University of Huddersfield Repository

Moore, David J. and Wakefield, Jonathan P.

Optimisation of the Localisation Performance of Irregular Ambisonic Decoders for Multiple Off-Centre Listeners

Original Citation

Moore, David J. and Wakefield, Jonathan P. (2010) Optimisation of the Localisation Performance of Irregular Ambisonic Decoders for Multiple Off-Centre Listeners. In: 128th Audio Engineering Society Convention, 22-25 May 2010, London, UK.

Audio Engineering Society

Convention Paper

Presented at the 128th Convention
2010 May 22–25 London, UK

Optimisation of the Localisation Performance of Irregular Ambisonic Decoders for Multiple Off-Centre Listeners

David Moore[1], Jonathan Wakefield[2]

[1] Glasgow Caledonian University, Glasgow, Lanarkshire, G4 0BA, UK
j.d.moore@gca.ac.uk

[2] University of Huddersfield, Huddersfield, West Yorkshire, HD1 3DH
j.p.wakefield@hud.ac.uk

ABSTRACT

This paper presents a method for optimising the performance of irregular Ambisonic decoders for multiple off-centre listeners. New off-centre evaluation criteria are added to a multi-objective fitness function, based on auditory localization theory, which guides a heuristic search algorithm to derive decoder parameter sets for the ITU 5-speaker layout. The new evaluation criteria are based upon Gerzon's Metatheory of Auditory Localisation and have been modified to take into account off-centre listening positions. The derived decoders exhibit improved theoretical localisation performance for off-centre listeners. The theoretical results are supported by initial listening test results.

1. INTRODUCTION

There has been much research aimed at improving the localisation performance of Ambisonic systems at the central listening position. However, few studies exist which look at improving the localisation performance of Ambisonic systems in multiple off-centre listening positions. There is clearly a need for research in this area, as many systems will be used for playing audio to a distributed audience – for example an audience in a cinema or auditorium. Improving the audience experience at multiple off-centre listening positions is a step towards increasing the size of the sweet spot. This paper describes a method for optimising the localisation performance of Ambisonic decoders for multiple off-centre listeners. The work specifically focuses on developing decoders for the current industry standard ITU 5-speaker system where loudspeakers are at a constant radial distance from the central listening point. It should be noted, however, that although the focus is on the ITU layout, the method presented could easily be extended to optimising decoders for other regular or irregular loudspeaker layouts.

2. REVIEW OF RELEVANT LITERATURE

Arguably the most commonly referenced work on off-centre Ambisonic surround sound is by Malham [1]. Malham describes informally several personal experiences of using Ambisonics for playback over different large-scale regularly spaced surround sound rigs. One of the major problems he identifies with delivering surround sound in this way is that at non-central listening positions the sound image is drawn towards the nearest loudspeaker. This happens because a listener in an off-centre position will be nearer or further away from some loudspeakers resulting in unintended time differences and level differences between sound waves arriving from each loudspeaker. This leads to the loss of temporal synchronisation of the contributing sound waves and also a sound intensity bias in the direction of the nearest loudspeaker. As a result phantom images can be distorted, or in worst case scenarios, lost completely.

Malham also identified another problem specific to off-centre Ambisonic playback. He observed that first order Ambisonic decoders designed according to one of Gerzon's theorems had poor localisation performance in off-centre positions. He noted the reason for this was because first order decoders play sound out of all loudspeakers simultaneously. As a result of this, listeners in off-centre listening positions perceived what Malham terms a "bounce back" effect where sound would effectively be heard in two different locations. In order to remove this effect, Malham later devised the 'Cardioid' decoder where the secondary lobe of the virtual microphone polar response is removed [2]. However, although this decoder removes the problems of bounce back, it leads to a significant decrease in overall localisation performance at the sweet spot. For example, studies have reported Cardioid decoders as having poor localisation performance with sound images sounding too diffuse [3, 4]. An alternative to using cardioid decoders is to use higher order decoders. Higher order decoders reduce the sound sent to the opposite speakers thus reducing the bounce back effect [5]. In this work fourth order is adopted for this reason.

Recent work by Poletti has introduced a method of improving the performance of surround sound systems away from the central listening position [6]. Poletti's work involves using a least-squares pressure matching method for approximating an optimal fourth order decoder for the ITU 5-speaker layout. Basically, the least-squares approach involves matching the sound pressure at several points in the listening area between an ideal soundfield and the decoded soundfield. One of the advantages of this method is soundfields can be analysed over an area rather than a single point. However, although the pressure matching approach is able to produce theoretically robust solutions, it does not take into account what the listener will perceive. In this work the decoder is designed using theory of what a listener will perceive at different points in the listening area.

