

Data Integration in Neuroimaging Studies

Benjamin B. Risk

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

benjamin.risk@emory.edu

github.com/thebrisklab



Table of Contents

- 1 Overview of Data Integration
- 2 Probabilistic Joint and Individual Variation Explained (ProJIVE)
- 3 Simultaneous Non-Gaussian Component Analysis (SING) for Data Integration in Neuroimaging

Overview of Data Integration

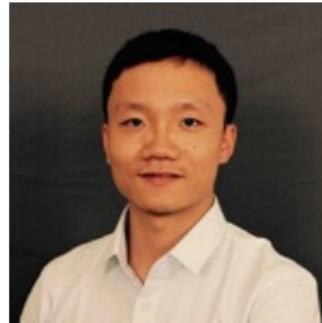
Data fusion

- Goal: find information shared by multiple datasets collected on the same participants. Multimodal data, multiview data, multiblock data.
- **Subject scores** can be used to summarize a subject's phenotype from multimodal data (dimension reduction).
- Subject scores can be investigated as possible biomarkers in behavior and neurological disorders [Sui et al., 2011].
- Features derived from neuroimaging and behavioral/clinical data.
 - Part 1: brain morphometry and cognitive test batteries from the Alzheimer's Disease Neuroimaging Initiative.
- Multimodal data in imaging: task fMRI, resting-state fMRI, diffusion-weighted MRI, T1 and T2 structural images.
 - Part 2: working memory task fMRI and resting-state fMRI from the Human Connectome Project.

Probabilistic Joint and Individual Variation Explained (ProJIVE)

Joint work

Working manuscript: Raphiel Murden (Biostatistics, Emory), Ganzhong Gavin Tian (Biostatistics, Emory), Deqiang Qiu (Department of Radiology, Emory University School of Medicine; Department of Biomedical Engineering, Georgia Tech) and Benjamin Risk.



Joint & Individual Variation Explained

[Lock et al., 2013, Feng et al., 2018]: shared information in subject score subspaces $\in \mathbb{R}^n$.

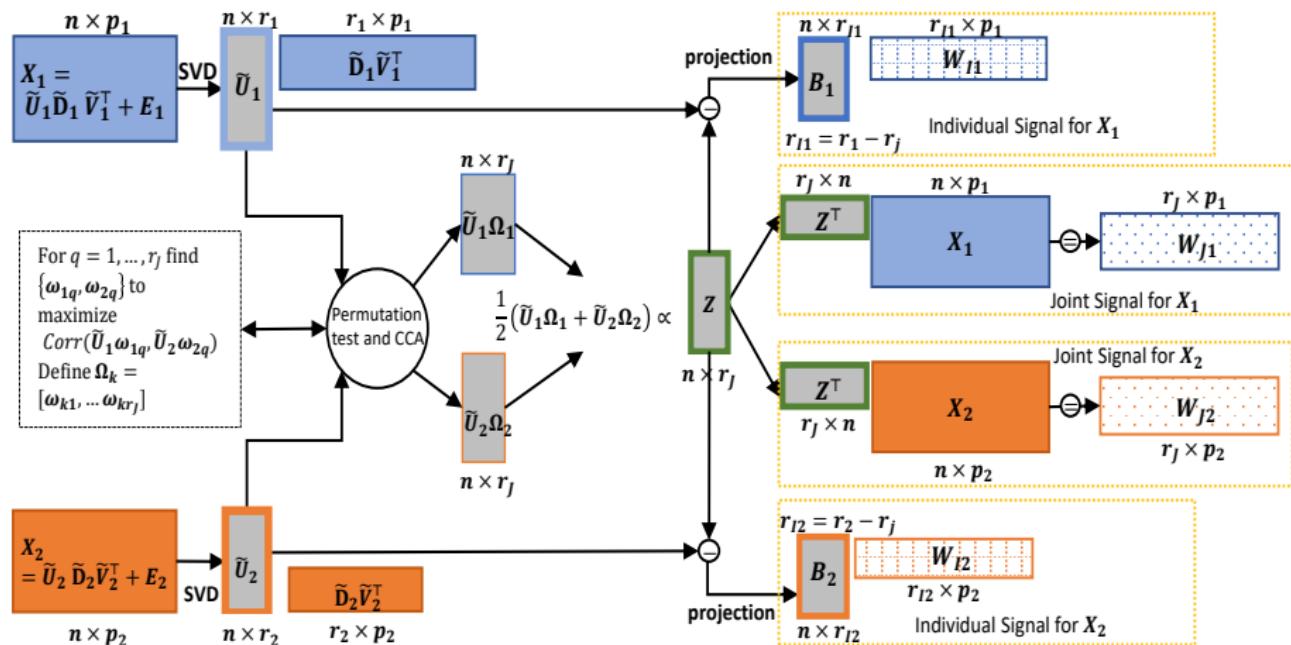
Consider $\mathbf{X}_k \in \mathbb{R}^{n \times p_k}$, where p_k is the number of features/variables in the k^{th} data set for $k = 1, \dots, K$.

$$\mathbf{X}_k = \mathbf{J}_k + \mathbf{A}_k + \mathbf{E}_k, \text{ for } k = 1, \dots, K.$$

- Signal $\mathbf{B}_k = \mathbf{J}_k + \mathbf{A}_k$ is low rank
- $\text{col}(\mathbf{J}_k) = \text{col}(\mathbf{J}_{k'})$ for all $k, k' \in \{1, \dots, K\}$
- $\text{col}(\mathbf{J}_k) \perp \text{col}(\mathbf{A}_k)$
- $\bigcap_{k=1}^K \text{col}(\mathbf{A}_k) = \mathbf{0}$.
- \mathbf{E}_K is isotropic (singular values are approximately equal).

CCA and AJIVE Schematic

Angle-based JIVE from [Feng et al., 2018] is closely related to CCA (Murden et al Risk in review):



JIVE in Neuroimaging

From Yu, Risk, Zhang, and Marron (2017) using AJIVE:

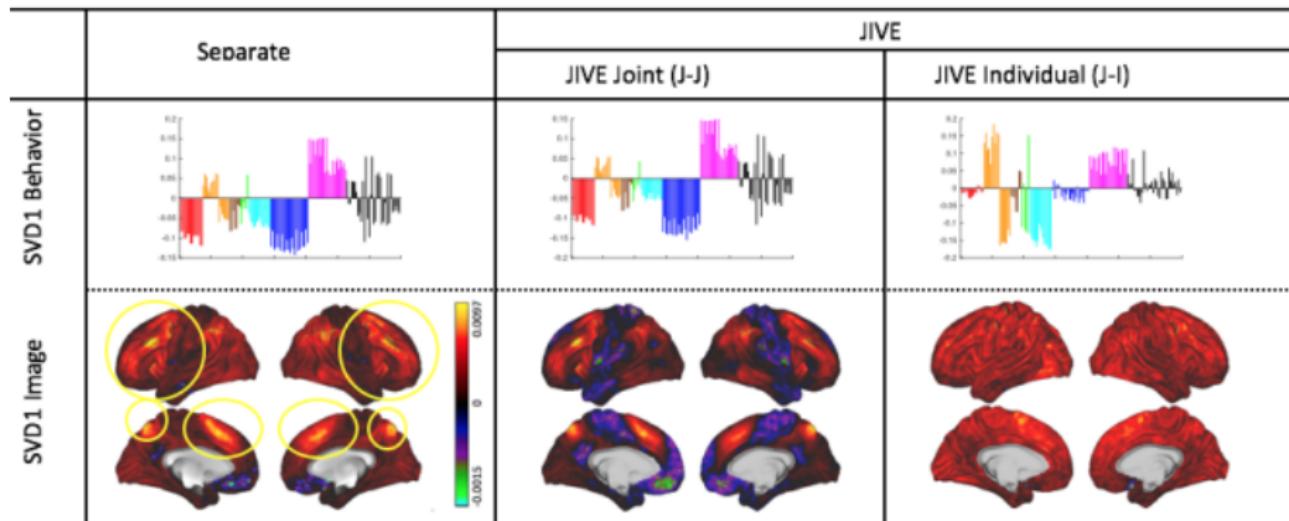


Figure: AJIVE of behavioral data and working memory maps. In behavior, blue and cyan are working memory task performance variables. Compared to conducting separate SVDs on each dataset, AJIVE results in stronger task signal in the working memory map. Unrelated task signal goes to the individual component.

