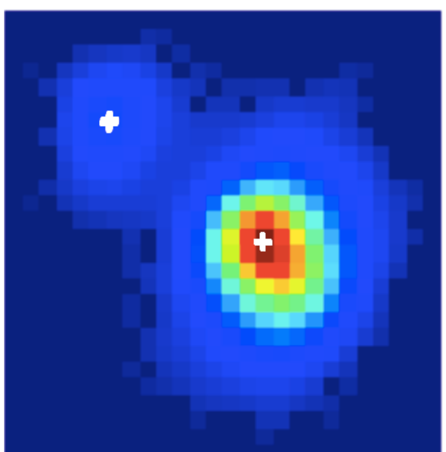
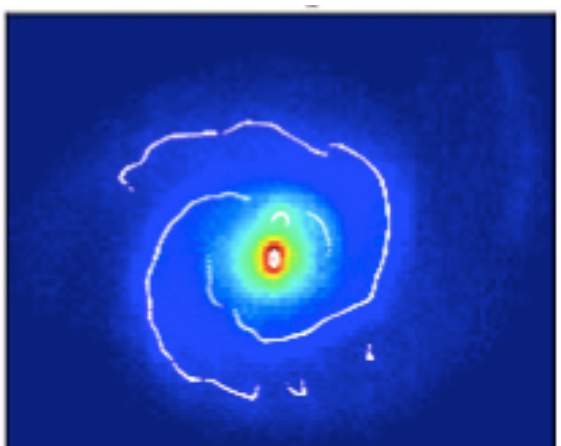
Deblending strategies for OU-MER Euclid pipeline

Valerio Roscani

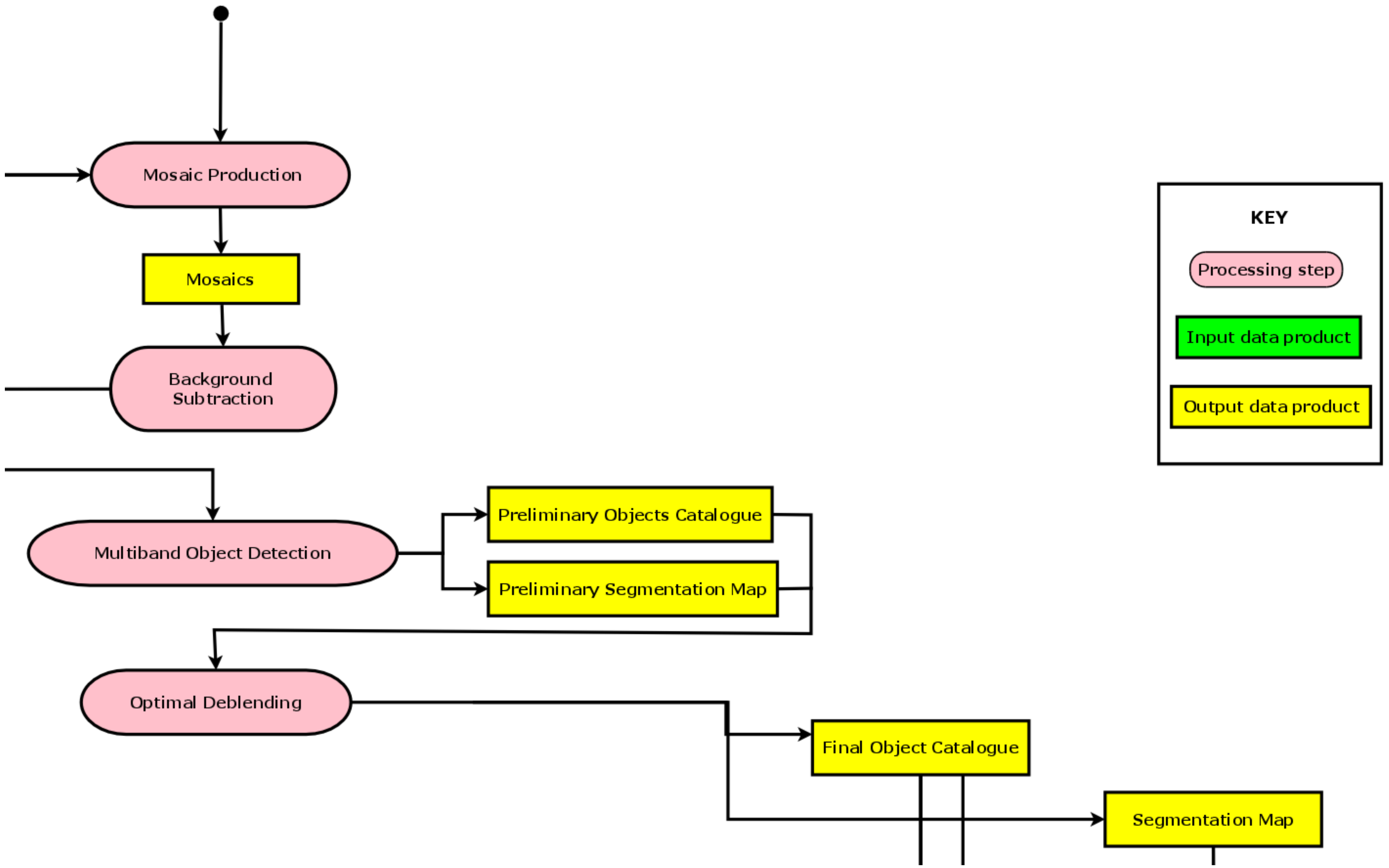
INAF - Osservatorio Astronomico di Roma

Collaborators: M. Castellano, A. Tramacere, E. Merlin, A. Fontana, S. Pilo + MER Team

