Is My Decoder Ambisonic?

Aaron J. Heller

SRI International, Menlo Park, CA, US

Richard Lee

Pandit Litoral, Cooktown, QLD, AU

Eric M. Benjamin

Dolby Labs, San Francisco, CA, US

125th AES Convention, San Francisco
Session P9, Multichannel Sound Reproduction

3 Oct 2008 -- 2:30PM

Ambisonics

- Provides a mathematical encapsulation of auditory localization models
- A single recording can be reproduced on a variety of speaker arrays

But…

- Decoder must be matched to the speaker array geometry and listening conditions

Why test decoders?

- No controlling interest currently
- Current decoders are software written by enthusiasts
  - Many adjustments
  - Scant guidance
  - Users expected to listen and “tune”
  - Difficult to diagnose faults
- Software difficult to validate by inspection

Ambiguity

"The precise definition [of Ambisonics] has been ignored, and the term 'ambisonic' is now applied loosely to any system that makes use of circular or spherical harmonics."

Peter Craven, "The `Hierarchical' Viewpoint,"

Illusions in Sound -- AES 22nd UK Conference, 2007

Consequences

- Quality of information on the web is mixed
- Decoder writers
  - Many defective or improperly used decoders
- Researchers
  - What were they using for their work?
- Listeners
  - Confusing or unpleasant results

Definitions

- Localization models
- Ambisonic criteria

- Used to drive decoder design, *evaluation and validation*

Localization Models

- Two primitive models
  - Velocity localization vector, r_V
    - *ITD -- Blumlein,Clark, et al.*
  - Energy localization vector, r_E
    - *ILD -- Fransen, Mertens, …*
- *Direction* indicates direction of localization perception
- *Magnitudes* indicates quality and stability
  - *In natural hearing, magnitude is 1*
- *Different approach needed for each regime*

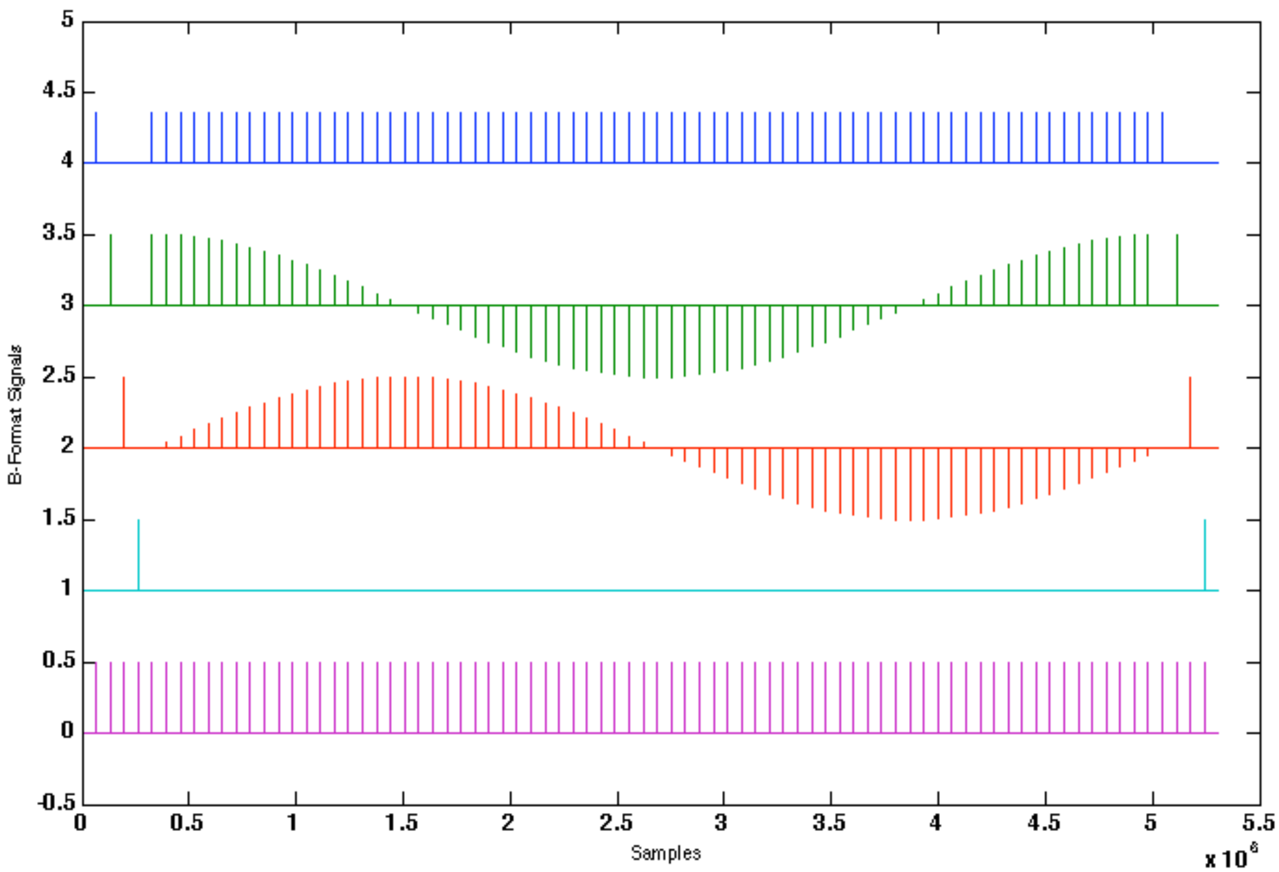
Ambisonic Criteria

- Gerzon's definition
  - Velocity and energy vector directions are the same up to around 4 kHz and are largely unchanged with frequency.
  - At low frequencies, the magnitude of the velocity vector is near 1 for all directions.
  - At mid/high frequencies the energy vector is maximized over as many directions as possible.
- Necessary (if perhaps not sufficient) for good surround sound reproduction
- Confirmed by listening tests

Test Procedure

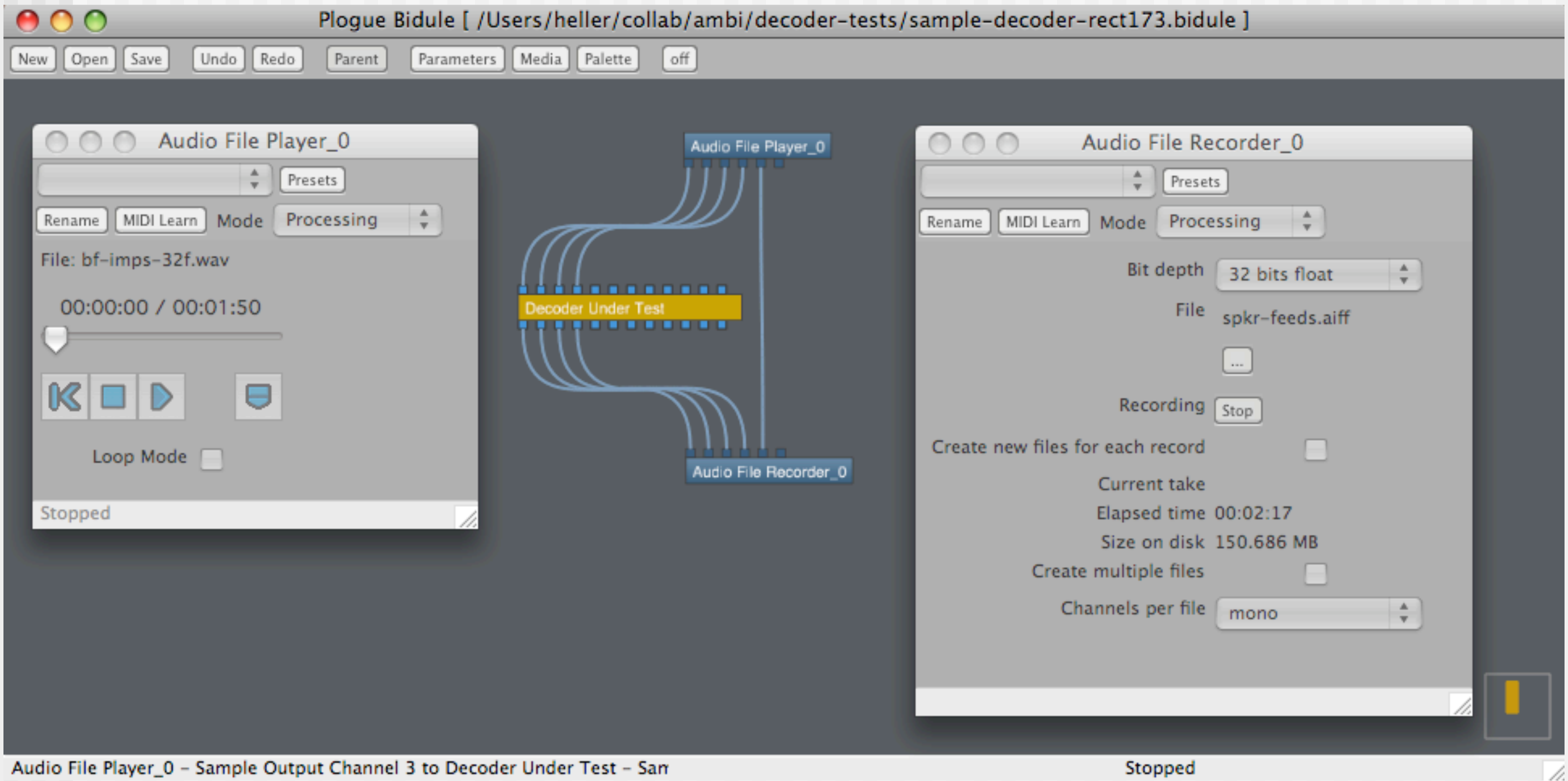
- Measure Impulse Response from a variety of directions
- Evaluate those against the Ambisonic criteria
- Current paper examines
  - A single speaker array ($\sqrt{3}$: 1 rectangle)
  - Four decoders
- Matlab code to generate test signals and analyze results

