Sparse, ℓ_1-Optimal Multi-Loudspeaker Panning and its Relation to Vector Base Amplitude Panning

Andreas Franck[1], Wenwu Wang[2], and Filippo Maria Fazi[1]

[1]Institute of Sound and Vibration Research, University of Southampton, Southampton, UK
[2]Centre for Vision, Speech, and Signal Processing, University of Surrey, Guildford, UK

This is a preprint of the paper:

Andreas Franck, Wenwu Wang, and Filippo Maria Fazi. "Sparse, ℓ_1-Optimal Multi-Loudspeaker Panning and its Relation to Vector Base Amplitude Panning"[1]

In: *IEEE/ACM Transactions on Audio, Speech and Language Processing, 2017*

DOI: 10.1109/TASLP.2017.2674975

[1]This work was supported by the EPSRC Programme Grant S3A: Future Spatial Audio for an Immersive Listener Experience at Home (EP/L000539/1) and the BBC as part of the BBC Audio Research Partnership. All code and data underlying the findings are fully available without restriction. Details of the data and how to request access are available through the DOI 10.15126/surreydata.00813551.

Abstract

Panning techniques, such as vector base amplitude panning (VBAP) are a widely-used practical approach for spatial sound reproduction using multiple loudspeakers. Although limited to a relatively small listening area, they are very efficient and offer good localisation accuracy, timbral quality as well as a graceful degradation of quality outside the sweet spot. The aim of this paper is to investigate optimal sound reproduction techniques that adopt some of the advantageous properties of VBAP, such as the sparsity and the locality of the active loudspeakers for the reproduction of a single audio object. To this end, we state the task of multi-loudspeaker panning as an ℓ_1 optimization problem. We demonstrate and prove that the resulting solutions are exactly sparse. Moreover, we show the effect of adding a nonnegativity constraint on the loudspeaker gains in order to preserve the locality of the panning solution. Adding this constraint, ℓ_1-optimal panning can be formulated as a linear program. Using this representation, we prove that unique ℓ_1-optimal panning solutions incorporating a nonnegativity constraint are identical to VBAP using a Delaunay triangulation for the loudspeaker setup. Using results from linear programming and duality theory, we describe properties and special cases, such as solution ambiguity, of the VBAP solution.

1. Introduction

Sound reproduction over multiple loudspeakers aims at recreating plausible spatial sound scenes, often consisting of multiple audio objects, for either a single listener or over extended listening areas. As summarized in the review paper [1], this is an area with a long history but also of very active research. Spatial sound reproduction approaches can be broadly classified into physically and perceptually motivated techniques. The methods that attempt to physically recreate an acoustic field are referred to as *sound field synthesis* in [1]. Examples of sound field synthesis techniques include wave field synthesis (WFS), e.g., [2, 3, 4, 5, 6], Higher Order Ambisonics (HOA) [7, 5, 8] , and sound field control techniques [9, 10, 11, 12, 13, 14]. For a more thorough review, the reader is referred to [1] and the references therein.

While the former approaches are based on analytic descriptions of the acoustic field, sound field control generally employs an optimization approach to minimize the difference between the desired and the synthesized field, most often using an ℓ_2 (least-squares) error norm.

In contrast to physical reproduction techniques, perceptually motivated techniques attempt to achieve a plausible spatial perception by providing the relevant psychoacoustic cues at the listener's ears. Panning laws, which apply amplitude changes or time delays to the audio object's signal [15, 16], form important classes of perceptually motivated reproduction techniques. Vector-base amplitude panning (VBAP) [17] is likely the most widely used perceptually motivated method for two- and three-dimensional multi-loudspeaker reproduction. It is an extension of amplitude panning for stereophonic reproduction, and its subjective properties have been evaluated extensively [18, 19, 20]. Although localization is accurate only in a small listening area, the *sweet spot*, VBAP has advantageous properties for practical application, including a low computational complexity, absence of destructive interference in the sweet spot, high timbral quality, e.g., [21], and a gradual degradation of sound quality outside the sweet spot. For these reasons, VBAP is used in numerous current transmission standards and reproduction systems for object-based audio, including reference rendering in the ISO/IEC GMPEG-H 3D Audio standard [22, 23].

An objective comparison between optimization-based physical and perceptually motivated reproduction techniques is hindered by the conceptual gap between these approaches. The design

of the VBAP algorithm, which consists of a geometric criterion to select a set of loudspeakers and a panning law to calculate their amplitude weights, further impedes a comparison to physical approaches. An objective of this paper is to establish a link between VBAP and optimization-based physical reproduction techniques.

Most of the advantageous properties of VBAP can be directly linked to properties of the loudspeaker driving signals, specifically the small number (i.e., one to three) of nonzero amplitude gains for each audio object. This corresponds to a sparse solution. Recent years have seen major advances in *sparsity-promoting* optimization and signal processing techniques, including the Lasso method [24], matching pursuit [25], orthogonal matching pursuit [26], basis pursuit [27], sparse reconstruction using the focal underdetermined system solver (FOCUSS) [28], sparse Bayesian learning [29], and compressed sensing [30, 31]. In particular, the use of the ℓ_1 norm to achieve sparse approximate or exact solutions has led to very efficient algorithms and significant improvements in several signal processing fields.

This paper considers the application of ℓ_1 optimization techniques to amplitude panning for two specific reasons. Firstly, ℓ_1-optimal amplitude panning problems can be efficiently solved using convex optimization methods because the underlying objective function is convex for static loudspeaker configurations. Secondly, in amplitude panning a minimal ℓ_1 norm corresponds to a maximally localized or "sharp" reproduction of an audio object, as described in Sec. 2.4.

Several research publications investigate the application of ℓ_1 minimization and/or compressive sensing to the analysis and reconstruction of spatial sound fields. Epain et al. [32] consider the use of an ℓ_1 minimization for the loudspeaker gains subject to a least-squares constraint on the reproduction gain, where the Lasso method [24] is used to analyze and reproduce sound fields consisting of a small number of plane wave sources. In [33], this technique is extended to time-domain sound field reconstruction. Lilis et al. [34] propose the use of the Lasso operator to sound field control over multiple, spatially distributed sampling points. This technique generates optimized complex-valued loudspeaker gains over a grid of frequencies and enables superior reproduction quality for undersampled sound fields as well as a judicious selection of loudspeaker positions. Koyama et. al. [35] consider sparse decomposition of a sound field within a recording area to achieve wave field reconstruction with reduced aliasing artifacts. Radmanesh et al. [36] propose a two-stage Lasso least-squares method to optimize loudspeaker locations and weightings for multizone reproduction. A method for joint optimization of loudspeaker placement and weights using a constrained matching pursuit approach is described in [37].

In [38], authors of the present paper consider the application of convex optimization techniques to listener-centric sound field control, and demonstrate the similarity between ℓ_1-optimal and amplitude panning methods by means of numerical examples. However, these approaches generally involve a numerical optimization step to calculate the sparse loudspeaker driving functions, hence they are significantly more complex than established techniques such as VBAP.

In contrast, the main contribution of the present paper is to express multi-loudspeaker amplitude panning in the framework of ℓ_1 optimization. More specifically, we use this framework to characterize ℓ_1-optimal solutions of the amplitude panning problem, for instance their exact sparsity and conditions for solution uniqueness. Here we use "exact" to denote sparse solutions that have only a few nonzero values and are exactly zero otherwise [31]. Based on these properties, we show that VBAP is identical to the ℓ_1-optimal solution if three basic requirements are fulfilled: a) the ℓ_1 approach incorporates a nonnegativity constraint on the panning gains, b) the loudspeaker selection of the VBAP algorithm is based on a Delaunay triangulation, and c) this Delaunay triangulation is unique. Most practical VBAP implementations meet these conditions. The results are then generalized to ℓ_1-optimal solutions without the nonnegativity

constraint. In this way we demonstrate that ℓ_1-optimal amplitude panning, with and without nonnegativity constraints for the panning gains, can be computed with basically the same effort as VBAP.

The second main resultof this paper is that the interpretation of amplitude panning as an ℓ_1 optimization problem enables new insight into real-world problems of current VBAP algorithms. For example, it is shown later that asymmetries or ambiguities reported in [39, 40, 22] correspond to nonunique solutions of the ℓ_1 optimization problem. This reveals that they are not implementation problems, but are inherent to the design objective underlying amplitude panning. Although resolving this ambiguity remains an open research question, the present paper provides a full characterization of the set of optimal solutions, which is a valuable starting point to further improve panning algorithms.

The remainder of this paper is outlined as follows. Section 2 reviews amplitude panning techniques for multi-loudspeaker reproduction, in particular VBAP. The proposed idea of expressing amplitude panning as a global ℓ_1 optimization problem is presented in Sec. 3. An additional nonnegativity constraint is introduced in Sec. 4, and conditions for equivalence between this formulation and VBAP are established. Based on this result, Sec. 5 characterizes the ℓ_1 optimal panning solution without this nonnegativity constraint.Different panning methods are evaluated and compared in Sec. 6 using objective and psychoacoustic performance measures, and Sec. 7 summarizes the main outcomes of this paper.

2. Multichannel Amplitude Panning Techniques

This section reviews amplitude panning techniques, their objectives and properties. In particular, it describes VBAP [17], the predominantly used technique for panning virtual sources in three-dimensional loudspeaker setups. This description also establishes the nomenclature used to derive the novel sparse, optimal panning techniques in the subsequent sections.

2.1. Amplitude Panning

Panning is one of the principal and most widely used techniques for spatial sound reproduction. It creates phantom images in the direction of the virtual source by providing auditory cues to a listener within a confined sweet spot [18, 21]. The main auditory cues used in panning are the interaural level difference (ILD) and the interaural time difference (ITD). To this end, the source signal is reproduced over multiple loudspeakers, whereby level differences and/or different time delays are applied to the loudspeaker signals. These techniques are referred to as *amplitude*, *level*, or *intensity panning* and *delay/time-delay panning*, respectively. The computation of the amplitude or delay values is governed by *panning laws* such as the law of sines or the tangent law [15, 16, 20]. While in amplitude panning level differences translate to reliable ITD cues in the frequency range relevant for ITD localization [18, 1], the localization performance of time delay panning is more frequency-dependent [41, 20].

2.2. Spherical Geometry Preliminaries

Throughout this paper, we make extensive use of geometric relations on sphere surfaces to describe 3D amplitude panning techniques as well as ℓ_1-optimal panning approaches. Therefore, here we briefly outline the necessary concepts of spherical geometry and the notation used within this paper. For more detail, the reader is referred to textbooks such as [42].

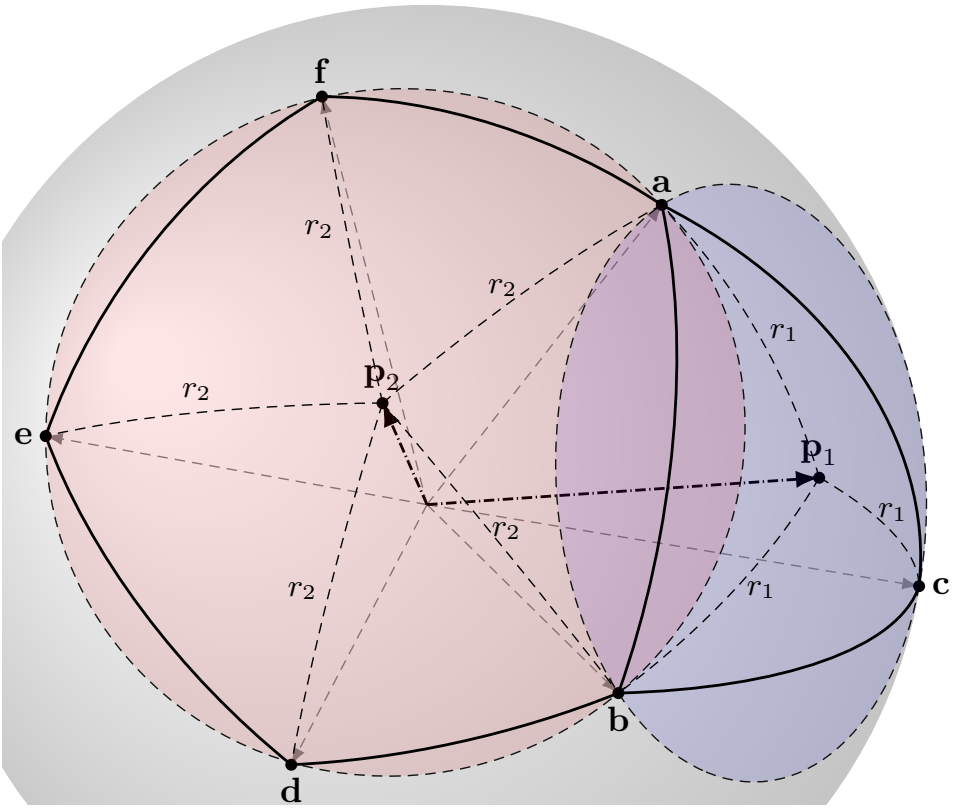

Figure 1: Basic elements of spherical geometry.