At the time of writing this paper, the work by Poletti is the only work that details the optimisation of surround sound systems for the ITU 5-speaker layout away from the central listening position.

3. OFF-CENTRE OPTIMISATION

In this work a heuristic search algorithm known as the Tabu Search is employed to find a set of Ambisonic decoder parameters that maximise the localisation performance of a decoder according to a multi-objective fitness function. This approach has been used previously when optimising decoders for a single central listening position [7-8].

The novel element of this research is the fact that new optimisation criteria have been added to the fitness function to allow localisation performance to be measured in off-centre listening positions (in addition to the central position). The newly developed criteria are based upon two psychoacoustic models defined by Gerzon (i.e. the velocity vector and energy vector) [9]. In this work, both vectors are mathematically adjusted to take into account the fact that the loudspeakers are at different distances to an off-centre listener and also at different angles.

The following sections provide a definition of the velocity and energy vectors and their modification.

3.1. Velocity and energy models

A decoder's localisation performance can be measured at the central position using the velocity vector and the energy vector. Basically, the velocity vector can be used for predicting the low frequency localisation performance of a sound reproduction system, and the energy vector can be used for predicting the mid to high frequency localisation performance. The vector magnitudes indicate the "quality" of the reproduced sound image and the vector angles indicate the reproduced sound source's angular position. A magnitude of unity is optimal for both vectors.

$$r_V^x = \sum_{i=1}^{n} S_i \cos(\theta_i)/P \quad (1.)$$

$$r_V^y = \sum_{i=1}^{n} S_i \sin(\theta_i)/P \quad (2.)$$

$$r_E^x = \sum_{i=1}^{n} S_i^2 \cos(\theta_i)/E \quad (3.)$$

$$r_E^y = \sum_{i=1}^{n} S_i^2 \sin(\theta_i)/E \quad (4.)$$

where

$$P = \sum_{i=1}^{n} S_i \quad (5.)$$

$$E = \sum_{i=1}^{n} S_i^2 \quad (6.)$$

where P is the pressure, E is the energy, r_V^x is the velocity vector in the x direction, r_V^y is the velocity vector in the y direction, r_E^x is the energy vector in the x direction, r_E^y is the energy vector in the y direction, n is the number of loudspeakers, θ_i is the angular position of the i^{th} loudspeaker and S_i represents the gain of the i^{th} loudspeaker.

3.2. Off-centre optimisation criteria

In order for sound localisation performance to be measured in off-centre listening positions, the velocity vector and energy vector were adjusted. This adjustment takes into account the fact that the loudspeakers are at different distances to an off-centre listener and also at different angles

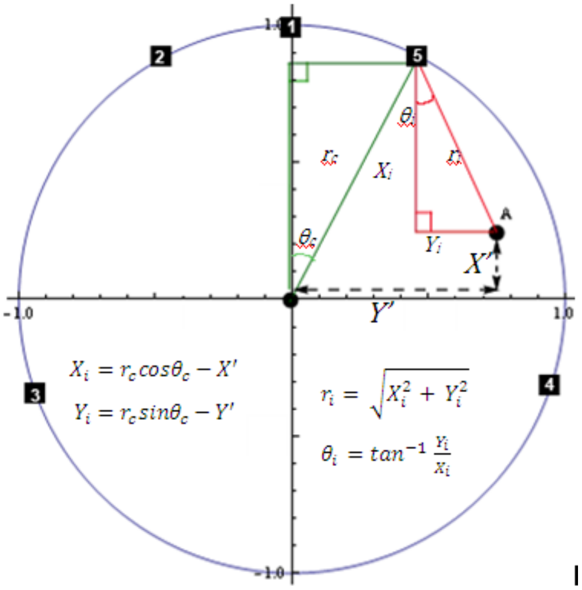

Figure 1 The distance and angle of each loudspeaker changes according to position

It is clear from figure 1 that sound arriving at the off-centre position A from loudspeakers 5 and 4 will be louder than sound emitted at the same level from loudspeakers 2 and 3. This change in sound level with distance can be modelled using the inverse square law. The inverse square law says that sound intensity decreases as the distance to the source increases. This translates to an inverse relationship between gain and distance. To take account of this the following gain factor was used:

$$g_i = \frac{1}{r_i} \quad (7.)$$

where g_i is the difference in sound pressure level for the *i*th loudspeaker and r_i is the distance to the *i*th loudspeaker.

