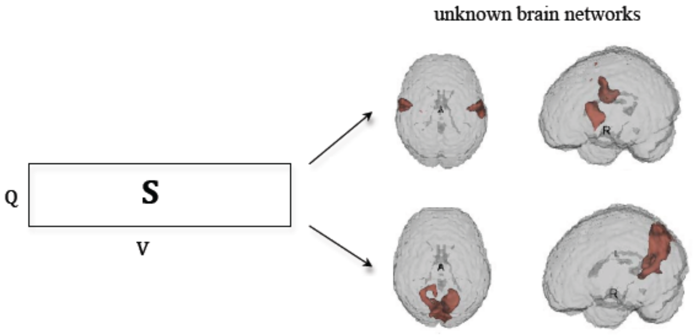
Statistical analysis of brain images using matrix decompositions

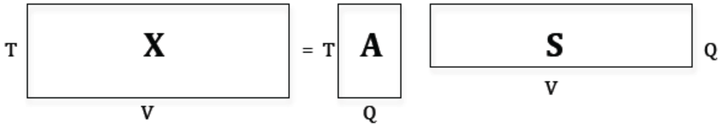
Ani Eloyan, PhD

Assistant Professor
Department of Biostatistics
School of Public Health
Brown University

April 9, 2016

Independent Component Analysis





A - temporal mixtures. Many methods assume $Q = T$.

Group Independent Component Analysis

$$
\begin{array}{lcccc}
\text{Subject 1} & \boxed{\quad \boldsymbol{X}_1 \quad} & & \boxed{\boldsymbol{A}_1} & \boxed{\quad \boldsymbol{S} \quad} \\
\text{Subject 2} & \boxed{\quad \boldsymbol{X}_2 \quad} & = & \boxed{\boldsymbol{A}_2} & \\
 & \dots & & \dots & \\
\text{Subject I} & \boxed{\quad \boldsymbol{X}_{\mathrm{I}} \quad} & & \boxed{\boldsymbol{A}_{\mathrm{I}}} &
\end{array}
$$

- Reconstruct each row of $\boldsymbol{S}$ in 3D.
- Each 3D image is a brain network (Calhoun, 2001).

ICA in Brain Research

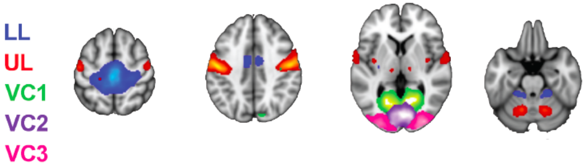
Children with Autism Spectrum Disorder (ASD) have difficulties performing motor tasks.

- Autism trait severity using total Raw SRS score.
- Imitation ability.
- Overall skilled gesture performance using praxis exam scores.

Goals:

- Is visual-motor synchrony different in ASD?
- Is visual-motor synchrony associated with imitation ability?

ICA based Connectivity Analysis - KKI



Motor system

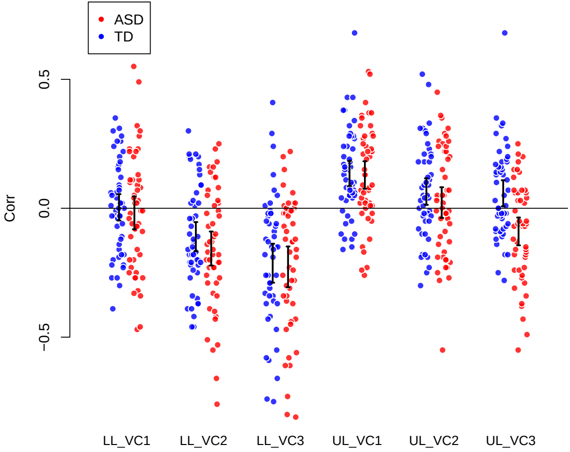
- dorsomedial lower limb areas ("LL")
- more lateral upper limb areas ("UL")

Visual components

- visual processing areas ("VC1" and "VC2")
- lateral occipital cortex ("VC3")

