UKD-S9 —
Cross-Matching and Astrometry at LSST Depths

Macauff

birnam

banquo

Tom J. Wilson & Tim Naylor
with the UK DAC team

Matching Across Catalogues using the Astrometric Uncertainty Function and Flux

macauff is a cross-match algorithm that:

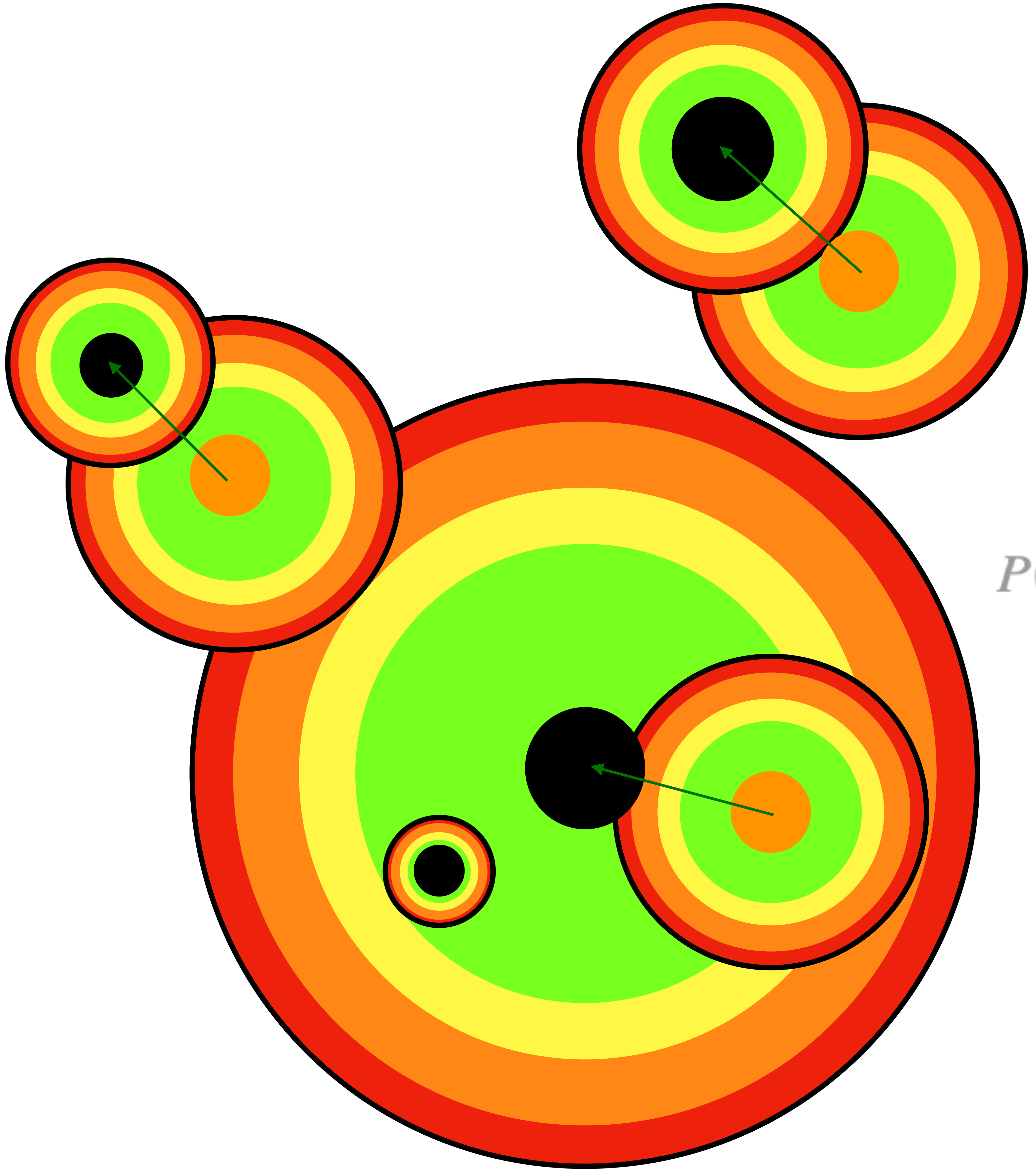
- **is fully Bayesian, providing relative strength of match pairings for the user**
- **includes the astrometric uncertainties of all objects in considering matches on positional grounds**
- **uses the measured colours of potential matches as a further discriminator of true and false matches**
- **is a "many-to-many" algorithm, solving for all potential competing counterparts simultaneously**
- **extends the "Astrometric Uncertainty Function" to include more than just the *centroid* uncertainty, adding e.g.**
  - **unknown or weakly-known proper motion of stars**
  - **<u>the "Perturbation Component" of the AUF — the effect of unresolved contaminants in measured positions</u>**

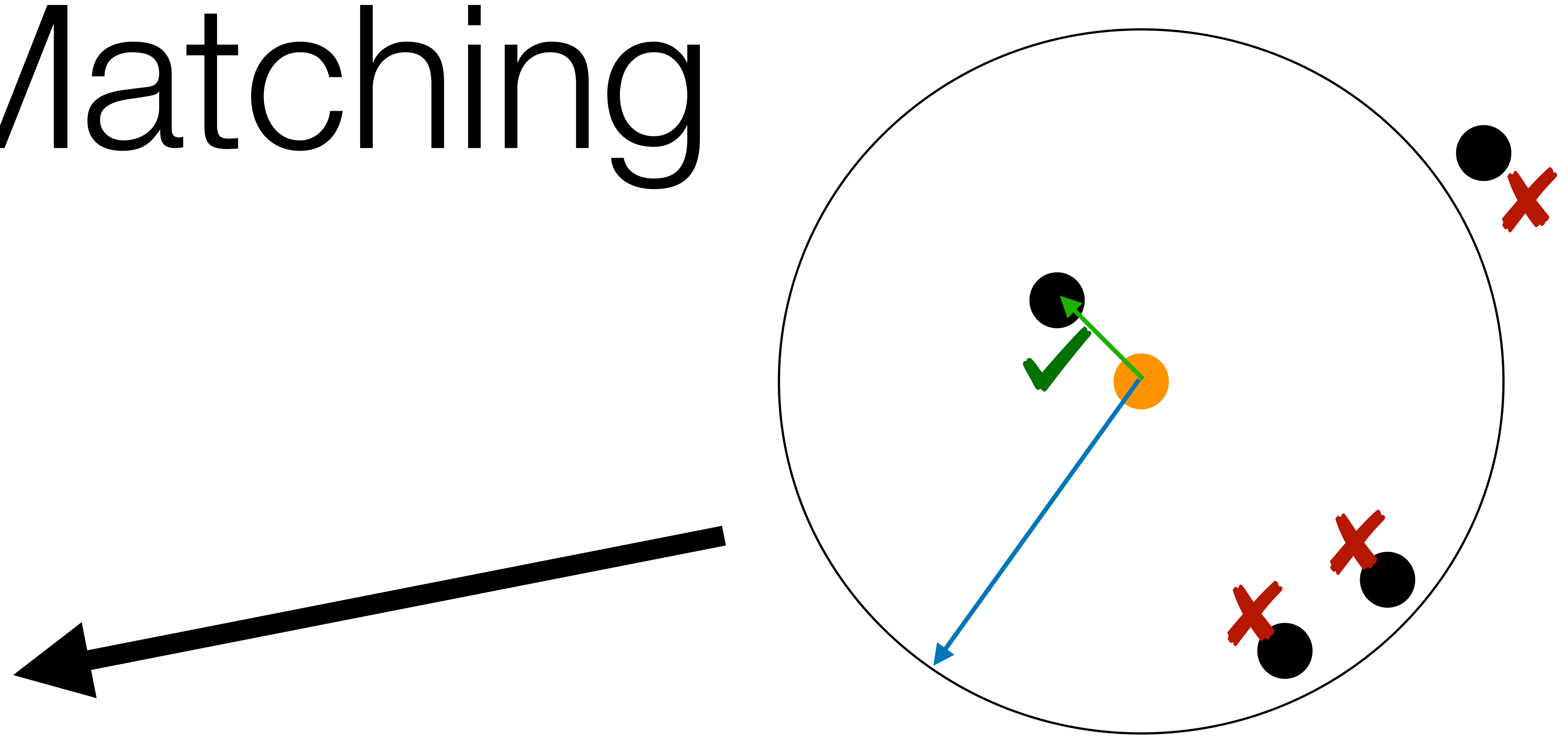
(While *we* consider basically the entire LSST sky to be populated enough that unresolved contaminants will affect just about every faint LSST object, the Galactic plane is, unsurprisingly, more densely packed than the extragalactic sky, and where this last point makes us the most gains over other algorithms!)

Our first, and main, In-Kind Contribution is then to run this code on a number (N~10) of cross-matches of various photometric catalogues with Rubin for all DRs, and host those results on the UK IDAC for access by data-rights holders.

https://github.com/macauff/macauff

Probabilistic Cross-Matching





Probability of two sources having their <u>on-sky separation</u> given the hypothesis they are counterparts

$$P(\zeta, \lambda, k|\gamma, \phi) = \frac{1}{K} \times \prod_{\delta \notin \zeta \cap \delta \in \gamma} N_\gamma f_\gamma^\delta \prod_{\omega \notin \lambda \cap \omega \in \phi} N_\phi f_\phi^\omega \prod_{i=1}^{k} N_\mathrm{c} G_{\gamma\phi}^{\zeta_i \lambda_i} c_{\gamma\phi}^{\zeta_i \lambda_i}$$

Probability of sources having their <u>brightnesses</u> given they are unrelated to one another (“field stars”)