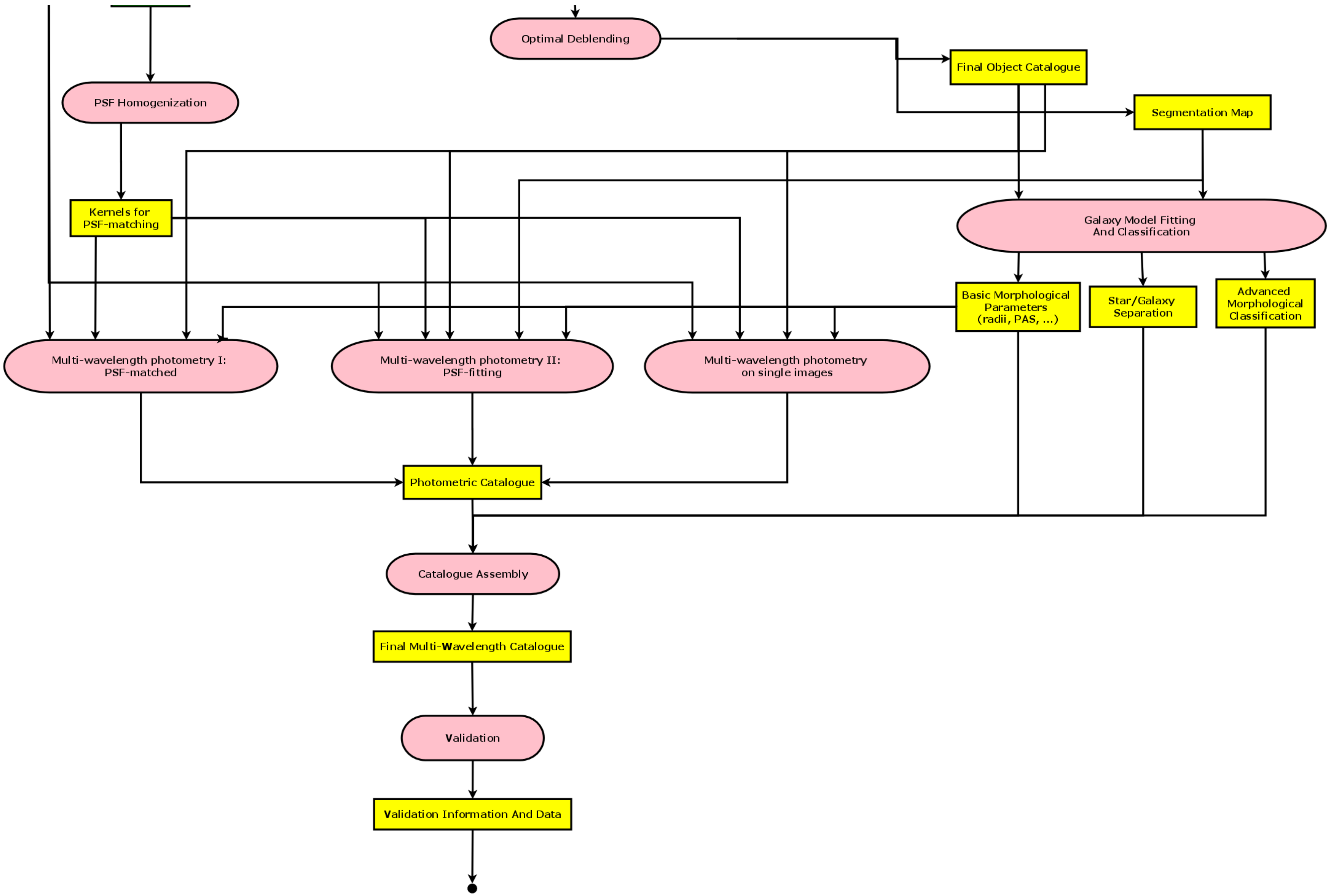




OU-MER PF Workflow



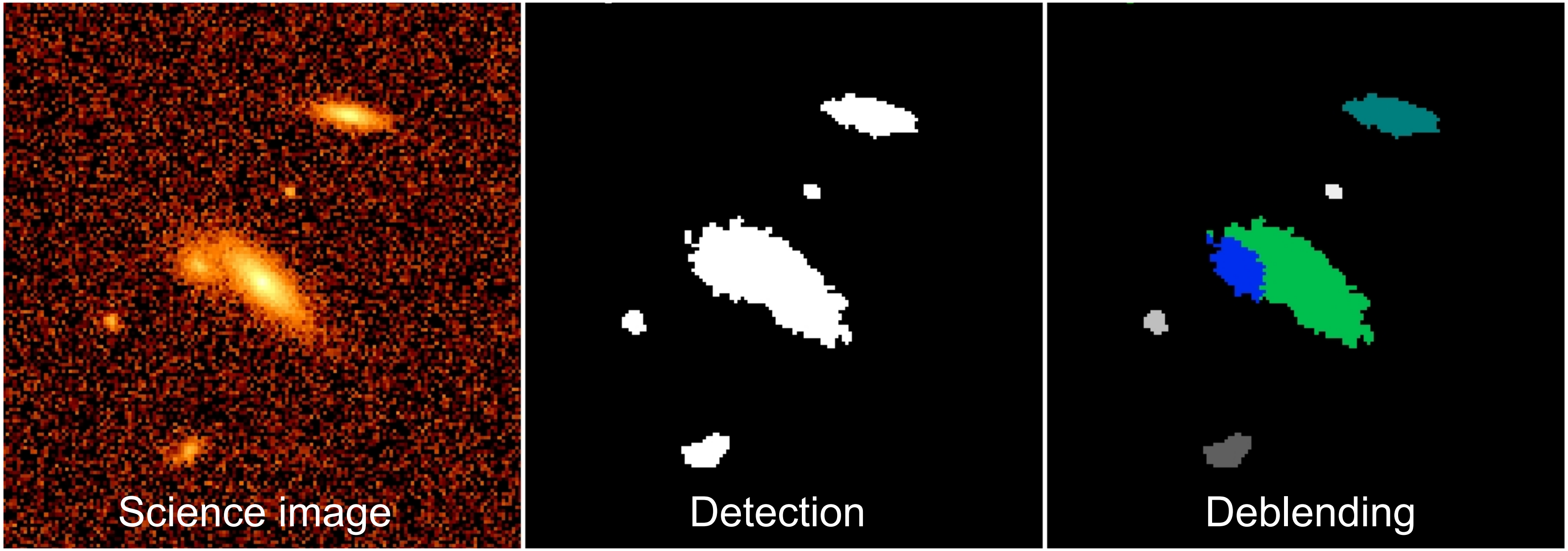
OU-MER PF Workflow



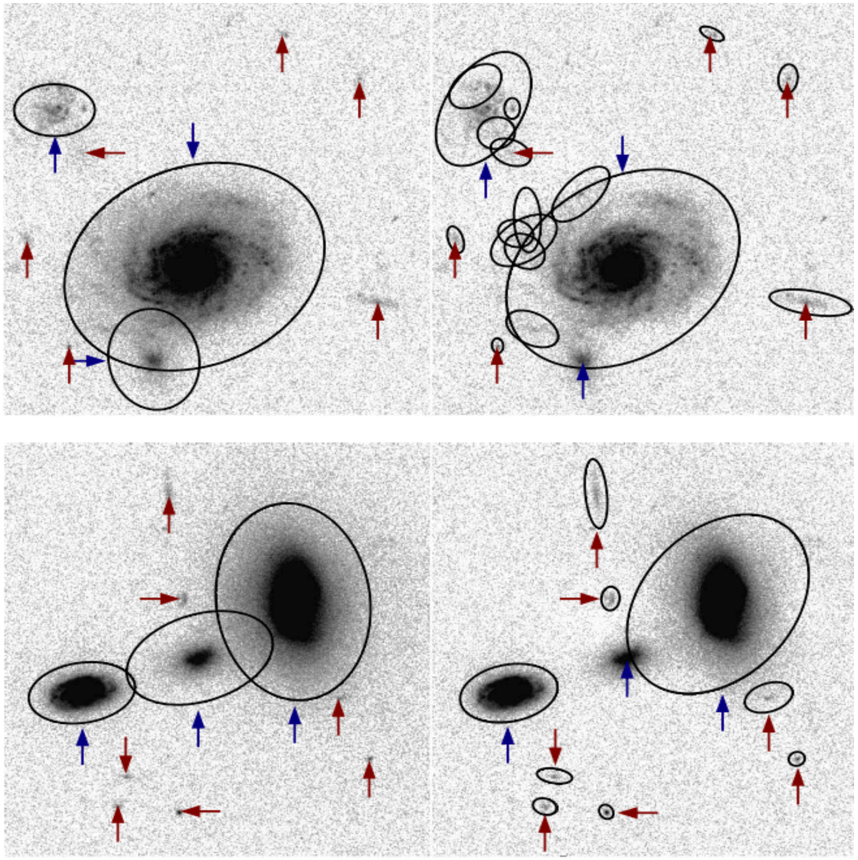
Optimal Deblending (PE-MER-05)

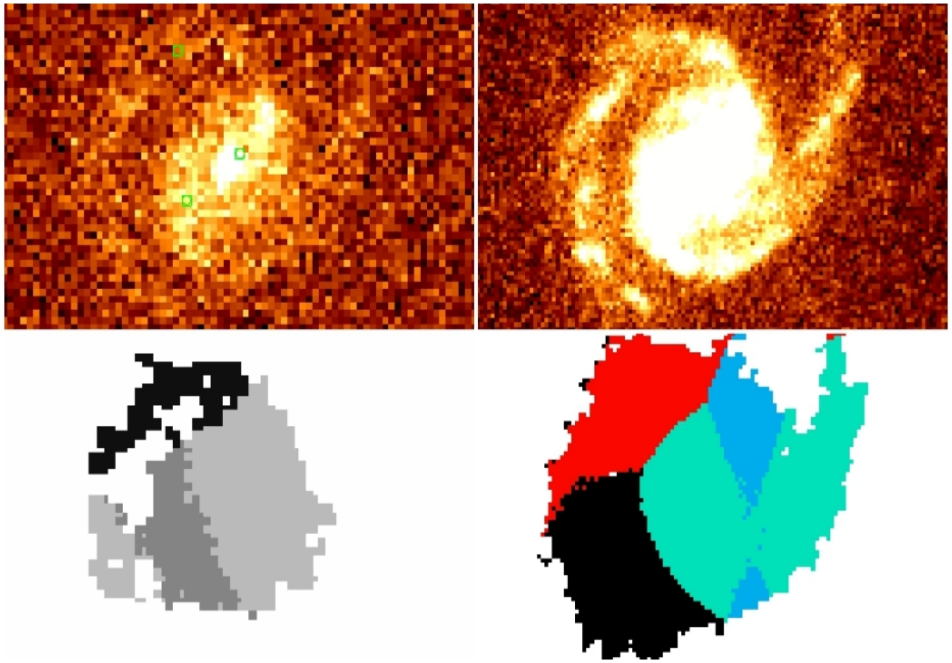
- Develop an algorithm for efficient object deblending (separation of close sources) in the VIS+NIR bands and explore the possibility of extending the approach to VIS+NIR+EXT bands.
- PE inputs: science, r.m.s. and flag images, segmentation map (= map of significant pixels from PE "Detection")

Internal OU-MER VIS simulations (E. Merlin, S. Pilo)



Source deblending: an ill-posed problem?



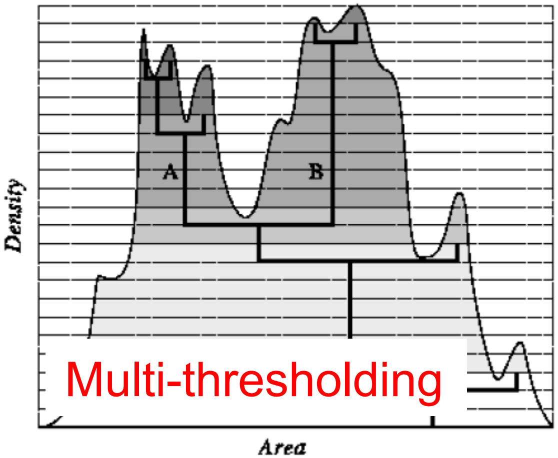


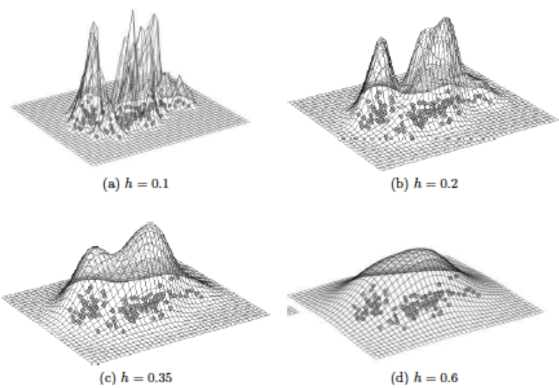
HOT+COLD SExtractor routines (e.g. Galametz+13, and all CANDELS catalogs)

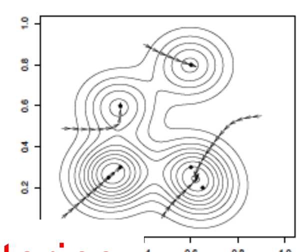
Problem of tuning deblending so to avoid over-splitting extended sources

Status

- Status: First version implemented in the current MER pipeline, tested on MER-PHZ Data-challenges and running in Scientific Challenge 3
- Current implementation: Multi-threshold approach (SExtractor) with parameters optimized after tests on Euclid-like images
- Alternative approach : Topological clustering algorithms DBSCAN-DENCLUE (Tramacere&Vecchio 2013)



