



MRC Cognition
and Brain
Sciences Unit



UNIVERSITY OF
CAMBRIDGE

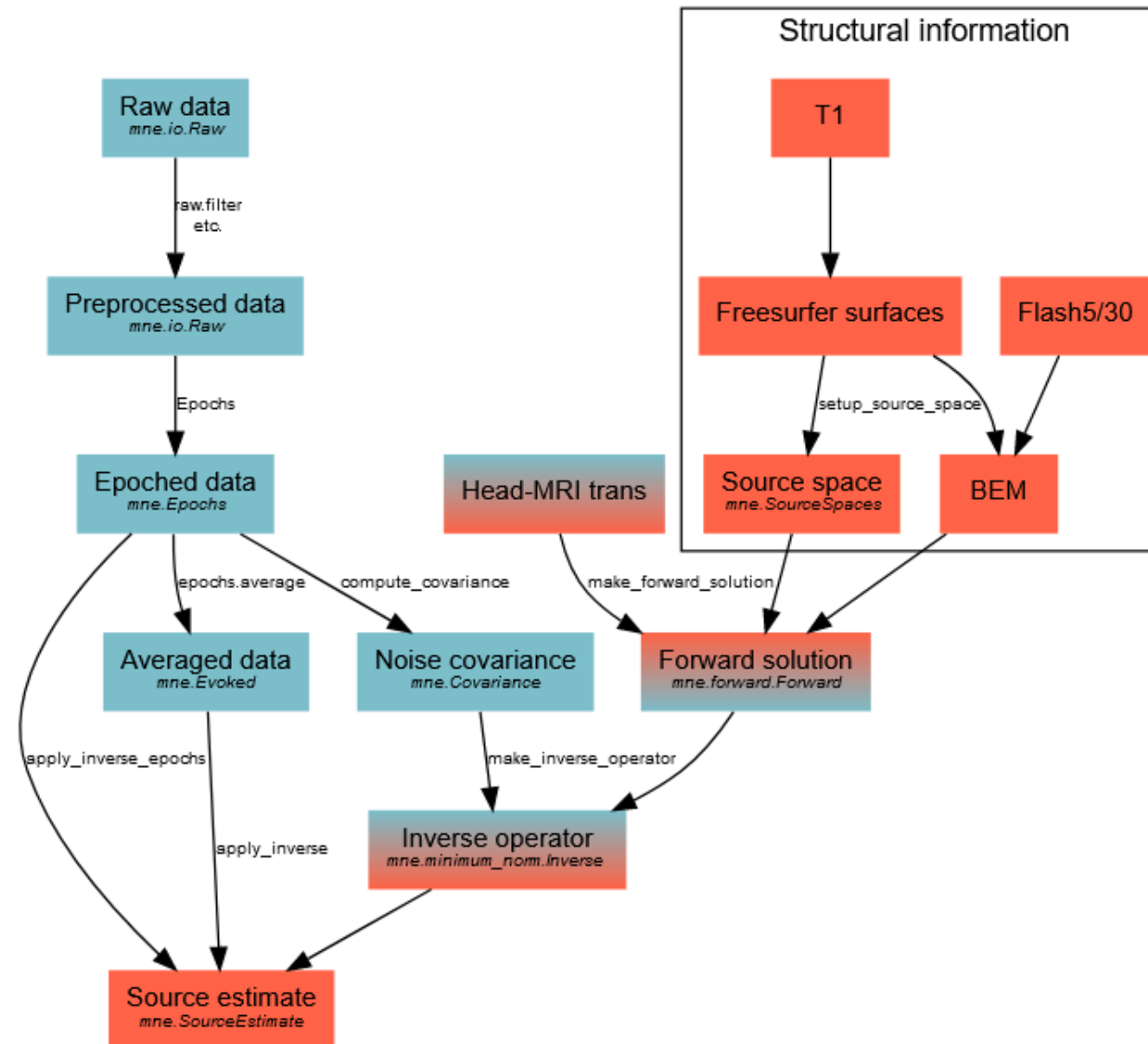
EEG/MEG 1:

Pre-processing

Olaf Hauk

olaf.hauk@mrc-cbu.cam.ac.uk

Typical EEG/MEG Analysis Pipeline



Data Pre-Processing - Artefacts

Artefacts

- **non-physiological**, i.e. from outside the body (sensor-intrinsic noise, line noise, moving objects, vibrations)
=> Maxfilter (SSS), Frequency-Filtering, SSP, PCA/ICA
- **Physiological but non-brain**, e.g. eye movements, muscles
=> SSP, PCA/ICA, H/L-Filtering
- **Physiological from the brain**, i.e. brain sources that are not of interest or not included in your source model
=> choose appropriate source estimation, regularisation

Wisdoms:

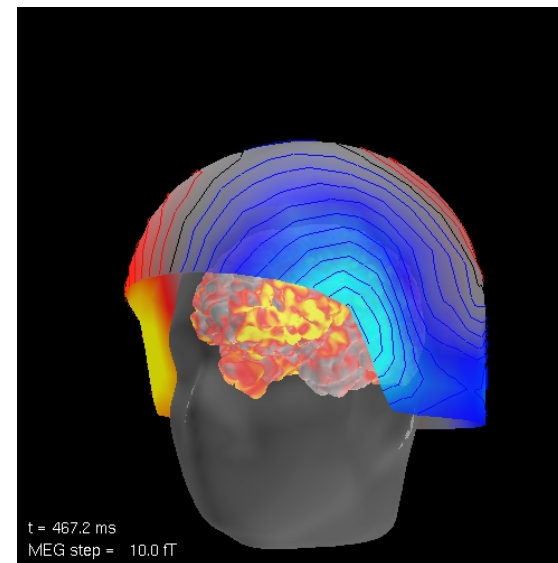
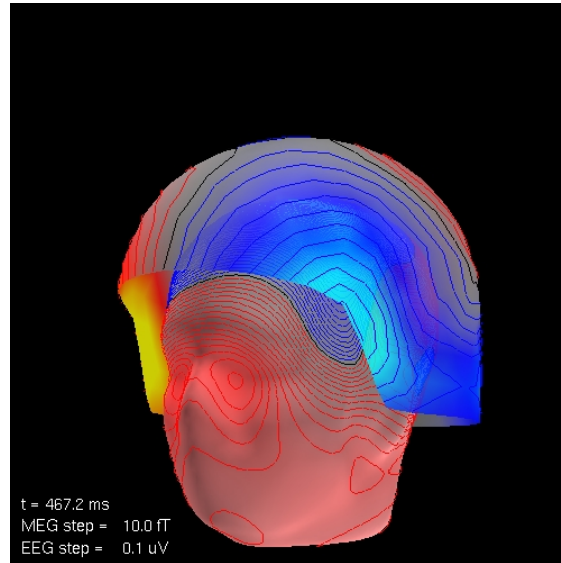
“Some people’s signal is other people’s noise.”

Unfortunately, you cannot just choose what’s signals and what’s noise.

It’s always better to avoid artefacts than to correct them.

Artefacts in EEG and MEG Will End Up in Source Space

Eye Blink



This will affect all source estimation methods –
get rid of your artefacts beforehand.

Separating Signal and Noise Components - “Decompositions”

You can decompose your signal in any way you want – the question is whether it can separate “interesting” from “non-so-interesting” parts of your data, either in space or time.

- 1) Define a set of “basis functions”, e.g.
 - a) A priori: sines/cosines, wavelets, polynomials.
 - b) Data driven: PCA/SVD/ICA, empirical modes.
- 2) Fit these basis functions to your data to explain (most of) its variance.
- 3) Select (or remove) those basis functions that you are (not) interested in, e.g. signal vs noise.
- 4) Reconstruct your signal with the components of interest.

If the basis functions are meaningful, you may be successful: cleaner data.

