Moore, David J. and Wakefield, Jonathan P.

Designing Ambisonic Decoders for Improved Surround Sound Playback in Constrained Listening Spaces

Original Citation

Moore, David J. and Wakefield, Jonathan P. (2011) Designing Ambisonic Decoders for Improved Surround Sound Playback in Constrained Listening Spaces. In: 130th Audio Engineering Society AES Convention, 13-16 May 2011, London, UK.

Audio Engineering Society

Convention Paper

Presented at the 130th Convention
2011 May 13–16 London, UK

Designing Ambisonic Decoders for Improved Surround Sound Playback in Constrained Listening Spaces

David Moore[1], Jonathan Wakefield[2]

[1] Department of Computing and Creative Technologies, Glasgow Caledonian University, Glasgow, UK
j.d.moore@gcu.ac.uk

[2] School of Computing and Engineering, University of Huddersfield, Huddersfield, UK
j.p.wakefield@hud.ac.uk

ABSTRACT

Much research has been undertaken to optimise irregular 5-speaker Ambisonic decoders for idealised listening environments. In such environments speaker placement is not restricted and can conform to the ITU 5.1 standard. In domestic settings, the room shape, furniture and television positioning may restrict speaker placement. It is often the case that a compromised speaker layout is enforced by other domestic requirements. This paper seeks to derive Ambisonic decoders to optimise perceived localisation performance for these constrained asymmetrical speaker layouts. This work uses a heuristic search algorithm to derive decoder coefficients and simultaneously optimise speaker angle within specified bounds. Theoretical results are shown for different orders of newly derived Ambisonic decoders for typical domestic scenarios.

1. INTRODUCTION

It is clear that few people follow the ITU guidelines when setting up a surround sound loudspeaker arrangement in a domestic environment. The loudspeakers are often arranged in a manner convenient for the listener(s) and which does not cause safety risks with trailing cables. One of the main issues in setting up a surround sound system according to the ITU standard is the placement of the rear loudspeakers [1]. In a domestic environment, walls or furniture usually prevent the user from placing the rear loudspeakers in the correct positions. As a solution, users typically opt to fit them in convenient positions around the furniture. Often, especially in older buildings with protruding fireplaces, a convenient position for a television is the corner of the living room. Given that the most common use for surround sound listening is to accompany film and television images, the television placement impacts on the speaker placement.

Considering this, it may be concluded that a technique for reproducing surround sound in a domestic environment must be robust enough to cope with asymmetrical loudspeaker placement since the placement of loudspeakers according to standards is generally not user friendly. That is unless the user has a dedicated space for setting up the surround sound system.

Several recent papers have looked at the optimisation of Ambisonic decoders for irregular left-right symmetrical loudspeaker layouts, such as the ITU 5.1 configuration [2-4]. Clearly using Ambisonic decoders which are derived to give best performance for ITU 5-speaker layouts in asymmetric domestic speaker configurations will lead to degraded performance. In this paper we look at extending work in the area of decoder optimisation by deriving decoders for asymmetrical layouts that may typically be found in the home environment.

2. AMBISONIC DECODERS

2.1. Decoding

To playback Ambisonic audio, a re-composition of encoded audio components is made that takes into account the location of each loudspeaker. For example, for first-order horizontal Ambisonics the output of each loudspeaker is a weighted sum of the encoded audio:

$$S_i = \alpha_i W + \beta_i X + \gamma_i Y \tag{1}$$

where S_i is the gain of the ith loudspeaker, W, X, and Y are the encoded audio signals (captured by SoundField microphone for example), and α_i, β_i, and γ_i are the constant gain decoder parameters for the ith loudspeaker. Note that this equation assumes the decoder is frequency independent (i.e. a single set of coefficients is used). Frequency dependent decoders require a different set of parameters for low and high frequencies.

For higher order Ambisonic decoders the above equation needs to be expanded to take account of 2 additional encoded components per system order (e.g. second order Ambisonic decoders for horizontal loudspeaker layouts require 5 constant gain decoder parameters per speaker).

2.2. Deriving decoder parameters

It is well known that decoder parameters for regular arrangements of loudspeakers (e.g. square, hexagon etc.) can be derived analytically by matrix inversion [2]. However, for irregular arrangements, such as the ITU 5-speaker array, it becomes a more complicated matter. A non-linear system of equations needs to be mathematically solved in order to produce a set of decoder parameters with good localisation performance around the listener. An alternative method of producing decoder parameters for irregular layouts is to use a search algorithm to find a suitable set. In previous work the authors used this approach when deriving decoders for the ITU 5-speaker layout [5, 6].

When deriving decoders for the ITU 5-speaker array, it is possible to take account of the left-right symmetrical positioning of the loudspeakers in order to reduce the number of parameters that the search needs to adjust. For asymmetrical layouts, however, it is not possible to reduce the number of parameters - each speaker is independent. For a first order decoder for an asymmetric layout a total of 15 decoder parameters are required. For second order this increases to 25.

3. DECODER OPTIMISATION

3.1. Search and fitness function

In this work, multiple runs of a heuristic search algorithm known as the Tabu Search were employed to find a set of Ambisonic decoder parameters that maximise the localisation performance of a decoder according to a multi-objective fitness function. The fitness function objectives are identical to those used in previous work [5, 6] and are based upon the velocity vector and energy vector, which are able to quantify a decoder's localisation performance at low and mid to high frequencies respectively [7]. The magnitude of the vectors indicates the perceived quality of a phantom sound source, whereas the angle of the vectors indicates its perceived direction.

In summary, the objectives used in this work aim to meet the following:

- Velocity vector magnitude is as close to the optimum magnitude as possible
- Energy vector magnitude is as close to the optimum magnitude as possible

- Velocity vector angle is as close to the correct sound source angle as possible
- Energy vector angle is as close to the correct sound source angle as possible
- The velocity vector and energy vector angles are as closely matched as possible
- Mid/high frequency volume is equal around the listener

When optimising decoders for left-right symmetrical layouts such as the ITU array the objectives need only be checked on one side of the sound stage (i.e. between 0° and 180°). However, when optimising for asymmetrical layouts, as in this work, the objectives needed to be checked on both sides of the sound stage to account for different performance values on the left and right sides. No weightings were assigned to the objectives (i.e. they were given equal importance in the search) but range-removal was used to ensure the search was not biased towards objectives with a larger range of possible values (as in [5]).

