ICA-based artifact removal in EEG

John J.B. Allen

http://jallen.faculty.arizona.edu/ICA_Workshop

Würzburg Workshop, 26 June, 2015

ICA-based artifact removal in EEG

John J.B. Allen

http://jallen.faculty.arizona.edu/ICA_Workshop

Würzburg Workshop, 26 June, 2015

“Clinical” vs Actuarial Approaches

Clinical Versus Actuarial Judgment

ROBYN M. DAWES, DAVID FAUST, PAUL E. MEEHL

Professionals are frequently consulted to diagnose and predict human behavior; optimal treatment and planning often hinge on the consultant's judgmental accuracy. The consultant may rely on one of two contrasting approaches to decision-making—the clinical and actuarial methods. Research comparing these two approaches shows the actuarial method to be superior. Factors underlying the greater accuracy of actuarial methods, sources of resistance to the scientific findings, and the benefits of increased reliance on actuarial approaches are discussed.

a clinical practitioner. A clinician in psychiatry or medicine may use the clinical or actuarial method. Conversely, the actuarial method should not be equated with automated decision rules alone. For example, computers can automate clinical judgments. The computer can be programmed to yield the description “dependency traits,” just as the clinical judge would, whenever a certain response appears on a psychological test. To be truly actuarial, interpretations must be both automatic (that is, prespecified or routinized) and based on empirically established relations.

Virtually any type of data is amenable to actuarial interpretation. For example, interview observations can be coded quantitatively (patient appears withdrawn: [1] yes, [2] no). It is thereby possible to incorporate qualitative observations and quantitative data into

<u>Dawes, R.M., Faust, D., & Meehl, P.E.</u>(1989). *Science*, *243*, 1668-1674.

“Clinical” vs Actuarial Approaches

- Human raters
  - Good source of possible algorithms
  - Lousy at reliably implementing them
    - Inter-rater
    - Intra-rater
- Actuarial methods
  - Always arrive at the same conclusion
  - Weight variables according to *actual* predictive power

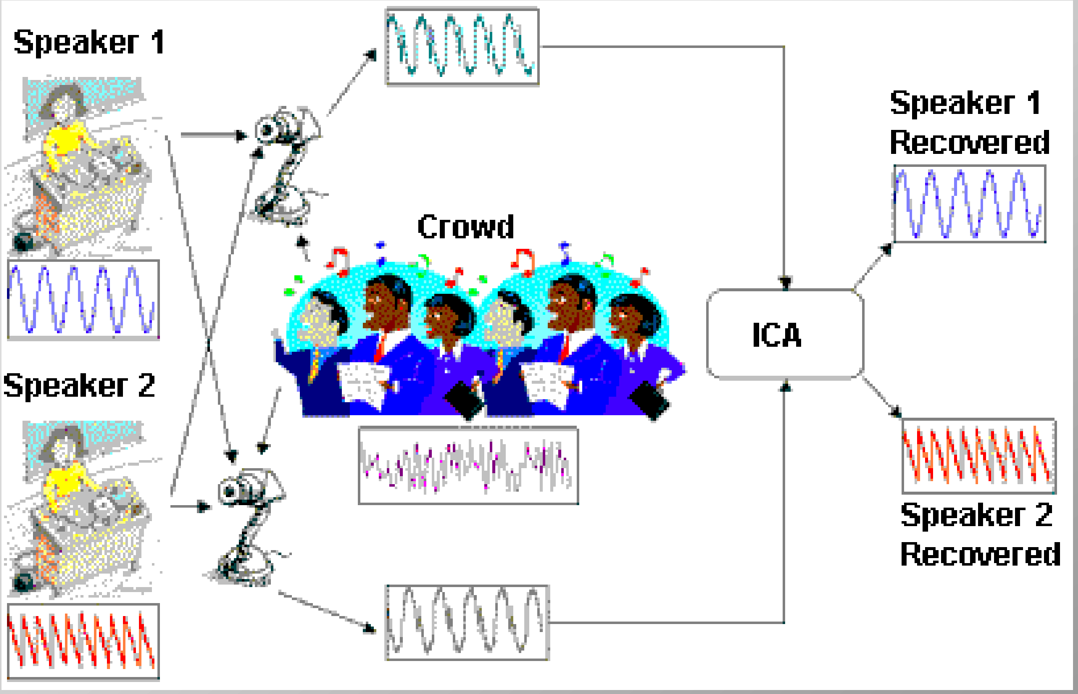
Dawes, R.M., Faust, D., & Meehl, P.E.(1989). *Science*, *243*, 1668-1674.

Algorithmic EEG Artifact Approaches

- Amplitude criteria
  - Too sudden: 100uV between samples
- Activity criteria
  - Too big: 250uV/2sec epoch
  - Too small: <0.5 uV range across 100ms
- Regression-based EOG correction
- Can we make ICA-based correction algorithmic?

What is ICA?

- ICA is a "blind source separation" technique.
- ICA separates sources of activity that are mixed together at recording electrodes.

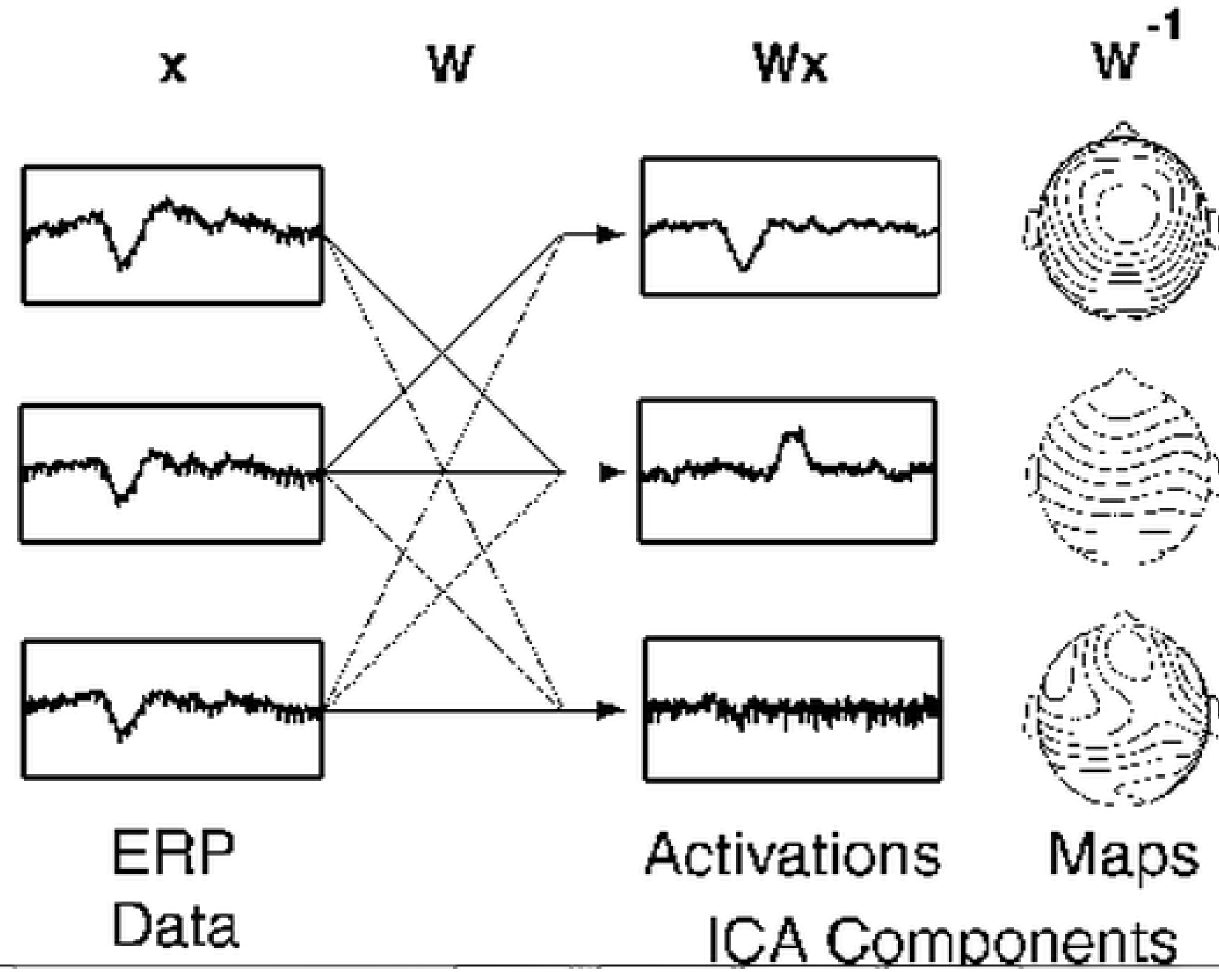


What is ICA?

- For EEG data
  - Channel data (X) can be thought of as a weighted (W) combination of independent component activations (Wx), each of which has a scalp projection (W^{-1}).

- You can think of ICs as putative sources of the scalp-recorded EEG.

ICA Decomposition



Psychophysiology, *48* (2011), 229–240. Wiley Periodicals, Inc. Printed in the USA.

DOI: 10.1111/j.1469-8986.2010.01061.x

ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features

ANDREA MOGNON,[a,b] JORGE JOVICICH,[a] LORENZO BRUZZONE,[c] AND MARCO BUIATTI[a,d,e,f]

[a]Functional NeuroImaging Laboratory, Center for Mind/Brain Sciences, Department of Cognitive and Education Sciences, University of Trento, Trento, Italy
[b]NILab, Neuroinformatics Laboratory, Fondazione Bruno Kessler, Trento, Italy
[c]Department of Information Engineering and Computer Science, University of Trento, Trento, Italy
[d]INSERM, U992, Cognitive Neuroimaging Unit, Gif/Yvette, France
[e]CEA, DSV/I2BM, NeuroSpin Center, Gif/Yvette, France
[f]Université Paris-Sud, Cognitive Neuroimaging Unit, Gif/Yvette, France

Abstract

A successful method for removing artifacts from electroencephalogram (EEG) recordings is Independent Component Analysis (ICA), but its implementation remains largely user-dependent. Here, we propose a completely automatic algorithm (ADJUST) that identifies artifacted independent components by combining stereotyped artifact-specific spatial and temporal features. Features were optimized to capture blinks, eye movements, and generic discontinuities on a feature selection dataset. Validation on a totally different EEG dataset shows that (1) ADJUST's classification of independent components largely matches a manual one by experts (agreement on 95.2% of the data variance), and (2) Removal of the artifacted components detected by ADJUST leads to neat reconstruction of visual and auditory event-related potentials from heavily artifacted data. These results demonstrate that ADJUST provides a fast, efficient, and automatic way to use ICA for artifact removal.

Winkler et al. Behavioral and Brain Functions 2011, 7:30
http://www.behavioralandbrainfunctions.com/content/7/1/30

BEHAVIORAL AND BRAIN FUNCTIONS

METHODOLOGY Open Access

Automatic Classification of Artifactual ICA-Components for Artifact Removal in EEG Signals

Irene Winkler*, Stefan Haufe and Michael Tangermann

Abstract

Background: Artifacts contained in EEG recordings hamper both, the visual interpretation by experts as well as the algorithmic processing and analysis (e.g. for Brain-Computer Interfaces (BCI) or for Mental State Monitoring). While hand-optimized selection of source components derived from Independent Component Analysis (ICA) to clean EEG data is widespread, the field could greatly profit from automated solutions based on Machine Learning methods. Existing ICA-based removal strategies depend on explicit recordings of an individual's artifacts or have not been shown to reliably identify muscle artifacts.

Methods: We propose an automatic method for the classification of general artifactual source components. They are estimated by TDSEP, an ICA method that takes temporal correlations into account. The linear classifier is based on an optimized feature subset determined by a Linear Programming Machine (LPM). The subset is composed of features from the frequency-, the spatial- and temporal domain. A subject independent classifier was trained on 640 TDSEP components (reaction time (RT) study, n = 12) that were hand labeled by experts as artifactual or brain sources and tested on 1080 new components of RT data of the same study. Generalization was tested on new data from two studies (auditory Event Related Potential (ERP) paradigm, n = 18; motor imagery BCI paradigm, n = 80) that used data with different channel setups and from new subjects.

Results: Based on six features only, the optimized linear classifier performed on level with the inter-expert disagreement (<10% Mean Squared Error (MSE)) on the RT data. On data of the auditory ERP study, the same pre-calculated classifier generalized well and achieved 15% MSE. On data of the motor imagery paradigm, we demonstrate that the discriminant information used for BCI is preserved when removing up to 60% of the most artifactual source components.

Conclusions: We propose a universal and efficient classifier of ICA components for the subject independent removal of artifacts from EEG data. Based on linear methods, it is applicable for different electrode placements and supports the introspection of results. Trained on expert ratings of large data sets, it is not restricted to the detection of eye- and muscle artifacts. Its performance and generalization ability is demonstrated on data of different EEG studies.

Journal of Neuroscience Methods 250 (2015) 47–63

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

Research article

A practical guide to the selection of independent components of the electroencephalogram for artifact correction

Maximilien Chaumon[a,b,*], Dorothy V.M. Bishop[c], Niko A. Busch[a,b]

[a] Berlin School of Mind and Brain, Luisenstraße 56, 10117 Berlin, Germany
[b] Institute of Medical Psychology, Charité University Medicine, Luisenstraße 57, 10117 Berlin, Germany
[c] Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK

ARTICLE INFO

Article history:
Received 2 July 2014
Received in revised form 18 February 2015
Accepted 19 February 2015
Available online 16 March 2015

Keywords:
EEG
ICA
Artifact
Pre-processing
EEGLAB plugin

ABSTRACT

Background: Electroencephalographic data are easily contaminated by signals of non-neural origin. Independent component analysis (ICA) can help correct EEG data for such artifacts. Artifact independent components (ICs) can be identified by experts via visual inspection. But artifact features are sometimes ambiguous or difficult to notice, and even experts may disagree about how to categorise a particular component. It is therefore important to inform users on artifact properties, and give them the opportunity to intervene.
New Method: Here we first describe artifacts captured by ICA. We review current methods to automatically select artifactual components for rejection, and introduce the SASICA software, implementing several novel selection algorithms as well as two previously described automated methods (ADJUST, Mognon et al. Psychophysiology 2011;48(2):229; and FASTER, Nolan et al. J Neurosci Methods 2010;48(1):152).
Results: We evaluate these algorithms by comparing selections suggested by SASICA and other methods to manual rejections by experts. The results show that these methods can inform observers to improve rejections. However, no automated method can accurately isolate artifacts without supervision. The comprehensive and interactive plots produced by SASICA therefore constitute a helpful guide for human users for making final decisions.
Conclusions: Rejecting ICs before EEG data analysis unavoidably requires some level of supervision. SASICA offers observers detailed information to guide selection of artifact ICs. Because it uses quantitative parameters and thresholds, it improves objectivity and reproducibility in reporting pre-processing procedures. SASICA is also a didactic tool that allows users to quickly understand what signal features captured by ICs make them likely to reflect artifacts.

Table 1
Measures computed by the three automated tools evaluated here. Abbreviations refer to those used in figures and throughout the paper.

