
**https://swissroboticsday.ch/**


**Discount for students on registration and train tickets**

**https://swissroboticsday.ch/**

# APPLIED MACHINE LEARNING

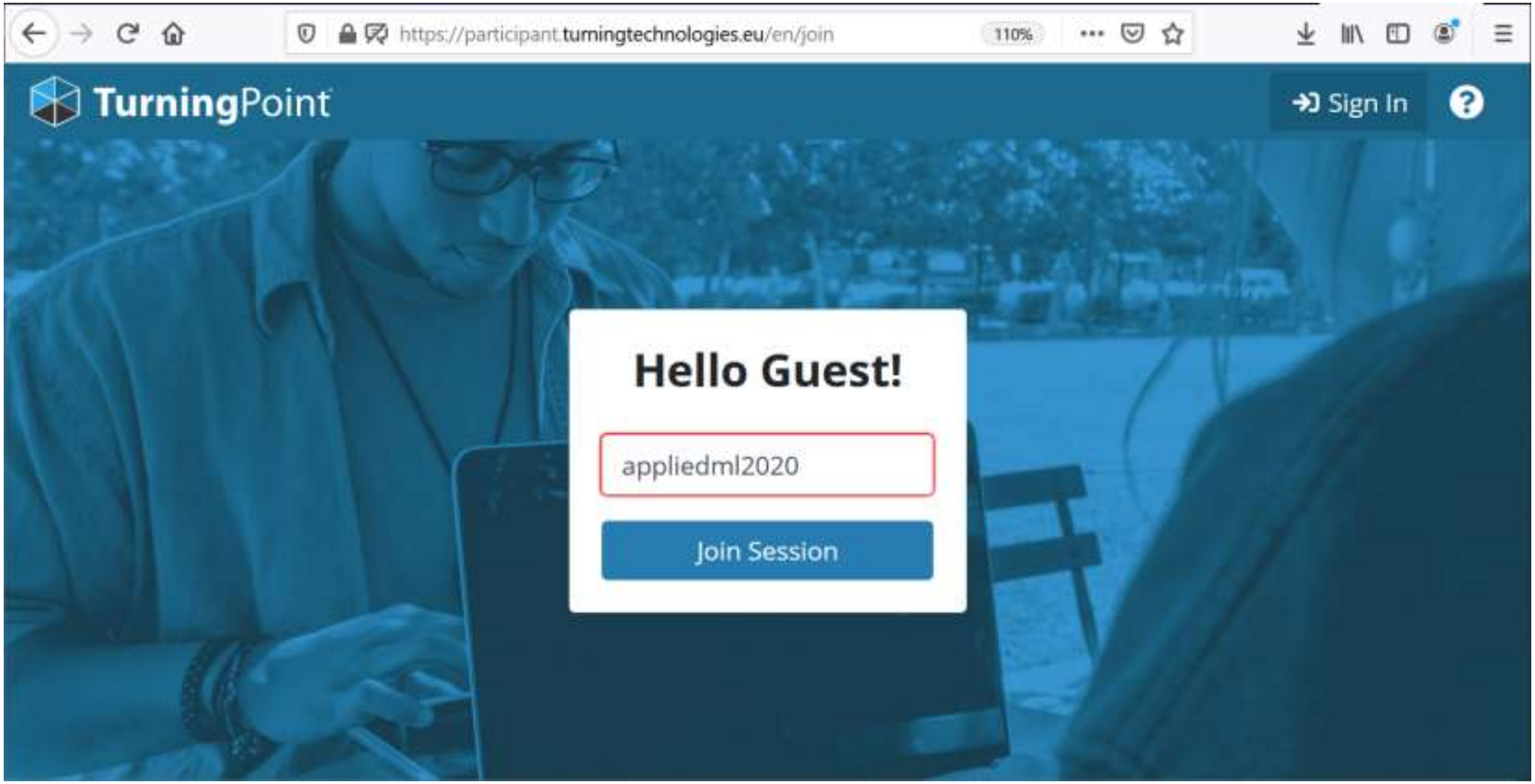
## *Independent Component Analysis (ICA)*

