MELODIC
Multivariate Exploratory Linear Optimised Decomposition into Independent Components

- Resting state fMRI
- Independent Component Analysis
  - Independence
  - Overfitting & thresholding
  - ICA cleanup
- Multi-subject ICA
  - Tensor ICA
  - Concatenated ICA
  - Dual regression
Resting state fMRI
Model-based (GLM) analysis

- model each measured time-series as a linear combination of signal and noise
- If the design matrix does not capture every signal, we typically get wrong inferences!
Data Analysis

Confirmatory

- “How well does my model fit to the data?”

Problem → Data → Model → Analysis → Results

- results depend on the model

Exploratory

- “Is there anything interesting in the data?”

Problem → Data → Analysis → Model → Results

- can give unexpected results
FMRI inferential path

Interpretation of final results

Analysis

Experiment

Physiology

MR Physics
Variability in FMRI

Interpretation of final results

Experiment
- suboptimal event timing
- inefficient design, etc.

Physiology
- secondary activation
- ill-defined baseline
- resting-fluctuations etc.

Analysis
- filtering & sampling artefacts
- design misspecification
- stats & thresholding issues etc.

MR Physics
- MR noise
- field inhomogeneity
- MR artefacts etc.
Resting state methods

**ICA**
- Multivariate voxel-based approach
- Finds interesting structure in the data
- Exploratory “model-free” method
- Spatial approach

**Network modelling**
- Node-based approach (first need to parcellate the brain into functional regions)
- Map connections between specific brain regions (connectomics)
- Temporal approach
Independent Component Analysis
Model-free?

There is no explicit time-series model of assumed ‘activity’
There is an underlying mathematical (generative) model

\[ Y^i = S^i A^i + E^i, \quad \text{where} \quad E^i_j \sim \mathcal{N}(0, \sigma^2_{Y^i} I) \]
Exploratory techniques

- try to ‘explain’ / represent the data
  - by calculating quantities that summarise the data
  - by extracting underlying ‘hidden’ features that are ‘interesting’
- differ in what is considered ‘interesting’
  - are localised in time and/or space (Clustering)
  - explain observed data variance (PCA, FDA, FA)
  - are maximally independent (ICA)
Melodic

multivariate linear decomposition:

Data is represented as a 2D matrix and decomposed into components.
What are components?

- express observed data as linear combination of spatio-temporal processes

- techniques differ in the way data is represented by components
Spatial ICA for FMRI

- data is decomposed into a set of spatially independent maps and a set of time-courses

McKeown et al. HBM 1998
Independence
PCA vs. ICA?

Simulated Data
(2 components, slightly different timecourses)

PCA
- Timecourses orthogonal
- Spatial maps and timecourses “wrong”

ICA
- Timecourses non-co-linear
- Spatial maps and timecourses “right”

Simulated Data (2 components, slightly different timecourses)

PCA
- Timecourses orthogonal
- Spatial maps and timecourses “wrong”

ICA
- Timecourses non-co-linear
- Spatial maps and timecourses “right”
PCA vs. ICA

- PCA finds projections of maximum amount of variance in Gaussian data (uses 2nd order statistics only)
- Independent Component Analysis (ICA) finds projections of maximal independence in non-Gaussian data (using higher-order statistics)
PCA vs. ICA

- PCA finds projections of maximum amount of variance in Gaussian data (uses 2nd order statistics only)

- Independent Component Analysis (ICA) finds projections of maximal independence in non-Gaussian data (using higher-order statistics)
Correlation vs. independence

- de-correlated signals can still be dependent

- higher-order statistics (beyond mean and variance) can reveal these dependencies

Stone et al. 2002
Non-Gaussianity

sources

mixing

mixtures
Non-Gaussianity

non-Gaussian

Gaussian

mixing
ICA estimation

- Random mixing results in more Gaussian-shaped PDFs (Central Limit Theorem)

- Conversely:

  if mixing matrix produces less Gaussian-shaped PDFs this is unlikely to be a random result

  ➡ measure non-Gaussianity

- Can use neg-entropy as a measure of non-Gaussianity

Hyvärinen & Oja 1997
ICA estimation

- need to find an unmixing matrix such that the dependency between estimated sources is minimised

