

# SOUND SOURCE LOCALIZATION USING SECTOR-BASED ANALYSIS WITH MULTIPLE RECEIVERS

Michael MCCREA<sup>1</sup>, Leo MCCORMACK<sup>1</sup>, and Ville PULKKI<sup>1</sup>

<sup>1</sup>Department of Signal Processing and Acoustics, Aalto University, Espoo, Finland

## ABSTRACT

The success of parametric approaches to spatial sound reproduction and sound field navigation depend on the accuracy of the initial analysis and decomposition of the sound field. In this work, the sector-based high-order extension to intensimetric sound field analysis is evaluated in the context of 3D source localization. The evaluation is performed with simulations of ideal spherical harmonic receiver signals using two intensimetric estimators: source direction-of-arrival and sound field diffuseness. The technique is first assessed for a single receiver with regard to influential factors of analysis order, source incidence angle, and the presence of diffuse noise. The technique is then applied to 3D source localization, utilizing concurrent analyses from multiple receivers. Results for different analysis orders are compared and mitigating factors for robust localization over a broad spatial region are discussed. Optimization strategies targeting specific conditions are proposed, tested, and found to improve localization accuracy.

## 1. INTRODUCTION

In recent years, rapid advances in augmented reality and remote presence applications has brought renewed attention to virtual 3D sound reproduction. This is due to the fact that accurate spatial sound is requisite to create a convincing sense of virtual presence, which is reinforced by an auditory environment that is responsive to a listener that is fully mobile. As such, there is increasing demand to accurately capture and reconstruct spatial sound scenes across navigable space. A prerequisite to auditory scene reconstruction is effective *sound field decomposition*—inferring the spatial organization and behavior of recorded sound sources—which remains an open challenge.

Spherical microphone arrays (SMAs) have proven to be a central tool in sound field capture and analysis, as they offer an efficient and directionally consistent means of encoding incident sounds from all directions via a truncated spherical harmonic (SH) series [1]. The resultant signals comprise a spatial sound format, known as Higher Order Ambisonics (HOA), representing a sound scene from one point in space which can be rendered binaurally to headphones and rotated with a listener’s motion-tracked head

movements.

Despite the utility of HOA, the listener’s auditory perspective is co-located with the SMA and cannot accurately be shifted far beyond the recorded position in space [1], which presents a problem for virtual sound field navigation. Numerous authors have sought to address these limitations by *parametric decomposition* of the recorded sound field, whereby salient properties of the sound field are analyzed and parameterized to enable selective reconstruction of the sound scene [2, 3]. Indeed, parametric approaches have been used to extend the navigable range of single-perspective [4, 5] and multiperspective recordings [6].

Intensimetric sound field analysis offers an intuitive decomposition of the sound field into directional and a diffuse sound components [7] and is widely used in parametric spatial audio [2, 8]. Classical first-order intensimetric analysis offers only “global” sound field measures so is of limited use in complex and multisource sound scenes. However, the method has been adapted to exploit the full spatial resolution and beamforming capabilities of HOA signals to provide estimates of sound field parameters within confined angular regions, called *sectors* [9–11]. So-called *sector-based* parametric sound field decomposition has been shown to mitigate the adverse effects of concurrent sound sources to improve single-perspective sound field recordings [2, 12].

Sector-based analysis is a promising approach to complex scene analysis, though its performance across multiple distributed receivers in 3D space has not been explored and is the topic of this work. In what follows, the performance of sector-based sound scene analysis will be characterized from the perspective of source localization. In Section 2, intensimetric sound field analysis in the SH domain and its extension to higher orders are reviewed. In Section 3, single-receiver DoA estimation using sector processing is evaluated for prototype sound fields, which is extended to 3D source localization by multiple concurrent analyses. Conclusions are discussed in Section 4.

