Robust Cross-Matches with *Herschel* (and beyond): Overcoming the Effect of Unresolved Contaminant Objects and False Positive Matches

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

National Astronomy Meeting 2023 — 10 Years After Herschel, 6/Jul/23

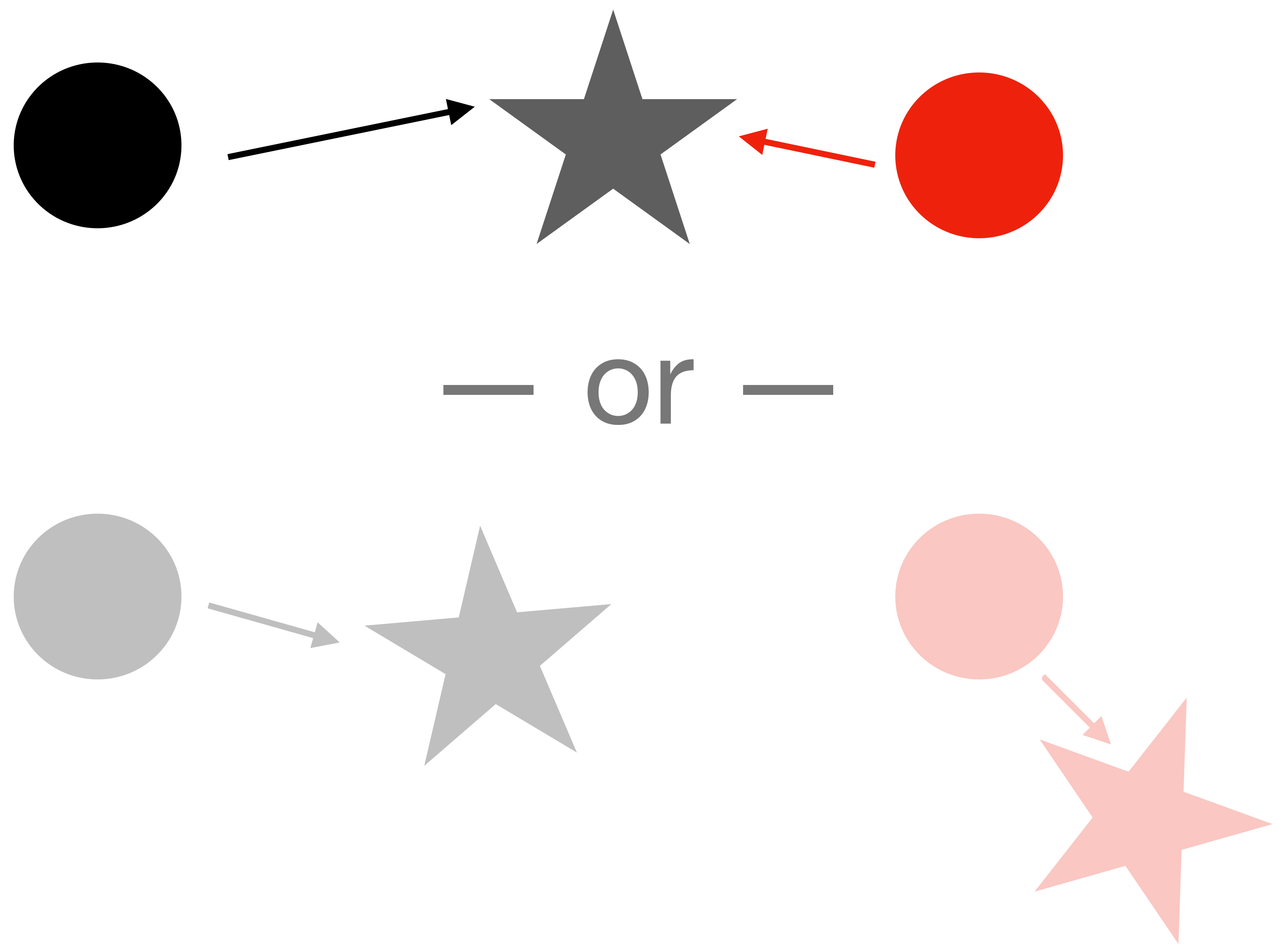
@Onoddil @pm.me
.github.io

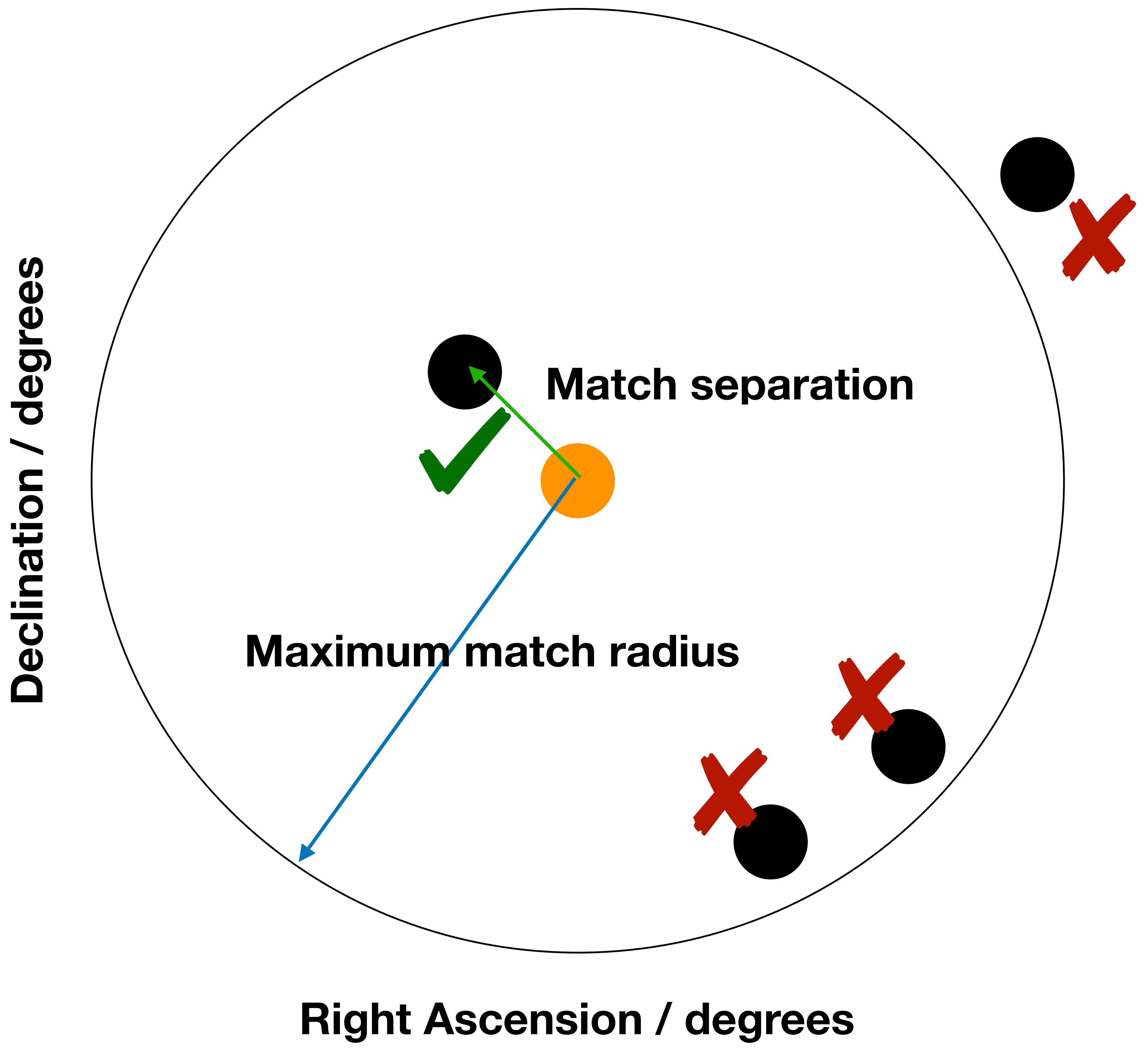
University of Exeter

LSST:UK Consortium

Science and Technology Facilities Council

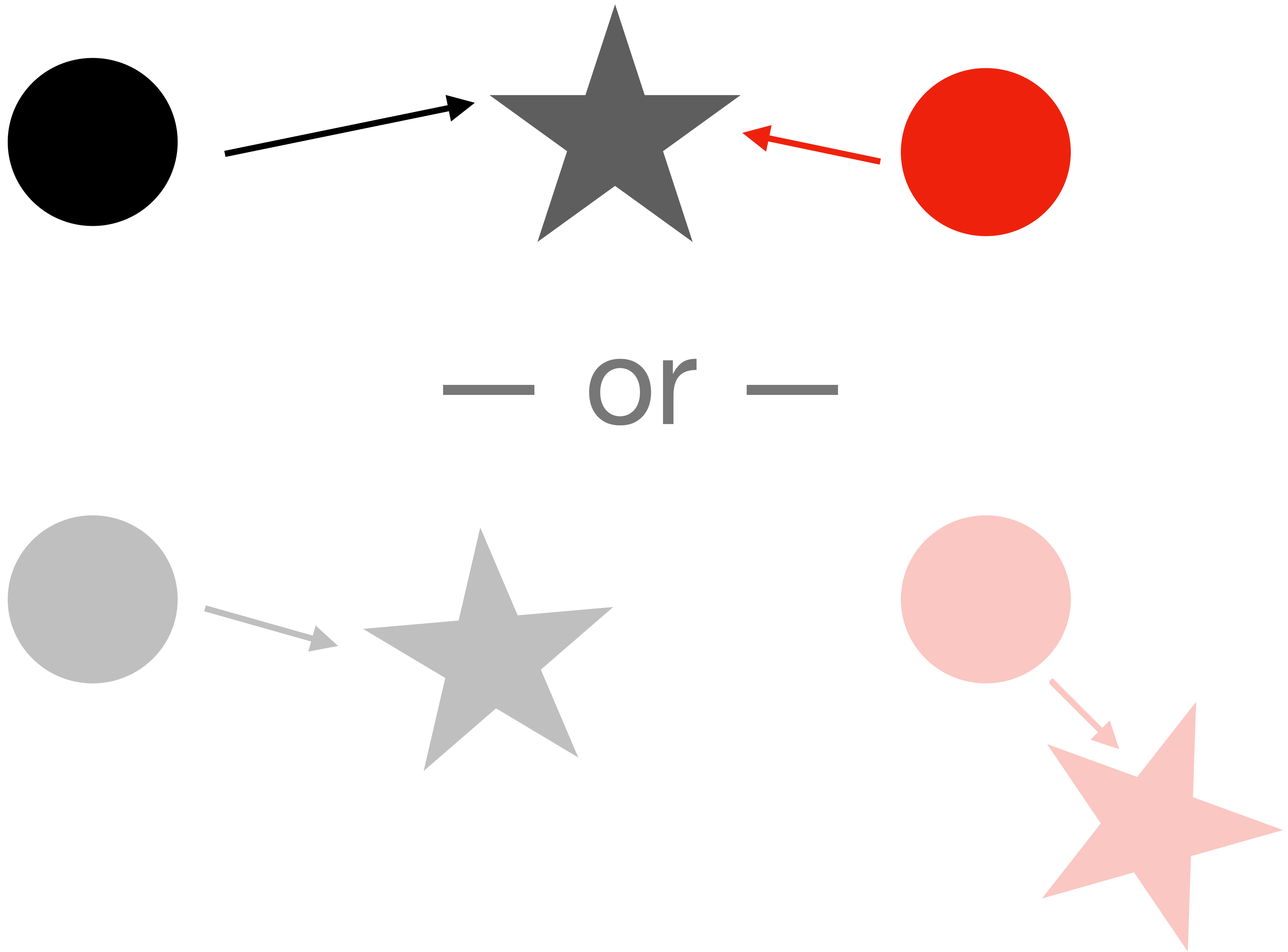
“Simple” Cross-Matching

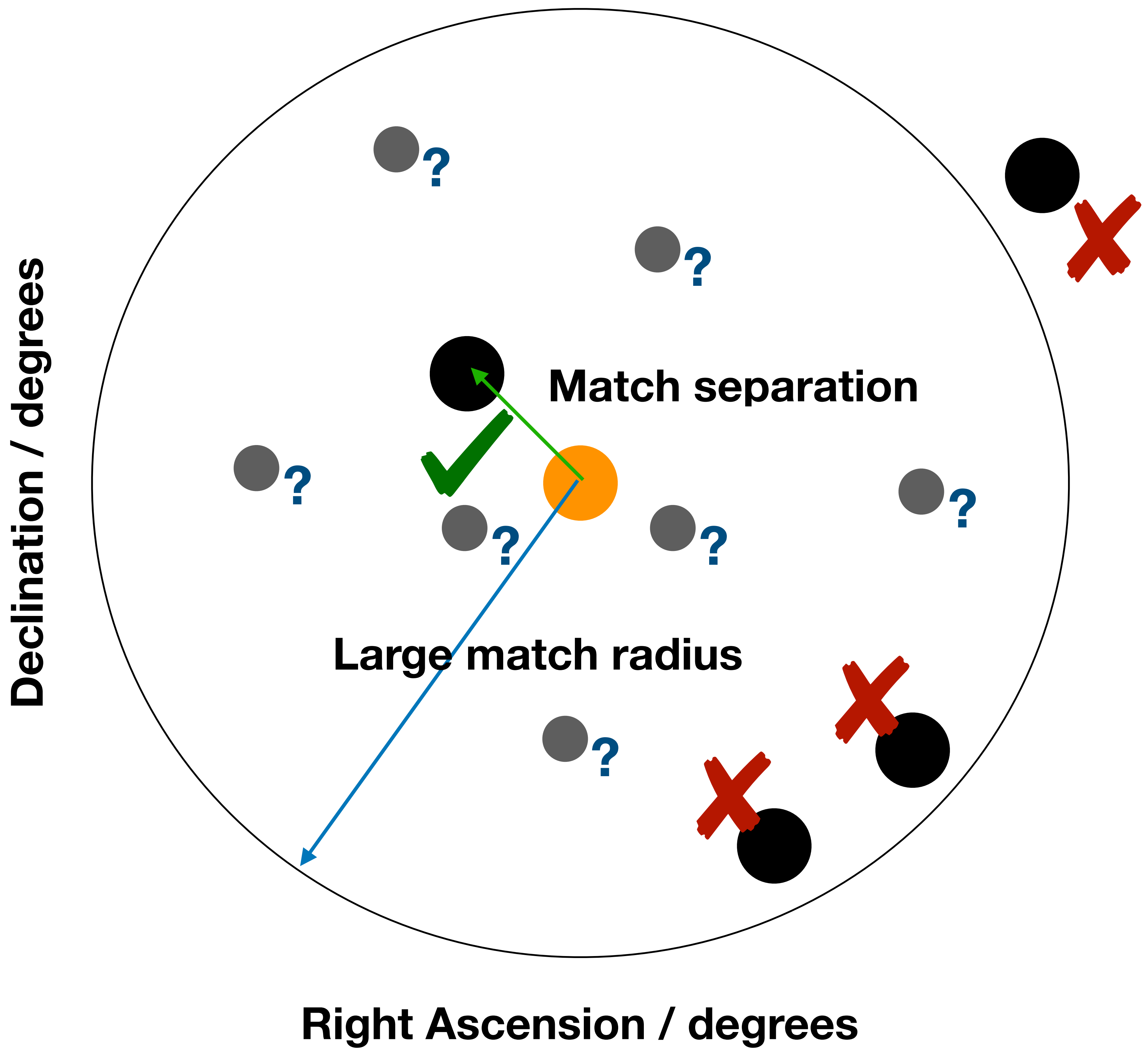


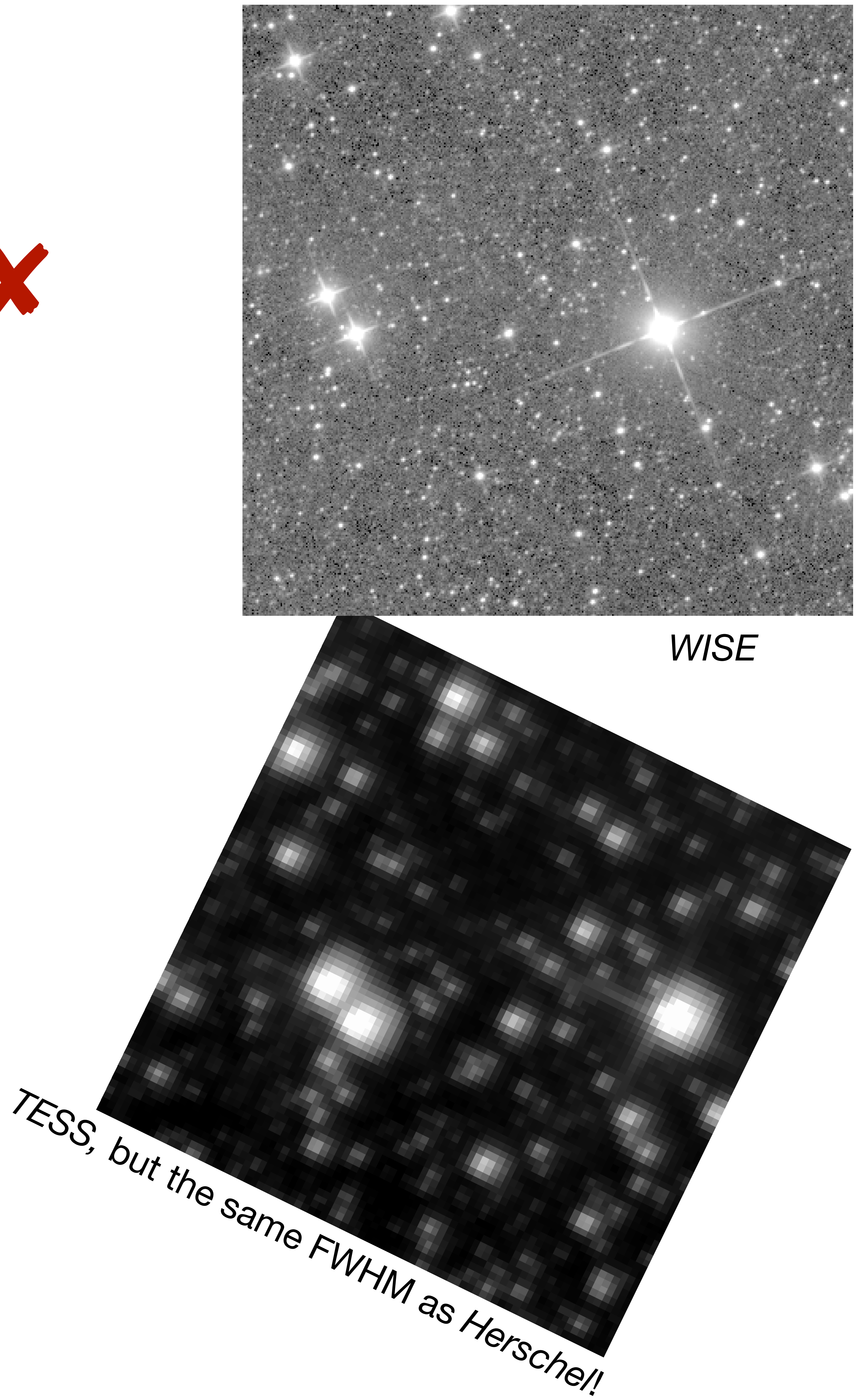


Tom J Wilson @onoddil

“Simple” Cross-Matching with *Herschel*







