Adaptive Removal of the Transcranial Alternating Current Stimulation Artifact from the Electroencephalogram

Robert Guggenberger[1,*] and N.N.[1]

[1]Department for Translational Neurosurgery, University Hospital Tübingen
[*]Corresponding author: robert.guggenberger@posteo.eu

May 12, 2017

Abstract

Your abstract.

1 Introduction

The combination of transcranial alternating current stimulation (tACS) and electroencephalogram (EEG) has been explored in several recent studies. While the analysis of EEG before or after stimulation posits limited technical challenges, the EEG recording during stimulation is heavily affected by the stimulation artifact.

1.1 Non-Stationary Amplitude Modulation

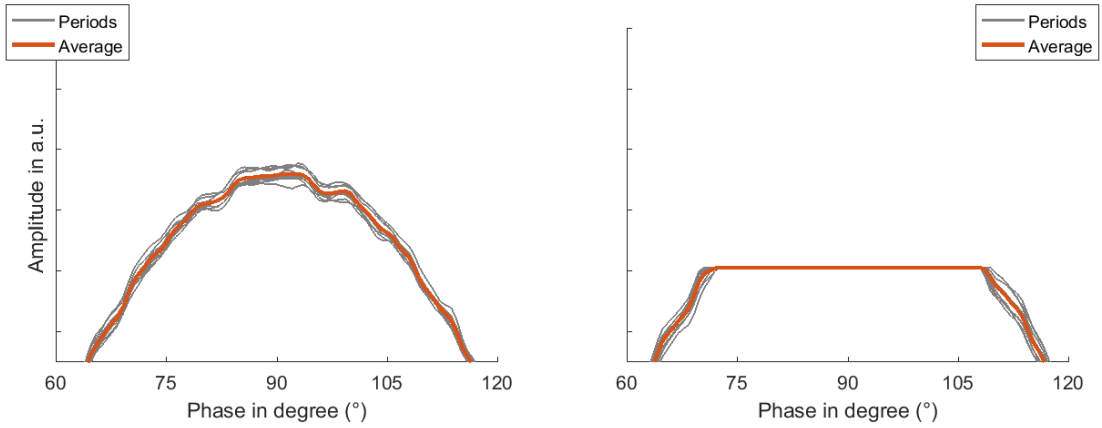
An approach to tackle this issue is to assess the time-course of the EEG signal. Consider the assumption that the artifact is stationary and superpositioned on the physiological signal. Then, modulations in the amplitude of the recorded EEG-signal must be caused by changes in the underlying physiology. This would be the case, even if frequency and phase are matched to the stimulation signal. Approaches assuming such stationarity of the stimulation artifact have been used e.g. by Pogosyan et al. (2009).

Yet, detailed analysis of the stimulation artifact provides evidence that the artifact amplitude is actually not stationary. Instead, the amplitude is modulated by heart-beat and respiration (Noury et al., 2016). Recently, non-linearities in how stimulators control the applied current has been suggest as further source of modulation (Neuling et al., 2017). It has been argued that unregularized spatial filters might be able to remove this amplitude modulation (Neuling et al., 2017). But if only few channels are recorded, the method can fail, as the estimation of the spatial covariance is insufficient, or impossible in the single-channel-case. Consider furthermore that event-related responses like modulation of skin impedance can also affect the scalp conductance at stimulation electrodes. This would introduce event-related amplitude modulation of the stimulation artifact. In that regard, disentangling true signal from the stimulation artifact stays technically challenging.

1.2 Artifact Distortion

Ideally, the stimulation artifict of tACS resembles a sinusoid. Yet, practical experience suggests that the signal is usually distorted to various degrees. Figure 1 shows examples of distortion and saturation in two recordings of tACS-EEG. The gray traces indicate nine invididual periods, while the red trace indicates their average. In figure 1a, note the periodic, yet non-sinusoidal waveform. In figure 1b, note the saturation.

The temporally and spatially uneven impedance distribution has been suggested as cause of distortion, rendering the resulting waveform periodic, but non-sinusoidal. A major problem is amplifier saturation, i.e. the stimulation artifact exhibiting an amplitude to large for the dynamic range of the amplifier, causing



(a) Distortion (b) Saturation

Figure 1: Examples for loss of sinusoidal fidelity

the signal to be cut off and information to be lost. Additionally, non-linearites in the amplifier slew rate can distort the shape even when the signal is close to the saturation threshold.

