Independent component analysis for brain fMRI does indeed select for maximal independence

Vince D Calhoun[1,2,3], Vamsi K Potluru[1,3], Ronald Phlypo[5], Rogers F Silva[1,2], Barak A Pearlmutter[4], Arvind Caprihan[1], Sergey M Plis[1], Tülay Adalı[5]

[1]The Mind Research Network, Albuquerque, New Mexico 87106.
[2]Dept. of Electrical and Computer Engineering, University of New Mexico, Albuquerque, New Mexico 87131.
[3]Dept. of Computer Science, University of New Mexico, Albuquerque, New Mexico 87131.
[4]Hamilton Institute & Dept. of Computer Science, NUI Maynooth, Co. Kildare, Ireland
[5]Dept. of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, MD 21250

4000 characters (w/ spaces)
10 refs

Introduction: A recent paper by Daubechies et al. claims that two independent component analysis (ICA) algorithms, Infomax and FastICA, widely used for fMRI analysis[2], select for sparsity and not independence[1]. We show that their synthetic data based experiments fall short of proving this claim and that these ICA algorithms are doing what they are designed to do: identify maximally independent sources.

Methods: Daubechies et al.[1] shows results in which (1) ICA algorithm performance suffers when the assumptions on the sources are violated, and (2) ICA algorithms can separate sources in certain cases even if the sources are not strictly independent. The two points above, already widely known in the ICA community at the time, are not sufficient evidence to support the claim that ICA selects for sparsity and not independence. In addition, Daubechies et al.[1] presents a case in which the sources are somewhat dependent but also very sparse, and ICA does well. This result is used to claim that it is sparsity rather than independence that matters. We augment this experiment with new evidence which shows that the same ICA algorithms perform equally well in the case of minimum sparsity, suggesting that the role of sparsity (if any) is minor in the separation performance.

Additional evidence involves a claim that ICA separates Gaussian sources which are also sparse, which ICA should not be able to do. We show that the sources they generated are far from Gaussian, and the sparsity mentioned in [1] does not even refer to the sources. Thus the results do not support the claim being made.

Results: Daubechies et al. argues, based largely on results from synthetic datasets using different sized boxes to represent activated regions of a component, that it is sparsity and not independence that is important[1]. In Figure 1, we see that excess kurtosis of the simulated sources for the two cases where Infomax and FastICA are noted to fail (but it is claimed they should not) is very close to zero. Moreover, at these points the distributions are bimodal, far from the unimodal super-Gaussian assumptions that underpin the nonlinearities of Infomax and FastICA[1] creating very challenging scenarios for Infomax and FastICA. In actuality, the sources are not even close to the “ideal” setup for these algorithms, contrary to the claim on p.10418[1]. It is thus expected that these algorithms would do poorly.

The paper[1] notes that mixtures of independent Gaussian random variables cannot be recovered by ICA, which is true if each source comes from a single Gaussian distribution, and then shows that ICA can recover sources in this case. The sources generated are claimed to be sparse and Gaussian and the successful recovery of the sources is claimed as evidence that sparsity is the driving force helping ICA to recover Gaussian sources. However, there are two problems with this argument. First, the sparsity mentioned is not in the relevant domain. And secondly, this example utilizes a mixture of Gaussians as the sources and hence the sources themselves are in fact super-Gaussian (i.e. they have positive excess kurtosis). Infomax and FastICA are expected to successfully separate such sources (see Figure 2). This example again points to a confusion with respect to the definition of the underlying ICA sources, i.e., what is actually being simulated and what is assumed in Daubechies et al.[1].