**Interactive Lecture**

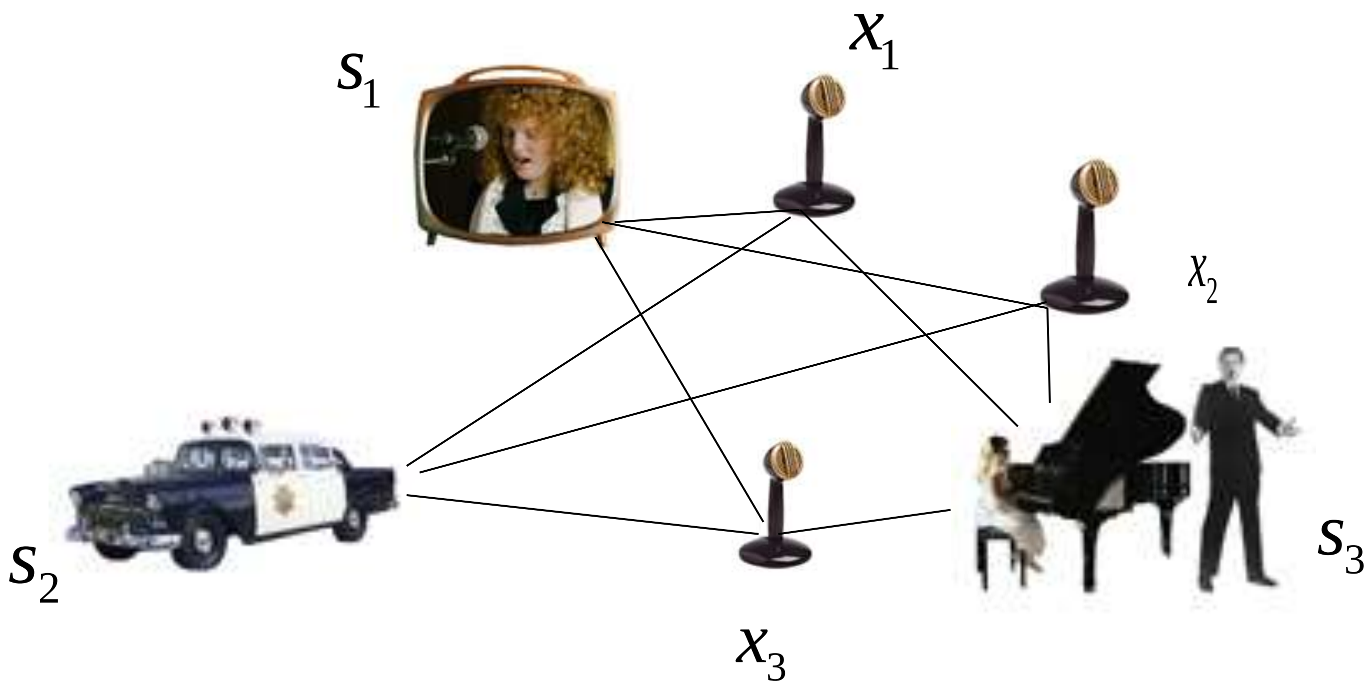
# Launch polling system

https://participant.turningtechnologies.eu/en/join

Acces as GUEST and enter the session id: *appliedml2020*



# ICA: Formalism



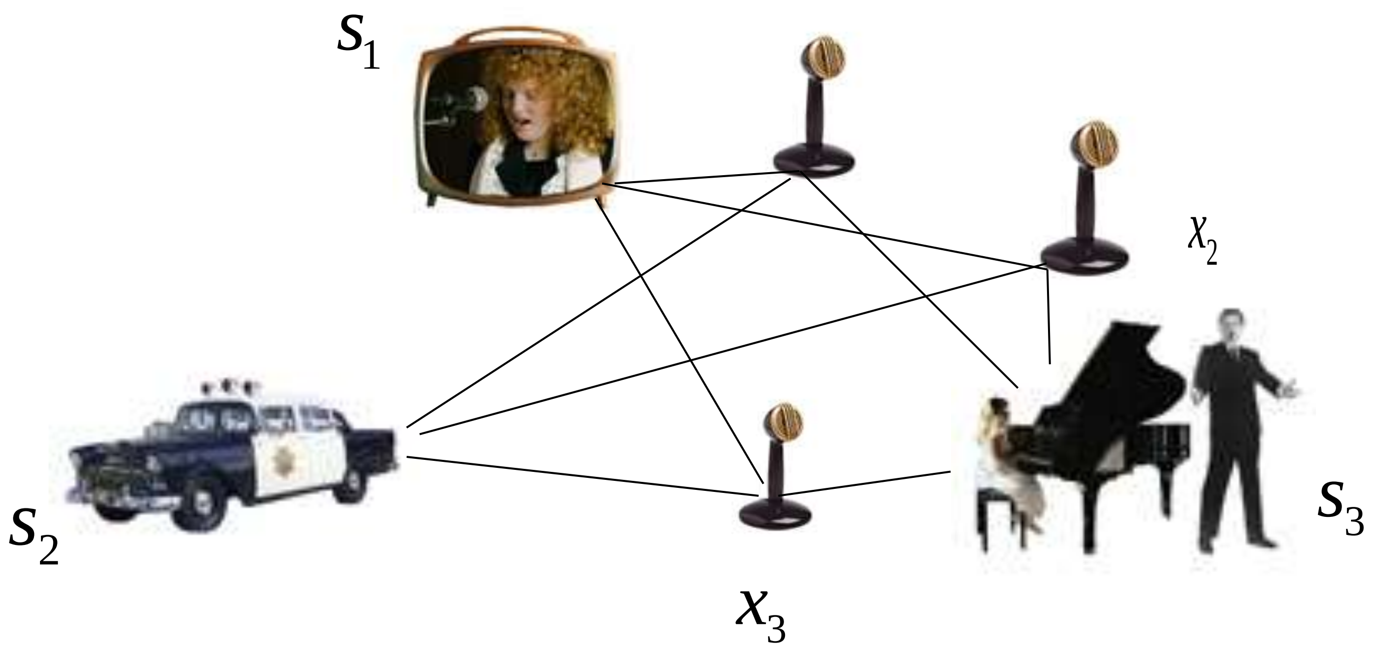

N-dimensional observation vector $x \in \mathbb{R}^N$, $N = 3$.

$x$ was generated by a linear combination of $N$ sources, $s \in \mathbb{R}^N$.

$x = As$, mixing matrix $A : N \times N$

*ICA* uncovers both $A$ and $s$. $s$ are called the independent components.

N-dimensional observation vector $x \in \mathbb{R}^N$ and sources s $\in \mathbb{R}^N$, $N = 3$.

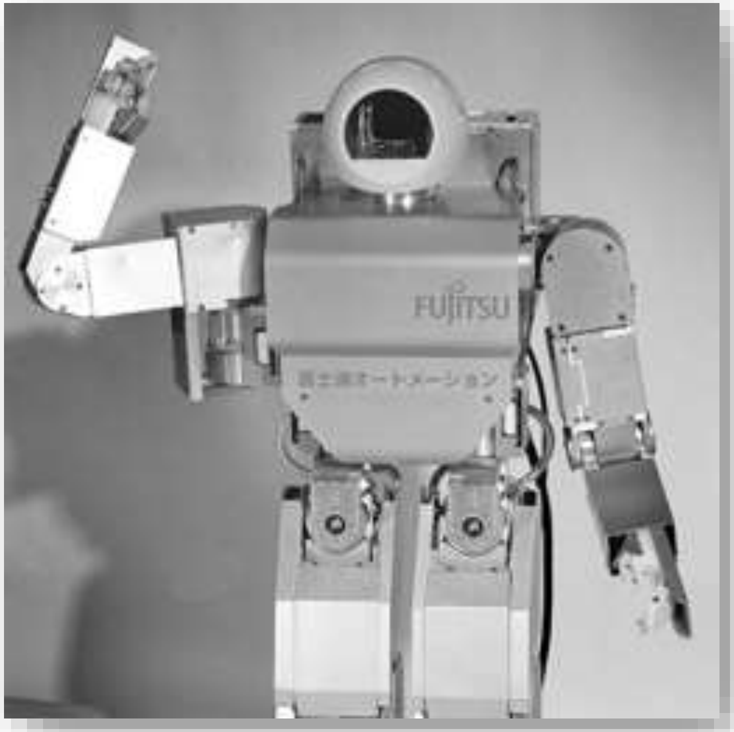



But at each time step of the recording, we obtain a new observation of both mixtures and sources.