Test signal



B-Format Signals
Samples
x 10^6
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
-0.5
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5

Typical Test Harness



Speaker Array Geometry

- Regular polygons and polyhedra
  - Often difficult to fit into real rooms
- Irregular, but diametric opposite pairs
  - Rectangles, bi- and tri-rectangles
- General irregular arrays
  - ITU 5.1, hemispheres

- *Assumption that all arrays can be treated as regular polygonal is the most common error*

Components of a decoder

- Decoder matrix matched to speaker array geometry
- Phase-matched dual-band processing
- Near-field compensation

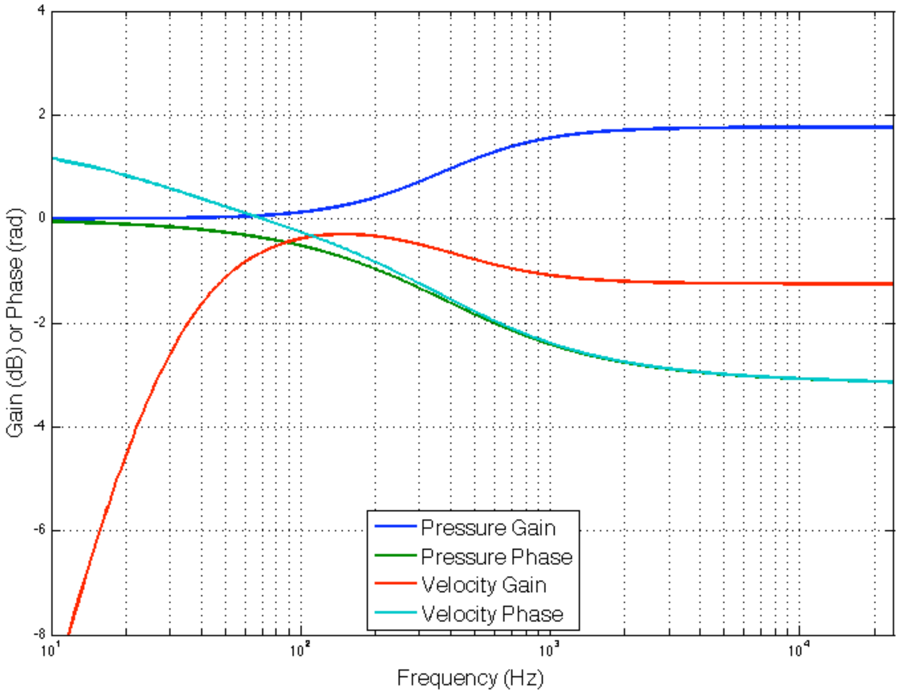
- Cookbook design procedures for all three components in Appendix.

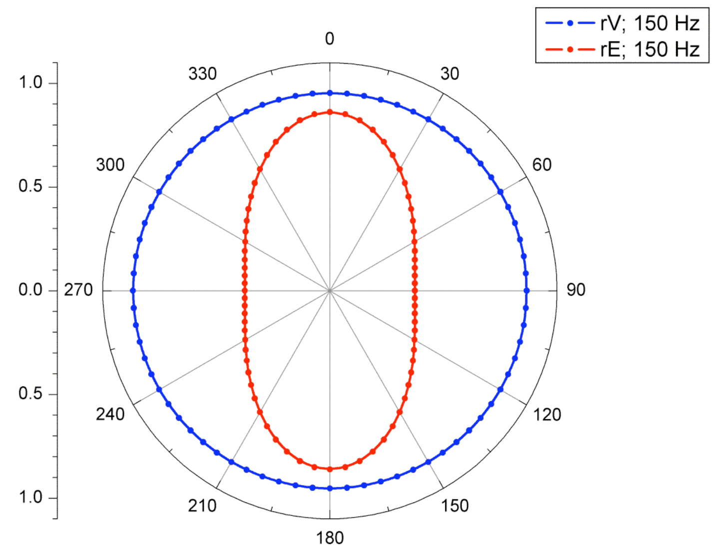
- *Lack of dual-band processing is another common problem*
  - *Poor localization or comb filter artifacts*

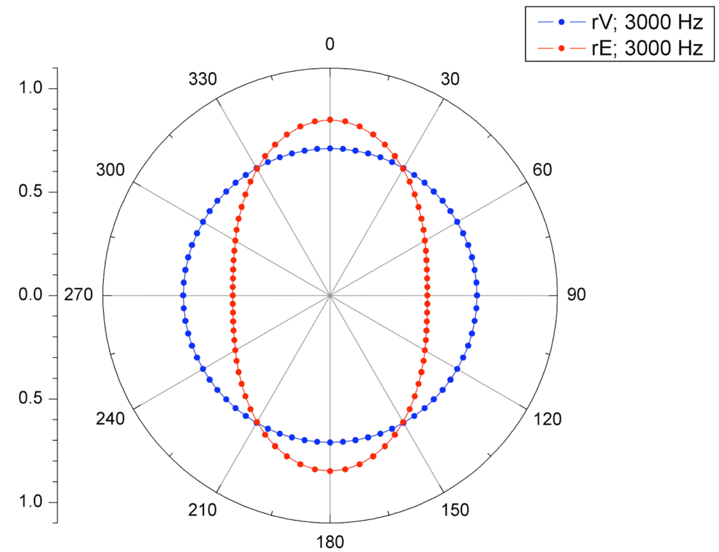
Types of Decoders

- Matrix and other parameters entered directly
  - Adriaensen's *AmbDec*
- Presets for various array geometries
  - Csound, CMT, …
- Virtual Microphones
  - Many VST and AU plugins

