DDSP logbook

Initial 'toy' system

I have implemented a modular system in pytorch, with the following:

1. **FIRFilter1D**:
    - Implements an arbitrary phase FIR filter using a 1D convolutional layer.
    - The filter taps are learnable parameters.
2. **FIRFilter1DLinearPhaseI**:
    - Implements a linear phase FIR type I filter using 1D convolution.
    - Only allows an odd number of taps and enforces type I symmetry in the filter taps.
    - Taps are learnable parameters (only from the mid-index and up).
3. **GammaToneFilter**:
    - Impulse response of the filter is computed based on specified parameters like center frequency, sampling frequency, and impairment factor
        - !Impairing like this makes no sense just something I did for the toy example.
    - The filter is applied to the input signal using a 1D convolutional layer.
4. **GammaToneFilterbank**:
    - Uses the `GammaToneFilter` module to create individual filters and applies them to the input signal as a filterbank.
5. **NormalModel**:
    - Uses a `GammaToneFilterBank` to process the input with no pre-gain (FIR filter).
6. **ImpairedModel**:
    - Uses a `FIRFilter1D` to apply a learnable gain filter to process which then gets processed by `GammaToneFilterBank`.
    - The gain filter is implemented using `FIRFilter1D`, I have not tried using `FIRFilter1DLinearPhaseI`, although it should work with no issues.
7. **MyModel_v1**:
    - A trainable model composed of a normal hearing model and an impaired hearing model.
    - Uses the `NormalModel` and `ImpairedModel` to process the input and produce outputs for normal and impaired hearing, respectively.

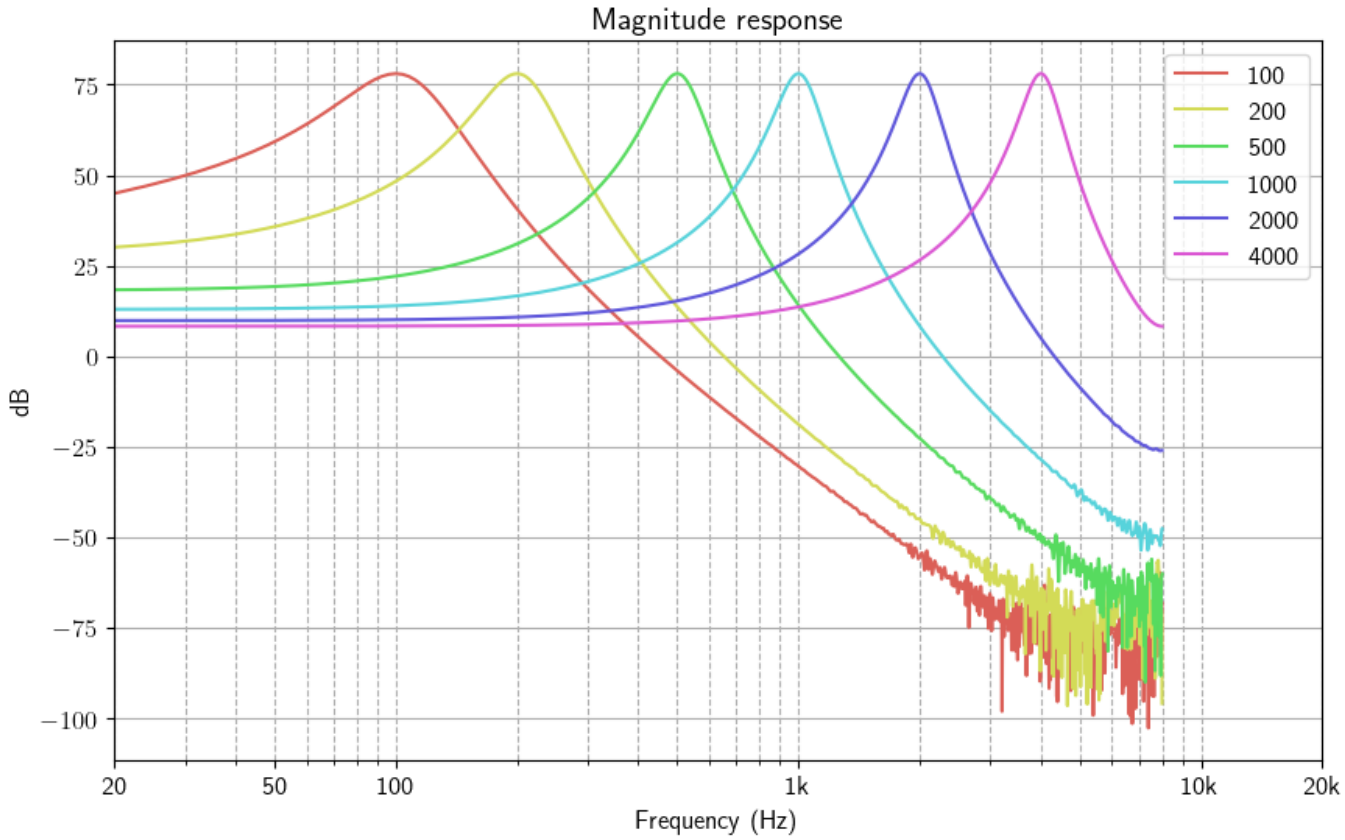
I have tried training a `NormalModel` using SGD with the SI-SDR loss between normal and impaired outputs as the loss function, which was successful. The results are presented under RESULTS below.

