ADASS XXVIII, 11-15 Nov 2018
Marco C Lam :: c.y.lam@ljmu.ac.uk

Data-Driven Pixelation with Voronoi Tessellation

Marco C Lam[1,2], Paul R McWhirter[1]

1 Liverpool Telescope, Astrophysics Research Institute, Liverpool JM University

2 Institute for Astronomy, University of Edinburgh

Outline

- Motivation: Pan-STARRS 1 3π Survey
- Motivation: Density of a Proper Motion Limited Sample
- Pixelation Schemes in Astronomy
- Potential Applications & Extensions

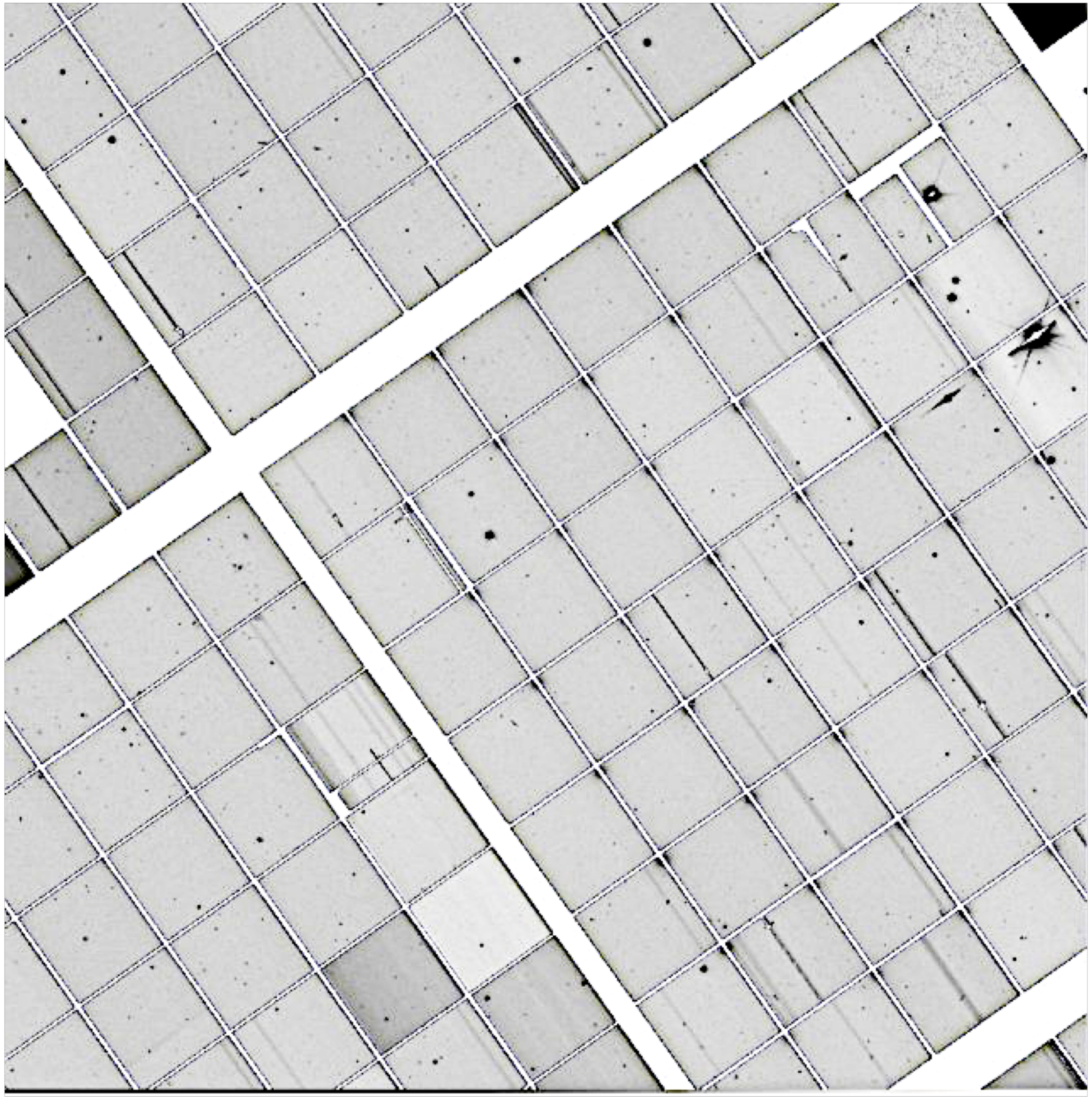
Pan-STARRS 1

- 1.8 m on Haleakala, Hawai'I
- 3π Survey (δ > -30º)
- 12 epochs in each of grizy
  (i.e. 60 epochs in total on average)
- 10^5 pointings over ~3.5 year
- 60 CCDs each with 64 regions

GCRF
Global Challenges Research Fund

Single Exposure

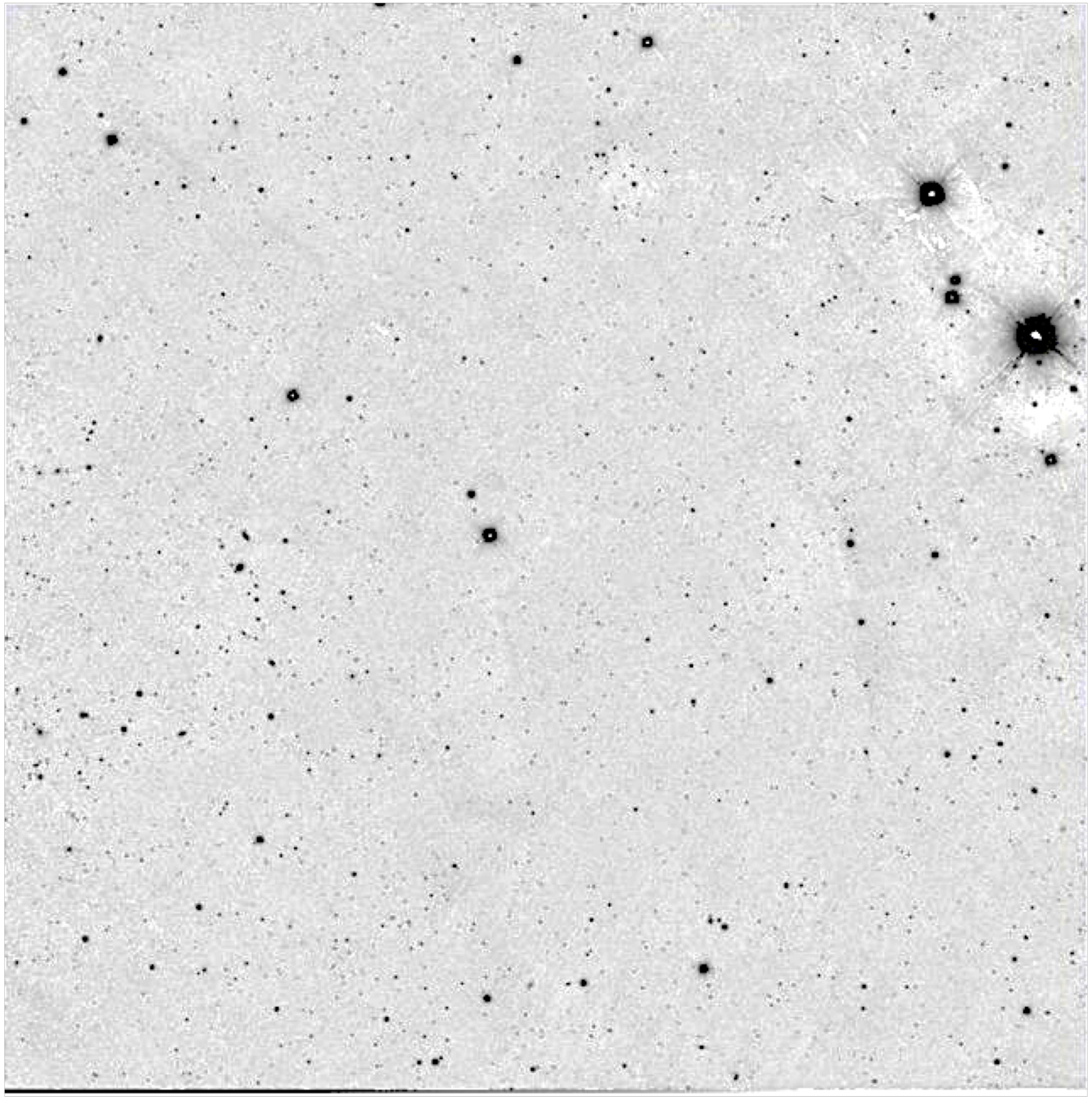
30s z-filter
skycell.2047.005
(α, δ) = (179.763, 32.1899)
Exposure o4985g0073o
2009-06-03



~0.5°

Figure 14 of Waters et al. 2016

Total Stack
(25 images)

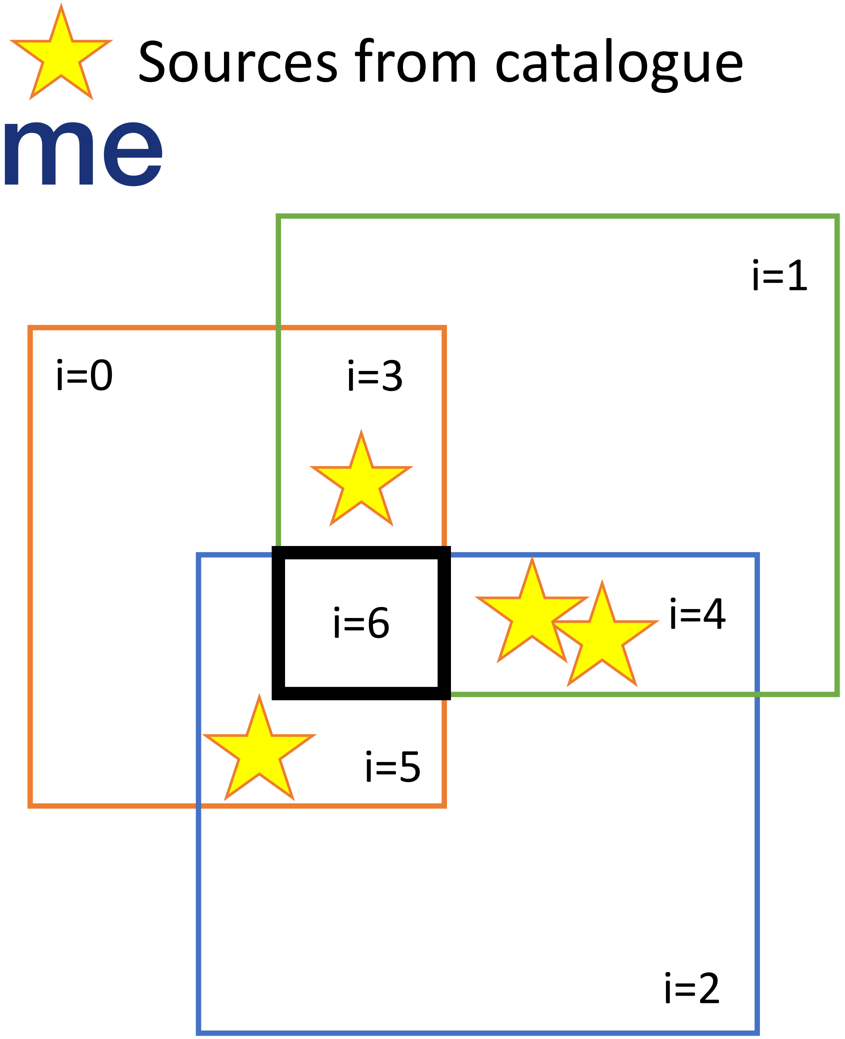


~0.5°

Motivations: Maximum Volume

- Maximum Volume : $V_{max} = \Omega \int_{D_{min}}^{D_{max}} \rho(r) r^2 dr$
- Dmin and Dmax found by testing the observability of a source