When calculating the pressure, velocity vector, energy and energy vector in an off-centre position this gain factor is directly applied to all loudspeaker gains i.e.

$$S_i = g_i S_i^{original} \quad (8.)$$

When estimating sound localisation from the centre point, the optimum length of the velocity vector and energy vector is unit magnitude. However, in off-centre position the optimal length of both vectors changes according to the vector from the origin, and also the vector to the sound source. The optimal vector angles will also be different at each listening position (see θ_i in figure 1).

This adjustment enables a 'local' velocity vector and energy vector to be derived that predicts the perceived location and quality of a virtual sound source for an off-centre listener.

Please note that this approach is only concerned with adjusting each of the loudspeaker gains, consequently time delay compensation is considered outside the scope of this work.

3.3. Fitness function objectives

Seven objectives were used in the fitness function. The objectives were checked at each listening position and at each sound source angle between 0 degrees and 180 degrees in steps of 1 degree. Only one side of the soundfield need be tested as the speaker setup is left-right symmetrical.

In summary, the objectives 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
- Low frequency volume is equal around the listener
- Mid/high frequency volume is equal around the listener

A full mathematical definition of the objectives is provided in previous papers by the authors [10–11].

When implementing this fitness function a technique known as 'range-removal' was used to scale all of the objectives to the same range of values [0, 1]. This was to prevent certain objectives biasing the search [12]. The method for achieving this range-removal was to store the maximum and minimum values encounter for each objective during the search.

In this implementation range-removal is 'position dependent' meaning different scaling is used at each evaluated position to take account of the different possible objective ranges at each position.

It should be noted that the runtime performance of this fitness function is highly dependent on the number of listening positions checked and the number of angles checked around the sound stage.

4. EVALUATION

4.1. Theoretical results

To test the new optimisation criteria a number of searches for decoder parameters were undertaken. The goal was to produce a fourth order frequency-independent decoder with improved performance in off-centre listening positions. When deriving the decoder, 9 evenly distributed listening positions were evaluated in the fitness function: the centre point and 8 surrounding positions (see figure 2).

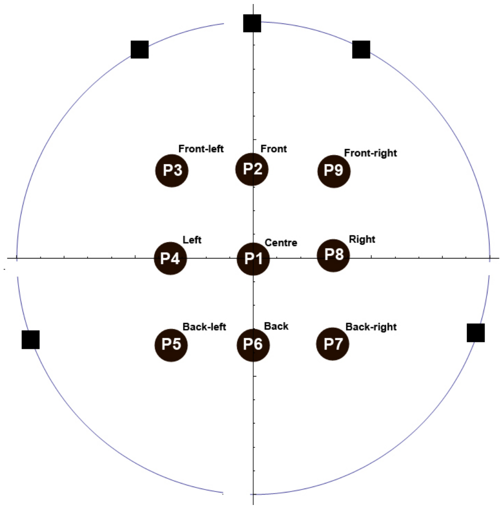

Figure 2 The distance and angle of each loudspeaker changes according to position

Positions 2, 4, 6 and 8 are at 35% of the loudspeaker rig radius whereas positions 3, 5, 7 and 9 are at 50% of the loudspeaker rig radius.

It is important to note that there is a direct performance trade-off when using any technique to improve off centre localisation performance. Improving the velocity vector or energy vector at one position can have an adverse effect on performance at another position because of the change in loudspeaker level.

Figure 3 plots the mean error of the velocity vector and energy vector magnitude and angle for each position. In each subplot the mean errors for a fourth order decoder and a first order decoder derived without the off-centre criteria are shown for comparison. This figure demonstrates that the off-centre optimised decoder is able to produce better performance at a greater number of positions than the other decoders. Of particular note are the consistently low vector magnitude errors across all positions and the improved vector angles at listening position on the left side of the system and the right side of the system.

Figure 4 and figure 5 plot the local velocity vector for the fourth order off-centre optimised decoder at the 9 listening points evaluated in the improved fitness function. The vectors are shown at 0°, 30°, 60° and 90° in figure 4 while the vectors are shown at 120°, 150° and 180° in figure 5. An ideal vector is also indicated at each position.

The velocity vector performance of the off-centre optimised decoder is better at most positions and for most angles. Take, for example, when a source is panned to 120°. The local velocity vectors are closer to the ideal vectors (in terms of magnitude and angle) in nearly all listening positions.

