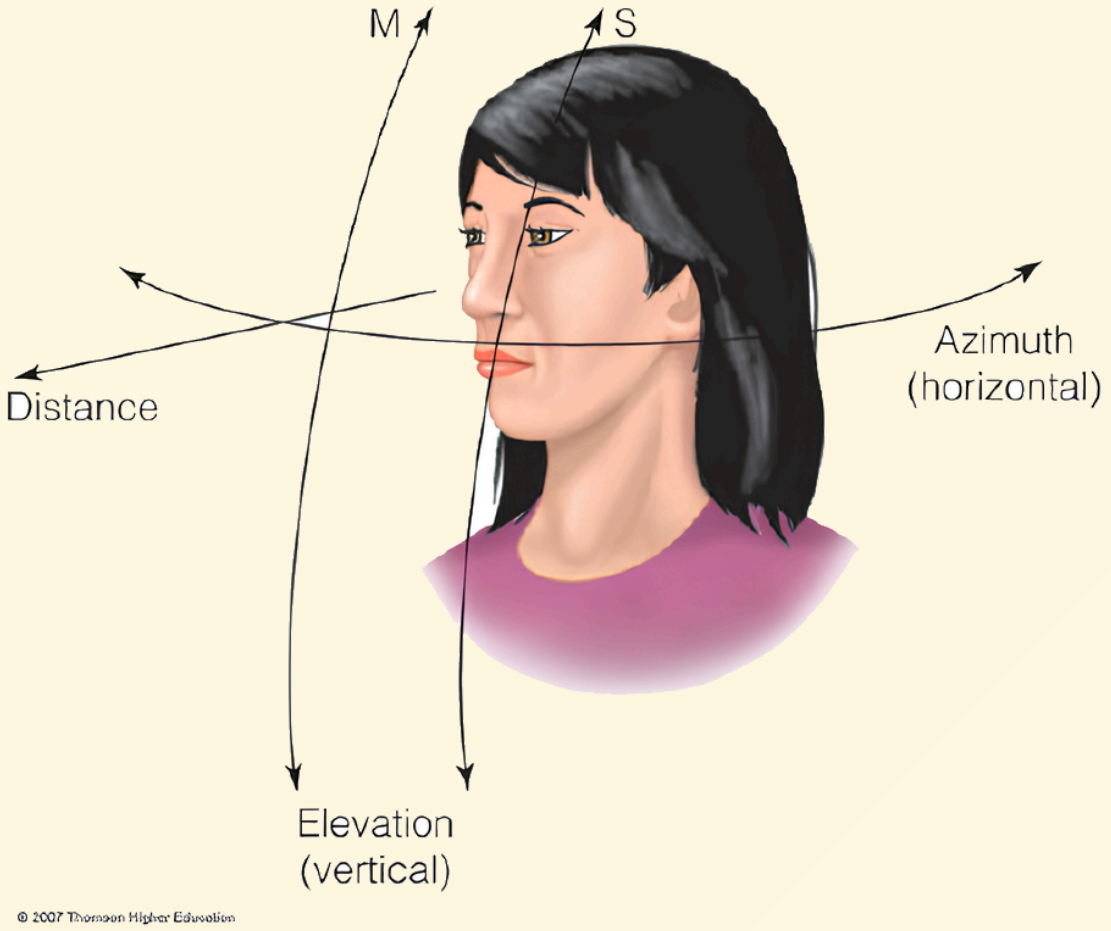
Sound Localization Update

author: Paolo Marzolo
advisor: Alberto Antonietti
co-adv: Francesco De Santis

Table of contents

Today's objectives

1. show current state
2. foster discussion on where to go from here
3. practice!



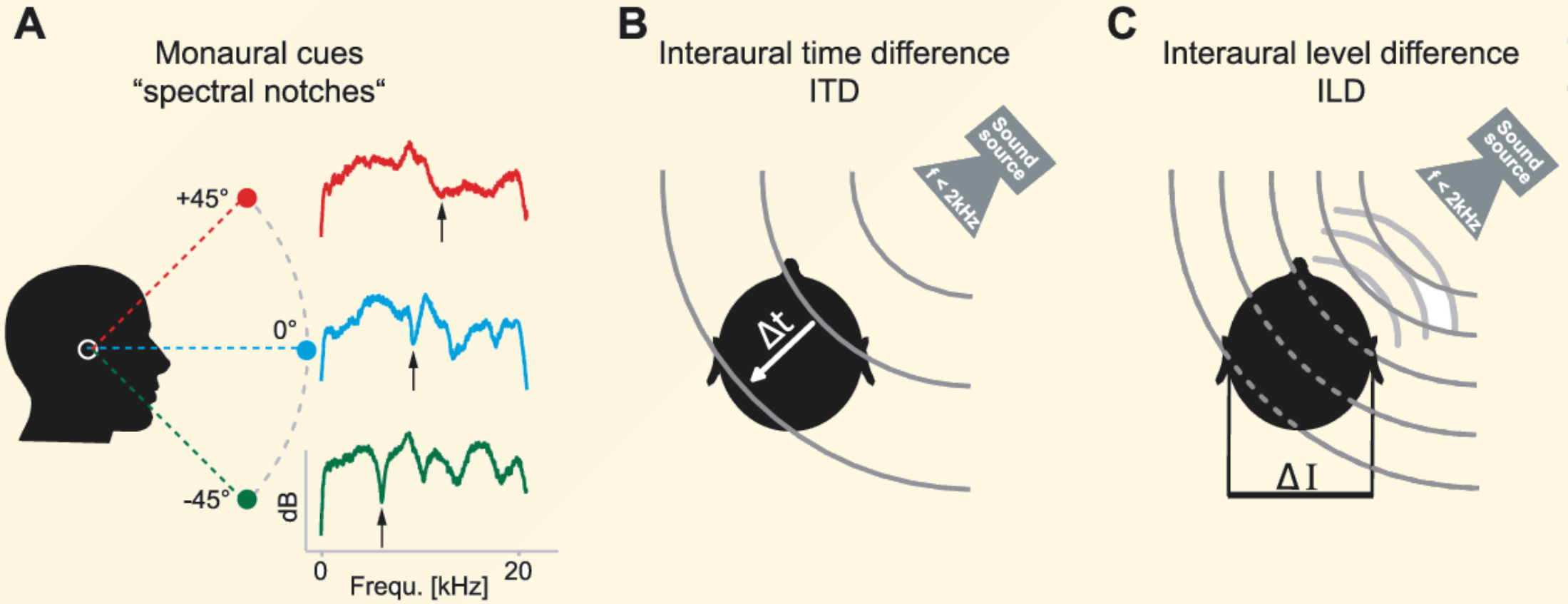
Task definition

> “ The ability to identify the location of a sound source in a sound field. ”
>
> (Jutras et al., 2020)

Why it's interesting

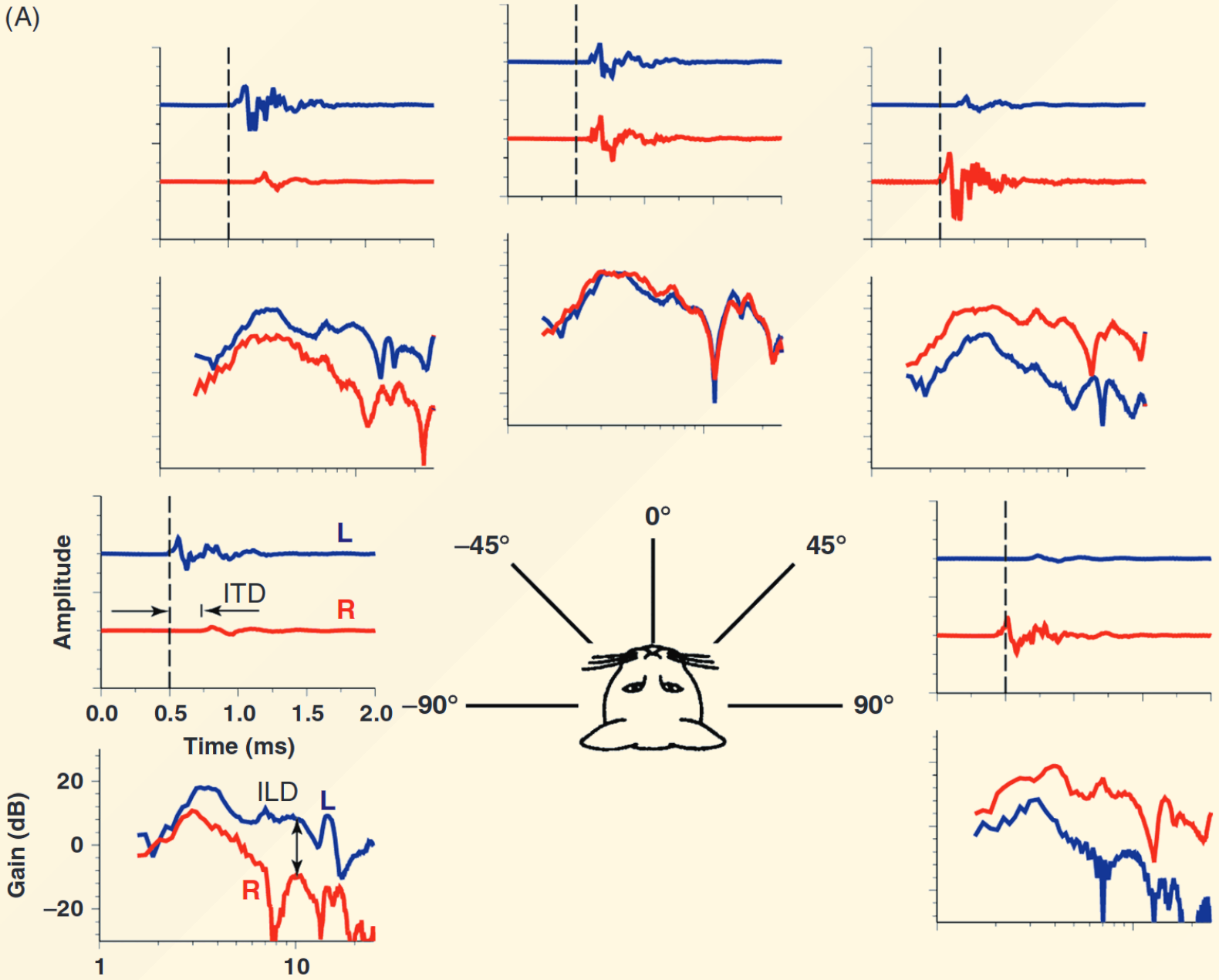
1. point-like sensor to spatial
2. localizing is instrumental to selectively improve SNR
3. features special neural features

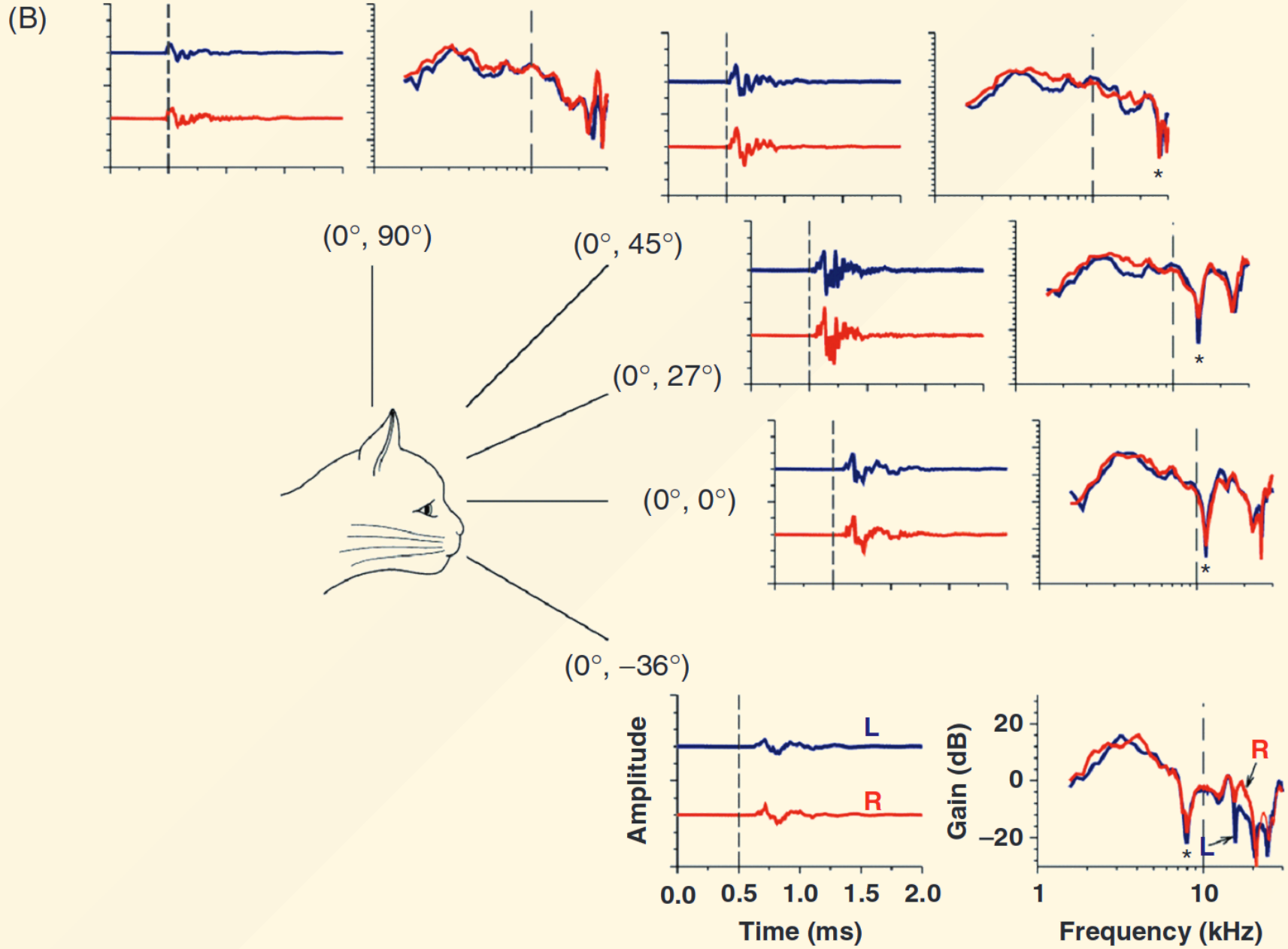
Hearing cues



(Grothe, 2010)

HRTFs

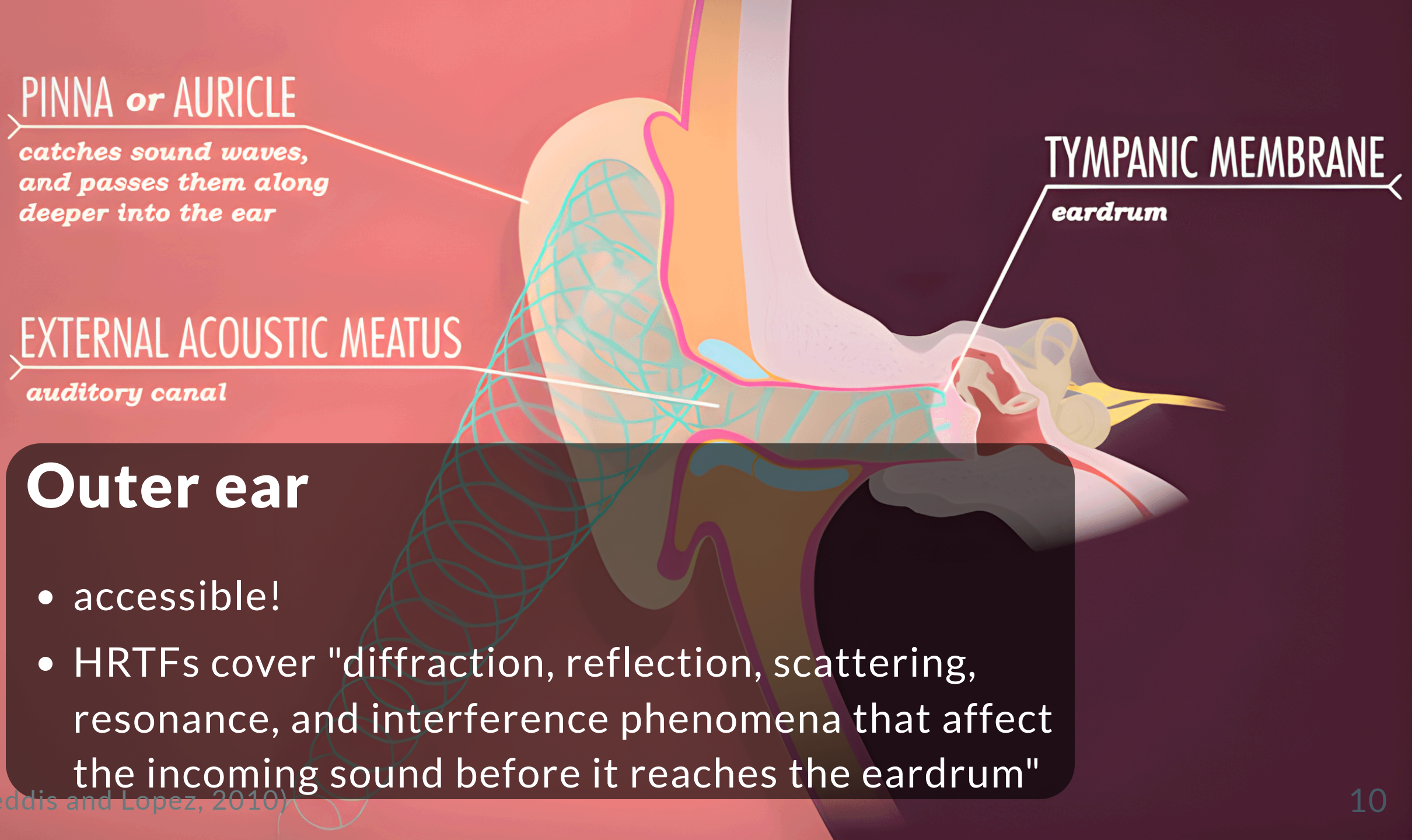

(A)
L
ITD
R
Amplitude
0.0
0.5
1.0
1.5
2.0
Time (ms)
ILD
L
R
20
0
–20
Gain (dB)
1
10
0°
–45°
45°
–90°
90°


(B)
(0°, 90°)
(0°, 45°)
(0°, 27°)
(0°, 0°)
(0°, –36°)
L
R
Amplitude
0.0
0.5
1.0
1.5
2.0
Time (ms)
20
0
–20
Gain (dB)
1
10
Frequency (kHz)
*

Path to auditory nerve

1. Outer ear
2. Middle ear
3. Basilar membrane
4. Inner Hair Cells
5. Active components (Outer Hair Cells)
6. Auditory nerve synapse

