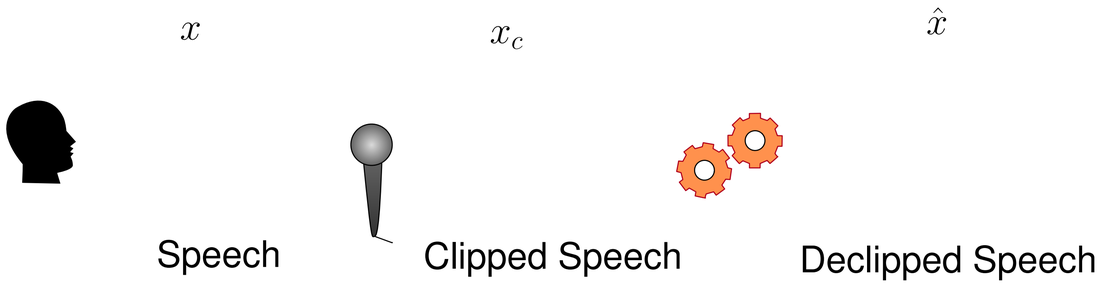
INTERNATIONAL AUDIO LABORATORIES ERLANGEN
A joint institution of Fraunhofer IIS and Universität Erlangen-Nürnberg

Declipping Speech Using Deep Filtering

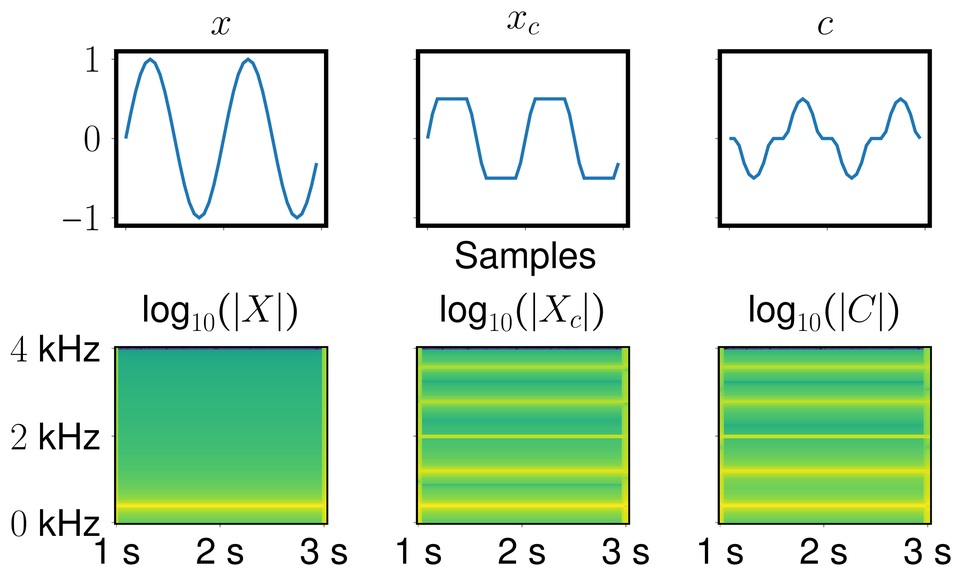
Wolfgang Mack, Emanuël A. P. Habets

1. Introduction



2. Problem Formulation

- Clipping reduces the overall energy.
- ... introduces higher-order harmonics.



- Declipping requires amplification of the fundamental frequency.
- ... attenuation of the harmonics introduced by clipping.
- We formulate declipping as an extraction problem,

