REPRODUCIBILITY IN JOINT BLIND SOURCE SEPARATION: APPLICATION TO FMRI ANALYSIS

Francisco Laport [1, 2], Trung Vu [1], Hanlu Yang [1], Vince Calhoun [3] and Tülay Adali [1]

[1]Dept. of CSEE, UMBC, Baltimore, MD [2]CITIC Research Center, UDC, A Coruña, Spain
[3]TReNDS, Georgia State University, Georgia Institute of Technology, Emory University, Atlanta, GA

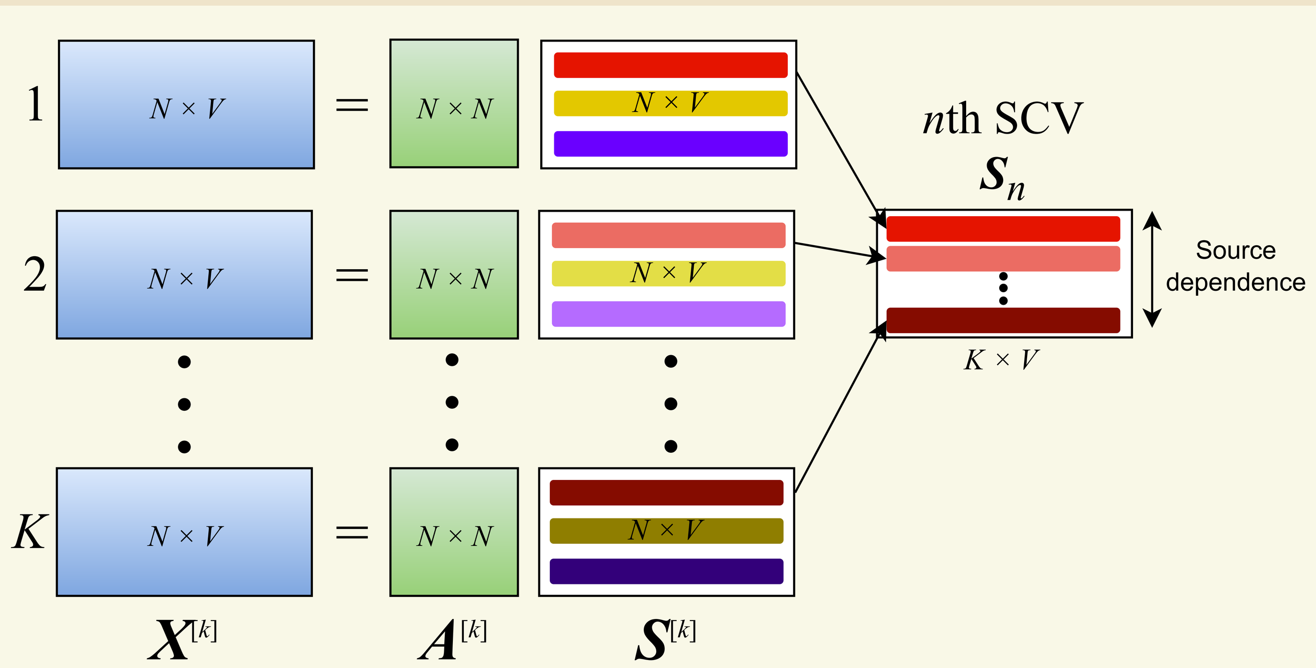
Introduction

- Blind source separation (BSS) techniques have been successfully applied to a wide array of domains including fMRI data analysis
- Joint BSS (JBSS) techniques are implemented to leverage the joint information across multiple datasets
- JBSS results present high variability:
  - Cost functions of most JBSS algorithms are non-convex
  - Closed-form solutions do not exist for these problems – iterative solutions
  - No unique and perfect initialization – random initialization
- Reproducibility assessment of JBSS has been limited in the literature
- Highly consistent results do not guarantee a low bias in the estimates