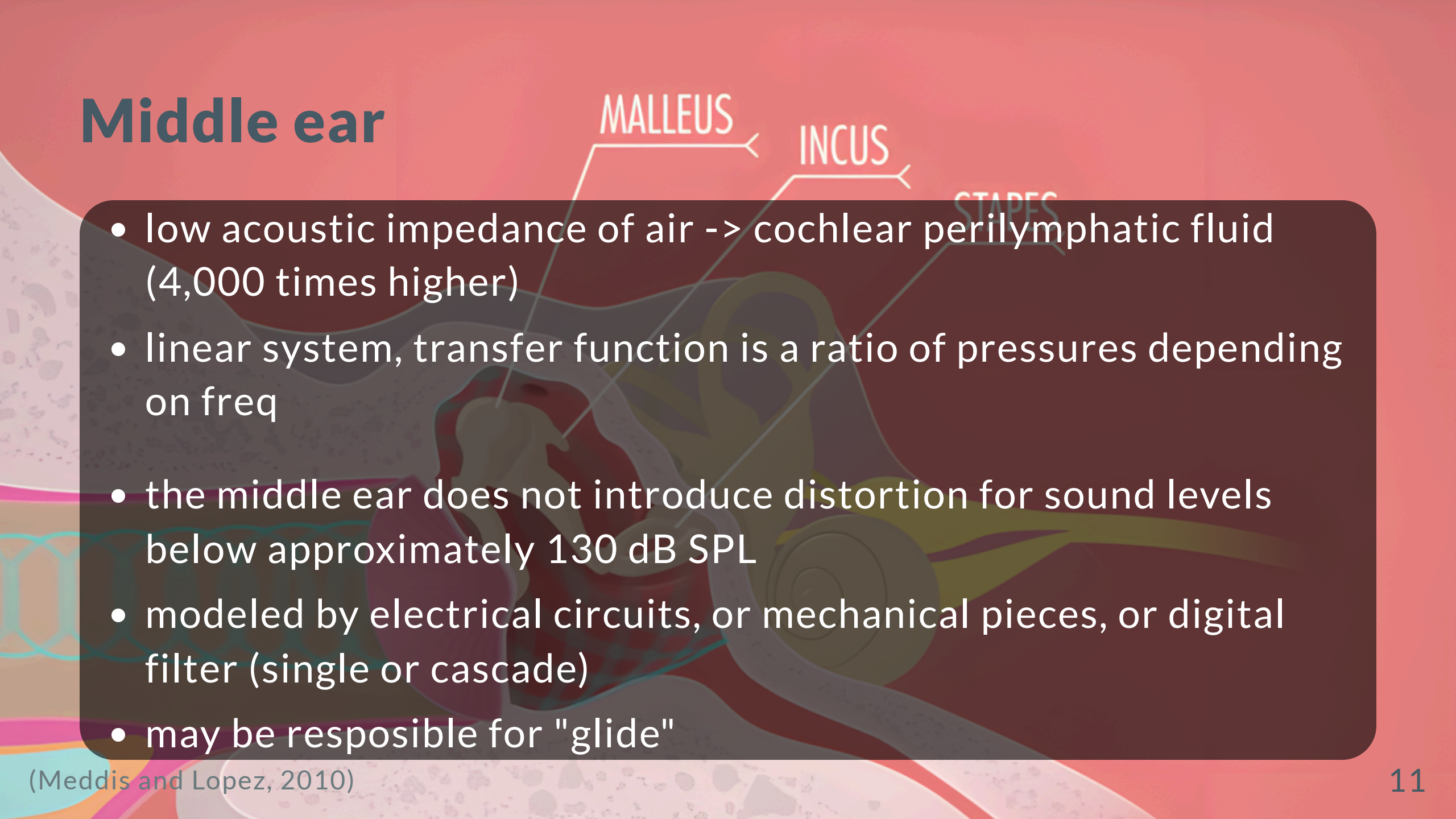
(Meddis and Lopez, 2010)



Outer ear

- accessible!
- HRTFs cover "diffraction, reflection, scattering, resonance, and interference phenomena that affect the incoming sound before it reaches the eardrum" (Meddis and Lopez, 2010)

Middle ear

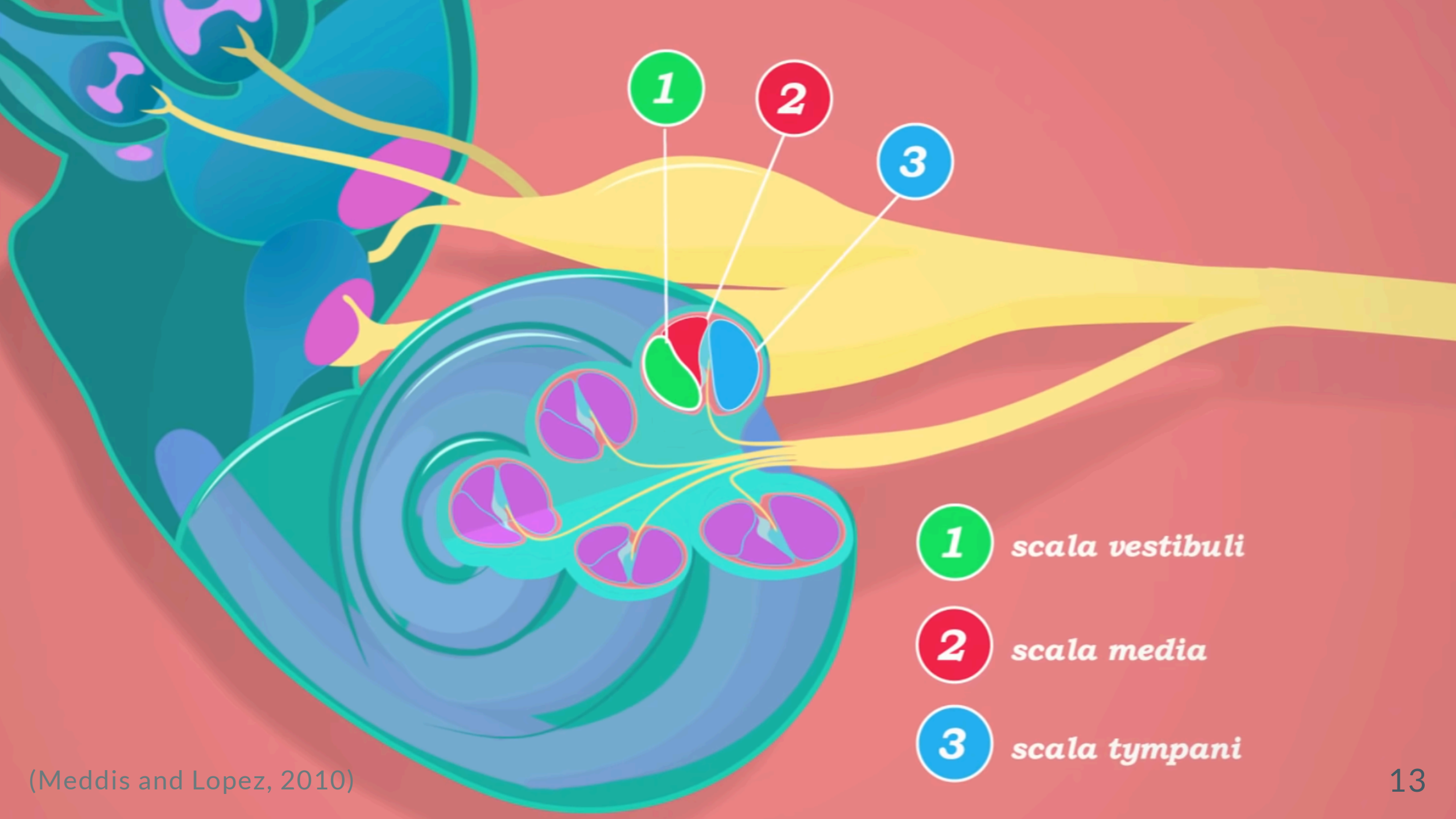


- low acoustic impedance of air -> cochlear perilymphatic fluid (4,000 times higher)
- linear system, transfer function is a ratio of pressures depending on freq
- the middle ear does not introduce distortion for sound levels below approximately 130 dB SPL
- modeled by electrical circuits, or mechanical pieces, or digital filter (single or cascade)
- may be resposible for "glide"

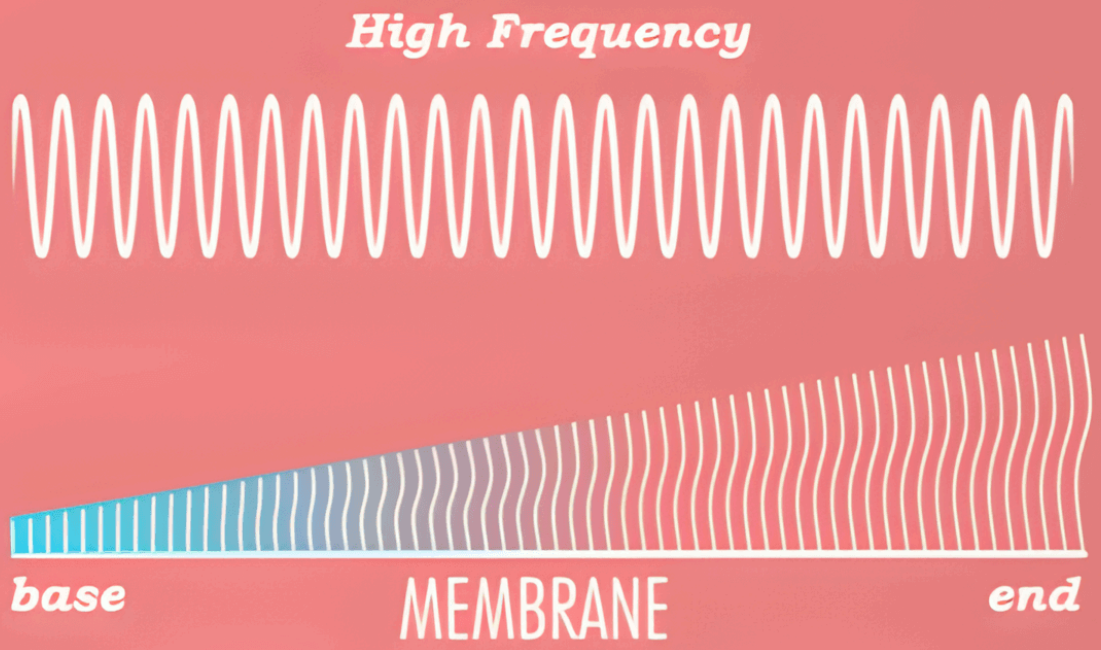
(Meddis and Lopez, 2010)

“ **the output signal must be multiplied by an appropriate scalar to achieve a realistic gain** „

(Meddis and Lopez, 2010)



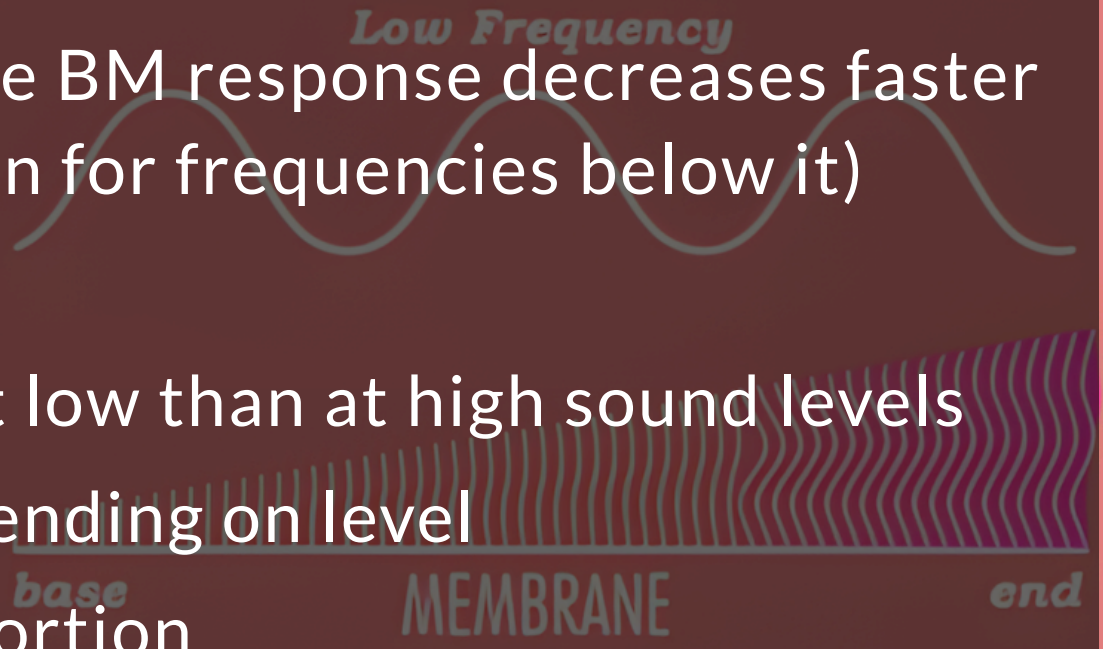
(Meddis and Lopez, 2010)



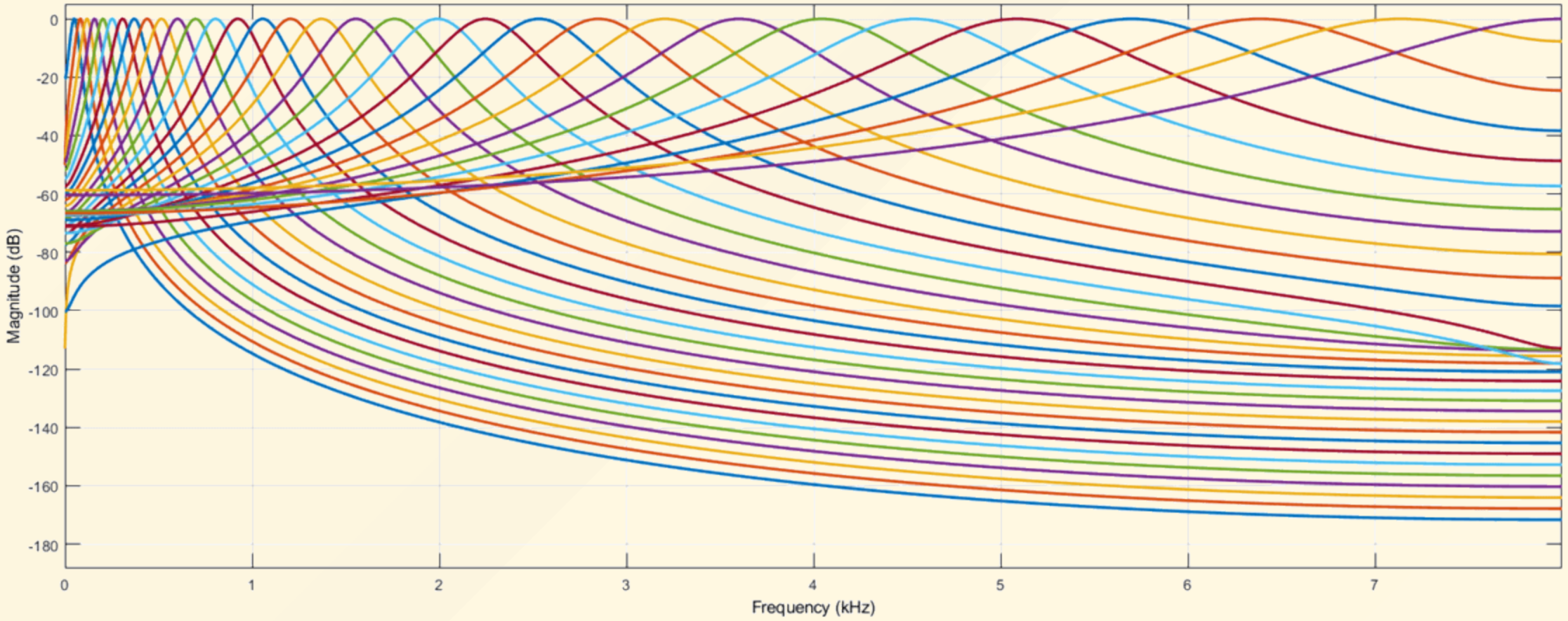
(Meddis and Lopez, 2010)

Basilar membrane

- asymmetric (the magnitude of the BM response decreases faster for frequencies above the BF than for frequencies below it)
- nonlinear!
- BM responses show more gain at low than at high sound levels
- BF of a cochlear site change depending on level
- also causes suppression and distortion
- responses are fully linear postmortem