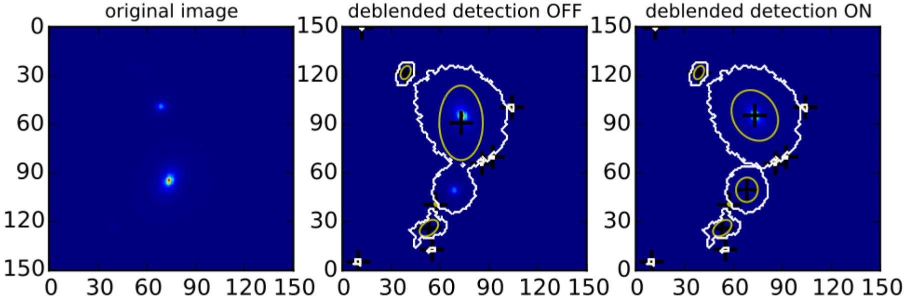




fast hill climbing

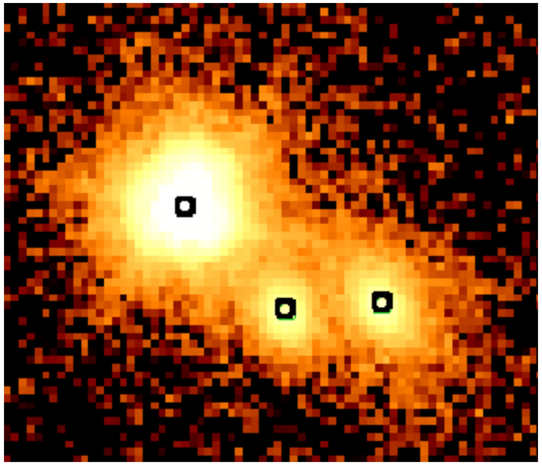
$$x_{t+1} = \frac{\sum_{i=1}^{n} K\left(\frac{x_t - x_i}{h}\right) x_i}{\sum_{i=1}^{n} K\left(\frac{x_t - x_i}{h}\right)}$$

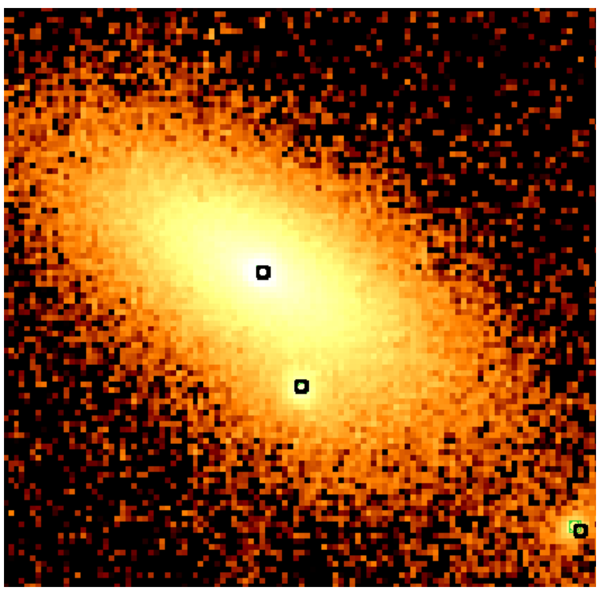
$x_t - x_{t+1} < \varepsilon$

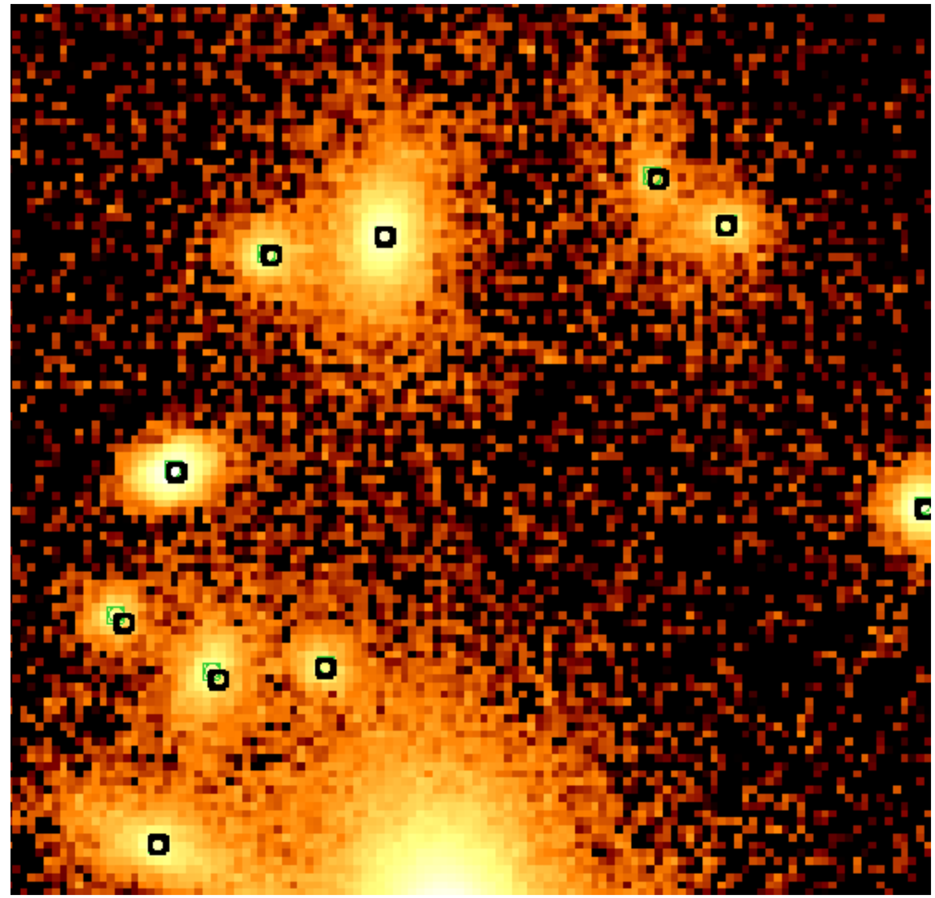


Topological clustering

Euclid Scientific Challenge 3



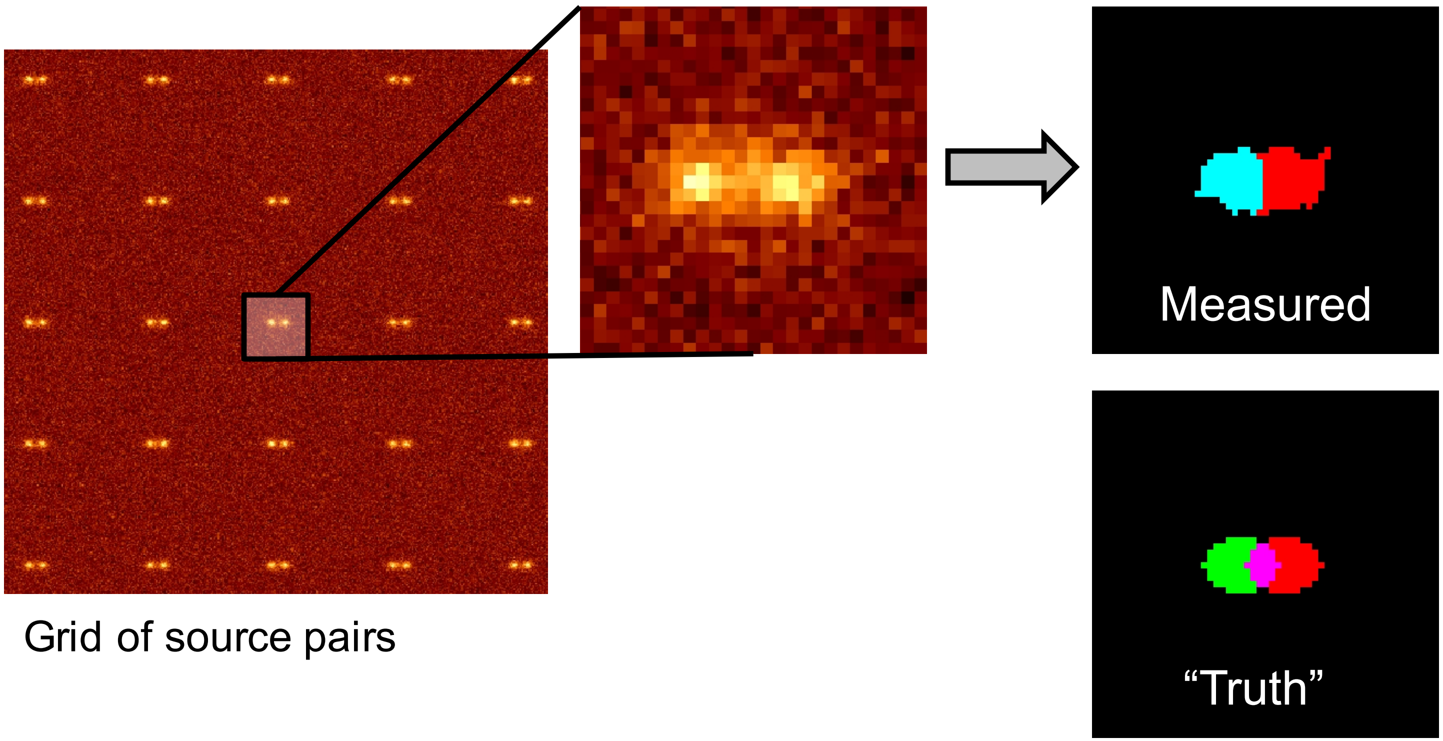




Basic deblending with tuned parameters is accurate enough to properly identify the bulk of the Euclid sources: check on field 53877 of SC3 (MER PF at work these days)

Test: characterization of deblending efficiency

- Baseline solution tested on MER-PHZ data-challenges and characterized on MER simulations



