PSFEx:a generic tool for extracting the PSF in astronomical images

E.Bertin, IAP & Obs. de Paris/LERMA

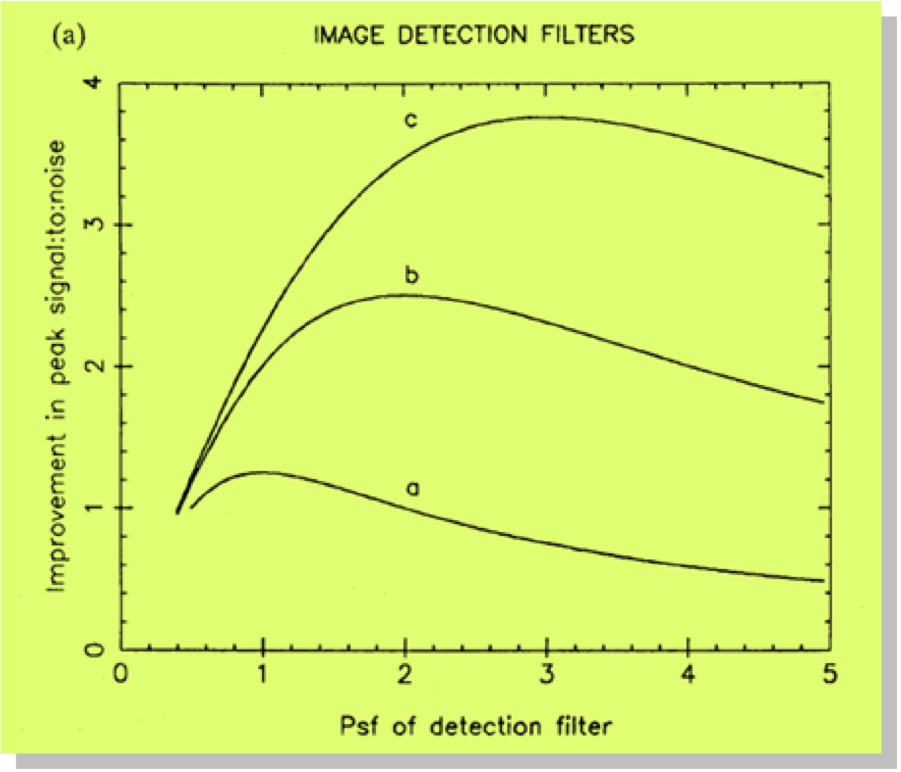
Outline

- Use of the PSF for detection, measurement and classification of astronomical sources
- Modeling the PSF with PSFEx
- Finding « prototype » stars
- Quality control at TERAPIX
- PSF-fitting with SExtractor

Detection and the PSF

- Matched (optimum) filtering for detection
  - Stationary noise with power spectrum $P(\boldsymbol{k})$ and isolated point-sources: convolve with

$$h = \phi^* * \mathrm{F}(P^{-1})$$



Irwin 1985

Source-deblending and profile-fitting in crowded star fields

- The PSF profile $\phi(\boldsymbol{x})$ can be quickly centered on isolated stars using a simple gradient descent
  - At each step, derive a profile offset $\Delta\boldsymbol{x}$ by fitting

$$F.\left(\phi + (\nabla\phi).\Delta\boldsymbol{x}\right)$$

  - Clumps of overlapping stars can be fitted using the same simple technique with additional constraints (no negative flux,minimum distance between stars)

Astrometry

- Effects of crowding
- The definition of a star position can be ambiguous for asymmetric PSFs
  - Flux-dependency when centroiding thresholded profiles

Point-source photometry

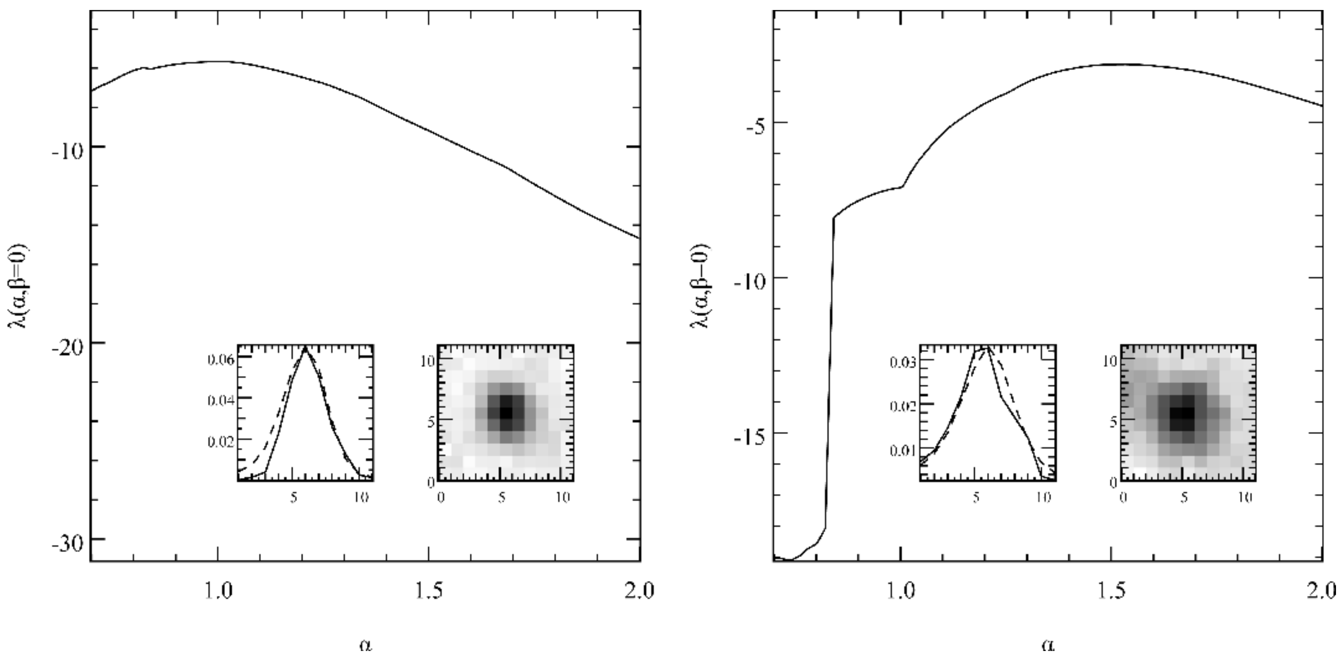
- Profile-fitting photometry always optimum in terms of SNR:

$$F \propto \sum_i \frac{\phi_i f_i}{\sigma_i^2}$$

- On photon-noise limited images with negligible background
  - $\sigma_i^2 \propto \phi_i$: profile-fitting equivalent to integration of pixel values within an aperture
- On photon-noise limited images with dominant background
  - $\sigma_i^2 \propto cste$: profile-fitting equivalent to a profile-weighted sum of pixel values

Star/galaxy separation

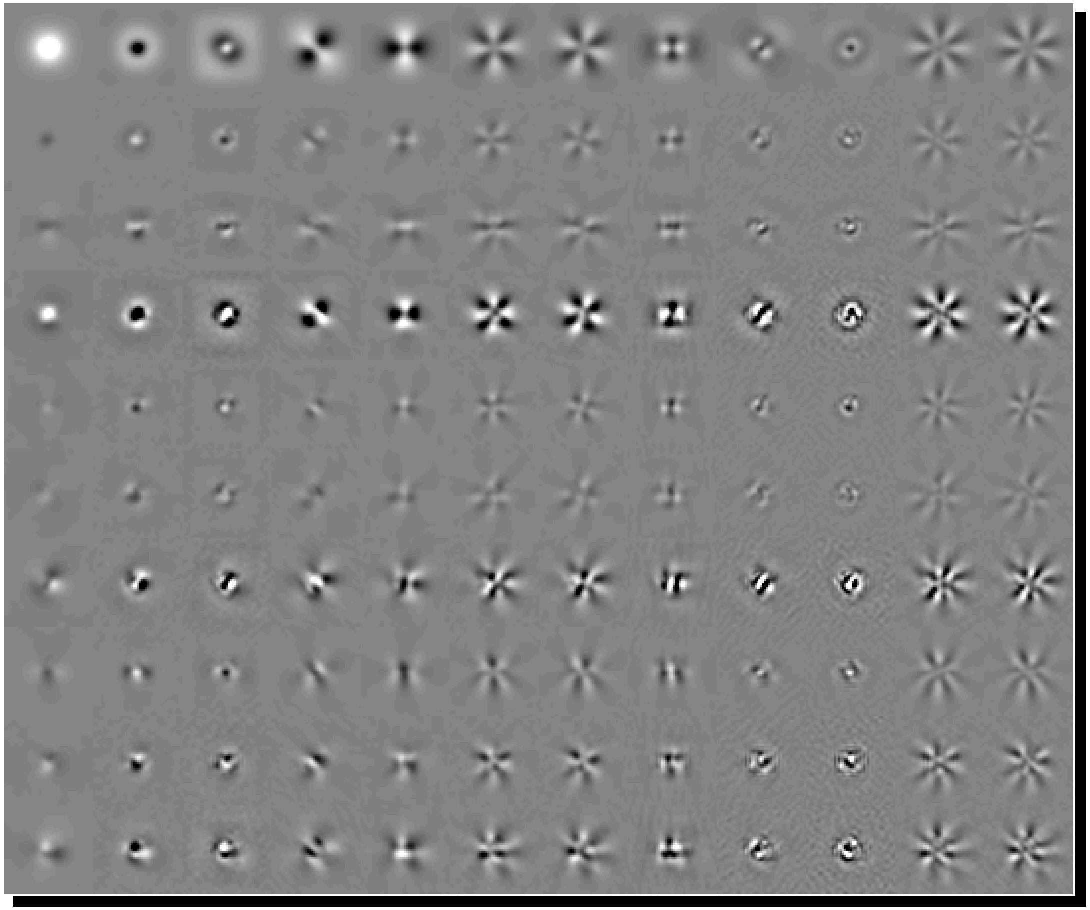
- Local PSF used as a reference for computing the likelihoods $p(\boldsymbol{y}|S)$ and $p(\boldsymbol{y}|G)$ of a star/galaxy Bayesian classifier (Sebok 1979, Valdes 1982 and followers)