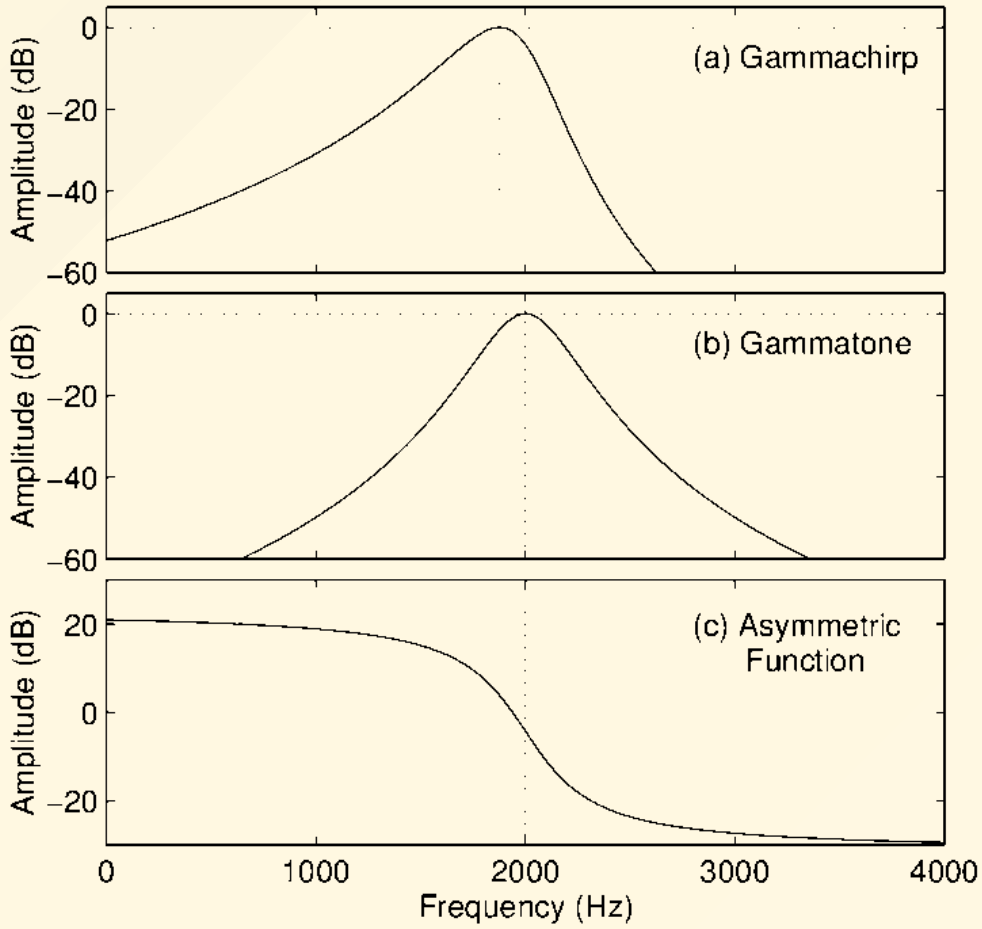
Modelling the BM



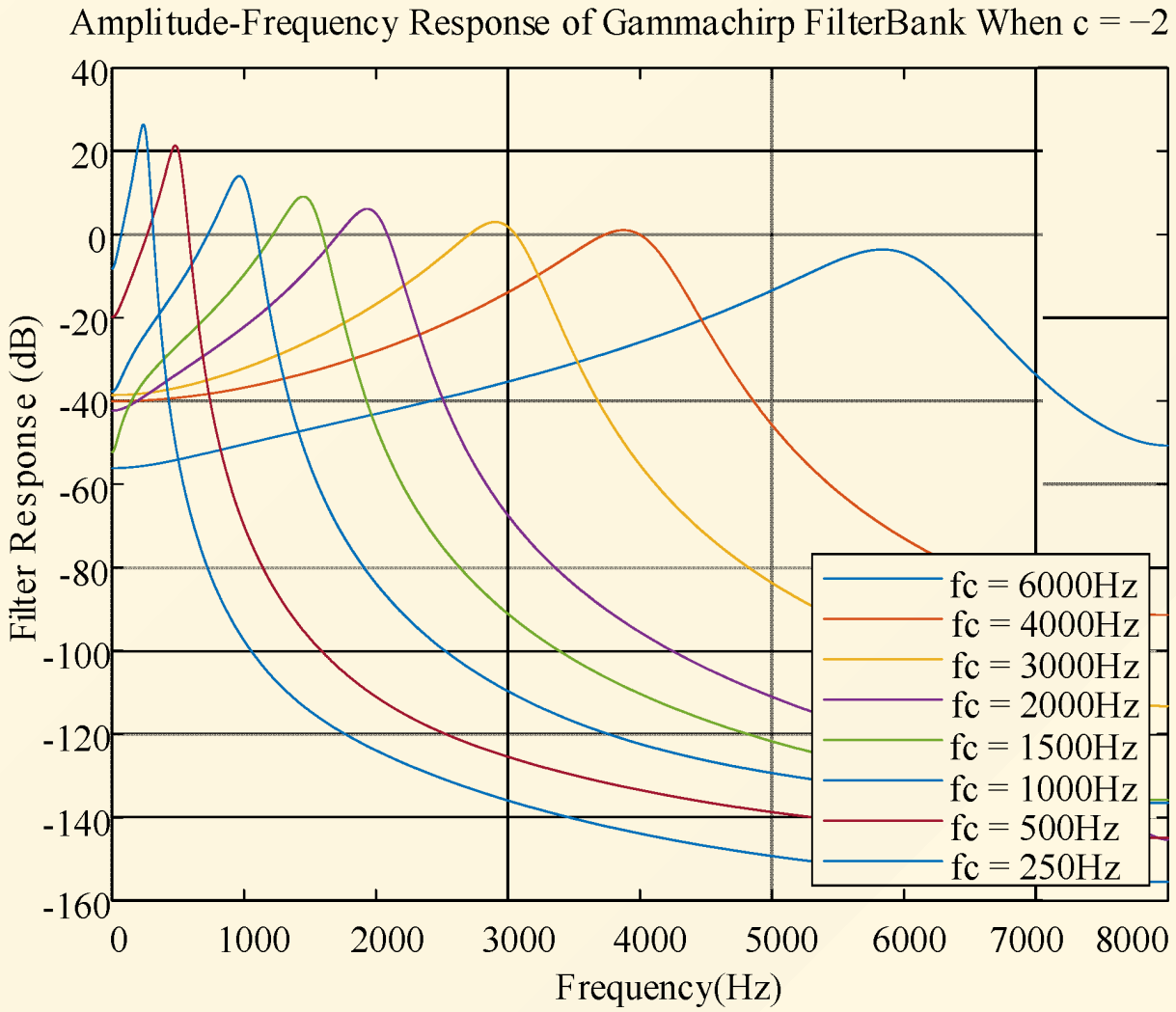
gammatone filterbank (linear, symmetric frequency response)

Gammatone issues

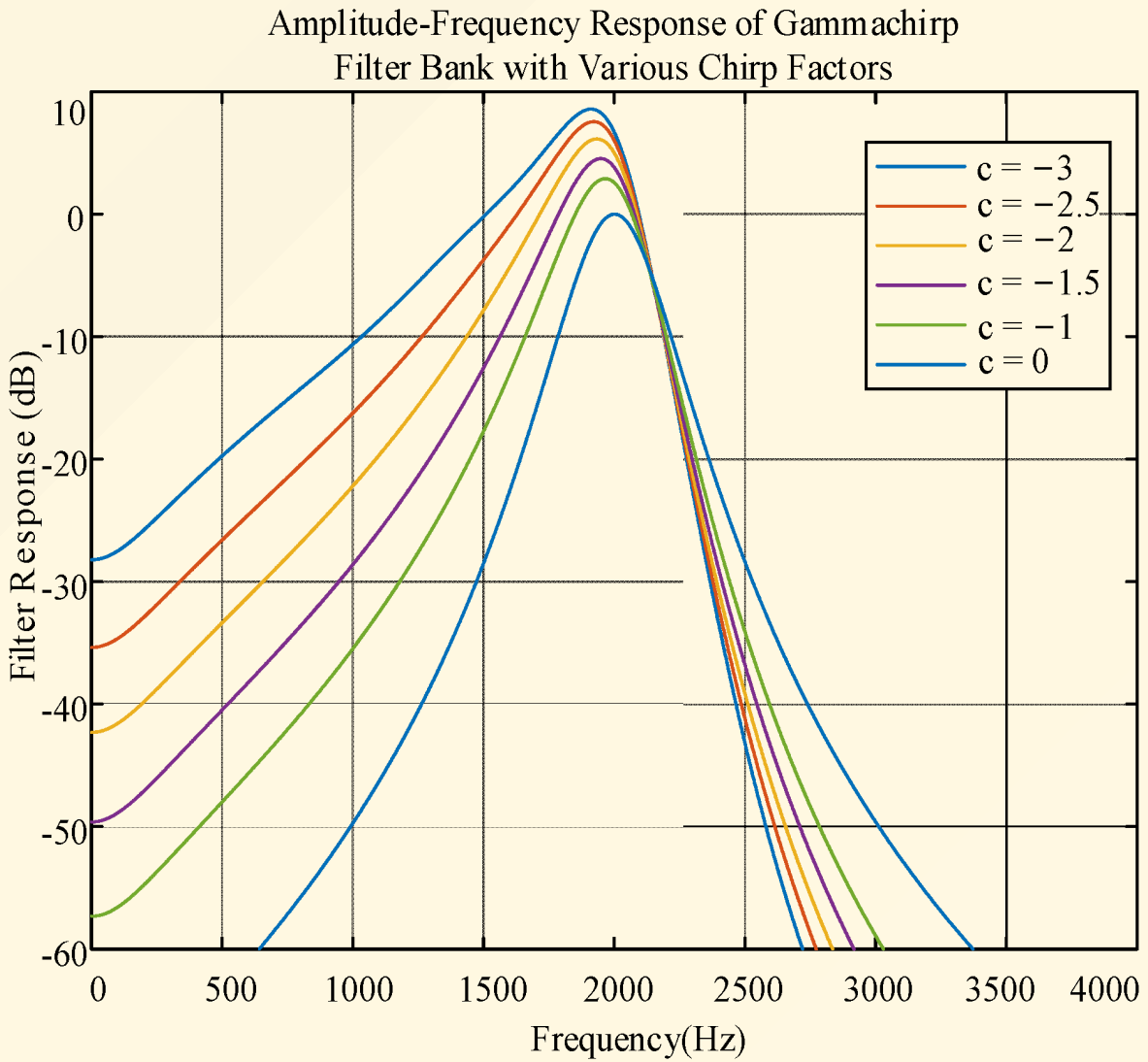
- linear (wrt level)
- symmetric



https://www.jstage.jst.go.jp/article/ast1980/20/6/2



(c)



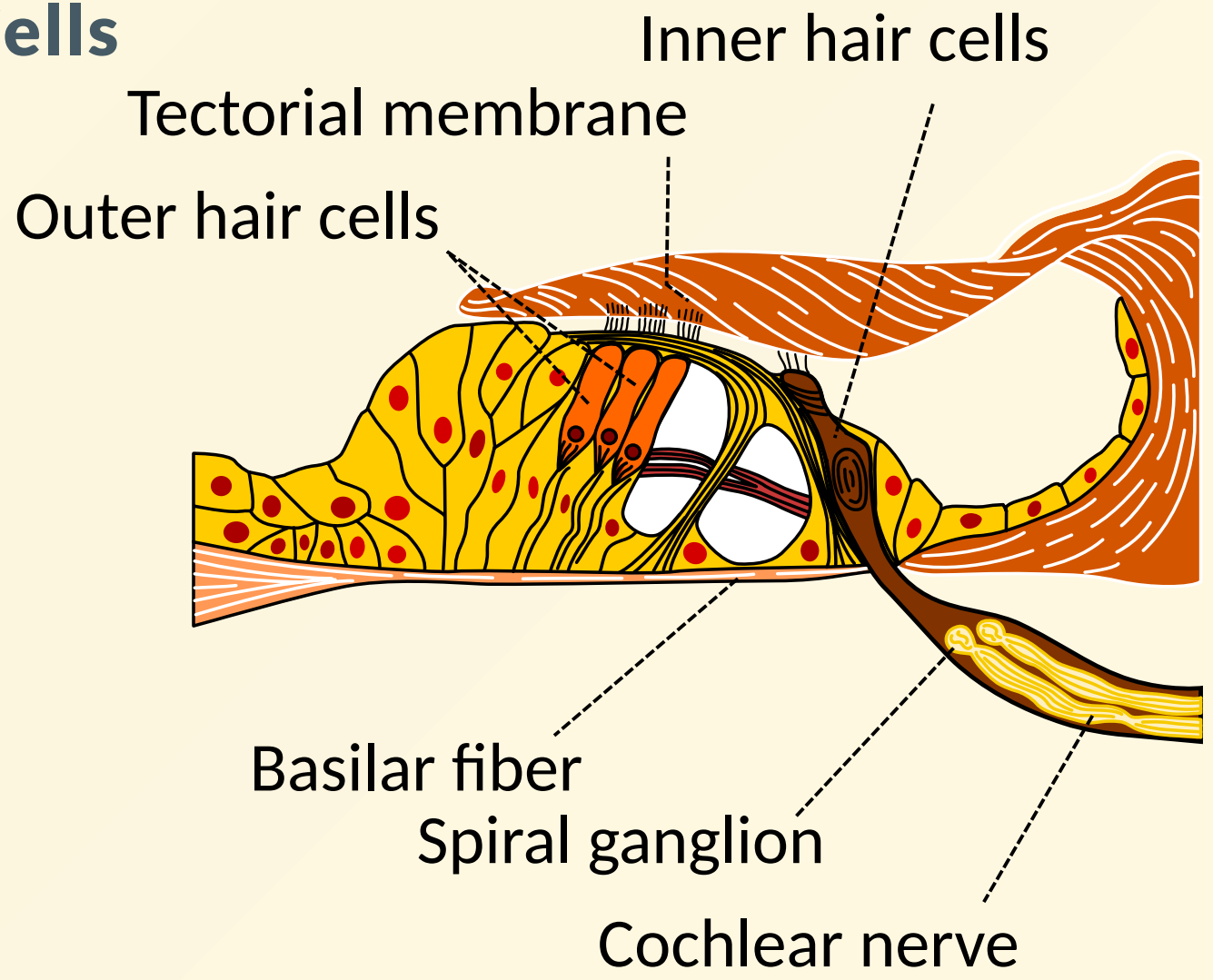
(d)

(Ruxue Guo et al, 2022)

and beyond!

- Tan-Carney model:
  - level independent chirps
  - gain and bandwidth vary dynamically depending on level
  - accounts for two-tone suppression
- DRNL (Dual Resonance Non Linear):
  - simulates the velocity of vibration of a given site on the BM
  - linear responses at low levels, compressive for moderate levels
  - predicts AN representation of stimuli with complex spectra, used to drive models of brain stem units, and as the basis for a speech processor for cochlear implants

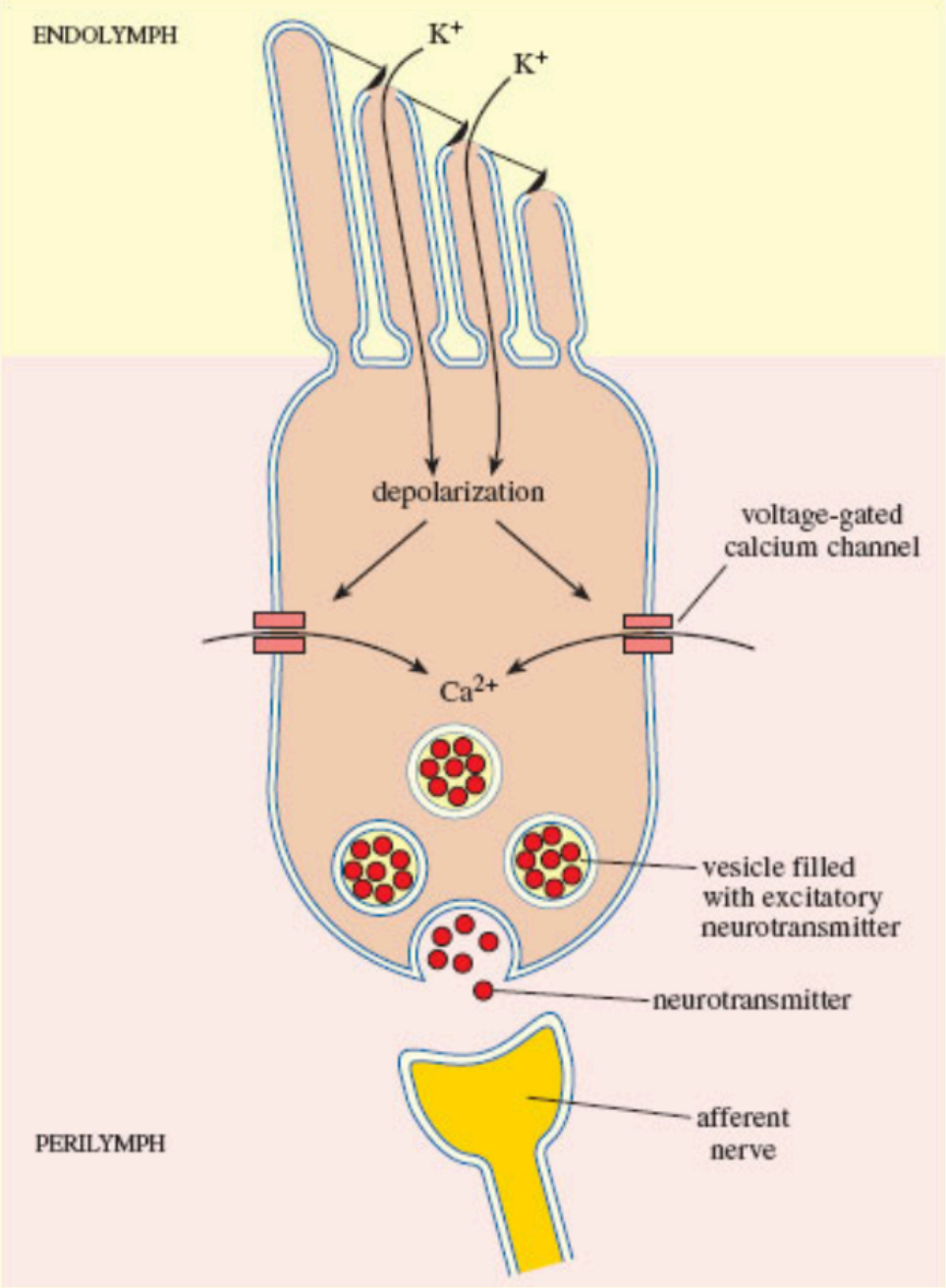
Inner Hair Cells



from IHC to ANF

- although inner hair cells depolarize, they do not produce an action potential, but a *graded potential*
- to simulate correctly, should distinguish between vesicle *release probability* and vesicle *availability*

image from here



Current modeling elements

feature	modelling
ln tonotopic organization	erbspace
approximate frequency	gammatone
OHC/active hearing	compression
amplification (gain)	scalar

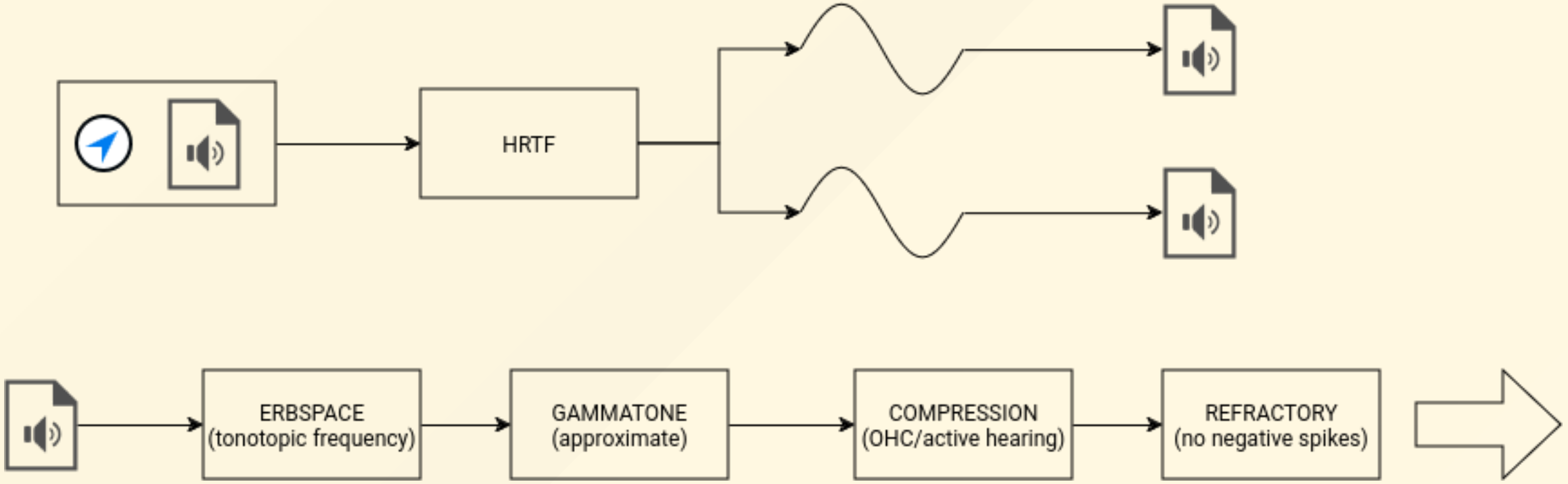
Current modeling elements

feature	modelling
In tonotopic organization	erbspace
~~approximate frequency~~	~~gammatone~~
~~OHC/active hearing~~	~~compression~~
~~amplification (gain)~~	~~scalar~~

