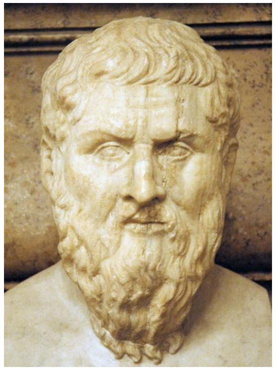
Plato: On board and on ground algorithms of data processing

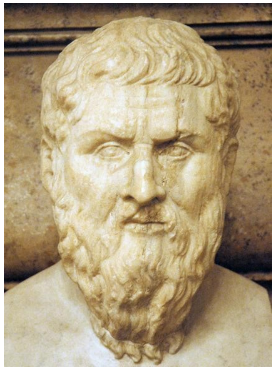


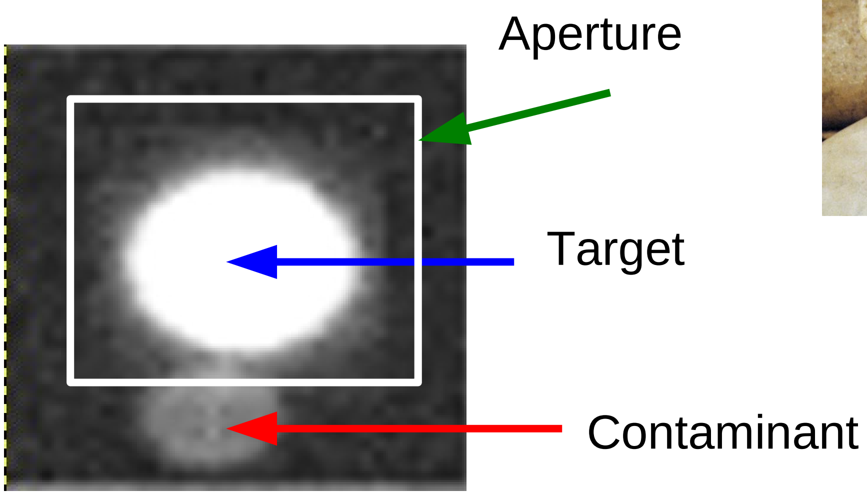
Réza Samadi

LESIA, Observatoire de Paris

- The sources of perturbation and their correction
- Assessment of the performances
- The configuration mode
- Organization and schedule

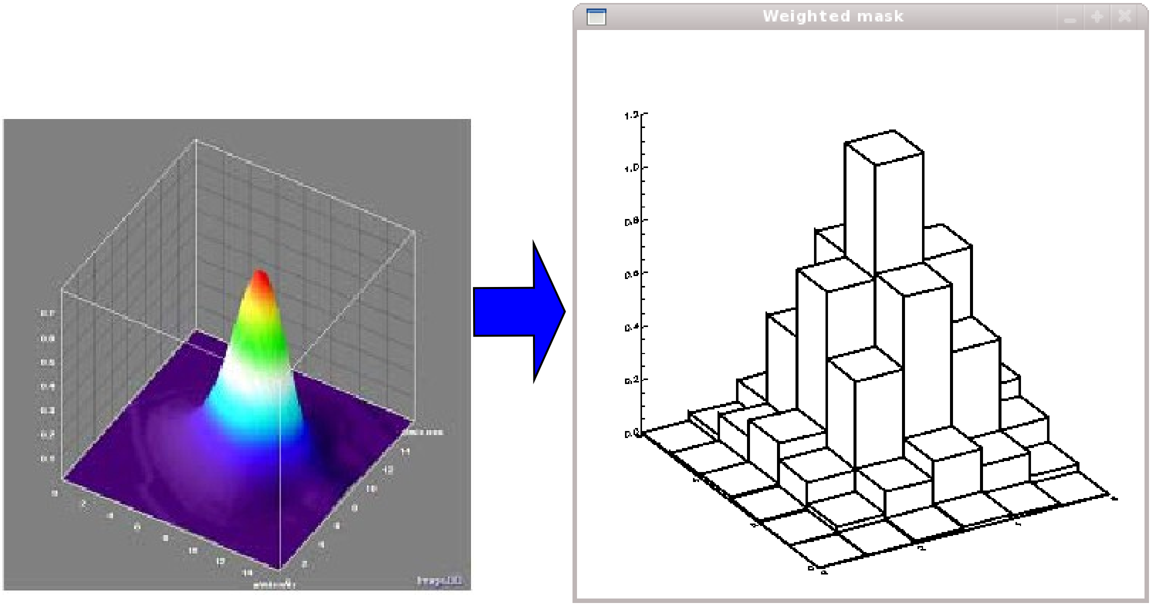
The problem of confusion





- To avoid confusion : use of a weighted mask
- But:
  - We need to know the PSF
  - If too narrow: We can lost significant part of the star flux

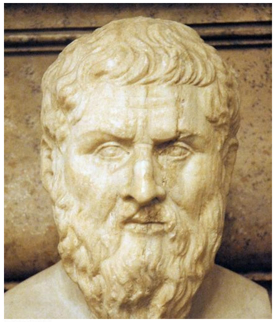
- Thanks to GAIA: positions and intensities of the contaminants known a priori
- optimization of the width of the mask

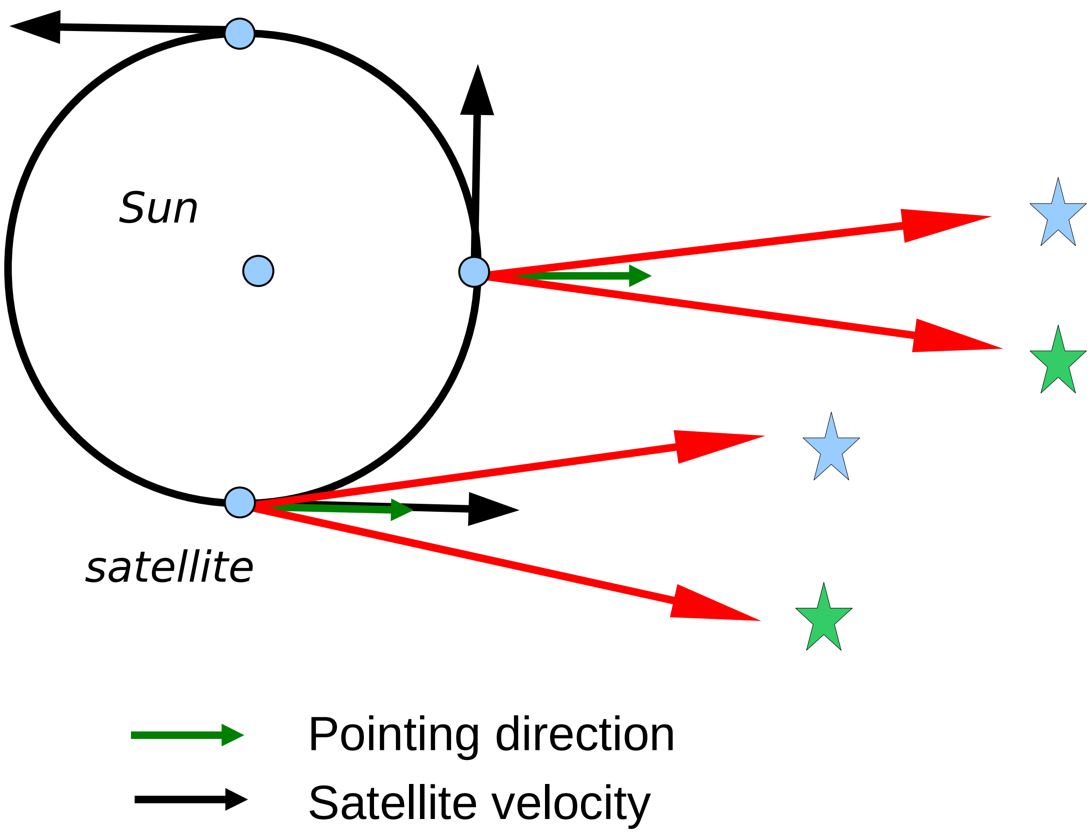


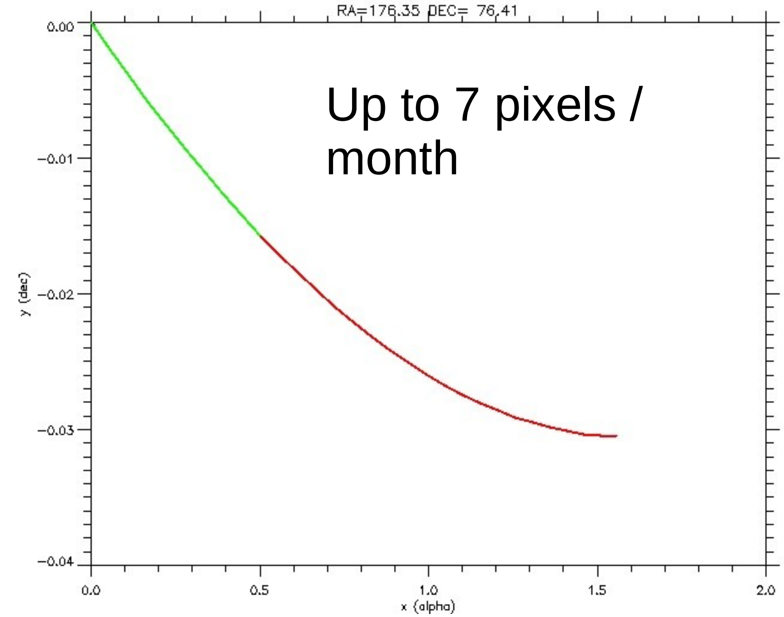
weighted mask

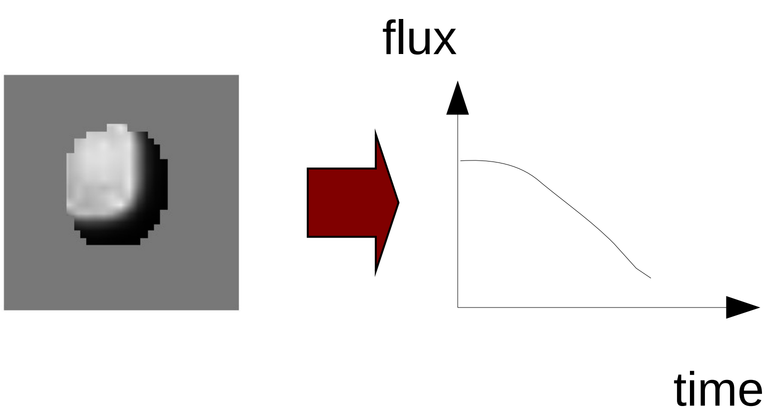
Correction of instrumental and environmental perturbations

Differential (kinematic) aberration



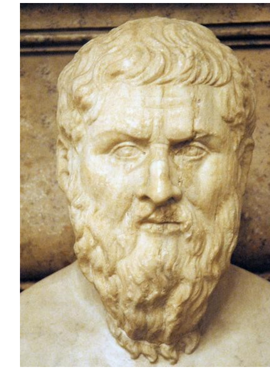






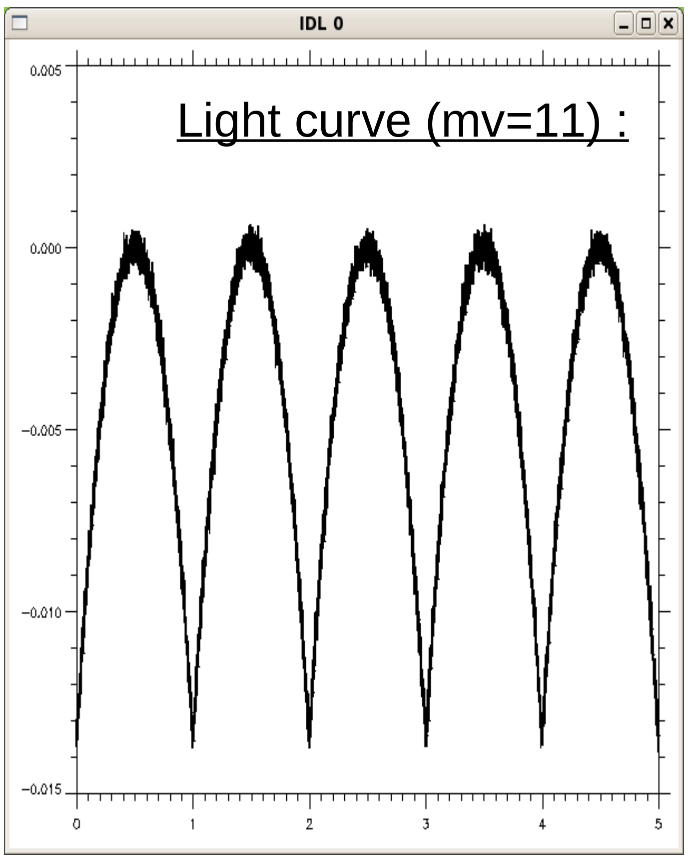
- *In addition*: Thermoelastic variations of the telescope pointing direction

