#### Can wavelets help computers listen and focus their attention ?

# An introduction to source separation with sparse decompositions

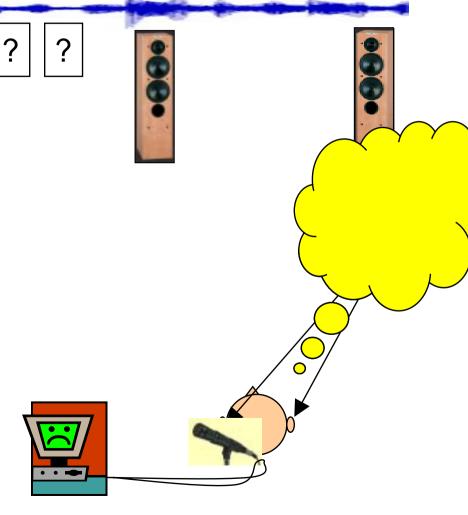
#### Rémi Gribonval (INRIA) Projet METISS, IRISA

Simon Arberet Sacha Krstulovic Sylvain Lesage Alexey Ozerov

Morten Nielsen (Univ. Aalborg) Pierre Vandergheynst (EPFL)

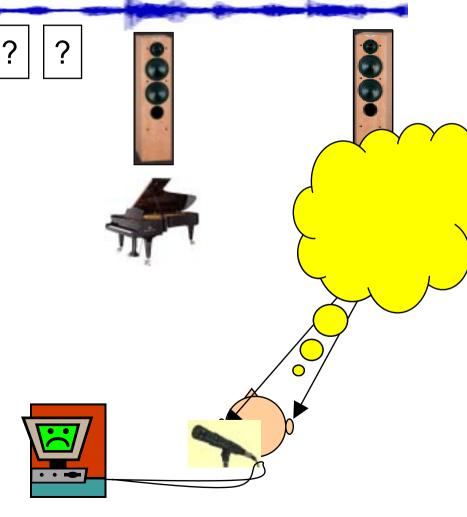
19752005

?



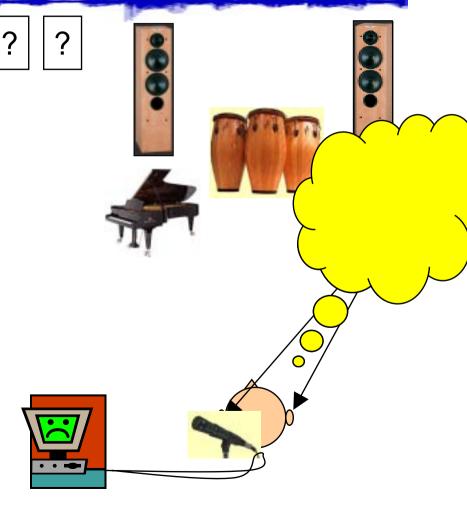


?



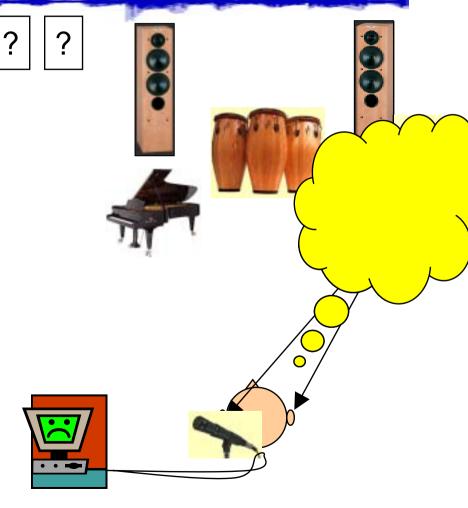


?



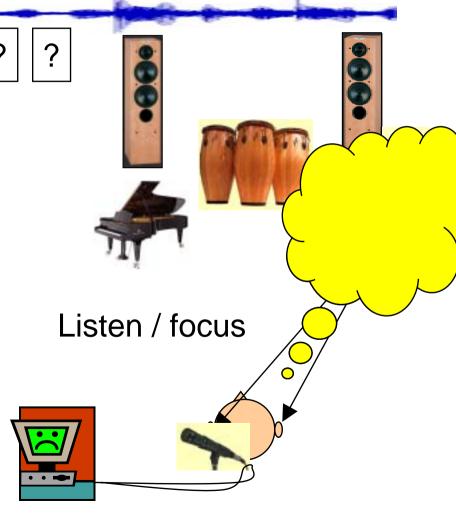


?



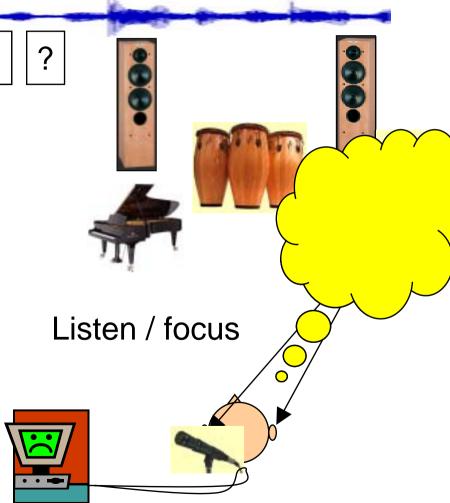


?



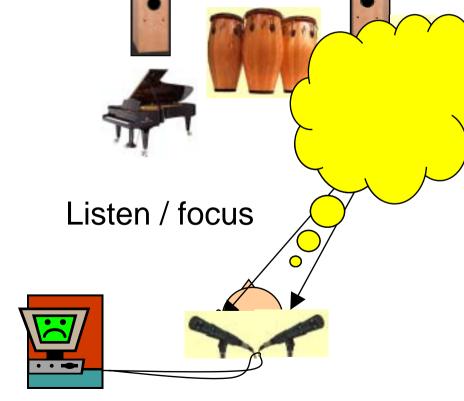


- Where is it ?
- Listening cues What : timbre
- **Pitch/frequency/patterns** 
  - Where : direction
- Time difference
- Intensity difference





- Where is it ?
- Listening cues What : timbre
- **Pitch/frequency/patterns** 
  - Where : direction
- Time difference Intensity difference





