Comparison of energy-preserving and all-round Ambisonic decoders

Franz Zotter, Matthias Frank, and Hannes Pomberger
University of Music and Performing Arts, Graz, Institute of Electronic Music and Acoustics,
Email: {zotter,frank,pomberger}@iem.at

Introduction

Ambisonics is driven by the idea that its underlying multichannel audio representation works nicely on any loudspeaker arrangement, if it is decoded suitably. Only a little of this promising idea became true in the past as decoding was an absolutely non-trivial thing for any general, irregular loudspeaker arrangement that spans a three dimensional surface. In fact, classical sampling and mode-matching decoders are well-behaved on special, regular arrangements only. The authors recently presented two decoding methods that cope with loudspeaker arrangements of any kind: energy-preserving and all-round Ambisonic decoding. Which of them is better, and how good are they compared with the classical sampling and mode-matching decoders? This contribution sketches their particularities and shows an exemplary comparison using various estimators visualizing the smoothness of the phantom source in terms of loudness, directional mapping error, and width.

Virtual Panning Function in Circular and Spherical Harmonics

The technique we call Ambisonics expresses the excitation of the sound field as a continuous function at some specified radius R where the loudspeakers are placed at. We call this a *virtual panning function* as it is continuous and therefore an idealization of what can be controlled by real loudspeakers. Nevertheless, in Ambisonics this function is conveniently expanded into orthonormal circular or spherical harmonics of limited order $n, |m| \leq \mathrm{N}$. Stacking these functions into a vector $\boldsymbol{y}_\mathrm{N}(\boldsymbol{\theta})$ and their expansion coefficients into another vector $\boldsymbol{\varepsilon}_\mathrm{N}$, we obtain

$$g(\boldsymbol{\theta}) = \boldsymbol{y}_\mathrm{N}(\boldsymbol{\theta})^\mathrm{T}\, \boldsymbol{\varepsilon}_\mathrm{N}, \tag{1}$$

where the vector $\boldsymbol{\theta} = [\cos\varphi \sin\vartheta, \sin\varphi\sin\vartheta, \cos\vartheta]^\mathrm{T}$ depends on azimuth φ and zenith ϑ. For virtual sound sources, a useful shape of $g(\boldsymbol{\theta})$ is obtained by the expansion coefficients of a Dirac delta function at the panning direction $\boldsymbol{\theta}_\mathrm{s}$ and optional weights $\boldsymbol{a}_\mathrm{N}$

$$\boldsymbol{\varepsilon}_\mathrm{N}(\boldsymbol{\theta}_\mathrm{s}) = \mathrm{diag}\{\boldsymbol{a}_\mathrm{N}\}\; \boldsymbol{y}_\mathrm{N}(\boldsymbol{\theta}_\mathrm{s}). \tag{2}$$

As the order is limited, the angular resolution of this function is strictly limited. In circular systems, $\boldsymbol{y}_\mathrm{N}$ contains circular harmonics Φ_n of limited $-\mathrm{N} \leq n \leq \mathrm{N}$: $\cos(n\varphi)/\sqrt{\pi}$ for $n > 0$, $\sin(n\varphi)/\sqrt{\pi}$ for $n < 0$, and $1/\sqrt{2\pi}$ for $n = 0$. In spherical systems, $\boldsymbol{y}_\mathrm{N}$ contains spherical harmonics Y_n^m: $\sqrt{\frac{(n-|m|)!(2n+1)}{(n+|m|)!\,2}}\, P_n^{|m|}(\cos\vartheta)\, \Phi_m(\varphi)$ employing the circular harmonics and the associated Legendre functions P_n^m with $0 \leq n \leq \mathrm{N}$, $-n \leq m \leq n$.

The weights $\mathrm{diag}\{\boldsymbol{a}_\mathrm{N}\}$ suppress the side lobes (Gibbs phenomenon) of the virtual panning function. Suitable max-r_E weights are, cf. [3], $a_n = \cos(n\,90°/(\mathrm{N}+1))$ for circular and $a_n = P_n(139°/(\mathrm{N}+1.51))$ for spherical systems. (Other than here, but equivalently, weights are classically part of decoding Eq. (4) not encoding Eq. (3).)

