Simultaneous Non-Gaussian Component Analysis (SING) for Data Integration in Neuroimaging

Benjamin B. Risk

Department of Biostatistics & Bioinformatics
Rollins School of Public Health
Emory University

benjamin.risk@emory.edu

Joint work

- Joint work with Irina Gaynanova, Department of Statistics, Texas A&M.

Data integration

- Motivating principle: combining information across datasets leads to a more accurate understanding of the underlying biology.
- Subject scores: summarize a subject's brain phenotype from multiple neuroimaging data types.
- Scientific aim: Data integration in cognition studies [Lerman-Sinkoff et al., 2017].

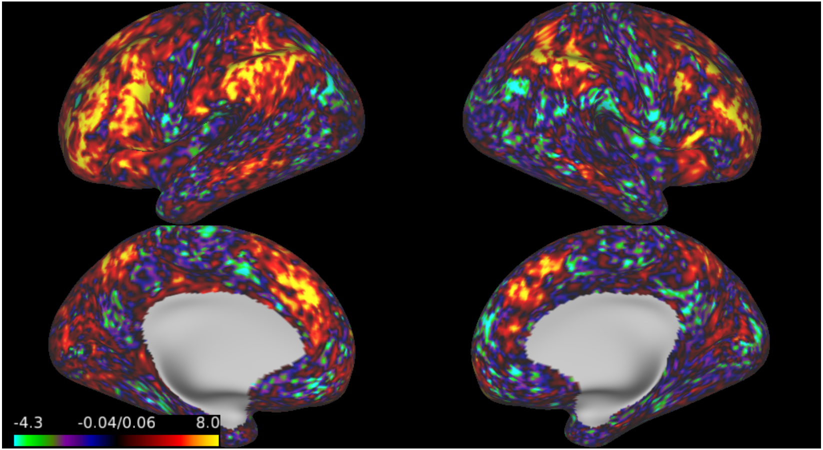
  Are cognitive task regions related to spontaneous brain activity?

Application: Human Connectome Project

- Utilize data from the Human Connectome Project (HCP).
- We integrate working memory task fMRI activation maps for 60,000 vertices with resting-state fMRI correlation matrices for 379 regions.
- Scientific questions:
  1. Joint structure: Are there associations between working memory task activation maps and resting-state functional connectivity?
  2. Relationship to external data: Can we predict fluid intelligence from the joint subject scores?

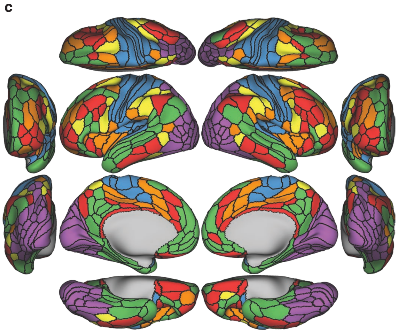
Example activation maps

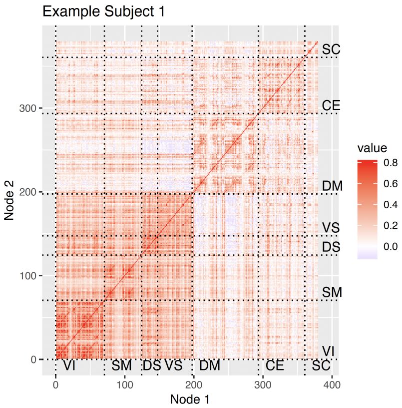
Input X: z-stat at each vertex from 2-back versus 0-back task contrast [Barch et al., 2013]. Vectorized activation maps: $p_x \approx 60{,}000$, $n = 996$.



Functional Connectivity: Resting-state correlations

Input Y: Fisher-transformed rs correlations between regions defined by multi-modal parcellation [Glasser et al., 2016, Akiki and Abdallah, 2018].





