Enabling Early Rubin Science with Robust Cross-Matches in the Crowded LSST Sky

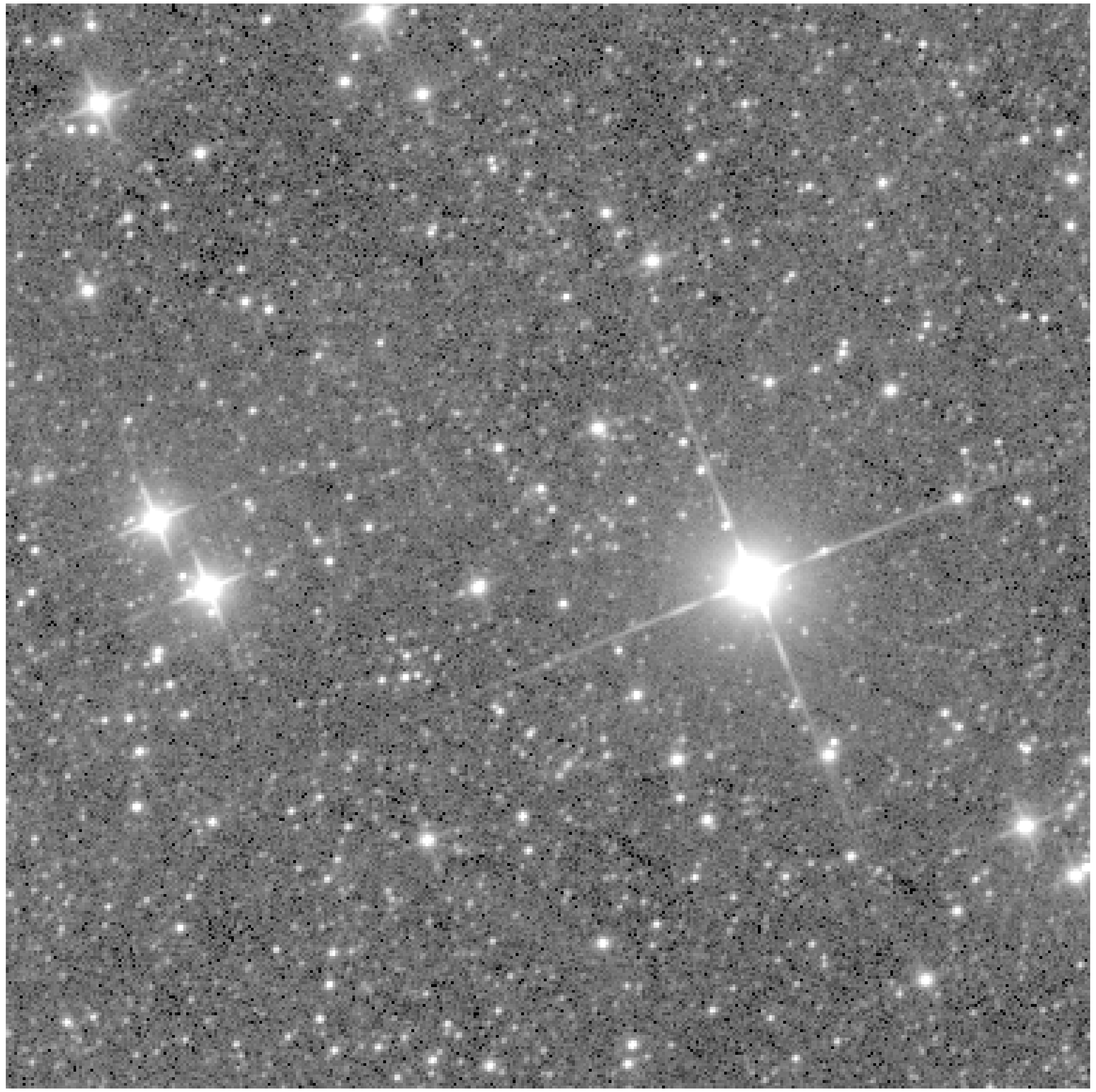
Tom J Wilson (he/him) and Tim Naylor
t.j.wilson@exeter.ac.uk
University of Exeter

UNIVERSITY OF EXETER

LSST:UK Consortium

VERA C. RUBIN OBSERVATORY

Photometric Observations

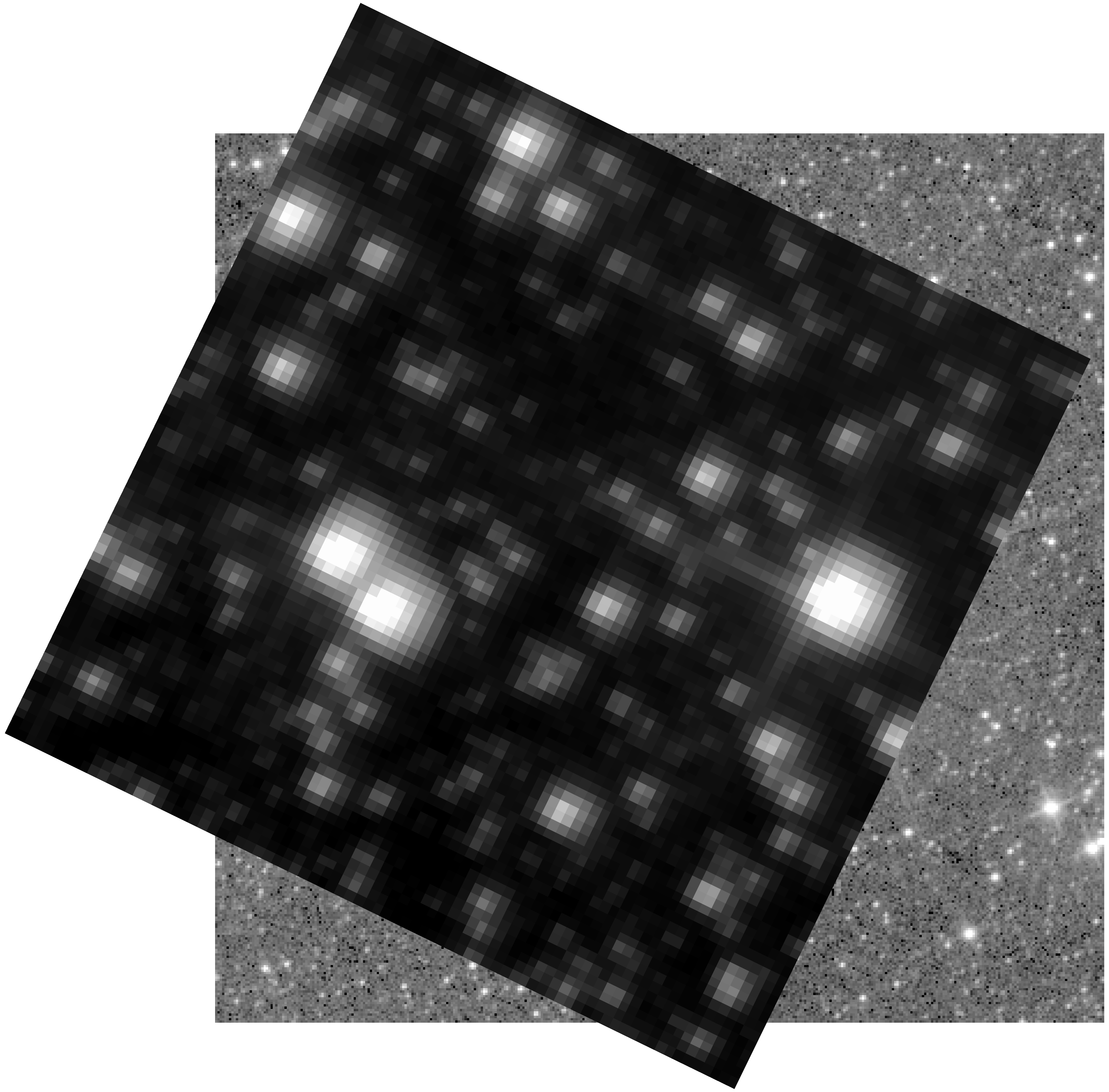


WISE - Wright et al. (2010)

WISE W1

Tom J Wilson @onoddil

Photometric Observations

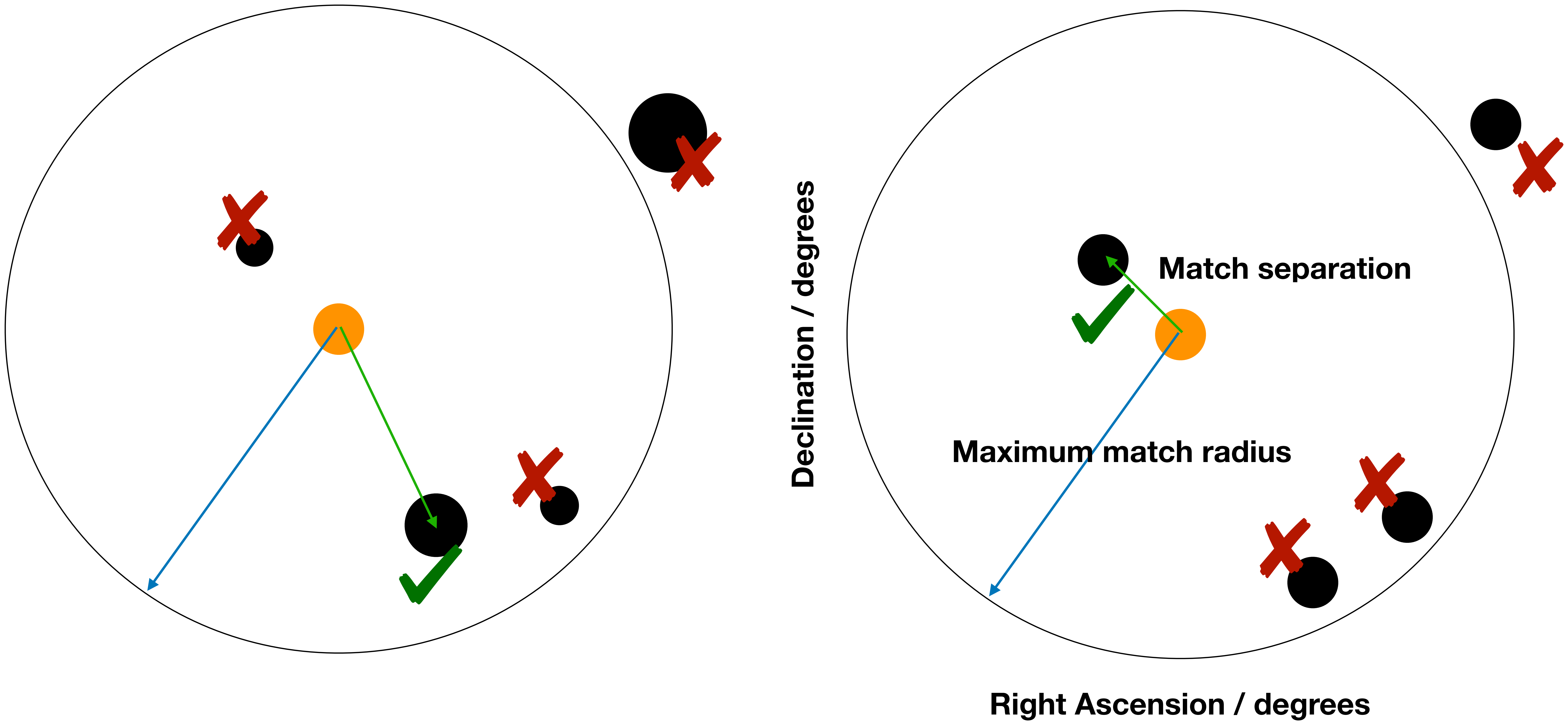


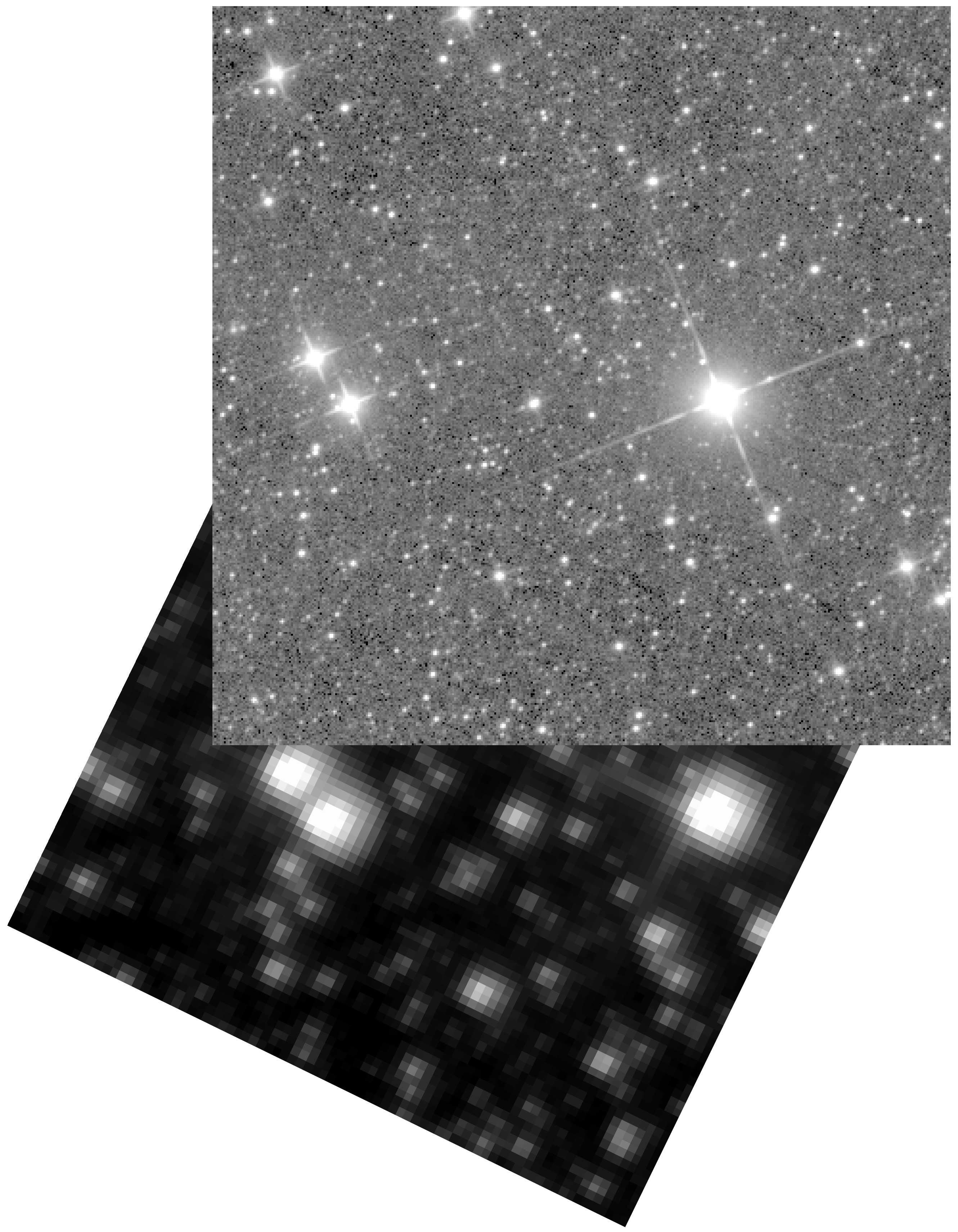
WISE - Wright et al. (2010)
TESS - Ricker et al. (2015)