Limitations of JIVE

- Defining joint and individual structure in terms of subspaces can be challenging to understand.
- A probabilistic model may improve interpretation.
- AJIVE estimates individual subspaces after estimation of the joint subspace.
- A framework for simultaneously estimating joint and individual subspaces may improve accuracy.

Probabilistic PCA

Probabilistic PCA [Tipping and Bishop, 1999].

Let $\mathbf{x}_i \in \mathbb{R}^P$, $i = 1, \dots, n$:

$$\mathbf{x}_i = \boldsymbol{\mu} + \mathbf{W}\mathbf{z}_i + \mathbf{e}_i$$

$$\mathbf{z}_i \stackrel{iid}{\sim} N(0, \mathbf{I})$$

$$\mathbf{e}_i \stackrel{iid}{\sim} N(0, \sigma^2 \mathbf{I})$$

Low rank signal.

Factor analysis with isotropic noise.

The MLE corresponds to the classic PCA solution. Even though classic PCA doesn't explicitly "assume" Gaussianity, its solution can be derived from Gaussian assumptions. However, it applies much more generally.

Probabilistic JIVE

We propose Probabilistic JIVE (ProJIVE).

Let $\mathbf{x}_{ik} \in \mathbb{R}^{p_k}$, $k = 1, \dots, K$, $i = 1, \dots, n$.

$$\mathbf{x}_{ik} = \boldsymbol{\mu}_k + \mathbf{W}_{Jk}\mathbf{z}_i + \mathbf{W}_{Ik}\mathbf{b}_{ik} + \boldsymbol{\epsilon}_{ik},$$

$$[\mathbf{z}_i^\top, \mathbf{b}_{i1}^\top, \dots, \mathbf{b}_{iK}^\top]^\top \stackrel{iid}{\sim} N(\mathbf{0}, \mathbf{I}),$$

$$\boldsymbol{\epsilon}_{ik} \stackrel{iid}{\sim} N(\mathbf{0}, \sigma_k^2 \mathbf{I}).$$

Inter-battery factor analysis

It turns out a similar model was proposed in [Tucker, 1958]: inter-battery factor analysis.

An MLE for the joint signal ignoring individual components was derived in [Browne, 1979].

More recently, Bayesian Canonical Correlation analysis with variational inference was proposed [Klami et al., 2013] and the related group factor analysis [Klami et al., 2015].

BCCA includes an inter-battery factor version with isotropic noise, which is the same generative model. They use an automatic relevance determination prior and shrinkage to approximate the block-wise sparsity.

For the case of $K = 2$:

$$\begin{aligned}\text{Cov} \begin{pmatrix} \mathbf{x}_{i1} \\ \mathbf{x}_{i2} \end{pmatrix} &= \begin{pmatrix} \mathbf{W}_{J1}\mathbf{W}_{J1}^\top + \mathbf{W}_{I1}\mathbf{W}_{I1}^\top + \sigma_1^2\mathbf{I} & \mathbf{W}_{J1}\mathbf{W}_{J2}^\top \\ \mathbf{W}_{J2}\mathbf{W}_{J1}^\top & \mathbf{W}_{J2}\mathbf{W}_{J2}^\top + \mathbf{W}_{I2}\mathbf{W}_{I2}^\top + \sigma_2^2\mathbf{I} \end{pmatrix} \\ &= \mathbf{C}.\end{aligned}$$

For $\mu_k = \mathbf{0}$, the log-likelihood of the data is

$$\ell = -\frac{n}{2} \left\{ (p_1 + p_2) \log(2\pi) + \log(|\mathbf{C}|) + \text{tr}(\mathbf{C}^{-1}\mathbf{S}) \right\} \quad (1)$$

where $\mathbf{S} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^\top$.

We derived an EM algorithm to fit this model.

Selecting number of components: use a permutation test on the joint scores from the AJIVE model. LRTs or BIC possible but computationally costly.

ProjIVE Uniqueness

$$\mathbf{W} = \begin{pmatrix} \mathbf{W}_{J1} & \mathbf{W}_{I1} & 0 \\ \mathbf{W}_{J2} & 0 & \mathbf{W}_{I2} \end{pmatrix}$$

Theorem

Suppose $k = 1, 2$ data blocks. Consider an $r_J \times r_J$ non-singular matrix \mathbf{T}_1 and define $\mathbf{W}_{J1}^* = \mathbf{W}_{J1}\mathbf{T}_1^\top$ and $\mathbf{W}_{J2}^* = \mathbf{W}_{J2}\mathbf{T}_1^{-1}$. Then there exists a transformation of the individual + noise components such that $\ell(\mathbf{W}) = \ell(\mathbf{W}^*)$.

Lemma

Suppose $k = 1, \dots, K$ data blocks with $K > 2$. Consider some orthogonal $r_J \times r_J$ matrix \mathbf{O}_J . Consider orthogonal matrices $\mathbf{O}_k \in \mathbb{R}^{r_{Ik} \times r_{Ik}}$. Define $\mathbf{W}_{Jk}^* = \mathbf{W}_{Jk}\mathbf{O}_J$ and $\mathbf{W}_{Ik}^* = \mathbf{W}_{Ik}\mathbf{O}_k$. Then $\ell(\mathbf{W}) = \ell(\mathbf{W}^*)$.

Simulations

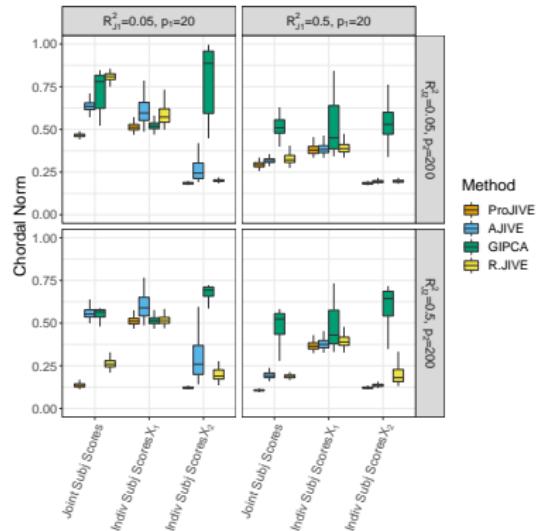
Factors

- ① (a) $p_2 = 20$ and (b) $p_2 = 200$, ($p_1=20$).
- ② Joint Variation Explained in \mathbf{X}_1 : (a) $R_{J1}^2 = 0.05$ and (b) $R_{J1}^2 = 0.5$.
- ③ Joint Variation Explained in \mathbf{X}_2 : (a) $R_{J2}^2 = 0.05$ and (b) $R_{J2}^2 = 0.5$.
- ④ Data generating distributions: (a) Gaussian scores and loadings and (b) mixture of Gaussian joint scores ($\pi_1 = 0.2$, $\mu_1 = -4$, unit variance; $\pi_2 = 0.50$, $\mu_2 = 0$; and $\pi_3 = 0.30$, $\mu_3 = 4$) and Rademacher loadings (joint and individual).

