



Capabilities of inverse scheme for acoustic source localization at low frequencies

Stefan Gombots^{1,*}, Manfred Kaltenbacher², and Barbara Kaltenbacher³

¹Institute of Mechanics and Mechatronics, 1040 Vienna, Austria

²Institute of Fundamentals and Theory in Electrical Engineering, 8010 Graz, Austria

³Department of Mathematics, Alpen-Adria-Universität Klagenfurt, 9020 Klagenfurt, Austria

Received 22 January 2021, Accepted 1 September 2021

Abstract – We present the capabilities of a recently developed inverse scheme for source localization at low frequencies within an arbitrary acoustic environment. The inverse scheme is based on minimizing a Tikhonov functional matching measured microphone signals with simulated ones. We discuss the sensitivity of all involved parameters, the precision of geometry and physical boundary modeling for the numerical simulation using the finite element (FE) method, and the automatic determination of the positions of the recording microphones being distributed around the object of investigation. Finally, we apply the inverse scheme to a real-world scenario and compare the obtained results to state-of-the-art signal processing approaches, e.g. Clean-SC.

Keywords: Source Identification, Inverse Scheme, Beamforming

1 Introduction

The knowledge about position and distribution of sound sources is necessary for taking actions in noise reduction. To obtain this information, a microphone array method can be applied. Thereby, a common technique is acoustic beamforming which can be used to determine the source position and distribution as well as the source strength. This technique is based on evaluating simultaneously recorded sound pressure data from microphone array measurements. The sound pressure received at the different microphone positions is mapped onto an image of the acoustic source field. The results are given in so called source maps (beamform maps), which are used for visualization. Thereby, the localization of acoustic sources, especially in the low frequency range, is still a topic of ongoing research with strong practical applications, e.g. wind turbines, heat pumps, transformers. The fundamental processing method, frequency domain beamforming [1], is robust and fast. Herein, the source map is computed by

$$\sigma(\tilde{\mathbf{w}}) = \sqrt{\tilde{\mathbf{w}}^H \tilde{\mathbf{C}} \tilde{\mathbf{w}}},$$

with the (weighted) steering vector $\tilde{\mathbf{w}}$ and the cross spectral matrix $\tilde{\mathbf{C}}$ of the microphone signals.¹ Here, the steering vector accounts for the phase shift and the amplitude

correction as well as microphone weighting. In this simple method, the limitations regarding resolution and dynamic range are caused by the array response function (point spread function, PSF). A modification of this classic approach is delivered by functional beamforming (FuncBF) [2, 3], which leads to an improvement of resolution and dynamic range, while the computational cost remains almost the same as in the standard approach. Furthermore, deconvolution techniques (e.g. DAMAS [4], Clean-SC [5], etc.) can be used, which attempt to eliminate the influence of the PSF on the raw source map. In addition, there are beamforming independent inverse methods (like e.g. L1 generalized inverse beamforming [6], cross-spectral matrix fitting [7], etc.), which aim to solve an inverse problem considering the presence of all acoustic sources at once in the localization process. Resolution and dynamic range are greatly improved by these advanced methods at the costs of computational time. In literature, many different comparisons of beamforming methods can be found. Exemplarily, in [8, 9] and [10], simulated data was used as input, and in [11, 12], data coming from experiments. A comprehensive overview of different acoustic imaging methods can be found in [8, 13, 14].

Despite these advances in beamforming techniques, it has to be mentioned that major limitations are caused by the source model. Most beamforming algorithms model the acoustic sources as monopoles or/and dipoles. Moreover, the steering vector $\tilde{\mathbf{w}}$, describing the transfer function between source and microphone, is usually modeled by Green's function for free radiation. Different formulations

*Corresponding author: stefan.gombots@tuwien.ac.at

¹ with \square^H the Hermitian operation (complex conjugation and transposition) and \square denoting complex values.

of the steering vector exist [15], all of them exhibiting a trade-off between correct reconstruction of the source strength and the location. Further challenges are posed to these methods by source localization at low frequencies and in environments with partially or fully reflecting surfaces, for which beamforming techniques do not provide physically reasonable source maps. Furthermore, obstacles in the sound propagation path that distort the phase information can not be considered. To take a reverberant environment into account, the steering vectors have to be adapted. In [16] two approaches to consider reflective surfaces are envisaged, namely modeling the reflections by a set of monopoles located at the image source positions and an experimentally based identification of the transfer function.

Beamforming is mainly used for acoustic source localization rather than for obtaining quantitative source information. The qualitative statements given by the source map provide information about the distribution and position, and allow a relative comparison of the source strength at the considered map. In many cases, this information may already be sufficient to determine the origin of the sound emission. However, sometimes quantitative source information is also needed. The estimation of quantitative source spectra is not straightforward [17], but can be obtained through integration methods. To this end, the source map is integrated over a certain region to obtain a pressure spectrum for this specific area. Thus, the source map is required before integration. A distinction must be made between integrating the raw and deconvolved map. The deconvolved maps can be seen as ideal images of the source distribution and therefore, the integration may be done without further processing [18]. In [19], an overview of different integration methods is given.

To overcome the above mentioned limitations, an inverse scheme has been developed [20], which fully solves the Helmholtz equation using the finite element (FE) method with the correct boundary conditions. This approach is used in combination with microphone array measurements to localize and quantify low-frequency sound sources. The main advantage of this approach is given by the fact that a detailed source distribution, both in amplitude and phase is identified, and finally with this information a numerical simulation to obtain the acoustic field can be performed. Here, our main focus is on the sensitivity analysis of the parameters involved in the inverse scheme and its application to real-world sound source identification. Thereby, the challenges in the identification process of acoustic sources at low frequencies in an arbitrary environment are discussed, and the advantages and disadvantages of such an inverse scheme compared to advanced signal processing approaches (e.g., Clean-SC) are pointed out.

The rest of the paper is structured as follows: In Section 2 we summarize our developed inverse scheme and highlight the main features. In addition, a sensitivity analysis of the key algorithmic parameters is performed. In Section 3, the experimental environment, where the sound pressure measurements take place, is described and characterized. Furthermore, we discuss the challenges of

the physical modeling and numerical simulation of the experimental environment. In addition, the algorithm and practical realization of an automatic determination of the positions of the recording microphones are presented. In Section 4, we evaluate the quality of the numerical FE model towards measurements, when the experimental environment is excited by an omni-directional loudspeaker. Section 5 presents and discusses in detail all results of the inverse scheme for sound source localization at low frequencies and compares them to state-of-the-art methods. Finally, Section 6 concludes the main findings and provides an outlook.

2 Inverse scheme

The inverse scheme is based on a sparsity promoting Tikhonov functional to match measured and simulated microphone signals. A main part of the scheme is the solution of the wave equation in the frequency domain (Helmholtz equation), which allows to fully consider realistic geometry and boundary condition scenarios.

2.1 Physical and mathematical model

Assuming that the original geometry of the setup including the boundary conditions and the Fourier-transformed acoustic pressure signals $p_m^{\text{ms}}(\omega)$ (ω being the angular sound frequency, $m = 1, \dots, M$) at the microphone positions \mathbf{x}_m are given, the physical model is represented by the Helmholtz equation. Here, we consider the following generalized form of the Helmholtz equation in the computational domain $\Omega = \Omega_{\text{acou}} \cup \Omega_{\text{damp}}$

$$\nabla \cdot \frac{1}{\varrho} \nabla p + \frac{\omega^2}{K} p = \sigma^{\text{in}} \quad \text{in } \Omega, \quad (1)$$

with the searched for interior σ^{in} acoustic sources, and

$$\varrho(\mathbf{x}) = \begin{cases} \tilde{\varrho}^{\text{eff}} & \text{in } \Omega_{\text{damp}} \\ \varrho_0 & \text{in } \Omega_{\text{air}} \end{cases},$$

$$K(\mathbf{x}) = \begin{cases} \tilde{K}^{\text{eff}} & \text{in } \Omega_{\text{damp}} \\ c_0^2 \varrho_0 & \text{in } \Omega_{\text{acou}} \end{cases},$$

including speed of sound c_0 as well as density ρ_0 . In this way, poroelastic materials (e.g., porous absorbers) can be considered as a layer of an equivalent complex fluid having a frequency-dependent effective density $\tilde{\varrho}^{\text{eff}}$ and bulk modulus \tilde{K}^{eff} . With this formulation the acoustic properties of materials can be adjusted to a certain absorption coefficient. For this purpose a large number of models for characterization have been established for poroelastic materials, see e.g. [21]. Furthermore, the sound sources on the surface are modeled by

$$\mathbf{n} \cdot \nabla p = -\sigma^{\text{bd}} \quad \text{on } \Gamma_{\text{src}}. \quad (2)$$

Since the identification is done separately for each frequency ω , the dependence on ω is neglected in the notation. Now, the considered inverse problem is to reconstruct σ^{in} and/or σ^{bd} from pressure measurements

$$p_m^{\text{ms}} = p(\mathbf{x}_m), \quad m = 1, \dots, M, \quad (3)$$

at the microphone positions $\mathbf{x}_1, \dots, \mathbf{x}_M$. For the acoustic sources the following ansatz is made,

$$\sigma^{\text{in}} + \sigma^{\text{bd}} = \sum_{n=1}^N a_n e^{i\varphi_n} \delta_{\mathbf{x}_n}, \quad (4)$$

with the searched for amplitudes $a_1, a_2, \dots, a_N \in \mathbb{R}$ and phases $\varphi_1, \varphi_2, \dots, \varphi_N \in [-\pi/2, \pi/2]$. Here, N denotes the number of possible sources and $\delta_{\mathbf{x}_n}$ the delta function at position \mathbf{x}_n .

