ICA EEG Analysis
Basics and Applications

Jeng-Ren Duann, PhD

Associate Director, Biomedical Engineering Research Center CMU&H and

Associate Scientist, Institute for Neural Computation, University of California San Diego

What's ICA?

- Independent Component Analysis
- Blind Source Separation
  - Speech recognition
  - Cocktail party problem

Historical Remarks

- Herault & Jutten, Space or time adaptive signal processing by neural network model, Neural Nets for Computing Meeting, Snowbird, 1986 (seminal paper)
- Camon (1994): Approximation of mutual information by 4th order statistics
- Bell & Sejnowski (1995): Information maximization (infomax)
- Amari et al., (1996): Natural Gradient Learning
- Cardoso (1996): JADE

ICA learning rule

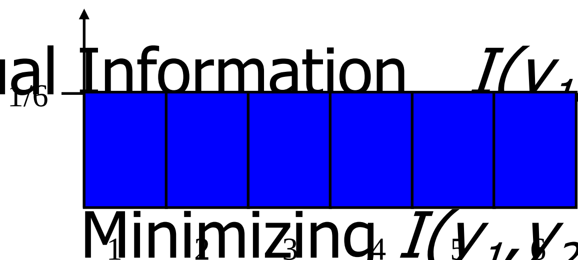
How to make the outputs statistically independent?

Minimize their redundancy or mutual information.

Entropy: $$H(X) = -\sum_{x\in X} p(x)\log(p(x))$$

Joint entropy $$H(X,Y) = -\sum_{(x,y)\in X\times Y} p(x,y)\log(p(x,y))$$

Dice: 1/6



$$H = 6\left(-\frac{1}{6}\log_2\left(\frac{1}{6}\right)\right) = 2.58$$

Mutual Information $I(y_1,y_2) = H(y_1) + H(y_2) - H(y_1,y_2)$

Minimizing $I(y_1,y_2)$ → **Maximizing** $\boldsymbol{H(y_1,y_2)}$

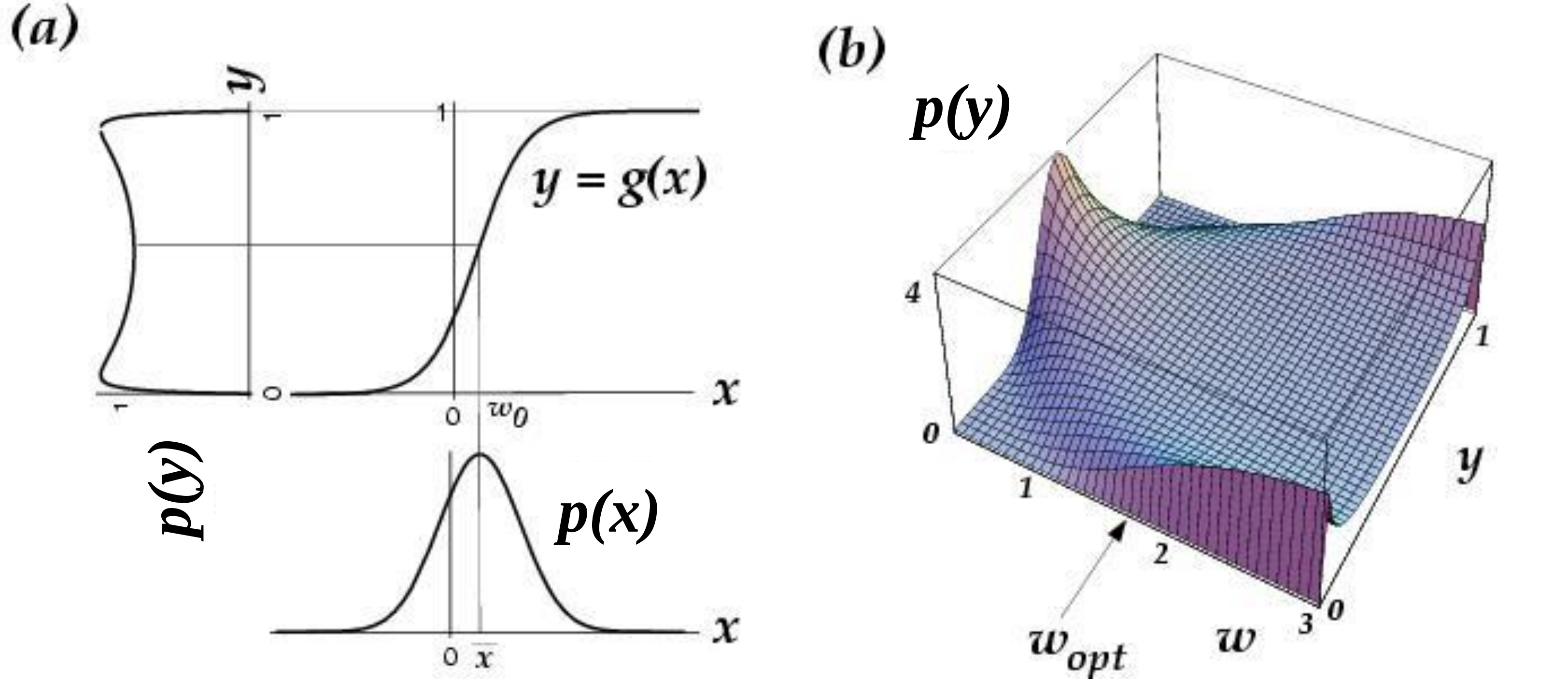
=0 if the two variables are independent

ICA learning rule

$$\Delta W = \frac{\partial H(y_1, y_2, \ldots)}{\partial W}$$

InfoMax(Bell & Sejnowski, 1995)

The non-linear function provides all the higher-order statistics necessary to establish independence.

