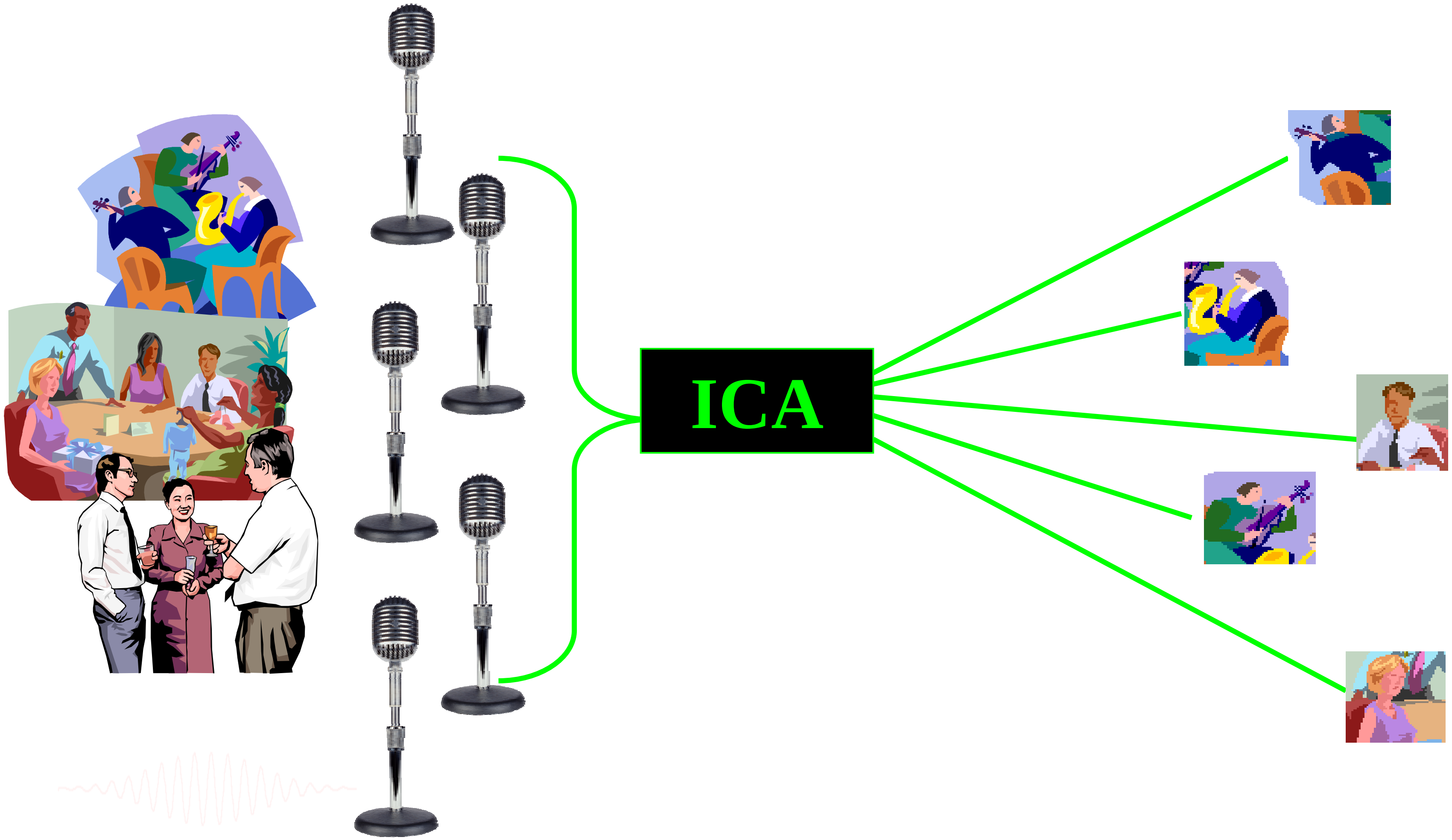
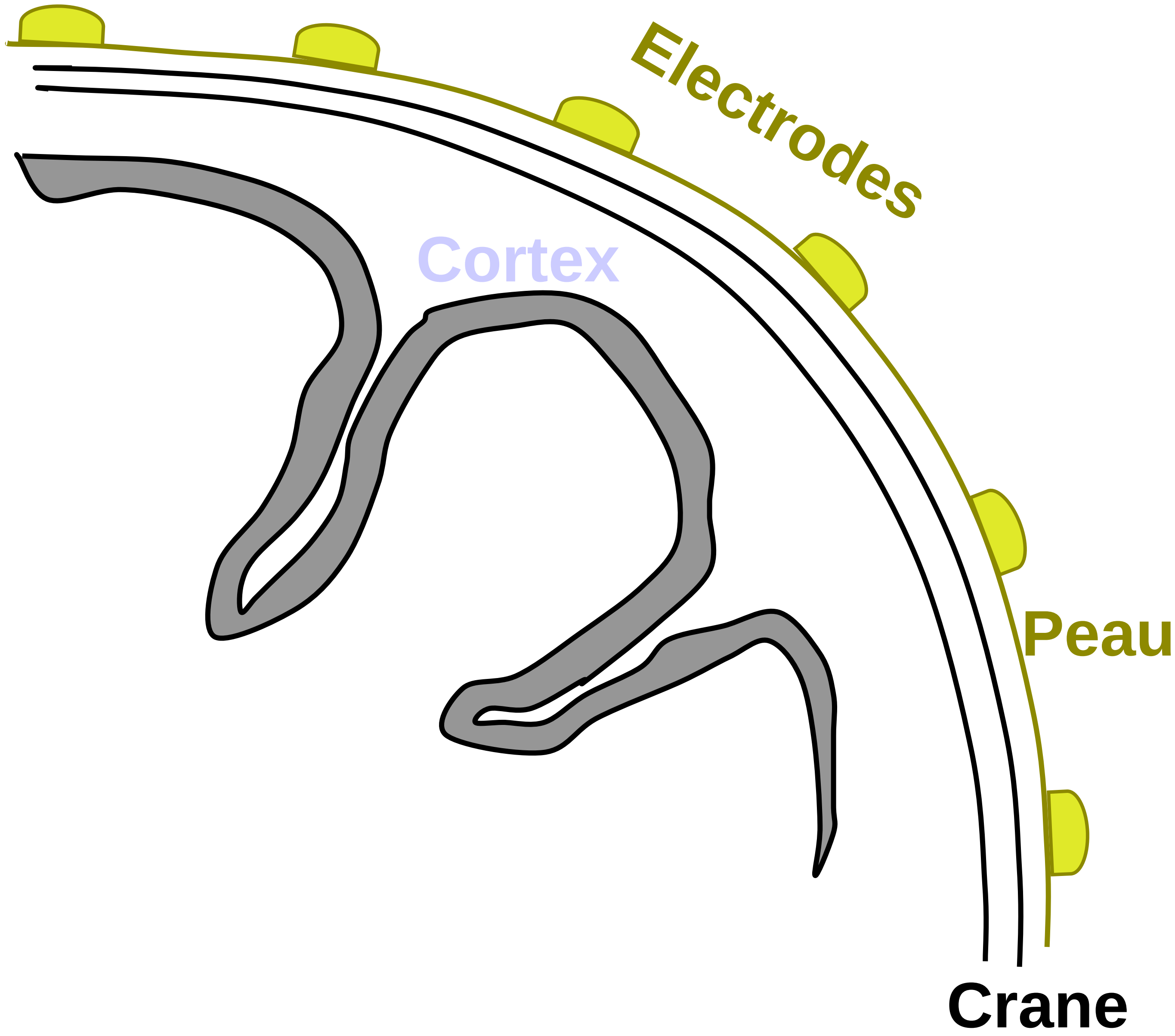
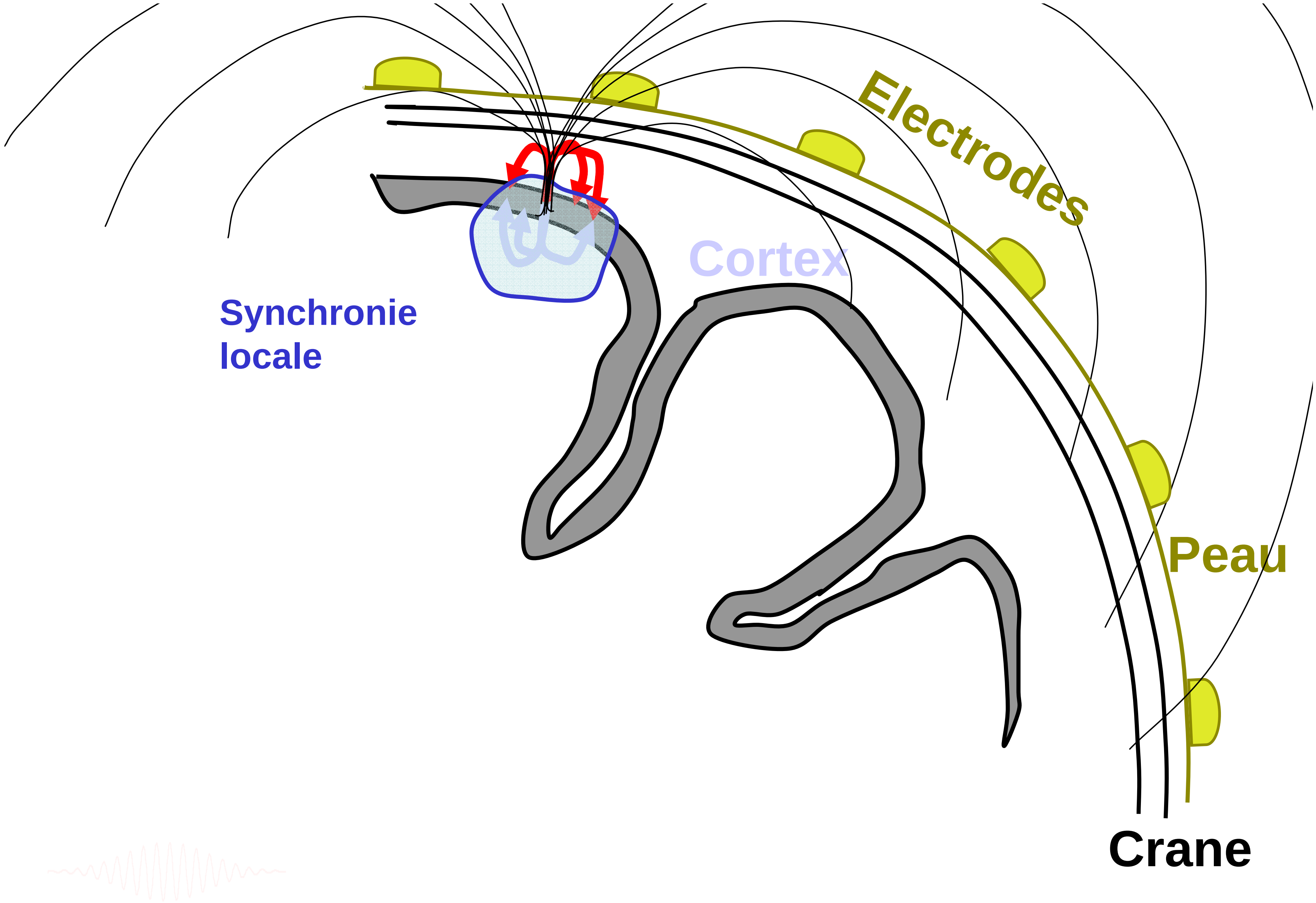
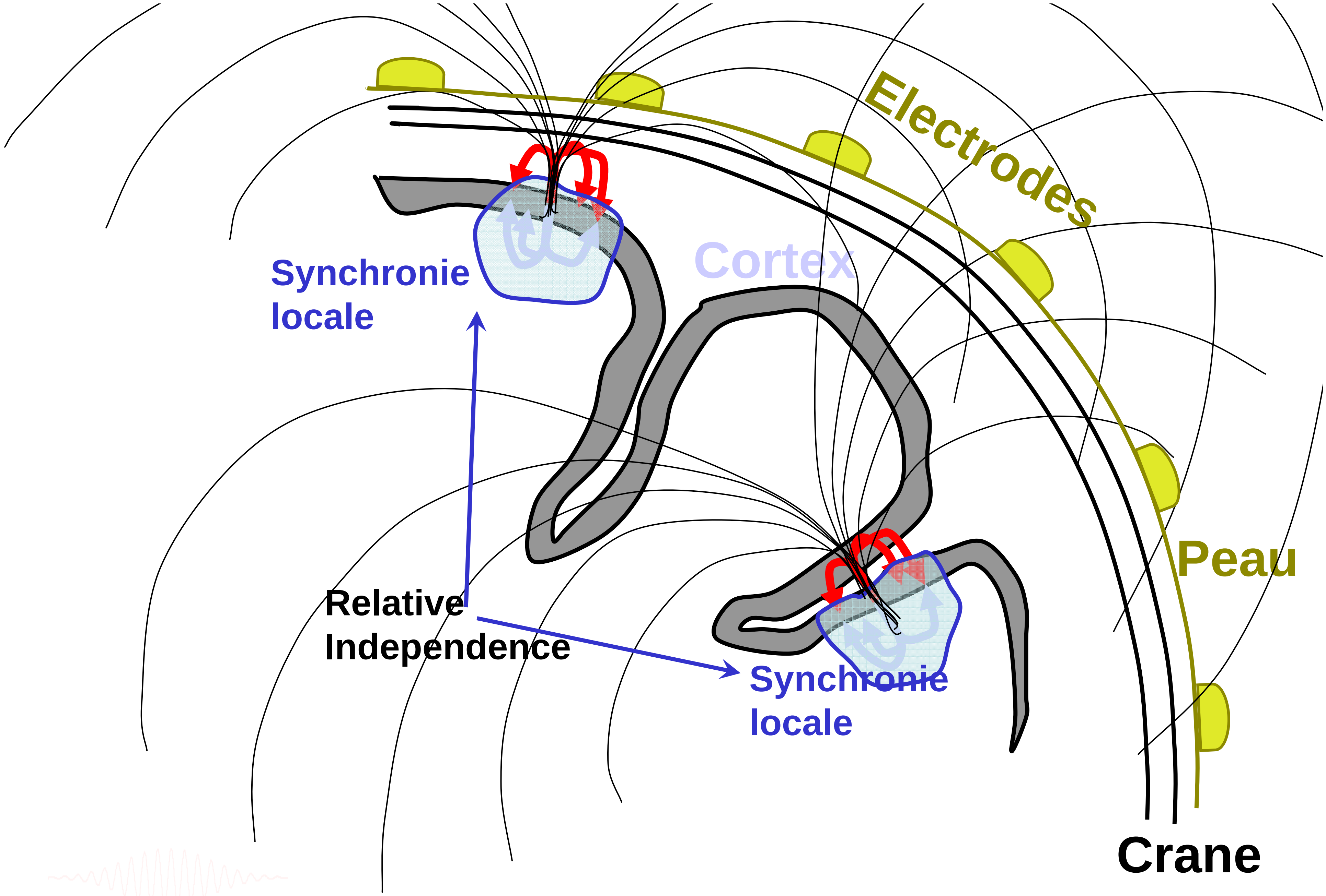
Independent component analysis applied to biophysical time series and EEG

Example: Speech Separation



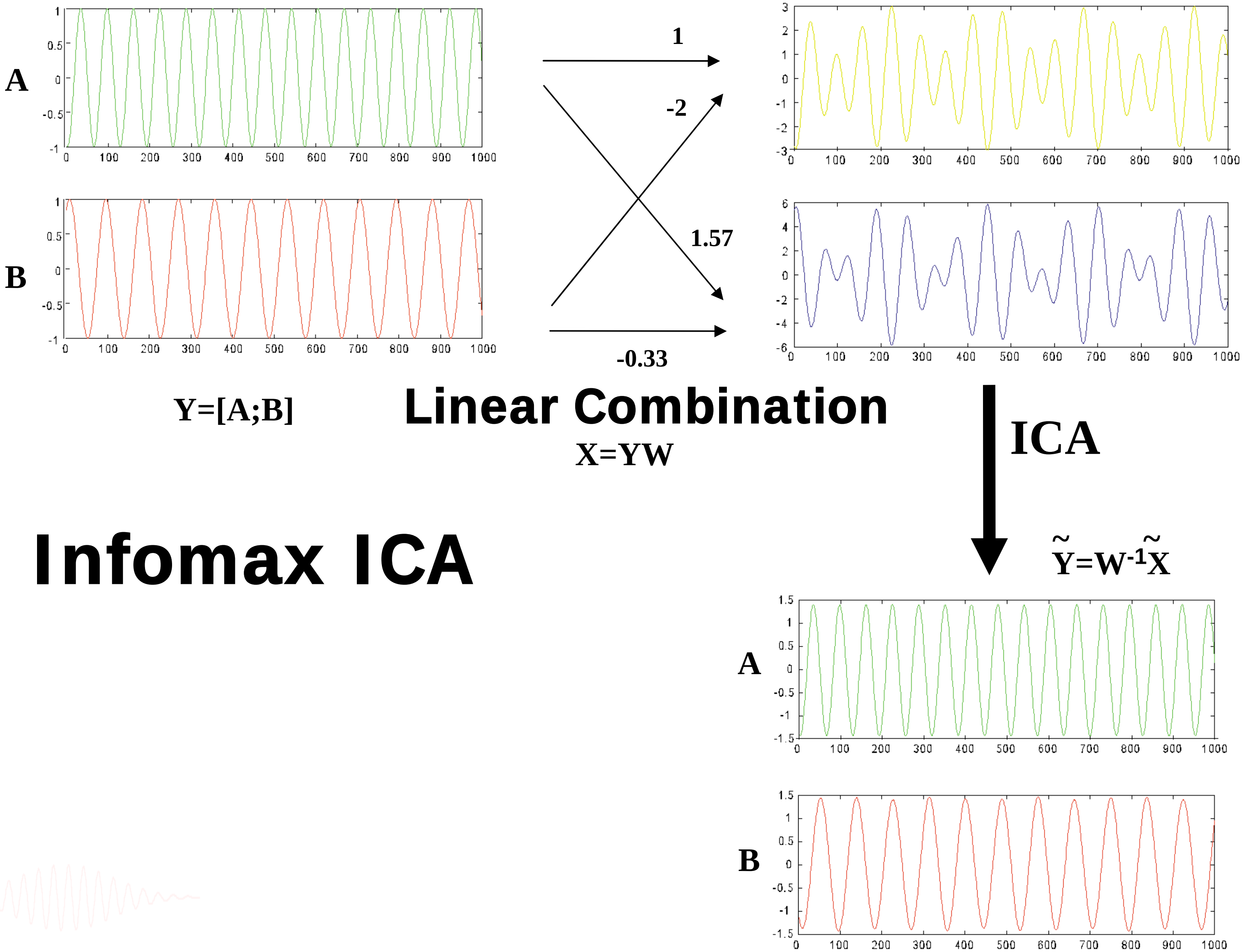

Electrodes
Cortex
Peau
Crane


Electrodes
Cortex
Synchronie
locale
Peau
Crane



Electrodes
Synchronie
locale
Cortex
Relative
Independence
Synchronie
locale
Peau
Crane

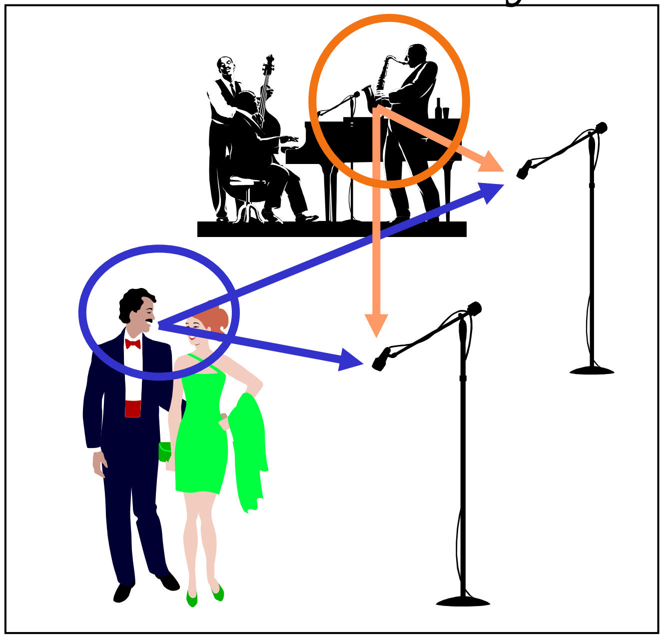
Infomax ICA



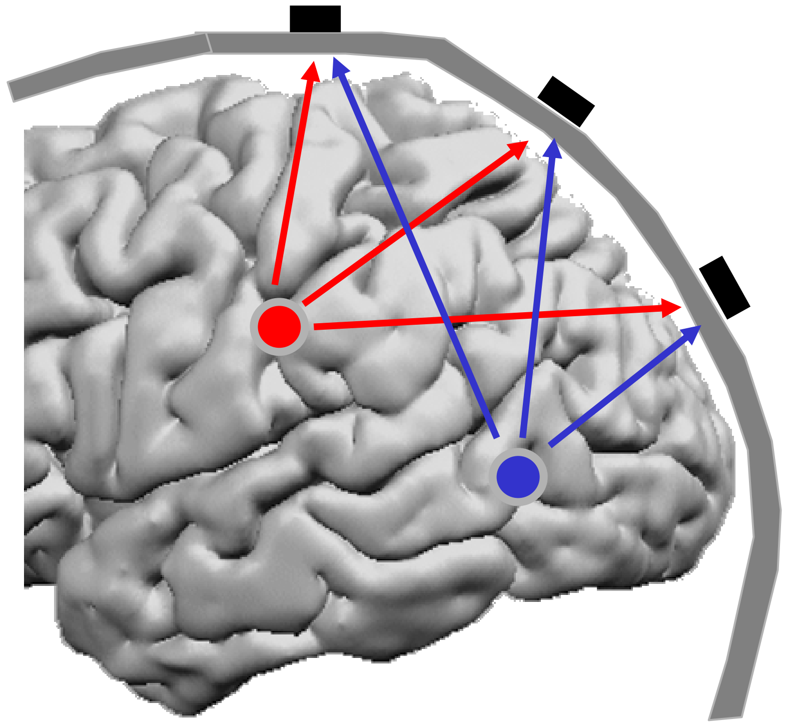
A
B
1
-2
1.57
-0.33
Y=[A;B]
Linear Combination
X=YW
ICA
$\tilde{Y}=W^{-1}\tilde{X}$
A
B

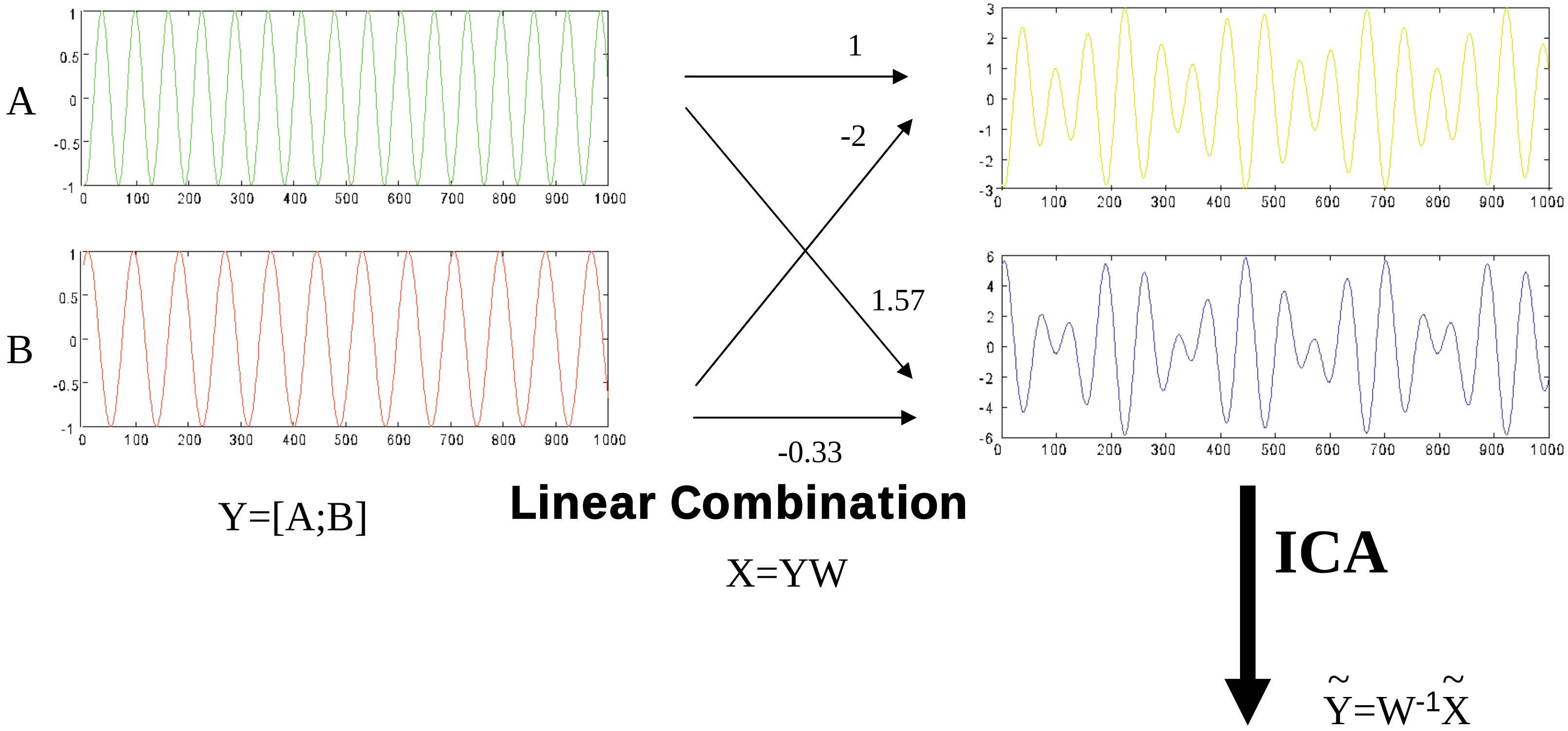
Independent component analysis

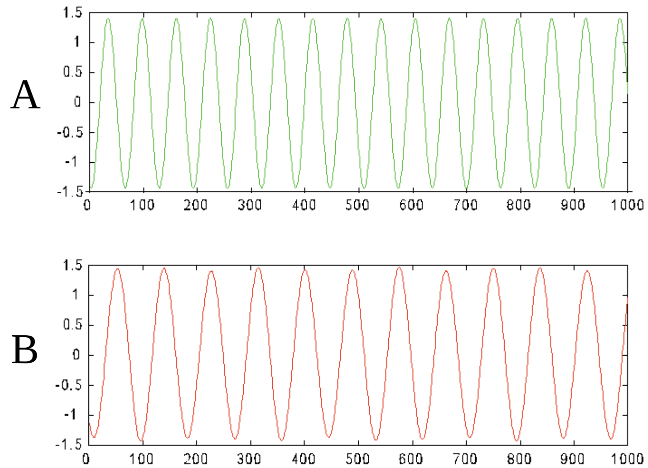
Cocktail Party



Mixture of Brain source activity







ICA is a method to recover a version, of the original sources by multiplying the data by a unmixing matrix

$$\mathbf{U} = \mathbf{W}\mathbf{X}$$

ICA activity

Data

Data X

$$\begin{array}{l}\left[\begin{array}{ccccccc} 3 & 2 & 5 & 4 & 3 & 2 & \\ 0 & -2 & -5 & -1 & 1 & -1 & \ldots \\ -1 & 2 & 0 & 1 & 0 & -3 & \end{array}\right] \end{array} \begin{array}{l} \leftarrow \text{Channel 1} \\ \leftarrow \text{Channel 2} \\ \leftarrow \text{Channel 3} \end{array}$$

$$\begin{bmatrix} 5 & 3 & -2 \\ 1 & 2 & 4 \\ 0 & -1 & 3 \end{bmatrix} \longrightarrow \left[\begin{array}{ccc} 3*5+0*3-1*(-2) & 2*5+(-2)*3+2*(-2) & \\ 3*1+0*2-1*4 & 2*1+(-2)*2+2*4 & \ldots \\ 5*1-5*2+0*4 & 5*1-5*2+0*4 & \end{array}\right] \begin{array}{l} \leftarrow \text{Comp. 1} \\ \leftarrow \text{Comp. 2} \\ \leftarrow \text{Comp. 3} \end{array}$$

Weight matrix W

ICA activity U

$$\mathbf{X} = \mathbf{W}^{-1}\mathbf{U}$$

Data

ICA activity U

$$\begin{array}{cl} \begin{bmatrix} 3 & 2 & 5 & 4 & 3 & 2 & \\ 0 & -2 & -5 & -1 & 1 & -1 & \ldots \\ -1 & 2 & 0 & 1 & 0 & -3 & \end{bmatrix} & \begin{array}{l} \leftarrow \text{Comp. 1} \\ \leftarrow \text{Comp. 2} \\ \leftarrow \text{Comp. 3} \end{array} \end{array}$$

$$\begin{bmatrix} 5 & 3 & -2 \\ 1 & 2 & 4 \\ 0 & -1 & 3 \end{bmatrix} \longrightarrow \begin{array}{cl} \begin{bmatrix} 3*5+0*3-1*(-2) & 2*5+(-2)*3+2*(-2) & \\ 3*1+0*2-1*4 & 2*1+(-2)*2+2*4 & \ldots \\ 5*1-5*2+0*4 & 5*1-5*2+0*4 & \end{bmatrix} & \begin{array}{l} \leftarrow \text{Chan 1} \\ \leftarrow \text{Chan 2} \\ \leftarrow \text{Chan 3} \end{array} \end{array}$$

Inverse weight matrix W^{-1}

Data X

Historical Remarks

- Herault & Jutten ("Space or time adaptive signal processing by neural network models“, *Neural Nets for Computing Meeting*, Snowbird, Utah, 1986): Seminal paper, neural network
- Bell & Sejnowski (1995): Information Maximization
- Amari et al. (1996): Natural Gradient Learning
- Cardoso (1996): JADE

- **Applications of ICA to biomedical signals**
  - EEG/ERP analysis (Makeig, Bell, Jung & Sejnowski, 1996).
  - fMRI analysis (McKeown et al. 1998)

ICA Theory – Cost Functions

Family of BSS algorithms

- Information theory (Infomax)
- Bayesian probability theory (Maximum likelihood estimation)
- Negentropy maximization
- Nonlinear PCA
- Statistical signal processing (cumulant maximization, JADE)

A unifying Information-theoretic framework for ICA

- Pearlmutter & Parra showed that InfoMax, ML estimation are equivalent.
- Lee et al. (1999) showed negentropy has the equivalent property to InfoMax.
- Girolami & Fyfe showed nonlinear PCA can be viewed from information-theoretic principle.

