forum acusticum 2026
8-12 September · Graz, Austria

# CEPSTRAL DECOMPOSITION OF MICROPHONE DIRECTIVITY IN ARRAYS FOR ACOUSTIC DIRECTION FINDING

Jakob Tschavoll[1], Ennes Sarradj[1]
[1]*Engineering Acoustics, TU Berlin, Germany*





## Abstract

Many acoustic direction-of-arrival (DoA) algorithms rely on digital beamforming with omnidirectional receivers, which requires evaluating dense full-sphere focus grids. Conventional methods often neglect the spatial filter characteristics inherent in passive acoustic gain patterns, which are standard in studio microphones. This work introduces a phaseless DoA estimation method for a full-sphere focus grid, using the comb-like filter characteristics of first-order directive receivers, here synthesized from omnidirectional endfire pairs. Cepstral decomposition informed by domain knowledge of the spatial filter yields an analytical relationship between the rotation-ambiguous angle of arrival and the filter parameters observed in the spectrum. For an array of orthogonally positioned receivers, superposing relative direction estimators can yield one or more sound source directions of comparable strength in 3D space. The method is first simulated for far-field noise sources and then experimentally validated under free-space conditions using real noise sources. Classification metrics and spatial error for source detection are presented, with comparisons across different array geometries. As the proposed method derives deterministic, phaseless cepstral scores from each receiver, it remains based on simple digital signal processing steps, making it suitable for autonomous low-power sensor networks in noise monitoring applications.

**Keywords:** *Acoustic DoA, Direction Finding, Cepstral Analysis, Spatial Filter*

***Corresponding author:*** *j.tschavoll@tu-berlin.de*


## 1. Introduction

Acoustic source detection for arbitrarily spaced sources, rather than a known focus grid, requires consideration of every possible angle. With omnidirectional microphones and classical beamforming algorithms, this makes full-sphere searches for one or more sparse sources inefficient, because most steering vectors point to source-free regions. Standard omnidirectional microphones offer flat magnitude and predictable phase responses for a rich ecosystem of beamformers, but they do not offload direction-dependent filtering to passive analog acoustics. Directivity was explored 100 years ago in both the RF [1], [2] and acoustic domains [3], the latter being used for studio and artistic applications and gaining a mature scientific understanding during the 20th century. In its simplest form, directivity can act as a received signal strength (RSS) maximizer over angle, which remains common in RF direction finding [4], [5], [6] and has also been used before digital beamforming in acoustics [7]. While this may appear archaic today, the DoA-related metric is already present in the unprocessed signal, meaning zero processing overhead. This motivates hybrid approaches, which use analog spatial filtering as a preprocessing step for digital extraction of DoA-related parameters [8], [9], [10]. However, RF applications have strict band limitations and usually operate in narrowband regions, which is not generally true for acoustic applications, where prior knowledge about source spectra is often unavailable. For this reason, RF hybrid approaches cannot be translated directly into broadband acoustic direction finding and remain reserved for specialized use cases [11], [12]. This work presents a method for detecting the DoA of one or more acoustic sources by deriving the spatial filter system from broadband noise excitation of first-order directive receivers, here realized by omnidirectional microphone pairs. To extract the angle-dependent, phaseless spectral structure of the filter, cepstral analysis is used. While cepstral analysis is established for identifying periodic spectral structures and has been used in acoustic spatial characterization before [13], [14], this work uses the resulting spacing as a deterministic parameter of a first-order directive receiver, yielding relative DoA estimates whose orthogonal angular mappings are added to obtain global source directions. A previous work explored the simulation results of this approach [15], and this work adds experimental validation.

## 2. Methods

The simplest 1[st]-order directivity, the dipole, can be modeled via two monopoles spaced a distance $d$ apart as a pressure-gradient or broadside receiver in the far field of a source by

$$p = \hat{p}e^{-\mathrm{j}\omega t} \mp \hat{p}e^{-\mathrm{j}\omega(t-\tau)}, \tag{1}$$

where the free-field path delay $\tau = \frac{d}{c}\cos(\theta)$ is the only cause for a "figure-8" pattern at the characteristic frequency $f_c = \frac{c}{2d}$. The resulting symmetric transfer function

$$G_{\Delta/\Sigma}(\theta, f) = 1 \mp e^{\mathrm{j}\omega\tau}, \theta \in [0, \pi] \tag{2}$$

