THREE-DIMENSIONAL GRIDLESS SOURCE MAPPING USING A SIGNAL SUBSPACE APPROACH

Ennes Sarradj

Technische Universität Berlin, FG Technische Akustik
Einsteinufer 25, 10587 Berlin, Germany

Abstract

If a microphone array is used for acoustic source mapping and characterization, sources are of mapped to a plane. However, in many cases even the major sources have a three-dimensional spatial distribution and thus a three-dimensional mapping appears to be appropriate. Direct beamforming methods fail in three-dimensional mapping because of the limited depth-wise resolution, so that deconvolution or inverse methods have to be used. Most of these are grid-based and computationally expensive. This means that three-dimensional mapping requiring a grid with many points is computationally very demanding if feasible at all. This contribution addresses the problem by introducing a variant of the signal subspace approach (sometimes coined: 'orthogonal beamforming') that does not need a grid at all. In this method, for each eigenpair from the signal subspace, the maximum in the three-dimensional map is sought using an efficient simplical homology global optimization method that typically needs less than 10^3 function values instead of the more than 10^5 available in a typical three-dimensional grid. The theory and the limitations of the approach are explained. The method is compared to other gridless methods. Finally the application is demonstrated using both synthetic and measured data. The results show that performance of the method variant using a grid is also met with the gridless approach. In certain situations, the gridless approach yield more precise localization results.

1 INTRODUCTION

Many practical applications of microphone array based acoustic source mapping and characterization concern applications where the spatial distribution of sources is inherently three-dimensional. Classical beamforming cannot consider this properly, because it has a generally a very poor depth-wise spatial resolution. Therefore, it must resort to two-dimensional mapping. Other methods can resolve this problem. This was demonstrated for a number of different methods and applications. One example is aeroacoustic noise from a pantograph that was mapped

using Clean-SC [2]. Other examples are Clean-SC, DAMAS, General Inverse Beamforming and SC-DAMAS applied to distributions of point sources [9], two IRLS based methods applied to an airfoil in an open jet [1]. Recently Clean-SC was also applied to three-dimensional underwater acoustics [10]. Many more examples can be found in literature.

Nearly all methods used for source mapping use a search grid as a basis for estimating the position of sources. The grid contains all N_G possible source positions with a region of interest (ROI). The distance between possible source positions limits the spatial resolution of the source map. If the ROI is three-dimensional, much more such positions are of interest and therefore N_G becomes considerably larger as in the case of a two-dimensional ROI. Thus, instead of $N = 10^{3...4}$ as in the two-dimensional case, $N = 10^{5...6}$ must be considered. This fact has a severe impact on the computational effort. Even in the optimistic case that the necessary number of computing operations scale only linearly with N_G, long computation times are needed. Depending on the method, very large N_G can lead to a computational effort that is no longer acceptable, see the detailed analysis given in [4].

Actual sources are often estimated only at very few positions in the grid. In other words, the spatial distribution of sources is sparse. Therefore it is obvious that one could think of an approach that does not rely on a grid at all, but estimates the position of the source in a different way. One example for such an approach is [8], where a global optimization problem for both source positions and strengths is solved using genetic programming. It turns out that this works well if the global optimization problem is not too difficult to solve, which is no longer the case at shorter wavelengths and higher frequencies. Another example is [7], where a convolutional neural network is applied to directly estimate both position and source strength. However, the network needs extensive training and cannot readily be applied.

Some of the more classical grid-based methods can in principle also be reformulated into optimization problems where the solutions gives both the coordinates and strength of sources. In present contribution this exercise shall be performed for the method proposed in [11], sometimes coined "orthogonal deconvolution". In what follows, first a short review of the grid-based version of that method is given, then the gridless reformulation is proposed and the solution strategy using the simplical homology global optimization method is explained. Finally the application is demonstrated for synthetic data case as well as measured data from an experiment in an aeroacoustic windtunnel.