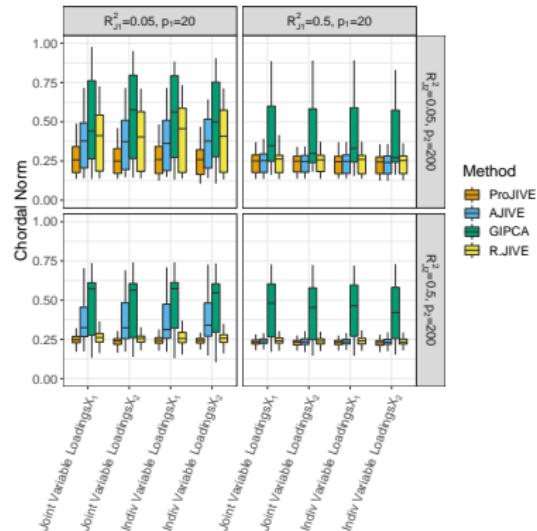
Data generating scheme:

- ① $R_{I1}^2 = R_{I2}^2 = 0.25$.
- ② $(\mathbf{z}_i, \mathbf{b}_{i1}, \mathbf{b}_{i2})$ for $i = 1, \dots, 1000$ according to factor 4.
- ③ $\mathbf{W}_{Jk} = \mathbf{Q}_J \mathbf{R}_{Jk}$, $\mathbf{W}_{Ik} = \mathbf{Q}_{Ik} \mathbf{R}_{Ik}$.
- ④ Rescale such that $\mathbf{X}_k = d_k \mathbf{J}_K + c_k \mathbf{A}_k + \mathbf{E}_k$ for appropriate constants c_k and d_k to achieve factor 2.

Simulations: Gaussian assumptions met



(a)



(b)

Figure: Gaussian scores and loadings (generating mechanism = model). a) subject scores for $p_1 = 20$ and $p_2 = 200$; b) variable loadings. ProJIVE: Our method. AJIVE: [Feng et al., 2018]. R.JIVE: [Lock et al., 2013]. GIPCA: [Zhu et al., 2020].

Simulations: robustness

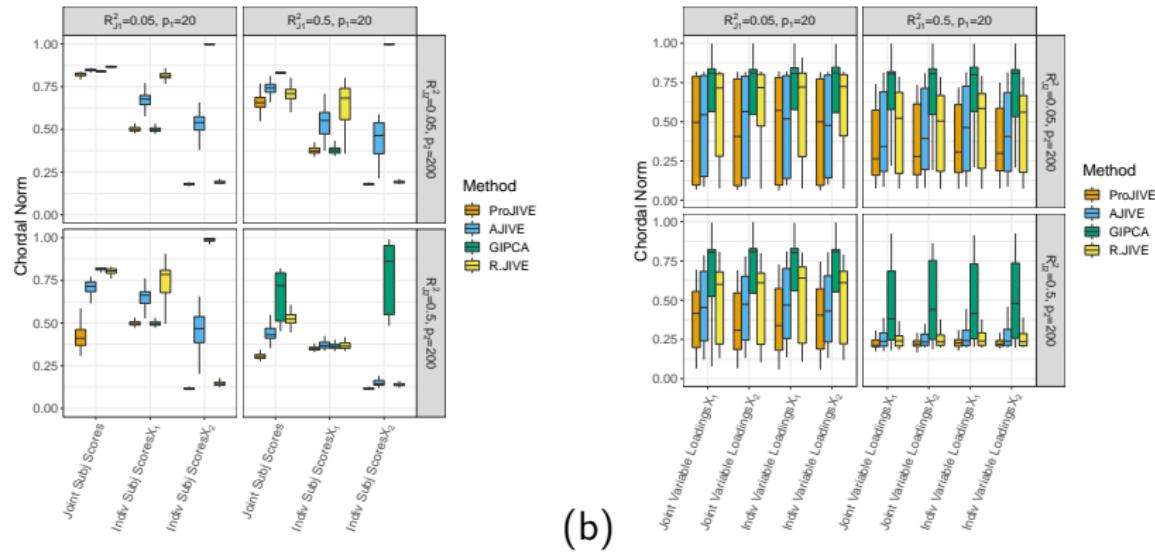


Figure: MOG scores and rademacher loadings (generating mechanism \neq model).
a) subject scores for $p_1 = 20$ and $p_2 = 200$; b) variable loadings.

Brain morphometry and cognition in ADNI

- Preprocessed data from the Alzheimer's Disease Prediction of Longitudinal Evolution (TADPOLE) Challenge (Alzheimer's Disease Neuroimaging Initiative data) in 2018.
- Single time point with $n=587$ adults with full battery of cognition. $p_1 = 22$.
- Dataset 1: standardized cognition variables (CDR-SB, ADAS, MMSE, RAVLT, MOCA, ECOG).
- Dataset 2: cortical thickness, surface area, and cortical volume for 34 ROIs per hemisphere. [Desikan et al., 2006], along with cortical volumes for other regions/structures. $p_2 = 245$.
- For each feature, regressed out age and sex, then standardized residuals.

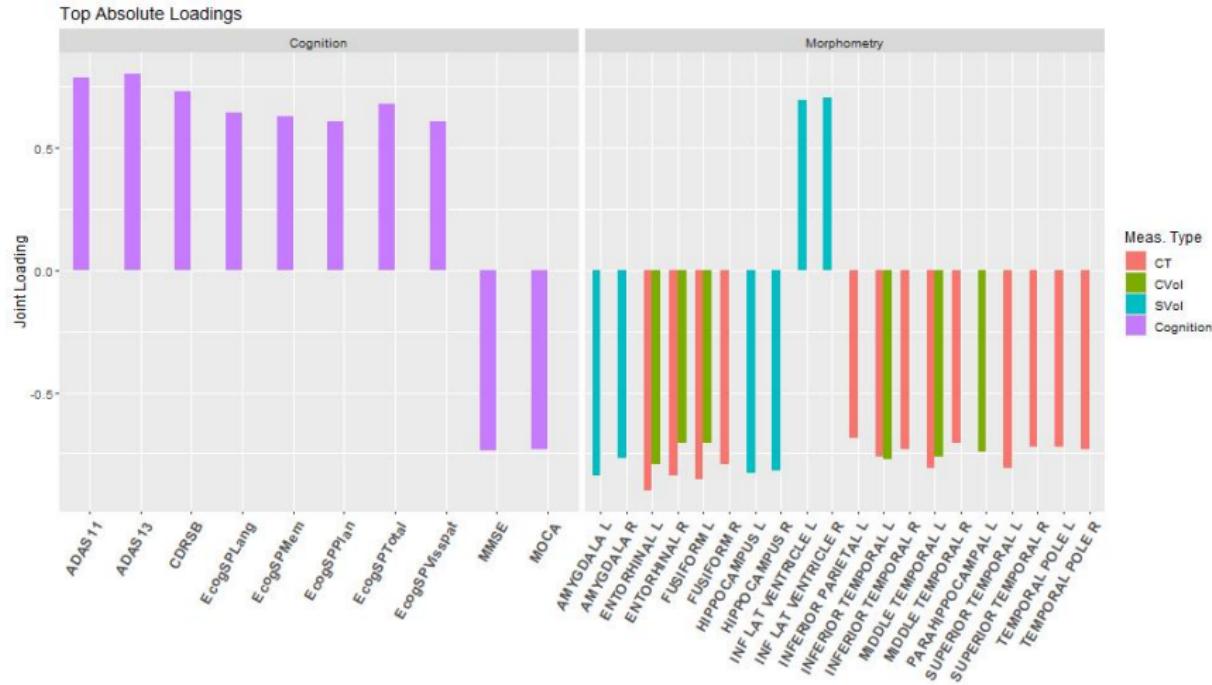
Table: Summary statistics for selected covariates of participants in ADNI-GO and ADNI2.