Figure 6 and figure 7 show the local energy vectors for the decoders. The difference in performance is again clear. For instance, when a source is panned to the front (0°) the local energy vector is closer to the ideal vector at all positions for the off-centre optimised decoder. For the other decoders, the vector angles are biased towards the front left loudspeaker when evaluated from positions 3, 4 and 5, and the front right loudspeaker when evaluating from positions 7, 8 and 9.

The most problematic area of the sound stage for all decoders is at the rear (see 150° and 180°). In off-centre positions the local velocity vectors and energy vector pull away from their ideal direction towards the nearest loudspeaker. This result was expected considering the large angular spacing between the rear loudspeakers.

4.2. Preliminary listening tests

A series of listening tests was undertaken to test the decoder produced by the off-centre optimisation method. In the tests listeners were required to localise different low and high frequency sound sources from different listening positions. The results from the test show that the decoder optimised using the method presented in this paper gives better off-centre performance than decoders optimised only for the central listening position. Furthermore, the off-centre decoder gives comparable performance to the decoder derived by Poletti. For a full account of these experiments and their results the reader is referred to [13].

5. CONCLUSIONS

A new method was developed to enable the localisation performance of Ambisonic decoders to be evaluated in off-centre listening positions.

This method was incorporated into a fitness function allowing a search algorithm to produce decoders for irregular loudspeaker layouts with improved off-centre localisation performance.

The new off-centre decoder derived in this work shows improved theoretical performance over decoders just optimised for the centre position. Listening tests described elsewhere support the results [13].

6. REFERENCES

[1] Malham, D.G, “Experience with Large Area 3-D Ambisonic Sound Systems”, Proceedings of the Insitute of Acoustics, 14(5) pp.209 – 216, 1992

[2] Furse, R. and Malham, D.G. http://www.muse.demon.co.uk/3daudio.html, last accessed 12th March 2010

[3] Benjamin, E., Lee, R. & Hellar, A. “Localization in Horizontal-Only Ambisonic Systems.” Presented at the 121st Audio Engineering Society Convention, San Francisco, USA, 2006

[4] Guastavino, C., Larcher, V., Catusseaus G. & Boussards, P. "Spatial Audio Quality Evaluation: Comparing Transaural, Ambisonics and Stereo." Presented at the 13th International Conference on Auditory Display, Montreal, Canada, 2007.

[5] Daniel, J. & Moreau, S. "Further Study of Sound Field Coding with Higher Order Ambisonics." Presented at the 116th Audio Engineering Society Convention, Berlin, Germany, 2004.

[6] Poletti, M.A. "Robust Two-Dimensional Surround Sound Reproduction for Nonuniform Loudspeaker Layouts." Journal of the Audio Engineering Society, 55(7/8), pp.598- 610, 2007.

[7] Moore, D. & Wakefield, J.P. "The Design and Detailed Analysis of First Order Ambisonic Decoders for the ITU layout." Presented at the 122th Audio Engineering Society Convention, Vienna, Austria, 2007.

[8] Wiggins, B., Paterson-Stephens, I., Lowndes, V. & Berry, S. "The Design and Optimisation of Surround Sound Decoders Using Heuristic Methods." Proceedings of UKSim 2003, Conference of the UK Simulation Society p.106-114, 2003.

[9] Gerzon, M.A. "General Metatheory of Auditory Localisation." Presented at the 92nd Audio Engineering Society Convention, Vienna, Austria, 1992.

[10] Moore, D. & Wakefield, J.P. "The Design of Ambisonic Decoders for the ITU 5.1 Layout with Even Performance Characteristics." Presented at the 124th Audio Engineering Society Convention, Amsterdam, The Netherlands, 2008.

[11] Moore, D. & Wakefield, J.P. "Exploiting Human Spatial Resolution in Surround Sound Decoder Design." Presented at the 125th Audio Engineering Society Convention, San Francisco, California, USA, 2008.

[12] 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 Audio Engineering Society Conference, London, UK, 2007.

[13] Moore, D., "The Development of a Design Tool for 5-Speaker Surround Sound Decoders", PhD Thesis, University of Huddersfield, 2009.