2 SIGNAL SUBSPACE METHOD

The method is based on the following model of the sound field. Consider a sound source that cause a complex-valued sound pressure q at some reference location $\boldsymbol{x}_0$. Then, the sound pressure at at the i-th microphone of an array of N microphones is given by

$$p(\boldsymbol{x}_i) = a(\boldsymbol{x}_i, \boldsymbol{x}_0, \boldsymbol{x}_s) q(\boldsymbol{x}_s). \tag{1}$$

where $\boldsymbol{x}_s$ is the location of the source. The transfer function $a(\boldsymbol{x}_i, \boldsymbol{x}_0, \boldsymbol{x}_s)$ depends on the acoustic environment and may also consider the presence of flow. If M sources are present, then their contributions superpose at each microphone. The vector containing the sound pressures at all

microphones is then

$$p = Aq + n, \tag{2}$$

with the matrix A containing all transfer functions between source and microphone positions and n adds possible contributions from noise or non-acoustic pressure fluctuations.

Assuming that the noise signals at all microphones are of equal power n^2 and mutually uncorrelated, the cross spectral matrix of the microphone signals is given by

$$G = ASA^H + n^2 I. \tag{3}$$

If the analysis is restricted to uncorrelated source signals, the source cross spectral matrix S is a diagonal matrix whose elements correspond to the power of the source signals measured at x_0 and representing the contribution of individual sources.

Conventional beamforming seeks to estimate this by filtering out the apparent contribution from a possible source at location x_t within a grid of N_G possible source positions. This is done by multiplying the cross spectral matrix of microphone signal with a vector of filter coefficients:

$$B(x_t) = h^H(x_t) G h(x_t). \tag{4}$$

The steering vector h depends on the vector a of all transfer functions that relate x_t to any of the microphones. The result for the case that x_t from the grid coincides with a source position is either a local maximum or it equals the contribution of this source. Which of both is applicable depends on the chosen formulation of the steering vector [3, 12]. A source map on the grid can be constructed by evaluating $B(x_t)$ for all N_G locations x_t in the grid.

The cross spectral matrix is positive semidefinite. Its eigendecomposition yields

$$G = VLV^H. \tag{5}$$

with the diagonal matrix L containing the real-valued and positive eigenvalues and V the eigenvectors. The eigenvalues can be split into the M largest, with the corresponding eigenvectors spanning the signal subspace and the $N - M$ smallest, with the corresponding eigenvectors spanning the noise subspace.

It can be shown [11] that each of the eigenvalues Λ_{Sii} from the signal subspace approximates the contribution from one of the sources. Then, using

$$B_i(x_t) = h(x_t)^H G_i h(x_t). \tag{6}$$

a map for just source i can be constructed when $G_i = v_i \Lambda_{Sii} v_i^H$ is the rank-1 cross spectral matrix reconstructed using just one eigenvalue / eigenvector pair. Note the (6) can be efficiently evaluated by computing just one vector - vector product per grid point. If an appropriate formulation for the steering vector is used, e.g. formulation IV from [12], then the maximum value in the map corresponds to the source position. In principle, this position can be found with high accuracy if the spatial resolution of the grid is high. This corresponds to large number of grid points N_G, especially in the three-dimensional case.

To construct a global map of all sources, a map is constructed for each i from the signal subspace. Then the x_t for the maximum $B(x_t)$ in map i is taken as the location of source i, and

Λ_{Sii} as its strength. Thus, the global map is sparse and contains only as many nonzero elements as there are sources.

3 GRIDLESS FORMULATION

The maps for individual sources need only to be evaluated the find the maximum. No further use is made of these maps. This leads to the conclusion that while the maps themselves are not needed, it could be more efficient not to compute them at all, but just finding the location of global maximum

$$\underset{\boldsymbol{x}_t \in \mathbb{R}^3 \cap \boldsymbol{x}_t \in ROI}{\arg\max} \quad B_i(\boldsymbol{x}_t). \tag{7}$$

The constraints restrict the solution to the region of interest (ROI) where sources are expected. This is an optimization problem. Because $B_i(\boldsymbol{x}_t)$ may have many local maxima, a global optimizer is needed to solve this problem. Moreover, as the gradient of the objective function is not known, the optimizer cannot make use of it.

