

Definition of the problem

The general problem can be formalised this way. Suppose we have some random variables  $x \in \mathbb{R}^p$  which correspond to a mix of some primitive sources  $s \in \mathbb{R}^n$ . The aim is to extract from  $x$  every source  $s_i$ . To do so, we suppose here that:

- the sources are independents
- the mix is linear and instantaneous
- at most one source has a gaussian distribution.

We write:

$$x = As \text{ and } y = Wx, \tag{1}$$

where  $A$  is the mixing matrix,  $W$  the separation matrix and  $y$  the estimation of the sources. The goal is then to find the matrix  $W$  that maximises the independence of  $y$ . )

Measure of independance

As a measure of independence, we consider, for theoretical purpose, the mutual information:

$$I(Y) = \int_{\mathbb{R}^p} P(Y) \log \frac{P(Y)}{\prod_i P_i(Y_i)} dY. \tag{2}$$

However, as it is too hard to compute, we consider other contrast functions, invariant by permutation, scaling on coordinates and maximal for independant ones. For practical purpose, we consider other contrast function than the mutual information (too hard to compute), which are invariant by permutation and scaling on coordinates and “minimal” for independent coordinates.

Results

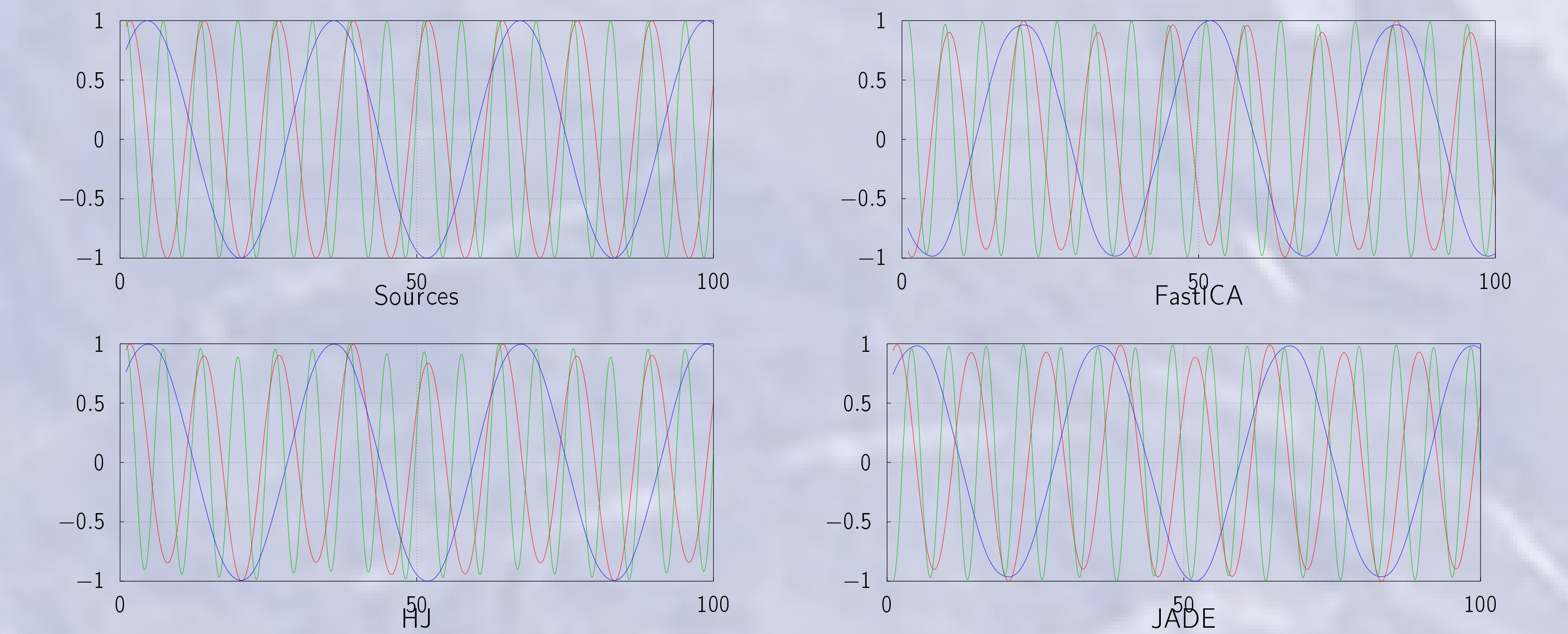


Figure: Results of the main ICA algorithms on simulated data. Each colour corresponds to a signal. The “Sources” graph shows the unmixed signal  $s$  and the other ones the results of the algorithms.

Kernel ICA

The idea here is to have a contrast function defined by the correlation of a range of functions over a function vector space  $\mathcal{F}$ . Thus defining the  $\mathcal{F}$ -correlation :

$$\rho_{\mathcal{F}}(x_1, x_2) = \max_{f_1, f_2 \in \mathcal{F}} \text{corr}(f_1(x_1), f_2(x_2))$$

For  $\mathcal{F}$  “large enough”, we have the following equivalence :

$\rho_{\mathcal{F}}(x_1, x_2) = 0 \iff x_1 \text{ \& } x_2 \text{ are independant}$ , hence the “vector space contrast function” idea. We then use the *reproducing kernel Hilbert space* to obtain tractable computations over the vector space  $\mathcal{F}$ .

Methods benchmark

We now wish to compare the different algorithms performing ICA we kept :