Motivations: Maximum Volume

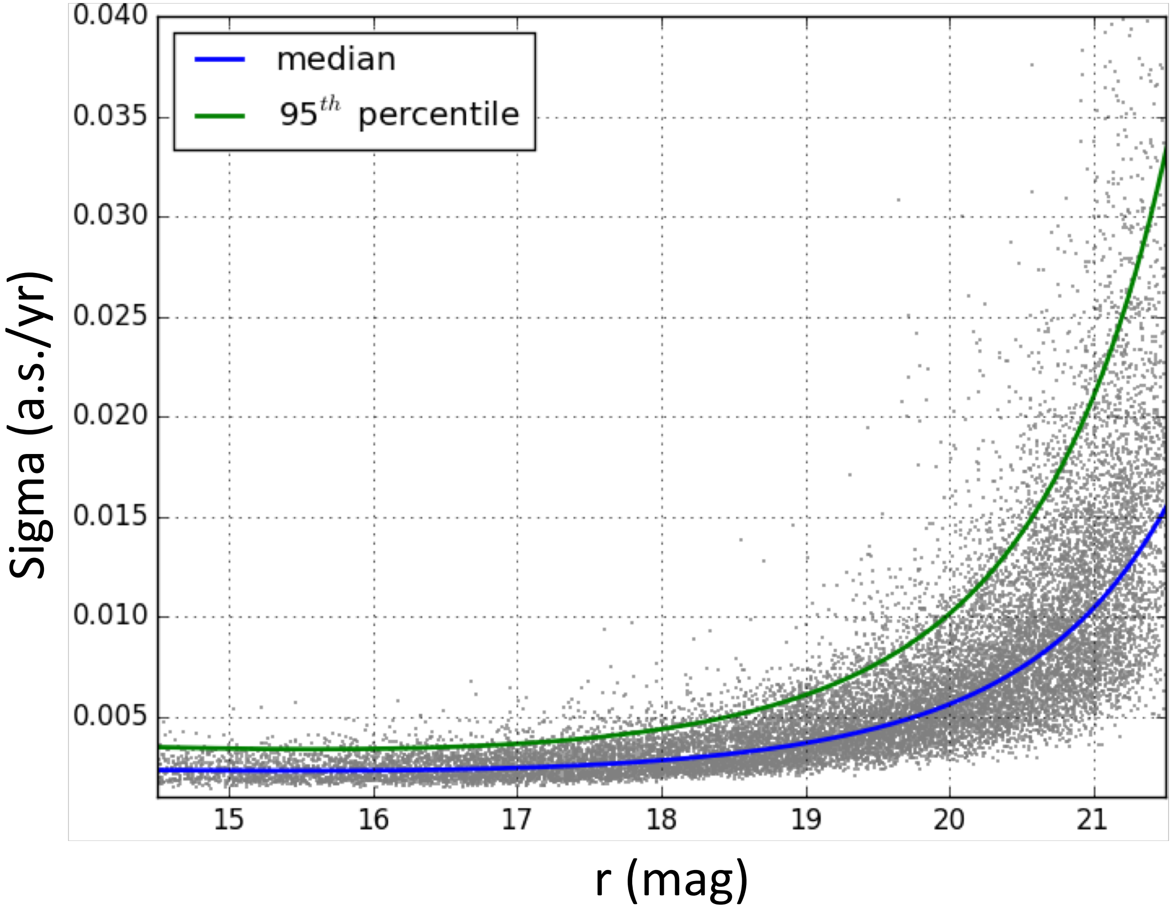


- Maximum Volume : $V_{max} = \Omega \int_{D_{min}}^{D_{max}} \rho(r) r^2 dr$
- Dmin and Dmax found by testing the observability of a source

- Density of source j : $\psi_j = \sum_i \left[\frac{1}{V_{max,i}}\right]$

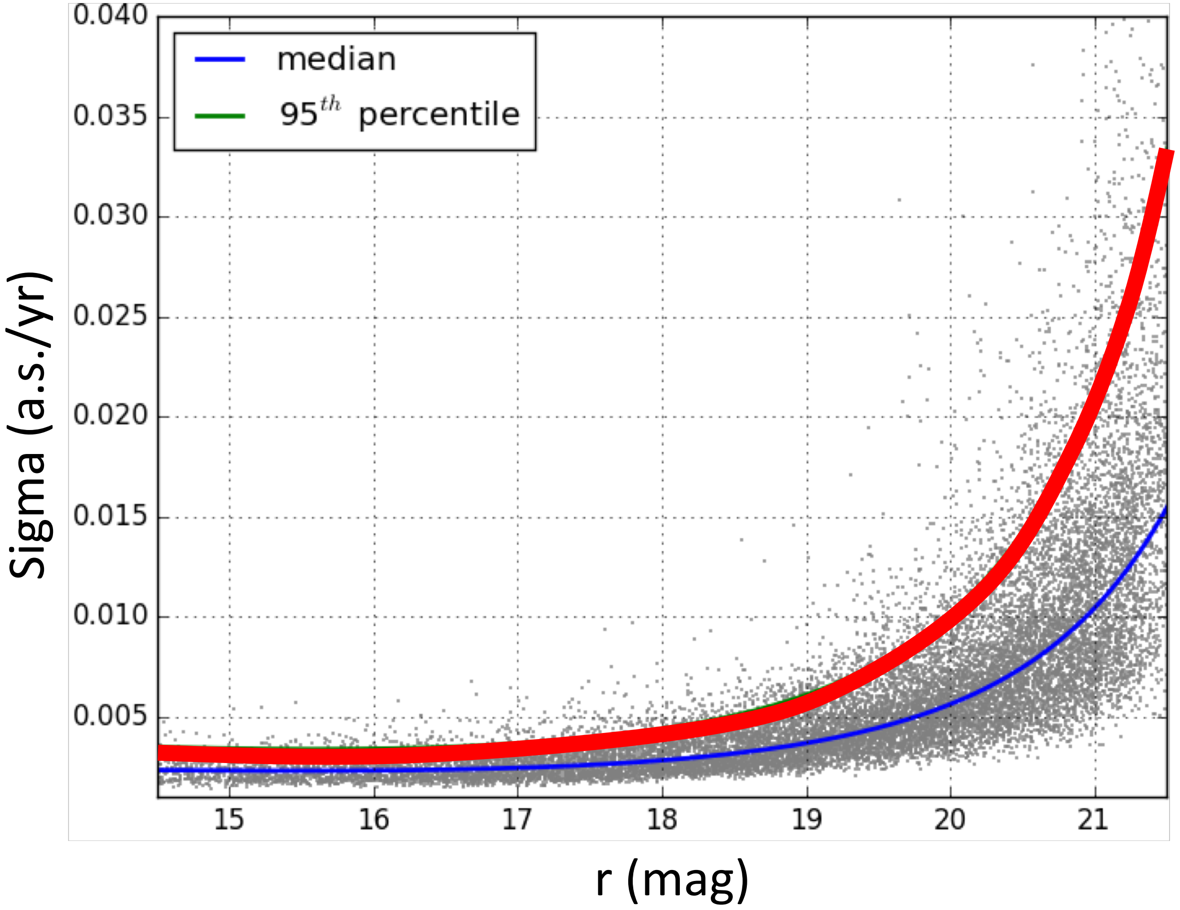
Motivations: Proper Motion Limit

- Scatter in PM uncertainties as a function of magnitude
- PM limits found from a large sample



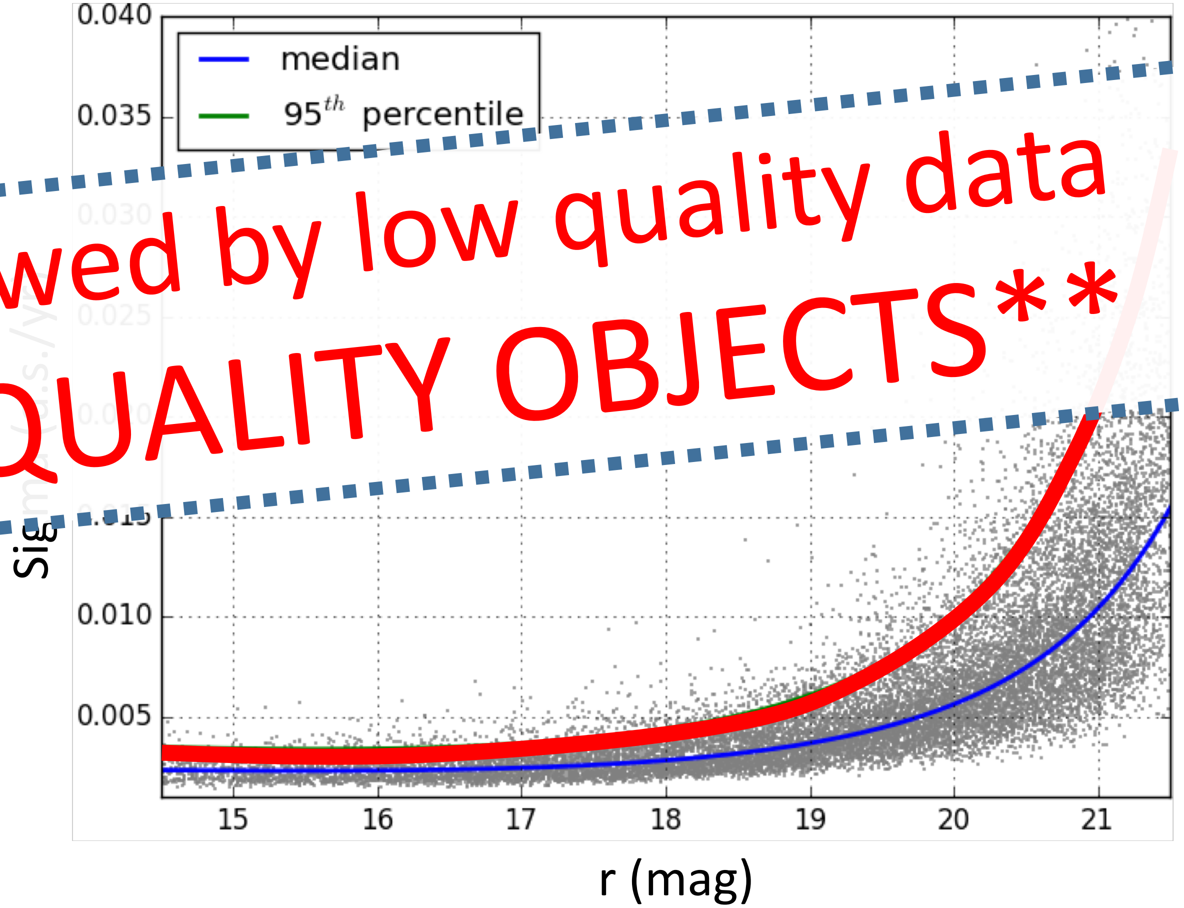
Motivations: Proper Motion Limit

- Scatter in PM uncertainties as a function of magnitude
- PM limits found from a large sample
- PM uncertainties are REASSIGNED to the PM limits



Motivations: Proper Motion Limit

- Scatter in PM uncertainties as a function of magnitude
- PM limits found from a large sample
- PM uncertainties are REASSIGNED to the PM limits



Selection Function skewed by low quality data

LOSING GOOD QUALITY OBJECTS

GCRF
Global Challenges Research Fund

Target Solution

- Re-register the survey
  - ➢ Relatable to the data or the survey
  - ➢ Low computing cost (hours/days of CPU time, not months/years)
  - ➢ Simple/Well established concept

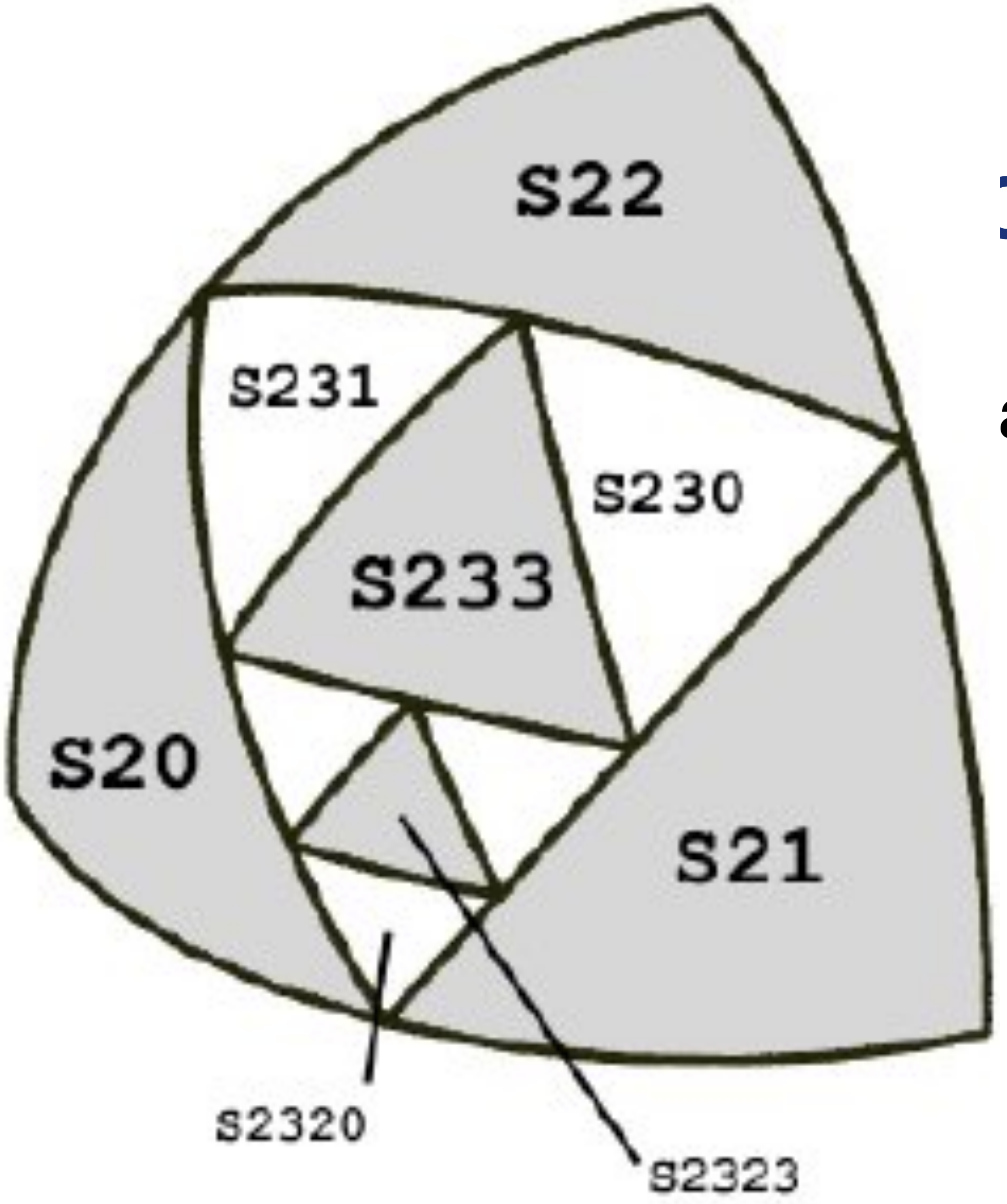
Global Challenges Research Fund

Pixelation Schemes in Astronomy

- HTM – Hierarchical Triangular Mesh (Kunszt et al. 2000)

