Paris-Saclay
Center for Data Science

Workshop deblending with Deep Learning

Alexandre Boucaud

Paris-Saclay Center for Data Science / LAL

aboucaud@lal.in2p3.fr

Outline

1. Recent deblending progress from the literature

2. Why this « workshop »

Deblending (recent work)

Forward modeling — MuSCADeT

Morpho Spectral Component Analysis (MCA)

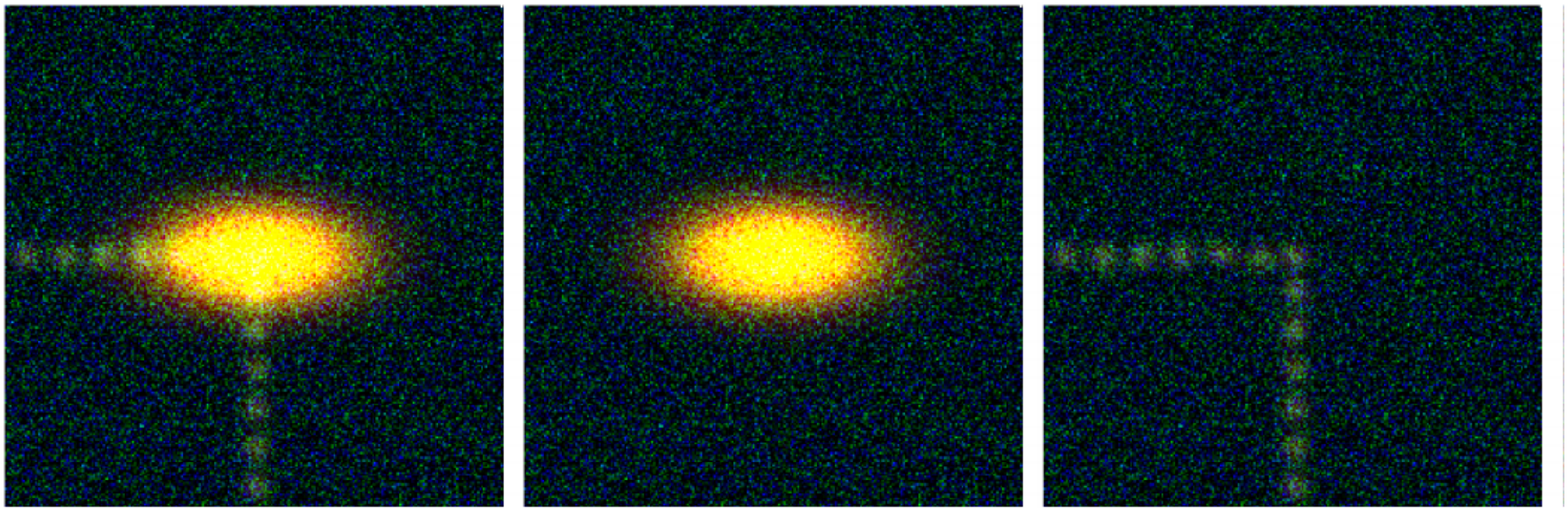


Fig. B.1. Separation of blended sources at low S/N. From left to right are shown the original simulated images, the original image after subtraction of the blue component as estimated from `MuSCADeT`, the original image after subtraction of the red component and the residual image after subtraction of both components.

Joseph+16

Forward modeling — NMF

- Color should be useful, photo-z are dangerous
- Star/Galaxy separation is not obvious: non-parametric
- Objects are somehow "compact", mostly symmetric