Contributions

- Evaluate the computational reproducibility of a JBSS algorithm: **constrained independent vector analysis (cIVA)**
- Propose a normalized measure related to the cost function to evaluate cIVA performance in practical scenarios
- Present a new mechanism for selecting the model complexity based on reproducibility and accuracy metrics
- Show that the model orders that balance accuracy and reproducibility metrics provide the most meaningful and interpretable results from analyzing real fMRI data

Independent Vector Analysis (IVA)



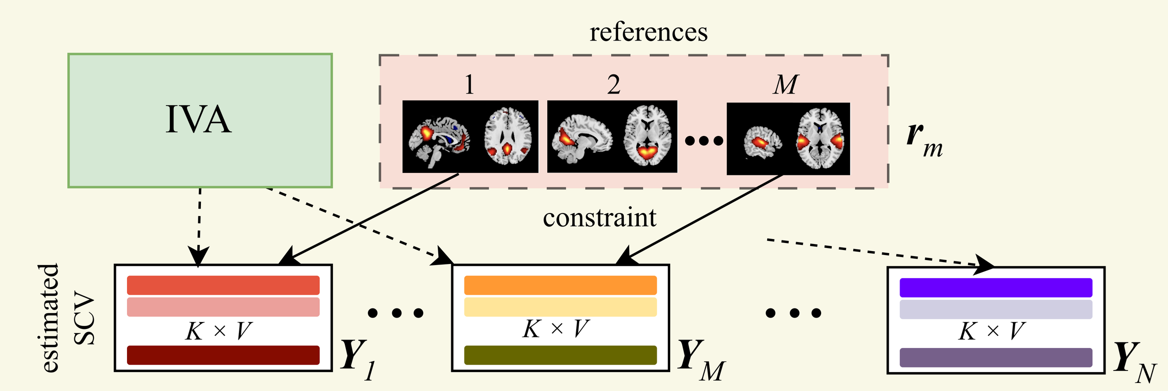
- IVA cost function

$$\mathcal{J}_{\text{IVA}}(\mathcal{W}) = \sum_{n=1}^{N}\left(\sum_{k=1}^{K}\mathcal{H}(y_n^{[k]}) - \mathcal{I}(y_n)\right) - \sum_{k=1}^{K}\log|\det(W^{[k]})|$$

where $\mathcal{W} = \{W^{[1]}, \ldots, W^{[K]}\}$ are the demixing matrices of the K datasets, y_n is the nth estimated SCV, $\mathcal{I}(y_n)$ is the mutual information of y_n and $\mathcal{H}(y_n^{[k]})$ is the entropy of the nth estimated source for the kth dataset

This work is supported in part by the grants NIH R01MH118695, NIH R01MH123610, NIH R01AG073949, NSF 2316420, and Xunta de Galicia – Fulbright ED481B 2022/012

Constrained IVA



- The augmented cost function

$$\mathcal{L}_\lambda(\mathcal{W}) = \mathcal{J}_{\text{IVA}}(\mathcal{W}) + \frac{\lambda}{2}\sum_{m=1}^{M}\sum_{k=1}^{K}\left(\sum_{\substack{n=1\\n\neq m}}^{M}\epsilon^2(r_m, y_n^{[k]}) - \epsilon^2(r_m, y_m^{[k]})\right)$$

Reproducibility and Accuracy Metrics

- In real applications the ground truth is unknown
- **Reproducibilty: cross-joint-ISI**
  - Let $G^{[k]} = A^{[k]}W^{[k]}$ and $G = 1/K\sum_{k=1}^{K}|G^{[k]}|$ then

$$\text{ISI}(G) = \frac{\sum_{i=1}^{N}\left(\frac{\sum_{j=1}^{N}|G_{ij}|}{\max_p|G_{ip}|} - 1\right) + \sum_{j=1}^{N}\left(\frac{\sum_{i=1}^{N}|G_{ij}|}{\max_p|G_{pj}|} - 1\right)}{2N(N-1)}$$

  - Let $W_r^{[k]}$ be the kth demixing matrix of the rth run and $P_{ij} = 1/K\sum_{k=1}^{K}|(W_i^{[k]})^{-1}W_j^{[k]}|$ then