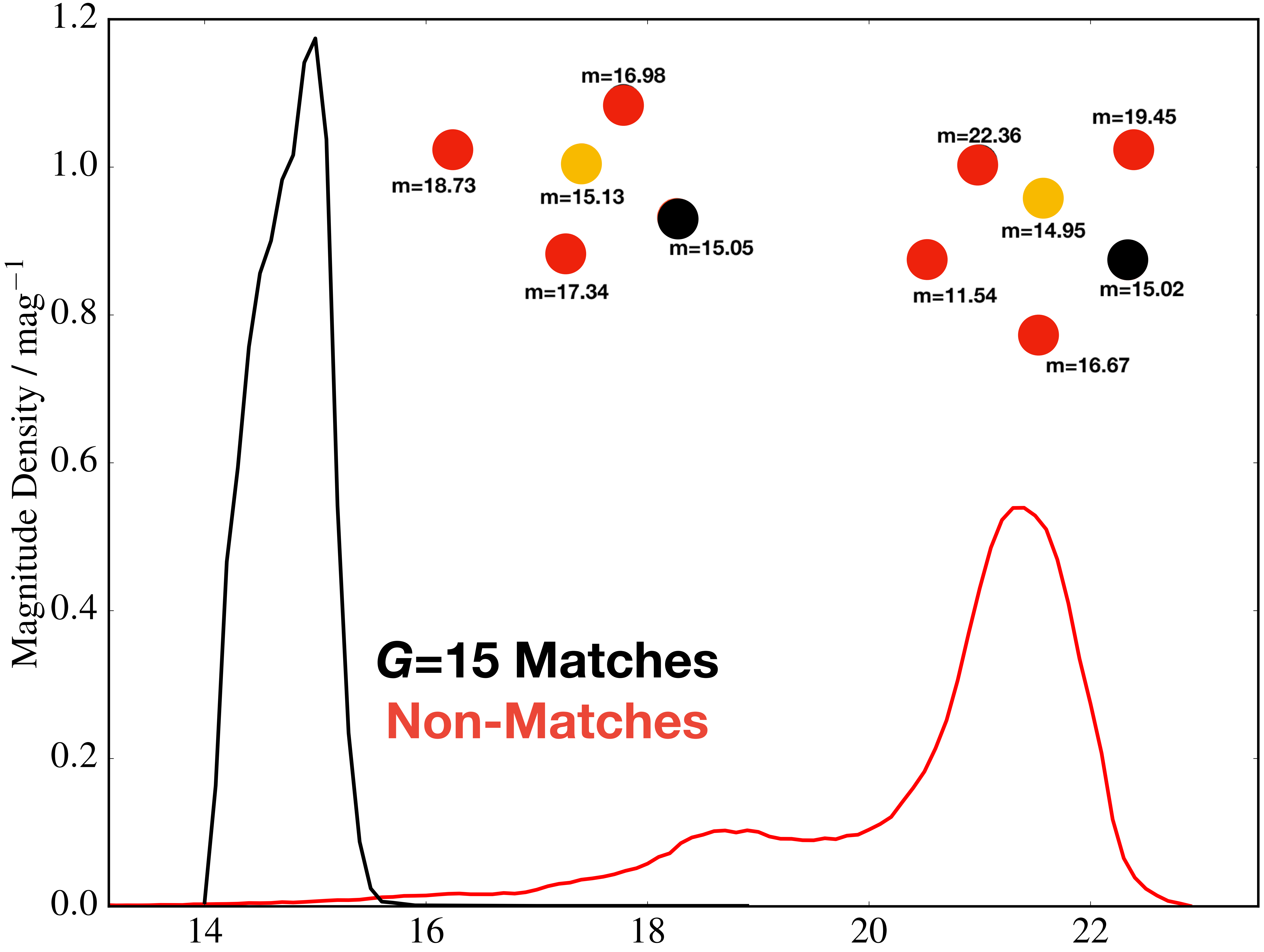
Probability of sources having their <u>brightnesses</u> given they are counterparts

Wilson & Naylor (2018a)

Tom J Wilson @onoddil

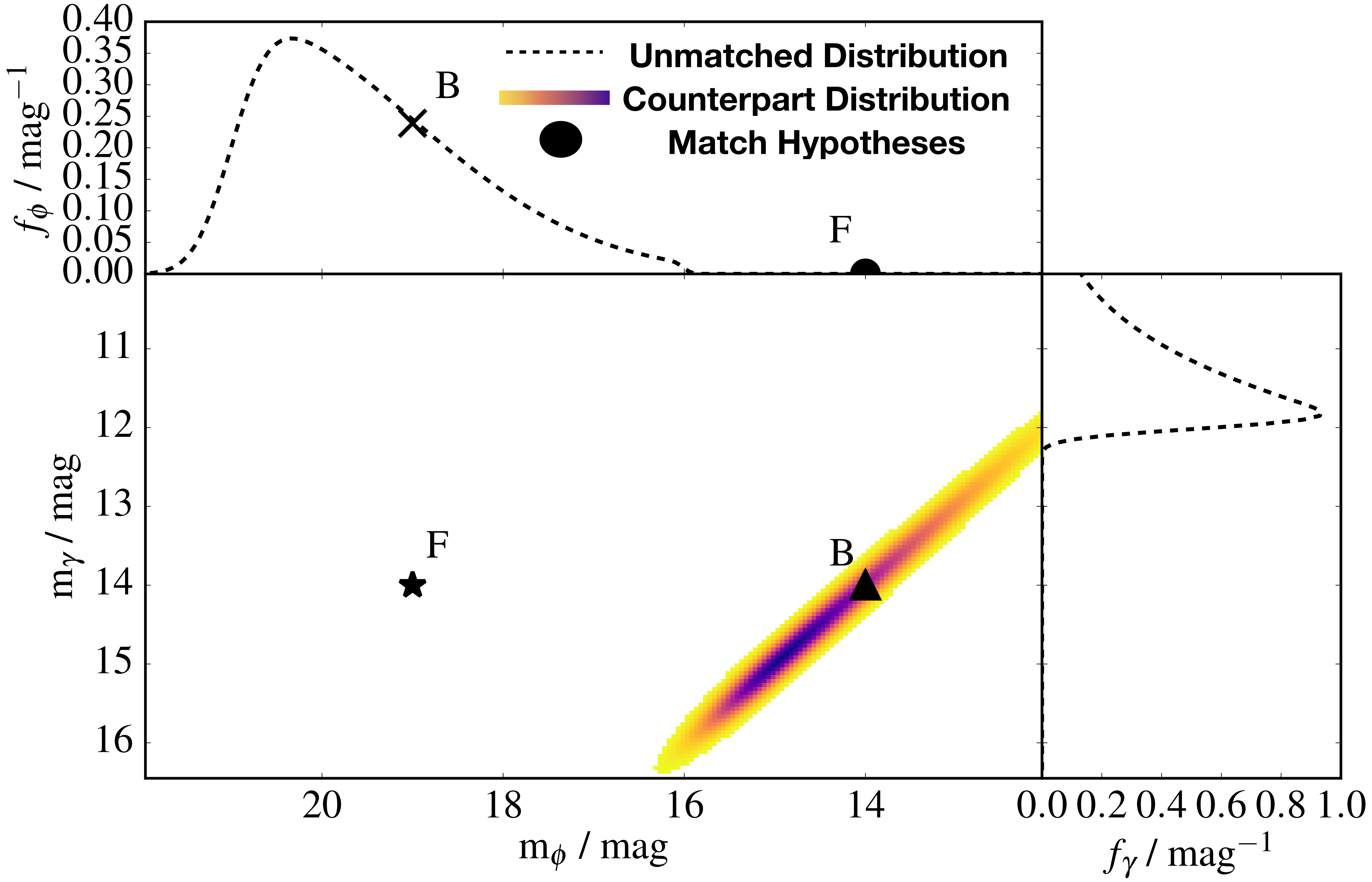
Photometry: Rejecting False Positives

$$P(\zeta, \lambda, k|\gamma, \phi) = \frac{1}{K} \times \prod_{\delta\notin\zeta\cap\delta\in\gamma} N_\gamma f_\gamma^\delta \prod_{\omega\notin\lambda\cap\omega\in\phi} N_\phi f_\phi^\omega \prod_{i=1}^{k} N_c G_{\gamma\phi}^{\zeta_i\lambda_i} c_{\gamma\phi}^{\zeta_i\lambda_i}$$



IPHAS - Barentsen et al. (2014)

Gaia DR2 - Gaia Collaboration, Brown A. G. A., et al. (2018)

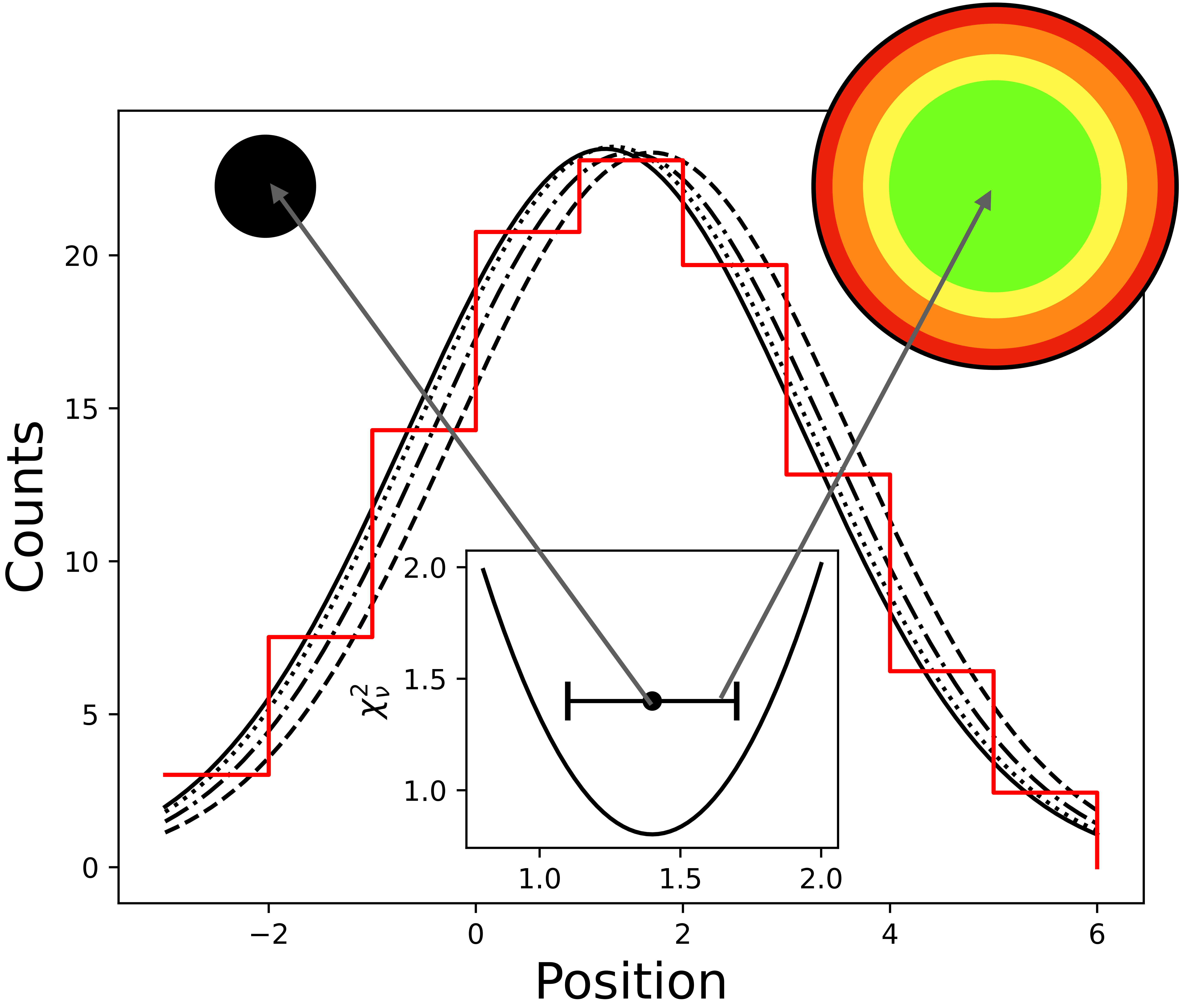


Wilson & Naylor (2018a)

