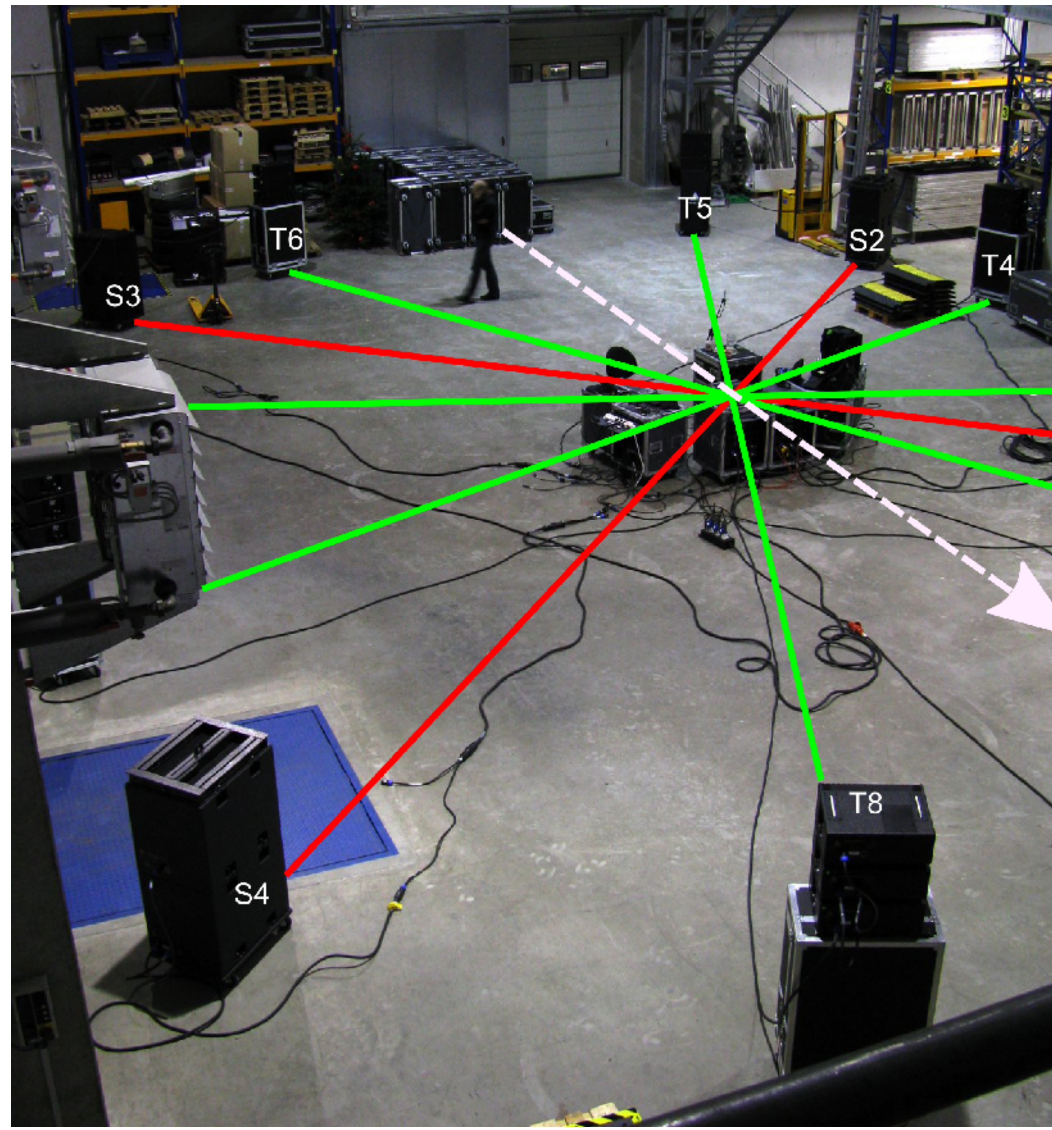
Large-scale Higher Order Ambisonic Sound Reinforcement Systems

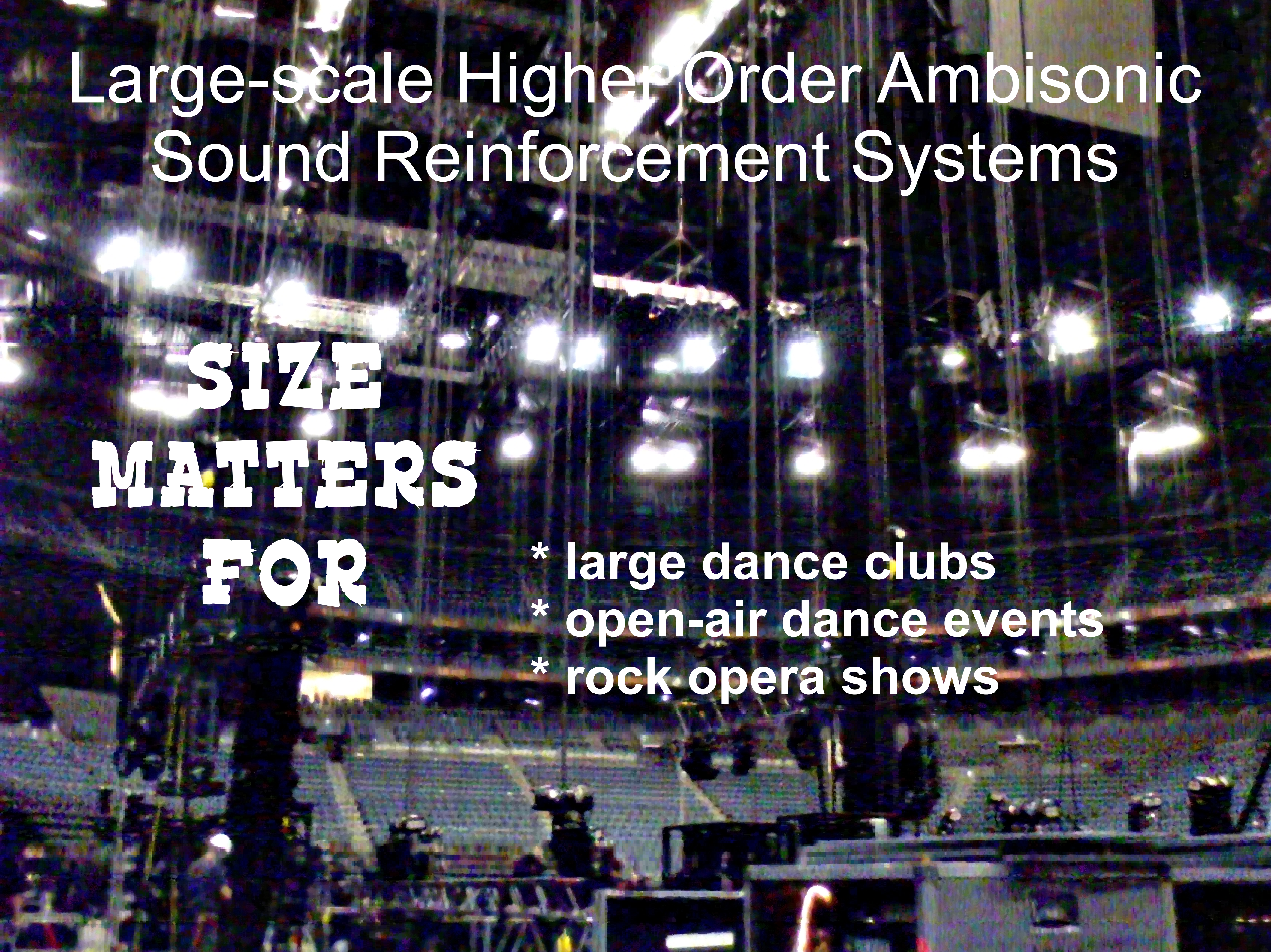
or:

SIZE DOES MATTER

Jörn Nettingsmeier
<nettings@stackingdwarves.net>

David Dohrmann
<david@adamsonsystems.com>




Large-scale Higher Order Ambisonic
Sound Reinforcement Systems
SIZE
MATTERS
FOR
* large dance clubs
* open-air dance events
* rock opera shows

Introduction:

- Modern line arrays can be designed to deliver almost constant SPL over a wide area. Therefore, the only limiting constraint for large systems is run-time.
- Can we scale HOA systems to hundreds or thousands of listeners?
- What happens very far outside the sweet spot?

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Large-scale Higher Order Ambisonic Sound Reinforcement Systems

Overview:

- Previous large-scale Ambi implementations
- Test system configuration
- Listening impressions
- Theoretical analysis
- Conclusion

Previous implementations:

Malham et al. 1990, “Mediamix 90”, York, UK:
8 horizontal speakers
4 elevated speakers
first-order decoder
max. radius: ca. 5m

Malham et al. 1992, “York Mystery Plays”:
quasi-spherical: 1-4-6-4-1
first-order in-phase decode
audience outside the sphere!
max. radius: ca. 7m

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Previous implementations:

