NeuroImage 207 (2020) 116356

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

ZapLine: A simple and effective method to remove power line artifacts

Alain de Cheveigné [a,b,c,*]

[a] Laboratoire des Systèmes Perceptifs, UMR 8248, CNRS, France
[b] Département d'Etudes Cognitives, Ecole Normale Supérieure, PSL, France
[c] UCL Ear Institute, United Kingdom

ABSTRACT

Power line artifacts are the bane of animal and human electrophysiology. A number of methods are available to help attenuate or eliminate them, but each has its own set of drawbacks. In this brief note I present a simple method that combines the advantages of spectral and spatial filtering, while minimizing their downsides. A perfect-reconstruction filterbank is used to split the data into two parts, one noise-free and the other contaminated by line artifact. The artifact-contaminated stream is processed by a spatial filter to project out line components, and added to the noise-free part to obtain clean data. This method is applicable to multichannel data such as electroencephalography (EEG), magnetoencephalography (MEG), or multichannel local field potentials (LFP). I briefly review past methods, pointing out their drawbacks, describe the new method, and evaluate the outcome using synthetic and real data.

1. Introduction

Most laboratories are criss-crossed with power lines, and peppered with line-powered lights and electronic equipment. The electric and magnetic fields that these create can interfere with the tiny potentials and fields that are measured from the brain with electroencephalography (EEG), magnetoencephalography (MEG) or other electrophysiological techniques. Power line artifacts are manifest in the spectrum as a series of narrow peaks at multiples of the power line frequency (50 Hz or 60 Hz depending on the geographical region) (Fig. 1). The mechanisms by which that power makes its way into the rig can be puzzling for the experimentalist (ground loops, capacitive coupling, etc.), and the presence of some amount of line artifact in the data is often accepted as a fact of life. Line interference is usually unavoidable in MEG, which is sensitive to fields from distant sources, and for preexisting datasets the question of 'avoiding artifacts at the source' is moot.

A wide range of approaches have been proposed to attenuate or remove line artifacts, including *lowpass* and *notch filters* (Luck, 2005), *frequency domain filters* (Mitra and Pesaran, 1999; Mullen, 2012; Keshtkaran and Yang, 2014; Leske and Dalal, 2019), *regression based on a reference signal* (Vrba and Robinson, 2001; de Cheveigné and Simon, 2007), *independent component analysis (ICA)* (Barbati et al., 2004; Escudero et al., 2007) or other *spatial filtering* techniques (de Cheveigné and Parra, 2014). These methods are well known and widely used, and there are ongoing efforts to develop new ones (Leske and Dalal, 2019). Many of these methods are listed in standard guidelines and textbooks (Picton et al., 2000; Widmann and Schröger, 2012; Luck, 2005; Gross et al., 2013), implemented in widely-used data analysis toolboxes (Delorme and Makeig, 2004; Oostenveld et al., 2011; Gramfort et al., 2014), or integrated in automated processing pipelines (Bigdely-Shamlo et al., 2015; Gabard-Durnam et al., 2018).

A straightforward solution is simply to low-pass filter the data with a cutoff below the line frequency (50 or 60 Hz). The primary motivation is often to enhance slowly varying features relative to fast fluctuations judged less relevant, line artifact reduction being an added benefit. A cutoff in the range 10–30 Hz is typical (Fig. 2 top). Downsides are (a) loss of high frequency information, (b) limited attenuation at line frequency unless the cutoff is steep, and (c) waveform distortion, particularly if the cutoff is steep and the filter impulse response consequently long (Widmann and Schröger, 2012; Cohen, 2014; de Cheveigné and Nelken, 2019).

One effective version of this idea convolves the data with a square-shaped impulse response with a duration exactly equal to the line period. The transfer function of this filter has zeros at the line frequency and all harmonics, and thus the filter *perfectly* removes the artifact (Fig. 2 panel 2), and the impulse response is short so waveform smearing or ringing are minimized. In the event that the sampling rate is not multiple of the line frequency, a reasonably effective 'square' impulse response can be approximated by interpolation.