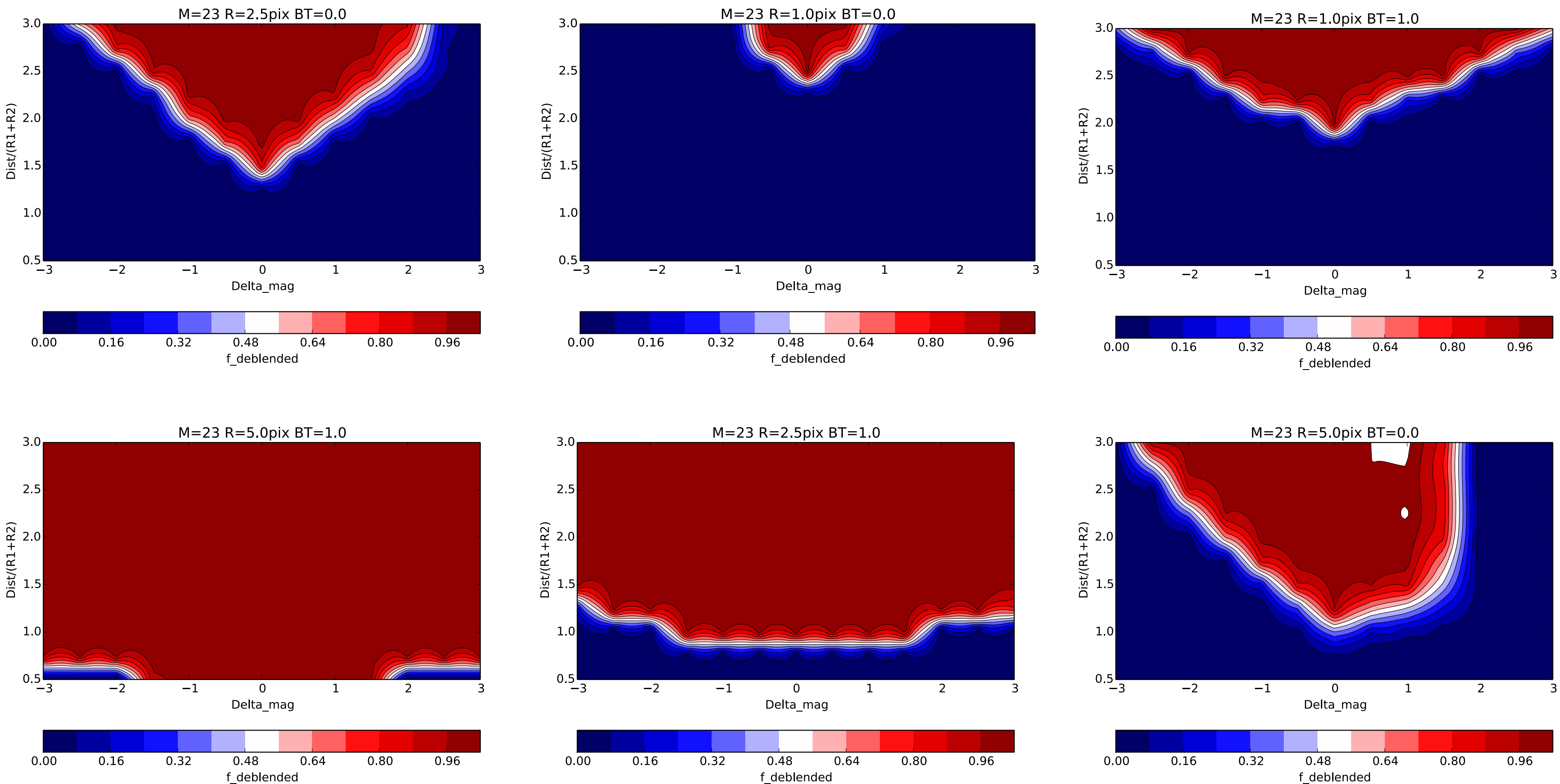
Grid of source pairs

Same tests performed for multi-threshold and topological approach

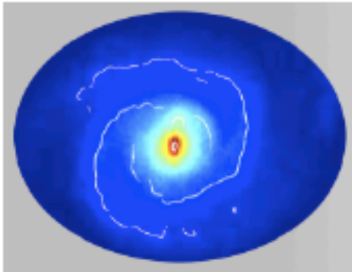
Test Status and Results

Test: characterization of deblending efficiency

- Baseline solution tested on MER-PHZ data-challenges and characterized on MER simulations



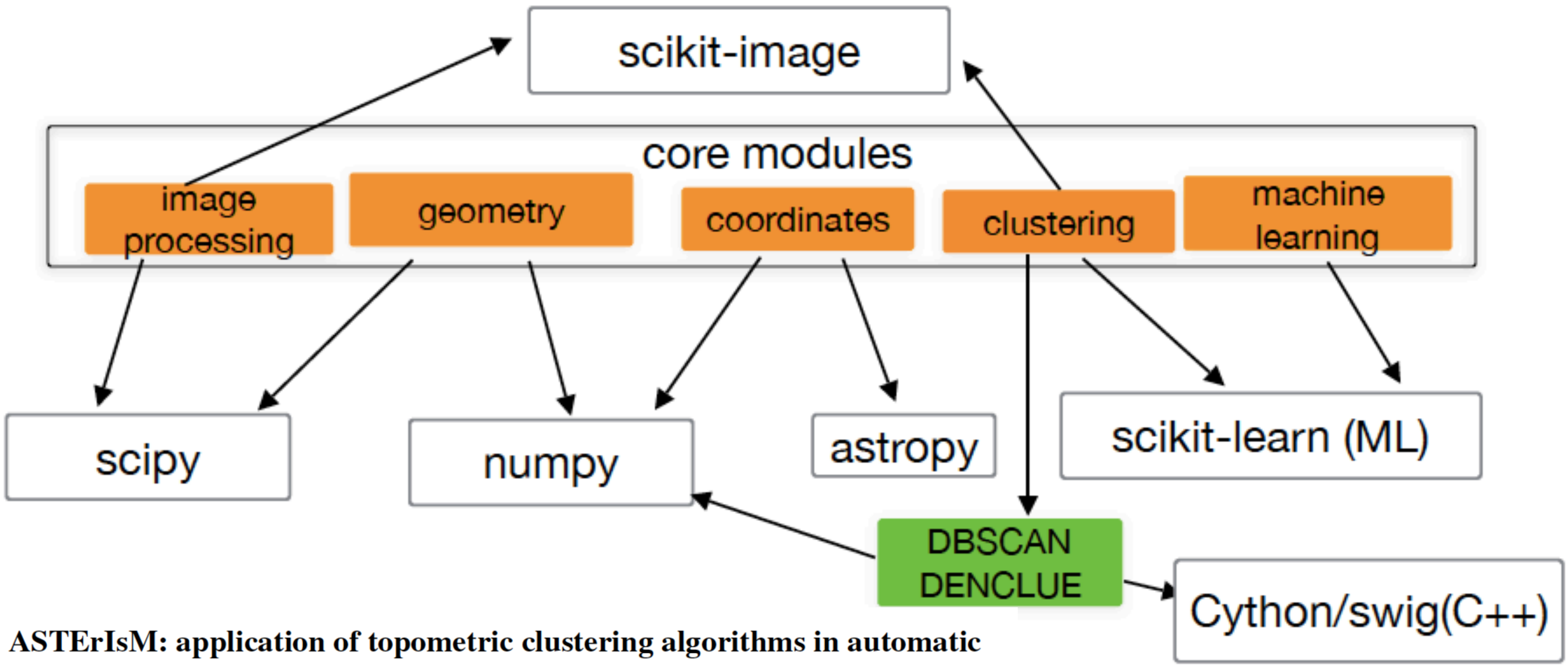
Deblending: refined solutions



ASTErIsM

python AStronomical Tools for clustering-based dEtectIon and Morphometry

ASTErIsM is a python framework for application of topometric clustering algorithms in automatic source detection morphometry and classification



ASTErIsM: application of topometric clustering algorithms in automatic galaxy detection and classification