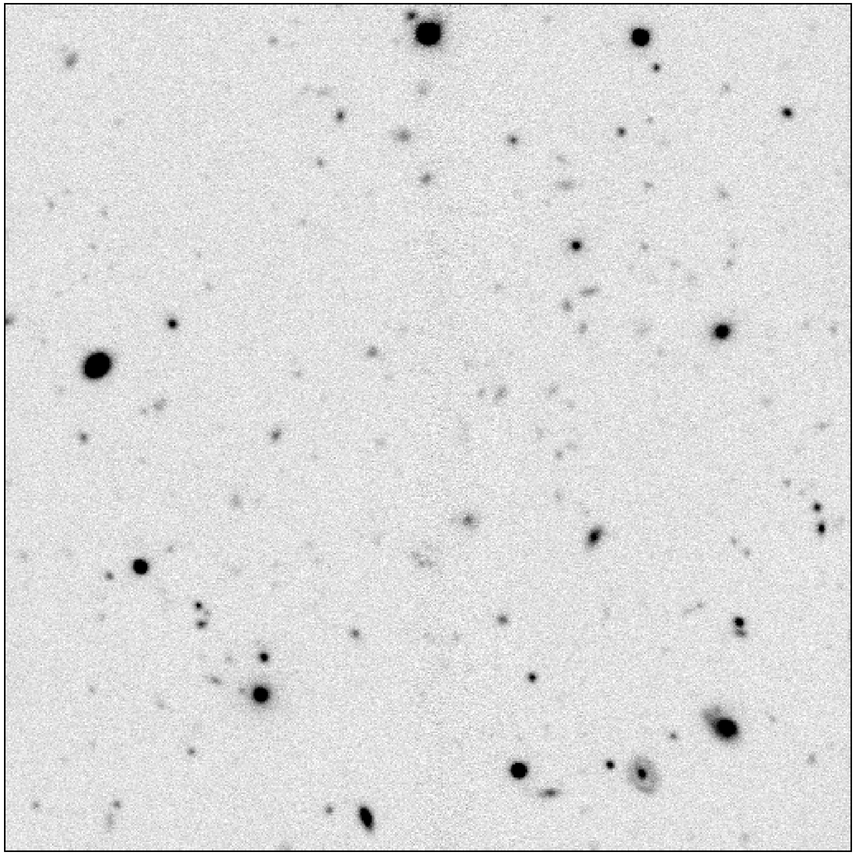
Drory 2002

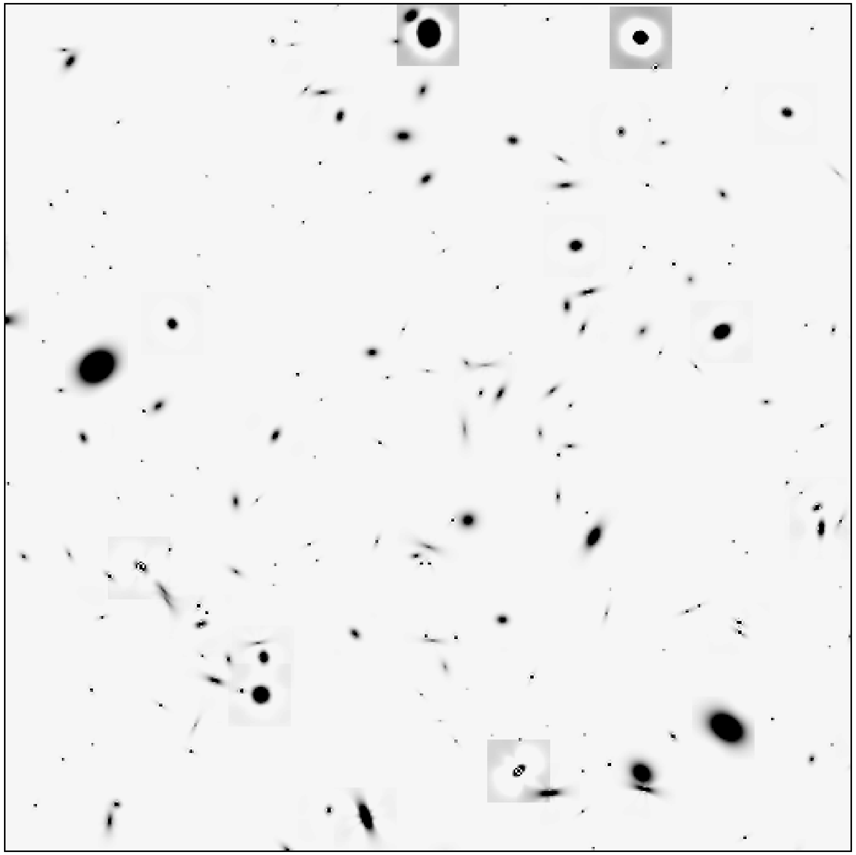
Morphology of extended sources

- Non-linear galaxy profile-fitting (e.g. GIM2D)
  - Reconvolution with the local PSF needed at each iteration
- Decomposition on basis functions (PCA, shapelets)
  - Basis functions are convolved with the local PSF before fitting



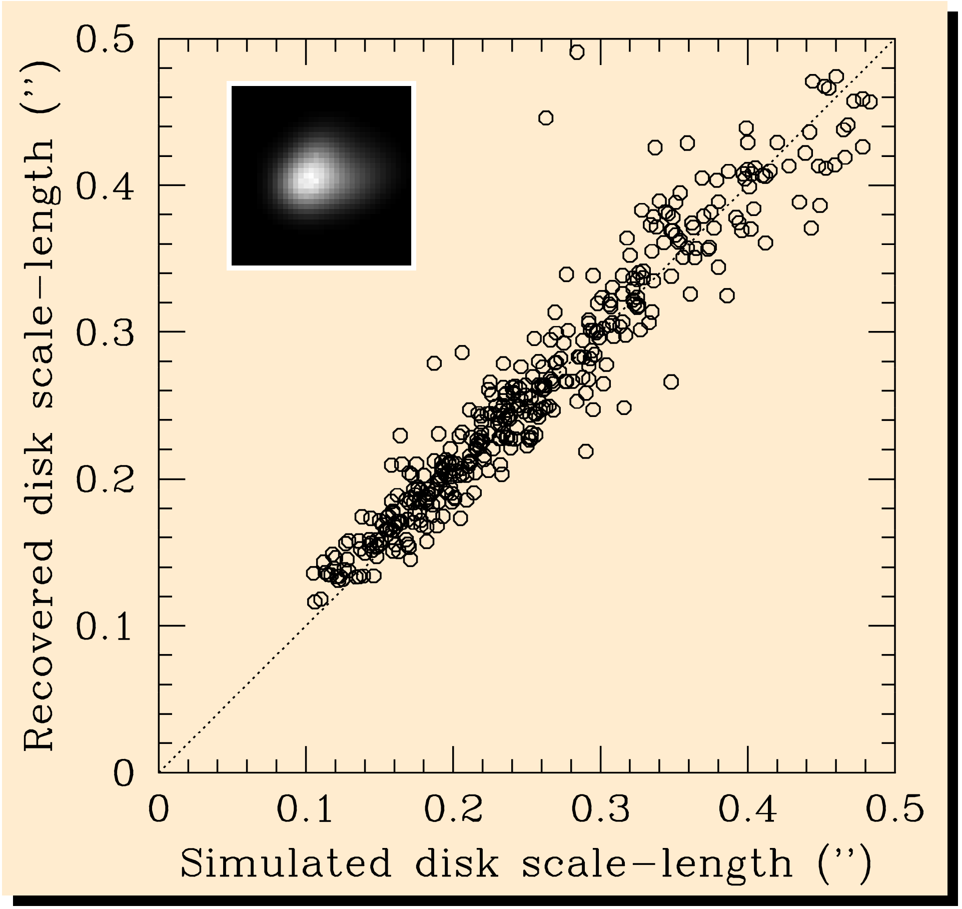
Parametric deconvolution of galaxies

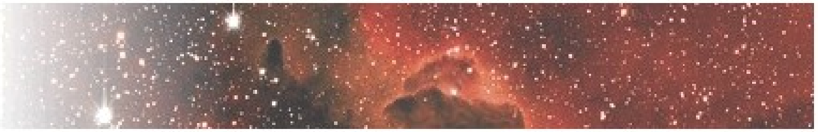




Measuring morphological parameters

- I<23
- 3h exposure with 0.7” seeing (ground-based) on a 3.6m telescope





Building a model of the PSF

- Software written in 1998 for SExtractor
  - Not publicly available yet
- Requirements:
  - Model variations across the field
  - Be able to deal with (moderate) undersampling
  - Number of degrees of freedom as small as possible while being capable of modeling any arbitrary (optical) PSF

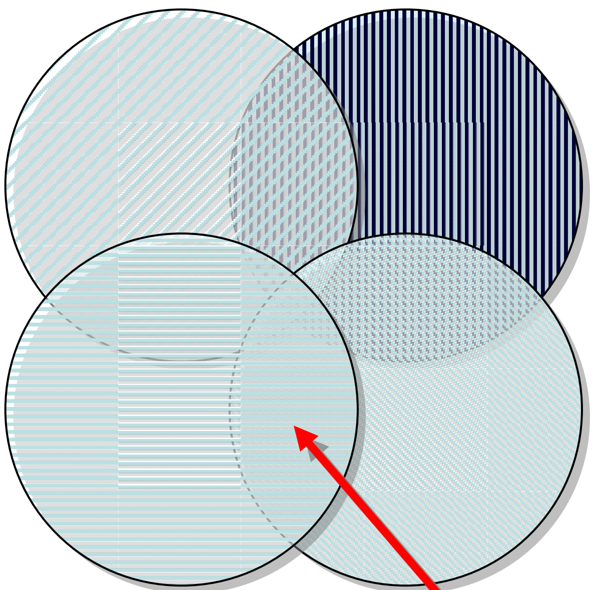
PSF models

- Analytical *vs* tabulated models
  - Analytical models are simpler to implement and can deal with undersampling “naturally”
    - BUT: simple (not instrument-dependent) models have trouble handling PSF features like diffraction effects (spikes and rings)
      - ☞ Such features can be tabulated provided that the data are correctly sampled, but this is not always the case (ex: WFPC2, NICMOS,…)
  - Tabulated models don’t have these limitations
    - BUT: over- and under-sampling are not properly handled.