TESS T

Tom J Wilson @onoddil

“Traditional”/“Simple” Cross-Matching

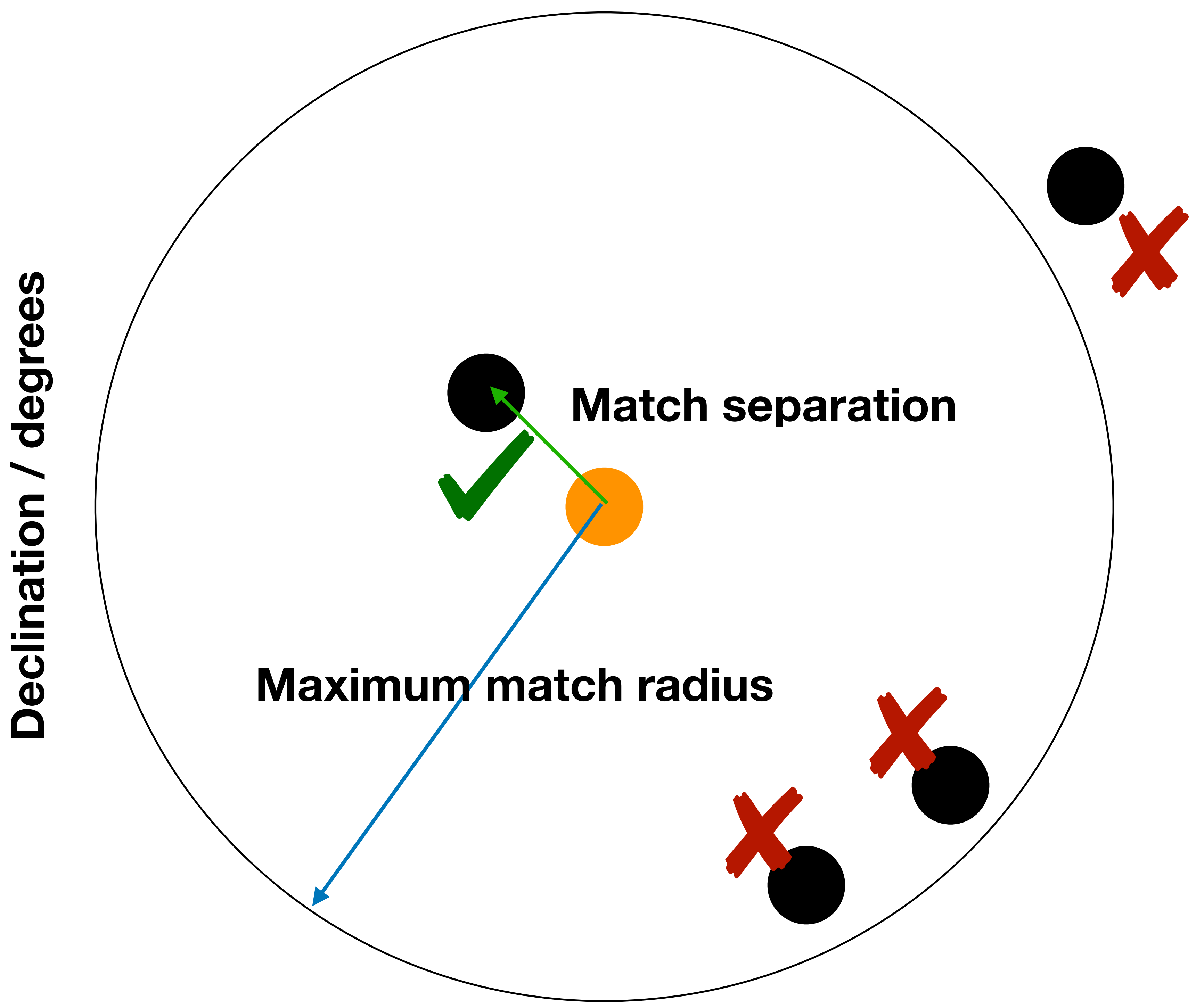


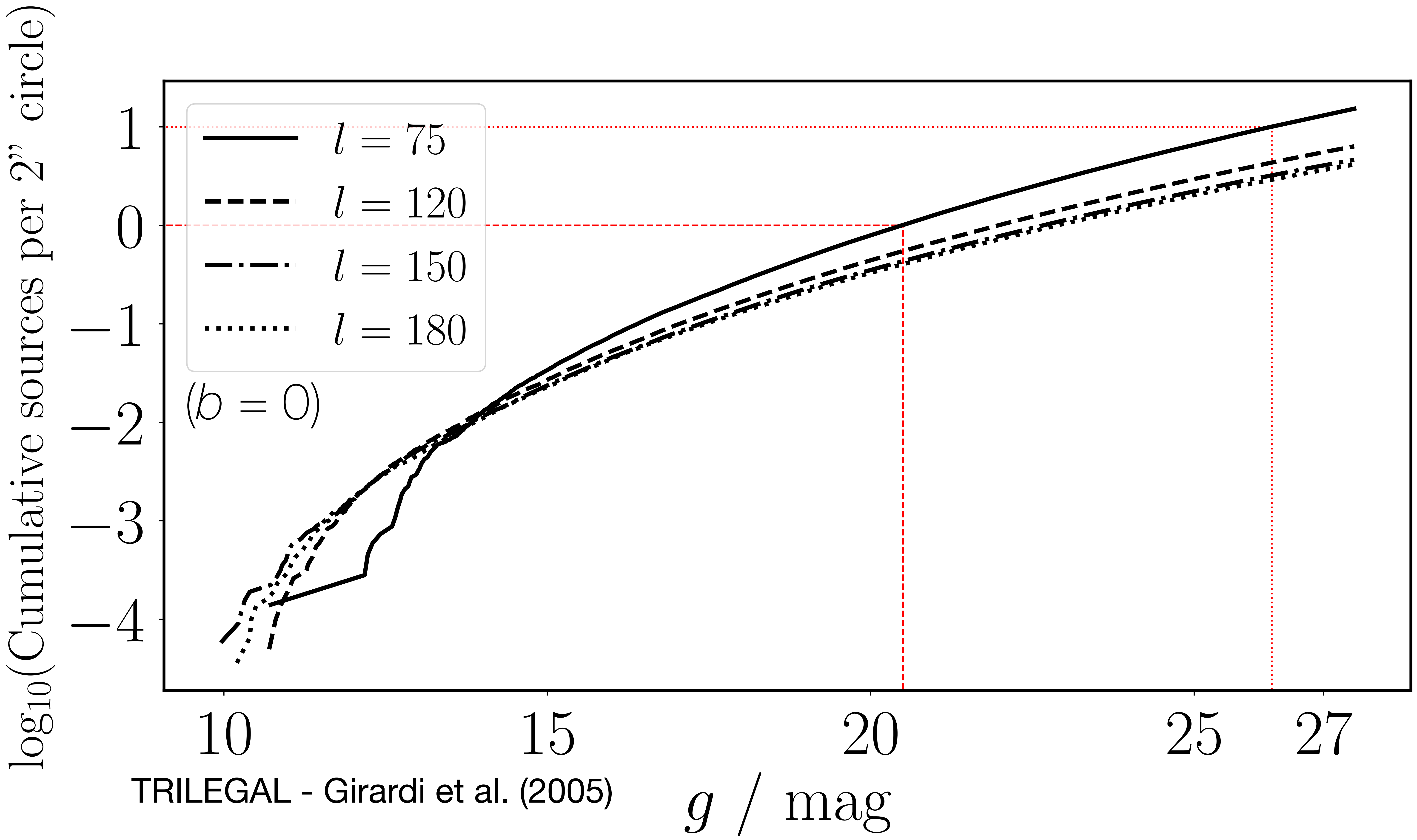


Tom J Wilson @onoddil

The Problem With Rubin Obs.'s LSST

Nearest neighbour/ proximity matching

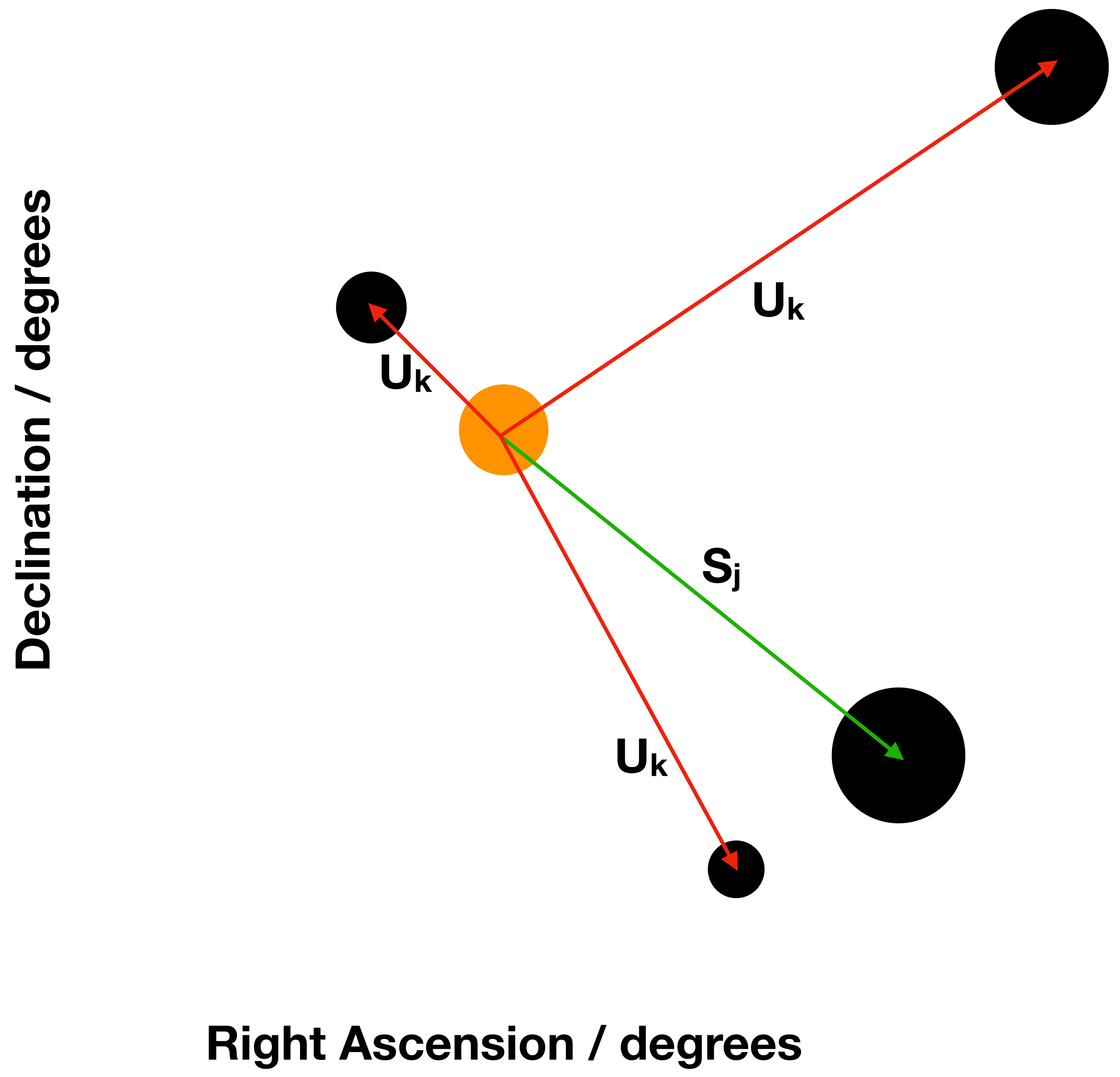




TRILEGAL - Girardi et al. (2005)

Tom J Wilson @onoddil

Probabilistic Cross-Matching



$$L = \frac{q(m, c)\, f(x, y)}{n(m, c)}$$

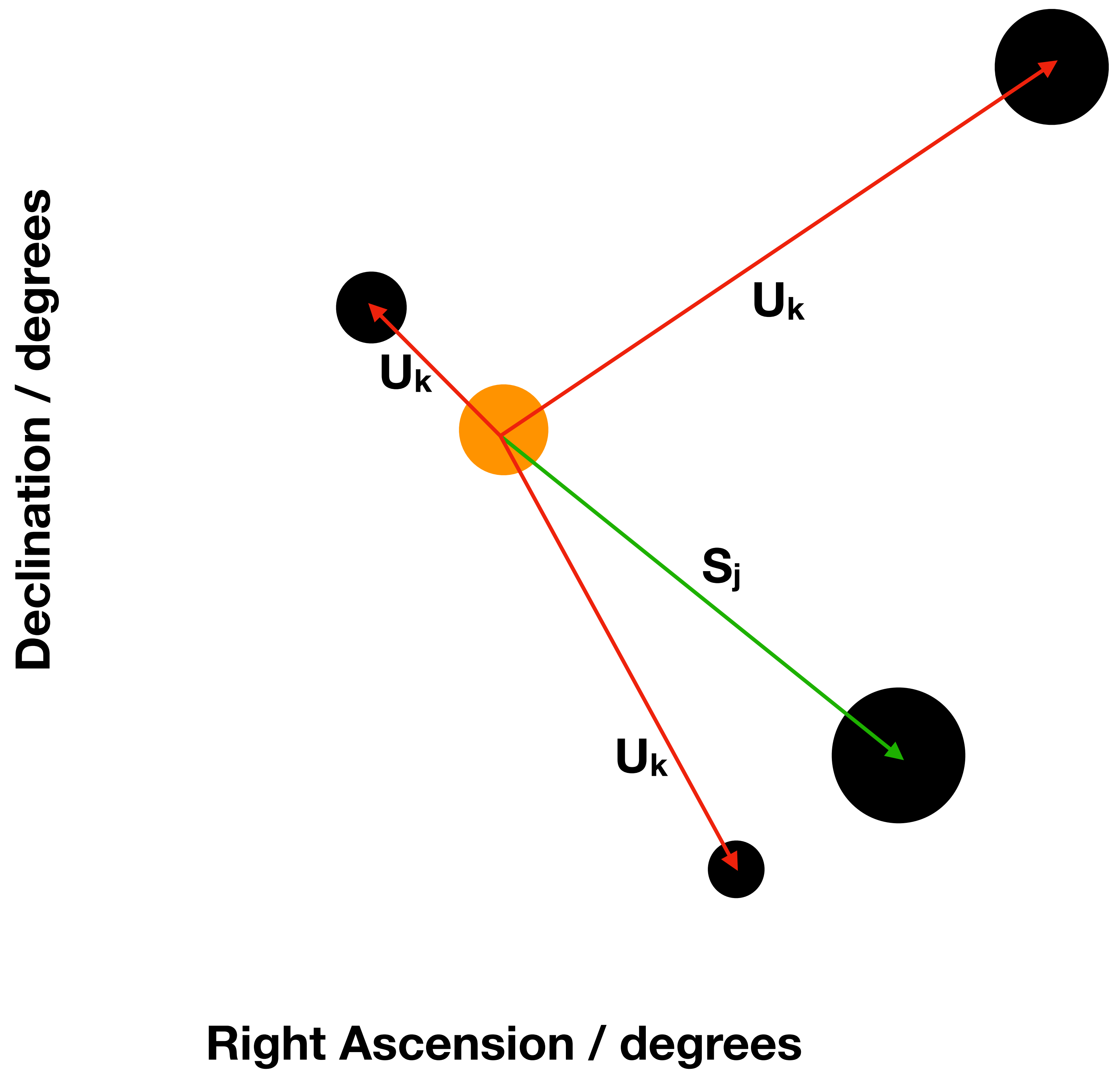
$$R_j = \frac{\Pr\left[S_j \cap \left(\bigcap_{k \neq j} U_k\right) \cap \left(\bigcap_{k'} E_{k'}\right)\right]}{\sum_i \Pr\left[S_i \cap \left(\bigcap_{k \neq i} U_k\right) \cap \left(\bigcap_{k'} E_{k'}\right)\right] + \Pr\left[(m_S > m_{\mathrm{lim}}) \cap \left(\bigcap_{k} U_k\right) \cap \left(\bigcap_{k'} E_{k'}\right)\right]} = \frac{L_j}{\sum_i L_i + (1 - Q)}$$

