Astronomical source detection and background separation via hierarchical Bayesian nonparametric mixtures

Andrea Sottosanti

October 2, 2018

University of Padua - Department of Statistical Sciences
sottosanti@stat.unipd.it

Collaborators

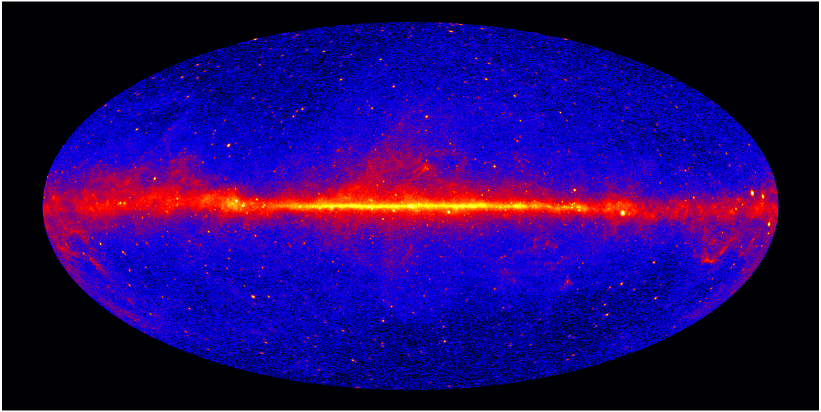
From **University of Padua**:

- Mauro Bernardi
- Alessandra R. Brazzale

from **Imperial College London**:

- David van Dyk
- Roberto Trotta
- Alex Geringer-Sameth
- David Stenning

High-energy astronomical count maps



(Image Credit: NASA/DOE/Fermi LAT Collaboration)

High-energy astronomical count maps

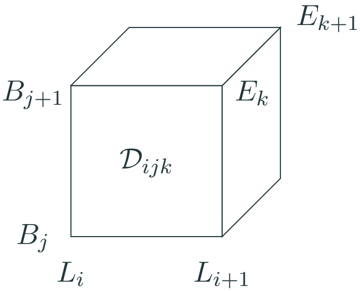
- Let $\mathcal{D}$ be a $I \times J \times K$ array of pixels:

  $\mathcal{D}_{ijk} = \#$photons spatially projected into ij with energy level k.

High-energy astronomical count maps

- Let $\mathcal{D}$ be a $I \times J \times K$ array of pixels:

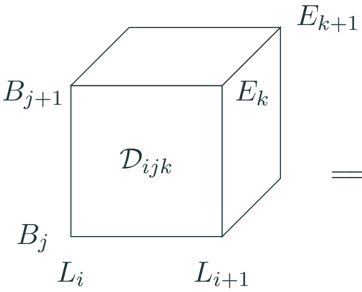
  $\mathcal{D}_{ijk} = $ #photons spatially projected into ij with energy level k.



High-energy astronomical count maps

- Let $\mathcal{D}$ be a $I \times J \times K$ array of pixels:

  $\mathcal{D}_{ijk} = \#$photons spatially projected into ij with energy level k.



$\Longrightarrow$

i	L_DEG	B_DEG	Energy
1	$\frac{L_i+L_{i+1}}{2}$	$\frac{B_j+B_{j+1}}{2}$	$\frac{E_k+E_{k+1}}{2}$
2	$\frac{L_i+L_{i+1}}{2}$	$\frac{B_j+B_{j+1}}{2}$	$\frac{E_k+E_{k+1}}{2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathcal{D}_{ijk}$	$\frac{L_i+L_{i+1}}{2}$	$\frac{B_j+B_{j+1}}{2}$	$\frac{E_k+E_{k+1}}{2}$

The astronomical maps

The statistical unit: Photons $i = 1, \dots, n$ with directions

$$\mathbf{x}_i = (x_i, y_i) \quad \in \quad \mathcal{X} = (x_{min}, x_{max}) \times (y_{min}, y_{max})$$

and energy level

$$E_i \in (E_{min},\ E_{max}).$$

Two main relevant sources of information:

- Astronomical sources at different levels of energies;
- background contamination.

Main Goals

Start from $\{(x_i, y_i, E_i)\}_{i=1}^{n}$

Main Goals

Start from $\{(x_i, y_i, E_i)\}_{i=1}^n \implies$ Infer on the sources in the map

Main Goals

Start from $\{(x_i, y_i, E_i)\}_{i=1}^{n} \implies$ Infer on the sources in the map

1. **discover** and locate **the high-energy astronomical sources** in a sky map;
2. **quantify** their **intensities**;
3. **distinguish** them from **the irregular background contamination** spread over the analysed area.

A very general statistical model

- $\mathbf{x}_i = (x_i, y_i), E_i$ with $i = 1, \ldots, n$ and $Z \in \{0, 1\}$,

$$\mathbf{x}_i, E_i | Z_i = z \sim \begin{cases} s(\mathbf{x}_i, E_i) & z = 1 \implies \text{Source model} \\ b(\mathbf{x}_i, E_i) & z = 0 \implies \text{Background model.} \end{cases}$$

A very general statistical model

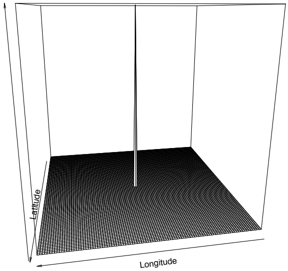
- $\mathbf{x}_i = (x_i, y_i), E_i$ with $i = 1, \dots, n$ and $Z \in \{0, 1\}$,

$$\mathbf{x}_i, E_i | Z_i = z \sim \begin{cases} s(\mathbf{x}_i, E_i) & z = 1 \implies \text{Source model} \\ b(\mathbf{x}_i, E_i) & z = 0 \implies \text{Background model.} \end{cases}$$