Pixelation S… …omy

- HTM – Hierarchi… …al. 2000)



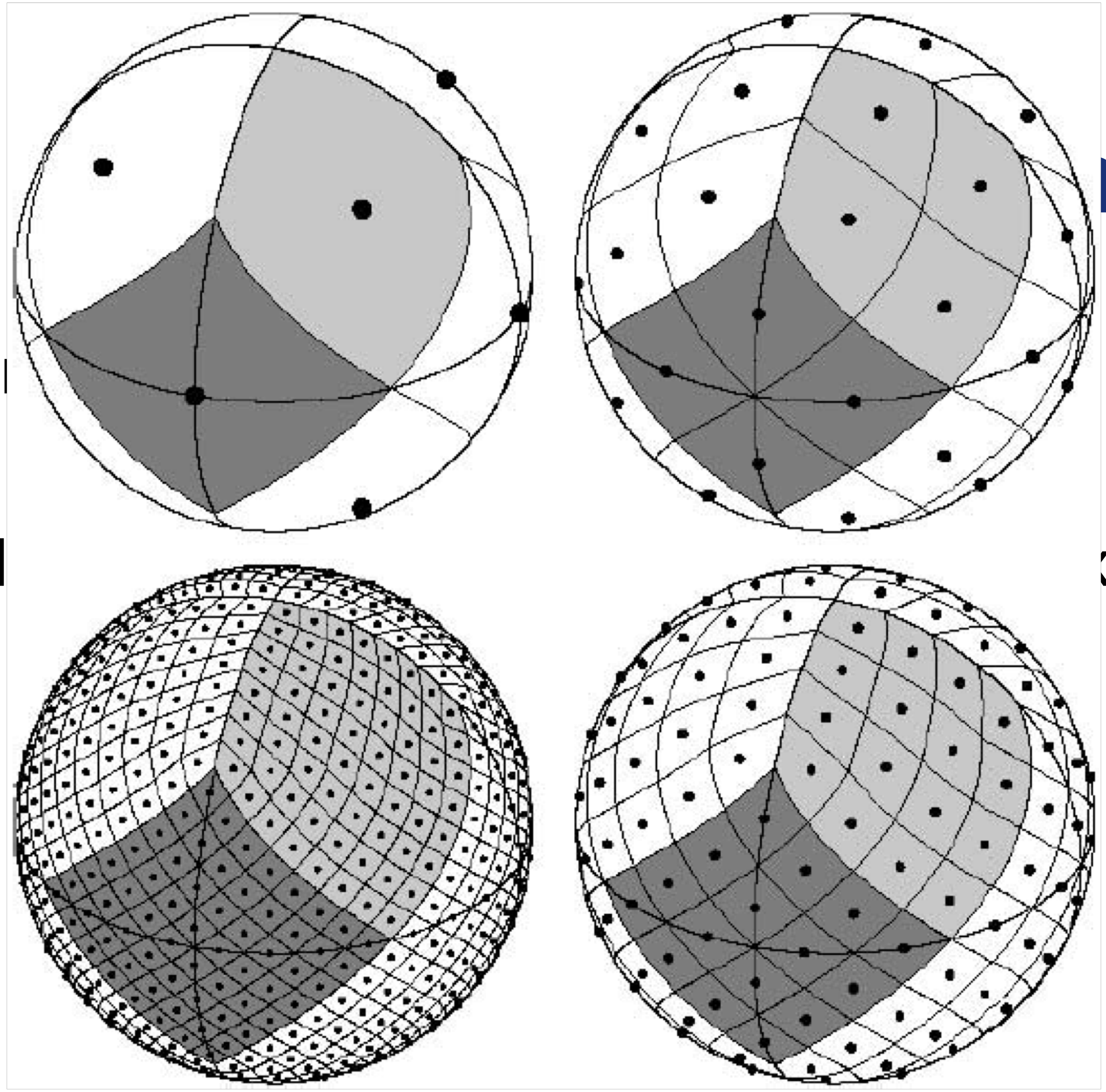
Pixelation Schemes in Astronomy

- HTM – Hierarchical Triangular Mesh (Kunszt et al. 2000)

- HEALPix – Hierarchical Equal Area isoLatitude Pixelisation (Gorski et al. 2005)

Pixelatio ny

- HTM – Hiera 2000)
- HEALPix – H elisation (Gorski et al.



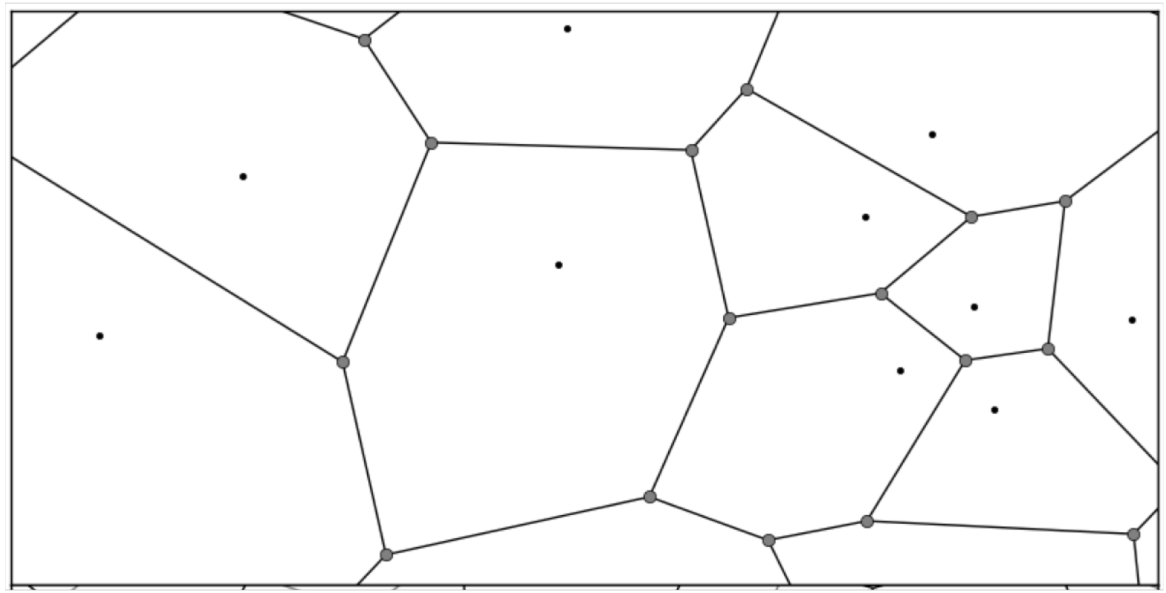
Global Challenges Research Fund

Pixelate the sky based on the Data

GCRF
Global Challenges Research Fund

Voronoi Tessellation

- Partitioning a plane with n points into n convex polygons
- Any position inside the polygon is closest to the governing point



Marco Lam :: c.y.lam@ljmu.ac.uk

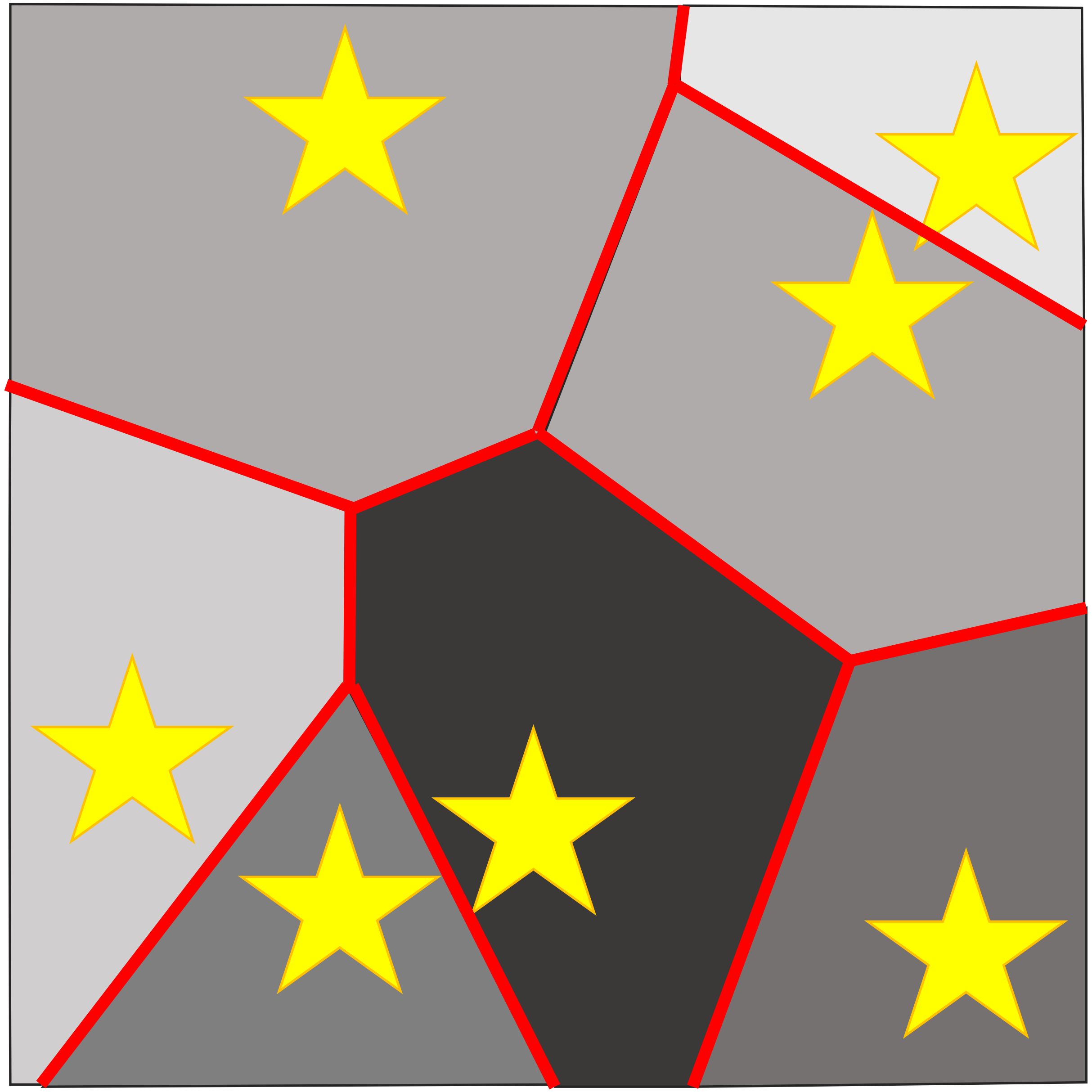
Global Challenges Research Fund

Voronoi Tessellation

- On a 2-sphere
- Scipy.spatial.SphericalVoronoi (scipy 0.18+)
- github.com/tylerjereddy/py_sphere_Voronoi

Marco Lam :: c.y.lam@ljmu.ac.uk

Global Challenges Research Fund



~0.5°

Consequence to the Maximum Volume

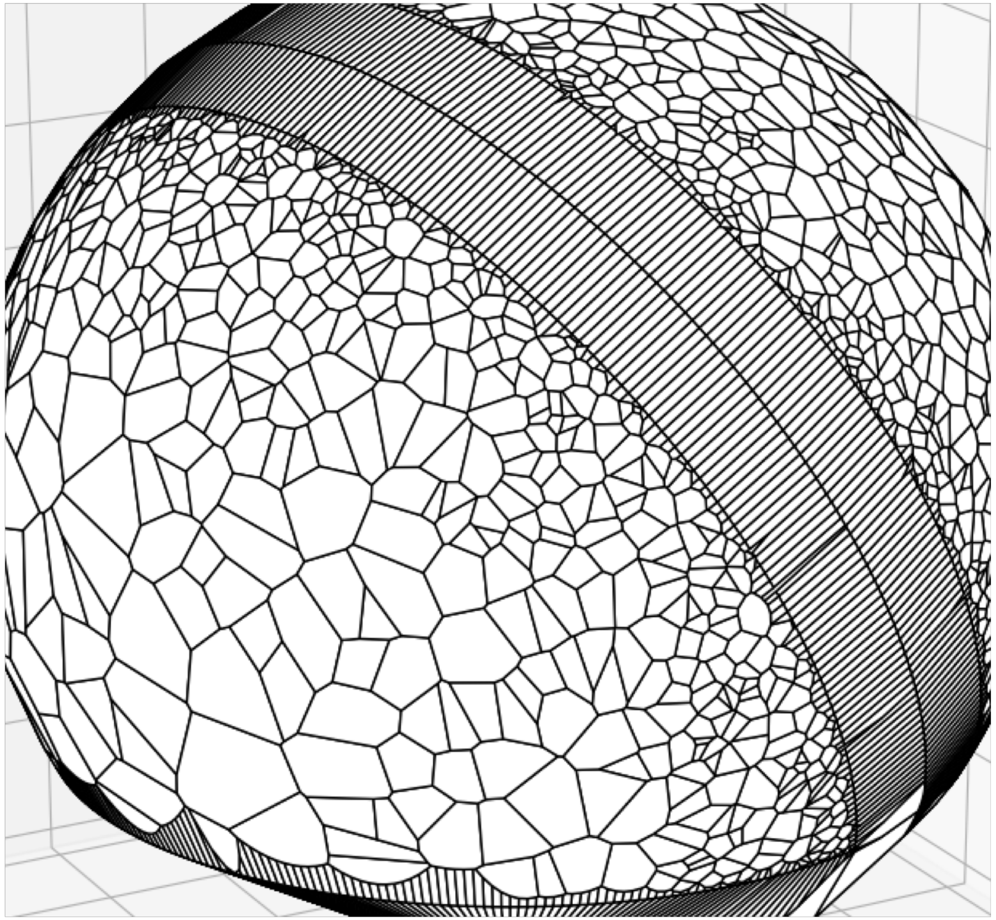
- Density for object j, summing over pixels i

$$\psi_j(mag) = \sum_i \left[\frac{1}{\Omega_i \int_{D_{i,j,min}}^{D_{i,j,max}} \rho(r) r^2 dr} \right] (mag)$$