Differential aberration and mask updates

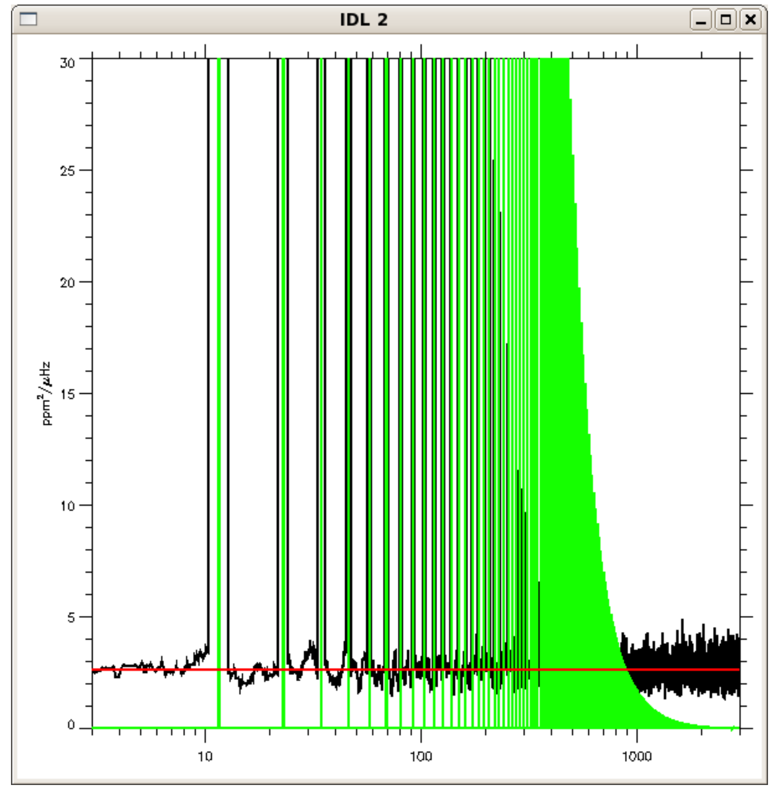


Worst case: 7 pixels / month
= 0.23 pixels / days

The mask is updated every day



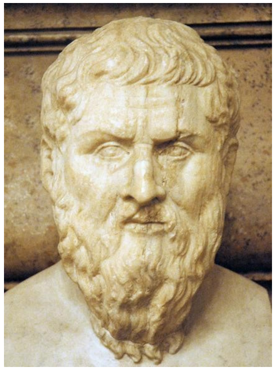
Power Density spectrum (5 months)



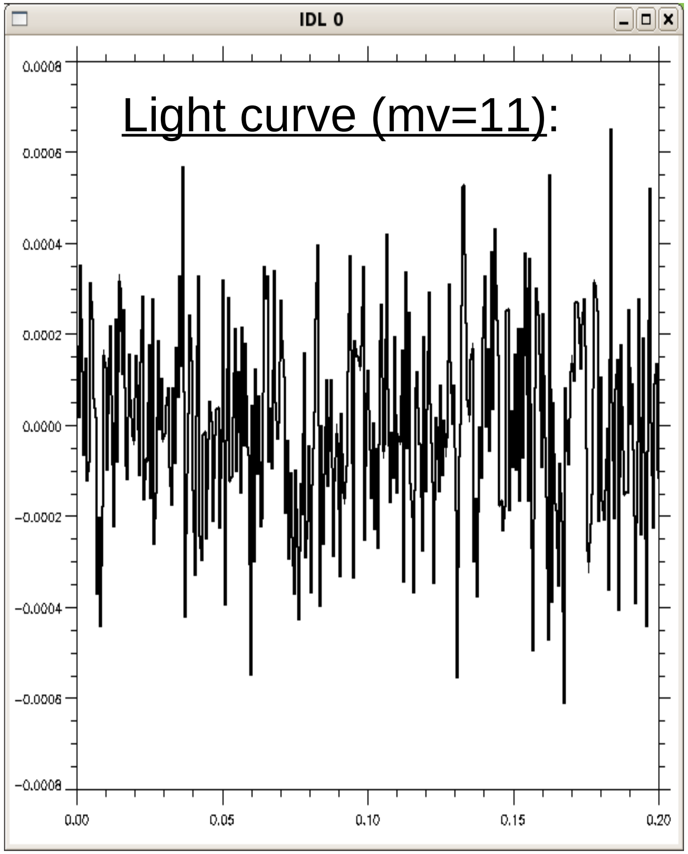
red : white noise level (40 telescopes)
black : signal + photon noise
green : signal only (no photon noise)

⇒ ~ **80 peaks** are above the photon noise level

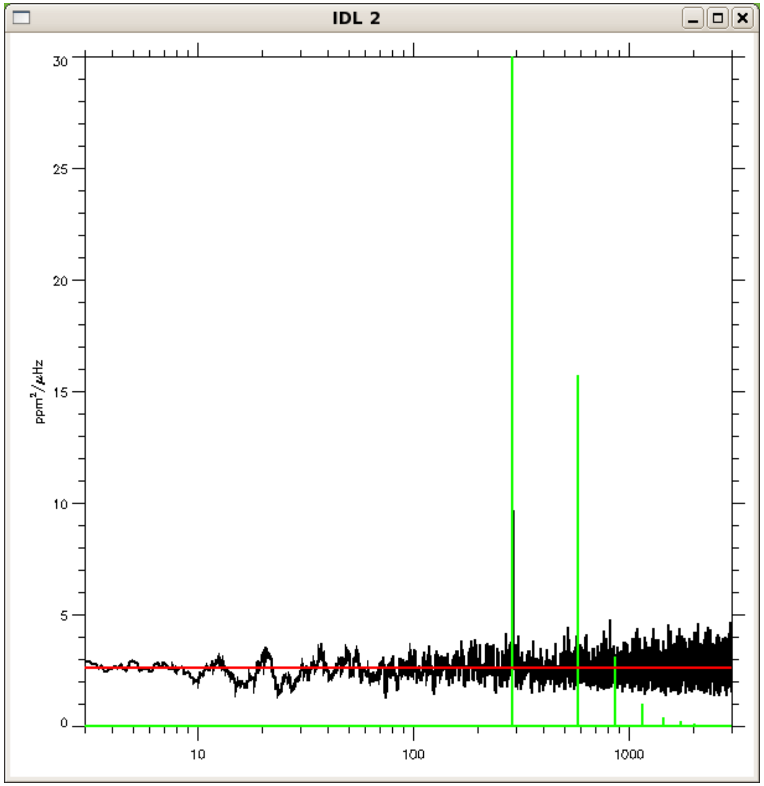
Differential aberration and mask updates



The mask is updated <u>every 1h</u>



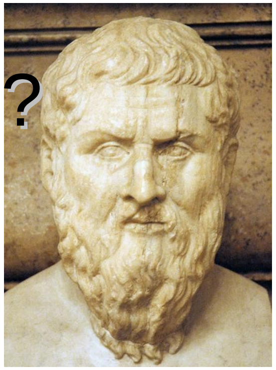
<u>Power Density spectrum</u> (<u>5 months</u>)



⇒ **<u>3 peaks</u>** are above the photon noise level

- **<u>Updates every ~ 1 000 s</u>** (displacement ~ 1/400 pixels) ⇒ flux variation ~ 1.8 ppm
- For star with mv>11 ⇒ **<u>NO peaks</u>** above the photon noise level
- For brighter stars: we increase even more the frequency of the updates

Updates of the masks: how we proceed (on board) ?



$$PSF(x,y) = F(x-x_{0,}y-y_0)$$

(X_0,Y_0) : star centroid at a given instant

$$x_0 = f(t) \qquad y_0 = g(t)$$

➔ We assume to have available an analytic model of the PSF

The star centroid (x_0,y_0) moves due to:

- The kinematic differential aberration – *fully predictable*
- The movements of the satellite (jitter)
- The thermoelastic differential aberration

How to derive the star displacements at any instants ?

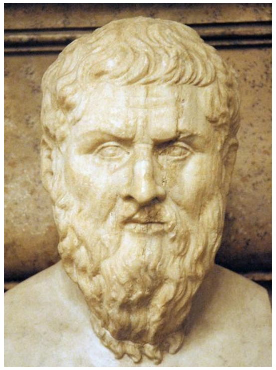
1) Imagettes of 1 000 reference stars (the brightest non saturated stars) :

⇒ variations of the pointing direction of the normal telescope

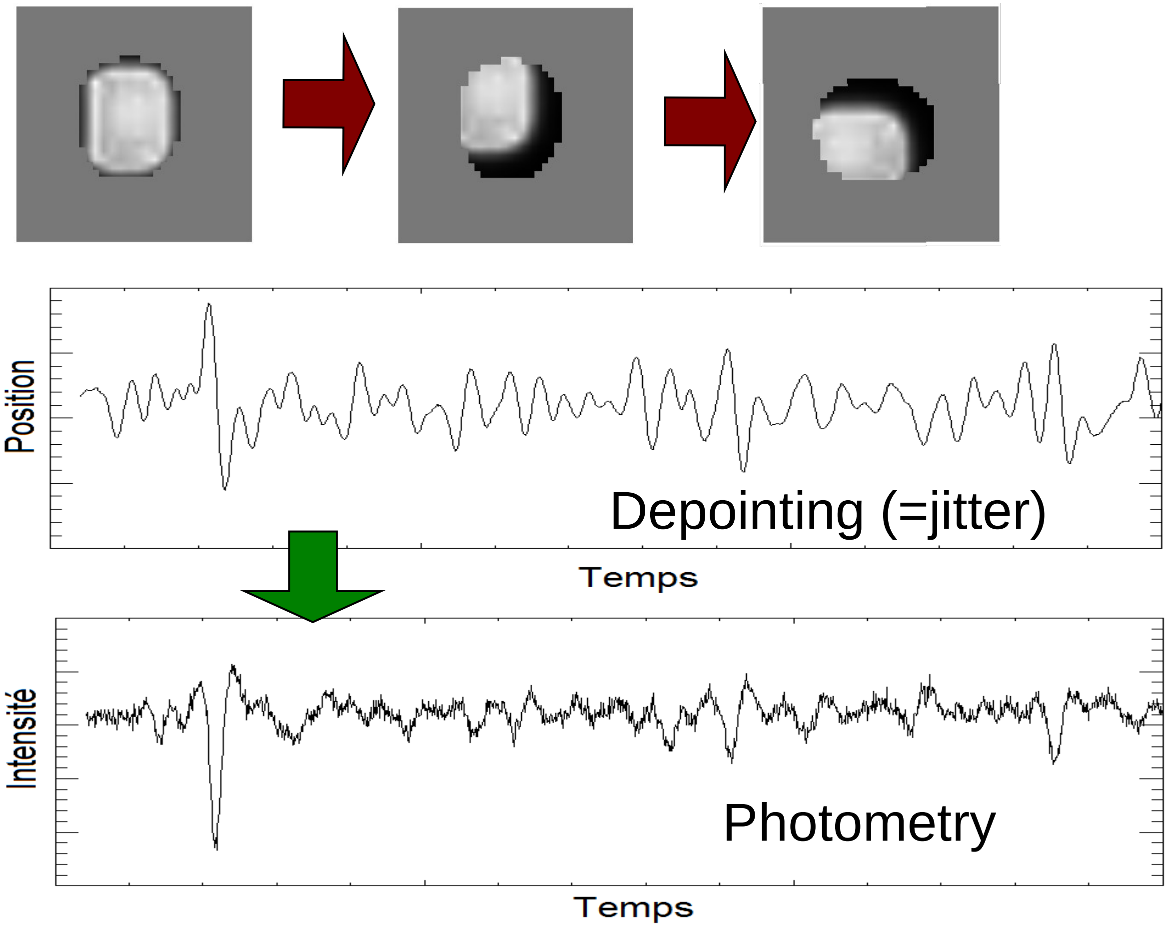
⇒ we can finally derive the **actual** displacements of any stars within the FoV of the telescope

2) The measured barycenter of the stars

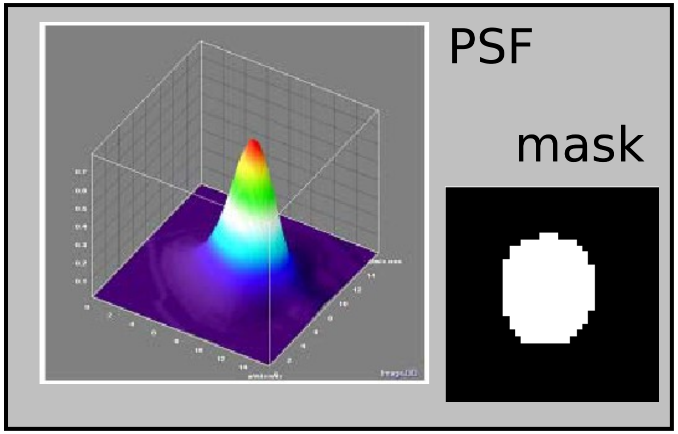
Noise dues to the satellite jitter

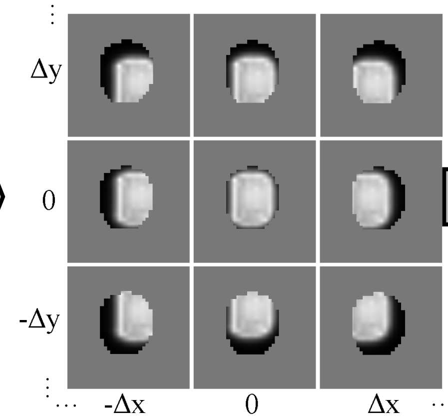


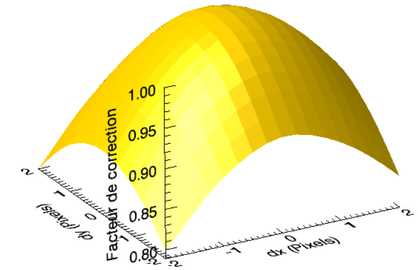
The satellite moves ! (=jitter)