A solution: “super-tabulation”

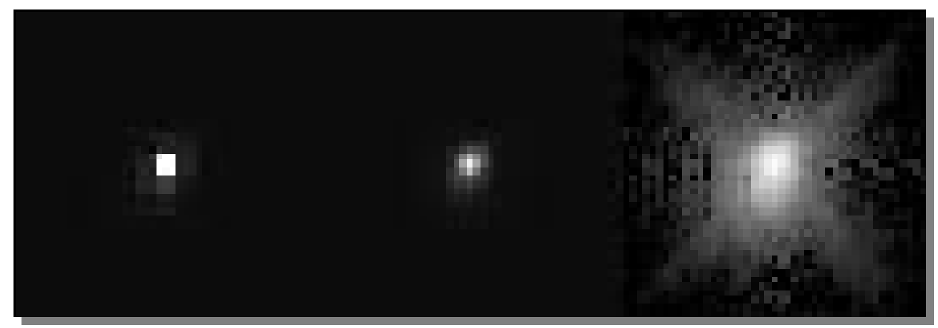
- The PSF is tabulated at a resolution which depends on the stellar FWHM (typically 3 pixels/FWHM)
  - Minimize redundancy in cases of bad seeing
  - Handle undersampled data by building a “super-tabulated” PSF model
  - Work with diffraction-limited images (images are band-limited by the autocorrelation of the pupil)
  - Find the sample values by solving a system using stars at different positions on the pixel grid
    - Intuitive approach: solve in Fourier space. Easy but suboptimum (no weighting)
    - Working in direct space would give much more robust results

Solving in Fourier space



Aliased portion of the spectrum

Reconstructed NICMOS PSF



Lauer 1999

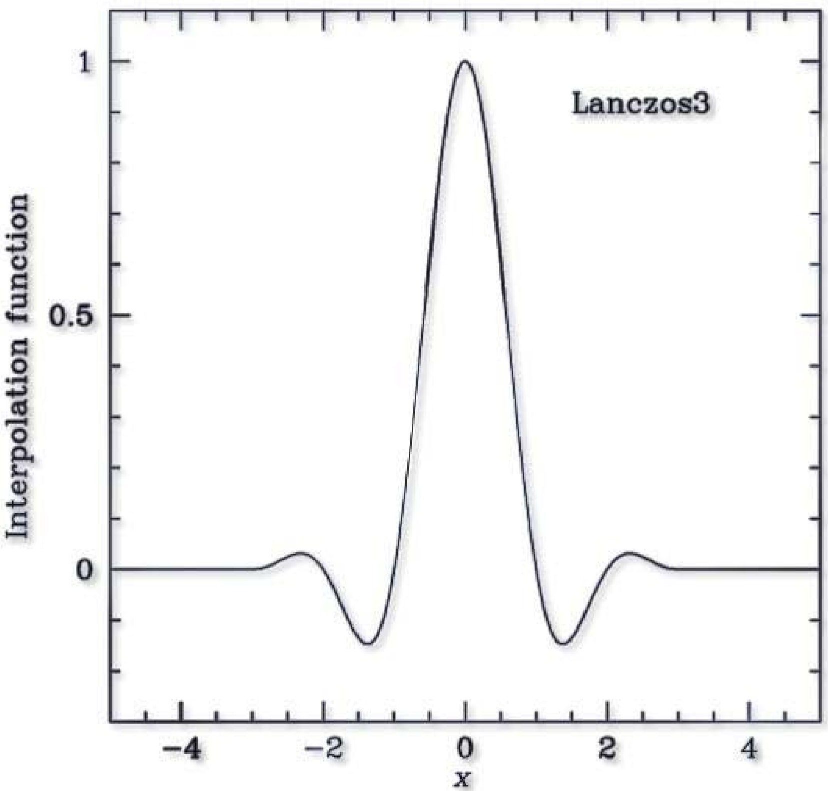
Solving in direct space

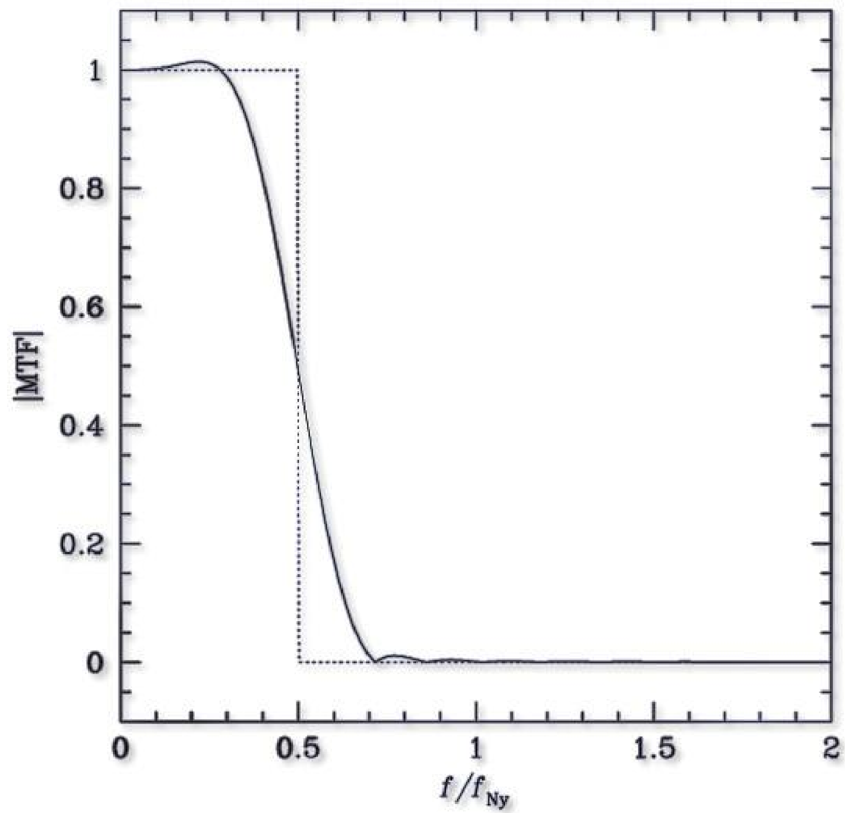
- A resampling kernel *h*, based on a compact interpolating function (*Lanczos3*), links the "super-tabulated" PSF to the real data: the pixel *i* of star *j* can be written as

$$P_{ij} = \sum_k h_j\left(\vec{\mathbf{x}}_k - \vec{\mathbf{x}}_i\right)\varphi_k$$

- The φ_k's are derived using a weighted χ^2 minimization.
  - Lots of computations involved:
    - ☞ Sparse matrix processing might prove useful for large models
    - ☞ In practice the oversampling of faint peripheral pixels can be dropped.