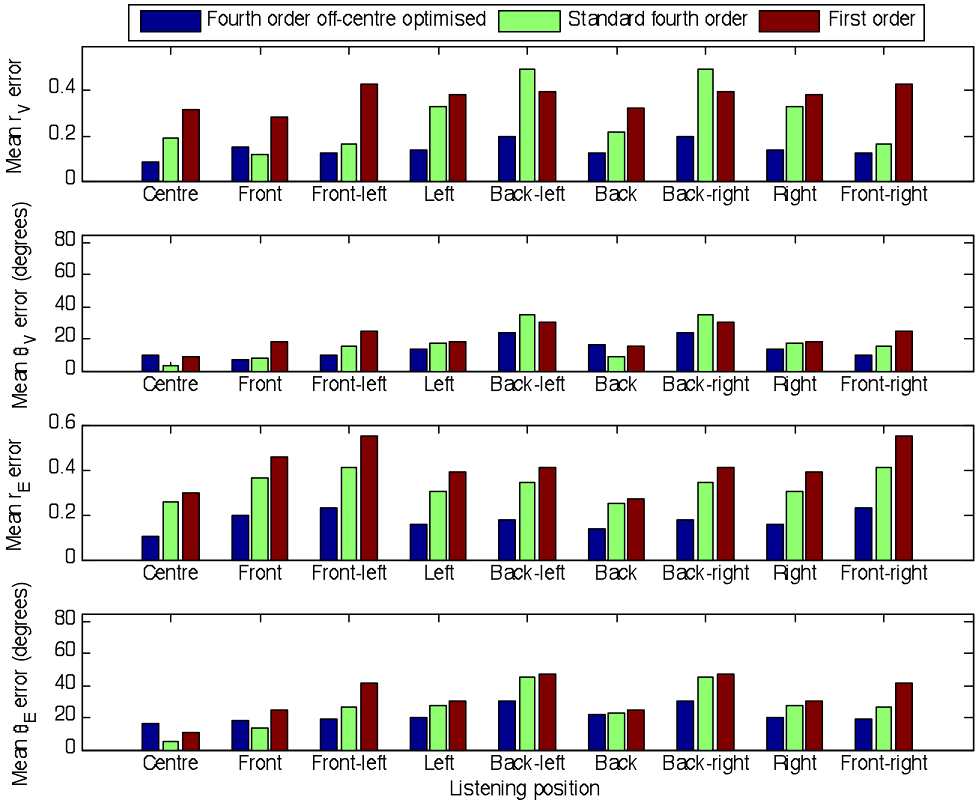

Figure 3 Mean velocity vector and energy vector errors for each decoder at each listening position

