PSF-matched photometry with T-PHOT: state of the art, future improvements, conceptual issues

Emiliano Merlin

ASTRODEEP meeting

La Sapienza - Sept 8th-9th 2014

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

Outline

- What's new and present status
- Future improvements
- T-PHOT v1.6 public release

- Accuracy of the method
- What are we measuring?
- Error budget (to be continued)

- What's new and present status

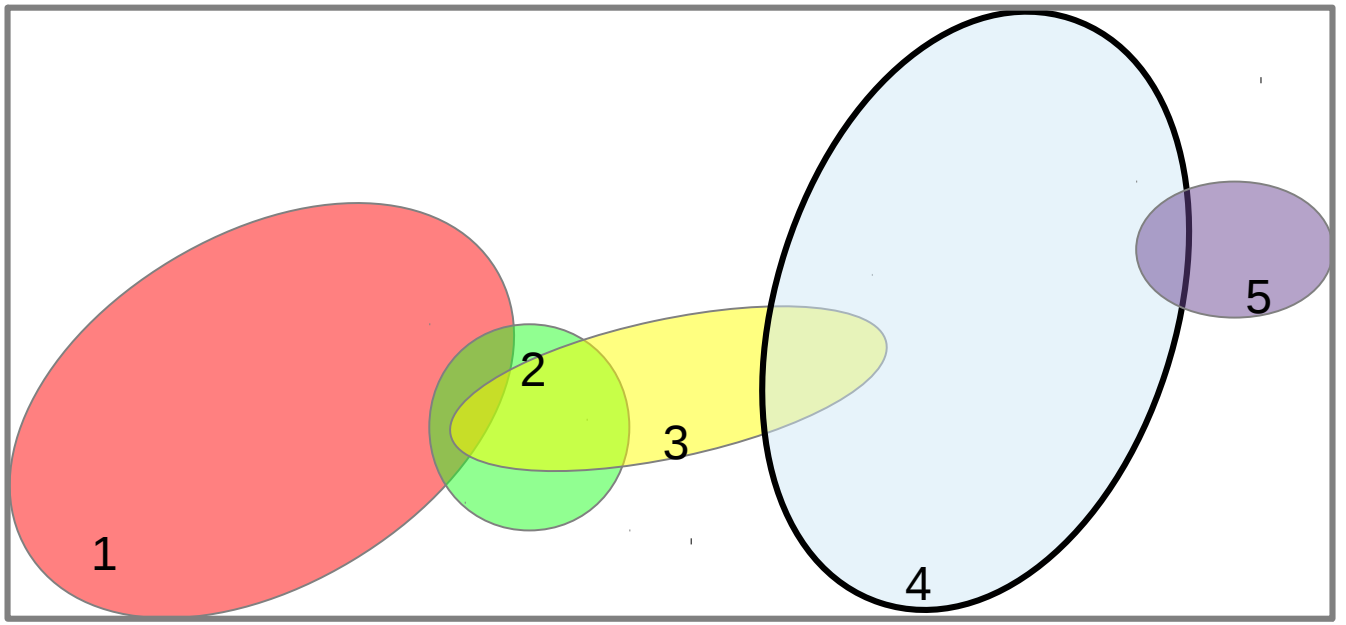
- T-PHOT has been tested by:
Marco, Stefano, Konstantina, Nathan, Tao, Xinwen, Koryo, Corentin; Andrea Grazian, Kuang-Han Huang, Henry Ferguson, Mimi Song, Alice Mortlock, Hooshang Nayyeri, Mauro Stefanon

- Recently (or now-being) added features:

* Cells-on-object method

* Mixed priors routines

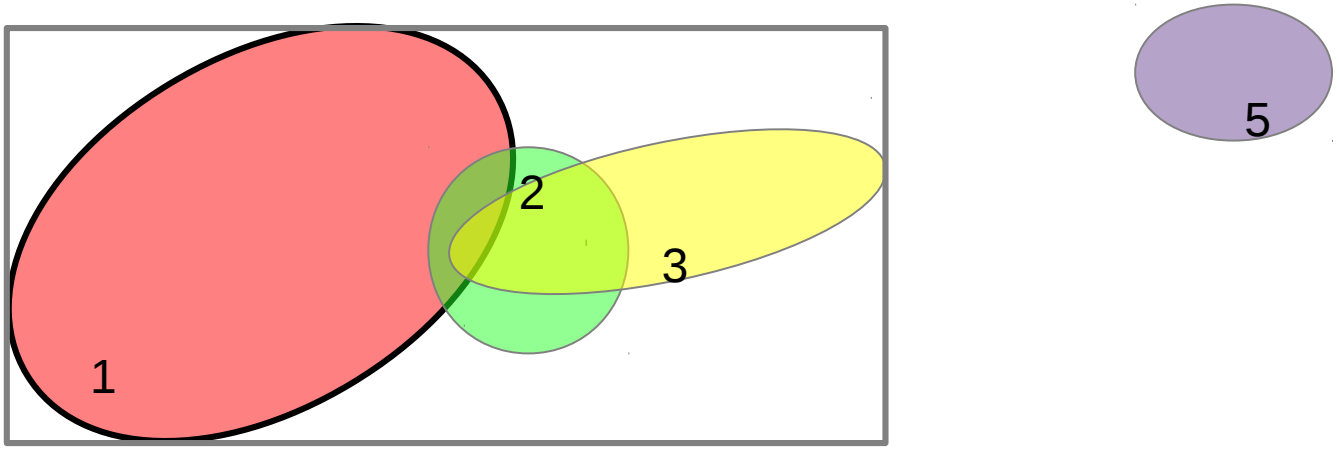
* Flags for “bad” sources (see discussion on errors)



TPHOT CELLS-ON-OBJECTS method:

* One cell per object, ordered by their HRI flux
* Include all first order contaminants
* Higher order contaminants are all included unless:
  - they are fainter than given fraction (0.9) of the total flux of the central object
  - 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

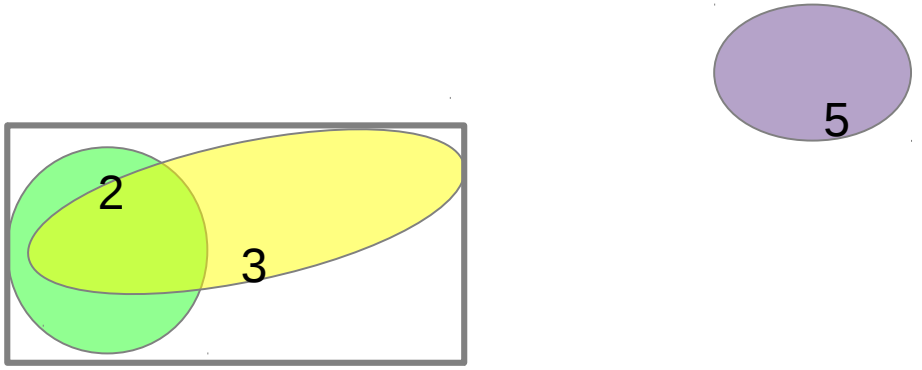
* Saving of CPU time depending on the level of blending in LRI



TPHOT CELLS-ON-OBJECTS method:

* One cell per object, ordered by their HRI flux
* Include all first order contaminants
* Higher order contaminants are all included unless:
  - they are fainter than given fraction (0.9) of the total flux of the central object
  - 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

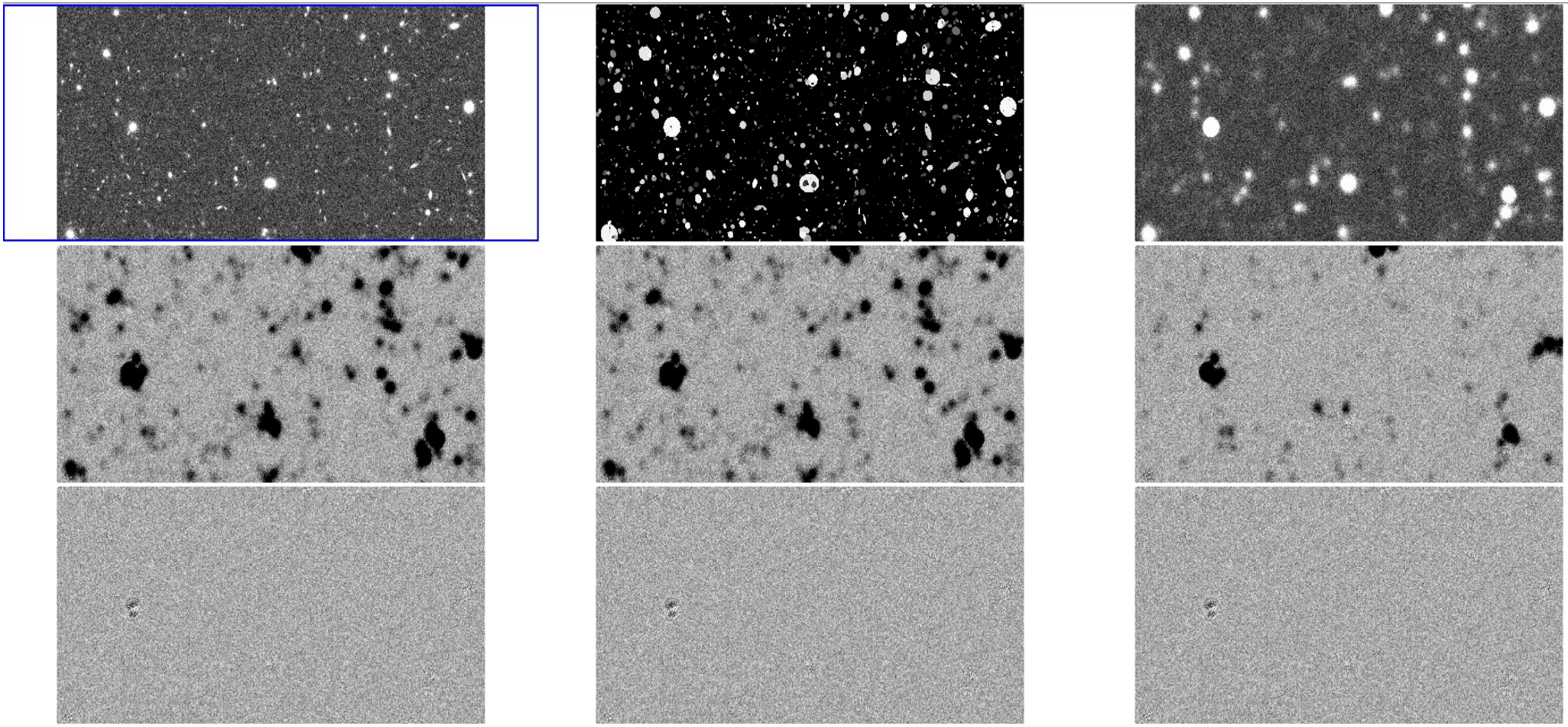
* Saving of CPU time depending on the level of blending in LRI