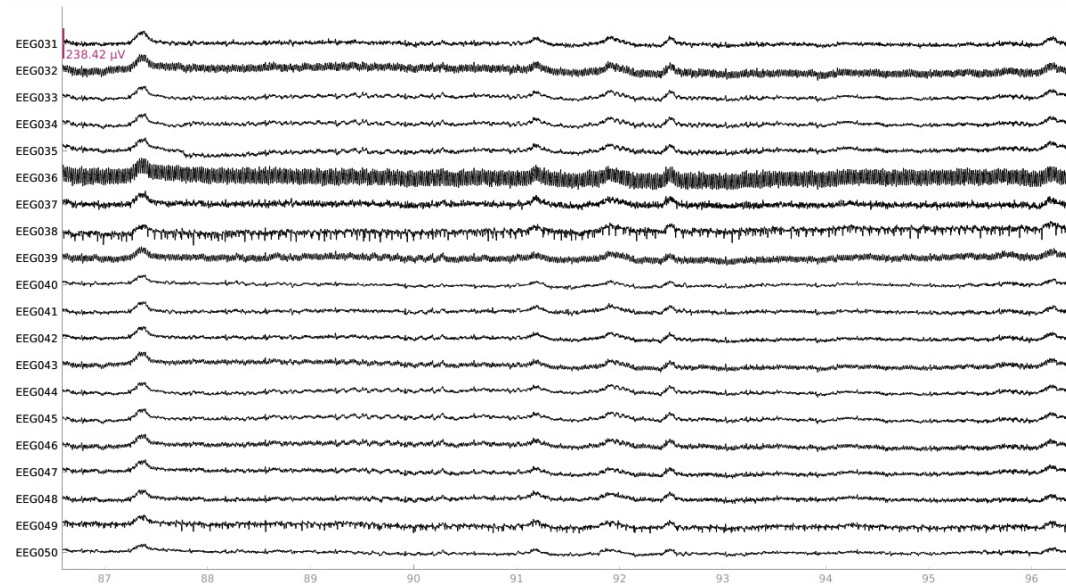
If the basis functions do not correspond to signal and noise parts of your data – you will get a result, but it will be distorted.

Think about what artefacts are the most relevant in your data, then decide which method is best to remove them:

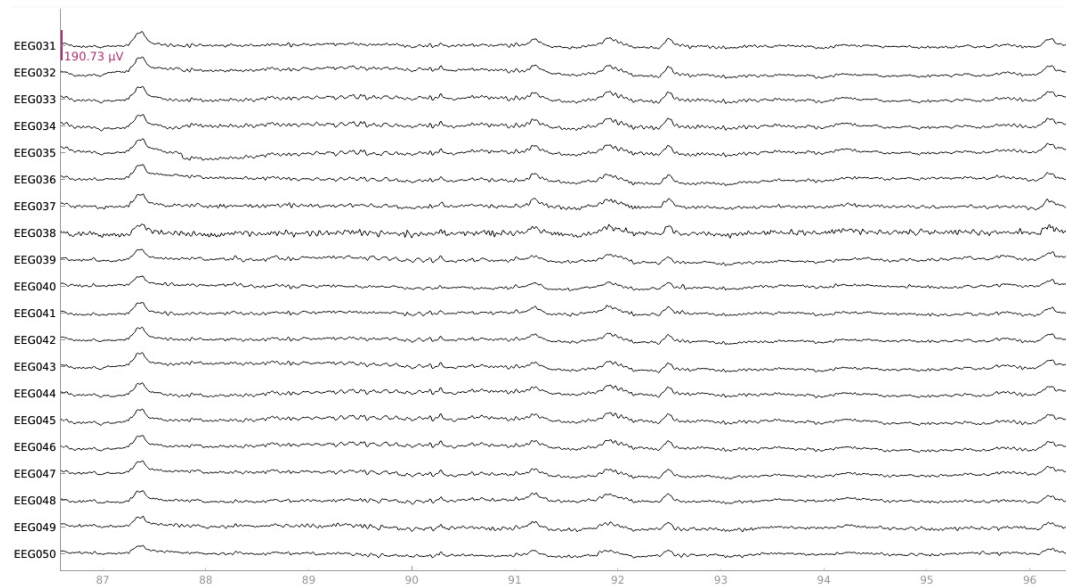
1. How frequent?
2. How large? Time course and topography?
3. Are they related to your variables of interest?

Frequency and Time-Domain Filtering

Raw:

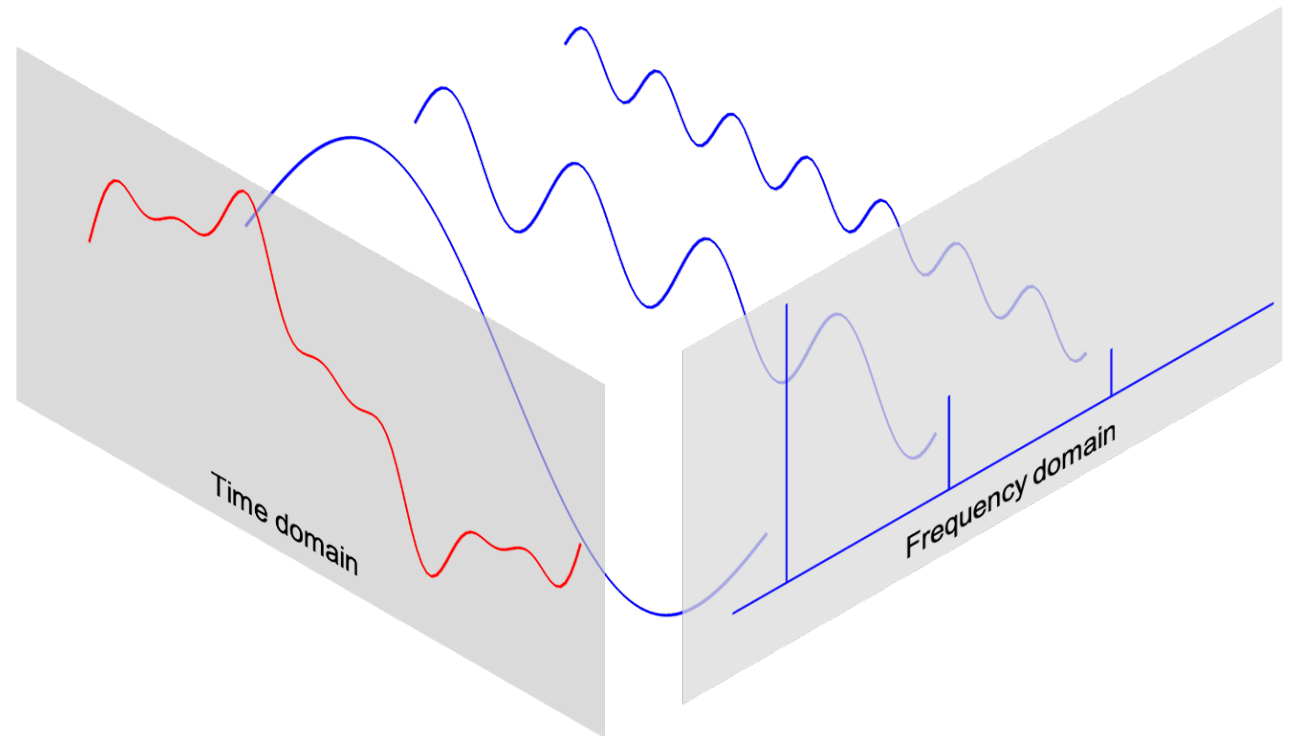
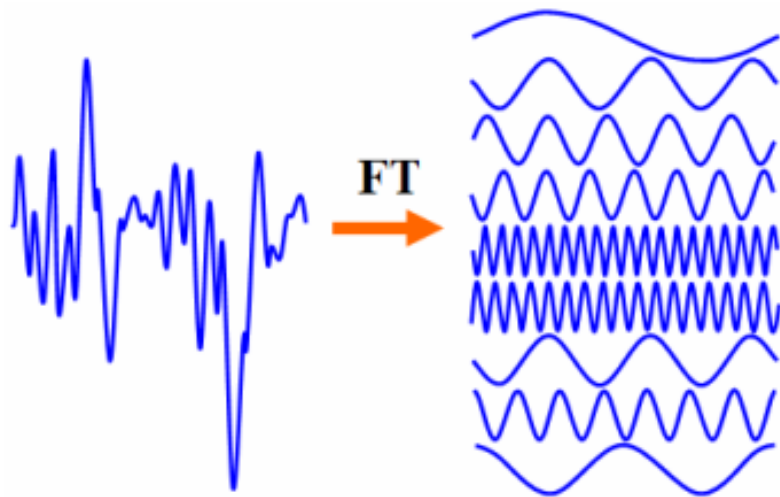


Filtered:



Time-Domain Signals Can Be Represented in the Frequency Domain - and Vice Versa

Decomposing signals
into sine/cosine terms



Basic Principals of Frequency Filtering

If you signals of interest and artefacts occur in separate frequency bands:

- Decompose your signal into its frequency spectrum
- Remove the part of the frequency spectrum that represents artefacts
- Recompose your time domain signal from the remaining frequency spectrum

Examples:

- Line Noise from electrical equipment (50 or 60 Hz): Notch filter
- Muscle artefacts are commonly high frequency (> 30 Hz): Low-pass filter

Basic Principals of Frequency Filtering

Unfortunately, signal and artefacts often overlap in the frequency domain (e.g. eye movements, head movements, heart activity). Thus:

- Avoid artefacts wherever possible
- Control for artefacts (e.g. EOG electrodes, movement parameters)
- Use spatial artefact correction methods where appropriate (ICA, Maxfilter, etc.)