Tom J Wilson @onoddil

Probabilistic Cross-Matching

The Likelihood Ratio

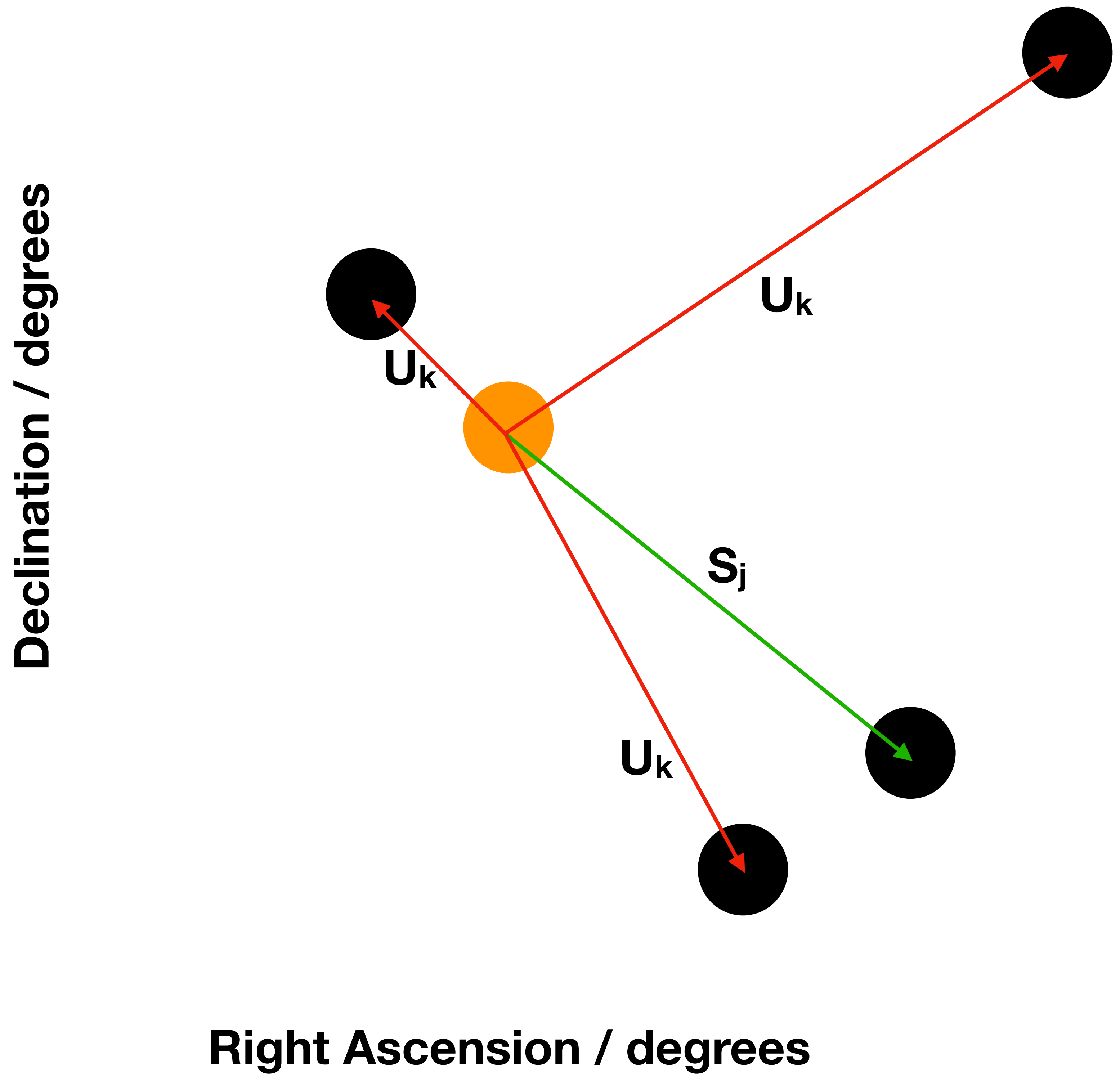
$$dp(r|id) = r \times e^{-r^2/2}\, dr.$$
$$dp(r|c) = 2\lambda r \times e^{-\lambda r^2}\, dr$$
$$LR(r) = dp(r|id)/dp(r|c) = \frac{1}{2\lambda}\exp\left\{\frac{r^2}{2}(2\lambda - 1)\right\}$$

de Ruiter, Willis, & Arp (1977)

$$dp_{\mathrm{id}} = Qr\exp\left(\frac{-r^2}{2}\right)dr. \quad dp_{\mathrm{uo}} = 2\lambda r\, dr$$
$$LR(r) = \frac{dp_{\mathrm{id}}}{dp_{\mathrm{uo}}} = \frac{Q\exp(-r^2/2)}{2\lambda}$$

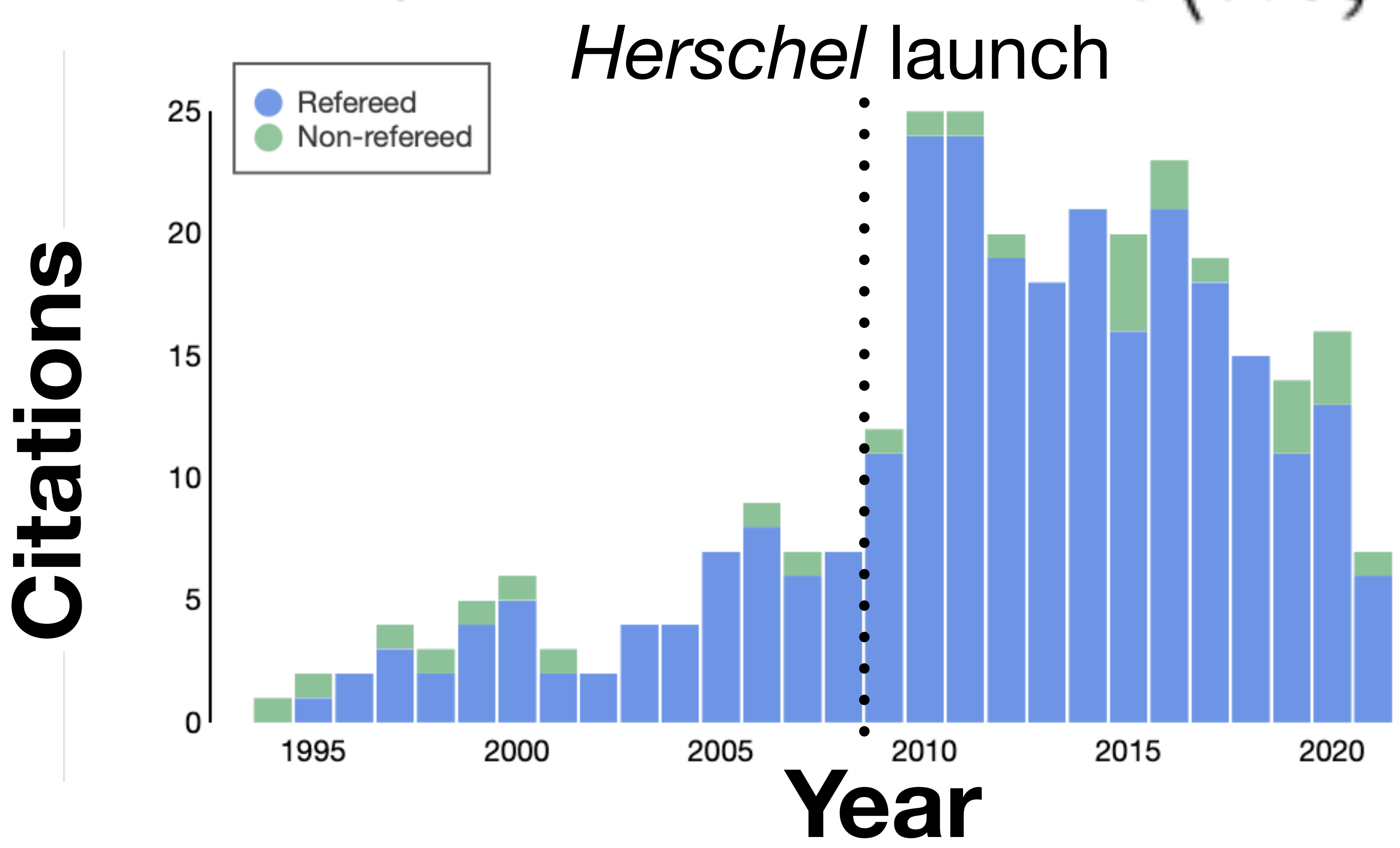
Wolstencroft et al. (1986)

The "Reliability" — Sutherland & Saunders (1992)