- need (i) a contrast (objective/cost) function to drive the unmixing which measures statistical independence and (ii) an optimisation technique:
  - kurtosis or cumulants & gradient descent (Jade)
  - maximum entropy & gradient descent (Infomax)
  - neg-entropy & fixed point iteration (FastICA)
Overfitting & thresholding
The ‘overfitting’ problem

fitting a noise-free model to noisy observations:

- no control over signal vs. noise (non-interpretable results)
- statistical significance testing not possible

GLM analysis  standard ICA (unconstrained)
Probabilistic ICA model

Statistical “latent variables” model: we observe linear mixtures of hidden sources in the presence of Gaussian noise

\[ Y = XB + E \]

Issues:

- Model Order Selection: how many components?
- Inference: how to threshold ICs?
Model Order Selection

‘How many components’?

*under-fitting*: the amount of explained data variance is insufficient to obtain good estimates of the signals.

*optimal fitting*: the amount of explained data variance is sufficient to obtain good estimates of the signals while preventing further splits into spurious components.

*over-fitting*: the inclusion of too many components leads to fragmentation of signal across multiple component maps, reducing the ability to identify the signals of interest.
Model Order Selection

- observed Eigenspectrum of the data covariance matrix
- Laplace approximation of the posterior probability of the model order
- theoretical Eigenspectrum from Gaussian noise

optimal fit

under-fitting  over-fitting
Variance Normalisation

- we might choose to ignore temporal auto-correlation in the EPI time-series but:
- generally need to normalise by the voxel-wise variance

Estimated voxel-wise std. deviation (log-scale) from FMRI data obtained under rest condition.
Thresholding
Thresholding

- classical null-hypothesis testing is invalid
- After estimation, the spatial maps are a projection of data on to the ‘unmixing matrix’
- data is assumed to be a linear combination of signals and noise
- the distribution of the estimated spatial maps is a mixture distribution!
Alternative Hypothesis Test

- use Gaussian/Gamma mixture model fitted to the histogram of intensity values (using EM)
What about overlap?
What about overlap?

\[ \rho = 0.5 \]
What about overlap?

Sources $\prec 0.1$

Sources + noise

$\rho = 0.5$  $\rho < 0.1$
What about overlap?

- Sources
- Sources + noise
- ICA solution

\[
\rho = 0.5 \quad \rho < 0.1 \quad \rho = 0
\]
What about overlap?

Sources

Sources + noise

ICA solution

after thresholding

\[ \rho = 0.5 \]

\[ \rho < 0.1 \]

\[ \rho = 0 \]

\[ \rho \approx 0.5 \]
What about overlap?

- Spatial correlation

$$\rho(s_1, s_2) = \frac{s_1^t s_2}{N \sqrt{\text{Var}(s_1)} \sqrt{\text{Var}(s_2)}}$$

- in the presence of noise

$$\rho(s_1 + \eta_1, s_2 + \eta_2) = \frac{s_1^t s_2}{N \sqrt{\text{Var}(s_1) + \sigma^2} \sqrt{\text{Var}(s_2) + \sigma^2}}$$
ICA cleanup
Artefact detection

- FMRI data contain a variety of source processes
- Artifactual sources typically have unknown spatial and temporal extent and cannot easily be modelled accurately
- Exploratory techniques do not require a priori knowledge of time-courses and spatial maps
motion
cardiac
susceptibility motion
multiband
signal
effects of scan parameters
manual classification

https://doi.org/10.1016/j.neuroimage.2016.12.036
semi-automatic classification

- **FIX** ([fsl.fmrib.ox.ac.uk/fsl/fslwiki/FIX](fsl.fmrib.ox.ac.uk/fsl/fslwiki/FIX))
  - Classifier with many features
  - Requires manually labelled training data
  - 99% accuracy on high-quality data
- **ICA-AROMA** ([github.com/rhr-pruim/ICA-AROMA](github.com/rhr-pruim/ICA-AROMA))
  - Simple classifier with only 4 features
  - No training data required
  - Mainly designed for motion artefacts
Structured Noise and the GLM

