©Elbaz2014

T-PHOT:
Advanced techniques of precision photometry for present and future multiwavelength surveys

Emiliano Merlin

The Spectral Energy Distribution of High Redshift Galaxies

Sexten, Jan 26th 2015

INAF- Osservatorio Astronomico di Roma
INAF - Astronomical Observatory of Rome

ASTRODEEP

ASTRODEEP

INAF – OAR: A. Fontana, E. Merlin, R. Amorin, M. Castellano, P. Santini, K. Boutsia

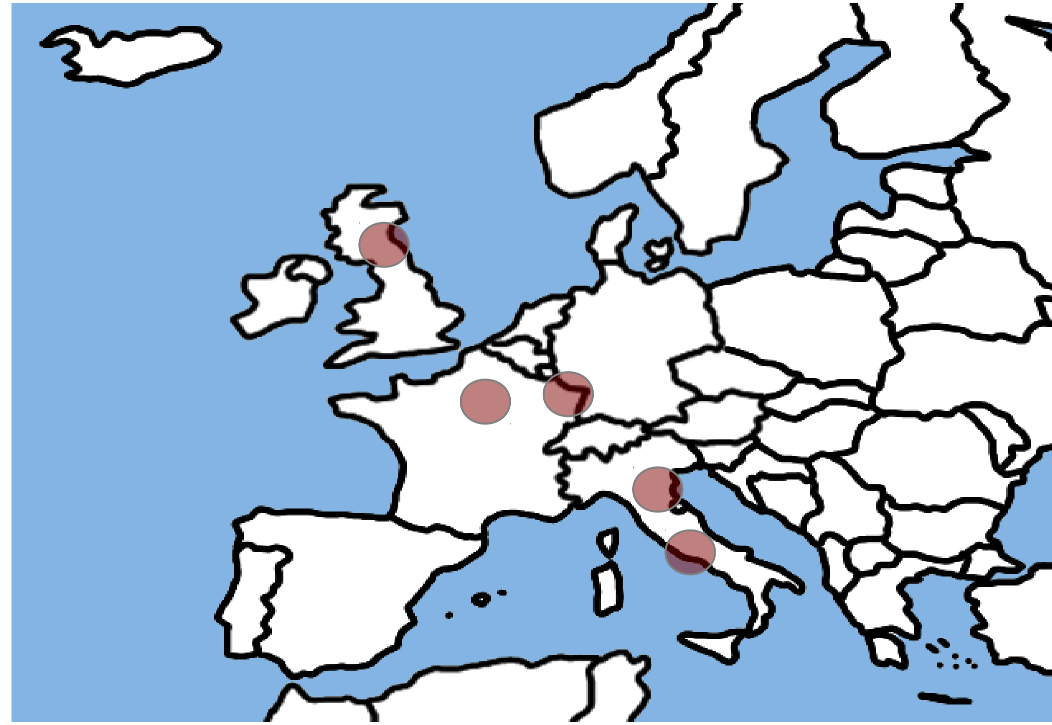
INAF – OABO: A. Comastri+

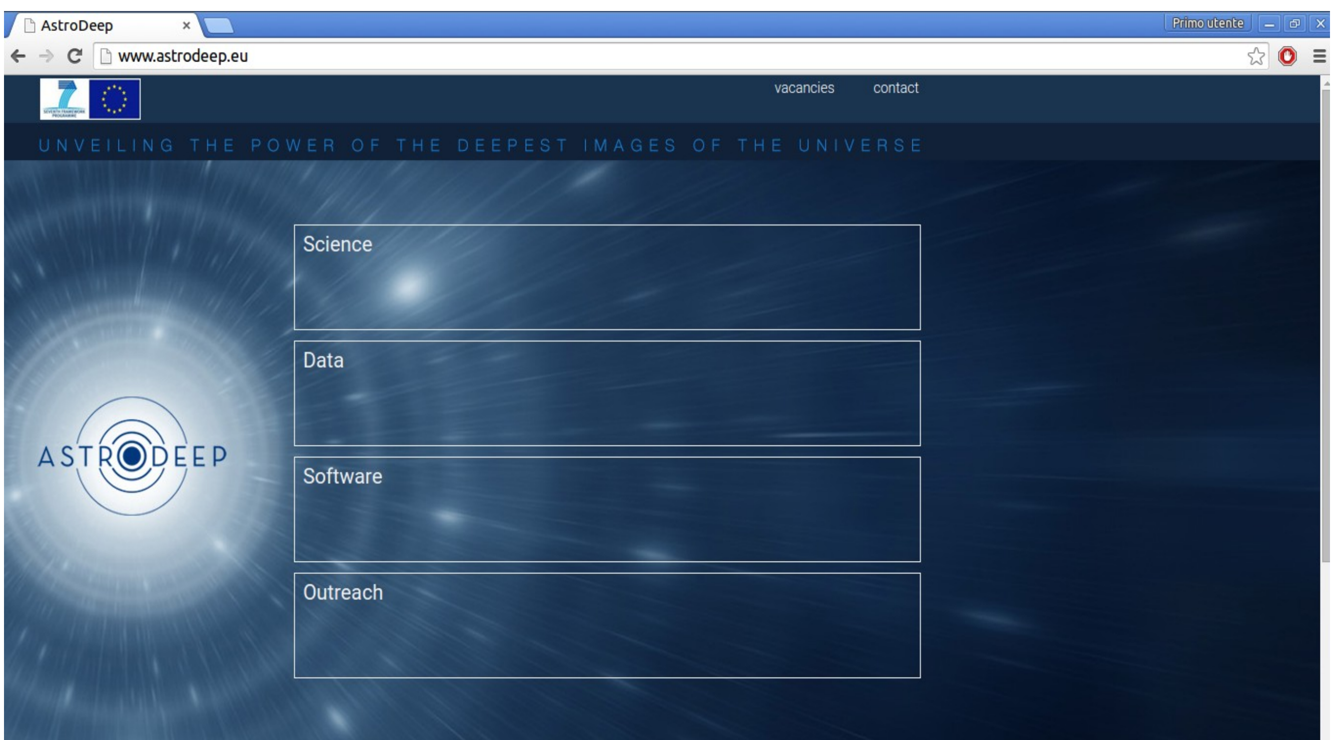
CEA – Saclay: D. Elbaz+

Univ. of Edinburgh: J. Dunlop+

CNRS – Strasbourg: S. Derriere

STScI – Baltimore: H. Ferguson+

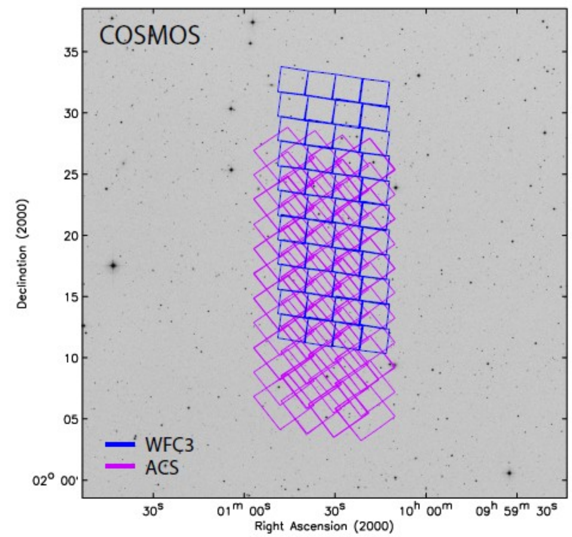


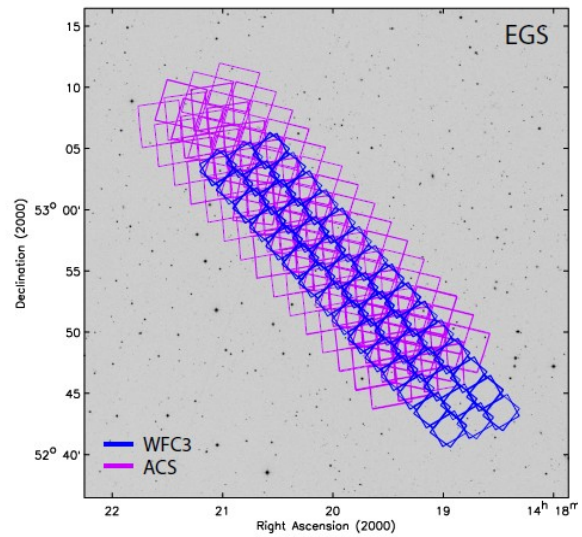


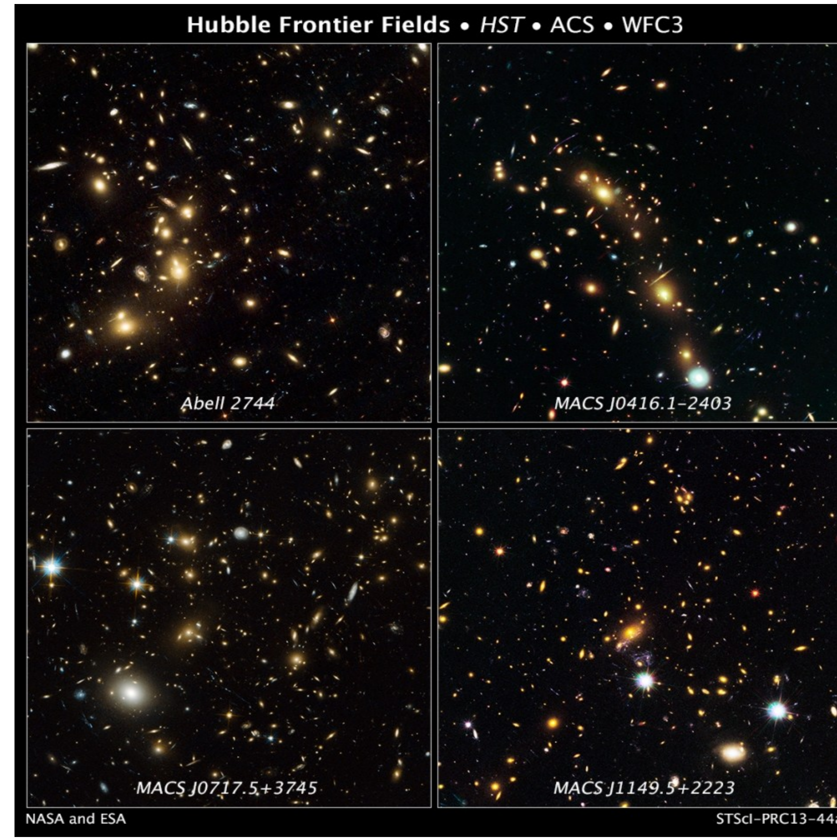
2013 - 2016

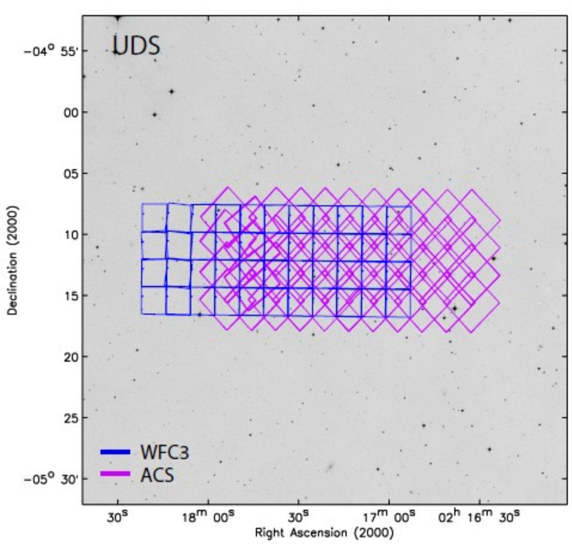
www.astrodeep.eu

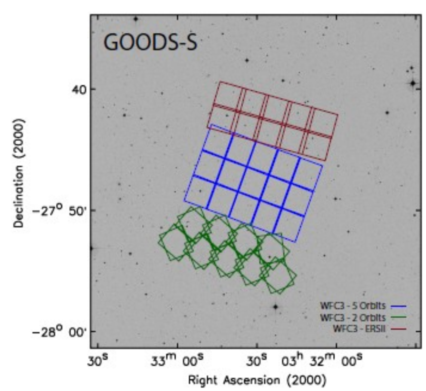
ASTRODEEP goal #1: produce complete, up to date multiwavelength photometric catalogs of available deep fields

