Adjoint-based identification of moving sound sources considering rigid walls

Mario Sroka[1,2], Ennes Sarradj[2], Mathias Lemke[1]
[1] Numerical Fluid Dynamics, TU Berlin, 10623 Berlin, Email: mario.sroka@tnt.tu-berlin.de
[2] Engineering Acoustics, TU Berlin, 10587 Berlin

Introduction

Sound source identification is a common challenge in acoustics, where the goal is to determine various parameters of one or more sound sources, such as their signals and positions. Microphone arrays, using beamforming techniques, are typically used for this purpose. However, identifying moving sound sources - such as airplanes, cars, trains, or drones - poses a particular challenge. This is especially true when these sources are within the array's detection range for only a short time. Recent research has introduced an adjoint-based method that uses computational aeroacoustics to identify and characterize free-moving sound sources in the time domain [1, 3]. Recent developments have extended this method to include the effects of sound-reflecting rigid walls in the computational domain.

Adjoint Approach

A Brinkman penalization approach is introduced to the Euler equations [2]. An embedded geometry is defined with a mask function ϕ

$$\phi(x_i) = \begin{cases} 1 & \text{free flow} \\ \ll 1 & \text{inside walls.} \end{cases} \tag{1}$$

Each term in the Euler equations is modified, i.e., for the time derivatives, ϕ is a prefactor, and for the space derivatives, ϕ is an extra factor in the derivatives. An additional Darcy term in the momentum equation dampens numerical instabilities within the rigid walls. The derivation of the adjoint equations is analogous to the case without volume penalization.
First, the scalar objective function J is defined in space and time as

$$J = \frac{1}{2} \int \underbrace{\left(q - q^{\text{target}}\right)}_{g}{}^2 \mathrm{d}V \mathrm{d}t. \tag{2}$$

It describes the difference g of the state q (sound field with an unknown source distribution s) and a target state q^{target} (e.g., microphone measurements). The state q is the solution of the Euler equations $Eq = s$. Minimization of the objective function J by fitting the source distribution s yields the optimal source distribution and the corresponding signal, thus, identifying the sound source from the measurements. Using a Lagrangian approach with the Euler equations as constraints, as well as the Lagrange multiplier q^*, the change in the objective function can be written as

$$\delta J = {q^*}^{\mathrm{T}} \delta s + \delta q^{\mathrm{T}} \left(g - E_{\text{lin}}^{\mathrm{T}} q^*\right). \tag{3}$$

The solution of the adjoint equation $E_{\text{lin}}^{\mathrm{T}} q^* = g$ yields the adjoint state vector. The gradient of J with respect to the source terms s can then be interpreted as $\nabla_s J = q^*$. It is shown that it is not necessary to achieve the optimal source distribution s and, thus, the minimization of the J in order to identify sound sources, but that determining the sensitivity is sufficient for many applications.

Validation

Before applying volume penalization to sound source identification, the method is validated by solving the adjoint equations with a test case. A three-dimensional volume is defined, where a sound source and a corresponding mirror source are located, see Fig. 1. The initial condition is a simple Gaussian pulse propagating spherically in the computational domain. After a particular time, the pulses from the source and the mirror source have penetrated each other, as shown on the left in Fig. 2.
If the lower part of the volume (the mirror source) is reflected by a rigid wall, the initial pulse is reflected at this wall. The error can be evaluated by comparing the case with a wall and the reference solution using the mirror source. In Fig. 2 on the right, the resulting reflection at the wall is shown at the same time step.

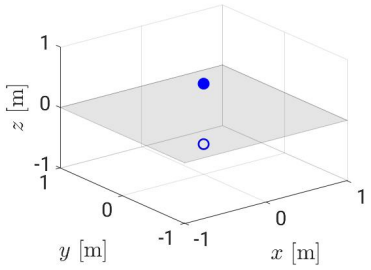

Figure 1: Validation case arrangement: • sound source (Gaussian pulse), ○ mirror pulse, rigid wall at $z \leq 0$ m.