3.2. Speaker angle optimisation

In this work we also consider the optimisation of the loudspeaker angles in the search. Rather than specifying an asymmetrical layout with fixed loudspeaker positions, we restrict each speaker position within an upper and lower angle bound which allows the search to take account of some flexibility in speaker placement within the domestic environment. Our implementation, therefore, simultaneously optimises the decoder coefficients and determines the best speaker angles within the imposed restrictions of the environment.

Please note that variations in distance of speakers from the listener position can be dealt with via gain compensation and time delays (in fact current commercial systems already provide this feature). Therefore in this work we only consider optimising placement of speakers by angle.

4. RESULTS

A series of searches was undertaken for four different surround sound loudspeaker setups. Each setup might typically be found in a domestic space with furniture constraints.

In all scenarios, the search was permitted to move the front-left loudspeaker (L) between 10° and 20°, and the front-right loudspeaker (R) between 340° and 350°. The centre speaker (C) was always fixed at 0° because this speaker is normally positioned below or above the television (i.e. in the direction the listener is facing). Walls as well as chairs and other furniture often determine positioning of the rear loudspeakers so four contrasting configurations were investigated.

In all four scenarios Ambisonic decoders for first order and second order were derived. When analysing performance in each scenario, we compare all asymmetric optimised decoders with a decoder of the same order optimised for the standard ITU 5-speaker layout allowing us to evaluate the potential improvement over an ITU optimised decoder.

4.1. Scenario 1

In scenario 1 the living room sofa is positioned away from the lower wall and the left wall, which allows maximum freedom when placing the rear loudspeakers behind the listening position (see Figure 1). In the search, the left-surround speaker (LS) was permitted to move between 70° and 165°, and the right-surround speaker (RS) between 210° and 330°.

A first order decoder was produced by the search with speakers positions: C(0°), L(20°), LS(134°), RS(226°) and R(340°). Overall this configuration is not too far away from the standard left-right symmetrical ITU configuration. The regular spacing of the rear speakers is not a surprising result given that optimum consistent localisation performance around the listener can be achieved when the loudspeakers are evenly distributed. Similarly, it is not surprising the left and right speakers have been pushed as far from the centre speaker as possible. The derived first order decoder gives an 18% overall improvement according to the fitness function.

The figures included at the end of this paper show the performance of the decoders for each scenario. The new decoders are shown as bold lines on the plots and, for comparison purposes, the standard ITU 5-speaker decoder is shown on the plots as faded lines. Speaker positions are illustrated as black circles, and circular lines represent magnitudes.

When examining the velocity vector and energy vector magnitudes in Figure 2 it is clear that improvements have been made at the front of the system, at the cost of performance at the rear. For the velocity vector angle a slight loss in performance is apparent overall but the energy vector angle is much improved at the rear. Furthermore, the overall match between the velocity vector angles and energy vector angles is much better around the 360° sound stage for the decoder optimised for this scenario. A small improvement has been made for the perceived volume (see Figure 3).

For second order, the search produced a decoder with speakers at: C(0°), L(20°), LS(130°), RS(230°) and R(340°). Overall only minor improvements have been made to the localisation performance of this decoder. The most significant improvement has been with the velocity vector and energy vector angles which are closer to their target angles, and match more closely with each other. The derived second order decoder gives a 6% overall improvement.

When comparing first order and second order it is clear that second order gives the best overall performance for the velocity vector and energy vector. It is also apparent from inspection of the speaker gains (Figure 3 and Figure 5) that the centre speaker is used more by the second order decoder because of the narrower responses that can be generated for second order.

4.2. Scenario 2

In scenario 2 the sofa is away from the left wall but is up against the lower wall (see Figure 6). This allows freedom for placement of LS but limits the area that RS can be placed. In this scenario the RS is always in front of a forward facing listener. In the search, LS was allowed to move between 70° and 165°, and RS between 300° and 330°.

A first order decoder was produced by the search with speakers at: C(0°), L(20°), LS(163°), RS(300°) and R(343°). In this case the search has pushed the LS and RS speakers almost as far back as possible to reduce the gaps at the rear and sides of the sound stage. The derived first order decoder gives a 37% overall improvement according to our function.

When comparing this decoder to the decoder optimised for the standard ITU layout, the energy vector magnitude is similar but improved to the sides. However, the velocity vector magnitude is reduced which must be the result of a tradeoff with improving other performance characteristic in the search (see Figure 7). The vector angles are more closely matched with the desired angles and with each other. A significant improvement has also been made for the perceived volume (see Figure 8).

The search produced a second order decoder with speakers at almost identical positions to the first order decoder: C(0°), L(20°), LS(160°), RS(300°) and R(345°). The overall localisation performance characteristics are quite different, however. The derived second order decoder gives a 32% overall improvement according to the fitness function.

When compared to the standard ITU decoder the velocity vector and energy vector magnitudes are substantially better at the front of the system (see figure 9). This is also the case for the vector angles and the perceived volume is much more even around the listener (see figure 10).

4.3. Scenario 3

In scenario 3 the sofa is up against the left wall but moved away from the lower wall (see Figure 11). This allows freedom for placement of RS but limits the area that LS can be placed. In this scenario LS is able to move just into the rear half of the 360° sound stage if required. In the search LS was permitted to move between 70° and 95°, and RS between 210° and 330°.

A first order decoder was produced by the search with speakers at: C(0°), L(20°), LS(95°), RS(210°) and R(340°). Like scenario 2, the search has pushed LS and RS as far back as possible to reduce the gaps at the rear and sides of the 360° sound stage. The derived first order decoder gives a 25% overall improvement according to the fitness function.

The first order decoder in this scenario gives similar velocity vector magnitude performance overall. There has been a loss in performance at the left side but an improvement at the right side (see Figure 12). Interestingly, the velocity vector angle is better overall - performance at the direct front is slightly worse but better at the sides and rear (see Figure 13). The energy vector magnitude is improved at the front-right of the system. The perceived volume is also more consistent overall (see Figure 13).

For the second order decoder the search placed the loudspeakers at: C(0°), L(20°), LS(95°), RS(214°) and R(340°). This is almost identical to first order decoder apart from RS where there is a small difference of 4°.

The derived second order decoder gives a 20% overall improvement according to the fitness function.

The velocity vector magnitude for the asymmetric optimised decoder is better at the right rear at the cost of performance to the front-left (see Figure 14). The velocity vector angle around the left side of the 360° sound stage is significantly improved. The energy vector angle is improved at the front-left and both vectors match more closely overall. Figure 15 shows a marginal improvement in terms of perceived volume.