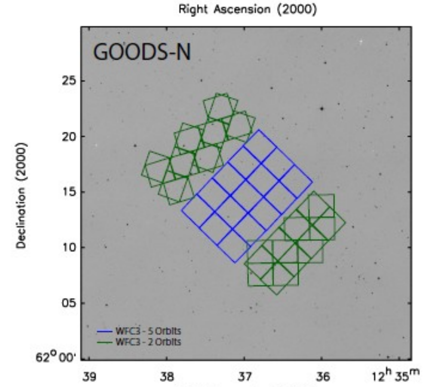








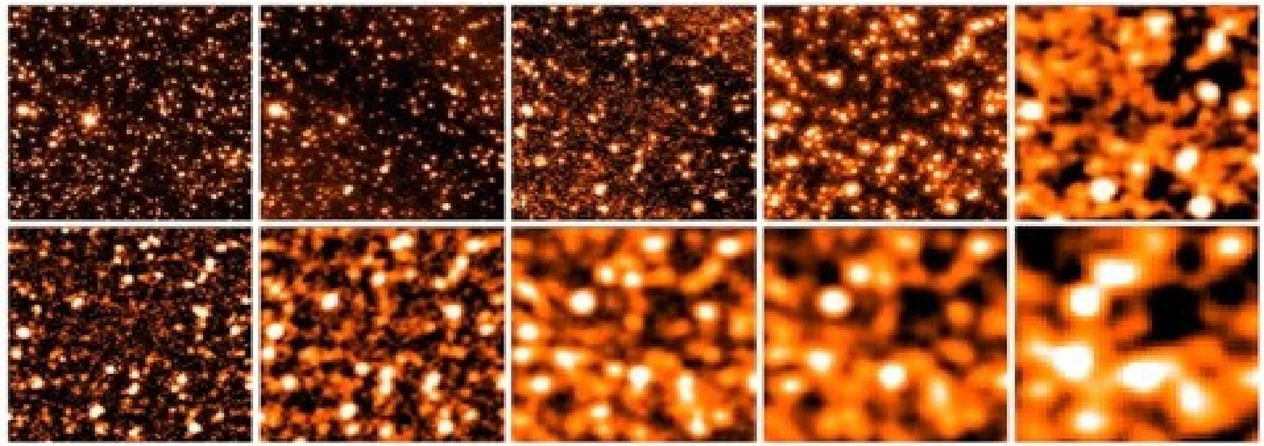




5 CANDELS fields + 4x2 FRONTIER fields

FIR: Herschel
MIR / NIR: Spitzer
NIR / Opt / UV: HST, ground (Subaru, Hawk-I, ...)
X: Chandra

ASTRODEEP goal #2: set a “best” standard procedure, develop and publicly release dedicated software tools



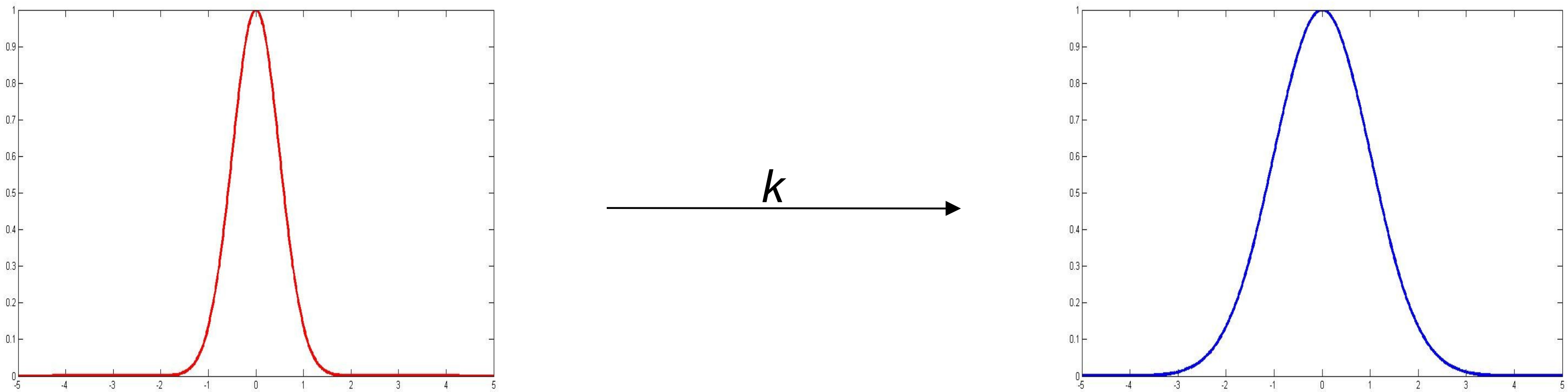
Main concern: *confusion/blending/overlapping* of sources at decreasing resolution and increasing wavelength

T-PHOT (Merlin+2015, in prep.): A code for PSF-matched photometric analysis of multiwavelength data using priors

Acknowledgment:
Adriano Fontana, Marco Castellano, Stefano Pilo, Konstantina Boutsia, Nathan Bourne, Tao Wang, Xinwen Shu, Koryo Okumura, Corentin Schreiber, Andrea Grazian, Kuang-Han Huang, Henry Ferguson, Mimi Song, Alice Mortlock, Hooshang Nayyeri, Mauro Stefanon

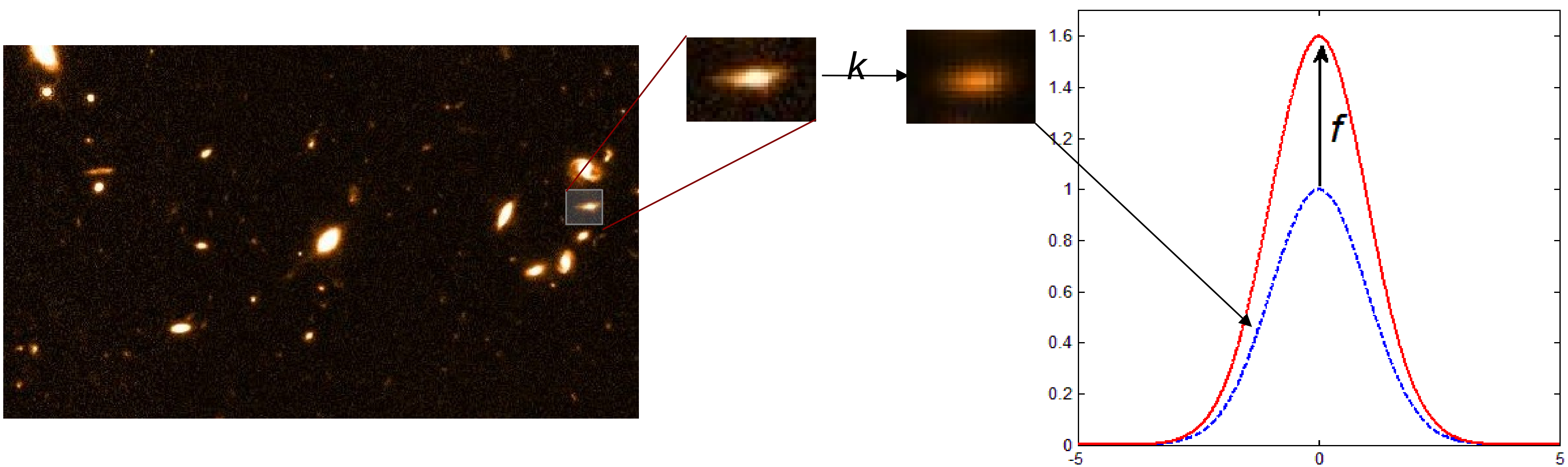
PSF-MATCHED MULTIWAVELENGTH PHOTOMETRY: BASIC METHOD

- Convolve PSFs and obtain convolution Kernel

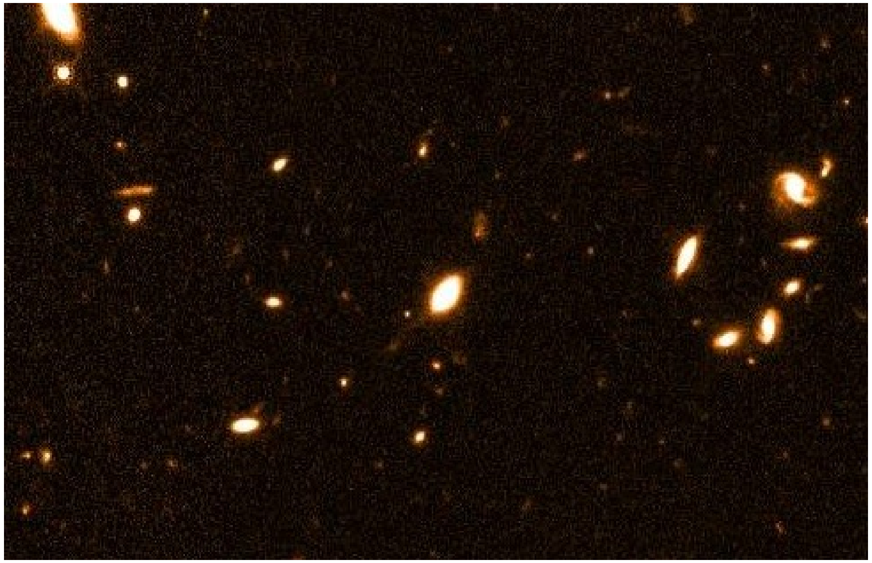


*PSF_LRI = k * PSF_HRI*

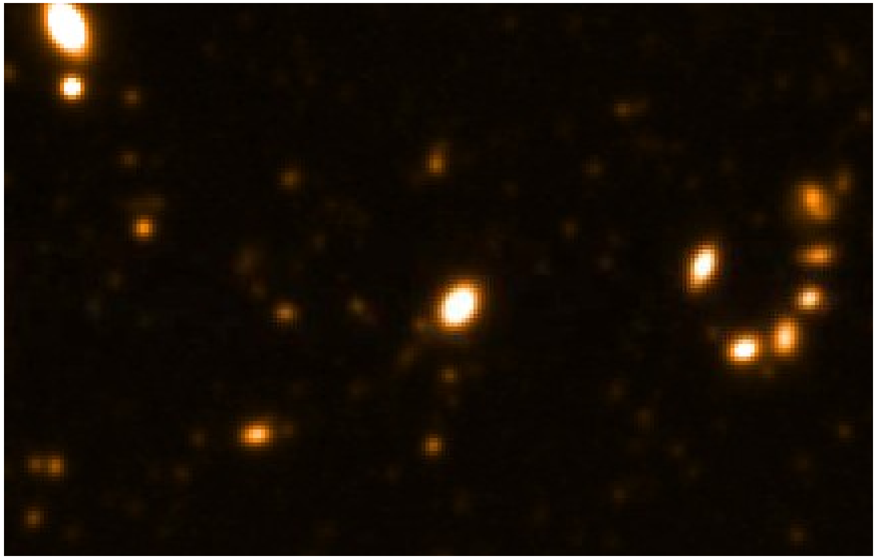
k = F⁻¹ [F(PSF_LRI) / F(PSF_HRI)]

