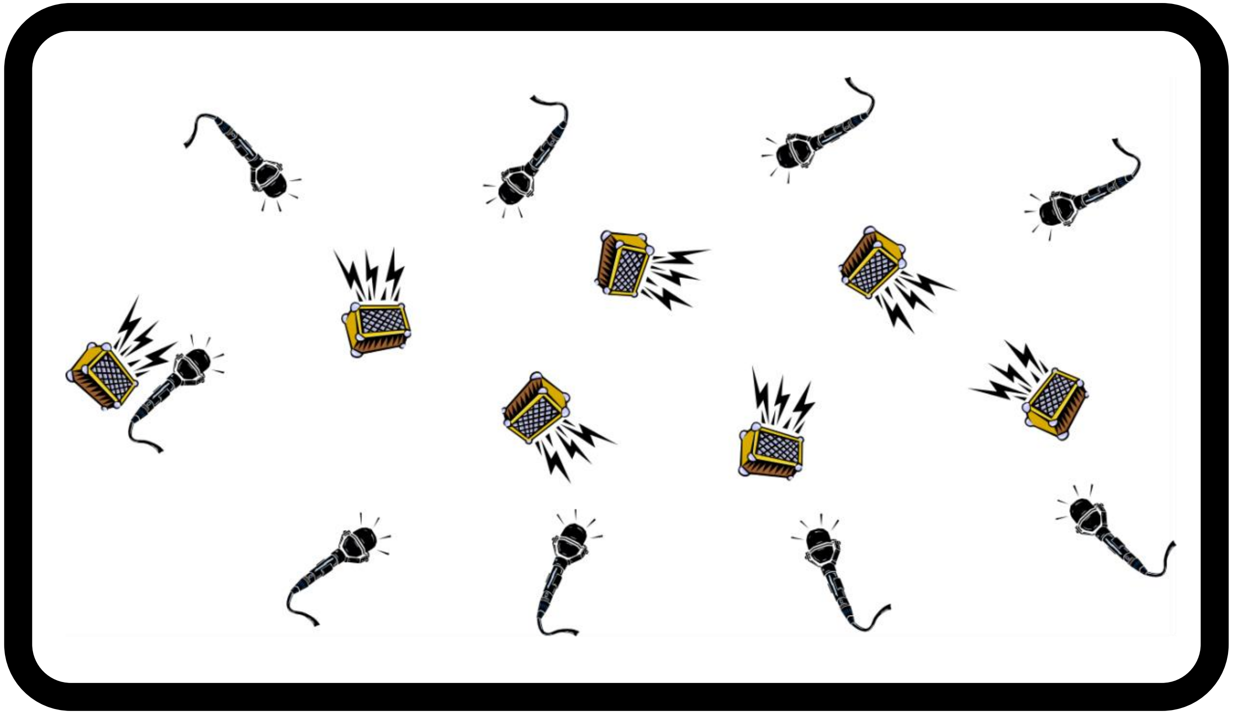
Microphone Calibration

Sensor positions from sound

General Definition

Given a set of M microphones (sensors) and N sources, determine the 3D position of both microphones and sources.



Problem assumptions

Free space assumption

Requires: no walls or reflection from walls is negligible

Time of emission (TOE) is known

Requires: timestamp from the emitted source and common clock

Sources position is known

Requires: tracking system of the sources (not practical indoor, not accurate outdoor)

Microphones position is known

Requires: in general a known microphones setup (rigid array)

Room Geometry is known

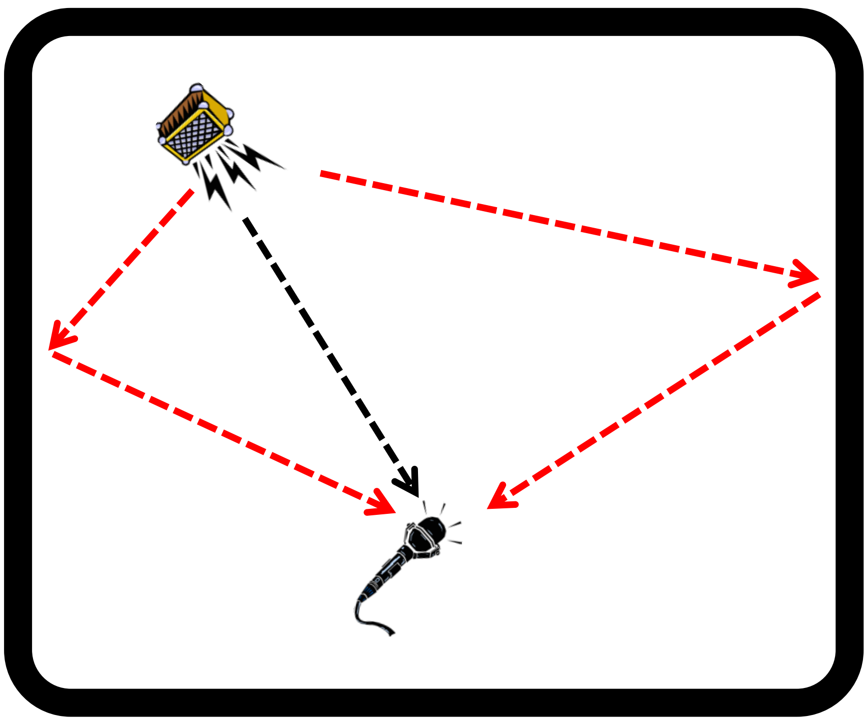
Requires: manual room measurements or available CAD model

Free space assumption in a room

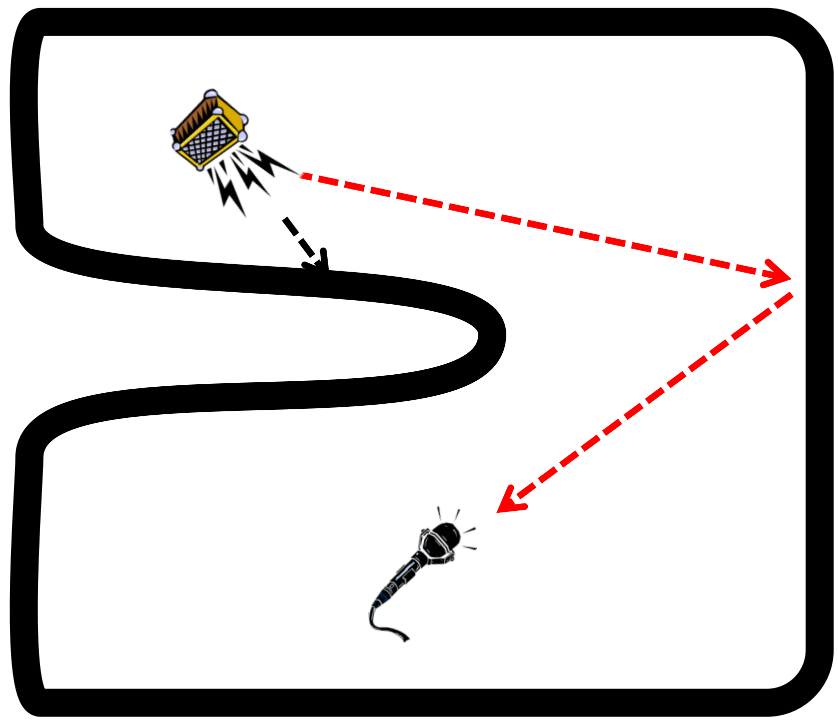
Free space assumption

<u>Requires</u>: no walls or reflection from walls is negligible

If the room is convex, by considering the direct path delay only we can approximate the problem as being in free space

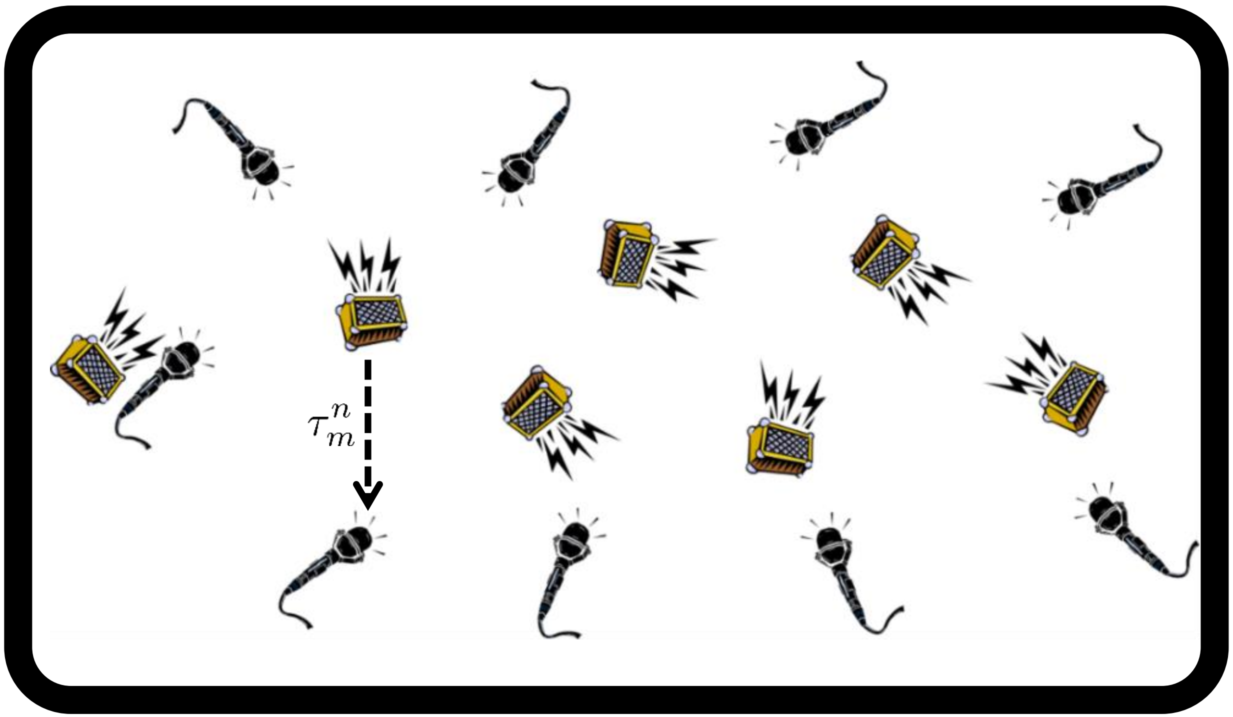


This is no longer valid for non-convex rooms since direct path delay might not exist