Tool	Artifact type	Measure	Abbreviation
SASICA	Blinks/vertical eye movements	Correlation with vertical EOG electrodes	CorrV
	Horizontal eye movements	Correlation with horizontal EOG electrodes	CorrH
	Muscle	Low autocorrelation of time-course	LoAC or AutoCorr
	Bad channel	Focal channel topography	FocCh
	Rare event	Focal trial activity	FocTr
	Non dipolar component	Residual variance	ResVar
	Bad channel	Correlation with Bad channel	CorrCh
FASTER	Eye blinks/saccades	Correlation with EOG electrodes	EOGcorr
	"Pop-Off"	Spatial Kurtosis	SK
	White noise	Slope of the power spectrum	SpecSl
	White noise	Hurst exponent	HE
	White noise	Median slope of time-course	MedGrad
ADJUST	Eye blinks	Temporal Kurtosis	TK
	Eye blinks	Spatial average difference	SAD
	Eye blinks	Spatial variance difference	SVD
	Vertical Eye Movements	Maximum epoch variance	MEV
	Horizontal Eye Movements	Spatial eye difference	SED
	Generic Discontinuities	Generic discontinuity spatial feature	GDSF

Chaumon et al., 2015

Neural components

A

Expected properties

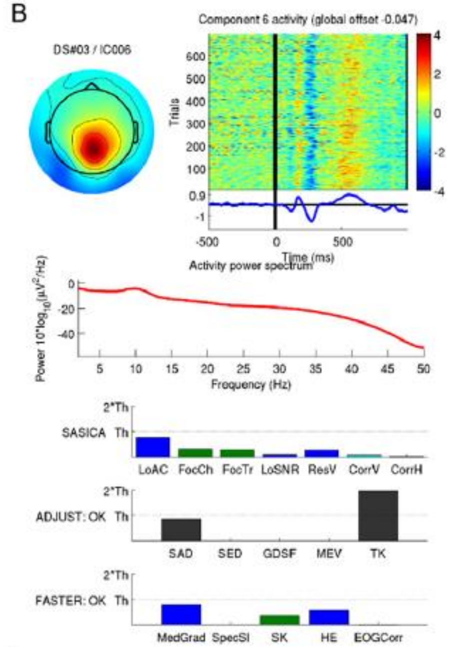
Smooth/dipolar topography

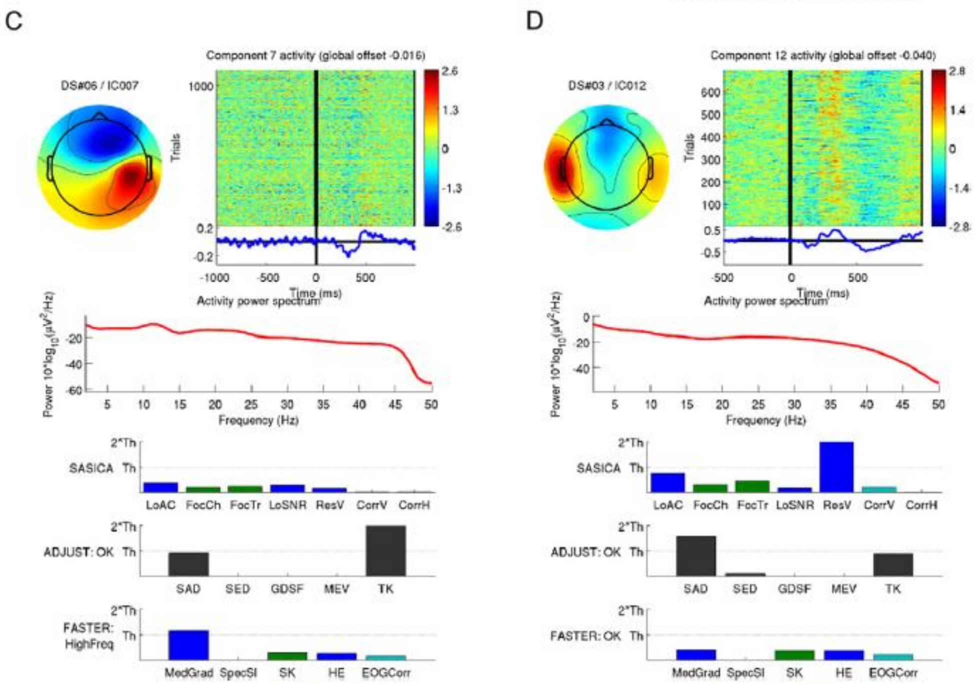
Large amplitude

Strong evoked activity

Power peak at physiological frequency

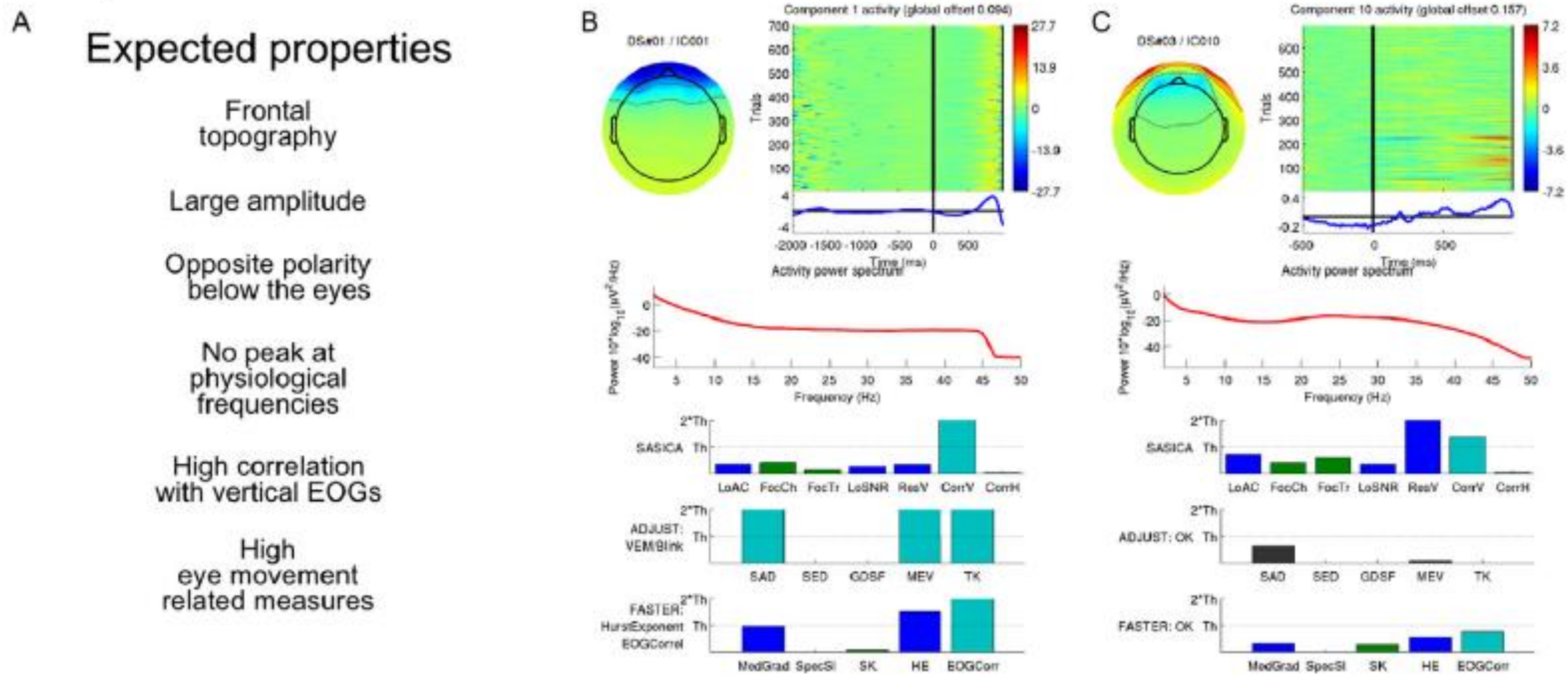
Low artefact measures





Chaumon et al., 2015

Blink components



Chaumon et al., 2015

Horizontal eye movement components

D

Expected properties

