A novel method for component separation of extended sources in X-rays

MPE seminar, October 2022

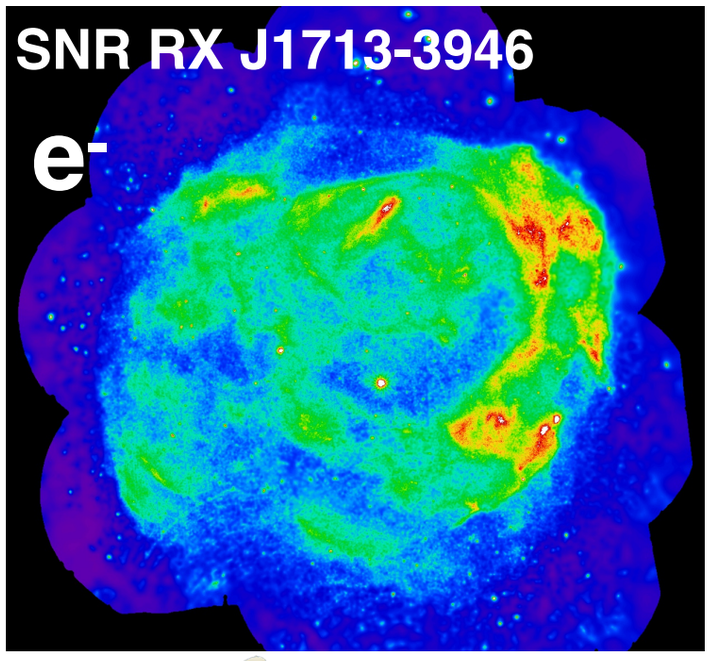
Fabio Acero
A.Picquenot (PhD)
L. Godinaud (PhD)
J. Bobin, J. Ballet, G. Pratt, P. Maggi

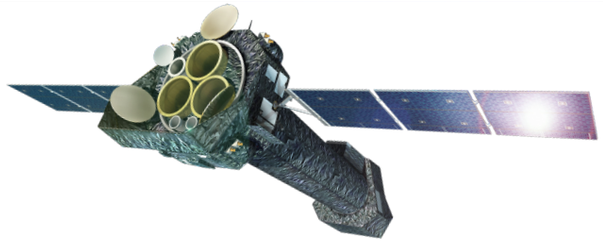
Astronomy department, CEA/Saclay
@fabacero

Chandra SNR gallery

My field of expertise : SNRs & PWNe from keV to TeV

keV

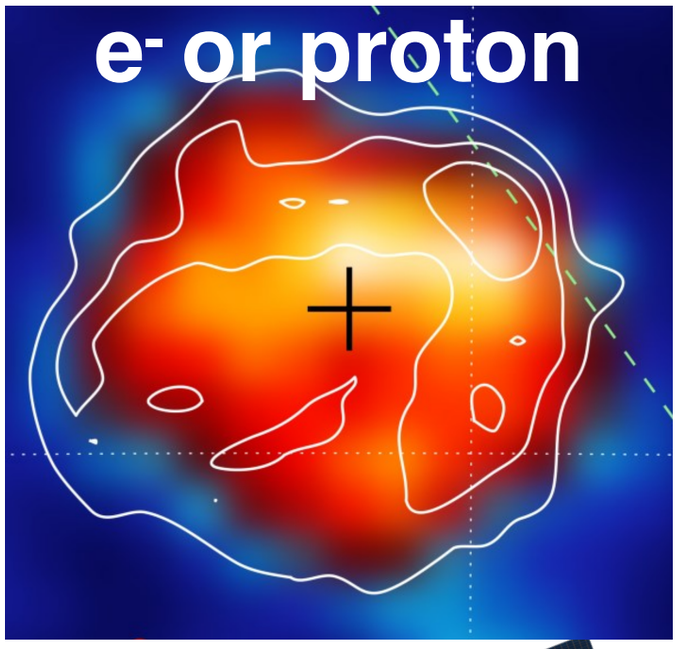




XMM-Newton

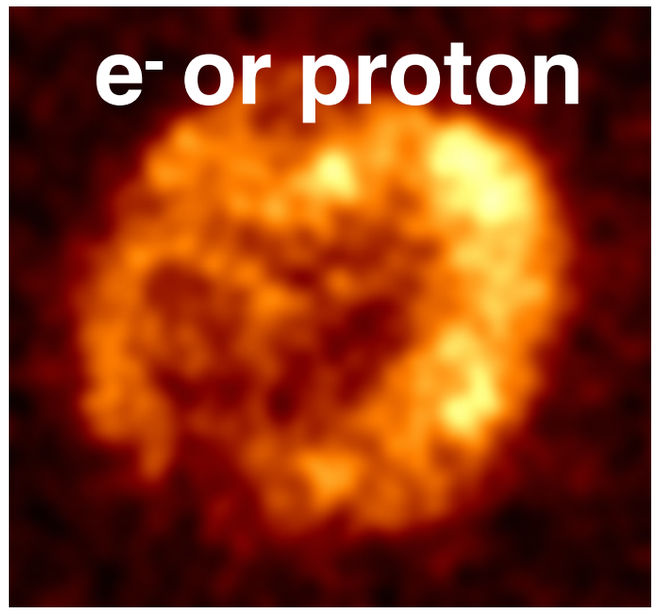
NewAathena

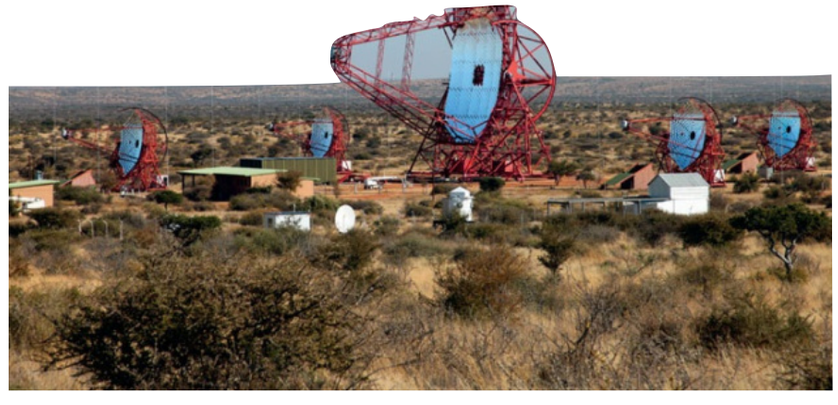
GeV gamma-ray



Fermi

TeV gamma-ray





HESS + **CTA**

Synchrotron

Inverse Compton + Pi0 decay

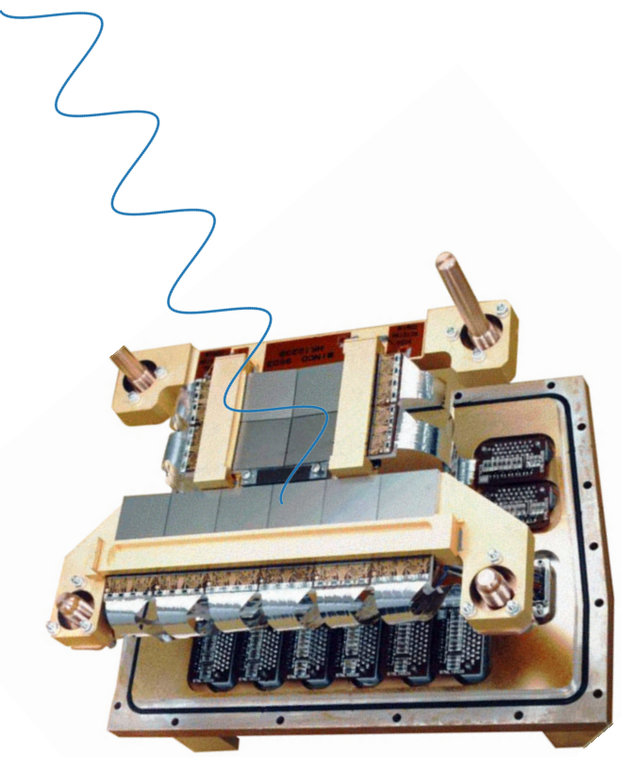
Outline

- **A few words on supernova remnants (SNRs)**
- **Science with spectro-imagers**
- **Challenges in X-ray data analysis & different existing workflows**
- **A new data analysis method using the 3D (X,Y,E) data**
- **Panorama of applications: Cassiopeia A & Tycho SNRs, Perseus cluster**
- **A benchmarking Athena dataset for testing new methods**

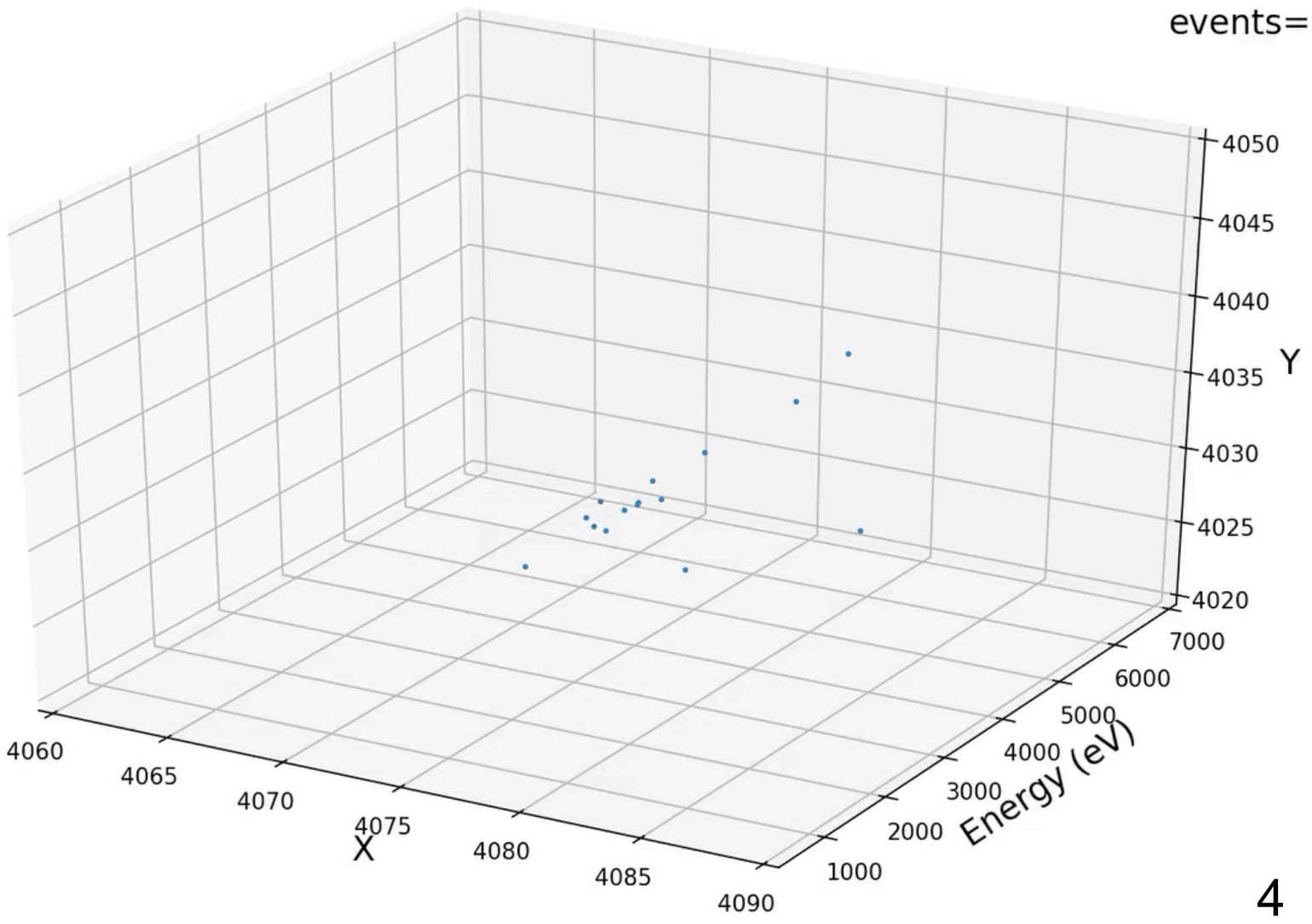
Spectro-imagers in high-energy

- **X-ray spectro-imagers since Einstein (1978) and ROSAT (90)**
- **CCDs revolution with ASCA (93)**

- **Measure photon by photon:**
  - **X, Y, Energy, Time**

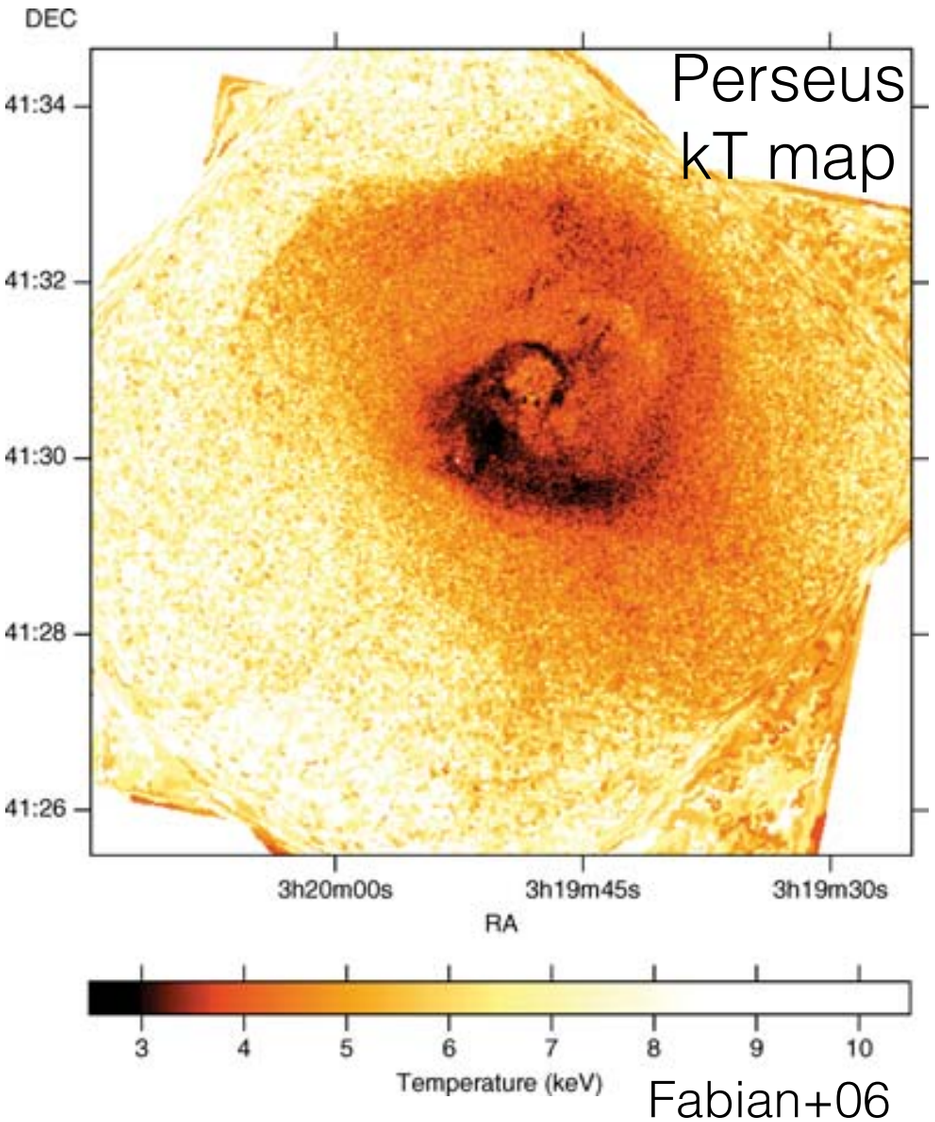


Chandra ACIS

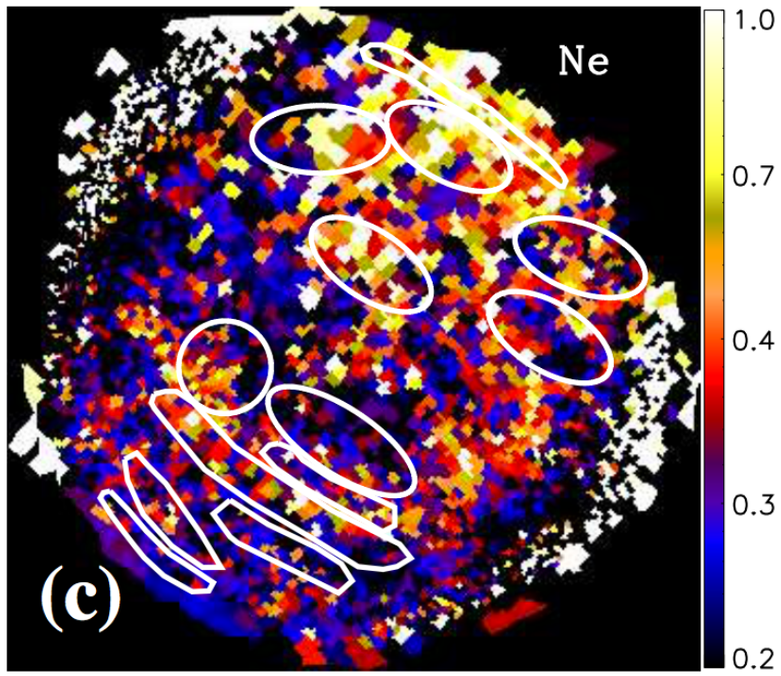


Science with spectro-imagers

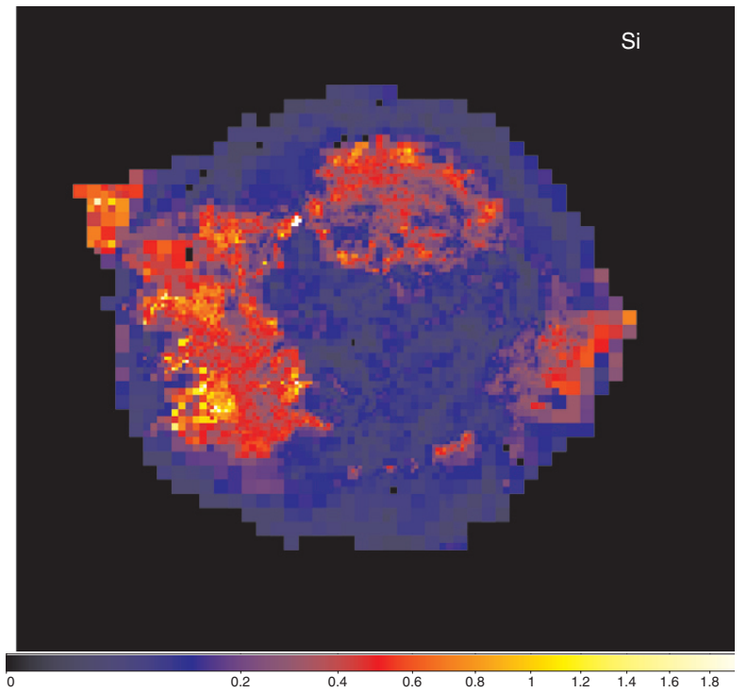
- **Allow us to isolate point sources or features from extended src**
- **Study spatial variations of physical parameters (kT, Index, …) in :**
  - **galaxy clusters, galaxies, ISM, SNRs, PWNe**



Fabian+06

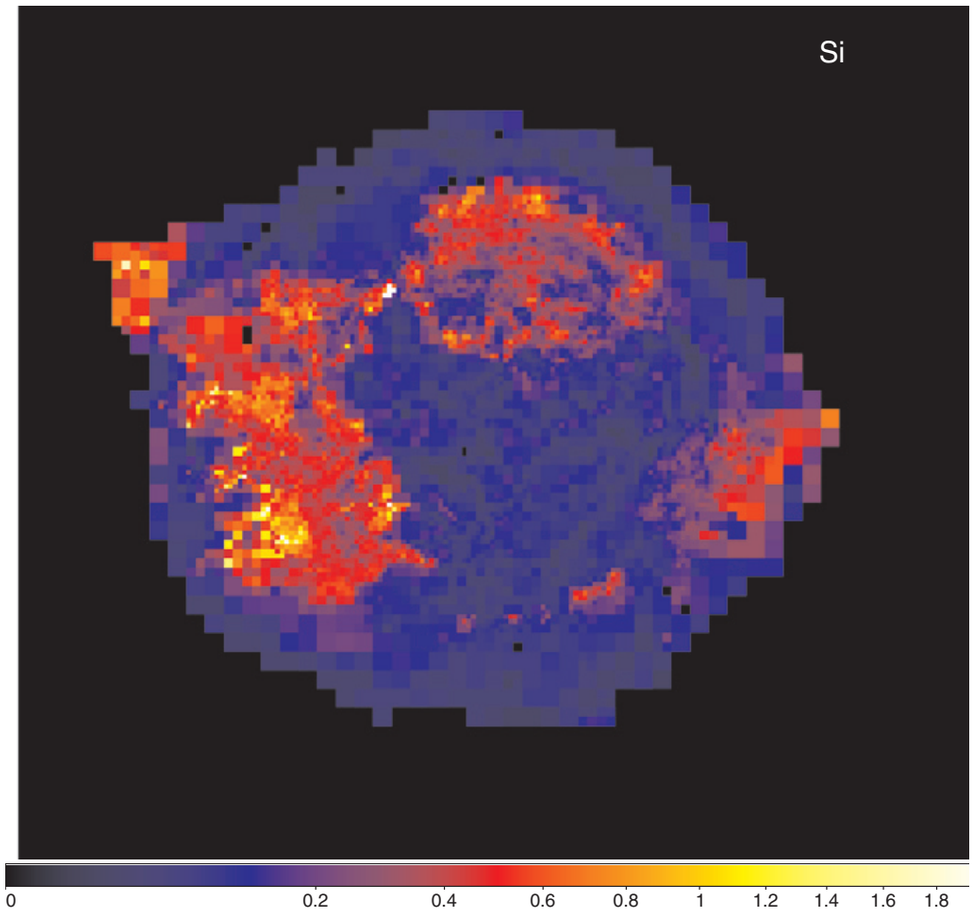


Ne abund SN1006, Li+15



Si abund CasA, Hwang+12

Challenges in data analysis

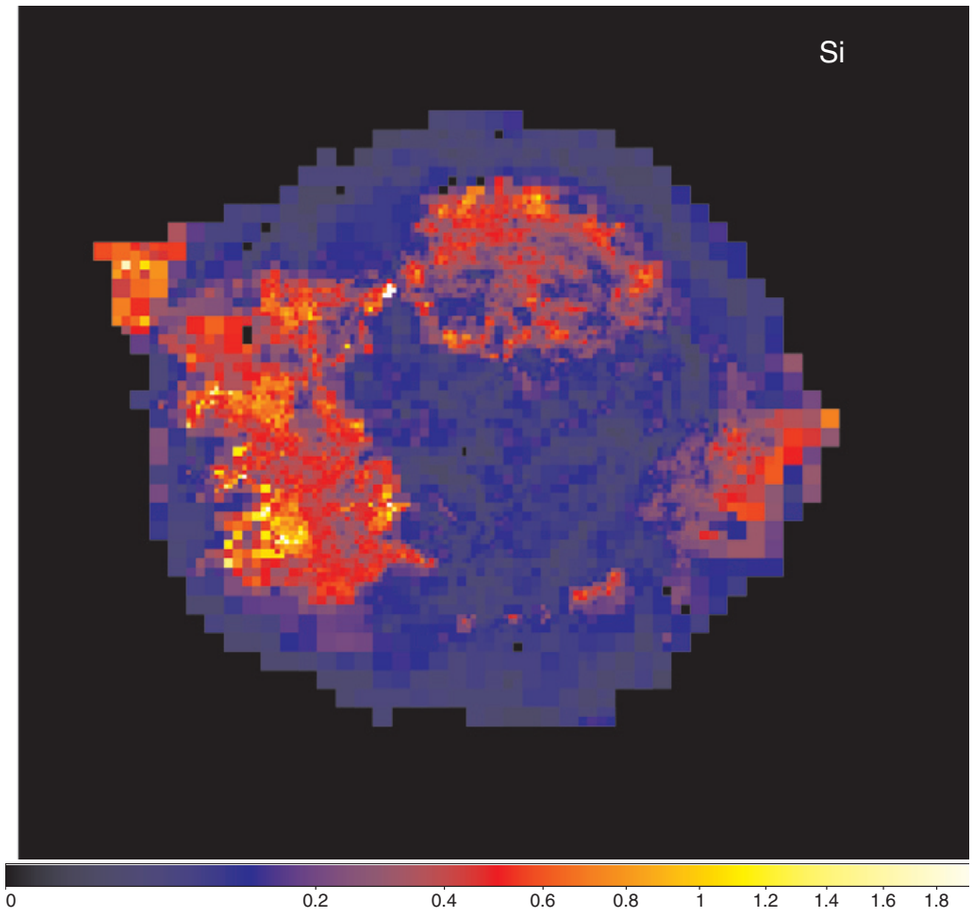


CasA, Hwang+12

CasA dataset: >1 billion photons !

- **Multiple physical compents entangled:**
  - **Mixed/nested ejecta, synchrotron**
  - **projection effect**
  - **region definition impacts spectra**
- **Adaptative tiling (Voronoï, contours):**
  - **cells defined on surface brightness not physical components spatial distribution**
  - **cell definition for Si is not ideal for synchrotron emission**
  - **Painful analysis (>10^4 spectra to fit)**

Challenges in data analysis

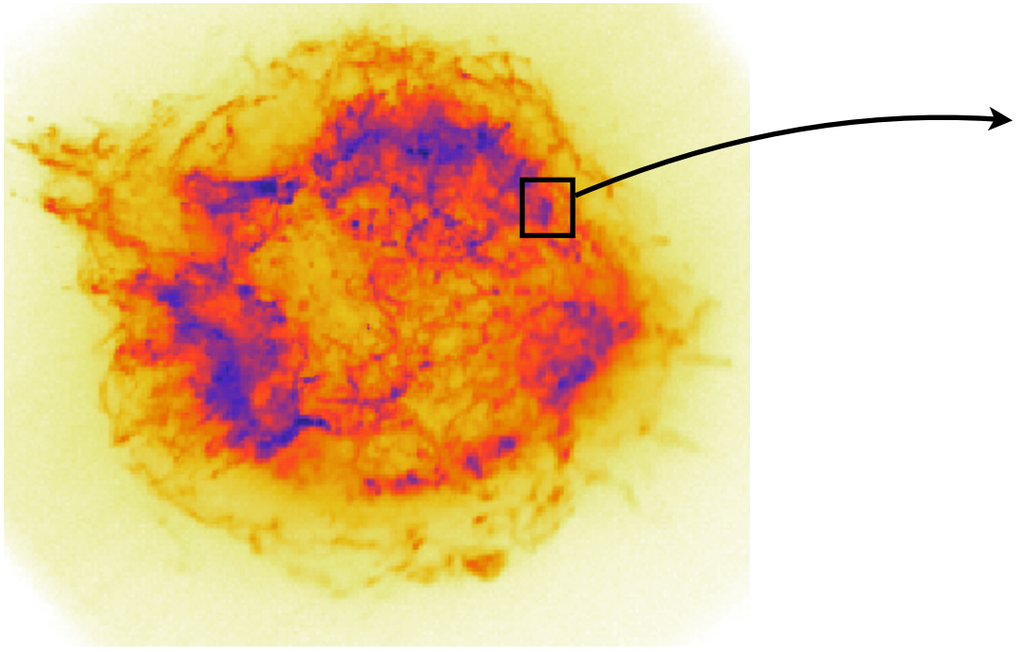


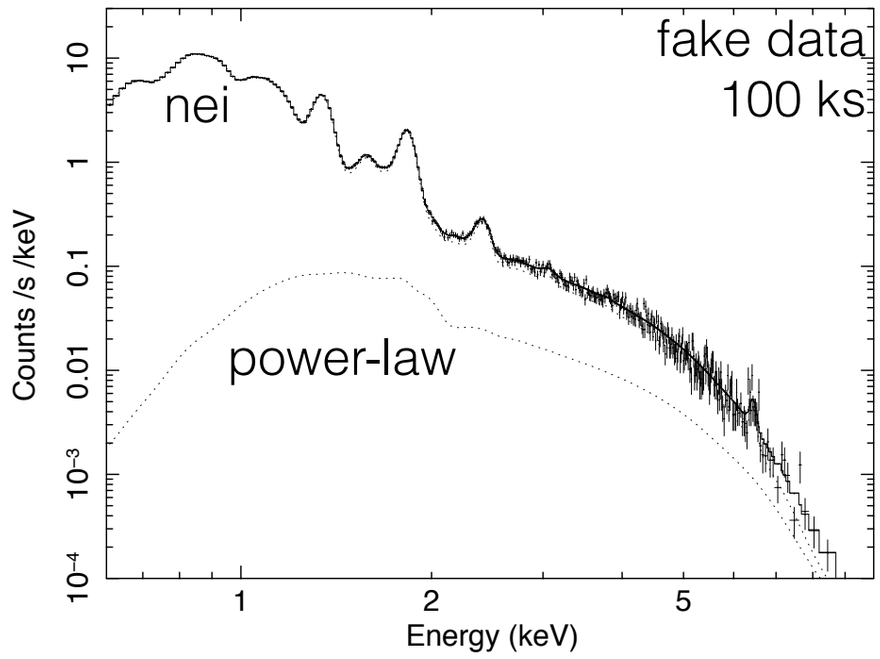
CasA, Hwang+12

CasA dataset: >1 billion photons !