The general equation in free space

We assume here the delays being already extracted from the previous TOF estimation stage.



$\tau_{m0}^n \dashrightarrow \tau_m^n$ To simplify notation we remove 0, all the delays are direct path

The general equation in free space

Given the delays for the direct paths we can define the distance by knowing:

τ_e^n Time of emission for event *n*
τ_m^o Sampling offset time at microphone *m* (= 0 if mics share same clock)
τ_m^n Time of flight from event *n* to microphone *m*

$$\tau_m^{an} = \tau_m^n + \tau_m^o + \tau_e^n$$

where τ_m^{an} is the time of arrival of the signal at the microphone.

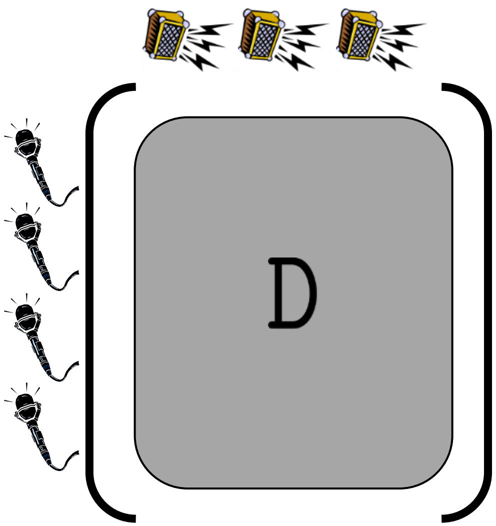
If these delays are known, we can compute the distance d_m^n using the constant *c* representing sound velocity as:

$$\tau_m^n = d_m^n / c$$

Distance is then defined by: $d_m^n = |\mathbf{b}^n - \mathbf{s}_m|$ where:

 $\mathbf{b}^n$ is a vector with the x, y, z coordinates of the microphones.
 $\mathbf{s}_m$ is a vector with the x, y, z coordinates of the sources.

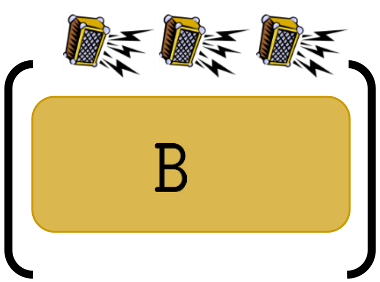
In matrix form...



$$= \begin{bmatrix} d_1^1 & \cdots & d_1^N \\ d_2^1 & \cdots & d_2^N \\ \vdots & \ddots & \vdots \\ d_M^1 & \cdots & d_M^N \end{bmatrix}$$

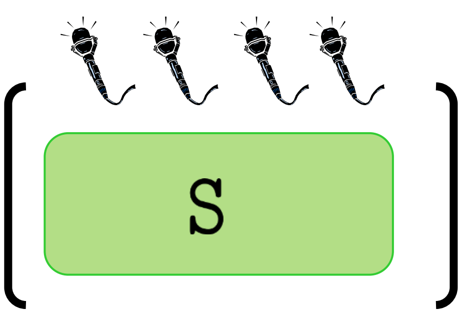
M X N matrix

From now on to ease notation, the matrix $\mathtt{D}$ contains the squared distances instead of the standard distances.



$$= \begin{bmatrix} \mathbf{b}^1 & \mathbf{b}^2 & \cdots & \mathbf{b}^N \end{bmatrix}$$

3 x N matrix



$$= \begin{bmatrix} \mathbf{s}_1 & \mathbf{s}_2 & \cdots & \mathbf{s}_M \end{bmatrix}$$

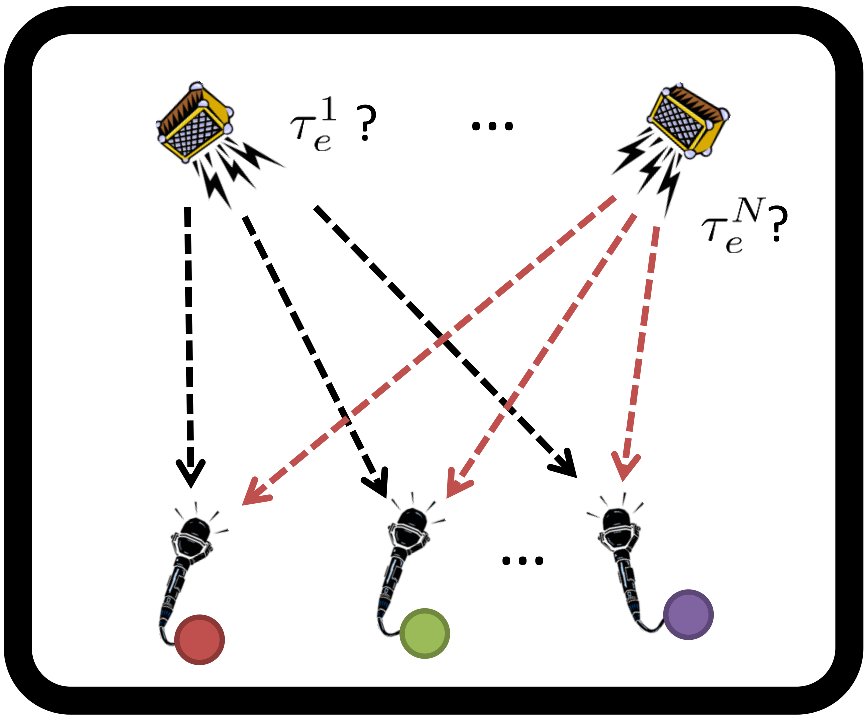
3 x M matrix

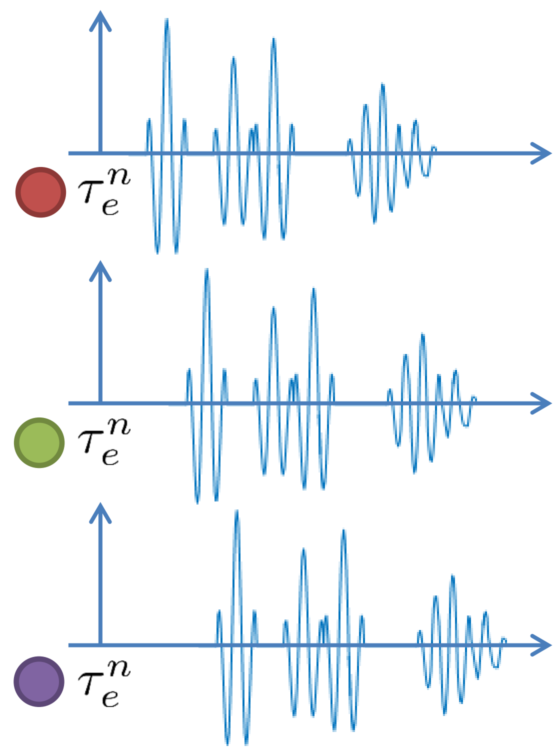
Signal acquisition synchronization

Time of emission (TOE) is known

<u>Requires</u>: timestamp from the emitted source and common clock

In general, the emission time of the source is unknown, moreover, the sensors might not share a common clock when sampling, i.e., there is a new unknown variable τ_e^n



