Pattern Recognition Lab

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
FACULTY OF ENGINEERING

CLCNet: Deep learning-based noise reduction for hearing aids using Complex Linear Coding

H. Schröter, T. Rosenkranz, A. N. Escalante-B, M. Aubreville, A. Maier
Pattern Recognition Lab, Friedrich-Alexander University of Erlangen-Nürnberg
45th International Conference on Acoustics, Speech, and Signal Processing, May 4th - 8th 2020

Motivation: Aging society

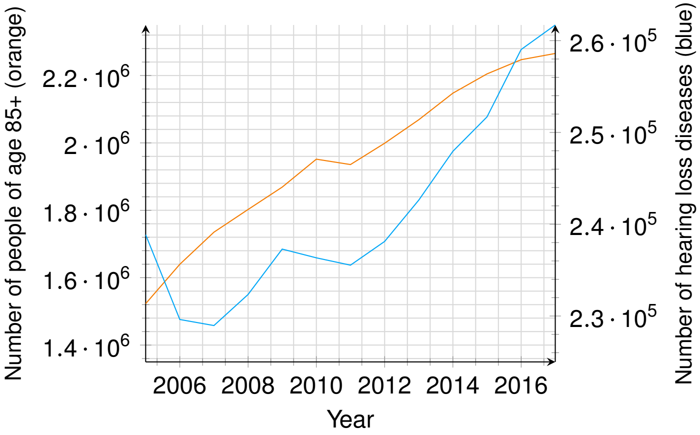
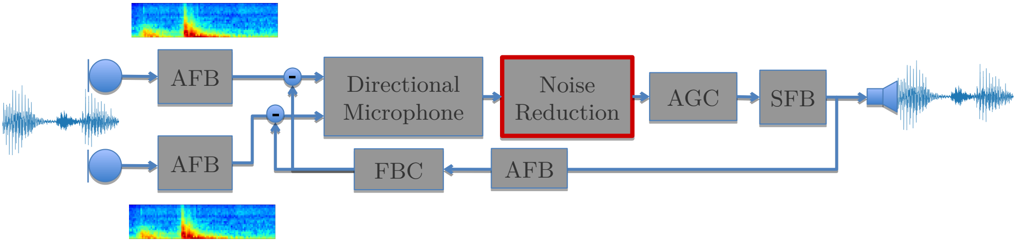


Figure: Correlation between hearing loss diseases and an aging society. [2]

Modern Hearing Aids

Figure: signia hearing aid [3]

Hearing Aid Pipeline



AFB: Analysis Filterbank
AGC: Automatic Gain Control
SFB: Synthesis Filterbank
FBC: Feedback Canceler

Figure: Typical hearing aid pipeline. [4]

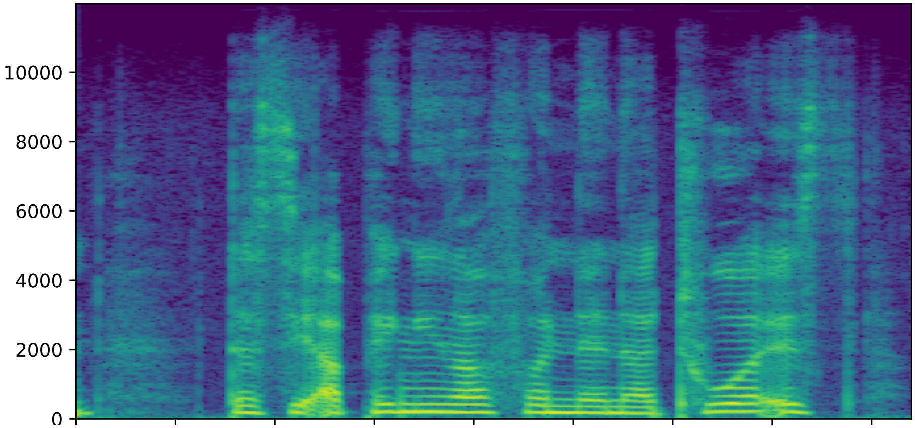
Modern Hearing Aids

Hearing aid constraints:

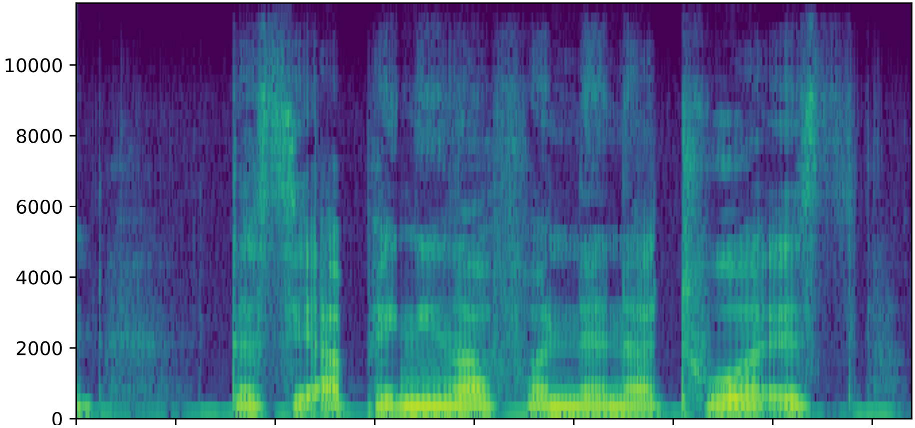
- On-line processing: Only limited temporal context in the future
- Limited frequency resolution (about 1kHz band width)
- (Limited computational power and power consumption)

Figure: signia hearing aid [3]

Narrowband power spectrogram of clean speech



Broadband filter bank representation



State of the art

Wiener Filter

Aubreville et al. [5] used DNN based Wiener Gain estimation.

The Wiener filter is defined as

$$\hat{y}_g[k] = g[k] * x[k]$$

where $\hat{y}_g$ is the predicted signal, g the Wiener filter and x the noisy signal. $*$ denotes the convolution operator.

In the frequency domain this can be expressed as

$$\hat{S}_{XY}[\nu] = G[\nu] \cdot S_{XX}[\nu] \,.$$

Wiener Filter

The ideal Wiener gains G can be estimated given the PSD S_{xx} of the noisy signal and the PSD S_{xy}

$$G[\nu] = \frac{S_{YY}[\nu]}{S_{XX}[\nu]} .$$

This assumes uncorrelated signals $S_{XY} = S_{YY}$ i.e. the additive noise model