4.4. Scenario 4

In the final scenario the sofa was placed against both the left wall and lower wall (see Figure 16). The represents the most constrained configuration to be tested in terms of placement of the rear loudspeakers. In the search, LS was permitted to move between 70° and 95°, and RS between 300° and 330°.

For the first order decoder and second order decoders produced by the search the speakers are at: C(0°), L(18°), LS(95°), RS(300°) and R(340°). In this case the search has pushed LS and RS as far back as possible in order to try and minimise the performance loss at the rear. The derived first order decoder gives a 20% overall improvement according out function, whereas the second order decoder gives a 23% improvement

Both vector angles have been improved at the front. As expected, the velocity and energy vectors indicate that sound sources cannot be positioned at the rear of the listener. This was the case for both orders of decoder derived (see Figure 17 and Figure 19). The vector angles for the rear point towards the front so sound that should have been at the rear will pan appropriately from left to right. It is not obvious why this would represent a good solution and requires further investigation.

Overall performance is better for both orders of decoder and the vectors angles in particular match more closely with the desired sources angles and each other. The volume is more consistent around the listener in both cases (see Figure 18 and Figure 20).

5. CONCLUSIONS

This paper has focused on the optimisation of Ambisonic decoders for first and second order in constrained domestic listening spaces. It presented results for four contrasting surround sound setups that might typically be found in the home environment.

The results show that in all scenarios tested, the decoder specifically optimised for the asymmetric layout performs better than a decoder for the ITU layout.

6. FUTURE WORK

In future work we plan to evaluate these decoders in listening tests and to look at optimising decoders for multiple listener positions in constrained listening environments.

7. REFERENCES

[1] Theile, G. “The New Sound Format "3/2-Stereo". Presented at the 10th International Audio Engineering Society Conference, Berlin, (1993).

[2] Heller, A. J., Benjamin, E. and Lee, R., “Design of Ambisonic Decoders for Irregular Arrays of Loudspeakers by Non Linear Optimization”, Presented at the 129th Audio Engineering Society Convention, San Francisco (2010).

[3] Wiggins, B., “The Generation of Panning Laws for Irregular Speaker Arrays using Heuristic Methods”, Presented at the 31st International Audio Engineering Society Conference, London (2007).

[4] Moore, D., Wakefield, J. P., “A Design Tool To Produce Optimised Ambisonic Decoders”, Presented at the 40th International Audio Engineering Society Conference, Tokyo (2010).

[5] Moore, D. Wakefield, J. P., “The Design of Improved First Order Ambisonic Decoders by the Application of Range-Removal and Importance in a Heuristic Search Algorithm”, Presented at the 31st International AES Conference, London (2007).

[6] Moore, D., Wakefield, J. P., “Exploiting Human Spatial Resolution in Surround Sound Decoder Design”, Presented at the 125th Audio Engineering Society Convention, San Francisco (2008).

[7] Gerzon, M. A., “General Metatheory of Auditory Localization,” Presented at the 92nd Audio Engineering Society Convention, Vienna (1992).

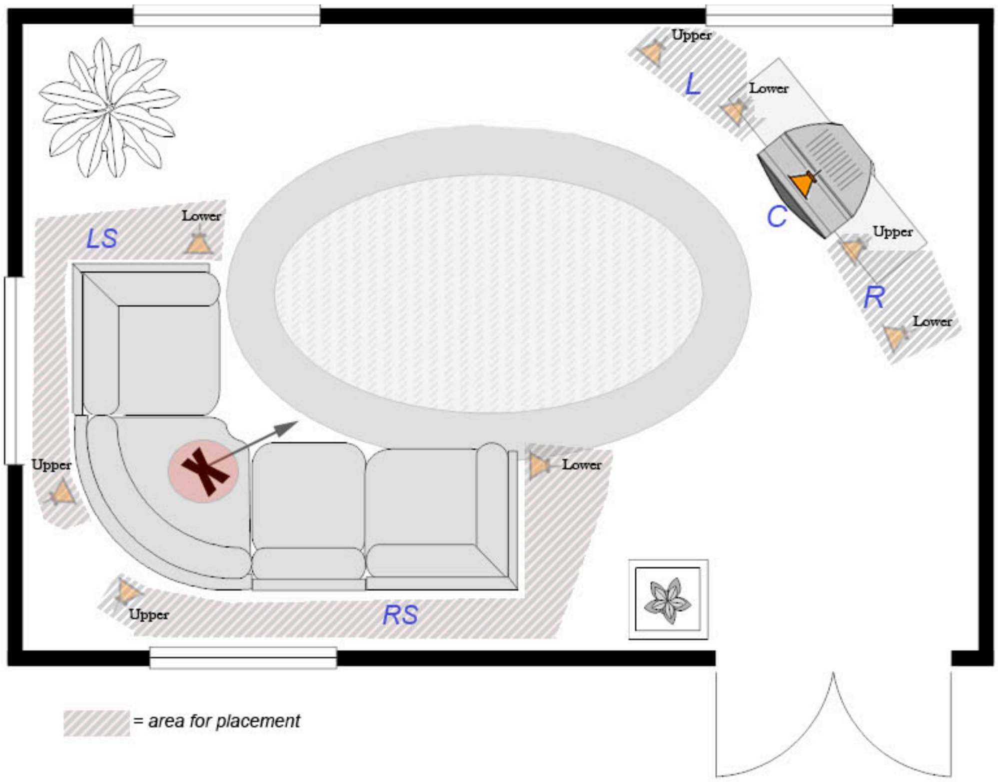

Figure 1: The loudspeaker layout for scenario 1. The listener position is marked with a cross and the area for potential speaker placement is shaded. The sofa is away from the walls in this scenario allowing the speakers to be placed at several different positions at the sides and rear if required.