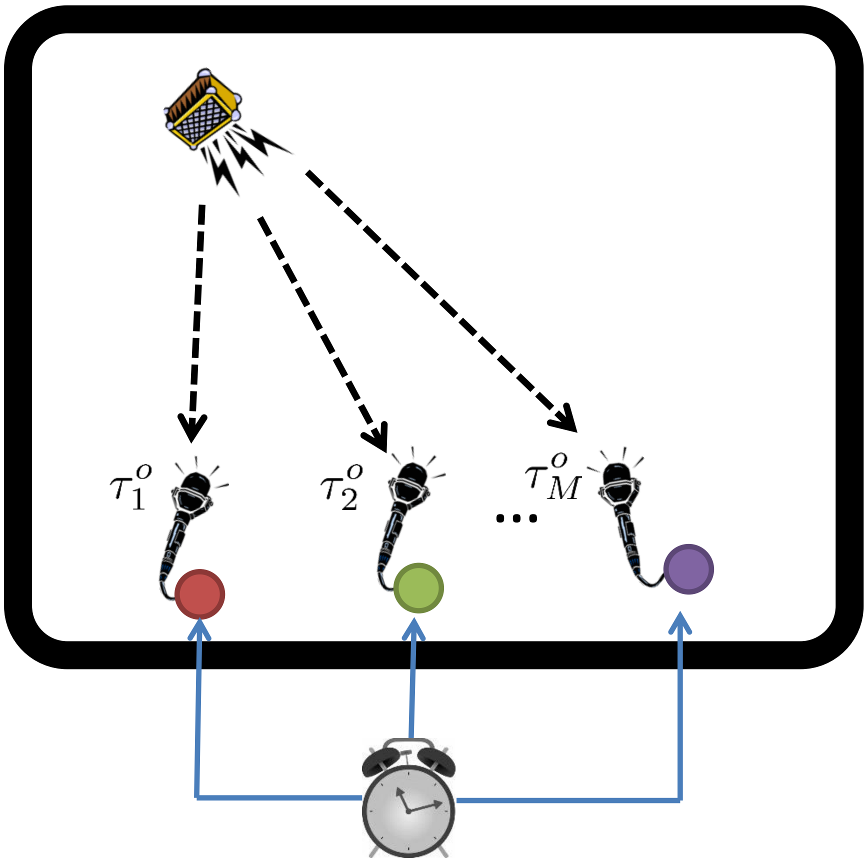


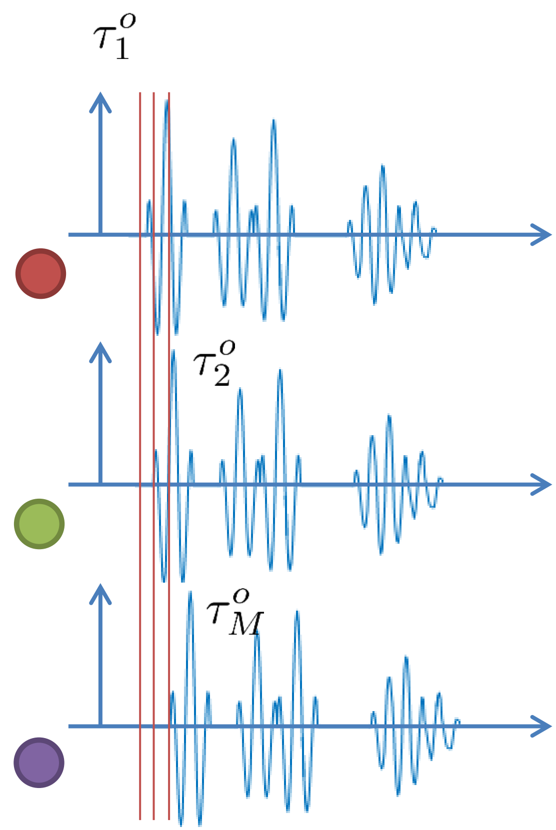
Signal acquisition synchronization

All the mics shares the same clock

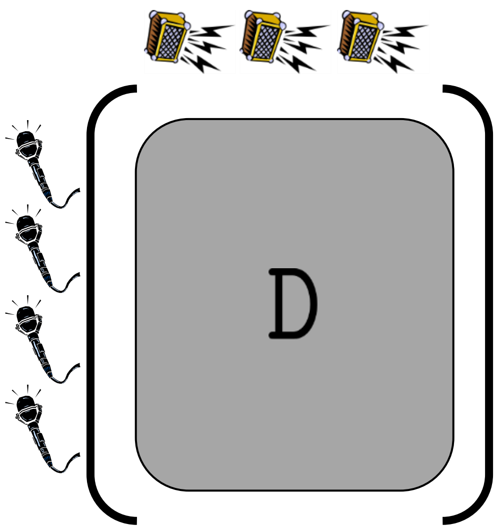
Requires: acquisition time equal or offset known (hardware based)

Same acquisition platform with a common clock for sampling or clock from each sensor given with high accuracy (e.g. using gps if outdoor)





Different unknowns but similar problems



M X N matrix

$$
= \begin{array}{cc}
\boxed{\tau_e^1 ? \quad \cdots \quad \tau_e^N ?} & \\
\begin{bmatrix} d_1^1 & \cdots & d_1^N \\ d_2^1 & \cdots & d_2^N \\ \vdots & \ddots & \vdots \\ d_M^1 & \cdots & d_M^N \end{bmatrix} & \boxed{\begin{array}{c} \tau_1^o \; ? \\ \tau_2^o \; ? \\ \vdots \\ \tau_M^o \; ? \end{array}}
\end{array}
$$

Finding the time of emission (TOE)

The problem is to estimate the time of emission (or the off-set in sampling) only from the measured direct path delays.

This is a hard problem **if we do not know signal attenuation** but there exists solutions only using the matrix $\mathbb{D}$ of distances. The principle used is that if the distances are computed from the exact time of flights τ_m^n, **the distance matrix has rank 5.**

We make this explicit by expressing the distance between mic-source explicitly:

$$\mathbf{b}^n = (b_1^n, b_2^n, b_3^n)^\top \qquad v = 1/c$$

$$\mathbf{s}_m = (s_{1m}, s_{2m}, s_{3m})^\top \qquad \tau_m^o = 0$$

Assuming clock being synchronized

$$(d_m^n)^2 = (s_{1m} - b_1^n)^2 + (s_{2m} - b_2^n)^2 + (s_{3m} - b_3^n)^2 = v\,(\tau_m^{an} - t_e^n)^2$$

Measured time of arrival

Unknown time of emission

Finding the time of emission (TOE)

$$\left(d_m^n\right)^2 = \left(s_{1m} - b_1^n\right)^2 + \left(s_{2m} - b_2^n\right)^2 + \left(s_{3m} - b_3^n\right)^2 = v\left(\tau_m^{an} - t_e^n\right)^2$$

$$\left(d_m^n\right)^2 = \tilde{\mathbf{s}}_m^\top \tilde{\mathbf{b}}^n = v^2\left(\left(\tau_m^{an}\right)^2 - 2\tau_m^{an} t_e^n + \left(t_e^n\right)^2\right)$$

Where:

$$\tilde{\mathbf{s}}_m^\top = \begin{pmatrix} \mathbf{s}_m^\top \mathbf{s}_m & -2s_{1m} & -2s_{2m} & -2s_{3m} & 1 \end{pmatrix}^\top$$ 5 x 1 vector

$$\left(\tilde{\mathbf{b}}^n\right)^\top = \begin{pmatrix} 1 & -2b_1^n & -2b_2^n & -2b_3^n & \left(\mathbf{b}^n\right)^\top \mathbf{b}^n \end{pmatrix}^\top$$ 5 x 1 vector

Finding the time of emission - Back to matrices

$$(d_m^n)^2 = \tilde{\mathbf{s}}_m^\top \tilde{\mathbf{b}}^n = v^2 \left((\tau_m^{an})^2 - 2\tau_m^{an} t_e^n + (t_e^n)^2 \right)$$

$$\tilde{\mathbf{s}}_m^\top = \begin{pmatrix} \mathbf{s}_m^\top \mathbf{s}_m & -2s_{1m} & -2s_{2m} & -2s_{3m} & 1 \end{pmatrix}^\top$$ 5 x 1 vector

$$\left(\tilde{\mathbf{b}}^n\right)^\top = \begin{pmatrix} 1 & -2b_1^n & -2b_2^n & -2b_3^n & (\mathbf{b}^n)^\top \mathbf{b}^n \end{pmatrix}^\top$$ 5 x 1 vector

$$\mathrm{D} = \underbrace{\begin{bmatrix} \{(d_m^n)^2\} \end{bmatrix}}_{\text{M x N}} = \underbrace{\begin{bmatrix} \tilde{\mathbf{s}}_1^\top \\ \tilde{\mathbf{s}}_2^\top \\ \vdots \\ \tilde{\mathbf{s}}_M^\top \end{bmatrix}}_{\substack{\tilde{\mathrm{S}} \\ \text{M x 5}}} \underbrace{\begin{bmatrix} \tilde{\mathbf{b}}^1 \, \tilde{\mathbf{b}}^2 \cdots \, \tilde{\mathbf{b}}^N \end{bmatrix}}_{\substack{\tilde{\mathrm{B}} \\ \text{5 x N}}}$$

Finding the time of emission - Back to matrices

$$\mathtt{D} = \begin{pmatrix} \{(d_m^n)^2\} \end{pmatrix}$$

M x N

The matrix $\mathtt{D}$ has a rank $\leq$ 5 since being the product of two rank constrained matrices $\tilde{\mathtt{S}}$ and $\tilde{\mathtt{B}}$

However the rank 5 constraint is valid only if the time of emission is correctly estimated, i.e. **the matrix** $\mathtt{D}$ **contains the real time of flights** τ_m^n .

