ADA8

Sparsity Tutorial

Samuel Farrens

22nd May 2016 - Χανιά

OVERVIEW

Inverse Problems

Sparsity

Wavelets

Blind Source Separation

Inverse Problems

INVERSE PROBLEMS

Data $\longrightarrow$ Model Parameters

$$Y = MX$$

$$X = M^{-1}Y$$

$$X = (M^TM)^{-1}M^TY$$

OVERVIEW

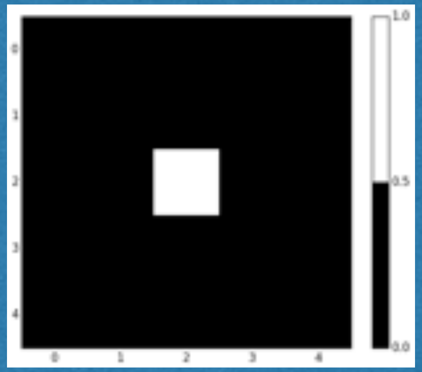
- Inverse Problems
- Sparsity
- Wavelets
- Blind Source Separation

Sparsity

SPARSITY

A sparse signal is one comprised mostly of zeros.

$$X = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$



A signal maybe sparse in a different domain.

$$X = \begin{bmatrix} 0.04 & -0.032 & 0.012 & 0.012 & -0.032 \\ -0.032 & 0.012 & 0.012 & -0.032 & 0.04 \\ 0.012 & 0.012 & -0.032 & 0.04 & -0.032 \\ 0.012 & -0.032 & 0.04 & -0.032 & 0.012 \\ -0.032 & 0.04 & -0.032 & 0.012 & 0.012 \end{bmatrix} \rightarrow \mathcal{F}(X) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