$$\text{scene} = \sum_k \text{SED}_k \times \text{Morphology}_k + \text{noise}$$

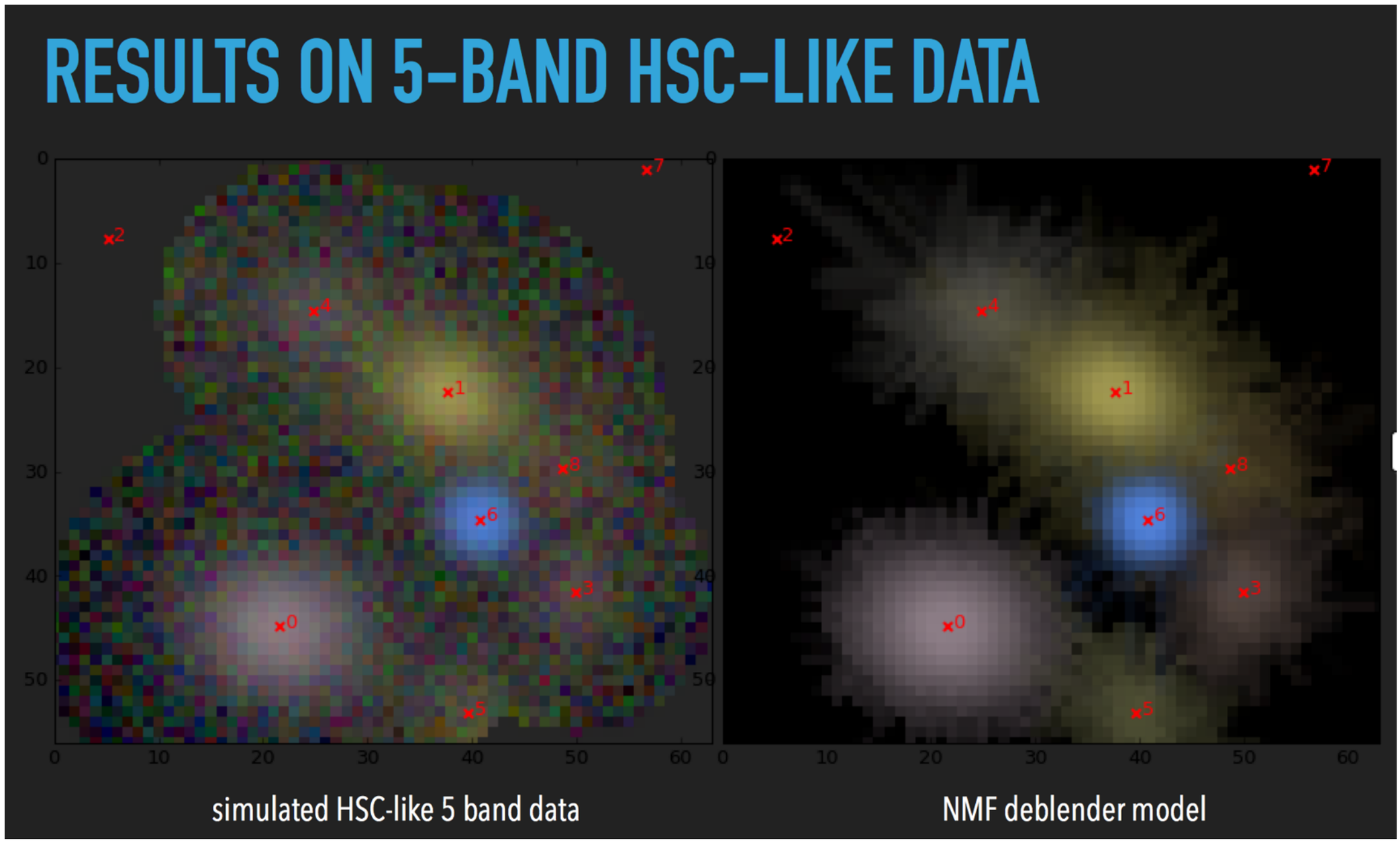
$$Y = A \cdot S + \text{noise}$$

$$(Y \in \mathbb{R}^{B \times N}, A \in \mathbb{R}^{B \times K}, \ S \in \mathbb{R}^{K \times N})$$

$$||Y - A \cdot S||_2^2 + g(A, S)$$

Melchior+17 in prep.

