Shared Independent Component Analysis for Multi-Subject Neuroimaging

Hugo Richard[1], Pierre Ablin[3], Alexandre Gramfort[1], Bertrand Thirion[1], Aapo Hyvarinen[2]

[1]Inria, CEA, Université Paris-Saclay, France [2]Department of Computer Science HIIT, University of Helsinki, Finland [3]CNRS and DMA, Ecole Normale Supérieure - PSL University, France

https://github.com/hugorichard/ShICA

Problem

- Uncover the shared neural responses of multiple subjects exposed to the same naturalistic stimuli (e.g. movie watching).
- Current well principled approaches need non-Gaussianity of the common sources to be identifiable.

Solution: Shared ICA (ShICA)

ShICA performs multi-subjects ICA while modeling inter-subject variability yielding an identifiable model even when common sources are Gaussian. In practice, it yields better results than competitive methods

ShICA

Given m subjects, we model the data $\mathbf{x}_i \in \mathbb{R}^k$ of subject i as

$$\mathbf{x}_i = A_i(\mathbf{s} + \mathbf{n}_i), i = 1, \dots, m \quad (1)$$

- H1: $\mathbf{s}$ are independent components some of which may be Gaussian
- H2: $\mathbf{n}_i \sim \mathcal{N}(0, \Sigma_i)$ where Σ_i is diagonal positive and $\mathbf{n}_i$ independent from $\mathbf{s}$
- H3: $\mathbb{E}[\mathbf{x}_i] = 0$, A_i invertible, $\mathbb{E}[\mathbf{s}\mathbf{s}^\top] = I_k$ and $m \geq 3$

Definition (Noise diversity): Let $\mathcal{G}$ be the set of Gaussian components. For all $j, j' \in \mathcal{G}, j \neq j'$, the sequences $(\Sigma_{ij})_{i=1 \dots m}$ and $(\Sigma_{ij'})_{i=1 \dots m}$ are different where Σ_{ij} is the j, j entry of Σ_i.

Theorem (Identifiability): Assuming noise diversity, ShICA is identifiable up to a sign and permutation matrix.

Estimation by ShICA-ML

We assume the ShICA model with:

$$s_j \sim \frac{1}{2} \Sigma_{\alpha \in \{\frac{1}{2}, \frac{3}{2}\}} \mathcal{N}(0, \alpha)$$

Optimization Optimized via an EM algorithm.
E-step:

$$\mathbb{E}[s_j|\mathbf{x}] = \frac{\Sigma_{\alpha \in \{\frac{1}{2}, \frac{3}{2}\}} \theta_\alpha h_\alpha}{\Sigma_{\alpha \in \{\frac{1}{2}, \frac{3}{2}\}} \theta_\alpha}, \quad \mathbb{V}[s_j|\mathbf{x}] = \frac{\Sigma_{\alpha \in \{\frac{1}{2}, \frac{3}{2}\}} \theta_\alpha g_\alpha}{\Sigma_{\alpha \in \{\frac{1}{2}, \frac{3}{2}\}} \theta_\alpha}.$$

M-step: Closed form updates for noise variances and quasi-Newton updates for unmixing matrices.

Estimation by ShICA-J

Background (Multiset CCA): CCA of $(\mathbf{x}_i)_{i=1}^m$ given by solving $C\mathbf{u} = \lambda D \mathbf{u}$ where block i, j of C is $\mathbb{E}[\mathbf{x}_i \mathbf{x}_j^\top]$ and D given by diagonal blocks of C.

Theorem (MultisetCCA solves ShICA: Assume $\mathbf{x}_i$ follows ShICA. Let $U = [\mathbf{u}_1 \dots \mathbf{u}_k]$ (first k eigenvectors of CCA problem) and $\lambda_1, \dots, \lambda_k$ (first k eigenvalues). Set $\begin{bmatrix} W_1^\top \\ \vdots \\ W_m^\top \end{bmatrix} = U$ where $W_i \in \mathbb{R}^{k,k}$.
Then <u>if $\lambda_1 \dots \lambda_k$ are distinct</u>, $W_i = P \Gamma_i A_i^{-1}$ where P is a permutation matrix and Γ_i a scaling matrix.

Serious problem:

- We only have access to empirical covariance estimates
- The mapping from matrices to eigenvectors is highly non smooth

so <u>MultisetCCA fails in practice</u> (especially when the eigenvalues are close).

