Smoothing filters for EEG signals

Nick Porter

Abstract— **The purpose of this paper is to detail the preprocessing of raw EEG data with digital filters to both smooth the signal and remove offsets and drift. A high order high pass FIR filter with low frequency cutoff was used to remove the offset and smooth drift artifacts. A Savitzky–Golay filter was used for smoothing as well as notching out the 50Hz noise from power.**

I. Introduction

The use of electroencephalogram (EEG) data is becoming more and more common as the technology needed to capture these signals becomes more accessible. EEG data is gathered by the use of small electrodes placed at various locations on the scalp. There are numerous sources of noise, for which this paper only addresses a few. There can be signal drift, due to variable impedance between the sensor and the subject, as well as DC offsets. There is often (as is the case in the dataset used for this paper) a spike at the power frequency. There can be noise from various other bodily processes, such as eye movement, muscle contractions, and the heart.

The data used in this paper [1] was gathered from an experiment on "resting state and auditory stimulation experiments", so the majority of brain wave activity was expected to be around the 10Hz alpha wave range, as these are associated with a conscious, closed-eye resting state.

The Savitzky–Golay filter is used to smooth signals without making big alterations to the original form of the signal. It is a go-to filter for smoothing time-series data this reason, and is often cited for use on EEG signals.

The purpose of this filtering design was to remove the DC offset, reduce the noise from power, and smooth the overall signal in the time domain.

II. Methodology

A causal implementation of the filtering process was used, to prevent distortions that might appear at the front edges of an event, which could trigger a recognition of the event before it actually happened. Before each filter was applied, the signal was mirrored at each edge by the length of the filter. After filtering, those values were discarded. This was to prevent edge effects.

A high pass filter was used first to remove the DC offset and very low frequency drift. Then a Savitzky–Golay filter was applied to smooth the signal, while also reducing the noise from power at 50Hz.

The specific reasons for choosing this filtering process are due to the situation in which it is being used. Since it is only being applied in post-processing, at the leisure of the data analyst, delay and computation time are not a concern. For this reason, FIR filters could be used for both, in order to achieve a linear phase response. The FIR filter used causes a constant group delay, which means the output can simply be shifted back by the same amount to ensure that time stamps still line up with the original data. This could not be done with real time processed data.

A. High-Pass filtering

The data was first processed with a high pass filter to discard drift and DC offset. This filter was chosen not to offset the benefits of the choices for the smoothing filter, so it was also linear phase and causal. EEG filtering is often high pass filtered at about 0.05Hz, since brainwave spectral content goes down to around 1Hz. This was not a concern for this paper, since the relevant frequencies from this dataset are in the 8-12 Hz range (alpha waves).

Using filter design tools provided by the MNE library for python, a linear phase high-pass FIR filter with a cutoff at 3Hz, 2Hz transition band, and a Hamming window was applied. A Hamming window is often used in EEG preprocessing for high pass filtering at low frequency cutoff values as it has high stop-band attenuation with a relatively low transition band.

These parameters forced the filter to be high order, 331 samples, with a group delay of 156 samples. The delay after which the bulk of the impulse has passed is about 180 samples. For this reason, this filter design would not work well in a time-sensitive process, but is acceptable for working with pre-recorded data.

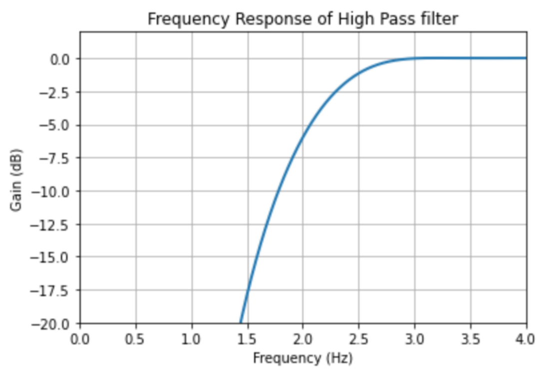

Fig. 1. High pass filter frequency response

B. *Savitzky–Golay filtering*

The parameters of the Savitzky–Golay filter were chosen to provide moderate smoothing, while also attempting to place the first zero of the frequency response near the 50Hz noise spike. The filter generally works like a low pass filter, unless it is designed to also take the derivation. As the length of the filter, M, increases, the cutoff frequency decreases. But as the polynomial order increases, the cutoff frequency also increases. A balancing between the two can achieve some room for design. A length of 7 samples was chosen, and a polynomial order of 3. The derivative properties of this filter were not used for this design.