Wiggins/Funktion One 2006-2010, Glastonbury Festival, UK:

six horizontal horn arrays
second-order decode
max. radius: ~20m

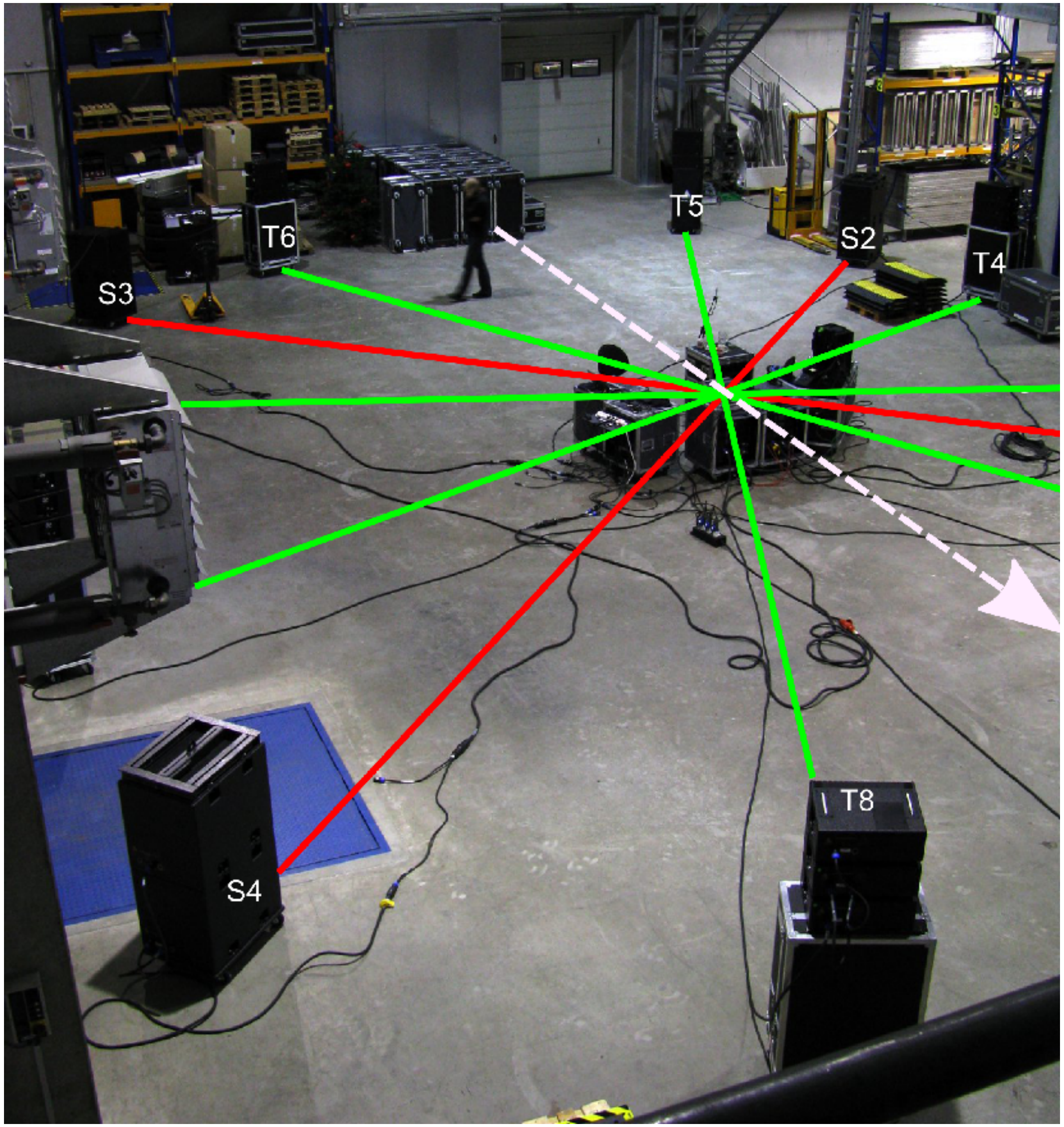
Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Configuration of the test system:

8 mid-hi stacks:
3x Adamson Metrix
8.5” LF, 3”->1.4” HF

4 sub stacks:
2x Metrix Sub
2x15”

X-over 100Hz,
24dB/Oct Linkwitz-Riley



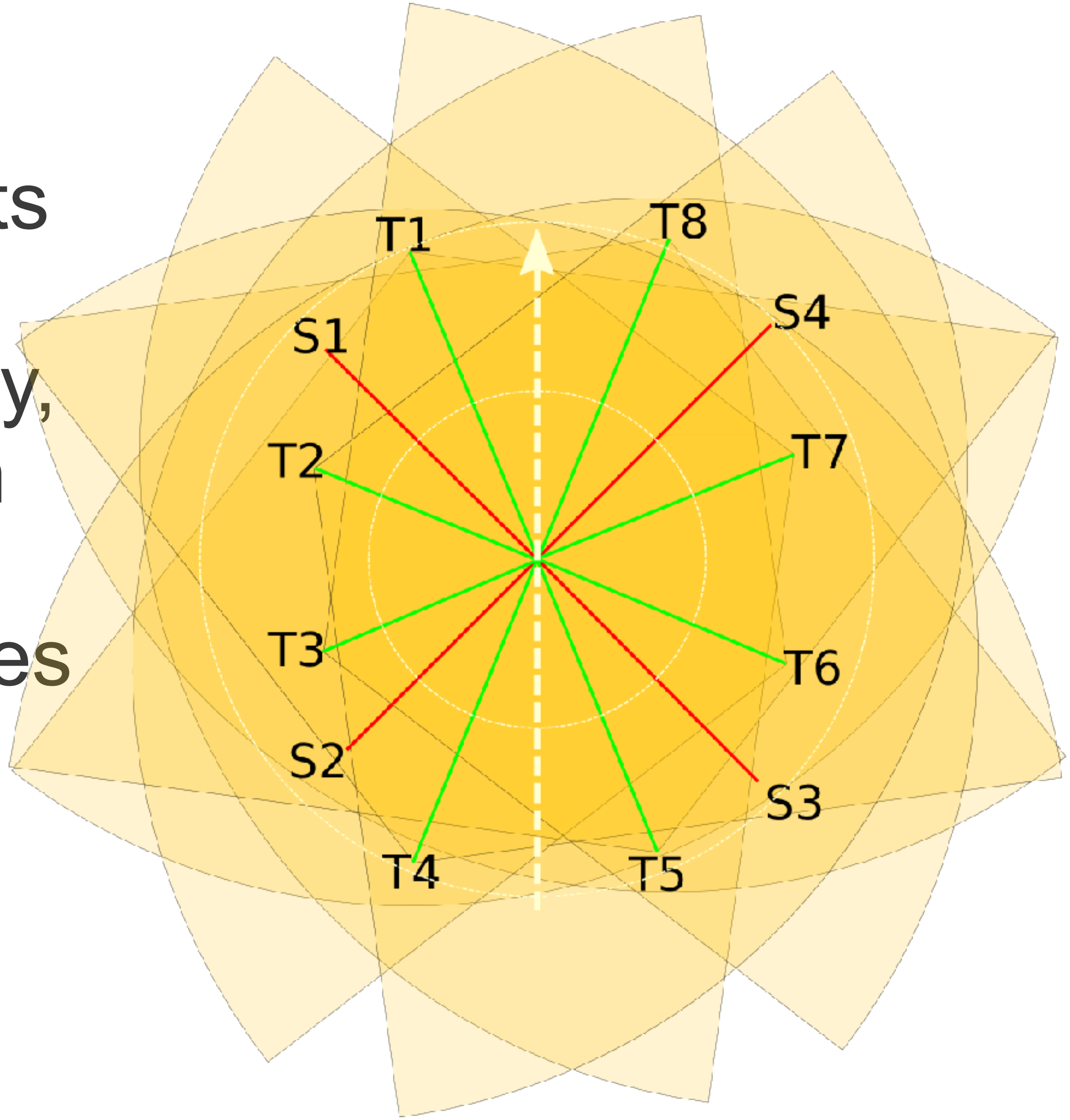
Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Configuration of the test system:

slightly oblong layout
due to space constraints

compensated with delay,
effective radius 10.25m

nominal coverage angles
mid-hi: 120°
subs: omni



Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

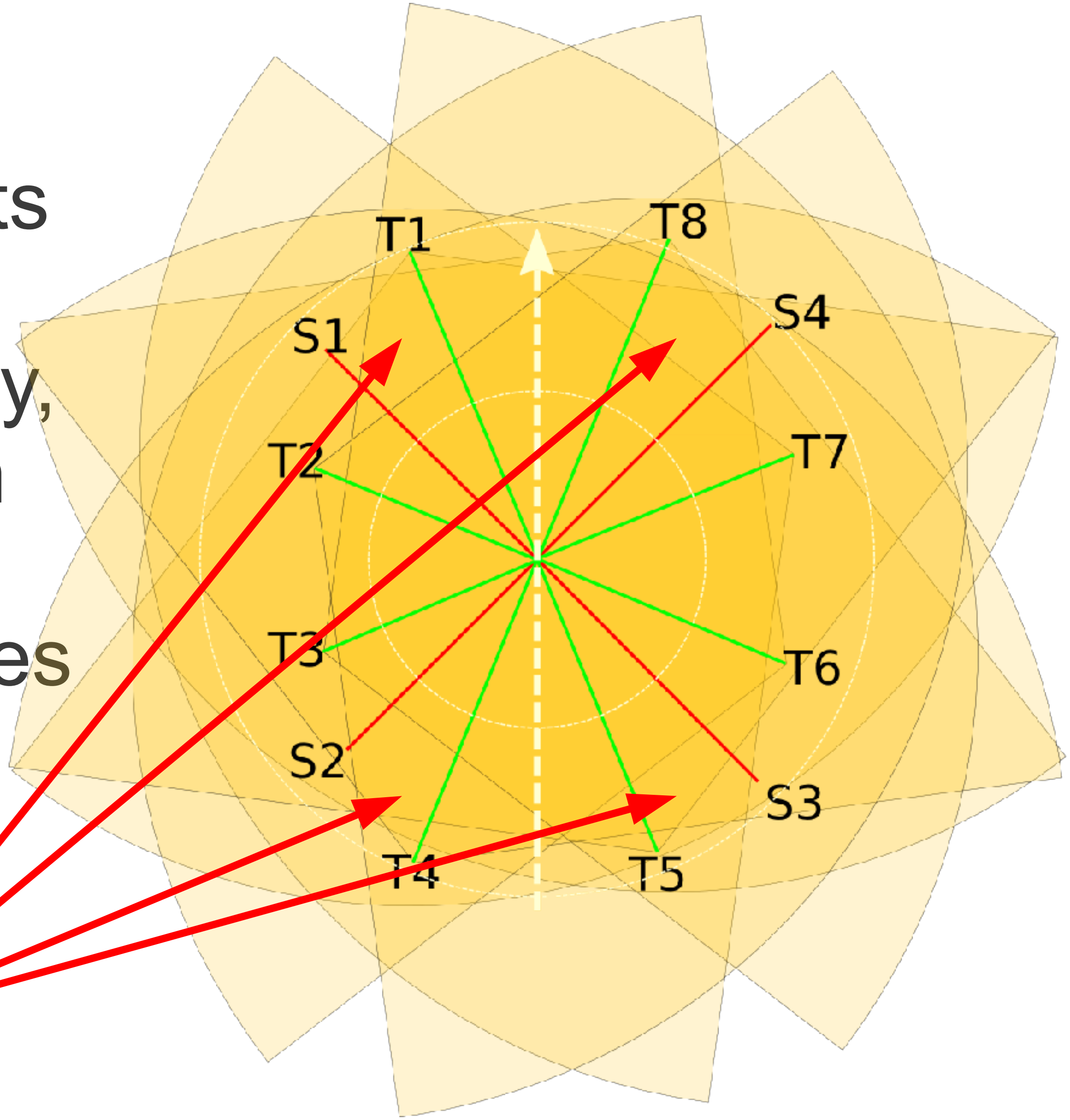
Configuration of the test system:

slightly oblong layout
due to space constraints

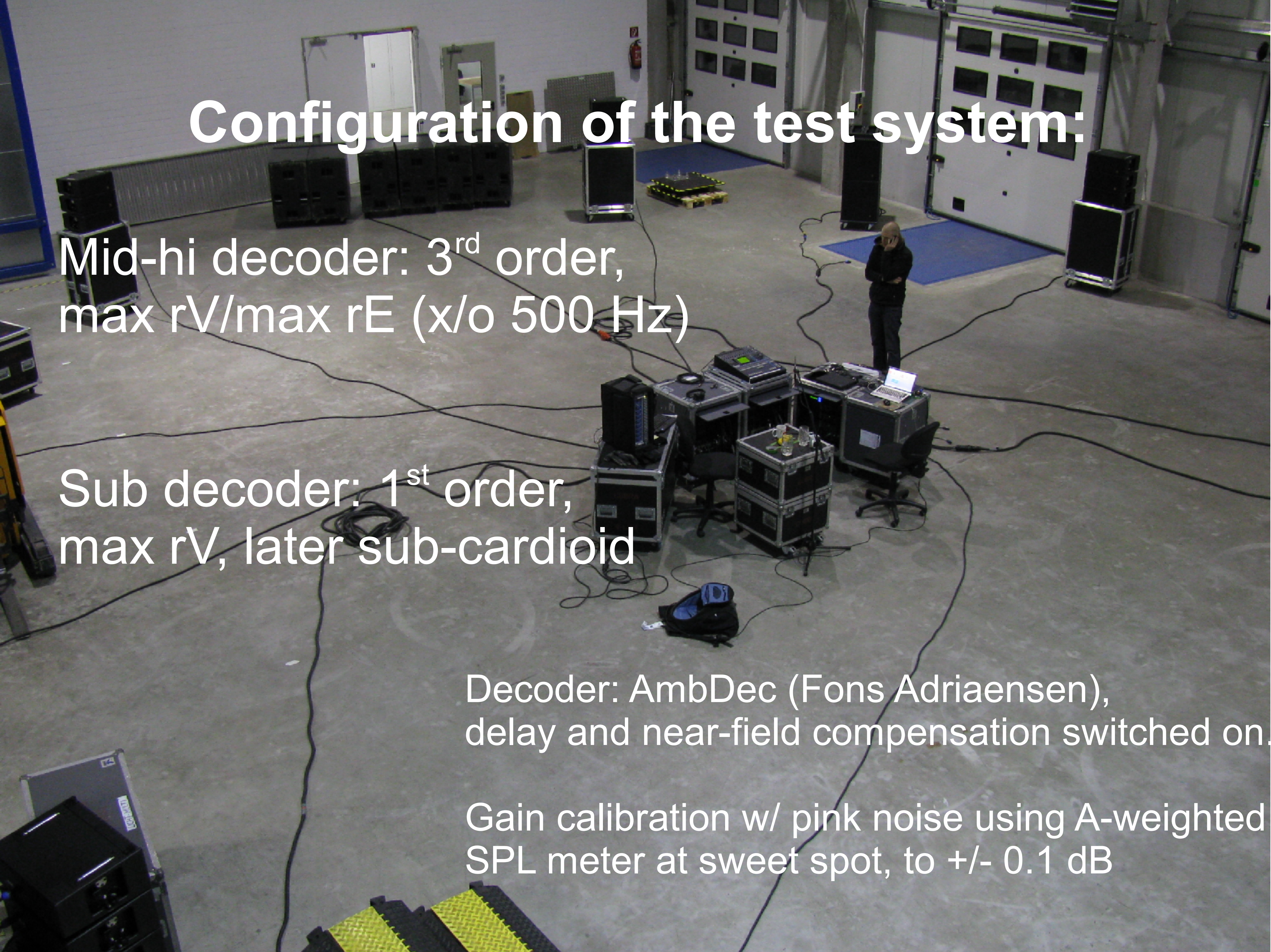
compensated with delay,
effective radius 10.25m

nominal coverage angles
mid-hi: 120°
subs: omni

coverage issues!



Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>



Configuration of the test system:

Mid-hi decoder: 3rd order,
max rV/max rE (x/o 500 Hz)

Sub decoder: 1st order,
max rV, later sub-cardioid

Decoder: AmbDec (Fons Adriaensen),
delay and near-field compensation switched on.

Gain calibration w/ pink noise using A-weighted
SPL meter at sweet spot, to +/- 0.1 dB

Listening impressions:

NB: Two subjects only (the authors)!

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Listening impressions:

NB: Two subjects only (the authors)!

Yeah, that's unscientific. But try to get people to attend a listening session two days before Christmas, with three days' notice, in a cozy warehouse heated to 12°, from 20-24:00h.

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Listening impressions:

NB: Two subjects only (the authors)!

Yeah, that's unscientific. But try to get people to attend a listening session two days before Christmas, with three days' notice, in a cozy warehouse heated to 12°, from 20-24:00h.

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Listening impressions:

First-order material:
ok'ish for background music, not focused enough to draw attention to stage or performers

Third-order material:
stable and focused over a wide area; tested with single sources and hybrid HOA recordings

very large perceived source size - irritating with nature recordings, but kind of nice for rock shows
← not specific to Ambisonics

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

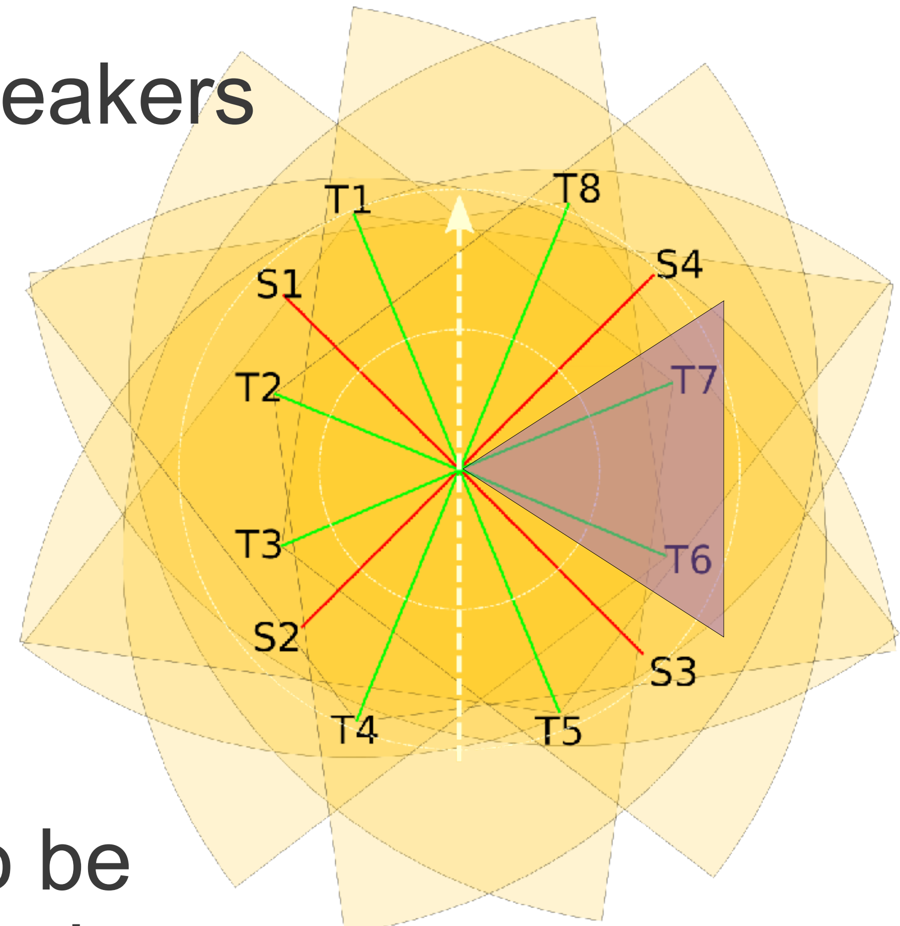
Listening impressions:

A/B comparison:
native Stereo vs. virtual speakers

Base angle 45°

Stacks T7 and T6
with S4 and S3 used
for symmetry reasons.

Virtual speakers panned to be
congruent with physical stacks.

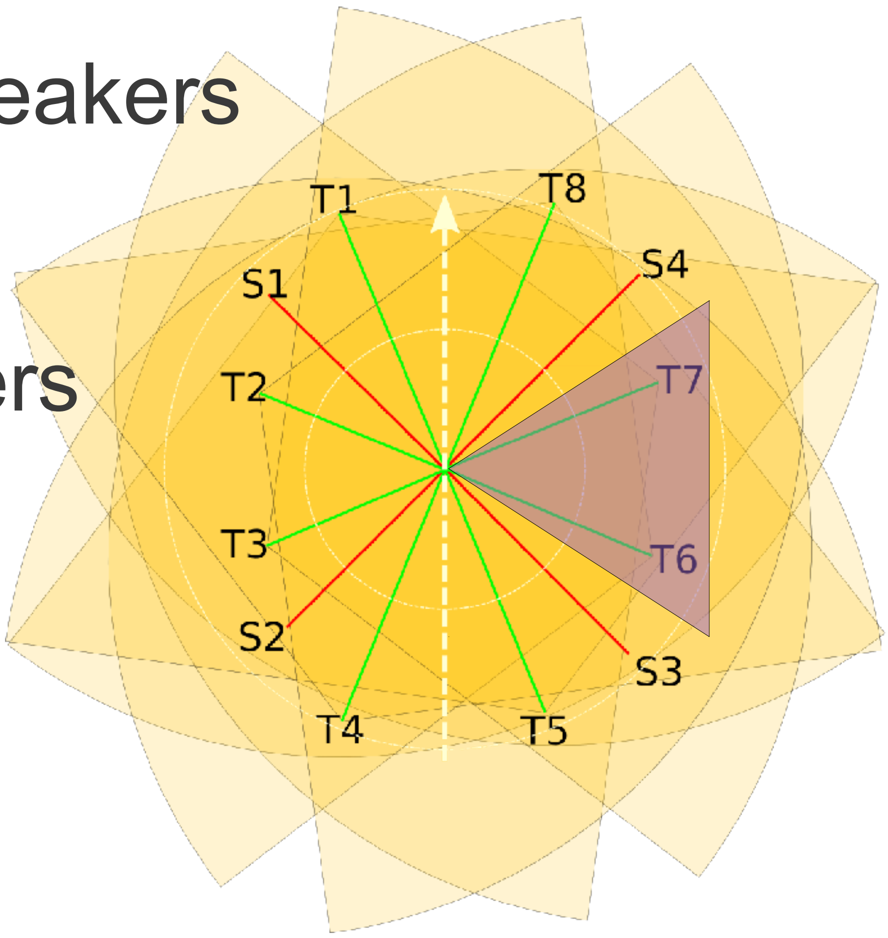


Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Listening impressions:

A/B comparison:
native Stereo vs. virtual speakers

slight loss of presence and
dryness with virtual speakers
in direct comparison;
still very enjoyable

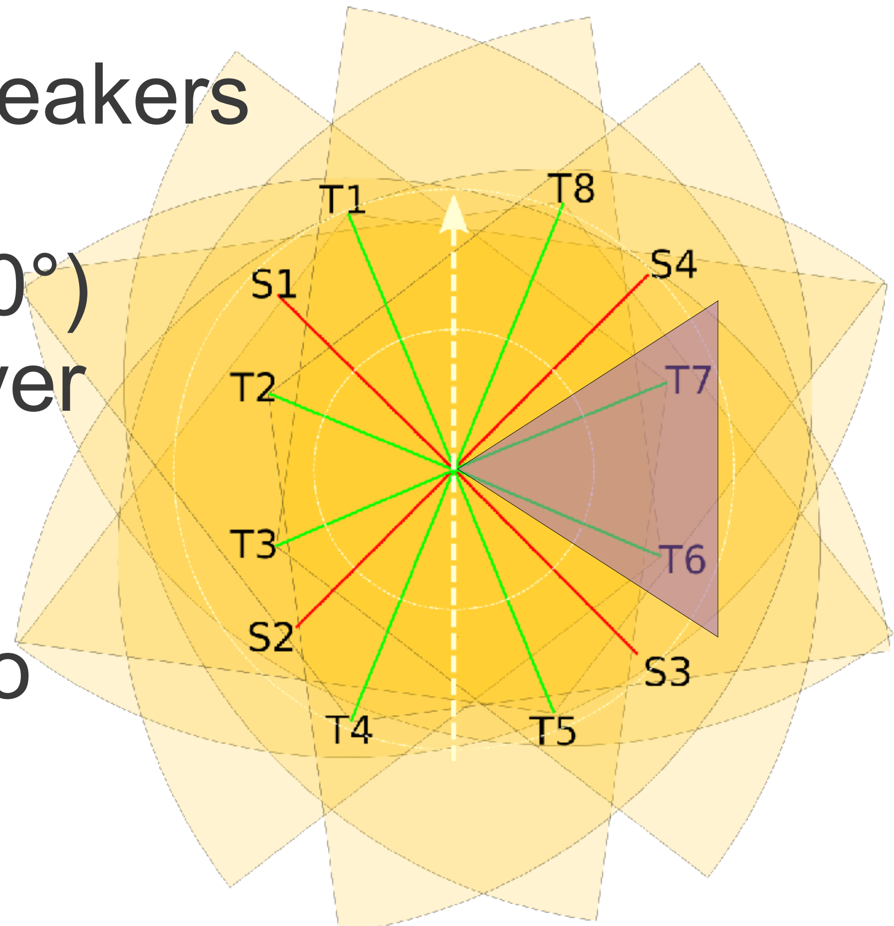


Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Listening impressions:

A/B comparison:
native Stereo vs. virtual speakers

wider virtual stereo (90-120°) prefered by both testers over native stereo;
very immersive, no “hole-in-the-middle” effect, stereo localization remains stable



→ **an improvement even for standard stereo material**

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Listening impressions:

Sub decode:

Max rV means out-of-phase components:
→ loss of “punch”, less efficient
→ no perceptible gain in localization

Sub-cardioid or even mono decode is better suited to rock'n'roll settings, unless
→ the crossover frequency is very high, or
→ the system is installed in a small room (less excitation of room modes with hypercard decode).

Listening impressions:

* very good localization over a wide area
* no evident coloration
* only very subtle phasing artifacts
* no collapse until within 1-1.5m of the opposite speaker even with hypercardioid decode

possibly masked by room reverberation!

* confirmed usable area: 180m²
→ 360 tickets for dance floor, easily more if you provide lounge area around the listening zone.

We are confident that 600 people will have a satisfactory listening experience with adequate localisation.

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Listening impressions:

* very good localization over a wide area
* no evident coloration
* only very subtle phasing artifacts

possibly masked by room reverberation!

* no collap [illegible] speaker
  even with hypercardioid decode

Slight problem:

* confirmed usable area: 180m²
  → 360 tickets for dance floor, easily more if you
  provi [illegible]

This is too good to be true.

We are confident that 500 people will have a satisfactory listening experience with adequate localisation.

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Theoretical considerations:

Almost everybody is outside the sweet spot. Therefore, the **ITD** cues will be incorrect.

As the listener moves away from the center, the ITDs will shift out of phase.

It seems those inconsistent cues are being discarded by the brain and do not cause noticeable confusion, as long as robust **ILD** cues are available.

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Theoretical considerations:

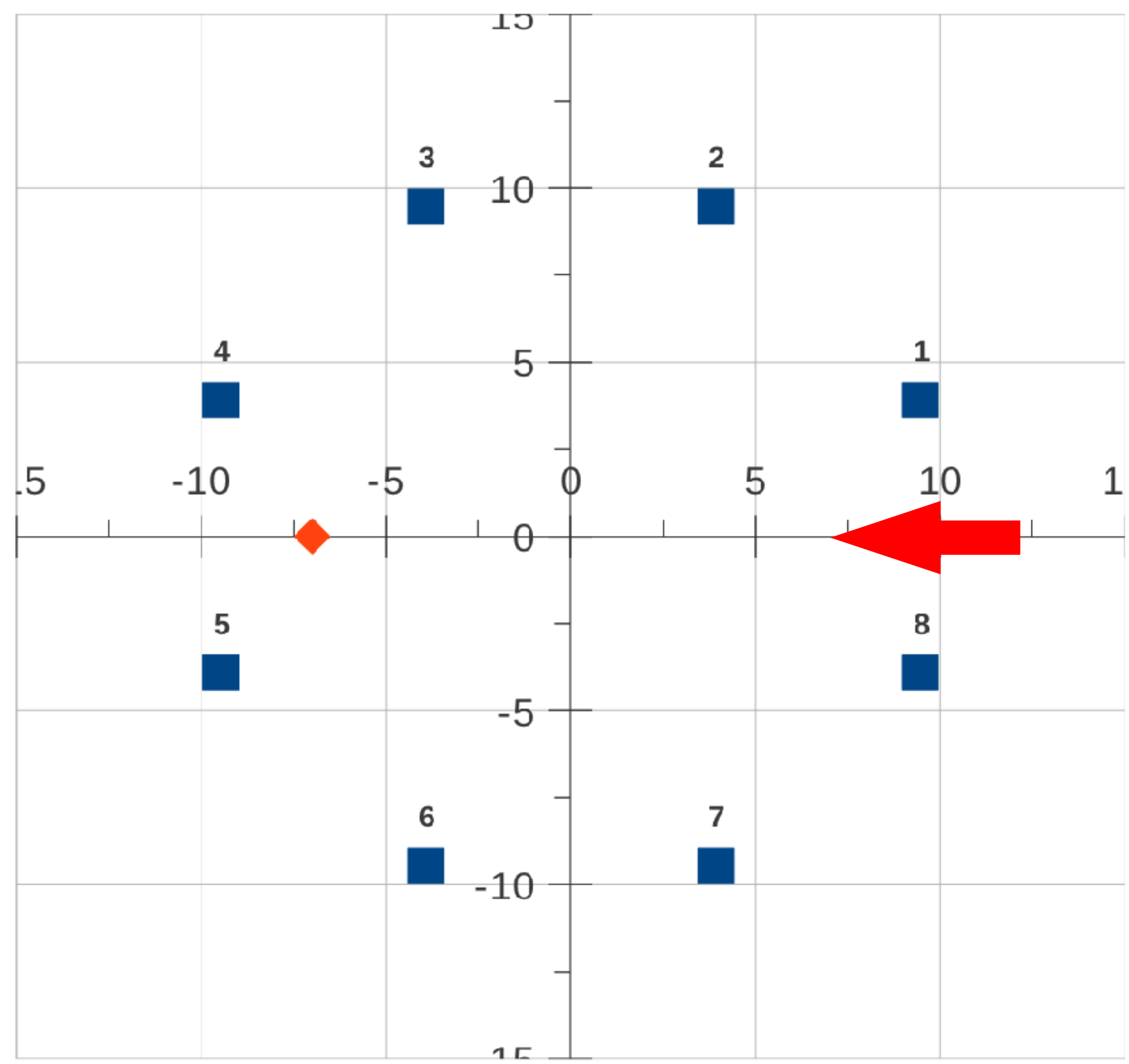
Our best metric is therefore rE.

Consequently, it might be best to use a max rE decode throughout the spectrum, perhaps excluding the subs (not tested).