produces side lobes for rising frequency in both cases. The spectrum therefore exhibits a comb-filter effect, with constant minimum spacing determined by $\theta$, reaching its smallest possible minimum frequency $f_0 = f_c$ when $\theta_\Delta \in \left\{\frac{\pi}{2}, 3\frac{\pi}{2}\right\}$ and $\theta_\Sigma \in \{0, \pi\}$. Since the dipole is symmetric in two axes, a unipolar directivity such as the cardioid is advantageous. With a fixed one-sided acoustic or digital path extension $l$, the angle $\theta \in [0, \pi]$ is measured from the side opposite to the extension, giving

$$G_{\Delta/\Sigma}(\theta, f) = 1 \mp e^{\mathrm{j}\omega\tau_l}, \tau_l = \frac{l + d\cos\theta}{c}. \tag{3}$$

For the cardioid case $l = d$, the added delay and geometric delay sum to $2\frac{d}{c}$ at $\theta = 0$ and compensate at $\theta = \pi$. The first minimum occurs at $f_c = \frac{c}{4d}$ and the inverse under $G_\Sigma \approx 0$ is

$$G^{-1}_{\Sigma,0}(f_0) = \theta_0(f_0) = \arccos\left(\frac{c}{2df_0} - 1\right), \tag{4}$$

where $f_0(\theta_0)$ is the frequency of the first occurring minimum in the spectrum, yielding a symmetric pair of relative source directions[1]. Due to the remaining ambiguity under ideal radial symmetry, each relative direction corresponds to a circle of latitude on the unit sphere, whose normal vector is parallel to the rotated null axis of the sensor.

To obtain the broadband filter excitation for $f_0$, the recorded signal of sensor $m$, $x_m[n]$, is split into $K$ snapshots with windowed spectra $X_{m,k}(f)$ to satisfy the Welch power spectral density estimation

$$\hat{P}_m(f) = \frac{1}{K}\sum_{k=0}^{K-1}\left|X_{m,k}(f)\right|^2. \tag{5}$$

The measured magnitude response is estimated as $M_m(f) = \sqrt{\hat{P}_m(f)} \propto G_\Sigma$, or as a normalized ratio when a reference channel is available. For multiple incoherent noise sources, energy sums up in the spectrum, leading to an addition of comb filters. This causes the minima to interfere, but not to disappear completely, as their periodic structure is still preserved for more than one source direction if $\theta_1, \theta_2 \in [0, \pi], \theta_1 \neq \pm\theta_2$. As the broadband excitation leads to periodic structures in the spectrum, a cepstral score can be used to obtain the periodic spacing of minima, which is constant across occurrences for each existing source direction. Since $G_\Sigma(\theta, f_0) \propto |\cos(\theta)|$, the minimum spacing follows $\Delta f = 2f_0$. For each sensor, the one-sided real cepstral score is computed from the measured magnitude response as

$$c_m(q) = |\mathcal{F}\{\log|M_m(f)|\}|, \tag{6}$$

where the null-quefrency $q_0 = \arg\max c_m(q)$ indicates the frequency spacing $\Delta f = \frac{1}{q_0}$ and therefore $f_0 = \frac{1}{2q_0}$. To find multiple sources, $q_0$ is not reduced to a single maximum. Instead, the full magnitude cepstrum is reparameterized from quefrency to first-minimum frequency by $f_0(q) = \frac{1}{2q}$ and then to relative angle by Eq. (4). This yields the relative-angle ordinate

$$\tilde{c}_m(\theta) = |c_m(q(\theta))|, \quad q(\theta) = \frac{1}{2f_0(\theta)}. \tag{7}$$

The measured one-dimensional score is then $C_m(\theta) = w(\theta)\tilde{c}_m(\theta)$, where $w(\theta)$ is the empirical reliability weighting shown in Figure 2. The weighting penalizes angles whose mapped minimum spacing produces only few spectral minima within the analyzed band, because the corresponding cepstral peaks are less reliable. The relative scores $C_m$ can now be used to derive the global DoA for known sensor axis orientations. Let $\boldsymbol{b}_m \in \mathbb{S}^2$ denote the known and fixed null-axis direction of sensor $m$ in the array coordinate frame. For a direction $\boldsymbol{r} \in \mathbb{S}^2$, the corresponding relative angle of sensor $m$ is

$$\theta_m(\boldsymbol{r}) = \arccos(\boldsymbol{b}_m^T\boldsymbol{r}). \tag{8}$$