Ambisonic Encoding and Decoding

In Ambisonics, a mono signal $s(t)$ can be *encoded* as virtual source at the direction $\boldsymbol{\theta}_\mathrm{s}$ by the vector of Eq. (2),

$$\boldsymbol{\chi}_\mathrm{N}(t) = \boldsymbol{\varepsilon}_\mathrm{N}(\boldsymbol{\theta}_\mathrm{s})\; s(t). \tag{3}$$

Typical Ambisonic signals $\boldsymbol{\chi}_\mathrm{N}(t)$ contain several linearly simperimposed virtual sources or recordings. The signals are distributed to a discrete arrangement of L loudspeakers at the directions $\{\theta_l\}_{l=1\ldots\mathrm{L}}$ by a *decoder* $\boldsymbol{D}$ that needs to be determined somehow. The loudspeaker signals $\boldsymbol{x}(t)$ are obtained by matrix multiplication with this decoder

$$\boldsymbol{x}(t) = \boldsymbol{D}\; \boldsymbol{\chi}_\mathrm{N}(t). \tag{4}$$

Decoder design aims at finding $\boldsymbol{D}$ to optimally present the Ambisonic signals Eq. (3) to the human ears using the discrete loudspeakers. In fact, decoder design has been one of the biggest challenges in Ambisonics. Three decoder designs are discussed below, which are decribed in the articles [1, 2]: sampling Ambisonic decoding (SAD), all-round Ambisonic decoding (AllRAD), and energy-preserving Ambisonic decoding (EPAD). Because mode-matching Ambisonc decoding (MMAD) can be unstable, we only discuss it in the appendix.

Sampling Ambisonic Decoding (SAD)

The simplest decoding is sampling the virtual panning function Eq. (1) at the L loudspeaker directions $\{\theta_l\}_{l=1\ldots\mathrm{L}}$, cf. [6]. The sampled harmonics are

$$\boldsymbol{Y}_\mathrm{N} = [\boldsymbol{y}_\mathrm{N}(\theta_1),\, \ldots,\, \boldsymbol{y}_\mathrm{N}(\theta_\mathrm{L})], \tag{5}$$

and the transpose of this matrix evaluates $\boldsymbol{\varepsilon}_\mathrm{N}$. Normalized by the mean angular segment of each loudspeaker, SAD for a semi-circular system is

$$\boldsymbol{D}_\mathrm{SAD} = \frac{\pi}{\mathrm{L}} \boldsymbol{Y}_\mathrm{N}^\mathrm{T}. \tag{6}$$

SAD is optimal for t-design layouts with $t \geq 2\mathrm{N}+1$, cf. [2], i.e. regular $\geq 2\mathrm{N}{+}2$-polygons for a circular system.

All-Round Ambisonic Decoding (AllRAD)

All-round Ambisonic decoding, cf. [2], is designed in two steps. First, an optimal *virtual* loudspeaker layout is considered, e.g., a regular 50-polygon, for which the sampling decoder $\boldsymbol{D}_\mathrm{SAD}^{(\mathrm{opt})}$ is optimal (energy-preserving and mode-matching). Secondly, the signals of these virtual