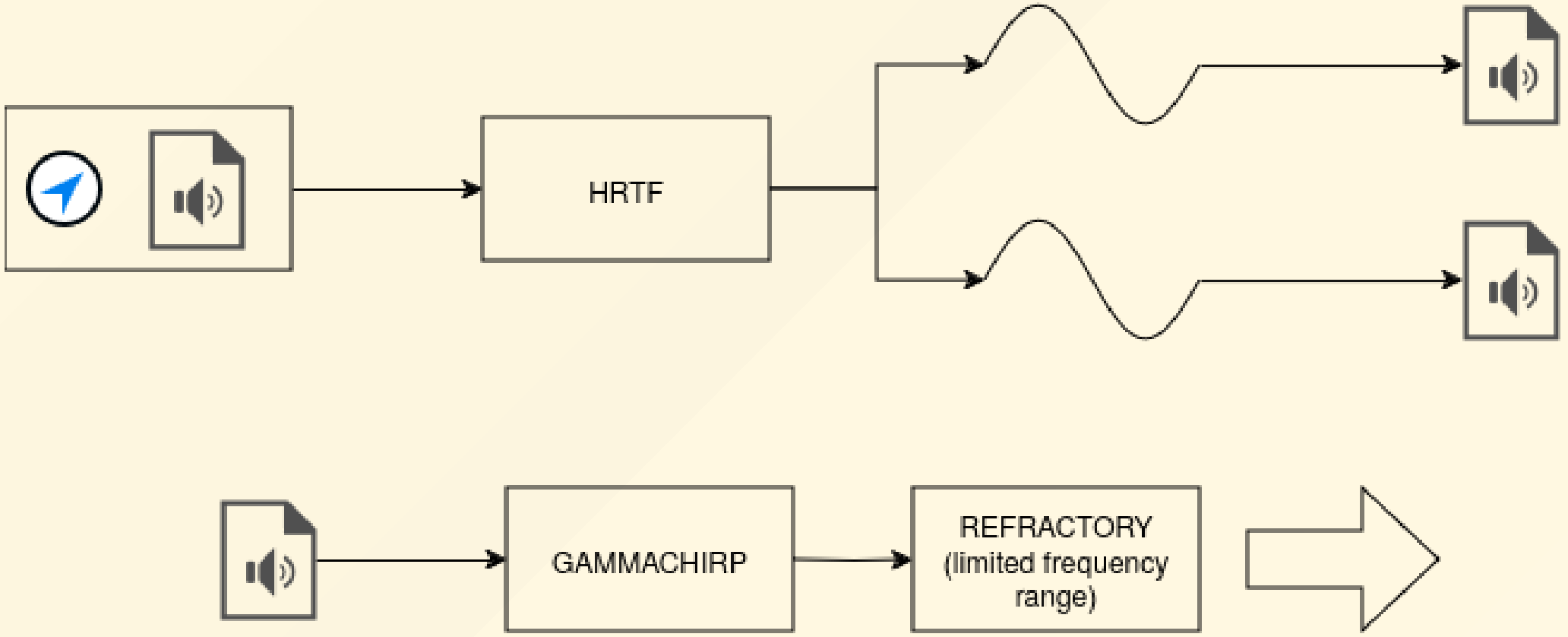
Future modeling elements?

feature	modelling
ln tonotopic organization	erbspace
- approximate, asymmetric frequency - dynamic, justified gain - active hearing, compression	gammachirp

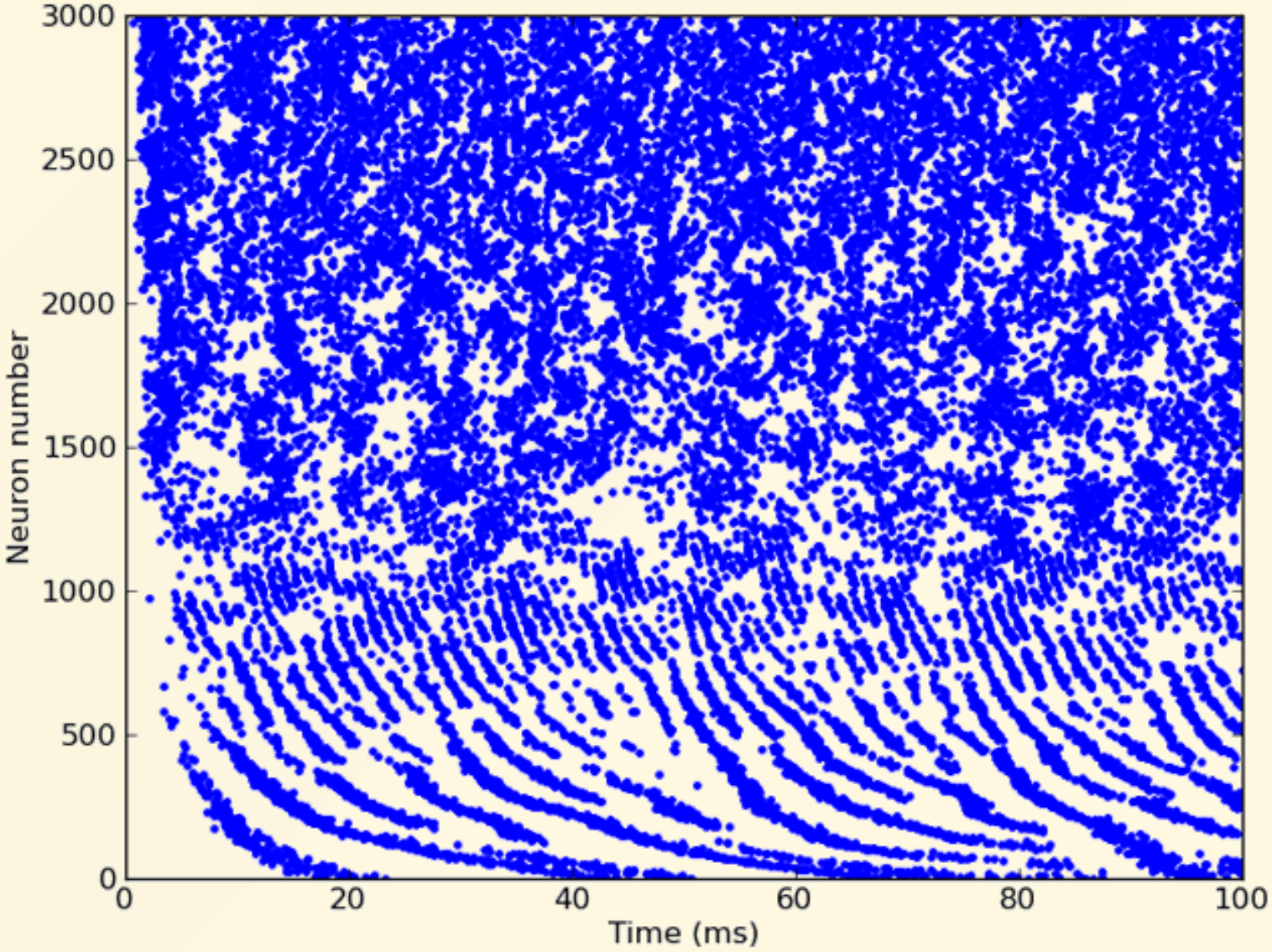
Existing scheme



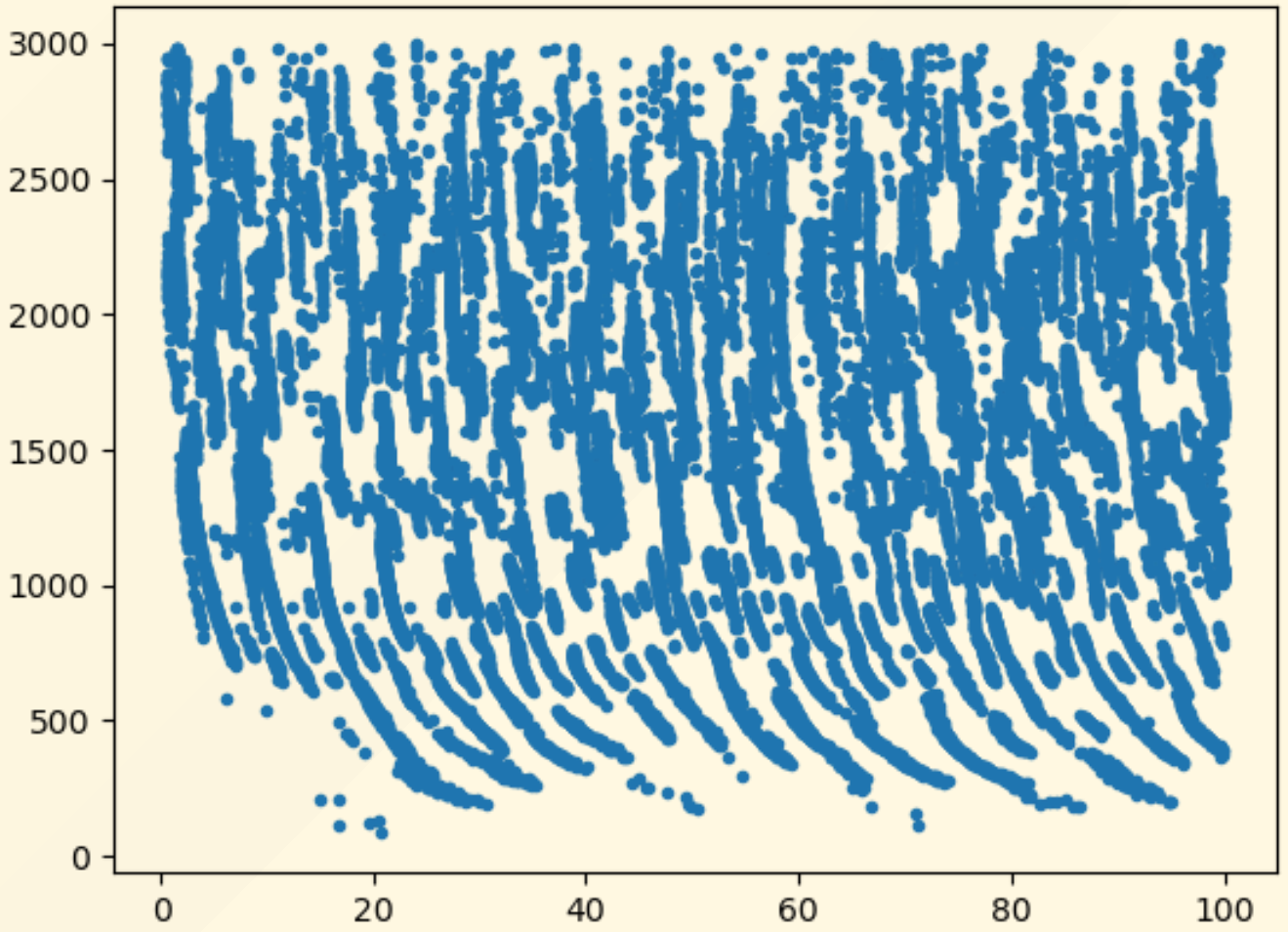
Updated scheme

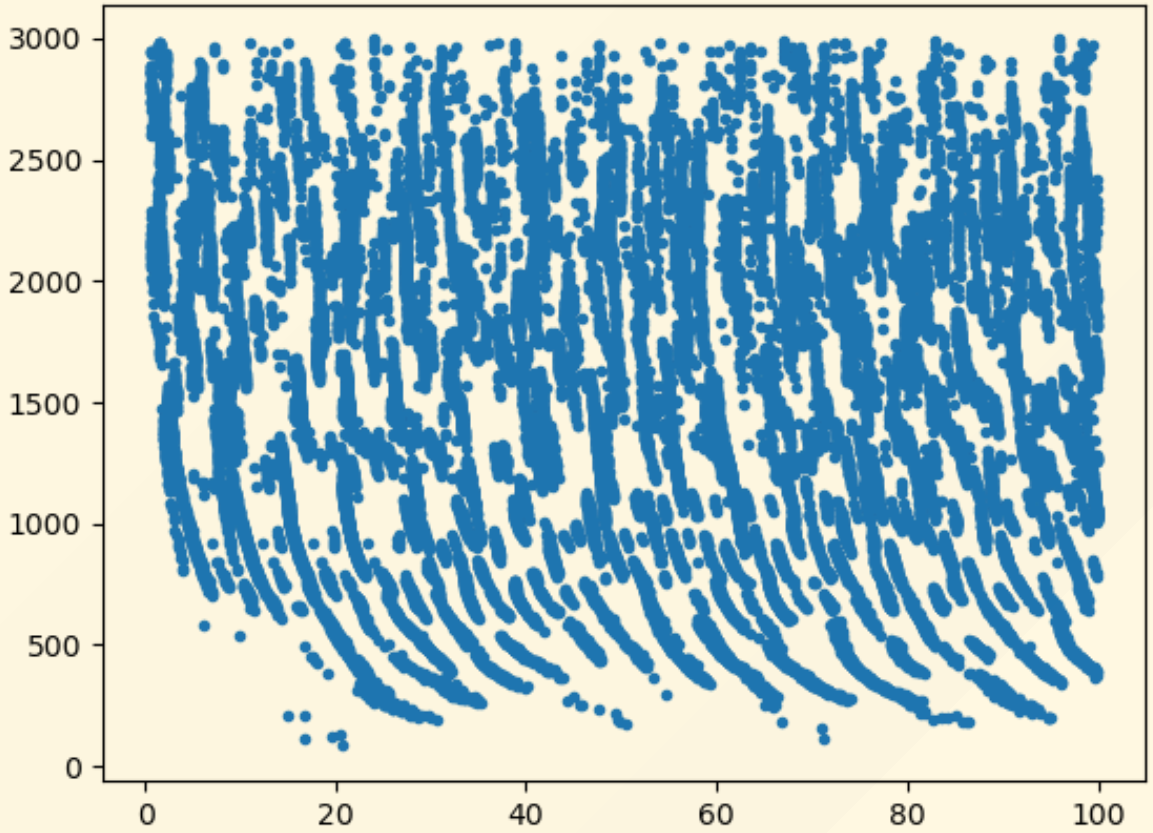


results:


3000
2500
2000
1500
1000
500
0
Neuron number
0
20
40
60
80
100
Time (ms)

results:





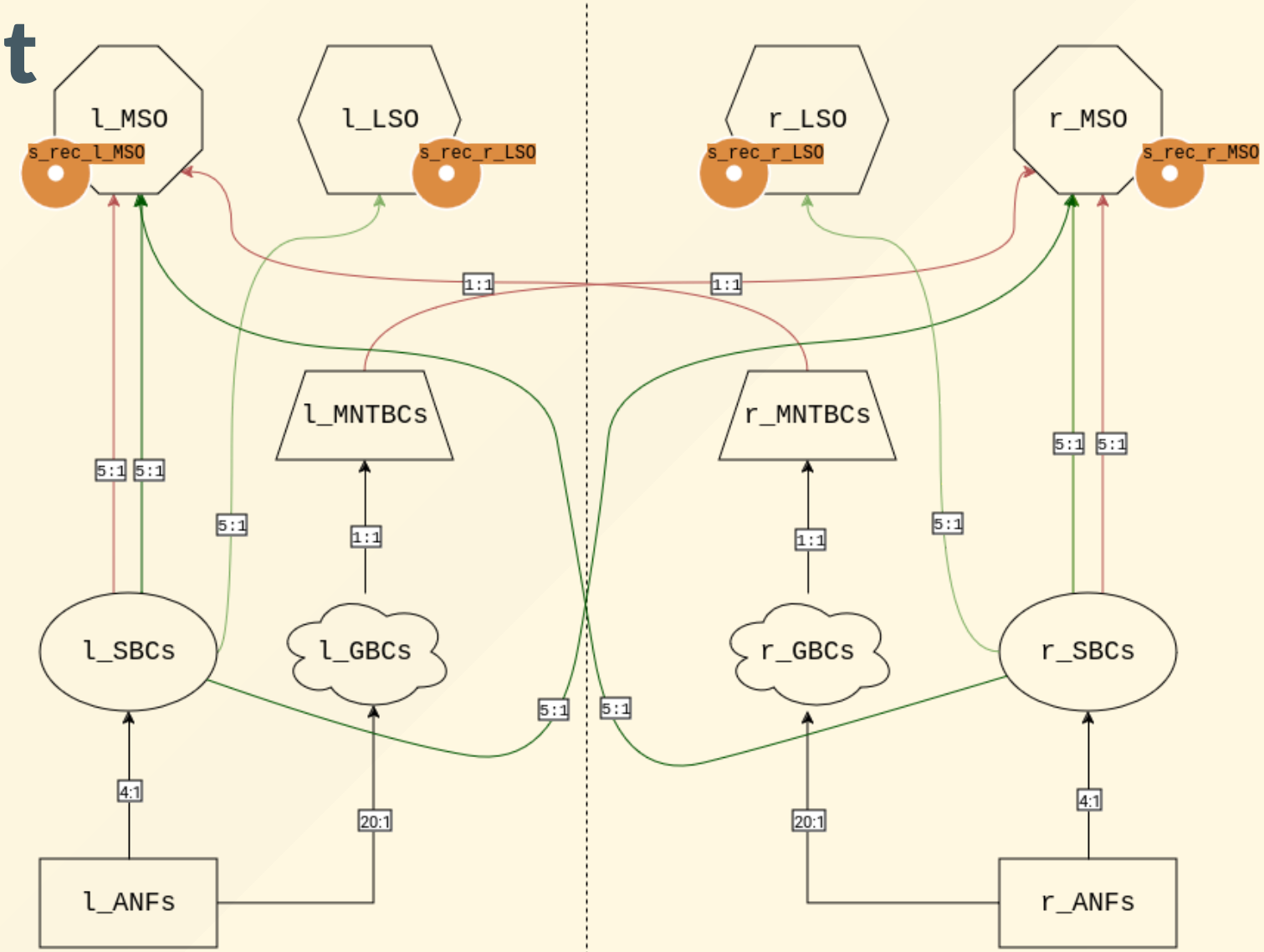
The bad part:

- current results with entire net based on long ablation
- gammachirp shows:
  - phase locking
  - more realistic (?) gain
  - less spikes
  - lower engagement of low freq ANFs
  - needs more ablation

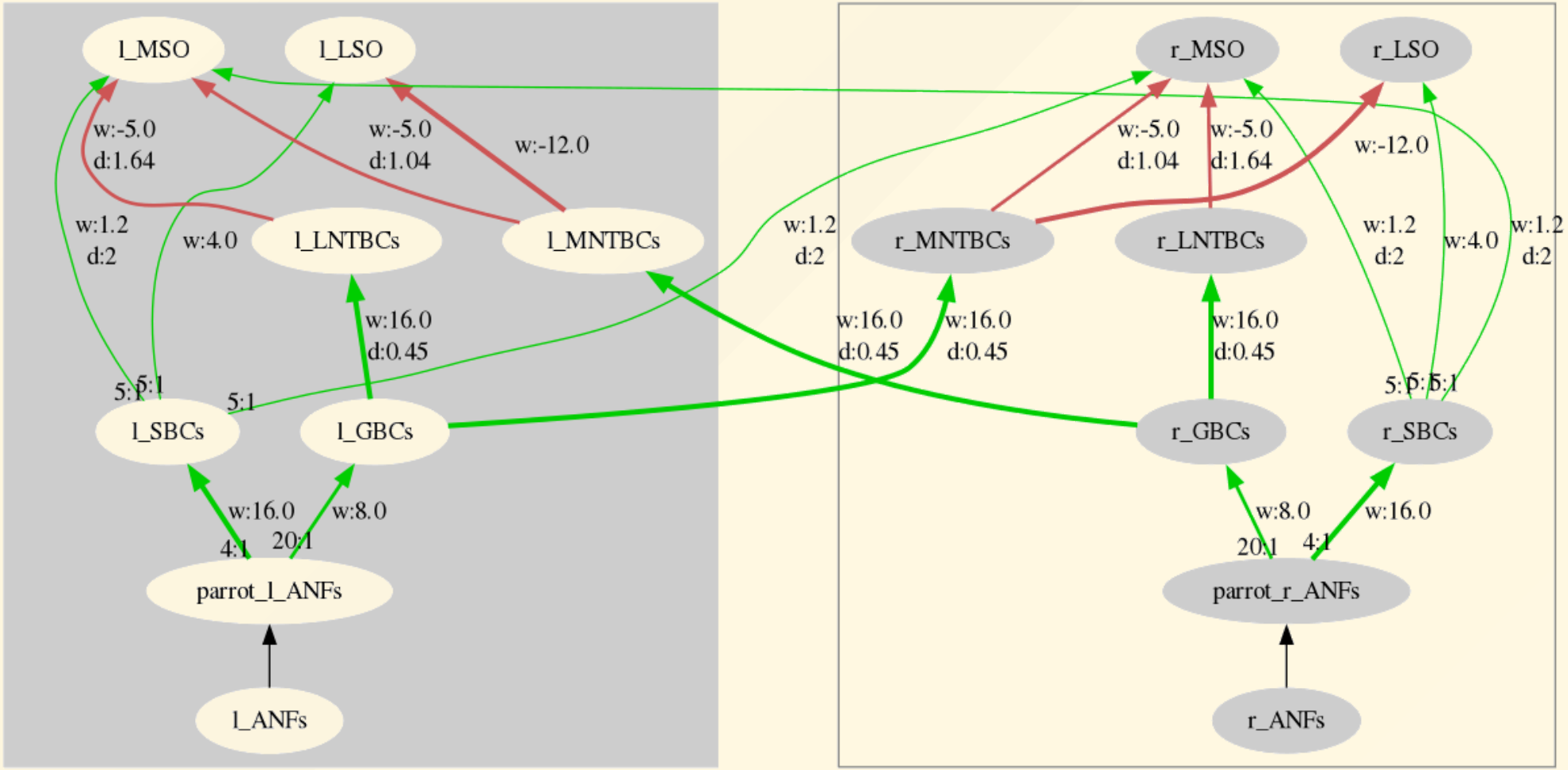
MSO and Plasticity

- so far, we've looked at inputs
- what about the rest of the network?
- net overview
- current issue explanation
- plastic delays?
- plastic inhibition?