	AD (N=88)	MCI (N=340)	CN (N=159)	Total (N=587)
	Mean (S.D.) or N (%)			
Age	74.0 (7.92)	71.5 (7.57)	72.8 (5.85)	72.2 (7.25)
Gender				
Female	28 (31.8%)	150 (44.1%)	84 (52.8%)	262 (44.6%)
ApoE4				
0	21 (23.9%)	178 (52.4%)	111 (69.8%)	310 (52.8%)
1	45 (51.1%)	126 (37.1%)	46 (28.9%)	217 (37.0%)
2	22 (25.0%)	36 (10.6%)	2 (1.3%)	60 (10.2%)

Joint and individual variance

- Screeplots: cognition $r_1 = 5$; morphometry $r_2 = 10$.
- Joint rank from permutation test: $r_J = 1$.
- Proportion of variance explained by the joint signal: 0.21 for cognition; 0.10 for morphometry.
- Individual proportions: 0.50 for cognition; 0.36 for morphometry.

Joint Loadings



(a)

Figure: (a) Ten most extreme joint cognition loadings and 90th percentile of absolute joint brain loadings estimated via ProJIVE.

Joint loadings: morphometry

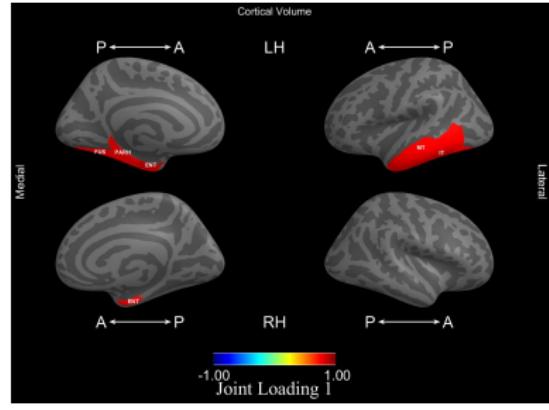
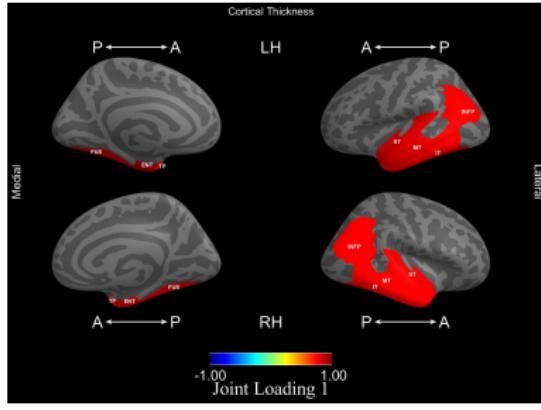
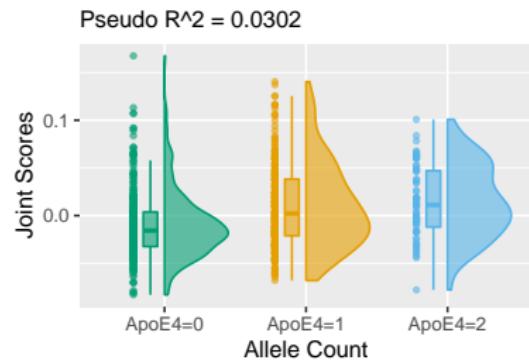


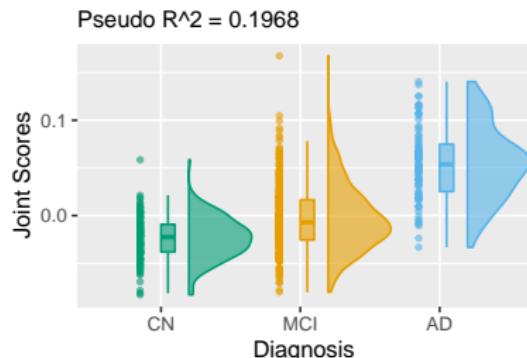
Figure: (b) 90th percentile of absolute brain loadings which occur in measures of cortical thickness. (c) 90th percentile of absolute brain loadings which occur in measures of cortical volume.

Joint scores and external variables

To gain insight into joint scores, relate to other measures:



(a)



(b)

Figure: Joint subject scores estimated via ProJIVE show separation by (a) the count of ApoE4 allele counts and (b) diagnosis via raincloud plots.

Joint scores and external variables

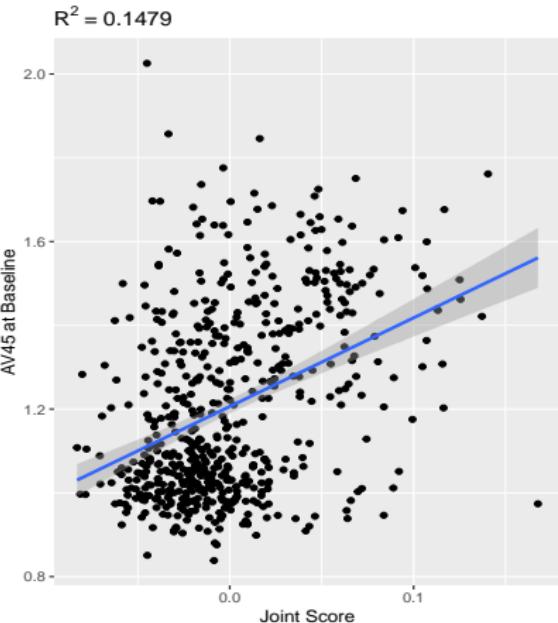
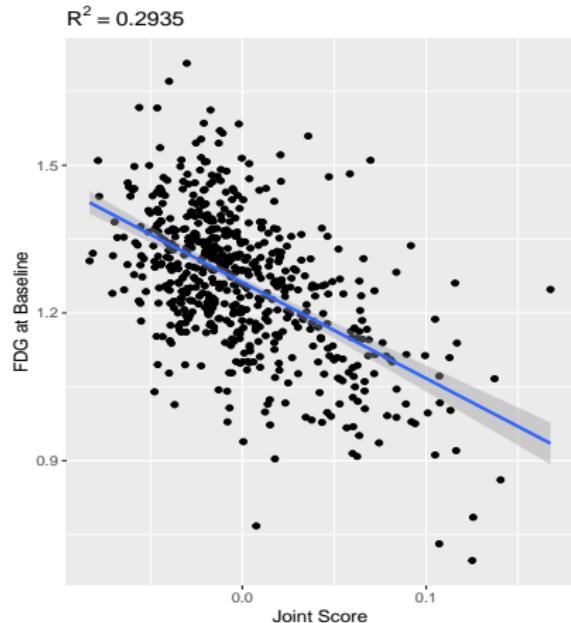


Figure: One question is whether we can use cheap, non-invasive MRIs instead of expensive PET that uses radioactive tracers. FDG PET: measure of brain metabolism and indicative of neurodegeneration. AV45 PET: measures amyloid beta.

Conclusions

- We propose a probabilistic model for JIVE, called ProJIVE.
- Intuitive latent variable formulation.
- Extends probabilistic PCA to multiple datasets.
- EM algorithm to simultaneously estimate joint and individual subspaces.
- In simulations, improves accuracy of both joint and individual scores over existing JIVE methods.
- Joint loadings extract brain regions associated with cognition and dementia.
- Joint loadings are related to biomarkers of AD and dementia, including APOE4, FDG PET, and AV45 PET.

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

Part 2: SING

Part 2: Simultaneous Non-Gaussian Component Analysis (SING) for Data Integration in Neuroimaging

- Co-first author with Irina Gaynanova, Department of Statistics, Texas A&M.
Annals of Applied Statistics
15(3):1431-1454. open access



Integrating rs-fMRI and task fMRI

- 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.
- Here, we extract subspaces using information measured by higher order moments, whereas JIVE uses variance.
- Scientific aim: Data integration in cognition studies of healthy adults [Lerman-Sinkoff et al., 2017].

Are cognitive task regions related to spontaneous brain activity?

Application: Human Connectome Project

- Integrate working memory task fMRI activation maps for 60,000 vertices with resting-state fMRI correlation matrices for 379 regions from the HCP. 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?