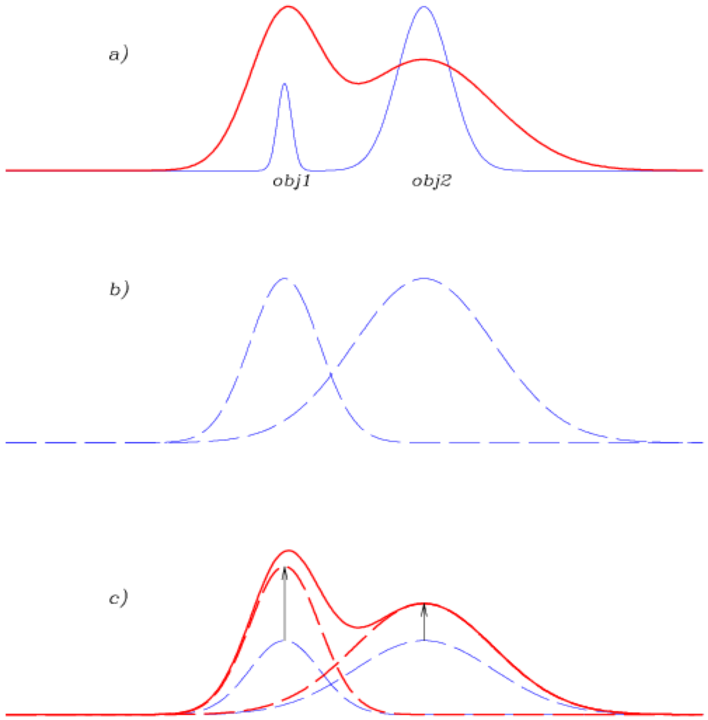


PSF-MATCHED MULTIWAVELENGTH PHOTOMETRY: BASIC METHOD



$\xrightarrow{k}$





$$\chi^2 = \sum_{m,n} \left[\frac{I(m,n) \qquad - M(m,n)}{\sigma(m,n)} \right]^2$$

$$M(m,n) = \sum_{i=1}^{N_{\rm obj}} F_i P_i(m,n)$$

$$\chi^2 = \sum \left[\frac{I(m,n) - \Sigma_i F_i P_i(m,n)}{\sigma(m,n)} \right]^2$$

$$\frac{\partial \chi^2}{\partial F_i} = 0, \quad i = 1, \ldots, N_{sources}$$

$$\frac{\partial \chi^2}{\partial F_i} = -2 \sum \frac{IP_i}{\sigma^2} + 2 F_i \sum \frac{P_i^2}{\sigma^2} + 2 \sum F_j \sum \frac{P_i P_j}{\sigma^2}$$

$$\boxed{AF = B}$$

$$A = \begin{bmatrix} A_{11} & A_{12} & \ldots & A_{1i} \\ A_{21} & A_{22} & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ A_{i1} & \ldots & \ldots & A_{ii} \end{bmatrix}, \quad A_{ij} = \sum \frac{P_i P_j}{\sigma^2}$$

$$B = \begin{bmatrix} B_1 \\ B_2 \\ \ldots \\ B_i \end{bmatrix}, \quad B_i = \sum \frac{IP_i}{\sigma^2}$$

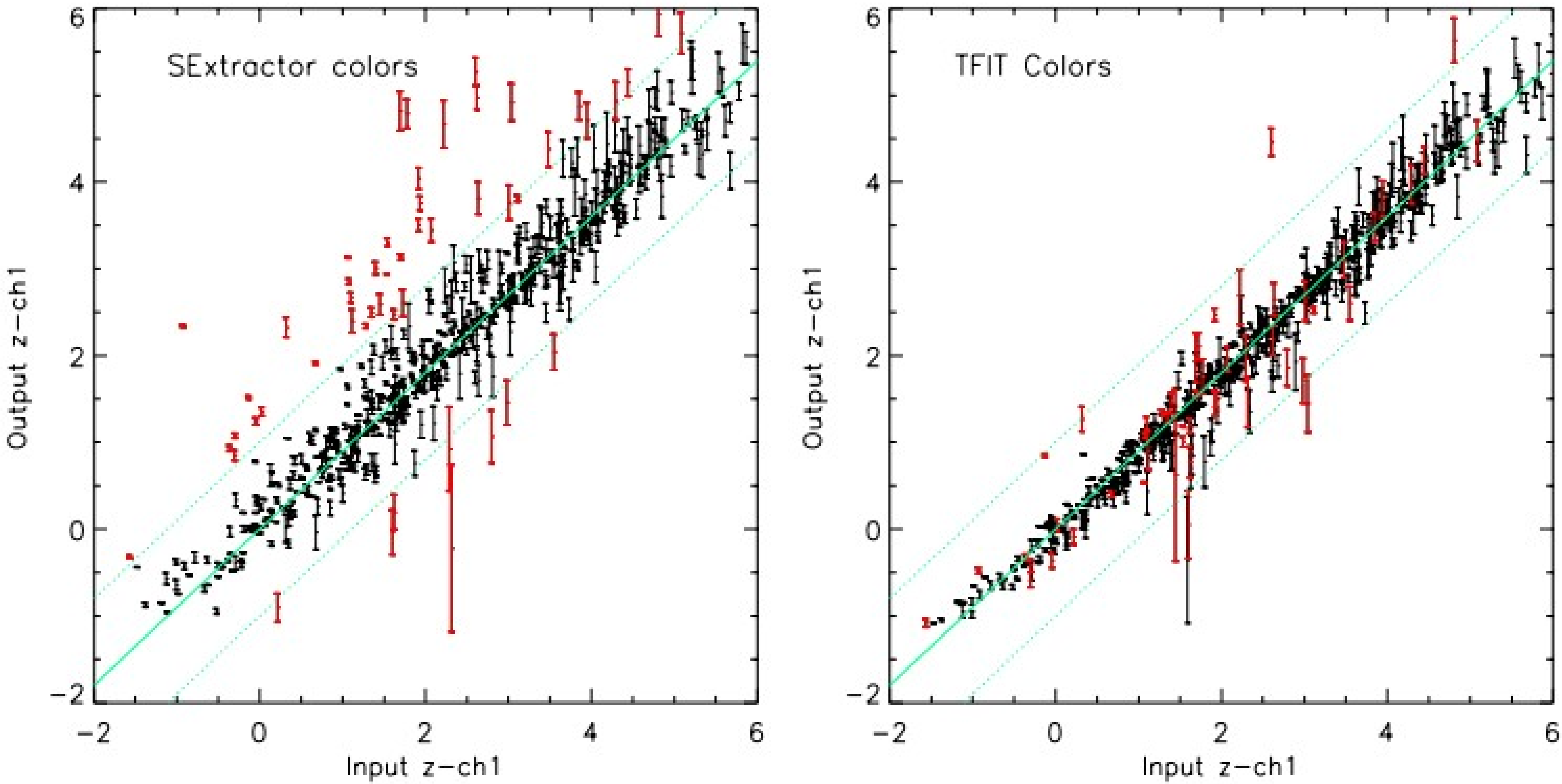


Figure 1. Optimal multi-wavelength photometry by TFIT. Extended sources with a range of colors are simulated to compare the performance of conventional aperture photometry and TFIT. Left: the $z - ch1$ (ACS z_{850} band − *Spitzer* [3.6 μm]) colors measured by SExtractor are compared with the input colors for 600 simulated galaxies. Large scatter is expected due to source blending and confusion. Right: the colors of the same sources measured by TFIT. The sources with the "aperture color" bias >1 mag are shown in red on both panels. Note that most of the same sources are successfully recovered in TFIT-derived colors.

TFIT (STScI – Papovich+ 1999, Laidler+ 2007):

- C++ core (fitting) + Python envelop
- 10,500 lines .py + .cc (plus external libraries)
- Requires many external tools (Python modules, IRAF, etc.)
- Cell fitting + Dithering; best flux chosen geometrically

- 24 hours on a typical field (mostly because of Python slowness in preparation and post-fit stages)

CONVPHOT (OAR – DeSantis+ 2007):

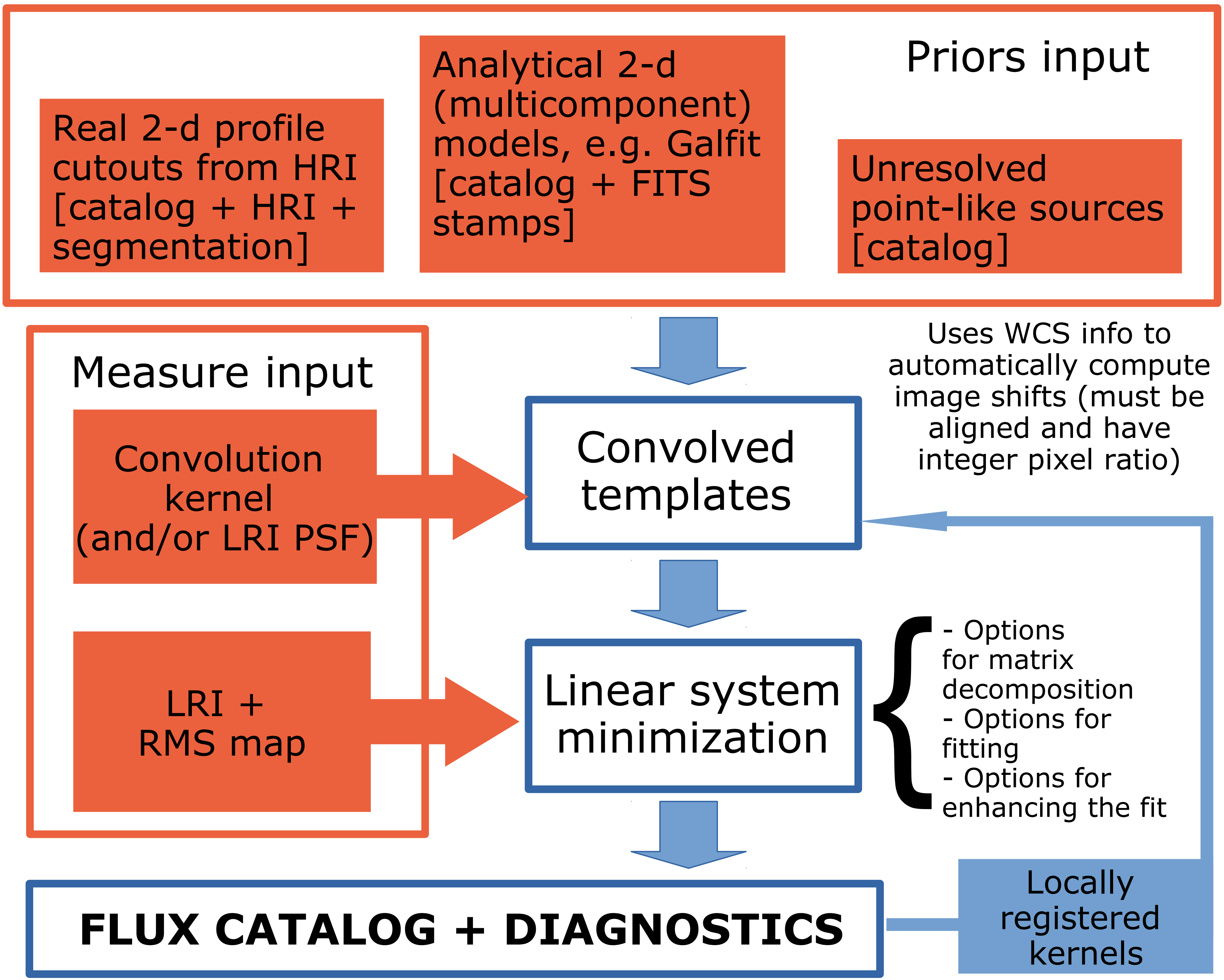
- C
- 4,200 lines .c (plus external libraries)
- Single fit on whole image
- No FFT convolution