Spherical geometry describes geometric relations on the two-dimensional surface of a three-dimensional sphere. Without loss of generality, the radius of the sphere is assumed to be 1. Thus, any 3D unit vector, denoted

$$\mathbf{v} = \begin{bmatrix} x_v & y_v & z_v \end{bmatrix}^T \quad \text{with } \|\mathbf{v}\|_2 = 1\,,$$

corresponds to a *point* on the sphere. In Fig. 1, they are represented as $\mathbf{a}, \mathbf{b}, \ldots$. An *arc* $\overset{\frown}{\mathbf{ab}}$ is a segment of a great circle connecting the points $\mathbf{a}$ and $\mathbf{b}$. On a unit sphere the *arc length*, denoted here as $\sphericalangle(\mathbf{a}, \mathbf{b})$, equals the angle between the vectors $\mathbf{a}$ and $\mathbf{b}$ and is related to their dot product $\langle \mathbf{a}, \mathbf{b} \rangle$ by

$$\sphericalangle(\mathbf{a}, \mathbf{b}) = \cos^{-1} \langle \mathbf{a}, \mathbf{b} \rangle\,. \tag{1}$$

A *spherical polygon* is a connected, closed chain of arcs formed of three or more points. Fig. 1 shows a spherical triangle **abc** and a polygon **abdef** consisting of five points. A circle on the sphere surface that passes through all points of a spherical polygon is denoted as its *circumcircle*, and polygons that have a circumcircle are termed *cyclic polygons*. While all spherical triangles are cyclic, this does not generally hold for polygons consisting of four or more points. The circumcircle of a cyclic polygon is determined by its *circumcenter* $\mathbf{p}_k$, a unit vector, and its *radius* r_k such that the arc length between $\mathbf{p}_k$ and each polygon point equals r_k. Note that there are two points on the opposite sides of the sphere fulfilling this property. Unless stated otherwise, we refer to $\mathbf{p}_k$ corresponding to the smaller radius r_k. Fig. 1 depicts the circumcircles of the cyclic polygons **abc** and **abdef**.

Given a set of points on a surface, a *triangulation* is a subdivision of that surface into triangles formed by edges between these points such that the triangles are not intersecting, e.g., [43]. While often defined for straight-line edges, it is straightforwardly extended to sphere surfaces and spherical triangles. Triangulations form a subset of *tessellations*, i.e., subdivisions of a surface into a set of nonoverlapping geometric shapes. Among the various existing triangulation strategies, the *Delaunay triangulation* is of particular importance for the panning methods considered here. Delaunay triangulations maximize the minimum angle over all triangles. The defining condition for the Delaunay triangulation is the *circumcircle condition*, e.g., [44]:

Definition 1 (Circumcircle condition for Delaunay triangulations) *A triangulation $\mathcal{T}$ is a Delaunay triangulation if and only if no triangle of $\mathcal{T}$ contains any other point within its circumcircle.*

This implies that the Delaunay triangulation is nonunique if there is a circumcircle passing through more than three points with no other points in the interior of the circumcircle. In this paper, we consider triangulations on the unit sphere, i.e., spherical triangles and circumcircles on the sphere surface.

2.3. Vector Base Amplitude Panning (VBAP)

VBAP [17] expresses the tangent law for amplitude panning in a vector formulation and extends it to three-dimensional source directions and 3D loudspeaker setups. In the classification of [1], VBAP is considered as a local panning technique, because it only drives a small number of loudspeakers (at most three) close to the source direction, as opposed to global panning techniques such as Ambisonics amplitude pannning, e.g., [45], which activates loudspeakers all over the setup.

A 3D VBAP configuration is shown in Fig. 2. In VBAP, audio objects are modeled as plane waves. They are represented by the source direction vector $\mathbf{p}$ which is a unit vector pointing to the intended location of the audio object. The loudspeaker locations, represented by unit vectors

$$\mathbf{l}_l = \begin{bmatrix} x_l & y_l & z_l \end{bmatrix}^T \text{ with } \|\mathbf{l}_l\|_2 = 1 \text{ for } l = 1, 2, \ldots, L\,, \tag{2}$$

are assumed to lie on the unit sphere. In case of non-spherical configurations, the vectors $\mathbf{l}_l$ are determined by projecting the actual positions onto the unit sphere, and appropriate gain and delay compensations are applied to the loudspeaker signals. The direction vectors of all L loudspeakers of a setup are compactly represented by the loudspeaker direction matrix

$$\mathbf{L} = \begin{bmatrix} \mathbf{l}_1 & \mathbf{l}_2 & \cdots & \mathbf{l}_L \end{bmatrix} \in \mathbb{R}^{d \times L}\,, \tag{3}$$

where d denotes the dimension of the panning configuration, i.e., $d = 3$ for 3D VBAP.

2.3.1. Panning Gain Calculation

The VBAP method comprises two distinct stages: Firstly, three active loudspeakers, denoted by indices i, j, and k, are selected. To this end, the unit sphere is partitioned into a set of nonoverlapping spherical triangles whose vertices are formed by the loudspeaker direction vectors. Then the active loudspeakers are chosen such that the corresponding spherical triangle contains the source $\mathbf{p}$. These steps are described in more detail in Sec. 2.3.2 and 2.3.3. In the

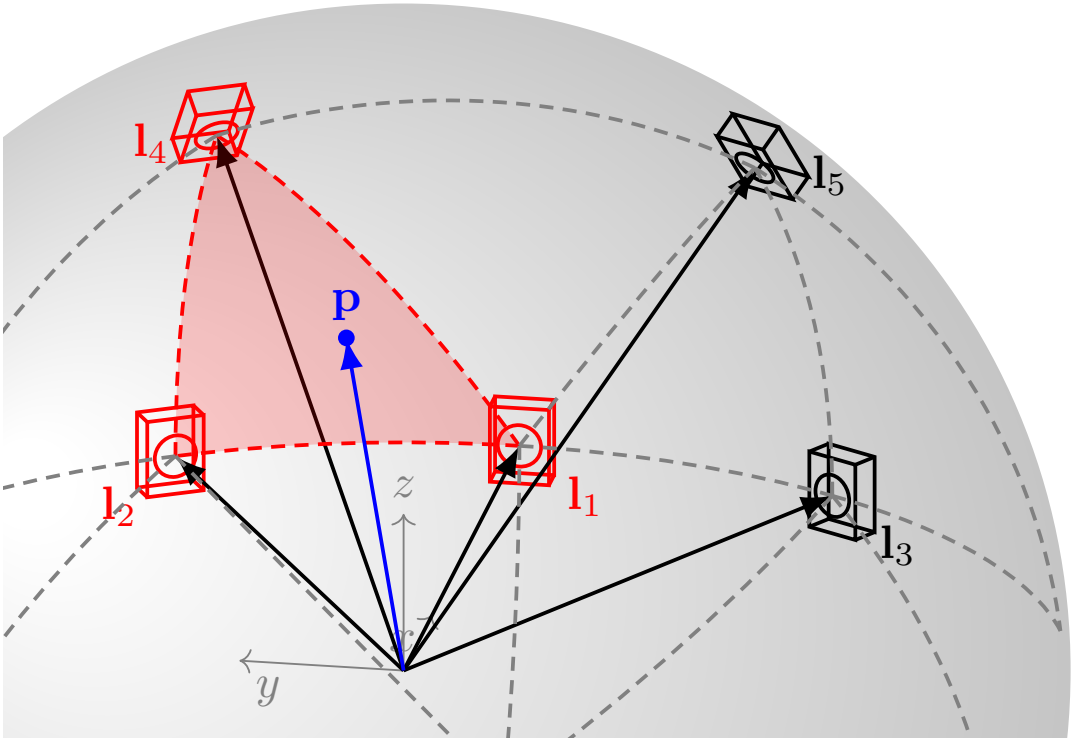

Figure 2: Exemplary 3D amplitude panning configuration. The active loudspeaker triangle selected by VBAP is marked in red.

second stage, the panning gain vector $\mathbf{g}_{ijk} = \begin{bmatrix} g_i & g_j & g_k \end{bmatrix}^T$ for the three active loudspeakers $\mathbf{l}_i$, $\mathbf{l}_j$, and $\mathbf{l}_k$ is obtained as

$$\mathbf{g}_{ijk} = \mathbf{L}_{ijk}^{-1}\mathbf{p} \qquad \text{with } \mathbf{L}_{ijk} = \begin{bmatrix} \mathbf{l}_i & \mathbf{l}_j & \mathbf{l}_k \end{bmatrix}, \tag{4}$$

where the matrix $\mathbf{L}_{ijk}$ is formed of the loudspeaker direction matrix $\mathbf{L}$ defined in (3) by selecting the columns corresponding to the loudspeaker indices i, j, and k. The global gain vector $\mathbf{g} \in \mathbb{R}^{L\times 1}$ for the complete loudspeaker setup contains the weights g_i, g_j, g_k at the indices i, j, and k, respectively, and zeros otherwise.

Eq. (4) implies that the panning weights are determined such that the weighted sum of the active loudspeaker's direction vector matches the source direction vector

$$\mathbf{p} = \sum_{l\in\{ijk\}} g_l \mathbf{l}_l \,. \tag{5}$$

2.3.2. Triangulation

As described above, the selection of the active loudspeakers is based on a triangulation of the unit sphere into spherical triangles formed by the direction vectors $\mathbf{l}_l$. In the original description of VBAP [17], triangulation is performed manually based on empirical criteria: a) the triangles should not intersect and b) they should be selected such that the localization accuracy in every direction is maximized. The latter objective can be interpreted as minimizing the size of the individual triangles. An automated algorithm to generate such triangulations is proposed in

[46]. As described in [20], this algorithm aims at minimizing the length of the triangle edges, although it does not specify whether this refers to the length of individual edges or the sum of edge lengths. It also states that this algorithm is similar to a greedy triangulation, e.g., [47], which is an approximation of a minimum-weight triangulation that minimizes the sum of the lengths of the triangle edges [43]. Current VBAP implementations, for instance [45, 40, 22] typically use a Delaunay triangulation, e.g., [43], owing to its properties (see Sec. 2.2) and the availability of efficient algorithms and implementations such as Quickhull [48]. The Delaunay triangulation is another approximation of the minimum-weight triangulation. Its property of maximizing the minimum angle over all triangles effectively prevents triangles with long sides and acute angles.

If the Delaunay triangulation is nonunique, i.e., if it contains a cyclic polygon with more than three loudspeakers, standard VBAP implementations select an arbitrary valid triangulation. As reported in [39, 40, 22], this may lead to artifacts as reproduction asymmetries or uneven virtual source movements.

2.3.3. Loudspeaker Selection

The active loudspeaker triangle $\{i, j, k\}$ is selected such that the source position $\mathbf{p}$ lies within the spherical triangle spanned by $\mathbf{l}_i$, $\mathbf{l}_j$, and $\mathbf{l}_k$. This corresponds to the triangle for which $\mathbf{p}$ in (5) can be formed as a *conical combination* of the active loudspeakers $\mathbf{l}_l$

$$\mathbf{p} = \sum_{l \in \{ijk\}} g_l \mathbf{l}_l \qquad \text{with} \quad g_l \geq 0\,, \tag{6}$$

that is, a linear combination with nonnegative gains g_l. In practical implementations, the selection of the active triangle is performed by evaluating the unnormalized panning gains according to (4) for all triangles of the triangulation, and selecting the triangle that fulfills the nonnegativity condition (6), e.g., [17, 46]. That is, the computational effort is determined by a solution of a 3×3 linear system for each triangle of the setup. Thus the complexity of the VBAP gain calculation is linear with respect to the number of triangles, which is proportional to the number of loudspeakers, i.e., $O(L)$.

Fig. 2 depicts the active loudspeaker triangle $\{\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_4\}$ for an exemplary source position $\mathbf{p}$. For audio object positions strictly in the interior of a triangle, this criterion is unambiguous if the triangles of the triangulation are not overlapping. Special cases occur if an object $\mathbf{p}$ lies on a triangle edge or coincides with a loudspeaker location. In these cases, only two or one loudspeakers are active, respectively. This also means that either the two triangles sharing this edge or all triangles containing the given loudspeaker meet the selection criterion (6). However, this ambiguity is not critical, because the gains calculated for all triangles fulfilling (6) are identical.

2.3.4. Gain Normalization

To maintain a constant loudness at the listener position independent of the source direction, the panning gains $\mathbf{g}$ are normalized [17, 49]

$$\mathbf{g}' = \frac{1}{\|\mathbf{g}\|_p} \mathbf{g} \tag{7}$$

to obtain the final panning weights $\mathbf{g}'$. Here $\|\cdot\|_p$ represents the ℓ_p norm $\|\mathbf{x}\|_p = \sqrt[p]{\sum |x_i|^p}$. The present paper uses the ℓ_2 norm, i.e., power normalization in accordance with [17]. Because

normalization is applied uniformly to the complete gain vector $\mathbf{g}$, it does not change the properties of the panning solution apart from the total sound pressure.

2.4. Properties of VBAP

As the psychoacoustic attributes of VBAP are extensively described, e.g., in [17, 19, 20], we focus on the objective properties and their links to perceptive features.