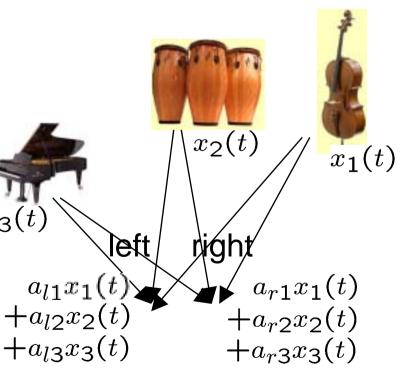


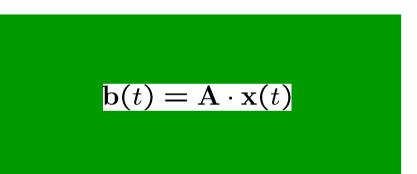
Linear algebra

Time-frequency masking

Conclusion and perspectives

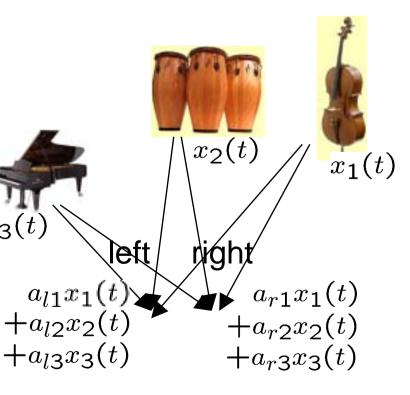


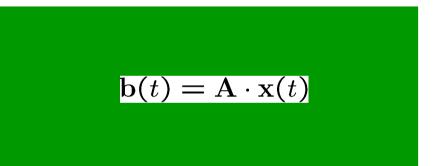






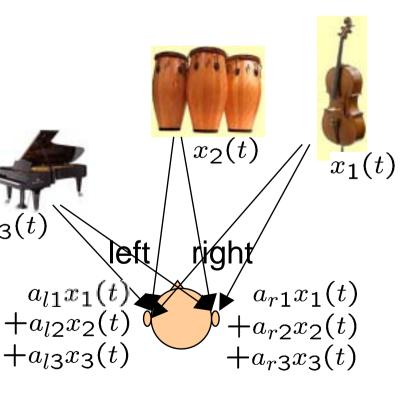
#### Linear instantaneous mixture

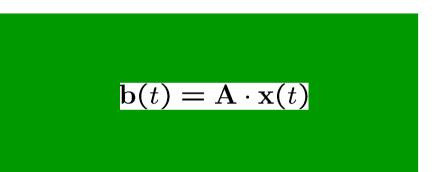






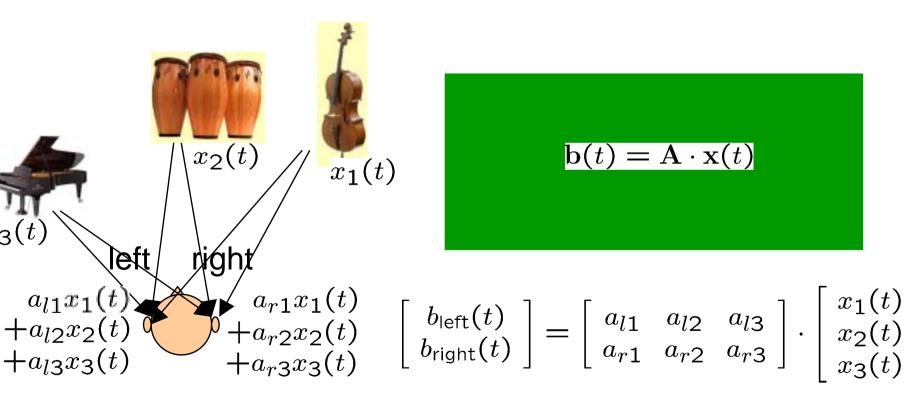
#### Linear instantaneous mixture





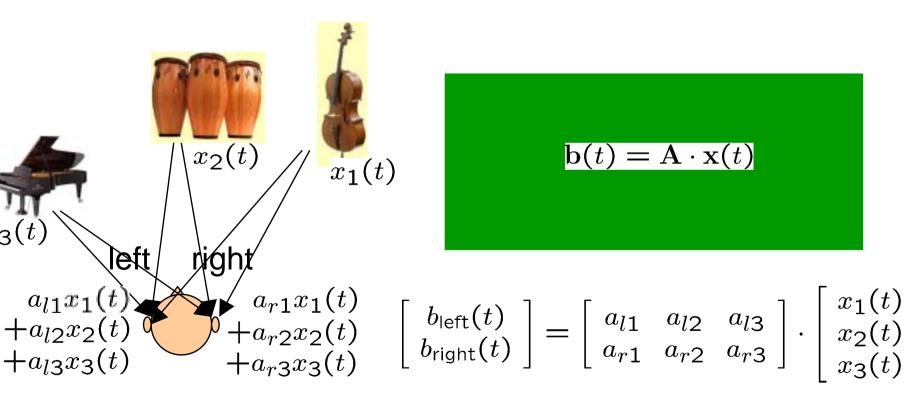


#### Linear instantaneous mixture

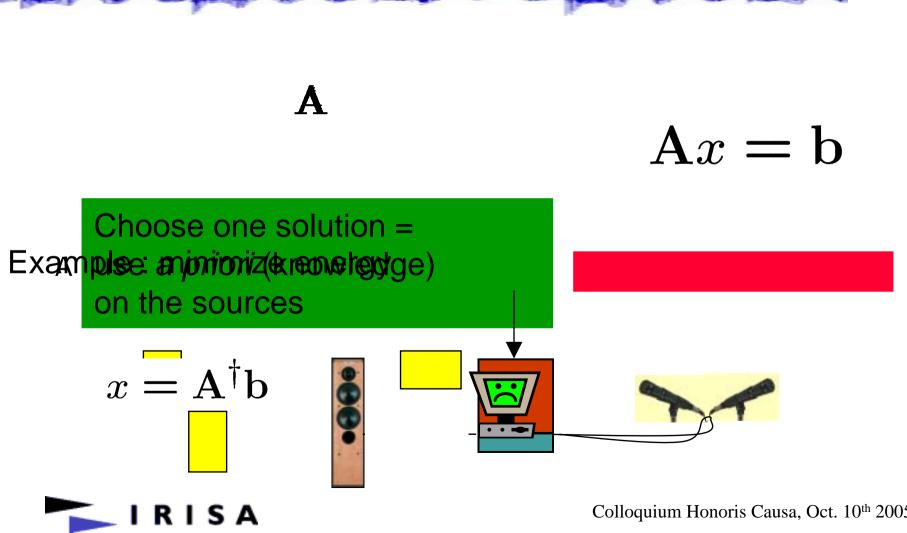


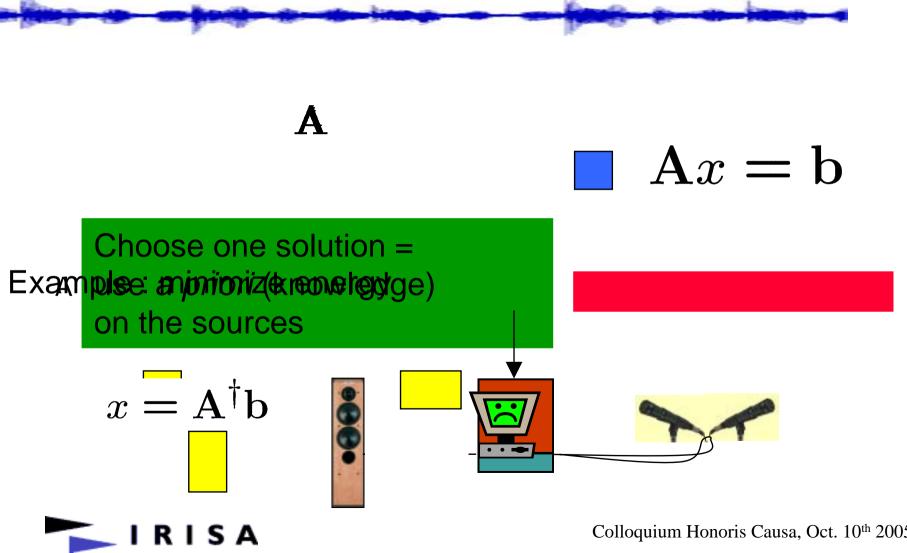


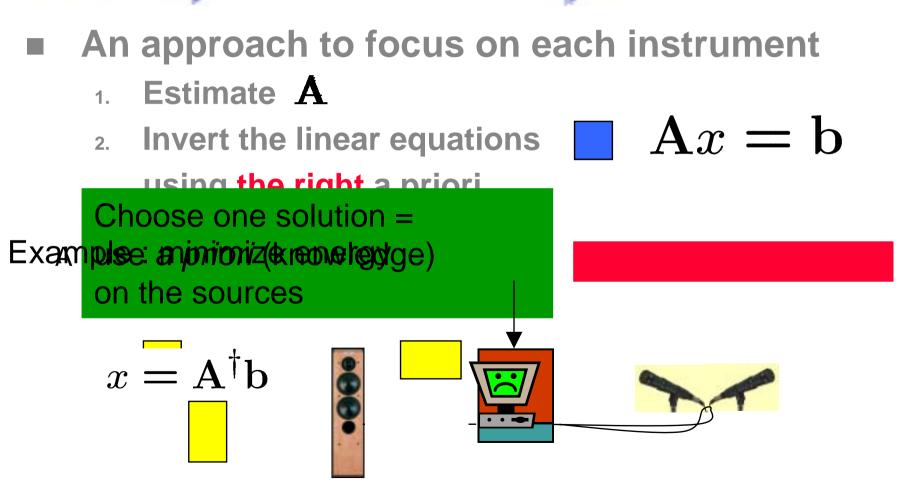
#### Linear instantaneous mixture



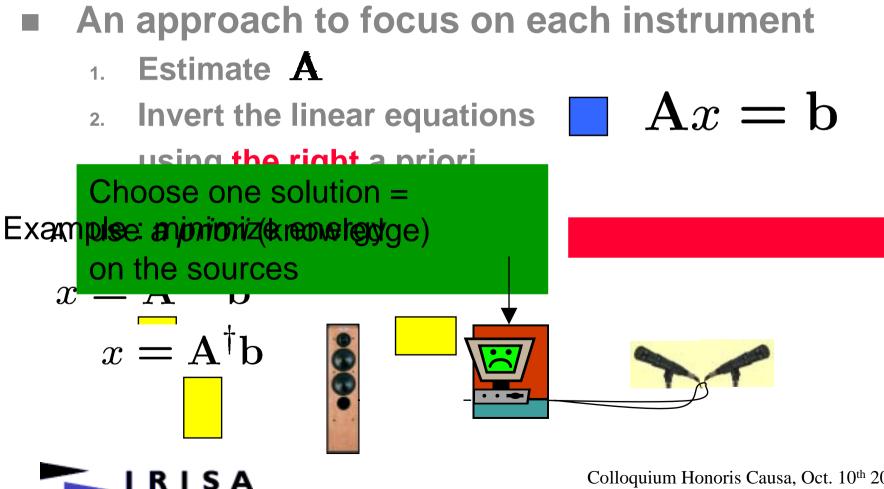




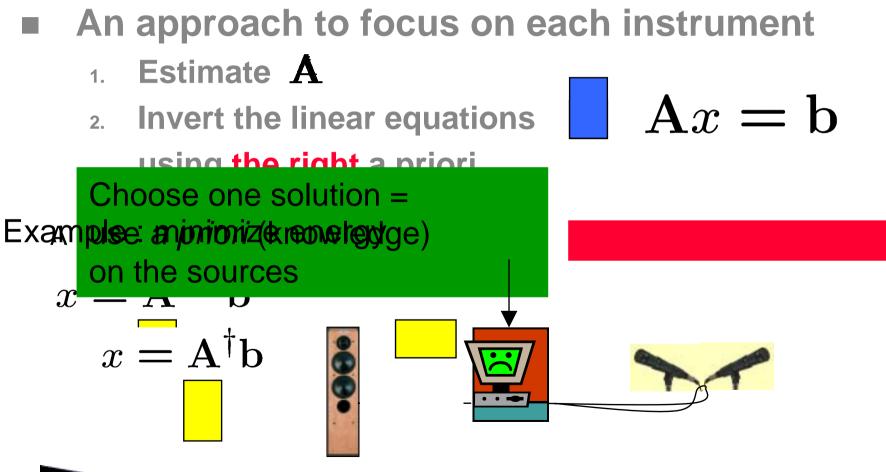


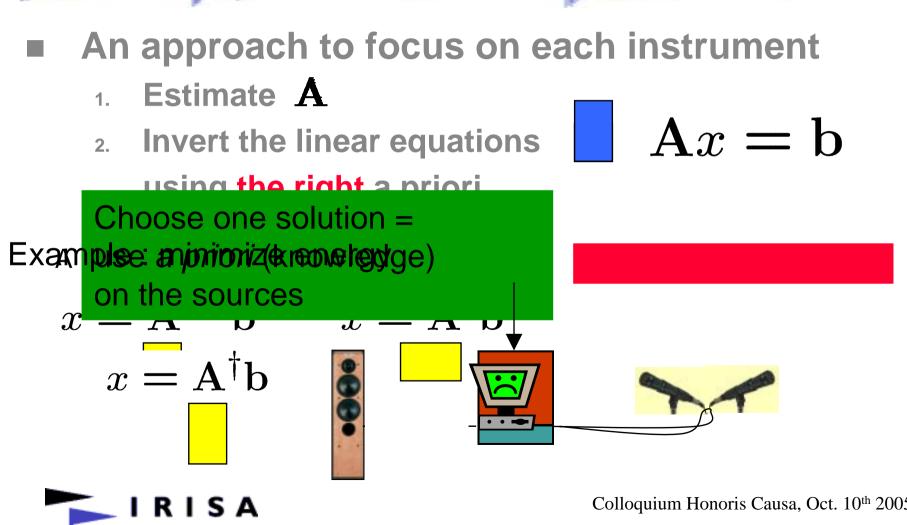


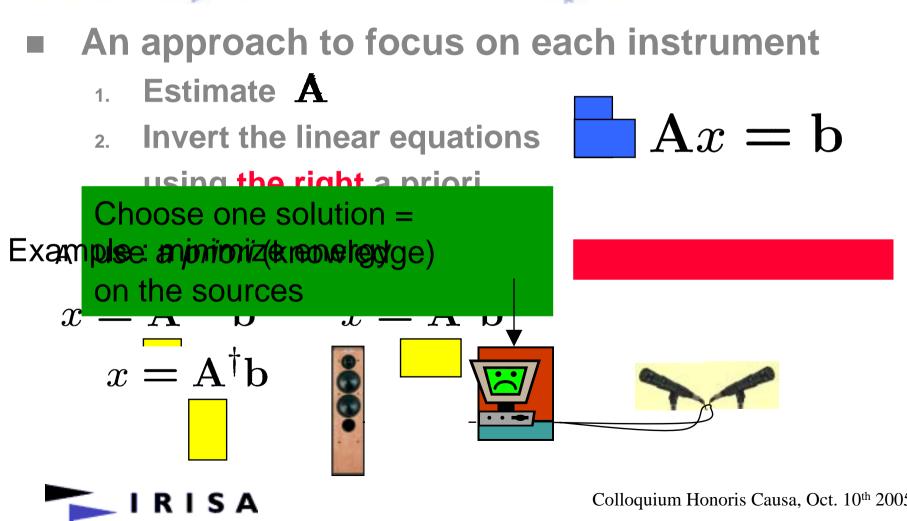