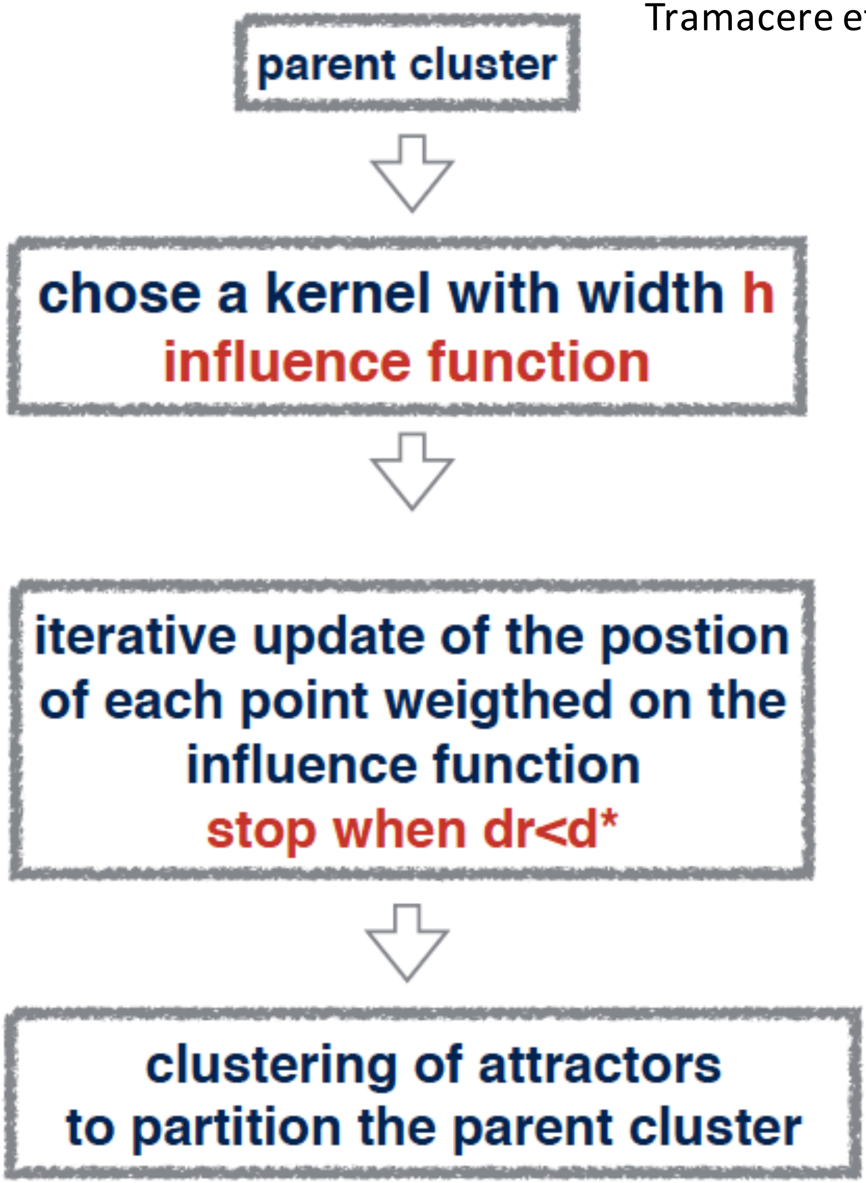
A. Tramacere,[1★] D. Paraficz,[2] P. Dubath,[1] J.-P. Kneib[2] and F. Courbin[2]

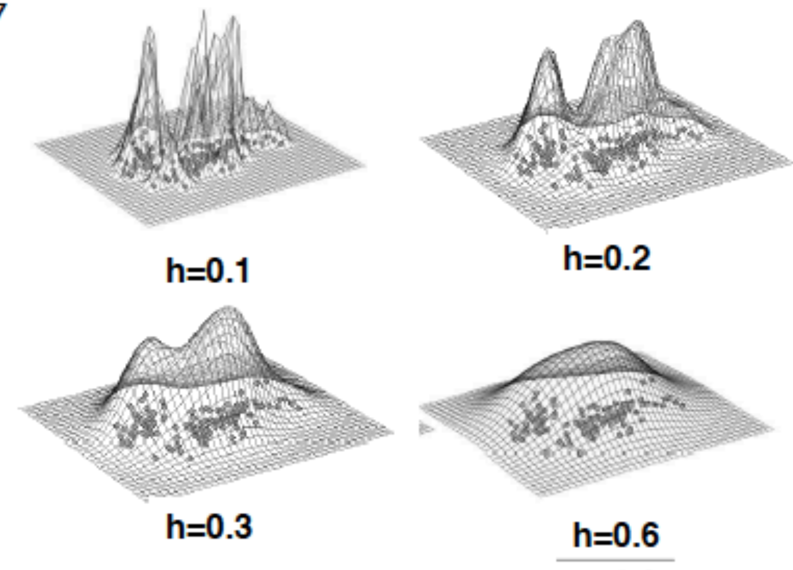
[1] *ISDC Data Centre for Astrophysics, Astronomy Department of the University of Geneva, Chemin d'Ecogia 16, CH-1290 Versoix, Switzerland*

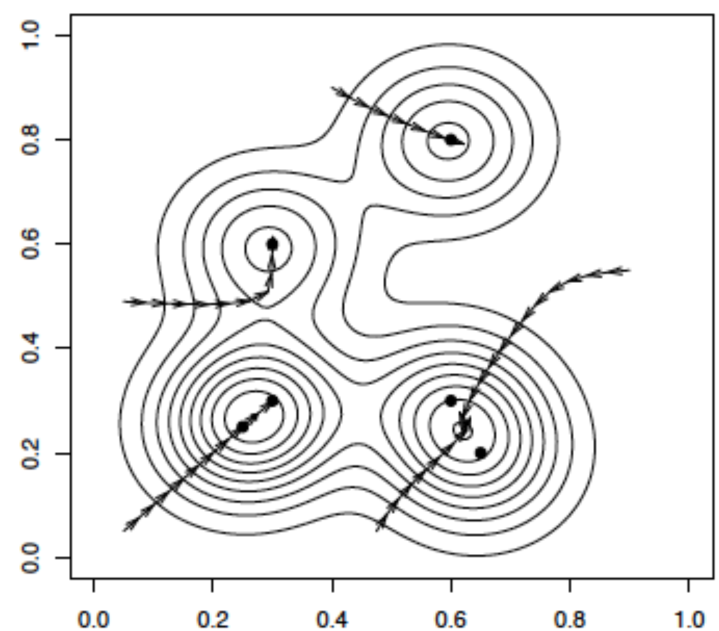
[2] *Laboratoire d'Astrophysique, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, CH-1290 Versoix, Switzerland*

Deblending & DENCLUE

Hinneburg &Gabriel 2007
Tramacere et. al 2016







The convolved image f is defined through the kernel G, q are the pixels coordinates :

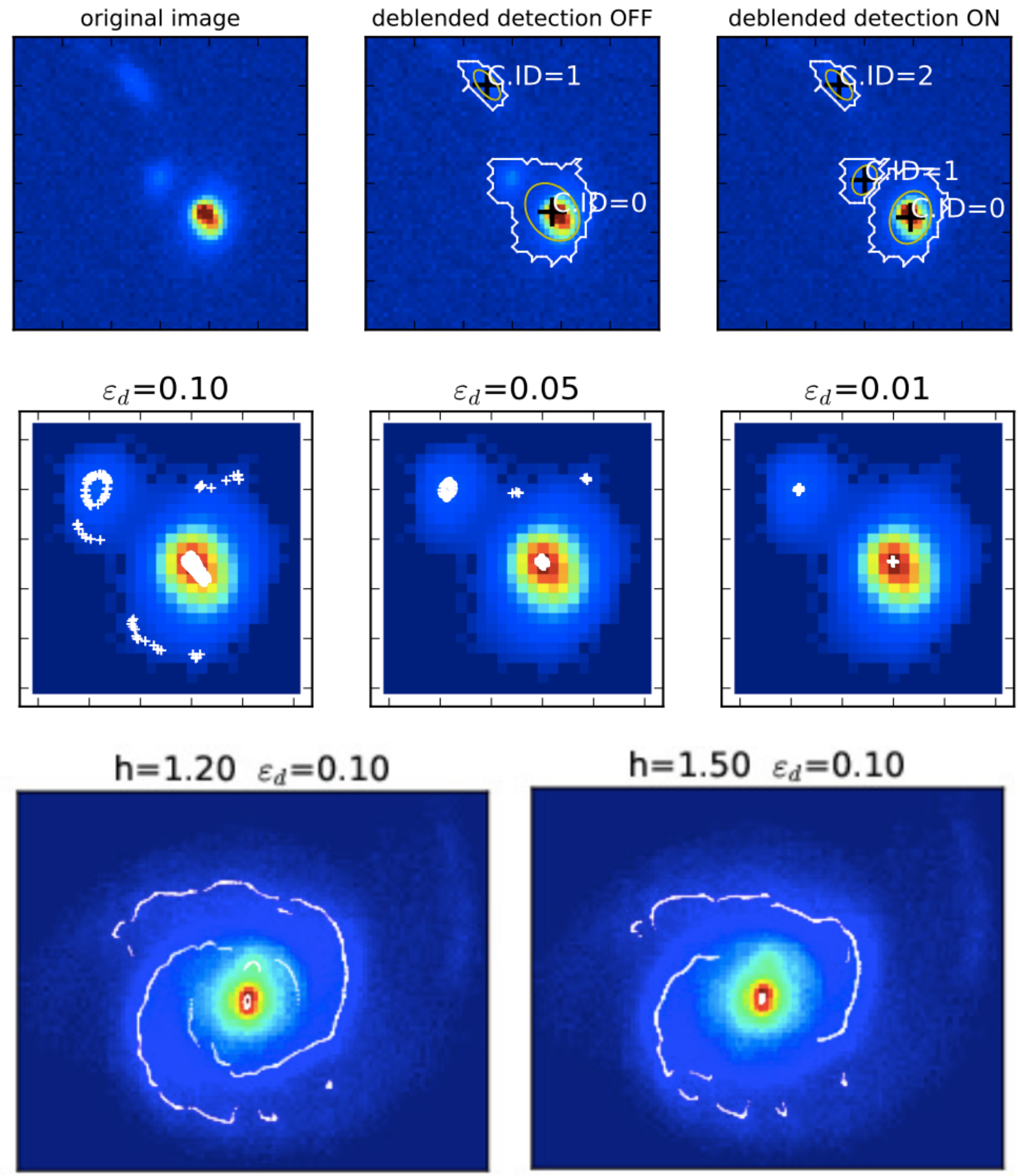
Tramacere et. al 2016

$$f(p_j) \propto \sum_{i=1}^{n} G\left(\frac{\mathbf{q}_j - \mathbf{q}_i}{h}\right) I(p_i)$$