OVERVIEW

Inverse Problems

Sparsity

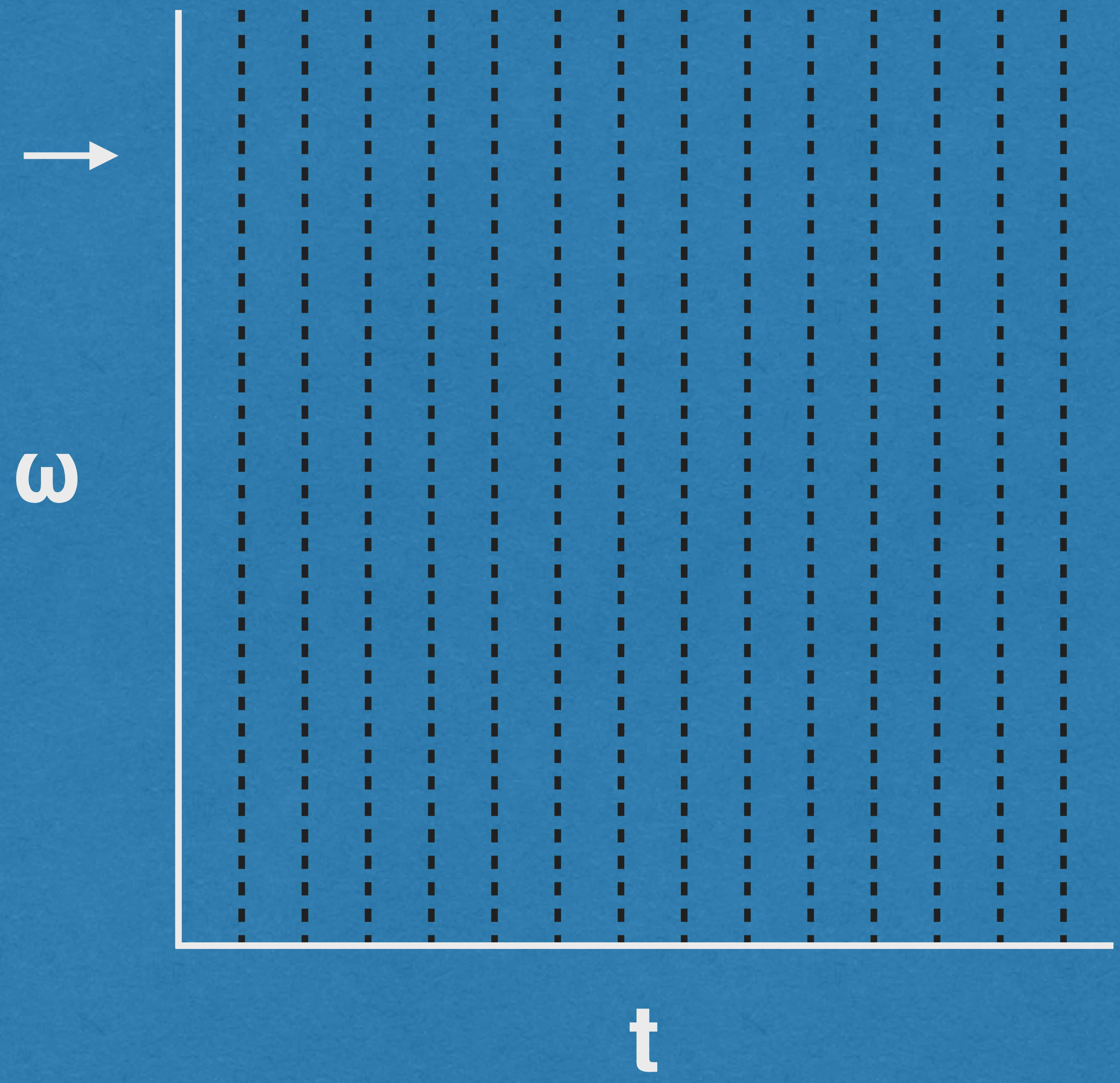
Wavelets

Blind Source Separation

Wavelets

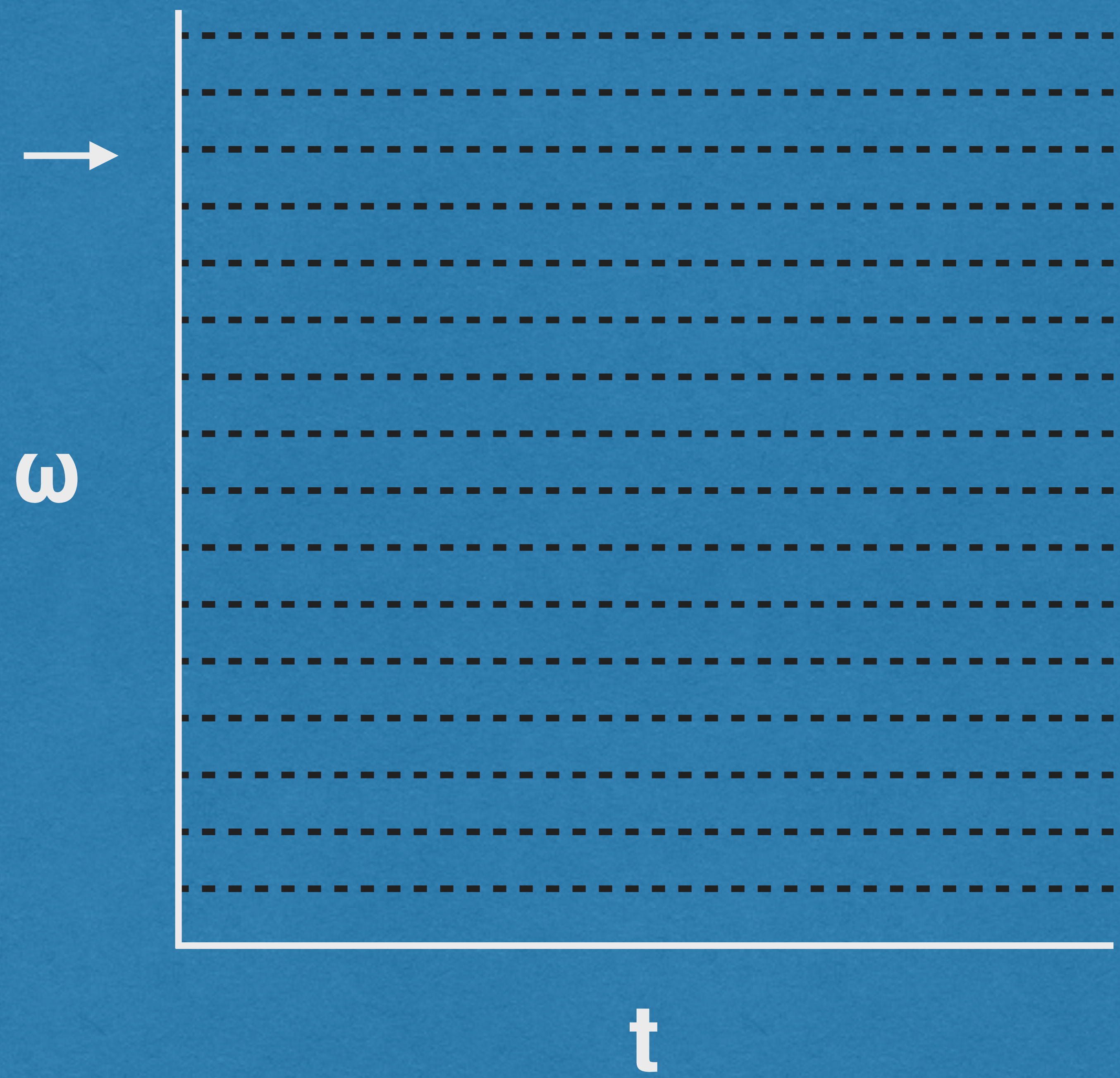
WAVELETS

No frequency information



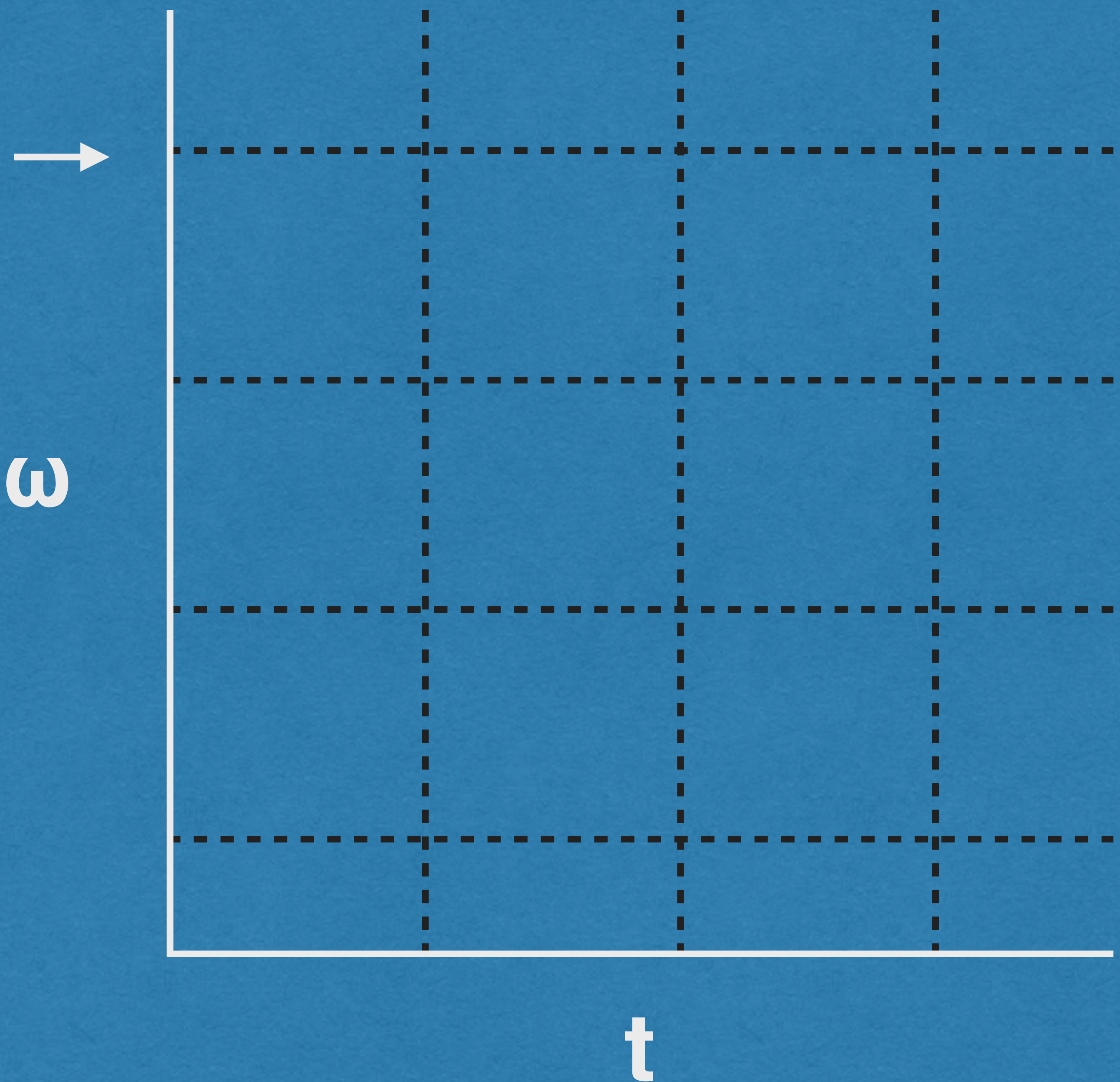
WAVELETS

No time information



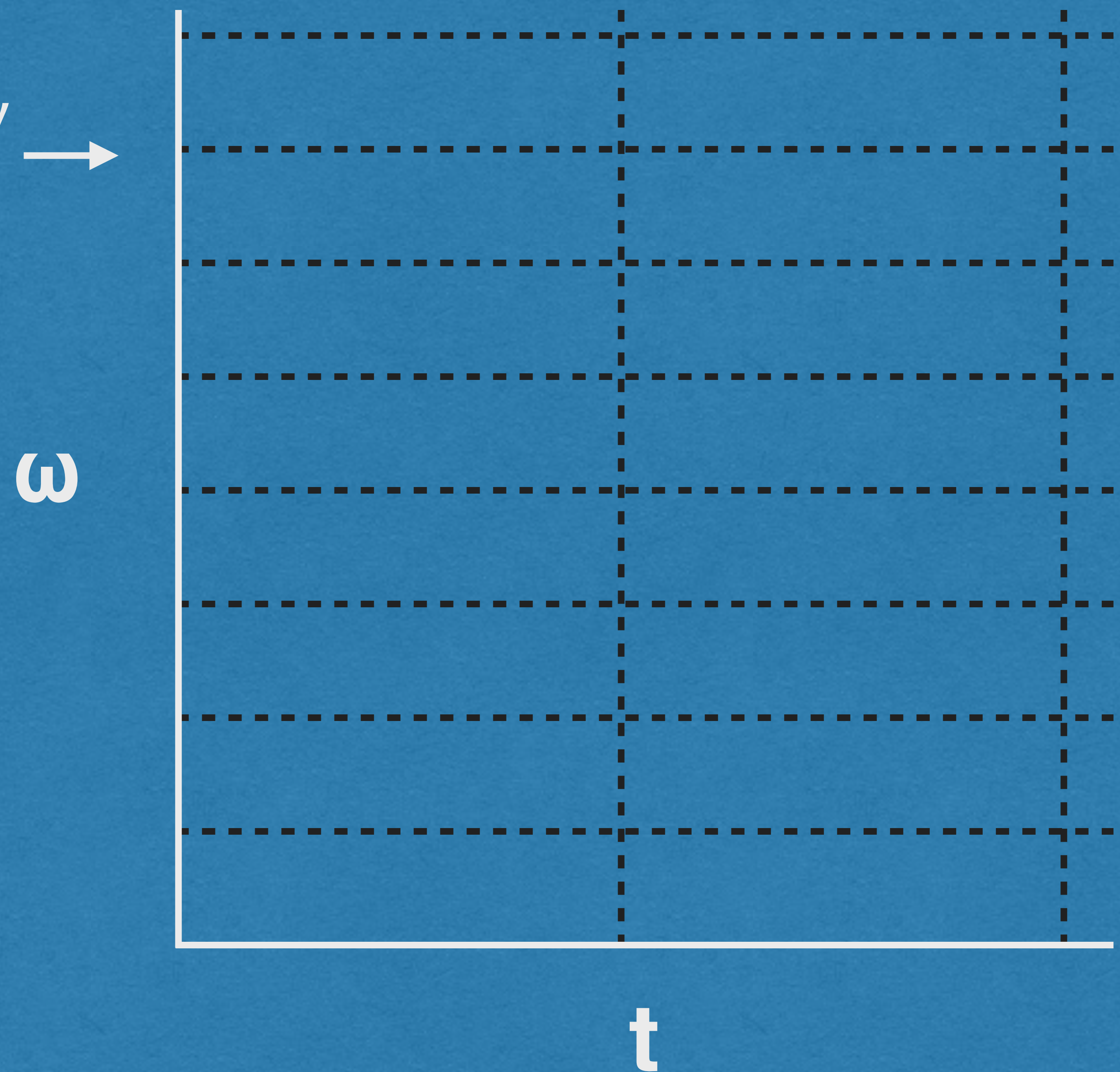
WAVELETS

Some time and frequency information



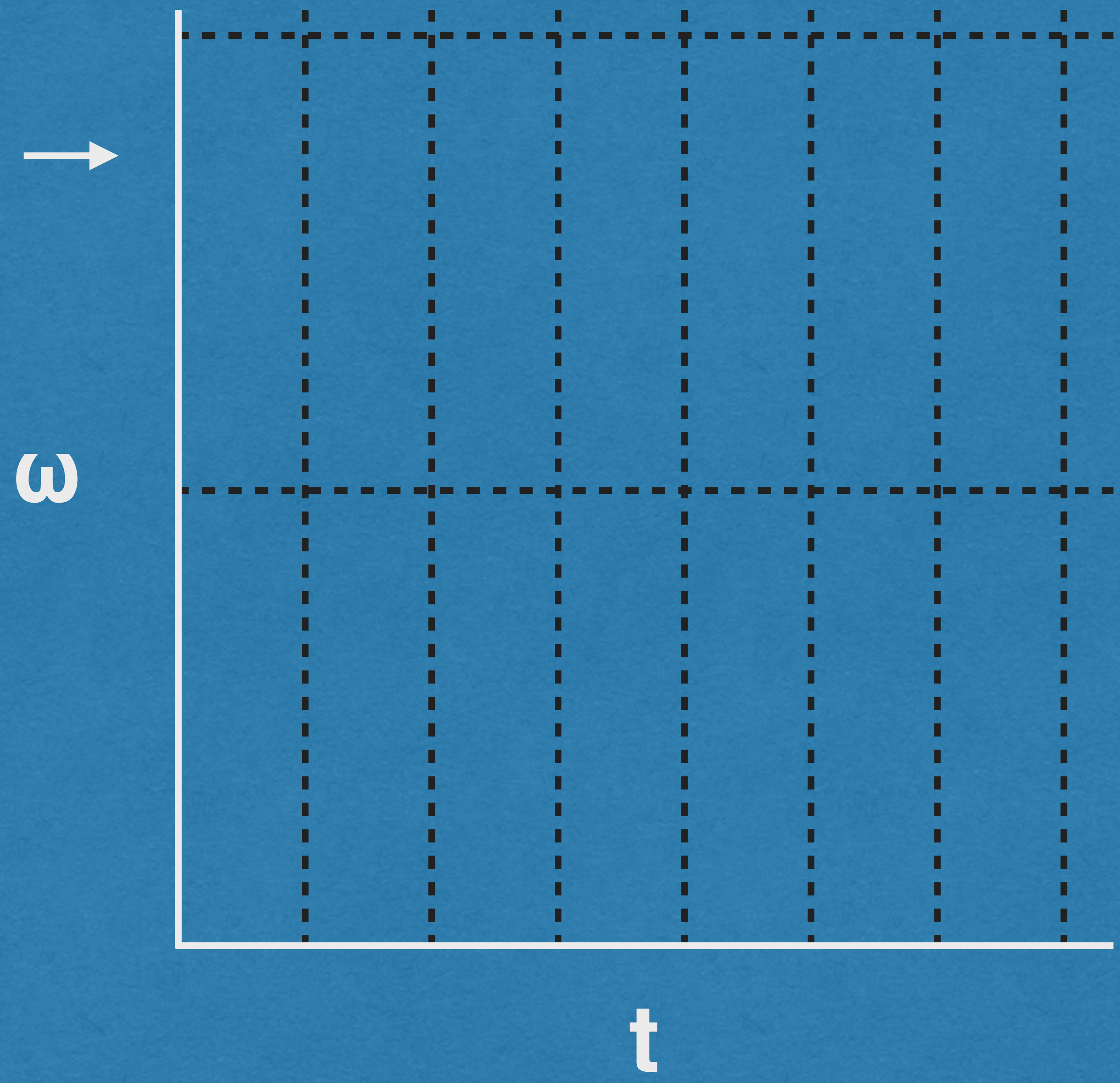
WAVELETS

More frequency
less time



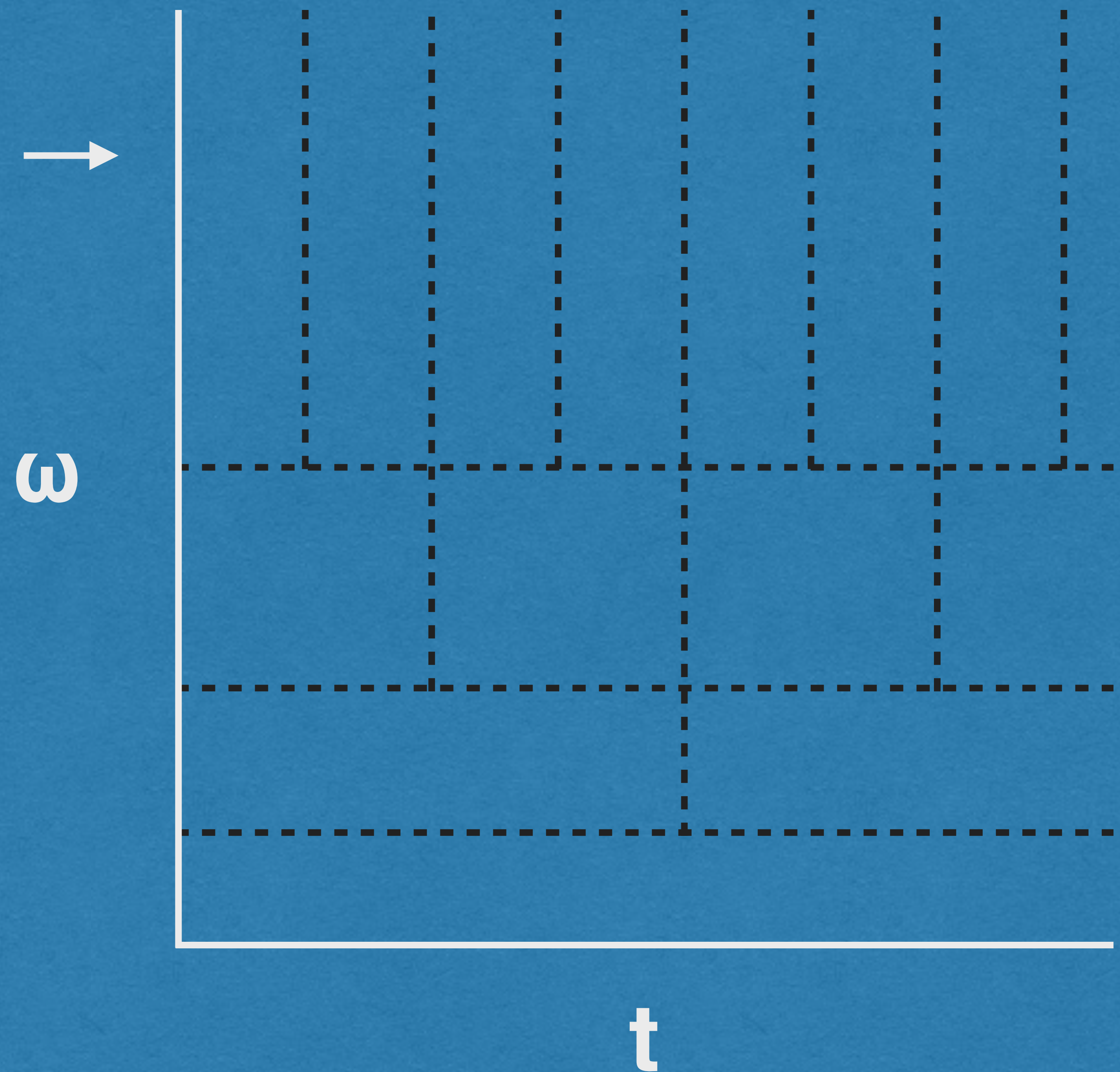
WAVELETS

More time
less frequency



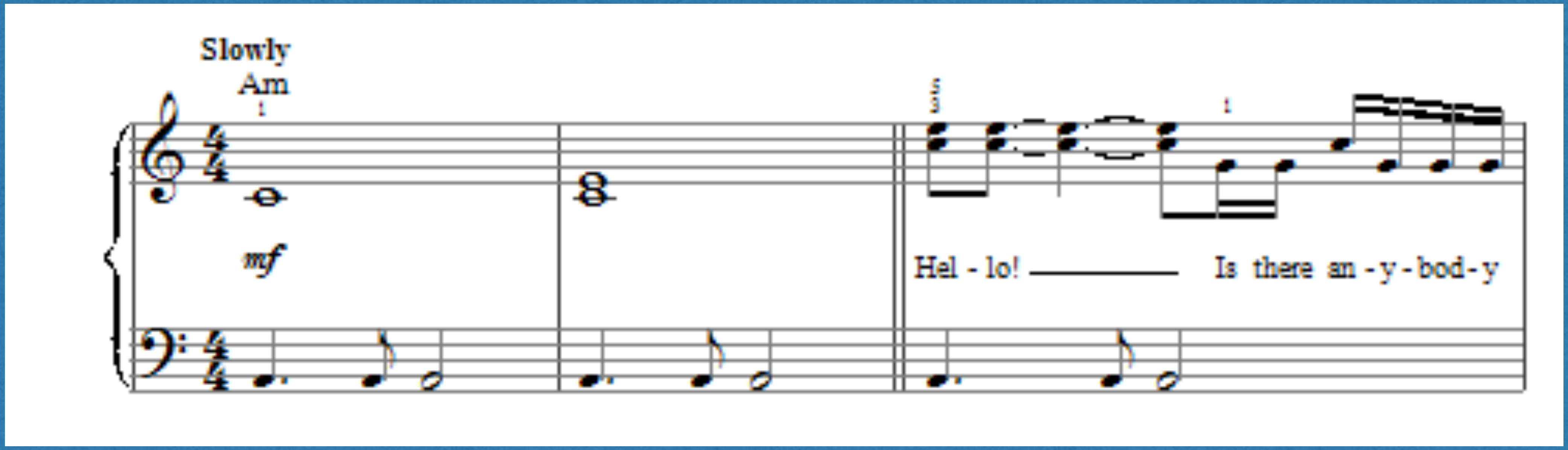
WAVELETS

Multiesolution Analysis



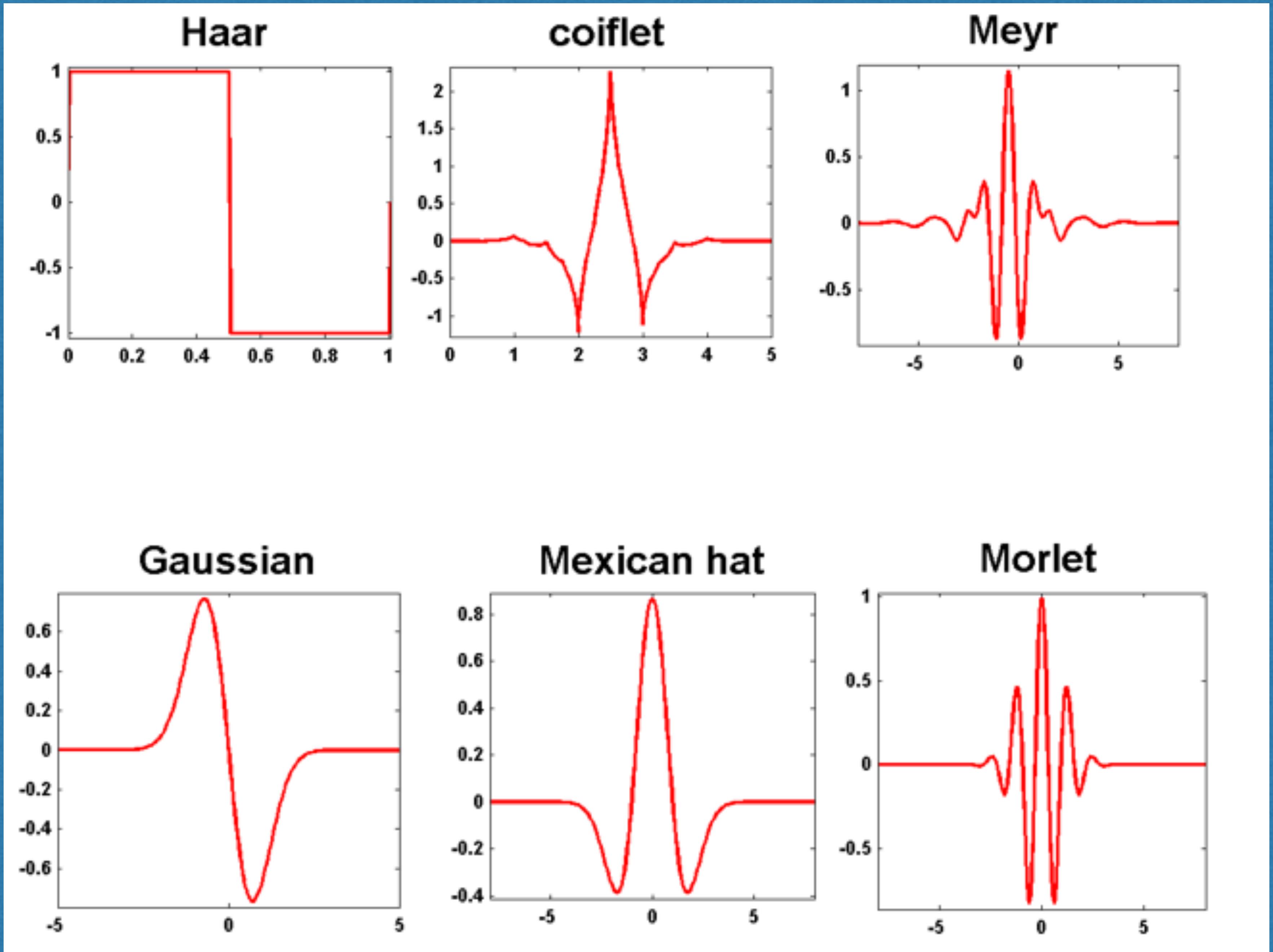
WAVELETS

Time-Frequency Trade-Off



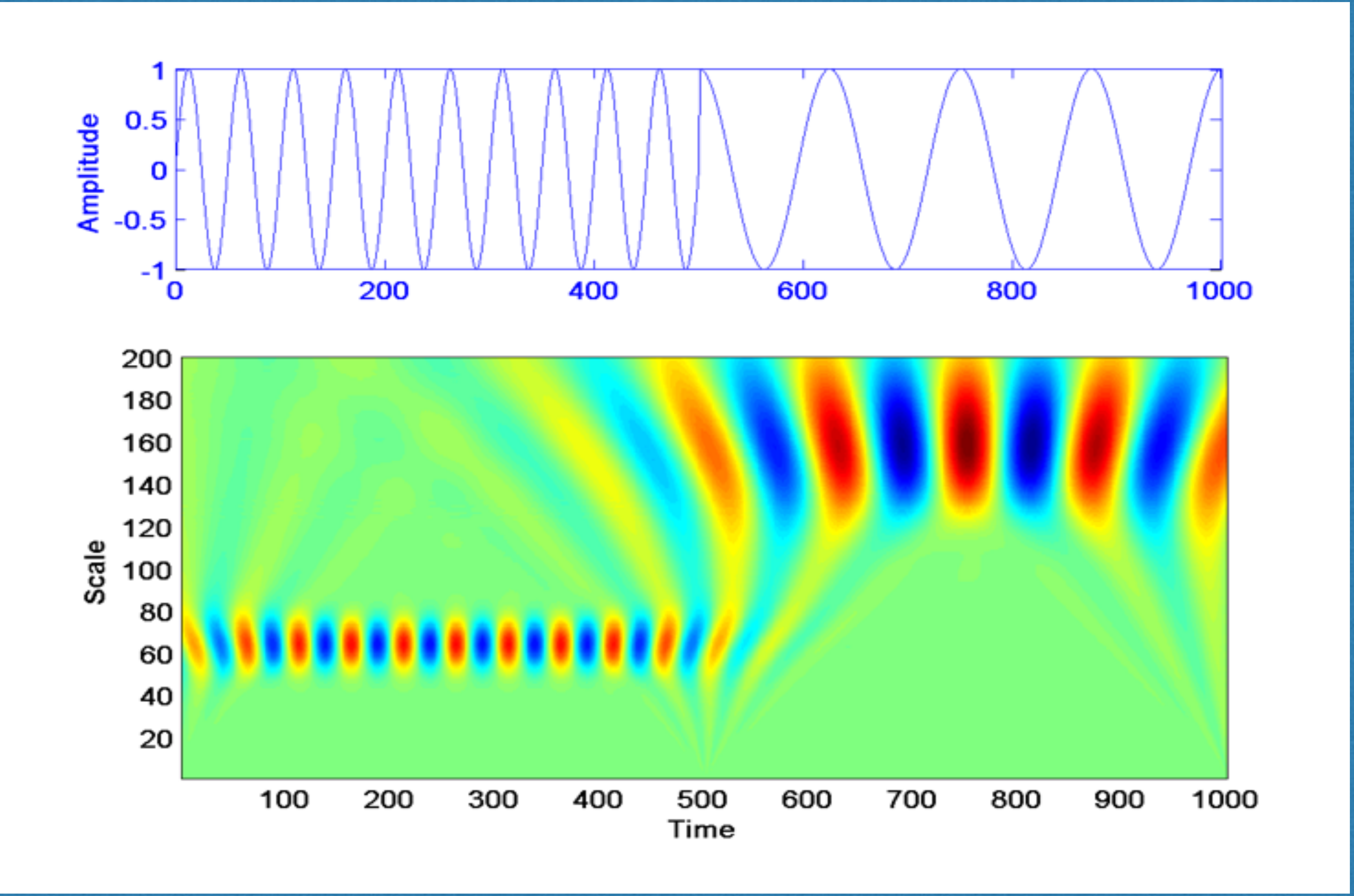
WAVELETS

Wavelet Shapes

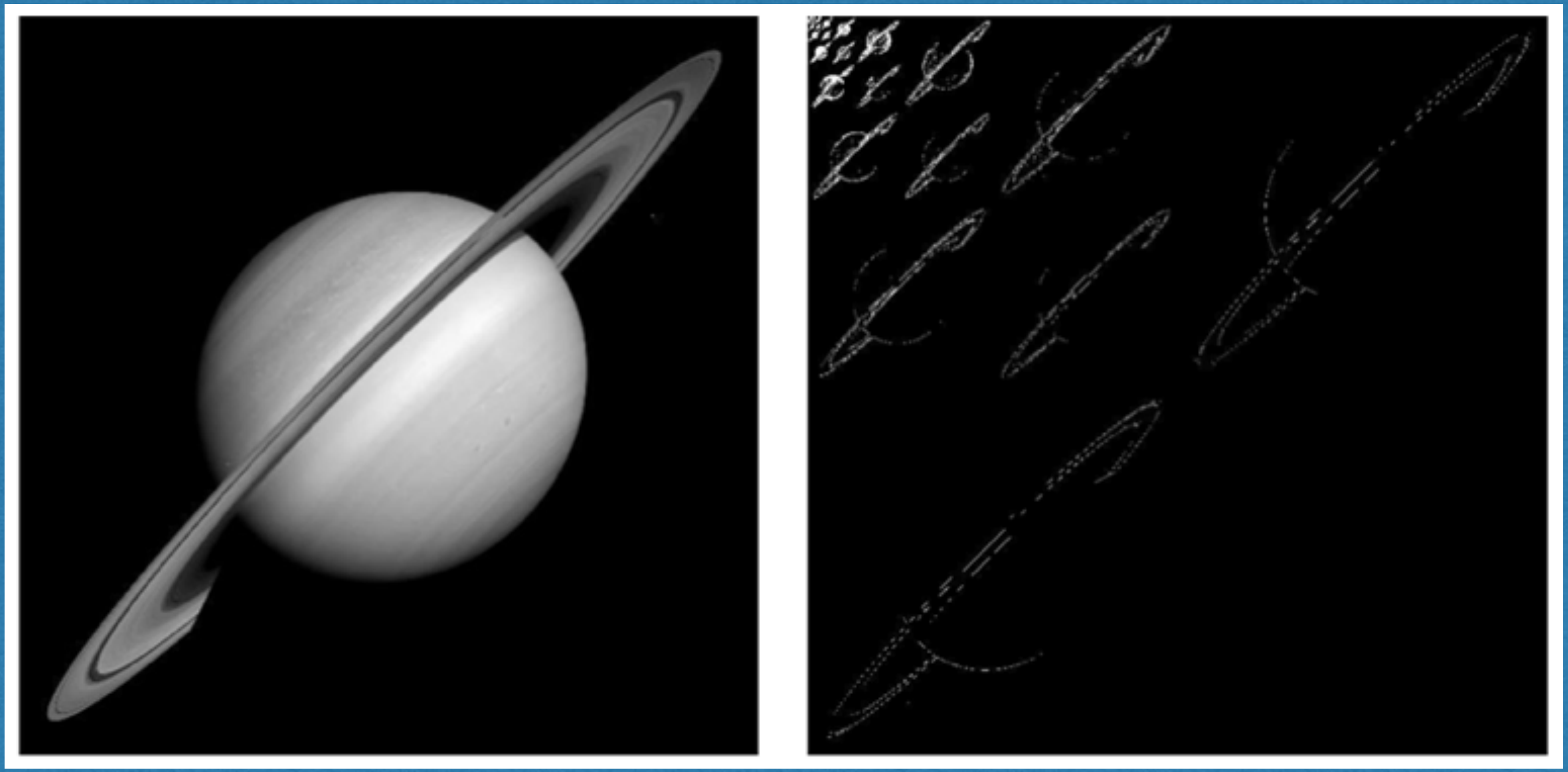


WAVELETS

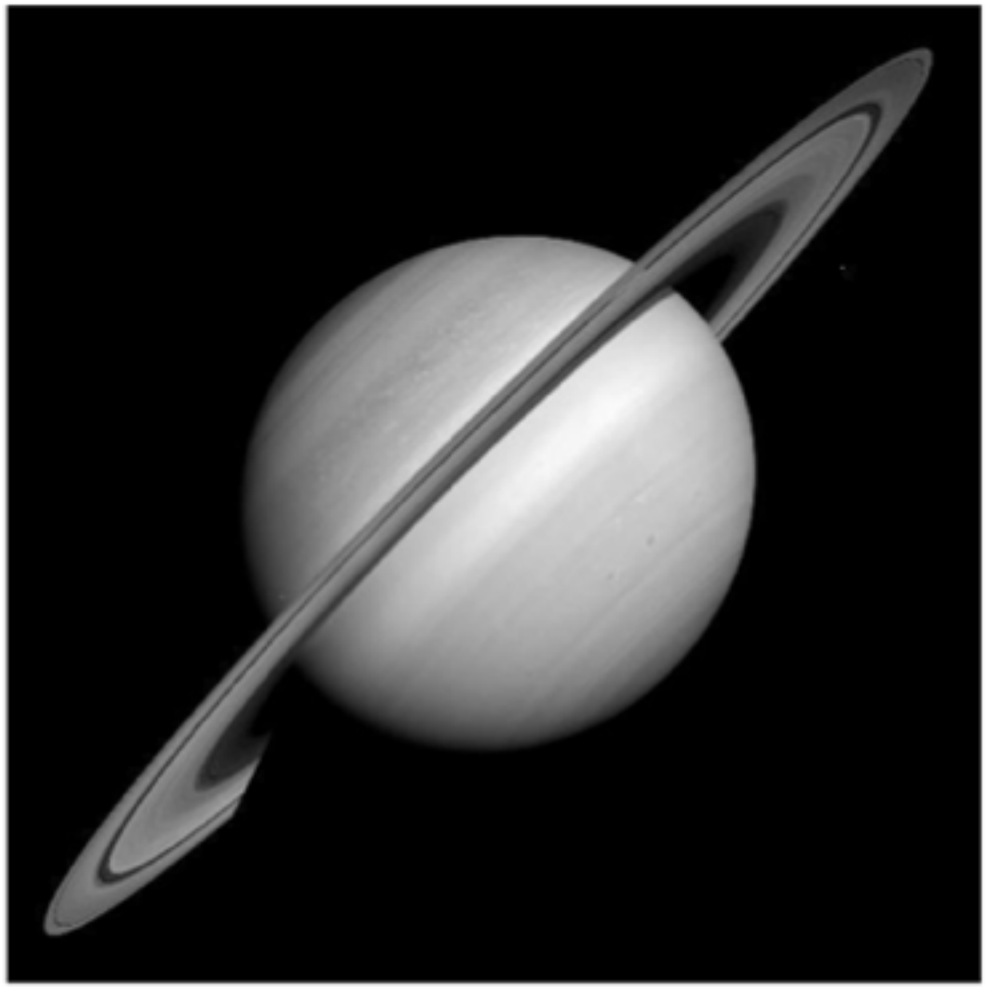
Spectrogram



WAVELETS



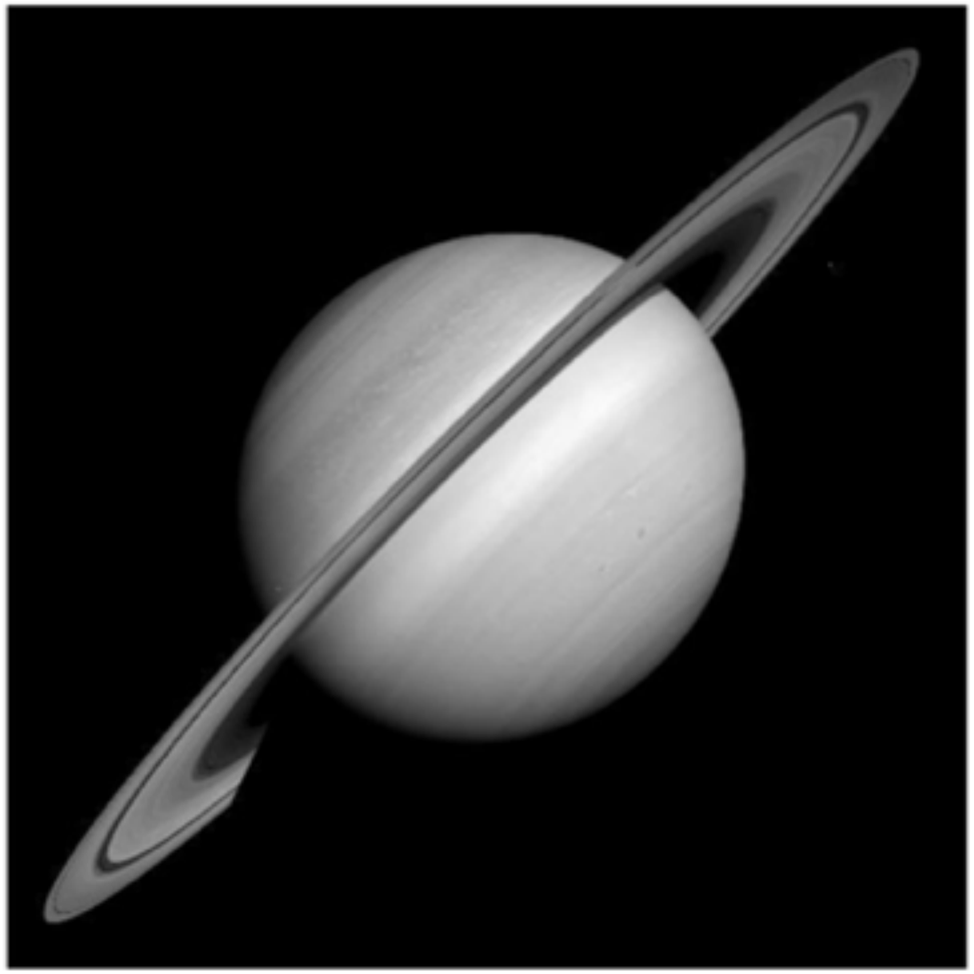
WAVELETS



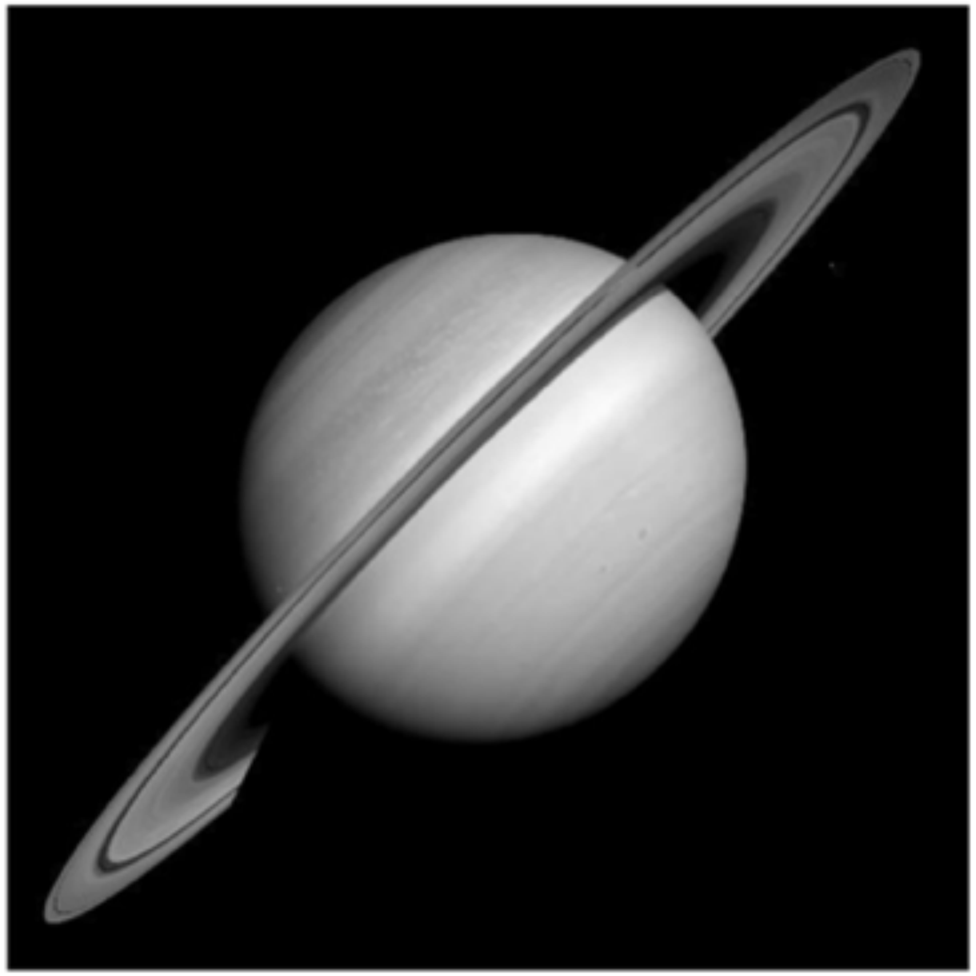
The top 1% of the coefficients concentrate only 8.66% of the energy. Not sparse...