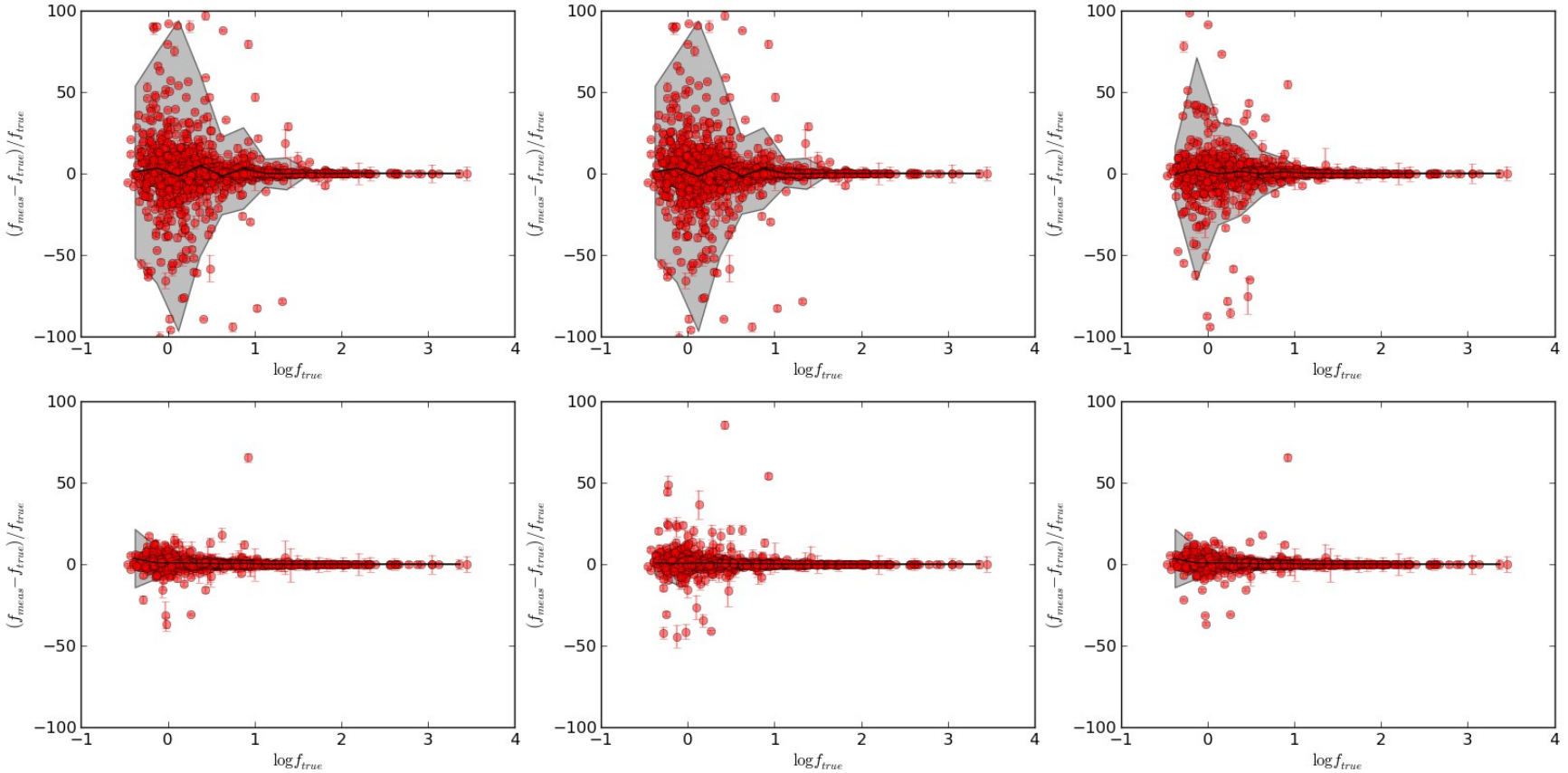
TPHOT CELLS-ON-OBJECTS method:

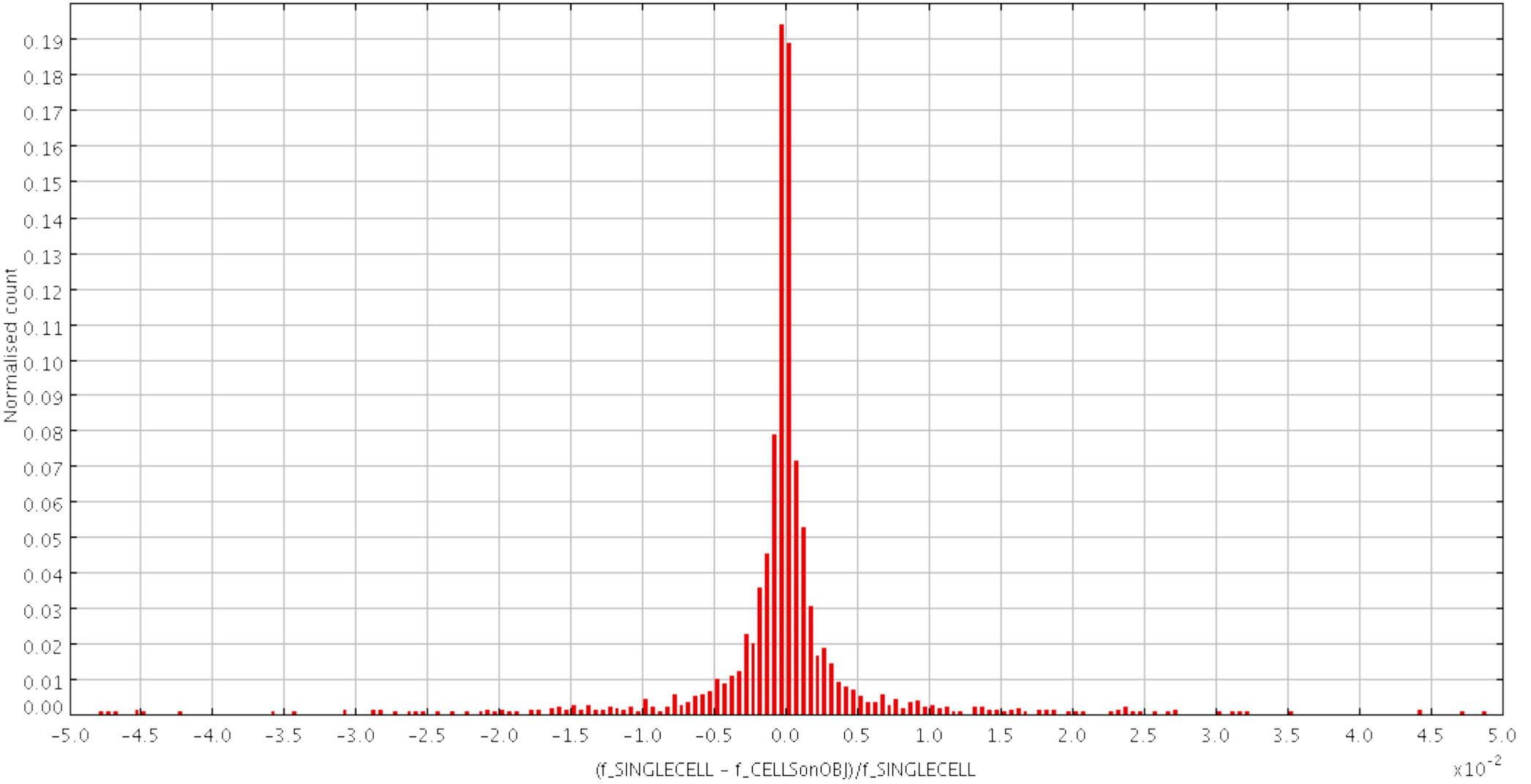
* One cell per object, ordered by their HRI flux
* Include all first order contaminants
* Higher order contaminants are all included unless:
  - they are fainter than given fraction (0.9) of the total flux of the central object
  - OR they only overlap to the previous level contaminant with a small fraction 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

* Saving of CPU time depending on the level of blending in LRI



Comparison between residual images for an IRAC simulation. Left to right, top to bottom: Detection simulated image (HST F160W); Segmentation image; Measure simulated image (IRAC ch1); Residual images (TFIT with 100x100 pixels cells; TPHOT with 100x100 pixels cells, identical to TFIT; TPHOT 200x200 pixels cells; CONVPHOT; TPHOT with single cell fitting; TPHOT with cells-on-objects fitting).


$(f_{meas} - f_{true})/f_{true}$
$\log f_{true}$
100
50
0
−50
−100
−1
0
1
2
3
4


Normalised count
0.19
0.18
0.17
0.16
0.15
0.14
0.13
0.12
0.11
0.10
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
-5.0
-4.5
-4.0
-3.5
-3.0
-2.5
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
(f_SINGLECELL – f_CELLSonOBJ)/f_SINGLECELL
x10-2

T-PHOT MIXED PRIORS:

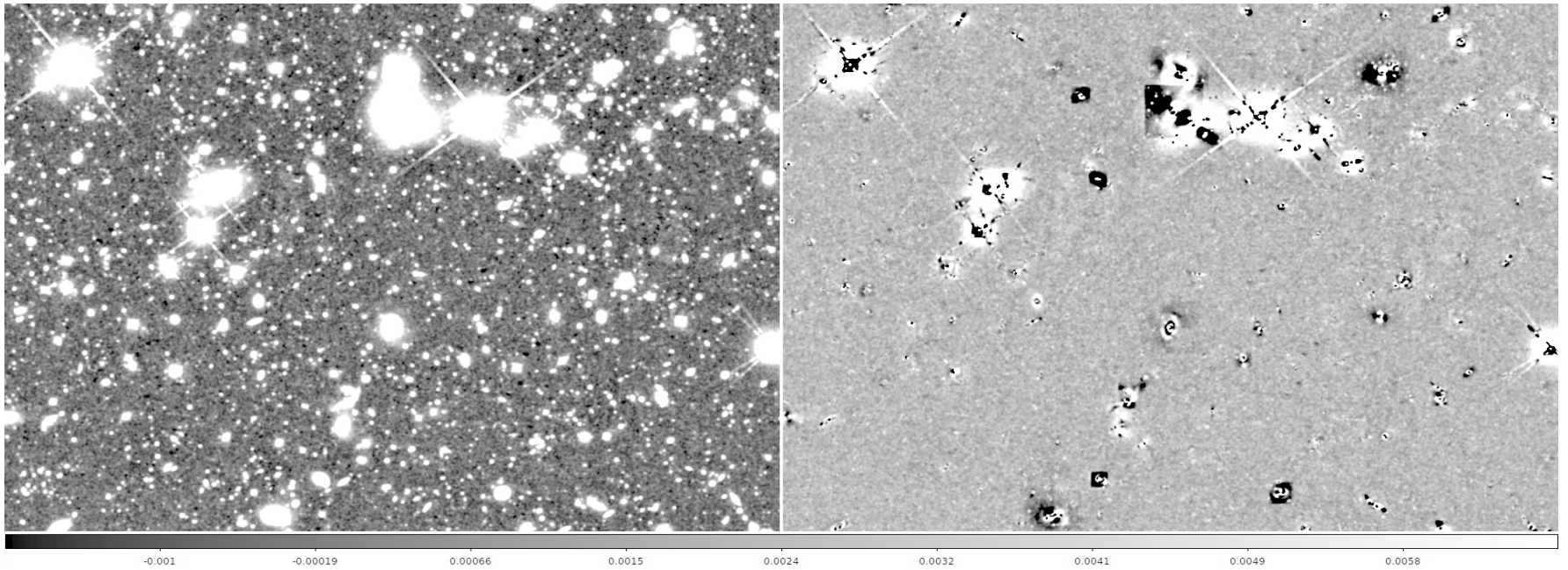
* **True cutouts from HRI** (standard old-style input). Needs: HRI, SEx segmentation map, SEx catalog

* **Galfit (multi-component) models** (or equivalent). Needs: models stamps (one per object per component), models catalog(s) (each component to be considered as a single object; positions of objects can be referred to LRI or to Galfit stamps)