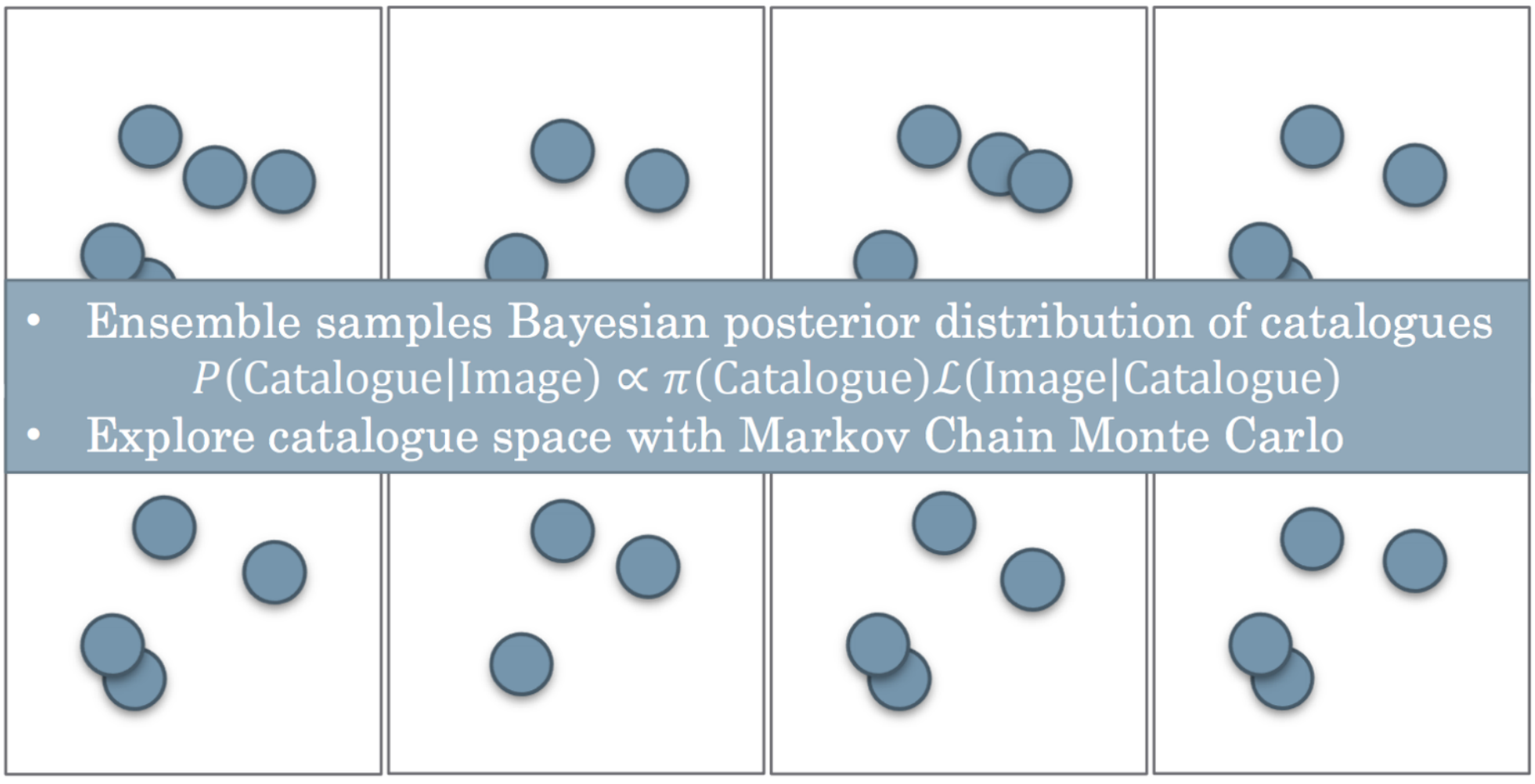
Forward modeling — NMF



Melchior+17 in prep.