$$X_\mathrm{c} = X + C,$$

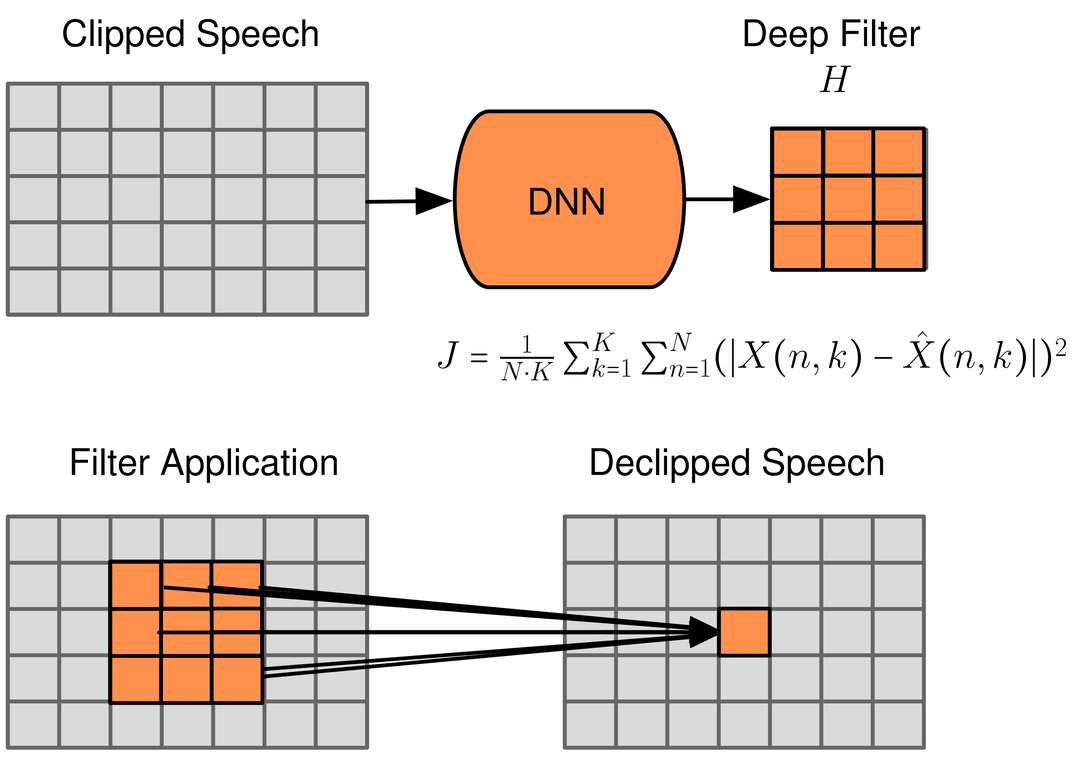
where X is to be extracted from X_c.

3. Proposed Method [1]

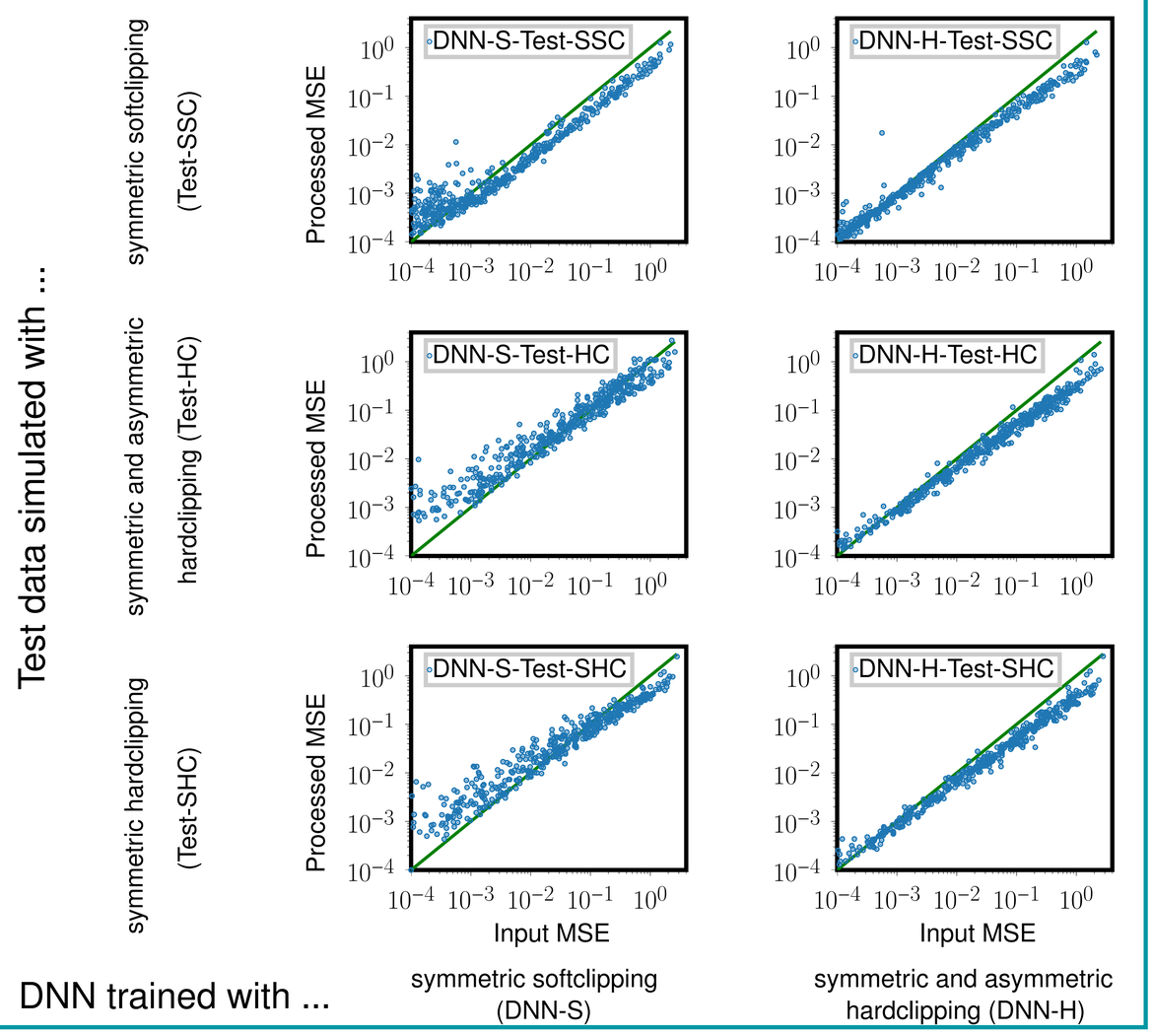
- The declipped speech estimate $\hat{X}$ is obtained from X_c by