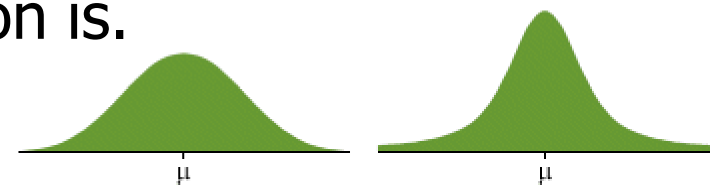


From Bell & Sejnowski, *Neural Compu. 1995.*

Kurtosis, Super- and Sub-Gaussian

Kurtosis: a measure of how peaked or flat of a probability distribution is.

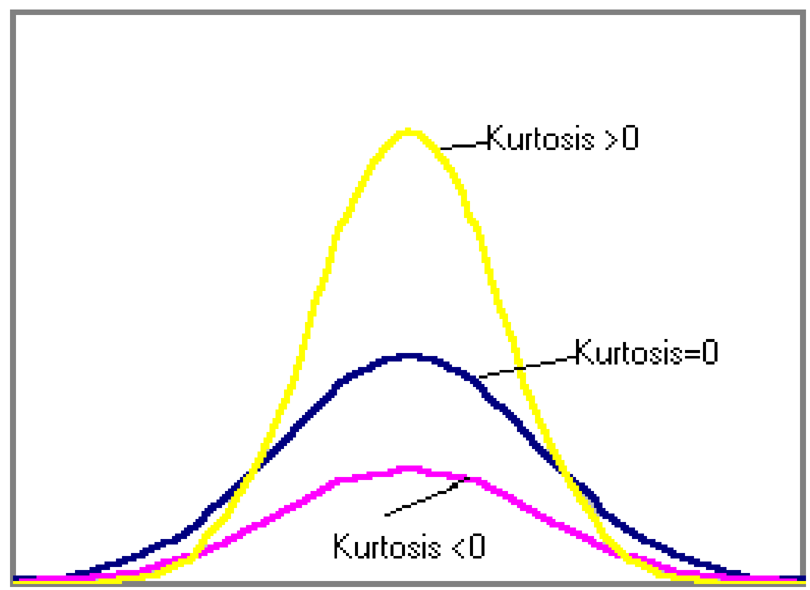
$$kurt(X) = \frac{E\left[(X-\mu)^4\right]}{\sigma^4} - 3$$



Gaussian Dis. Kurtosis = 0

Super-Gaussian: kurtosis > 0

Sub-Gaussian: kurtosis < 0



- Remove the mean

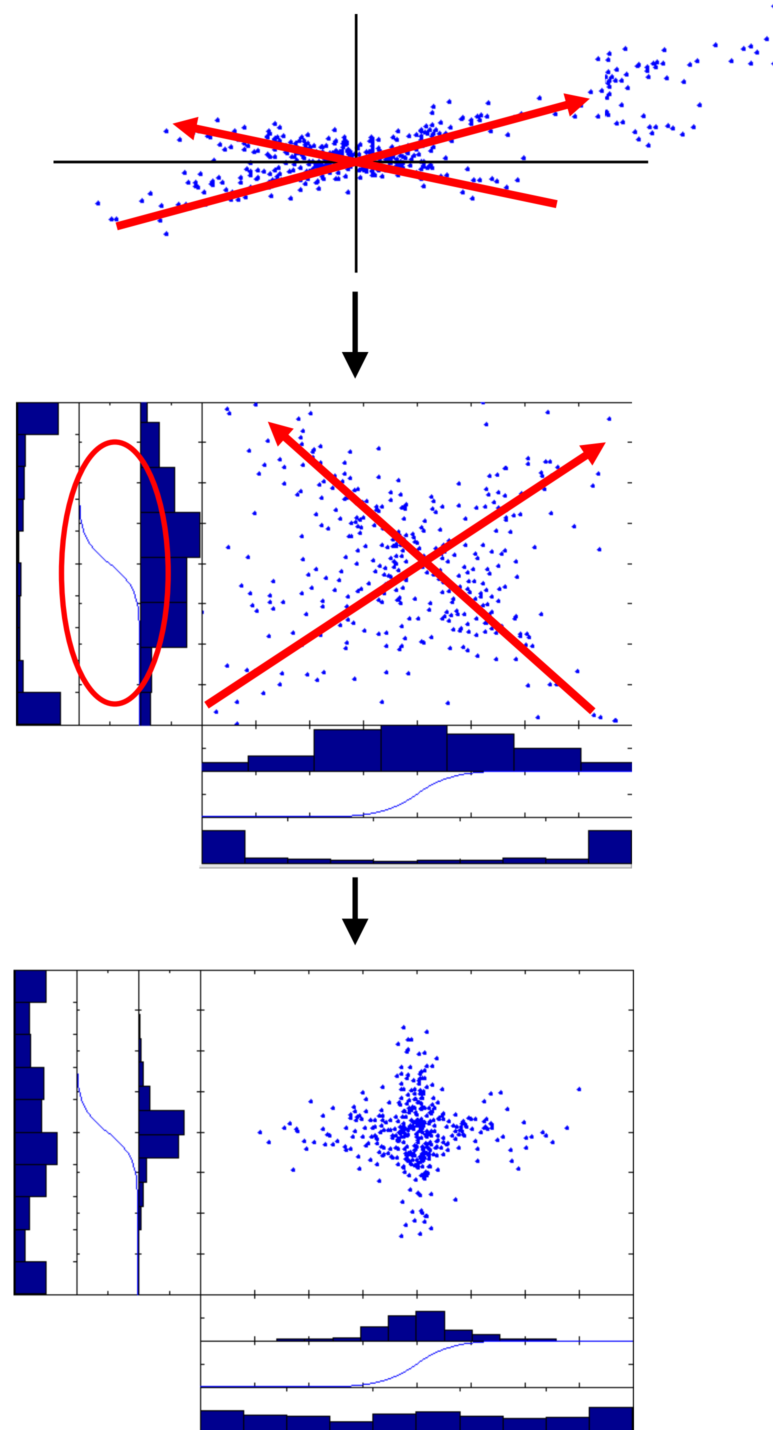
  $x = x - \langle x \rangle$.

