AMBISONICS SYMPOSIUM 2009
June 25-27, Graz

A NEW MIXED-ORDER SCHEME FOR AMBISONIC SIGNALS

Chris Travis

Sonopsis Ltd (Chris.Travis@Sonopsis.Ltd.UK)

Abstract: *Traditionally the directional resolution of a 3D Ambisonic signal is uniform over the sphere. It is determined by a single scaling parameter, the periphonic order P. Recently there has been increasing interest in mixed-order schemes that provide higher resolution in the horizontal plane than at the poles. The most widely known is a two-parameter scheme (#H#P) in which the signal is the union of a higher-order horizontal-only component set and a lower-order fully-periphonic component set. We present an alternative two-parameter scheme (#H#V) which truncates the spherical harmonic expansion in a different way. It gives resolution-versus-elevation curves that are flatter in and near the horizontal plane. The paper includes simulation results for various mixed-order signals and speaker layouts. On the basis of these result the author recommends deprecating #H#P signals with P greater than 1.*

1 INTRODUCTION

For most people, surround sound continues to be 2D. Even many cutting-edge projects in Wave Field Synthesis and Higher Order Ambisonics run shy of the difficulties of rendering soundscapes in all three spatial dimensions. Yet even fairly simple 3D setups can give good results.

The author has been impressed by several 3D audio demonstrations, and has come to believe that capturing and rendering at-least some height information is very worthwhile. The prospect is of a more-natural rendition, with spatial unmasking letting listeners hear more detail. On the other hand, for the author at least, some 3D systems have a distinct localization blur problem.

With conventional Ambisonics there is another problem. In horizontal-only systems the number of speakers scales linearly with the signal order, but in fully-periphonic systems it scales with the square of the order. Hence there is in-practice a rather wide gulf between the 2D and conventional 3D playback options.

The author was introduced to mixed-order possibilities by Dave Malham in 1994. The appeal of mixed-order schemes and strategies is that they will let us balance the good and bad points of horizontal-only systems against the good and bad points of fully-periphonic systems. They should also let us find better matches between Ambisonics technologies and the recent with-height proposals from NHK and others.

To the author's best knowledge, the #H#V mixed-order scheme considered in this paper was first proposed by Jerome Daniel in 2001 [1]. It was also independently arrived at by Fons Adriaensen and Chris Travis in 2008.

2 HORIZONTAL-ONLY FIRST-ORDER

To set the scene it is convenient to explore the performance of horizontal-only first-order Ambisonics. The corresponding B-format signal has just three components. When the conventional scalings apply, these can be referred to as W, X and Y. Figure 1 shows simulation results for such a signal. The emphasis here is on what happens outside of the horizontal plane, i.e. on how elevated sources are rendered.

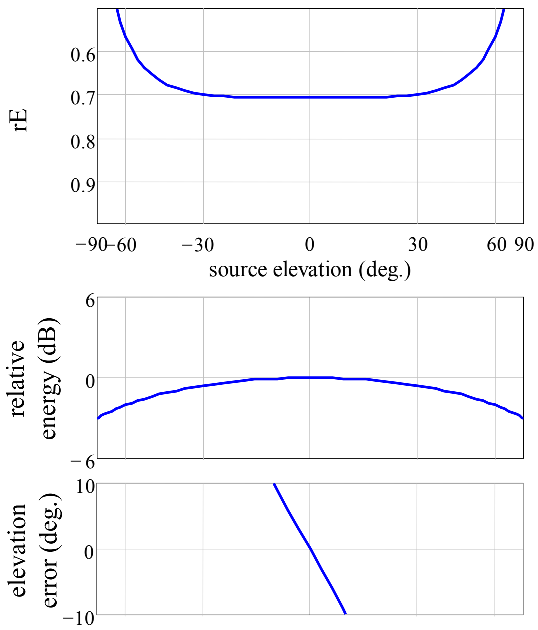

Figure 1: Typical performance with 1H systems

The simulation is of "max rE" decoding to a horizontal speaker rig, ignoring near-field aspects. The rig has four or more speakers with uniform azimuthal distribution.

The first plot is of the energy vector magnitude rE. A good non-mathematical description of this measure is given in [2]. Note that rE=1 has been put at the bottom of the graph, so the height of the trace above the X axis is representative of the energy localization blur 1−rE. Lower height means less blur. Note also that the source elevation scale has been warped with a sine law. This is analogous to equal-area mapping of the globe in geography. Exactly one-half of all possible source directions are contained in the range -30 to +30 degrees.