$$R_j = \frac{L_j}{\sum_i L_i + (1 - Q)}$$

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

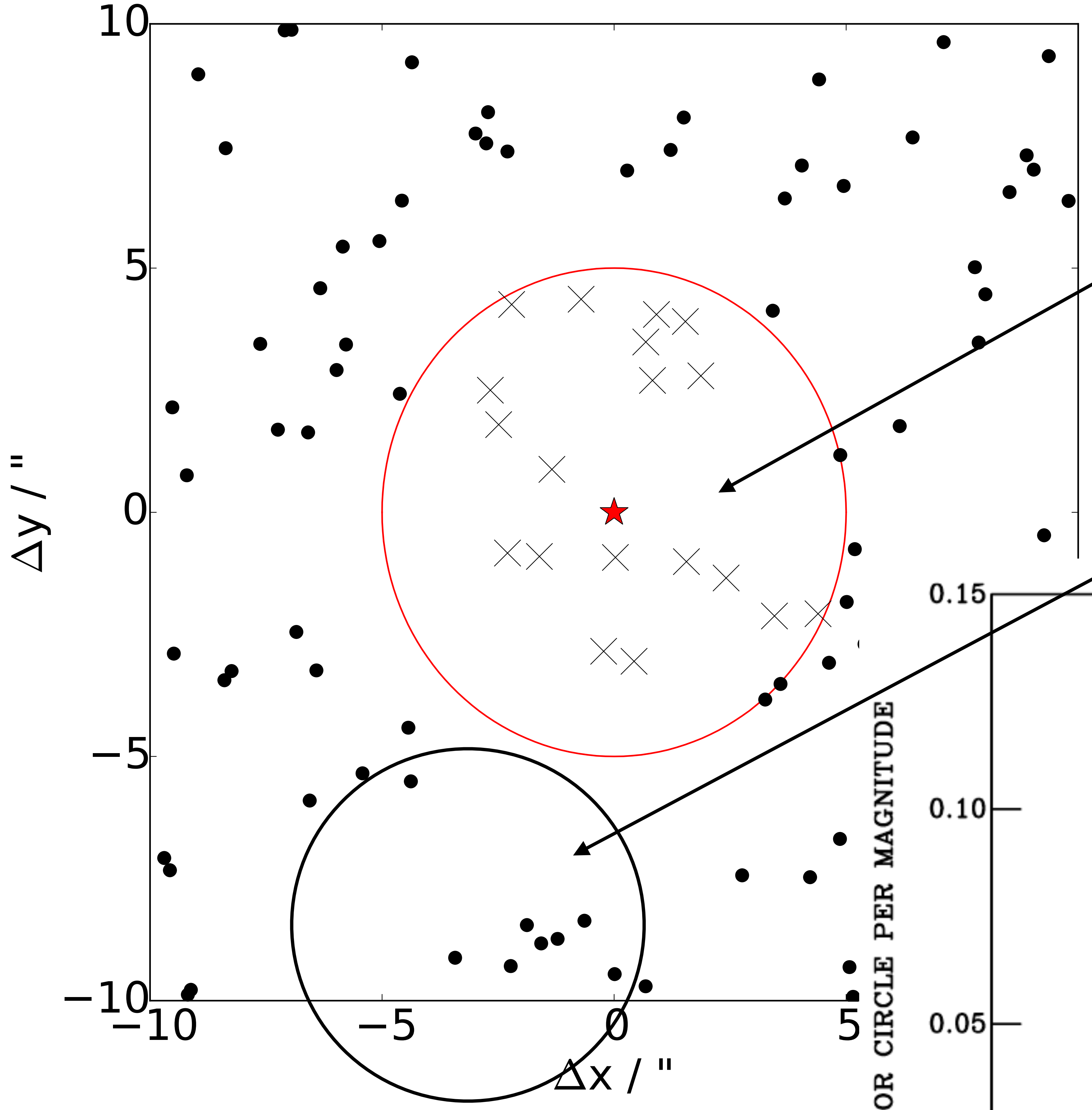


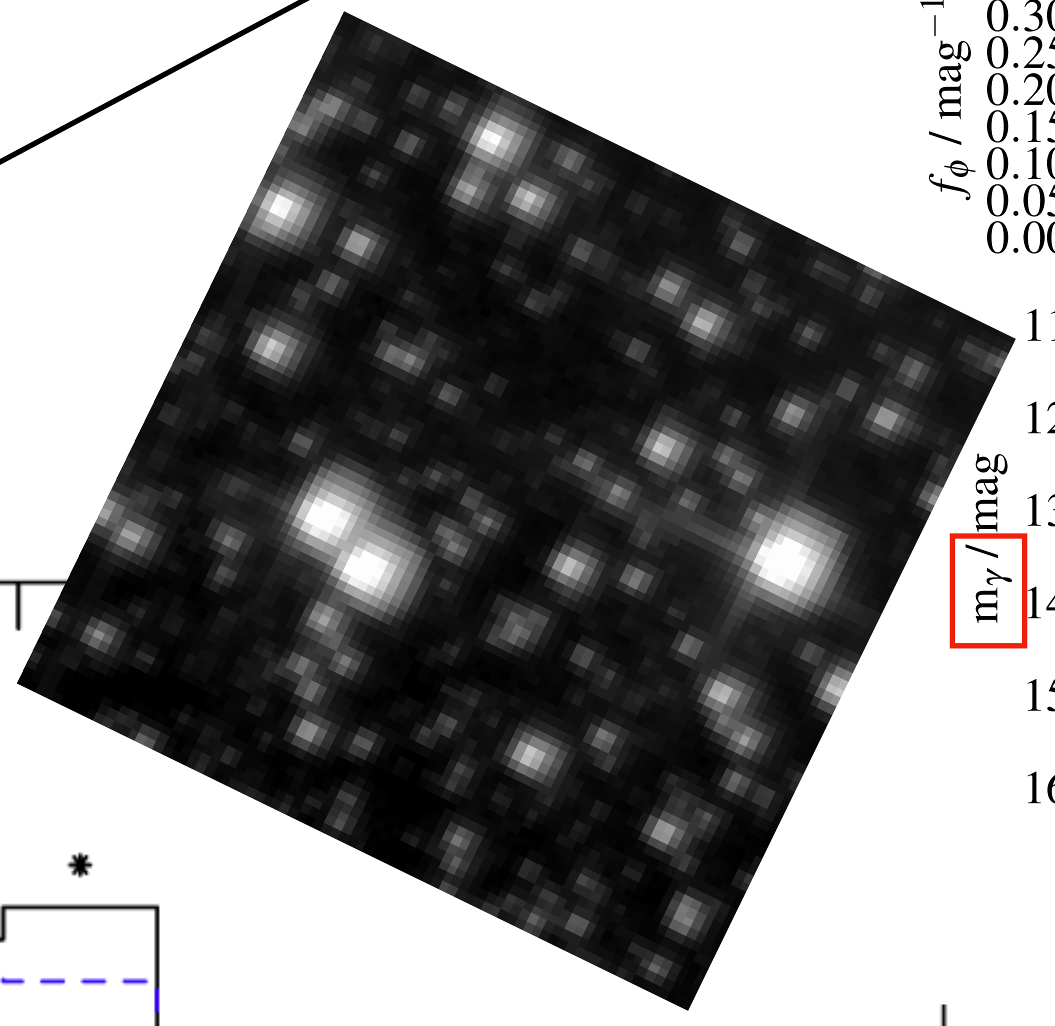
Tom J Wilson @onoddil

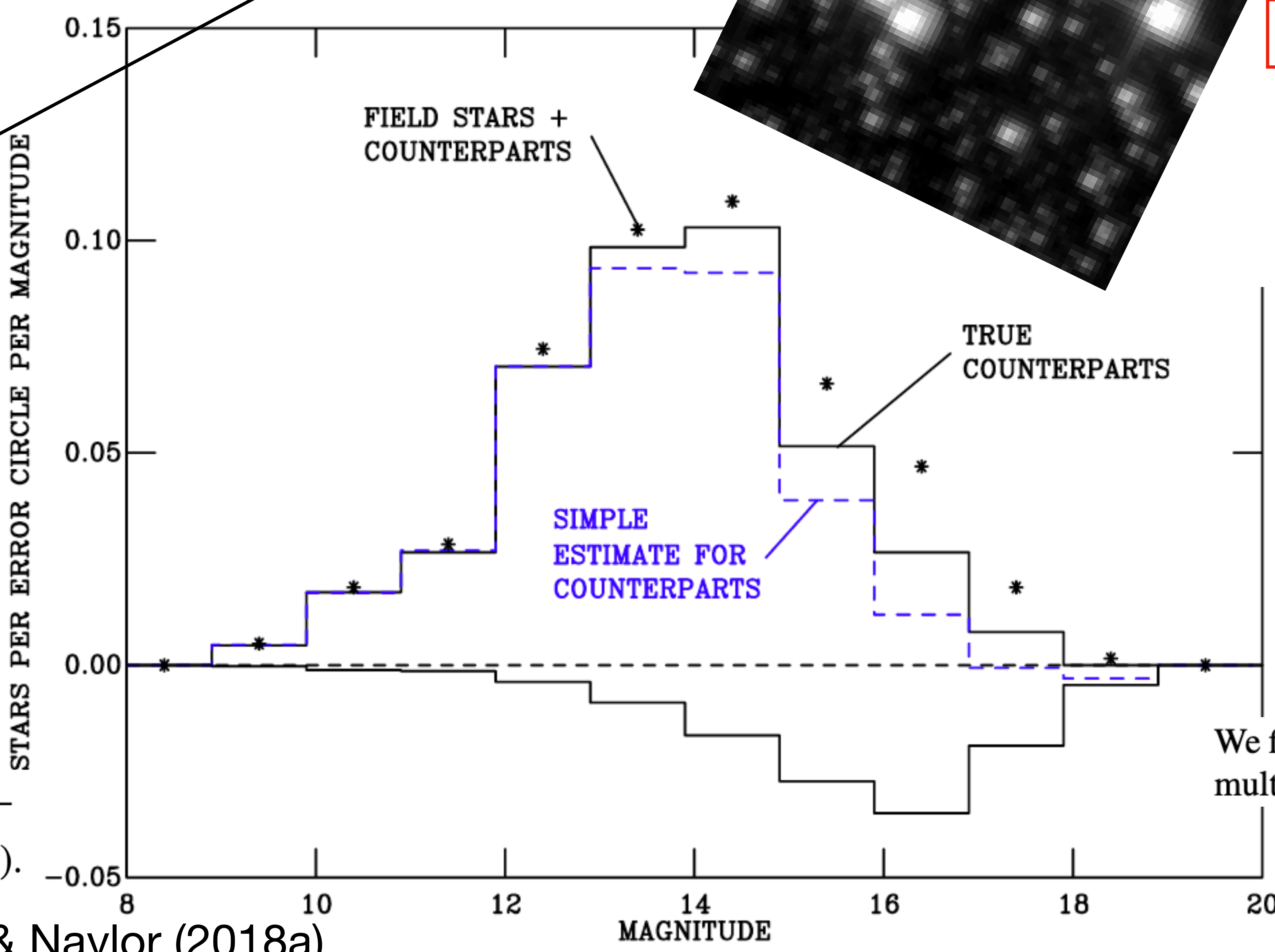
Photometry: Rejecting False Positives

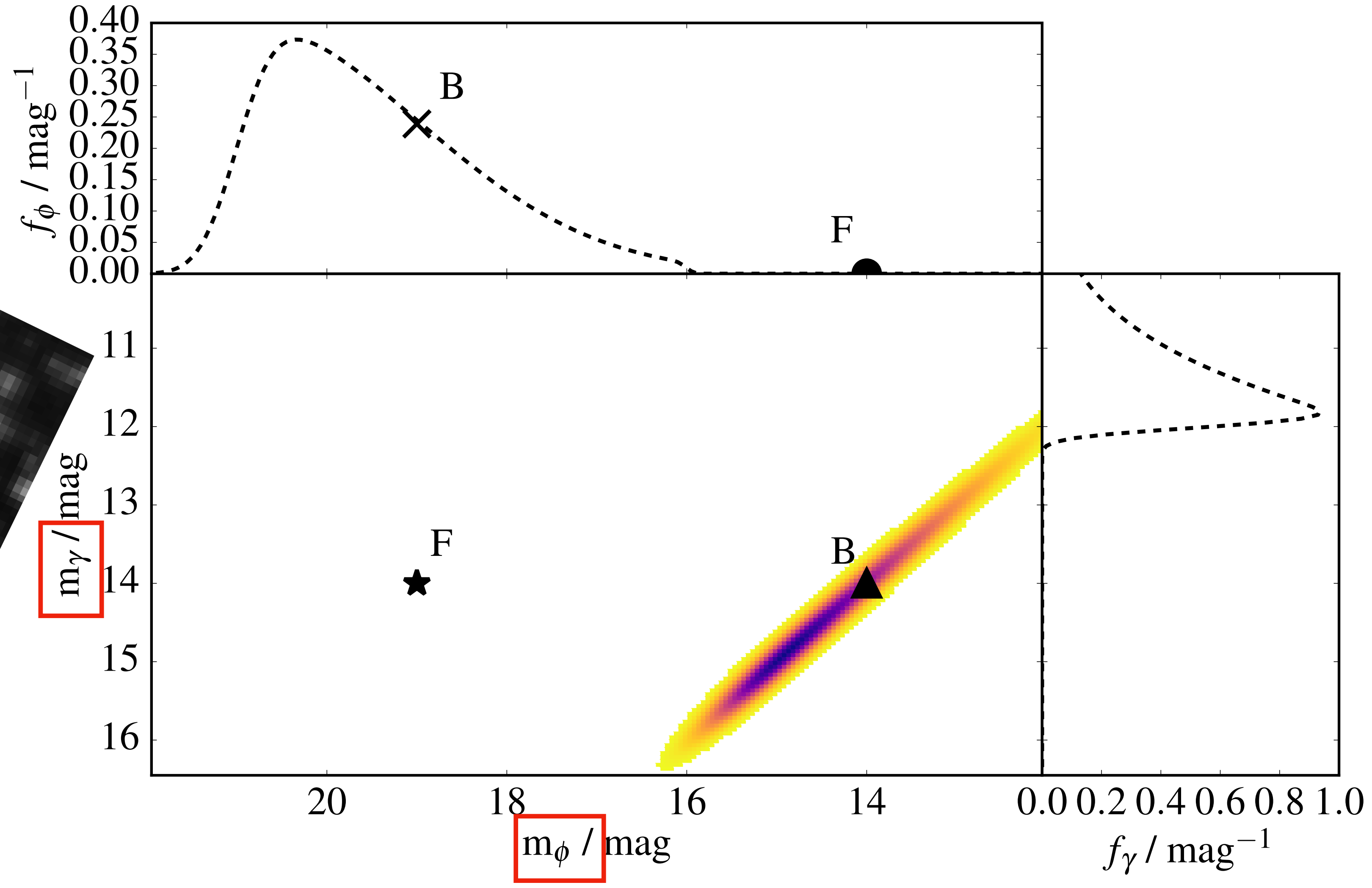
$$L = \frac{q(m, c)\, f(x, y)}{n(m, c)}$$

Counterparts = ("Counterparts" + field) - field









Where possible, <u>two-sided</u> photometry-based likelihoods (c and f) allow us to mitigate high false positive rate in crowded and confused fields

$$Z_{c\gamma} \cdot c_\gamma(m_\phi|m_\gamma) = Z_\gamma b_\gamma(m_\phi|m_\gamma) \exp\left(A_\gamma N_\phi F_\phi(m_\phi)\right) - \left(1 - Z_{c\gamma} C_\gamma(m_\phi|m_\gamma)\right) A_\gamma N_\phi f_\phi(m_\phi).$$

Naylor, Broos, & Feigelson (2013); Wilson & Naylor (2018a)

We find that 179 096 (92.5%) of the 250 μm sources in the SGP have multiple VIKING objects within 15 arcsec. Ward et al. (2022)

Tom J Wilson @onoddil

Probabilistic Cross-Matching

The Likelihood Ratio

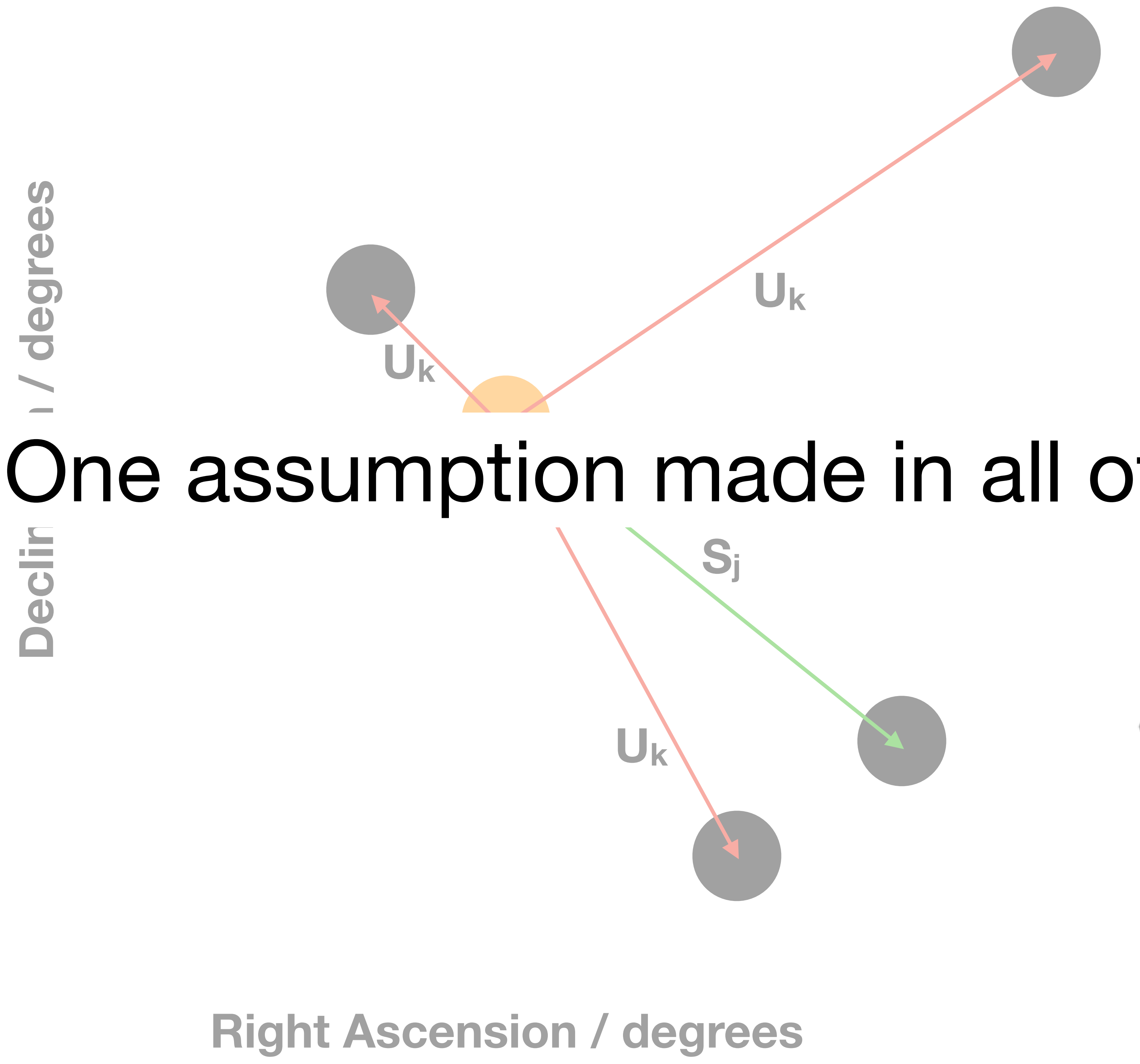
$$dp(r|id) = r \times e^{-r^2/2}\,dr.$$
$$dp(r|c) = 2\lambda r \times e^{-\lambda r^2}\,dr$$
$$LR(r) = dp(r|id)/dp(r|c) = \frac{1}{2\lambda}\exp\left\{\frac{r^2}{2}(2\lambda - 1)\right\}$$

de Ruiter, Willis, & Arp (1977)

$$dp_{\mathrm{id}} = Qr\exp\left(\frac{-r^2}{2}\right)dr. \quad dp_{\mathrm{uo}} = 2\lambda r\,dr$$
$$LR(r) = \frac{dp_{\mathrm{id}}}{dp_{\mathrm{uo}}} = \frac{Q\exp(-r^2/2)}{2\lambda}$$

Wolstencroft et al. (1986)

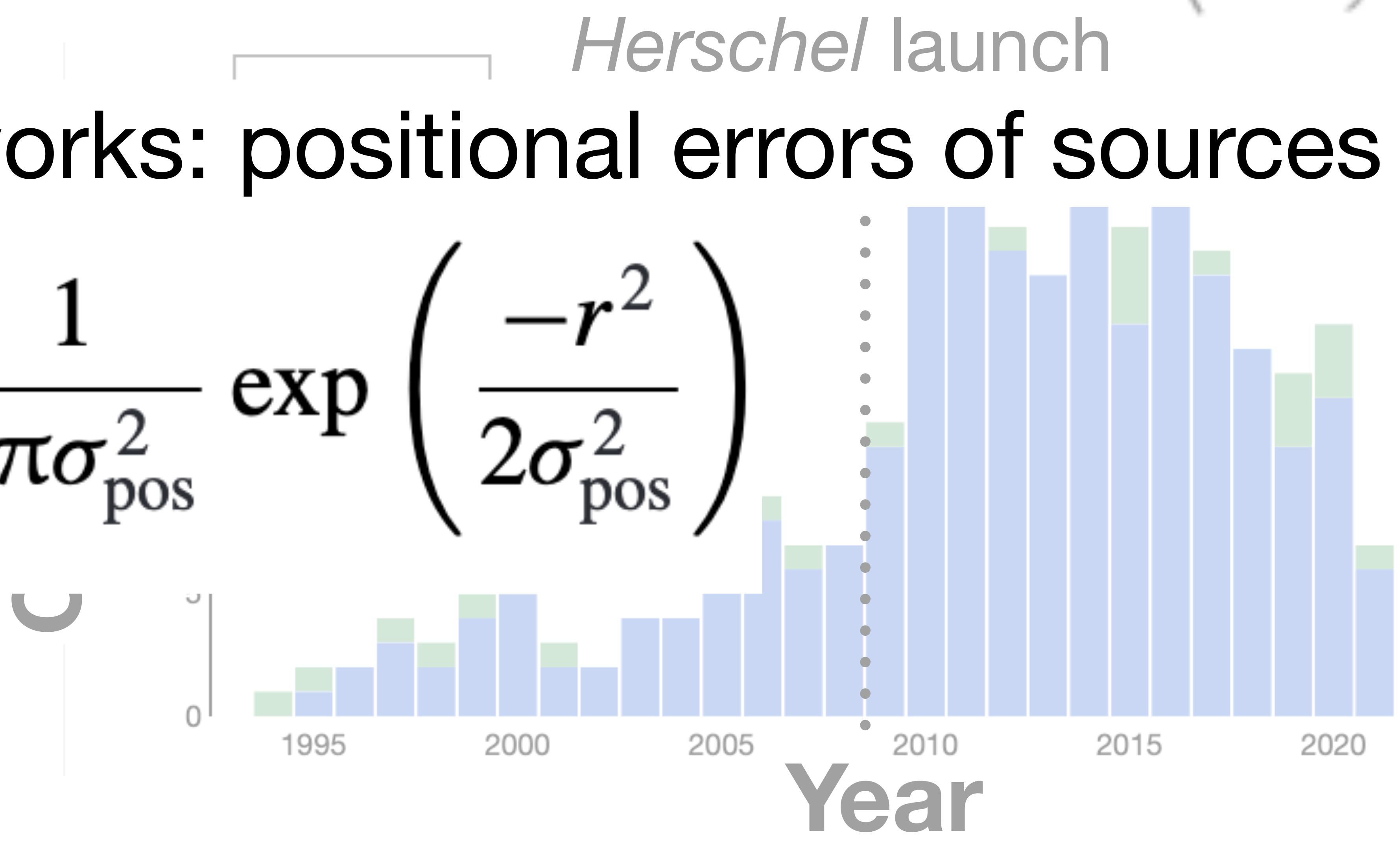
The "Reliability" — Sutherland & Saunders (1992)



$$R_j = \frac{L_j}{\sum_i L_i + (1 - Q)}$$
$$L = \frac{q(m, c)\, f(x, y)}{n(m, c)}$$