Alternatively, a notch filter, or set of notch filters, can be applied to attenuate the signal at the line frequency and harmonics. Attenuation at other frequencies is minimal if the notches are narrow (Fig. 2 panel 3). The downside is that the very long impulse response of a notch filter can lead to artifacts triggered by the onset, offset, or transitions within the

* Laboratoire des Systèmes Perceptifs, UMR 8248, CNRS, France.
E-mail address: alain.de.cheveigne@ens.fr.

https://doi.org/10.1016/j.neuroimage.2019.116356
Received 13 September 2019; Received in revised form 22 October 2019; Accepted 11 November 2019
Available online 28 November 2019

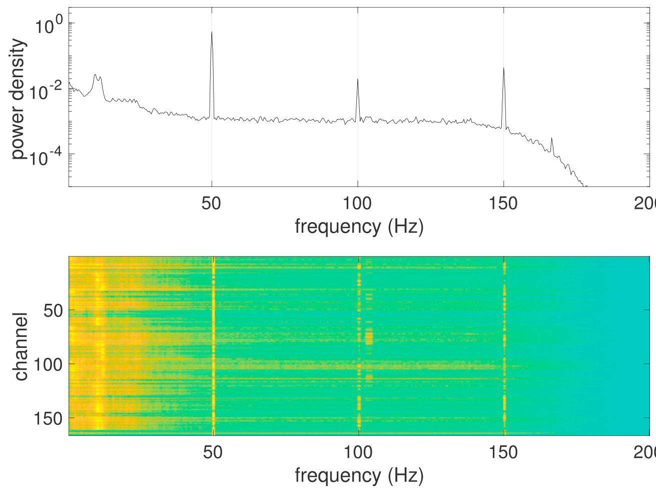

Fig. 1. Normalized power spectral density (PSD) of a MEG data file. Power line artifacts are manifest as spectral peaks at the power line frequency (here 50 Hz) and multiples. The bottom panel shows the same information per channel as a colour plot. All channels are contaminated to some degree. In this plot and others the PSD is normalized so its sum over frequencies is one.

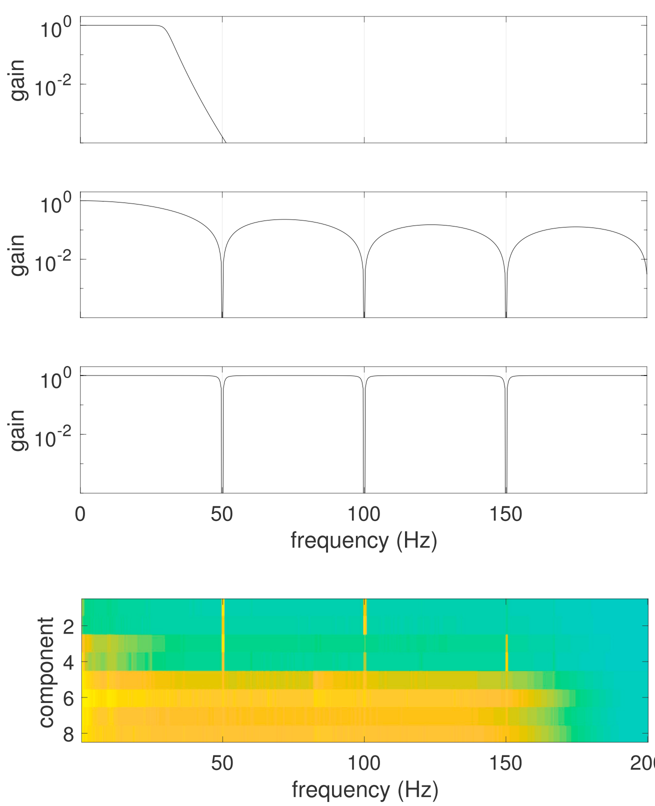

Fig. 2. Top panel: magnitude transfer function of a lowpass filter (Butterworth, order 16, cutoff 30 Hz). Second panel: transfer function of a smoothing filter with square kernel of duration 1/50 Hz. Third panel: transfer function of a multiple notch filter with notches at 50 Hz and multiples. Each notch is implemented as a 2nd order IIR filter with 0.01 relative bandwidth. Bottom panel: PSDs of the first 8 components of a DSS analysis of a 166-channel MEG data file designed to find dimensions dominated by power line noise. The first 4 components, clearly dominated by power line noise, can be projected out of the data to obtain clean data.