2.2 Optimization based source identification

The fitting of the parameters by means of Tikhonov regularization amounts to solving the following constrained optimization problem

$$\begin{aligned} \min_{p \in U, a \in \mathbb{R}^N, \varphi \in [-\frac{\pi}{2}, \frac{\pi}{2}]^N} J(p, a, \varphi), \\ \text{s. t. Eq. (1) is fulfilled} \end{aligned} \quad (5)$$

where $a = (a_1, \dots, a_N)$, $\varphi = (\varphi_1, \dots, \varphi_N)$ and

$$\begin{aligned} J(p, a, \varphi) = \frac{\Psi}{2} \sum_{m=1}^M |p(\mathbf{x}_m) - p_m^{\text{ms}}|^2 + \alpha \sum_{n=1}^N |a_n|^q + \beta \sum_{n=1}^N \varphi_n^2 \\ - \rho \sum_{n=1}^N \left(\ln \left(\frac{\pi}{2} + \varphi_n \right) + \ln \left(\frac{\pi}{2} - \varphi_n \right) \right). \end{aligned} \quad (6)$$

In here, the box constraints on the phases φ_n are realized by a barrier term with some penalty parameter $\rho > 0$. This also helps to avoid phase wrapping artefacts. The penalty parameter ρ and the regularization parameters α and β are chosen according to the sequential discrepancy principle [22]

$$\rho = \frac{\rho_0}{2^x}, \quad \alpha = \frac{\alpha_0}{2^x}, \quad \beta = \frac{\beta_0}{2^x}. \quad (7)$$

with x the smallest exponent such that following inequality

$$\sqrt{\sum_{m=1}^M (p(\mathbf{x}_m) - p_m^{\text{ms}})^2} \leq \varepsilon,$$

is fulfilled, with ε being the measurement error.

Sparsity of the reconstruction is desired to pick the few true source locations from a large number N of trial sources. By choosing $q \in (1, 2]$ close to one, an enhanced sparsity can be obtained. The optimization uses a gradient method with Armijo line search exploiting the adjoint method to efficiently obtain the gradient of the objective function. Hence, the computational time does not depend on the

number of microphones M nor on the assumed number of possible sources N . Since the optimization algorithm uses different stopping criteria, a scaling factor

$$\Psi = \frac{\psi}{\max(|p_m^{\text{ms}}|)}, \quad (8)$$

is introduced in (6), where the maximum absolute value of the measured pressure $|p_m^{\text{ms}}|$ is scaled to an amplitude of ψ . Further details about the inverse scheme can be found in [20].

In the current implementation, the FE method is applied for solving the Helmholtz equation (1). Hence, the applicability of the inverse scheme towards computational time is mainly restricted to the low frequency range, since the discretization effort and therefore the number of degree of freedoms in 3D is of the order $O(e_{\text{size}}^{-3})$, where e_{size} is the mesh size being determined by

$$e_{\text{size}} \approx \frac{\lambda_{\text{min}}}{N_e} = \frac{c_0}{N_e f_{\text{max}}}. \quad (9)$$

In (9) f_{max} denotes the highest frequency of the acoustic sources, c_0 the speed of sound and N_e should be at least 10 for linear FE basis functions [23]. Please note that any other numerical method, e.g., the boundary element method (BEM) can be applied, which may be even more efficient with respect to computation time, depending on the particular scenario.

2.3 Sensitivity analysis

For identifying important parameters that dominate the model behavior, sensitivity analysis is a commonly used technique. In order to gain an understanding of the relationship between the input variables and the identification result, a global sensitivity analysis (gSA) using the Sobol' method [24] is carried out. It is a quantitative method which allows for an assessment on how much the input parameter X_i is more important than another.

As a demonstrative example a dipole source above a sound hard floor is chosen (see Fig. 1a). To approximate free radiation, a perfectly matched layer is used on the remaining sides. The chosen frequency is 1000 Hz, so that the wavelength λ is 0.34 m. The sound pressure field as well as the true source distribution can be found in Figure 1b and one of the results of the inverse scheme is shown in Figure 1c.

To emulate the sound pressure measurement a forward run on a fine FE mesh is done ($e_{\text{size}} = \lambda/20$), whereas the inverse run for the identification is carried out on a coarser grid with $e_{\text{size}} = \lambda/12$ in order to avoid an inverse crime (forward and inverse run on the same grid). The obtained results are used to investigate the dependence of the identification result on the inversion parameters α_0 , β_0 , ρ_0 , q and ψ . For reasons of simplicity and reduced calculation effort, the simulations are initially performed in 2D. The conclusions of these results are adopted for 3D examples as well as for the real-world situation presented in Section 3.

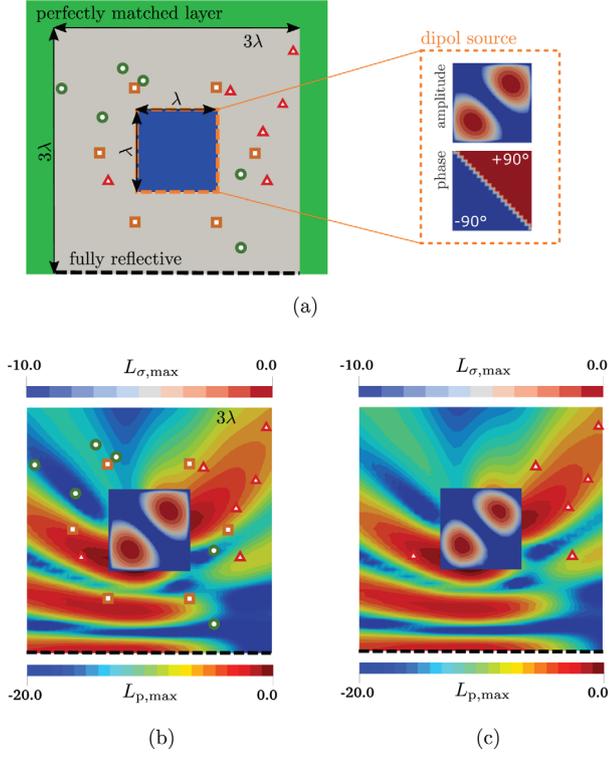


Figure 1. (a) Setup (wavelength λ), source distribution (amplitude and phase) as well as the microphone arrangements (indicated by Δ , \square , \circ , each with $M = 6$). (b) Computational setup with the normalized sound pressure level $L_{p,\max}$ and source level $L_{\sigma,\max}$ including the microphone arrangements. (c) Result of the inverse scheme with microphone positions at Δ (same reference value as in (b)).

For the quantification of the quality of the identification, three different error measures were used:

1. the relative L2 error between measured p_m^{ms} and simulated pressure values p_m^{inv} at M microphone positions

$$p_{\text{err}} = \sqrt{\frac{\sum_m^M (p_m^{\text{inv}} - p_m^{\text{ms}})^2}{\sum_m^M (p_m^{\text{ms}})^2}}, \quad (10)$$

2. the relative L2 error between the true source strengths σ_n^{ms} and the reconstructed σ_n^{inv} at all points N_{src} in the source region Ω_{src}

$$\sigma_{\text{err}} = \sqrt{\frac{\sum_n^{N_{\text{src}}} (\sigma_n^{\text{inv}} - \sigma_n^{\text{ms}})^2}{\sum_n^{N_{\text{src}}} (\sigma_n^{\text{ms}})^2}}, \quad (11)$$

3. and the relative L2 error between the true source strengths σ_n^{ms} and the reconstructed σ_n^{inv} at points in a circle with radius $\lambda/10$ around the actual source position in the source region

$$\sigma_{\text{err,loc}} = \sqrt{\frac{\sum_n^{N_{\text{circ}}} (\sigma_n^{\text{inv}} - \sigma_n^{\text{ms}})^2}{\sum_n^{N_{\text{circ}}} (\sigma_n^{\text{ms}})^2}}. \quad (12)$$

It should be stressed that sound pressure p and source strength σ are complex quantities. In real-life situations only the error between the sound pressure at the microphone positions p_{err} , equation (10), can be determined.

Three different microphone arrangements, which are used here as an example to demonstrate the identification, are shown in Figure 1b (indicated by Δ , \square , \circ , each with $M = 6$). Two arrangements (\square and Δ) show good source identification capability (low σ_{err} and $\sigma_{\text{err,loc}}$), whereas configuration \circ results in a bad identification result (see Tab. 1). Next, this setup, including the three different microphone layouts, is used to show the sensitivity to the parameters of the inverse scheme. Additionally, this setup was used to study the effect of the microphone positions and the number of microphones M (see [25]). In [20] and [25] it was shown that the inverse scheme performs well, also at high SNR levels (up to 7 dB).