2.4.1. Preservation of Velocity Direction

The source direction $\mathbf{p}$ and the loudspeaker direction vectors $\mathbf{l}_l$ are proportional to the particle velocity vectors of the source object and the loudspeaker wave fronts, respectively. Thus, (5) implies that VBAP synthesizes the correct particle velocity direction by a weighted superposition of the particle velocity vectors of the active loudspeakers. This provides a good localization at low frequencies (up to $\approx 700\,\mathrm{Hz}$), e.g., [16, 50, 51].

2.4.2. Locality

By construction, the loudspeaker triangulation ensures that only loudspeakers close to the source direction $\mathbf{p}$ are active. This ensures a graceful degradation of directional quality for listener outside the reference position [20].

2.4.3. Sparsity

By construction, VBAP uses the minimal number of nonzero panning gains to correctly synthesize the velocity direction of the virtual source $\mathbf{p}$, that is, at most three for a general position in a 3D setup, and two or one loudspeakers if $\mathbf{p}$ coincides with a triangle edge of loudspeaker direction, respectively. Combined with locality, this implies that a sound source is reproduced as "sharp", i.e., with minimal directional spread, as possible with the given loudspeaker configuration, as discussed in [17]. For low frequencies, this spread is quantified by the *velocity vector magnitude*, also termed *velocity magnitude* [16, 50] or *velocity factor* [51]

$$r_\mathrm{v} = \frac{1}{\sum_{l=1}^{L} g_l} \tag{8}$$

if (6) is met. For nonnegative panning gains, r_v is also the reciprocal of $\|\mathbf{g}\|_1$, thus establishing a relation between the properties of $\mathbf{g}$ and the source spread. Large values of r_v correspond to sharp image localization, whereas low values of r_v yield less localized, spatially spread sound images. Thus, the spread of a virtual sound source depends on its position relative to the loudspeakers, leading to a nonuniform spread distribution [52, 45].

2.4.4. Nonnegativity

Because VBAP synthesizes the desired source direction as a conical representation of loudspeaker direction vectors (6), all loudspeaker gains are nonnegative. As described in Sec. 2.3.3, most VBAP implementations utilize this nonnegativity property to select the active triangle. The nonnegativity constraint has also a positive impact on the perceived quality, as it avoids anti-phase signals resulting in destructive interference at the listener position, which degrades spatial fidelity. For the same reasons, nonnegative panning gains are enforced by *in-phase* Ambisonic decoders [51, 7].

3. ℓ_1-Optimal Amplitude Panning

In the preceding section we characterized VBAP as a practical approach to multichannel sound reproduction. In particular, VBAP is a combination of a geometric approach to determine the triangle of active loudspeakers and an algebraic solution to compute the panning gains such that the velocity vector of the synthesized sound field matches the direction of the virtual source. In the following we consider amplitude panning as a global optimization problem to generate panning gain vectors $\mathbf{g}$ without resorting to an intermediate loudspeaker selection step. Nonetheless, we aim to retain the advantages of amplitude panning techniques as VBAP, such as a correct particle velocity direction at the listener position, minimal spread of the source image, and a small number of active loudspeakers close to the source direction.

3.1. The ℓ_1 Optimization Problem

Because of the sparsity-promoting nature of the ℓ_1 norm, see e.g., [31], we formulate the multi-loudspeaker amplitude panning problem as an ℓ_1 optimization problem

$$\underset{\mathbf{g}}{\operatorname{argmin}} \ \|\mathbf{g}\|_1 \tag{9a}$$

$$\text{subject to } \mathbf{L}\mathbf{g} = \mathbf{p}\,, \tag{9b}$$

which is an equality-constrained convex optimization problem (e.g., [53]). Here, the equality constraint (9b) ensures the desired particle velocity direction analogous to (4) but applied to the complete loudspeaker direction matrix $\mathbf{L}$, while (9a) ensures sparsity of the panning gain vector. Alternatively, the objective (9a) can be considered as to maximize the velocity vector magnitude (8), which can be interpreted as creating the sharpest possible sound image at low frequencies.

Problem (9) follows from the Basis Pursuit optimization principle proposed in [27]. It can also be considered as a limiting case of the Lasso method [24, 31]

$$\underset{\mathbf{g}}{\operatorname{argmin}} \|\mathbf{g}\|_1 \qquad \text{subject to } \|\mathbf{L}\mathbf{g} - \mathbf{p}\|_2 \leq \epsilon \tag{10}$$

for $\lim \epsilon \to 0$, which can always be met in case of an under-determined problem such as amplitude panning.

3.2. Characterization of the ℓ_1-Optimal Solution

The framework of ℓ_1 minimization and compressive sampling enables us to describe the solution of the optimization problem (9), in particular its uniqueness and sparsity properties, e.g, [54, 55, 56, 30, 57]. Here we focus on a recent result that establishes necessary and sufficient conditions for the solution uniqueness of ℓ_1 minimization problems:

Theorem 1 (ℓ_1 uniqueness [58]) *Let $\mathbf{g}^*$ denote a solution to (9). Also, let $\mathbf{I}$ denote the index set of the nonzero elements of $\mathbf{g}^*$, and $\mathbf{s} = \operatorname{sgn}(\mathbf{g}^*_\mathbf{I})$ the signs of the nonzero elements of $\mathbf{g}^*$. Then $\mathbf{g}^*$ is the unique solution if and only if the following conditions hold:*

- *The submatrix $\mathbf{L}_\mathbf{I}$ containing the columns corresponding to the index set $\mathbf{I}$ has full column rank.*

- *There exists a vector* $\mathbf{y} \in \mathbb{R}^d$ *such that* $\mathbf{L}_\mathbf{I}^T \mathbf{y} = \mathbf{s}$ *and*

$$\left\| \mathbf{L}_{\mathbf{I}^c}^T \mathbf{y} \right\|_\infty < 1 \,, \tag{11}$$

  where $\mathbf{L}_{\mathbf{I}^c}$ *denotes the submatrix containing the columns of* $\mathbf{L}$ *corresponding to the zero entries of* $\mathbf{g}$.

Several properties of ℓ_1-optimal amplitude panning solutions follow directly from Theorem 1. The column rank of the matrix $\mathbf{L}_\mathbf{I}$, corresponding to the number of active, linearly independent loudspeakers, cannot exceed the maximum row rank d of this matrix because $\mathbf{L} \in \mathbb{R}^{d \times L}$. That is, if a unique solution exists, it contains at most d nonzero weights, i.e., at most three active loudspeakers for a 3D setup. Therefore, the global ℓ_1-optimal panning solution preserves the sparsity properties of VBAP. While it is possible to completely characterize the solution of (9), including the selection of active loudspeakers, in terms of Theorem 1 and the proofs in [58], we choose a different, more intuitive approach here that is based on the features of amplitude panning. To this end, the following section considers problem (9) with an additional nonnegativity constraint on the panning gains. After describing the optimal solution of this restricted problem, a generalization to unconstrained panning gains is established in Sec. 5.

4. ℓ_1^+: ℓ_1-Optimal Panning with Nonnegative Gains

As discussed in Sec. 2.4.4, the limitation to nonnegative panning gains is an important feature of VBAP, but also other sound reproduction techniques. Adding this constraint to the ℓ_1-optimal panning problem (9) leads to the following convex optimization problem

$$\underset{\mathbf{g}}{\operatorname{argmin}} \ \|\mathbf{g}\|_1 \tag{12a}$$

$$\text{subject to } \mathbf{L}\mathbf{g} = \mathbf{p} \tag{12b}$$

$$\mathbf{g} \geq 0 \,. \tag{12c}$$

This is referred to as ℓ_1^+-*optimal amplitude panning* in the following. In this section, we use the framework of linear programming (LP) to study the solution of this problem, first in terms of the original (primal) LP and later, in Sec. 4.2, using the corresponding dual LP. Based on these results, the equivalence between VBAP and the ℓ_1^+ problem is proven in Sec. 4.3.

4.1. Representation as a Linear Program

LP is a widely used framework for modeling and solving optimization problems with a linear objective function subject to linear equality and inequality constraints, see, e.g., [59, 60, 61]. With this framework, problem (12) can be expressed as an LP in the so-called *standard form* as follows

$$\underset{\mathbf{g}}{\operatorname{argmin}} \ \mathbf{c}^T \mathbf{g} \tag{13a}$$

$$\text{subject to } \mathbf{L}\mathbf{g} = \mathbf{p} \tag{13b}$$

$$\mathbf{g} \geq 0 \,, \tag{13c}$$

where $\mathbf{c} = [1, \cdots, 1]^T = \mathbf{1}_{L \times 1}$ is a column vector of ones. In this way, the objective function reduces to the sum of the elements of $\mathbf{g}$, which is equivalent to the ℓ_1 norm due to the

nonnegativity condition. Because $\sum_{l=1}^{L} g_l$ represents the sound pressure at the listener position, (13) can be interpreted as minimizing the sound pressure while synthesizing a desired particle velocity vector $\mathbf{p}$. It is worth noting that the representation of (12) as (13) differs from the standard transformation of an ℓ_1 optimization problem into an LP, e.g., [27, 58]. While the latter essentially doubles the number of variables and constraints, (13) has the same dimensions as (12).

In the following, we use basic concepts of the LP framework to interpret the solutions of the nonnegative panning problem.

4.1.1. Existence of the Solution

A vector $\mathbf{g}$ is a *feasible*, i.e., valid, solution of problem (13) if all equality constraints (13b) and inequality constraints (13c) are satisfied. A problem is feasible if at least one feasible solution exists. The *optimal value* is the minimum value of the objective function (13a) over all feasible solutions. The set of feasible solutions for which the objective function attains the optimal value forms the set of optimal solutions of (13), whose elements are denoted as $\mathbf{g}^*$. If the minimum and maximum objective values over the set of feasible solutions are finite, the problem is *bounded below* or *bounded above*, respectively. Obviously, the nonnegativity constraint (13c) establishes a trivial minimum lower bound $\mathbf{c}^T\mathbf{g} \geq 0$ for the panning problem (13).

For general LP problems, the decision whether it is feasible and bounded has a complexity comparable to the solution of the LP itself (e.g., [61]). Thus, no general rules for solution existence can be deduced from (13). However, as shown in Sec. 4.2.3, feasibility conditions for the panning problem can be established by using the dual linear program.

4.1.2. Vertex Solutions and Number of Nonzero Panning Gains

The number of active loudspeakers of the optimal panning solution can be directly linked to the property of the LP. *Vertex solutions* (e.g., [61]), also termed *vertices*, *basic solutions* [59], or *basic feasible solutions* [60] are a basic concept in the LP framework. A vertex solution is a feasible solution $\mathbf{g}$ for which at least L linearly independent constraints are active. Each row of the vector-valued inequality constraint (13c) for which the "$\geq$" relation holds with equality "$=$" forms an active constraint. For problem (13), this number is identical to the number of zero-valued gains g_l. In case of equality constraints, each row of the matrix (13b) represents an active constraint, resulting in d active equality constraints in case of problem (13). Therefore, a vertex solution of (13) contains at least $L - d$ zeros, i.e., $\mathbf{g}$ has at most d active loudspeakers, for instance $d = 3$ for a 3D setup.

A vertex solution is termed *nondegenerate* if there are exactly L active constraints, and degenerate if more than L constraints are active, that is, less than three active speakers. This distinction bears a close resemblance to 3D VBAP, where a solution contains either $d = 3$ or fewer active loudspeakers (see Sec. 2.4.3).

4.1.3. Optimal Solution Set

The fundamental theorem of linear programming, e.g., [60], establishes a relation between vertices and optimal solutions:

Theorem 2 (Fundamental theorem of linear programming) *If a LP is bounded and feasible, it has an optimal solution. In this case, it has at least one vertex solution. Furthermore, the optimal value is attained at at least one vertex solution.*

Such a vertex is termed an *optimal vertex*. That is, a feasible LP has either one or multiple optimal vertices. In the former case, the optimal panning problem is unique, and the corresponding gain vector $\mathbf{g}^*$ has at most d nonzero entries. In the latter case, there are multiple optimal vertices $\{\mathbf{g}_1^*, \ldots, \mathbf{g}_S^*\}$, and the set of optimal panning gain vectors consists of all convex combinations of these vectors, which is a corollary of Theorem 2, e.g., [61]

$$\mathbf{g}^* = \sum_{s=1}^{S} \alpha_s \mathbf{g}_s^* \quad \text{with} \quad \sum_{s=1}^{S} \alpha_s = 1 \text{ and } \alpha_s \geq 0\,. \tag{14}$$

In this case, an optimal solution can have more than d nonzero gains. Applied to 3D amplitude panning, this implies that if a valid solution exists, there is an optimal solution with at most three active loudspeakers. Optimal solutions with more than three active loudspeakers exist only if the LP is nonunique.

