Independent Component Analysis of Electrophysiological Data

Scott Makeig
Institute for Neural Computation
University of California San Diego

29th EEGLAB Workshop
Bois Perché, Aspet, France

June 17-21, 2019
ICA can find distinct EEG source activities -- and their ‘simple’ scalp maps!
ICA Assumptions

- Mixing is **linear** at electrodes
- Propagation **delays** are negligible
- Component **locations** are fixed
- Component time courses are independent
- # components <= # scalp channels

\[ \text{Contribution to EEG} \ \Rightarrow \ # \text{Scalp channels} \]

\[ \text{# Effective sources} \rightarrow \]
Are EEG effective source signals independent?

Independent Domains of Local Synchrony

Cortex

Thalamus

Freeman - phase cones
Plenz - avalanches

S. Makeig (2007)
Properties of EEG Independent Components

- Maximally Temporally Independent
- Concurrently Active and Spatially Overlapping
- *Dipolar* Scalp Maps  (Delorme et al., 2012)
- Functionally Distinct
- Between-Subject Similarity / Complexity
ICA vs. PCA

PCA simply decorrelates the outputs using an **orthogonal mixing matrix**.

PCA makes each successive component account for as much **variance** in the data as possible.

ICA makes each component account for as much **temporally independent information** in the data as possible, with no constraints on the mixing matrix.

**PCA lumps – ICA splits!**

Arnaud Delorme & S. Makeig, 2016
ICA in practice

Onton & Makeig, 2006
Classifying ICs

Non-brain sources

EOG

EMG

Brain (bilateral)

Effective brain sources

Brain

http://ICLabel.ucsd.edu

J. Onton & S. Makeig, 2005
A P300' visual target response at a vertex channel

What sources contribute to this potential?

No scalp response in these trials ... Why not?

The vertex channel response sums 238 independent sources.
Scalp EEG signals are strong mixtures of brain sources.

In this sense scalp channel signals are *epiphenomena*. Source signals are the EEG phenomena of real interest!

No more than ~30% of any scalp channel variance is produced by any one brain source!
ICA finds Non-Brain Independent Component (IC) Processes ...

... separates them from the remainder of the data ...
... and also separates cortical brain IC processes
Single Session - Two Maximally Independent Alpha Processes
EEG Source Localization

LORETA = Low-Resolution Electrical Tomography
Sparse Compact Smooth (SCS)

Distributed IC source location estimation using SCALE-optimal head model.

Z. Akalin Acar et al., 2016
ICA is a linear data decomposition method.

Makeig & Onton, 2011

\[ \text{Act} = W \times \text{Data} \]

\[ \text{Data} = W^{-1} \times \text{Act} \]

\[ \text{Data} = W^{-1} \times (W \times \text{Data}) \]
Infomax ICA learning approach

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 \( H(y_1, y_2) \) \iff \( \text{minimizing } I(y_1, y_2) \).

Infomax

The learning rule:

\[ \Delta W \propto \frac{\partial H(y)}{\partial W} W^T W \]

Natural gradient normalization (Amari)

Is 0 if the two variables are independent.
Historical Remarks

• Herault & Jutten ("Space or time adaptive signal processing by neural network models“, *Neural Nets for Computing Meeting*, Snowbird, Utah, 1986): Seminal paper

• Bell & Sejnowski (1995): Information maximization (Infomax)

• Makeig, Bell, Jung, Sejnowski (1996); ICA decomposition of EEG

• Amari et al. (1996): Natural gradient learning

• Cardoso (1996): Joint approximate diagonalization (JADE)

• Hyvarinen (1999): (fastICA)

• Lee/Girolami (1999): Mixture model ICA (Extended Infomax)

• Palmer (2006-9): Adaptive mixture ICA (AMICA)

Applications of ICA to biomedical signals

– EEG/ERP analysis (Makeig, Bell, Jung & Sejnowski, *NIPS 1996*)
– fMRI analysis (McKeown et al., 1998)
– Fetal/mother ECG separation (Cardoso, 1998)
– Electrocorticography (ECoG) (Whitmer, 2010)
Delorme et al., *PLOS One*, 2012

S. Makeig, 2011
Important Recent Result (2012)

Those linear decompositions of multi-channel EEG data that find ICs whose time courses are more temporally independent …

Also find more ICs whose scalp maps are highly ‘dipolar’ – i.e., ICs compatible with the spatial projection of a single local cortical (or non-brain, ‘artifactual’) source process – whose location and distribution can be more or less accurately estimated.