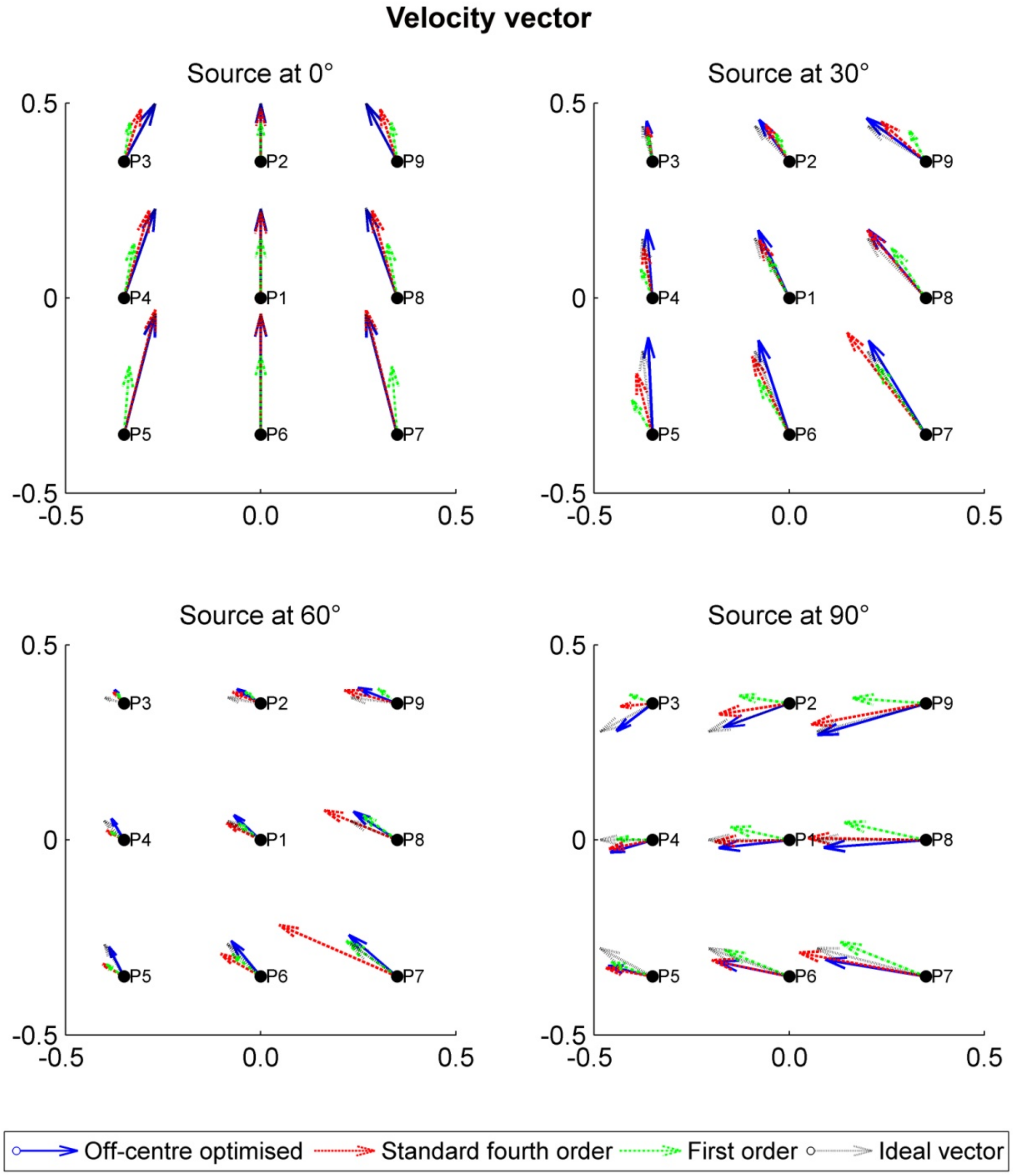

Figure 4 Velocity vector for each decoder and position for the angle of 0°, 30°, 60° and 90°