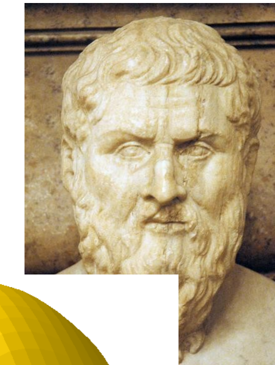


Jitter noise : correction (on ground)









Surface for the jitter correction

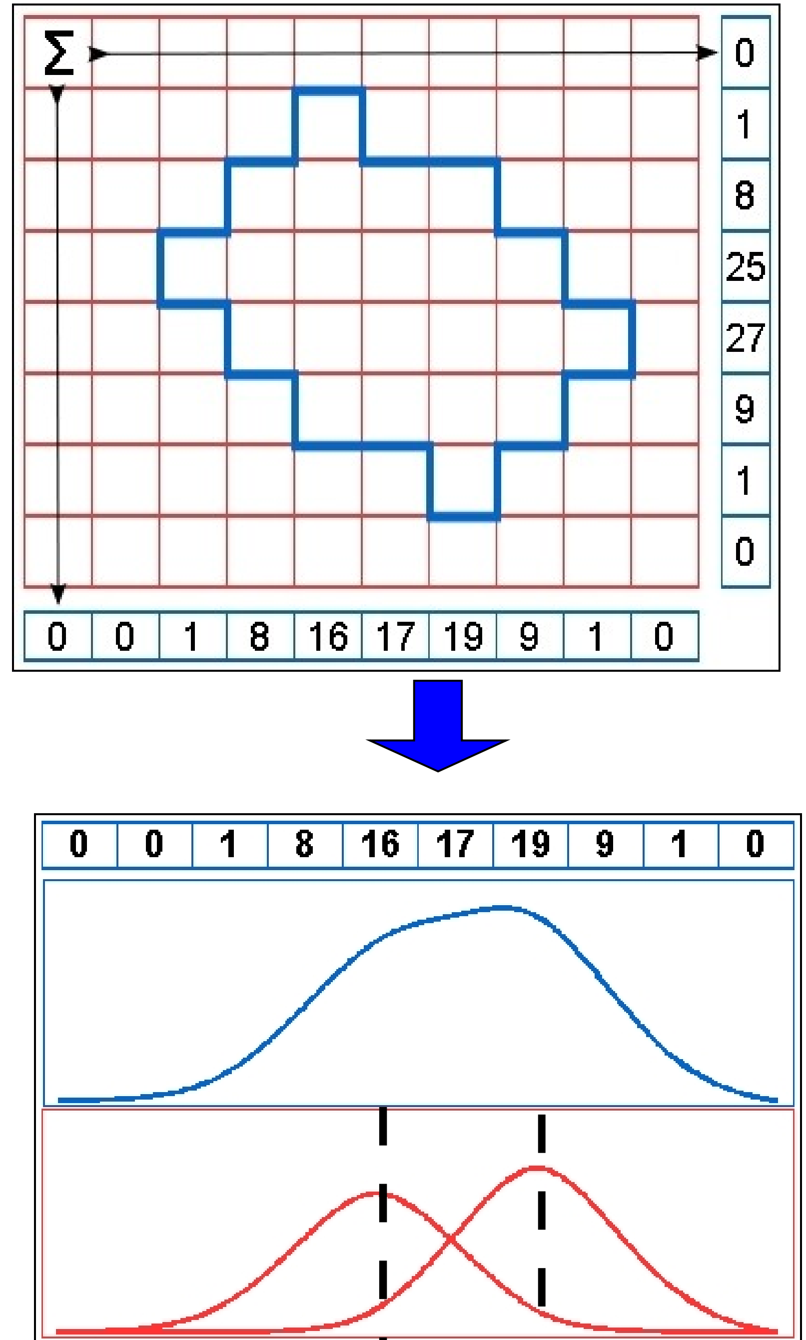
- From the PSF, we can predict the perturbations induced by any displacements :

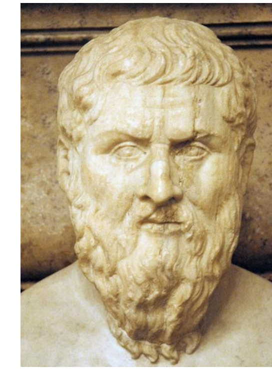
$$F_{c_i} = K(\Delta x_i, \Delta y_i).F_{m_i}$$

Fialho et al (2007, PASP)

- This method also corrects the differential aberration
- But we need to derive accurately the star displacements (Δx, Δy) as well as the PSF !
- The surface used for jitter correction must take the presence of contaminants into account.
- Thanks to GAIA with can a priori know the positions and intensities of the contaminants

An alternative photometry methods: *Line Spread Function fitting*



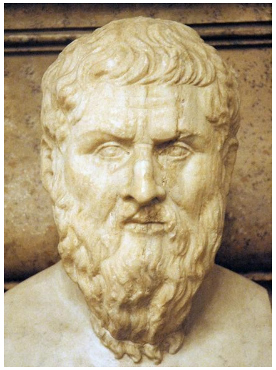


- **LSF-fitting**: flux estimation of individual components in compound objects
- Advantages:
  - Improved management of confusion
  - No sensitive to jitter
  - No need to update the mask ⇒ continuous photometry

➔ Need for a representative LSF

-

Performances of the photometry methods



Method	Noise level (ppm/1h)	
	PSF 0°	PSF 14°
Binary mask	29.2	32.7
Binary mask + jitter correction	28.6	32.5
Weighted mask	28.2	32.4
Weighted mask + jitter correction	**27.9**	**32.2**
LSF - Gauss	*28.4*	*33.6*
LSF - PSF	*31.8*	*36.7*

Time series of simulated images

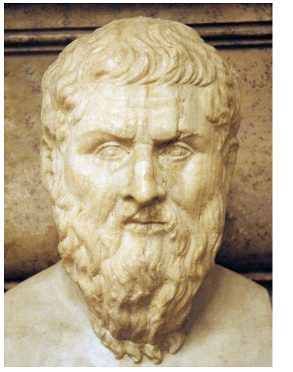
Target: mag =11

A single contaminant:
- Mag=13
- 1 pixel far from the target

Gaussian weighted mask

In all cases:
best performances with the weighted mask

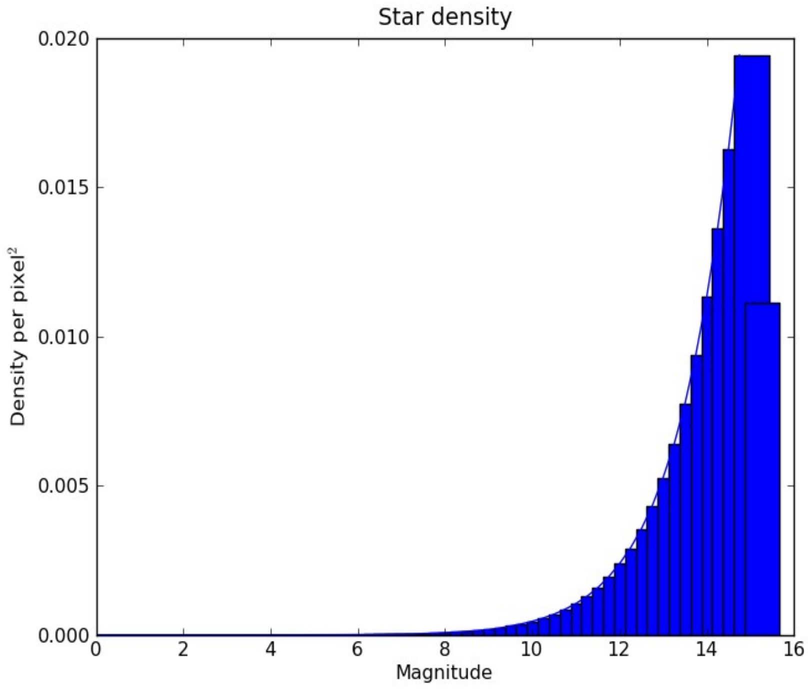
A tool to assess the global performances



Included perturbations:

- Photon noise target
- Photon noise contaminants
- Sky background (constant)
- Readout noise
- Quantification noise
- Jitter noise:
  - Target
  - Contaminants
- Jiitter correction (residues):
  - Target
  - Contmaninants

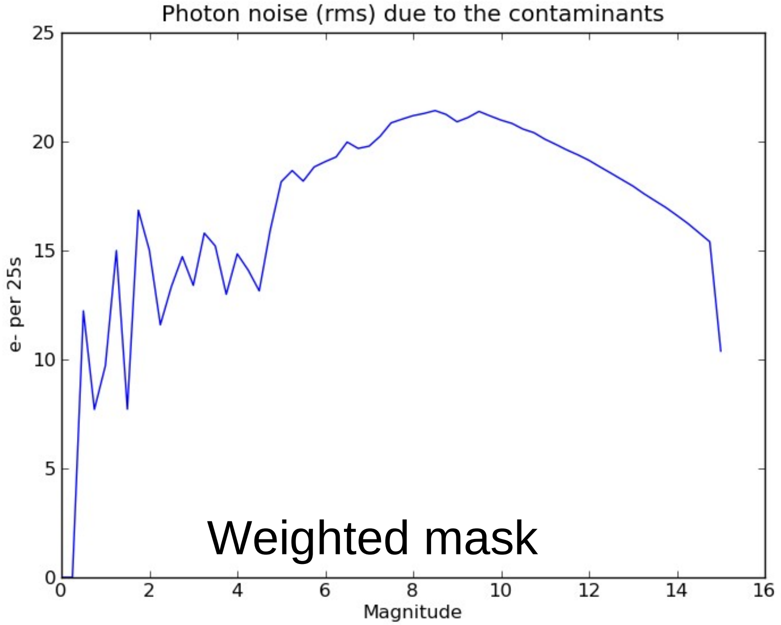
PRNU: neglected

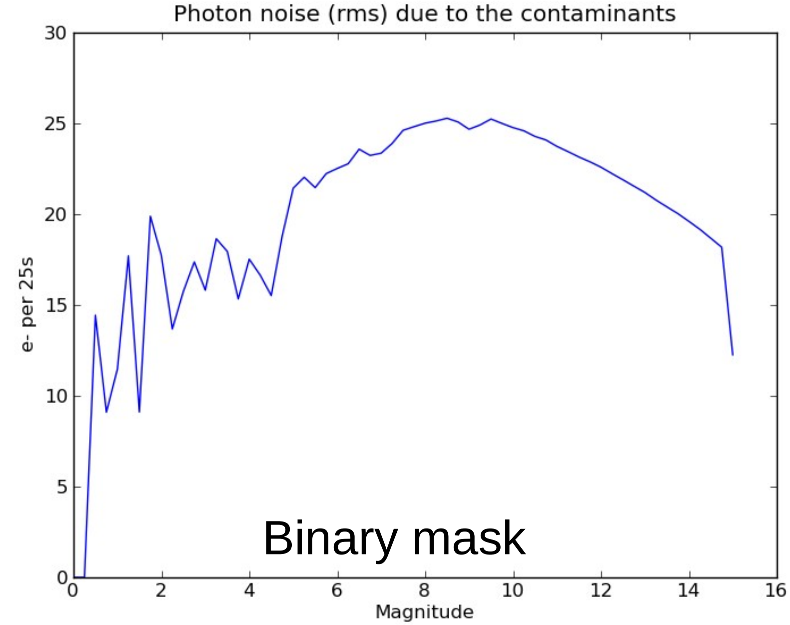


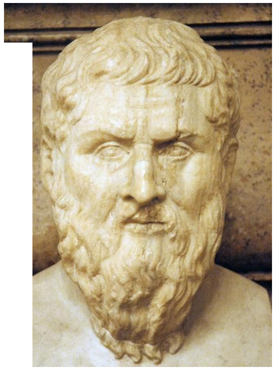
Inputs:

- Star density (star number per pixel²)
- PSF (e.g from the optic model)
- Mask (e.g. binary or weighted)
- PDF of the jitter (e.g. normal distribution)