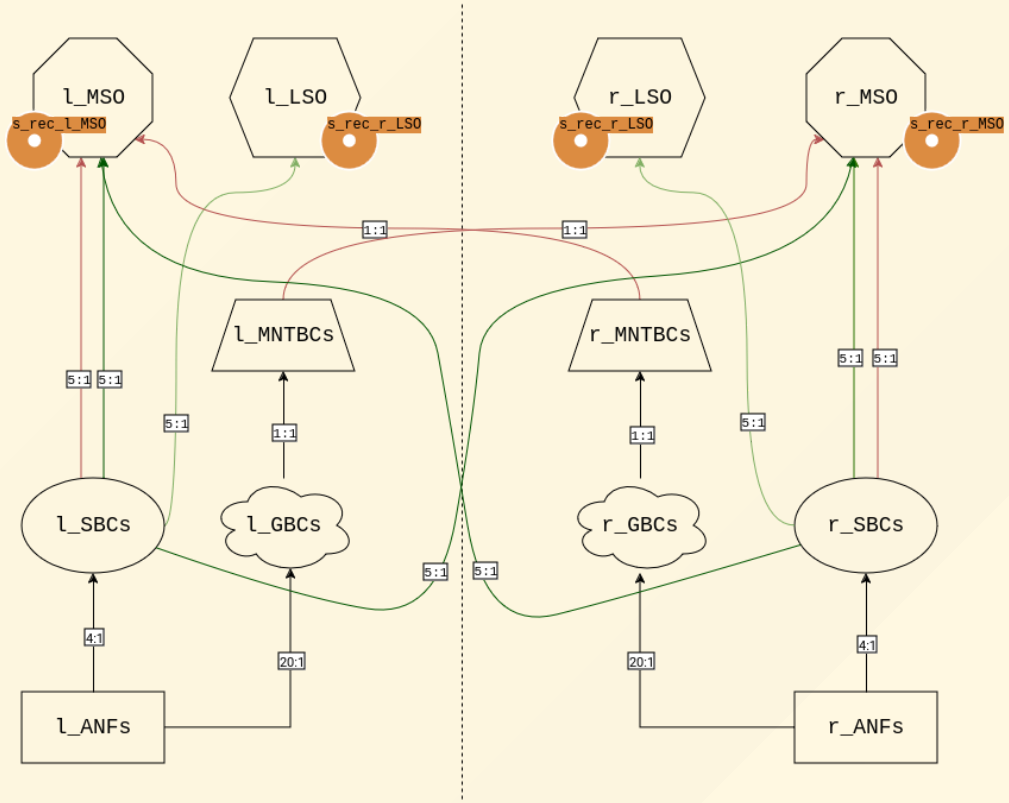
First net

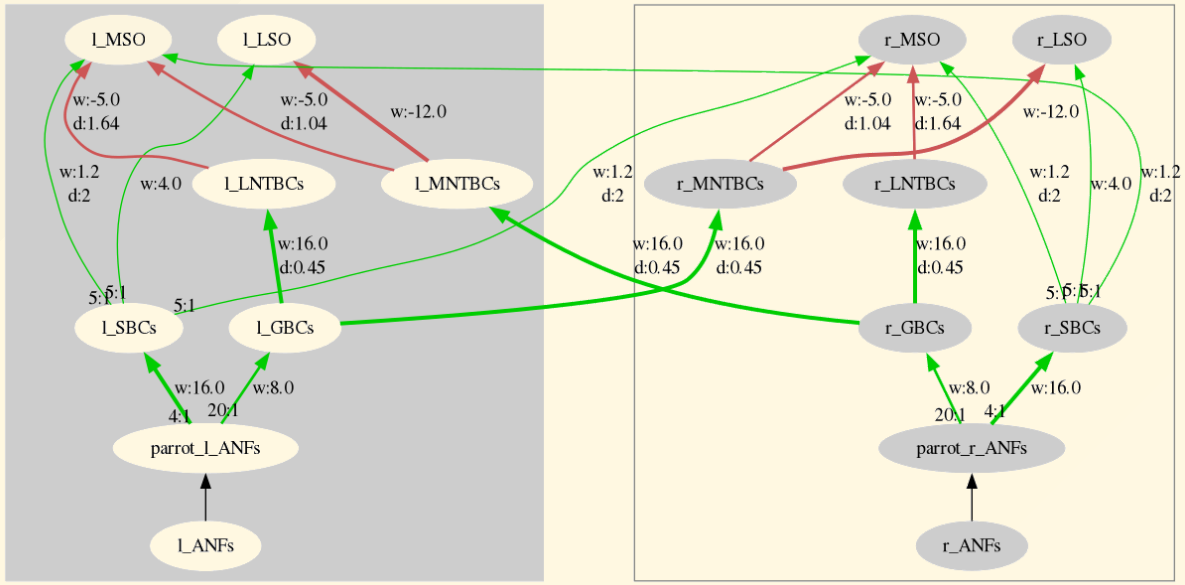

l_MSO
s_rec_l_MSO
l_LSO
s_rec_r_LSO
r_LSO
s_rec_r_LSO
r_MSO
s_rec_r_MSO
1:1
1:1
l_MNTBCs
r_MNTBCs
5:1
5:1
5:1
5:1
5:1
5:1
1:1
1:1
l_SBCs
l_GBCs
r_GBCs
r_SBCs
5:1
5:1
4:1
20:1
20:1
4:1
l_ANFs
r_ANFs

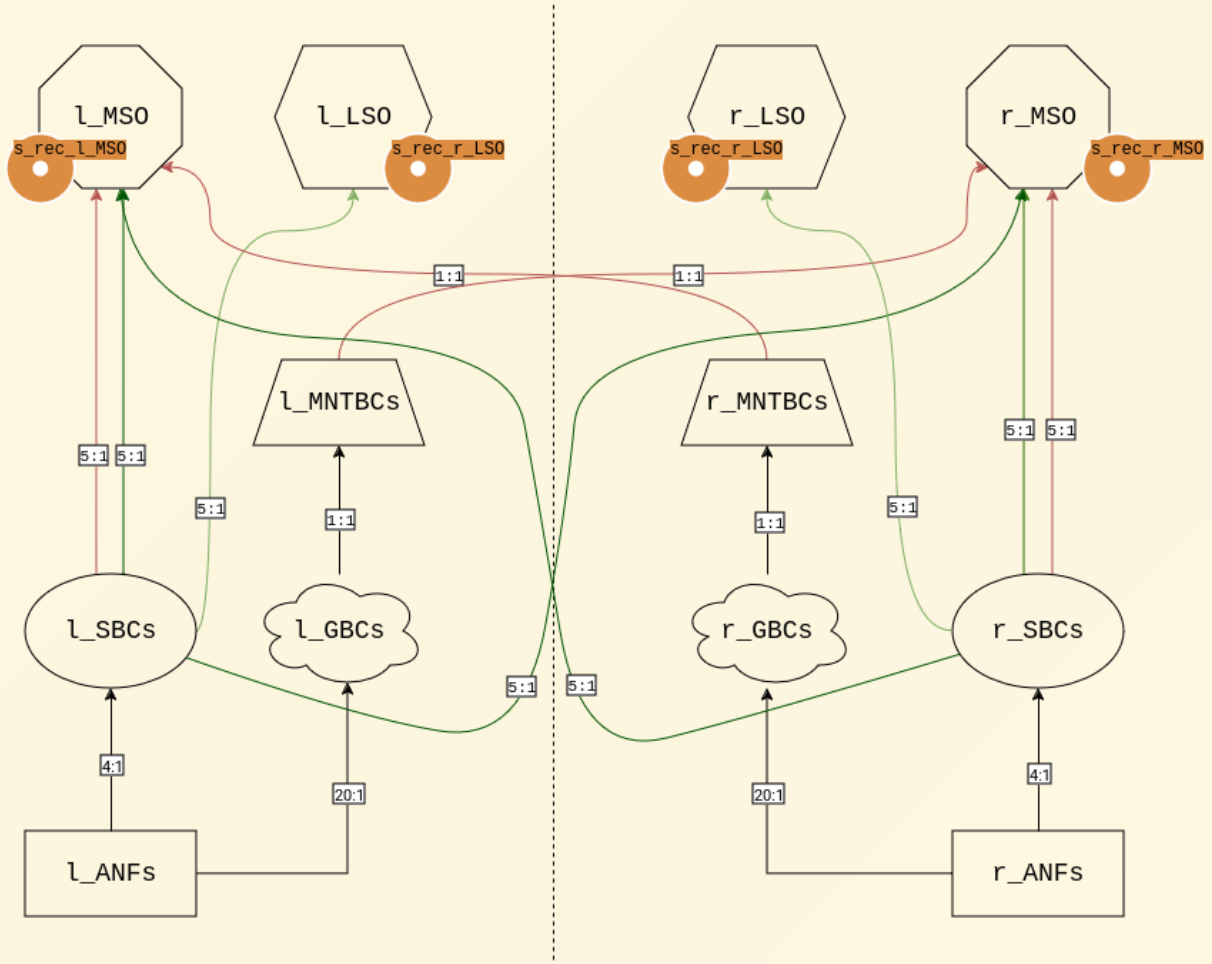
Updated design



Differences



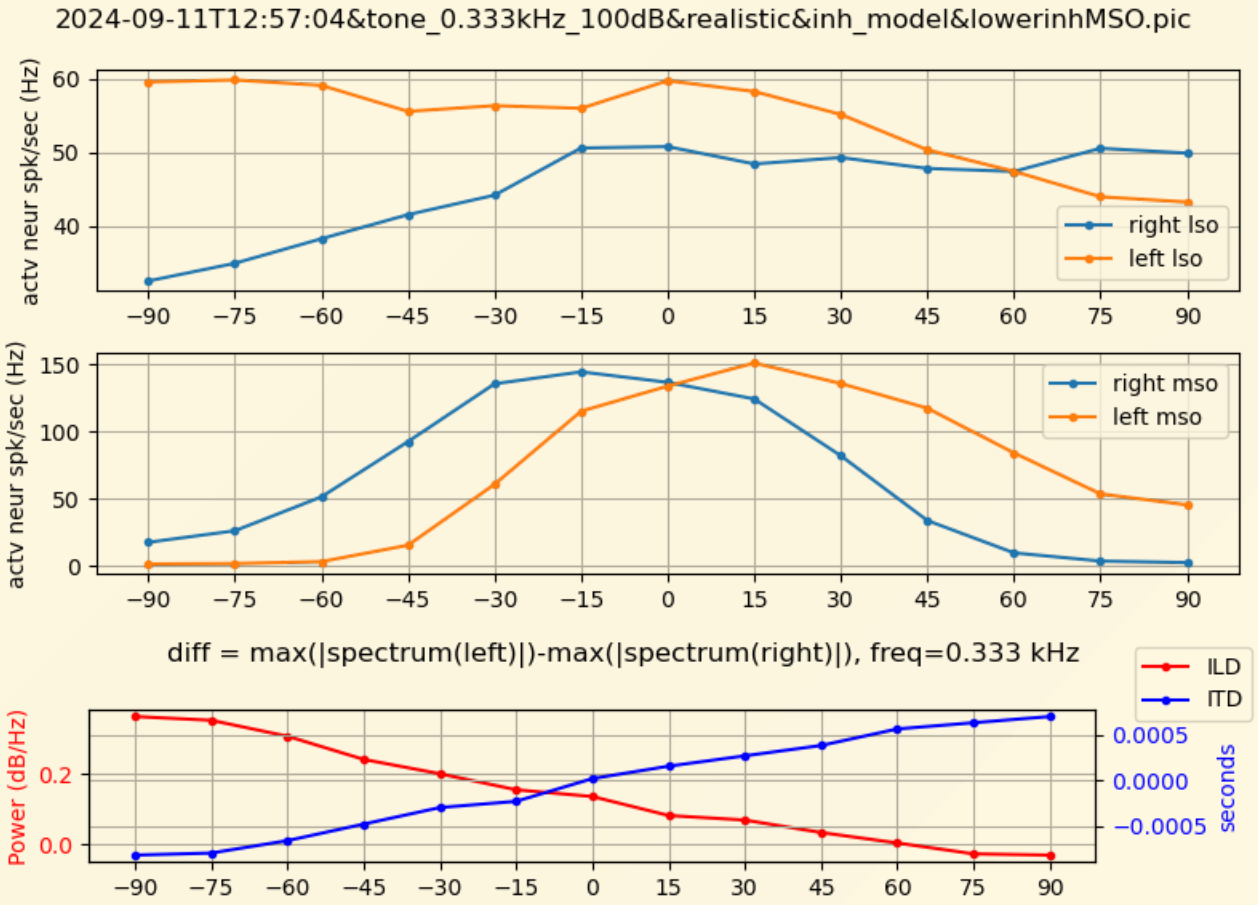




Changes

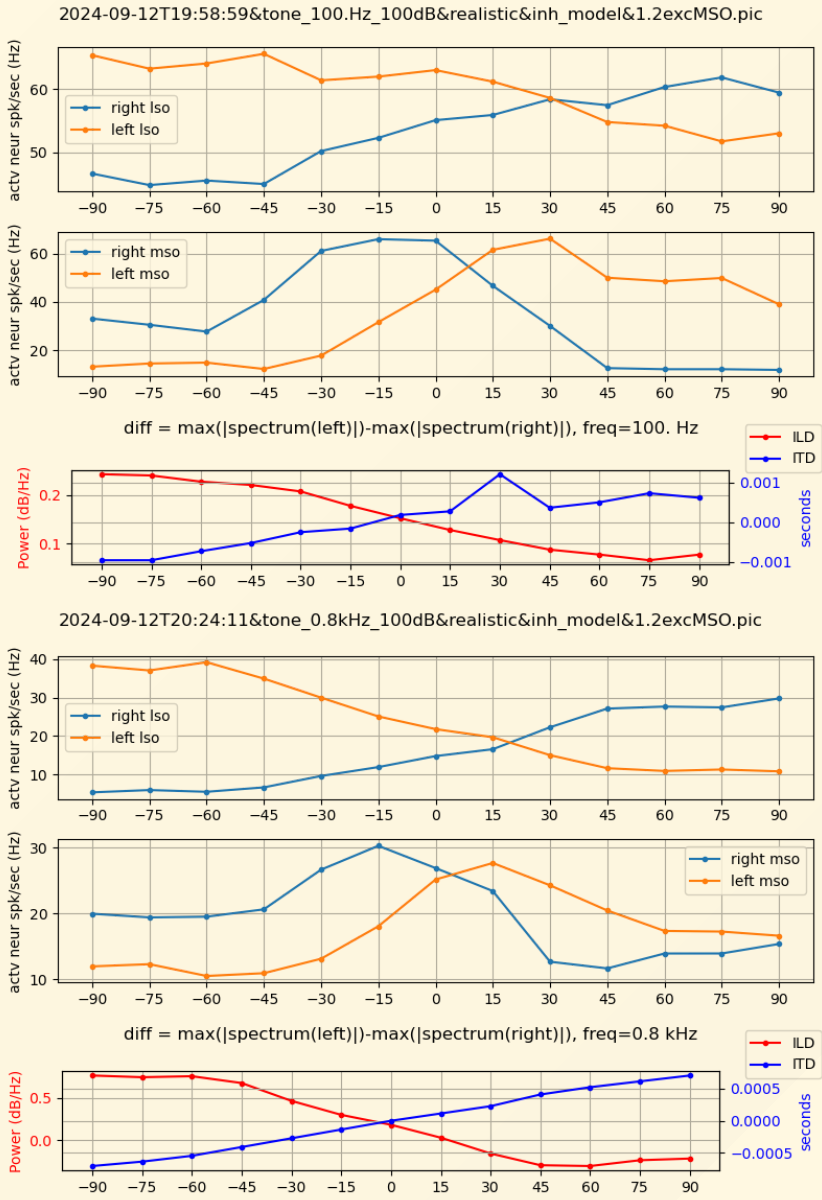
- Design-wise:
  - Updated weights
  - Included LNTB for inh
  - Updated delays (we'll see effect later)
- Technical:
  - Batch processing
  - Cache on disk
  - Batched results