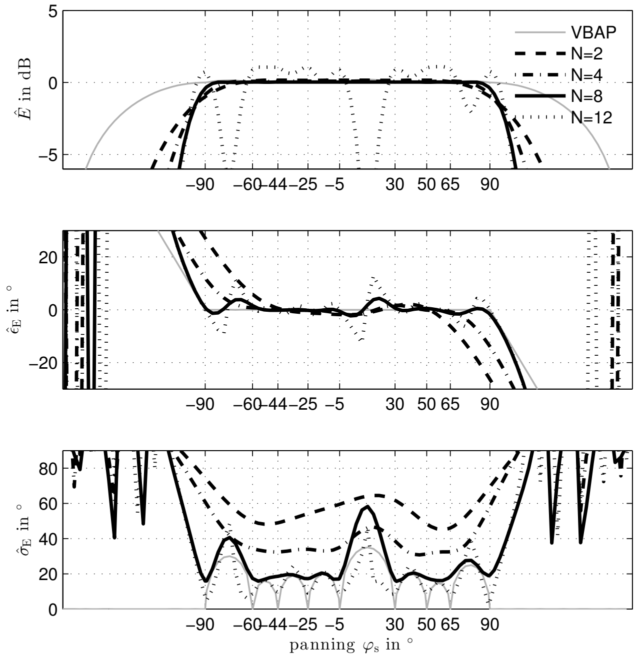

Figure 1: Energy, directional energy mapping error, energy spread of EPAD (energy-preserving decoder) for examplary semi-circular loudspeaker layout and the orders $\mathrm{N} = \{2, 4, 8, 12\}$; EPAD only works up $\mathrm{N} = 8$ in the example.

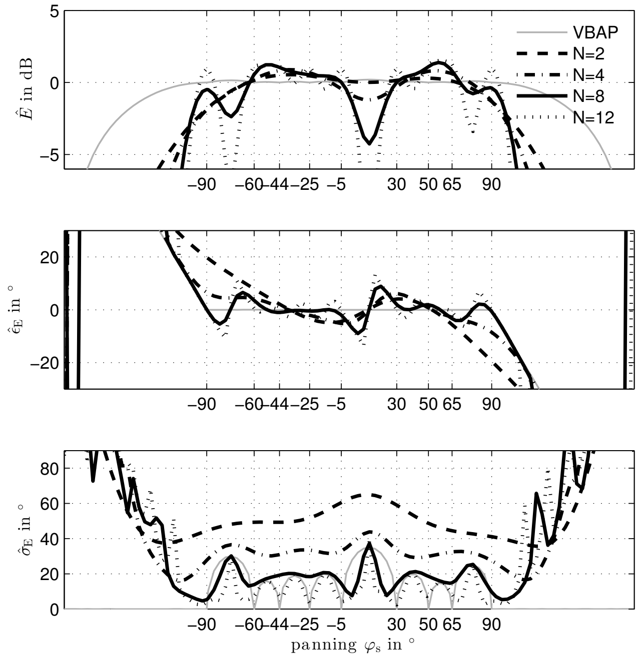

Figure 2: Energy, directional energy mapping error, energy spread of the SAD (sampling decoder) for examplary semi-circular loudspeaker layout and the orders $\mathrm{N} = \{2, 4, 8, 12\}$. $\hat{E}$ curves were normalized in the mean between $-90°$ and $90°$.

loudspeakers are mapped to the *real* loudspeakers by a matrix $\boldsymbol{G}$ obtained by vector-base amplitude panning (VBAP, [7]). Combining both steps we get

$$\boldsymbol{D}_{\mathrm{AllRAD}} = \boldsymbol{G}_{1:\mathrm{L},:}\, \boldsymbol{D}_{\mathrm{SAD}}^{(\mathrm{opt})}. \tag{7}$$

For incomplete circular/spherical layouts, $\boldsymbol{G}$ involves an *imaginary* $(\mathrm{L}+1)^{\mathrm{st}}$ dummy loudspeaker (or more) to control the signal loss at missing directions. An interesting equivalence of AllRAD to the most VBAP-like Ambisonic decoding is described in the appendix section.

Energy-Preserving Ambisonic Decoding (EPAD)

On the semi circle, the number of required basis functions can be reduced to push the loss of energy out of the panning range and to maintain a high definition elsewhere. For a semi circle of the range $|\varphi| \leq 90°$, an SVD is done of a matrix that contains uniformly sampled harmonics of the slightly over-sized range, cf. [1],

$$\boldsymbol{U}\,\boldsymbol{S}\,\boldsymbol{V}^{\mathrm{T}} = \mathrm{svd}\{[\boldsymbol{y}_{\mathrm{N}}(-105°), \boldsymbol{y}_{\mathrm{N}}(-104°), \dots, \boldsymbol{y}_{\mathrm{N}}(105°)]\}.$$