Basic Principals of Frequency Filtering

Filtering changes the time course of your data. Thus:

“Filter as much as necessary but as little as possible.”

Common types of filters:

“High-pass”: Lets higher frequencies pass, suppresses lower frequencies (incl. “detrending”)

“Low-pass”: Lets lower frequencies pass, removes higher frequencies

“Band-pass”: Lets frequencies within a frequency band pass, suppresses frequencies above and below the band

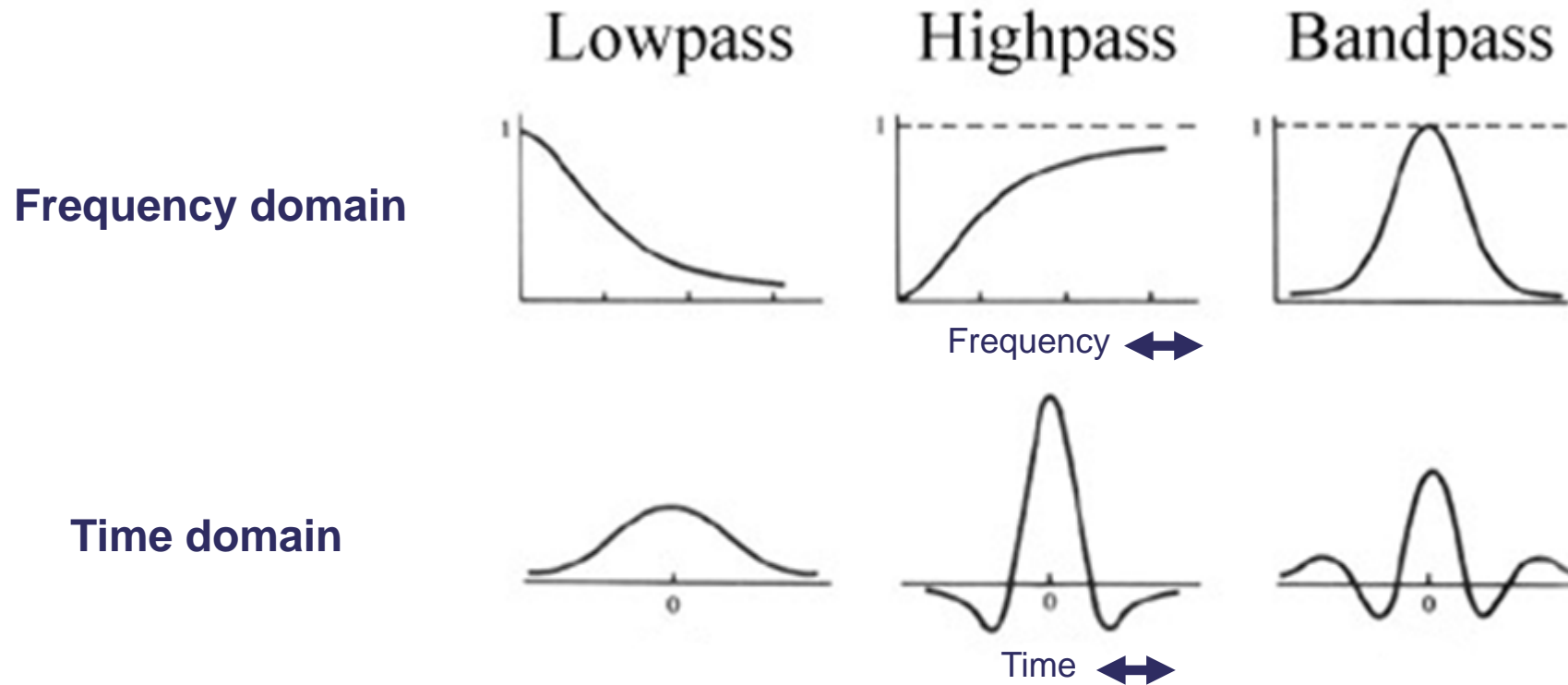
“Notch” filter: A very sharp band-pass filter, e.g. for 50 or 60 Hz line noise

(e.g. Cheveigen & Nelken, Neuron 2019, <https://www.sciencedirect.com/science/article/pii/S0896627319301746>), Widmann et al., Journal of Neuroscience Methods 2015, <https://www.sciencedirect.com/science/article/pii/S0165027014002866>, Tanner et al., Psychophysiology 2016, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4506207/>).

Basic Principals of Frequency Filtering

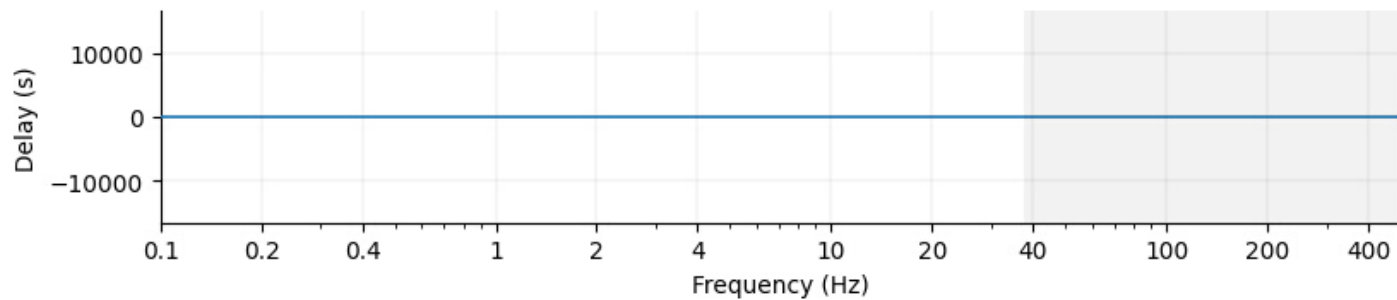
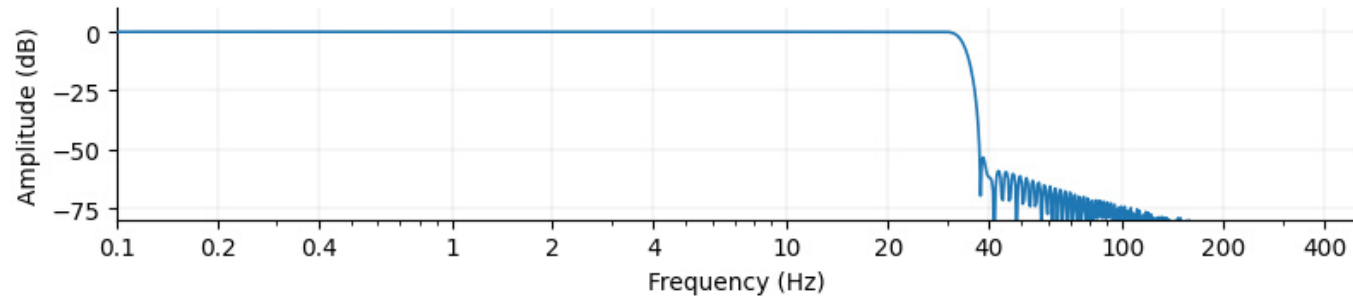
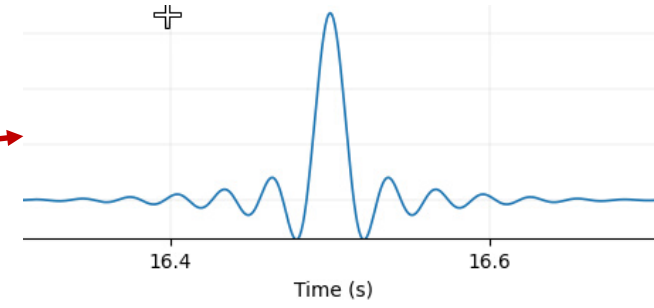
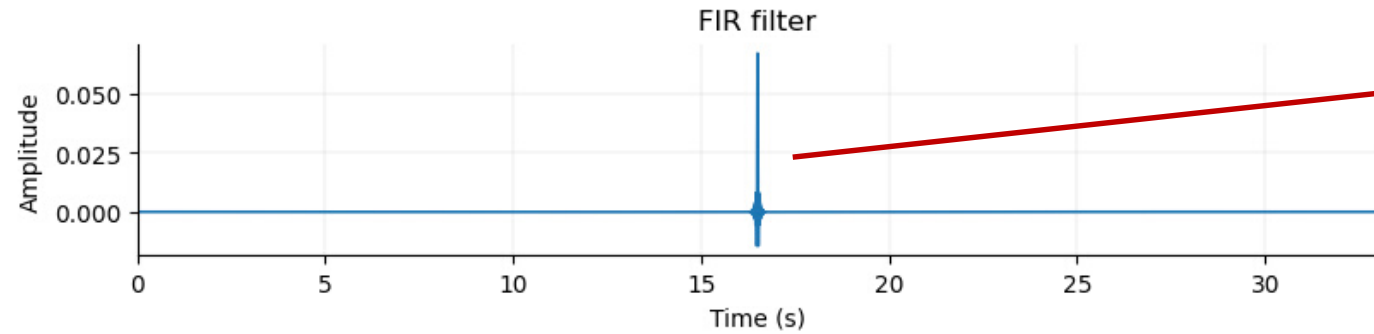
Time-domain and frequency-domain filtering are two sides of the same coin:

One type of frequency-domain filtering corresponds to one type of time-domain filtering.