data (Vale-Cardoso and Guimarães, 2009; Kıraç et al., 2015). Analogous to applying a notch filter is to fit to the waveform a quadrature pair of sinusoids at the line frequency, and subtracting the fit. Additional pairs can be fit to all harmonics. The same result is obtained in the frequency domain by zeroing the corresponding components, as long as the sampling rate is a multiple of the line frequency, but the fit-and-subtract approach is more flexible in that it allows arbitrary frequency ratios, and the fit can be weighted to avoid contamination by glitches (de Cheveigné and Arzounian, 2018). Regardless of the method, fitting a sinusoid is ineffective if the power line components fluctuate in amplitude and/or phase.

An alternative to time-domain or frequency-domain filtering is to project the data on a reference channel that picks up only line components, and subtract the projection. Certain MEG systems are equipped with reference channels to support this approach, which also allows suppressing a wider range of artifacts (e.g. slow fields due to machinery) (Vrba and Robinson, 2001; de Cheveigné and Simon, 2007). If no reference channel is available, one can be 'created' from multichannel data by applying a spatial filter designed to isolate line-dominated components. ICA has been proposed for this purpose (Barbati et al., 2004; Escudero et al., 2007), and other data-driven techniques are also effective, for example the JD (Joint Diagonalization) method also referred to as DSS (Denoising Source Separation) (de Cheveigné and Parra, 2014). Fig. 2 (bottom) shows, as a colour plot, the PSD of the first few components of a JD/DSS analysis designed for this purpose. The first 4 of these, strongly dominated by line components, may be used as 'reference channels' to project out line artifacts (de Cheveigné and Parra, 2014).

The drawback of the spatial filtering approach is that the power line components are not always linearly separable from brain activity, so projecting out the line-dominated components may lead to loss of useful data. It may also happen that those components are contaminated by other artifacts (such as glitches or drifts), that are then injected into previously uncontaminated channels.

The Algorithm described in this paper addresses these issues, combining the benefits of both spatial and temporal filtering.

2. Methods

Data model. Data consist of a time series matrix $\mathbf{X}$ of dimensions T (time) $\times$ J (channels) that is assumed to be the sum

$$\mathbf{X} = \mathbf{X}_b + \mathbf{X}_n \tag{1}$$

of linear transforms of brain sources $\mathbf{B}$ and power line noise sources $\mathbf{N}$. These sources and transforms are unknown. Linear analysis methods such as ICA try to find a 'demixing matrix' to apply to the data to reverse the mixing, but this is only possible if they are linearly separable within the data. Here we assume that this is *not* the case: there exists no scalar demixing matrix $\mathbf{D}$ such that $\mathbf{X}_b\mathbf{D} = \mathbf{X}_b$ and $\mathbf{X}_n\mathbf{D} = 0$. Likewise, spectral filtering methods try to find a filter that isolates sources one from another, but here we assume that they are *not* separable in the spectral domain, i.e. there exists no filter $\mathscr{F}$ such that $\mathbf{X}_b * \mathscr{F} = \mathbf{X}_b$ and $\mathbf{X}_n * \mathscr{F} = 0$. However, we do assume that there exists a *spatio-spectral* transform in the form of a multichannel filter $\mathscr{T}$ such that brain and line noise *are* separable, $\mathbf{X}_b * \mathscr{T} = \mathbf{X}_b$ and $\mathbf{X}_n * \mathscr{T} = 0$, at least approximately. Our aim is to find such a transform.