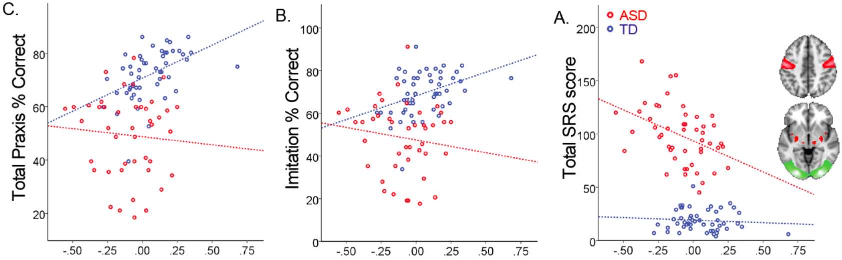
Estimated by ICA for 50 children with ASD and 50 controls.
Age 8-12 years.

ICA based Connectivity Analysis - KKI



Nebel, M.B., Eloyan, A., Nettles, C., Ament, K., Sweeney, K., Ward, R., Barber, A.D., Choe. A., Pekar, J.J., and Mostofsky, S.H. (2016) Reduced intrinsic visual-motor synchrony relates to autism severity. Biological Psychiatry.

ICA based Connectivity Analysis - KKI



Nebel, M.B., Eloyan, A., Nettles, C., Ament, K., Sweeney, K., Ward, R., Barber, A.D., Choe. A., Pekar, J.J., and Mostofsky, S.H. (2016) Reduced intrinsic visual-motor synchrony relates to autism severity. Biological Psychiatry.

Validation

ABIDE
Autism Brain Imaging
Data Exchange

indi
International Neuroimaging
Data-Sharing Initiative

Data from UM was used for validation.

Independent Component Analysis

The ICA model and assumptions.

$$\boldsymbol{X} = \boldsymbol{AS} + \boldsymbol{E}$$

- The components $S_1, \ldots, S_Q$ are statistically independent.
- The mixing matrix $\boldsymbol{A}$ is nonsingular.
- At most one of the components S_q is Gaussian.

When $\boldsymbol{E} = 0$ the model is called noise-free.

Contrast Function

A mapping $\phi(\cdot)$ from the set of densities $\{f_s, \boldsymbol{s} \in \mathbb{R}^n\}$ to $\mathbb{R}$ is called a contrast function (Comon (1994)) if it satisfies the following requirements

- $\phi(f_s) = \phi(f_{\boldsymbol{P}s})$ where $\boldsymbol{P}$ is a permutation matrix.
- $\phi(f_s) = \phi(f_{\Lambda s})$ where Λ is a diagonal invertible matrix.
- $\phi(f_{As}) \leq \phi(f_s)$ if all elements of s are independent and the matrix $\boldsymbol{A}$ is invertible.

Commonly used contrast functions involve kurtosis, negentropy and mutual information.

Contrast Function

Negentropy as a measure of nongaussianity.

$$J(f_x) = H(f_z) - H(f_x),$$

where $E(\boldsymbol{z}) = E(\boldsymbol{x})$ and $\mathrm{cov}(\boldsymbol{z}) = \mathrm{cov}(\boldsymbol{x})$, $H(\cdot)$ differential entropy.

Hyvarinen (1997) negentropy of $\boldsymbol{s}$ can be approximated by

$$J_G(f_s) = [E_s(G(s)) - E_z(G(z))]^2,$$

where z is a Gaussian random variable with mean zero and variance one, $G(\cdot)$ is a nonquadratic function.

FastICA

The R package `fastICA` is based on this method using

$$G_1(u) = \log \ \cosh \ au, \qquad\qquad G_2(u) = \exp(-\frac{u^2}{2}),$$

where a is a constant such that $1 \leq a \leq 2$.

More information in Hyvarinen, Karhunen, and Oja (2003).

Unified Framework for Group ICA

Guo and Pagnoni (2008) and Guo (2011) - EM based algorithm.

- Assume the mixing matrix $\boldsymbol{A}$ is square,
- Define a structure for the mixing matrix,
- Model densities of underlying sources using Gaussian mixtures,
- Parameter estimation via EM-algorithm.

Shi and Guo (2016) incorporate covariates within group ICA.