1.3 Computational Demands

Methods based on adaptive template construction and temporal principal component analysis (tPCA) (Niazy et al., 2005) have been explored for removal of non-stationary and misshaped tACS artifacts (Helfrich et al., 2014). Consider that the process of template construction, the estimation of accurate weights for removal by template subtraction and the suqsequent removal of residual artifacts using tPCA is computationally cumbersome. Additionally, it often requires off-line analysis supported by visual inspection. Such a multi-staged template-approach is therefore of limited utility for on-line artifact removal. Furthermore, critical evaluation has suggested that the residual artifact spans several principal components, and a sufficient artifact removal is therefore not possible with tPCA (Noury et al., 2016).

1.4 Motivation

We were interested in development of a computationally fast approach, feasible for online artifact removal. At the same time, the approach was required to account for the dynamical modulation of the artifact amplitude, and the possibility of non-sinusoidal distortion and saturation. Ideally, the approach would allow to derive physiological signals at the frequency of stimulation, even if physiological oscillations were phase-locked to the stimulation signal. Furthermore, the approach should be usable offline, as to allow recovery of information from already collected datasets.

2 Approach

The main idea is that at any given time point t, the recorded signal $r(t)$ is a linear superposition of a neurophysiological signal $n(t)$, the stimulation artifact $a(t)$ and a white noise term $e(t)$. The task is to recover $n(t)$ by estimating $\hat{a}(t)$ and $e(t)$ and subtracting from $r(t)$.

$$r(t) = n(t) + a(t) + \epsilon(t) \tag{1}$$

$$n(t) = r(t) - \hat{a} - \epsilon(t) \tag{2}$$

2.1 Periodic Estimation

Assume that the tACS artifact were *non-sinusoidal*, but *stationary and periodic*. At the same time, assume that neurophysiological signals n and noise ϵ were absent. Then, we could estimate the amplitude of a at any time-point t by using the signal r recorded from any time-point, as long as this time-point is an integer multiple of the artifacts period length p earlier (3). Subtraction of a delayed version of the signal is also known as comb filter. Please note that for discretely sampled signals, this approach natively supports only frequencies which are integer divisibles of the sampling frequency. If this is not the case, upsampling the recorded signals to an appropriate sampling rate, comb filtering, and downfiltering is a viable approach.

$$\hat{a}(t) = r(t - np) \tag{3}$$

$$n \in \mathbb{Z} \tag{4}$$

2.1.1 Uniform Comb Filter

Consider that the noise term ϵ is still superpositioned on r. If the noise term were white, and because the expectation of white noise $\langle\epsilon\rangle$ converges asymptotically to zero with increased sample size, an approach to estimate a bias-free artifact amplitude would be to average across as many earlier periods as possible (5). Subsequently, this estimate can be used to remove the artifact from r. In real applications, stimulation duration is limited and computational constraints exist. This is reflected by the fact that we have to use a finite number for N.

$$\hat{a}(t) = \sum_{n=1}^{N} \frac{r(t - np)}{N} \tag{5}$$

2.1.2 Superposition of Moving Averages

Please note that averaging across neighbouring periods M (6) has been suggested before and termed superposition of moving averages (SMA) by Kohli and Casson (2015).

$$\hat{a}(t) = \sum_{n-M/2}^{n+M/2} \frac{r(t - np)}{M + 1} \tag{6}$$

Consider that the approach using only past values (5) returns a causal filter. Applied online, a causal filter would be able to remove the artifact without the delay of $(Mp)/2$ necessary for SMA. Furthermore, SMA is well-defined only for even M. This motivates the exploration of causal filters for artifact removal.

2.2 Temporal Weighting

Consider that the amplitude of the artifact has been described to be non-stationary and dynamically modulated (Noury et al., 2016; Neuling et al., 2017). Although it has been suggested that there are event-dependent components of the amplitude modulation (e.g. by heartbeat or respiration Noury et al. (2016) or stimulator impedance check Neuling et al. (2017)), the parameters of the dynamical system governing event-independent amplitude modulation are usually not known a priori.This can render online artifact removal problematic.