Opposite sign bilateral frontal topography

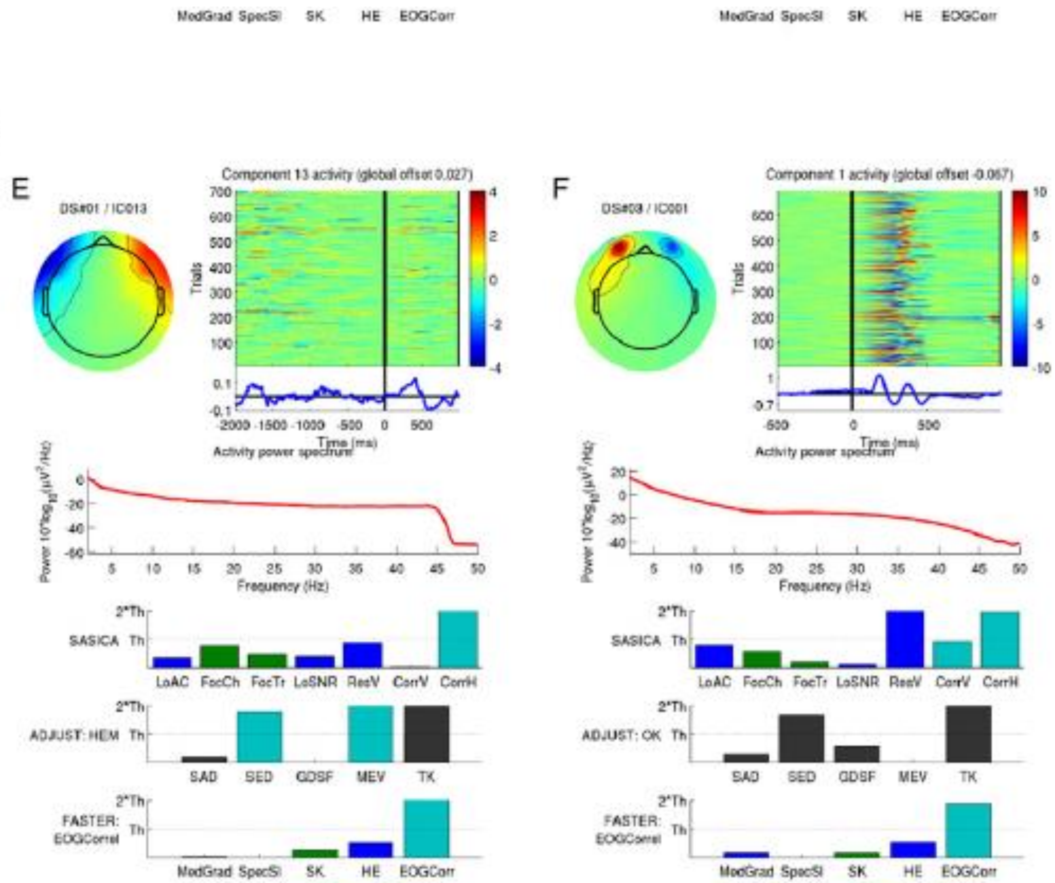
Step-like events

Opposite polarity around the eyes

No peak at physiological frequencies

High correlation with vertical/horizontal EOGs

High eye movement related measures



Chaumon et al., 2015

Non-artifact components may be mistaken for ocular components

G

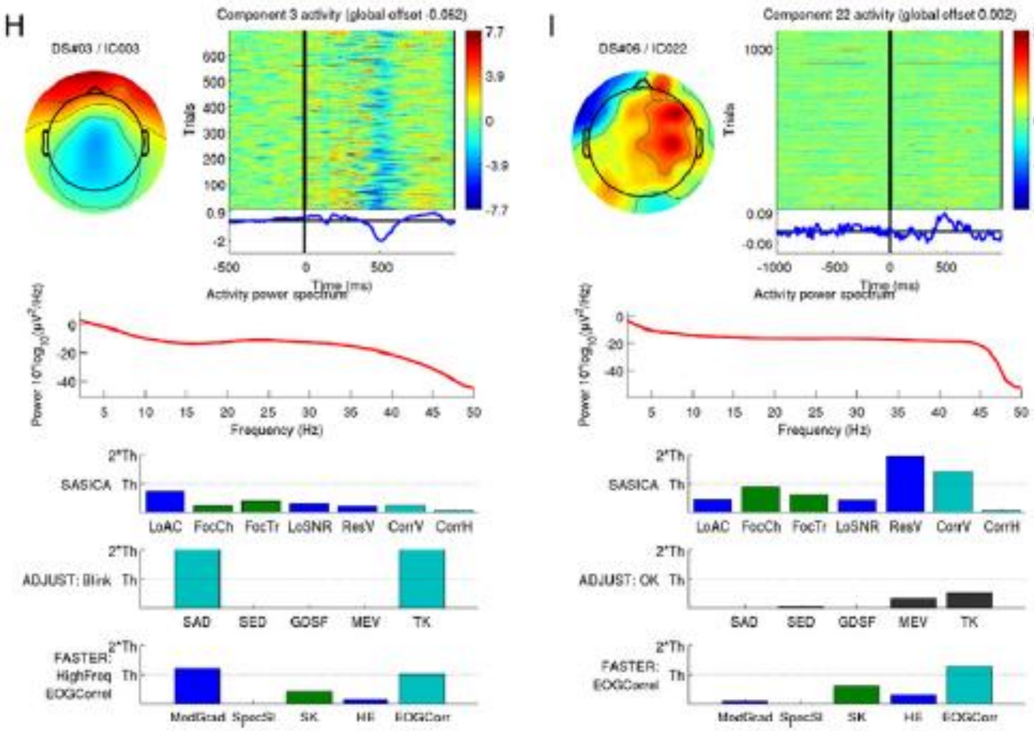
Expected properties

Inverse weight at posterior channels

Noisy time course

No opposite polarity around the eyes

Weak correlation with EOGs



Chaumon et al., 2015

A

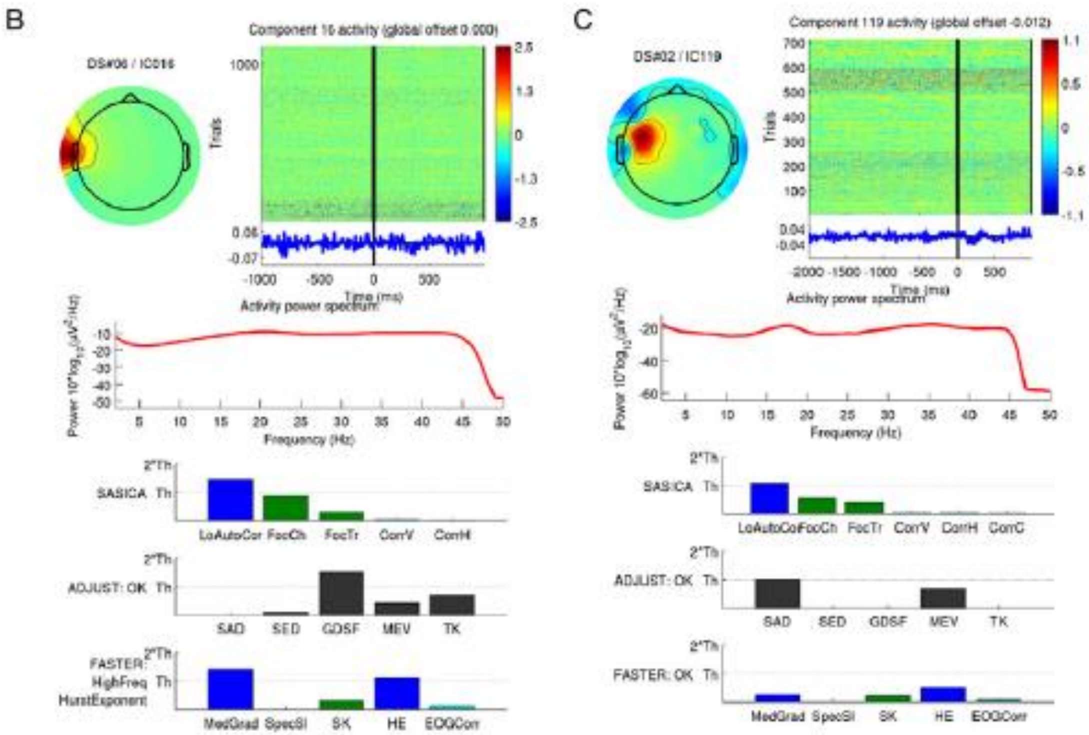
Muscle components
Expected properties

Focal topography

Steady noisy time courses dissipating / building up across trials

Power at high frequencies

High noise measures



Chaumon et al., 2015

Other types of artifacts may be mistaken for muscle components

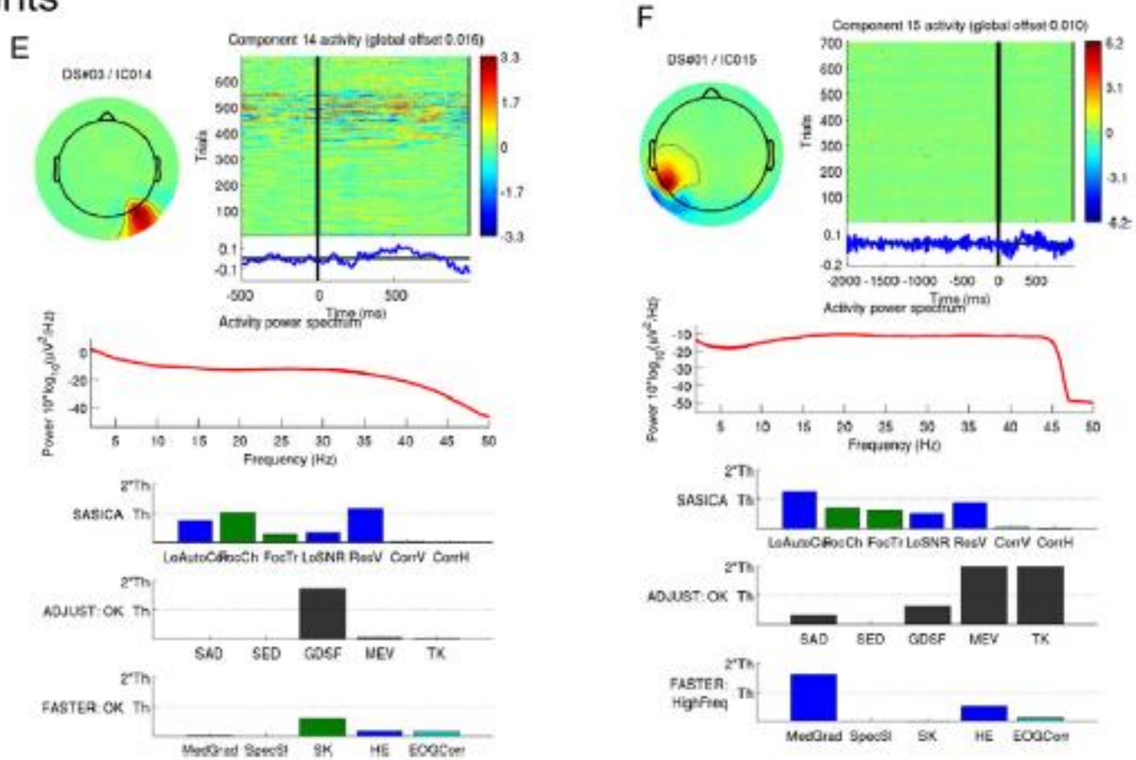
D

Expected properties

Irregular/patchy topography

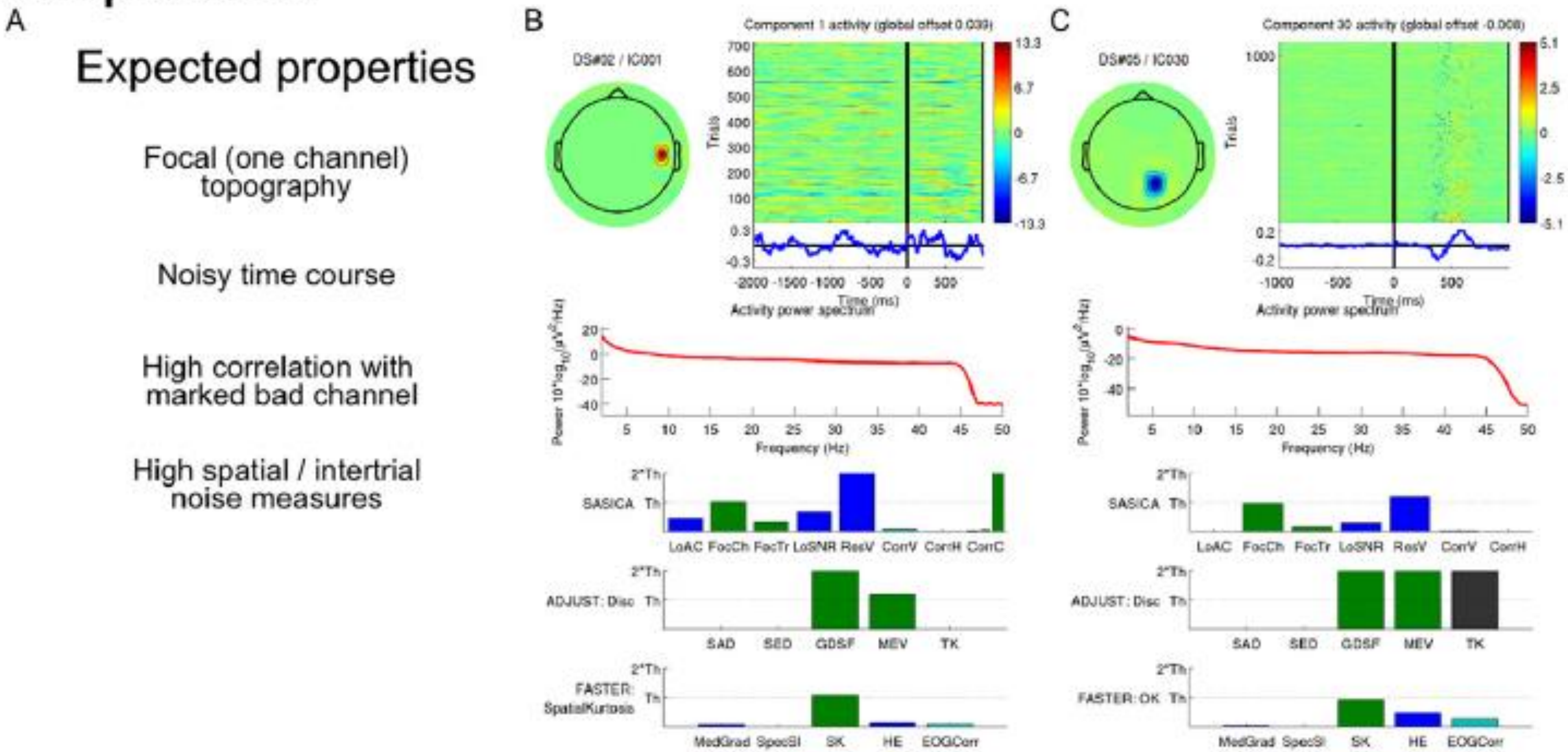
Irregular / low frequency noise

Stimulus evoked response



Chaumon et al., 2015

Bad Channel components



Chaumon et al., 2015

Ambiguous mixture components

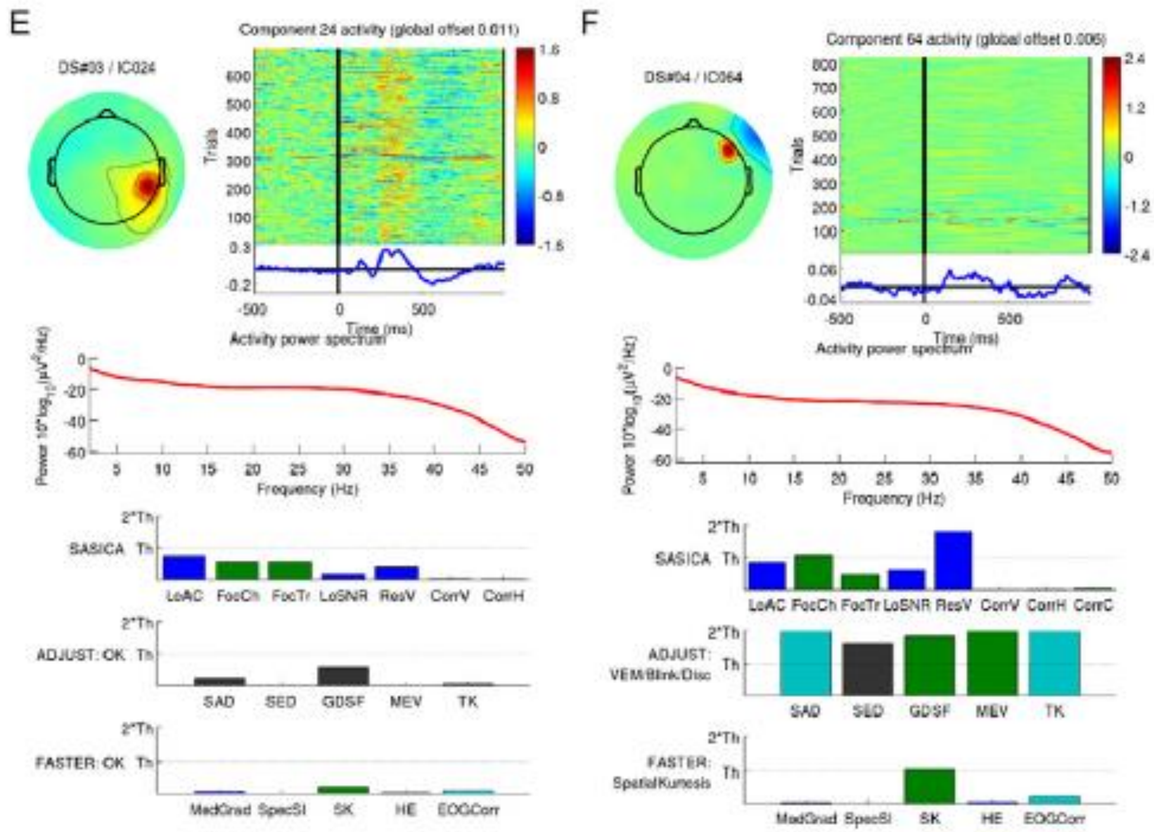
D

Expected properties

More spread-out topography

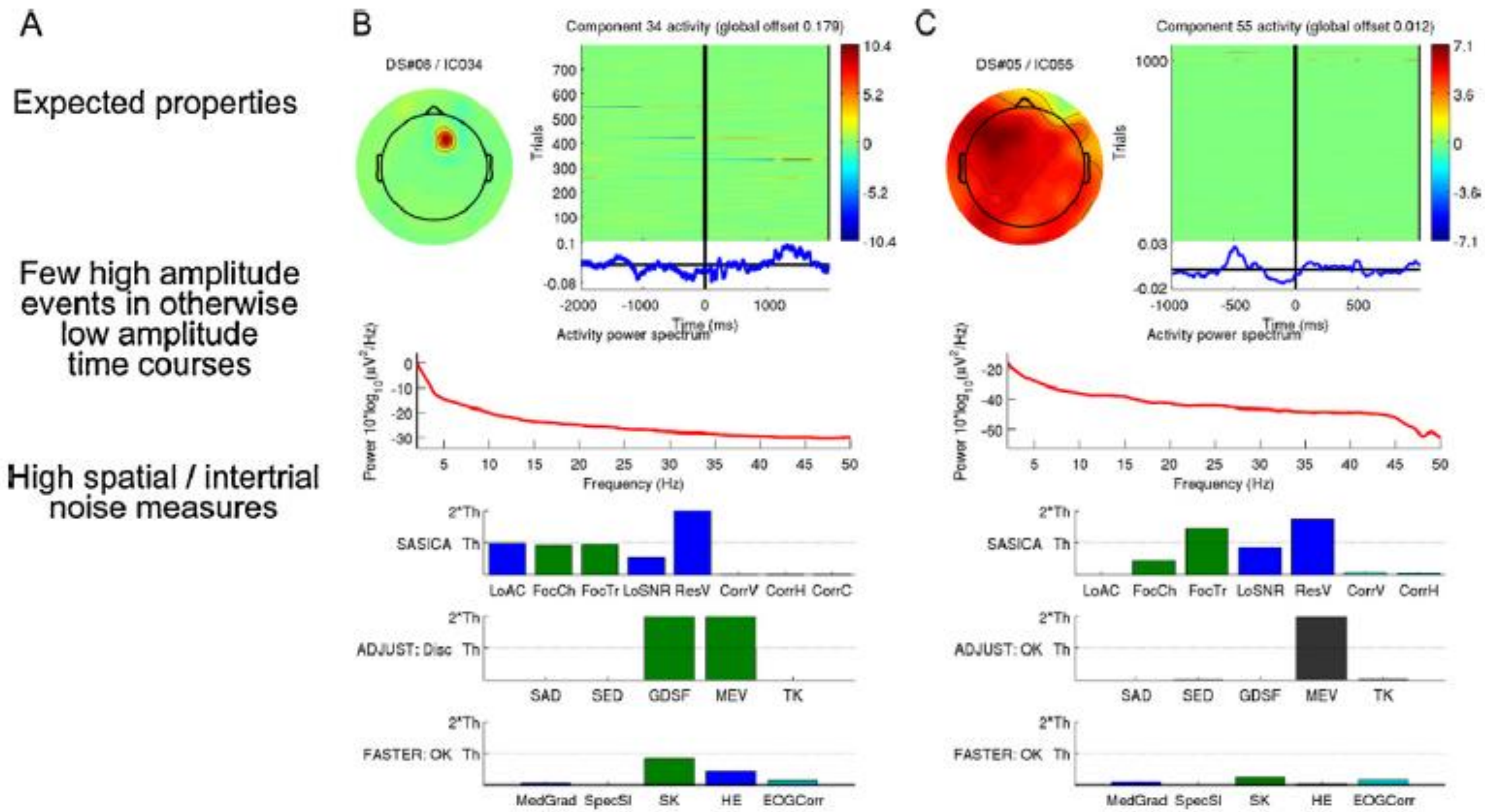
Stimulus evoked response

Transient noise activity



Chaumon et al., 2015

Rare Events



Chaumon et al., 2015

What is ADJUST?