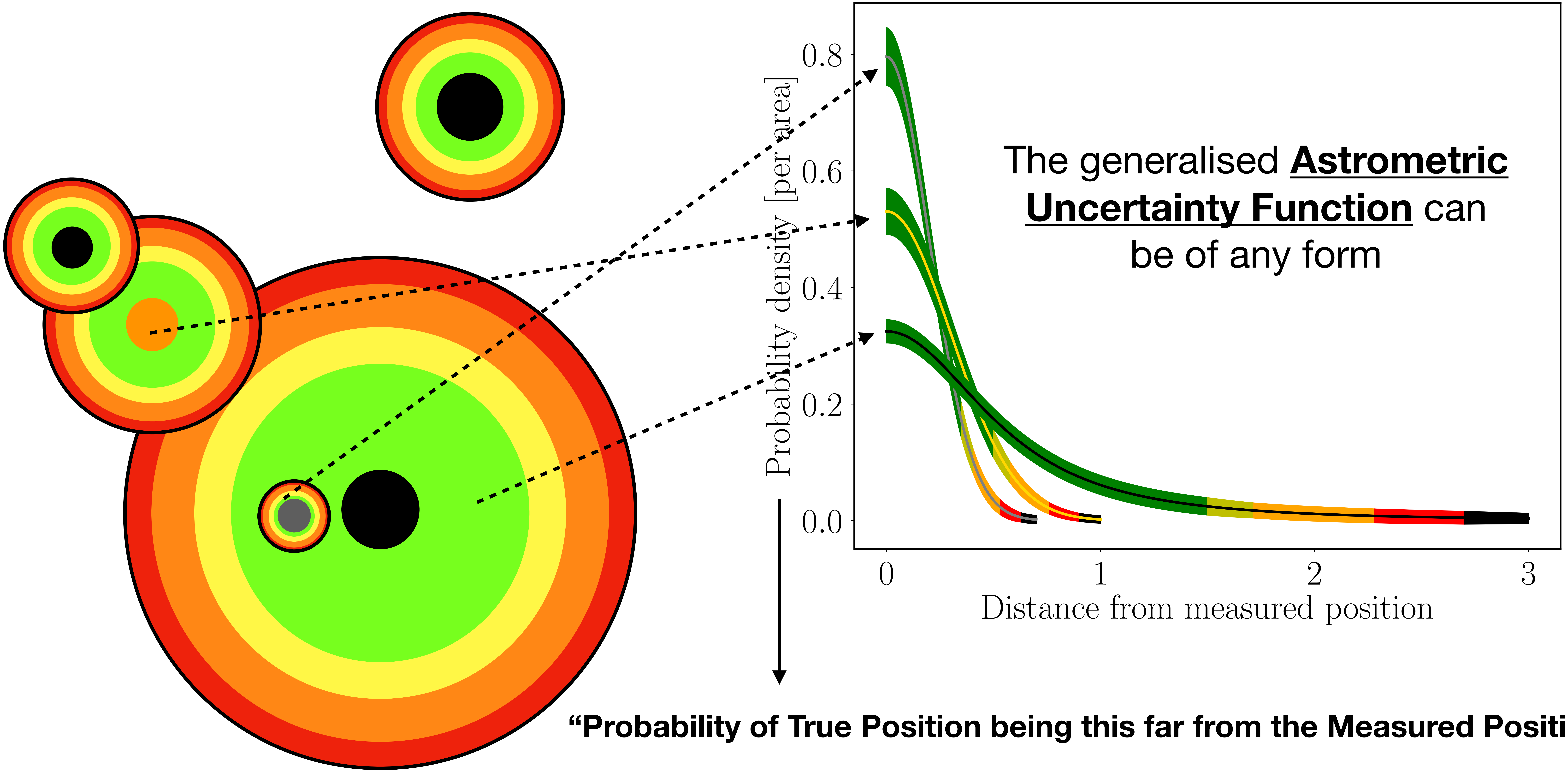
One assumption made in all of these works: positional errors of sources are Gaussian!

$$f(r) = \frac{1}{2\pi\sigma_{\mathrm{pos}}^2}\exp\left(\frac{-r^2}{2\sigma_{\mathrm{pos}}^2}\right)$$



Tom J Wilson @onoddil

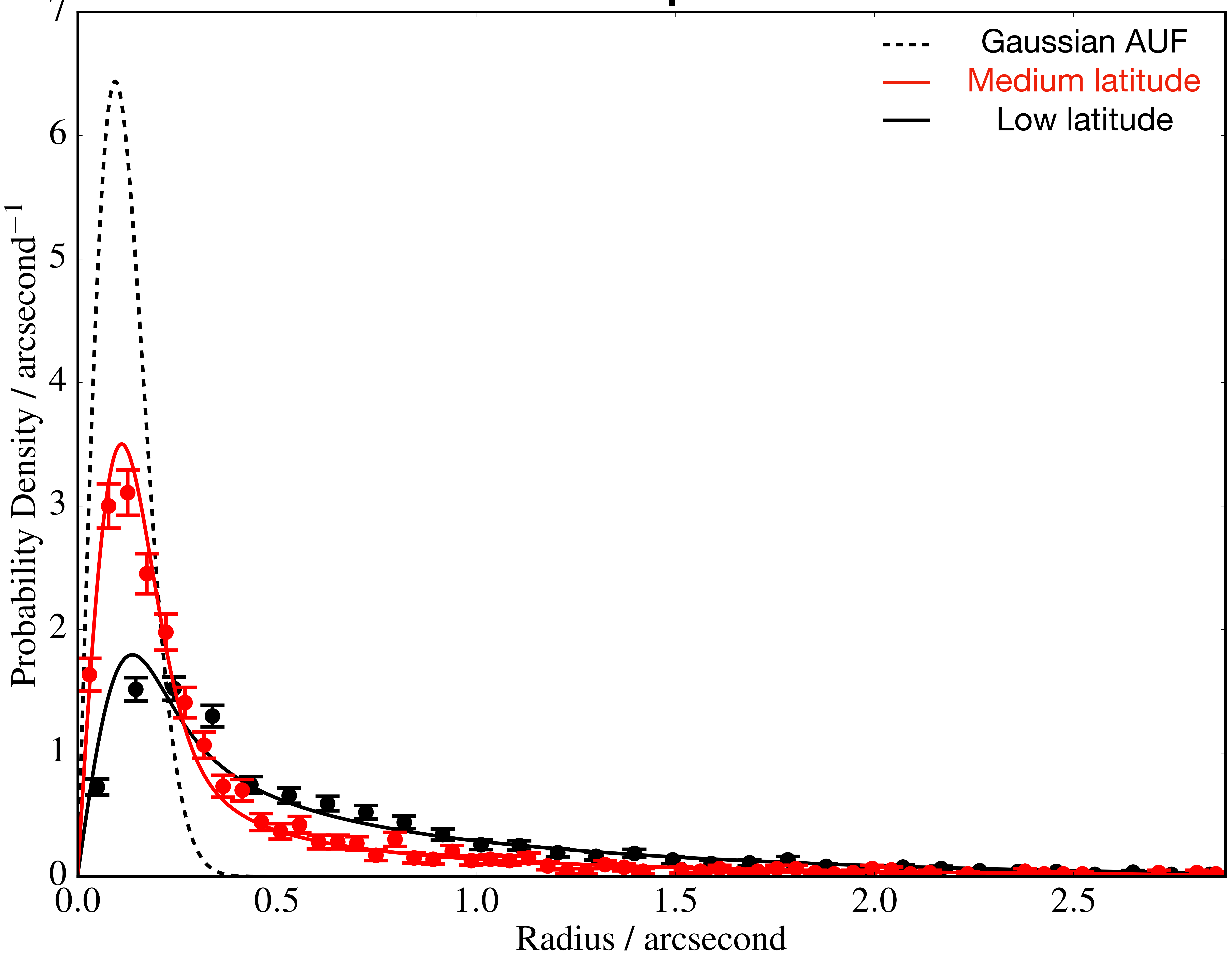
Probabilistic Cross-Matching: the AUF



"Probability of True Position being this far from the Measured Position"

Tom J Wilson @onoddil

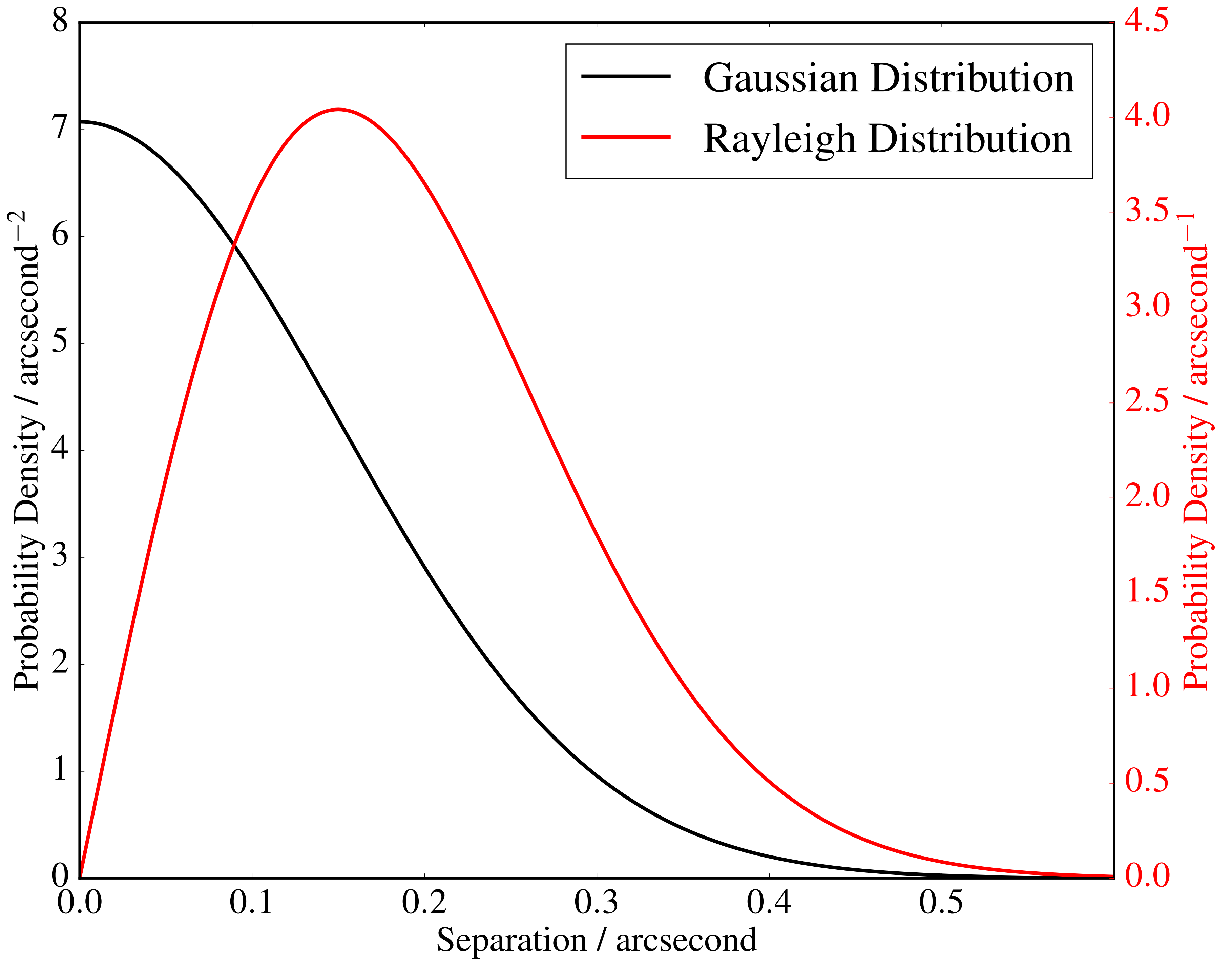
Additional Components of the AUF



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

$$\downarrow$$

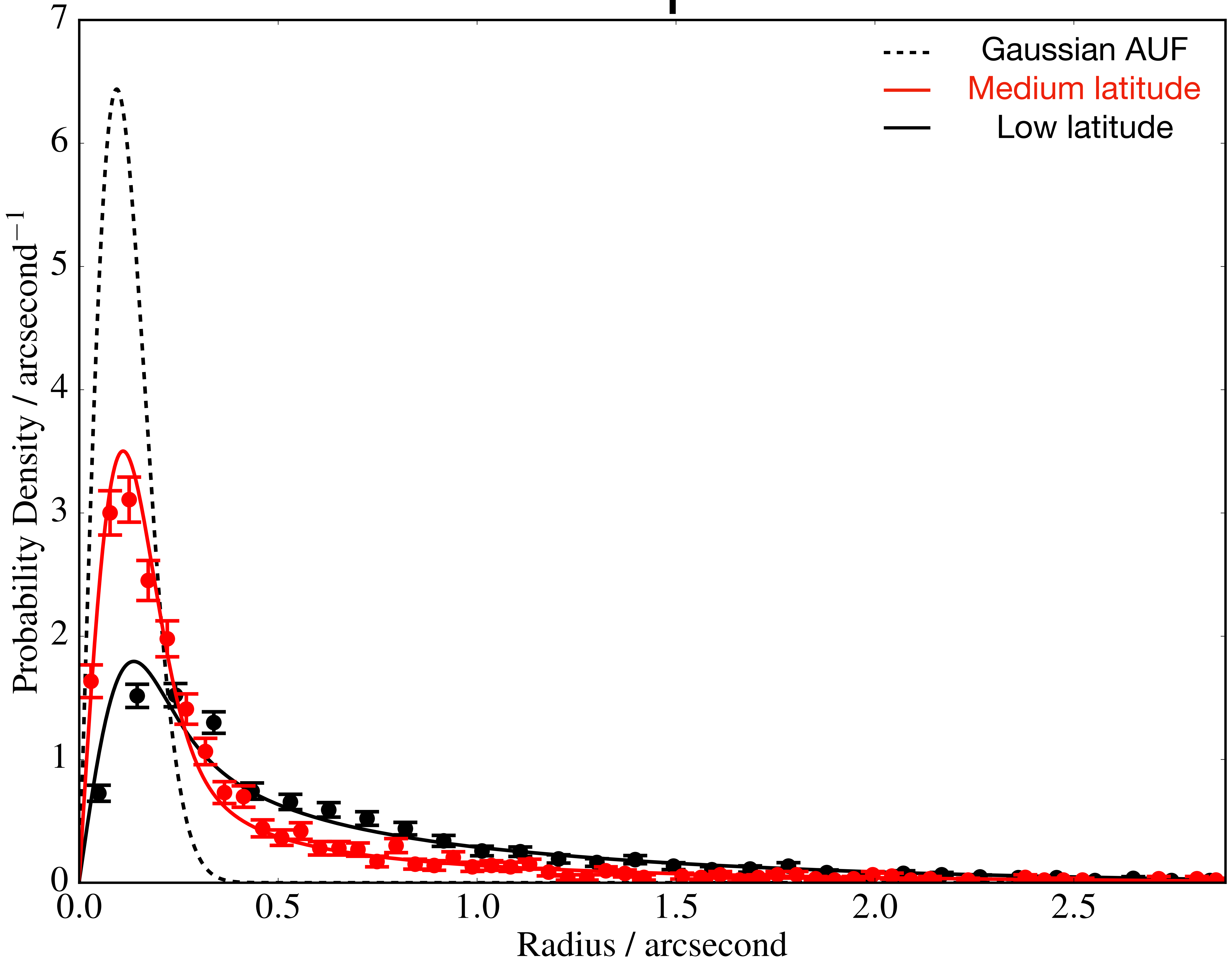
$$g(r, \sigma) = \frac{r}{\sigma^2}\exp\left(-\frac{1}{2}\frac{r^2}{\sigma^2}\right)$$

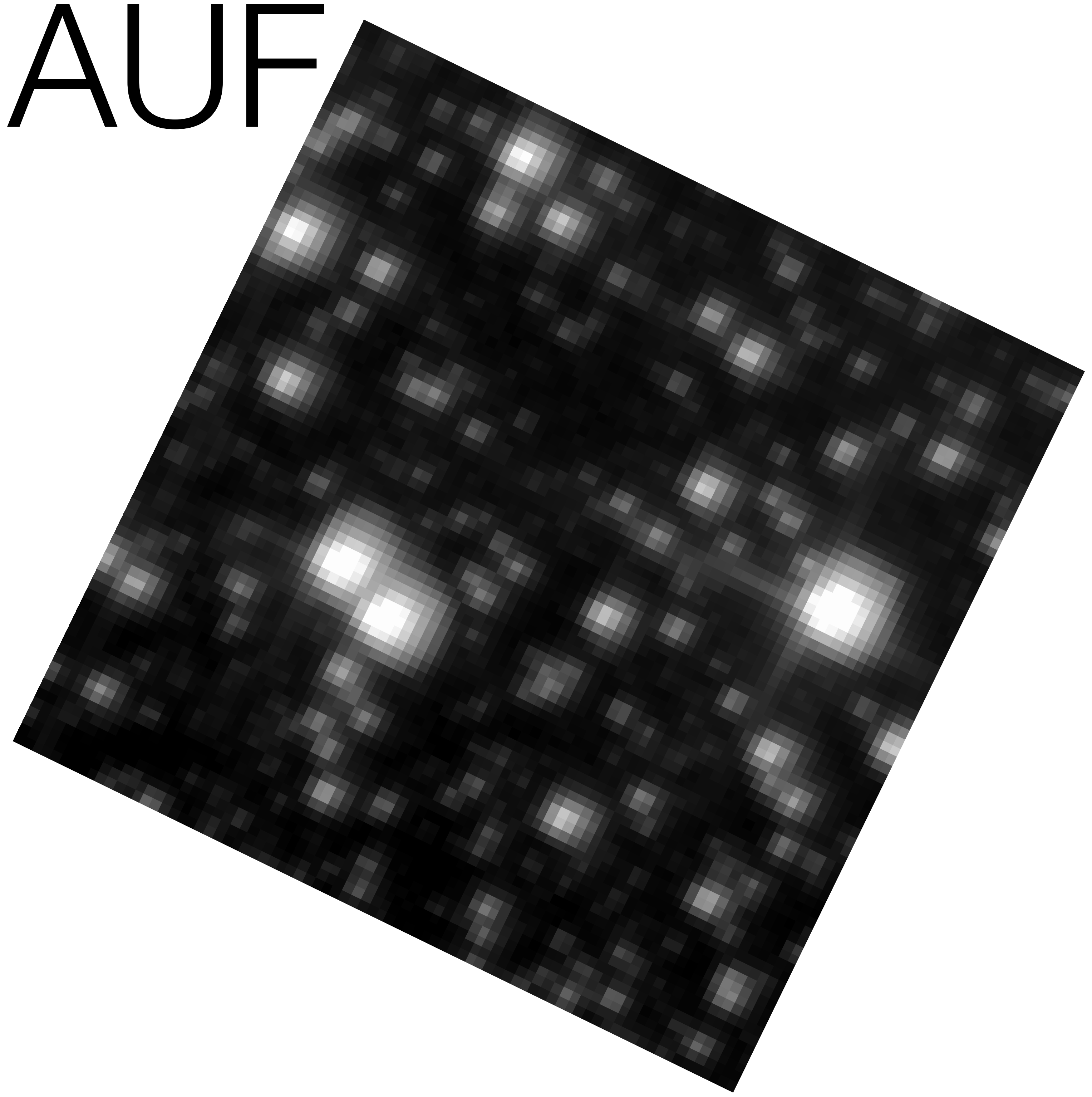


WISE - Wright et al. (2010)
Gaia DR2 - Gaia Collaboration, Brown A. G. A., et al. (2018)