If we do T measurements, the total dimension of the dataset is X: $N \times T$

# ICA: Application to Image Decomposition

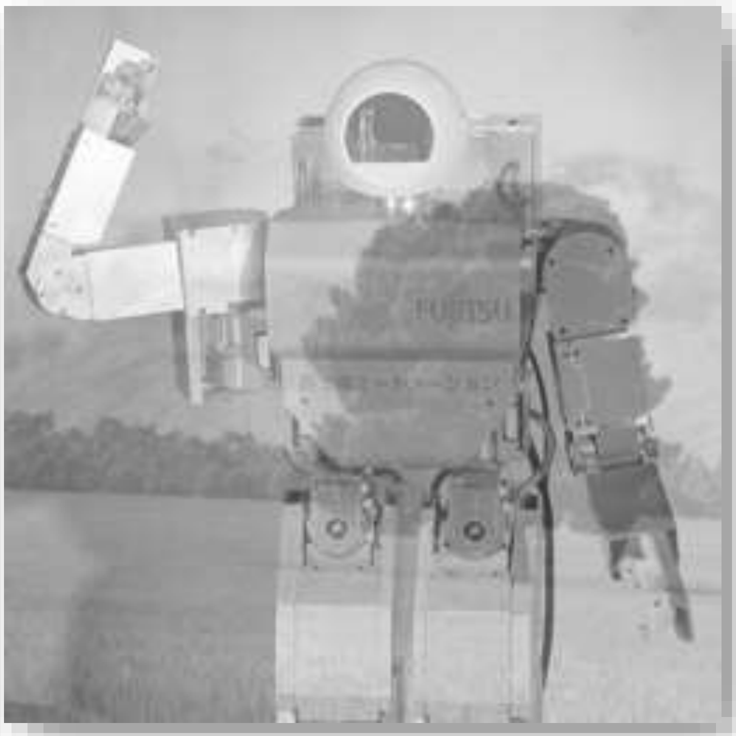
Source 1: S1




Source 2: S2

# ICA: Application to Image Decomposition

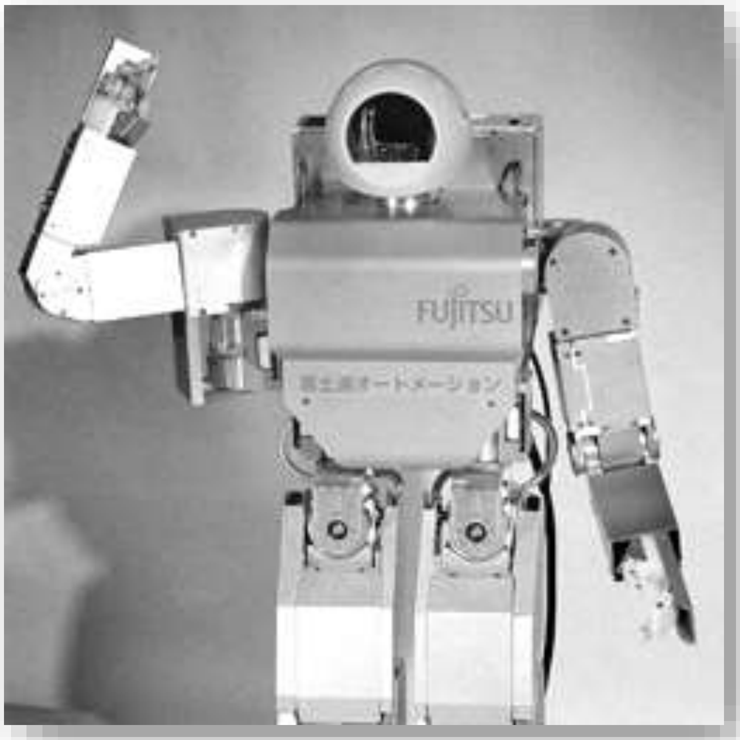
Mixture 1: X1



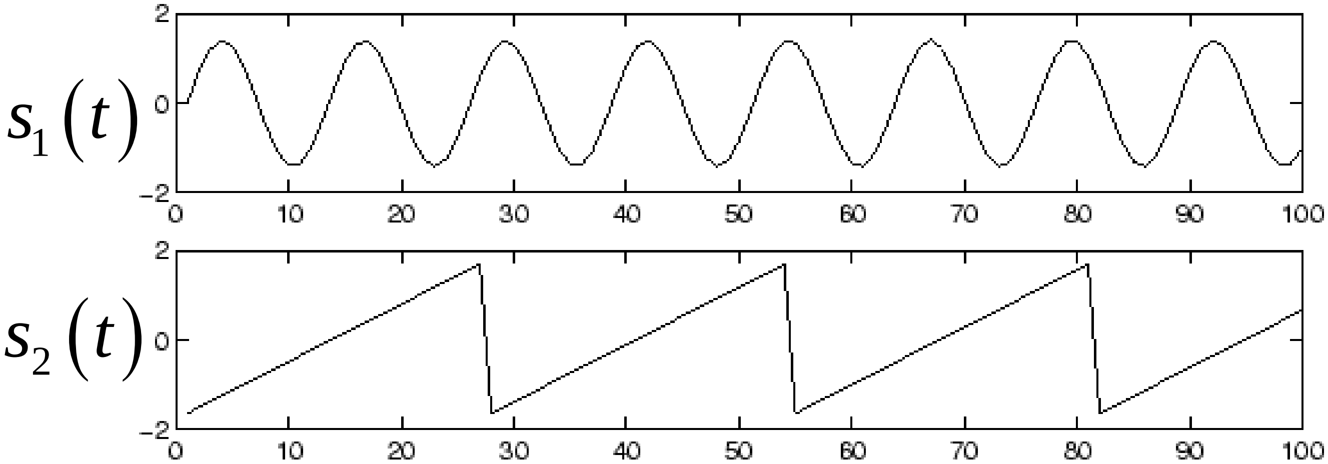
Mixture 2: X2

# ICA: Application to Image Decomposition

Reconstructed source S1