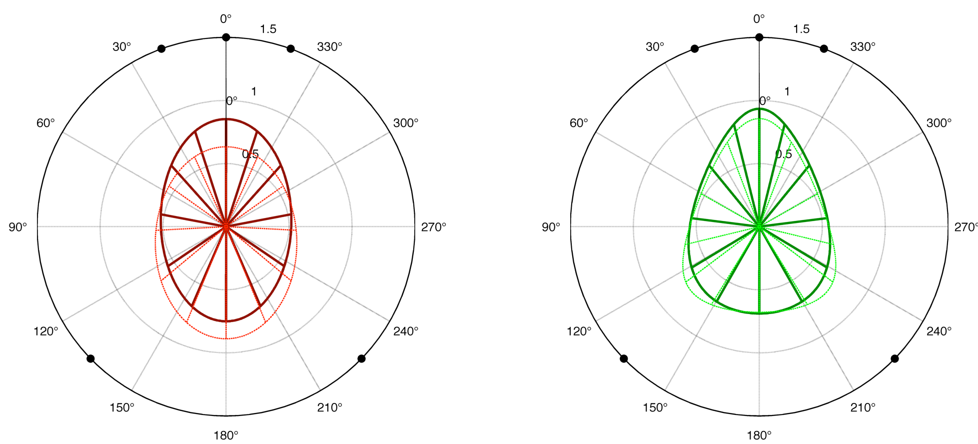

Figure 2: The velocity vector and energy vector for the first order decoders in scenario 1. Vector angles are displayed every 30 degrees and the loudspeakers are shown as black circles.

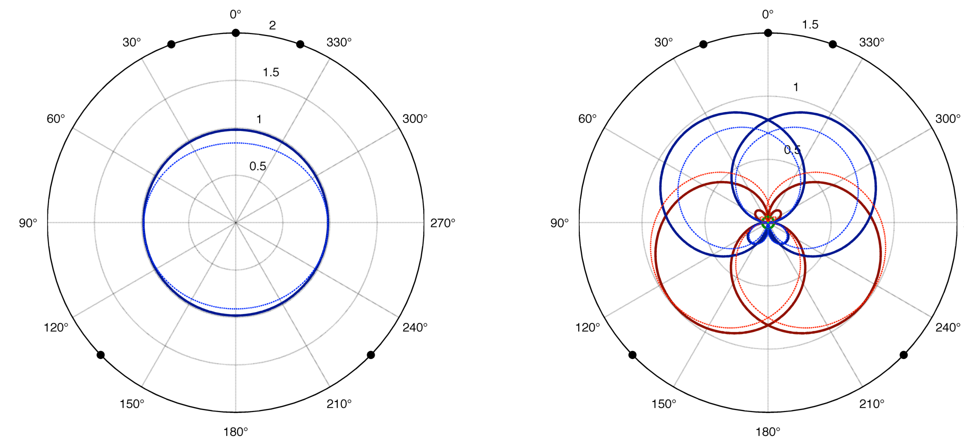

Figure 3: Perceived volume (energy) and speaker gains by angle for the first order decoders scenario 1

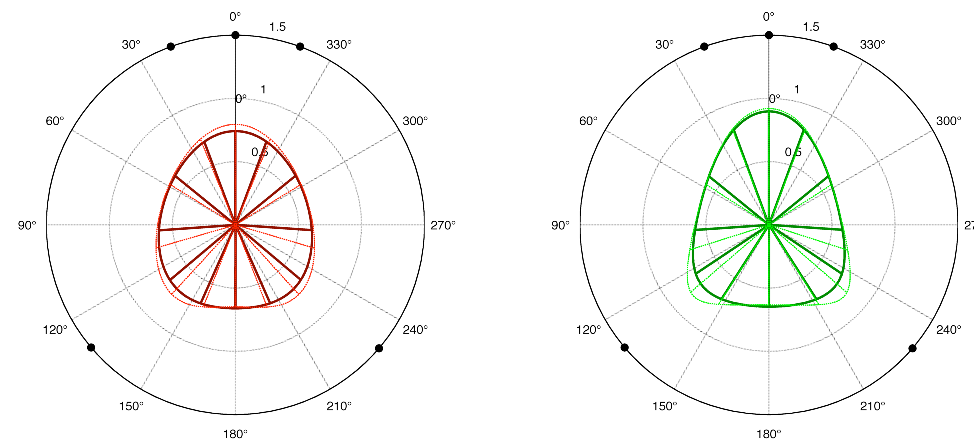

Figure 4: The velocity vector and energy vector for the second order decoders in scenario 1

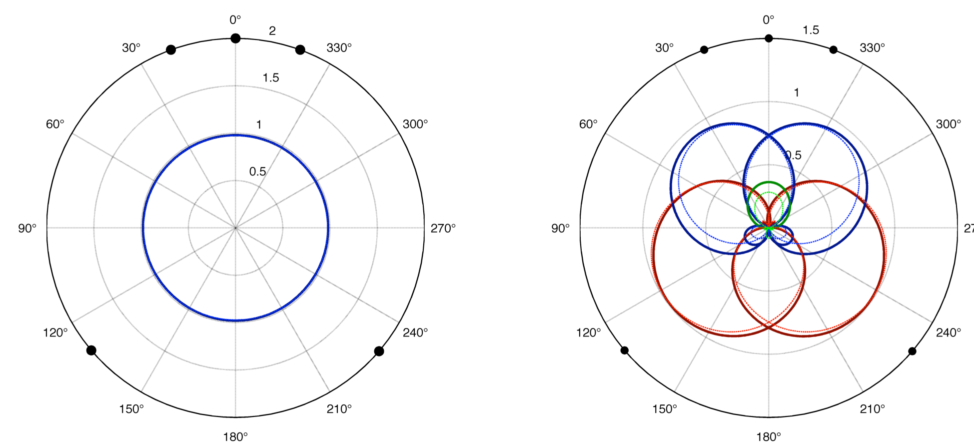

Figure 5: Perceived volume (energy) and speaker gains by angle for the second order decoders scenario 1

