Binaural Audio Externalization Processing

Jean-Marc Jot, Alexey Lukin, Kurt James Werner, Evan Allen

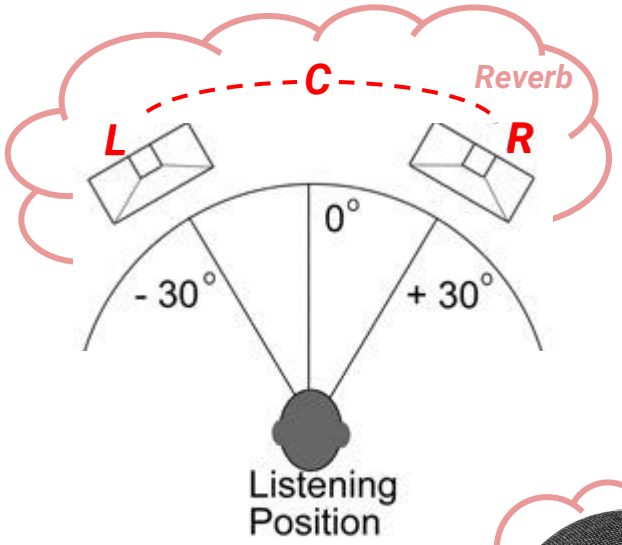
iZotope, Inc.

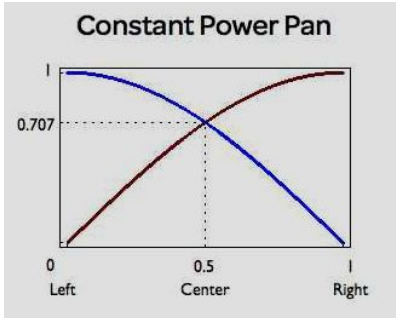
Audio examples are posted at izotope.com/tech/aes_extern .

Correspondence should be addressed to jjot@izotope.com.

Spatial reproduction of 2-channel stereo recordings

Over loudspeakers





Over headphones

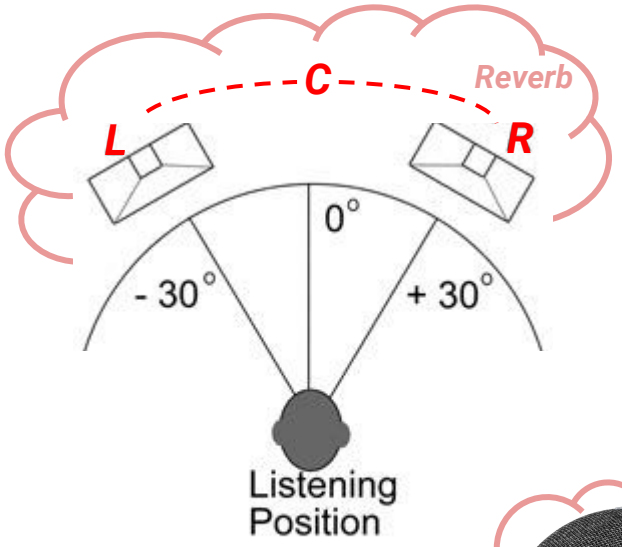
Sounds localized "in head" and in the frontal plane

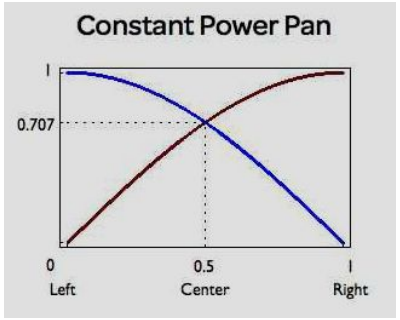
(the vertical plane joining the ears).

Listening fatigue.

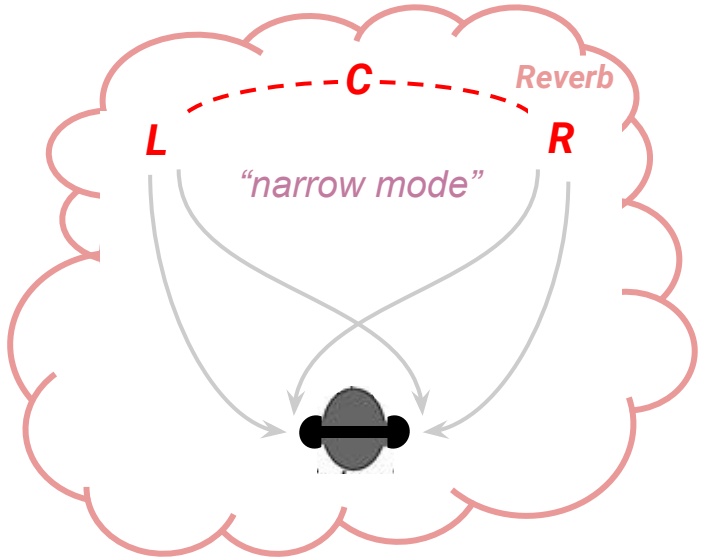
Aspirational binaural listening experience

Over loudspeakers





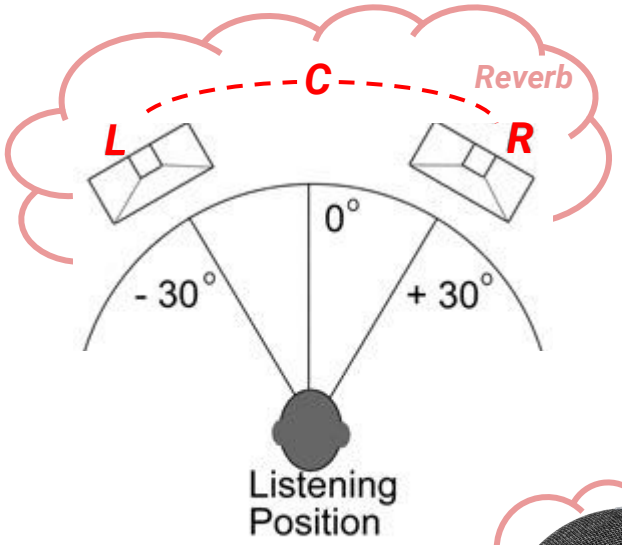
Over headphones

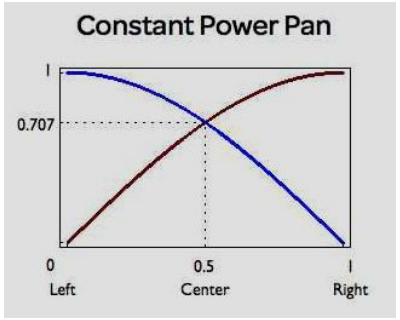


Sounds appear "externalized" and localized in front of listener

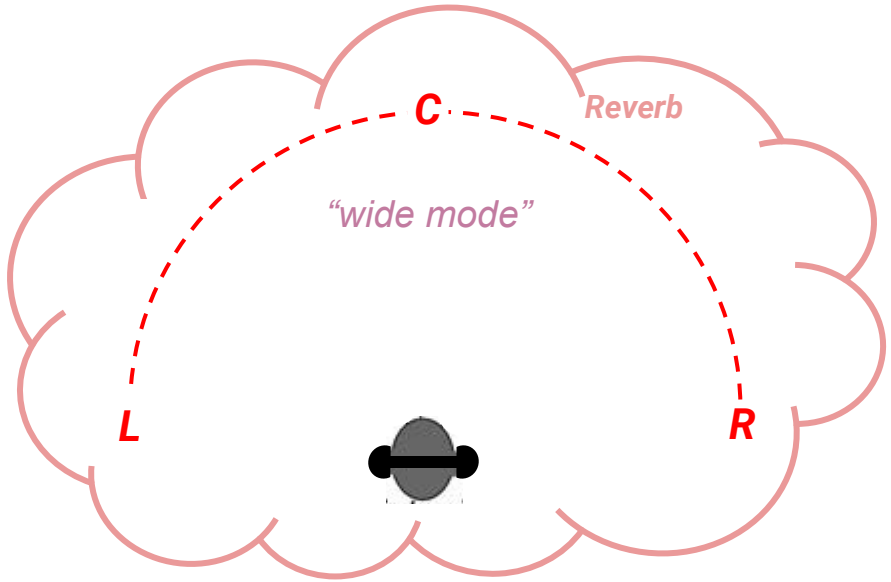
Aspirational binaural listening experience