- Losing less survey volume than any global approach
- Retrieve all good quality sources

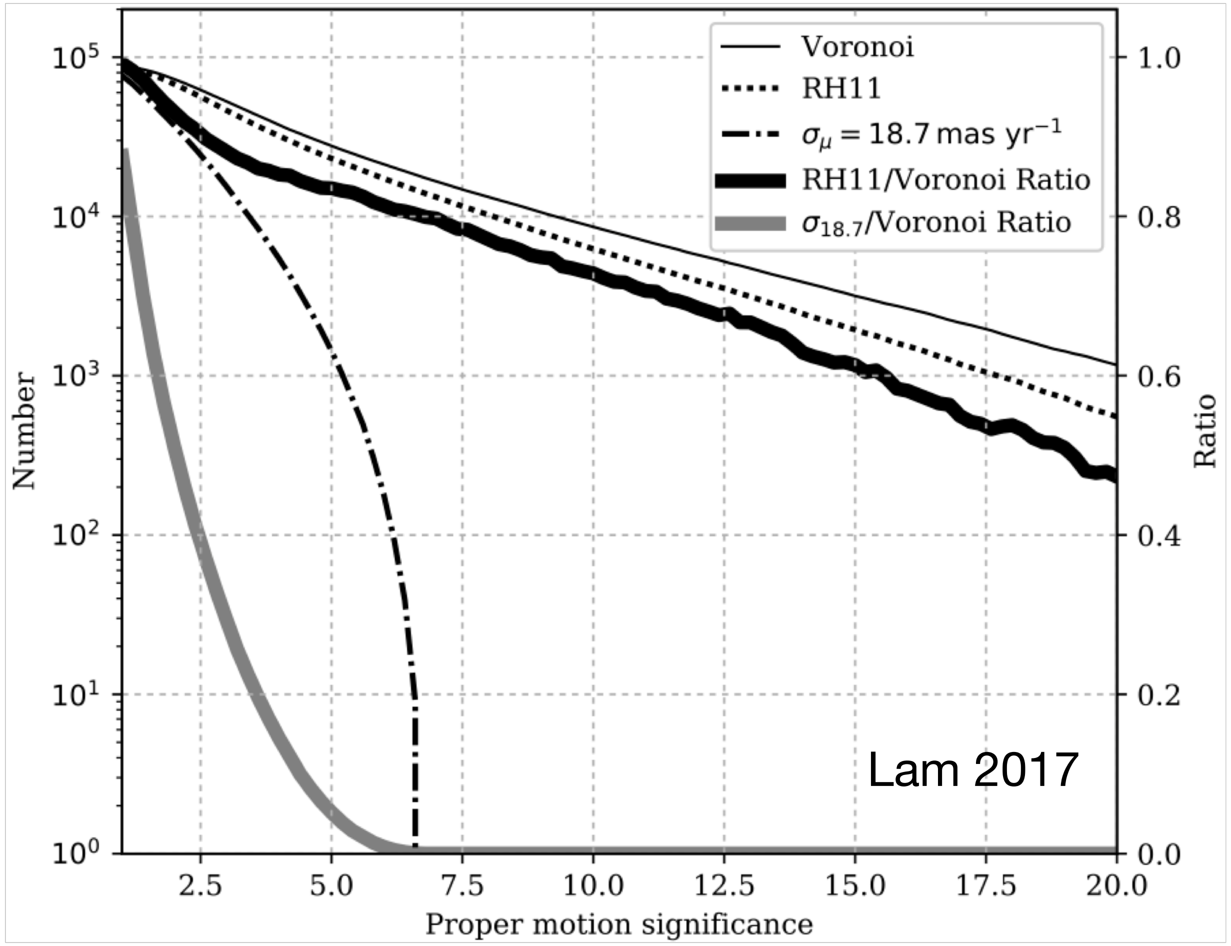
Footprint Area

- Adding artificial points along both sides the boundary
- Summing the area of the artificial cells to the parent cells



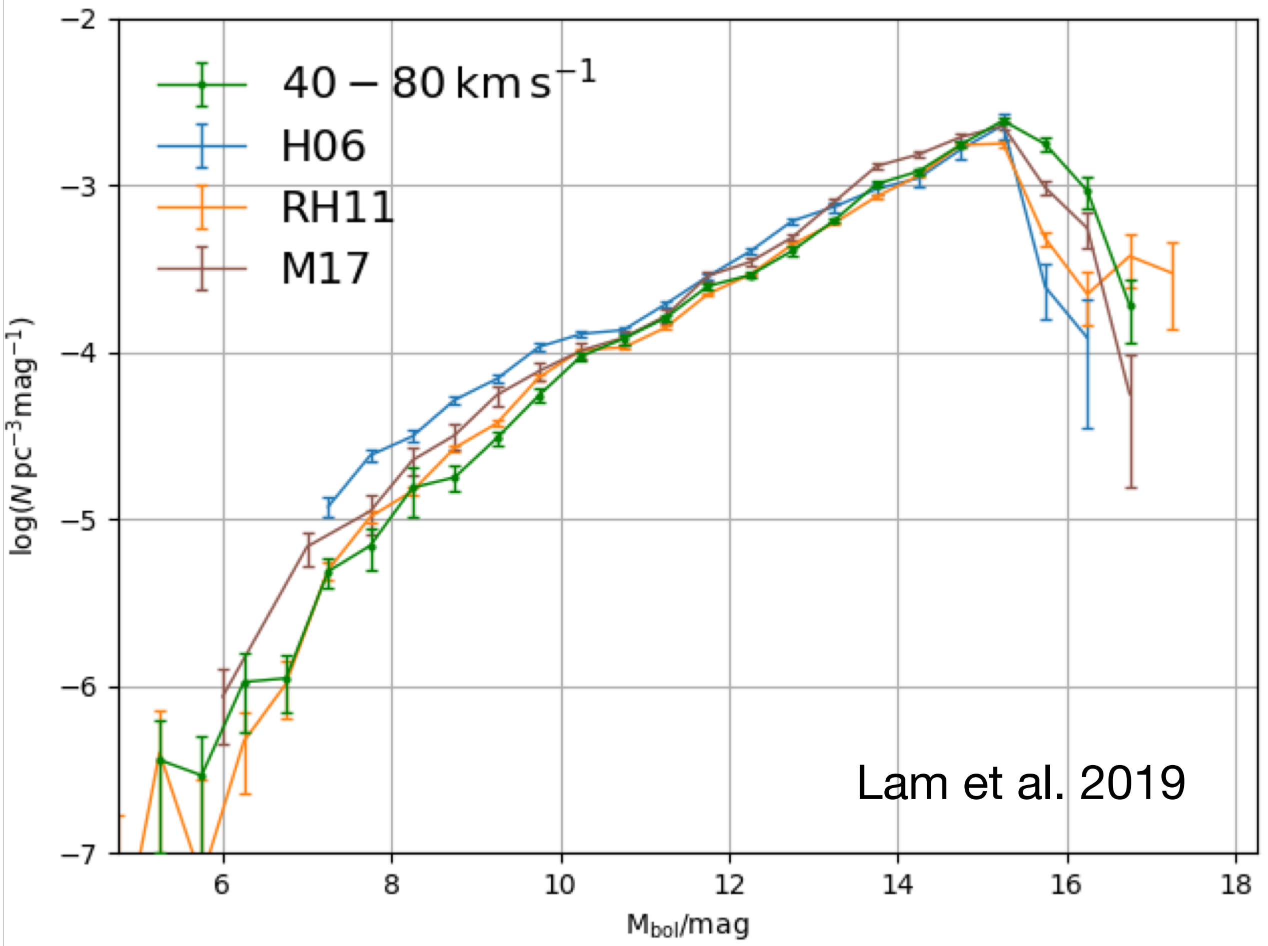
Yield (simulation)

- ~10 times more sources compared to global limit
- 25% more compared to a 5x5 degrees tile method (Rowell et al. 2011)
- Most significant gain at the faint end



Yield (observed)

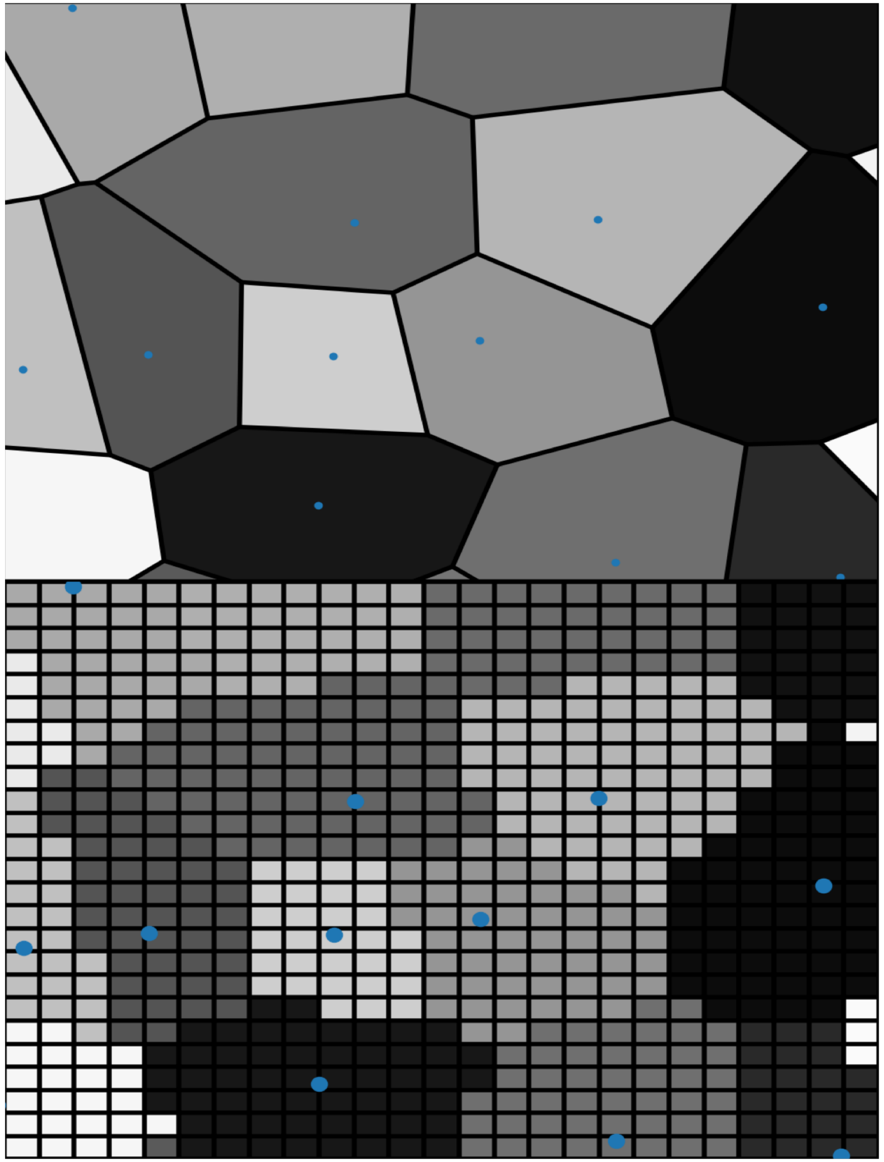
- ~15% more compared to adopting the "average" statistics from ~3°x3° tiles
- Good agreements with previous work