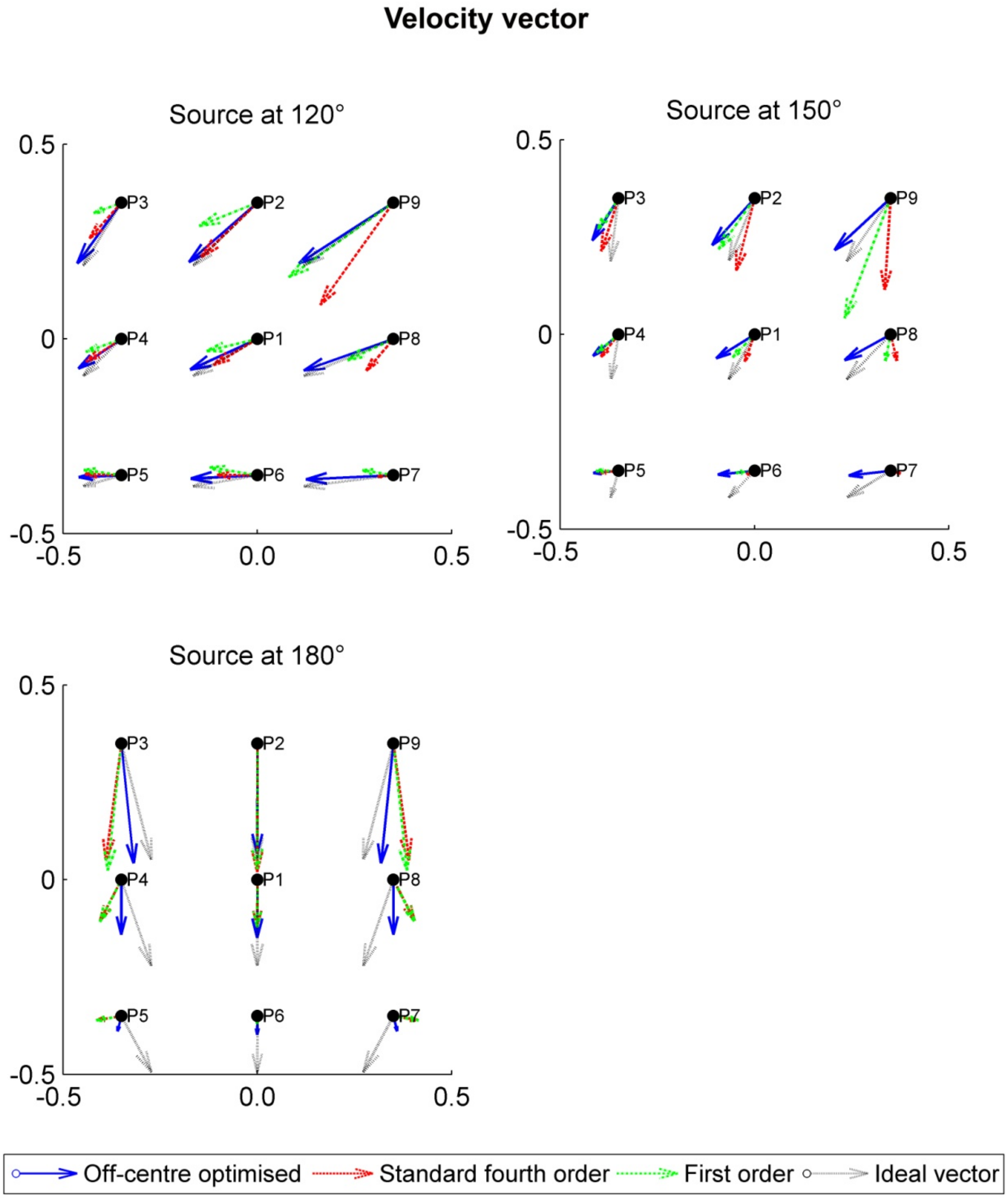

Figure 5 Velocity vector for each decoder and position for the angle of 120°, 150° and 180°

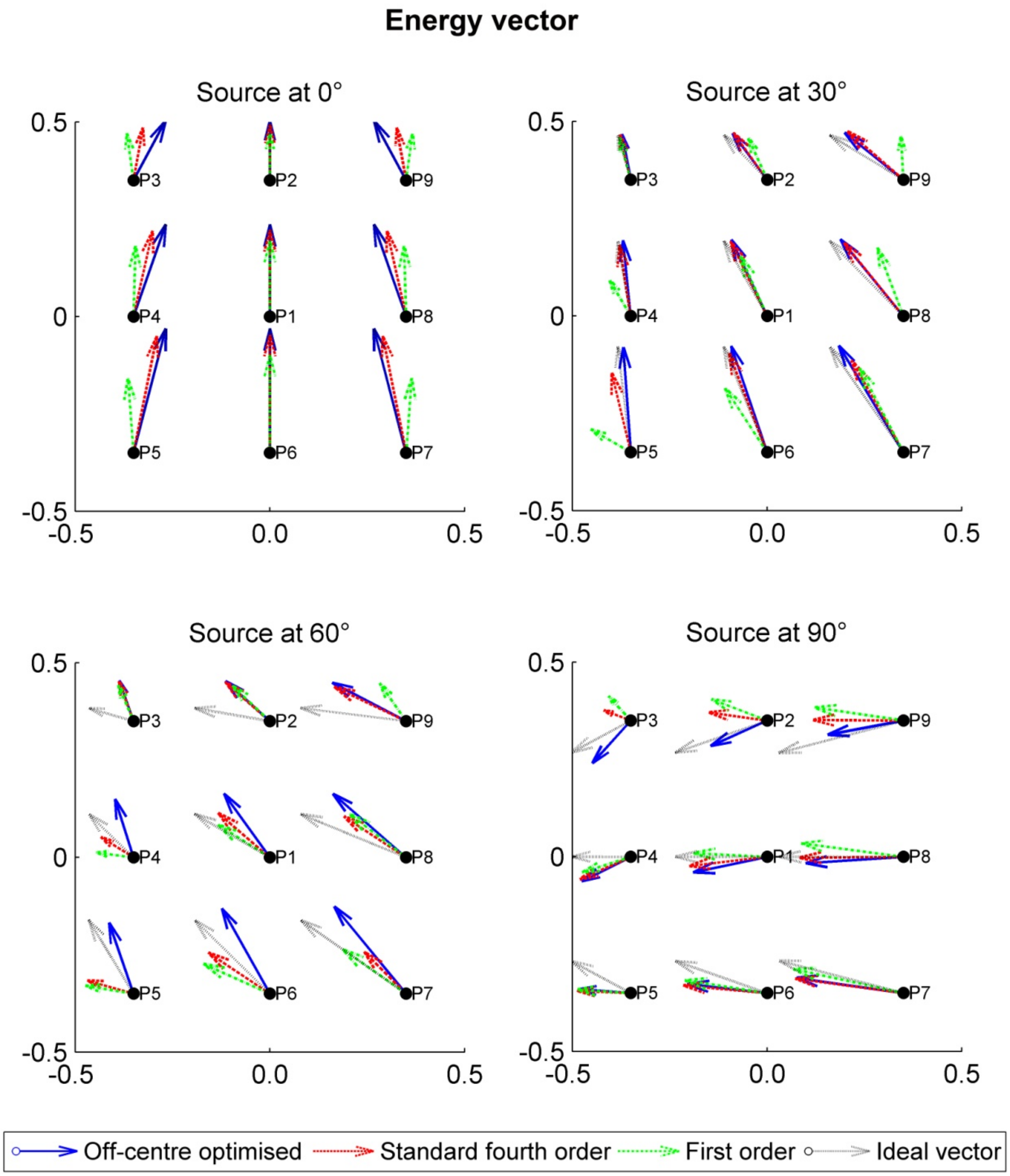

Figure 6 Energy vector for each decoder and position for the angle of 0°, 30°, 60° and 90°