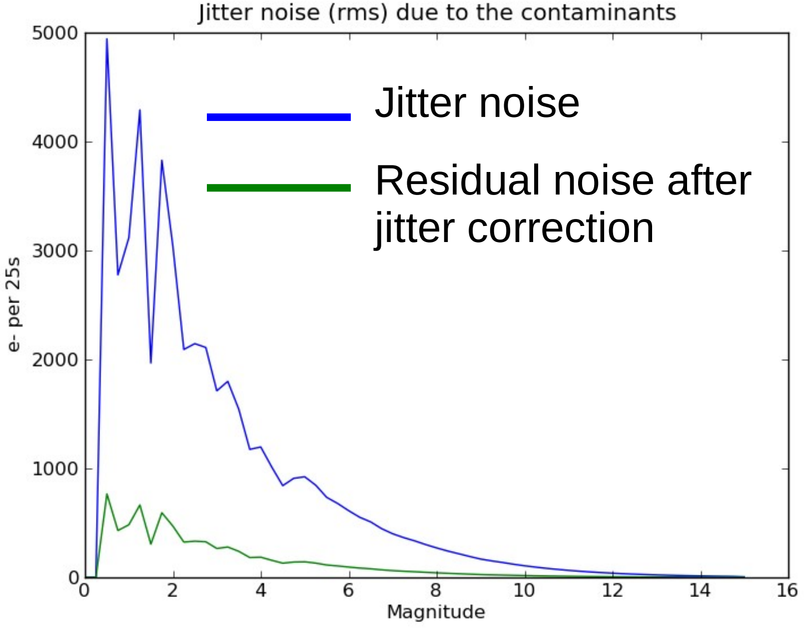
PSF 0°

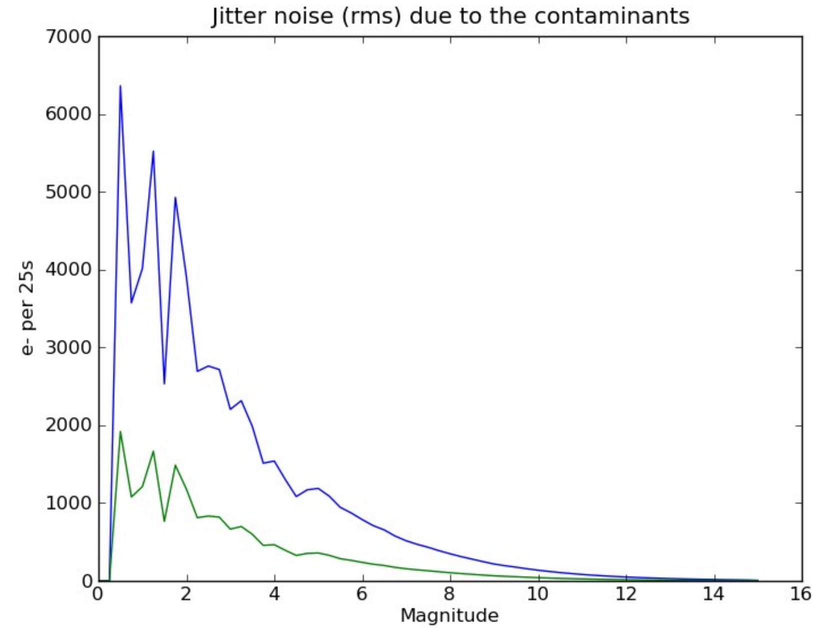




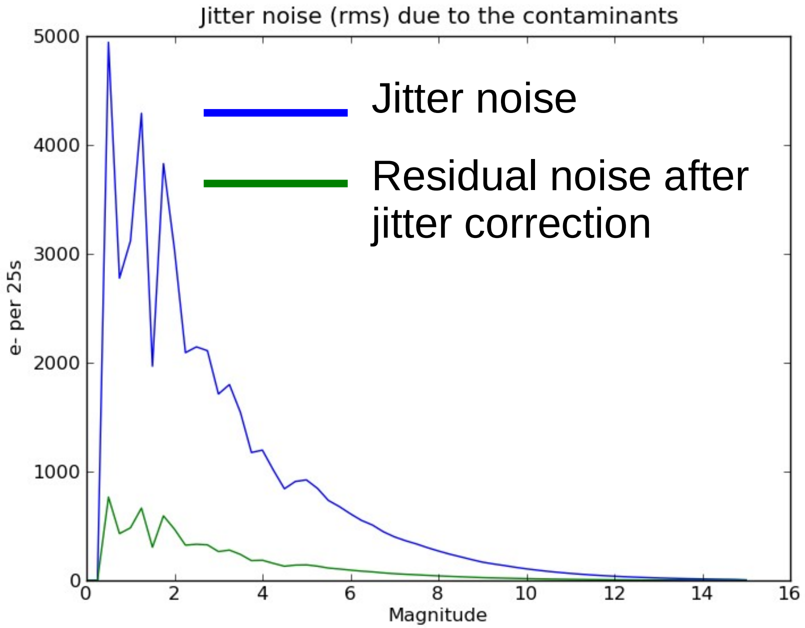


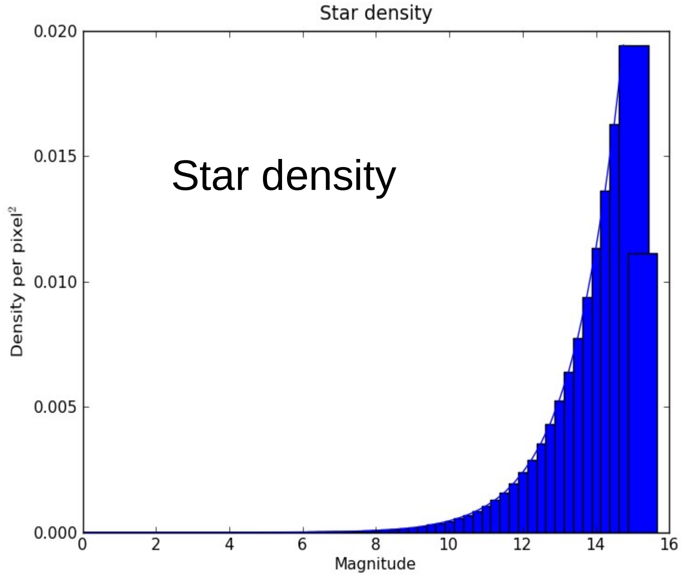
Photon noise

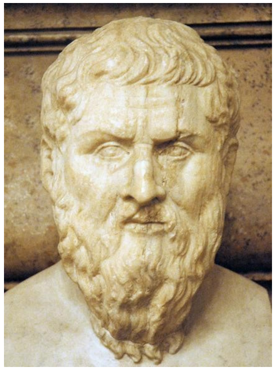


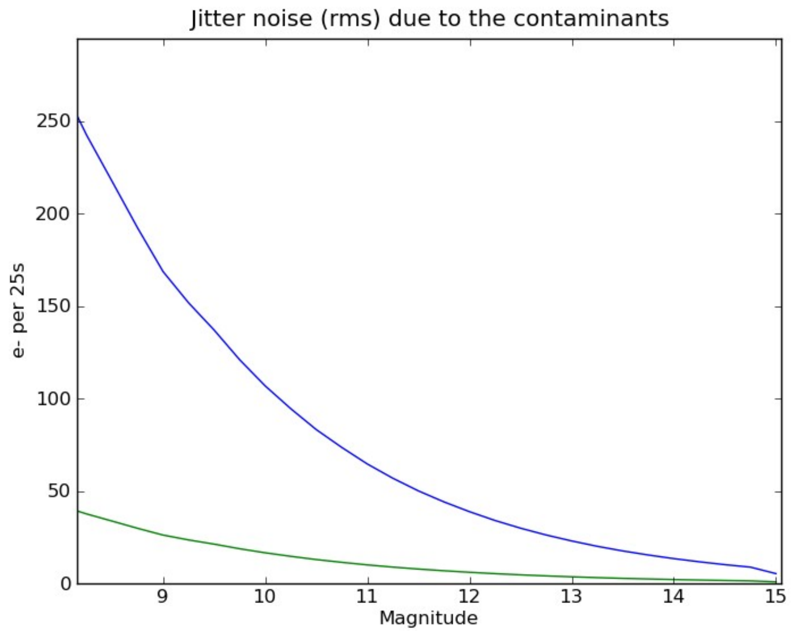


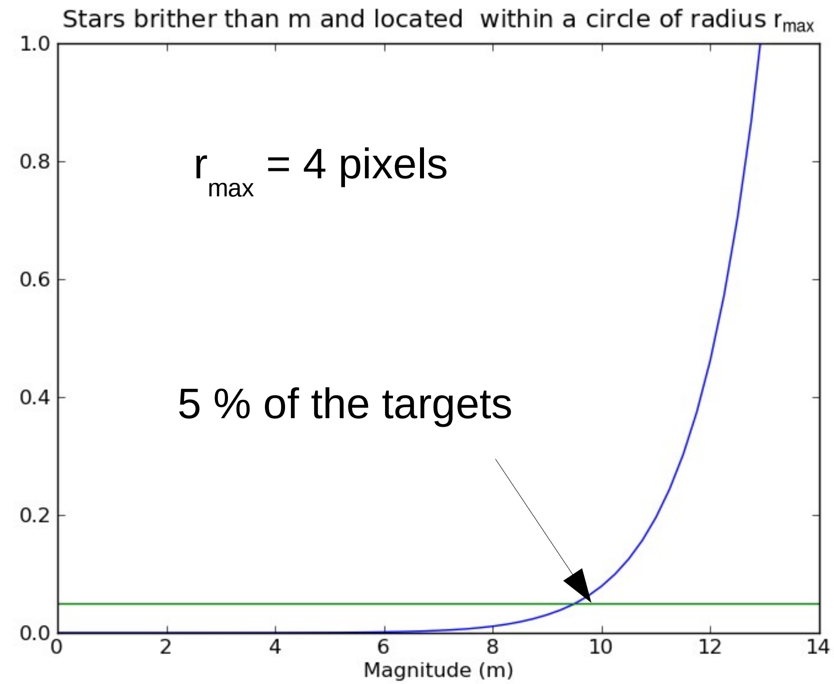
Jitter noise


Jitter noise (rms) due to the contaminants
Jitter noise
Residual noise after jitter correction
e- per 25s
Magnitude

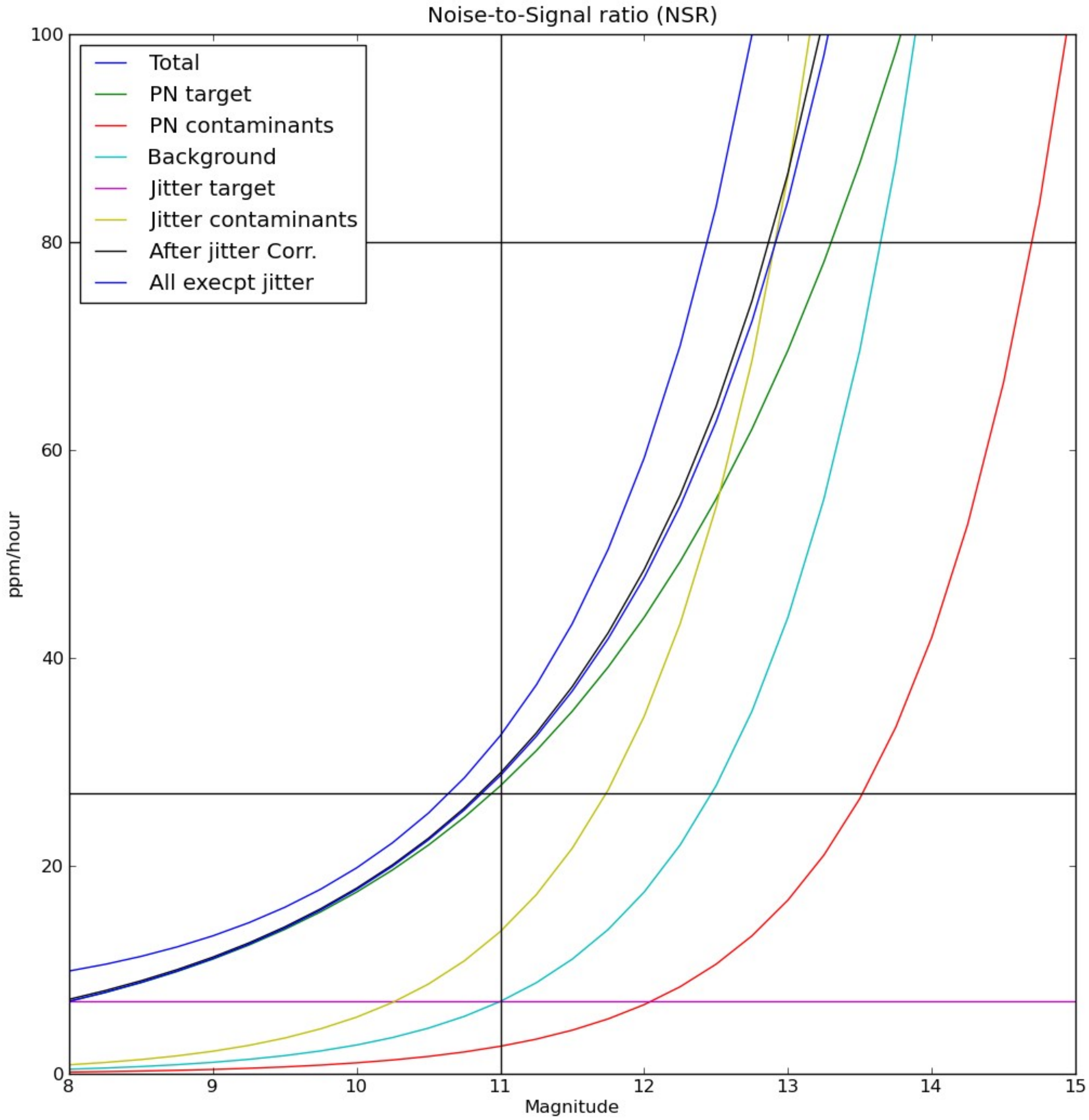

Star density
Star density
Density per pixel²
Magnitude

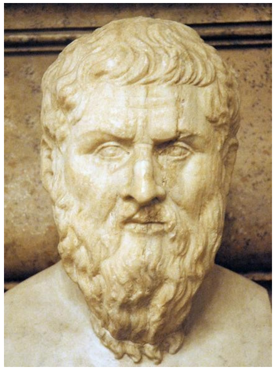



Jitter noise (rms) due to the contaminants
e- per 25s
Magnitude


Stars brither than m and located within a circle of radius r_max
r_max = 4 pixels
5 % of the targets
Magnitude (m)