Still the rank constraint can be used as a principle to estimate the correct time of emission, i.e. we find the τ_e^n that makes the $\mathtt{D}$ matrix rank 5.

Pollefeys, Marc, and David Nister. "Direct computation of sound and microphone locations from time-difference-of-arrival data." ICASSP 2008.

Jiang, F., Kuang, Y., Astrom, K. “Time delay estimation for TDOA self-calibration using truncated nuclear norm regularization.” ICASSP 2013.

Gaubitch, N. D., Kleijn, W. B., Heusdens, R. “Auto-localization in ad-hoc microphone arrays.” ICASSP 2013

Solution with 10 mics and 5 sources

Pollefeys and Nister solution is based on a rearrangement of the matrix D that separates unknown from known data. This creates a new matrix with rank 10.

$$\mathtt{A} = \{(\tau_m^{an})^2\} \quad \mathtt{C} = \{\tau_m^{an}\} \quad \mathtt{Z} = diag(\tau_1^e, \ldots, \tau_N^e)$$

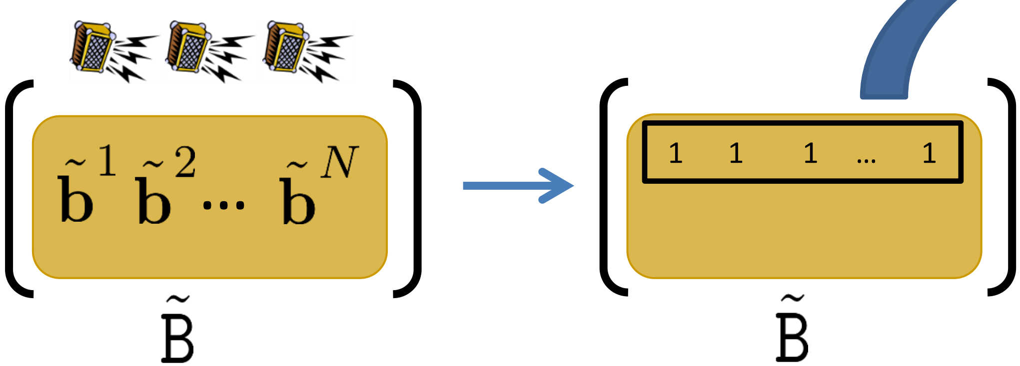
$$\mathtt{D} = \left\{ v^2 \left((\tau_m^{an})^2 - 2\tau_m^{an} t_e^n + (t_e^n)^2 \right) \right\} \longrightarrow \mathtt{T} = \{2\tau_m^{an} t_e^n + (\tau_m^{an})^2\}$$

$$\tilde{\mathbf{s}}_m^\top = \tilde{\mathbf{s}}_m^\top - \left[\ (\boldsymbol{\tau}_e^n)^2 \quad 0 \quad 0 \quad 0 \quad 0\ \right]^\top$$

Then by matrix product:

$$\mathtt{T} = \mathtt{A} + \mathtt{ZC} = \left[\ \mathtt{I} \mid \mathtt{Z}\ \right] \begin{bmatrix} \mathtt{A} \\ \mathtt{C} \end{bmatrix}$$

Notice that:



Enforce the presence of a row of 1s in the 5 x N matrix

Pollefeys, Marc, and David Nister. "Direct computation of sound and microphone locations from time-difference-of-arrival data." ICASSP 2008.

Solution with 10 mics and 5 sources

IDEA: find a linear system that enforces the row of ones, i.e. there is a linear combination of five **independent rows** of $\mathtt{T}$ that gives the row of ones.

$$\bar{\mathtt{T}} = \bar{\mathtt{A}} + \bar{\mathtt{Z}}\bar{\mathtt{C}} = \left[\begin{array}{c|c} \mathtt{I} & \bar{\mathtt{Z}} \end{array}\right] \left[\begin{array}{c} \bar{\mathtt{A}} \\ \bar{\mathtt{C}} \end{array}\right]$$

Reduced system with 5 selected rows

Problem: Find the vector **v** which linear combination is a row of ones

$$\mathbf{v}^\top \bar{\mathtt{T}} = \mathbf{v}^\top \bar{\mathtt{A}} + \mathbf{v}^\top \bar{\mathtt{Z}}\bar{\mathtt{C}} = \mathbf{v}^\top \left[\begin{array}{c|c} \mathtt{I} & \bar{\mathtt{Z}} \end{array}\right] \left[\begin{array}{c} \bar{\mathtt{A}} \\ \bar{\mathtt{C}} \end{array}\right] = \left[\begin{array}{cccc} 1 & 1 & \ldots & 1 \end{array}\right]$$

Since $\bar{\mathtt{Z}}$ is an unknown we instead estimate the 10 vector **x** given by:

$$\mathbf{x}^\top = \mathbf{v}^\top \left[\begin{array}{c|c} \mathtt{I} & \bar{\mathtt{Z}} \end{array}\right]$$

Thus giving:

$$\mathbf{x}^\top \bar{\mathtt{T}} = \mathbf{x}^\top \left[\begin{array}{c} \bar{\mathtt{A}} \\ \bar{\mathtt{C}} \end{array}\right] = \left[\begin{array}{cccc} 1 & 1 & \ldots & 1 \end{array}\right]$$

Pollefeys, Marc, and David Nister. "Direct computation of sound and microphone locations from time-difference-of-arrival data." ICASSP 2008.

Recovering time of emission (TOE)

$$\mathbf{x}^\top \bar{\mathtt{T}} = \mathbf{x}^\top \begin{bmatrix} \bar{\mathtt{A}} \\ \bar{\mathtt{C}} \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}$$

The equation can be solved with Least Squares, after obtaining the solution we recover back τ_e^n but this implies **having 10 microphones**.

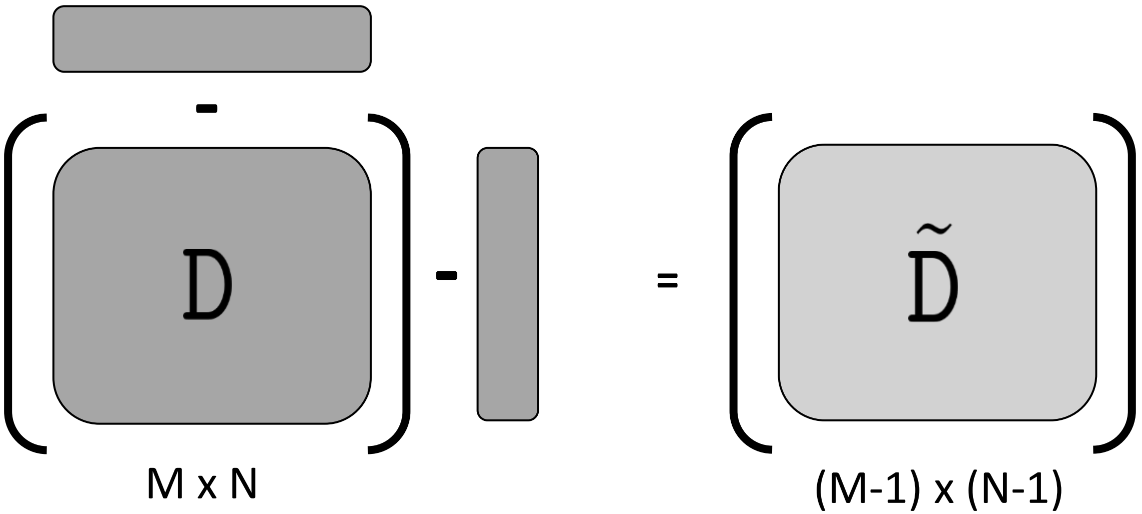
Moreover **the solution is sub-optimal** because it needs a sub-grouping of the sources in groups of 5. Each group has to have independent rows, something that might be difficult to assess in the presence of noise and degenerate mic/sources configurations.