Multimodal methods using ICA

Methods for multimodal analysis based on ICA:

- **Joint ICA** [Calhoun et al., 2009]: PCA on $[X_s \; Y_s] \in \mathbb{R}^{n \times (p_x + p_y)}$, then ICA:

$$
\begin{aligned}
[X_s \; Y_s] &= U_{(r_J)} D_{(r_J)} V_{(r_J)}^\top + U_{(-r_J)} D_{(-r_J)} V_{(-r_J)}^\top \\
[X_s \; Y_s] &= (U_{(r_J)} D_{(r_J)} W^\top)(W V_{(r_J)}^\top) + U_{(-r_J)} D_{(-r_J)} V_{(-r_J)}^\top \\
[X_s \; Y_s] &= M_J [S_x \; S_y] + U_{(-r_J)} D_{(-r_J)} V_{(-r_J)}^\top
\end{aligned}
$$

- Estimates joint subspace, then an additional rotation of joint subspace to maximize "independence" (non-Gaussianity).
- **Multimodal CCA with Joint ICA** [Sui et al., 2011]: PCA on separate datasets, CCA on PC scores, then joint ICA on concatenated loadings, M_{Jx}, M_{Jy}.
- Both these methods use PCA to estimate the subspace – **removes low variance information**.

Joint & **I**ndividual **V**ariation **E**xplained [Lock et al., 2013, Feng et al., 2018]: shared information in subject score subspaces, matrix decomposition focuses on variance.

We focus on non-Gaussian component analysis – information measured using third and fourth moments.

SImultaneous **N**on-**G**aussian component model (SING):

$$\mathrm{X}_c = \mathrm{M}_J \mathrm{D}_x \mathrm{S}_{Jx} + \mathrm{M}_{Ix}\mathrm{S}_{Ix} + \mathrm{M}_{Nx}\mathrm{N}_x,$$
$$\mathrm{Y}_c = \mathrm{M}_J \mathrm{D}_y \mathrm{S}_{Jy} + \mathrm{M}_{Iy}\mathrm{S}_{Iy} + \mathrm{M}_{Ny}\mathrm{N}_y.$$

- $\mathrm{M}_J \in \mathbb{R}^{n \times r_J}$, $||\mathrm{m}_{J\ell}||_2 = 1$, D_x diagonal.
- $\mathrm{S}_{Jx} : r_J \times p_x$, $\mathrm{S}_{Jx}\mathrm{S}_{Jx}^\top = p_x \mathrm{I}_{r_J}$, $\mathrm{S}_{Jx} 1_{p_x} = 0$.
- $\mathrm{S}_{Ix} : (r_x - r_J) \times p_x$.
- Gaussian components: $\mathrm{N}_x : (n - r_x) \times p_x$.

Objective function to SING

Let X_w and Y_w be whitened X and Y, respectively. Let $\widehat{L}_X$, $\widehat{L}_Y$ be the corresponding whitening matrices. We consider

$$\underset{U_X, U_Y}{\text{argmin}} - \sum_{\ell=1}^{r_x} -f(u_{X\ell}^\top X_w) - \sum_{\ell=1}^{r_y} f(u_{Y\ell}^\top Y_w) + \rho \sum_{\ell=1}^{r_J} d(\widehat{L}_X^{-1} u_{X\ell}, \widehat{L}_Y^{-1} u_{Y\ell})$$

subject to $\quad U_X U_X^\top = I_{r_x},\ U_Y U_Y^\top = I_{r_y},$

$$d(x, y) = \left\| \frac{xx^\top}{\|x\|_2^2} - \frac{yy^\top}{\|y\|_2^2} \right\|_F^2.$$

We use the Jarque-Bera test statistic [Virta et al., 2016]:

$$f(S_\ell) = 0.8(\mathbb{E}_p\, s_{j\ell}^3)^2 + 0.2(\mathbb{E}_p\, s_{j\ell}^4 - 3)^2.$$

Use an algorithm based on [Wen and Yin, 2012] for feasible updates.

