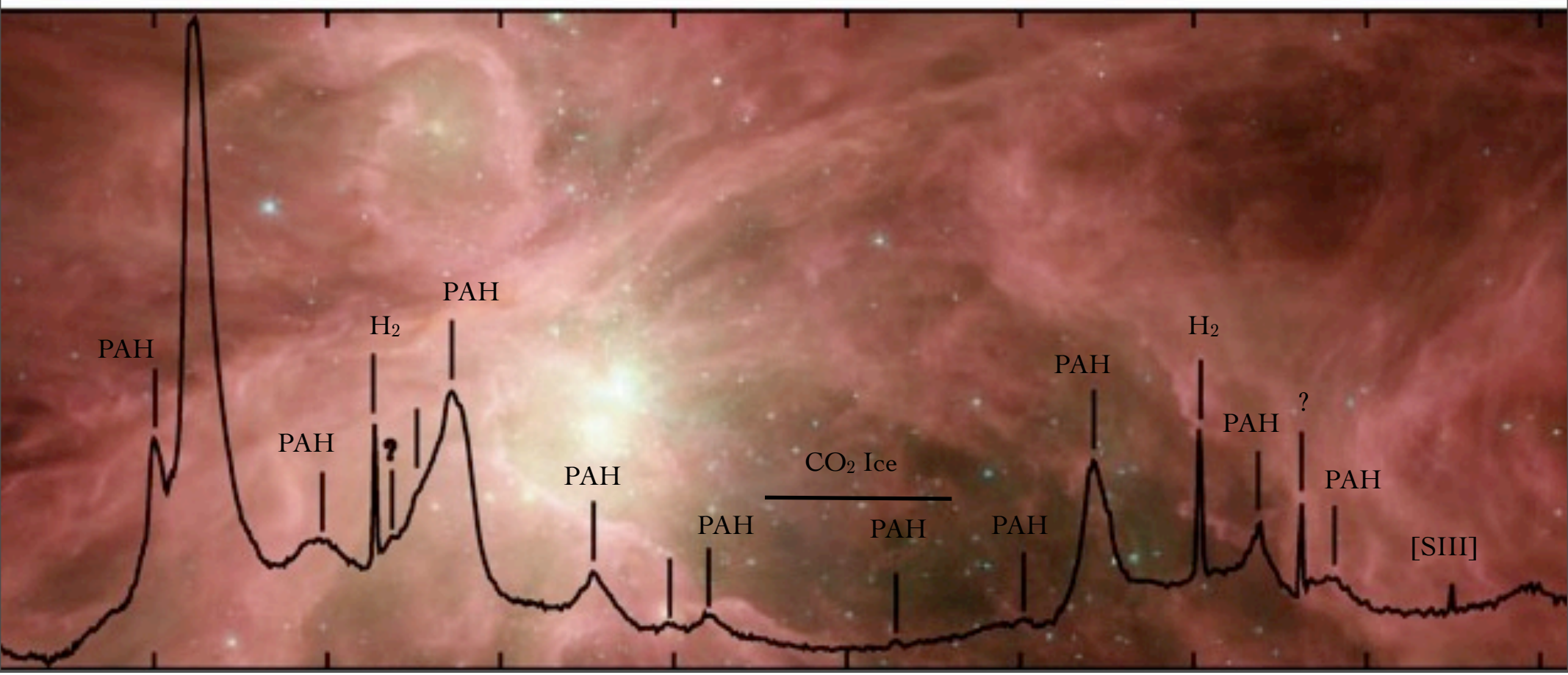
Méthodes de séparation aveugle de sources (BSS) en astronomie IR et radio-millimétrique

Olivier Berné (IRAP, CNRS/Univ. Toulouse),

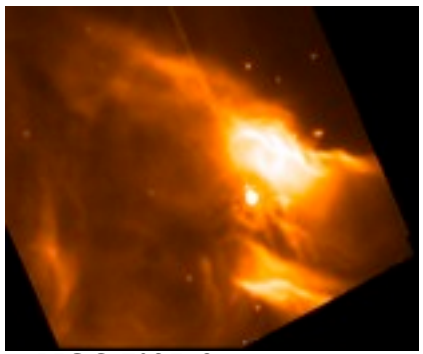
C. Joblin (IRAP), Y. Deville (IRAP), P. Pilleri (OAN, Madrid), J. Pety (IRAM), D. Teyssier (ESA, Madrid), M. Gerin (LERMA), A. Fuente (OAN, Madrid)



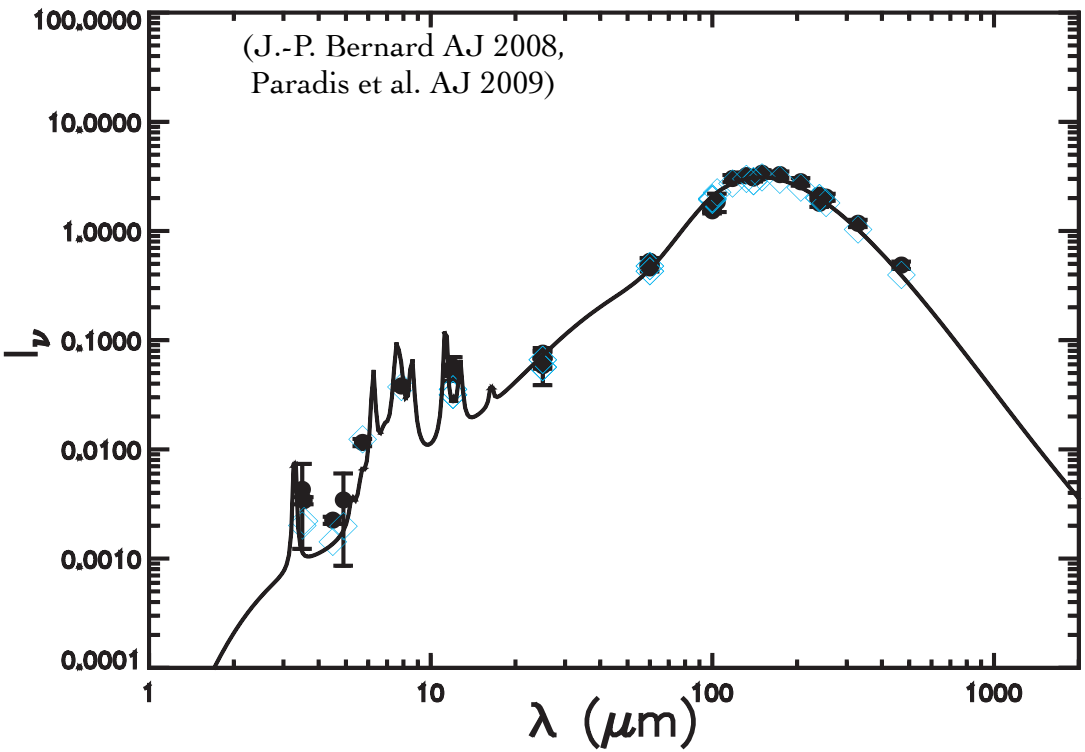
Outline of the presentation

- ☆ The variation of the mid-infrared spectrum in galaxies
- ☆ Unveiling the origins of these variations with BSS
- ☆ Hyperspectral radioastronomy and BSS methods
- ☆ Limitations of BSS for hyperspectral radio-astronomy
- ☆ Conclusions and perspectives

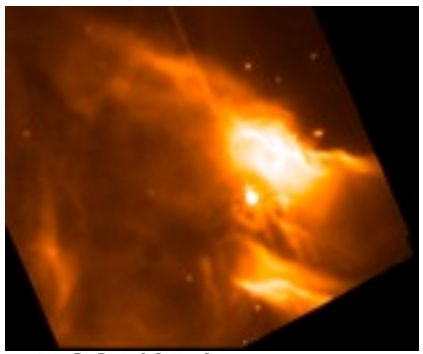
The infrared spectrum of galaxies



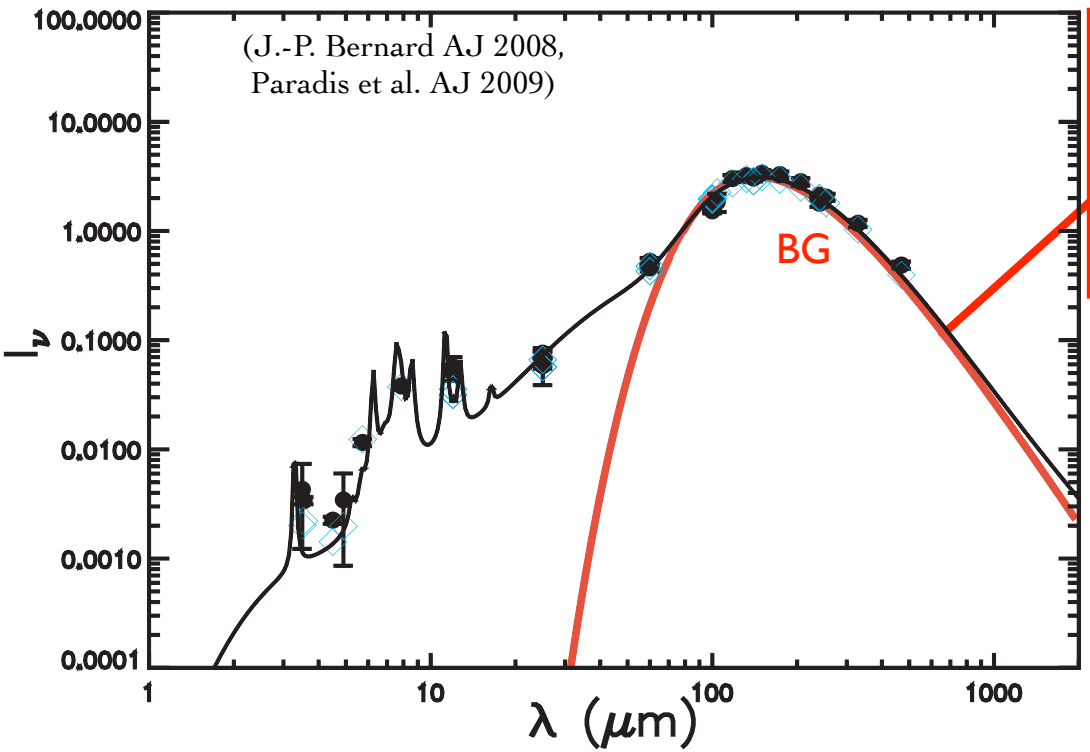
NGC 7023, 8 μm



The infrared spectrum of galaxies



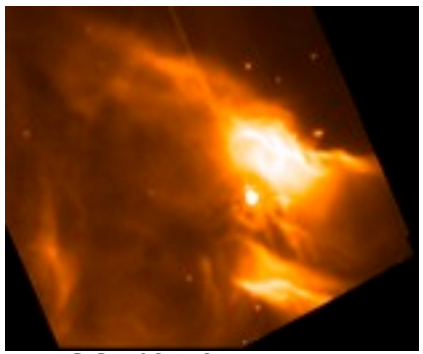
NGC 7023, 8 μm



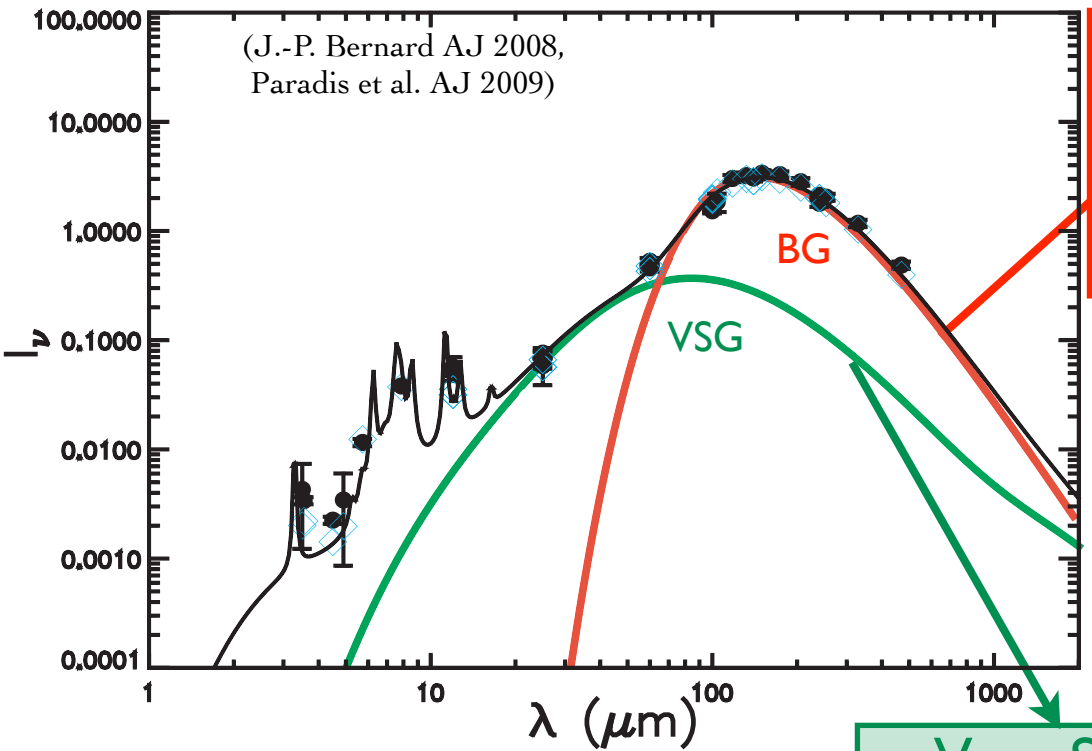
Big Grains (BG)

Silicate
size > 0.1 μm

The infrared spectrum of galaxies



NGC 7023, 8 μm



Big Grains (BG)

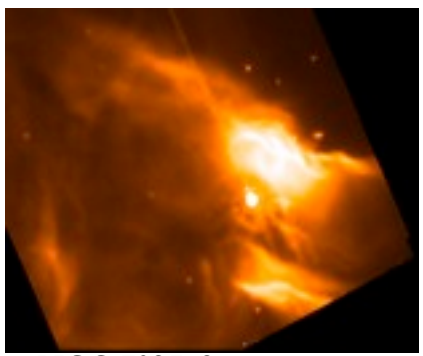
Silicate
size > 0.1 μm

Very Small Grains (VSGs)