4.2. Solution Properties Based on the Dual LP Problem

Further insight into the ℓ_1 optimal panning problem can be gained by considering the dual LP of (13) [59, 60, 61, 53]. For the LP (13), termed the primal problem, the dual program is

$$\underset{\pi}{\text{argmax}}\ \mathbf{p}^T \boldsymbol{\pi} \tag{15a}$$

$$\text{subject to } \mathbf{L}^T \boldsymbol{\pi} \leq \mathbf{c}\,, \tag{15b}$$

where $\boldsymbol{\pi} \in \mathbb{R}^d$ is the dual solution. The optimal (maximum) value of the objective function (15a) is identical to the optimal value of the primal problem, i.e.,

$$\mathbf{p}^T \boldsymbol{\pi}^* = \mathbf{c}^T \mathbf{g}^* = \|\mathbf{g}^*\|_1\,, \tag{16}$$

where $\boldsymbol{\pi}^*$ and $\mathbf{g}^*$ denote optimal solutions of the dual and primal problem, respectively. The elements of the dual solution $\boldsymbol{\pi}$ relate to the Lagrange multipliers of the inequality constraint (13c) of the primal problem [61].

There are numerous connections between the primal and the corresponding dual problem, see, e.g, [60, 61]. Here we introduce two relations that are used throughout this paper.

Theorem 3 (Dual Degeneracy and Primal Nonuniqueness) *If the optimal solution of the dual is degenerate, then the solution of the primal problem is nonunique, provided that the primal is nondegenerate.*

Theorem 4 (Primal Feasibility and Dual Boundedness) *If the primal problem is feasible, the dual is feasible if and only if the primal is bounded.*

LP duality is a symmetric relation, that is, the dual of a dual problem is the primal. In this way, the same properties can be inferred from the dual to the primal.

4.2.1. Geometric Interpretation of the Dual Solution $\boldsymbol{\pi}$

Without loss of generality, the dual solution vector $\boldsymbol{\pi}$ can be separated into a unit vector $\mathbf{p}_\pi$ and a nonnegative factor c_π provided that $\boldsymbol{\pi} \neq \mathbf{0}$, namely

$$\boldsymbol{\pi} = c_\pi \mathbf{p}_\pi \qquad \text{with } \|\mathbf{p}_\pi\|_2 = 1\,,\, c_\pi > 0\,. \tag{17}$$

Dividing (15) by c_π, the dual problem can be expressed as

$$\underset{\mathbf{p}_\pi}{\operatorname{argmax}}\ \mathbf{p}^T \mathbf{p}_\pi \tag{18a}$$

$$\text{subject to } \mathbf{L}^T \mathbf{p}_\pi \leq \frac{1}{c_\pi}\mathbf{c}\,. \tag{18b}$$

Each row i of the inequality constraint (18b) is a dot product (1) of the unit vectors $\mathbf{l}_i$ and $\mathbf{p}_\pi$

$$\mathbf{l}_i^T \mathbf{p}_\pi = \langle \mathbf{l}_i, \mathbf{p}_\pi \rangle = \cos \sphericalangle (\mathbf{l}_i, \mathbf{p}_\pi) \leq \frac{1}{c_\pi}\,, \tag{19}$$

which can be interpreted as a minimum angle constraint

$$\sphericalangle (\mathbf{l}_i, \mathbf{p}_\pi) \geq r_\pi \qquad \text{with } r_\pi = \cos^{-1} \frac{1}{c_\pi}\,. \tag{20}$$

Thus, condition (20) corresponds to a circle on the unit sphere with center (or axis [42]) $\mathbf{p}_\pi$ and radius r_π, such that there are no loudspeakers within the surface area enclosed by the circle, but potentially on its boundary.

Here we use "radius" to denote the angular distance from $\mathbf{p}_\pi$ to a point on the circle, which is identical to the angle between the corresponding direction vectors in case of a unit sphere.

4.2.2. Vertex Solutions

According to Theorem 2, the optimal value of an LP is attained at at least one vertex. As the solution vector $\boldsymbol{\pi}$ of the dual problem has d components, i.e., $d = 3$ for 3D setups, the active constraint matrix $\mathbf{L}_\mathbf{I}^T$ must have d linearly independent rows for an optimal solution. It is readily verified that the active constraint matrix $\mathbf{L}_\mathbf{I}^T$ attains the maximum column rank d for every possible combination of loudspeaker vectors $\mathbf{l}_i$, $i \in \mathbf{I}$. Matrix $\mathbf{L}_\mathbf{I}^T$ can be rank-deficient only if at least $d = 3$ loudspeaker vectors lie on a common plane. As all vectors $\mathbf{l}_i$, $i \in \mathbf{I}$ lie on a common circle, these vectors span a space with a dimension lower than d only if at least two direction vectors coincide. Consequently, a vertex solution of the dual LP defines a circle with center $\mathbf{p}_\pi$ and radius $r_\pi = \cos^{-1}(1/c_\pi)$ such that there are no loudspeakers inside the circle and at least three loudspeakers on the boundary.

Only loudspeakers $\mathbf{l}_i$ on the circumcircle correspond to nonzero gains g_i in the primal problem. This follows from the condition of complementary slackness, e.g., [61], which states that a solution variable can be nonzero only if the Lagrange multiplier corresponding to the respective element of the dual solution is zero, i.e., if the corresponding inequality is active.

4.2.3. Solution Existence

The dual problem enables a direct geometric interpretation of the existence of a panning solution in form of an angle limit

$$r_\pi < \tfrac{\pi}{2}\,. \tag{21}$$

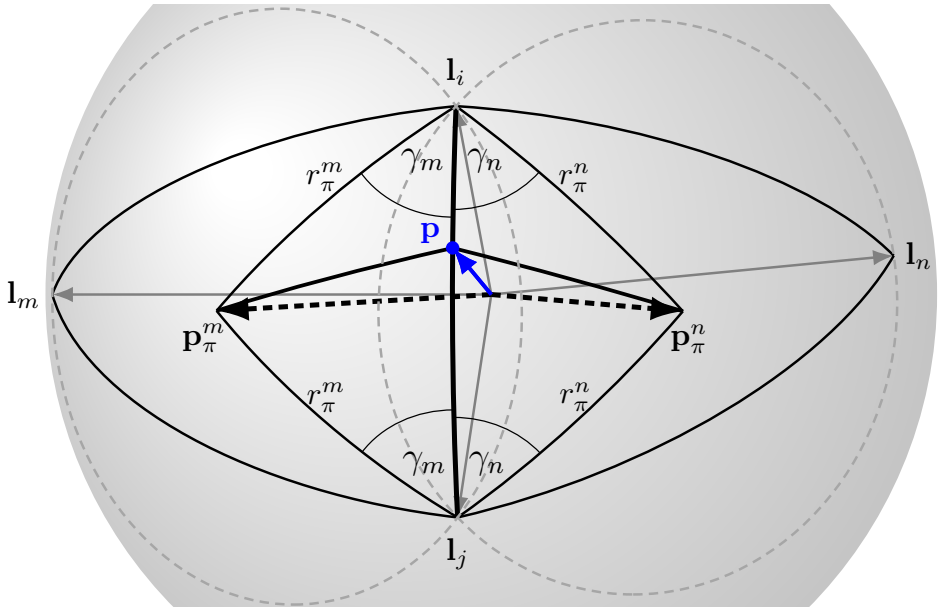

Figure 3: Construction of the objective of the dual LP for a source direction $\mathbf{p}$ on an arc between two loudspeaker directions.

Assume a loudspeaker setup such that $r_\pi \geq \pi/2$ for some center $\mathbf{p}_{\boldsymbol{\pi}}$ and a source position $\mathbf{p}$ such that $\sphericalangle(\mathbf{p}, \mathbf{p}_\pi) < \pi/2$. In this case, all dot products $\langle \mathbf{l}_l, \boldsymbol{\pi} \rangle$ of the inequality constraint (15b) are negative, and therefore these constraints are never active. This means that the dual solution $\boldsymbol{\pi}$ and thus the objective value $\mathbf{p}^T\boldsymbol{\pi}$ can be made arbitrarily large without violating these constraints. Thus, the dual problem is unbounded, and Theorem 4 implies that the corresponding primal problem is infeasible. This means that if the 3D setup contains a zone such that the minimum loudspeaker distance from a central point is greater or equal to $\pi/2$, then there is no nonnegative ℓ_1 panning solution for virtual sources in this zone. In a way, this provides a quantitive interpretation to the qualitative statement for VBAP [45] which states that the loudspeaker aperture should not exceed roughly 90°.

4.3. Equivalence to VBAP

In the previous section we described the optimal dual solution and related it to geometrical conditions on the unit sphere. In the following we extend this to a full geometrical characterization of the ℓ_1^+-optimal solution and derive conditions under which this solution is equivalent to VBAP.

4.3.1. Delaunay Tessellation Imposed by Dual Vertex Solutions

In a first step we demonstrate that the dual vertex solutions correspond to a Delaunay triangulation, or, more general, a tessellation of the sphere surface. As shown in Sec. 4.2.2, a vertex $\boldsymbol{\pi}$ can be interpreted as a circle on the unit sphere surface with center $\mathbf{p}_{\boldsymbol{\pi}}$ and radius $r_{\boldsymbol{\pi}}$ such that there are no loudspeaker vectors within in the interior of the circle and at least d loudspeakers on the boundary. This condition is equivalent with the circumcircle condition of the Delaunay triangulation (Definition 1), with the exception that it allows general cyclic

polygons instead of only triangles. Thus, this construction can be regarded as the tessellation conforming to the Delaunay circumcircle condition. If all cyclic polygons have exactly $d = 3$ points, this tessellation is identical to the unique Delaunay triangulation for this setup. In case of cyclic polygons consisting of more than three loudspeakers, a Delaunay triangulation can be constructed by adding nonintersecting arcs between loudspeaker vectors on the circumcircle. As remarked in Sec. 2.2, this renders the Delaunay triangulation nonunique. In this way, the dual vertex solutions partition the unit sphere surface into a finite number of cyclic polygons, each associated with a center $\mathbf{p}_{\boldsymbol{\pi}}$. This partitioning depends only on the loudspeaker configuration $\mathbf{L}$, but not on the source position $\mathbf{p}$.

4.3.2. Optimal Dual Vertex Solution

In a second step we show that the optimal vertex solution of the dual problem is attained when the source position $\mathbf{p}$ is located within the cyclic polygon corresponding to this vertex. The objective function of the dual LP (15a), which is to be maximized, can be expressed using (17) and (20) as

$$\mathbf{p}^T \boldsymbol{\pi} = \frac{\cos \sphericalangle (\mathbf{p}, \mathbf{p}_\pi)}{\cos r_\pi} \,. \tag{22}$$

In order to find the global maximum of this function we consider two vertex solutions, represented by center vectors $\mathbf{p}_{\boldsymbol{\pi}}^m$ and $\mathbf{p}_{\boldsymbol{\pi}}^n$ such that corresponding cyclic polygons share a common arc $\widehat{\mathbf{l}_i \mathbf{l}_j}$. This configuration is depicted in Fig. 3. For a source position $\mathbf{p}$ on the spherical arc $\widehat{\mathbf{l}_i \mathbf{l}_j}$ the value of the objective function for a vertex $\boldsymbol{\pi}^k$, $k \in \{m, n\}$ is

$$\mathbf{p}^T \boldsymbol{\pi}^k = \cos \sphericalangle (\mathbf{l}_i, \mathbf{p}) - \frac{\sin \sphericalangle (\mathbf{l}_i, \mathbf{p})}{\sin \sphericalangle (\mathbf{l}_i, \mathbf{l}_j)} \left[\cos \sphericalangle (\mathbf{l}_i, \mathbf{l}_j) - 1\right] \,, \tag{23}$$

as derived in Appendix A. It is apparent that the objective function is independent of the chosen vertex $\boldsymbol{\pi}^k$. Consequently, the vertex solutions have identical objective values for sources on the arc $\widehat{\mathbf{l}_i \mathbf{l}_j}$.

Next we consider a source position $\mathbf{p}$ not on $\widehat{\mathbf{l}_i \mathbf{l}_j}$. For a given vertex $\boldsymbol{\pi}^k$, the dual objective value (22) decreases monotonically with increasing distance between $\mathbf{p}$ and the circumcenter $\mathbf{p}_\pi^k$. Combined with (23), this implies that the objective value for a source position $\mathbf{p}$ is larger for a vertex $\boldsymbol{\pi}^k$ that lies on the same side of the arc $\widehat{\mathbf{l}_i \mathbf{l}_j}$ as $\mathbf{p}$. Applying this argument to all arcs of the cyclic polygon enclosing $\mathbf{p}$, it follows that the objective function reaches its optimum for the vertex $\boldsymbol{\pi}^*$ if $\mathbf{p}$ lies inside the cyclic polygon defined by circumcenter $\mathbf{p}_{\pi^*}$ and radius r_{π^*}. Thus, the selection of active loudspeakers is identical to VBAP except that the ℓ_1^+ method facilitates not only triangles, but also cyclic polygons with more than three loudspeakers. This case is discussed in Sec. 4.3.5 below.

As described in Sec. 2, VBAP solutions can be distinguished into three cases. In the following we characterize these cases in terms of the corresponding dual LP to show their equivalence to ℓ_1^+-optimal panning.

4.3.3. Unique Panning Solutions With Three Active Loudspeakers

The VBAP panning weights are unique if the Delaunay triangulation around the source direction $\mathbf{p}$ is unambiguous, i.e., no circumcircle contains more than three loudspeakers. Applied to the dual LP, this means that there are exactly $d = 3$ active constraints corresponding to the

same active loudspeakers as for VBAP. That is, the dual LP is nondegenerate. Consequently, Theorem 3 implies that the primal LP has a unique solution. Thus, the equality constraint (12b) reduces to the same uniquely solvable linear system (4) as for VBAP. This confirms the equivalence of both methods for this case.