Over loudspeakers





Over headphones

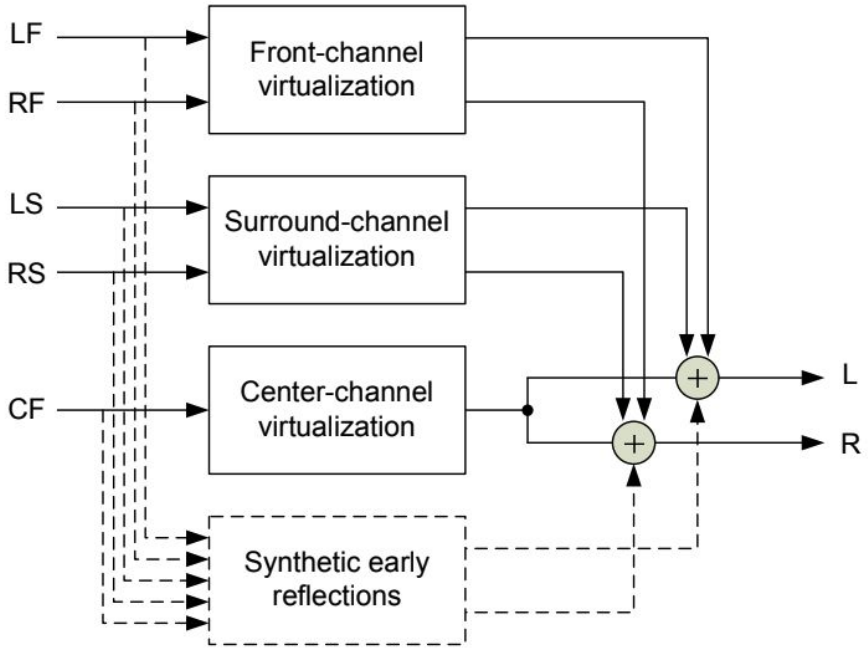


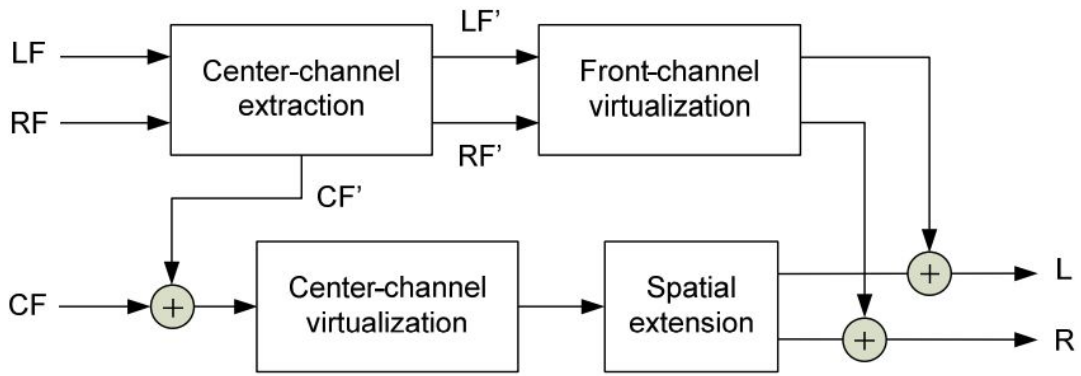
Sounds appear "externalized" and localized in front and sides

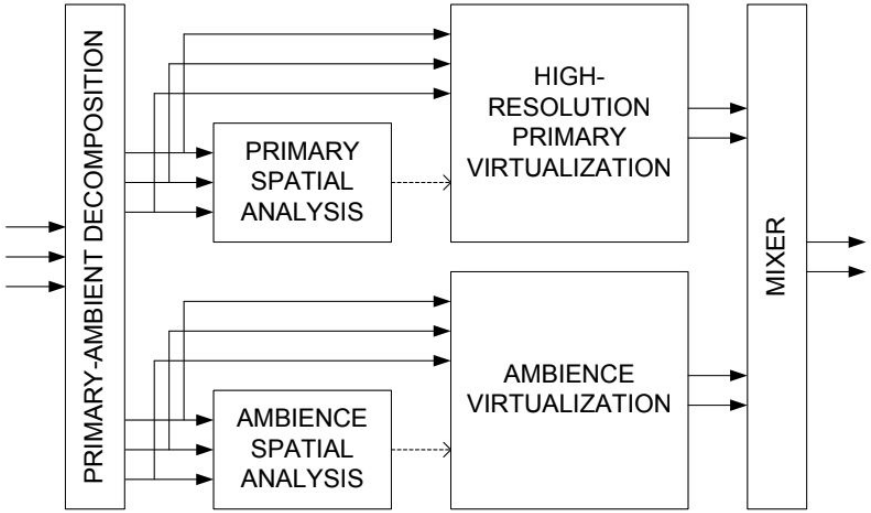
Review of approaches to binaural externalization processing

Method	Rationale / motivation	Observations	References
Virtual speakers processing	Restore natural binaural cues conveyed in stereo loudspeaker playback conditions.	In-head and/or elevated. Image width compromise.	Rubak (1991) Jot et al. (1995) Kirkeby (2002) Merimaa (2009)
Artificial reflections	Simulate a more natural listening experience.	Timbre coloration.	Jot & Avendano (2003) Davidson et al. (2016)
Audio AR reverberation	Simulate the local room's reverb properties for virtual/real cognitive congruence.	Requires reverb fingerprint detection technology.	Jot & Lee (2017) Murgai et al. (2018)
Direct-diffuse decomposition	Enhance the spatial discrimination of direct vs. diffuse components in the source material.	Frequency-domain processing complexity. Risk of artifacts.	Jot & Avendano (2003)
Up-mixing techniques	Mitigate the incorrect matching of natural HRTF cues for panned ("phantom") sources	Frequency-domain processing complexity. Risk of artifacts.	Goodwin & Jot (2007) Breebaart & Schuijers (2008)
Decorrelation techniques	Mitigate localization and timbre preservation issues for center-panned sound components	Frontal externalization of phantom center sounds.	Jot & Walsh (2010)

Review of approaches to binaural externalization processing







A new approach to binaural externalization processing

Some key remaining challenges:

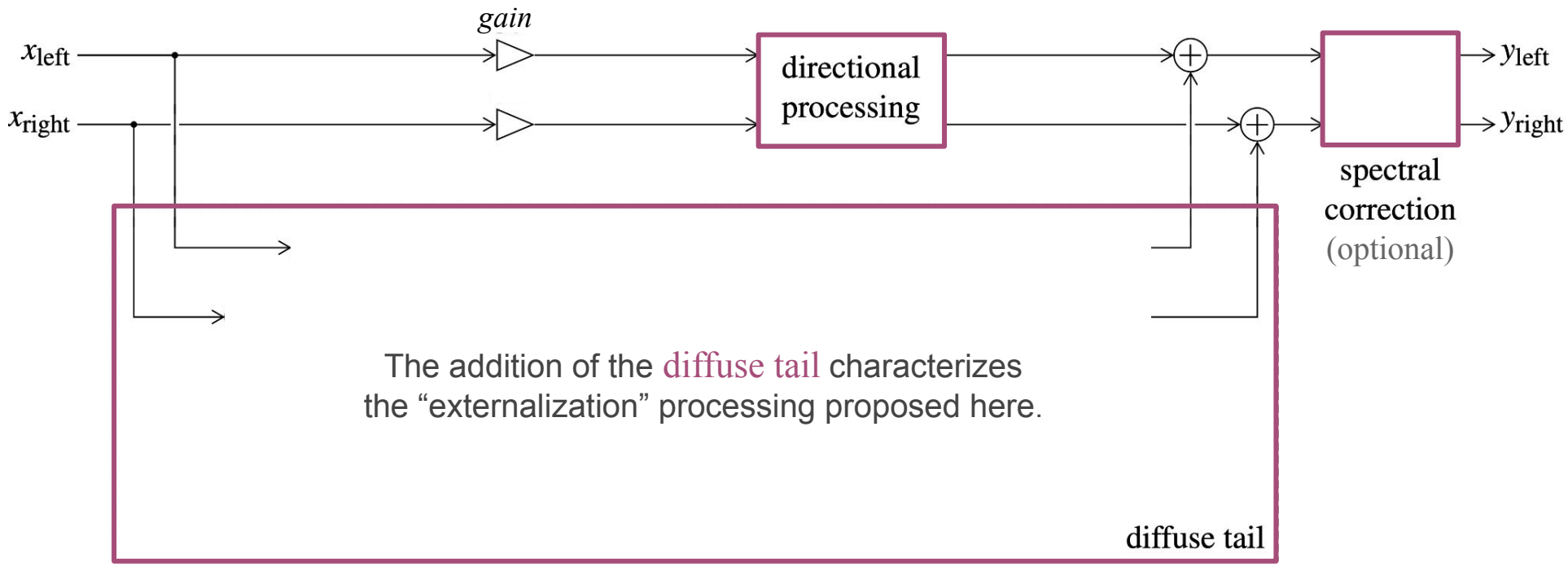
- preservation of the perceived overall width and timbre of the original recording
- frontal localization and timbre preservation of center-panned sound elements.