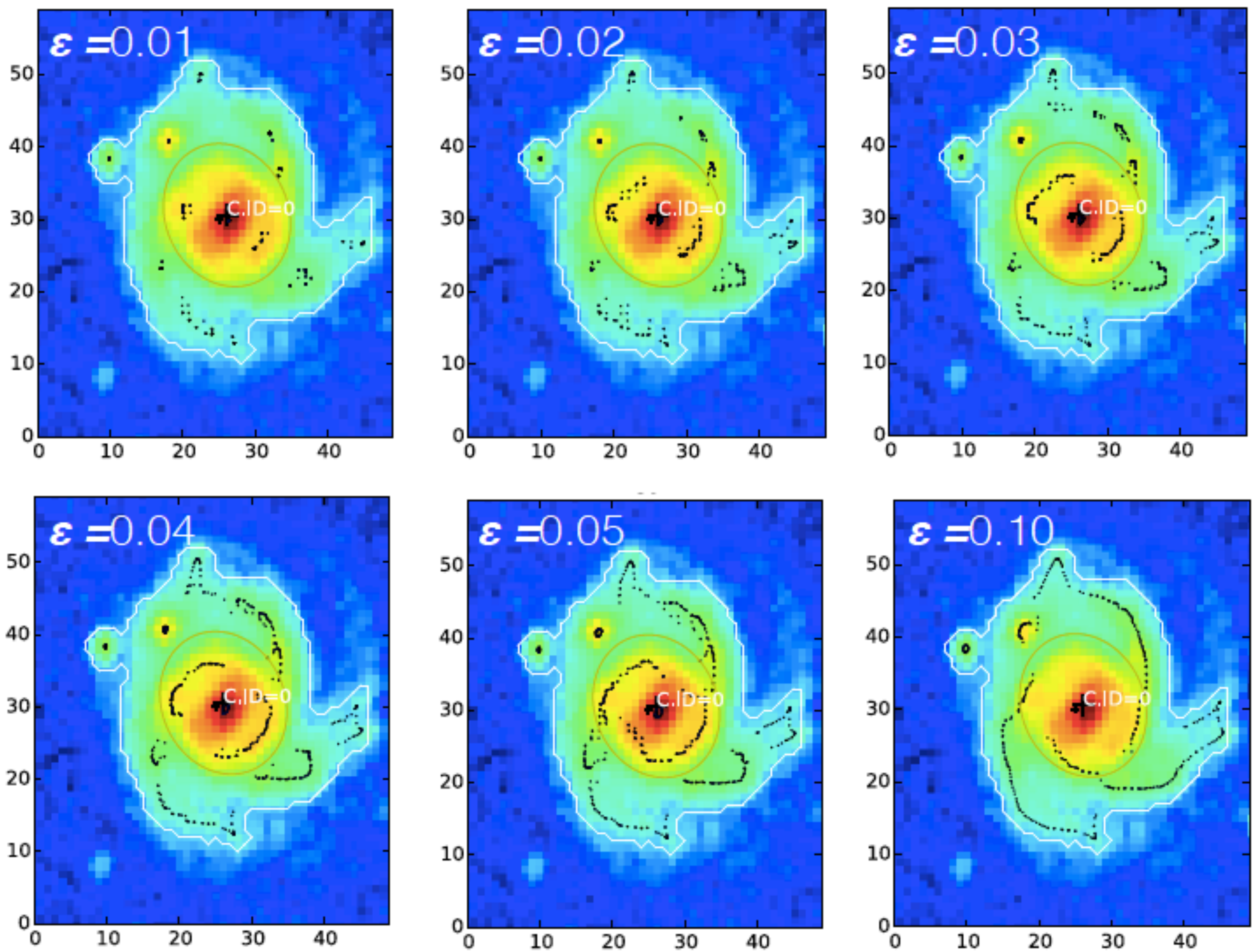
$$G(\mathbf{q}) \propto \exp(-\mathbf{z}\mathbf{z}^{\mathbf{T}}) \qquad \mathbf{z} = \frac{\mathbf{q} - \mathbf{q}_i}{h}$$

$$\mathbf{q}_{t+1} = \frac{\sum_{i=1}^{n} G\left(\frac{\mathbf{q}_t - \mathbf{q}_i}{h}\right) \mathbf{q}_i I(p_i)}{\sum_{i=1}^{n} G\left(\frac{\mathbf{q}_t - \mathbf{q}_i}{h}\right) I(p_i)}$$

$$\|\mathbf{q}_t - \mathbf{q}_{t+1}\| \leqslant \varepsilon_d$$

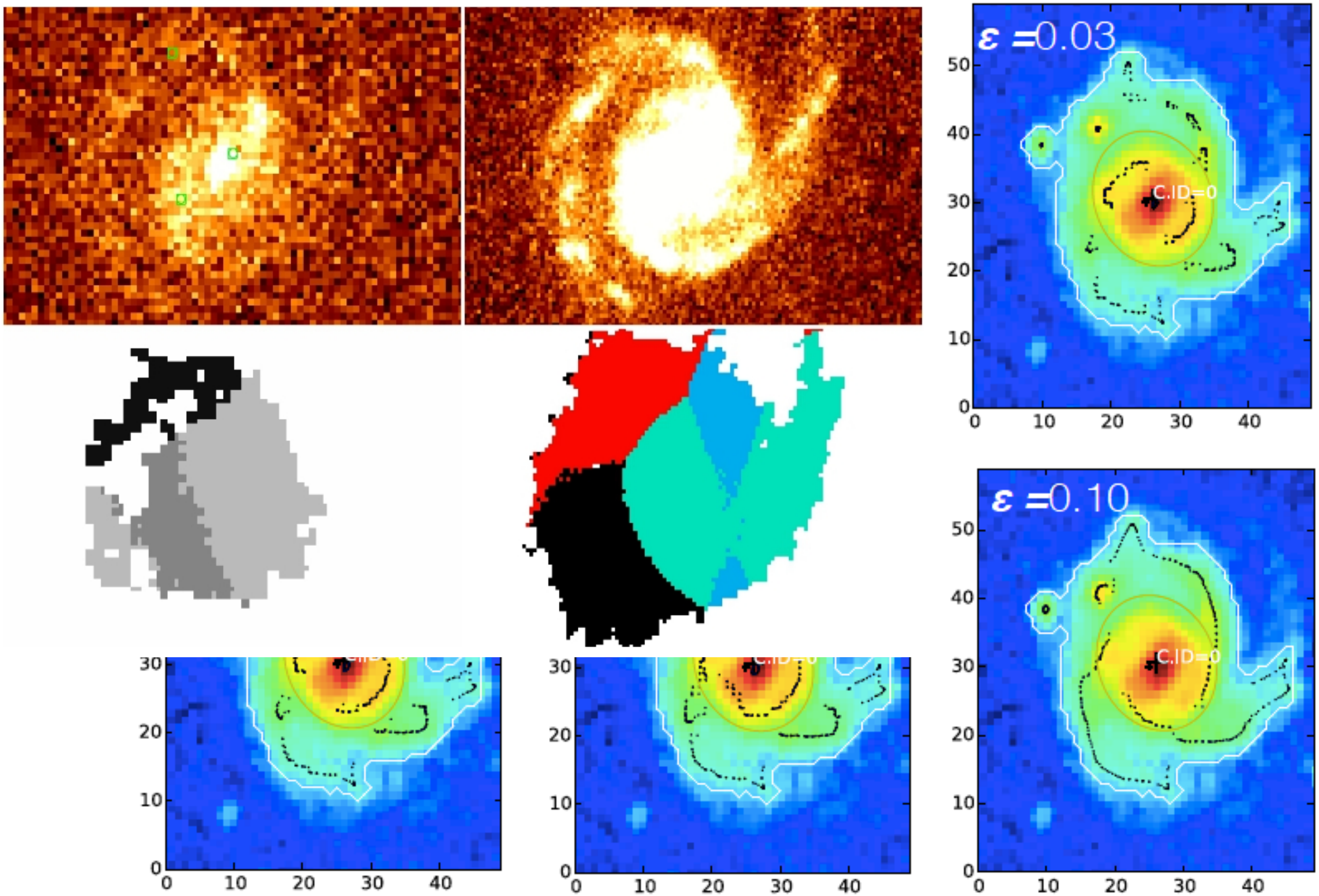


variations of **ε** and ***h*** allow to find more local maxima that follow (if present) spiral patterns, plus the core



