Summary NeuroEngineering [F.Cincotti]

(2020-2021)

Electroencephalogram – 1

Everything started in the first years of '900 when Hans Berger was able to demonstrate a large and persistent 10 Hz rhythm that he named (ongoing activity).

In 1930 Lord Adrian observed that *alpha* oscillations appeared when the cortex had a little to do (this stands only for *alpha* band signals).

Years later other scientists experimented the **event related potential**, measured as response of a stimulus. This potential is still mixed with the **ongoing activity**, so it is necessary to take more measures and average them to remove then the ongoing activity.

Alpha rhythm is generated in the occipital region of our cortex, which is the first to process images (**primary visual areas**). Whenever we close our eyes, it is decreased.

There exist other rhythms, for instance representing the somatosensory functions. Motor area and somatosensory area can be dealt together to produce the μ rhythm. It is originated in the central area of the brain, and it also oscillates around 10 Hz, still in waveform shape but in a different way.

α is a pure sinewave, μ is a sort of $\sin^2 x$, the so called arc-shaped sinewave.

Note that with the word rhythm we can mean both the functionality and the signal in the spectrum.

As anticipated, there are many rhythms:

- δ <3.5 Hz
- θ 4 – 7.5 Hz
- α 8 – 13 Hz ($\mu\, around\, 10Hz$)
- β 14 – 30 Hz
- γ > 30 Hz

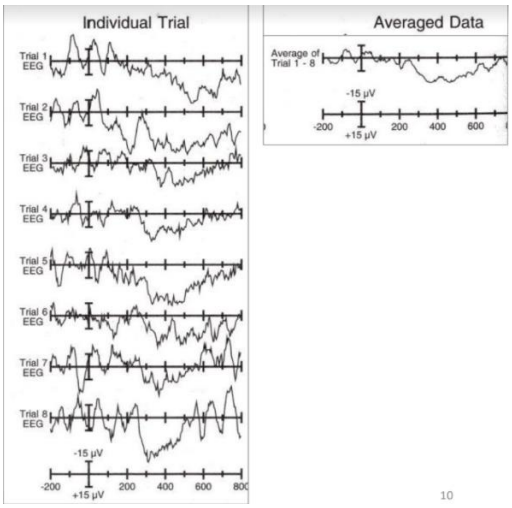
Besides spontaneous brain activity, an EEG can show time-locked(evoked) responses to various sensory stimuli.

Evoked Potentials are **deflection** of the EEG signal following the presentation of a sensory input.

The latency after the stimulus can range from a few milliseconds to several hundreds of milliseconds. Event related potentials can be recorded during tasks involving aspects of motor planning and other cognitive operations.

When you deliver a stimulus a time window around it is called **trial**. **Segmentation** is the activity of extracting trials from the ongoing EEG, so that then it is possible to compare the trials.

To do so, all the trials have to be aligned in the same **latency** (distance between samples). By definition the stimulus is at **zero latency**.



Instrumentation:

In biological systems electricity is conducted through **ions**. The most commonly used EEG electrodes are made of **silver** and **silver chloride**.

Issues on the choice:

- Non-reactivity with biological tissues.
- Accurate transduction.
- Accurate reproduction of extremely slowly changing potentials.
- Low polarization, drift, noise.

An **electrolyte gel** helps the conduction between skin and electrodes, indeed there is an exchange of **electrons** between the gel and the metal of the electrode.

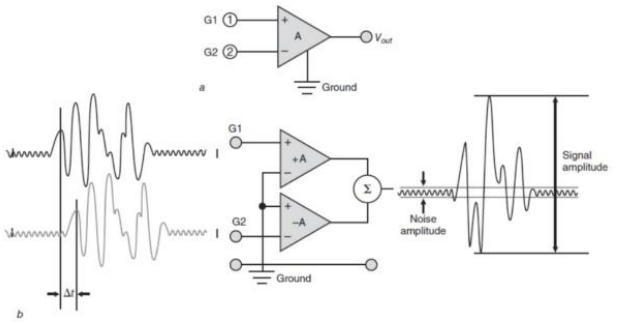
Impedance is a measure of the quality of the contact between the scalp, the electrode, and the conducting medium. It must be measured with alternating current and to improve it is important to **scrub** the gel.

Low values of impedance can be due to:

- Voltage drops.

- Circuit of electrodes with big difference of impedance between electrode. Unbalanced circuit cause noise.

To remove this noise, it is possible to use a third electrode as **ground** and two differential amplifiers processing the two electrode signals. Then sum the outputs.



As it is easy to note, the ground value is **cancelled**, since it enters in the amplifiers once with the plus sign and once with the minus sign (it is only useful to the amplifiers to have a reference). The **G1** [Fig 2] electrode is the **exploring one**, is the one we put on the head to catch neural activity. The **G2** electrode instead is the **reference one**, and as we will see later on, it is usually put in areas in which the neural activity is about zero, but is never really zero (ears, nose…).

The requirements for a good EEG amplifier are:

- Low noise/high gain
- High input **impedance**
- High **CMRR**

Examples of unwanted common-mode signals could be power-line noise or noise arising from a cardiac or nerve stimulator in another part of the body.

- A good amplifier should reject, or strongly suppress, all common-mode signals and amplify only the signals of interest.

CMRR (common-mode rejection ratio) is a measure of interest of the amplifiers. It represents the ratio between the differential gain (how much the difference between the inputs is amplified) and the common-mode gain (how much the common signal in both inputs is amplified). It is expressed in decibels.

$$CMRR = 20\log_{10}\left(\frac{gain\ for\ difference\ signal}{gain\ for\ common-mode\ signal}\right)$$

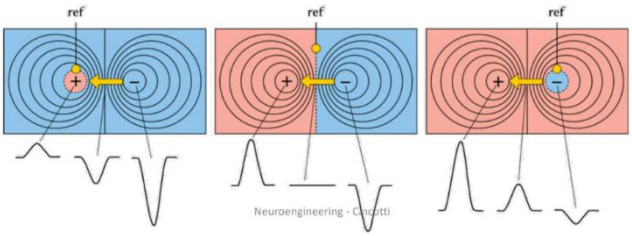
An important feature of the amplifier is its input impedance, which should be many orders of magnitude higher than the electrode one, so that the amplifier remains as insensitive as possible to small changes in electrode impedances.

- Impedances should be low so that the biological potential is not dampened (voltage divider).
- impedances should be comparable across electrodes, because CMRR is reduced by any asymmetry in the measuring circuit.

EEG guidelines suggest to keep impedance below 5 kOhm, but modern amplifiers with input impedance of hundreds of kOhm allow to relax this requirement.

Reference electrodes are typically sited in places that are assumed to be **far from** the putative activity of interest, and traditionally popular places for reference electrodes have been earlobes, mastoids, and the nose.

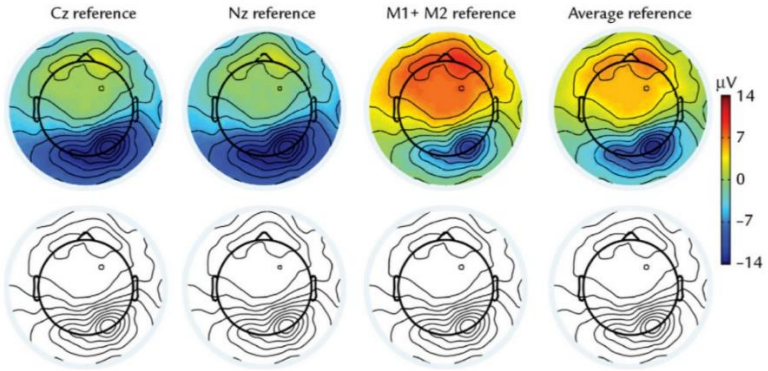
The shape of the **EEG waveforms may change** depending on where the reference electrode is located, whereas the corresponding spatial scalp voltage topographies are not altered.



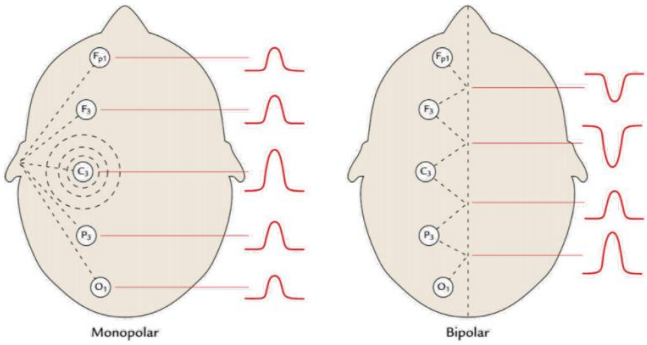
Electroencephalogram – 2

If we are just interested in the profile of the potential, it does not matter where we put the reference electrode(G2). Indeed, it only represents the value we are going to subtract to the exploring electrode(G1), affecting only the absolute value of the potential (just offsetting it by a constant) but not the profile.