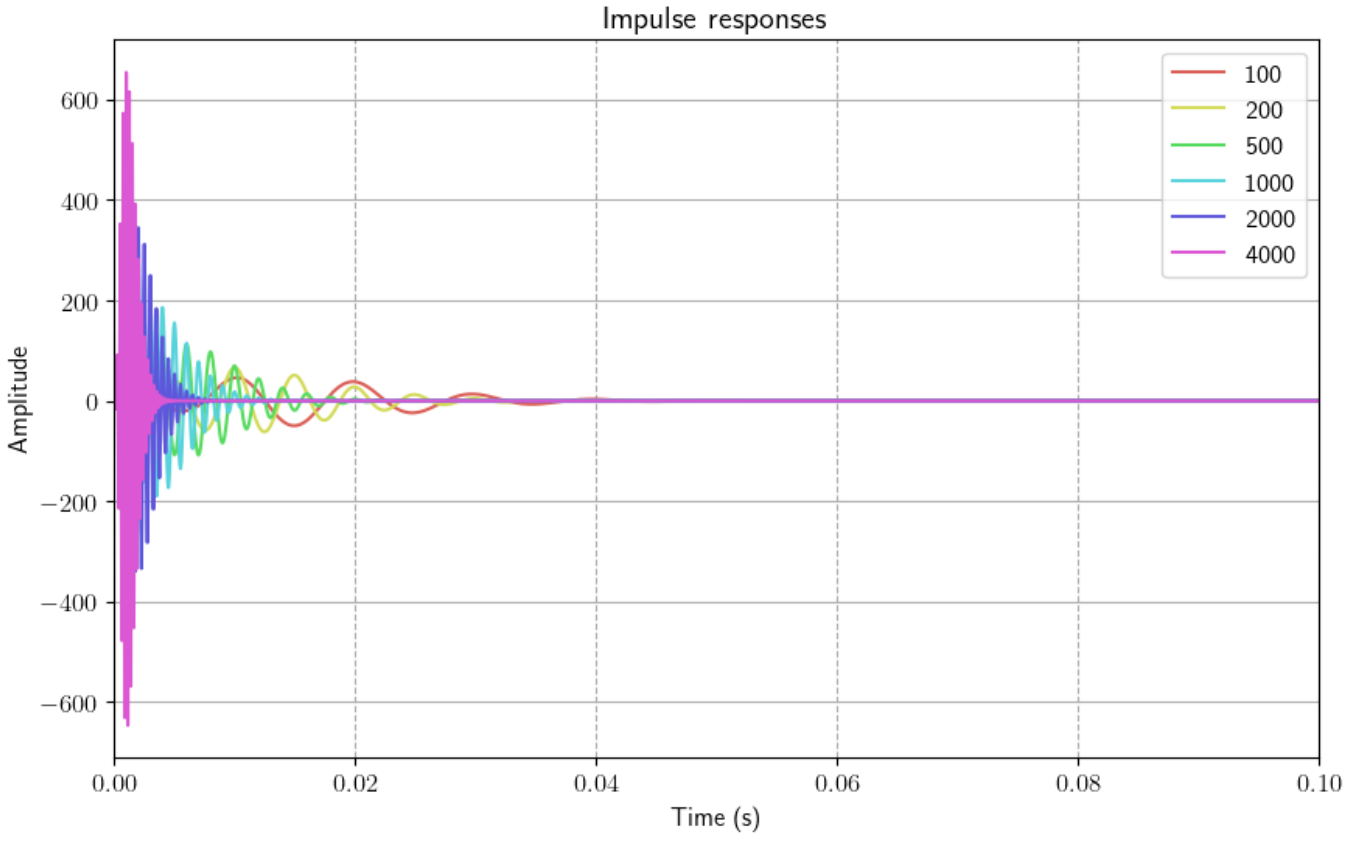
Next steps

- [ ] Try to mimic how the DHASP paper implemented the diffrentiable perceptual model.
- [ ] Try to mimic how the DHASP paper implemented the diffrentiable hearing-aid processing model.
- [ ] Introduce a perceptual loss to optimize the NN.

