NONLINEARITY DETECTION AND COMPENSATION FOR EEG-BASED SPEECH TRACKING

Johanna Wilroth, Emina Alickovic, Martin A. Skoglund, Martin Enqvist.

Automatic Control, Linköping University, Linköping, Sweden
{firstname.lastname}@liu.se

ABSTRACT

Clusters of neurons generate electrical signals which propagate in all directions through brain tissue, skull, and scalp of different conductivity. Measuring these signals with electroencephalography (EEG) sensors placed on the scalp results in noisy data. This can have severe impact on estimation, such as, source localization and temporal response functions (TRFs). We hypothesize that some of the noise is due to a Wiener-structured signal propagation with both linear and nonlinear components. We have developed a simple nonlinearity detection and compensation method for EEG data analysis and utilize a model for estimating source-level (SL) TRFs for evaluation. Our results indicate that the nonlinearity compensation method produce more precise and synchronized SL TRFs compared to the original EEG data.

Index Terms— Auditory Processing, EEG, Nonlinearity Compensation, Temporal Response Function, Source Localization

1. INTRODUCTION

Nonlinear science has been humorously likened to studying all animals except elephants [1]. One interpretation is that linearity is an important exception to the rule of nonlinear natural phenomena. Another interpretation is that finding the correct nonlinear system which describes the world perfectly is extremely difficult due to the uncountable combinations of functions and variables. A prime example of such complexity is the human brain, composed of about 100 billion neurons constantly interacting to manage daily tasks, yet with puzzling underlying dynamics [2]. Notably, studies reveal that electroencephalography (EEG) signals—arising from electrodes on the scalp and measures neuronal cluster electrical activity—are also nonlinear, non-stationary, high-dimensional, and exceedingly noisy [3, 4, 5, 6, 7].

The authors gratefully acknowledge the financial support from the *Excellence Center at Linköping – Lund in Information Technology (ELLIIT)*, project *Brain-Based Monitoring of Sound*, and from the *William Demant Foundation*. The authors would like to thank Joshua P. Kulasingham and Oskar Keding for their helpful discussions of the methods. EA and MAS are also with Eriksholm Research Centre, Snekkersten, Denmark (e-mail: {eali, mnsk}@eriksholm.com)

Neural speech tracking involves predicting EEG responses from speech features using linear filters called Temporal Response Functions (TRFs) [8]. These TRFs, typically estimated at the sensor level, predict EEG responses and exhibit notable peaks at specific time intervals (e.g., around 50-70 ms for P1 and 100-180 ms for N1). The Neuro-Current Response Function (NCRF) model, a one-step approach developed for higher signal-to-noise ratio (SNR) magnetoencephalography (MEG) data, extends the concept to estimating TRFs at source level [9]. However, NCRFs face certain limitations in accurately accounting for complex nonlinear behaviors within EEG channels.

This study focuses on a novel nonlinearity detection and compensation method for EEG data analysis and utilizes the NCRF model [10] for evaluation. We explore nonlinear effects in estimating sources and source-level TRFs (SL TRFs) from EEG data of a listener with hearing impairment in a cocktail party-like environment. To address the complexity of modeling nonlinear brain behaviors, we present a binning-based method to counter various nonlinear behaviors in the EEG sensors.

The results support our hypothesis of Wiener-structured signal propagation, i.e., a static nonlinearity on output. Moreover, we evaluate the precision in estimating a non-parametric nonlinearity, revealing that the number of bins used per sensor indicates the sensor-specific degree of nonlinearity. Building upon the insights gained from the binning method, we address some of these sensor-specific nonlinear behaviors. Here, the prime emphasis is on employing the NCRF model to evaluate the effectiveness of our compensation approach. Furthermore, We analyze how the nonlinearities are distributed over the scalp. The study's findings indicate that our compensated EEG data yield notably more precise and synchronized SL TRFs compared to the original EEG data.