One approach to tackle this problem is to use instead of a constant weight $1/N$ (5), a time-dependent weighting function w_n (7).

$$\hat{a}(t) = \sum_{n=1}^{N} w_n r(t - np) \tag{7}$$

By controlling the parameters of the weight functions, can we attempt to match the process governing the amplitude modulation This might allow us to achieve a better artifact estimation and subsequent removal. Please note that if the reliability of past periods can be estimated, we might select weights empirically. 2.4.3 Yet, these approaches require extended calculation periods and/or offline analysis. In the following sections, we will discuss and justify three predeterminable weighting functions feasible for fast online-filtering.

2.2.1 Justification by Sampling

Consider for example the simple one-step comb filter, where we remove the artifact by subtracting a value sampled from any earlier period. Assume now that this value were not drawn from the last period, but instead drawn at random from the last N periods with uniform distribution. If the system governing amplitude modulation were fully stationary for the last N periods, performance would in law be virtually identical to the comb filter based on averaging (5). If it were instead non-stationary, we would expect that the accuracy of the estimate degrades with longer delay while the precision improves by increasing the number of periods on which we base our estimate. For example, in the case of the uniform filter, we have a shape parameter N, defining how far back we trust a measurement to have the same accuracy as earlier samples. This rationale justifies non-uniform weights, and to use weights that are designed to increase precision and accuracy of the estimate.

2.2.2 Justification by AR (1) process

Consider that the system governing the amplitude of the stimulation artifact could be decomposed into the constant amplitude c controlled by the stimulator and a dynamical, *event-unrelated* process governing the amplitude modulation. If we model this process as an AR (1) process, this would return for N approaching infinity:

$$X_t = c + \sum_{n=0}^{\infty} \phi_n \epsilon_{t-k} \tag{8}$$

If the kernel Φ behind the modulation of the artifacts amplitude were known, or could be estimated sufficiently, we could construct an optimal weighting function as a deconvolution filter. This line of reasoning is based on the similarity between the generic weighted comb filter (7) and the generic discrete-time AR (1) process (8). Note that, for practical applications, we would need finite N. This would limit the application to kernels decaying sufficiently fast.

2.3 Temporal Weighting Examples

2.3.1 Linear Weighting

One straight-forward approach is using a linear decreasing weighting function. The assumption of local linearity is widely used in the analyis of non-linear dynamical systems, which could render the kernel sufficiently flexible. Additionally, the function is simple and the necessary normalization can easily be calculated by using the triangular number for a given N as normalizing constant k (10). Hence, equation (9) returns weights for earlier periods based on a linear temporal weight decay.

$$w_n = \frac{N - n + 1}{k} \tag{9}$$

with the following normalization

$$k = \sum_{n=1}^{n} n = \frac{N(N+1)}{2} \tag{10}$$

2.3.2 Exponential Weighting

Motivated by the fact that the autocorrelation function of an AR (1) process can be expressed as a decaying exponential, an alternative approach would be an exponential weighting function. The time constant τ of an exponential controls its temporal decay. To maintain the shape across different N, we consider it reasonable to normalize n by N. Hence, equation (11) returns weights for earlier periods based on their exponential temporal weight decay.

$$w_n = \frac{1}{k} e^{\tau - \tau(n/N)} \tag{11}$$

with the following normalization

$$k = \sum_{n=1}^{N} e^{\tau - \tau(n/N)} \tag{12}$$

2.3.3 Gaussian Weighting

Consider that the amplitude modulation might be governed by more than one AR (1) process, or the process might be non-stationary itself. Motivated by the fact that such sums often converge to a normal distribution, an alternative approach would be a Gaussian weighting function. Using a suitable parameterization, and centering on zero, the inverse of the standard deviation $\frac{1}{\sigma^2}$ defines the time constant τ of a Gaussian distribution, which controls its temporal decay. To maintain the shape across different N, we consider it again reasonable to normalize n by N. Hence, equation (13) returns weights for earlier periods based on their Gaussian temporal weight decay.

$$w_n = \frac{1}{k} f(n/N) \tag{13}$$

with the following generating function and normalization

$$f(x) = \sqrt{\frac{\tau}{2\pi}} e^{-(\tau x^2)/2} \tag{14}$$

$$k = \sum_{n=1}^{N} f(n/N) \tag{15}$$

2.4 Alternative Filter Approaches

Besides the discussed real-valued causal comb filters, several other approaches for filtering the tACS-artifact are possible. For example, one could derive the weight at each period from the autocorrelation across periods 2.4.1, or remove periodic principial components until the artifacts power is sufficiently suppressed 2.4.3, or use complex-valued kernels which would natively support non-integer frequencies 2.4.2.