ProDenICA, Distance Covariance, LCA

ProDen ICA proposed by Hastie and Tibshirani (2002)

- Model densities of underlying sources using exponentially tilted Gaussian densities,
- Estimate the mixing matrix using a fixed point algorithm.

ICA via Distance Covariance, Matteson and Tsay (2011)

- Estimate components by targeting independence,
- Define independence via Distance Covariance.

Likelihood Component Analysis, Risk, et. al (2016)

- Allows for non-square mixing matrices,
- Options for modeling densities of underlying sources.

Group ICA

$$\boldsymbol{S}(q, v) = \boldsymbol{W}(q, .)\boldsymbol{X}(., v),$$

$\boldsymbol{W} = \boldsymbol{A}^{-1}$, $\boldsymbol{W}(q, .)$ - *qth* row of $\boldsymbol{W}$, $\boldsymbol{X}(., v)$ - *vth* column of $\boldsymbol{X}$.

$$\boldsymbol{S}(q, 1), \ldots, \boldsymbol{S}(q, V) \sim f_q(\cdot).$$

The likelihood function for ICA model

$$L(\boldsymbol{S}) = \prod_{v=1}^{V} \prod_{q=1}^{Q} f_q[\boldsymbol{S}(q, v)],$$

$$L(\boldsymbol{W}, f) = \prod_{v=1}^{V} \prod_{q=1}^{Q} f_q[\sum_{l=1}^{Q} \boldsymbol{W}(q, l)\boldsymbol{X}(l, v)].$$

Estimate the matrix $\boldsymbol{W}$ and densities f_q given the observed $\boldsymbol{X}$.

Mixtures for Estimating Densities

We parameterize the density of $\boldsymbol{S}_q$ as a mixture density:

$$f_q(s) = \sum_{j=1}^{N} \theta_{qj} \phi\left(\frac{s - \mu_{qj}}{\sigma_q}\right) \frac{1}{\sigma_q},$$

where $\phi(\cdot)$ is the standard normal density function.

- The means μ_{qj} and the standard deviations σ_j - fixed.
- The estimation of $\theta_{q1}, \ldots, \theta_{qN}$ is performed via a modified EM algorithm.

Eloyan, A. and Ghosh, S.K. (2011) *Smooth Density Estimation with Moment Constraints Using Mixture Distributions.* J. of Nonparametric Statistics. 23, 2, 513-531.

Independent Component Analysis

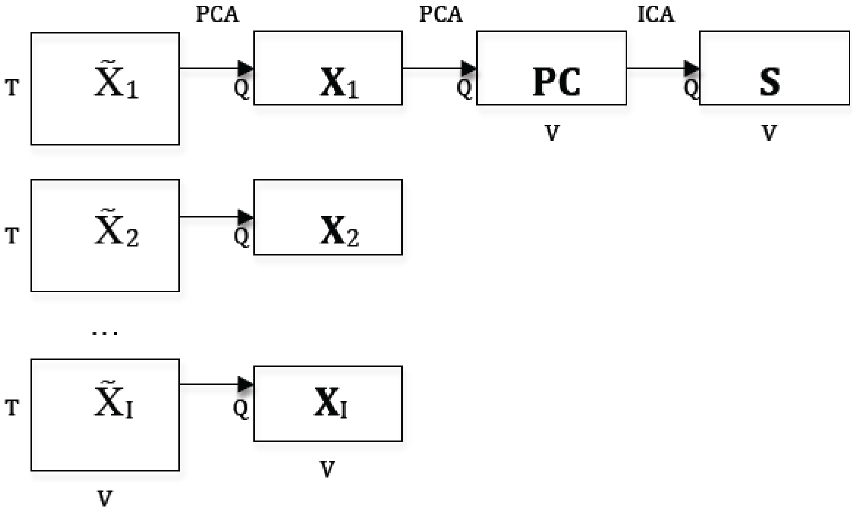
The log-likelihood of ICA is obtained as

$$l(\boldsymbol{W}, \widehat{f}) = \sum_{v=1}^{V} \sum_{q=1}^{Q} \log\{\hat{f}_q(\sum_{l=1}^{Q} x_{vl} w_{lq})\} + V \log|\det \boldsymbol{W}|.$$

Estimate the mixing matrix $\boldsymbol{W} = \boldsymbol{A}^{-1}$ and the densities $\hat{f}_q$ via an iterative optimization algorithm.