For this data, there is a 50Hz spike from power which can be filtered out to make the time series data more transparent. Setting the Savitzky–Golay filter parameters such that the first zero aligns (nearly) with this spike can yield better suppression of this 50Hz noise than using an FIR bandstop with a narrow band centered at 50Hz.

III. Results

The results of this process are not groundbreaking by any means, but they do show some

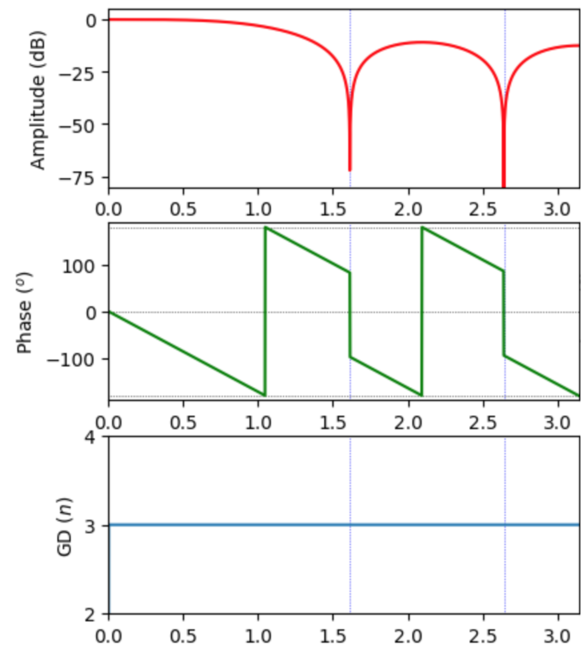

Fig. 2. Amplitude, wrapped phase, and group delay of the Savitzky–Golay filter used.

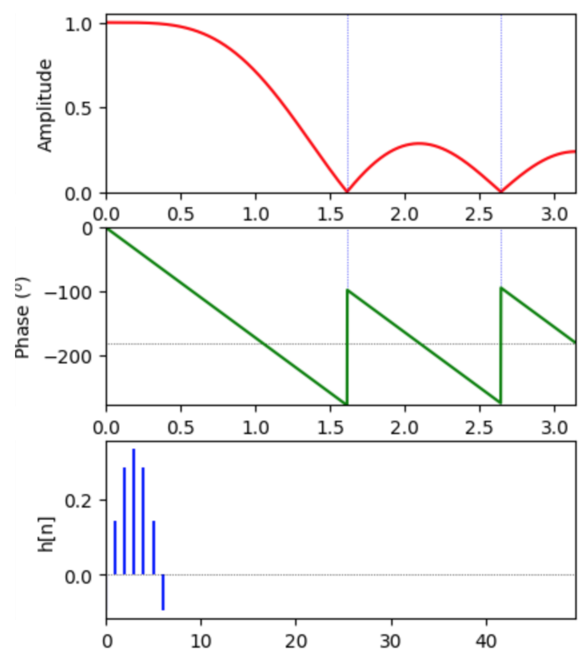

Fig. 3. Amplitude, unwrapped phase, and impulse delay of the Savitzky–Golay filter used.

clear benefits. For the sake of brevity, the figures used in this paper focus on only one of the sensors from a single subject in the dataset. The sensor is titled ’P4’. The high pass filtering succeeds in removing DC offset, and helps to remove some

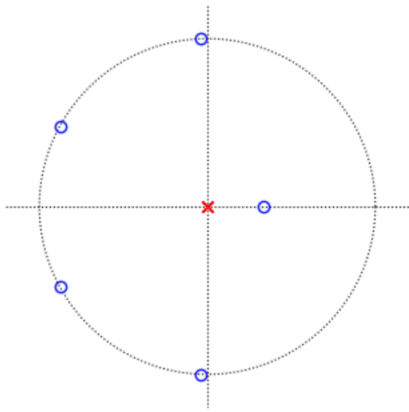

Fig. 4. Pole-Zero plot of the Savitzky–Golay filter used.

low frequency drift, as shown in Figures 5 and 6.