Photometry-based likelihoods allow us to mitigate high false positive rate in crowded fields

Tom J Wilson @onoddil

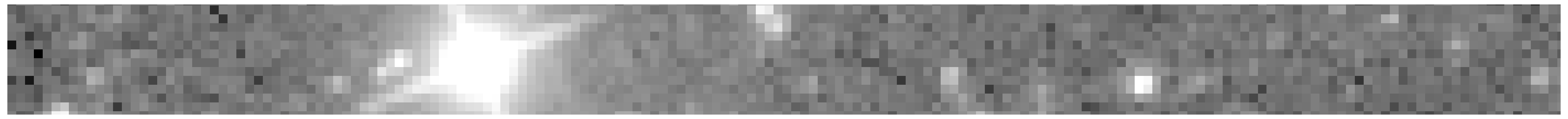
Centroid Positions and Uncertainties



$$p(D\,|\,M) \propto \frac{\exp\left(-\frac{1}{2}\left(\mathbf{x}-\boldsymbol{\mu}\right)^T \boldsymbol{\Sigma}^{-1}\left(\mathbf{x}-\boldsymbol{\mu}\right)\right)}{\sqrt{(2\pi)^k\,|\,\boldsymbol{\Sigma}\,|}}$$

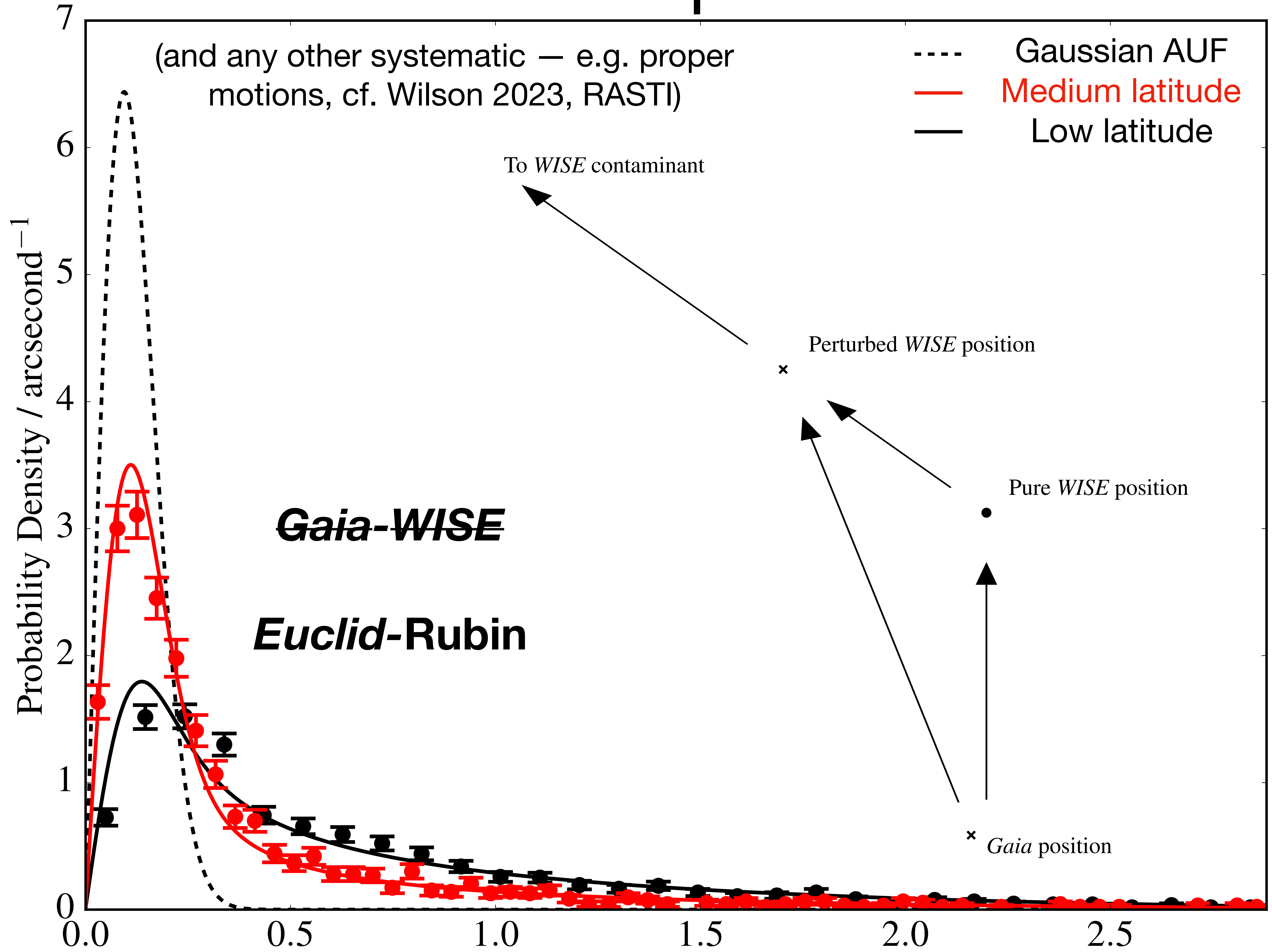
$$\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \boldsymbol{\mu} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}$$

$$g(x, y, \sigma) = (2\pi\sigma^2)^{-1}\exp\left(-\frac{1}{2}\frac{x^2+y^2}{\sigma^2}\right)$$

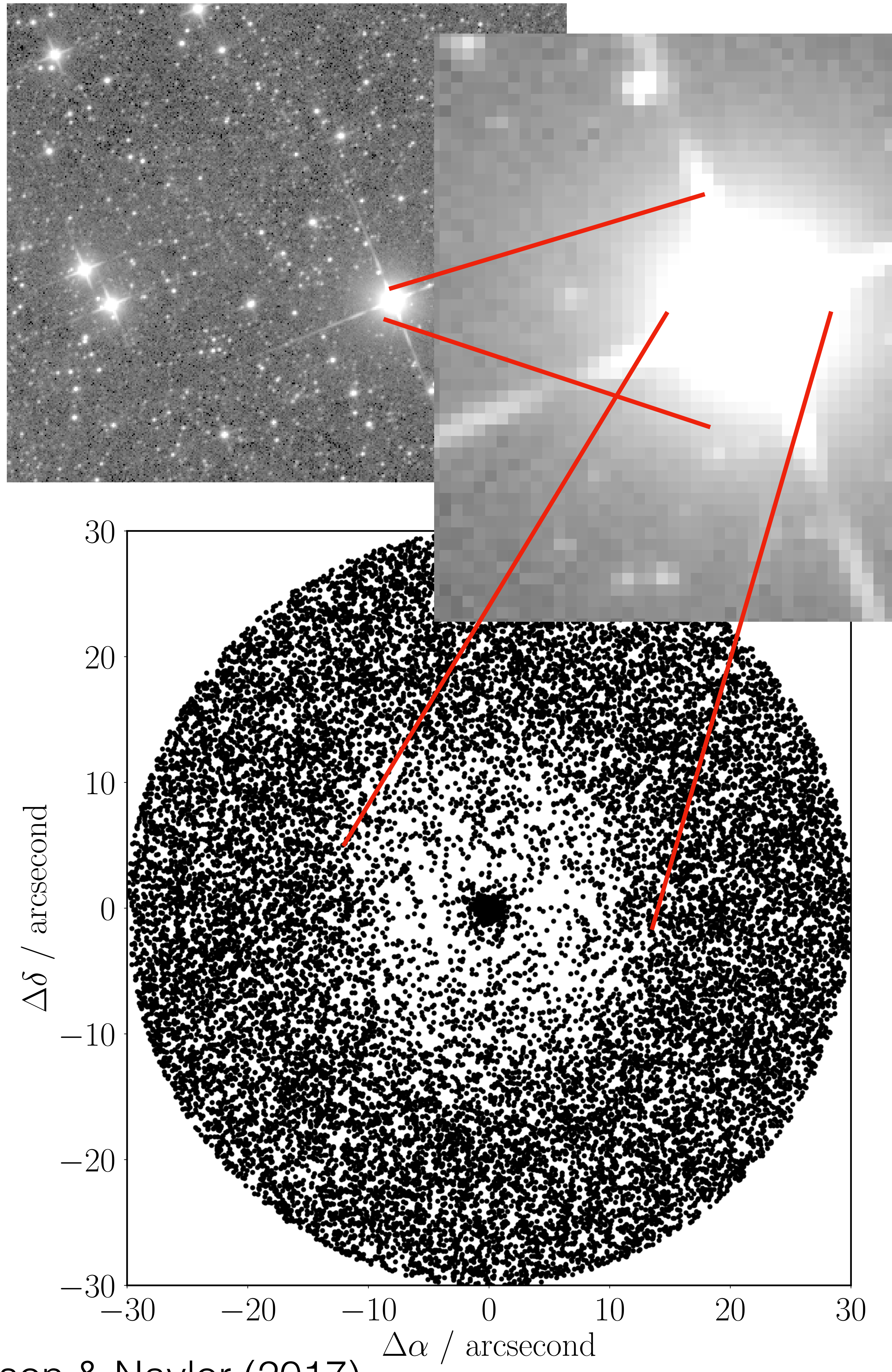


Tom J Wilson @onoddil

Additional Components of the AUF



Wilson & Naylor (2018b)



Wilson & Naylor (2017)