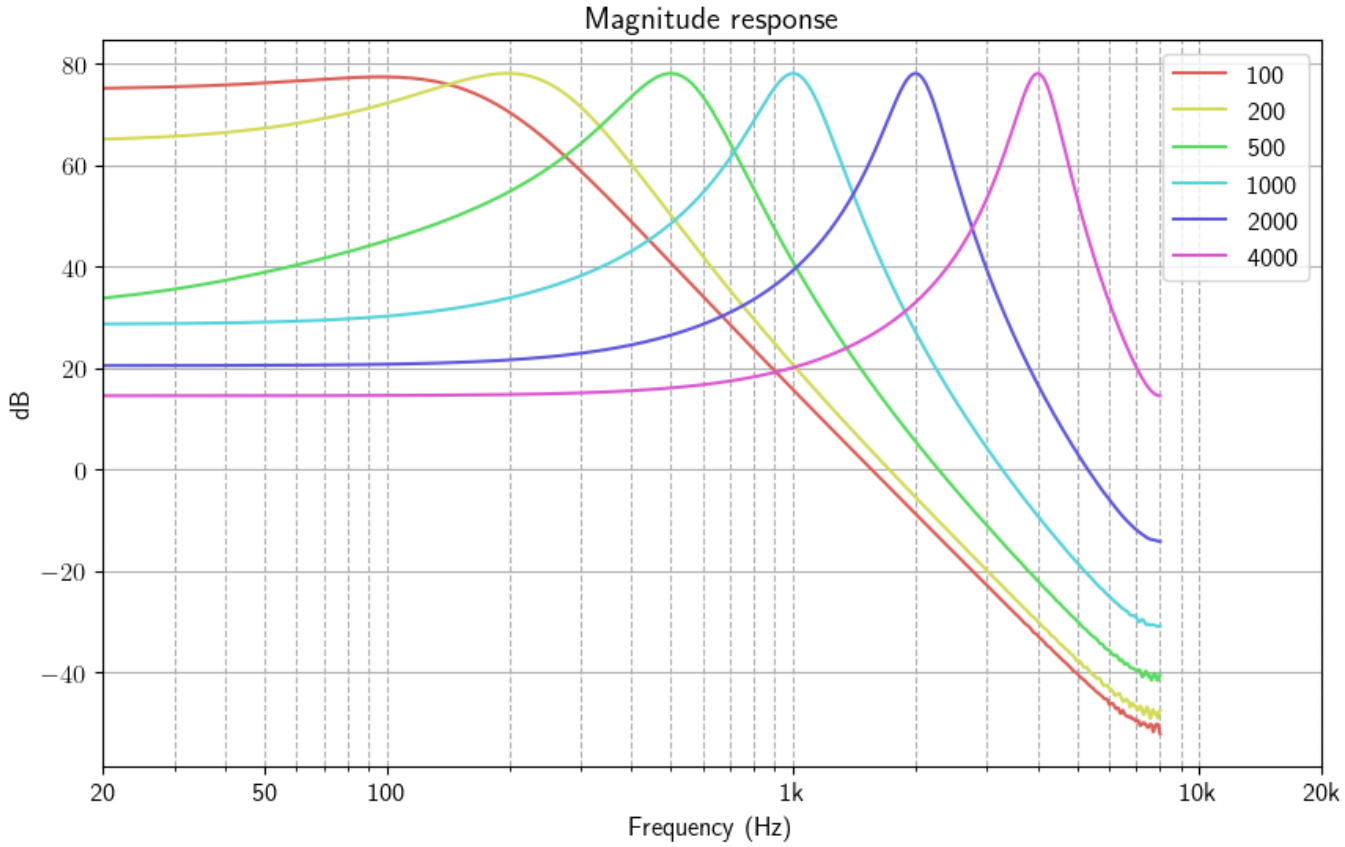
RESULTS

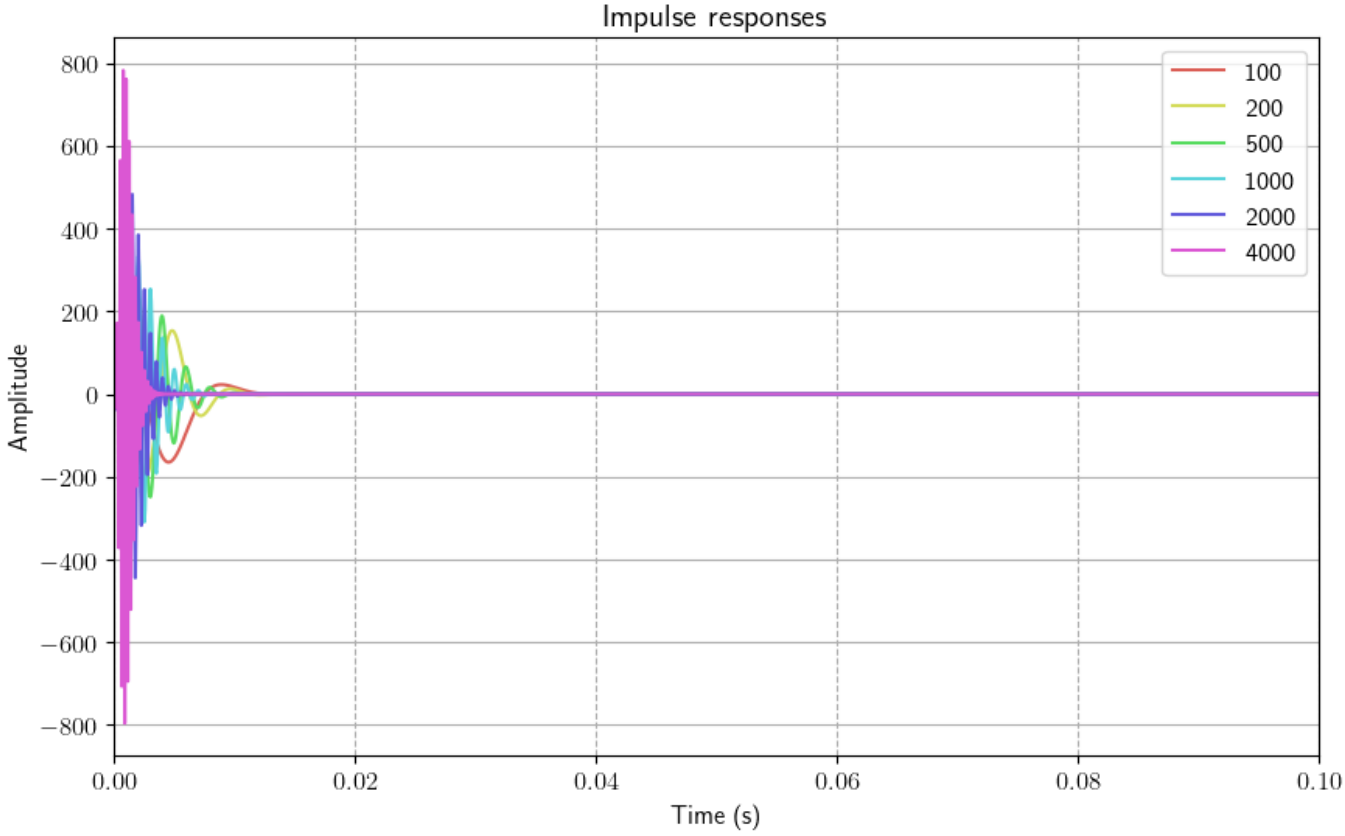
Normal Hearing filterbank used:



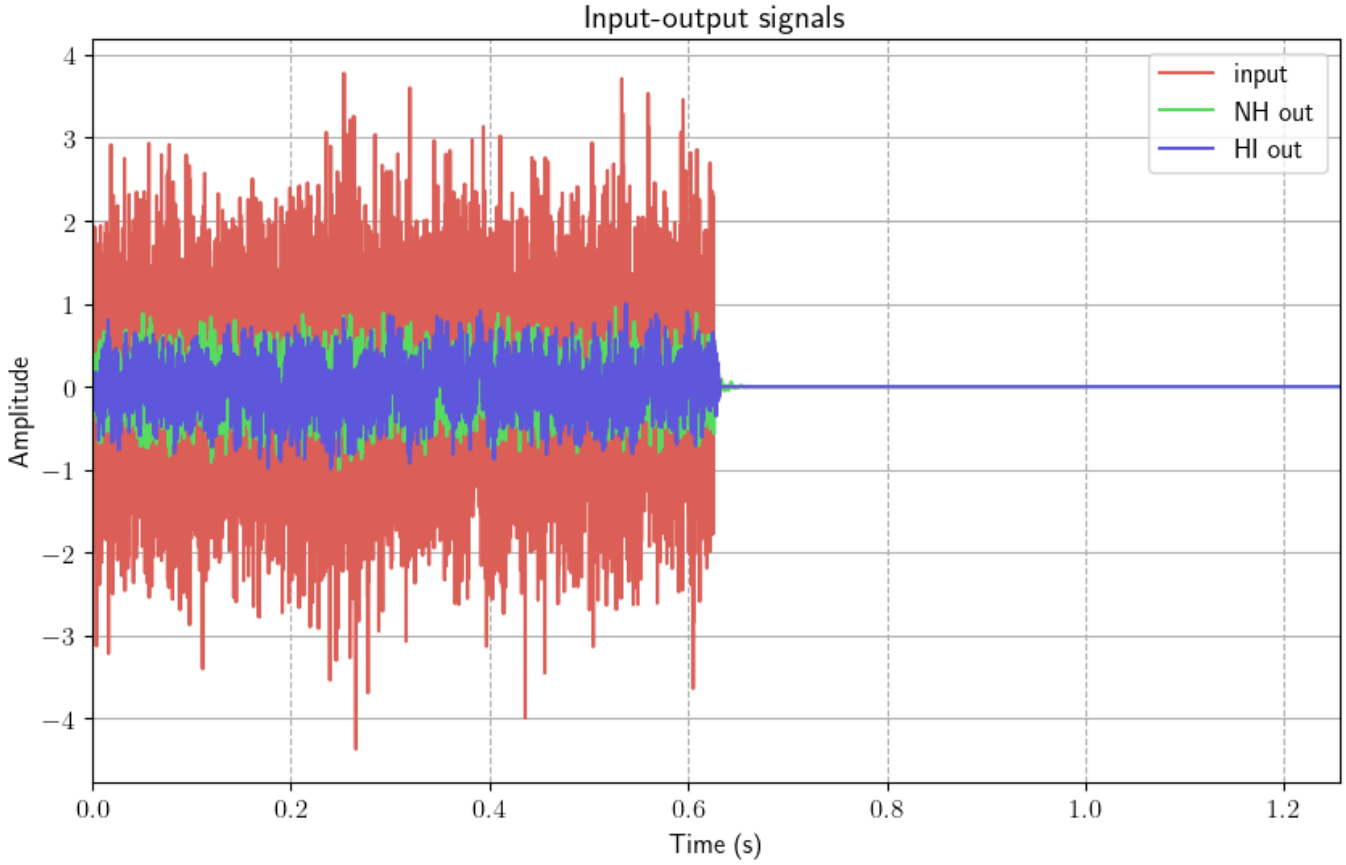


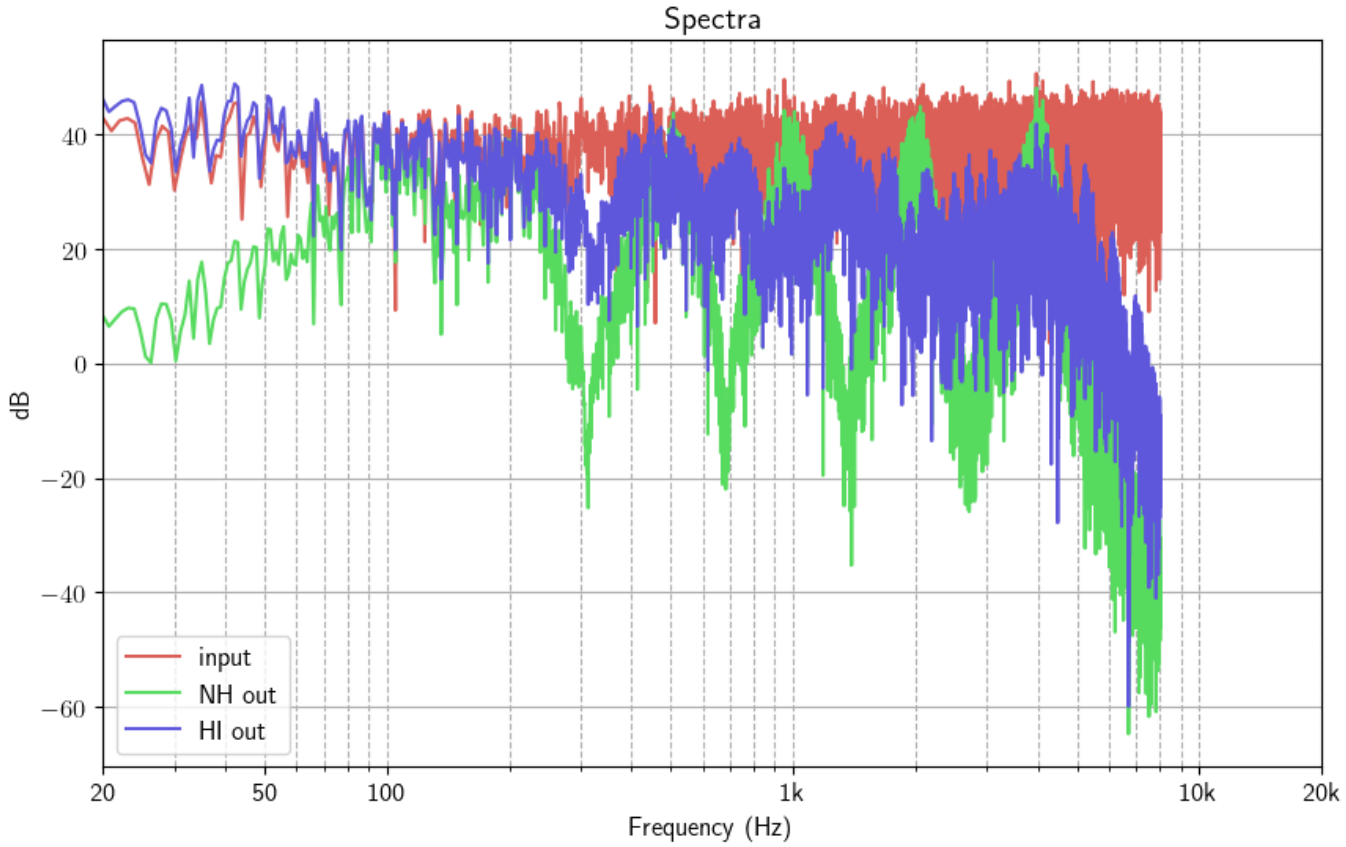
Impaired filterbank used:





Input-output of untrained model

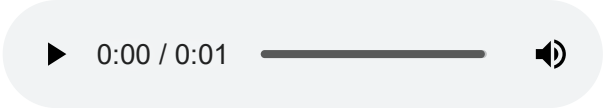




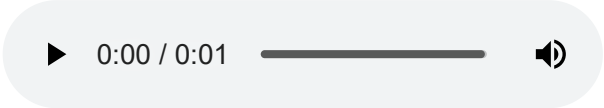
input

0:00 / 0:01

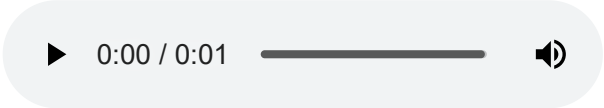
normal out (target)