The mask function ϕ can be described with the two parameters: width of the step ϕ_{dx} and height of the step ϕ_h. Theoretically, a narrow step that is as high as possible has the smallest reflection error. However, numerical stability (stiffness of the equation system) limits the choice of these two parameters. Fig. 3 shows the error, defined as the Euclidean norm of the deviation from the reference solution for different settings.
Except for the case of a very small step (ϕ_h = 1e-1), the error decreases continuously for smaller ϕ_{dx}. The height of the step and the grid resolution have a minor influence on the error. It is important to note that the error evaluates (sums) both amplitude deviations and phase

shifts together. The order of convergence is approximately $\mathcal{O}(\triangle x^{2.27})$, which corresponds to the error decay for the direct Euler equations [4]. In addition to the results in Fig. 3, further error analyses confirm that the volume penalization in the adjoint equations leads to a physically plausible acoustic behavior at the embedded walls.

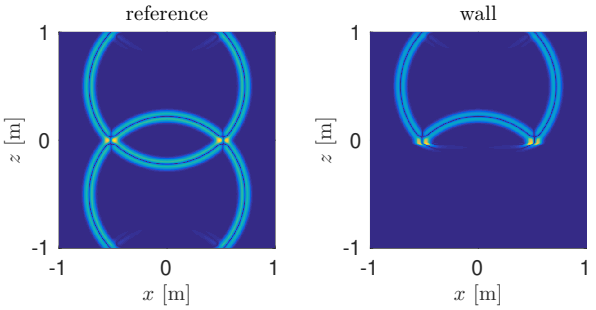

Figure 2: Maps of the adjoint pressure p^*, (left) reference case with mirror sound source, (right) rigid wall instead of mirror source, with $\phi_{dx} = 1.5$ dx and $\phi_h =$ 1e-4.

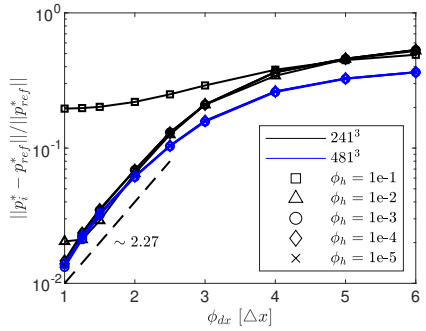

Figure 3: Error between reference solution and rigid wall cases for different parameters ϕ_{dx} and ϕ_h.

Results

Hidden sound source

As the first application, a hidden sound source is identified. The arrangement shown in Fig. 4 is analyzed as an example. The domain is three-dimensional and consists of two connected non-symmetrical volumes. In the upper volume is a spiral array consisting of 64 microphones with a diameter of $\approx$ 1.43 m. In the lower volume is a static sound source that emits white noise. The position of the source is gradually shifted towards a wall so that for the case $x = -0.8$ m, there is no direct line of sight between the source and the microphones. This means that the sound of the source reaches the microphones only by diffraction at the edge to the upper volume and by reflection at the walls. The wall parameters are $\phi_h =$ 1e-4 and $\phi_{dx} = 1.25dx$. The direct Euler equations are solved to generate synthetic measurement data. These are then used as sources in the adjoint equations.
The sensitivity $\bar{p}$ is used in the adjoint framework to localize the sound sources. $\bar{p}$ is calculated according to Eq. 4 from the adjoint pressure p^*. The adjoint pressure field consists locally of fluctuations whose temporal change is integrated over the entire computed time. Local maxima indicate the most likely locations for a sound source to cause the observation (microphone data).

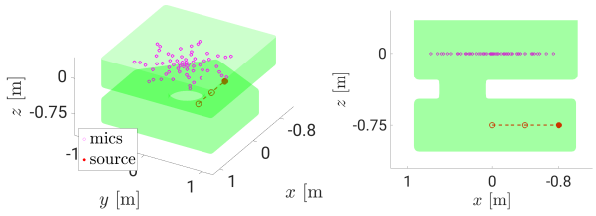

Figure 4: Hidden sound source arrangement: ●, ○ sound source (white noise), ○ microphone array, (acoustical) free field $\phi = 1$.

$$\bar{p} = \int_t \partial_t \left(p^*\left(x, y, z, t\right)\right) \mathrm{d}t \tag{4}$$

Fig. 5 shows the sensitivity of $\bar{p}$ in the plane of the source ($z = -0.75$ m). On the left is the case where the source is placed directly below the connection between the volumes. The white dashed line indicates this cylindrical connection. Note that the junction starts at $z = -0.5$ m. In the second case on the right side, the source is close to the lower wall at $x = -0.8$ m. The wall is located at $x = -0.9$ m. The color scale in both plots is logarithmic with a range of 6 dB.
In principle, the source position is found in all cases.