- 24 hours on a typical field (mostly because of pixels summation convolution and of fitting procedure)

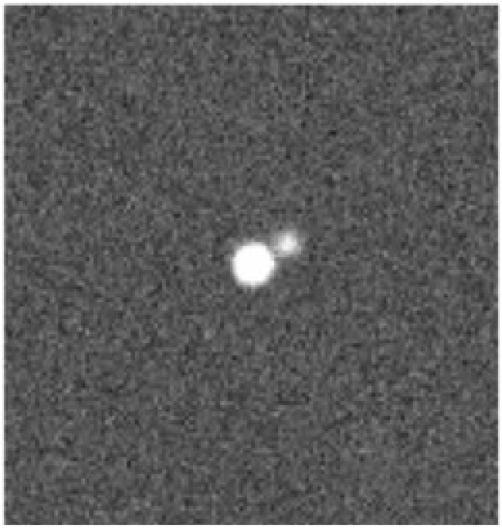
T-PHOT (Merlin et al. 2015, in prep.)

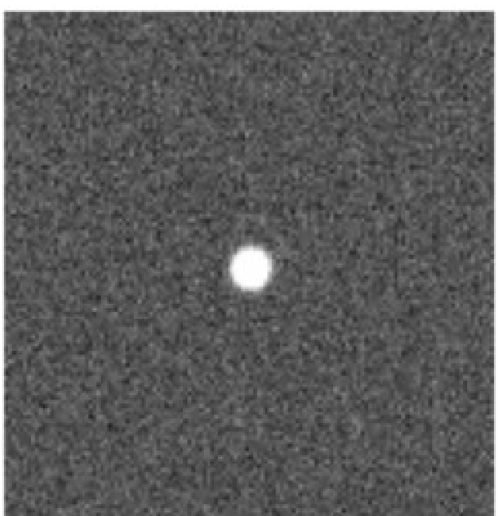
* **Python envelop, C/C++ cores**
* Clearly organized in "stages" (similarly to TFIT)
* **Fast**: ca. 30 mins. on a "standard" CANDELS field with TFIT parameters
* Robust, and can handle large datasets with smart memory allocation
* Only needs Python modules Numpy, Astropy and Matplotlib, plus CFITSIO and FFTW3 (no IRAF, STDAS, anfft)
* **Versatile**: includes all different choices and methods already present in TFIT and CONVPHOT concerning smoothing (pixel summation or FFT), fitting (cells vs. single fit, three methods for matrix solving, threshold, clipping of negative sources), dance stage for kernel registration

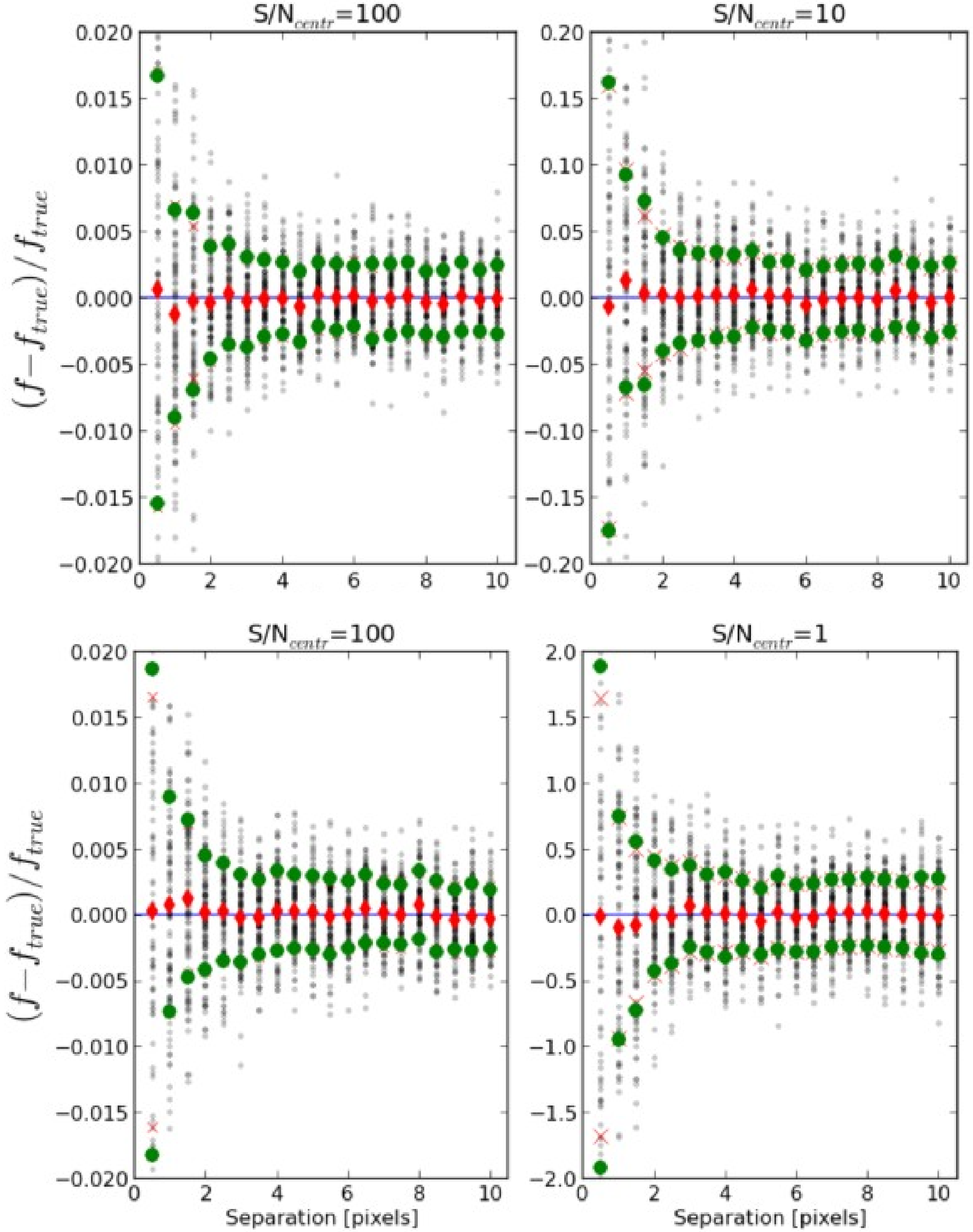
* Includes a **cells-on-objects method**, which combines the computational efficiency of TFIT cells approach and the robustness of CONVPHOT single fit method
* Can operate with **three different types of priors: real 2-d cutouts from HRI, analytical models, or unresolved point-like sources**


Priors input
Real 2-d profile cutouts from HRI [catalog + HRI + segmentation]
Analytical 2-d (multicomponent) models, e.g. Galfit [catalog + FITS stamps]
Unresolved point-like sources [catalog]
Uses WCS info to automatically compute image shifts (must be aligned and have integer pixel ratio)
Measure input
Convolution kernel (and/or LRI PSF)
Convolved templates
LRI + RMS map
Linear system minimization
- Options for matrix decomposition
- Options for fitting
- Options for enhancing the fit
FLUX CATALOG + DIAGNOSTICS
Locally registered kernels

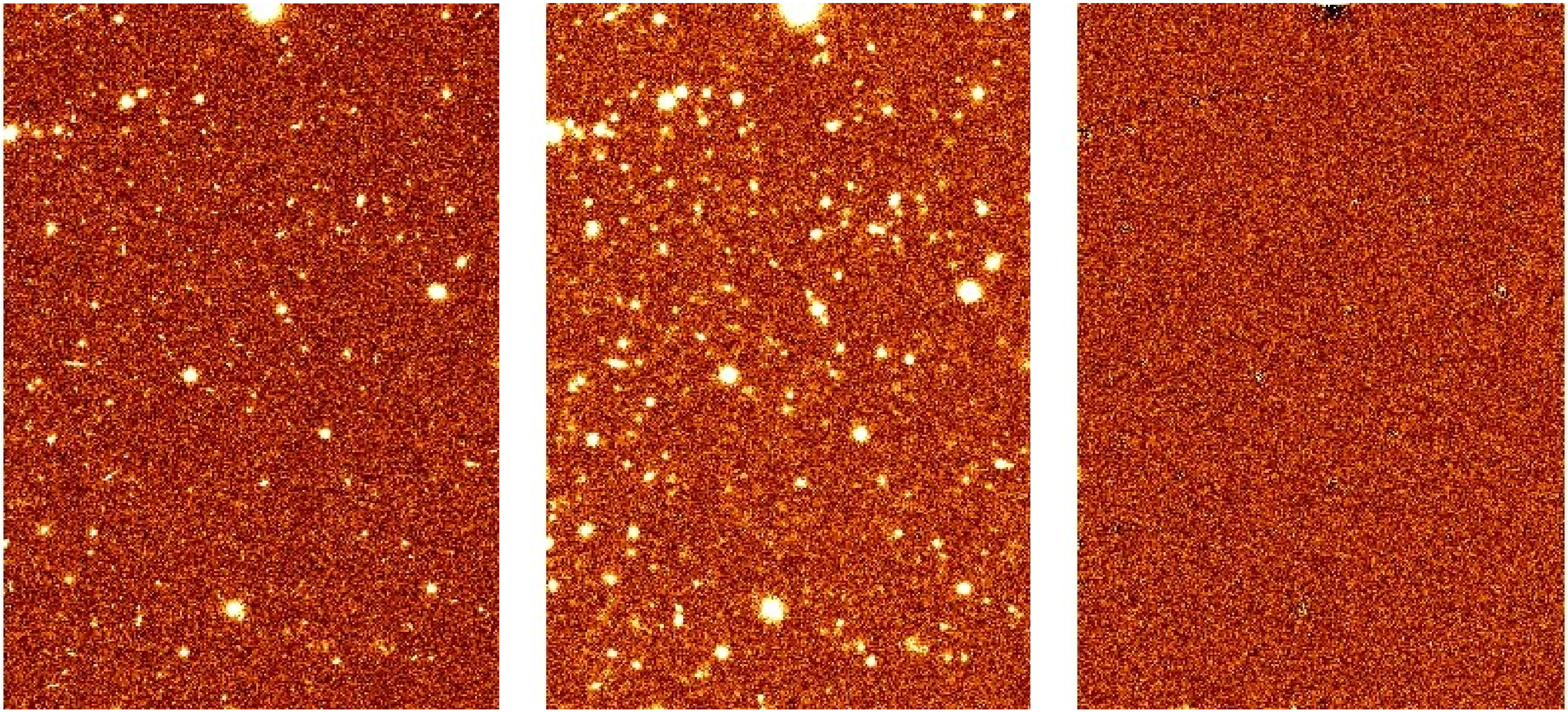
Basic testing
(with Koryo Okumura)