WISE - Wright et al. (2010)

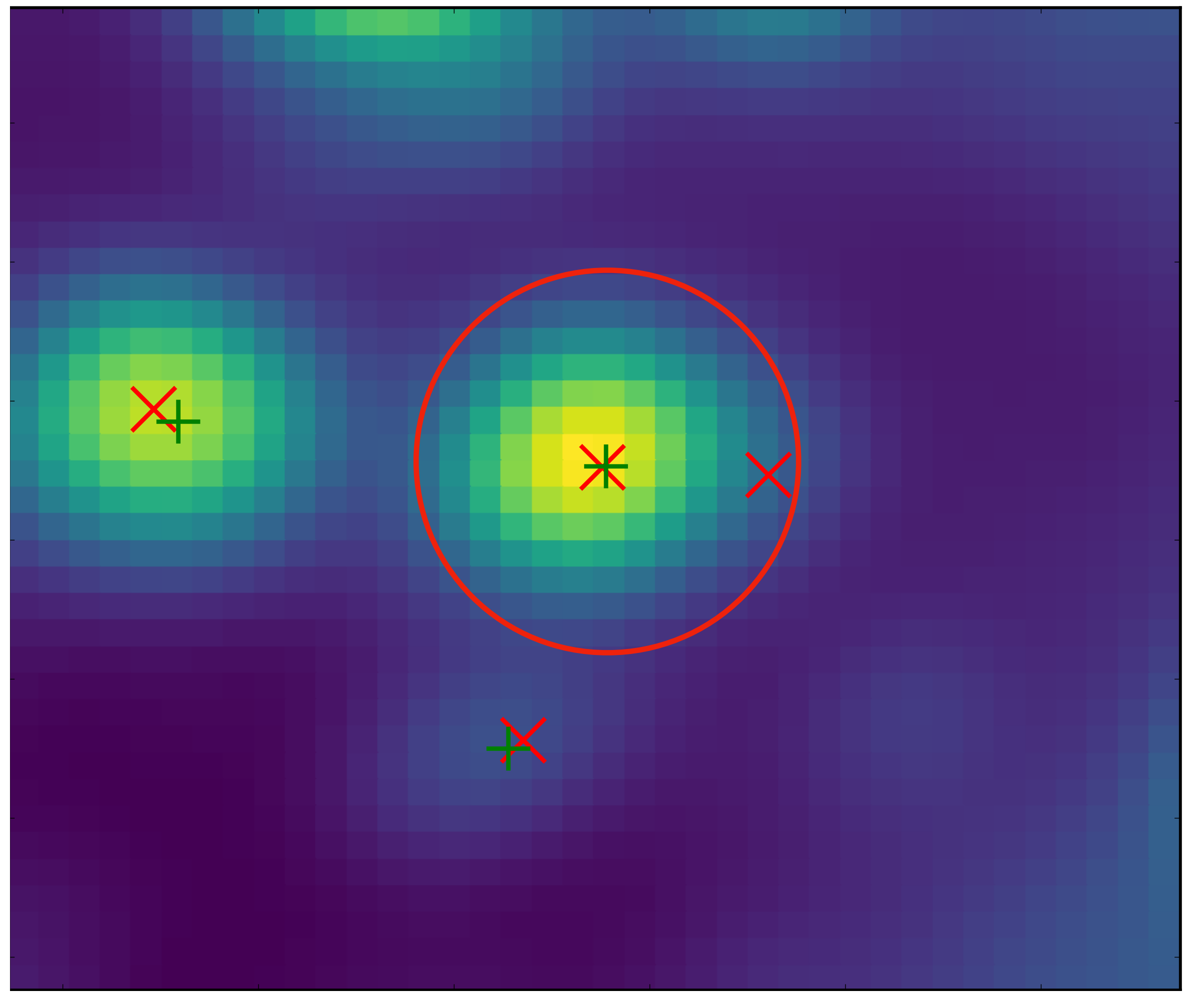
Gaia DR2 - Gaia Collaboration, Brown A. G. A., et al. (2018)

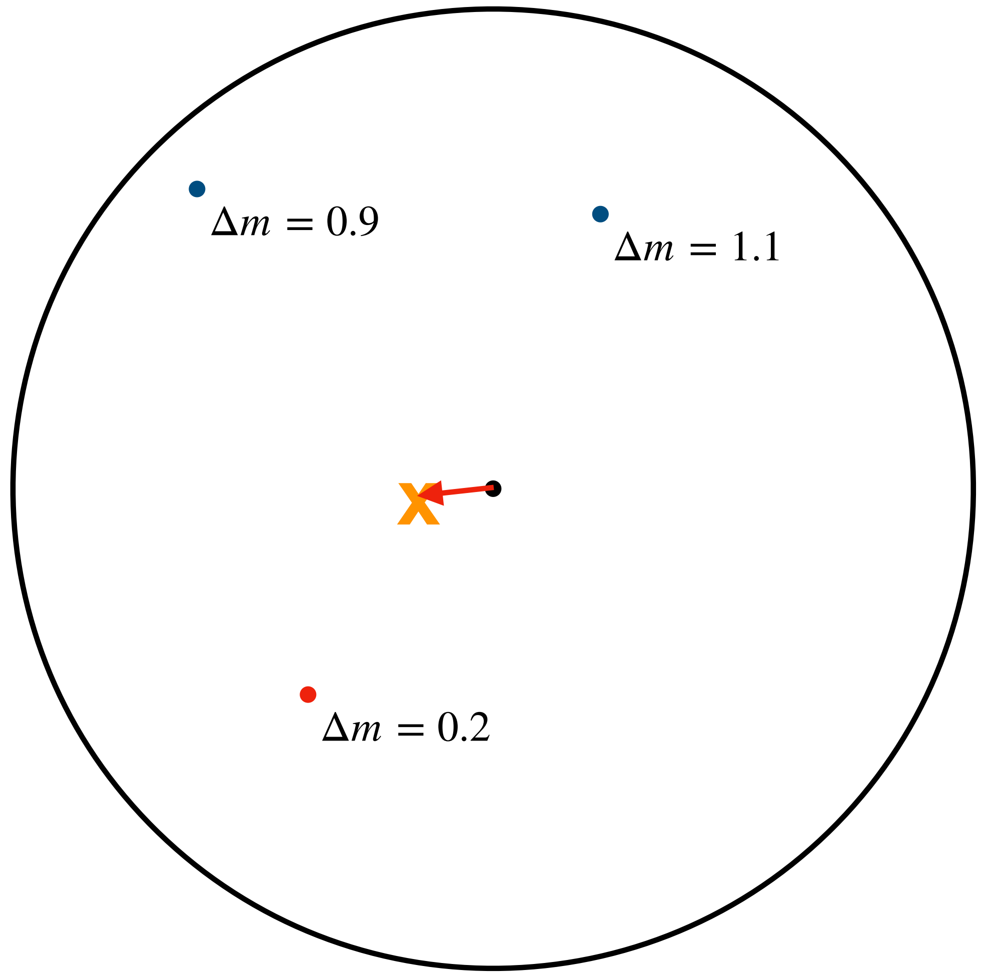
Tom J Wilson @onoddil

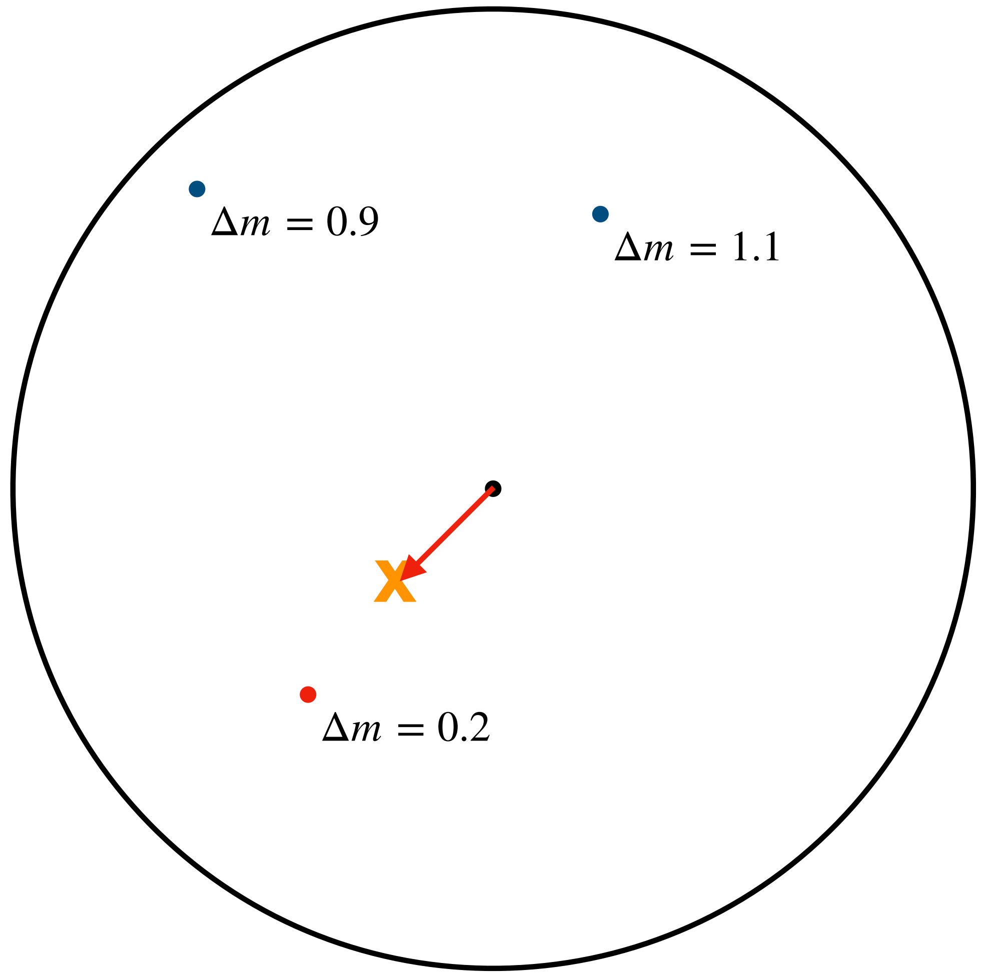
Modelling Crowded-Fields: Flux Brightening