- HJ
- JADE
- FastICA
- KernelICA

One way to compare these algorithms is by their estimated mixing (ou unmixig) matrices : given numerous generated i.i.d. sources and a mixing matrix, all of these algorithms will return the best estimated mixing matrix according to their corresponding method. Thus we could compare the estimations with our original mixing matrix with any norm whatsoever, but as we previously said, the algorithms are not able to find out the original order of the sources or their amplitude.

Amari distance

Fortunately, the “Amari distance” (which is unfortunately not a distance), gives a criterion of proximity between two matrices regardless to these two problems (amplitude and permutation). If  $U$  and  $V$  are two  $n$ -by- $n$  matrices, the Amari distance is defined by:

$$d(U, V) = \frac{1}{2n} \sum_{i=1}^n \left( \frac{\sum_{j=1}^n |a_{ij}|}{\max_j |a_{ij}|} - 1 \right) + \frac{1}{2n} \sum_{j=1}^n \left( \frac{\sum_{i=1}^n |a_{ij}|}{\max_i |a_{ij}|} - 1 \right)$$

with  $a_{ij} = (UV^{-1})_{ij}$ . The closest it is to 0, the more  $U$  and  $V$  represents the same components.

Modus operandi

We will thereby analyze the results of the algorithms in terms of Amari distance over large number of i.i.d generated sources from several different distributions.

Introduction

- some items and  $\alpha = \gamma, \sum_i$
  - some items
  - some items
  - some items
- $\alpha = \gamma, \sum_i$

Introduction

- some items and  $\alpha = \gamma, \sum_i$
  - some items
  - some items
  - some items
- $\alpha = \gamma, \sum_i$

Introduction

- some items and  $\alpha = \gamma, \sum_i$
  - some items
  - some items
  - some items
- $\alpha = \gamma, \sum_i$

Introduction

- some items and  $\alpha = \gamma, \sum_i$
  - some items
  - some items
  - some items
- $\alpha = \gamma, \sum_i$

Introduction

- some items
- some items
- some items
- some items

Introduction

- some items and  $\alpha = \gamma, \sum_i$
  - some items
  - some items
  - some items
- $\alpha = \gamma, \sum_i$

Introduction

- some items and  $\alpha = \gamma, \sum_i$
  - some items
  - some items
  - some items
- $\alpha = \gamma, \sum_i$