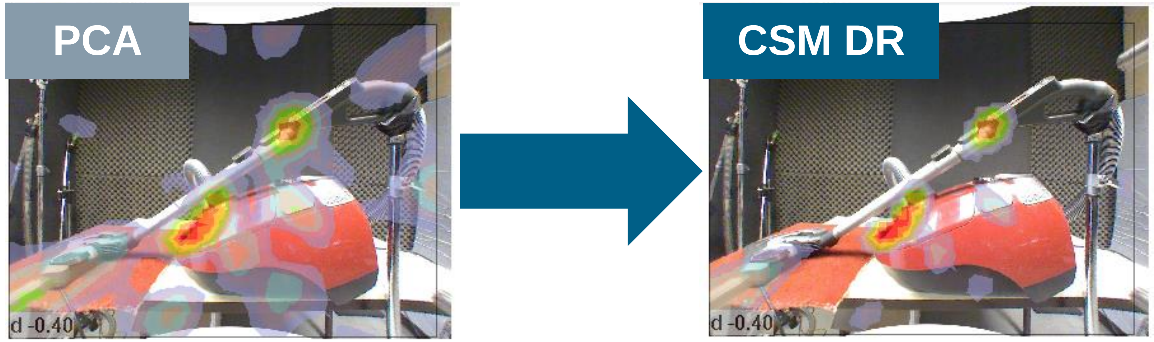
- Removal of microphone **self-noise** (e.g. wind noise, sensor noise, etc.)
- Extends Hologram averaging, significantly improves dynamic range (especially **aero-acoustic** sources)

CSM DR averaging

- Cross-Spectrum Matrix formulation of classical Beamforming with N microphones at point k:

$$B_k(f) = w_k'(f) CSM(f) w_k(f), CSM(f) = \left[p_i(f)\, p_j^*(f)\right]_{NxN}$$

- Observation: microphone autopowers on diagonal are very sensitive to **self-noise** → dynamic range reduction
- Removal of diagonal components improves quality of Beamforming result by reducing overall noise:



Restricted © Siemens AG 2018

Map averaging methods
Summary

SIEMENS
Ingenuity for life

	Spectrum	PCA	Hologram	CSM
Correlated sources	☺	☺	😐	☺
Uncorrelated sources	☹	☺	😐	☺
Non-stationary sources	😐	😐	☺	😐
Calculation time	☺	😐	😐	☺