Tom J Wilson @onoddil

Additional Components of the AUF



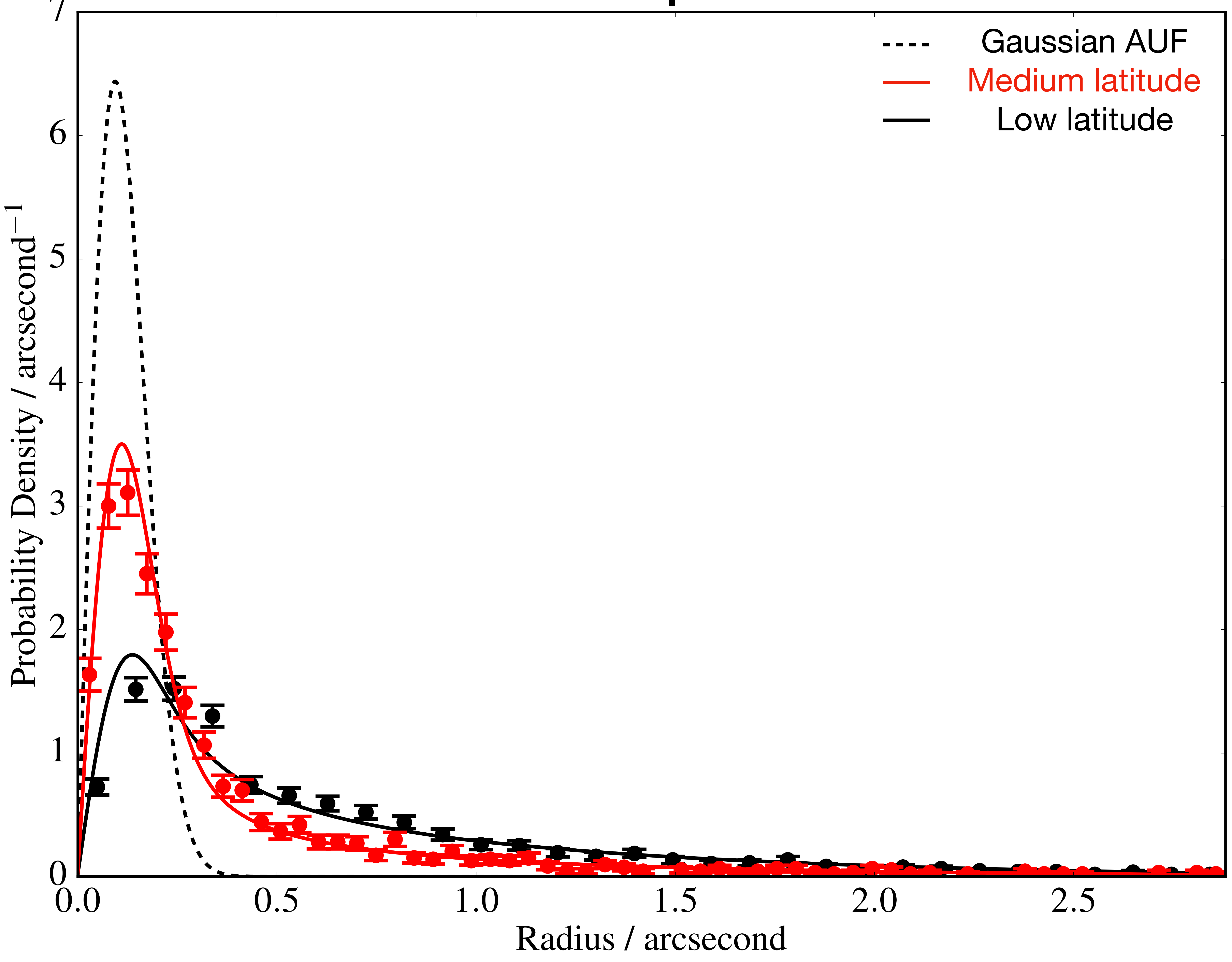


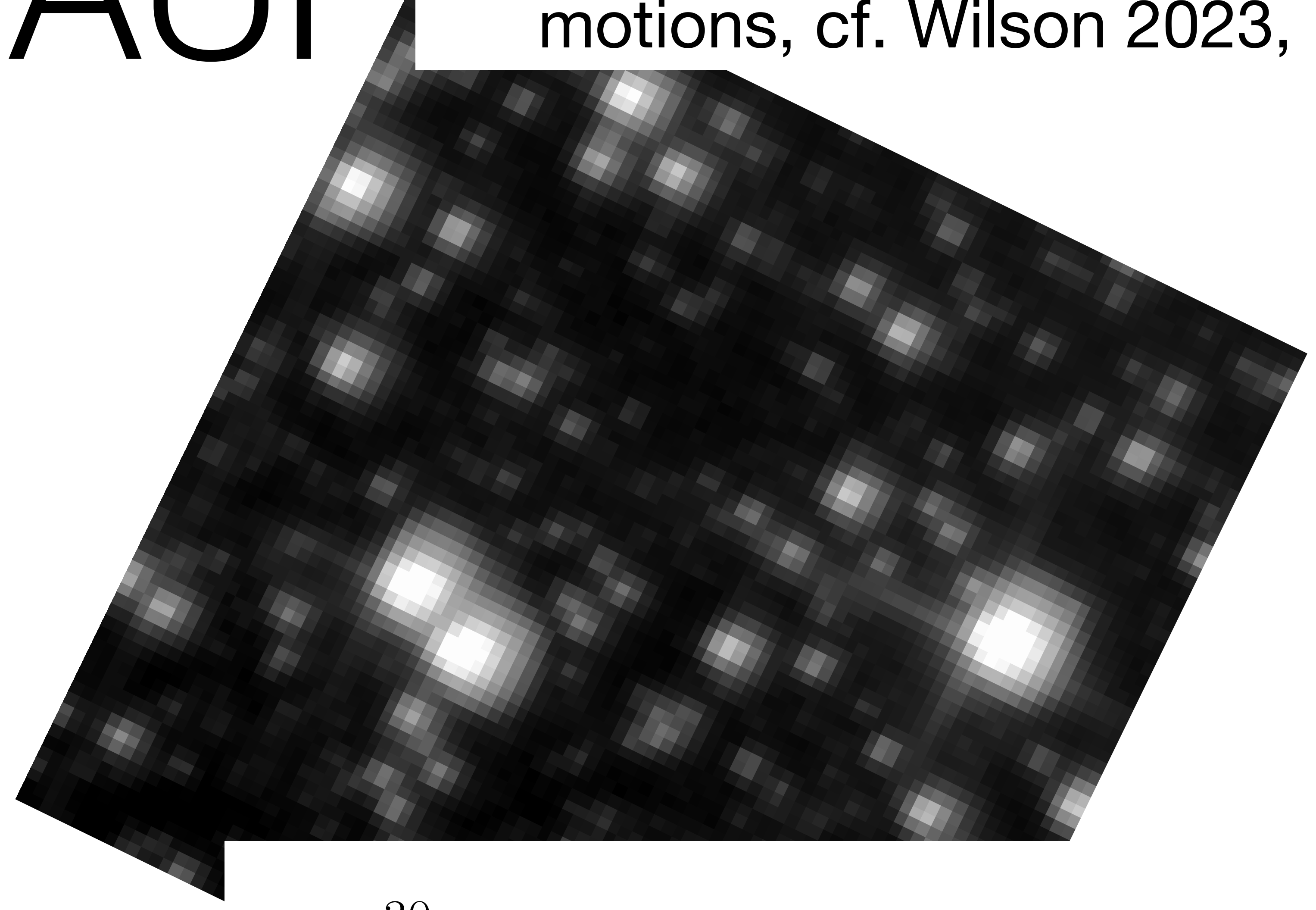
WISE - Wright et al. (2010)
Gaia DR2 - Gaia Collaboration, Brown A. G. A., et al. (2018)

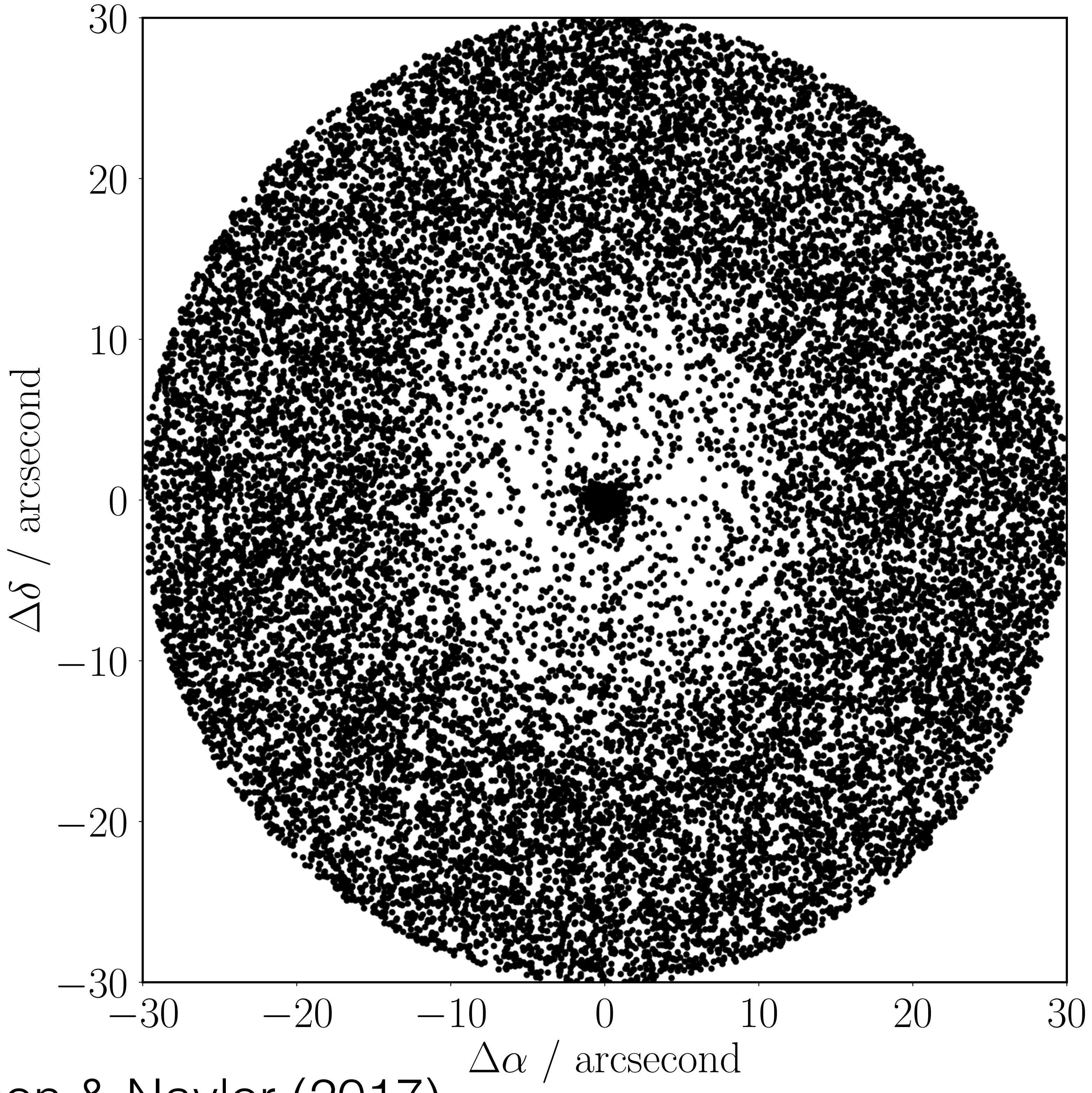
Tom J Wilson @onoddil

Additional Components of the AUF

(and any other systematic — e.g. proper motions, cf. Wilson 2023, RASTI)







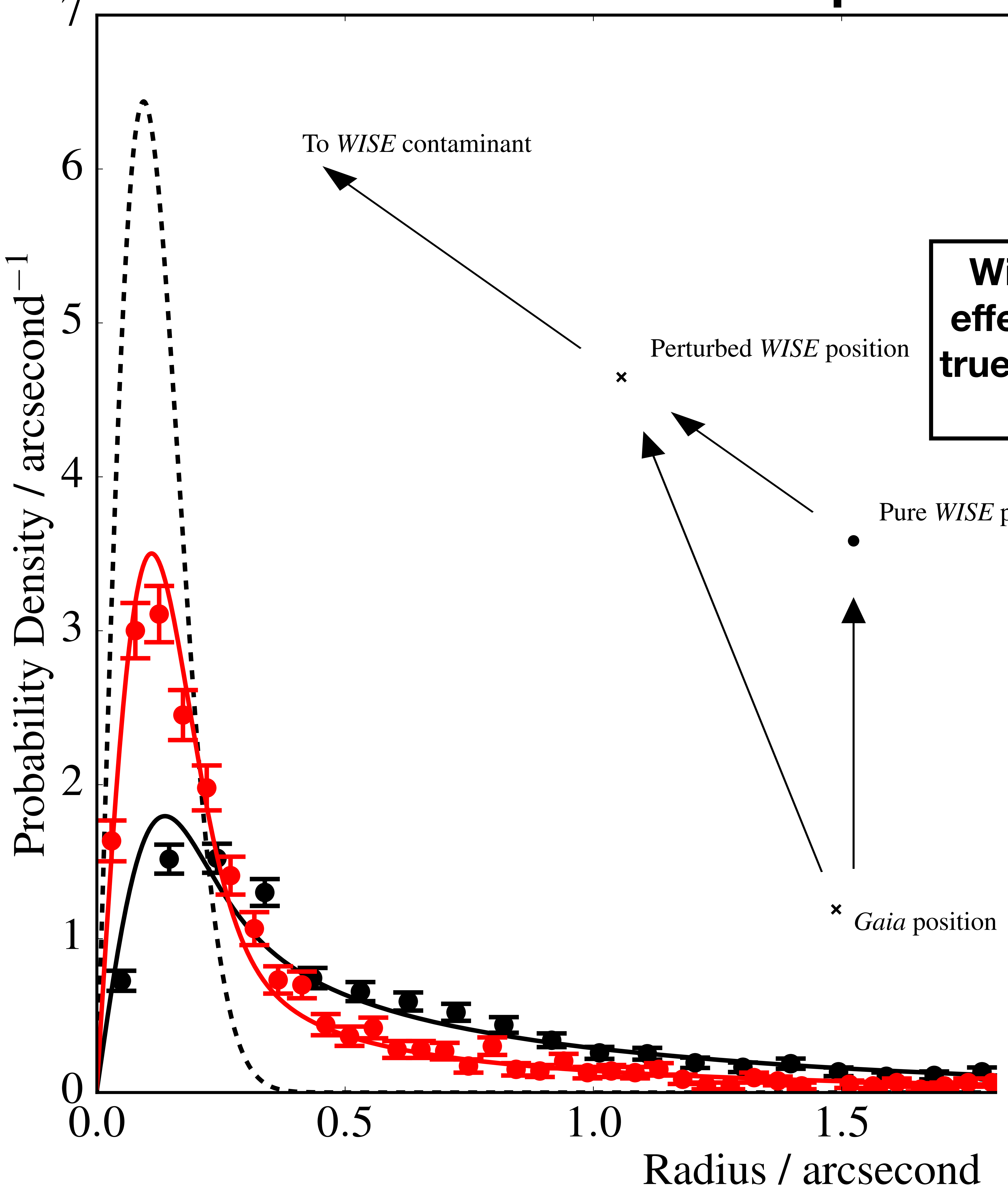
WISE - Wright et al. (2010)
Gaia DR2 - Gaia Collaboration, Brown A. G. A., et al. (2018)

Wilson & Naylor (2017)