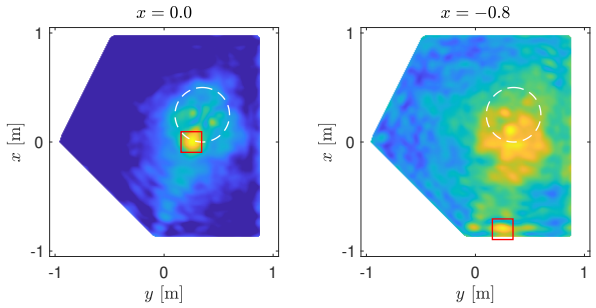

Figure 5: Maps of the normalized adjoint-based sensitivity $\bar{p}$ at $z = -0.75$ m, white dashed line: projected connection between both volumes, red square: correct sound source position, logarithmic color scaling with a range of 6 dB, yellow $\bar{p} = 1$.

$\bar{p}$ has a local maximum at the locations of the sound sources (red colored square in Fig. 4). The source is clearly identifiable in the left map because the sensitivity at the source location is several dB higher than in the rest of the domain. However, the source is directly observable for some microphones in this case.
If the source is placed closer to the wall and is therefore no longer directly visible, the difference in sensitivity between the source and its surroundings is significantly smaller. In Fig. 4 on the right side, the areas around the source only have a deviation of about $1 - 3$ dB. It can also be seen that other areas of the domain also have local sensitivity maxima. Without information

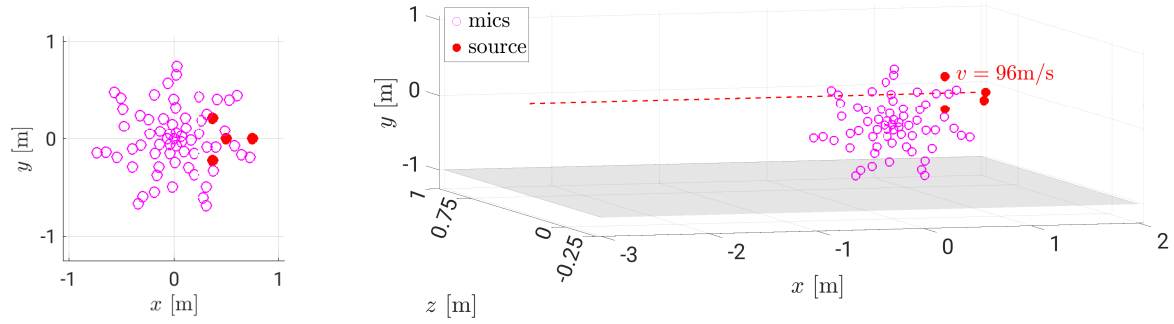

Figure 6: Moving sound sources arrangement: • moving sound sources (white noise) with velocity $v = 96$ m/s, ○ microphone array, gray plane: rigid wall (ground).

about the exact location of the source, it would be challenging to identify the sound source in this case. In a practical application, this would mean that these results would have to be discussed further. Nevertheless, this observation aligns with expectations since solving the adjoint equations cannot yield a unique source distribution. The inverse acoustic problem is not unique. From a practical point of view, a possible discussion of the source distribution for this case may not be complicated. Because there should always be some information about the sources available, purely from observation during an experiment. Alternatively, an entire optimization loop can be computed in the adjoint framework. The sources should then be identified much more precisely. For this purpose, the objective function J in Eq. 2 must be defined and assumptions have to be made about the location of the source.