AmbDec

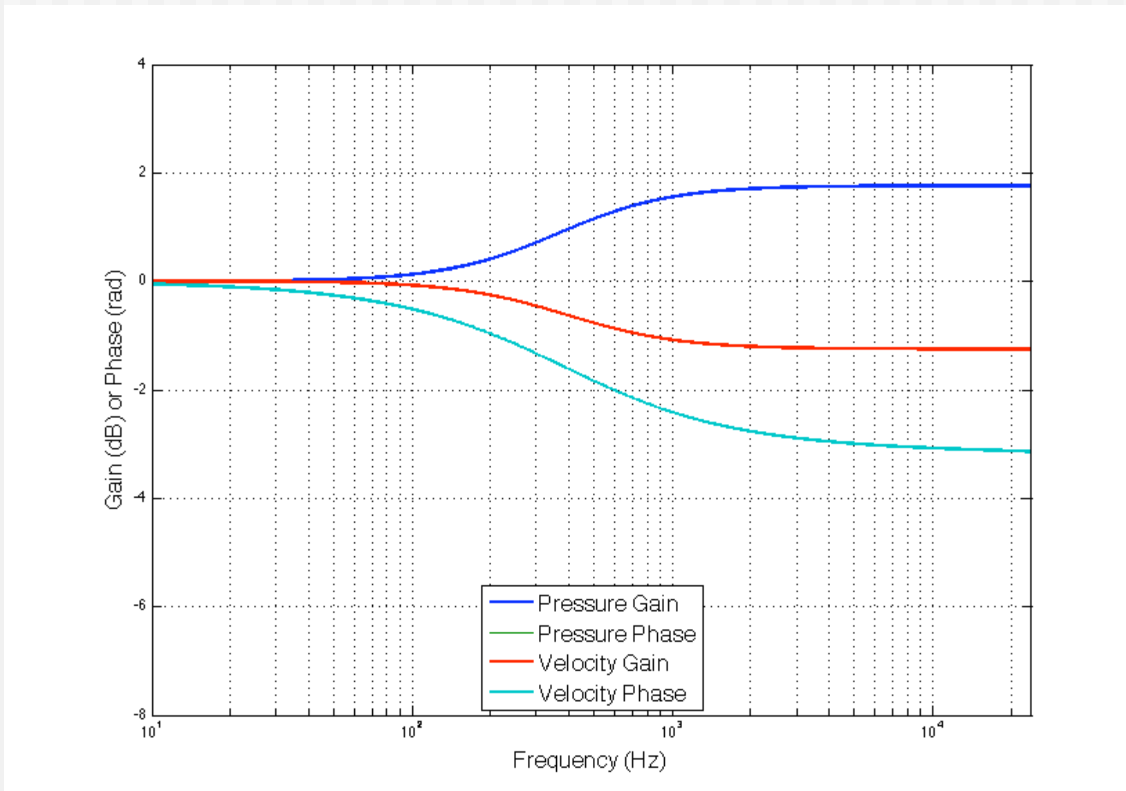


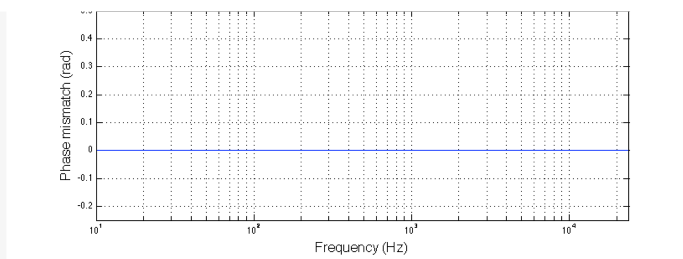




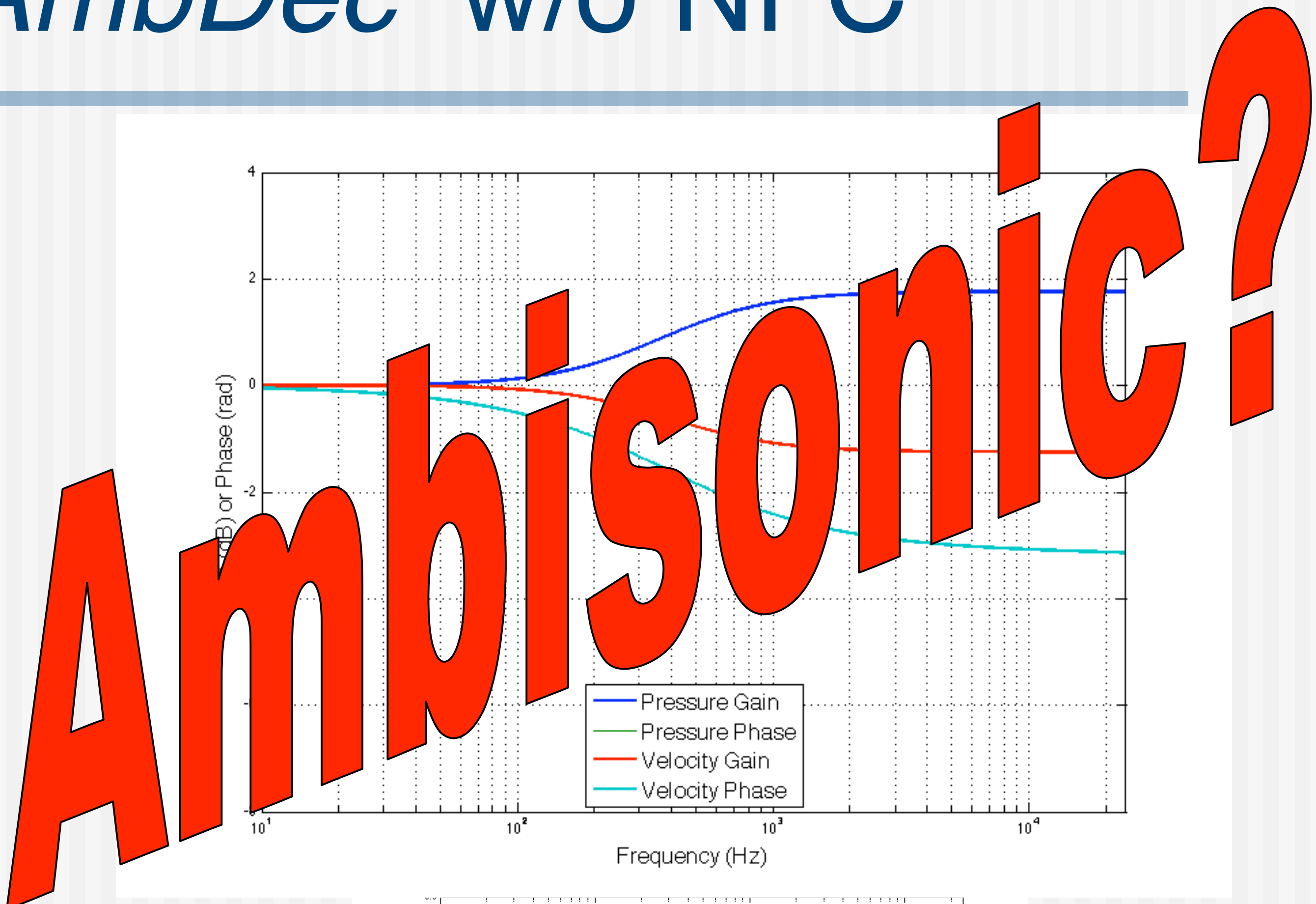
Decoder matrix and parameters derived by procedures in appendix