For instance, the difference in putting the reference electrode on the ear or on the nose, is just the constant value that changes between these two areas.



This approach based on the presence of a single reference electrode is called **monopolar**. But often, especially in clinical recordings, potentials are measured using neighboring electrodes as reference, this is called **bipolar** configuration.



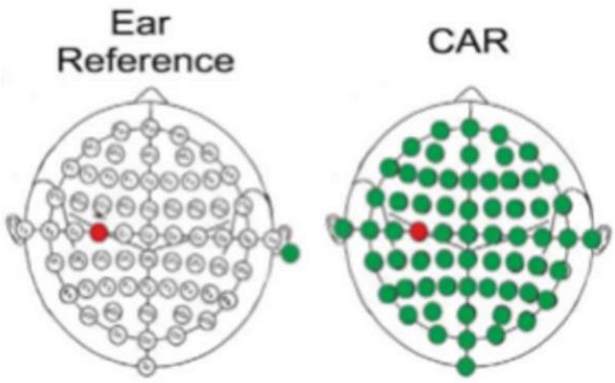
In the right figure in particular is a **longitudinal** bipolar configuration, other configurations are possible.

The reason why this approach is used is that it allows to identify more easily the source of neurological phenomenon.

CAR:

To prevent (reference) electrodes from catching neural activity, we could think to put them outside the head, potentially in an infinite position, but this would generate a lot of noise. So, there is not a position on the head that is free of this problem. An alternative can be derived from the Gauss theorem: *the integral of the potential over the surface of a sphere that contains only concentric inhomogeneities is* ***zero***. So, we can think to use a combination of electrodes.

A problem with the Gauss approach in practice, is that we do not have access to the integral of the potential on the whole head surface; we can substitute it with an average (sampling), summing the potentials and dividing by their number. Moreover, we should integrate (or average in our case) on the **whole** surface, but we have access only to the upper part of the head.



Accepted these two approximations, we can define the **common average reference.**

According to this approach, instead of using the monopolar approach, we consider the exploring electrode and subtract the average of all the others.

General principles of good experimentation

- Record EEG data from alert and cooperative subjects in as artifact and noise-free conditions as possible.
- Divide the recording session to smaller "runs," or blocks.
- Remind subjects about relaxing their muscles, minimizing blinking, and avoiding head and body movements.
- Monitor the brain signals continually online for any artifacts or technical issues.

Electrodes, skin preparation and electrode-impedance measurement

• General subject preparation: ask subjects to wash their hair (and skin) prior to arriving for an EEG study and to refrain from applying any cosmetic products on the skin or hair/scalp,

• Skin preparation and electrode application:

  - Most electrodes have a hole at the top where conductive gel can be inserted with a blunted needle/syringe
  - The blunted needle also allows light abrasion of the skin of the scalp to be performed, if needed.
  - Following skin preparation, impedances of less than 5-10 kΩ can be easily obtained

• Electrode-impedance measurement.

Artifacts

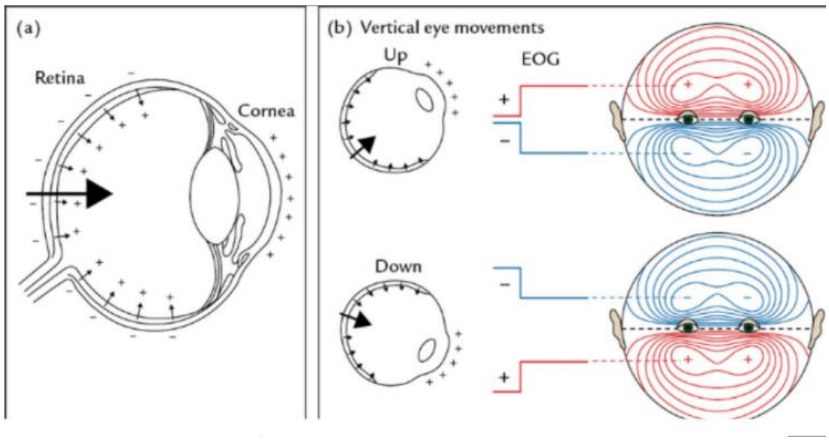
An artifact in EEG can be defined as any potential difference due to an **extracerebral** source. Two big families:

- Biological, like eye blinks, eye movements (EOG), muscles activity (EMG)…
- Technical, not due to physiological causes, like power supply, spurious electrical, noise from electrical engines…

Moreover, body and head movements may induce not only muscle electrical activity, also slow potential shift due to the physical displacement of the ions double layer on the electrode surface.

Eye related artifacts:

Ocular artifacts arise because the eye is an electrical dipole, where the cornea is positive charged with respect to retina at the back of the eyeball.



During **eye movements**, the eyeball moves within the volume conductor generating artifacts of the order of 0.5mV (note that EEG is measured in microvolts)

During **eye blinks**, the volume conductor changes because of movements of the eyelids that due to their moist inner surface provide a well conducting pathway for current flow. This generates a 0.5mV, monophasic, 200-400 ms.

Simple rule: the electrode you are looking at becomes more positive.

Muscle artifacts:

Muscle contractions are seen as artifacts in the 100μV or 1mV range, with a wide frequency spectrum from tens of Hz to a few kHz (most of the power is around 500Hz). EMG is more and more fast than the normal EEG and the EOG (eyes). It is at the same frequency range as beta(14-30Hz) and gamma(>30Hz) band signals.

Cardiac artifacts:

Even if it is in the order of mV, since heart is not in the head, if we apply all the counter measures to avoid common mode potentials, we do not see so much ECG artifacts. We instead can see the effect of **ballistocardiogram**, that reflects small electrode movements due to a nearby blood vessel in the scalp.

Sweating: it can be a problem in EEG recordings as it is accompanied by electrodermal responses. It has high amplitude slow potentials (<0.5 Hz).

Line noise:

"Line" (power-line) noise arises from the capacitive coupling of alternating current supplied to electrical wall outlets.

The line noise occurs either at 50 Hz (or 60 Hz) and at their harmonics.

- Thus, in the time domain, the signal is not necessarily a sinewave

It is accentuated in EEG electrode pairs that have an impedance mismatch (reduced CMRR)

Unless there are no other alternatives, a digital notch filter set to the appropriate line frequency (and its harmonics, if necessary) can remove this noise.

Electrode-Related Artifacts:

• A number of different artifacts can arise from improperly applied EEG electrodes.

- "Bad electrode" artifacts as very slow baseline drifts.
- Electrode "pop" is an intermittent phenomenon that appears as a sharp, sudden rise followed by a gradual fall in the EEG signal of a single electrode.

• In EEG recordings, artifacts can be created by the experimenter's movements around the subject, and this problem can be exaggerated in dry air environment creating a problem with static electricity.

Movement artifacts:

• At the interface between metal (electrode) and ionic conductor (gel/skin) there is a displacement of ions that resembles a capacitor

• Movement of the electrode (e.g. produced by movements of the subject's head) disturbs the geometry of the «capacitor», producing a relatively strong potential.

Electroencephalogram – 3

First thing to do is to assure the quality of the recordings. So, it is necessary to look at the data as first thing, in order to remove eventually some segments if too much corrupted.