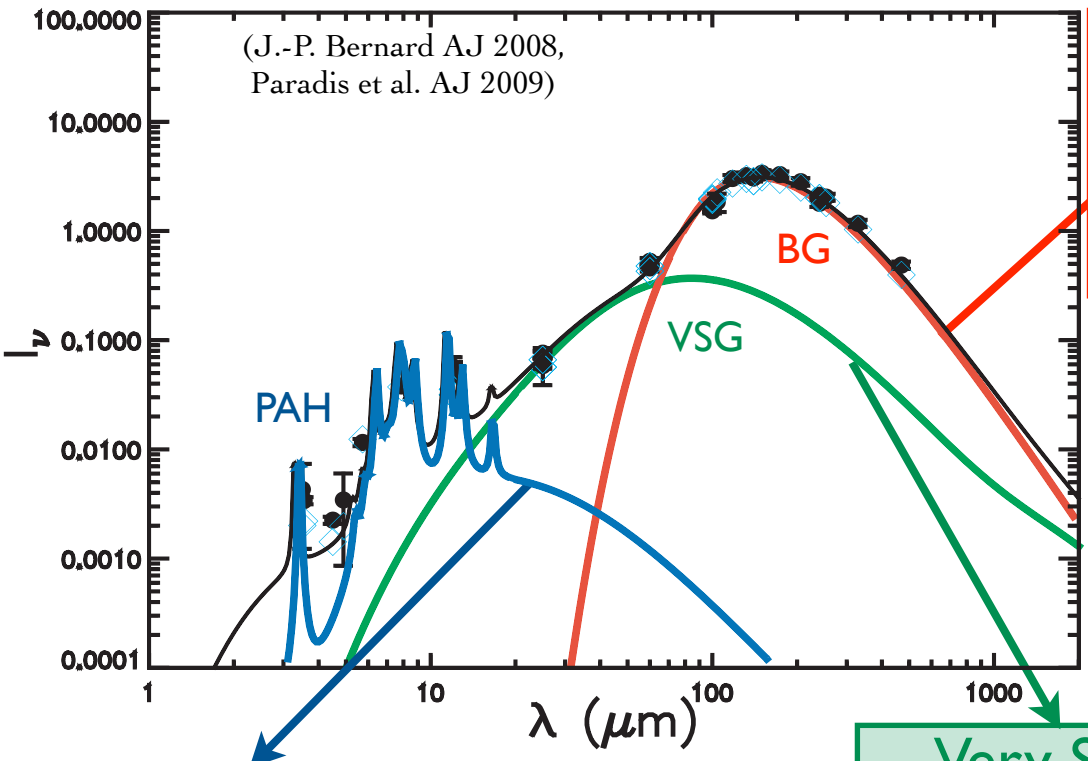
carbonaceous,
10-100 nm...

(Désert et al. A&A, 1990; Draine & Li, ApJ, 2001)

The infrared spectrum of galaxies



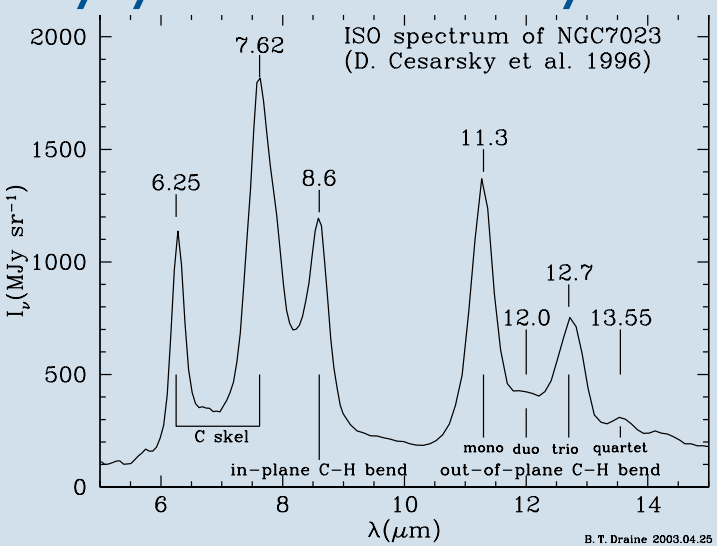
NGC 7023, 8 μm



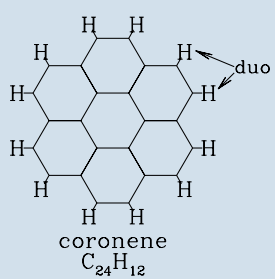
Big Grains (BG)

Silicate
size > 0.1 μm

Polycyclic Aromatic Hydrocarbons (PAH)



No specific PAH identified yet



(Léger & Pujet A&A 1984;
Allamandola, Tielens, Barker ApJ 1985;
Draine ARA&A 2003)

Very Small Grains (VSGs)

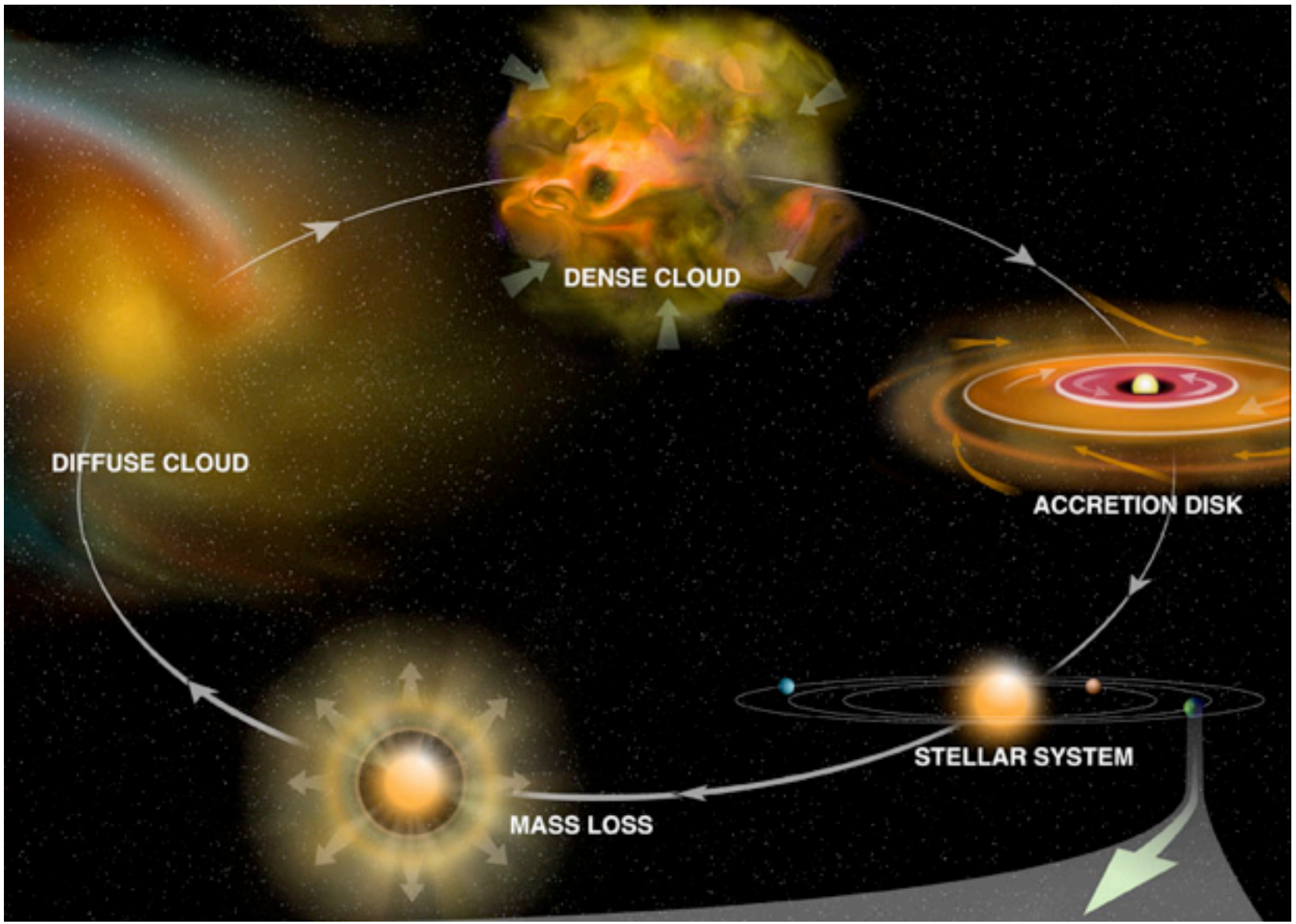
carbonaceous,
10-100 nm...

(Désert et al. A&A, 1990; Draine & Li, ApJ, 2001)

Some examples

[Peeters et al. 2002, Van Didenhoven 2004, Sloan et al. 2007]

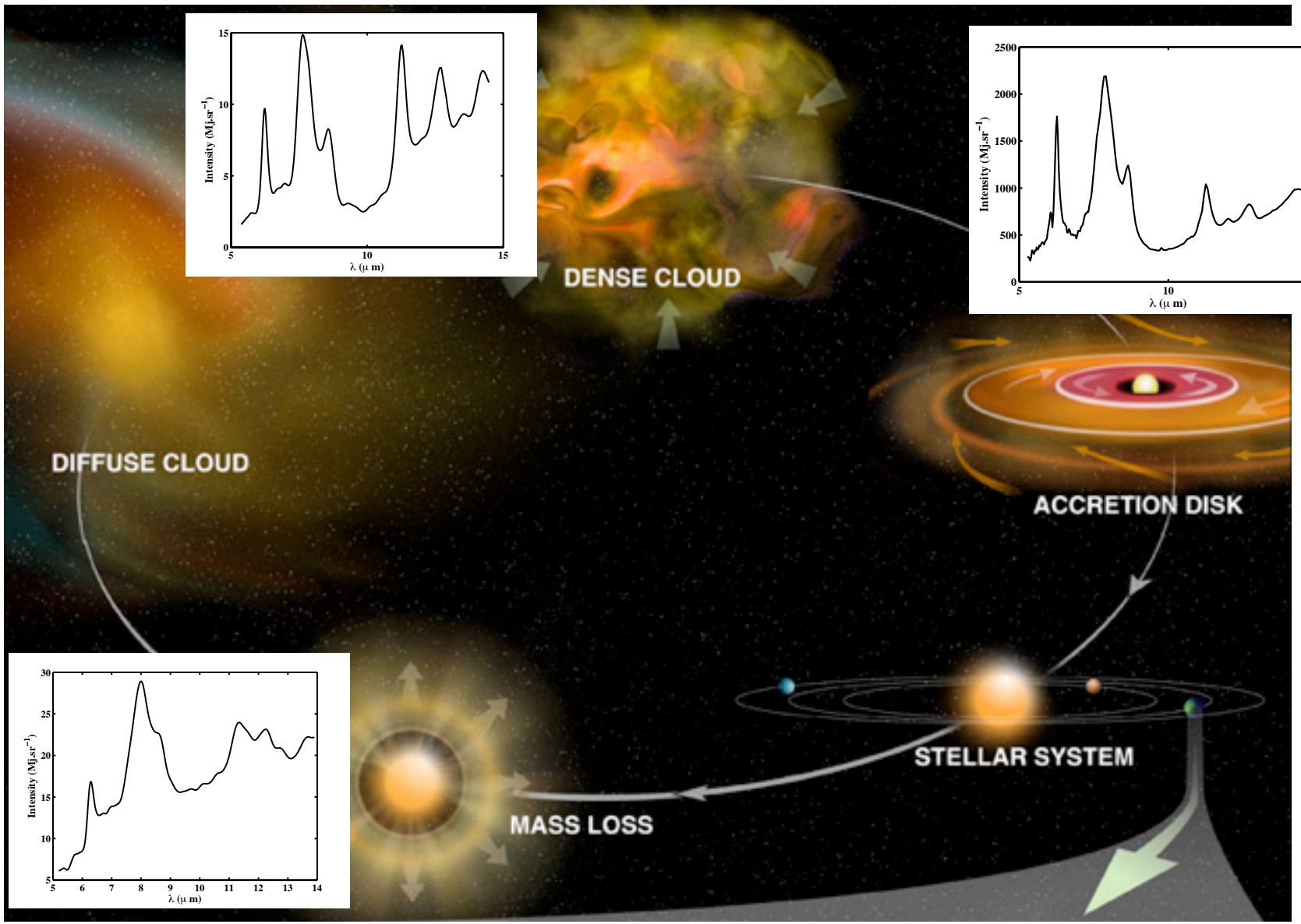
Life of stars $M<8M_{\odot}$



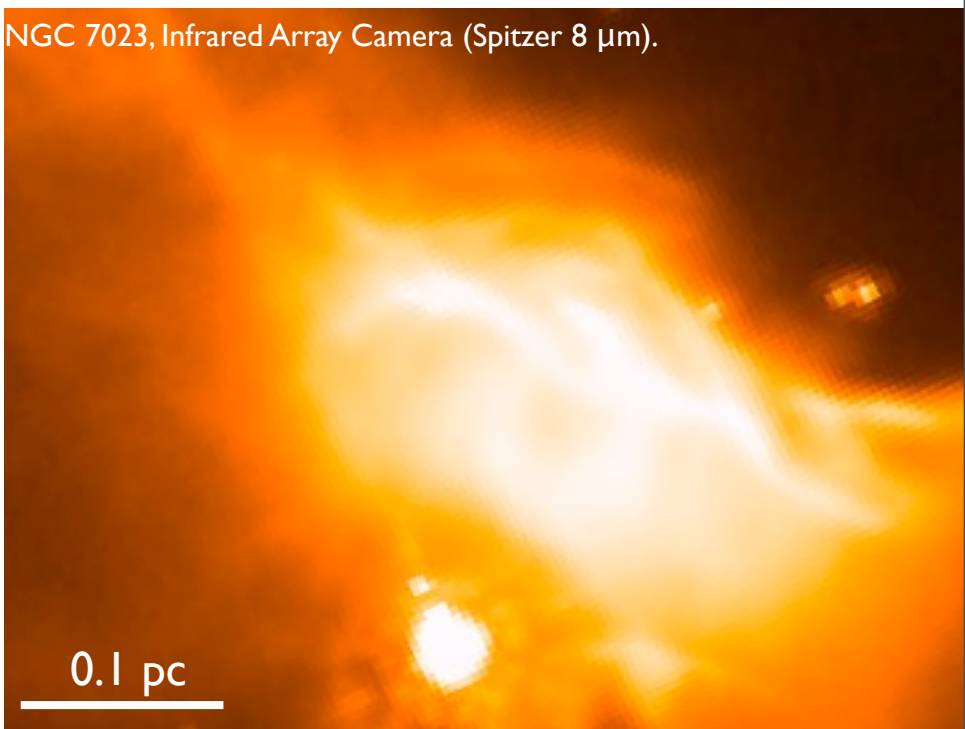
Some examples

[Peeters et al. 2002, Van Didenhoven 2004, Sloan et al. 2007]