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Theoretical considerations:

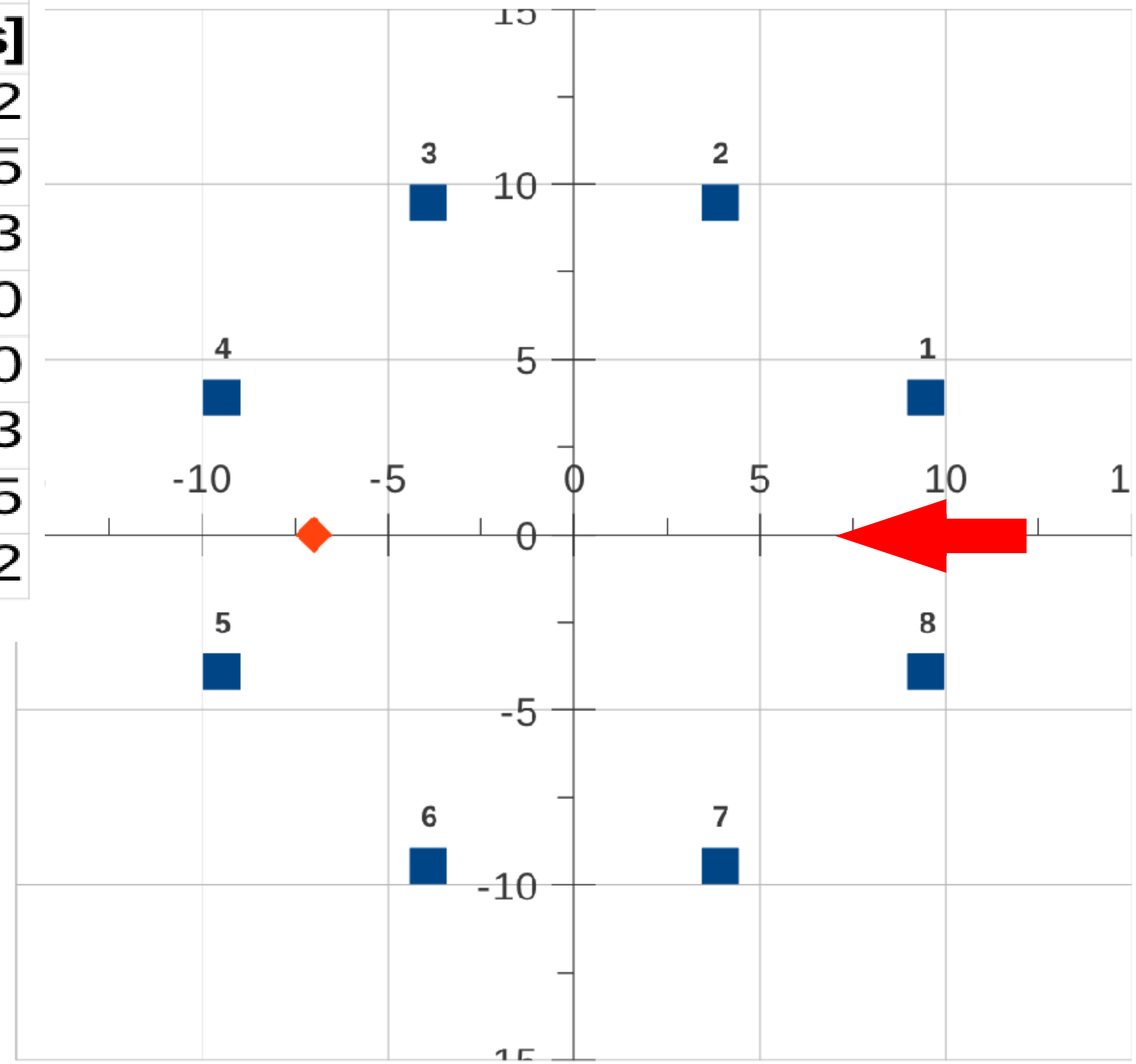
Worst-case scenario:
a source from 0°,
listener at 180°,
7m away from center



Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Theoretical considerations:

	relative level	relative delay
i	Δp(i) [dB]	Δt(i) [ms]
1	0.000	36.162
2	-10.000	28.885
3	-15.000	15.653
4	-18.000	0.000
5	-18.000	0.000
6	-15.000	15.653
7	-10.000	28.885
8	0.000	36.162

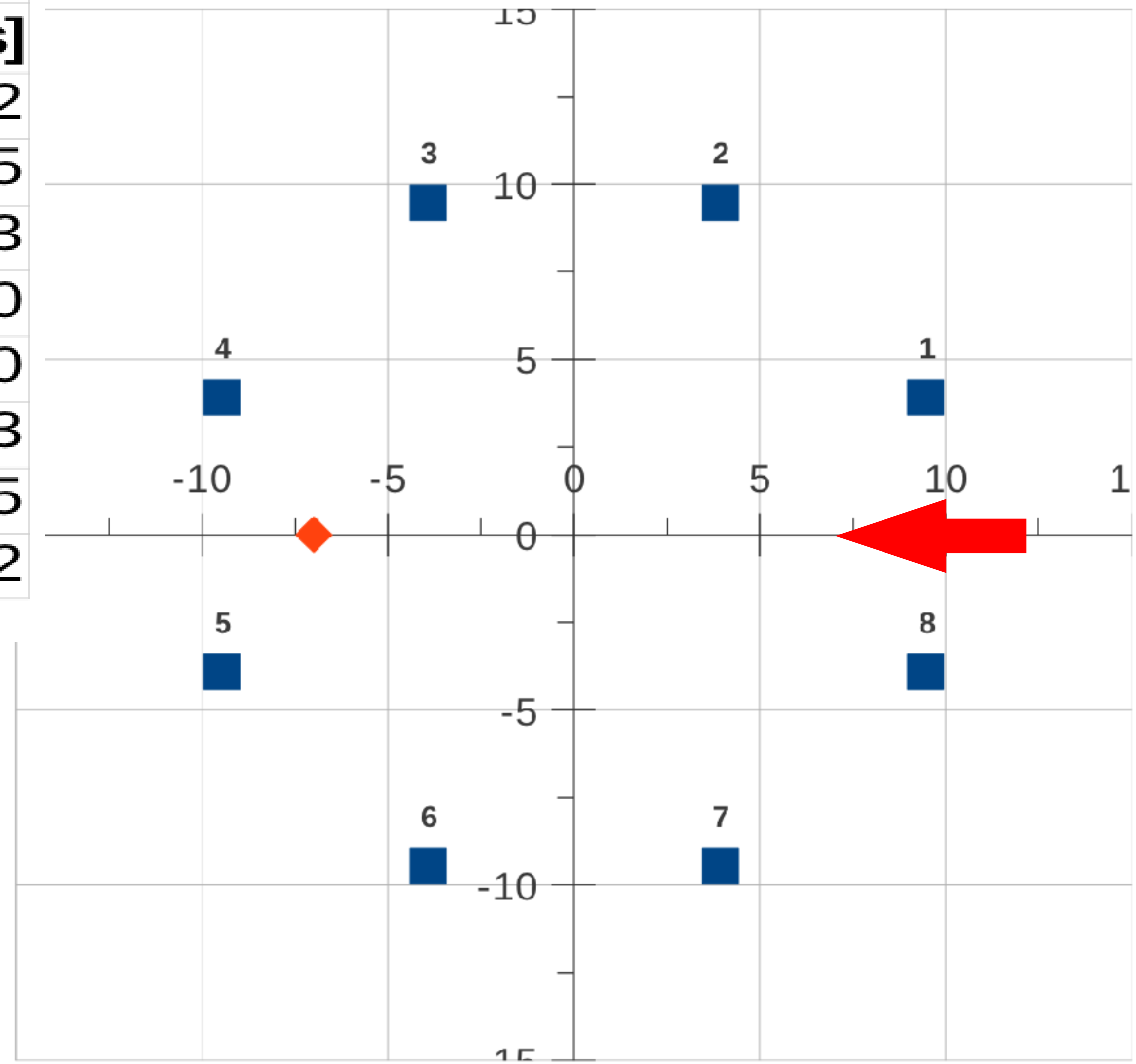


relative levels and timings at listening position (assuming constant SPL, normalized to first incoming wave front)

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Theoretical considerations:

	relative level	relative delay
i	Δp(i) [dB]	Δt(i) [ms]
1	0.000	36.162
2	-10.000	28.885
3	-15.000	15.653
4	-18.000	0.000
5	-18.000	0.000
6	-15.000	15.653
7	-10.000	28.885
8	0.000	36.162

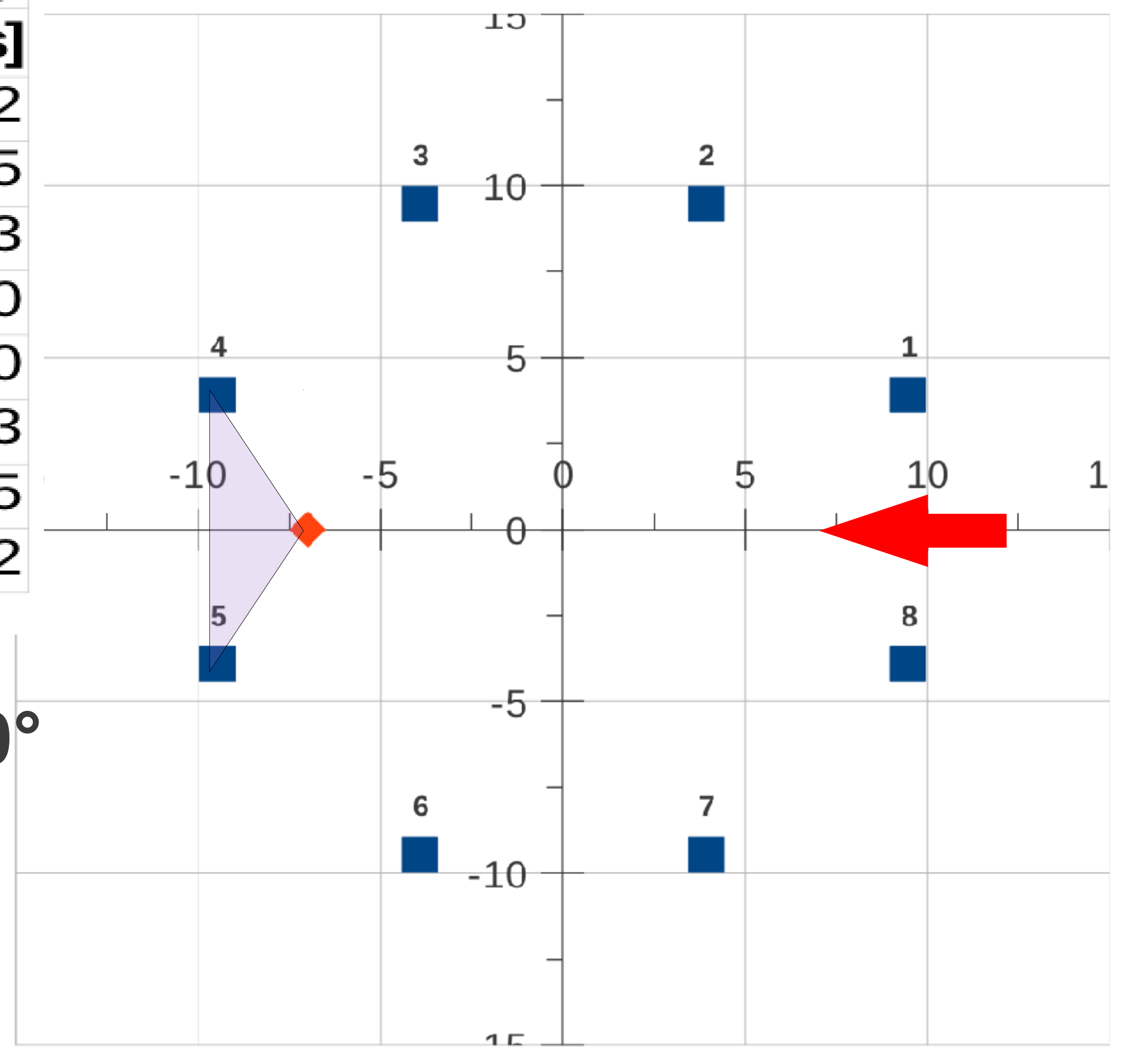


Applying stereophonic theory (summing localization, precedence effect, echo threshold):