## 2. SOUND FIELD ANALYSIS

### 2.1 Sound Field Model

The pressure at a point in the sound field can be represented as the sum of a directional amplitude distribution defined across the unit sphere as

$$p = \int_{\Omega} a(\Omega) d\Omega, \quad (1)$$

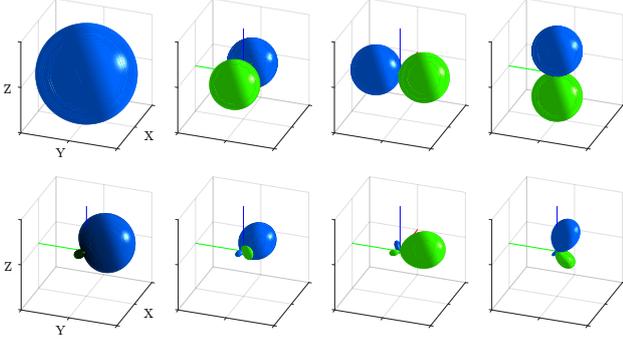


Figure 1. Pressure-velocity beam patterns (top row), spatially weighted by a second-order max- $\mathbf{r}_E$  sector steered to  $\Omega = (75^\circ, -50^\circ)$  (bottom row).

where the direction  $\Omega = (\theta, \varphi)$  is comprised of inclination  $\theta \in [0, \pi]$  and azimuth  $\varphi \in [-\pi, \pi]$  and  $\int_{\Omega} d\Omega = \int_{-\pi}^{\pi} d\varphi \int_0^{\pi} \sin \theta d\theta$  denotes integration over the unit sphere. The acoustic particle velocity at that point in the sound field is defined as [10]

$$\mathbf{u} = -\frac{1}{Z_0} \int_{\Omega} \mathbf{d}(\Omega) a(\Omega) d\Omega = -\frac{1}{Z_0} \mathbf{v}, \quad (2)$$

where three orthogonal dipole patterns  $\mathbf{d}(\Omega) = [d_x(\Omega), d_y(\Omega), d_z(\Omega)]^T = [\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta]^T$  spatially weight the amplitude distribution to produce the unnormalized Cartesian components of the particle velocity, represented by the signal vector  $\mathbf{v}$ . Normalizing by the characteristic impedance of air  $Z_0 = c\rho_0$  yields the acoustic particle velocity.

The time-varying amplitude distribution of a sound field may be encoded into a vector of spherical harmonic (SH) signals by the Spherical Harmonic Transform (SHT)

$$\mathbf{a}(t) = \mathcal{SHT}\{a(t, \Omega)\} = \int_{\Omega} a(t, \Omega) \mathbf{y}(\Omega) d\Omega, \quad (3)$$

where  $\mathbf{y}(\Omega)$  is a vector of real SHs evaluated at  $\Omega$ .

## 2.2 Intensity Analysis

By proper normalization,  $\mathbf{a}(t)$  represents a conventional ambisonic signal such that the zeroth-degree component corresponds to an omnidirectional pressure signal at the origin and the first-degree components correspond to three orthogonal dipoles. These signals can be transformed into the pressure and unnormalized particle velocity of the sound field (omitting the time dependence) by

$$\mathbf{a}_{pv} = \begin{bmatrix} p \\ \mathbf{v} \end{bmatrix} = \mathbf{W}_{pv,1} \mathbf{a} \quad (4)$$

where  $\mathbf{W}_{pv,1}$  is a transformation matrix that scales the first-degree components by  $1/\sqrt{3}$  and re-orders them to match Cartesian component ordering. The pressure-velocity signals derived from the SH domain can then produce correlative measures of the sound field active inten-

sity and energy by [13]

$$\mathbf{I}_a = -\Re\{p\mathbf{v}^H\}, \quad (5)$$

$$E = \frac{1}{2} [|p|^2 + \mathbf{v}^H \mathbf{v}], \quad (6)$$

where  $(\cdot)^H$  is the Hermitian transpose for complex signals, or transpose for real signals, and  $\Re\{\cdot\}$  is the real operator.

The active intensity indicates the direction of net energy transport, from which the direction-of-arrival (DoA) of a sound source may be inferred. The magnitude of the active intensity in proportion to the overall energy,  $\|\mathbf{I}_a\|/E$ , indicates the normalized speed of energy propagation and has been formulated into a correlative metric of sound field *diffuseness*, defined as [14]

$$\psi = 1 - \frac{\|\mathbf{I}_a\|}{E} = 1 - \frac{2\|\Re\{p\mathbf{v}^H\}\|}{|p|^2 + \mathbf{v}^H \mathbf{v}}. \quad (7)$$

This formulation of diffuseness  $\psi \in [0, 1]$  has been defined as the ratio of locally confined energy [14], wherein  $\psi = 1$  indicates a purely directional sound field, as with a single plane wave, and  $\psi = 0$  may indicate solely oscillatory, non-propagating energy, as would be observed in a standing wave field. Diffuseness may also be interpreted as the degree to which sound arrives from all directions with equal probability [7], thereby  $\psi = 0$  may indicate a sound field comprised of mutually uncorrelated, isotropically distributed plane waves, given a sufficient spatial or temporal averaging. Numerous related intensimetric formulations of diffuseness are compared in [7].