Non-linear optimization in the form of alternating Least Squares is then used to refine the initial solution

$$\left\| \left(\{(d_m^n)^2\} \right) - \left(\tilde{\mathtt{S}} \right) \left(\tilde{\mathtt{B}} \right) \right\|_F^2$$

A solution with nuclear norm optimization

From rank 5 to rank 3



$$\underset{x,\tau_e^n}{\text{minimize}} \quad \|\mathtt{X}\|_*$$

$$\text{subject to} \quad \tilde{\mathtt{D}} + \sum_{n=1}^{N-1} \tau_e^n \tilde{\mathtt{D}}_n = \mathtt{X}$$

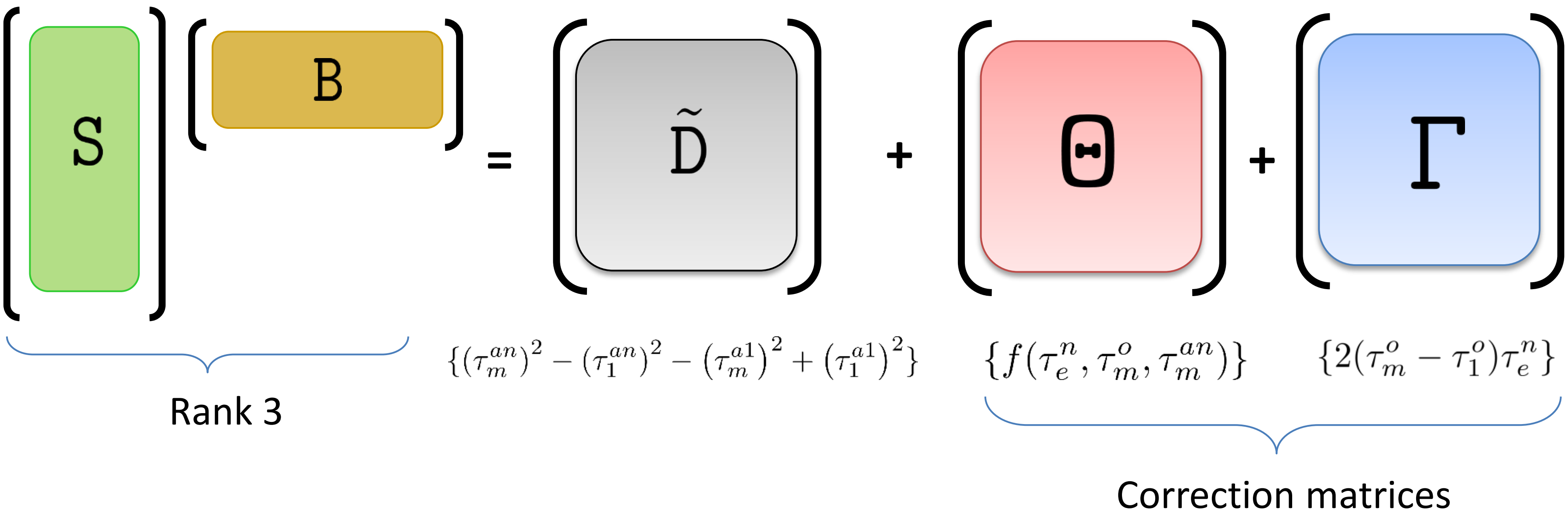
where $\tilde{\mathtt{D}}_1, \tilde{\mathtt{D}}_2, \ldots, \tilde{\mathtt{D}}_N$ are constant matrices that contain delays at each sound event. When τ_e^n is correctly estimated, $\mathtt{X}$ contains the correct times of flight

A two-step iterative scheme solves for the nuclear norm optimization (derived from Zhang et al., CVPR 2012).

Jiang, F., Kuang, Y., Astrom, K. "Time delay estimation for TDOA self-calibration using truncated nuclear norm regularization." ICASSP 2013.

Solving for both offset and TOE

Gaubitch et al. instead solve for both the TOE and offset delays again using the rank-3 constraint of the reduced matrix .



The correction matrices make the matrix of time of arrivals being of rank 3 so it is possible to extract directly microphone and source positions (rank 3 instead of rank 5)

Gaubitch, N. D., Kleijn, W. B., Heusdens, R. “Auto-localization in ad-hoc microphone arrays.” ICASSP 2013

Solving for both offset and TOE

The principle is now to optimize again the rank of the time of arrivals using the two correction matrices. We have that:

$$\underset{\Gamma,\Theta}{\text{minimize}} \quad \hat{\mathtt{D}} = \tilde{\mathtt{D}} + \Gamma + \Theta$$

$$\text{subject to} \quad rank(\hat{\mathtt{D}}) = 3$$

Solution to this complex optimization is given by steps after initializing τ_e^n, τ_m^o to random values:

$\tilde{\mathtt{D}}_3 = rank_3(\tilde{\mathtt{D}} + \Theta + \Gamma)$	Do a rank-3 approximation (put to 0 singular values after the 3rd)
$\mathtt{E} = \tilde{\mathtt{D}} - \tilde{\mathtt{D}}_3$	Find the residual from the approximation
$min_{\tau_e^n, \tau_m^o} \left\Vert \mathtt{E} - \Theta - \Gamma \right\Vert_F^2$	Re-estimate the parameters from the residual
$\tau_e^n, \tau_m^o \rightarrow \Theta, \Gamma$	Update the correction matrices
$\text{iterate until convergence}$	

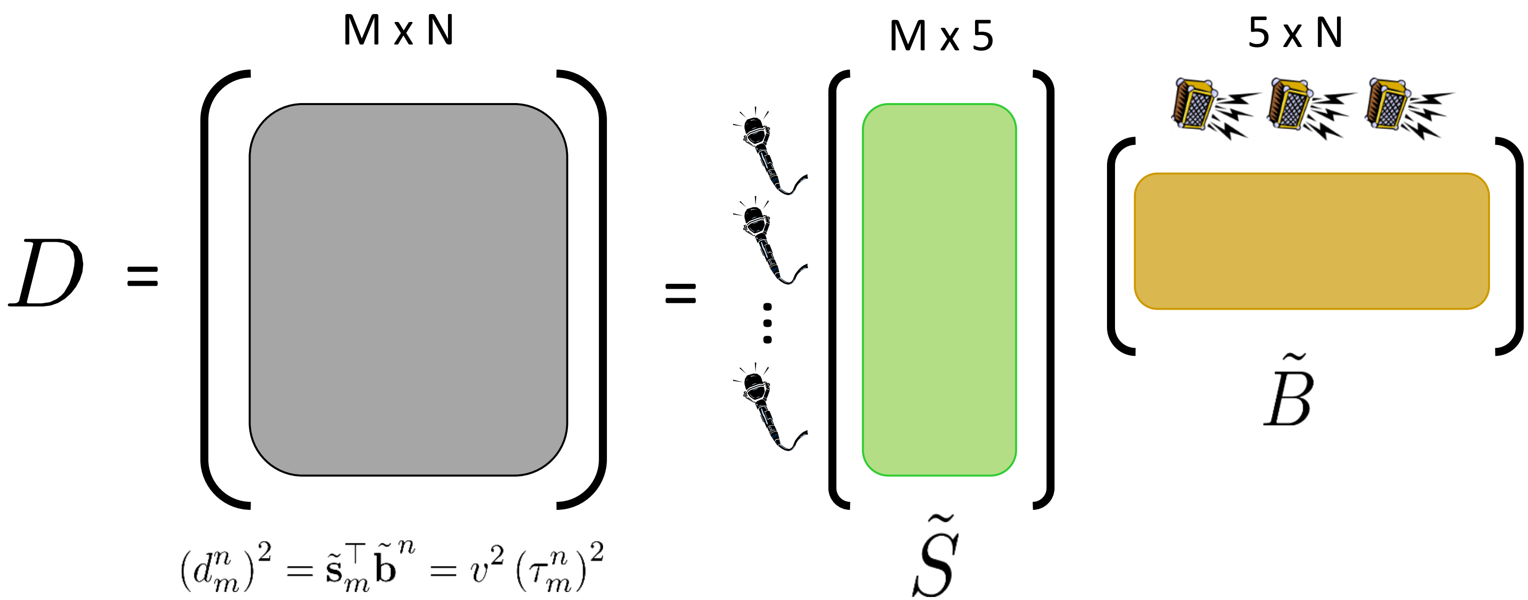
The algorithm requires a minimum of 5 microphones and 13 sound source events

Gaubitch, N. D., Kleijn, W. B., Heusdens, R. "Auto-localization in ad-hoc microphone arrays." ICASSP 2013

Uncalibrated case

Known the TOEs and offsets, is it possible to estimate both mics and sources positions?

By exploiting the rank constraint in the D matrix, we can factorise out mics and sound sources 3D positions. This is both valid for far-field and near-field cases.



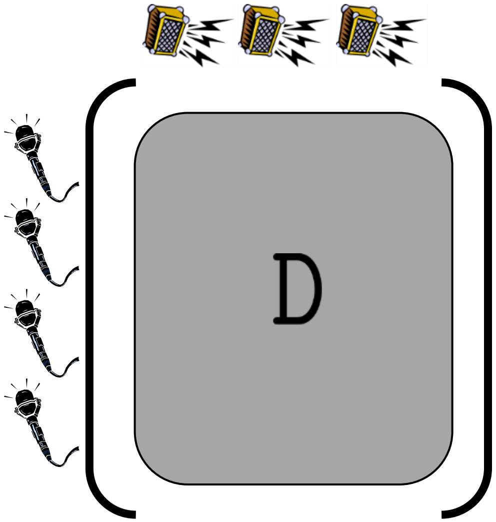
$$(d_m^n)^2 = \tilde{\mathbf{s}}_m^\top \tilde{\mathbf{b}}^n = v^2 \, (\tau_m^n)^2$$

From now on, the matrix D contains only the time of flights, i.e. $rank(D) = 5$

Thrun, “Affine structure from sound”, NIPS 2005

M. Crocco et al. “A Bilinear Approach to the Position Self-Calibration of Multiple Sensors,” IEEE TSP, 2012.

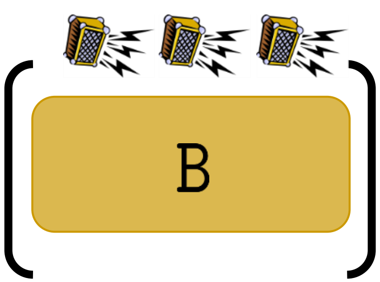
Bilinear Matrix form



$$= \begin{bmatrix} d_1^1 & \cdots & d_1^N \\ d_2^1 & \cdots & d_2^N \\ \vdots & \ddots & \vdots \\ d_M^1 & \cdots & d_M^N \end{bmatrix}$$