ADJUST= Automatic EEG artifact **D**etection based on the **J**oint **U**se of **S**patial and **T**emporal features

- Automatic ICA-based algorithm that identifies artifact-related IC components
- Uses both spatial and temporal distributions
- Combines stereotyped features to efficiently and systematically reject an artifact

Mognon, Jovicich, Bruzzone, & Buiatti, 2010

How does it work?

- EEG is decomposed into ICs (done in EEGlab)
  - ICs defined only by statistical relationships.
  - It knows nothing about where electrodes are
- Detectors are applied for 4 types of artifacts
  - Computes class-specific spatial and temporal features on all ICs
  - Each feature has a threshold dividing artifacts from non-artifacts
- For each detector, ICs identified as artifacts if features associated with the artifact exceed their respective threshold.

Mognon, Jovicich, Bruzzone, & Buiatti, 2010

Artifacts

- Eye blinks
- Vertical Eye Movement
- Horizontal Eye Movement
- Generic Discontinuity

Features

- Spatial Average Difference (SAD)
- Temporal Kurtosis (TK)
- Maximum Epoch Variance (MEV)
- Spatial Eye Difference (SED)
- Generic Discontinuities Spatial Feature (GDSF)

MARA: Six features derived from machine learning algorithm

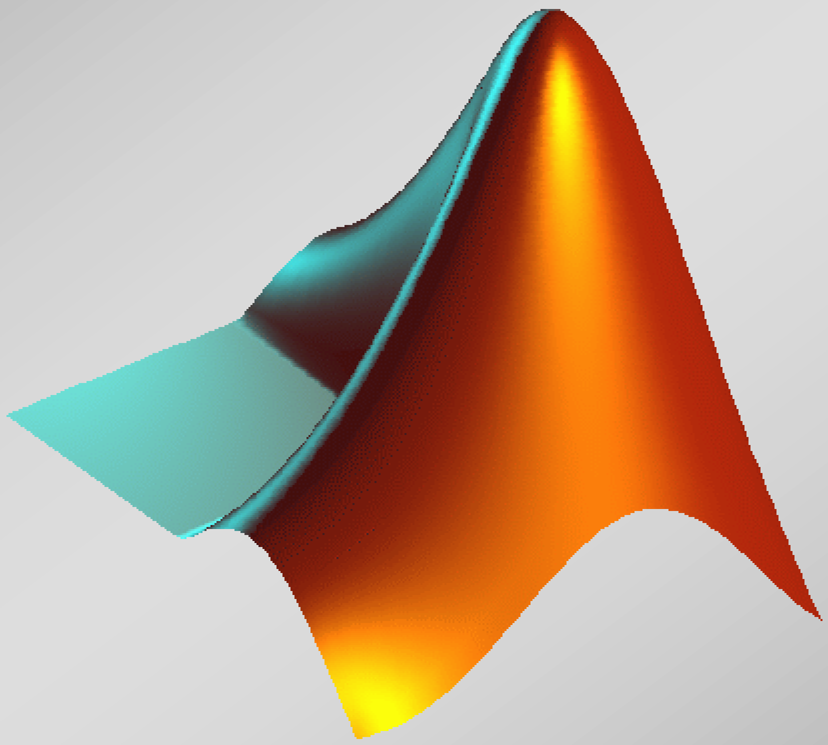
1. **Current Density Norm**: ICA scalp maps can be interpreted as EEG potentials for which the location of the sources can be estimated. It is natural that artifactual signals originating outside the brain can only be modeled by rather complicated sources. Those are characterized by a large l2-norm, which MARA uses as a feature.
2. **Range Within Pattern**: The logarithm of the difference between the minimal and the maximal activation in a scalp map. Spatially localized scalp maps stemming from e.g muscle artifacts or loose electrodes are typically characterized by a high Range Within Pattern.
3. **Mean Local Skewness**: The mean absolute local skewness of time intervals of 15s duration. This feature aims to detect outliers in the time series.
4. **λ and Fit Error**: These two features describe the deviation of a component's spectrum from a prototypical 1/f curve and its shape.
5. **8-13 Hz**: The average log band power of the α band (8–13Hz). This feature aims to detect the typical α peak in components of neural origin.

Winkler et al., 2011

Pre-Processing Steps for ICA Artifact Rejection

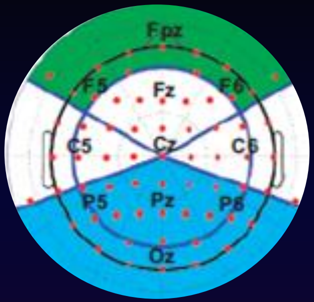
1. **A rough pre-cleaning of the data** by e.g. channel rejection and trial rejection may be performed. This step is usually helpful for obtaining a good ICA decomposition.
2. **Filter:** As ICA decomposition is known to be sensitive to slow drifts, high pass filtering the data (at 0.5 Hz or even 2 Hz) can sometimes improve the quality of the decomposition. Note that MARA might lead to suboptimal results on narrow band-passed filtered data, because its spectrum features are calculated on the power spectrum between 2 Hz and 39 Hz.
3. **Run ICA!** Tools>Run ICA calculates ICA decomposition: The option 'pca' can be set to perform a dimensionaliy reduction prior to IC computation. Such a step may be helpful, in order to reduces the noise level and avoid an unnatural splitting of sources. (It also makes IC computation faster and reduces and the number of components that have to be labeled.)

Winkler et al., 2011

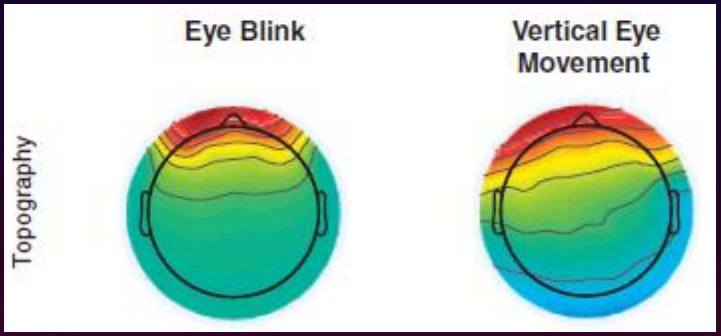


The following slides were not used, but may be helpful, and derive from Laura Zambrano-Vazquez’s talk about ADJUST

Features



- Spatial Average Difference (SAD)
  - Spatial topography of blink ICs
  - Looks for higher amplitude in frontal vs. posterior areas

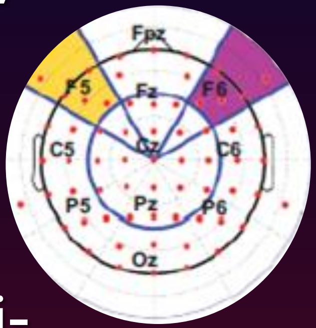


- Temporal Kurtosis (TK)
  - Kurtosis over the IC time course
    - Kurtosis is "peakedness" of the distribution (i.e. distribution of timepoints in the epoch)
  - Looks for outliers in amplitude distribution typical of blinks

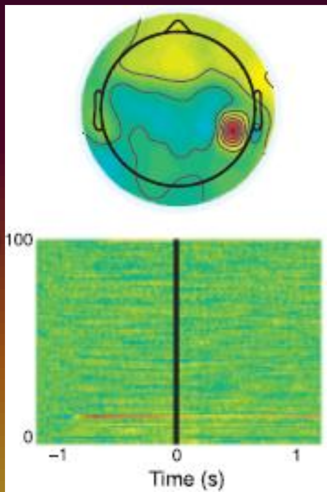
Mognon, Jovicich, Bruzzone, & Buiatti, 2010

Features

- Maximum Epoch Variance (MEV)
  - Is a ratio of variance in epoch with most variance compared to mean variance over all epochs
  - Looks for slower fluctuations typical of vertical eye movement



- Spatial Eye Difference (SED)
  - Looks for large amplitudes in frontal areas in anti-phase typical of horizontal eye movement
- Generic Discontinuities Spatial Feature (GDSF)
  - Looks for local spatial discontinuities

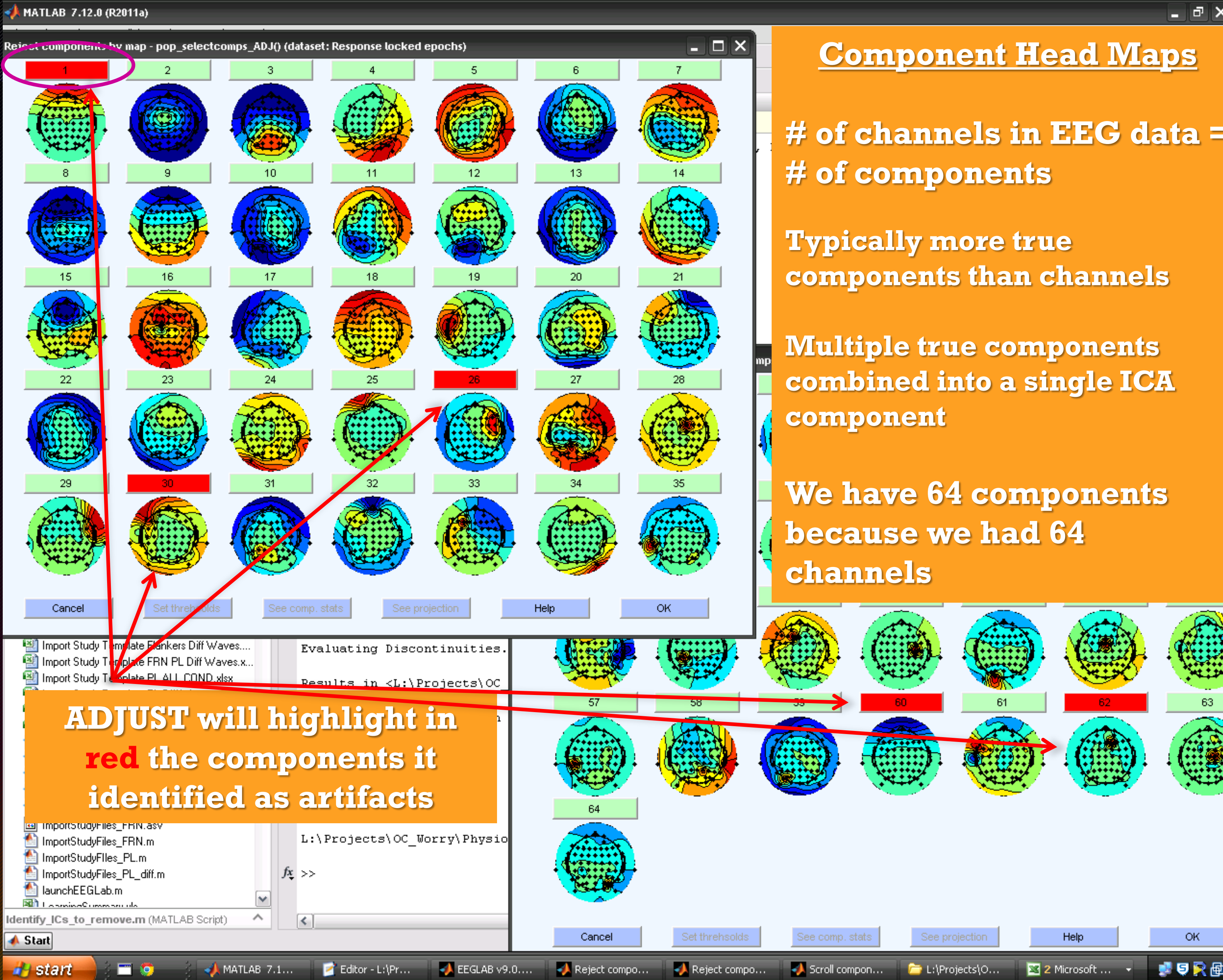


Mognon, Jovicich, Bruzzone, & Buiatti, 2010

Where to begin?

- Pre-processing
  - Clean file for non-stereotyped "gross artifacts"
    - AKA Muscle activity and other external factors
    - Variable spatial distribution that could take up a lot of ICs
    - Low-Pass filtering, if appropriate for your data, can remove some artifacts and prevent so many ICs from capturing these higher-frequency noise artifacts
- ICA
  - From EEGlab or, if not performed already, can be called from ADJUST
- Run ICA'd files with ADJUST
  - Import dataset into EEGLab→ Tools → ADJUST OR
  - Use script and select file