$$S_{XX} = S_{NN} + S_{YY} .$$

Wiener Filter

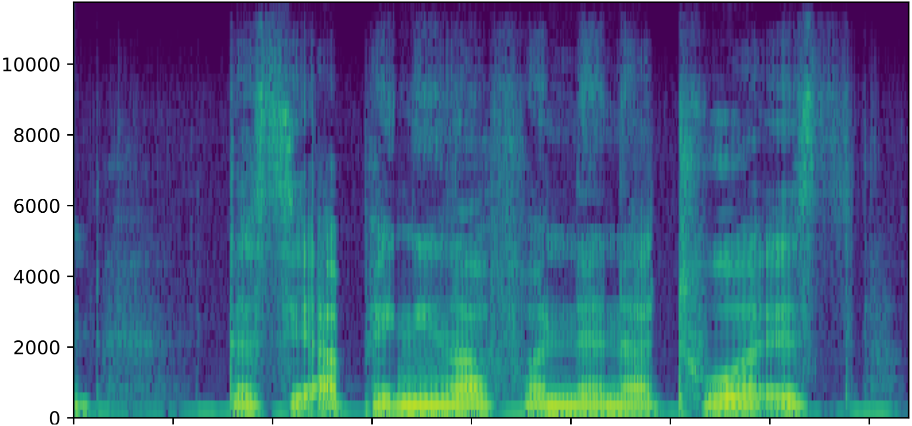


Figure: Filter bank representation of the clean speech

Wiener Filter

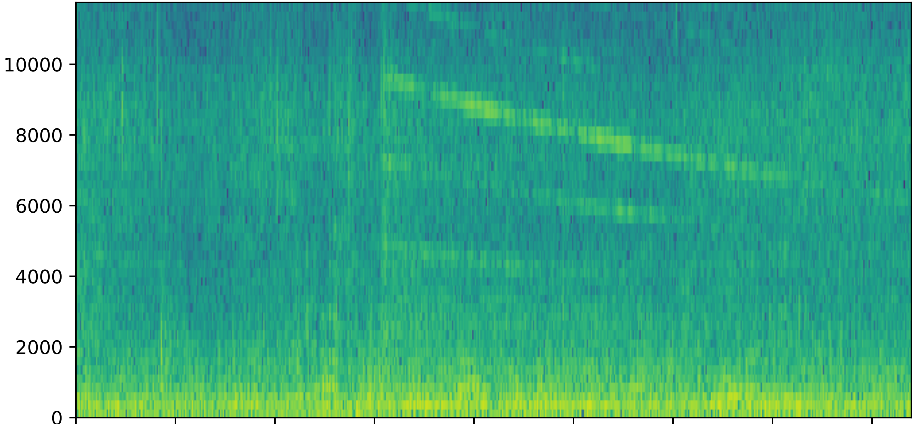
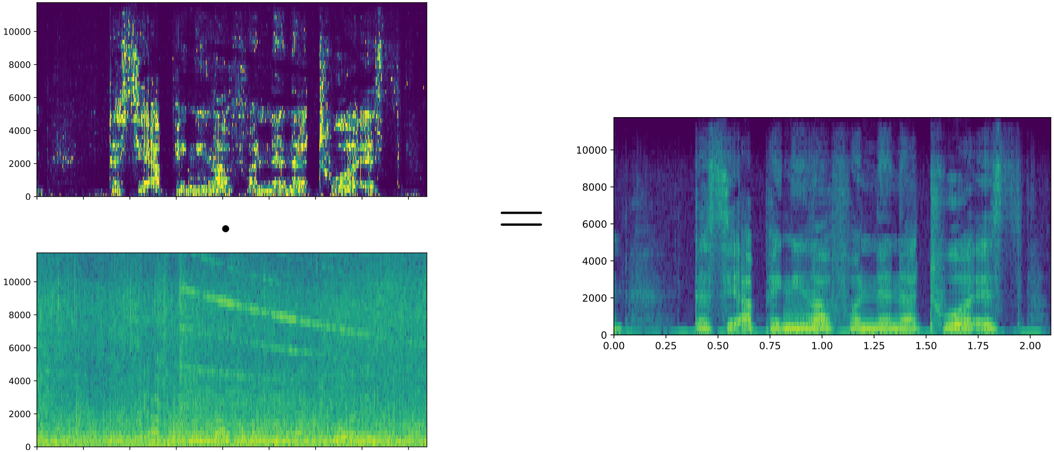