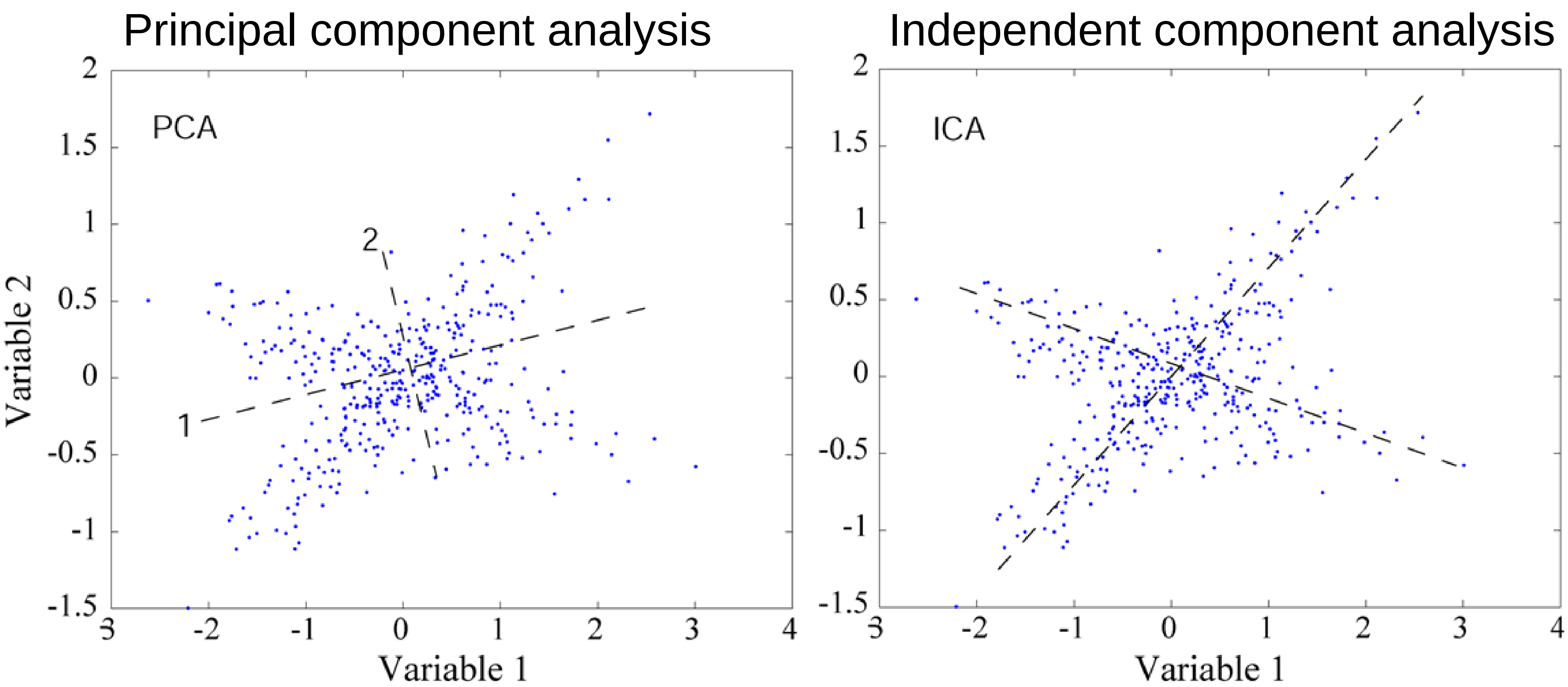
Independent Component Analysis

ICA is a method to recover a version, of the original sources by multiplying the data by a unmixing matrix,

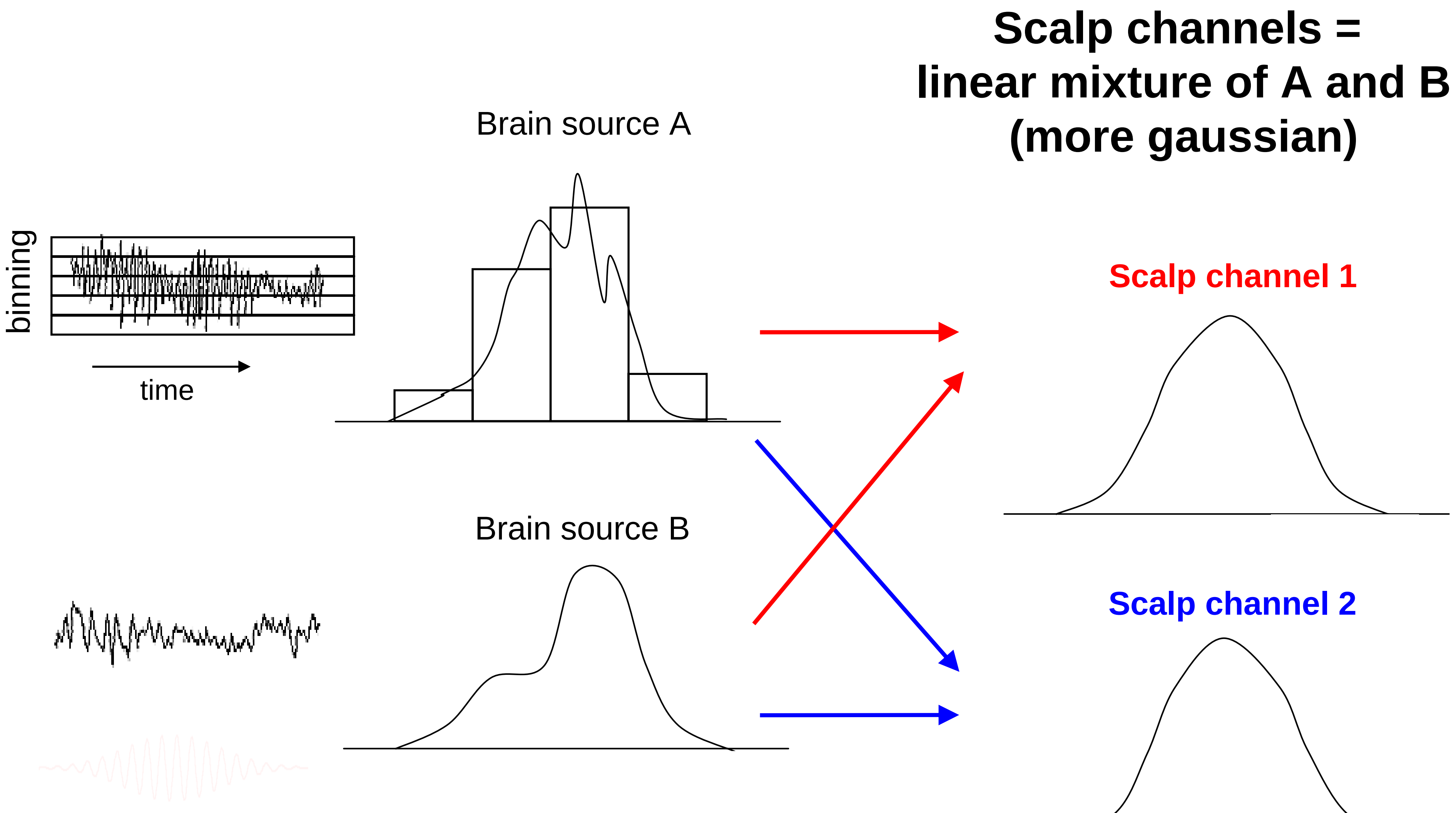
$$\mathbf{U} = \mathbf{WX},$$

While 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**.

ICA and PCA

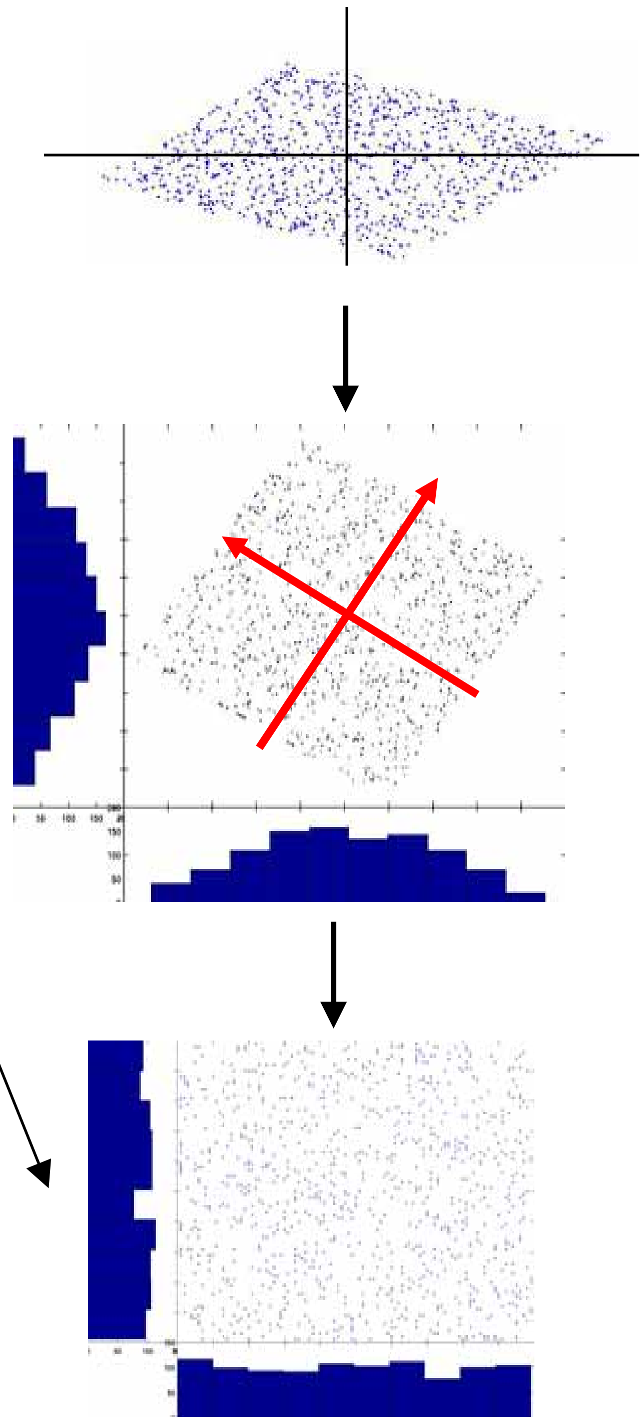

Principal component analysis
PCA
Variable 1
Variable 2
1
2
Independent component analysis
ICA
Variable 1
-3
-2
-1
0
1
2
3
4
-1.5
-1
-0.5
0
0.5
1
1.5
2

Central limit theorem



ICA Training Process

Central limit theorem



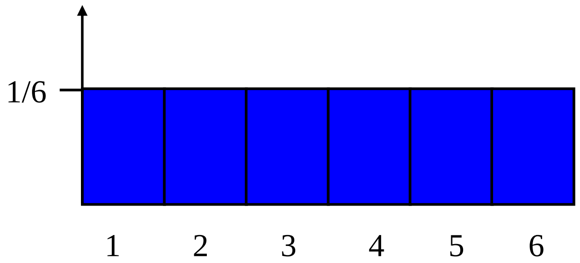
- Remove the mean
  $x = x - <x>$
- ‘Sphere’ the data by diagonalizing its covariance matrix,
  $x = <xx^T>^{-1/2}(x-<x>)$.
- Update W according to

$$\Delta \mathbf{w} \propto \frac{\partial H(\mathbf{y})}{\partial \mathbf{w}} \mathbf{w}^T \mathbf{w}$$

Entropy

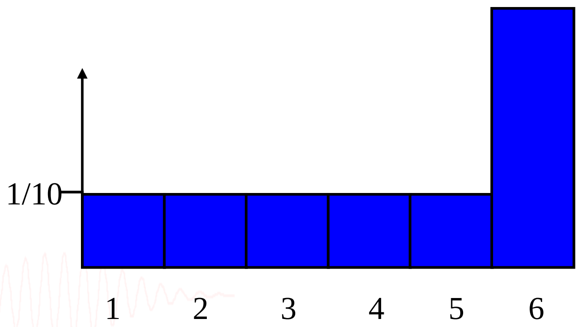
$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \log_b p(x).$$

Dice: 1/6



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

Fake dice (make a 6 half of the time): entropy 2.16 (base 2)



$$H = 5\left(-\frac{1}{10}\log_2\left(\frac{1}{10}\right)\right) - \frac{1}{2}\log_2\left(\frac{1}{2}\right) = 2.16$$

Entropy

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \log_b p(x)$$

Joint entropy

$$H(X,Y) = -\sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} p(x,y) \log_b p(x,y);$$

Mutual Information

$$H(y_1, y_2) = H(y_1) + H(y_2) - I(y_1, y_2).$$

Shannon in his landmark 1948 paper ``A Mathematical Theory of Communication.''

From http://planetmath.org/encyclopedia/ShannonsTheoremEntropy.html

Contingency table for stress and emotionality

	STRE						
	1	**2**	**3**	**4**	**5**	**6**	**Total**
EMOT= **1**	19	4					23
2	11	63	64	3	1		142
3	2	16	18	20	2	2	60
4	1	4	1	9	6	2	23
5			1	2	4	3	10
6				1	1	1	3
Total	33	87	84	35	13	8	

From http://tecfa.unige.ch/~lemay/thesis/THX-Doctorat/node149.html

Contingency frequencies for stress and emotionality

	STRE					
	1	**2**	**3**	**4**	**5**	**6**
EMOT= **1**	0.07	0.02				
2	0.04	0.24	0.25	0.01		
3	0.01	0.06	0.07	0.08	0.01	0.01
4		0.02		0.03	0.02	0.01
5				0.01	0.02	0.01
6						

Joint entropy 3.46; exercise: compute mutual information

$$H(X,Y) = -\sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} p(x,y)\log_b p(x,y)$$

ICA learning rule

How to make the outputs statistical independent?

Minimize their redundancy or mutual information.

Consider the joint entropy of two components,

$$H(y_1, y_2) = H(y_1) + H(y_2) - I(y_1, y_2).$$

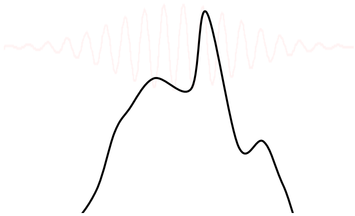
Maximizing $H(y_1, y_2) \Longrightarrow$ minimizing $I(y_1, y_2)$.

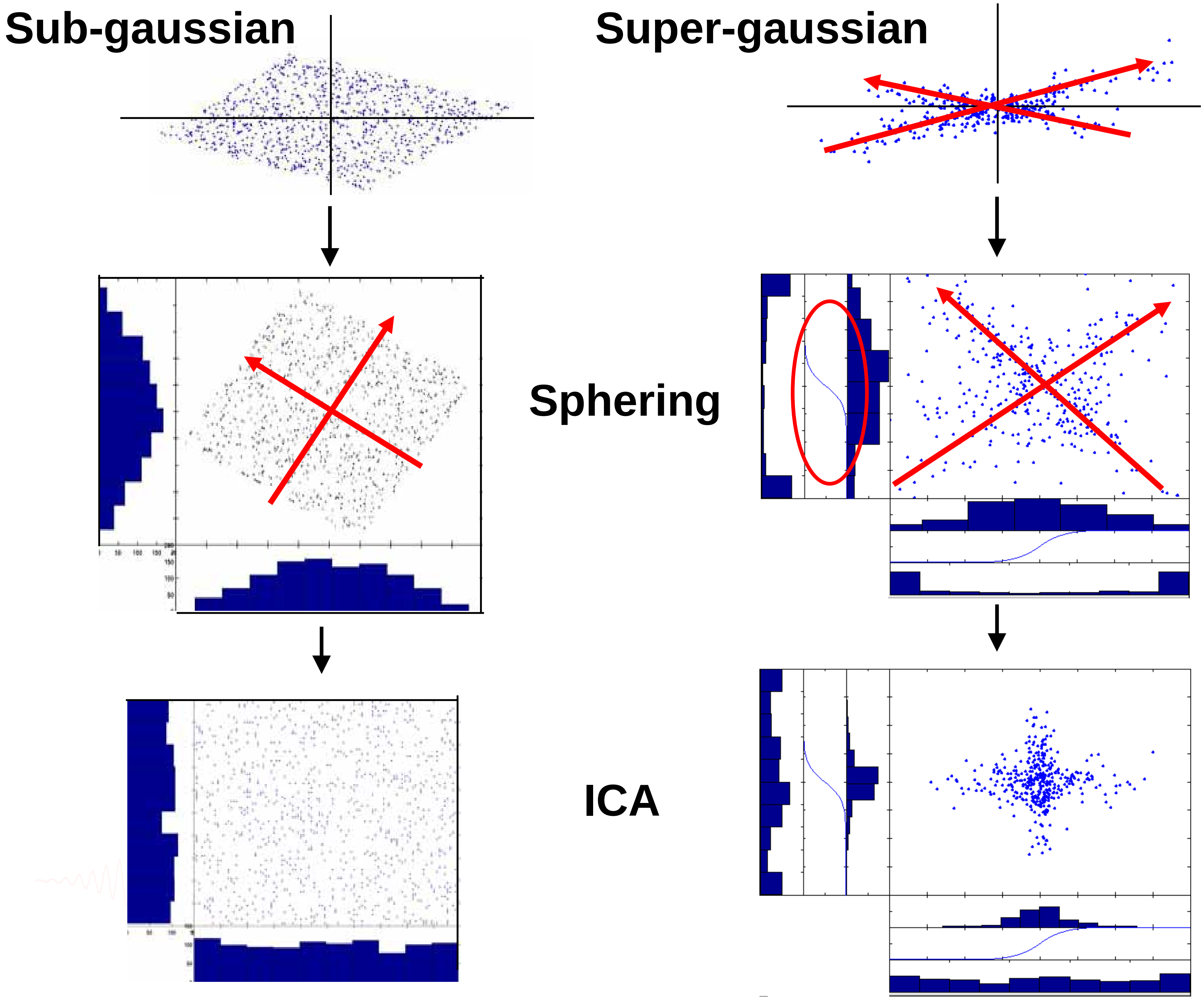
The learning rule:

=0 if the two variables are independent

$$\Delta \mathbf{W} \propto \frac{\partial H(\mathbf{y})}{\partial \mathbf{W}} \underbrace{\mathbf{W}^T \mathbf{W}}_{\text{Natural gradient (Amari)}}$$

Entropy extremum

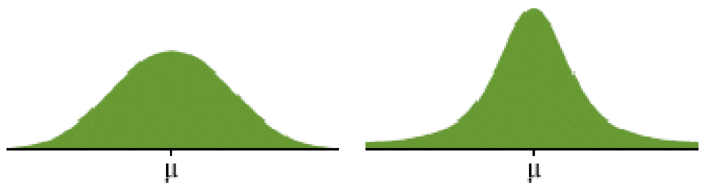



Sub-gaussian
Super-gaussian
Sphering
ICA

Kurtosis, Super- and Sub-Gaussian

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

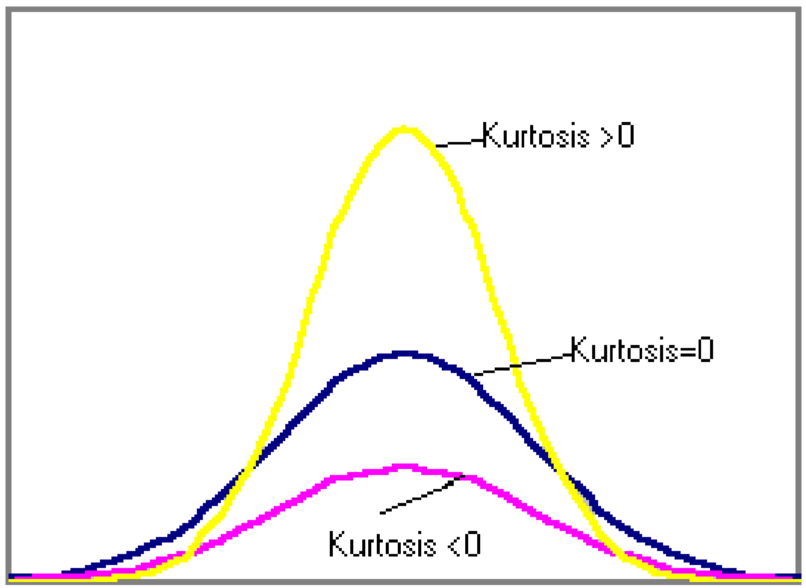
$$kurt(X) = \frac{E\left[(X-\mu)^4\right]}{\sigma^4}$$



Gaussian Dist. Kurtosis = 0

Super-Gaussian: kurtosis > 0

Sub-Gaussian: kurtosis < 0



Moments, Cumulants

Moments $\mu_x(n) = E\{x^n\}$

Central moments $m_x(n) = E\{(x - m_x)^n\}$

Cumulants

$$\begin{aligned} c_1 &= m_1 = \mu \\ c_2 &= m_2 = \sigma^2 \\ c_3 &= m_3 \\ c_4 &= m_4 - 3m_2^2 \end{aligned}$$

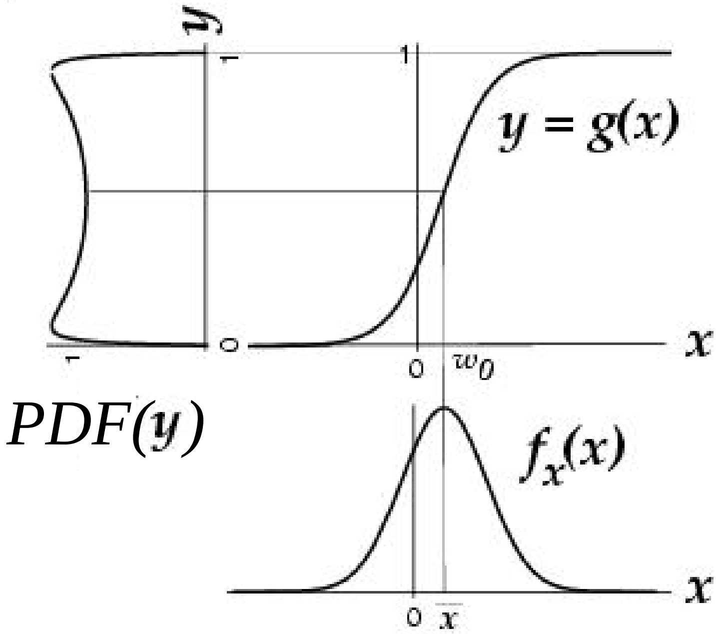
mean
variance
skewness
kurtosis

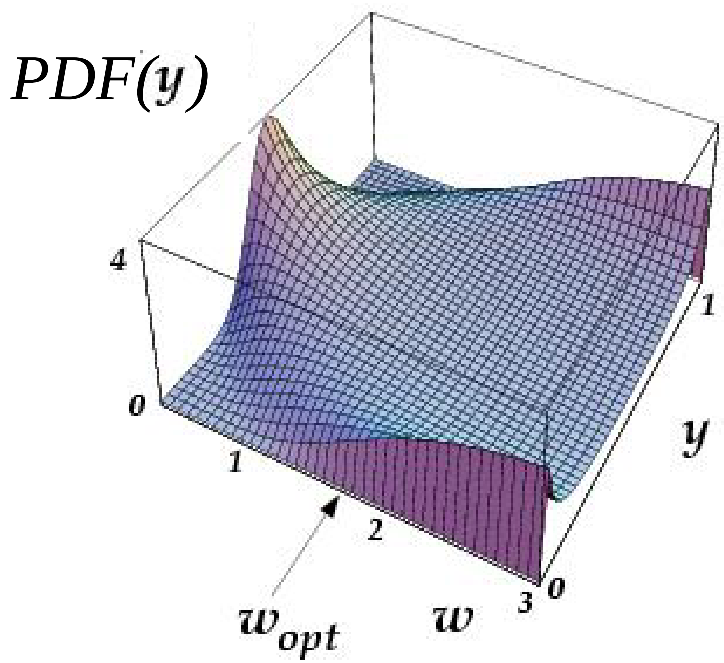
InfoMax (Bell & Sejnowski, 1995)

To make the u_i independent, we need to operate on non-linear transformed output variables, $\mathbf{y} = g(\mathbf{u})$, such as

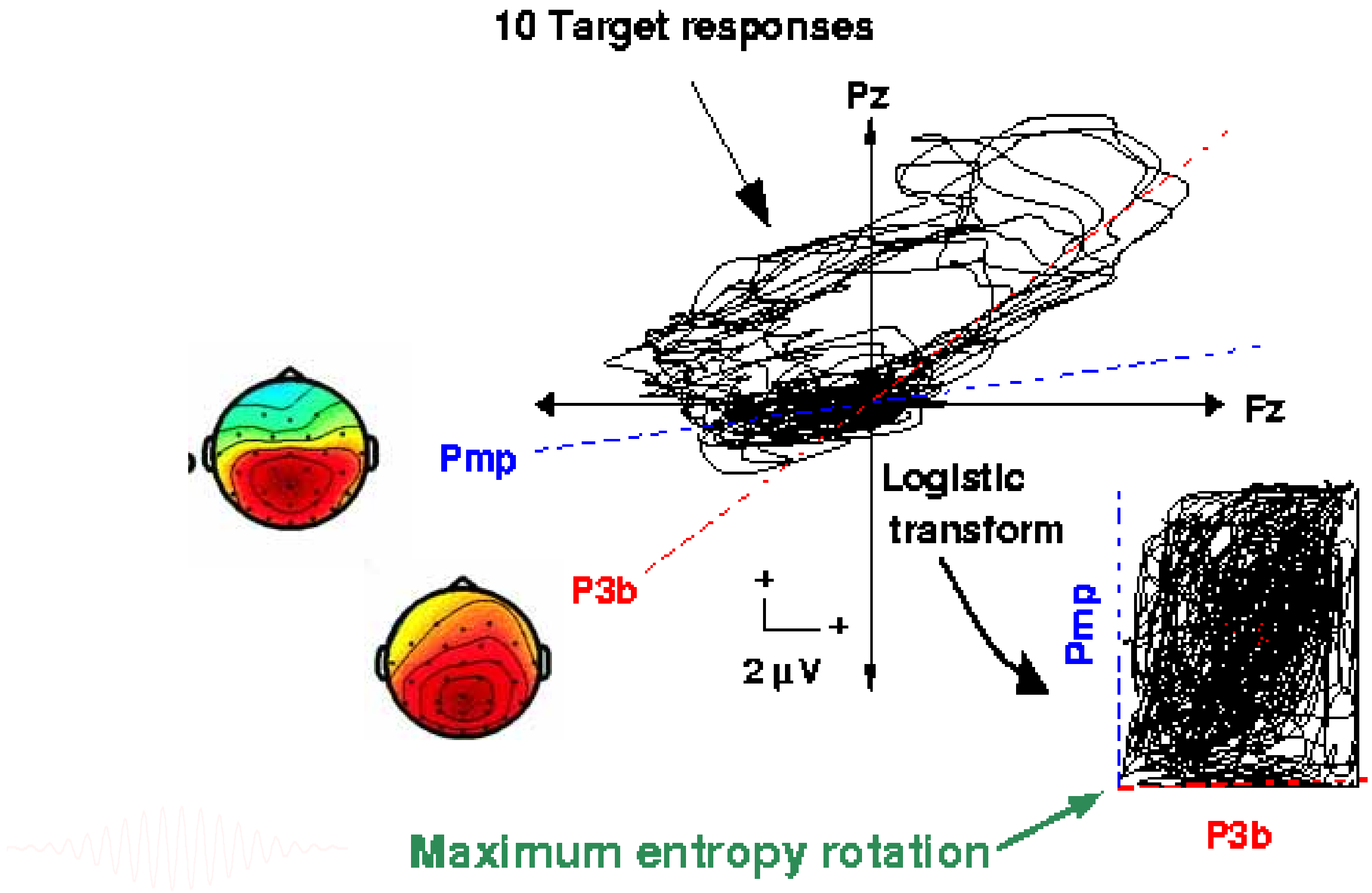
$$\mathbf{y} = \frac{1}{1+e^{-\mathbf{u}}}, \qquad \mathbf{u} = \mathbf{W}\mathbf{x} + \mathbf{w}_0$$

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



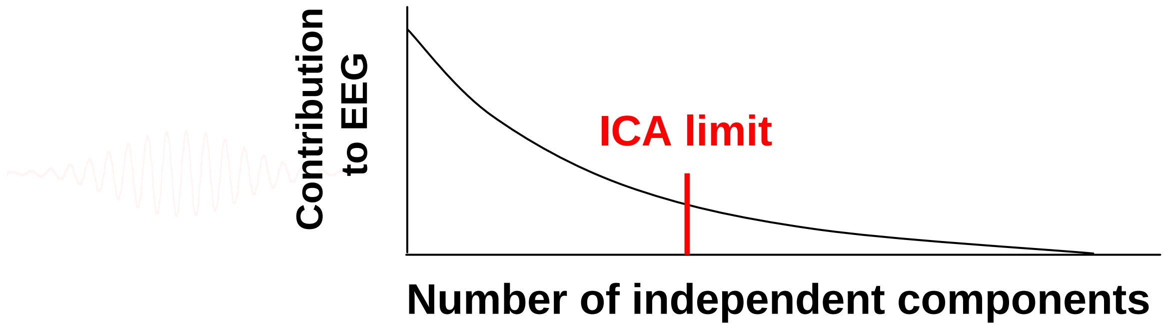


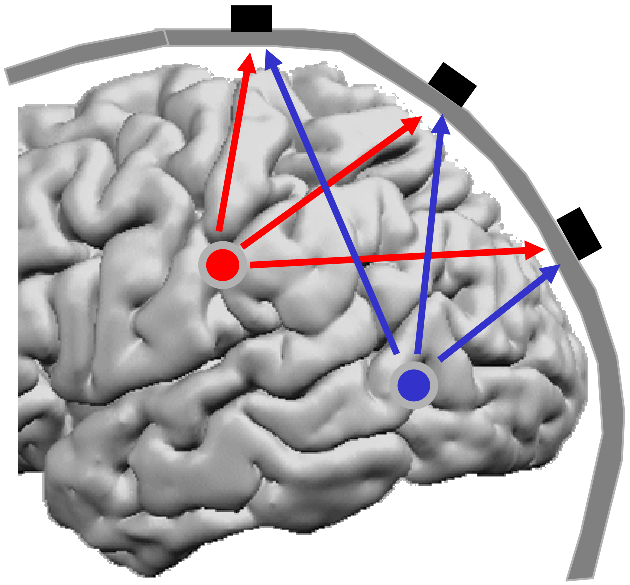
Independent components of EEG/ERP



ICA/EEG Assumptions

- Mixing is linear at electrodes **OK**
- Propagation delays are negligible **OK**
- Component time courses are independent **~**
- Number of components less than the number of channels.