This relation defines an extension of the one-dimensional score $C_m : [0, \pi] \to \mathbb{R}$ to a spherical score field $\tilde{C}_m : \mathbb{S}^2 \to \mathbb{R}$,

$$\tilde{C}_m(\boldsymbol{r}) = C_m(\theta_m(\boldsymbol{r})). \tag{9}$$

In a sampled elevation–azimuth map, the extension copies each relative-angle score value to all azimuthal positions on the corresponding circle around $\boldsymbol{b}_m$. The extended sensor fields are then summed as

$$\hat{C}(\boldsymbol{r}) = \sum_{m=1}^{M}\tilde{C}_m(\boldsymbol{r}). \tag{10}$$

The DoA estimate is obtained from local maxima of $\hat{C}(\boldsymbol{r})$. For simulation and measurement post-processing, this continuous field is evaluated on finite spherical grid points $\boldsymbol{r}_i \in \mathbb{S}^2$, giving $\hat{C}_i = \hat{C}(\boldsymbol{r}_i)$ before thresholding and clustering, see Figure 4.

[1]Here and subsequently, the additive transfer function $G_\Sigma$ is used for ease of demonstration, as it does not exhibit a high-pass effect. However, it requires active analog or digital summation, while the gradient can operate passively.

For testing, first-order receivers are synthesized as two-component endfire pairs of omnidirectional microphones; the model is equivalent to a passive first-order receiver with a one-sided path delay. Octahedral and tetrahedral arrays with $f_c = 1$ kHz are used, see Figure 1. Because fixed geometry limits the possible range of $\Delta f$, Eq. (6) is evaluated with a zoom FFT. Simulations use 3 s signals at $f_s = 48$ kHz and source positions on a 5000-point Fibonacci sphere of 100 m radius. Experiments use 10 s anechoic recordings at $f_s = 51.2$ kHz with white and pink noise. Welch PSD estimates use 4096-sample Hann windows, 50% overlap, and center-microphone source normalization when available. Source directions are obtained by thresholding the spherical score map and clustering the remaining unit-sphere points with DBSCAN and a haversine metric. Simulation metrics use a −1 dB global-maximum threshold, 8° clustering radius, and 10° elevation-and-azimuth matching tolerance; experimental plots use −1.5 dB and 20°.

## 3. Results

The simulated and measured spectra contain the expected direction-dependent minima. For the octahedron geometry, Figure 2 aligns spectral minima and cepstral maxima on the frequency axis. Eq. (4) can either compute a discrete $f_0$ or produce the cepstral score $C_m(\theta)$. Spatial resolution depends on frequency bins and therefore on the sampling rate. For a highlighted far-field source, the summed score $\hat{C}(\theta, \varphi)$ has its maximum at the likely source direction, as shown in Figure 4. With $f_s = 48$ kHz, Eq. (4) cannot produce angular values $\theta > 145°$; these values are therefore set to the neutral element 0.
The zoom FFT suppresses narrowband peaks and valleys, giving similar cepstral scores for tonal and

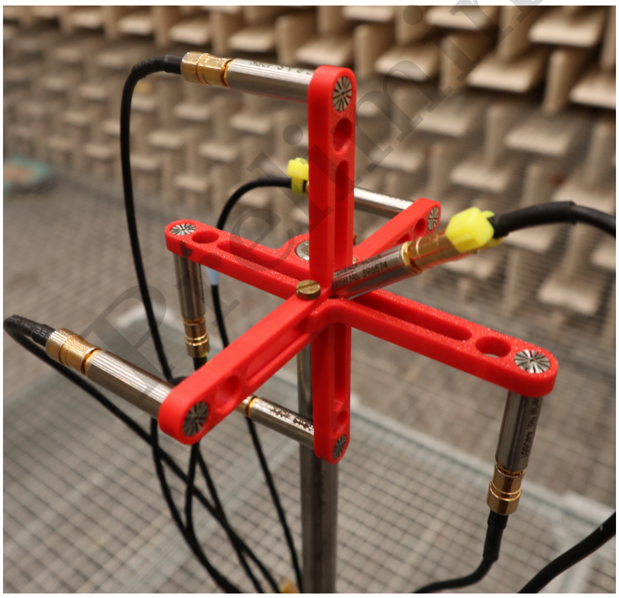

**Figure 1:** 3D-printed octahedral array with omnidirectional microphones. Opposing microphone pairs are processed as first-order endfire receivers; the center microphone is used as a reference.