2. METHOD

The study's workflow is presented in Fig. 1. It shows the original EEG signal $\mathbf{Y}$, serving as an input to the NCRF model (briefly discussed in Section 2.1), with subsequent evaluation of SL TRF results. Following this, the predicted EEG $\hat{\mathbf{Y}}$ from the model is compared to the original EEG signal $\mathbf{Y}$ within the nonlinear compensation method. This aims to ad-

dress nonlinear components within the data and is outlined in Section 2.2. The resulting compensated output $\mathbf{Y}_{new}$ is then subsequently fed back into the NCRF model, facilitating the evaluation of SL TRFs.

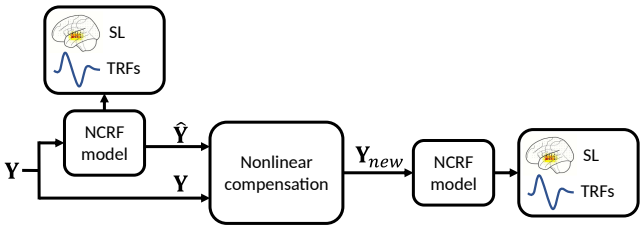

Fig. 1: Block diagram of the workflow. The original EEG data $\mathbf{Y}$ enters the NCRF-model and SL TRFs are evaluated by visual inspection. The predicted EEG data $\hat{\mathbf{Y}}$ from the NCRF-model is compared to $\mathbf{Y}$ in the nonlinear compensation method and $\mathbf{Y}_{new}$ is produced after compensation of nonlinear components. SL TRF estimations from the NCRF-model are evaluated.

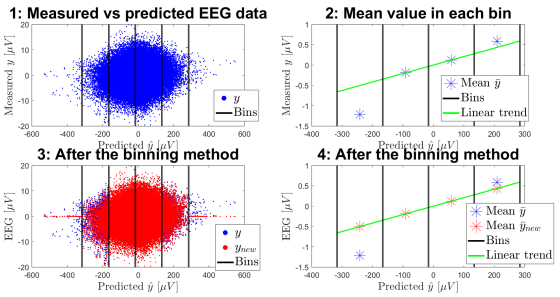

Fig. 2: Illustration of the nonlinear compensation method for one channel with four bins. Top left: Measured EEG $\mathbf{Y}$ and predicted EEG $\hat{\mathbf{Y}}$ after the NCRF-model divided into four bins. Top right: The mean in each bin and a line fitted between the means in the two middle bins. Bottom left: New EEG data $\mathbf{Y}_{new}$ after the datapoints in the outer bins have been compensated such that the new mean lies on the line (bottom right).

2.1. The NCRF-model

Initially, the analysis of SL TRFs using the NCRF model was introduced for MEG data collected from individuals instructed to listen to a single continuous speech stream [9]. Our preliminary findings, as outlined in [10], indicate that the NCRF model may also be applicable to EEG data, even when the SNR is considerably lower compared to MEG data. The NCRF model solves a nonlinear regression task by iteratively solving two linear systems, see [9].

2.2. The binning-based nonlinear compensation method

Ideally, the predicted EEG $\hat{\mathbf{Y}}$ using the linear SL TRFs should be correlated with the measured EEG data $\mathbf{Y}$. However, the top left panel of Fig. 2 shows a representative sensor without any apparent correlation. As an example of the method, we partitioned the data into equally spaced bins, excluding the first and last 100 samples with the largest and smallest $\hat{\mathbf{Y}}$, respectively. The mean in each bin was computed, as demonstrated in the top right panel. Assuming that the middle bins encompass the majority of data points and are consequently less susceptible to noise, we fitted a line between the mean values of these two bins. Notably, the mean values of the outer bins deviating from the green line indicate the presence of nonlinearities.

Our focus lies in compensating for the deviations. Each sample in each outer bin is weighted with the value of the line at the center of each bin divided by the mean of the bin. The outcomes of this compensation are displayed in the bottom panels of Fig. 2, where red datapoints denote the new EEG data $\mathbf{Y}_{new}$. SL TRF estimates were compared with those derived from the original EEG data $\mathbf{Y}$. Pseudo-code outlining the nonlinear compensation method is provided in Algorithm 1.