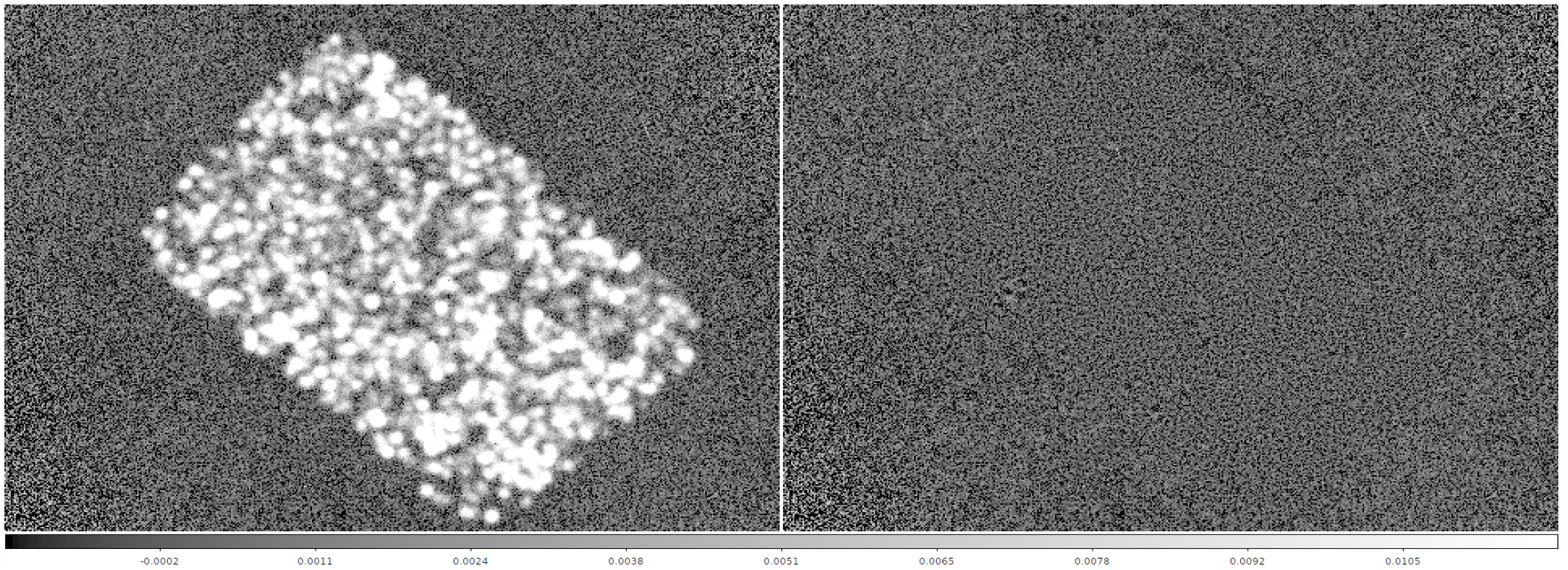
* List of **point-like unresolved sources**. Needs: *id x y* catalog in pixel reference frame

They can be used simultaneously (not ALL combinations tested, though!).

Useful e.g. when dealing with clusters (see Marco's slides about FF)



TPHOT using Galfit priors with two components – Test on COSMOS r band (with Fernando)



TPHOT using list of poit-like sources – Test on Herschel simulation (with Tao and Xinwen)

T-PHOT FUTURE IMPROVEMENTS:

Near future:

- Complete tests and checks on mixed priors
- Background fitting in LRI image (Nathan)
- PSF and kernel estimation and production
- Introduction of flags for bad sources (see discussion on errors)
- Thoroughly check effect of threshold fitting
- Quantify effects of PSF errors

Far future:

- Parallel computing / GPU?
- Other suggestions?

- T-PHOT v1.6 public release

- T-PHOT v1.6 to be uploaded and made publicly available from ASTRODEEP site

- To be presented at ADASS XXIV conference (Calgary, Oct 4-9)

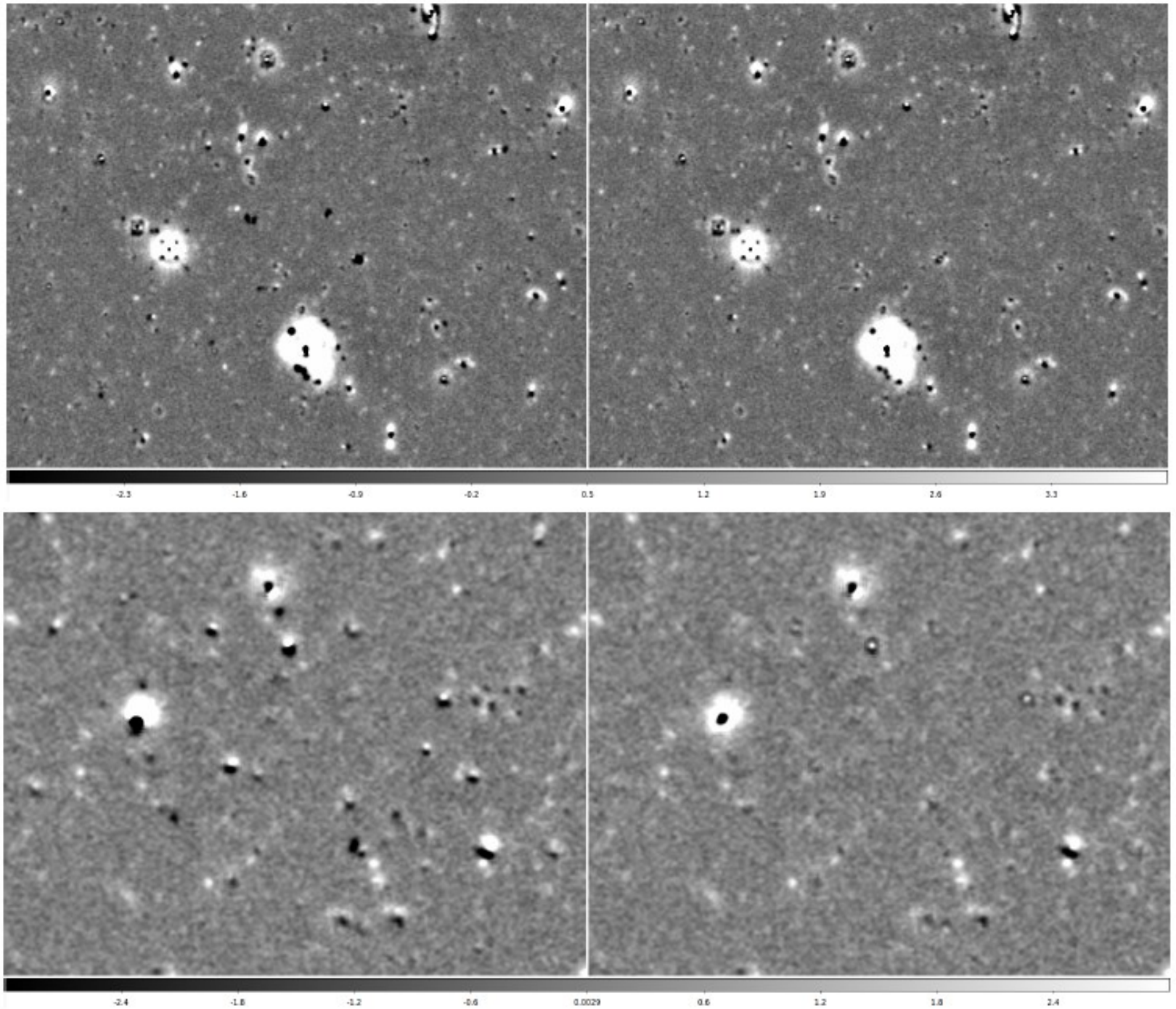
- Paper (still) in preparation. Guidelines:

1) General presentation of the code

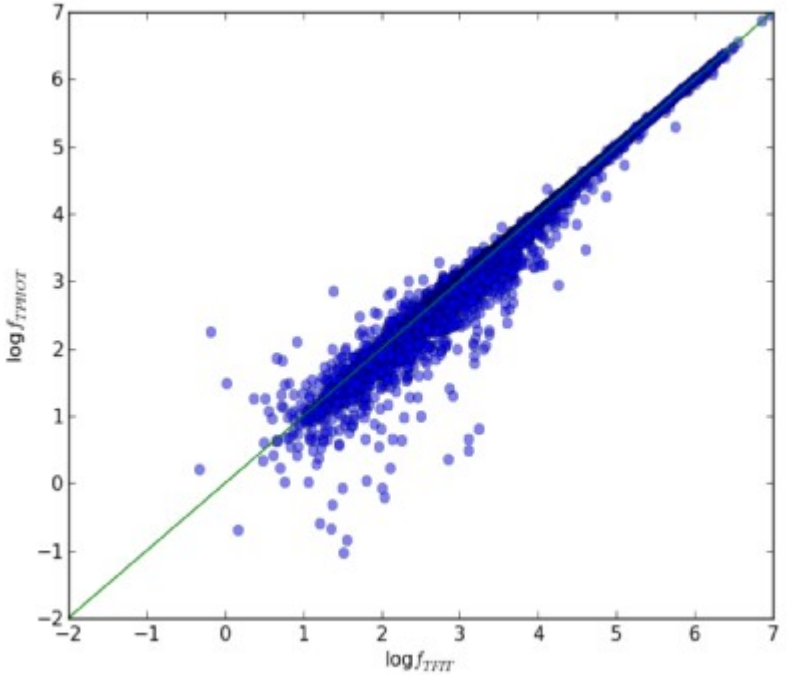
2) Advantages wrt similar codes: speed (also using cells-on-objects method; e.g., Konstantina Goods-S K selected catalog), robustness, versatility

3) Compared accuracy: wrt SExtractor (on simulations), TFIT (on simulations and on real data: UDS CANDELS catalog), Fastphot?

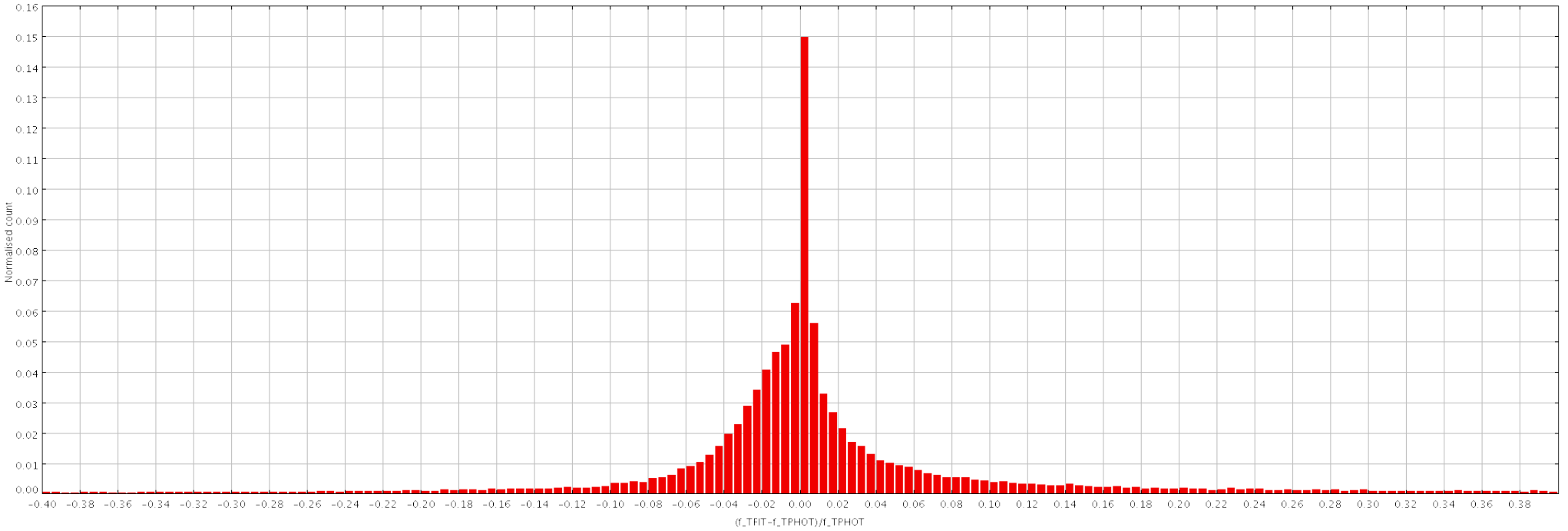
4) Overall reliability of the method: accuracy, error estimates



UDS Subaru I band – Left: official catalog residual; right: TPHOT cells-on-obj + dance residual