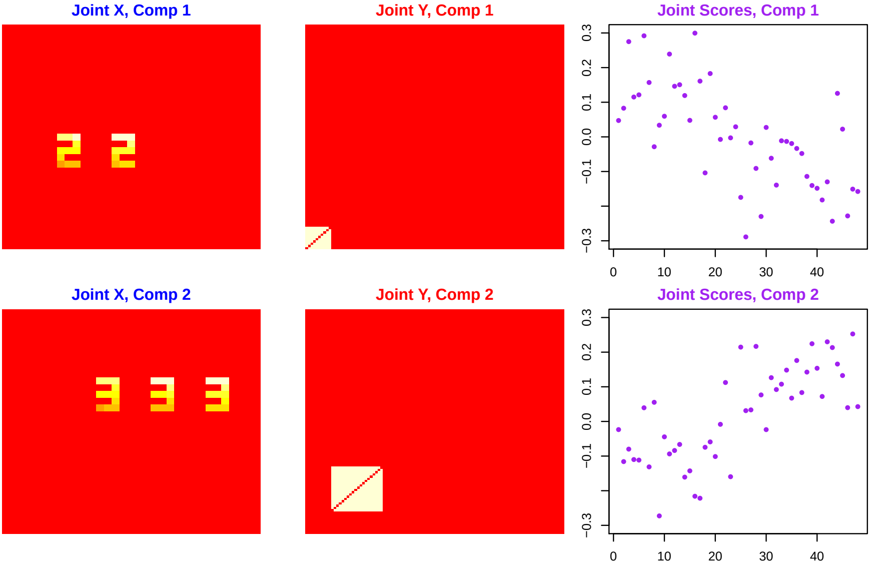
Estimating joint rank

Remark

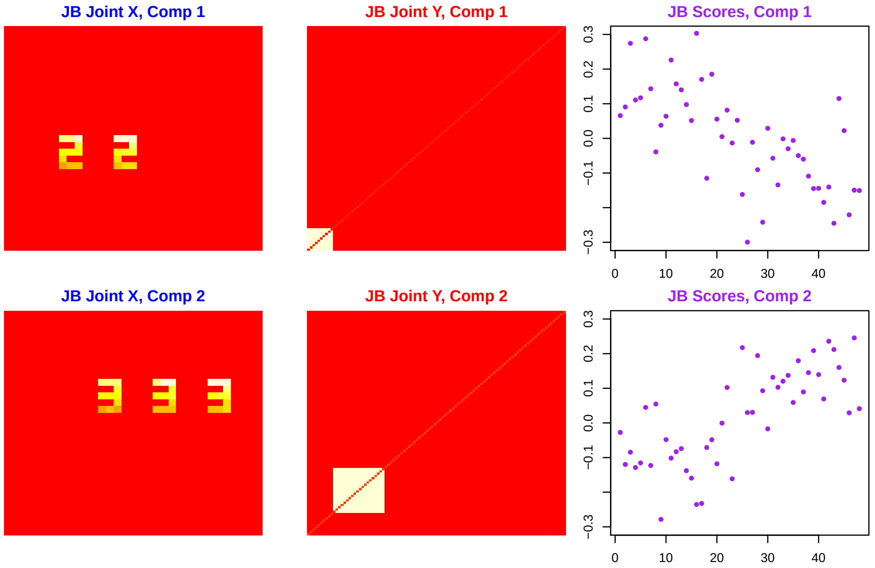
Due to the uniqueness of the linear non-Gaussian component analysis decompositions (up to signed permutations), the joint components are a subset of the components from the separate decompositions.

1. Perform separate LNGCA.
2. Match columns of M_x and M_y based on their correlation.
3. Estimate r_J using an FWER-controlled permutation test based on the correlation between M_x and permuted rows of M_y.