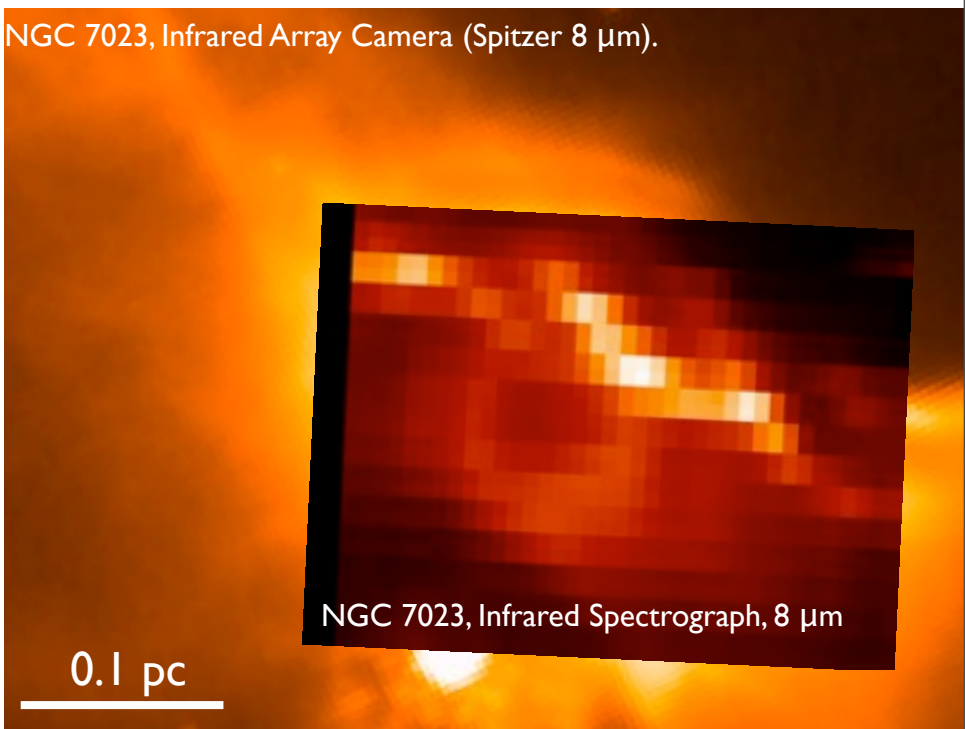
Life of stars $M<8M_{\odot}$



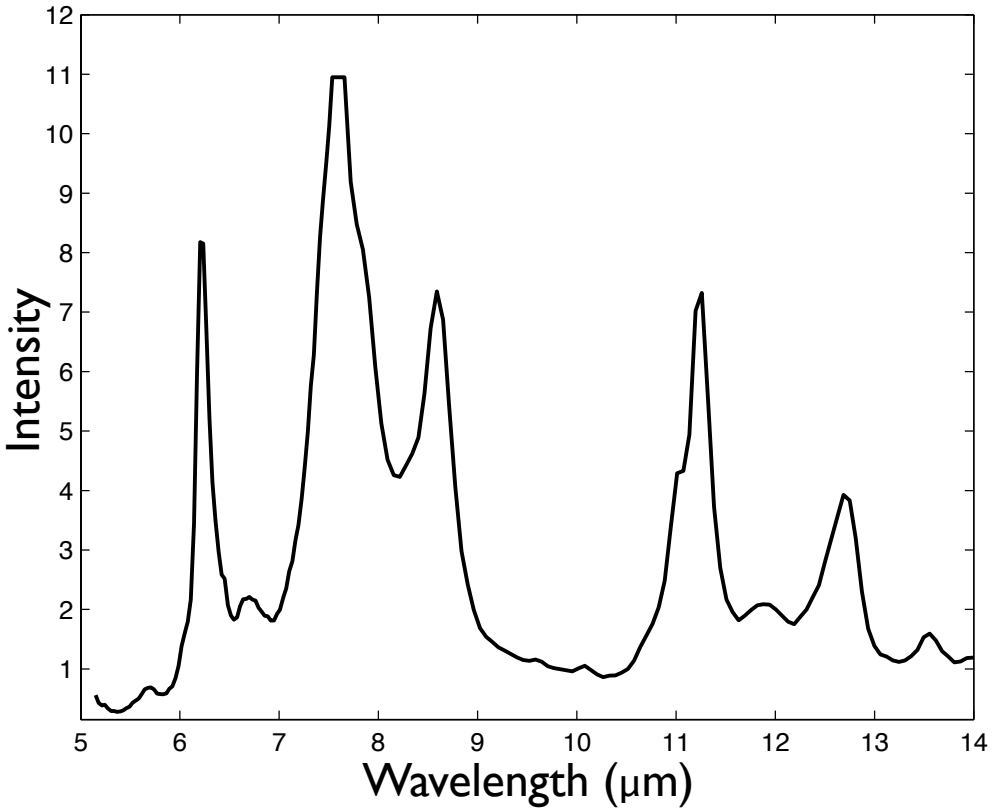
Using spectral mapping of Photodissociation regions (PDRs)

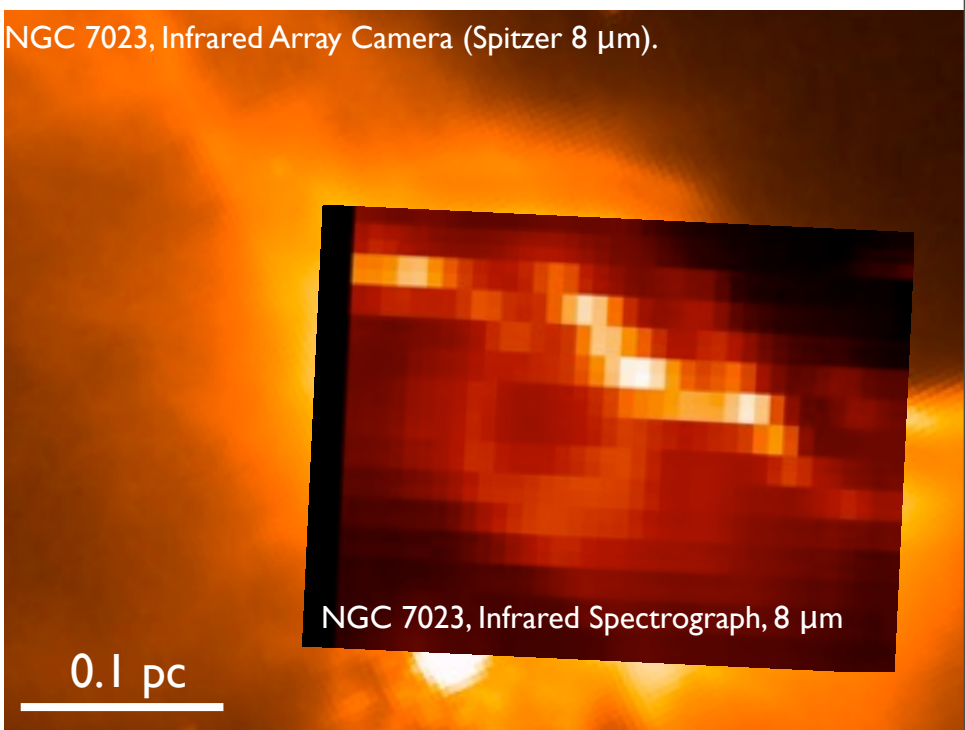


Using spectral mapping of Photodissociation regions (PDRs)

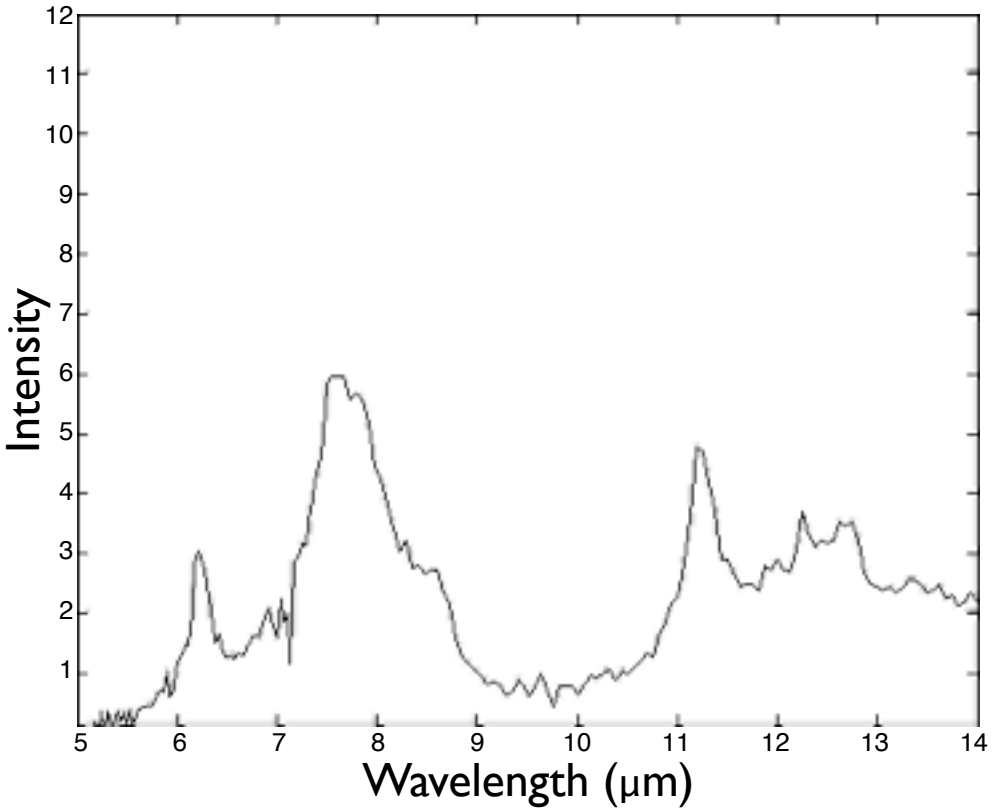


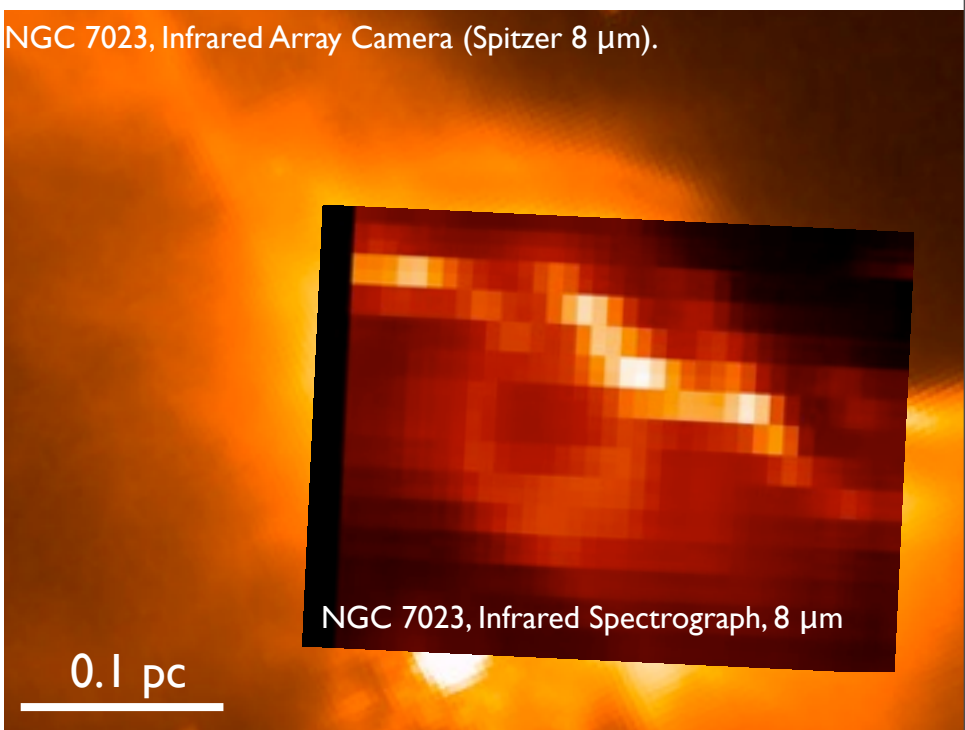
Using spectral mapping of Photodissociation regions (PDRs)