$$\hat{X}(n,k) = \sum_{i=-I}^{I} \sum_{l=-L}^{L} H^{\star}_{n,k}(l+L, i+I) \cdot X_\mathrm{c}(n-l, k-i),$$

where L and I define the filter spread in time and frequency.

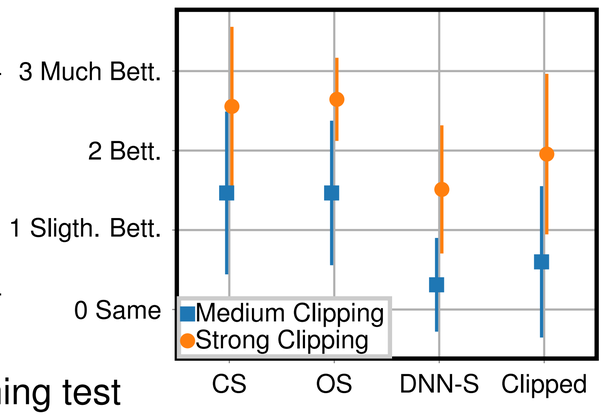


4. Performance Evaluation - Simulated Clipped Speech



5. Performance Evaluation - Measured Clipped Speech

	Medium Clipping		Strong Clipping	
	p (‰)	Effect size	p (‰)	Effect size
OS	3.3	.50	0.2	.65
CS	0.4	.61	0.2	.64
DNN-S	6.5	.45	0.3	.63
Clipped	4.9	.47	0.3	.63



- DNN-H is compared to all other Methods with a preference listening test using five seconds long measured clipped speech from 2 males and 1 female.
- The evaluation is performed using a Wilcoxon signed-rank test (15 participants).

6. Conclusions

Artificial and real clipped speech can be declipped with a multi-dimensional complex STFT filter (deep filter), which is applied to the clipped speech STFT. The deep filter is obtained by a DNN optimized to minimize the reconstruction MSE.

Audio examples are available on
`https://www.audiolabs-erlangen.de/resources/2019-WASPAA-Declipping`

[1] W. Mack and E. A. P. Habets, Deep filtering: Signal extraction using complex time-frequency filters, https://arxiv.org/abs/1904.08369, Apr. 19.

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
Fraunhofer IIS