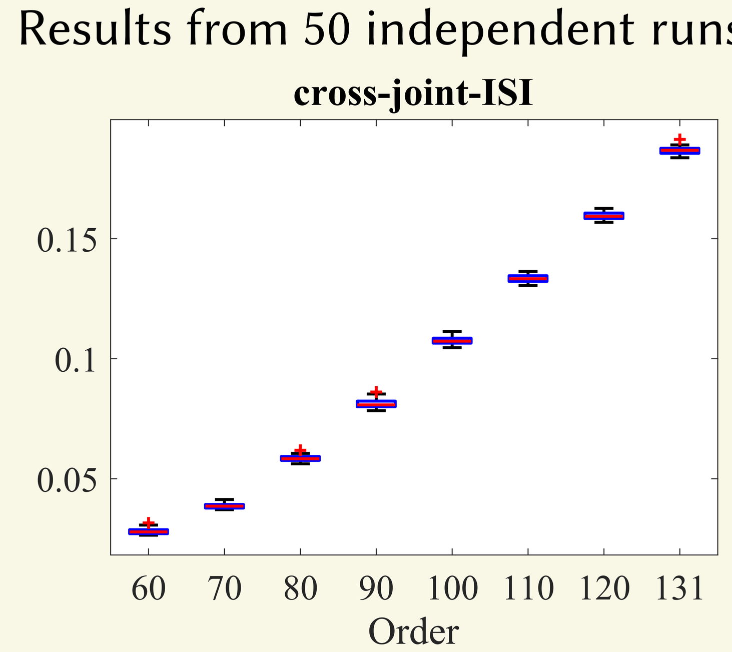
$$\text{cross-joint-ISI}_i = \frac{1}{R}\sum_{j=1, j\neq i}^{R}\text{ISI}(P_{ij})$$

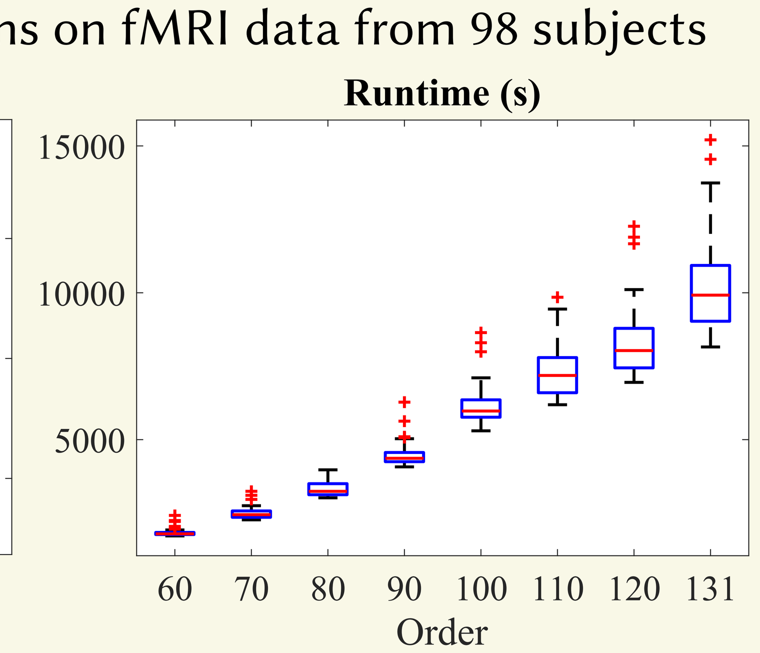
- **Accuracy: Pairwise normalized mutual information**
  - Let $I(y_i^{[k]}, y_j^{[k]})$ be the mutual information between two estimated components

$$\text{N-MI}_r = \frac{2}{KN(N-1)}\sum_{k=1}^{K}\sum_{i=1, i\neq j}^{N}\sum_{j>i}^{N}\frac{2I(y_i^{[k]}, y_j^{[k]})}{I(y_i^{[k]}, y_i^{[k]}) + I(y_j^{[k]}, y_j^{[k]})}$$