Probabilistic cataloguing

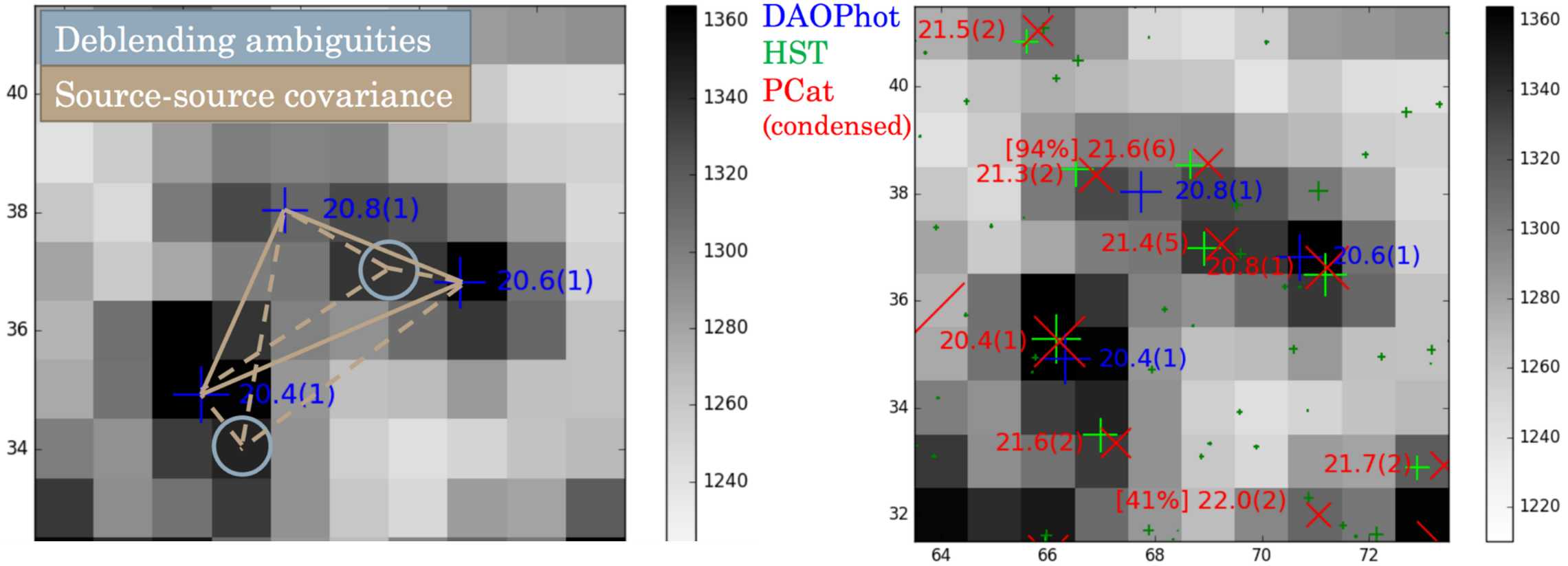
- Infer an ***ensemble of catalogues***
- Naturally handles deblending ambiguities and source-source covariance