Algorithm. As a first step, the data are fed through a two-channel 'perfect reconstruction filterbank', where the first channel convolves them with a square-shaped kernel $\mathscr{K}$ of duration $1/f_{line}$, where f_{line} is the line frequency, and the second with the complementary filter $\delta(0) - \mathscr{K}$. Denoting as $\mathbf{X}' = \mathbf{X} * \mathscr{K}$ and $\mathbf{X}'' = \mathbf{X} * (\delta(0) - \mathscr{K})$, we have $\mathbf{X} = \mathbf{X}' + \mathbf{X}''$: the data can be perfectly reconstructed by adding the filter outputs. The filter in the first branch has zeros at f_{line} and all its multiples (as illustrated in Fig. 2 panel 2), so $\mathbf{X}'$ is perfectly devoid of line artifacts.

As a second step, the power-line-contaminated branch $\mathbf{X}''$ is multipled by a denoising matrix $\mathbf{D}$ designed to project out the power-line artifacts. Assuming this is effective, $\mathbf{X}''\mathbf{D}$ too is devoid of line artifacts.

As a final step, the two branches are summed to obtain the clean data:

$$\tilde{\mathbf{X}} = \mathbf{X}' + \mathbf{X}''\mathbf{D}. \quad (2)$$

The denoising matrix **D** is designed using the JD/DSS method (de Cheveigné and Parra, 2014). More precisely, an artifact-enhancing bias filter $\mathscr{B}$ is defined in the frequency domain by setting to one all Fourier coefficients multiple of f_{line}, and to zero all other coefficients. The covariance matrices $\mathbf{C}_0 = \mathbf{X}''^{\top}\mathbf{X}''$ and $\mathbf{C}_1 = (\mathbf{X}''*\mathscr{B})^{\top}(\mathbf{X}''*\mathscr{B})$ are calculated and jointly diagonalized to find **M** such that

$$\mathbf{C}_0\mathbf{M} = \Lambda\mathbf{C}_1\mathbf{M} \quad (3)$$

where Λ is a diagonal matrix. The matrix **M** defines a transform of **X**'' to a column matrix **X**'' **M** of time series ('components') ordered in terms of decreasing artifact power. Assuming that the first d of these components capture all the artifact, the denoising matrix **D** can be obtained as the product of the last $J - d$ colums of **M** by the last $J - d$ rows of its pseudoinverse $\mathbf{M}^{+}$.

It is easy to see that $\tilde{\mathbf{X}}$ contains *no* power line components, as those components are perfectly cancelled in the first term of Eq. (2) by a spectral filter and in the second term by a spatial filter. Other than that, the data are barely affected: the $1/f_{line}$ square-shaped smoothing kernel is applied only to d out of J spatial components, the others $J- d$ remaining unfiltered (Fig. 3, bottom), while the spatial filter **D** is applied only to the line-dominated, high pass part portion **X**'' of the data, the line-free low pass part remaining untouched (Fig. 3, top). In contrast to spectral filtering methods, the clean data are *full-band*. In contrast to spatial filtering methods, they are *full rank*.

Implementation. The method is implemented by the nt_zapline() function of the NoiseTools toolbox (audition.ens.fr/adc/NoiseTools/). Example scripts are available at http://audition.ens.fr/adc/NoiseTools/src/NoiseTools/EXAMPLES/zapline/. The bias filter $\mathscr{B}$ is implemented by a 1024-point FFT, selecting bins at the line frequency and multiples, and setting all others to zero. That filter serves only as a bias: it does not affect the cleaned data. The Algorithm has only one main parameter: d.

3. Results

Simulated data. Simulated multichannel data consisted of a 10000 × 100 matrix of normally distributed random samples, taken to represent a wide-band 'signal', to which was added a periodic waveform (sinusoid half wave rectified and raised to power 3) multiplied by a 1 × 100 random matrix, taken to represent 'line noise'. Signal and noise were scaled to approximately the same power. The Algorithm was applied with $d = 1$. Fig. 4 (top) shows the normalized PSD of the mixture (left), and of the denoised data (right, green). The part removed, attributed to line noise, is plotted in red. The PSD of the denoised data is flat, as expected for uncorrelated noise, and contains no trace of the simulated line noise. The line noise instead shows prominently in the part removed. The power spectral 'floor' of the part removed is about two orders sof magnitude below the denoised signal. This indicates that the amount by which the signal was changed at frequencies other than line and harmonics is minimal, so any deleterious effect of processing is small.