Figure: Filter bank representation of a noisy mixture

Wiener Filter: $\hat{S}_{XY}[\nu] = G[\nu] \cdot S_{XX}[\nu]$



Wiener Filter

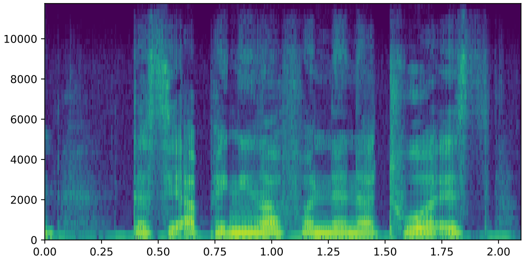


Figure: Original clean speech filter bank representation

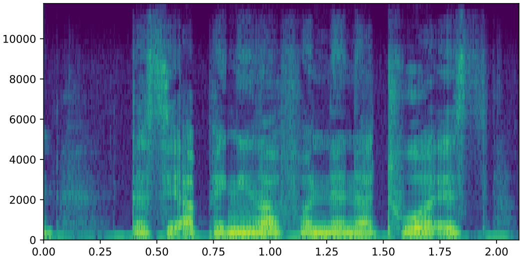


Figure: Enhanced noisy filter bank representation

Wiener Filter

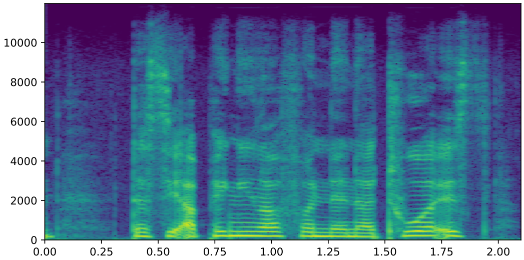


Figure: Original clean speech high resolution spectrogram

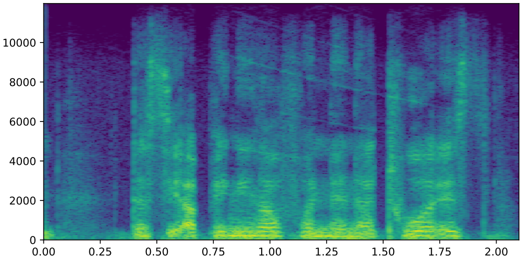


Figure: Enhanced noisy filter high resolution spectrogram

Linear Predictive Coding

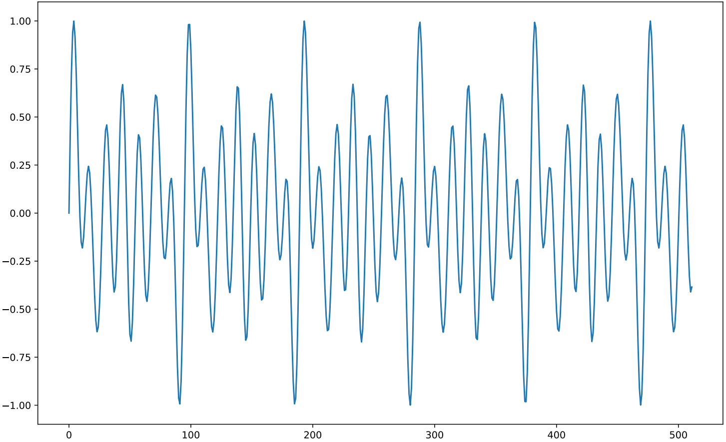


Figure: Synthezised *Bb3*

Linear Predictive Coding

Linear predictive coding (LPC) predicts the next time step k using a linear combination of the LPC coefficients a and the previous frames up to $k-1$.

$$\hat{x}_k = \sum_{i=1}^{N} a_i x_{k-i} \,, \tag{1}$$

Linear Predictive Coding

Linear predictive coding (LPC) predicts the next time step k using a linear combination of the LPC coefficients a and the previous frames up to $k-1$.

$$\hat{x}_k = \sum_{i=1}^{N} a_i x_{k-i} \,, \tag{1}$$

Let d_k be the prediction error

$$d_k = x_k - \hat{x}_k = x_k - \sum_{i=1}^{N} a_i x_{k-i} \,. \tag{2}$$

Linear Predictive Coding

Linear predictive coding (LPC) predicts the next time step k using a linear combination of the LPC coefficients a and the previous frames up to $k-1$.

$$\hat{x}_k = \sum_{i=1}^{N} a_i x_{k-i} \, , \tag{1}$$

Let d_k be the prediction error

$$d_k = x_k - \hat{x}_k = x_k - \sum_{i=1}^{N} a_i x_{k-i} \, . \tag{2}$$

$\rightarrow$ Closed-form solution for a_i via

- Autocorrelation of x_k
- Solve system of linear equations using the Levinson-Durbin algorithm