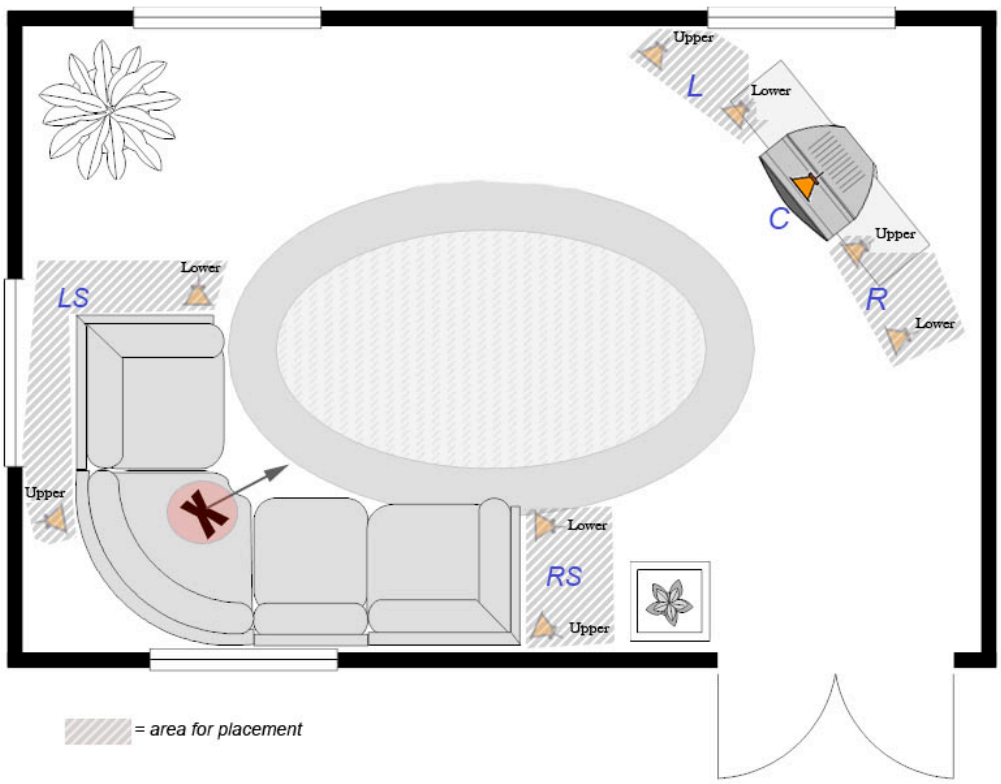

Figure 6: The loudspeaker layout for scenario 2. The listener position is marked with a cross and the area for potential speaker placement is shaded. The sofa is away from the left wall in this scenario allowing the left surround speaker to be placed at several different positions at the side and behind the listener. The right surround speaker is limited in movement and is in front of the listener.

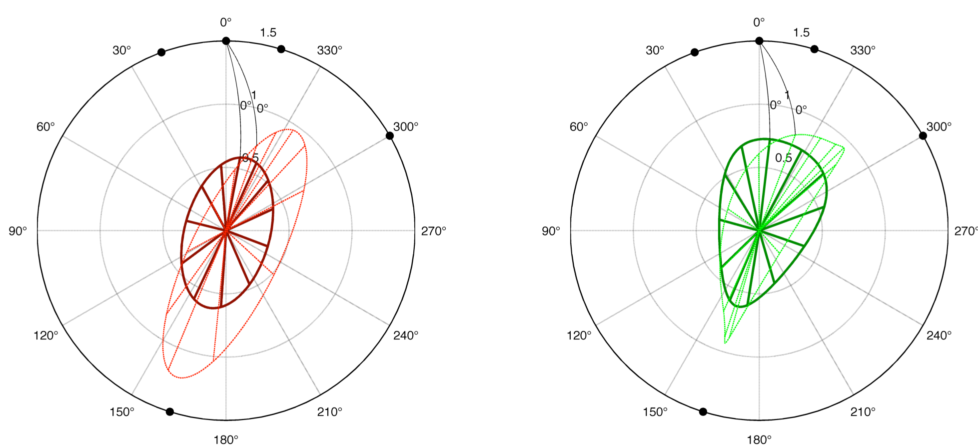

Figure 7: The velocity vector and energy vector for the first order decoders in scenario 2.

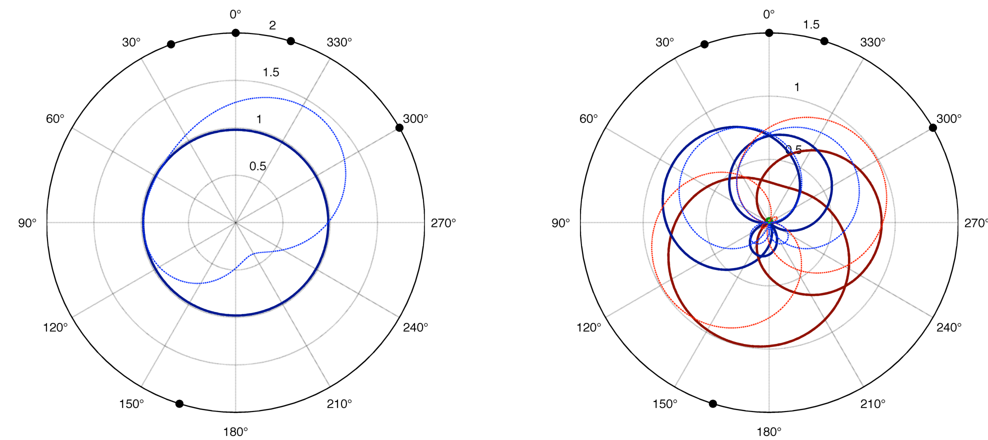

Figure 8: Perceived volume (energy) and speaker gains by angle for the first order decoders scenario 2

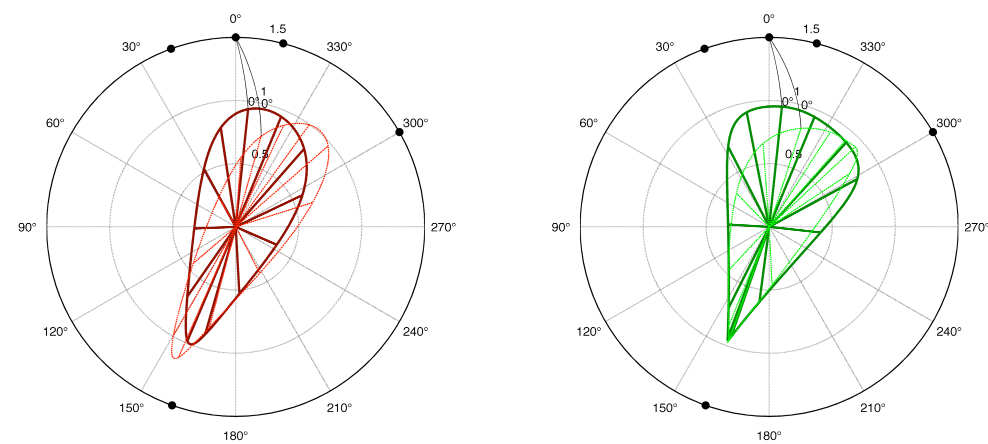

Figure 9: The velocity vector and energy vector for the second order decoders in scenario 2