EEG EEG data consisted of 136 channels of data from a BioSemi system sampled at 512 Hz, taken from a study on auditory perception of speech (Di Liberto et al., 2015). The normalized PSD (Fig. 4 middle left) shows a peak at 50 Hz and odd harmonics, that are prominent in the part removed (right, red) but absent in the denoised data (right, green). The Algorithm here was applied with $d = 2$ (two spatial dimensions were removed from **X**'').

MEG MEG data consisted of 166 channels of data from a Yokogawa system downsampled to 200 Hz. These data are typical of MEG data in that they are strongly contaminated by power line noise (Fig. 4 bottom left). Power line components are absent in the denoised data (right, green) but prominent in the part removed (right, red). Interestingly, the analysis reveals additional weak spectral lines, some of which seem to reflect aliasing, that were not visible in the original data. The Algorithm here was applied with $d = 4$ (four spatial dimensions were removed from **X**'').

Additional examples, together with code and links to data, are provided at http://audition.ens.fr/adc/NoiseTools/src/NoiseTools/EXAMPLES/zapline/. They help make the point that the method is effective and safe for a wide range of data.

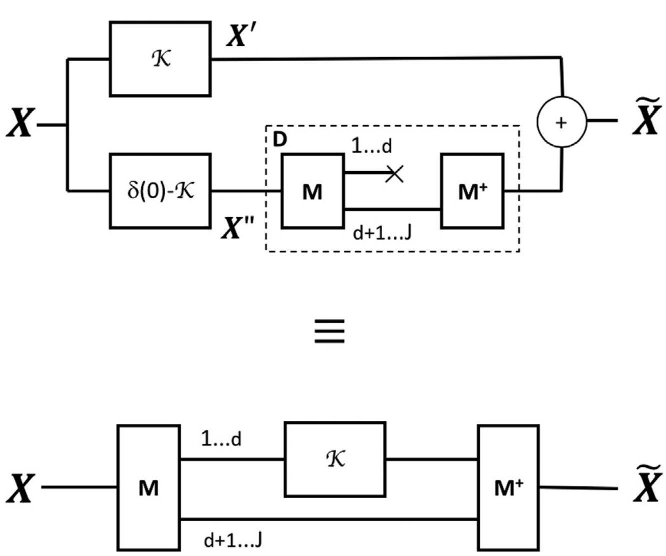

Fig. 3. Top: the signal is split into two branches, one that is spectrally filtered to remove line artifact (upper), the other that is spatially filtered to remove line artifact (lower), then the two branches are combined. Bottom: swapping linear operations ($\mathscr{K}$ and **M**) leads to an equivalent pipeline where it is obvious that d dimensions are spectrally filtered, while the rest are intact.

4. Discussion

The Algorithm is highly effective at removing power line artifacts while preserving non-artifactual parts of the signal. It is computationally cheap and easy to apply, with only one main parameter, d, that specifies the number of spatial components to reject from the artifact-contaminated branch **X**''. Multiple power line components imply multiple sources of artifact, each with a different time course and spatial signature. One might have thought that a single component would suffice to represent the contribution of the power line to brain signals, but empirically values of d between 2 and 4 often give better results. For EEG, it is plausible that different phase shifts arise in capacitive coupling to different electrodes, due to different values of electrode impedance and coupling capacitance. Equipment with different harmonic content (power supplies or fluorescent lights) might contribute additional components. MEG tends to be sensitive to distant sources, and thus often requires larger values of d.

Choosing d too small makes the Algorithm less effective because not all contaminated dimensions are removed. Choosing it too large will cause more spatial dimensions to be affected by the line-suppression filter than needed, which is usually of little consequence. The method is thus attractive as a 'no-nonsense' line-suppression stage within an automated proceessing pipeline. A simple interactive procedure is to start with $d = 1$ and increment until the clean data PSD shows no trace of artifact.