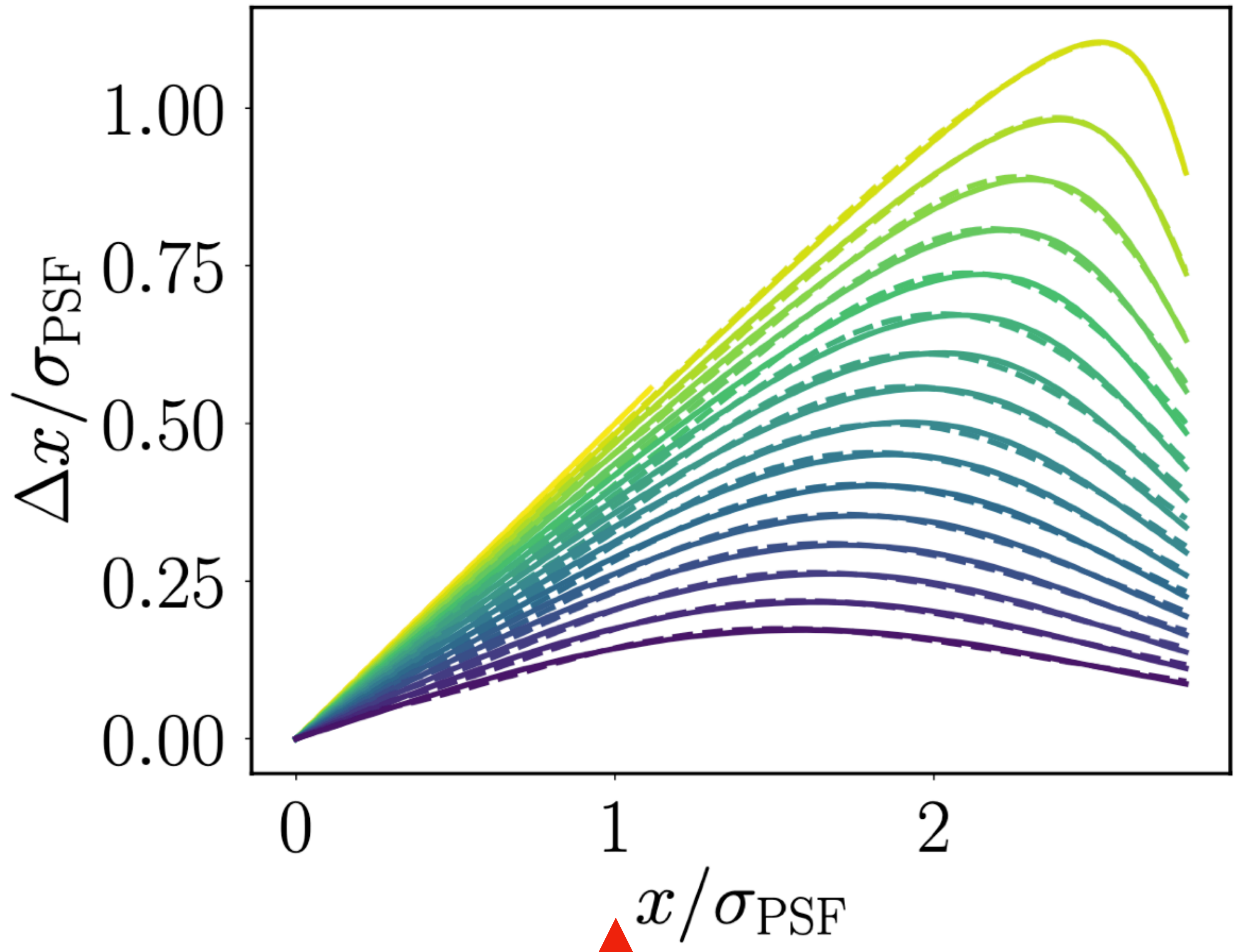
High SNR PSF or Aperture Photometry

Low SNR PSF Photometry







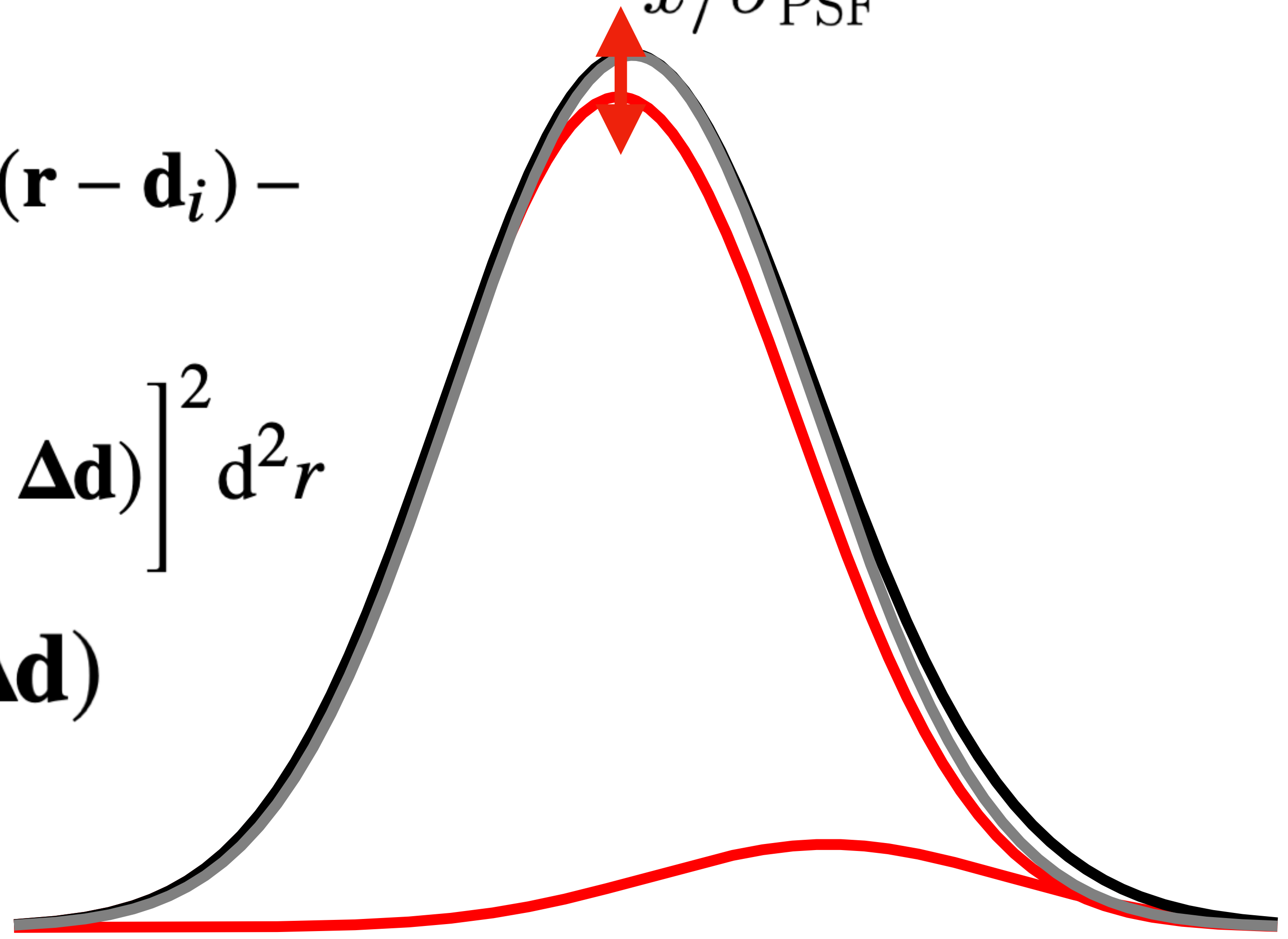


$$\Delta x = \frac{\sum_i f_i x_i}{1 + \boxed{\sum_i f_i}}$$

$$\Delta f = \sum_i f_i$$

$$\log \mathcal{L} = -\frac{1}{2} \times L \int_{-\infty}^{\infty} \left[\phi(\mathbf{r}) + \sum_i f_i \phi(\mathbf{r} - \mathbf{d}_i) - \boxed{(1 + \Delta f)\phi}(\mathbf{r} - \mathbf{\Delta d}) \right]^2 \mathrm{d}^2 r$$

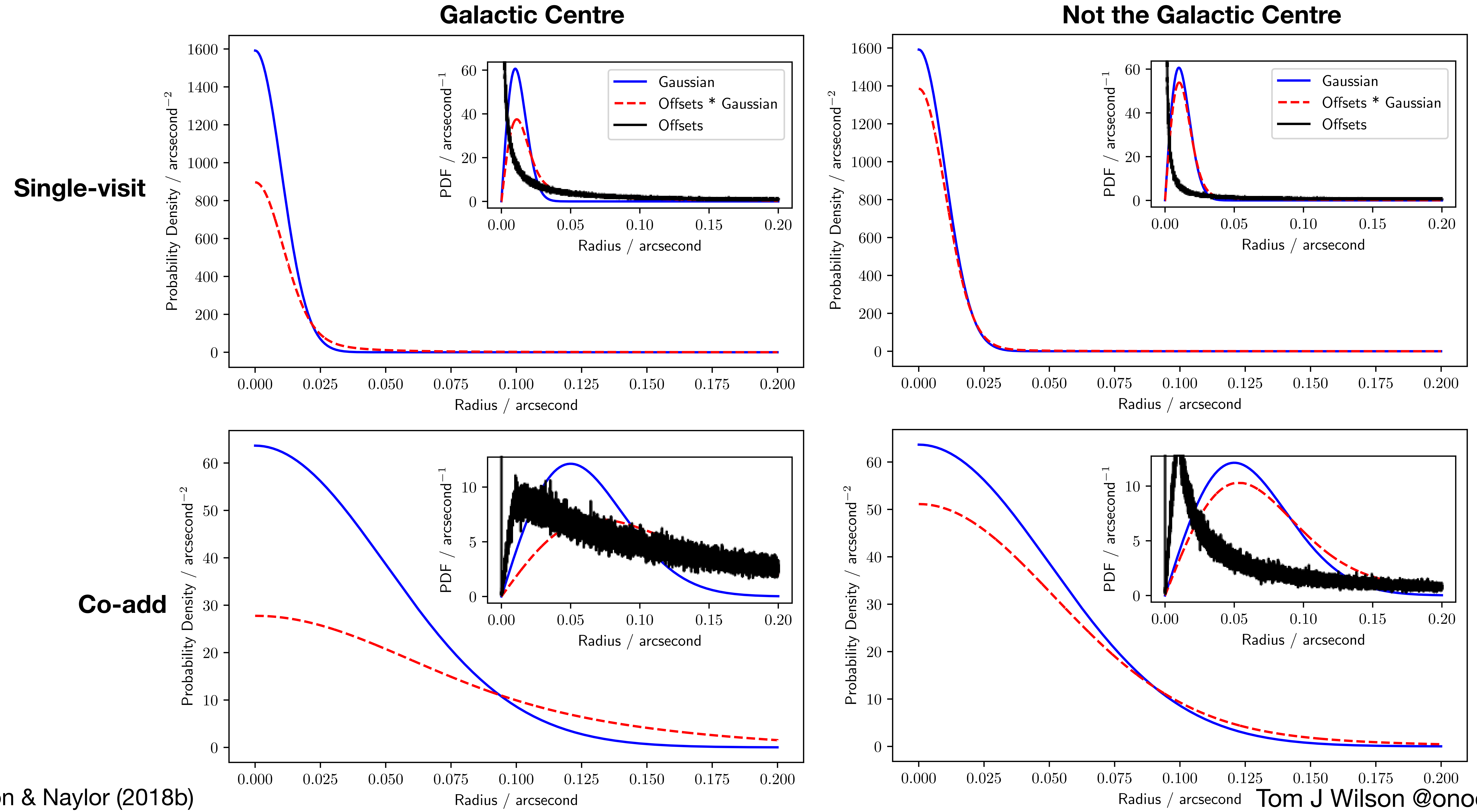
$$\Delta f = \psi'(\mathbf{\Delta d}) - 1 + \sum_i f_i \psi'(\mathbf{d}_i - \mathbf{\Delta d})$$