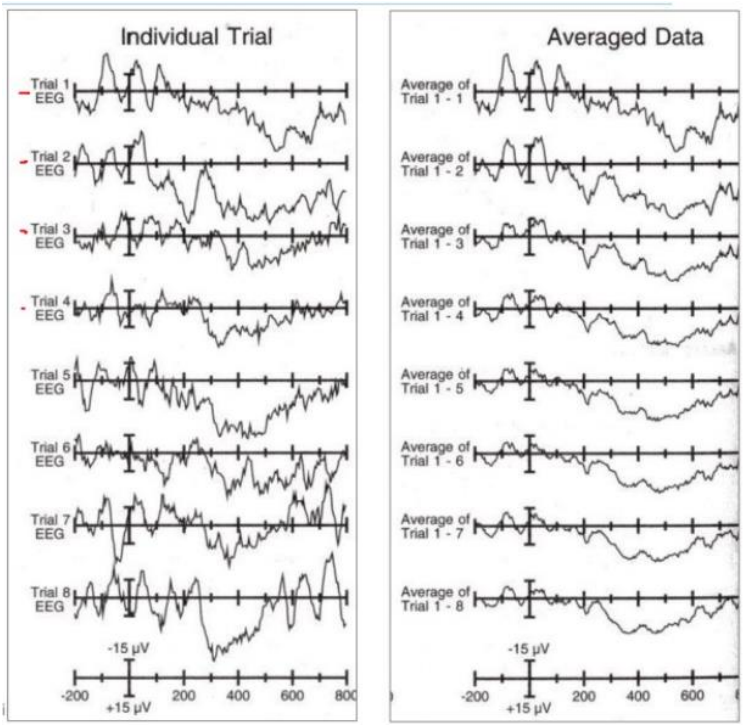
Averaging: is a signal processing technique used while dealing with event related potential. It is used because the potential generated by a stimulus can be hidden behind the spontaneous brain activity, so **it is not sufficient to give the stimulus and see the reaction, because probably we would not see anything.** What we have to do is to repeat the experiment many times and **average** them.

This kind of potential is called **evoked potential** (it is referred to a specific stimulus), and it is a stimulus-driven activity. It belongs to the family of **event-related potential**, a term used to generally describe changes in EEG signals triggered by either external stimuli or related to internal mental or task-related events.

To do this the continuous record must first be **cut up** into epochs or segments, consisting of a pre-stimulus, followed by a time suitably long enough to be able to highlight the activity of interest.

This method is called synchronized because all trials share the same time axis. We will call the time on this axis **latency.** The stimulus delivery is considered at time zero (200ms of latency means 200ms after the stimulus).

Averaging the trials means taking any latency, measure the values of the trials at that latency and average them.

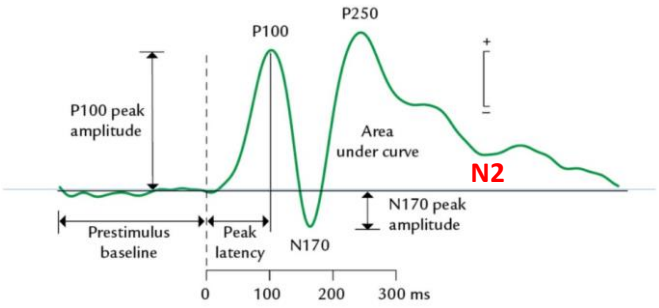


The average of the signals, which is a mix of spontaneous activity and evoked potential would at the end highlight the evoked potential, because it is almost deterministic; at each trial it is there, even if barred, but it is there. What changes is the spontaneous activity, uncorrelated with the stimulus, making high the probability that positive waves in a trial cancel out with negative one in another trial.

Averaging N responses would improve the **signal-to-noise** ratio by $\sqrt{N}$

An important thing to do, while analyzing an evoked potential, is to assure that the pre-stimulus average is zero.

The peak amplitudes, with respect to pre-stimulus baseline, and latency are commonly measured.



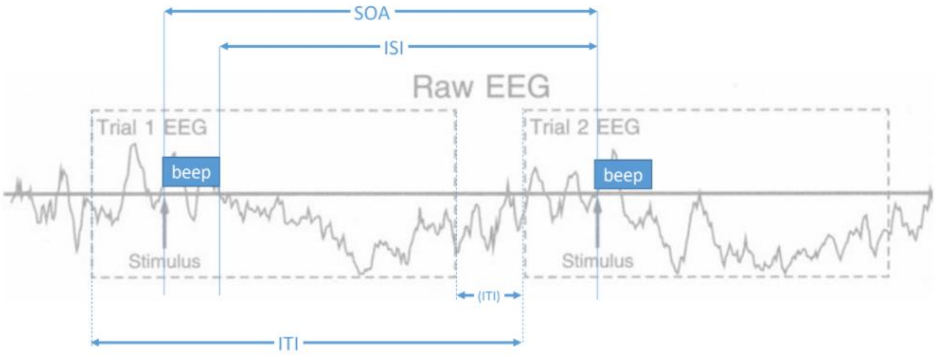
A pick is considered negative (or positive) not only because its amplitude is negative (or positive), but also because it can represent a negative (or positive) inflection with respect to the previous part of the graph (**N2** *in figure*)

When we find a pick named, for instance, N100, we have to not give for granted that its latency is exactly 100. It could also be that on average on the population that pick occurs at 100 of latency, and for this reason it is called N100 even if on a certain person occurs at latency 90.

Effect on stimulus timing:

- **Stimulus onset asynchrony** (SOA) refers to the time between two successive stimulus onsets(beginning of the stimulus).
- **interstimulus interval** (ISI) refers to the time between the offset(end of the stimulus) of one stimulus and the onset of another.
  Note that $\mathrm{ISI} = (\mathrm{SOA} - \text{stimulus duration})$
- **intertrial interval** (ITI) refers to the time between the beginning of subsequent trials.
  ITI and SOA can be equal if in each trial there is just one stimulus, otherwise they are usually different.

some authors define the ITI as the pause period between the end of a trial and the beginning of the next one!



An important aspect to be taken into account is that the shape and amplitude of ERP is dependent on the duration of the SOA. If we take a too short SOA, we could record an ERP which is different from the one we would have recorded with a longer SOA.

Why should we reduce the SOA? Because the SOA represents in a certain way the duration of the experiment, and it is obviously better to make the experiment as short as possible even if in some cases we necessarily need to make the experiment long to acquire more data. In fact, more data means more samples, and more samples we acquire, more the noise will be reduced (by a factor $\sqrt{N}$).

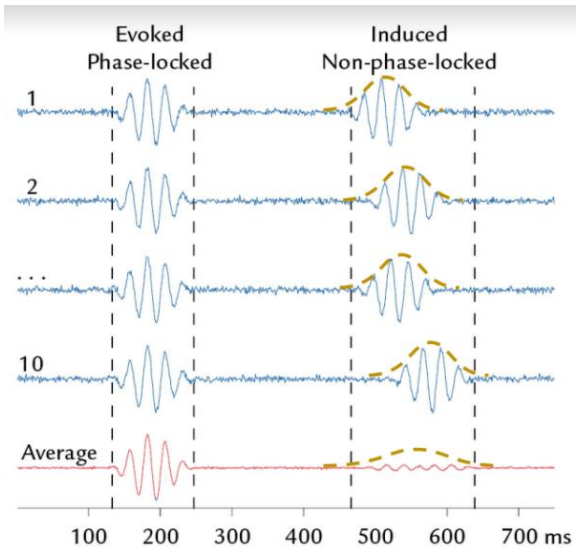
So, we need to find the trade off between the number of stimuli and the SOA.

If we take more samples, we would obtain a decrease of the noise (SNR increases). If we decide to reduce the SOA (or ISI), and we reduce it too much, the amplitude of the signal can decrease (SNR decreases), but in this way we would also have acquired more data. So, it is a tradeoff.

Mapping is a procedure useful to attribute components of the evoked potential to the area which plausibly generated them.

Spontaneous EEG and single trials

Until now we have talked about deterministic responses driven by a stimulus. In particular, we have considered time-locked (**phase-locked**) evoked activity, meaning that it starts, ends, and has its picks at exactly the same latency, following always the same wave form.



But it is also possible that each stimulus will induce an activity, which, despite being each time of the same frequency, is not exactly phase-locked to the stimulus.

This cause that the average evoked potential is almost zero. So, instead of taking the average sample by sample, we can average the envelopes of the various trials (**brown** lines in figure); the result we will obtain is not exactly the original evoked potential but better than almost zero.