ADJUST Environment



Component Head Maps

of channels in EEG data = # of components

Typically more true components than channels

Multiple true components combined into a single ICA component

We have 64 components because we had 64 channels

ADJUST will highlight in red the components it identified as artifacts

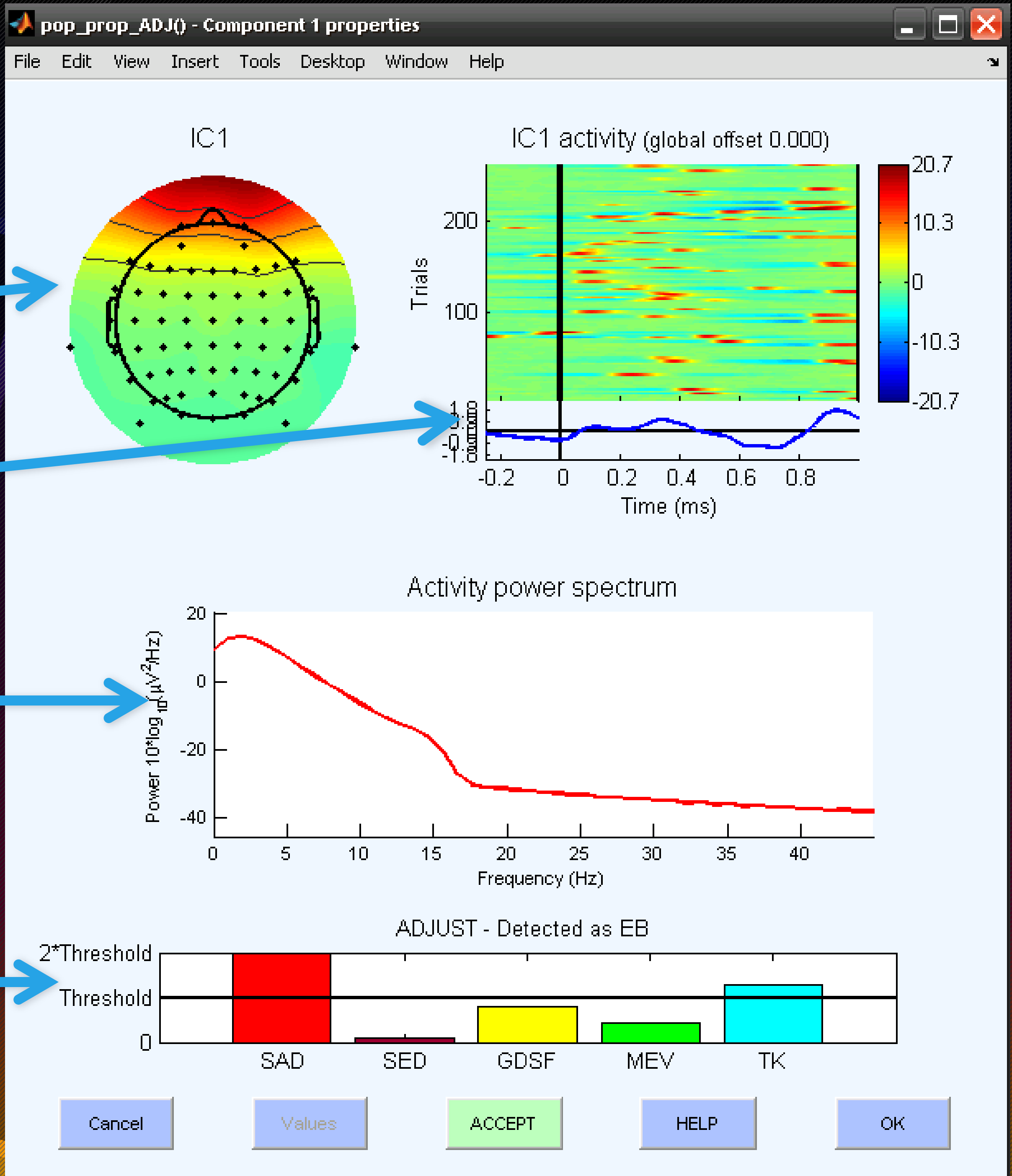
Looking at an individual IC

Head map

IC activity plotted against trial

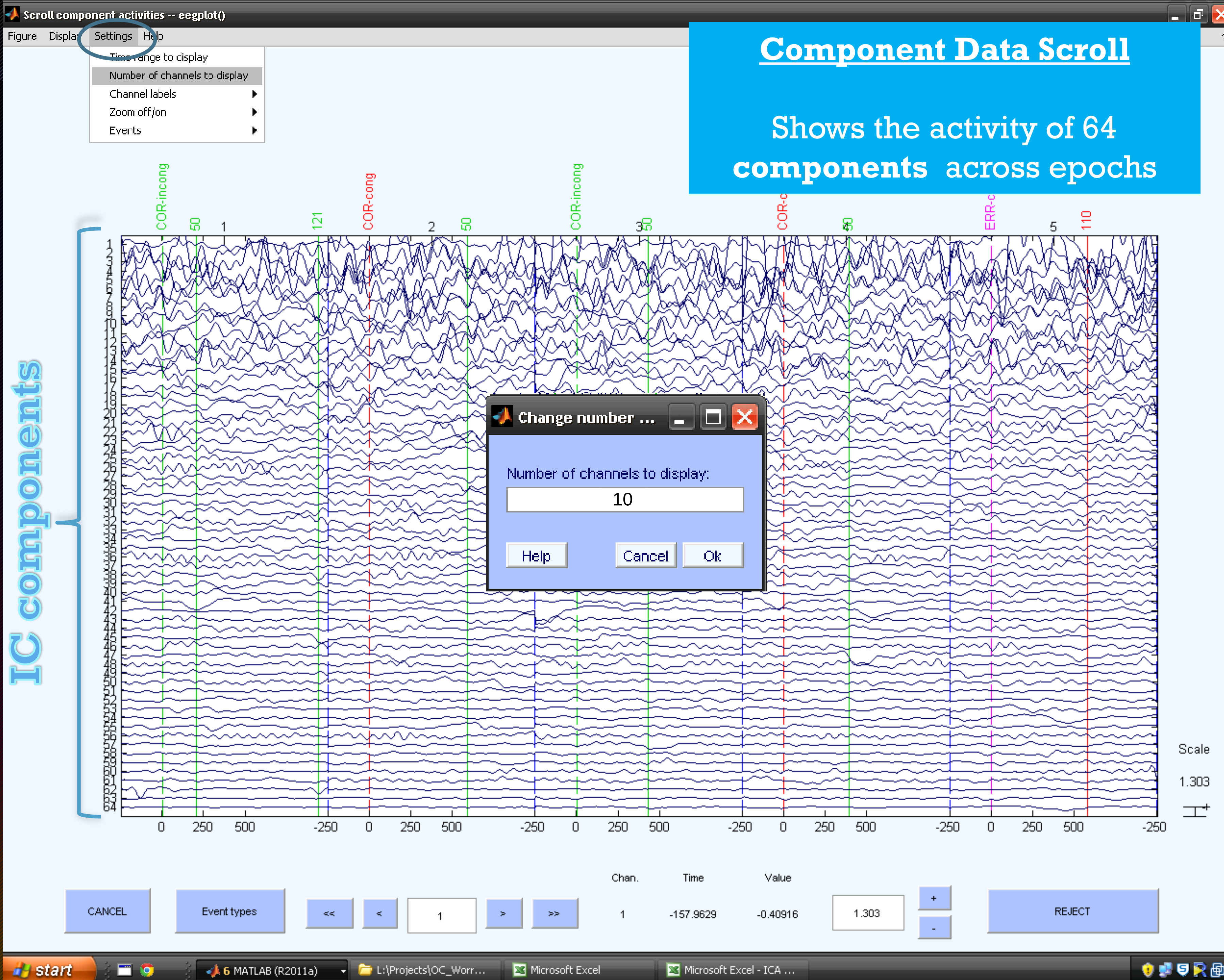
Activity power spectrum

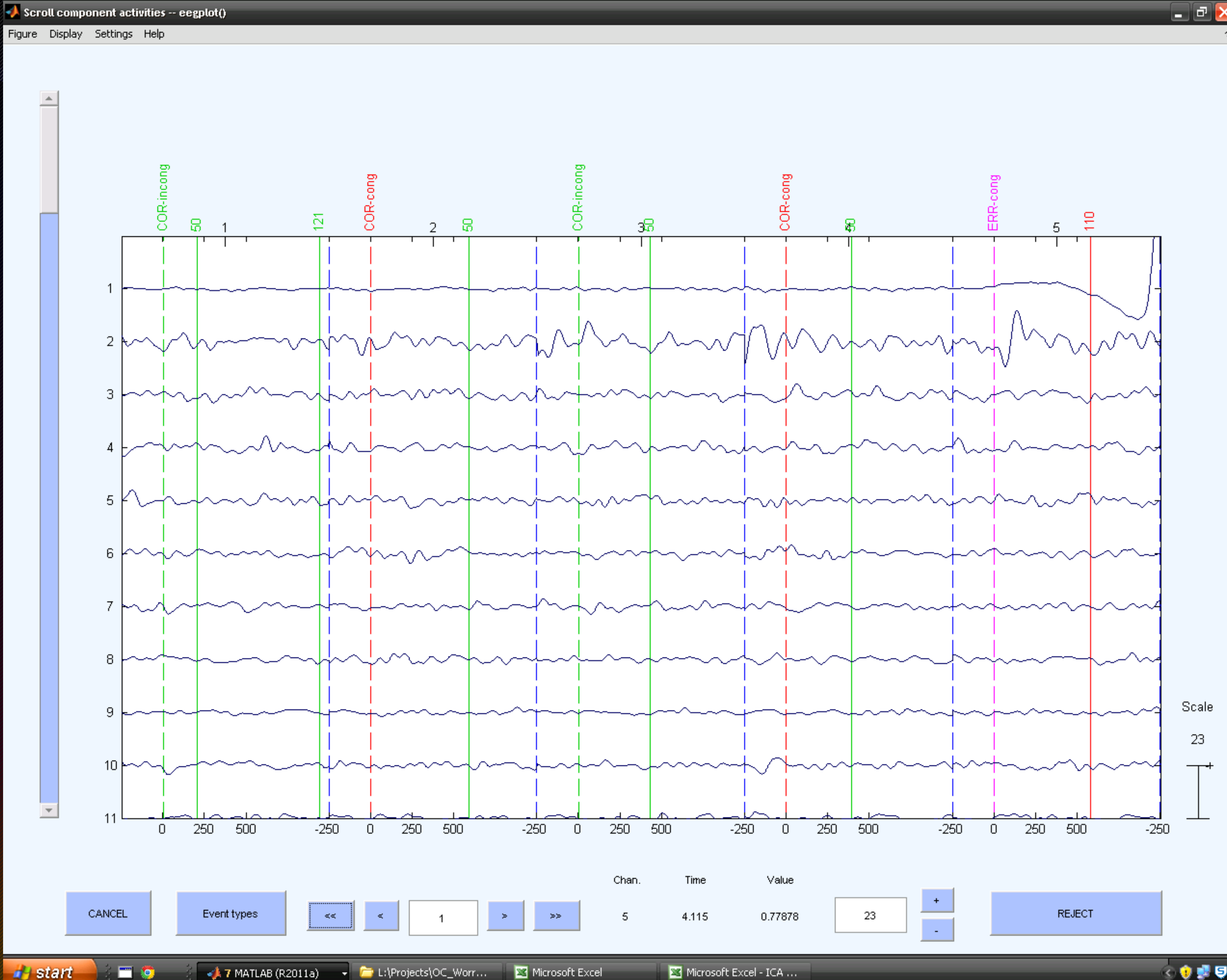
Features and thresholds



Component Data Scroll

Shows the activity of 64 **components** across epochs



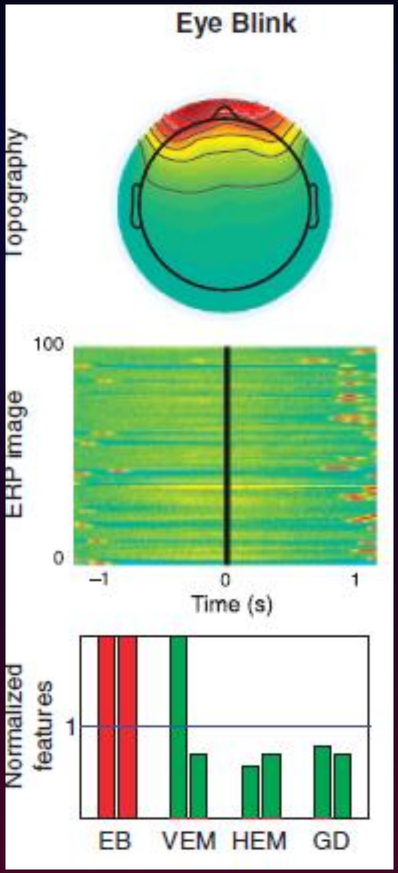

Scroll component activities -- eegplot()
Figure
Display
Settings
Help
COR-incong
50
121
COR-cong
50
COR-incong
50
COR-cong
50
ERR-cong
110
1
2
3
4
5
1
2
3
4
5
6
7
8
9
10
11
0
250
500
-250
Scale
23
Chan.
Time
Value
CANCEL
Event types
<<
<
1
>
>>
5
4.115
0.77878
23
+
-
REJECT
start
7 MATLAB (R2011a)
L:\Projects\OC_Worr...
Microsoft Excel
Microsoft Excel - ICA ...

Artifact Detection

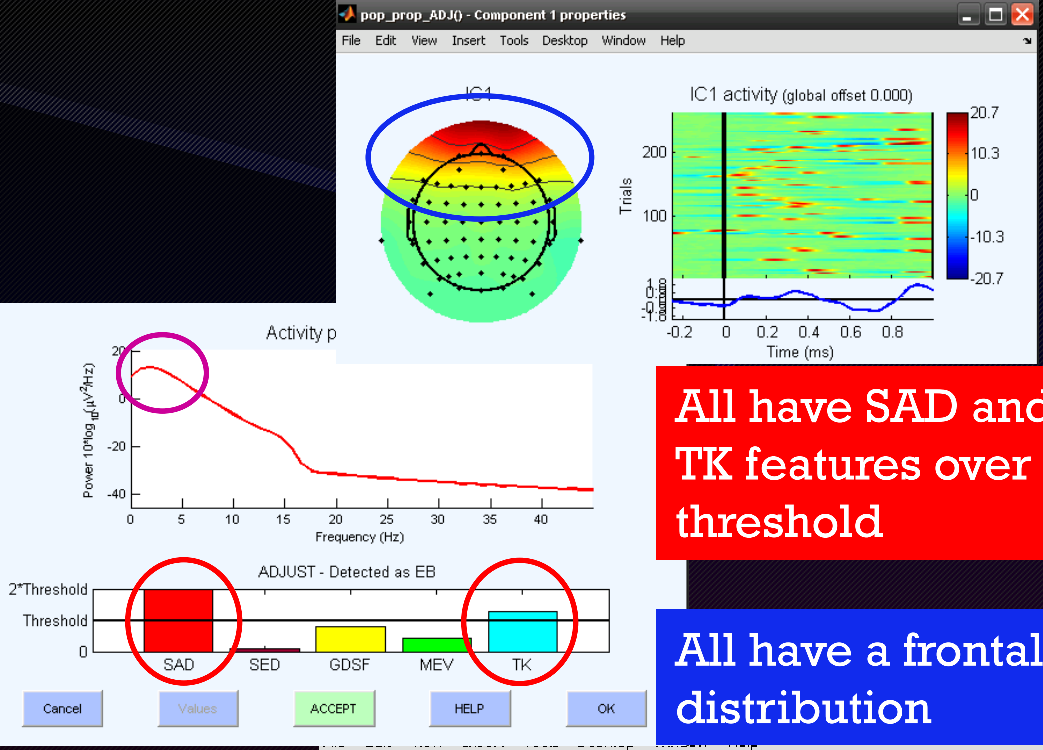
Things to look for

Eye blinks

- Features used
  - Spatial Average Difference (SAD)
  - Temporal Kurtosis (TK)
- Frontal distribution
- High power in delta frequency band

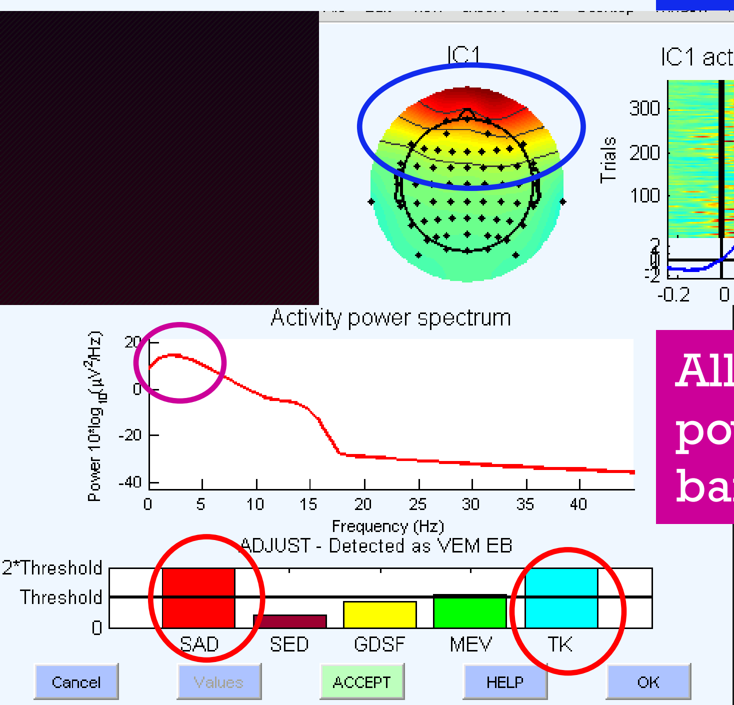


- In component data scroll high potentials with morphology of eye blink (like in EEG) can be observed