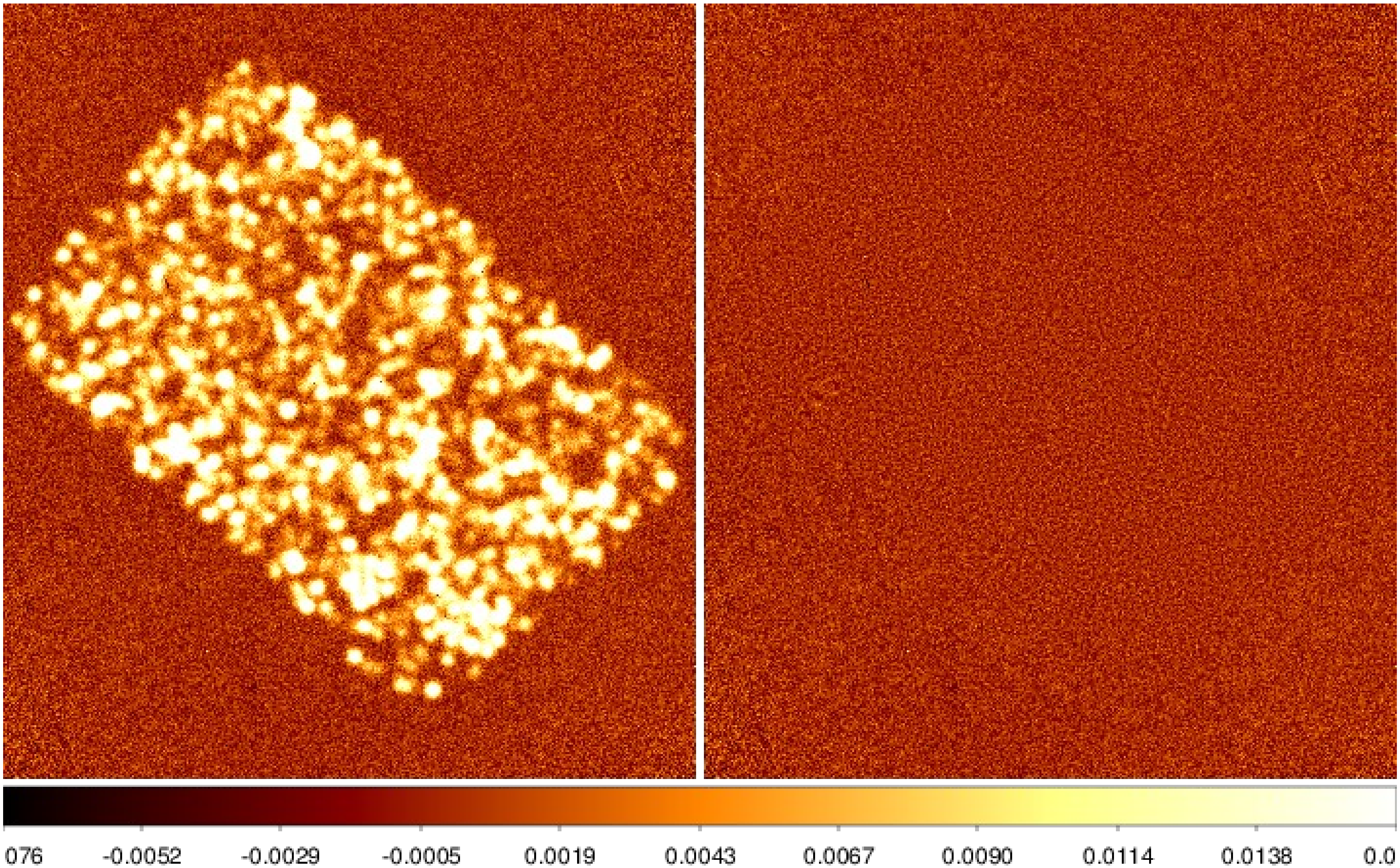


Testing: simulated datasets – Extended sources (E. Bertin's Stuff+SkyMaker used to produce realistic images)



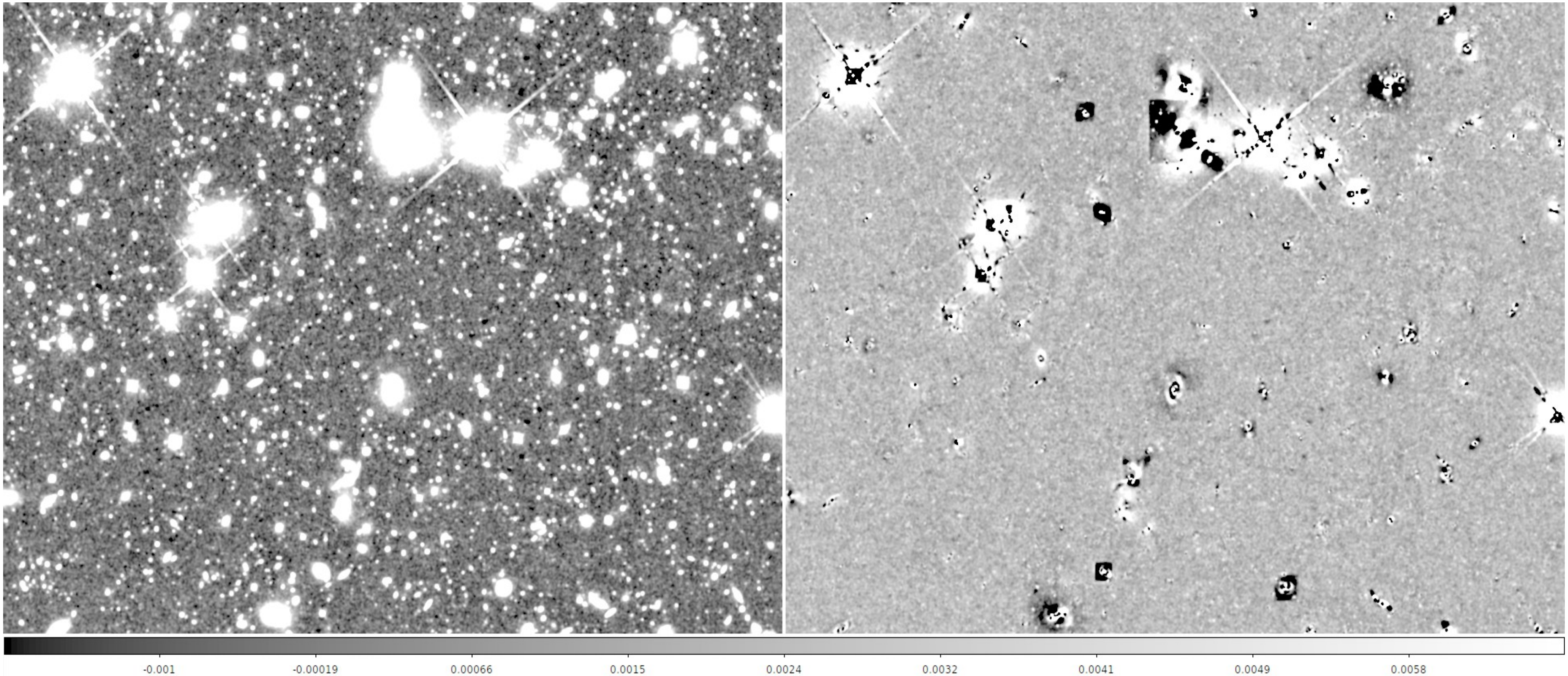
Left: simulated HRI (fwhm=0.2”). Center: simulated LRI (fwhm=1”).
Right: residuals image [T-PHOT **whole image fit using “real” priors**]

Testing: simulated datasets – PSF-shaped sources (with Xinwen Shu and Tao Wang)

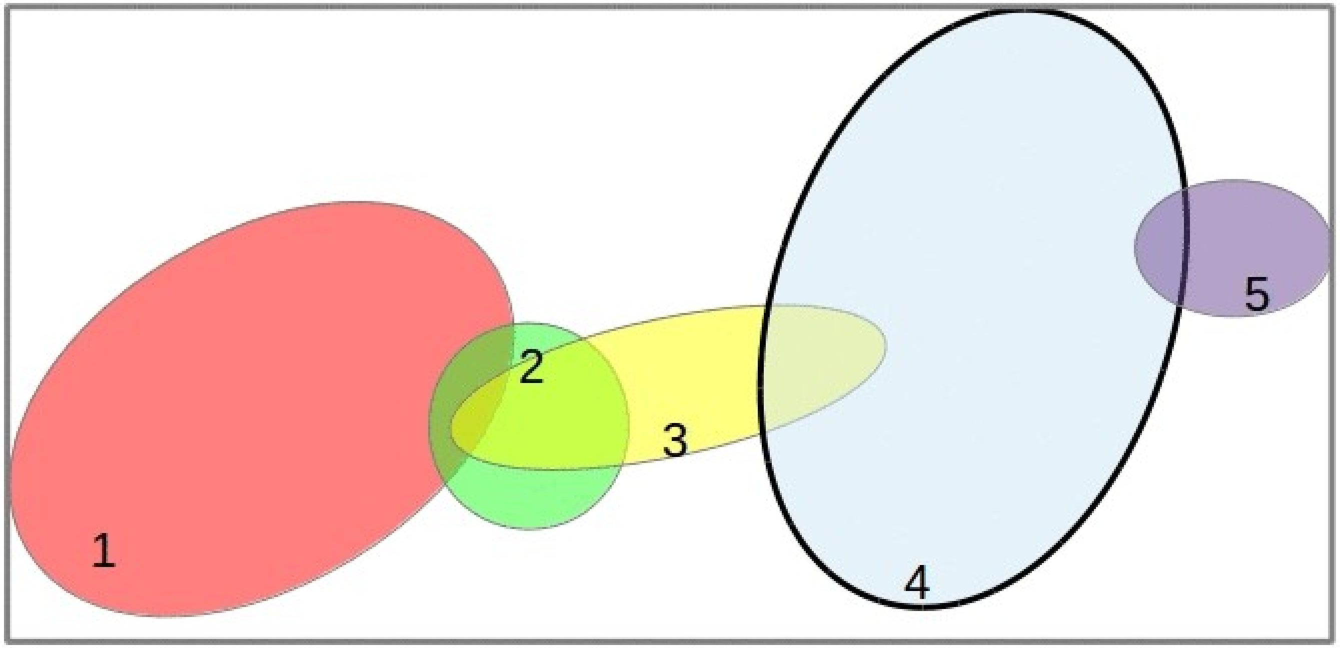


Left: simulated LRI (SPIRE). Right: residuals image
[T-PHOT **whole image fit using unresolved point-like priors**]

Testing: simulated datasets – using analytical models (with Fernando Buitrago)