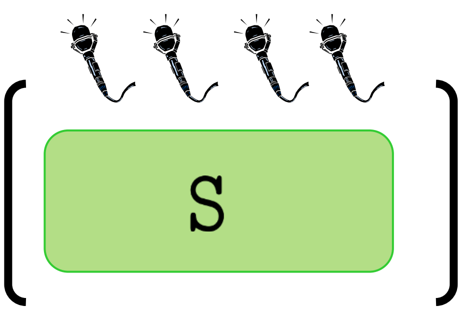
M X N matrix

From now on to ease notation, the matrix D contains the squared distances instead of the standard distances.



$$= \begin{bmatrix} \mathbf{b}^1 & \mathbf{b}^2 & \cdots & \mathbf{b}^N \end{bmatrix}$$

3 x N matrix

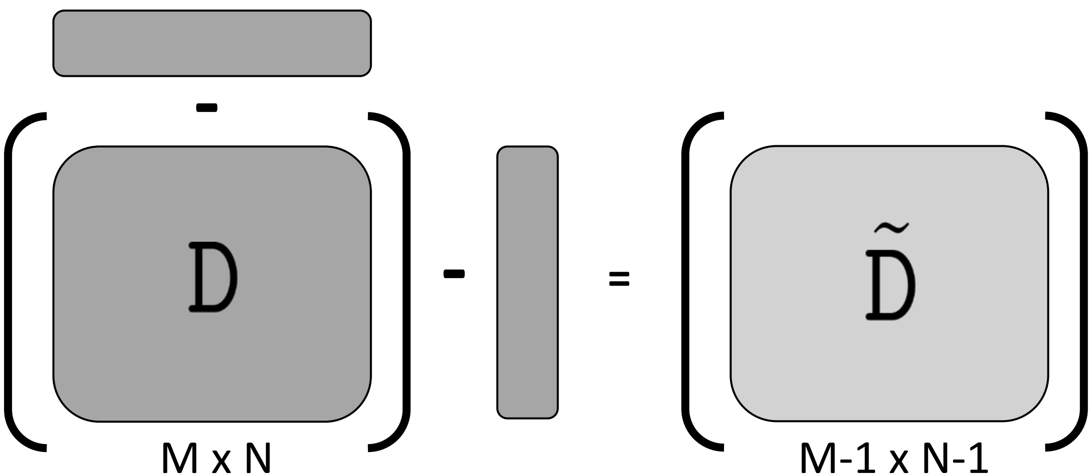


$$= \begin{bmatrix} \mathbf{s}_1 & \mathbf{s}_2 & \cdots & \mathbf{s}_M \end{bmatrix}$$

3 x M matrix

Distances and mic/sources positions

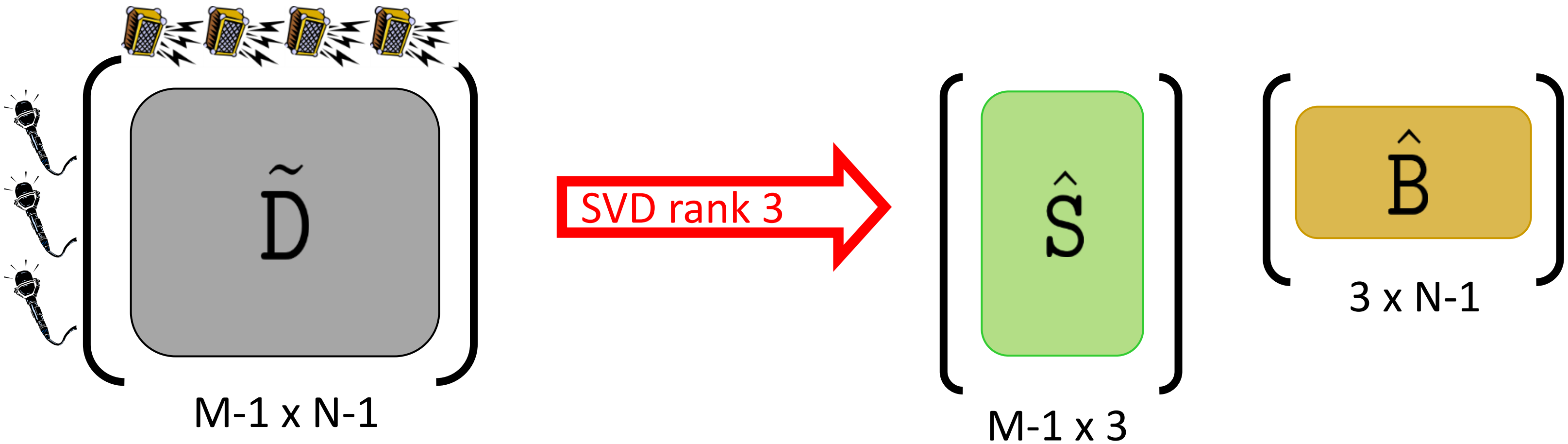
$$d_m^n = \left(\left\| \begin{pmatrix} s_{1m} \\ s_{2m} \\ s_{3m} \end{pmatrix} - \begin{pmatrix} b_1^n \\ b_2^n \\ b_3^n \end{pmatrix} \right\| \right)^2 \qquad d_m^n = \mathbf{s}_m^\top \mathbf{s}_m - 2\mathbf{s}_m^\top \mathbf{b}^n + (\mathbf{b}^n)^2$$



$$\tilde{d}_m^n = -2(s_{1m} - s_{11})(b_1^n - b_1^1) - 2(s_{2m} - s_{21})(b_2^n - b_2^1) - 2(s_{3m} - s_{31})(b_3^n - b_3^1)$$
$$\tilde{d}_m^n = d_m^n - d_1^n - d_m^1 + d_1^1$$

- This new matrix has an overall **rank 3!**
- The quadratic terms have disappeared, so we can express the product as a matrix multiplication.

Extracting a solution with SVD

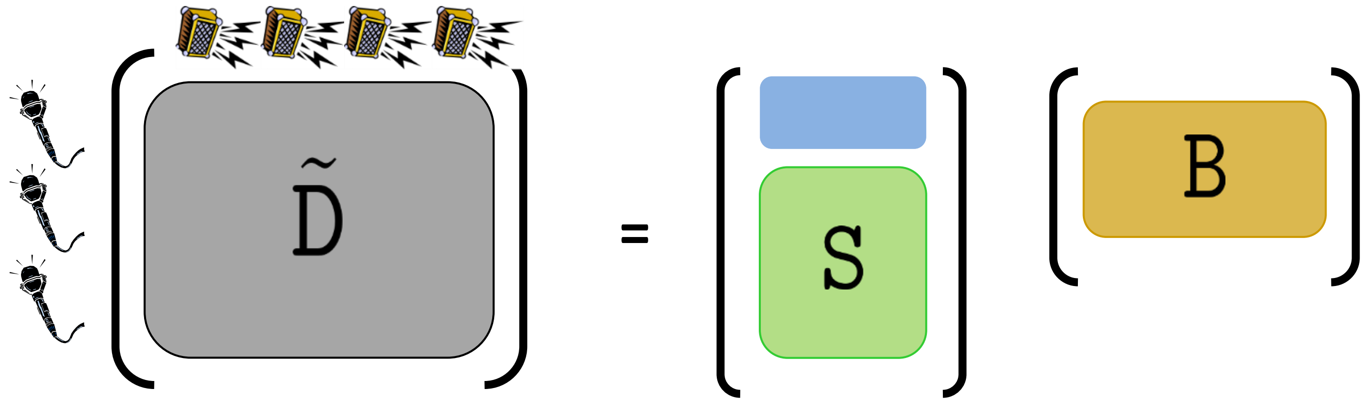


However, this factorisation is not unique. There is a general ambiguity in the form of a 3 x 3 full-rank matrix C.

$$\tilde{D} = \underbrace{\hat{S}\, C}_{S}\; \underbrace{C^{-1}\, \hat{B}}_{B}$$

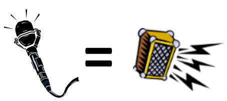
How to solve for the ambiguity?

1.



If a set of **anchors** exist (i.e. mics with a known a priori 3D position), it is possible to solve the problem in closed form.

2.



If a source and a mic have the same position, it is possible to solve the problem in closed form.

M. Crocco, A. Del Bue, V. Murino. “A Bilinear Approach to the Position Self-Calibration of Multiple Sensors," IEEE Transactions on Signal Processing, 2012.

3.

If the assumption of a microphone and source with same position is too restrictive, there exists a closed-form solution using 9 sources and 4 microphones.

Trung-Kien LE and Nobutaka ONO, “Closed-form and Near closed-form Solutions for TOA-based Joint Source and Sensor Localization” To appear.

Same location for 1 Mic, 1 Source

2.  =  If a source and a mic have the same position, it is possible to solve the problem in closed form.