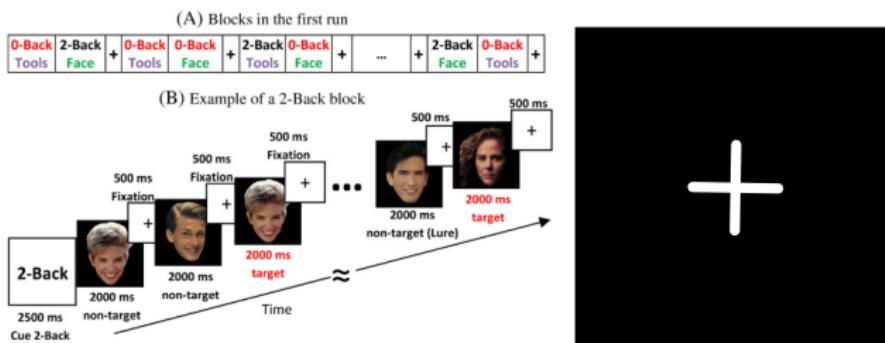
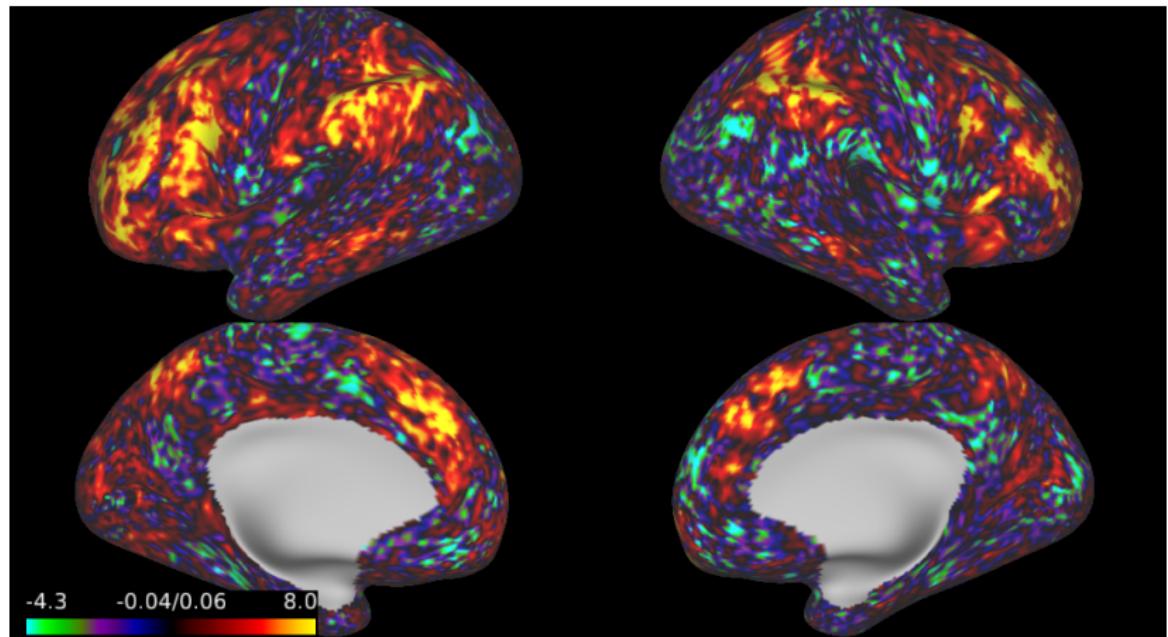


Figure: Example of 2-back working memory task (left) and a crosshairs stared at during resting-state fMRI (right).

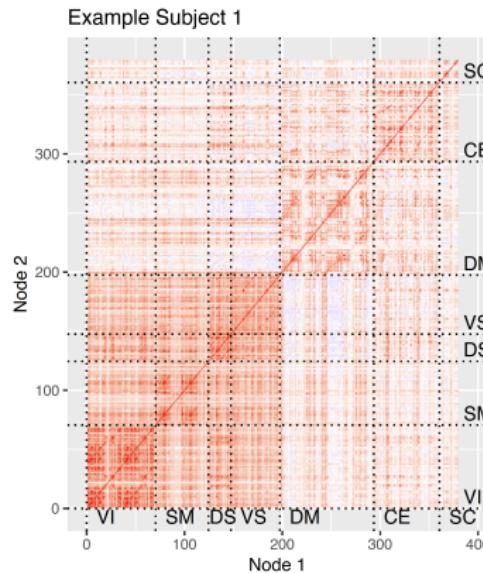
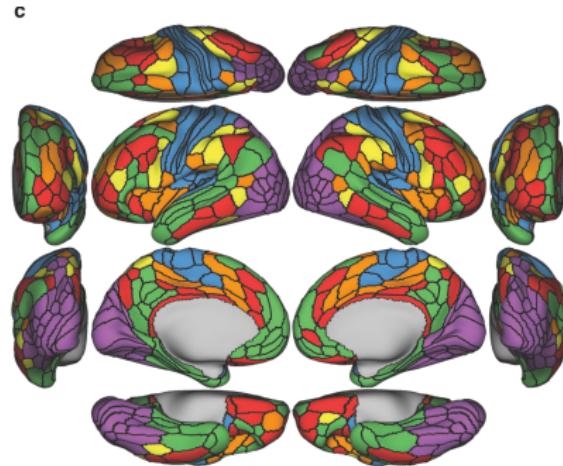
Example activation maps

Input \mathbf{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]: $[\mathbf{X}_s \ \mathbf{Y}_s] \in \mathbb{R}^{n \times (p_x + p_y)}$:

$$[\mathbf{X}_s \ \mathbf{Y}_s] = \mathbf{U}_{(r_J)} \mathbf{D}_{(r_J)} \mathbf{V}_{(r_J)}^\top + \mathbf{U}_{(-r_J)} \mathbf{D}_{(-r_J)} \mathbf{V}_{(-r_J)}^\top$$

$$[\mathbf{X}_s \ \mathbf{Y}_s] = (\mathbf{U}_{(r_J)} \mathbf{D}_{(r_J)} \mathbf{W}^\top) (\mathbf{W} \mathbf{V}_{(r_J)}^\top) + \mathbf{U}_{(-r_J)} \mathbf{D}_{(-r_J)} \mathbf{V}_{(-r_J)}^\top$$

$$[\mathbf{X}_s \ \mathbf{Y}_s] = \mathbf{M}_J [\mathbf{S}_x \ \mathbf{S}_y] + \mathbf{U}_{(-r_J)} \mathbf{D}_{(-r_J)} \mathbf{V}_{(-r_J)}^\top$$

- Estimates **joint** subspace, then an additional rotation of joint subspace to maximize “independence” (non-Gaussianity).
- **Parallel ICA** [Vergara et al., 2014]: PCA on separate datasets, iterate between ICA step and maximizing the correlation between matched columns of \mathbf{M}_y and \mathbf{M}_x .
- **Multimodal CCA with Joint ICA** [Sui et al., 2011]: PCA on each dataset, CCA on PC scores, joint ICA concatenated loadings.
- These methods **use PCA** to estimate the subspace – **removes low variance information**.

JIVE and SING

JIVE: information measured using variance. We focus on non-Gaussian component analysis – information measured using 3rd and 4th moments.