- ‘Sphere’ the data by diagonalizing its covariance matrix,

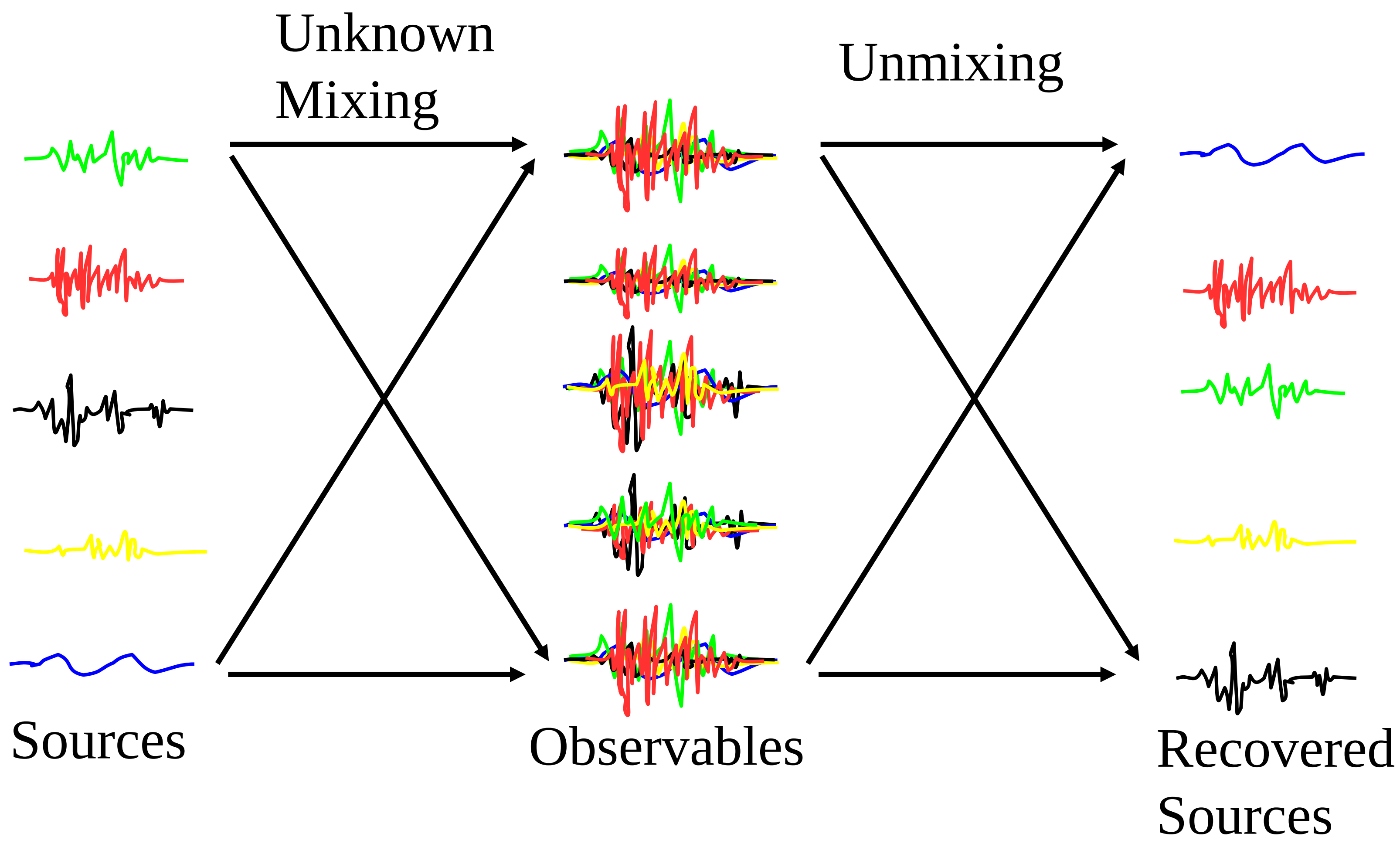
  $x = 2\langle xx^T \rangle^{-1/2}(x-\langle x \rangle)$.