A typical filter

Impulse response



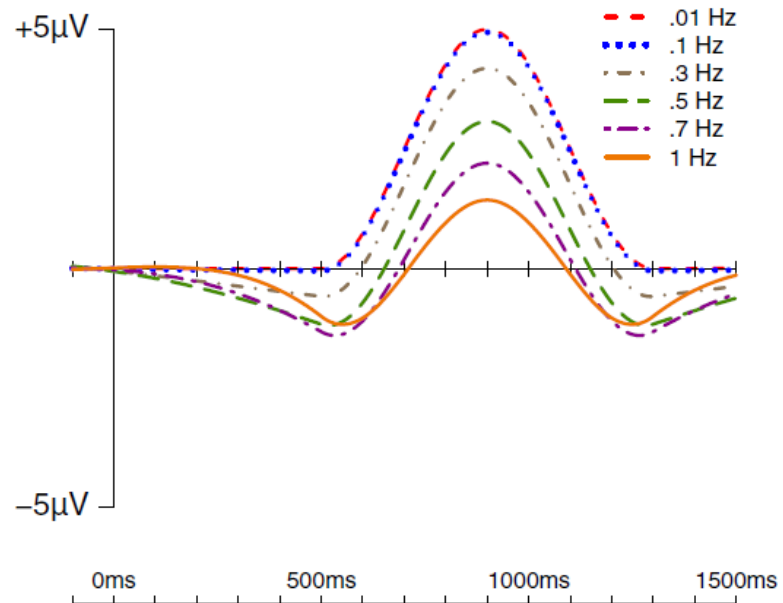
Filter settings:

- Frequency range 0.1-30 Hz
- Sampling frequency 1000 Hz
- Zero-phase FIR
- 4th-order Butterworth
- Hamming window
- ...

Filtering can affect both signal and artefact

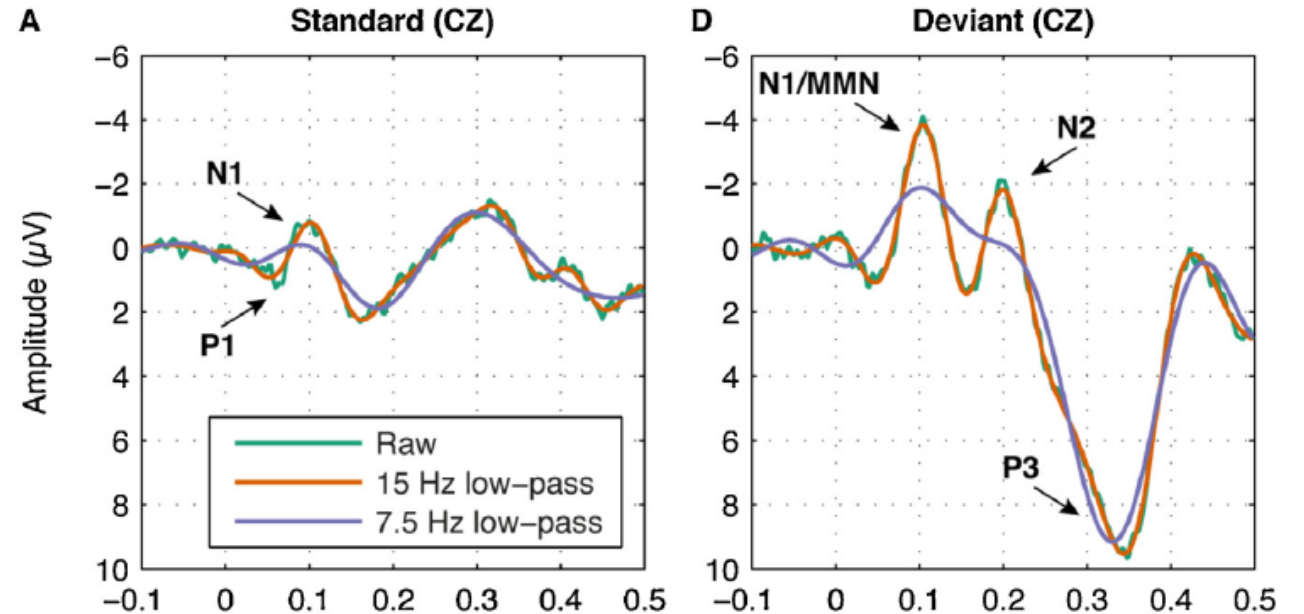
High-pass filtering:

“(linear/polynomial) Detrending”
“Removing slow drifts”



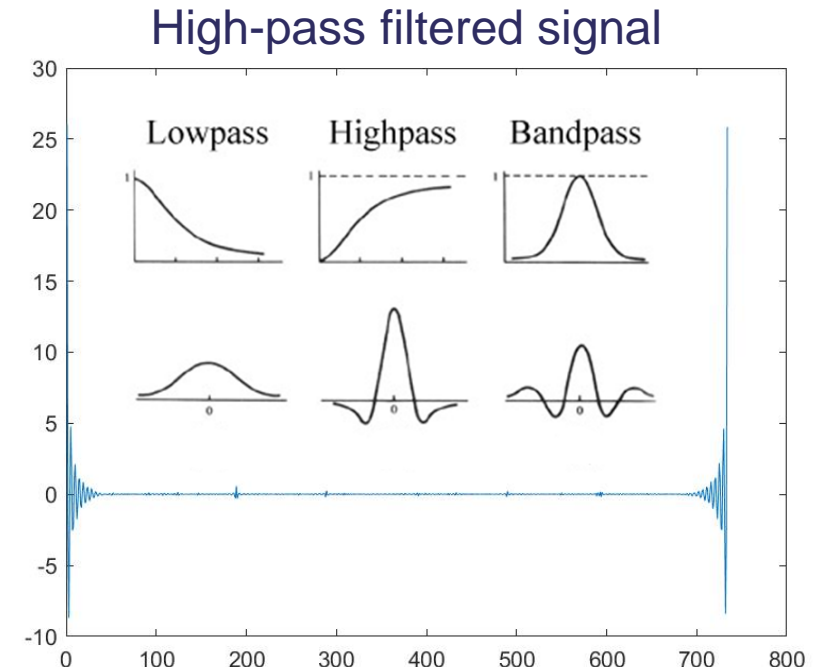
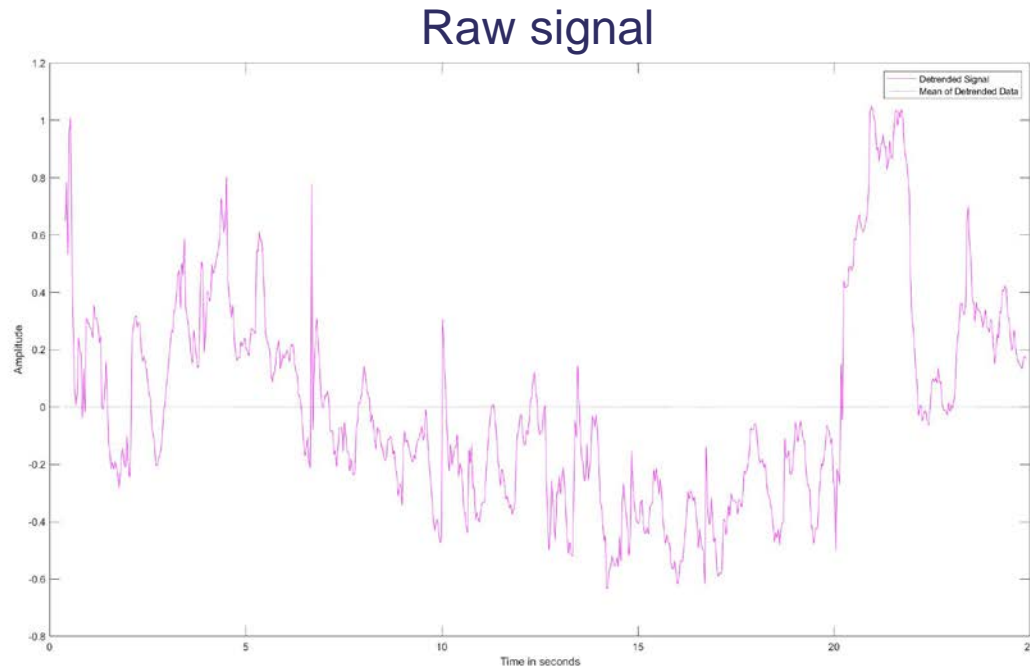
Low-pass filtering:

“Smoothing”



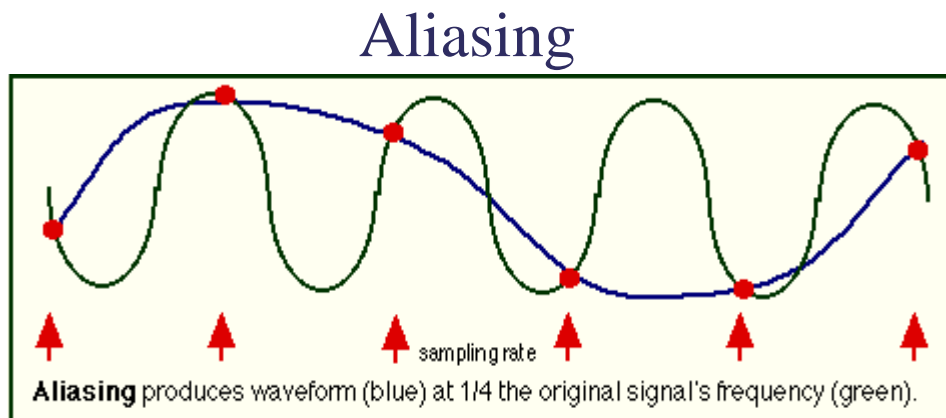
Edge Artefacts of Filters

- Filtering artefacts occur at signal discontinuities, e.g. at the beginning and the end of the data.
- Thus, filter the “longest possible data segment”, ideally the raw data as early as possible.
- If you have to filter epochs, consider filtering longer epochs than you actually need.
- Be careful with “effects” close to the border of epochs.



Filtering and Downsampling: “Aliasing”

- Downsampling can lead to “aliasing” if the data are not filtered appropriately (Nyquist theorem):
Filter at least below half of the sampling frequency before downsampling.



Also watch:

<https://www.youtube.com/watch?v=R-IVw8OKjvQ>

Thanks to Alessandro.

Decibels

- In signal processing, the effect of filters are often described in “decibels”.
- The decibel describes the relative strength of signals (usually their power) on a logarithmic scale.
- It is often used to describe the attenuation effect of filters relative to the original signals (across frequencies).

$$L(p) = 10\log_{10}\left(\frac{P}{P_0}\right) = 20\log_{10}\left(\frac{\sqrt{P}}{\sqrt{P_0}}\right)$$

Logarithms turn multiplication into addition: $L(\alpha p_1) = L(\alpha) + L(p_1)$

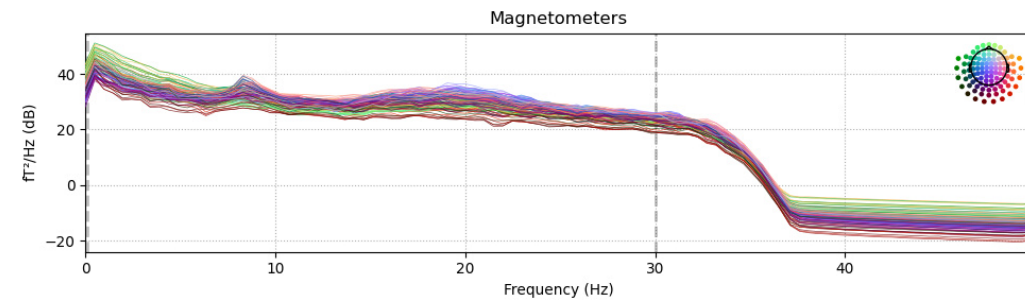
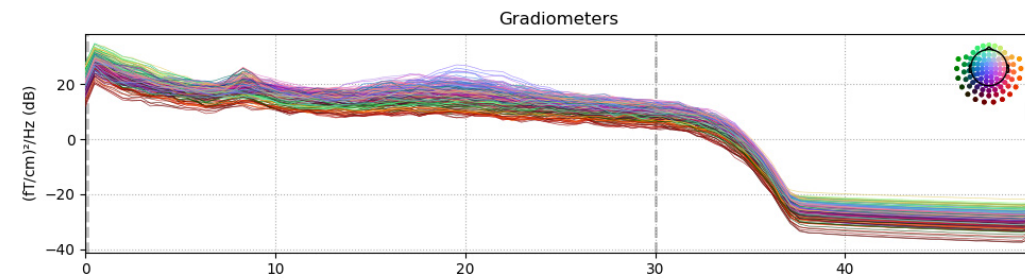
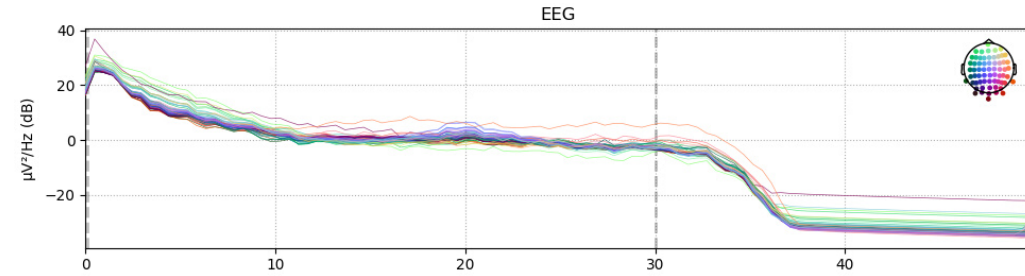
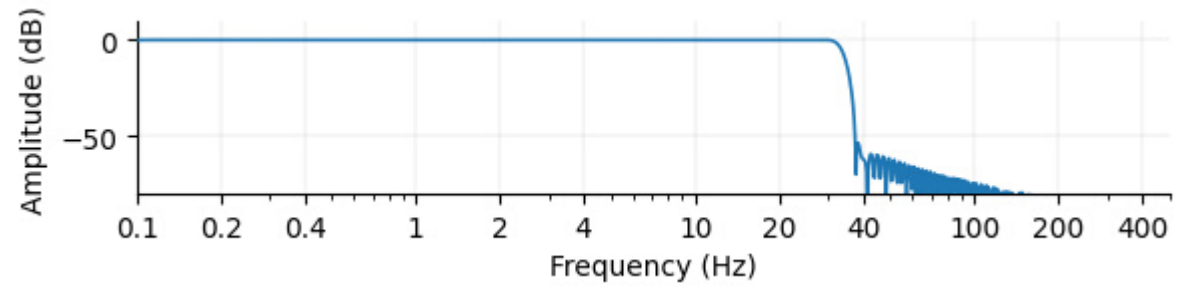
Doubling the power ratio $\frac{P}{P_0}$ increases L_P by 6.

What about a 10-fold increase?

dB	Power ratio	Amplitude ratio
100	10000000000	100000
90	1000000000	31623
80	100000000	10000
70	10000000	3162
60	1000000	1000
50	100000	316 .2
40	10000	100
30	1000	31 .62
20	100	10
10	10	3 .162
6	3 .981 ≈ 4	1 .995 ≈ 2
3	1 .995 ≈ 2	1 .413 ≈ √2
1	1 .259	1 .122
0	1	1
-1	0 .794	0 .891
-3	0 .501 ≈ 1/2	0 .708 ≈ 1/√2
-6	0 .251 ≈ 1/4	0 .501 ≈ 1/2
-10	0 .1	0 .3162
-20	0 .01	0 .1
-30	0 .001	0 .03162
-40	0 .0001	0 .01
-50	0 .00001	0 .003162
-60	0 .000001	0 .001
-70	0 .0000001	0 .0003162
-80	0 .00000001	0 .0001
-90	0 .000000001	0 .00003162
-100	0 .0000000001	0 .00001

An example scale showing power ratios x , amplitude ratios \sqrt{x} , and dB equivalents $10 \log_{10} x$.