The Sobol' method is a variance-based decomposition technique which provides a quantitative measure of how each input parameter and its interactions contributes to the output variance. In this popular method, the variance V of the model output $Y = f(X_1, \dots, X_P)$ is decomposed into contributions from effects of single parameters, combined effects of pairs of parameters, and so on (see [26]) leading to

$$V(Y) = \sum_{i=1}^P V_i + \sum_{i=1}^{P-1} \sum_{j=i+1}^P V_{ij} + \dots + V_{1,\dots,P}, \quad (13)$$

where $V(Y) = \text{Var}(Y)$, $V_i = \text{Var}(\text{IE}(Y|X_i))$, $V_{ij} = \text{Cov}(\text{IE}(Y|X_i), \text{IE}(Y|X_j))$, etc. The variance contribution of each input parameter, and their interacting contributions on the total output variance, respectively, are expressed as so-called Sobol' sensitivity indices (SI) [24]:

$$S_i = \frac{V_i}{V(Y)}, \quad S_{ij} = \frac{V_{ij}}{V(Y)}, \dots \quad (14)$$

The first order SI S_i provides the main effect of the parameter X_i namely the contribution of the i th input parameter to the total model variance. The second order SI S_{ij} is a measure of the impact of the interaction between parameter X_i and X_j . Higher order SI are usually not of practical interest [27]. Furthermore, total order SI were introduced in [28]

$$S_{Ti} = S_i + \sum_{j \neq i}^P S_{ij} + \dots \quad (15)$$

which takes into account the main effect of X_i and all its interactions with the other parameters (second-order and higher order effects). Therefore, S_{Ti} is always greater than S_i

Table 1. Error values of the three microphone arrangements. Inversion parameters: $\alpha_0 = 10^{-4}$, $\beta_0 = 10^{-3}$, $\rho_0 = 1$, $q = 1.1$, $\psi = 40$.

	p_{err} (%)	σ_{err} (%)	$\sigma_{\text{err,loc}}$ (%)
\triangle	3.55	27.17	5.90
\square	8.15	20.64	4.71
\circ	48.04	89.96	87.69

$$S_{T_i} \geq S_i. \quad (16)$$

If the considered model is additive, the following relations hold

$$\sum_{i=1}^P S_i = \sum_{i=1}^P S_{T_i} = 1, \quad \text{i.e.} \quad S_i = S_{T_i}. \quad (17)$$

For additive models, the effect of the independent parameter X_i does not depend on the values of other independent parameters X_j , hence no parameter interaction exists. Table 2 summarizes the interpretations that can be made on the SI. There are several tools freely available for performing sensitivity analyses, e.g., SAFE Matlab toolbox [29], SIMLAB a stand-alone software package [30], the C++ based PSUADE software [31, 32], the GUI-HDMR Matlab package [33] or the sensitivity analysis library in python (SALib) [34]. These tools offer different sampling techniques and SA methods. Here, to perform the SA, the tool PSUADE is used.

Before a SA can be carried out, the input (i.e. α_0 , β_0 , ρ_0 , q , ψ) and output parameters (i.e. p_{err} , σ_{err} , $\sigma_{\text{err,loc}}$) have to be defined. Moreover, the probability density functions and value ranges of the input parameters have to be specified (see Tab. 3 for the definition).

Furthermore, the number of model evaluations N_{eval} (also called sample points) must be chosen appropriately, since the amount of information revealed via SA depends heavily on the number of sample points [35]. The amount of model evaluations N_{eval} that are sufficient for a reliable Sobol' sensitivity analysis depend on the number of input parameters and model complexity [36]. In principle, $N_{\text{eval}} = Q(2P + 1)$ are necessary for obtaining the Sobol' SI (Q is the size of an initial Monte Carlo sample). Using the computational strategy proposed in [37], the computational costs can be reduced to $N_{\text{eval}} = Q(P + 2)$. However, there is no general agreement on the optimal sample size Q . In general, the higher the number of input parameters, the larger the sample size should be [36]. According to [38], the sample size Q was set to 1000, which should provide accurate first order SI (main effects). For obtaining the 95%-confidence intervals on the first and total SI, bootstrapping with re-sampling was applied [39, 40]. This can be used to check the appropriateness of the chosen sample size. To this end, the parameters should have narrow confidence intervals [36]. The generation of the input matrix as well as the computation of the Sobol' SI is performed in PSUADE. The model evaluations are made with openCFS [23], a FE-based multi-physics modeling and simulation tool.

Table 2. Interpretations regarding the Sobol' SI.

$S_i \approx 0$	Parameter X_i has no significant effect on the model output*
$S_{T_i} \approx S_i$	Parameter X_i has no significant interactions with others
$\sum S_i = 1$	Model is additive (no parameter interactions)
$\sum S_{T_i} = 1$	
$1 \gg \sum S_i$	Significant parameter interactions in the model exists
$\sum S_{T_i} \gg 1$	

*In [36], the threshold value is defined by 0.05.

Table 3. Value ranges of the input parameters for the SA (the ranges were chosen empirically). The parameters were assumed to be uniformly distributed.

	min	max
α_0	0	1
β_0	0	10
ρ_0	0	10
q	1	2
ψ	1	50

In the plots of the sensitivity indices (SI) (Figs. 2a–2c) also the bootstrapped result (bootstrap sample size was 100) is shown (dark color bars). This was done to check the appropriateness of the sample size. Since, all SI have narrow confidence intervals, the sample size for the Sobol' method can be assumed to be sufficient. Considering the results it can be observed that β_0 and ρ_0 have a SI < 0.05 in all settings, which is an indication that these two parameters have no significant effect on the output. Since the sum of total order SI is higher than one in all settings, significant parameter interactions exist. When considering p_{err} (Fig. 2a), the main effect is primarily determined by the parameters α_0 and ψ . A similar conclusion can be drawn for $\sigma_{\text{err,loc}}$ (Fig. 2c), where the main effect is due to ψ . For σ_{err} (Fig. 2b), a clear statement cannot be given, but ψ or α_0 are suspected to determine the main effect. The parameter q has a small first order SI in all settings, but a high total order SI which indicates that it has significant interactions with the other parameters.

Hence, the following conclusions can be drawn from the sensitivity analysis:

- ψ and α_0 are the most significant parameters for all model outputs,
- q has significant interactions with others,
- β_0 and ρ_0 have neither significant effect on the model output, nor strong interactions with other parameters.

A further, very important point for a successful source identification is the positioning and number of the microphones. In [20] and [25] it is shown that a small p_{err} does not necessarily lead to a good source identification if the number of microphones is too small or the microphone positions are not sensitive enough. For precise source identification, the microphones should be positioned at locations,

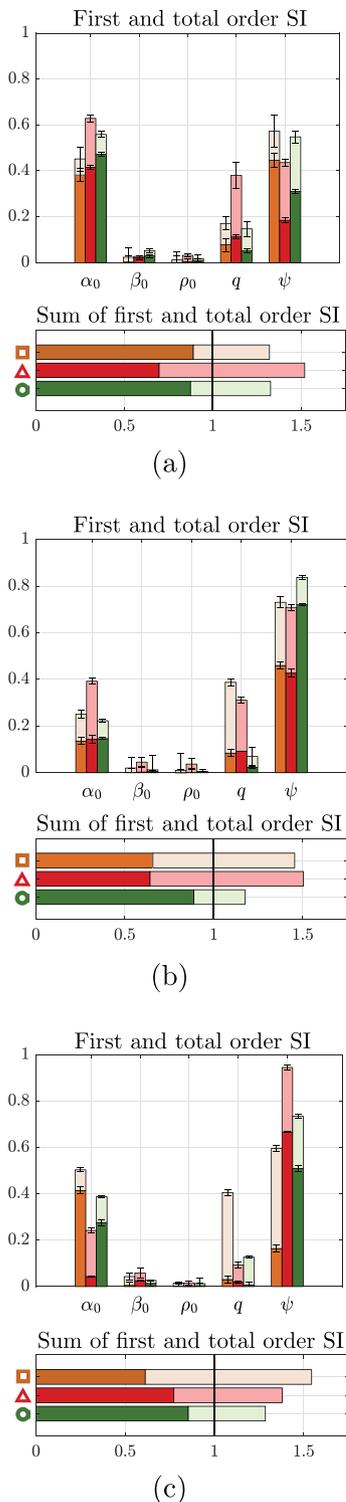


Figure 2. Sobol' SI (first and total order) for the different model outputs (a) to (c).

where the acoustic pressure has a maximum. Since these optimal microphone positions cannot be determined before the actual measurement, a successful source localization may need re-positioning of microphones. It is also preferable to surround the sound source with microphones and to allow for a spatial distribution of microphones near the

source. Furthermore, the microphone-to-source points ratio $\chi = M/N$ should be at least 4%, see [25]. Consequently, a higher discretization of the source region will lead to a higher number of microphones. Ratios higher than $\chi > 35\%$ should be avoided, since an enhancement in the source identification is not expected, due to numerical noise.

3 Real-world situation

3.1 Experimental setup

To demonstrate the applicability of the inverse scheme in real-world scenarios, microphone array measurements were performed in a room where a generic sound source was located. The room is partially lined with porous absorbers (Baso Plan 100) on the walls and ceiling, leading to the reverberation time shown in Figure 3. A ground plan of the room is shown in Figure 4 including the sound source location. There is also a cupboard in the room and the DAQ (data acquisition) unit for recording the sound pressure. The sound source is a box (dimensions: $50 \times 100 \times 55$ cm) made of Doka formwork sheet, see Figure 5a. The sound can be generated by three separately excitable loudspeakers (VISATON WS 20 E, $\theta = 8^\circ$). In order to characterize the source, the normal velocity is measured with a laser scanning vibrometer (LSV) Polytec PSV-500. In the measurement, first speaker L1 and afterwards L2 was active. The excitation frequency of the speakers was 250 Hz. The measured normal velocity level L_v is depicted in Figures 5b and 5c. Here, just the side with the active speaker is illustrated.

3.2 Physical and numerical modeling

The application of the inverse scheme needs physical and geometrical modeling. In doing so, one has to deal with some challenges concerning real versus ideal world, which will be discussed in this section. Three main challenges will be considered in detail:

- geometrical modeling needed for the application of FEM,
- physical modeling of boundary conditions, and
- the determination of the microphone positions.

The presented inverse scheme requires an appropriate FE model of the measurement environment. The model was created with as few details as necessary to keep the model as simple as possible. The geometric modeling is not completely independent of the physical modeling (boundary conditions), which should be kept in mind.

The created FE model is depicted in Figure 6. The cupboard and the DAQ unit in the room are modeled by simple blocks. Moreover, the I-beam on the ceiling was considered in the FE model. The supports for the microphones (microphone trees, see Fig. 9) and the mounting frame (see Fig. 5) of the generic sound source were neglected in the modeling process, since the dimensions are small compared to the wavelength according to the excitation frequency. The grey region Ω_{damp} represents the porous absorber on the walls

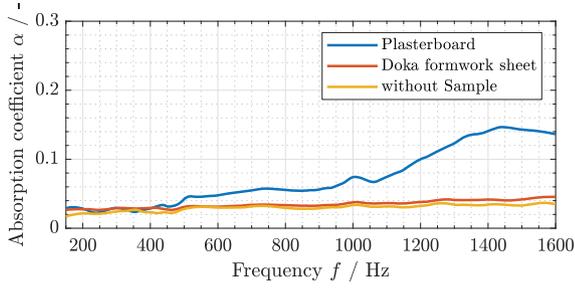


Figure 7. Absorption curves for plasterboard and for the Doka formwork sheet obtained by the 2p-method (ISO 10534-2) [41]. To see whether an absorption effect is present, the curve obtained when no sample is placed in the tube is also shown.