4.3.4. Panning Solutions with Less Than Three Active Loudspeakers

As shown in Sec. 2.3.3, VBAP uses only one or two active loudspeakers if the source direction $\mathbf{p}$ coincides with a loudspeaker position or lies on an arc of the triangulation, respectively. Sec. 4.1.2 explained that such cases correspond to degenerate vertices of the primal LP. Theorem (3) implies, by interchanging the role of primal and dual LP, that the corresponding dual LP is nonunique. This case is depicted in Fig. 3, where the source position $\mathbf{p}$ lies on the spherical arc between the two loudspeakers $\mathbf{l}_i$ and $\mathbf{l}_j$. Thus, both $\boldsymbol{\pi}^m$ and $\boldsymbol{\pi}^n$, corresponding to the loudspeaker-free triangles $\{\mathbf{l}_i, \mathbf{l}_j, \mathbf{l}_m\}$ and $\{\mathbf{l}_i, \mathbf{l}_j, \mathbf{l}_n\}$ with cirumcenters $\mathbf{p}_\pi^m$ and $\mathbf{p}_\pi^m$, respectively, are vertex solutions of the dual LP. According to (23), in this case the objective value of a vertex solution depends neither on $\mathbf{p}_\pi^k$ nor on r_π^k, and thus the objective values of the two vertex solutions $\boldsymbol{\pi}^m$ and $\boldsymbol{\pi}^n$ are identical. Theorem 2 confirms that these vertices are the optimal vertex solutions of the dual LP. It also implies that the set of optimal solutions of the dual LP consists of all convex combinations of $\boldsymbol{\pi}^m$ and $\boldsymbol{\pi}^n$.

This characterization is straightforwardly extended to cases where the source direction $\mathbf{p}$ coincides with a loudspeaker direction vector $\mathbf{l}_i$. In this case, all cyclic polygons that contain the loudspeaker $\mathbf{l}_i$ on its boundary are vertex solutions of the dual LP. As the distance between the circumcenter $\mathbf{p}_\pi^k$ of this polygon to $\mathbf{l}_i$ equals the radius r_π^k of this polygon, (22) implies that all these dual vertex solutions have the same objective value 1. Thus they are identical and therefore all vertices of this set are optimal vertex solutions of the dual LP. The vertex solutions of the dual LP correspond to the multiple valid VBAP triangle selections if the source direction lies on a loudspeaker or an arc connecting loudspeakers, i.e., they completely contain the VBAP solution. Furthermore, this implies that all these solutions have the same objective value.

4.3.5. Nonunique Panning Solutions

As described in Sec. 2.3.2, the VBAP panning gains are nonunique if the underlying triangulation is ambiguous. In case of the Delaunay triangulation, this corresponds to configurations with more than three loudspeakers on a common circumcircle. In the LP framework, this implies that more than three inequality constraints (15b) are active for the optimal solution $\boldsymbol{\pi}^*$ of the dual LP (15). Thus, the optimal solution of the dual is degenerate. As reasoned above, this implies that the ℓ_1^+ panning problem, i.e., the corresponding primal LP, is nonunique provided that the primal is nondegenerate. The latter condition holds because a degenerate optimal vertex solution of the primal would mean that less than $d = 3$ loudspeakers were active, i.e., that the source direction is a linear combination of two or less loudspeaker directions. This case has already been handled in the preceding section.

Theorem 2 ensures that there is at least one optimal vertex solution. As reasoned above, these vertex solutions are nondegenerate. At the same time, the nonuniqueness property implies that there are multiple optimal vertex solutions. These are denoted as $\mathbf{g}_1^*$, $\mathbf{g}_2^*$,... $\mathbf{g}_S^*$, and each of these S solutions $\mathbf{g}_s^*$ has exactly $d = 3$ nonzero elements. All optimal vertex solutions $\mathbf{g}_s^*$ of the primal have the same (degenerate) dual solution $\boldsymbol{\pi}^*$. Consequently, the optimal vertex solutions are formed by all subsets of $d = 3$ loudspeakers on the circumcircle that attain the optimal

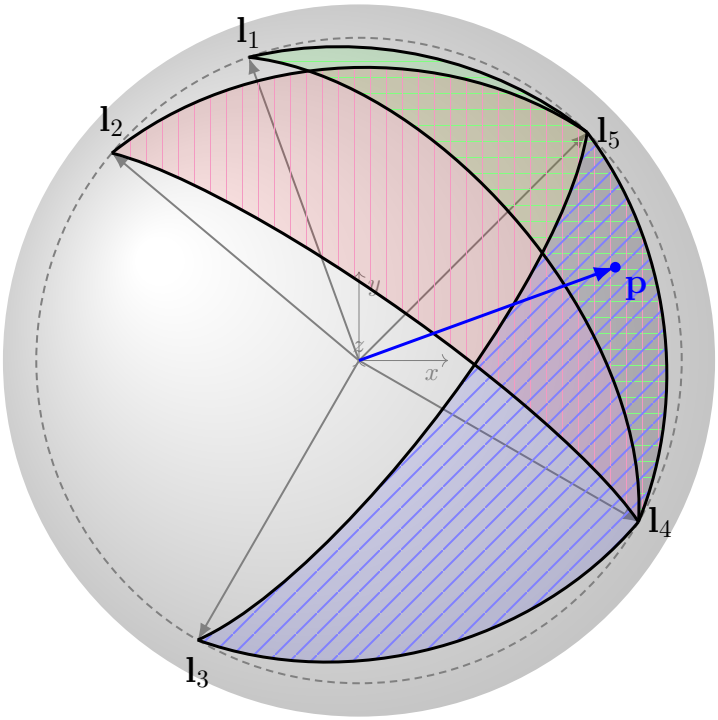

Figure 4: Nonunique ℓ_1 panning with five loudspeakers $\mathbf{l}_1, \ldots, \mathbf{l}_5$ on a common circumcircle. The loudspeaker triangles corresponding to optimal vertex solutions for the source direction $\mathbf{p}$ are marked by patterns.

objective value. As shown in Sec. 4.3, the optimal solution of the dual is attained if the polygon spanned by the active loudspeakers includes the source direction $\mathbf{p}$. Thus, the set of optimal vertex solutions consists of all three-element sets of active loudspeakers on the circumcircle such that the spherical triangle formed by these loudspeakers contains the source $\mathbf{p}$. This is illustrated in Fig. 4 for a cyclic polygon formed by five loudspeakers $\mathbf{l}_1, \ldots, \mathbf{l}_5$. In this case each optimal vertex $\mathbf{g}_s^*$ corresponds to the selected triangle of a valid Delaunay triangulation of this polygon.

Moreover, as expressed by (14), the set of optimal solutions is formed by all convex combinations of the optimal vertices $\mathbf{g}_1^*$, $\mathbf{g}_2^*$,…,$\mathbf{g}_S^*$. Thus the VBAP solutions are a strict subset of the valid ℓ_1^+ solutions for nonunique cases. It is worth noting that some practical panning algorithms apply convex combinations of the vertex solutions. For instance, [39] averages the VBAP gains of all valid triangulations to improve the smoothness of the panning for ambiguous loudspeaker setups.

5. Optimal ℓ_1 Panning Without Nonnegativity

As shown in the previous section, a nonnegativity constraint imposed on the panning gains enables the ℓ_1 panning problem to be expressed as a linear program which yields identical solutions to VBAP, thus preserving the beneficial sparsity and locality properties of amplitude panning techniques. In this section, we demonstrate how the same LP framework can be used to solve the ℓ_1 problem without the nonnegativity constraint, and characterize the resulting

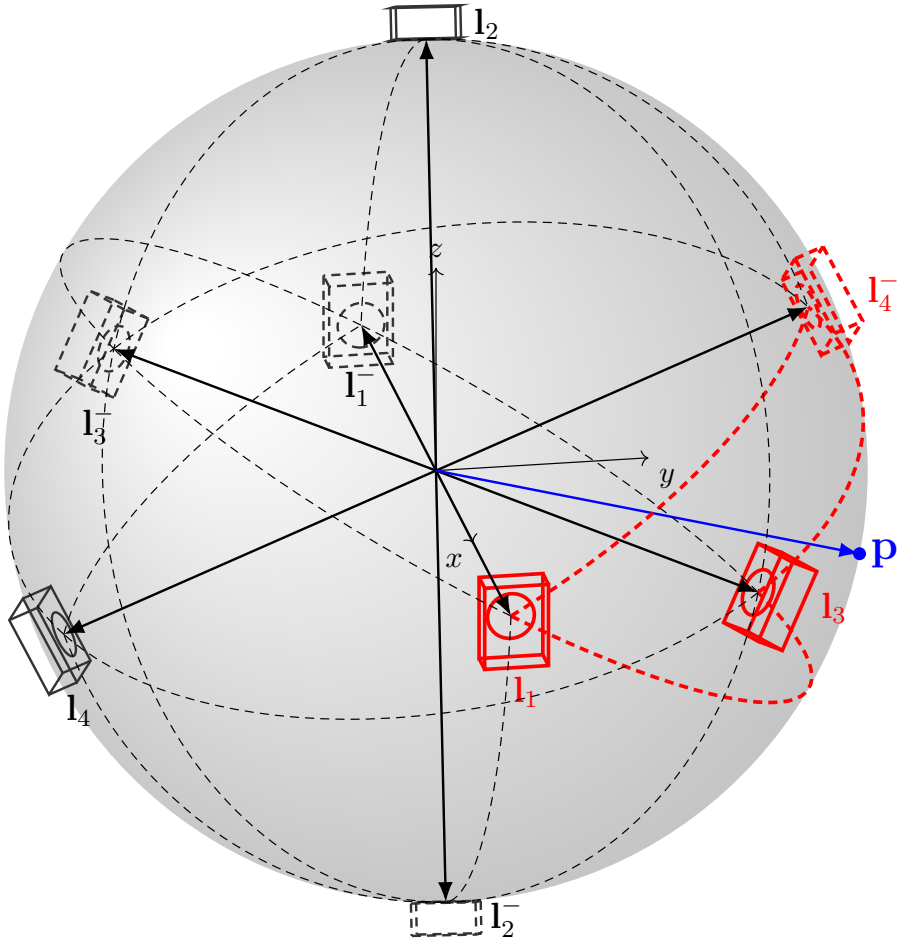

Figure 5: Augmented loudspeaker setup for ℓ_1 optimal panning without nonnegativity constraints. Virtual loudspeakers $\mathbf{l}_i^-$ are dashed, and the active loudspeaker triangle in red.

panning solutions.

The ℓ_1 optimization problem (9) can be translated into an LP in standard form [27, 58] as follows

$$\underset{\mathbf{g}^\pm}{\operatorname{argmin}}\ \mathbf{c}^T\mathbf{g}^\pm \text{ subject to } \mathbf{L}^\pm\mathbf{g}^\pm = \mathbf{p} \text{ and } \mathbf{g}^\pm \geq 0 \tag{24a}$$

where

$$\mathbf{g}^\pm = \begin{bmatrix}\mathbf{g}^+ & \mathbf{g}^-\end{bmatrix} \qquad \in \mathbb{R}^{2L\times 1} \tag{24b}$$

$$\mathbf{L}^\pm = \begin{bmatrix}\mathbf{L} & -\mathbf{L}\end{bmatrix} \qquad \in \mathbb{R}^{d\times 2L} \tag{24c}$$

$$\mathbf{c}^T = \begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix} \quad \in \mathbb{R}^{1\times 2L}, \tag{24d}$$

which doubles the sizes of the optimization variable $\mathbf{g}^\pm$ and the constraint matrix $\mathbf{L}^\pm$. This can be interpreted as augmenting the loudspeaker setup represented by the direction matrix $\mathbf{L}$ by a set of *mirror loudspeakers* $\mathbf{l}_i^- = -\mathbf{l}_i$, $1 \leq i \leq L$ pointing in the opposite directions to form the complete loudspeaker direction matrix $\mathbf{L}^\pm$ and the corresponding gain vector $\mathbf{g}^\pm \in \mathbb{R}^{2L\times 1}$,

$\mathbf{g}^{\pm} \geq 0$. The final panning gains for the physical loudspeaker configuration, that is, problem (9), are obtained from $\mathbf{g}^{\pm}$ through

$$\mathbf{g} = \mathbf{g}^{+} - \mathbf{g}^{-} \,. \tag{25}$$

Fig. 5 shows an augmented tetrahedral loudspeaker setup containing real and mirror loudspeakers.