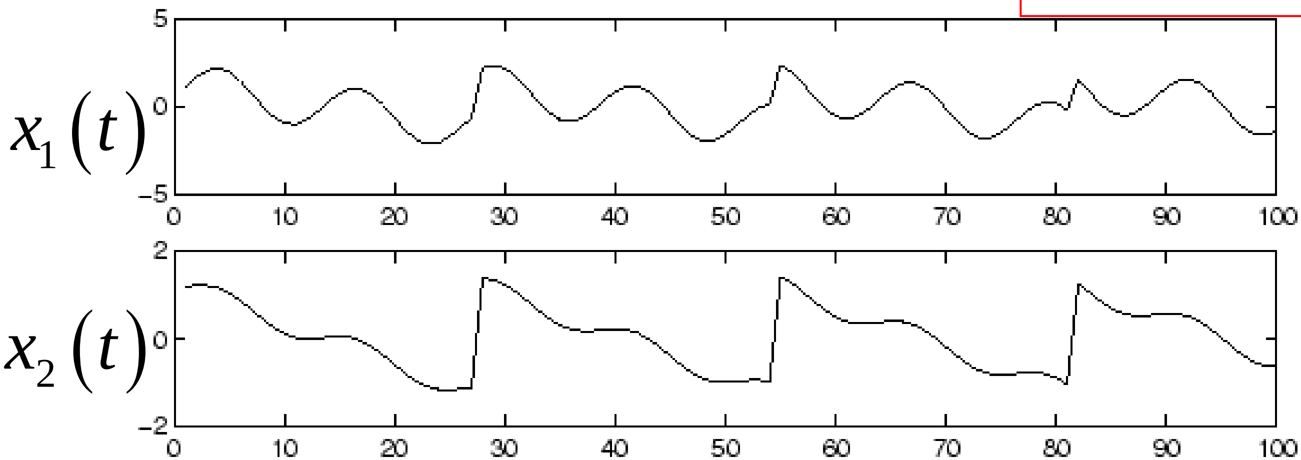



Reconstructed source S2




**Sources**

$$x(t) = As(t)$$

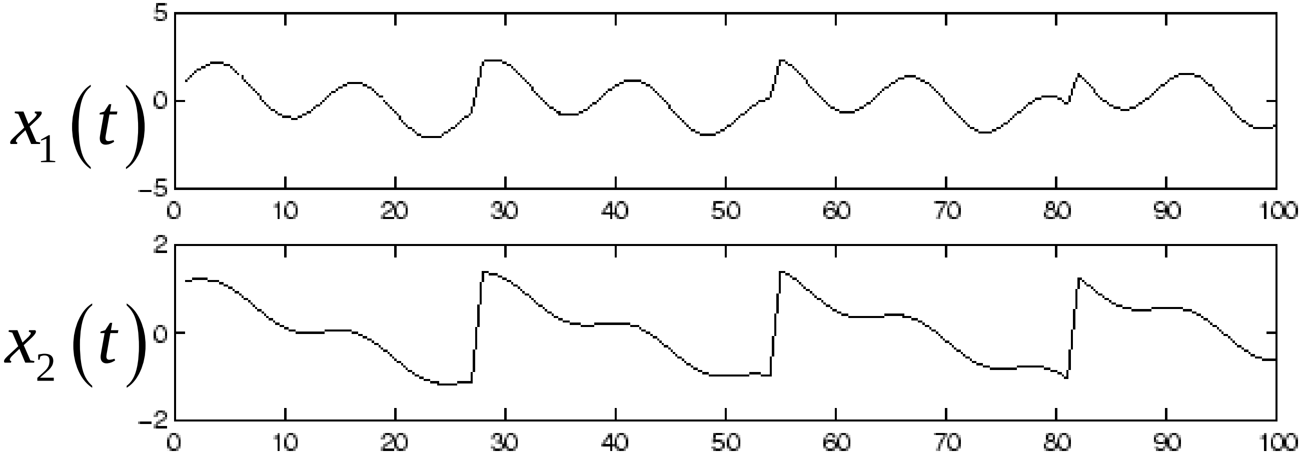



**Mixed Signal**

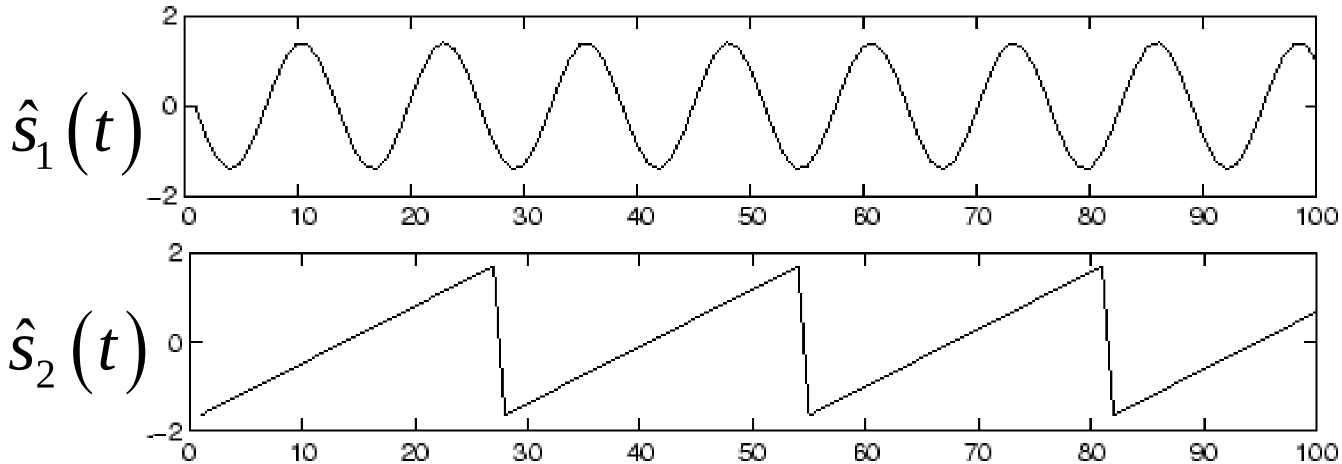
$$s(t) = A^{-1}x(t)$$

$$s(t)?$$

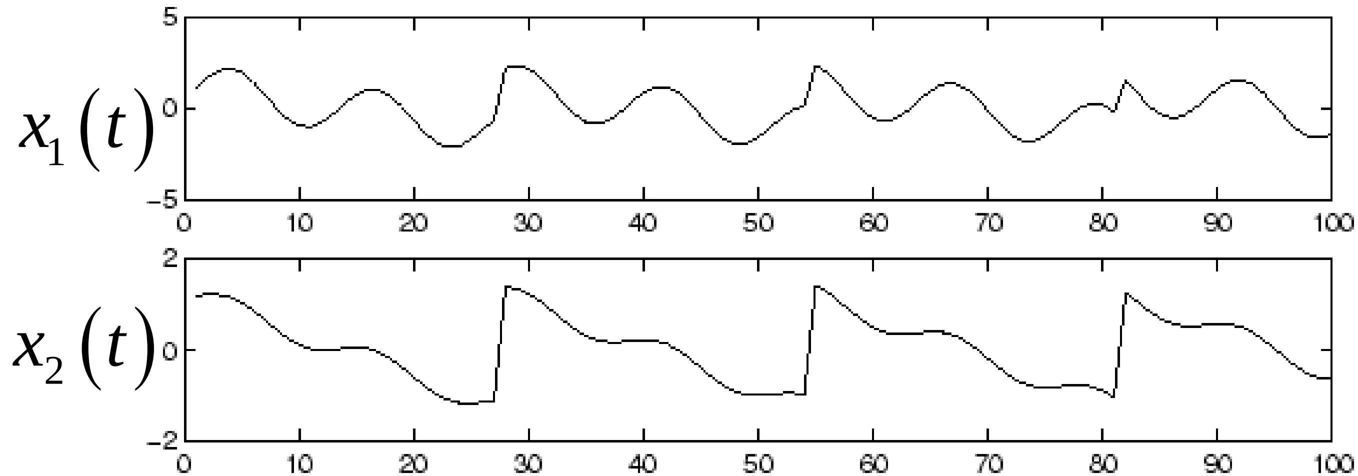
$$A^{-1}?$$




**Mixed Signal**




**Estimated Sources**

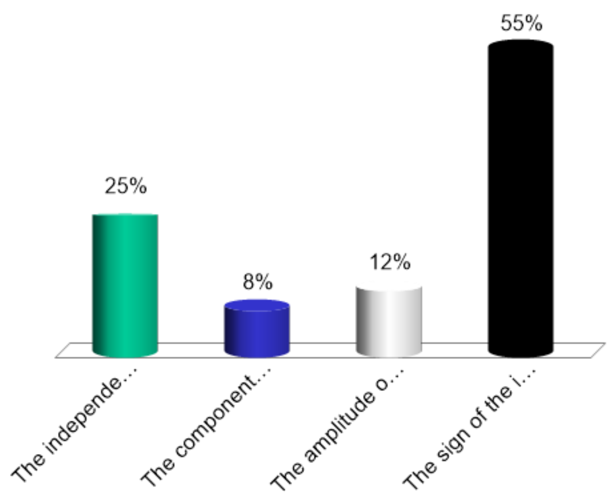



**Mixed Signal**

# *ICA: Limitations*

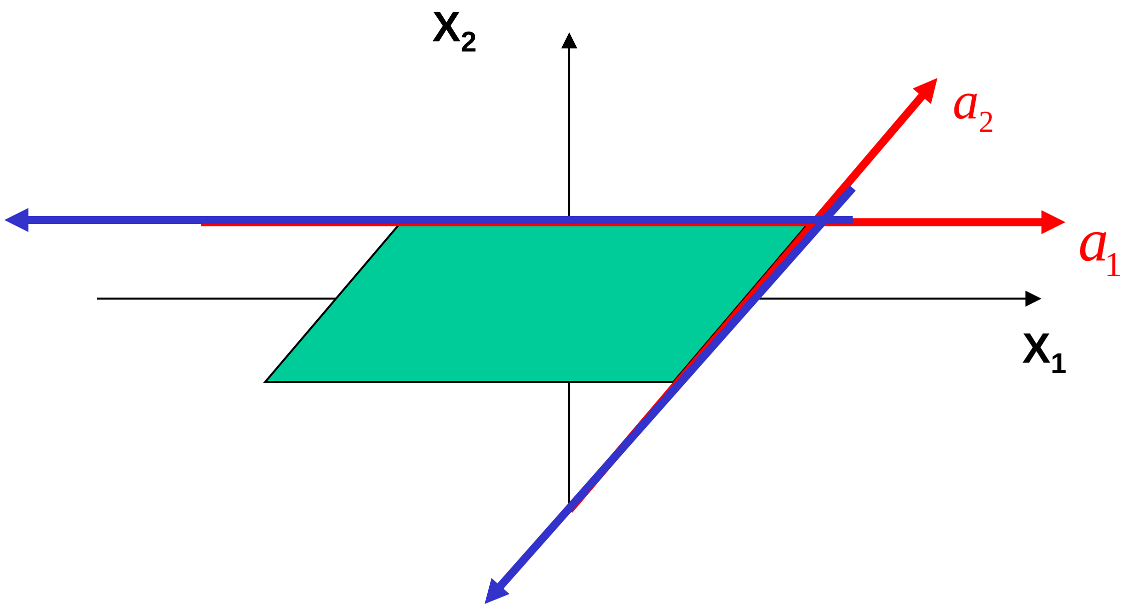