- Update W according to

$$\Delta \mathbf{W} \propto \frac{\partial H(\mathbf{y})}{\partial \mathbf{W}} \mathbf{W}^T \mathbf{W} = \left[\mathbf{I} + \phi \mathbf{u}^T\right] \mathbf{W}$$



Blind Source Separation



Solving a BSS problem

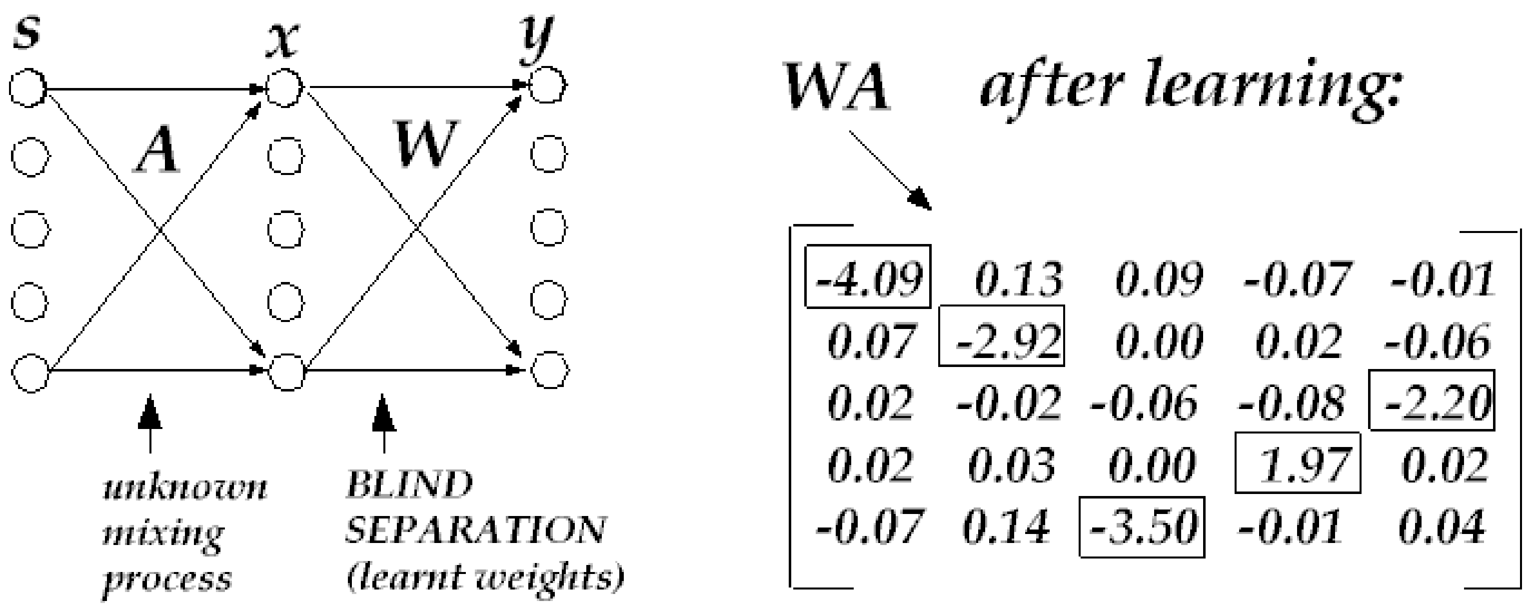
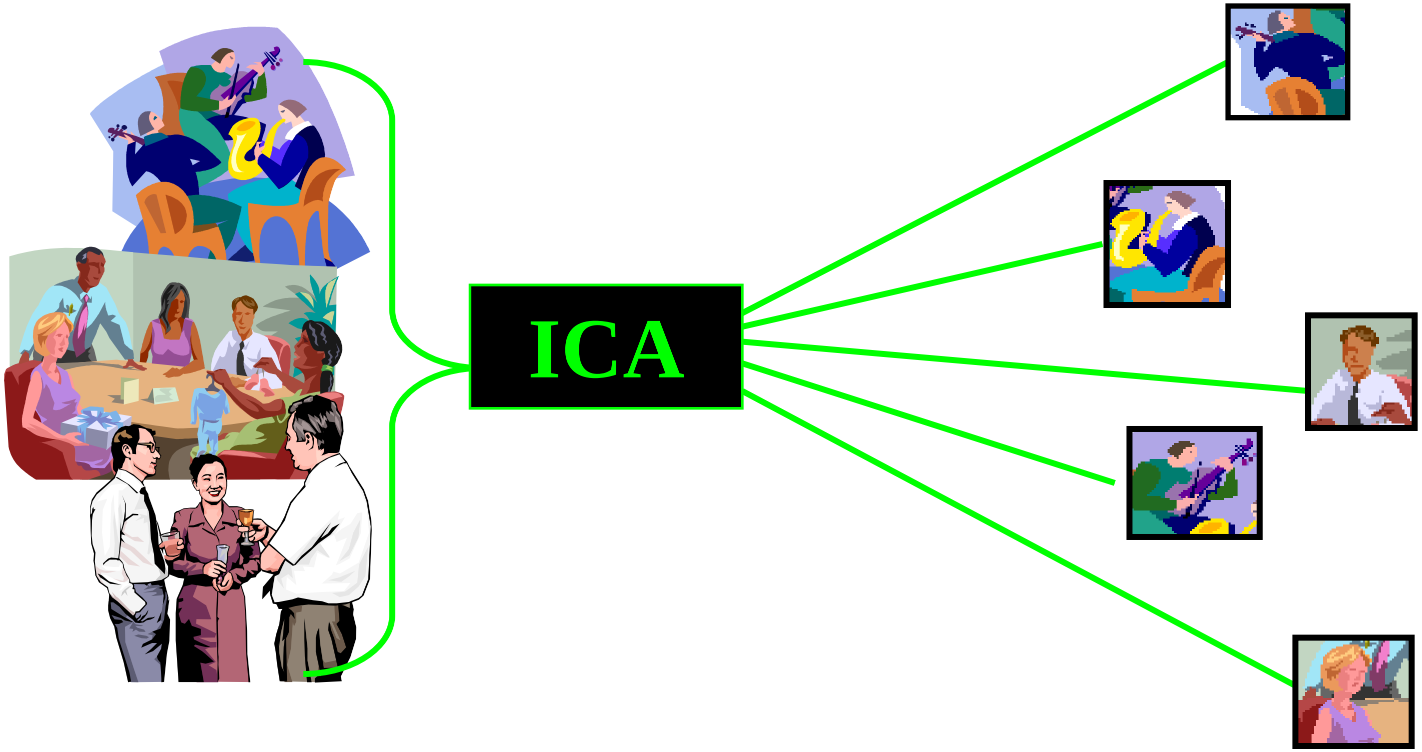