Moving sound sources above a rigid wall

In a second application, four moving sound sources are examined. As shown in Fig. 6, four tetrahedrally arranged sources are moved at a high velocity of $v = 96$ m/s over a fixed wall (ground). The sources move between $0.55 - 0.75$ m in front of a spiral microphone array. The array is identical to the hidden source case. The velocity of the sources is explicitly set very high to test the limits of the adjoint method. The sources emit white noise and have an amplitude difference. The amplitude level is set to $[L_0, L_0 - 3\text{ dB}, L_0 - 6\text{ dB}, L_0 - 3\text{ dB}]$. The Euler equations are solved for 2000 time steps with $\triangle t = 2.08\text{e-}5$ s. Thus, the sources are only 731 time steps directly in front of the array.

Fig. 7 shows the sensitivity $\bar{p}$ for $z = 0.75$ in the plane with three (of four) sound sources. $\bar{p}$ is calculated accordingly to Eq. 4 and therefore includes the motion of the sources. This is shown by the fact that the sources are not clearly visible in the map in Fig. 7. However, there is an area of maximum sensitivity on and near the trajectories, shown as a red dotted line. The region of maximum sensitivity is roughly divided into three parts. As discussed in [3], the trajectory can generally be found using the adjoint method. However, in the map in Fig. 7 this is not observable. This can be explained by the high velocity and the small number of samples

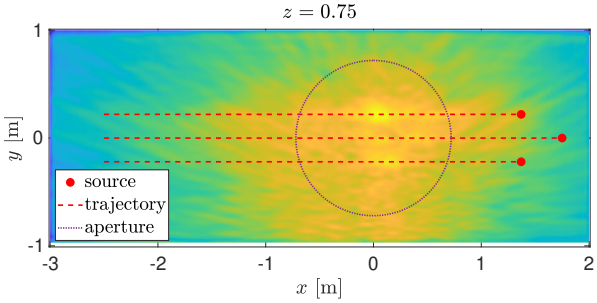

Figure 7: Maps of the normalized adjoint-based sensitivity $\bar{p}$ at $z = 0.75$ m, • sound sources, dashed line trajectory, dotted line projected microphone array, logarithmic color scaling with a range of 6 dB, yellow $\bar{p} = 1$.

directly in front of the array. At lower speeds, the three trajectories would be visible as separated lines.
Another observation in Fig. 7 is that the sensitivity decreases significantly in the direct vicinity of the array. At a distance of 0.5 m, the amplitude of the sensitivity is already less than 3 dB. This shows the directional characteristic of the array. A planar array forms a main lobe directly perpendicular to the array. This is confirmed by the $\bar{p}$ map, thus validating the numerical investigation additionally.
Also discussed in [3] was how to calculate the sensitivity field from the computational data, in which the motion of the sources is taken into account. This is done by a suitable back-transformation of the p^* field into a fixed reference frame. If the actual motion of the sources (trajectory and velocity) is known, then the transformation can be performed directly. If little or no knowledge of the motion exists, then appropriate information can be obtained from the results of the adjoint solution. This is discussed further in [5].

The transformed sensitivity $\bar{p}$ at $z = 0.75$ m is shown in the left plot in Fig. 8. All three sources in this plane can be identified at the correct positions. The local $\bar{p}$ maxima can be clearly distinguished from the surrounding and other domain parts. However, besides the three sources, there is also a locally increased sensitivity at the projected position of the fourth source (magenta cross). This can be explained by the three-dimensional structure of the sensitivity, which is shown in the right

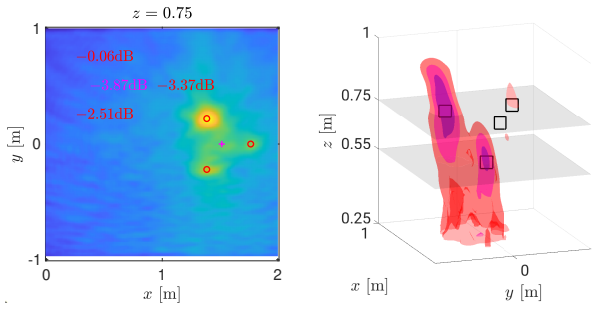

Figure 8: (left) Maps of the normalized adjoint-based sensitivity $\bar{p}$ at $z = 0.75$ m, ∘, + sound sources, logarithmic color scaling with a range of 6 dB, yellow $\bar{p} = 1$. (right) Iso-Surfaces of the normalized adjoint-based sensitivity $\bar{p}$, surfaces at $[-2, -1, -0.5]$ dB, (red, magenta, blue), squares: correct sound source position.