Event related desynchronization (ERD) or synchronization:

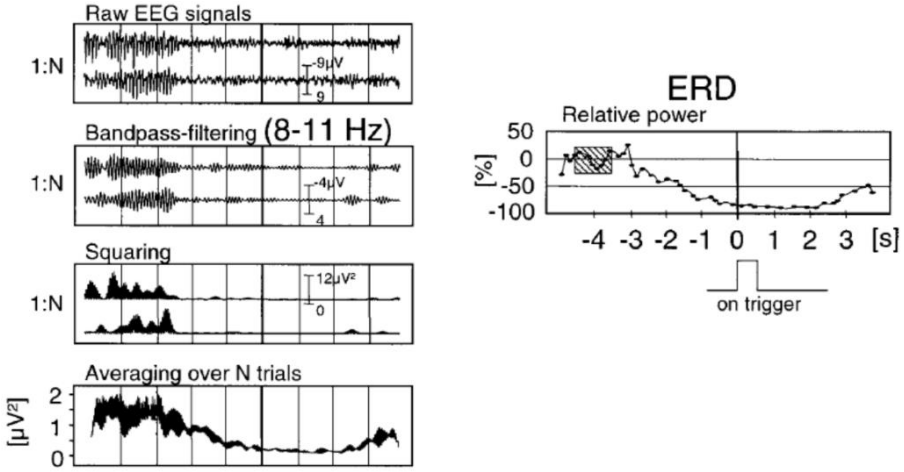
Desynchronization means a **reduction** of amplitude; **synchronization** stands for an **increase** of the amplitude.

Consider a cortical network which generates one rhythm, and this rhythm as a consequence of an event goes down, we talk about desynchronization. Vice versa, if the rhythm increases, we talk about event related synchronization.

To compute them starting from the EEG, the signals are:

- first **bandpass-filtered** to the frequency band of interest
- their amplitudes are **squared** across all epochs to have all positives samples (this avoids signals to cancel out while averaging)
- then, a moving window is used to calculate the **average** power as a function of time (average of the envelopes?)
  As an optional passage it is possible to do also a time average, and not only one on the trials, to smooth the curve.
- Last, normalize, expressing everything as a percentage of the average in the baseline and remove 100%. In this way we have something around zero in the baseline, decreases while preparing the movement, gets to the minimum while executing the

movement, and then comes back to its normal value (see figure which is referred to an ERD, same procedure but opposite result for an ERS).



ERD and ERS are expressed as percentage power changes relative to a baseline period preceding the event, be the event a sensory stimulus or a motor action.

Basics of signal processing - 1

Analog to Digital Conversion (ADC)

A signal is the variation of a physical entity (magnitude) in time.

Whenever we are acquiring a signal, we are using a **transducer**. So, we have an **electrode** that **transduces** electrical ionic activity (energy), into electrical energy carried by electrons, but this is an **analog** signal. We have to **convert** it into a **digital** representation.

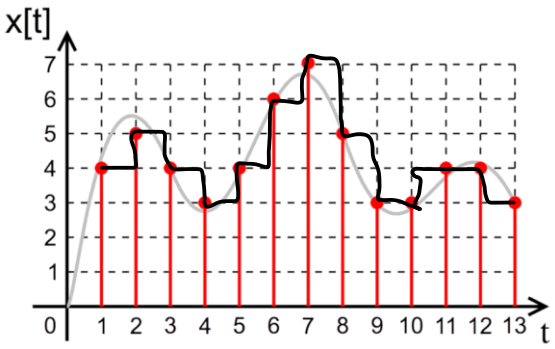
Analog means that the signal is **continuous** in time.

Converting a signal to continuous we are **discretizing** both the axes, meaning that the function will be defined only in some discrete values (on the independent variable axis) and can assume only some allowed values (on the dependent variable axis).

$$x(t) => x[n]$$

A continuous function of continuous time become a discrete function in discrete time.

We will use a sort of **grid** approach. **Sampling + Quantization.**



Quantization is the process in which, keeping the time continuous, we discretize the signal. The result of quantization is a **step-like** signal. Obviously, we are introducing an error, **noise** *n(t)*.

The noise is uniformly distributed in a specific range whose maximum and minimum values are half of the distance two consecutive levels $[-\frac{1}{2}V_{LSB}, \frac{1}{2}V_{LSB}]$. LSB is the acronym for **least significant bit** and represent the distance between two consecutive levels. It can be obtained by the formula

$$V_{LSB} = {Range}/{2^{bits}}$$

The **standard deviation** of the noise is:

$$\sigma = \frac{1}{\sqrt{12}} LSB$$

Where *Range* is the maximum value that the signal can assume, and *bits* is the number of levels we are going to use to discretize it.

The *LSB* **must be small enough** so that the **quantization noise** is **smaller** than the **input signal's intrinsic noise**, in such a way that the **quantization noise** becomes negligible with respect to the signal one.

$$noise_{tot} = \sqrt{noise_{signal}^2 + noise_{quantization}^2}$$

Sampling is the operation with which we discretize the x-axis, usually the time's axis.

Nyquist-Shannon theorem: A continuous signal *x(t)* can be "properly" sampled if it does not contain frequency components above one-half of the sampling rate.

Known that a signal can assume a maximum frequency value of *x*, according to this theorem I **must** sample it with a sampling frequency of *2*x* or bigger. **For instance**, a scalp EEG contains frequency until a maximum value (gamma band) of 100 Hz, thus, I know that the sampling frequency must be bigger or at least equal to 200 Hz.

"properly" means that: we are able to reconstruct the analog signal from the sampled signal with no error.

Nyquist frequency: one-half the sampling rate. If I'm sampling a signal and it contains frequency above the Nyquist frequency, I'm **making a mistake**.

Aliasing is a phenomenon that occurs when we do not sample properly a signal. It is an effect that turns samples that are not sufficiently near to each other into a wave form which is actually in a different part of the spectrum. In particular, since we can sample only below the Nyquist frequency, if we try to sample a signal that overcomes this value, we will obtain a reconstructed signal equal to one below the threshold. If Nyquist frequency is 50 Hz, sampling a 49Hz signal will be equal to a 51Hz one. A 95Hz equal to a 5 Hz signal.

Real **ADC** are limited by:

- Input range
- Maximum sampling frequency

The analog signal must be conditioned to fit the ADC we are using. In particular so that the **range** (amplitude) and **bandwidth** (frequency) do not exceed what is strictly needed for a faithful representation of the signal of interest. We have to **filter** the

signal to remove non-interesting part of the signal (noise). **Anti-aliasing** filters.

Basics of signal processing – 2

Statistics, probability and noise:

There are **deterministic** signals, but not all signals behave in this way, for instance noise is not deterministic. These signals are called **stochastic** or **random** signals. We will use probability theory to define them.

Basic measures for signal characterization:

Mean: $\bar{x} = \frac{1}{N}\sum_{i=0}^{N-1} x_i \rightarrow \mu_x$

In our applications x_i will be the values of the signal. $\bar{x}$ represents the empirical concept of mean, while μ_x the statistical concept. The mean describes the **one-value** which is representative of the whole signal. On the other hand, we might be interested in how much we deviate from this value.

Now there will be a list of indices measuring how the signal is dispersed around its average value.

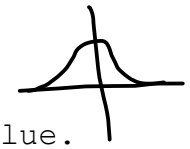
Average Rectified Value (ARV): $ARV_x = \frac{1}{N}\sum_{i=0}^{N-1} |x_i|$

Represents the average deviation from zero. A sort of measure of the signal's amplitude.

Root Mean Square (RMS): $RMS_x = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1} {x_i}^2}$

Computing the square of samples, we are taking into account the power the signal, for this reason we then compute the square root. Mathematically this approach is easier because square is easier to be derived with respect to the absolute value.

These two indices measure how much the signals deviates from zero, but sometimes we could be interested in the oscillation around its average value. Thus:



Average deviation: $AD_x = \frac{1}{N}\sum_{i=0}^{N-1} |x_i - \bar{x}|$

This is a sort of ARV but after having removed the mean value.

Variance and Standard deviation:

$s_x^2 = \frac{1}{N-1}\sum_{i=0}^{N-1}(x_i - \bar{x})^2 \rightarrow \sigma_x^2$ this is the **variance** and it is strictly linked to the power, the energy of the signal.

As before σ_x^2 is the mathematical interpretation, while s_x^2 is the empirical one. If we are using the sample mean, the empirical one, we have to put *N-1* at the denominator.

$s_x = \sqrt{s_x^2} = \sqrt{\frac{1}{N-1}\sum_{i=0}^{N-1}(x_i - \bar{x})^2} \approx\rightarrow \sigma_x$ is the **standard deviation**

Usually, a signal can be divided into a **constant** part, which carries no information, and a **variable** part, in which the whole information content is contained. All the indices defined after removing the mean from the samples (AD, Variance and Standard deviation) are focused on the variable part.

Gaussian white noise wave is a particular signal in which the knowledge of a sample is totally uncorrelated to the next one, which value is unpredictable.