Linear Predictive Coding

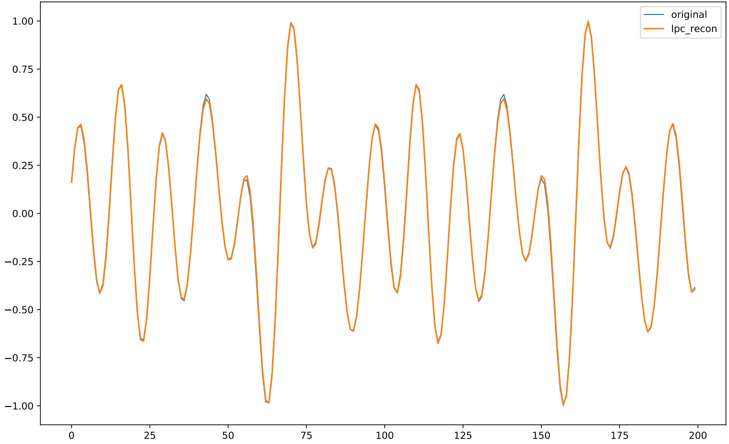


Figure: Reconstructed *Bb3*

Linear Predictive Coding

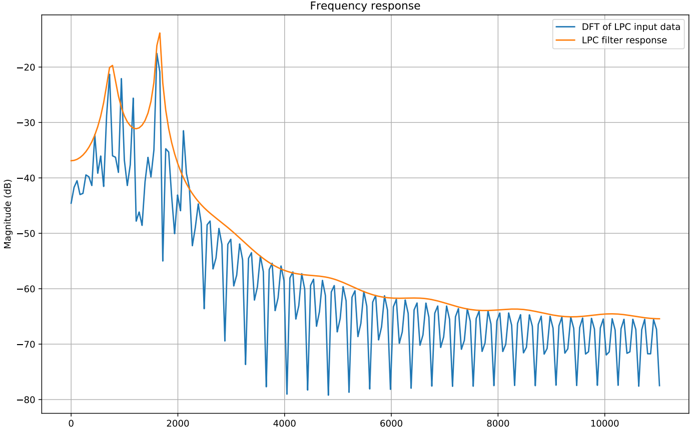


Figure: Frequency response of the LPC filter

Complex Spectrogram

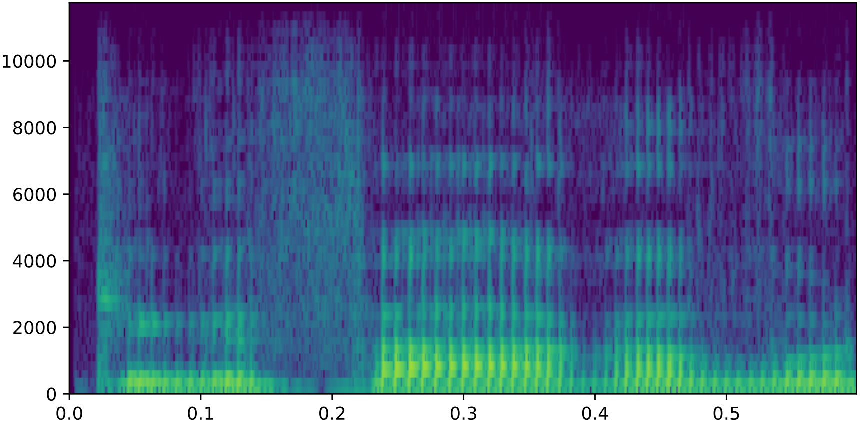
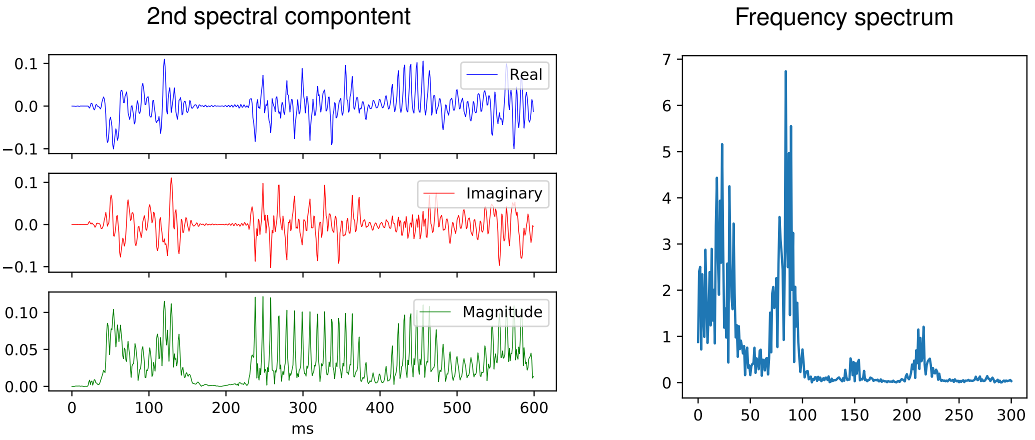


Figure: Filter bank representation.

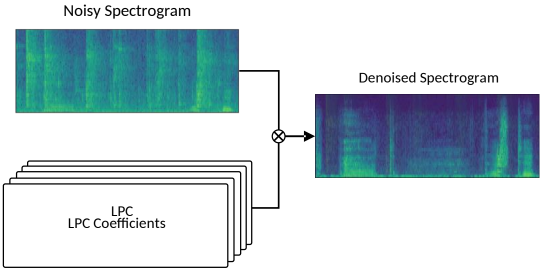
Spectral Component



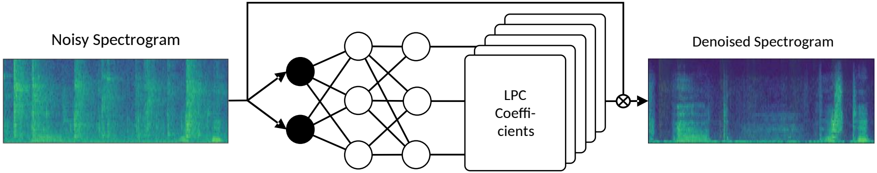
Linear Predictive Coding

For each frequency bin f at position k, predict the next time step based on the N previous frames of a noisy spectrogram X via LPC.