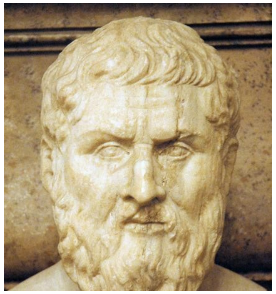
Global performances : results





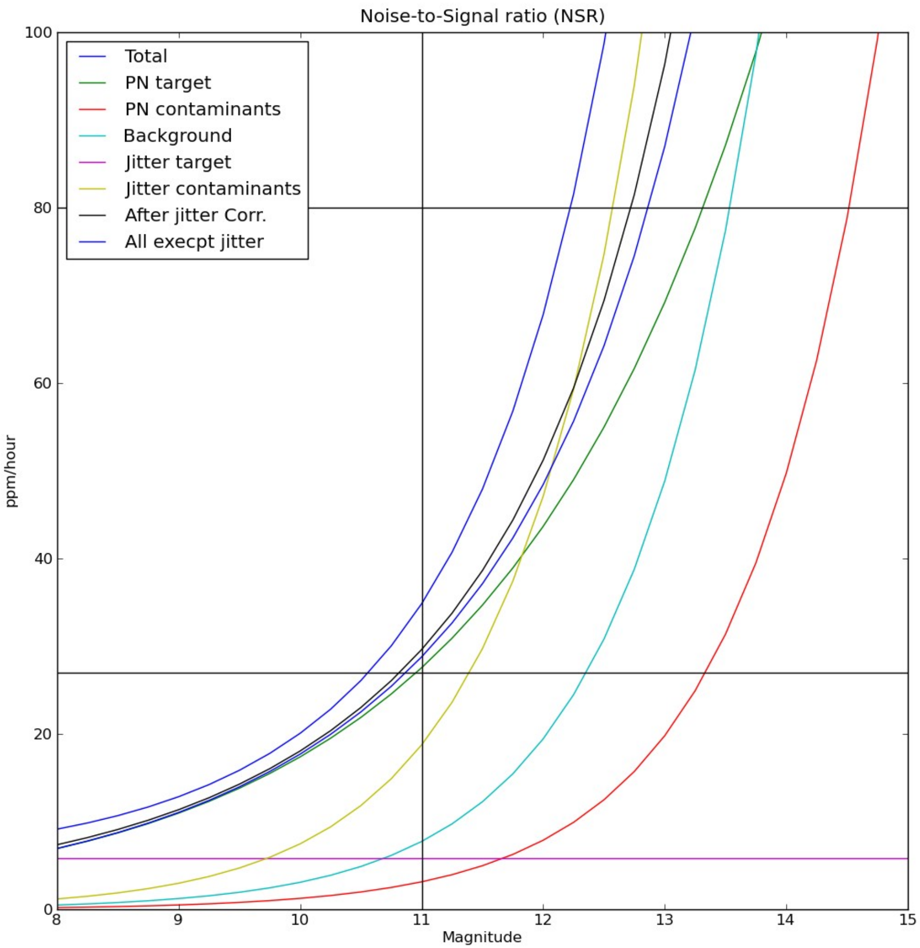
- Weighted mask (width: 1 pix)
- PSF 0° (center)
- 32 telescopes

Global performances : results

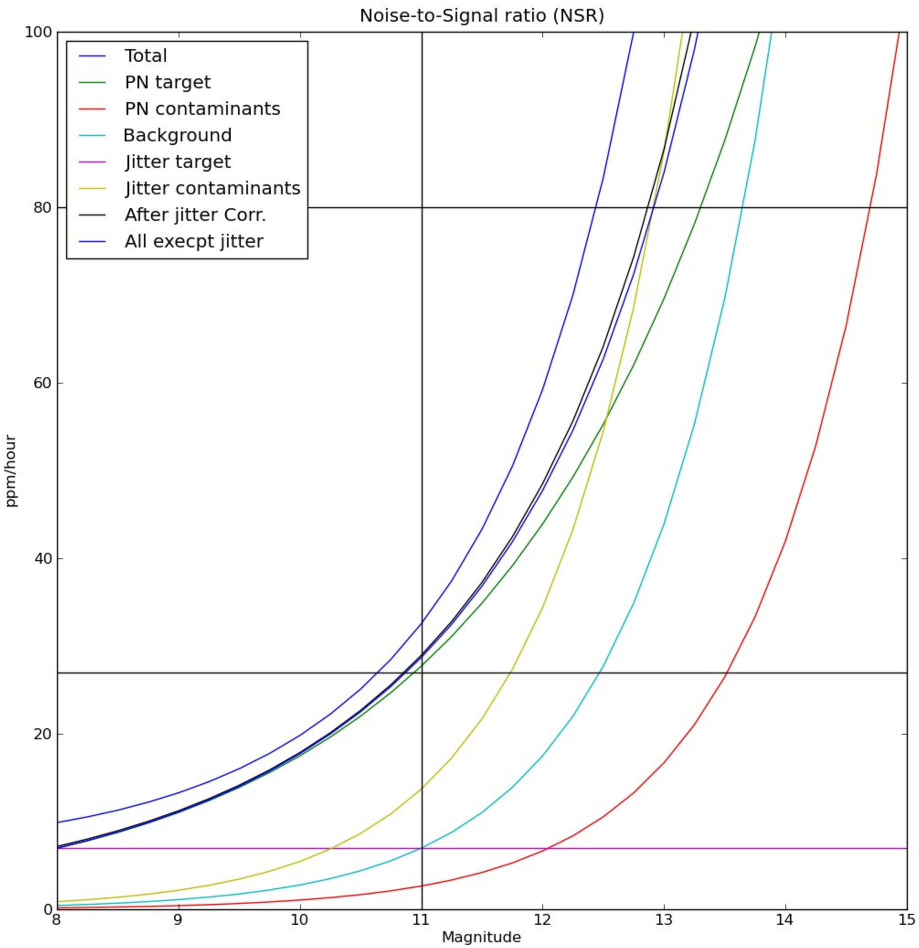


PSF 0°

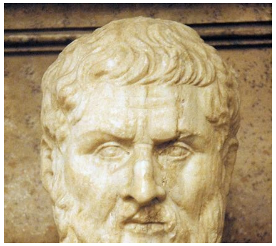
Binary mask



Weighted mask

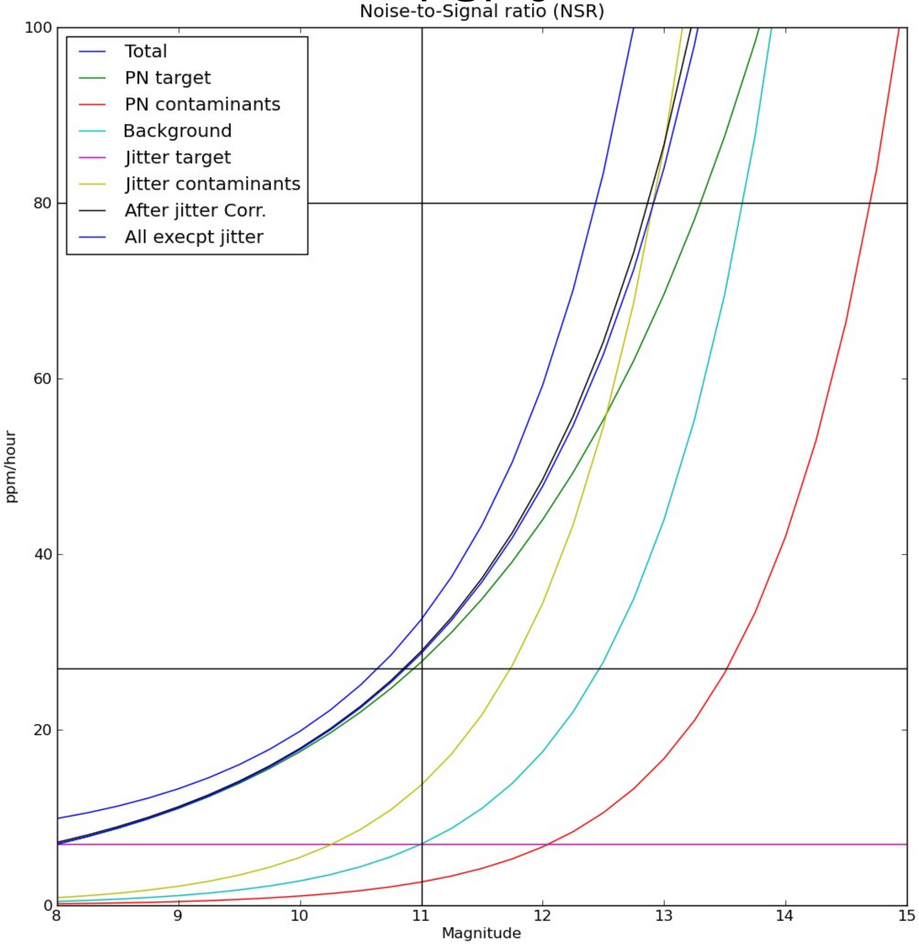


Global performances : results

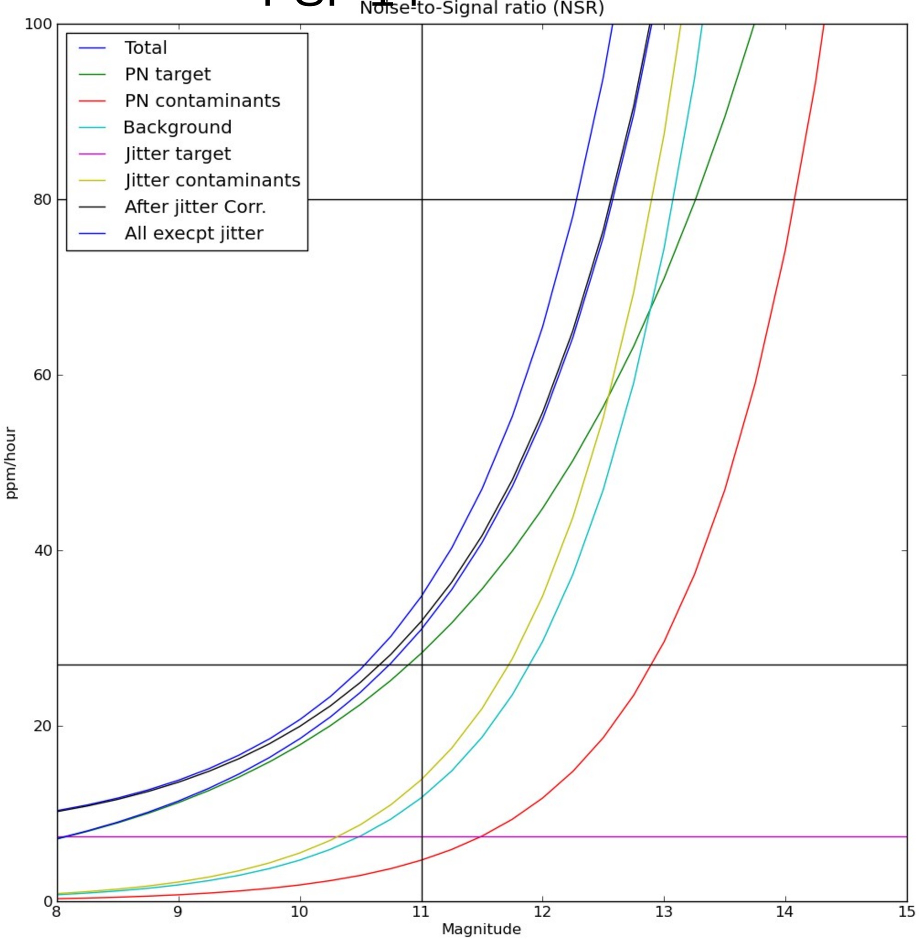


Weighted mask

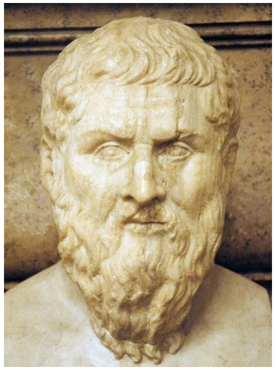
PSF 0°



PSF 14°

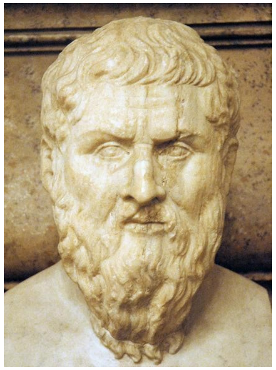


Global performances : conclusion



- Dominant contribution to the noise :
  - Below mag. 8 : jitter noise associated with the target
  - Between mag. 8 – 12 : photon noise of the target
  - Above mag. 12 : jitter noise associated with the contaminants
- Performances slightly degraded in the edge of the field of view
- In all cases, best performances with the weighted mask
- (on ground) jitter correction is in any cases required

The configuration mode

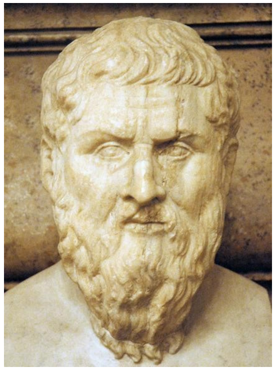


- The observation sequence can started as soon as the windows and the masks are attributed and the background estimated

Requirements:

- Recognition of the field of view and identification of the targets
- For each star :
  - Determine initial position of the centroid
  - Derive a representative PSF
  - Derivation of the initial parameters of the LSF
- Calibration of the background model

Reconstitution of the PSF across the field



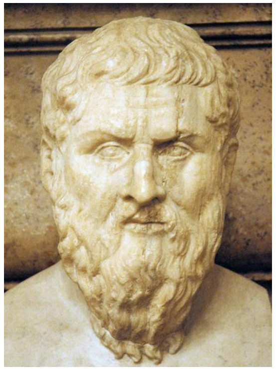
<u>Assumptions, for each telescopes</u> :

- The PSF varies slowly across the field of view
- We have available N (=1600) reference stars with associated image time series (**n** images)
- We have a <u>functional form</u> of the PSF as a function of K parameters a_i (eg. center x_0 and y_0, width σ , skewness ... etc)

Illustrative case of a *Gaussian* PSF:

$$PSF(x,y) = A \exp\left[-\frac{1}{2}\left(\left(\frac{x-x_0}{\sigma_x}\right)^2 + \left(\frac{y-y_0}{\sigma_y}\right)^2\right)\right]$$

Reconstitution of the PSF across the field



Illustrative case of a *Gaussian* PSF:

$$PSF(x,y) = A \exp\left[-\frac{1}{2}\left(\left(\frac{x-x_0}{\sigma_x}\right)^2 + \left(\frac{y-y_0}{\sigma_y}\right)^2\right)\right]$$

Step #1: For each reference stars (~ 1600), for each telescopes:

- We constrain the parameters using the imagettes time-series.

