Bayesian inference using Gaia data

Coryn Bailer-Jones
Max Planck Institute for Astronomy, Heidelberg

What is Bayes?

- an approach to learning (= inference)
- given data on a phenomenon, determine how well a model explains the data
- "how well" is quantified using probabilities

Priors and subjectivity

- learning: reconcile new data with existing knowledge
- "existing knowledge" = prior
  - e.g. limits, monotonicity, smoothness
- subjective, just like many other data analysis steps
  - what data do we decide to collect?
  - what data do we discard?
  - what assumptions and approximation do we make?
- smooth transition from data-dominated to prior-dominated

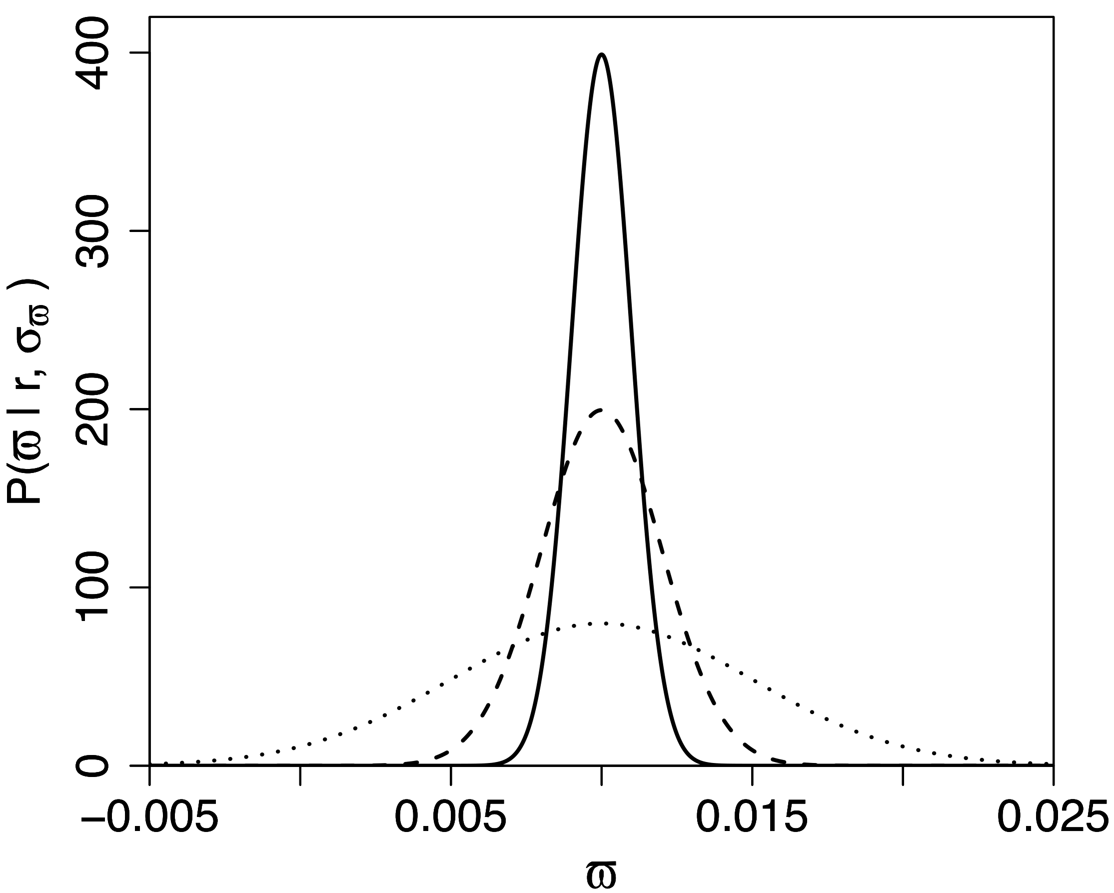
Inference of distance from a parallax