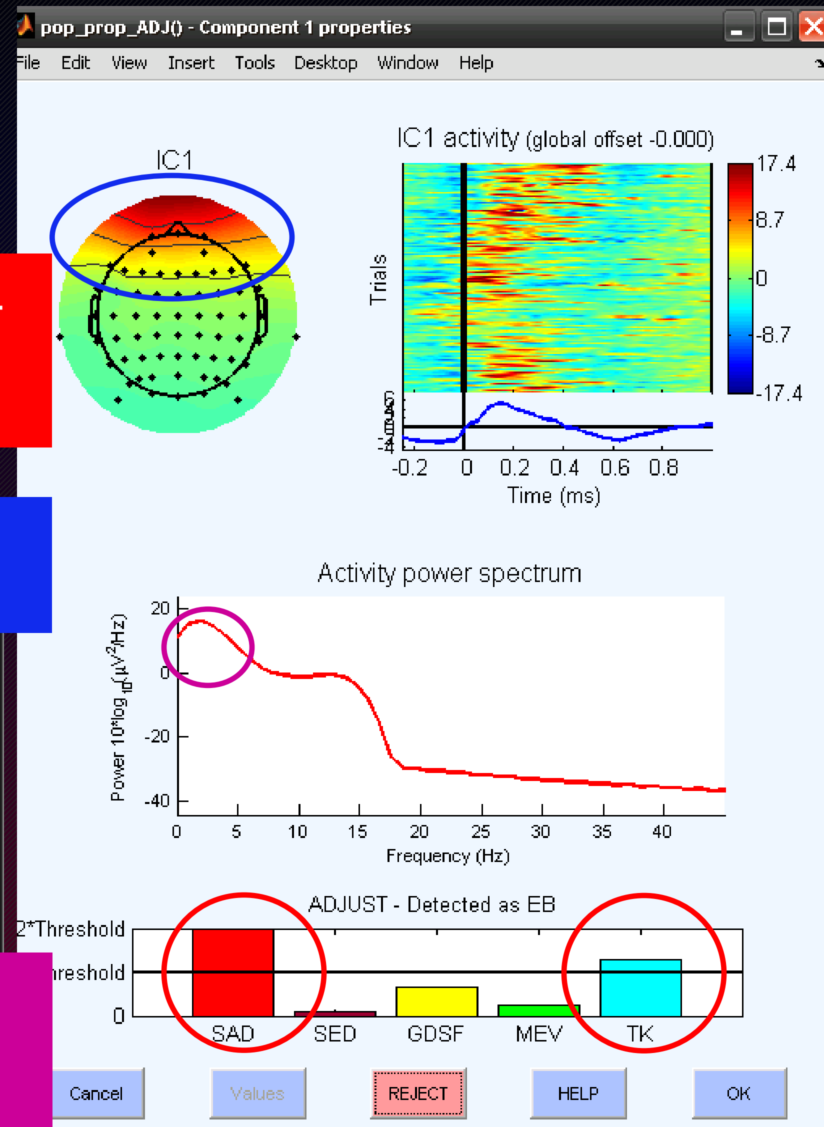


All have SAD and TK features over threshold

All have a frontal distribution

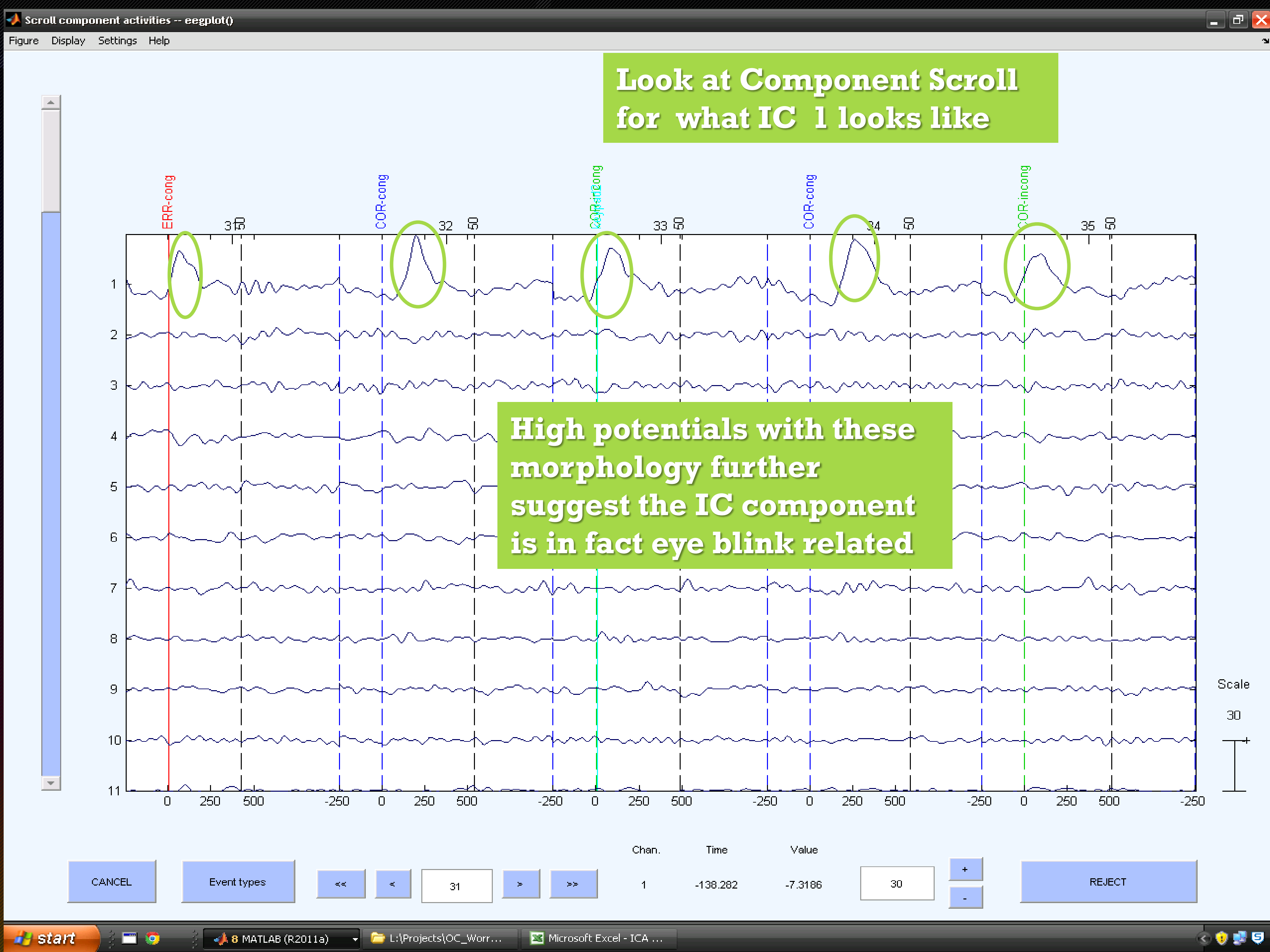


All have higher power in delta band



3 examples of eye blink artifacts

Look at Component Scroll for what IC 1 looks like

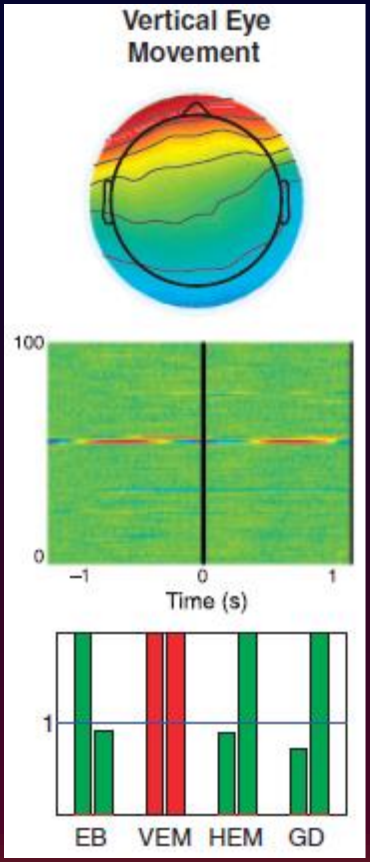


High potentials with these morphology further suggest the IC component is in fact eye blink related

Vertical Eye Movement

- Features used
  - Spatial Average Difference (SAD)
  - Maximum Epoch Variance (MEV)

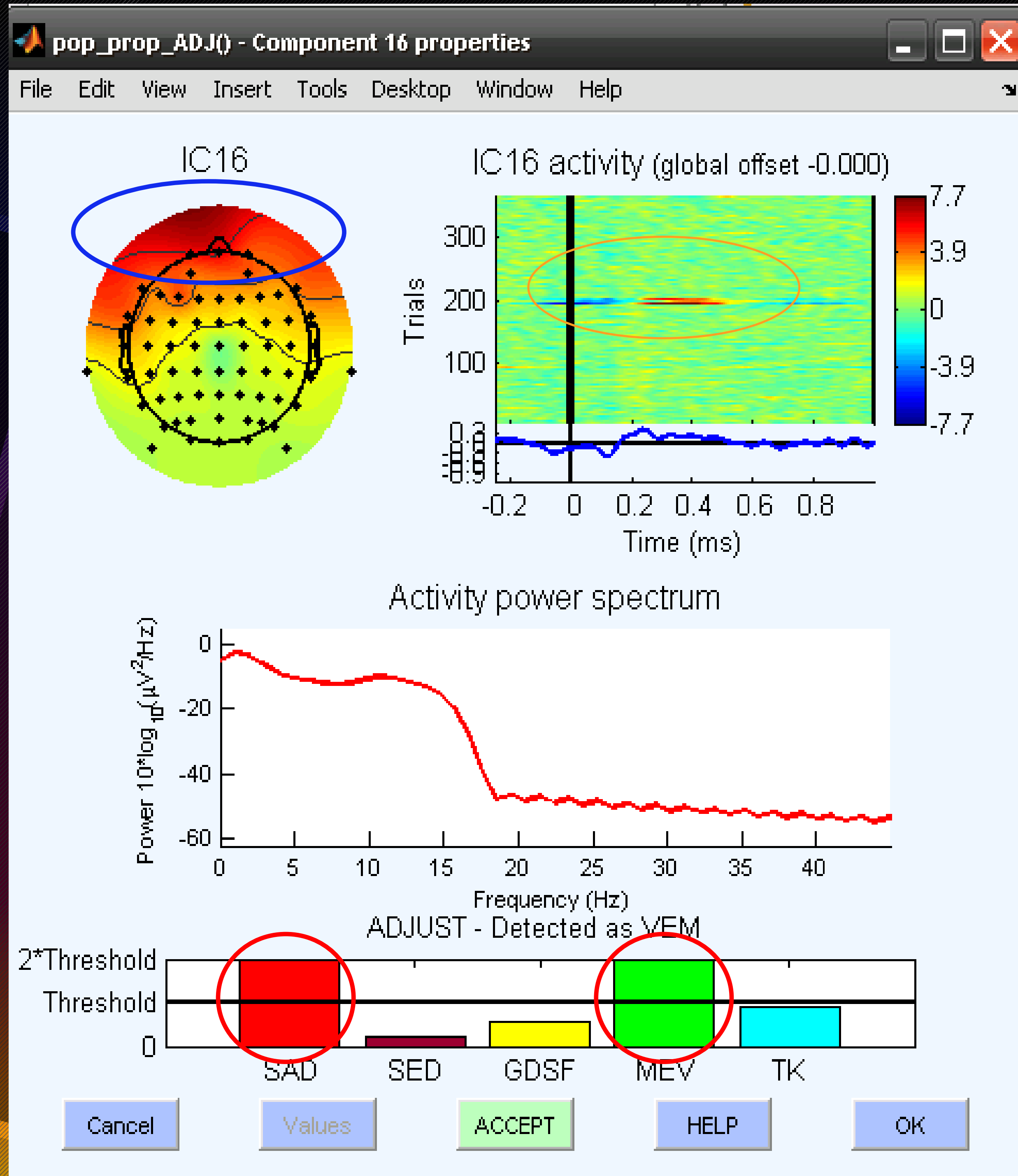
- Frontal distribution similar to that of an eye blink

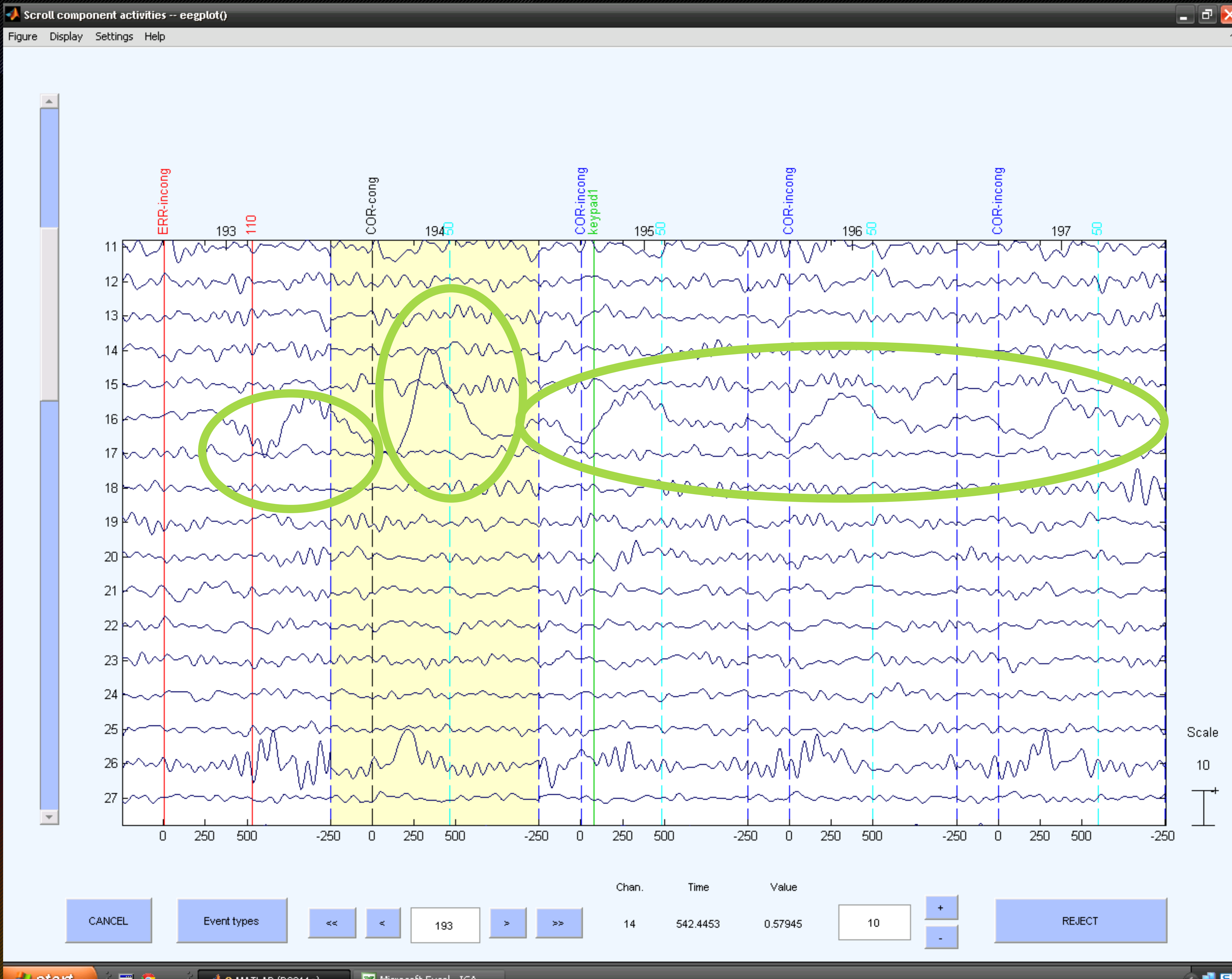


Frontal distribution

It appears that the artifact is mostly driven by what is happening around trial 200