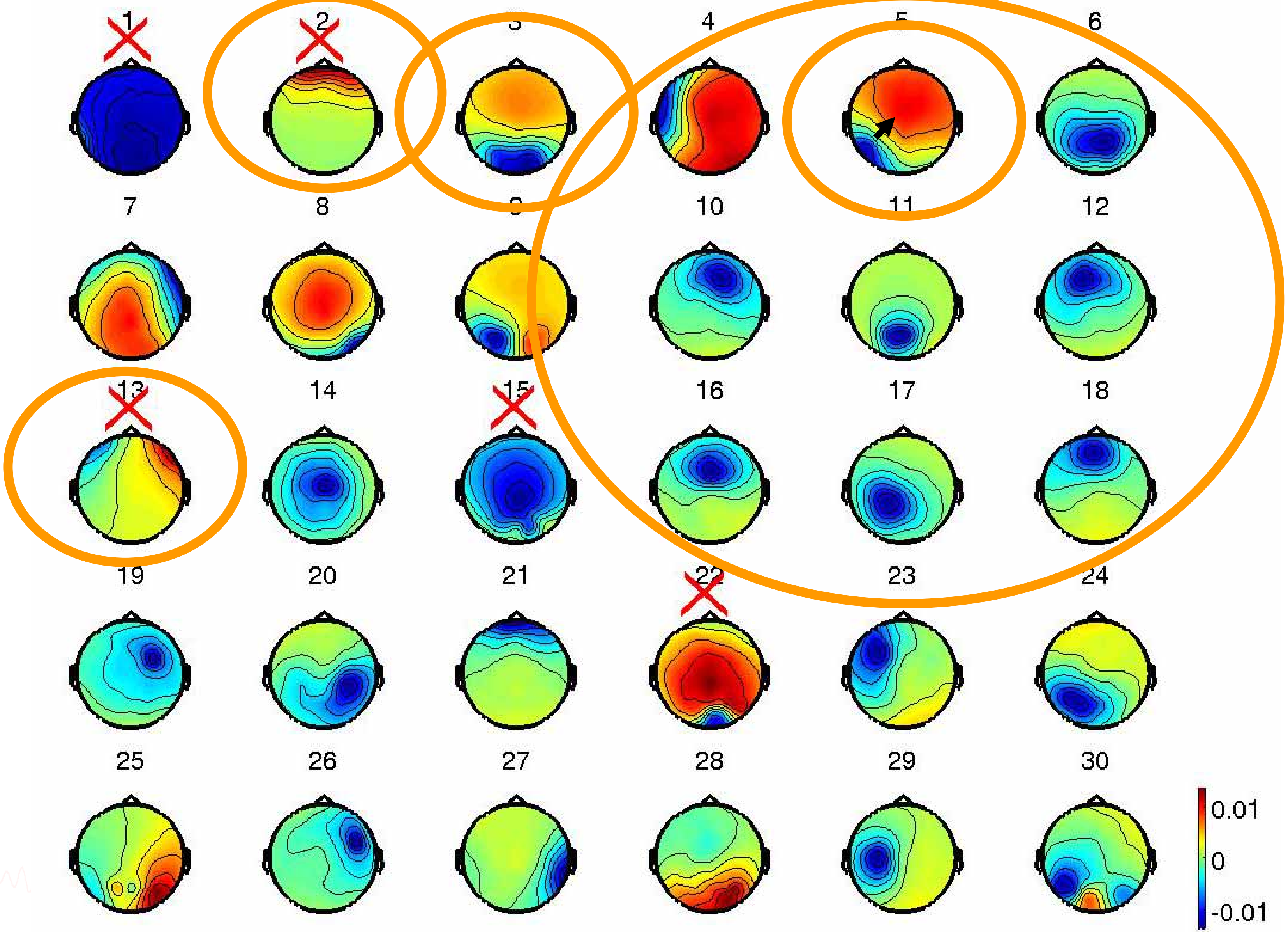
Independent Component Categories

- Artifacts
- Stimulus-locked activity
- Response-locked activity
- Non-phase locked activity
- Event-modulated oscillatory activity

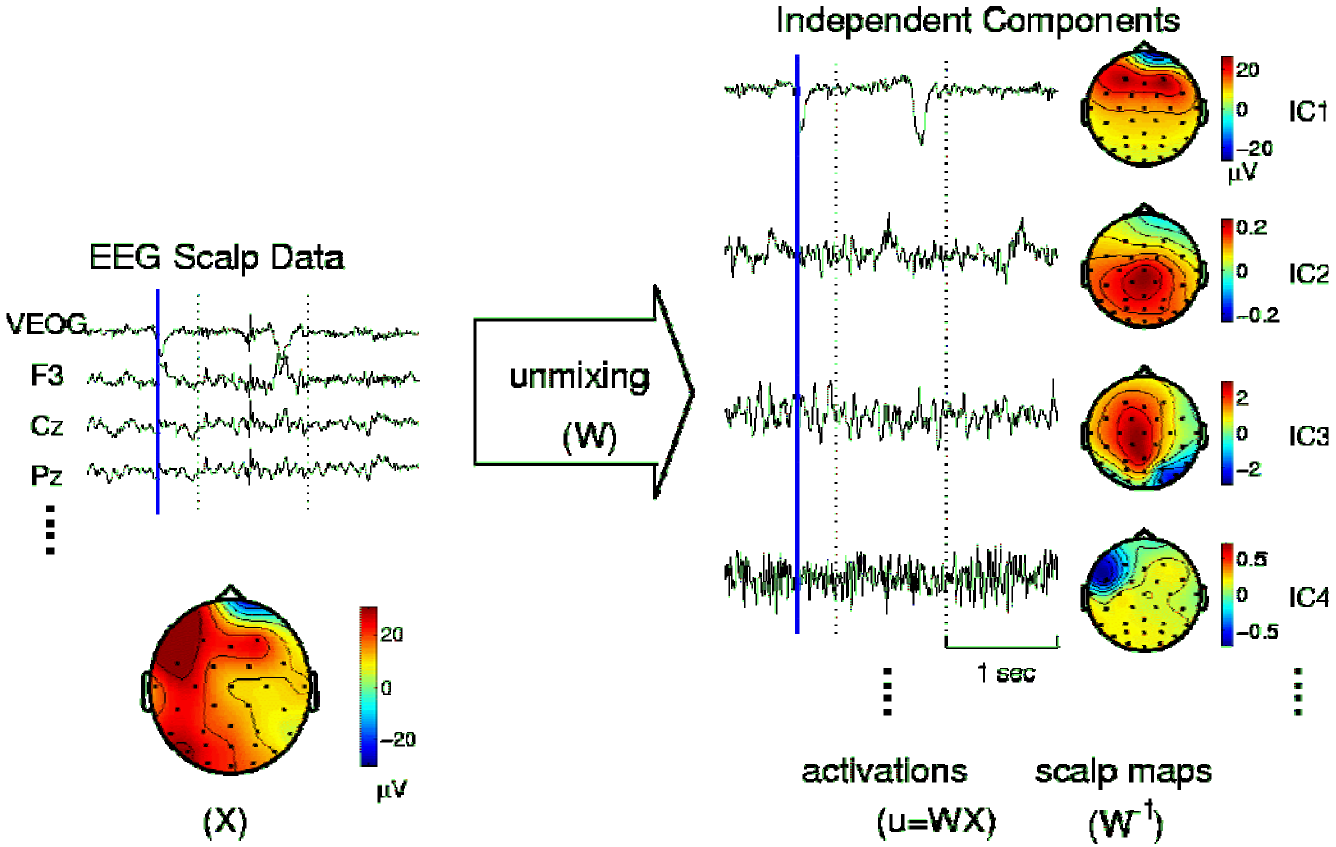
Characteristics of Independent Component of the EEG

- Concurrent Activity
- Maximally Temporally Independent
- Overlapping Maps and Spectra
- Dipolar Scalp Maps
- Functionally Independent
- Between-Subject Regularity

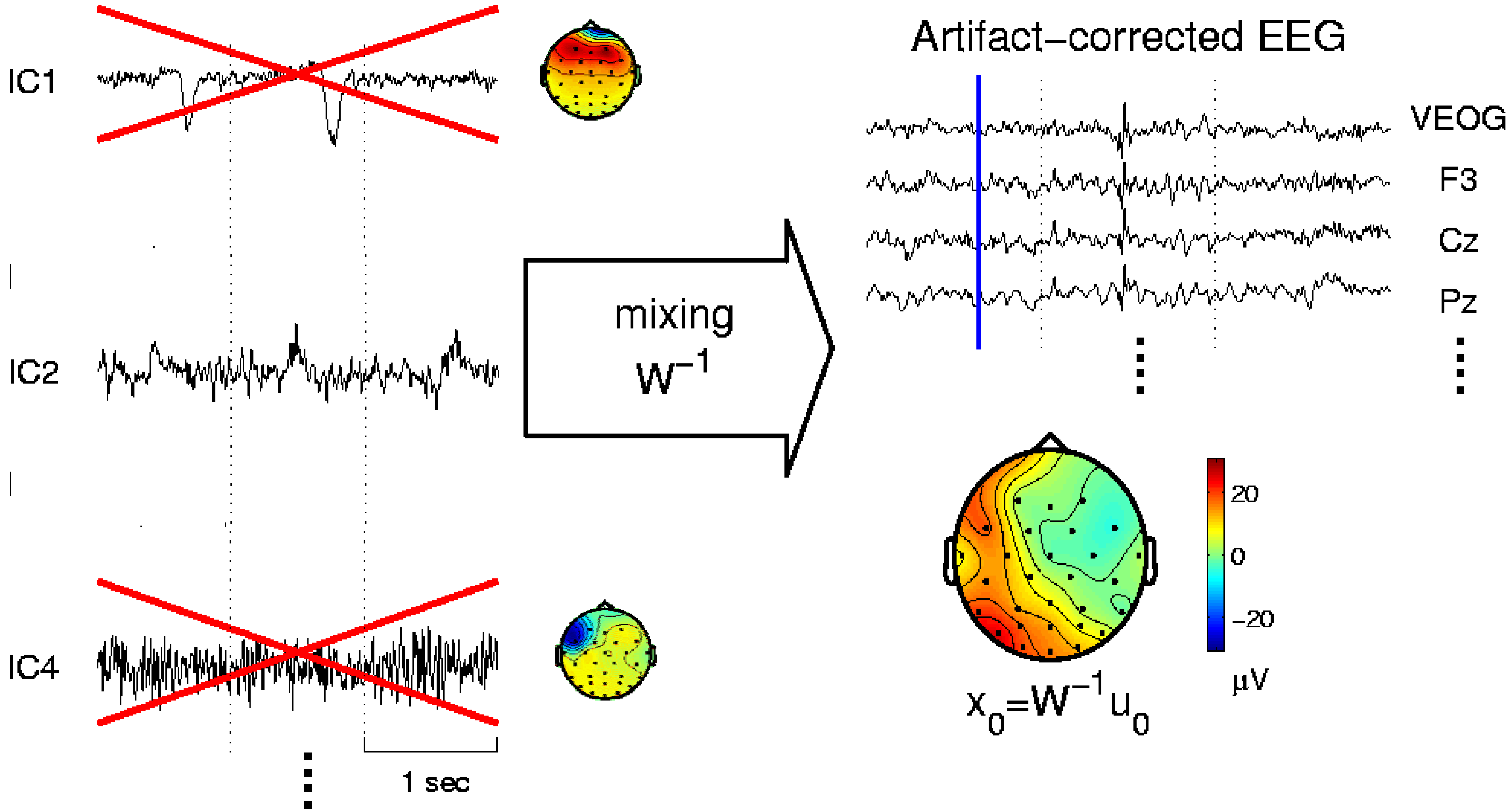
Largest 30 Independent Components
(single subject)



ICA Decomposition into Independent Components



Selective Projection onto Scalp Channels



$$X = W^{-1}U$$

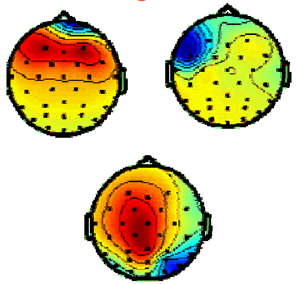
Data

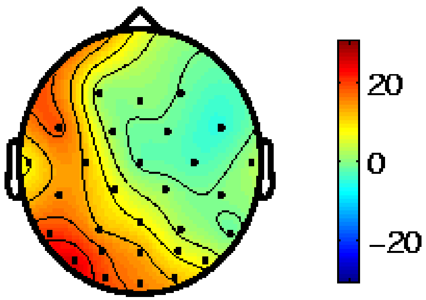
ICA activity U

$$
\begin{array}{l}
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & \dots \\
0 & -2 & -5 & -1 & 1 & -1 & \dots \\
0 & 0 & 0 & 0 & 0 & 0 & \dots \\
\dots & & & & & &
\end{bmatrix}
\end{array}
\begin{array}{l}
\leftarrow \text{Comp. 1} \\
\leftarrow \text{Comp. 2} \\
\leftarrow \text{Comp. 3} \\
\
\end{array}
$$

$$
\begin{bmatrix}
5 & 3 & -2 & \dots \\
1 & 2 & 4 & \dots \\
0 & -1 & 3 & \dots \\
\dots & & &
\end{bmatrix}
\rightarrow
\begin{bmatrix}
3*5+0*3-1*(-2) & 2*5+(-2)*3+2*(-2) & \dots \\
3*1+0*2-1*4 & 2*1+(-2)*2+2*4 & \dots \\
5*1-5*2+0*4 & 5*1-5*2+0*4 & \dots \\
\dots & \dots &
\end{bmatrix}
\begin{array}{l}
\leftarrow \text{Chan 1} \\
\leftarrow \text{Chan 2} \\
\leftarrow \text{Chan 3} \\
\
\end{array}
$$

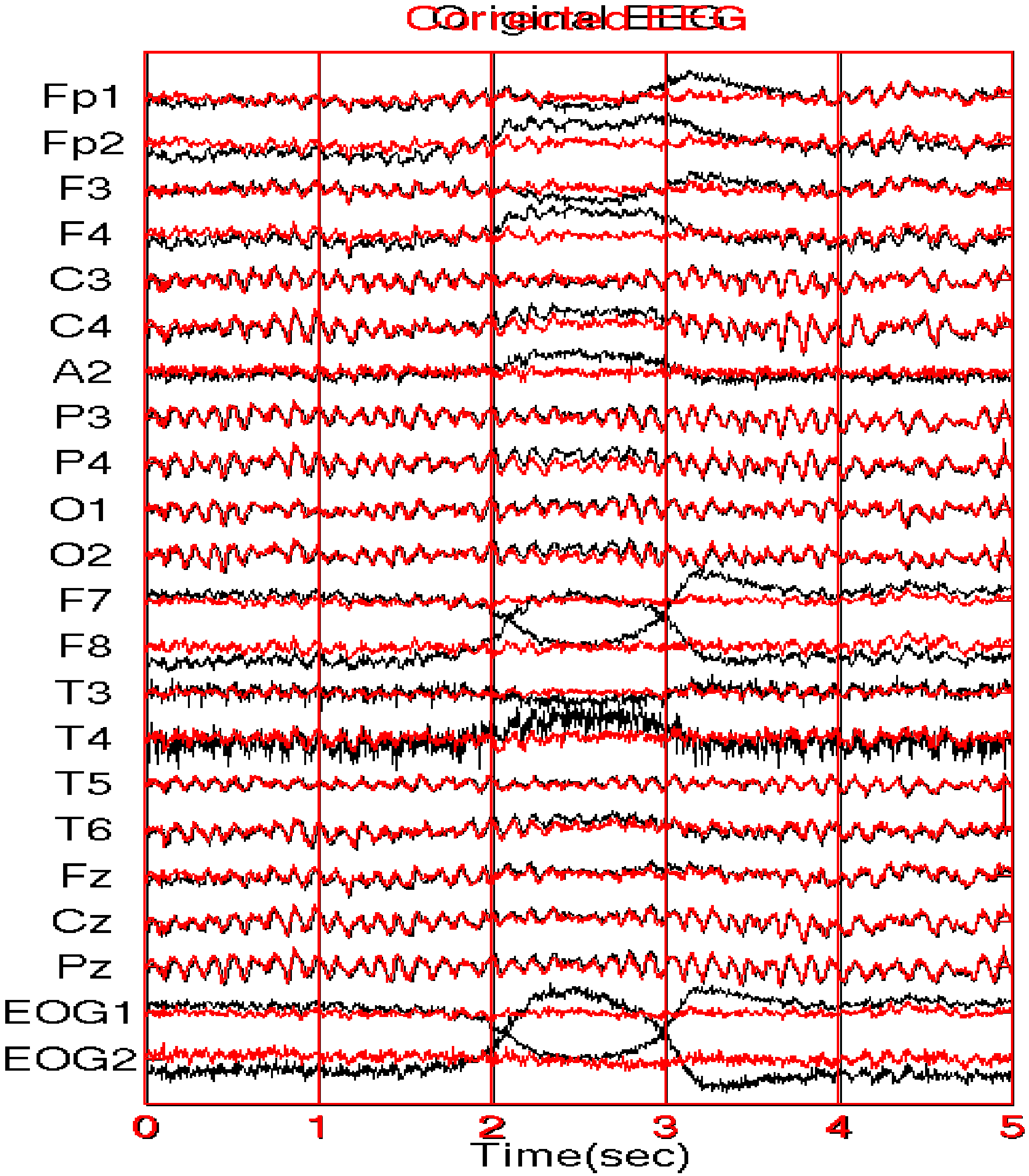
Data X



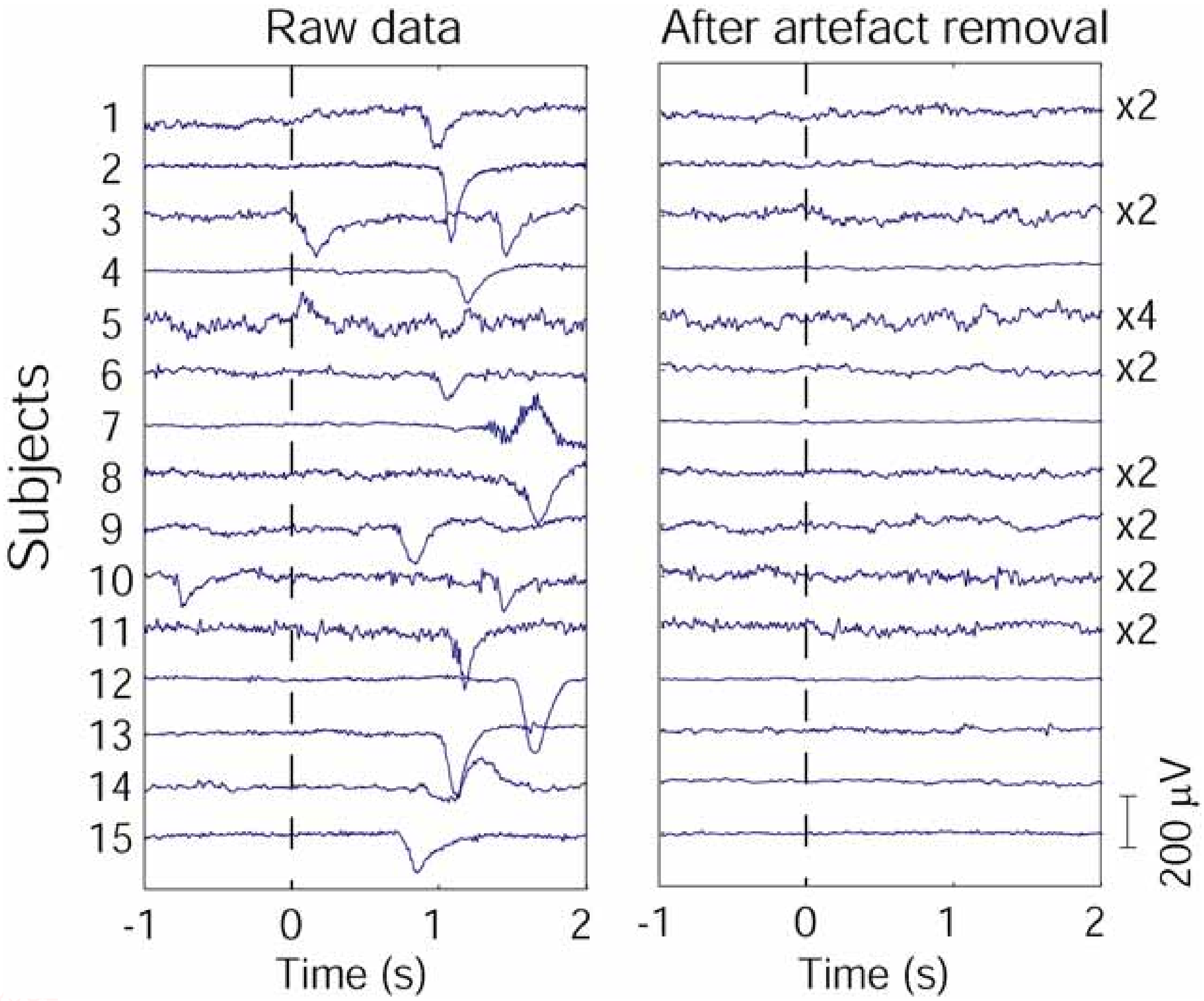


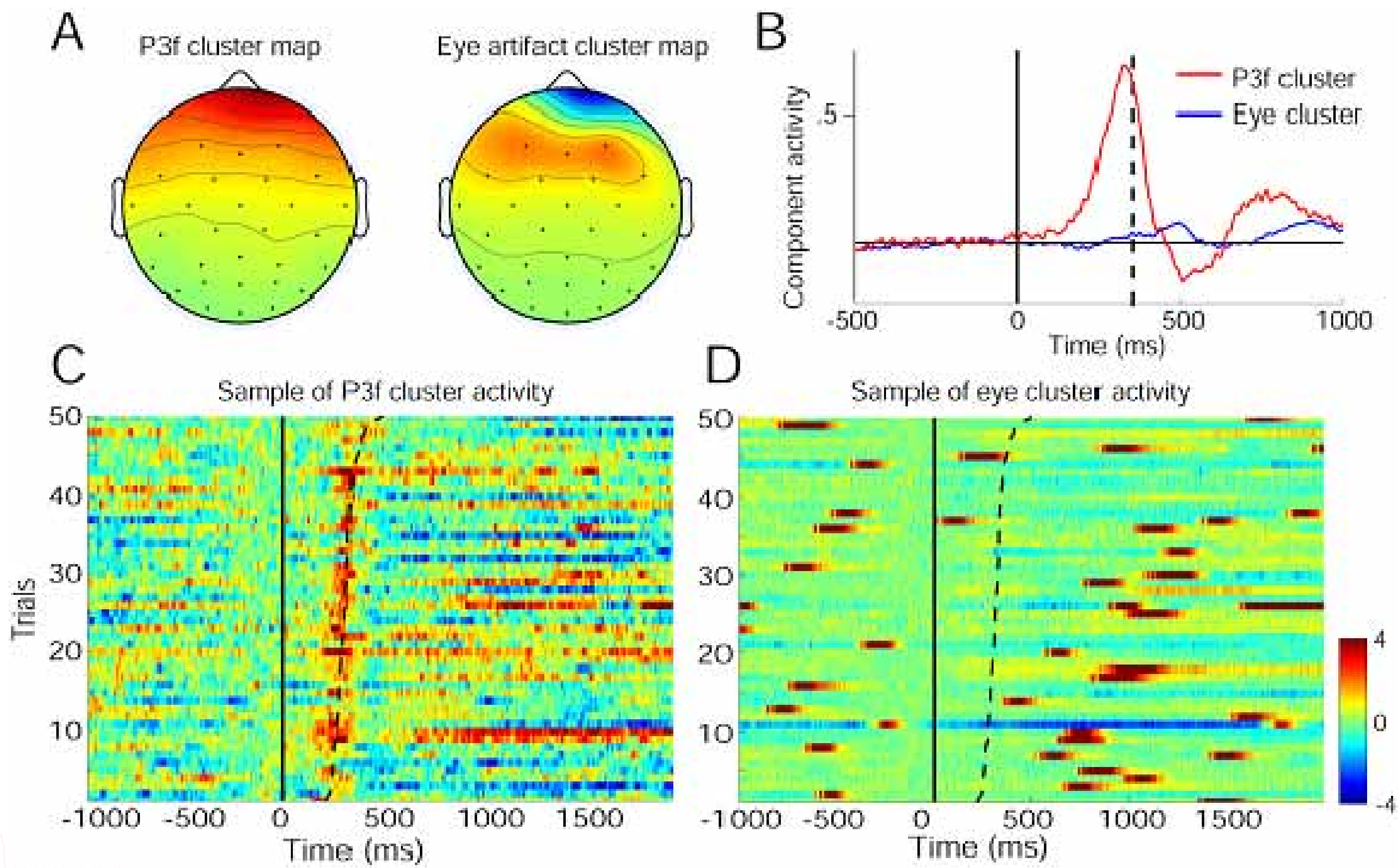
Inverse weight matrix W^{-1}

ICA-based Artifact Removal



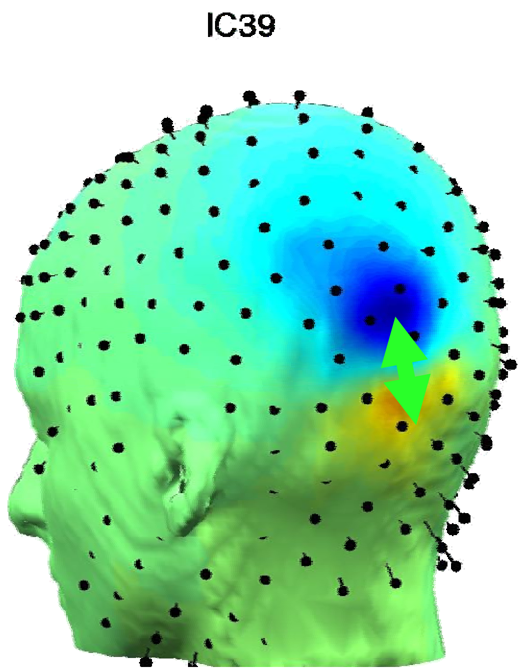
Artifact removal using ICA

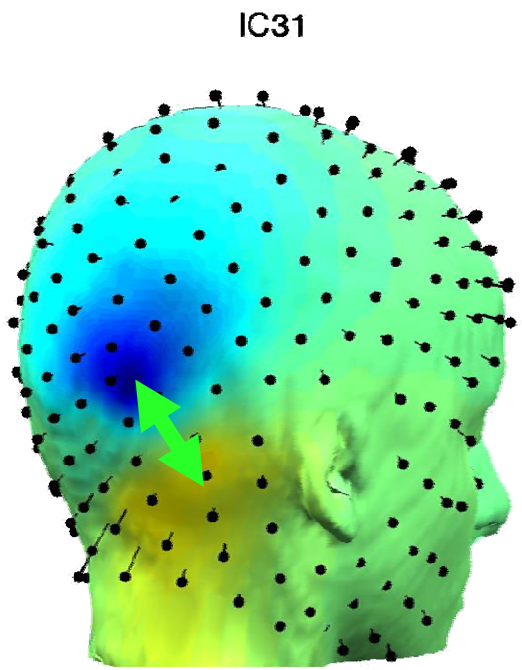


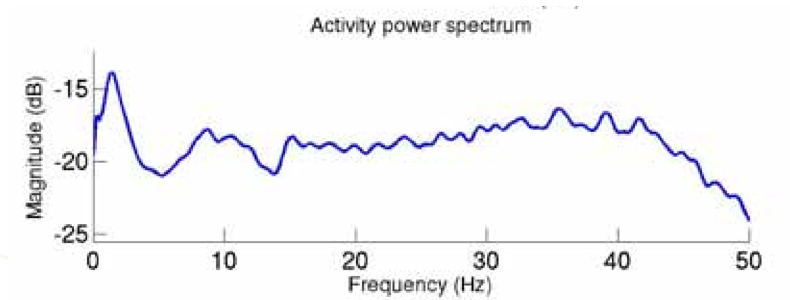


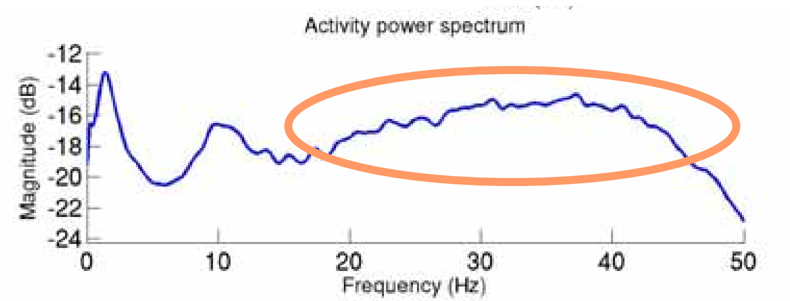
A
P3f cluster map
Eye artifact cluster map
B
Component activity
.5
P3f cluster
Eye cluster
-500
0
500
1000
Time (ms)
C
Sample of P3f cluster activity
Trials
50
40
30
20
10
-1000
-500
0
500
1000
1500
Time (ms)
D
Sample of eye cluster activity
4
0
-4