Which of the following statements are true?

A. The independent components are unique.
B. The components can be ordered according to their (statistical) importance.
C. The amplitude of the independent components (vectors norm) is unknown.
D. The sign of the independent components (vectors) is unknown.



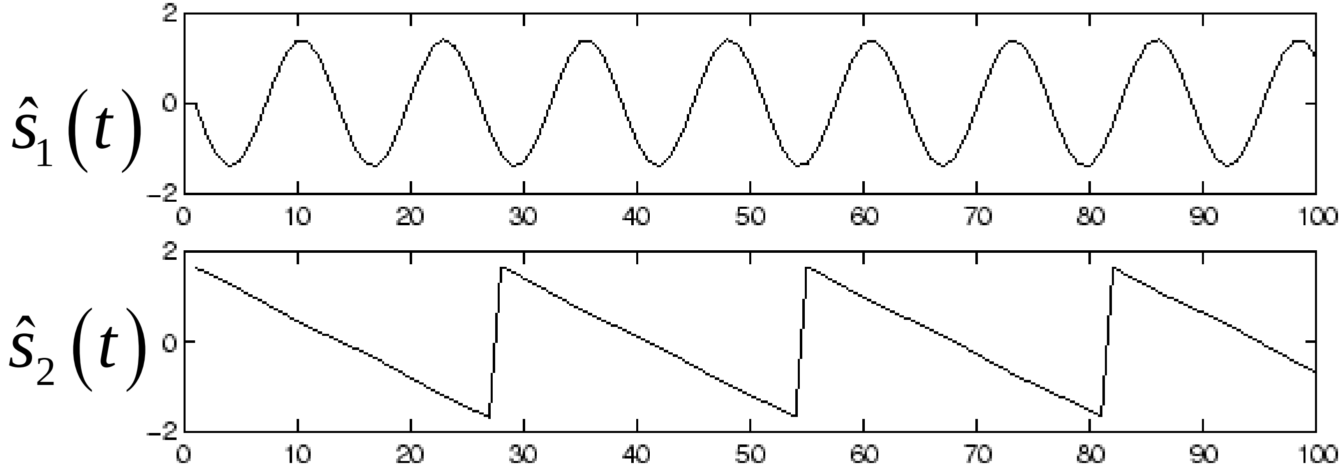
# ICA: Finding mixing matrix - intuition

$x = A \cdot s, \;\; A$ and $s$?