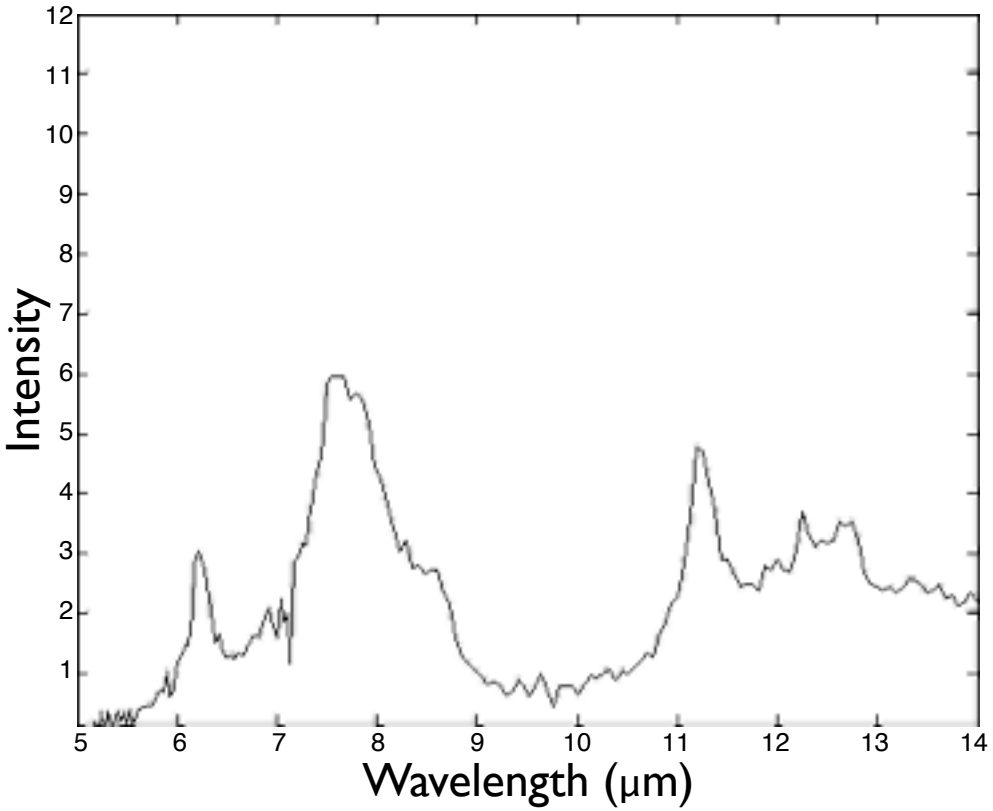


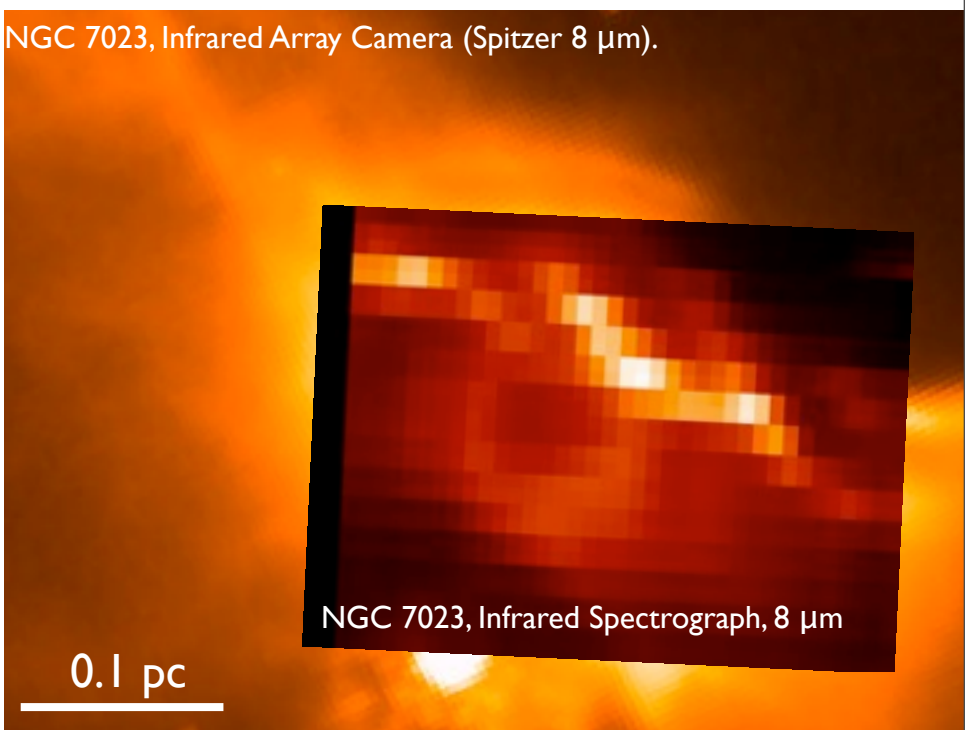
Using spectral mapping of Photodissociation regions (PDRs)





Using spectral mapping of Photodissociation regions (PDRs)





- Each observed spectrum is a linear combination of elementary spectra
- We observe different mixtures of the same elementary spectra

The evolution of spectral shape seen as a linear combination of a limited number of spectra

$$x_i(\lambda) = a_{i1} \times S_1(\lambda) + a_{i2} \times S_2(\lambda) + a_{i3} \times S_3(\lambda) \ldots$$

$$\begin{pmatrix} x_i(\lambda) \\ \vdots \\ x_m(\lambda) \end{pmatrix} = \begin{pmatrix} a_{1,1} & \ldots & a_{1,r} \\ \vdots & & \vdots \\ a_{m,1} & \ldots & a_{m,r} \end{pmatrix} \times \begin{pmatrix} S_1(\lambda) \\ \vdots \\ S_m(\lambda) \end{pmatrix}$$

The evolution of spectral shape seen as a linear combination of a limited number of spectra

$$x_i(\lambda) = a_{i1} \times S_1(\lambda) + a_{i2} \times S_2(\lambda) + a_{i3} \times S_3(\lambda) \dots$$

$$\begin{pmatrix} x_i(\lambda) \\ \vdots \\ x_m(\lambda) \end{pmatrix} = \begin{pmatrix} a_{1,1} & \dots & a_{1,r} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,r} \end{pmatrix} \times \begin{pmatrix} S_1(\lambda) \\ \vdots \\ S_m(\lambda) \end{pmatrix}$$

Observation matrix

The evolution of spectral shape seen as a linear combination of a limited number of spectra

$$x_i(\lambda) = a_{i1} \times S_1(\lambda) + a_{i2} \times S_2(\lambda) + a_{i3} \times S_3(\lambda) \dots$$

$$\begin{pmatrix} x_i(\lambda) \\ \vdots \\ x_m(\lambda) \end{pmatrix} = \begin{pmatrix} a_{1,1} & \dots & a_{1,r} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,r} \end{pmatrix} \times \begin{pmatrix} S_1(\lambda) \\ \vdots \\ S_m(\lambda) \end{pmatrix}$$

Observation matrix

Mixing matrix

The evolution of spectral shape seen as a linear combination of a limited number of spectra

$$x_i(\lambda) = a_{i1} \times S_1(\lambda) + a_{i2} \times S_2(\lambda) + a_{i3} \times S_3(\lambda) \ldots$$

$$\begin{pmatrix} x_i(\lambda) \\ \vdots \\ x_m(\lambda) \end{pmatrix} = \begin{pmatrix} a_{1,1} & \ldots & a_{1,r} \\ \vdots & & \vdots \\ a_{m,1} & \ldots & a_{m,r} \end{pmatrix} \times \begin{pmatrix} S_1(\lambda) \\ \vdots \\ S_m(\lambda) \end{pmatrix}$$

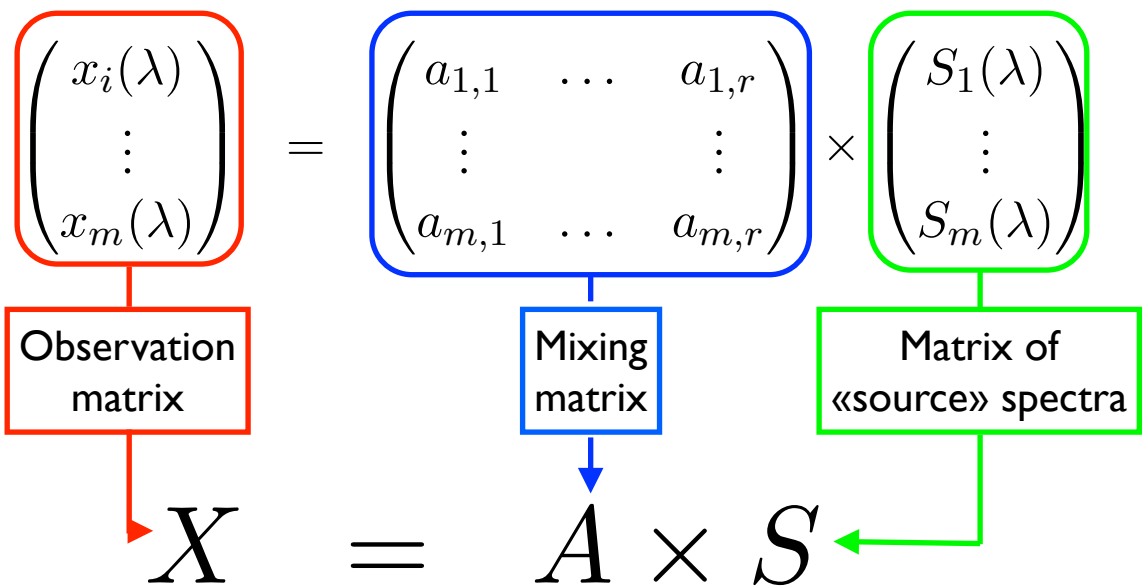
Observation matrix

Mixing matrix

Matrix of «source» spectra

The evolution of spectral shape seen as a linear combination of a limited number of spectra

$$x_i(\lambda) = a_{i1} \times S_1(\lambda) + a_{i2} \times S_2(\lambda) + a_{i3} \times S_3(\lambda) \dots$$



The evolution of spectral shape seen as a linear combination of a limited number of spectra

$$x_i(\lambda) = a_{i1} \times S_1(\lambda) + a_{i2} \times S_2(\lambda) + a_{i3} \times S_3(\lambda) \ldots$$

$$\begin{pmatrix} x_i(\lambda) \\ \vdots \\ x_m(\lambda) \end{pmatrix} = \begin{pmatrix} a_{1,1} & \ldots & a_{1,r} \\ \vdots & & \vdots \\ a_{m,1} & \ldots & a_{m,r} \end{pmatrix} \times \begin{pmatrix} S_1(\lambda) \\ \vdots \\ S_m(\lambda) \end{pmatrix}$$

Observation matrix — Mixing matrix — Matrix of «source» spectra

$$X = A \times S$$

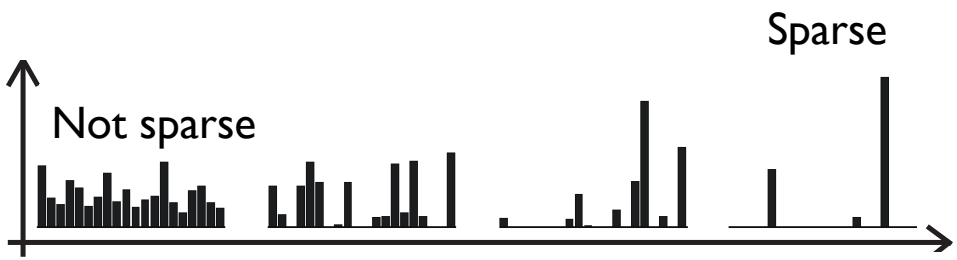
Goal identifying A *and* S, from X

Blind signal separation (BSS)

- 3 classes of methods:

  - Methods based on independence between the *sources* (Independent Component Analysis, ICA)
  - Methods based on positivity of *sources* (non-Negative Matrix Factorization)
  - Methods based on sparsity of *sources* (Sparse Component Analysis, SCA)

- In our case ?



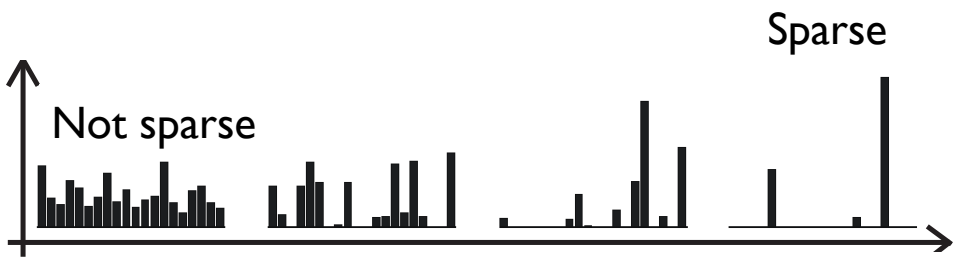
	Issues	Advantages
ICA	Sources slightly correlated	Unique solution
NMF	Sources must be sparse for unique solution	Sources are positive
SCA	Need to find a «space» in which source are sparse	No independence required

Blind signal separation (BSS)

☆ 3 classes of methods:

- Methods based on independence between the *sources* (Independent Component Analysis, ICA)
- Methods based on positivity of *sources* (non-Negative Matrix Factorization)
- Methods based on sparsity of *sources* (Sparse Component Analysis, SCA)

☆ In our case ?



	Issues	Advantages
ICA	Sources slightly correlated	Unique solution
NMF	Sources must be sparse for unique solution	Sources are positive
SCA	Need to find a «space» in which source are sparse	No independence required

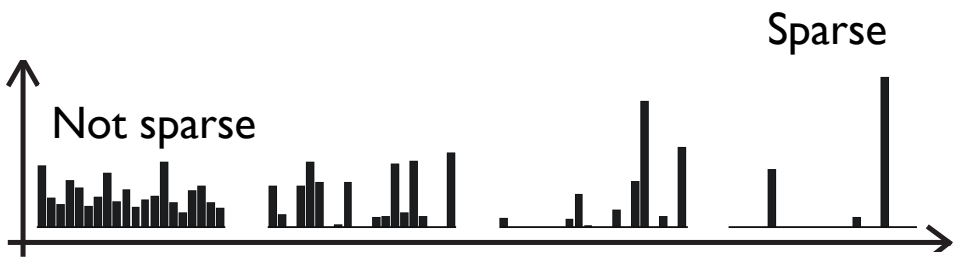
Similar results on real and artificial data

Blind signal separation (BSS)

☆ 3 classes of methods:

- Methods based on independence between the *sources* (Independent Component Analysis, ICA)
- Methods based on positivity of *sources* (non-Negative Matrix Factorization)
- Methods based on sparsity of *sources* (Sparse Component Analysis, SCA)

☆ In our case ?