1% largest coefficients in real space (the others are set to 0)

WAVELETS



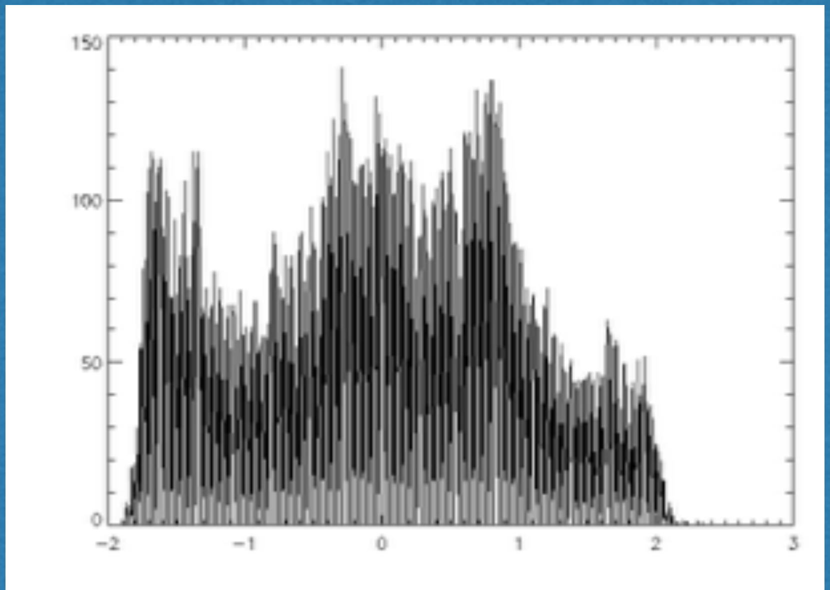
1% of the wavelet coefficients concentrate 99.96% of the energy: This can be used as a *prior.*



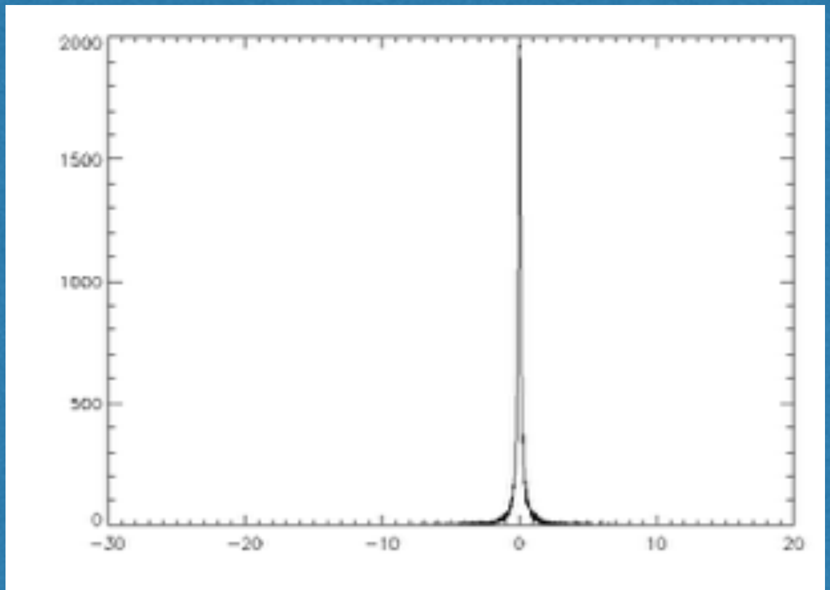
Reconstruction, after throwing away 99% of the wavelet coefficients

WAVELETS

Direct Space



Curvelet Space



OVERVIEW

Inverse Problems

Sparsity

Wavelets

Blind Source Separation

Blind Source Separation

Blind Source Separation

Definition

Blind source separation (BSS) is the separation of a set of source signals from a set of mixed signals, with little to no information about the source signals or the mixing process.

This technique is used in various fields such as astrophysics, biology, biomedicine, chemistry, geology, physics, remote sensing, …

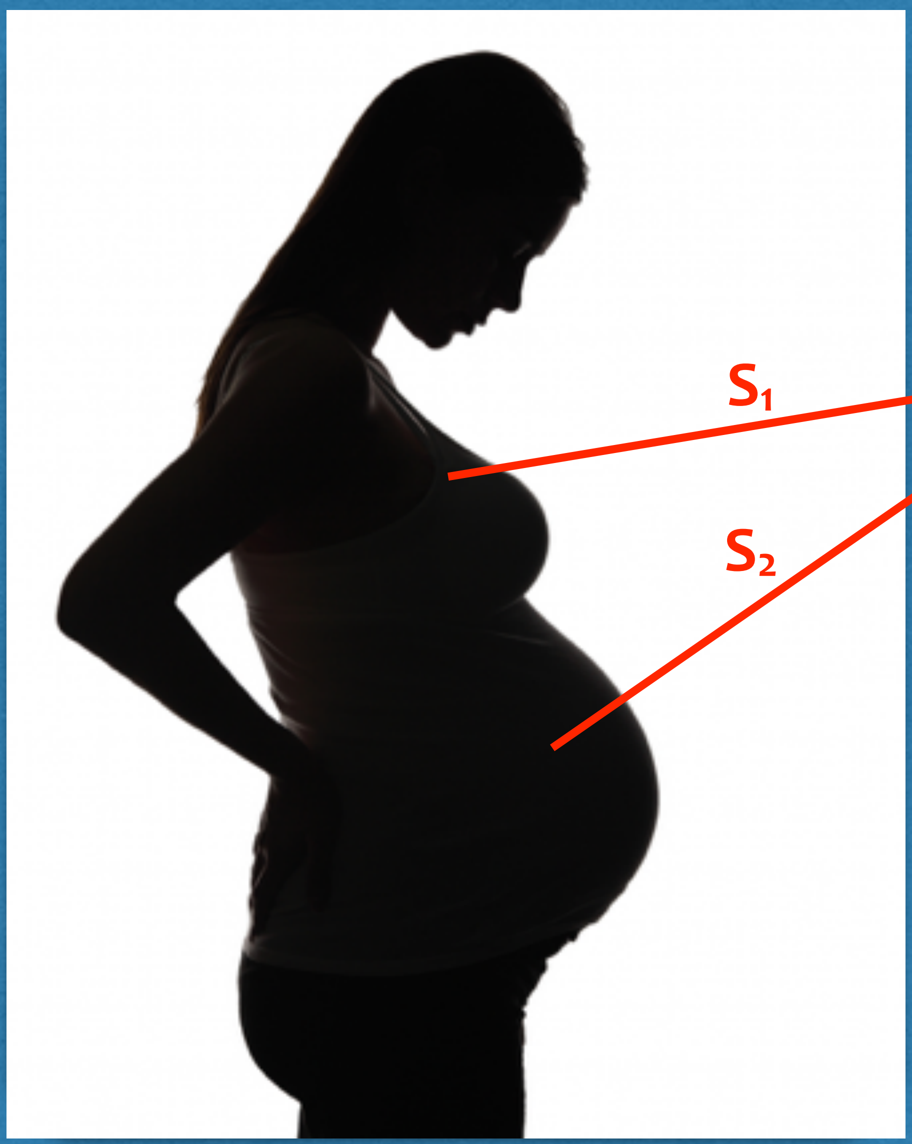
Blind Source Separation

Electrocardiography (ECG)

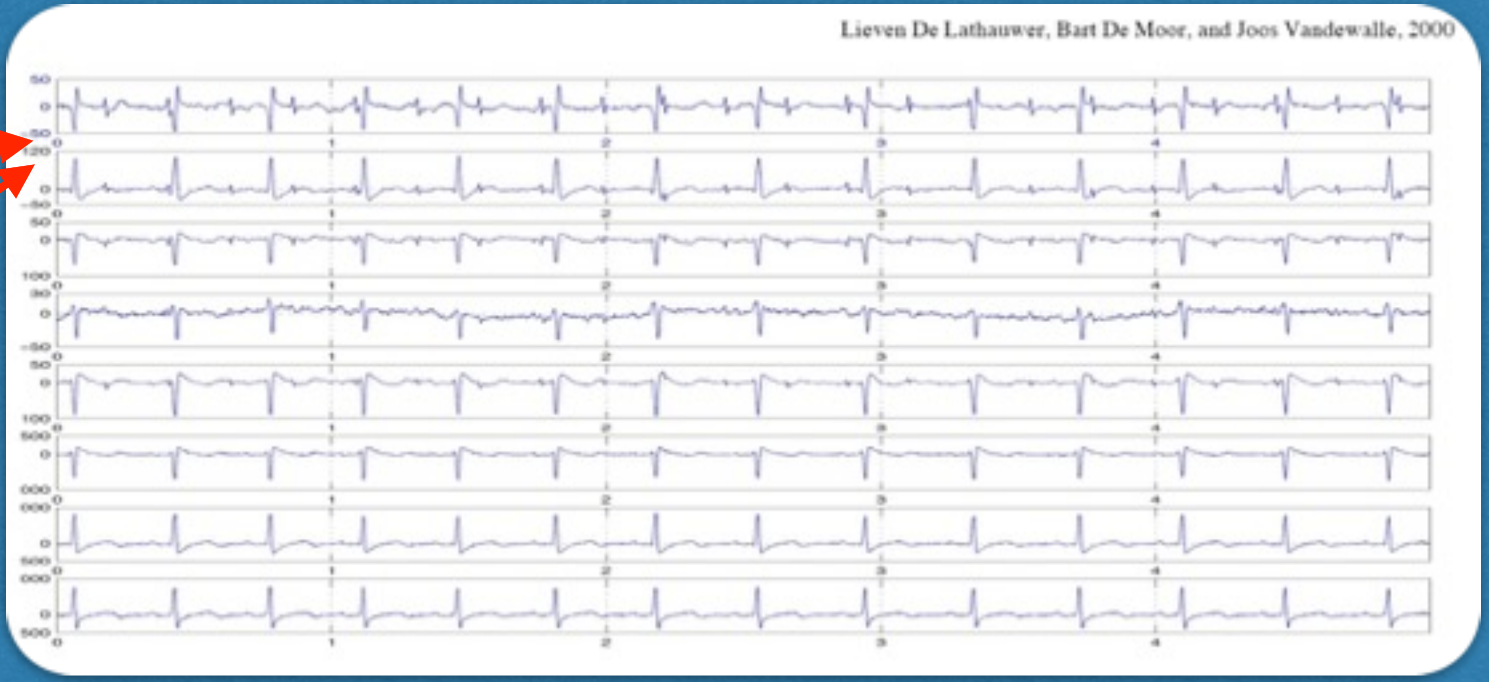
Extract the pulse of a foetus

BLIND SOURCE SEPARATION

Electrocardiography (ECG)