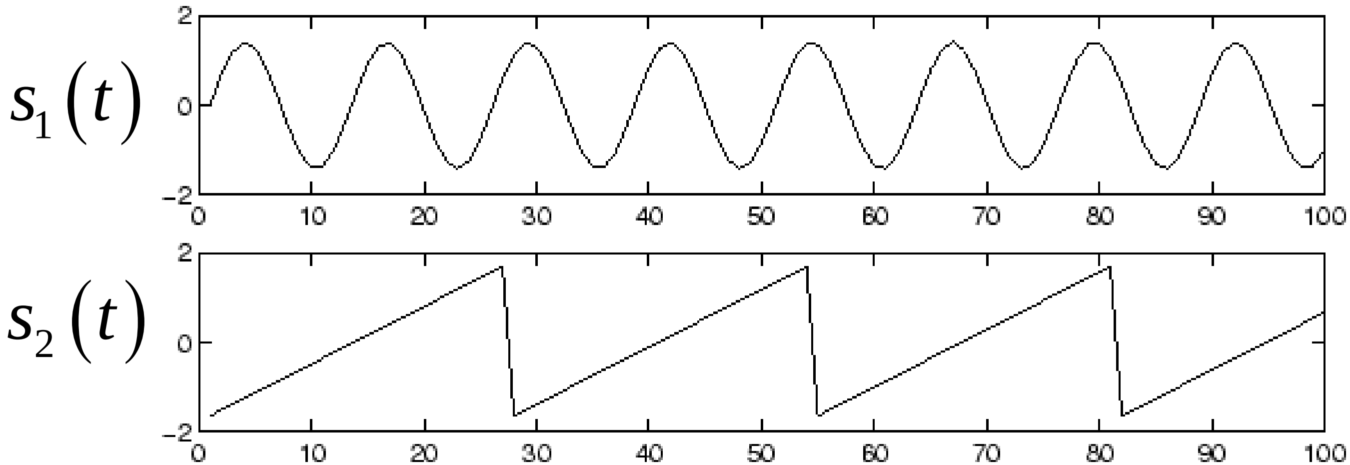



$s$ and $A$: known up to a scaling factor

## Estimated Sources




**One of the original and estimated source is inverted!**



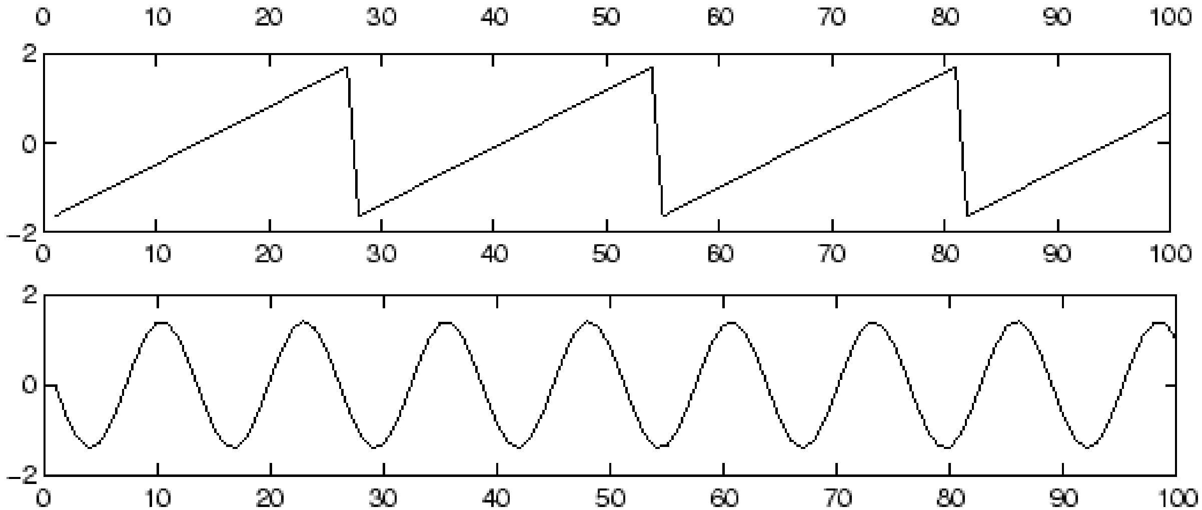

## Original Sources

## Solution after another ICA run

**Estimated Sources**

$\hat{s}_1(t)$

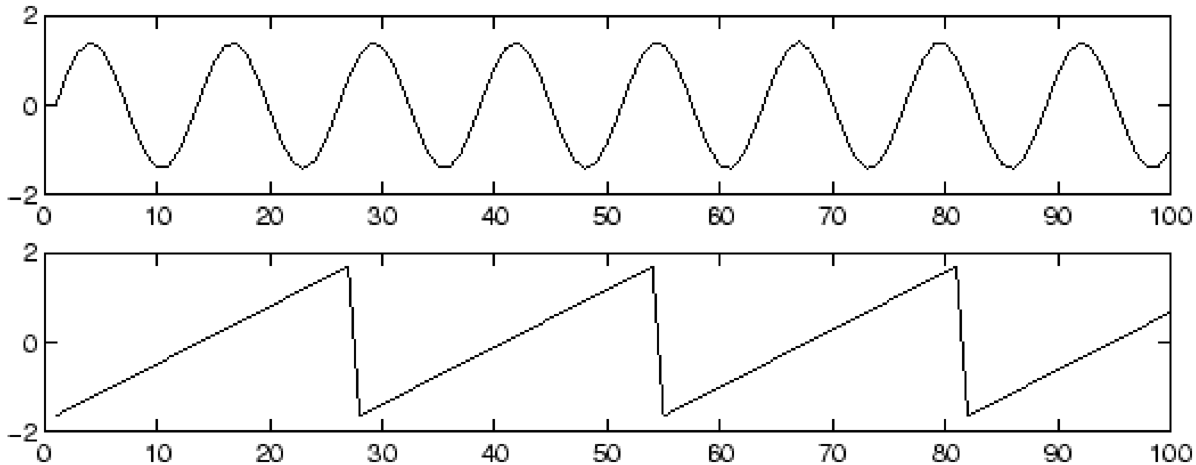
$\hat{s}_2(t)$




**The ordering is inverted!**

$s_1(t)$

$s_2(t)$




**Original Sources**

# *ICA: Optimization*

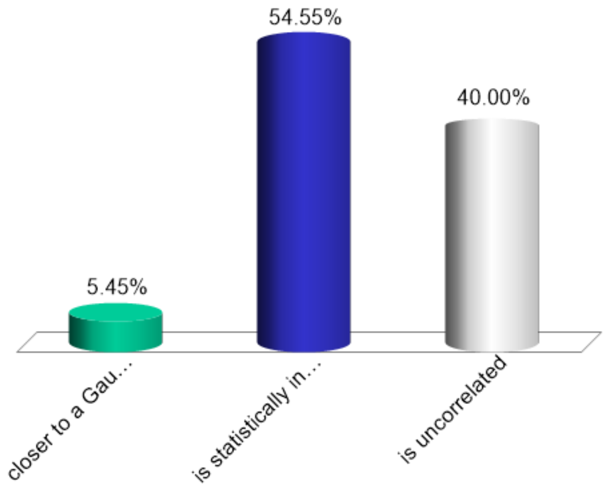
Which of the following statements are true?