Portillo+17

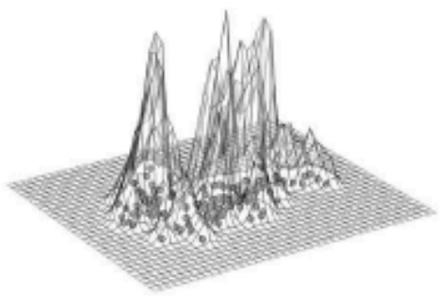
Probabilistic cataloguing

Condensed catalog

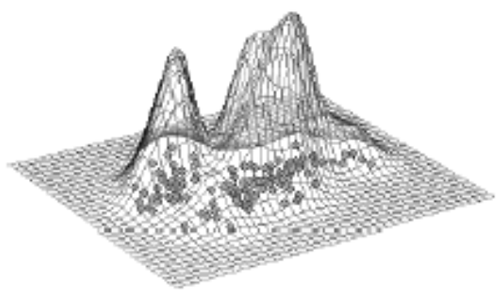


Portillo+17

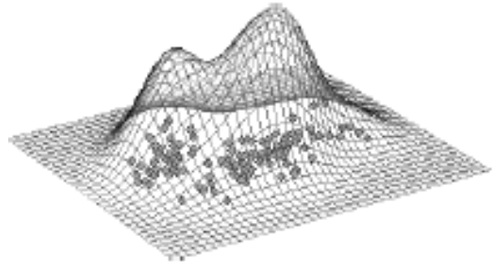
Topological clustering — ASTErIsM



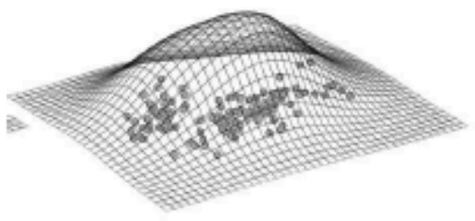
(a) $h = 0.1$



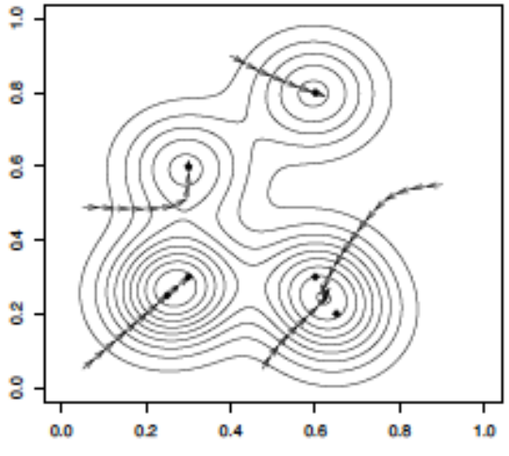
(b) $h = 0.2$



(c) $h = 0.35$



(d) $h = 0.6$



fast hill climbing

$$\mathbf{x}_{t+1} = \frac{\sum_{i=1}^{n} K\left(\frac{\mathbf{x}_t - \mathbf{x}_i}{h}\right) \mathbf{x}_i}{\sum_{i=1}^{n} K\left(\frac{\mathbf{x}_t - \mathbf{x}_i}{h}\right)}$$

$x_t - x_{t+1} < \varepsilon$ ←

See talk from Valerio

Tramacere+

SExtractor ++

soon ?!

Deep learning

see today’s presentations

Why are you here ?