	Issues	Advantages
ICA	Sources slightly correlated	Unique solution
NMF	Sources must be sparse for unique solution	Sources are positive
SCA	Need to find a «space» in which source are sparse	No independence required

Similar results on real and artificial data

Non-Negative Matrix Factorization NMF

The problem

$$X = A \times S$$

[Lee and Seung, Nature, 2001]

Non-Negative Matrix Factorization NMF

The problem

$$X = A \times S$$

Approach:

$$X \approx W \times H$$

[Lee and Seung, Nature, 2001]

Non-Negative Matrix Factorization NMF

The problem

$$X = A \times S$$

Approach:

$$X \approx W \times H$$

The criteria, Euclidian distance / Kullback's divergence :

$$\|X - WH\|^2 = \sum_{ij} (X_{ij} - (WH)_{ij})^2$$

$$D(X|WH) = \sum_{ij} (X_{ij} \log \frac{X_{ij}}{(WH)_{ij}} - X_{ij} + (WH)_{ij})$$

[Lee and Seung, Nature, 2001]

Non-Negative Matrix Factorization NMF

The problem

$$X = A \times S$$

Approach:

$$X \approx W \times H$$

The criteria, Euclidian distance / Kullback's divergence :

$$\|X - WH\|^2 = \sum_{ij} (X_{ij} - (WH)_{ij})^2$$

$$D(X|WH) = \sum_{ij} (X_{ij} \log \frac{X_{ij}}{(WH)_{ij}} - X_{ij} + (WH)_{ij})$$

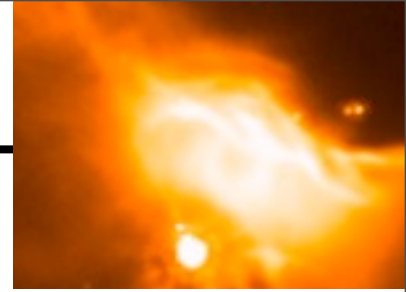
The algorithm

$$H_{a\mu} \leftarrow H_{a\mu} \frac{\sum_i W_{ia} X_{i\mu} / (WH)_{i\mu}}{\sum_k W_{ka}}, W_{ia} \leftarrow W_{ia} \frac{\sum_\mu H_{a\mu} X_{i\mu} / (WH)_{i\mu}}{\sum_\nu H_{a\nu}}$$

- We set the numbers or rows of **H**
- **W** and **H** must be positive
- We start iteration with random **W** and **H**
- Results do not depend on initialization

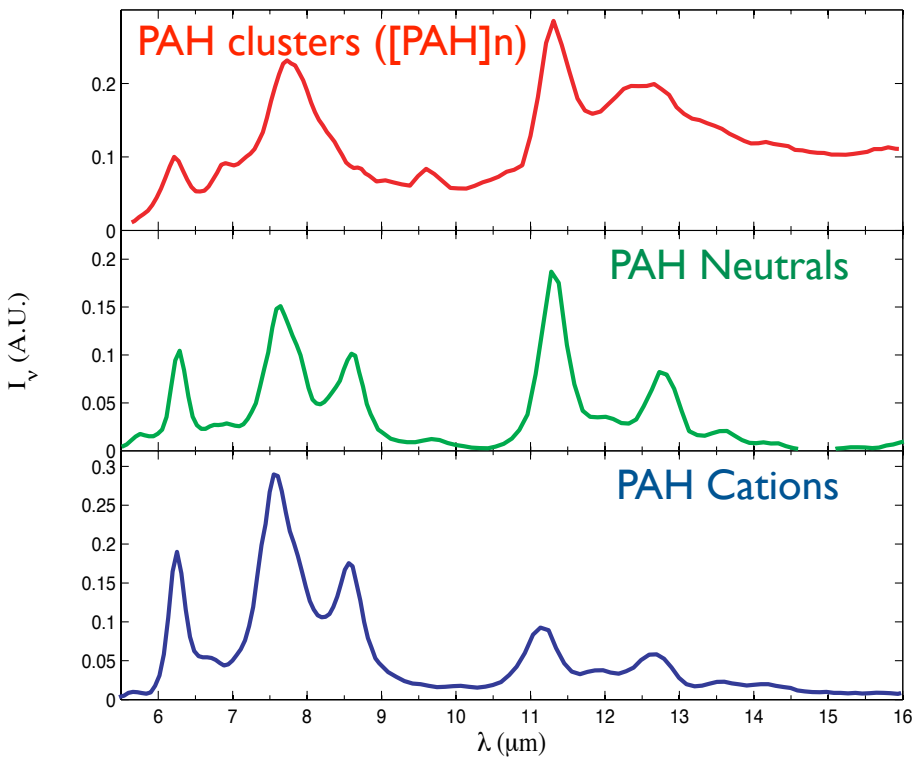
[Lee and Seung, Nature, 2001]

At the origin of spectral variations of PAHs

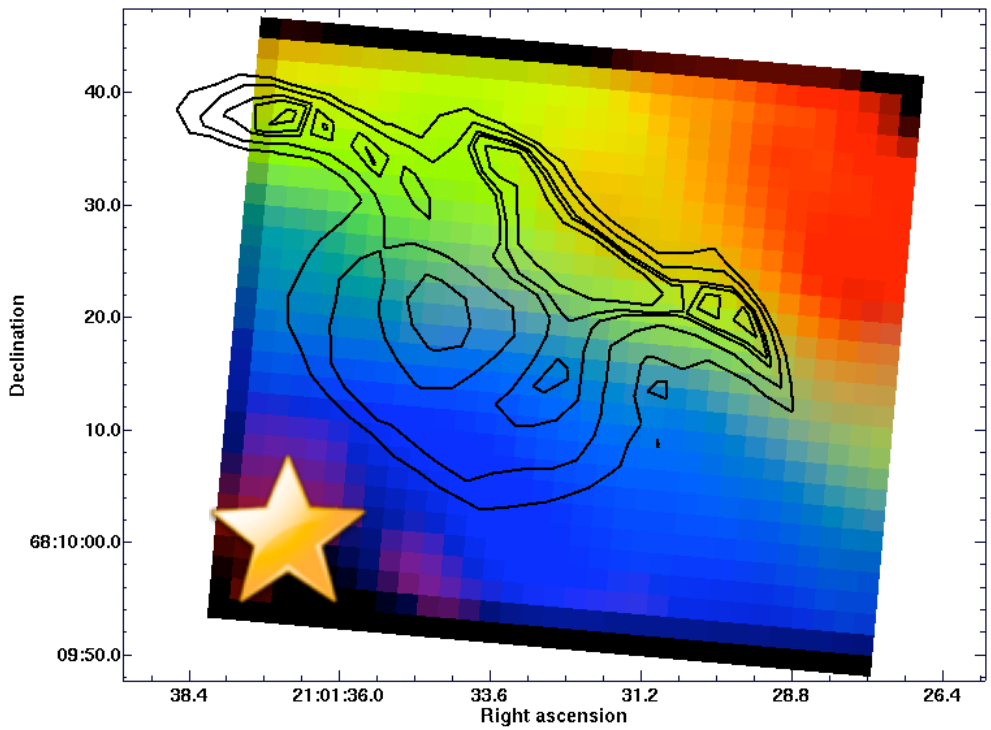


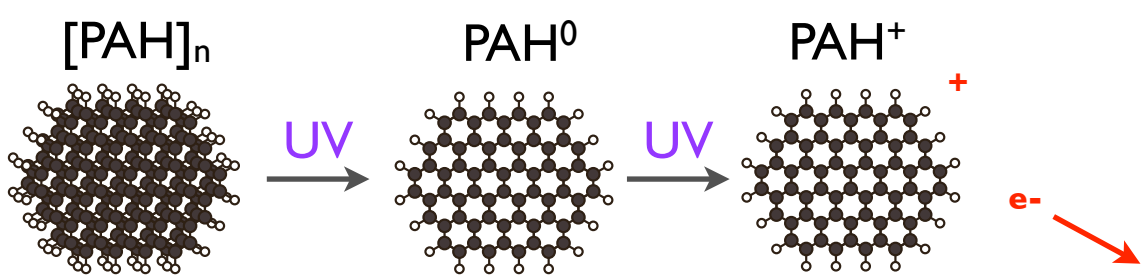
[Rapacioli et al. *Astron. Astrophys.* 2005, Berné et al. *Astron. Astrophys.* 2007]

Extracted spectra H



Spatial distribution W

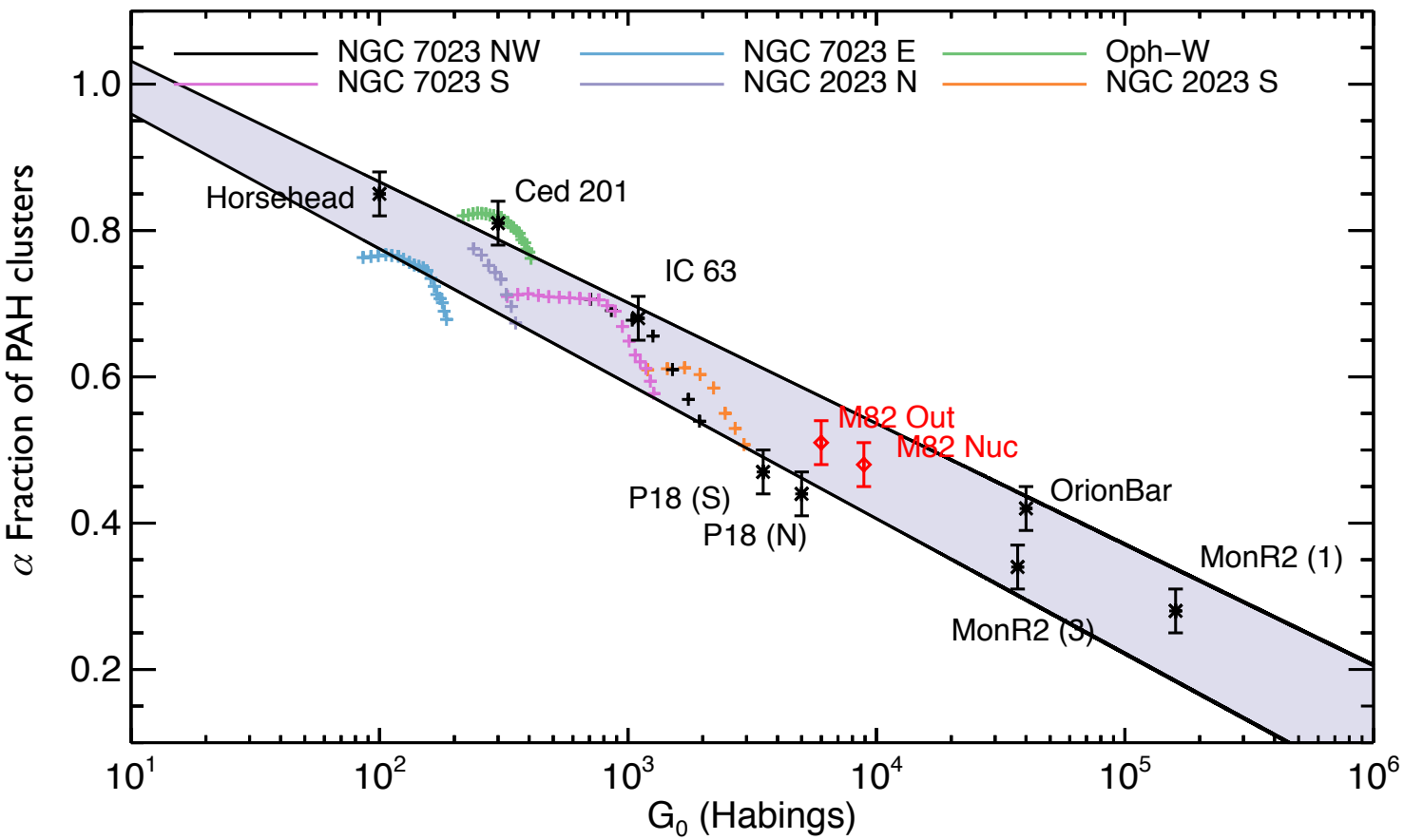




Astrophysical implications

[Pilleri et al. 2012 A&A]

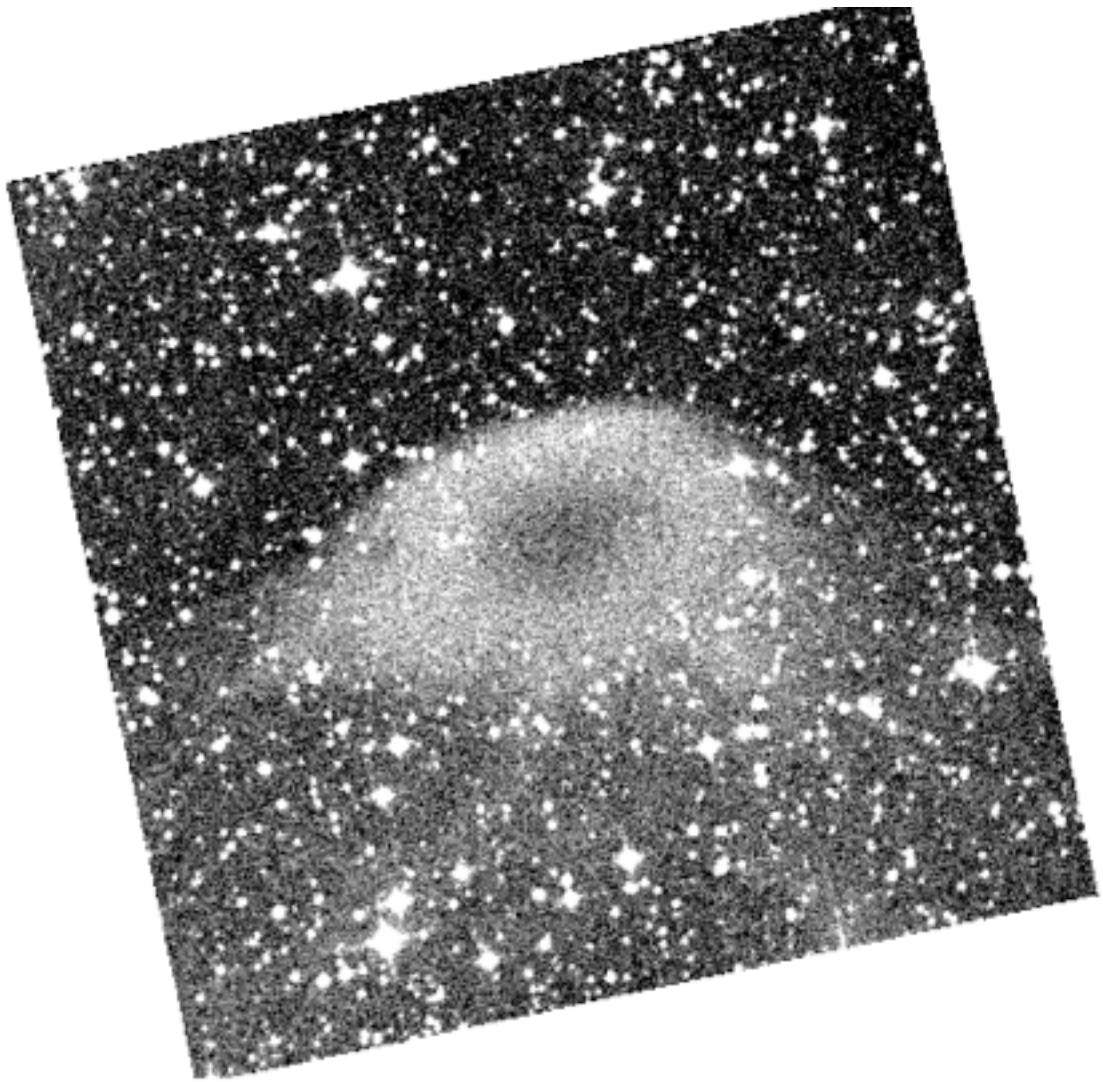
Based on the BSS results a mid-IR emission model was constructed which relates radiation field to abundance of PAH clusters