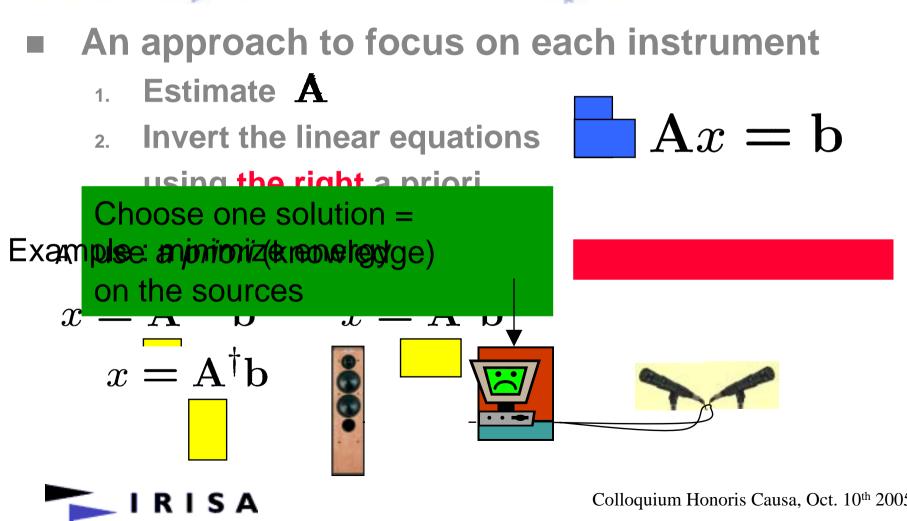


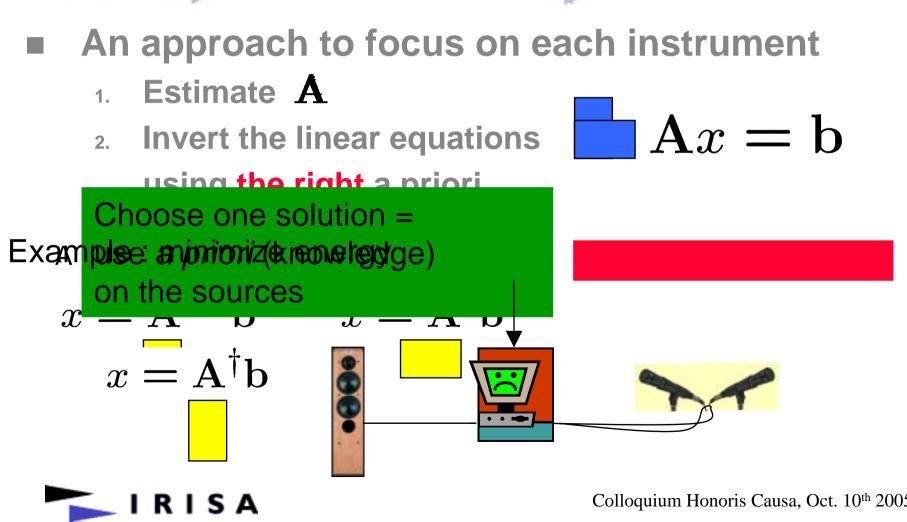
RISA



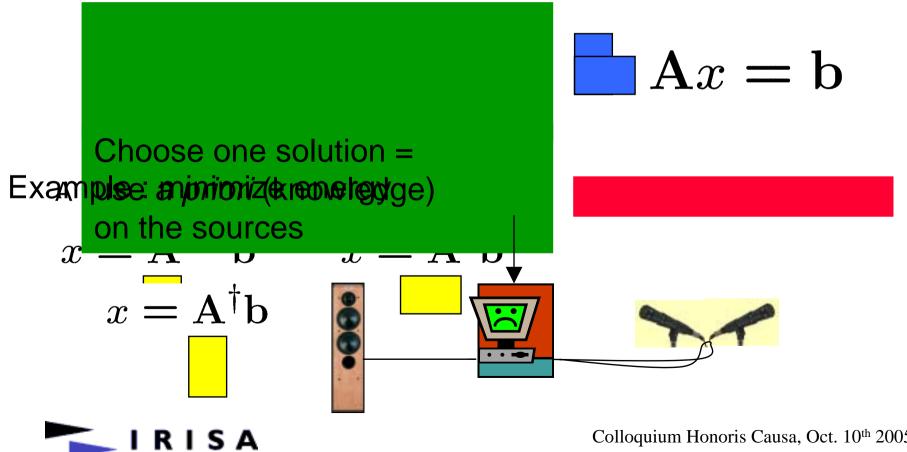




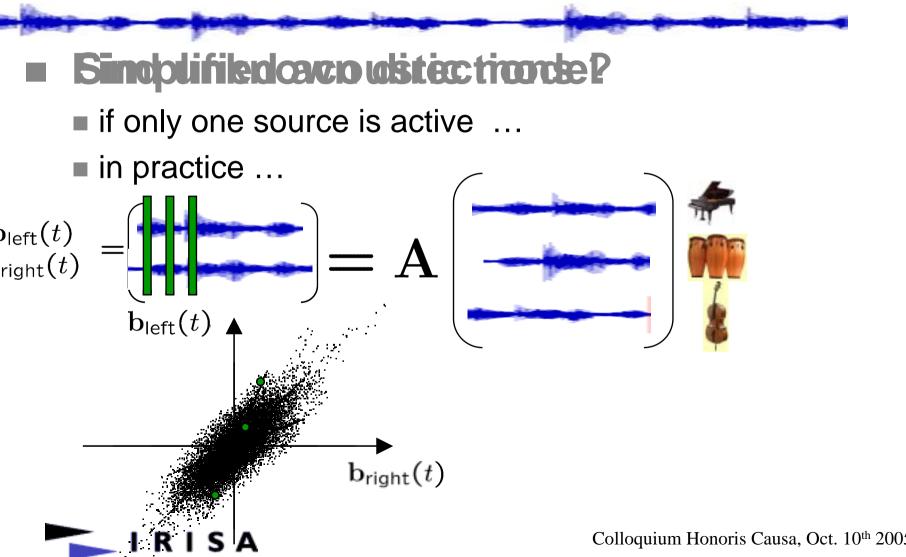




#### An approach to focus on each instrument

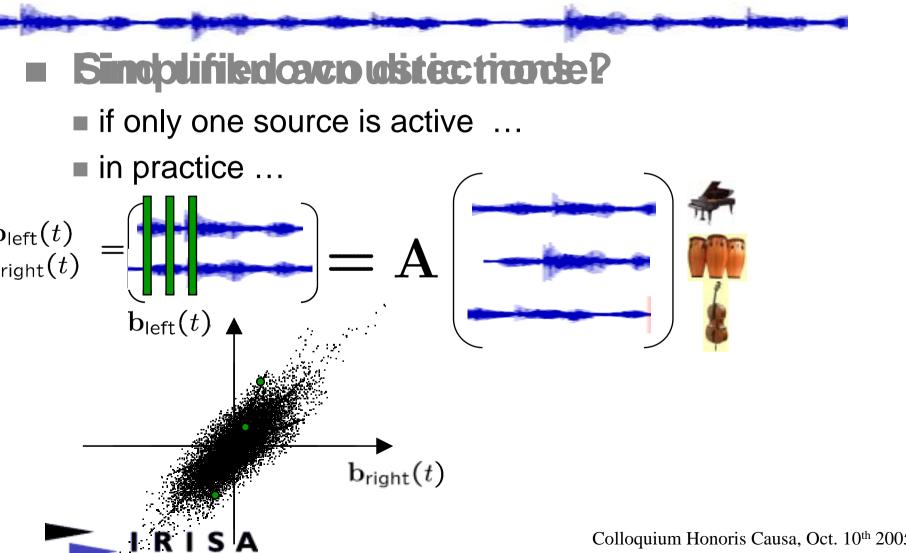


# find the directions invert the equations



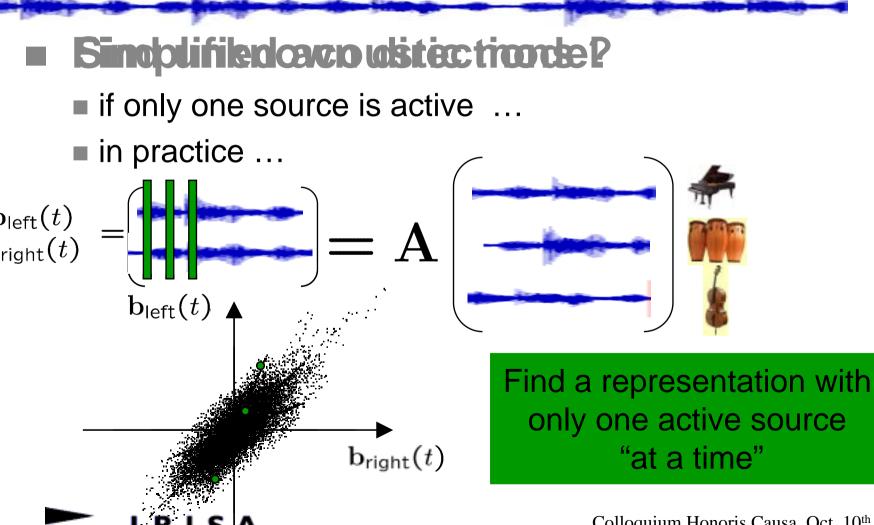
Agenda :

# find the directions invert the equations



Agenda :

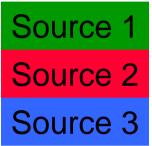
# find the directions invert the equations



Agenda :











#### Segmentation

Source 1 Source 2 Source 3





#### Segmentation

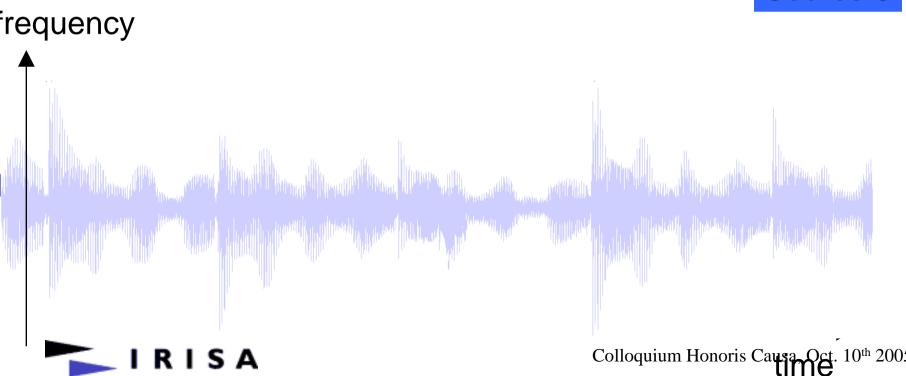
Source 1 Source 2 Source 3





#### Demultiplexing

Source 1 Source 2 Source 3



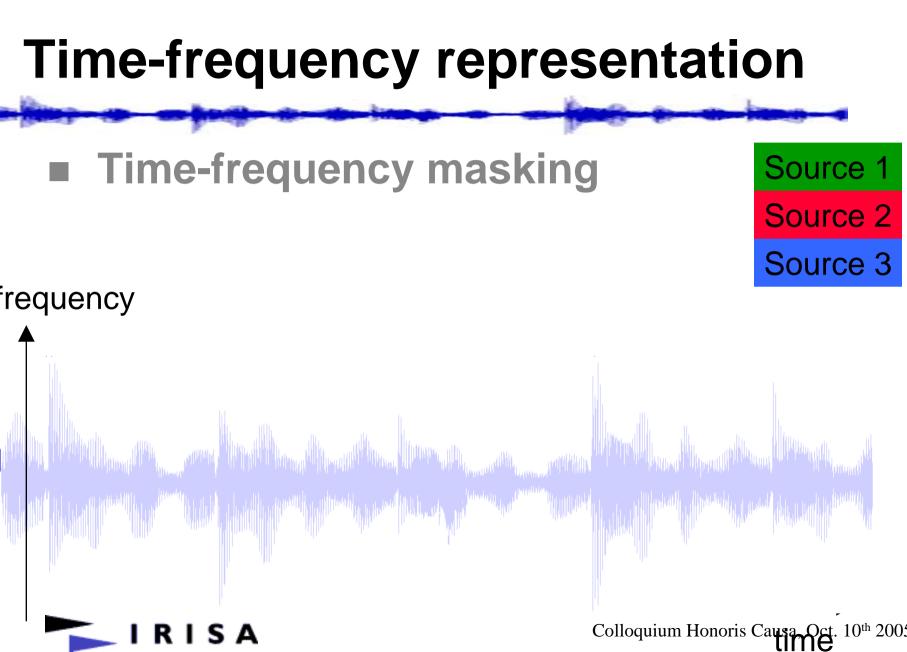


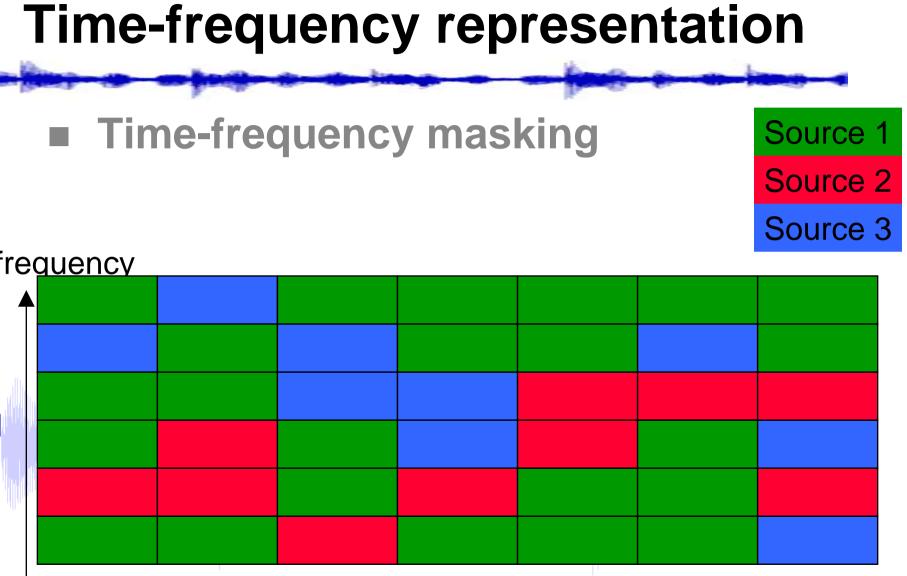
#### Demultiplexing

Source 1 Source 2 Source 3

#### requency

	JU