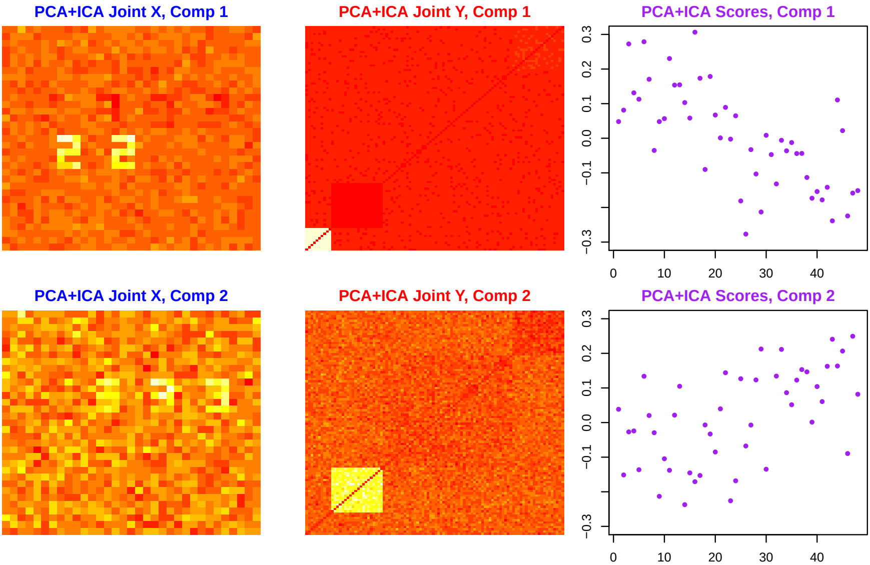
Toy Example: Truth



Toy Example: SING Estimate



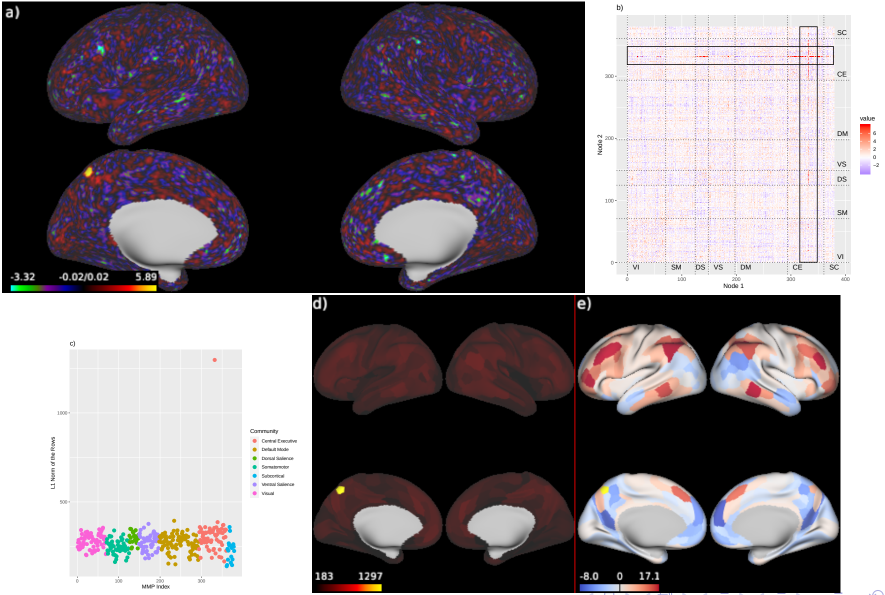
Toy Example: Joint ICA Estimate



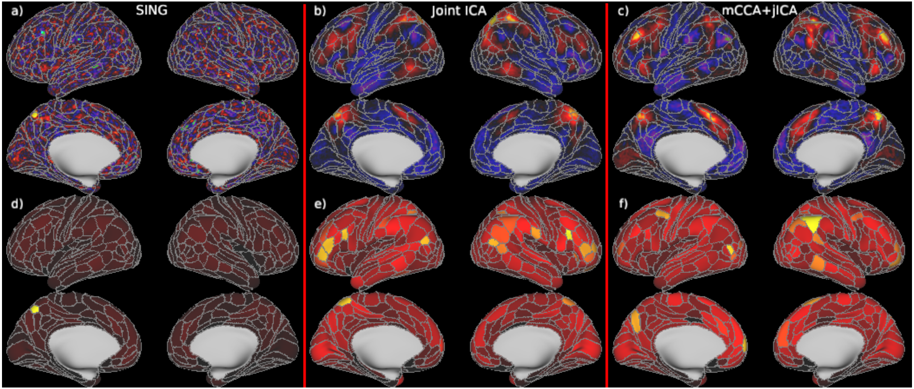
Activation Maps and Networks

- Scientific questions:
  1. Joint structure: Are there associations between working memory task activation maps and resting-state functional connectivity?
     - Estimate with 100 initializations, choose argmax, does this a second time, retain reliable components.
     - 26 joint components were selected.
     - In SING, ρ chosen to result in correlations ranging > 0.99 (here, $\rho = JB/10$).
  2. Relationship to external data: Can we predict fluid intelligence from the joint subject scores?
     - Component 24 most strongly related to fluid intelligence (t=5.10).