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Theoretical considerations:

	relative level	relative delay
i	Δp(i) [dB]	Δt(i) [ms]
1	0.000	36.162
2	-10.000	28.885
3	-15.000	15.653
4	-18.000	0.000
5	-18.000	0.000
6	-15.000	15.653
7	-10.000	28.885
8	0.000	36.162

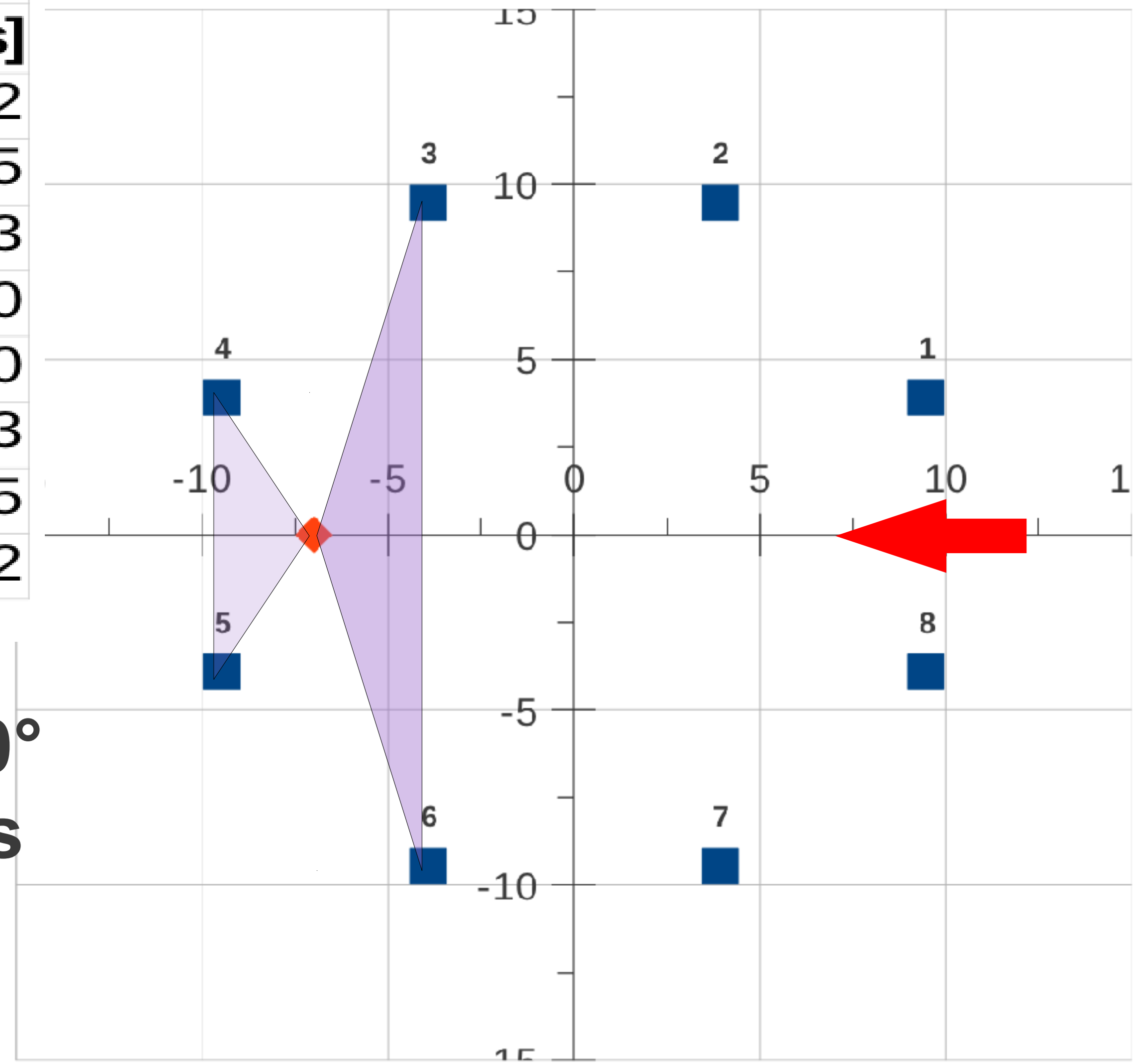


1st phantom source at 180°

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Theoretical considerations:

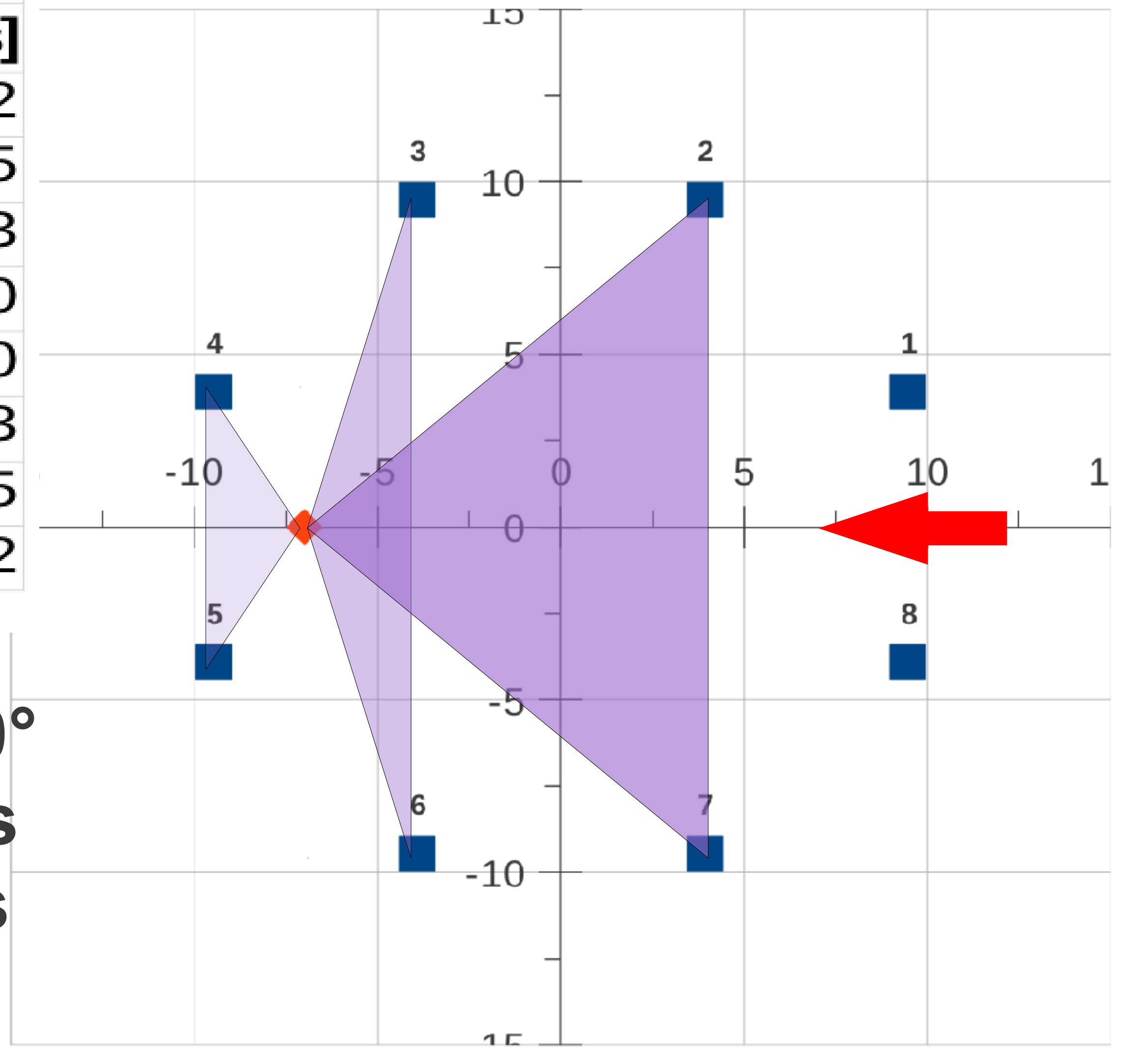
	relative level	relative delay
i	Δp(i) [dB]	Δt(i) [ms]
1	0.000	36.162
2	-10.000	28.885
3	-15.000	15.653
4	-18.000	0.000
5	-18.000	0.000
6	-15.000	15.653
7	-10.000	28.885
8	0.000	36.162



1st phantom source at 180°
2nd p.s. at 0°, +3dB, +16ms

Theoretical considerations:

	relative level	relative delay
i	Δp(i) [dB]	Δt(i) [ms]
1	0.000	36.162
2	-10.000	28.885
3	-15.000	15.653
4	-18.000	0.000
5	-18.000	0.000
6	-15.000	15.653
7	-10.000	28.885
8	0.000	36.162