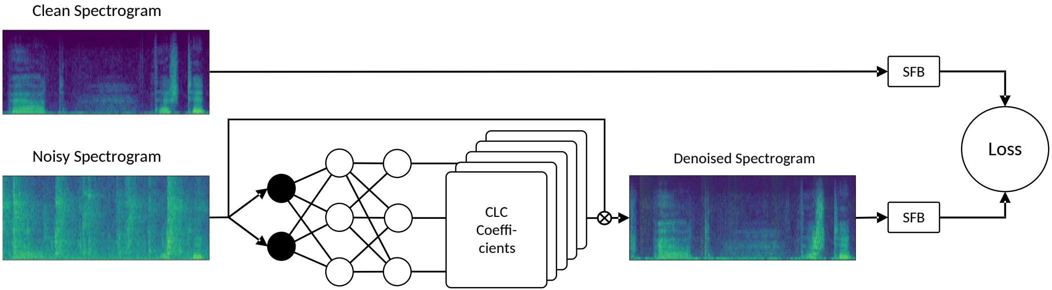
$$\hat{\boldsymbol{S}}(k,f) = \sum_{i=0}^{N} \boldsymbol{A}(k,i,f) \cdot \boldsymbol{X}(k-i-1,f)\ , \tag{3}$$



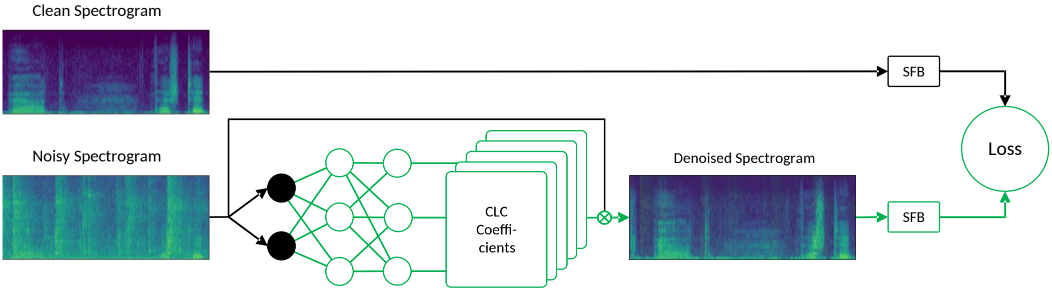
Training objective



CLCNet: Training



CLCNet: Training



CLCNet: Complex Linear Coding Network

Similar to LPC: CLC uses a complex valued linear combination to estimate a clean time-frequency bin.

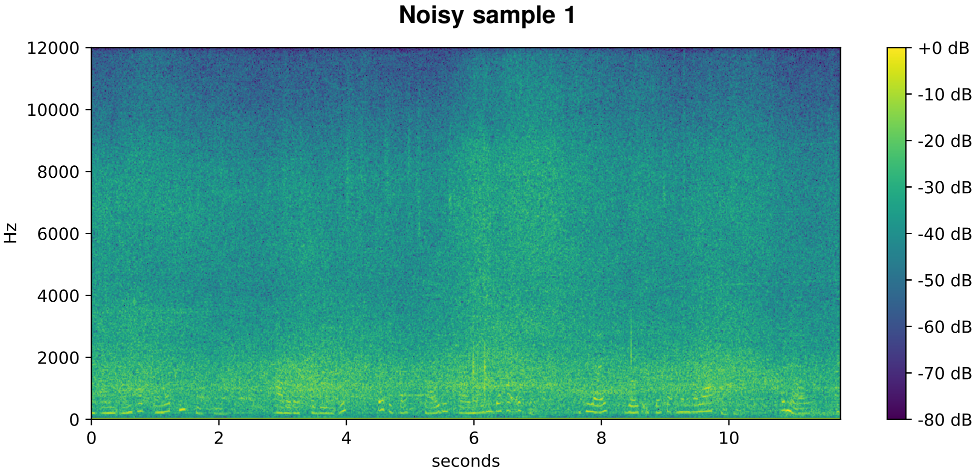
$$\hat{\boldsymbol{S}}(k,f) = \sum_{i=0}^{N} \boldsymbol{A}(k,i,f) \cdot \boldsymbol{X}(k-i+l,f) \;, \tag{4}$$

where,

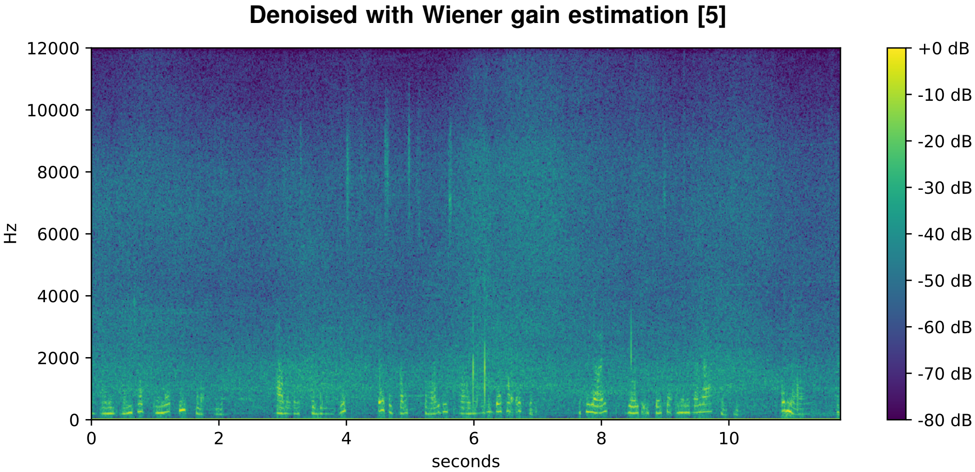
- k is the time-index
- f is the frequency-index
- N is the order of the linear combination
- $\boldsymbol{A}(k,i,f)$ are the output coefficients with $i = 0, \ldots, N$
- l is an offset