An ideal system would have an energy response that is uniform over all source directions. Significant deviations from uniformity seem to disrupt the listener's sense of the recording venue. This might be for reasons relating to the Craven hypothesis on distance perception. Such considerations justify the second plot in figure 1.

The main spatial limitation of horizontal-only systems is of-course that they don't render height. This is reflected in the third plot of figure 1, which shows the difference between the elevation of the energy vector (zero degrees in this case) and that of the source. Elevated sources alias into the horizontal plane. But note that this aliasing generally does not sound 'bad' or 'wrong'. So it should perhaps be thought of as the removal and absence of spatial information, rather than as the substitution of incorrect spatial information. On the plus side, figure 1 shows that low-order horizontal-only systems tend to have reasonable rE and energy balance over quite a broad range of source elevations. The rE value is greater than 0.70 over most of the sphere. This contrasts with full 3D playback of a first-order signal on a regular polyhedral rig, which delivers an rE of only 0.58 [1].

3 THE #H#V MIXED-ORDER SCHEME

The #H#V mixed-order scheme is illustrated in figure 2. The figure shows all the components, and two cut lines. The chosen positions of the cut lines determine which components are included in the signal. The case shown has parameter H equal to 4 and parameter V equal to 1. The resulting signal can be referred to as a 4H1V signal.

The organization of the components in this figure is unusual, but is well suited to Ambisonics. It can easily be arrived at from the more-common harmonic triangle, in which index m runs from negative values on the left through to positive values on the right. You simply fold the right-hand side of that triangle over onto the left.

The component naming scheme embodied in figure 2 is experimental. In it 'D' relates to degree, 'E' relates to elevational wavenumber, and 'c' and 's' relate to cosine and sine ($m > 0$ and $m < 0$).

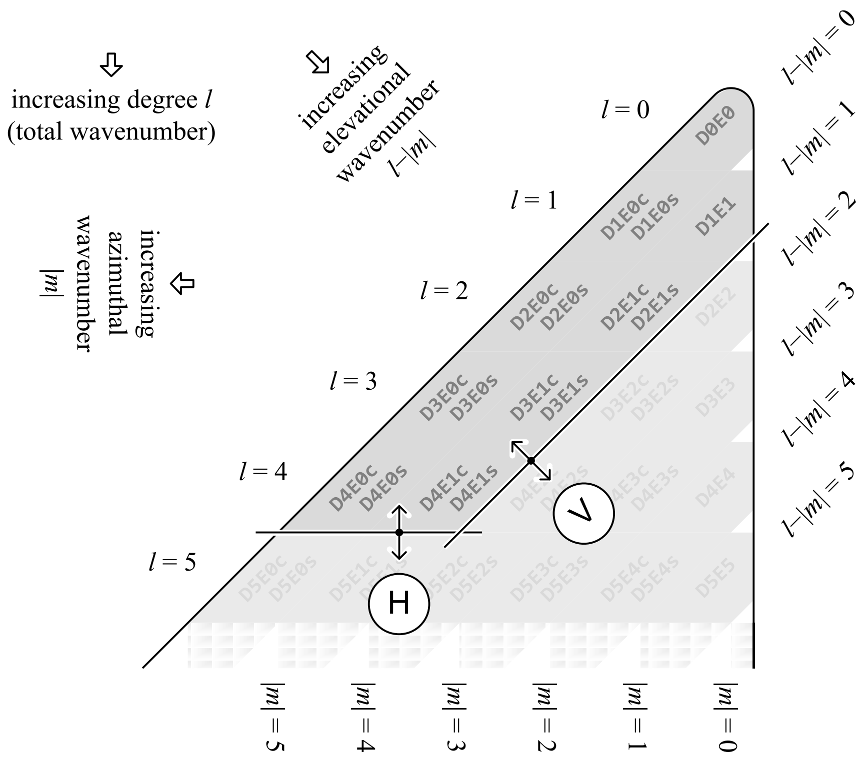

Figure 2: The Ambisonic component triangle, showing truncation to a 4H1V signal

A fully-periphonic signal of order P consists of all the components of degrees 0 to P. This is shown on the right-hand side of figure 3. Truncating the spherical harmonic expansion in this way gives signals that have uniform resolution over the sphere. There is only one parameter, and only one cut-line.