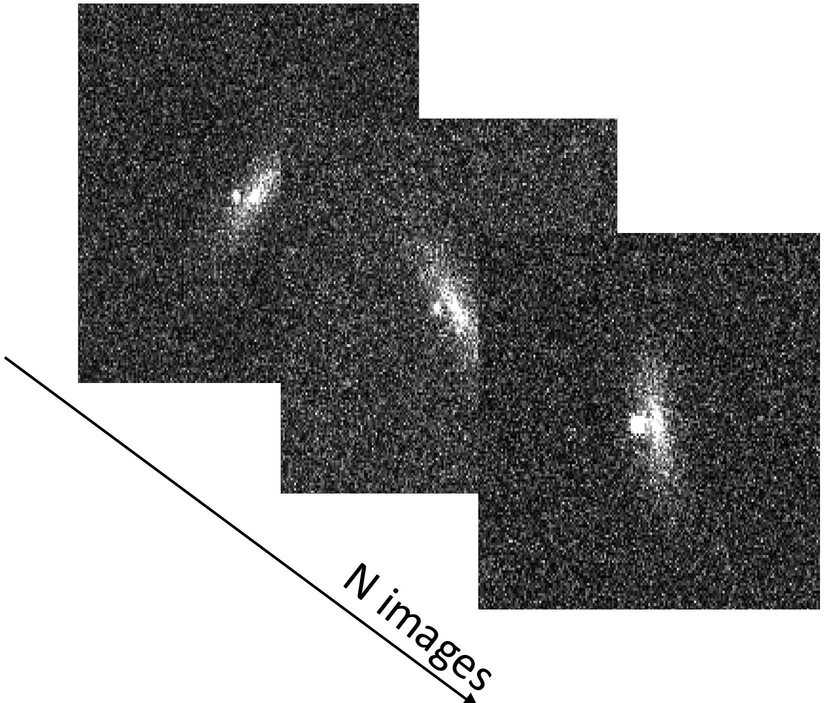
Tramacere et. al 2016

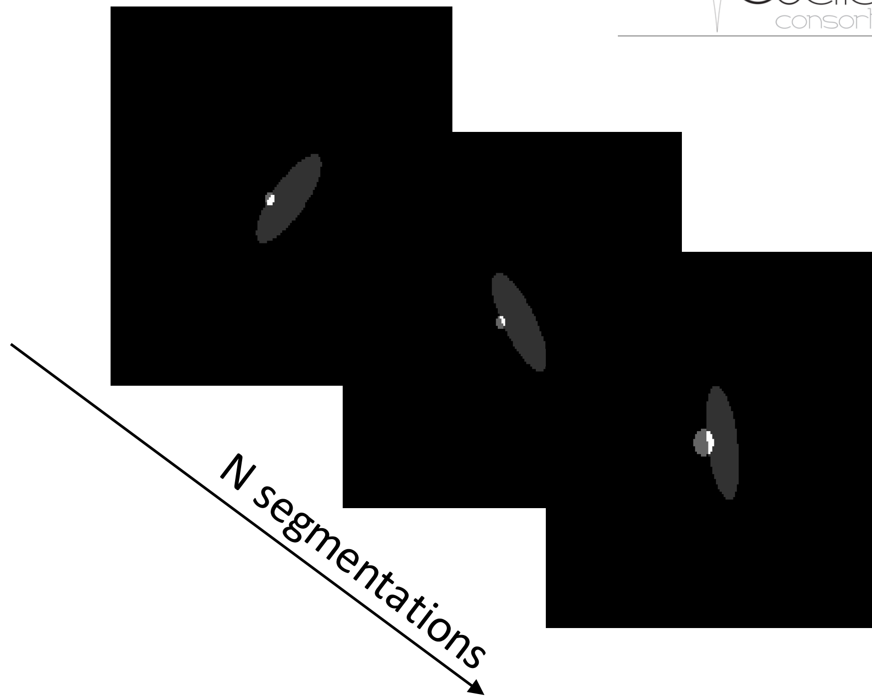
variations of $\boldsymbol{\varepsilon}$ and $\boldsymbol{h}$ allow to find more local maxima that follow (if present) spiral patterns, plus the core



Tramacere et. al 2016

Extensive tests ongoing



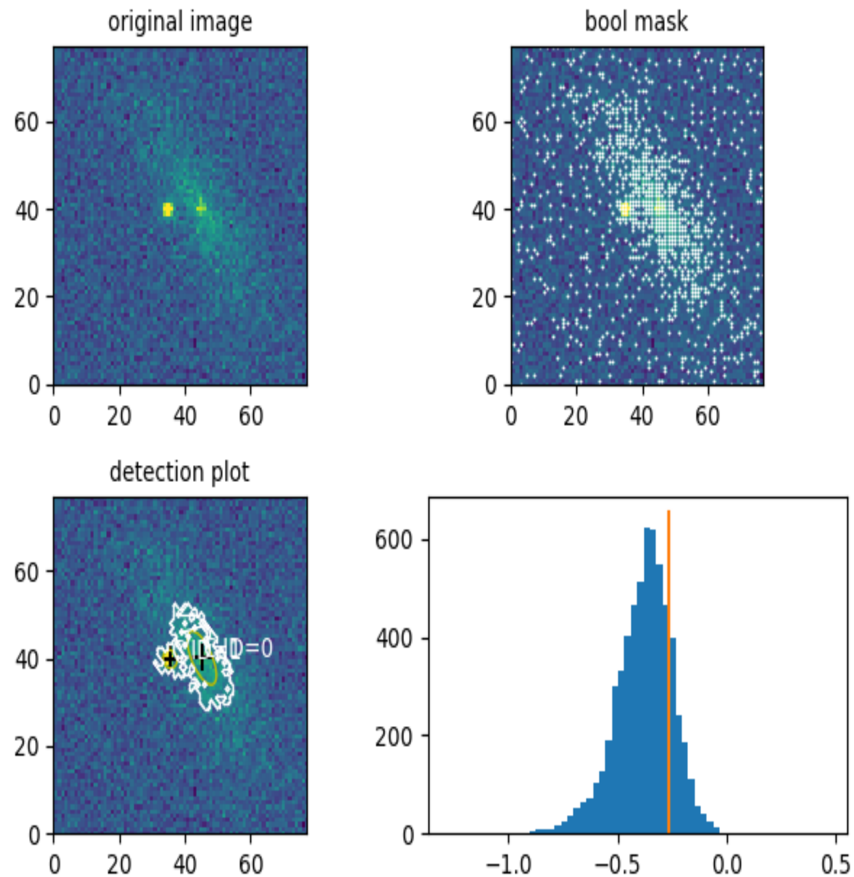


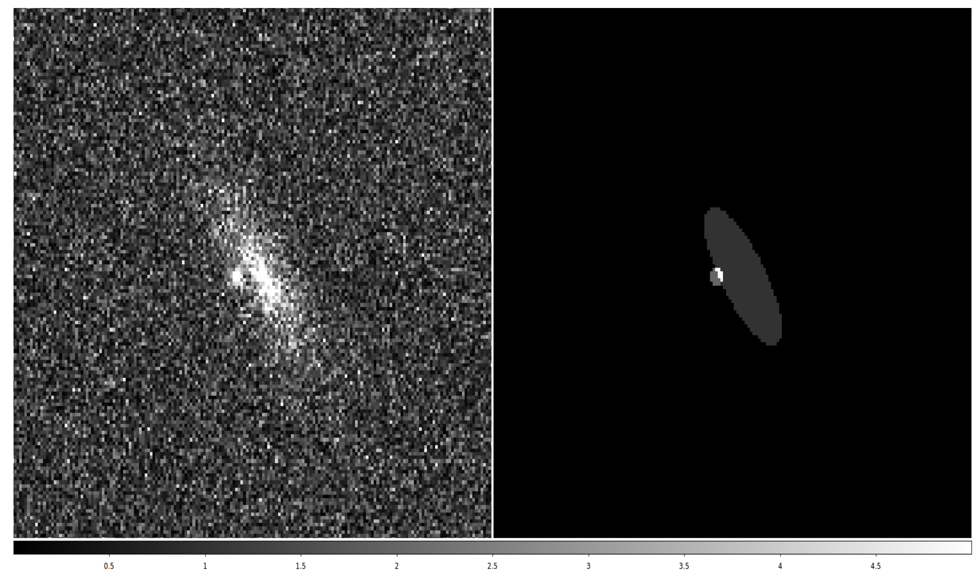
We generate a cube of images with relevant "true" segmentations

Each image is the result of a Skymaker run in which we explore a fair set of parameters

Parameter	Min value	Max value
Magnitude	20	26
Disk radius	1 px	100 px
Bulge/Total ratio	0.01	0.99
Distance between Objects	5 px	30 px
Bulge radius	1 px	100 px