The measured and predicted EEG data were divided into equally sized estimation and validation sets. The nonlinear compensation method was applied to each sensor, and the mean values of the two sets were contrasted.

Given our hypothesis that nonlinear components could differ across sensors, we investigated how different number of bins effect the results. The Euclidean distance between the estimation and the validation datasets was computed for each number of bins and for each sensor. Subsequently, the number of bins yielding the smallest norm for each sensor was identified. This approach leads to a sensor-specific nonlinear compensation, as the optimal number of bins may vary between sensors.

3. EXPERIMENTAL DATASET

This section explains the experimental paradigm and data processing of the EEG and audio data.

3.1. Experimental paradigm

A subject with mild-to-moderate hearing loss was instructed to selectively attend to one out of the two concurrent frontal speech streams ($\pm 30°$), while 16-talker babble noise played in the background (+3 dB SNR). The subject was wearing hearing aids with noise reduction algorithms switched on, aiming at reducing the impact of the background babble noise. Each trial consisted of 33 s, preceded by 5 s of the background babble noise before the onset of the frontal speakers and EEG data was recorded with a 64-channel BioSemi ActiveTwo system at a sampling rate of 1024 Hz. This dataset has been used for different analysis and a complete description of the

Algorithm 1 The nonlinear compensation method

Input: Original EEG $\mathbf{Y} \in \mathbb{R}^{N\times T}$ and predicted EEG $\hat{\mathbf{Y}} \in \mathbb{R}^{N\times T}$ for N channels and T samples; k number of bins; d number of discarded samples.
Output: Modified EEG $\mathbf{Y}_{new} \in \mathbb{R}^{N\times T}$

```
procedure
  for each channel ch do
    Define B bins s.t. j: S_j = {} with lower bound b_j
    and upper bound u_j
    for each sample s do
      Set (Ŷ(ch,s), Y(ch,s)) into the appropriate
      bin based on the value of Ŷ
    end for
    for each bin j do
      m̄_j =mean(Y) for Y s.t. (Ŷ, Y) ∈ S_j
    end for
    Fit a line f between the mean values
    m̄_j s.t. j ∈ the two middle bins
    for each sample s do
      if (Ŷ(ch,s), Y(ch,s)) ∈ bin k then
        x_k as the mid Ŷ-value in bin k
        Y_new(ch,s) = Y(ch,s) f(x_k)/m̄_k
      else
        Y_new(ch,s) = 0
      end if
    end for
  end for
end procedure
```

experimental setup is provided in [11]. This study comprised of 40 trials and was reviewed and approved by The Swedish Ethical Review Authority, Sweden (DNR: 2022-05129-01).

3.2. Data preprocessing

The envelopes of the target and masker speech stimulus in each trial were computed with the absolute values of the Hilbert transform. These envelopes were then band-pass filtered between 1–8 Hz by a 6$^{\text{th}}$ order Butterworth filter, downsampled to 100 Hz and cut between the first and last second to remove the edge effects. EEG data was re-referenced to the two mastoid channels, 50 Hz line noise was filtered out followed by a 6$^{\text{th}}$ order Butterworth filter pass-band filtering at 0.5–70 Hz. Manual inspection did not identify any bad channels. Artifacts such as eye blinks, heart beats and muscle movements were removed with independent component analysis [12]. The data was then filtered with a pass-band of 1-8 Hz using a 6$^{\text{th}}$ order Butterworth filter, downsampled to 100 Hz and cut to match the speech envelopes. For more details, see [11, 13, 14].

The standard 10-20 sensor locations for the 64 EEG channels used in the Biosemi-64 cap [15] was used in this study. An average brain model 'fsaverage' and a volume source space with 10 mm spacing between voxels were used to compute the lead field matrix which consisted of 1024 voxel sources in three-dimensional space [16]. The noise covariance matrix was computed from pre-stimulus EEG recordings from the subject, preprocessed the same way as for the speech-evoked EEG data recordings [17]. Only the diagonal elements (the variance measured at each sensor) were kept as the final approximation of the noise covariance matrix [18].