Solved by joint diagonalization:

- Large gap between the first k eigenvalues and other: the span of the p leading eigenvectors is preserved: $W_i \approx Q\tilde{W}_i$.
- Recover Q by joint diagonalization of $Q\tilde{W}_i \frac{1}{n} X_i X_i^\top \tilde{W}_i^\top Q^\top$

Last steps: Find scalings and noise variances (see paper)

Related work: SRM [2], MVICA [5], CanICA [6], IVA [1].

Separation performance depending on the density of sources

Data are generated with $m = 4$ views, $k = 5$ components using the ShICA model. Non-Gaussian sources follow a Laplace density.

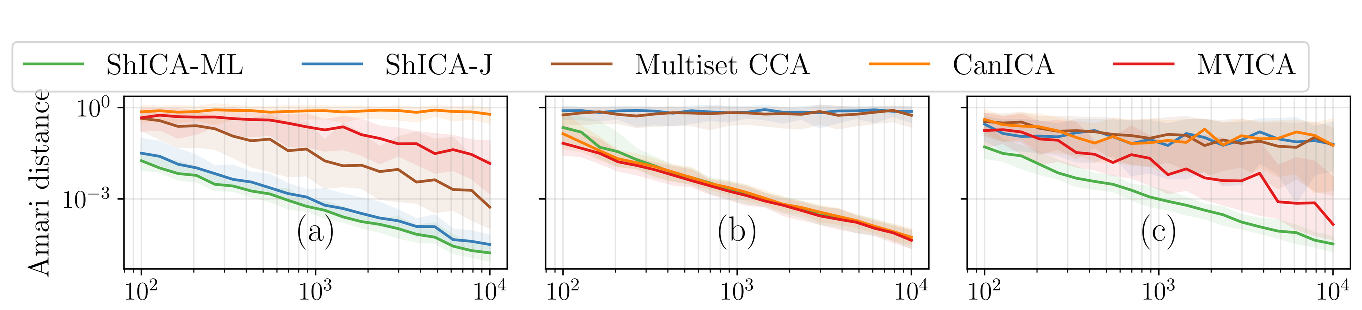
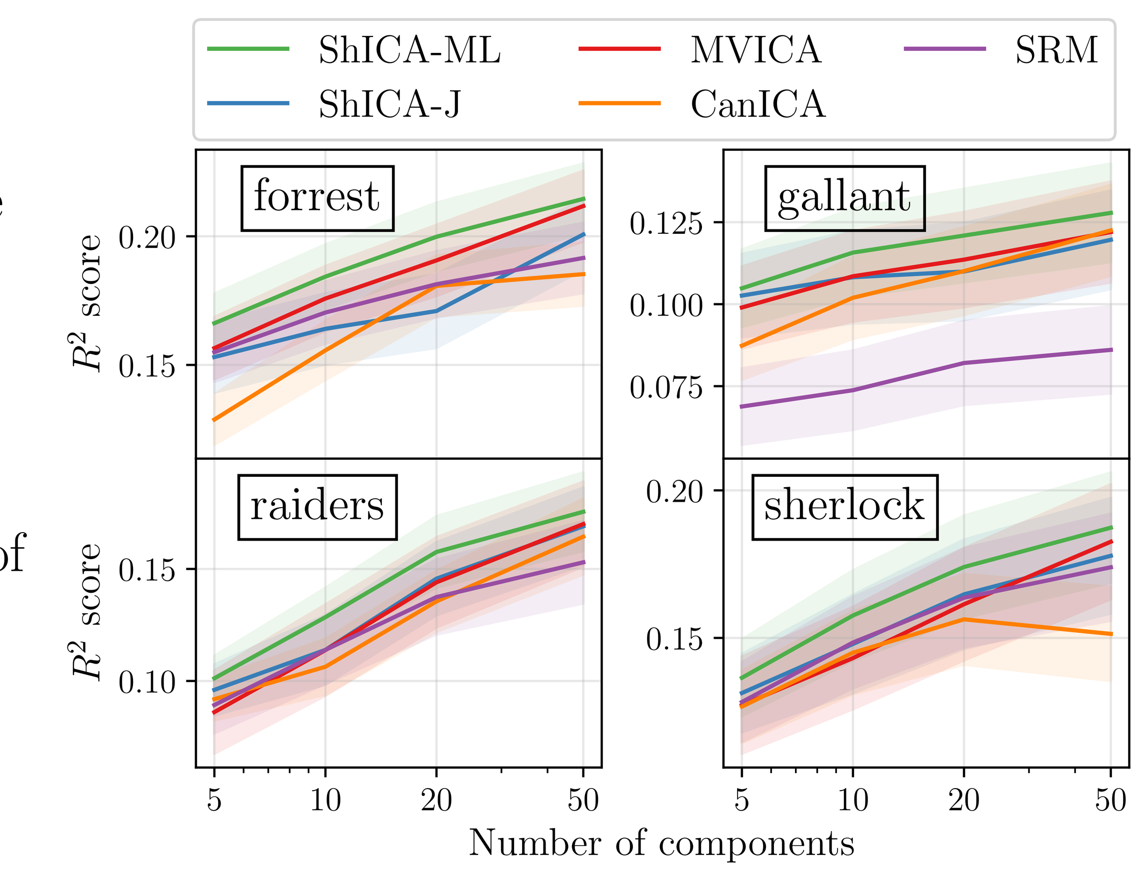