Wilson & Naylor (2018b; in prep.)
Plewa & Sari (2018)

Tom J Wilson @onoddil

The Rubin AUF

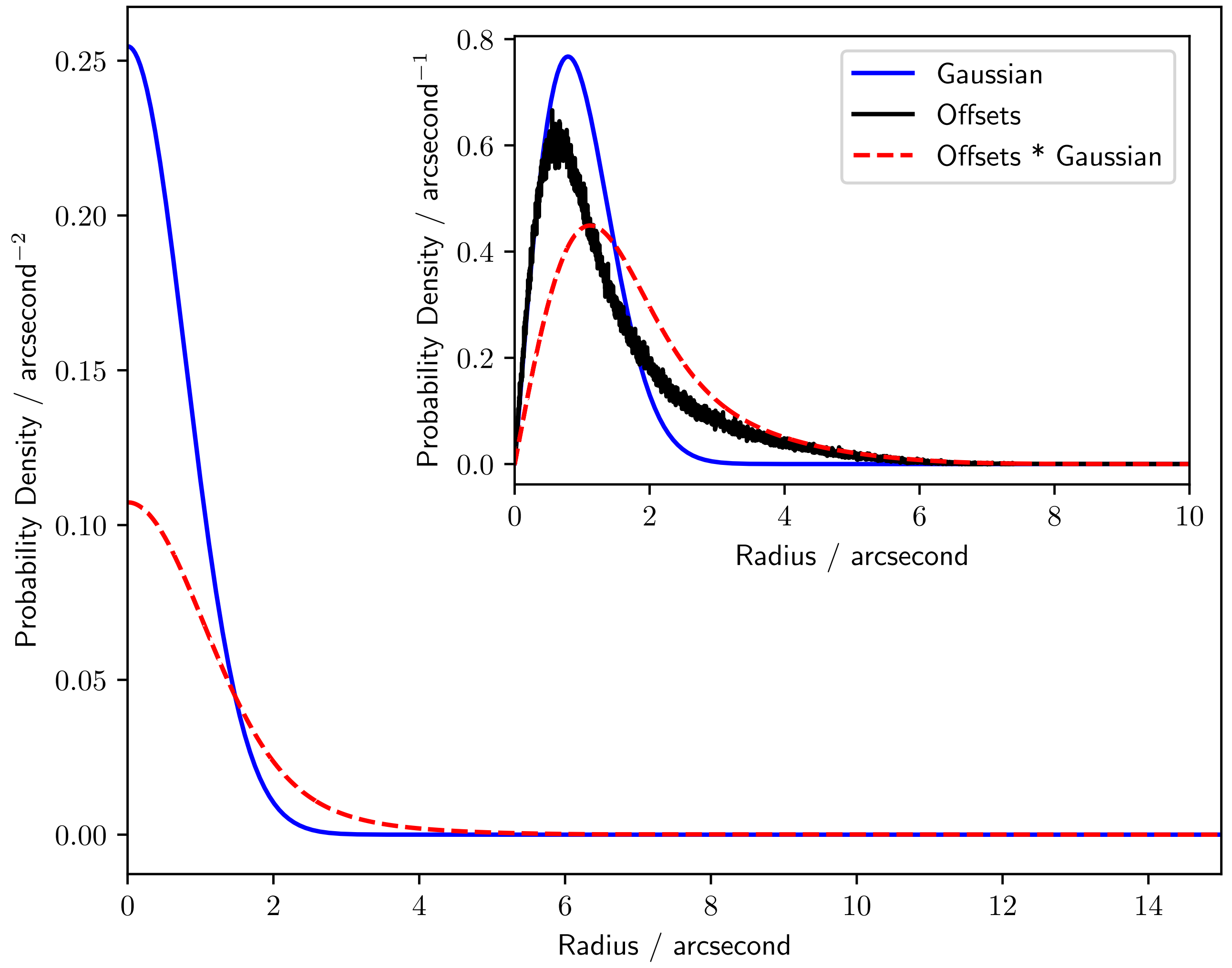


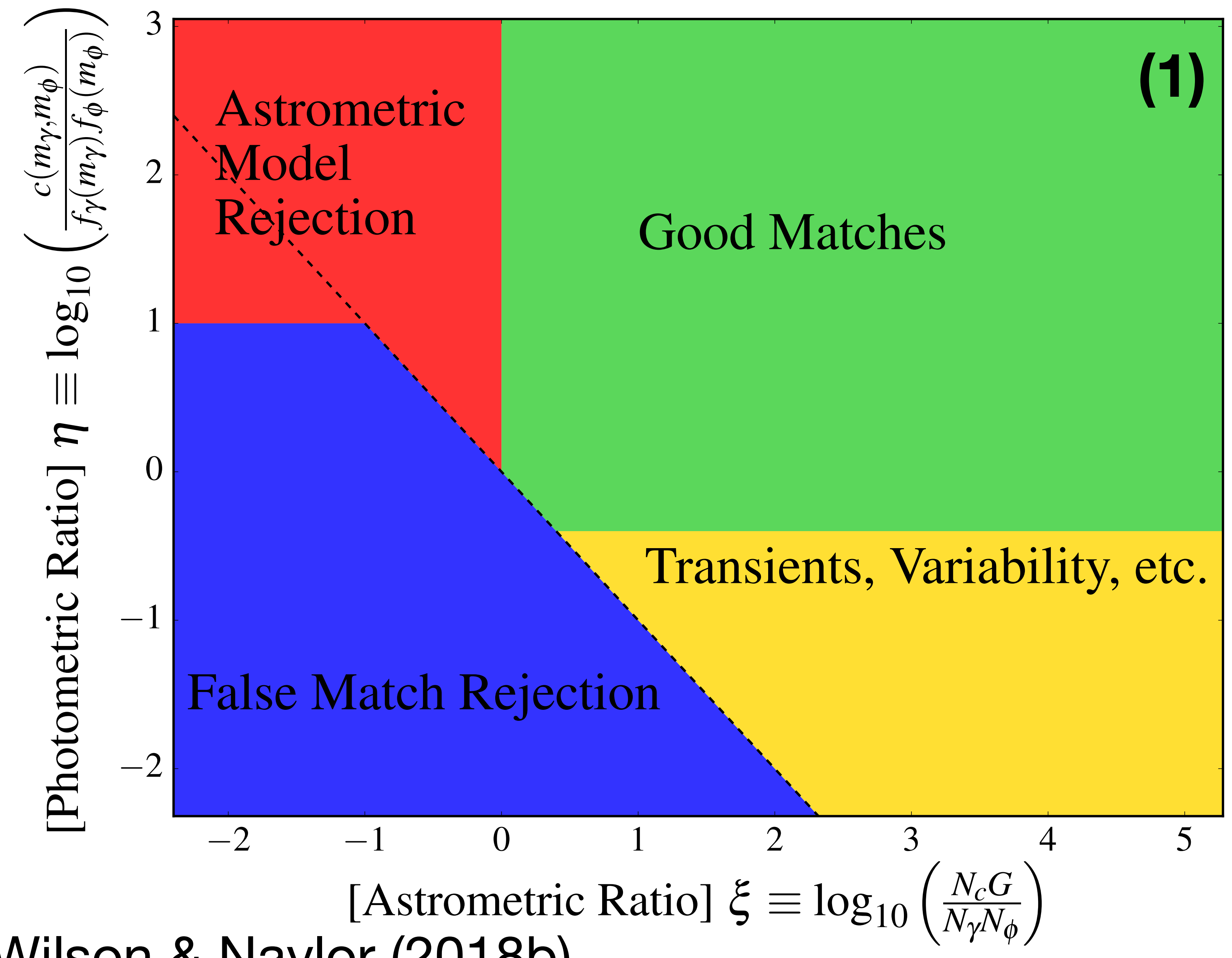
Wilson & Naylor (2018b)