Figure 2: (a) In blind separation, sources, **s**, have been linearly scrambled by a matrix, **A**, to form the inputs to the network, **x**. We must recover the sources at our output **y**, by somehow inverting the mapping **A** with our weight matrix **W**. The problem: we know nothing about **A** or the sources. (b) A successful 'unscrambling' occurs when **WA** is a 'permutation' matrix. This one resulted from separating five speech signals with our algorithm.

Bell & Sejnowski, NIPS 1994

Cocktail Party Problem

Solving a Cocktail Party Problem



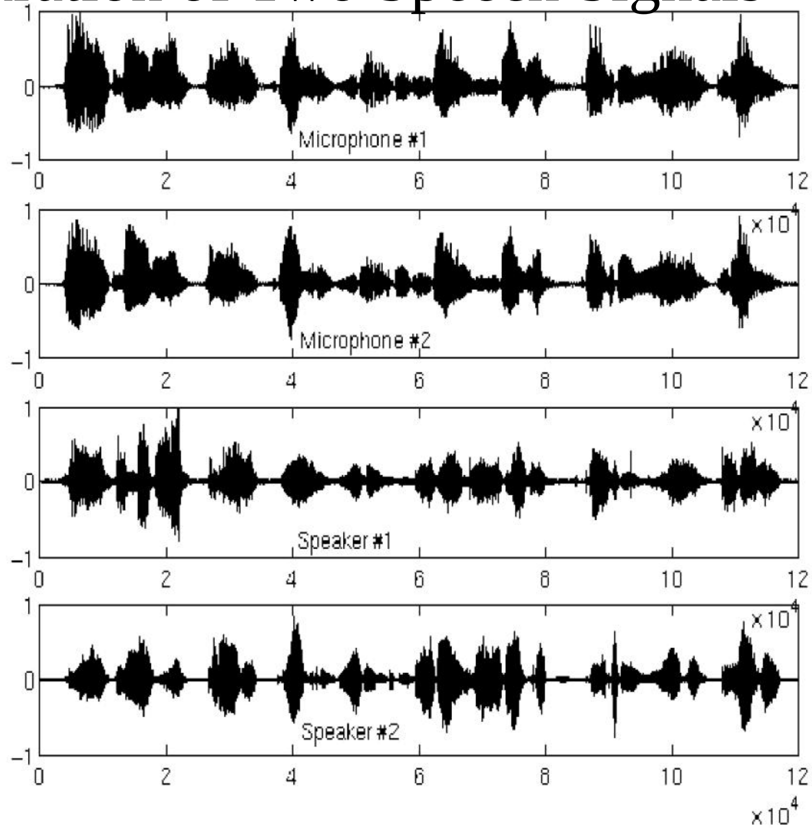
Speech Enhancement & Recognition

Improves speech recognition rate after separation
Algorithm works for various sounds in different environments.

