What is an EEG?

* Electroencephalogram
  + Measures electricity in brain and records it
* Electrodes are placed outside the head (non-invasively) or surgically inside head (invasively)
* Exploitation of dipole
  + Two different charges separated by difference
  + Electrodes can generate read out of difference in charge between two different areas
* Resting neuron has negative charge relative to outside neuron
  + Excitatory Post-Synaptic Potential
    - Neurotransmitters binding causes positive ions to flow into neuron and make the neuron more positive
    - Extra cellular space directly outside the neuron therefore becomes relatively more negative
    - The space that is further from neuron also becomes relatively more positive
    - Two different charges are separated by a distance; we have a dipole
  + Inhibitory Post-Synaptic Potential
    - A dipole that is the other way around!
    - Neurotransmitters binding causes negative ions to flow into neuron and make the neuron more negative
    - Extra cellular space directly outside the neuron therefore becomes relatively more positive
    - The space that is further from neuron also becomes relatively more negative
  + These create differences in charge at opposite end of neuron
* Billions of large neurons lie parallel to each other in cerebral cortex and perpendicular to scalp
  + Clusters of these pyramidal neurons are stimulated together
  + Individual dipoles are created and summate to a larger dipole
  + So, an electrode can measure the difference in charge between these two areas
    - This is what the EEG signal is!
    - Electrical difference between pair of electrodes on two different parts of brain. It’s created by the cluster of neurons being excited and inhibited
* 32 electrodes are placed across brain, and need to pair up
  + Common reference montage is where each electrode is compared to the same electrode
    - Each horizontal line shows the difference from each electrode to the same electrode
    - Identified from the last hand side of the montage
  + Another approach is comparing every electrode to electrode next to it
    - Bipolar montage
    - Each horizontal line has the electrical difference between an electrode and one next to it
* Things to look for
  + Amplitude
    - How much signal goes up and down from midline
  + Frequency
    - Number of waves in given amount of time
* Bands of Frequency
  + Sleep, Delta Waves, <4 Hz
  + Relaxing, Theta Waves, 4-8 Hz
  + Attention, Alpha Waves, 8-13 Hz
  + Active Mind, Beta Waves, 13-35 Hz
  + High Concentration, Gamma Waves, >35 Hz
* Establishing frequency of signals can help detect abnormal brain patterns
  + Diagnosis of different diseases, etc.

What is ICA

* Independent Component Analysis is a signal processing method to separate independent sources linearly mixed in several sensors
* Microphone example
  + Multiple microphones recording a situation, ICA can separate the sounds into the various channels
  + All of the sounds are in original signal, and ICA separates them
* Looking at EEG
  + Multiple electrodes are present
  + If there’s a patch that projects to many electrodes, it’s like having a conversation projecting to different microphones
* How does this work?
  + ICA recovers version of the original signal
  + U = WX
    - X is the data (channels x time)
    - U are the ICA source activities (component x time)
      * Different sounds or different brain sources
    - W is the ICA unmixing matrix (components x channels)
  + PCA simply decorrelates the outputs (using an orthogonal matrix W)
  + ICA attempts to make the outputs statistically independent, while placing no constraints on the matrix W
* Example
  + Linear mixture of two sources: a and b
  + ICA is able to unmix these and separate the sources

A graph of a graph of a graph

Description automatically generated with medium confidence

How does ICA Work?

* PCA finds the axis which projection of data has maximum variance
  + A graph with dots and lines

    Description automatically generated
* ICA finds projection of data points that maximize independence
  + A graph with dots and lines

    Description automatically generated
* ICA is a method to recover a version of the original sources by multiplying the data by a unmixing matrix
* Central Limit Theorem (CLT)
  + What we record on scalp is a mixture of two different sources
    - Scalp channels will be by definition a mixture
  + CLT tells us that mixture of these sources will converse to a Gaussian distribution
    - ICA starts from scalp channels and rotates axis to make projection as non-Gaussian as possible

ICA Training Process

* Maximize entropy to rotate axis

A diagram of a training process

Description automatically generated

* Entropy is presence of randomness, the more uniform means that the results are more random, therefore more entropy
  + A screenshot of a math problem

    Description automatically generated
* Don’t just maximize entropy, also maximize joint entropy
  + This is related to mutual information
  + Joint Entropy = Entropy of first process + Entropy of second process – mutual information between the two processes
* ICA learning rule
  + How to make outputs statistically independent?
    - Minimize their redundancy or mutual information
    - Maximizing joint entropy means minimizing mutual information
      * Mutual information is 0 if two variables are independent