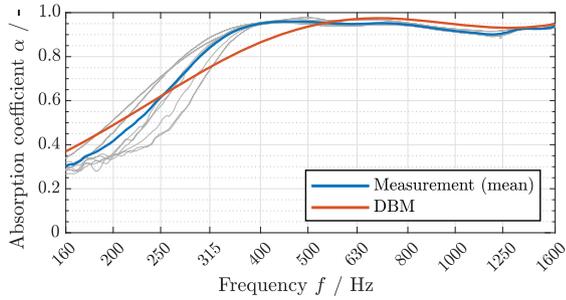


Figure 8. Absorption curves for Baso Plan 100 obtained by measurement (grey curves, mean value shown in blue) and by the DBM model.

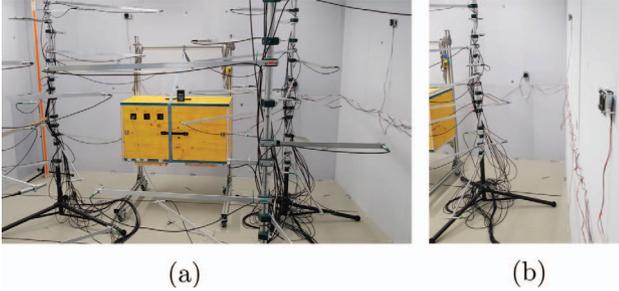


Figure 9. (a) Room with the generic sound source (yellow box) and the microphone trees as well as (b) the mounted speakers on the wall.

$$\nabla p \cdot \mathbf{n} = 0 \quad \text{on} \quad \partial\Omega \cup \Gamma_{\text{plaster}} \cup \Gamma_{\text{src}} \cup \Gamma_{\text{floor}} \cup \Gamma_{\text{cb}} \cup \Gamma_{\text{daq}} \cup \Gamma_{\text{lbeam}}. \quad (18)$$

Next, the modeling of the porous absorber (Baso Plan 100), denoted by Ω_{damp} , on the walls and on the ceiling is discussed. Therefore, first, impedance tube measurements were conducted to characterize the properties of the absorber (see Fig. 8 for the obtained absorption curves). The measurements with the impedance tube (grey curves, mean value shown in blue) spread considerably in the lower frequency range, whereas nearly a constant absorption behavior above 400 Hz is observed.

For the modeling of the porous absorber an equivalent fluid model (assuming isotropic and volume averaged features) is used. For this purpose the Delany-Bazely-Miki (DBM) model is used, which is purely empirical and derived from measurements on many highly porous materials. The basis of the model is presented in [42] and was adapted by [43]. Following formulas are given for the characteristic impedance \tilde{Z}_c and the wave number \tilde{k}

$$\begin{aligned} \tilde{Z}_c &= \varrho_0 c_0 \left[1 + 5.50 \left(10^3 \frac{f}{\Xi} \right)^{-0.632} - j 8.43 \left(10^3 \frac{f}{\Xi} \right)^{-0.632} \right], \\ \tilde{k} &= \frac{\omega}{c_0} \left[1 + 7.81 \left(10^3 \frac{f}{\Xi} \right)^{-0.618} - j 11.41 \left(10^3 \frac{f}{\Xi} \right)^{-0.618} \right]. \end{aligned} \quad (19)$$

Here, the only material parameter is the static air flow resistivity Ξ . The limits for the validity of the model is given by [42]

$$0.01 < \frac{f}{\Xi} < 1, \quad (20)$$

whereas the lower limit is not that critical for the quality, which means that the model behaves well below this limit [43]. The static air flow resistivity can be measured according to [44], but here an inverse determination was chosen by adapting the material model to the measured absorption values. For this process, the sample thickness of the absorber was measured ($h = 100$ mm). All other material parameters were fitted with a least squares algorithm by comparing the absorption coefficient measurements α_M to the respective absorption curve α_A of the material model.

This approach is based on measurements for which analytical solutions (models) for the measured quantity exists. In the considered case, the measurement is done in an impedance tube using the 2p-method. For this purpose a sample of the material to be characterized is placed inside the tube. This sample influences the sound wave propagation (plane wave propagation) in the tube for which mathematical models exist. Thus, the absorption coefficient can be expressed considering the measurement conditions at the 2p-method

$$\alpha_A = 1 - |\tilde{R}|^2 = 1 - \left| \frac{\tilde{Z}_S - Z_{c,\text{air}}}{\tilde{Z}_S + Z_{c,\text{air}}} \right|^2, \quad (21)$$

with \tilde{Z}_S the surface impedance at the transition between air and absorber, and $Z_{c,\text{air}}$ the characteristic impedance of air. For perpendicular sound incidence and due to the fact that the sample is backed by a rigid wall in the tube, the surface impedance can be determined by [45]

$$\tilde{Z}_S = -j \tilde{Z}_c \cot(\tilde{k}h). \quad (22)$$

resulting in

$$\alpha_A = 1 - |\tilde{R}|^2 = 1 - \left| \frac{-j \tilde{Z}_c \cot(\tilde{k}h) - Z_{c,\text{air}}}{-j \tilde{Z}_c \cot(\tilde{k}h) + Z_{c,\text{air}}} \right|^2. \quad (23)$$

The attenuation effect in the absorber is considered by the complex quantities for the wave number \tilde{k} and the characteristic impedance \tilde{Z}_c (j denotes the imaginary unit). These quantities result from the corresponding material model, here the DBM model (19).

Next, the material parameters, which minimize the error between measured α_M and analytical absorption curve α_A are searched. The measured absorption curve consists of F measurement points $(f_i, \alpha_{M,i})$ with $i = 1 \dots, F$. Here, the squared L2 error is used as error measure

$$J_{\text{err}} = \sum_{i=1}^F [\alpha_{M,i} - \alpha_A(f_i)]^2. \quad (24)$$

Hence, an optimization problem has to be solved to find those material parameters that minimize the error J_{err} . Normally, and especially for nonlinear problems, such optimization problems are solved numerically, where the specification of a starting point (initial values) is necessary. For the optimization problem the gradient based Levenberg–Marquardt [46, 47] algorithm was applied. Since gradient based algorithms may converge only locally, a combination with a genetic optimization algorithm was used. Here, the genetic algorithm “optimizes” the initial model parameters that were used afterwards in the Levenberg-Marquardt algorithm. In Figure 8, the fitted absorption curves of the DBM model is compared with the measurements. The material model was fitted to the mean value of the seven absorption curve measurements and resulted in a $\Xi = 8.92\text{e}3 \text{ Nsm}^{-4}$ for the DBM model. The fit with the model shows a good agreement with the measurements.

To consider the porous absorber in the FE simulation, the generalized form of the Helmholtz equation (1) is used, which is suitable for an inhomogeneous medium (spatially varying density). Here, the searched for interior σ^{in} acoustic sources will be set to zero, since no sources are assumed to be in a volume. With the formulation (1), the porous absorber on the wall and ceiling is considered as a layer of an equivalent complex fluid. The frequency-dependent effective parameters \tilde{q}_{eff} and \tilde{K}_{eff} are obtained by

$$\begin{aligned} \tilde{q}_{\text{eff}}(\omega) &= \frac{\tilde{k}(\omega)\tilde{Z}_c(\omega)}{\omega}, \\ \tilde{K}_{\text{eff}}(\omega) &= \omega \frac{\tilde{Z}_c(\omega)}{\tilde{k}(\omega)}, \end{aligned} \quad (25)$$

employing the expressions for the characteristic impedance \tilde{Z}_c and the wave number \tilde{k} of the DBM model (19). The sound sources on the surface of the generic source are modeled by (2).

3.3 Microphone positions

The presented approach for sound source localization require not only a suitable FE model but also precisely determined positions of the microphones with respect to a reference position. Through many simulations, it has been shown that spatially distributed microphone positions lead to the best identification results. This may be due to a higher sensitivity of the microphones compared to, e.g. a planar array. For the positioning within the room so-called

microphone trees have been constructed (see Fig. 9) to achieve a spatial distribution of the microphones. These trees can have different branches with different lengths at various heights to which the microphones are attached. In order to determine the microphone locations in the room, an acoustic positioning system was developed. The system is based on the principle of multi-lateration, where distances between an unknown position (= microphone) and several known points (= loudspeakers at the walls, see Fig. 9) are used to determine the location. This setting is similar to the well-known global-positioning system (GPS) [48], which is mainly used for outdoor positioning purposes. If the distances r_i between microphone and speakers \mathbf{s}_i ($i = 1, \dots, N_S$) are known, the position of the microphone can be obtained by using the sphere equation

$$(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2 = r_i^2, \quad (26)$$

with the speaker coordinates $\mathbf{s}_i = (x_i, y_i, z_i)^T$ and the searched for microphone coordinates $\mathbf{y} = (x, y, z)^T$ (Cartesian coordinate system). For a unique solution of the set of sphere equations (26), four speakers are used, three of which are essential. In trilateration, three equations are used to solve (26) which provides two possible solutions. For many situations it is clear which of the two solutions can only be correct (e.g., at the GPS the second solution would not be on earth but in outer space). Since this is not clear in the present case, multi-lateration is applied.