M. Crocco, A. Del Bue, V. Murino. "A Bilinear Approach to the Position Self-Calibration of Multiple Sensors," IEEE Transactions on Signal Processing, 2012.

If $\mathbf{s}_1 = \mathbf{b}^1$ the factorization directly relates to the 3D coordinates of mic/sources by imposing the origin of coordinate system at $\mathbf{s}_1 = \mathbf{b}^1$

$$\mathtt{B} = \begin{bmatrix} \mathbf{b}^2 - \mathbf{b}^1 & \cdots & \mathbf{b}^N - \mathbf{b}^1 \end{bmatrix} \qquad \mathtt{S} = \begin{bmatrix} \mathbf{s}_2 - \mathbf{s}_1 & \cdots & \mathbf{s}_M - \mathbf{s}_1 \end{bmatrix}$$

$$\mathtt{B} = \begin{bmatrix} \mathbf{b}^2 & \cdots & \mathbf{b}^N \end{bmatrix} \qquad \mathtt{S} = \begin{bmatrix} \mathbf{s}_2 & \cdots & \mathbf{s}_M \end{bmatrix}$$

$$\tilde{\mathtt{D}} \longrightarrow \mathtt{U}_3\ \mathtt{V}_3\ \mathtt{W}_3$$

SVD truncation to rank 3

$$\mathtt{S} = \mathtt{UC}$$ 3 x N-1 matrix

$$\mathtt{B} = \mathtt{C}^{-1}\mathtt{VW}$$ 3 x M-1 matrix

Factorization ambiguity

$$\begin{aligned} \mathtt{S} &= \mathtt{UC} \\ \mathtt{B} &= \mathtt{C}^{-1}\mathtt{VW} \end{aligned}$$

Every matrix can be decomposed with a QR factorisation where Q is an orthogonal matrix and R an upper triangular matrix.

$$\mathtt{C} = \mathtt{Q}\,\mathtt{R} \qquad\qquad \mathtt{R} = \begin{pmatrix} r_1 & r_2 & r_3 \\ 0 & r_4 & r_5 \\ 0 & 0 & r_6 \end{pmatrix},$$

Since the factorization is up to an arbitrary rotation we can impose without loss of generality that Q is equal to the identity:

$$\begin{aligned} \mathtt{S} &= \mathtt{UR} \\ \mathtt{B} &= \mathtt{R}^{-1}\mathtt{VW} \end{aligned}$$

This means solving for 6 parameters only instead of 3(M-1) + 3(N-1) parameters for the 3D positions of mic and sources.

From many variables to few variables

In this way we can substitute back R into the original cost function with quadratic terms:

$$\mathrm{R}^* = \arg\min_R \sum_{m=1}^{M-1}\sum_{n=1}^{N-1} [u_{m1}^2(r_1^2 + r_2^2 + r_3^2) + u_{m2}^2(r_4^2 + r_5^2) +$$

$$+ u_{m3}^2 r_6^2 + 2u_{m1}u_{m2}(r_2 r_4 + r_3 r_5) + 2u_{m1}u_{m3}(r_3 r_6) +$$

$$+ 2u_{m2}u_{m3}(r_5 r_6) - k_{mn}]^2,$$

with:

$$k_{mn} = -u_{m1}v_{11}w_{1n} - u_{m2}v_{22}w_{2n} - u_{m3}v_{33}w_{3n} + d_{m+1,n+1}^2 - d_{1,n+1}^2.$$

We can use non-linear optimisation over the 6 parameters of R. Alternatively there is a closed-form solution obtained by grouping all the unknowns as:

$$\mathbf{f}^\top = \begin{pmatrix} r_1^2 + r_2^2 + r_3^2 & r_4^2 + r_5^2 & r_6^2 & r_2 r_4 + r_3 r_5 & r_3 r_6 & r_5 r_6 \end{pmatrix}$$

With:

$$\mathbf{s}_i^\top = \begin{pmatrix} u_{i1}^2 & u_{i2}^2 & u_{i3}^2 & 2u_{i1}u_{i2} & 2u_{i1}u_{i3} & 2u_{i2}u_{i3} \end{pmatrix}$$

$$\mathbf{k}^\top = \begin{pmatrix} k_{1,1} & k_{2,1} & \dots & k_{N-1,1} & k_{1,2} & \dots & k_{N-1,M-1} \end{pmatrix}$$

$$\mathrm{S}^\top = \begin{pmatrix} \mathbf{s}_1 & \mathbf{s}_2 & \dots & \mathbf{s}_{N-1} \end{pmatrix} \qquad \mathrm{P} = \begin{pmatrix} \mathrm{S} & \mathrm{S} & \dots & \mathrm{S} \end{pmatrix}$$

$$\mathbf{f}^* = \arg\min_{\mathbf{f}} = (|\mathrm{P}\mathbf{f} - \mathbf{k}|)^2$$

Closed-form solution

$$\mathbf{f}^* = \arg\min_{\mathbf{f}} = (|\mathrm{P}\mathbf{f} - \mathbf{k}|)^2$$

Solution can be obtained with a pseudo-inverse over P

$$\mathbf{f}^* = (\mathrm{P}^\top \mathrm{P})^{-1} \mathrm{P}^\top \mathbf{k},$$

and the values of R can be recovered as:

$$r_6 = \pm\sqrt{f_3}; \qquad r_5 = f_6 / r_6;$$
$$r_4 = \pm\sqrt{f_2 - r_5^2}; \qquad r_3 = f_5 / r_6;$$
$$r_2 = (f_4 - r_3 r_5)/r_4; \quad r_1 = \pm\sqrt{f_1 - r_2^2 - r_3^2};$$

This solution requires a minimum of 4 microphones and 4 sources.

M. Crocco, A. Del Bue, V. Murino. “A Bilinear Approach to the Position Self-Calibration of Multiple Sensors," IEEE TSP, 2012.

Non-linear optimization over *C*

If $\mathbf{s}_1 \neq \mathbf{b}^1$ we can always optimise for the original 3 x 3 matrix $\mathtt{C}$ giving the following cost function:

$$\begin{aligned} S &= UC \\ B &= C^{-1}VW \end{aligned}$$

$$\begin{aligned} \mathtt{C}^* = \arg\min_{\mathtt{C}} \sum_{m=1}^{M-1} \sum_{n=1}^{N-1} [&(u_{m1}c_{11} + u_{m2}c_{21} + u_{m3}c_{31})^2 + \\ &+ (u_{m1}c_{12} + u_{m2}c_{22} + u_{m3}c_{32})^2 + \\ &+ (u_{m1}c_{13} + u_{m2}c_{23} + u_{m3}c_{33})^2 + \\ &+ (u_{m1}c_{11} + u_{m2}c_{21} + u_{m3}c_{31})a_1 + u_{m1}v_{11}w_{1n} + \\ &+ u_{m2}v_{22}w_{2n} + u_{m3}v_{33}w_{3n} - d^2_{m+1,n+1} + d^2_{1n+1}]^2, \end{aligned}$$

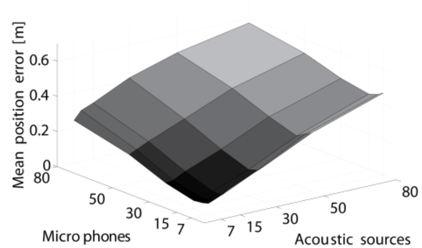
This optimization can be implemented with a gradient descent optimizing 9 parameters regardless the number of microphone and sources.

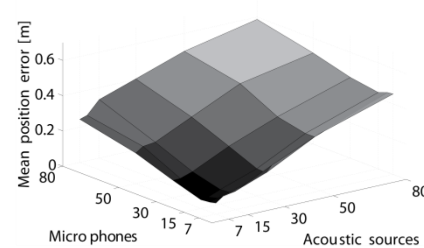
Synthetic experiments

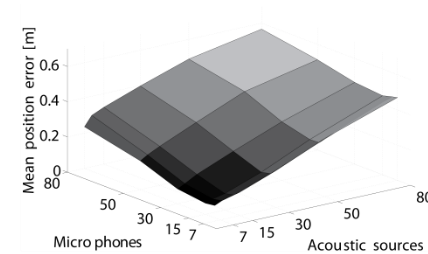
Variable number of sources and microphones displaced in a cubic region. Distance measurements corrupted with 0 mean Gaussian error and variable standard deviation σ.