One possible algorithm to solve the problem is the simplical homology global optimizer [5]. If the problem is reformulated so that the sought $\boldsymbol{x}_t$ is the global minimum, this optimizer is known to find it under condition of Lipschitz smoothness that is fulfilled in the case of $B_i(\boldsymbol{x}_t)$. It works by sampling the ROI and constructing a simplical complex. From this, locally convex subdomains are identified and within these an appropriate local minimizer can be used to find the minima. Details on how the subdomains are identified using homology theory see [5].

For the initial sampling, a uniform distribution of sampling points is needed. Following [5], the low-discrepancy Sobol sequence [15] is applied. This gives a quasirandom distribution of 2^k points within the bounds of the ROI. Other sampling options turned out to be less efficient for the problem considered here and needed more evaluations of the objective function to arrive at a similar result. The Nelder-Mead algorithm [6] is chosen as local minimizer despite not handling constraints in the available implementation. This turned out to be no problem in the present case. Other local minimizers that handle constraints do so at the cost of additional evaluations of the objective function.

4 DATA

The gridless method is evaluated here using two different data sets. One data set is from a synthetic test case and the other from an experiment in an aeroacoustic wind tunnel. Both data sets are analyzed using the proposed method as well as the signal subspace method using a three-dimensional grid. This allows for the comparison of the results.

4.1 Three point sources

The synthetic test case simulates three point sources emitting white noise signals in front of a 64 microphone array. It resembles the test case from [12], see Fig. 1. The positions of the sources are chosen randomly within a mapping region in front of the array. All coordinates and lengths are nondimensionalized by the aperture d of the array. This means that also the frequency has to be nondimensionalized as a Helmholz number $He = fd/c$, where c is the speed of sound.

Source levels were chosen to be different with source A having the highest level, and source B and C having a level 10 dB and 22 dB less, respectively.

Simulation of the microphone signals was performed for 512000 samples using a sampling frequency of He = 150. The sampled microphone signal were partitioned into blocks of 1024 samples with an overlap of 50%. Following the application of a von Hann window and a fast Fourier transform, the cross spectral matrix was computed for frequency bins with a width of He $\approx$ 0.145 as an average of 999 blocks.

The data was processed using both the signal subspace method with a grid and the gridless variant proposed here. The grid is covering the mapping region of $1.25 \times 1.25 \times 1.0$ with a resolution of 0.02, thus having a total of 208896 grid points. For the gridless approach, the ROI was set to the same extent as the grid. The simplical homology global optimizer was provided with 256 initial sampling points and needed approximately 400 overall function evaluations at each frequency bin and eigenvalue considered to find the maximum.
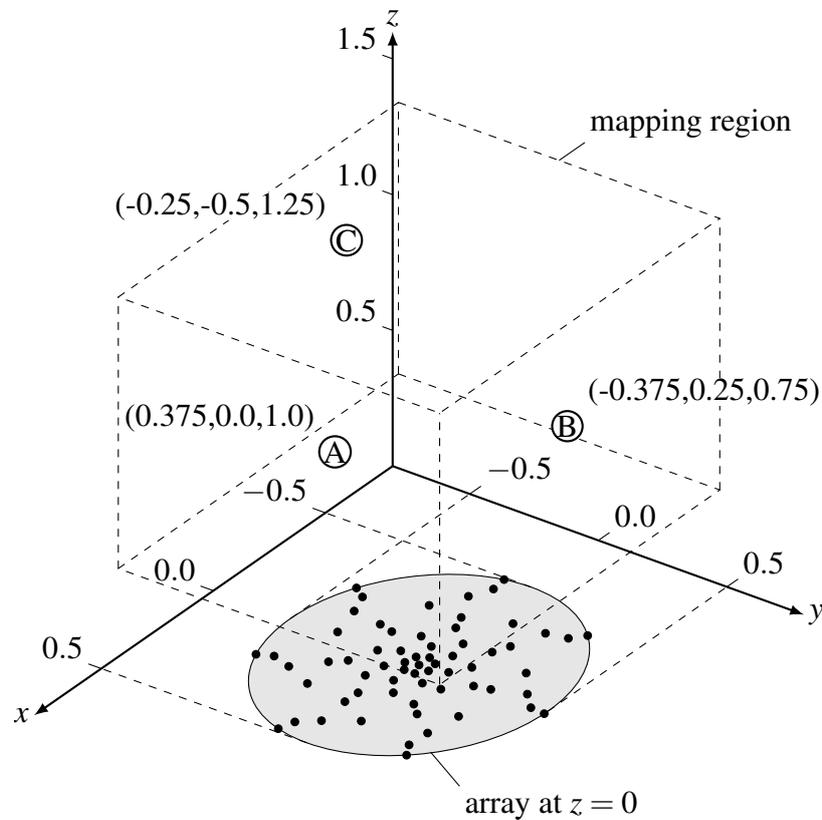
Figure 1: Scene with three point sources and 64 microphone array (adapted from [12]). The mapping region dimensions are 1.25 × 1.25 × 1.0 apertures.