The $(2\mathrm{N}+1)$-sized basis of the circular harmonics can be reduced to a set of $\mathrm{N}+1$ basis functions. Given singular values in $\boldsymbol{S}$ sorted in descending order, the reduction matrix can be written (in MATLAB notation) as $\boldsymbol{U}_{:,1:\mathrm{N}+1}^{\mathrm{T}}$. Left multiplication is used to re-expand any Ambisonics-encoded function. All loudspeaker directions are encoded without weights by $\boldsymbol{Y}_{\mathrm{N}}$, cf. Eqs. (5) and (2). The energy-preserving decoder takes another SVD of the re-expanded loudspeaker encoding

$$\check{\boldsymbol{U}}\,\check{\boldsymbol{S}}\,\check{\boldsymbol{V}}^{\mathrm{T}} = \mathrm{svd}\{\boldsymbol{U}_{:,1:\mathrm{N}+1}^{\mathrm{T}}\,\boldsymbol{Y}_{\mathrm{N}}\}.$$

After transposition, omission of the singular values, and re-expansion of the Ambisonic signals, energy-preserving decoding becomes, cf. [1]:

$$\boldsymbol{D}_{\mathrm{EPAD}} = \check{\boldsymbol{V}}_{:,1:\mathrm{N}+1}\check{\boldsymbol{U}}^{\mathrm{T}}\,\boldsymbol{U}_{:,1:\mathrm{N}+1}^{\mathrm{T}}. \tag{8}$$

Example: semi-circular layout

Let's assume an irregular semi-circular loudspeaker layout with loudspeakers located at $\{\varphi_l\} = \{-90°, -60°, -44°, -25°, -5°, 30°, 50°, 65°, 90°\}$ and the decoders $\boldsymbol{D}_{\mathrm{SAD}}$, Eq. (6), $\boldsymbol{D}_{\mathrm{AllRAD}}$, Eq. (7), $\boldsymbol{D}_{\mathrm{EPAD}}$, Eq. (8). In particular, regarding Eqs. (2), (3), and (4), we see that Ambisonic encoding and decoding allows the interpretation in terms of amplitude panning gains

$$\boldsymbol{g}(\boldsymbol{\theta}_{\mathrm{s}}) = \boldsymbol{D}\,\boldsymbol{\varepsilon}_{\mathrm{N}}(\boldsymbol{\theta}_{\mathrm{s}}) \tag{9}$$

to distribute signals to the loudspeakers. To evaluate the quality of these gains for panning, we employ the quality measures below.

Quality measures

The *energy* measure, cf. [4], $\hat{E}(\varphi_{\mathrm{s}})$ estimates the loudness fluctuation of the decoder using Eq. (9)

$$\hat{E}(\varphi_{\mathrm{s}}) = \sum_{l=1}^{\mathrm{L}} g_l^2(\varphi_{\mathrm{s}}). \tag{10}$$

The $\hat{\boldsymbol{r}}_{\mathrm{E}}$ measure, cf. [4], $\hat{\boldsymbol{r}}_{\mathrm{E}}(\varphi_{\mathrm{s}}) = \frac{\sum_{l=1}^{\mathrm{L}} g_l^2(\varphi_{\mathrm{s}})\boldsymbol{\theta}_l}{\hat{E}(\varphi_{\mathrm{s}})}$ is a vector estimating the directional mapping of a decoder. Its direction is used to estimate the angular mapping error, cf. [1]. For a circular system, this is

$$\hat{\epsilon}_{\mathrm{E}}(\varphi_{\mathrm{s}}) = \left[\arctan\frac{\hat{r}_{\mathrm{E,y}}(\varphi_{\mathrm{s}})}{\hat{r}_{\mathrm{E,x}}(\varphi_{\mathrm{s}})} - \varphi_{\mathrm{s}} + 180°\right]_{\mathrm{mod}\,360°} - 180°. \tag{11}$$

The length of $\boldsymbol{r}_{\mathrm{E}}$ is used to estimate the angular spread of a decoded virtual source

$$\hat{\sigma}_{\mathrm{E}}(\varphi_{\mathrm{s}}) = 2\arccos\|\hat{\boldsymbol{r}}_{\mathrm{E}}(\varphi_{\mathrm{s}})\|. \tag{12}$$