2.4.1 Automatic Weight Estimation

As mentioned, weighted comb filters can be justfied by using weights to maximize accuracy and precision of the artifact estimate based on sampling from earlier periods. This process can be put on an empirical foundation, e.g. by calculating the autocorrelation of the recording, which is dominated by the artifact. Weights are then based on the autocorrelation coefficients at periodic lags.

2.4.2 Causal Discrete Fourier Transformation

We might also tackle the non-stationarity of the artifact by estimating the artifacts power in a specific frequency for each time-point, using a window of N periods based on the frequency of tACS. This approach can be implemented as convolution of the recording with a complex kernel at the artifacts frequency, followed by subtraction. While it has the advantage that it supports natively frequencies which are not integer

divisibles of the sampling frequency, we need to assume sinusoidality of the artifact. Additionally, to remove the artifact at harmonic frequencies, the computation has to be performed several times for each harmonic frequency. Whether this approach is sufficient, will therefore depends on whether the artifact can be sufficiently catched by these few frequencies.

2.4.3 Adaptive Removal of Principal Components

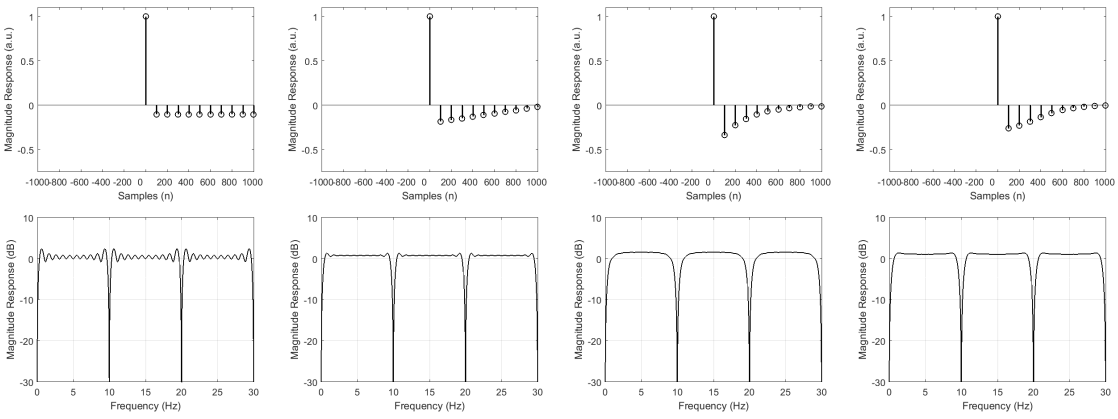
Instead of understanding the artifacts in sinusoidal terms, we might focus on its periodicity and estimate the modulation of the artifacts periodic components across time. This can be implemented by cutting a recording into epochs with period length, followed by prinical component analysis. If the trial duration is significantly larger than the duration of the physiological signal, we can safely assume that the physiological signal will only span a small subspace with small eigenvalues. Removing the largest principal components is then more likely to reduce the power of the artifact, but not of the signal. This approach can be performed iteratively until a specific condition is achieved. We implemented such an iterative approach based on the power in the artifacts frequency, removing components until artifact power is sufficiently reduced to the level of nearby, non-artifacted frequencies. From an algorithmic perspective, we found that a naive approach of simply cutting the recording into epochs returns boundary artifacts at the start and end of each epoch. Running the algorithm in parallel for a set of random initial lags before cutting, and subsequently averaging across these parallel processes, sufficiently removed such boundary artifacts.

3 Evaluation

We implemented all of the discussed approaches for kernel creation and artifact removal in Matlab 2016b. The code can be accessed freely, in open-source and online at `https://github.com/agricolab/ARtACS` and has been licensed under a X11-license. We selected several kernel approaches and evaluated their frequency response characteristics (see 3.1). Additonally,we evaluated their performance in comparison also to alternative approaches on simulated (see 3.2) and real data (see 3.3).