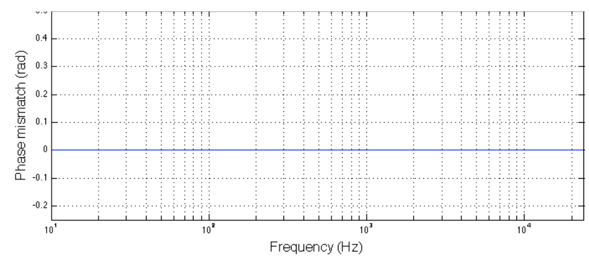
AmbDec w/o NFC



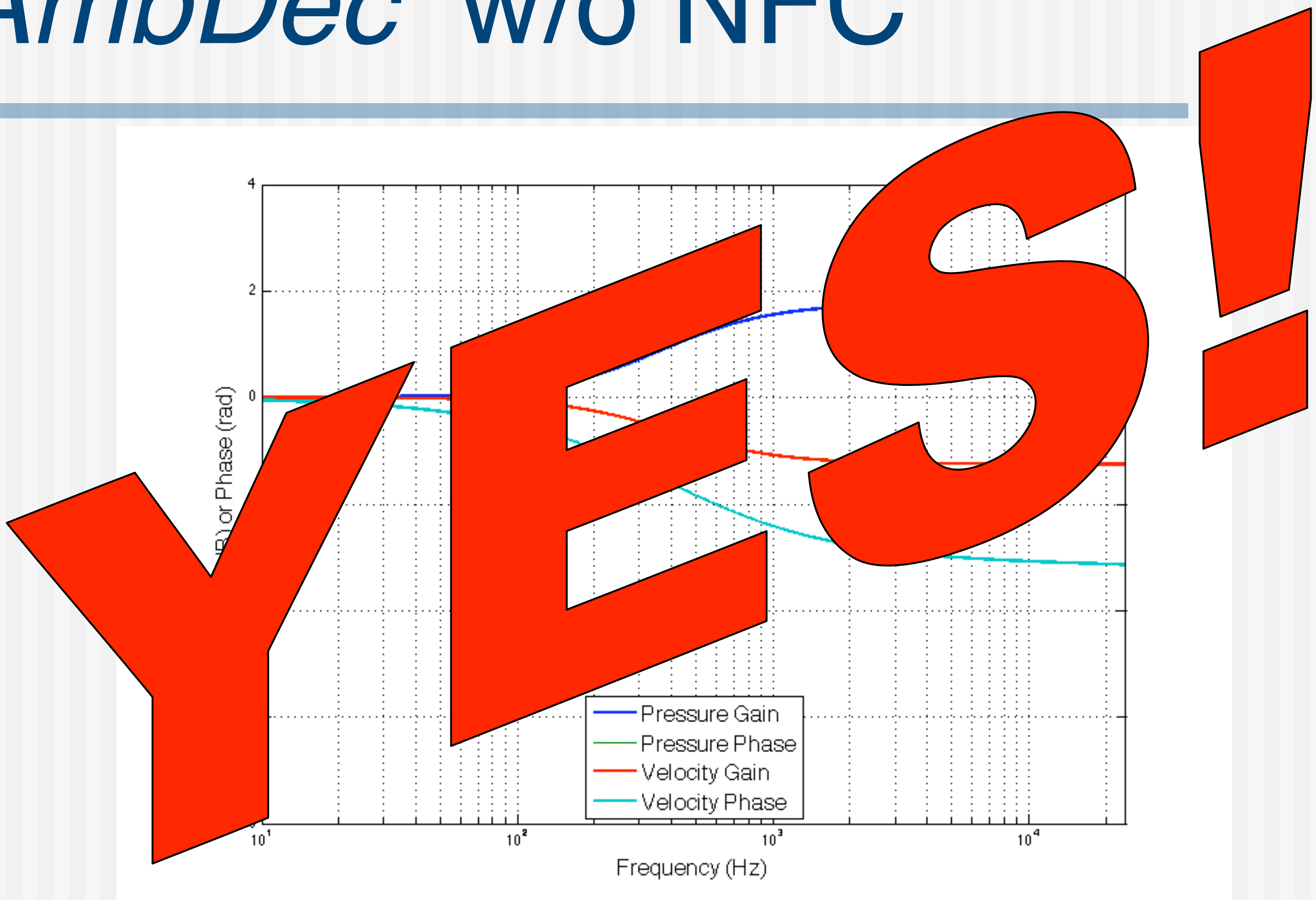


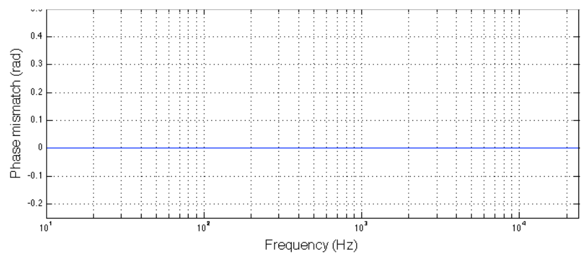
AmbDec w/o NFC





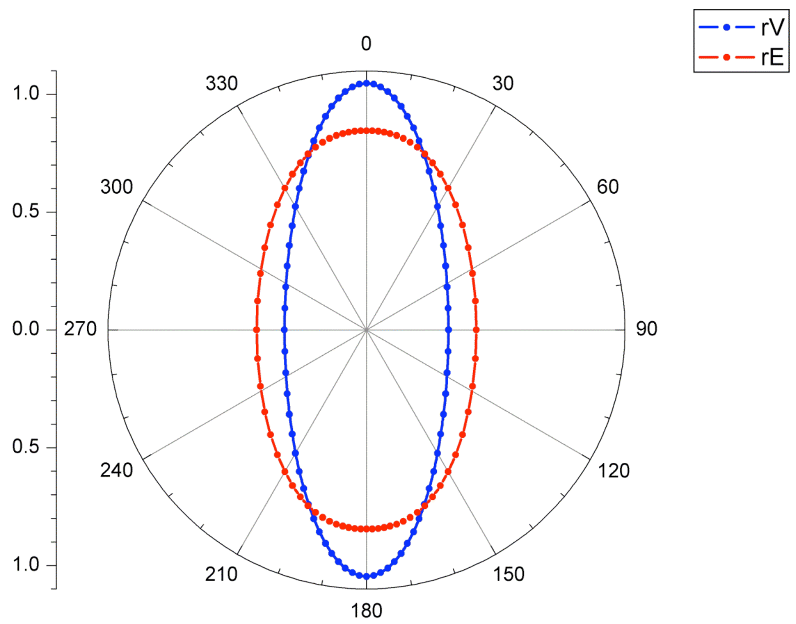
AmbDec w/o NFC



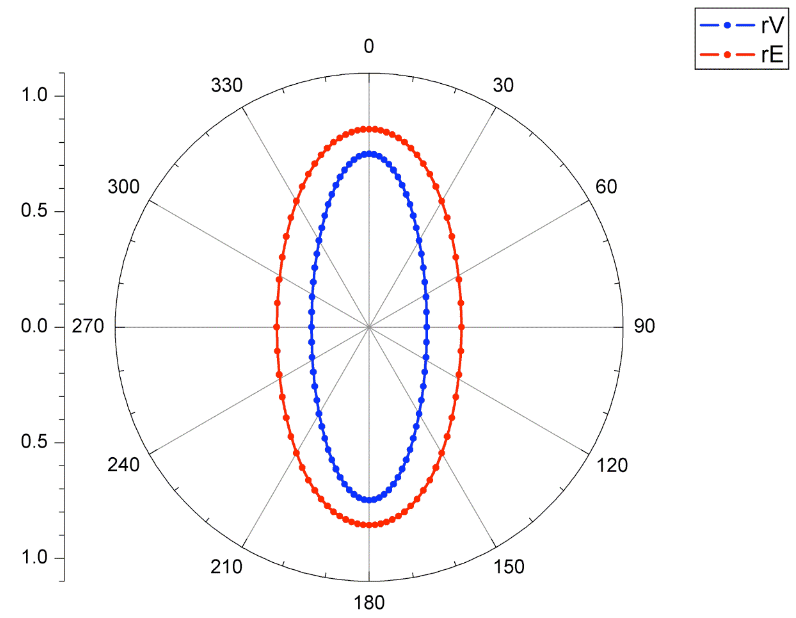


VST Plugin (virtual mic type)

150 Hz



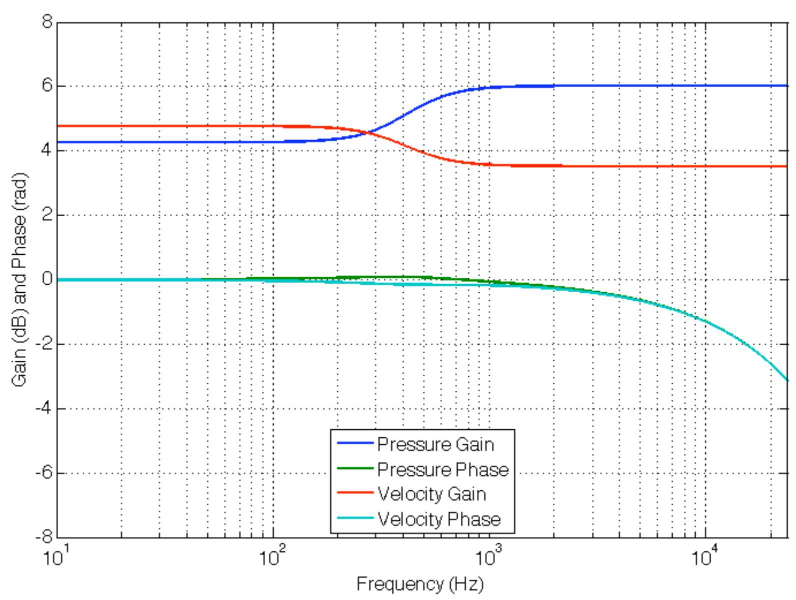
3 kHz



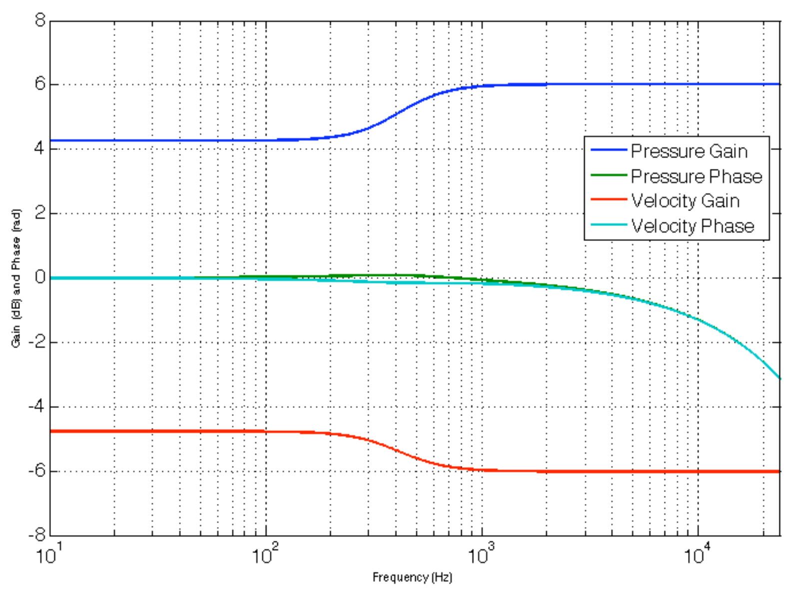
Virtual mics pointed at loudspeakers per directions. Other parameters left at default settings.