The answer to the question "What is the probability that the amplitude of a signal is exactly one number" is always zero. We have always to reason in terms of **intervals**.

The probability of a sample to fall in a particular interval is proportional to the **width** of this interval.

We will compute the probability as the number of favorable samples divided by the total number of samples.

Probability density function is a measure of probability distribution of the values of a signal into an interval. (Mia interpretazione). It is computed by taking smaller and smaller intervals and rather then measuring the probability of a sample falling in this interval, we are considering the ratio between the probability and the width of the interval. By doing so and taking the limit of the width interval approaching to zero, then you will obtain the probability density function.

The integral of the **probability density function** is 1 by definition.

PDF for a triangular wave is constant in the range of the function, zero outside.

Basics of signal processing – 3

A **random variable** is a variable of which you do not know the exact value you will see when you observe it in a certain point, you only know some statistical properties describing it.

Central limit theorem (CLT):

*When N **independent*** (meaning that you cannot predict one starting from another, even partially) *and **identically distributed*** (same **probability density function**) *random variables* (samples(?)) X_i *are averaged, the resulting variable* $Z = \frac{1}{N}\sum_{i=1}^{N} X_i$ *tends toward a **normal distribution** even if the original variables themselves are not normally distributed.*

(Normal distributions are sometimes called Gaussian distributions)

The mean of Z equals the mean of X_i, the variance decreases by a factor N, the standard deviation of a factor $\sqrt{N}$

$$\mu_Z = \mu_X$$

$$\sigma_Z^2 = \frac{1}{N}\,\sigma_X^2$$

$$\sigma_Z = \frac{1}{\sqrt{N}}\,\sigma_X$$

According to the Fourier theorem a **periodic signal** in time can be expressed as the summation of a **fundamental** sinewave and its **harmonics**, which are other sinewaves having frequency multiple of the fundamental one.

A **sinewave** is a function of time with parameters: frequency, amplitude, and phase. Amplitude and phase can be represented as a complex number.

$C = A * e^{j\varphi} \leftrightarrow A = |C|, \varphi = \ phase(C)$

Basics of signal processing – 4

Discrete Fourier transform allows to represent a time series in the frequency domain. In particular:

Transformation from time (N real samples, taken every $1/f_s$ seconds) to frequency domain (N complex samples spanning [0, f_s) Hz)

$$\underset{\text{HARMONIC}}{X[m]} = \sum_{n=1}^{N} \underset{\text{SAMPLE}}{x[n]} e^{-\frac{j2\pi mn}{N}}$$

$x[n]$ is one sample of the time series.

$X[m]$ represents one of the **harmonics**, one of the components. Each one is a complex number with an amplitude and a phase, representing a sinewave.

So, you take a signal in time, you compute its frequency component and obtain its frequency domain components.

Just by summing these components it is possible to approximately (take care of discontinuities) reconstruct the signal.

Since the fundamental frequency sinewave repeats only one in the time window we are considering, its frequency is exactly $\frac{1}{time\_window}$

For non-periodic signals we are focusing on the time window they are in, but after having transformed and anti-transformed them, we will obtain a periodic signal due to the periodic behaviour of the sinewaves we have used to reconstruct it.

Given the sequence of complex numbers obtained by transforming the signal using DFT, we can build a frequency representation of the original signal split in two parts:

- **Amplitude**, represented by the magnitude of the complex numbers. (Interpolate them to obtain a graph)
- **Phase**, represented by the complex part of the complex numbers. (Interpolate them to obtain a graph)

Nice symmetry

Time domain periodic <-> frequency domain discrete

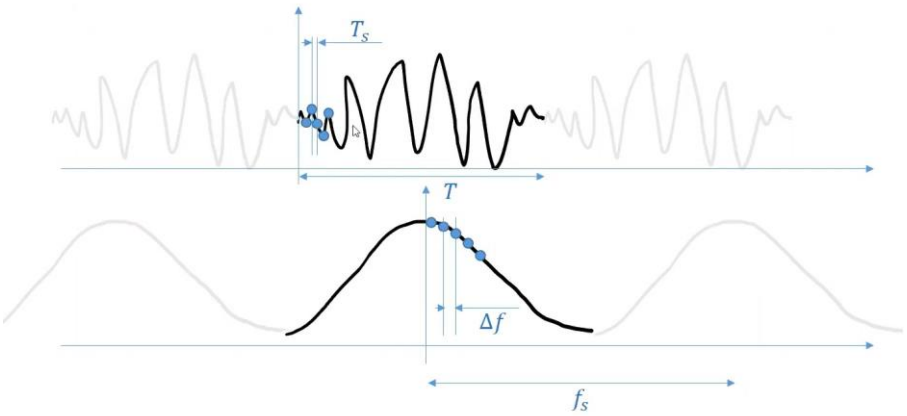
Time domain discrete <-> frequency domain periodic

Moreover, if the time domain is real, the frequency domain is symmetric: *X(-f) = X*(f)*. And this is the reason why to represent it I just need half of the samples, the positive axis half.

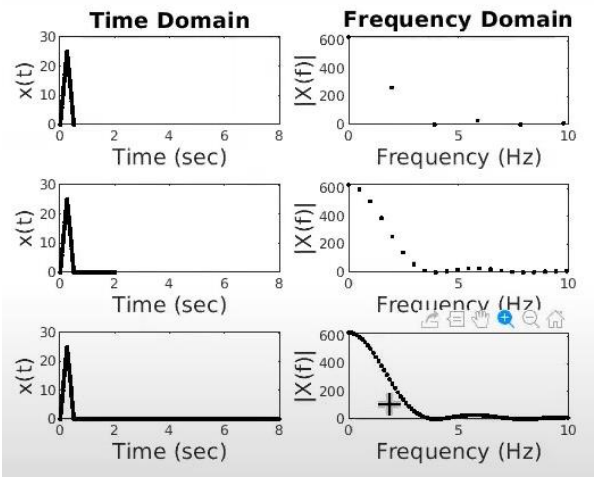
The replicants are distant exactly one sampling frequency.

Basics of signal processing – 5

- Time domain T-periodic <–> Frequency domain Δf-discrete ($\Delta f = 1/T$)
- Time domain T_s-discrete <–> Frequency domain f_s-periodic($f_s = 1/T_s$)



Since the distance between two samples in the frequency domain depends on the time signal period, we can enlarge it with fictitious data in order to obtain a better resolution of the frequency signal. (**Zero padding**)



Since we are respecting the Shannon theorem all those three graphs in the frequency domain carry the same information, so mathematically nothing changes, we would still be able to reconstruct the signal. However, it helps to visualize better the result.

Zero padding lies on the duality between time and frequency domain, according to which:

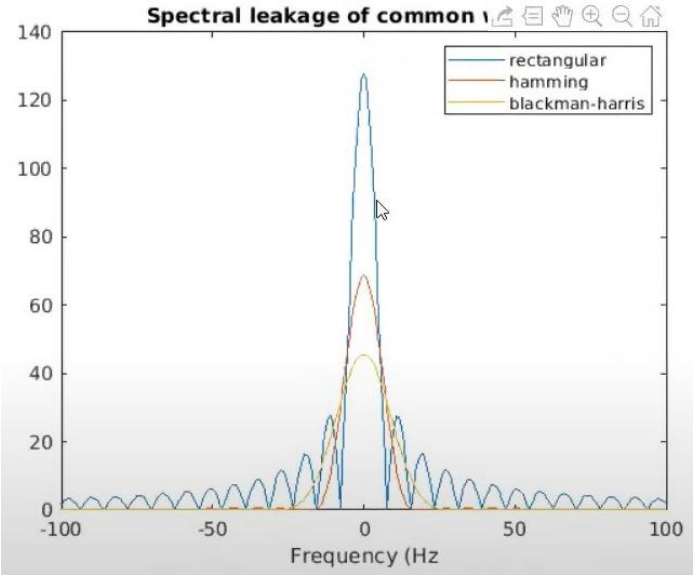
Higher sampling rate in t <-> broader spectrum in f

Longer signal in t <-> higher resolution in f

This technique consists in adding a row of zeros at the end of the signal and allow to, artificially, increase the spectral resolution, even dealing with short signals.

Windowing is another process useful to take a finite span of a signal. This is mathematically equivalent to multiplying by a rect() function.