Lanczos interpolation kernel





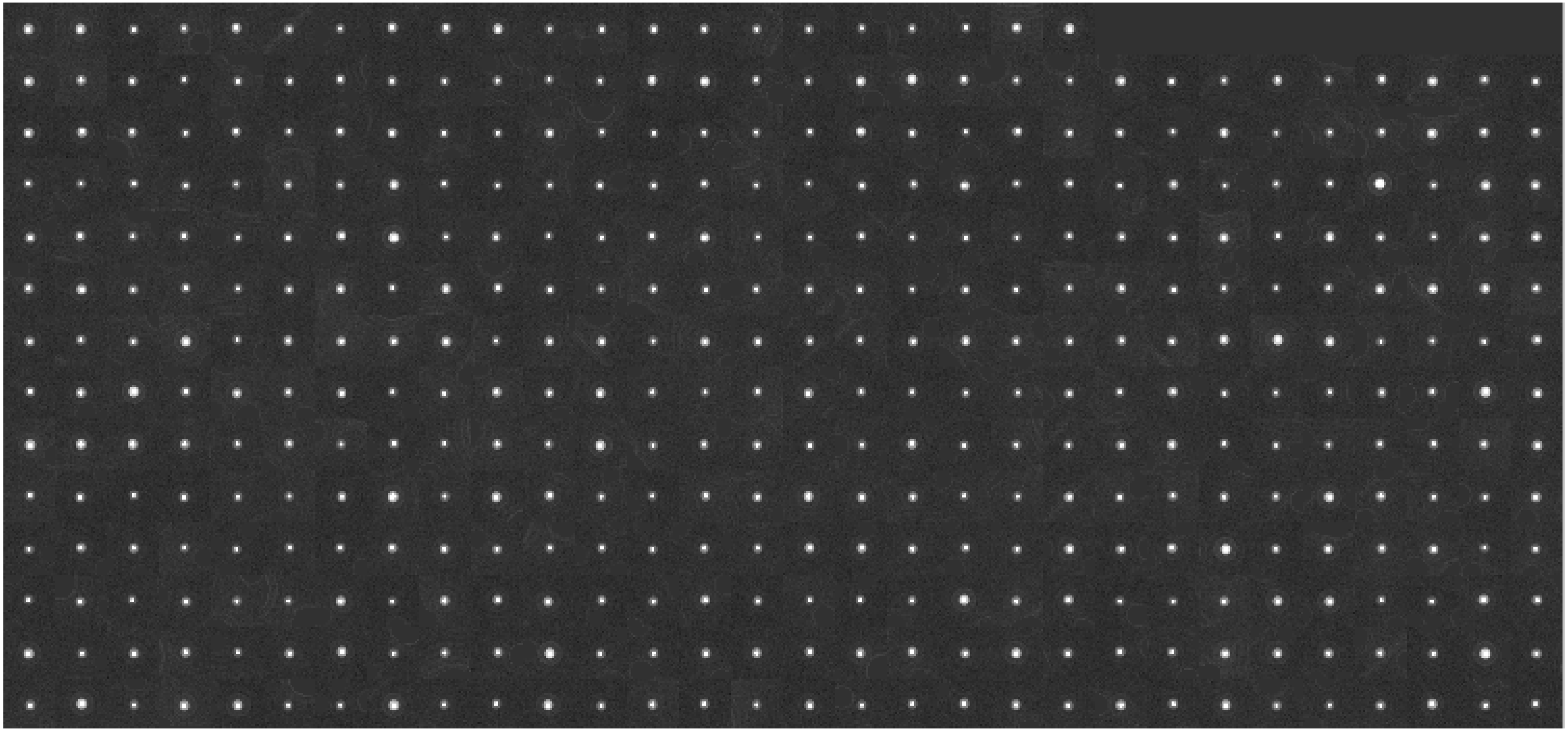
Testing on simulated, undersampled data

Diffraction-limited

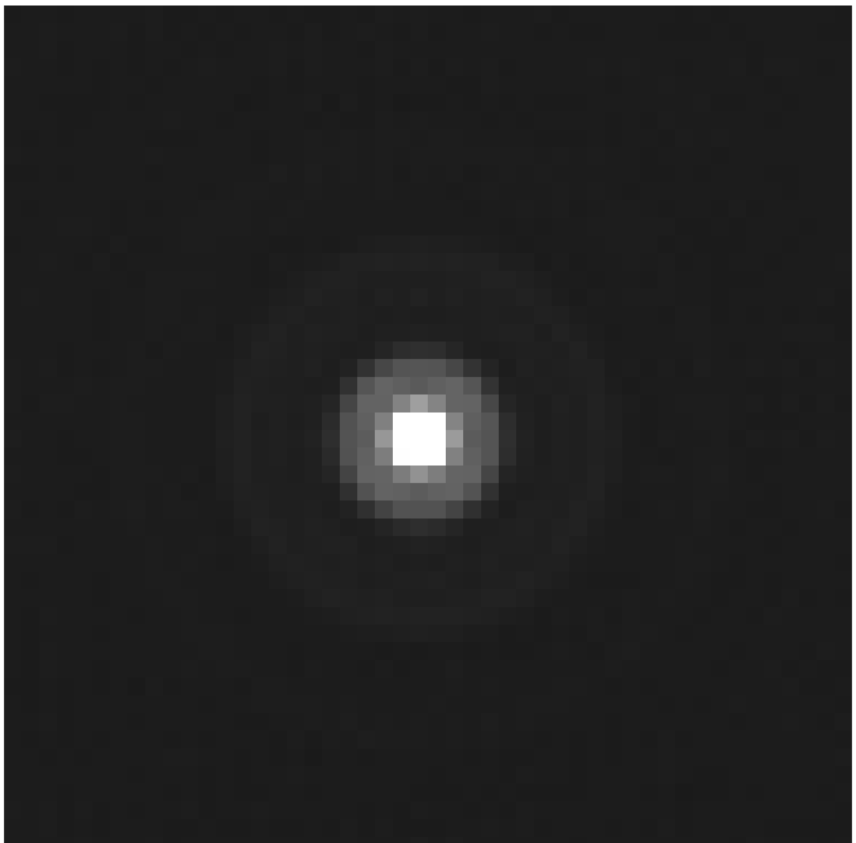
FWHM ≈ 1pixel

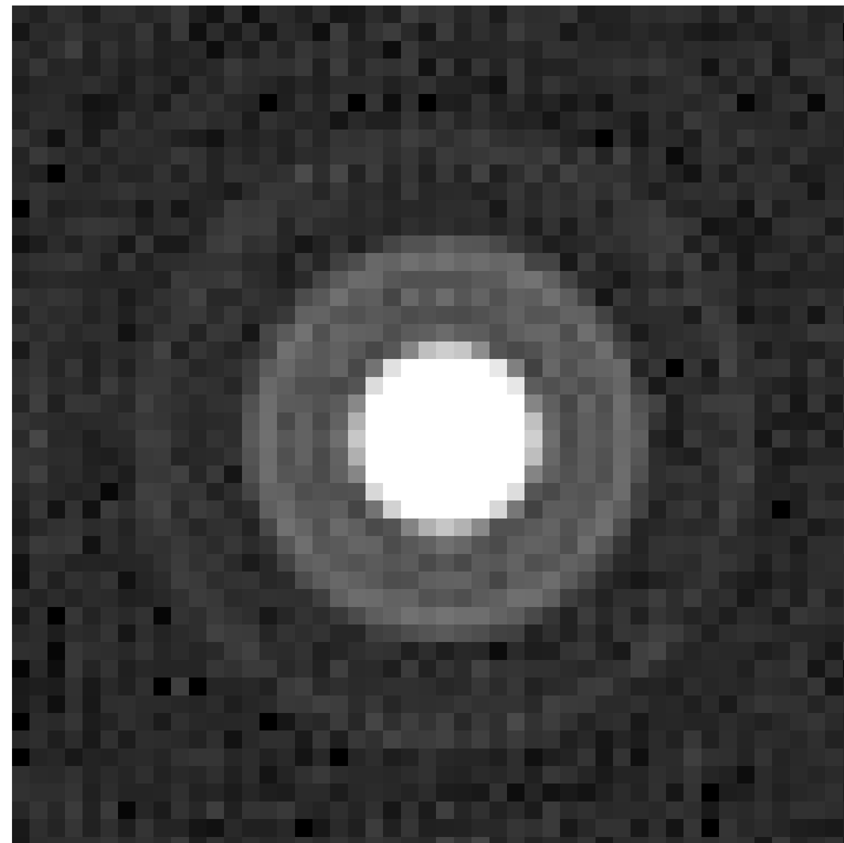
Moderately crowded

Automatic candidate selection

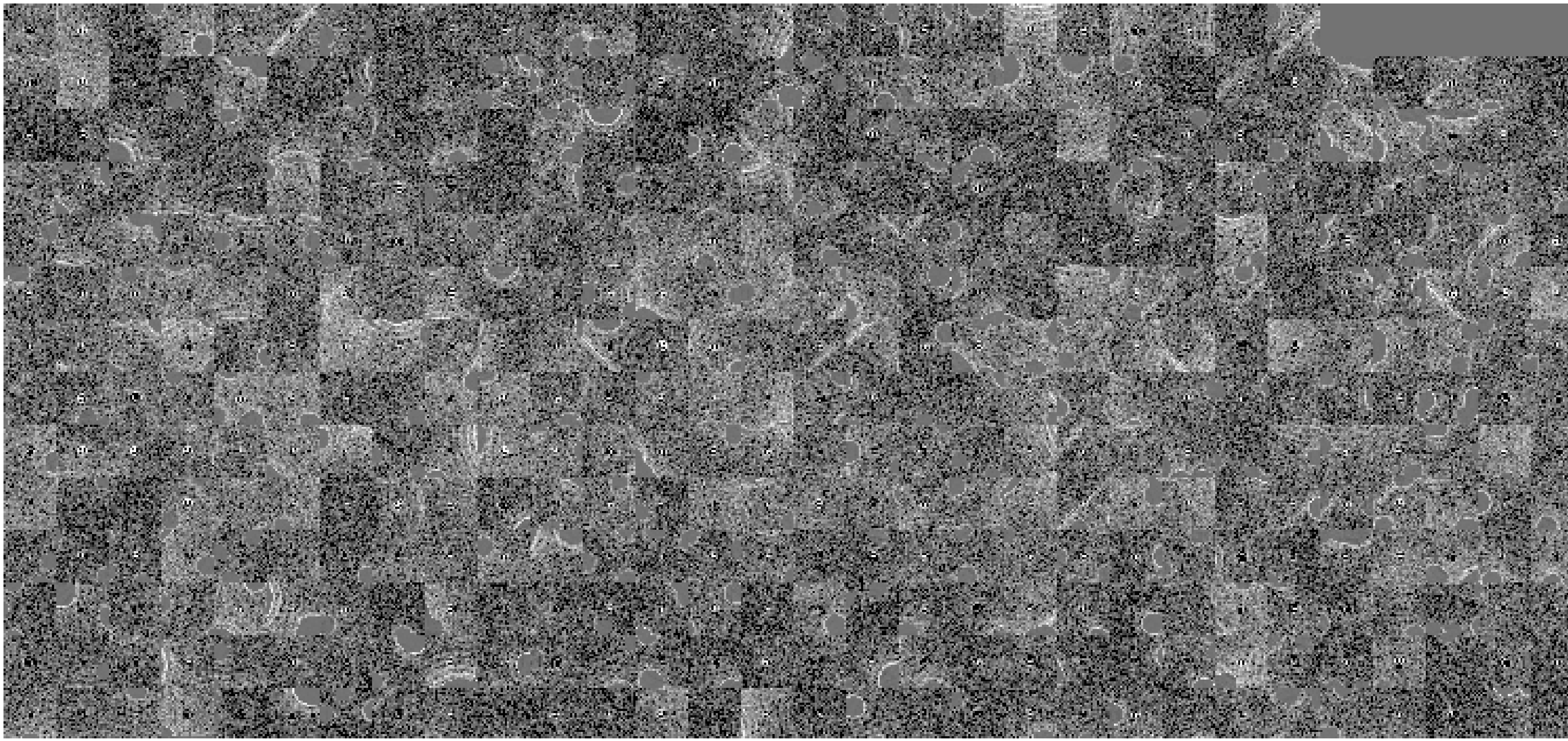


Recovered PSF with simulated, undersampled data

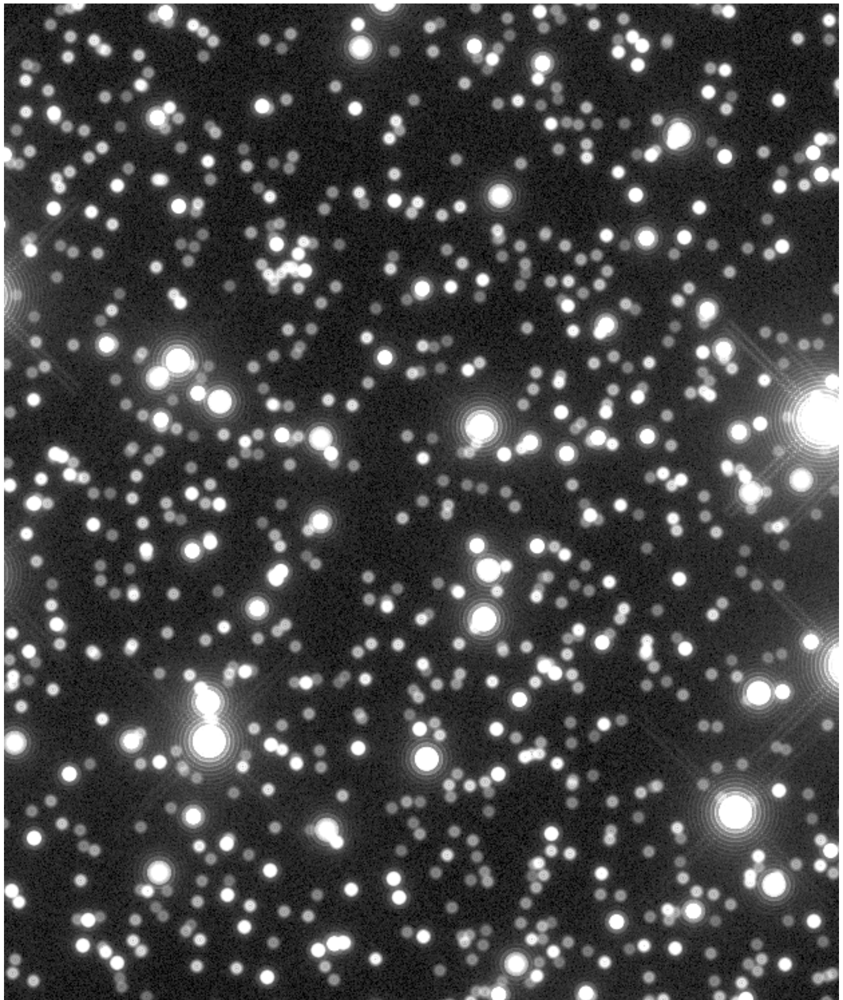




Residuals on simulated, undersampled data



Simulated, defocused data

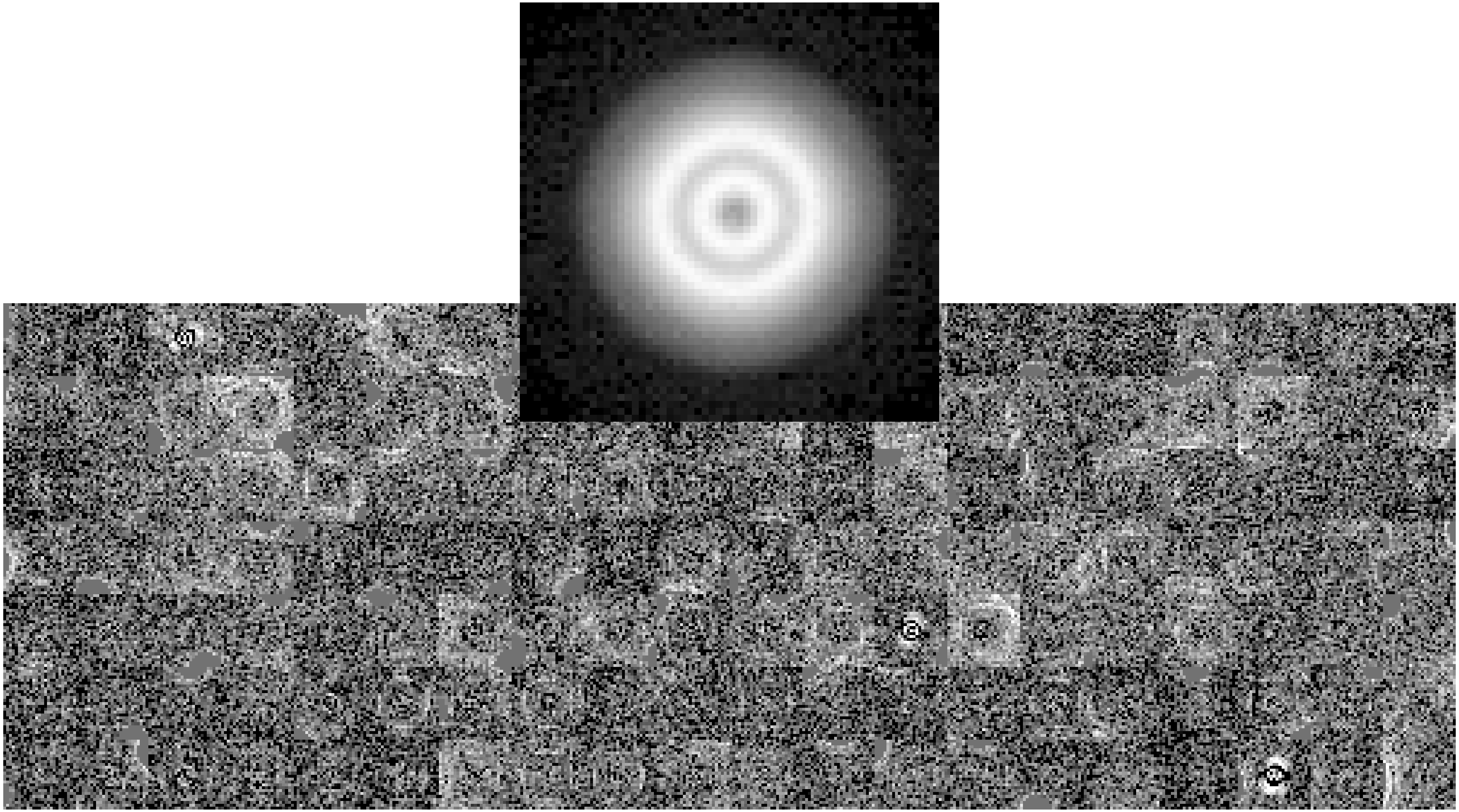


Diffraction-limited

FWHM ≈ 7 pixels

Moderately crowded

Results with simulated, defocused data