ICA searches projections such that the distribution of the projected dataset is

A. closer to a Gauss distribution

B. is statistically independent

C. is uncorrelated



# Statistical Independence and and uncorrelatedness

$$\textbf{Independent} \quad \underset{\not\longleftarrow}{\longrightarrow} \quad \textbf{Uncorrelated}$$

$$p(x_1, x_2) = p(x_1)\,p(x_2) \quad \Rightarrow \quad E\{x_1, x_2\} = E\{x_1\}E\{x_2\}$$

$$p(x_1, x_2) = p(x_1)\,p(x_2) \quad \not\Leftarrow \quad E\{x_1, x_2\} = E\{x_1\}E\{x_2\}$$

Statistical independence ensures uncorrelatedness.
The converse is not true

# *ICA Optimization: Assumptions*

Which of the following statements are true?

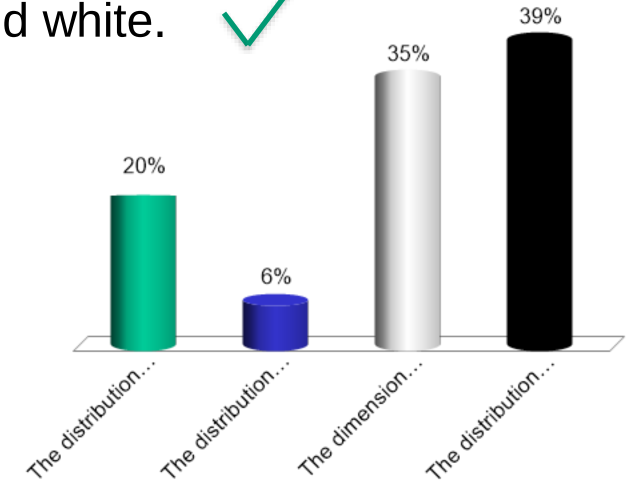
A. The distribution of the observables (mixed variables) follows a Gauss distribution.

B. The distribution of the sources follows a Gauss distribution.

C. The dimension of the sources equals the dimension of the observables.

D. The distribution of the observables is centered and white.

*B: If it is known that number of sources is smaller, apply PCA on the observables and reduce dimensionality to expected sources' dimension.*

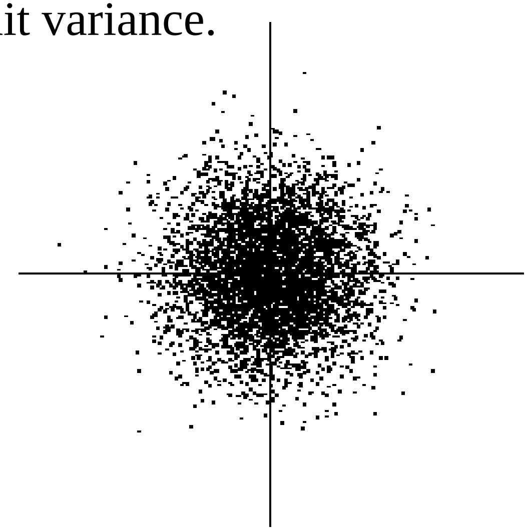


# Why Gaussian distributions for sources are forbidden

A fundamental restriction in ICA is that
the independent components must be non-Gaussian.

Assume that the mixing matrix is orthogonal and the $\boldsymbol{s}_i$ are Gaussian.
Then $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are also Gaussian, uncorrelated, and of unit variance.
Their joint density is given by:

$$p\left(x_1, x_2\right) = \frac{1}{\sqrt{2\pi}} e^{-\left(x_1^2 + x_2^2\right)^2}$$



The joint density is completely symmetric. Therefore, it does not contain any information on the directions of the columns of the mixing matrix!

→ The mixing matrix cannot be estimated

# ICA: Preprocessing Steps – Summary

- ❖ **Centering →** Both the dataset and the sources (latent components) are zero mean.

- ❖ **Whitening →** Find out projections that embed correlations through PCA. Project onto PCA basis and scale data to unit variance. The dataset is now uncorrelated; its variance is 1 along all dimensions.

→ Simplifies computation of the independent components (Kurtosis computation)

# ICA: Hypotheses & Identification of Independent Components – Summary

- ❖ Sources and data have the *same dimension*.

  → The unknown mixing matrix is square. If this is not the case, use PCA to reduce the dimension of the sources and by extension the ICA components.

- ❖ The sources are assumed to be *statistically independent*. They also must follow a *non-Gaussian* distribution. If so, a measure of non-Gaussianity is an indication that the sources are close to be statistically independent sources.

  → Optimize non-Gaussianity measure for the distribution of the sources.

  → Find the sources iteratively, one projection at a time.

# Kurtosis / Entropy Value for Gauss Distributions

ICA uses either Kurtosis or Negentropy to find independent components. It uses the fact that these two measures have explicit bounds for the equivalent Gauss distribution.

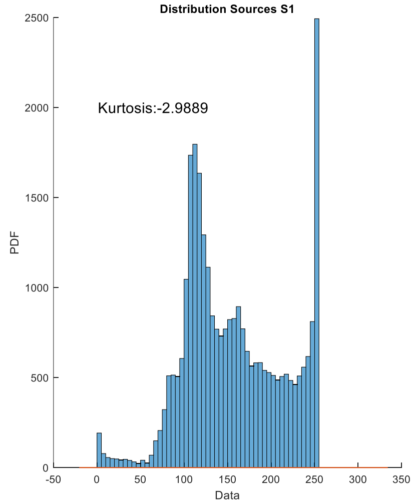
**Kurtosis:**
The kurtosis for the Gauss distribution has a closed-form expression and is equal to 3.0.
*Note that, in this course, we follow computer vision standard and assign zero for the kurtosis of the normal distribution by substracting 3.0 from the kurtosis.*

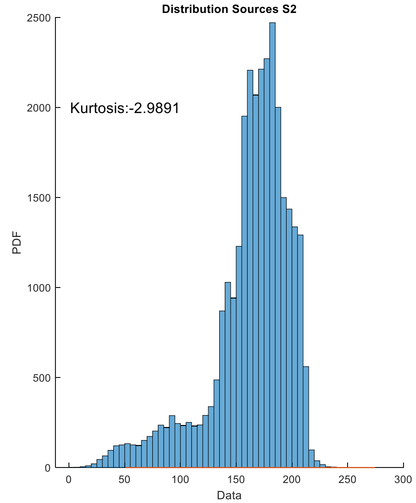
**Entropy:**
The entropy of a Gauss distribution is larger than that of any other distribution. Moreover, it has a closed-form solution, see exercise session.

Hence, ICA Maximizes Neg-Entropy!