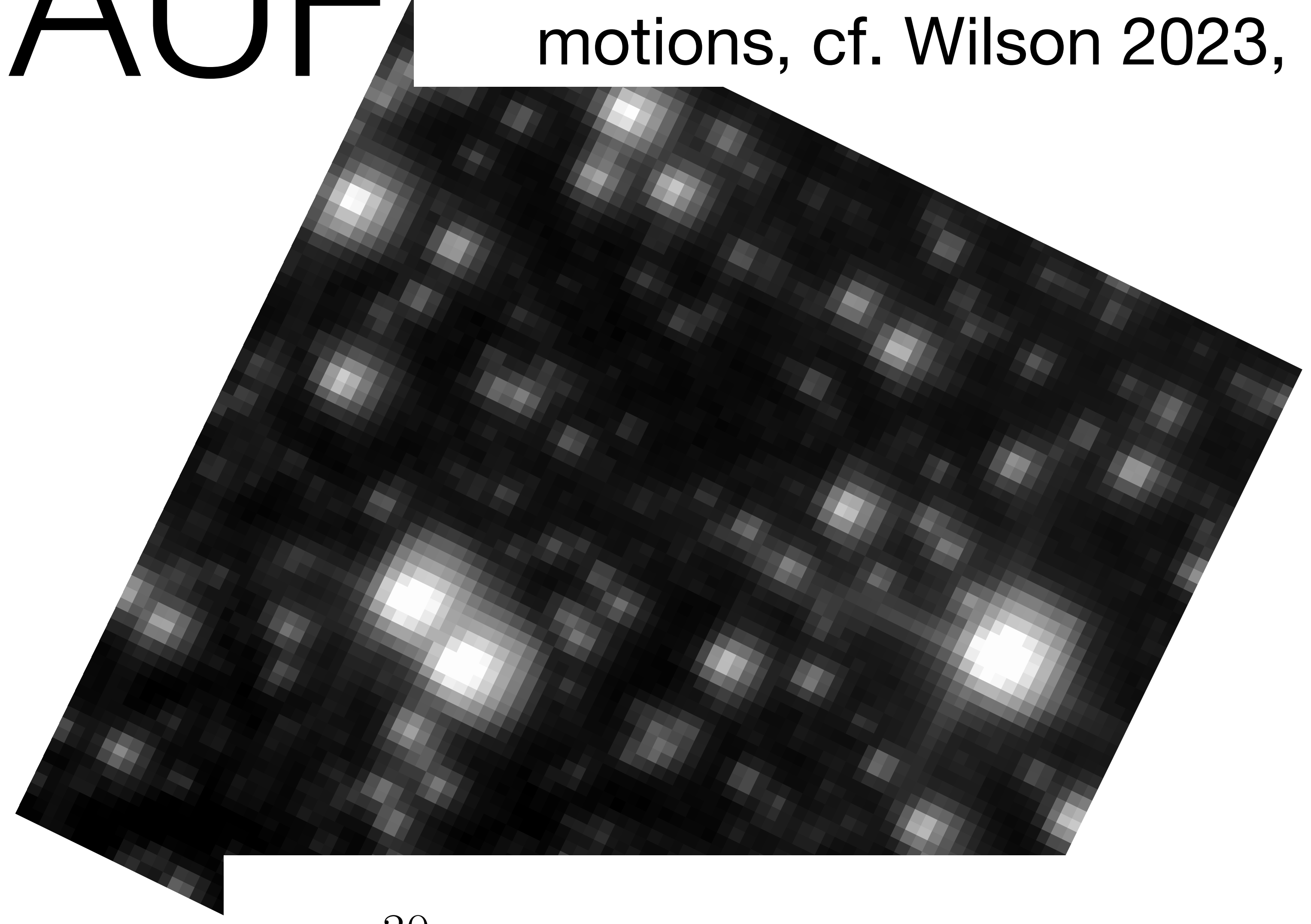
Tom J Wilson @onoddil

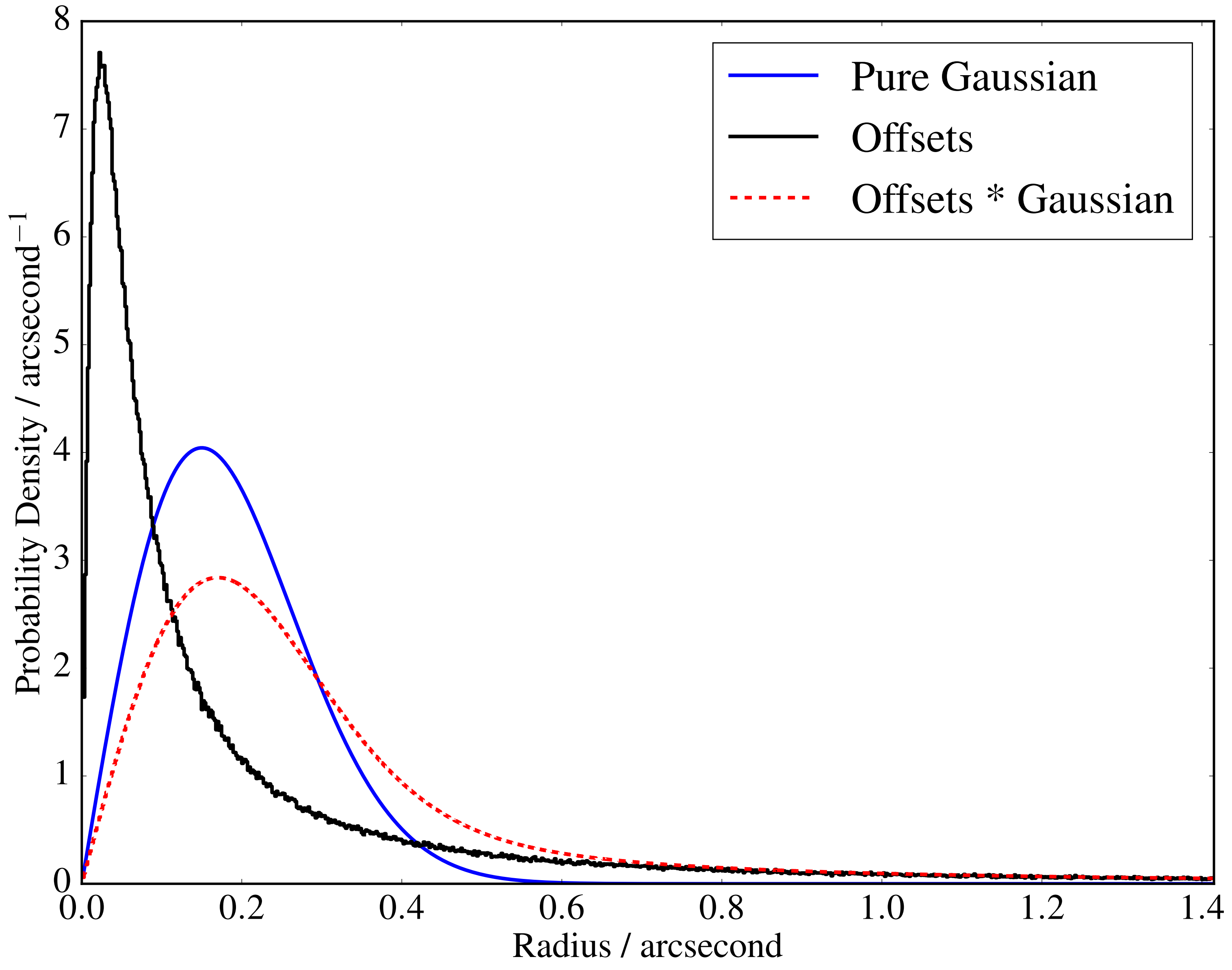
Additional Components of the AUF

(and any other systematic — e.g. proper motions, cf. Wilson 2023, RASTI)

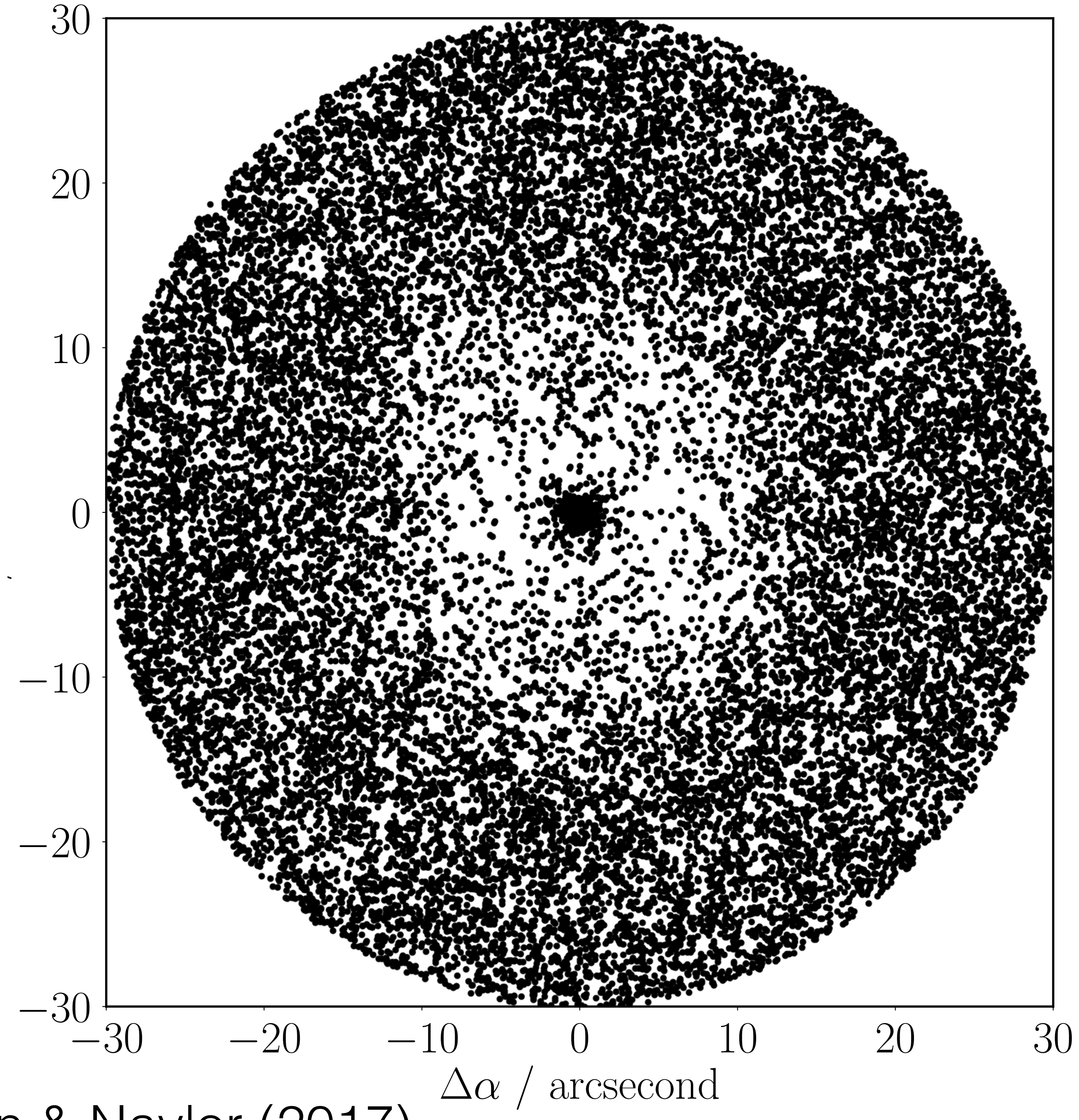


Without modelling this extra effect, we fail to recover many true pairings, with an artificially high false negative rate!





Wilson & Naylor (2018b)

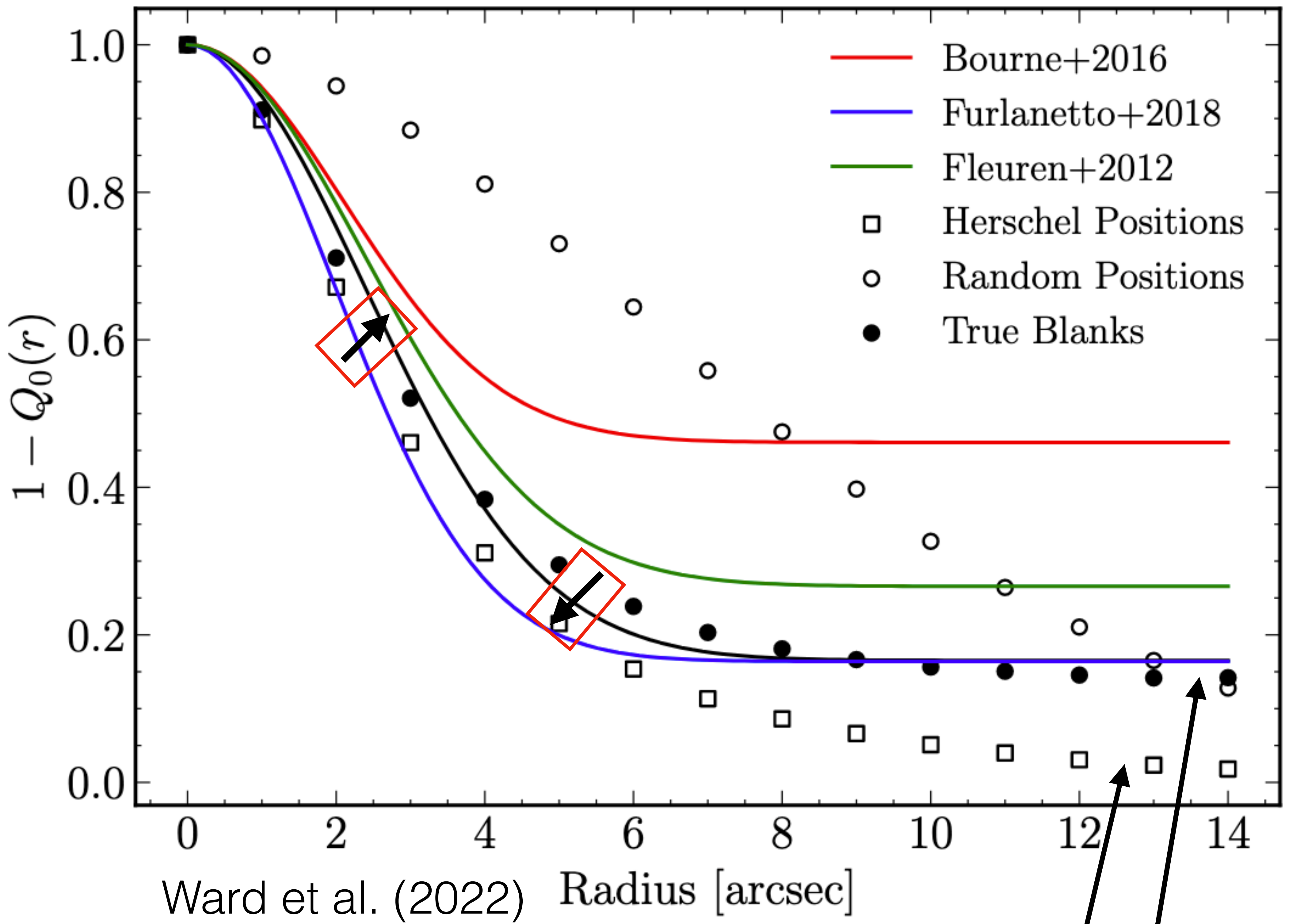


Wilson & Naylor (2017)

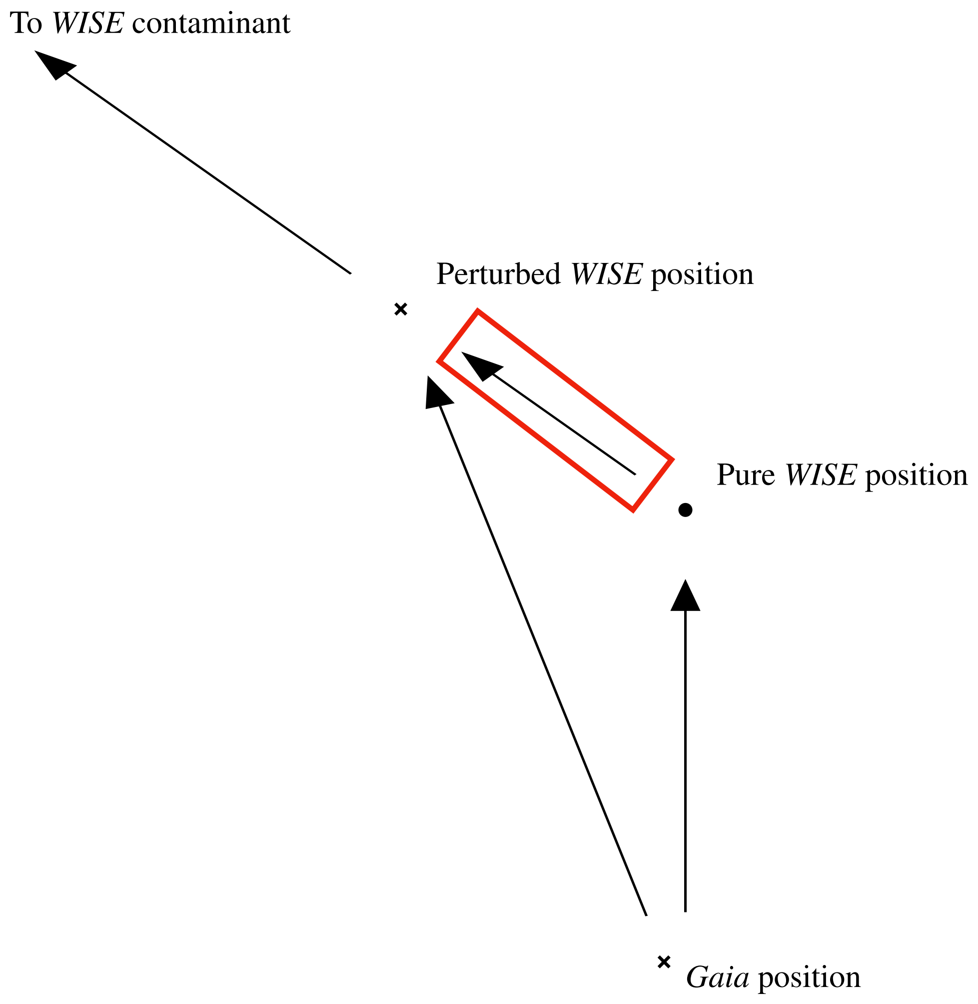
WISE - Wright et al. (2010)
Gaia DR2 - Gaia Collaboration, Brown A. G. A., et al. (2018)