The diffuseness measure and DoA estimation for a mixed field, comprised of a plane wave amid a diffuse field, will vary according to the ratio of powers of the plane wave and diffuse sound, called the direct-to-diffuse ratio (DDR) [7]

$$\Gamma = \frac{\mathbb{E}\{|p_{pw}|^2\}}{\mathbb{E}\{|p_{df}|^2\}}, \quad (8)$$

where  $\mathbb{E}\{\cdot\}$  is the expectation operator.

### 2.2.1 High-order Sector-based Intensity Analysis

Beamforming spatial filters can be designed in the SH domain to extract a direction-dependent signal from a recorded sound field. Applying such a filter to a spatial amplitude distribution generates a spatially-weighted pressure signal by

$$p_w = \int_{\Omega} w(\Omega) a(\Omega) d\Omega = \mathbf{w}^T \mathbf{a}, \quad (9)$$

where  $\mathbf{w}$  is a vector of order-limited beamforming coefficients, corresponding to a directional gain pattern  $w(\Omega)$ . Functions that generate the coefficients of common axisymmetric patterns, such as the ‘max- $\mathbf{r}_E$ ’ pattern used throughout this work, are found in [13]. Applying this directional gain on the dipole patterns of Equation (2) imposes further spatial selectivity on the resultant sound field velocity components. The spatially-filtered pressure and velocity components are considered to be local to a sub-region of the sound field, a so-called *sector*, defined by the

the *sector pattern*  $w(\Omega)$  [9]. The acoustic particle velocity of the sound field sector is then

$$\mathbf{u}_w = -\frac{1}{Z_0} \int_{\Omega} w(\Omega) \mathbf{d}(\Omega) a(\Omega) d\Omega = -\frac{1}{Z_0} \mathbf{v}_w. \quad (10)$$

The resulting weighted dipole patterns are visualized in Fig. 1. The *sector components*, signals corresponding to the pressure and velocity local to the sector, are formed by [10, 11]

$$\mathbf{a}_{pv} = \begin{bmatrix} p_w \\ \mathbf{v}_w \end{bmatrix} = \mathbf{W}_{pv,L} \mathbf{a}. \quad (11)$$

The matrix  $\mathbf{W}_{pv,L} = [\hat{\mathbf{w}}, \mathbf{w}_x, \mathbf{w}_y, \mathbf{w}_z]^\top$  is comprised of four beamforming coefficient vectors. The velocity-beamforming coefficients  $\mathbf{w}_i$  correspond to the spatially-weighted dipole pattern of Equation (10) and are constructed through a transformation of the sector pattern coefficients  $\mathbf{w}$  by

$$\mathbf{w}_i = \mathbf{A}_i \mathbf{w} \quad \text{for } i = \{x, y, z\}, \quad (12)$$

where  $\mathbf{A}_i$  are sparse, deterministic matrices with dimension  $(L+1)^2 \times L^2$  and derive from the spherical expansion of the product of two spherical functions, as detailed in [10]; which are independent of the sector pattern, so may be pre-computed for the desired input SH order in practice.

Equations (9) and (11) show that both the velocity- and pressure-beamforming weights will need to be of the same order as the input HOA signal. However, because the velocity-beamformers are the product of the sector pattern and a first-order dipole pattern,  $\mathbf{w}_i$  will be one order higher than  $\mathbf{w}$ . It follows that, for a sound field signal of order  $L$ , the sector pattern will need to be designed at order  $L-1$ . As such,  $\hat{\mathbf{w}}$  denotes the sector pattern coefficient vector  $\mathbf{w}$  zero-padded from length  $L^2$  to the length  $(L+1)^2$  to match the input order. Finally, the weighted pressure and velocity signals of Equation (11) can be substituted into Equations (5)-(7) to produce spatially-localized estimates of intensimetric sound field quantities, such as the diffuseness and sound source DoA.