ϖ parallax
σ_ϖ parallax uncertainty
r distance

$$\varpi/\mathrm{as} = \frac{1}{r/\mathrm{pc}}$$

$$f_{\mathrm{obs}} = \frac{\sigma_\varpi}{\varpi}$$

$r = 100$

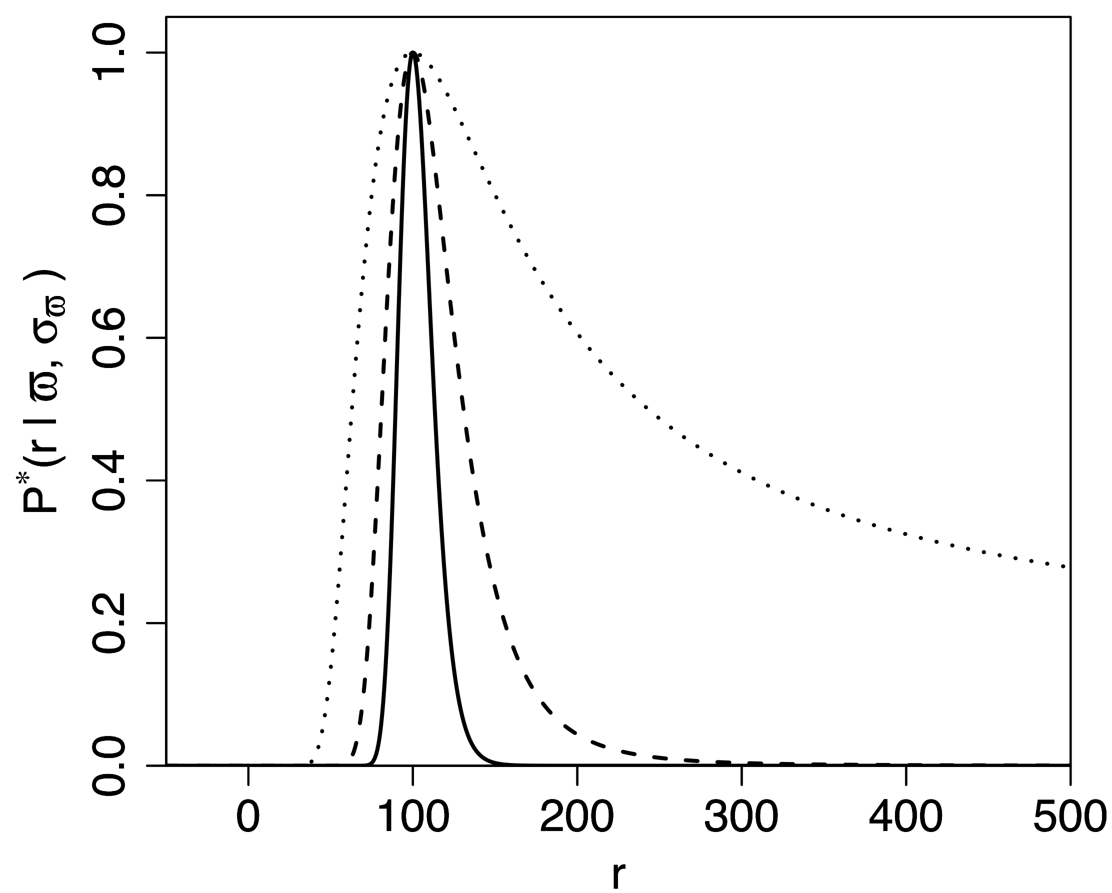


$f_{\mathrm{obs}} = 0.1$

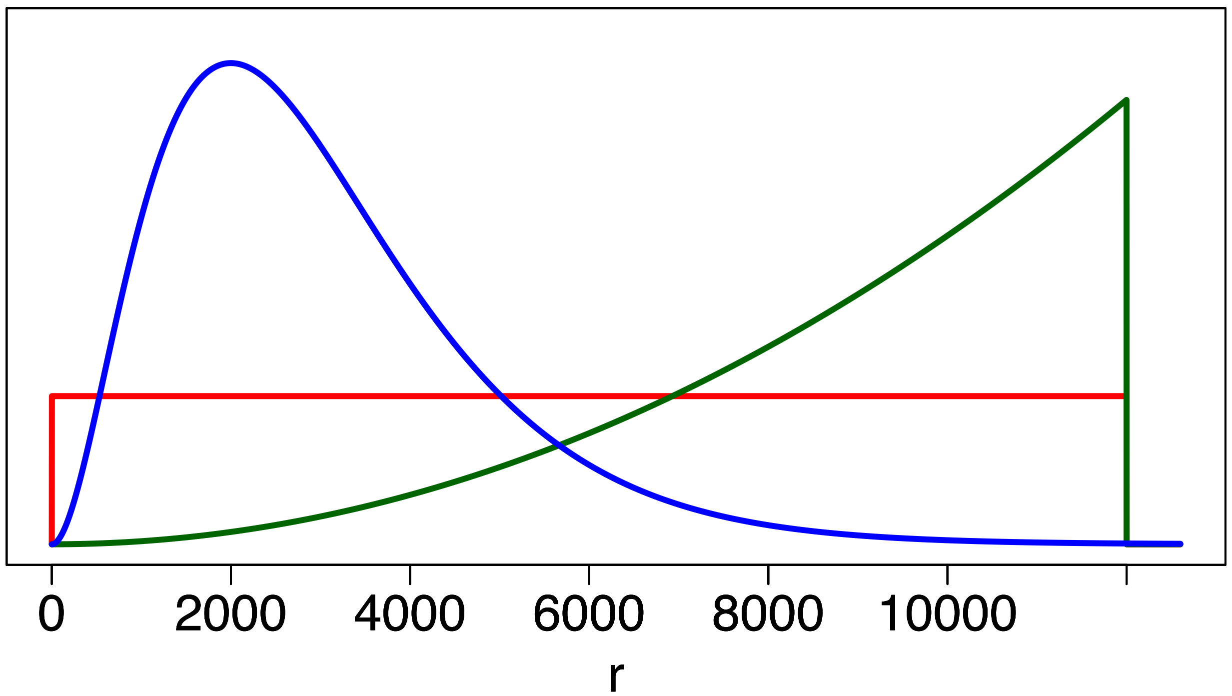
$f_{\mathrm{obs}} = 0.2$

$f_{\mathrm{obs}} = 0.5$

uniform prior $\varpi = 0.01$