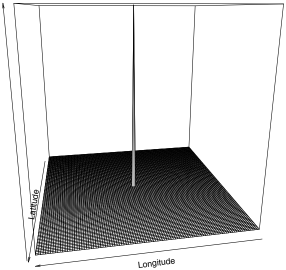
- Given $\delta \in (0, 1)$,

$$f(\mathbf{x}_i, E_i) = \delta s(\mathbf{x}_i, E_i) + (1 - \delta) b(\mathbf{x}_i, E_i).$$

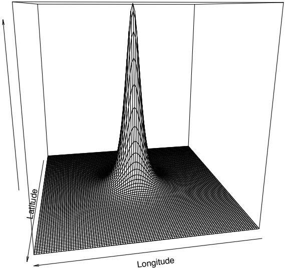
Single source model



Single source model



$\Longrightarrow$



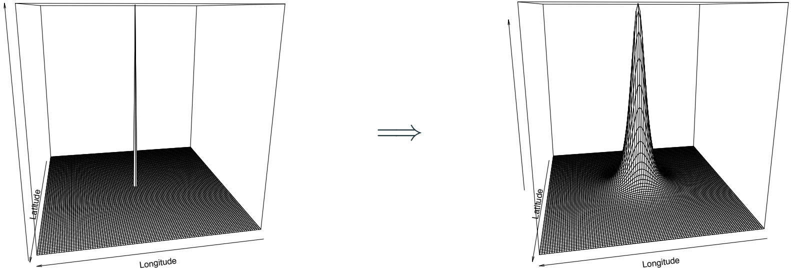
- A photon i from a source j spreads around it as

$$\mathbf{x}_i|E_i \sim \mathcal{N}(\boldsymbol{\mu}_j, \sigma^2_{E_i} I_2), \qquad \boldsymbol{\mu}_j \in \mathcal{X}.$$

- The spectral information from the sources is taken as

$$E_i \sim Par_t(\lambda_s, E_{\min}, E_{\max}), \qquad \lambda_s \in (1, 4).$$

Single source model



- A photon i from a source j spreads around it as

$$\mathbf{x}_i | E_i \sim \mathcal{N}(\boldsymbol{\mu}_j, \sigma^2_{E_i} I_2), \qquad \boldsymbol{\mu}_j \in \mathcal{X}.$$

- The spectral information from the sources is taken as

$$E_i \sim Par_t(\lambda_s, E_{\min}, E_{\max}), \qquad \lambda_s \in (1, 4).$$

Unknown Parameters	***Known Parameters***
$\boldsymbol{\mu}_j$: location of the source j	$\sigma^2_{E_i}$: variance parameter
λ_s: spectral parameter	

Modelling $s(\cdot)$: some aspects to consider

- Unknown number of sources in the map.

Modelling $s(\cdot)$: some aspects to consider

- Unknown number of sources in the map.
- Each source can be anywhere in the sky.

Modelling $s(\cdot)$: some aspects to consider

- Unknown number of sources in the map.
- Each source can be anywhere in the sky.
- Each source has a different intensity.

Modelling $s(\cdot)$: some aspects to consider

- Unknown number of sources in the map.
- Each source can be anywhere in the sky.
- Each source has a different intensity.

Multiple sources model

$$s(\mathbf{x}_i, E_i|\mathcal{F}, \lambda_s) = p(\mathbf{x}_i|E_i, \mathcal{F})g(E_i|\lambda_s),$$

where

$$p(\mathbf{x}_i|E_i, \mathcal{F}) = \int \phi(\mathbf{x}_i|\boldsymbol{\mu}, \sigma^2_{E_i} I_2)\mathcal{F}(d\boldsymbol{\mu}),$$

$$\mathcal{F} \sim \mathcal{DP}(\alpha_s, \mathcal{F}_0),$$

with

$$\mathcal{F}_0(\boldsymbol{\mu}) = \mathcal{U}(x_{min}, x_{max}) \times \mathcal{U}(y_{min}, y_{max})$$

.

Modelling $s(\cdot)$: an alternative representation

- The model can be rewritten as

$$s(\mathbf{x}_i, E_i|\boldsymbol{\mu}, \lambda_s, \boldsymbol{\pi}^s) = \sum_{j=1}^{\infty} \pi_j^s \phi(\mathbf{x}_i|\boldsymbol{\mu}_j, \sigma_{E_i}^2) p(E_i|\lambda_s),$$

$$V_j \sim Beta(1, \alpha_s), \qquad \pi_1^s = V_1, \qquad \pi_j^s = V_j \prod_{k=1}^{j-1}(1 - V_k),$$

$$\boldsymbol{\mu}_j \sim \mathcal{U}(\mathcal{X}), \qquad \lambda_s \sim \mathcal{U}(1, 4).$$

Nonparametric Bayes: why?

- Why should we have an infinite mixture?

Nonparametric Bayes: why?

- Why should we have an infinite mixture?
- *Jones et al (2015)* propose to model $s(\mathbf{x}_i, E_i)$ as a finite mixture, and the number of clusters is selected through reversible jump MCMC.

Nonparametric Bayes: why?

- Why should we have an infinite mixture?
- *Jones et al (2015)* propose to model $s(\mathbf{x}_i, E_i)$ as a finite mixture, and the number of clusters is selected through reversible jump MCMC.
  1. with BNP, $K \to \infty$ as $n \to \infty$: the number of clusters grows with the sample size.

Nonparametric Bayes: why?