VST Plugin

0 degrees



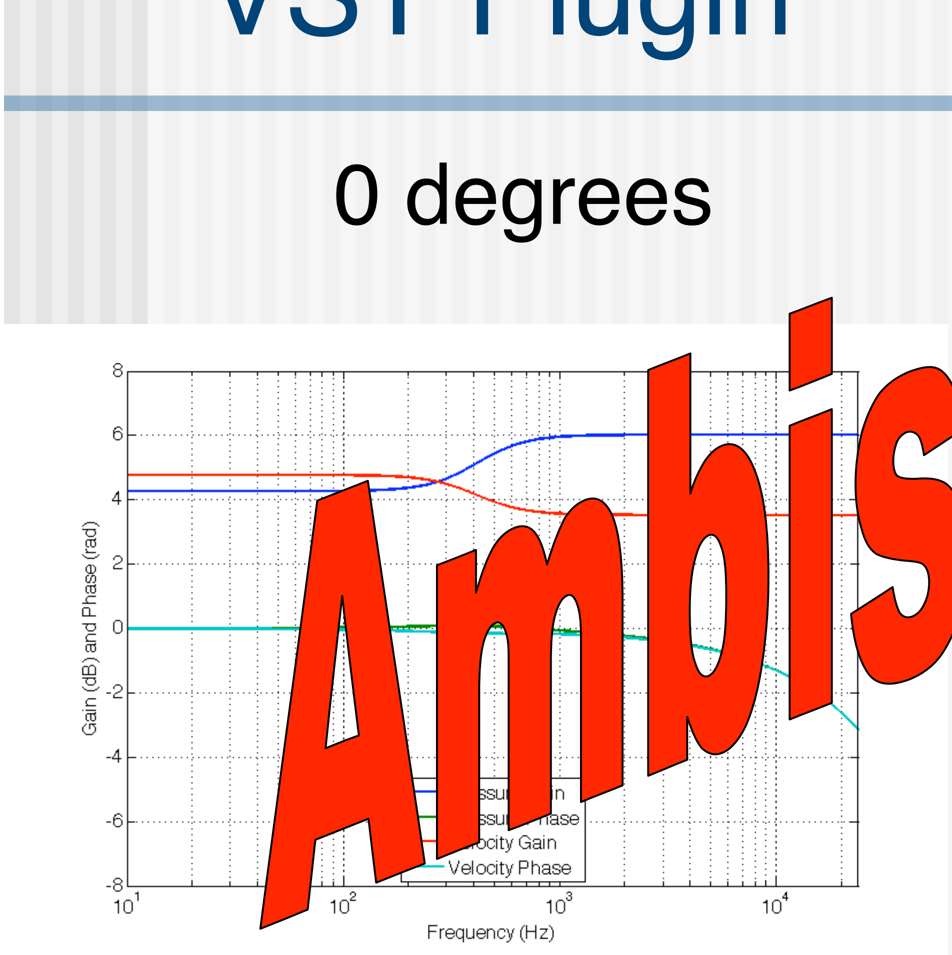
90 degrees

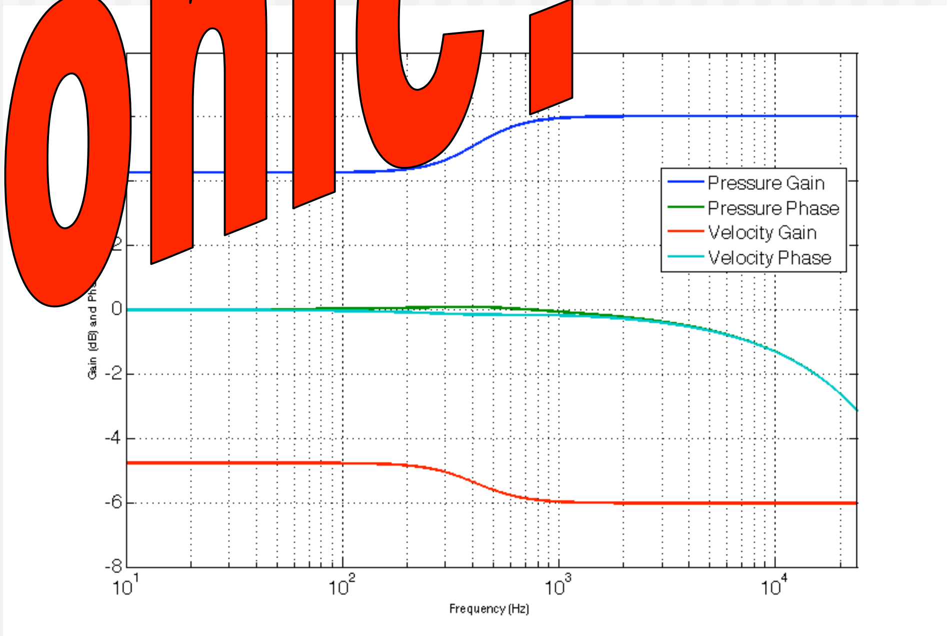


VST Plugin

0 degrees

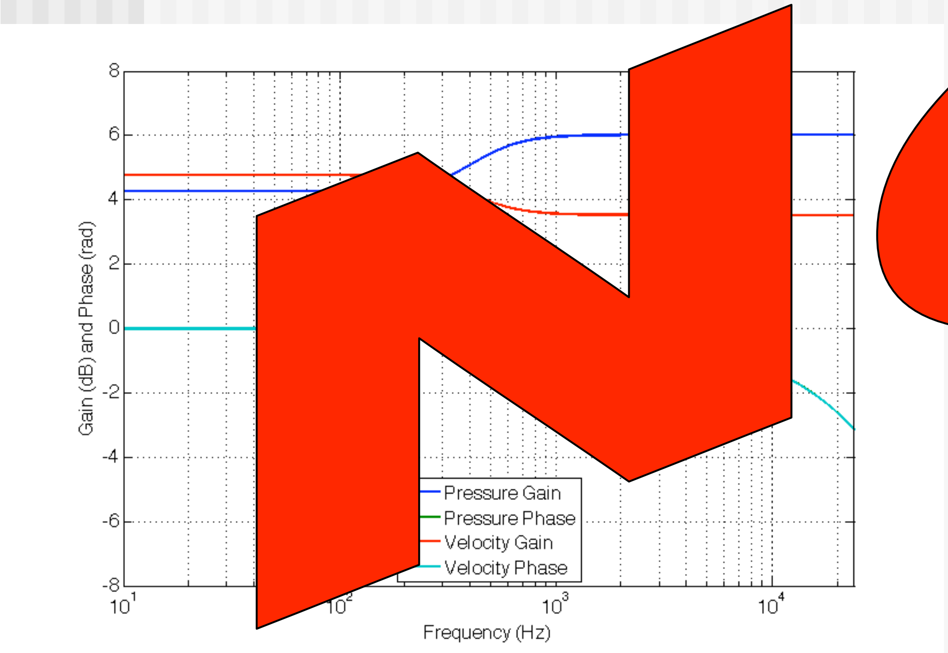
90 degrees



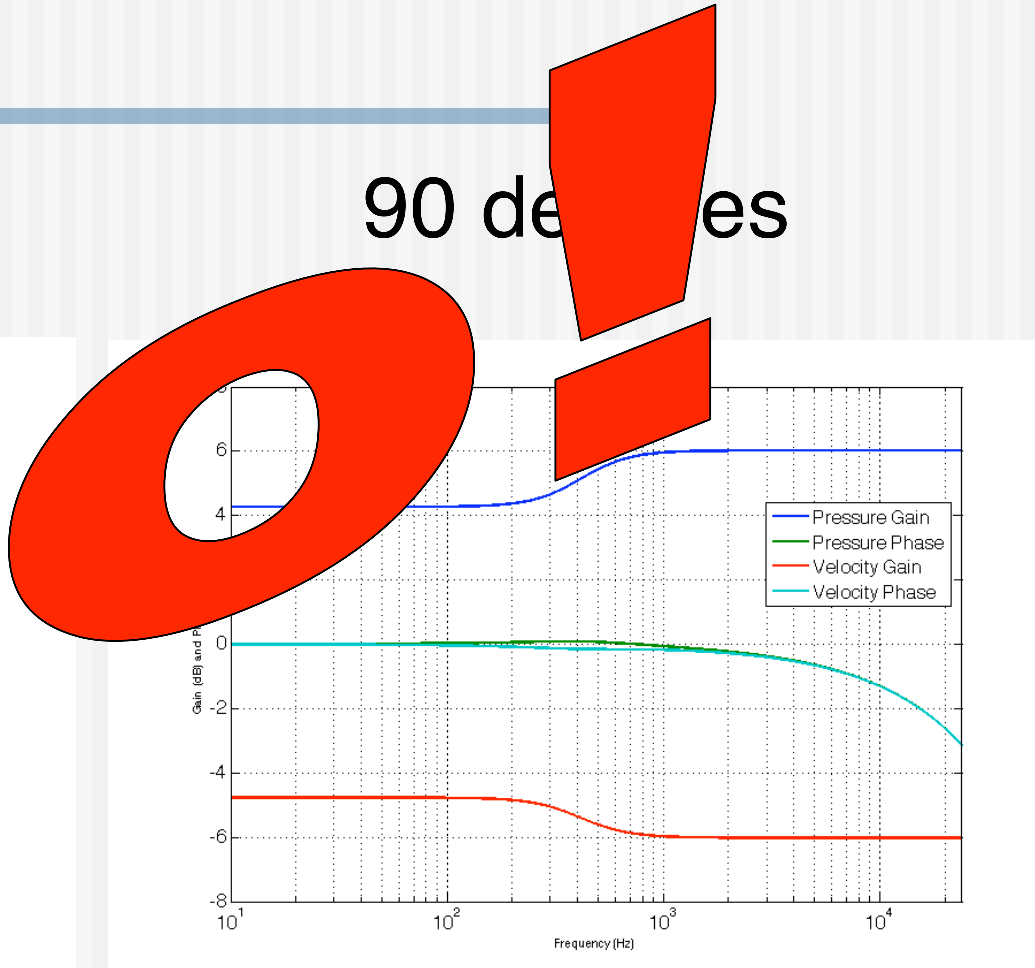


VST Plugin

0 degrees

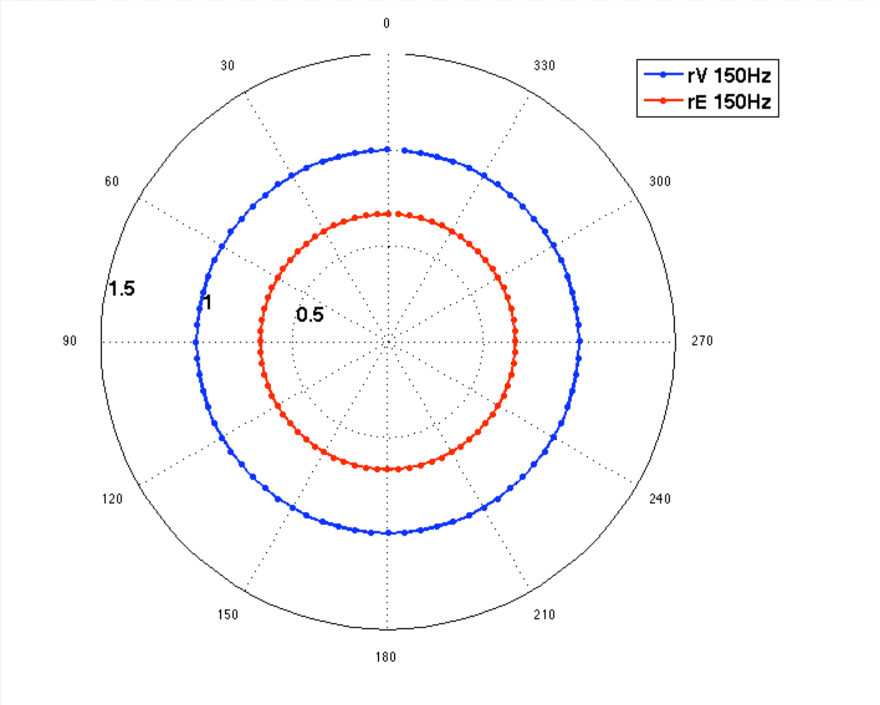


90 degrees

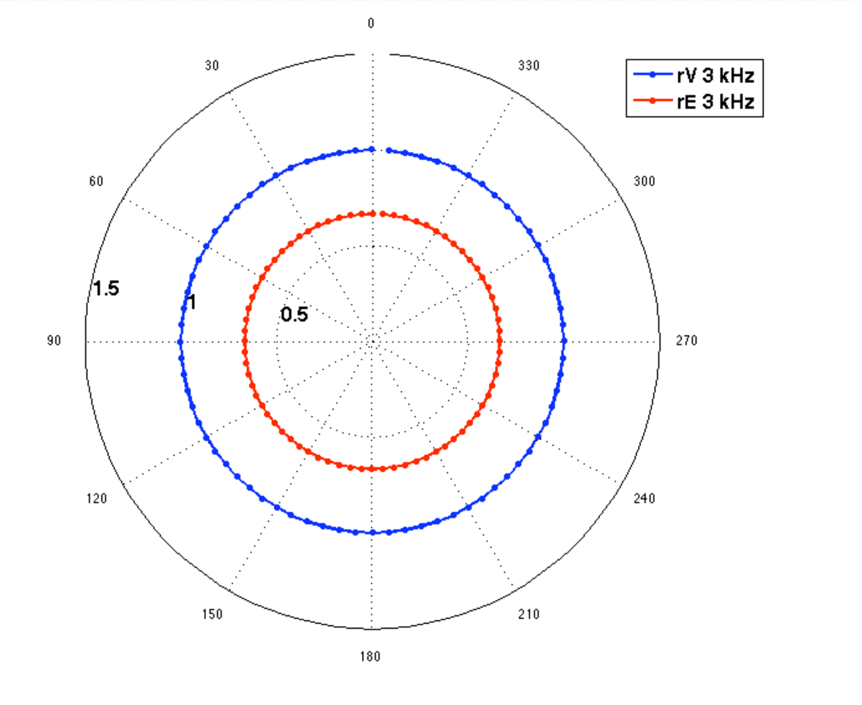


Csound “bformdec” opcode

150 Hz

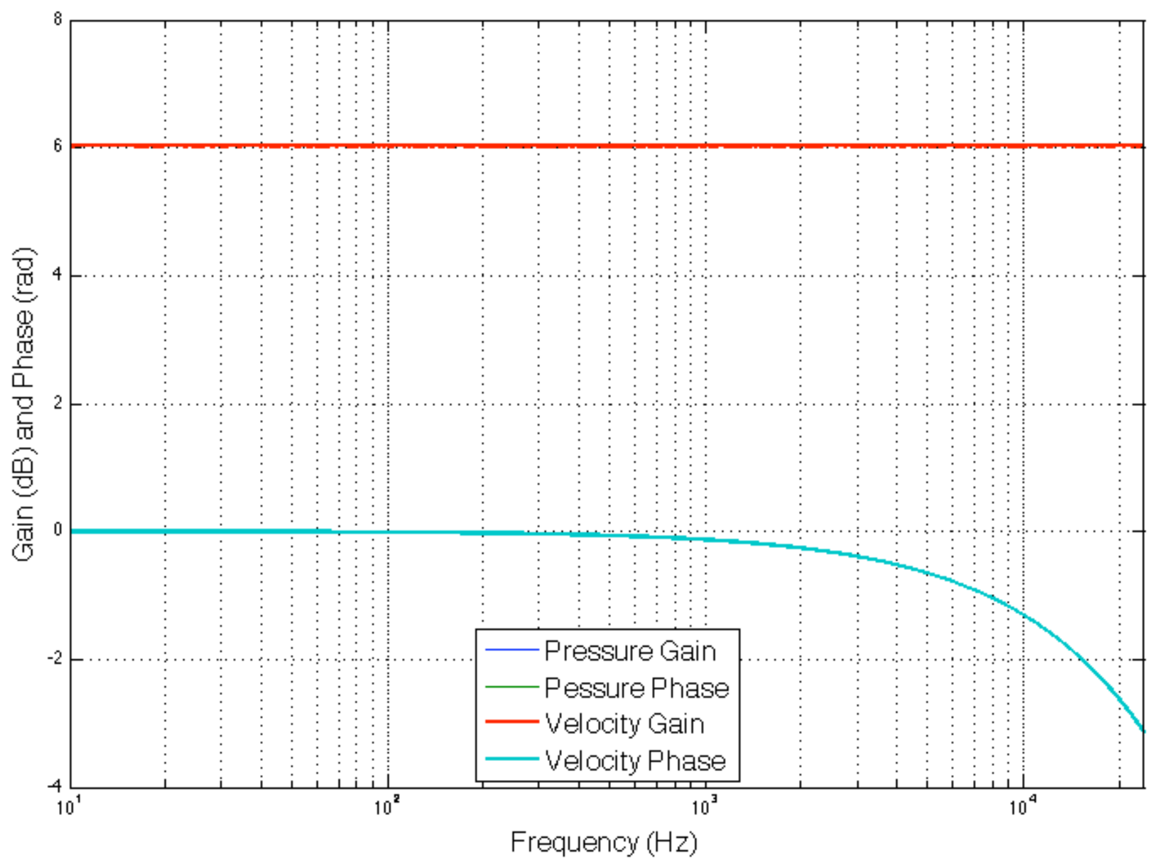


3 kHz



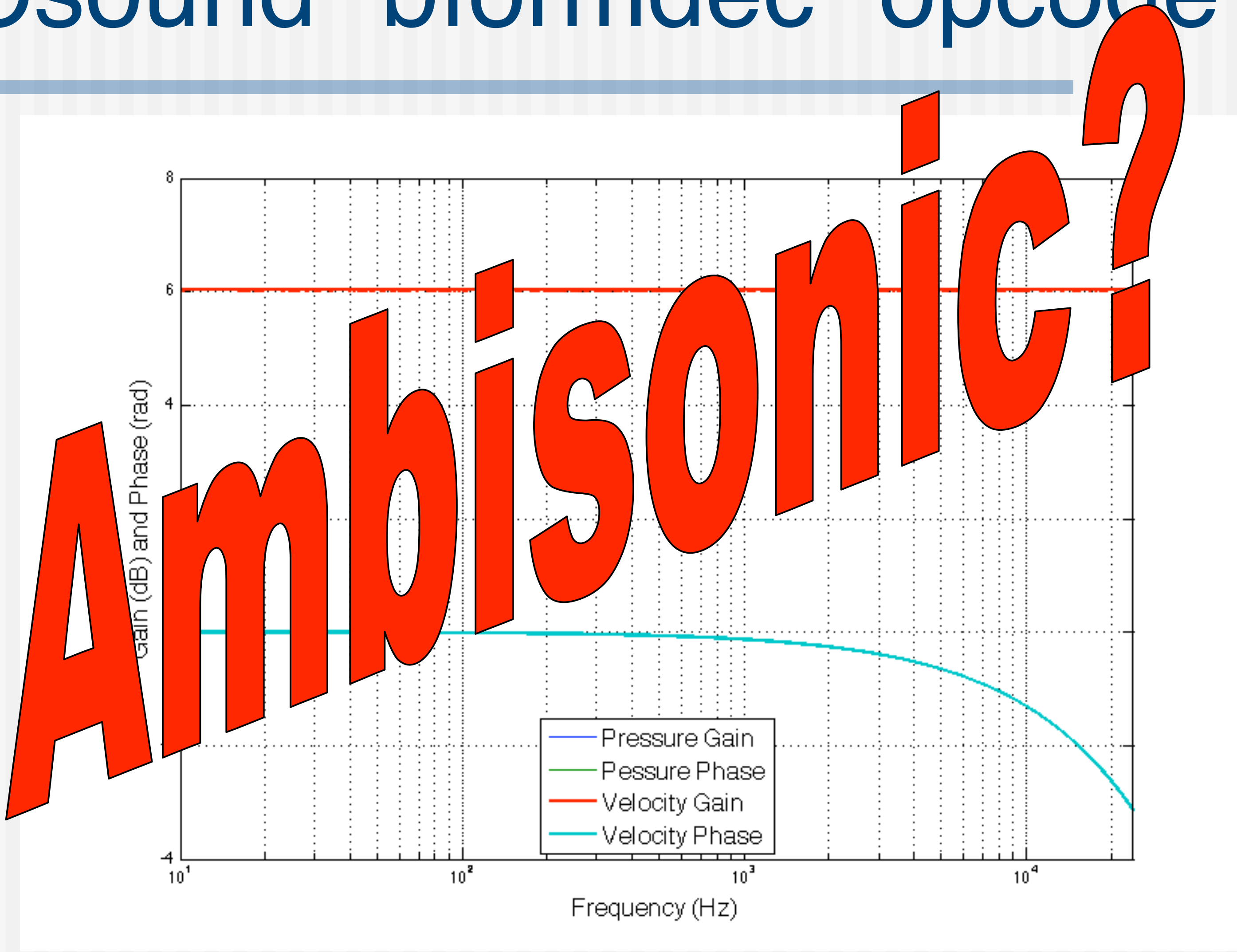
Tested square decoder

Csound "bformdec" opcode



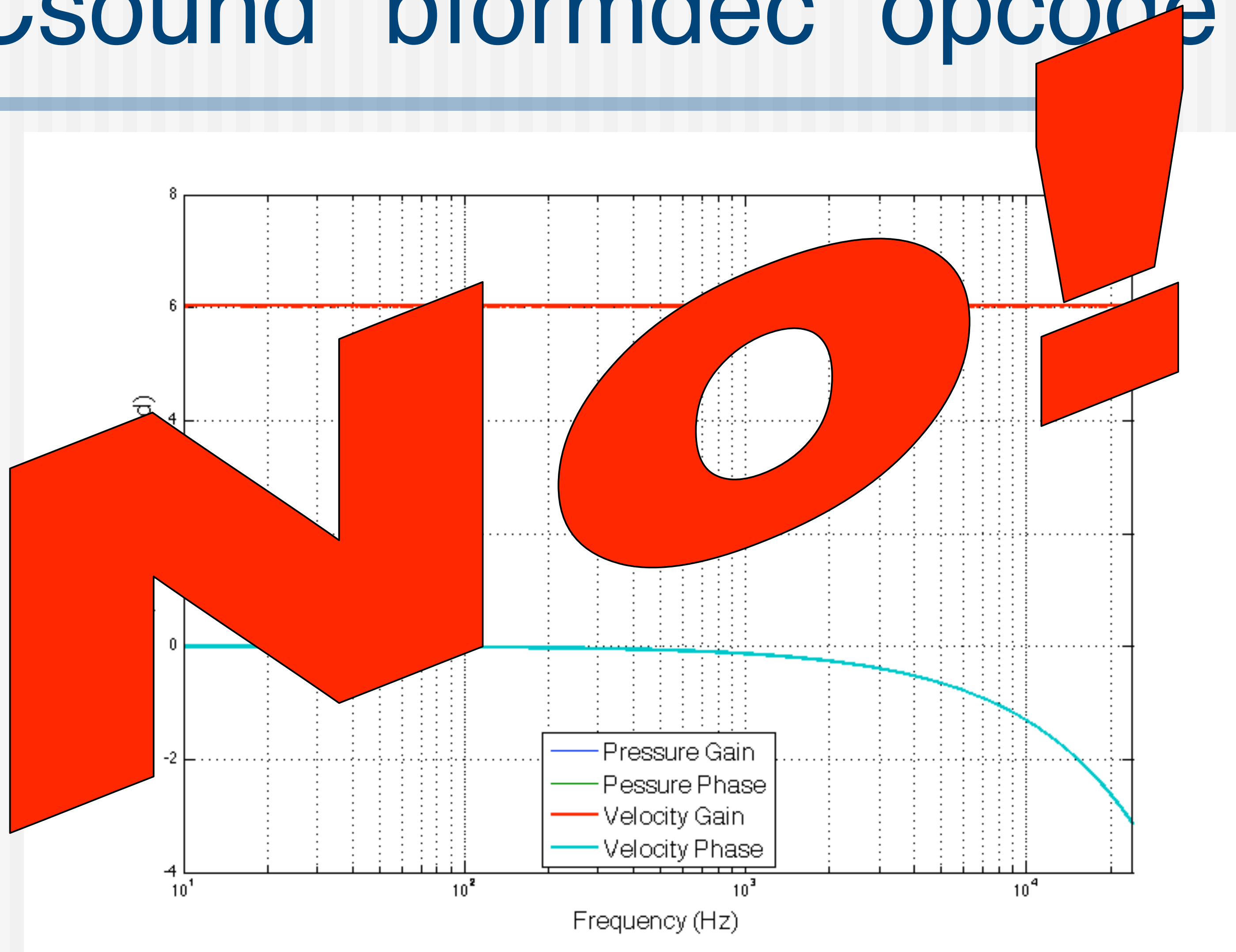
Tested square decoder

Csound “bformdec” opcode



Tested square decoder

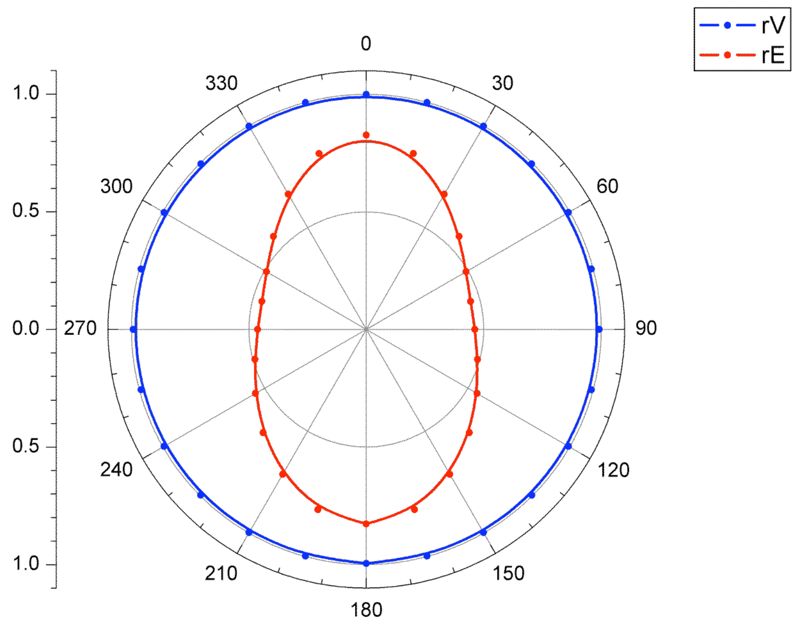
Csound “bformdec” opcode



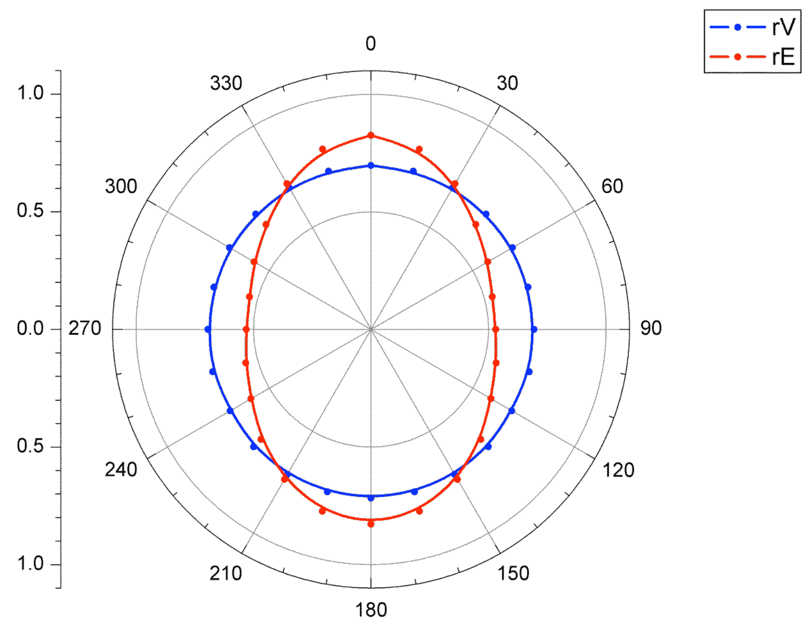
Tested square decoder

Minim AD-10

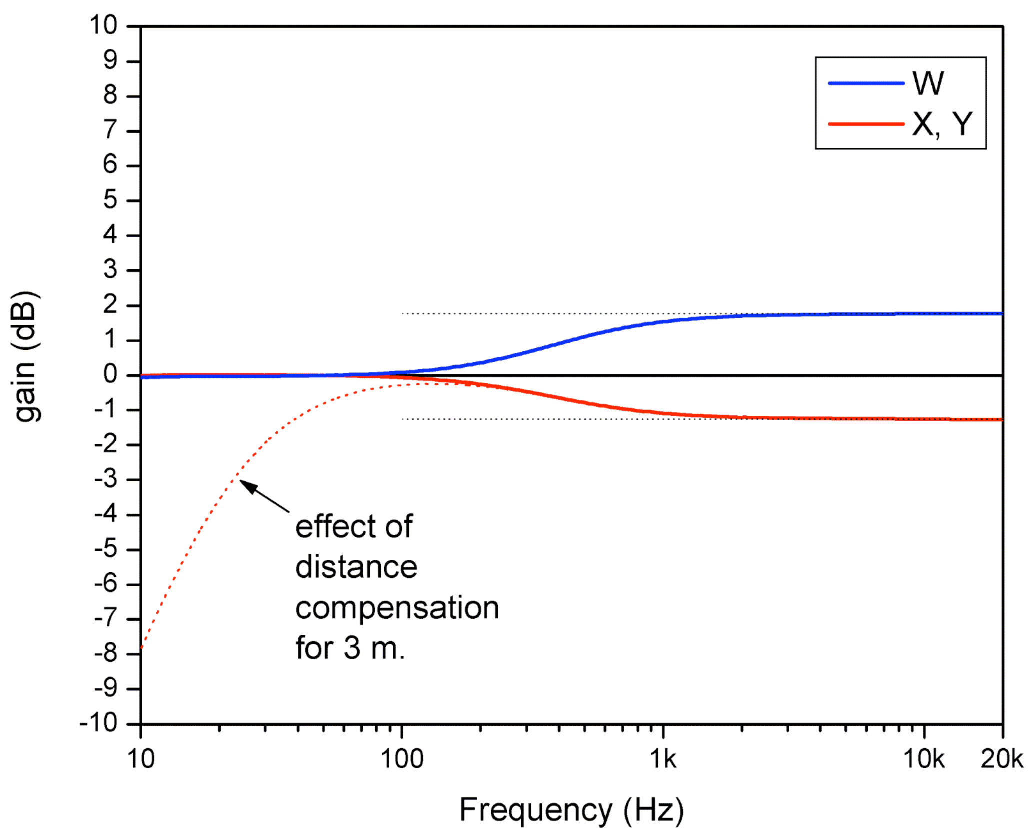