SImultaneous **N**on-Gaussian component model (SING):

$$\mathbf{X}_c = \mathbf{M}_J \mathbf{D}_x \mathbf{S}_{Jx} + \mathbf{M}_{Ix} \mathbf{S}_{Ix} + \mathbf{M}_{Nx} \mathbf{N}_x,$$

$$\mathbf{Y}_c = \mathbf{M}_J \mathbf{D}_y \mathbf{S}_{Jy} + \mathbf{M}_{Ly} \mathbf{S}_{Ly} + \mathbf{M}_{Ny} \mathbf{N}_y.$$

- $\mathbf{M}_J \in \mathbb{R}^{n \times r_J}$, $\|\mathbf{m}_{J\ell}\|_2 = 1$, \mathbf{D}_x diagonal.
- $\mathbf{S}_{Jx} : r_J \times p_x$, $\mathbf{S}_{Jx} \mathbf{S}_{Jx}^\top = p_x \mathbf{I}_{r_J}$, $\mathbf{S}_{Jx} \mathbf{1} = \mathbf{0}$.
- $\mathbf{S}_{Ix} : (r_x - r_J) \times p_x$, $\mathbf{S}_{Ix} \mathbf{S}_{Ix}^\top = p_x \mathbf{I}$, $\mathbf{S}_{Ix} \mathbf{1} = \mathbf{0}$.
- Gaussian components: $\mathbf{N}_x : (n - r_x) \times p_x$.
- $\mathbf{S}_{Jx} \mathbf{S}_{Ix}^\top = \mathbf{0}$, $\mathbf{N}_x \mathbf{S}_{Jx}^\top = \mathbf{0}$, $\mathbf{N}_x \mathbf{S}_{Ix}^\top = \mathbf{0}$.
- Rows of \mathbf{S}_{Jx} , \mathbf{S}_{Ix} , \mathbf{S}_{Jy} , and \mathbf{S}_{Ly} have maximum non-Gaussianity.

Objective function to SING

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

$$\operatorname{argmin}_{\mathbf{U}_X, \mathbf{U}_Y} - \sum_{\ell=1}^{r_X} -f(\mathbf{u}_{X\ell}^\top \mathbf{X}_w) - \sum_{\ell=1}^{r_Y} f(\mathbf{u}_{Y\ell}^\top \mathbf{Y}_w) + \rho \sum_{\ell=1}^{r_J} d(\widehat{\mathbf{L}}_X^{-1} \mathbf{u}_{X\ell}, \widehat{\mathbf{L}}_Y^{-1} \mathbf{u}_{Y\ell})$$

subject to $\mathbf{U}_X \mathbf{U}_X^\top = \mathbf{I}_{r_X}$, $\mathbf{U}_Y \mathbf{U}_Y^\top = \mathbf{I}_{r_Y}$,

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

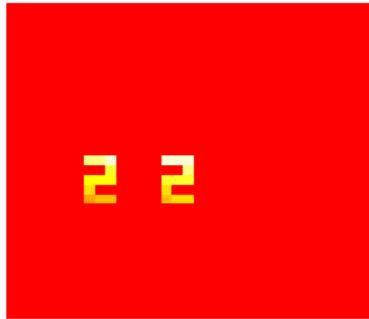
We use the Jarque-Bera test statistic [Virta et al., 2016]:

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

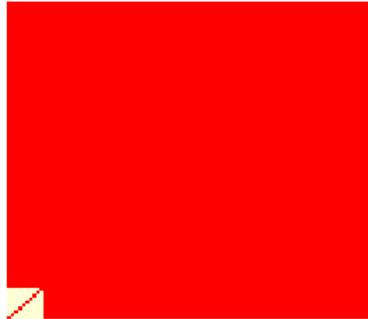
Use an algorithm based on [Wen and Yin, 2012] for feasible updates.