Decibels



Spatial Data Decompositions: Spatial Filtering

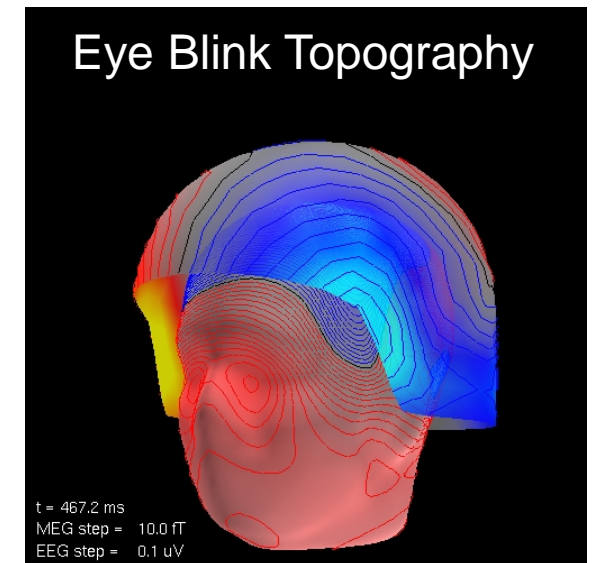
Spatial Data Decompositions: SSP and ICA

If artefacts have characteristic topographies, several methods can be applied to remove (some) noise or extract signals:

- **SSP: Signal Space Projection** (needs pre-defined topographies)

The following often go under the term “blind source separation”, because the topographies are not pre-defined, and found by the methods themselves (under certain assumptions):

- PCA: Principal Component Analysis
- SVD: Singular Value Decomposition
- **ICA: Independent Component Analysis**



Signal Space Projection (SSP)

You know the artefact topography **N** and regress it out of your data.

You decompose your data **D**, such that

$$\mathbf{D} = \mathbf{a} * \mathbf{N} + \mathbf{Signal}$$

You only analyse **Signal**.

This can work well with eye-movement and blink artefacts.

Artefact topographies have to be extracted from the data, or measured separately.

Note:

Brain signals whose topographies are highly correlated with **N** will also be removed or attenuated.

PCA and SVD

- Decompose data into **orthogonal** components \mathbf{T}_1 , \mathbf{T}_2 , etc. (topographies or time courses), i.e. data $\mathbf{D} = a*\mathbf{T}_1 + b*\mathbf{T}_2 + \dots$
- Find the components you don't like (e.g. correlate highly with EOG and ECG, or components that explain little variance).
- Reconstitute your data only with the “good” components,
e.g. $\mathbf{D} = a*\mathbf{T}_1 + c*\mathbf{T}_3 + \dots$ if component 2 reflects eye blinks.

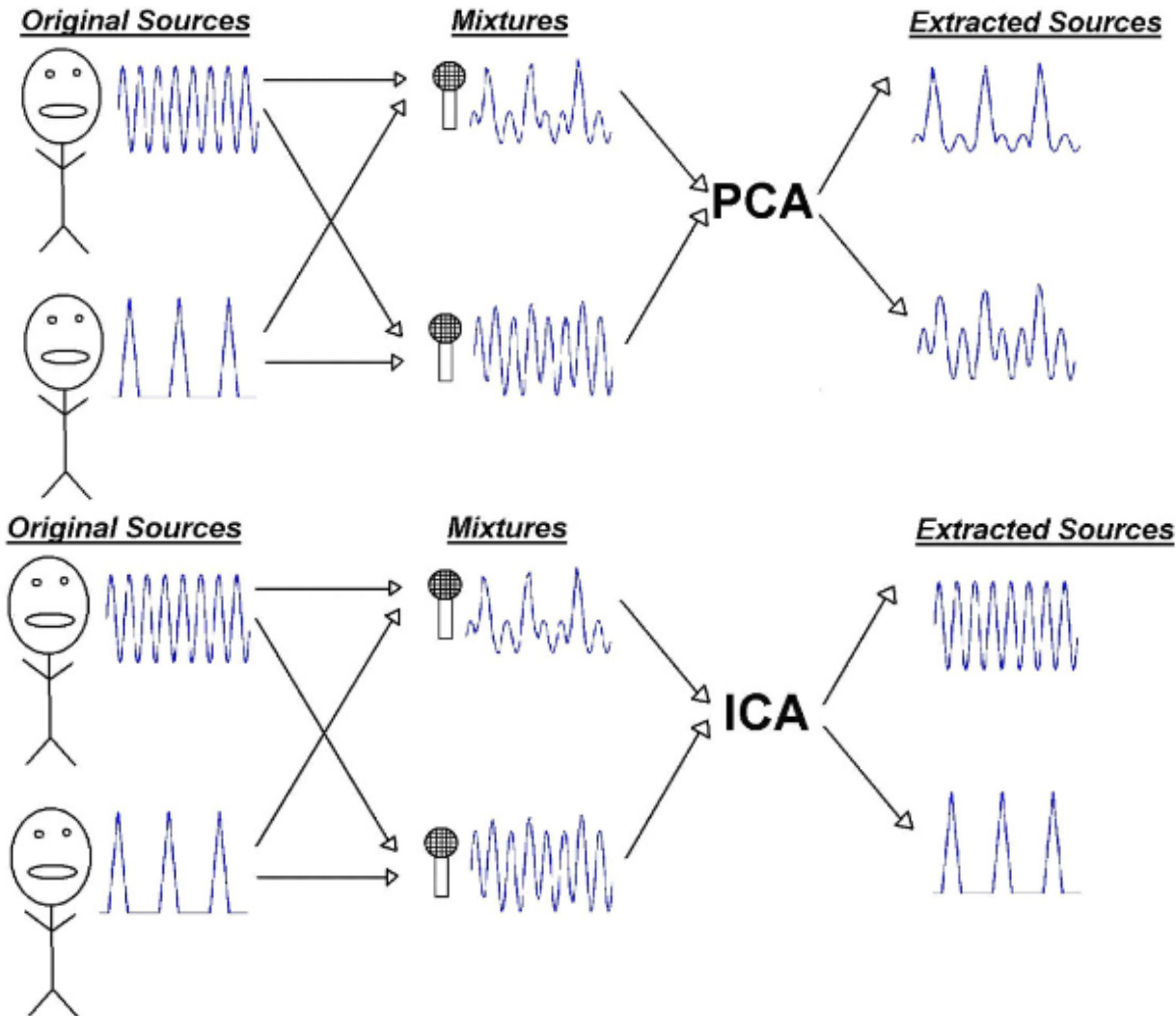
Also:

- Components have an order according to the variance they explain (e.g. $\text{var}(\mathbf{T}_1) > \text{var}(\mathbf{T}_2) > \dots$). Can be useful for exploratory purposes.
- Can be used to determine the number of independent components (according to specified criteria).
- Relatively fast.

- **Unfortunately: Orthogonality and variance ordering is not physiologically plausible.**

Independent Component Analysis

Example: (De-)mixing of sources in the cocktail party effect



Independent Component Analysis

Basic idea is similar to PCA and SVD:

Decompose data into components \mathbf{T}_1 , \mathbf{T}_2 , etc. (topographies or time courses), i.e. data