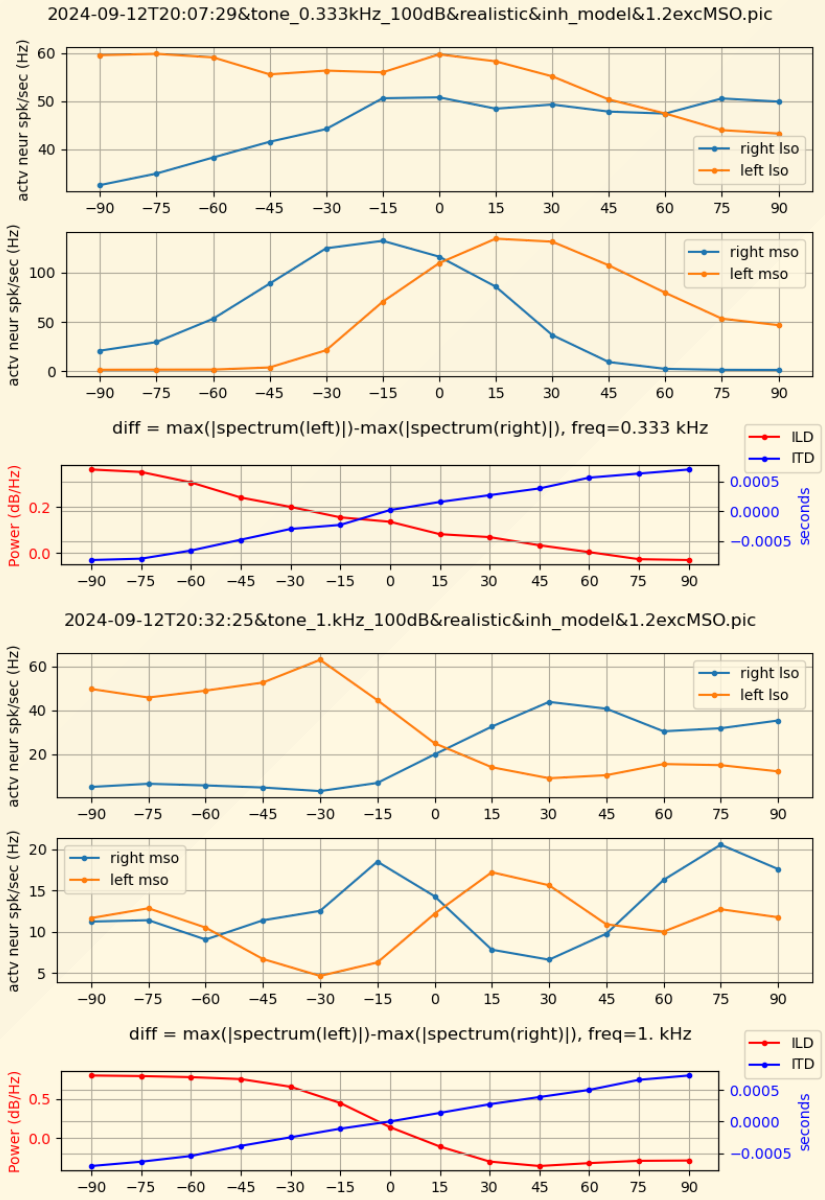
sidenote: results

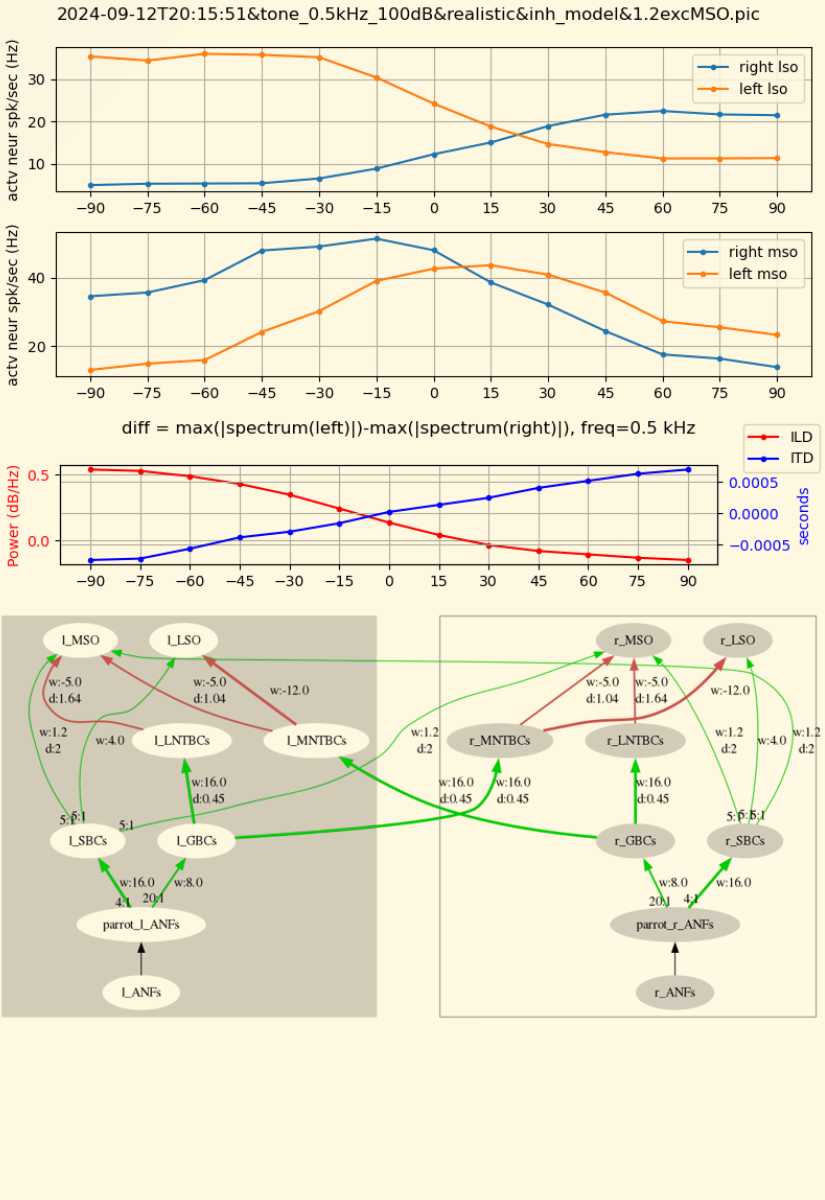


what we look for:

- LSO ipsi, MSO contra
- MSO preferring specific angles, different from zero
- MSO good for low freqs
- LSO crossing around zero
- HRTF results


2024-09-12T19:58:59&tone_100.Hz_100dB&realistic&inh_model&1.2excMSO.pic
actv neur spk/sec (Hz)
right lso
left lso
right mso
left mso
diff = max(|spectrum(left)|)-max(|spectrum(right)|), freq=100. Hz
ILD
ITD
Power (dB/Hz)
seconds
2024-09-12T20:24:11&tone_0.8kHz_100dB&realistic&inh_model&1.2excMSO.pic
diff = max(|spectrum(left)|)-max(|spectrum(right)|), freq=0.8 kHz


2024-09-12T20:07:29&tone_0.333kHz_100dB&realistic&inh_model&1.2excMSO.pic
actv neur spk/sec (Hz)
right lso
left lso
right mso
left mso
diff = max(|spectrum(left)|)-max(|spectrum(right)|), freq=0.333 kHz
ILD
ITD
Power (dB/Hz)
seconds
2024-09-12T20:32:25&tone_1.kHz_100dB&realistic&inh_model&1.2excMSO.pic
diff = max(|spectrum(left)|)-max(|spectrum(right)|), freq=1. kHz


2024-09-12T20:15:51&tone_0.5kHz_100dB&realistic&inh_model&1.2excMSO.pic
actv neur spk/sec (Hz)
right lso
left lso
right mso
left mso
diff = max(|spectrum(left)|)-max(|spectrum(right)|), freq=0.5 kHz
ILD
ITD
Power (dB/Hz)
seconds

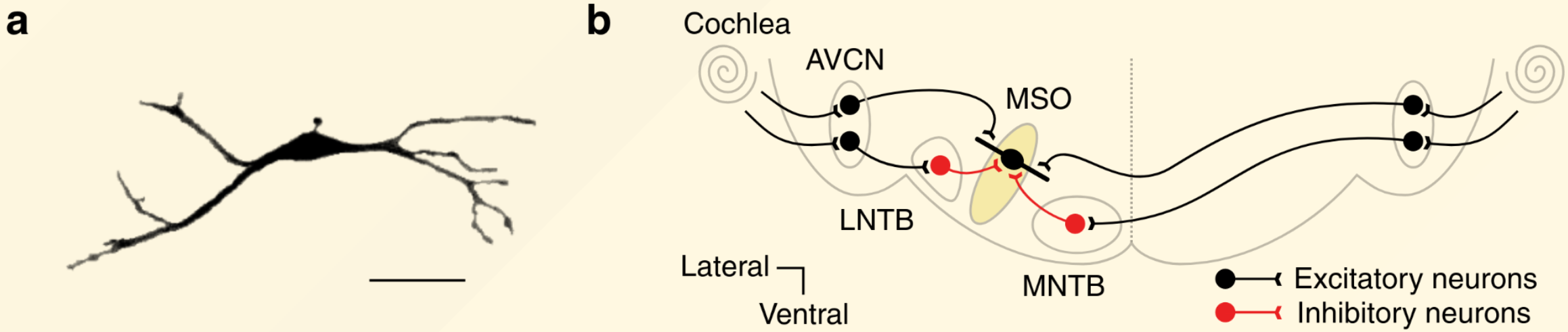
l_MSO
l_LSO
l_LNTBCs
l_MNTBCs
l_SBCs
l_GBCs
parrot_l_ANFs
l_ANFs
r_MSO
r_LSO
r_MNTBCs
r_LNTBCs
r_GBCs
r_SBCs
parrot_r_ANFs
r_ANFs

What's behind the MSO?

- to process ITDs, you need single-spike sensitivity
- inhibition is clearly very important:
  - calyx of Held
  - bilateral inhibition
  - phase locked
- the answer must be with inhibition (*)

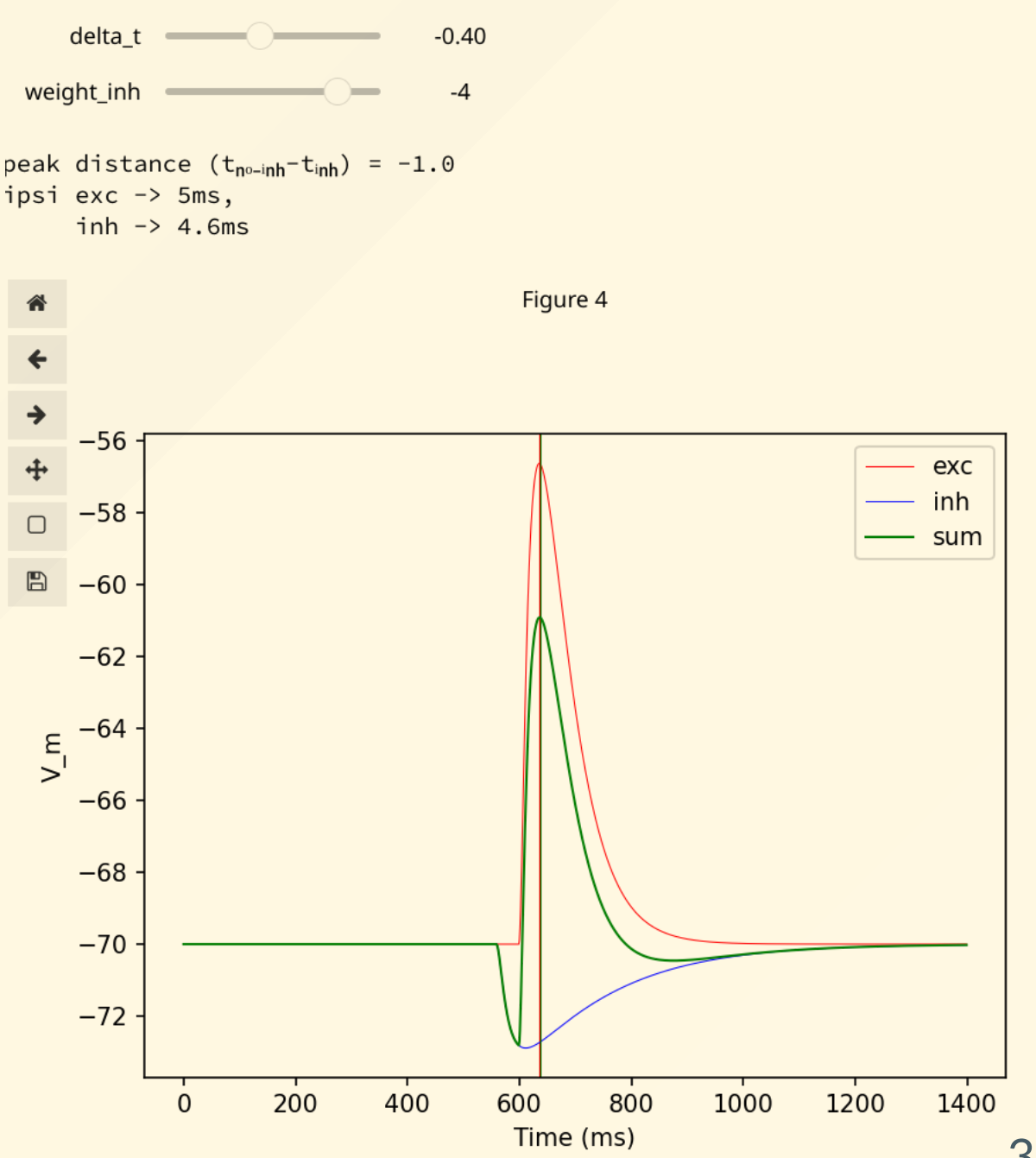
MSO structure

- bipolar
- bilateral inputs
- 2 inh, 2 exc



Start easy: two inputs

- peak voltage time is shifted
- shift possible in both directions
- synaptic weight matters... a bit