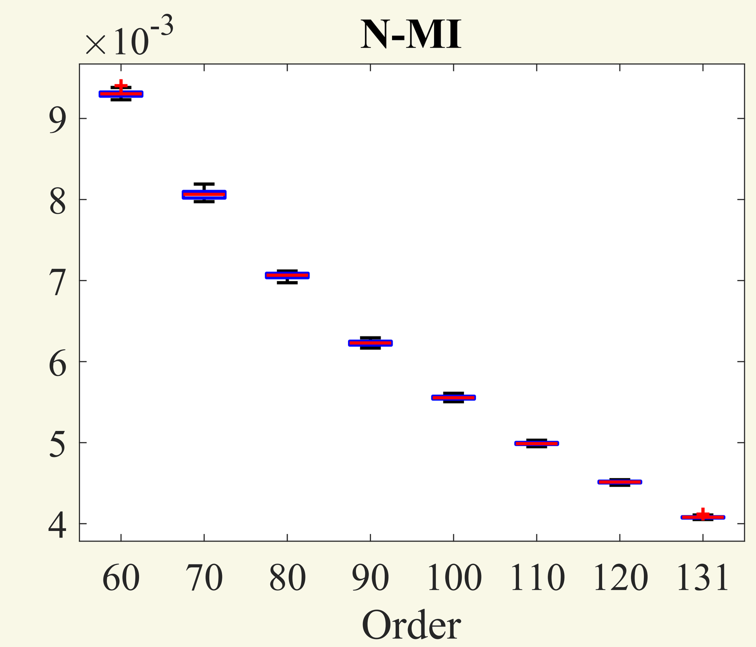
Experimental Results

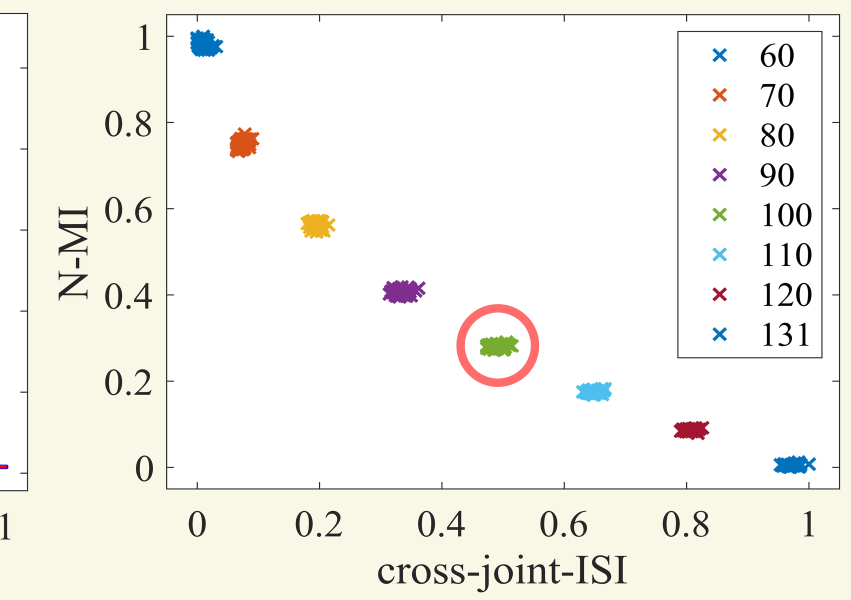
- Results from 50 independent runs on fMRI data from 98 subjects





Experimental Results (continued)

