ICA Decomposition of EEG Data & Evaluating ICA Components

EEGLAB Workshop XXI
Santa Margherita Ligure, Italy
Day 1
Independent Component Analysis

\[ x = \text{scalp EEG} \]

\[ W = \text{unmixing matrix} \]

\[ u = \text{sources} \]

\[ W^{-1} (\text{scalp projections}) \]

\[ W^{-1} x = u \]

\[ x = W^{-1} u \]
Source activation = unmixing * Channel data

Channel data = mixing (topo) * Source activation
ICA and PCA

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

While PCA simply decorrelates the outputs (using an orthogonal mixing matrix), ICA attempts to make the outputs \textit{statistically independent}, while placing no constraints on the mixing matrix.
Finally: ICA options

<table>
<thead>
<tr>
<th>Option</th>
<th>Default</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘extended’</td>
<td>0</td>
<td>1 is recommended to find sub-gaussians</td>
</tr>
<tr>
<td>‘stop’</td>
<td>1e-7</td>
<td>final weight change → stop</td>
</tr>
<tr>
<td>‘lrate’</td>
<td>determined from data</td>
<td>too small → too long… too large → wts blow up</td>
</tr>
<tr>
<td>‘maxsteps’</td>
<td>512</td>
<td>more channels → more steps</td>
</tr>
<tr>
<td>‘pca’</td>
<td>0 or EEG.nbchan</td>
<td>Decompose only a principal data subspace</td>
</tr>
</tbody>
</table>

Other algorithms: binica, amica, cudaica, beamica

EEGLAB Workshop XXI, April 4-8, 2016, Italy –John Iversen– Evaluating ICA components
Runica progress...

Input data size [33,133175] = 33 channels, 133175 frames
Finding 33 ICA components using extended ICA. Kurtosis will be calculated initially every 1 blocks using 6000 data points.
Decomposing 122 frames per ICA weight ((1089)^2 = 133175 weights. Initial learning rate will be 0.001, block size Learning rate will be multiplied by 0.98 whenever angledelta >= 80 deg.
More than 32 channels: default stopping weight change 1E-7
Training will end when wchange < 1e-07 or after 512 steps.
Online bias adjustment will be used.
Removing mean of each channel ...
Final training data range: -171.806 to 179.094
Computing the spherering matrix ...
Starting weights are the identity matrix ...
Spherering the data ...
Beginning ICA training ... first training step may be slow ...
step 1 - Irace 0.001000, wchange 16.85061324, angledelta 0.0 deg
step 2 - Irace 0.001000, wchange 0.26760405, angledelta 0.0 deg
step 3 - Irace 0.001000, wchange 0.79058323, angledela 104.0 deg
step 4 - Irace 0.000890, wchange 0.6700031, angledela 147.2 deg
step 5 - Irace 0.000890, wchange 0.62849071, angledela 146.5 deg
step 6 - Irace 0.000890, wchange 0.73967955, angledela 150.7 deg
step 7 - Irace 0.000890, wchange 0.73727229, angledela 151.6 deg
step 8 - Irace 0.000890, wchange 0.74051387, angledela 137.9 deg
step 9 - Irace 0.000886, wchange 0.74563137, angledela 156.0 deg
step 10 - Irace 0.000886, wchange 0.72101402, angledela 143.7 deg
step 11 - Irace 0.000851, wchange 0.14650114, angledela 102.5 deg
step 12 - Irace 0.000854, wchange 0.11822100, angledela 114.3 deg
step 13 - Irace 0.000817, wchange 0.75552966, angledela 100.6 deg
step 14 - Irace 0.000801, wchange 0.26739750, angledela 109.1 deg
step 15 - Irace 0.000785, wchange 0.12123251, angledela 94.2 deg
step 16 - Irace 0.000769, wchange 0.10285606, angledela 110.7 deg
step 17 - Irace 0.000754, wchange 0.09770499, angledela 118.6 deg
step 18 - Irace 0.000739, wchange 0.09544428, angledela 117.1 deg

Sorting components in descending order of mean projected variance ...
Permuting the activation wave forms ...