4.2 Aeroacoustic experiment

The experiment concerns a NACA0012 airfoil with a 0.2 m chord and a span of 0.4 m in a open jet with 0.2 diameter and a speed of 86.9 m/s, which was already used as a test case in [13]. Sound is generated from the interaction of the jet shear layer and the airfoil as well as at the
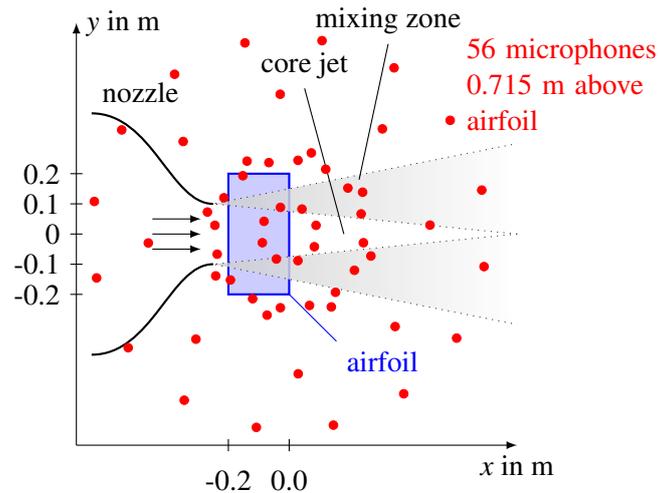
Figure 2: Schematic of setup with NACA0012 airfoil in windtunnel

trailing edge where the turbulent boundary layers from the flow over both sides over the airfoils interact.

A 56 microphone array with an aperture of 1.3 m was mounted in a plane in 0.715 m distance from the airfoil plane. The airfoil leading edge was 0.05 m downstream of the nozzle. Microphone signals were sampled at 51200 Hz for a duration of 40 s and the cross spectral matrix was computed from 50% overlapping blocks of 1024 samples each and using the von Hann window. The steering vectors were computed considering the refraction in the thick shear layer using a ray casting method [13].

Again, both gridded and gridless variants of the signal subspace method were used to process the data. The ROI was stretching from $x = -0.3 \ldots 0.1$ m, $y = -0.2 \ldots 0.2$ m and $z = 0.5 \ldots 0.9$ m. The grid (same extend as the ROI) had a total of 531441 grid points. The optimizer for the gridless method was supplied with the same parameters as in the three source case.