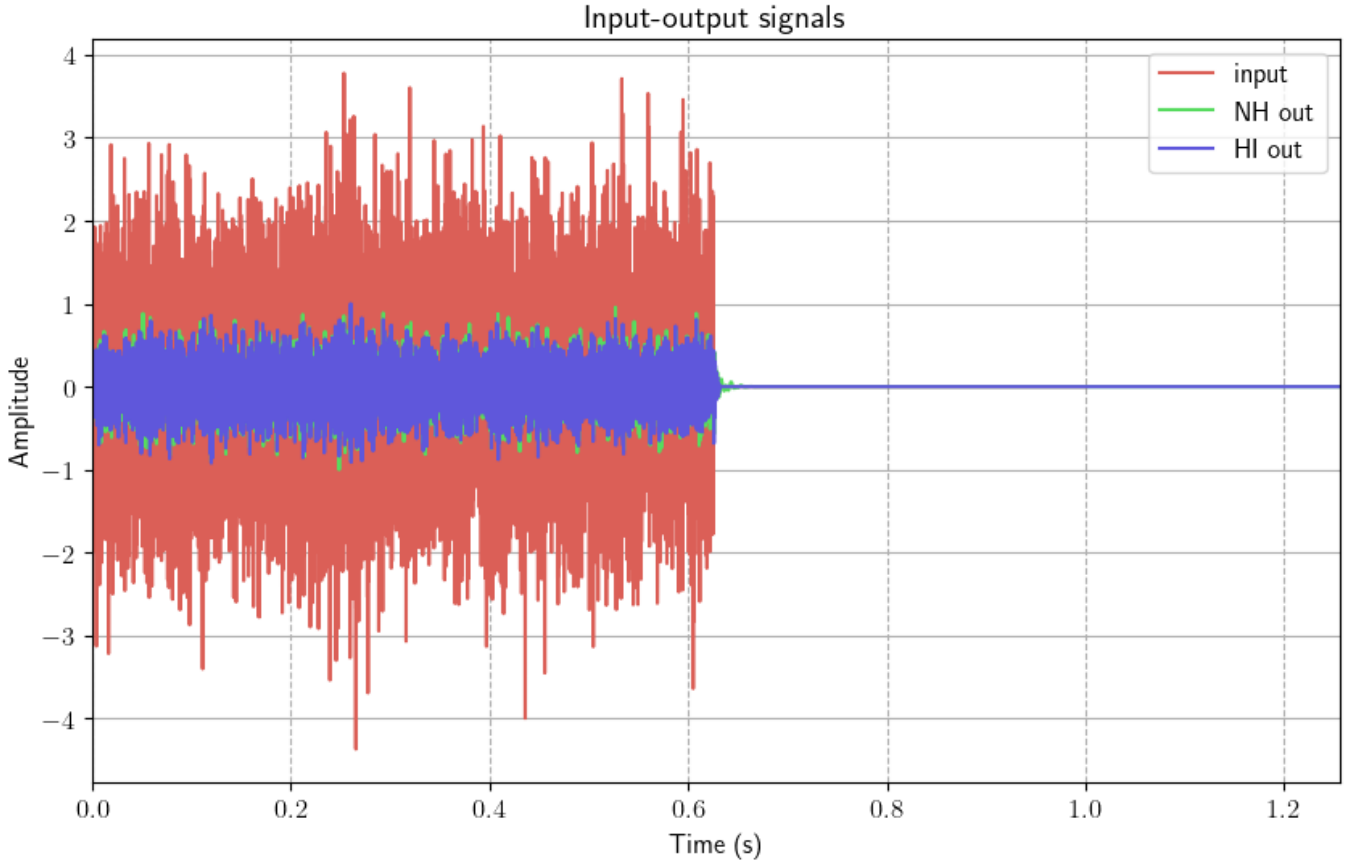
untrained impaired out

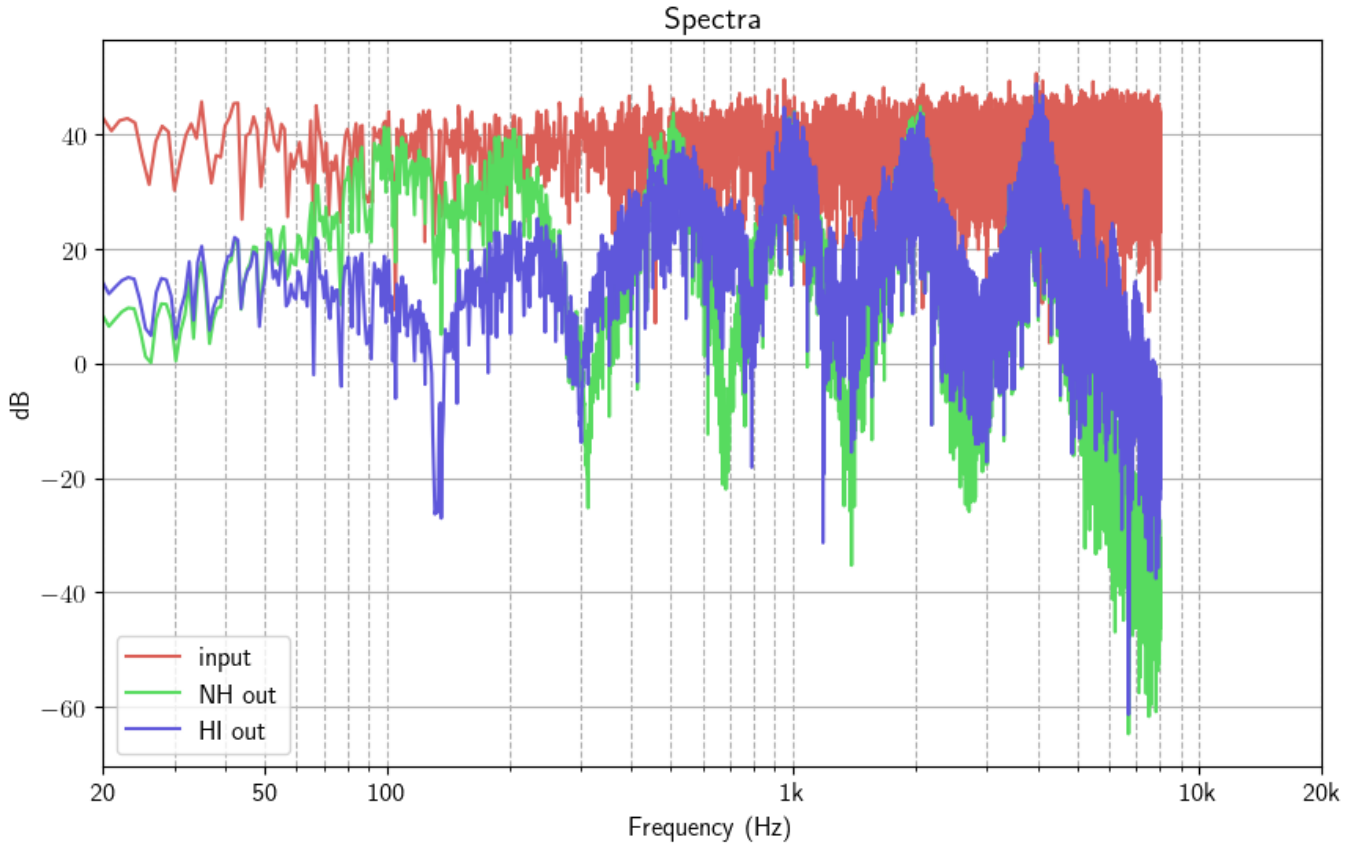


trained impaired out