UDS B band: TFIT official catalog vs. TPHOT cells-on-obj + dance

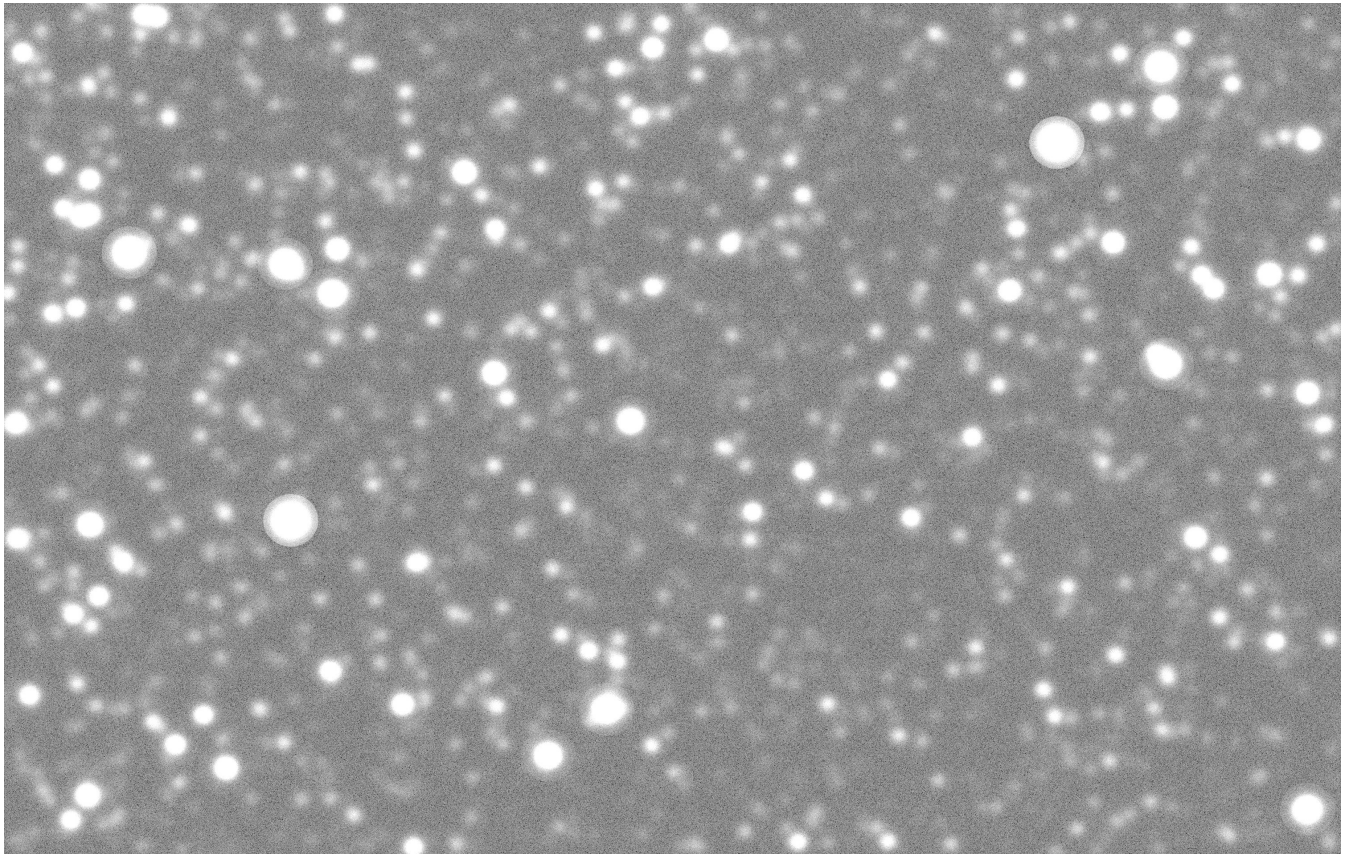


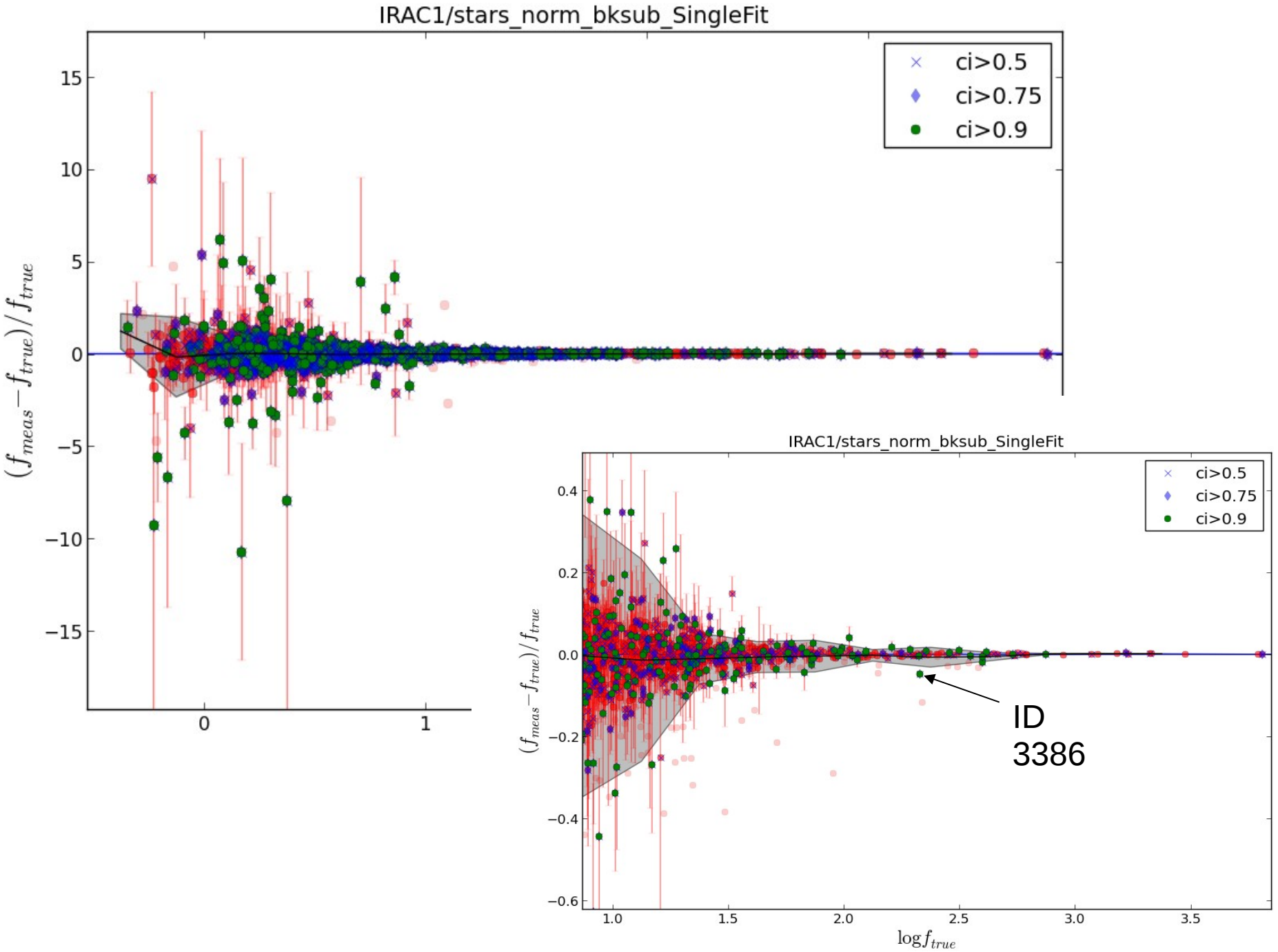
Accuracy

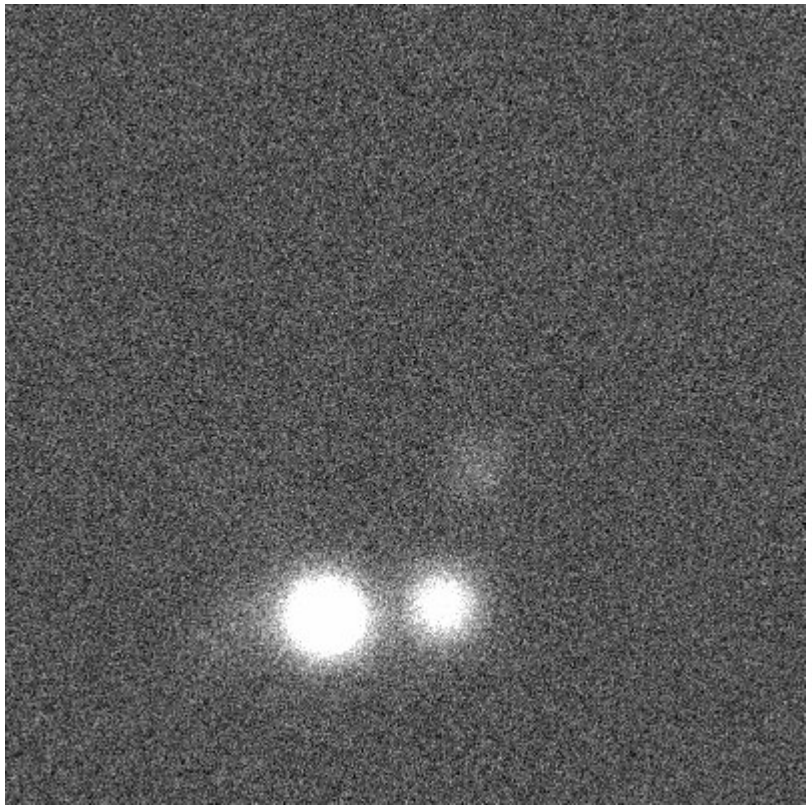
- Started with simplest possible cases:

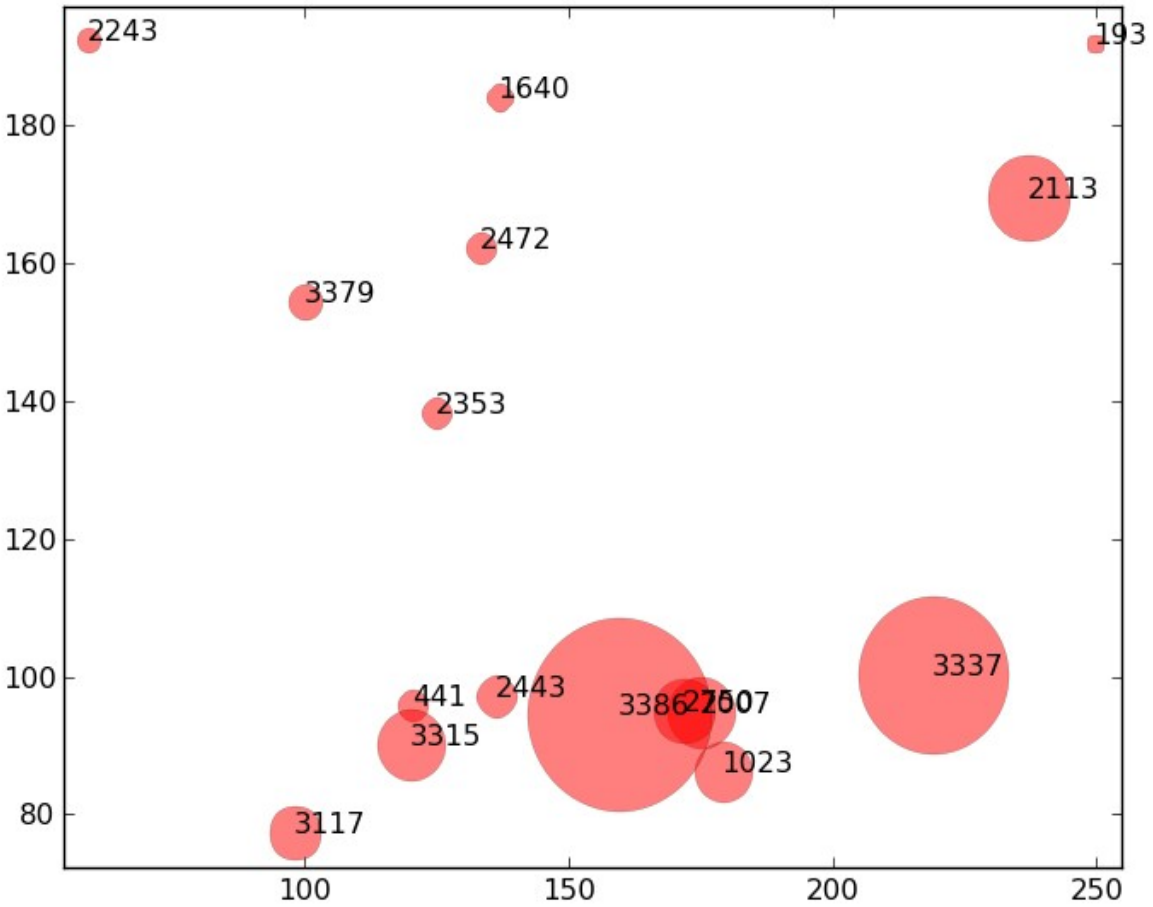
* Two point-like sources at different distances and different S/N (Koryo)