### 3. EVALUATION OF SECTOR-BASED SOURCE LOCALIZATION

To evaluate the performance of sound field parameter estimation using sector processing, an acoustic scene is simulated<sup>1</sup> in MATLAB. Plane waves are encoded directly in the SH domain accounting for the source DoA and distance to an ideal spherical harmonic receiver. Gaussian white noise is used to model uncorrelated source signals. The isotropic diffuse noise component is modeled by plane waves of white noise encoded in the directions of a 216-point  $t$ -design and normalized to the target DDR using Equation (8). The diffuse field level is set equally in all receivers, according to the DDR of a reference source-receiver arrangement: a source in the ‘‘focal position’’ and the Receiver 1, as depicted in Fig. 3a. Sources and receivers are located on a horizontal plane spanning a 24 m<sup>2</sup> area. When higher-order analysis is performed, a sector from each receiver is steered toward the focal position,

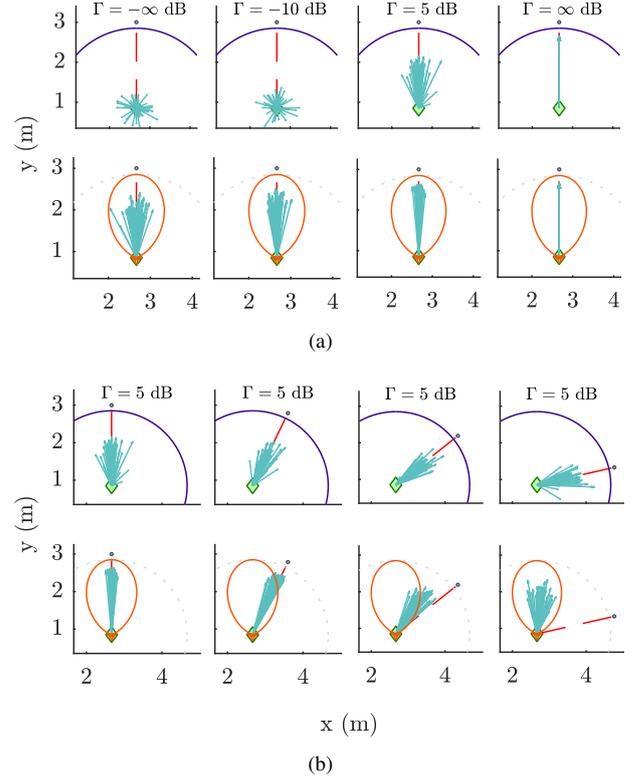


Figure 2. DoA vectors ( $\rightarrow$ ) from an acoustic scene analysed at orders  $L = [1, 4]$  with (a) an on-axis source, varying DDR, and (b) a source moving off-axis,  $\Gamma = 5$  dB. DoA vectors and overlaid sector patterns are normalized to the distance between source ( $\circ$ ) and receiver ( $\diamond$ ).

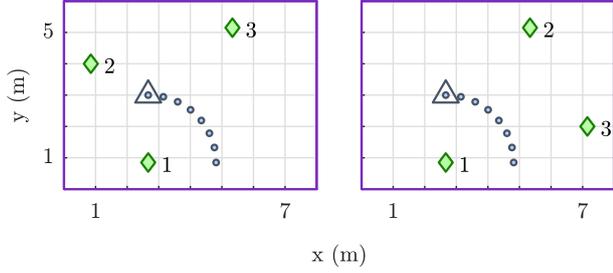
thereby relative source proximity is accounted for in the constituent analyses. A temporal average of 3 ms is taken of the power- and cross-spectral density terms of Equations (5)-(7) to exploit the time-ergodicity of the random processes comprising the source and diffuse field signals, which approximates the expectation operator in the statistical sound field model of Equation (8) [7].

#### 3.1 Single Receiver DoA Estimation

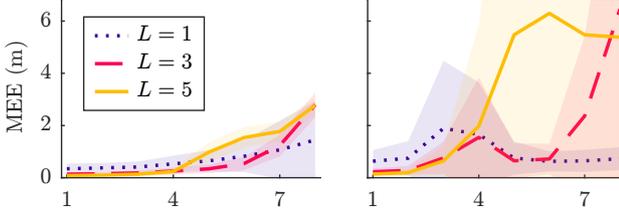
The proportion of directional to diffuse energy observed at the receiver location is central to DoA estimation. This ratio is expressed by the vector  $\mathbf{I}_a/E$  in which the active intensity, as a measure of net energy transport, will comprise a large proportion of the total energy in the case of a single prominent sound source. Negating the vector gives the *DoA vector* which points in the DoA of the propagating energy. Its length is proportional to the sound field directivity, where low magnitude indicates high diffuseness.