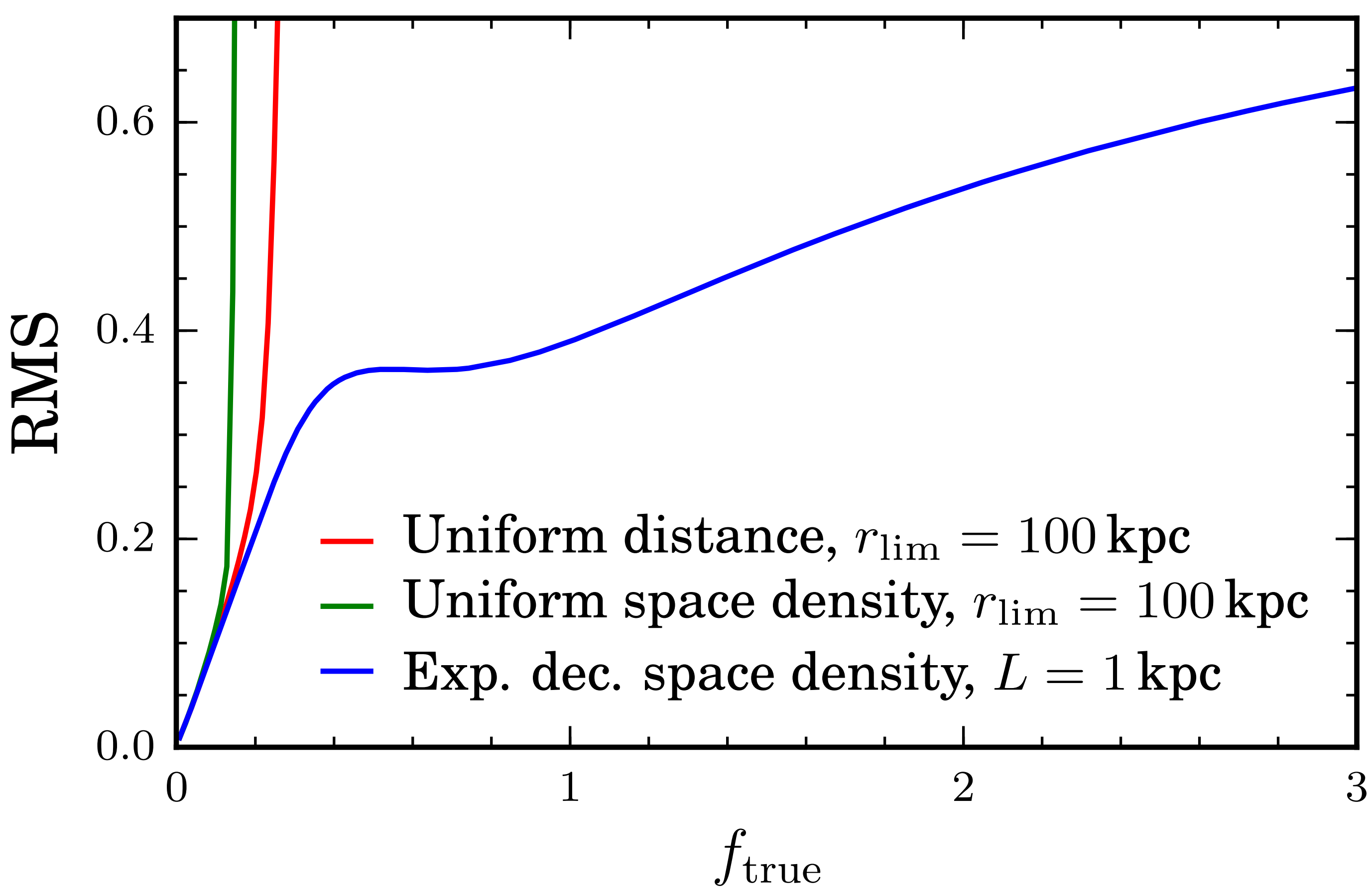


Possible distance priors



Prior PDF

uniform in r
uniform space density
exponentially decreasing space density

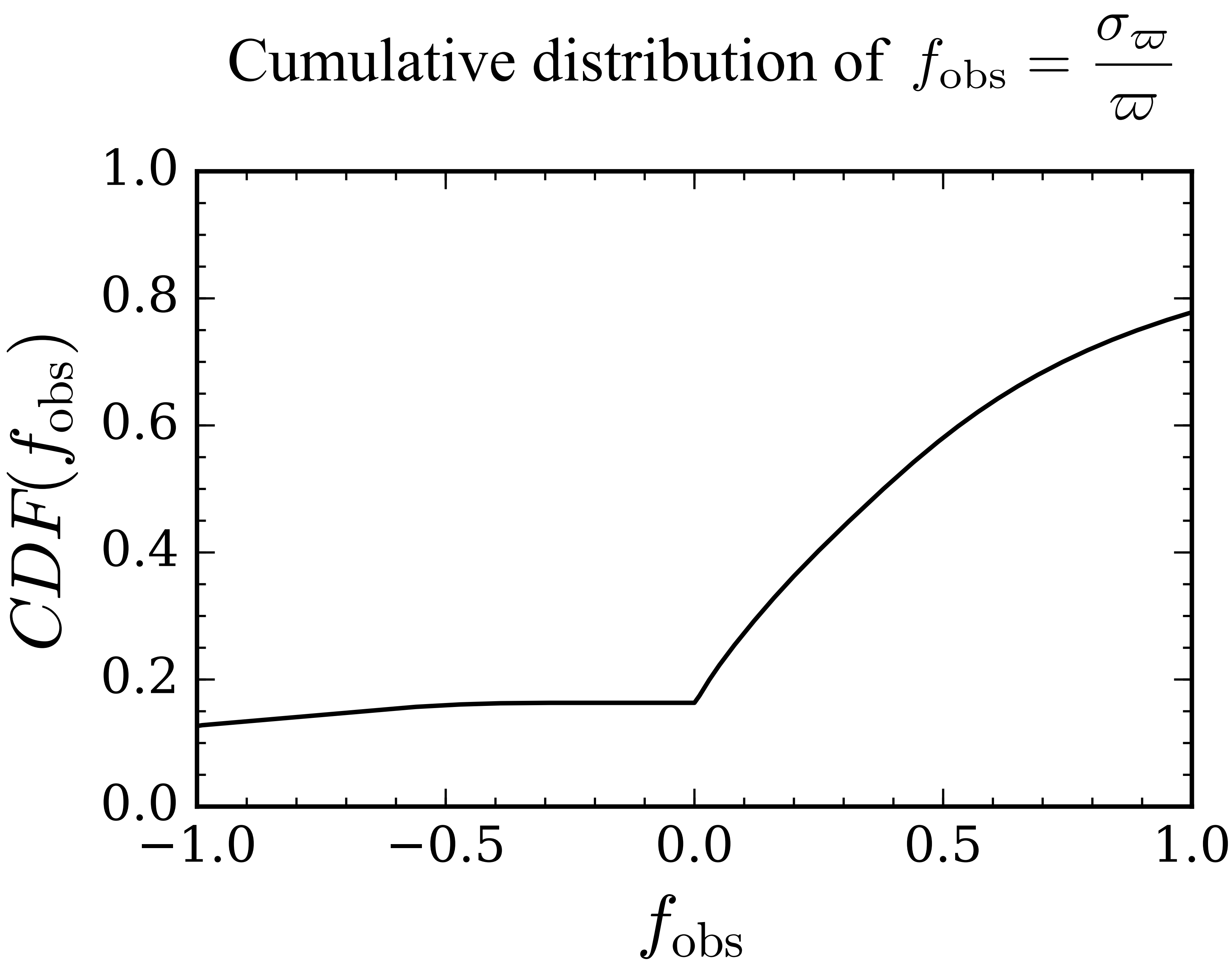


Astraatmadja & CBJ 2016 a

Test using simulations (GUMS)

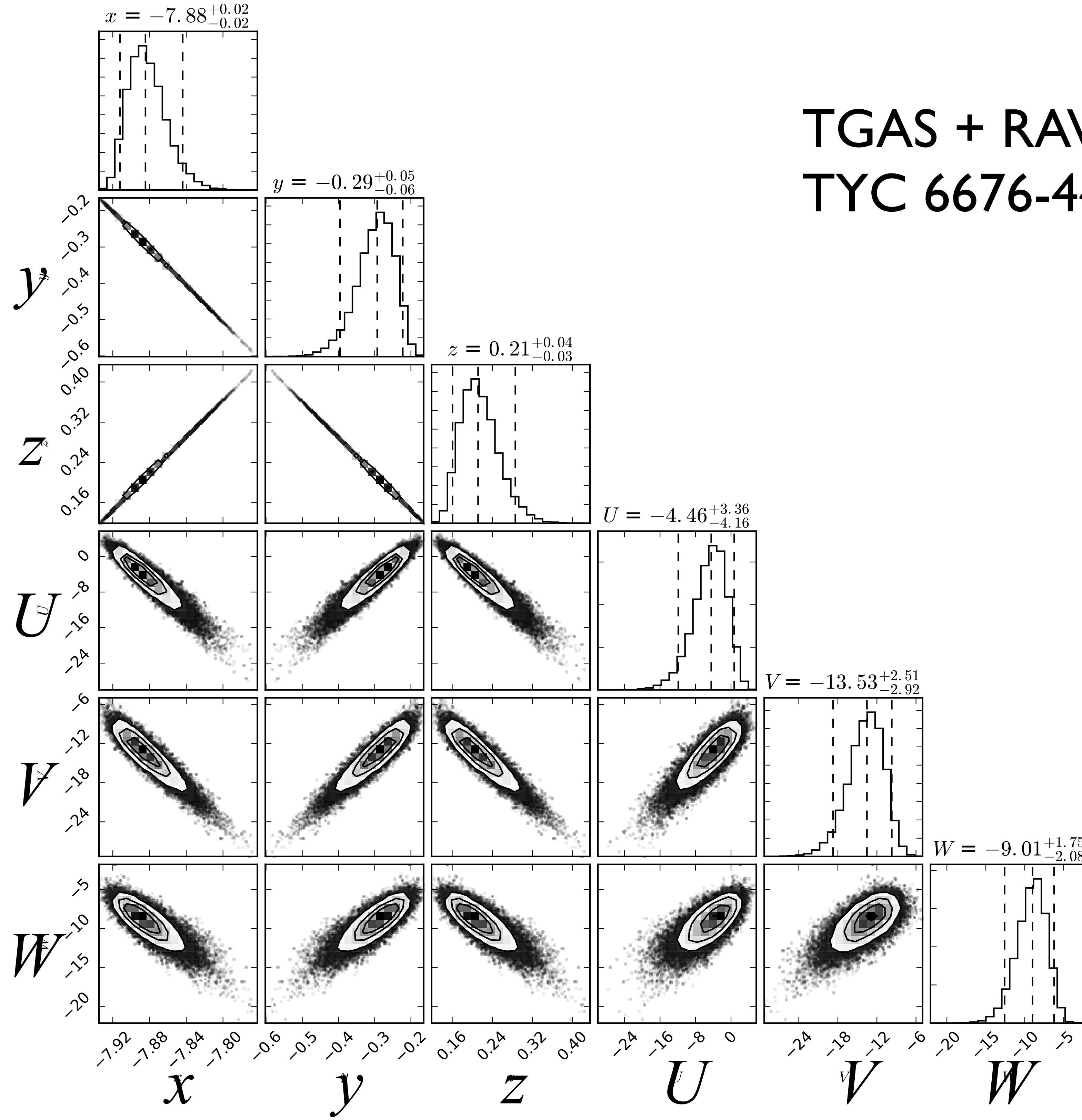
Fractional distance error vs. fractional parallax uncertainty