3.1 Evaluation of Frequency Response

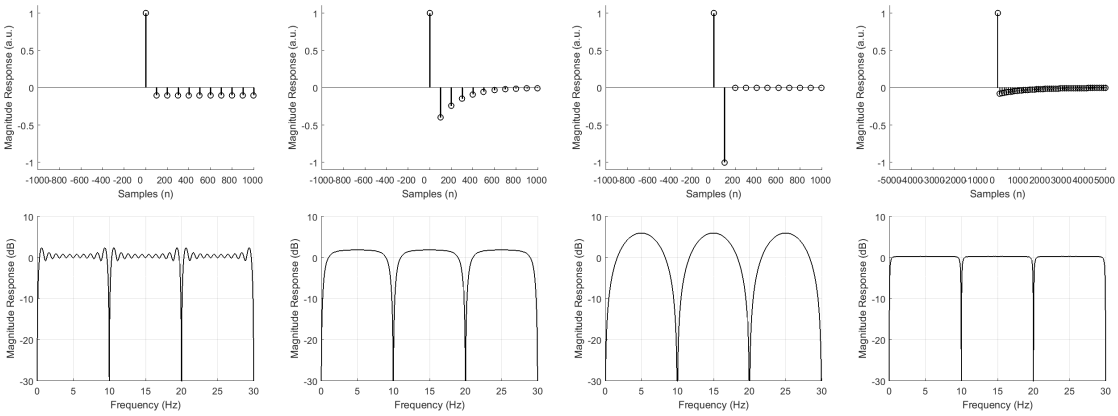
Examine the following exemplary kernels constructed for a sampling rate of 1KHz, a stimulation frequency of 10 Hz and a memory of 10 (see figure 2).



(a) Uniform (b) Linear (c) Exponential $\tau = 4$ (d) Gaussian $\tau = 3$

Figure 2: Exemplary Causal Kernels

As you can see from figure 2, the (numerically calculated) magnitude responses of the four approaches are highly similar. Their key characteristic is the strong supression of the target frequency and its integer multiples. Yet, note the difference in passbands. We find strong ringing in the passband for the uniform (2a)



(a) $\tau = 0, N = 10$ (b) $\tau = 5, N = 10$ (c) $\tau = 100, N = 10$ (d) $\tau = 4, N = 50$

Figure 3: Exemplary τ and N Tuning

and linear kernel (2b), especially compared to the smooth transitions of the exponential (2c) or Gaussian kernel (2d).

We constructed a set of exponential kernels constructed with different τ and N to explore their behavior (see figure 3). We find it of note that that in the limiting case of $\tau = 0$, the exponential kernel virtually converges with the uniform kernel (see figure 3a). Using a τ equal to the artifacts period length, the exponential kernel almost fully converges with the simple comb filter (see figure 3c). Note also that very high τ would return an impulse response, and therefore just pass all signals. This is also true for the limiting case of $N = 1$, were we would again acquire a simple comb filter. By increasing N we achieve a flattening of the pass-band gain (see figure 3d).

3.2 Evaluation on Simulated Signals

To evaluate the filters, we tested them on simulated signals. The simulated signal was created as superposition of a 10 Hz tACS-artifact, and an event-related potential (ERP). The ERP was simulated as the gradient of a flat top window. The amplitude of the artifact was simulated as a distorted, non-stationary sinusoidal. The initial stimulation signal was modelled as sinusoidal signal. We distorted it by adding periodic white noise. Additionally, the stimulation amplitude was driven by an Ornstein–Uhlenbeck process with a stiffness 0.5, and a Hanning window time-locked to the event- related potential. In that way, we were able to simulate event-independent modulations of the artifacts amplitude, and modulations locked to an event. [1] Such effects were described by (Noury et al., 2016), as having a strong potential to mask the true event-related neurophysiological activity.