$$\mathbf{D} = \mathbf{a} * \mathbf{T}_1 + \mathbf{b} * \mathbf{T}_2 + \dots$$

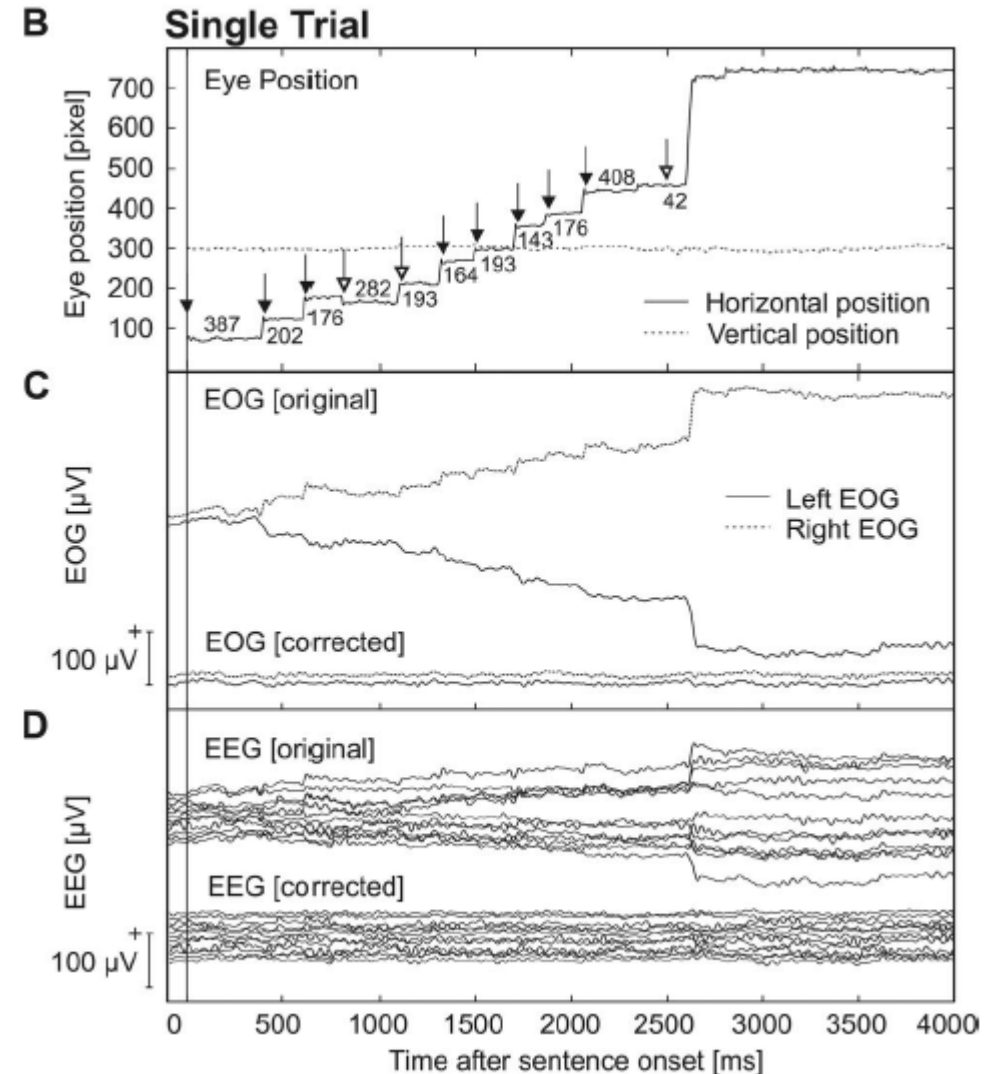
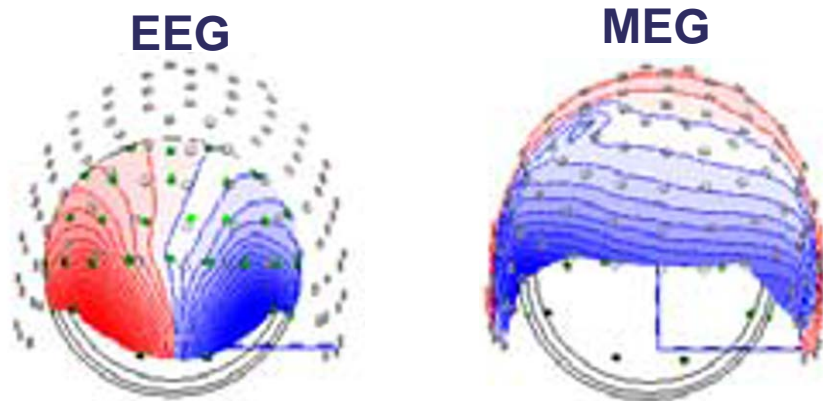
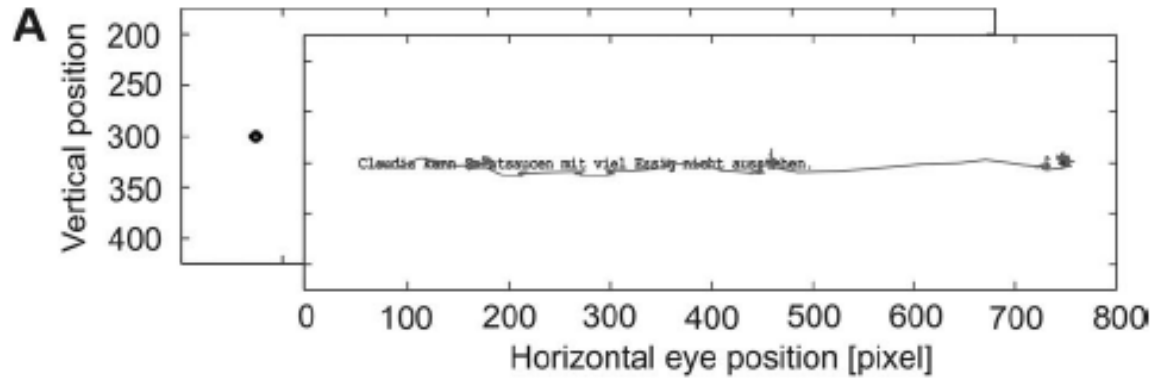
But:

ICA does not produce orthogonal components, and does not assume Gaussianity of signals.

There are number of ICA algorithms available that have been optimised for EEG/MEG data.

They usually work well for example to remove eye movement and heart beat artefacts.

Saccades

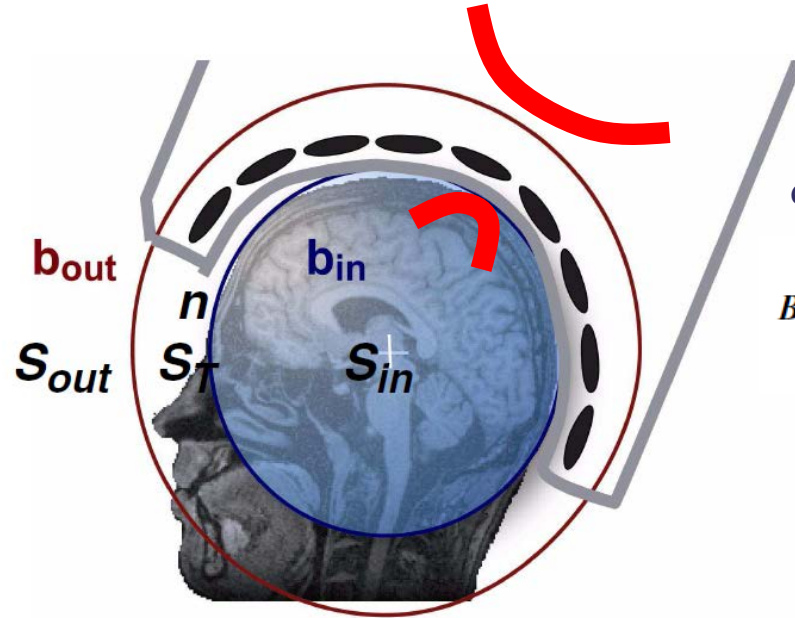


Dimigen, JEP-G 2011, <https://pubmed.ncbi.nlm.nih.gov/21744985/>

“Optimizing the ICA-based removal of ocular EEG artifacts from free viewing experiments”
Dimigen, Neuroimage 2020, <https://www.sciencedirect.com/science/article/pii/S1053811919307086>

“Maxfilter” – Maxwell Filtering

Suppressing Signals From Distant Sources (MEG only)



$$\mathbf{b} = \mathbf{b}_{in} + \mathbf{b}_{out} + \mathbf{n}$$

The mathematical basis of Maxfilter:

decomposition of magnetic fields into spherical harmonics

$$\mathbf{B}(r) = -\mu_o \sum_{n=1}^{\infty} \sum_{m=-n}^n \alpha_{nm} \frac{v_{nm}(\theta, \varphi)}{r^{n+2}} - \mu_o \sum_{n=1}^{\infty} \sum_{m=-n}^n \beta_{nm} r^{n-1} \omega_{nm}(\theta, \varphi).$$

$$v_{nm}(\theta, \varphi) = -(n+1)Y_{nm}e_r + \frac{\partial Y_{nm}}{\partial \theta}e_{\theta} + \frac{imY_{nm}}{\sin \theta}e_{\varphi},$$