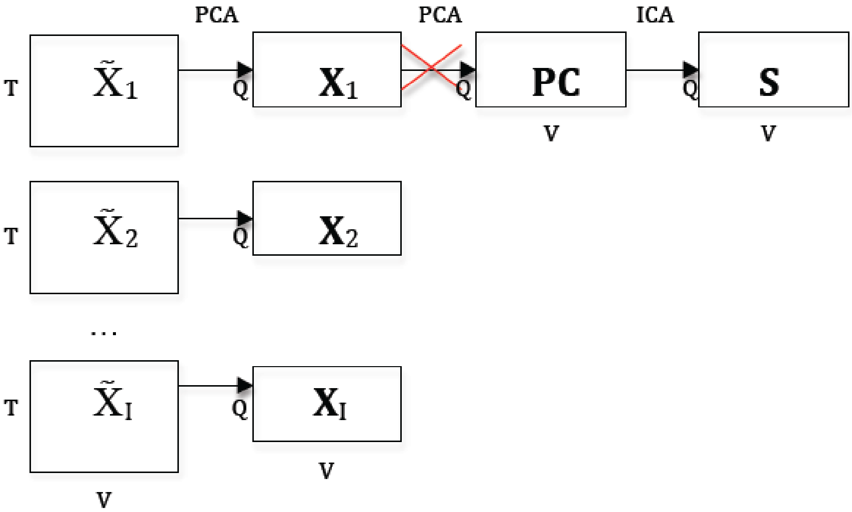
Eloyan, A. and Ghosh, S.K. (2013) *A Semiparametric Approach to Source Separation using Independent Component Analysis.* Comp. Stat. and Data Analysis. 58, 383-396.

Group Independent Component Analysis



Two-stage singular value decomposition.

Group Independent Component Analysis



Two-stage singular value decomposition.

Iterative Algorithm, HDICA

1. $\boldsymbol{S}_i = \widehat{\boldsymbol{W}_i}\boldsymbol{X}_i$
2. Estimate the weights of the density for each independent component. $\boldsymbol{\theta}_q$ is estimated by an EM algorithm.
3. Compute the derivative and Hessian for the log-likelihood.

$$L(\widehat{\boldsymbol{W}}) = \sum_{i=1}^{\mathrm{I}} \sum_{v=1}^{V} \sum_{q=1}^{Q} \log[f_q(\widehat{\boldsymbol{W}}_{iq}\boldsymbol{X}_{iv})] + V \log|\det \widehat{\boldsymbol{W}_i}|.$$

4. $\widehat{\boldsymbol{W}_i}^{new} = \widehat{\boldsymbol{W}_i} - L''(\widehat{\boldsymbol{W}_i})^{-1} L'(\widehat{\boldsymbol{W}_i})$
5. Stopping rule

$$\max \|\widehat{\boldsymbol{W}_i} - \widehat{\boldsymbol{W}_i}^{new}\| < \delta$$

Eloyan, A., Crainiceanu, C.M., and Caffo, B.S. (2013) Likelihood based population independent component analysis. *Biostatistics*. 14, 3, 514-527.
Chen, S. Huang, L., Qui, H., Nebel, M.B., Mostofsky, S.H., Pekar, J.J., Lindquist, M.A., Eloyan, A., and Caffo, B.S. (submitted) Parallel Group Independent Component Analysis for Massive fMRI Data Sets.