DoA vectors resulting from first- and fourth-order analysis are shown in Fig. 5 for successive analysis frames over a one second steady state. Sector processing is clearly advantageous in the case of an on-axis target source (Fig. 2a), wherein off-axis noise is attenuated and the increased DDR equates to an increased signal-to-noise ratio and improved DoA estimates. However, the *source* energy is attenuated if it moves off-axis from the sector (Fig. 2b) and the DoA

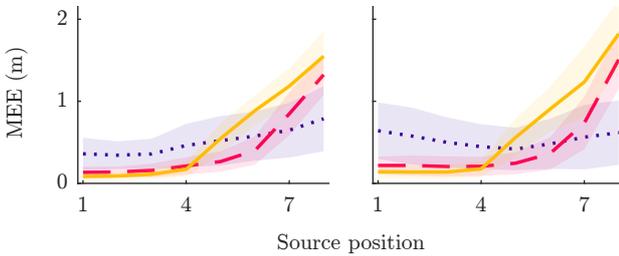
<sup>1</sup> <https://github.com/polarch/shoebbox-roomsim>



(a) Scene layout: source (o), receiver (◇), sector focal positions (△).



(b) Localization performance by Receivers 1 and 2.



(c) Localization performance by all three receivers.

Figure 3. Mean estimation error (MEE) of single-source localization, with one standard deviation shaded. The first position is at the focal position. Note the difference in range between (b) and (c).  $\Gamma = 10$  dB.

estimates are drawn away, toward sector pattern maximum, as the diffuse field begins to dominate the spatially-filtered DDR. There is a clear trade-off between the increased accuracy of on-axis high-order estimation and the angular width in which the accuracy is maintained.

### 3.2 Multiple Receiver Source Localization

The advantages and caveats of high-order sector-based processing carry forward into three-dimensional (3D) source position estimation, here called *source localization*. The task employs intensimetric DoA estimation across multiple distributed SH receivers, requiring additional considerations of the number and arrangement of receivers to optimize localization. Source position estimation is performed by a least squares solution to the intersection of DoA projection lines

$$\mathbf{R}\mathbf{p} = \mathbf{q}, \quad (13)$$

with

$$\mathbf{R} = \sum_{j=1}^K w_j (\mathbf{I} - \gamma_j \gamma_j^\top), \quad \mathbf{q} = \sum_{j=1}^K w_j (\mathbf{I} - \gamma_j \gamma_j^\top) \mathbf{r}_j,$$

where  $\mathbf{r}_j$  are the positions of  $K$  receivers, and  $\gamma_j$  are the corresponding unit-length DoA estimate vectors. The

method minimizes the sum of squared distance between the estimated source position  $\hat{\mathbf{p}}$  and the DoA projection lines, with an optional weight  $w_j$  which biases the estimate toward the  $j$ -th DoA vector. The system of equations is solved directly or via

$$\hat{\mathbf{p}} = \mathbf{R}^\dagger \mathbf{q}, \quad (14)$$

where  $^\dagger$  denotes the Moore-Penrose pseudoinverse.

#### 3.2.1 Single Source Localization

Localization performance is evaluated for eight source positions starting in the focal position and arcing outward to a  $90^\circ$  incidence angle with Receiver 1. In this way the DDR at the reference receiver will be constant, while the DDR will vary with source position for the other receivers. Figures 3b and 3c show the mean estimation error of localization for two- and three-receiver configurations, respectively. The performance of the two-receiver configuration shows both higher error and larger variance in estimation, when first-order processing is used (i.e. without spatial filtering). The error is consistently lower at higher orders in positions nearest to the focal position, but performance degrades as the source moves further away. In the left configurations, for example, DoA estimation degrades in Receiver 1 as the source moves off-axis. A corresponding degradation occurs in Receiver 2 as the DDR decreases with the receding source. The error therefore compounds for source positions farthest from the focal point. The right two-receiver configuration of Fig. 3b is ill-conditioned for the localization algorithm; for a source positioned between the receivers, the DoA vectors approach collinear orientations and small errors result in large variance of intersection points. In this unstable condition the MEE peaks at different source positions for different orders, and thus lower-order processing is more preferable.