A horizontal-only signal of order H consists of the components of degrees 0 to H that have an elevational wavenumber of zero. This is shown on the left-hand side of figure 3. There is only one parameter, but there are two cut lines.

The #H#V mixed-order scheme bridges the space between horizontal-only signals and fully periphonic signals. As is evident from figure 3, horizontal-only signals can be thought of as degenerate #H#V signals in which V = 0. Raising V to 1, 2, 3 etcetera introduces progressively more components. That is until V = H, at which point we have a fully periphonic signal.

The prior scheme #H#P also bridges the space between horizontal-only signals and fully periphonic signals. This is shown in figure 4. An #H#P signal is the union of a higher-order horizontal-only component set and a lower-order fully-periphonic component set. There are two parameters, but three cut lines. #H#P signals can be thought of as degenerate cases in a three-parameter scheme, #H#V#P, with V = 0.

To make the various constructions more concrete figure 5 lists all the resulting signals with up to 16 components.

The signals are organized into families. In the #H#V scheme the first step above horizontal-only Ambisonics is provided by the #H1V family. Its signals 2H1V, 3H1V and 4H1V have 8, 12 and 16 components respectively. At higher horizontal orders the emphasis can sensibly shift to the #H2V and #H3V families. But at moderate horizontal orders it is the #H1V signals that are of particular interest.