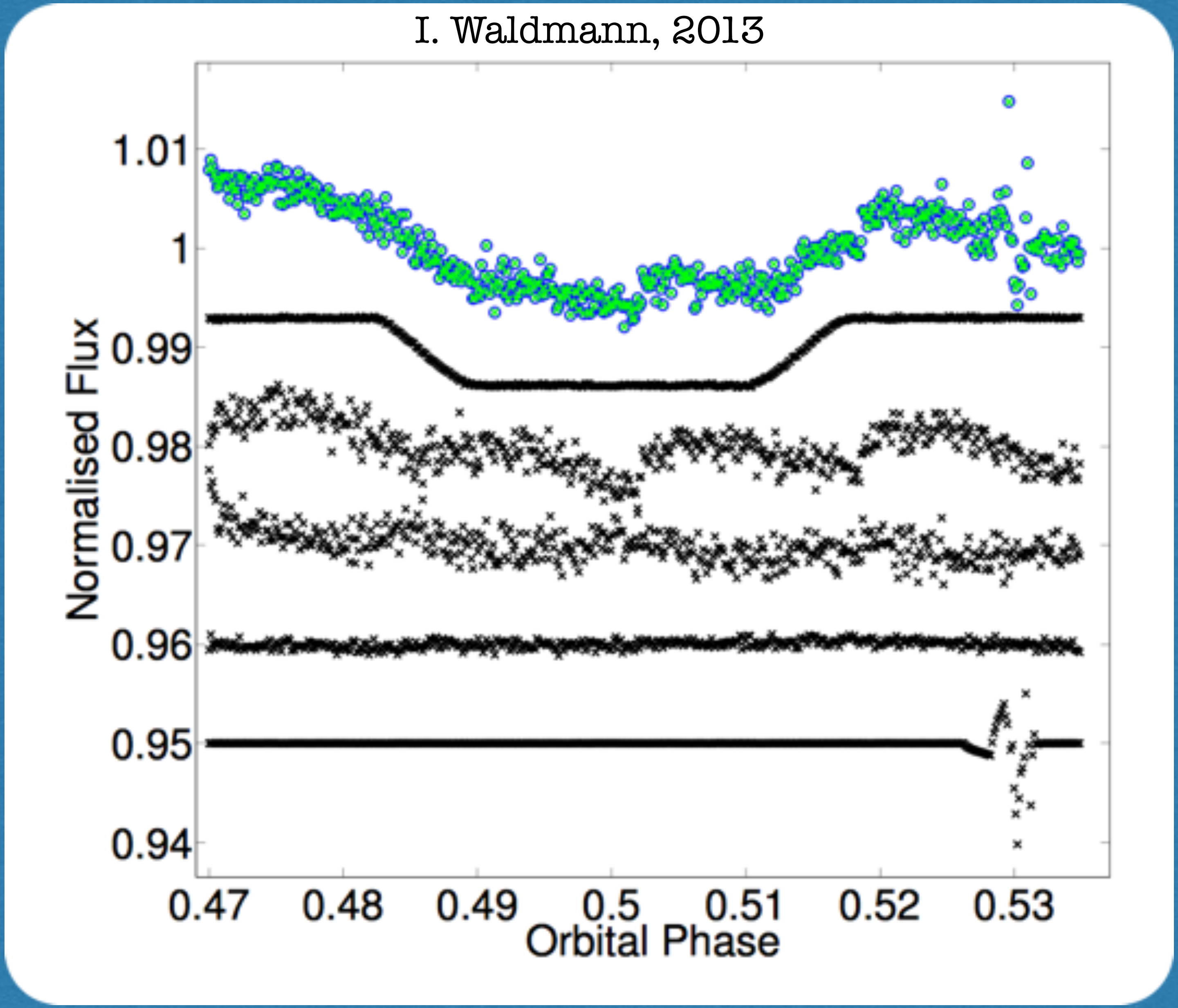


Extract the pulse of a foetus



Blind Source Separation

Exoplanet Analysis



Extraction/de-trending of transit time series

Blind Source Separation

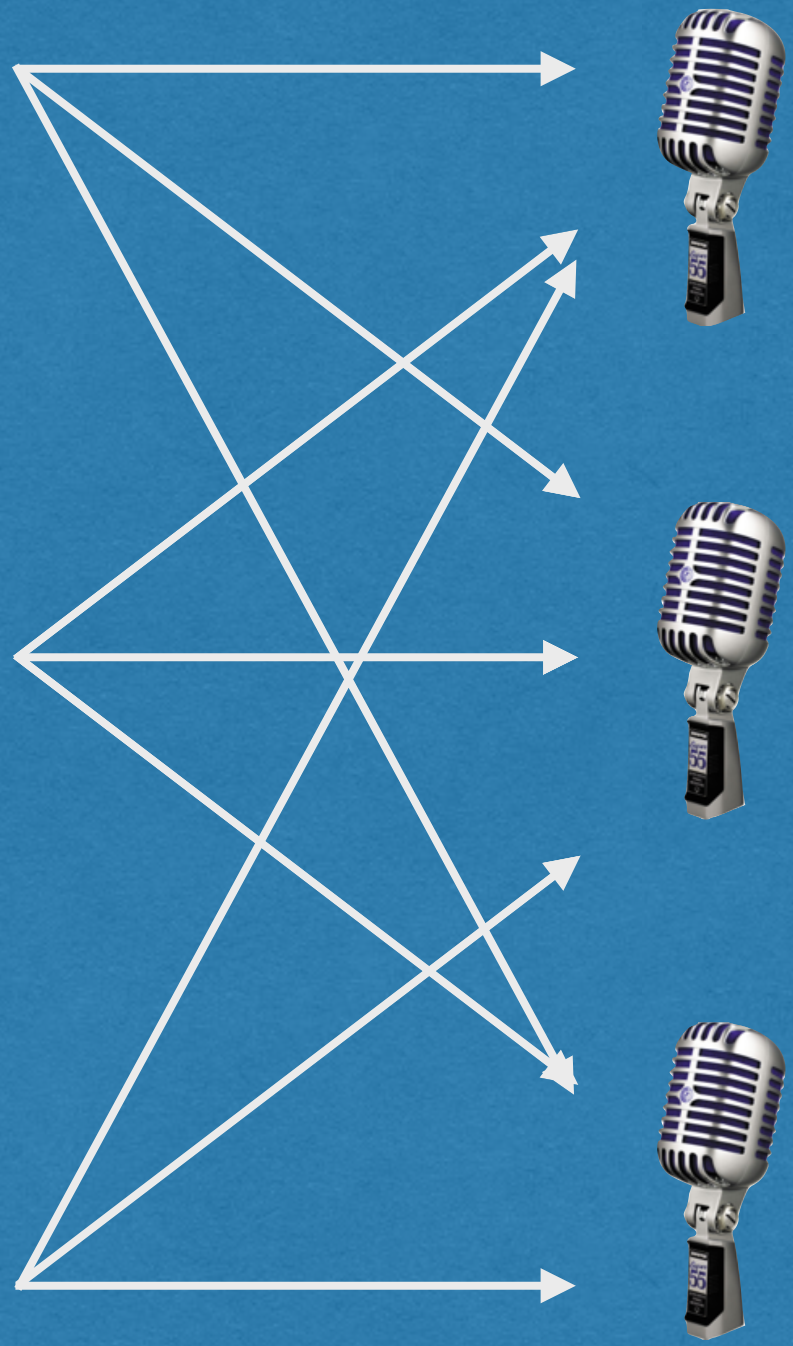
Cocktail Party Problem

Blind Source Separation

Separation is essentially impossible

$$x[t] = a_1 s_1[t] + a_2 s_2[t] + a_3 s_3[t]$$

Blind Source Separation



Now it may be possible

$$x_1[t] = a_{11}s_1[t] + a_{12}s_2[t] + a_{13}s_3[t]$$

$$x_2[t] = a_{21}s_1[t] + a_{22}s_2[t] + a_{23}s_3[t]$$

$$x_3[t] = a_{31}s_1[t] + a_{32}s_2[t] + a_{33}s_3[t]$$

Different linear combinations

Blind Source Separation

Common Theme

- Multivalued data provide simultaneous observations of the same physical phenomena.
- Each observation is composed of the combination of the same elementary components/sources/etc …

Blind Source Separation

The Model

assumed to be known

sources

$$\mathbf{X} = \sum_{k=1}^{n} a^k s_k = \mathbf{AS}$$

mixing matrix

Number of observations > number of sources

Blind Source Separation

The Model

assumed to be known

sources

$$\mathbf{X} = \sum_{k=1}^{n} a^k s_k = \mathbf{A}\boxed{\mathbf{S}}$$

mixing matrix

Number of observations > number of sources

Blind Source Separation

The Model

assumed to be known

sources

$$\mathbf{X} = \sum_{k=1}^{n} a^k s_k = \boxed{\mathbf{AS}}$$

mixing matrix

Number of observations > number of sources

Blind Source Separation

Ill-Posed Problem

There exists an infinite number of solutions such that :

$$X = AS = A'S'$$

Specifically for any invertible matrix U :

$$A' = AU \;\; \& \;\; S' = U^{-1}S$$

So, how do we solve this kind of problem?

Blind Source Separation

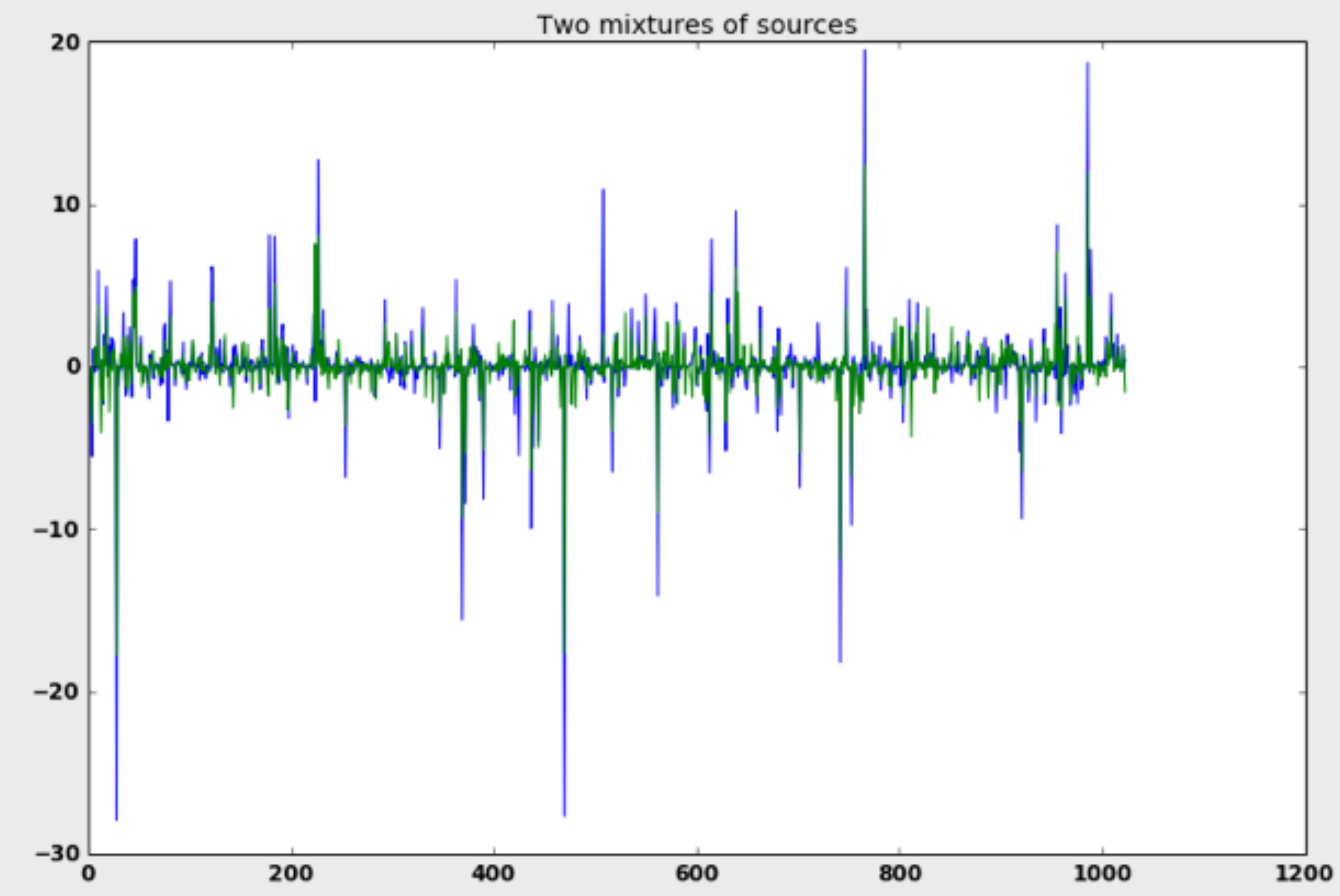
Source Properties

To solve these problems we need to exploit additional information about the sources such as:

- **Decorrelation**
- **Independence**
- **Sparsity**

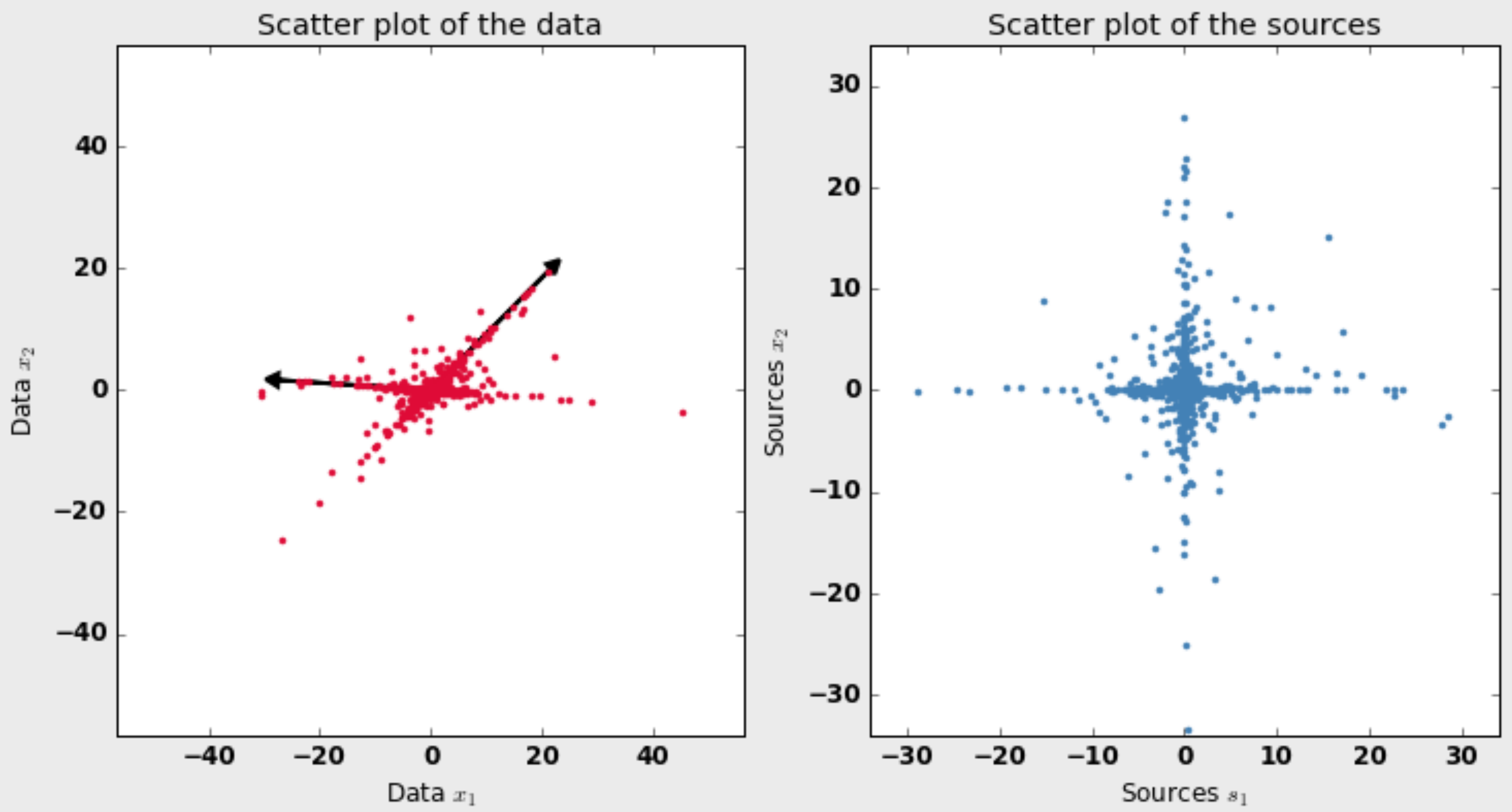
BLIND SOURCE SEPARATION

Source Properties



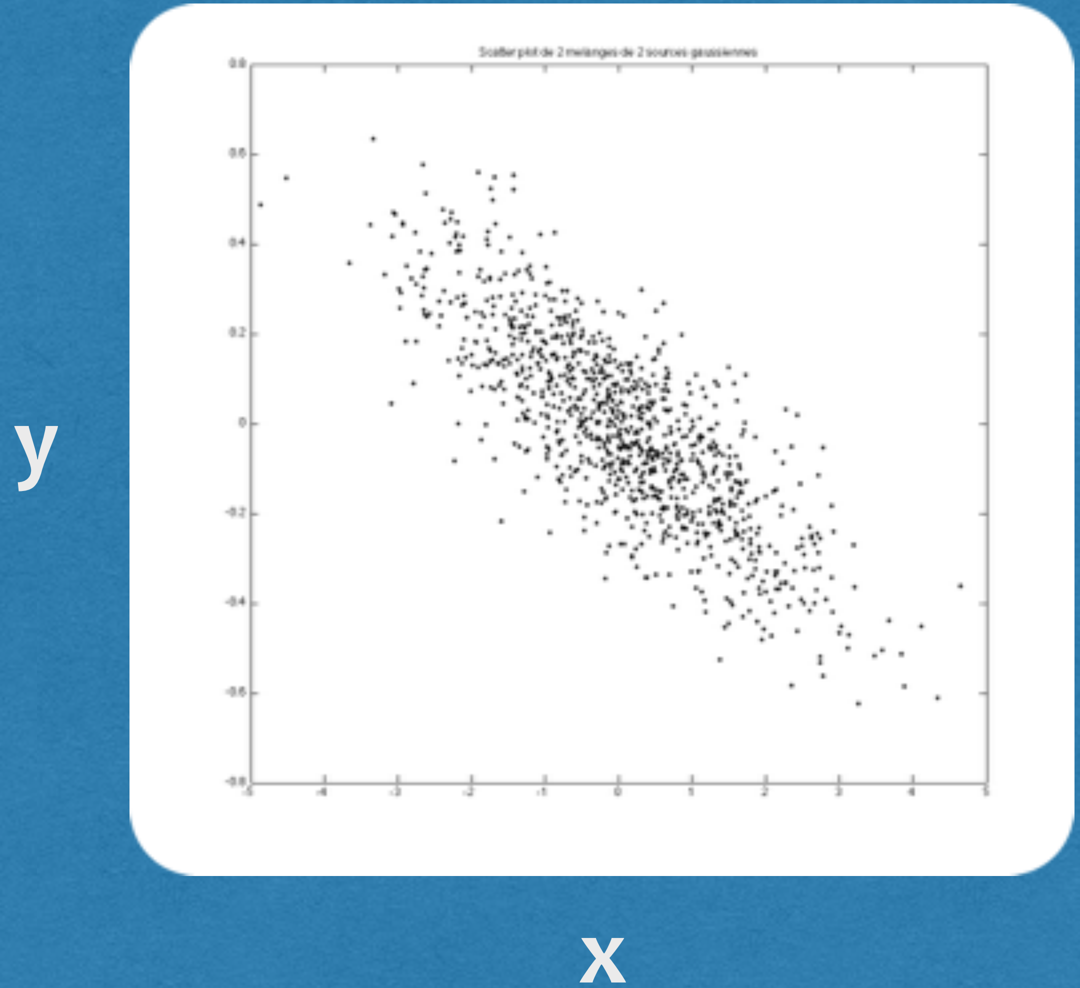
BLIND SOURCE SEPARATION

Source Properties



Blind Source Separation

Principal Component Analysis (PCA)