Demonstration #1:

Katrina Marie - "*Clocktower*" (excerpt)

1. *Externalized* (with spectral correction)
2. Original
3. *Externalized* (without spectral correction)
4. Directional processing only

A new approach to binaural externalization processing



A new approach to binaural externalization processing

Some key remaining challenges:

- preservation of the perceived overall width and timbre of the original recording
- frontal localization and timbre preservation of center-panned sound elements.

Demonstration #1:

Katrina Marie - "*Clocktower*" (excerpt)

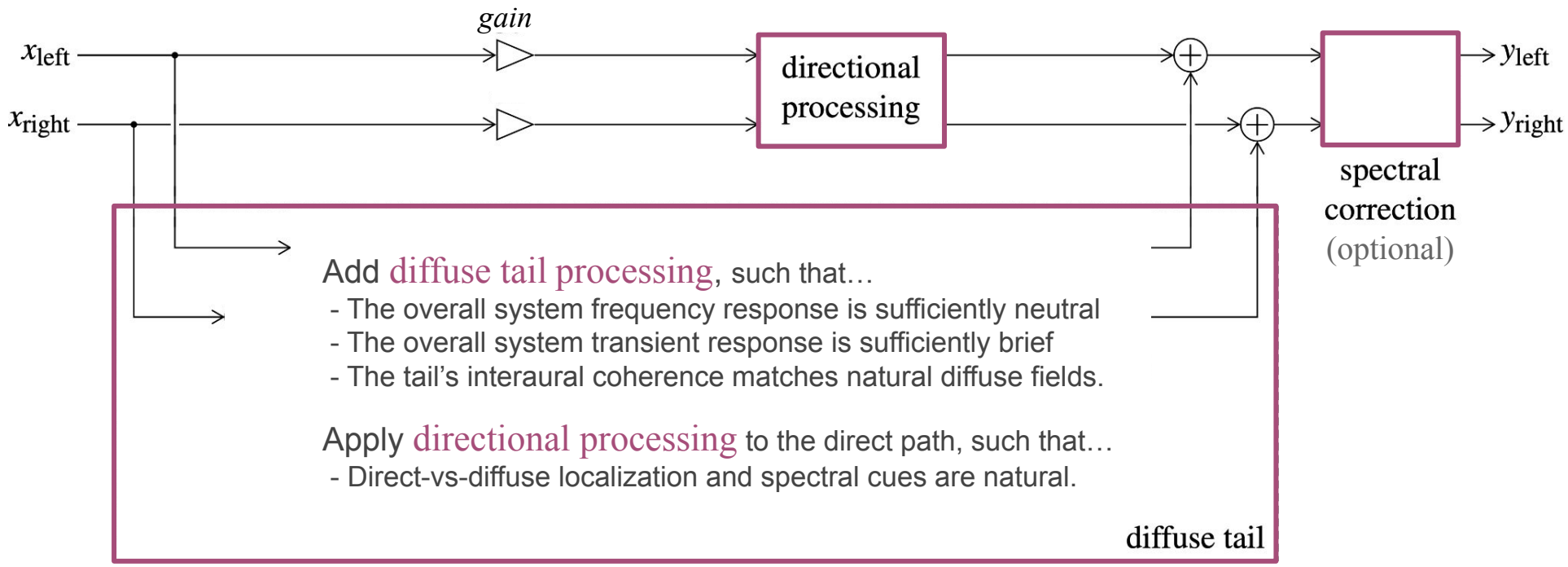
1. *Externalized* (with spectral correction)
2. Original
3. *Externalized* (without spectral correction)
4. Directional processing only

Demonstration #2:

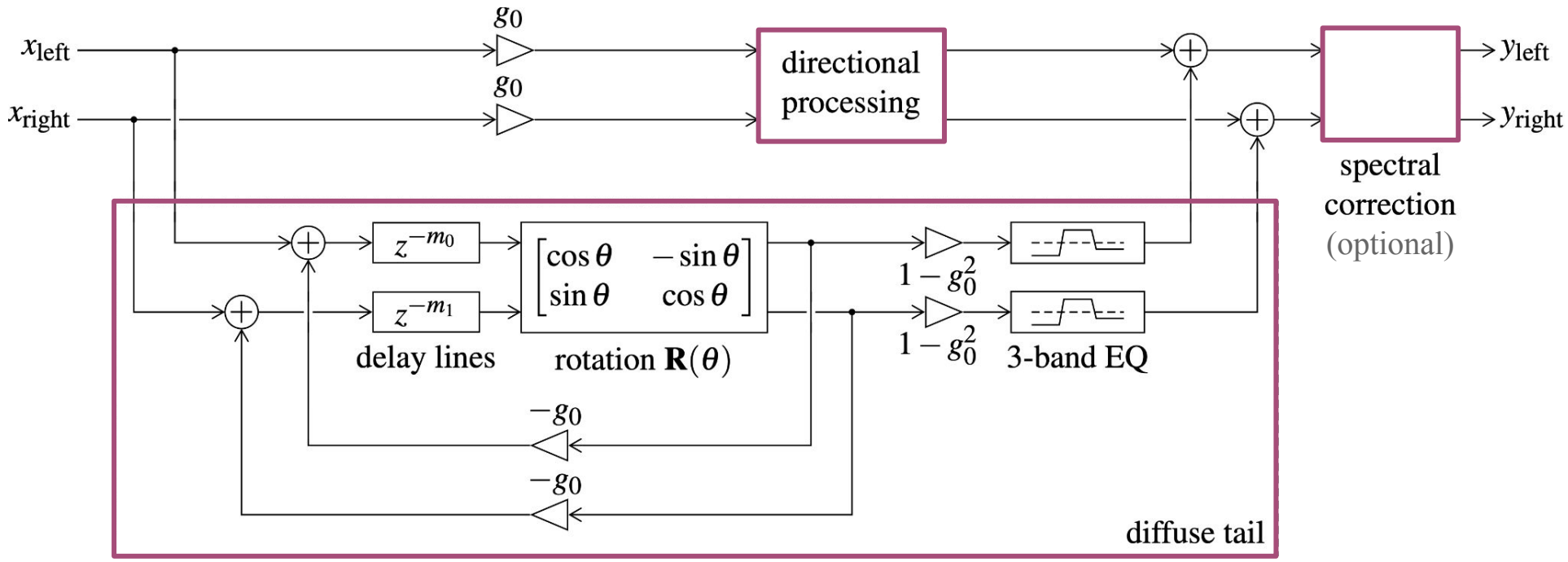
Caleb Hawley - *"Tell Me What It's Like To Have a Dream Come True"* (excerpt)

1. *Externalized* (with spectral correction)
2. Original
3. *Externalized* (without spectral correction)
4. Directional processing only

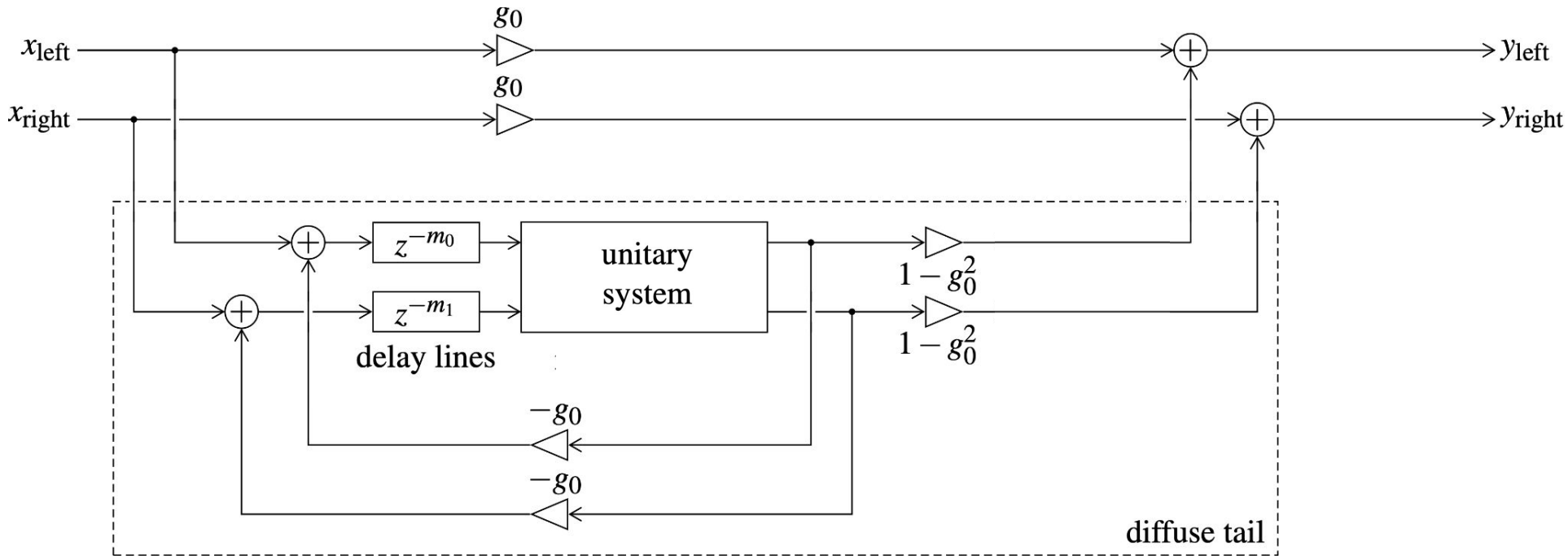
A new approach to binaural externalization processing