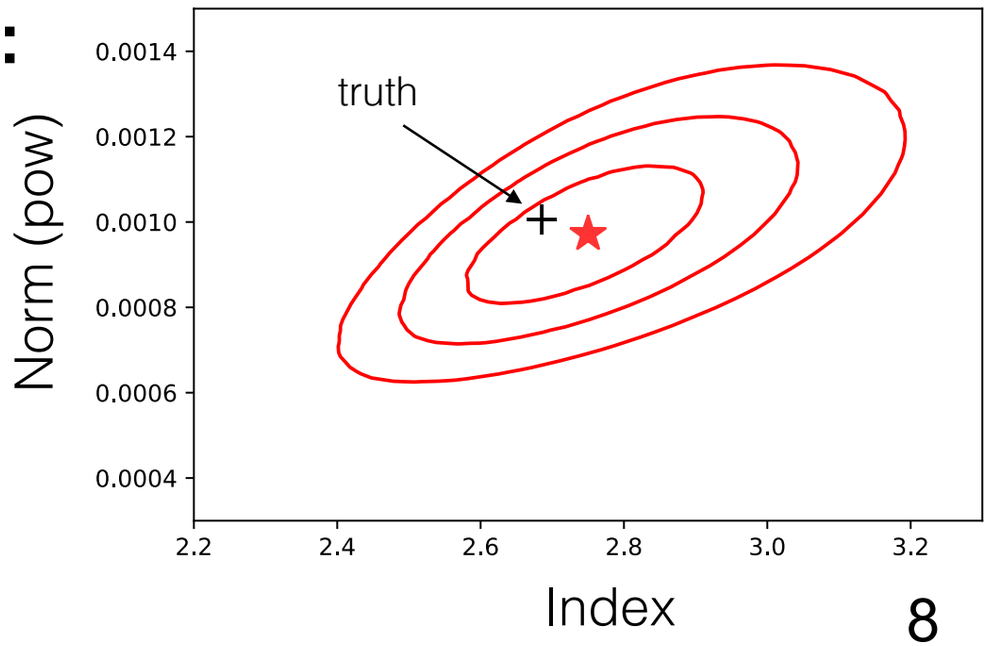
- **Multiple physical compents entangled:**
  - **Mixed/nested ejecta, synchrotron**
  - **projection effect**
  - **region definition impacts spectra**
- **Adaptative tiling (Voronoï, contours):**
  - **cells defined on surface brightness not physical components spatial distribution**
  - **cell definition for Si is not ideal for synchrotron emission**
  - **Painful analysis (>10^4 spectra to fit)**
- **Issue: each cell is treated independently from it neighboors !**
  - **2D then 1D, not fully exploiting the data**

X-ray traditional workflow

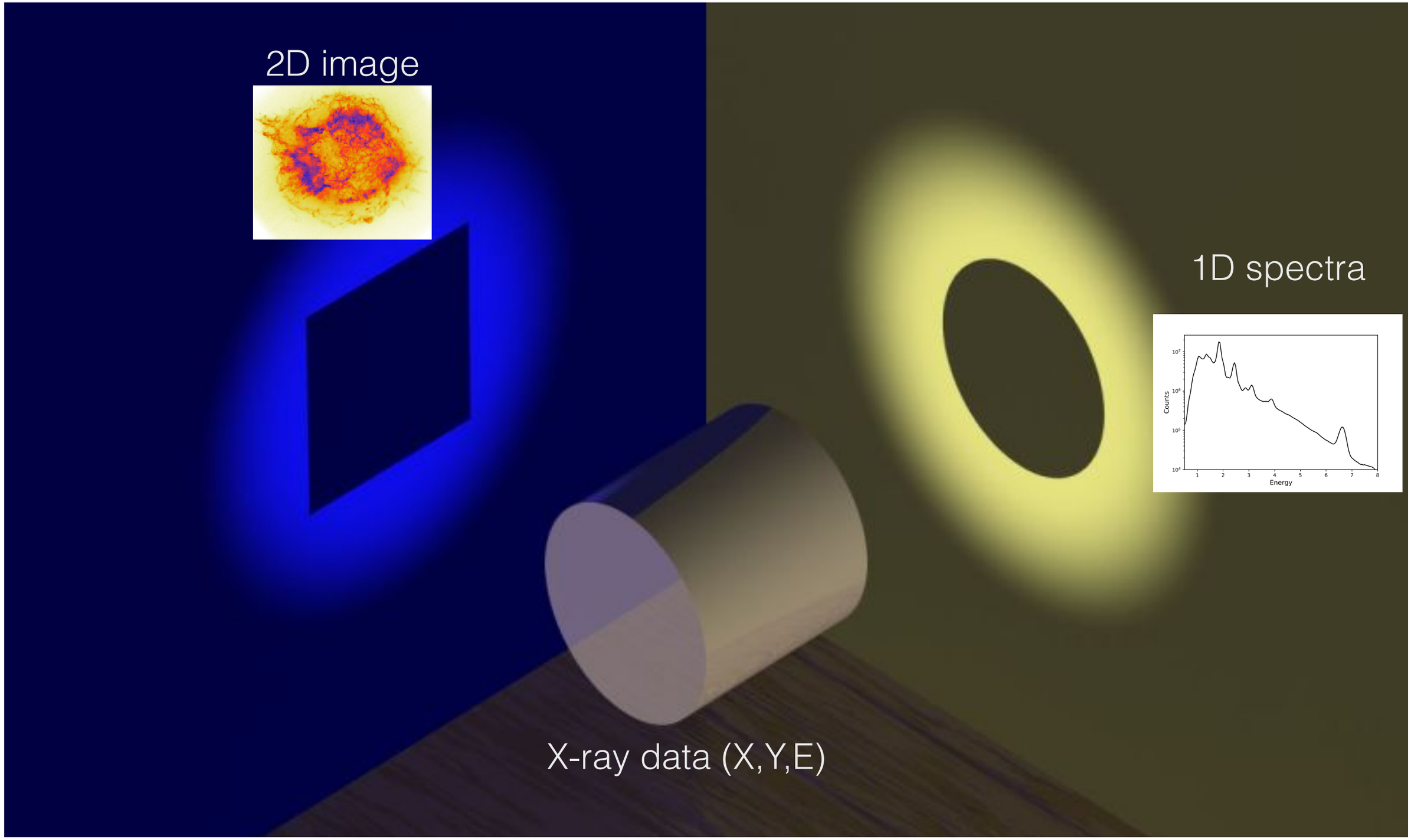




- **Parametric fit. Compare models : χ^2(nei) vs χ^2(nei+pow)**
- **Disentangling process is done in 1D using spectral signatures**
- **Hard to characterize faint components:**
  - **Degeneracy between Norm & Index**
- **Components have spatial signatures !**

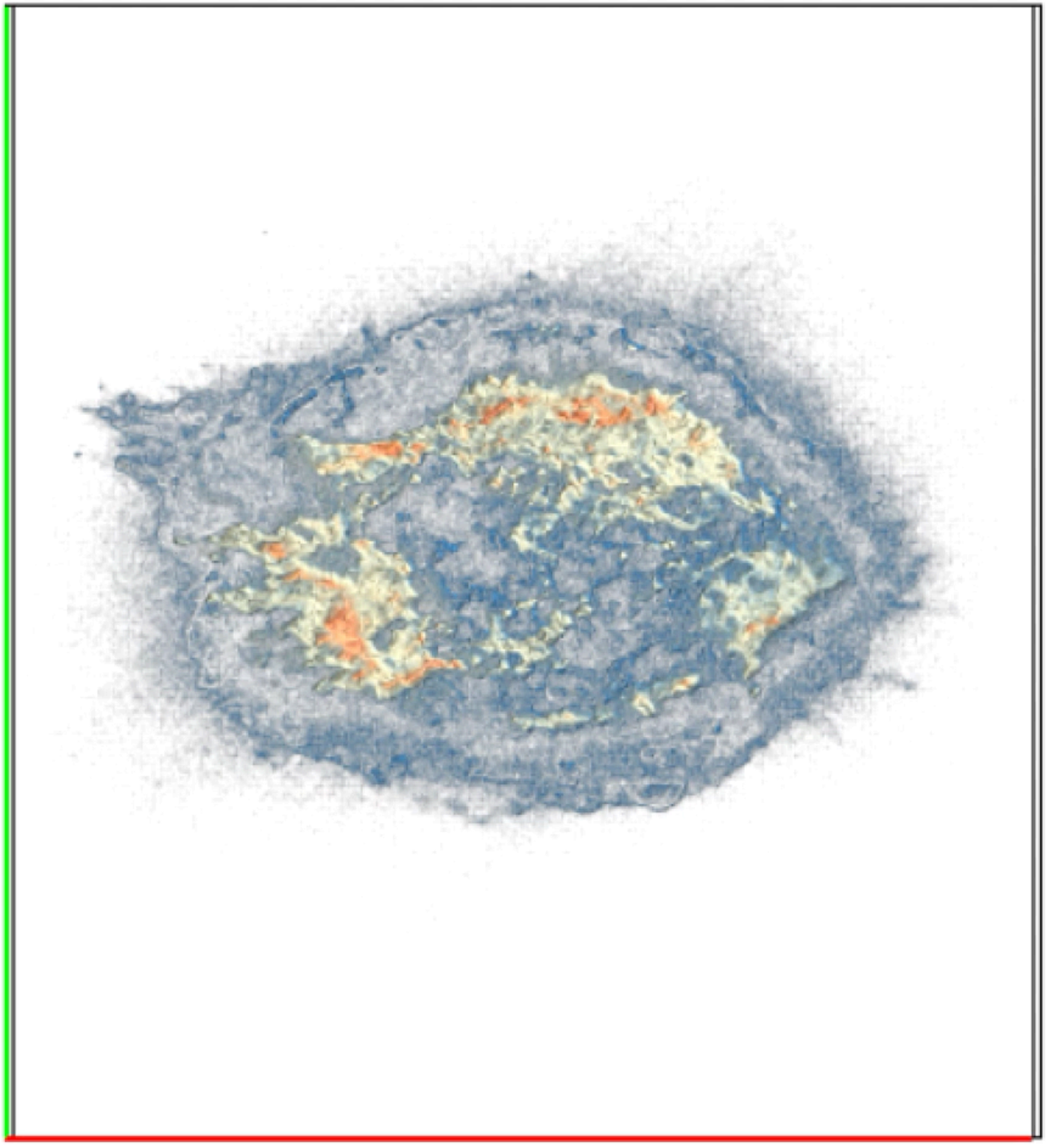


The spectro-imagers schizophrenia



A 3D view of CasA data

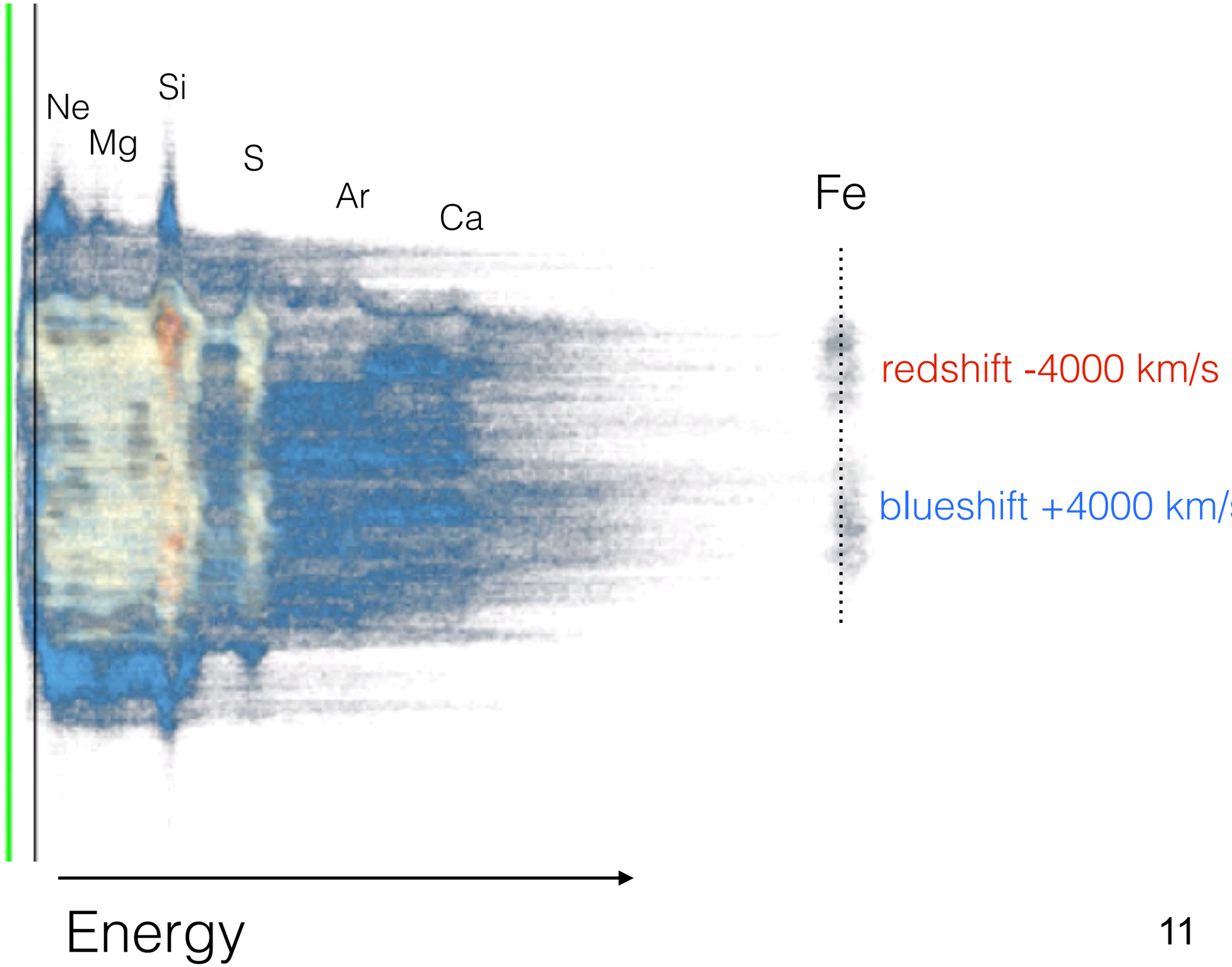
CasA Chandra X,Y,E cube



Data visualization
with vaex

A 3D view of CasA data

CasA Chandra X,Y,E cube



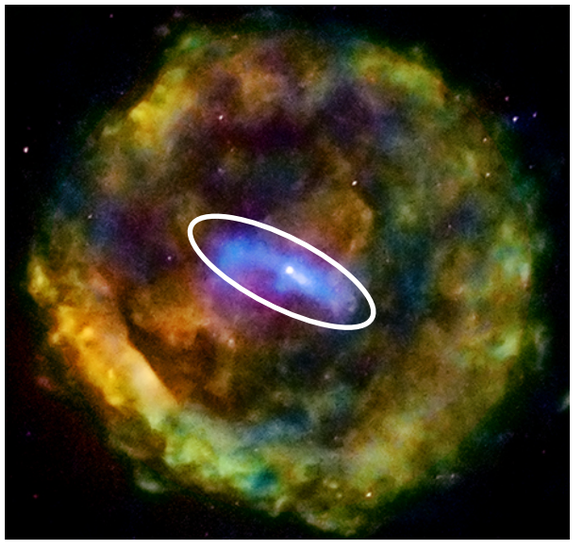
Data visualization with vaex

Status of high-energy workflows

2D then 1D

Spatial selection
then spectral analysis

a)



SNR G11.2-0.3

Spectro-morphology drawn by *eye*

X-ray, TeV analysis

Issue:
ON region contains SNR + PWN

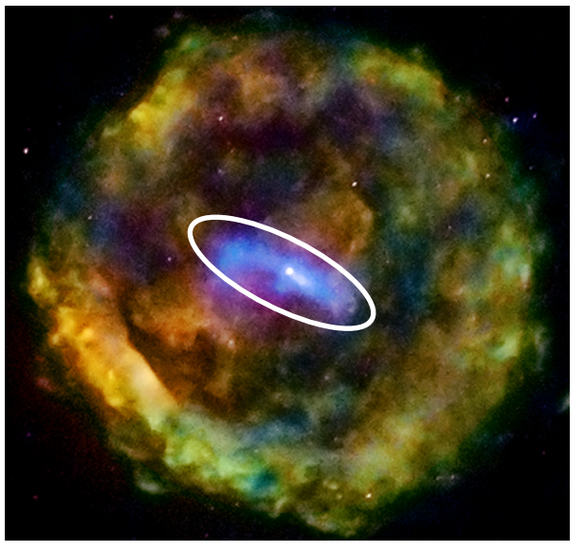
Status of high-energy workflows

2D then 1D

Spatial selection
then spectral analysis

2D + 1D
spatial + spectral
parametric analysis

a)



SNR G11.2-0.3

b)

Spectro-morphology drawn by *eye*

parametric morphology (gaussian, disk)
parametric spectra (Power-Law, etc)

X-ray, TeV analysis

GeV analysis

Issue:
ON region contains SNR + PWN

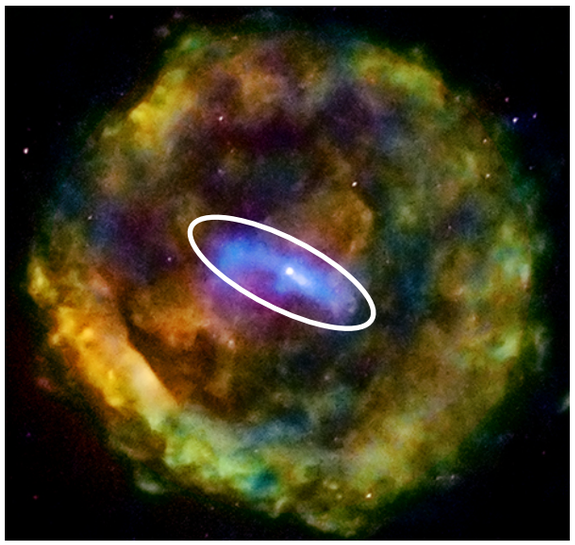
Issue:
Which morphology ?
Reality is not disk & gaussians

Status of high-energy workflows

2D then 1D

Spatial selection
then spectral analysis

a)



SNR G11.2-0.3

Spectro-morphology drawn by *eye*

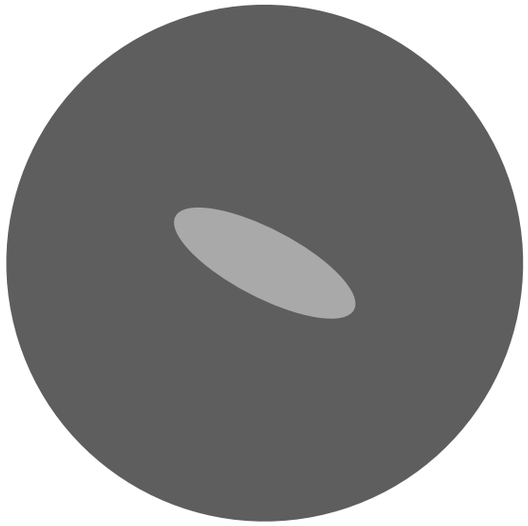
X-ray, TeV analysis

Issue:
ON region contains SNR + PWN

2D + 1D

spatial + spectral
parametric analysis

b)



parametric morphology (gaussian, disk)
parametric spectra (Power-Law, etc)

GeV analysis

Issue:
Which morphology ?
Reality is not disk & gaussians

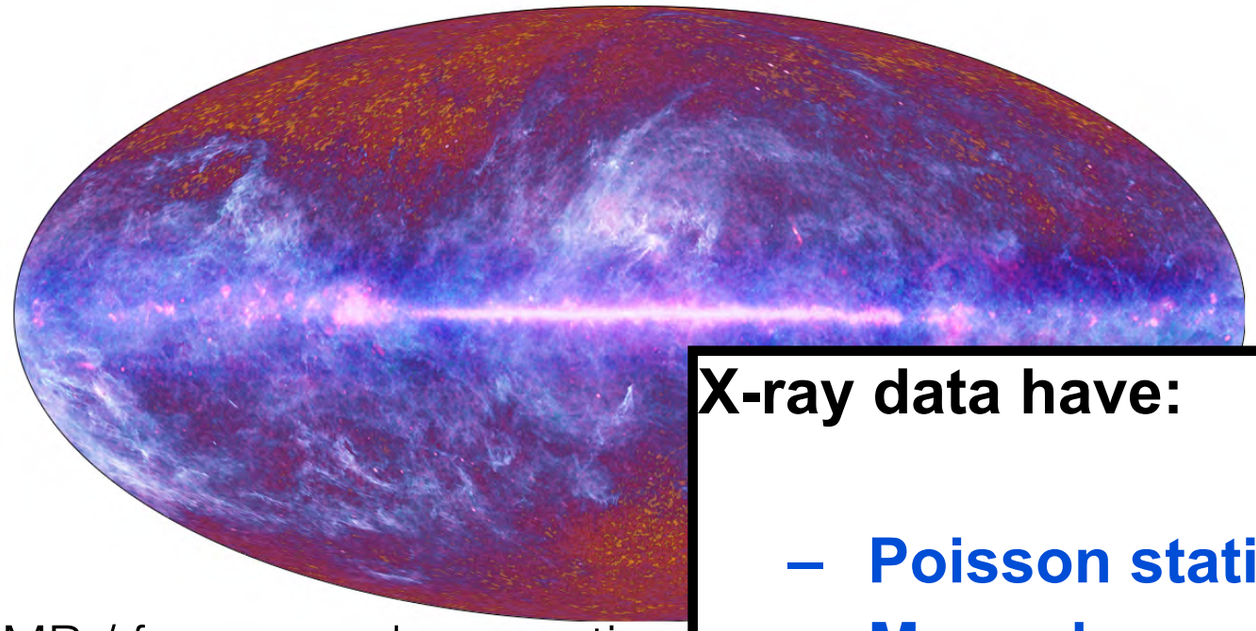
Component separation:
2D + 1D non parametric

c)

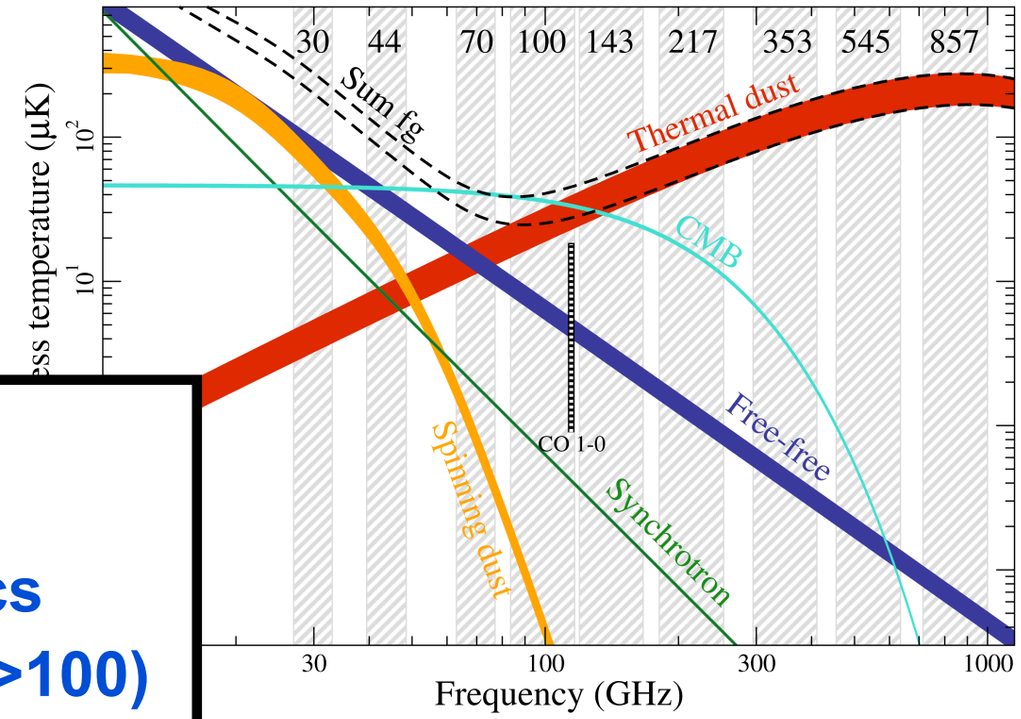
Non-parametric morphology of spectral features

New method
disentangle components with no morphology hypothesis

How to disentangle CMB from foregrounds

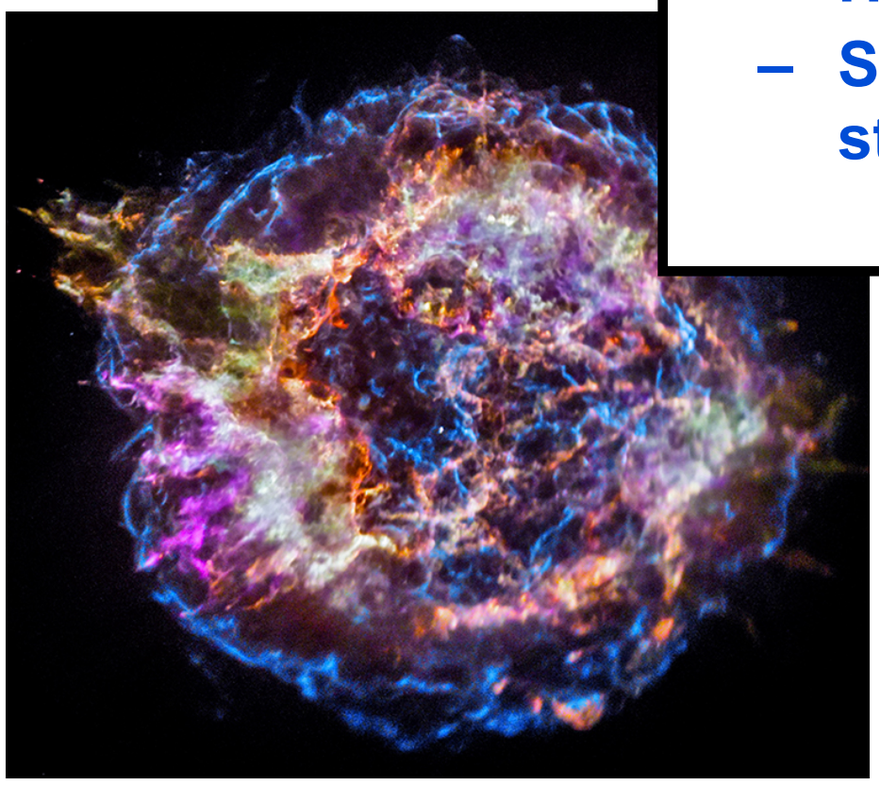


CMB / foreground separation

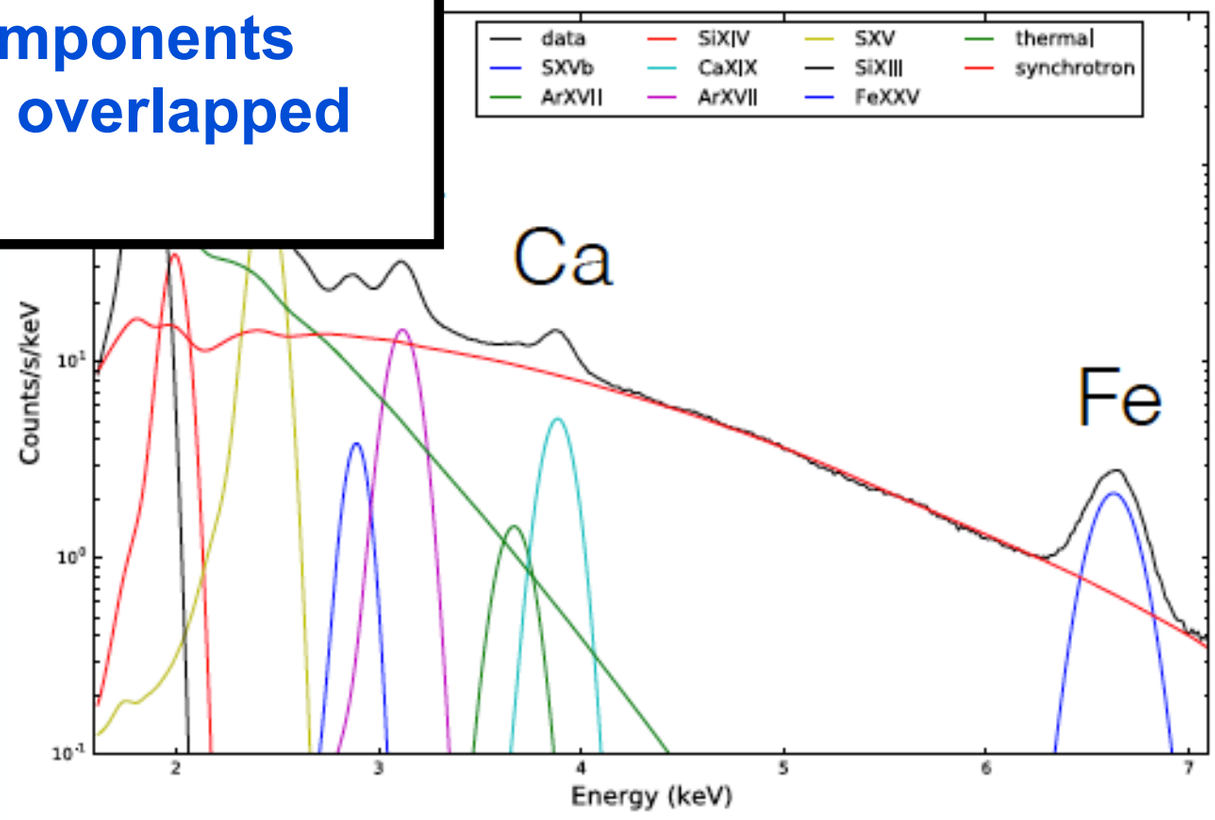


X-ray data have:

- **Poisson statistics**
- **More channels (>100)**
- **High dynamic range**
- **SNR: components strongly overlapped**

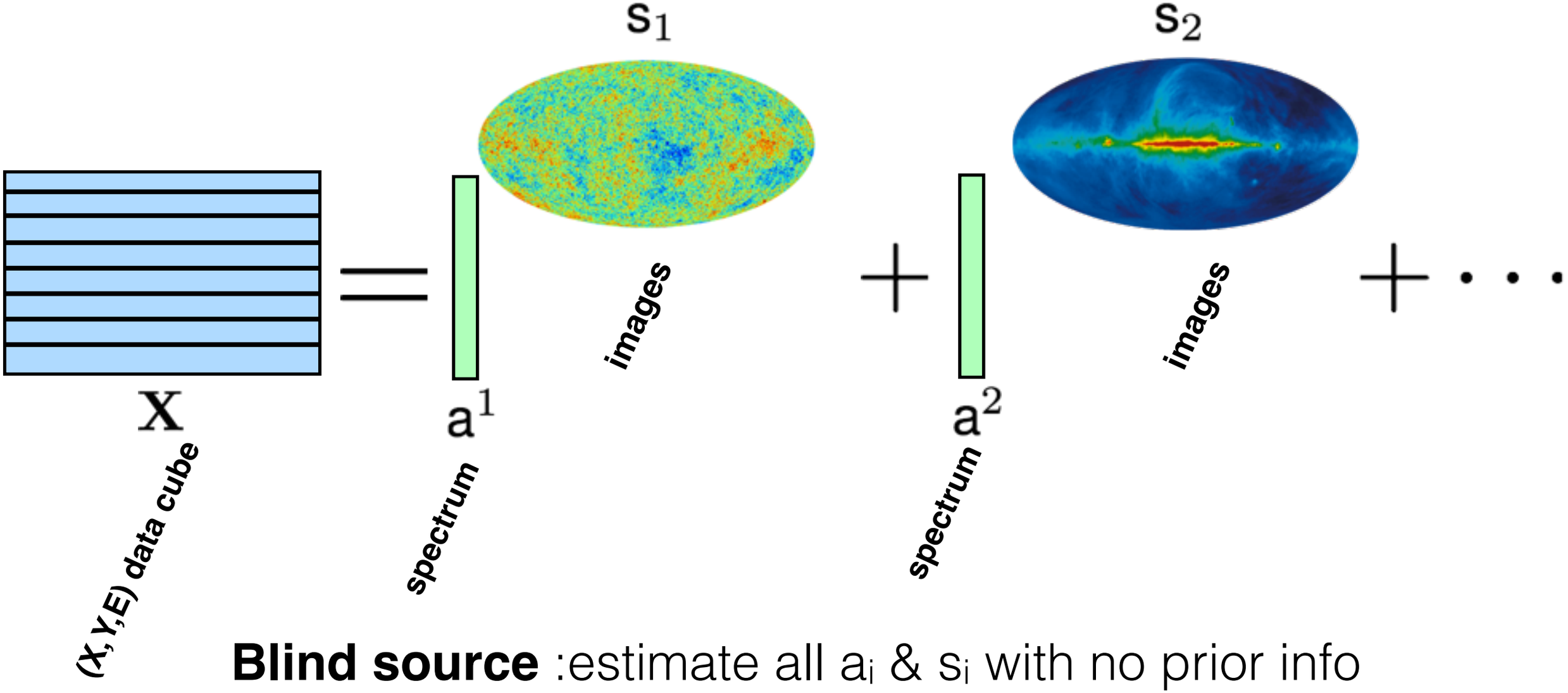


Core Collapse SN : Cassiopeia A



Generalized Morphological Component Analysis (Bobin et al. 2016)

Assumption: Linear combination of Spec*Image



Blind source :estimate all a_i & s_i with no prior info
Semi-blind : fix one spectral component (e.g. CMB)

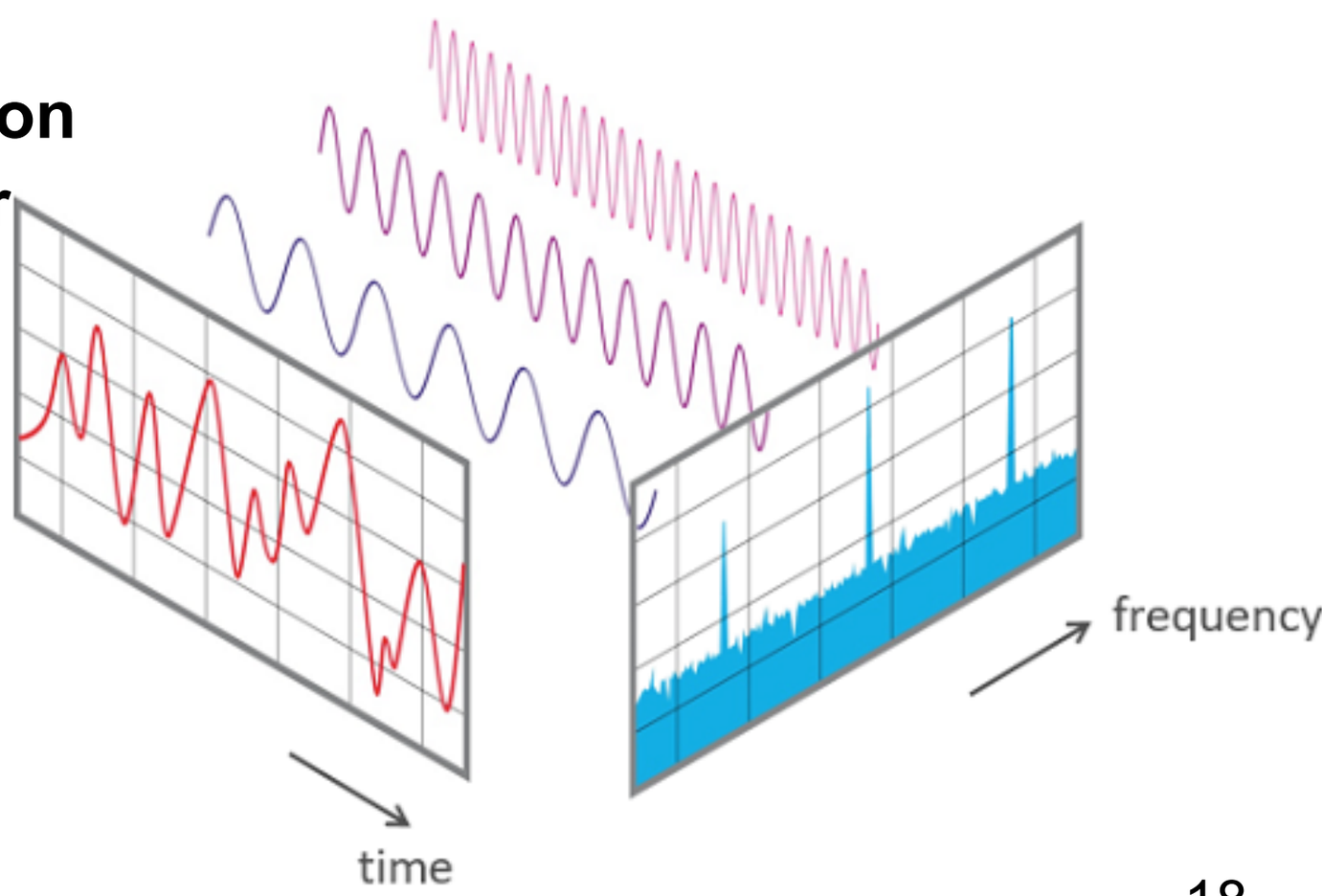
Limitation: morphology doesn't change with energy

Blind source separation: where is the magic ?

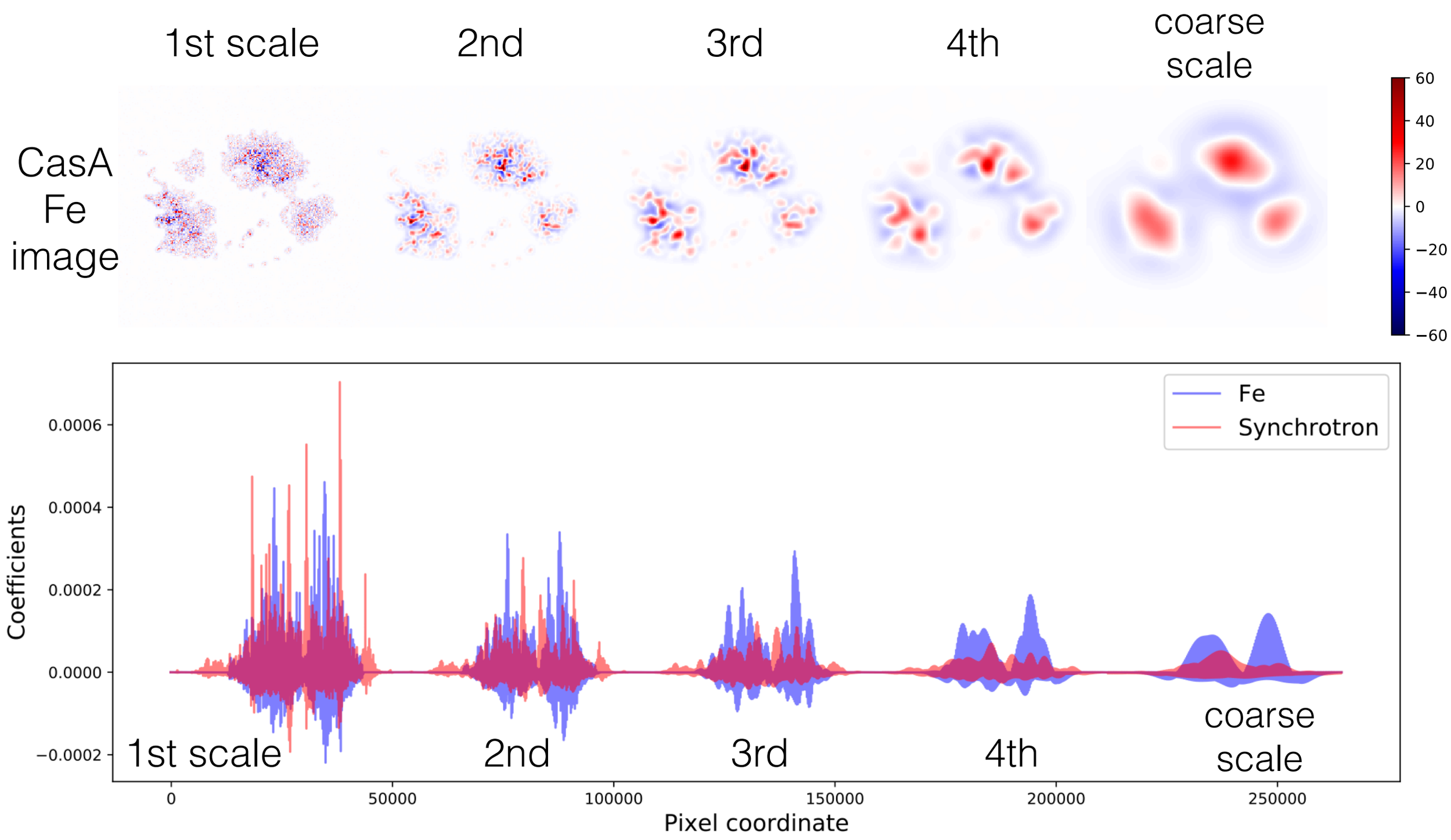
- **Specify NO prior information (no spectral model, instrument response)**
- **Input :**
  - **Cube: 2D-1D data (X,Y,E)**
- **Output:**
  - **N cubes ($Image_i$*$spectra_i$), N fixed by user**

Blind source separation: where is the magic ?

- **Specify NO prior information (no spectral model, instrument response)**
- **Input :**
  - **Cube: 2D-1D data (X,Y,E)**
- **Output:**
  - **N cubes ($Image_i$*$spectra_i$), N fixed by user**
- **Where is the magic ?**
- **Find a sparse data representation**
- **A 1D analogy would be Fourrier**



Wavelet Starlet transform



GMCA: algorithm

$$X = AS + N = \sum_{i=1}^{n} A_i S_i + N$$

data cube — spectrum — image — additive noise

- **Equation has an infinite number of solution (ill-posed inverse problem)**
  - **Example: 4 = a * b cannot be solved without constraints**
  - **Constraint: a & b are positive integers. Solution: (2,2) & (4,1)**

- **Each cube slice (image) is transposed in the wavelet domain (X: 4D array)**

GMCA: algorithm

$$\min_{A,S} \sum_{i=1}^{n} \lambda_i \|S_i\|_p + \|X - AS\|_F^2$$

sparsity constraint term

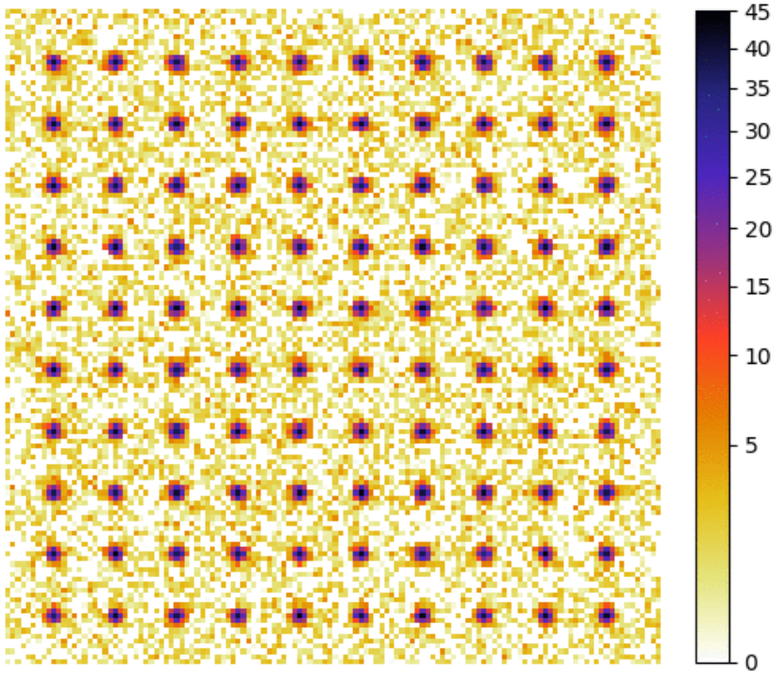
data fidelity term (Poisson likelihood for X-rays)

- **Here we constrain sparsity of the images (S_i) in the wavelet domain**
- **Alternatively estimate S for fixed A, then A fox fixed S**

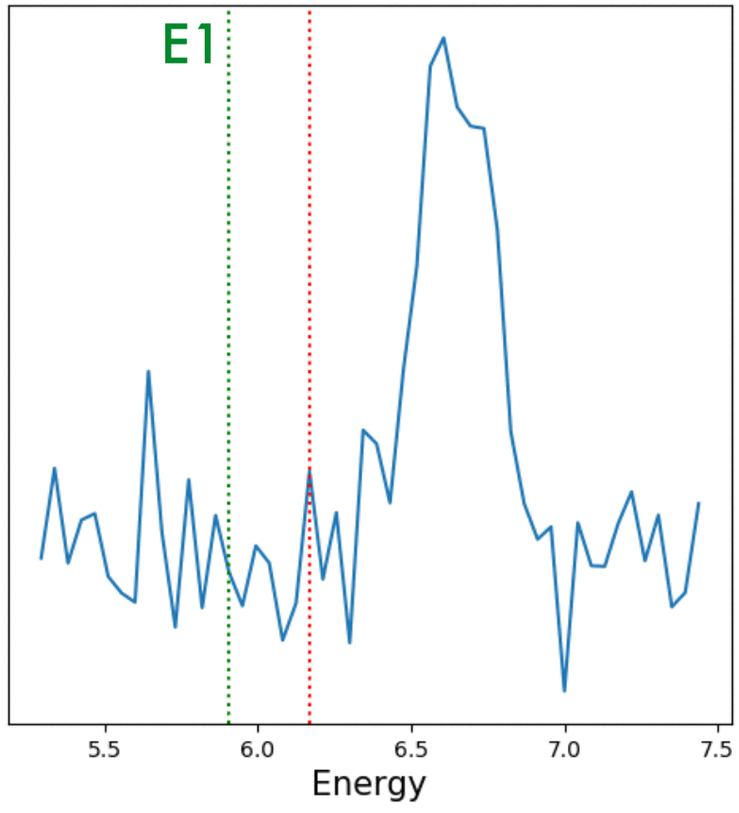
GMCA : example

- **In a way GMCA is an unsupervised ML clustering algorithm**

Cube slice



E= 6.17 - 6.21 keV



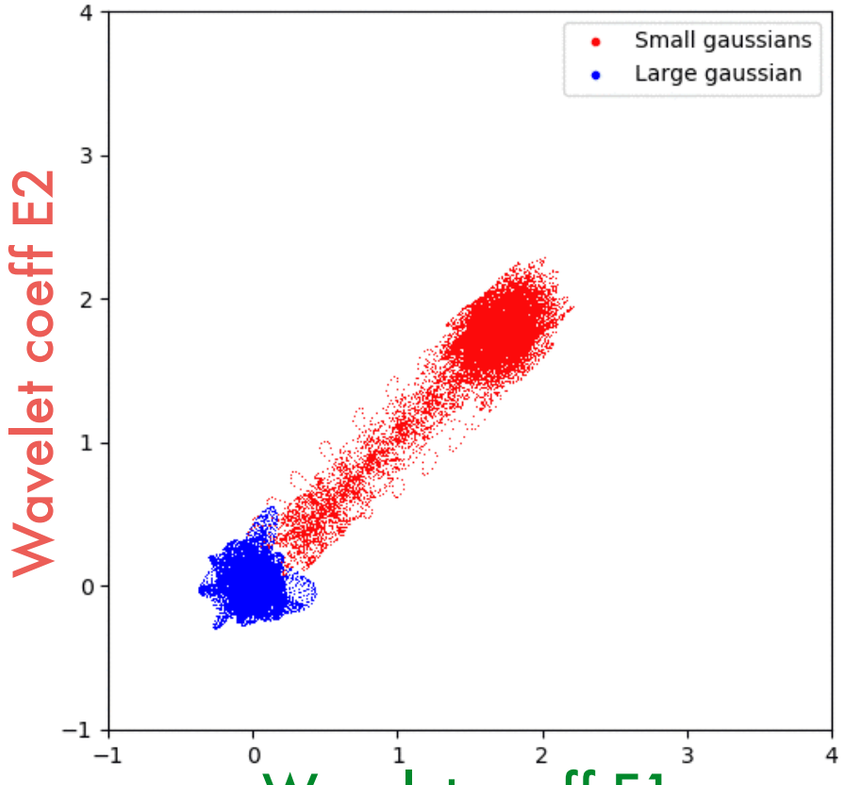
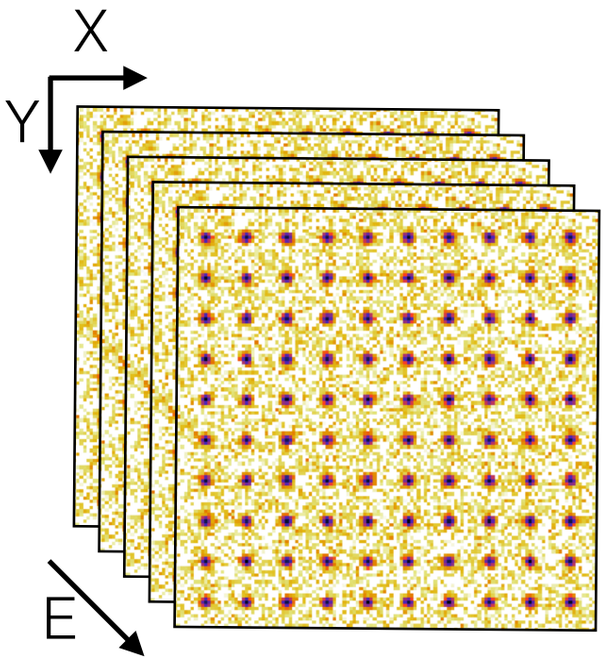
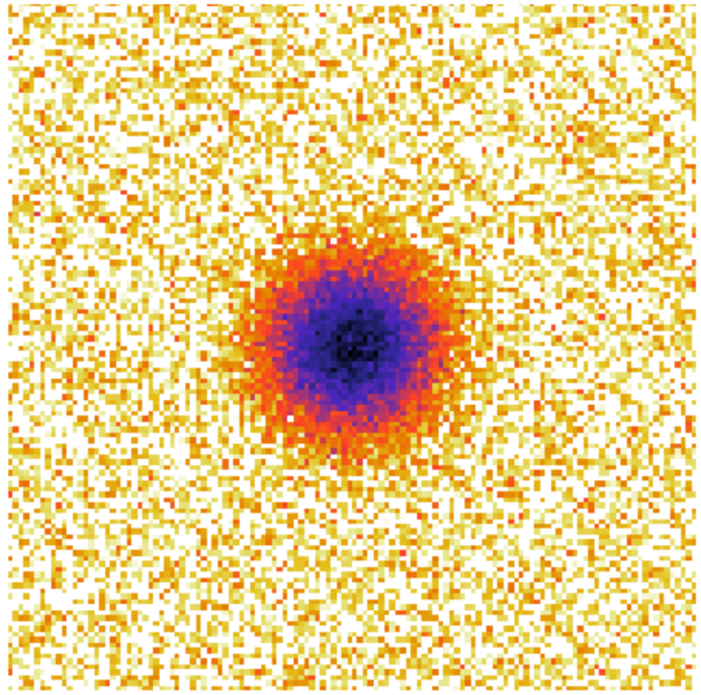


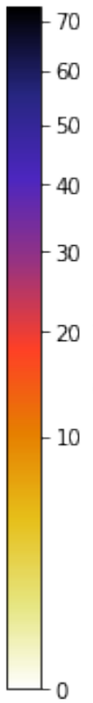
Image Wavelet transform (third scale) coefficients of each component