Many Gaia parallaxes will be poor



Astraatmadja & CBJ 2016 a

Inference of all six Galactic coordinates



TGAS + RAVE data
TYC 6676-44-1

Burger 2017 (BSc thesis, Heidelberg)

Coryn Bailer-Jones, MPIA

Distance information from spectra, colours, and magnitudes

Improved distances to stars common to TGAS and RAVE

Paul J. McMillan[1⋆], Georges Kordopatis[2], Andrea Kunder[3], James Binney[4], Jennifer Wojno[3], Tomaž Zwitter[5], Matthias Steinmetz[3], Joss Bland-Hawthorn[6], Brad K. Gibson[7], Gerard Gilmore[8], Eva K. Grebel[9], Amina Helmi[10], Ulisse Munari[11], Julio F. Navarro[12], Quentin A. Parker[13,14], George Seabroke[15], Rosemary F. G. Wyse[16]

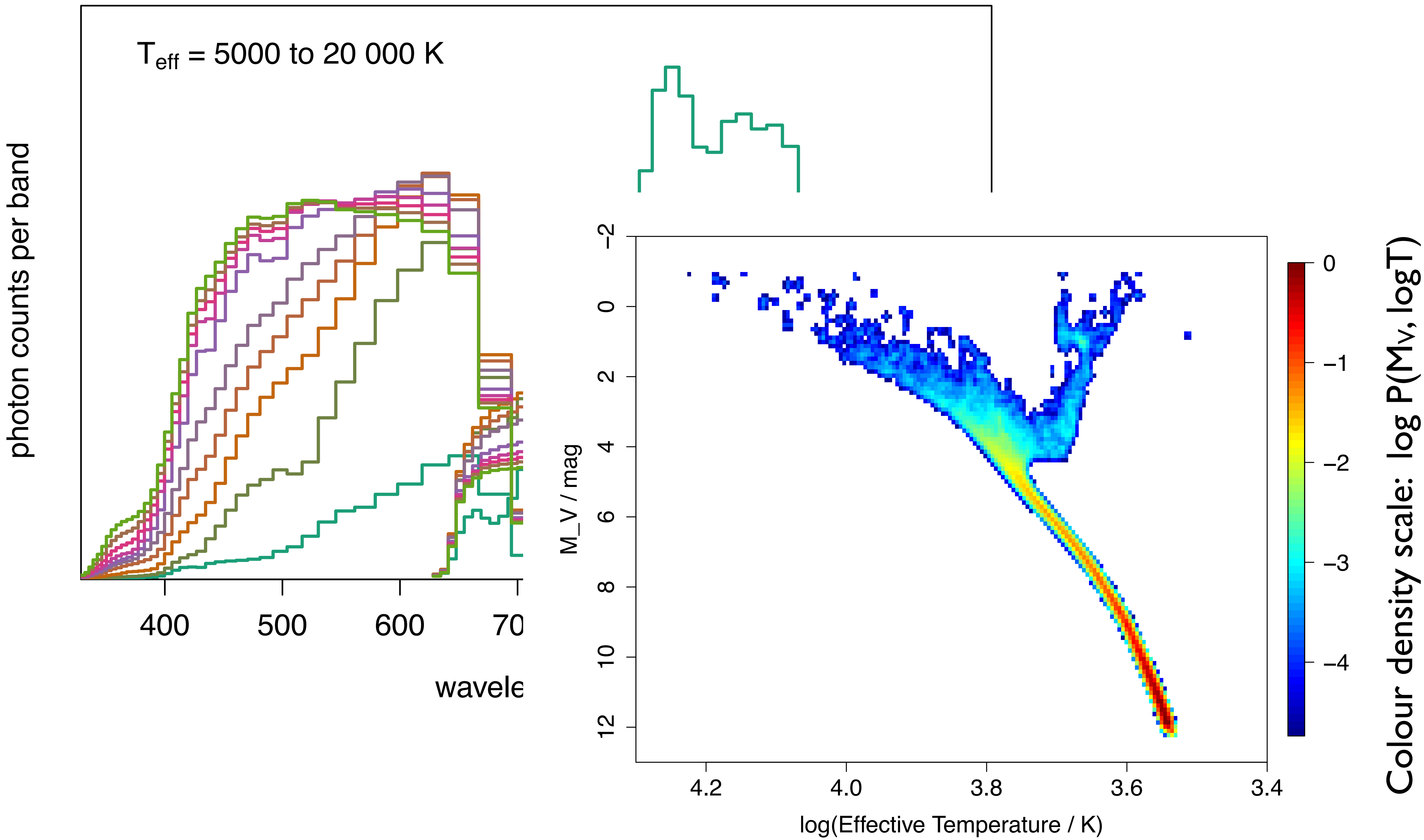
IMPROVING *GAIA* PARALLAX PRECISION
WITH A DATA-DRIVEN MODEL OF STARS

Lauren Anderson,[1] David W. Hogg,[1,2,3,4] Boris Leistedt,[2,5]
Adrian M. Price-Whelan,[6] and Jo Bovy[1,7,8]