The distances r_i in (26) can be rewritten using the time delay τ_i between speaker and microphone, and the propagation speed c_0 (= speed of sound)

$$r_i = \tau_i c_0. \quad (27)$$

Hence, to obtain the distance r_i , a sound signal (sin-burst, 16 kHz, 6 periods) is emitted from a speaker and the time delay between emitting and receiving this signal (at the microphone) is determined. An overview of different time delay estimation techniques in room acoustics is given in [49]. Furthermore, the speed of sound c_0 (ϑ , Φ) must be known. For this purpose, the relative humidity of air Φ and the temperature ϑ (in °C) is measured and the speed of sound is estimated using [50]

$$c_0(\vartheta, \Phi) = 47.6215 \frac{\text{m}}{\text{s}} \sqrt{1 + \frac{\vartheta}{273.15 \text{ °C}}} \sqrt{\frac{\kappa_w(\vartheta, \Phi)}{M_w(\vartheta, \Phi)}}. \quad (28)$$

$$\kappa_w(\vartheta, \Phi) = \frac{700p_0 + e(\vartheta)\Phi}{500p_0 + e(\vartheta)\Phi},$$

$$M_w(\vartheta, \Phi) = 0.0289 - \frac{0.00011e(\vartheta)\Phi}{p_0},$$

with p_0 the static pressure and $e(\vartheta)$ the temperature dependent vapor pressure of air which can be calculated using the Antoine-equation [51]

$$e(\vartheta) = 133.322 \times 10^{8.07131 - \frac{1730.63}{233.426 + \vartheta}}. \quad (29)$$

Then, the distances between transmitter and receiver can be obtained by the time delay τ and by knowing the propagation speed c_0 (27). Here, the time delay is obtained by cross correlation (CCO). In order to improve robustness against reverberation usually generalized cross correlation with phase transform (GCC-PHAT) may be applied [52]. However, an improvement using GCC-PHAT could not be observed here. For the correlation, a high sampling rate is advantageous. At the used DAQ unit, the highest possible sampling frequency was 192 kHz.

Besides the determination of the distances and the measurement of the temperature ϑ and relative humidity Φ , also the set of equations (26) has to be solved to obtain the microphone coordinates. There exist several algorithms belonging to multi-lateration for solving (26) [53–55]. In the actual system, a combination including a Kalman-filter (KF) [56] is applied. First, the system (26) is rewritten as an optimization problem

$$\hat{\mathbf{y}} = \min_{\mathbf{x}} \sum_{i=1}^{N_S} \text{res}_i(\mathbf{x})^2, \quad (30)$$

with

$$\text{res}_i(\mathbf{x}) = [(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2] - r_i^2,$$

where $\mathbf{x} = [x, y, z]^T$ and $\hat{\mathbf{y}}$ is the optimum, i.e. the searched position of the microphone. For this, one of the sphere equations (marked by \square^*) is subtracted from the remaining $I = N_S - 1$. This process leads to the following equation

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b} \quad (31)$$

with

$$\mathbf{A} = \begin{bmatrix} x_1 - x_* & y_1 - y_* & z_1 - z_* \\ x_2 - x_* & y_2 - y_* & z_2 - z_* \\ \vdots & \vdots & \vdots \\ x_I - x_* & y_I - y_* & z_I - z_* \end{bmatrix}$$

$$\mathbf{b} = \frac{1}{2} \begin{bmatrix} x_1^2 - x_*^2 + y_1^2 - y_*^2 + z_1^2 - z_*^2 + d_*^2 - d_1^2 \\ x_2^2 - x_*^2 + y_2^2 - y_*^2 + z_2^2 - z_*^2 + d_*^2 - d_2^2 \\ \vdots \\ x_I^2 - x_*^2 + y_I^2 - y_*^2 + z_I^2 - z_*^2 + d_*^2 - d_I^2 \end{bmatrix},$$

which may be solved by

$$\hat{\mathbf{y}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}. \quad (32)$$

The question which of the N_S equations is used for subtraction is still open. Here, no fixed equation is used. Since the searched position should be independent of the chosen reference equation \square^* , the process is performed by choosing each equation once as reference and calculating the microphone position $\hat{\mathbf{y}}$. Moreover, also the number of used equations N_S for obtaining $\hat{\mathbf{y}}$ could be varied, where – as pointed out above – four are necessary. The total number of possible combinations can be calculated by

$$\sum_{N_S=4}^{N_{S,\max}} \frac{N_{S,\max}!}{N_S!(N_{S,\max} - N_S)!}, \quad (33)$$

with the maximum number of available speakers on the walls $N_{S,\max} = 8$. Hence, this procedure results in 163 estimations of the microphone position \mathbf{y} . Next, the KF is applied to this data. The KF is an optimal estimation algorithm [57] and can be used for the state estimation (= searched microphone position) of a system with uncertain measurements. The process of the application is depicted in Figure 10. It is a recursive process in which the state estimation is improved with each new input measurement (= obtained position through (32)), leading in an optimal estimation $\hat{\mathbf{y}}$ for \mathbf{y} . The following definitions have been made for the KF

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{R} = \begin{bmatrix} \sigma_r^2 & 0 & 0 \\ 0 & \sigma_r^2 & 0 \\ 0 & 0 & \sigma_r^2 \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} s_q^2 & 0 & 0 \\ 0 & s_q^2 & 0 \\ 0 & 0 & s_q^2 \end{bmatrix},$$

with the state transition matrix \mathbf{A} , the observation matrix \mathbf{H} , measurement (noise) covariance matrix \mathbf{R} , and process (noise) covariance matrix \mathbf{Q} . The starting value for the error covariance \mathbf{P} of the state was set to \mathbf{R} . The standard deviation for the measurement was assumed to be $s_r = 0.1$ m and for the process $s_q = 0.1$ m (empirically chosen). The overall positioning error $\|\mathbf{y} - \hat{\mathbf{y}}\|$ is 37.3 ± 44.7 mm (determined at 70 locations in the room). This accuracy seems to be sufficient for now, since the identification by the inverse scheme is currently limited to low frequencies in the actual implementation (frequencies lower than 500 Hz seem to be practicable). It has to be noted that the determined position near the floor and ceiling have the highest deviations compared to the true position of the microphone. Moreover, the positions of the speakers on the walls will also have an influence on the localization result of the microphones.

4 Model validation

In order to check the appropriateness of the applied modeling approaches, the FE model needs to be validated. Therefore, a loudspeaker (Bruel & Kjaer OmniPower Sound Source Type 4292) was placed in the room and excited by white noise (bandwidth 1000 Hz). In the simulation, the source is modeled as point source (omni-directional). Since sound sources show such behavior at low frequencies and also, due to the fact that a special omni-directional loudspeaker was used, this approach has been followed.

To assess the suitability of the FE model, the relative L2 error p_{err} (10) and the spatial distribution of the sound

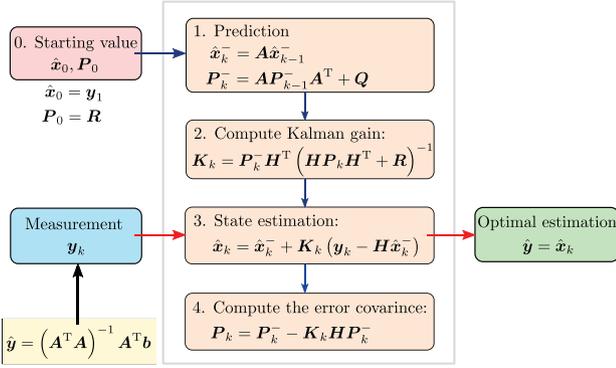


Figure 10. Application of the Kalman-Filter (KF) to improve the position obtained by the linearized sphere equations leading to an optimal estimation of the microphone position.

pressure for each frequency at the microphone positions were compared with the measurement. Here, the regression coefficient R^2 was calculated to quantify the model adequacy for the spatial distribution, which is determined by

$$R^2 = 1 - \frac{\sum_{m=1}^M (p_m^{\text{ms}} - p_m)^2}{\sum_{m=1}^M (p_m^{\text{ms}} - \text{mean}(p_m^{\text{ms}}))^2}, \quad (34)$$

with p_m^{ms} the measured pressure and p_m the pressure obtained from the simulation. Since the source strength is not known, a scaling factor “sc” for the calculation is determined by means of least squares minimization ($\min_{sc} \sum (p_m^{\text{ms}} - p_m)^2$). This scaling factor is used to scale the simulation before the comparisons were done, to account for the unknown source strength. This approach is possible, since a linear acoustic field is assumed. In Figure 11, the obtained regression coefficient and the relative L2 error are shown. It can be observed, that the model is quite good up to 500 Hz, but for higher frequencies both assessment criteria become worse. The investigations have shown that a regression coefficient R^2 of approximately 0.5 or higher is needed for a useful localization result. Figure 11 shows that this requirement is just fulfilled up to 350 Hz. Note, the microphone positions that were used for validation are different from those used for the source identification (see Fig. 12).

5 Source identification

In comparison to the inverse scheme, three common beamforming algorithms are used for the source identification. ConvBF [1] and FuncBF [2, 3] were applied, which can also handle coherent acoustic sources. Furthermore, the Clean-SC [5] deconvolution algorithm was employed. This algorithm removes the side lobes from the raw source map based on spatial source coherence. However, it will not work satisfactorily with several tonal (i.e. coherent) sound sources. Steering vector formulation II (according to [15]) was used for the beamforming algorithms providing the

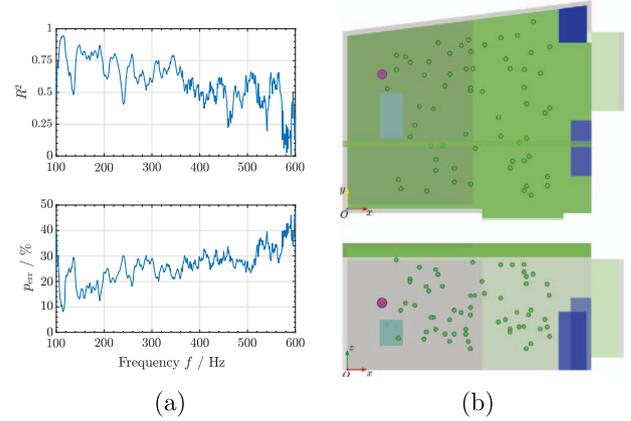


Figure 11. Model validation using an omni-directional loudspeaker: (a) R^2 and p_{err} -values (b) microphone positions (magenta dots) and source (magenta dot) position in the room.