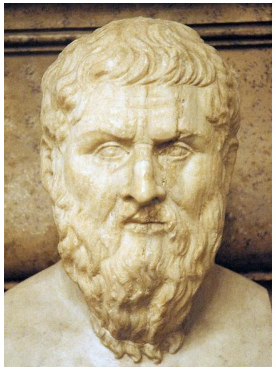
The fitted parameters a_i (*j*) (e.g. width σ , skewness ... etc) are then considered as a function of the position [x_0(*j*) and y_0(*j*)] of the star *j*.

Step #2:

- A 2D polynomial interpolation is then performed to derive the values of the parameters at any position across the field of the telescope

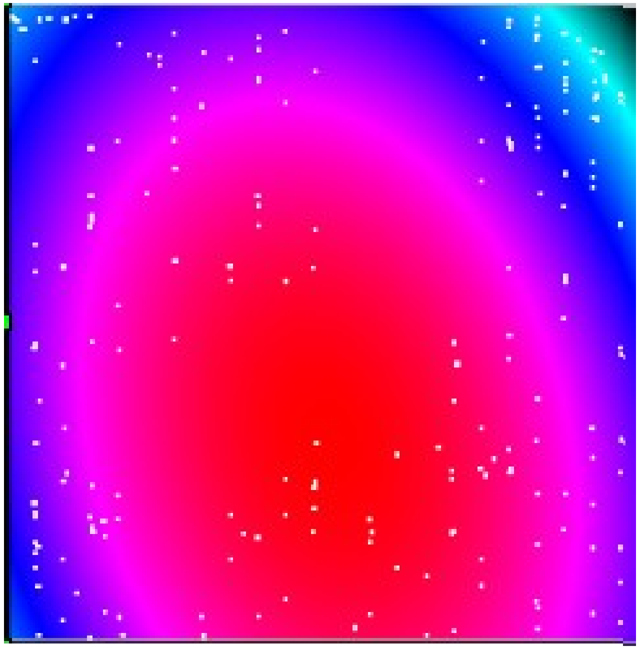
PSF can depend on the color of the star ⇒ 3D polynomial interpolation w.r.t. the color of the star

Modeling the sky background



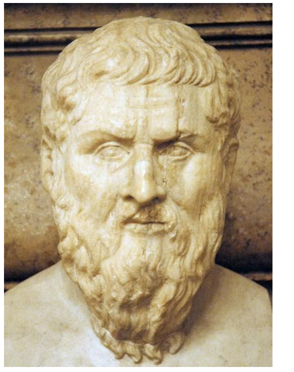
- We set ~ 400 background windows per telescope (100 per CCD)

- **During the configuration mode**
  - we collect a long enough time series of background measurements
  - We model te background using a 2D polynomial fit



- The sky background level can then be estimated at any position, then for any target

Organization and planning at the system level

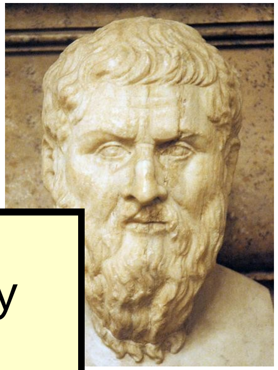


- Phase A: *until June 2011*
  - Specifications and development ⇒ sharing between board and ground
  - Implementation (Python or IDL)

- Phase B1: *from June to December 2011*
  - Optimization
  - Implementation (in C++) within ***PLATOsim*** (= PLATO simulator)

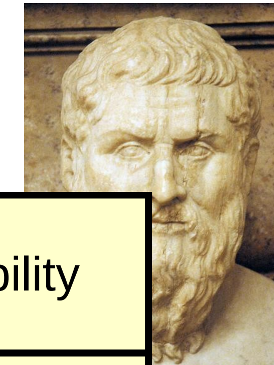
Works split into 14 work packages

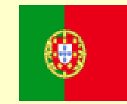
Work packages at the system level



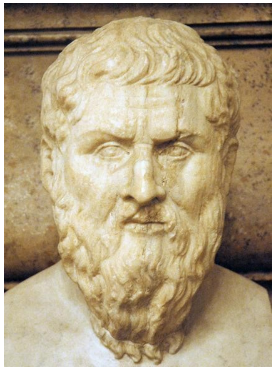
WP #	Contain	Responsibility
1	Time series of simulated images (using ***PLATOsim***)	LESIA
2 & 3	Modeling the PSF across the field of view	LESIA
4 & 5	Mask based photometry (weighted or binary)	LESIA
6	LSF based photmetry (LSF = 1D PSF fitting)	LESIA
7	Determiniation of the star centroids	FCUL 
8	Modeling the sky bacground	LESIA

Work packages at the system level (continue)



WP #	Contain	Responsibility
9	Field recoginition and determination of the Line Of Sight	FCUL 
10	Time series of simulated light curve	LESIA
11	Determination of the star displacements (ground)	FCUL 
12	Jitter correction (ground)	Brésil / LESIA 
13	Correction of the discontinuities (board & ground)	LESIA
14	Correction of the outliers (e.g. proton or cosmic impacts) (board & ground)	LESIA

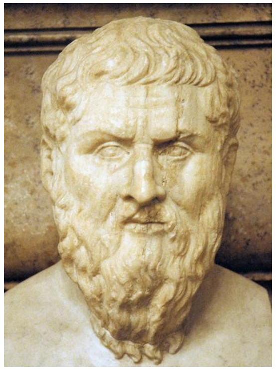
Data validation and treatments at the Ground Data Center level



Work packages :

- WP3: Pipeline, workflow management system
- WP4: Management of data flow, network
- WP5: Simulation of data stream
- WP6: Development of software for validation of L0 data
- WP7: Validation of L0 data (operational task)
- WP8: Development of software for the calibration of L1 data
- WP9: L1 Data processing

Data validation and treatments at the PLATO Data Center level (PDC)



WP5: Simulation of data stream

simulations of the data stream, from the telemetry to the end data products

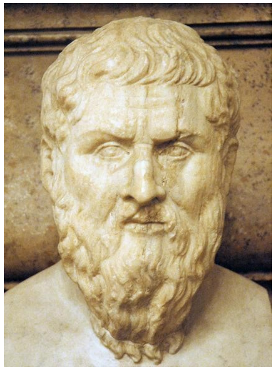
WP6: Development of the software for validation of L0 data

software to validate the L0 data, monitor the data quality and integrity, and provide support for the on board processing

WP8: Development of software for the calibration of L1 data

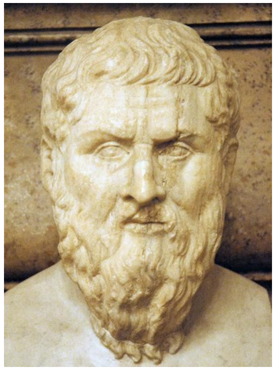
production of the flux-calibrated light curves and their averages (Level 1 data)

- WP5 : rely on **PLATOSIM**
- WP6 & WP8 : rely on the work done at **the system level during the definition phase**



The on ground treatments

- Correction of the (residual) differential aberration and satellite jitter
- Integration time correction
- Sampling time correction (including heliocentric correction)
- Long term detrenting
- Detection and removal of the outliers (eg. Cosmics rays, hot pixels)
- Treatment of the imagettes:
  - Offset, smearing (trailing) and background subtraction
  - Photometry (PSF fitting or mask based)
  - Jitter correction (if mask based photometry)



Crucial open questions:

- Jitter correction : **crucial** for the performances. The **efficient** of the correction **must be demonstrated** → **WP 12** (resp. : Brazil / LESIA)

- Model for the PSF ? Resolution required for the jitter correction ? → **WP 2 & 3** (resp. : LESIA)

PDC activities in support of the SOC

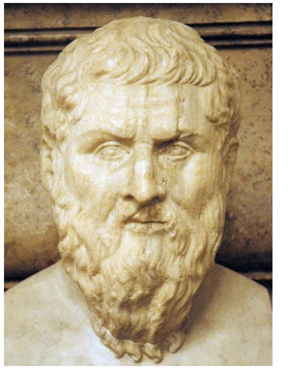
- Implementation and test of the data algorithms defined at the system level

- Study and define the treatments that are not taken into account at the system level (e.g. long term detrenting, time correction, calibration, treatment of the imagettes ...)

- .... ?

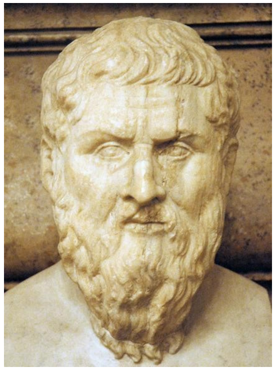
➢ All these activities: must be undertaken in close collaboration with the persons in charge of WPs at the system level

➢ Interfaces and responsibilities must be, in term, clearly be defined

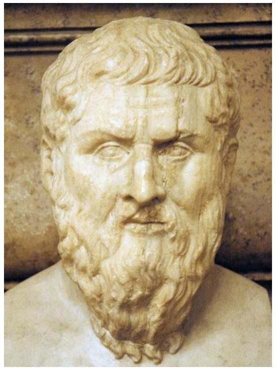


Present documentation:

- Assessment phase PPLC design report (FDR)
- PLATO data processing algorithms (appendix to the FDR)
- PLATO Normal telescope DPU data processing and hardware assessment report (appendix to the FDR)
- PLATO definition phase: Data processing work packages
- Alternative concept



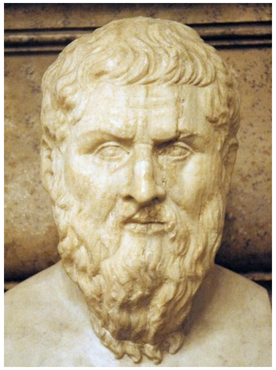
- Pointing performances ? Level and nature of the jitter noise ? → we have set our requirements on the AOCS
- Jitter correction : **crucial** for the performances. The **efficient** of the correction **must be demonstrated → WP 12** (resp. : Brazil / LESIA)
- Model for the PSF ? → **WP 2 & 3** (resp. : LESIA)
  - Resolution required for the jitter correction ?
  - Resolution required for the calculation of the weighted mask ?
- Photometry of the saturated stars ? Down to which magnitude ?
- Calculation of the barycenter : thresholding ? simple mask ? Weighted mask ? → **WP 7** (resp. : Portugal)



Gaussian PSF

Method	Noise level (ppm/1h)
Width (pix)	0.9
Sub-pixel resolution (1/pix)	64
Window size	6x6
Binary mask	30.4
Binary mask + jitter correction	31.7
Weighted mask (width in pix)	29.4 (w1.5)
Weighted mask + jitter correction	**29.8 (w1.5)**
LSF - Gauss	29.8
LSF - PSF	33.7
Fit 2D PSF	36.7

Method	Noise level (ppm/1h)
Width (pix)	1.8
Sub-pixel resolution (1/pix)	64
Window size	8x8
Binary mask	37.5
Binary mask + jitter correction	37.7
Weighted mask (width in pix)	36.7 (w3.0)
Weighted mask + jitter correction	36.9 (w3.0)
LSF - Gauss	39
LSF - PSF	40.6