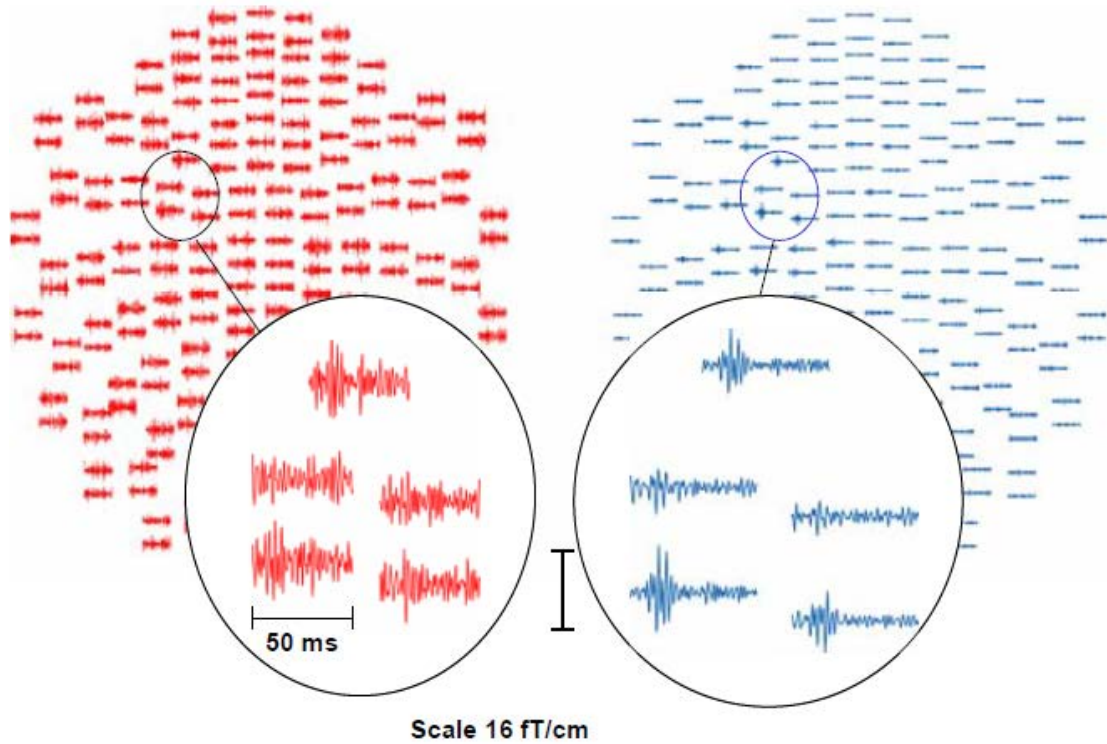
$$\omega_{nm}(\theta, \varphi) = nY_{nm}e_r + \frac{\partial Y_{nm}}{\partial \theta}e_{\theta} + \frac{imY_{nm}}{\sin \theta}e_{\varphi},$$

The measured magnetic field distribution is decomposed into “inside” (the helmet) and “outside” components, and the outside components are removed.

Maxfilter – Signal Space Separation

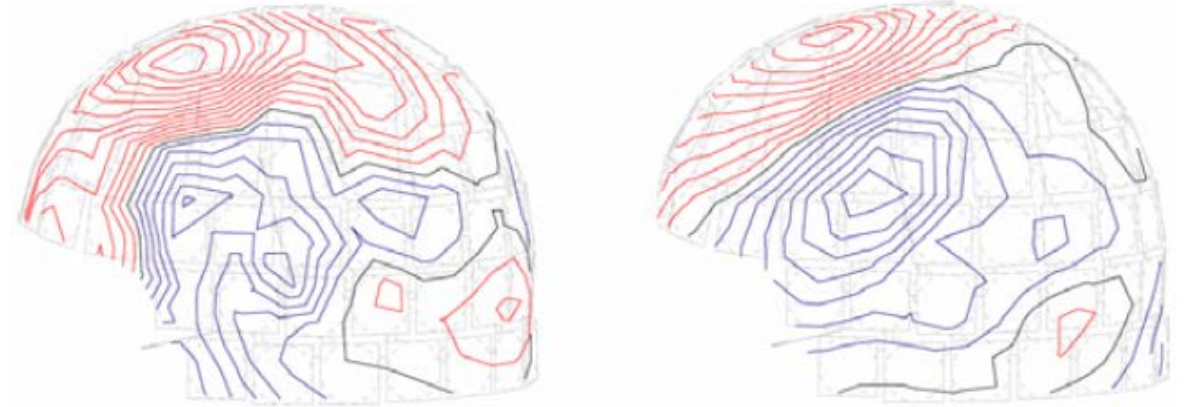
Without

With



Without

With



Original Field Map

SSS Reconstructed Field Map

Latency 20 ms
 $Q = 2 \text{ nAm}$

Maxfilter Software

Software shielding (Signal Space Separation, SSS)

By subtracting the outer SSS components from measured signals, the program suppresses artifacts from distance sources.

Automated detection of bad channels

By comparing the reconstructed sum with measured signals, the program can automatically detect if there are MEG channels with bad data that need to be excluded from Maxwell-filtering.

Spatio-temporal suppression of artifacts (“-st”)

By correlation the time courses of SSS artefact components with the cleaned signal, the program can identify and suppress further artefacts that arise close to the sensor array.

Notch Filter to remove 50/60 Hz line noise.

Transformation of MEG data between different head positions (“-trans”)

By transforming the inner components into harmonic amplitudes (i.e. virtual channels), MEG signals in a different head position can be estimated easily.

Compensation of disturbances caused by head movements (“-movecomp”)

By extracting head position indicator (HPI) signals applied continuously during a measurement, the data transformation capability is utilized to estimate the corresponding MEG signals in a static reference head position.

Maxfilter – Movement Compensation

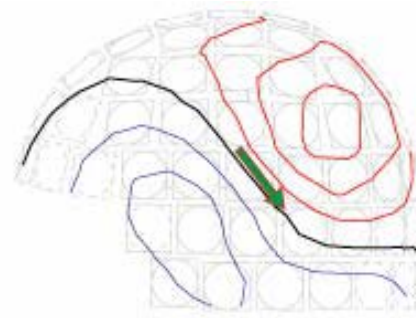
Head movement is tracked continuously (well, every 200 ms) via HPI (Head Position Indicator) coils.

We can take Maxfilter parameters from any time point t , and estimate the MEG signals at sensor positions of time point t_0 .

This compensates – to some degree – for spatial variation caused by head movements.



Stable subject



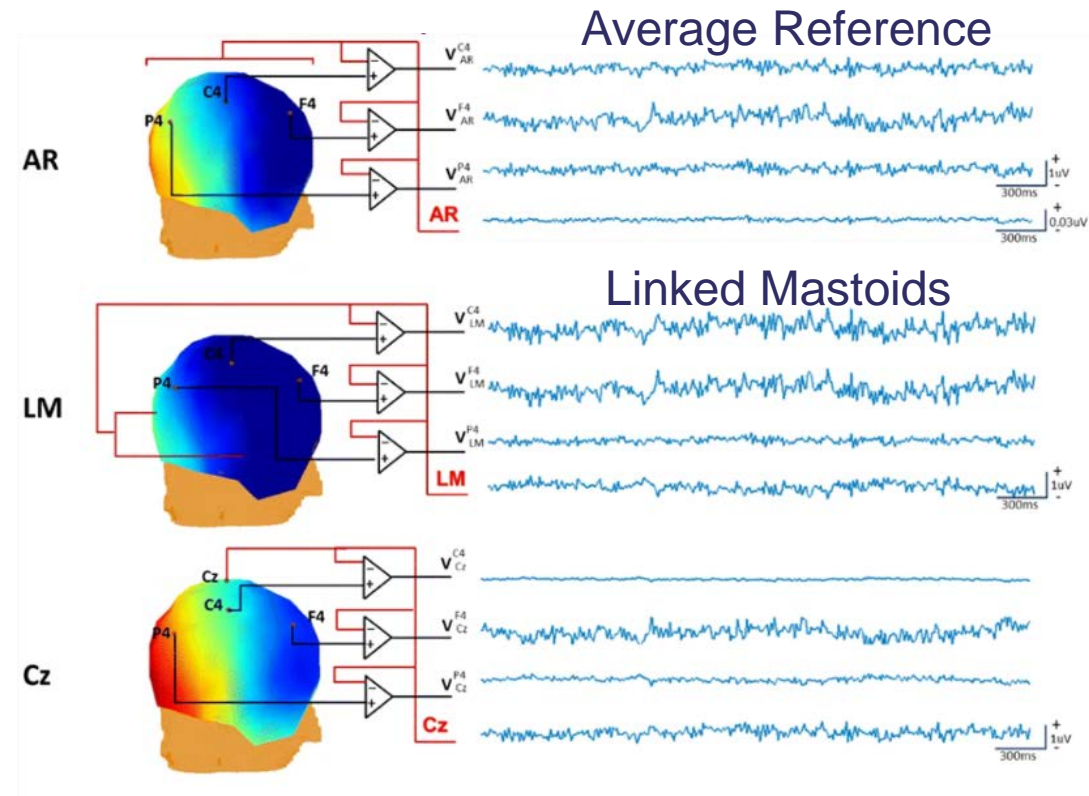
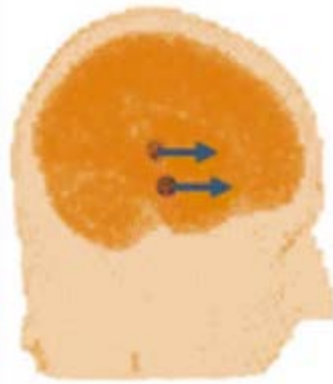
Moving subject,
No compensation



Moving subject,
with compensation

EEG only: Choice of reference site

Data from
two simulated dipoles



The choice of reference changes time course and topography. For high-density recordings (> 65 channels), average reference is recommended. Note: Source estimates do not depend on the reference.



MRC Cognition
and Brain
Sciences Unit



UNIVERSITY OF
CAMBRIDGE

Thank you