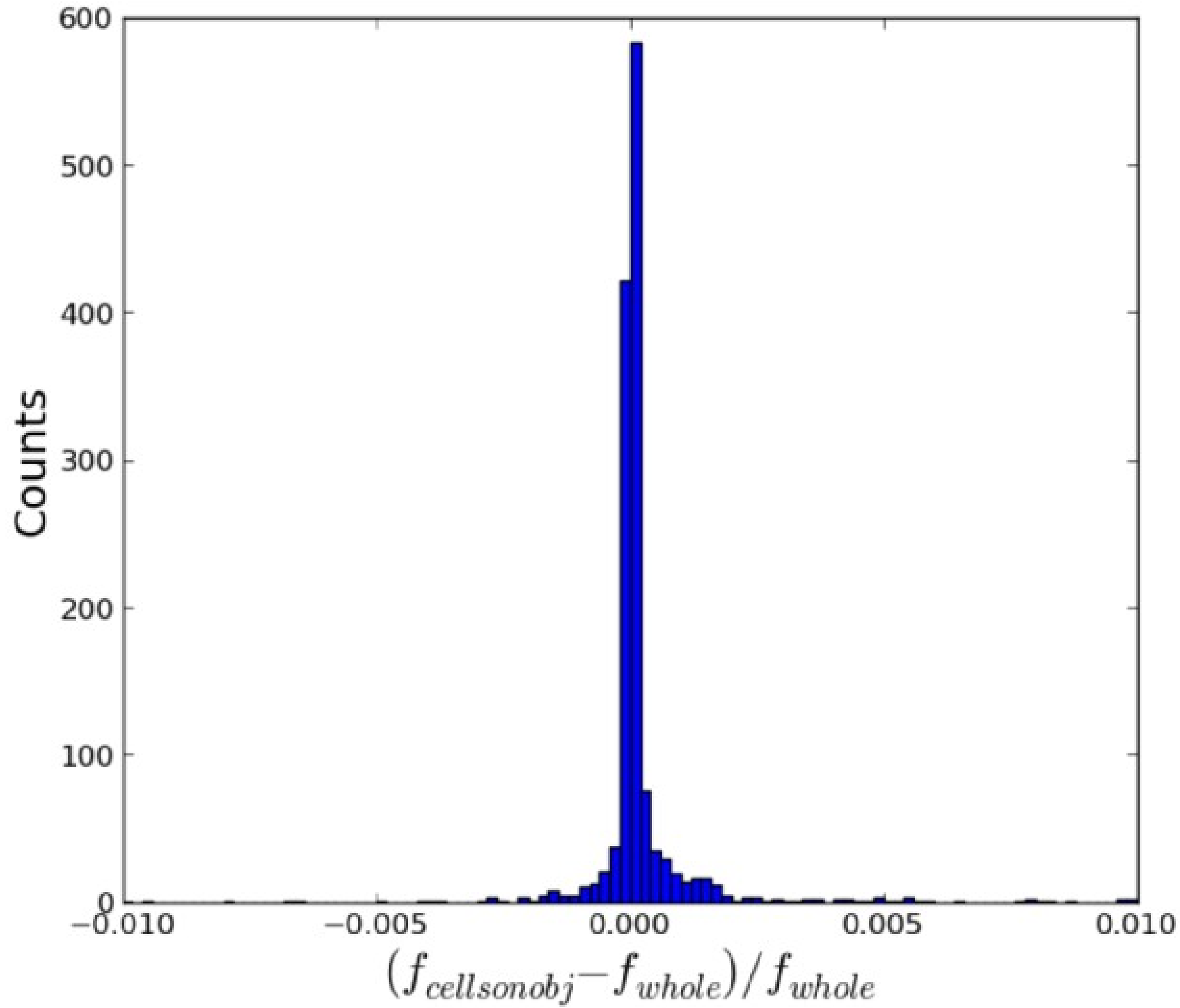


Left: simulated LRI (COSMOS H band smoothed to R).
Right: residuals image
[T-PHOT **whole image fit using MEGAD Galfit models**]



CELLS-ON-OBJECTS method:

* One cell per object
* Include all first order contaminants
* Higher order contaminants are all included unless:
  - they are fainter than given fraction of the total flux of the central object (use the Detection image flux as a proxy)
  - OR they only overlap to the previous level contaminant with a small fraction of their total area
* Once the central object is fitted, it is subtracted from the Measure image; catasptrophic contamination is excluded
* NOTE: trying to keep fits for other “non central” objects in a given cell proves unsatisfactory

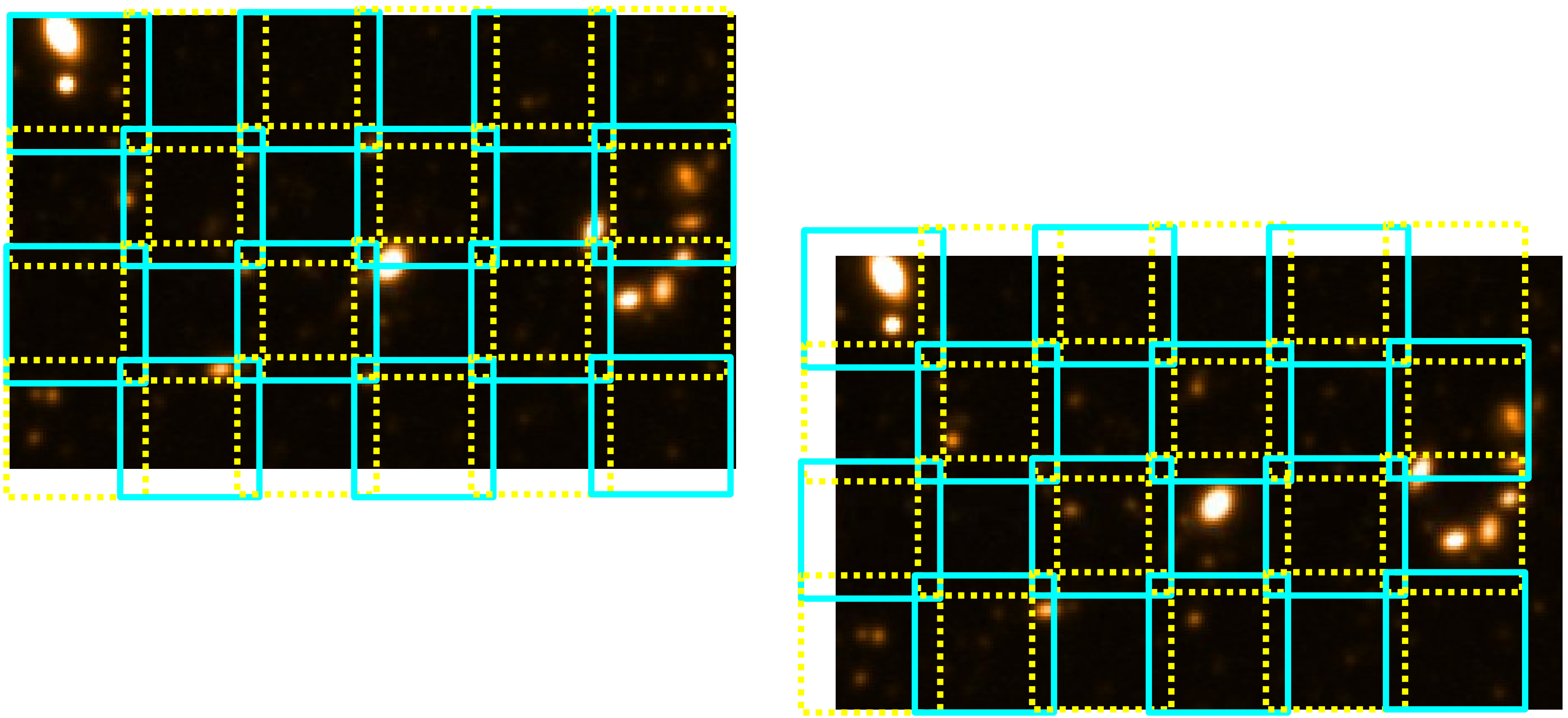


Comparison of measured flux using single image fitting and *cells-on-object* method on the same simulated field

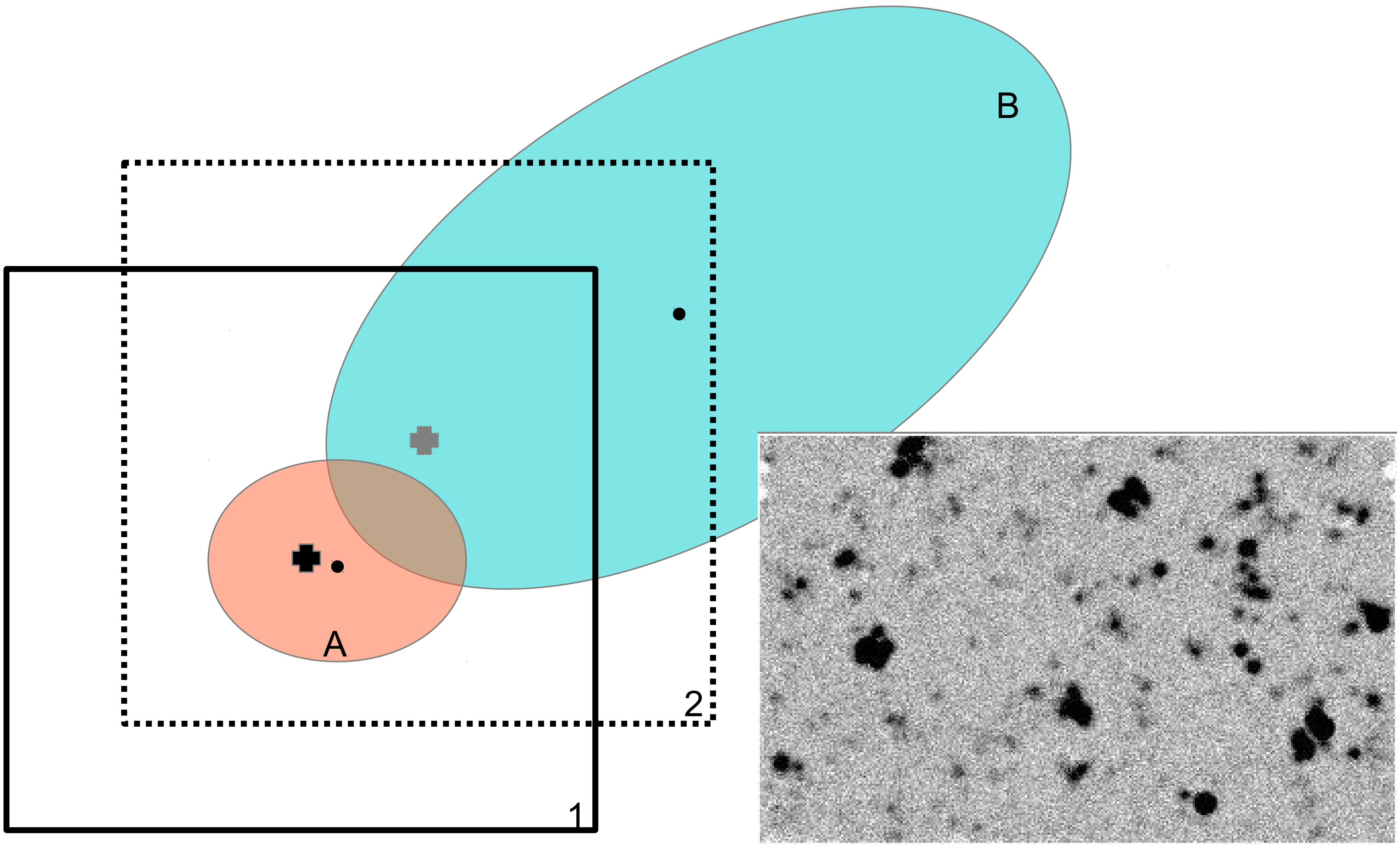
TFIT **(STScI – Laidler+ 2007):**

Quite fast fitting, thanks to cell fitting:

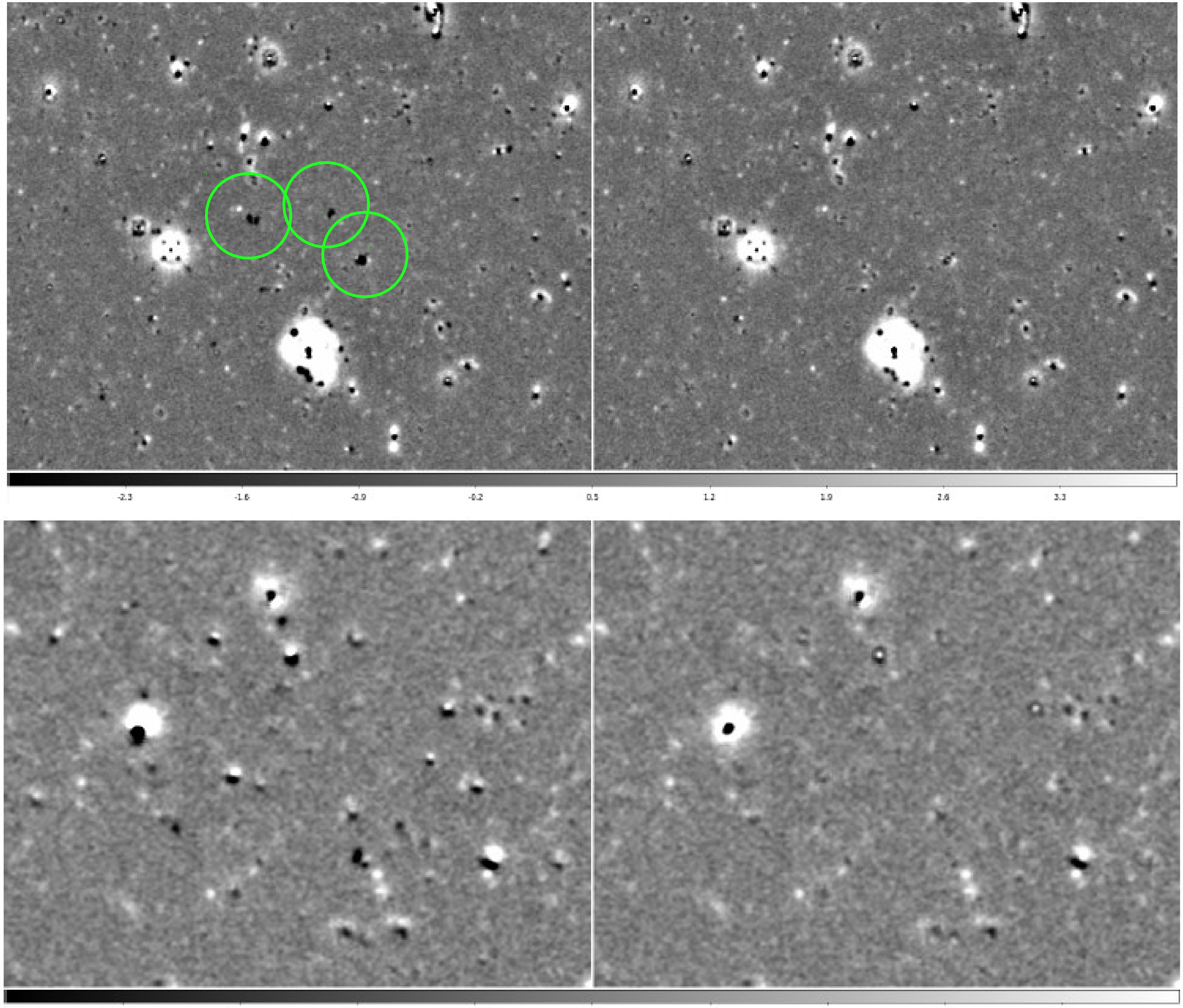
Cell fitting + Dithering; best flux chosen geometrically



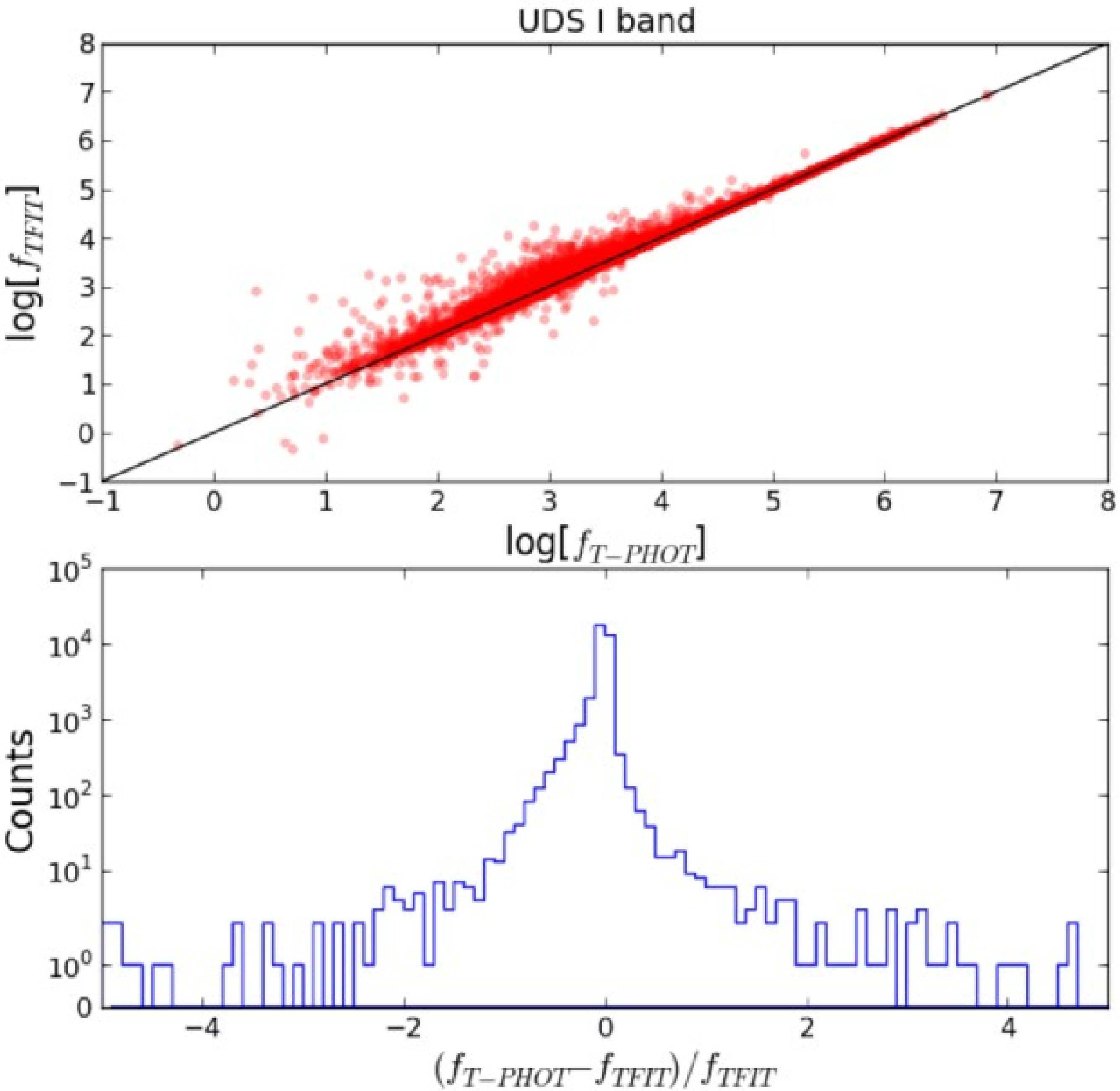
Problem in TFIT-like cell fitting approach



Possible solutions:
- fit on the whole image at once (CONVPHOT approach; drawbacks: memory and computing time limitations)
- cells-on-object approach (as in McLure's 2011 code)



Regions from UDS I band residual images. Left: TFIT “official” catalog; right: T-PHOT with *cells-on-object* method and revised kernel registration

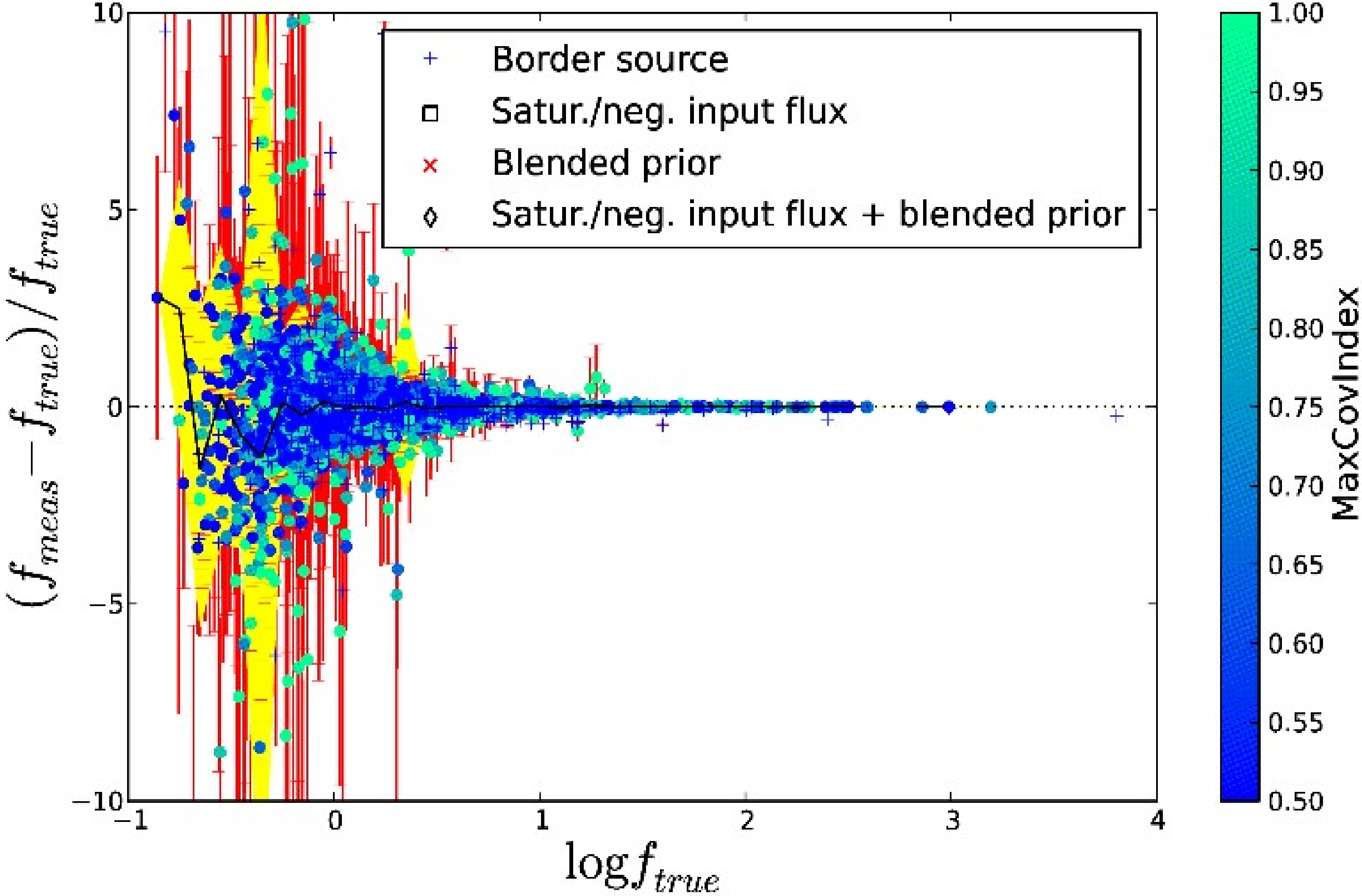


UDS I band
log[f_{TFIT}]
log[f_{T-PHOT}]
Counts
$(f_{T-PHOT} - f_{TFIT})/f_{TFIT}$