SAD and MEV features are over threshold

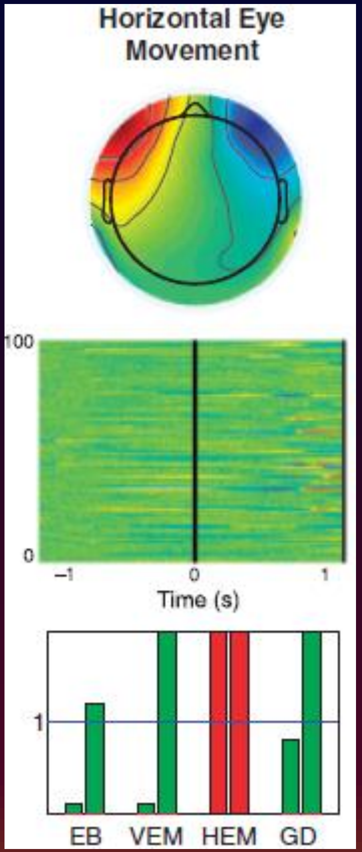


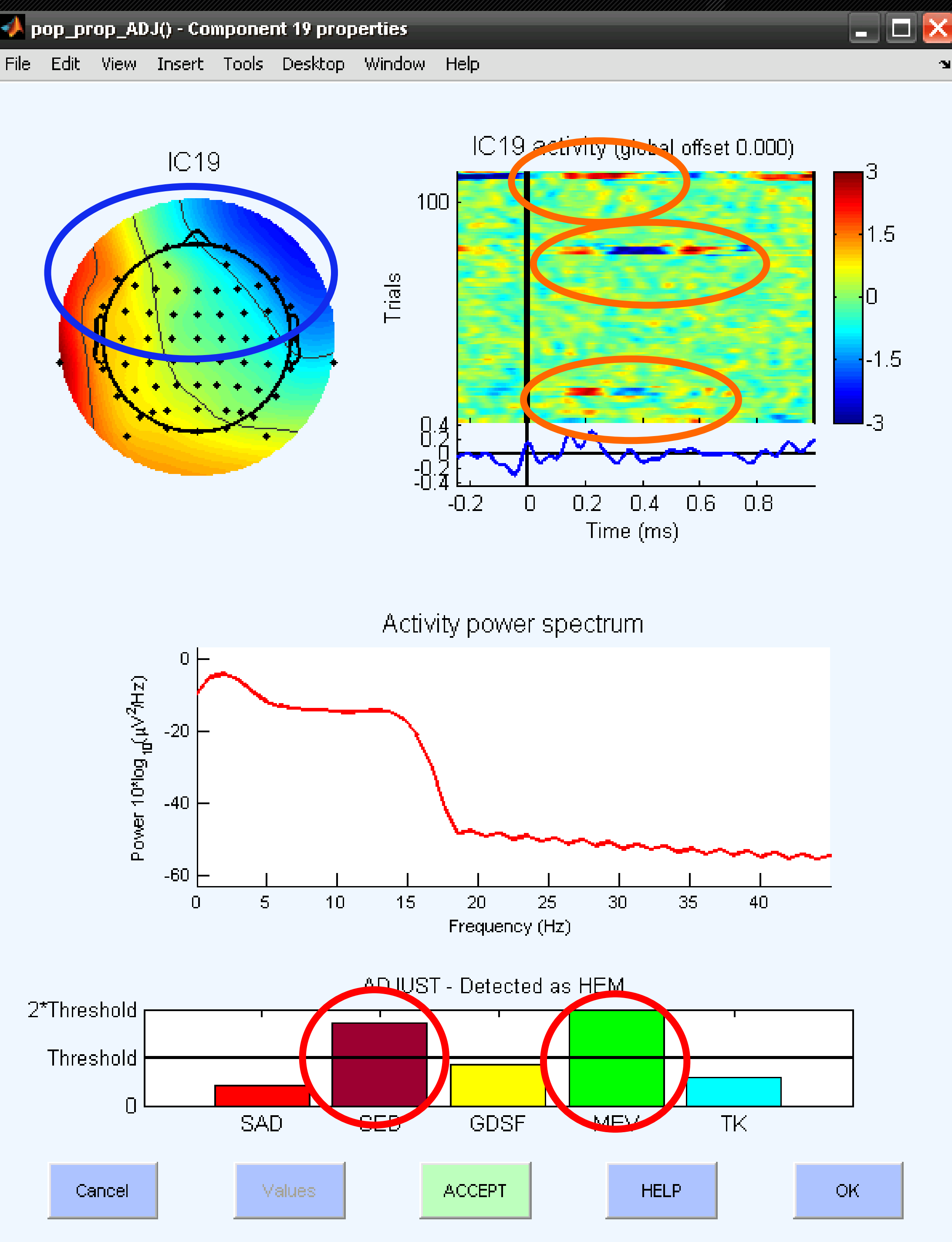

Scroll component activities -- eegplot()
Figure
Display
Settings
Help
ERR-incong
110
COR-cong
50
COR-incong
keypad1
50
COR-incong
50
COR-incong
50
193
194
195
196
197
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
0 250 500 -250 0 250 500 -250 0 250 500 -250 0 250 500 -250 0 250 500 -250
Scale
10
CANCEL
Event types
<<
<
193
>
>>
Chan.
14
Time
542.4453
Value
0.57945
10
+
-
REJECT
start
8 MATLAB (R2011a)
Microsoft Excel - ICA ...

Horizontal Eye Movement

- Features used
  - Spatial Eye Difference (SED)
  - Maximum Epoch Variance (MEV)

- Frontal distribution in anti-phase (one positive and one negative)



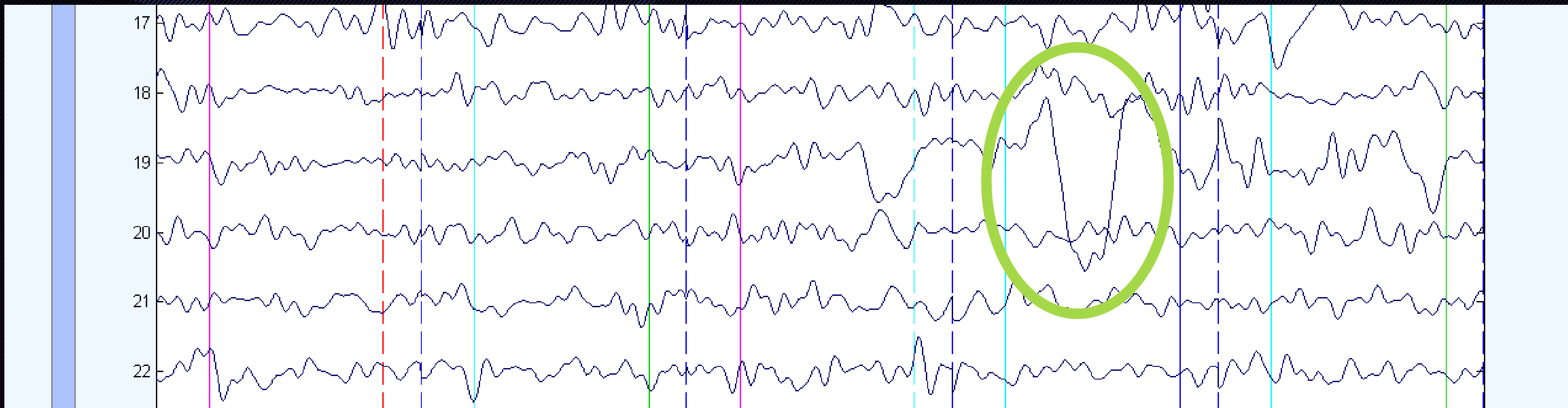


Anti-phase, primarily frontal distribution

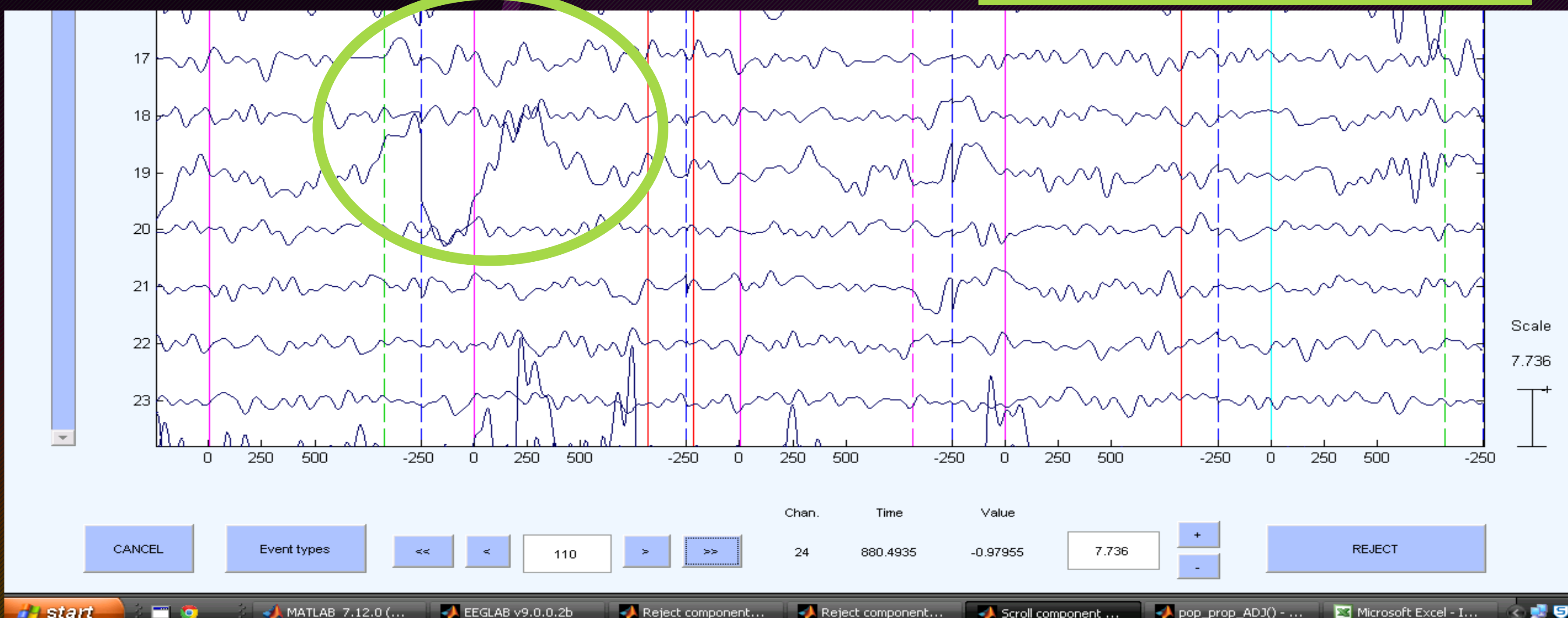
In the IC Activity by trial, three sections stand out