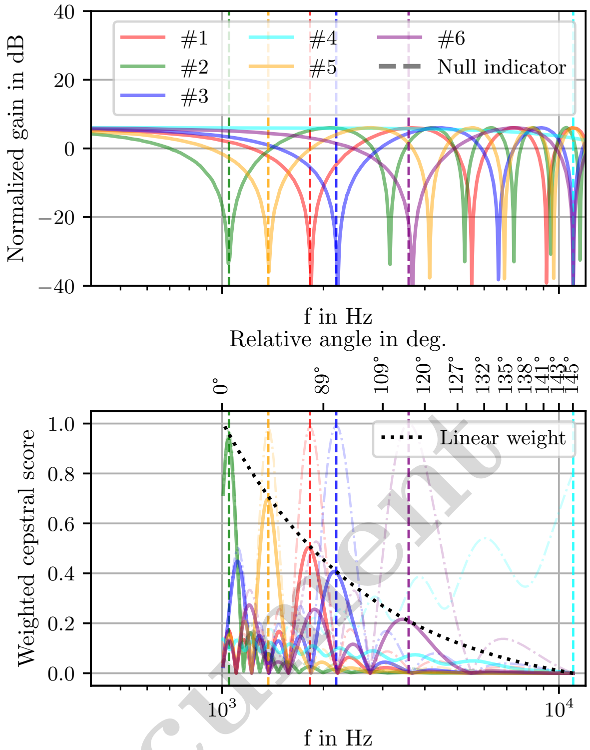


**Figure 2:** Simulated spectra and linearly weighted cepstral scores for six sensors in an octahedron configuration. Spectral minima align with cepstral maxima; semi-transparent curves show unweighted scores.

white-noise excitation from the same angle. Figure 3 compares an ideal white-noise source with a simulated drone signal. The drone signal uses a recording, so spatial radiation patterns are not retained; nevertheless, the cepstral scores remain similar despite different spectra.
Single-source simulations find the source reliably across the 5000-point Fibonacci sphere, with RMSE and classification metrics shown in Figure 5 and Table 1. Here and below, RMSE denotes the root-mean-square angular distance between true and detected directions. Table 1 also reports two-source simulations for two array geometries. In these scenarios, one source iterates in a spiral pattern as in the single-source case, and the other stays constant at $\theta = 0°, \varphi = 90°$. TP, FP, and FN are accumulated over all source-position iterations.
Experiments find the single source in all tested white- and pink-noise cases, with angular errors differing by at most 0.5° between source signal types, see Figure 6. The measured and simulated score maps form nearly identical clusters but show slightly different per-axis projection spacing, as shown in Figure 7. In the turntable experiment, Figure 8 shows the expected azimuthal RMSE of 1.1° from the 5° resolution step and an experimental RMSE of 2.1°. The directional error pattern is caused by a measurement error, where the array center was not aligned with the turntable

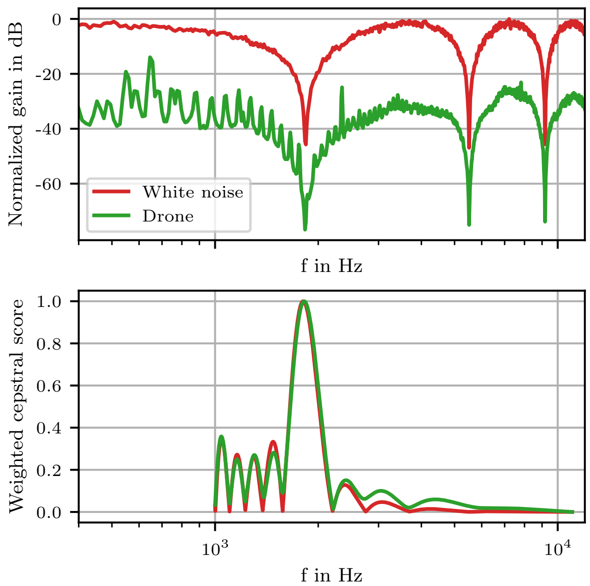


**Figure 3:** White-noise and drone signals from the same angle yield similar cepstral scores after zoom-FFT processing. The drone signal is attenuated to avoid visual overlap with the white-noise spectrum.

center because the physical mounting point is not co-located with the azimuthal center.