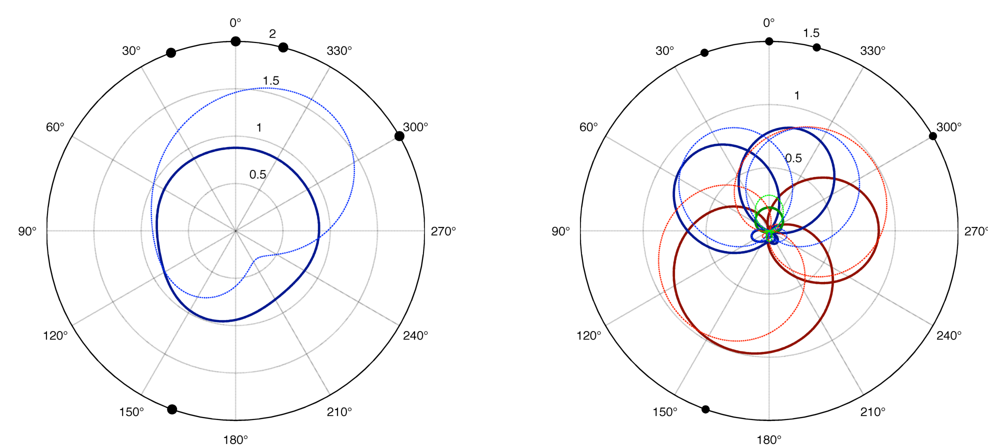

Figure 10: Perceived volume (energy) and speaker gains by angle for the second order decoders scenario 2

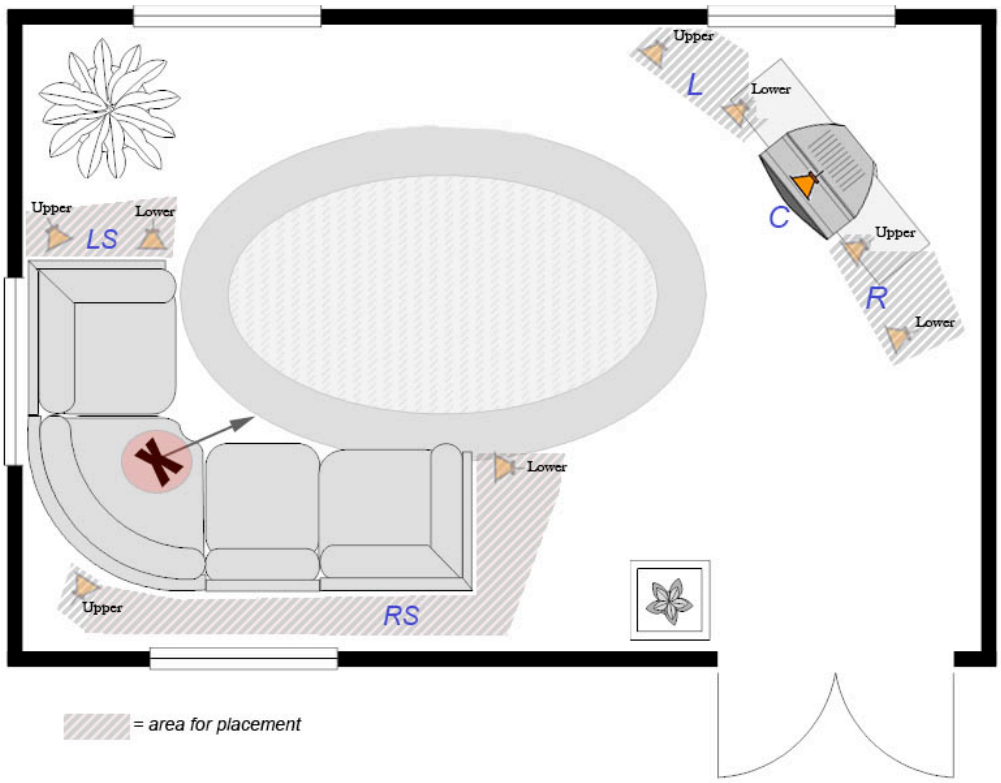

Figure 11: The loudspeaker layout for scenario 3. The listener position is marked with a cross and the area for potential speaker placement is shaded. The sofa is away from the bottom wall in this scenario allowing the right surround speaker to be placed at several different positions to the side and behind the listener. The left surround speaker is limited in movement and is in front of the listener.

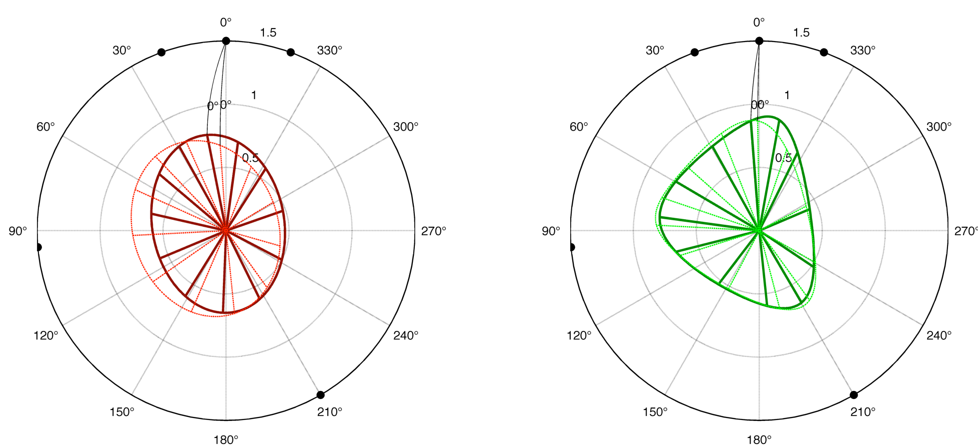

Figure 12: The velocity vector and energy vector for the first order decoders in scenario 3