Toy Example: Truth

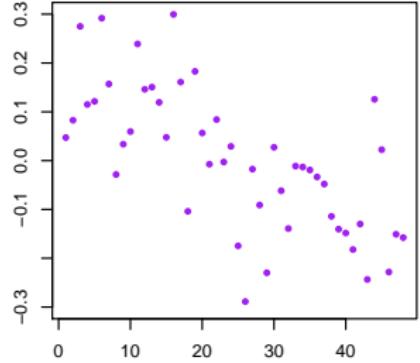
Joint X, Comp 1



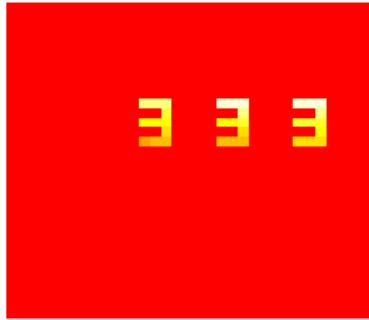
Joint Y, Comp 1



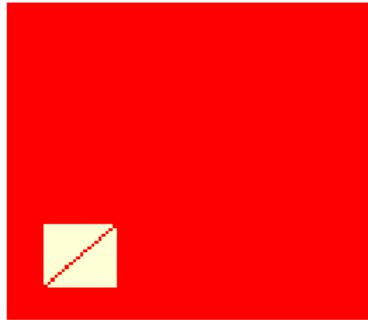
Joint Scores, Comp 1



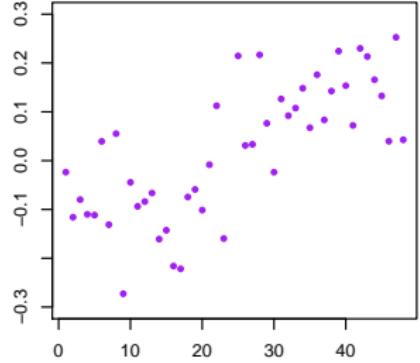
Joint X, Comp 2



Joint Y, Comp 2

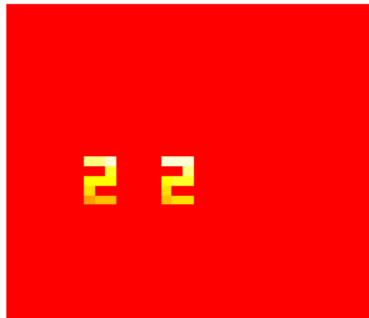


Joint Scores, Comp 2



Toy Example: SING Estimate

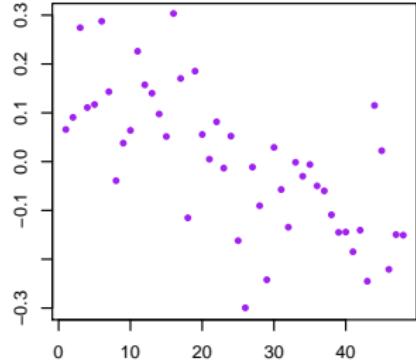
JB Joint X, Comp 1



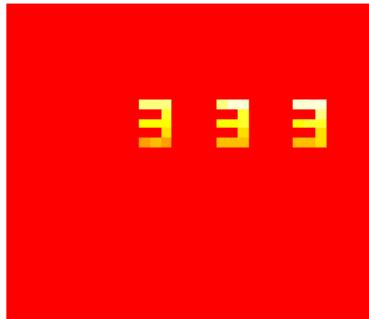
JB Joint Y, Comp 1



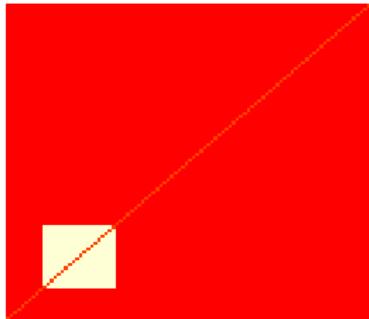
JB Scores, Comp 1



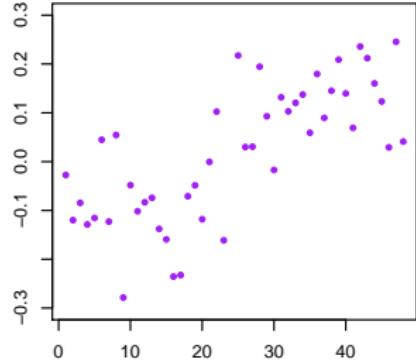
JB Joint X, Comp 2



JB Joint Y, Comp 2

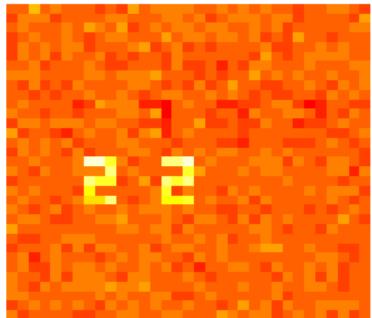


JB Scores, Comp 2



Toy Example: Joint ICA Estimate

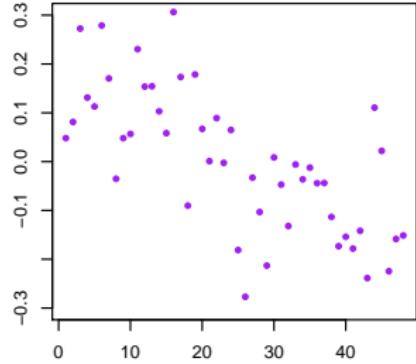
PCA+ICA Joint X, Comp 1



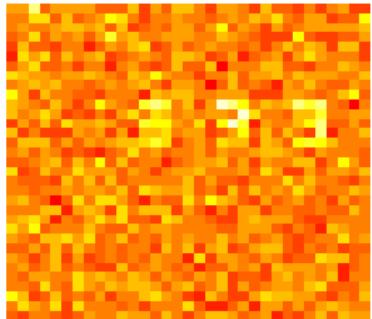
PCA+ICA Joint Y, Comp 1



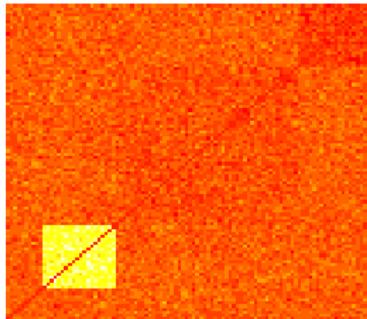
PCA+ICA Scores, Comp 1



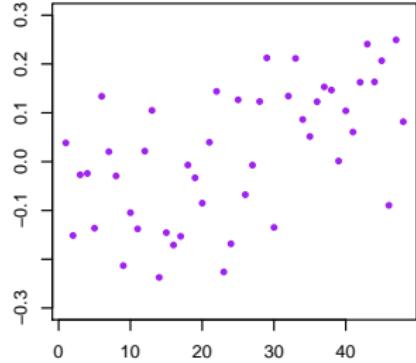
PCA+ICA Joint X, Comp 2



PCA+ICA Joint Y, Comp 2



PCA+ICA Scores, Comp 2



Estimating joint rank

To estimate the joint rank, we can use a preliminary estimate of the components and see if their correlation is significantly greater than chance.

Remark

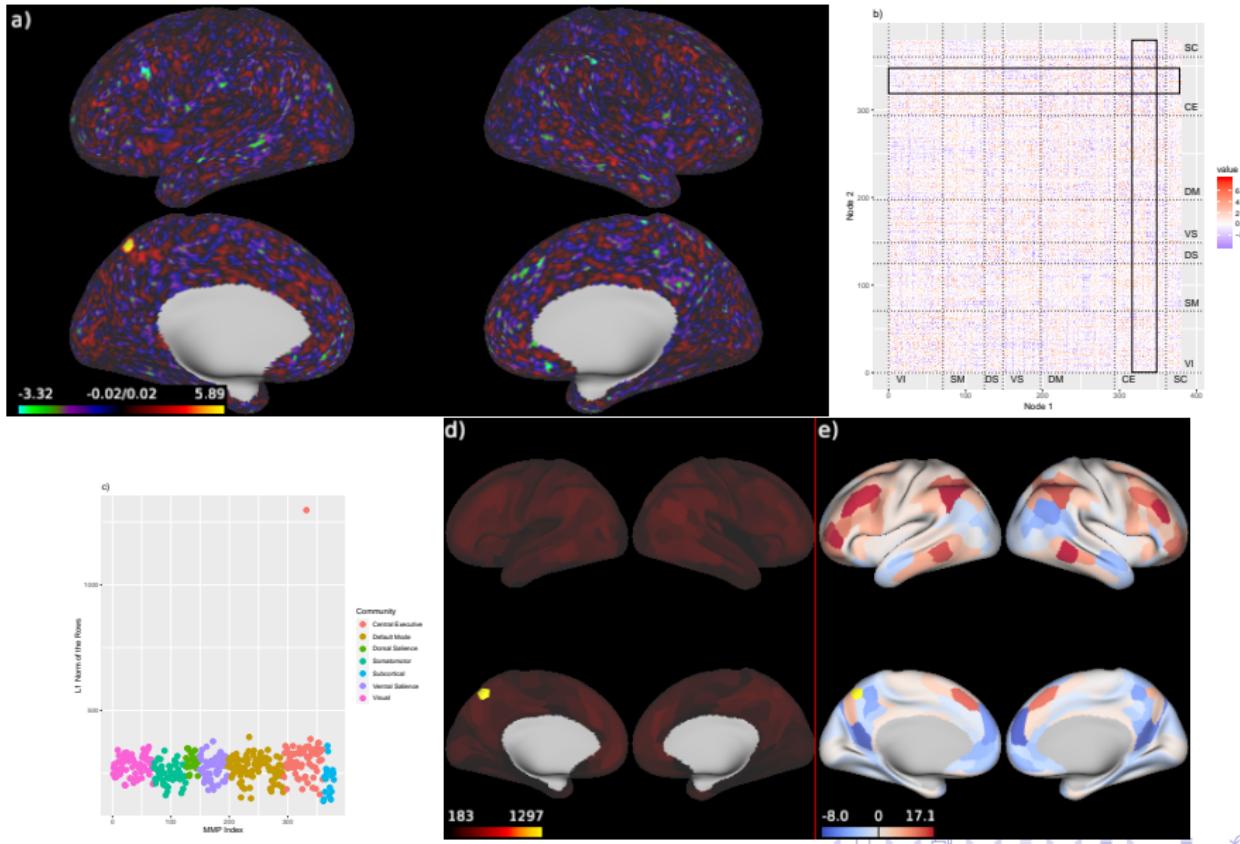
Due to the uniqueness of the linear non-Gaussian component analysis decompositions (up to signed permutations [Risk et al., 2019]), the joint components are a subset of the components from the separate decompositions.

- ① Perform separate LNGCA.
- ② Match columns of \mathbf{M}_x and \mathbf{M}_y based on their correlation.
- ③ Estimate r_J using an FWER-controlled permutation test based on the correlation between \mathbf{M}_x and permuted rows of \mathbf{M}_y .

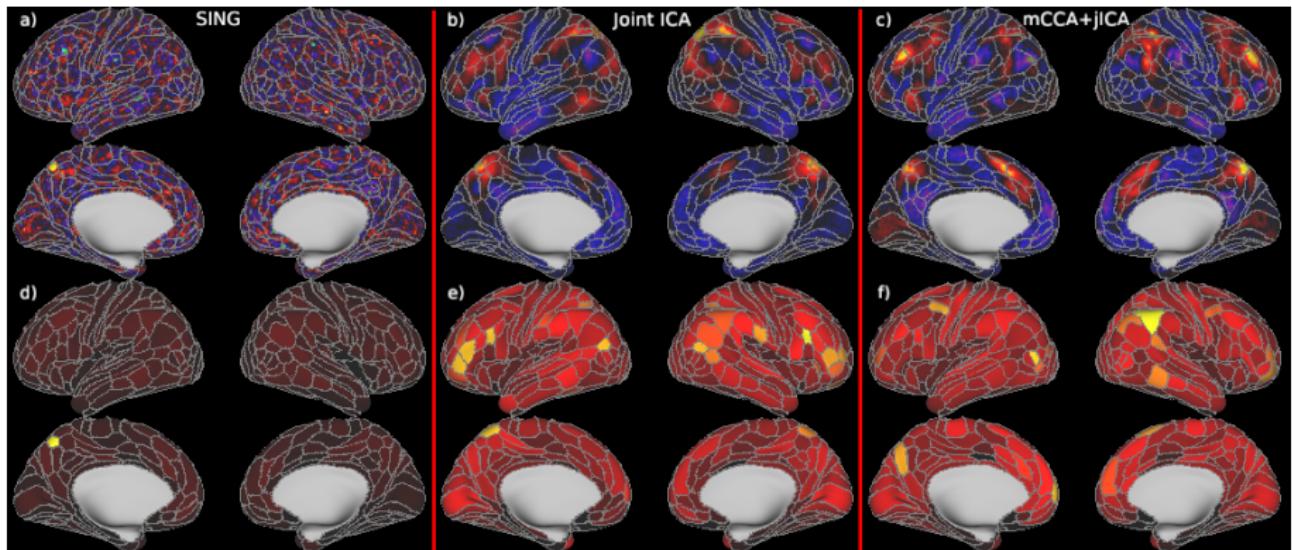
Activation Maps and Networks

- 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, penalty 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 ($t=5.10$).