Two Neck Muscle Processes

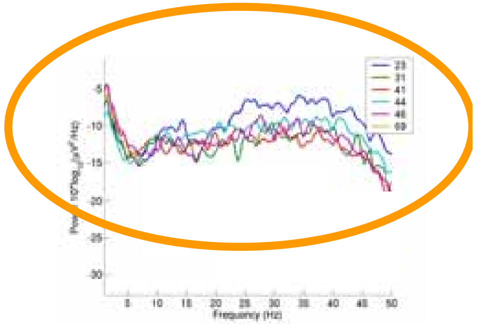


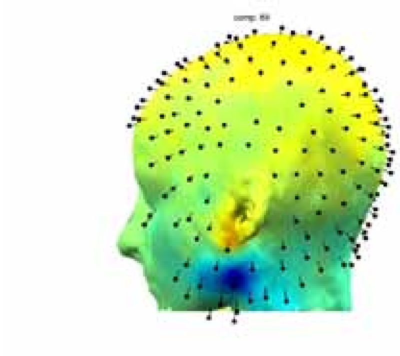


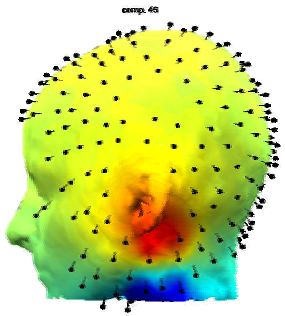


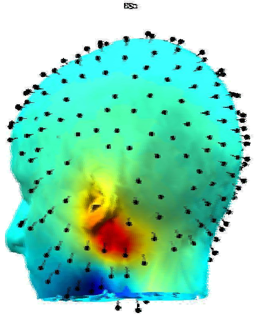


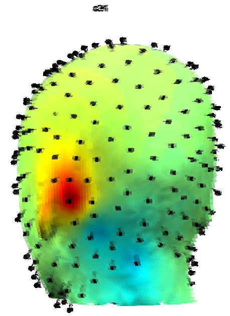
Some Independent EMG Components

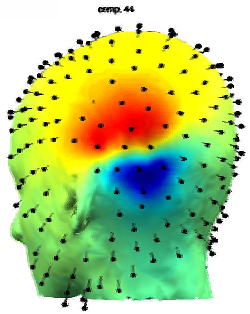


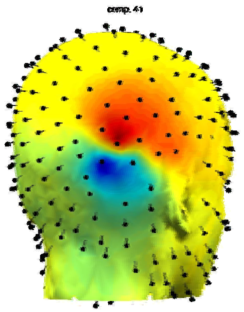


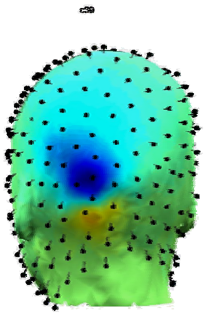


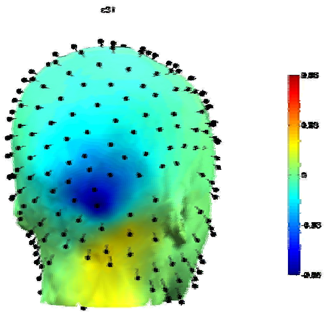


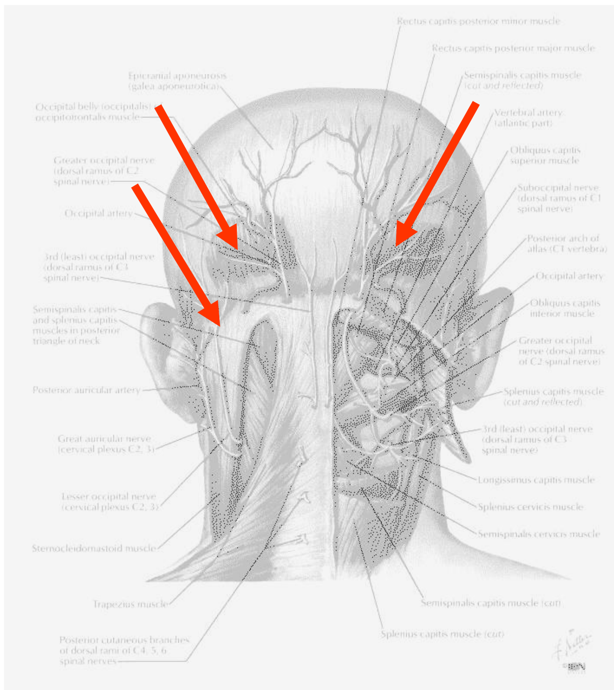


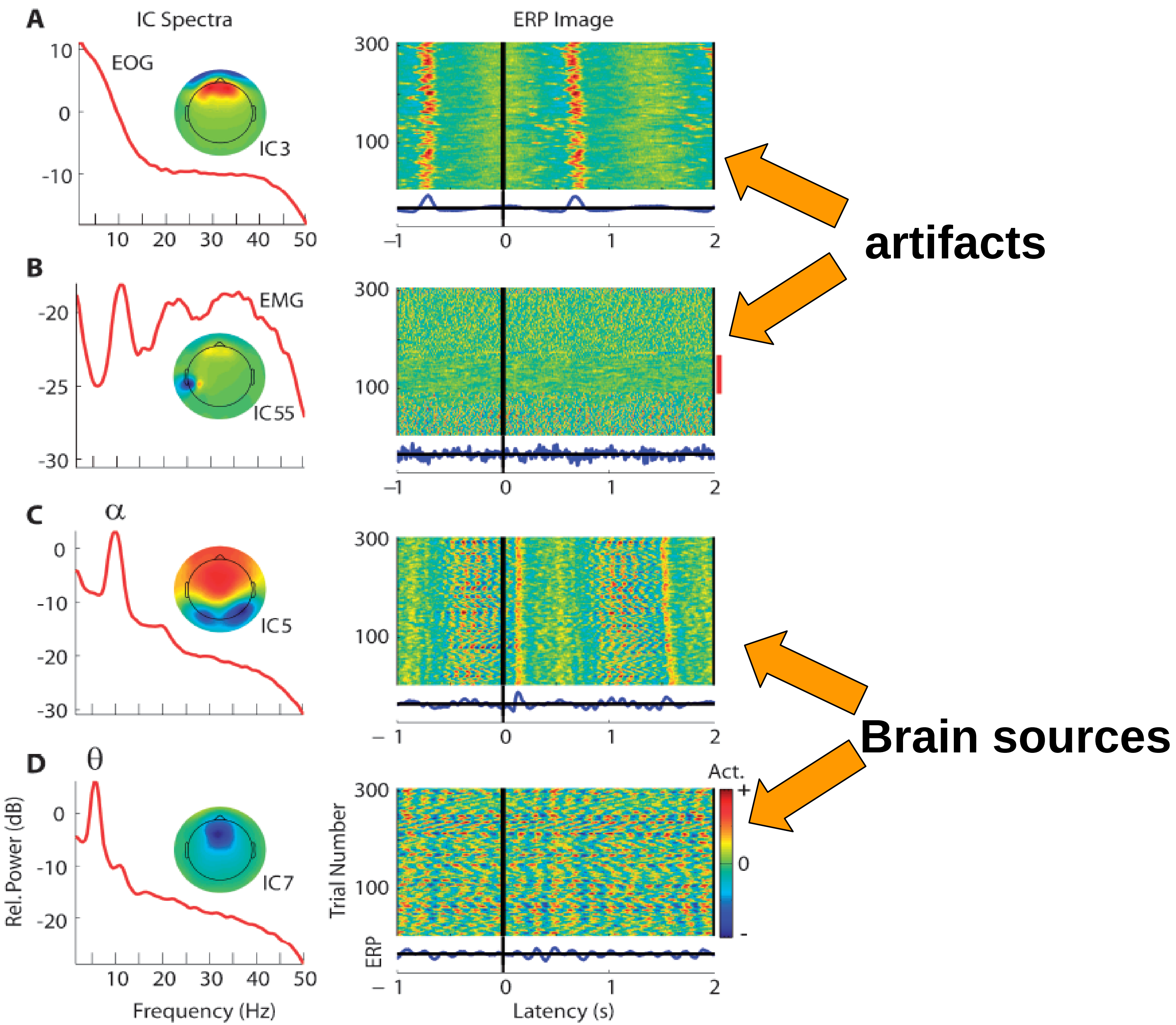






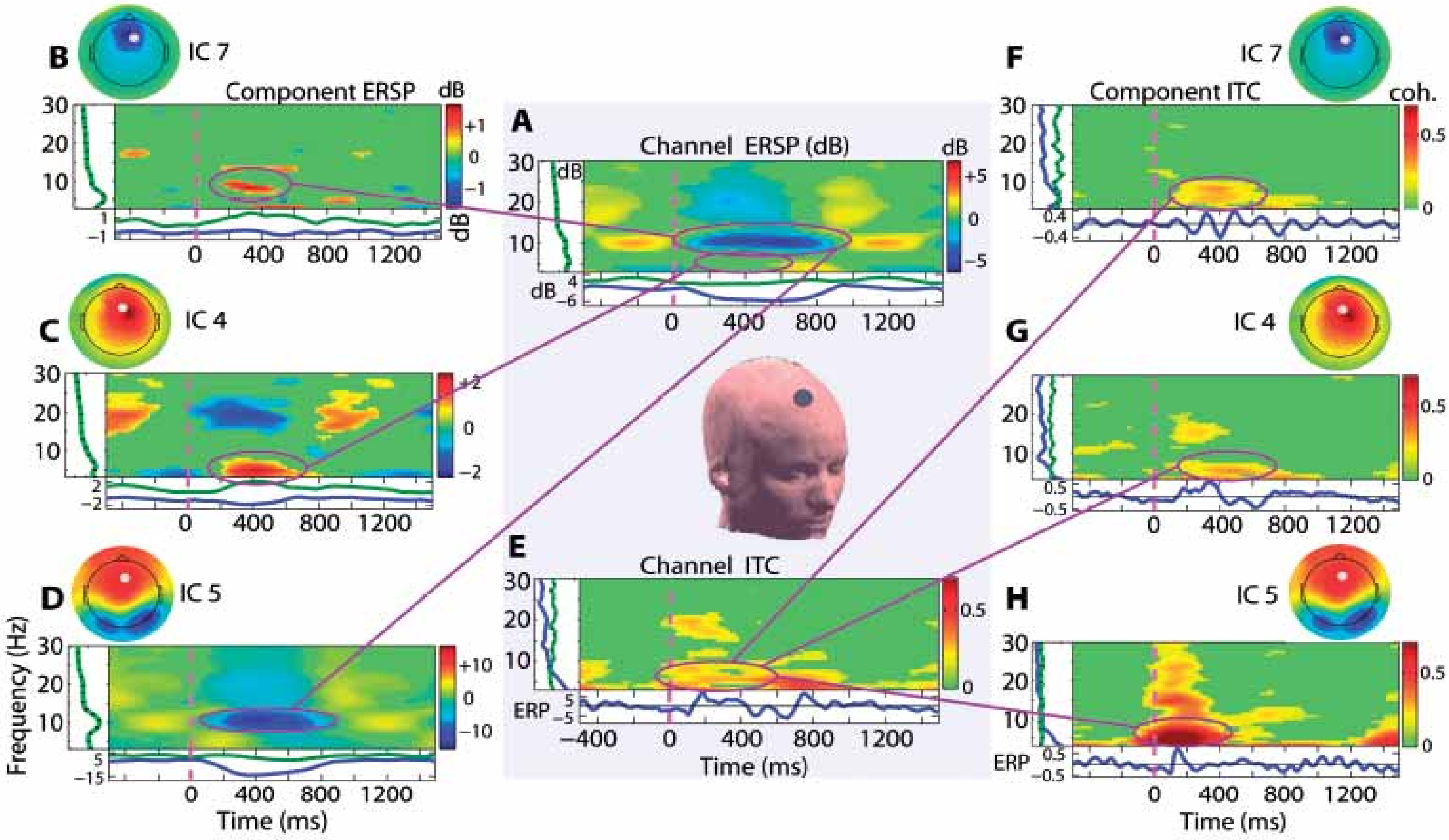






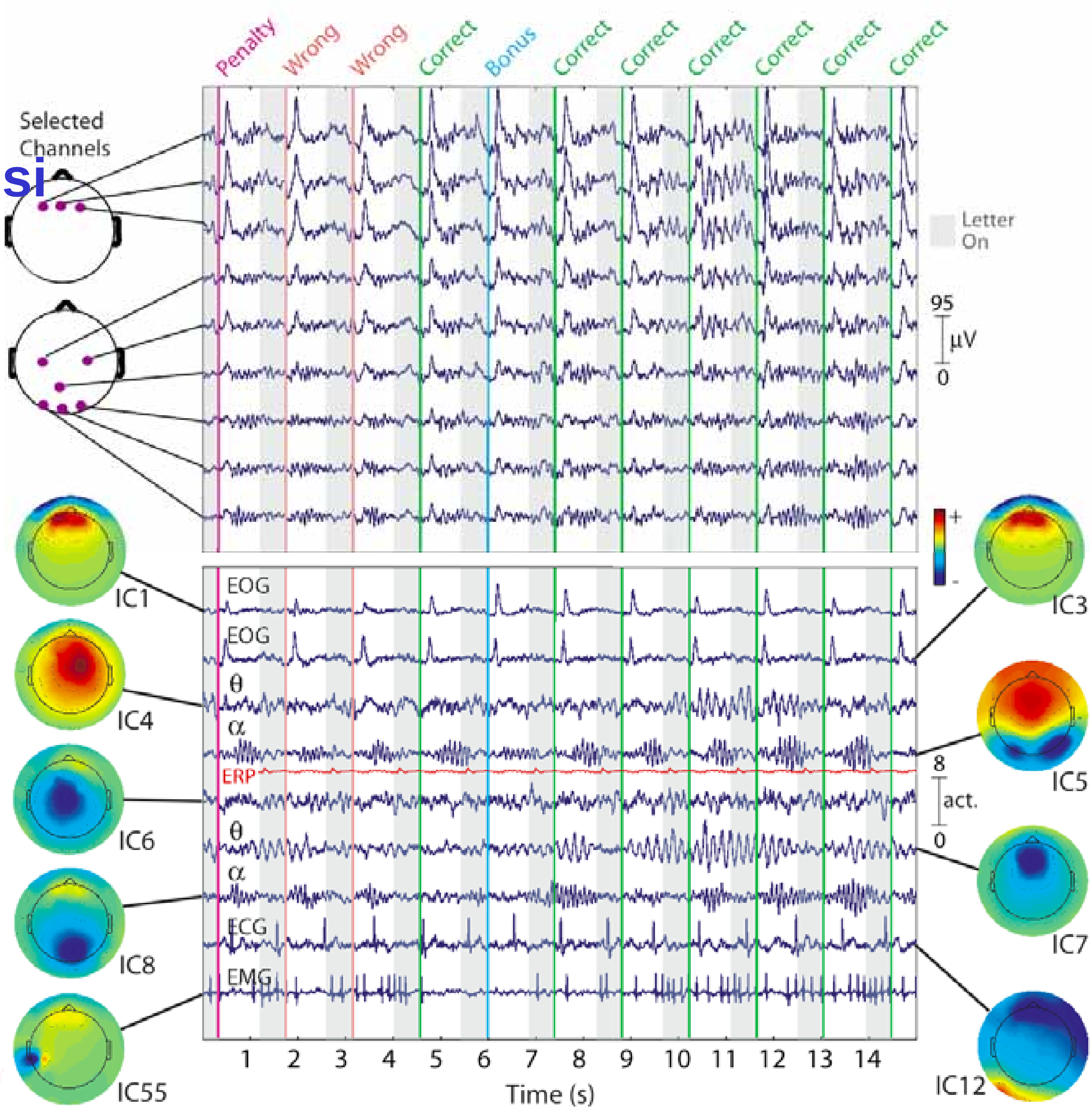
Credit: J. Onto

Why analyze source activity instead of channels?

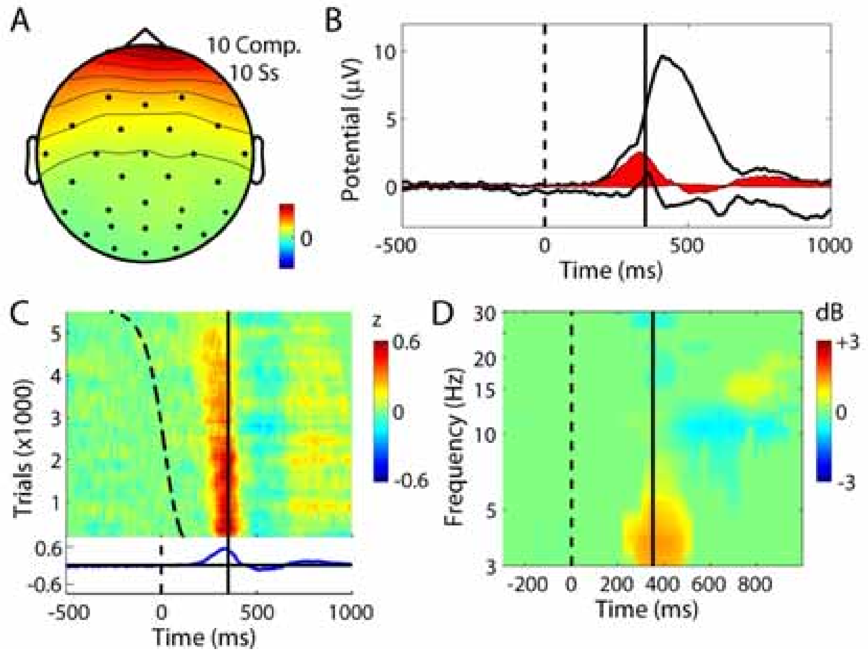


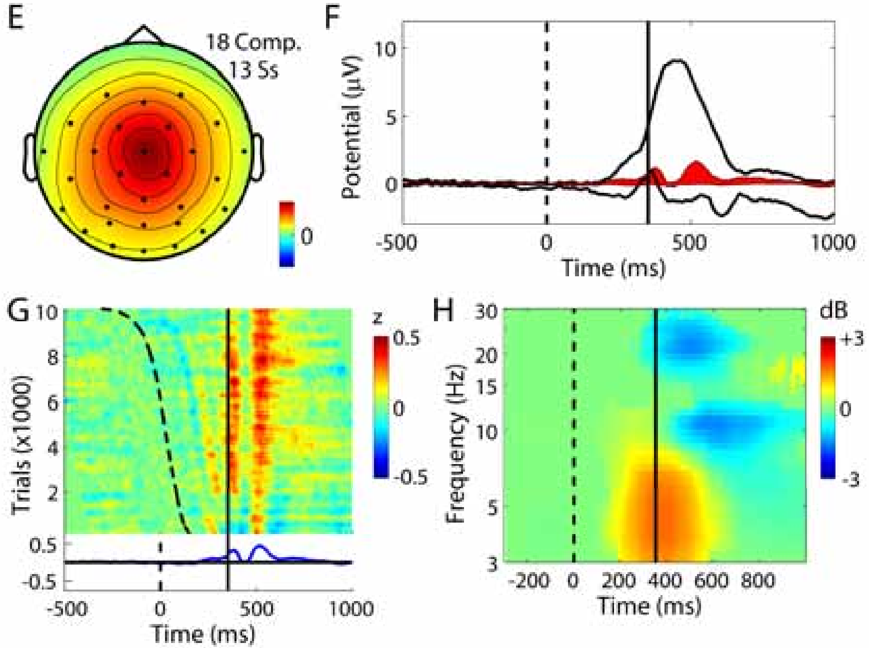
Credit: J. Onto

Sample EEG Decomposition

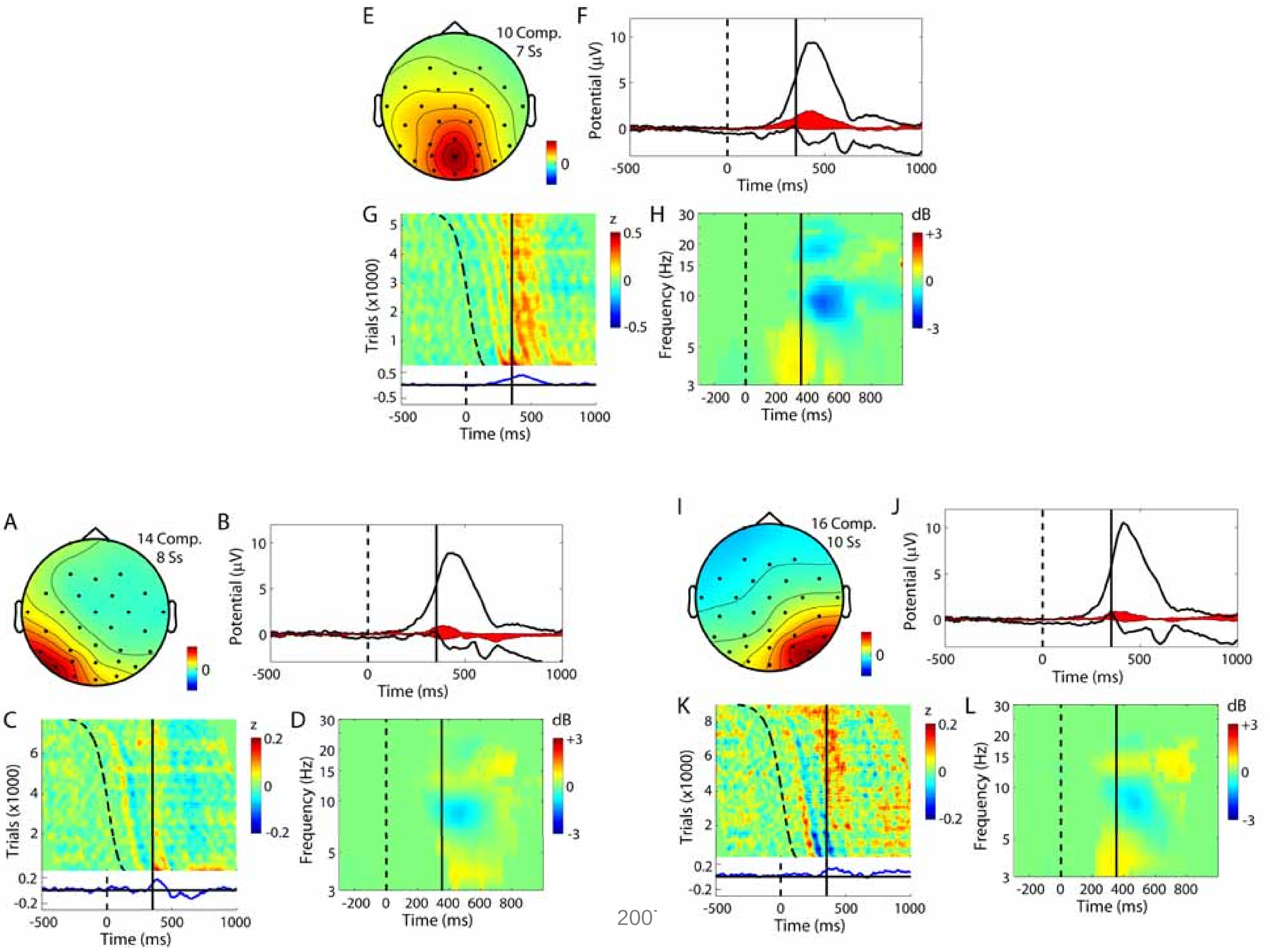


Credit: J. Onto


A
10 Comp.
10 Ss
0
B
Potential (μV)
10
5
0
-500
0
500
1000
Time (ms)
C
Trials (x1000)
5
4
3
2
1
z
0.6
0
-0.6
0.6
-0.6
-500
0
500
1000
Time (ms)
D
Frequency (Hz)
30
20
15
10
5
3
-200
0
200
400
600
800
Time (ms)
dB
+3
0
-3

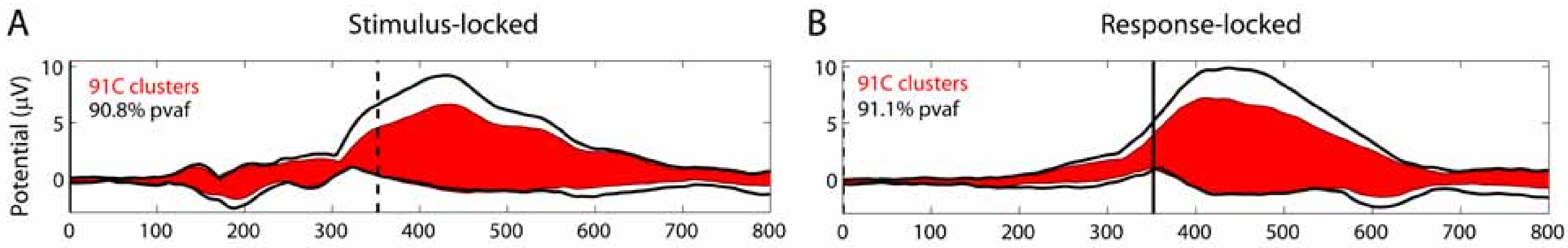


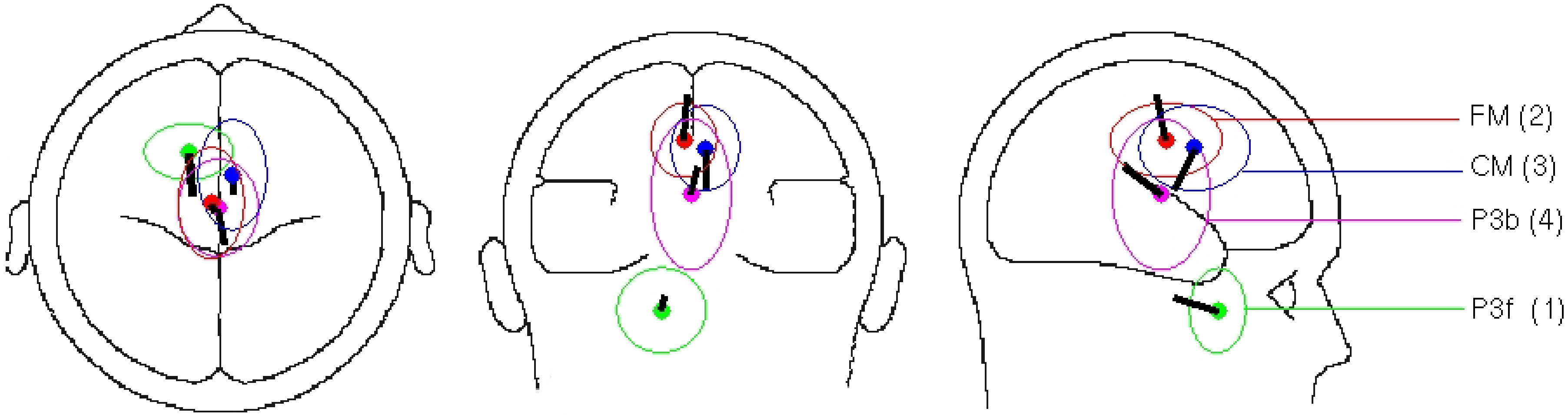
E
18 Comp.
13 Ss
0
F
Potential (μV)
10
5
0
-500
0
500
1000
Time (ms)
G
Trials (x1000)
10
8
6
4
2
0.5
-0.5
z
0.5
0
-0.5
-500
0
500
1000
Time (ms)
H
Frequency (Hz)
30
20
15
10
5
3
dB
+3
0
-3
-200
0
200
400
600
800
Time (ms)

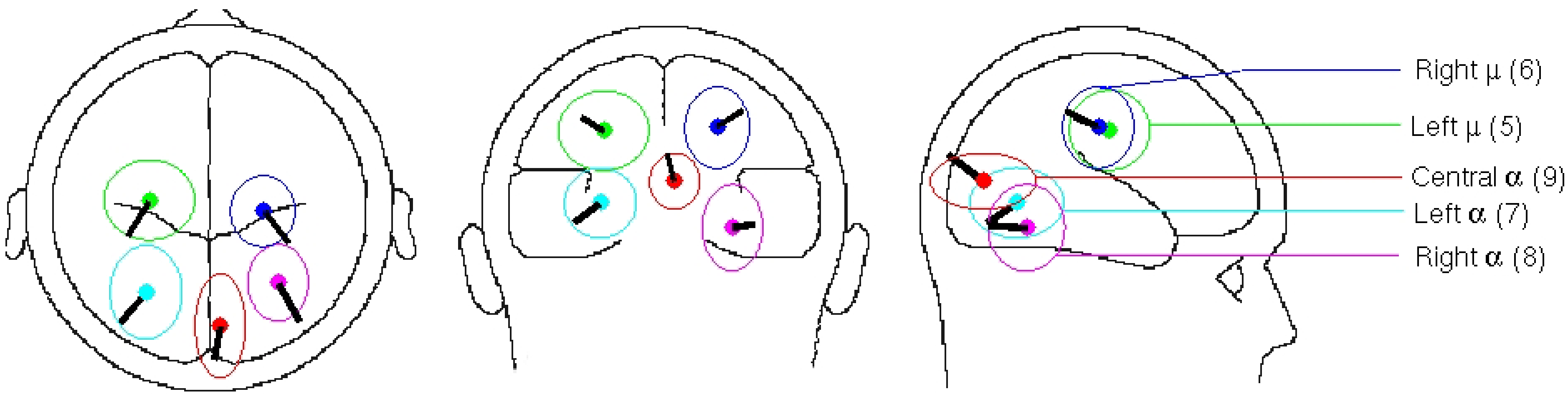


E
10 Comp.
7 Ss
0
F
Potential (μV)
Time (ms)
G
Trials (x1000)
z
0.5
-0.5
Time (ms)
H
Frequency (Hz)
dB
+3
-3
Time (ms)
A
14 Comp.
8 Ss
B
Potential (μV)
Time (ms)
C
Trials (x1000)
z
0.2
-0.2
Time (ms)
D
Frequency (Hz)
dB
Time (ms)
I
16 Comp.
10 Ss
J
Potential (μV)
Time (ms)
K
Trials (x1000)
Time (ms)
L
Frequency (Hz)
dB
Time (ms)