Go: intensity of UV radiation field emitted by stars

Hyperspectral radioastronomy

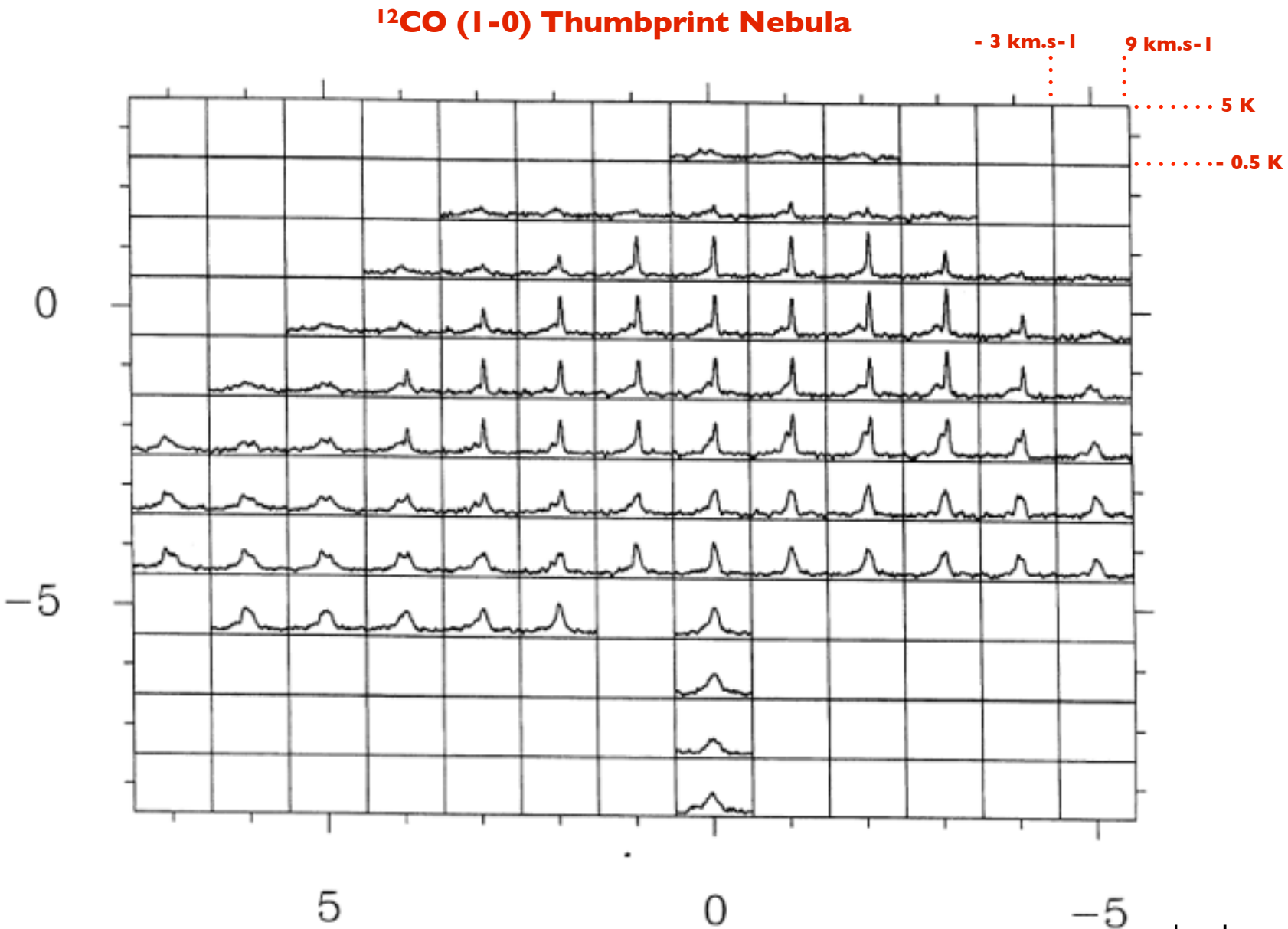
- High resolution spectral maps of line emission from atoms/molecules



[Juvela et al. 1996]

Hyperspectral radioastronomy

- High resolution spectral maps of line emission from atoms/molecules



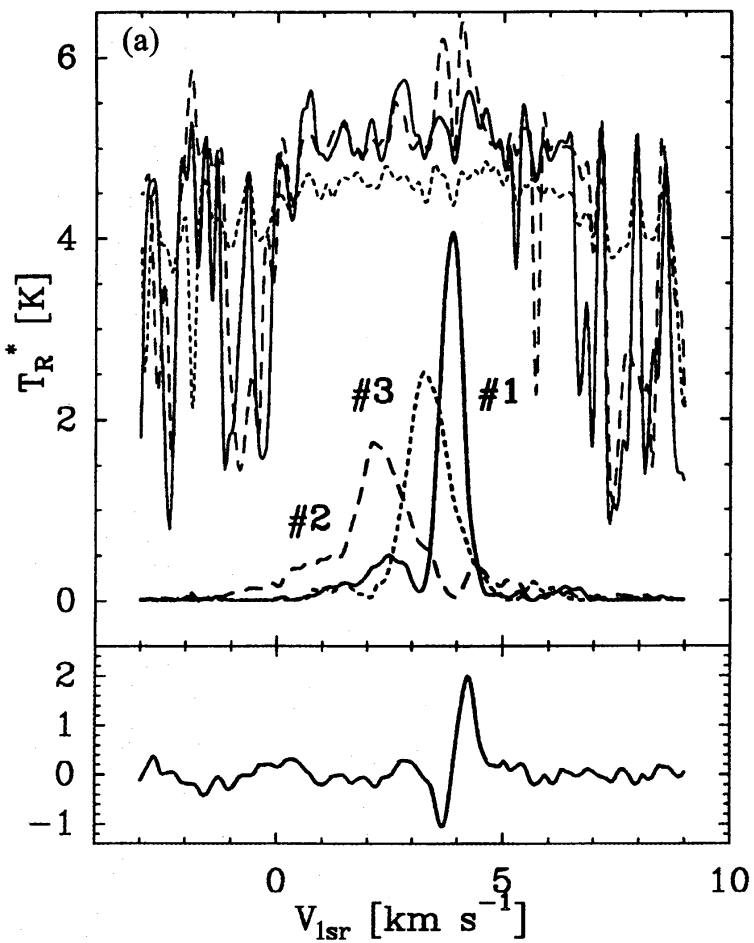
[Juvela et al. 1996]

Results of BSS method on this data

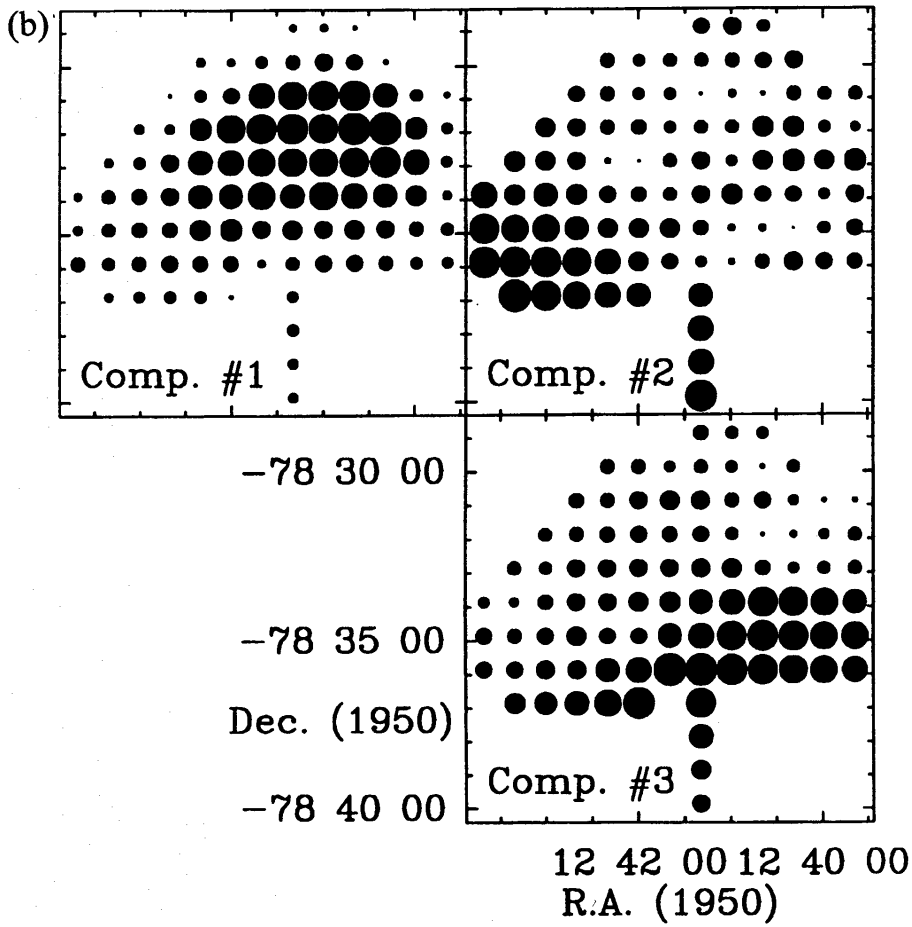
- Decomposition of the set into «velocity components»

^{12}CO (1-0) Thumbprint Nebula

Extracted spectra H



Spatial distribution W

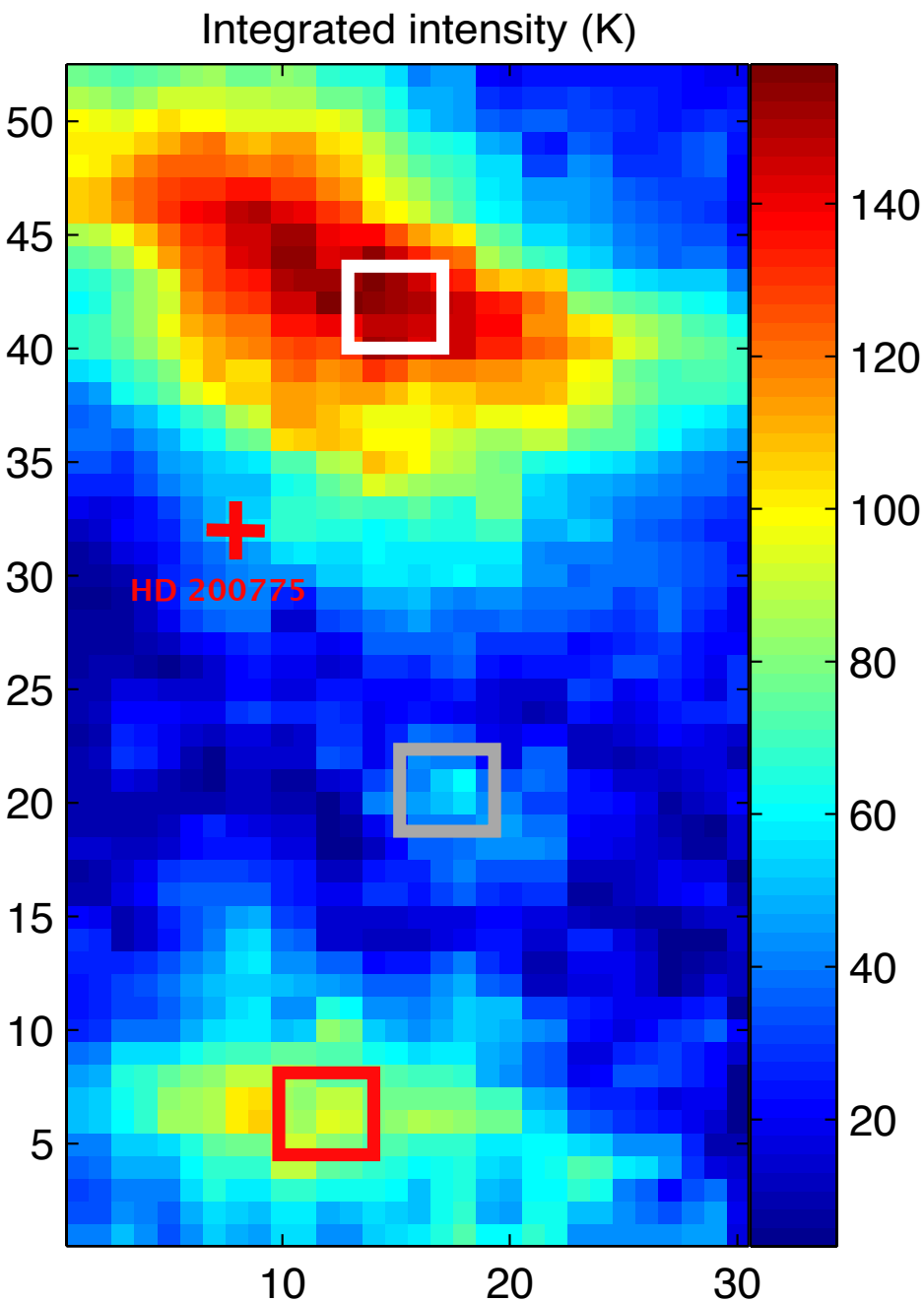


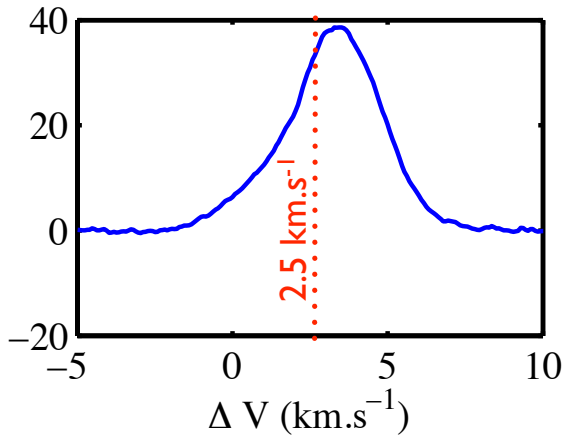
[Juvela et al. 1996]