HIERARCHICAL PROBABILISTIC INFERENCE
OF THE COLOR-MAGNITUDE DIAGRAM
AND SHRINKAGE OF STELLAR DISTANCE UNCERTAINTIES

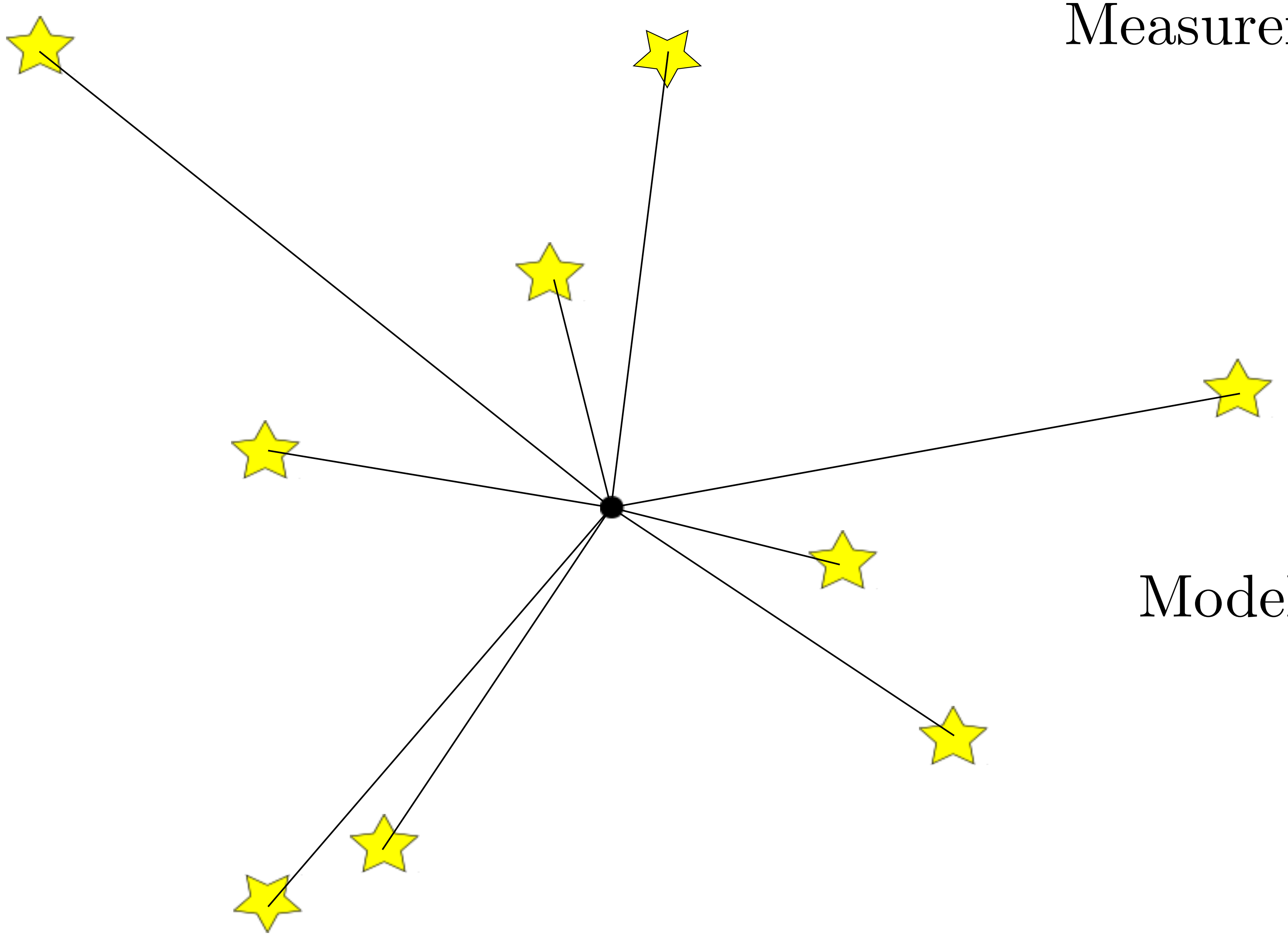
Boris Leistedt[1,2], David W. Hogg[1,3,4]