Power line artifacts are detrimental in that they may mask true brain activity, or interfere with processing methods designed to reveal that activity. For example, in an evoked-response experiment, if the timing of trials is inadvertently set such that each trial begins at the same phase of the power line cycle, the line artifact adds up in phase instead of washing out. Worse, if an Algorithm such as DSS is used to isolate repeatable directions within the data, it may return artifact-contaminated directions.

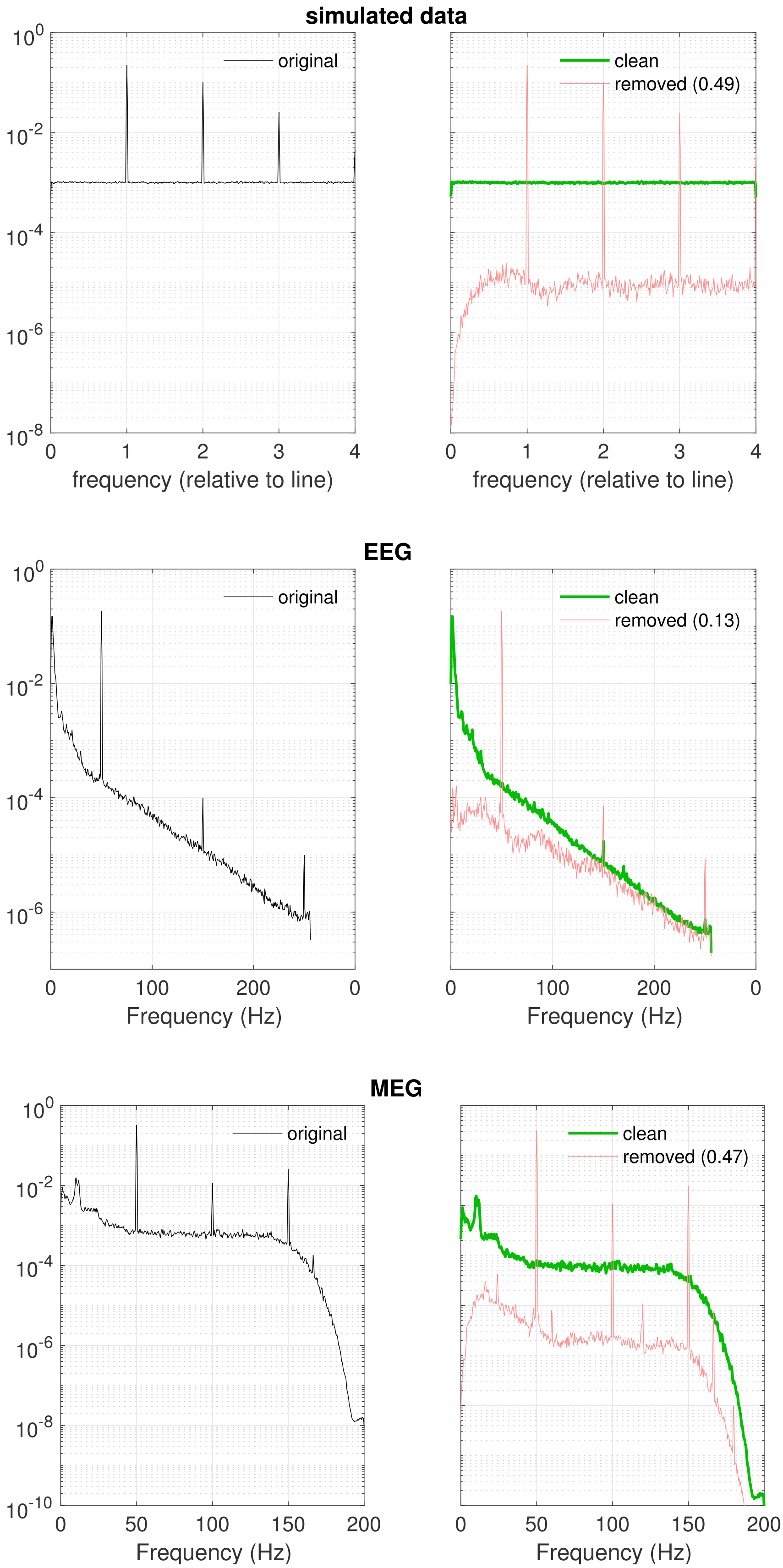

Fig. 4. Normalized PSD of line-contaminated data (left), cleaned data (right, green), and residual (right, red). Top: simulated data, middle: EEG, bottom: MEG.