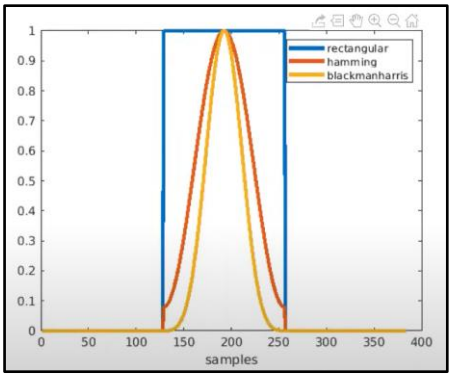
A problem with this is the so-called leakage effect. This is visible with sinewaves. Transforming a sinewave, we would expect a vertical line (delta?) because it has power at just one frequency. But if we transform a part of it, obtained by windowing it, we will obtain a primary lobe, which is not more a line but a lobe, and instead of having a zero in all the other frequencies we will have lots of secondary lobes.



So, the figure above is the spectrum of the windowing function, but, since the spectrum of a sinewave is a delta, the result of the spectrum of the two multiplied signals (in frequency convolved) will be of that shape but translated where the delta is centred.

The distance between two consecutives zeros of that spectrum (the duration of secondary lobes, is exactly 1/(*windowing function width)*

The cause of the presence of the side lobes are the discontinuities in the rectangular function corners. Thus, solution is to use more gentle windowing functions.



Tapering the corners of the rectangular function we can approximately delete the presence of secondary lobes, but as a cost for this we will also broaden the primary lobe.

It is important, while tapering the corners of the windowing function, to take care of the trade-off between the primary lobe width and the height of the secondary ones.

Basics of signal processing – 6

The frequency representation of a signal allows us to know what the power is and, what is the phase of each sinewave component of the original signal.

We are interested in knowing the power of the signal at each frequency.

Power spectrum or **power spectrum density**:

It describes the power of the signal in each frequency band of the spectrum.

$$PSD(f) = \left|DFT\left(x(t)\right)\right|^2$$

Two problems with this:

- Since we do not have infinite data, we are just estimating the real value of the PSD.
- On the other hand, recording the signal for a long-time interval will generate a very high-resolution spectrum, which is not necessary.

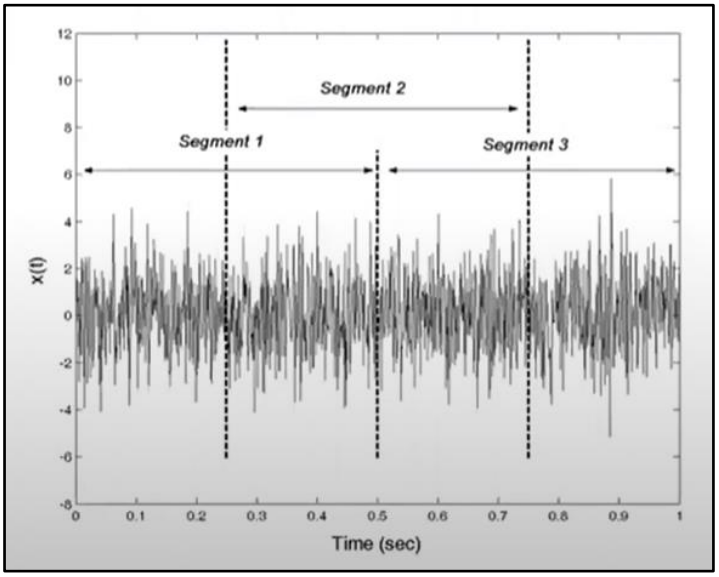
We have to make a trade off. We can trade off spectral resolution by reducing the estimation error. This is done through the **average periodogram** method (**Welch's** method).

Since measuring an EEG two times will not give the same signal, we can consider it as a stochastic signal. PSD of stationary stochastic process (non-deterministic signal) is an **estimate** subject to errors. We are trying to estimate the spectrum of the process that has generated the signal, and this is why we can repeat the measurement.

A long signal yield to a very high spectral resolution.

This procedure follows some steps:

- Take a short segment of the signal of length n-samples
- Apply a windowing function *win(t)* (optional)
- Compute the **periodogram**, i.e. $\left|DFT\left(win(t) * x(t)\right)\right|^2$ and store the resulting PSD
- Shift (samples do not move, it is you that shift your attention to a segment shifted right of a certain distance and consider then a new segment starting from there) n-samples to the right (or a fraction of this amount, if overlapping segments are desired)
- Repeat steps 1-4 until the end of the signal is reached (num-segments)
- **Average** all PSDs. The result will be a better estimation of the process.



The **spectral resolution** is reduced, determined by the length of the segment. (1/length)

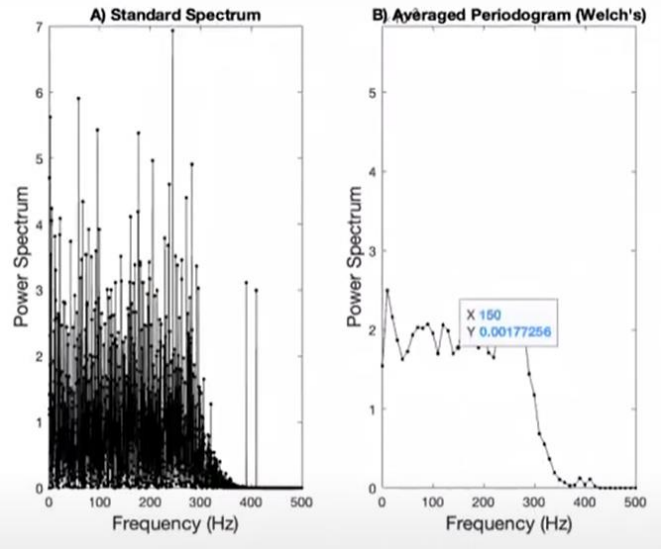
Since, whenever you multiply uncorrelated signals which have a mean of 0, the power is reduced by a factor n, so the amplitude goes down of a factor $\sqrt{n}$

In this case, the variability (estimation error) of the PSD is reduced by a factor $\sqrt{num\_segments}$

So, we have obtained the tradeoff we were looking for.

The shorter are the segments, the higher is the reduction in estimation error. But the shorter is the segment the broader is the spectral resolution.

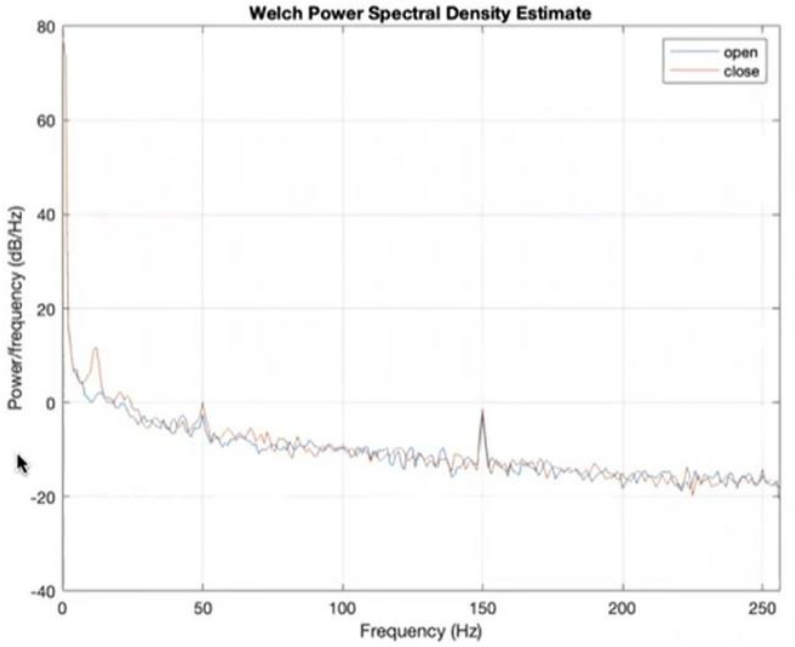
The overlapping between segments should not be more than the 50% because otherwise the reduction of $\sqrt{num\_segments}$ does not apply anymore because of the too high correlation between segments.



The strong assumption we have made is that every segment is a partial estimation of the process, this is true only if the signal is stationary (if you close your eyes EEG is not stationary anymore).

If the signal is **not stationary**, we are not allowed to use the Welch's method, unless we segment the signal in stationary sub-processes.