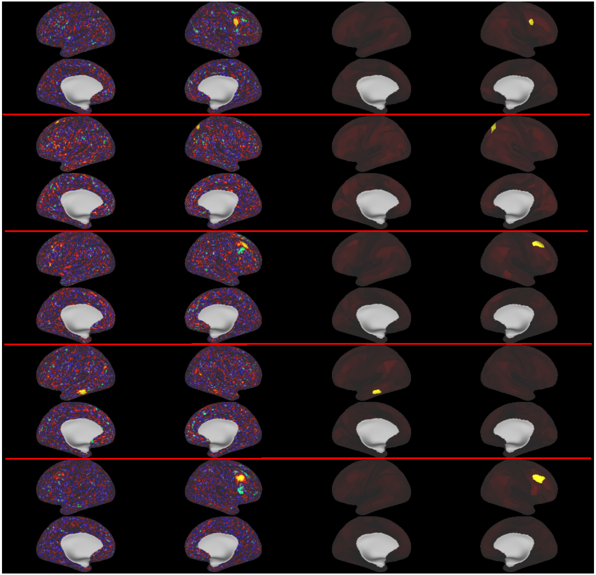
Joint component and fluid intelligence



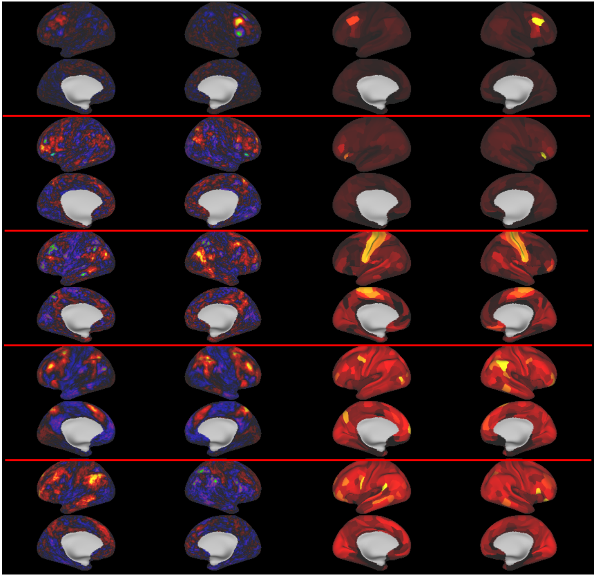
Comparing SING, Joint ICA, mCCA+jICA



Additional SING component loadings



Additional mCCA+jICA component loadings



Take home

1. We propose a new matrix decomposition for shared structure across datasets with non-Gaussian features.
2. Improves estimation of joint subject scores and loadings compared to popular methods in neuroimaging.
3. In contrast to existing approaches, joint loadings in SING capture spatially coinciding features in the working memory task and rs-fMRI.

Acknowledgments

- Thank you!
- Data were provided (in part) by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657).

References I

Akiki, T. J. and Abdallah, C. G. (2018).
Determining the Hierarchical Architecture of the Human Brain Using Subject-Level Clustering of Functional Networks.
bioRxiv, page 350462.

Barch, D. M., Burgess, G. C., Harms, M. P., Petersen, S. E., Schlaggar, B. L., Corbetta, M., Glasser, M. F., Curtiss, S., Dixit, S., Feldt, C., and others (2013).
Function in the human connectome: Task-{fMRI} and individual differences in behavior.
NeuroImage, 80:169–189.

Calhoun, V. D., Liu, J., and Adali, T. (2009).
A review of group {ICA} for {fMRI} data and {ICA} for joint inference of imaging, genetic, and {ERP} data.
Neuroimage, 45(1):S163–S172.

Feng, Q., Jiang, M., Hannig, J., and Marron, J. (2018).
Angle-based joint and individual variation explained.
Journal of multivariate analysis, 166:241–265.

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C. F., Jenkinson, M., and others (2016).
A multi-modal parcellation of human cerebral cortex.
Nature, 536(7615):171–178.

Lerman-Sinkoff, D. B., Sui, J., Rachakonda, S., Kandala, S., Calhoun, V. D., and Barch, D. M. (2017).
Multimodal neural correlates of cognitive control in the Human Connectome Project.
NeuroImage, 163.

Lock, E. F., Hoadley, K. A., Marron, J. S., and Nobel, A. B. (2013).
Joint and individual variation explained (JIVE) for integrated analysis of multiple data types.
The annals of applied statistics, 7(1):523.

References II

Sui, J., Pearlson, G., Caprihan, A., Adali, T., Kiehl, K. A., Liu, J., Yamamoto, J., and Calhoun, V. D. (2011).
Discriminating schizophrenia and bipolar disorder by fusing fmri and dti in a multimodal cca+ joint ica model.
Neuroimage, 57(3):839–855.

Virta, J., Nordhausen, K., and Oja, H. (2016).
Projection Pursuit for non-Gaussian Independent Components.
arXiv preprint arXiv:1612.05445.

Wen, Z. and Yin, W. (2012).
A feasible method for optimization with orthogonality constraints.
Mathematical Programming, 142(1-2):397–434.