150 Hz



3 kHz

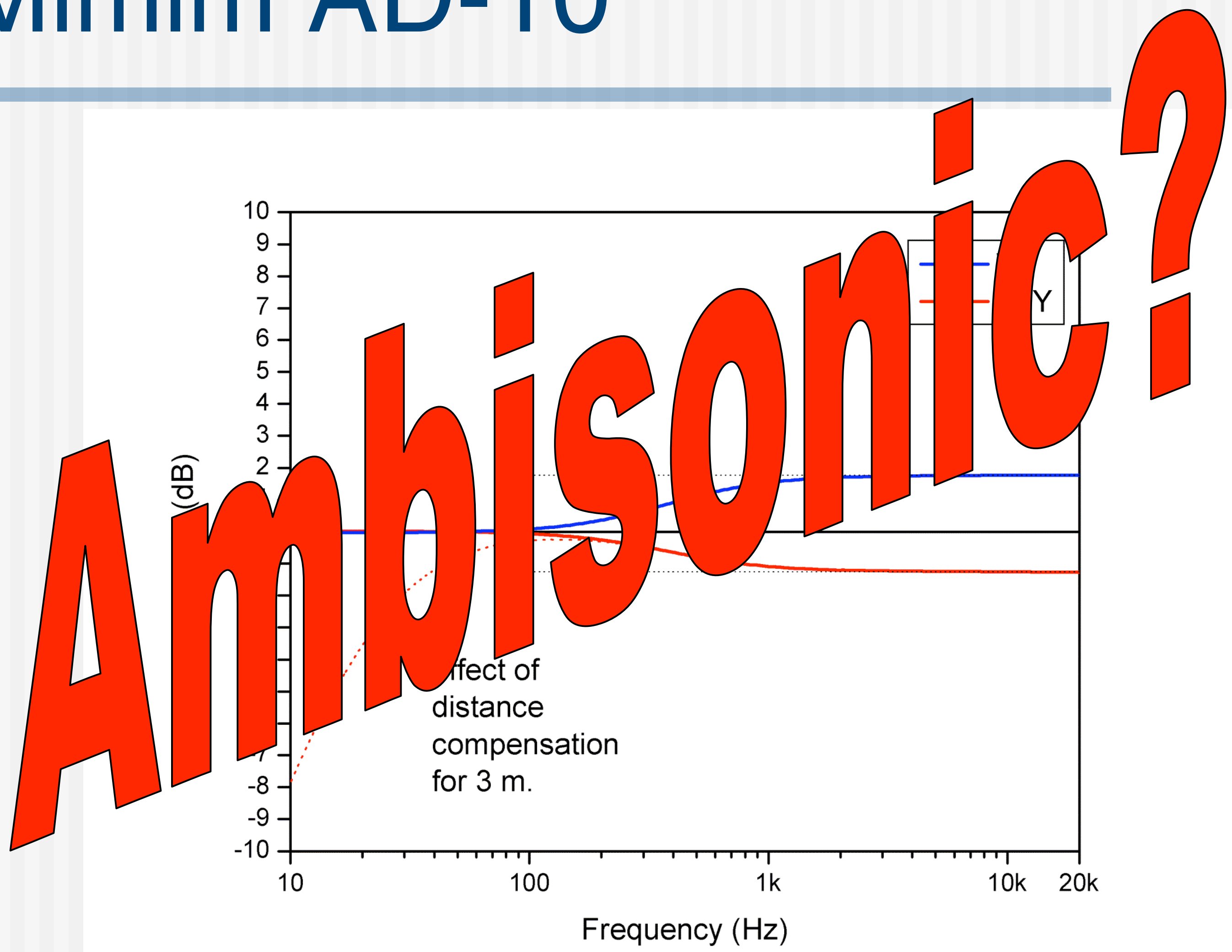


Mimim AD-10



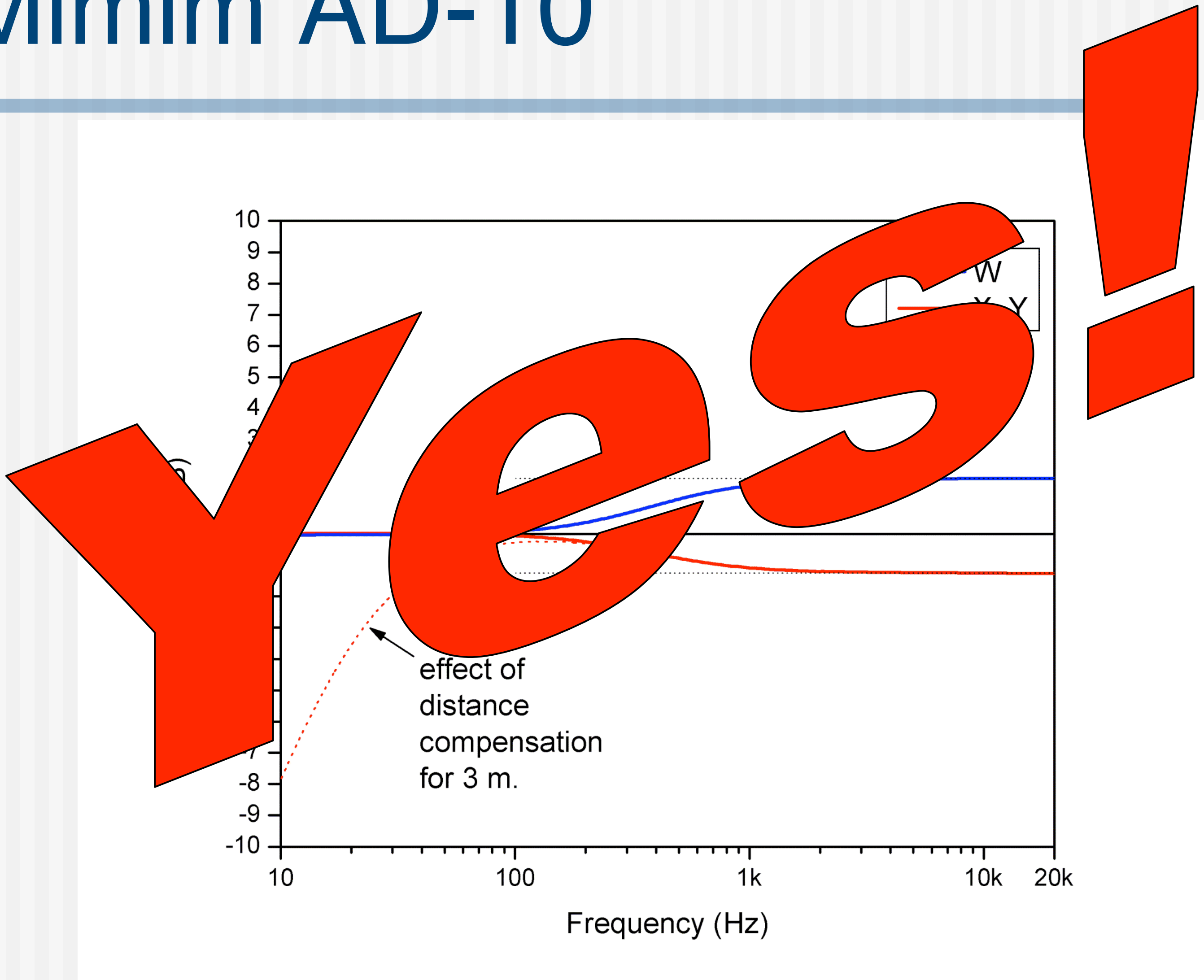
W
X, Y
effect of distance compensation for 3 m.
gain (dB)
Frequency (Hz)
10
100
1k
10k
20k

Mimim AD-10



Ambisonic?
10
9
8
7
6
5
4
3
2
(dB)
Y
-8
-9
-10
10
100
1k
10k
20k
Frequency (Hz)
distance
compensation
for 3 m.

Mimim AD-10



Yes!
10
9
8
7
6
5
4
-8
-9
-10
10
100
1k
10k
20k
Frequency (Hz)
W
effect of
distance
compensation
for 3 m.

Informal Listening Tests

- Same material as earlier tests
- Ambdec
  - Good localization and envelopment
  - No audible artifacts
- Decoder 2 (VST Plugin)
  - Front and rear localization only
- Csound “bformdec” (simulated)
  - Comb filtering and in-head localization artifacts

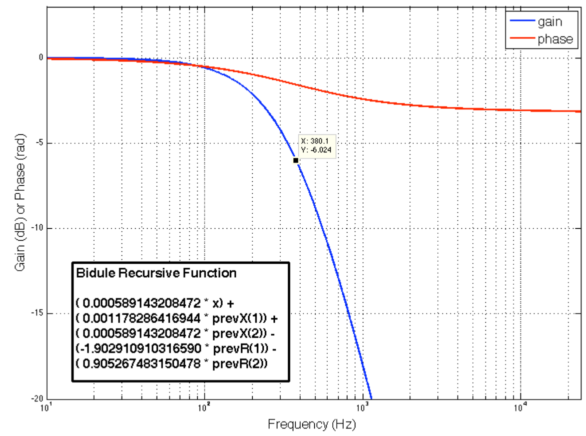
Decoder Design

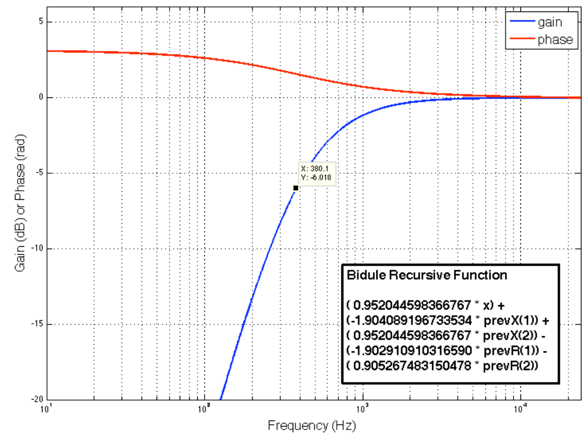