SED and MEV features are over threshold

Trial 75-79

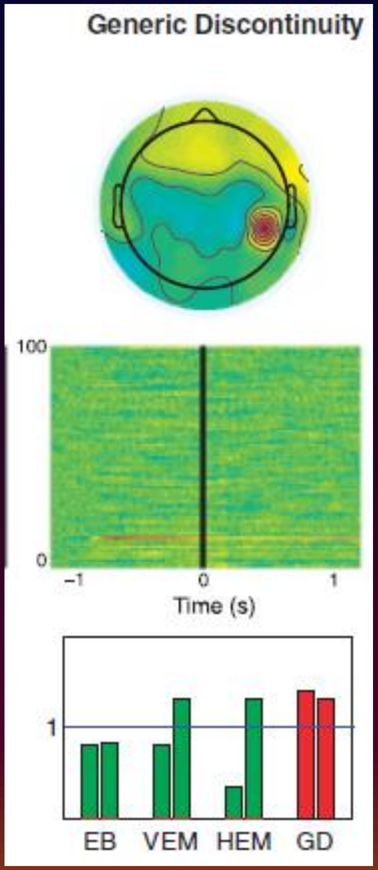


Trials 110-114



Generic Discontinuities

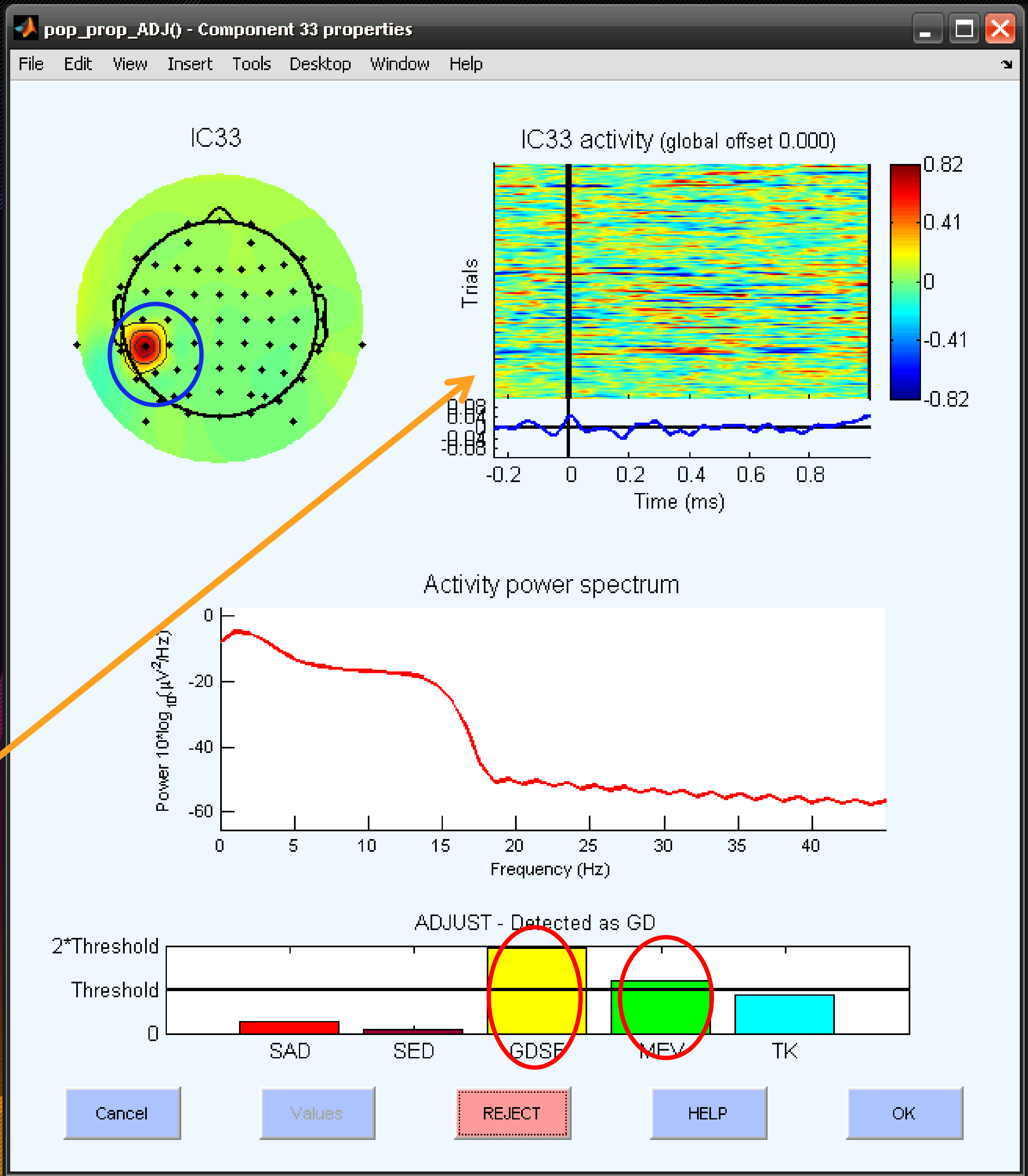
- Features used
  - Generic Discontinuities Spatial Feature (GDSF)
  - Maximum Epoch Variance (MEV)
- Variable distribution
- Sudden amplitude fluctuations with no spatial preference
  - Could be present in as little as one or 2 trials, and limited to 1 channel
- In component data scroll weird activity in the trial plotted on the IC activity

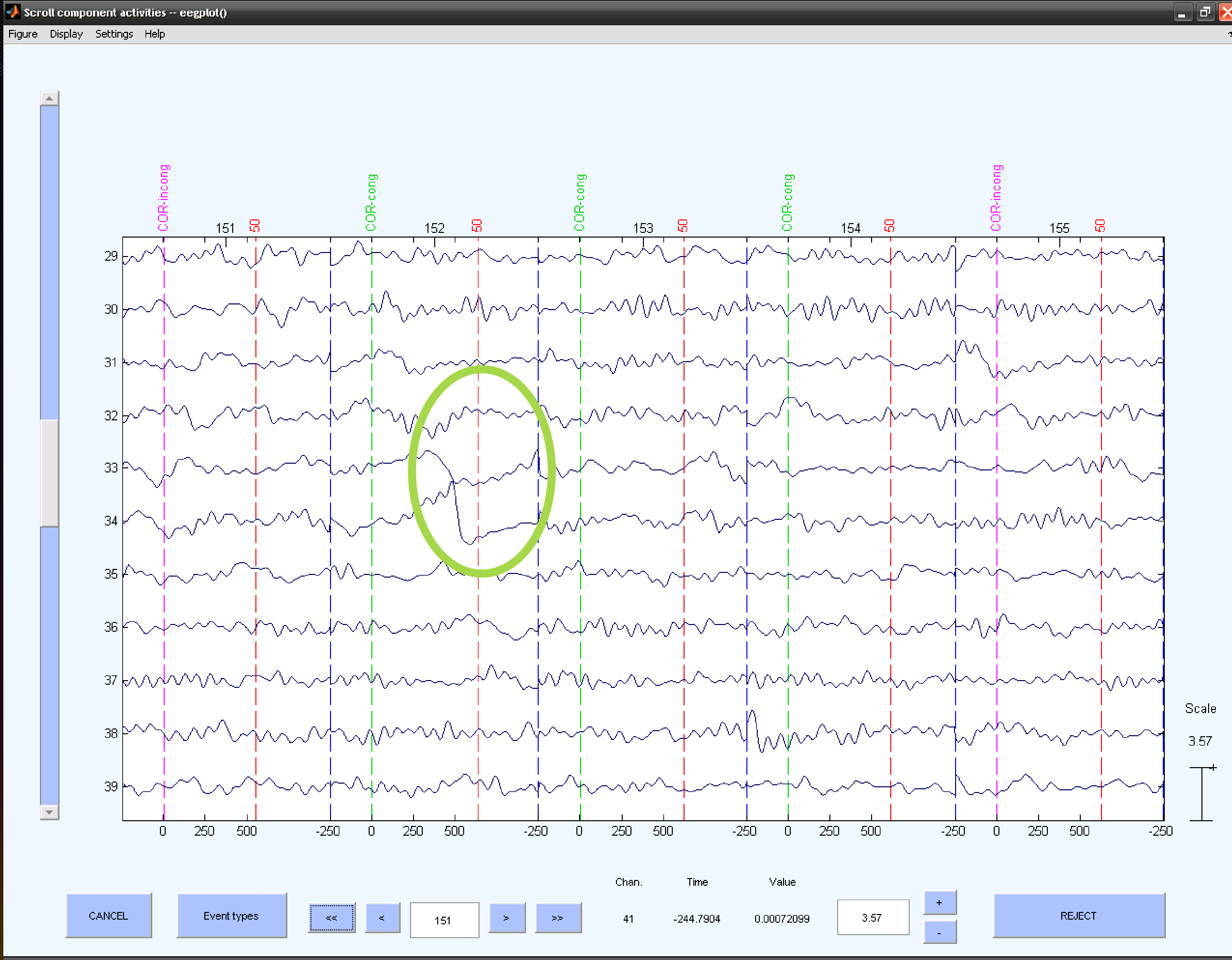


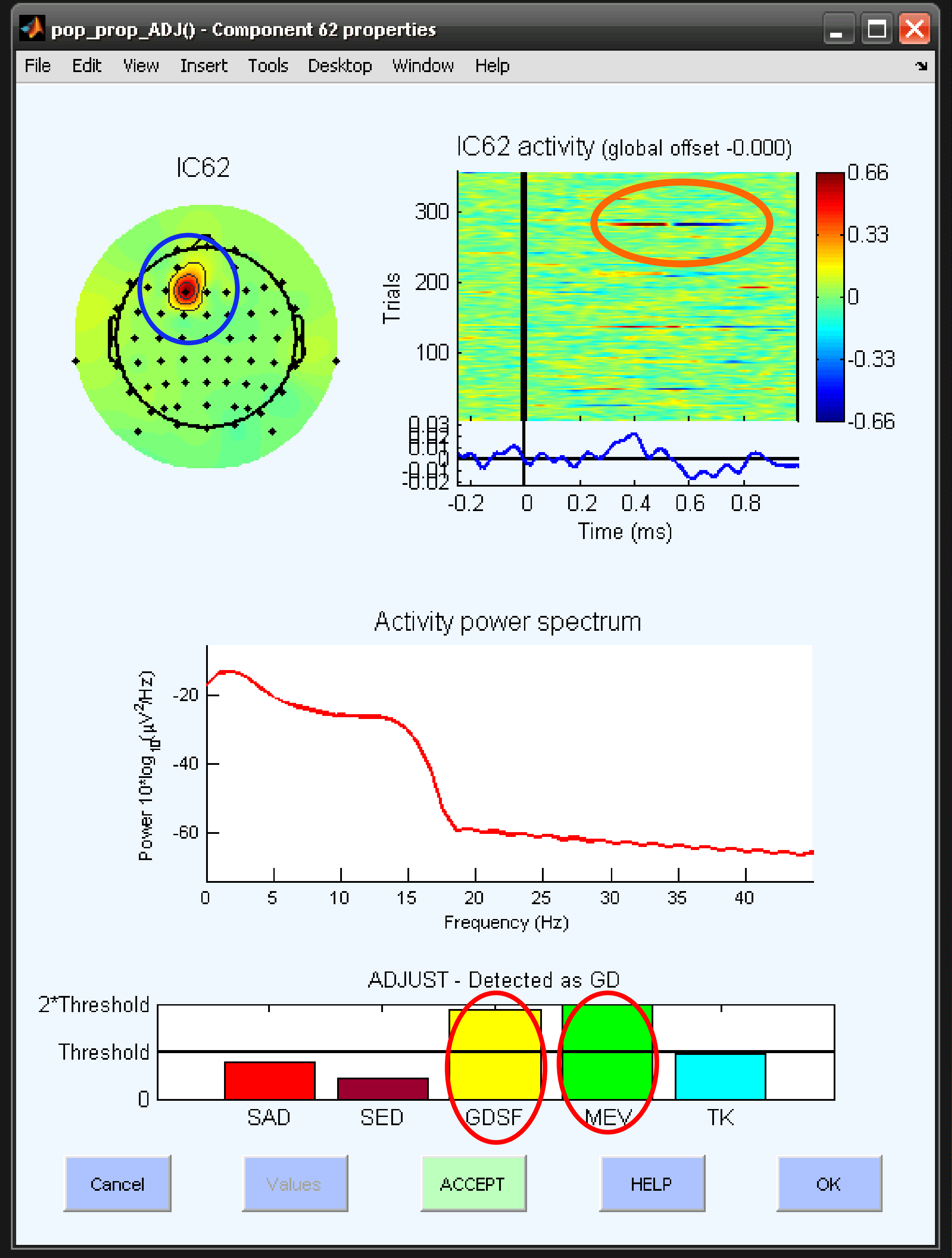
GDSF and MEV features over threshold

Variable distribution even at a single channel

IC activity shows a lot of variability across epochs and doesn't show one as responsible



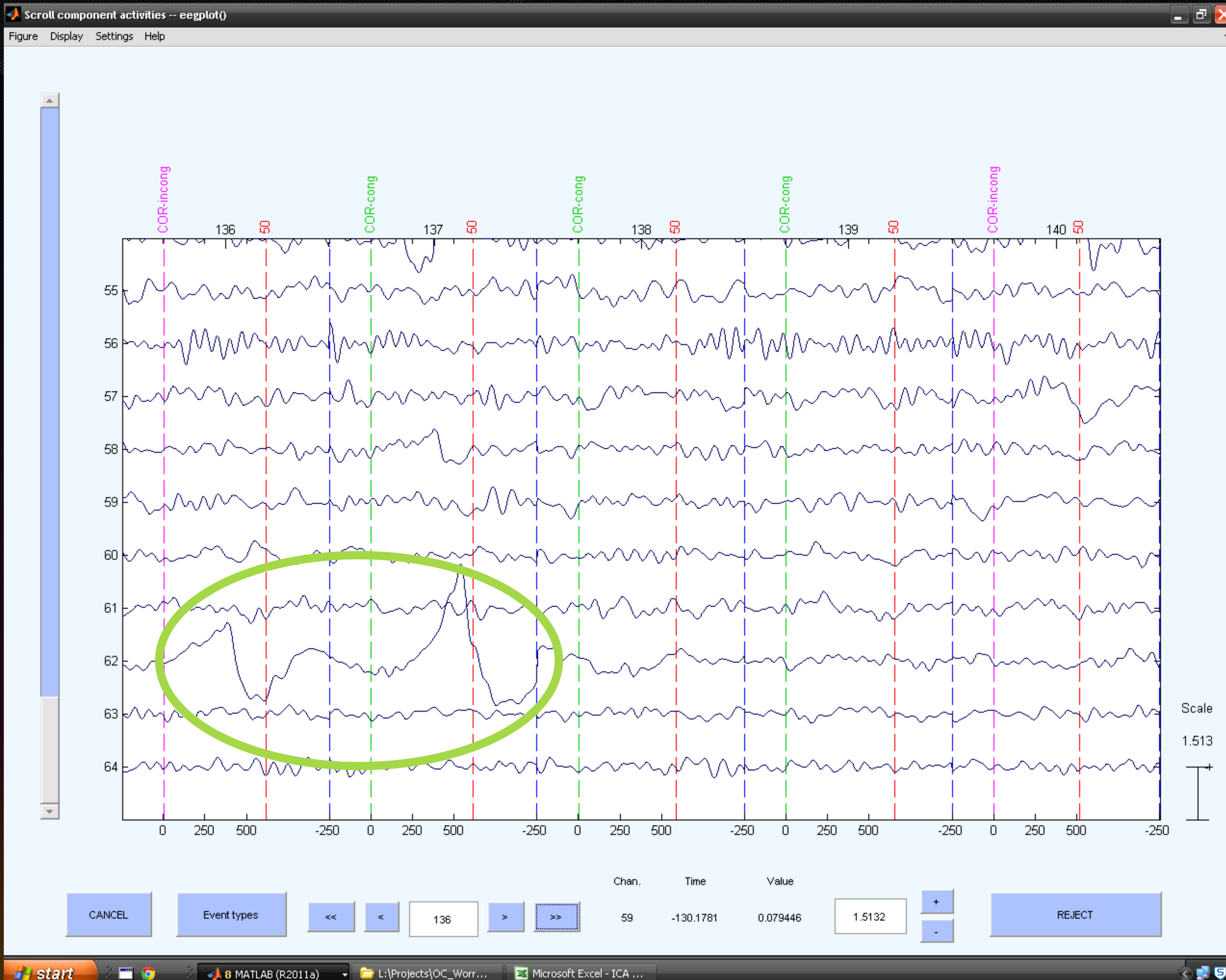

Scroll component activities -- eegplot()
Figure
Display
Settings
Help
COR-incong
COR-cong
COR-cong
COR-cong
COR-incong
151
152
153
154
155
50
29
30
31
32
33
34
35
36
37
38
39
0
250
500
-250
Scale
3.57
CANCEL
Event types
<<
<
151
>
>>
Chan.
Time
Value
41
-244.7904
0.00072099
3.57
+
-
REJECT



GDSF and MEV features over threshold

Variable distribution.
In this case, present at a single channel

In the IC Activity by trial, one section stands out



Scroll component activities -- eegplot()
Figure Display Settings Help
COR-incong
COR-cong
136
137
138
139
140
50
55
56
57
58
59
60
61
62
63
64
0 250 500 -250
Scale
1.513
Chan.
Time
Value
59
-130.1781
0.079446
CANCEL
Event types
<<
<
136
>
>>
1.5132
+
-
REJECT
start
8 MATLAB (R2011a)
L:\Projects\OC_Worr...
Microsoft Excel - ICA ...

Sample Files

Under L:\Projects\OC_Worry\Physiodata

- Few ICs
  - 714flankers.cnt.ICA.MAT
- A lot of ICs
  - 770flankers.cnt.ICA.MAT
  - 701pl.cnt.ICA.MAT