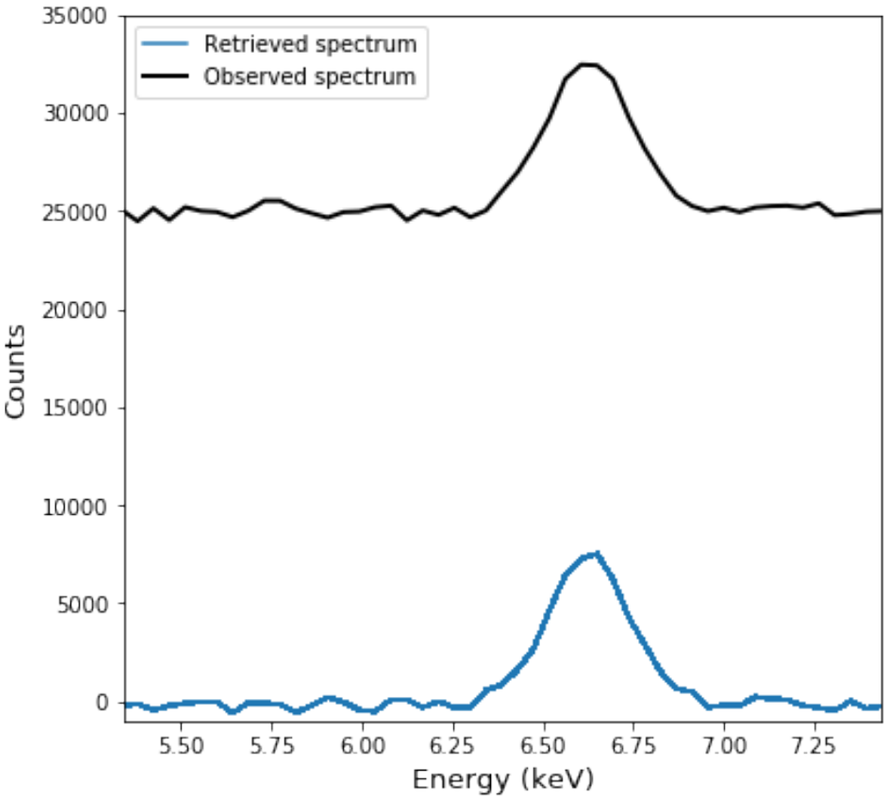
GMCA : example

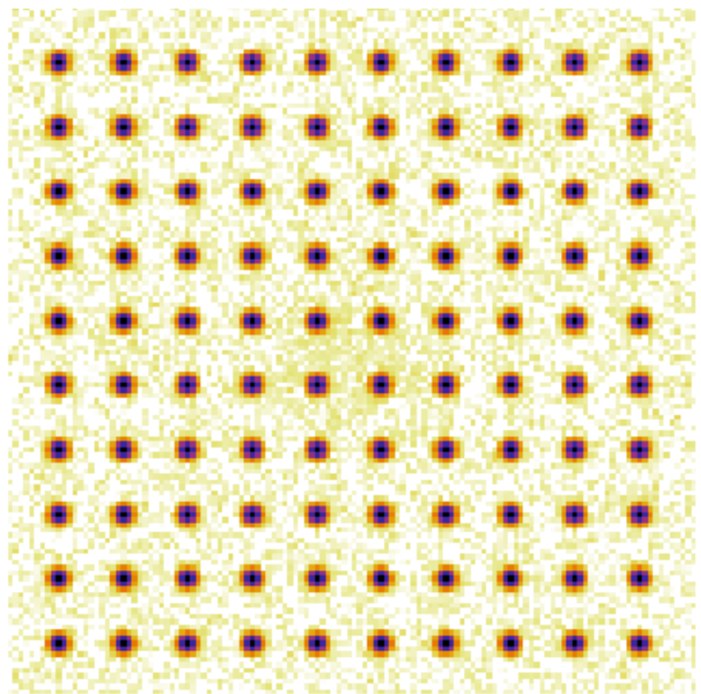


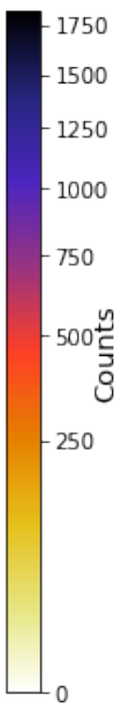
GMCA

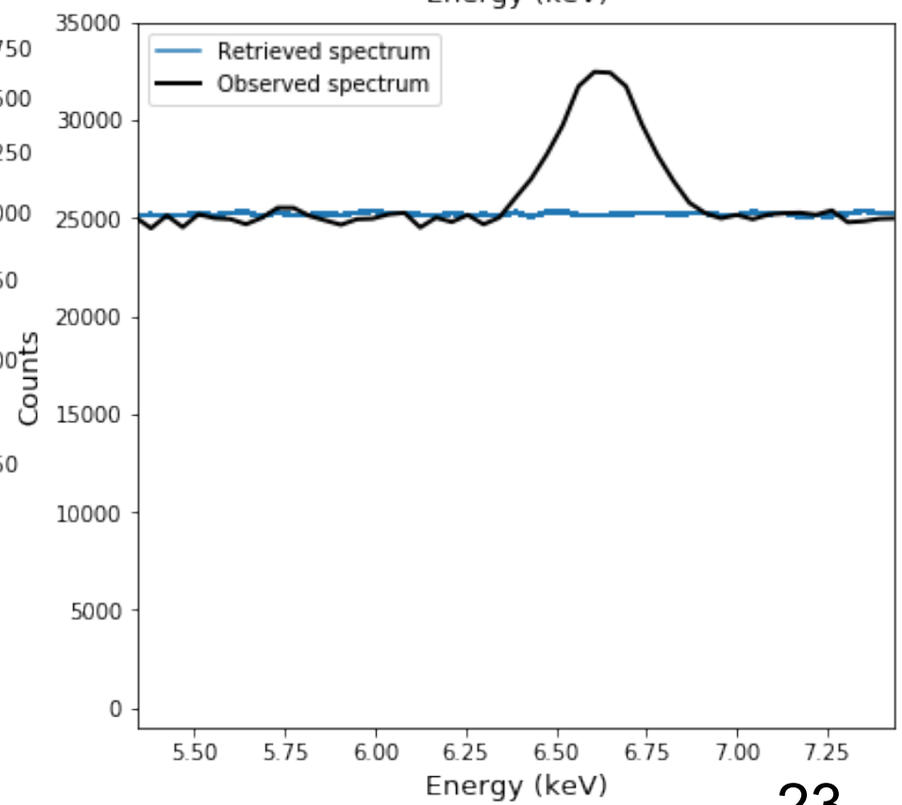




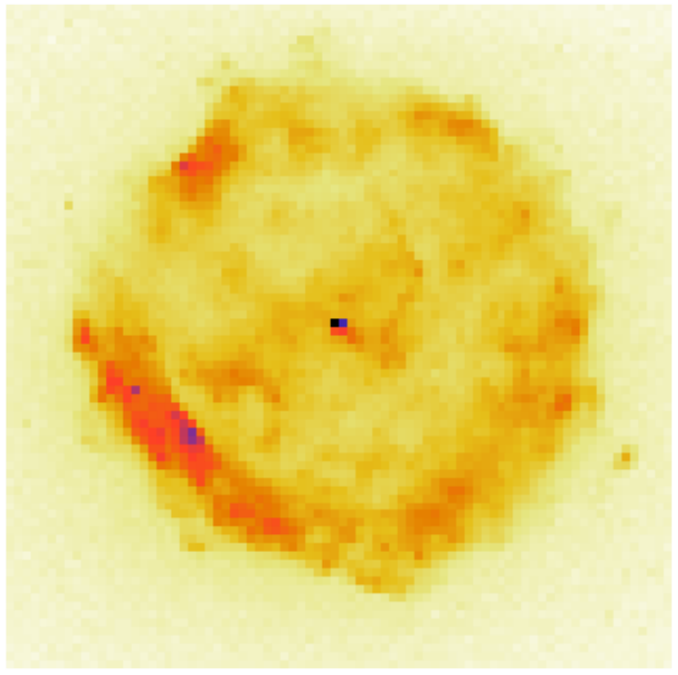


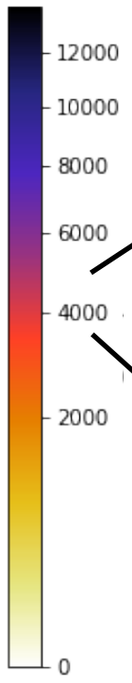




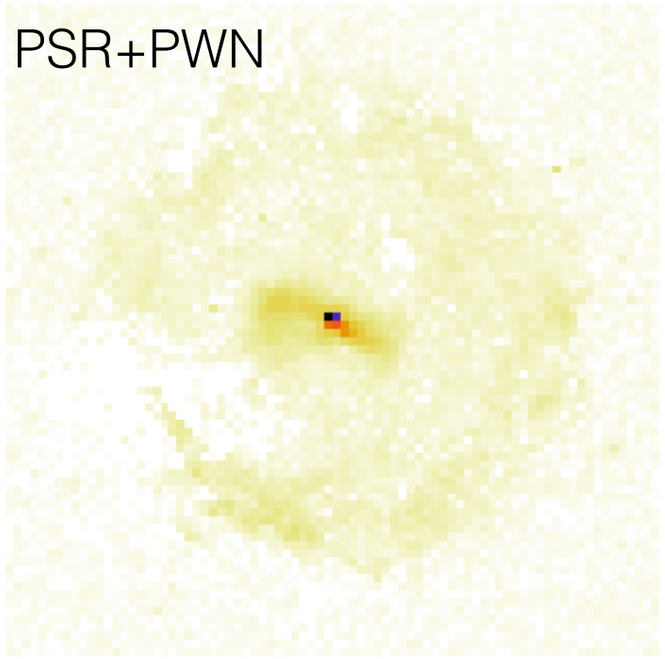


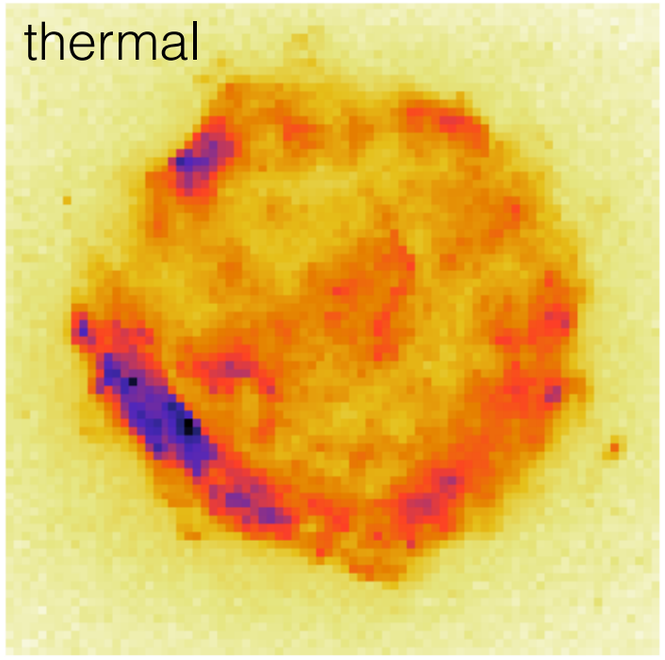
Real case applications: Chandra G11.2-0.3

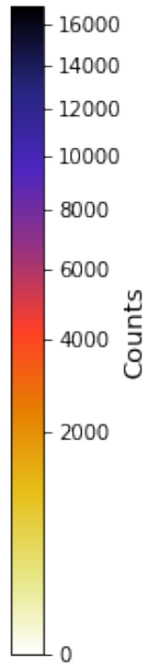


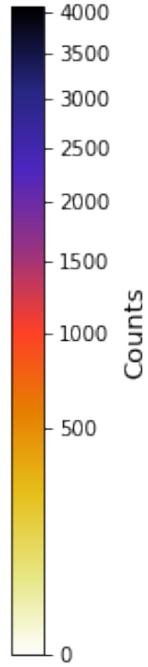


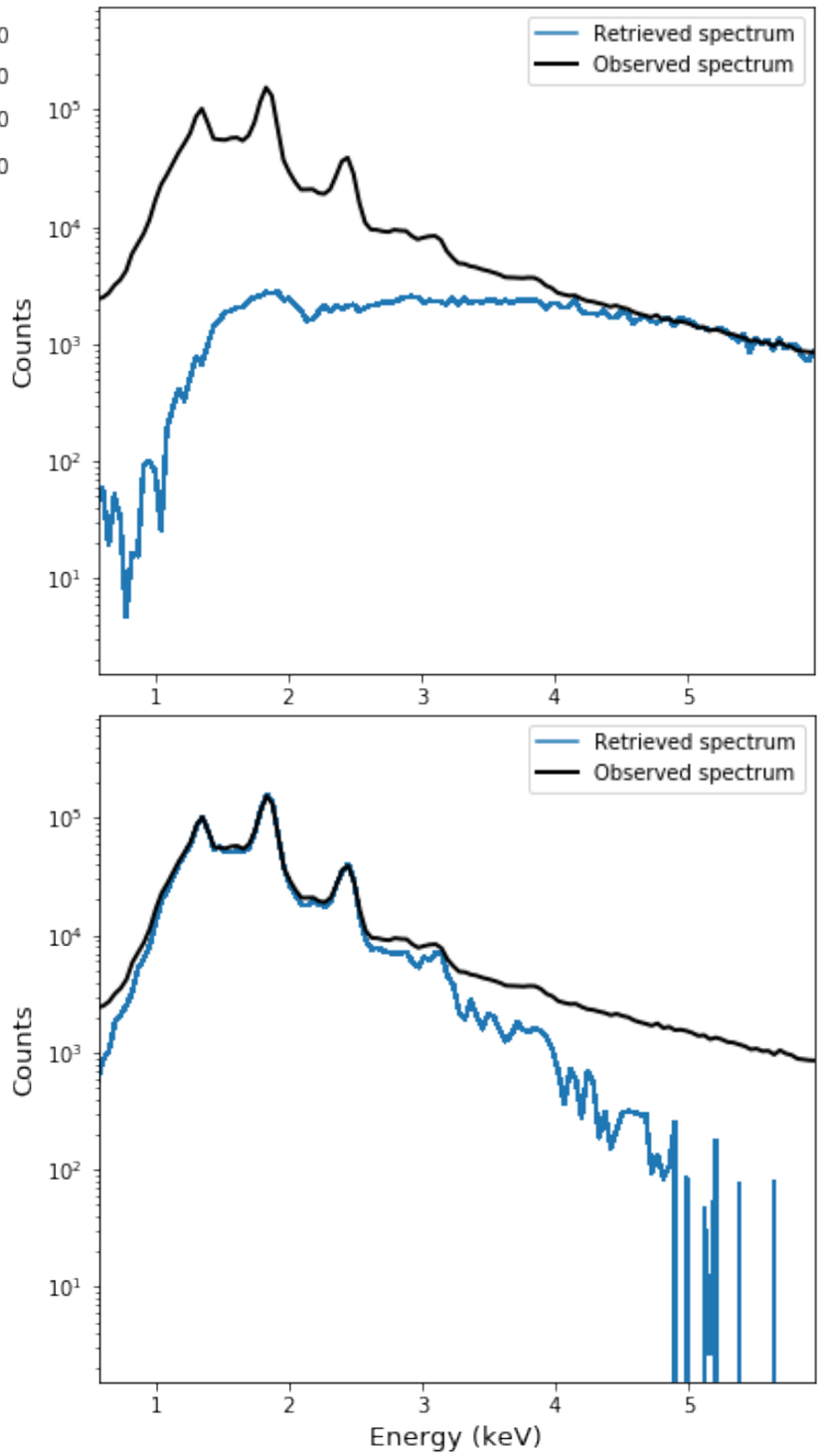
0.5-6 keV cube



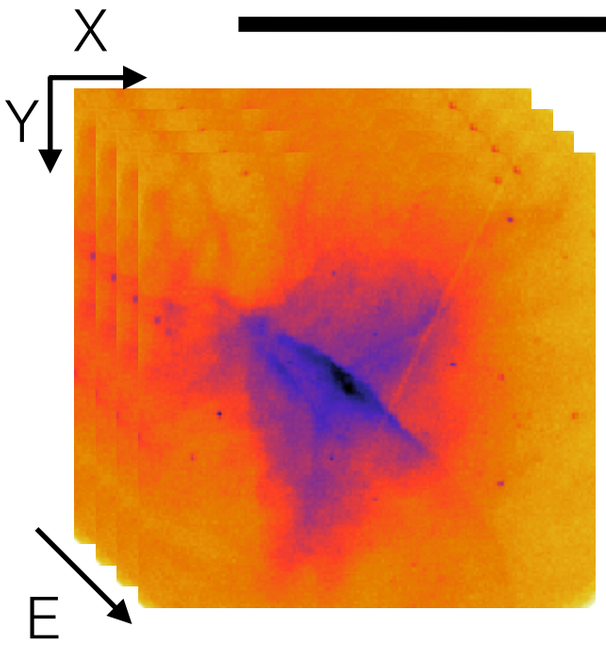




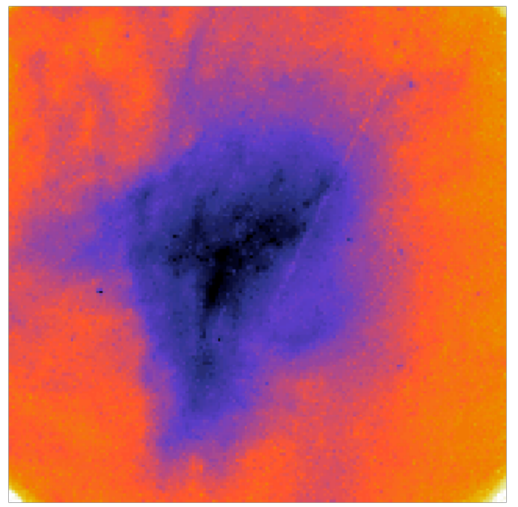


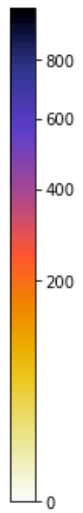


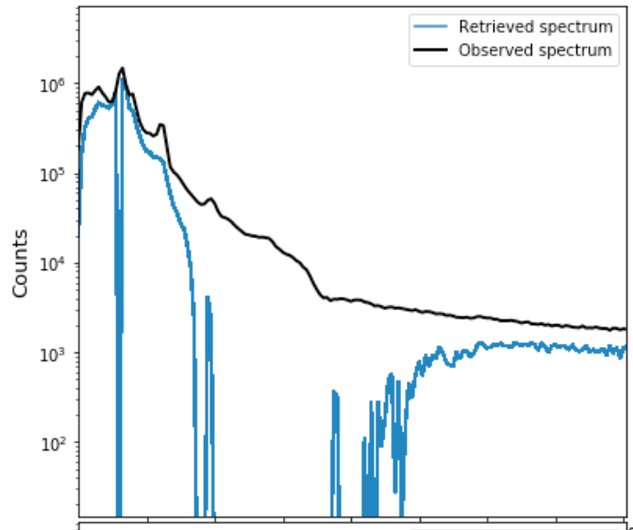
Real case applications: eROSITA EDR Vela region