EEGLAB Workshop XXI, April 4-8, 2016, Italy –John Iversen– Evaluating ICA components
Alternatives to runica

Infomax ICA
runica matlab implementation
binica compiled version; fast
cudaica GPU version


AMICA
Best at extracting dipolar ICs
Multiple-model support
Now what…?

Part 1
Getting an overview of your ICs

Part 2
Classifying/Evaluating ICs

Part 3
Detailed look at IC properties
ERP
Spectrum
ERP images
ERSP
Now what…?

Part 1
Getting an overview of your ICs

Part 2
Classifying/Evaluating ICs

Part 3
Detailed look at IC properties
- ERP
- Spectrum
- ERP images
- ERSP
Load: EEG_data/faces_4.set

Tools → Extract Epochs

Subtract Baseline [-1000 0]

(Some other examples use stern_125Hz.set)
Edit → Dataset Info → Enter Comments

Edit comments of current dataset

Parent dataset: faces_4 continuous

Parent dataset "faces_4 continuous":
Data acquired by: Stefan Debener
Data acquired on: Oct 15, 2005

Data:
33 channel EEG
nose-tip reference
sampling rate: 250 Hz
filtered: .5 - 100 Hz
16 bit, BrainAmps

Task:
speeded discrimination between objects and faces
500 ms presentation duration
ISI: 500-1900 ms
362 trials

CANCEL  SAVE
Results of ICA Decomposition in EEG struct
Source activation = unmixing * Channel data

Channel data = mixing (topo) * Source activation

EEG.icaact = (EEG.icaweights*EEG.icasphere) * EEG.data

EEG.data = EEG.icawinv * EEG.icaact
A convenient ‘trick’…

Use ‘Reject components by map’ to survey components
An interactive overview of ICs
Step 0: Quality of Decomposition

BAD ICA Components

ICA Components
Examining IC Properties

[Image of EEG data analysis interface showing IC properties and waveforms]
IC Properties

**IC Topography**

`topoplot()`

**ERP Image & ERP**

`erpimage()`

**Power Spectrum**

`spectopo()`

EEGLAB Workshop XXI, April 4-8, 2016, Italy – John Iversen – Evaluating ICA components
Many plot panels in EEGLAB will expand when clicked on!
Part 1
Getting an overview of your ICs

Part 2
Classifying/Evaluating ICs
   Eye Artifacts
   Muscle Artifacts
   Other Artifacts
   Brain ICs

Part 3
Detailed look at IC properties
   ERP
   Spectrum
   ERP images
   ERSP
Evaluating ICs

Over time, most EEGLAB users develop a heuristic sense of which ICs might be brain vs. artifact.

Heuristics are generally based on:
- Topography
- Component Activities (scroll)
- ERP
- Power Spectrum

IC Classification is typically used to ‘clean’ data—study likely brain activity without artifacts

There are some efforts to automate this process, but doing it by hand is a good place to start to build intuition.
Topography

EEGLAB Workshop XXI, April 4-8, 2016, Italy – John Iversen – Evaluating ICA components
IC 4 – eyeblink

Classic frontal eye-blink topography

Sporadic large biphasic pulses

That may yet be task related!

But: highly variable (erpalpha=0.01)
Plot → Component Activations (scroll)
IC 4 Activation – eyeblink

Sporadic large biphasic pulses
IC 9 – lateral eye movement

Classic frontal eye-movement topography
IC 9 Activation – lateral eye movement
IC 12, 18 – Muscle

Narrowly spaced dipolar topography (consistent with superficial source)

Noisy ERP/ERP Image

High frequencies dominate power spectrum
IC 12, 18 Activation – Muscle
IC 17, 25 – Bad channels

Punctate topography (single channel)

Sporadic epoch activity (sometimes just a single large spike)
IC 2, 7 – Cardiac

Unusual, peaky spectrum (often peaks ~5, 10 Hz)