Using different basis functions

- The array of “super-pixels” can be replaced by a combination of ad-hoc basis functions ψ_b (the c_b are the parameters to determine)

$$P_{ij} = \sum_b \sum_k h_j\left(\vec{\boldsymbol{x}}_k - \vec{\boldsymbol{x}}_i\right) c_b \psi_{bk}$$

- Should be more robust in many cases
- One might use PCA components of the theoretical PSF aberrations for diffraction-limited instruments.

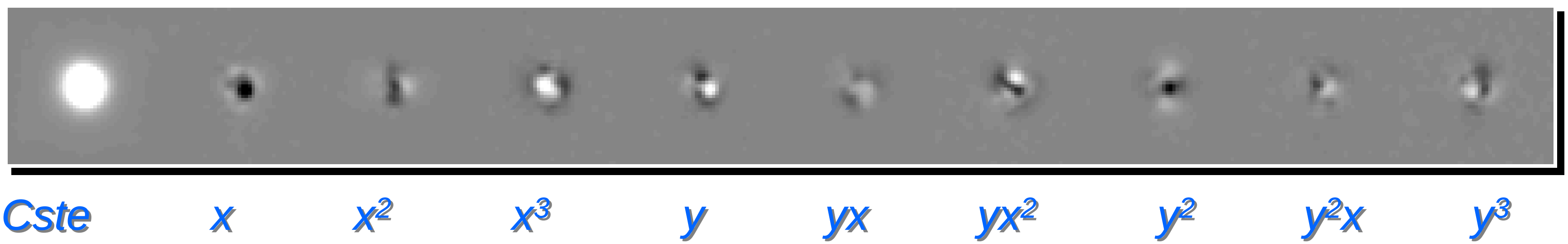
Handling PSF variations

- PSF variations are assumed to be a smooth function of object coordinates
  - ☞ The variations can be decomposed on a polynomial basis X_l