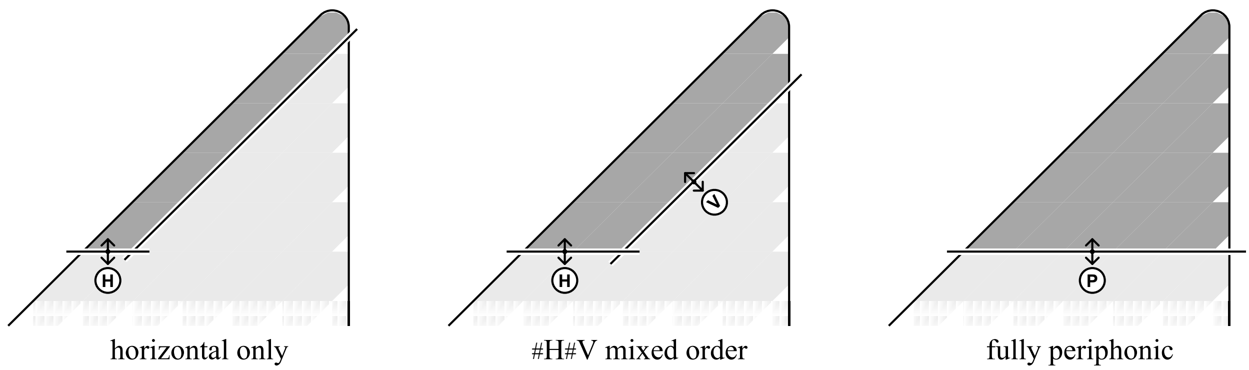

Figure 3: Progression according to the new mixed-order scheme

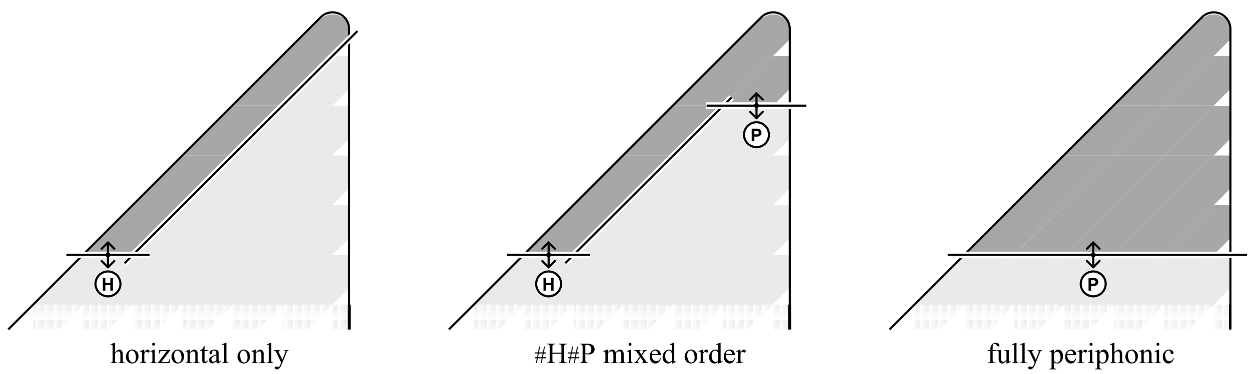

Figure 4: Progression according to the prior mixed-order scheme

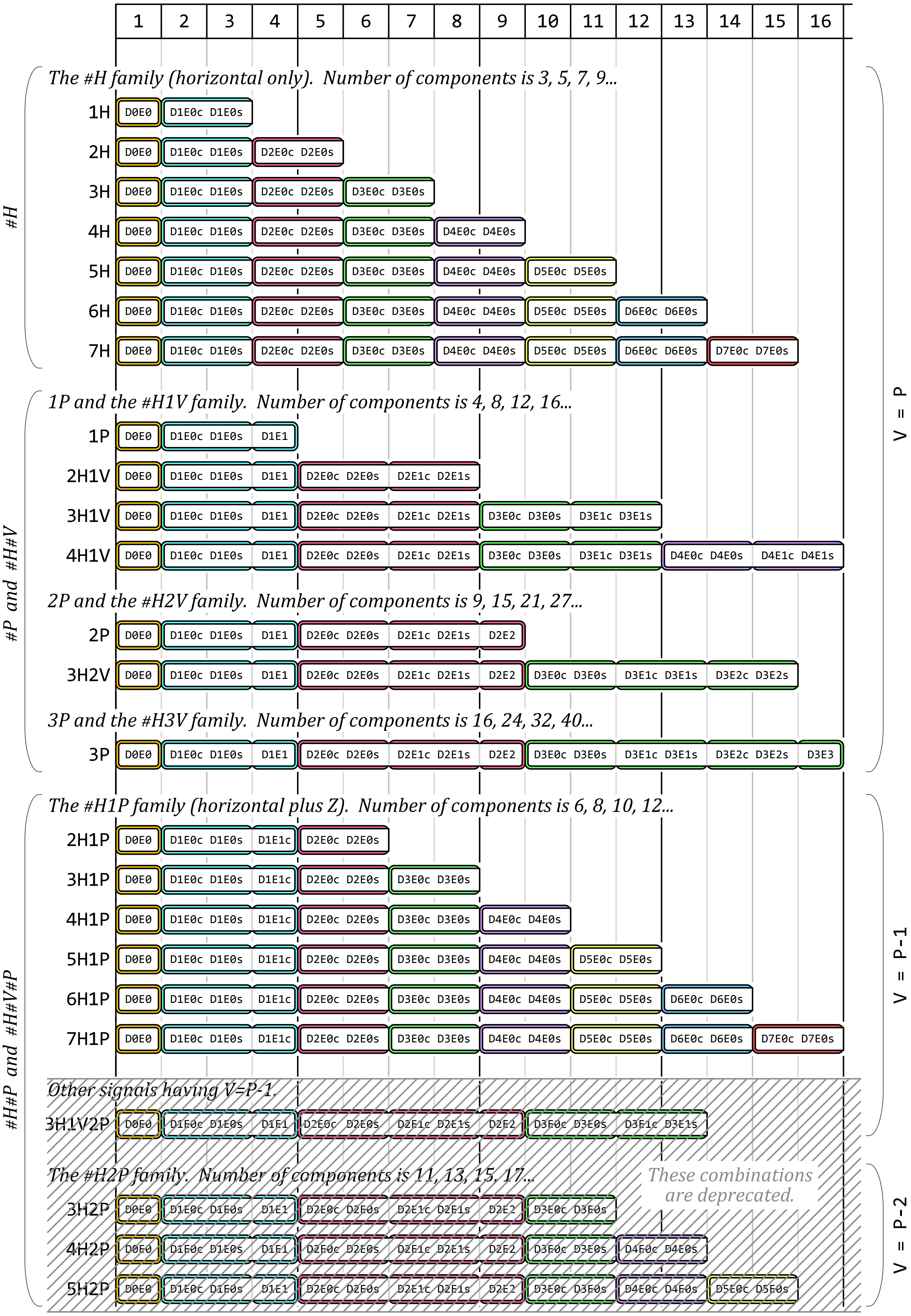

Figure 5: Mappings for all combinations with up to 16 components

4 SIMULATIONS

4.1. Comparing 2H1V and 3H1P signals

To get a feeling for the merits and demerits of the #H#V and #H#P mixed-order schemes we ran some simulations. The first compares a 2H1V signal with a 3H1P signal. These both have eight components. The 2H1V signal can be thought of as a second-order fully-periphonic signal with one component omitted. (In Furse-Malham notation the omitted component is 'R'.) The 3H1P signal can be thought of as a third-order horizontal signal with the Z component added.