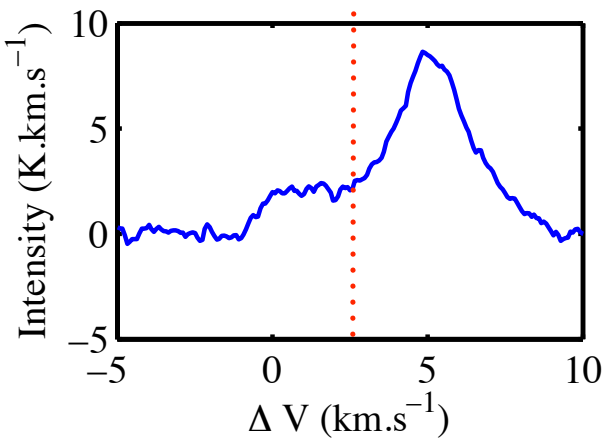
Herschel-HIFI mapping (OT1_cjoblin_1)

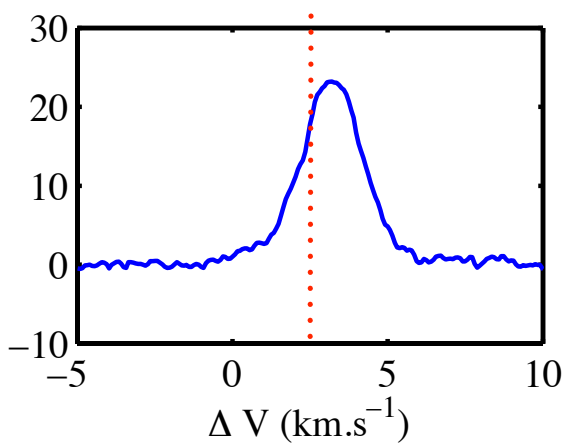
[CII] 158 μm 0.08 km.s^{-1} / channel (12.5" beam)



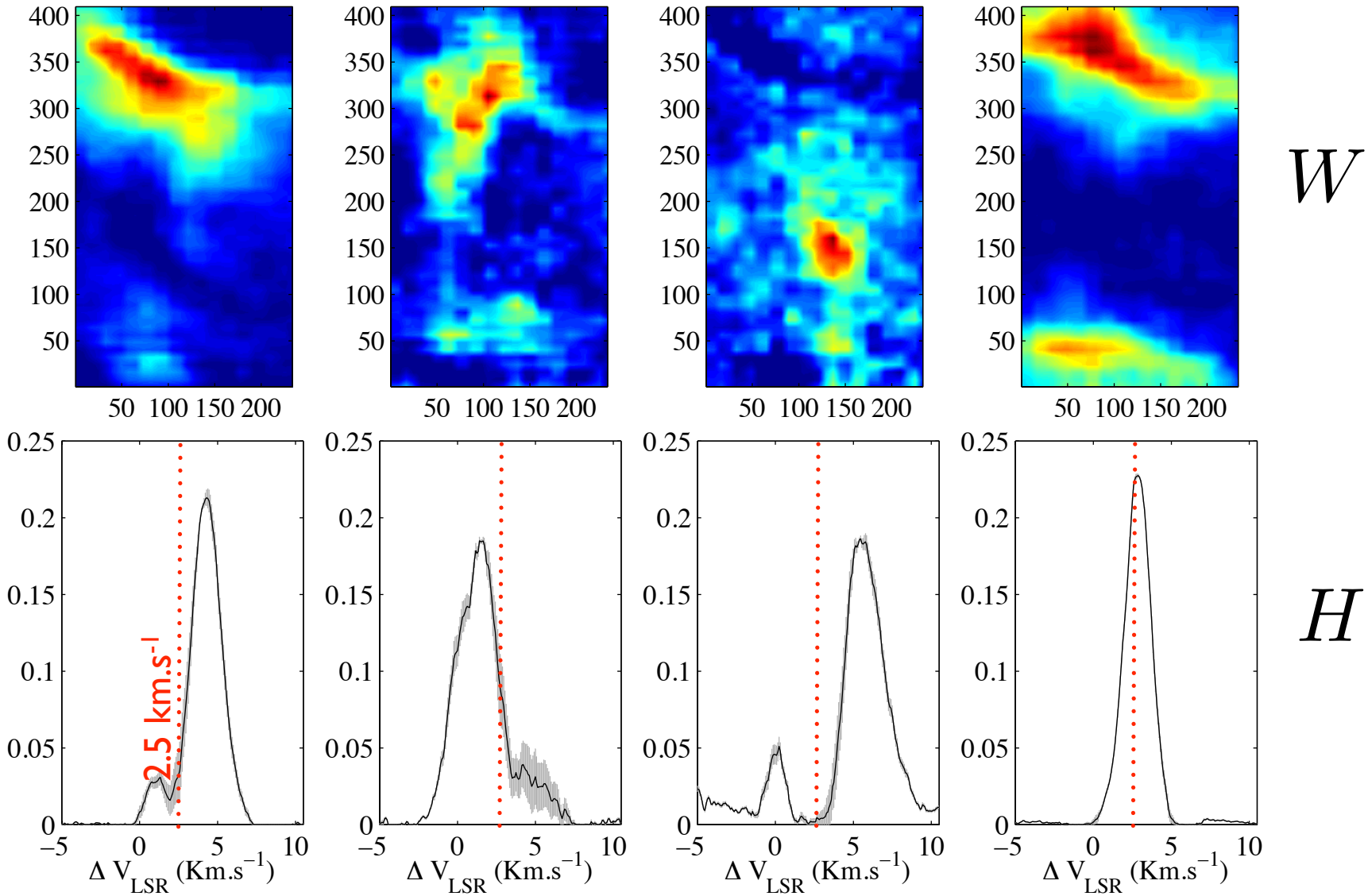








Herschel-HIFI mapping: application of NMF



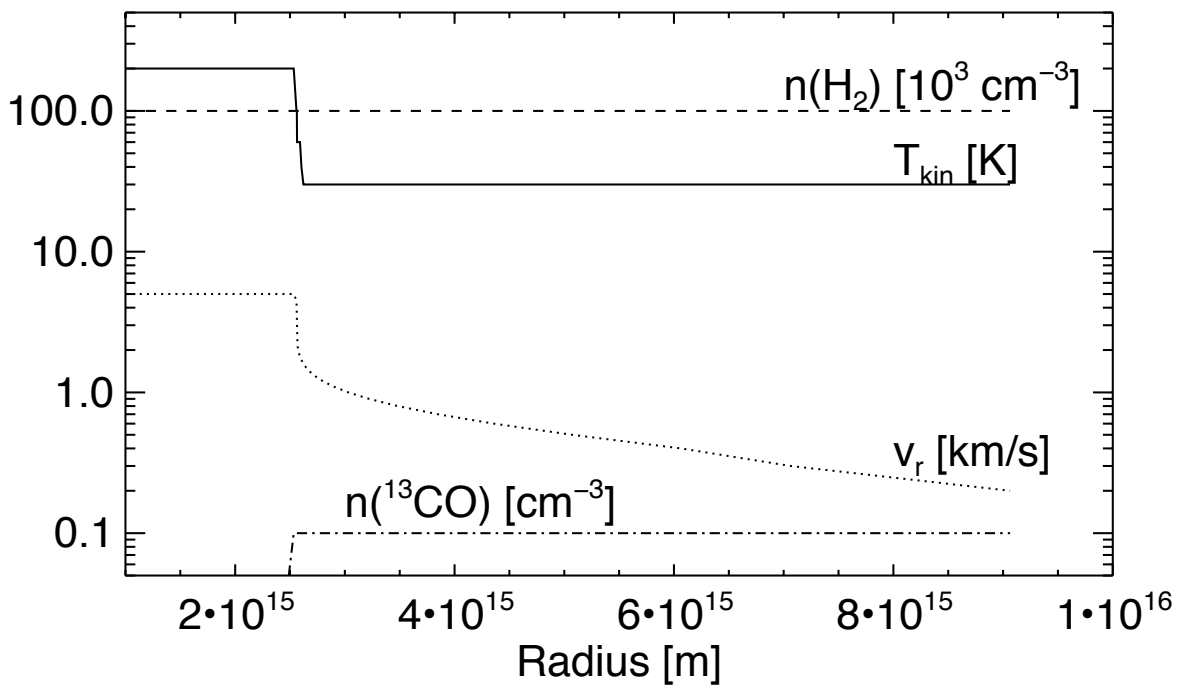
- Upgraded NMF using Monte Carlo and clustering to estimate error-bars
- We find different, velocity components
- Convergence is stable (no dependance on initial conditions) i.e. unique minimum

Testing on a spherical model

- Is the mixture linear ?
- Are the radiative transfer effects ? cf. Nuzillard & Bijaoui (2001)