- Why should we have an infinite mixture?
- *Jones et al (2015)* propose to model $s(\mathbf{x}_i, E_i)$ as a finite mixture, and the number of clusters is selected through reversible jump MCMC.
    1. with BNP, $K \to \infty$ as $n \to \infty$: the number of clusters grows with the sample size.
    2. Simulation algorithm for BNP are simple to implement and explore the parameter space faster than the reversible jump MCMC.

Modelling the background

Some characteristics of the background contamination:

Modelling the background

Some characteristics of the background contamination:

- it has an **irregular and unpredictable behaviour**;

Modelling the background

Some characteristics of the background contamination:

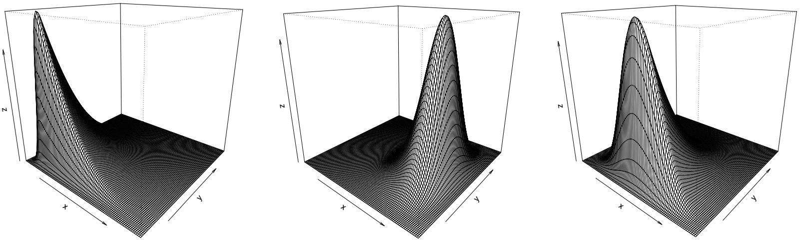
- it has an **irregular and unpredictable behaviour**;
- it tends to be **smoother than the sources**;

Modelling the background

Some characteristics of the background contamination:

- it has an **irregular and unpredictable behaviour**;
- it tends to be **smoother than the sources**;
- **no parametric models are available** to account for it.

The B-spline basis function



- For $x \in \mathbb{R}$, a **B-spline basis function** of order m is defined as

$$B_m(x|\boldsymbol{\xi}) = \frac{x-\xi_1}{\xi_{m+1}-\xi_1} B_{m-1}(x|\boldsymbol{\xi}_{1:m}) + \frac{\xi_{m+1}-x}{\xi_{m+1}-\xi_2} B_{m-1}(x|\boldsymbol{\xi}_{2:(m+1)}),$$

  where $B_1(x|a,b) = I(a \le x \le b)$.
- $B_m(\cdot|\cdot)$ is always positive, unimodal and simple to normalize.

Modelling $b(\cdot)$: a (Bayesian) nonparametric approach

- Given $\boldsymbol{l}_j = (l_{j1}, \ldots, l_{j5})$ and $\boldsymbol{b}_j = (b_{j1}, \ldots, b_{j5})$, we define

$$\mathcal{P}(\mathbf{x}_i | \boldsymbol{l}_j, \boldsymbol{b}_j) \propto B_4(x_i | \boldsymbol{l}_j) B_4(y_i | \boldsymbol{b}_j).$$

Modelling $b(\cdot)$: a (Bayesian) nonparametric approach

- Given $\boldsymbol{l}_j = (l_{j1}, \ldots, l_{j5})$ and $\boldsymbol{b}_j = (b_{j1}, \ldots, b_{j5})$, we define

$$\mathcal{P}(\mathbf{x}_i | \boldsymbol{l}_j, \boldsymbol{b}_j) \propto B_4(x_i | \boldsymbol{l}_j) B_4(y_i | \boldsymbol{b}_j).$$

- The energy is modelled as

$$E_i | \lambda_b \sim Par_t(\lambda_b, E_{\min}, E_{\max})$$

Modelling $b(\cdot)$: a (Bayesian) nonparametric approach

- Given $\boldsymbol{l}_j = (l_{j1}, \ldots, l_{j5})$ and $\boldsymbol{b}_j = (b_{j1}, \ldots, b_{j5})$, we define

$$\mathcal{P}(\mathbf{x}_i|\boldsymbol{l}_j, \boldsymbol{b}_j) \propto B_4(x_i|\boldsymbol{l}_j)B_4(y_i|\boldsymbol{b}_j).$$

- The energy is modelled as

$$E_i|\lambda_b \sim Par_t(\lambda_b, E_{\min}, E_{\max})$$

Our proposal

$$b(\mathbf{x}_i, E_i|\mathcal{G}, \lambda_b) = k(\mathbf{x}_i|\mathcal{G})g(E_i|\lambda_b),$$

where

$$k(\mathbf{x}_i|\mathcal{G}) = \int \mathcal{P}(\mathbf{x}_i; \boldsymbol{l}, \boldsymbol{b})\mathcal{G}(d\boldsymbol{l}, d\boldsymbol{b}), \qquad \mathcal{G} \sim \mathcal{DP}(\alpha_b, \mathcal{G}_0).$$

Background model: an alternative representation

- The model can be rewritten as

$$b(\mathbf{x}_i, E_i | \boldsymbol{l}, \boldsymbol{b}, \lambda_b, \boldsymbol{\pi}^b) = \sum_{j=1}^{\infty} \pi_j^b \mathcal{B}_4(x_i | \boldsymbol{l}_j) \mathcal{B}_4(y_i | \boldsymbol{b}_j) g(E_i | \lambda_b),$$

$$U_j \sim Beta(1, \alpha_b), \qquad \pi_1^b = U_1, \qquad \pi_j^b = U_j \prod_{k=1}^{j-1} (1 - U_k),$$

Background model: an alternative representation

- The model can be rewritten as

$$b(\mathbf{x}_i, E_i|\boldsymbol{l}, \boldsymbol{b}, \lambda_b, \boldsymbol{\pi}^b) = \sum_{j=1}^{\infty} \pi_j^b \mathcal{B}_4(x_i|\boldsymbol{l}_j)\mathcal{B}_4(y_i|\boldsymbol{b}_j)g(E_i|\lambda_b),$$

$$U_j \sim Beta(1, \alpha_b), \qquad \pi_1^b = U_1, \qquad \pi_j^b = U_j \prod_{k=1}^{j-1}(1 - U_k),$$

$$\begin{aligned} l_{j1} &\sim \mathcal{U}(x_{\min}, x_{\max}), \qquad l_{jk} \sim \mathcal{U}(l_{j(k-1)}, x_{\max}), \\ b_{j1} &\sim \mathcal{U}(y_{\min}, y_{\max}), \qquad b_{jk} \sim \mathcal{U}(b_{j(k-1)}, y_{\max}), \quad k = 2, \ldots, 5 \end{aligned}$$