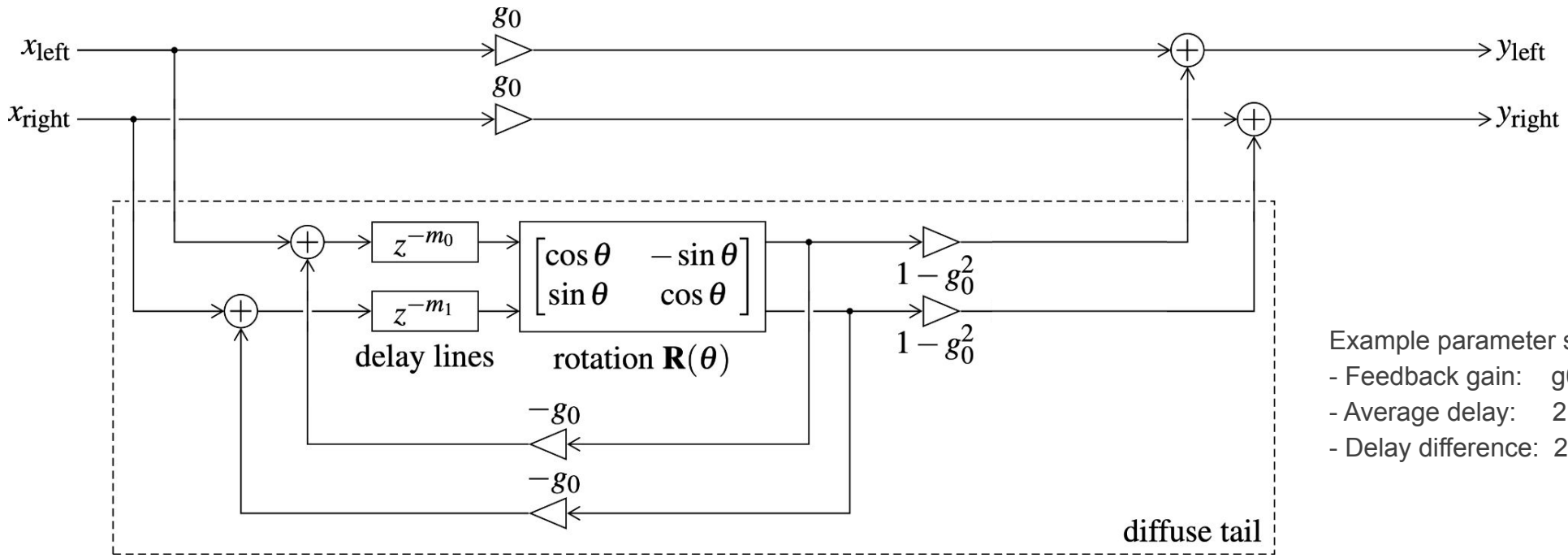
Example realization: modifying a 2-channel all-pass IIR filter