Sutherland & Saunders (1992)

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

Wilson & Naylor (2018a)

Probabilistic Cross-Matching



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

Probability of sources having their brightnesses given they are unrelated to one another ("field stars")

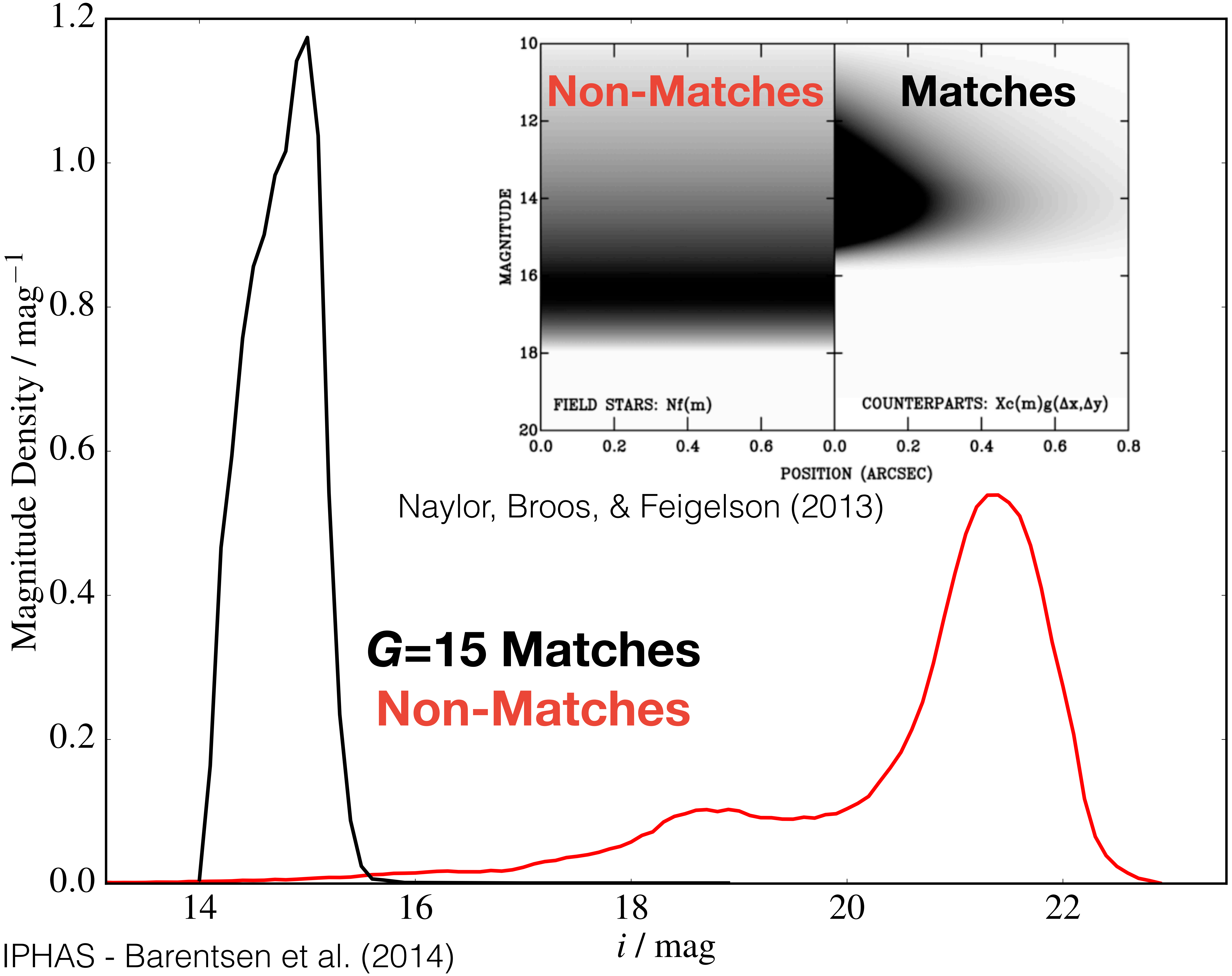
Probability of sources having their brightnesses given they are counterparts

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

Wilson & Naylor (2018a)

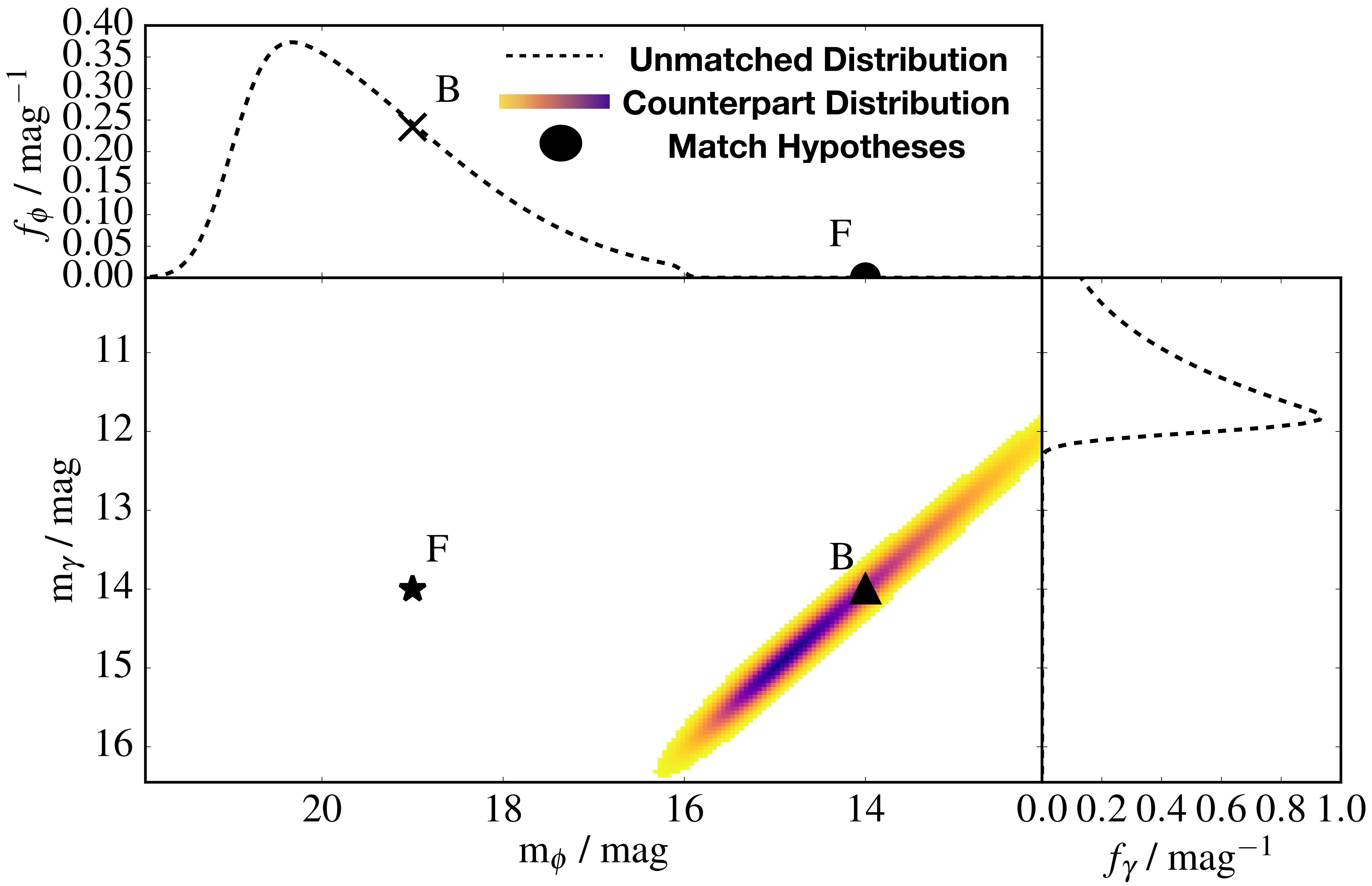
Including Magnitude Information

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



IPHAS - Barentsen et al. (2014)

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

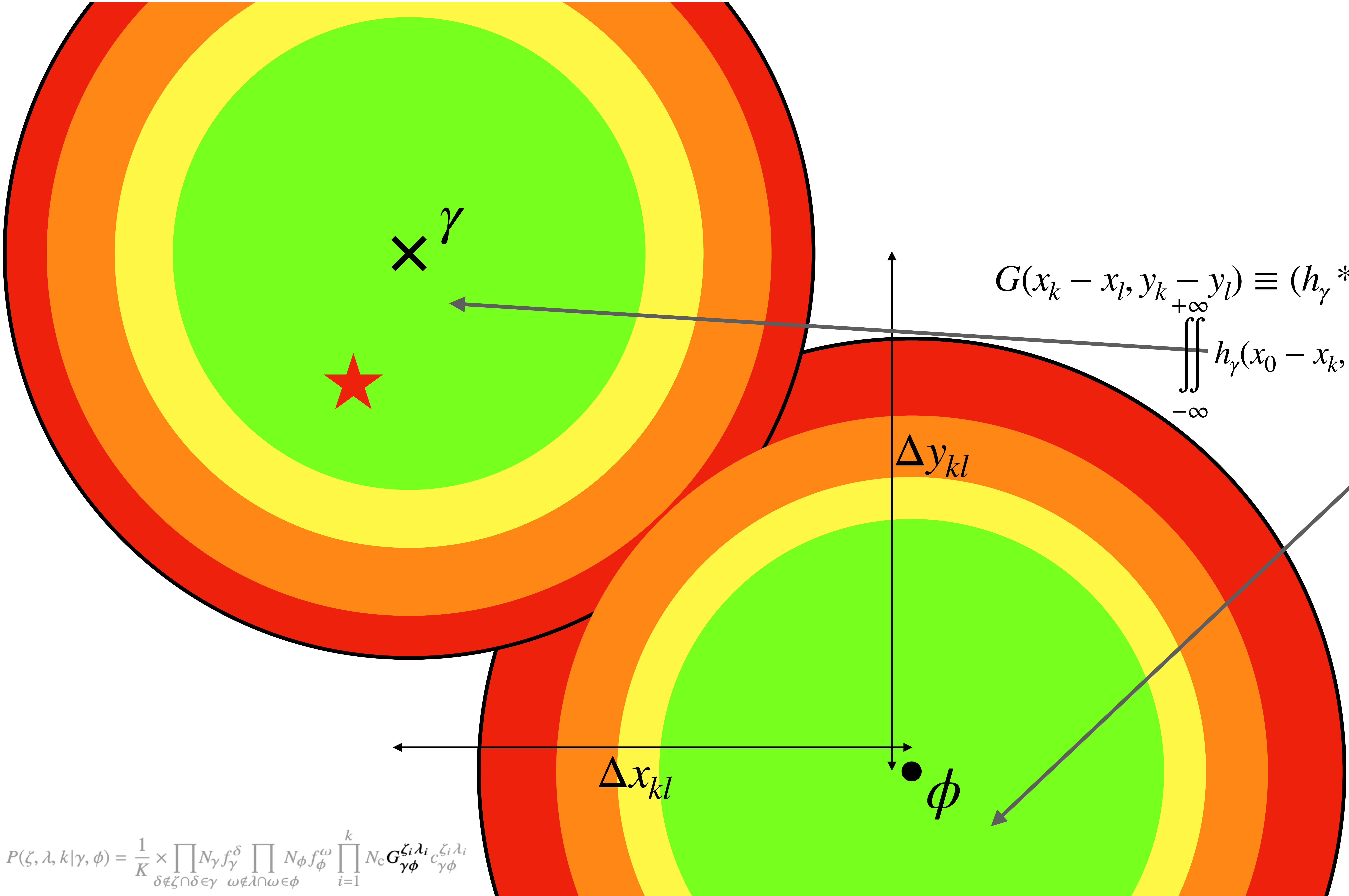


Wilson & Naylor (2018a)

The photometry-based likelihoods (c and f) allow us to reject some matches in crowded fields, but now we need the position-based likelihood G