Periodic spikes (~1/ sec)

Cardiac-like topographies:
Shallow gradient = extremely distant source
Artifacts

Eye
Muscle
Cardiac
Badchan
Brain ICs

Classic occipital topography

Strongly task-related ERP Image & ERP

10 Hz Alpha peak
Dipole orientation matters
Brain ICs

Classic radial-dipole source topography

Strongly task-related ERP Image & ERP
Brain ICs

Classic tangential-dipole source topography
two peaks, not as closely spaced as muscle: deeper

Task-related ERP Image & ERP
IC Classification...so far

Eye
Muscle
Cardiac
Badchan
Brain
Now what…?

Part 1
Getting an overview of your ICs

Part 2
Classifying/Evaluating ICs

Part 3
Detailed look at IC properties
- ERP
- Spectrum
- ERP images
- ERSP
Component ERPs

EEGLAB v7.1.7.13b

#2: Stereo 10/20

Component ERPs

- Channel locations
- Channel data (scroll)
- Channel spectra and maps
- Channel properties
- Channel ERP image
- Channel ERPs
- ERP map series
- Sum/Compare ERPs
- Component activations (scroll)
- Component spectra and maps
- Component maps
- Component properties
- Component ERP image

Component ERPs

- In rectangular array

Sternberg: Memorize epochs ERP

Figure 2: Plot > Component ERPs > In rect. array -- plotdata()
A step back: Electrode-level ERP

EEGLAB Workshop XXI, April 4-8, 2016, Italy – John Iversen – Evaluating ICA components
Traditional ERP: Time-locked activity at each channel

Topography at latency of two peaks

EEG peaks are sometimes called “components.”
Not to be confused with ICAs Independent Components (ICs)!
ERP at two channels

![EEGLAB Workshop XXI, April 4-8, 2016, Italy –John Iversen– Evaluating ICA components](image)
Definition: The data envelope

Data (all channels)

Data envelope (max/min traces)
Definition: IC Envelope

IC 10 Topography

Data Envelope

IC Envelope

IC 10 Activation ERP back-projected to channels
Key: Scalp ERP peaks are often the sum of multiple independent source processes.

PPAF/PVAF
“% Variance Accounted For”

The variance of scalp EEG accounted for by this component.

This component accounts for all of the negative scalp ERP peak at ~150 ms, but only some of other ERP peaks.

Net PVAF: 37%
Component ERP envelope
ERP peak- and IC Component-topographies
Component 3 ERP envelope
Component 1 ERP envelope

Note: IC Envelope can exceed the data envelope: Other component(s) have opposite sign at this latency.
Component 1 + 3 ERP envelope

Else plot these component numbers only (Ex: 2:4,7): 1 3
pvaftopo plugin (Makoto Miyakoshi)

Max PVAF ~ 40%
Max PVAF ~ 70%

Note: However, PVAF is calculated over entire signal duration. PVAF at times of peaks often higher, but still typically not 100%.
Top 6 IC contributions to data ERP envelope
Non-artifact IC contrib. to data ERP envelope

Else plot these component numbers only (Ex. 2, 4, 7).

Component numbers to **remove** from data before plotting:

Plot title:

Largest non-artifact ERP component
Non-artifact IC contrib. to data ERP envelope

Else plot these component numbers only (ex: 2, 4, 7).

Component numbers to remove from data before plotting:

Plot title:

Artifact Components

Figure 2

Largest non-artifact ERP components of faces_4 face epochs

pop_envtopo(EEG, [-500 1500], 'limcontrib', [0 1500], ...
'compsplot', [6], 'subcomps', [2 4 5 7 9 12 17 18 25], ...
'title', 'Largest non-artifact ERP components of faces_4 face epochs', ...
'electrodes', 'off');
Compare: Effect of removing artifacts
What is the IC ERP difference between these 2 conditions?

(Data: stern_125Hz.set)
IC ERP difference
IC ERP difference

Largest ERP components of Memorize-Ignore epochs

-6, 23, 10, 5, 24, 8
\[ \text{ppaf 36.68\%} \]
Now what…?