General 2-channel all-pass IIR filter topology



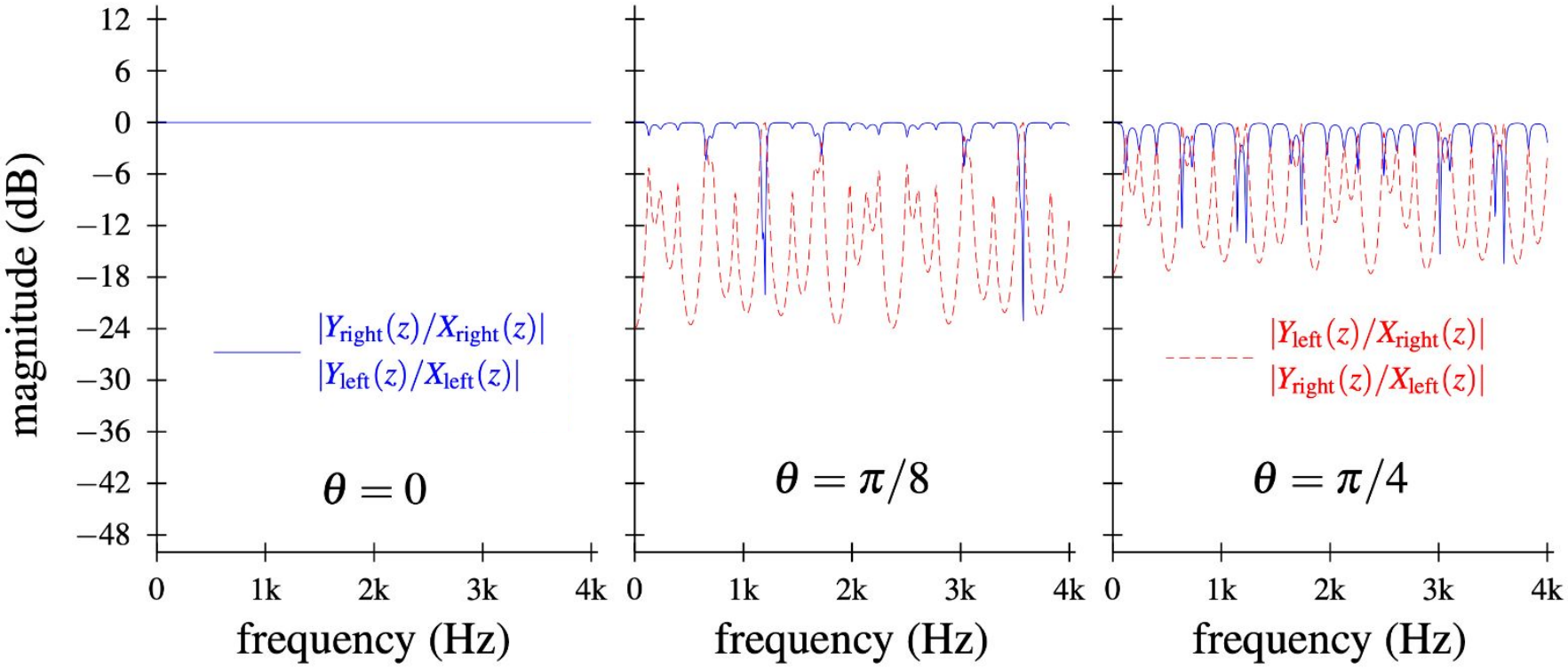
Simple 2-channel all-pass IIR filter



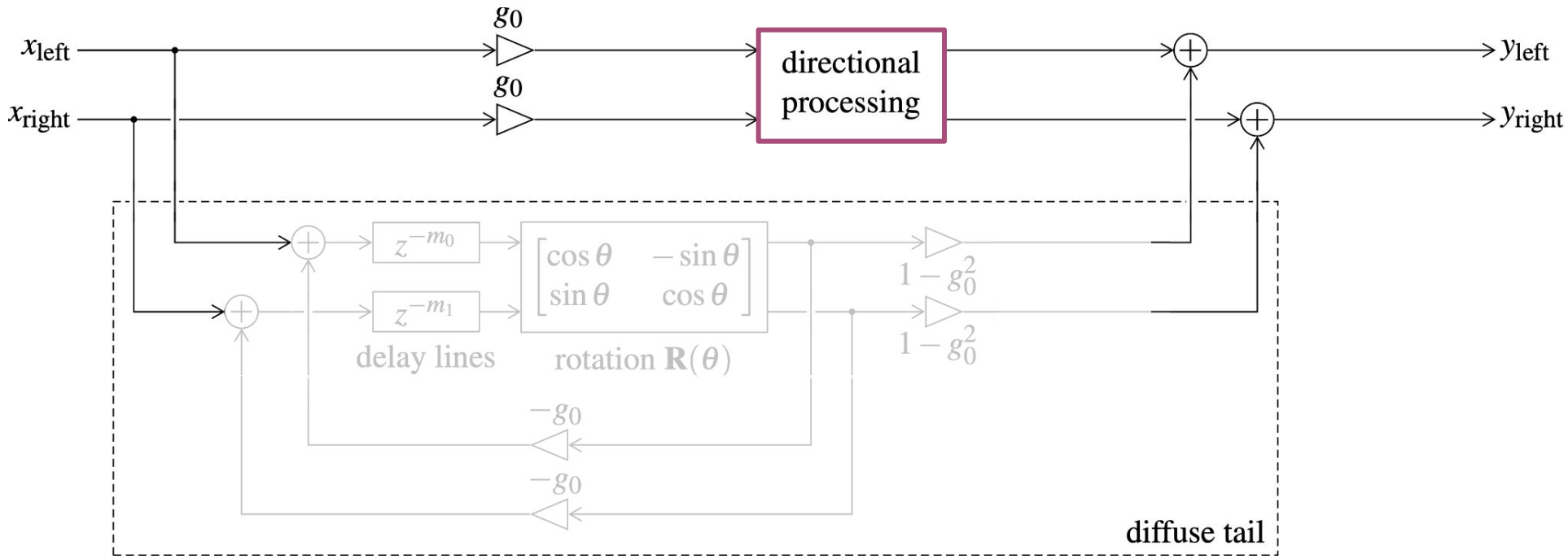
Example parameter settings:
- Feedback gain: g0 = 0.72
- Average delay: 2.943 ms
- Delay difference: 28.74 %

Simple 2-channel all-pass IIR filter

System frequency response for three settings of the rotation angle

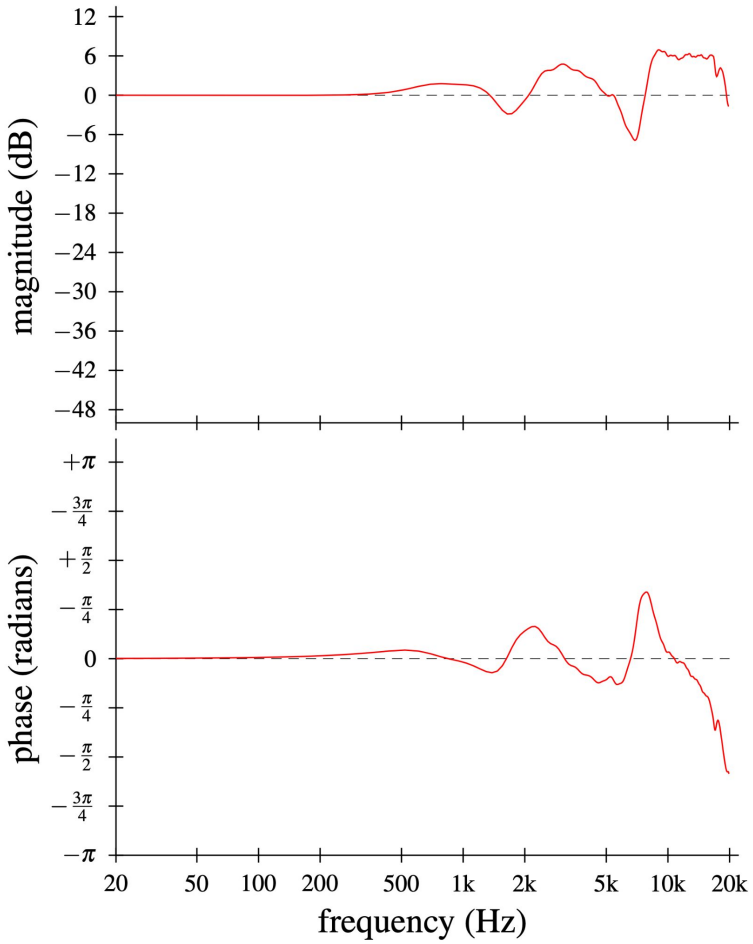


Inserting a directional processing filter



Inserting a directional processing filter

Example of directional processing filter



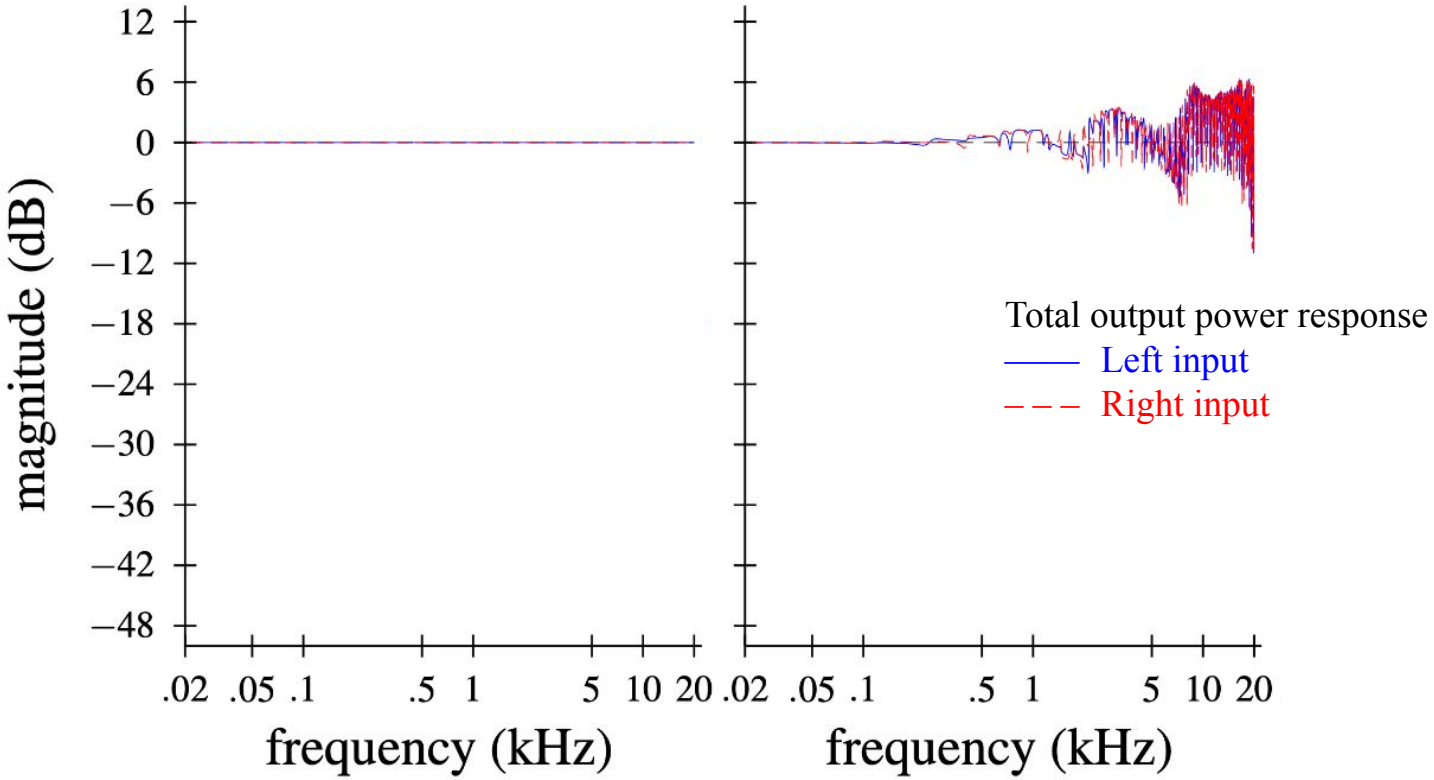
Head-Related Transfer Function (HRTF) measured on Neumann KU100 dummy head at front direction (0 deg azimuth, 0 deg elevation). Source: SADIE II database (York Univ.)

Correction of magnitude frequency response:
- diffuse-field compensation.
- neutralization at low frequencies.

Conversion to a minimum phase filter.