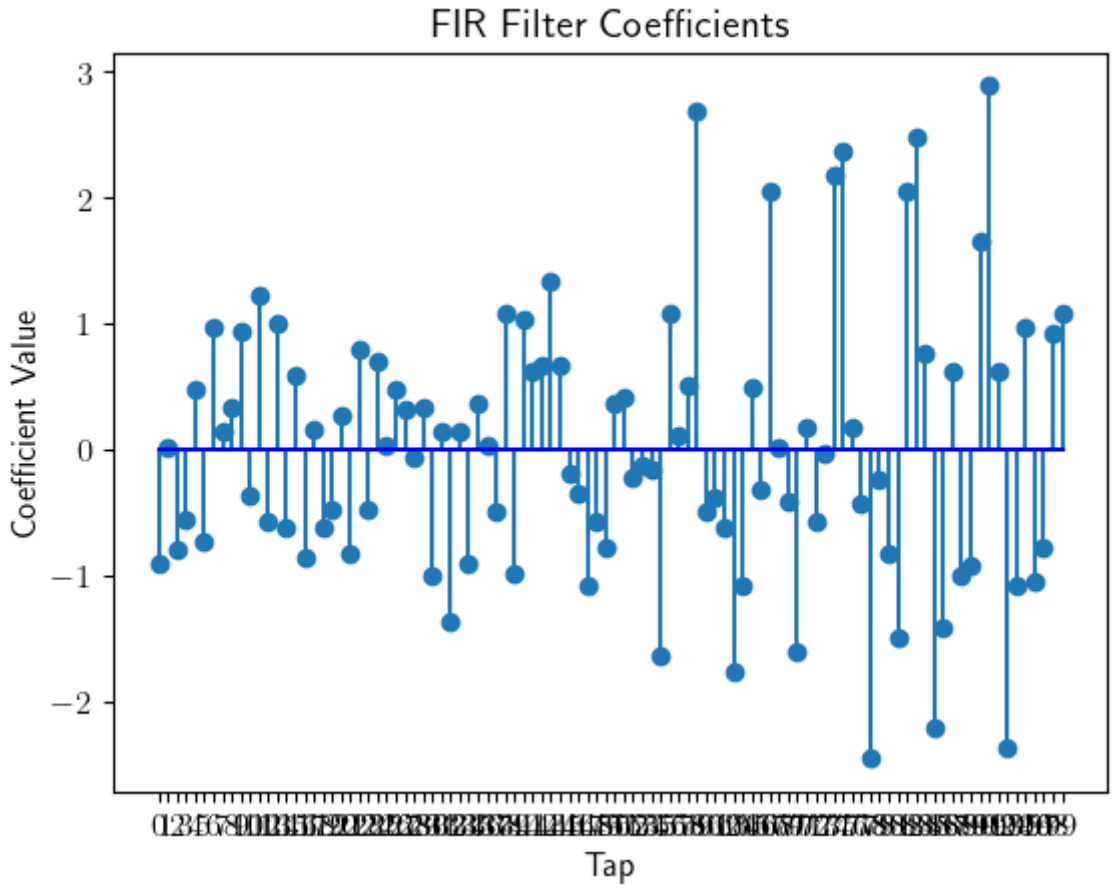


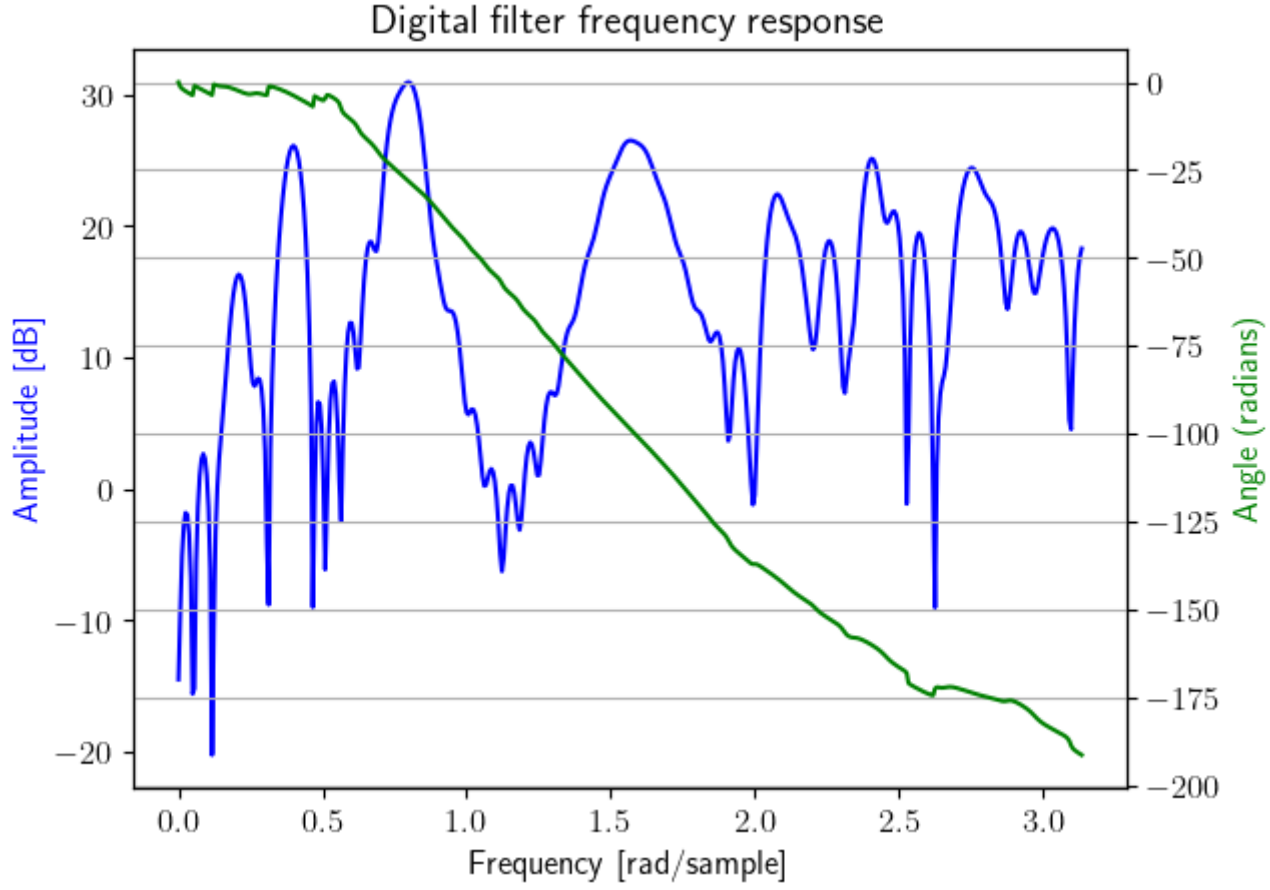
Input-output of trained model