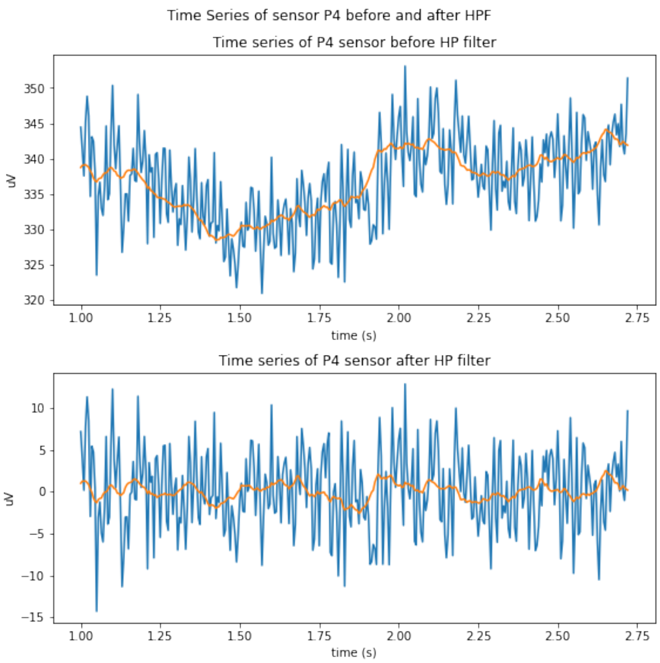

Fig. 5. Time series of a single sensor's values from the raw data (top) and after being filtered by the high pass filter (bottom). A moving average is overlaid each to emphasize the effect.

The results of the Savitzky–Golay filter were mixed. Even though it isn't the best tool for the job, it worked well to suppress the spike at 50Hz, as shown in figure 7. It also suppressed higher frequency noise, although the zeros of the savgol filter make the data above 60Hz more distorted. For the sake of this paper, that is not an issue, though.

Figure 8 shows the effect of suppressing the 50Hz noise in the time domain.

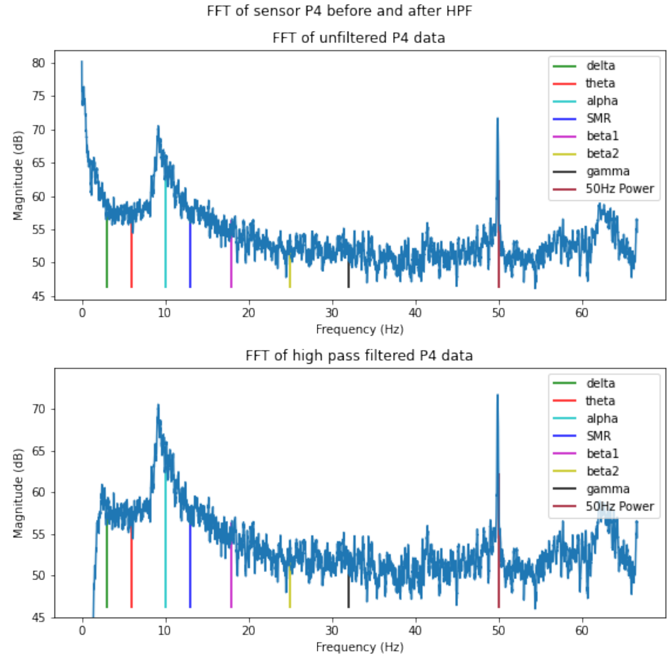

Fig. 6. A comparison between the FFT of the raw data and the FFT of the high pass filtered data. A moving average was used on the FFT data in order to make these plots clearer.