5 Results

For the three point source case, the number of sources implies that the largest three eigenvalues have to be used. That means for all frequency bins, both methods are mapping three sources. Fig. 3 shows the results for the estimation of sound pressure level contributions of the three sources from both methods. The estimation is done by considering any contribution mapped to within 0.02 from the true position of the source. For the gridded method this is equivalent to at least one grid point spacing. Generally, both methods produce similar results. For some frequency bins, no contribution is estimated as the sources are not mapped within 0.02, but further away. The possibility of somewhat wrong mapping is a property of the signal subspace method and the conditions for this are explained in [11] in more detail. It appears to happen more often for the gridded variant, which can be explained by the fact that the grid does not allow to map to arbitrary positions, but only to discrete positions given by the grid. Thus the precision of the estimated position is less than for the gridless method.
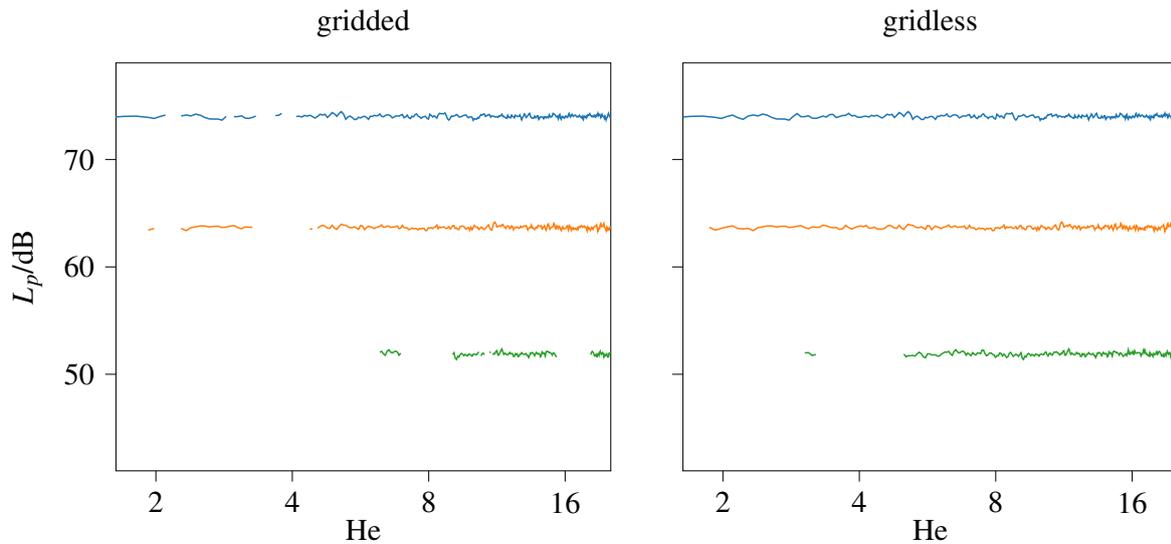
Figure 3: Sound pressure levels at the array from the three sources as a function of (dimensionless frequency): contributions from a sector with a radius of 0.02 around the true positions of the sources were considered, — source A, — source B, — source C

This can be seen in more detail in Fig. 4. The estimated positions of the sources deviate somewhat from the true position. This deviation is continuous for the gridless method, while the gridded method is bound to the grid resolution and discrete values. Apart from this, both methods show similar behavior. Other array methods deliver better results, especially at low frequencies, but require often more computational effort.

Both the gridded and the gridless method are implemented in Python as part of Acoular [14]. While the gridded method is optimized for a short run time, the implementation of the gridless method is less efficient. Still, the run time of the gridless method is considerably shorter (35 ms per frequency bin) than that for the gridless method (850 ms). These run times were achieved on a i7-10510U CPU at 1.80GHz. The comparison shows that the gridless method can be considerably faster while achieving similar results but with a better spatial resolution than the gridded variant.