Colloquium Honoris Causa Oct. 10	0 <sup>th</sup> 2005







#### find the directions Agenda : invert the equations Short Time Fourier Transform (STFT) $\mathbf{b}_{\mathsf{left}}(t)$ $\mathbf{b}_{\text{left}}(\tau, f)$ $\mathbf{b}_{right}(t)$ $\mathbf{b}_{\mathsf{right}}(\tau, f)$ $\mathbf{b}_{\mathsf{left}}( au, f)$ Scatter plot Clustering $\mathbf{b}_{\mathsf{right}}(\tau, f)$ (Simon Arberet) Mask + inverse STFT IRISA Colloquium Honoris Causa, Oct. 10th 2005

#### find the directions Agenda : invert the equations Short Time Fourier Transform (STFT) $\mathbf{b}_{\mathsf{left}}(t)$ $\mathbf{b}_{\text{left}}(\tau, f)$ $\mathbf{b}_{right}(t)$ $\mathbf{b}_{\mathsf{right}}(\tau, f)$ $\mathbf{b}_{\mathsf{left}}( au, f)$ Scatter plot Clustering $\mathbf{b}_{\mathsf{right}}(\tau, f)$ (Simon Arberet) Mask + inverse STFT IRISA Colloquium Honoris Causa, Oct. 10th 2005