Artifacts also have an indirect effect in that they motivate processing steps (filtering, etc.) that may distort brain activity patterns.

Other methods to deal with line artifacts also have drawbacks. Low-pass filtering has the obvious effect of suppressing high-frequency patterns, as well as a more insidious effect of distorting brain waveforms (de Cheveigné and Nelken, 2019). Notch filtering has less of a spectral footprint, but the very long impulse response may trigger artifacts at onsets, offsets, and changes in the signal (Vale-Cardoso and Guimarães, 2009; Kıraç et al., 2015). Regression and frequency-domain methods that fit sinusoids to the data are also sensitive to glitches and variations of phase and amplitude of the power line signals.

Spatial filtering methods face a different set of issues. ICA exploits non-Gaussianity, non stationarity, or non-whiteness of sources mixed within the data to find a demixing matrix that isolates them. Isolated components then need to be classified according to the kind of source they represent, for example based on their spectral properties, and removed to obtain clean data. However, the selection process is hard to automate, and furthermore line artifacts may not always be perfectly resolved among the components. Spatial filtering based on DSS/JD yields components directly optimized to maximize power at line frequencies, rather that merely non-gaussianity (de Cheveigné and Parra, 2014). Regardless of which method is used, the cleaned data are rank-deficient, which may be a problem for certain algorithms that expect data to be full-rank. Brain activity that happens to fall within those dimensions is lost, and artifacts that occur within those dimensions may be injected into otherwise clean data.

The method proposed here draws on both spatial and spectral filtering. The problems of the former are mitigated by restricting it to a limited part of the spectrum, and those of the latter by limiting it to a small subset of dimensions within the data. The method draws on a simple artifact rejection filter ($\mathscr{K}$) with the shortest possible impulse response; it can be replaced by some other line-suppressing filter, e.g. low-pass or notch, in the (unlikely) event that a different spectral response is needed.

The principle behind the Algorithm (filter into two complementary paths and apply different spatial filtering to each) is possibly of wider applicability, for example to suppress slow environmental fields in MEG, or attenuate narrowband activity such as alpha while minimizing effects on other brain components. Processing amounts to application of a particular multichannel FIR (finite impulse response) filter that the algorithm discovers. More general ways to find such filters have been suggested (de Cheveigné, 2010), but their greater flexibility is not necessarily an advantage as it can lead to overfitting. The present method is more tightly constrained to 'do the right thing'. Overfitting may nonetheless occur if the number of channels is large, if so it can be avoided by applying PCA to $\mathbf{X}''$ and truncating the series of PCs, before applying DSS.

There is recent interest in methods to process EEG for real-time applications such as steering a hearing aid based on brain signals. The present method is well suited to that context: it involves filtering with a FIR filter of very short impulse response (hence low latency), and a spatial filter that can easily be learned on past data, and if necessary updated to accommodate changes in mixing weights of the line artifact.

5. Conclusion

I presented a simple, effective, and efficient method to remove power line artifacts from multichannel data, applicable to EEG, MEG and other multichannel electrophysiology techniques. The method combines spectral and spatial filtering in such a way as to attain perfect artifact rejection while minimizing deleterious effects. The method requires only minimal tuning, and is therefore an ideal component to include in an automatized data analysis pipeline.

Acknowledgements

This work was supported by grants ANR-10-LABX-0087 IEC, ANR-10-IDEX-0001-02 PSL, and ANR-17-EURE-0017. The manuscript benefitted from comments by Malcolm Slaney and two anonymous reviewers.

References

Barbati, G., Porcaro, C., Zappasodi, F., Rossini, P.M., Tecchio, F., 2004. Optimization of an independent component analysis approach for artifact identification and removal in magnetoencephalographic signals. Clin. Neurophysiol. 115, 1220–1232.