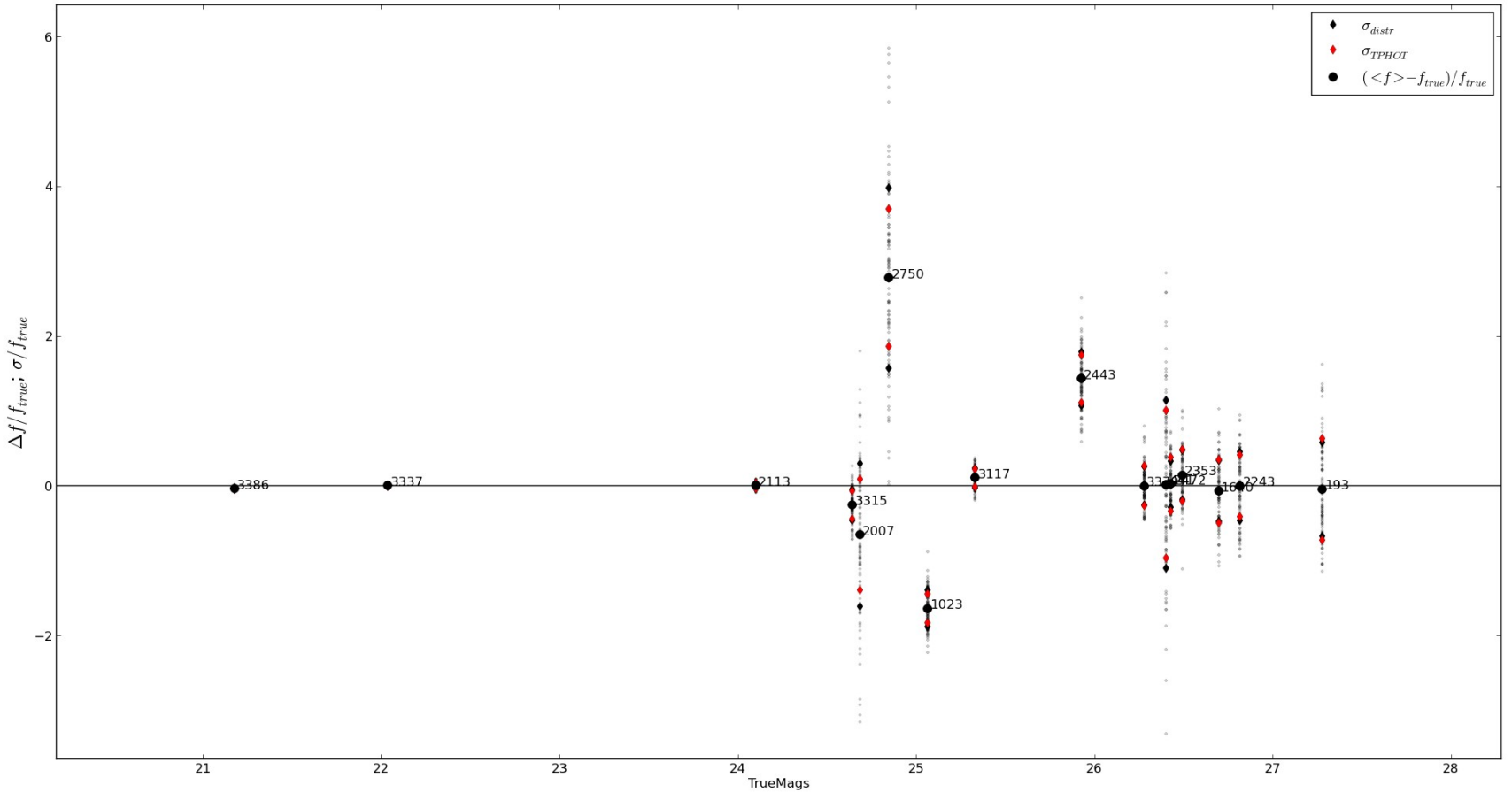
* “Realistic” field of point-like sources (SkyMaker simulation):
saturation, border effects, blending

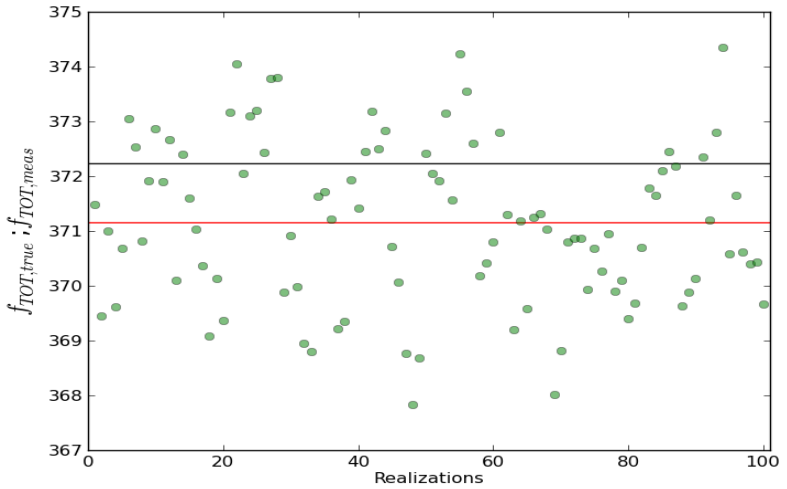


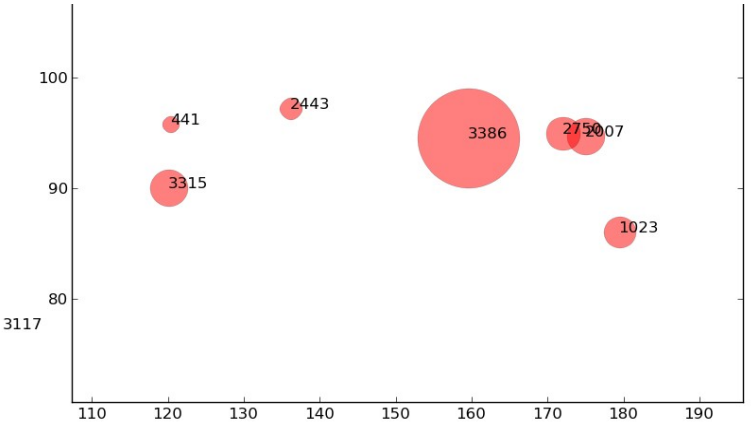

IRAC1/stars_norm_bksub_SingleFit
ci>0.5
ci>0.75
ci>0.9
$(f_{meas} - f_{true})/f_{true}$
15
10
5
0
−5
−10
−15
0
1
IRAC1/stars_norm_bksub_SingleFit
ci>0.5
ci>0.75
ci>0.9
$(f_{meas} - f_{true})/f_{true}$
0.4
0.2
0.0
−0.2
−0.4
−0.6
1.0
1.5
2.0
2.5
3.0
3.5
$\log f_{true}$
ID
3386

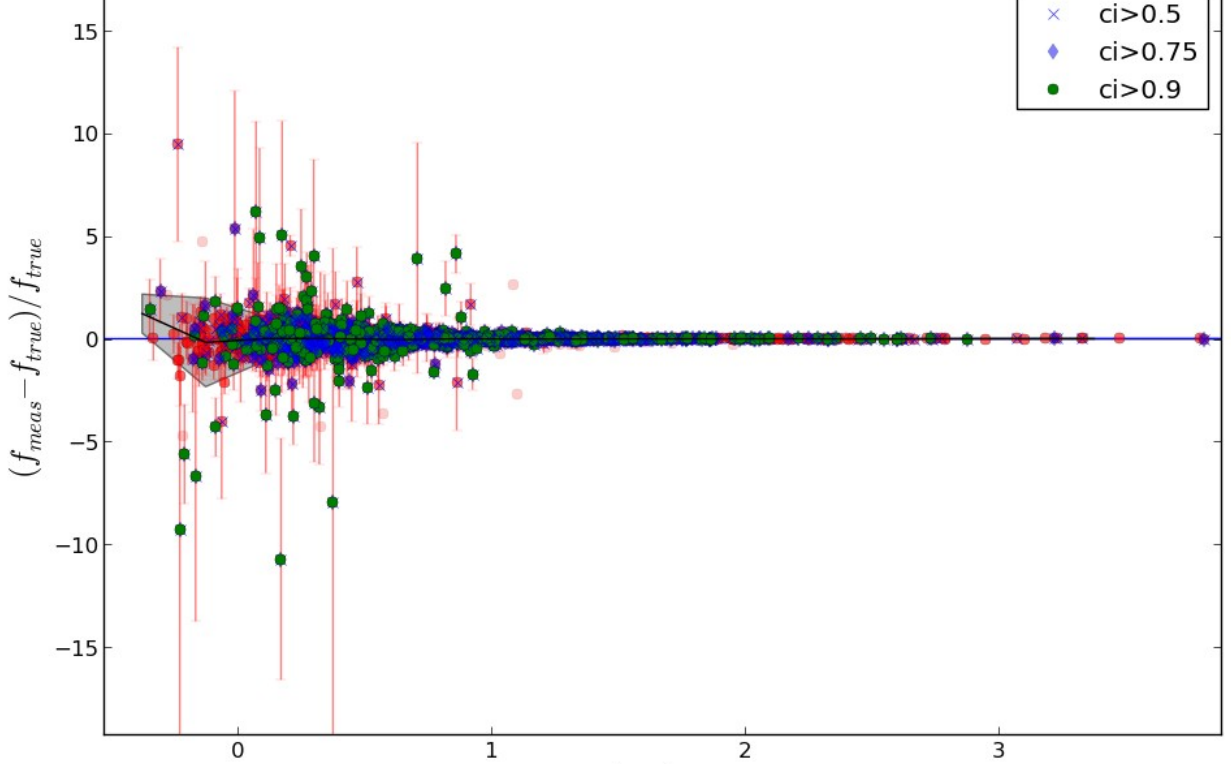


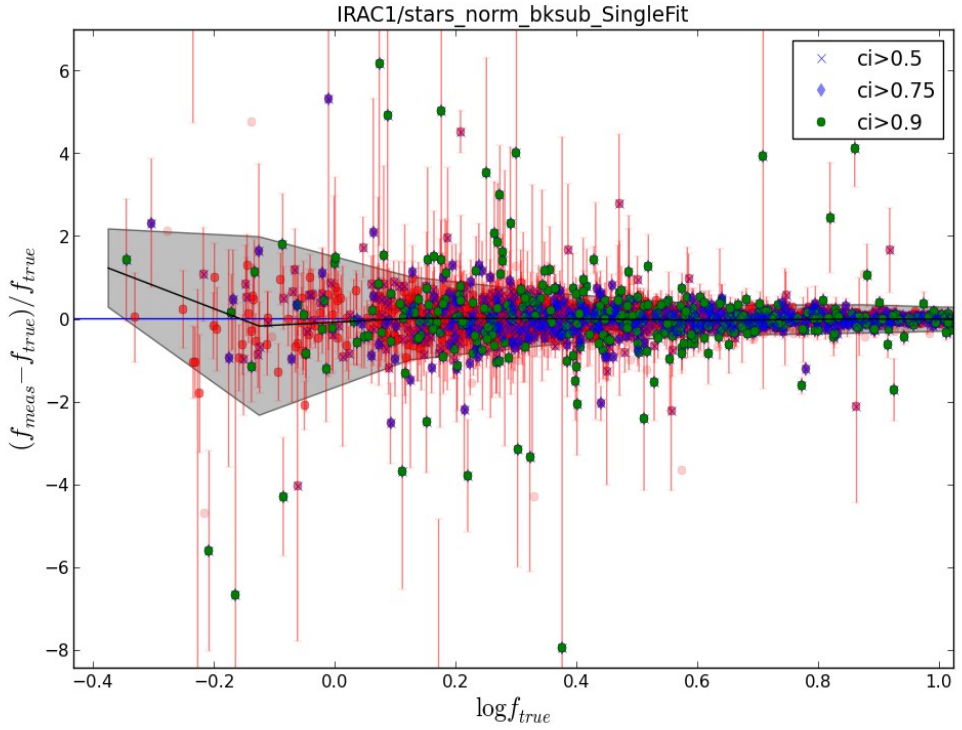

2243
193
1640
180
2113
2472
160
3379
140
2353
120
3337
100
441
2443
3315
1023
80
3117
100
150
200
250