#### find the directions Agenda : invert the equations Short Time Fourier Transform (STFT) $\mathbf{b}_{\mathsf{left}}(t)$ $\mathbf{b}_{\text{left}}(\tau, f)$ $\mathbf{b}_{right}(t)$ $\mathbf{b}_{\mathsf{right}}(\tau, f)$ $\mathbf{b}_{\mathsf{left}}( au,f)$ Scatter plot Clustering $\mathbf{b}_{\mathsf{right}}(\tau, f)$ (Simon Arberet) Mask + inverse STFT Colloquium Honoris Causa, Oct. 10th 2005 IRISA

# Agenda : I find the directions invert the equations



 $\mathbf{b}_{\mathsf{left}}( au,f)$ 



### Scatter plot

- Clustering
  (Simon Arberet)
  - Mask + inverse STFT

Colloquium Honoris Causa, Oct. 10th 2005

 $\mathbf{b}_{\text{left}}(\tau, f)$ 

 $\mathbf{b}_{\mathsf{right}}(\tau, f)$ 

 $\mathbf{b}_{\mathsf{right}}(\tau, f)$ 

# Agenda : I find the directions invert the equations



 $\mathbf{b}_{\mathsf{left}}( au,f)$ 



### Scatter plot

- Clustering
  (Simon Arberet)
  - Mask + inverse STFT

Colloquium Honoris Causa, Oct. 10th 2005

 $\mathbf{b}_{\text{left}}(\tau, f)$ 

 $\mathbf{b}_{\mathsf{right}}(\tau, f)$ 

 $\mathbf{b}_{\mathsf{right}}(\tau, f)$ 

# Agenda : Ind the directionsInvert the equations



 $\mathbf{b}_{\mathsf{left}}( au,f)$ 



### Scatter plot

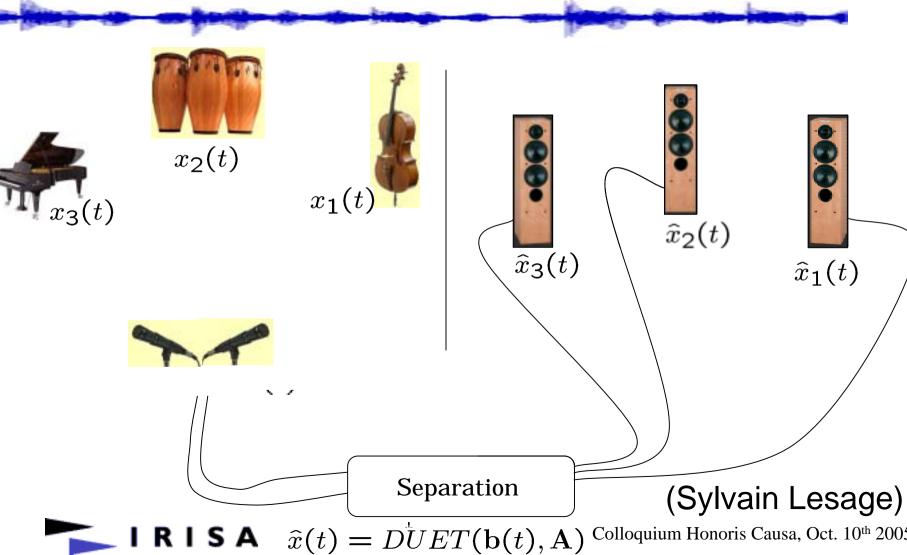
- Clustering
  (Simon Arberet)
  - Mask + inverse STFT

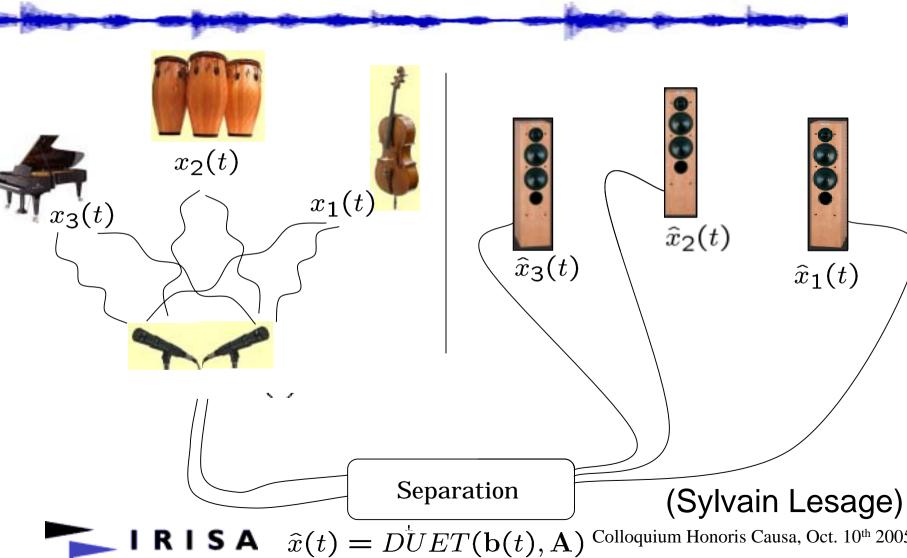
Colloquium Honoris Causa, Oct. 10th 2005