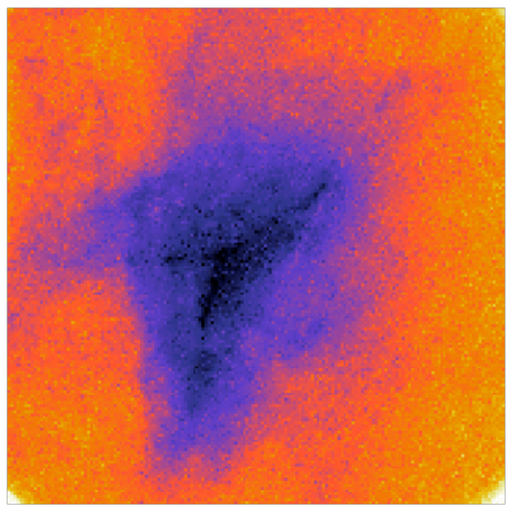


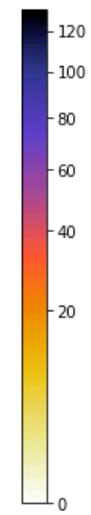
GMCA

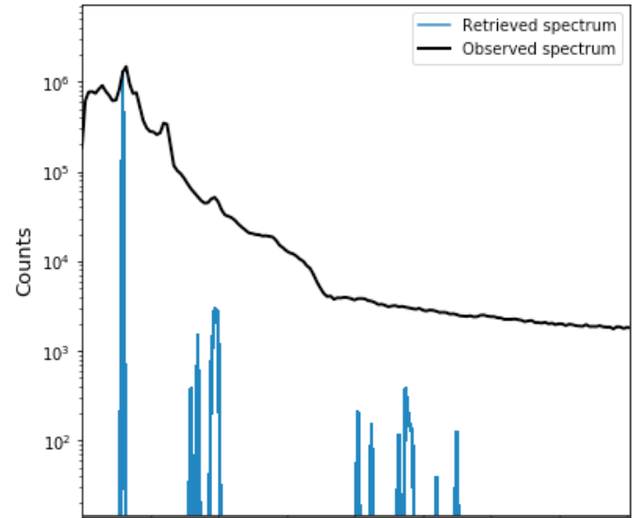


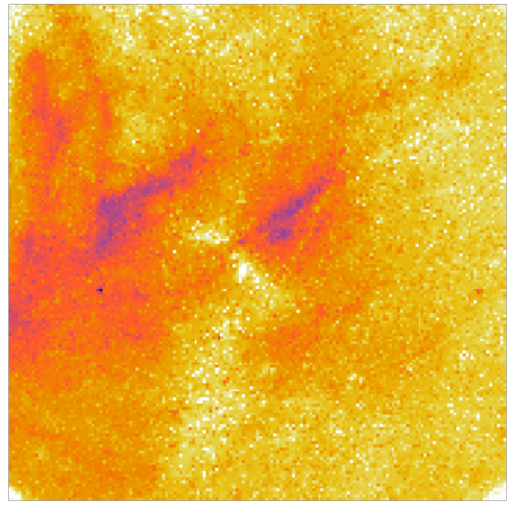


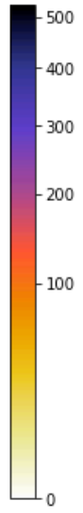


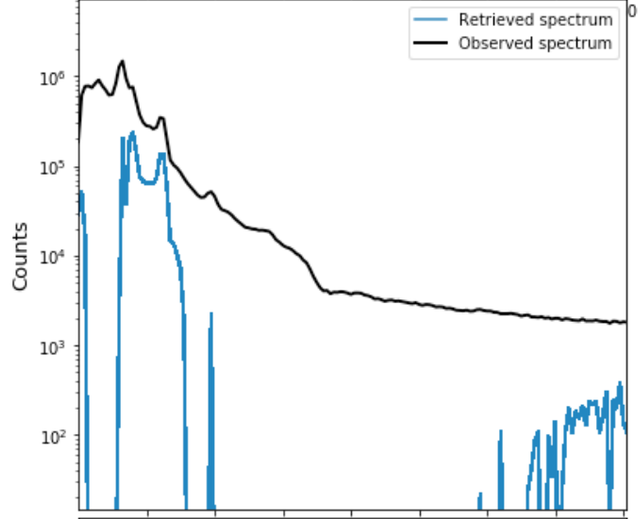


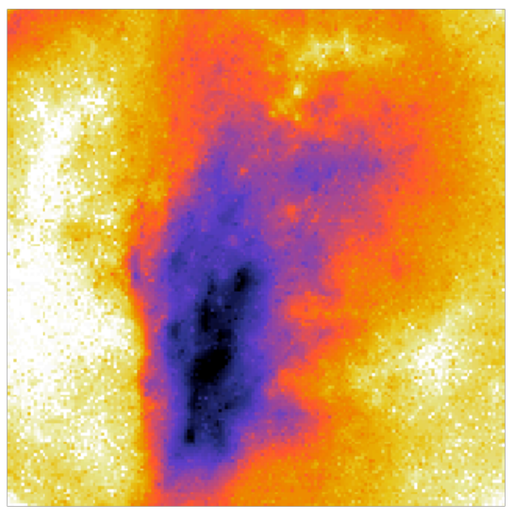


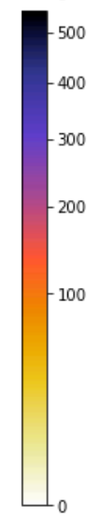


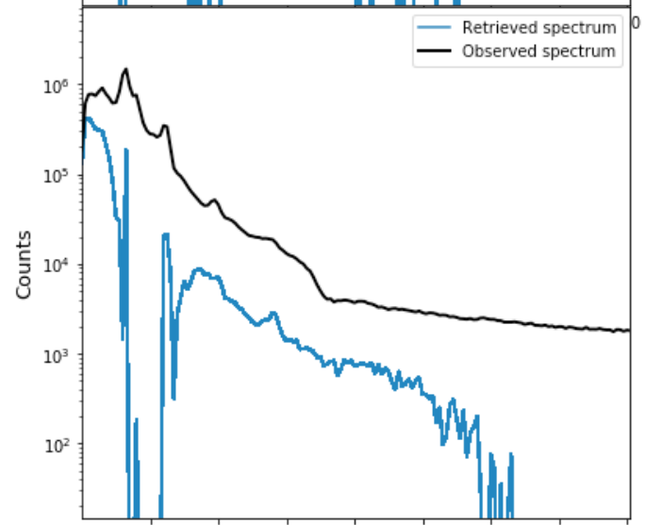


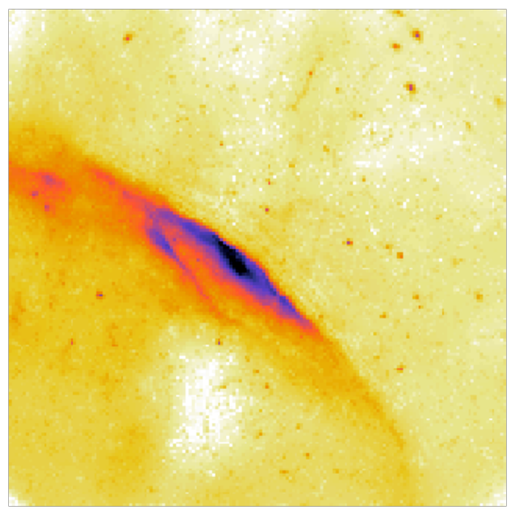


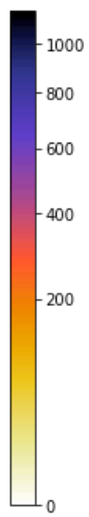


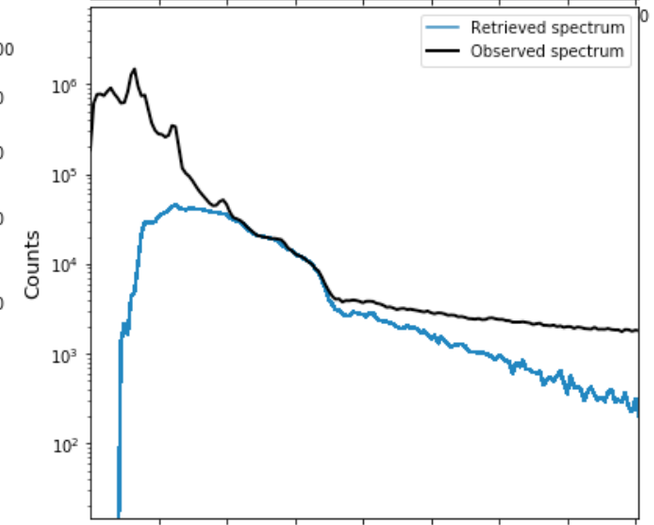






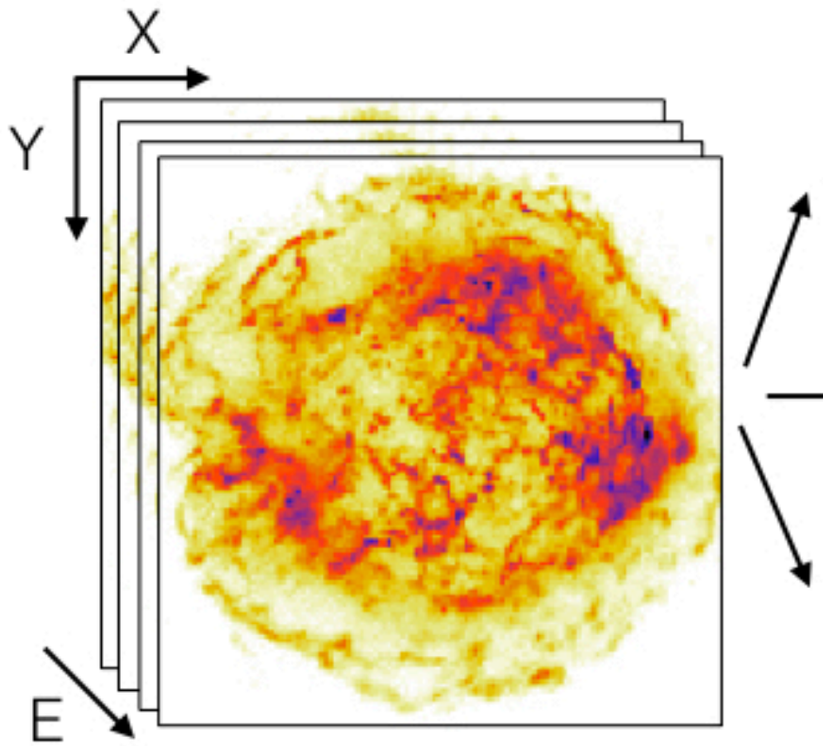




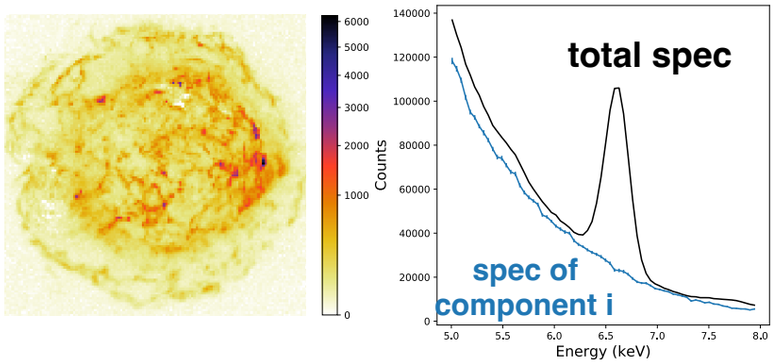


Application to real data: Cassiopeia A

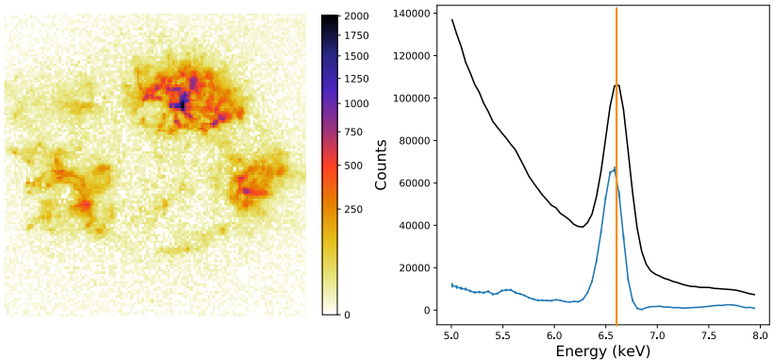
- **Chandra 1 Ms 5-8 keV:**



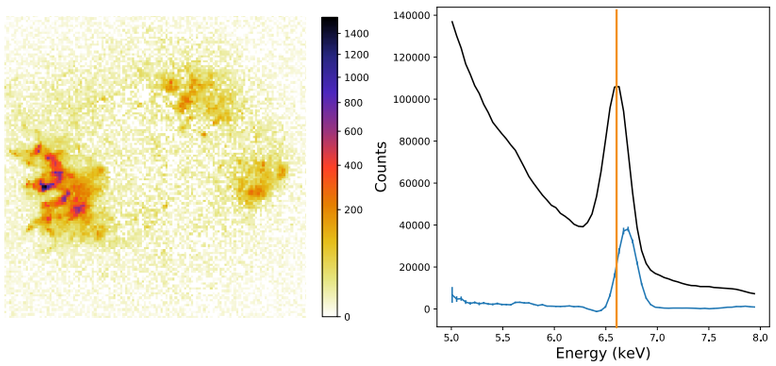
Synchrotron



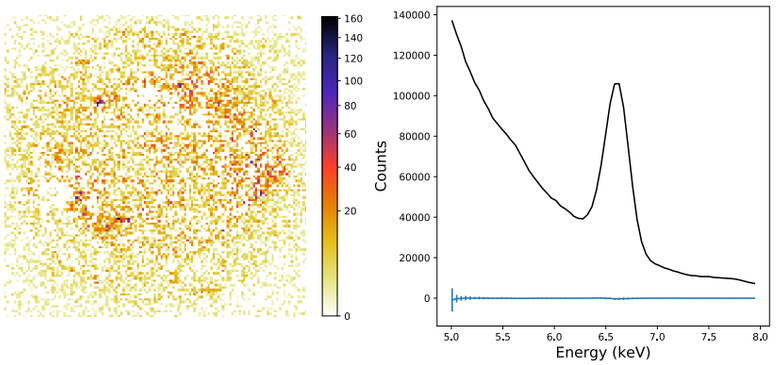
Red-shifted Fe structure



Blue-shifted Fe structure

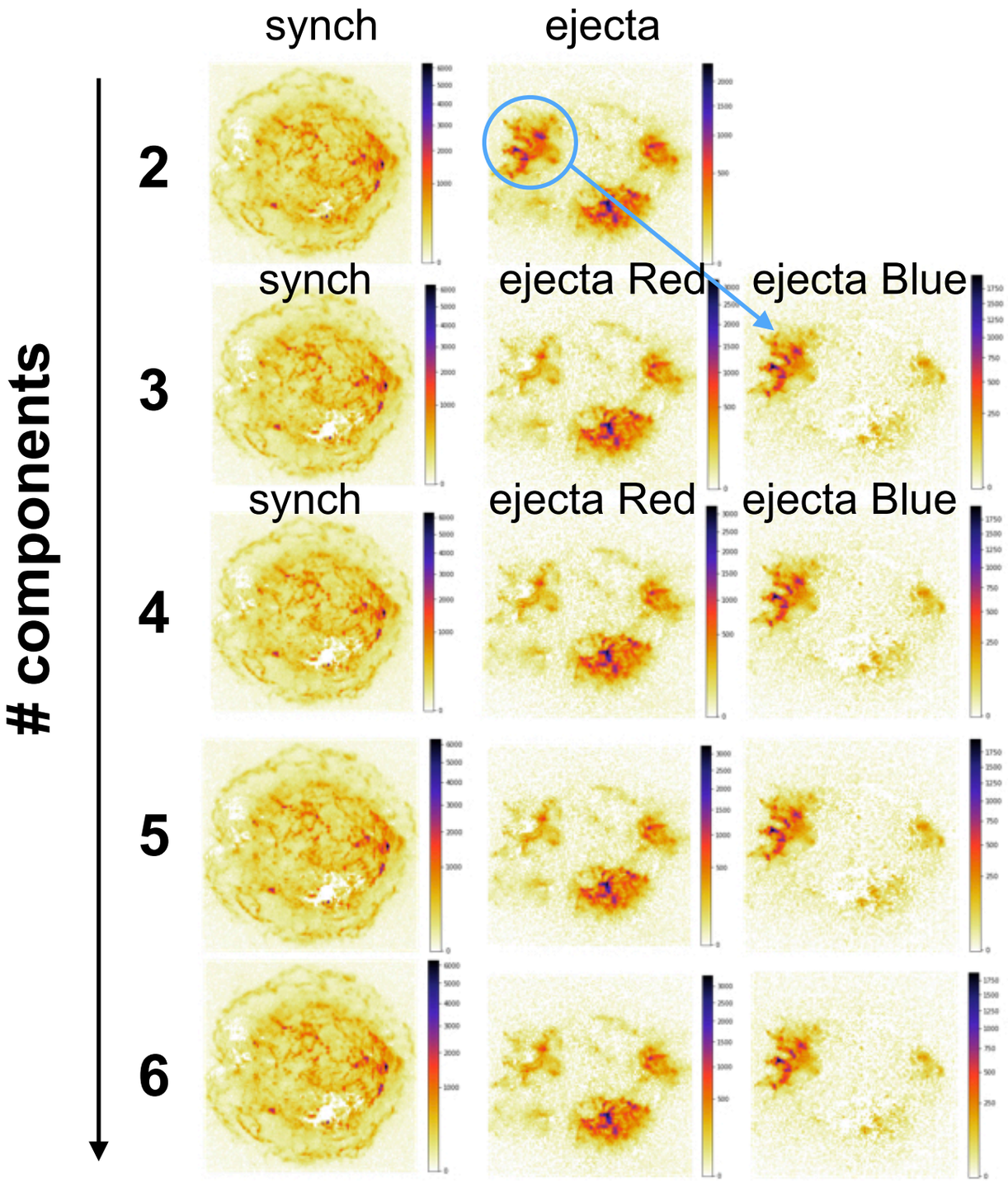


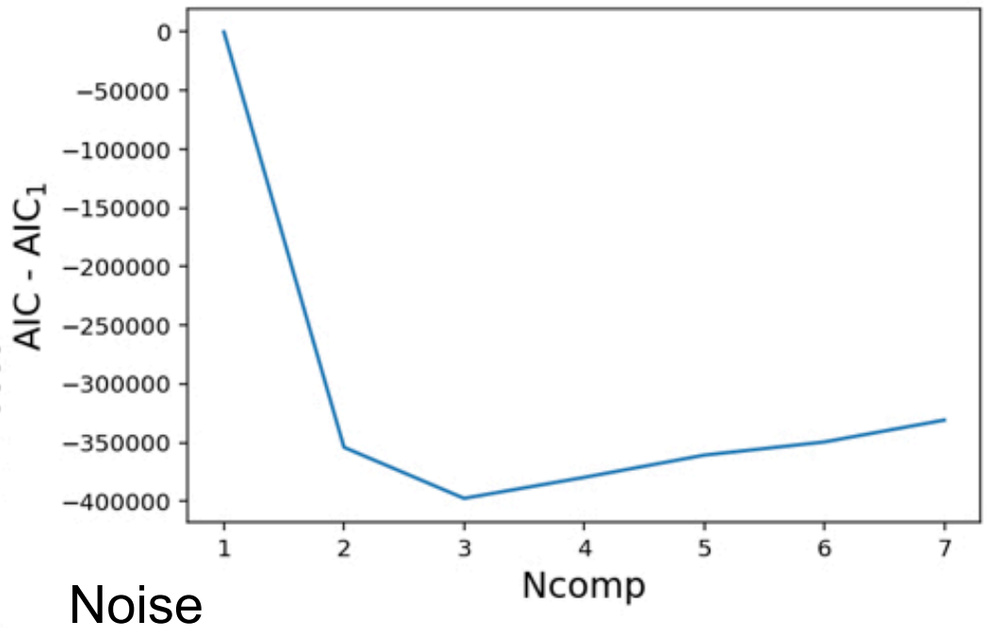
Noise



Akaike information criteria for choosing N component

- **Is my separation better with more components ?**

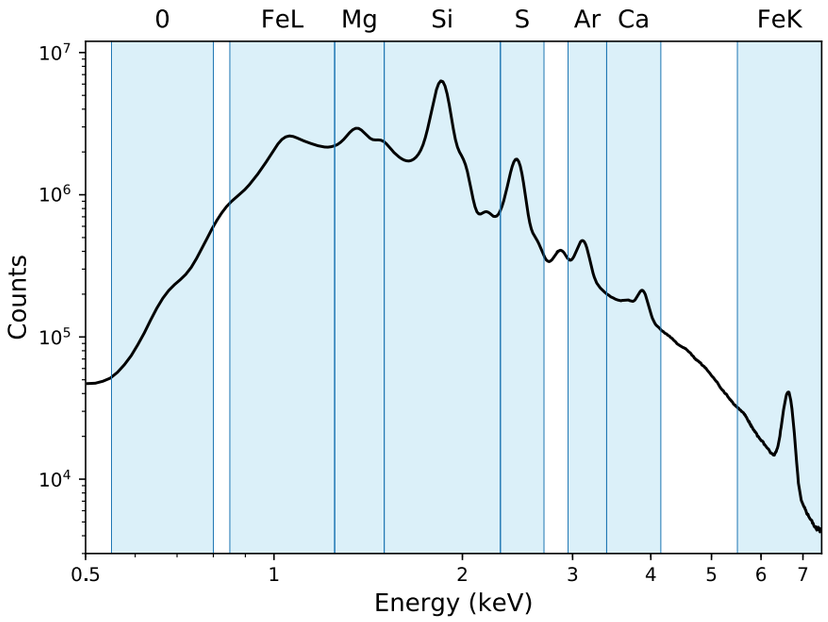




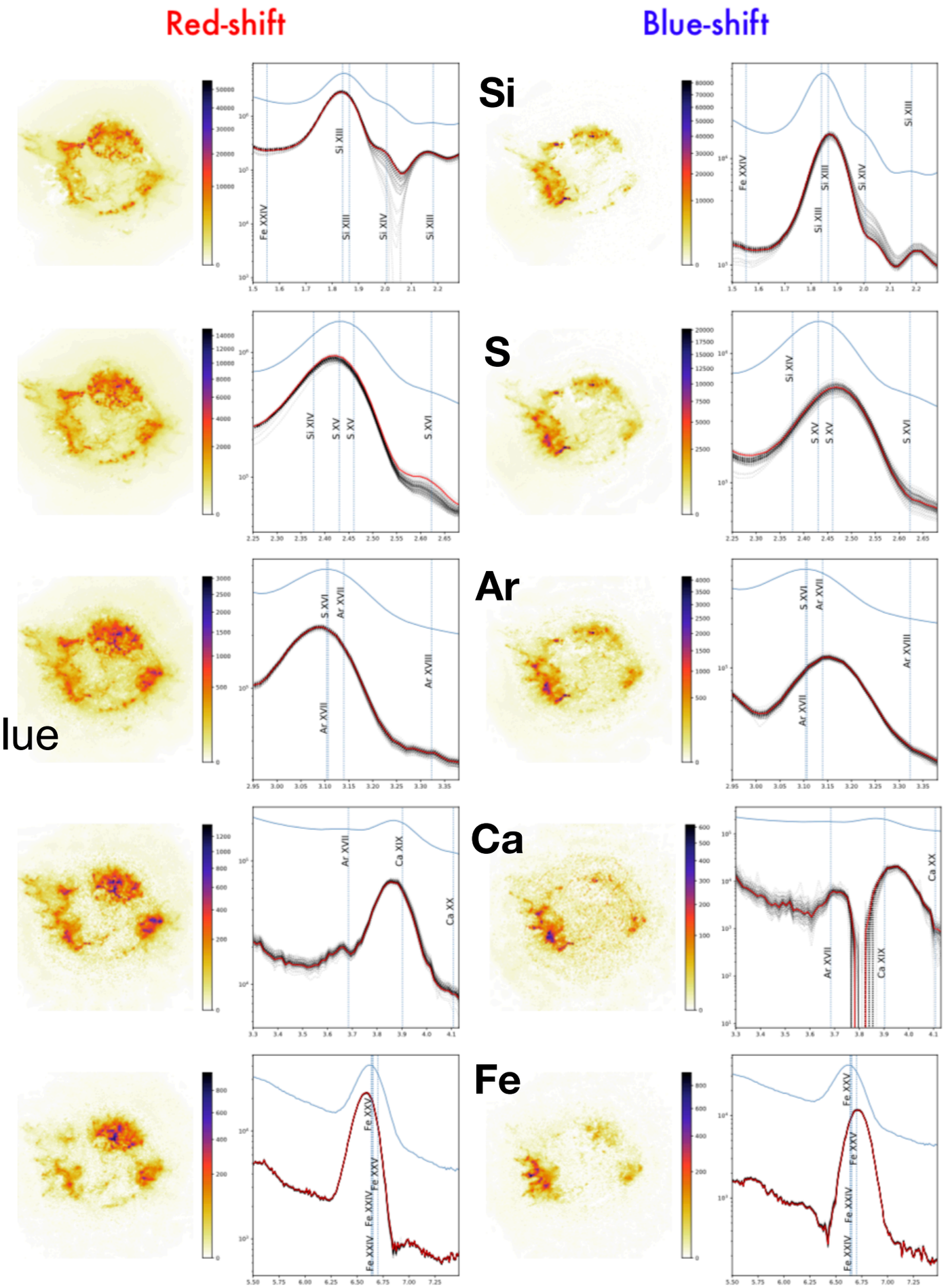
3D assymetries in CasA

Picquenot, Acero, et al., 2021

GMCA on Chandra 1 Ms subcubes

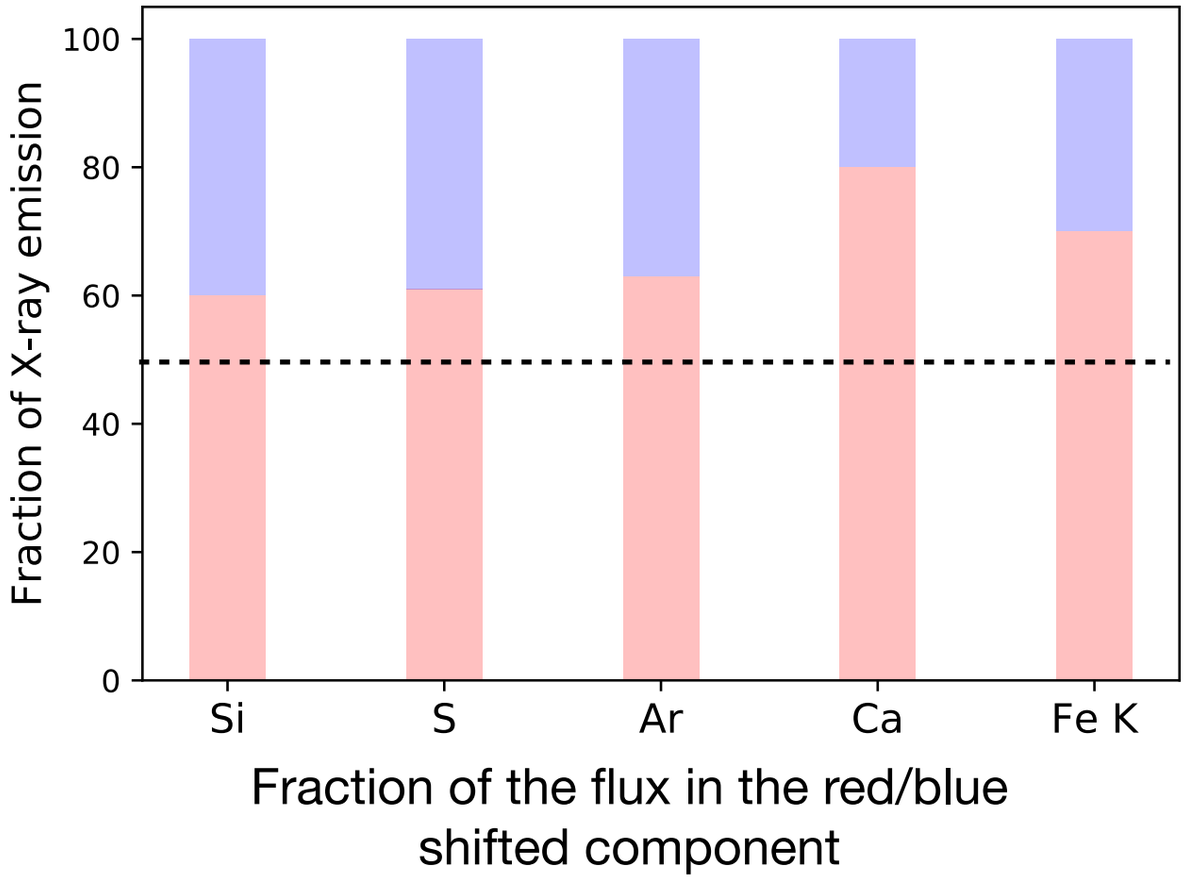


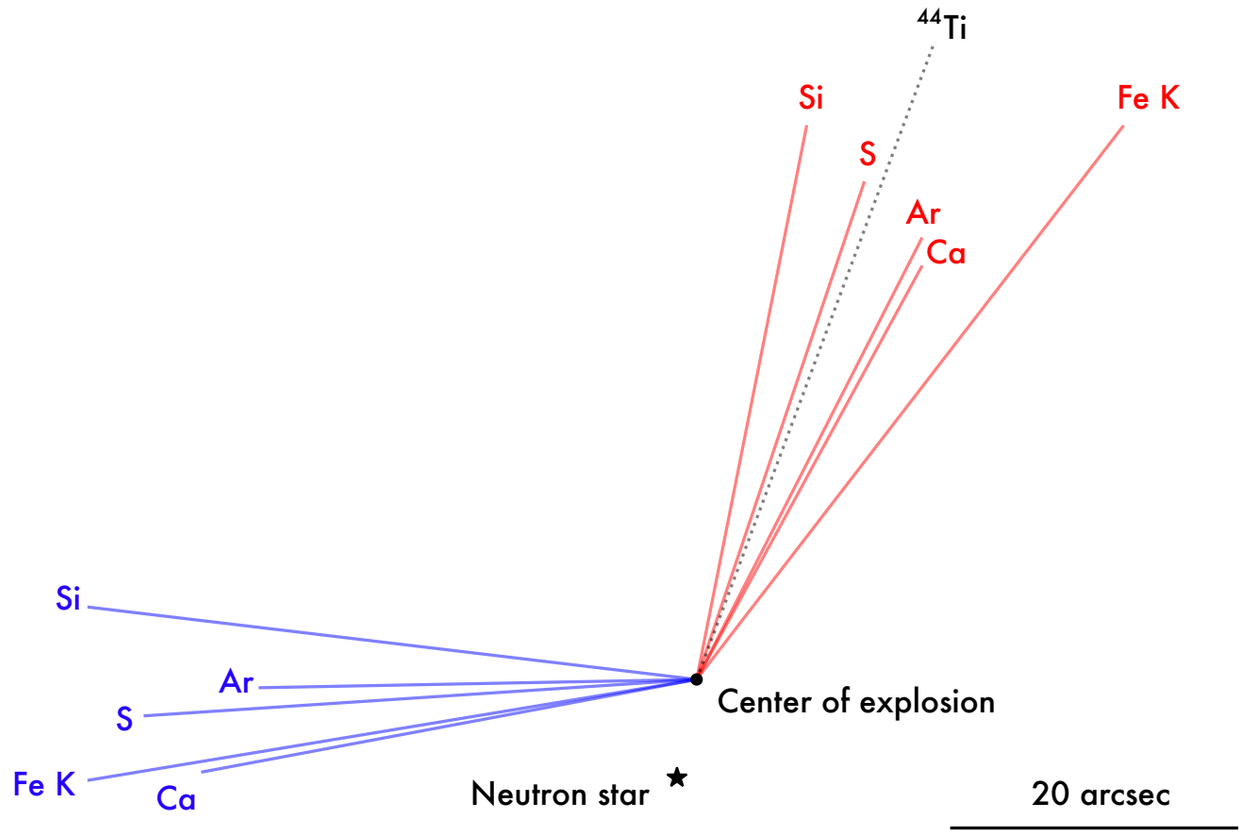
- For the Si, S, Ar, Ca, FeK bands two components that we interpret as red/blue shifted line emission from ejecta
- 3D X-ray maps element by element
- Study assymetries in the explosion



3D asymmetries in CasA

- **Most of the X-ray emission is red-shifted. Similar to ^{44}Ti**
- **Most of the material is expelled on the opposite direction of NS**

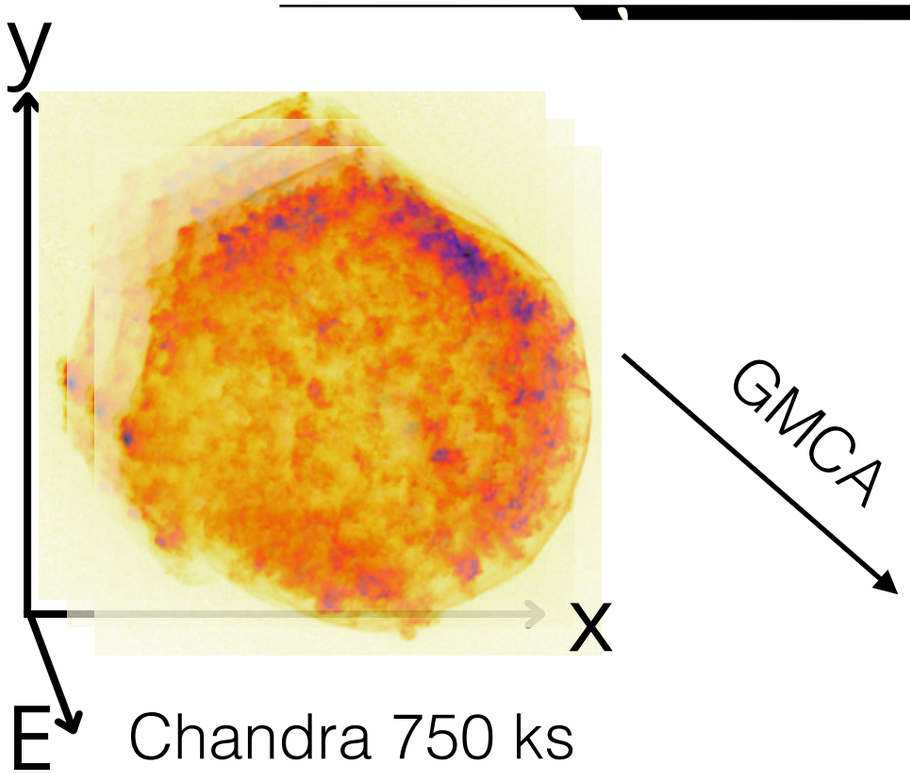




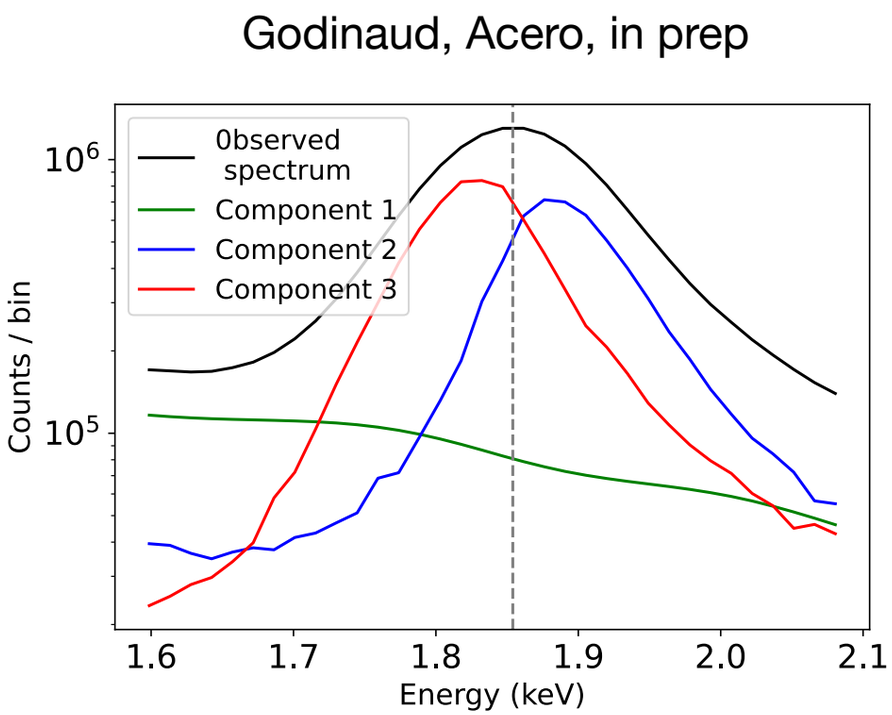
Picquenot, Acero, et al., 2021

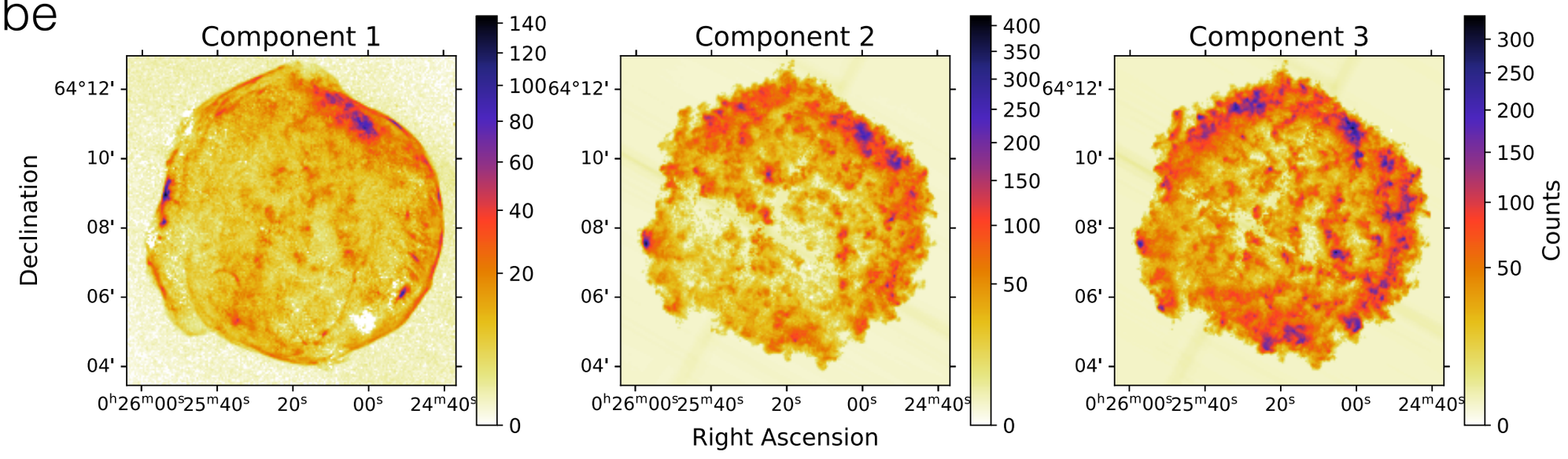
Red/blue shifted components are not in opposite directions disfavoring jet/counter jet explosion

Tycho SNR: asymmetries with 3D vector fields



Chandra 750 ks
1.6-2.1 keV cube





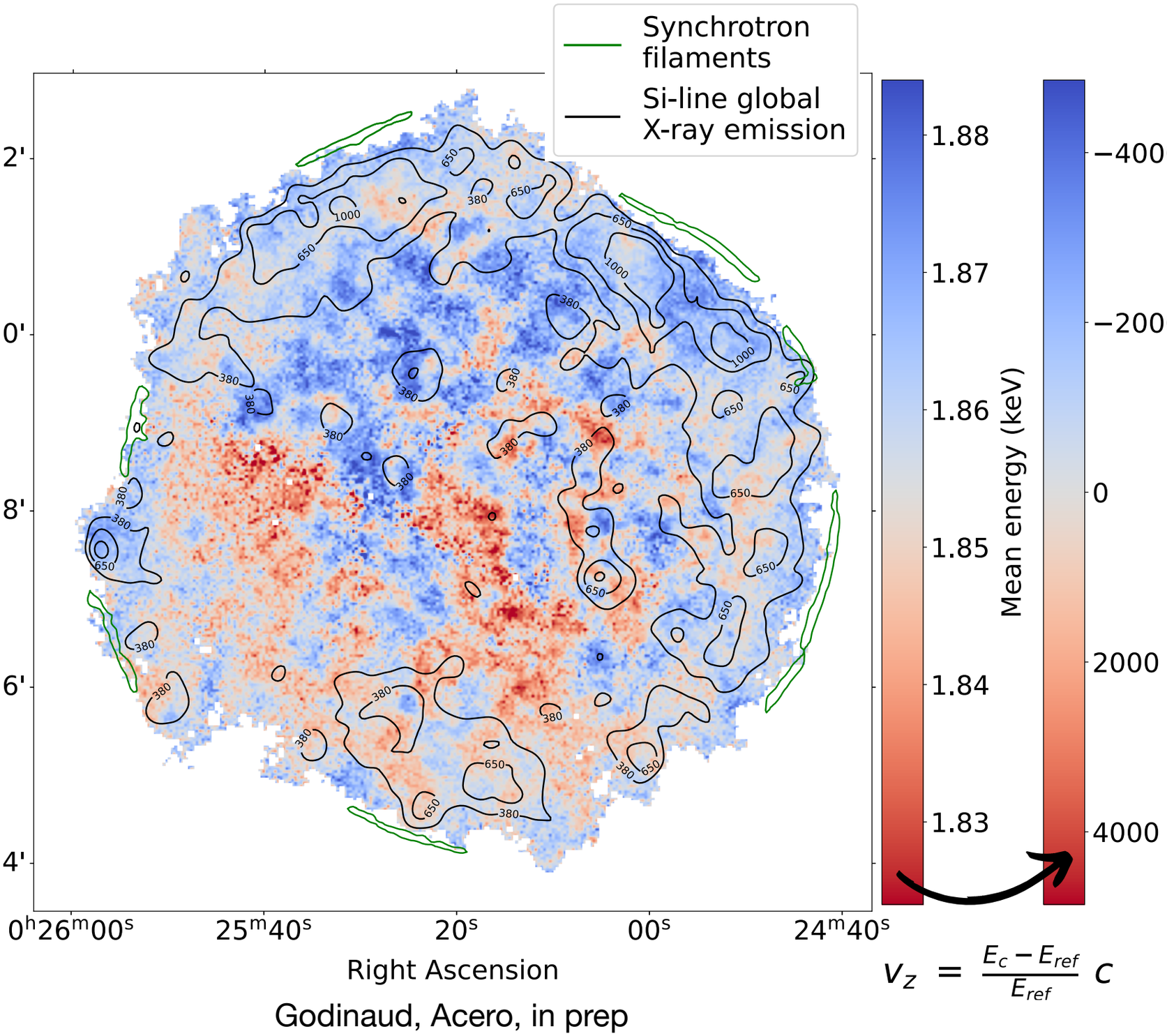
Components that we interpret as:

Synchrotron

Blue-shifted ejecta

Red-shifted ejecta

Tycho SNR: ejecta velocity along the line of sight (V_z)



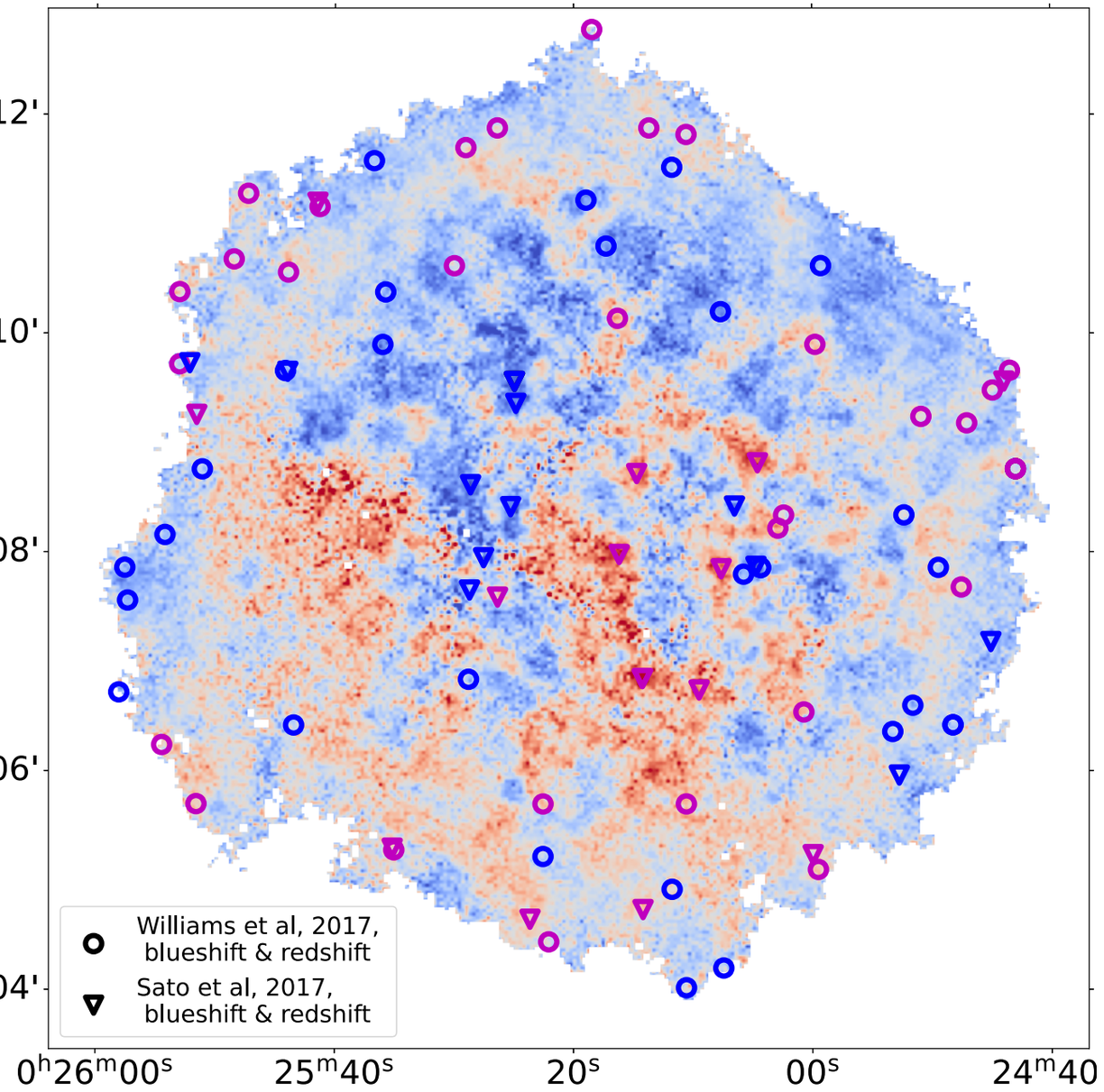
$$v_z = \frac{E_c - E_{ref}}{E_{ref}} c$$

Godinaud, Acero, in prep

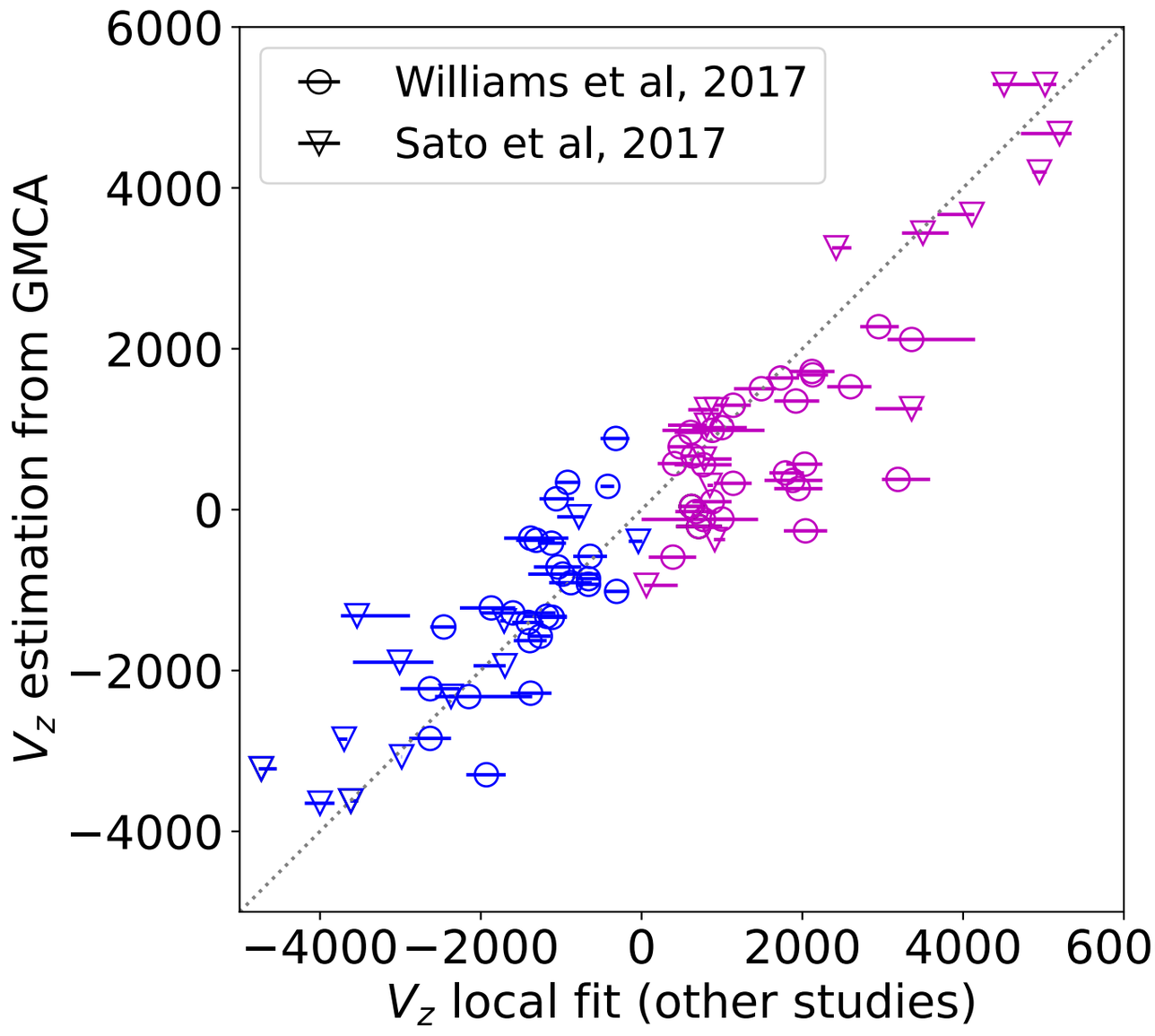
- **Idea: Use GMCA components (red/blue) as a basis to reconstruct the energy centroid in each pixel**
- **Complete coverage + pixel level**
- **Clear asymmetry of ejecta North/South**

Cross checking V_z with other spectral studies

- **Comparison of our method with dedicated spectral analyses**



Previous studies have ~50 regions

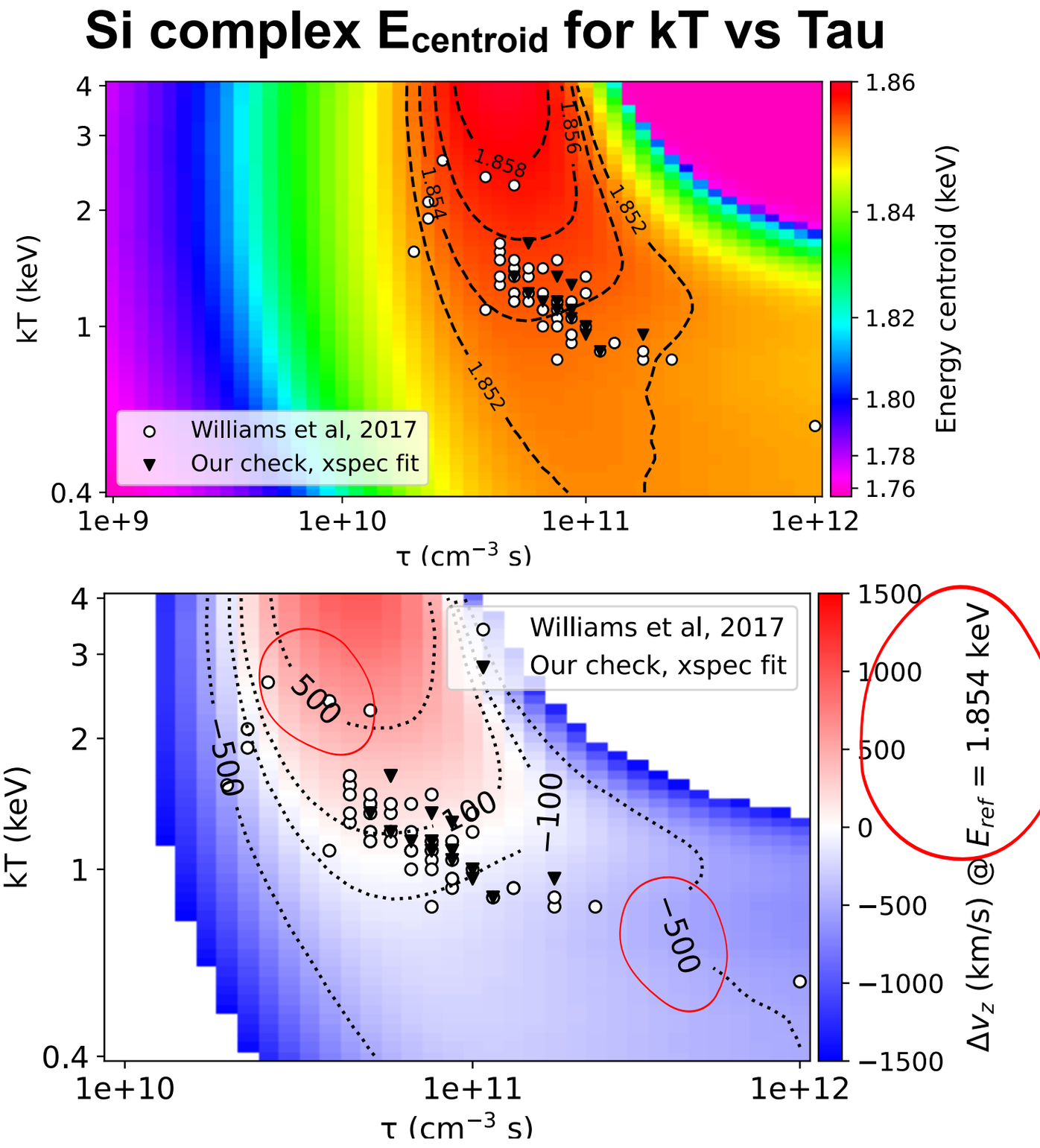


Good agreement between our V_z proxy and dedicated spectral fitting

Impact of choosing an energy of reference vs kT and Tau

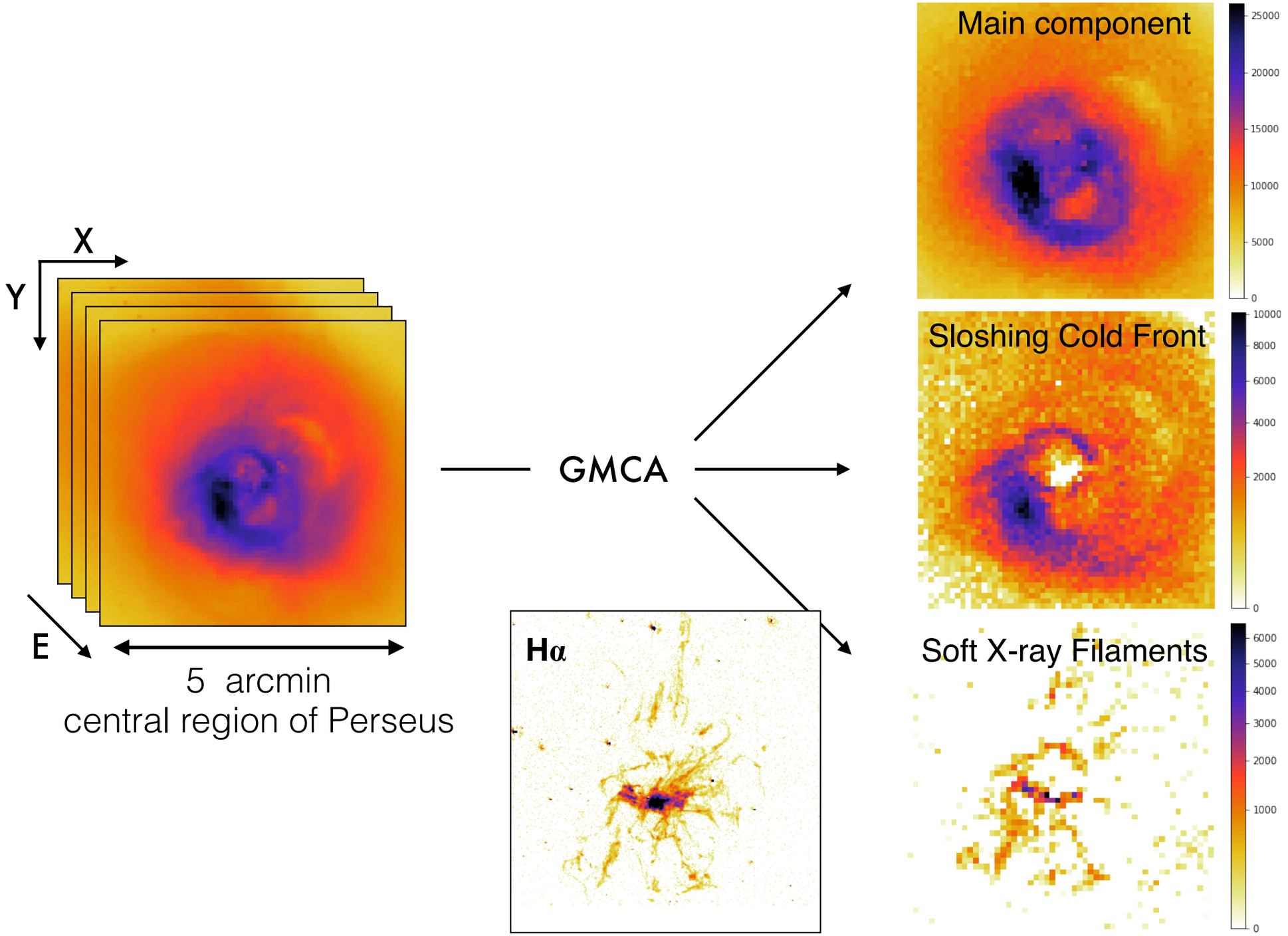
- **We take the mean corresponding energy centroid at 1854 eV**

- **Small impact of the variation of E_{ref} for V_z (maximal values of 500 km/s)**



Godinaud, Acero, in prep

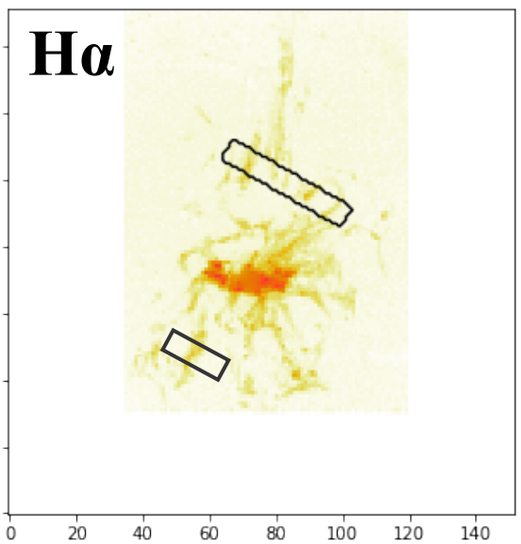
Galaxy cluster: Perseus 0.5-3 keV

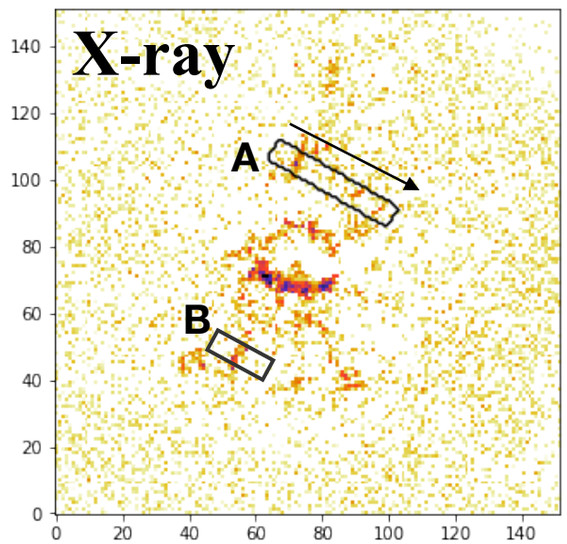


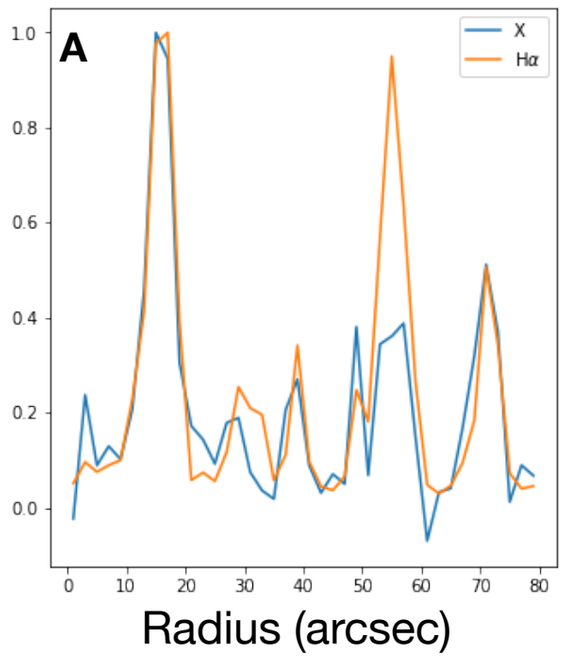
First time imaging of X-ray filaments
Ability to obtain image+spectrum

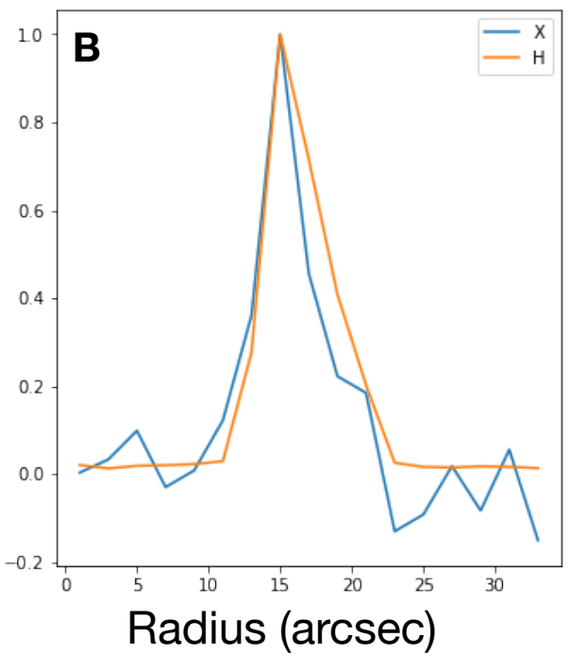
X-ray/Halpha morphological comparison

- **Same width filament width in Hα and X-rays**

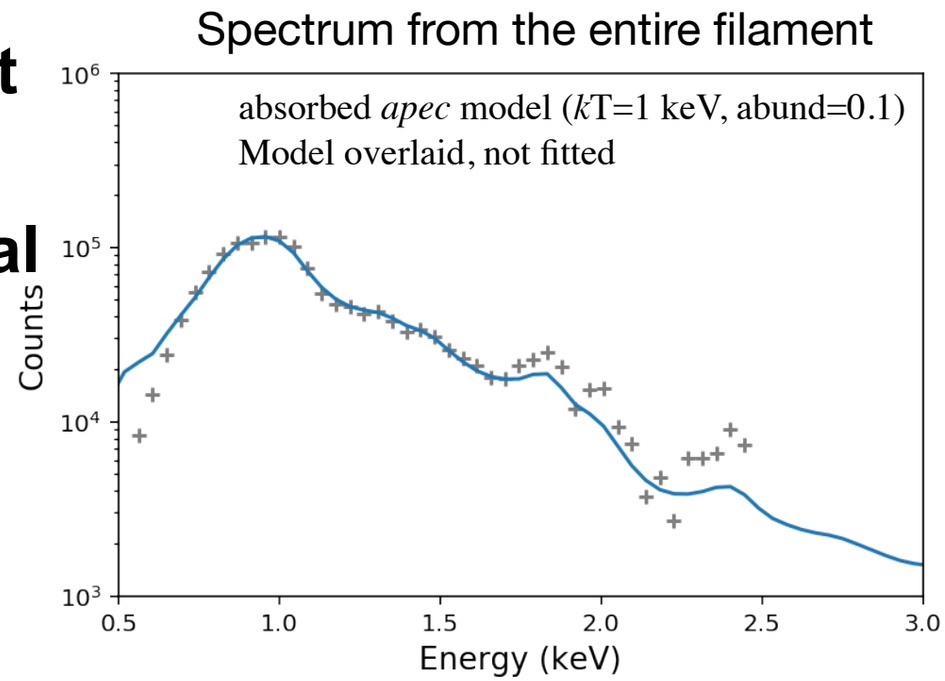








- **Mist of condensing gas. Not a stratified filament**
- **Similar abundance to intracluster gas**
- **Infalling matter from cluster gas not from central Galaxy outflow**
- **Thermal radiation from the cooling gas**

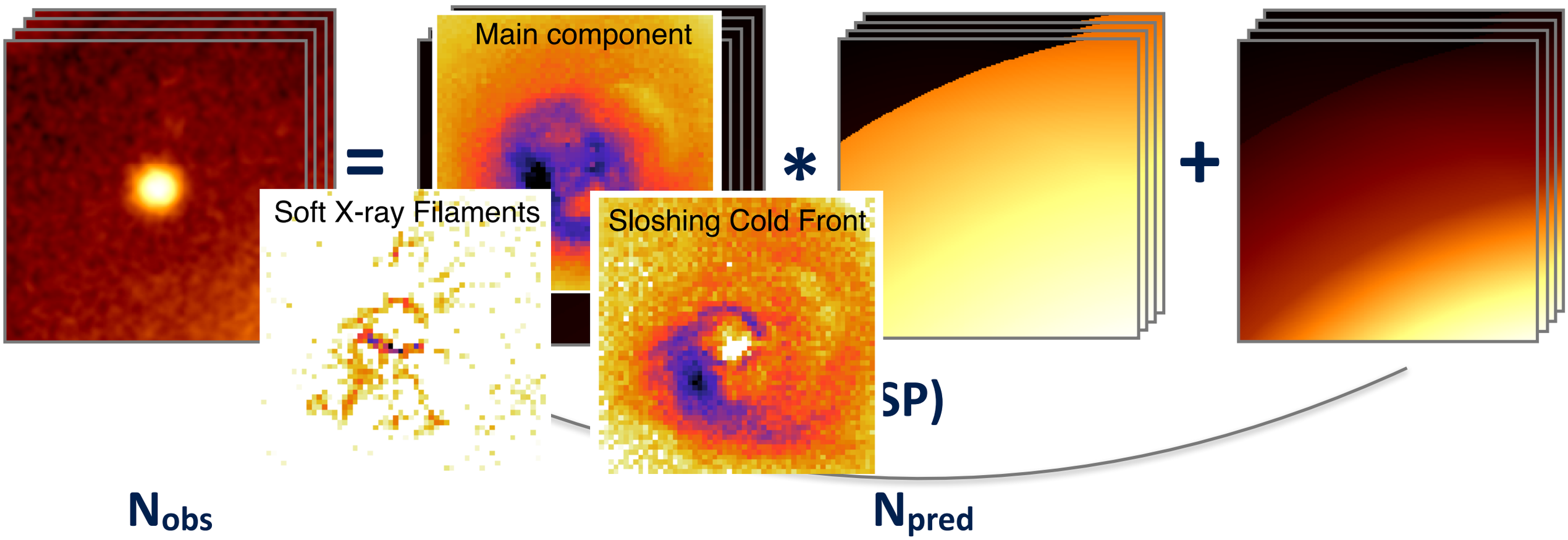


A 3D analysis for classical fitting of each Perseus component

Counts cube (X,Y,E) | Parametric source model | Exposure cube | Bkg cube

Free params:

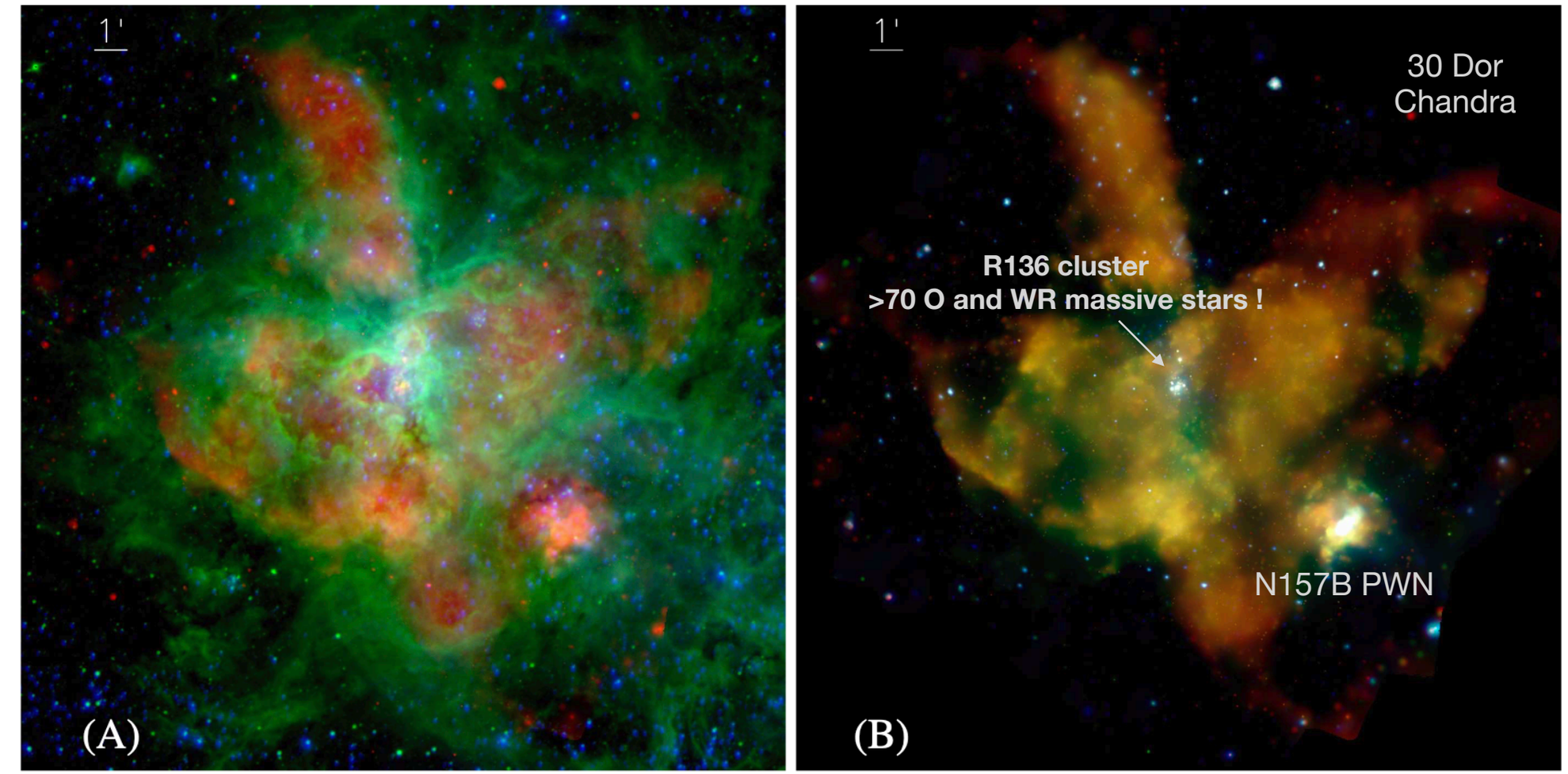
X, Y, Sigma, Norm, Γ



- **Minimization of -LogLikelihood(Npred, Nobs)**
- **Replacing spatial parametric model with GMCA spatial templates**

eROSITA analysis of EDR 30 Doradus

Collaboration with C. Maitra, M. Sasaki, J. Knies, P. Maggi



Yingjie Cheng 2021

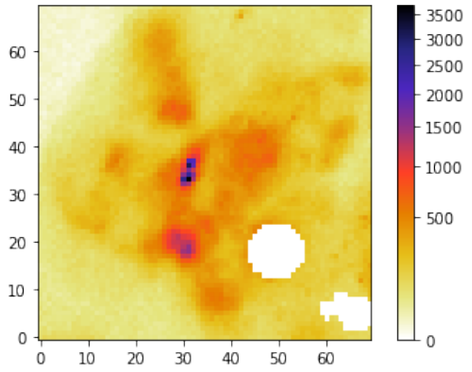
Figure 1. Multi-wavelength montages of 30 Doradus: (A) 0.5-2 keV (red), Hα (green), and UV (blue); (B) *Chandra* images in the 0.5-1 keV (red), 1-2 keV (green), and 2-8 keV (blue) bands. The Hα image was taken with the 4m telescope at Cerro Tololo Inter-American Observatory (Chu & Kennicutt 1994), while the UV image was from the GALEX survey.