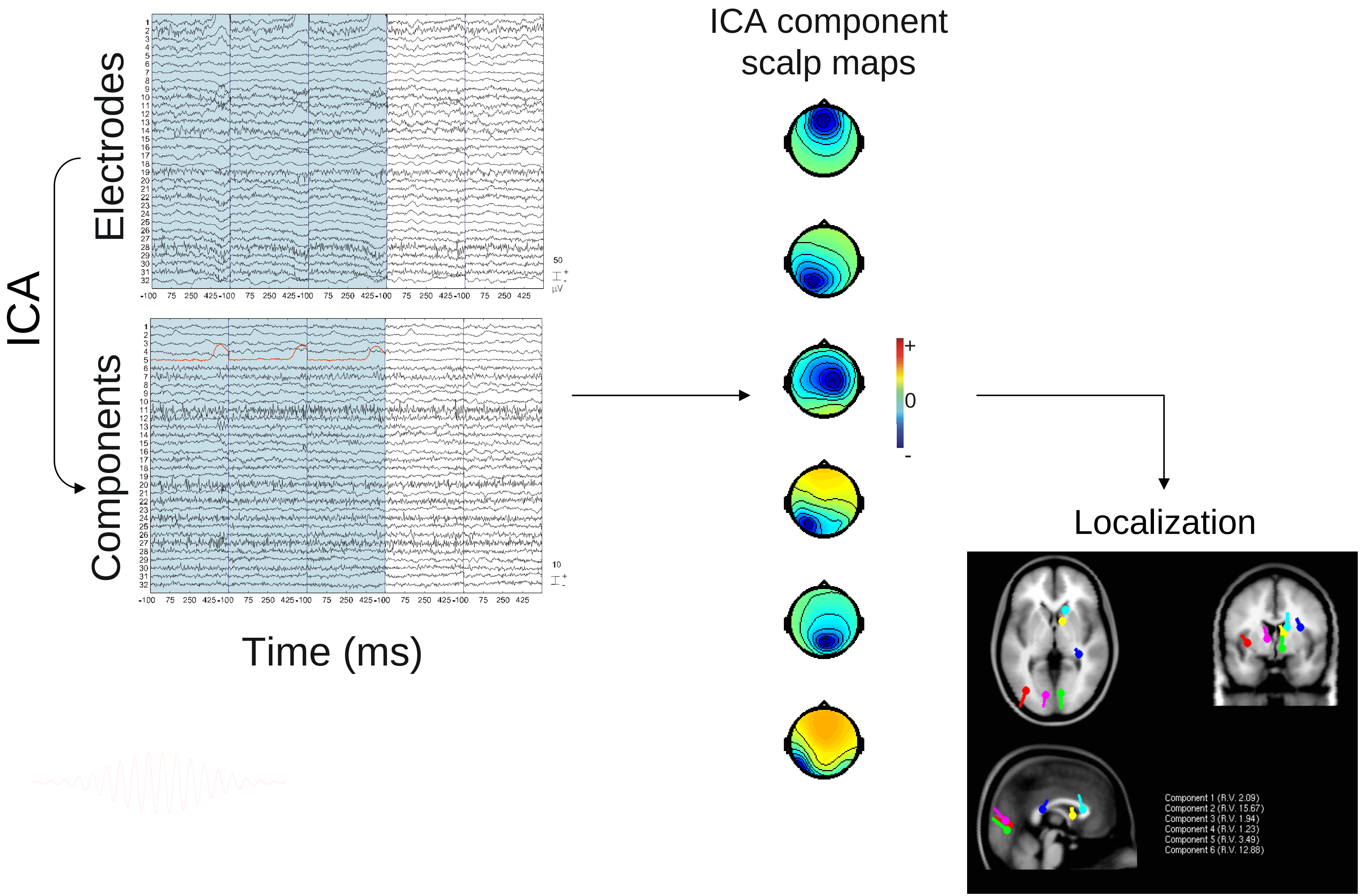
200


A
Stimulus-locked
91C clusters
90.8% pvaf
Potential (μV)
10
5
0
0
100
200
300
400
500
600
700
800
B
Response-locked
91C clusters
91.1% pvaf
10
5
0
0
100
200
300
400
500
600
700
800

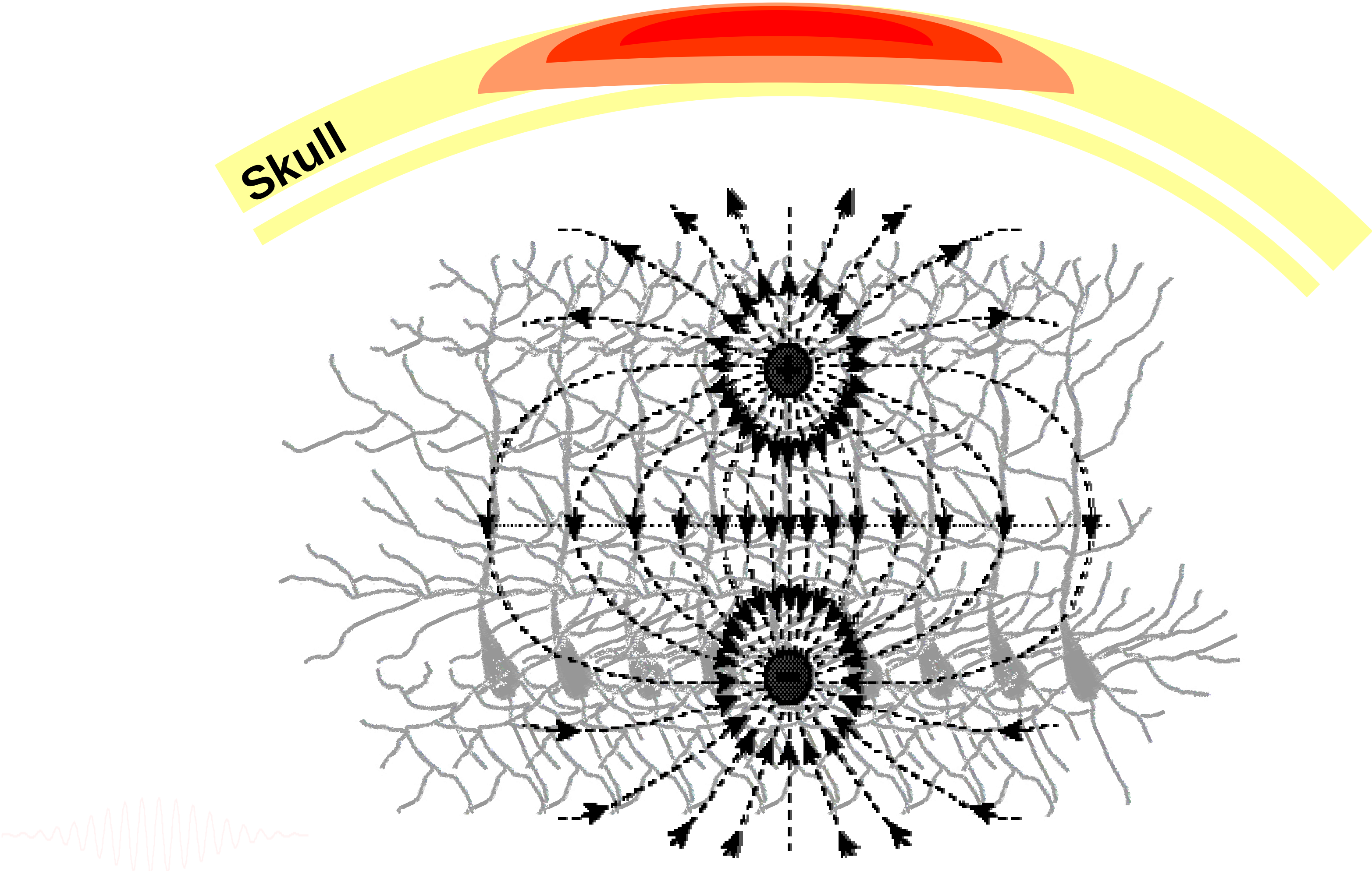

FM (2)
CM (3)
P3b (4)
P3f (1)


Right μ (6)
Left μ (5)
Central α (9)
Left α (7)
Right α (8)

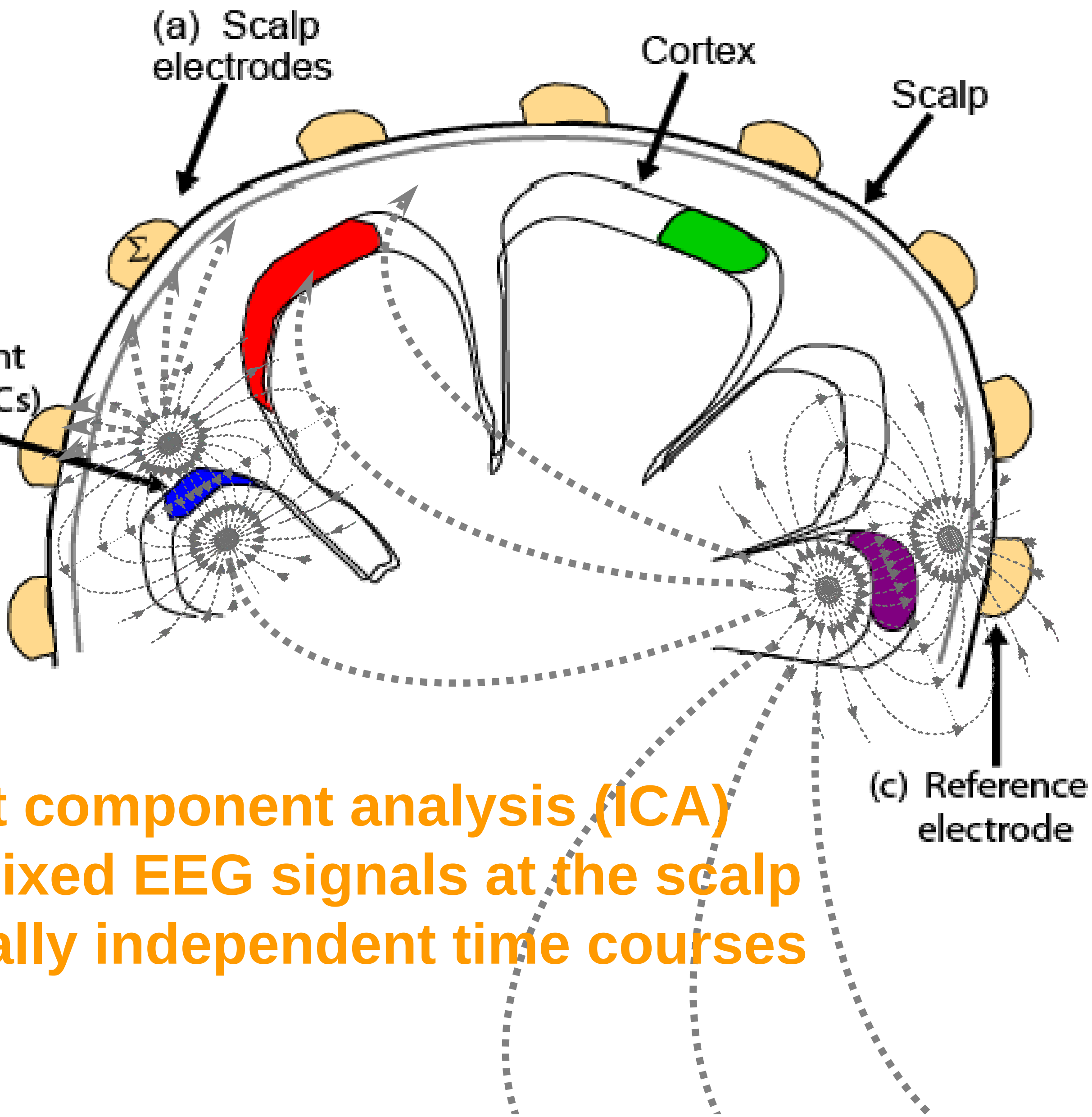
Localization



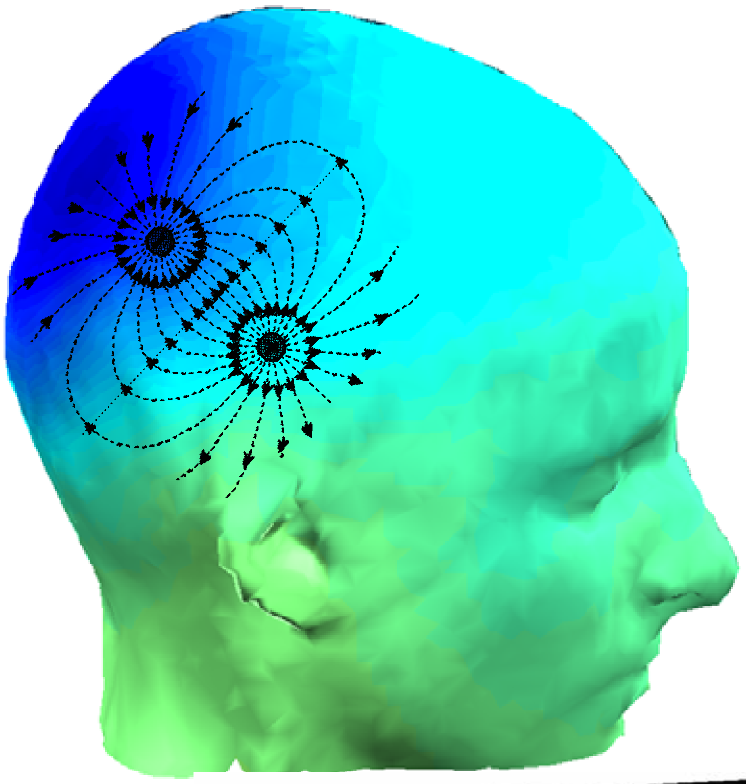
Patch of Cortex Acting as a Dipole



Separating EEG source activities

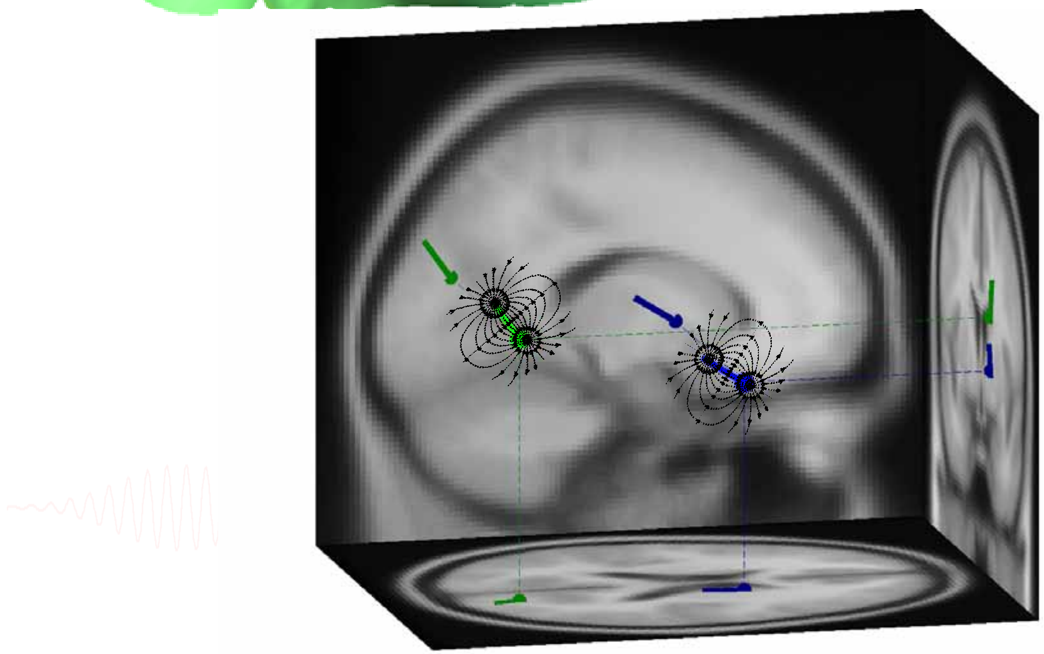


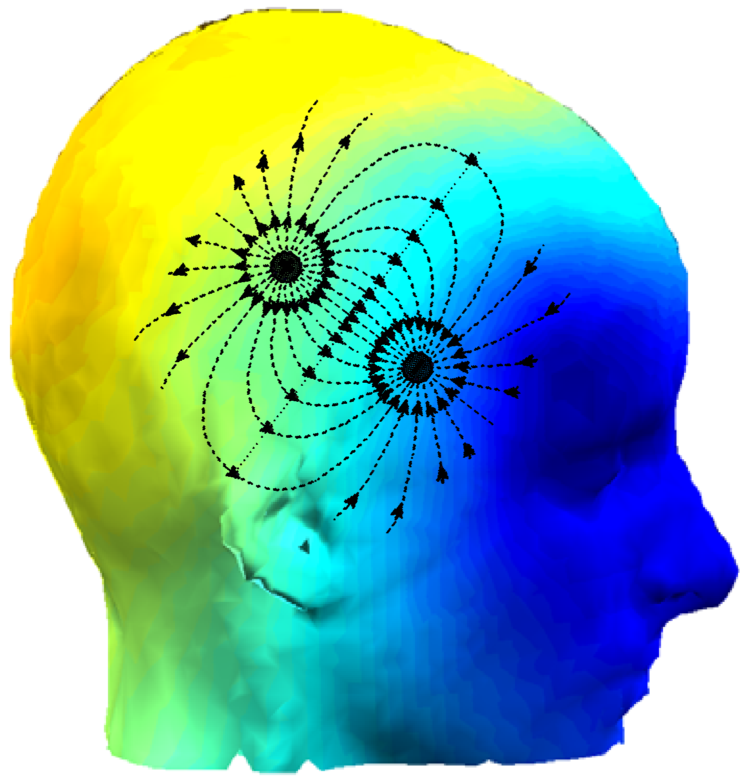
Independent component analysis (ICA) separates mixed EEG signals at the scalp into temporally independent time courses

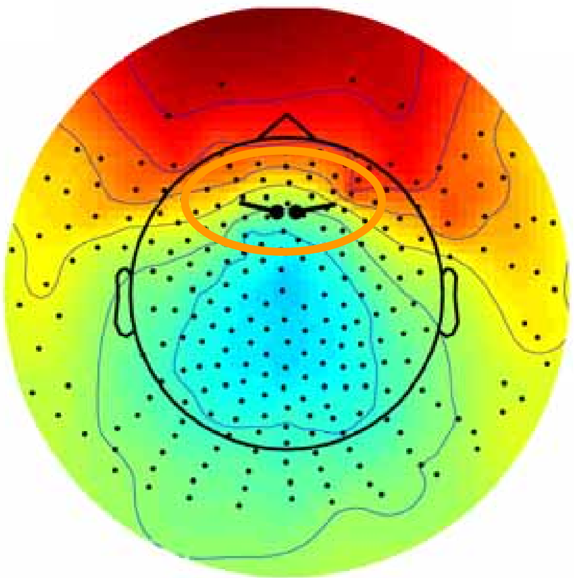


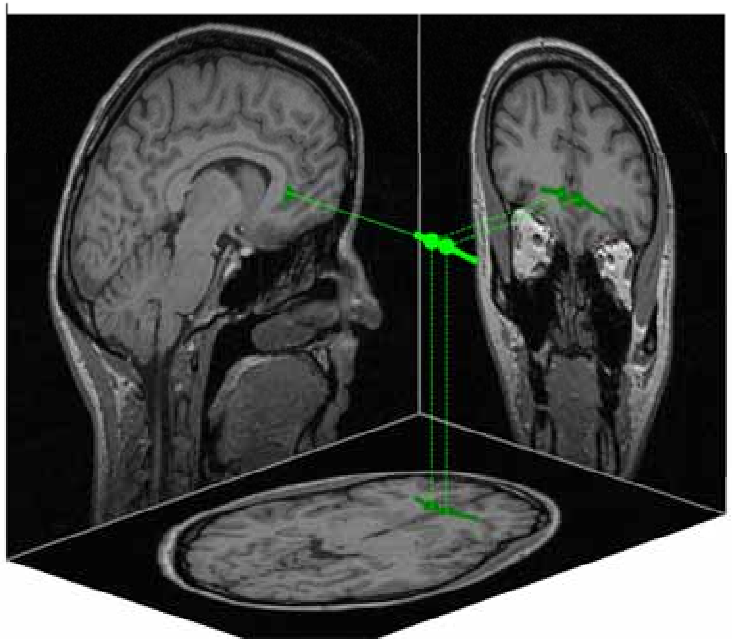
Dipolar Scalp Projections

ICA creates a spatial filter for each temporally independent source

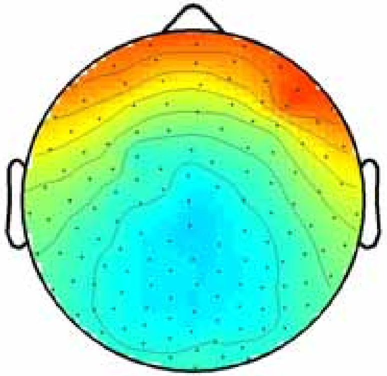




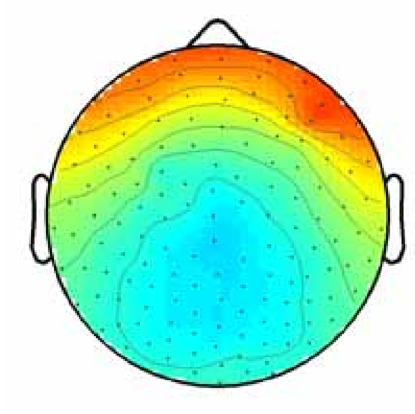


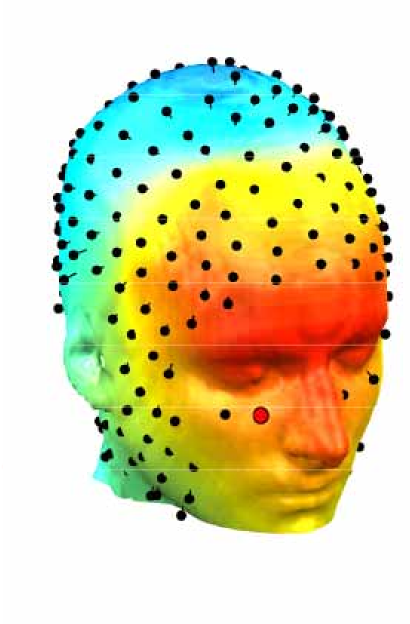


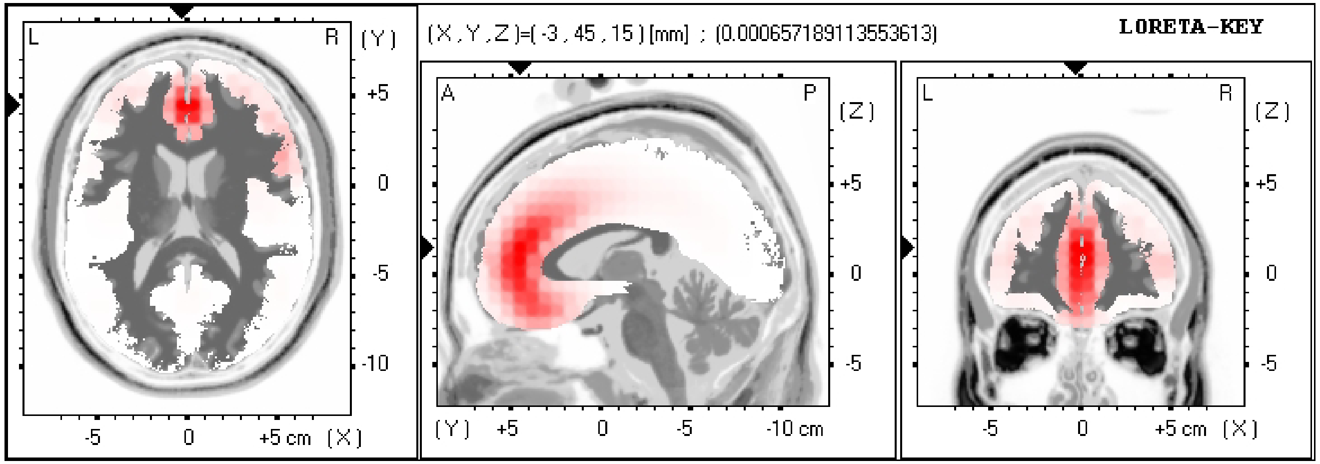
Localization of activity

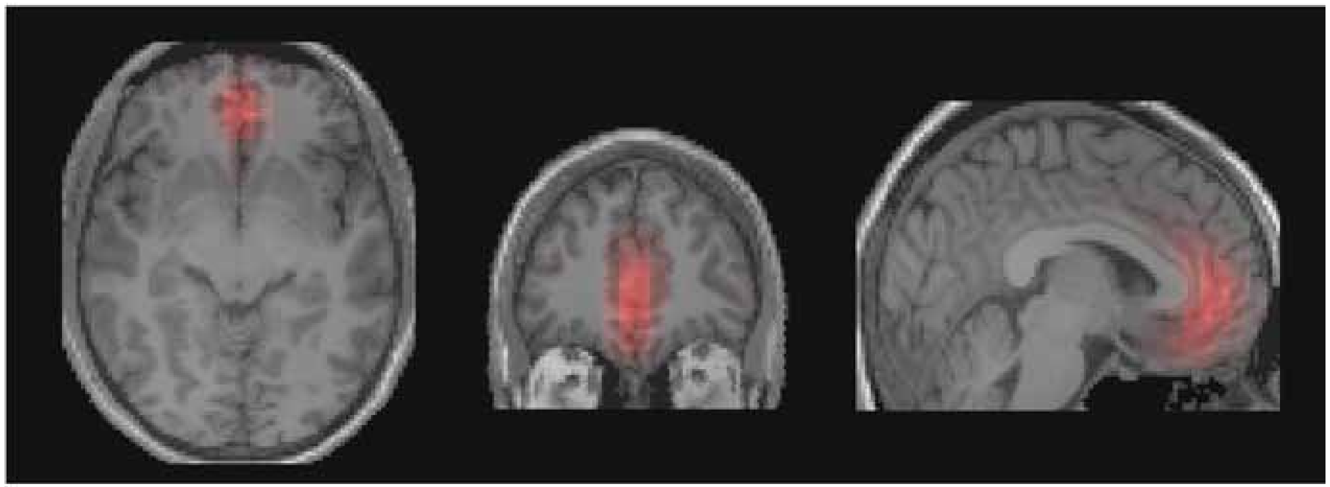


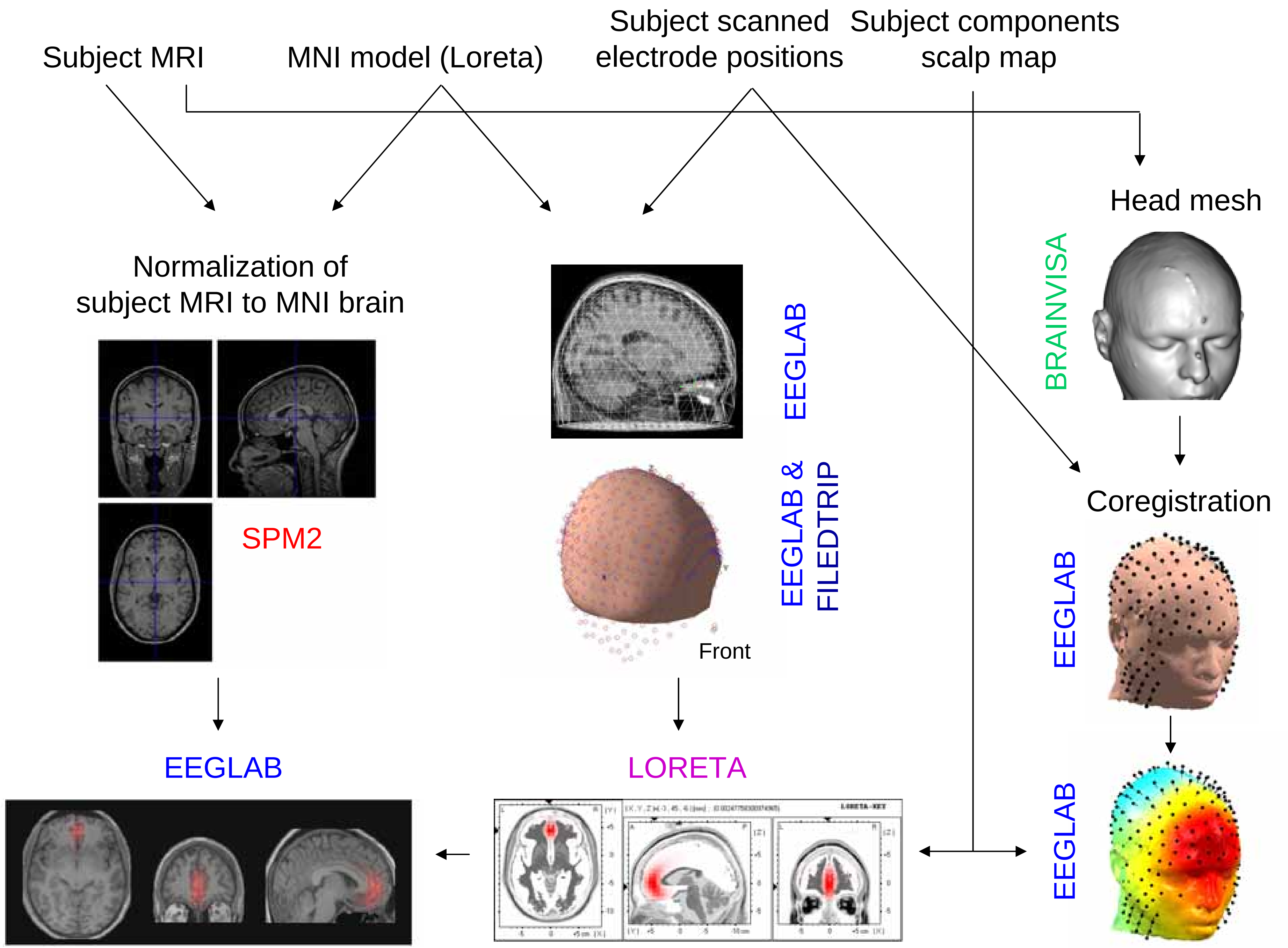
Localization of activity










Subject MRI
MNI model (Loreta)
Subject scanned electrode positions
Subject components scalp map
Head mesh
BRAINVISA
Normalization of subject MRI to MNI brain
EEGLAB
EEGLAB & FILEDTRIP
Front
SPM2
Coregistration
EEGLAB
EEGLAB
LORETA
EEGLAB