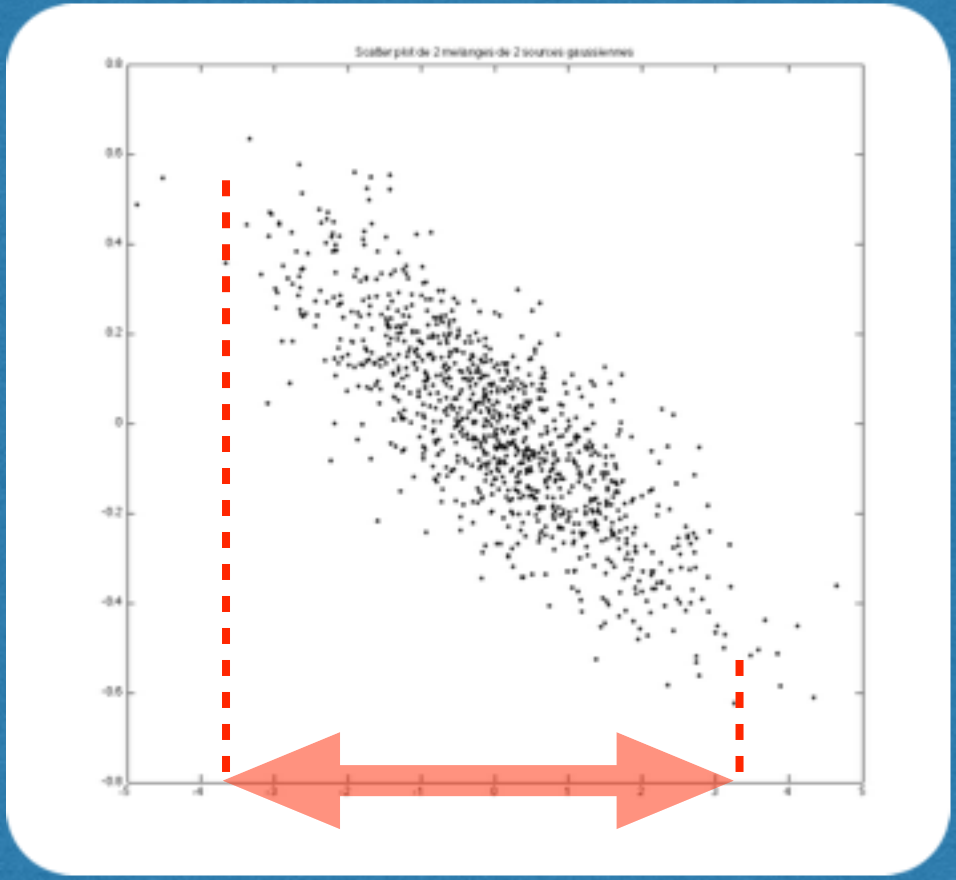


Find direction of maximal variance (2nd order statistic) in the data that gives the best reconstruction error.

Blind Source Separation

Principal Component Analysis (PCA)

y

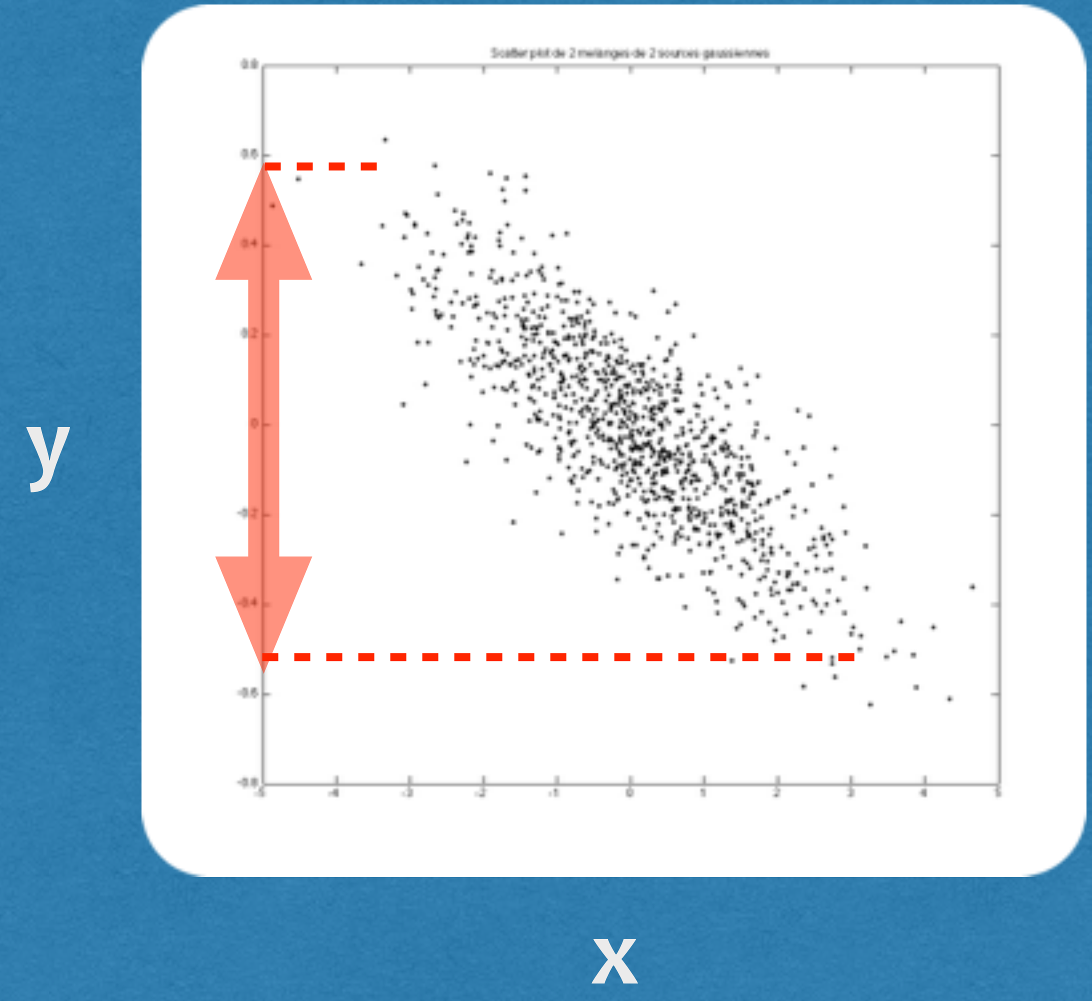


x

Find direction of maximal variance (2nd order statistic) in the data that gives the best reconstruction error.

Blind Source Separation

Principal Component Analysis (PCA)

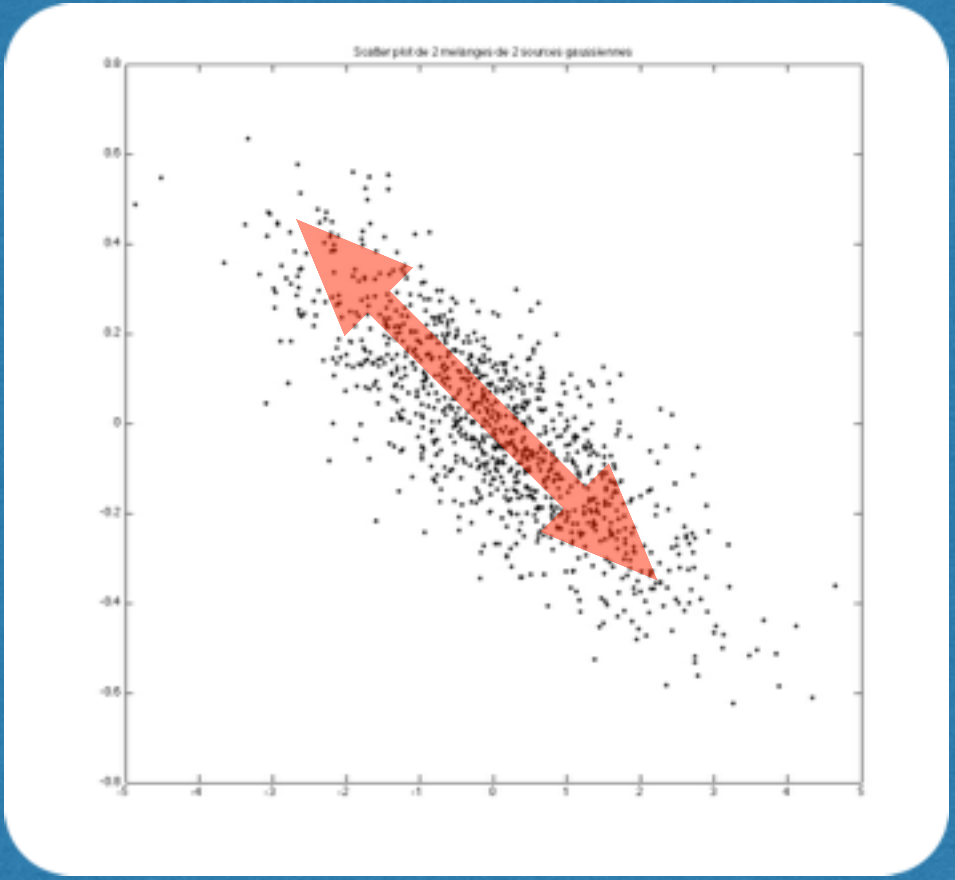


Find direction of maximal variance (2nd order statistic) in the data that gives the best reconstruction error.

BLIND SOURCE SEPARATION

Principal Component Analysis (PCA)

y



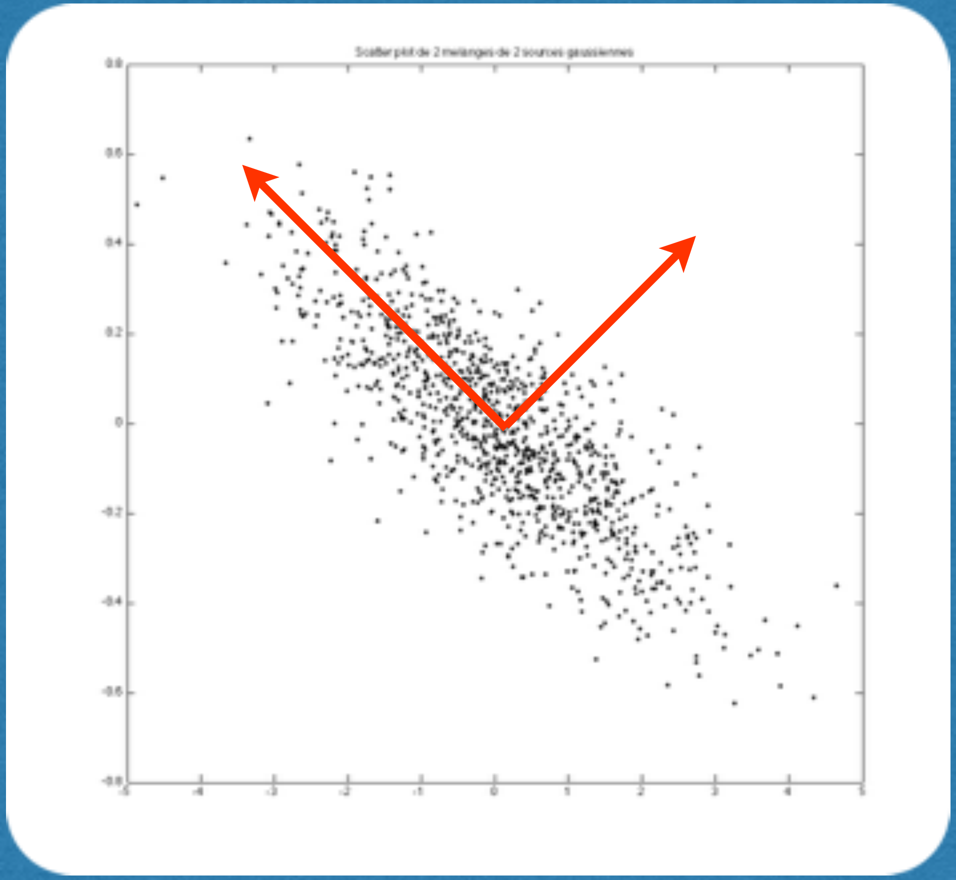
x

Find direction of maximal variance (2^{nd} order statistic) in the data that gives the best reconstruction error.

Blind Source Separation

Principal Component Analysis (PCA)

y



x

Find direction of maximal variance (2^{nd} order statistic) in the data that gives the best reconstruction error.

Blind Source Separation

Decorrelation

Decorrelation can be measured via the **covariance matrix** of the estimated sources : it should be diagonal .