Playback was over two rings of speakers at elevations of +30 and -30 degrees. With such an arrangement, the 3H1P signal needs 16 speakers while the 2H1V signal can manage on 12. Both decoders operated in a "max rE" manner for horizontal sources, with the additional freedom(s) being used to get flatness near zero elevation.

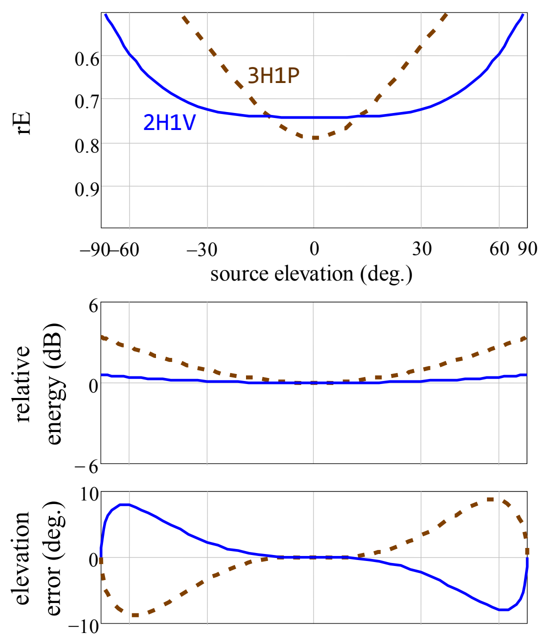

Figure 6: Comparing two 8-component signals

Looking first at rE: The 3H1P signal does very well with sources in the horizontal plane, but its performance degrades quite rapidly with elevation. This might not be a problem for soundscapes that have been constructed with this limitation in mind. But for real recordings, the rE curve of the 2H1V system might be more appealing.

Now looking at the second graph: In the 3H1P system, sounds from near the poles will play back louder than one would like. Such an emphasis is unfortunate. The 2H1V system gives significantly better energy balance.

4.2. Comparing 4H1V, 7H1P and 3P signals

To explore behaviour at higher orders, we next compared three sixteen-component signals. These are easy to find in figure 5. The 4H1V and 7H1P tests had 20 and 32 speakers respectively, with elevations of +22.5 and -22.5 degrees. For the 3P signal we used 20 speakers in an icosahedral arrangement. The 4H1V decoder was set up for minimum elevation error near the horizontal plane, but was not otherwise optimized.

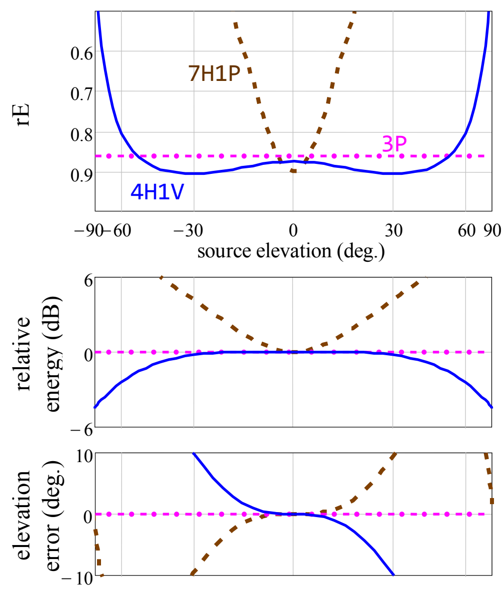

Figure 7: Comparing three 16-component signals

From the rE and relative energy plots it is clear that 7H1P is inappropriate for any system in which the sources can be panned away from the horizontal plane. Its Z channel is really only for ambience and reverb.

Furthermore, the contents of the Z channel play back much louder than one would like. This could be fixed by attenuating Z, but that would worsen the elevation errors.

It seems likely that material tailored to 7H1P signals would work better on a horizontal-only speaker rig!