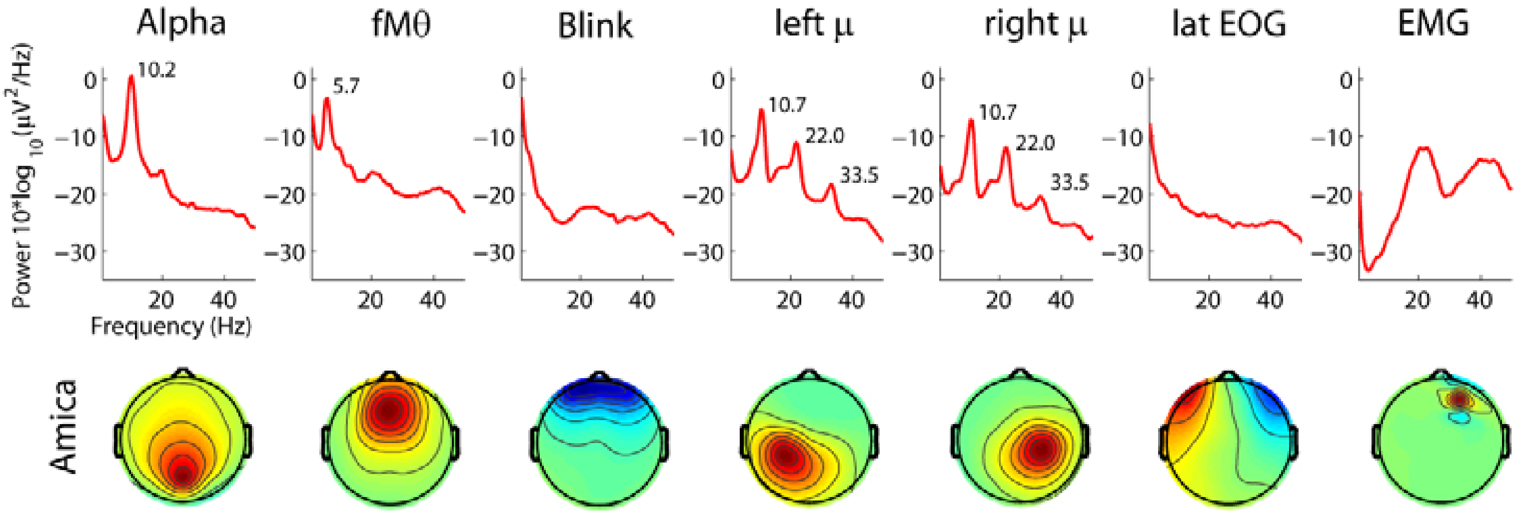
ICA component comparison

Data
- 13 subjects performing a memory task
- 71 electrodes including EOGs
- more than 300,000 data points/subject

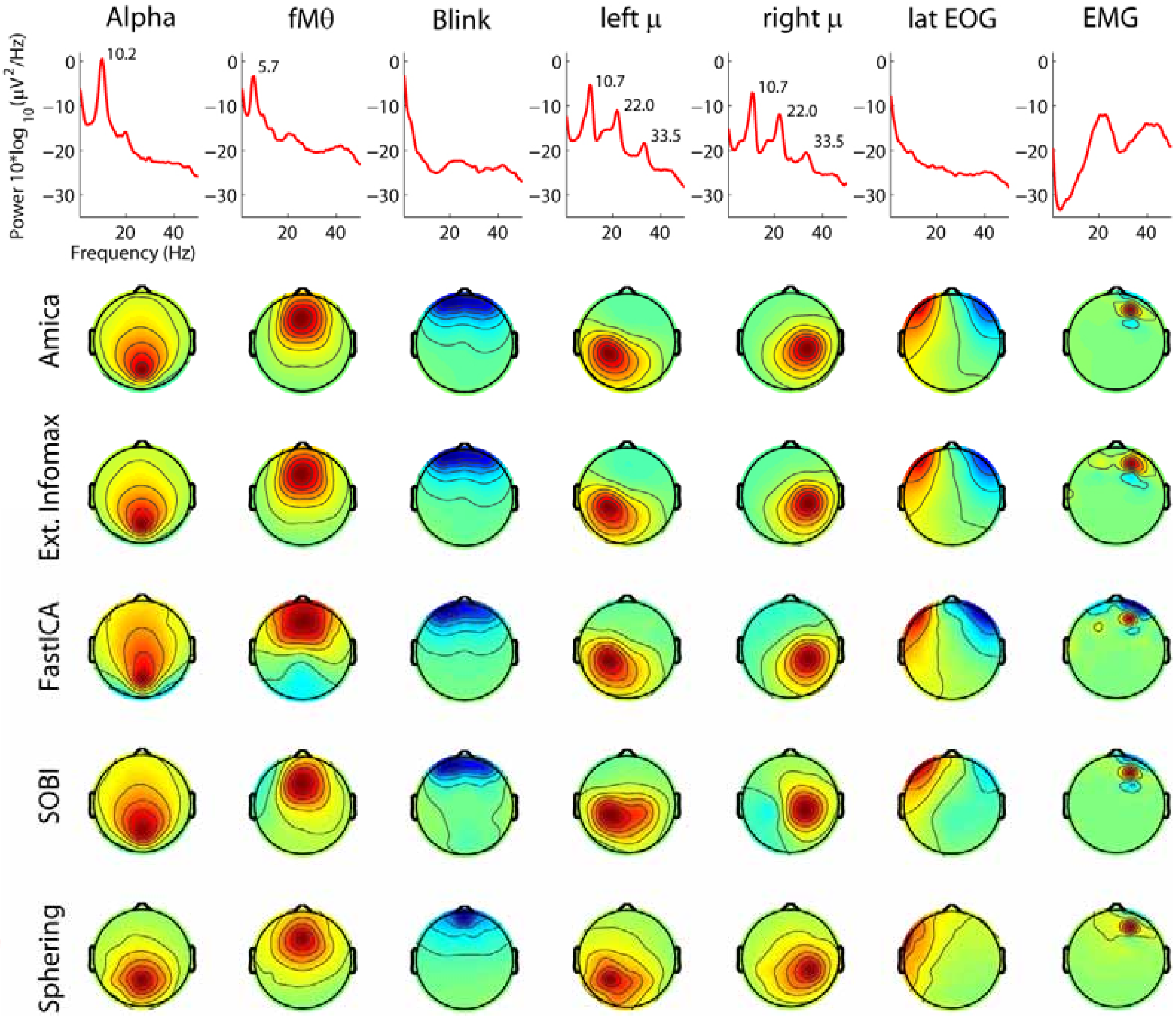
Decomposition
- 23 ICA algorithms plus PCA and Promax

Analysis
- Localization of all components with a single dipole (4-shell spherical model)

Algorithm (Matlab func.).	D%	MIR	Origin
Extended Infomax (runica).	29.9	178	EEGLAB 4.515
Pearson	29.1	169	ICAcentral (6).
Infomax (runica).	28.2	160	EEGLAB 4.515
ERICA	26.9	184	ICALAB 1.5.2
SONS	25.4	183	ICALAB 1.5.2
SHIBBS	23.7	169	ICAcentral (5).
FastICA*	23.5	169	ICAcentral (2).
JADE (jader).	23.4	169	EEGLAB 4.515
TICA	23.4	169	ICALAB 1.5.2
JADE optimized (jade_op).	21.4	169	ICALAB 1.5.2
JADE w/ time delay (jade_td).	20.2	169	ICALAB 1.5.2
eeA	19.0	305	ICAcentral (8).
Infomax (icaML) †	18.8	212	ICA DTU Tbox
FOBI	18.6	169	ICALAB 1.5.2
SOBIRO (acsobiro).	17.9	167	EEGLAB 4.515
EVD 24	17.7	169	ICALAB 1.5.2
EVD	17.0	169	ICALAB 1.5.2
SOBI	16.1	583	EEGLAB 4.515
icaMS†	10.6	169	ICA DTU Tbox
AMUSE	8.5	169	ICALAB 1.5.2
PCA	3.1	583	EEGLAB 4.515
Promax	33.7	467	EEGLAB 4.515
Whitening/Sphering	57.6	164	EEGLAB 4.515

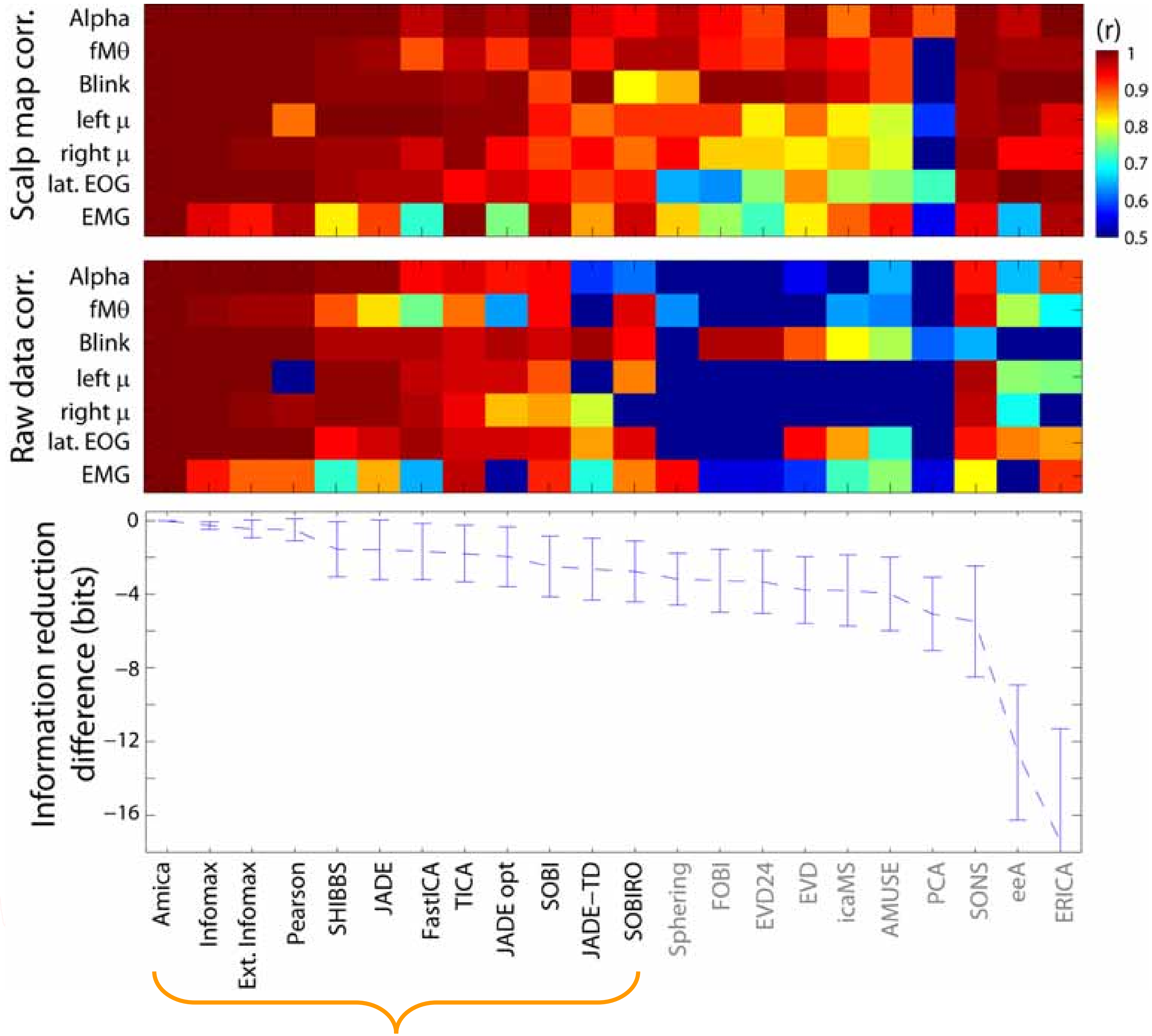


Alpha
fMθ
Blink
left μ
right μ
lat EOG
EMG
Power 10*log 10(μV2/Hz)
Frequency (Hz)
10.2
5.7
10.7
22.0
33.5
10.7
22.0
33.5
0
−10
−20
−30
20
40
Amica

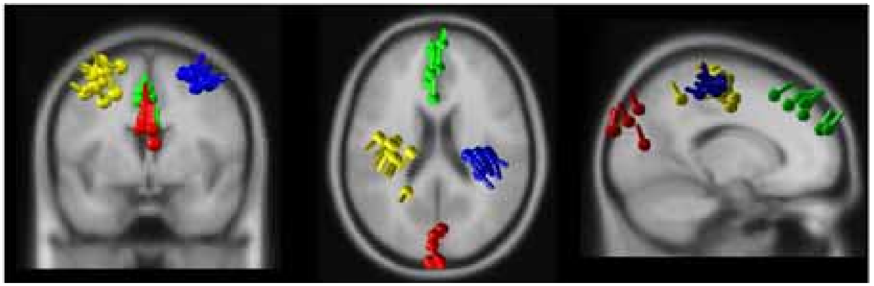


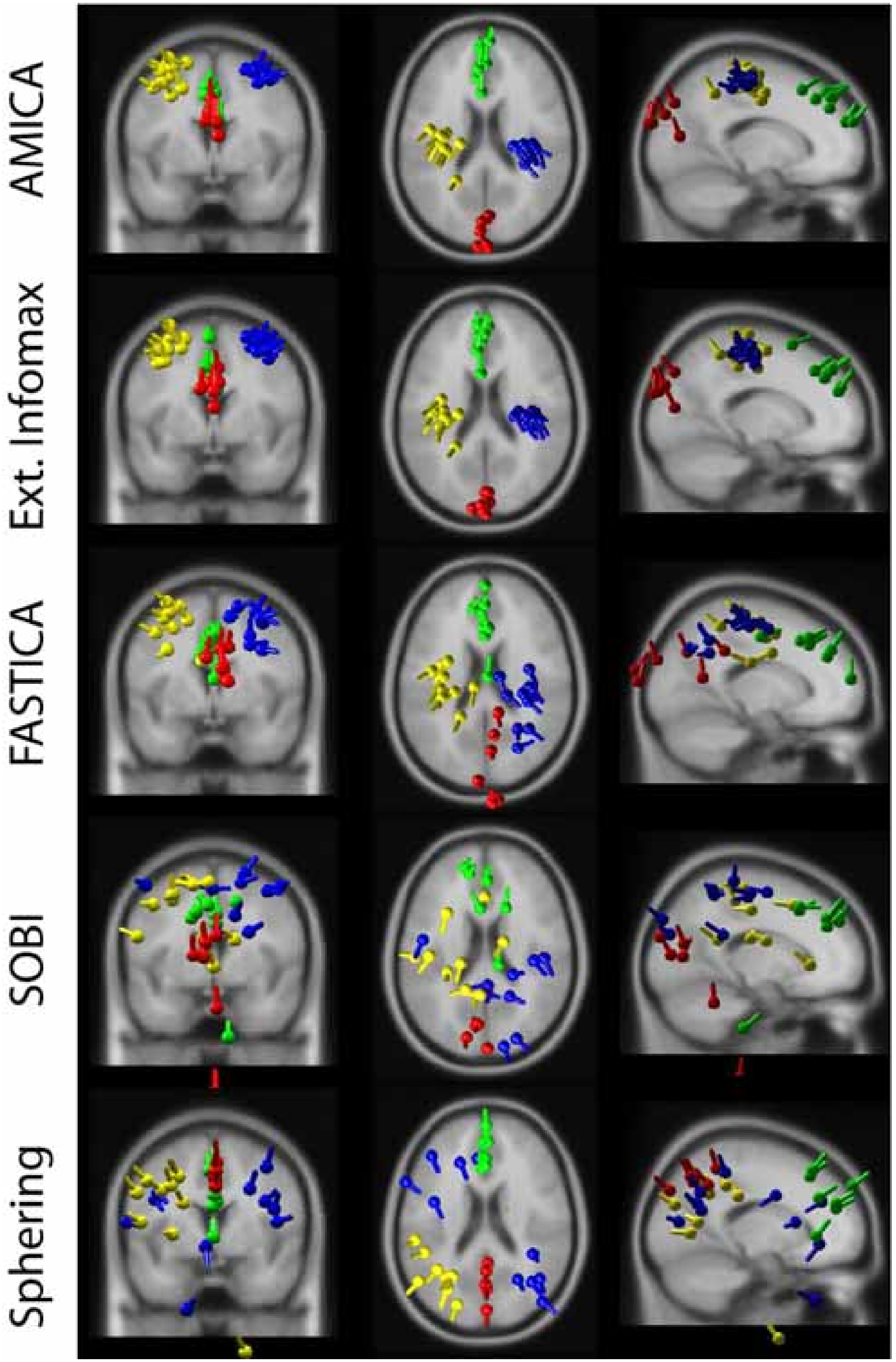
Alpha
fMθ
Blink
left μ
right μ
lat EOG
EMG
Power 10*log 10(μV²/Hz)
Frequency (Hz)
10.2
5.7
10.7
22.0
33.5
Amica
Ext. Infomax
FastICA
SOBI
Sphering

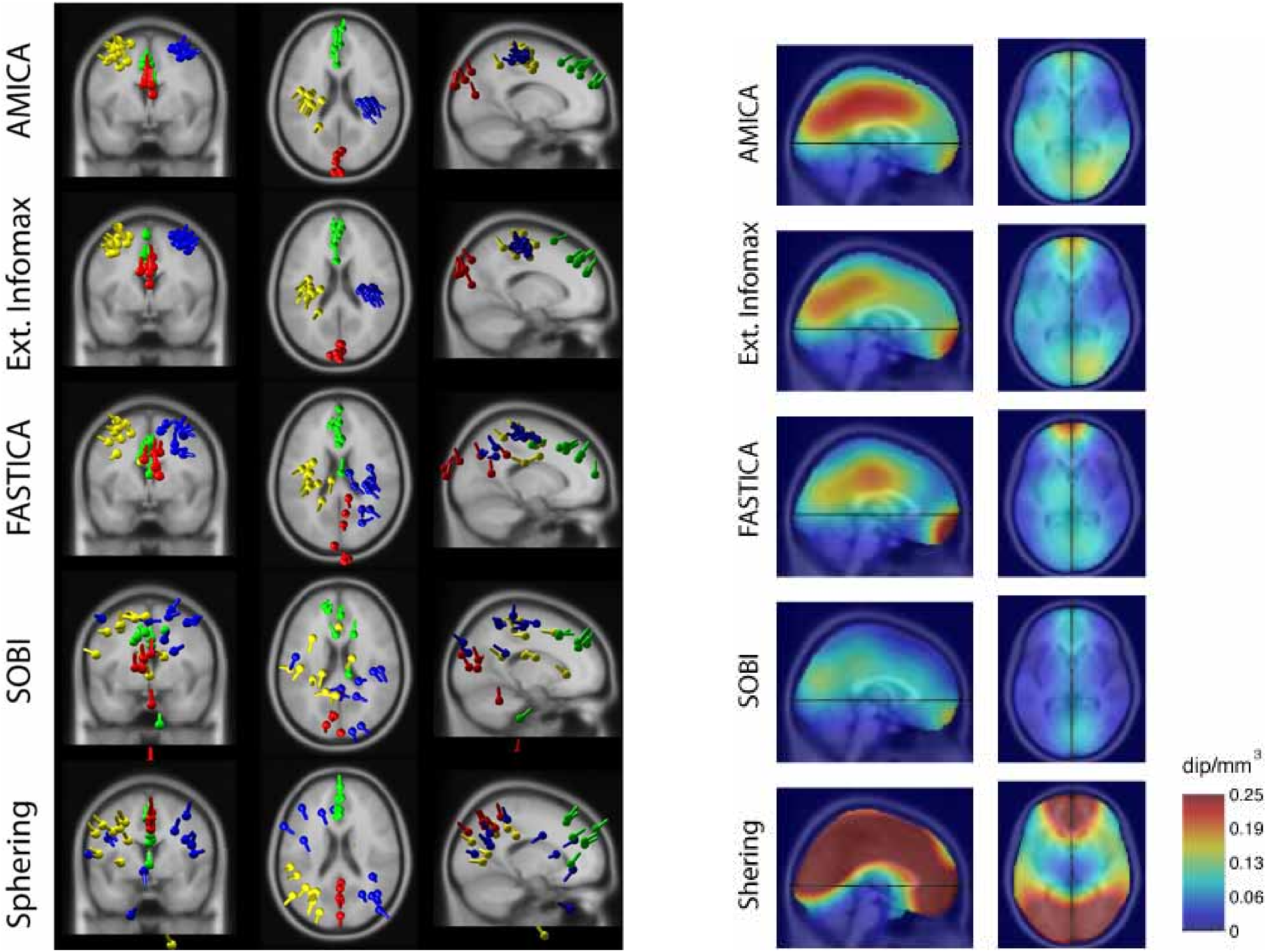
Correlations entre les décompositions



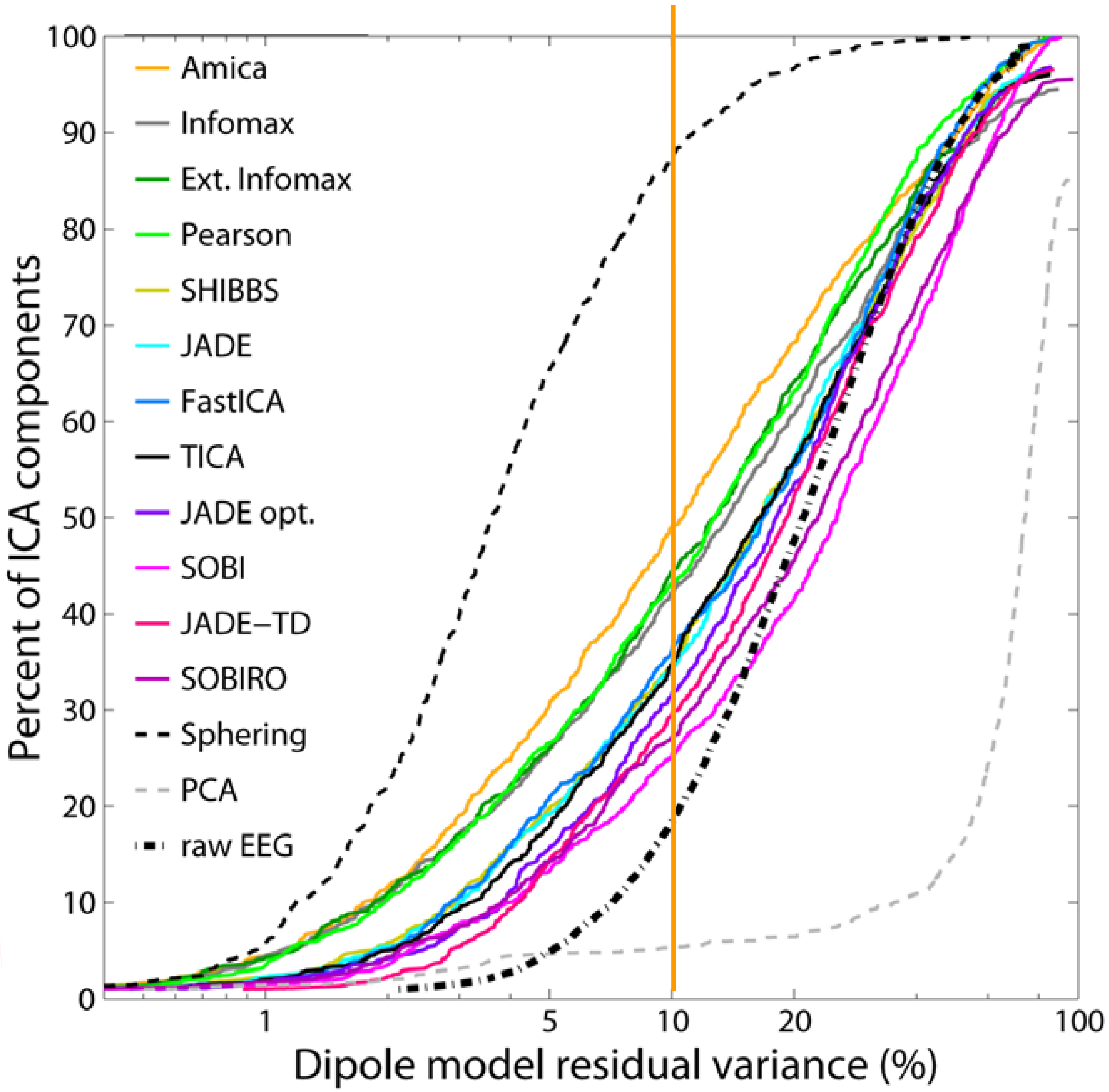
AMICA



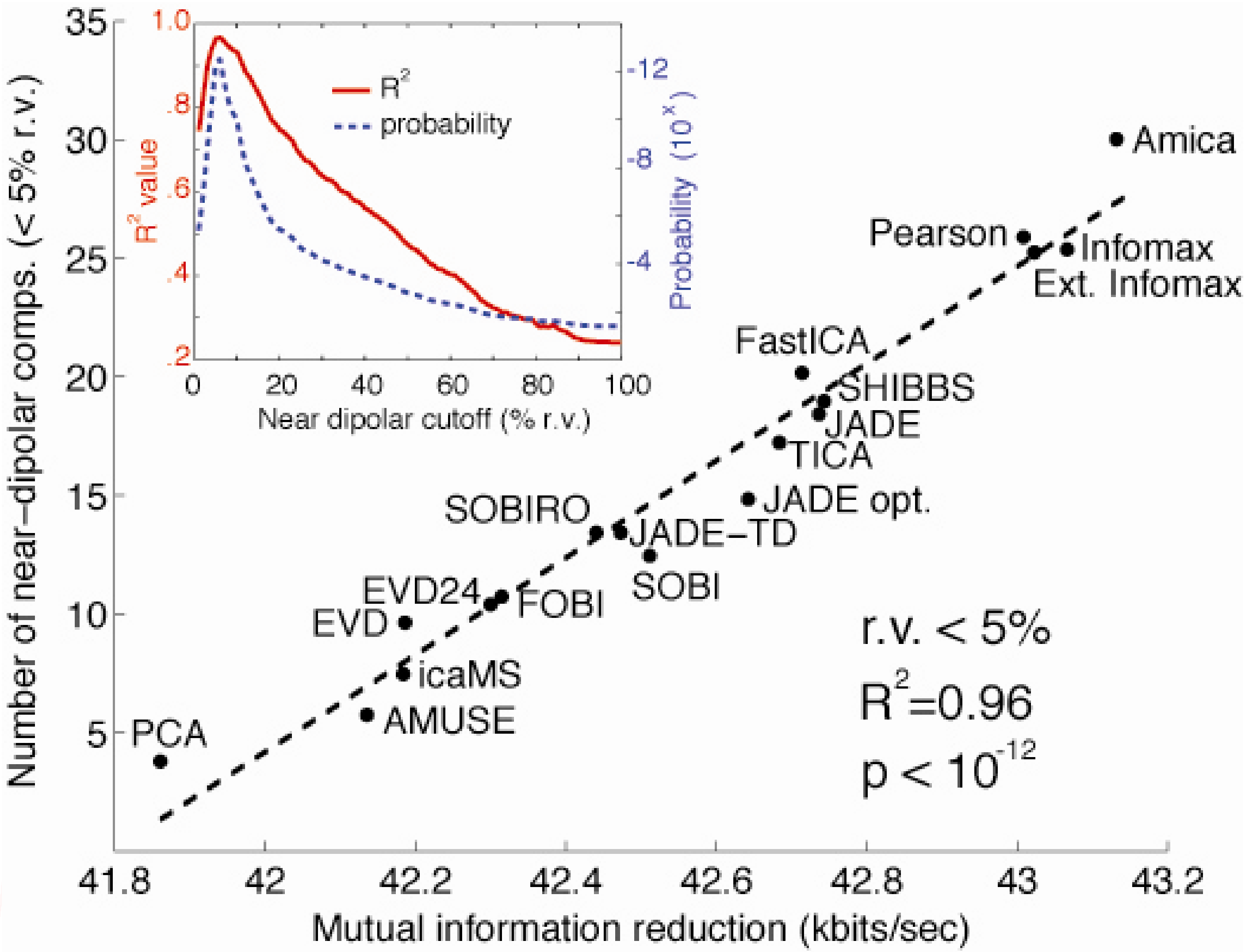

AMICA
Ext. Infomax
FASTICA
SOBI
Sphering

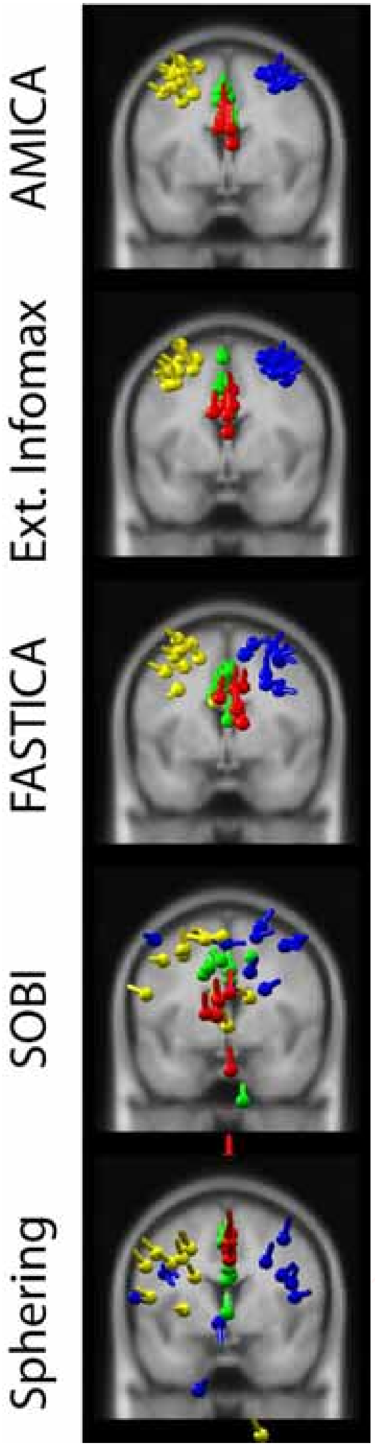

AMICA
Ext. Infomax
FASTICA
SOBI
Sphering
AMICA
Ext. Infomax
FASTICA
SOBI
Shering
dip/mm^3
0.25
0.19
0.13
0.06
0