$$R_X = XX^T$$ **is the covariance matrix of the data**

Don't forget that : $$X = AS$$

$$R_X = ASS^T A^T = A \underbrace{R_S}_{\text{Diagonal}} A^T$$

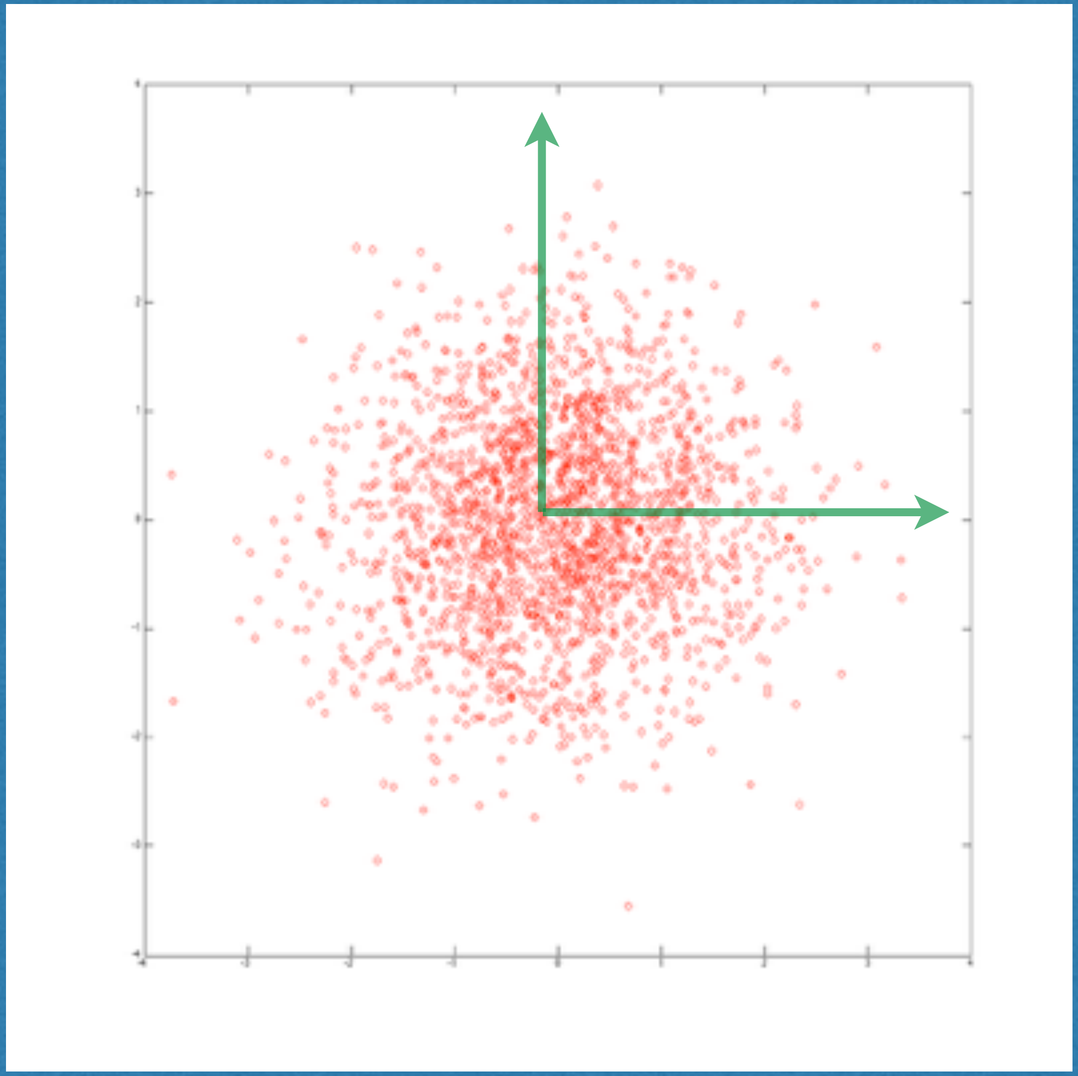
Blind Source Separation

Decorrelation

Rotational Invariance

Second-order statistics provide a "Gaussian" approximation of the sources

S_2



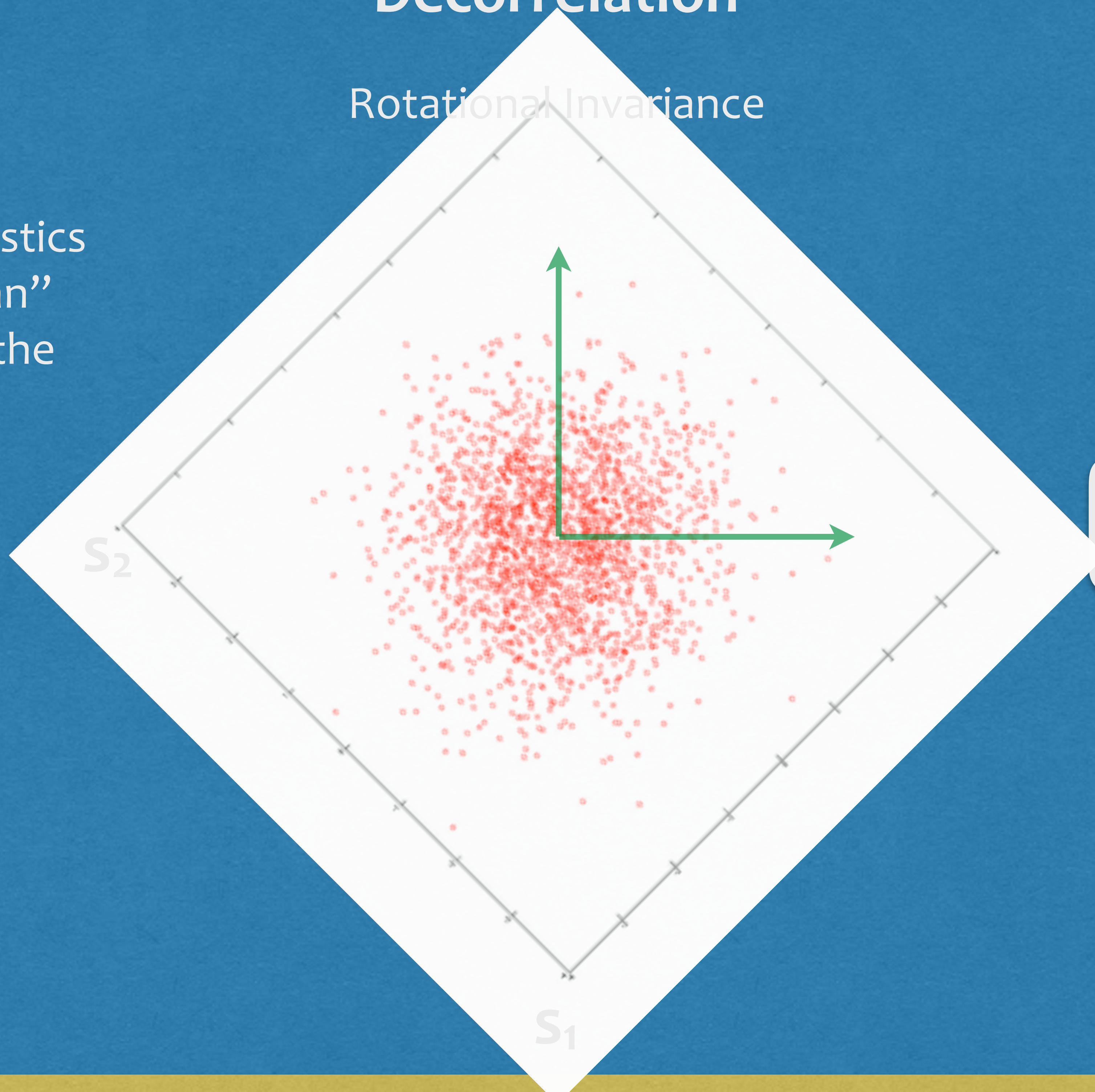
S_1

$$R_{S'} = R_S$$

Blind Source Separation

Decorrelation

Second-order statistics provide a "Gaussian" approximation of the sources



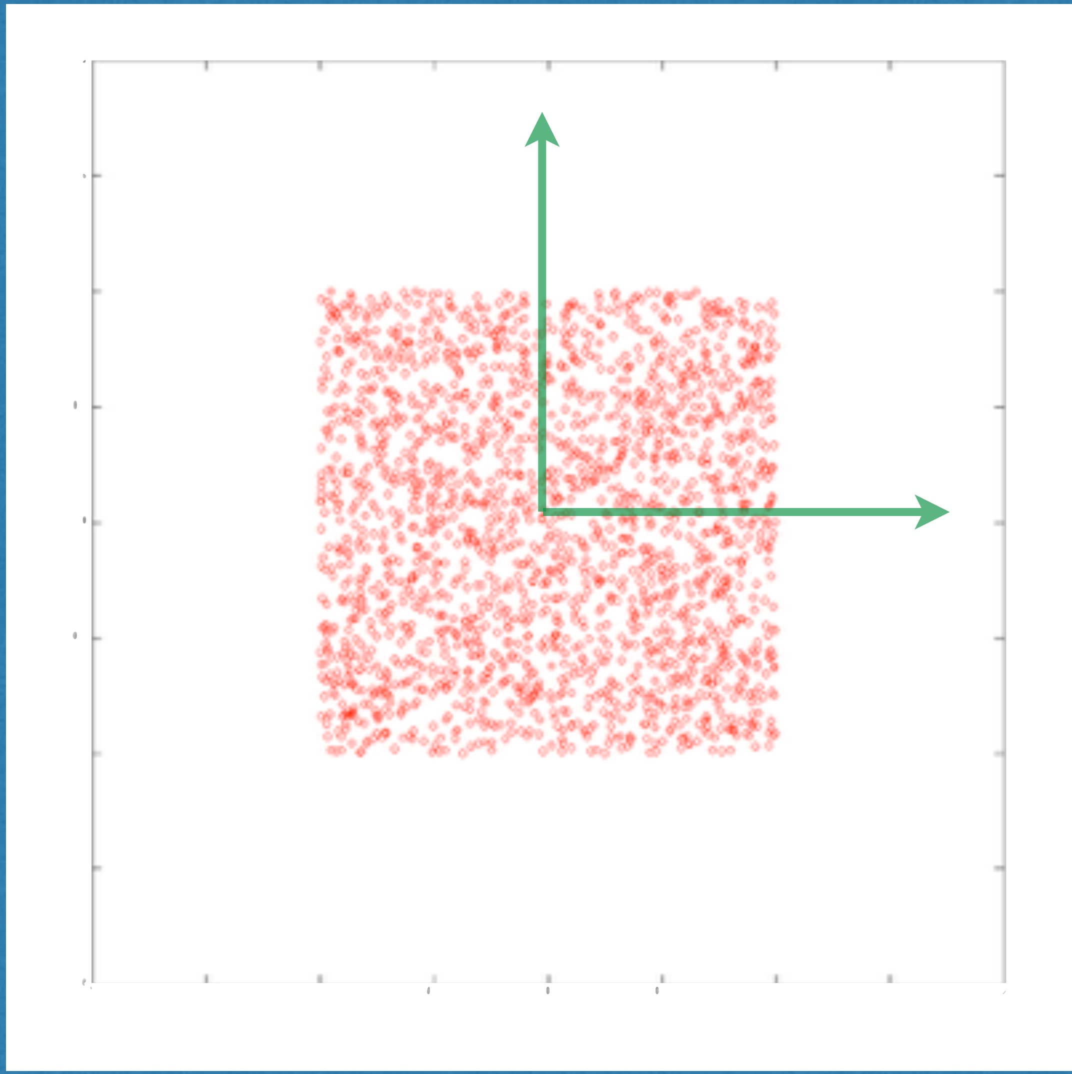
$$R_{S'} = R_S$$

Blind Source Separation

Independence

Rotational Variance

Uniform (non-Gaussian) sources



S_2

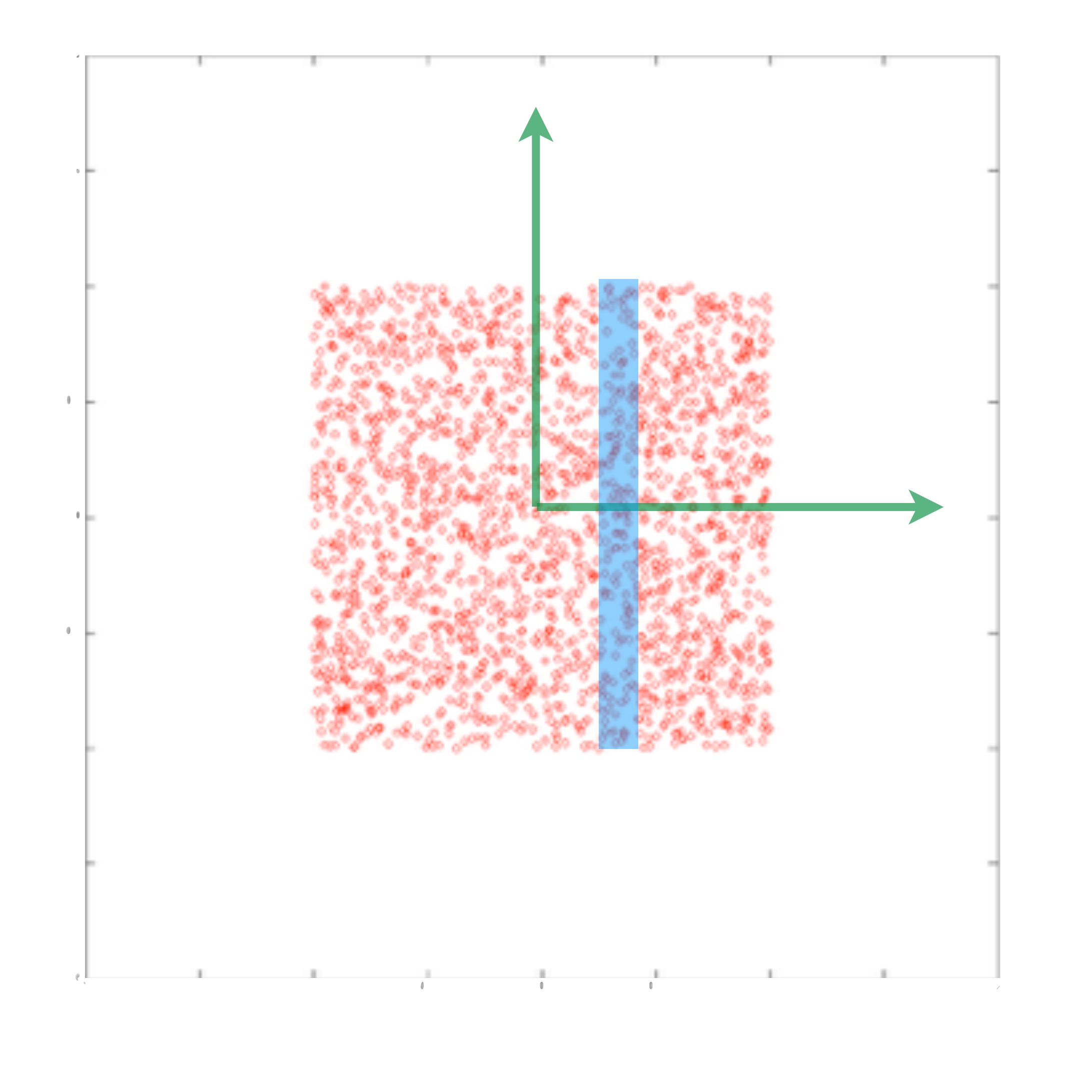
S_1

BLIND SOURCE SEPARATION

Independence

Rotational Variance

Uniform (non-Gaussian) sources



s_2

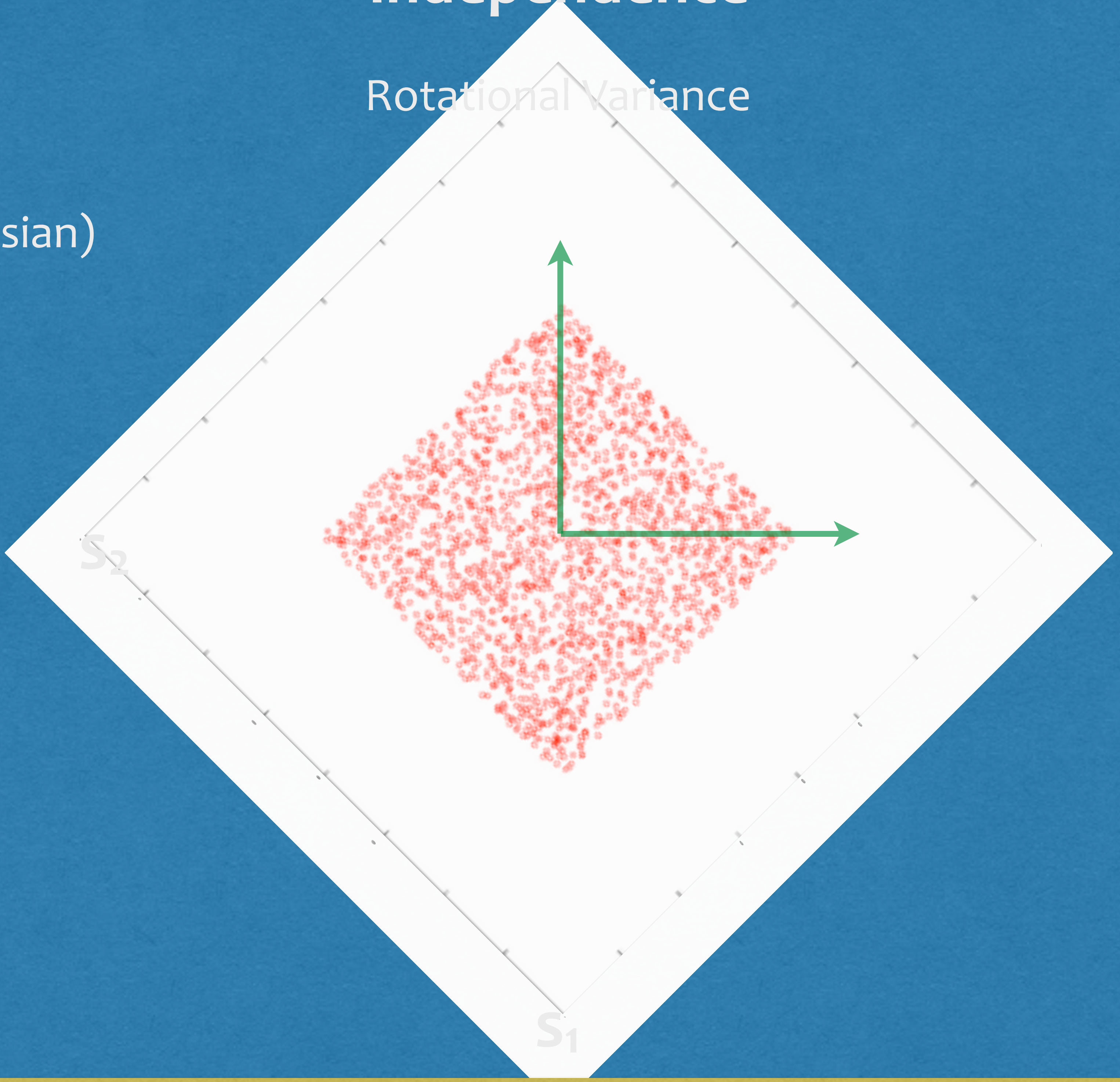
s_1

BLIND SOURCE SEPARATION

Independence

Rotational Variance

Uniform (non-Gaussian) sources



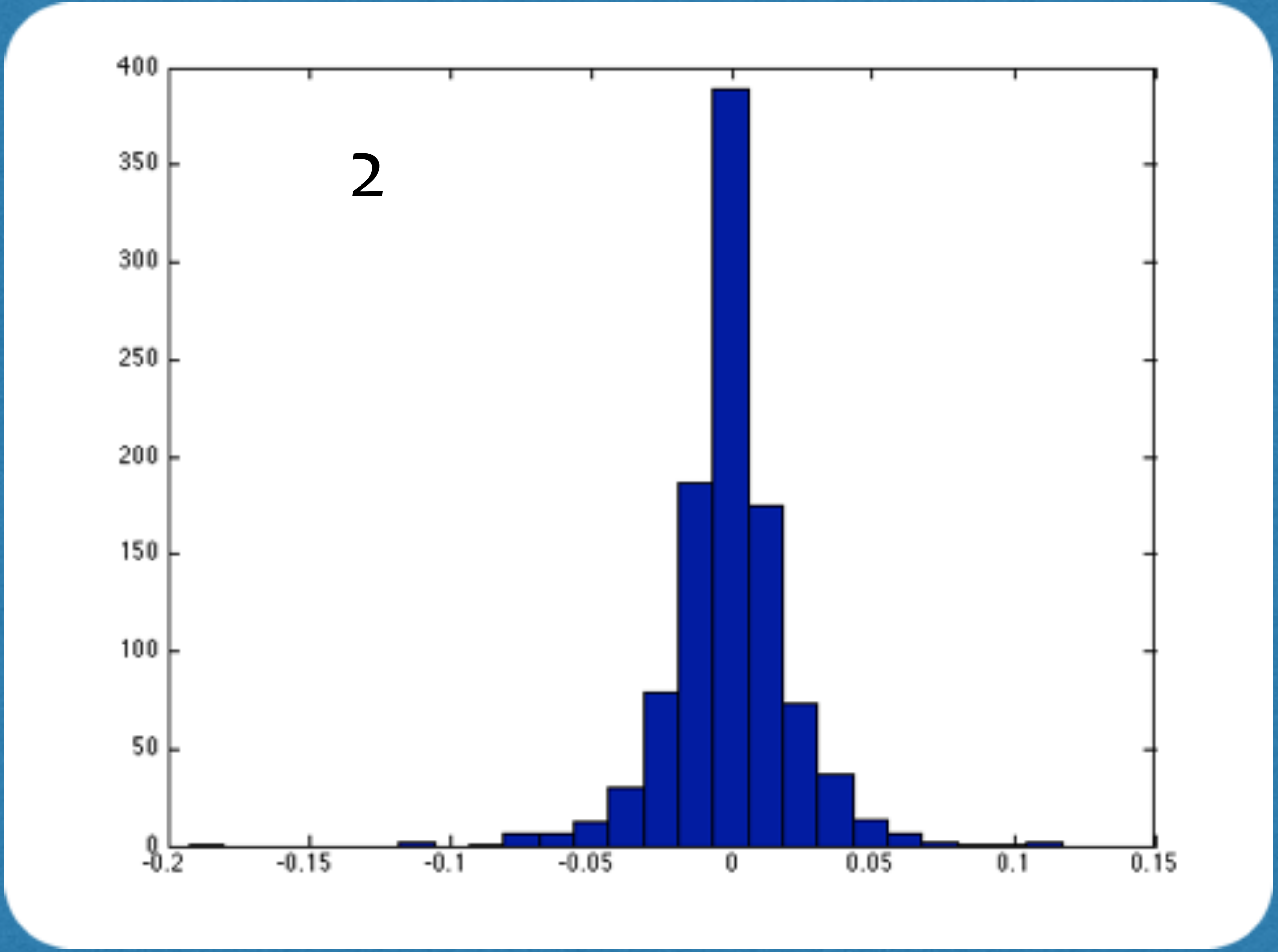
Blind Source Separation

Independence

$$\forall t; \quad \frac{1}{n} x_i[t] = \frac{1}{n} \sum_{j=1}^{n} s_j \longrightarrow x \sim \mathcal{N}\left(0, \sigma^2\right)$$

Central limit theorem

As more independent sources are mixed the signal becomes more Gaussian.