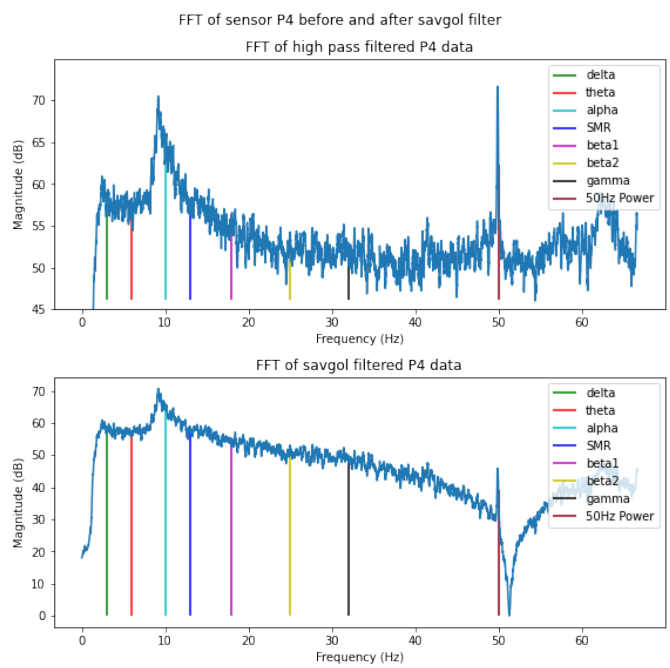

Fig. 7. A comparison between the FFT of the high pass filtered data and the FFT of the Savitzky–Golay and HP filtered data. A moving average was used on the FFT data in order to make these plots clearer.

IV. Discussion

The Savitzky–Golay was not originally intended to be used as a notch filter, as it is here. But

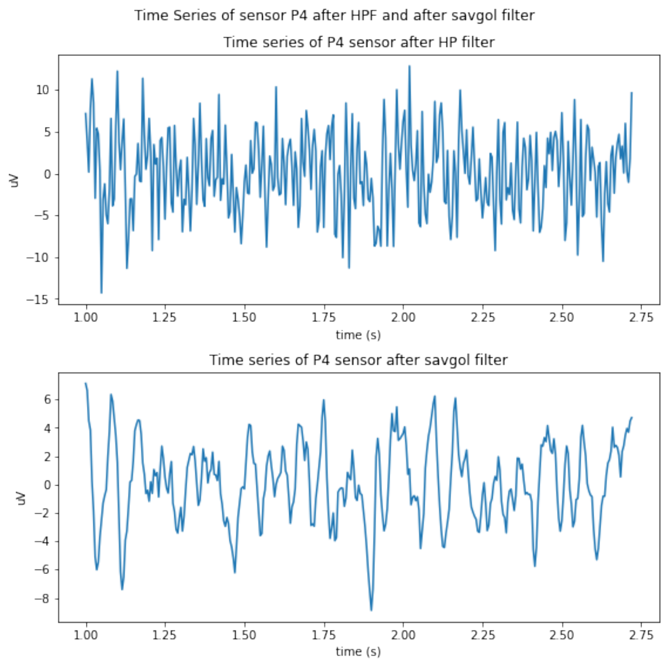

Fig. 8. Time series of the HP filtered data (top) and after being filtered by the Savitzky–Golay and HP filter (bottom). A moving average is overlaid each to emphasize the effect.

in combination with smoothing higher frequency artifacts, it works moderately well at attenuating the 50Hz noise for this purpose.

An IIR filter would be better suited for notching a single frequency, however, and the smoothing effects shown here have not been proven to be more beneficial than other smoothing methods. The transition band around the notches would be tighter with an IIR notch filter, which would allow easier detection of higher frequency brainwaves in the gamma range. This would result in a non-linear phase response, but that might be a justified trade-off.

The results of this filtering process do show a significantly increased clarity of the relevant data. This post-processed data could be more effectively used to develop an algorithm for detecting alpha wave content in EEG signals.

V. Conclusion

The combination of high pass and Savitzky–Golay appeared to effectively clean the raw data in order to more easily detect the alpha waves present in this particular dataset. This combination would not work well for detecting signals below 4Hz or above 40Hz, so it is not useful for all EEG signal processing purposes.

References

[1] Abo Alzahab, N., Di Iorio, A., Apollonio, L., Alshalak, M., Gravina, A., Antognoli, L., Baldi, M., Scalise, L., and Alchalabi, B. (2021). Auditory evoked potential EEG-Biometric dataset (version 1.0.0). PhysioNet. https://doi.org/10.13026/ps31-fc50.

[2] Jiang X, Bian GB, Tian Z. Removal of Artifacts from EEG Signals: A Review. Sensors (Basel). 2019;19(5):987. Published 2019 Feb 26. doi:10.3390/s19050987

[3] Kawala-Sterniuk A, Podpora M, Pelc M, et al. Comparison of Smoothing Filters in Analysis of EEG Data for the Medical Diagnostics Purposes. Sensors (Basel). 2020;20(3):807. Published 2020 Feb 2. doi:10.3390/s20030807

[4] Gallagher, Neal. (2020). Savitzky-Golay Smoothing and Differentiation Filter.