Background model: an alternative representation

- The model can be rewritten as

$$b(\mathbf{x}_i, E_i | \boldsymbol{l}, \boldsymbol{b}, \lambda_b, \boldsymbol{\pi}^b) = \sum_{j=1}^{\infty} \pi_j^b \mathcal{B}_4(x_i | \boldsymbol{l}_j) \mathcal{B}_4(y_i | \boldsymbol{b}_j) g(E_i | \lambda_b),$$

$$U_j \sim Beta(1, \alpha_b), \qquad \pi_1^b = U_1, \qquad \pi_j^b = U_j \prod_{k=1}^{j-1} (1 - U_k),$$

$$\begin{aligned} l_{j1} &\sim \mathcal{U}(x_{\min}, x_{\max}), \quad & l_{jk} &\sim \mathcal{U}(l_{j(k-1)}, x_{\max}), \\ b_{j1} &\sim \mathcal{U}(y_{\min}, y_{\max}), \quad & b_{jk} &\sim \mathcal{U}(b_{j(k-1)}, y_{\max}), \quad k = 2, \ldots, 5 \end{aligned}$$

$$\lambda_b \sim Unif(1, 4).$$

Distinguish between sources and background

Proposition

Let X be a random variable with a density function corresponding to $B_m(\cdot|\boldsymbol{\xi})$. Then

$$Var(X) = \frac{\sum_{p=1}^{m} \sum_{q=p+1}^{m+1} (\xi_p - \xi_q)^2}{(m+1)^2(m+2)}.$$

Distinguish between sources and background

Proposition

Let X be a random variable with a density function corresponding to $B_m(\cdot|\boldsymbol{\xi})$. Then

$$Var(X) = \frac{\sum_{p=1}^{m}\sum_{q=p+1}^{m+1}(\xi_p - \xi_q)^2}{(m+1)^2(m+2)}.$$

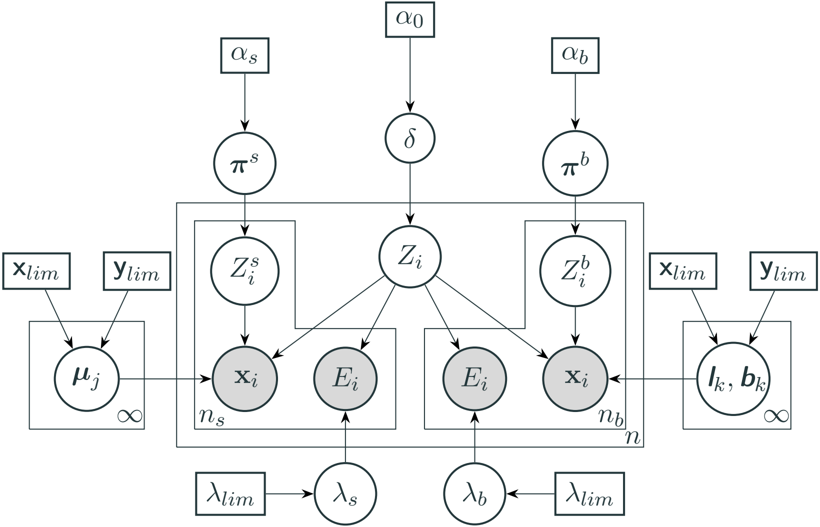
- *We impose a constraint $\forall j$ such that*

$$\frac{\sum_{p=1}^{m}\sum_{q=p+1}^{m+1}(\xi_{jp} - \xi_{jq})^2}{(m+1)^2(m+2)} > \psi,$$

with $\boldsymbol{\xi}_j = \{\boldsymbol{l}_j, \boldsymbol{b}_j\}$ and, given $c > 1$,

$$\psi = c \cdot \max_i \sigma^2_{E_i}.$$

The statistical model: a graphical representation



A suitable MCMC algorithm

- Let

  $\mathcal{S}^{(t)} = \{i = 1, \ldots, n : Z_i^{(t)} = 1\}$ and $\mathcal{B}^{(t)} = \{i = 1, \ldots, n : Z_i^{(t)} = 0\}$.
- We admit at most k_s components for s and k_b components for b.

A suitable MCMC algorithm

- Let

  $\mathcal{S}^{(t)} = \{i = 1, \ldots, n : Z_i^{(t)} = 1\}$ and $\mathcal{B}^{(t)} = \{i = 1, \ldots, n : Z_i^{(t)} = 0\}$.

- We admit at most k_s components for s and k_b components for b.
- For $t = 1, \ldots, T$, iterate the followings:

A suitable MCMC algorithm

- Let

  $$\mathcal{S}^{(t)} = \{i = 1, \ldots, n : Z_i^{(t)} = 1\} \text{ and } \mathcal{B}^{(t)} = \{i = 1, \ldots, n : Z_i^{(t)} = 0\}.$$

- We admit at most k_s components for s and k_b components for b.
- For $t = 1, \ldots, T$, iterate the followings:
  1. pull photons into $\mathcal{S}^{(t)}$ and $\mathcal{B}^{(t)}$ and update
     $\delta^{(t)} | \cdots \sim Beta(\alpha_0 + |\mathcal{S}^{(t)}|, \alpha_0 + |\mathcal{B}^{(t)}|)$;