Restricted © Siemens AG 2018

www.siemens.plm/simcenter

Fast new processing kernel in HDCAM

SIEMENS
Ingenuity for life

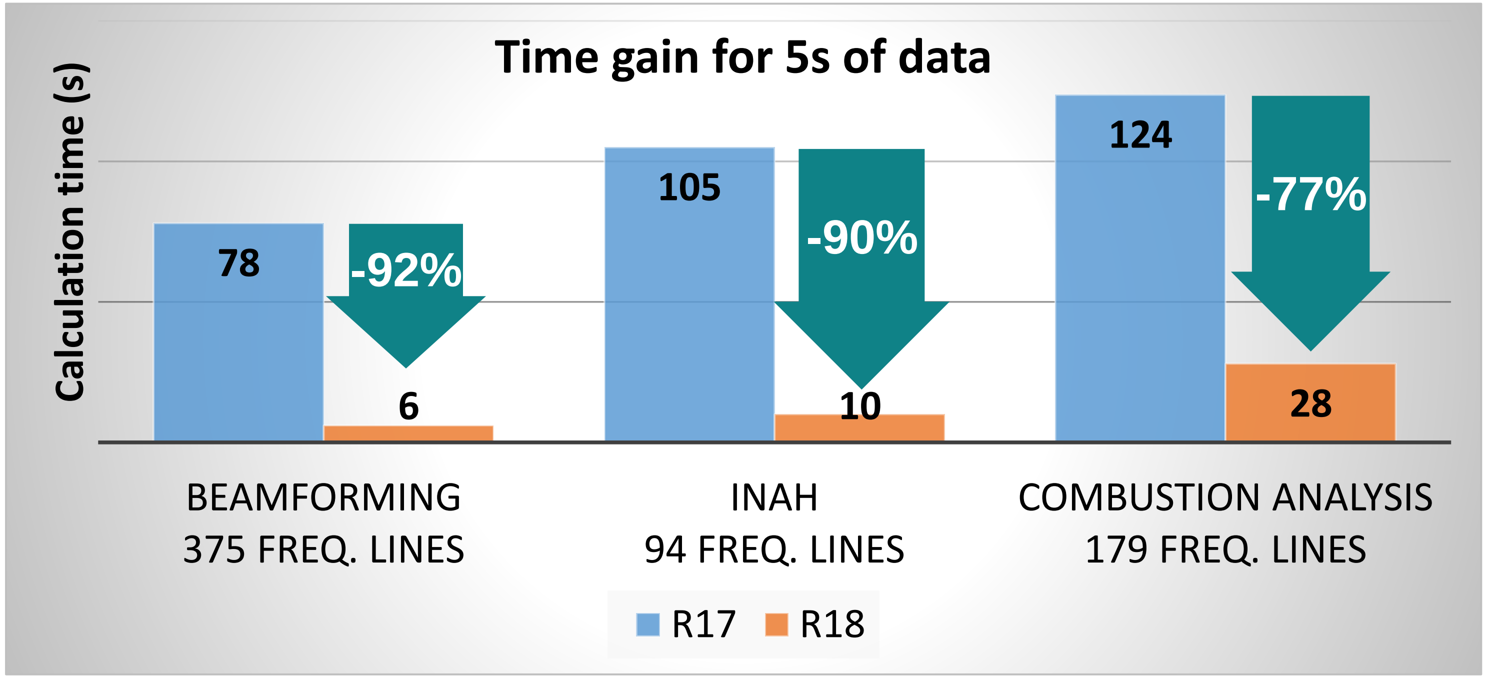
Up to 10x faster results for all methods

Supports all methods: beamforming, focalization, iNAH, deconvolution

Benchmark shows improvement by factor 7-10 for all methods

Fast processing using state of the art GPU and multi-core computing.

Supports NVidia cards with CUDA GPUs: (GeForce 9xx, Quadro Mxxx).