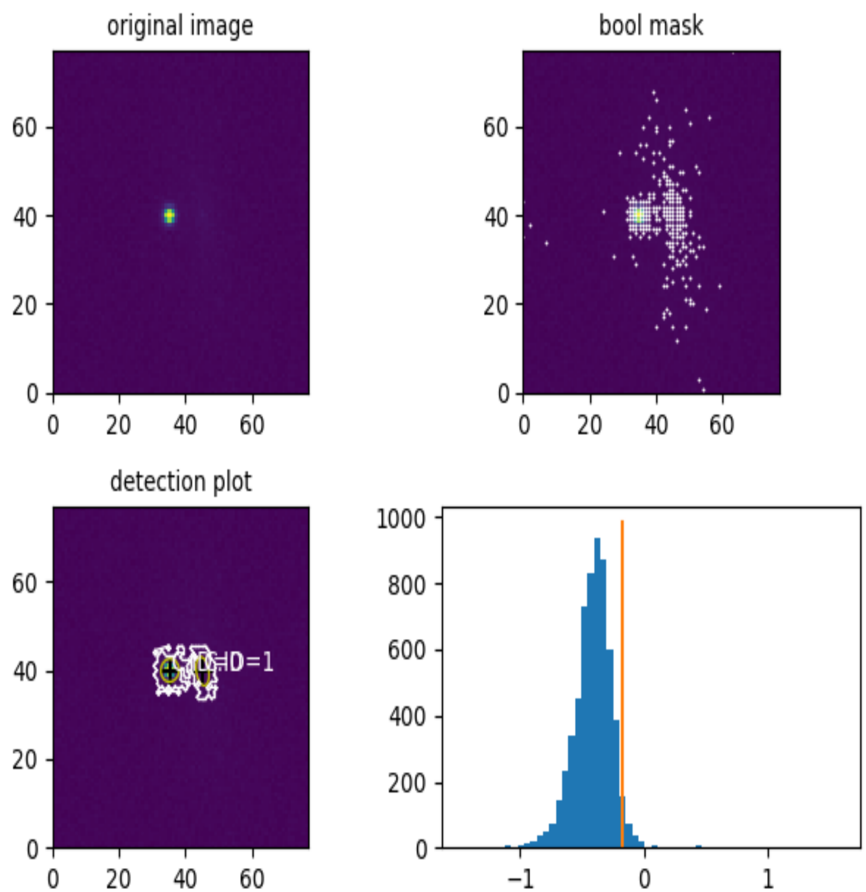
Extensive tests ongoing

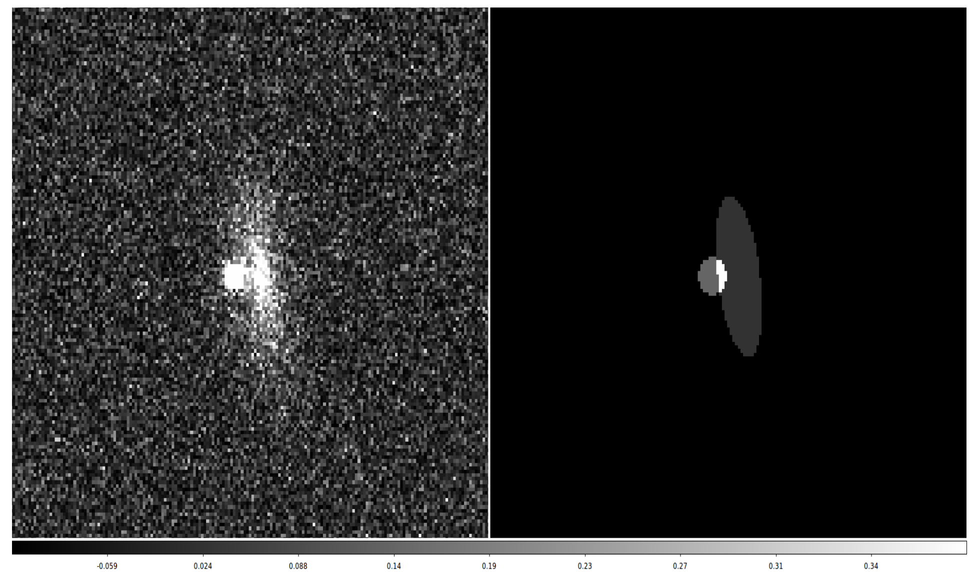




Using the images as input to Asterism we can estimate how well the deblending works for different galaxy types/properties/fluxes

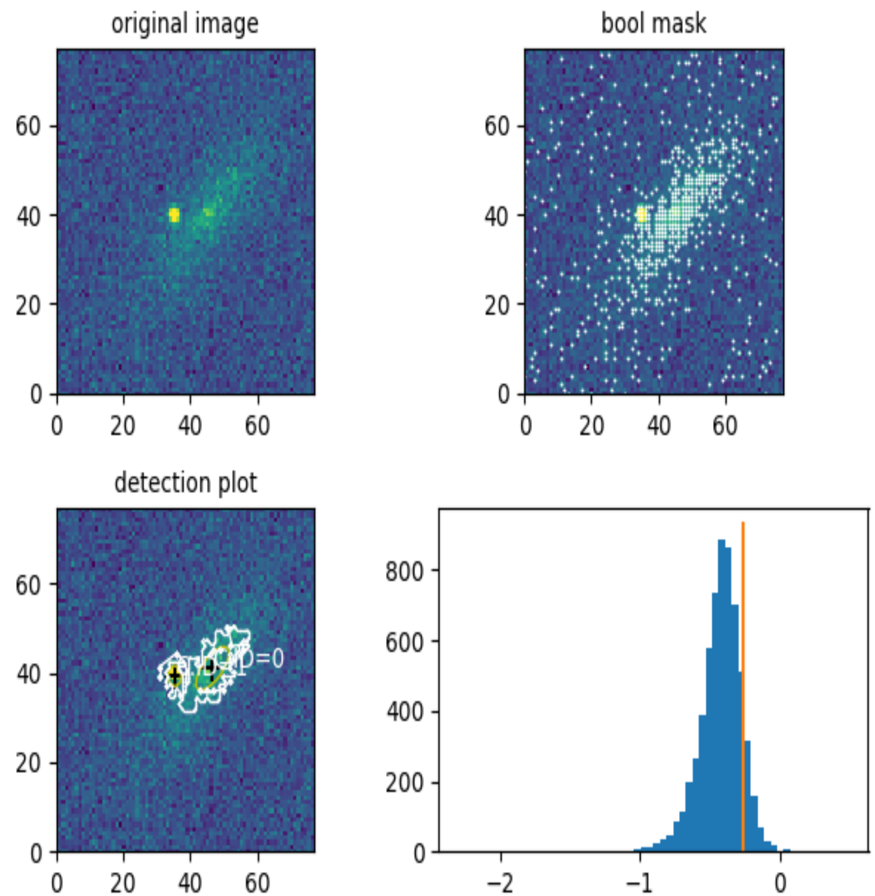
Extensive tests ongoing

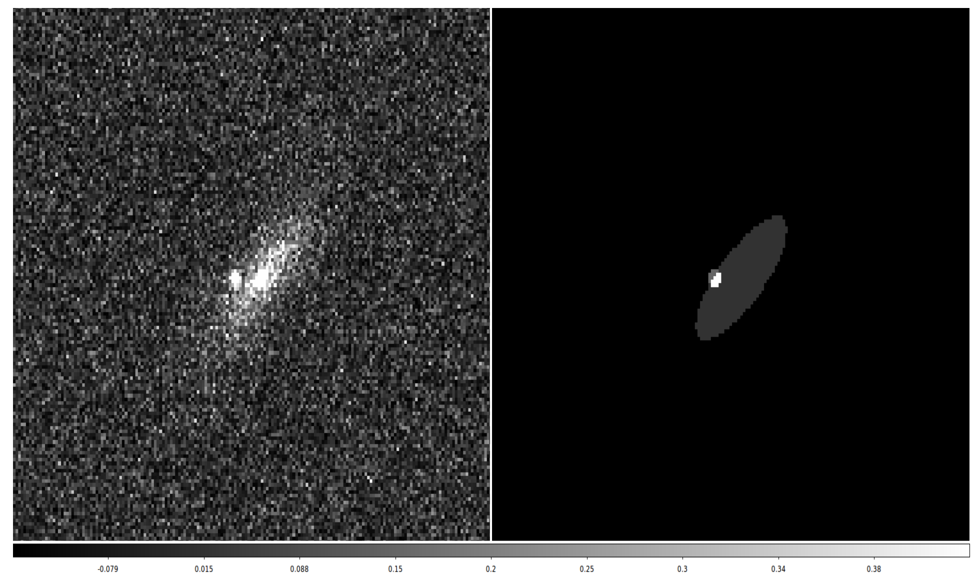




Using the images as input to Asterism we can estimate how well the deblending works for different galaxy types/properties/fluxes

Extensive tests ongoing



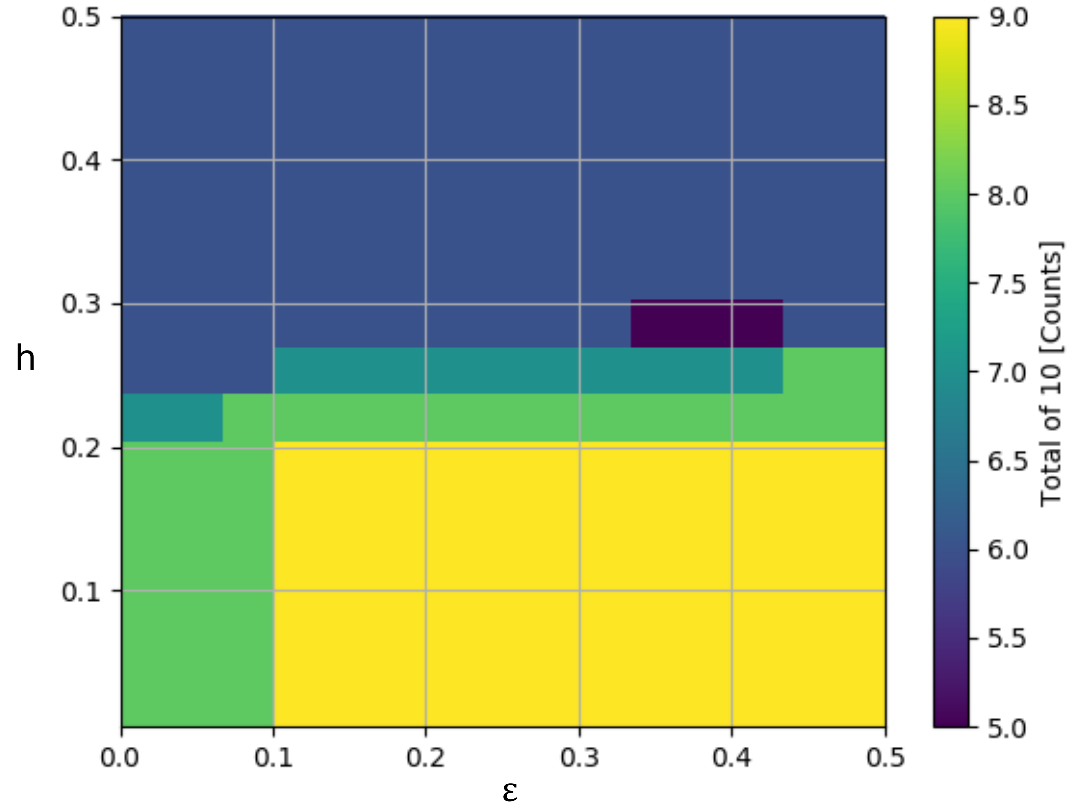


Using the images as input to Asterism we can estimate how well the deblending works for different galaxy types/properties/fluxes

Extensive tests ongoing

We are exploring a range of parameters to find the ones that give the best results on our simulated images

We are generating different datasets with the idea of testing the efficiency of Asterism in many possible cases



Conclusions

- Deblending is the key step leading to the final Euclid source list from a map of “significant” pixels
- The balance between over- and under-deblending is a thorny problem
- First multi-threshold approach in place for Data Challenges and Scientific Challenge 3: good for the bulk of the sources
- Need to move beyond basic approach to address the full range of scales/fluxes/shapes Euclid will observe
- Topometric clustering can provide the required improvement: we are including ASTErIsM algorithms (Tramacere et al. 2016) in the MER PF
- Final PF implementation might require a combination of parametric and machine learning/deep learning approaches