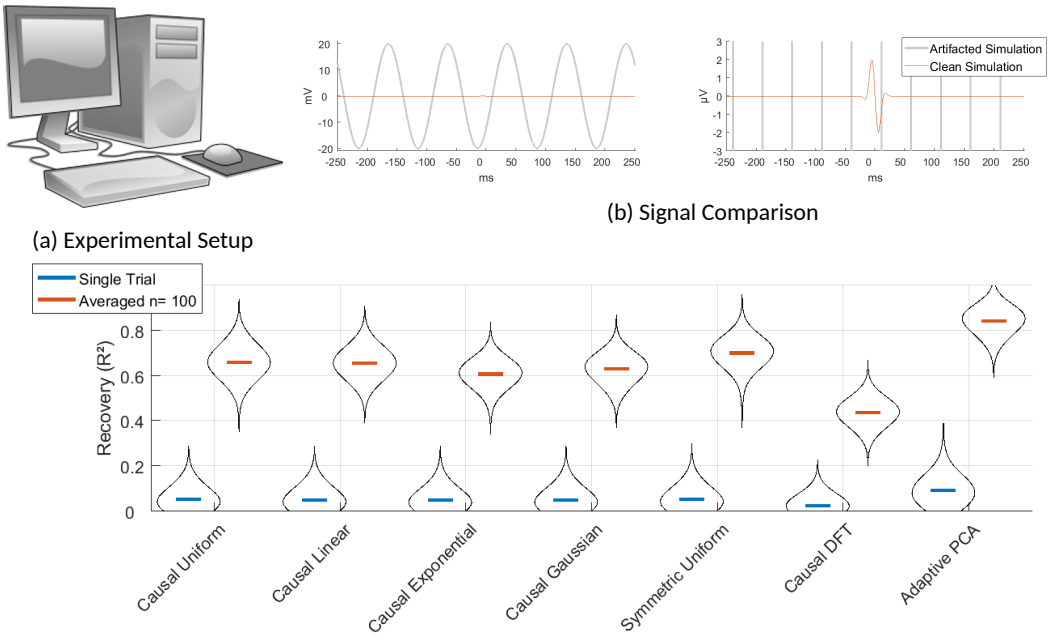
We simulated 100 trials with a length of 8000 samples. Subsequently, we explored different filters for 10 Hz, all with a period number of 10. The causal DFT filter was applied recursively not only at 10, but also at 20, 30, and 40 Hz to reduce harmonics.

Based on the correlation of the filtered, artifacted recording with the known shape of the simulated ERP, we calculated the R^2 value. Blue bars indicate the average R^2 for single-trial reconstructions, and red bars the average R^2 for boot-strapped grand-average ERPs (see figure 4c). Visual inspection of the R^2 estimates shows that all filtering approaches are able to recover at least some information about the ERP. Differences between the comb filters is small.

3.3 Evaluation on real data

To evaluate the filters further, we tested them on real physiological signals. Consider that tACS can have physiological effects on cortical ERP. To properly evaluate the recovery using the filtering approaches, we needed a physiological signal unlikely to be affected by tACS.

[1] Code for signal generation is included in the toolbox at `https://github.com/agricolab/ARtACS`



(a) Experimental Setup

(b) Signal Comparison

(c) Kernel Density Estimation of R^2 for different filter approaches

Figure 4: Overview for Artifact Removal from Simulated Signal

We therefore measured electrocardiogram (ECG) at the flexor digitorum of the left upper limb, while we stimulated distal and proximal with tACS to the recording electrodes (see figure 5a). We also recorded ECG at the chest, and detected the R-peak for epoching the data. Please note that no tACS-artifact was visible in the chest ECG. Physiological signals were acquired using BrainProducts amplifiers at a sampling rate of 1 kHz and low-pass filtered below 35 Hz. Stimulation as delivered using a NeuroConn DC Stimulator Plus with carbon rubber electrodes and Ten20 electrode gel with an intensity of 1 mA at a frequency of 11 Hz.