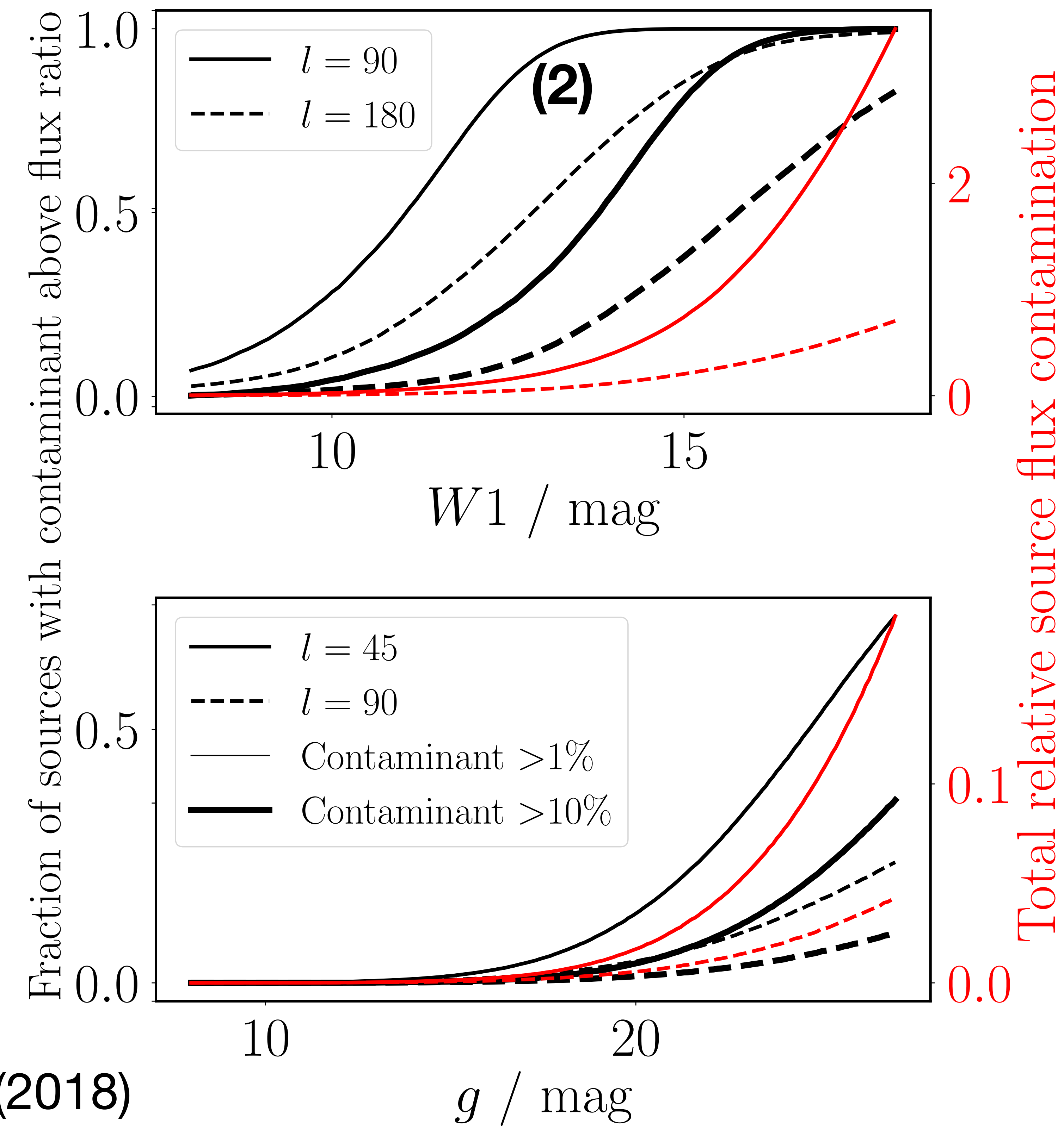
Tom J Wilson @onoddil

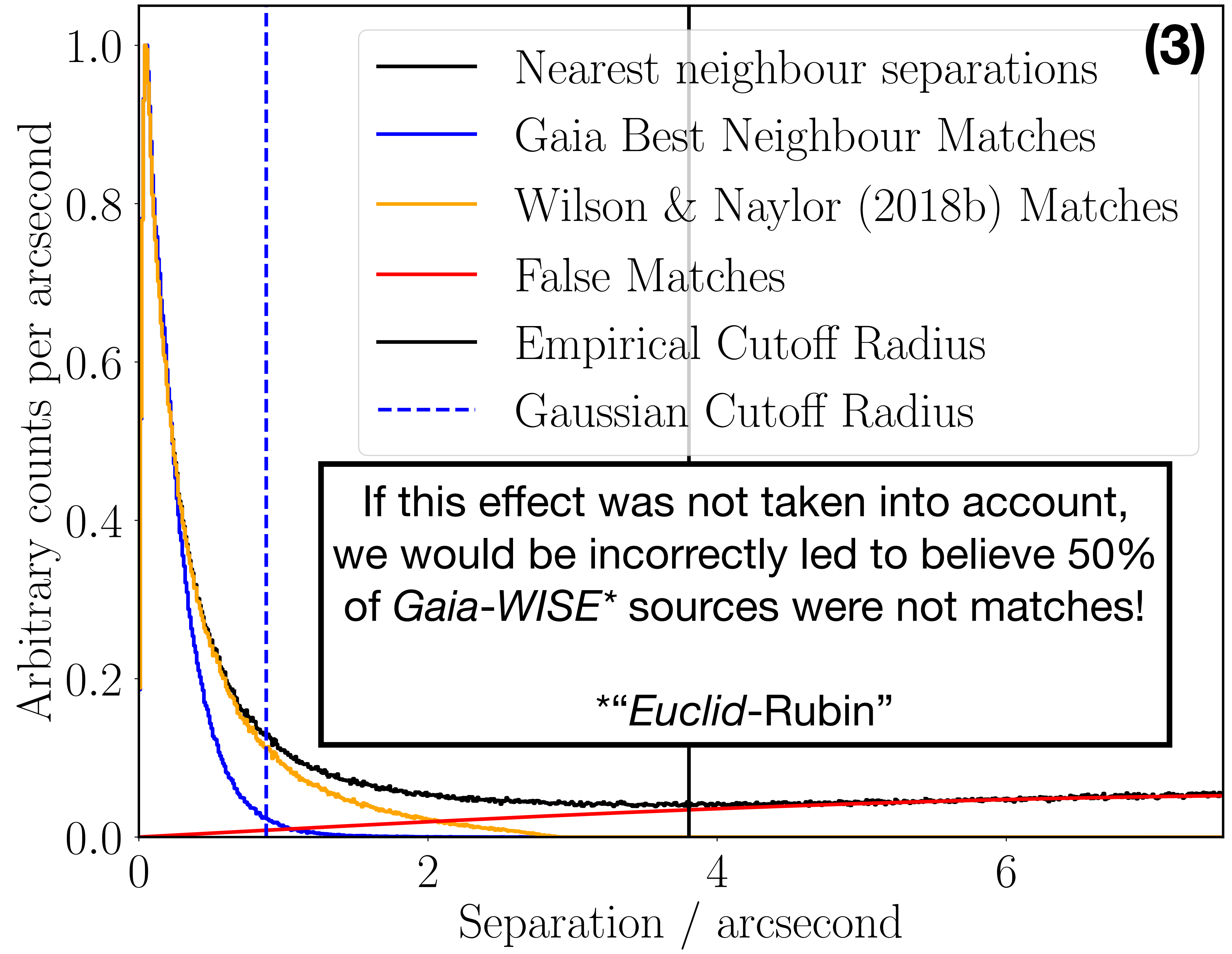
Why Use Our Cross-Matches (and Extensions)?

0) Getting cross-matches, even for “well behaved” fields
1) Finding “odd” objects, either using the inclusion vs non-inclusion of the photometry in the two match runs, or via the likelihood ratio space — separately-planned “real time” matching service for transient objects
2) Removing e.g. IR excess or correcting for extinction-like crowding brightening, through Average Contamination; crucial for “1% photometry” in both precision *and* accuracy
3) Recovering additional sources missed by other match services — either in crowded fields (we recover up to twice as many *Gaia-WISE* matches than the *Gaia* best neighbour matches), or with our extension to unknown proper motion modelling as an extra systematic









Wilson & Naylor (2018b)
WISE - Wright et al. (2010)
Gaia matches - Marrese et al. (2019)
Gaia DR2 - Gaia Collaboration, Brown A. G. A., et al. (2018)

Tom J Wilson @onoddil

The Rubin “Super-Match”

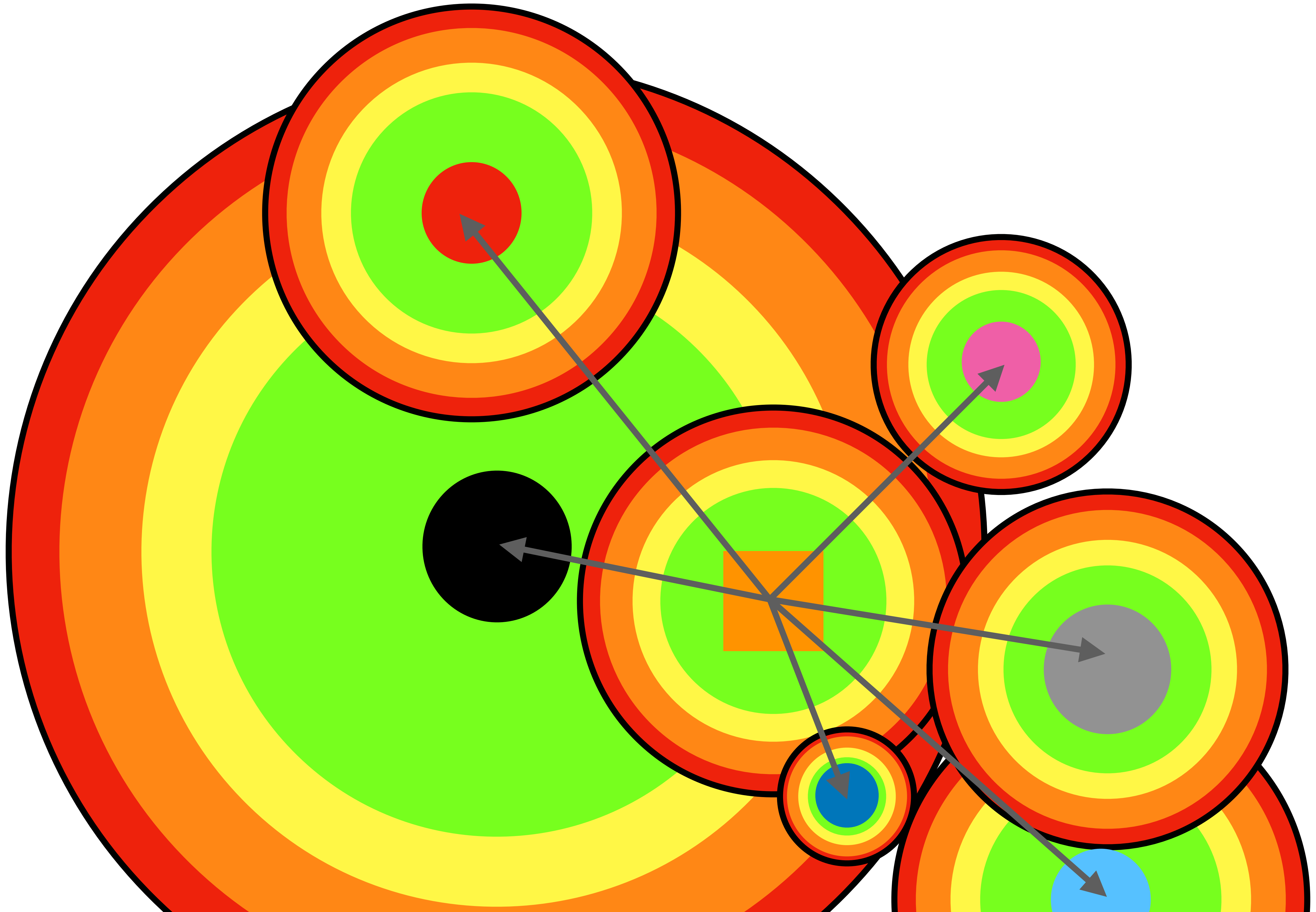
Bringing Independent Results together to Notify of Associations across Multiple catalogues