universite PARIS-SACLAY

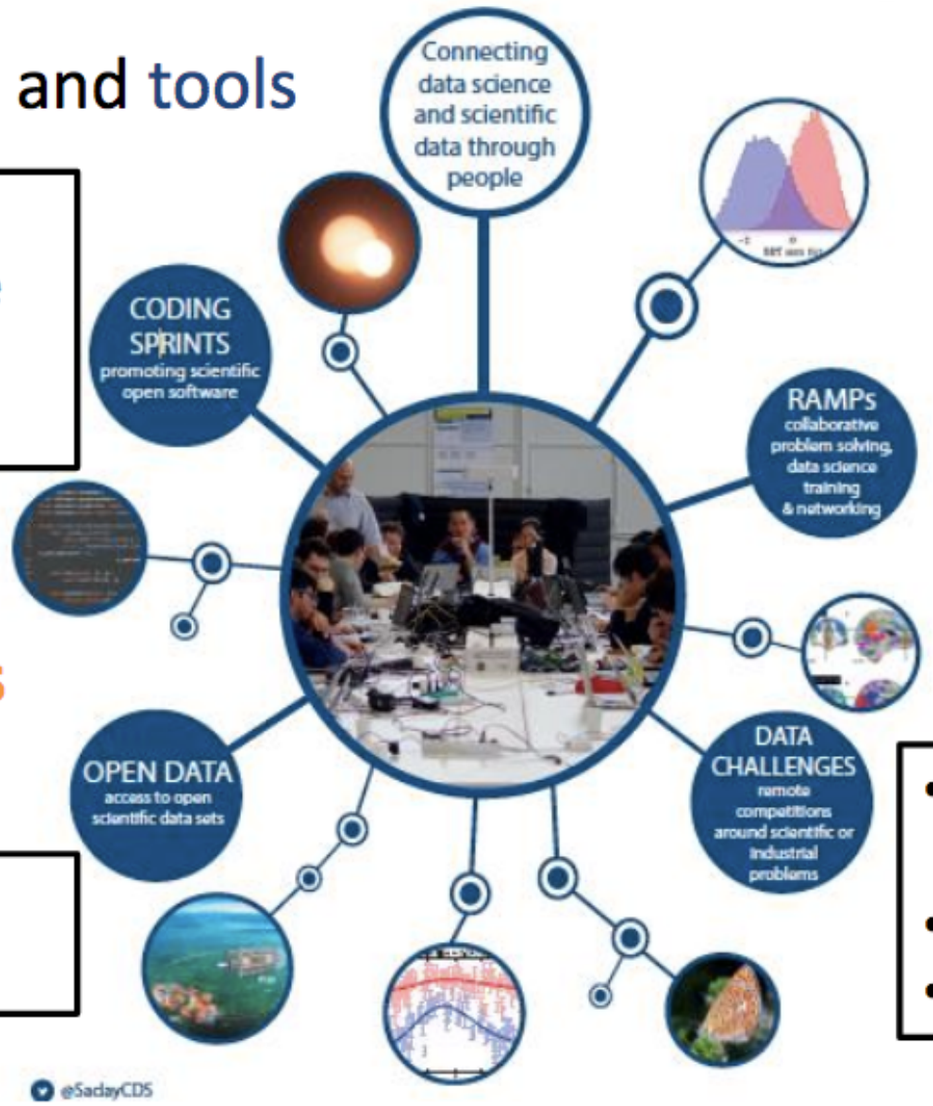
Paris-Saclay Center for Data Science

Connecting experts and tools

- State-of-the-art data science in easy-to-use tools
- High-quality software

Connecting experts and problems

- Prototyping
- Training
- Collaboration building



Connecting experts and data

- Data as a Service
- Linked (Open) Data

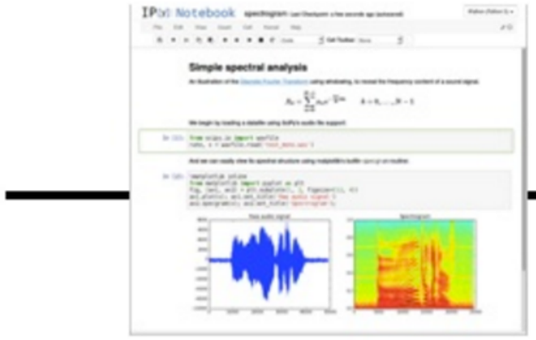
- Impact on science
- Visibility
- Benchmarks

www.datascience-paris-saclay.fr

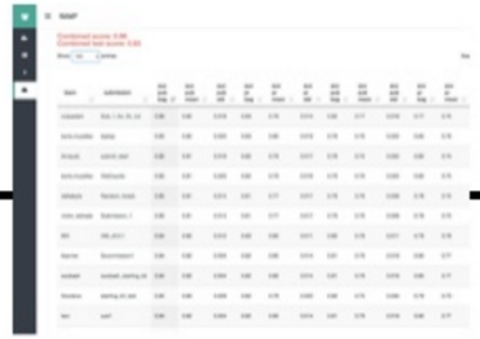
Use creativity to find the best solution

Rapid Analytics and Model Prototyping

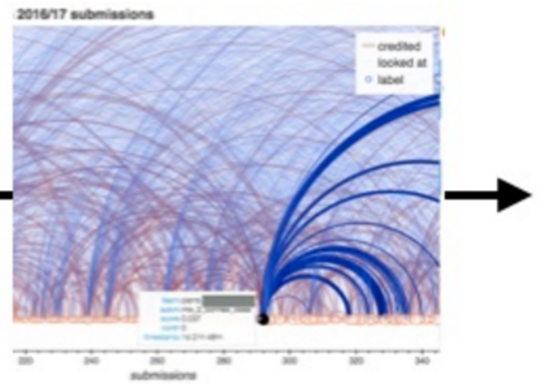
starting kit



competition



collaboration



a baseline solution
- usually built by an expert
- explains the domain problem, variables and the objectives
- enable other data scientists to become operational

a leaderboard
- gamifies model development
- boosts effort & exploration
- creates a diversity of models
- enhance problem understanding

propagating best ideas
- all models are open
- a wealth of information becomes accessible
- ideas & models are combined
- synergy & learning & creativity

see Mehdi and Yetkin talks

Why a DL challenge on deblending ?

- **Deblending** is a good topic for a challenge
  - difficult but not impossible
  - needed for current and upcoming surveys
- **Deep Learning** is the outsider tool
  - growing interest
  - has recently proven to beat all methods on SL challenge
  - huge possibilities but no recipe to create the best architecture

Our needs

- A deblending score, which is the definition of the problem itself

- A set of labeled data for training and testing