Numerical PSF from the optic model

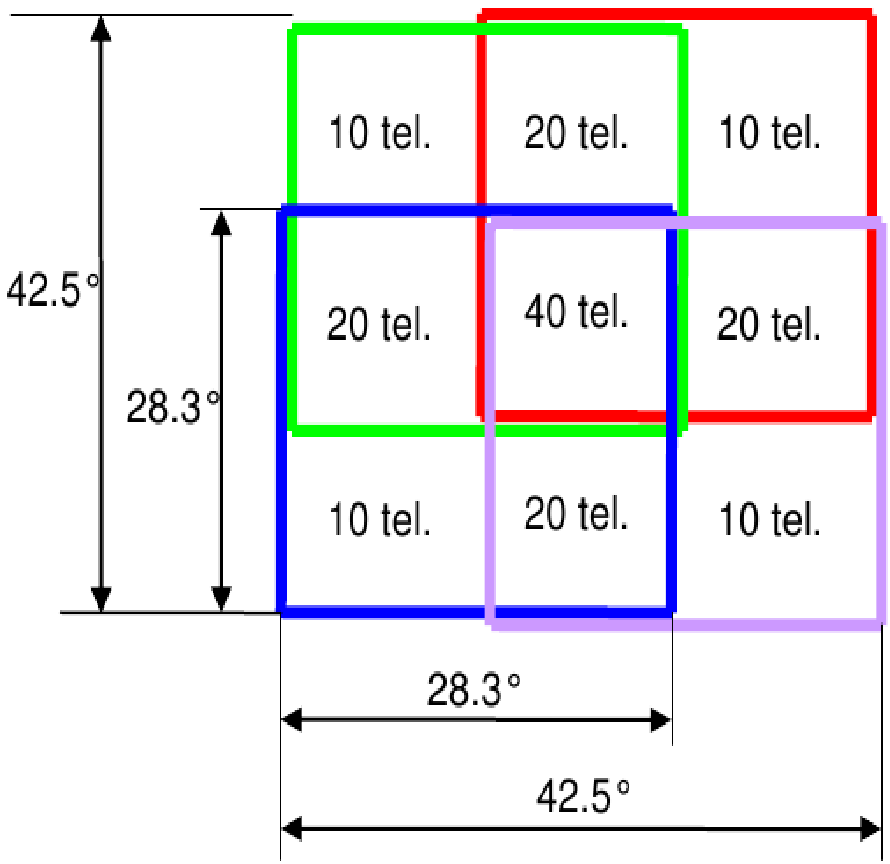
Nominal background level: 150 e/pix/s

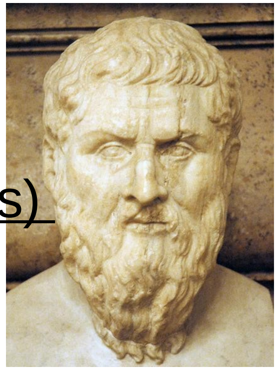
Method	Noise level (ppm/1h)
Sub-pixel resolution (1/pix)	64
Window size	6x6
Binary mask	32.7
Binary mask + jitter correction	32.5
Weighted mask (w=width in pix)	32.4 (w1.5)
Weighted mask + jitter correction	**32.2 (w1.5)**
LSF - Gauss	*33.6*
LSF - PSF	*36.7*
Fit 2D PSF	*39.9*

Low background level: 15 e/pix/s

Method	Noise level (ppm/1h)
Sub-pixel resolution (1/pix)	64
Window size	6x6
Binary mask	31
Binary mask + jitter correction	31
Weighted mask (w=width in pix)	31
Weighted mask + jitter correction	31
LSF - Gauss	*30.8*
LSF - PSF	*34.6*

The new field of view





Requirements: (cool dwarfs)

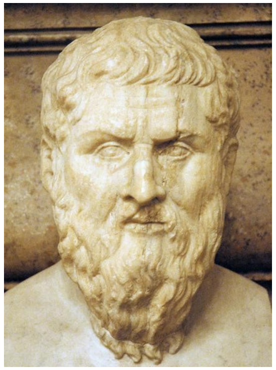
Sample P1 : 10 000 stars
Sample P2 : 40 000 stars
Sample P3 : 1 000 stars
Sample P4 : 2 000 stars
Sample P5 : 125 000 stars

(star count per pointing)

Noise level (in 1h)	Tel. number	Fraction of the FOV	Limit Mag.	Stars
27 ppm	10	4/9	9.60	2,450
27 ppm	20	4/9	10.40	6,400
27 ppm	40	1/9	11.15	3,600
			Total:	12,450
80 ppm	10	4/9	12.00	34,000
80 ppm	20	4/9	12.80	80,000
80 ppm	40	1/9	13.50	46,000
			Total:	160,000

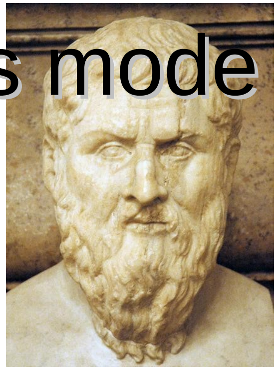
Sample P1 : we are not 100% sure
they are all cool dwarfs
we double their number
10 000 stars -> 20 000 stars

Star samples (per telescope, per pointing)



- Sample P1 : mv < 9.6 - 11.15 ; noise level < 27 ppm/h
  - 10 000 stars : photometry @ 50s , centroids @ 600 s
  - Subset : N = 1000 references stars, mv= 8.6-9.6, individual light curve
  - Sub-images (imagettes) : **n** = 400 stars @ 25 s sampling
- Sample P2 : mv < 12 ;   noise level < 80 ppm/h
  - 20 000 stars @ 600s
  - Oversampled : 400 stars @ 50s sampling
- Sample P3 (P4) : 4.75 < mv < 7.3   noise level < 27 ppm/h
  - 500 (1 000) stars @ 50s
  - Subset: 100 stars centroids @ 2.5 s
  - Sub-images (imagettes) : m = 100 @ 50 s
- Sample P5 : mv < 13.5 ; noise level 80 ppm/h ; no centroids measured
  - 80 000 stars @ 600s
  - Oversampled : 1000 stars @ 50s with
- Background windows : 400

The onboard processing : the observations mode



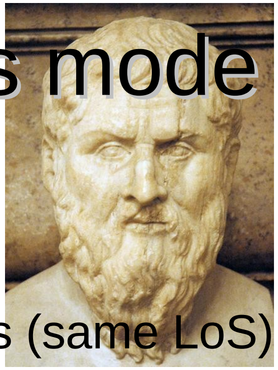
Normal DPU : at 25s sampling (P1,P2,P5)

- At each 25s:
- Smearing subtraction
- Background subtraction
- Update the mask position (TBS)
- Apply the mask and compute the flux
  - **Weighted** mask for **mv>9** and **aperture** mask (binary mask) for **mv<9**
- For samples P1-P4 : We compute the star barycenter (TBS)
- Correction of the jitter and differential aberration (TBC and TBS)
- Update the mask position (TBS)
- Transmit data to ICU

Note:

- Need for the correction of the jitter and differential aberration must be confirmed
- To be done on board ***if*** the individual light-curves are not downloaded

The onboard processing : the observations mode



ICU : at 25s sampling (P1,P2,P5)

- Gain correction
- Compute the median and the standard deviation associated with N telescopes (same LoS)
- Detect the outliers using the median and the standard deviation
- Compute the mean flux of the *k* valid measurements ($k <= N$)
- Stack the flux/centroid:
  - Up to 2 values stacked for flux with 50 s sampling
  - Up to 24 values stacked for flux & centroids with 600 s sampling

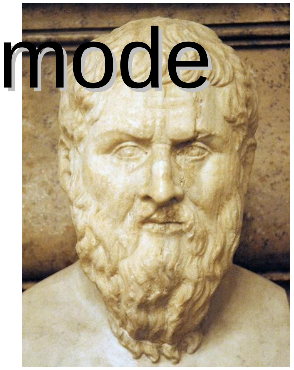
ICU : at 50s sampling (P1,P2)

- Compute the mean and standard deviation of the *p* ($p<=2$) valid stacked measurements
- Temporary bufferization
- Compress the data, send the data to SVM

ICU : at 600s sampling (P1,P2,P5)

- Compute the mean and standard deviation of the *p* ($p<=24$) valid stacked measurements
- Temporary bufferization
- Compress the data, send the data to SVM

The onboard processing : the observations mode



Fast DPU : at 2.5 s sampling (P3,P4)

- Smearing subtraction
- Background subtraction
- Apply the mask (binary mask) and compute the flux
- Compute the star barycenter (TBS)
- Compute angle error using the centroids of n=100 references stars
- Update mask position
- Transmit data (flux, barycenter positions) to ICU
- Transmit data (angle error) to VSM

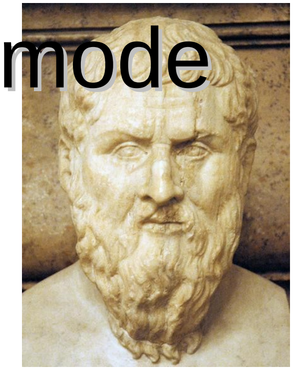
ICU: at 2.5s sampling (for each fast-telescope)

- Stack the flux/centroids of 20 measurements (50 s sampling)

ICU: at 50s sampling (for each fast-telescope)

- Compute median and standard deviation associated with the 20 last measurements
- Detect the outliers using the median and the standard deviation
- Compute mean and standard deviation of the *k* valid measurements ($k <= 20$)
- Temporary bufferization
- Compress the data, send the data to SVM

The onboard processing : the observations mode



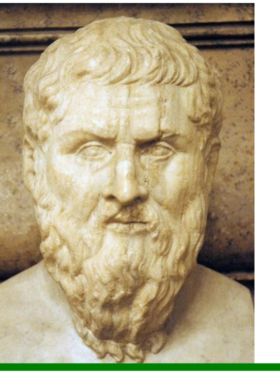
telemetry budget

- Case 1 : only 1000 LCs from Sample P1 are downloaded :
  - 31 Gb/days (with compression)
- Case 2 : all LCs are downloaded :
  - 71 Gb/days (with compression)

Predominant factor : the weight of the imagettes
Could possible to reduce by transmitting imagette accumulations at a lower cadence than 25 sec for the normal telescope and 2.5 sec for the fast telescope.

- Case 1 : correction of the jitter and differential aberration to be done onboard ! Puts strong constraints on the onboard software
- Case 2 : correction of the jitter and differential aberration can be done onground !

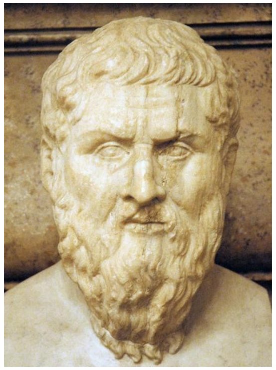
The configuration mode

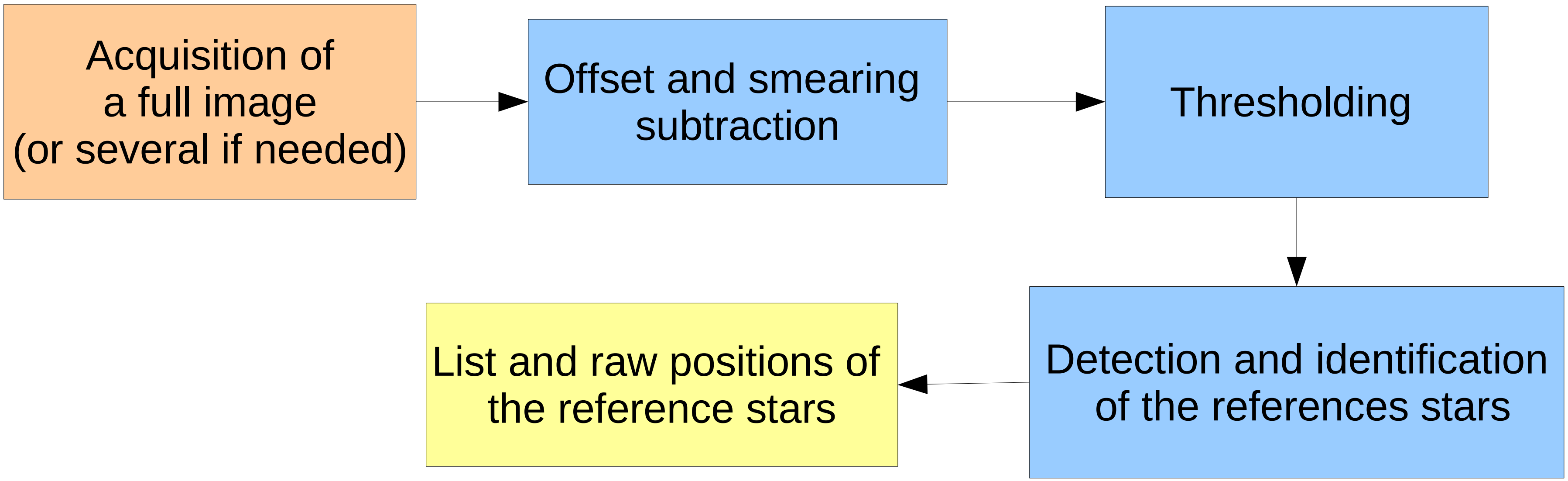


- The photometry mode can started as soon as the windows and the masks are attributed.
- Requirements:
  - Identify the stars
  - For each stars : derive a representative PSF .