Localization is improved considerably with the addition of a third receiver, as in Fig. 3c. In addition to improved MEE overall, the error variance is made largely consistent across source positions. Estimation accuracy degrades with increasing order for those source positions which lie substantially off-axis from one or more receivers. In Fig. 4a localization performance is extrapolated across the broad regions. The region of highest localization accuracy anchors to the focal position, further drawing toward the centroid of the receiver arrangement where DoA vectors would be maximally orthogonal.

#### 3.2.2 With an Interfering Source

Localization error is evaluated for a second sound scene in which the target source remains fixed in the focal position and an equal-power interferer traverses the region enclosed by the receivers. Figure 4b should be read as the error in localizing a source at the focal position while an interferer is located at the plotted position. As such, it can be more informative to consider the areas in which the localization deteriorates. Higher error is observed in regions where the interferer approaches a receiver, but from a direction that is misaligned with the target DoA. In such a case the energetic DDR for the nearby receiver will be dominated by the

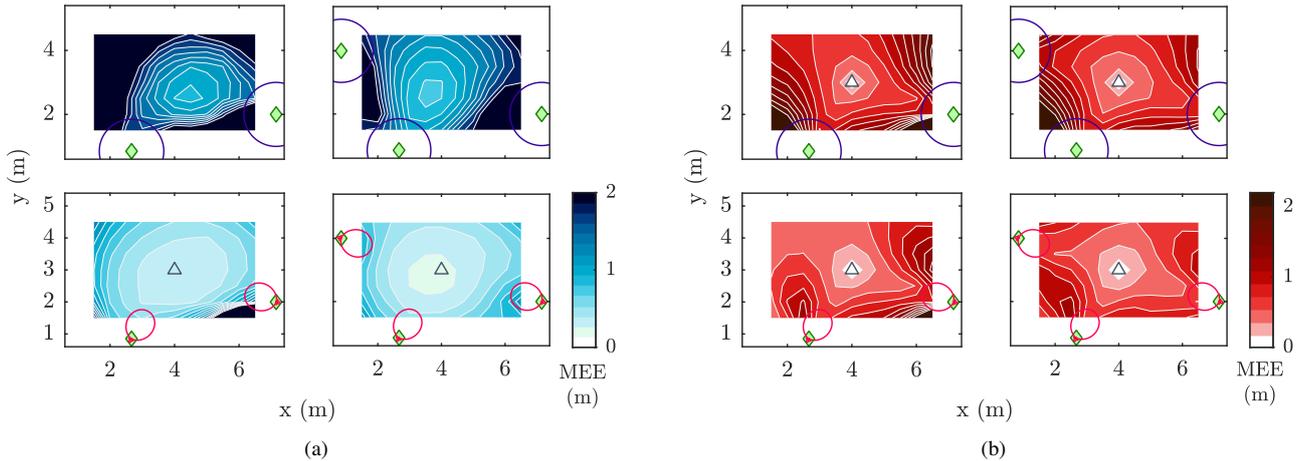


Figure 4. Mean estimation error (MEE) for localization of (a) a source traversing the observation region, and (b) a source in the focal position ( $\Delta$ ) while an equal-power interferer moves throughout the same region.  $L = [1, 4]$ ,  $\Gamma = 10$  dB.

proximate interfering source, drawing the DoA estimates toward it, away from the target. This is especially apparent in the two-receiver configurations, where DoA bias more strongly distorts the resulting localization. Shallower error contours with increasing order indicate the advantage of sector processing to mitigate off-axis interference. The peak accuracy at the focal position is naturally high on account of the interferer coinciding with the target source.

### 3.3 Estimation Refinement

#### 3.3.1 Distance Weighting

Considering the impact of the DDR on localization accuracy, it follows that as a target source traverses an observed region, the nearest receivers will likely have better DoA estimates. This motivates an estimation weighting scheme that modifies the contribution of each DoA vector to the resultant localization by the weighting term of Equation (13) [15]. The weighting scheme chosen for demonstration is

$$w_j = 1 - \frac{d_j}{d_{\max}}(1 - \vartheta), \quad (15)$$

where  $d_j$  is the Euclidean distance between the initial source position estimate and the  $j$ -th receiver, for  $j \in [1, K]$ , and  $d_{\max}$  is the maximum of  $K$  distances.  $\vartheta \in [0, 1]$  represents a user-defined compression of the applied weight range, whereby  $\vartheta = 0$  would fully diminish the contribution of the farthest receiver and  $\vartheta = 1$  would remove distance-weighting.