Park and Lee (1999):

SNR [dB]	*W/o sep.*	*With sep.*
15 dB	87.8%	90.8%
10 dB	68.9%	87.9%
5 dB	37.0%	79.9%

Separation of Two Speech Signals



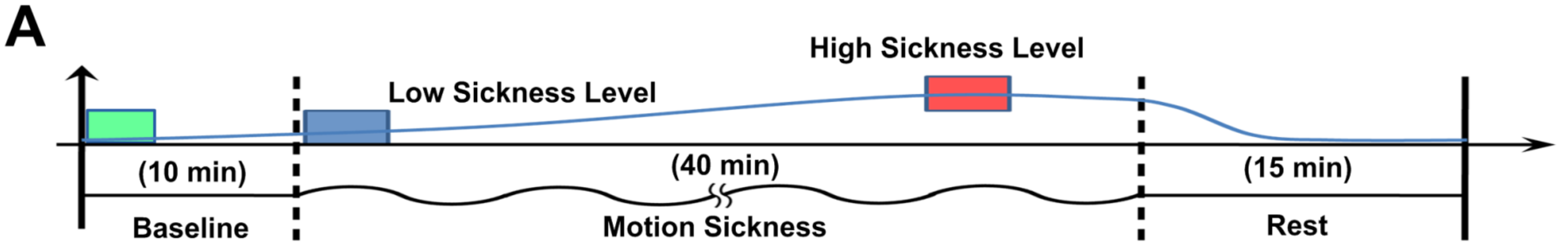
ICA EEG Analysis

Examples

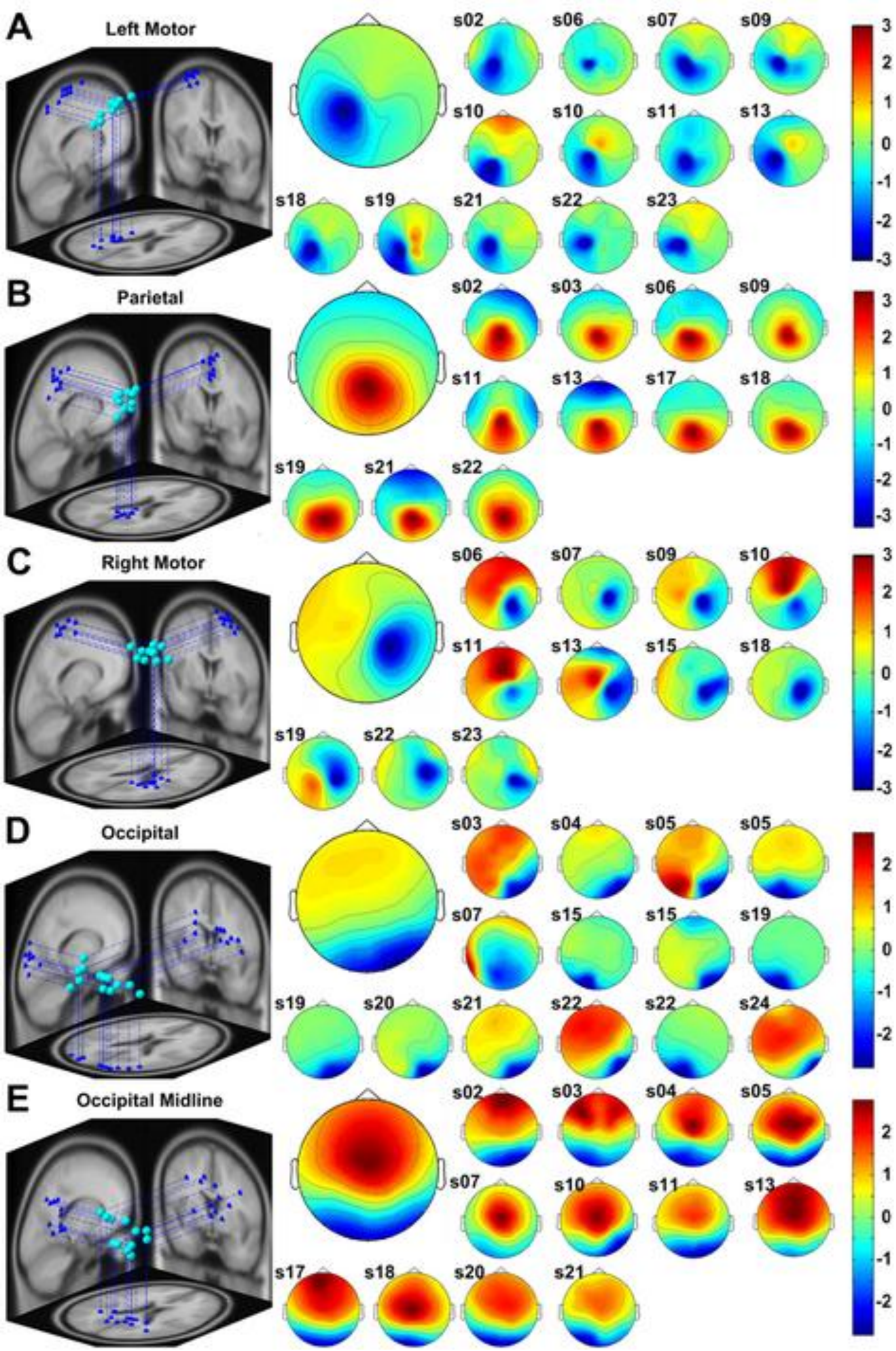
