Advances in the analysis of spontaneous EEG and MEG by independent component analysis

Aapo Hyvärinen

Dept of Mathematics and Statistics & Dept of Computer Science
University of Helsinki, Finland

with
Pavan Ramkumar, Riitta Hari, Lauri Parkkonen (Aalto University, Finland)
Kun Zhang (Univ. of Helsinki, MPI Tübingen)
Abstract

- Why is independent component analysis (ICA) needed?
- Improving ICA of “spontaneous” EEG/MEG
  - Applying ICA on time-frequency decompositions
  - Spatial version of independent component analysis (ICA)
- Testing components: Are they just random effects?
  - Intersubject consistency provides an plausible null hypothesis
- Causal analysis / effective connectivity
  - Analyze correlations of the envelopes/amplitudes of sources
  - Structural equation models better estimated using non-Gaussianity
What is EEG and MEG?

- Measurements of electrical activity in the brain
  - High temporal accuracy (millisecond scale)
  - Not so high spatial accuracy (less than in fMRI)
- Typically characterized by oscillations, e.g. at 10 Hz
- Up to 306 time series (signals), $10^4 \ldots 10^5$ time points.
Why use independent component analysis (ICA)?

- Exploratory analysis like ICA can give information about internal dynamics not related to stimulation
- Necessary to analyze resting-state data: no stimulation at all
  - With fMRI, resting-state networks found, how about EEG/MEG?
- Natural stimulation (i.e. watching a movie) too complex to be described as a box-car function
- “Two-person neuroscience” (R. Hari) presents similar analysis challenges
Consider EEG or MEG recording $x_i(t)$ of “spontaneous” (resting / natural stimulation) activity.

We assume a linear mixing model

$$x_i(t) = \sum_j a_{ij} s_j(t)$$

$i$ is EEG/MEG channel index, $t$ is time, $s_j(t)$ are “hidden” sources, $a_{ij}$ are unknown mixing parameters.

Goal: find sources and mixing based on statistical criteria.

Sources found by maximizing a statistic of linear combinations $\sum_i w_i x_i(t)$. 
Blind separation obtained by maximizing sparseness

- Independent component analysis (ICA) finds sources based on assumption of statistical independence.
- Deep theoretical result: Sources found by maximizing sparseness, which is computationally much simpler.
- Sparseness means amplitude distribution (histogram) has heavy tails and a peak at zero:
- Separation criterion should ideally also define which sources are interesting.
Using time-frequency decompositions to better find rhythmic sources

- Problem: Rhythmic sources (oscillations) may not be sparse
- ICA algorithms mainly find artifacts, which are sparse
Using time-frequency decompositions to better find rhythmic sources

- Problem: Rhythmic sources (oscillations) may not be sparse
- ICA algorithms mainly find artifacts, which are sparse

- Solution: Perform ICA on short-time Fourier transforms:
  - Divide each channel into time windows e.g. 1 sec long
  - Fourier transform each window
Application on spontaneous MEG data

- **MEG data**
  - 204 gradiometer channels measured at sampling rate of 600 Hz
  - Subject resting, eyes closed
  - Signal space separation for noise reduction
  - Downsampling to 150 Hz

- **Analysis**
  - Half-overlapping windows of 1 second extracted
  - Fast Fourier Transform computed of each window
  - Only frequencies in the range 5–30 Hz used
  - Dimension reduced by PCA to 25
  - Complex-valued FastICA used to estimate 10 sources
  - Reliability analysis performed by ICASSO
Results on resting-state MEG data

Temporal envelope (arbitrary units)  Fourier amplitude (arbitrary units)  Distribution over channels  Phase differences

1 2 3 4 5 6 7 8 9

0 100 200 300

Time (seconds)

5 10 15 20 25 30

Frequency (Hz)
Spatial ICA in fMRI (background)

- Assume we observe several brain images at different time points.
- ICA expresses observed images as linear sums of “source images”:
  - $= a_{11}$
  - $= a_{21}$
  - $= a_{n1}$

- Reverses the roles of observations and variables.
Spatial ICA for MEG: methods

- Spatial ICA possible for MEG by projecting data on the cortex
  - E.g. minimum norm estimate, a linear pseudoinverse operator
- We combine this with the short-time Fourier transforms above
- Finds sources which are maximally spatially localized and bandpass
  - Independence assumed spatially, “non-systematically overlapping” (F. Esposito)
  - No assumption on temporal independence
Spatial ICA for MEG: results

(Here, not resting data but with “naturalistic stimulation”)

Aapo Hyvärinen

Advances in the analysis of spontaneous EEG and MEG by indep
How do we know that an estimated component is not just a random effect?