Tom J Wilson @onoddil

Match Separation Probability



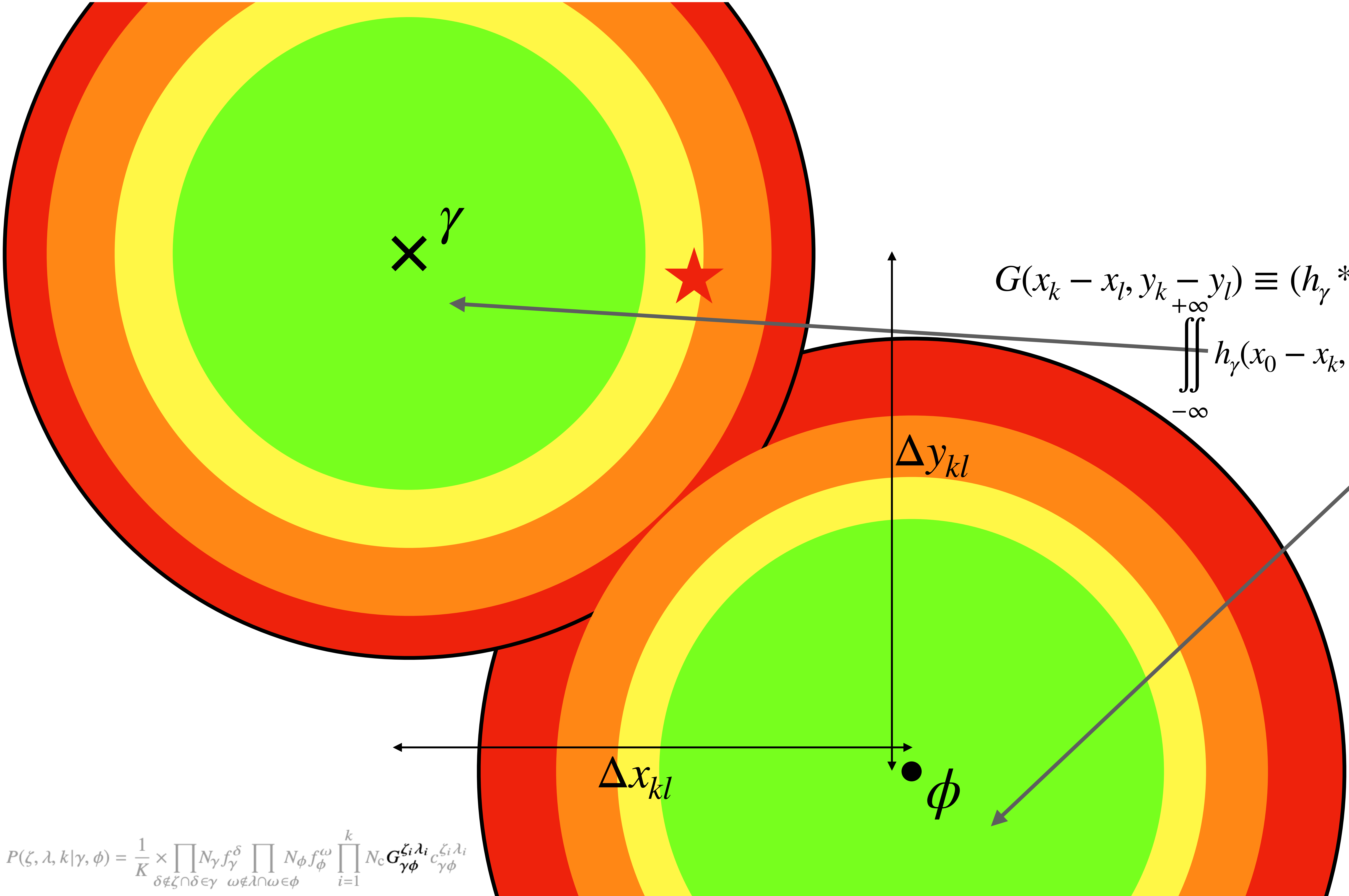
$$G(x_k - x_l, y_k - y_l) \equiv (h_\gamma * h_\phi)(\Delta x_{kl}, \Delta y_{kl}) = \iint_{-\infty}^{+\infty} h_\gamma(x_0 - x_k, y_0 - y_k) h_\phi(x_l - x_0, y_l - y_0)\, \mathrm{d}x_0\, \mathrm{d}y_0$$

Wilson & Naylor (2018a)

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

Tom J Wilson @onoddil

Match Separation Probability



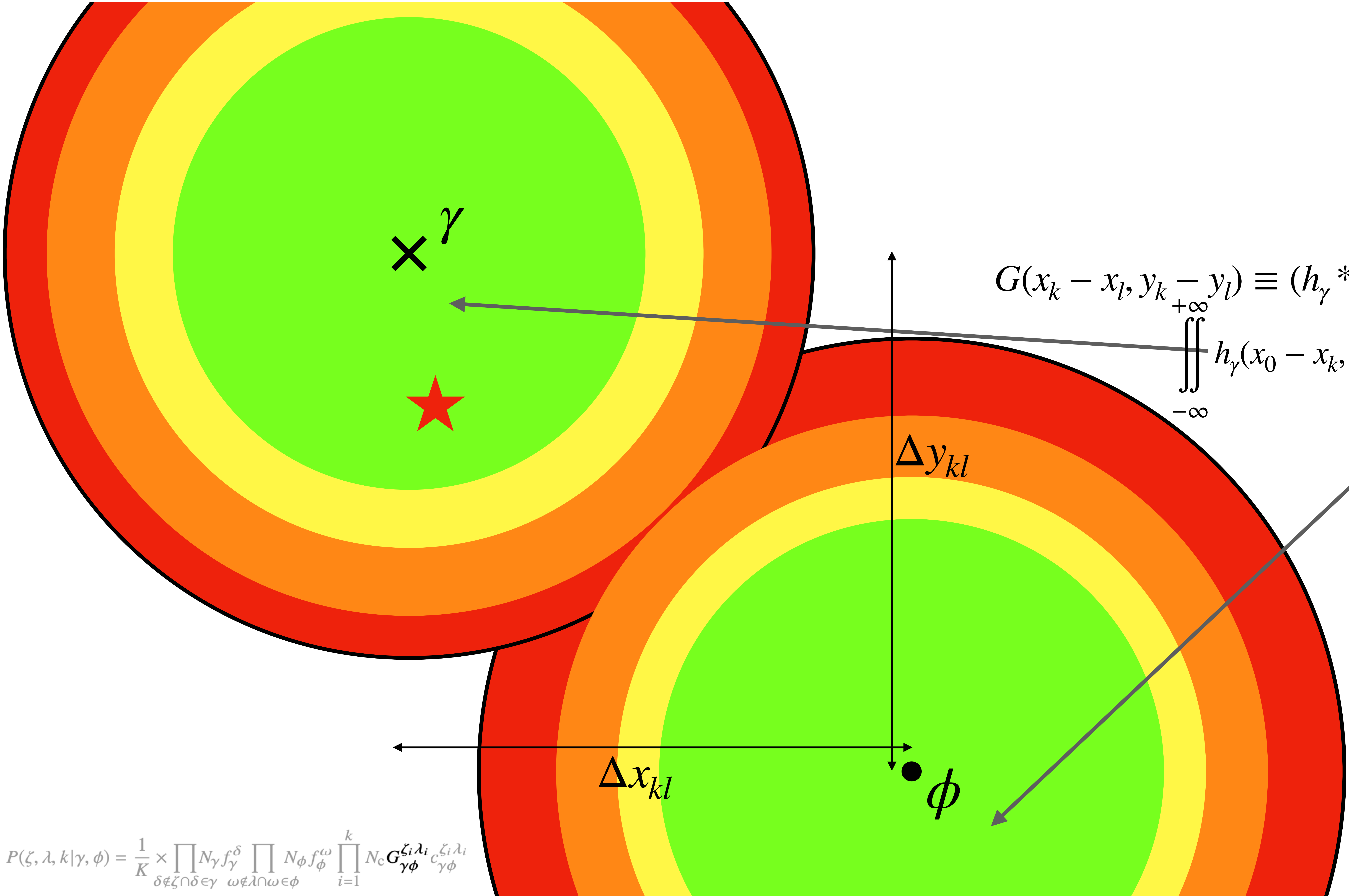
$$G(x_k - x_l, y_k - y_l) \equiv (h_\gamma * h_\phi)(\Delta x_{kl}, \Delta y_{kl}) = \iint_{-\infty}^{+\infty} h_\gamma(x_0 - x_k, y_0 - y_k) h_\phi(x_l - x_0, y_l - y_0)\, \mathrm{d}x_0\, \mathrm{d}y_0$$

Wilson & Naylor (2018a)

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

Tom J Wilson @onoddil

Match Separation Probability



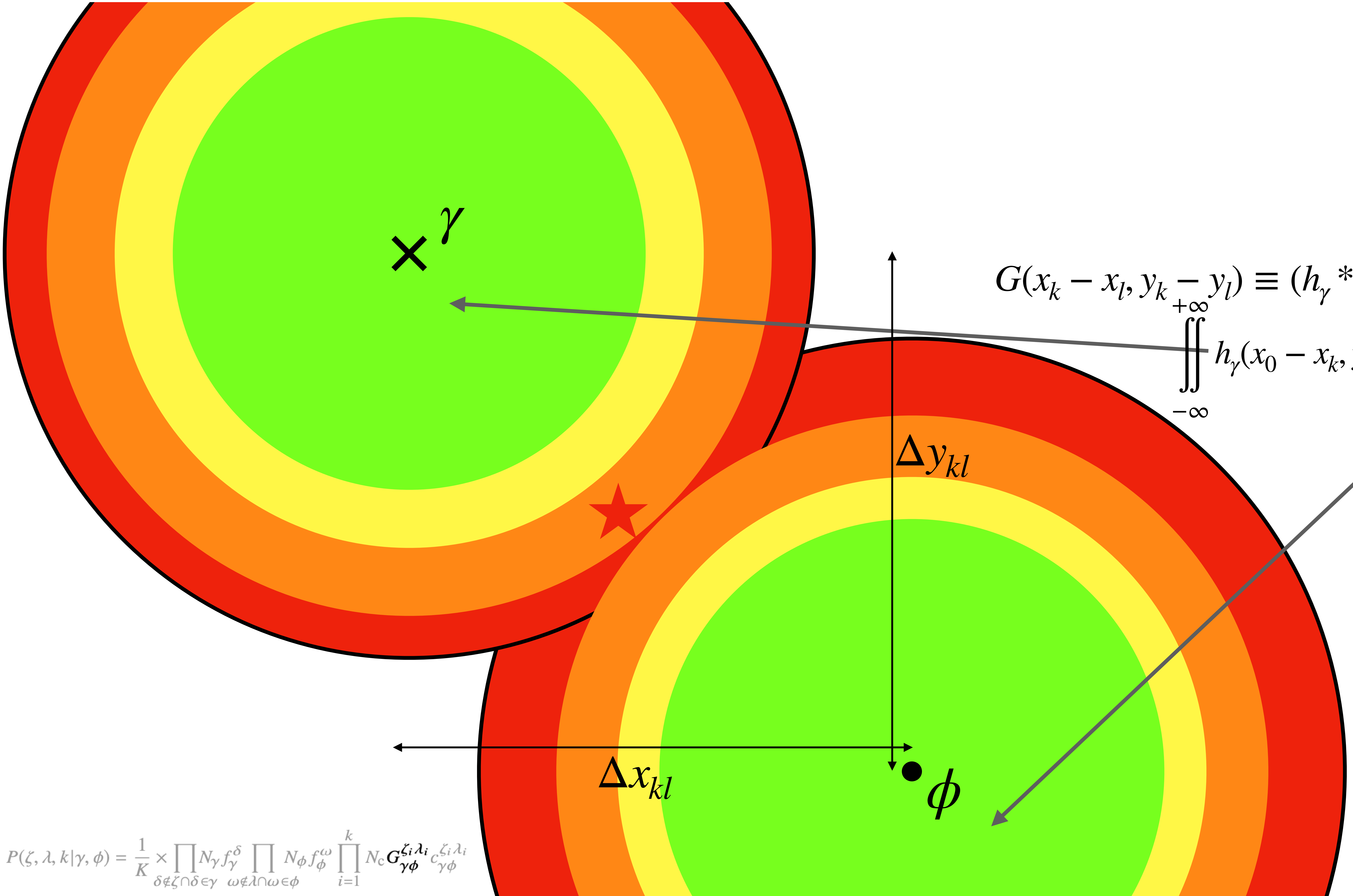
$$G(x_k - x_l, y_k - y_l) \equiv (h_\gamma * h_\phi)(\Delta x_{kl}, \Delta y_{kl}) = \iint_{-\infty}^{+\infty} h_\gamma(x_0 - x_k, y_0 - y_k) h_\phi(x_l - x_0, y_l - y_0)\, \mathrm{d}x_0\, \mathrm{d}y_0$$

Wilson & Naylor (2018a)

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

Tom J Wilson @onoddil

Match Separation Probability



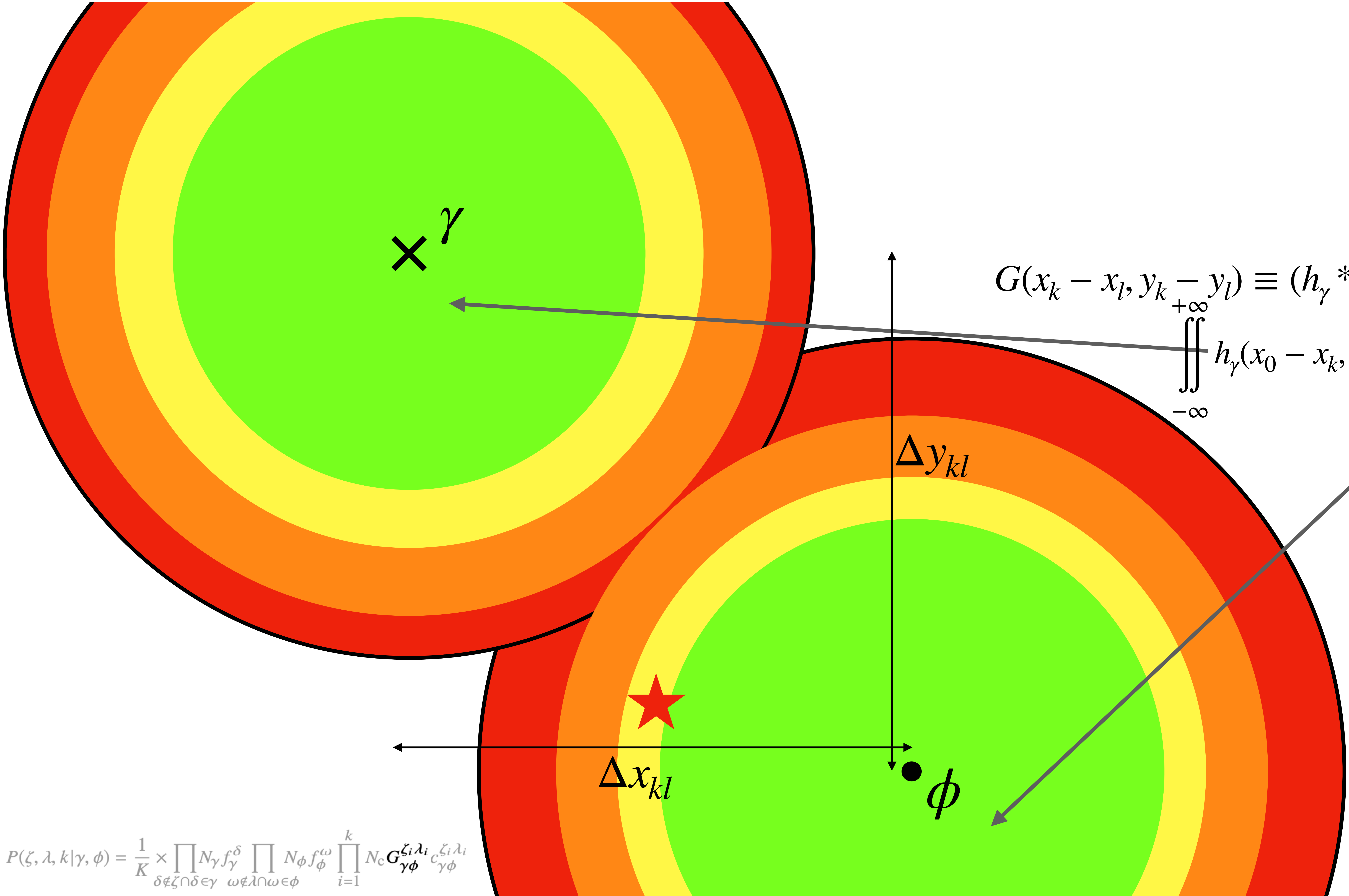
$$G(x_k - x_l, y_k - y_l) \equiv (h_\gamma * h_\phi)(\Delta x_{kl}, \Delta y_{kl}) = \iint_{-\infty}^{+\infty} h_\gamma(x_0 - x_k, y_0 - y_k) h_\phi(x_l - x_0, y_l - y_0)\, \mathrm{d}x_0\, \mathrm{d}y_0$$