Using this construction, we can apply the results for ℓ_1^+-optimal panning derived in the preceding section to the ℓ_1 panning problem without a nonnegativity constraint. Firstly, the ℓ_1 solutions preserve the same sparsity properties as in ℓ_1^+ panning, i.e., there always exists an optimal solution with at most d nonzero gains. Secondly, the optimal solution can be found using the same geometric construction based on the dual LP described in Sec. 4.2.1. That is, the active loudspeakers are selected based on a Delaunay triangulation of the complete augmented direction matrix $\mathbf{L}^{\pm}$ containing both real and mirror loudspeakers. This is exemplified in Fig. 5. For the source direction $\mathbf{p}$, the ℓ_1-optimal solution corresponds to the loudspeaker triangle $\{\mathbf{l}_1, \mathbf{l}_3, \mathbf{l}_4^-\}$, including the mirror loudspeaker $\mathbf{l}_4^-$. Using (25) to translate this solution into the gain vector of the real setup, this means that loudspeaker $\mathbf{l}_4$ is activated with a negative gain. This construction demonstrates that without a nonnegativity constraint, ℓ_1 optimal panning does not maintain the locality property of VBAP. Instead, loudspeakers close to the opposite of the source direction might become active with negative panning gains, creating antiphase sound field components from these directions. As argued in Secs. 2.4.2 and 2.4.4, such contributions typically degrade the quality of panning-based reproduction methods. Thirdly, augmentation by a set of mirror loudspeakers may lead to special cases caused by nonunique or degenerate VBAP solutions as described above. For instance, the optimization problem becomes ambiguous if the real setup contains diametrical loudspeakers, because in this case a mirror loudspeaker coincides with the opposite real loudspeaker.

Notwithstanding these potential drawbacks of omitting the nonnegativity constraint, this construction also represents an efficient algorithm to compute the globally optimal ℓ_1 panning solution. That is, the VBAP algorithm is applied to the augmented loudspeaker matrix (24c), and (25) is used to obtain the panning gains $\mathbf{g}$. That is, the algorithm does not require an explicit optimization step and its complexity is comparable to the very efficient VBAP algorithm.

6. Evaluation

In this section, we evaluate the properties of the ℓ_1 and ℓ_1^+ amplitude panning techniques and their equivalence to VBAP, and compare it to VBAP extensions to resolve nonunique traingulations, specifically an averaging technique proposed in [39] and a strategy using additional virtual loudspeakers that are downmixed to neighboring speakers specified in MPEG-H [23, 22]. These methods aim at a more symmetric reproduction and smoother source movements. Objective performance metrics are presented in Sec. 6.1 while ITD and ILD localization cues are used to estimate the subjective localization performance in Sec. 6.2.

To this end, we choose a practical 3D loudspeaker layout defined as Layout 15 in [23] and shown in Fig. 6. It consists of a total of ten loudspeakers in a spherical configuration, seven in the horizontal plane and three at an elevation angle of $\theta = 35°$. The loudspeaker labels have the form "`CH_{M,U}_{R,L}NNN`", where "`M`" and "`U`" denote a position in the horizontal (middle) and upper layer, respectively, "`L`" and "`R`" represent angles to the left and right, and "`NNN`" is the azimuth angle in degree. In the following, the panning solutions of the different algorithms are shown for three source positions that highlight different cases of 3D multichannel amplitude

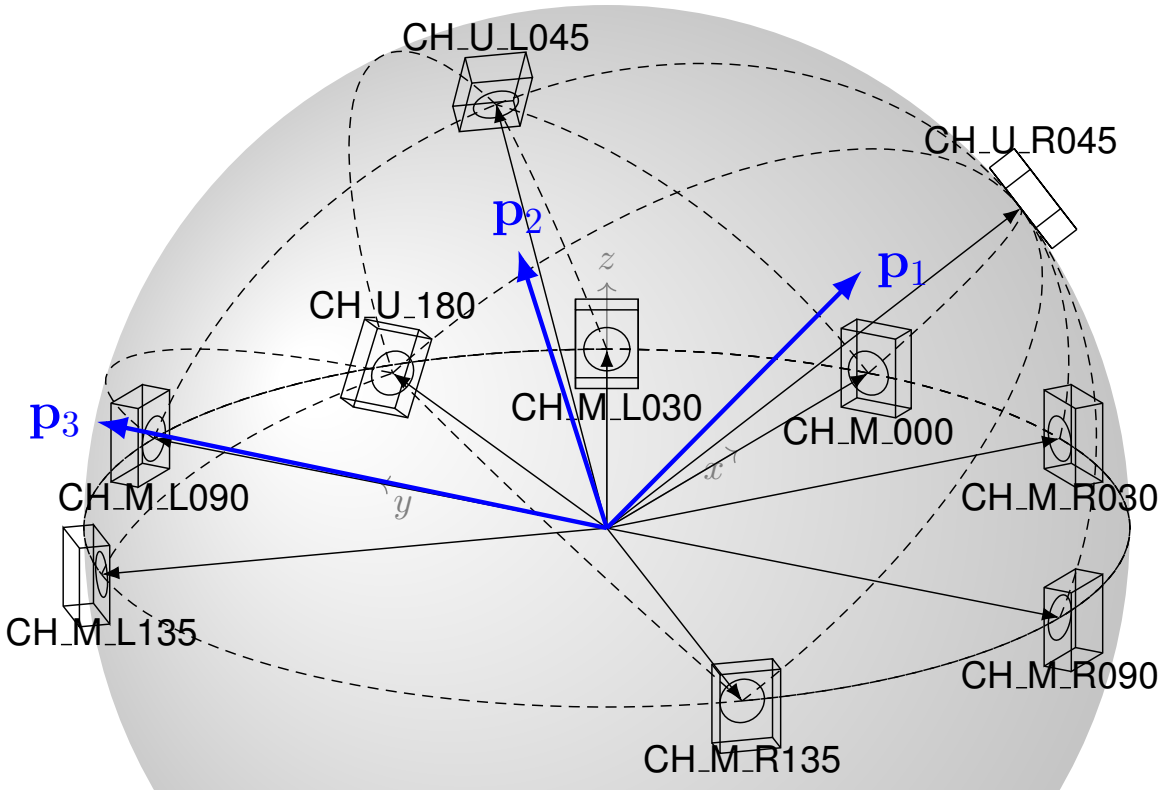

Figure 6: Example 3D loudspeaker setup according to Layout 15 in MPEG-H 3D audio [23], also showing the loudspeaker labels and the Delaunay triangulation for use with VBAP and source positions $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$.

panning. The CVX modeling framework [62, 63] is used for the proposed ℓ_1 and ℓ_1^+ panning methods.

6.1. Objective Performance Measures

For the objective evaluation, we evaluate the loudspeaker gain distribution and measures such as the ℓ_1 norm, the deviation of the velocity vector direction $\sphericalangle(\mathbf{p}, \mathbf{p}')$, the velocity vector magnitude r_v, and the number of nonzero gains, i.e., the ℓ_0 norm $\|\mathbf{g}\|_0$. These results are summarized in Table 1.

6.1.1. Unique Panning Solutions

As a first example, a source with direction $\mathbf{p}_1 = (\phi_1, \theta_1) = (0°, 12.5°)$ is chosen, where ϕ_i and θ_i denote source azimuth and elevation, respectively. VBAP reproduces this direction with the active loudspeaker triangle {`CH_M_000`, `CH_U_L045`, `CH_U_R045`}. The corresponding panning gains are shown in Fig. 7(a). This figure also shows that the gains obtained by the ℓ_1-optimal panning technique are identical to the VBAP case, both with respect to the active loudspeaker selection and gains. In fact, the maximum differences are in the order of $< 1 \cdot 10^{-15}$, which is within the accuracy of the numerical optimization algorithm (precision setting cvx_precision best). Thus, the ℓ_1-optimal solution retains the advantageous sparsity

Table 1: Objective performance measures for amplitude panning examples.

	$\mathbf{p}_1 = (0°, 12.5°)$				$\mathbf{p}_2 = (155°, 12.5°)$				$\mathbf{p}_3 = (100°, 12.5°)$			
Method	$\|\mathbf{g}\|_1$	$\|\mathbf{g}\|_0$	$\sphericalangle(\mathbf{p}, \mathbf{p}')$	r_v	$\|\mathbf{g}\|_1$	$\|\mathbf{g}\|_0$	$\sphericalangle(\mathbf{p}, \mathbf{p}')$	r_v	$\|\mathbf{g}\|_1$	$\|\mathbf{g}\|_0$	$\sphericalangle(\mathbf{p}, \mathbf{p}')$	r_v
VBAP	1.135	3	0°	0.881	1.192	3	0°	0.839	1.281	3	0°	0.781
ℓ_1	1.135	3	0°	0.881	1.091	3	0°	4.998	1.281	5	0°	1.615
ℓ_1^+	1.135	3	0°	0.881	1.192	3	0°	0.839	1.281	4	0°	0.781
Averaging [39]	1.135	3	0°	0.881	1.192	3	0°	0.839	1.281	4	0°	0.781
Virtual loudspeakers [23]	1.135	3	0°	0.881	1.192	3	0°	0.839	1.406	4	4.351°	0.818

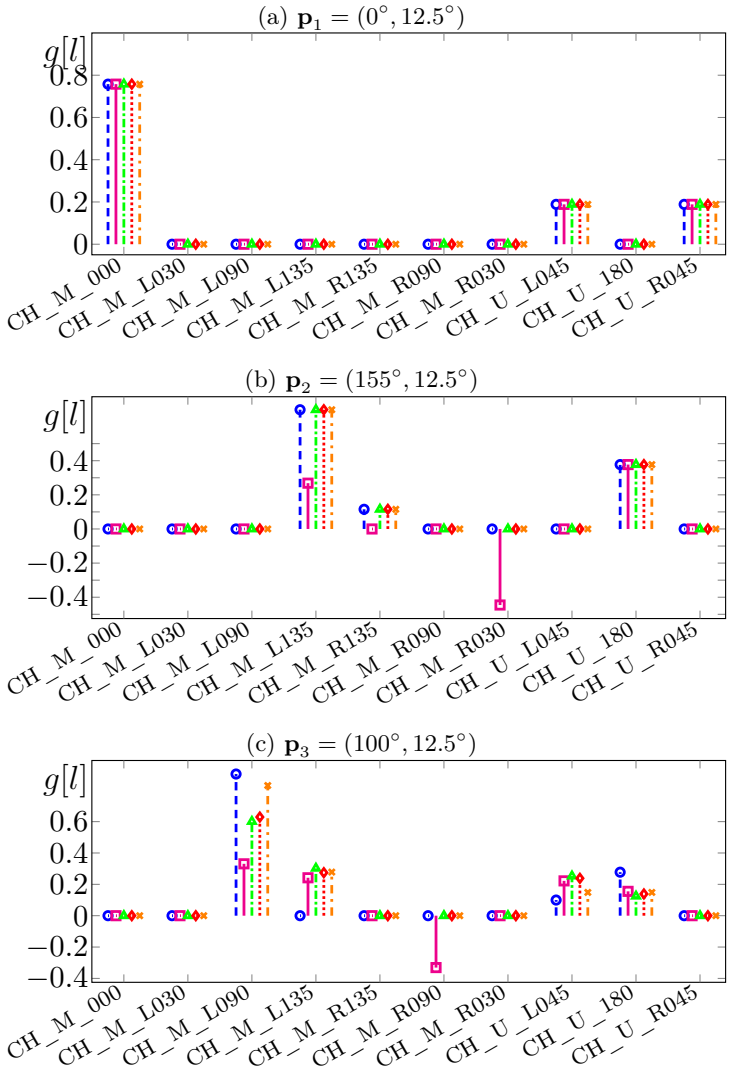

Figure 7: Panning weights for the loudspeaker configuration of Fig. 6 for source directions (ϕ, θ). : VBAP, : ℓ_1-optimal, : ℓ_1^+-optimal, : Averaging [39], : Virtual loudspeaker [23].

and locality properties of VBAP. Because the ℓ_1 panning gains are nonnegative in this case, it is clear that the results of the ℓ_1^+ method are identical to the ℓ_1 and VBAP cases. Likewise, the methods based on averaging and virtual loudspeakers are identical to all these solutions, since they differ from these methods only for nonunique panning cases. Table 1 summarizes the performance measures for the different methods. For $\mathbf{p}_1$, the velocity direction of the panned source matches the desired direction, i.e., $\sphericalangle(\mathbf{p}, \mathbf{p}') = 0°$ for all methods. Likewise, since the panning gains are identical, the ℓ_1 norms $\|\mathbf{g}\|_1$ and the velocity factors r_v are equal.

6.1.2. Nonnegativity Constraints

The panning weights for a second source direction $\mathbf{p}_2 = (155°, 12.5°)$ are displayed in Fig. 7(b). In this case, the ℓ_1 solution differs from the VBAP panning weights in that it contains a negative weight, namely from loudspeaker `CH_M_R090`. As reasoned in Sec. 5, the corresponding

mirror loudspeaker vector CH_M_R030$^-$ is included in the triangle {CH_M_L135, CH_M_R030$^-$, CH_U_L180} which fulfills the Delaunay circumcircle condition (Definition 1). For this reason, this triangle is chosen over the VBAP solution {CH_M_L135, CH_M_R135, CH_U_180}. As in the previous example, the averaging and virtual loudspeaker downmix solutions are identical to VBAP, because the panning is unique. The performance measures for this case are summarized in the second column block of Table 1. With the exception of the ℓ_1 solution, all measures are identical. The latter method achieves a smaller ℓ_1 norm , which is due to the selection of the mirror loudspeaker CH_M_R135$^-$. At the same time, the velocity vector magnitude is increased significantly because of the effect of negative gains on the denominator of (8). While larger values of r_v in theory indicate sharper image locations at low frequencies, this does not correspond with perception (see Sec. 6.2).