A suitable MCMC algorithm

- Let

  $\mathcal{S}^{(t)} = \{i = 1, \ldots, n : Z_i^{(t)} = 1\}$ and $\mathcal{B}^{(t)} = \{i = 1, \ldots, n : Z_i^{(t)} = 0\}$.
- We admit at most k_s components for s and k_b components for b.
- For $t = 1, \ldots, T$, iterate the followings:
  1. pull photons into $\mathcal{S}^{(t)}$ and $\mathcal{B}^{(t)}$ and update
     $\delta^{(t)} | \cdots \sim Beta(\alpha_0 + |\mathcal{S}^{(t)}|, \alpha_0 + |\mathcal{B}^{(t)}|)$;
  2. update the weights

$$V_j | \mathcal{S}^{(t)} \sim Beta(n_j^{(t)} + 1, \sum_{k>j} n_k^{(t)} + \alpha_s) \Longrightarrow {\pi^s}_j^{(t)} | \mathcal{S}^{(t)}, \quad j = 1, \ldots, k_s,$$

$$U_j | \mathcal{B}^{(t)} \sim Beta(n_j^{(t)} + 1, \sum_{k>j} n_k^{(t)} + \alpha_b) \Longrightarrow {\pi^b}_j^{(t)} | \mathcal{B}^{(t)}, \quad j = 1, \ldots, k_b.$$

A suitable MCMC algorithm

- Let

  $$\mathcal{S}^{(t)} = \{i = 1, \dots, n : Z_i^{(t)} = 1\} \text{ and } \mathcal{B}^{(t)} = \{i = 1, \dots, n : Z_i^{(t)} = 0\}.$$

- We admit at most k_s components for s and k_b components for b.
- For $t = 1, \dots, T$, iterate the followings:
  1. pull photons into $\mathcal{S}^{(t)}$ and $\mathcal{B}^{(t)}$ and update
     $\delta^{(t)} | \dots \sim Beta(\alpha_0 + |\mathcal{S}^{(t)}|, \alpha_0 + |\mathcal{B}^{(t)}|)$;
  2. update the weights

$$V_j|\mathcal{S}^{(t)} \sim Beta(n_j^{(t)} + 1, \sum_{k>j} n_k^{(t)} + \alpha_s) \Longrightarrow {\pi^s}_j^{(t)} | \mathcal{S}^{(t)}, \quad j = 1, \dots, k_s,$$

$$U_j|\mathcal{B}^{(t)} \sim Beta(n_j^{(t)} + 1, \sum_{k>j} n_k^{(t)} + \alpha_b) \Longrightarrow {\pi^b}_j^{(t)} | \mathcal{B}^{(t)}, \quad j = 1, \dots, k_b.$$

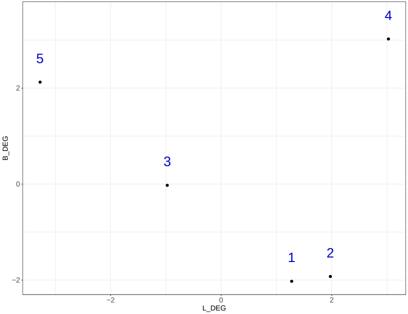
  3. update

$$(\boldsymbol{\mu}_j, \lambda_s)^{(t)} | \mathcal{S}^{(t)}, \dots \qquad \forall j = 1, \dots, k_s,$$

$$(\boldsymbol{l}_j, \boldsymbol{b}_j, \lambda_b)^{(t)} | \mathcal{B}^{(t)}, \dots \qquad \forall j = 1, \dots, k_b.$$

Application on a simulated dataset

- Let $\mathcal{D}$ be a $200 \times 200 \times 25$ array:
  - > $\mathcal{X} = (-4.975, 5.025) \times (-4.975, 5.025)$; each spatial bin is large 0.05.
  - > Energy divided into 25 log10-spaced bins in the range $(1GeV,\ 316.2278GeV)$
- The dataset consists in a background component and 5 sources in the following locations:



Application on a simulated dataset

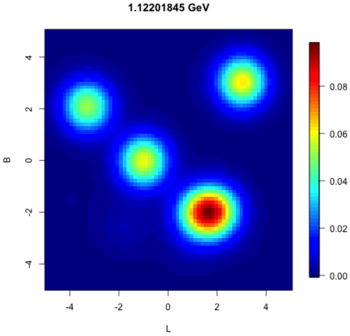
- We sample the counts from each source using a tabulated Point Spread Function and differential flux

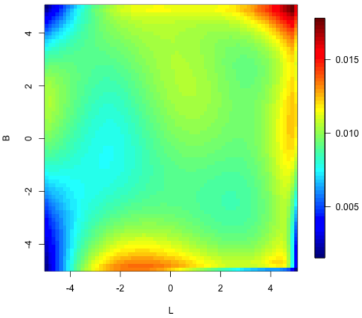
$$\frac{\partial F}{\partial E} = F_0 \left(\frac{E}{E_0} \right)^{-\lambda}, \qquad F_0 = 1 \cdot 10^{-9},\ \lambda = 2,\ E_0 = 1GeV.$$