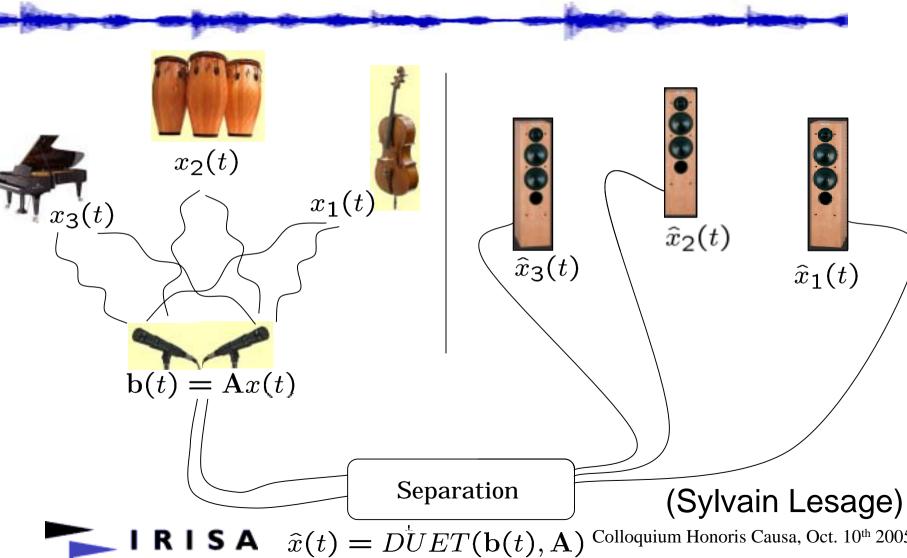
 $\mathbf{b}_{\text{left}}(\tau, f)$ 

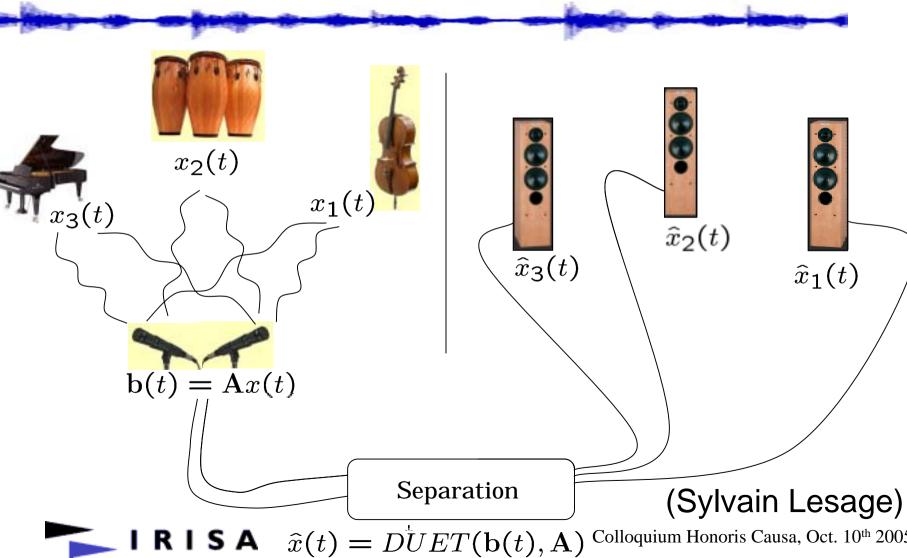
 $\mathbf{b}_{\mathsf{right}}(\tau, f)$ 

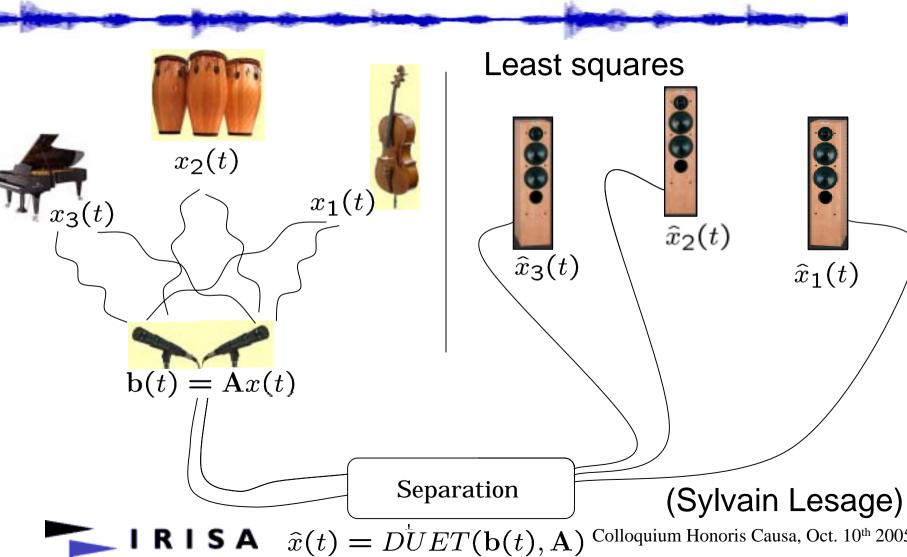
 $\mathbf{b}_{\mathsf{right}}(\tau, f)$ 

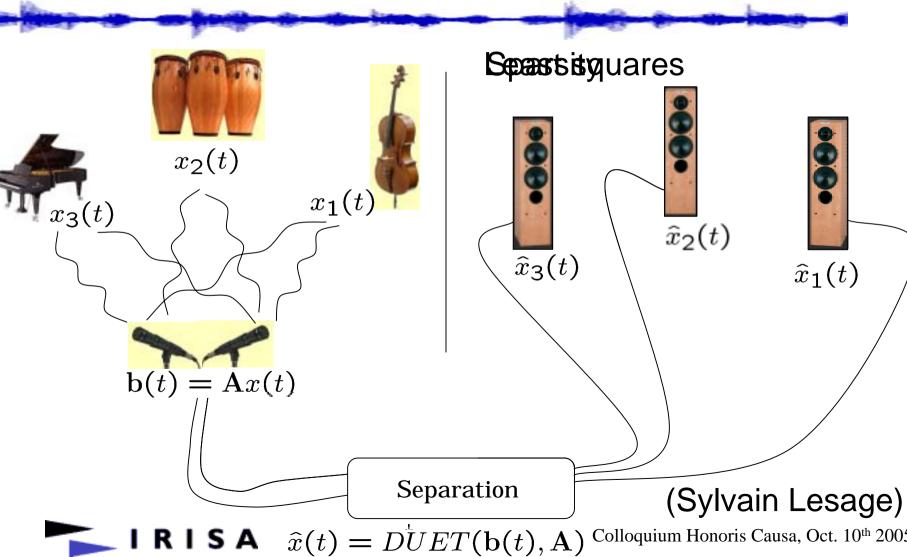




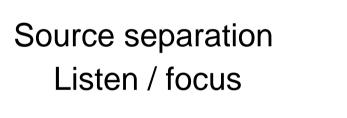


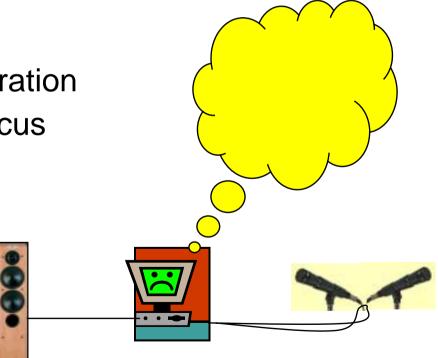




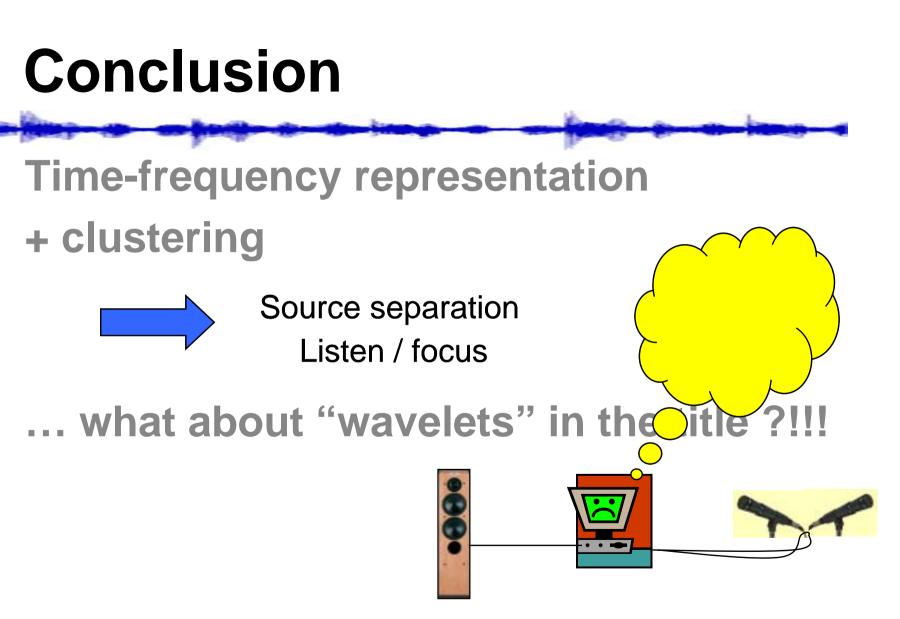














### What about wavelets ?

x(t)

Many nonzero values ~Gaussian histogram x(T, J)Mostly zeroes = sparse ~Laplacian histogram





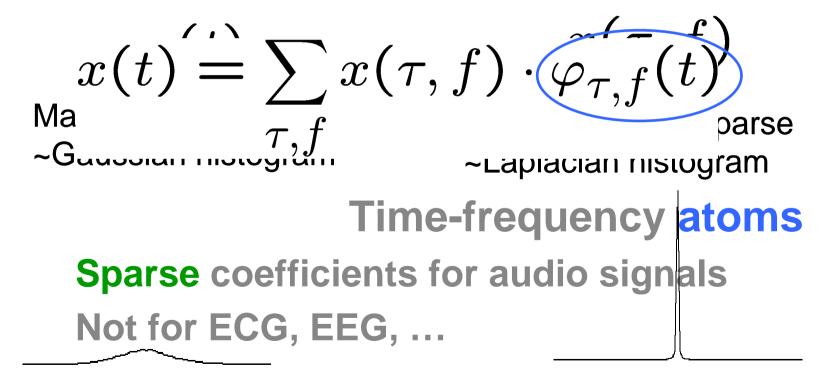
### What about wavelets ?

$$x(t) = \sum_{\substack{\tau, f \\ \text{-Guassian motory, and}}} x(\tau, f) \cdot \varphi_{\tau, f}(t)$$
parse  
-Lapiacian mistogram



## What about wavelets ?

#### Role of time-frequency representation





### Some other possible waveforms

- Gabor atomsBandeletsDiracPointlets
- Local cosine basis
  - Wavelets
- Wavelet packets

RISA

Curvelets

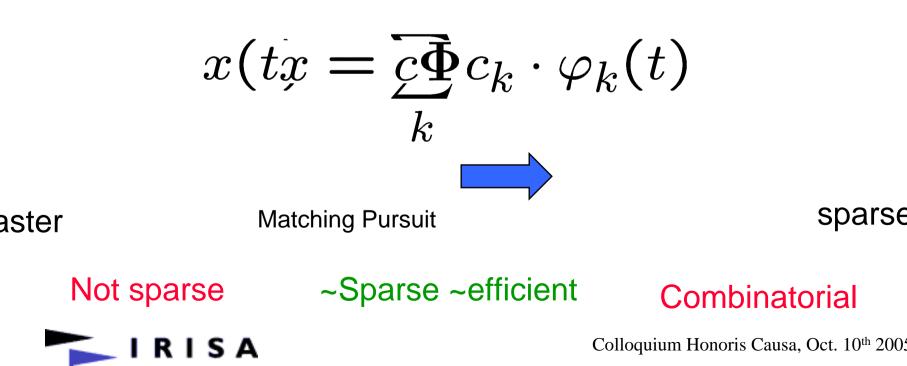
Brushlets

Coiflets

Ridgelets

## Sparse decompositions with redundant dictionaries

 $\{\varphi_k(t)\}$ 



# Sparse decompositions with redundant dictionaries

- Redundant dictionary of waveforms  $\{\varphi_k(t)\}$
- Sparse representation

**Matching Pursuit** 

- $x(tx = c\Phi c_k \cdot \varphi_k(t)$ Redun  $k \quad ed$ too many solutions compute one
- aster

Not sparse

IRISA

~Sparse ~efficient

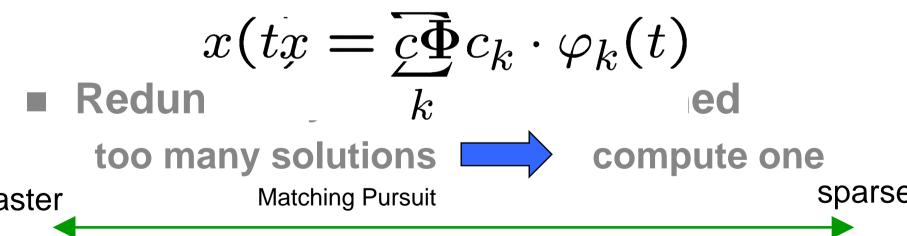
#### Combinatorial

sparse





- Redundant dictionary of waveforms  $\{\varphi_k(t)\}$
- Sparse representation



Not sparse

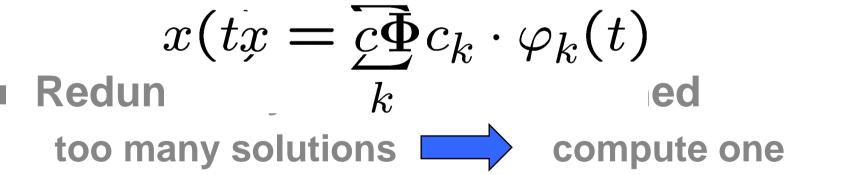
RISA

~Sparse ~efficient

**Combinatorial** Colloquium Honoris Causa, Oct. 10<sup>th</sup> 2003

## **Sparse decompositions** with redundant dictionaries

- **Redundant dictionary of waveforms**  $\{\varphi_k(t)\}$
- **Sparse representation**