- ‘structured noise’ appears:
  - in the GLM residuals - inflates variance estimates (more false negatives)
  - in the parameter estimates (more false positives and/or false negatives)
- In either case leads to suboptimal estimates and wrong inference!
Denoising FMRI

- Example: left vs right hand finger tapping

**LEFT - RIGHT contrast**
Multi-subject ICA
Different ICA models

**Single-Session ICA**
- Each ICA component comprises:
  - Spatial map & timecourse

**Multi-Session or Multi-Subject ICA: Concatenation approach**
- Each ICA component comprises:
  - Spatial map & timecourse
  (that can be split up into subject-specific chunks)

**Multi-Session or Multi-Subject ICA: Tensor-ICA approach**
- Each ICA component comprises:
  - Spatial map, session-long-timecourse
  & subject-strength plot
Different ICA models

Single-Session ICA
each ICA component comprises:
spatial map & timecourse

Multi-Session or Multi-Subject ICA:
Concatenation approach
good when:
each subject has DIFFERENT timeseries
e.g. resting-state FMRI

Multi-Session or Multi-Subject ICA:
Tensor-ICA approach
good when:
each subject has SAME timeseries
e.g. activation FMRI
Tensor-ICA Group analysis

Extend single ICA to higher dimensions
Resting state multi-subject ICA

- Why not just run ICA on each subject separately?
  - Correspondence problem (eg RSNs across subjects)
  - Different splittings sometimes caused by small changes in the data (naughty ICA!)

- Instead - start with a “group-average” ICA
  - But then need to relate group maps back to the individual subjects
Concatenated ICA

- Concatenate all subjects’ data temporally
  - For lots of subjects use MIGP = iterative PCA to reduce memory requirements
  - Then run ICA
  - More appropriate than tensor ICA (for RSNs)

\[
\text{Subject 1} \quad \text{Subject 2} \\
\vdots \quad \vdots
\]

\[
\begin{align*}
\text{voxels} & \quad \#\text{components} \\
\text{time} & \quad \text{time} \\
\end{align*}
\]

\[
\begin{align*}
\text{group mixing} & \quad \times \quad \text{group ICA maps} \\
\end{align*}
\]
Resting state networks
Resting state multi-subject ICA

Group ICA map

Example subject maps derived from dual regression
Two steps that both involve multiple regression:

1. Extract subject timeseries
2. Extract subject maps
1. Regress group maps into each subject’s 4D data to find subject-specific timecourses.

2. Regress these timecourses back into the 4D data to find subject-specific spatial maps.
Dual Regression

Group ICA map

Example subject maps derived from dual regression
Running dual_regression

FSL command line tool, combining:

- DR to create subject-wise estimates
- Group comparison using randomise
Group comparison

- Collect maps and perform voxel-wise test (e.g. randomisation test on GLM)

- Can now do voxelwise testing across subjects, separately for each original group ICA map

- Can choose to look at strength-and-shape differences
Group analysis on maps

- can use the Glm tool (Glm_gui on mac) to create GLM design and contrast matrices
Dual regression outputs

- **dr_stage1_subject[#SUB].txt** - the timeseries outputs of stage 1 of the dual-regression.
Dual regression outputs

- dr\_stage1\_subject[#SUB].txt - the timeseries outputs of stage 1 of the dual-regression.

- dr\_stage2\_subject[#SUB].nii.gz - the spatial maps outputs of stage 2 of the dual-regression.

- dr\_stage2\_ic[#ICA].nii.gz - the re-organised parameter estimate images
Dual regression outputs

- dr_stage1_subject[#SUB].txt - the timeseries outputs of stage 1 of the dual-regression.

- dr_stage2_subject[#SUB].nii.gz - the spatial maps outputs of stage 2 of the dual-regression.

- dr_stage2_ic[#ICA].nii.gz - the re-organised parameter estimate images

- dr_stage3_ic[#ICA]_tstat[#CON].nii.gz - the output from randomise (corrected for multiple comparisons across voxels but not across #components!!)
Using template maps

- from the data, using ICA
- use all data to get unbiased templates
- use independent control group
- will model signals and artefacts
- use existing template

http://www.fmrib.ox.ac.uk/analysis/research
That’s all folks