The 4H1V system does much better, giving high rE over a wide range of elevations. It even compares well with the traditional 3P system, delivering comfortably higher rE over most of the sphere.

A weak point of the simulated 4H1V setup is its increasing elevation error, as sources move beyond the speaker elevations. On the other hand, as mentioned in section 2, we seem to be much more tolerant of elevation errors than of azimuth errors. Also, these errors are perhaps as-much to do with the chosen speaker arrangement as with the 4H1V signal.

4.3. Comparing traditional and dual-ring playback

In this section we take a 3P signal as our starting point, and focus on two different rendering strategies. The first uses 20 speakers in an icosahedral arrangement. It has great elegance but low practicality. The second strategy involves throwing away four of the sixteen components to get a 3H1V signal, and then rendering that signal via a dual-ring arrangement of 16 speakers. The speakers are at elevations of +/-22.5 degrees.

We are discarding components and using fewer speakers. Can we preserve reasonable performance for source elevations between +22.5 and -22.5 degrees?

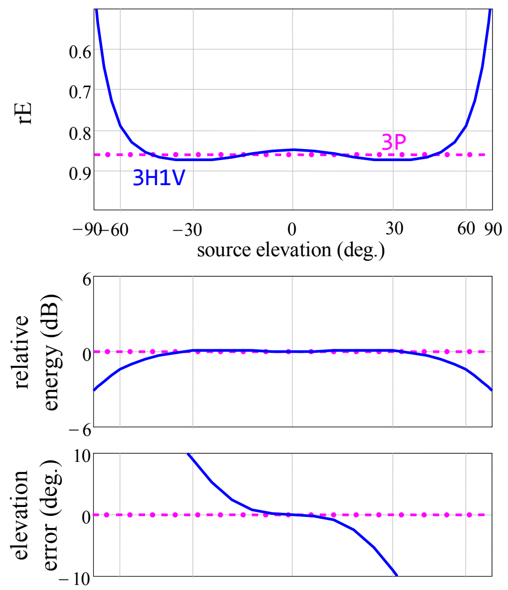

Figure 8: Comparing two playback strategies

Figure 8 show that the cut-down strategy is indeed capable of good rE and reasonable energy balance. The question-mark is again over its elevational errors.

Figures 7 and 8 both show a tendency to better rE values for source directions near speakers. This has also been seen with other speaker layouts. It bodes well for the performance of three-ring rigs.

5 RECOMMENDATIONS

The author believes that the results presented in this paper cast considerable doubt on the utility of the #H#P mixed-order scheme. At the same time, they confirm the utility of the #H#V scheme. One could consider adopting a three-parameter scheme #H#V#P, as a superset of #H#V and #H#P. The author instead recommends deprecating the #H#P scheme for values of P greater than one. For signals with up to 16 components, the results of such a restriction have been indicated in figure 5 by hashing. More generally, this restriction would leave us with the four categories of signals that are shown in table 1.

Signal category	Relevant parameters	Number of components
Fully periphonic	P	$(P+1)^2$
Mixed order	H and V	$(H+1)^2 - (H-V)^2$
Horizontal plus Z	H	2H + 2
Horizontal only	H	2H + 1

Table 1: Recommended Ambisonic signal categories

Standards such as the Ambisonics portion of MPEG-4 Part 11 include the option of constructing Ambisonic signals as arbitrary sets of components. In that light, Table 1 might seem to be something of a straightjacket. It is intended however as part of a practical framework for organizing, exchanging and interpreting Ambisonic files and streams. As such, further simplifications might be appropriate.

6 CONCLUSION

Mixed-order schemes give Ambisonics a way forward that is free of the limitations of horizontal-only systems yet relatively unencumbered by the costs and difficulties of conventional fully-periphonic systems. This paper has made an initial exploration of the mixed-order scheme that was proposed by Jerome Daniel in his thesis [1]. The scheme is found to have a good fit with the maths, and to have a natural synergy with some of the multi-ring speaker layouts explored by Eric Benjamin in [3].

REFERENCES

[1] Jérôme Daniel, "Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia", doctoral thesis, University of Paris, July 2001.

[2] Peter G. Craven, "Continuous Surround Panning for 5-Speaker Reproduction", AES 24th International Conference, June 2003.

[3] Eric Benjamin, "Ambisonic Loudspeaker Arrays", AES Convention Paper 7605, October 2008.