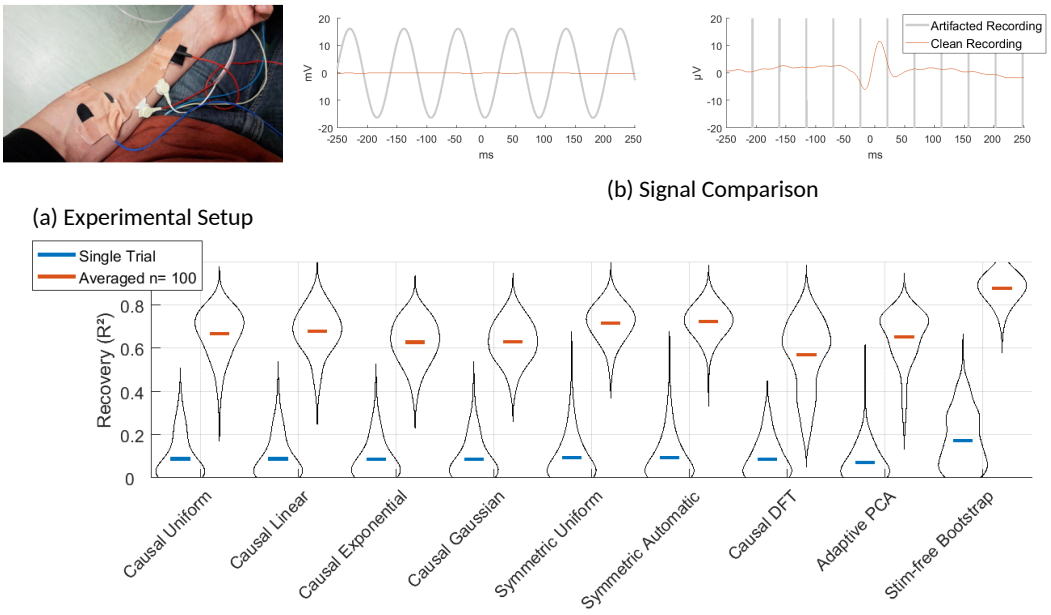
We detected the R-peak based on the chest-recordings, which allowed us to cut epochs of 8s duration for the recording at the upper limb, for a total of 134 R-peaks without stimulation and 136 with stimulation. We estimated the true shape of the R-peak by taking the average of the stimulation-free signal. Based on the correlation of the filtered, artifacted recordings with the averaged stimulation-free recording, we calculated the R^2 value. Blue bars indicate the average R^2 for single-trial reconstructions, and red bars the average R^2 for boot-strapped grand-average ERPs based on 100 epochs (see figure 5c).

4 Discussion

4.1 Echos

A key disadvantage of comb filters is that while it subtracts periodic signals, any underlying physiologal signal will cause a periodic echo, distributed according to the weights of the kernel. This means that strong weights close to the occurence of an event can increase the precision of the artifact amplitude estimate, but cause unwanted and very strong echos. If the duration of the physiological signal is longer than a period, the signal will be distorted by the echo. This could be addressed by *delaying* the kernel, i.e. kernels with zero weights for a number of periods before their estimation weights start, or *inverting* the kernels, i.e. kernels with increasing instead of decreasing weights with increasing delay, or applying causal filters *piece-wise* left and right of the signal. [2]

[2] If delayed or inverted kernels are used in a symmetric kernel, they appear *concave*. While such *concave, delayed or inverted* kernels might extend the window free of echos around physiological signals, they might not sufficiently account for the non-stationarity of the artifacts amplitude. The filtered signal might contain a residual artifact. Whether the trade-off between echo and residual artifact has an optimum for concave kernels or not depends very much on the individual signal under research. It might even vary between trials. Additionally, note that piece-wise filters can not be applied in real-time. Please note that at least on our datasets and simulations, these approaches were usually inferior to decreasing causal or symmetric kernels Nonetheless, concave methods have been implemented in the toolbox.



(a) Experimental Setup

(b) Signal Comparison

(c) Kernel Density Estimation of R^2 for different filter approaches

Figure 5: Overview for Artifact Removal from Upper Limb Electrocardiogram

4.2 Matched Phase and Frequency

Simulations suggested that the power of endogenous oscillations increases if the frequency of tACS matches the targets eigenfrequency (Kutchko and Fröhlich, 2013; Zaehle et al., 2010). This has been supported by evidence from animal studies (Schmidt et al., 2014), and human studies combining tACS with transcranial magnetic stimulation (TMS) (Guerra et al., 2016), or contrasting pre and post resting state power analysis (Zaehle et al., 2010). It has also been suggested that the phase of neuronal populations would be locked to the phase of the tACS signal (Reato et al., 2013). This has been supported by evidence from studies combining tACS with motor output (Brittain et al., 2013), TMS (Raco et al., 2016; Nakazono et al., 2016) or sensory perception (Gundlach et al., 2016). At the same time, event-locked modulation of skin impedance can mimic phase and/or frequency locked modulations. Events might even be endogenous, e.g. heart beats (Noury et al., 2016). This suggests that transient and event-locked modulations in frequency and phase matching to the tACS can make the distinction between artifact and neurophysiological effect difficult or impossible.