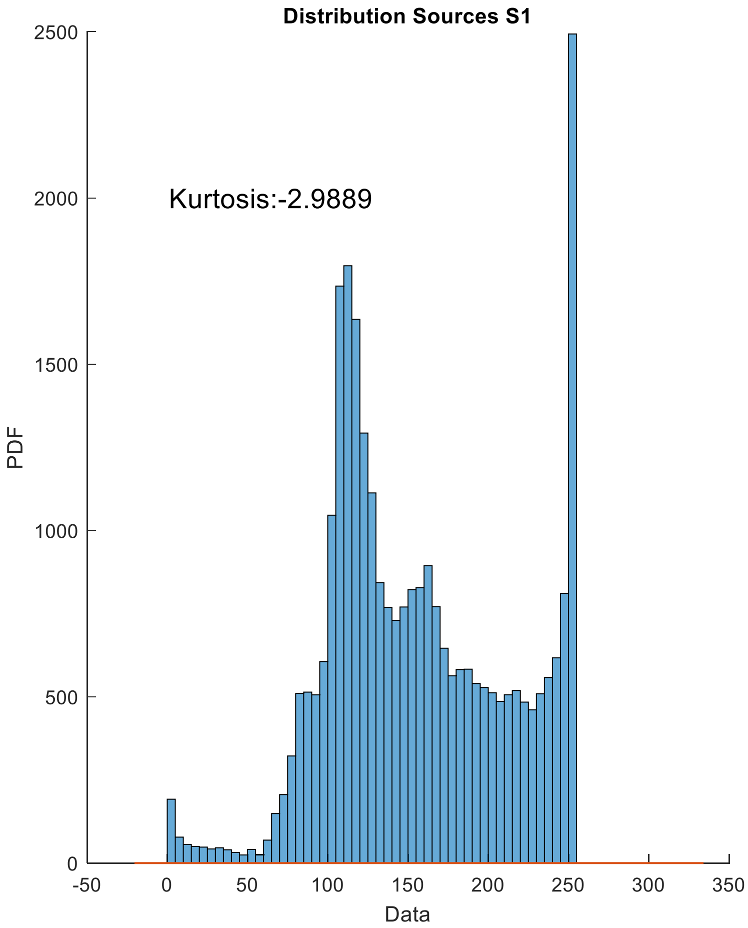
# Real sources
# pdf is non-Gaussian

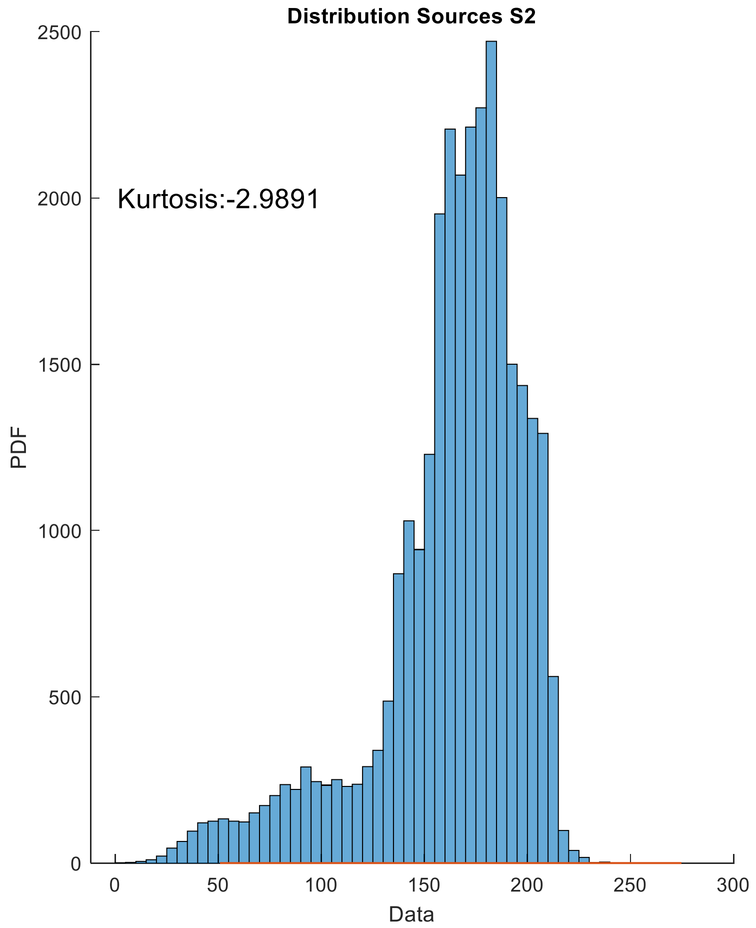
Sources S1, S2



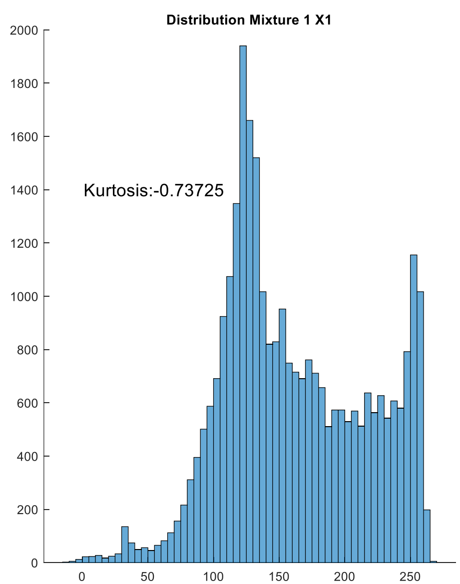
# Real sources
# pdf is non-Gaussian

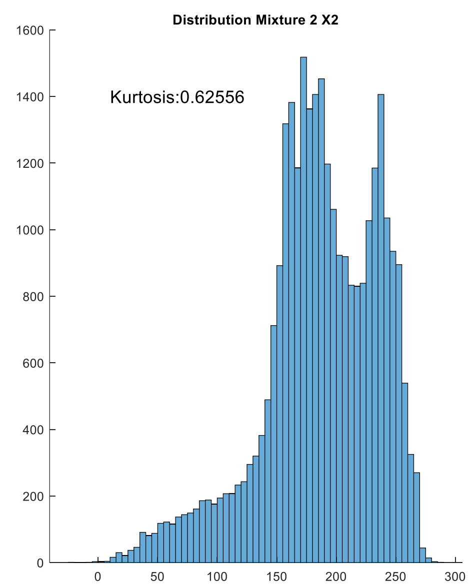


# Pdf of Mixtures are closer to a Gauss distribution

Mixtures x1, x2



# Pdf of Mixtures are closer to a Gauss distribution