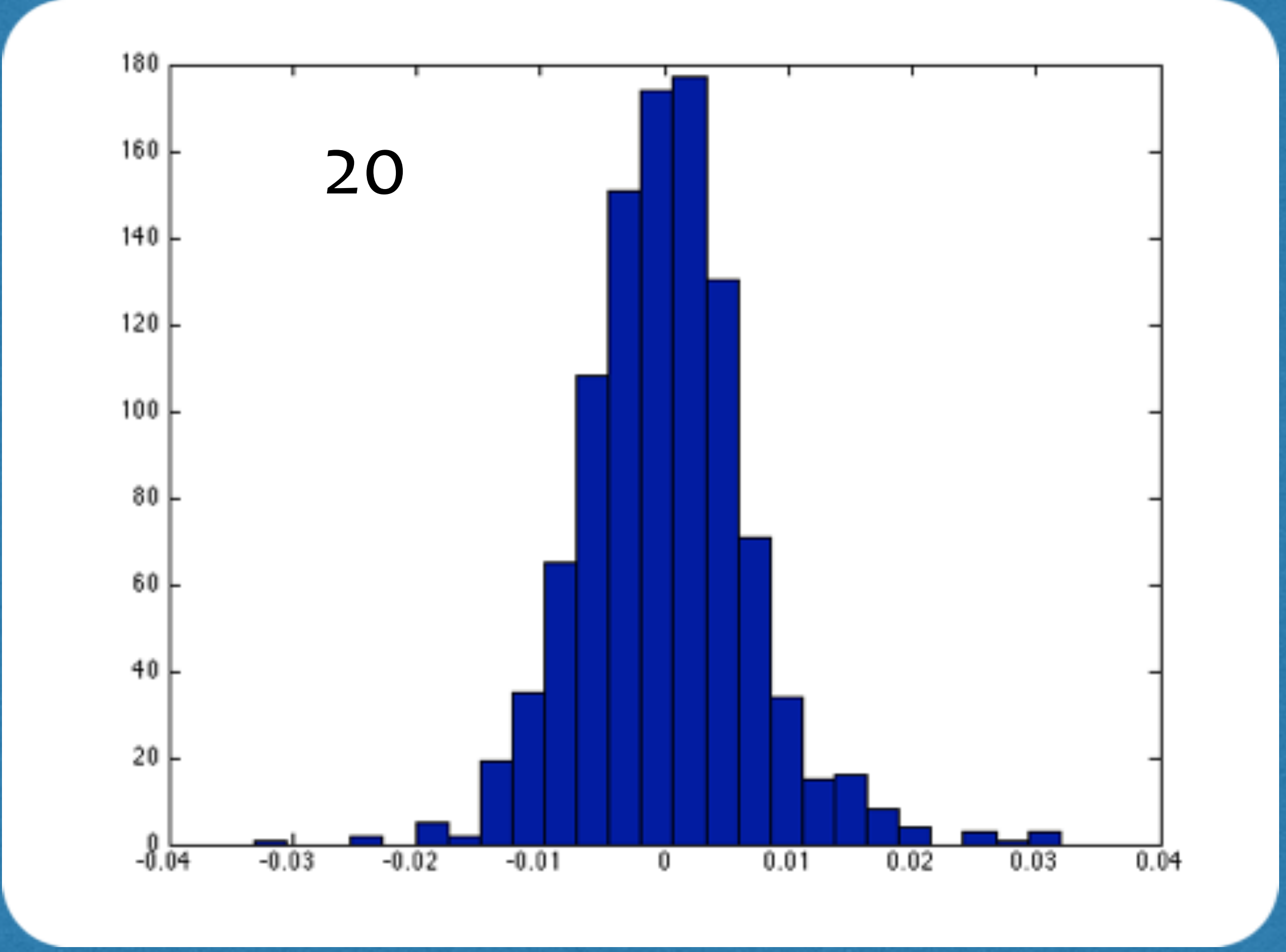
BLIND SOURCE SEPARATION

Independence

$$\forall t; \quad \frac{1}{n} x_i[t] = \frac{1}{n} \sum_{j=1}^{n} s_j \longrightarrow x \sim \mathcal{N}\left(0, \sigma^2\right)$$

Central limit theorem

As more independent sources are mixed the signal becomes more Gaussian.



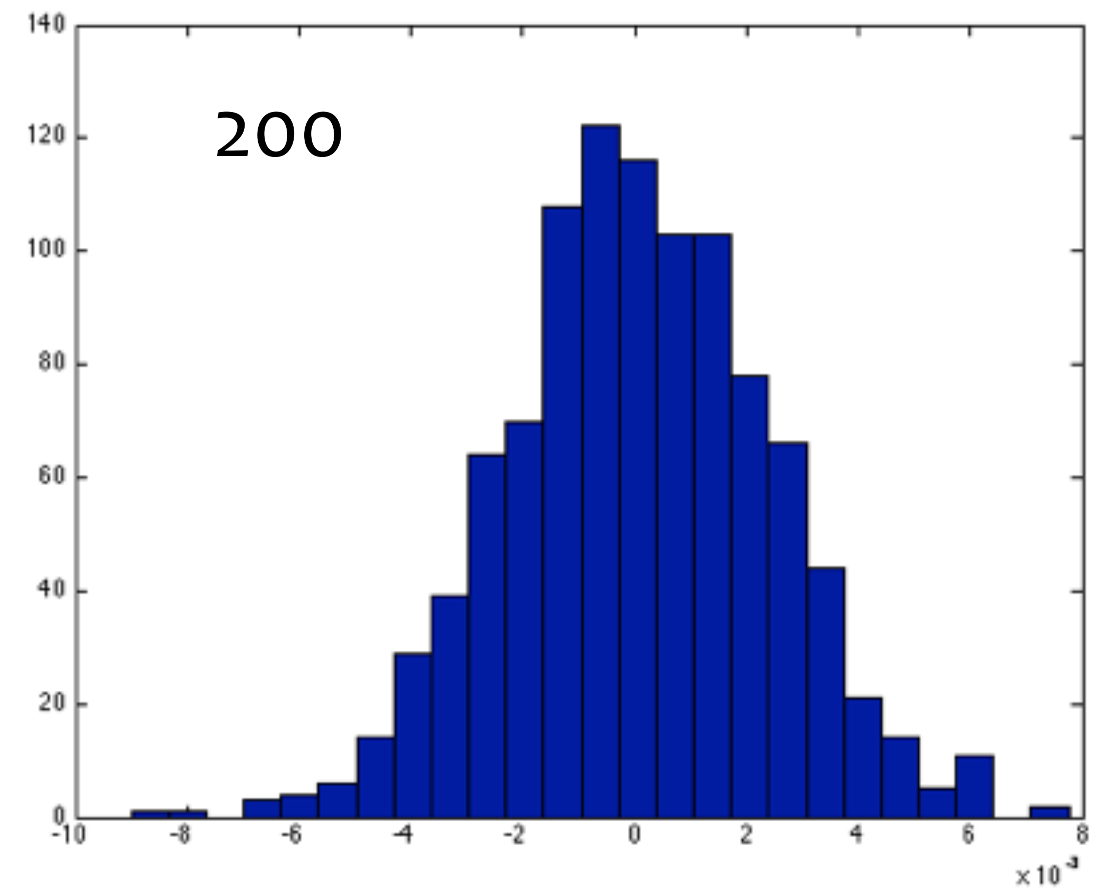
BLIND SOURCE SEPARATION

Independence

$$\forall t; \quad \frac{1}{n}x_i[t] = \frac{1}{n}\sum_{j=1}^{n} s_j \longrightarrow x \sim \mathcal{N}\left(0, \sigma^2\right)$$

Central limit theorem

As more independent sources are mixed the signal becomes more Gaussian.



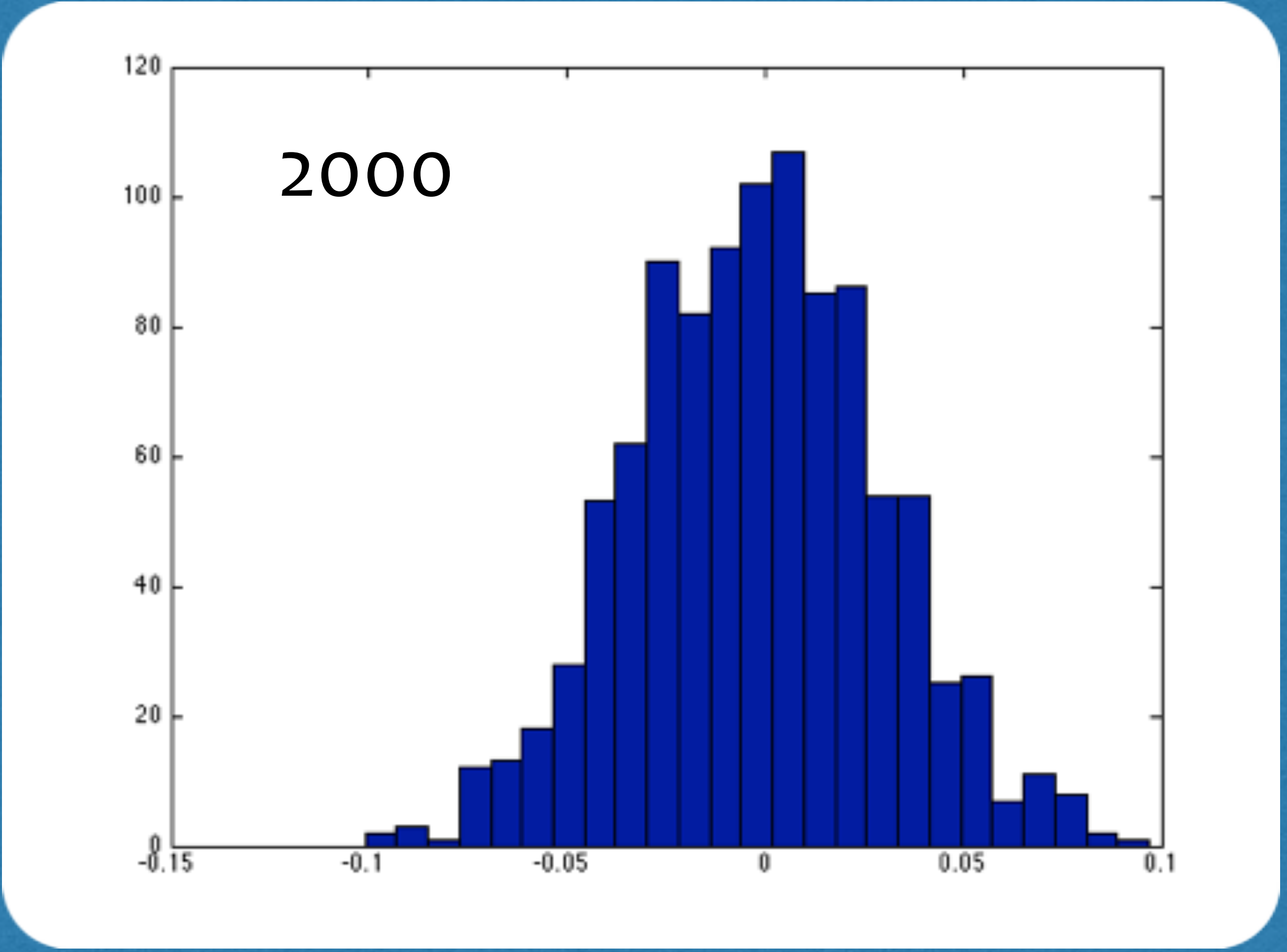
BLIND SOURCE SEPARATION

Independence

$$\forall t; \quad \frac{1}{n} x_i[t] = \frac{1}{n} \sum_{j=1}^{n} s_j \longrightarrow x \sim \mathcal{N}\left(0, \sigma^2\right)$$

Central limit theorem

As more independent sources are mixed the signal becomes more Gaussian.



BLIND SOURCE SEPARATION

Independent Component Analysis (ICA)

The separation of sources makes the signal less Gaussian. Therefore, by maximising the non-Gaussianity one can attempt to separate the sources.

We can measure the non-Gaussianity using higher order statistics (*i.e.* **skewness** and **kurtosis**)

Project the data onto a new feature space

$$X \to X'$$

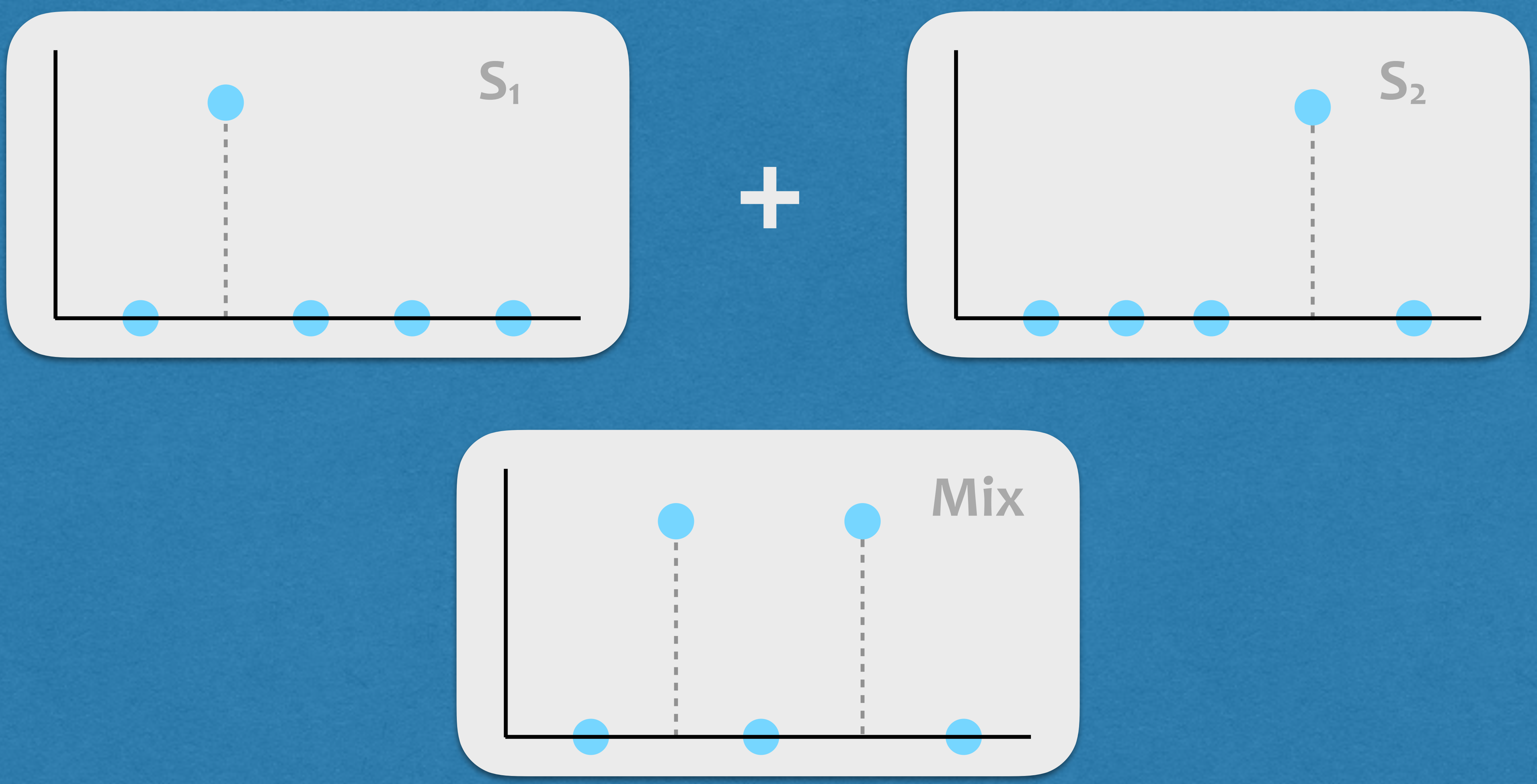
such that

$$x'_i \perp x'_j$$

BLIND SOURCE SEPARATION

Sparsity

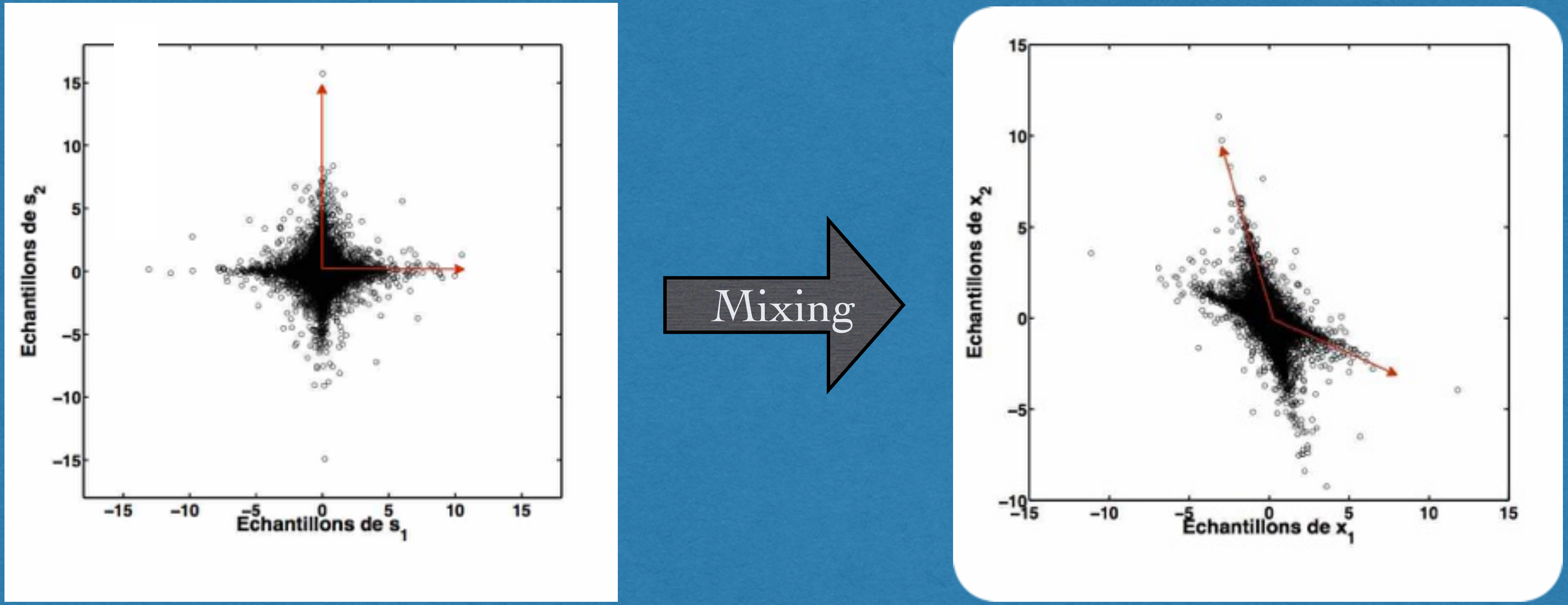
Mixtures are less sparse than sources!