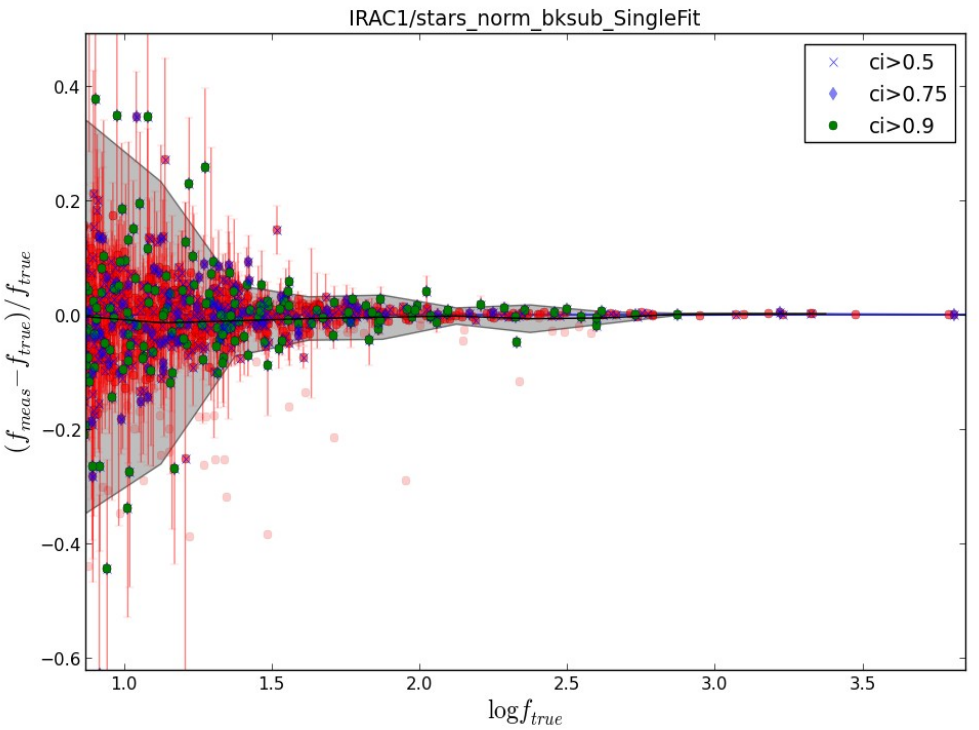

σ_{distr}
σ_{TPHOT}
$(<f>-f_{true})/f_{true}$
$\Delta f/f_{true}$; σ/f_{true}
TrueMags
3386
3337
2113
3315
2007
2750
1023
3117
2443
2353
2243
193


$f_{TOT;true}$; $f_{TOT;meas}$
Realizations


441
2443
3386
2750
2007
3315
1023
3117


ci>0.5
ci>0.75
ci>0.9
$(f_{meas} - f_{true})/f_{true}$


IRAC1/stars_norm_bksub_SingleFit
ci>0.5
ci>0.75
ci>0.9
$(f_{meas} - f_{true})/f_{true}$
$\log f_{true}$


IRAC1/stars_norm_bksub_SingleFit
ci>0.5
ci>0.75
ci>0.9
$(f_{meas} - f_{true})/f_{true}$
$\log f_{true}$

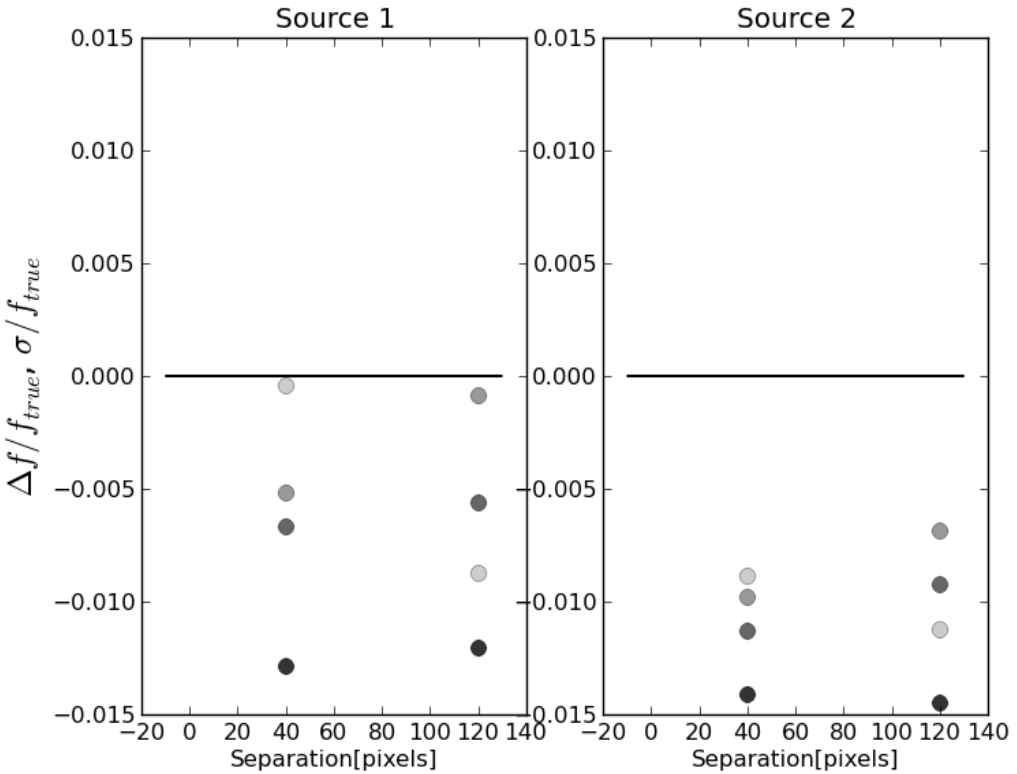
Accuracy

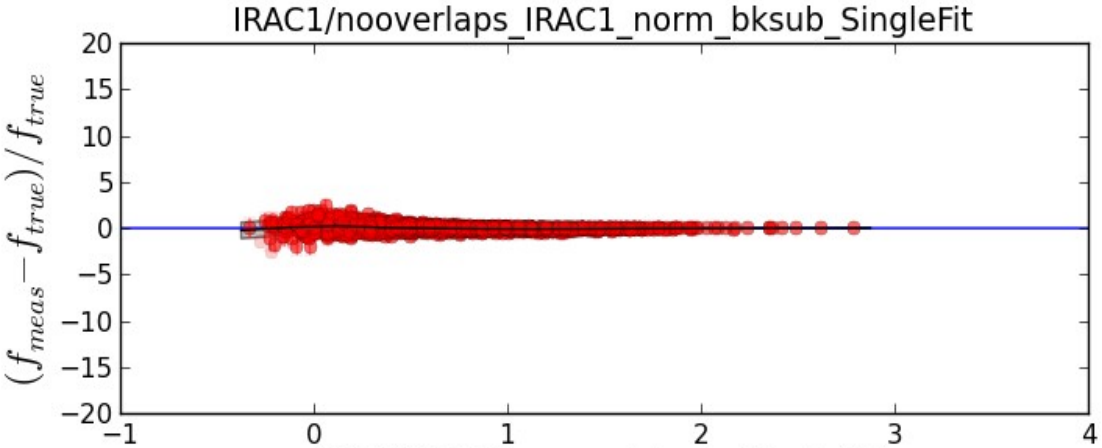
- Started with simplest possible cases:

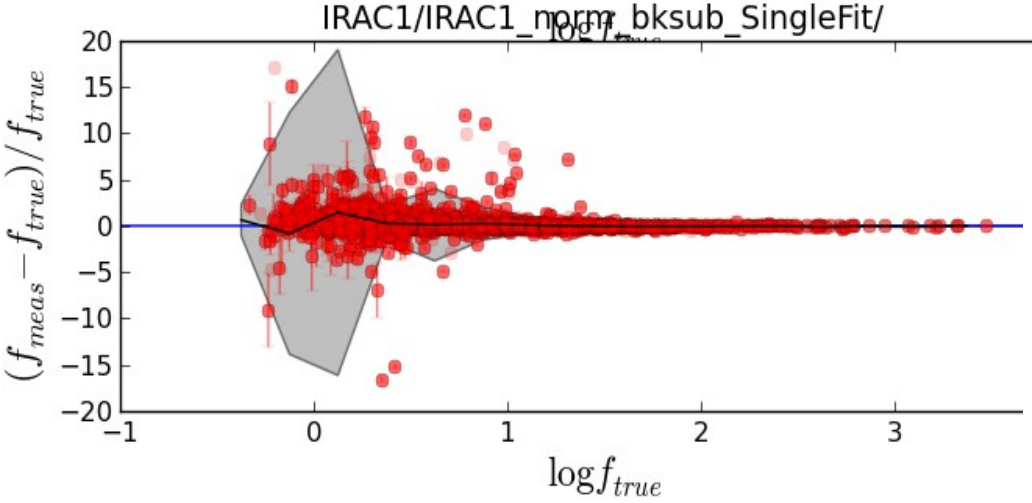
* Two point-like sources at different distances and different S/N (Koryo)

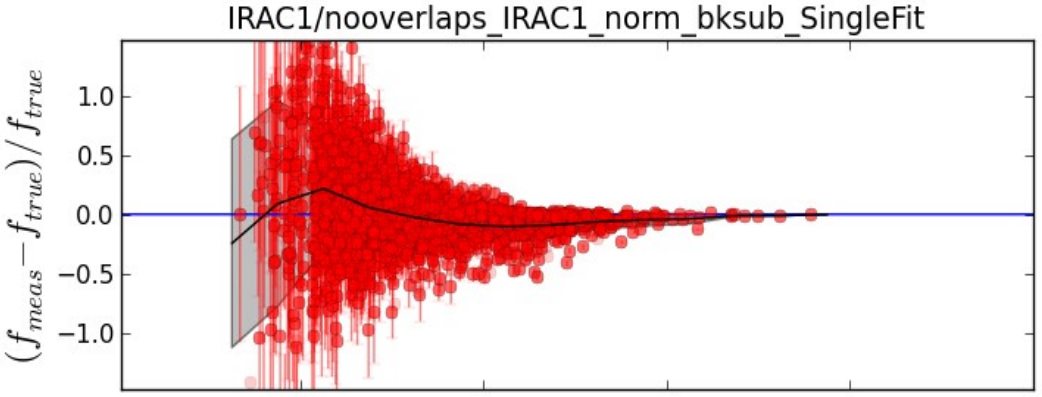
* “Realistic” field of point-like sources (SkyMaker simulation):
saturation, border effects, blending