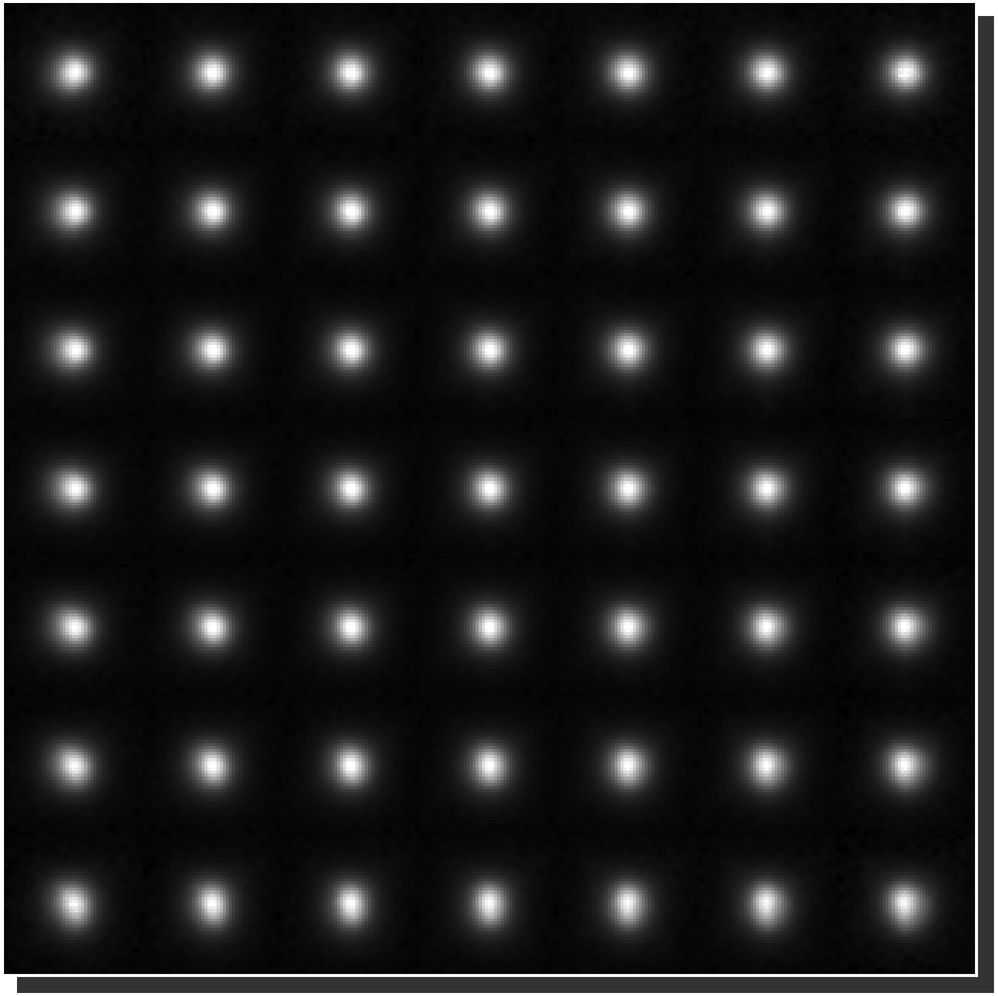
$$P_{ij} = \sum_{l} X_l \sum_{k} h_j\left(\boldsymbol{x}_k - \boldsymbol{x}_i\right)\varphi_{kl}$$

- A third order polynomial (l =10) is generally sufficient to describe the variation of the PSF with position in the field
- Different basis functions, with arbitrary parameters (flux, instrumental context, etc.) can be used for X_l
- In our case a KL decomposition (e.g. Lupton et al. 2OO1) was not beneficial (and in fact it makes the rejection of « bad » PSF prototypes harder).

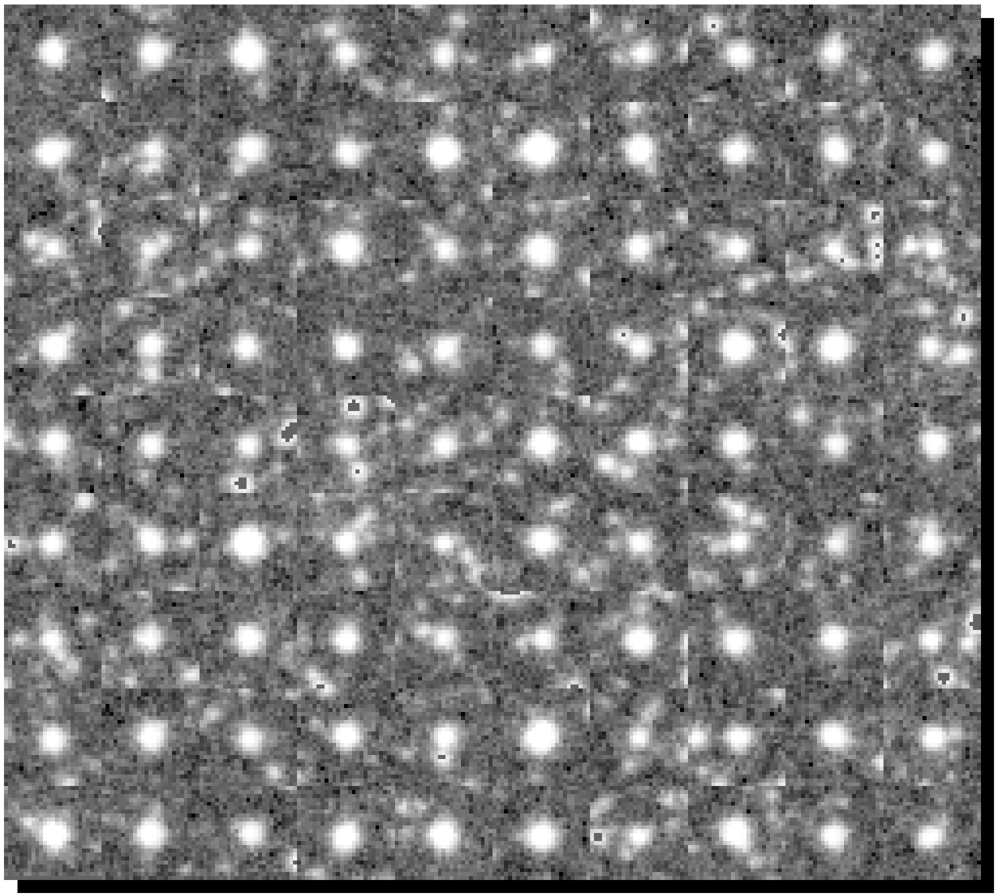
Example of φ_{lk} PSF components for a UH8k image



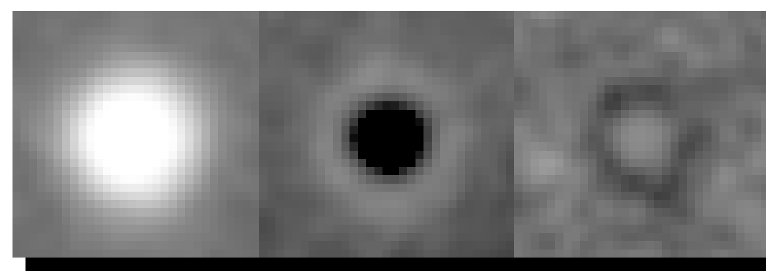
Reconstructed UH8k PSF



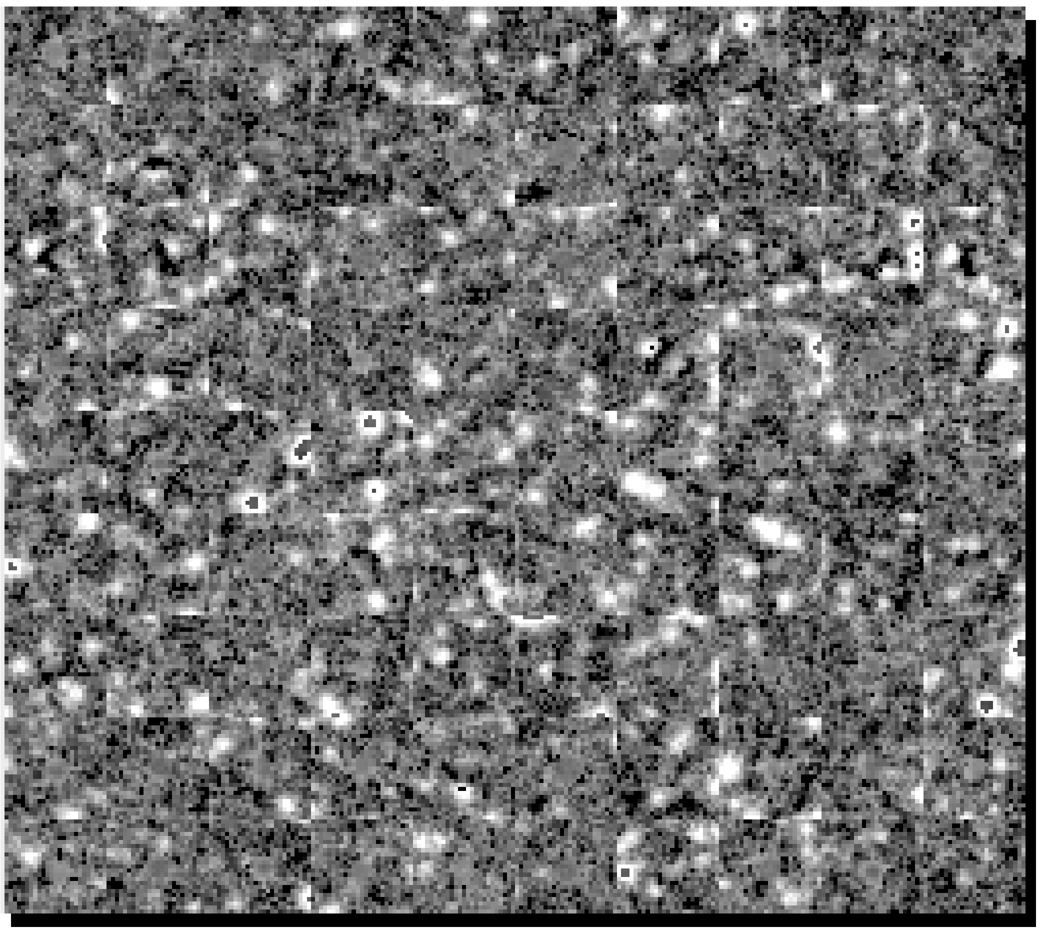
Testing on real, non-linear data



- Schmidt-plate exposures in the galactic plane
- FWHM ≈ 3pixel
- Second order polynomial of FLUX_AUTO