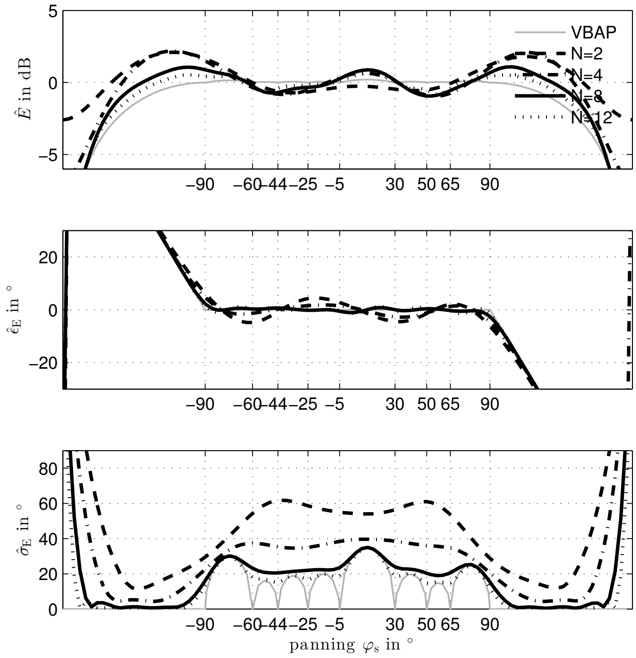

Figure 3: Energy, directional energy mapping error, energy spread of the AllRAD (all-round decoder) for examplary semi-circular loudspeaker layout and the orders $\mathrm{N} = \{2, 4, 8, 12\}$. $\hat{E}$ curves were normalized in the mean between $-90°$ and $90°$.

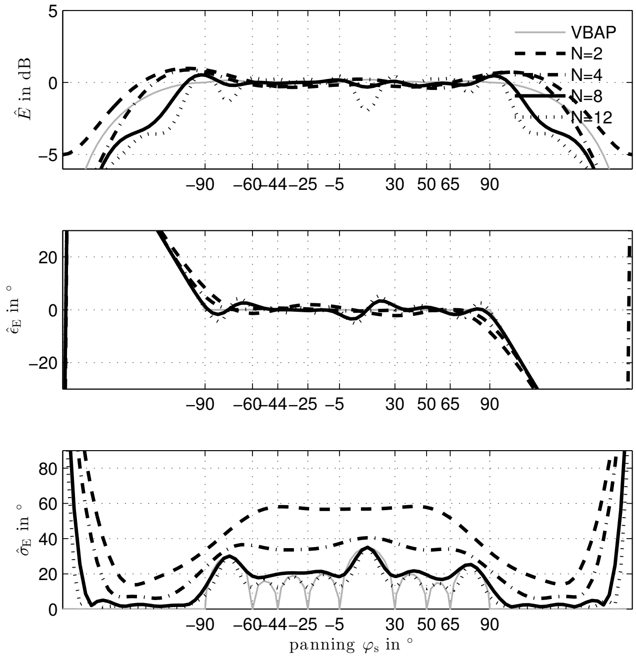

Figure 4: Energy, directional energy mapping error, energy spread of AllRAD+ (improved all-round decoder) for examplary semi-circular loudspeaker layout and the orders $\mathrm{N} = \{2, 4, 8, 12\}$. Resulting $\hat{E}$ curves were normalized in the mean between $-90°$ and $90°$.

All these three measures, $\hat{E}$, $\hat{\epsilon}_{\mathrm{E}}$, $\hat{\sigma}_{\mathrm{E}}$, should ideally be panning-independent, i.e. constant, cf. [1]. As the exemplary layout is non-ideal, we expect panning-dependent quality measures.

In all examples below, we included the measures for VBAP as a reference. In addition to the imaginary loudspeaker at 180°, the employed modified variant of VBAP used the square root of its unnormalized gains to ensure $\hat{\epsilon}_{\mathrm{E}} = 0$ in the panning range.

Properties of energy-preserving decoding. Fig. 1 shows the loudness measure $\hat{E}$ obtained by EPAD which is perfectly constant in all the panning range between $-90°$ and $90°$. The mislocalization $\hat{\epsilon}_{\mathrm{E}}$ is rather smooth, but sounds are pulled inwards at the $\pm 90°$ boundaries of the playback facility. The spread $\hat{\sigma}_{\mathrm{E}}$ is only wiggly in the largest gap between $-5°$ and $30°$.