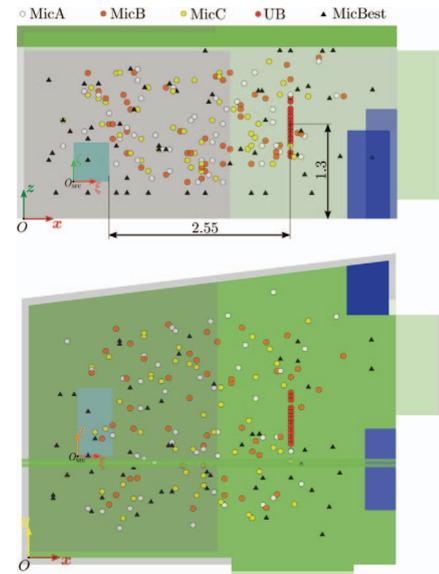


Figure 12. Microphone locations (MicA, MicB, MicC, UB and MicBest) in the considered measurement environment (dimensions in m).

correct source strength. It should be noted that the commonly used steering vector formulations do not use the available information of the room. To overcome this limitation, an approach combining measurement with simulation can be applied. Thereby, the transfer function between source and microphone positions is computed by a FE simulation leading to numerically calculated transfer functions (NCTF) as demonstrated in [25]. In doing so, the actual measurement setup is taken into account. Compared to the standard formulations, the NCTF leads to an improvement in resolution as well as dynamic range. However, the results at these low frequencies just showed minor improvements [25].

For the identification, four different microphone arrangements are considered. The first one is an Underbrink (UB) array (2D layout, aperture 1 m). At the other three

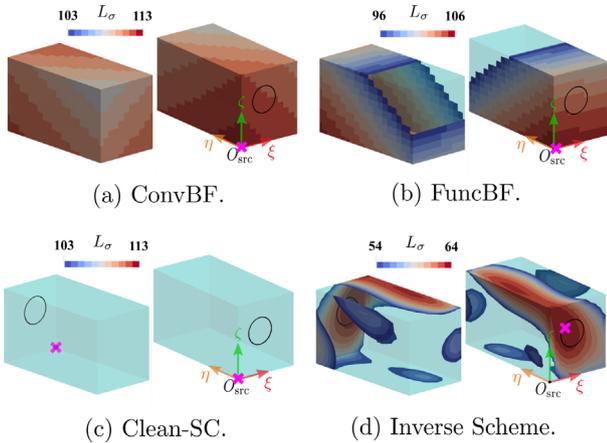


Figure 13. Localization results of speaker L1 obtained with different methods utilizing the UB array (the true source location is marked by a black circle). Two views of the generic source (Fig. 5) are shown for each result. The point with the highest source level is marked with a cross in magenta.

arrangements (MicA, MicB, MicC), the microphones are spatially distributed throughout the room without any special requirements for their positions. However, care was taken not to place them too close to the ceiling, the floor and the walls. The used microphone positions in the room are depicted in Figure 12. Furthermore, an optimized arrangement named MicBest was used for the inverse scheme, but only with synthesized measurement data (presented by black triangles). It should be noted that in the UB array 31, whereas for the others (MicA, MicB, MicC, MicBest) 50 microphones are used. This results in a microphone-to-source points ratio of $\chi \approx 3\%$ respectively $\chi \approx 4\%$. The parameters for the inverse scheme were set to: $\alpha_0 = \beta_0 = \rho_0 = 0.125$, $q = 1.1$, $\psi = 40$.

The temperature in the room was $\vartheta = 25^\circ \pm 2^\circ \text{C}$ and the relative humidity $\Phi = 25 \pm 10\%$ at the measurements. In order to consider the effects of these environmental conditions on c_0 , (28) is used. The considered source frequency is 250 Hz ($\lambda \approx 1.36 \text{ m}$).

5.1 Single speaker active

In the first example, speaker L1 on the left side of the generic sound source (see Fig. 5a) is active (the active speaker is marked by a black circle in the following source maps). First, the beamforming methods utilizing the UB array are applied (see Figs. 13a–13c). In the localization results (source maps) the source level L_σ (ref. $20 \mu\text{Pam}$) is shown. The localization results are rather poor, since the source position is detected at a wrong location. But the effect of improving the dynamic range through the different algorithms can be observed.

At low frequencies, beamforming methods usually do not perform well, especially when many sources are involved. Then it will hardly be possible to locate them individually. However, this is not the case in this example, since there is just one active sound source. Therefore, the

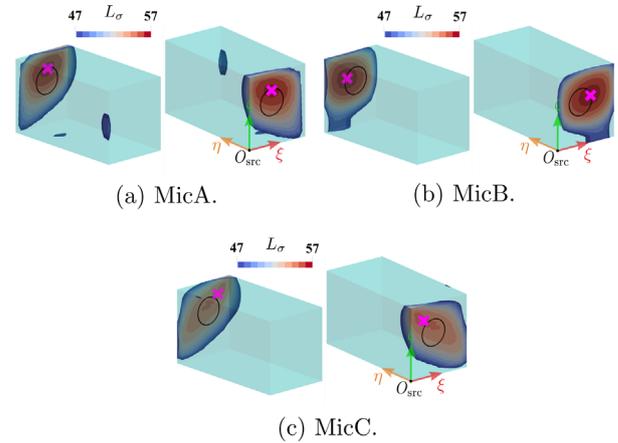


Figure 14. Localization results of speaker L1 obtained with the inverse scheme using different microphone arrangements (the true source location is marked by a black circle and the point with the highest source level with a cross in magenta).

standard methods should not perform poorly. However, the small aperture of the UB array and the fact that the sound waves from the source do not impinge directly on the array, the localization becomes increasingly difficult. In addition, the reflective surfaces in the room make localization even more challenging.

Next, the inverse scheme with the UB array measurement data was used. The area of the active sound source could be identified correctly. However, on the top of the source other spurious acoustic sources have been identified (see Fig. 13d). This can be an indication that the microphone positions are not sensitive enough or there are too few microphones. Considering the guidelines in [25] also a spatial distribution is desirable.

To further improve the localization result, other microphone arrangements (MicA, MicB, MicC) have been used (see Fig. 14). Compared to the previous results, there is a clear improvement of the localization, which indicates the importance of microphone positions. Thereby, MicA and MicB yield a similar localization result which outperforms those of MicC.

Beamforming was also done with the measured sound pressure data of these three arrangements, to have a comparison with the results of the inverse scheme. Due to the spatially distributed microphones a considerable improvement can also be expected there. Exemplarily, the localization performance of MicA is given in Figure 15.

So far, only localization results have been considered, which indicates whether the source was identified at the correct position. In order to make quantitative statements about the identified source distribution, the identified sources will be used to perform a sound field computation. This allows comparisons with the original acoustic field measured at the microphone positions. For the inverse scheme this is straightforward and no further steps need to be taken, since a detailed source distribution both in amplitude and phase is identified. Hence, a numerical simulation

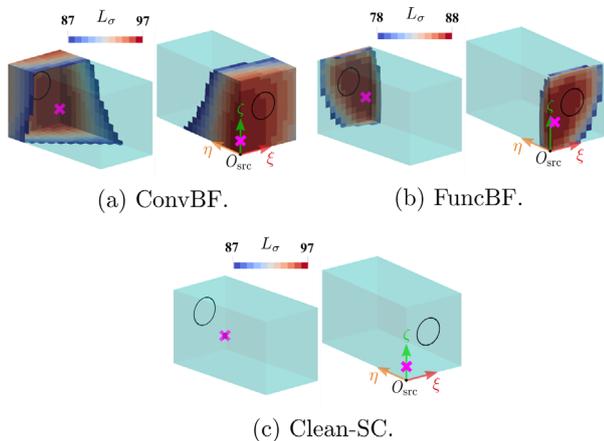


Figure 15. Localization results of the beamforming methods using MicA (the true source location is marked by a black circle and the point with the highest source level with a cross in magenta).

Table 4. Error value p_{err} in % using the identified sources depending on the microphone arrangement and the localization methods.

	ConvBF	FuncBF	Clean-SC	Inverse scheme
MicA	82	82	80	22
MicB	107	84	107	20
MicC	105	80	105	21
UB	218	135	219	11

was performed to obtain the acoustic field. For the quantification of the obtained result of the sound field computation the relative L2 error p_{err} between measured and simulated pressure values at the microphone positions (10) is used. The results are given in Table 4.

The best acoustic field reconstruction is achieved by the simple 2D UB array considering the p_{err} -values. However, this may lead to a wrong conclusion, if only this error value is considered. The source maps obtained with the UB array (Fig. 13) and the other 3D arrays (Fig. 14) show that the UB array does not provide a good localization result, although p_{err} is the smallest. This fact clearly indicates that other microphone positions with a higher sensitivity should be used or the number of microphones should be increased to improve the localization result. But also with the other arrangements, a good agreement with the measured pressure was obtained ($p_{\text{err}} \approx 20\%$). Considering the localization result and the source field reconstruction the arrangement MicB is the best.