Wilson & Naylor (2018a)

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

Tom J Wilson @onoddil

Match Separation Probability



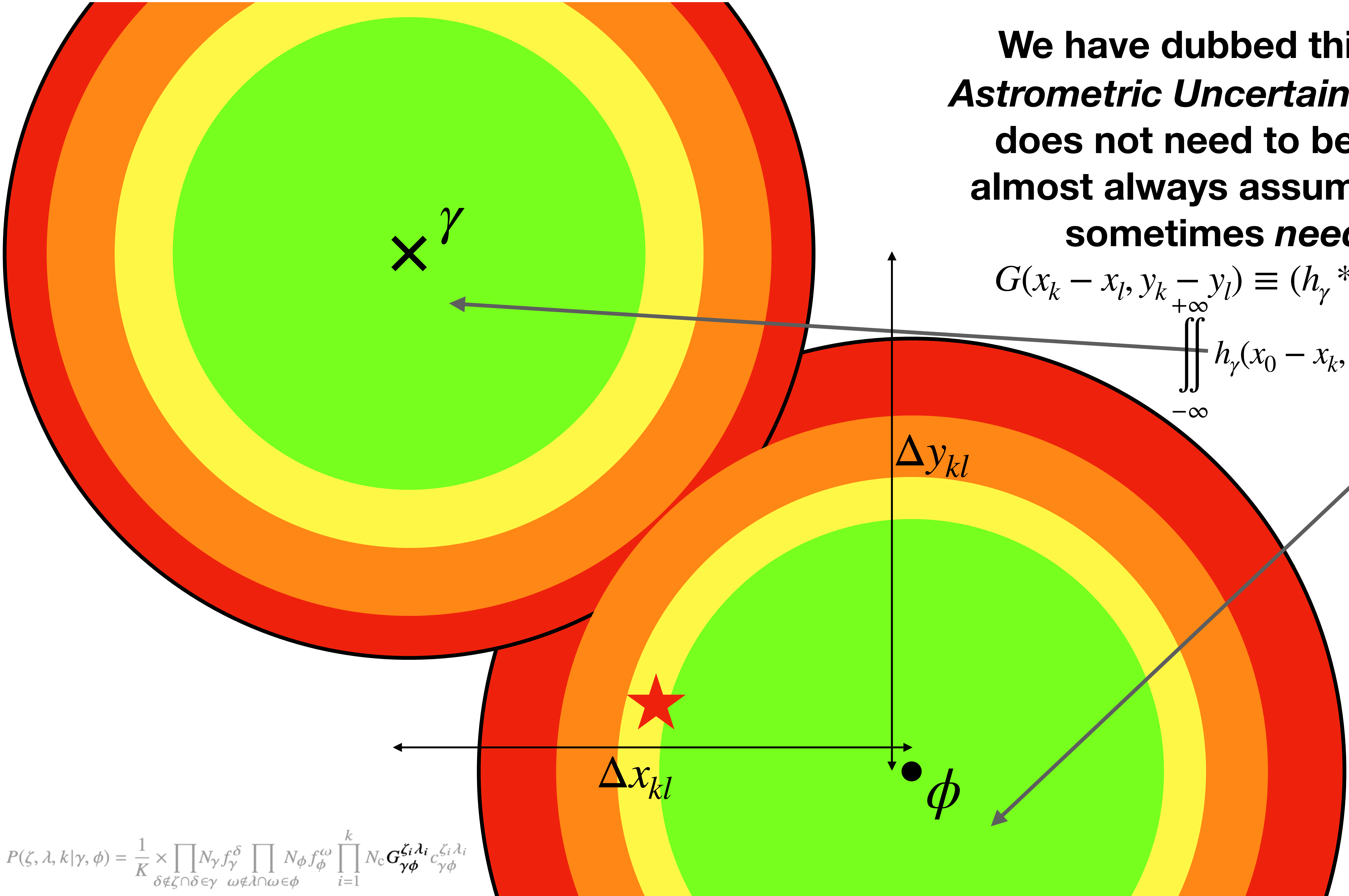
$$G(x_k - x_l, y_k - y_l) \equiv (h_\gamma * h_\phi)(\Delta x_{kl}, \Delta y_{kl}) = \iint_{-\infty}^{+\infty} h_\gamma(x_0 - x_k, y_0 - y_k) h_\phi(x_l - x_0, y_l - y_0)\, \mathrm{d}x_0\, \mathrm{d}y_0$$

Wilson & Naylor (2018a)

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

Tom J Wilson @onoddil

Match Separation Probability



We have dubbed this function h the *Astrometric Uncertainty Function,* which does not need to be Gaussian, as is almost always assumed — and indeed sometimes *needs* not to be!

$$G(x_k - x_l, y_k - y_l) \equiv (h_\gamma * h_\phi)(\Delta x_{kl}, \Delta y_{kl}) = \iint_{-\infty}^{+\infty} h_\gamma(x_0 - x_k, y_0 - y_k) h_\phi(x_l - x_0, y_l - y_0)\,\mathrm{d}x_0\,\mathrm{d}y_0$$

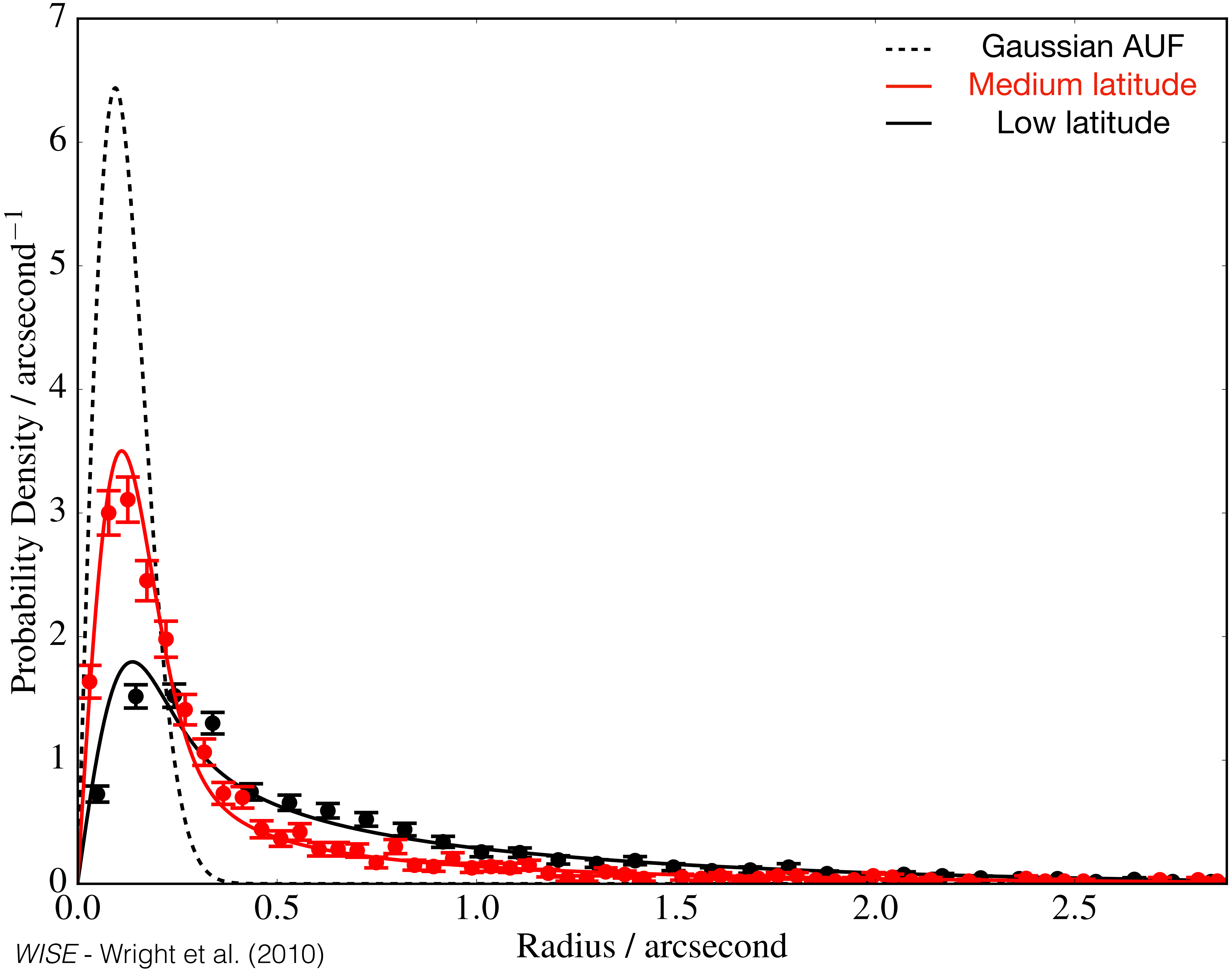
Wilson & Naylor (2018a)

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

Tom J Wilson @onoddil

Additional Components of the AUF

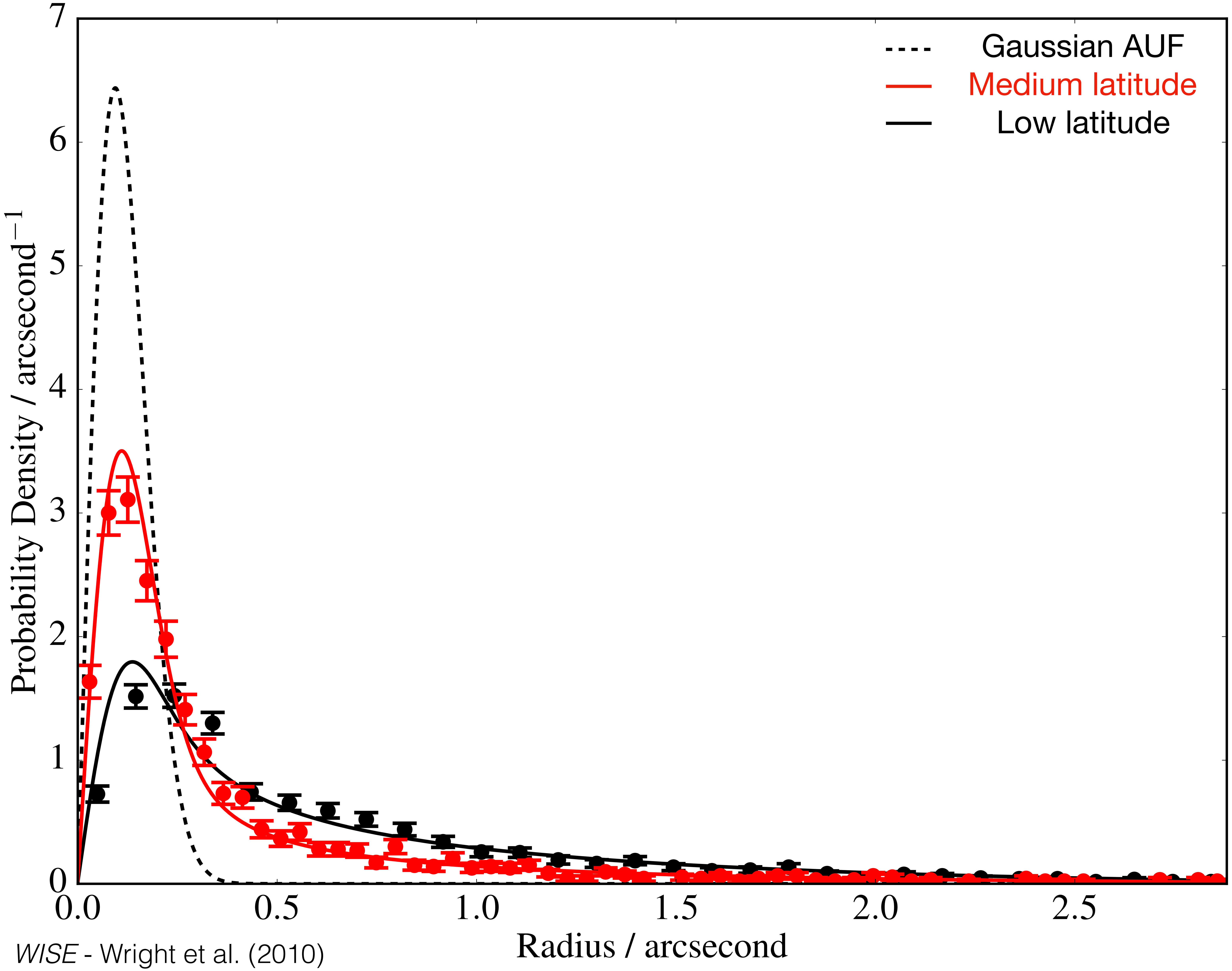
$$P(\zeta, \lambda, k|\gamma, \phi) = \frac{1}{K} \times \prod_{\delta \notin \zeta \cap \delta \in \gamma} N_\gamma f_\gamma^\delta \prod_{\omega \notin \lambda \cap \omega \in \phi} N_\phi f_\phi^\omega \prod_{i=1}^{k} N_c \boldsymbol{G}_{\gamma\phi}^{\zeta_i \lambda_i} c_{\gamma\phi}^{\zeta_i \lambda_i}$$