For the aeroacoustic experiment, the number of sources is generally unknown. Thus, the number of eigenvalues to use has to be assessed beforehand. From the data it appeared that the first ten eigenvalues account for at least 95% of the energy. Thus, the ten eigenvalues were used for both methods. Fig. 5 shows the jet and the airfoil along with three-dimensional maps of the sound pressure level contribution at the array center for the 8 kHz octave band. It is apparent that the main sources are sitting at the trailing edge where the shear layer crosses. Another region with sources is the trailing within the core jet. Upon comparing the maps from the gridded and the gridless method it becomes obvious that in the gridded map the source positions are restricted to the grid points and the arrangement appears to be more regular. The sources might also seem to be stronger than in the gridless case. The reason for this is that in the gridded case, the contributions from the frequency bins within the octave band are accumulated at a grid point if source position is found at the same grid point for multiple frequency bins. This never happens for the gridless method because the estimated source positions do never coincide for
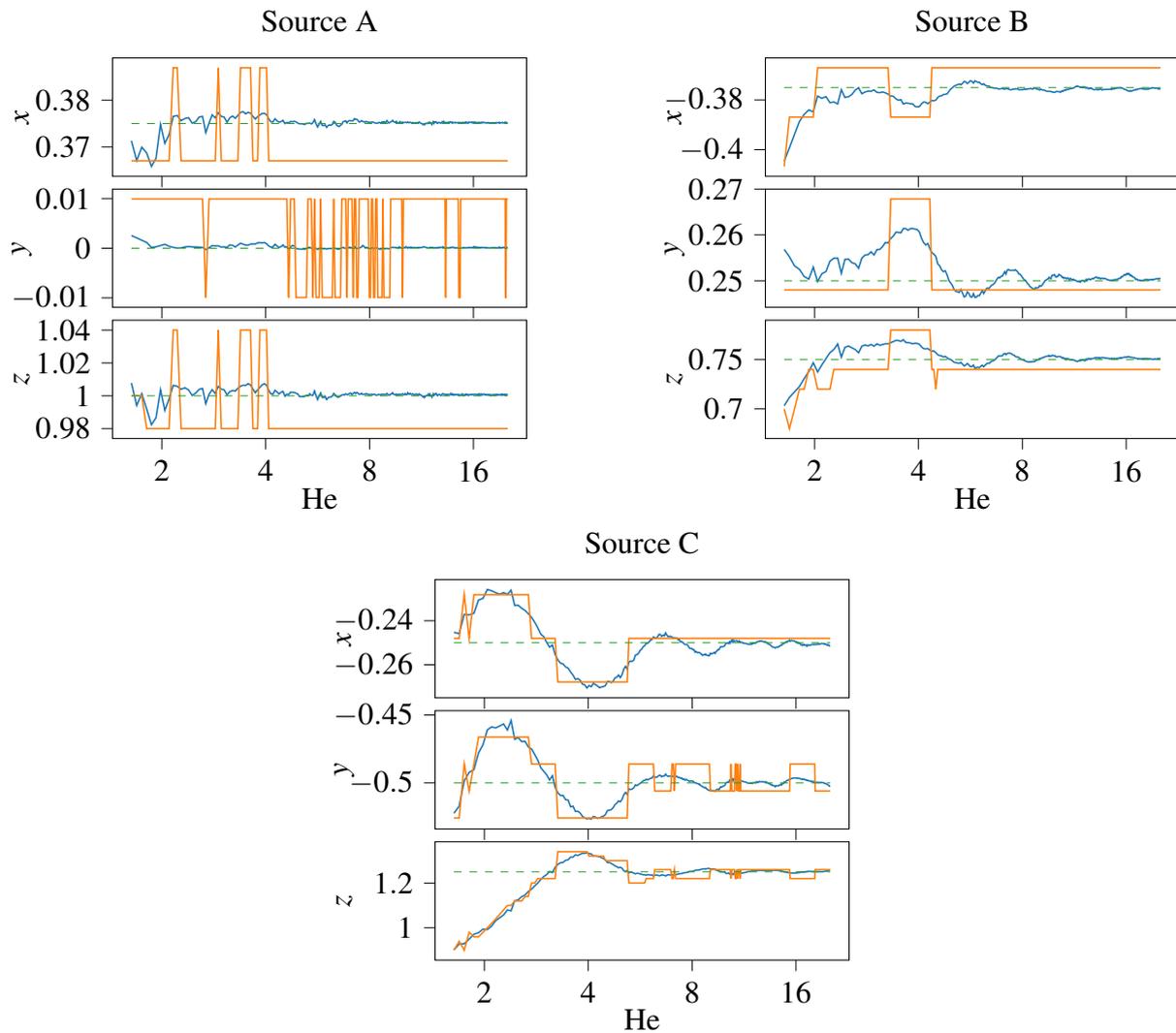
Figure 4: Estimated coordinates x, y, z for the position of the sources: — gridded method, — gridless method, - - true value

different frequency bands even if the distance between them may be extremely small.