The approach can detect more than one source, but the two-source case is sensitive to relative signal gain. In Figure 9, both simulation and experiment find the white-noise source but not the pink-noise source. With an additional +6 dB gain for the pink-noise source, the simulation also detects the second source. Implications are further discussed in Section 4.

## 4. Discussion

The demonstrated approach exploits the periodic comb-filter structure produced by first-order spatial filtering. A fast and deterministic cepstral score can

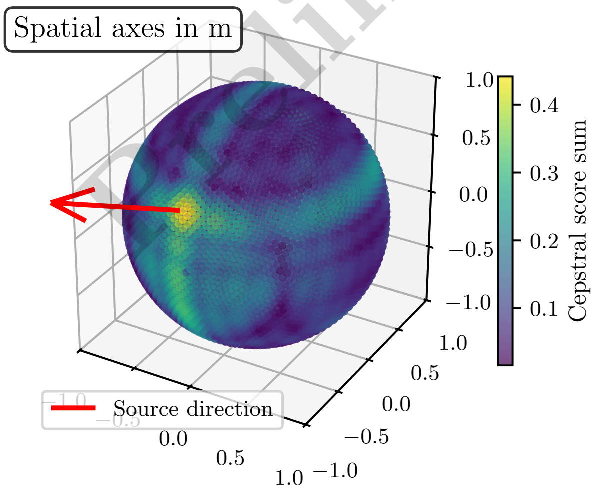


**Figure 4:** Global cepstral score for a single far-field noise source; the maximum marks the estimated source direction.

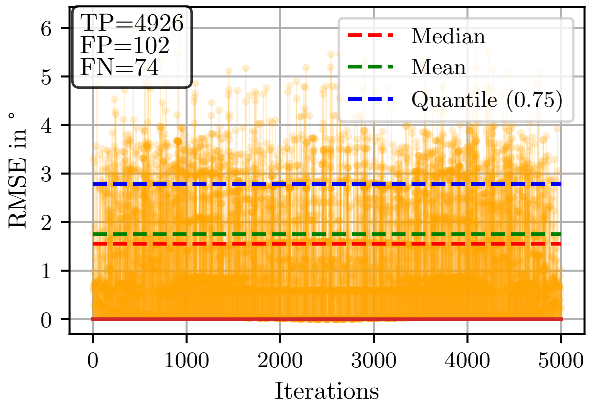


**Figure 5:** RMSE across single-source simulation iterations on the Fibonacci sphere.

**Table 1:** Detection performance and angular RMSE for simulated one- and two-source scenarios on the octahedral and tetrahedral array geometries.

<table>
<tr><th>Scenario</th><th>Oct. 1 src.</th><th>Oct. 2 src.</th><th>Tet. 1 src.</th><th>Tet. 2 src.</th></tr>
<tr><td>TP</td><td>4926</td><td>5280</td><td>4767</td><td>5549</td></tr>
<tr><td>FP</td><td>102</td><td>2288</td><td>1089</td><td>826</td></tr>
<tr><td>FN</td><td>74</td><td>4720</td><td>233</td><td>4451</td></tr>
<tr><td>Precision</td><td>0.98</td><td>0.70</td><td>0.81</td><td>0.87</td></tr>
<tr><td>Recall</td><td>0.99</td><td>0.53</td><td>0.95</td><td>0.55</td></tr>
<tr><td>F1</td><td>0.98</td><td>0.60</td><td>0.88</td><td>0.68</td></tr>
<tr><td>RMSE in deg</td><td></td><td>src. 1 | 2</td><td></td><td>src. 1 | 2</td></tr>
<tr><td>Max.</td><td>6.12</td><td>9.87 | 12.20</td><td>13.66</td><td>13.1 | 12.3</td></tr>
<tr><td>Med.</td><td>1.56</td><td>0.95 | 3.15</td><td>2.26</td><td>1.09 | 2.31</td></tr>
<tr><td>Mean</td><td>1.75</td><td>1.98 | 3.75</td><td>2.43</td><td>1.46 | 2.59</td></tr>
</table>

therefore act as a DoA indicator, in contrast to peak detection. For reflection-free single-source cases, the approach detects sources reliably with an angular RMSE of $\approx 2°$, as seen in columns 1 and 3 in Table 1, Figure 6, and Figure 8. The experimental results are similar to the simulations, and Figure 5 shows no bias toward individual sensor-axis placements within the geometry.