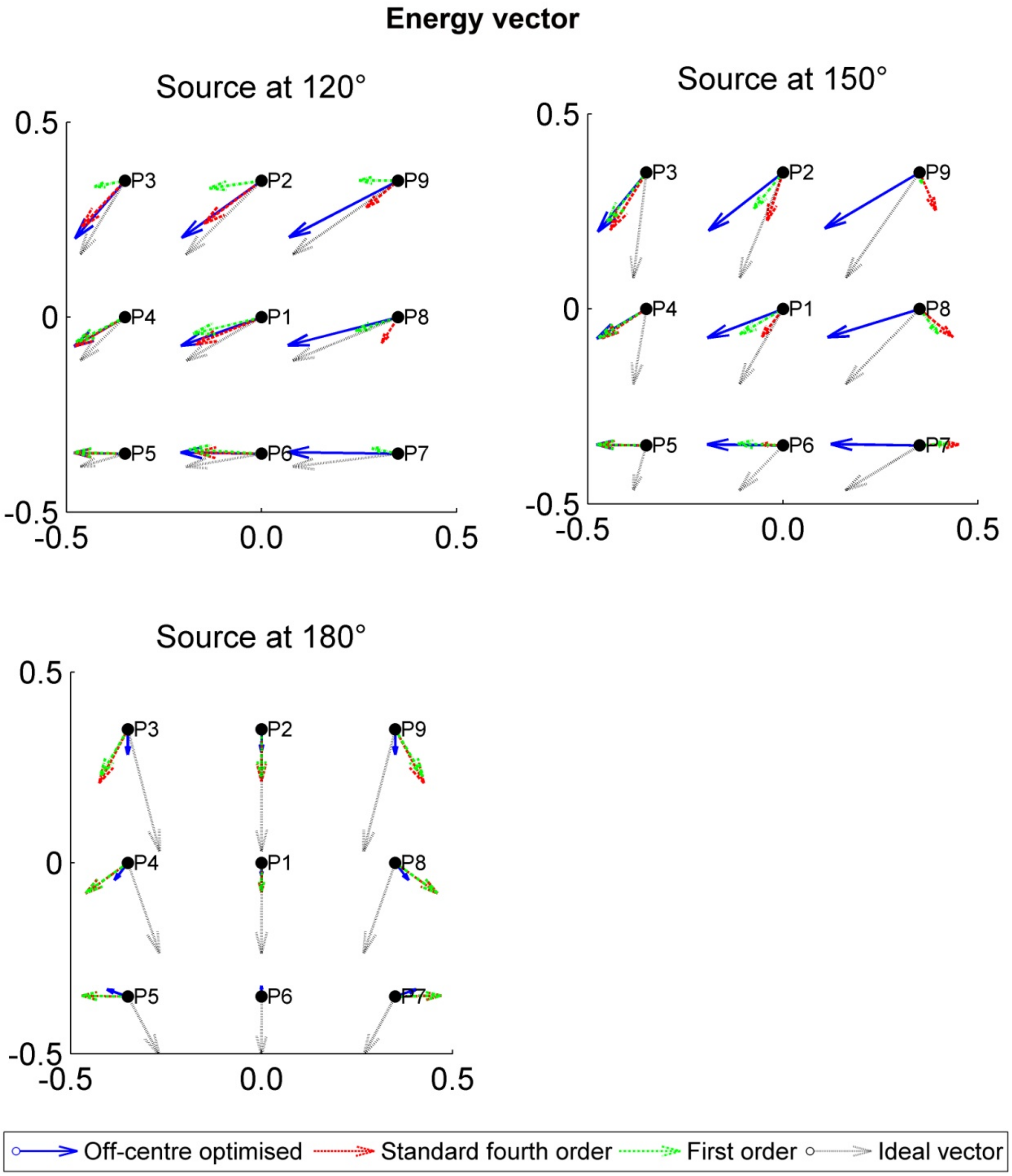

Figure 7 Energy vector for each decoder and position for the angle of 120°, 150° and 180°