GCRF
Global Challenges Research Fund

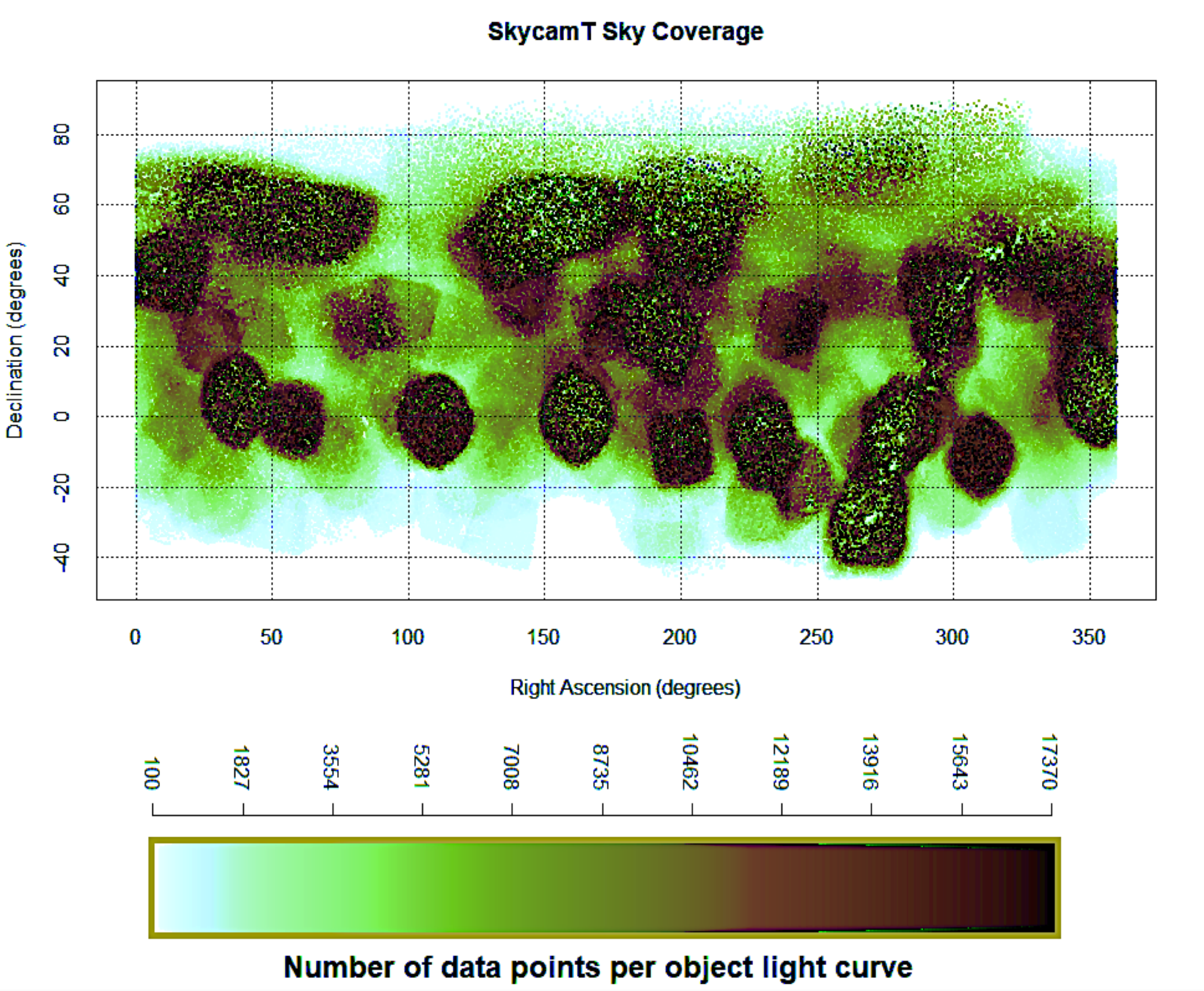
Extension

- Same Principle, different pixelation schemes
- Apply the concept of Voronoi Tessellation
- Masking useful pixels with a footprint service (Budavári 2007, 2010)
- Assign HEALPix pixels to the nearest sources
- or with HTM or Q3C
  (pixels do not have equal-area)
- Less Maths



Extension

- Skycam T at Liverpool Telescope with 600,000 light curves (Mawson et al. 2013)
- Identify density peaks to be the generating data points
- Aggregate multiple sources to create trend model



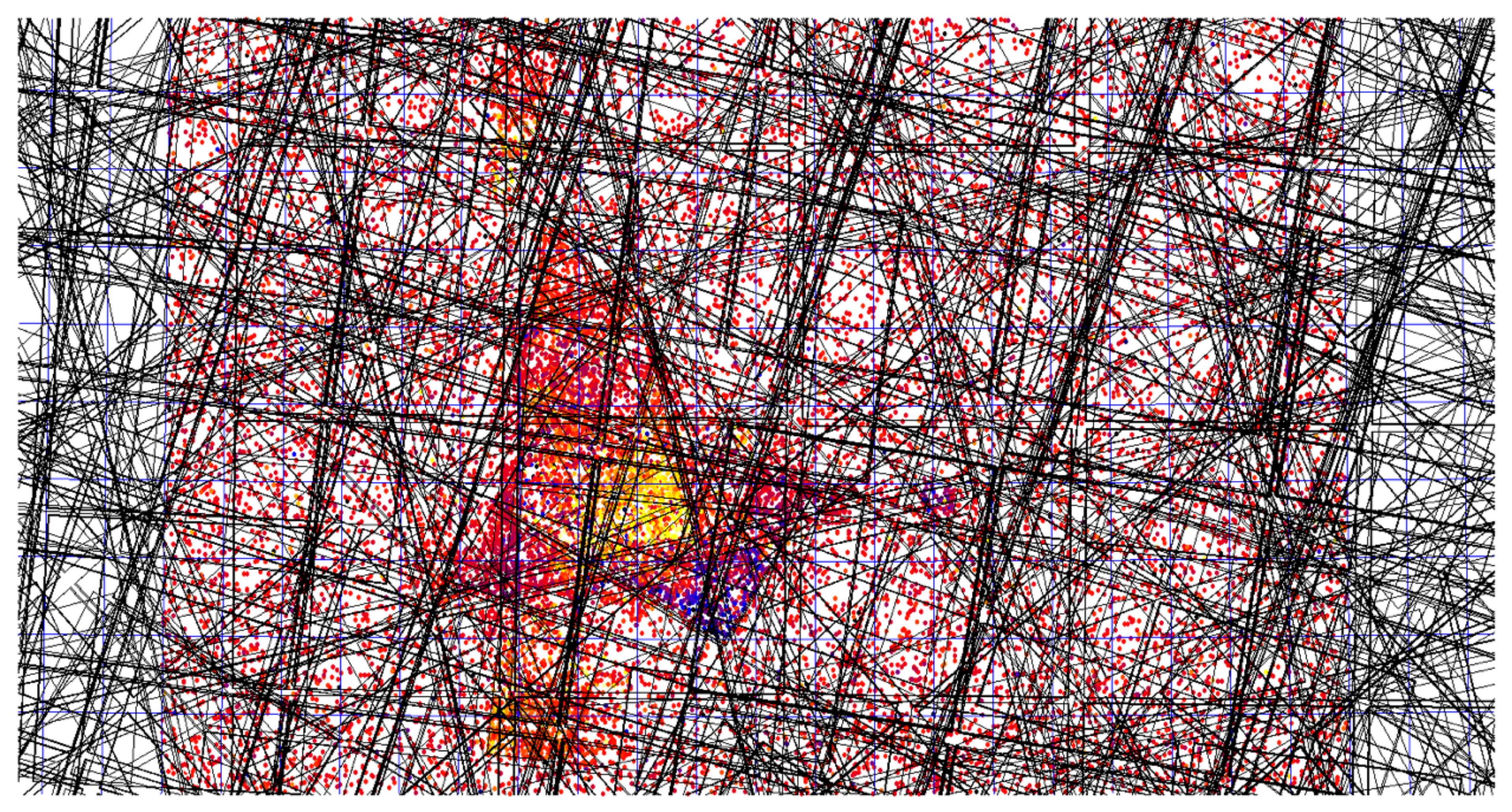
Other Work

- Browser-based FITS image viewer for school kids with the National Schools' Observatory
- Astroecology Machine Learning & Data Management System (P1-9)
- LT pipeline and web enhancement

Global Challenges Research Fund

Questions?

Global Challenges Research Fund



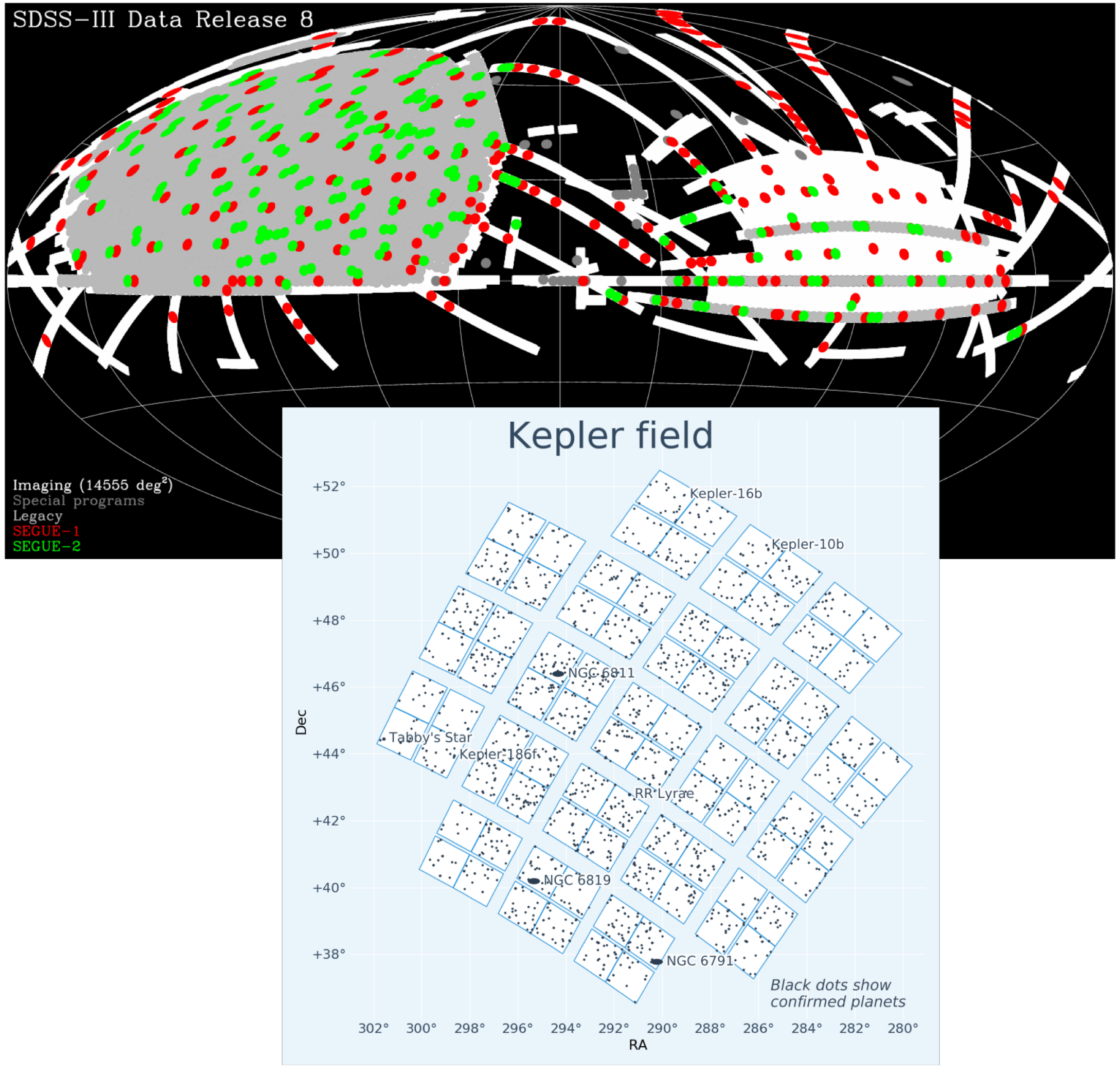
Multi-epoch Large Sky Area Survey

- Wide Field of View
- Many Repeated Observations
- Many CCD/IR Detectors on the focal plane
- SDSS (30), VISTA (16), VST (32), DES (74), Pan-STARRS (60), Gaia (63), LSST (189), EUCLID (32)