- A subset of 10000 photons is used.
- Hyperparameters are chosen to be as follows:

$$\alpha = 1, \qquad \alpha_s \sim Gamma(9, 3) \qquad \alpha_b = 1.$$

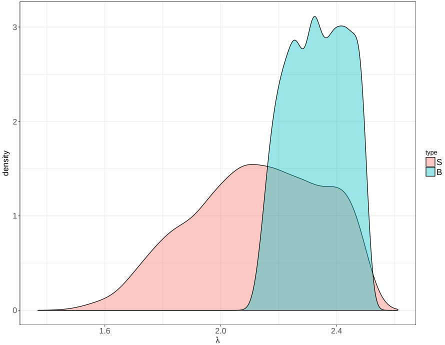
Density estimation





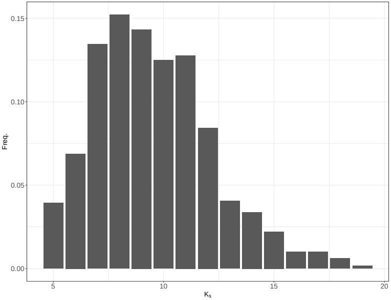
Posterior results: spectral parameters

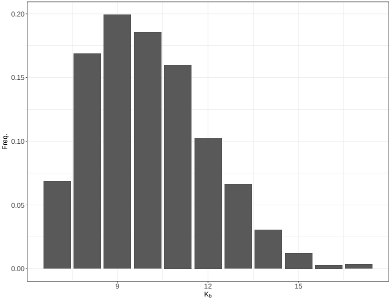
- The spectral parameters λ_s and λ_b are.



Posterior results

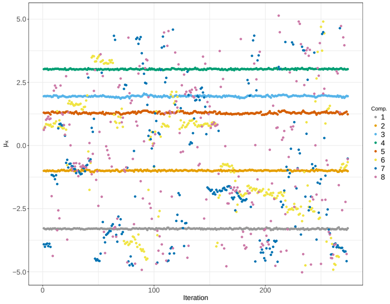
- Number of active components inside the two mixtures after burn-in:

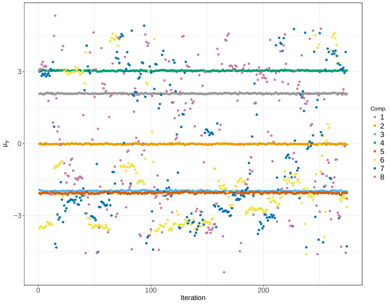




Posterior results

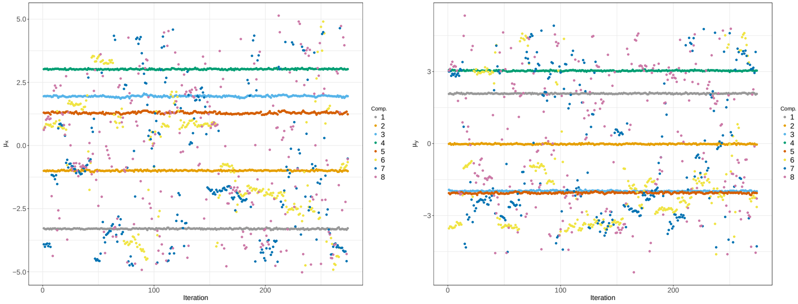
- Given $K_s = \arg\max_k \sum_t I(K_s^{(t)} = k)$, the draws from the posterior distribution of $\boldsymbol{\mu}$ under K_s are:





Posterior results

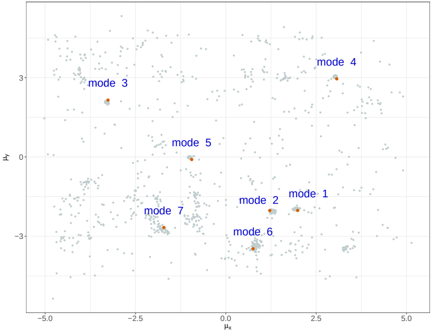
- Given $K_s = \arg\max_k \sum_t I(K_s^{(t)} = k)$, the draws from the posterior distribution of $\boldsymbol{\mu}$ under K_s are:



A post-processing algorithm to quantify the information from the posterior distribution of $\boldsymbol{\mu}$ is required.

Post processing algorithm: stage 1

- $\{\boldsymbol{\mu}_1^{(t)}, \ldots, \boldsymbol{\mu}_{K_s}^{(t)}\}$ be the set of draws from the posterior distribution of $\boldsymbol{\mu}$ when the number of active clusters is K_s.



- fit a nonparametric density (or alternatively a 3D-histogram) to determine the most relevant points in the previous map: $(\mathbf{m}_1, \ldots, \mathbf{m}_p)$.

Post processing algorithm: stage 2

- For each $\boldsymbol{\mu}_k^{(t)}$, find the point such that

$$\min_{j=1,\ldots,K_s} ||\boldsymbol{\mu}_k^{(t)} - \mathbf{m}_j|| < r, \tag{1}$$

where r is a given threshold.

Post processing algorithm: stage 2

- For each $\boldsymbol{\mu}_k^{(t)}$, find the point such that

$$\min_{j=1,\ldots,K_s} ||\boldsymbol{\mu}_k^{(t)} - \mathbf{m}_j|| < r, \tag{1}$$

  where r is a given threshold.
- If there exists j which satisfies (1), label $\boldsymbol{\mu}_k^{(t)}$ as j.

Post processing algorithm: stage 2

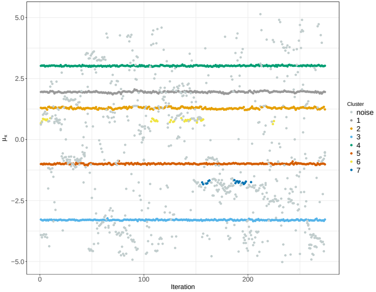
- For each $\boldsymbol{\mu}_k^{(t)}$, find the point such that

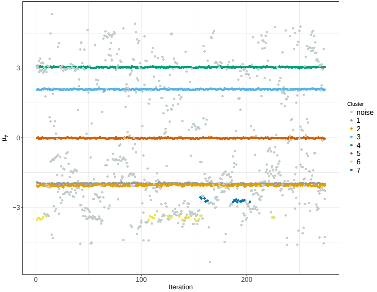
$$\min_{j=1,\ldots,K_s} ||\boldsymbol{\mu}_k^{(t)} - \mathbf{m}_j|| < r, \tag{1}$$

  where r is a given threshold.
- If there exists j which satisfies (1), label $\boldsymbol{\mu}_k^{(t)}$ as j.
- If no j satisfies (1), label $\boldsymbol{\mu}_k^{(t)}$ as *noise*.