Properties of sampling decoding. Fig. 2 indicates that SAD produces a loudness dip in $\hat{E}$ wherever the spacing of the loudspeakers is large. Moreover, SAD tends to create large mislocalization overshoots of $\hat{\epsilon}_{\mathrm{E}}$ at these positions. The spread $\hat{\sigma}_{\mathrm{E}}$ is rather smooth.

Properties of all-round decoding. Fig. 3 shows the $\hat{E}$ curve of AllRAD. The energy fluctuation seems rough, keeping in mind that the decoder to the underlying virtual layout is optimal. The mislocalization $\hat{\epsilon}_{\mathrm{E}}$ is highly smooth and very small, and it still is at the border of the $\pm 90°$ panning range. The spread $\hat{\sigma}_{\mathrm{E}}$ has the least wiggles and overshoots of all decoders and even decreases outside the panning range, which seems preferable.

Improved Definition AllRAD+

Inspecting Figs. 2 (SAD) and 3 (AllRAD), the opposing tendencies of the loudness measure $\hat{E}$ become obvious for $\mathrm{N} \leq 8$. We expect an improved performance of their combination

$$\boldsymbol{D}_{\mathrm{AllRAD+}} = g_{\mathrm{AllRAD}}\, \boldsymbol{D}_{\mathrm{AllRAD}} + g_{\mathrm{SAD}}\, \boldsymbol{D}_{\mathrm{SAD}}. \qquad (13)$$

In fact, constant energy that is achieved for the idealized virtual loudspeaker setup in AllRAD is corrupted by the VBAP stage as, per loudspeaker pair, all virtual sources are superimposed linearly instead of energetically. The prevailing linear superposition increases the energy wherever the loudspeaker spacing is large. Roughly, at such directions AllRAD doubles the energy, whereas it is halved at directions with dense loudspeaker spacing. Conversely, SAD might lose all energy where the loudspeaker spacing is large and roughly doubles it where the loudspeaker spacing is dense. For both cases, solving the estimated equation system

$$\begin{pmatrix} \sqrt{2} & 0 \\ 1/\sqrt{2} & \sqrt{2} \end{pmatrix} \begin{pmatrix} g_{\mathrm{AllRAD}} \\ g_{\mathrm{SAD}} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad (14)$$

for unity yields $g_{\mathrm{AllRAD}} = \frac{1}{\sqrt{2}}$, $g_{\mathrm{SAD}} = \frac{1}{\sqrt{8}}$. Fig. 4 shows the hereby-obtained improvement in the curve for $\hat{E}$. The favorable angular mapping characteristics of AllRAD is largely preserved in both $\hat{\epsilon}_{\mathrm{E}}$ and $\hat{\sigma}_{\mathrm{E}}$.

Conclusion

In this contribution, we compared the recent developments: the superior but complicated energy-preserving decoding (EPAD) and all-round Ambisonic decoding (AllRAD). We obtained an improved decoder AllRAD+ by combining AllRAD with sampling decoding (SAD). The loudness variation of AllRAD+ is competitive with EPAD and its angular mapping resembles AllRAD.

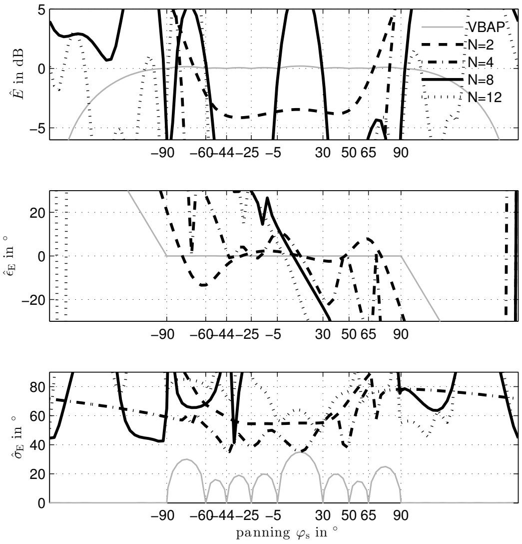

Figure 5: The MMAD (mode-matching decoder) with $\alpha = 0$ does not compete with other decoders.