ASSUMPTIONS, CAVEATS, ANALYSIS

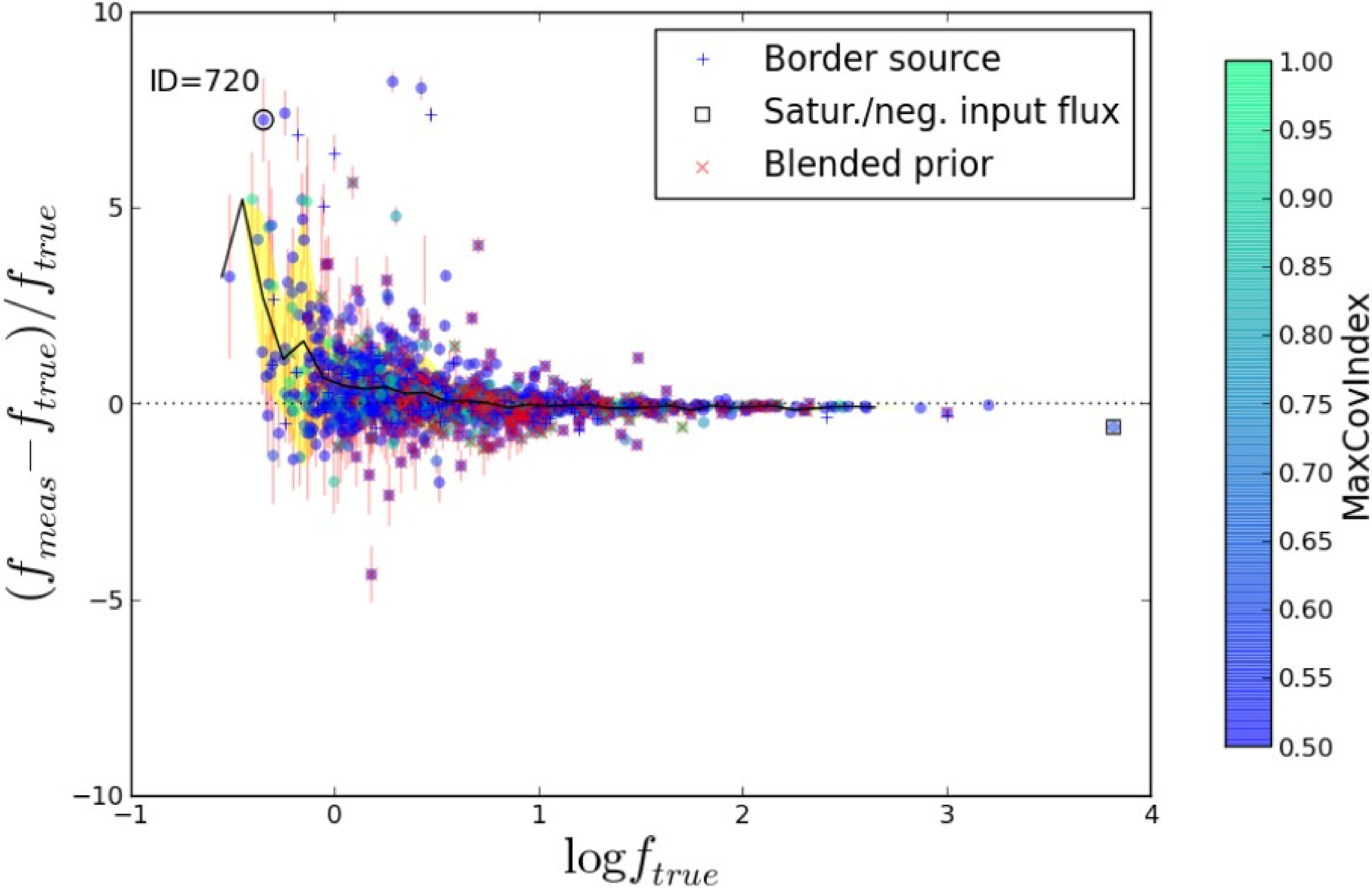
- Strong dependence on the accuracy of the PSF

- Prone to assumptions: no morphology dependence on wavelength (for real priors… how about multicomponent models?); no priors blending, etc.

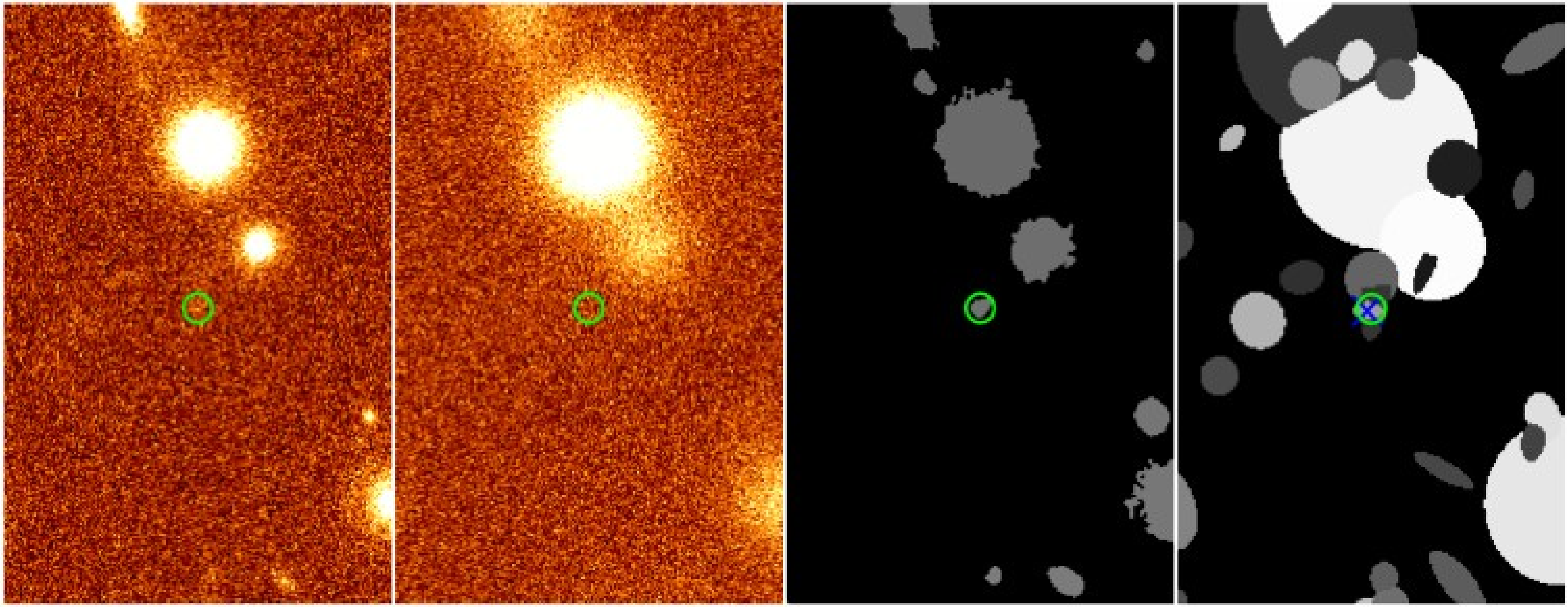


Point-sources simulation

ASSUMPTIONS, CAVEATS, ANALYSIS
- Strong dependence on the accuracy of the PSF
- Prone to assumptions: no morphology dependence on wavelength (for real priors... how about multicomponent models?); no priors blending, etc.

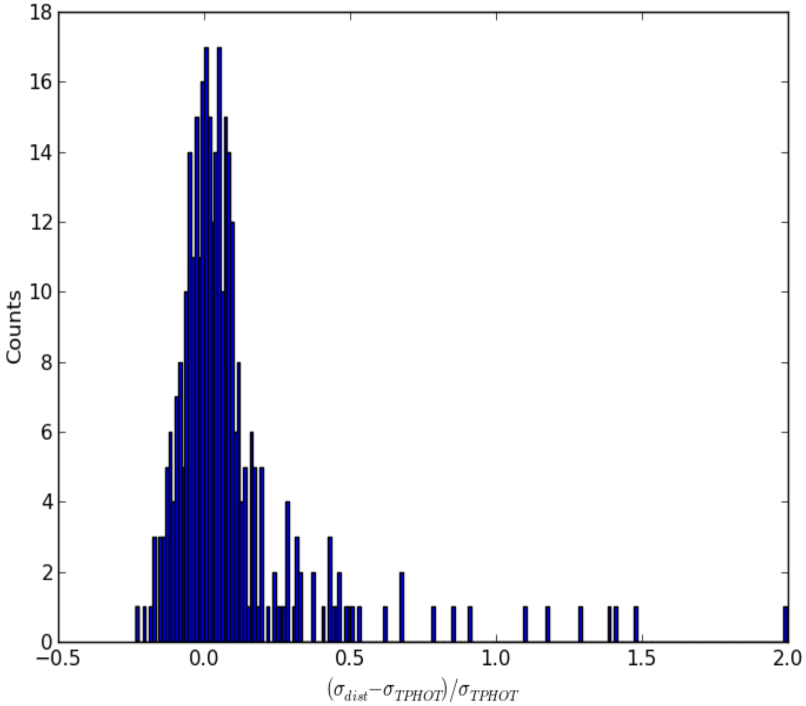


Extended sources simulation



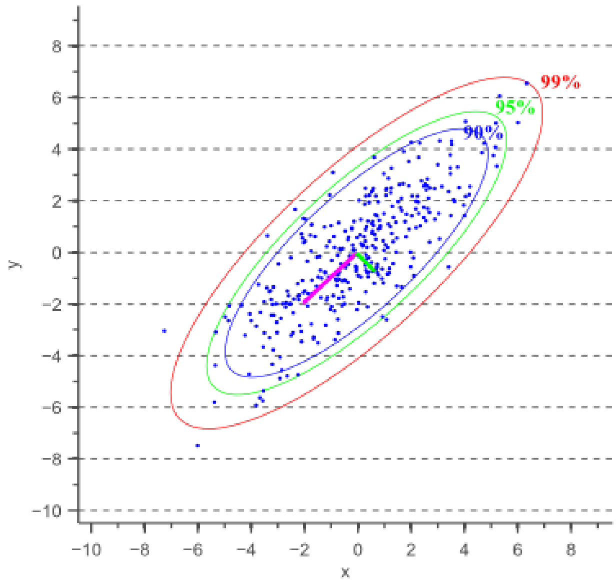
Bottom line:

The error budget computed via covariance matrix is a statistical error



Distribution of relative differences between nominal TPHOT errors and variance of the distributions of measurements in 100 realizations of the same field, for each object

$$C_{ij}=A_{ij}^{-1}=Adj(A)_{ji}/det[A]$$



*... but there are many possible causes of non-statistical, **systematical errors** (which would correspond to a <u>shift of the center of the error ellipse</u>)*

SUMMARY:

- T-PHOT is fast, robust, versatile and accurate :)
- Works fine on FIR to UV datasets, uses three types of priors
- It is promising as the weapon of choice for future (large, demanding) surveys (… Euclid?)

- Ongoing work at OAR using T-PHOT:
  * Goods-S K selected K+IRAC catalog
  * Frontier Fields IRAC catalogs
  * Extended "final" simulation set from FIR to UV

- To do:
  *parallelize fitting routine for very large datasets?

HOW TO GET T-PHOT:

http://www.astrodeep.eu/t-phot/

emiliano.merlin@oa-roma.inaf.it