Mean reconstruction error function of number of targets and sensors

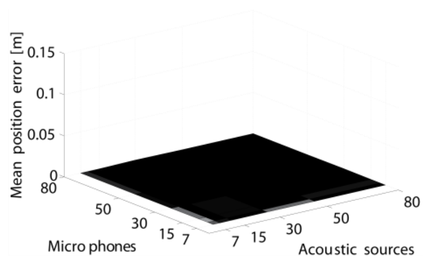
Gradient descent

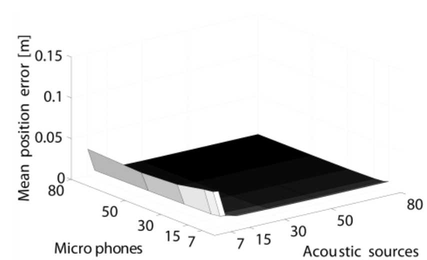


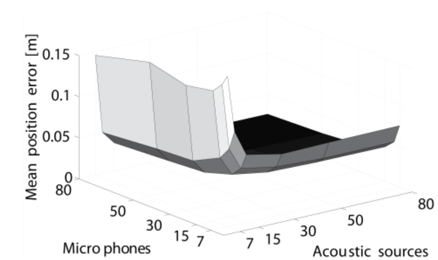




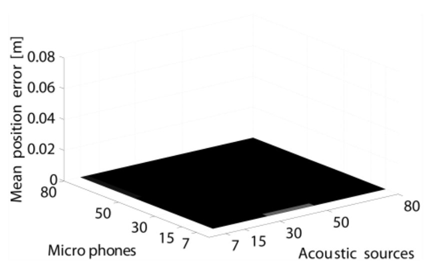
Proposed method

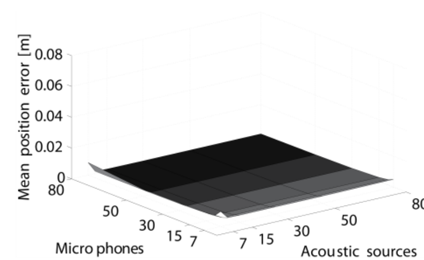


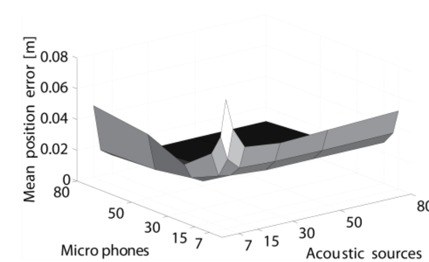




Proposed method refined by gradient descent







$\sigma = 0$ m $\qquad \sigma = 0.002$ m $\qquad \sigma = 0.02$ m

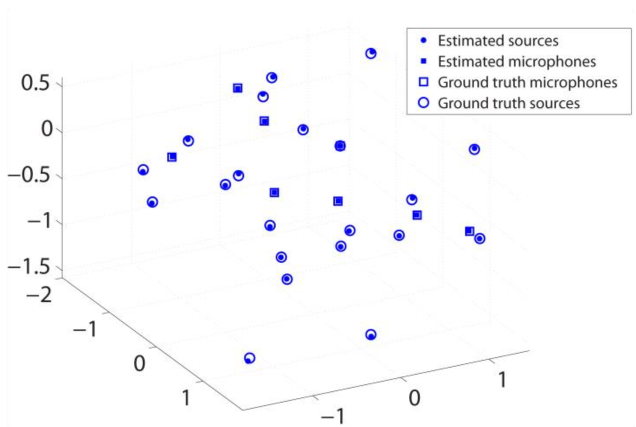
M. Crocco, A. Del Bue, V. Murino. "A Bilinear Approach to the Position Self-Calibration of Multiple Sensors," IEEE TSP 2012.

Real experiments

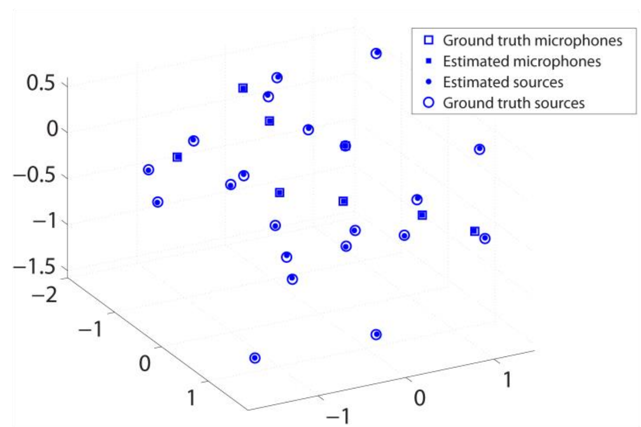
1) Room of 6 x 4 x 3 meters with 8 microphones and 21 sources obtained by moving a transducer in different positions. Distance is proportional to time of flight of the transmitted acoustic pulse.
2) Ground truth was measured by a VICON motion capture system.

Real experiments

1) Room of 6 x 4 x 3 meters with 8 microphones and 21 sources obtained by moving a transducer in different positions. Distance is proportional to time of flight of the transmitted acoustic pulse.
2) Ground truth was measured by a VICON motion capture system.



Closed form solution

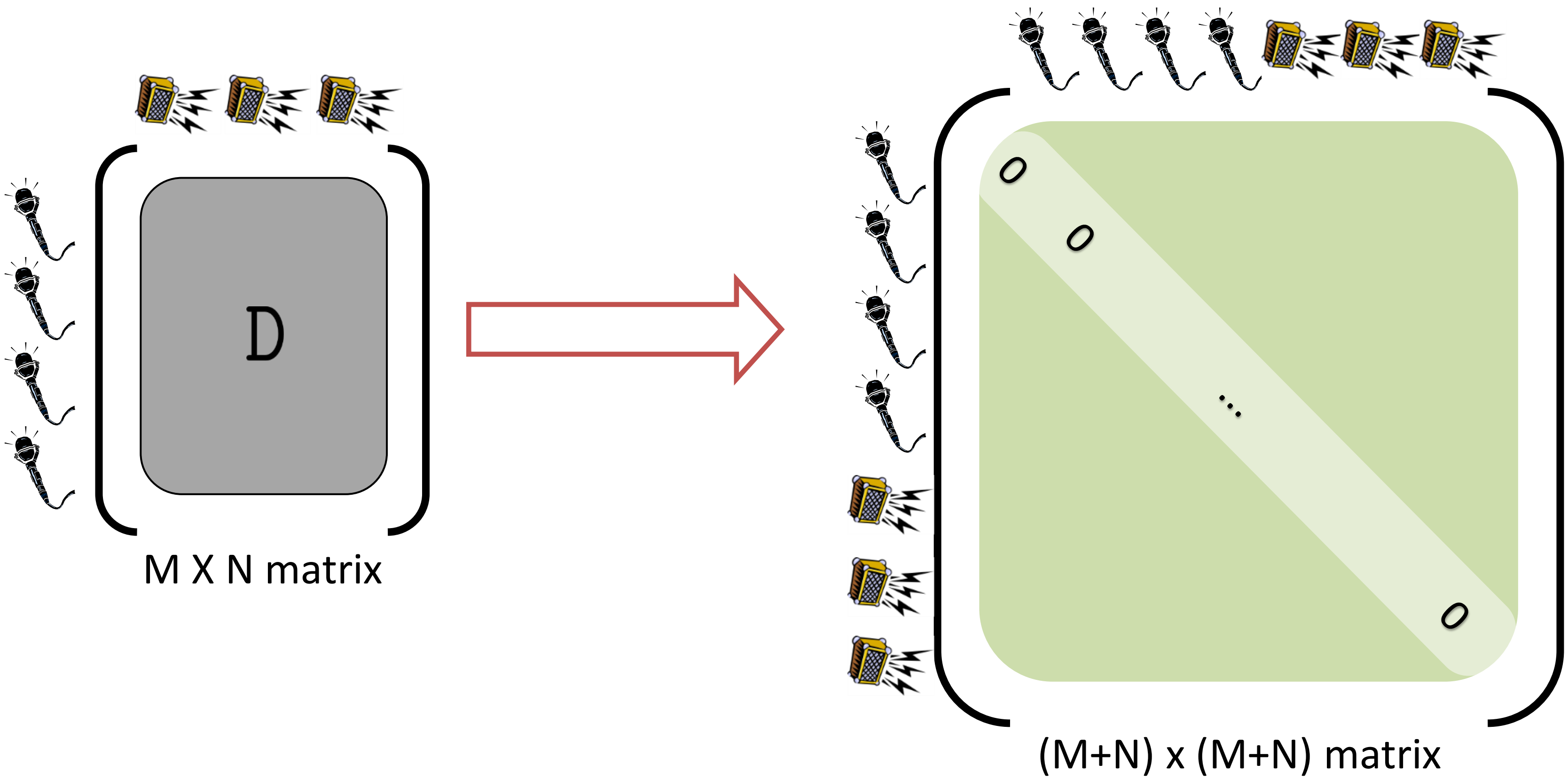


Solution refined with gradient descent

Mean reconstruction error of about 4 mm for microphones and about 1.2 cm for sources!

M. Crocco, A. Del Bue, V. Murino. "A Bilinear Approach to the Position Self-Calibration of Multiple Sensors," IEEE Transactions on Signal Processing, 2012.

From *D* matrices to $\mathbb{EDM}$



$\mathbb{EDM}$ matrices entail several theoretical properties (rank 5, symmetry, definiteness) and at the basis of the multidimensional scaling problem (MDS).

Thus, such theory and related computational tools can be linked to self-localisation problems even when missing entries are present in $\mathtt{D}$

Dokmanic et al., “Euclidean Distance Matrices: Essential theory, algorithms, and applications”, IEEE Signal Process. Mag. 2015.

Dokmanic et al., “Relax and unfold: Microphone localization with Euclidean distance matrices”, EUSIPCO 2015.

NEXT: Room geometry estimation