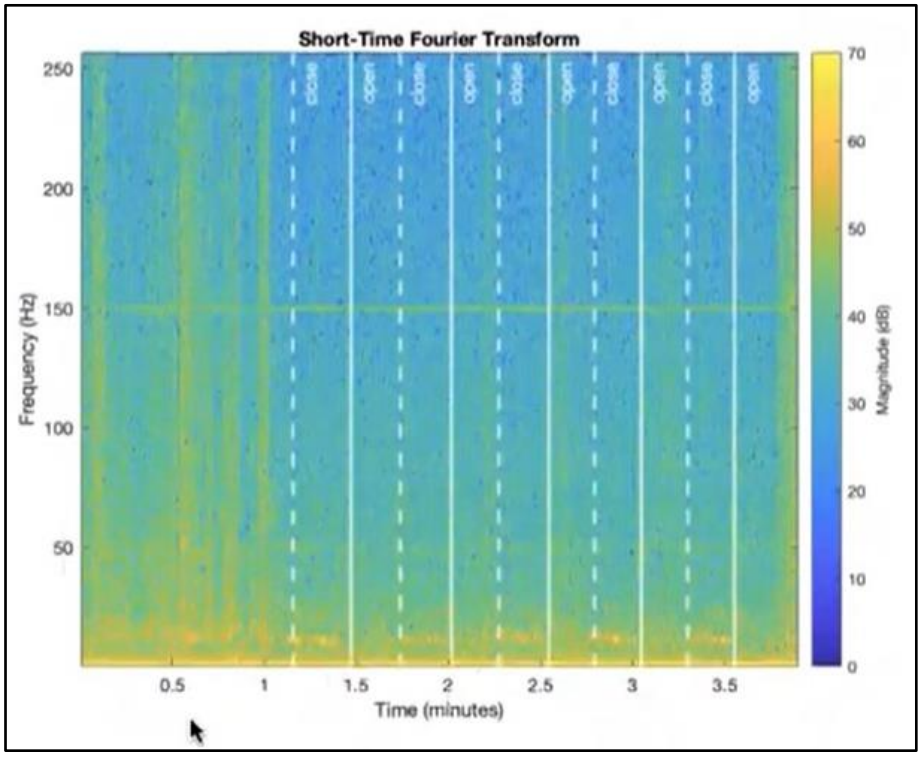
So, we have to separate the processes (eyes open, eyes closed…) consider them as different processes and compute their PSD and spectrum.



Those two picks are artifacts.

For not stationary signal PSD should not be used to describe the signal's spectrum. If short segments of the signal can be considered stationary (the signal is pseudo-stationary), the evolution over time can be displayed using a spectrogram. We have a full spectrum for each segment, not an average.

STFT (short-time Fourier transform) is the simplest method to achieve this goal.



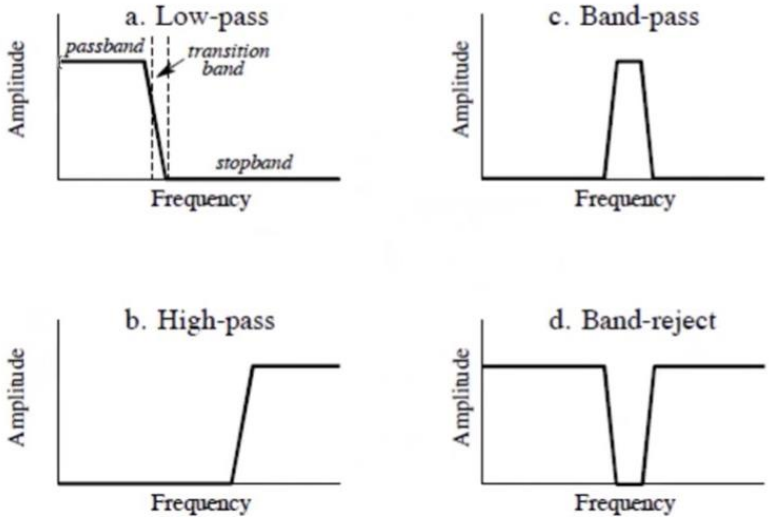
This allows to see the evolution of the spectrum over time.

Digital filters

The final purpose of a filter is to improve the **signal to noise** ratio, reducing as much as possible the part of the signal we consider noise. It allows some spectral component of a signal to pass (almost) unaltered, while (almost) blocking other spectral component. Even in this case is a matter of trade off between removing as much noise as possible and preserving the rest of the signal.

Main filters:

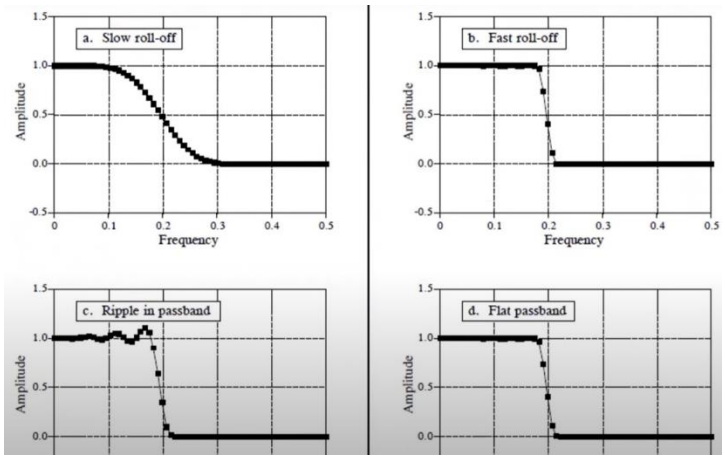
- **Low pass**, useful when we are interested in a low frequency signal
- **Band pass**, useful when we are interested in a specific frequency band signal
- **High pass**, useful when we are interested in a high frequency signal
- **Band-reject**, useful to remove the signal around a frequency.



The **passband** is the interval of frequencies that are passed (output equals to input), while **stopband** contains those frequencies that are blocked. The **transition band** is between.

The **cutoff frequency** is the division between the passband and transition band. Expressing the gain in dB, the cutoff frequency is a value in dB which express the exact point in which the passband becomes transition band. It usually defined to be where the amplitude is reduced to 0.707 times the input. Another cutoff frequency defines the division between transition band and stopband, defining with the other one the width of the transition band.

We will call **roll-off** the slope of the transition band. High roll-off means that the response of the filter in the transition band is almost vertical. Slow roll-off means that the slope is much smoother (less ideal filter).



A filter that has a non-monotonic (oscillatory) behavior either in the passband or in the stopband, it is said that it has a **ripple** in the passband or in the stopband.

There exist two main families of filter types:

- **Finite impulse response,** in which the output is computed by combining samples of the input

$$y[i] = \sum_{k=0}^{N} b_k x[i-k]$$

- **Infinite impulse response,** output computed as a combination of samples of the input and previous samples of the output (the output of these filters often does not go to zero because they take into account the previous outputs)

$$y[i] = \sum_{k=0}^{N} b_{kx}[i-k] - \sum_{k=1}^{N} a_k y[i-k]$$

The **order N** of a filter indicates how many recent samples are used to compute the output.

Types of IIR filters:

- Butterworth
- Chebyshev (type 1 or 2)
- Elliptic

Types of FIR filters:

- Windowed sinc
- Minimax (or Parks-Mcclellan)

IIR more efficient than the FIR (FIR would require higher N value to obtain the same result).

FIR filters introduce much less distortion on the time domain signal than IIR (very important for broad band signals).

Brain-computer interface – 1

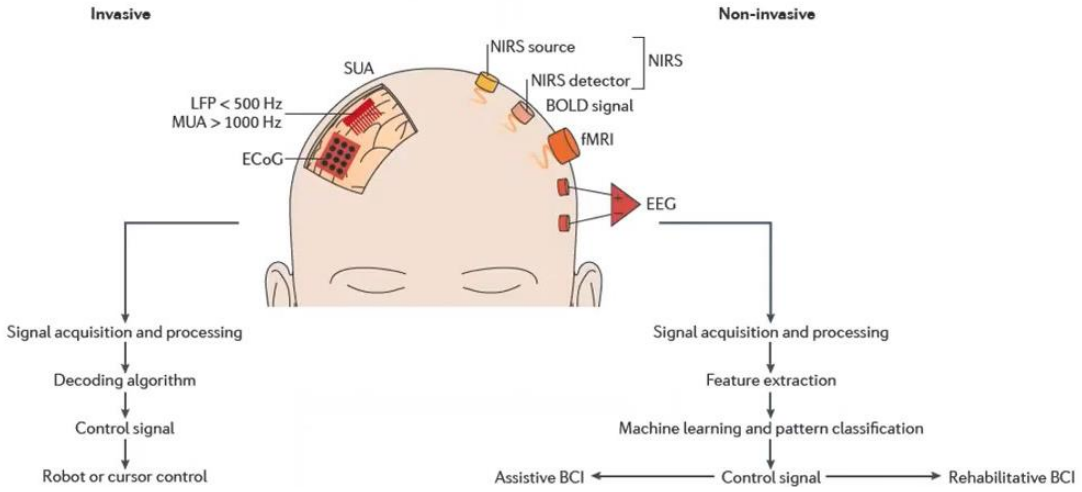
A BCI is a system that measures the brain activity, or more in general the central nervous system activity, and converts it into artificial output that replaces, restores, enhances, supplements or **improves** natura CNS output and thereby changes ongoing interactions between CNS and its external or internal environment.