For the beamforming source maps, the acoustic field computation is not as simple as for the source distributions obtained by the inverse scheme. Main problem is given through the point spread function (PSF) of the array, since the computed source map (raw map) is a convolution of the PSF with the real source distribution. Deconvolution algorithms (like Clean-SC) try to eliminate the influence of the PSF from the raw source map resulting in a deconvolved

map. Hence, the source maps obtained with Clean-SC can be integrated without further processing. For the ConvBF and FuncBF results, the source power integration technique [58–60] was applied to limit the effect of the PSF. To this end, the raw source map is normalized by the integrated PSF for a point source in the center of the integration area. Before the integration, an integration area must be specified. For this purpose, the maximum of the source map was taken as the center point of the integration area. From this center point, a sphere with radius 0.1 m was assumed and all surface points in this sphere were used for the integration. The results in Table 4 demonstrate the main advantage of the inverse scheme, namely an accurate identification of the source distribution with amplitude and phase for the computation of the acoustic field.

Next, the identified normal velocity level L_v by the inverse scheme utilizing MicB is considered in order to have a comparison with the LSV measurement data (see Fig. 5). Here, first the same situation as before, when speaker L1 on the left side of the source is active, is considered (Fig. 16a). The direct comparison with the LSV data shows a deviation of about 17 dB. In the next step, speaker L2 becomes the active sound source. This source radiates sound towards the floor. As before, the inverse scheme provides a good localization, but the amplitude again deviates by about 18 dB (Fig. 16b). To demonstrate the capability of the inverse scheme by using highly sensitive microphone positions (named by MicBest), we proceeded as follows. A forward simulation with prescribed normal velocity at the loudspeaker position (see Fig. 17a) was performed. The positions of the virtual microphones were determined according to the guidelines in [25] as the locations of maximal acoustic pressure. Since the acoustic field is known in the room through the forward simulation, the positions can be found easily. Therefore, first the maximum in the sound pressure field was searched and taken as microphone position, under the constraint of sufficient distance of the microphone from reflective surfaces. After this step the next position is determined. For this purpose, the region around the selected position is removed from the search space and the next maximum is searched, so that a spatial distribution is achieved as well. This procedure is followed until 50 positions are determined to have the same microphone number as MicB. In Figures 17b and 17c the various identification results using the two arrangements are depicted. Thereby, MicB (microphone arrangement as used in the measurements) shows a good localization result with a deviation of the velocity level of about 10 dB compared to the original one. However, using the microphone arrangement MicBest an almost perfect localization result as well as a good agreement in velocity level could be achieved (see Fig. 17c).

5.2 Two speakers active

The results achieved so far demonstrate the applicability of the proposed method for identifying low-frequency sound sources in real-world situations. An additional challenge is the localization (separation) of several active sound sources, especially in the low frequency range. To test

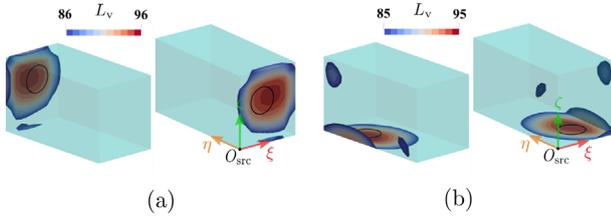


Figure 16. Identified normal velocity level L_v (ref. 50 nm/s) utilizing MicB: (a) speaker L1 and (b) speaker L2.

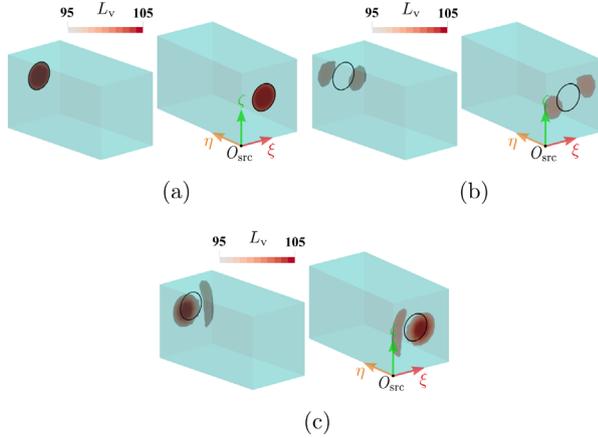


Figure 17. Identification results (normal velocity level L_v , ref. 50 nm/s) using different microphone arrangements (b) MicB, (c) MicBest as well as the (a) true velocity level.

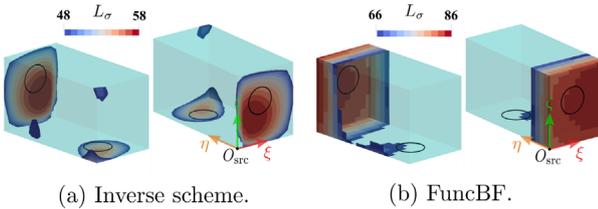


Figure 18. Comparison of the source separation between inverse scheme and FuncBF (case A).

the proposed method in the presence of more than one acoustic source, two setups have been considered: (case A) speaker L1 and L3 active and (case B) speaker L1 and L2 active. Both sources should have approximately the same source strength, since the excitation signal was the same for both. However, due to speaker tolerances, the same source level may not be achieved. By using the inverse scheme for localization a good identification especially for case A was achieved (see Fig. 18a). For case B (see Fig. 19a), the result is not as good, but we would like to note that this setup is more challenging, since the two sources are closer to each other which makes the separation harder.

The localization was also done with FuncBF (ConvBF is omitted, because FuncBF has the better ability for source

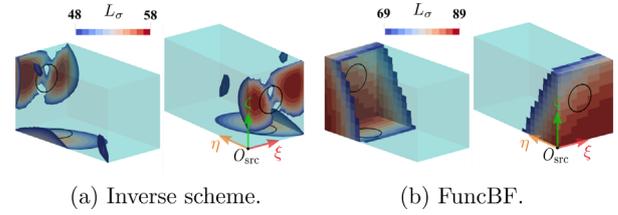


Figure 19. Comparison of the source separation between inverse scheme and FuncBF (case B).

separation). Since Clean-SC will not localize both sources (coherent sound sources), only the results of FuncBF are considered. It can be observed that in case A (Fig. 18b) FuncBF can separate the two sources, but the identified position of speaker L3 is more accurate with the inverse scheme. Moreover, the two source strengths of speaker L1 and L3 (which should be approximately equal) do not differ as much. For case B, FuncBF can not separate the two sources (see Fig. 19b).

6 Conclusion

In our developed inverse scheme the Helmholtz equation with the correct boundary conditions is solved. To recover the source locations, an inverse scheme based on a sparsity promoting Tikhonov functional to match measured (microphone signals) and simulated pressure is used. A clear advantage of such an inverse method is its ability to fully consider realistic geometry and boundary condition scenarios, as well as its straightforward generalizability to situations with convection and/or damping. The implemented optimization-based parameter identification algorithm relies on a gradient method with Armijo line search. For an efficient gradient computation of the objective function, the adjoint method is applied. By this approach, a detailed source distribution both in amplitude and phase is identified, and finally with this information a numerical simulation can be performed to obtain the acoustic field. State-of-the-art beamforming methods require an integration (question of the integration area arises) or more complex deconvolution algorithms must be applied.

Since a FE model is needed for the inverse scheme, the modeling was discussed in detail. Special attention is paid to the determination of boundary conditions, especially the modeling of the porous absorber as well as the determination of the microphone positions in the measurement environment. An equivalent fluid model is used for the modeling of partially absorbing surfaces. To obtain the model parameters for the complex fluid, an inverse parameter determination was applied, where the measurement of the absorption coefficient in an impedance tube served as starting point for the adjustment of the parameters. To obtain the microphone positions, a set of sphere equations (at least four equations are necessary) was used in combination with a Kalman-filter.

The localization results achieved with the inverse scheme demonstrate the applicability at low frequencies in

real-world scenarios. Furthermore, through simulations, it could be shown that a perfect reconstruction result of the acoustic sources can be achieved with the microphone positions at pressure maxima, which demonstrates the potential of the inverse scheme. Despite the superiority of the inverse scheme compared to advanced deconvolution based signal processing schemes one has to consider the high effort due to two challenges: (1) Geometry and boundary condition modeling for an accurate FE computation, and (2) the precise determination of the positions of the used microphones.