References

Brittain, J.-S., Probert-Smith, P., Aziz, T. Z., and Brown, P. (2013). Tremor suppression by rhythmic transcranial current stimulation. *Current Biology*, 23(5):436–440, DOI: 10.1016/j.cub.2013.01.068.

Guerra, A., Pogosyan, A., Nowak, M., Tan, H., Ferreri, F., Lazzaro, V. D., and Brown, P. (2016). Phase dependency of the human primary motor cortex and cholinergic inhibition cancelation during beta tACS. *Cerebral Cortex*, 26(10):3977–3990, DOI: 10.1093/cercor/bhw245.

Gundlach, C., Müller, M. M., Nierhaus, T., Villringer, A., and Sehm, B. (2016). Phasic modulation of human somatosensory perception by transcranially applied oscillating currents. *Brain Stimulation*, 9(5):712–719, DOI: 10.1016/j.brs.2016.04.014.

Helfrich, R. F., Schneider, T. R., Rach, S., Trautmann-Lengsfeld, S. A., Engel, A. K., and Herrmann, C. S. (2014). Entrainment of brain oscillations by transcranial alternating current stimulation. *Current Biology*, 24(3):333–339, DOI: 10.1016/j.cub.2013.12.041.

Kohli, S. and Casson, A. J. (2015). Removal of transcranial a.c. current stimulation artifact from simultaneous

EEG recordings by superposition of moving averages. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, DOI: 10.1109/embc.2015.7319131.

Kutchko, K. M. and Fröhlich, F. (2013). Emergence of metastable state dynamics in interconnected cortical networks with propagation delays. *PLoS Computational Biology*, 9(10):e1003304, DOI: 10.1371/journal.pcbi.1003304.

Nakazono, H., Ogata, K., Kuroda, T., and Tobimatsu, S. (2016). Phase and frequency-dependent effects of transcranial alternating current stimulation on motor cortical excitability. *PLOS ONE*, 11(9):e0162521, DOI: 10.1371/journal.pone.0162521.

Neuling, T., Ruhnau, P., Weisz, N., Herrmann, C. S., and Demarchi, G. (2017). Faith and oscillations recovered: On analyzing EEG/MEG signals during tACS. *NeuroImage*, 147:960–963, DOI: 10.1016/j.neuroimage.2016.11.022.

Niazy, R., Beckmann, C., Iannetti, G., Brady, J., and Smith, S. (2005). Removal of FMRI environment artifacts from EEG data using optimal basis sets. *NeuroImage*, 28(3):720–737, DOI: 10.1016/j.neuroimage.2005.06.067.

Noury, N., Hipp, J. F., and Siegel, M. (2016). Physiological processes non-linearly affect electrophysiological recordings during transcranial electric stimulation. *NeuroImage*, 140:99–109, DOI: 10.1016/j.neuroimage.2016.03.065.

Pogosyan, A., Gaynor, L. D., Eusebio, A., and Brown, P. (2009). Boosting cortical activity at beta-band frequencies slows movement in humans. *Current Biology*, 19(19):1637–1641, DOI: 10.1016/j.cub.2009.07.074.

Raco, V., Bauer, R., Tharsan, S., and Gharabaghi, A. (2016). Combining TMS and tACS for closed-loop phase-dependent modulation of corticospinal excitability: A feasibility study. *Frontiers in Cellular Neuroscience*, 10, DOI: 10.3389/fncel.2016.00143.

Reato, D., Rahman, A., Bikson, M., and Parra, L. C. (2013). Effects of weak transcranial alternating current stimulation on brain activity—a review of known mechanisms from animal studies. *Frontiers in Human Neuroscience*, 7, DOI: 10.3389/fnhum.2013.00687.

Schmidt, S. L., Iyengar, A. K., Foulser, A. A., Boyle, M. R., and Fröhlich, F. (2014). Endogenous cortical oscillations constrain neuromodulation by weak electric fields. *Brain Stimulation*, 7(6):878–889, DOI: 10.1016/j.brs.2014.07.033.

Zaehle, T., Rach, S., and Herrmann, C. S. (2010). Transcranial alternating current stimulation enhances individual alpha activity in human EEG. *PLoS ONE*, 5(11):e13766, DOI: 10.1371/journal.pone.0013766.