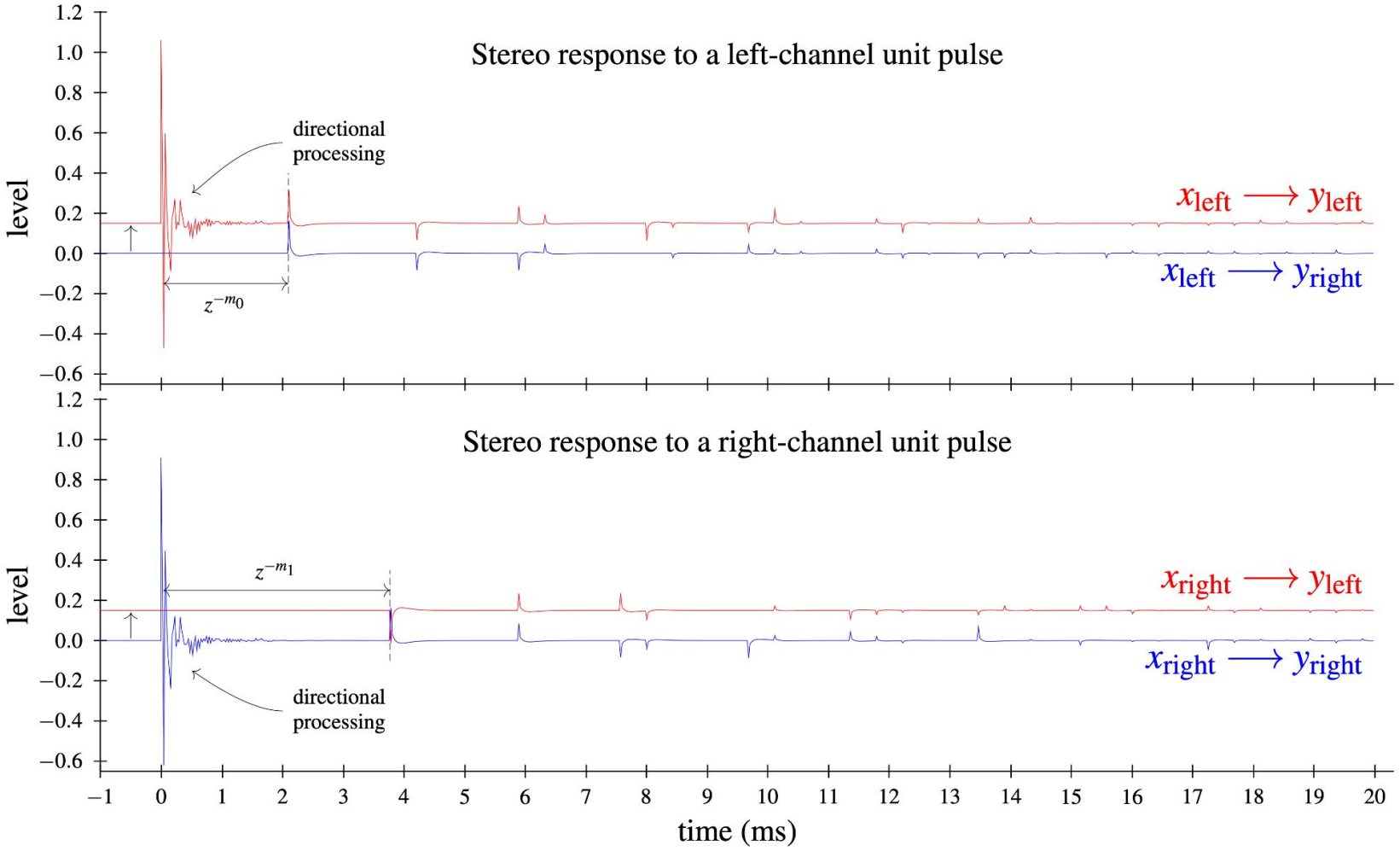
Inserting a directional processing filter