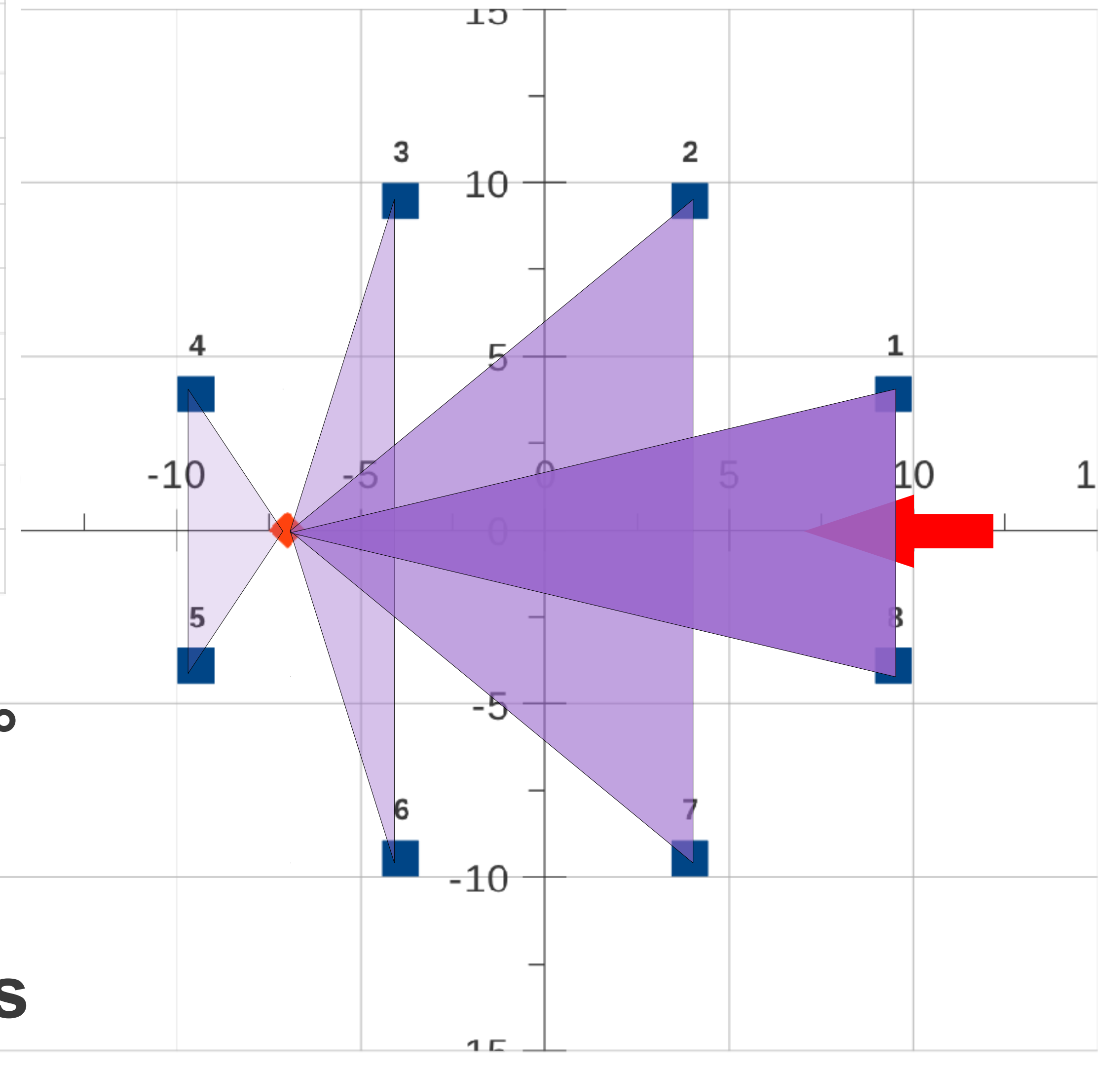


1st phantom source at 180°
2nd p.s. at 0°, +3dB, +16ms
3rd p.s. at 0°, +8dB, +29ms

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Theoretical considerations:

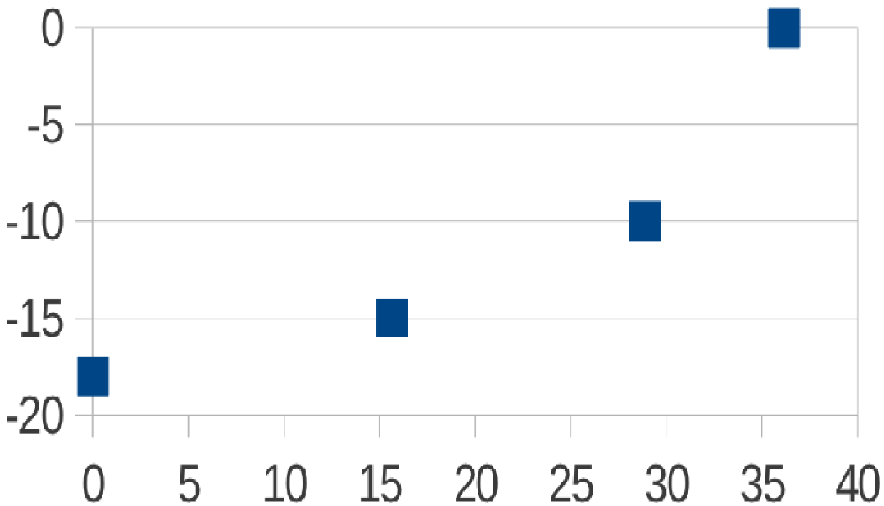
	relative level	relative delay
i	Δp(i) [dB]	Δt(i) [ms]
1	0.000	36.162
2	-10.000	28.885
3	-15.000	15.653
4	-18.000	0.000
5	-18.000	0.000
6	-15.000	15.653
7	-10.000	28.885
8	0.000	36.162

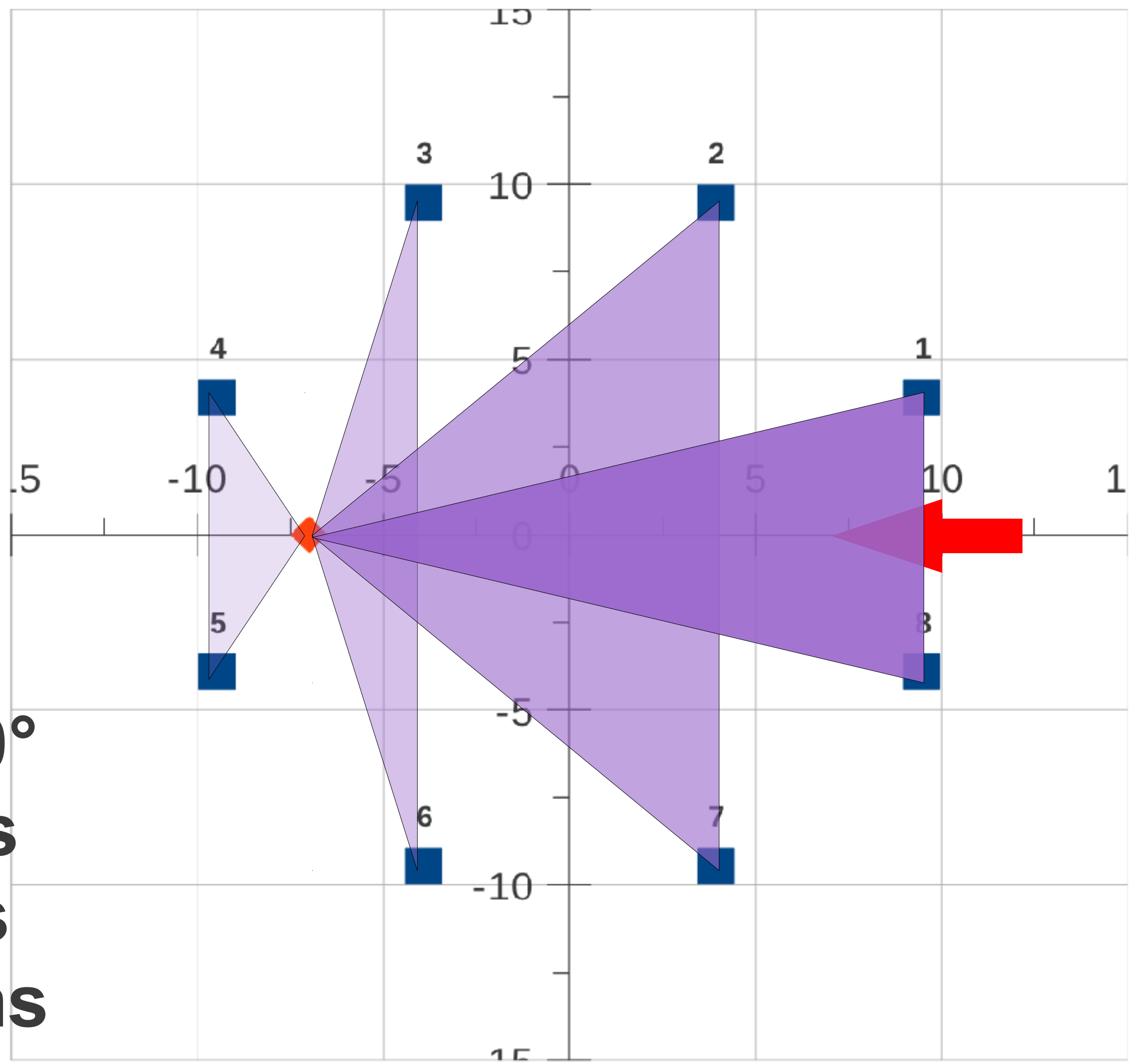


1st phantom source at 180°
2nd p.s. at 0°, +3dB, +16ms
3rd p.s. at 0°, +8dB, +29ms
4th p.s. at 0°, +18dB, +36ms

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Theoretical considerations:

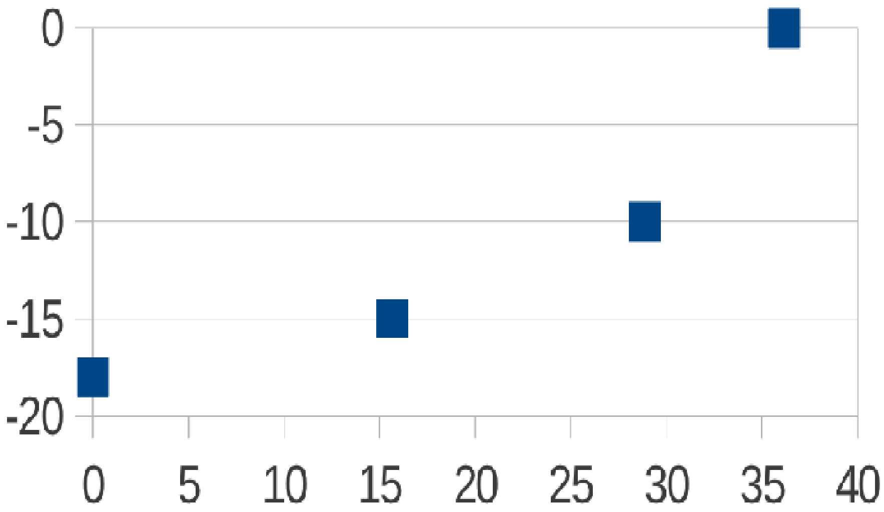




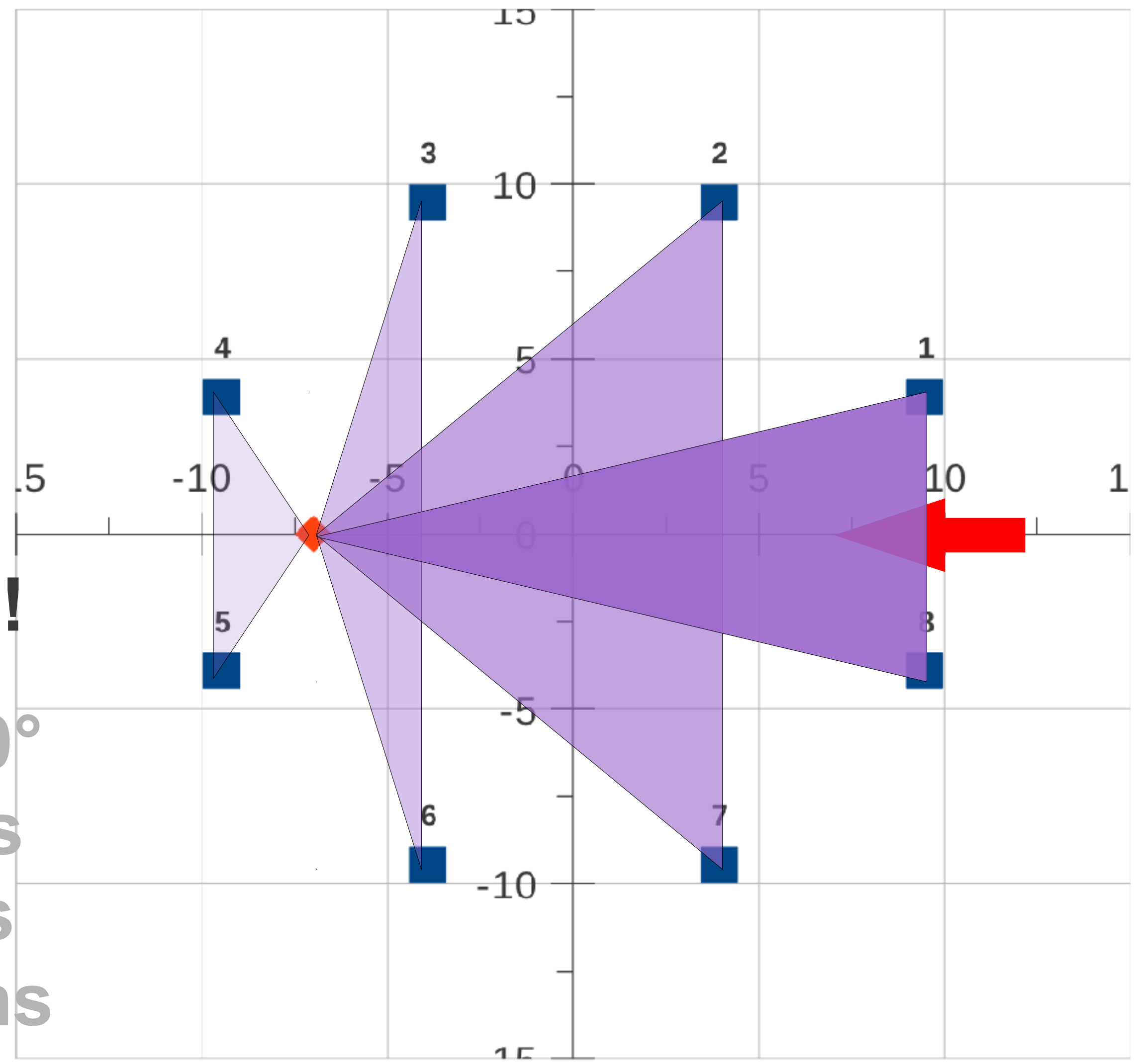
1st phantom source at 180°
2nd p.s. at 0°, +3dB, +16ms
3rd p.s. at 0°, +8dB, +29ms
4th p.s. at 0°, +18dB, +36ms

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Theoretical considerations:



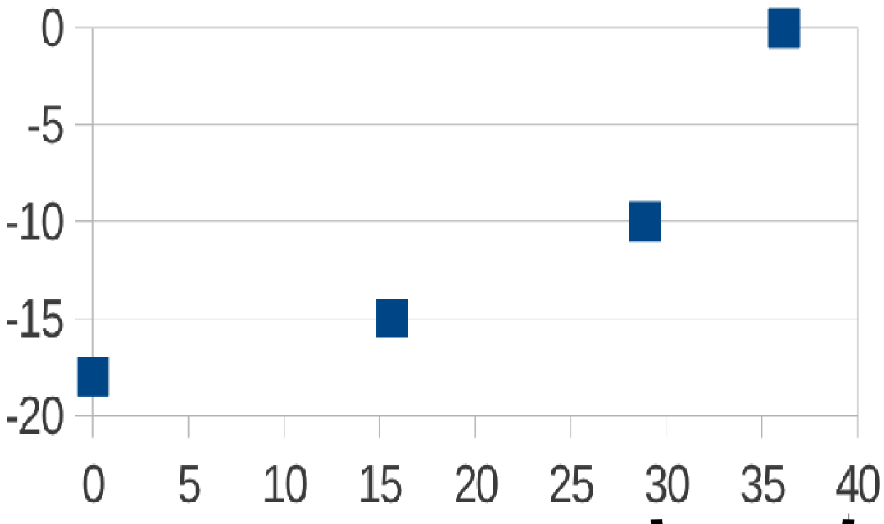
! ECHO threshold !



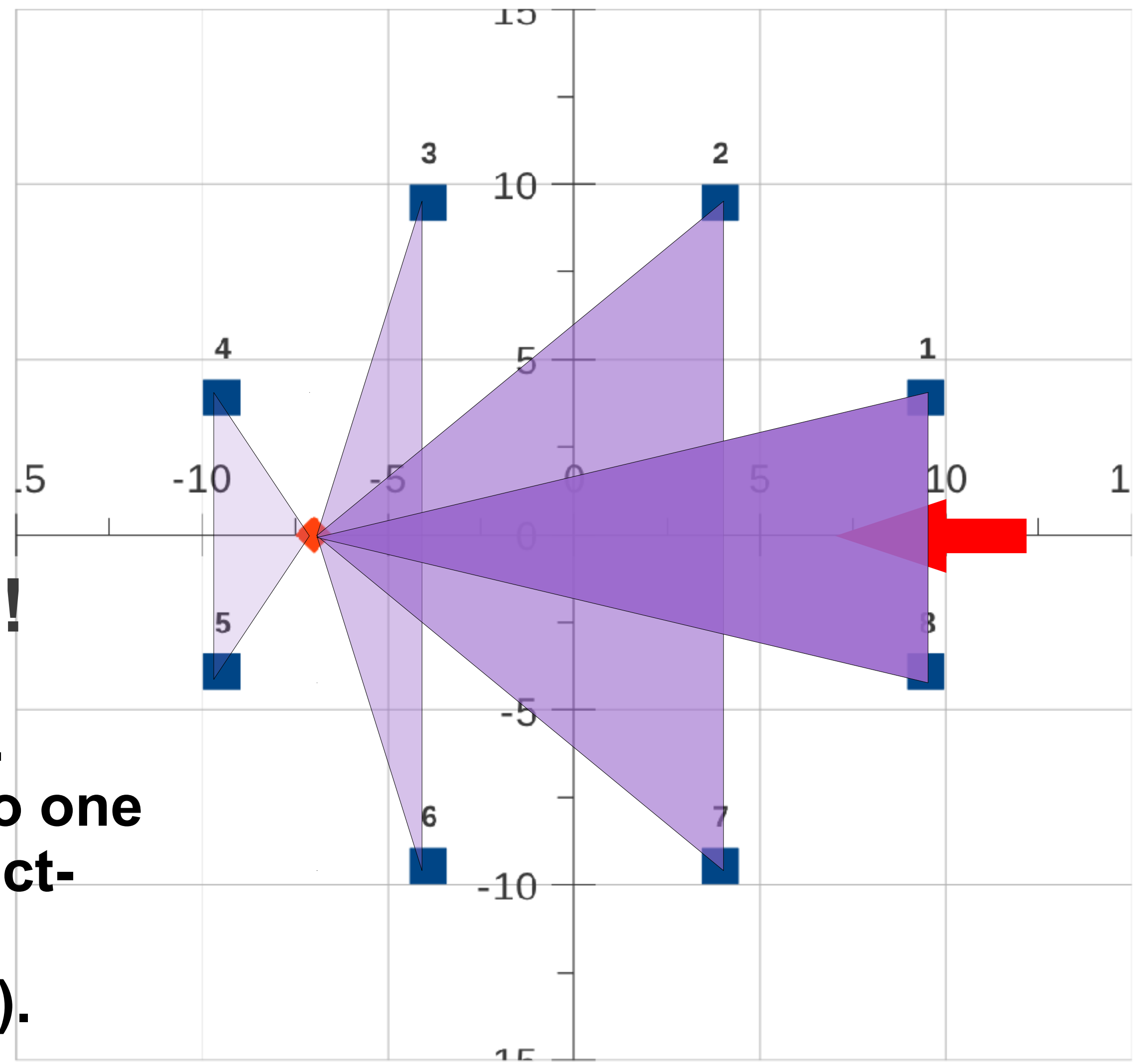
1st phantom source at 180°
2nd p.s. at 0°, +3dB, +16ms
3rd p.s. at 0°, +8dB, +29ms
4th p.s. at 0°, +18dB, +36ms

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Theoretical considerations:



! ECHO threshold !



We did not perceive any echoes. The sonic events were fused into one auditory event (→ listening expect-ation on large system is totally different from studio monitoring).

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Conclusion:

It works very well, although we don't quite understand why.

Backward masking?
Reduced precedence effect when confounded by multiple sonic events with rising level?
Trading effect of echo threshold vs. number of sonic events?

Further rigorous listening tests are necessary.

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

In practice, we feel confident that a 3rd-order Ambisonic system will deliver good results for at least 600 people ($300m^2$ or r~10m).

Reduced precedence effect when confounded by

Further rigorous listening tests are necessary.

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Acknowledgements:

Thanks to:

Adamson Europe
for the speaker systems
http://adamsonsystems.com

Cobra Sound Hamburg
for the space
http://cobrasound.de

The Linux Audio community
for the software.

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>

Thank you.

I'm looking forward to your questions.

The paper is available at
http://stackingdwarves.net/public_stuff/linux_audio/ambisonic_symposium_2011/AmbiSym2011-Nettingsmeier-Dohrmann_Large-scale_HOA_Systems.pdf

Jörn Nettingsmeier <nettings@stackingdwarves.net>, David Dohrmann <david@adamsonsystems.com>