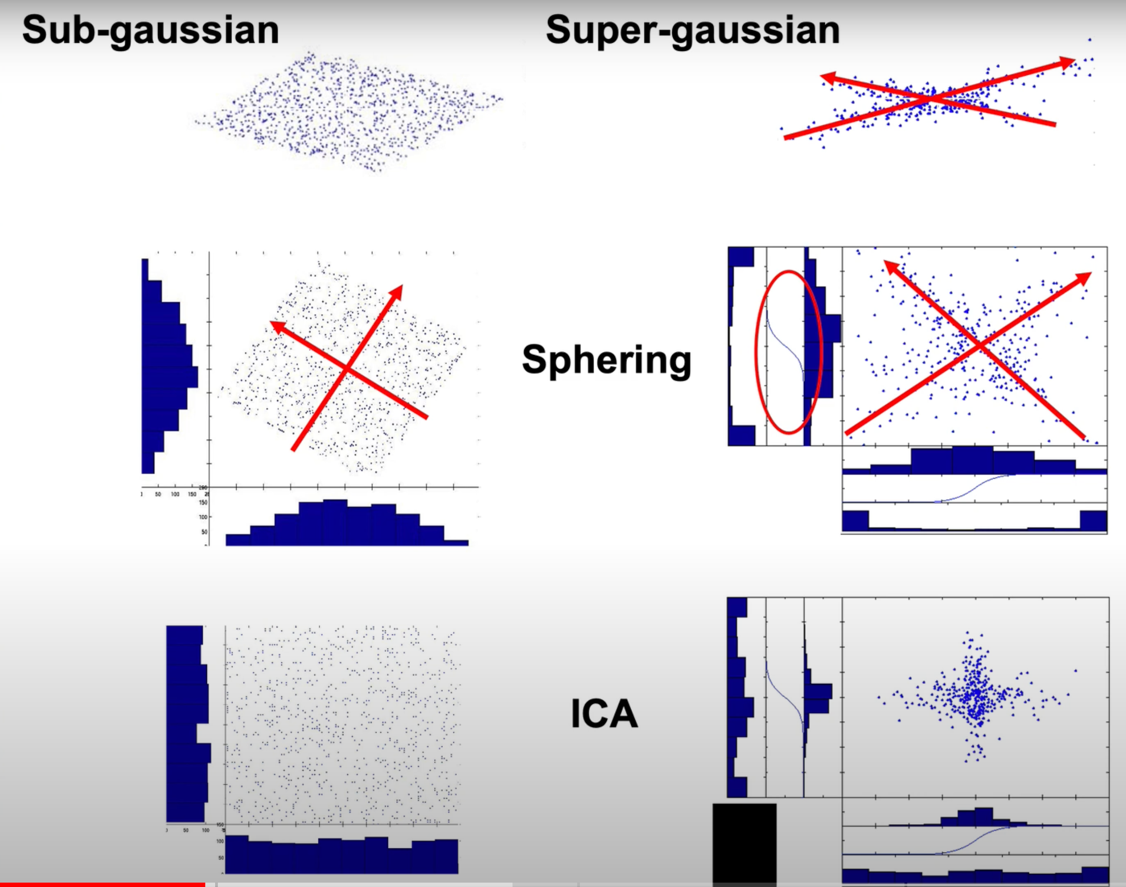
ICA Applied to EEG Data

* Super- and Sub-Gaussian sources

A screenshot of a graph

Description automatically generated

* EEG data is Super-gaussian
  + Need to apply transformation; distribution passed through non-linearity
  + Ultimately will transform so that entropy can be calculated; then original axis for source can be found



* Logistic transform is applied so that entropy can be calculated. Will have a relatively uniform distribution
  + A diagram of a nuclear explosion

    Description automatically generated with medium confidence
* Matrix Multiplication
  + Using Weight Matrix W (U = WX)
    - A screenshot of a white board

      Description automatically generated
  + Using Inverse Weight Matrix W-1 (X = W-1U)
    - A diagram of numbers and symbols

      Description automatically generated
  + By zeroing out the ICA activity of some artifacts, can remove these artifacts from EEG data
    - ICA is fundamentally matrix multiplication to separate various channels

ICA/EEG Assumptions

* Mixing is linear at electrodes
* Propagation delays are negligible
  + Propagation delays are close to speed of light!
  + At our time scale (1 Hz – 1000 Hz), we can ignore this
* Component time courses are independent
  + Truly independent in brain
  + Maximum projection algorithm
    - Solution that is *maximally* independent
* Number of components less than the number of channels
  + Find sources which contribute most to EEG signal, then apply some cutoff
  + Won’t see the ones that don’t contribute to EEG signal
  + A graph with red text

    Description automatically generated
* Characteristics of Independent Component of the EEG
  + Artifacts
  + Stimulus-locked activity
  + Response-locked activity
  + Non-phase locked activity
  + Event-modulated oscillatory activity
  + Overlapping maps and spectra
* Sample EEG ICA Decomposition
  + All components are mixed together on the top
  + ICA is able to separate into cognitive and artifactual processes
  + A diagram of a graph

    Description automatically generated with medium confidence
* ICA Applied to EEG vs ICA applied to fMRI
  + ICA activity is time course of components
  + Columns of inverse weight matrix are the scalp topographies

A screenshot of a diagram

Description automatically generated

ICA in MNE Specifically

* Cutoff of 1 Hz
* Four different methods
  + FastICA
    - Fixed-point iteration method to find an orthogonal rotation of pre-whitening data that maximizes a measure of the rotated components' non-Gaussianity
    - Corresponds with EEG data being super-Gaussian. Orthogonal rotation (transformation happens)
  + Picard
    - Preconditioned ICA for real data
    - Quickly solves maximum-likelihood ICA, typically faster than the other alternatives
    - Can work under orthogonal constraint and can separate super-Gaussian and sub-Gaussian signals
  + Infomax
    - Use of entropy to extract ICA signals
    - Looking for orthogonal rotation matrix that will maximize joint entropy
      * By maximizing joint entropy we are minimizing mutual information, i.e. separating channels
  + Extended-Infomax
    - Can separate sub and super-Gaussian signals but converges slowly as it uses stochastic gradient optimization