The more independent the component time courses ↔

The larger the number of ~dipolar component scalp maps.

Hypothesis: Dipolar ICs = Localized cortical source processes

Delorme et al., PLOS One, 2012
ICA transforms multiple correlated channels into maximally mutually independent component subspaces
Distinguishing IC types: ICLabel

→ https://labeling.ucsd.edu

Info: https://sccn.ucsd.edu/wiki/ICLabel

The ICLabel plug-in

L. Pion-Tonachini, 2018
Figure 4: The top panel shows the hypnogram, i.e., sleep stages annotated from the EEG record by a sleep expert, of a sleep session from a single subject. Bottom panels show mean probabilities, within each 30-sec sleep scoring interval, of ICA models learned by an 8-model AMICA decomposition applied to the EEG record. Red-shaded regions highlight changes in model probabilities for relevant models during transitions to and periods of deep sleep (N4). Gray-shaded regions highlight probability value changes for relevant models during REM sleep.

Figure 6: Scatter plot of window-mean model probabilities for AMICA model clusters A, B, and C (cf. Fig. 5), each point representing mean model probability within a 30-sec data segment from sleep recordings of 7 healthy subjects and 10 patients. Colors represent expert designated sleep-stage labels for the same data segments. Note the distinct deep sleep (N4) pattern and the relative closeness of wake and REM sleep characteristics.

Multi-model ICA: AMICA

Sleep stages:
Multi-model ICA: AMICA

Sleep stages:

(a) Time-locked to driving challenges
(b) Time-locked to start of car steering
(c) Time-locked to end of car steering

S. Hsu et al., Neuroimage, 2018
Are effective source *locations* similar across subjects?

Within task: Are source locations similar across participants?
Effective Source Density

Visual Working Memory

dipoledensity()

Onton & Makeig, 2005
Effective Source Density

Eyes-closed emotion imagination

>> dipoledensity()
Effective Source Density

Letter twoback with feedback

Onton & Makeig, 2005
Effective Source Density

Auditory novelty oddball
Effective Source Density

A. Old/new word memory

Onton et al., '05

Onton & Makeig, 2005
Effective Source Density

B. Visually cued selective response

Onton & Makeig, 2005
Are effective source *dynamics* similar across participants?
Visual Selective Attention Task

15 subjects
31 channels
Westerfield & Townsend

S. Makeig (2001)
Baseline power spectral shifts with direction of attention.

Baseline Spectrum (1 s before stimulus)
Example: frontal midline theta cluster

ERP image of 9.6-Hz alpha power

Load 3
Load 5
Load 7

-27.8 dB

Onton, Delorme & Makeig, 2005.
Why don’t all subjects contribute to every IC cluster?
Subject differences?

Significant ITC differences (by bootstrap) between the LOC and fLOC clusters immediately follow Probe presentation (5-11 Hz).
Subject differences?
Subject differences?
Properties of EEG Independent Components

- Maximally Temporally Independent
- Concurrently Active and Spatially Overlapping
- *Dipolar* Scalp Maps
- *Functionally* Distinct
- Between-Subject Similarity / Complexity
Important Recent Result

Those linear decompositions of multi-channel EEG data that find ICs whose time courses are **more** temporally independent …

Also find **more** ICs whose scalp maps are highly ‘**dipolar**’ – i.e., ICs compatible with the spatial projection of a single local cortical effective source (or non-brain artifact) – whose cortical location can be accurately estimated given a good forward-problem head model.

**More-independent IC time courses** ↔ **Larger # of dipolar IC scalp maps**

Delorme et al., *PLOS One*, 2012
Sleep spindles
Not the End...
Multi-model ICA: AMICA

S. Hsu et al., Neuroimage, 2018
The ‘receptive field’ of a bipolar EEG channel!
Multi-model ICA: AMICA

Reaction speed to simulated driving challenge:

Figure 8: The top panel shows reaction speed changes (inverse of reaction times) in response to lane-departure challenges in one simulated driving session. The three bottom panels show the 5-sec smoothed probabilities of the three ICA models learned by a three-model AMICA decomposition of the whole EEG data session before lane-departure events. Correlation coefficients ($r$) between each model probability time course and reaction speed are indicated. Black arrows in the lower panel mark brief (alert) periods when model M3 was dominate and reaction speed high.

S. Hsu et al., Neuroimage, 2018