Star subtraction on Schmidt-plate data

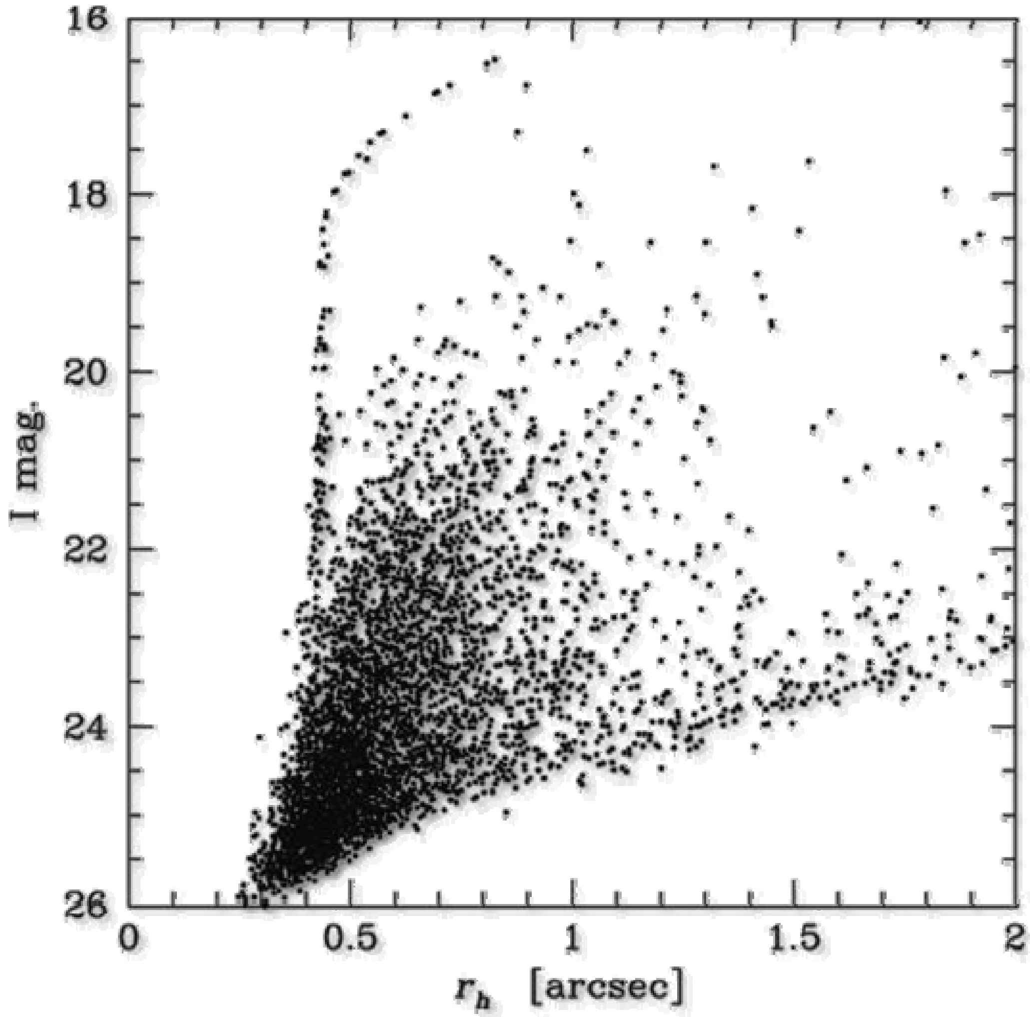


- Schmidt-plate exposures in the galactic plane
- FWHM ≈ 3pixel
- Second order polynomial of FLUX_AUTO

Finding prototype stars

- Basically we are looking for something we don't know yet
  - PSF variability makes the stellar locus “fuzzy” in feature space
  - Problems due to crowding at low galactic latitude
  - Confusion with galaxies in cluster areas
- Empirically designed automatic selection based on magnitude,half-light radius, ellipticity, crowding and saturation flags seems to work fine
  - Remaining configuration parameters for selection essentially consist of acceptable FWHM range and ellipticity
  - Iterative rejection procedure based on similarity between samples and a rough PSF estimate

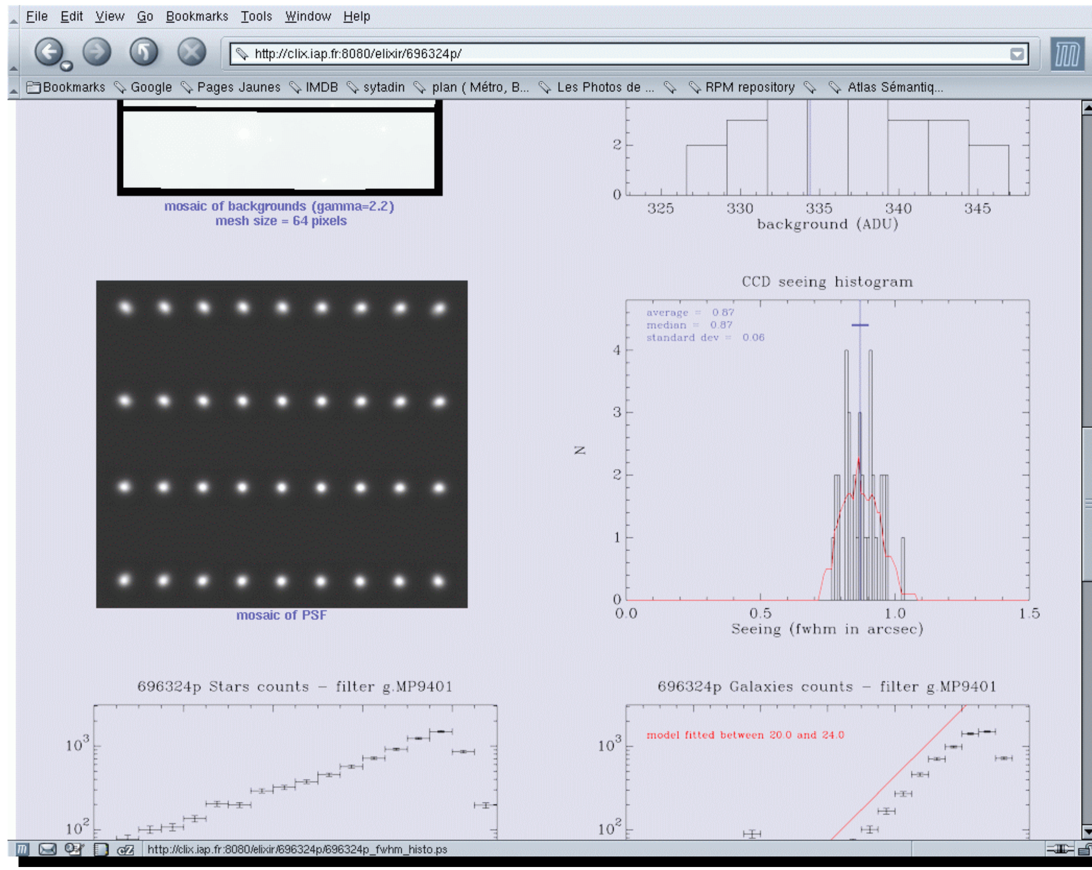
Half-light radius/magnitude diagram



TERAPIX

QualityFITS

- AstroWISE project developed at TERAPIX by F. Magnard
- Provides quality control for FITS images
  - Background homogeneity
  - PSF and variability
  - Source counts
  - Weight maps
- Diagnostic generated automatically for all incoming and outgoing MEGACAM survey images
  - FITS and XML formats
  - Access from [Spica]