Post processing algorithm: stage 2

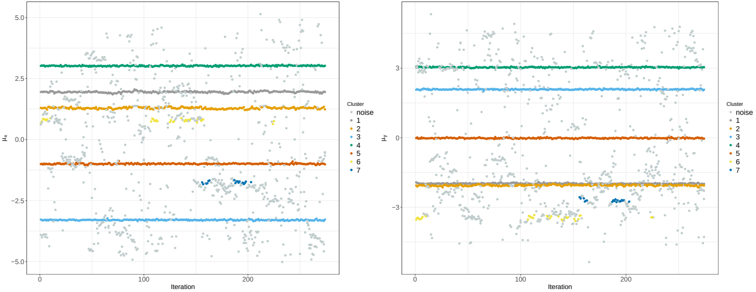
- The relabelled draws from $\boldsymbol{\mu}$ are:





Post processing algorithm: stage 2

- The relabelled draws from $\boldsymbol{\mu}$ are:



- The new labels can be applied also to $(\pi_1^s, \ldots, \pi_{K_s}^s)^{(t)}$ to have an estimate on the intensity of each cluster.

Post processing algorithm: stage 2

cluster	#counts	$\mathbb{E}(N\|\dots)$	2.5%	25%	75%	97.5%	$\mathbb{P}(source)$
1*	136	215.631	150.475	196.000	238.000	274.000	1.000
2*	166	182.073	123.000	162.250	205.750	237.950	0.956
3*	160	185.511	147.650	173.250	198.000	225.000	1.000
4*	138	231.967	178.825	211.250	252.000	282.350	1.000
5*	149	220.270	173.825	206.000	234.750	263.350	1.000
6	//	19.324	1.000	10.250	29.000	45.350	0.124
7	//	43.056	23.425	34.500	51.500	64.575	0.066

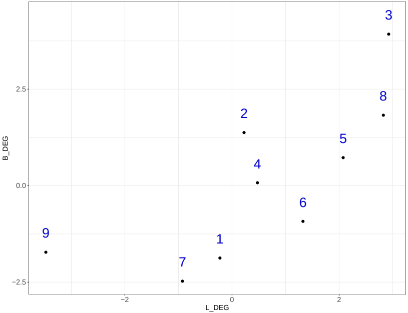
*: the cluster coincides with a real source.

A second simulated dataset

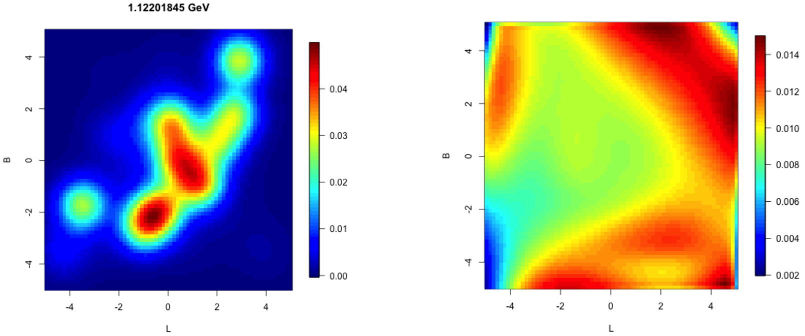
- A subset of 10000 photons is used.
- Hyperparameters are chosen to be as follows:

$$\alpha = 1, \qquad \alpha_s \sim Gamma(9,3) \qquad \alpha_b = 1.$$

- The dataset consists in a background component and 9 sources in the following locations:

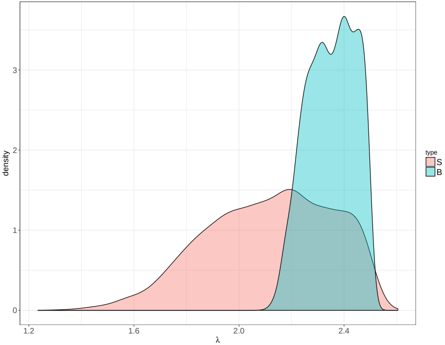


Density estimation



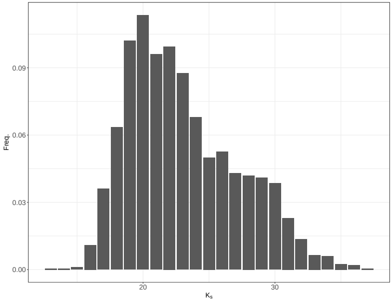
Posterior results: spectral parameters

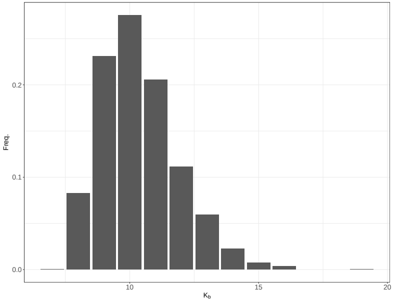
- The spectral parameters λ_s and λ_b are.