To asses the results also quantitatively, Fig. 6 shows the sound pressure level contribution integrated over the trailing edge (see sector in Fig. 5). Both methods deliver practically the same result. This shows that the gridless method may be considered as a replacement for the gridded method. Despite the huge grid, the run time advantage of the gridless method was less than for the case of three sources (about five time faster than gridded). This is due to implementation of the shear layer correction that is considerably faster when a large number of steering vectors is computed at once. For the gridless method this computation has to be done separately and sequentially for each value that is computed.
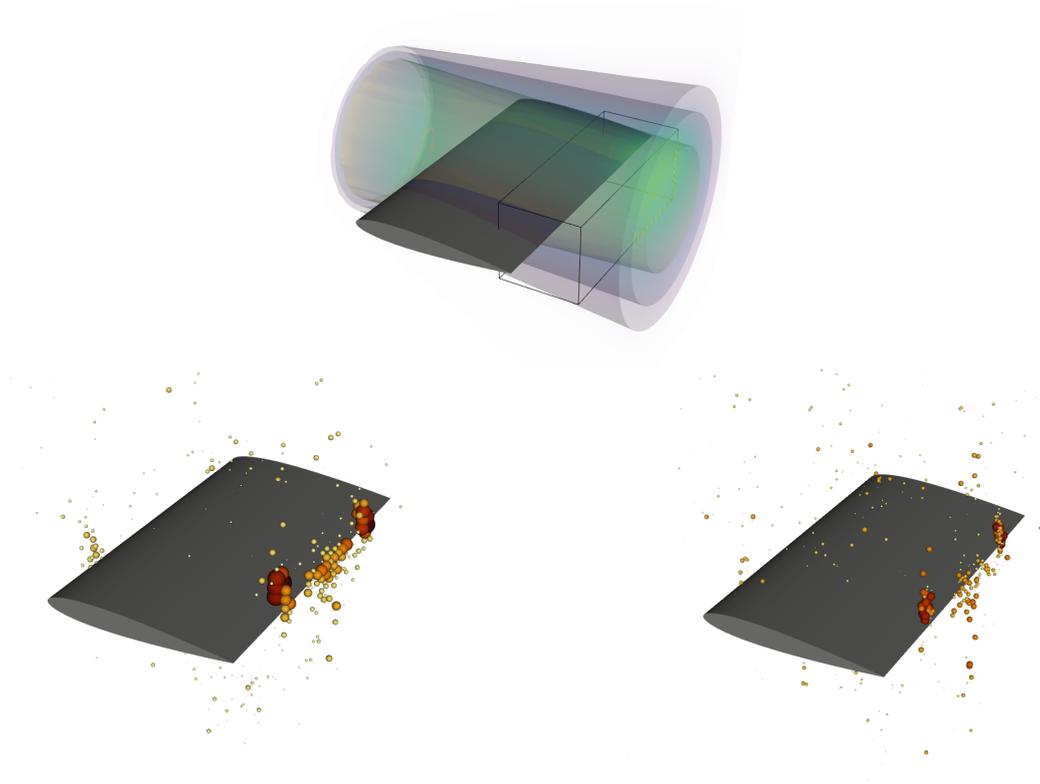
Figure 5: above: NACA 0012 airfoil and jet with iso-surfaces for the flow speed (10 m/s, 35 m/s, 80 m/s) and trailing edge integration sector, below: three-dimensional source mapping for the 8 kHz octave band, left: gridded, right: gridless, sphere color and size correspond to the sound pressure level contribution, range: 42...62 dB
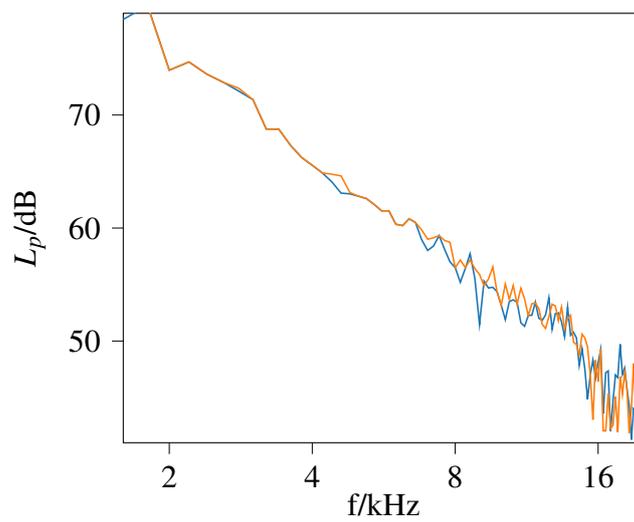
Figure 6: Sound pressure level contribution at array center integrated from the trailing edge sector shown in Fig. 5: — gridded method, — gridless method