- Decoder matrix derived by generalized inversion
  - Pick a basis set of spherical harmonics
  - "project" speaker locations onto basis set
  - Goal - reproduce basis set (exact solution)
    - Many solutions, want minimum radiated power
  - Use Moore-Penrose Pseudo-Inversion
    - Singular Value Decomposition

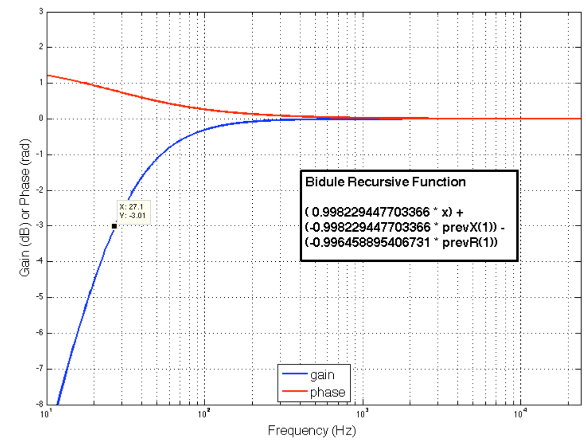
      $A = U \Sigma V^* \;\; \text{->} \;\; A^\dagger = V \Sigma^\dagger U^*$
  - pinv() in Matlab and Octave
  - Equivalent to Least-Squares solution
    - Minimum radiated power, highest average r_E

Decoder Design

- Phase-matched bandsplitting and NFC filters
- Cookbook procedures for design
- Sample implementation using Bidule recursive function block







Is My Encoder Ambisonic?

- Ambisonics can encode
  - Distance, diffuse fields, standing wave
- In fact, a properly aligned Ambisonic microphone must do this.
  - This is the proximity effect in all directional microphones
- Hence, Ambisonic panner/encoder should have these as well.
- See paper for details.

Conclusions

- Most decoders do not meet Ambisonic criteria
  - Incorrect coefficients for irregular arrays
  - Lack of dual-band decoding
  - Lack of near-field compensation
- Results in
  - Poor localization
  - Uncomfortable effects
- Good B-format material is now available
- Next, we need easy-to-use playback software

Further info

- Read the paper
- Web site URL
  - http://www.ai.sri.com/ajh/ambisonics
  - http://www.ambisonia.com
- Demonstration tonight, 6 - 9PM
  - Bubble, 73 Langton St, SF (3 blocks from Moscone)
  - 24-speaker hemispherical array
  - Decoder derived via techniques described here