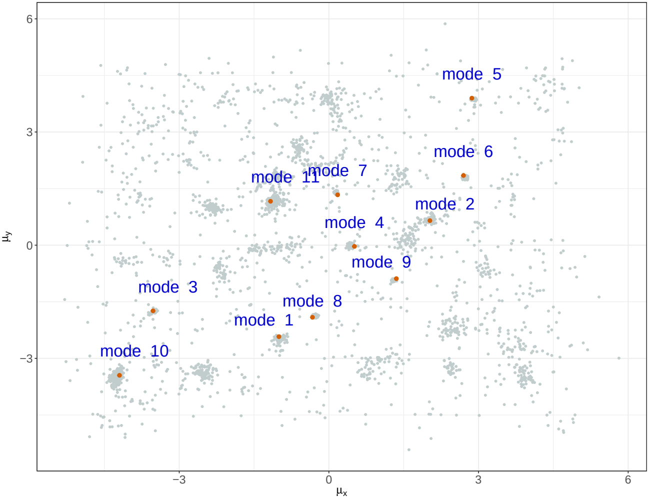
Posterior results: number of components

- Number of active components inside the two mixtures after burn-in:

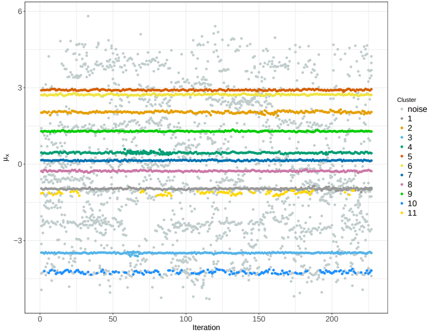


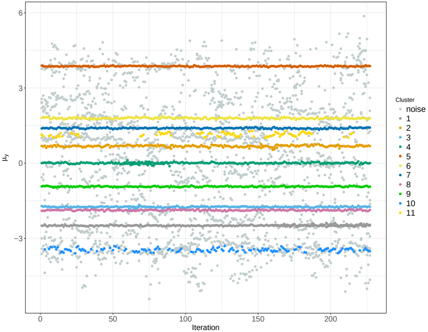


Applying the post processing algorithm, step 1



Applying the post processing algorithm, step 2





Results

cluster	#counts	$\mathbb{E}(N\|\dots)$	2.5%	25%	75%	97.5%	$\mathbb{P}(source)$
1*	138	202.295	142.000	186.000	222.000	252.000	1.000
2*	163	145.190	105.500	134.000	158.000	180.000	0.996
3*	138	174.956	135.650	163.000	186.500	206.700	1.000
4*	148	210.665	156.900	197.000	224.000	253.000	1.000
5*	126	184.093	131.000	167.000	201.500	234.000	1.000
6*	139	165.066	119.650	151.500	179.000	205.000	1.000
7*	141	191.802	140.650	178.500	208.000	241.350	1.000
8*	141	196.612	148.600	184.500	211.000	238.700	1.000
9*	160	214.907	158.300	202.000	230.000	258.700	1.000
10	//	36.884	10.125	31.000	45.000	62.000	0.643
11	//	41.057	14.350	27.750	51.750	91.300	0.388

*: the cluster coincides with a real source.

Conclusions

- We proposed a new approach to the analysis of astronomical 2D-images via hierarchical Bayesian nonparametric mixture modelling.

Conclusions

- We proposed a new approach to the analysis of astronomical 2D-images via hierarchical Bayesian nonparametric mixture modelling.
- Our method is capable of:

Conclusions

- We proposed a new approach to the analysis of astronomical 2D-images via hierarchical Bayesian nonparametric mixture modelling.
- Our method is capable of:
    1. reconstructing the background component;

Conclusions

- We proposed a new approach to the analysis of astronomical 2D-images via hierarchical Bayesian nonparametric mixture modelling.
- Our method is capable of:
    1. reconstructing the background component;
    2. locating the possible sources in the map;

Conclusions

- We proposed a new approach to the analysis of astronomical 2D-images via hierarchical Bayesian nonparametric mixture modelling.
- Our method is capable of:
    1. reconstructing the background component;
    2. locating the possible sources in the map;
    3. quantifying their intensities.

Conclusions

- We proposed a new approach to the analysis of astronomical 2D-images via hierarchical Bayesian nonparametric mixture modelling.
- Our method is capable of:
    1. reconstructing the background component;
    2. locating the possible sources in the map;
    3. quantifying their intensities.
- Even if the posterior distribution of the source model parameters is multimodal, we can quantify the intensity of each mode and estimate its probability of being a source.

Bibliography

- Ackermann, M. et al. (Fermi-LAT Collaboration) (2013). *Determination of the point-spread function for the Fermi Large Area Telescope from on-orbit data and limits of pair halos of Active Galactic Nuclei*. The Astrophysical Journal, 765:54–73.
- De Boor, C. (2001). *A Practical Guide to Splines (Revised Edition)*. Applied Mathematical Sciences 27, Springer.
- Hobson, M. P., Jaffe, A. H., Liddle, A. R., Mukherjee, P., and Parkinson, D.] (2010). *Bayesian Methods in Cosmology*. Cambridge University Press.
- Ishwaran, H. and Zarepour, M. (2000) *Markov chain Monte Carlo in approximate Dirichlet and beta two-parameter process hierarchical models*. Biometrika 87:371–390

Bibliography

- Ishwaran, H. and James, L. F. (2001) *Gibbs sampling methods for stick-breaking priors.* Journal of the American Statistical Association 96.453: 161–173.
- Jones, D. E., Kashyap, V. L., and van Dyk, D. A. (2015). *Disentangling overlapping astronomical sources using spatial and spectral information*. The Astrophysical Journal, 808:137–161.
- Sethuraman, J. (1994) *A* constructive definition of Dirichlet prior. Statistica sinica 2:639–650
- Sottosanti, A., Costantin, D., Bastieri, D., and Brazzale, A. R. *Discovering and Locating High-Energy Extra-Galactic Sources by Bayesian Mixture Modelling*. Working Paper Series, N. 2, May 2018, Department of Statistical Sciences, University of Padua.