GCRF
Global Challenges Research Fund

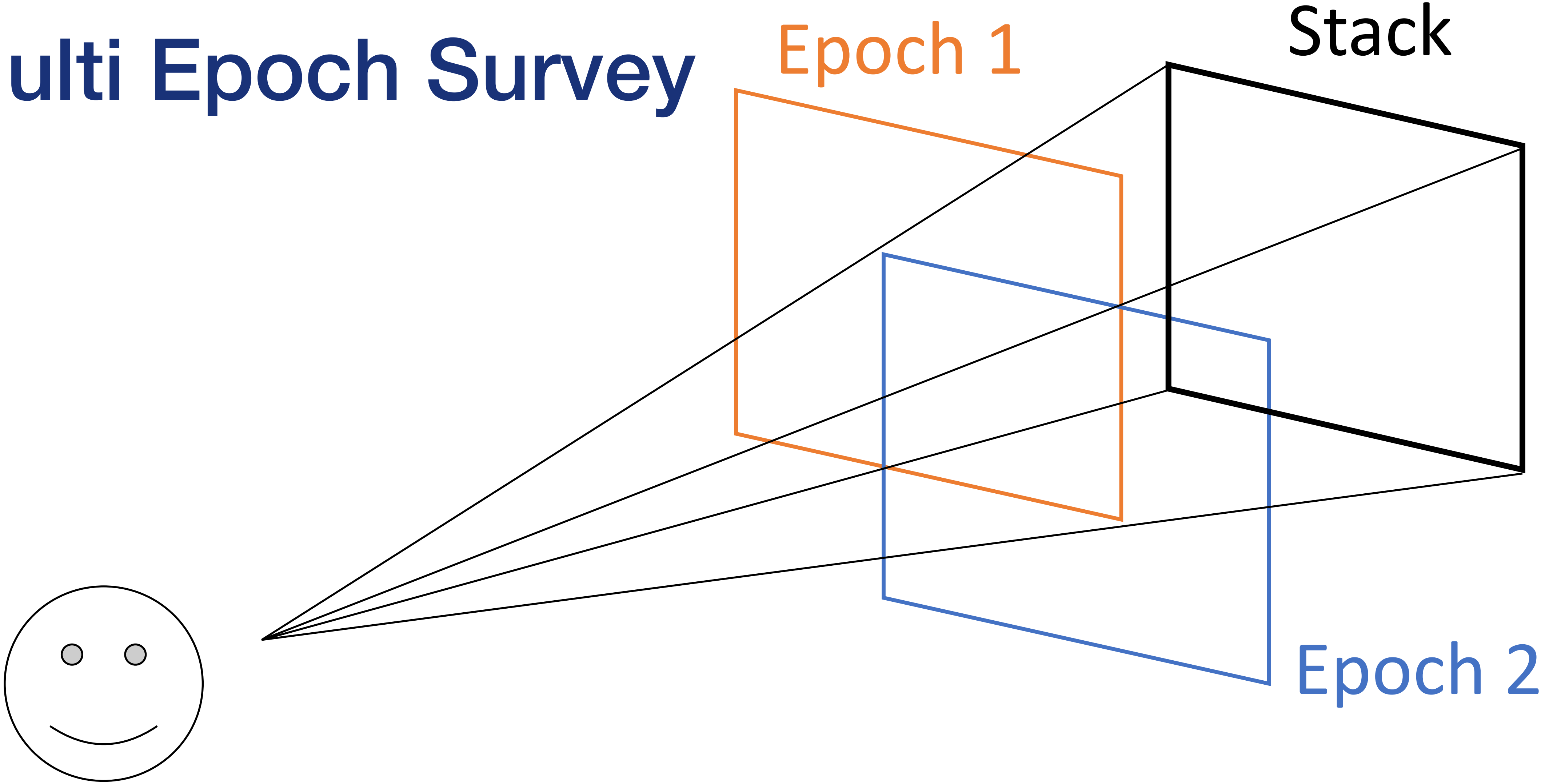
Footprint Area

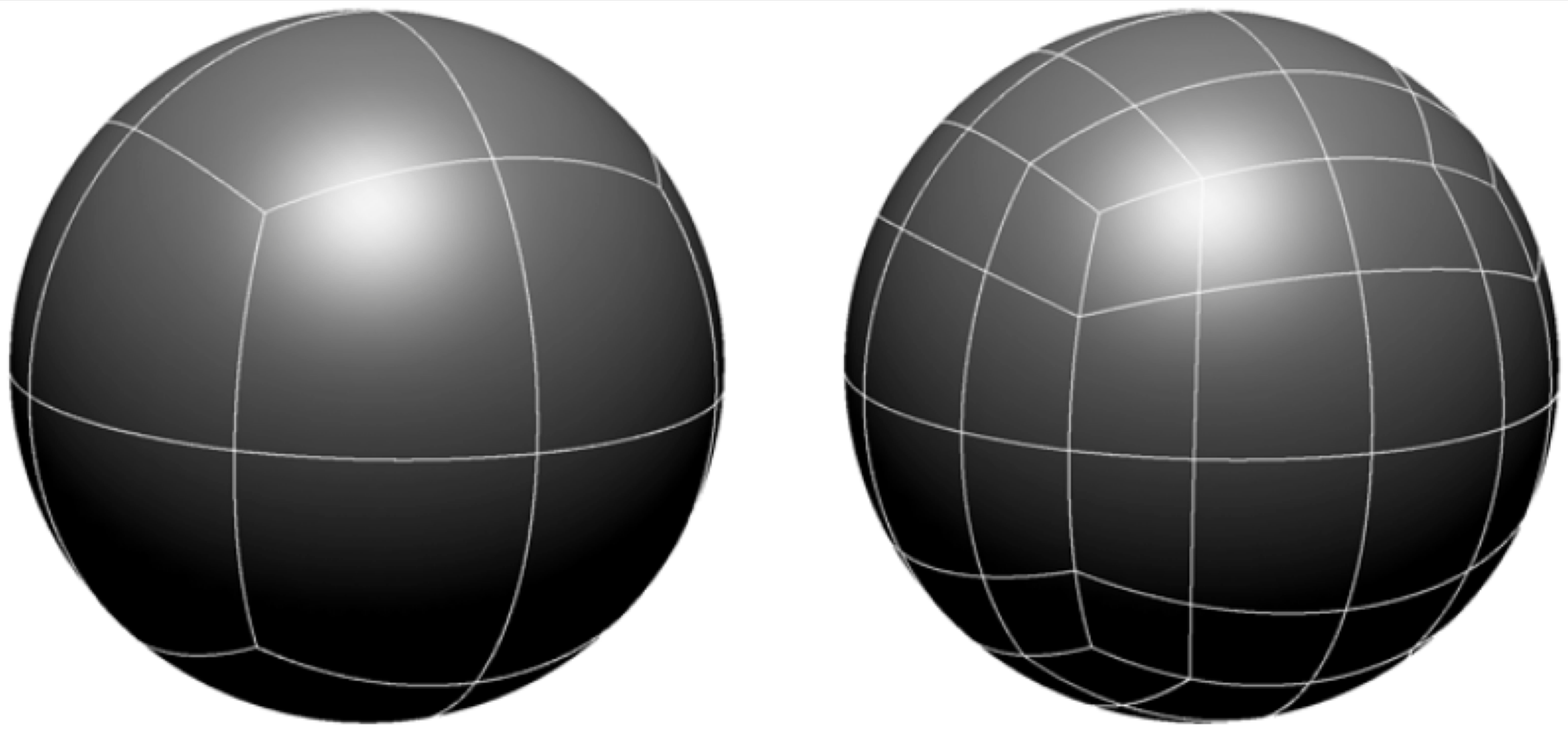
- Very difficult with complex footprint area (e.g. SDSS DR8)
- Completely unnecessary for survey like Kepler