Manami et al., 2021
arXiv:2106.14530

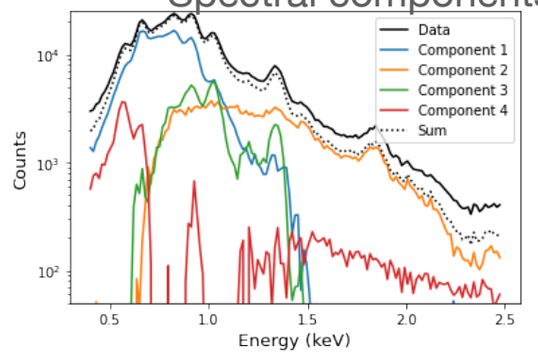
eROSITA analysis of 30 Doradus: 4 components

Collaboration with C. Maitra, M. Sasaki, J. Knies, P. Maggi

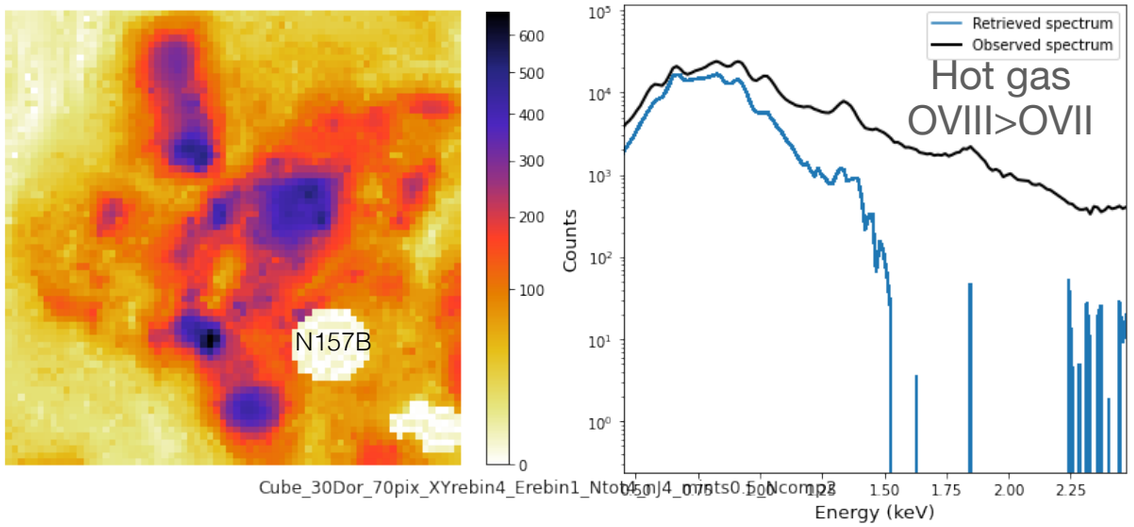
Summed data cube over all energies

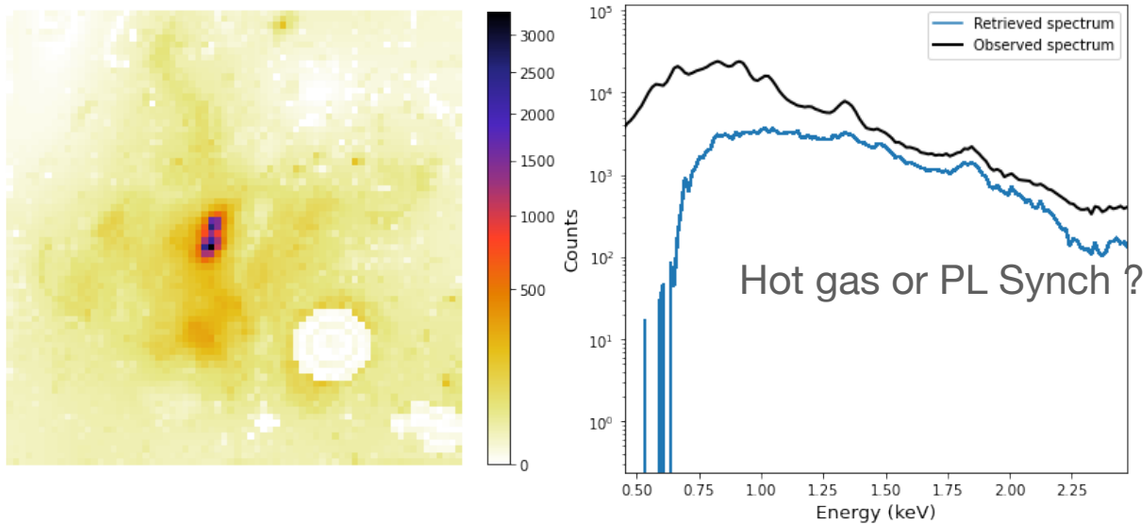


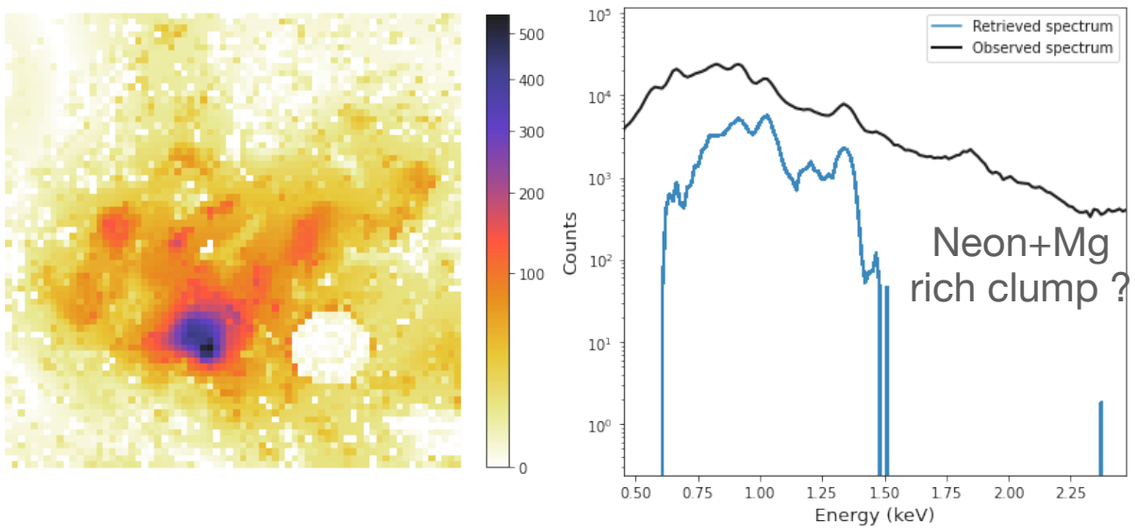
Spectral components

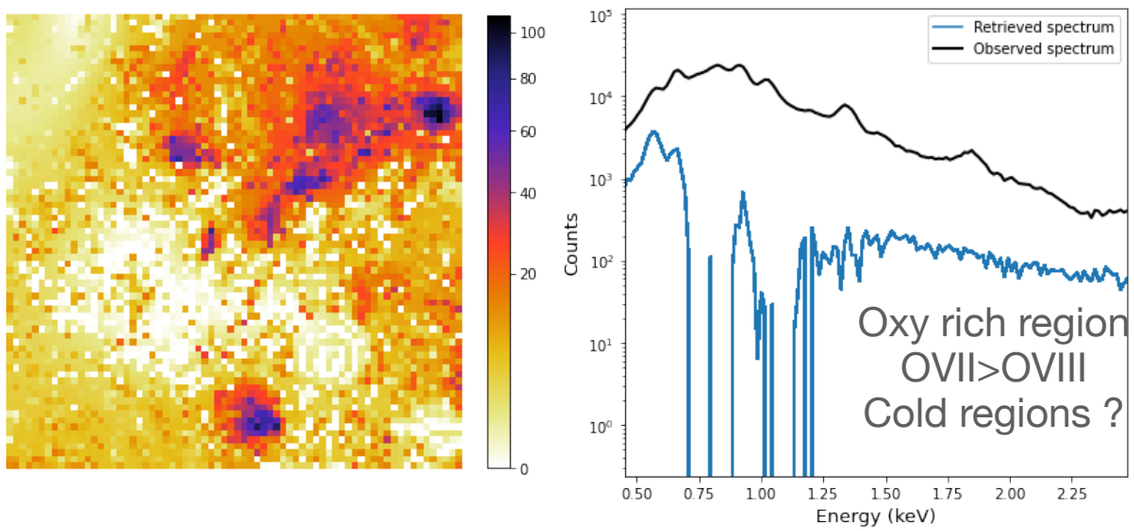


Cube_30Dor_70pix_XYrebin4_Erebin1_Ntot4_nJ4_mints0.5_Ncomp0



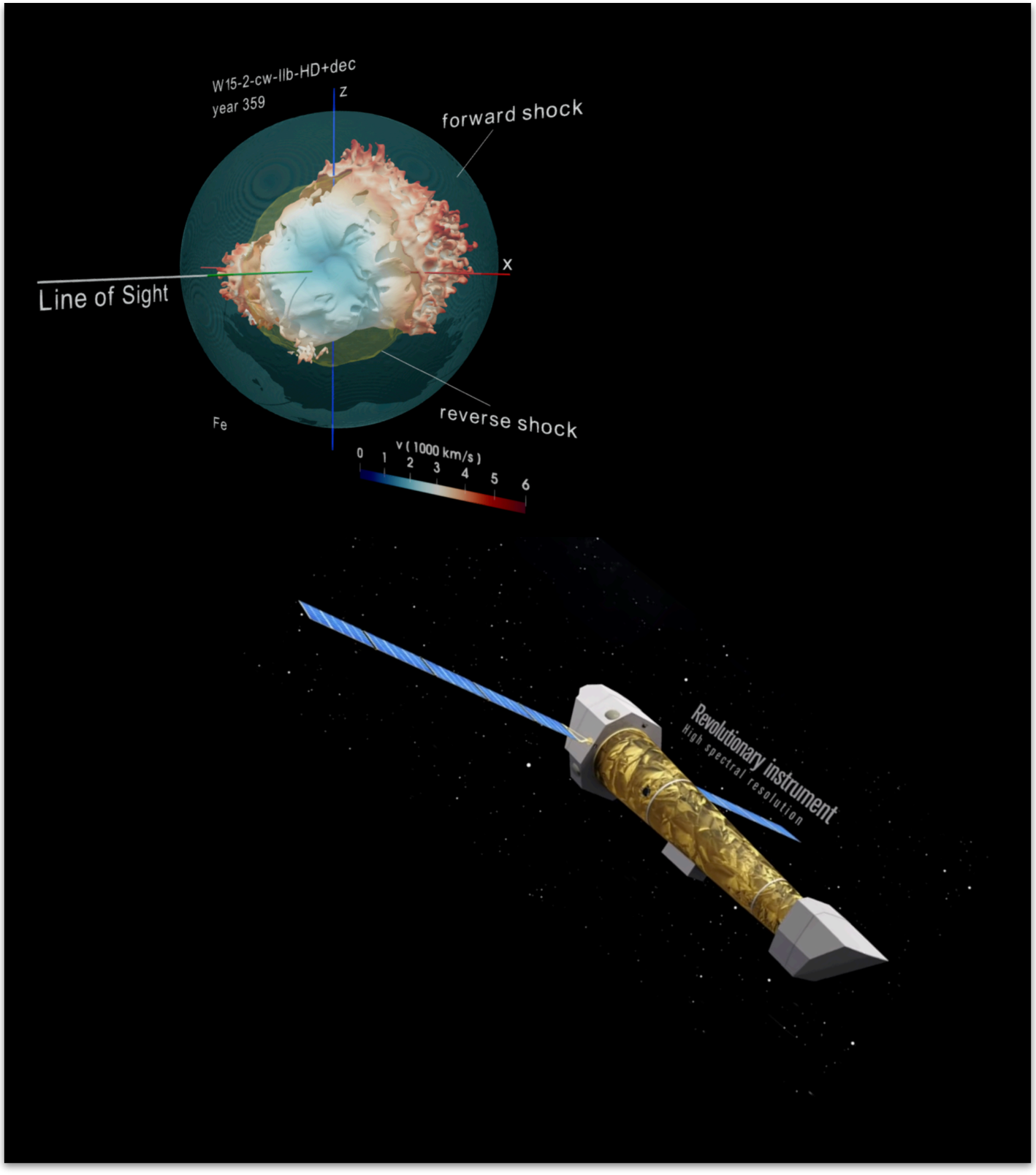






From numerical simulations to synthetic X-ray observations

The case of Athena X-IFU

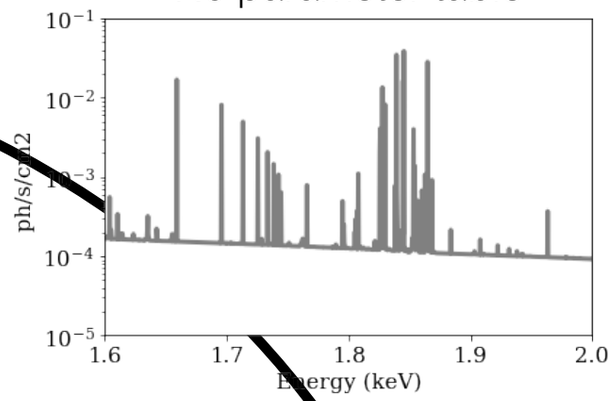


Pipeline for X-ray simulations of extended sources

Collaboration with C. Kirsch, J. Wilms

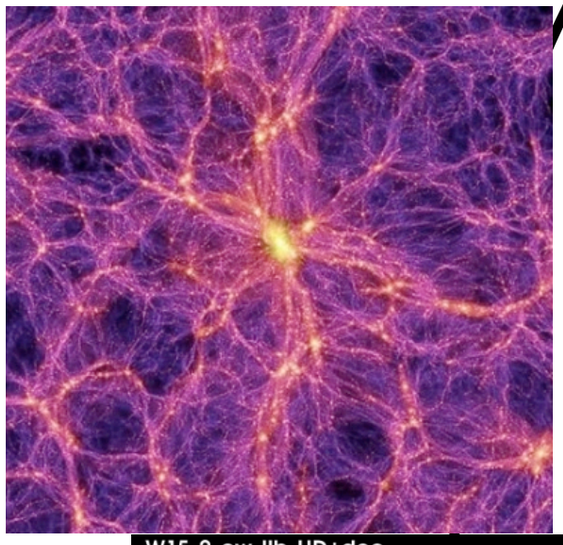
Plasma emission models

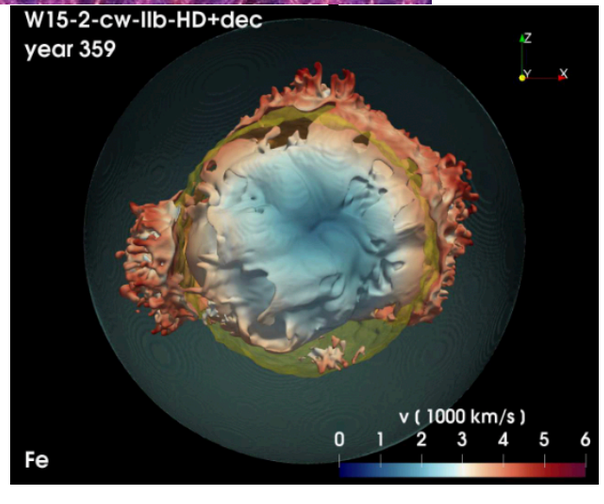
X-ray flux hypercubes (X,Y,Z,E)
No parameter table



Numerical simulations

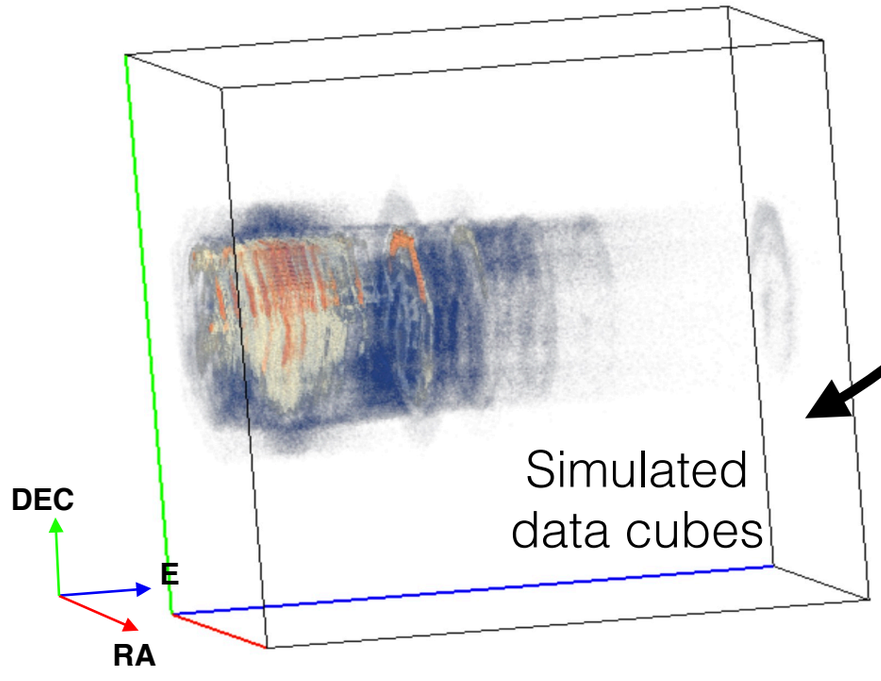
Temperature, density, velocity





SIXTE

Simulation of X-ray telescopes with instrument performance



Simulated data cubes

Tomography slicing through the cube @ Si lines

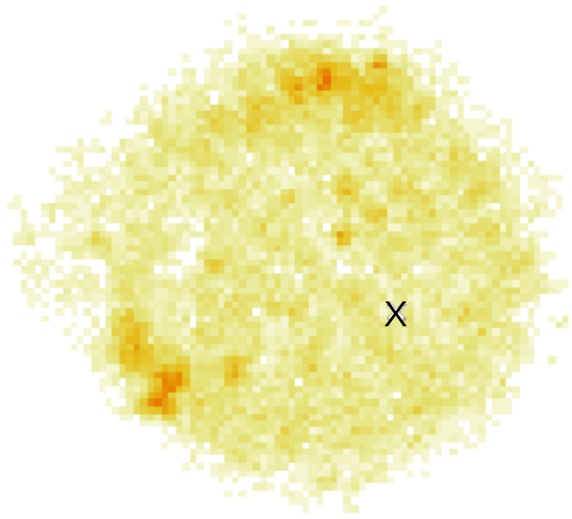
3D numerical simulation of a CasA-like SNR (Orlando et al.)

Viewed by in 10 ks by

XIFU 4” pixels

FWHM=5”

Athena X-IFU

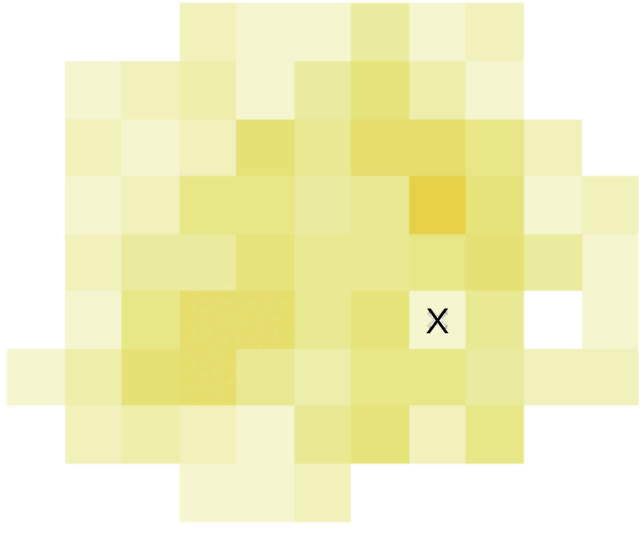


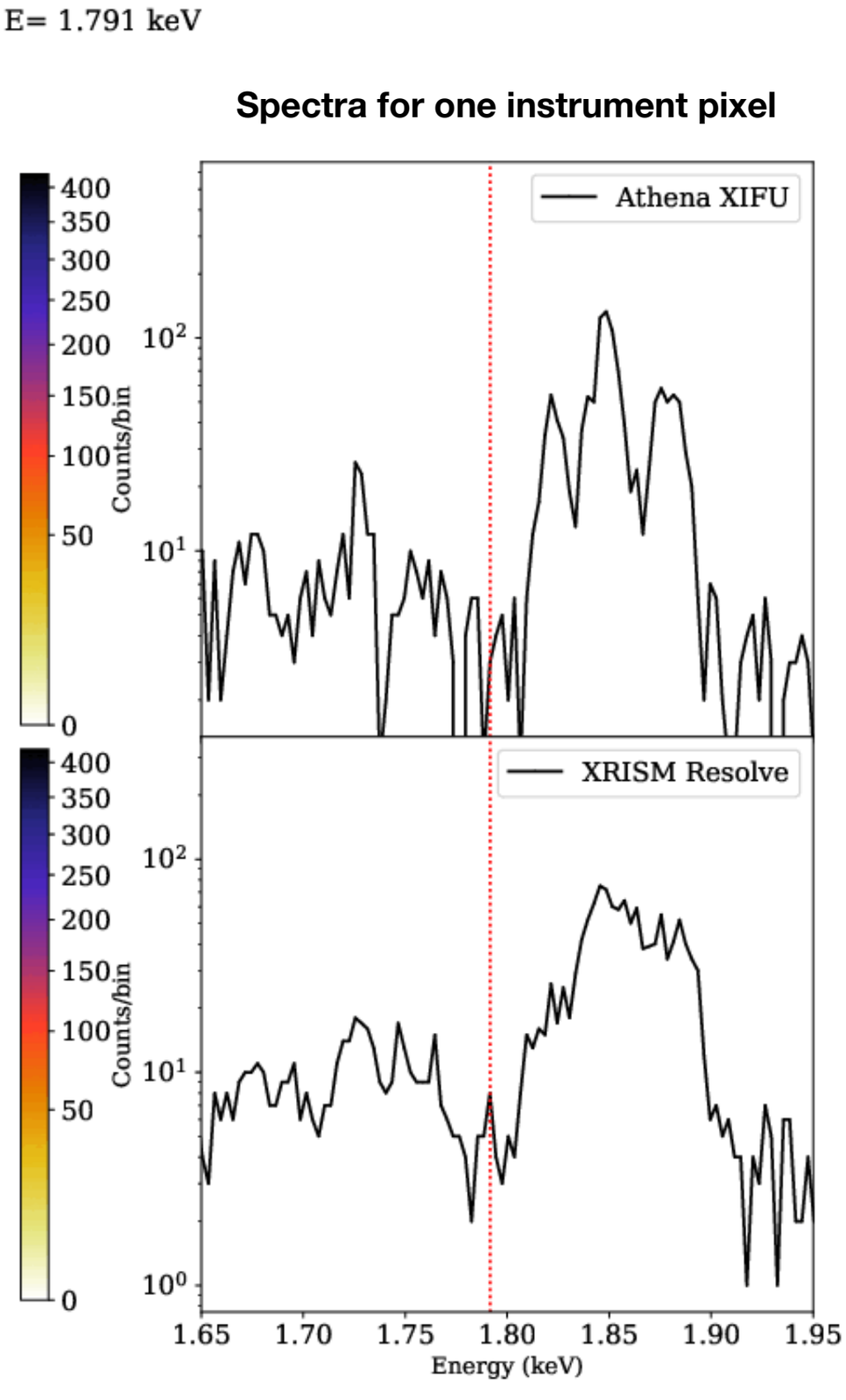
F. Acero

XRISM Resolve

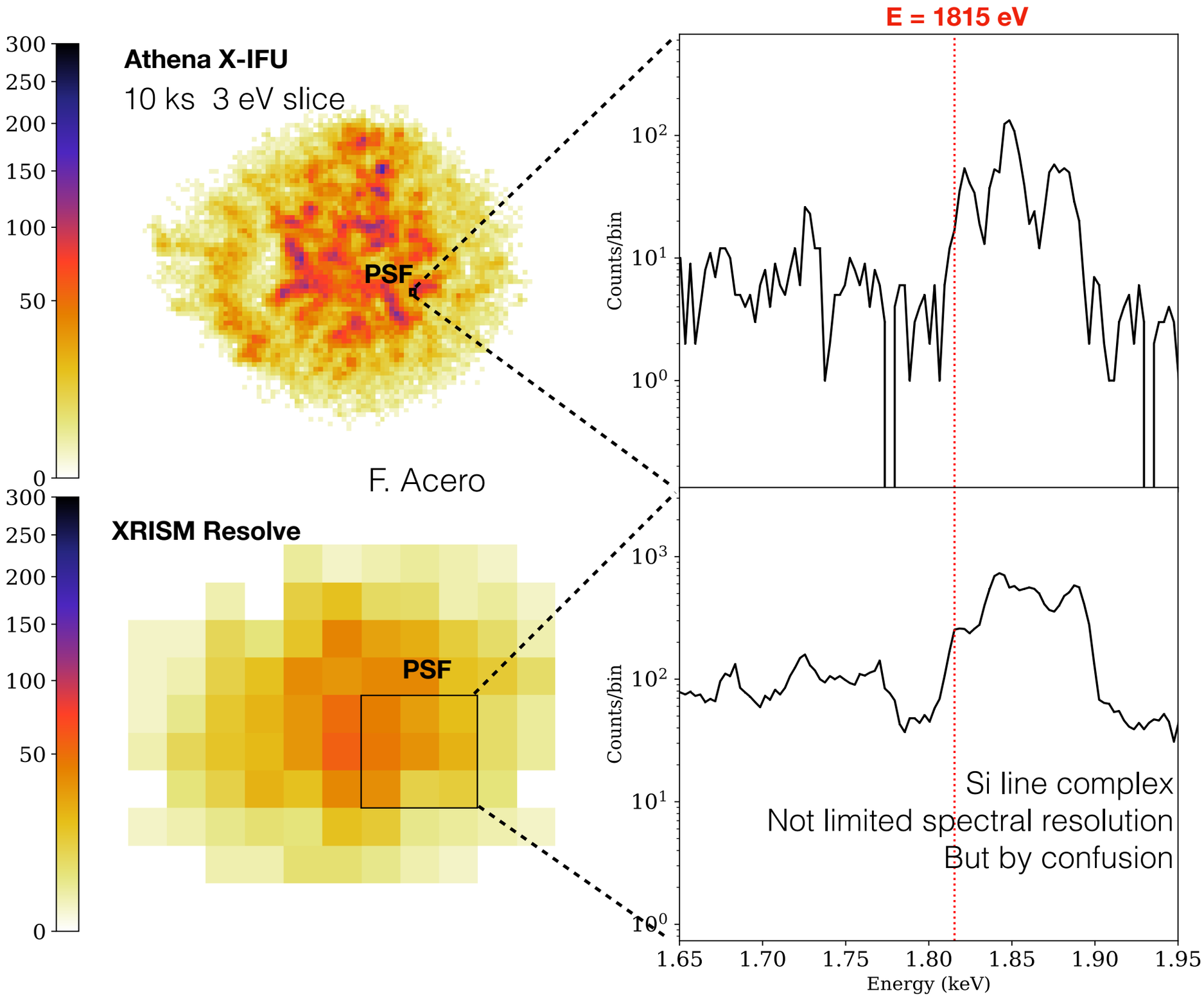
XRISM Resolve 30” pixels

FWHM=1.5 arcmin





The need for HR imaging for HR spectroscopy



Conclusion

- **Wealth of information in deep X-ray observations but analysis methods have stalled in the last 20 years**