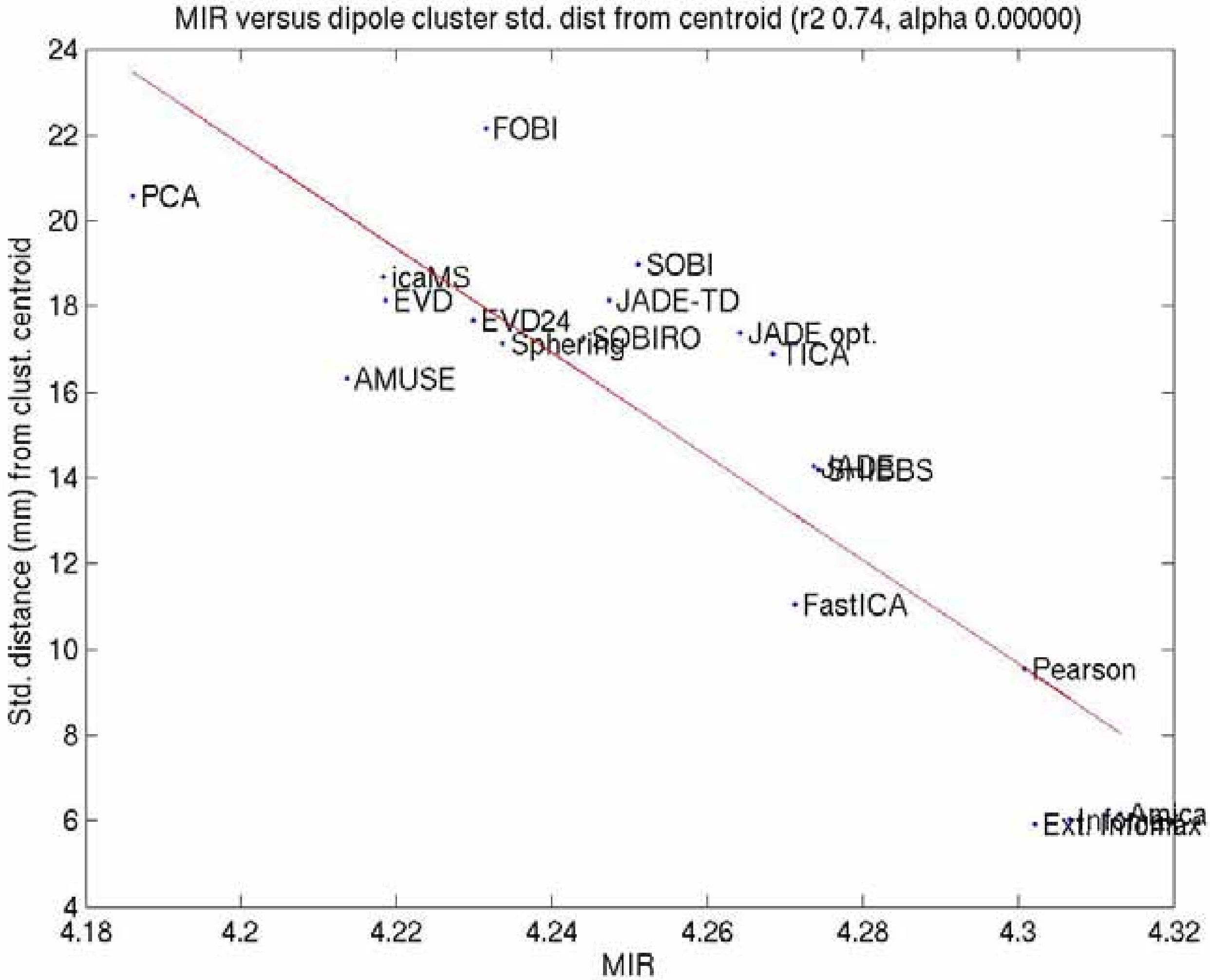
Nombre de composants inférieurs à chaque seuil de variance résiduelle



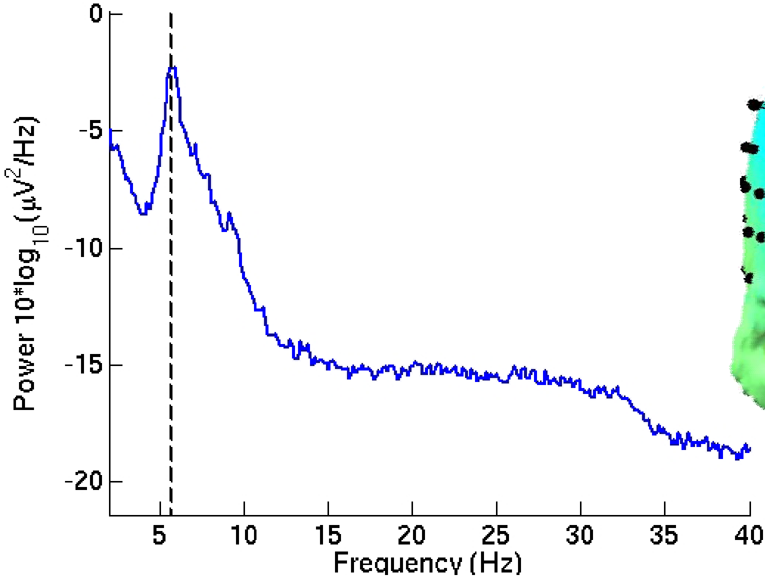
Plus d'indépendence implique plus de composants biologiquement plausibles

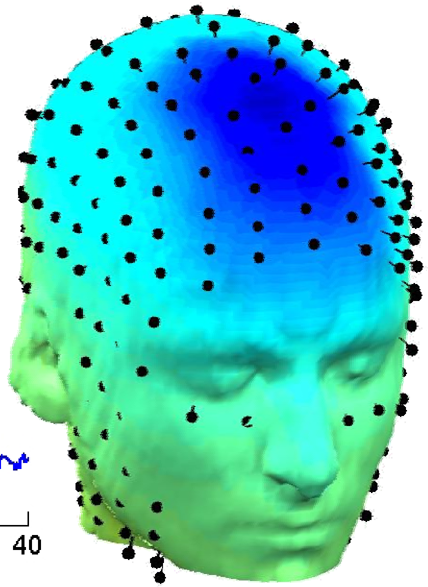


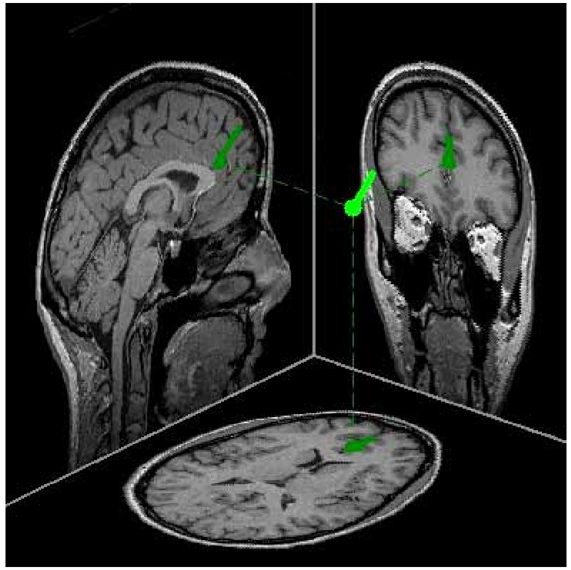

AMICA
Ext. Infomax
FASTICA
SOBI
Sphering


MIR versus dipole cluster std. dist from centroid (r2 0.74, alpha 0.00000)
Std. distance (mm) from clust. centroid
MIR
FOBI
PCA
SOBI
icaMS
EVD
JADE-TD
EVD24
SOBIRO
Sphering
JADE opt.
TICA
AMUSE
FastICA
Pearson
Ext. Infomax

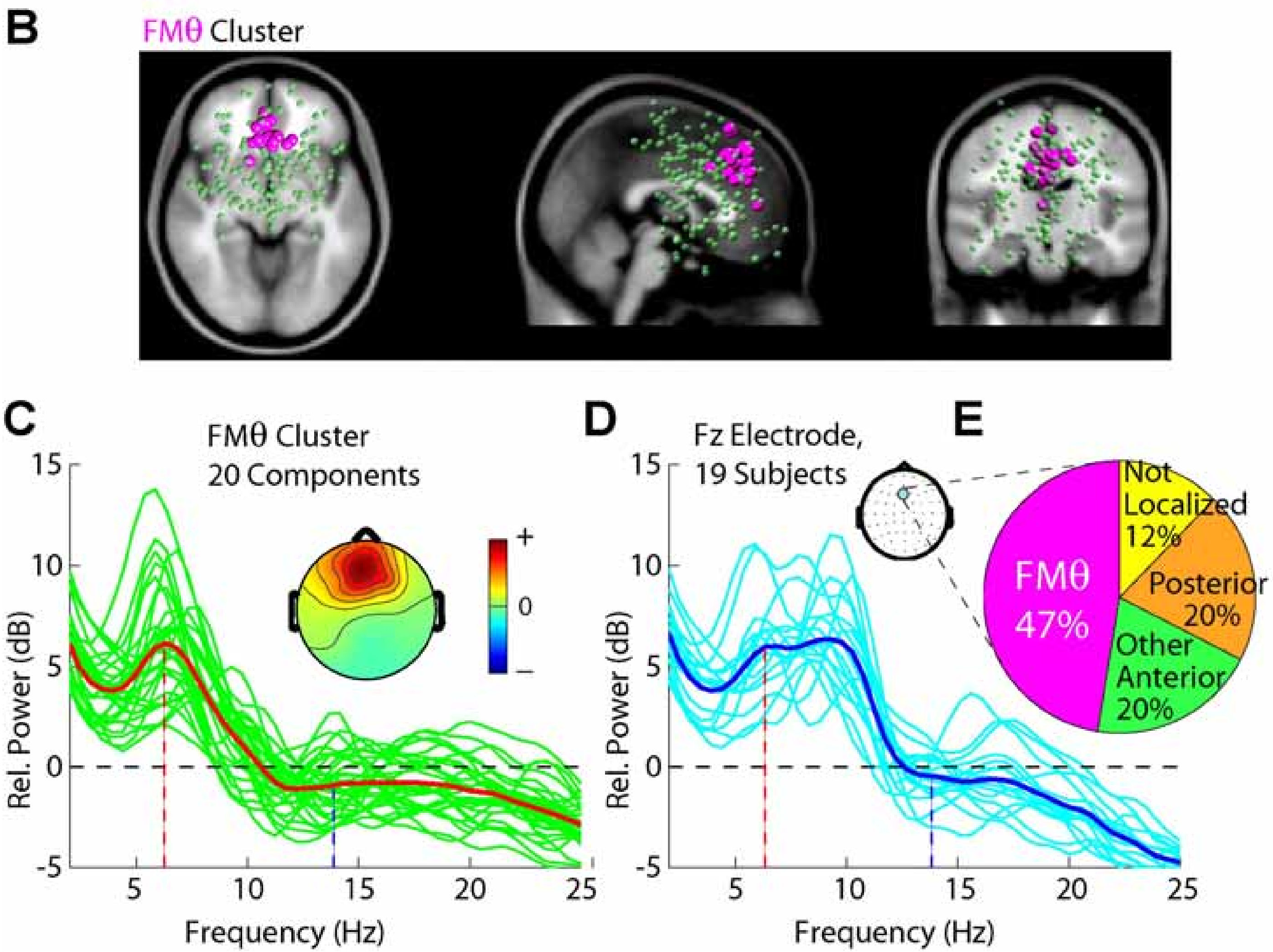
Frontal midline





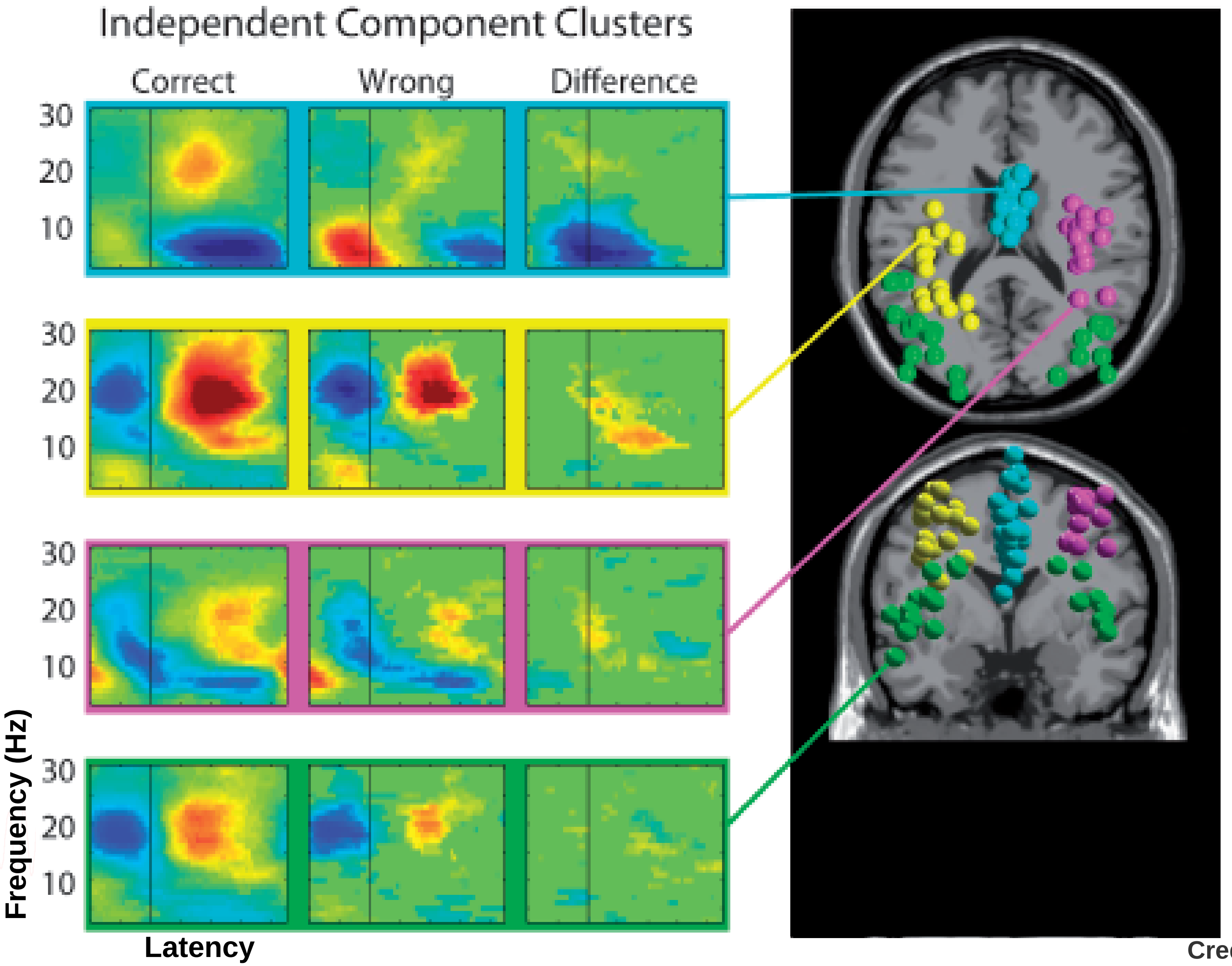


For example: frontal midline theta cluster



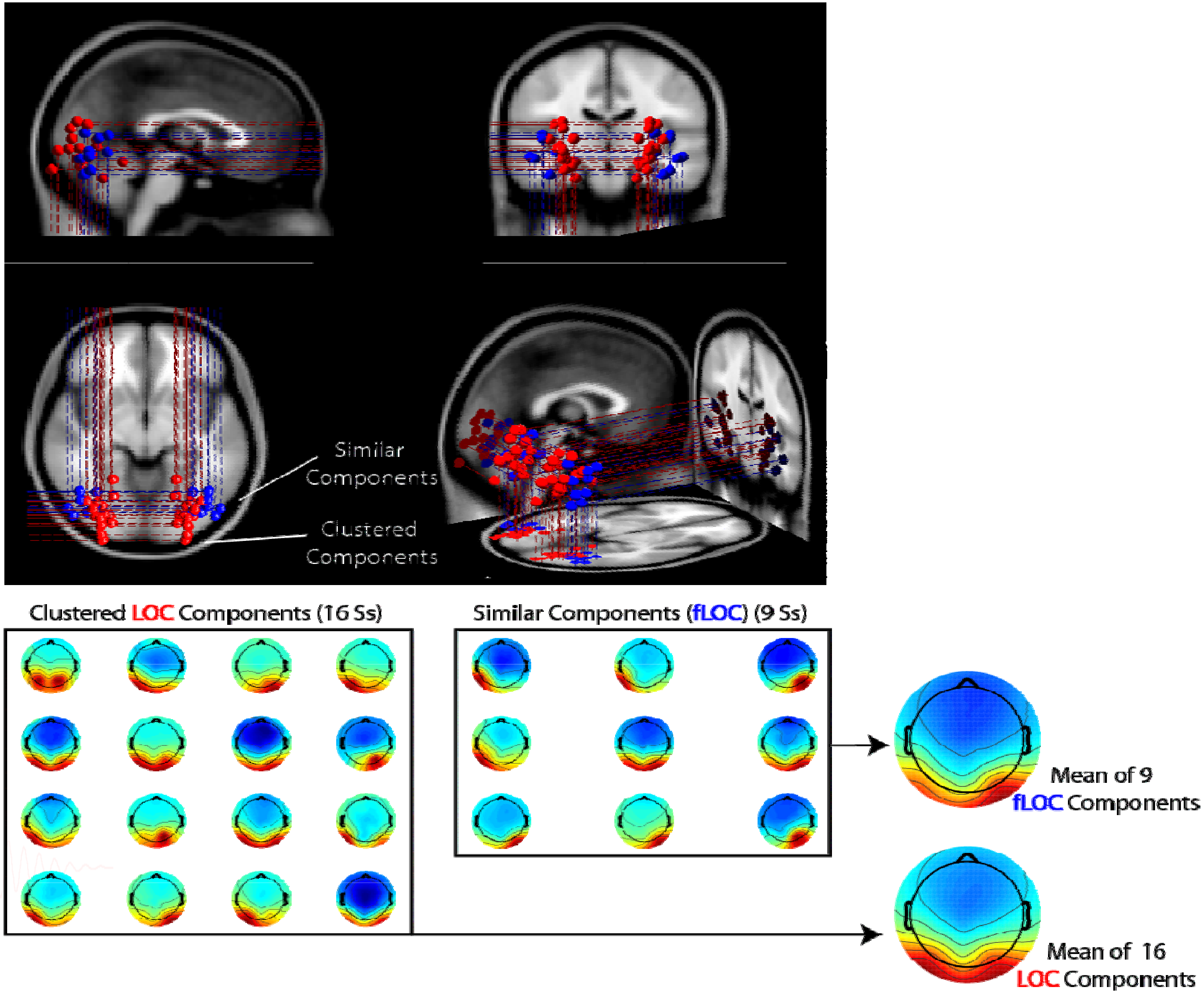
Onton, Delorme and Makeig, NeuroImage 27 (2005) 341 - 356

Goal: to cluster matching ICs across subjects



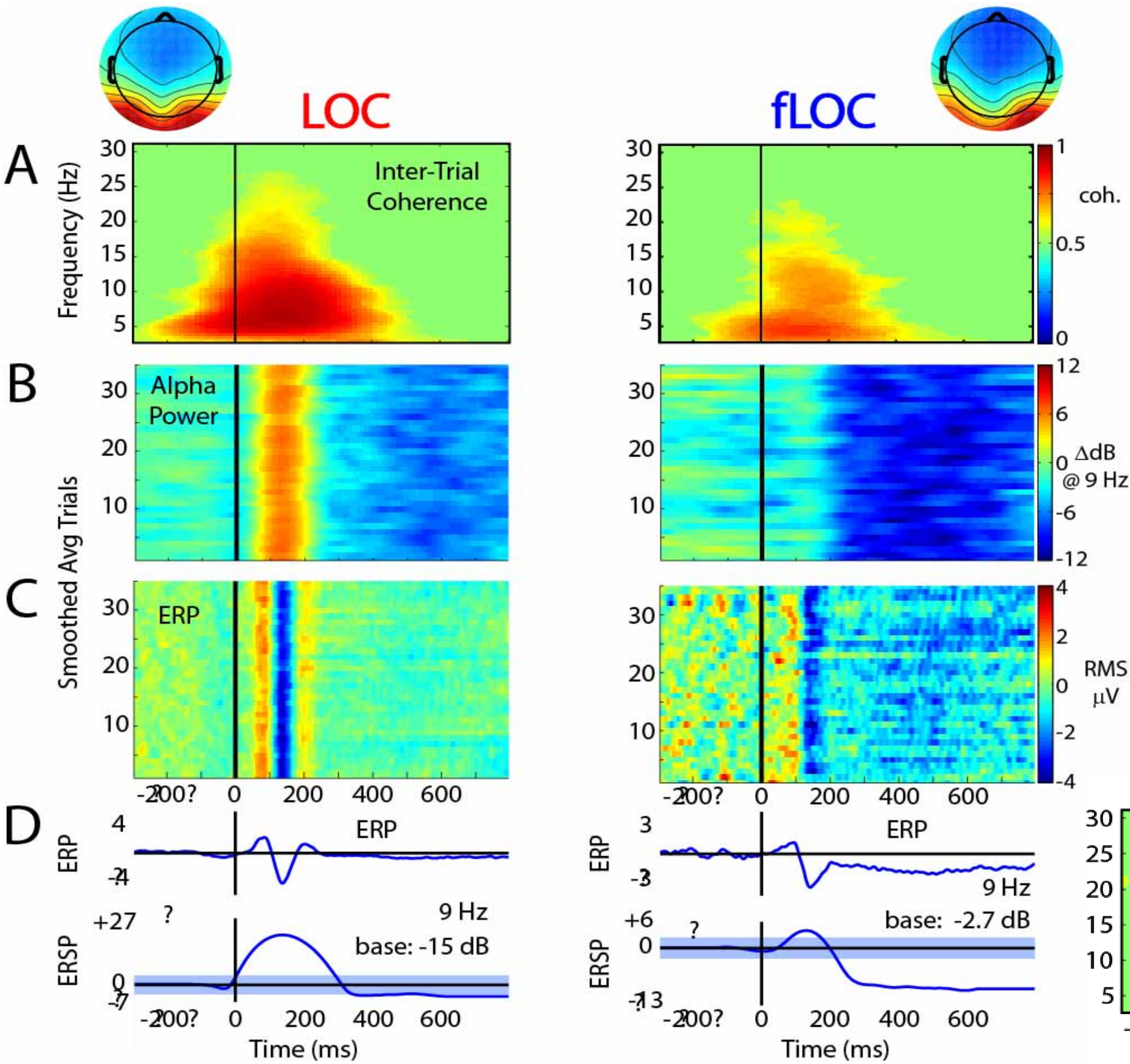
Credit J. Onto

What makes an ICA component?

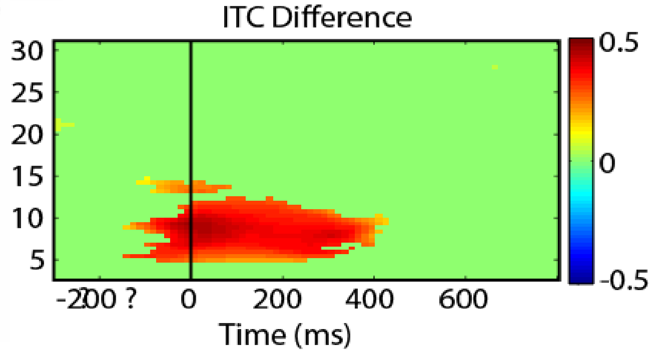


Credit: J. Onton

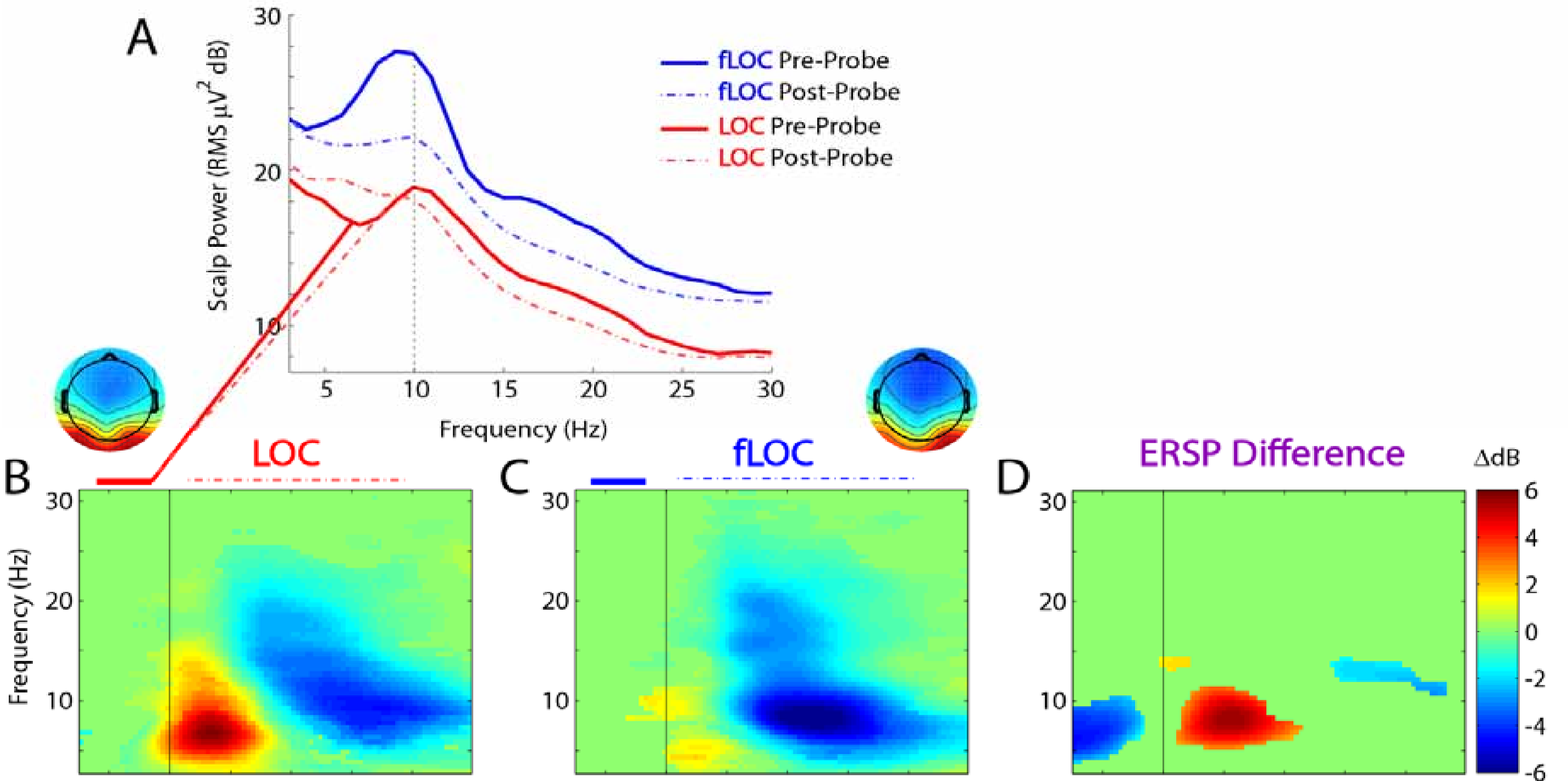
Subject differences?



Significant ITC differences (by bootstrap) between the LOC and fLOC clusters immediately follow Probe presentation (5-11 Hz).



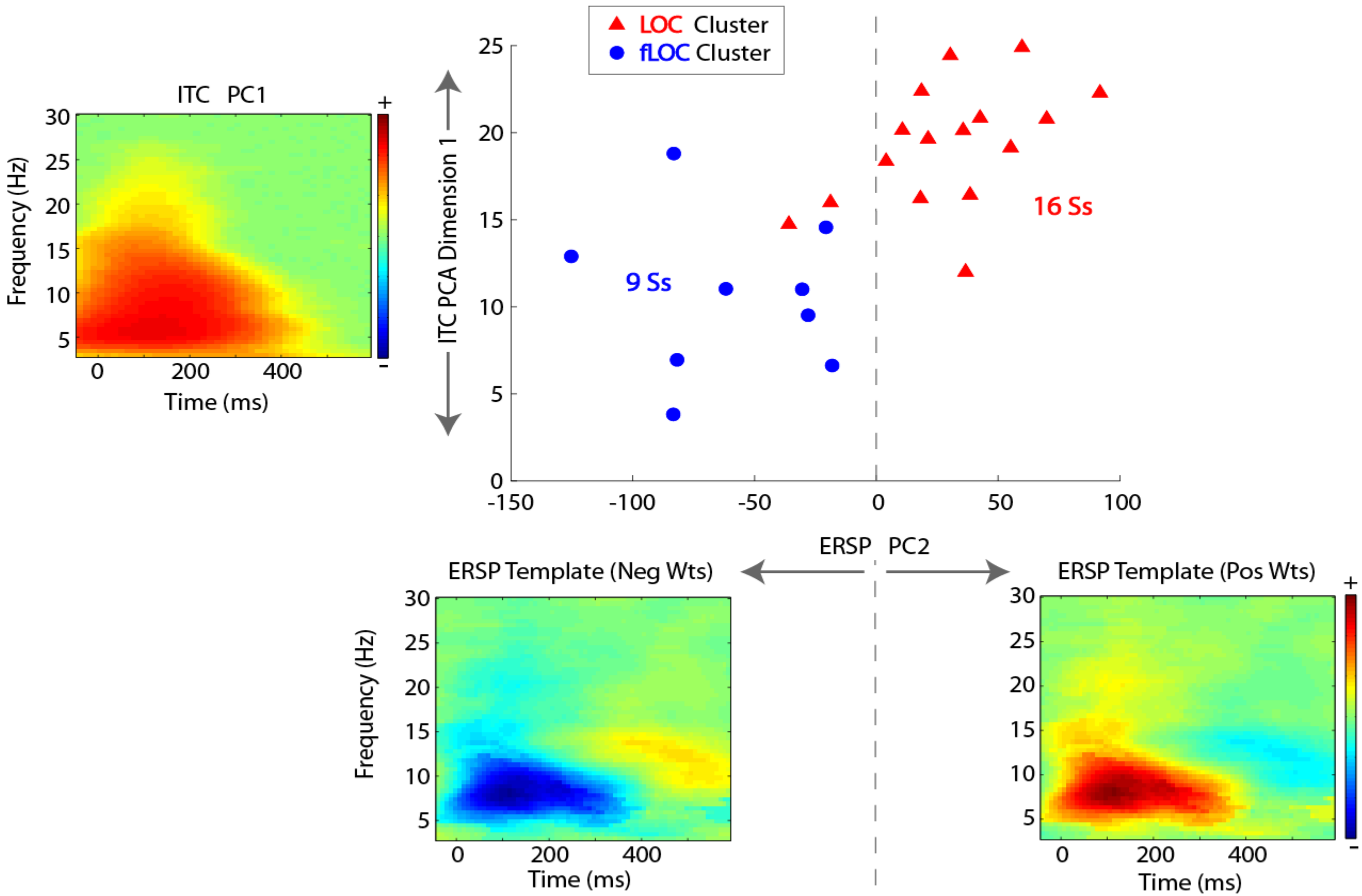
Subject differences?