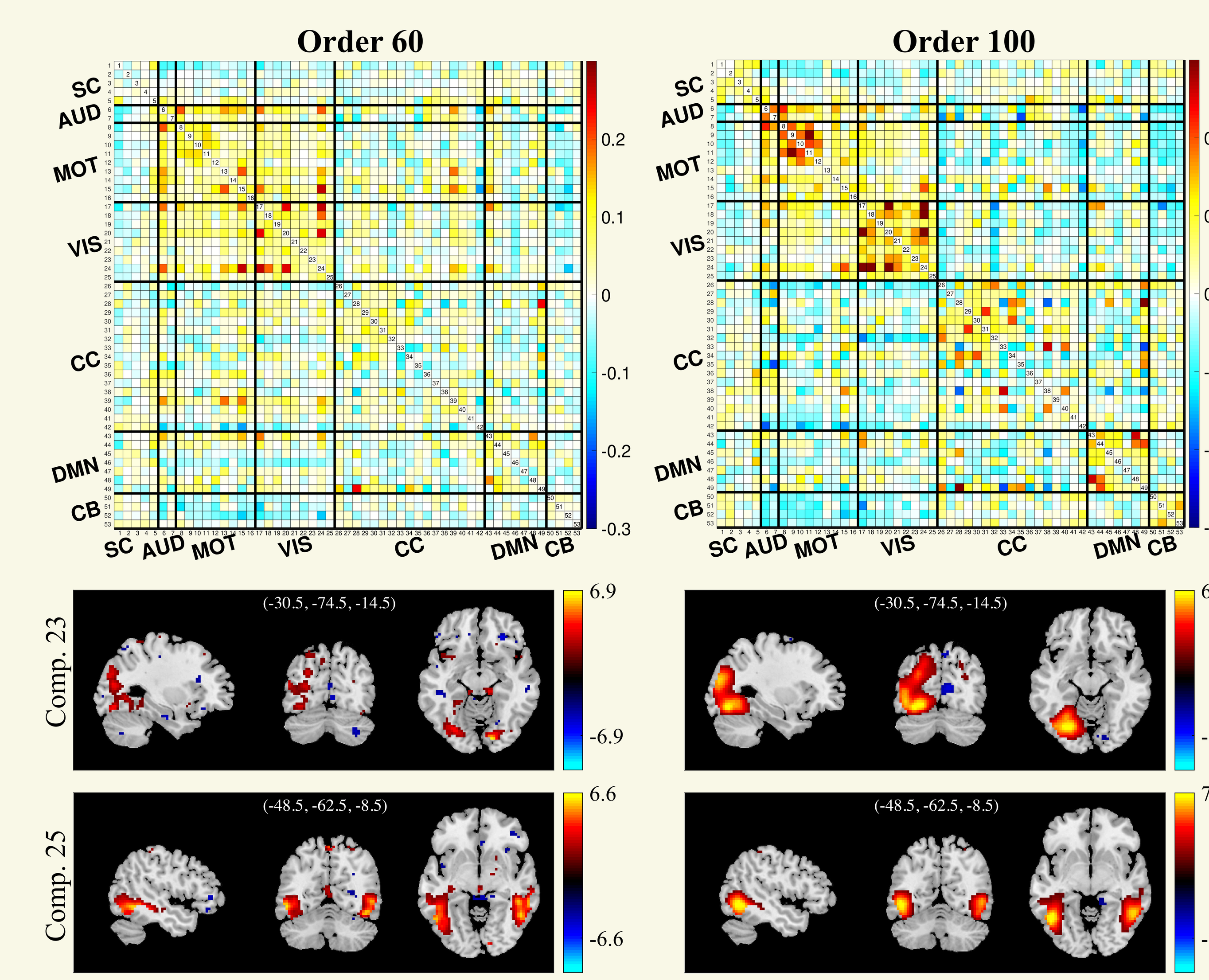




- Cross-joint-ISI increases with the model order while N-MI decreases – Bias and Variance dilemma
- Model orders balancing cross-joint-ISI and N-MI show clearer spatial maps, higher temporal correlations within functional domains, and higher power ratios

	60	80	90	100	120
Power Ratio	3.74±2	3.89±2.35	3.92±2.54	4±2.82	3.97±2.83

Table: Power ratio values for different model orders



Summary

- Model order selection should be guided by both reproducibility and accuracy metrics
- Model orders that balance the bias-variance tradeoff provide a better model match and more interpretable and meaningful results in fMRI data