4. RESULTS AND DISCUSSION

The middle panels in Fig. 3 show the estimated temporal profiles of the left hemispheres (lh; top) and right hemispheres (rh; bottom), respectively. Each brain source voxel is represented by the amplitude of its 3-dimensional vector TRF and contributes to a line in the panel. Fig. 3 (A) shows the results for the original EEG data $\mathbf{Y}$ and the target stimulus where the P1-peak around 60 ms is clearly visible in both hemispheres. A significant N1-peak around 175 ms along with a small P2-peak at 250 ms are visible in the left hemisphere. The N1-peak at 190 ms is smaller in the right hemisphere, as speech and language processing is mostly left hemisphere dominant [19]. The top and bottom brain plots show the source localization at the characteristic peaks. The amplitudes in the left hemisphere are close to the auditory cortex, while the location of the amplitudes in the right hemisphere is slightly more frontal than expected.

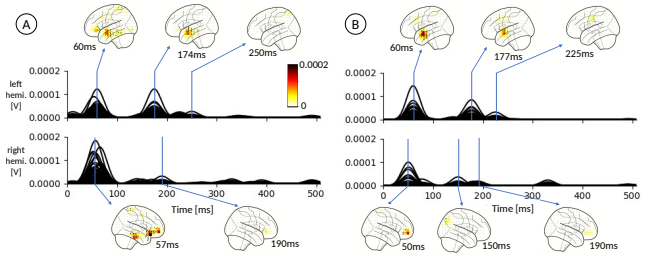

Fig. 3: (A) Estimated source TRFs for the original EEG data $\mathbf{Y}$ and the target stimulus. The middle panels show the temporal profiles of the left (lh; top) and right (rh; bottom) hemispheres respectively. Each voxel is represented by the amplitude of its 3-dimensional vector TRF and contributes to a line in the panel. The top and bottom brain plots show the source localization at the characteristic peaks. (B) Estimated SL TRFs after nonlinear compensation for the binned EEG data $\mathbf{Y}_{new}$ and the target stimulus.

Results of the validation procedure for five evenly distributed channels are presented in the left panel of Fig. 4. A higher bin-number implies fewer samples per bin, which increases the impact from noise. The bin-size with the lowest norm was picked for each channel. The right panel in Fig. 4 shows the number of bins for all channels and where they are placed on the scalp. Dividing the dataset into three bins seems to benefit most of the channels, although one bigger and one

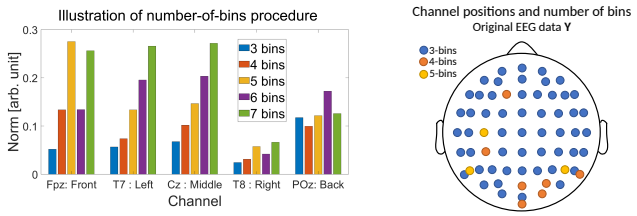

Fig. 4: Left: Illustration of the procedure of choosing the number of bins for five channels. The number with the lowest norm between estimation and validation datasets is chosen. Right: The channel positions on the scalp and the number of bins used for the original EEG data $\mathbf{Y}$.

smaller cluster of 4-5 bins are shown in the right back and mid-center part of the scalp.

The nonlinear compensation method with the original EEG data for 9 channels evenly distributed over the cortex is presented in Fig. 5. Each datapoint is the mean value of all samples in the bin. Both datasets follow the same pattern, where a positive correlation between $\mathbf{Y}$ and $\hat{\mathbf{Y}}$ is shown for several channels. However, channels P7 and POz positioned in the parietal area show some stronger nonlinear behaviors. In contrast to the other channels presented in Fig. 5, these two channels needed to be modeled with five respective four bins. This result reflects that larger number of bins are necessary to capture more pronounced nonlinear behaviors.