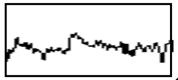
** Cluster ERSPs show significant activity determined by bootstrap statistics within subject and binomial probability between subjects ($p < 0.01$)

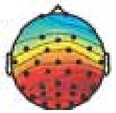
*** Difference ERSP shows significant differences between the two clusters by bootstrap statistics ($p < 0.001$)

Subject differences?



Temporal ICA





ICA activity U

$$\begin{array}{cl}\begin{bmatrix} u_{1,1} & u_{1,2} & u_{1,3} & u_{1,4} & \\ u_{2,1} & u_{2,2} & u_{2,3} & u_{2,4} & \cdots \\ u_{3,1} & u_{3,2} & u_{3,3} & u_{3,4} & \end{bmatrix} & \begin{array}{l}\leftarrow \text{Comp. 1}\\ \leftarrow \text{Comp. 2}\\ \leftarrow \text{Comp. 3}\end{array}\end{array}$$

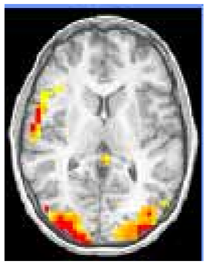
$$\mathbf{X}=\mathbf{W}^{-1}\mathbf{U}$$

$$\begin{bmatrix} w^{-1}_{1,1} & w^{-1}_{1,2} & w^{-1}_{1,3} \\ w^{-1}_{2,1} & w^{-1}_{2,2} & w^{-1}_{2,3} \\ w^{-1}_{3,1} & w^{-1}_{3,2} & w^{-1}_{3,3} \end{bmatrix} \longrightarrow \begin{array}{cl}\begin{bmatrix} u_{1,1}w^{-1}_{1,1}+u_{2,1}w^{-1}_{1,2}+u_{3,1}w^{-1}_{1,3} & u_{1,2}w^{-1}_{1,1}+u_{2,2}w^{-1}_{1,2}+u_{3,2}w^{-1}_{1,3} & \\ u_{1,1}w^{-1}_{2,1}+u_{2,1}w^{-1}_{2,2}+u_{3,1}w^{-1}_{2,3} & u_{1,2}w^{-1}_{2,1}+u_{2,2}w^{-1}_{2,2}+u_{3,2}w^{-1}_{2,3} & \cdots \\ u_{1,1}w^{-1}_{3,1}+u_{2,1}w^{-1}_{3,2}+u_{3,1}w^{-1}_{3,3} & u_{1,2}w^{-1}_{3,1}+u_{2,2}w^{-1}_{3,2}+u_{3,2}w^{-1}_{3,3} & \end{bmatrix} & \begin{array}{l}\leftarrow \text{Chan 1}\\ \leftarrow \text{Chan 2}\\ \leftarrow \text{Chan 3}\end{array}\end{array}$$

Inverse weight matrix W^{-1}

Data X (EEG/MEG time series)

Spatial ICA



ICA activity U

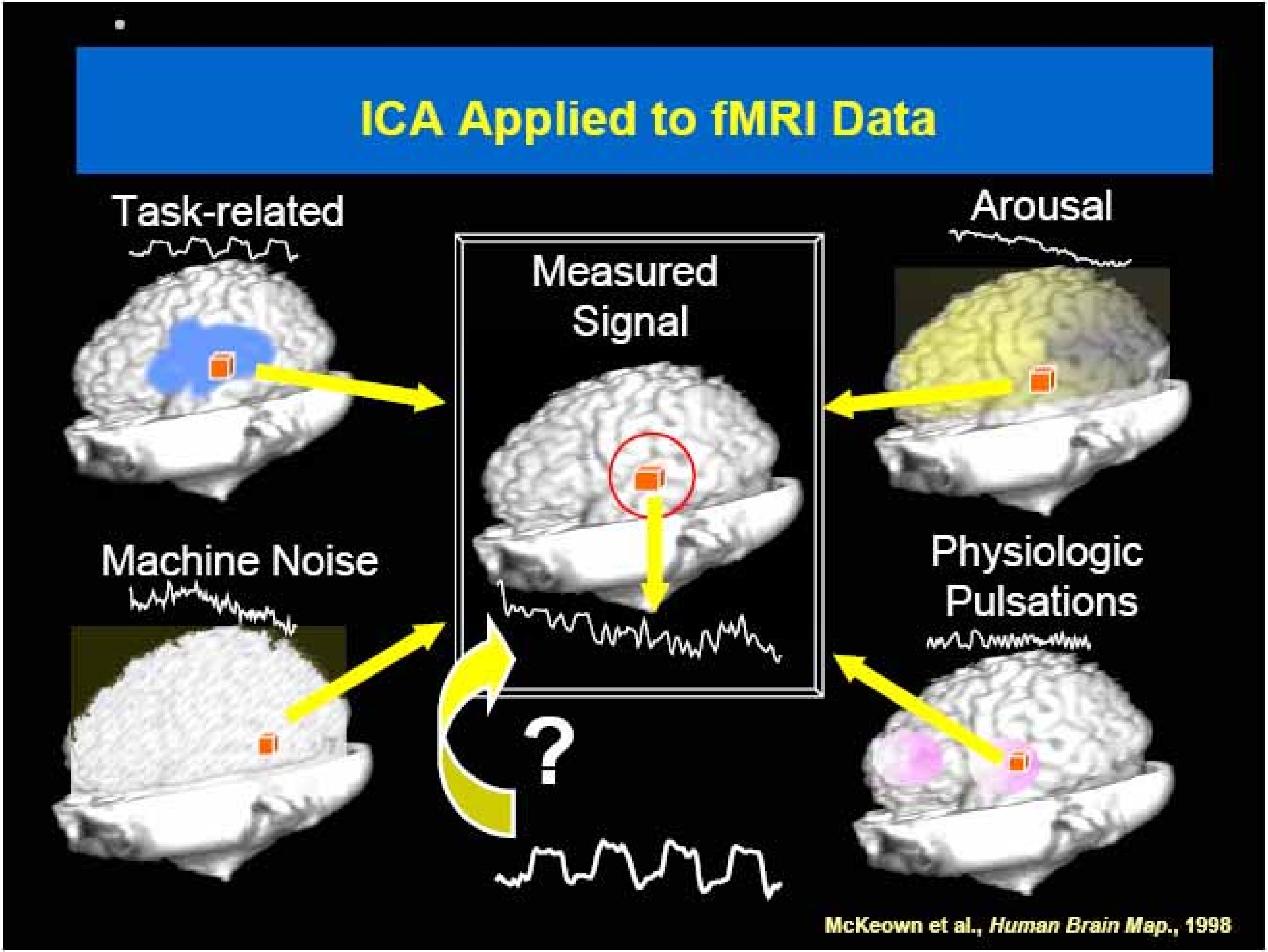
$$\begin{array}{cl}\begin{bmatrix} u_{1,1} & u_{1,2} & u_{1,3} & u_{1,4} & \\ u_{2,1} & u_{2,2} & u_{2,3} & u_{2,4} & \cdots \\ u_{3,1} & u_{3,2} & u_{3,3} & u_{3,4} & \end{bmatrix} & \begin{array}{l}\leftarrow \text{Comp. 1}\\ \leftarrow \text{Comp. 2}\\ \leftarrow \text{Comp. 3}\end{array}\end{array}$$

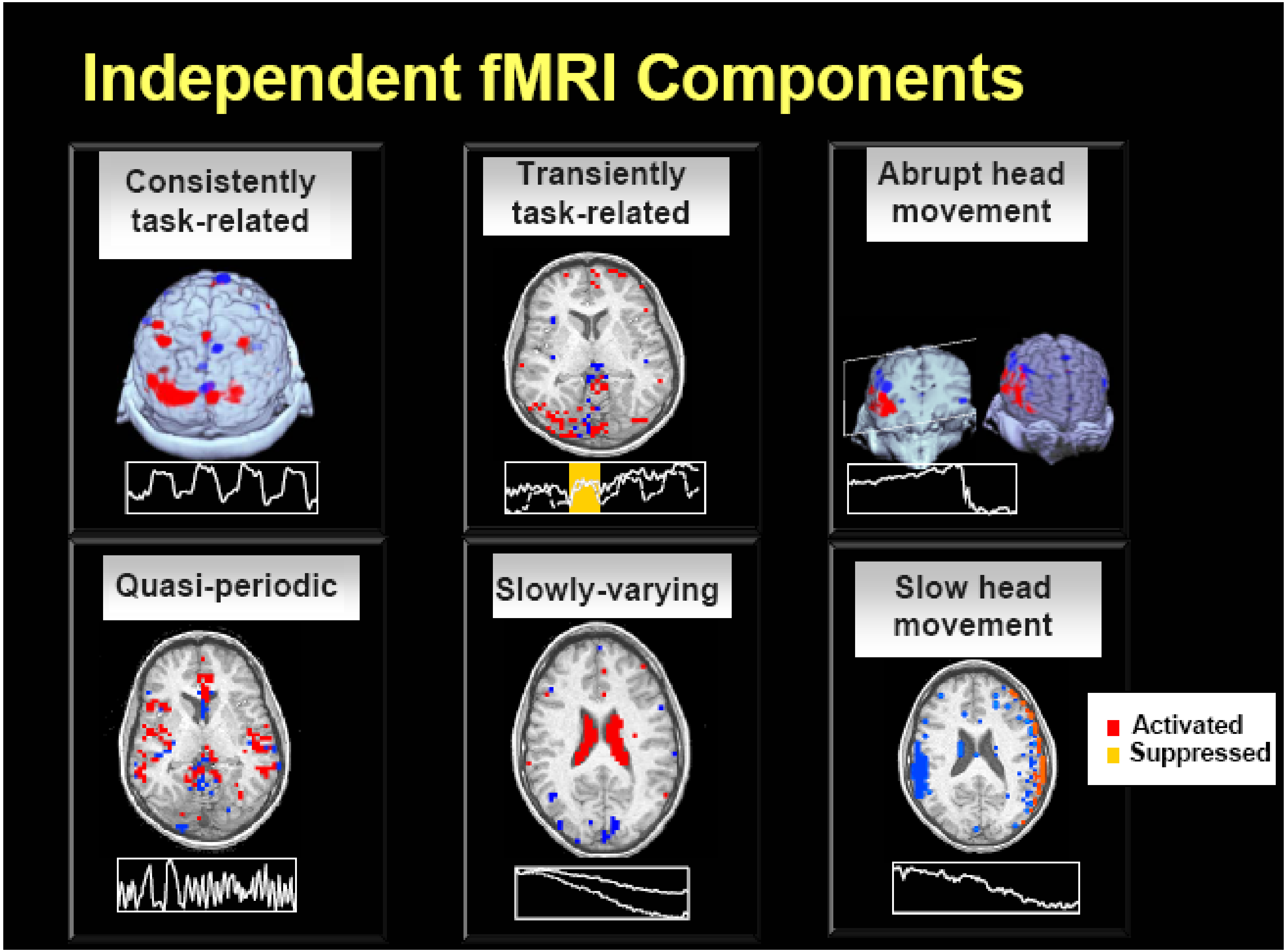
$$\mathbf{X}=\mathbf{W}^{-1}\mathbf{U}$$

$$\begin{bmatrix} w^{-1}_{1,1} & w^{-1}_{1,2} & w^{-1}_{1,3} \\ w^{-1}_{2,1} & w^{-1}_{2,2} & w^{-1}_{2,3} \\ w^{-1}_{3,1} & w^{-1}_{3,2} & w^{-1}_{3,3} \end{bmatrix} \longrightarrow \begin{array}{cl}\begin{bmatrix} u_{1,1}w^{-1}_{1,1}+u_{2,1}w^{-1}_{1,2}+u_{3,1}w^{-1}_{1,3} & u_{1,2}w^{-1}_{1,1}+u_{2,2}w^{-1}_{1,2}+u_{3,2}w^{-1}_{1,3} & \\ u_{1,1}w^{-1}_{2,1}+u_{2,1}w^{-1}_{2,2}+u_{3,1}w^{-1}_{2,3} & u_{1,2}w^{-1}_{2,1}+u_{2,2}w^{-1}_{2,2}+u_{3,2}w^{-1}_{2,3} & \cdots \\ u_{1,1}w^{-1}_{3,1}+u_{2,1}w^{-1}_{3,2}+u_{3,1}w^{-1}_{3,3} & u_{1,2}w^{-1}_{3,1}+u_{2,2}w^{-1}_{3,2}+u_{3,2}w^{-1}_{3,3} & \end{bmatrix} & \begin{array}{l}\leftarrow \text{Time 1}\\ \leftarrow \text{Time 2}\\ \leftarrow \text{Time 3}\end{array}\end{array}$$

Inverse weight matrix W^{-1}

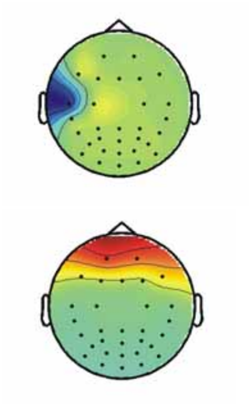
Data X (EEG/MEG voxel activities)


ICA Applied to fMRI Data
Task-related
Measured
Signal
Arousal
Machine Noise
Physiologic
Pulsations
?
McKeown et al., Human Brain Map., 1998


Independent fMRI Components
Consistently task-related
Transiently task-related
Abrupt head movement
Quasi-periodic
Slowly-varying
Slow head movement
Activated
Suppressed

Rejection: Raw Data vs. ICA

Artificial artifacts



1 - Transient high-frequency events (20-60 Hz) modeling temporal muscle artifacts

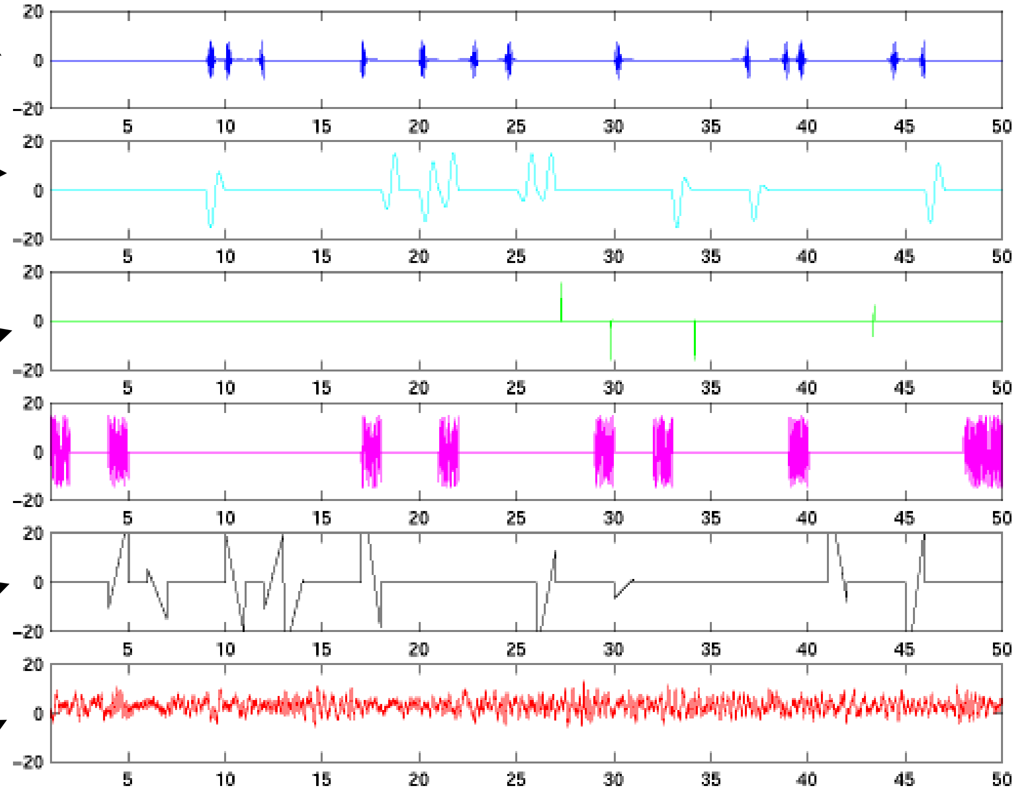
2 - Low-frequency events (1-3 Hz) modeling eye movement artifacts

3 - Signal discontinuities from electrical artifacts

4 - High noise EEG artifacts

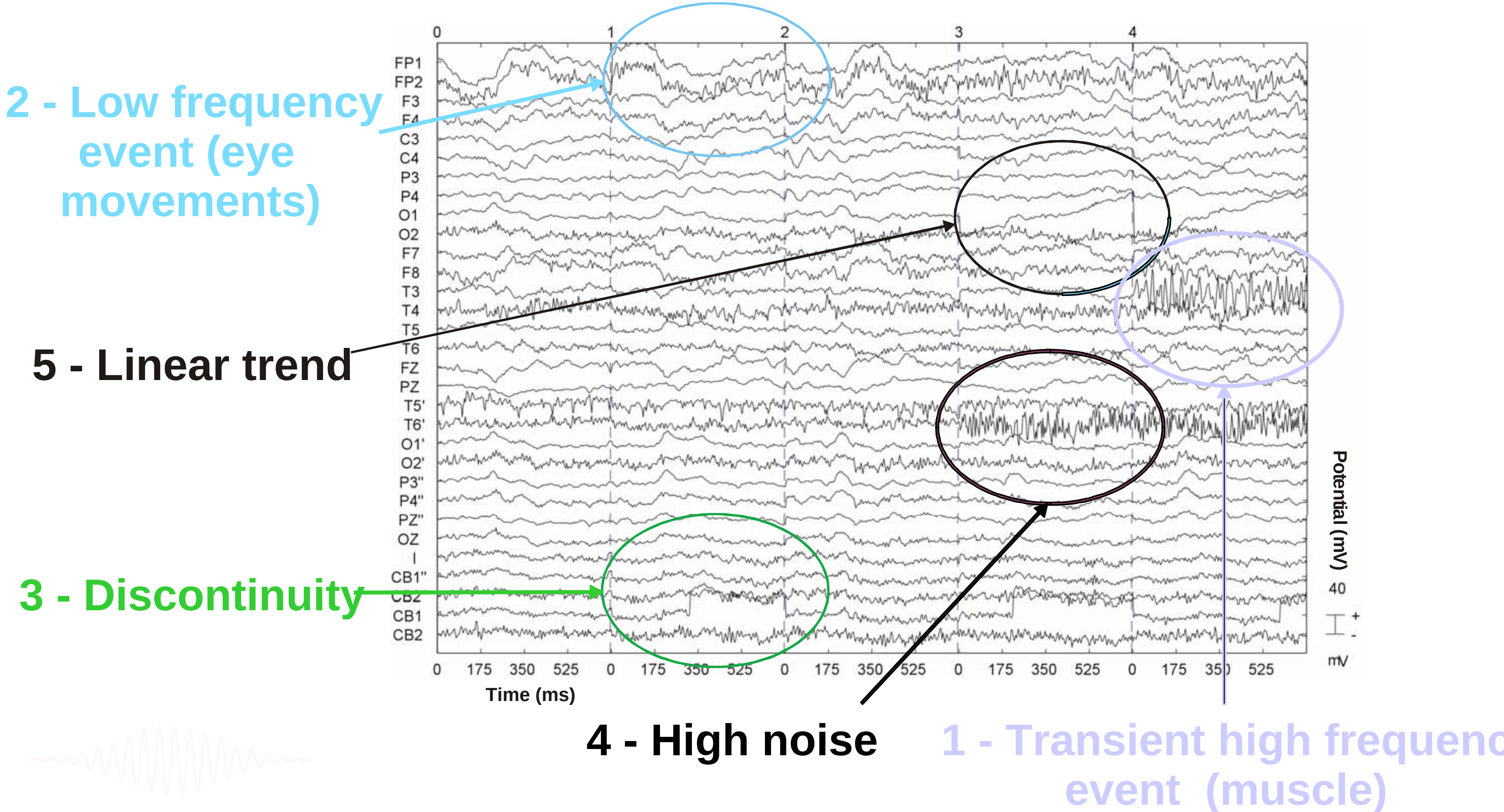
5 - Linear trends from electrical artifacts

EEG activity (31 channels x 100 trials of artifact-free EEG)



Rejection: Raw Data vs. ICA

Simulated data



Rejection: Raw Data vs. ICA

Rejection methods

- Detection of peaks of activity (thresholding)
- Detection of linear trends (R^2)
- Detection of improbable events (Joint Probability)
- Detection of peaky distributions of activity (Kurtosis)
- Detection of frequency peaks (frequency thresholding)

Vary artifacts' amplitude

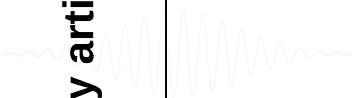
Known artifacts

↓

Optimize the free parameter for each method

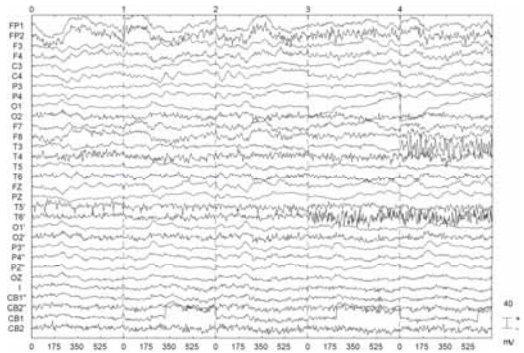
↓

Performance of each method

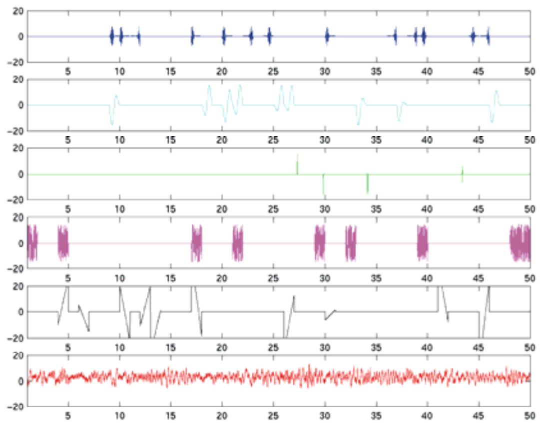


Rejection: Raw Data vs. ICA

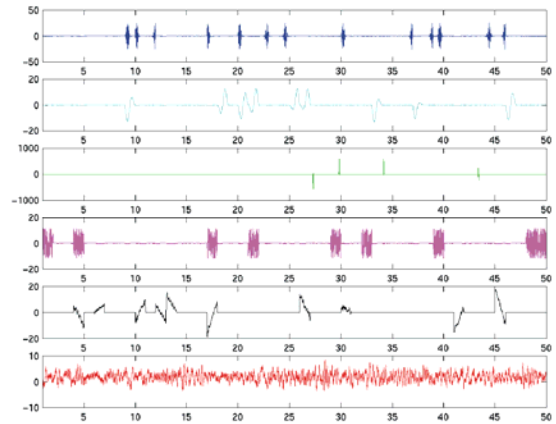
Mixture

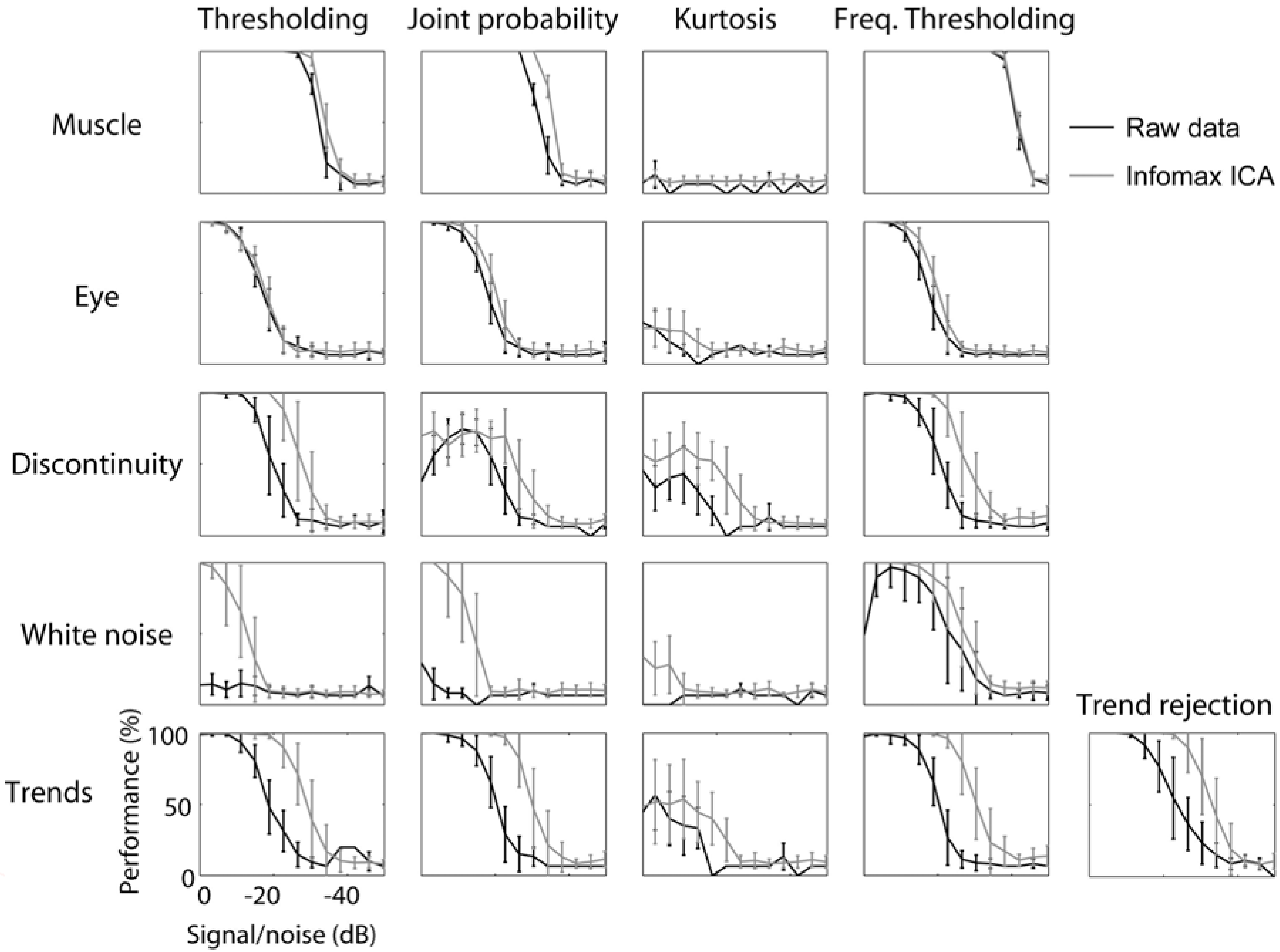


Original data

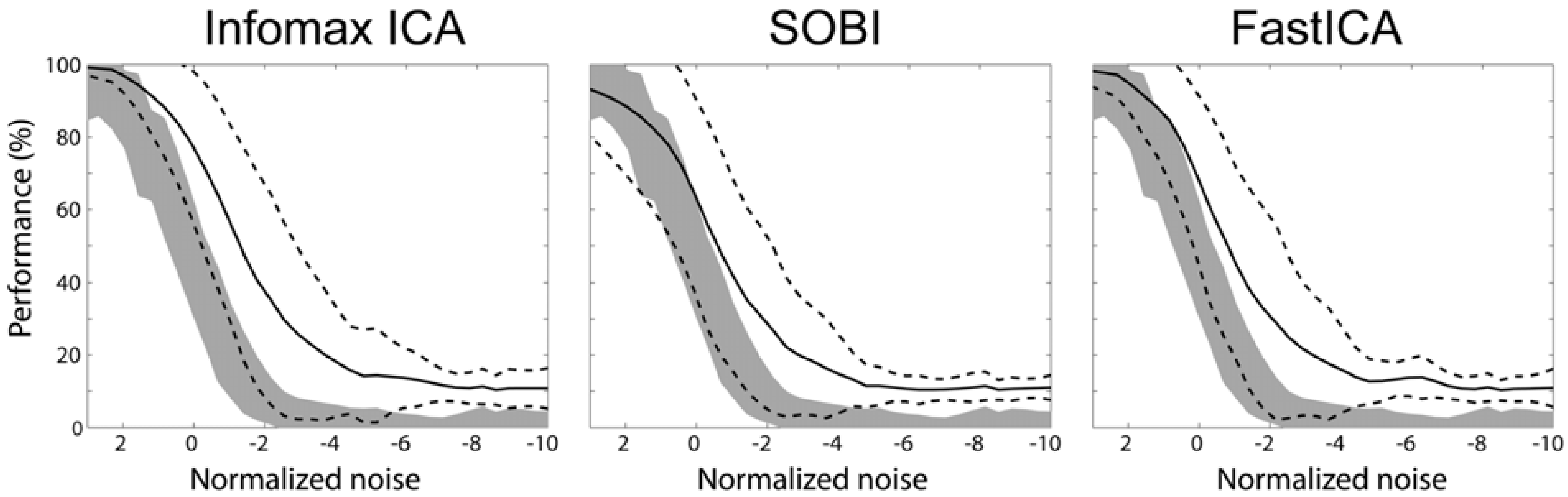


Infomax ICA





Thresholding
Joint probability
Kurtosis
Freq. Thresholding
Muscle
Eye
Discontinuity
White noise
Trends
Trend rejection
Raw data
Infomax ICA
Performance (%)
100
50
0
0
-20
-40
Signal/noise (dB)



Infomax ICA
SOBI
FastICA
Performance (%)
100
80
60
40
20
0
2
0
-2
-4
-6
-8
-10
Normalized noise