Bayesian Independent Component Analysis

The noisy ICA model:

$$\boldsymbol{X} = \boldsymbol{A}\boldsymbol{S} + \boldsymbol{E},$$

$$X_{tv}|\boldsymbol{A}, \boldsymbol{S}, \sigma_e \sim N(\boldsymbol{A}_{t.}\boldsymbol{S}_{.v}, \sigma_e^2)$$

Proposed priors:

$$\sigma_e^2|\alpha_e, \beta_e \sim \mathit{InverseGamma}(\alpha_e, \beta_e).$$

$$S_{qv}|Z_{qv} = k \sim N(\mu_k, \sigma_N^2),$$

$$P[Z_{qv} = k] = \theta_{qk},$$

$$\boldsymbol{\theta}_q = (\theta_{q1}, \theta_{q2}, \ldots, \theta_{qN})|\alpha \sim \mathit{Dirichlet}(\alpha, \alpha, \ldots, \alpha),$$

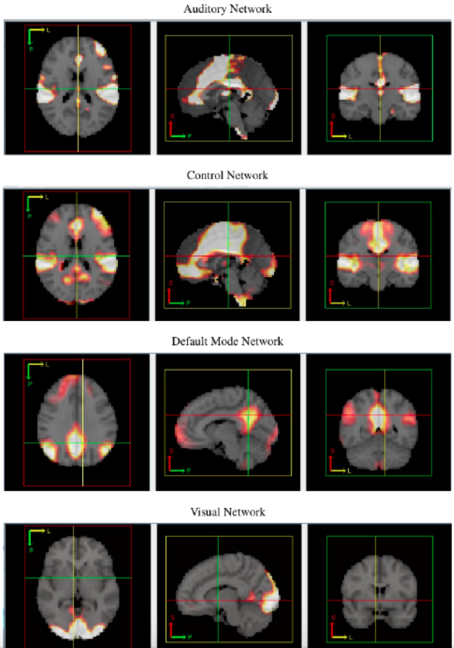
where $v = 1, \ldots, V$, $t = 1, \ldots, Q$ and $k = 1, 2, \ldots, N$.

The 1000 Functional Connectomes Project Dataset

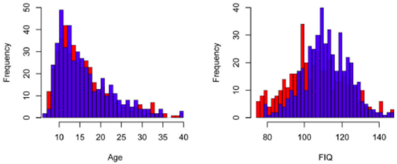
- More than 1400 scans available online.
- The scans are collected using a 3T scanner.
- For the subset used in this analysis the number of time points was $T = 119$.
- Standard image processing was performed to register the data to the MNI standard brain space.

$\boldsymbol{W}_i$ and $\boldsymbol{S}$ are estimated via the parallel HDICA algorithm.

Results for 301 Subjects

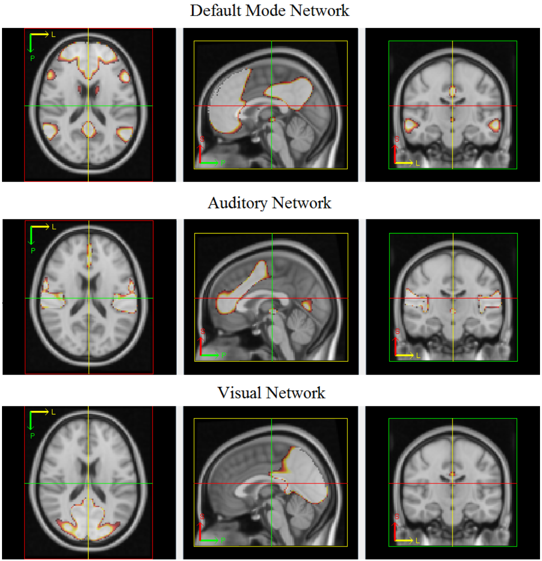


ICA based Connectivity Analysis - ABIDE



- 379 ASD.
- 400 typically developing
- For the subset used in this analysis the number of time points was $T = 220$.
- Standard image processing was performed to register the data to the MNI standard brain space.

ICA based Connectivity Analysis - ABIDE



Summary and Extensions

- Connectivity using ICA in Autism looking at motor function.
- New methods for finding brain networks for large groups of fMRI data.
- Extensions of existing methods for novel types of data.

Thank you and Acknowledgments.

SMART group
Taki Shinohara, PhD
Brian Caffo, PhD
Ciprian Crainiceanu, PhD
Daniel Reich, MD, PhD
Shaojie Chen
Elizabeth Sweeney
Stewart Mostofsky, MD
James Pekar, PhD
Mary Beth Nebel, PhD

BROWN
School of Public Health