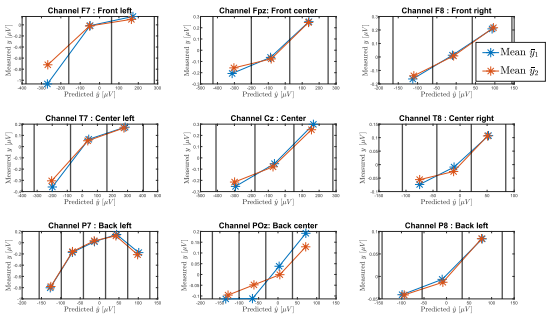

Fig. 5: Illustration of the validation procedure of the nonlinear compensation method for nine evenly distributed channels. All subplots show similar shapes between the equally sized datasets of the original EEG data $\mathbf{Y}$.

The temporal profiles and source localization of the nonlinear compensated EEG data are shown in Fig. 3 (B). Compared to the results of the original EEG data in Fig. 3 (A), the nonlinear compensated data seem to give smoother and more synchronized SL TRF:s. The source localization is more sparse and precise compared to Fig. 3 (A), which is especially noticeable, and relevant, for the important N1-peak in the left hemisphere.

The deviation of the mean value to the line in a bin can be viewed as a measure of nonlinearity. A sensor-specific additional analysis was made where three bins were used for all channels. A line was fitted between the means of the two outer bins, and the deviation of the mean in the middle bin was computed. A value above the line gives a positive residual, and a mean value below gives a negative residual. The topoplot on the right side of Fig. 6 shows the error of the mean value to the line for the middle bin. It shows an interesting pattern where positive values are located in the left hemisphere and negative values in the right hemisphere. The largest positive value was given in channel F1, which is plotted on the left side of Fig. 6. Further investigations on more data and subjects are needed to understand if this pattern is specific to the setup and/or subjects. Future work should also include statistical tests using more subjects to validate the nonlinear compensation method.

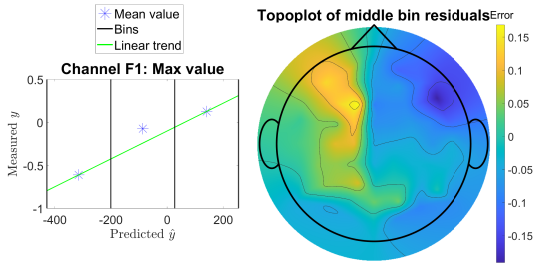

Fig. 6: Left: Illustration of the mean values in three bins and the fitted line. Channel F1 gave the largest positive residual. Right: Topoplot of the middle bin residuals when three bins are used for all sensors.

5. CONCLUSIONS

We have investigated the impact of nonlinear components in real data collected from EEG sensors measuring the electrical brain activity in response to competing speech. EEG data is known to be noisy, making the estimation of source level temporal responses (SL TRFs) a very challenging task. We hypothesized that some of the noise were due to Wiener-structured signal propagation from the firing neurons through brain tissue, skull, and scalp to the sensors placed on the outer scalp.

We have developed a simple nonlinearity detection and compensation method which compares the original EEG data with the predicted EEG data computed from the SL TRF:s. We estimated the nonlinearity of each sensor and were able to compensate for some of these nonlinear components. Our results indicate that the nonlinearity compensation method produced more precise and more synchronized SL TRFs when compared to the original EEG data.

6. REFERENCES

[1] David K. Campbell, "Nonlinear Science: from Paradigms to Practicalities," in *From Cardinals to Chaos: Reflections on the Life and Legacy of Stanislaw Ulam*. Cambridge University Press, 1989, pp. 218 – 262.

[2] Suzana Herculano-Houzel, "The Human Brain in Numbers: A Linearly Scaled-up Primate Brain," *Frontiers in human neuroscience*, vol. 3, pp. 31, 11 2009.