The configuration mode: step 1

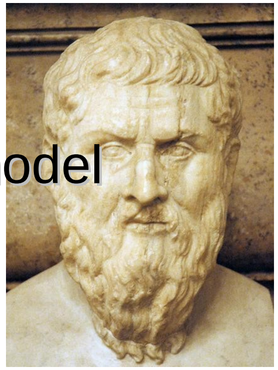
list and positions of the reference stars

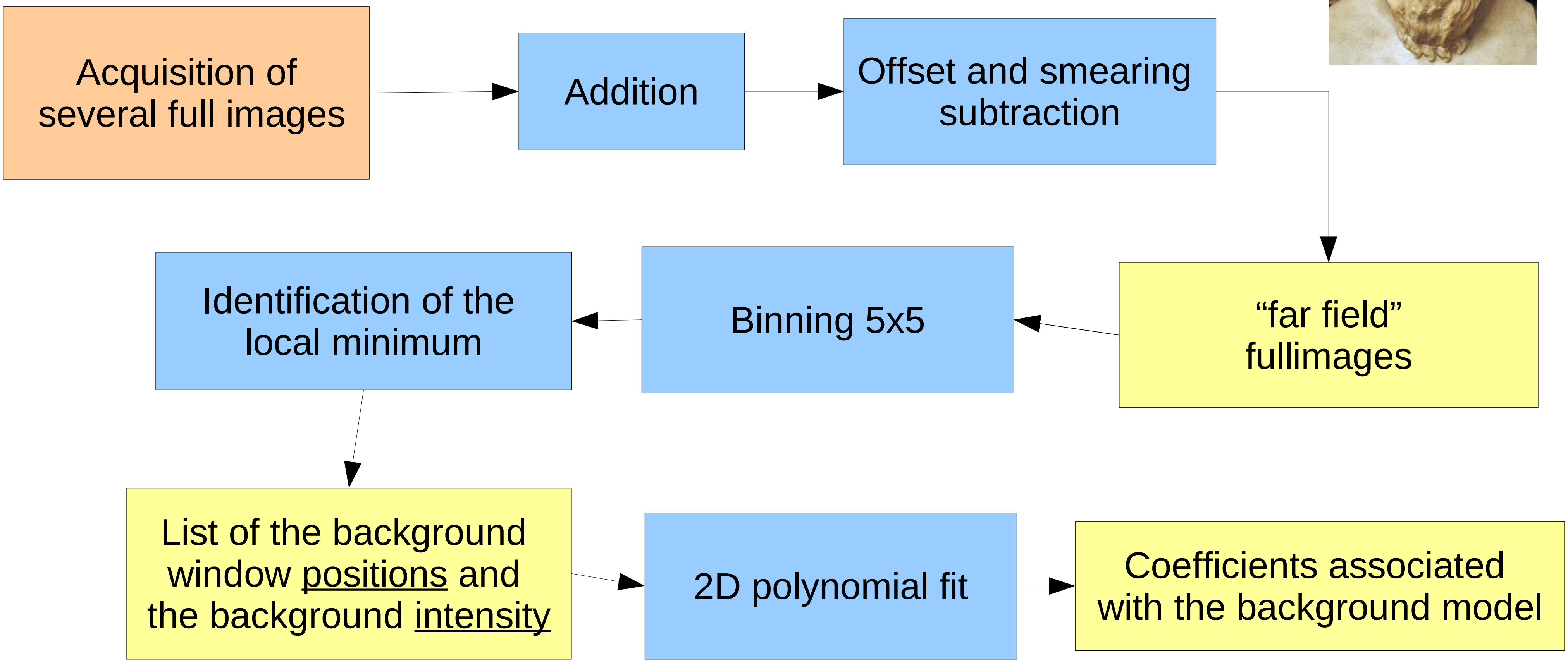




The configuration mode: step 2

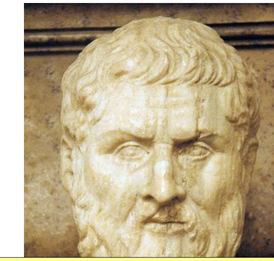
to set the background windows and derive the background model

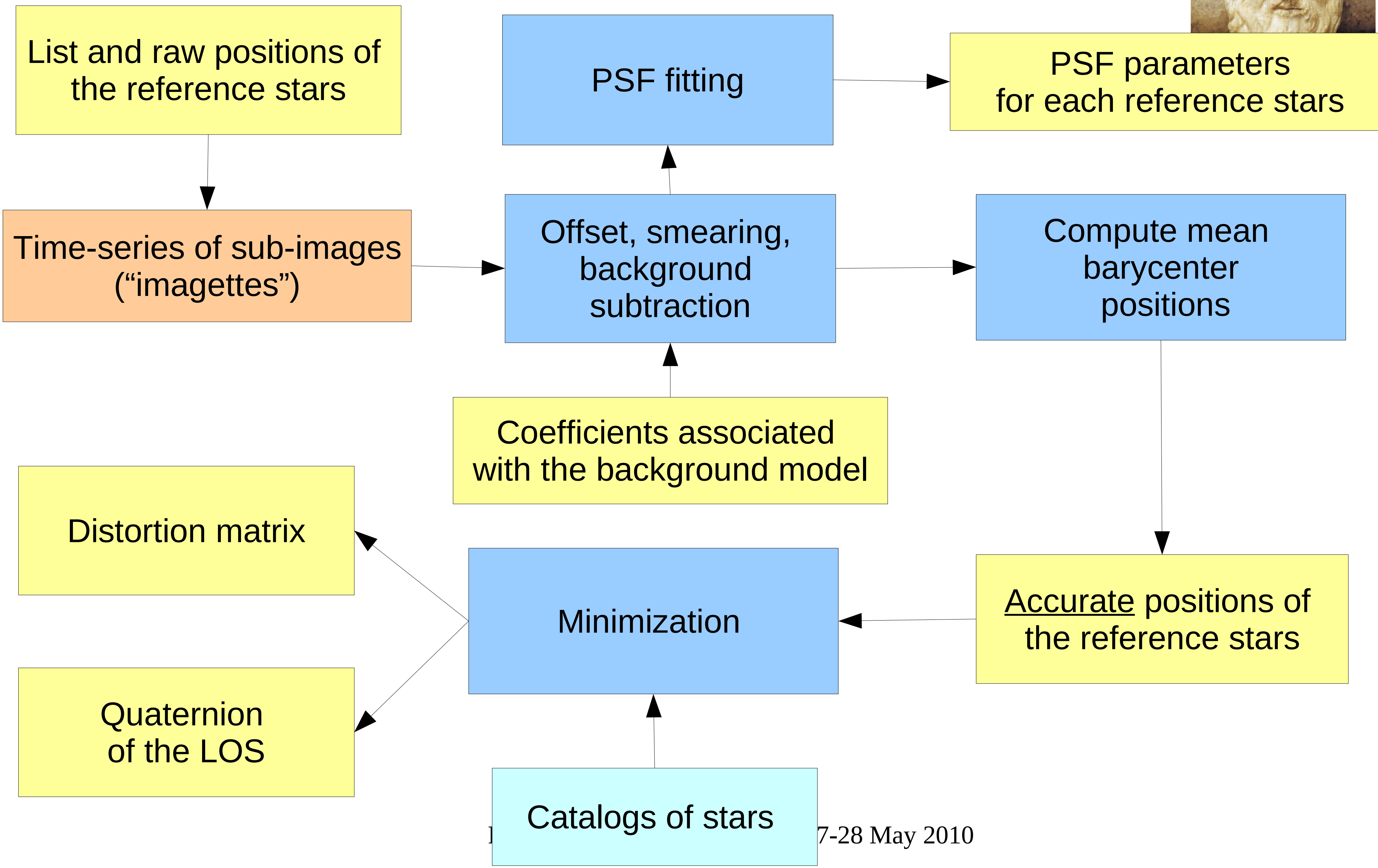




The configuration mode: step 3

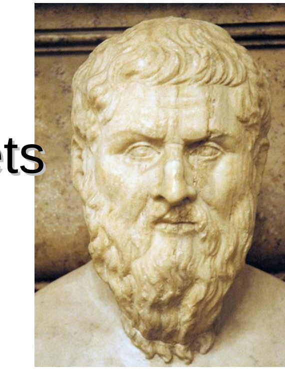
distortion matrix, PSF of the reference stars

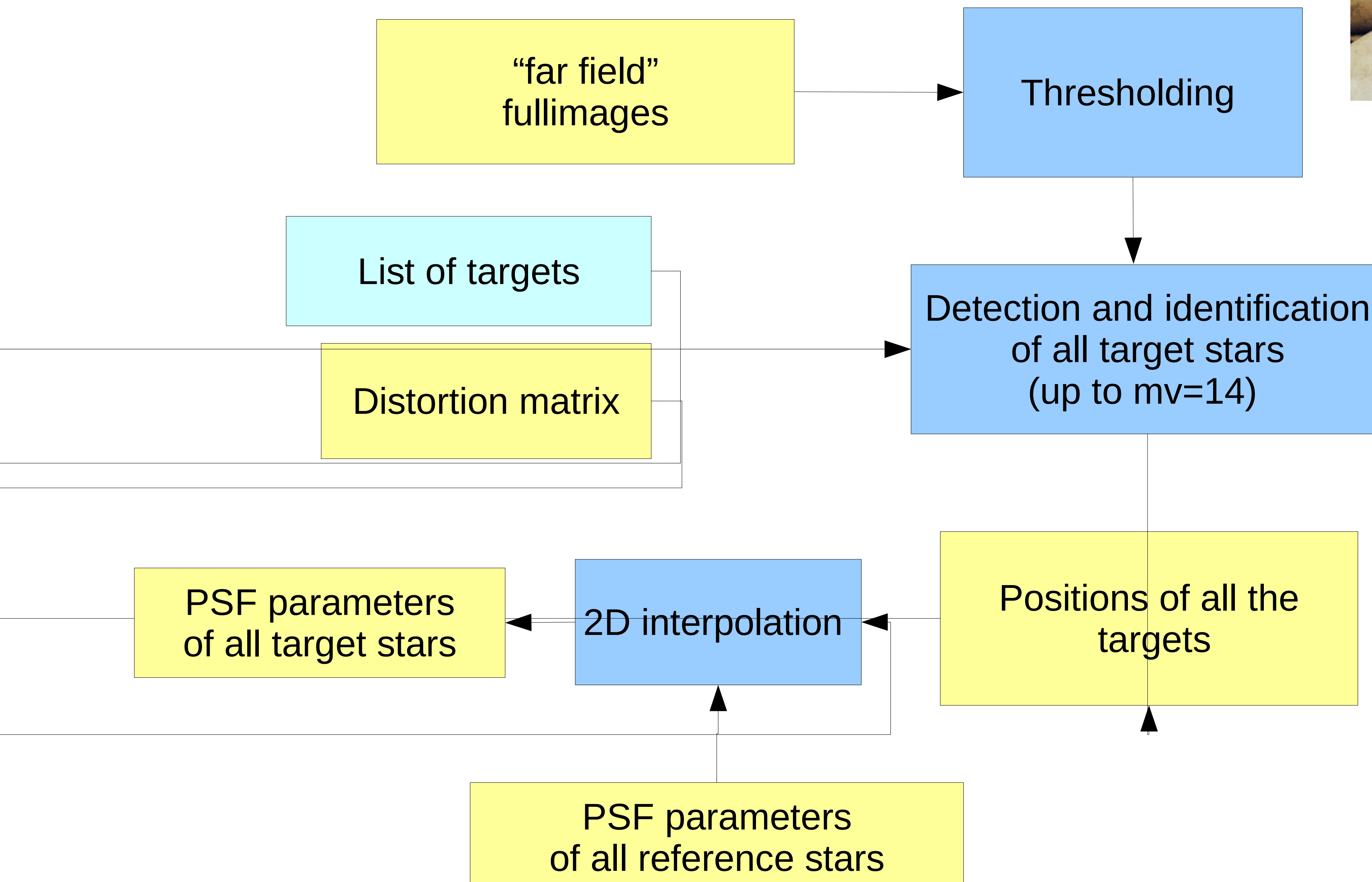




The configuration mode: step 4

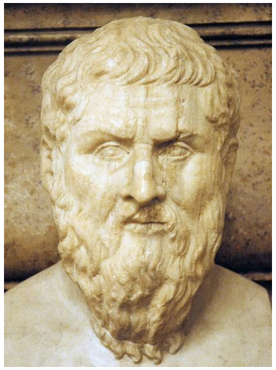
positions of all the targets, PSF parameter of all the targets

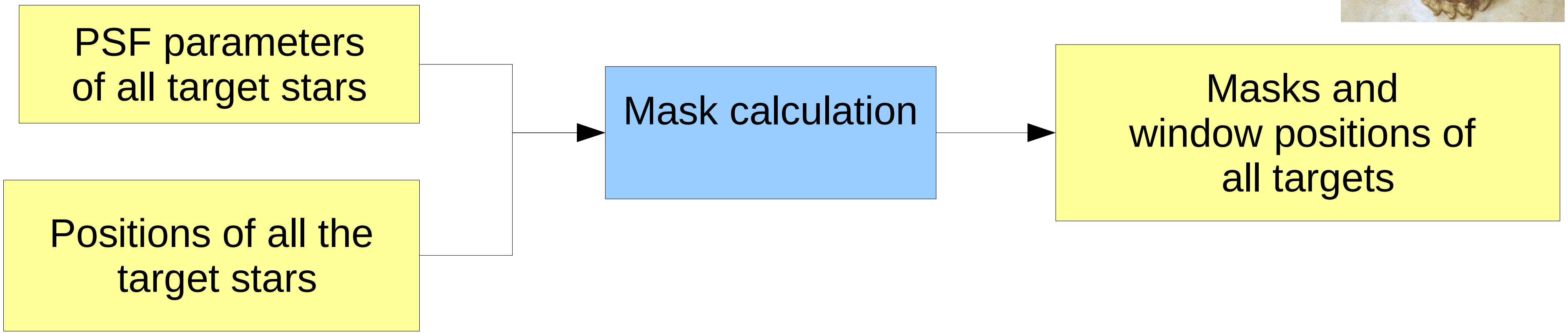




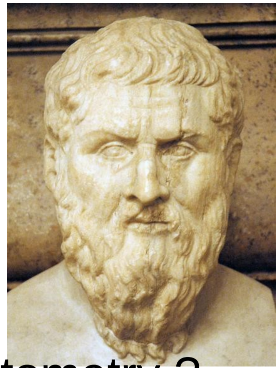
The configuration mode: step 5

masks and window positions of all the targets





Open questions



- Pointing performances ? Level and nature of the jitter noise ?
- Exact threshold in magnitude between weighted photometry and aperture photometry ?
- Model for the PSF ?
  - Resolution required for the jitter correction ?
  - Resolution required for the calculation of the weighted mask ?
- Photometry of the saturated stars ? Down to which magnitude ?
- Calculation of the barycenter : thresholding ? simple mask ?