Results for localization with distance weighting for a single source sound scene are shown in Fig. 5a. Comparing the unweighted and distance-weighted localization techniques shows that the accuracy is increased for a larger region, particularly extending into areas between adjacent receivers, as the estimates from those nearer receivers are given a greater weight. Note that this result would not hold in the presence of an interferer because the initial source position estimate will be drawn toward the inter-

ferer, thereby applying higher weights to receivers farther from the target.

#### 3.3.2 Diffuseness Weighting

For scenarios in which it is known that an interfering source or high DDR is degrading localization of a target source, diffuseness estimation can inform the weighting of Equation (13) and emphasize estimates with a higher degree of directional energy. The weighting scheme chosen here is

$$w_j = \max(\psi_j - \vartheta_{\min}, \epsilon), \quad (16)$$

where  $\vartheta_{\min}$  is a minimum expected diffuseness, accounting for the changing diffuseness bounds with sector order (refer to [10]). The  $\max(\cdot)$  operation ensures the weighted contribution is positive non-zero, clipped at the value  $\epsilon > 0$ , which can also serve as a minimum contribution weight. Figure 5b shows the results of applying diffuseness-weighted estimation to the interfering-source scenario (see Fig. 4). The right panel shows improved performance particularly in regions between adjacent receivers and where the interferer is proximate but off-axes to a receiver.

## 4. CONCLUSIONS

An evaluation was presented of high-order sector-based analysis of sound fields with distributed spherical harmonic receivers, in the context of 3D source localization. After a review of intensimetric sound field analysis using first- and higher-order SH signals, performance characteristics were evaluated for source direction-of-arrival (DoA) estimation using ideal SH receivers. This was followed by an evaluation of a 3D source localization scheme which employed DoA estimates derived from analyses across multiple distributed SH receivers. Across all evaluations, it is clear that a primary advantage of sector-based analysis in source localization is the improvement of the direct-to-diffuse (DDR) for a target source signal (or similarly,

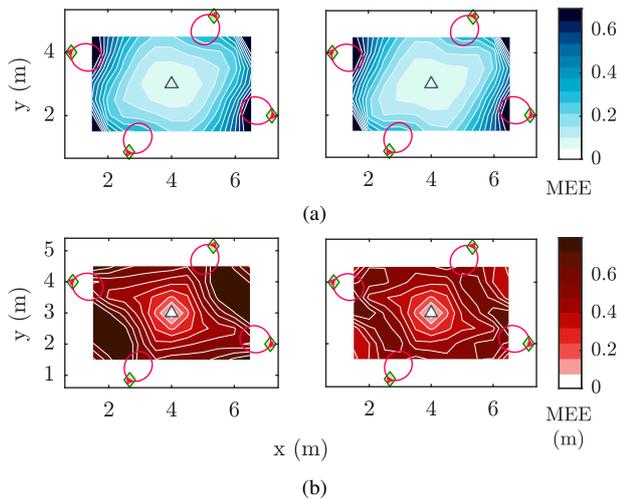


Figure 5. Mean estimation error (MEE) for localization with four receivers with estimation weight refinement by (a) estimated source distance in a single-source scene, and (b) diffuseness in the presence of an interferer. Left and right panels show unweighted and weighted results, respectively.  $\Gamma = 10$  dB.

the direct-to-interferer ratio), an effect which compounds when using multiple receivers. However, the advantages of higher-order processing diminish above a threshold angle-of-arrival of the target source, which decreases with increasing order, contingent upon the DDR. As such, the region of highest localization accuracy consistently anchors to the focal position of the sectors and further draws toward the centroid of the receiver arrangement where projected DoAs would be maximally orthogonal. A high receiver count can minimize the impact of the erroneous analyses, along with weighting the DoA estimates contributing to localization. Weighting schemes informed by the diffuseness metric and the initial source distance estimate were found to be beneficial under specific sound field conditions.