Since the spatial filter structures sum in the energy domain, their individual parts are preserved for sparse source scenarios until pathological cases excite the system such that the filter response becomes uniform. This allows a single first-order directive receiver to detect multiple relative, sparse, elevation-ambiguous source directions in principle, while additional receivers reduce the ambiguity and yield global DoA estimates. The two-source simulations and measurement in Figure 9 demonstrate this capability, but also show its current limitation. Different source levels can make an actual source direction weaker than a shadow source; in the shown case, the second source is detected only after increasing the pink-noise spectral

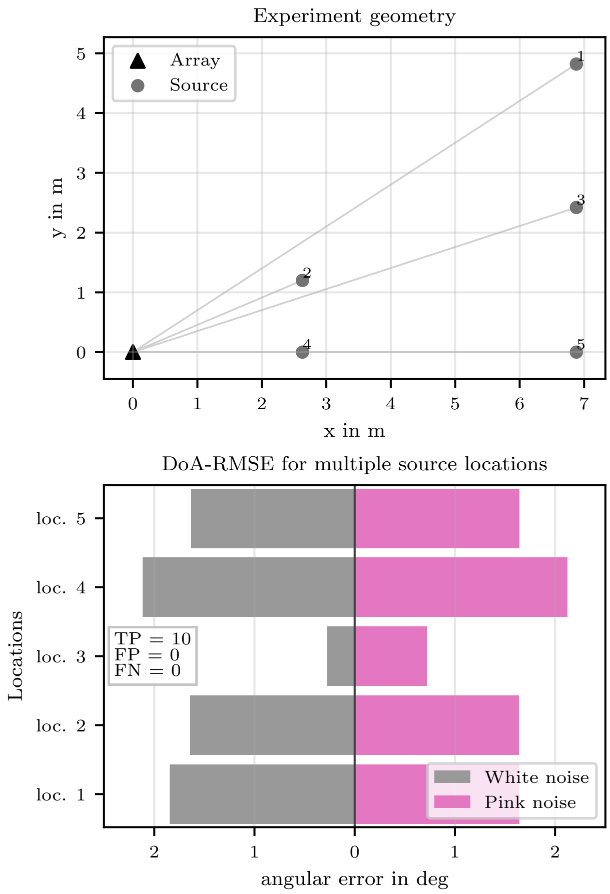


**Figure 6:** All independently tested white- and pink-noise source locations were detected. Coordinates and RMSE values are shown for each source location.

power by 6 dB, overcompensating for the −3 dB per octave power difference from pink to white noise.

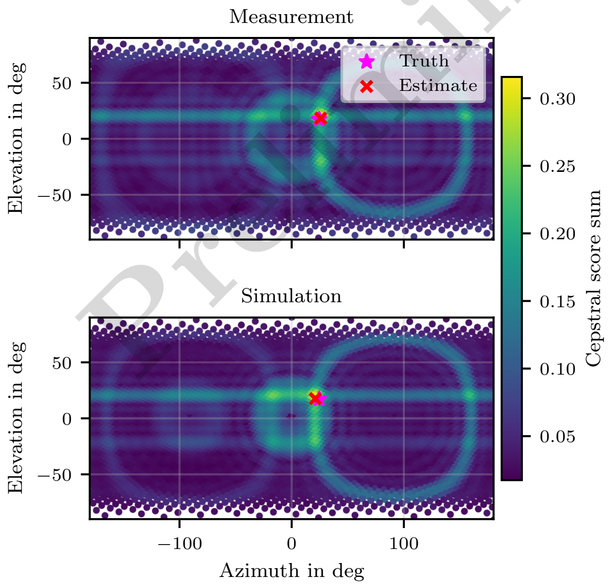


**Figure 7:** Measurement and simulation of a white-noise source at the same coordinates produce similar planar score projections. True and estimated locations are shown.