plot of Fig. 8.
As mentioned, the array is planar. This means, in particular, that the exact positions of a sound source perpendicular to the array are more difficult to detect because the distance to the array cannot be clearly determined. This can be clearly seen in the right plot in Fig. 8. This plot shows three-dimensional surfaces of equal sensitivity. These isosurfaces form lobe-shaped structures perpendicular to the microphone array. These surfaces enclose the sources. I.e., if the exact distance of the source to the array is unknown, the source's position can be approximated using the adjoint method. For example, in Fig. 8 (right), this would be the blue-magenta colored areas. Two of the four sources can be directly identified in this plot. The squares mark the exact positions of the sources. The two gray planes show the planes in which the sources are located. A third source is also identifiable in the isosurface plot in Fig. 8. The fourth is not recognizable. The reason for this is the choice of the iso values. Due to the amplitude differences, finding suitable values to identify all four sources simultaneously is difficult.
The amplitude differences in the sensitivity $\bar{p}$ are listed in the plot on the left side in Fig. 8. It can be seen that the differences are not precisely determined, in contrast to the test case in [3]. This observation needs to be further investigated in future work to understand how far the velocity of the sources influences this. However, it can be seen that, in principle, there are differences in the $\bar{p}$ amplitude, and the order of strength of the individual sources is found correctly.

Conclusion

An adjoint method for sound source localization was extended with a volume penalization approach. This approach was validated, and in particular, the error for reflection of an adjoint sound wave was examined. The order of convergence is approximately $\mathcal{O}(\triangle x^{2.27})$ and thus agrees well with the results for the direct Euler equations. It can be concluded that acoustic effects such as reflection and diffraction at the embedded geometry work properly in the adjoint framework.
The adjoint-based sound source identification can find hidden sound sources. This was demonstrated in an example case where a static sound source was found even without a direct line of sight to the microphones. However, it can be observed that the sensitivity is not a unique solution in the entire computational domain. This is to be expected and can be explained by the non-uniqueness of the inverse acoustic problem. The source position would have to be further discussed in a practical application. However, the immediate vicinity of the statistical source still shows a local sensitivity maximum.
A second test case showed that the adjoint method can also identify moving sound sources while considering rigid walls. It was shown that the determination of the trajectory at high speed and in a short time directly in front of the microphone array is not as clear as in previous work. In principle, however, the results are still analogous to previous observations [3]. In particular, all sources are correctly found after the transformation of the adjoint pressure into a fixed reference frame. The three-dimensional structure of the sensitivity is lobe-like and perpendicular to the array. Again, the position of the sources can be estimated without further knowledge of their exact position, which is particularly interesting for practical applications.

Acknowledgments

The research was funded by the German Research Foundation - project number 504470212.

References

[1] Lemke, M.: Adjoint based data assimilation in compressible flows with application to pressure determination from PIV data. PhD thesis, TU Berlin, (2015)

[2] Lemke, M., Reiss, J.: Approximate acoustic boundary conditions in the time-domain using volume penalization. The Journal of the Acoustical Society of America (2023), 153(2), 1219-1228

[3] Sroka, M., et al.: Adjungierten-basierte Identifikation sich frei bewegender Schallquellen. Fortschritte der Akustik: Tagungsband d. 50. DAGA (2024), 1065-1068

[4] Schubert, Y., Lemke, M.: Error Analysis of Boundary Conditions Using Volume Penalization in Acoustic Simulations. DAGA (2025), [Manuscript in preparation]

[5] Sroka, M., et al.: Moving source detection using microphone array measurements based on Euler equations. [Manuscript submitted for publication]