Modified system response

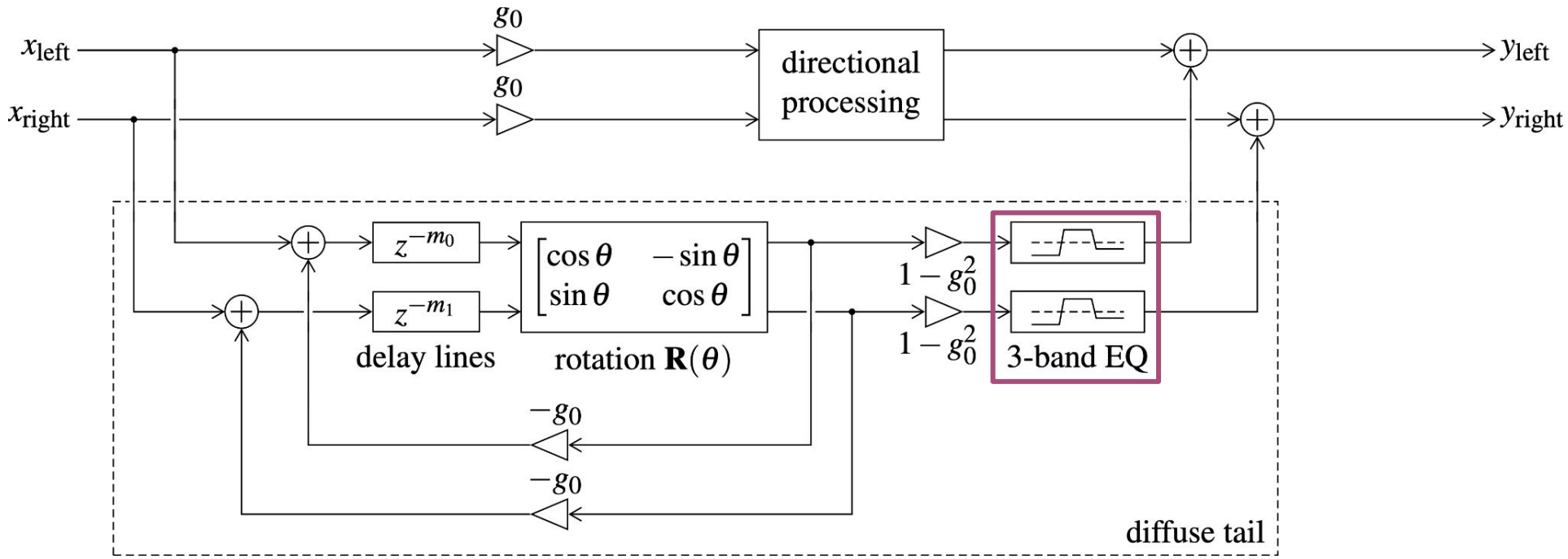


Inserting a directional processing filter

Modified system response

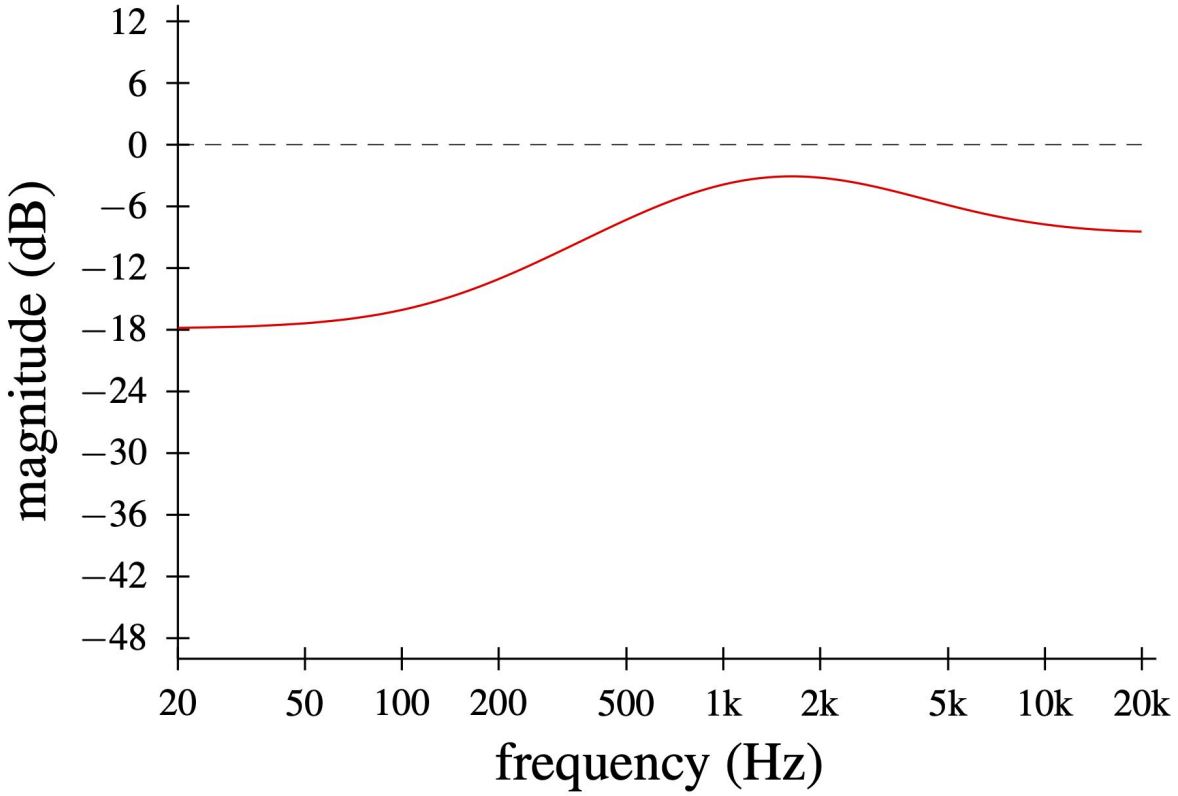


Inserting a Tail EQ filter

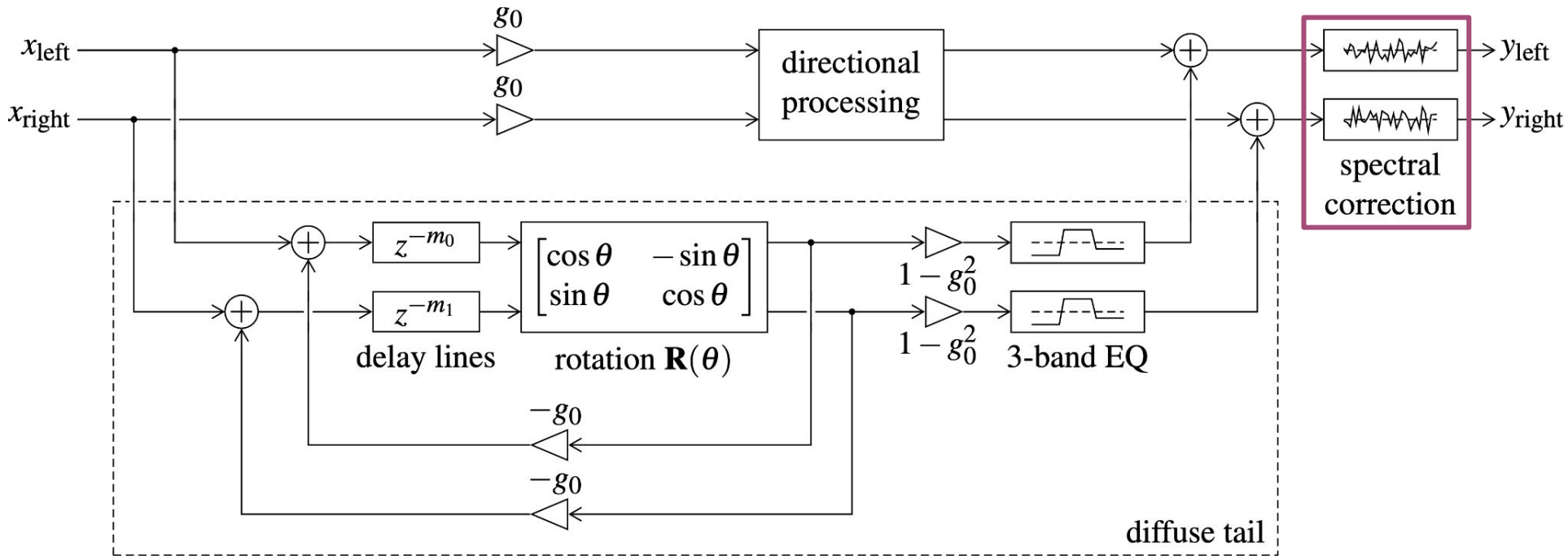


Inserting a Tail EQ filter

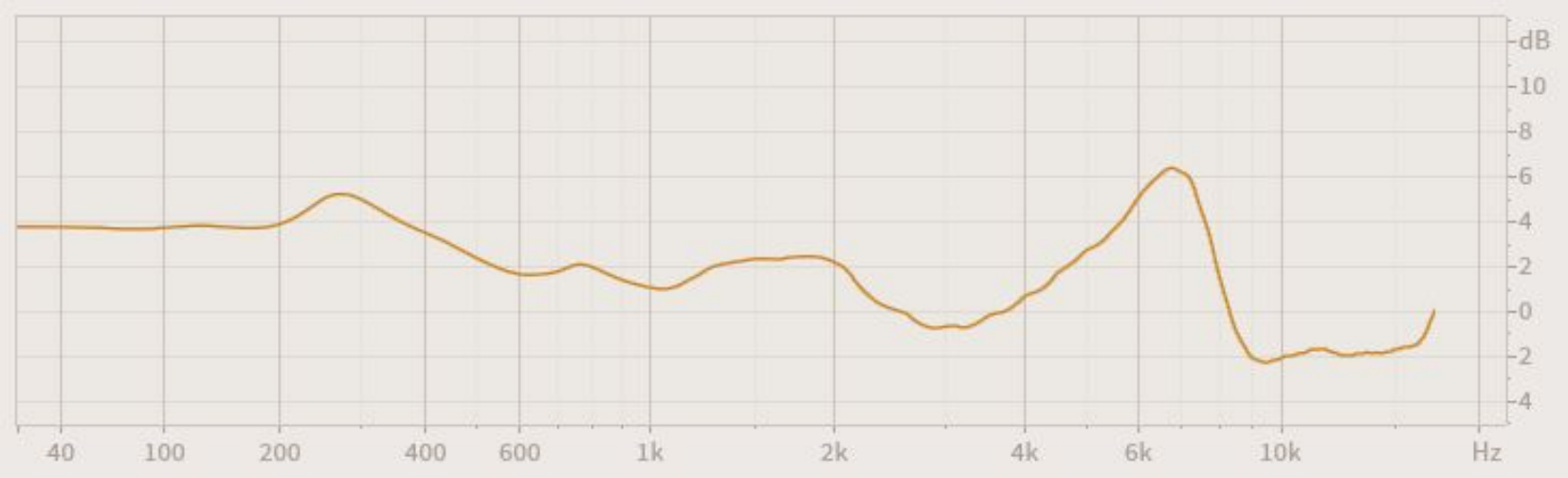
3-band EQ example



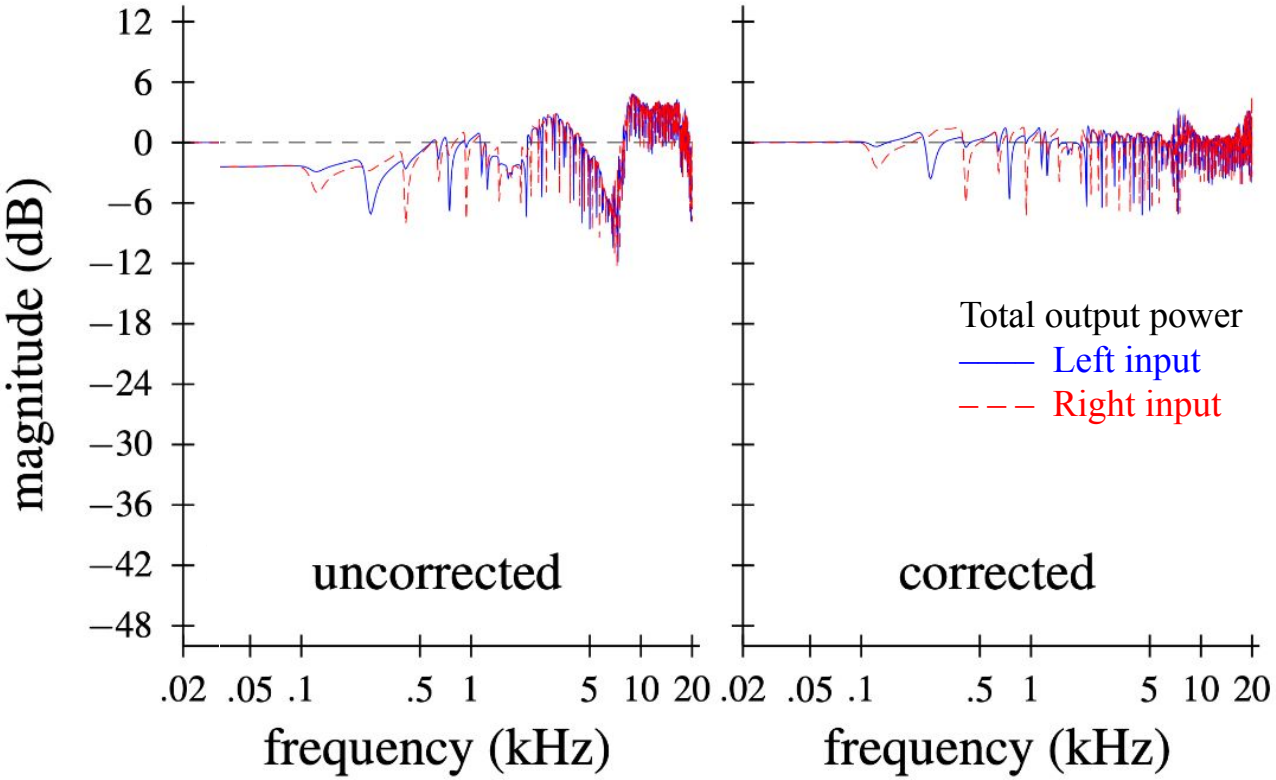
Overall spectral correction



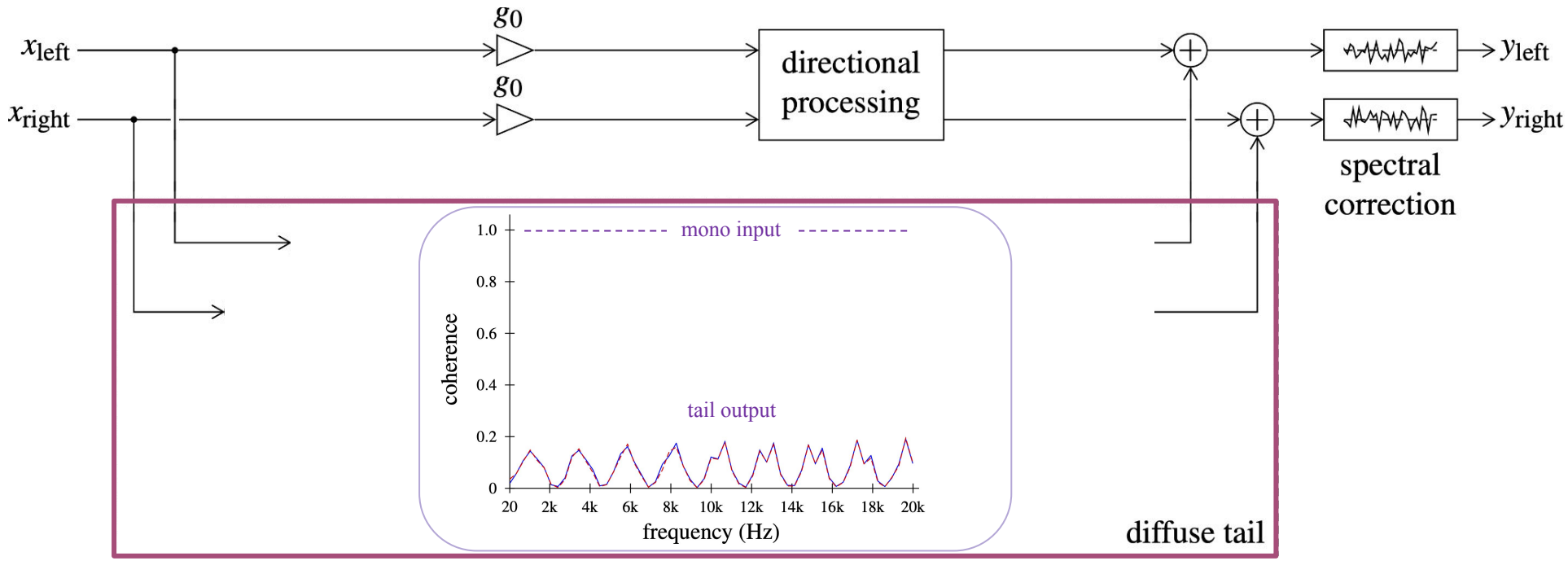
Overall spectral correction



Overall spectral correction

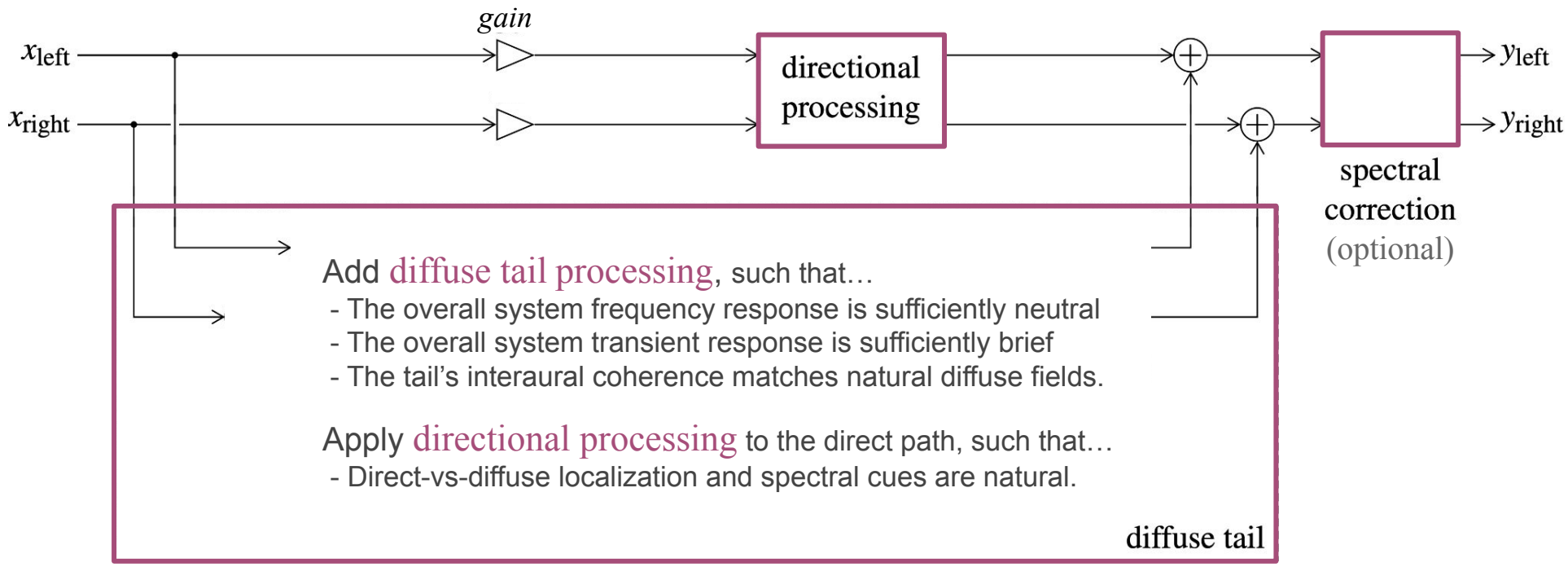


Inter-channel coherence in tail response



A new approach to binaural externalization processing

Recap:



A new approach to binaural externalization processing

Some possible extensions of this work...

Perceptual investigations: spatial audio perception of concurrent free-field & diffuse-field components

- Psycho-acoustical studies to verify the viability of the proposed externalization processing method.
  - Does it exploit a human evolutionary capability borne from listening experiences in reverberant environments?
- Can we establish objective “triggers” to externalization that are robust/consistent across listening conditions?
  - Discrimination of direct vs. diffuse components (coherence + spectral). Individual-dependent HRTF processing?.
- To which extent can we de-emphasize room-coloration cues without undermining perceived externalization?

Algorithm design extensions:

- Expose a set of intuitive knobs for fine-tuning and customization of this spatial audio effect.
- More complex all-pass IIR network designs (nested all-pass filters …), time-varying all-pass networks (Werner, 2020).
- Alternative approach based on Velvet Noise Decorrelators (Alary et al., 2020).
- Mono compatibility.

Applications to the production of 2-channel recordings:

- Production of stereo recordings that offer enhanced headphone playback and remain compatible with loudspeaker playback.
- Applying externalization processing selectively to some of the tracks or stems in a stereo mix.
- Producing binaural recordings that employ familiar 2-channel stereo production effects and techniques.
- Improving externalization in binaural panning/rendering methods, especially in the frontal localization sector.

Binaural Audio Externalization Processing

Jean-Marc Jot, Alexey Lukin, Kurt James Werner, Evan Allen




Audio examples are posted at izotope.com/tech/aes_extern .