6.1.3. Nonunique Panning Solutions

To demonstrate the behavior of the panning methods for a nonunique configuration, a third source direction $\mathbf{p}_3 = (100°, 12.5°)$ is chosen. This direction lies within the cyclic loudspeaker-free spherical polygon {CH_M_L090, CH_M_L135, CH_U_180, CH_U_L045}. In the VBAP case, the triangle {CH_U_L045, CH_M_L135, CH_U_180} has been selected arbitrarily, resulting in a vertex solution with three active loudspeakers. For the other methods, all four loudspeakers of this cyclic quadrilateral are active. In case of the ℓ_1 and ℓ_1^+ methods, the solutions generated by CVX are displayed, that is, arbitrary elements of the nonunique solution sets. As they are formed by a linear combination of two distinct vertex solutions, they have four nonzero gains. In case of the ℓ_1 approach, the cyclic polygon also contains the mirror loudspeaker CH_M_R090$^-$, which results in five nonzero loudspeaker gains. As observed in the third column block of Table 1, the ℓ_1 norm of VBAP, ℓ_1, ℓ_1^+, and the averaging method are identical. This means that the solutions are contained in the optimal solution set of the panning problem, i.e., they synthesize the correct velocity direction with the minimum objective value for the ℓ_1 norm. In contrast, the virtual loudspeaker downmix algorithm inserts a loudspeaker at the center of the cyclic polygon, and distributes the gain assigned to this virtual loudspeaker to the neighboring real speakers. As observed in Table 1, this results in a lower ℓ_1 norm, but also in a deviation from the target velocity direction of about 4.4°.

6.2. Psychoacoustic Localization Cues

To assess the subjective performance, we simulate the ITD and ILD as the predominant localization cues. Fig. 8 shows these measures for a varying azimuth angle, that is, a simulated circular horizontal movement of a virtual source at an elevation of 12.5° around the center of the setup. In this, the generated data also covers the source positions $\mathbf{p}_1$–$\mathbf{p}_3$ investigated above. The ITD calculated as the time of the maximum of the interaural cross correlation (IACC) of the synthesized binaural impulse reponses [64]. As the ITD cue is relevant for low frequencies, the impulse responses are lowpass filtered with cutoff frequency of 1 kHz. ILDs are computed by averaging octave-band sound pressure level differences.

Figures 8(a) and 8(b) show the ILD and ITD values for both the ideal virtual source and the panning methods under freefield conditions. This simulation uses a head related transfer functions (HRTFs) measured with a Neumann KU 100 dummy head [65]. The ITD and ILD trajectories of all methods except ℓ_1 panning without nonnegativity are very similar to those of the ideal virtual source. If the panning problem is unique (azimuth range approx. $\phi \in$

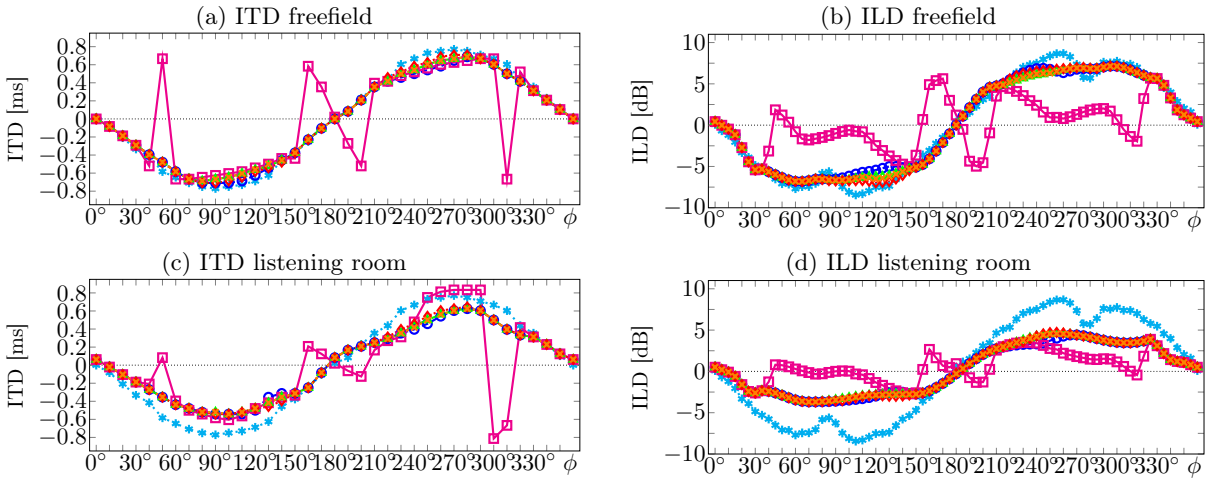

Figure 8: Simulated ITD and ILD values for freefield conditions and ITU-R BS.1116-3 listening room. ···✱···: Ideal freefield source, -○-: VBAP, —□—: ℓ_1-optimal, -·△·-: ℓ_1^+-optimal, ···◇···: Averaging [39], -·×·-: Virtual loudspeakers [23].

$\{0° \cdots 80°, 140° \cdots 220°, 280° \cdots 360°\}$), these methods are equivalent and therefore the ITDs and ILDs match exactly. For the remaining azimuth range (approx. $\phi \in \{90° \cdots 130°, 230° \cdots 270°\}$, the panning problem is nonunique and therefore the resulting ITDs and ILDs vary slightly, either due to the strategies used by the VBAP, the averaging, and the virtual loudspeaker downmix methods to resolve that ambiguity, or due to the arbitrary choice of one optimal solution returned by CVX in case of the ℓ_1^+ method. However, all cues are qualitatively similar and consistent with the ideal virtual source. Assessing the differences between these choices and designing perceptually optimal resolution strategies is a topic for future research.

In contrast, the ITD and ILD cues generated by the ℓ_1 approach differ significantly from the ideal values. In cases where the panning gains contain negative values as described in Sec. 6.1.2 (azimuth range approx. $\phi \in \{35° \cdots 45°, 150° \cdots 210°, 315° \cdots 325°\}$, ITD/ILD values differ significantly or are reversed compared to the ideal virtual source. For the ILD, this extends to directions that contain a negative gain contribution due to nonuniqueness ($\phi \in \{50° \cdots 145°, 215° \cdots 315°\}$). This is likely because of the sound energy from the opposing loudspeaker, which is significant at mid and high frequencies due to head shadowing.

To assess the performance of the proposed methods within a real room, the ITD/ILD evaluation is repeated using binaural room impulse responses (BRIRs). We use the BRIR dataset [66] of a multichannel reproduction system installed in a listening room ($RT_{60} \approx 0.22\,\mathrm{s}$) [67] that complies to the ITU-R BS.1116-3 standard. The resulting ITD and ILD trajectories are shown in Fig. 8(c) and 8(d). It is observed that the qualitative behavior is very similar to the freefield case. That is, the ITD/ILD cues of VBAP, ℓ_1^+, the averaging and the virtual loudspeaker method are similar to those of the ideal virtual source, while the ℓ_1-optimal solution without a nonnegativity constraint yields fluctuating or reversed ILD and ITD values. It is noted that reference ITD/ILD trajectories of the ideal virtual source are obtained from the freefield case, because the BRIR dataset used does not allow for arbitrary source positions. In contrast, the synthesized binaural impulse responses of the panning methods contain the reverberant field of the room. This might explain the lower absolute values of the ITD and ILD cues compared to the freefield reference.

The subjective sound localization performance of the different algorithms has been informally evaluated in an practical reproduction system and was found consistent with the ITD/ILD measures. While the ℓ_1 method without nonnegativity yields a fluctuating source localization, the other methods deliver a continuous, consistent source movements, and are very similar also in case of nonuniqueness. Binaural rendering based on both freefield HRTF data and BRIRs of a real listening room are provided as supplemental multimedia content. [2]

7. Conclusion

In this paper we have considered sparse, globally optimal solutions for multi-loudspeaker sound reproduction based on amplitude panning. To this end, we have proposed to formulate amplitude panning as an ℓ_1 optimization problem in order to retain the advantageous sparsity of amplitude panning methods as VBAP. We show that if the obtained solutions are unique, then they are exactly sparse with at most three nonzero loudspeaker gains, similar to the VBAP solution. It is shown that the ℓ_1 approach is in fact equivalent to VBAP if two conditions are fulfilled: 1) a nonnegativity constraint on the panning weights, and 2) the VBAP algorithm uses a Delaunay triangulation to determine the active speaker triangle. While the first condition is inherent to VBAP, the second is very close to the triangulation described in the original VBAP description, and actually used in the majority of existing implementations. By expressing this panning problem as a linear program, we utilize optimality conditions for LPs to characterize the optimal panning solutions. We show that the vertex solutions of the dual LP correspond to a Delaunay tessellation of the unit sphere surface. In particular, we prove that nonuniqueness of the panning solution results from degenerate vertex solutions of the dual LP, corresponding to more than three loudspeakers on a common circumcircle. We describe the shape of the solution set for these cases.

Utilizing the LP formulation, we show how the relaxation of the nonuniqueness constraint affects the full ℓ_1 solution by applying negative gains to loudspeakers opposite to the source direction, which contradicts the advantageous locality and constructive interference properties of amplitude panning methods. While such solutions are not desirable in most applications, we propose algorithms to solve the unconstrained ℓ_1-optimal panning problem by an inexpensive modification of the VBAP algorithm. In this way, we show that globally ℓ_1-optimal amplitude panning, with or without nonnegativity constraints, can be efficiently performed without runtime numerical optimization. This enables a linear complexity of $O(L)$ comparable to VBAP as opposed to $O(L^3)$ to $O(L^{3.5})$ for general-purpose ℓ_1 convex optimization methods [53, 68].

On a more conceptual level, this paper reduces the gap between perceptually motivated panning techniques (cf. [1]) and optimization-based physical sound field synthesis approaches. From a practical standpoint, we provide insight into the workings of VBAP-type algorithms. Specifically, we show how the properties of the Delaunay triangulation are linked to the optimality of the resulting panning solution, and that the ambiguities observed in practical VBAP implementations originate from the nonuniqueness of the solution of the underlying design objective. In this way, the present paper facilitates a better understanding of amplitude panning techniques, and thus paves the way to the development of more sophisticated and efficient panning algorithms.

[2] This paper has supplementary downloadable material available at http://ieeexplore.ieee.org and https://doi.org/10.15126/surreydata.00813551, provided by the authors. This includes eleven videos in MPEG-4 H.264 format and Matlab code. Contact a.franck@soton.ac.uk for further questions about this work.

Acknowledgment

The authors thank the reviewers for their constructive suggestions. We thank Dr. Jon Francombe and Dr. Russell Mason for providing the BRIR data of the listening room [66].

A. Objective Value of the Dual on a Spherical Arc Between Loudspeakers

In this appendix we derive the objective value

$$\mathbf{p}^T \boldsymbol{\pi}^k = \cos \sphericalangle (\mathbf{l}_i, \mathbf{p}) - \frac{\sin \sphericalangle (\mathbf{l}_i, \mathbf{p})}{\sin \sphericalangle (\mathbf{l}_i, \mathbf{l}_j)} \left[\cos \sphericalangle (\mathbf{l}_i, \mathbf{l}_j) - 1\right] \tag{26}$$

of the dual problem (22) for a source position $\mathbf{p}$ on a spherical arc connecting the loudspeakers $\mathbf{l}_i$ and $\mathbf{l}_j$ and a vertex $\boldsymbol{\pi}^k$. The spherical law of cosines. e.g., [42]

$$\cos(a) = \cos(b)\cos(c) - \sin(b)\sin(c)\cos(A) \tag{27}$$

relates the arc lengths a, b, and c of a spherical triangle to the angle A opposite to a. Applied to the geometry of Fig. 3, the cosine of angle γ_k is determined as

$$\cos \gamma_k = \cos r_\pi^k \frac{\cos \sphericalangle (\mathbf{l}_i, \mathbf{l}_j) - 1}{\sin r_\pi^k \sin \sphericalangle (\mathbf{l}_i, \mathbf{l}_j)} \,. \tag{28}$$

Using this result, $\cos \sphericalangle \left(\mathbf{p}^T \mathbf{p}_\pi^k\right)$ can be found by applying the law of cosines a second time

$$\begin{aligned} \cos \sphericalangle \left(\mathbf{p}^T \mathbf{p}_\pi^k\right) &= \cos r_\pi^k \cos \sphericalangle (\mathbf{l}_i, \mathbf{p}) - \sin r_\pi^k \sin \sphericalangle (\mathbf{l}_i, \mathbf{p}) \cos \gamma_k \\ &= \cos r_\pi^k \left(\cos \sphericalangle (\mathbf{l}_i, \mathbf{p}) - \frac{\sin \sphericalangle (\mathbf{l}_i, \mathbf{p})}{\sin \sphericalangle (\mathbf{l}_i, \mathbf{l}_j)} \left[\cos \sphericalangle (\mathbf{l}_i, \mathbf{l}_j) - 1\right] \right) . \end{aligned}$$

Inserting into (22) yields the final result (26).