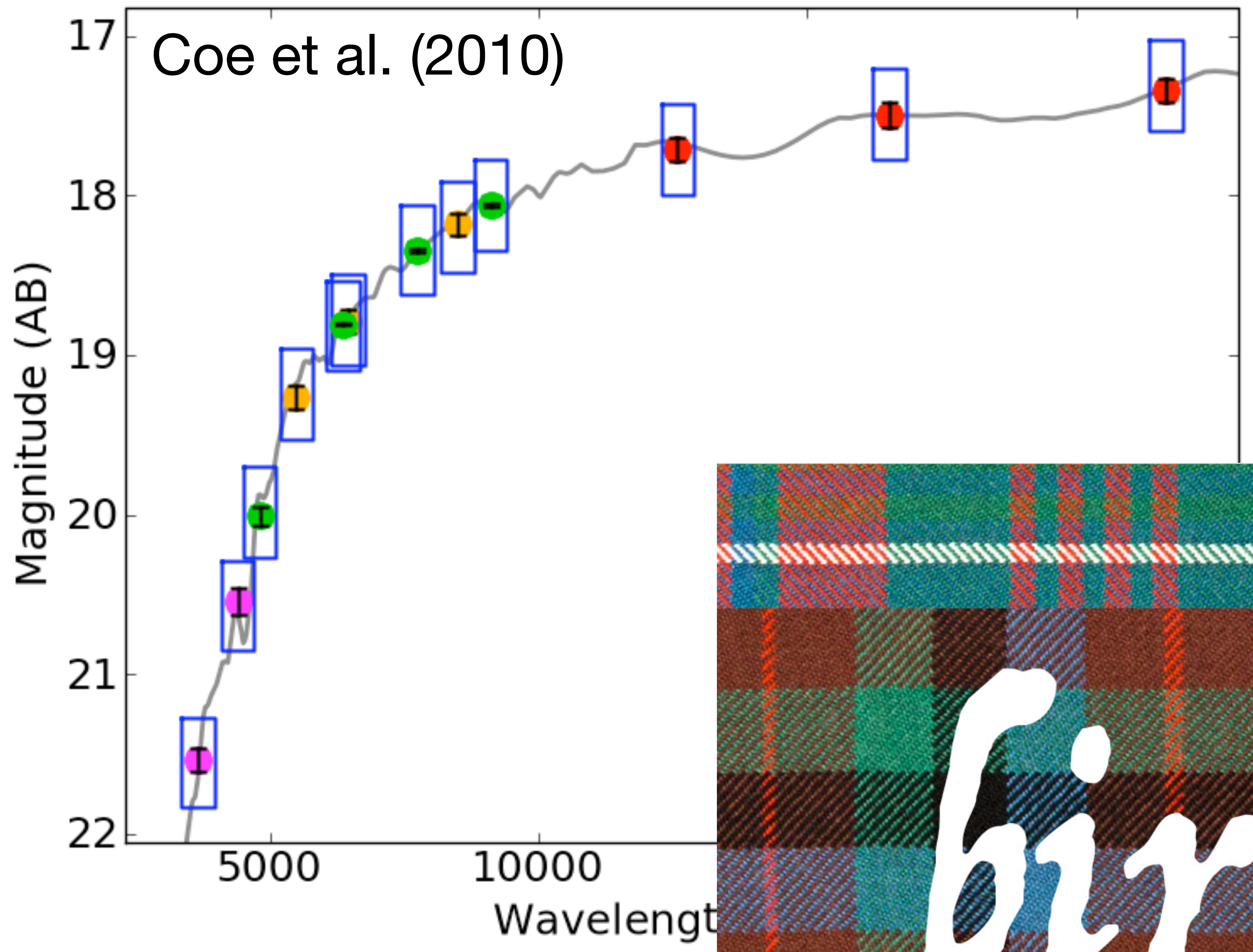
LSST -> *Gaia*, *WISE*, VISTA, *Euclid*, SDSS, … matches

Quick and easy construction of spectral energy distributions for each LSST source

Includes SED probabilities, individual match reliability, contamination statistics etc.

Create new composite dataset to host on the UK IDAC, as a second IKC.





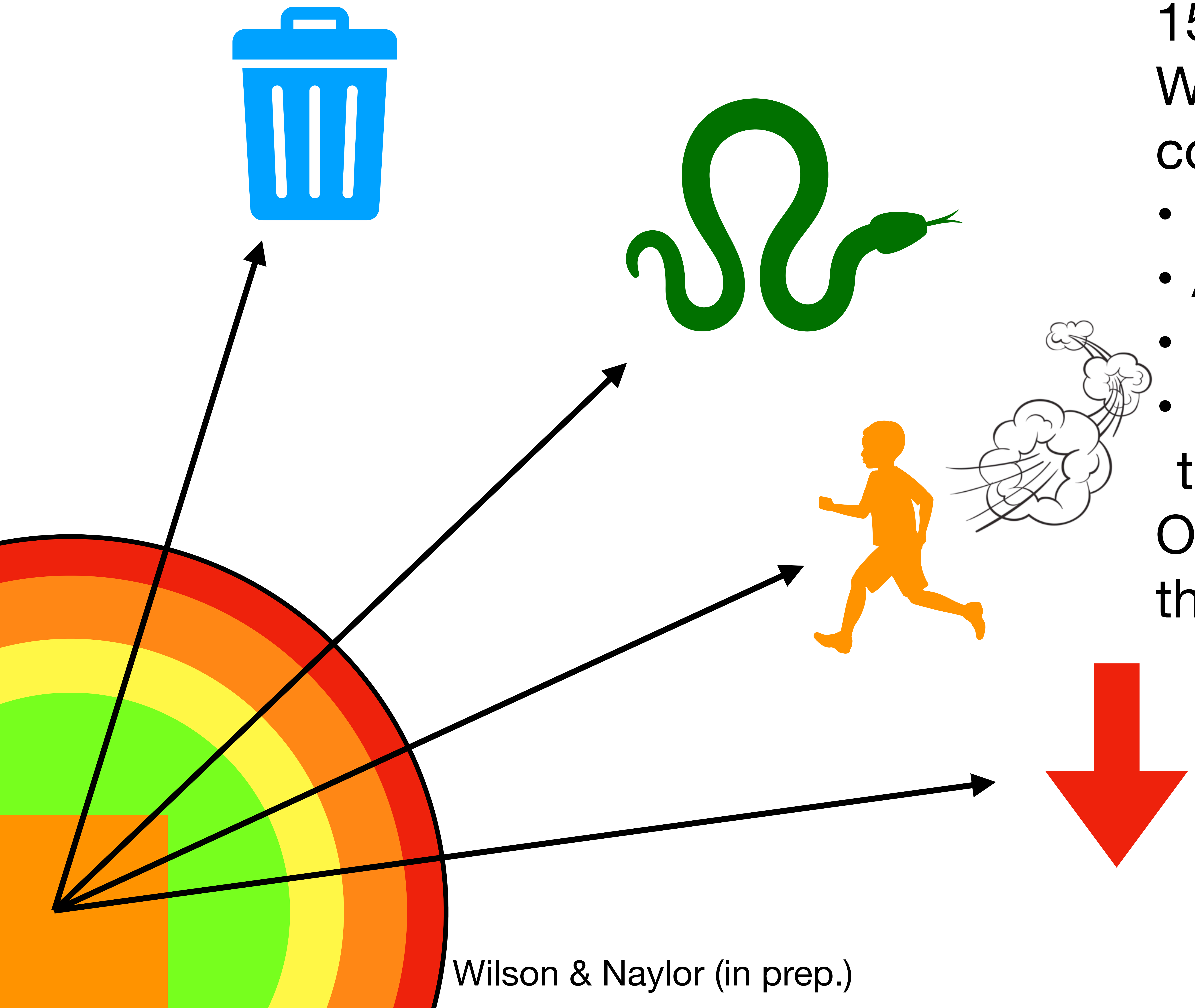
Wilson & Naylor (in prep.)
Pineau et al. (2017)

https://github.com/macauff/birnam

Tom J Wilson @onoddil

Confirming Lonely Rubin Sources

Blanks And Near-misses, Questionable sources, Upper-limits, and Objects of varying brightness



Most LSST sources will be “lonely” with 15x as many sources as the next dataset. We will follow up all non-matches, and confirm whether these objects are:

- Image artefacts
- Astrophysically variable objects
- High proper motion sources
- Regular objects that are simply too faint to be seen in the opposing catalogue

Once again, datasets would be hosted on the UK IDAC.

banquo

Wilson & Naylor (in prep.)

Tom J Wilson @onoddil

Re-Focussing UKD-S9 Efforts

We have (mostly) completed two of our stated packages, macauff and birnam. However, due to a lack of Rubin testing and timescales for effort (i.e., me!), we are considering what to do with the time that is given to us, and whether it would be better to descope banquo and focus on:

- Improving documentation — a few people other than me have tried to use the code, and are surprised at how much prep work is required. We can decrease barriers to entry and allow the community to pick up the codes after dedicated effort ceases.
- Running a large-scale “legacy” survey cross-match database for crowded fields. This would shake out the code (and the Rubin Science Platform, from which users would obtain the matches) but also provide useful early science for the community.
- Starting up a cross-match coordination group to optimise efforts and ensure algorithms are correctly paired with the right catalogues.

At this stage we are looking to get opinions from the recipient groups — i.e., you lot — on whether people feel it is better that we offer the “missing sources” package instead of ensuring our core products are vetted ahead of operations.