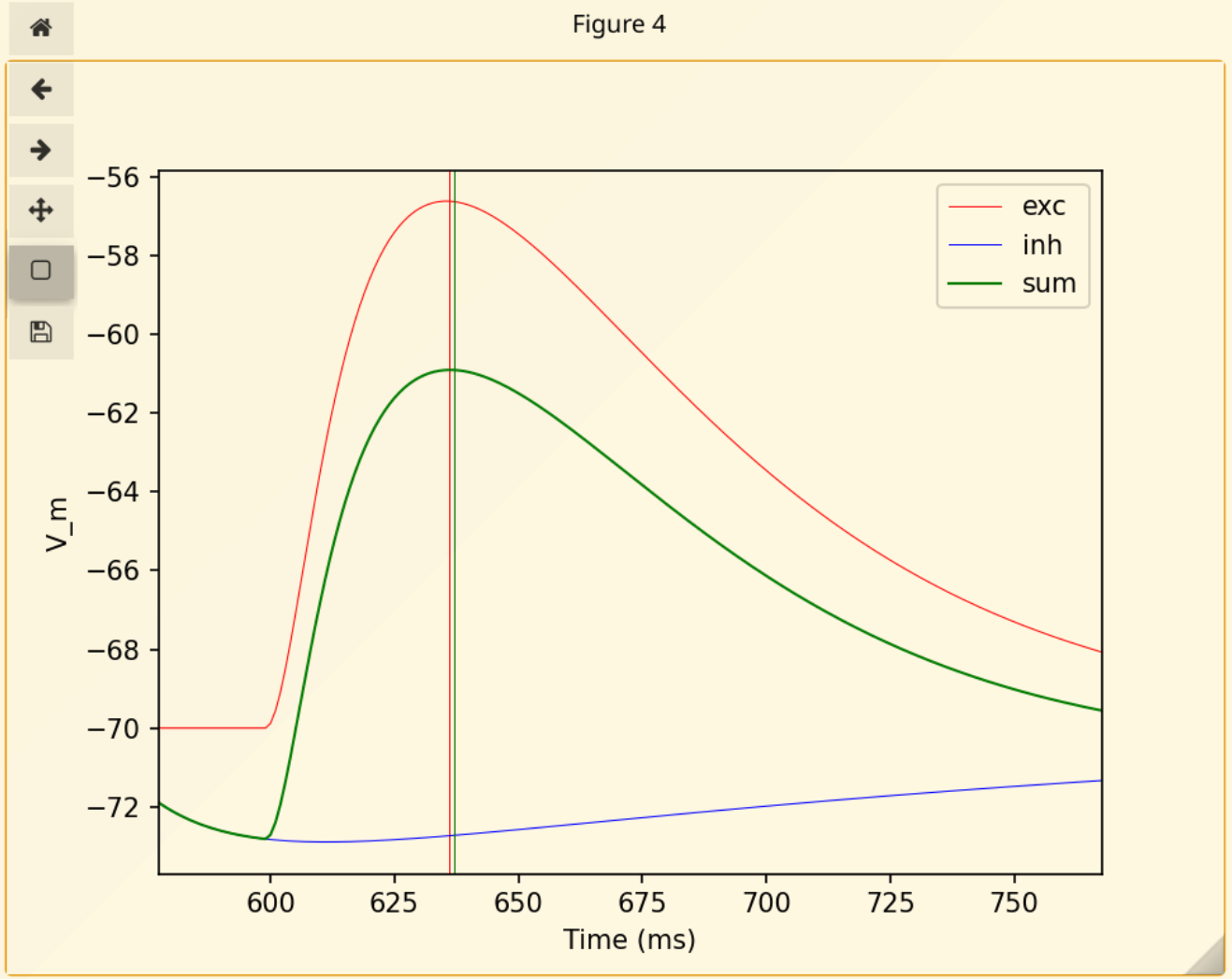


Zoomed

delta_t -0.40

weight_inh -4

peak distance (t_{no-inh}-t_{inh}) = -1.0
ipsi exc -> 5ms,
inh -> 4.6ms

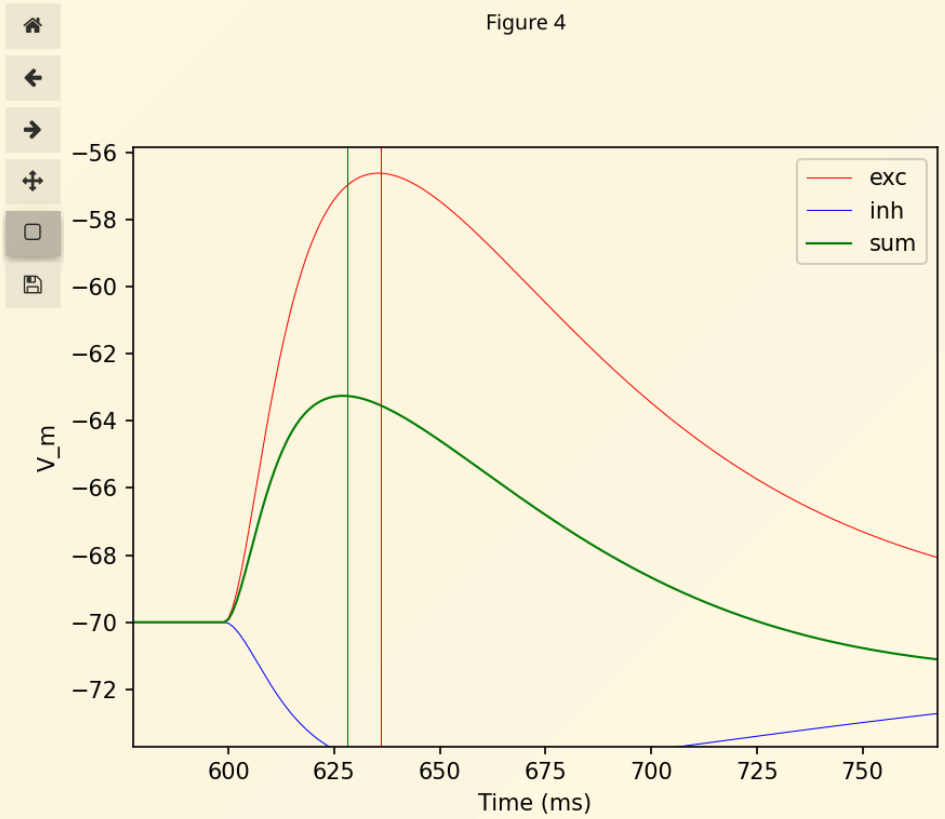


Peak delay

delta_t 0.00

weight_inh -7

peak distance (t_{no-inh}-t_{inh}) = 8.0
ipsi exc -> 5ms,
inh -> 5.0ms

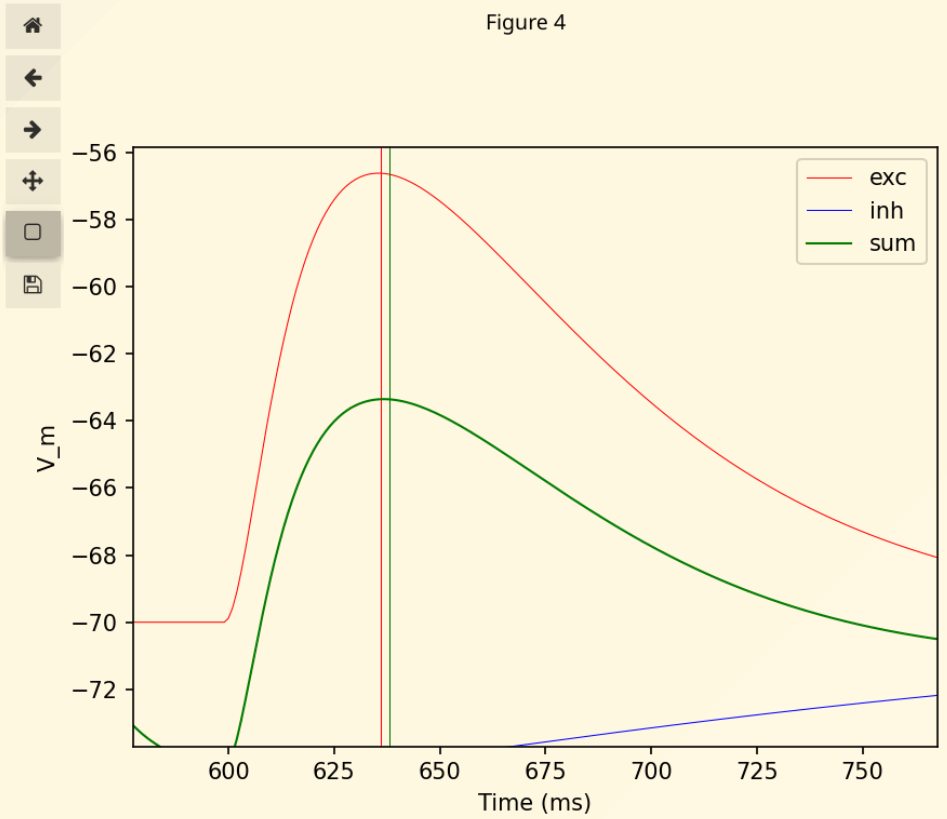


Peak advance

delta_t -0.40

weight_inh -7

peak distance (t_{no-inh}-t_{inh}) = -2.0
ipsi exc -> 5ms,
inh -> 4.6ms



Three inputs

- modulating only contra IPSP
- best ITD both delayed and advanced

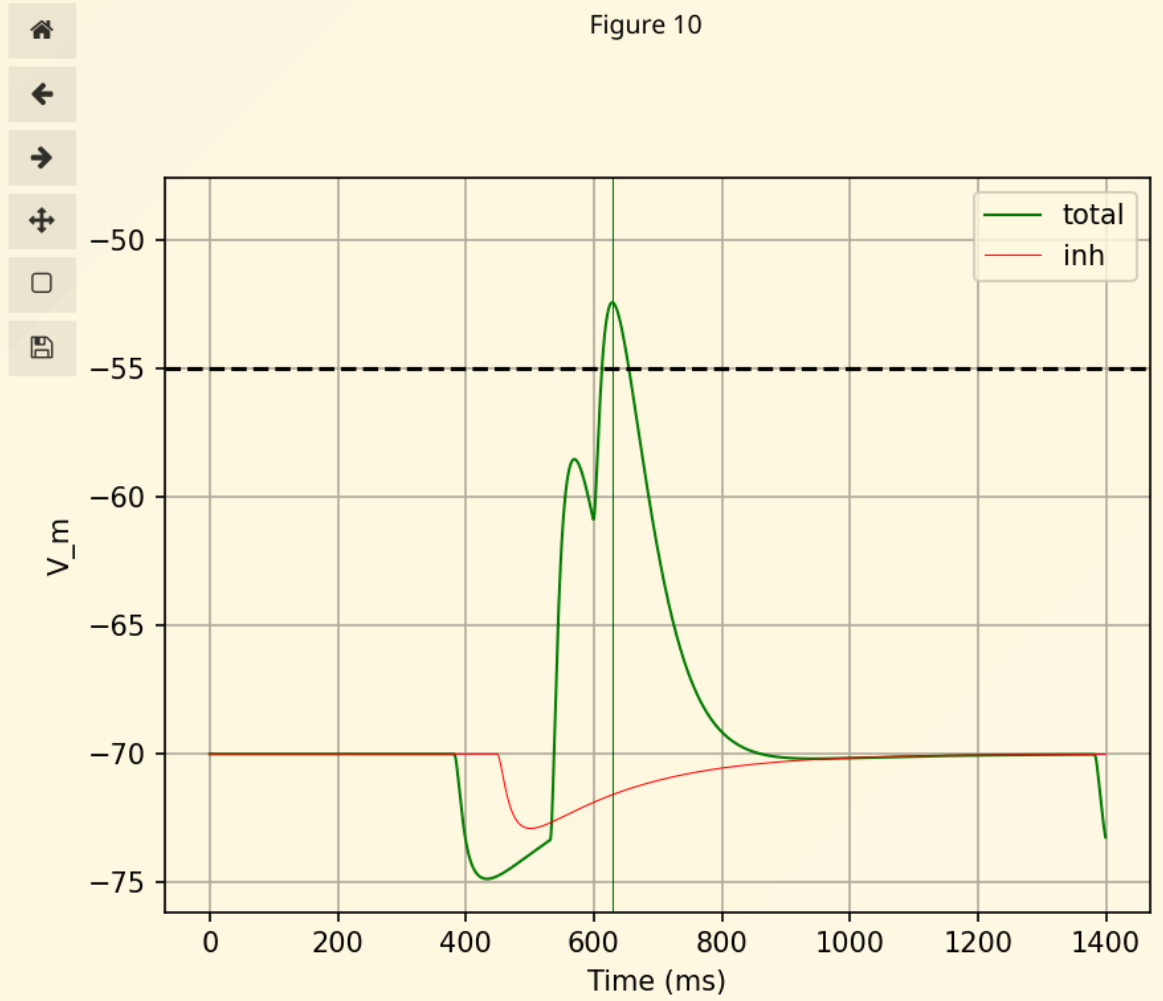
angle -90

deltat_inh -1.50

weight_inh -8

```
this is a right-side MSO cell
ipsi exc -> 5ms,        cntr exc -> 4.33ms,
                                inh -> 2.83ms
```

Figure 10



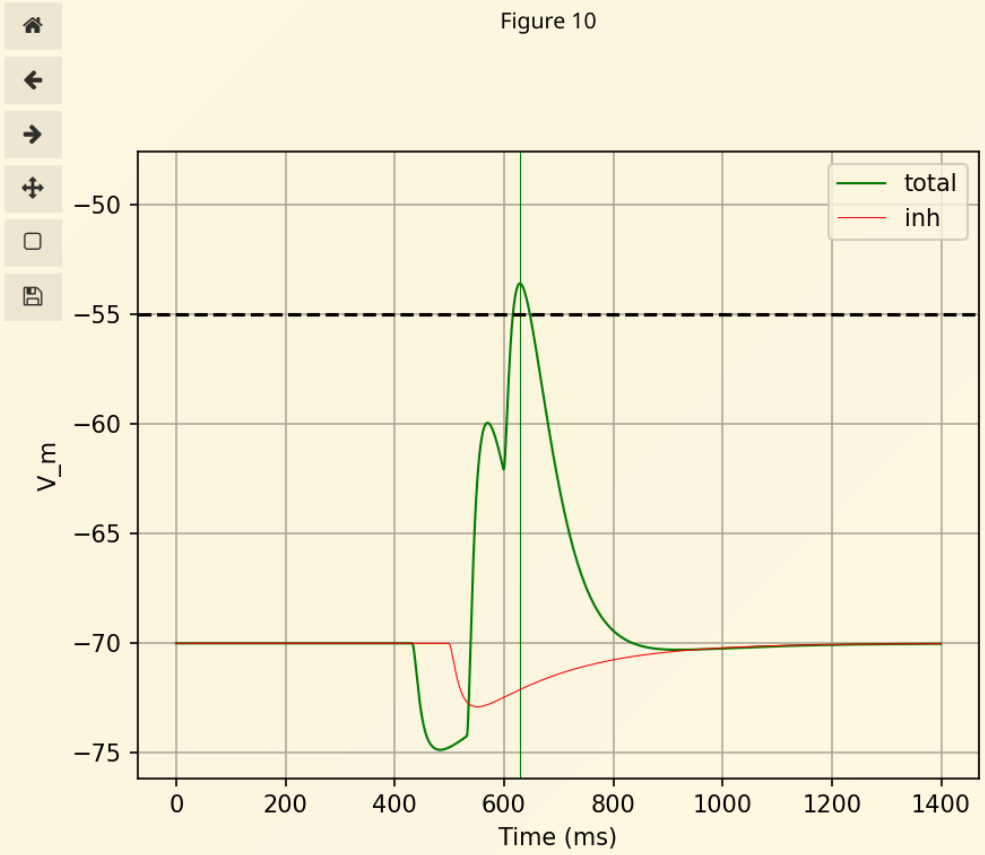
angle -90

deltat_inh -1.00

weight_inh -8

```
this is a right-side MSO cell
ipsi exc -> 5ms,        cntr exc -> 4.33ms,
                                inh -> 3.33ms
```

Figure 10



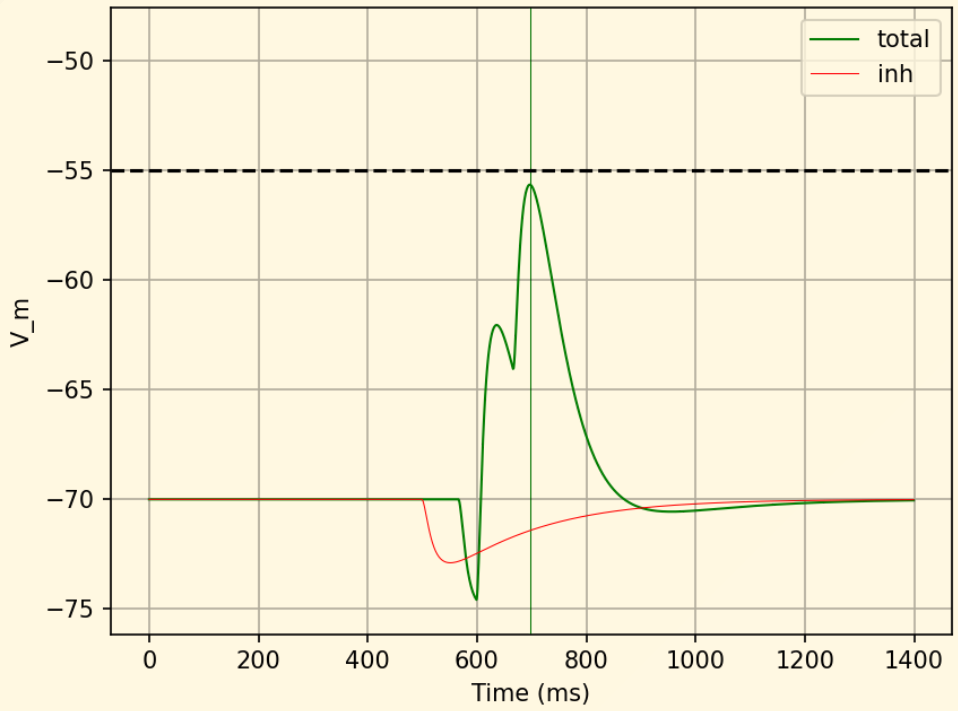
angle 90

deltat_inh -1.00

weight_inh -8

```
this is a right-side MSO cell
ipsi exc -> 5ms,        cntr exc -> 5.67ms,
                                inh -> 4.67ms
```

Figure 10



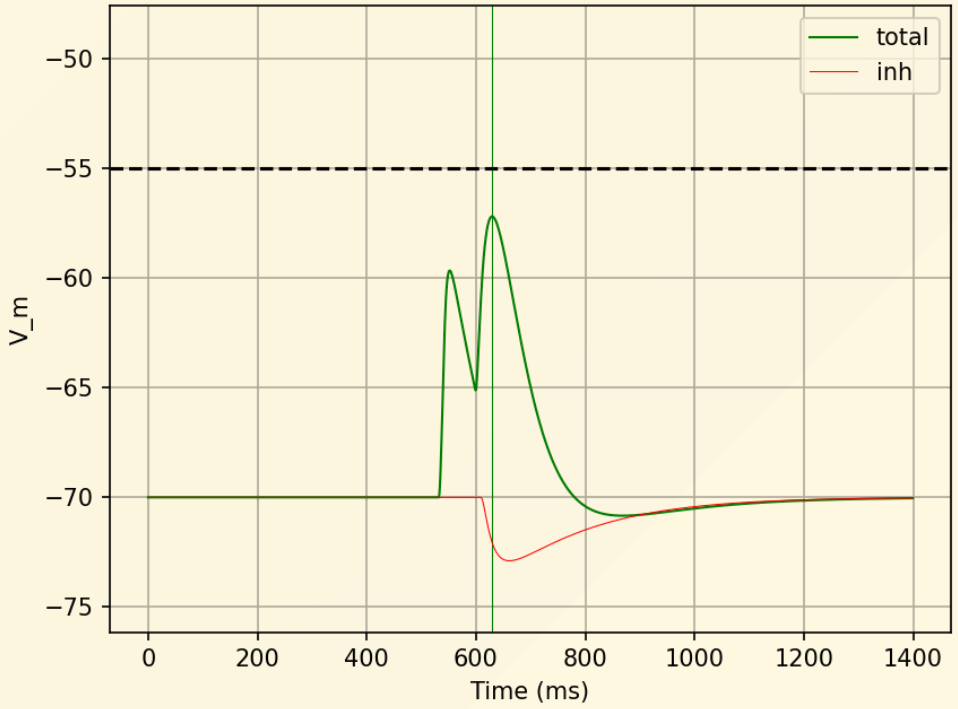
angle -90

deltat_inh 0.10

weight_inh -8

```
this is a right-side MSO cell
ipsi exc -> 5ms,        cntr exc -> 4.33ms,
                                inh -> 4.43ms
```

Figure 10



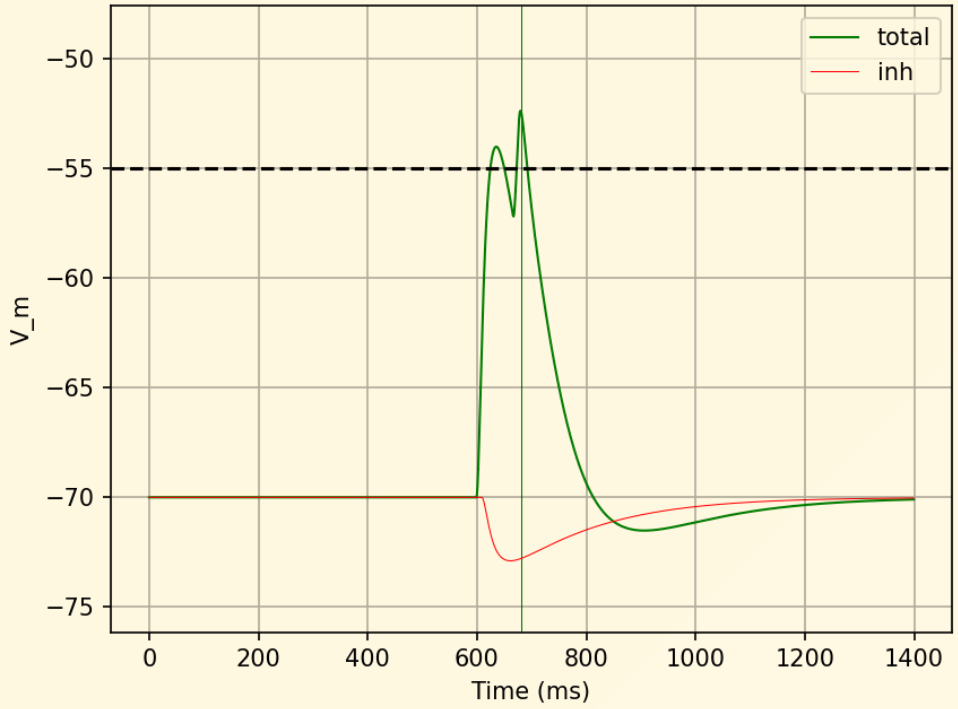
angle 90

deltat_inh 0.10

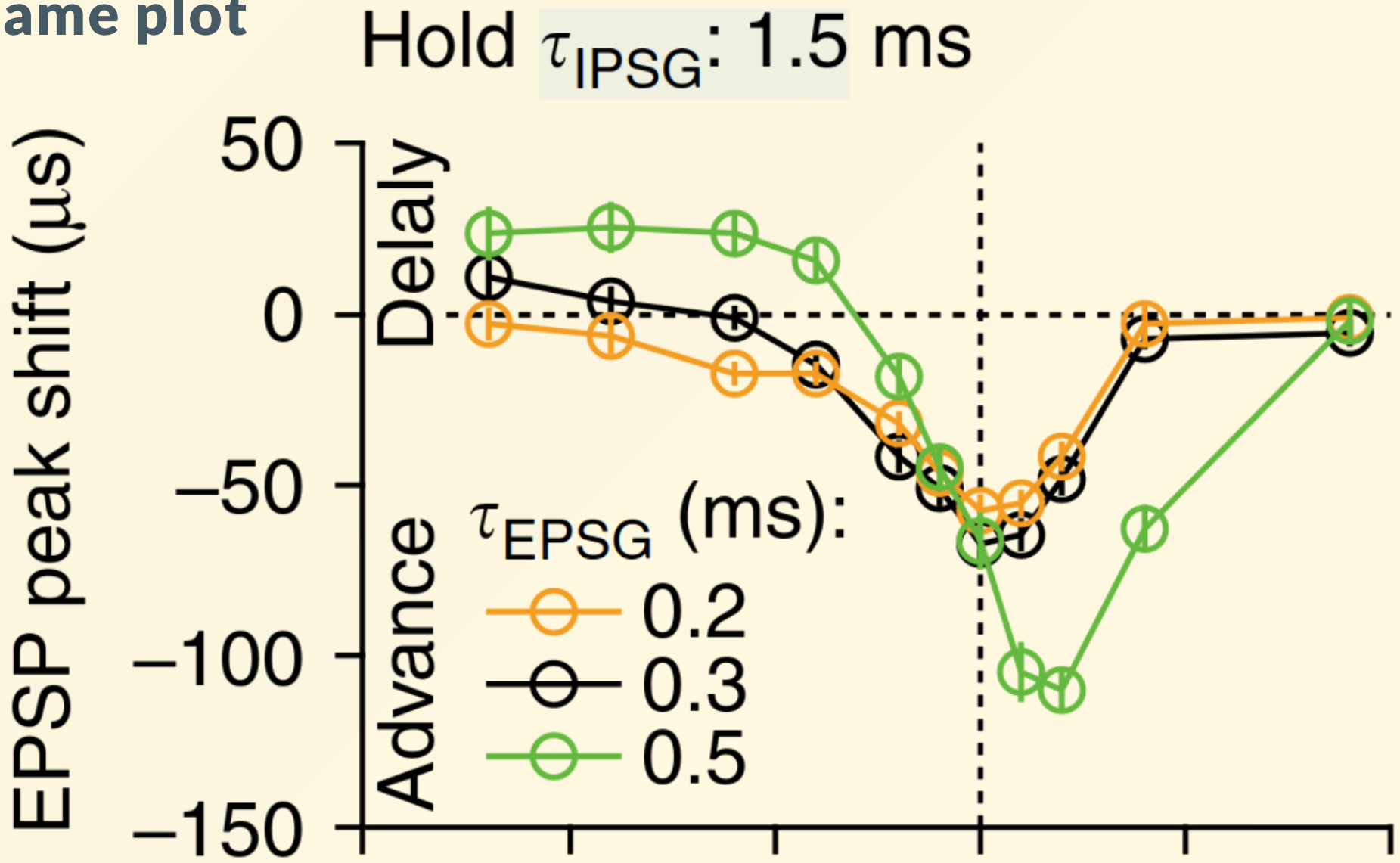
weight_inh -8

```
this is a right-side MSO cell
ipsi exc -> 5ms,        cntr exc -> 5.67ms,
                                inh -> 5.77ms
```