BLIND SOURCE SEPARATION

Sparsity

Mixtures are less sparse than sources!



Blind Source Separation

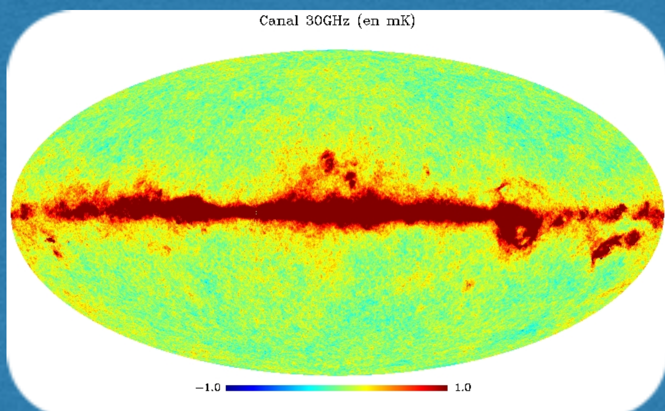
Sparsity

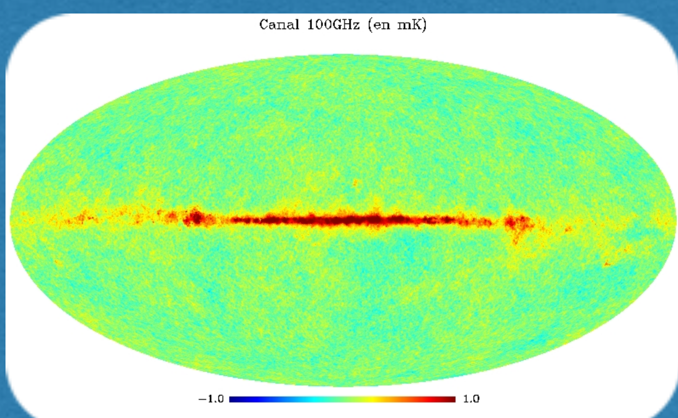
Since mixed sources are less sparse, we can estimate the sources by maximising the sparsity.

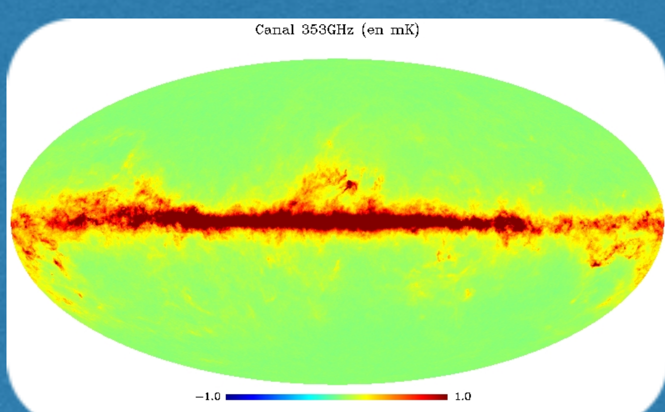
This method is less sensitive to noise than any other method which makes it more robust.

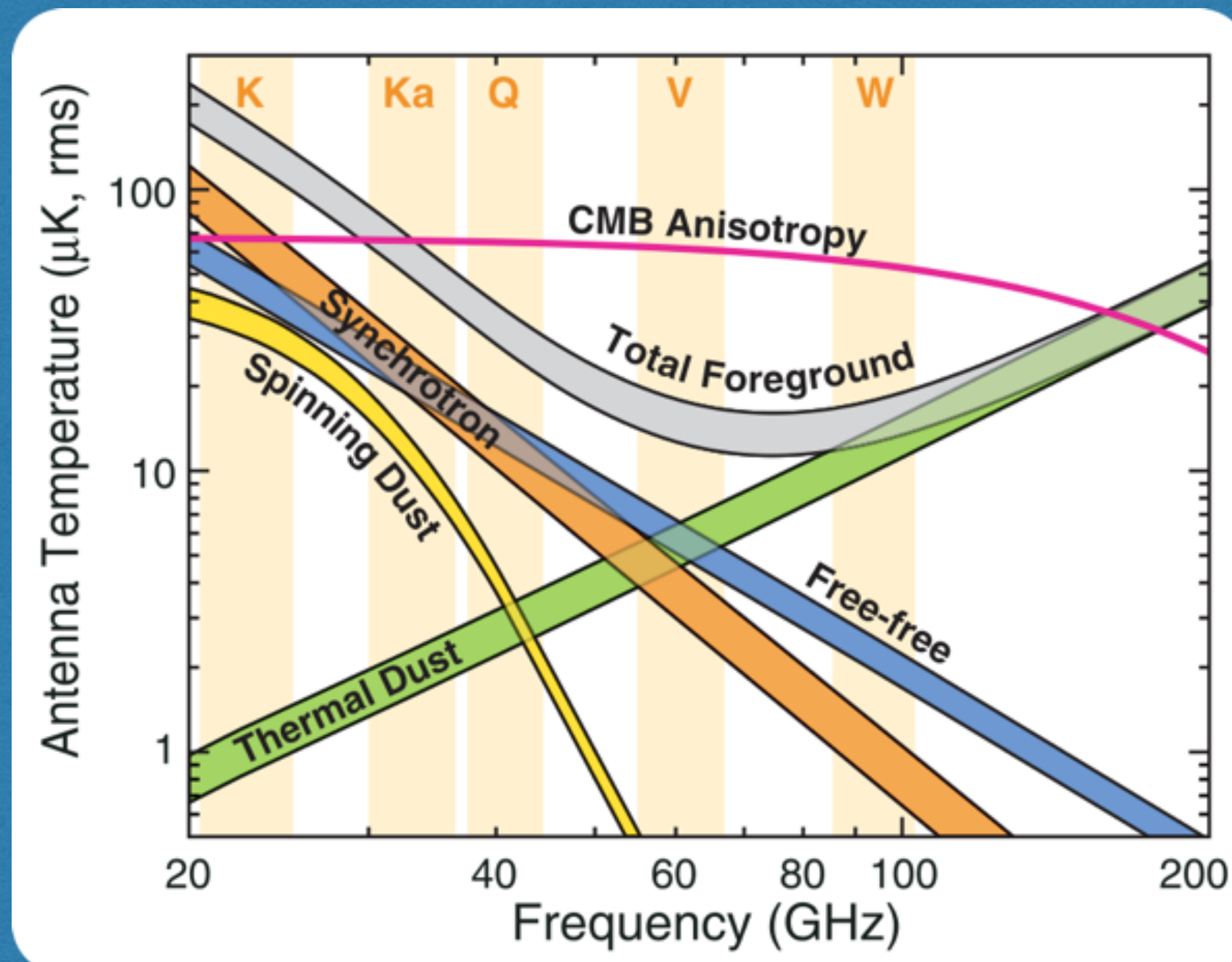
BLIND SOURCE SEPARATION

CMB Example









Blind Source Separation

CMB Example

Not a fully blind problem as we assume the CMB spectrum is know.

Can model the problem as follows

Observation

$$X = as + R$$

CMB spectrum

other components + noise

Other components have dominant contribution over CMB

$$||R|| \gg ||as||$$

Blind Source Separation

CMB Example

The industry standard

Internal Linear Combination (**ILC**)

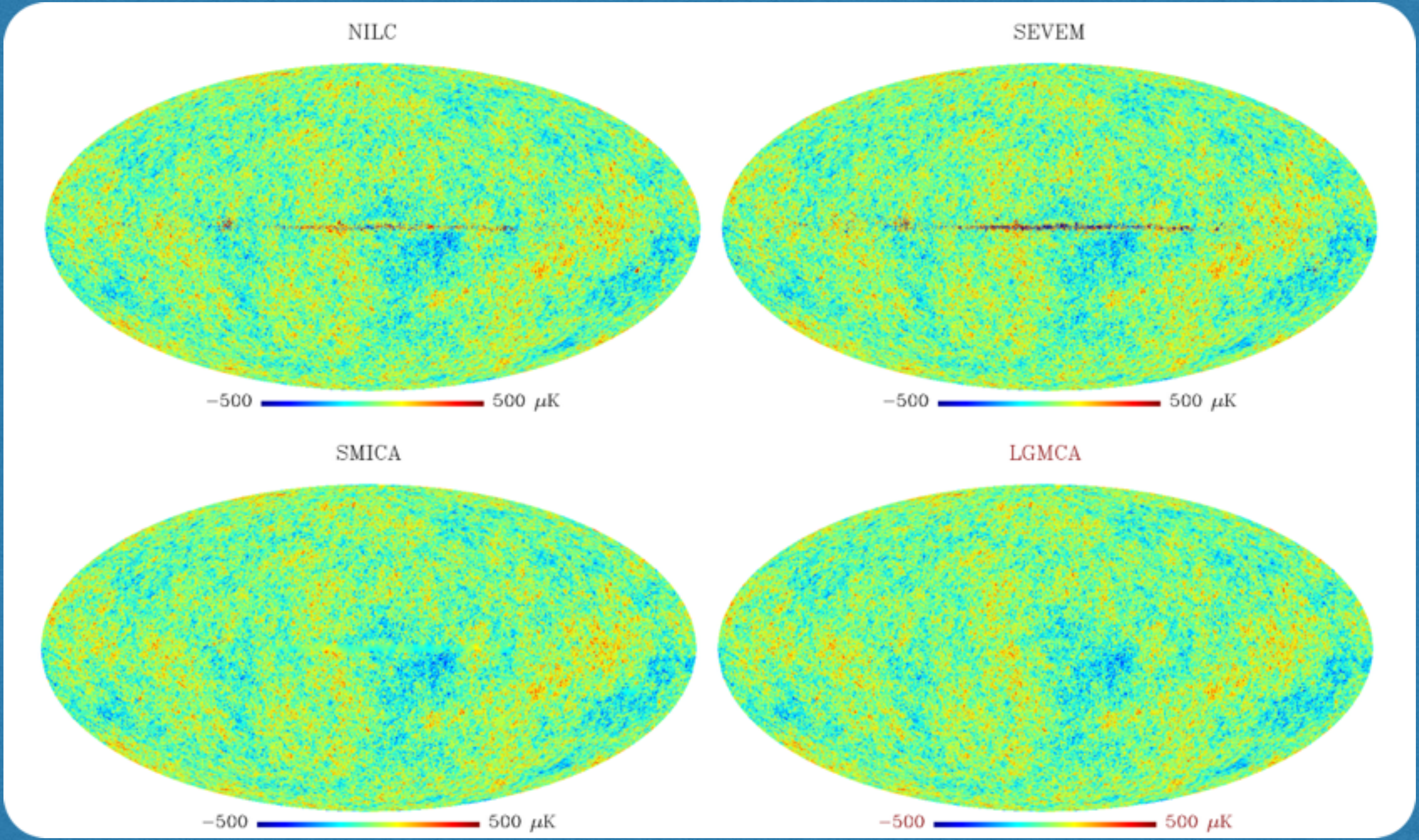
used officially on WMAP

$$\hat{s} = \operatorname{Argmin}_s \left(X - as\right) R_X^{-1} \left(X - as\right)^T$$

neglected

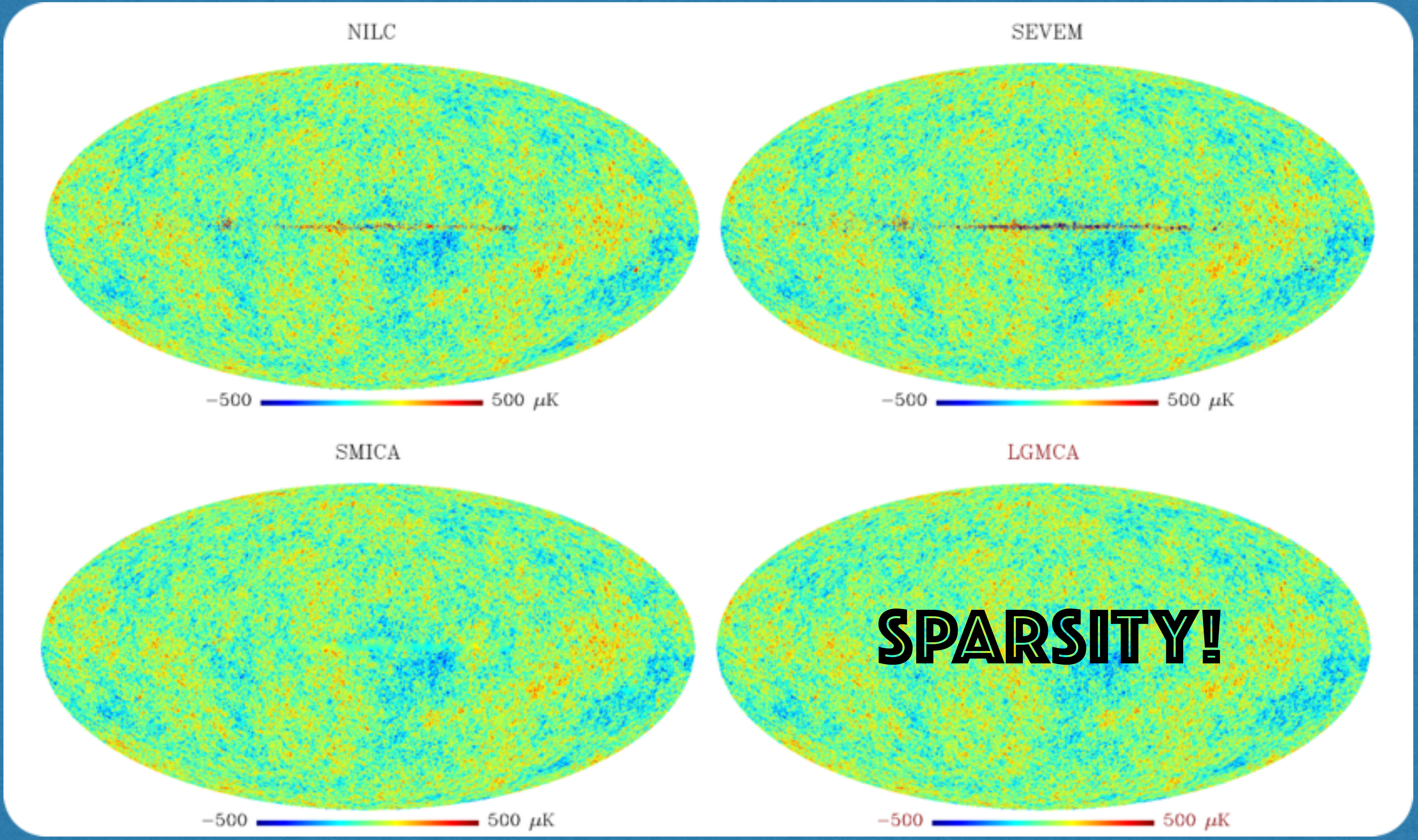
BLIND SOURCE SEPARATION

CMB Example



Blind Source Separation

CMB Example

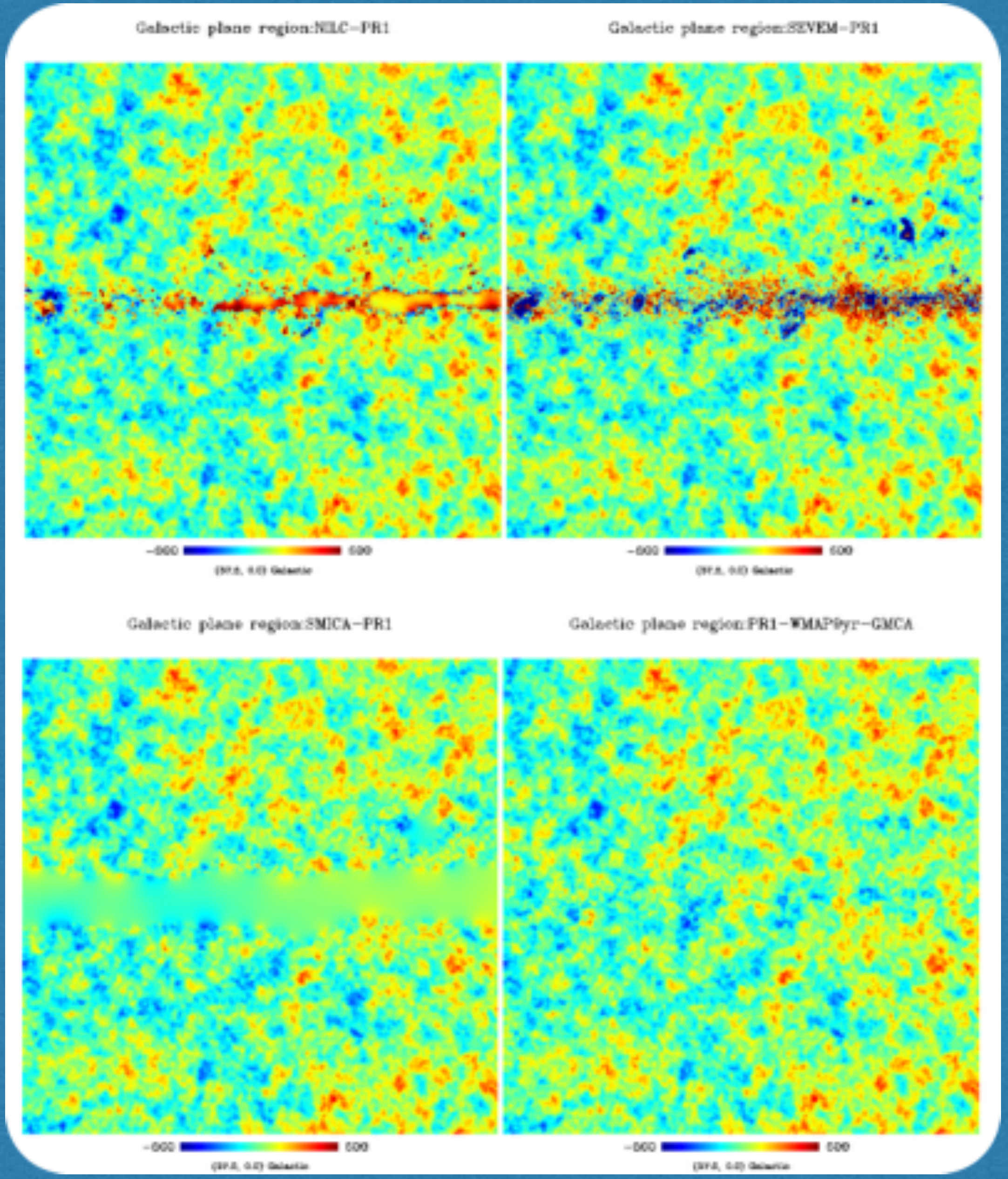


Blind Source Separation

CMB Example

Sparsity allows for better modelling of the foregrounds.

Other methods use second order statistics.



ΕΥΧΑΡΙΣΤΩ