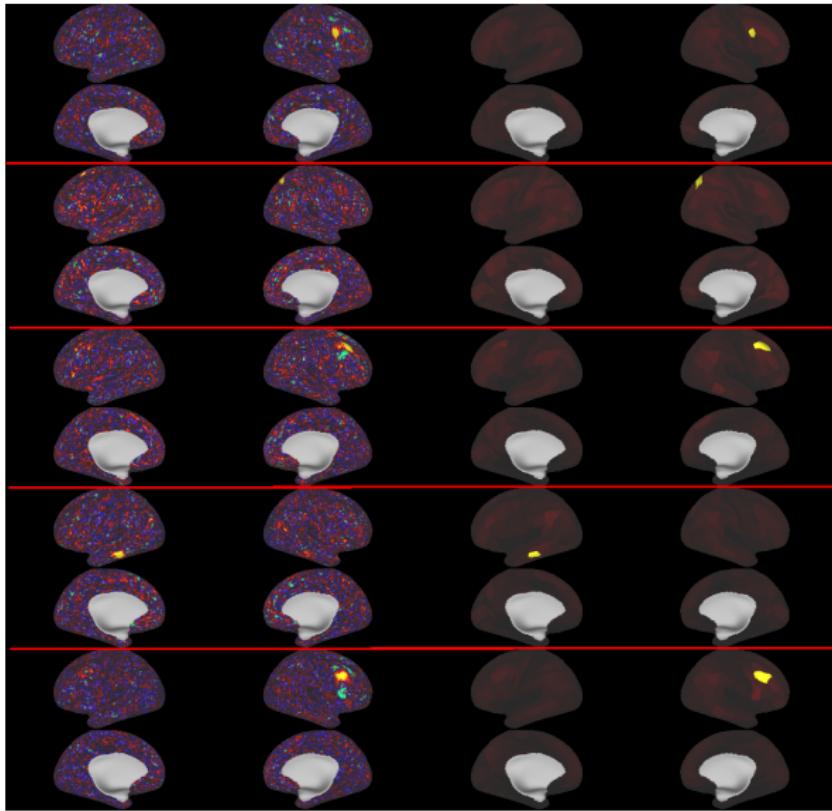
Joint component and fluid intelligence



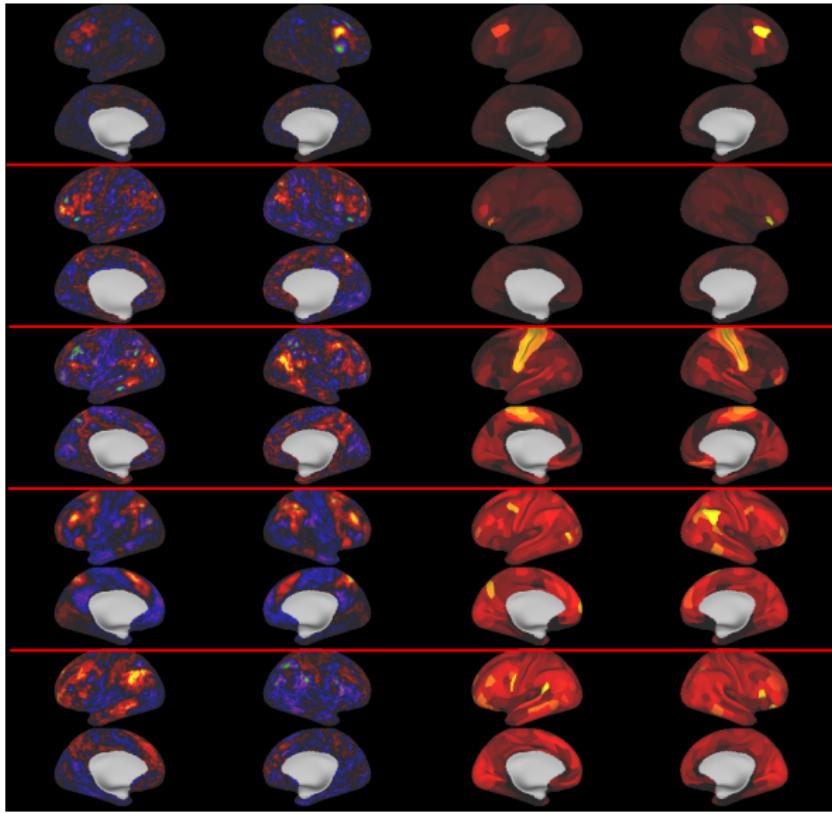
Comparing SING, Joint ICA, mCCA+jICA



Additional SING component loadings



Additional mCCA+jICA component loadings



Discussion

- ① We propose a new matrix decomposition for shared structure across datasets with non-Gaussian features.
- ② Can improve 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.
- ④ Future work: extending to > 2 datasets, models with different noise structures, missing data, sparsity, non-linear methods.
- ⑤ Future work: guidelines for variance versus non-Gaussianity for extracting subspaces.

Acknowledgments

- Thank you!
- Data were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at:
http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf
- Data were provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

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.



Browne, M. W. (1979).

The maximum-likelihood solution in inter-battery factor analysis.

British Journal of Mathematical and Statistical Psychology, 32(1):75–86.



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.



Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., and others (2006).
An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest.
Neuroimage, 31(3):968–980.



Feng, Q., Jiang, M., Hannig, J., and Marron, J. S. (2018).

Angle-based joint and individual variation explained.
Journal of Multivariate Analysis, 166:241–265.

References II

-  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.
-  Klami, A., Virtanen, S., and Kaski, S. (2013).
Bayesian Canonical Correlation Analysis.
Journal of Machine Learning Research, 14:965–1003.
-  Klami, A., Virtanen, S., Leppäaho, E., and Kaski, S. (2015).
Group Factor Analysis.
IEEE Transactions on Neural Networks and Learning Systems, 26(9):2136–2147.
-  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.
-  Risk, B. B., Matteson, D. S., and Ruppert, D. (2019).
Linear Non-Gaussian Component Analysis Via Maximum Likelihood.
Journal of the American Statistical Association.
-  Sui, J., Pearson, 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).

References III



Tipping, M. E. and Bishop, C. M. (1999).

Probabilistic principal component analysis.

Journal of the Royal Statistical Society: Series B (Statistical Methodology), 61(3):611–622.



Tucker, L. R. (1958).

An inter-battery method of factor analysis.

Psychometrika 1958 23:2, 23(2):111–136.



Vergara, V. M., Ulloa, A., Calhoun, V. D., Boutte, D., Chen, J., and Liu, J. (2014).

A three-way parallel ICA approach to analyze links among genetics, brain structure and brain function.

NeuroImage, 98:386–394.



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 2012 142:1, 142(1):397–434.



Yu, Q., Risk, B., Zhang, K., and Marron, J. (2017).

JIVE integration of imaging and behavioral data.

NeuroImage, 152.



Zhu, H., Li, G., and Lock, E. F. (2020).

Generalized integrative principal component analysis for multi-type data with block-wise missing structure.

Biostatistics, 21(2):302–318.