Figure: (a) Gaussian components with noise diversity (b) non-Gaussian components without noise diversity (c) Half of components are Gaussian with noise diversity, the other half is non-Gaussian without diversity

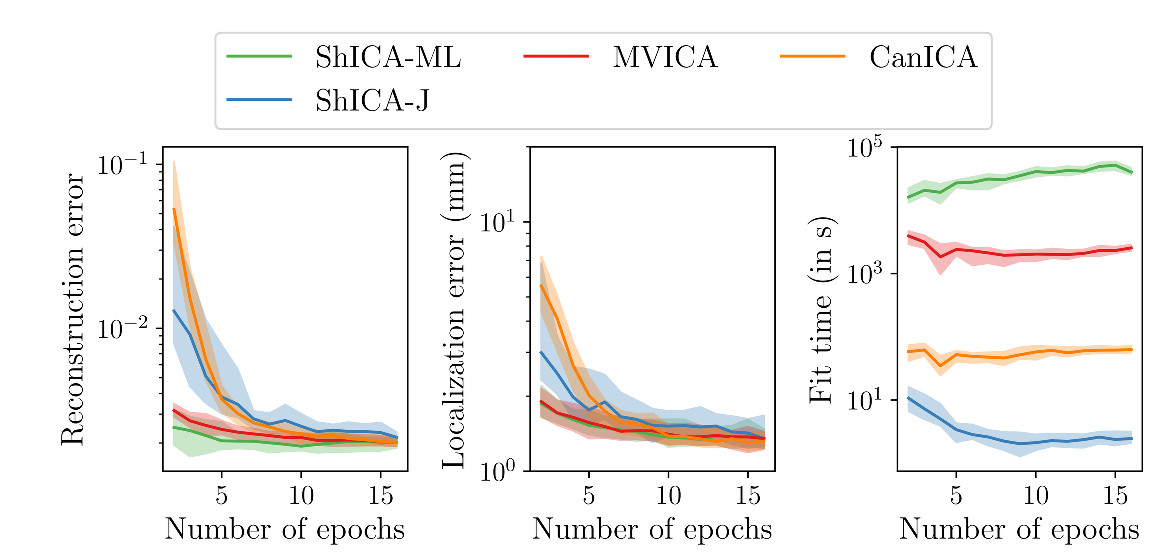
Reconstruction experiment fMRI

- Subjects exposed to the same stimuli
- 80% runs used to learn unmixing matrices
- In remaining runs we reconstruct the data of 20% of the subjects using the other subjects.



MEG Phantom experiment

- 8 dipoles in a plastic head at different locations separately emit signal S_{true} during n epochs
- 20 sources are estimated: the best one is compared with S_{true}.



References

[1] Matthew Anderson et al. "Independent vector analysis: Identification conditions and performance bounds". In: *IEEE Transactions on Signal Processing* 62.17 (2014), pp. 4399–4410.

[2] Po-Hsuan Cameron Chen et al. "A reduced-dimension fMRI shared response model". In: *Advances in Neural Information Processing Systems*. 2015, pp. 460–468.

[3] Yi-Ou Li et al. "Joint blind source separation by multiset canonical correlation analysis". In: *IEEE Transactions on Signal Processing* 57.10 (2009), pp. 3918–3929.

[4] Eric Moulines, Jean-Franccois Cardoso, and Elisabeth Gassiat. "Maximum likelihood for blind separation and deconvolution of noisy signals using mixture models". In: *1997 ieee international conference on acoustics, speech, and signal processing*. Vol. 5. IEEE. 1997, pp. 3617–3620.

[5] H. Richard et al. "Modeling Shared responses in Neuroimaging Studies through MultiView ICA". In: *Advances in Neural Information Processing Systems 33*. Dec. 2020.

[6] Gaël Varoquaux et al. "A group model for stable multi-subject ICA on fMRI datasets". In: *Neuroimage* 51.1 (2010), pp. 288–299.