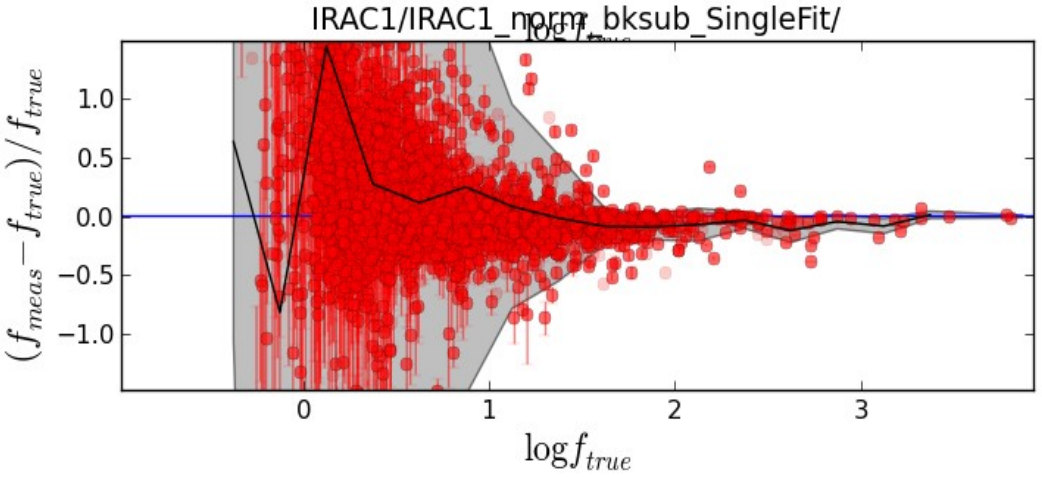
- Extended sources: segmentation issues




IRAC1/nooverlaps_IRAC1_norm_bksub_SingleFit
$(f_{meas} - f_{true})/f_{true}$


IRAC1/IRAC1_norm_bksub_SingleFit/
$(f_{meas} - f_{true})/f_{true}$
$\log f_{true}$


IRAC1/nooverlaps_IRAC1_norm_bksub_SingleFit
$(f_{meas} - f_{true})/f_{true}$


IRAC1/IRAC1_norm_bksub_SingleFit/
$(f_{meas} - f_{true})/f_{true}$
$\log f_{true}$

Causes of large, systematical offsets from true fluxes:

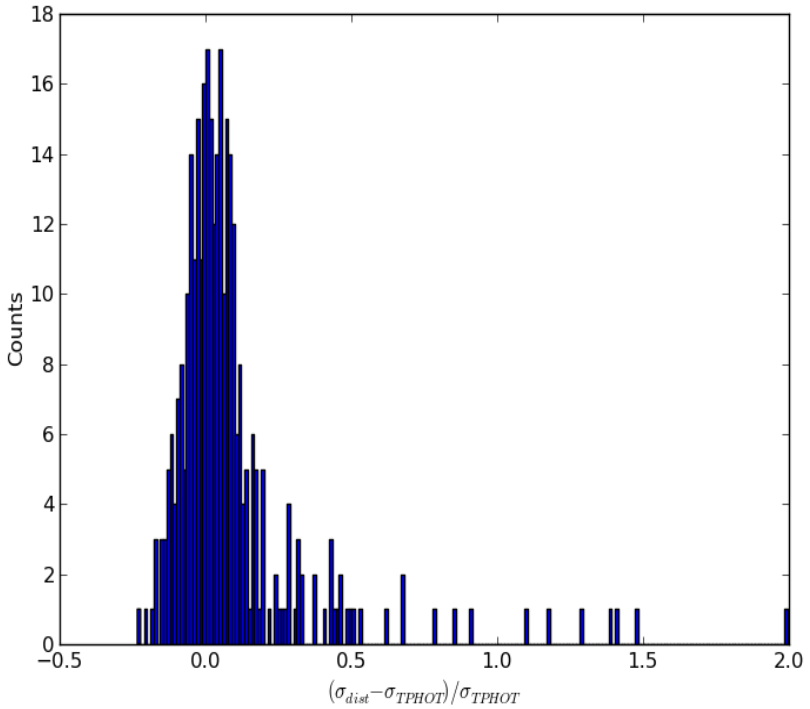
* Saturation of priors or of measure image (can be flagged)
* Cut, boundary object (can be flagged)
* Errors on prior position (... partly cured by PSF registration? Can be quantified?)

* Priors blending / confusion (can be flagged?)
* Wrong segmentation (too small, etc.). Flux is underestimated. Note that a few percent of the flux will quite always be lost (e.g. long tails of Sersic profiles...). Possible solutions: dilation, use models?

* High covariance in measure (dramatic measure blending). In this case, the covariance may give useful insights: *objects with large errors on their fluxes are likely to have at least one off-diag covariance matrix element of the same order of the variance*. Unfortunately, this is also true for many objects with little errors.

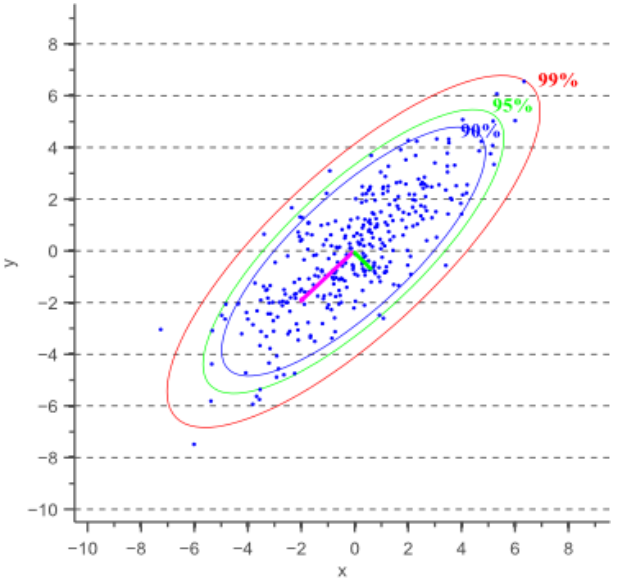
Summarizing:

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

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



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

Summarizing:

Causes of large, systematical offsets from true fluxes:

* Saturation of priors or of measure image (can be flagged)
* Cut, boundary object (can be flagged)
* Errors on prior position (partly cured by PSF registration?)

* Detection blending (can be flagged)
* Wrong segmentation (too small, etc.). Flux is underestimated. Note that a few percent of the flux will quite always be lost (e.g. long tails of Sersic profiles...). Possible solutions: dilation, use models?

* High covariance in measure (dramatic measure blending). In this case, the covariance may give useful insights: *objects with large errors on their fluxes are likely to have at least one off-diag covariance matrix element of the same order of the variance*. Unfortunately, this is also true for many objects with little errors.

What are we measuring?

At a first glance, the output quantity *f* should be, at first approximation, an **isophotal flux**, since it corresponds to the flux included in the high resolution cutout or model or PSF, convolved to a different resolution.
This may not be the whole flux of the object; a dilation of the latter can cure part of the issue.

However, a more precise comparison should be given obtaining an isophotal color, relating *f* with the flux within the (dilated) segmentation, and adding this term to a more robust estimation of the magnitude in the HR band;

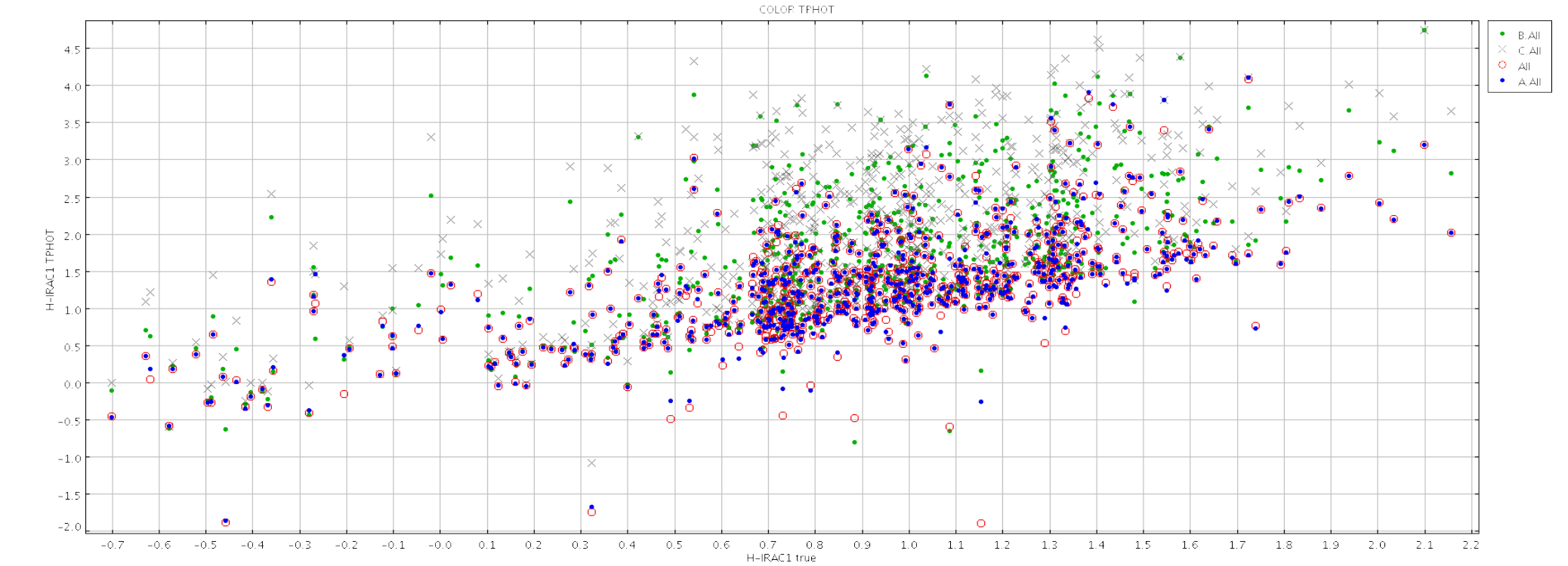
So a comparison with SExtractor based photometry should be given by:

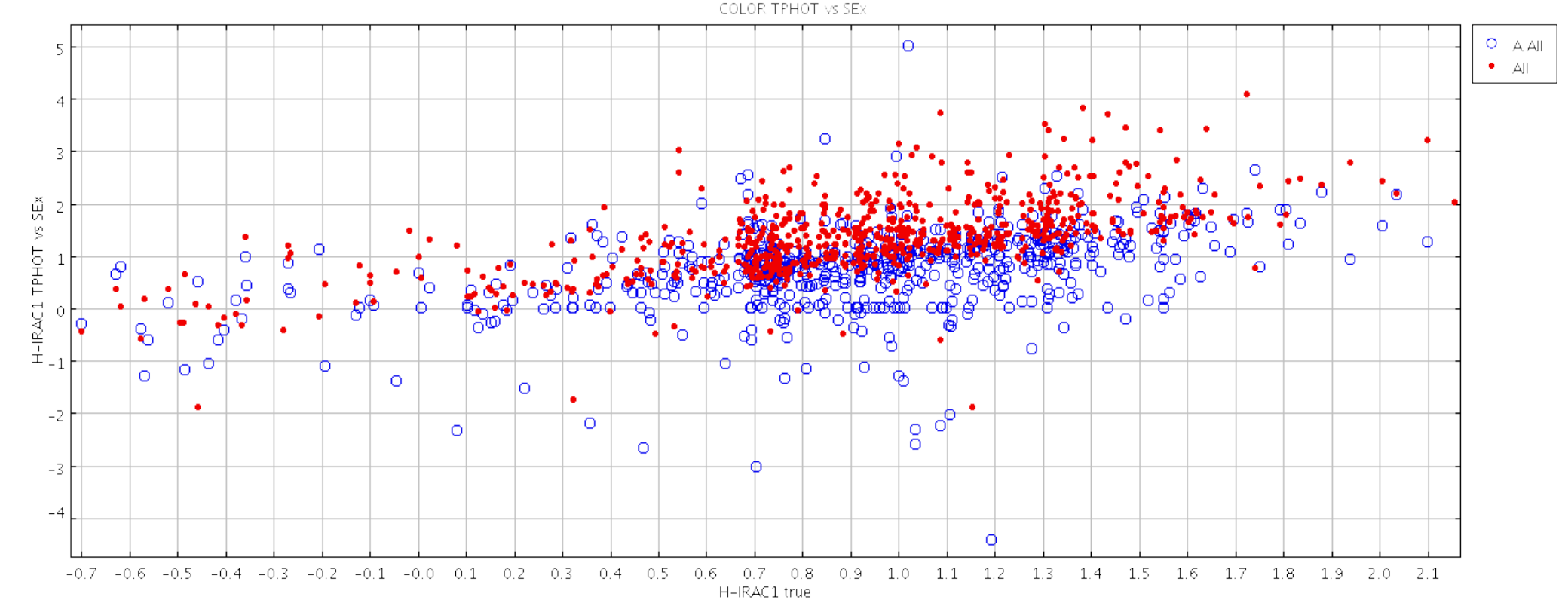
$$m_{best,LRI} = m_{best,HRI} + \left(m_{TPHOT,LRI} - m_{iso/dil,HRI} \right)$$