[3] Jie Sun, Bin Wang, Yan Niu, Yuan Tan, Chanjuan Fan, Nan Zhang, Jiayue Xue, Jing Wei, and Jie Xiang, "Complexity Analysis of EEG, MEG, and fMRI in Mild Cognitive Impairment and Alzheimer's Disease: A Review," *Entropy*, vol. 22, pp. 239, 02 2020.

[4] Tao Zhang and Wanzhong Chen, "LMD based features for the automatic seizure detection of EEG signals using SVM," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, pp. 1100–1108, 08 2017.

[5] Deon Garrett, David Peterson, Charles Anderson, and Michael Thaut, "Comparison of linear, nonlinear, and feature selection methods for eeg signal classification," *IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 11, pp. 141–4, 07 2003.

[6] Michel Le Van Quyen, Mario Chavez, David Rudrauf, and Jacques Martinerie, "Exploring the nonlinear dynamics of the brain," *Journal of physiology, Paris*, vol. 97, pp. 629–39, 07 2003.

[7] C.J. Stam, "Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field," *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, vol. 116, pp. 2266–301, 11 2005.

[8] Emina Alickovic, Thomas Lunner, and Lennart Ljung, "A Tutorial on Auditory Attention Identification Methods," *Frontiers in Neuroscience*, vol. 13, 03 2019.

[9] Proloy Das, Christian Brodbeck, and Behtash Babadi, "Neuro-Current Response Functions: A Unified Approach to MEG Source Analysis under the Continuous Stimuli Paradigm," *NeuroImage*, vol. 211, 01 2020.

[10] Johanna Wilroth, Joshua P. Kulasingham, Martin A. Skoglund, and Emina Alickovic, "Direct Estimation of Linear Filters for EEG Source-Localization in a Competing-Talker Scenario," 2023, pp. 7088–7095, to appear.

[11] Asger Andersen, Sébastien Santurette, Michael Pedersen, Emina Alickovic, Lorenz Fiedler, Jesper Jensen, and Thomas Behrens, "Creating Clarity in Noisy Environments by Using Deep Learning in Hearing Aids," *Seminars in Hearing*, vol. 42, pp. 260–281, 08 2021.

[12] Robert Oostenveld, Pascal Fries, Eric Maris, and Jan-Mathijs Schoffelen, "FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data," *Computational intelligence and neuroscience*, vol. 2011, pp. 156869, 01 2011.

[13] Emina Alickovic, Thomas Lunner, Dorothea Wendt, Lorenz Fiedler, Renskje Hietkamp, and Carina Graversen, "Neural Representation Enhanced for Speech and Reduced for Background Noise With a Hearing Aid Noise Reduction Scheme During a Selective Attention Task," *Frontiers in Neuroscience*, vol. 14, 09 2020.

[14] Emina Alickovic, Elaine Ng, Ning Ng, Lorenz Fiedler, Sébastien Santurette, Hamish Innes-Brown, and Carina Graversen, "Effects of Hearing Aid Noise Reduction on Early and Late Cortical Representations of Competing Talkers in Noise," *Frontiers in Neuroscience*, vol. 15, 03 2021.

[15] BioSemi, "The BioSemi headcap," https://www.biosemi.com/headcap.htm, [Online; accessed 01-September-2023].

[16] Bruce Fischl, "FreeSurfer," *NeuroImage*, vol. 62, no. 2, pp. 774–781, 2012, 20 YEARS OF fMRI.

[17] Denis Engemann, Daniel Strohmeier, Eric Larson, and Alexandre Gramfort, "Mind the Noise Covariance When Localizing Brain Sources with M/EEG," 06 2015, pp. 9–12.

[18] Denis Engemann and Alexandre Gramfort, "Automated model selection in covariance estimation and spatial whitening of MEG and EEG signals," *NeuroImage*, vol. 108, 12 2014.

[19] Yvonne Sininger and Anjali Bhatara, "Laterality of Basic Auditory Perception," *Laterality*, vol. 17, pp. 129–49, 03 2012.