Results

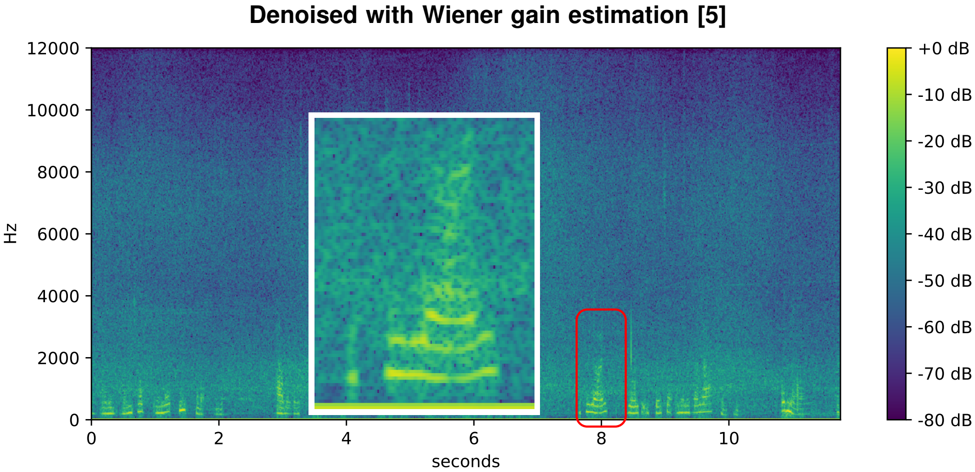
Complex Linear Coding: Example



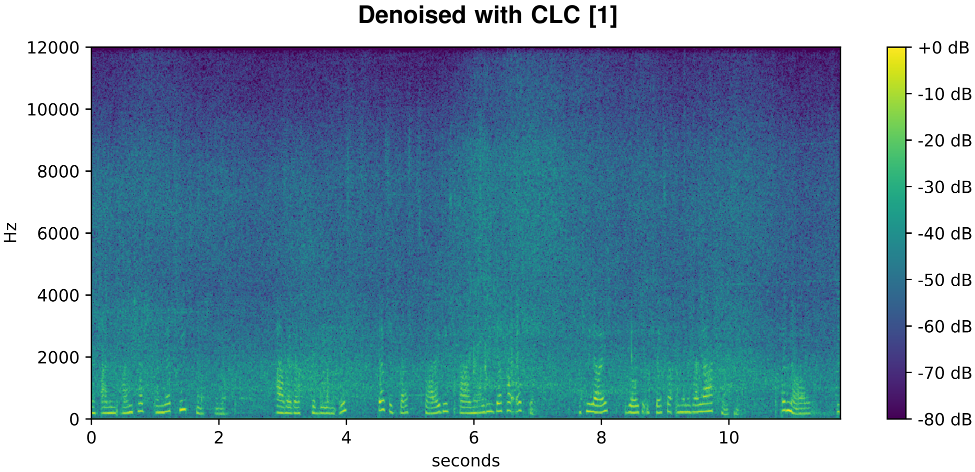
Complex Linear Coding: Example



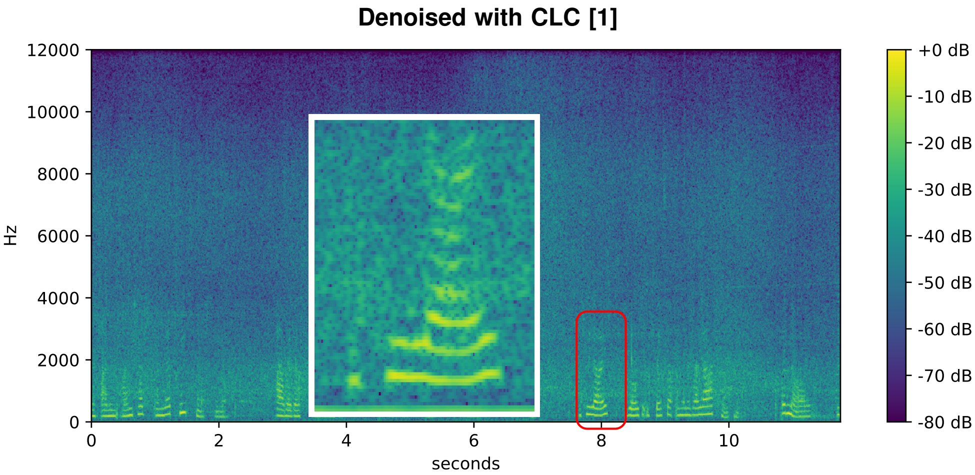
Complex Linear Coding: Example



Complex Linear Coding: Example

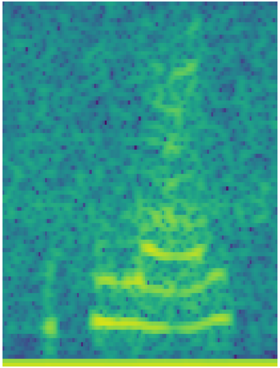


Complex Linear Coding: Example

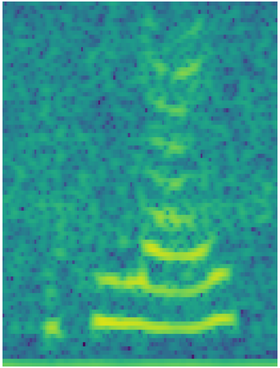


Complex Linear Coding - Wiener gains comparison: Example

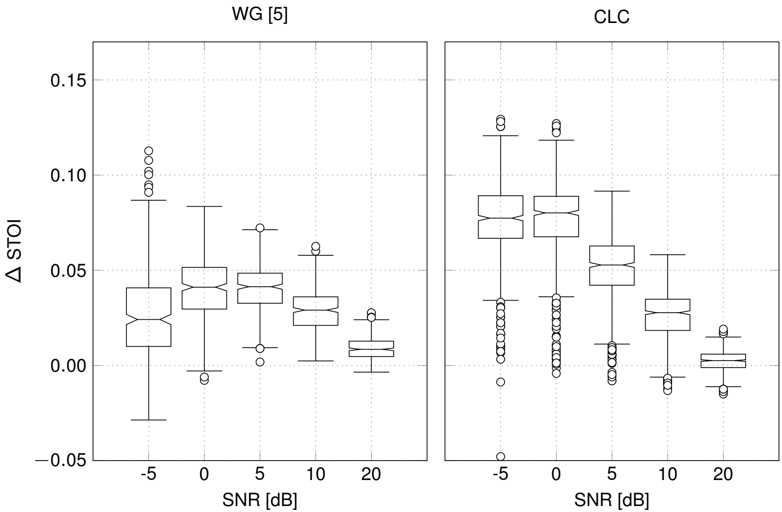
Wiener Gains



CLC



ΔSTOI metric comparison



Conclusion

- Complex Linear Coding: $\hat{\boldsymbol{S}}(k,f) = \sum_{i=0}^{N} \boldsymbol{A}(k,i,f) \cdot \boldsymbol{X}(k-i+l,f)$
- Indirectly estimate the phase
- Which makes it possible to remove noise between harmonics

Thank you for your attention

More samples and preprint at

https://rikorose.github.io/CLCNet-audio-samples.github.io/

References I

[1] H. Schröter, T. Rosenkranz, A. Escalante B., M. Aubreville, and A. Maier, “CLCNet: Deep learning-based Noise Reduction for Hearing Aids using Complex Linear Coding”, arXiv preprint arXiv:2001.10218 (2020).

[2] Statistisches Bundesamt Deutschland, *Nebendiagnosen der vollstationären Patienten (Code 23141-0001)*, 2016.

[3] Sivantos GmbH, *Signia Hearing aids*, (July 2019) `https://www.signia-hoergeraete.de/`.

[4] Ehrensperger, Kai, “Deep Learning-based Noise Reduction for Hearing Instrument Applications”, MA thesis (Friedrich-Alexander University Erlangen-Nürnberg", 2018).

[5] M. Aubreville, K. Ehrensperger, A Maier, T. Rosenkranz, B. Graf, and H. Puder, “Deep denoising for hearing aid applications”, in 2018 16th international workshop on acoustic signal enhancement (iwaenc) (IEEE, 2018), pp. 361–365.