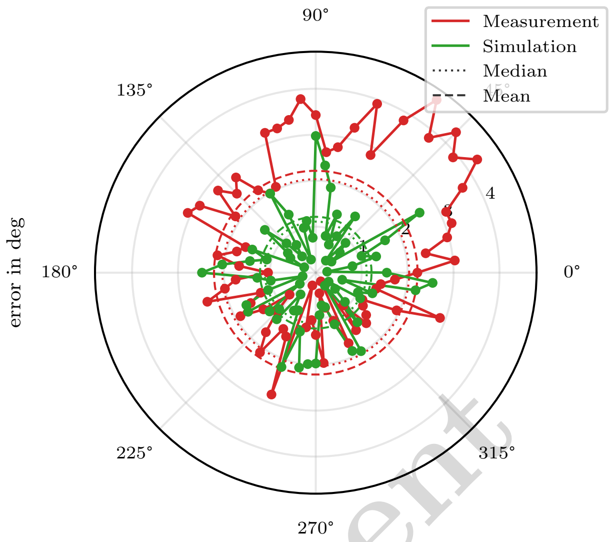


**Figure 8:** Azimuthal RMSE for simulation and turntable experiment in 5° steps. Mean and median values across azimuthal angle are also shown.

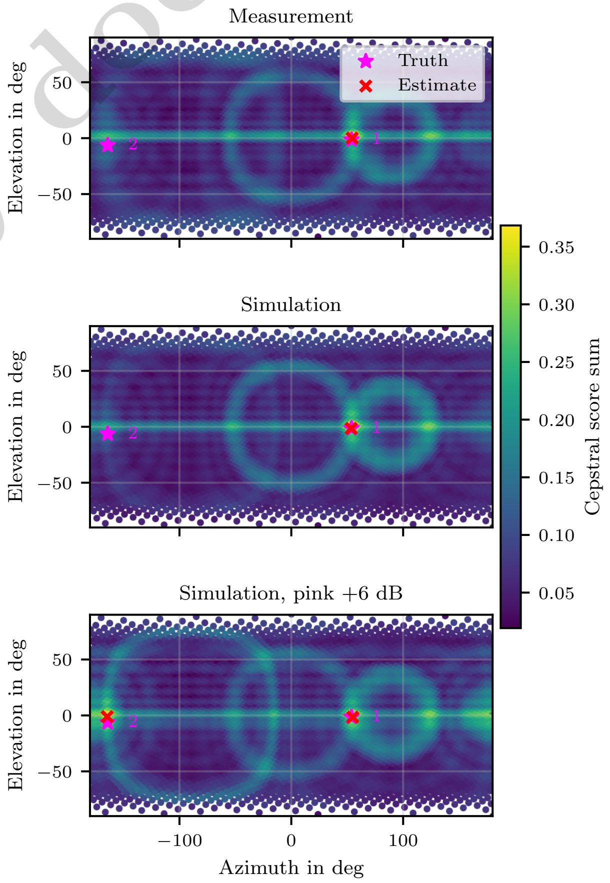


**Figure 9:** Two-source measurement and simulation at the same coordinates as planar projections. Source 1 is white noise; source 2 is pink noise. The bottom plot shows the same simulation with an additional gain of +6 dB for the pink-noise source.

The main limitation is therefore not single-source accuracy but score-map separation in multi-source cases. As visible in Figure 7, summing cepstral scores on the unit sphere creates "shadow sources", which can be ignored by global maximum thresholding for a single source but complicates multiple-source detection. Since all shadow sources are symmetrically correlated via sensor axes, spatial filtering across the cepstral score map, convolutional filters, or a CLEAN-like subtraction of a simulated single-source score map may improve separation. The source signal does not have to be noise, but it must excite the filtered band sufficiently evenly. Because the zoom FFT is fitted to the sensor path delay, sporadic spectral changes such as tonal components can be ignored, as shown in Figure 3.

## 5. Conclusion

Cepstral decomposition of first-order spatial filtering enables deterministic, phaseless DoA estimation from per-receiver spectra when sources emit a wideband signal. In reflection-free single-source cases, simulations and measurements reached about 2° angular RMSE. The results also show that more than one source direction can be recovered, although current score-map separation depends on relative source level and spectral balance. Future work should model suitable source spectra, suppress symmetric shadow sources on the spherical score map, and compare computational speed relative to classical beamforming methods. Code for simulation and experiment can be found at codeberg.org/jiep/ca-rss-doa-nod and codeberg.org/jiep/ca-doa-measurement.