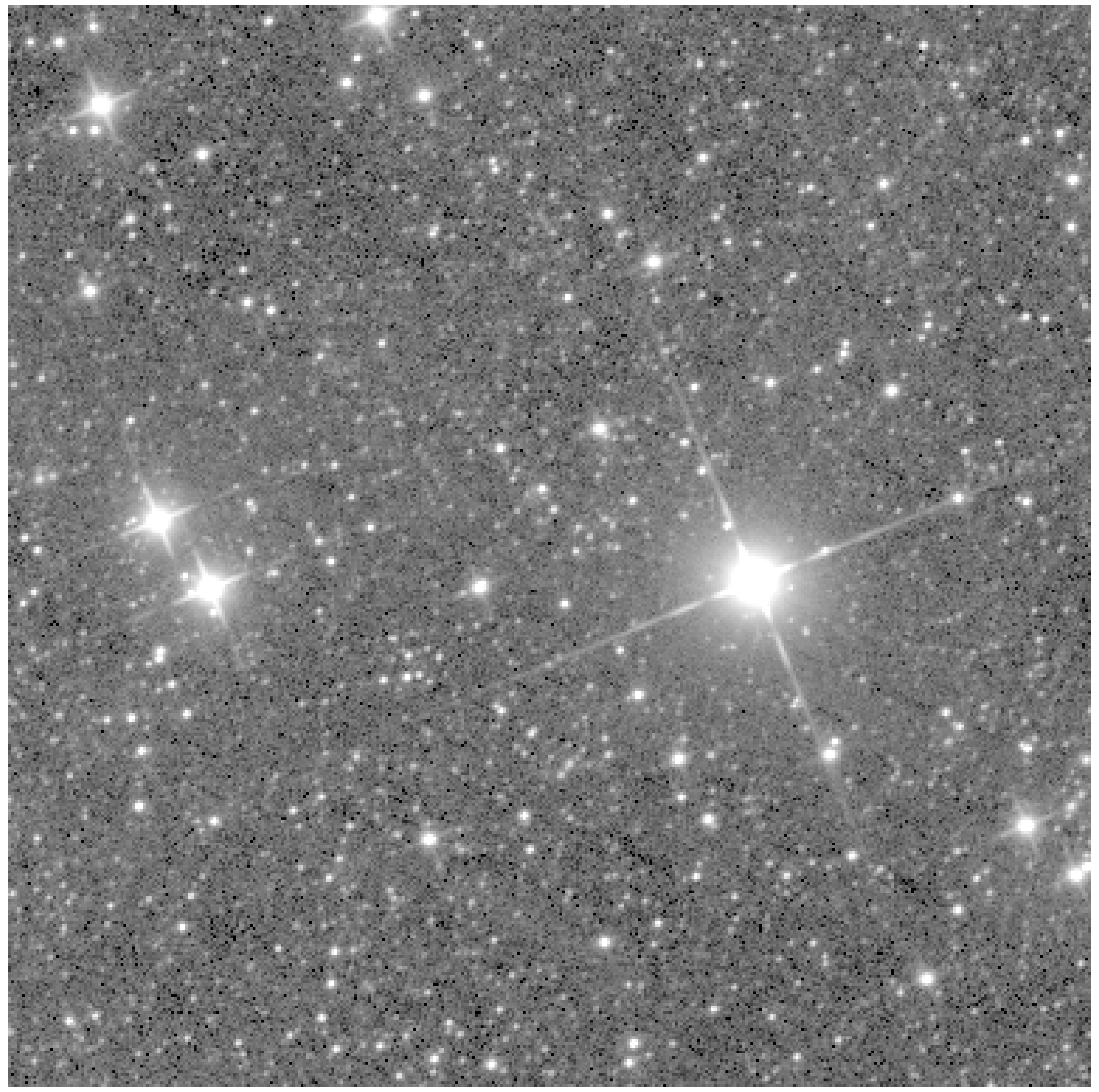


WISE - Wright et al. (2010)
Gaia DR2 - Gaia Collaboration, Brown A. G. A., et al. (2018)

Additional Components of the AUF

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



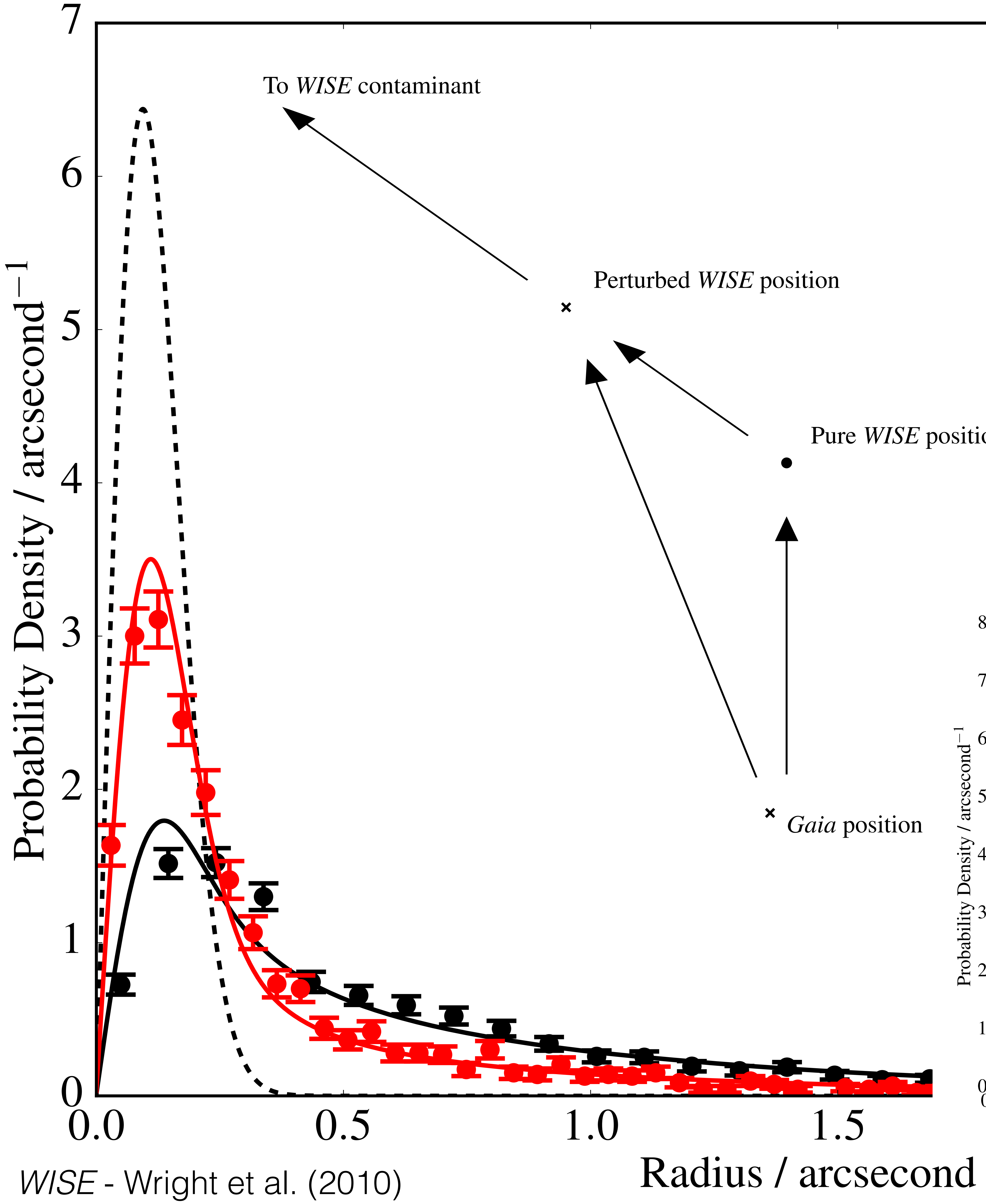


WISE - Wright et al. (2010)
Gaia DR2 - Gaia Collaboration, Brown A. G. A., et al. (2018)