It is based on three major steps:

- Signal acquisition
- Feature extraction
- Feature translation

The user not only can interact with the external environment, but thanks to the close loop structure with which these systems are built, he can receive feedback on what he is doing.

According to the way we have acquired the signal, BCI can be divided in **invasive** and **non-invasive**. The signal acquisition approach slightly changes also the procedures that follow the acquisition.



Rehabilitative BCI does not replace a lost function (as assistive BCI does) but support the patient.

Invasive BCI involves surgical implantation of electrodes or multi-electrode grids; it measures activity patterns of neurons which encode behaviorally relevant information.

With invasive BCI we can measure:

- Local field potentials (LFPs)
- Single-unit activity (SUA)
- Multi-unit activity (MUA)
- Electrocorticographic oscillations recorded from electrodes on the cortical surface (electrocorticography, EcoG
- Epidural or subdural field potentials

Non-invasive approach is mainly based on EEG recordings, with which we can record **rhythms** and **evoked potentials**.

To use a BCI, we start a **mental task**, like mental imagery, that changes the EEG activity. Then the BCI recognizes this change, transform it in **control features** and activate an **external device**.

The **P300** component of the ERPs can be extracted from the ongoing EEG (average). It is a large centro-parietal positive wave occurring approximately 300ms following a stimulus (visual or audio for instance). It has been associated with cognitive information processing (e.g. memory, attention, executive, function…). An example of real situation in which it is generated is when, being in a very noisy environment, someone pronounces our name. In BCI we simulate this using a sequence of sounds like bip, bip, bip randomly alternated with some bops and asking the patient to count the bops. Its attention would be focused on the bops and at each bop its brain will generate a P300.

The P300 ERP generated by attending a target stimulus is exploited to build virtual keyboards based on a BCI.

A type of BCI is based on imagery task that produce a change in the brain and control the BCI. For instance, rhythms that are useful to this goal are the **sensory motor rhythms**.

These are **oscillations** in magnetic or electric field recorded over the sensory-motor cortex in the mu, beta and gamma frequency bands. SMRs decrease (**ERD**) is associated with overt (execution) and covert (imagery) motor activity. This is related to an increase of corticospinal tract excitability accompanied with a decrease of GABAergic intracortical inhibition. (This is necessary to isolate the movement we are planning to only the muscles involved). SMRs-ERD covaries with BOLD signals deriving from the same cortical areas.

Movement and imagery generate a variation in the brain activity in the centro-parietal area.

The amplitude of **sensorimotor** rhythms can be voluntarily modulated through the exercise of motor imagery, to build a cursor control.

This learning is done using neurofeedback. It is a progenitor of BCIs, in which individuals aim to directly regulate external devices instead of neural substrates (heart frequency, breath).

Stroke (ictus) is one of the most common causes of brain disabilities. Some degree of **spontaneous motor recovery** is usually seen after a stroke. Constituting a remarkable evidence that innate physiological and anatomical **plasticity** are relevant processes underlying motor recovery after a stroke

- Remodeling of brain networks functionally close to those affected by stroke: **adaptability**
- Following similar rules that hold experience-dependent plasticity: **task-specific training and practice**

"Neural plasticity" refers to the capacity of the nervous system to modify itself, functionally and structurally, in response to experience and injury.

The reorganization after a stroke starts immediately after the injury and lasts for months.

BCI is so important in **neurorehabilitation** because:

- Can directly measures signals from damaged cortical areas and translate them into close-to normal motor outputs providing online feedback on performance. This satisfies some neurorehabilitation principles like:
  - Extensive task-oriented training program
  - Feedback on performance
  - Adherence to intervention
- BCI based training can thus, promote re-training of brain activity and induce adaptive plasticity. This satisfies some other neurorehabilitation principles like:
  - Retraining brain areas associated to functions (top down approach)
  - Enhance neuroplasticity

We can have three approaches:

- Brain to function
- Brain to limb
- Brain to brain

Brain-computer interface – 2

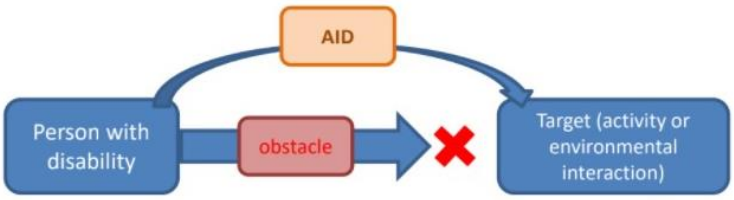
Whenever we want to interact with the environment, we have to use our nerves to control muscles. BCI connects our brain with the environment bypassing natural pathways. It requires the CNS to produce entirely new, **artificial** kinds of outputs, far from its natural function.

Augmentative and Alternative Communication (AAC) includes all forms of communications (other than oral speech) that are used to express thoughts, needs, wants and ideas.

An AAC can be:

- Aided if we use tools or equipment in addition to the user's body
- Unaided if we are using only our body to communicate

Assistive technology is an umbrella term indicating any product or technology-based service that **enables people** of all ages with activity limitation in their daily life, education, work or leisure. Sometimes it is possible to design assistive solutions by assembling mainstream technologies (tablet, pc…)



AT is any item, piece of equipment, software or product that is used to increase, maintain or improve the functional capabilities of persons with disabilities.

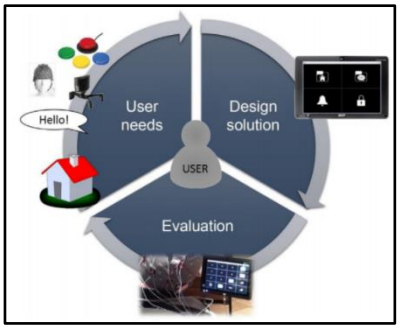
An assistive technology service is any service that directly assists an individual with a disability in the selection, acquisition, or use of an assistive technology device.

The process to choose the right aid for a person with a disability is long and accurate:

- evaluation of the assistive technology needs of the individual
- a service consisting of selecting, designing, fitting, customizing, adapting, applying, maintaining, repairing, replacing, or donating assistive technology devices;
- training or technical assistance for an individual with a disability or, where appropriate, the family members, guardians, advocates, or authorized representatives of such an individual;
- training or technical assistance for professionals

- a service consisting of expanding the availability of access to technology, including electronic and information technology, to individuals with disabilities.

A possible approach is the **user centered design (UCD)**

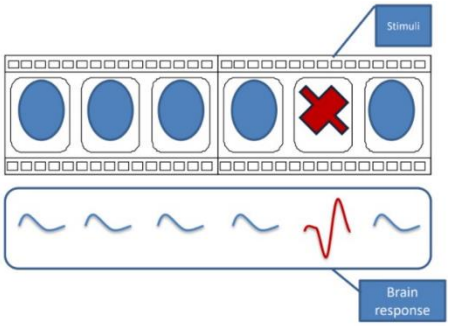


But how to evaluate the usability of a product? Three parameters:

- effectiveness
- efficiency
- satisfaction

Oddball paradigm

This is a version with different signals in which the less frequent is the target one.



Obviously the BCI must be calibrated on user's needs.

A BCI consists of three main phases:

- signal acquisition
  - 8 EEG channels
  - Sampling rate 200-250Hz
  - Labels about paradigm
    - Stimulation onset
    - Stimulation class
    - Stimulation type
  - Preprocessing
    - Bandpass filter
    - Notch filter
    - Spatial filter
      - CAR
      - Laplacian
- feature extraction and classification $y = \sum_{i=0}^{N} w f_i + b$
- application