## References


[1] R. Watson Watt and J. Herd, "An Instantaneous Direct-Reading Radiogoniometer," *Journal of the Institution of Electrical Engineers*, vol. 64, no. 353, pp. 611–617, May 1926, doi: 10.1049/jiee-1.1926.0051.

[2] H. De A. Donisthorpe, "The Marconi Marine Radio Direction Finder," *Proceedings of the Institute of Radio Engineers*, vol. 13, no. 1, pp. 29–47, Feb. 1925, doi: 10.1109/JRPROC.1925.220925.

[3] W. L. Dooley, "Ribbon microphones," *The Journal of the Acoustical Society of America*, vol. 136, no. 4, p. 2130, Oct. 2014, doi: 10.1121/1.4899679.

[4] P. i, R. Schwarz, G. Co, and Kg, "An Introduction to Direction Finding Methodologies White Paper," 2020.

[5] A. Cidronali *et al.*, "Comparison of Phase-less Direction of Arrival Estimation Methods for Switched Beam Antennas," in *2019 PhotonIcs & Electromagnetics Research Symposium - Spring (PIERS-Spring)*, Jun. 2019, pp. 2346–2354. doi: 10.1109/PIERS-Spring46901.2019.9017336.

[6] Q. Zheng, L. Luo, H. Song, G. Sheng, and X. Jiang, "A RSSI-AOA-Based UHF Partial Discharge Localization Method Using MUSIC Algorithm," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–9, 2021, doi: 10.1109/TIM.2021.3070617.

[7] T. Hayashi and Y. Kurosaki, "Noise Source Detection Using a Parabolic Reflector," *Journal of the Japan Society of Precision Engineering*, vol. 51, no. 8, pp. 1521–1526, 1985, doi: 10.2493/jjspe1933.51.1521.

[8] M. Passafiume, S. Maddio, A. Cidronali, and G. Manes, "MUSIC Algorithm for RSSI-based DoA Estimation on Standard IEEE 802.11/802.15.x Systems."

[9] F. Sohrabi and W. Yu, "Hybrid Digital and Analog Beamforming Design for Large-Scale Antenna Arrays," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 501–513, Apr. 2016, doi: 10.1109/JSTSP.2016.2520912.

[10] V. V. Ratnam, A. F. Molisch, O. Y. Bursalioglu, and H. C. Papadopoulos, "Hybrid Beamforming With Selection for Multiuser Massive MIMO Systems," *IEEE Transactions on Signal Processing*, vol. 66, no. 15, pp. 4105–4120, Aug. 2018, doi: 10.1109/TSP.2018.2838557.

[11] A. Politis, S. Delikaris-Manias, and V. Pulkki, "Direction-of-Arrival and Diffuseness Estimation above Spatial Aliasing for Symmetrical Directional Microphone Arrays," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2015, pp. 6–10. doi: 10.1109/ICASSP.2015.7177921.

[12] E. Tengan, T. Dietzen, F. Elvander, and T. van Waterschoot, "Direction-of-Arrival and Power Spectral Density Estimation Using a Single Directional Microphone and Group-Sparse Optimization," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2023, no. 1, p. 38, Oct. 2023, doi: 10.1186/s13636-023-00304-8.

[13] K. Imoto and N. Ono, "Spatial Cepstrum as a Spatial Feature Using a Distributed Microphone Array for Acoustic Scene Analysis," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 6, pp. 1335–1343, Jun. 2017, doi: 10.1109/TASLP.2017.2690559.

[14] R. Nagase, K. Oishi, and T. Furukawa, "A cepstrum prefiltering approach for DOA estimation of speech signal in reverberant environments," in *21st European Signal Processing Conference (EUSIPCO 2013)*, Sep. 2013, pp. 1–5. Accessed: Mar. 27, 2026. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6811509

[15] J. Tschavoll and E. Sarradj, "Ein Verfahren zur Schallquellenlokalisierung im Kugelraum basierend auf der akustischen Richtwirkung nicht-omnidirektionaler Mikrofonarrays und RSS," in *Fortschritte der akustik - DAGA 2026*, Dresden, Germany: Deutsche Gesellschaft für Akustik e.V. (DEGA), Mar. 2026, pp. 163–166. doi: 10.71568/dasdaga2025.511.