**Matching Pursuit** sparse aster  $\min \|c\|_2$ ~Sparse ~efficient Not sparse Combinatorial RISA

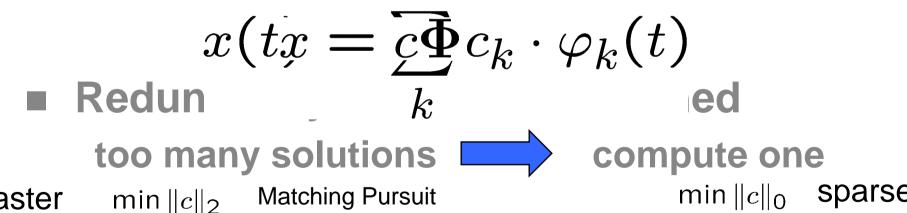
Colloquium Honoris Causa, Oct. 10<sup>th</sup> 2005

# Sparse decompositions with redundant dictionaries

- Redundant dictionary of waveforms  $\{\varphi_k(t)\}$
- Sparse representation

Not sparse

RISA



~Sparse ~efficient

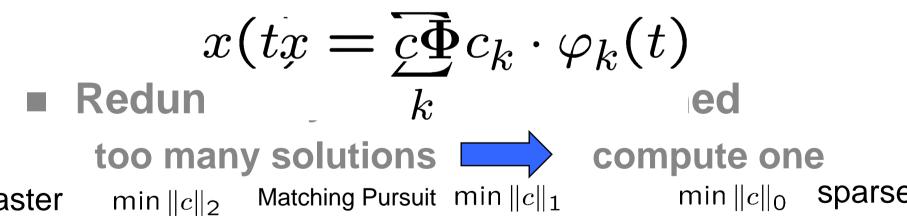
Combinatorial

# Sparse decompositions with redundant dictionaries

- Redundant dictionary of waveforms  $\{\varphi_k(t)\}$
- Sparse representation

Not sparse

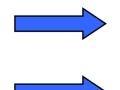
RISA



~Sparse ~efficient

Combinatorial







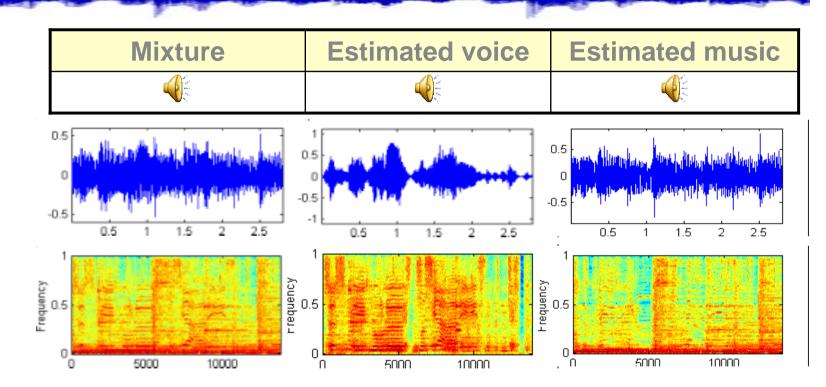
## Conclusion

- Sparse representations help for source separation
- Some other applications
  - Low bitrate image / signal compression (JPEG,MP3) Communications (MIMO) ?
- Main challenges
  - Adapted dictionaries
  - *(ECG, EEG, geology, financial data, ...)* Efficient decomposition algorithms
- What about computer audition ?
  - Simple acoustic model = approximate
    - room acoustics = convolution
  - Sparsity = rough source model

statistical source models



## Beyond sparsity : Monaural source separation



Tools: -time-frequency representations -statistical models

(Alexey Ozerov, FT R&D)



## To go further ... beyond sparsity

**Important dates**:  $b = SPARS05 \approx A \cdot [1, 1, 1, 1, 0, ...]^T$  **June : abstracts** = [Signal, Processing, Adaptive, Sparse, Structured, Representations, ... **September : registration** 

### http://spars05.irisa.fr spars05@irisa.fr







UNIVERSITE DE RENNES "





## To go further ... beyond sparsity

**Important dates**:  $b = SPARS05 \approx A \cdot [1, 1, 1, 1, 0, ...]^T$  **June : abstracts** = [Signal, Processing, Adaptive, Sparse, Structured, Representations, ... **Septemper : registration** 

isis

UNIVERSITE DE RENINES 1

SPARS05, IRISA, Rennes, 16-18 November 2005

http://spars05.irisa.fr spars05@irisa.fr

RINRIA





## To go further ... beyond sparsity

**Important dates**:  $b = SPARS05 \approx A \cdot [1, 1, 1, 1, 0, ...]^T$  **June : abstracts** = [Signal, Processing, Adaptive, Sparse, Structured, Representations, ... **September : registration** 

- SPARS05, IRISA, Rennes, 16-18 November 2005
- Invited speakers:

IRISA

S. MALLAT, Ecole Poytechnique, France J. TROPP, Michigan Univ., USA MIDAY//S.OBPS/057, Misaf Iqndon, UK M. ZIBULEVSKY, Technion Inst., Israël Spars05@irisa.fr











#### To go further . beyond sparsity . .

h =	SPARS05 $\approx \mathbf{A} \cdot [1, 1, 1, 1, 0,]^T$	
_ [Si	June : abstracts ignal, Processing, Adaptive, Sparse, Structured	Representations
_[0	Septemper : registration	
	SPARS05, IRISA, Rennes, 16-18 November 2005	
÷.,	Invited speakers:	
	S. MALLAT, Ecole Poytechnique, France	
	J. TROPP, Michigan Univ., USA	E
	M. ZIBULEVSKY, Technion Inst., Israël	Hit is the second
	Organizepars05@irisa.fr CNRS/MathSTIC "Sparse structured approximation	IS IS
	of audio signals"	
		and the second sec



IRISA







## To go further beyond sparsity

	P P	
b = = [S	Immortant dates · SPARS05 $\approx$ A · [1, 1, 1, 1, 1, 0,] <sup>T</sup> June : abstracts ignal, Processing, Adaptive, Sparse, Structured, September : registration	Representations, .
:	SPARS05, IRISA, Rennes, 16-18 November 2005 Invited speakers: S. MALLAT, Ecole Poytechnique, France J. TROPP, Michigan Univ., USA MIDAVIS, OBASONS, MISA Ifondon, UK M. ZIBULEVSKY, Technion Inst., Israël	
	Organizepars05@irisa.fr CNRS/MathSTIC "Sparse structured approximations	AREAS STREET

of audio signals"



RISA









## To go further Bibliography

- S. Mallat, Wavelet Tour of Signal Processing Academic Press, 1998
- Special issue EURASIP Signal Processing
  Sparse Approximations in Signal and Image Processing (R. Gribonval & M. Nielsen eds), in preparation
- Theory + humor : J. Tropp Greed is good: algorithmic results for sparse approximations, IEEE Trans. Inform. Th., Octobre 2004

http://www.irisa.fr/metiss/gribonval/



## To go further Matching Pursuit Toolkit

 $\Phi$  M



## To go further Matching Pursuit Toolkit

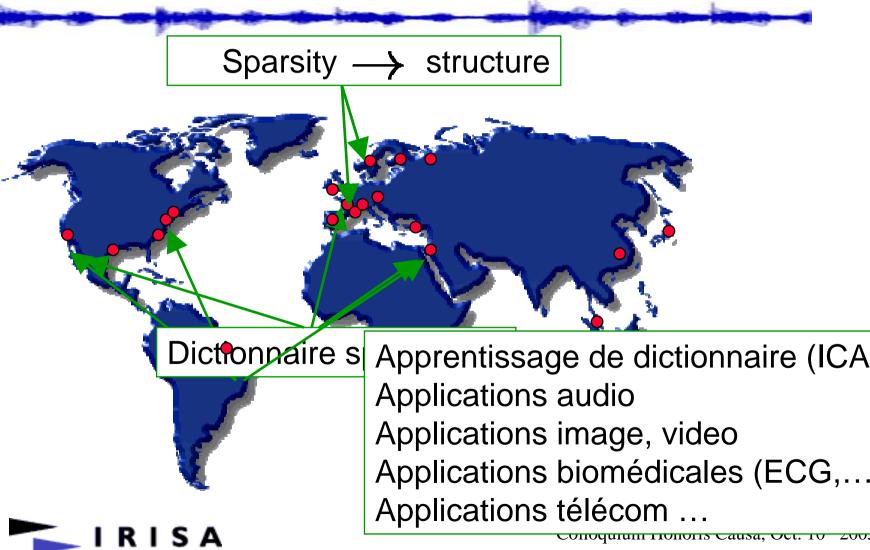
Flexible XML description of dictionaries

Independent of external interface (no Matlab, no LastWave)