II. Brain Dynamics and Motion Sickness

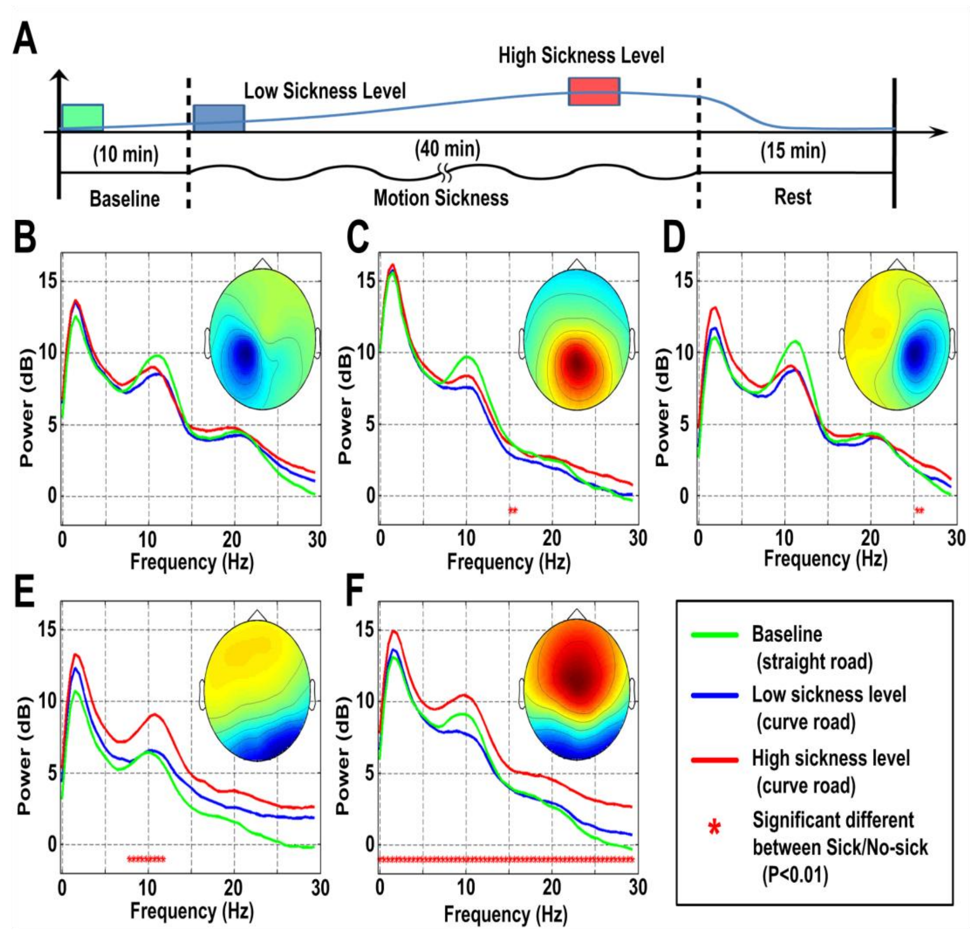
When single-trial analysis is NOT possible: estimating state changes

Experimental Paradigm



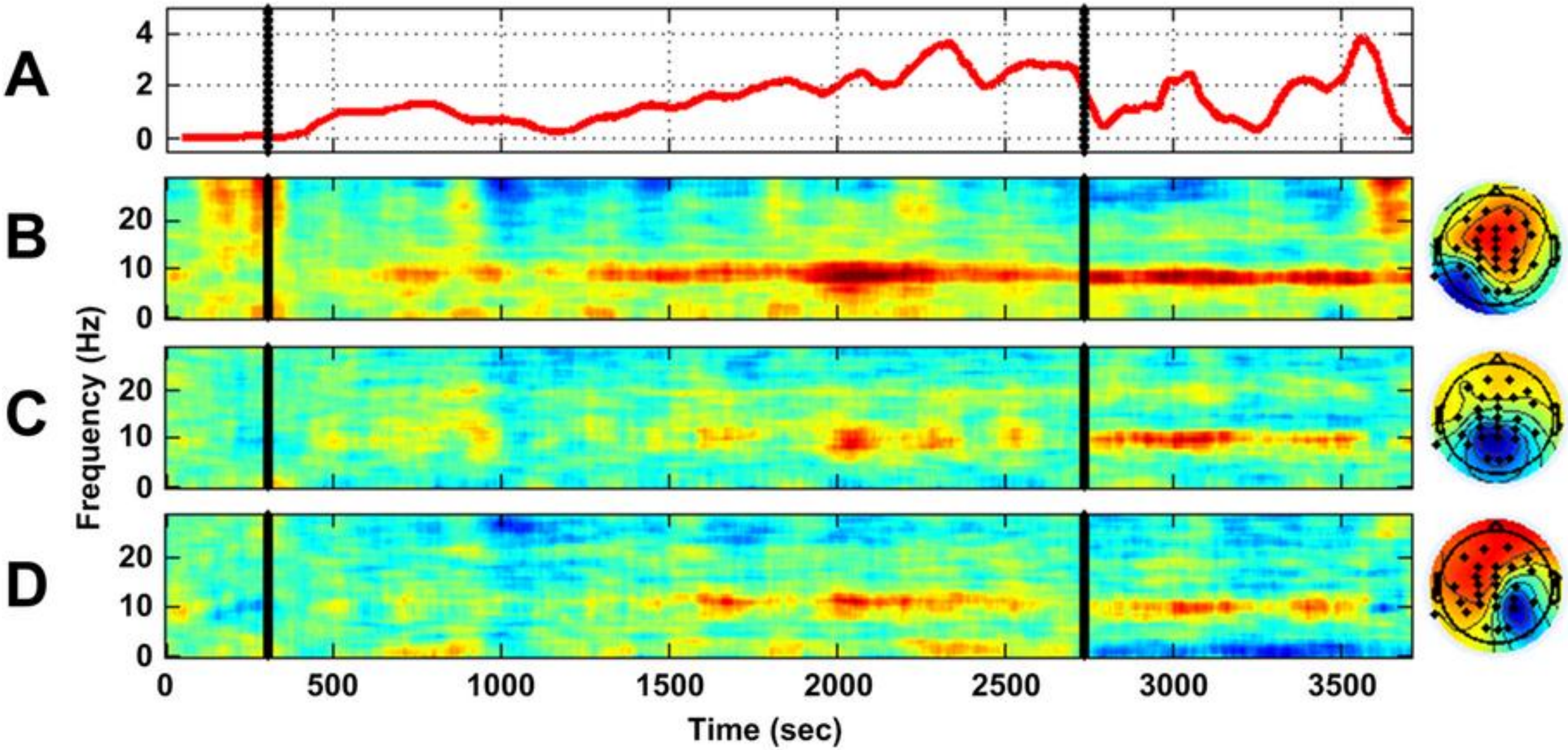
Chen et al., (2009)