Tom J Wilson @onoddil

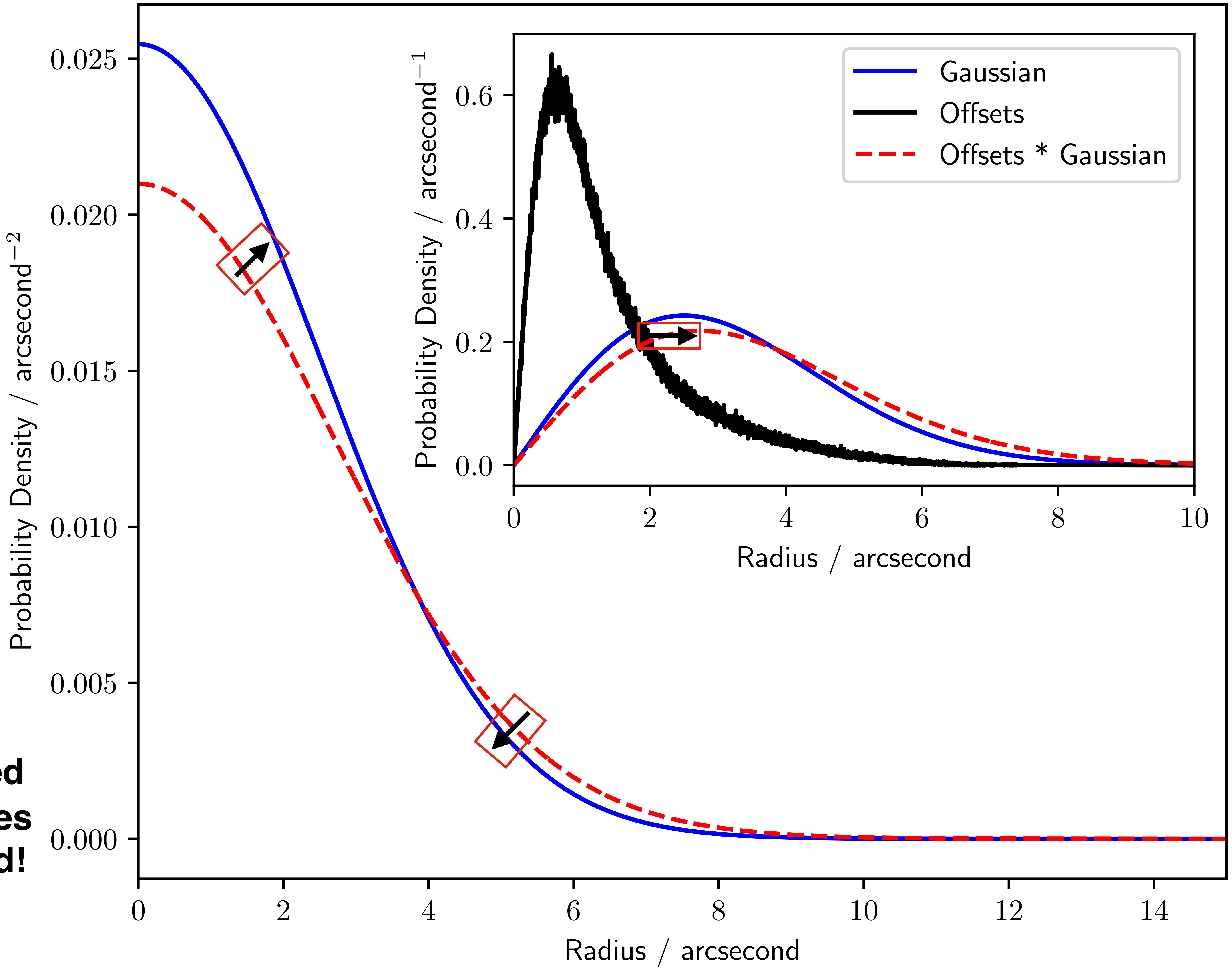
The Perturbation Component of the AUF



Ward et al. (2022)

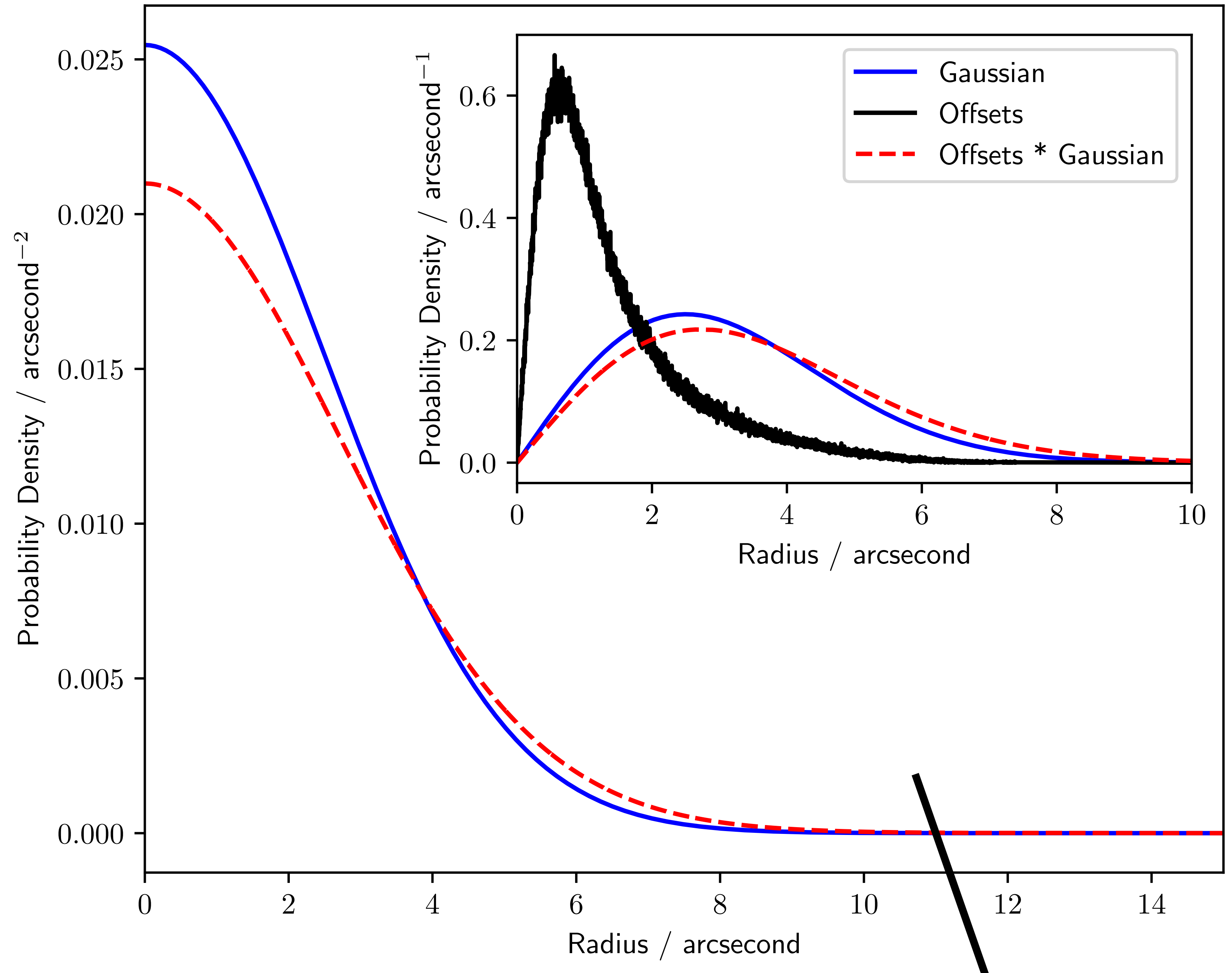


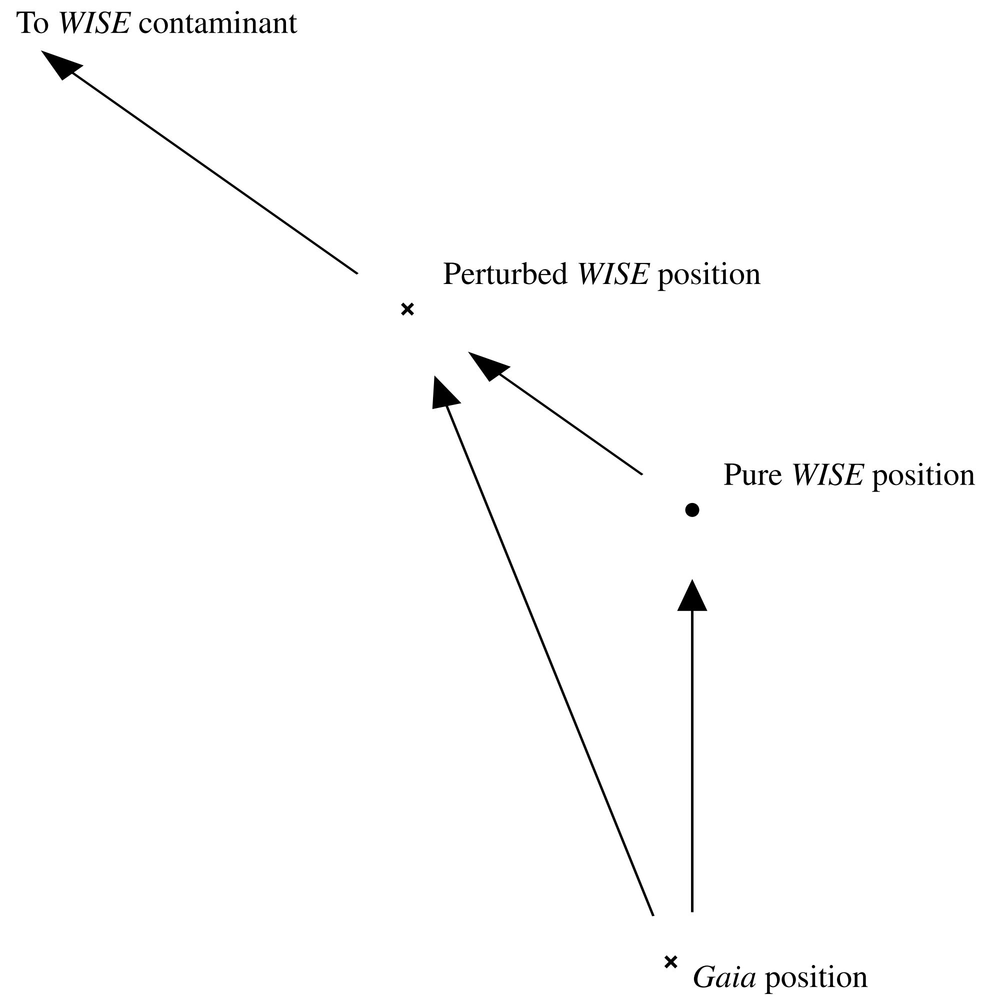
$1 - Q_0$ has not stabilised yet — fraction of matches is higher than calculated!