- **GMCA: a blind source separation method for X/Gamma-rays**
  - **Exploit the 2D-1D info from spectro-imagers (MUSE, Fermi, CTA)**
  - **Applicable to SNR, PWN, clusters, galactic center or time cubes (X, Y, T)**

- **GMCA retrieves the overlapping physical components with morphological spectral diversity**
  - **eROSITA provides very large mosaic and great spectral resolution. In several degree scale image, great spatial resolution & high statistics !**

- **Caveats:**
  - **Each component is Spec*Image. No spatial variation for now**
  - **Need deep observations of bright sources (> 1 millions counts)**
  - **Can only differentiate if structured emission (not ~gaussian blobs)**

Perspectives

- **Perspective :**
  - **Deep learning for a learnt representation of spectral shapes, ==> mapping physical parameters at pixel level**
  - **Component separation with spatial variation (e.g. synchrotron + varying redshift component)**

- **Athena :**
  - **new tools are needed to fully exploits the wealth of data**
  - **=> benchmark of new tools on realistic simulations**
  - **Side note for Athena Redesign : for extended sources HD imaging is needed for HR spectra to reduce confusion.**

Archival data & future high spectral resolution instrument are gold mines.
Need to explore next generation analysis tools.

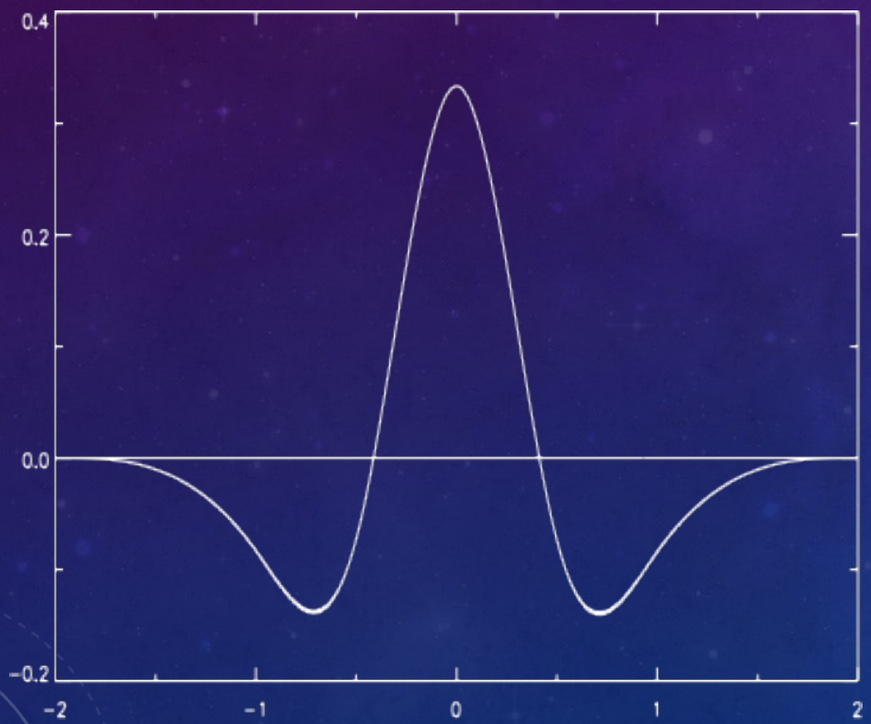
Back up slides

The starlet transform

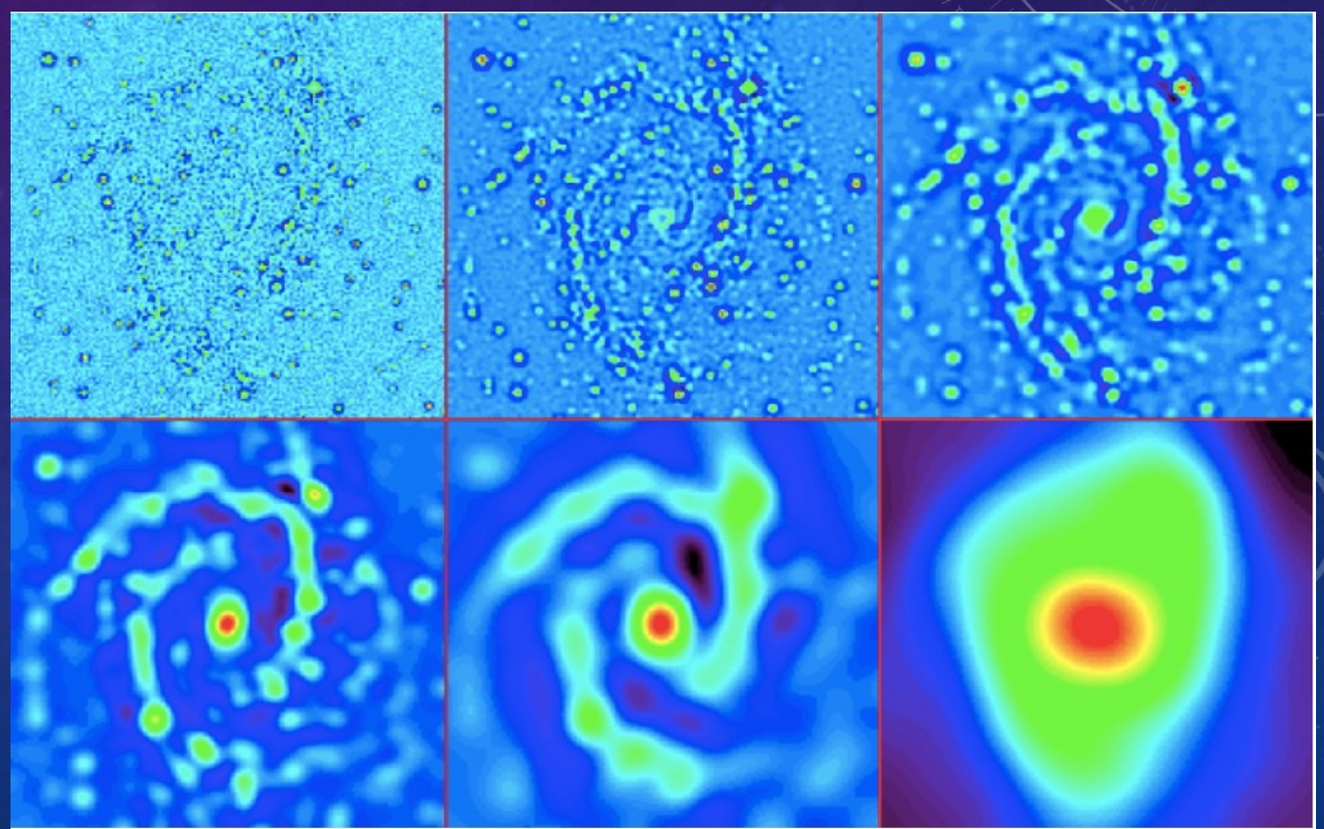
STARLET TRANSFORM

- Isotropic undecimated wavelet transform
- B3-spline for the scaling function

$$X = c_J + \sum_i^J w_i$$



the wavelet



from Starck [2011]

15

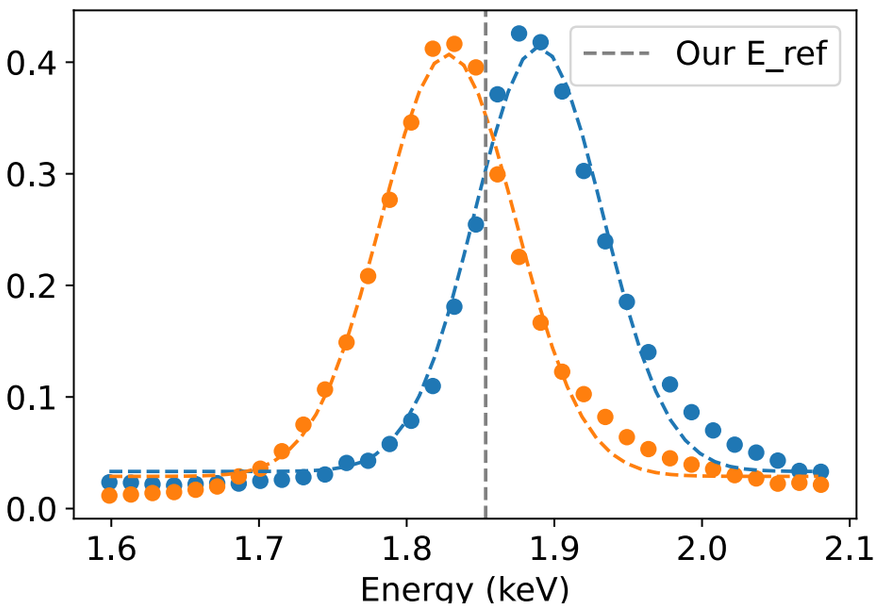
From Red/Blue maps to centroid energy maps

Method to obtain Ec map

1) We use GMCA's definition to "reconstruct" the spetrum in each pixel (i,j)

$$Spectrum_{i,j} = \sum_{Component\ k} Image_{GMCA,k,i,j}\ Spectrum_{GMCA,k}$$

2) We fit the GMCA spectrum with a gaussian.
We obtain an analytical expression of these spectra and so, of the spectrum in each pixel.



3) To find the maximum of the silicon line Ec in the spectrum of each pixel, we must to solve this equation :

$$\frac{d\ Spectrum_{i,j}\ (E_c)}{dE} = 0$$

4) And an analytical proxi of the solution is :

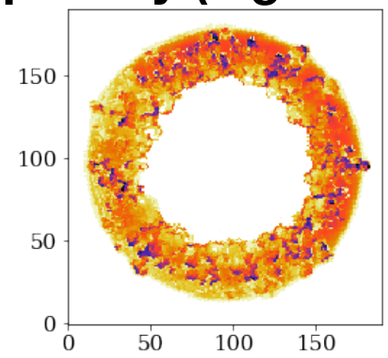
fit parameters from GMCA's spectrum red and blueshifted

$$E_c = \frac{Image_{ij,r}\frac{\alpha_r}{\sigma_r^2}\ E_{r,mean} + Image_{ij,b}\frac{\alpha_b}{\sigma_b^2}\ E_{b,mean}}{Image_{ij,r}\frac{\alpha_r}{\sigma_r^2} + Image_{ij,b}\frac{\alpha_b}{\sigma_b^2}}$$

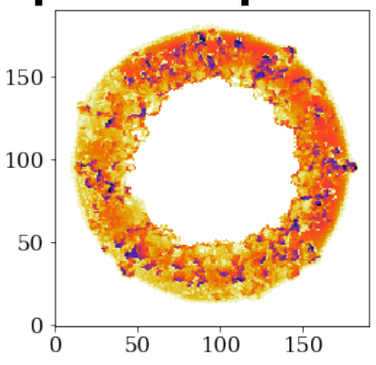
GMCA's image red and blueshifted

From simulation to observation :

Simulation hydro quantity (e.g. density)



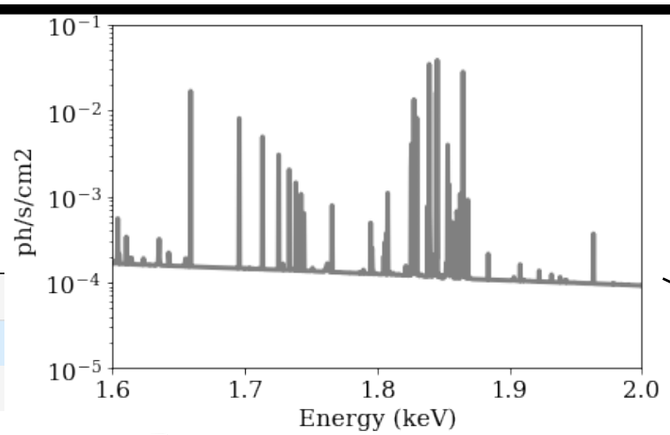
Post-processing for plasma quantities



Plasma spectral properties

Parameter	Thawed	Value
nH	✓	1.0
kT	✓	1.0
H		1.0
He		1.0
C		1.0
N		1.0
O		1.0
Ne		1.0
Mg		1.0
Si		1.0
S		1.0
Ar		1.0
Ca		1.0
Fe		1.0
Ni		1.0
Tau	✓	100000000000.0
redshift		0.0
norm	✓	1.0

Using spectral package as Xspec or SPEX

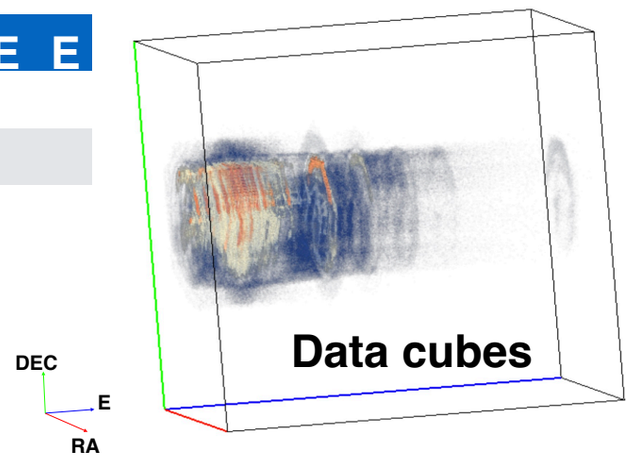


Synthetic X-ray spectrum for each cell

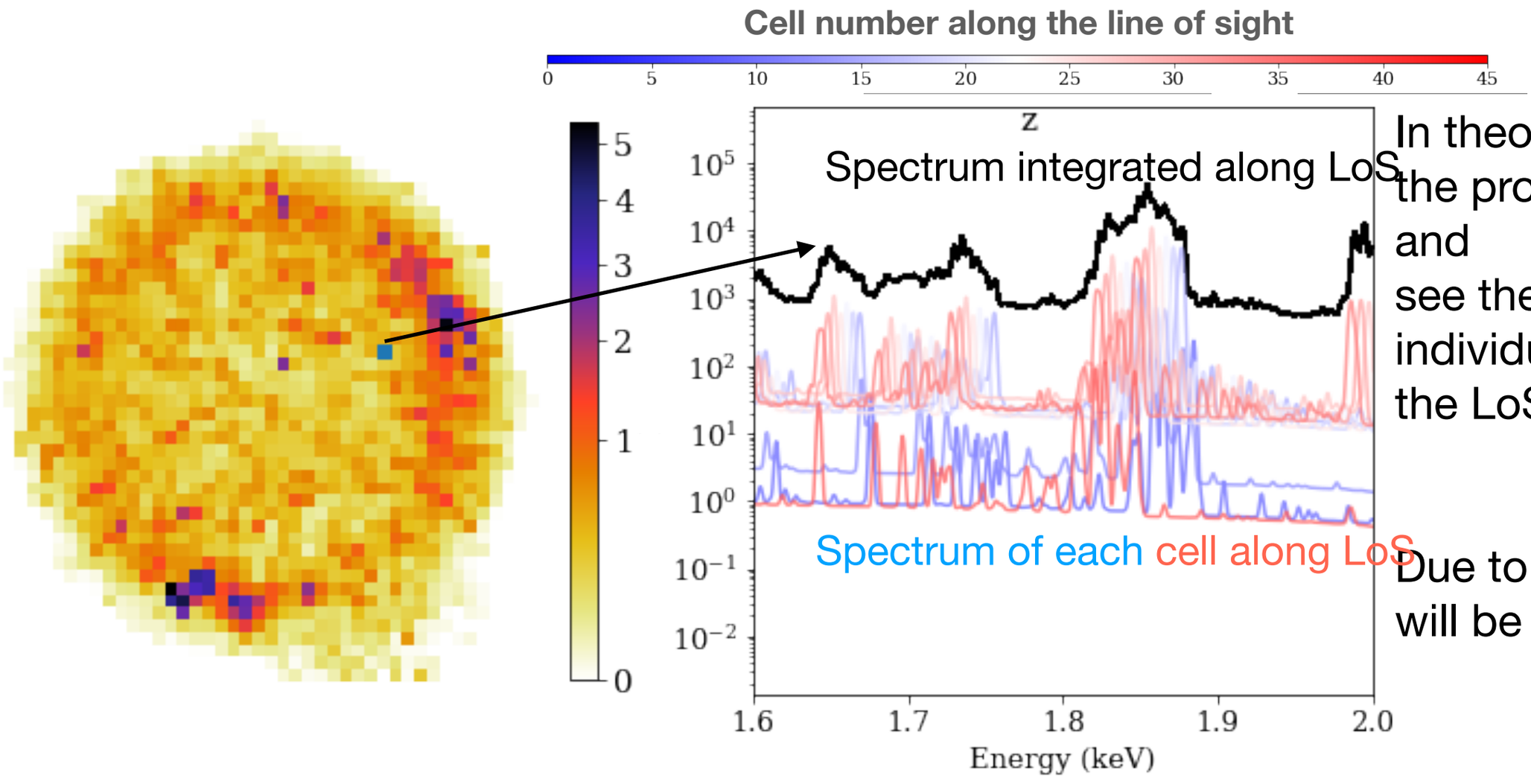
SIXTE +Instrument configuration

Event list

	R	DE	E
1			
2			
3			



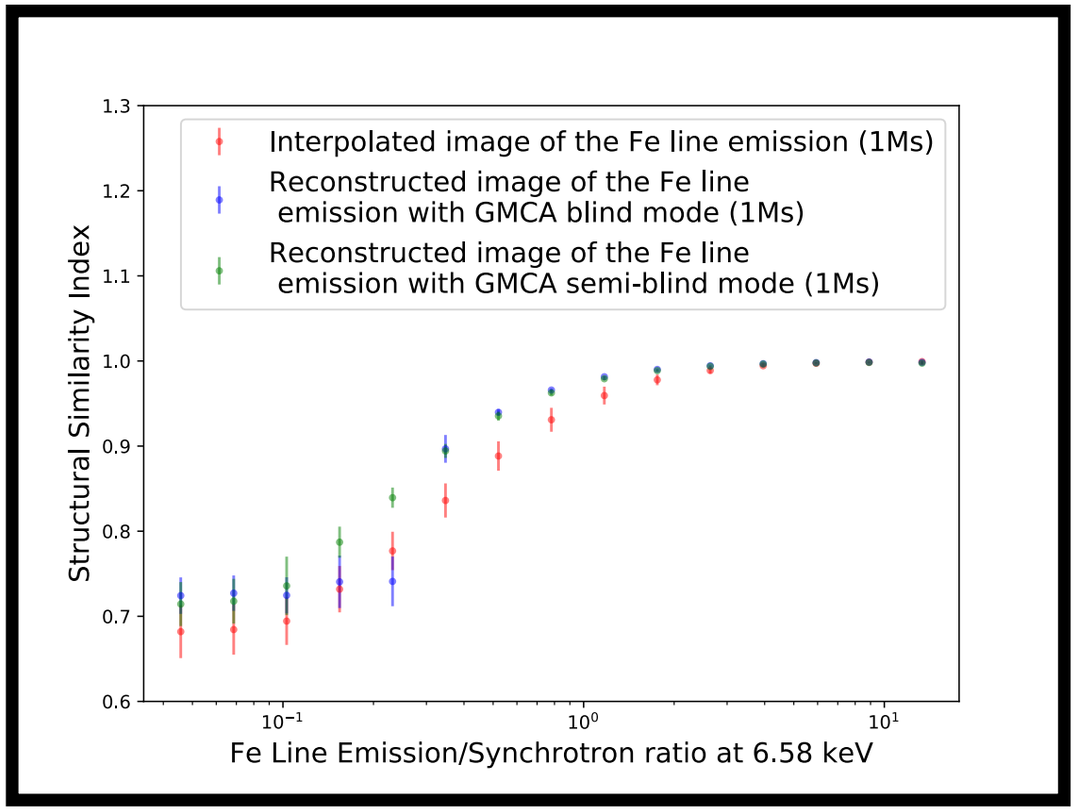
Studying the projection effects and mixing

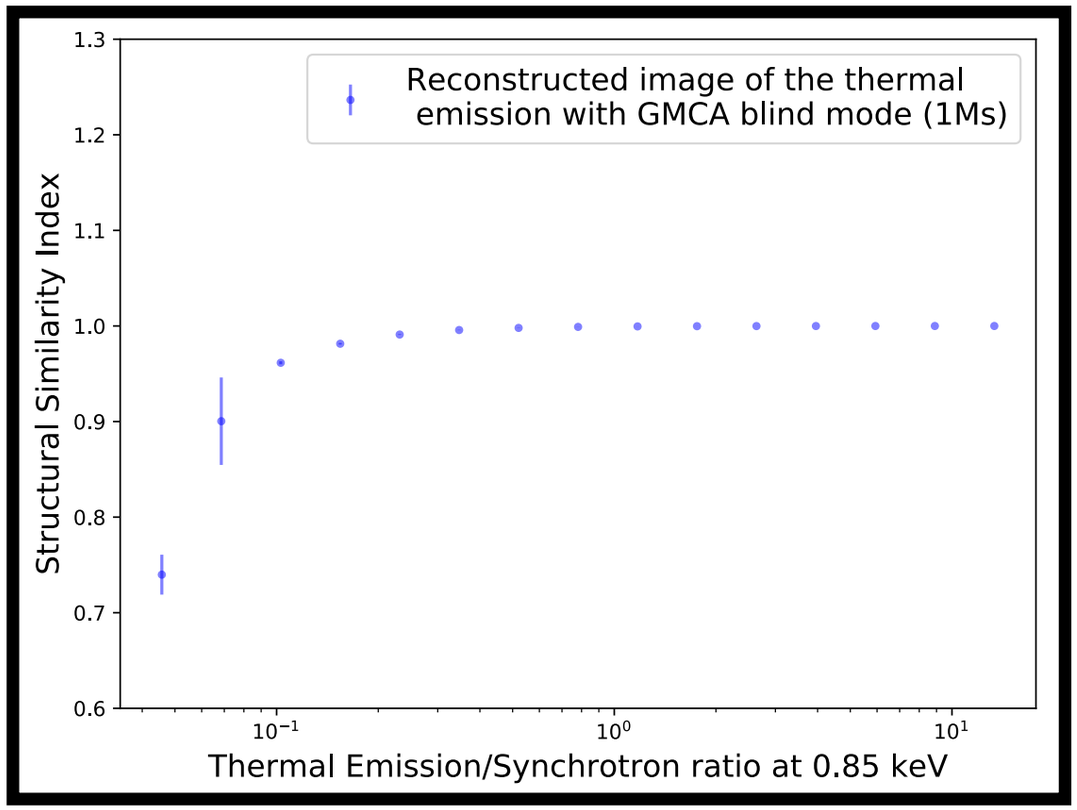


In theory we can probe the projection effects and
see the contribution from individual celles along the LoS

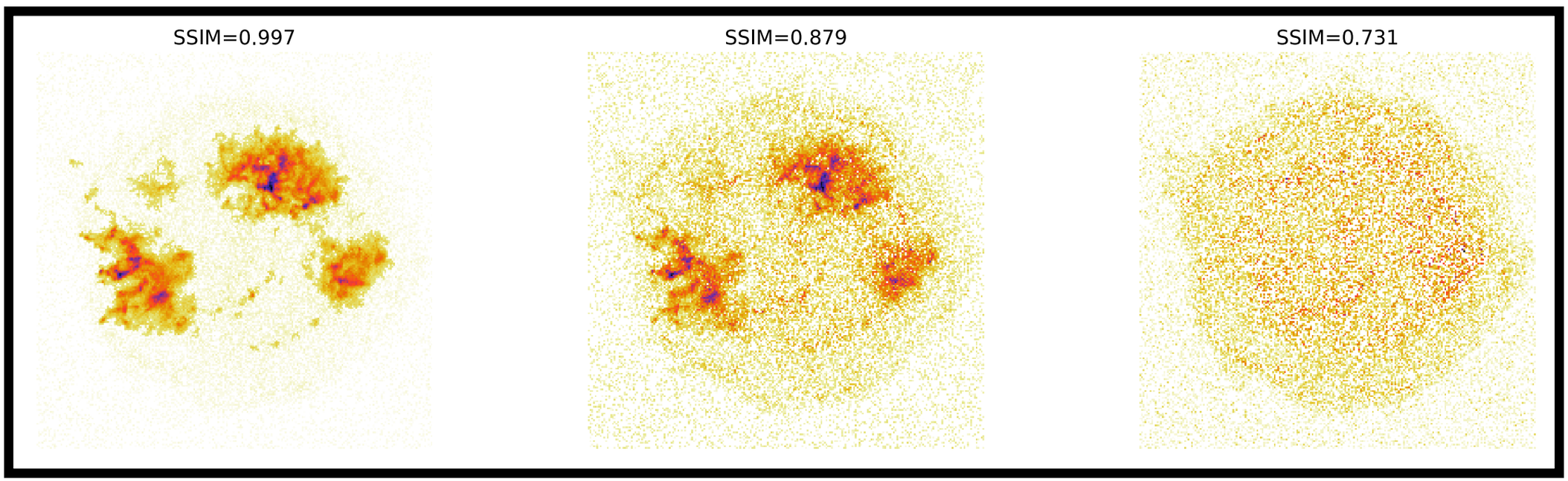
Due to LoS, line complex will be broadened.

Image accuracy





SSIM coefficients of the images of the retrieved second component in both toy models



Examples of SSIM coefficients associated with the corresponding images