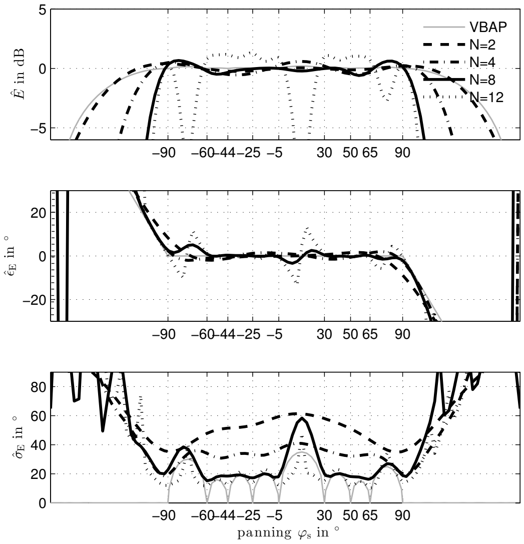

Figure 6: The MMAD is competitive after adding $\alpha = \frac{1}{4}\frac{\mathrm{L}}{\pi}$.

Acknowledgment

The authors thank Aaron Heller for a fruitful discussion concerning the properties of AllRAD and about improving this paper. Moreover, we thank Nicolas Epain for sharing his valuable perspective on AllRAD with us, which is described in the appendix.

Appendix: The most VBAP-like decoder

In a conversation at AIA-DAGA, Nicolas Epain proposed to regard his *most VBAP-like Ambisonic decoder*. According to Epain, this decoder, let's abbreviate it by MVLAD, is equivalent to AllRAD. This interesting persective is acknowledged here by a short formal derivation.

Let's write $g_l^{(\mathrm{v})}$ for the panning gain calculated by VBAP [7] for the l^{th} loudspeaker.
Unweighted Ambisonic panning/decoding to loudspeakers is $\boldsymbol{g}^{(\mathrm{a})} = \boldsymbol{D}\,\boldsymbol{y}_{\mathrm{N}}(\boldsymbol{\theta}_{\mathrm{s}})$. Defining the row of the l^{th} loudspeaker in the decoder as $\boldsymbol{d}_l^{\mathrm{T}}$, we obtain the gain $g_l^{(\mathrm{a})} = \boldsymbol{y}_{\mathrm{N}}(\boldsymbol{\theta}_{\mathrm{s}})^{\mathrm{T}}\,\boldsymbol{d}_l$ for the l^{th} loudspeaker.

The MVLAD minimizes the squared deviation $[g_l^{(\mathrm{a})} - g_l^{(\mathrm{v})}]^2$ to gains calculated with VBAP. We can define it by minimization across all panning directions $\boldsymbol{\theta}_{\mathrm{s}}$:

$$\int \Big[\boldsymbol{y}_{\mathrm{N}}(\boldsymbol{\theta}_{\mathrm{s}})^{\mathrm{T}}\,\boldsymbol{d}_l - g_l^{(\mathrm{v})}(\boldsymbol{\theta}_{\mathrm{s}})\Big]^2\,\mathrm{d}\boldsymbol{\theta}_{\mathrm{s}} \rightarrow \min\forall\,l. \tag{15}$$

By sampling all directions, e.g. by a numerical integration rule, we may minimize the vector norm instead of the integral $\Big\|\boldsymbol{Y}_{\mathrm{N}}^{\mathrm{T}}\,\boldsymbol{d}_l - \boldsymbol{g}_l^{(\mathrm{V})}\Big\|^2 \rightarrow \min\forall\,l.$. Here, $\boldsymbol{Y}_{\mathrm{N}} = [\boldsymbol{y}_{\mathrm{N}}(\theta_1),\ldots,\boldsymbol{y}_{\mathrm{N}}(\theta_{\mathrm{S}})]$ and $\boldsymbol{g}_l^{(\mathrm{v})} = [g_l(\theta_1),\ldots,g_l(\theta_{\mathrm{S}})]^{\mathrm{T}}$ contain values for various panning directions $\{\theta_s\}_{s=1\ldots\mathrm{S}}$. The least square error solution is $\boldsymbol{d}_l = (\boldsymbol{Y}_{\mathrm{N}}\,\boldsymbol{Y}_{\mathrm{N}}^{\mathrm{T}})^{-1}\boldsymbol{Y}_{\mathrm{N}}\;\boldsymbol{g}_l^{(\mathrm{v})}$. With VBAP and decoder matrices of all loudspeakers, and transposed back, it is