Teff, AV etc. from BP/RP, parallax with HRD prior



Coryn Bailer-Jones, MPIA

3D dust inference from stellar extinction



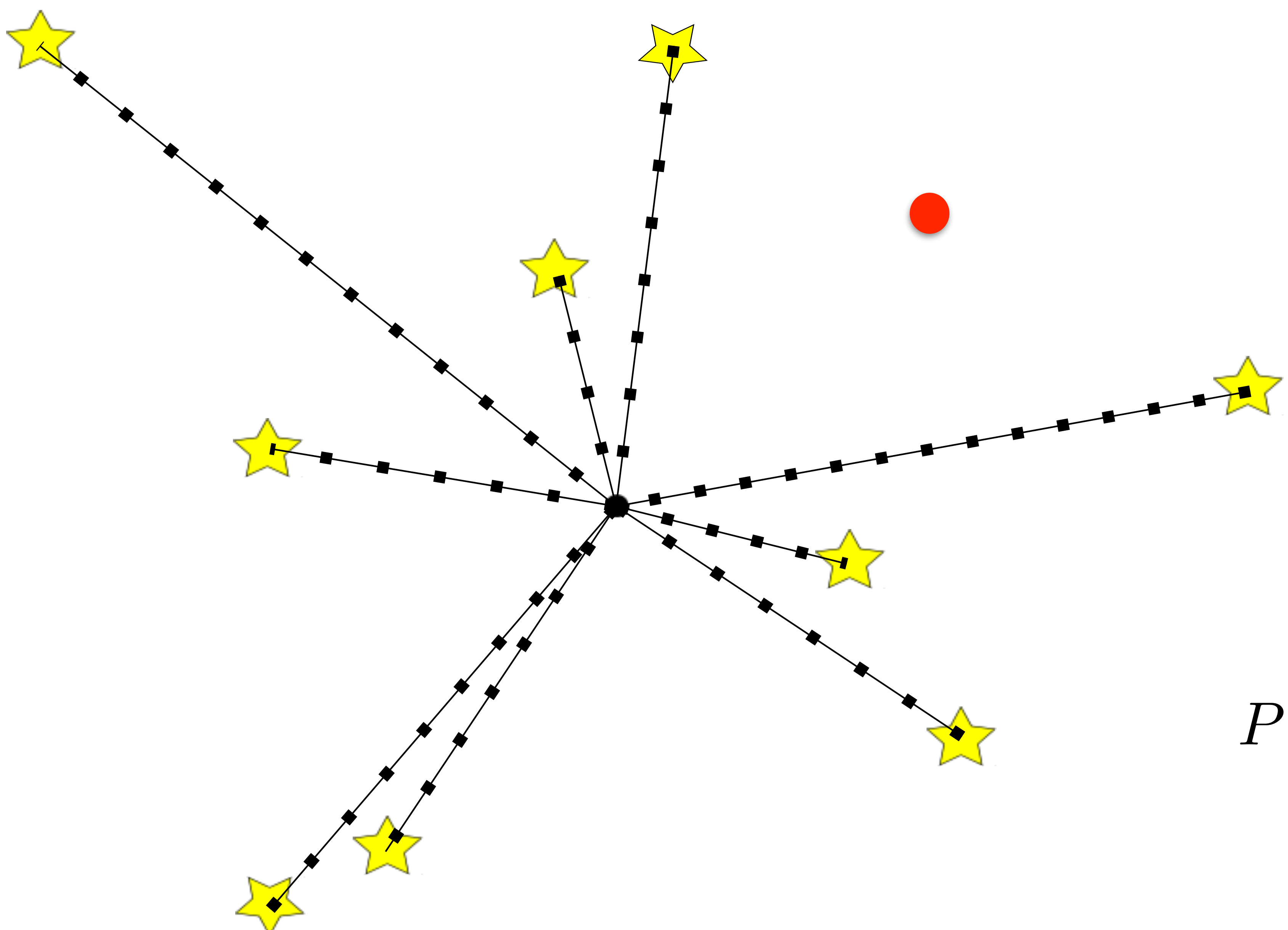
Measurements: $\{A_n(\mathbf{r}_n) \pm \sigma_n\}$

Model: $A_n = f_n + \mathcal{N}(0, \sigma_n^2)$

$$f_n \propto \int_{r=0}^{r=r_n} \rho(\mathbf{r}) dr$$

Coryn Bailer-Jones, MPIA

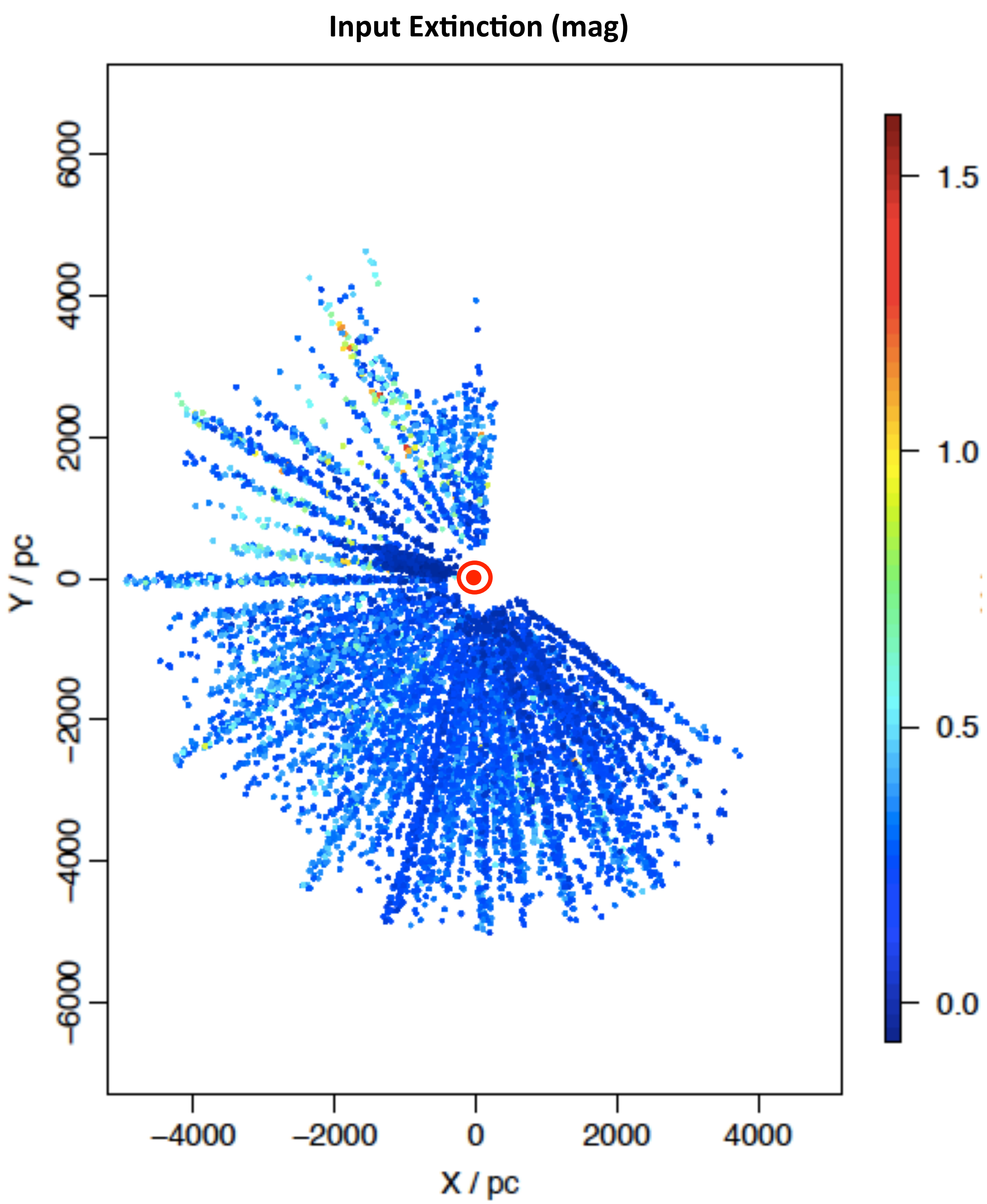
Smoothness constraint: Gaussian Process prior