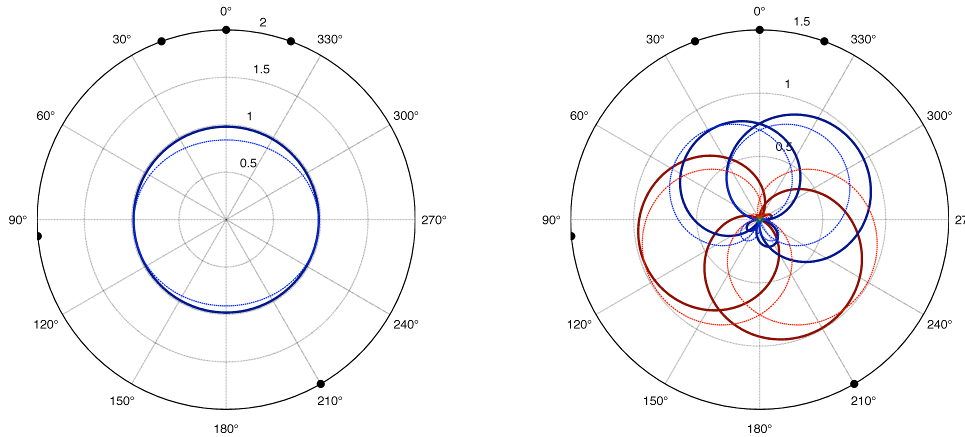

Figure 13: Perceived volume (energy) and speaker gains by angle for the first order decoders scenario 3

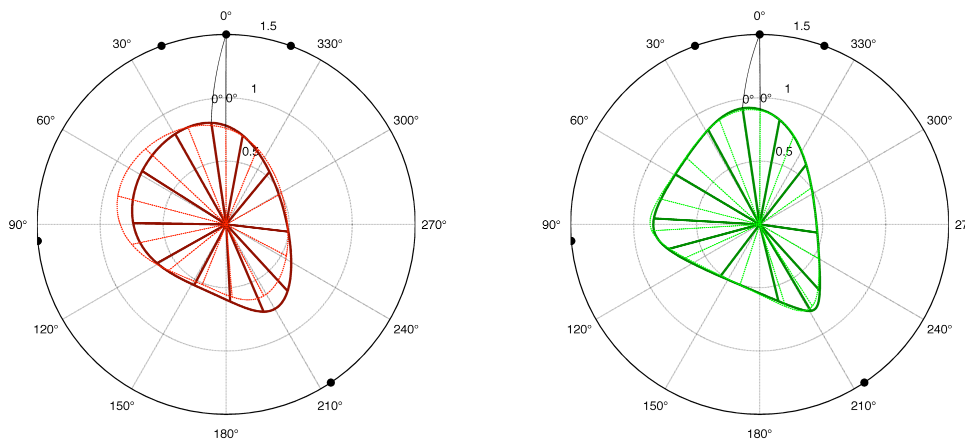

Figure 14: The velocity vector and energy vector for the second order decoders in scenario 3

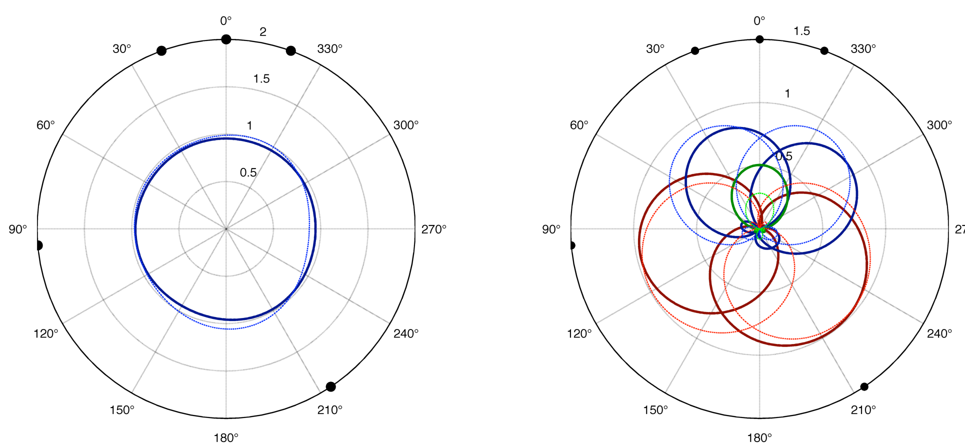

Figure 15: Perceived volume (energy) and speaker gains by angle for the second order decoders scenario 3

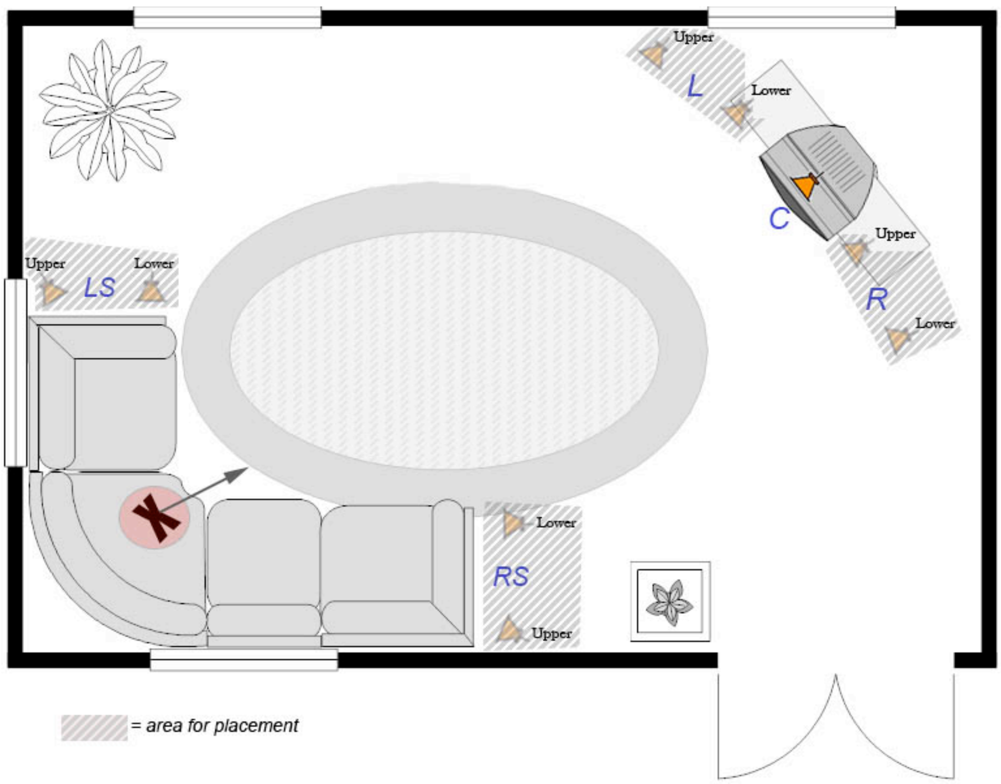

Figure 16: The loudspeaker layout for scenario 4. The listener position is marked with a cross and the area for potential speaker placement is shaded. The sofa is against both walls in this scenario so left and right surround speaker placement to the front of the listener.