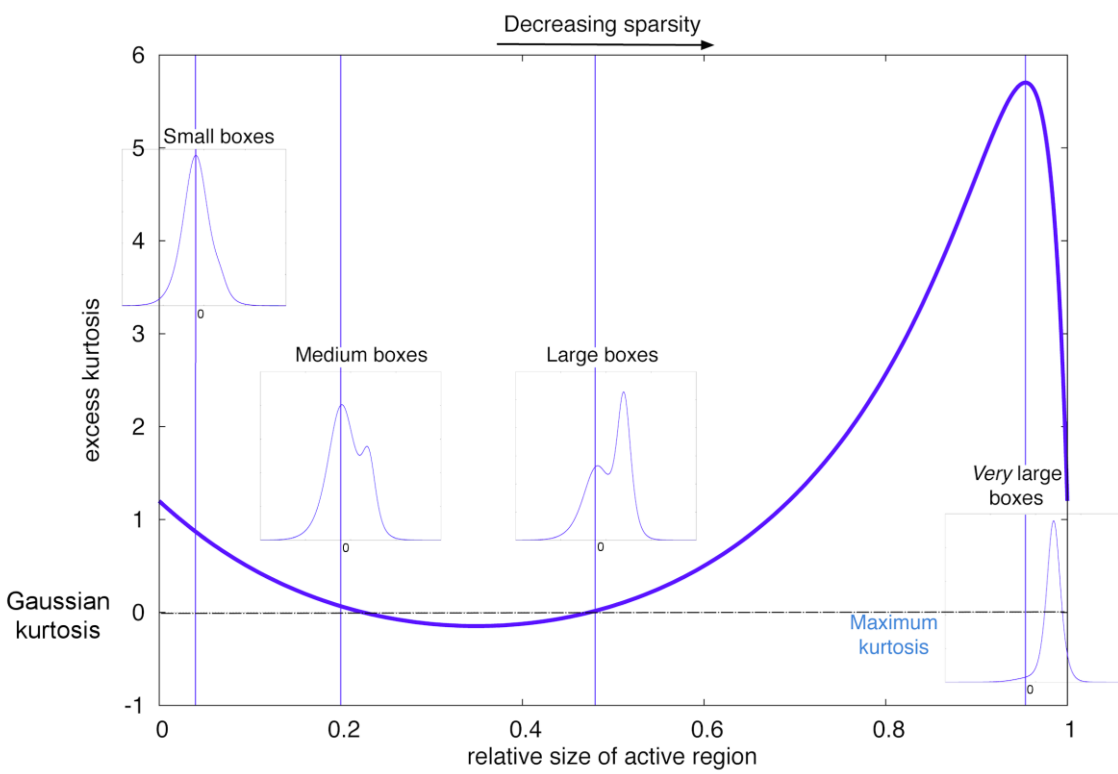
Conclusions: We conclusively show that the arguments in Daubechies et al.[1] fall short in supporting the claim that Infomax and FastICA select for sparsity and not for independence. While pointing out the use of other metrics for fMRI analysis such as sparsity is a reasonable goal, the claims used to justify this desire are misleading at best and in some cases are simply incorrect. In summary, FastICA and Infomax are doing what they were designed to do: maximize independence.

References:

[1] I. Daubechies, E. Roussos, S. Takerkart, M. Benharrosh, C. Golden, K. D'Ardenne, W. Richter, J. D. Cohen, and J. Haxby, "Independent component analysis for brain fMRI does not select for independence," *Proc Natl Acad Sci U S A,* vol. 106, pp. 10415-10422, Jun 30 2009.

[2] V. D. Calhoun and T. Adalı, "Multi-subject Independent Component Analysis of fMRI: A Decade of Intrinsic Networks, Default Mode, and Neurodiagnostic Discovery," *IEEE Reviews in Biomedical Engineering,* in press, PMC Journal - In Process.

Figure 1: The excess kurtosis (a Gaussian has zero excess kurtosis) of a source C_i as a function of the relative size of the active region. The four vertical lines at $q = 0.05,\ 0.20,\ 0.48,$ and 0.95 correspond to the relative sizes of the small box, the medium box V_2, the large box V_1, and a very large box corresponding to the maximal kurtosis case. Note that the medium and large box experiments have near zero excess kurtosis, *i.e., kurtosis value matching that of a Gaussian*. In addition, the pdfs of these sources are bimodal (see inset figures), ensuring that ICA algorithms designed for unimodal super-Gaussian distributions such as Infomax and FastICA with standard parameter settings, will likely fail. At the bottom of the figure are the ISI values (see Equation 1) for the various algorithms at those four points (see **Error! Reference source not found.** for full list). Also note the best separation performance of Infomax and FastICA for the maximum kurtosis case, which corresponds to almost the *lowest* level of sparsity.



FastICA (kurt)	0.04	0.15	0.31	0.02
Infomax (super)	0.02	0.74	0.92	0.01
Infomax (sub)		0.11	0.03	
ICA-EBM	0.04	0.05	0.01	0.01

good
poor
very poor

ISI for 4 algorithms

Figure 2: We plot (top) the distribution of sources, and (bottom) the scatter plot of mixtures for the case of $\gamma = 1$ for 30% of the time. Contrary to the claim made in Daubechies et al., the sources have in fact very peaky and heavy-tailed distributions and are not at all close to a Gaussian distribution. For comparison purposes we also present Gaussian distribution curves (blue)