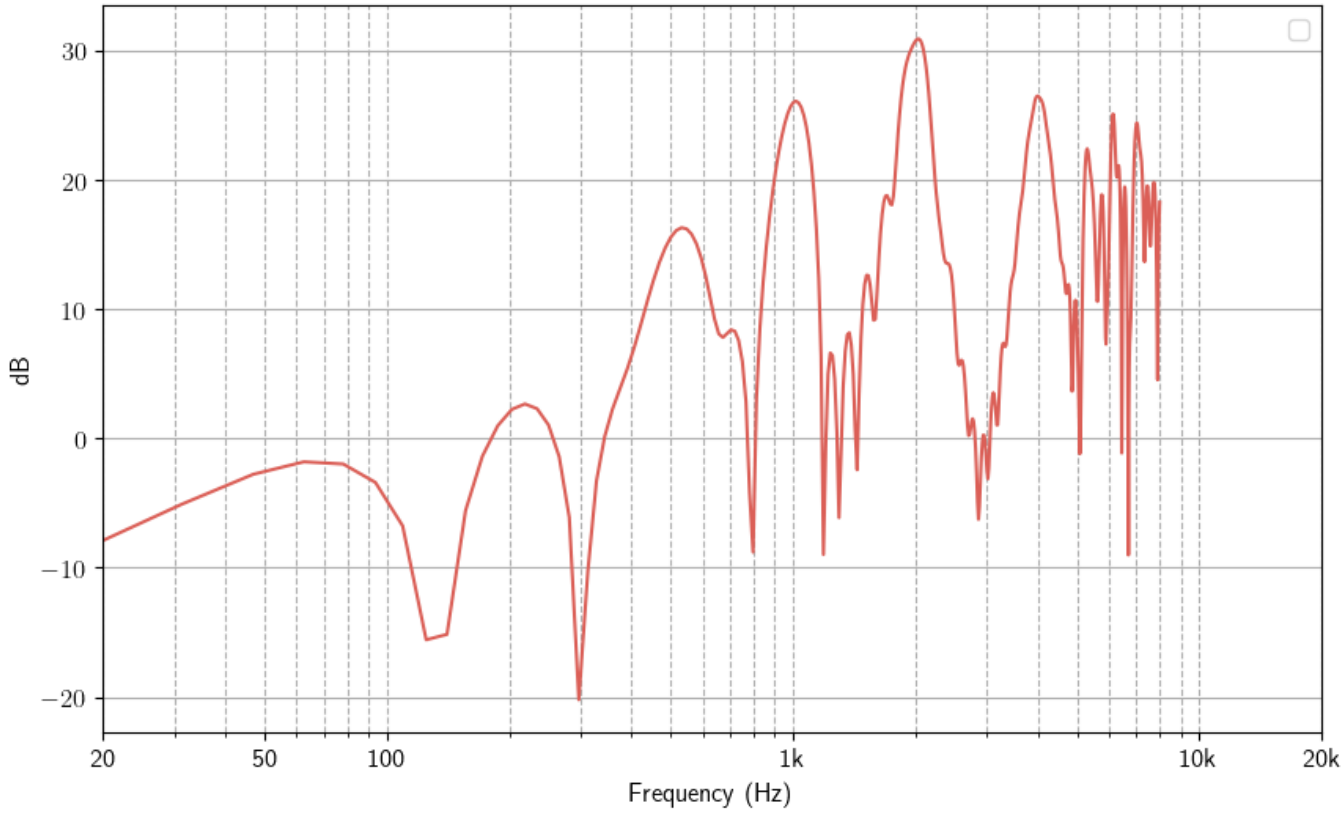


Learned filter





unnormalized frequency


30
20
10
0
−10
−20
dB
20
50
100
1k
10k
20k
Frequency (Hz)