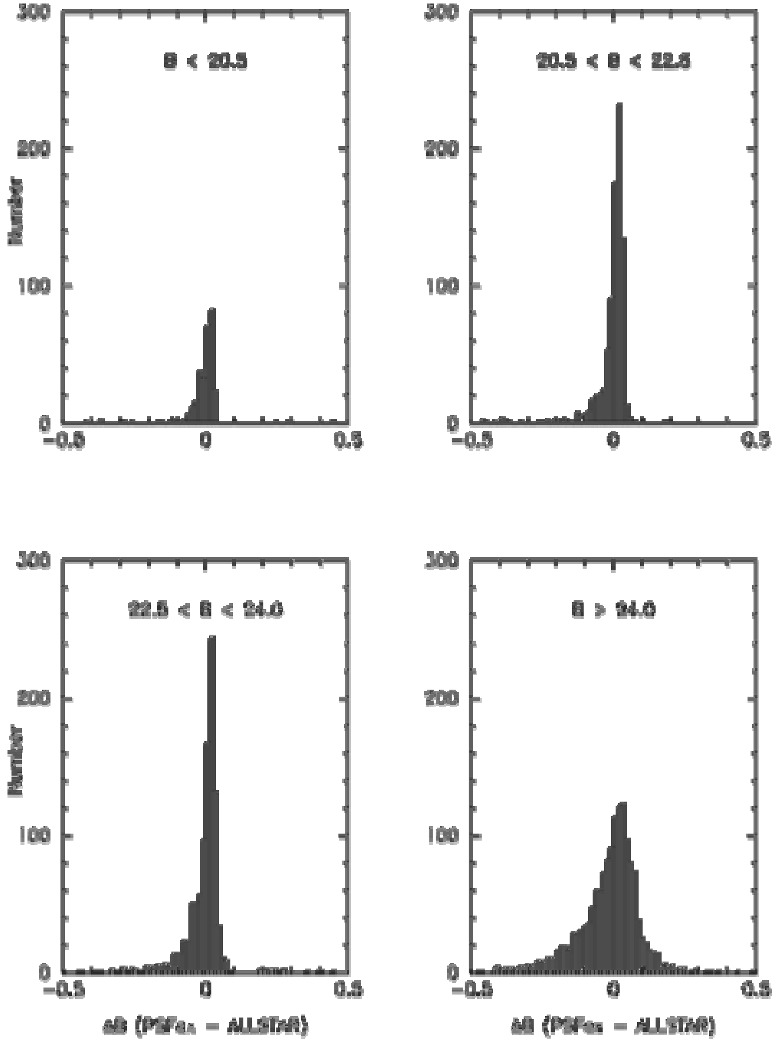
Fitting the PSF model

- Identify star "clusters", like in DAOPhot (*Stetson 1987*) and proceed interatively:
  - First a unique star is fitted
    - The basic centering algorithm is a modified gradient descent
  - The star is subtracted from the cluster and a local maximum sufficiently distant from the peak of the first star is identified
  - Two stars are fitted and subtracted, and a new maximum is found
    - Iterate up to 11 stars/cluster or
    - Stop if stars coalesce during the centering process

Current Performance

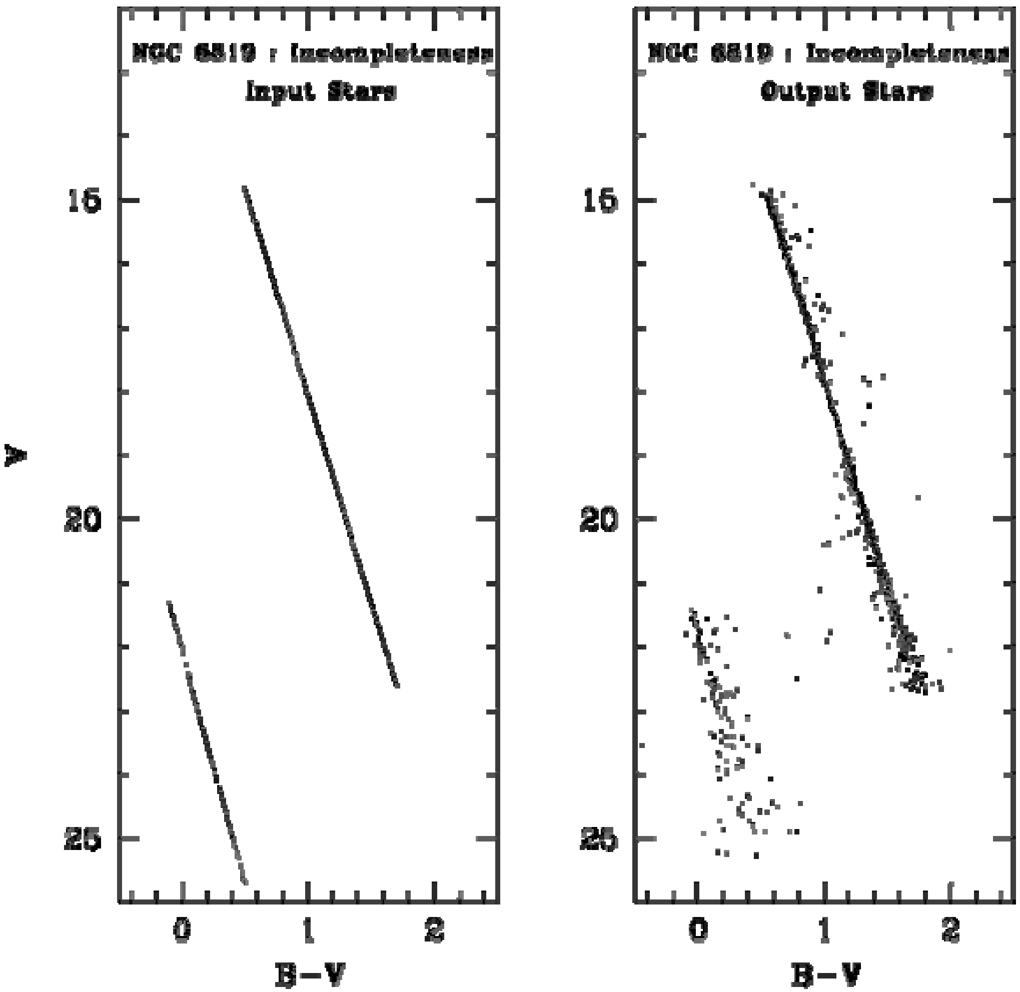
- Processing speed:
  - For building the PSF model: ~130 stars/second (Athlon 2GHz)
  - For the PSF-fitting: ~100-500 stars/second (Athlon 2GHz)
- Measurement accuracy:
  - Slightly better than DAOPhot on properly sampled, non-crowded fields
  - Slightly worse than DAOPhot (one pass) on properly sampled, crowded fields
  - Significantly better than DAOPhot on undersampled images
- Poor completeness (~99% for “obvious” detections) because of the underlying SExtractor detection scheme

Application: Comparison with DAOPhot on NGC 6819 (CFH12k)



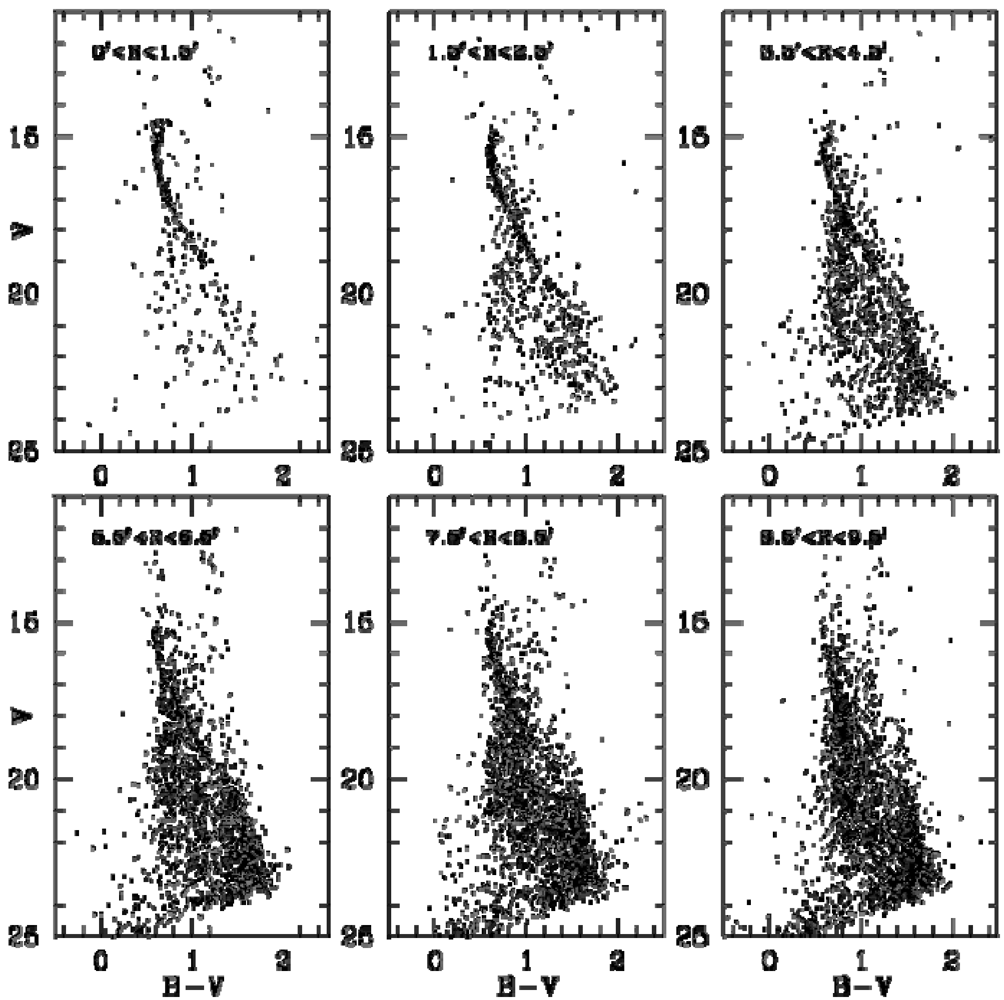
Kalirai et al. 2001a

Application: Photometric accuracy in NGC 6819 (CFH12k)

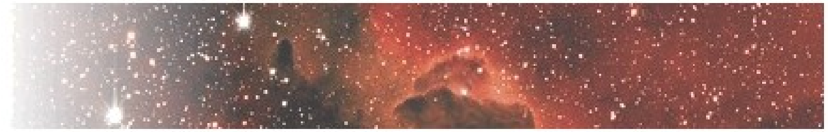


Kalirai et al. 2001b

Application: Colour-magnitude diagrams in NGC 6819 (CFH12k)



Kalirai et al. 2001b



Conclusions

- The PSFEx approach to PSF modeling gives reliables results
  - Undersampled data (down to 1 pixel FWHM)
  - Variability across the field
  - Moderately crowded fields
- Currently available as an external module: “PSFEx”
  - Soon to be publicly released (together with QualityFITS)
  - But not for PSF fitting in SExtractor
    - Mostly completeness issues
- Wait for SExtractor3
  - New detection scheme
  - Handling of variable noise ACF