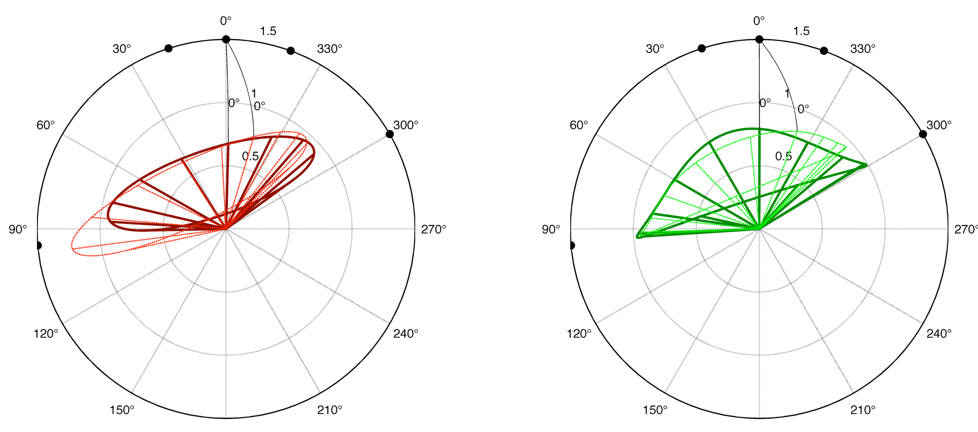

Figure 17: The velocity vector and energy vector for the first order decoders in scenario 4

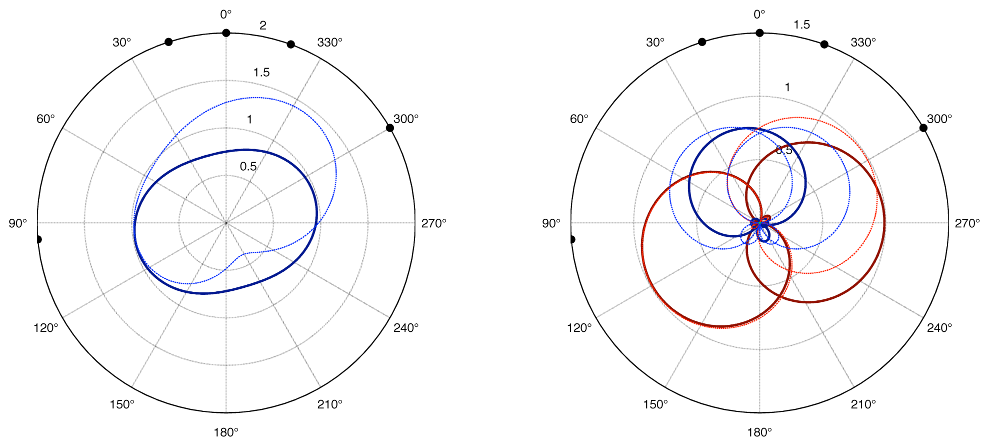

Figure 18: Perceived volume (energy) and speaker gains by angle for the first order decoders scenario 4

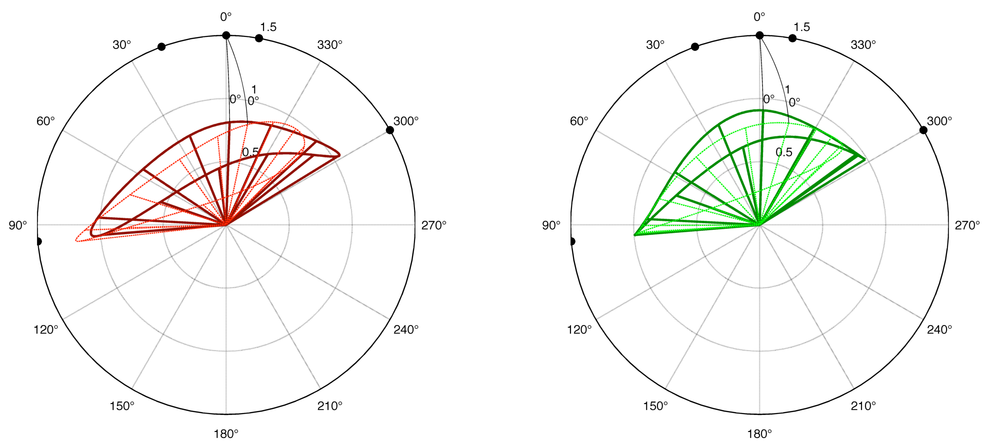

Figure 19: The velocity vector and energy vector for the second order decoders in scenario 4

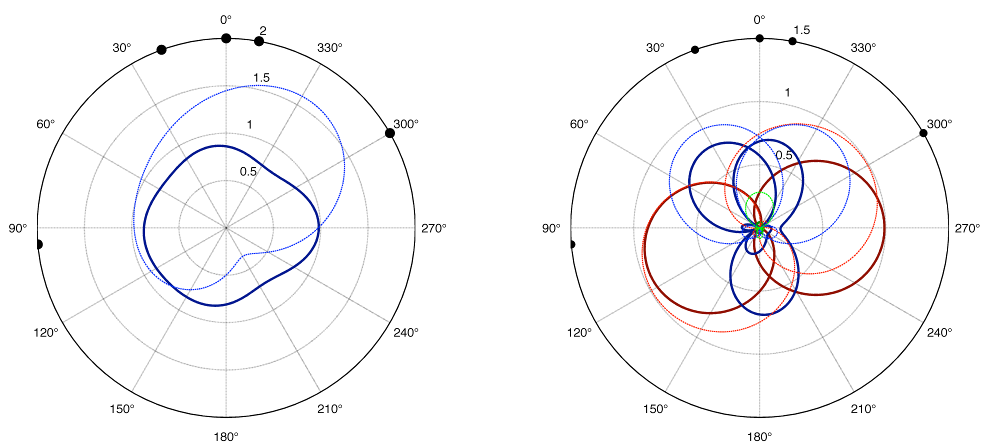

Figure 20: Perceived volume (energy) and speaker gains by angle for the second order decoders scenario 4