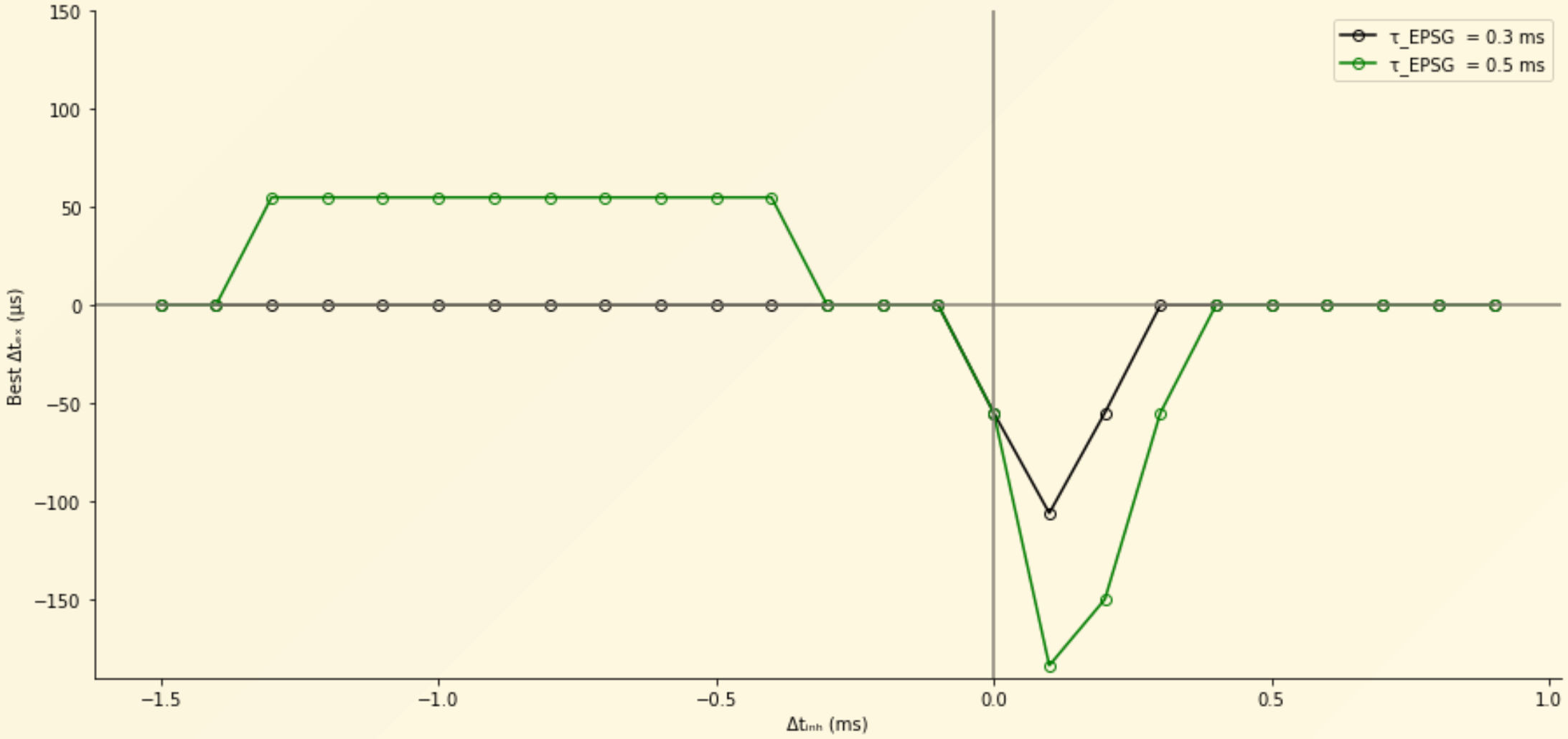
Figure 10



In the same plot



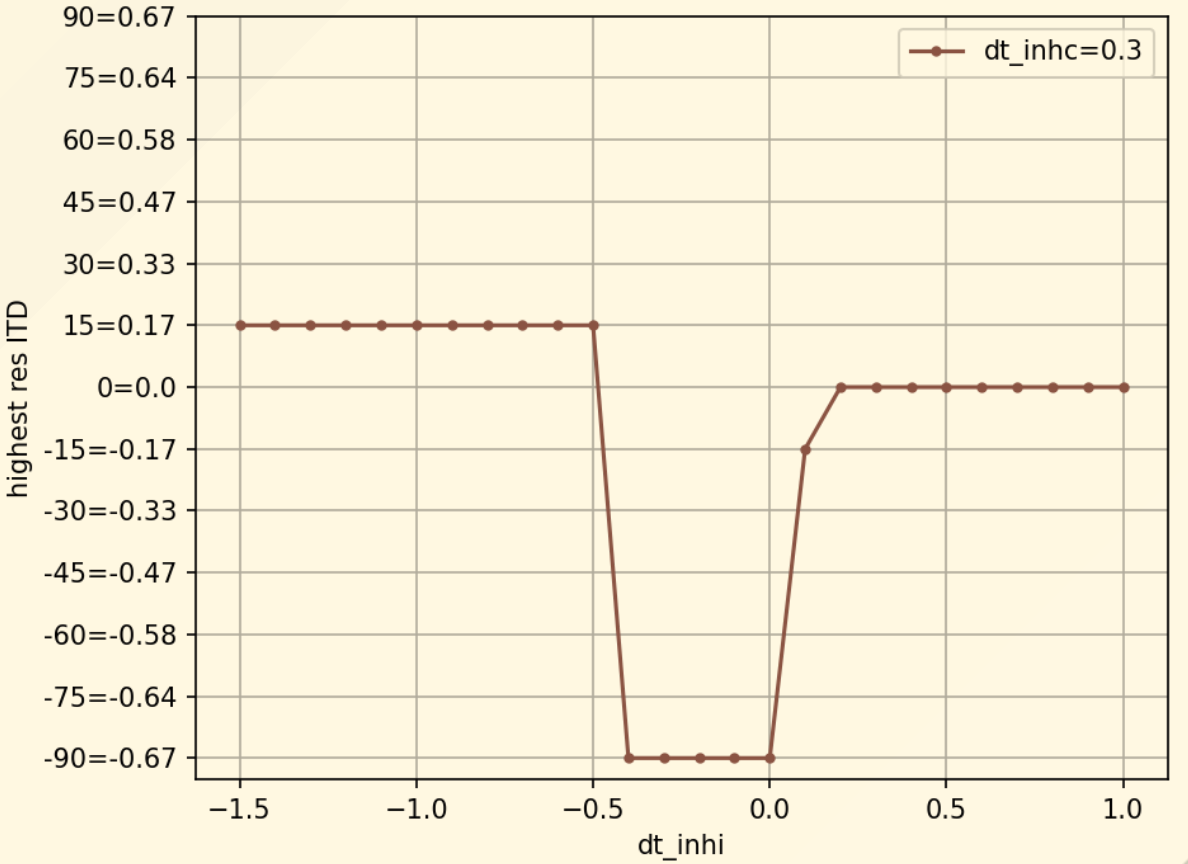
Replicated (desa!)



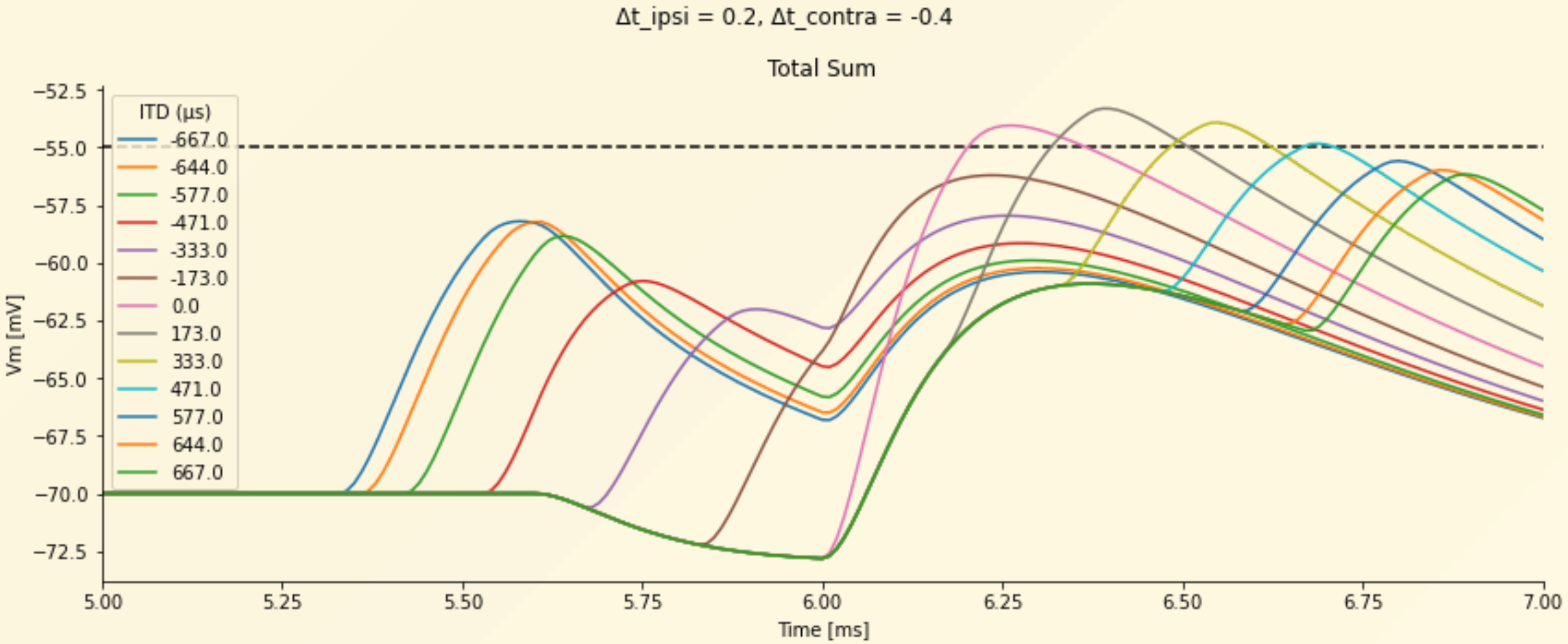
Four inputs

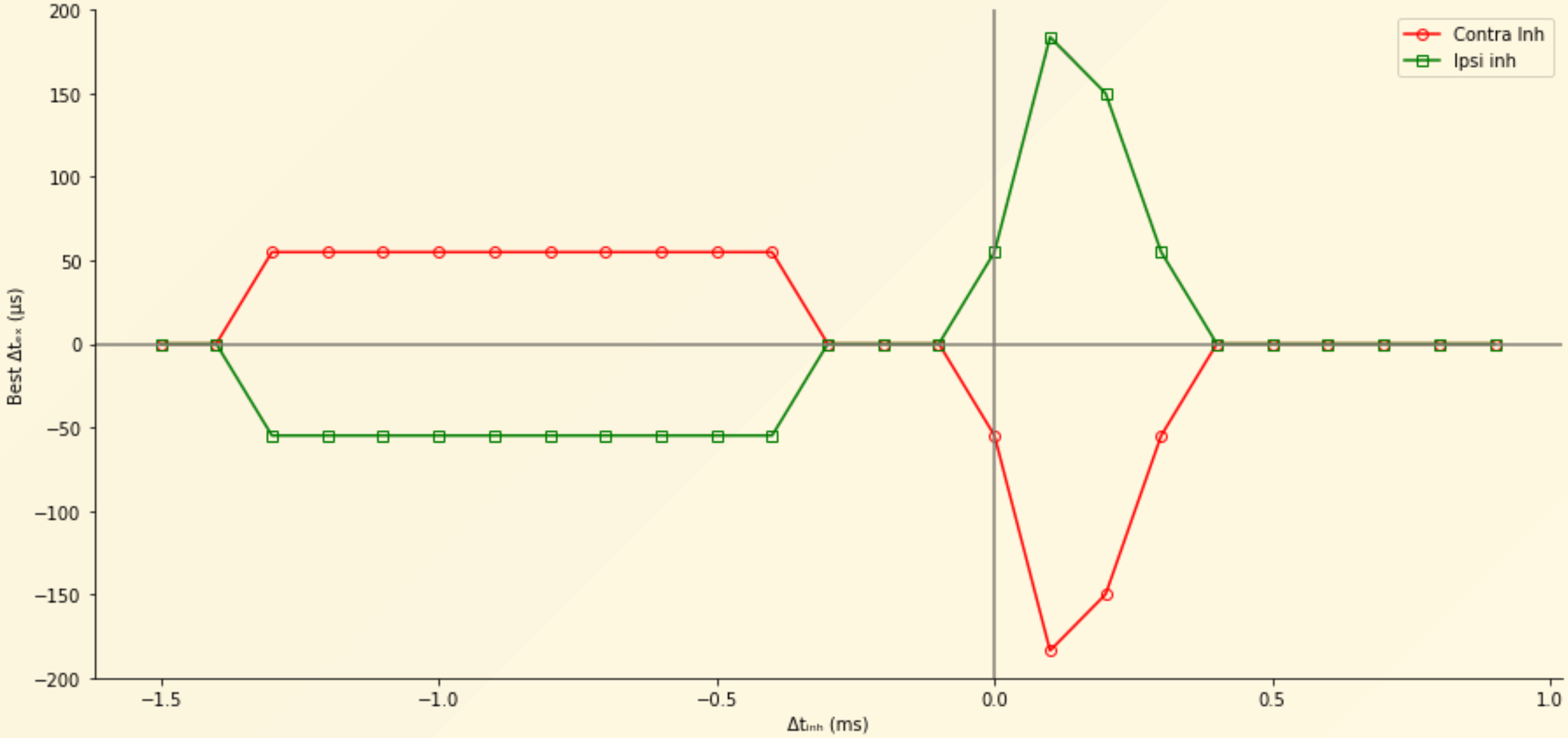
- similar results to three inputs: if you change one keeping the other stable you get a similar curve
- this means an extra knob to turn
- still needs timing plasticity

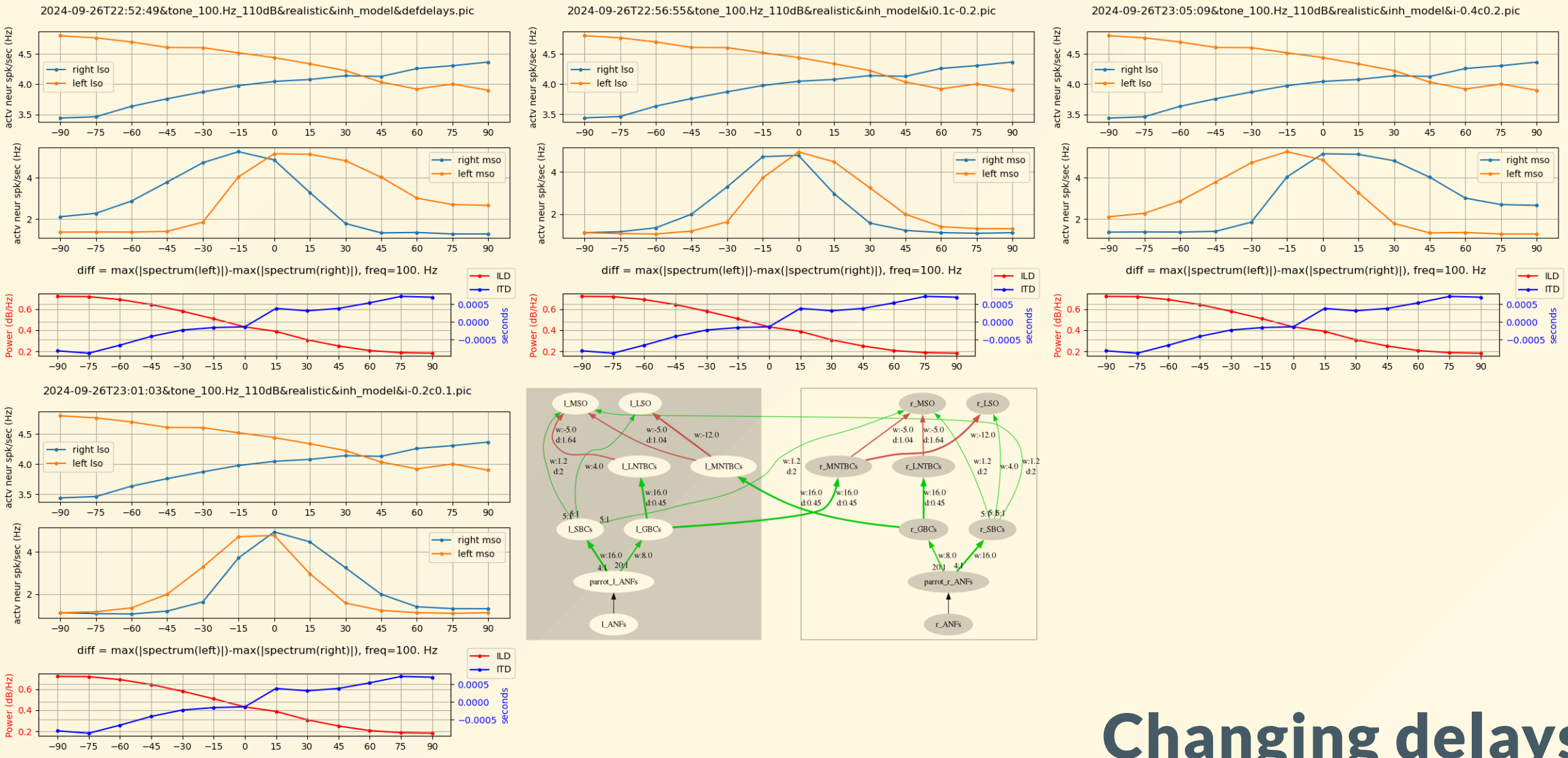
Figure 12



Desa result


Δt_ipsi = 0.2, Δt_contra = -0.4
Total Sum
ITD (µs)
-667.0
-644.0
-577.0
-471.0
-333.0
-173.0
0.0
173.0
333.0
471.0
577.0
644.0
667.0
Vm [mV]
−52.5
−55.0
−57.5
−60.0
−62.5
−65.0
−67.5
−70.0
−72.5
5.00
5.25
5.50
5.75
6.00
6.25
6.50
6.75
7.00
Time [ms]


Contra Inh
Ipsi inh
200
150
100
50
0
−50
−100
−150
−200
−1.5
−1.0
−0.5
0.0
0.5
1.0
Best Δtex (µs)
Δtinh (ms)



Changing delays

Issues

- bio plausibility:
  - how exactly do these timings change?
  - limited evidence on plasticity of timing
- technical:
  - angle error is a one-dimensional measure
  - differences in arrival time of bilateral inhibition are two dimensions
  - what to do?