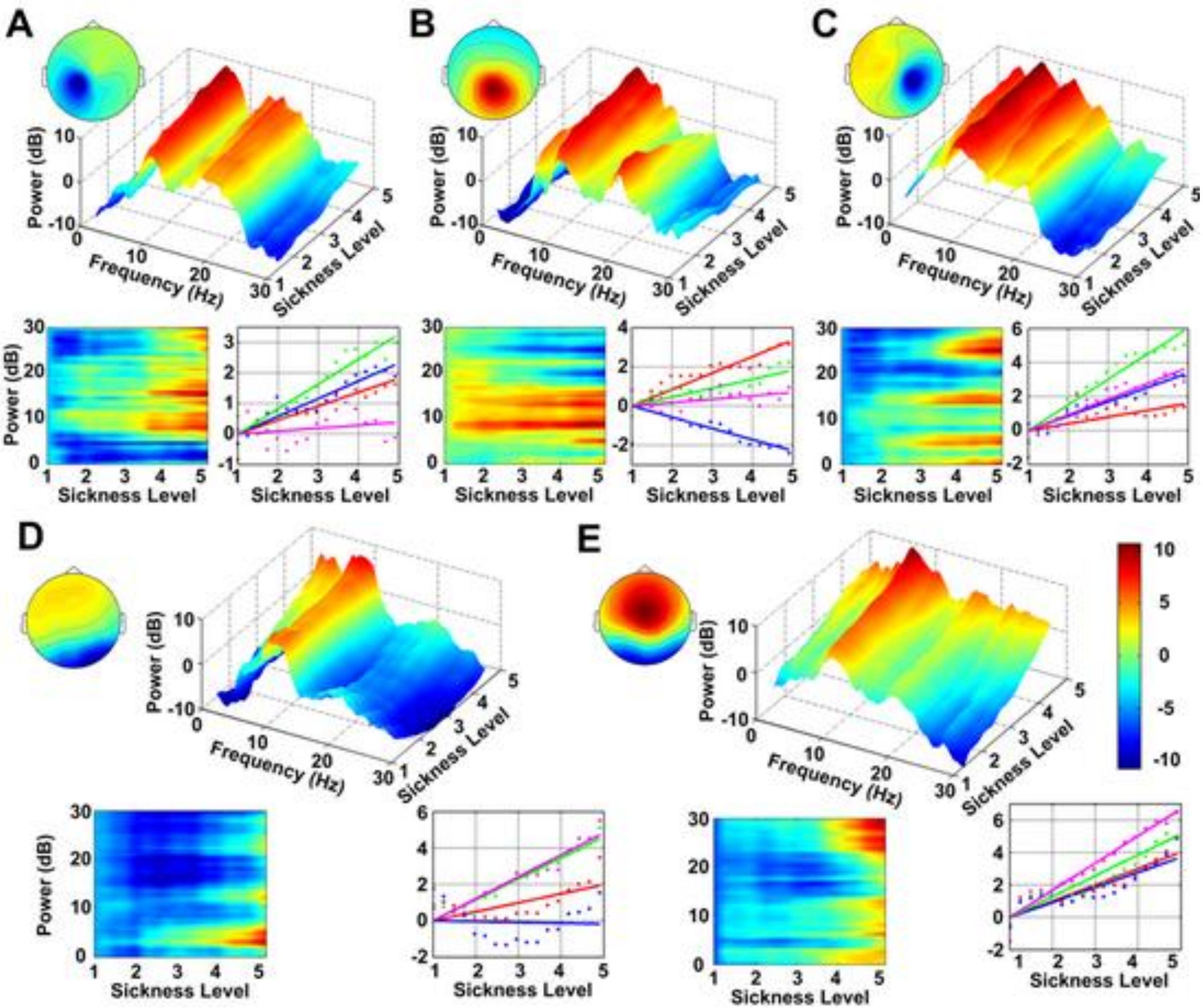




EEG Spectrum Changes

Group Motion-sickness Related Components




A
B
C
D
E
Power (dB)
Frequency (Hz)
Sickness Level
0
10
20
30
1
2
3
4
5
-10
10
5
0
-5
-10