ICA algorithms give a fixed number of components and do not tell which ones are reliable (statistically significant).

Algorithmic artifacts also possible (local minima).

We develop a statistical test based on inter-subject consistency:

- Do ICA separately on several subjects
- A component is significant if it appears in two or more subjects in a sufficiently similar form
- We propose a null hypothesis to quantify this idea, and a practical algorithm for doing the analysis
ICA is a rotation of whitened data

Under null hypothesis, spatial patterns of different subjects are “completely random” rotations in the PCA subspace (uniformly distributed in the set of orthogonal matrices).

(We also assume the PCA subspace same for all subjects)

Formulating a null hypothesis allows us to compute the chance level and control false positive rate
Testing ICs: grouping patterns over subjects

- Compute similarities of spatial patterns using Mahalanobis metric
  \[
  \gamma_{ij,kl} = \frac{|p_{ik}^T M p_{jl}|}{\sqrt{p_{ik}^T M p_{ik}} \sqrt{p_{jl}^T M p_{jl}}}
  \] (2)
  with \( M \) is (stabilized) inverse of covariance matrix of the patterns \( p \)

- Under null hypothesis, marginal distribution of \( \gamma \) can be obtained in closed form

- **Threshold** similarities based on desired false positive rate or false discovery rate.

- Group components of different subjects together with a method similar to hierarchical clustering
Testing ICs: results
Causal analysis: Introduction

- Model connections between the measured variables
- Two fundamental approaches
  - If time-resolution of measurements fast enough, we can use autoregressive modelling
  - Otherwise, we need structural equation models
- If measured variables are sensors, we should first localize sources
- After blind source separation, sources are uncorrelated
  ⇒ More meaningful to model dependencies of envelopes (amplitudes, variances)
Structural equation models

- How does an externally imposed change in one variable affect the others?

\[ x_i = \sum_{j \neq i} b_{ij} x_j + e_i \]

- Difficult to estimate, not simple regression
  - Classic methods fail in general
Structural equation models

- How does an externally imposed change in one variable affect the others?

\[ x_i = \sum_{j \neq i} b_{ij} x_j + e_i \]

- Difficult to estimate, not simple regression
  - Classic methods fail in general
  - Can be estimated if (Shimizu et al., JMLR, 2005)
    1. the \( e_i(t) \) are mutually independent
    2. the \( e_i(t) \) are non-Gaussian, e.g. sparse
    3. the \( b_{ij} \) are acyclic: There is an ordering of \( x_i \) where effects are all “forward”
Using generalized autoregressive heteroscedasticity (GARCH)

- Based on formalism widely used in econometrics
- We model causal influences between the variances $\sigma^2_{i,t}$ of the sources $s_{j,t}$ as

$$\sigma^2_{i,t} = \omega_i + \sum_j \sum_{\tau} \alpha_{ij,\tau} s^2_{j,t-\tau} + \beta_i \sigma^2_{i,t-1}. \quad (3)$$

where $\tau$ is the time lag of prediction.
- Parameters $\alpha_{ij}$ indicate the effective connectivity between sources (their amplitudes)
Results of GARCH model

Causality in the time-varying variances of the separated sources. The thickness of the lines indicates the strength of the causal effects. Undirected lines indicate bi-directed causal relations. Black (red) lines show positive (negative) effects. Note that the green line divides all sources into two groups; the effects between the groups are negative, while those inside each group are positive.
Exploratory data analysis by ICA can give information about:
- activity not directly related to stimulation
- responses when stimulation too complex
- internal dynamics (incl. resting-state data)

We present two stages of analysis:
- Finding sources by:
  - “Fourier-ICA” using short-time Fourier transforms, or
  - Spatial (Fourier) ICA
- Analyzing their effective connectivity:
  - Dependencies of envelopes (energies) of sources
  ⇒ Modifications of SEM and GARCH models

At some point, intersubject consistency should be analyzed:
- Makes significance tests possible