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.M., Robbins, K.A., 2015. The PREP pipeline: standardized preprocessing for large-scale EEG analysis. Front. Neuroinf. 9, 16.

Cohen, M.X., 2014. Analyzing Neural Time Series Data - Theory and Practice. MIT Press.

de Cheveigné, A., 2010. Time-shift denoising source separation. J. Neurosci. Methods 189, 113–120.

de Cheveigné, A., Arzounian, D., 2018. Robust detrending, rereferencing, outlier detection, and inpainting for multichannel data. Neuroimage 172, 903–912.

de Cheveigné, A., Nelken, I., 2019. Filters: when, why, and how (not) to use them. Neuron 102, 280–293.

de Cheveigné, A., Parra, L.C., 2014. Joint decorrelation, a versatile tool for multichannel data analysis. Neuroimage 98, 487–505.

de Cheveigné, A., Simon, J.Z., 2007. Denoising based on time-shift PCA. J. Neurosci. Methods 165, 297–305.

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134, 9–21.

Di Liberto, G.M., O'Sullivan, J.A., Lalor, E.C., 2015. Low-frequency cortical entrainment to speech reflects phoneme-level processing. Curr. Biol. : CB 25, 2457–2465.

Escudero, J., Hornero, R., Abasolo, D., Fernandez, A., Lopez-Coronado, M., 2007. Artifact removal in magnetoencephalogram background activity with independent component analysis. IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 54, 1965–1973.

Gabard-Durnam, L.J., Mendez Leal, A.S., Wilkinson, C.L., Levin, A.R., 2018. The Harvard automated processing pipeline for electroencephalography (HAPPE): standardized processing software for developmental and high-artifact data. Front. Neurosci. 12, 97.

Gramfort, A., Luessi, M., Larson, E., Engemann, D., Strohmeier, D., Brodbeck, C., Parkkonen, L., Hämäläinen, M., 2014. MNE software for processing MEG and EEG data. Neuroimage 86, 446–460.

Gross, J., Baillet, S., Barnes, G.R., Henson, R.N., Hillebrand, A., Jensen, O., Jerbi, K., Litvak, V., Maess, B., Oostenveld, R., Parkkonen, L., Taylor, J.R., van Wassenhove, V., Wibral, M., Schoffelen, J.M., 2013. Good practice for conducting and reporting MEG research. Neuroimage 65, 349–363.

Keshtkaran, M.R., Yang, Z., 2014. A fast, robust algorithm for power line interference cancellation in neural recording. J. Neural Eng. 11, 026017.

Kıraç, L., Vollmar, C., Rémi, J., Loesch, A., Noachtar, S., 2015. Notch filter artefact mimicking high frequency oscillation in epilepsy. Clin. Neurophysiol. 127.

Leske, S., Dalal, S.S., 2019. Reducing power line noise in EEG and MEG data via spectrum interpolation. Neuroimage 189, 763–776.

Luck, S.J., 2005. An Introduction to the Event-Related Potential Technique. MIT Press, Cambridge, MA.

Mitra, P., Pesaran, B., 1999. Analysis of dynamic brain imaging data. Biophys. J. 76, 691–708.

Mullen, T., 2012. NTRC: cleanline. https://www.nitrc.org/projects/cleanline/.

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 2011.

Picton, T.W., Bentin, S., Berg, P., Donchin, E., Hillyard, S.A., Johnson, R., Miller, G.A., Ritter, W., Ruchkin, D.S., Rugg, M.D., Taylor, M.J., 2000. Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. Psychophysiology 37, 127–152.

Vale-Cardoso, A.S., Guimarães, H.N., 2009. The effect of 50/60 hz notch filter application on human and rat ECG recordings. Physiol. Meas. 31, 45–58.

Vrba, J., Robinson, S.E., 2001. Signal processing in magnetoencephalography. Methods 25, 249–271.

Widmann, A., Schröger, E., 2012. Filter effects and filter artifacts in the analysis of electrophysiological data. Front. Psychol. 3, 1–5.