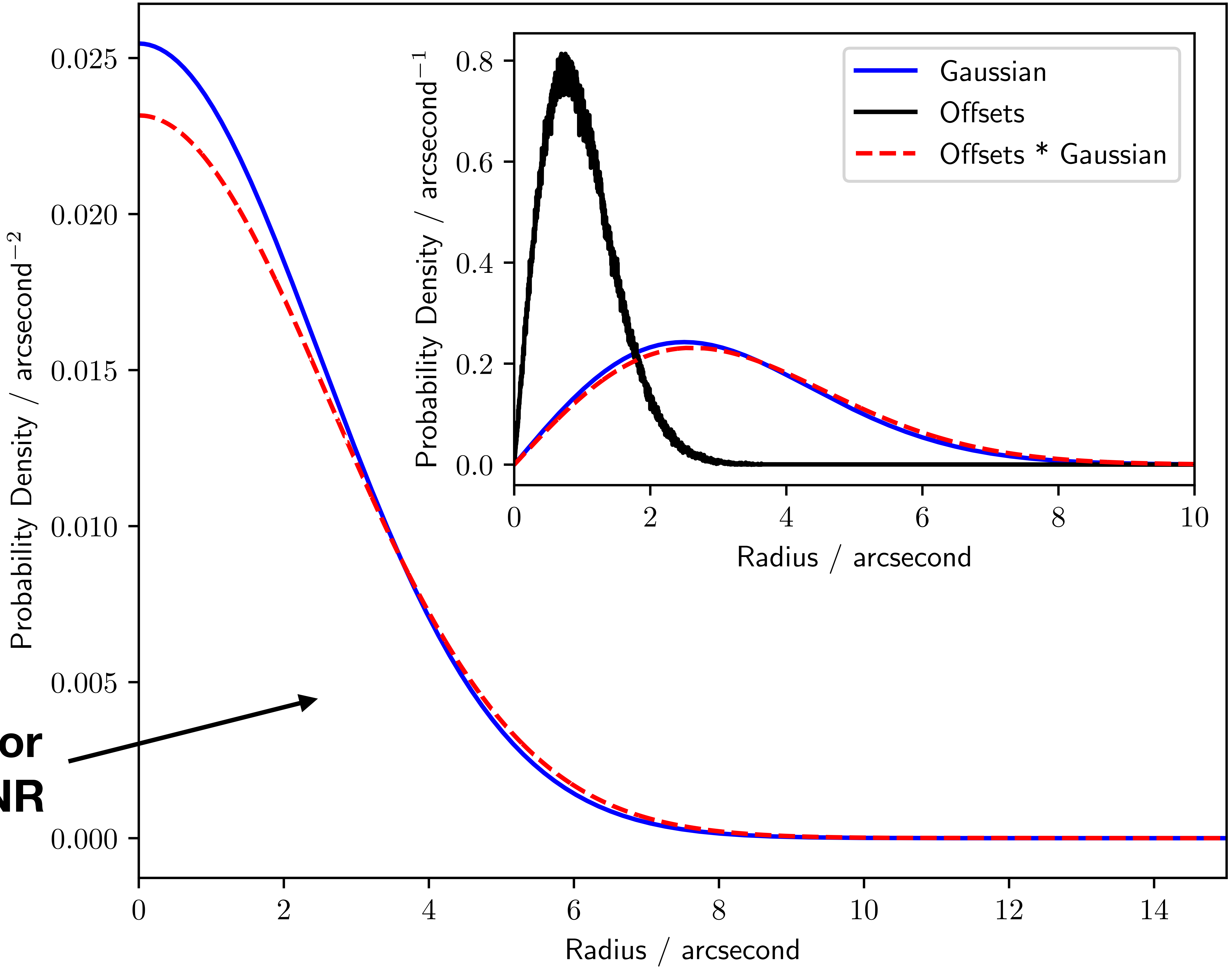
Tom J Wilson @onoddil

The Perturbation Component of the AUF



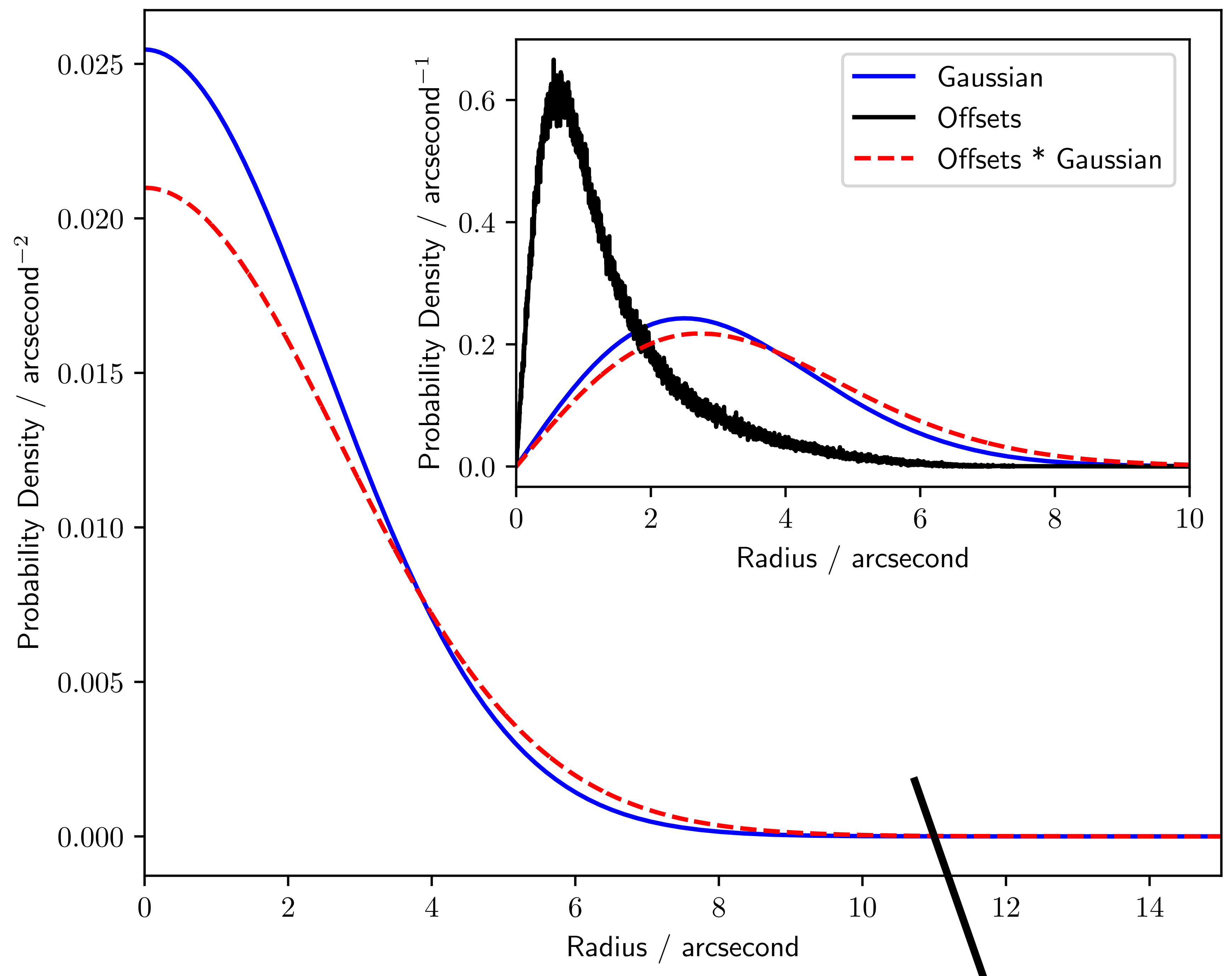


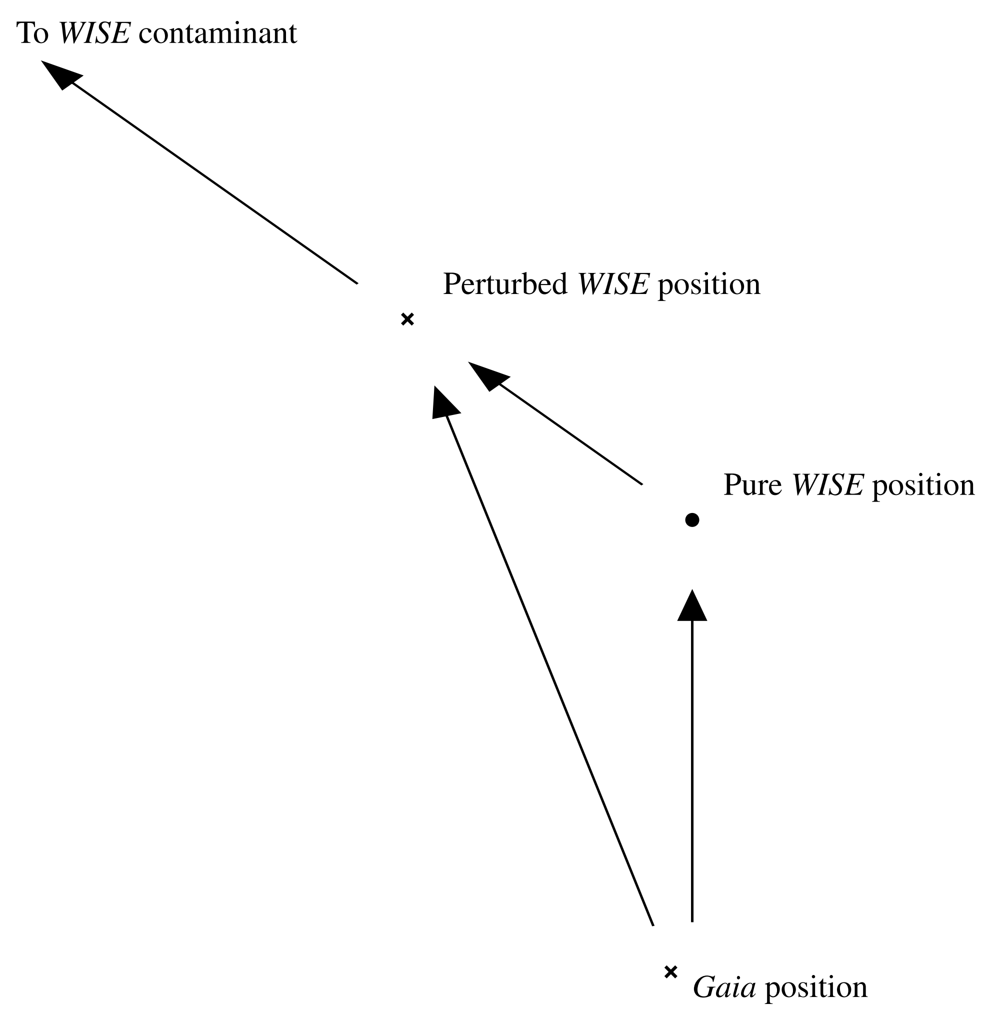
10x fainter for the same SNR



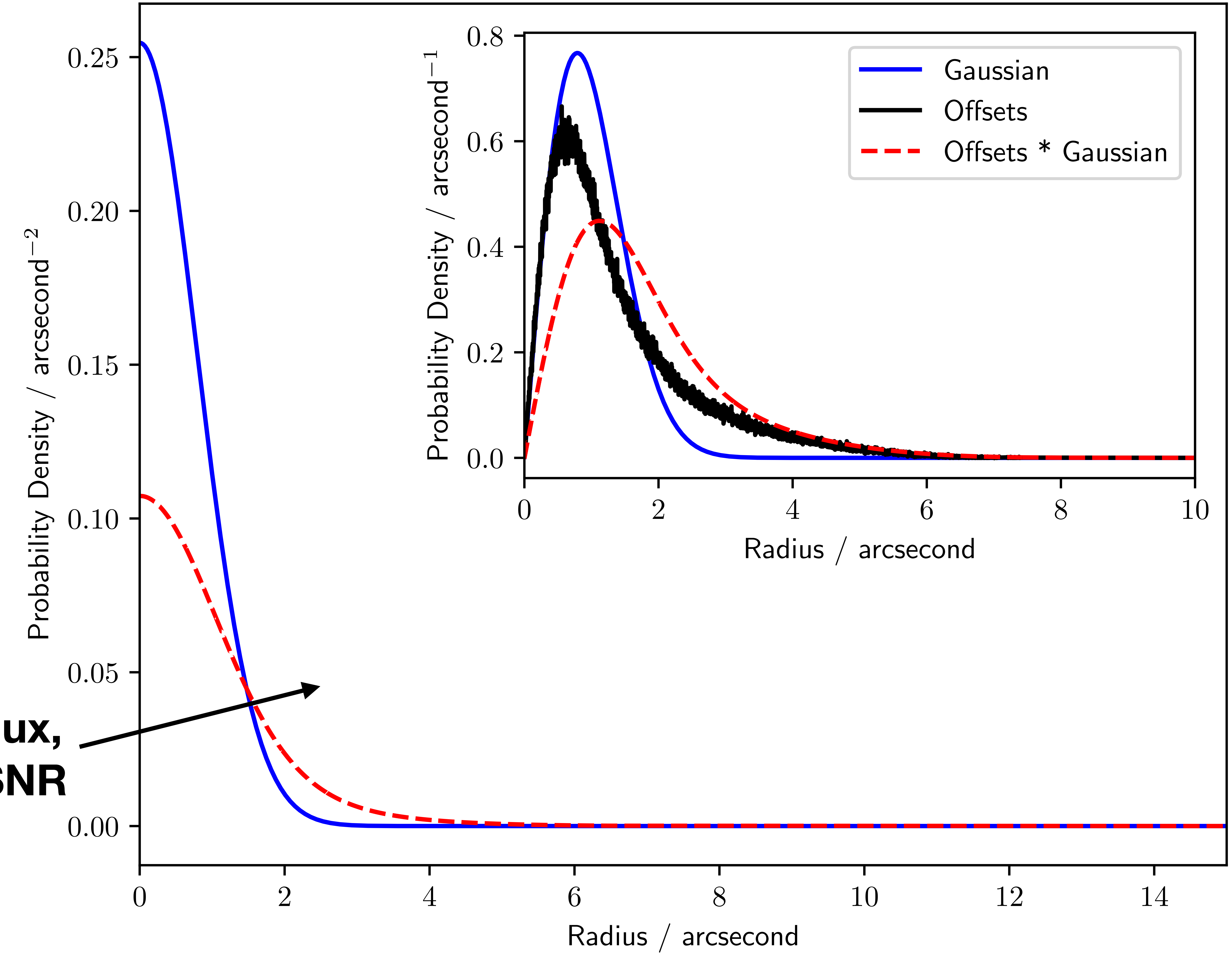
Tom J Wilson @onoddil

The Perturbation Component of the AUF



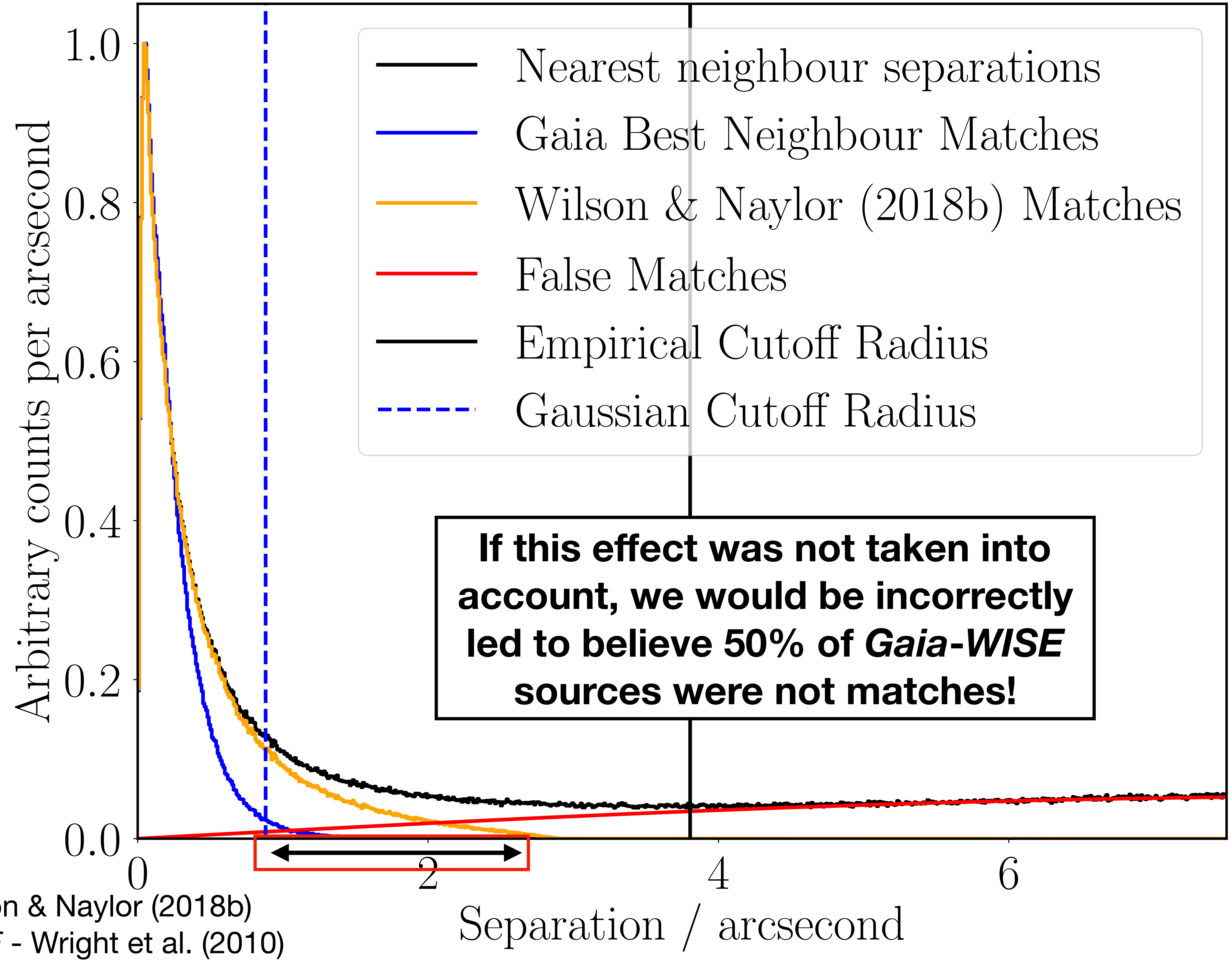


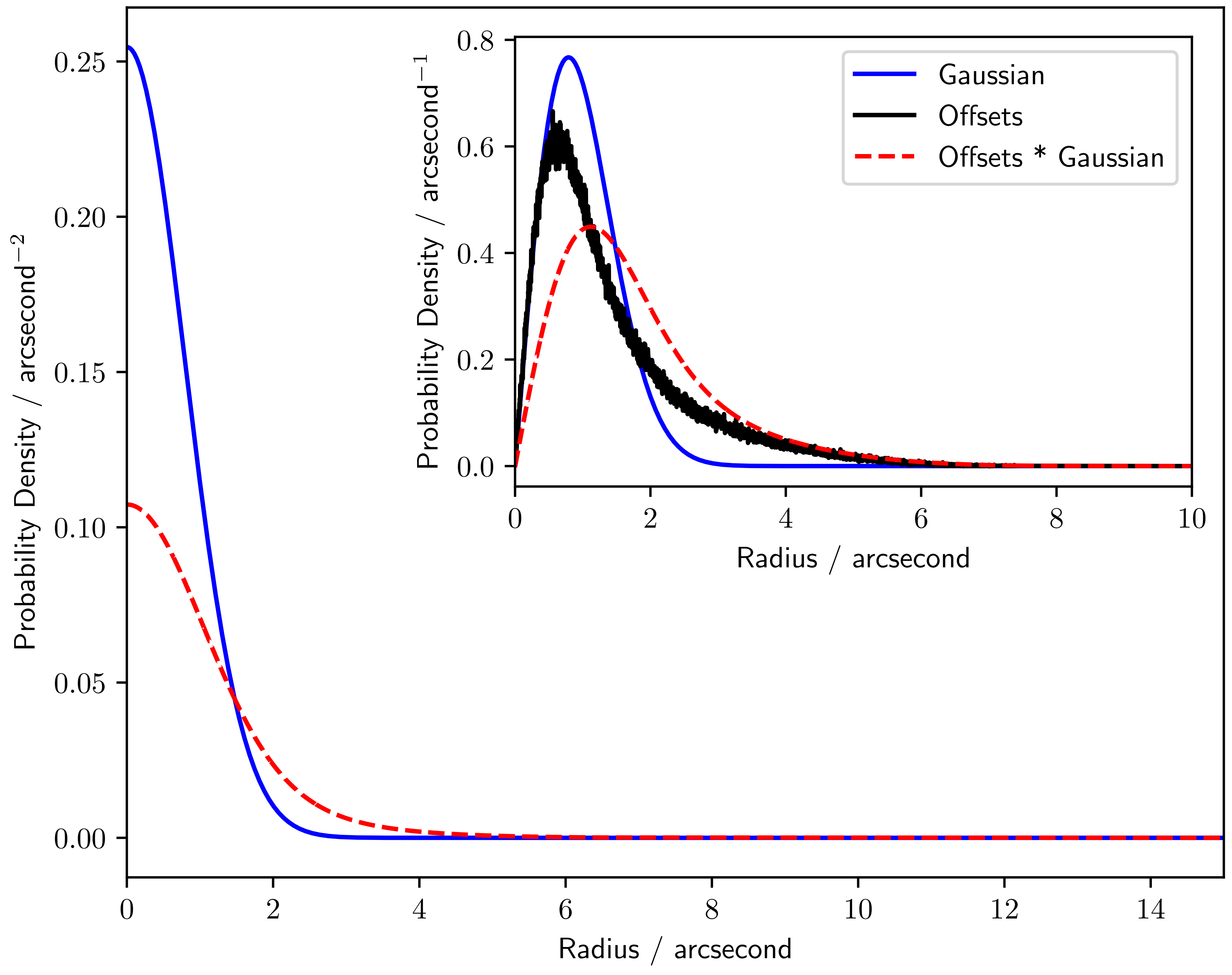
Same flux, better SNR

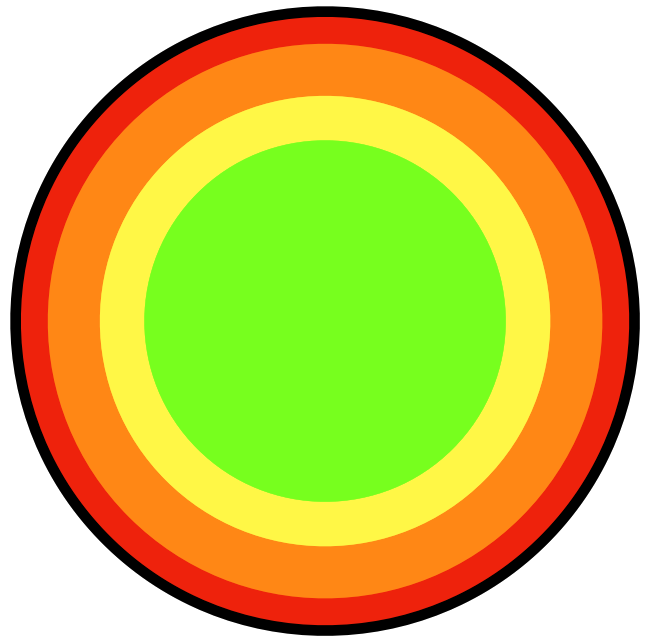


Tom J Wilson @onoddil

Match Separations







The AUF does not need to, and in fact quite often <u>should *not*</u>, be Gaussian!

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

Conclusions

- **Our cross-match algorithms include two key elements to avoid issues with crowded & confused data**
  - **A generalised approach to the Astrometric Uncertainty Function allows for the full inclusion of the effects of perturbation due to blended sources — reduce false -ves!**
  - **Use of (two-sided) photometry to sort out multiplicity of higher resolution data — reduce false +ves!**

- **Missing extra perturbation from blended sources has the effect of increasing $1 - Q_0$ and decreasing expected fraction of sub-mm/far-IR counterparts to shorter-wavelength datasets**

- **Software package macauff developed to cross-match catalogues, including the effect of unresolved contaminant sources and rejection of interloper objects using photometric information**
  - **Developed through Rubin/LSST:UK, with plans to match LSST to *Gaia*, *WISE*, VISTA, SDSS, …**
  - **We have compute time to cross-match datasets — let me know your favourite combo, and what you need matched (to LSST or otherwise)!**
- **Will be able to handle the increased effects of perturbation due to unresolved sources in the next-generation of far-IR data — crucial as source densities and sensitivities increase in future surveys**

University of Exeter

@Onoddil@pm.me

@Onoddil.github.io

LSST:UK Consortium

UKRI Science and Technology Facilities Council

Wilson & Naylor, 2017, MNRAS, 468, 2517

Wilson & Naylor, 2018a, MNRAS, 473, 5570

Wilson & Naylor, 2018b, MNRAS, 481, 2148

Wilson, 2022, RNAAS, 6, 60

Wilson, 2023, RASTI, 2, 1

https://github.com/Onoddil/macauff

Macauff