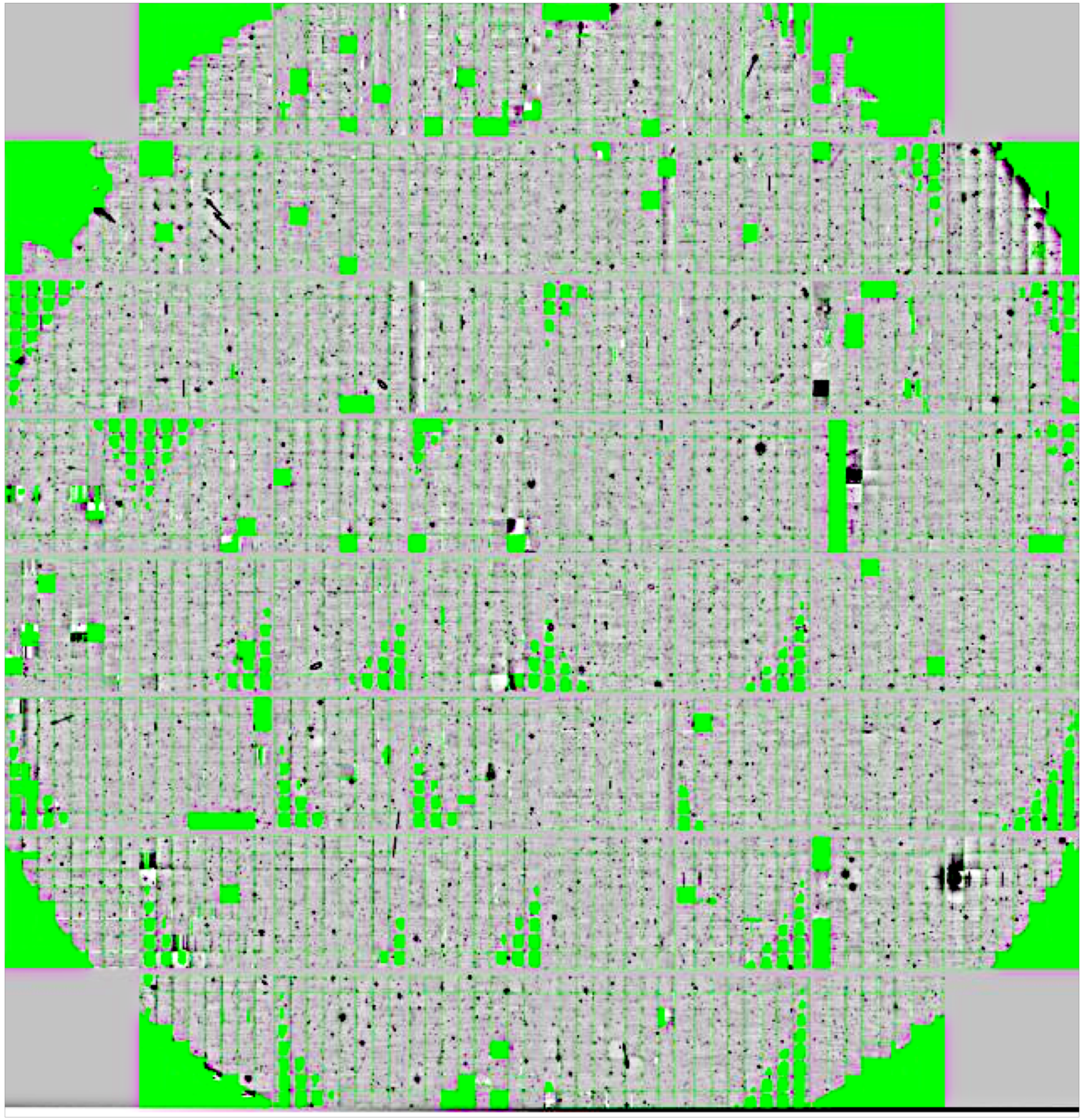
GCRF
Global Challenges Research Fund

Multi Epoch Survey



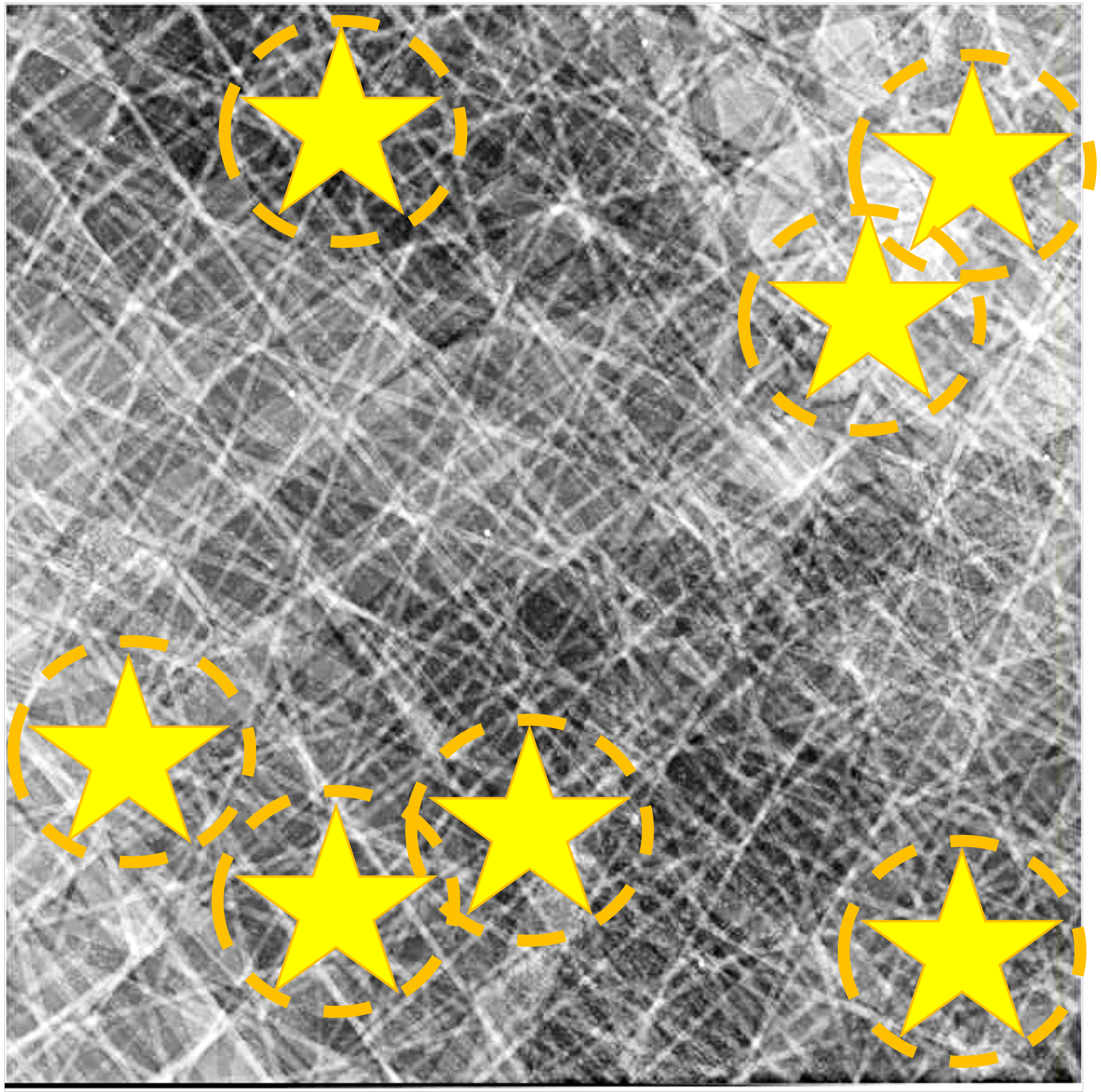


CCD Mask



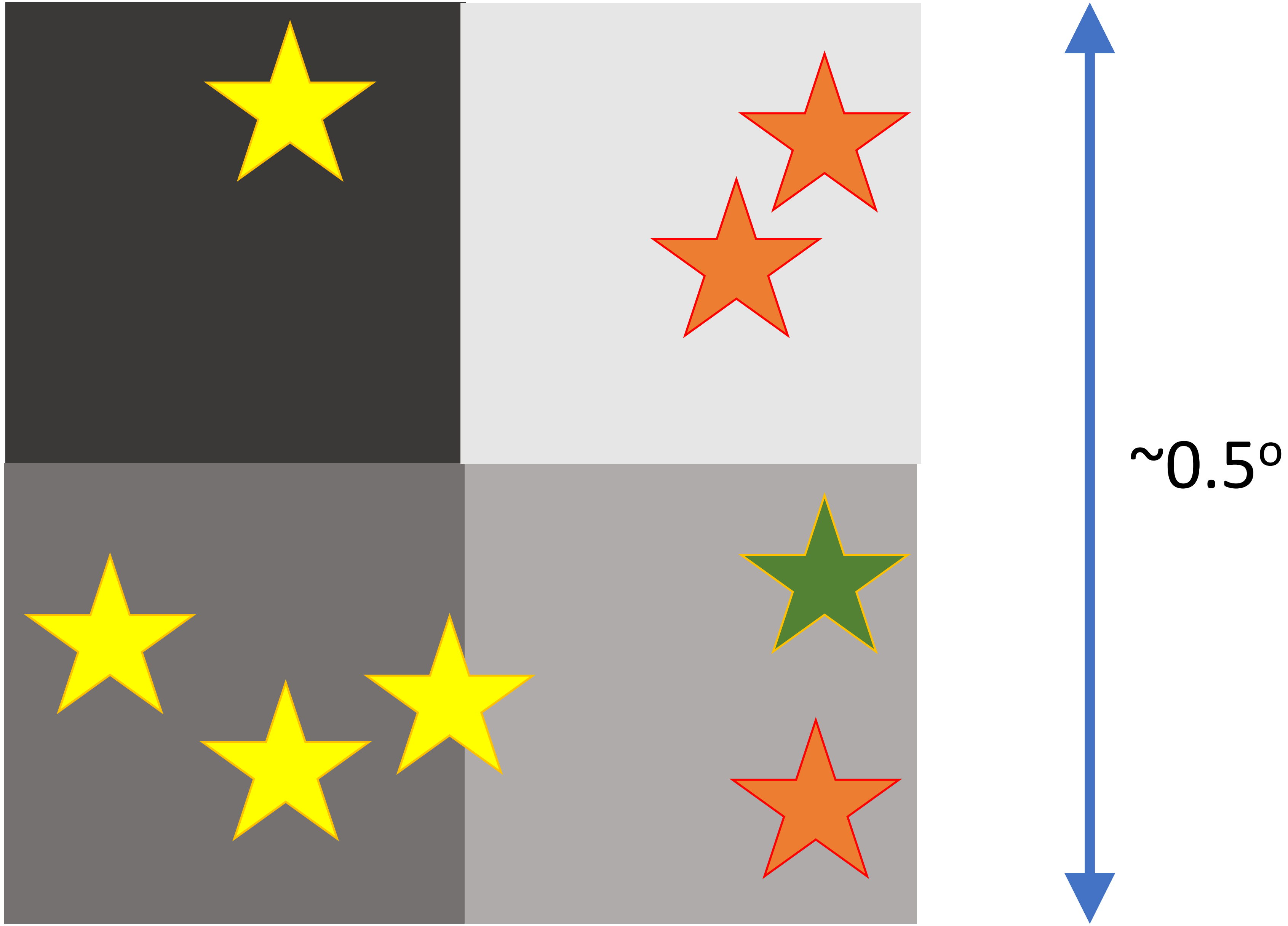
~3°

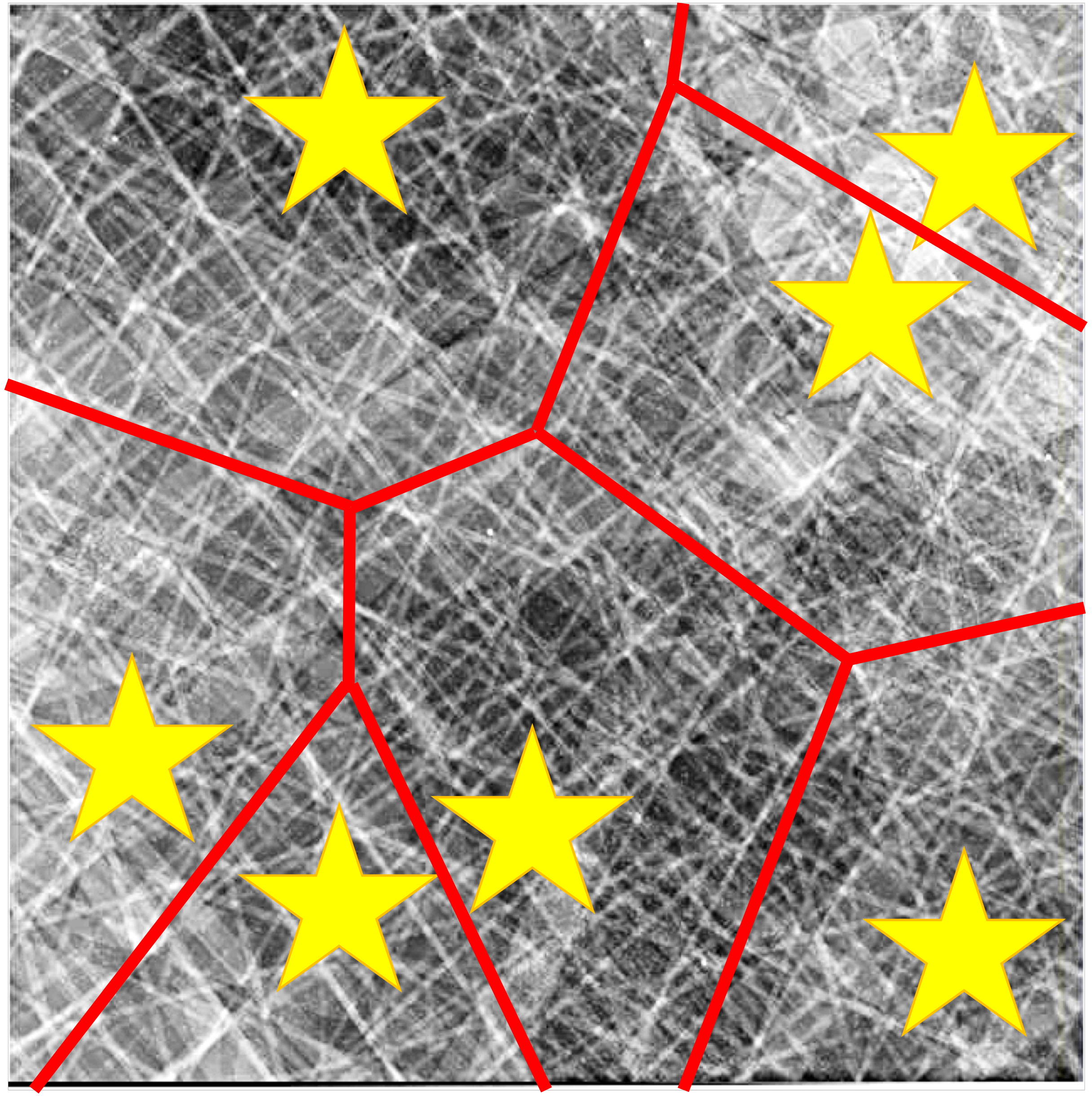
Individual
5σ
Photometric
+
Astrometric
Limits