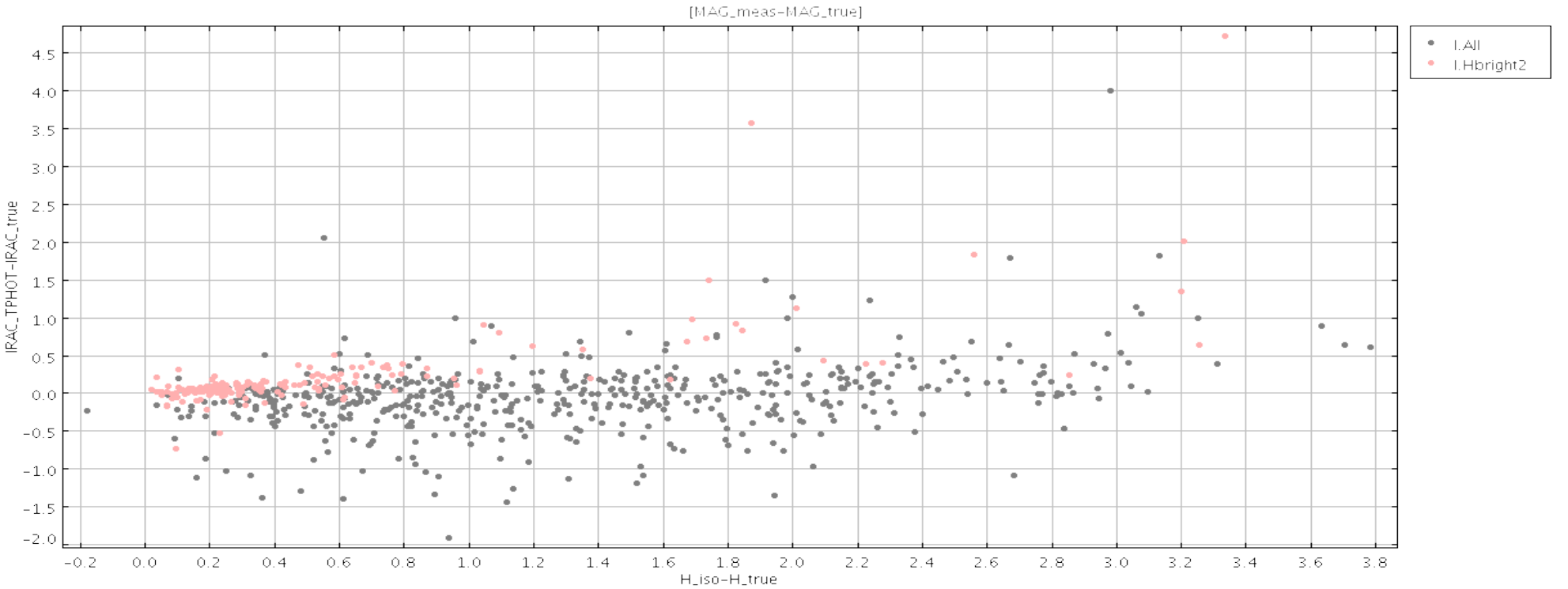
BUT...

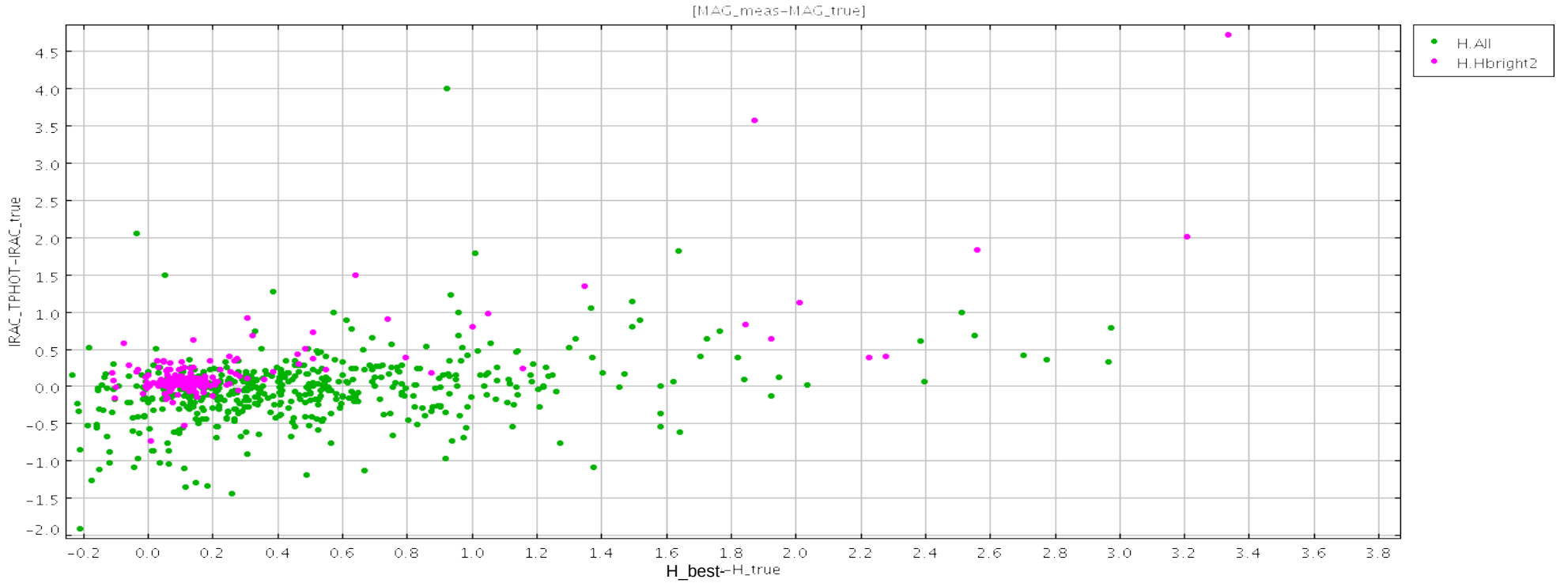
It seems that *f* actually gives a measure of the *total* flux, so that the color term should be given by

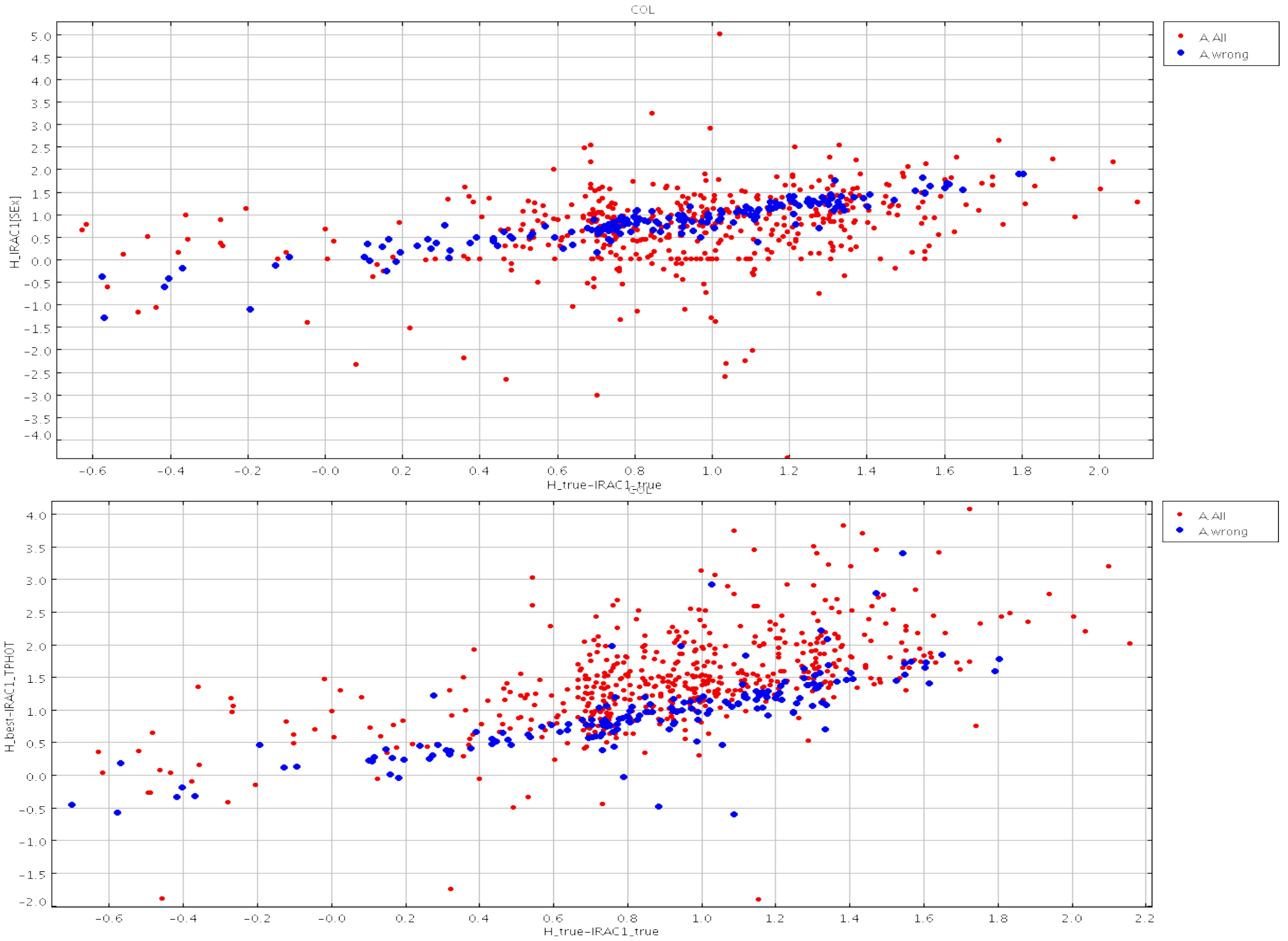
$$m_{TPHOT,LRI} - m_{best,HRI}$$


COLOR TPHOT
B_All
C_All
All
A_All
H-IRAC1 TPHOT
H-IRAC1 true


COLOR TPHOT vs SEx
A_All
All
H-IRAC1 TPHOT vs SEx
H-IRAC1 true


[MAG_meas-MAG_true]
I.All
I.Hbright2
IRAC_TPHOT-IRAC_true
H_iso-H_true


[MAG_meas-MAG_true]
H.All
H.Hbright2
IRAC_TPHOT-IRAC_true
H_best-H_true


COL
A_All
A_wrong
H_IRAC1[SEx]
H_true-IRAC1_true
COL
A_All
A_wrong
H_best-IRAC1_TPHOT
H_true-IRAC1_true