Part 1
Getting an overview of your ICs

Part 2
Classifying/Evaluating ICs

Part 3
Detailed look at IC properties
- ERP
- Spectrum
- ERP images
- ERSP
Plot component power spectrum

By default, plots topographies for frequency of largest peak.
Select the frequency for topographies
Now what…?

Part 1
Getting an overview of your ICs

Part 2
Classifying/Evaluating ICs

Part 3
Detailed look at IC properties
ERP
Spectrum
ERP images
ERSP
Component ERP image
ERP Image basics

by default, sorted by time-on-task
(1\text{st} trial, 2\text{nd} trial, ...)

EEGLAB Workshop XXI, April 4-8, 2016, Italy –John Iversen– Evaluating ICA components
ERP Image basics

Trial 1:
Trial 2:

No Smoothing

Smoothed across 10 Trials

EEGLAB Workshop XXI, April 4-8, 2016, Italy –John Iversen– Evaluating ICA components
ERP Images: smoothing across trials
Component ERP Image: Sort by RT
Component ERP Image: Sort by RT

Select fields:
- latency
- type
- epoch

Sort trials by epoch event values:
- Epoch-sorting field: latency
- Event type(s): bp1, bp4
- Event time range: 0-2000

Sort trials by phase:
- Frequency (Hz):
- Inter-trial coherence:

Other options:
- Plot spectrum (min Hz):
- Mark times (ms):
- More options (see >> help erpimage):
Component ERP Images: Sort by phase

Phase-sorted image
### Component ERP Images: ITC

#### Component ERP Image - `pop_erpimage()`

- **Component(s):** 3
- **Project to channel #:**
- **Smoothing:** 10
- **Downsampling:** 1
- **Time limits (ms):** -800 to 1000

#### Sort/align trials by epoch event values
- **Epoch-sorting field:**
- **Event type(s):**

#### Sort trials by phase
- **Frequency (Hz | minHz maxHz):** 10 12
- **Percent low-amp. trials:**

#### Inter-trial coherence options
- **Frequency (Hz | minHz maxHz):** 10 12
- **Signif. level (<0.2):** 0.01

#### Other options
- **Plot spectrum (minHz maxHz):**
- **Baseline ampl. (dB):**

---

#### Phase-sorted alpha power

- **ERSP**
- **ITC**

---

**ITC**

**ERSP**

**Phase-sorted alpha power**
Component ERP Images: Sort by amplitude

Phase-sorted alpha power

Same data: Sorted by alpha amplitude

'ampsort' = [center_ms, prcnt, freq, maxfreq]  Sort epochs by amplitude.
Component ERP Images: Amplitude vs. Activations

Same sorting order: Plotting Amplitude vs. Activations
Now what…?

Part 1
Getting an overview of your ICs

Part 2
Classifying/Evaluating ICs

Part 3
Detailed look at IC properties
  ERP
  Spectrum
  ERP images
  ERSP
Definition: ERSP

Event Related Spectral Perturbation

Change in power in different frequency bands relative to a baseline. ERS, ERD
Plot IC ERSP

Time points removed at beginning and end of epoch to avoid edge effects of wavelet transform.
Plot IC ERSP
Some attempts to automate the IC classification:

“Automatic Classification of Artifactual ICA-Components for Artifact Removal in EEG Signals”
Irene Winkler, Stefan Haufe and Michael Tangermann (2011)
http://www.behavioralandbrainfunctions.com/content/7/1/30

Bigdely-Shamlo’s EyeCatch (2013)
https://www.researchgate.net/publication/257602145_EyeCatch_Data-mining_over_half_a_million_EEG_independent_components_to_construct_a_fully-automated_eye-component_detector

Luca Pion-Tonachini (ongoing)
Crowd-sourcing heuristic knowledge about IC components to build automatic classifier
We'll play the game later: http://reaching.ucsd.edu:8000