$$\boldsymbol{D}_{\mathrm{MVLAD}} = \boldsymbol{G}\;\boldsymbol{Y}_{\mathrm{N}}^{\mathrm{T}}(\boldsymbol{Y}_{\mathrm{N}}\,\boldsymbol{Y}_{\mathrm{N}}^{\mathrm{T}})^{-1}. \tag{16}$$

Here, $\boldsymbol{Y}_{\mathrm{N}}^{\mathrm{T}}(\boldsymbol{Y}_{\mathrm{N}}\,\boldsymbol{Y}_{\mathrm{N}}^{\mathrm{T}})^{-1}$ is normally a well-behaved mode-matching decoder, especially if the integration rule/sampling $\{\theta_s\}_{s=1\ldots\mathrm{S}}$ uniformly covers a huge set of all directions on the unit circle or sphere. Then the approximation $\boldsymbol{Y}_{\mathrm{N}}^{\mathrm{T}}(\boldsymbol{Y}_{\mathrm{N}}\,\boldsymbol{Y}_{\mathrm{N}}^{\mathrm{T}})^{-1} \approx \frac{2\pi\text{ or }4\pi}{\mathrm{S}}\boldsymbol{Y}_{\mathrm{N}}^{\mathrm{T}}$ holds, and MVLAD, Eq. (16), is equivalent to AllRAD, Eq. (7).

Appendix: Mode-Matching Decoding

Mode-matching decoding (MMAD) superimposes the loudspeakers' encoding coefficients Eq. (5) to fit those of any virtual source, cf. [5], i.e. $\boldsymbol{Y}_{\mathrm{N}}\,\boldsymbol{D}\,\boldsymbol{\varepsilon}_{\mathrm{N}} = \boldsymbol{\varepsilon}_{\mathrm{N}}$ using

$$\boldsymbol{D}_{\mathrm{MMAD}} = \boldsymbol{Y}_{\mathrm{N}}^{\mathrm{T}}(\boldsymbol{Y}_{\mathrm{N}}\boldsymbol{Y}_{\mathrm{N}}^{\mathrm{T}} + \alpha\boldsymbol{I})^{-1}, \tag{17}$$

which is unstable for $\alpha = 0$ in a semi-circular system, cf. Fig. 5. It gets stable by regularization, i.e. after less accurately matching the modes, by setting, e.g., $\alpha = \frac{1}{4}\frac{\mathrm{L}}{\pi}$. This is the reciprocal of a suitable scalar times the mean angular segment covered by each loudspeaker, cf. Fig. 6.

References

[1] Zotter, F.; Pomberger, H.; Noisternig, M.: Energy-preserving Ambisonic Decoding.
Acta Acustica united with Acustica **98** (2012), 37-47.

[2] Zotter, F.; Frank, M.: All-Round Ambisonic Panning and Decoding.
J. Audio Eng. Soc. **60** (2012), 807-820.

[3] Daniel, J.: Répresentation de champs acoustiques, application à la transmission et à la reproduction de scénes sonores complexes dans un contexte multimédia. Phd Thesis, Université Paris 6, 2001.

[4] Gerzon, M. A.: General Metatheory of Auditory Localization. 92nd AES Conv, 1992.

[5] Poletti, M. A.: A Unified Theory of Horizontal Holographic Sound Systems.
J. Audio Eng. Soc. **48** (2000), 1155-1182.

[6] Malham, D. G.; Myatt, A.: 3-D Sound Spatialization using Ambisonic Techniques.
Computer Music Journal **19** (1995), 58-70.

[7] Pulkki, V.: Virtual Sound Source Positioning Using Vector Base Amplitude Panning.
J. Audio Eng. Soc. **45** (1997), 456-466.