### Acknowledgments

This work has been funded by the “Nordic Sound and Music Computing Network—NordicSMC”, NordForsk project number 86892.

## 5. REFERENCES

- [1] M. A. Poletti, “Three-Dimensional Surround Sound Systems Based on Spherical Harmonics,” *J. Audio Eng. Soc.*, vol. 53, no. 11, pp. 1004–1025, 2005.
- [2] V. Pulkki, S. Delikaris-Manias, and A. Politis, *Parametric time-frequency domain spatial audio*, 1st ed. Hoboken, New Jersey, USA: John Wiley & Sons, 2018.
- [3] S. Berge and N. Barrett, “High Angular Resolution Planewave Expansion,” in *2nd Int. Sympos. Ambisonics and Spherical Acoustics*, Paris, France, May 2010.
- [4] T. Pihlajamäki and V. Pulkki, “Projecting Simulated or Recorded Spatial Sound onto 3d-Surfaces,” in *45th Int. Conf. Audio Eng. Soc.: Applicat. of Time-Frequency Process. in Audio*, Mar 2012.
- [5] A. Plinge, S. J. Schlecht, O. Thiergart, T. Robotham, O. Rummukainen, and E. A. P. Habets, “Six-Degrees-of-Freedom Binaural Audio Reproduction of First-Order Ambisonics with Distance Information,” in *2018 Int. Conf. Audio Eng. Soc.: Audio for Virtual and Augmented Reality*, Redmond, WA, USA, Aug. 2018.
- [6] C. Schörkhuber, R. Holdrich, and F. Zotter, “Triplet-based variable-perspective (6DoF) audio rendering from simultaneous surround recordings taken at multiple perspectives,” in *Fortschritte der Akustik (DAGA)*, Hanover, Germany, Mar. 2020.
- [7] G. Del Galdo, M. Taseska, O. Thiergart, J. Ahonen, and V. Pulkki, “The diffuse sound field in energetic analysis,” *J. Acoust. Soc. of America*, vol. 131, no. 3, pp. 2141–2151, Mar. 2012.
- [8] J. Merimaa and V. Pulkki, “Spatial Impulse Response Rendering I: Analysis and Synthesis,” *J. Audio Eng. Soc.*, vol. 53, no. 12, pp. 1115–1127, 2005.
- [9] A. Politis, J. Vilkamo, and V. Pulkki, “Sector-Based Parametric Sound Field Reproduction in the Spherical Harmonic Domain,” *IEEE J. of Selected Topics in Signal Process.*, vol. 9, no. 5, pp. 852–866, Aug. 2015.
- [10] A. Politis and V. Pulkki, “Acoustic intensity, energy-density and diffuseness estimation in a directionally-constrained region,” arXiv:1609.03409, Tech. Rep., Sep. 2016.
- [11] L. McCormack, S. Delikaris-Manias, A. Politis, D. Pavlidi, A. Farina, D. Pinaridi, and V. Pulkki, “Applications of Spatially Localized Active-Intensity Vectors for Sound-Field visualization,” *J. Audio Eng. Soc.*, vol. 67, no. 11, pp. 840–854, Nov. 2019.
- [12] L. McCormack, V. Pulkki, A. Politis, O. Scheuregger, and M. Marschall, “Higher-Order Spatial Impulse Response Rendering: Investigating the Perceived Effects of Spherical Order, Dedicated Diffuse Rendering, and Frequency Resolution,” *J. Audio Eng. Soc.*, vol. 68, no. 5, pp. 338–354, Jun. 2020.
- [13] A. Politis and V. Pulkki, “Higher-Order Directional Audio Coding,” in *Parametric Time-Frequency Domain Spatial Audio*. John Wiley & Sons, Ltd, 2018, ch. 6, pp. 141–159.
- [14] J. Merimaa, “Analysis, Synthesis, and Perception of Spatial Sound - Binaural Localization Modeling and Multichannel Loudspeaker Reproduction,” Ph.D. dissertation, Helsinki University of Technology, Espoo, Finland, 2006.
- [15] C. Schörkhuber, P. Hack, M. Zaunschirm, F. Zotter, and A. Sontacchi, “Localization of Multiple Acoustic Sources with a Distributed Array of Unsynchronized First-Order Ambisonics Microphones,” in *Proc. 6th Congress of Alps-Adria Acoust. Association*, Graz, Austria, Oct. 2014.