6 CONCLUSION

Overall the results show that the gridless method works as good as the variant with grid. It has two main advantages. First, the estimated source positions have a high precision. With the gridded method, this can only be achieved use a grid with many densely spaced points and at the cost of higher computational effort. Second, a much smaller number of values of B_i need to be computed. This is the key for achieving short run times. However, as the implementation is quite cumbersome to optimize for speed, this advantage appears to be useful only if the gridded method would need to use a large or even huge grid with well over 10^4 grid points. This means that the use of the proposed gridless methods is predominantly advisable for three-dimensional case.

References

[1] G. Battista, G. Herold, E. Sarradj, P. Castellini, and P. Chiariotti. "IRLS based inverse methods tailored to volumetric acoustic source mapping." *Applied Acoustics*, 172, 107599, 2021.

[2] H. Brick, T. Kohrs, E. Sarradj, and T. Geyer. "Noise from high-speed trains: Experimental determination of the noise radiation of the pantograph." In *Forum Acusticum 2011, Aalborg*. 2011.

[3] G. Chardon. "Theoretical analysis of beamforming steering vector formulations for acoustic source localization." *Journal of Sound and Vibration*, 517, 116544, 2022. ISSN 0022-460X. doi:https://doi.org/10.1016/j.jsv.2021.116544.

[4] G. Chardon, J. Picheral, and F. Ollivier. "Theoretical analysis of the damas algorithm and efficient implementation of the covariance matrix fitting method for large-scale problems." *Journal of Sound and Vibration*, 508, 116208, 2021. ISSN 0022-460X. doi:https://doi.org/10.1016/j.jsv.2021.116208.

[5] S. C. Endres, C. Sandrock, and W. W. Focke. "A simplicial homology algorithm for lipschitz optimisation." *Journal of Global Optimization*, 72(2), 181–217, 2018.

[6] F. Gao and L. Han. "Implementing the nelder-mead simplex algorithm with adaptive parameters." *Computational Optimization and Applications*, 51(1), 259–277, 2012.

[7] A. Kujawski, G. Herold, and E. Sarradj. "A deep learning method for grid-free localization and quantification of sound sources." *The Journal of the Acoustical Society of America*, 146(3), EL225–EL231, 2019. ISSN 0001-4966. doi:10.1121/1.5126020.

[8] A. M. N. Malgoezar, M. Snellen, R. Merino-Martinez, D. G. Simons, and P. Sijtsma. "On the use of global optimization methods for acoustic source mapping." *The Journal of the Acoustical Society of America*, 141(1), 453–465, 2017. doi:10.1121/1.4973915.

[9] T. Padois and A. Berry. "Two and three-dimensional sound source localization with beamforming and several deconvolution techniques." *Acta Acustica united with Acustica*, 103(3), 392–400, 2017.

[10] M. Polichetti, V. Baron, J. I. Mars, and B. Nicolas. "Multiplane deconvolution in underwater acoustics: Simultaneous estimations of source level and position." *JASA Express Letters*, 1(7), 076001, 2021. doi:10.1121/10.0005513.

[11] E. Sarradj. "A fast signal subspace approach for the determination of absolute levels from phased microphone array measurements." *Journal of Sound and Vibration*, 329, 1553–1569, 2010.

[12] E. Sarradj. "Three-dimensional acoustic source mapping with different beamforming steering vector formulations." *Advances in Acoustics and Vibration*, 2012(292695), 1–12, 2012. doi:10.1155/2012/292695.

[13] E. Sarradj. "A fast ray casting method for sound refraction at shear layers." *International Journal of Aeroacoustics*, 16(1-2), 65–77, 2017.

[14] E. Sarradj and G. Herold. "A python framework for microphone array data processing." *Applied Acoustics*, 116, 50–58, 2017.

[15] I. M. Sobol'. "On the distribution of points in a cube and the approximate evaluation of integrals." *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki*, 7(4), 784–802, 1967.