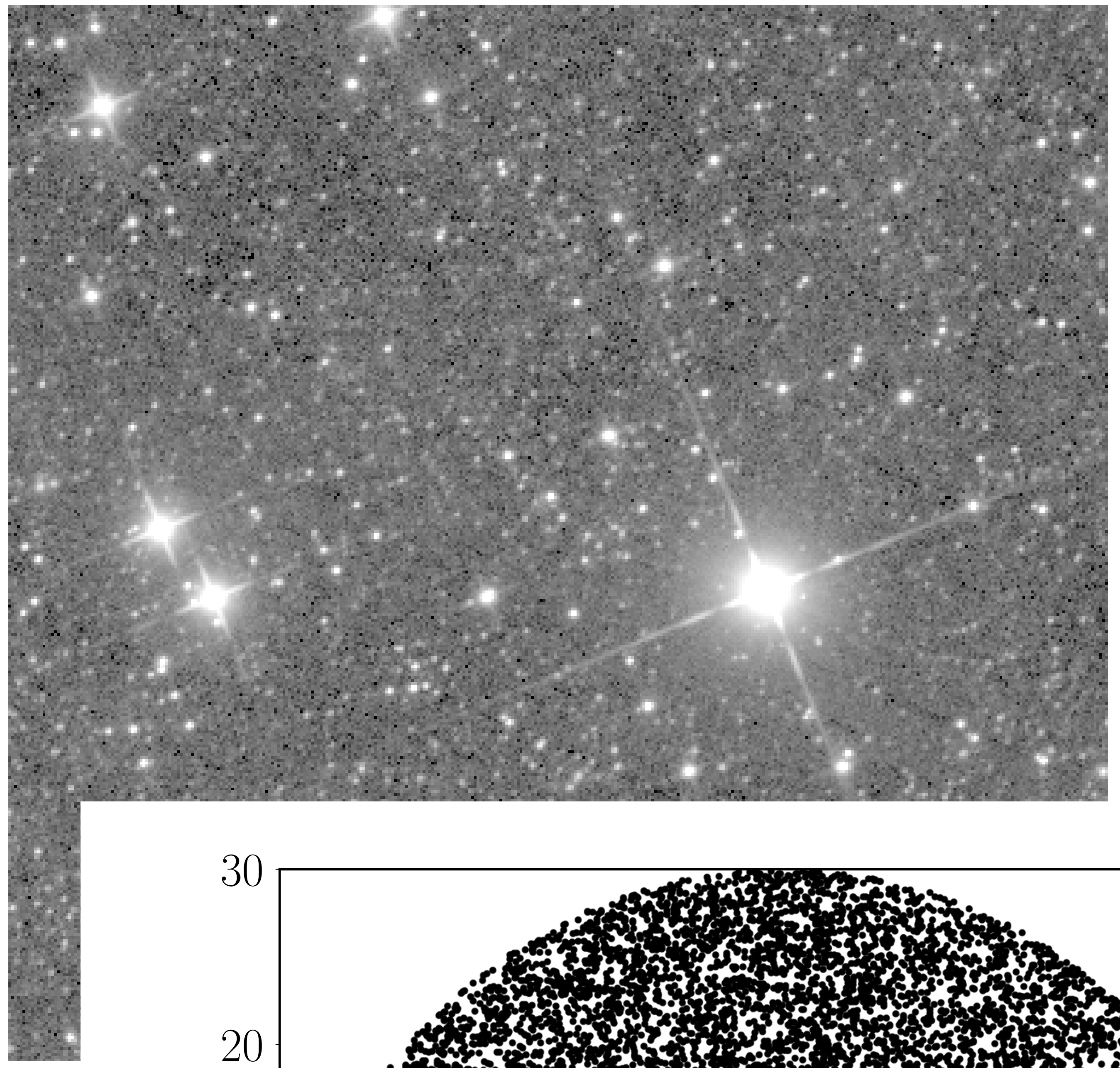
Tom J Wilson @onoddil

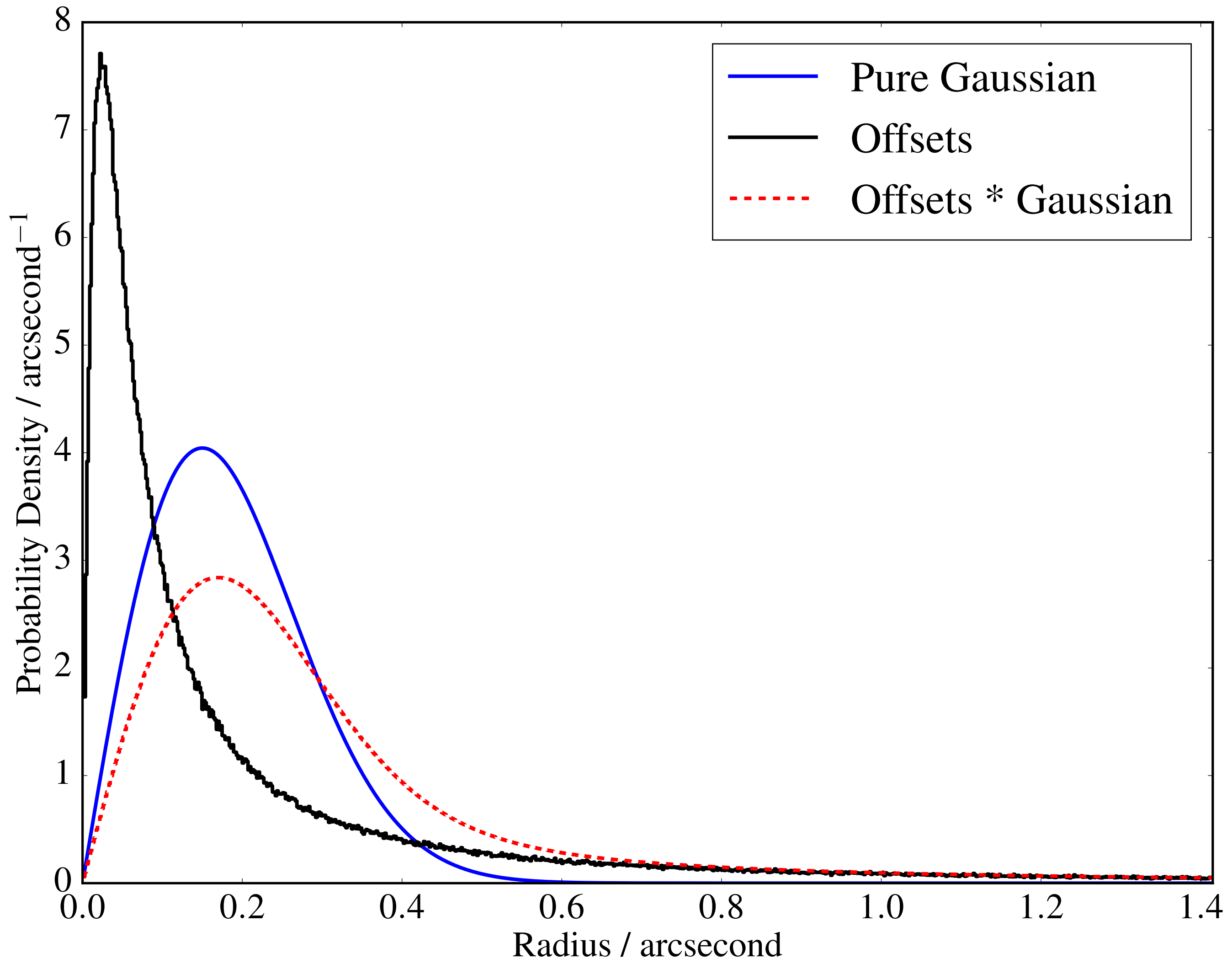
Additional Components of the AUF

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

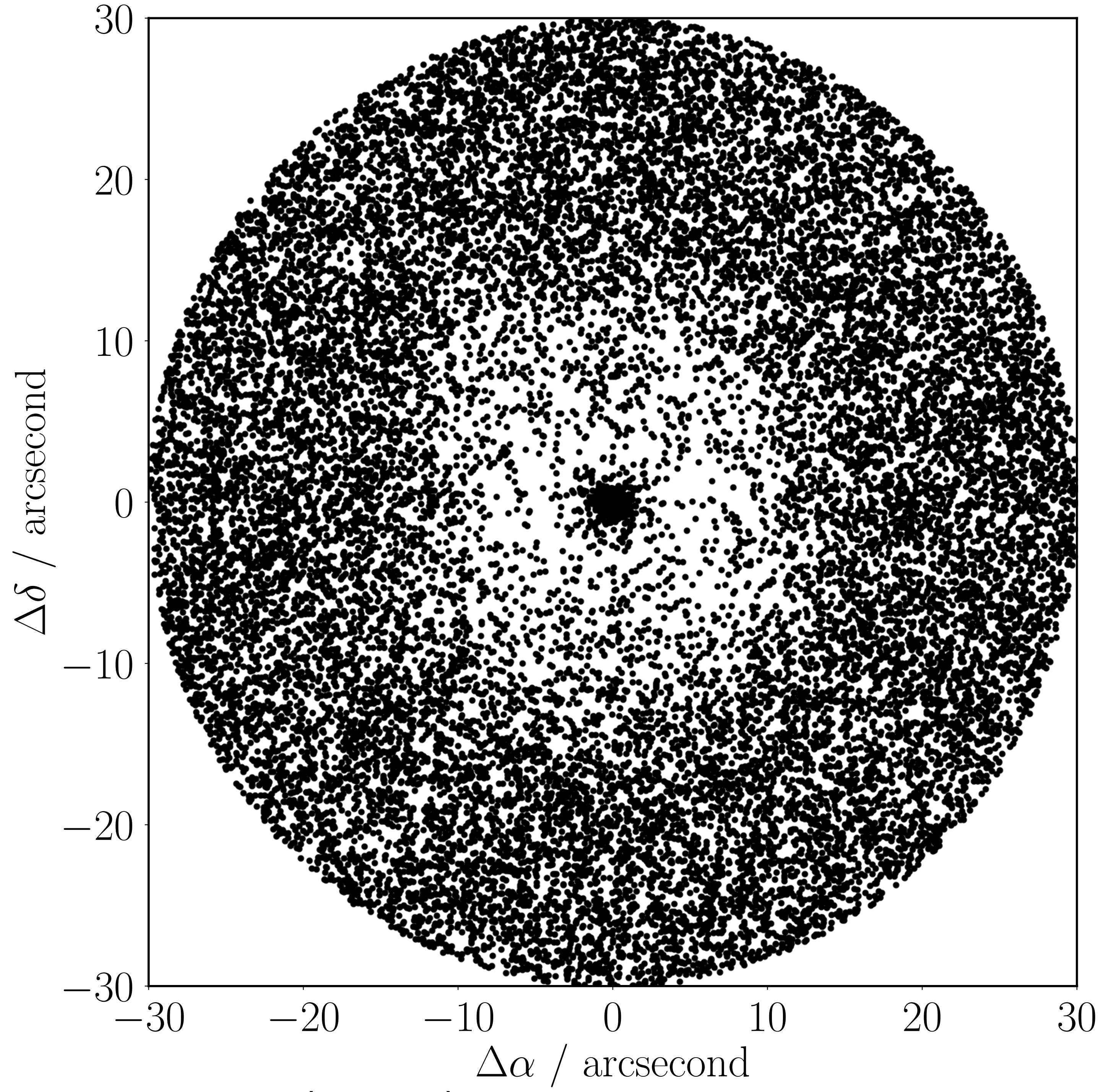


WISE - Wright et al. (2010)
Gaia DR2 - Gaia Collaboration, Brown A. G. A., et al. (2018)





Wilson & Naylor (2018b)

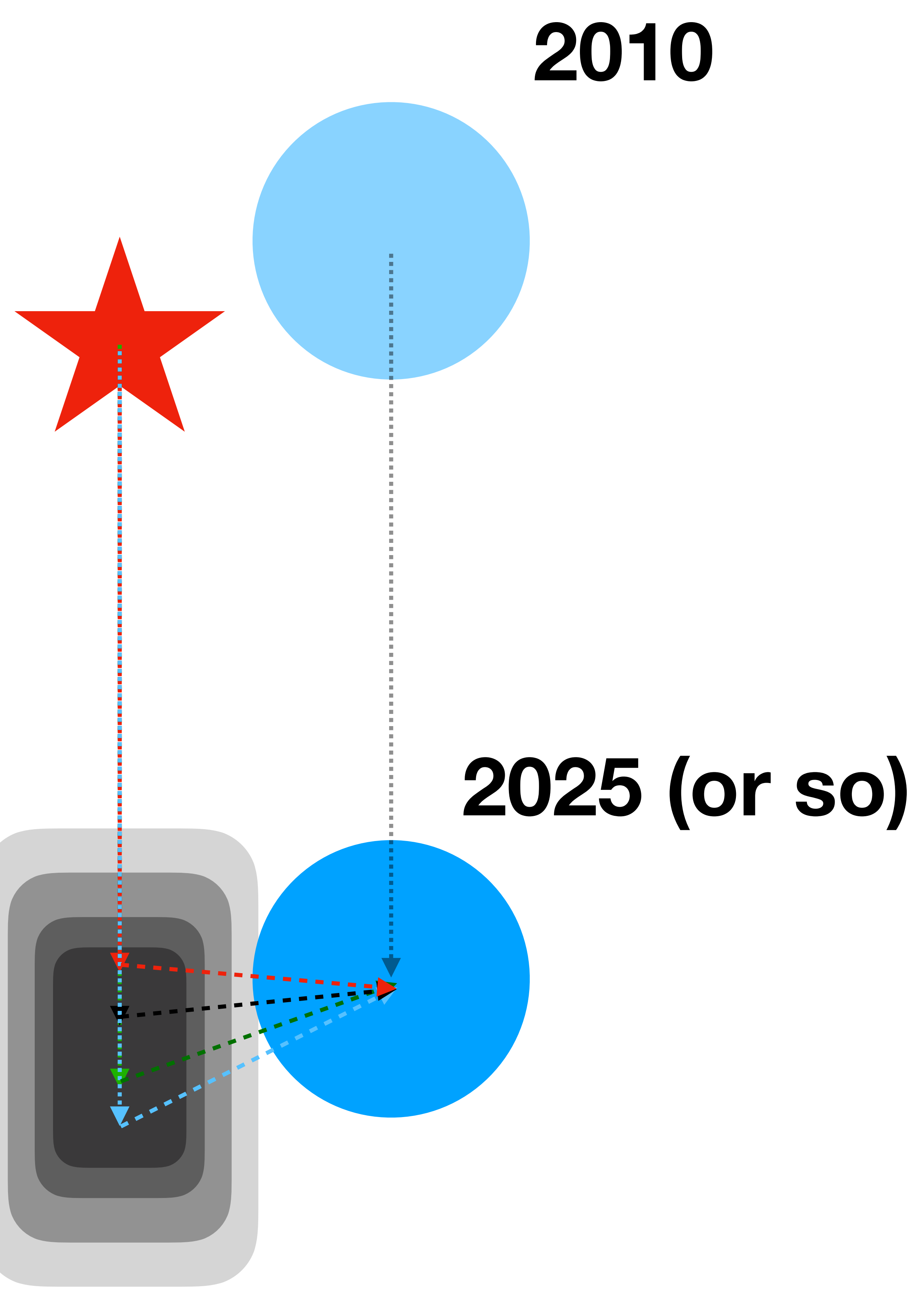


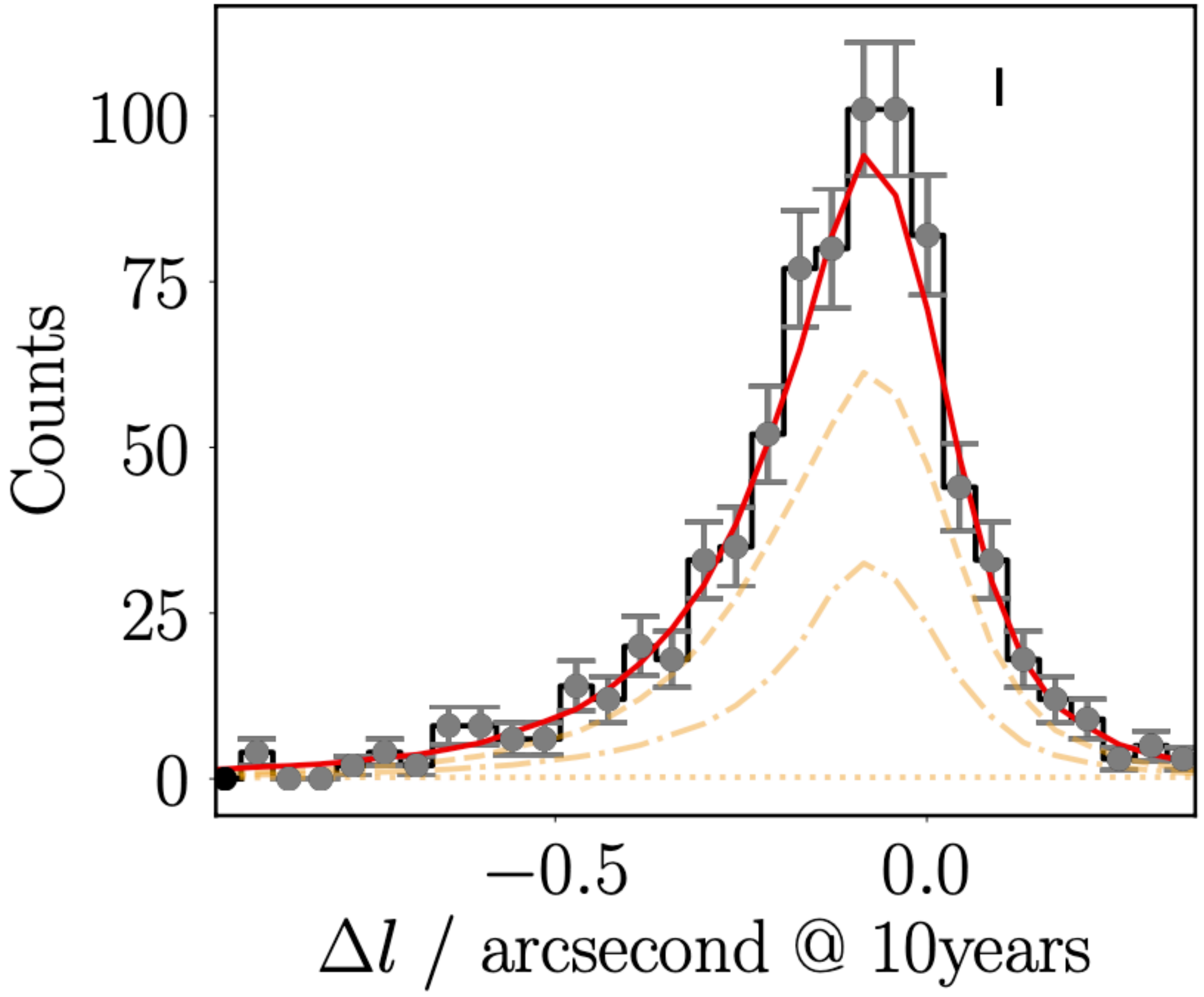
Wilson & Naylor (2017)

Tom J Wilson @onoddil

Including Unknown Proper Motions

Using a model for the distribution of potential proper motions, and hence astrometric drifts, of a source of a given sky position and brightness we can include "fast forwarding" of sources through time across different catalogues when individual proper motions are not known





Because this function works in *separation*, rather than pure *position*, space, we apply the distribution after calculating G.

$$G' = G * h'_{\mathrm{pm}} \qquad G = h_\gamma * h_\phi$$

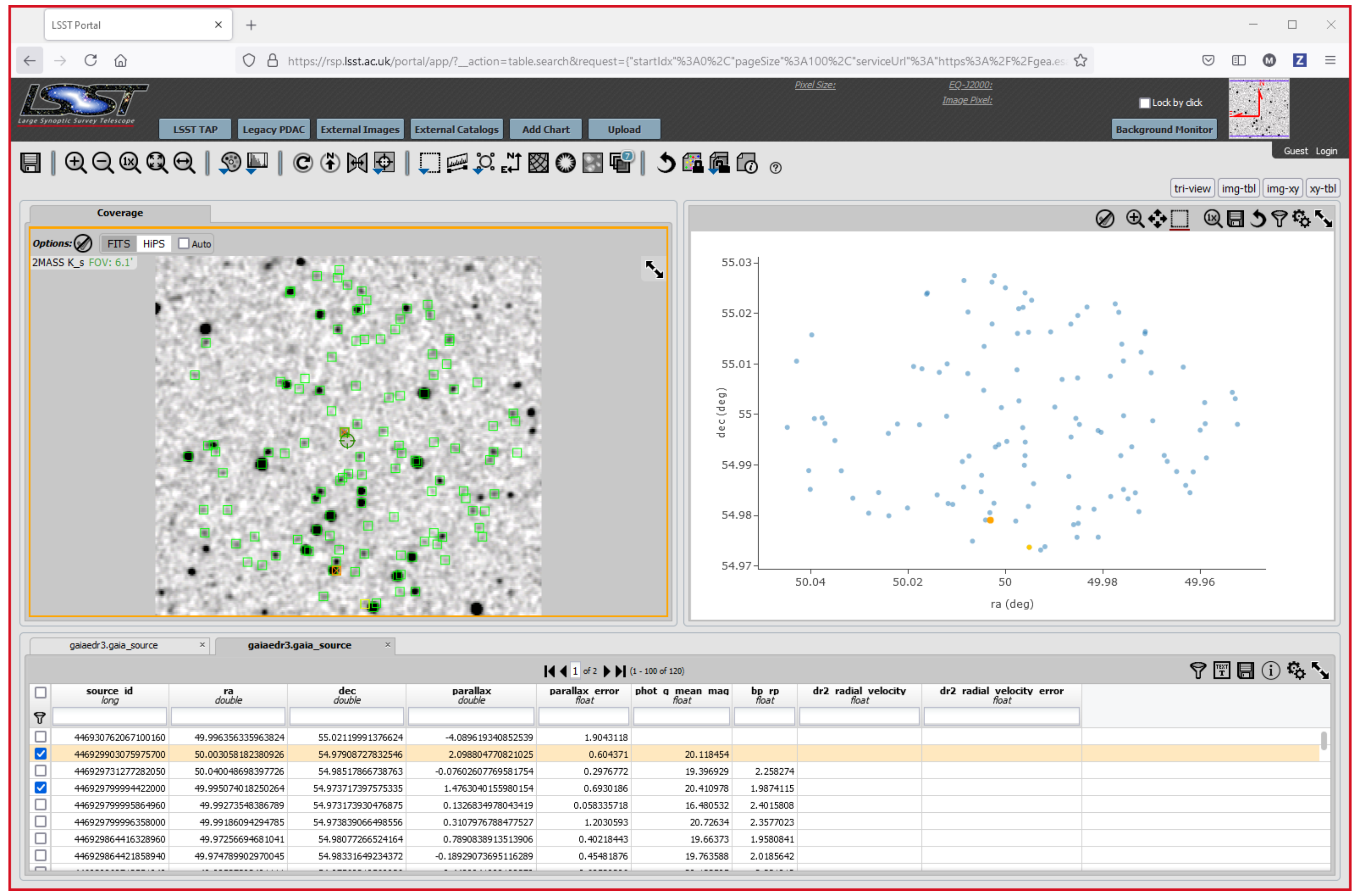
$$h_\gamma = h_{\gamma,\mathrm{centroiding}} * h_{\gamma,\mathrm{perturbation}}$$