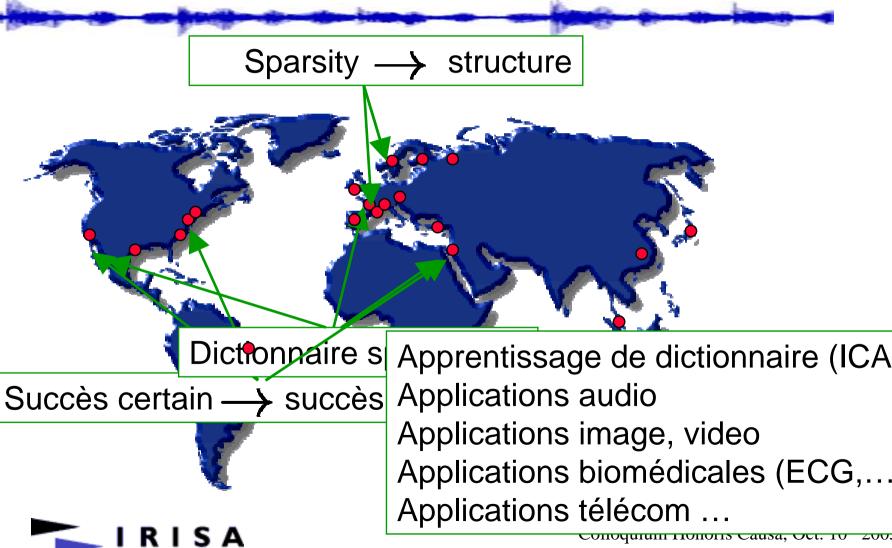
Multi-platform (Unix based), std C++, open source sacha@irisa.fr, remi.gribonval@irisa.fr



## To go further Some ongoing extensions

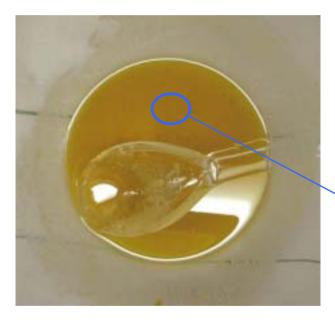


## To go further Some ongoing extensions



# **Time-frequency representation**

#### 1 mixture



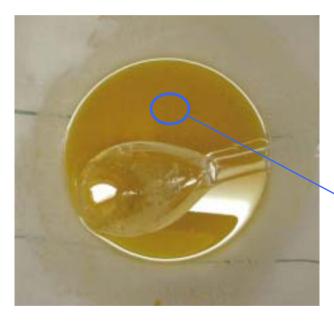






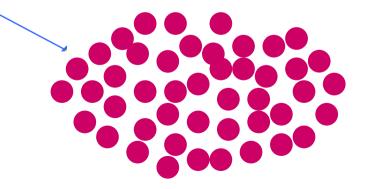
# **Time-frequency representation**

#### 1 mixture



#### ■2 sources

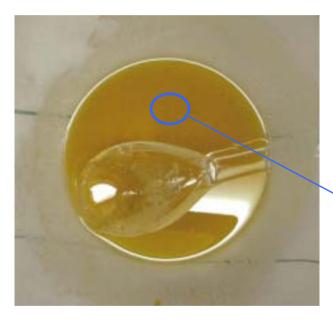






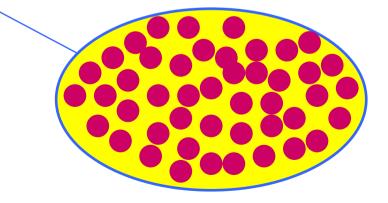
# **Time-frequency representation**

#### 1 mixture



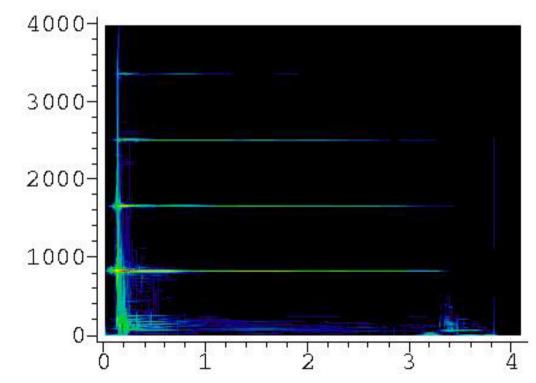
#### ■2 sources





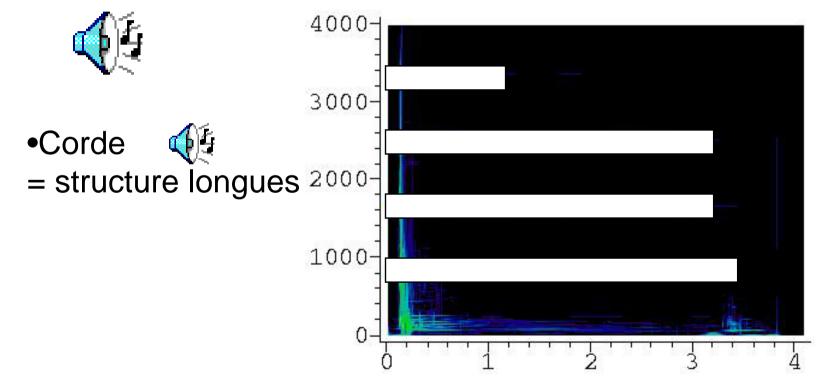


Une note de piano, 4 secondes à 8kHz



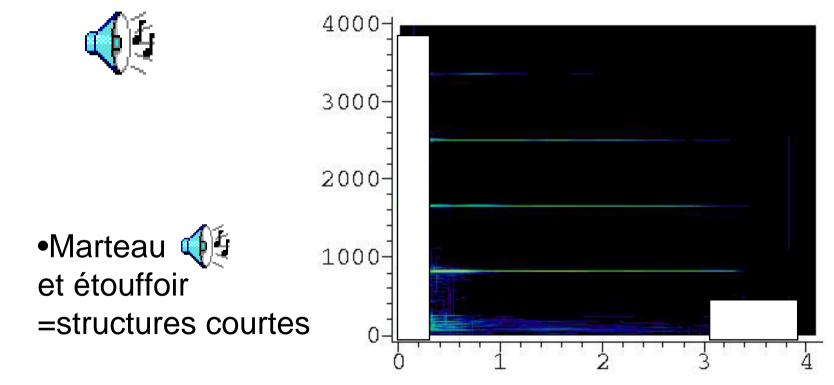


Une note de piano, 4 secondes à 8kHz



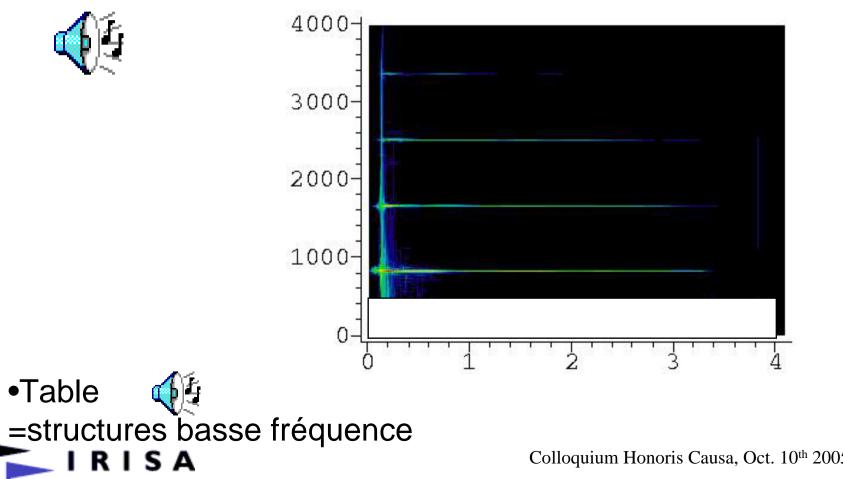


Une note de piano, 4 secondes à 8kHz

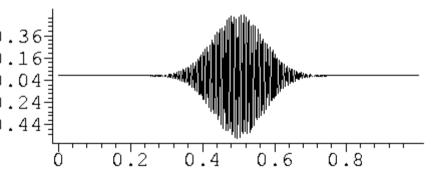


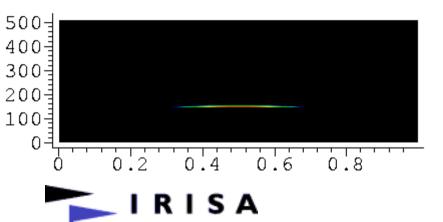


Une note de piano, 4 secondes à 8kHz



Time-frequency atom

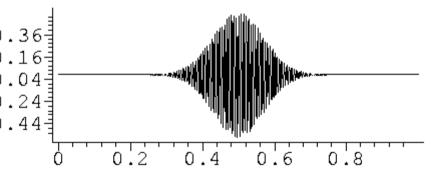


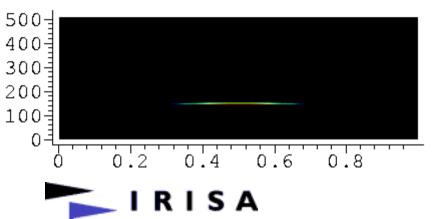


$$\varphi_{s,\tau,f}(t) = \frac{1}{\sqrt{s}} w\left(\frac{t-\tau}{s}\right) e^{2i\pi ft} e^{i\pi c(t-\tau)^2}$$

Time-frequency atom

Scale s, time t, frequency f





$$\varphi_{s,\tau,f}(t) = \frac{1}{\sqrt{s}} w\left(\frac{t-\tau}{s}\right) e^{2i\pi ft} e^{i\pi c(t-\tau)^2}$$