References

[1] Sascha Spors, Hagen Wierstorf, Alexander Raake, Frank Melchior, Matthias Frank, and Franz Zotter. Spatial sound with loudspeakers and its perception: A review of the current state. *Proc. IEEE*, 101(9):1920–1938, September 2013.

[2] A.J. Berkhout. A holographic approach to acoustic control. *J. Audio Eng. So.*, 36(12):977–995, December 1988.

[3] A. J. Berkhout, Diemer de Vries, and Peter Vogel. Acoustic control by wave field synthesis. *J. Acoust. Soc. Amer.*, 95(5):2764–2778, May 1993.

[4] Evert W. Start. *Direct Sound Enhancement by Wave Field Synthesis*. PhD thesis, Delft University of Technology, Delft, The Netherlands, 1997.

[5] Jérôme Daniel, Sébastien Moreau, and Rozenn Nicol. Further investigations of high-order ambisonics and wavefield synthesis for holophonic sound imaging. In *Proc. Audio Eng. Soc. 114th Conv.*, March 2003.

[6] Sascha Spors, Rudolf Rabenstein, and Jens Ahrens. The theory of wave field synthesis revisited. In *Proc. Audio Eng. Soc. 124th Conv.*, Munich, Germany, May 2008.

[7] Jérôme Daniel. *Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia.* PhD thesis, Université Paris 6, Paris, France, July 2001.

[8] Jens Ahrens and Sascha Spors. Analytical driving functions for higher order Ambisonics. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, pages 373–376, Las Vegas, NV, USA, March 2008.

[9] Ole Kirkeby and Philip A. Nelson. Reproduction of plane wave sound fields. *J. Acoust. Soc. Amer.*, 94(5):2992–3000, November 1993.

[10] Shiro Ise. A principle of sound field control based on the Kirchhoff-Helmholtz integral equation and the theorie of inverse systems. *Acustica-Acta Acustica*, 85(1):78–97, 1999.

[11] Philippe-Aubert Gauthier, Alain Berry, and Wieslaw Woszczyk. Sound-field reproduction in-room using optimal control techniques: Simulations in the frequency domain. *J. Acoust. Soc. Amer.*, 117(2):662–678, February 2005.

[12] Mark Poletti. Robust two-dimensional surround sound reproduction for nonuniform loudspeaker layouts. *J. Audio Eng. Soc.*, 55(7/8):598–610, 2007.

[13] Wan-Ho Cho, Jeong-Guon Ih, and Marinus M. Boone. Holographic design of a source array for achieving a desired sound field. In *Proc. Audio Eng. Soc. 124th Conv.*, Amsterdam, The Netherlands, May 2008.

[14] Filippo Maria Fazi, Mincheol Shin, Ferdinando Olivieri, Simone Fontana, and Yue Lang. Comparison of pressure-matching and mode-matching beamforming for methods for circular loudspeaker arrays. In *Proc. Audio Eng. Soc. 137th Conv.*, Los Angeles, CA, USA, October 2014. paper no. 9111.

[15] John C. Bennett, Keith Barker, and Frederick O. Edeko. A new approach to the assessment of stereophonic sound system performance. *J. Audio Eng. Soc.*, 33(5):314–321, May 1985.

[16] Michael A. Gerzon. Panpot laws for multispeaker stereo. In *Proc. Audio Eng. Soc. 92th Conv.*, Vienna, Austria, March 1992. paper no. 3309.

[17] Ville Pulkki. Virtual sound source positioning using vector base amplitude panning. *J. Audio Eng. Soc.*, 45(6):456–466, June 1997.

[18] Ville Pulkki and Matti Karjalainen. Localization of amplitude-panned virtual sources I: Stereophonic panning. *J. Audio Eng. Soc.*, 49(9):739–752, September 2001.

[19] Ville Pulkki. Localization of amplitude-panned virtual sources II: Two- and three-dimensional panning. *J. Audio Eng. Soc.*, 49(9):753–767, September 2001.

[20] Ville Pulkki. *Spatial Sound Generation and Perception by Amplitude Panning Techniques.* PhD thesis, Helsinki University of Technology, Espoo, Finland, August 2001.

[21] Helmut Wittek. *Perceptual Differences between Wavefield Synthesis and Stereophony.* PhD thesis, University of Surrey, Guildford, UK, October 2008.

[22] Jürgen Herre, Johannes Hilpert, Achim Kuntz, and Jan Plogsties. MPEG-H 3D audio - the new standard for coding of immersive spatial audio. *IEEE J. Sel. Topics Signal Process.*, 9(5), 2015.

[23] ISO/MPEG 23008–3/DIS 3D audio. International standard, ISO/IEC, Sapporo, Japan, July 2014. ISO/IEC JTC1/SC29/WG11 N14747.

[24] Robert Tibshirani. Regression shrinkage and selection via the Lasso. *J. R. Stat. Soc. Series. B Stat. Methodol.*, 58(1):267–288, 1996.

[25] Stéphane G. Mallat and Zhifeng Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Trans. Signal Process.*, 41(12):3397–3415, December 1993.

[26] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. In *Proc. 27th Asilomar Conf. Signals, Systems and Computers*, pages 40–44, November 1993.

[27] Scott Shaobing Chen, David L. Donoho, and Michael A. Saunders. Atomic decomposition by basis pursuit. *SIAM J. Sci. Comput.*, 20:33–61, 1998.

[28] Irina F. Gorodnitsky and Baskar D. Rao. Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm. *IEEE Trans. Signal Process.*, 45(3):600–616, March 1997.

[29] Michael E. Tipping. Sparse Bayesian learning and the relevance vector machine. *J. Mach. Learn. Res.*, 1:211–244, June 2001.

[30] David L. Donoho. Compressed sensing. *IEEE Trans. Inf. Theory*, 52(4):1289–1306, April 2006.

[31] E.J. Candes and M.B. Wakin. An introduction to compressive sampling. *IEEE Signal Process. Mag.*, 25(2):21–30, March 2008.

[32] Nicolas Epain, Craig Jin, and André Van Schaik. The application of compressive sampling to the analysis and synthesis of spatial sound fields. In *Proc. Audio Eng. Soc. 127th Conv.*, New York, NY, USA, October 2009.

[33] A. Wabnitz, N. Epain, A. van Schaik, and C. Jin. Time domain reconstruction of spatial sound fields using compressed sensing. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, pages 465–468, Prague, Czech Republic, May 2011.

[34] G.N. Lilis, D. Angelosante, and G.B. Giannakis. Sound field reproduction using the Lasso. *IEEE Trans. Audio, Speech, Language Process.*, 18(8):1902–1912, November 2010.

[35] Shoichi Koyama, Suehiro Shimauchi, and Hitoshi Ohmuro. Sparse sound field representation in recording and reproduction for reducing spatial aliasing artifacts. In *Proc IEEE Int. Conf. Acoust., Speech, Signal Process.*, pages 4443–4447, Florence, Italy, May 2014.

[36] N. Radmanesh, I. S. Burnett, and B. D. Rao. A Lasso-LS optimization with a frequency variable dictionary in a multizone sound system. *IEEE/ACM Trans. Audio, Speech, Language Process.*, 24(3):583–593, March 2016.

[37] Hanieh Khalilian, Ivan V. I. V. Bajić, and Rodney G. Vaughan. Joint optimization of loudspeaker placement and radiation patterns for sound field reproduction. In *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, pages 519–523, Brisbane, Australia, April 2015.

[38] Andreas Franck and Filippo Maria Fazi. Comparison of listener-centric sound field reproduction methods in a convex optimization framework. In *Proc. 2016 Audio Eng. Soc. Int. Conf. Sound Field Control*, Guildford, UK, July 2016.

[39] Pierre-Anthony Lemieux, Roger Dressler, and Jean-Marc Jot. Object-based audio system using vector base amplitude panning, 2013.

[40] Christian Borß. A polygon-based panning method for 3D loudspeaker setups. In *Proc. Audio Eng. Soc. 137th Conv.*, Los Angeles, CA, USA, October 2014.

[41] Stanley P. Lipshitz. Stereo microphone techniques: Are the purists wrong? *J. Audio Eng. Soc*, 34(9):716–744, September 1986.

[42] I. Todthunter. *Spherical Geometry*. Macmillan & Co., London, UK, 5th edition, 1886.

[43] Franco P. Preparata and Michael Ian Shamos. *Computational Geometry: An Introduction*. Springer, New York, NY, USA, 1985.

[44] Mark de Berk, Otfried Cheong, Marc von Krevald, and Mark Overmars. *Computational Geometry: Algorithms and Applications*. Springer-Verlag, Berlin, Germany, 3rd edition, 2008.

[45] Franz Zotter and Matthias Frank. All-round Ambisonic panning and decoding. *J. Audio Eng. Soc.*, 60(10):807–820, October 2012.

[46] Ville Pulkki and Tapio Lokki. Creating auditory displays with multiple loudspeakers using VBAP: A case study with DIVA project. In *Proc. Int. Conf. Auditory Display*, Glasgow, UK, 1998.

[47] Andrzej Lingas. The greedy and Delauney triangulations are not bad in the average case. *Inform. Process. Lett.*, 22(1):25 – 31, 1986.

[48] C. Bradford Barber, David P. Dobkin, and Hannu Huhdanpaa. The Quickhull algorithm for convex hulls. *ACM Trans. Math. Softw.*, 22(4):469–483, December 1996.

[49] Mikko-Ville Laitinen, Juha Vilkamo, Kai Jussila, Archontis Politis, and Ville Pulkki. Gain normalization in amplitude panning as a function of frequency and room reverberance. In *Proc. Audio Eng. Soc. 55th Int. Conf.*, Helsinki, Finland, August 2014.

[50] Michael A. Gerzon. General metatheory of auditory localisation. In *Proc. Audio Eng. Soc. 92th Conv.*, Vienna, Austria, March 1992. paper no. 3306.

[51] Jean-Marc Jot, Veronique Larcher, and Jean-Marie Pernaux. A comparative study of 3-D audio encoding and rendering techniques. In *Proc. Audio Eng. Soc. 16th Int. Conf.*, Rovaniemi, Finland, March 1999.

[52] Ville Pulkki. Uniform spreading of amplitude panned virtual sources. In *Proc. IEEE Workshop Applicat. Signal Process. Audio Acoust.*, pages 187–190, New Paltz, NY, USA, October 1999.

[53] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, UK, 2004.

[54] Jean-Jaques Fuchs. On sparse representations in arbitrary redundant bases. *IEEE Trans. Inf. Theory*, 50(6):1341–1344, June 2004.

[55] David L. Donoho and Michael Elad. Optimally sparse representation in general (nonorthogonal) dictionaries via l^1 minimization. *Proc. Natl. Acad. Sci. U.S.A.*, 100(5):2197–2202, 2003.

[56] Joel A. Tropp. Greed is good: Algorithmic results for sparse approximation. *IEEE Trans. Inf. Theory*, 50(10):2231–2242, October 2004.

[57] David L. Donoho and Yaakov Tsaig. Fast solution of ℓ_1-norm minimization problems when the solution may be sparse. *IEEE Trans. Inf. Theory*, 54(11):4789–4812, November 2008.

[58] Hui Zhang, Wotao Yin, and Lizhi Cheng. Necessary and sufficient conditions of solution uniqueness in 1-norm minimization. *J. Optim. Theory Appl.*, 164(1):109–122, January 2015.

[59] George R. Dantzig. *Linear Programming and Extensions*. Princeton University Press, Princeton, NJ, USA, 1963.

[60] Vašek Chvátal. *Linear Programming*. Freeman, New York, NY, USA, 1983.

[61] Philip E. Gill, Walter Murray, and Margaret H. Wright. *Numerical Linear Algebra and Optimization*, volume 1. Addison-Wesley, Redwood City, CA, USA, 1991.

[62] Michael Grant and Stephen Boyd. Graph implementations for nonsmooth convex programs. In V. Blondel, S. Boyd, and H. Kimura, editors, *Recent Advances in Learning and Control*, pages 95–110. Springer, 2008.

[63] Michael Grant and Stephen Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. http://cvxr.com/cvx, March 2014.

[64] Brian F. G. Katz and Markus Noisternig. A comparative study of interaural time delay estimation methods. *J. Acoust. Soc. Amer.*, 135(6):3530–3540, June 2014.

[65] Benjamin Bernschütz. A spherical far field HRIR/HRTF compilation of the Neumann KU 100. In *Proc. AIA-DAGA Conf. Acoust.*, Meran, Italy, March 2013.

[66] Jon Francombe. IoSR listening room multichannel BRIR dataset, 2016. DOI 10.15126/surreydata.00813511.

[67] Russell Mason. Installation of a flexible 3D audio reproduction system into a standardized listening room. In *Proc. Audio Eng. Soc. 140th Conv.*, Paris, France, June 2016. Engineering Brief.

[68] W. Xu and B. Hassibi. Efficient compressive sensing with deterministic guarantees using expander graphs. In *Proc. 2007 IEEE Information Theory Workshop*, pages 414–419, Sept 2007.