Wilson & Naylor (in prep.)
Gaia eDR3 - Gaia Collaboration, Brown A. G. A., et al. (2021)

Tom J Wilson @onoddil

How To Use Our Cross-Matches

(Or, how this impacts you on a day-to-day basis)



Three tables per cross-match: merged catalogue dataset, and 2x non-match dataset (one per catalogue)

Counterpart columns:

- Designations of the two sources (e.g., WISE J… and *Gaia* EDR3…)
- RA and Dec (or Galactic l/b) of the two sources
- Magnitudes (corrected for necessary effects, such as e.g. *Gaia*) in all bandpasses for both objects
- Match probability — probability of the most likely permutation (see equation 26 of Wilson & Naylor 2018a)
- Eta - Photometric likelihood ratio (counterpart vs non-match probability, just for brightnesses; see eq37 of WN18a)
- Xi - Astrometric likelihood ratio (just position match/non-match comparison; see eq38 of WN18a)
- Average contamination - simulated mean (percentile) brightening of the two sources, based on number density of catalogue
- Probability of sources having blended contaminant above e.g. 1% relative flux

Non-match columns:

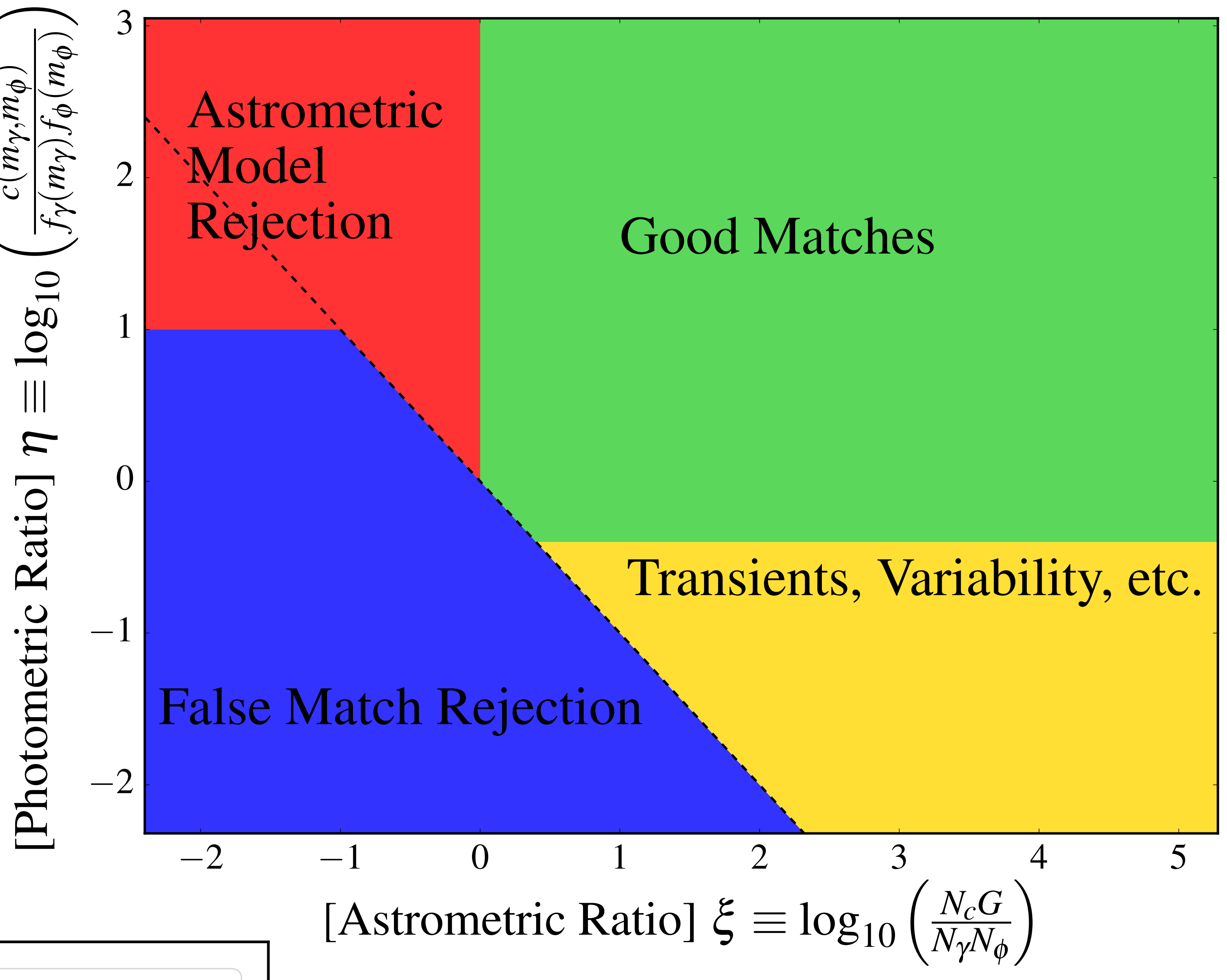
- Designations of the source (e.g., WISE J… or *Gaia* EDR3…)
- RA and Dec (or Galactic l/b) of the source
- Magnitudes (corrected for necessary effects, such as e.g. *Gaia*) in all bandpasses
- Match probability — probability of the most likely permutation (see equation 26 of WN18a)
- Average contamination - simulated mean (percentile) brightening of the source, based on number density of catalogue

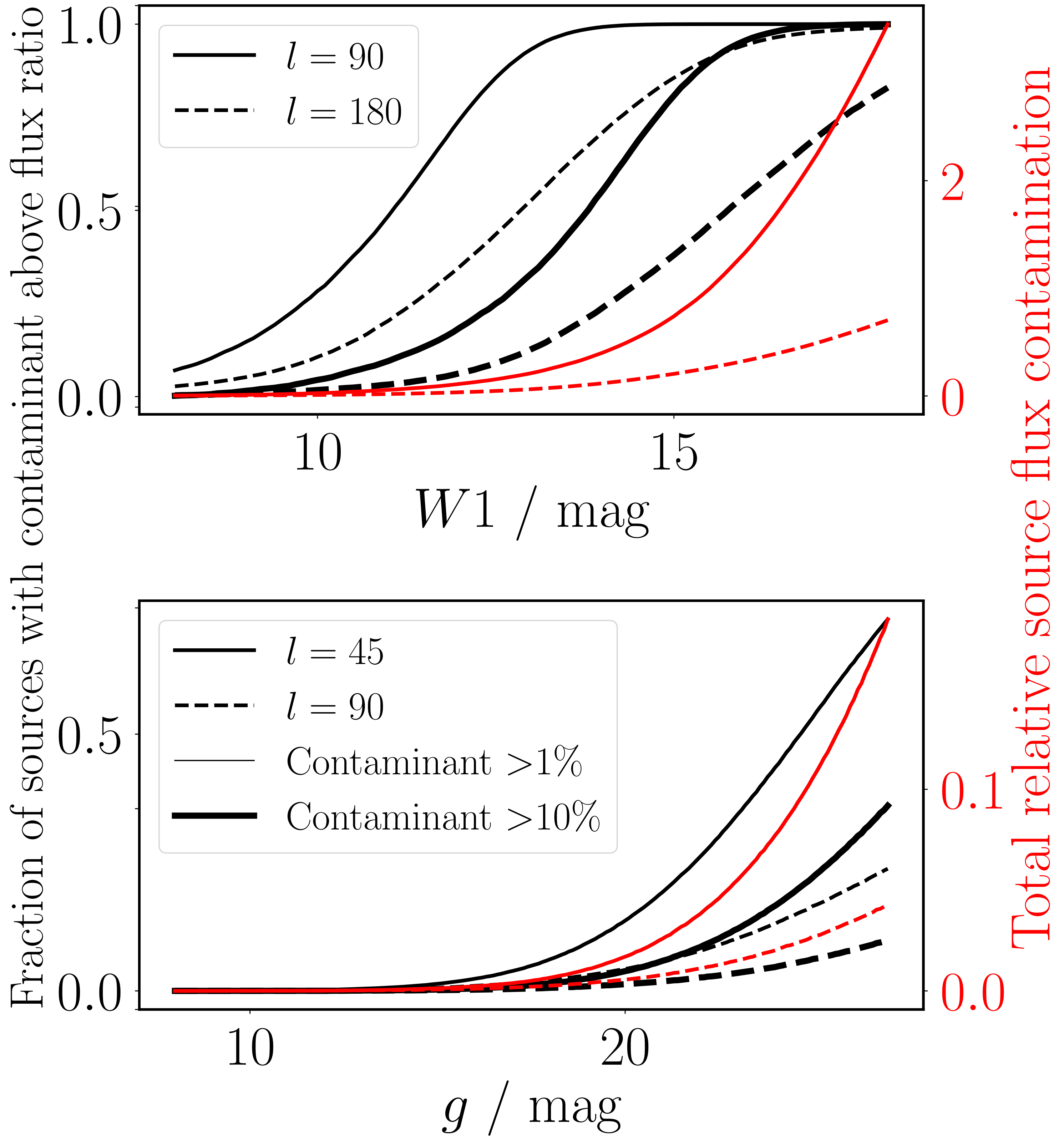
We will provide a two match runs per catalogue pair match: one with, and one without, the photometry considered, to allow for the recovery of sources with "weird" colours but otherwise agreeable astrometry

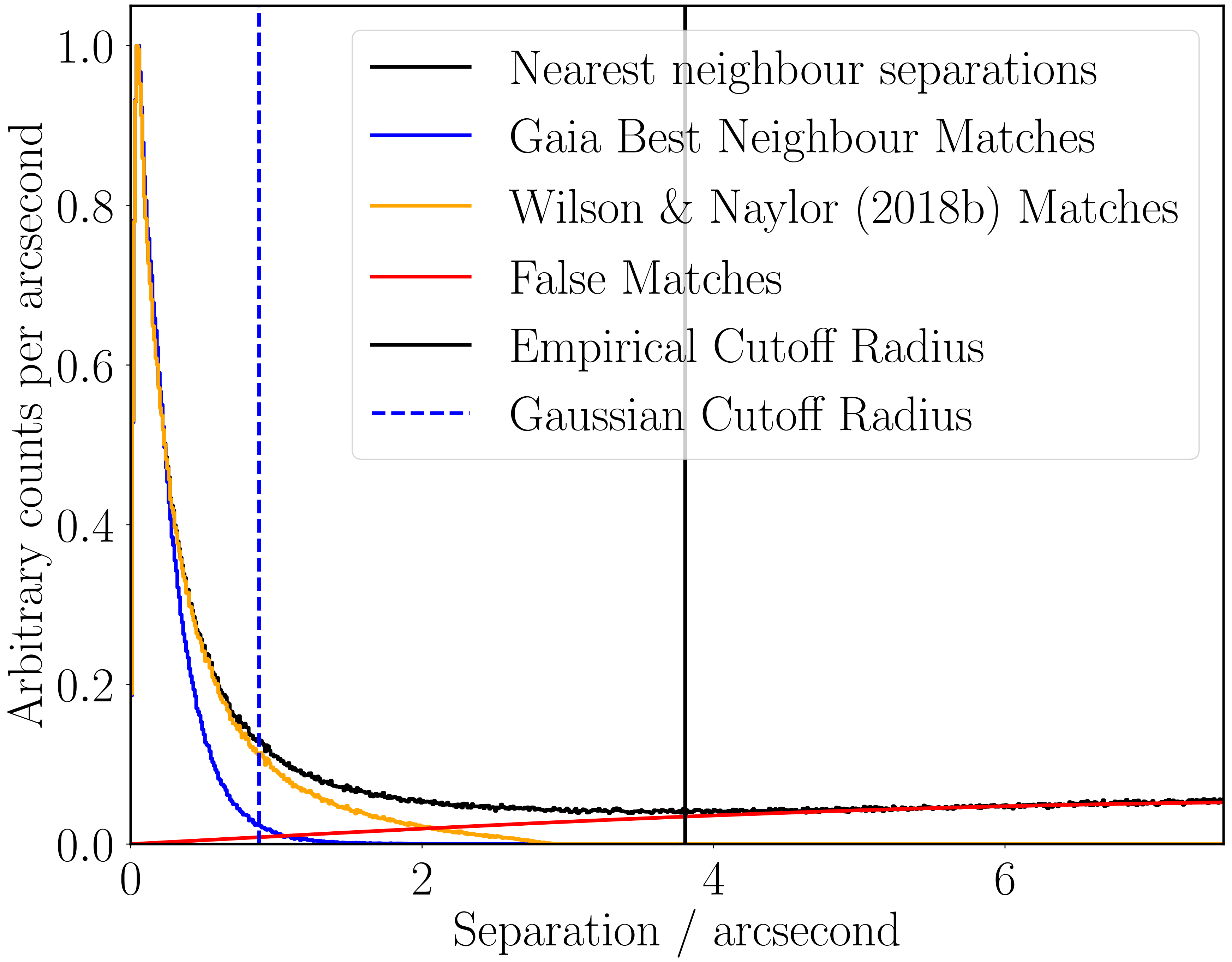
Tom J Wilson @onoddil

Why Use Our Cross-Matches?

- **Getting cross-matches, even for "well behaved" fields**
- **Finding "odd" objects, either using the inclusion vs non-inclusion of the photometry in the two match runs, or via the likelihood ratio space — planned "real time" matching service for transient objects**
- **Removing e.g. IR excess or correcting for extinction-like crowding brightening, through Average Contamination; crucial for removing completely unknown crowding of catalogues using aperture photometry**
- **Recovering additional sources missed by other match services — either in crowded fields (we recover 20-50% more *Gaia*-*WISE* matches than the *Gaia* best neighbour matches), or with our in-progress extension to unknown proper motion modelling**







Tom J Wilson @onoddil

Conclusions

- Upcoming LSST:UK cross-match service macauff — let me know your thoughts/needs/hopes/dreams
  - Provide tables of cross-matches between LSST and <your favourite catalogue here!>

- Our cross-matches include two key elements for avoiding issues with the crowded LSST sky
  - A generalised approach to the Astrometric Uncertainty Function allows for the inclusion of the effects of perturbation due to blended sources — and will be extended to unknown proper motions
  - Use of the photometry of sources allows for the rejection of false matches (with >1 "extra" source per 2 arc second circle in most of the LSST Galactic plane)
- Will include additional information on the crowding of sources, allowing for selection of uncontaminated objects, or modelling of excess flux — crucial for removal of red excess in SEDs
  - LSST will suffer of order 10% flux contamination, which could be confused with extinction

VERA C. RUBIN OBSERVATORY

UNIVERSITY OF EXETER

LSST

LSST:UK Consortium

@Onoddil @pm.me .github.io www

Wilson & Naylor, 2017, MNRAS, 468, 2517
Wilson & Naylor, 2018a, MNRAS, 473, 5570
Wilson & Naylor, 2018b, MNRAS, 481, 2148
https://github.com/Onoddil/macauff

Macauff