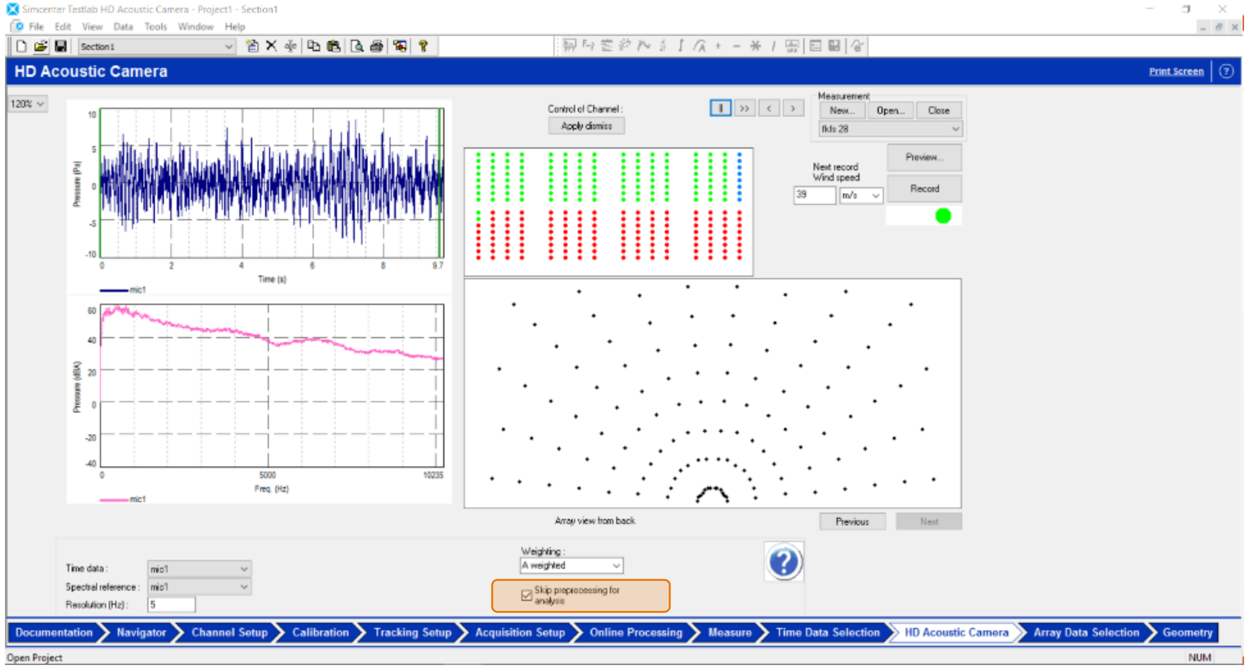
Restricted © Siemens AG 2018

Shortened time to next measurement in HDCAM

HDCAM enables faster repetitive measurements without having to wait for preprocessing for analysis

Improves time-to-next-measurement to 5-10s.

Repetitive sequential measurements do not require preprocessing for analysis after each recording



Restricted © Siemens AG 2018

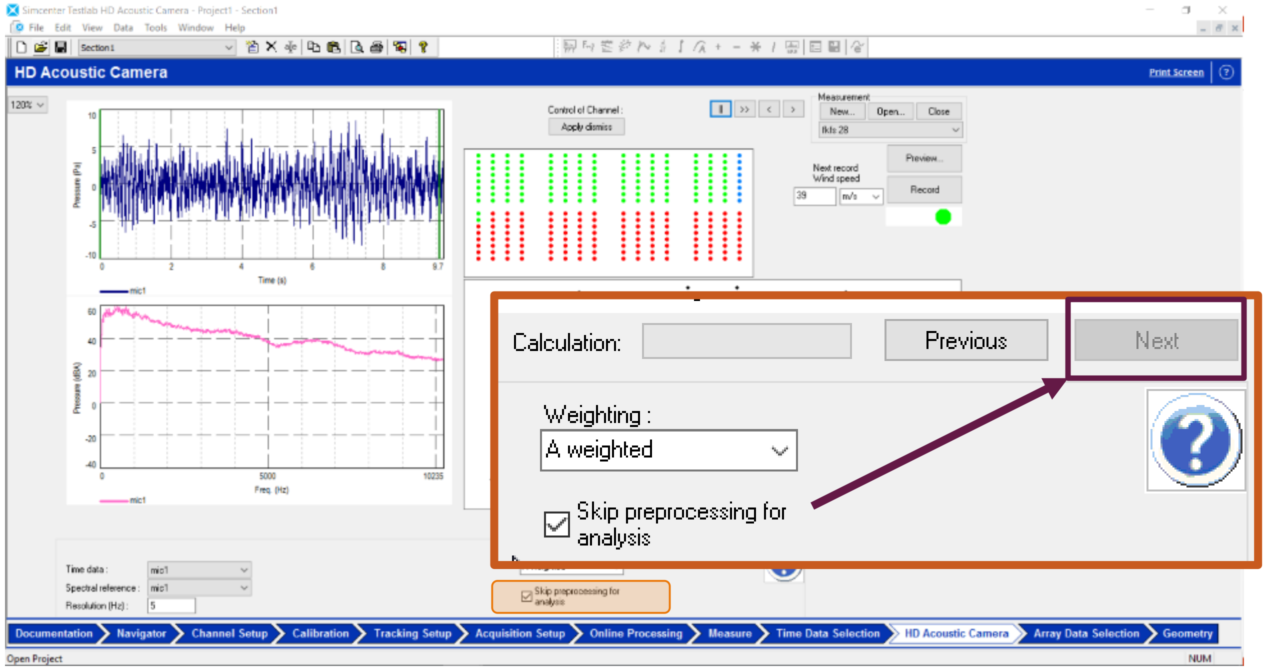
Shortened time to next measurement in HDCAM

SIEMENS
Ingenuity for life

HDCAM enables faster repetitive measurements without having to wait for preprocessing for analysis

Improves time-to-next-measurement to 5-10s.

Repetitive sequential measurements do not require preprocessing for analysis after each recording



Restricted © Siemens AG 2018