References

1. T.J. Mueller: Aeroacoustic measurements, Springer-Verlag, 2002; Beamforming in Acoustic Testing, 62–97.
2. R.P. Dougherty: Functional Beamforming, in: 5th Berlin Beamforming Conference, 2014, BeBeC-2014-01.
3. R.P. Dougherty: Functional beamforming for aeroacoustic source distributions, in: 20th AIAA/CEAS Aeroacoustics Conference, AIAA 2014-3066, 2014.
4. T.F. Brooks, W.M. Humphreys: A deconvolution approach for the mapping of acoustic sources (damas) determined from phased microphone arrays. *Journal of Sound and Vibration* 294 (2006) 856–879.
5. P. Sijtsma: *CLEAN based on spatial source coherence*. *International Journal of Aeroacoustics* 6, 4 (2007) 357–374.
6. T. Suzuki: L1 generalized inverse beam-forming algorithm resolving coherent/incoherent, distributed and multipole sources. *Journal of Sound and Vibration* 330 (2011) 5835–5851.
7. T. Yardibi, J. Li, P. Stoica, L.N. Cattafesta: Sparsity constrained deconvolution approaches for acoustic source mapping. *Journal of the Acoustical Society of America* 123 (2008) 2631–2642.
8. Q. Leclere, A. Pereira, C. Bailly, J. Antoni, C. Picard: A unified formalism for acoustic imaging based on microphone array measurements: *International Journal of Aeroacoustics* 164-5 (2017) 431–456.
9. G. Herold, E. Sarradj: Performance analysis of microphone array methods. *Journal of Sound and Vibration* 401 (2017) 152–168.
10. K. Ehrenfried, L. Koop: Comparison of iterative deconvolution algorithms for the mapping of acoustic sources. *AIAA Journal* 45, 7 (2007) 1584–1595.
11. Z. Chu, Y. Yang: Comparison of deconvolution methods for the visualization of acoustic sources based on cross-spectral imaging function beamforming. *Mechanical System and Signal Processing* 48 (2014) 404–422.
12. T. Yardibi, N.S. Zawodny, C. Bahr, F. Liu, L.N. Cattafesta, J. Li: Comparison of microphone array processing techniques for aeroacoustic measurements. *International Journal of Aeroacoustics* 9, 6 (2010) 733–761.
13. R. Merino-Martínez, P. Sijtsma, M. Snellen, T. Ahlefeldt, J. Antoni, C.J. Bahr, D. Blacodon, D. Ernst, A. Finez, S. Funke, T.F. Geyer, S. Haxter, G. Herold, X. Huang, W.M. Humphreys, Q. Leclère, A. Malgoezar, U. Michel, T. Padois, A. Pereira, C. Picard, E. Sarradj, H. Siller, D.G. Simons, C. Spehr: A review of acoustic imaging methods using phased microphone arrays. *CEAS Aeronautical Journal* 10 (2019) 197–230.
14. P. Chiariotti, M. Martarelli, P. Castellini: Acoustic beamforming for noise source localization: Reviews, methodology and applications. *Mechanical Systems and Signal Processing* 120 (2019) 422–448.
15. E. Sarradj: Three-dimensional acoustic sourcemapping with different beamforming steering vector formulations. *Advances in Acoustics and Vibration* 2012 (2012) 292695.
16. J. Fischer, C. Doolan: Beamforming in a reverberant environment using numerical and experimental steering vector formulations. *Mechanical Systems and Signal Processing* 91 (2017) 10–22.
17. E. Sarradj: Quantitative source spectra from acoustic array measurements, in: 2nd Berlin Beamforming Conference, 2008, BeBeC-2008-03.
18. E. Sarradj, T. Geyer, H. Brick, K.-R. Kirchner, T. Kohrs: Application of beamforming and deconvolution techniques to aeroacoustic sources at highspeed trains, in: NOVEM – Noise and Vibration: Emerging Methods, Sorrento, 2012.
19. R. Merino-Martínez, P. Sijtsma, A.R. Carpio, R. Zamponi, S. Luesuthiviboon, A.M.N. Malgoezar, M. Snellen, C. Schram, D.G. Simons: Integration methods for distributed sound sources. *International Journal of Aeroacoustics* 18, 4–5 (2019) 444–469.
20. M. Kaltenbacher, B. Kaltenbacher, S. Gombots: Inverse scheme for acoustic source localization using microphone measurements and finite element simulations. *Acta Acustica united with Acustica* 104 (2018) 647–656.
21. E. Deckers, S. Jonckheere, D. Vandepitte, W. Desmet: Modelling techniques for vibro-acoustic dynamics of poroelastic materials. *Archives of Computational Methods in Engineering* 22, 2 (2015) 183–236.
22. S.W. Anzengruber, B. Hofmann, P. Mathé: Regularization properties of the sequential discrepancy principle for Tikhonov regularization in Banach spaces. *Applicable Analysis* 93, 7 (2014) 1382–1400.
23. M. Kaltenbacher: Numerical simulation of mechatronic sensors and actuators: Finite elements for computational multiphysics, 3rd ed. Springer, 2015.
24. I.M. Sobol': Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling and Computational Experiments* 1 (1993) 407–414.
25. S. Gombots: Acoustic source localization at low frequencies using microphone arrays. PhD Thesis, TU Wien, 2020.
26. J. Nossent, P. Elsen, W. Bauwens: Sobol sensitivity analysis of a complex environmental model. *Environmental Modelling & Software* 26, 12 (2011) 1515–1525.
27. B. Iooss, P. Lemaître: A review on global sensitivity analysis methods, in: Uncertainty management in simulation-optimization of complex systems: Algorithms and applications, C. Meloni, G. Dellino, Editors. Springer, 2015.
28. T. Homma, A. Saltelli: *Importance measures in global sensitivity analysis of nonlinear models*. *Reliability Engineering & System Safety* 52, 1 (1996) 1–17.
29. F. Pianosi, F. Sarrazin, T. Wagener: A Matlab toolbox for global sensitivity analysis. *Environmental Modelling & Software* 70 (2015) 80–85.
30. S. Tarantola, W. Becker: SIMLAB software for uncertainty and sensitivity analysis, in: Handbook of Uncertainty Quantification, R. Ghanem, D. Higdon, H. Owhadi, Editors. Cham, Springer International Publishing, 2017, pp. 1695–1731.
31. C. Tong: Problem solving environment for uncertainty analysis and design exploration, in: Handbook of Uncertainty Quantification, R. Ghanem, D. Higdon, H. Owhadi, Editors. Cham, Springer International Publishing, 2017, pp. 1695–1731.
32. Y. Gan, Q. Duan, W. Gong, C. Tong, Y. Sun, W. Chu, A. Ye, C. Miao, Z. Di: A comprehensive evaluation of various sensitivity analysis methods: A case study with a hydrological model. *Environmental Modelling & Software* 51 (2014) 269–285.

33. T. Ziehn, A. Tomlin: GUI-HDMR – A software tool for global sensitivity analysis of complex models. *Environmental Modelling & Software* 24 (2009) 775–785.
34. J. Herman, W. Usher, SA Lib: An open-source python library for sensitivity analysis. *The Journal of Open Source Software* 2, 9 (2017) 97.
35. A. Saltelli, S. Tarantola, F. Campolongo, M. Ratto: *Sensitivity analysis in practice: a guide to assessing scientific models*, 2nd ed. Wiley Online Library, 2017.
36. X.-Y. Zhang, M. Trame, L. Lesko, S. Schmidt: Sobol sensitivity analysis: A tool to DE the development and evaluation of systems pharmacology model. *CPT: Pharmacometrics & Systems, Pharmacology* 4, 2 (2015) 69–79.
37. A. Saltelli: *Making best use of model evaluations to compute sensitivity indices*. *Computer Physics Communications* 145 (2002) 280–297.
38. A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola: *Global sensitivity analysis. The Primer*. John Wiley & Sons, 2008.
39. B. Efron, R. Tibshirani: *An Introduction to the Bootstrap*. Chapman & Hall/CRC, 1993.
40. R.W. Johnson: *An Introduction to the Bootstrap*. *Teaching Statistics* 23, 2 (2001) 49–54.
41. ISO 10534–2: Acoustics – Determination of sound absorption coefficient and impedance in impedance tubes – Part 2: Transfer-function method, 2001.
42. M.E. Delany, E.N. Bazley: Acoustical properties of fibrous absorbent materials. *Applied Acoustics* 3, 2 (1970) 105–116.
43. M.E. Delany, E.N. Bazley: Acoustical properties of porous materials. Modifications of Delany-Bazley models. *Journal of the Acoustical Society of Japan* 11, 1 (1990) 19–24.
44. ISO 9053–1: Acoustics – Determination of airflow resistance – Part 1: Static airflow method, 2019.
45. J.F. Allard, N. Atalla: *Propagation of sound in porous media: modelling sound absorbing materials*. 1st ed. John Wiley & Sons, 2009.
46. K. Levenberg: A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics* 2, 2 (1944) 164–168.
47. D.W. Marquardt: An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics* 11, 2 (1963) 431–441.
48. B. Hofmann-Wellenhof, H. Lichtenegger, J. Collins: *Global positioning system: theory and practice*. Springer Science & Business Media, 2012.
49. J. Chen, J. Benesty, Y. Huang: Time delay estimation in room acoustic environments: an overview. *EURASIP Journal on Advances in Signal Processing* 2006, 026503 (2006) 1–19.
50. D.A. Bohn: Environmental effects on the speed of sound. *Journal of the Audio Engineering Society* 36, 4 (1988) 223–231.
51. C. Antoine: *Thermodynamique – Tensions des vapeurs. Nouvelle relation entre les tensions et les températures*. *Comptes Rendus des Séances de l’Académie des Sciences* 107 (1888) 681–684.
52. M.S. Brandstein, H.F. Silverman: A robust method for speech signal time-delay estimation in reverberant rooms. *IEEE International Conference on Acoustics, Speech, and Signal Processing* 1 (1997) 375–378.
53. L. Cheng, C. Wu, Y. Zhang, H. Wu, M. Li, C. Maple: A survey of localization in wireless sensor network. *International Journal of Distributed Sensor Networks* 2012 (2012) 1–12.
54. J.M. Fresno, G. Robles, J. Martínez-Tarifa, B.G. Stewart: *Survey on the Performance of Source Localization Algorithms*. *Sensors (Basel, Switzerland)* 17, 11 (2017) 2666.
55. M. Cobos, F. Antonacci, A. Alexandridis, A. Mouchtaris, B. Lee: A survey of sound source localization methods in wireless acoustic sensor networks. *Wireless Communications and Mobile Computing* 2017 (2017) 1–24.
56. R.E. Kalman: *A new approach to linear filtering and prediction problems*. *Journal of Basic Engineering* 82, 1 (1960) 35–45.
57. D. Simon: *Optimal State Estimation*. Wiley-Interscience, 2006.
58. T.F. Brooks, W.M. Humphreys: Effect of directional array size on the measurement of airframe noise components, in: 5th AIAA/CEAS Aeroacoustics Conference and Exhibit. AIAA, 1999.
59. P. Sijtsma: Phased array beamforming applied to wind tunnel and fly-over tests, in: SAE Brasil International Noise and Vibration Congress, October 2010.
60. R. Merino-Martínez, E. Neri, M. Snellen, J. Kennedy, D. Simons, G.J. Bennett: Comparing flyover noise measurements to full-scale nose landing gear wind tunnel experiments for regional aircraft, in: 23rd AIAA/CEAS Aeroacoustics Conference, AIAA 2017-3006, 2017.

Cite this article as: Gombots S. Kaltenbacher M. & Kaltenbacher B. 2021. Capabilities of inverse scheme for acoustic source localization at low frequencies. *Acta Acustica*, 5, 44.