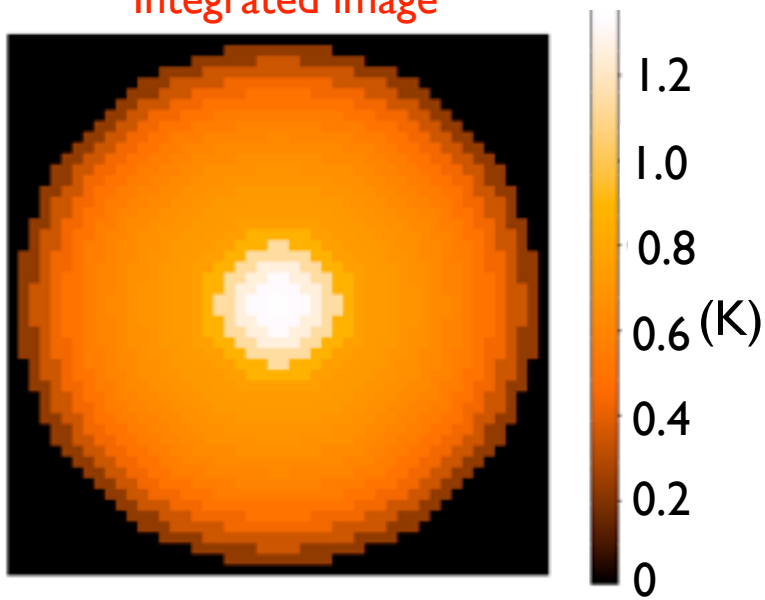
Model properties

Expanding shell e.g. massive star-forming region

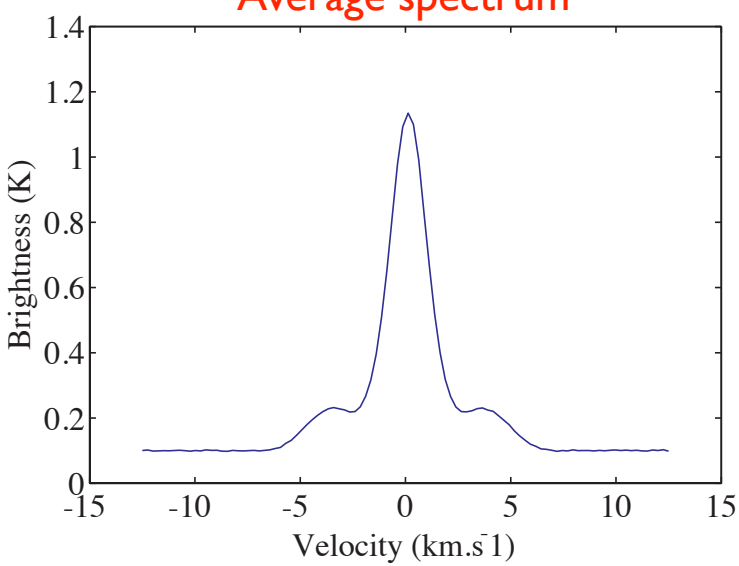


Model results

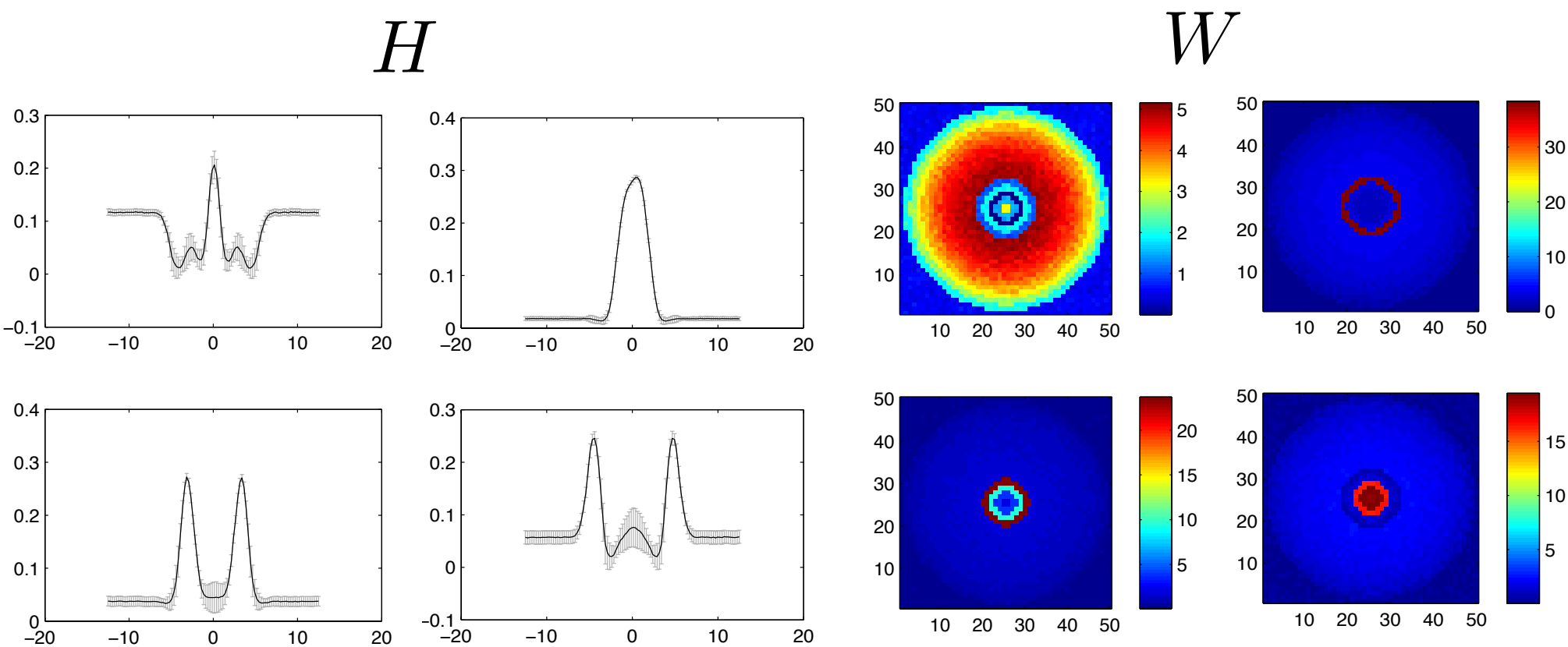
Integrated image



Average spectrum



Results of NMF on a spherical model



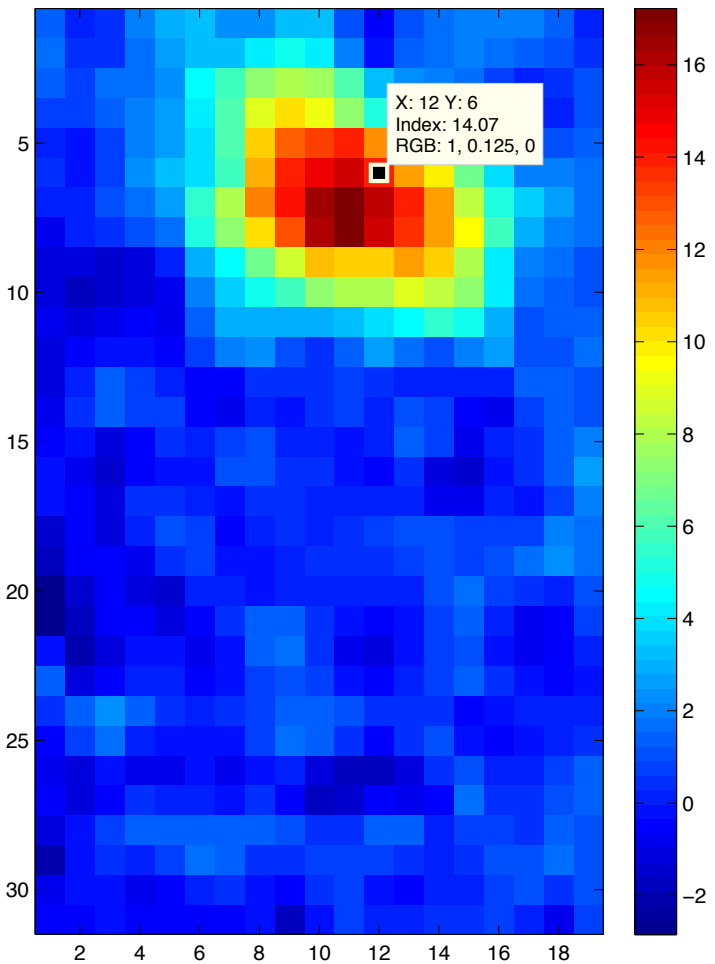
- We find different, symetric, velocity components
- Convergence is unstable (depends on initial conditions) i.e. no unique minimum
- In the model case at least BSS is not truly applicable

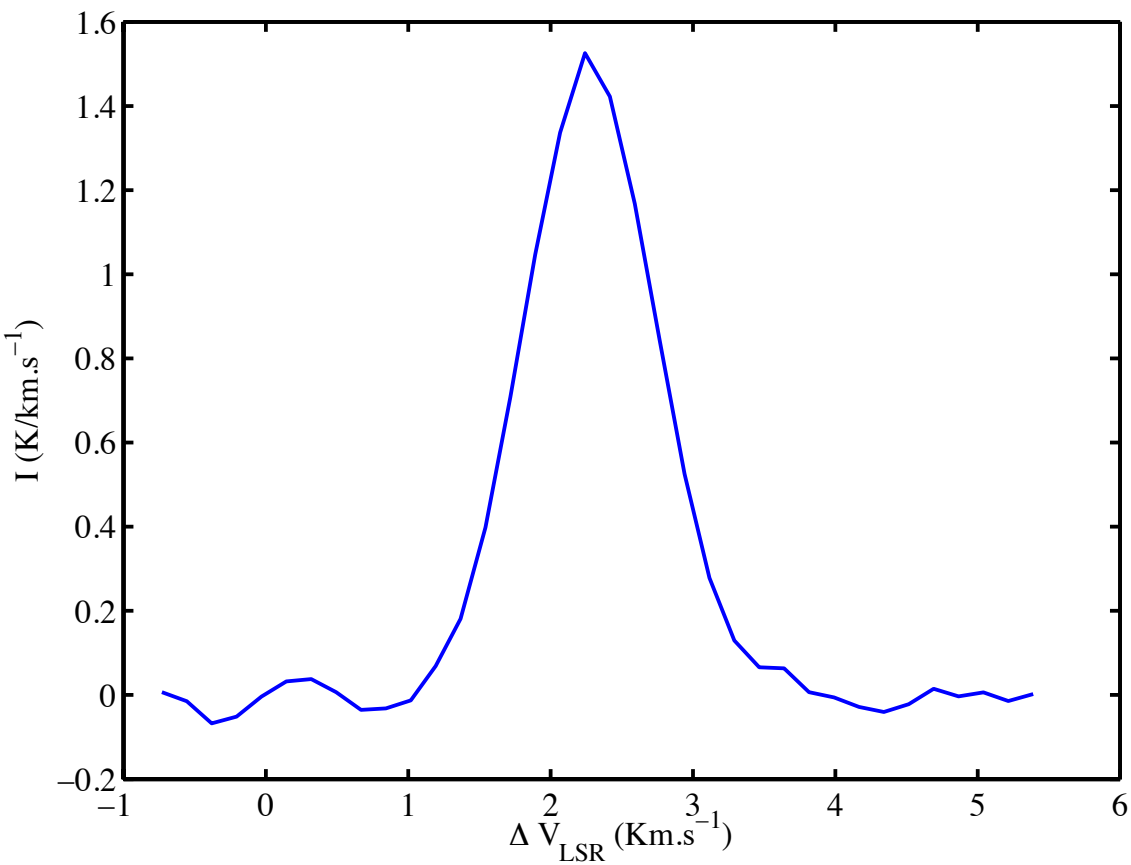
Conclusions

- In the mid-IR the mixing is linear and BSS provide the elementary spectra which represent distinct chemical populations
- In the radio-mm domain, it is unclear wether the mixing model remains linear, physical interpretation may be tricky
- BSS methods for radio-mm data can help «viewing» large datat-sets, e.g. this can be used for ALMA
- An additional use: systematic de-noising

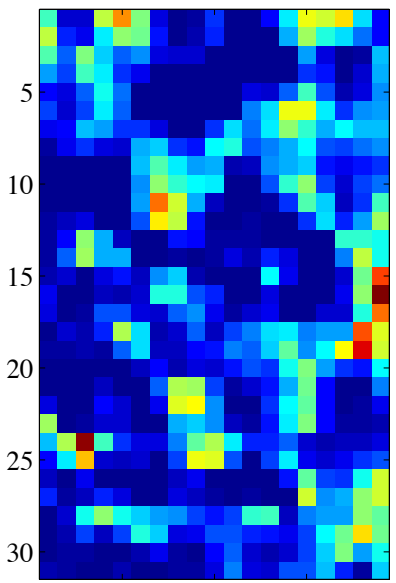
Thanks to PCMI

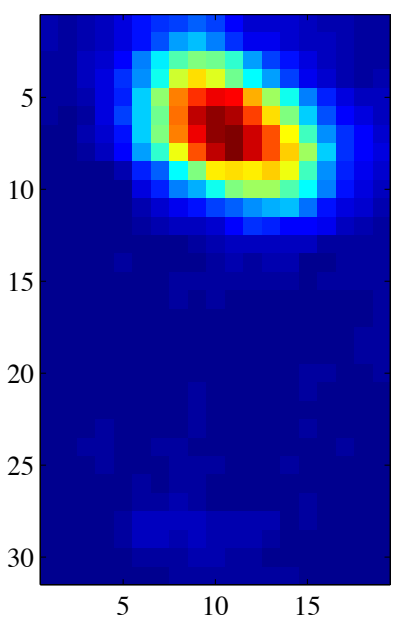
13CO(8-7) Maps

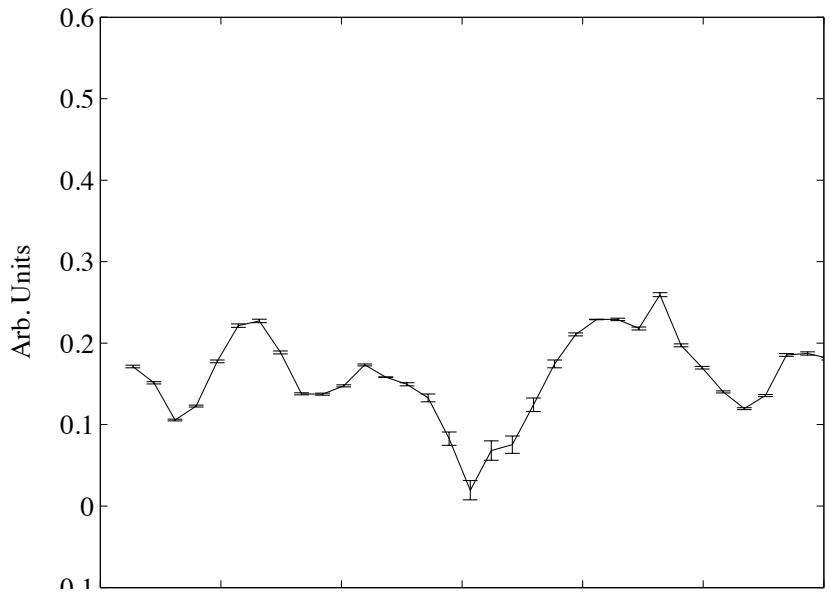

X: 12 Y: 6
Index: 14.07
RGB: 1, 0.125, 0
5
10
15
20
25
30
2
4
6
8
10
12
14
16
18
16
14
12
10
8
6
4
2
0
−2

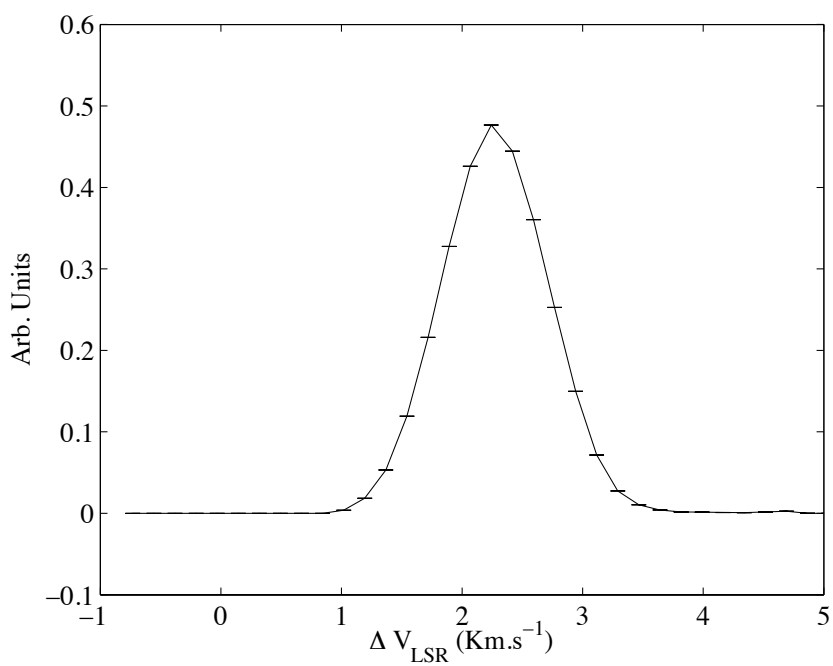

1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
−0.2
−1
0
1
2
3
4
5
6
I (K/km.s^{-1})
Δ V$_{LSR}$ (Km.s^{-1})

13CO(8-7) Results









Analysis with NMF

MC + clustering

Linear instantaneous model (optically thin)

$$x(p_x, p_y, v) = \sum_{i=1}^{r} a^i(p_x, p_y) s^i(v) + n(p_x, p_y, v)$$

$$X = AS + N$$

✦ Method (NMF)

- Linear mixing + noise: $X = A \times S + N$
- Approximation: $X \approx W \times H$
- Minimization of divergence using Lee & Seung

criterion $\delta(X|WH) = \sum_{ij} (X^{ij} \log \frac{X^{ij}}{(WH)^{ij}} - X^{ij} + (WH)^{ij})$

algorithm $H \leftarrow H \frac{\sum_i W^i X^i / (WH)^i}{\sum_j W^j} W \leftarrow W \frac{\sum_i H^i X^i / (WH)^i}{\sum_j H^j}$

✦ Estimation of number of sources

- We use a criterion based on noise: $\|\hat{X} - X\|_F < \|N\|_F$ (with $\hat{X} = W \times H$)
- We find that 3 sources spectra satisfy this

Analysis with NMF

MC + clustering

✦ Uncertainty estimation

- We run the algorithm 100 times

- Since NMF provides results with a degeneracy on the order of the 3 source spectra in we use *K-means* to form 3 clusters of 100 spectra minimizing the inter-cluster correlation

$$argmin \sum_{i=1,k} \sum_{h_j \in \sigma_i} 1 - corr(h_j, \mu_i),$$

- The uncertainty on source spectra is given by the variance in the spectra at a given wavelength amongst the 100 spectra in each cluster

$$\epsilon_n(v) = \sqrt{\frac{1}{k_n - 1} \sum_{i=1}^{k_n} (h_i(v) - \bar{h}(v))^2},$$

- The rows of the matrix of source spectra H_* are the mean of the 100 spectra in each of the 3 cluster

✦ Reconstruction of the mixing matrix

- We build the mixing matrix W_* by minimizing $||X - W_* H_*||$ using NNLS