~0.5°

Averaged 5σ Astrometric limit





~0.5°

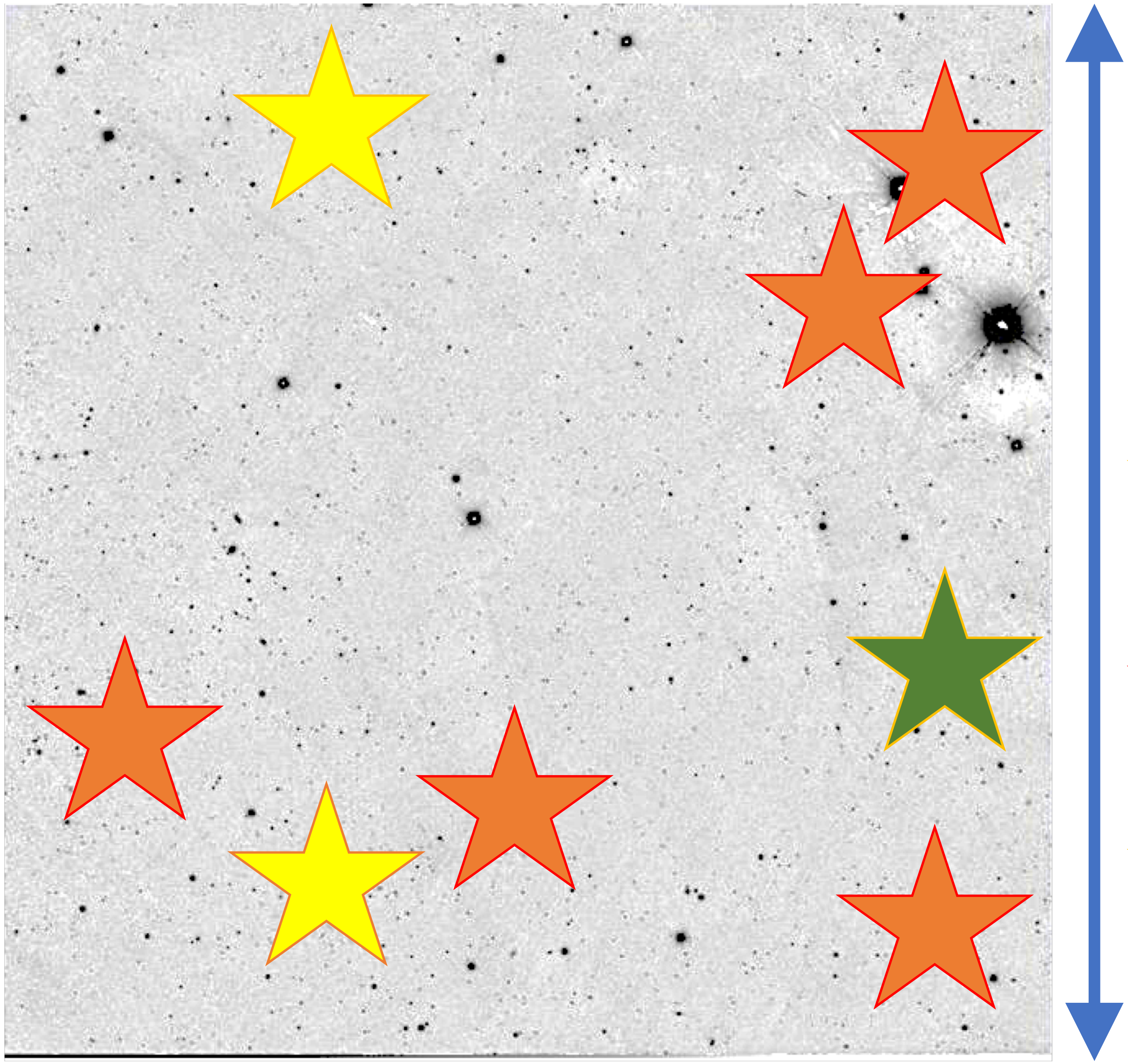
5σ

Photometric

+

Astrometric

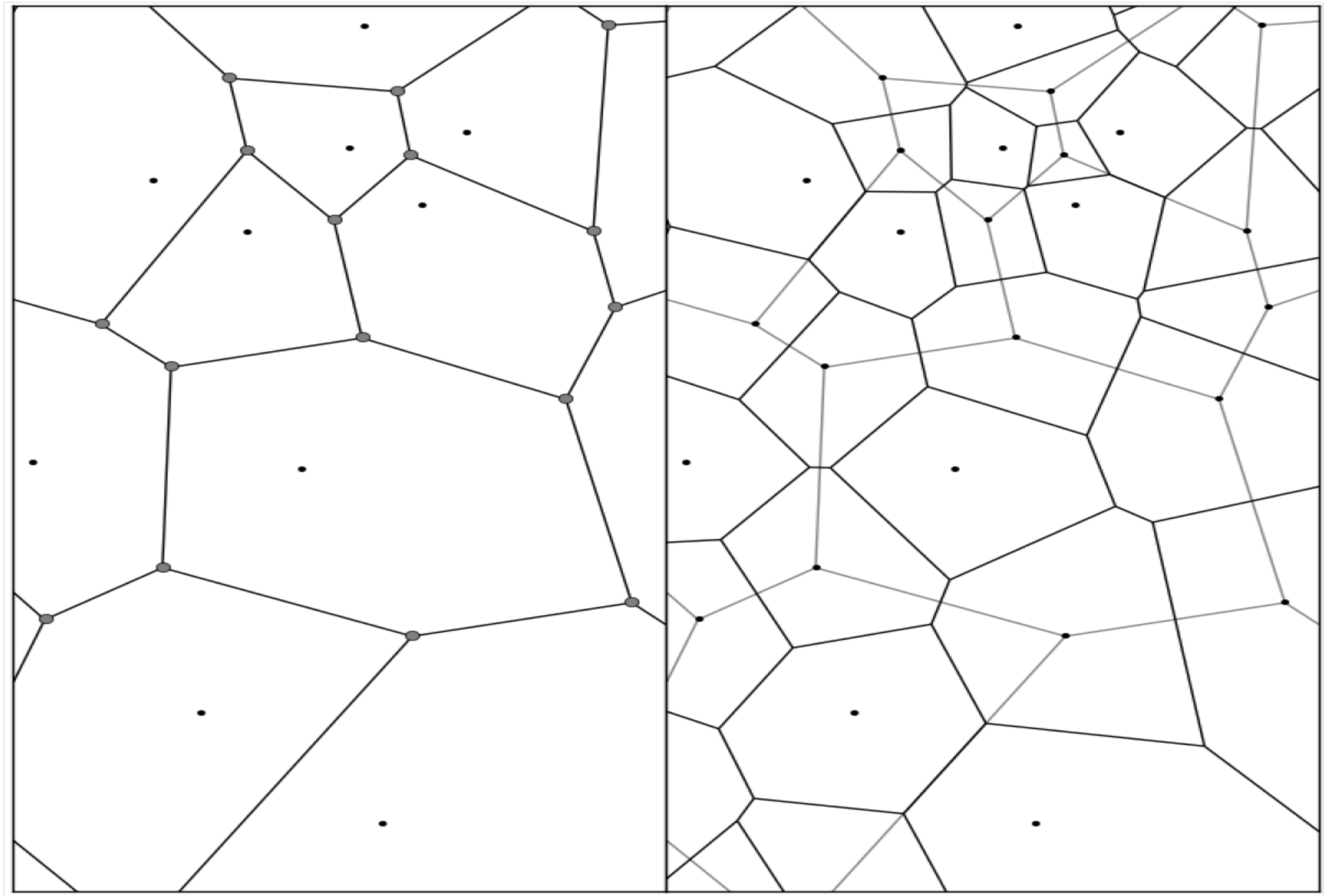
Limit



Good & Selected

Good & NOT Selected

Spurious

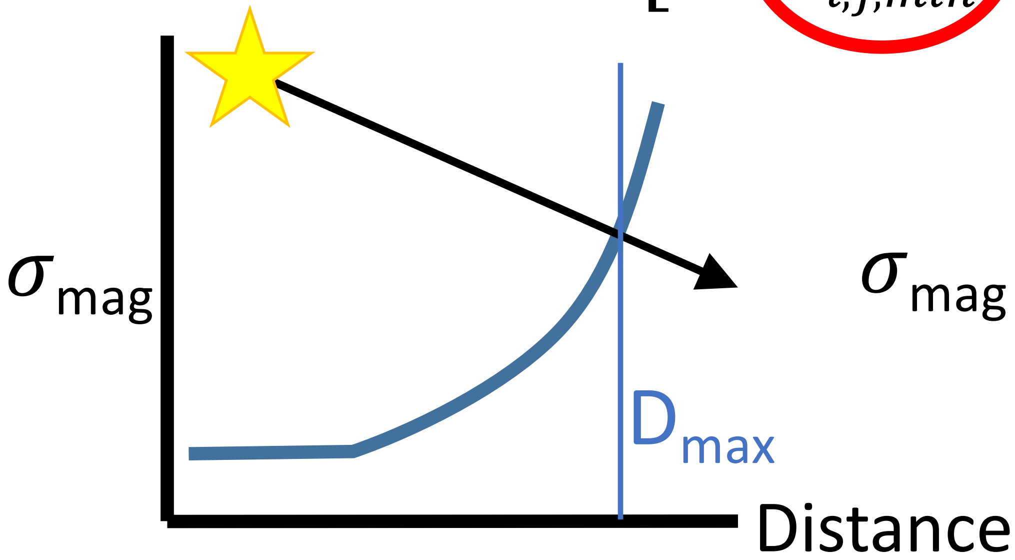


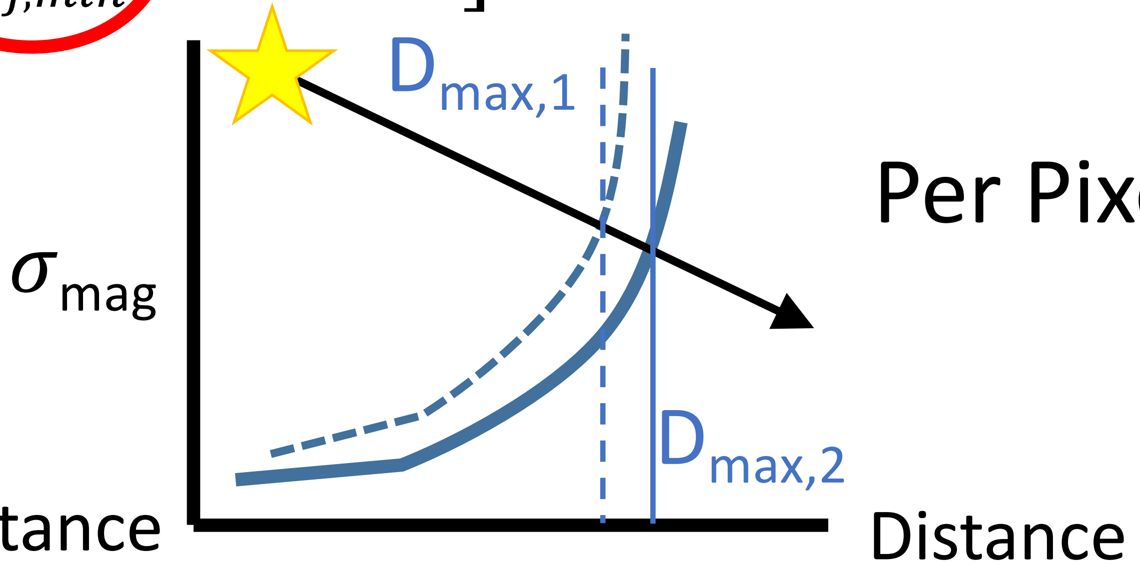
Consequence to the Maximum Volume

- Density for object j, summing over pixels i

$$\psi_j(mag) = \sum_i \left[\frac{1}{\Omega_i \int_{D_{i,j,min}}^{D_{i,j,max}} \rho(r) r^2 dr} \right] (mag)$$

Global



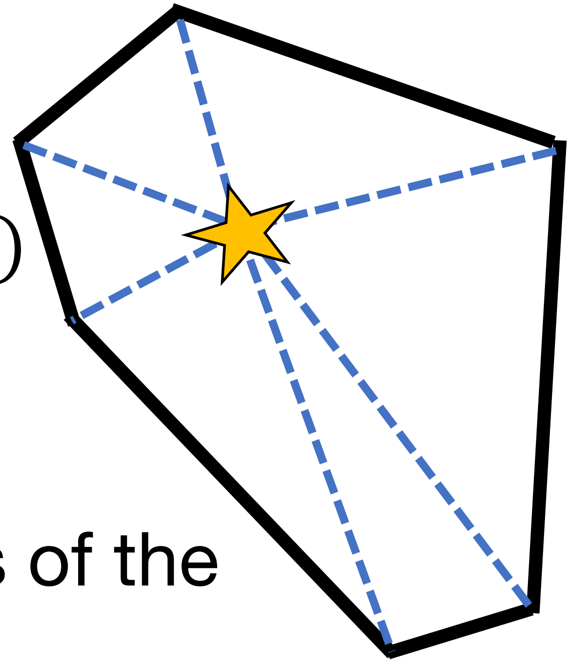


Per Pixel

Consequence to the Maximum Volume

- Density for object j, summing over pixels i

$$\psi_j(mag) = \sum_i \left[\frac{1}{\Omega_i \int_{D_{i,j,min}}^{D_{i,j,max}} \rho(r) r^2 dr} \right](mag)$$



- The area of a cell is the sum of the spherical excess of the constituent spherical triangles (L'Huilier's Theorem)

Propagate photometric uncertainties, observing conditions, instrumental noises from EACH epoch

2.2.1 Photometric Uncertainty

When a source is being tested for the observability, it is 'placed' at a different distance so the apparent brightness changes as a consequence. The background and other instrumental noises are constant, but the Poisson noise from the source changes with the measured flux, hence the photometric uncertainties are functions of distance. The total noise, N, of a photometric measurement can be estimated by

$$N = \sqrt{(F + d + s) \times t + r^2} \quad (1)$$

where F is the instrumental flux per unit time, d is the dark current per unit time, s is the sky background flux per unit time, t is the exposure time and r is the read noise. Among these quantities, d, s, t and r are fixed quantities in a given epoch, only the flux varies as a function of distance. We use F as the measured flux and $\mathcal{F}(D)$ as the flux at an arbitrary distance D. Therefore, in a Voronoi cell j at epoch k, the photometric noise of source i is

$$N_{i,j,k}(D) = \sqrt{(\mathcal{F}_i(D) + d_{j,k} + s_{j,k}) \times t_{j,k} + r^2} \quad (2)$$

where the flux at D is calculated from applying the inverse square law on the observed flux F_i and observed distance D_i,

$$\mathcal{F}_i(D) = F_i \times \left(\frac{D_i}{D}\right)^2. \quad (3)$$

The random photometric uncertainty of a source at an arbitrary distance in a given epoch is the inverse signal-to-noise ratio,

$$\delta\mathcal{F}_{i,j,k}(D) = \frac{N_{i,j,k}(D)}{\mathcal{F}_i(D)}. \quad (4)$$

The total photometric uncertainty of the source as a function of distance, combining with the systematic uncertainty, σ_s, coming from the absolute calibration of the detector, is therefore

$$\sigma_{i,j,k}(D) = \sqrt{\delta\mathcal{F}^2_{i,j,k}(D) + \sigma_s^2}, \quad (5)$$

which represents the photometric uncertainty as a function of the distance to the source.

Lam 2017

2.2.2 Astrometric Uncertainty

The least square solution of proper motion in one direction for source i can be expressed in the following matrix form, the epoch is labelled by the subscript from 1 to $M(j)$ where M is the number of epochs in cell j,

$$\underbrace{\begin{pmatrix} \frac{1}{\sigma_1} & \frac{\Delta t_1}{\sigma_1} \\ \vdots & \vdots \\ \frac{1}{\sigma_{M(j)}} & \frac{\Delta t_{M(j)}}{\sigma_{M(j)}} \end{pmatrix}}_{\mathsf{A}} \times \begin{pmatrix} \overline{\alpha} \\ \mu_\alpha \end{pmatrix} = \begin{pmatrix} \frac{\Delta\alpha_1}{\sigma_1} \\ \vdots \\ \frac{\Delta\alpha_{M(j)}}{\sigma_{M(j)}} \end{pmatrix} \quad (6)$$

where Δt_k is the time difference between the mean epoch and epoch k, $\Delta\alpha_k$ is the positional offset from the mean position, $\overline{\alpha}$, and proper motion, μ_α, in the direction of the right ascension. The associated uncertainties can be found from the diagonal terms of the normal matrix,

$$\mathsf{A}^{\mathrm{T}}\mathsf{A} = \begin{bmatrix} \sum_k \left(\frac{1}{\sigma_k}\right)^2 & \sum_k \left(\frac{1}{\sigma_k}\frac{\Delta t_k}{\sigma_k}\right) \\ \sum_k \left(\frac{1}{\sigma_k}\frac{\Delta t_k}{\sigma_k}\right) & \sum_k \left(\frac{\Delta t_k}{\sigma_k}\right)^2 \end{bmatrix} \quad (7)$$

so for each cell,

$$\frac{1}{\sigma^2_{\mu_\alpha \cos\delta}} = \sum_k \left(\frac{\Delta t_k}{\sigma_k}\right)^2 \quad (8)$$

and the total proper motion uncertainty is

$$\sigma_\mu = \sqrt{\sigma^2_{\mu_\alpha \cos\delta} + \sigma^2_{\mu_\delta}} = \sqrt{2}\sigma_{\mu_\alpha \cos\delta}. \quad (9)$$