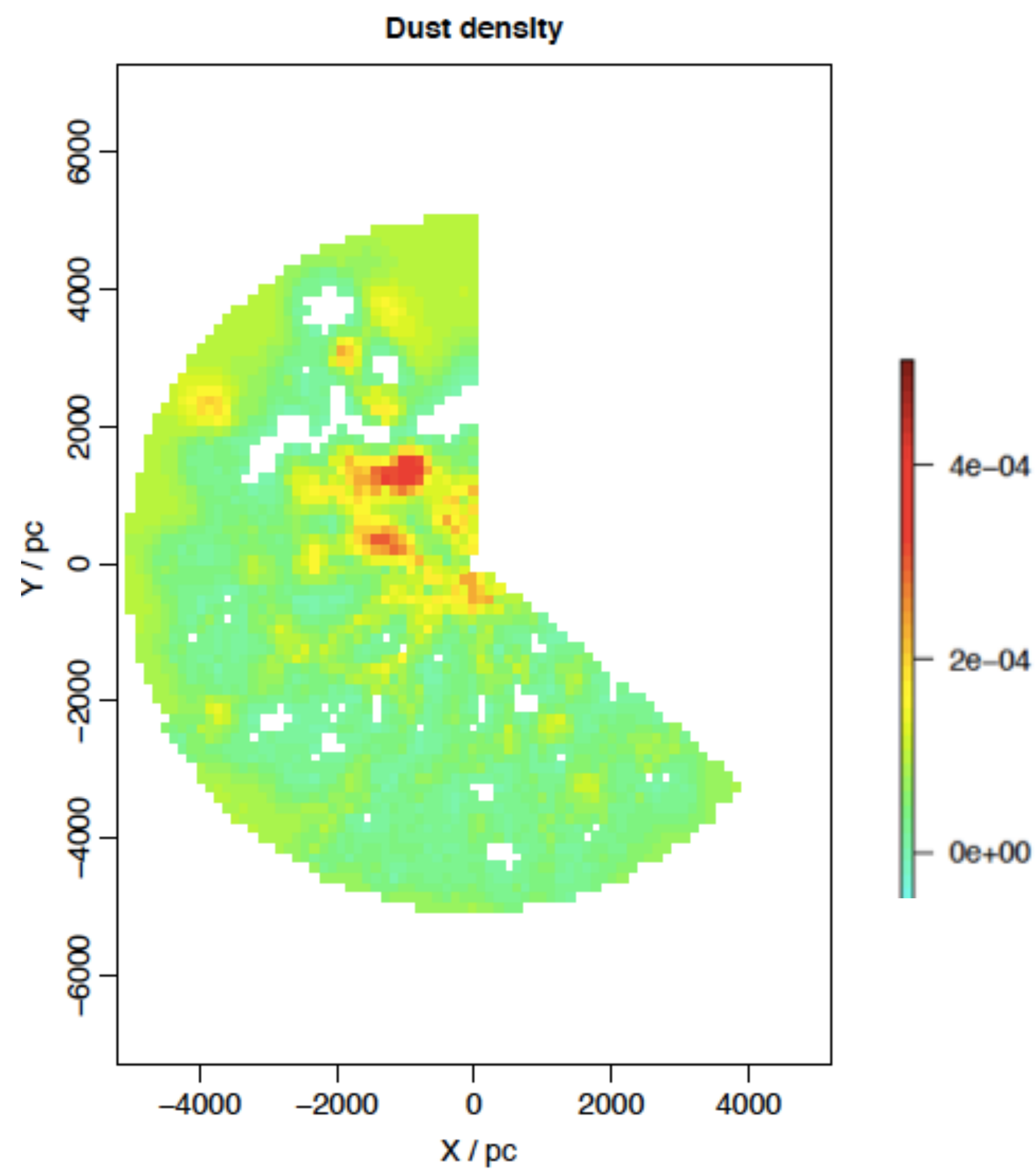


$$P(\rho_i, \rho_j) \sim \mathcal{N}(0, c_{i,j})$$

$$c_{i,j} = \theta \exp\left(-\frac{|\mathbf{r}_i - \mathbf{r}_j|^2}{\lambda^2}\right)$$

3D dust inference (APOGEE red clump stars)



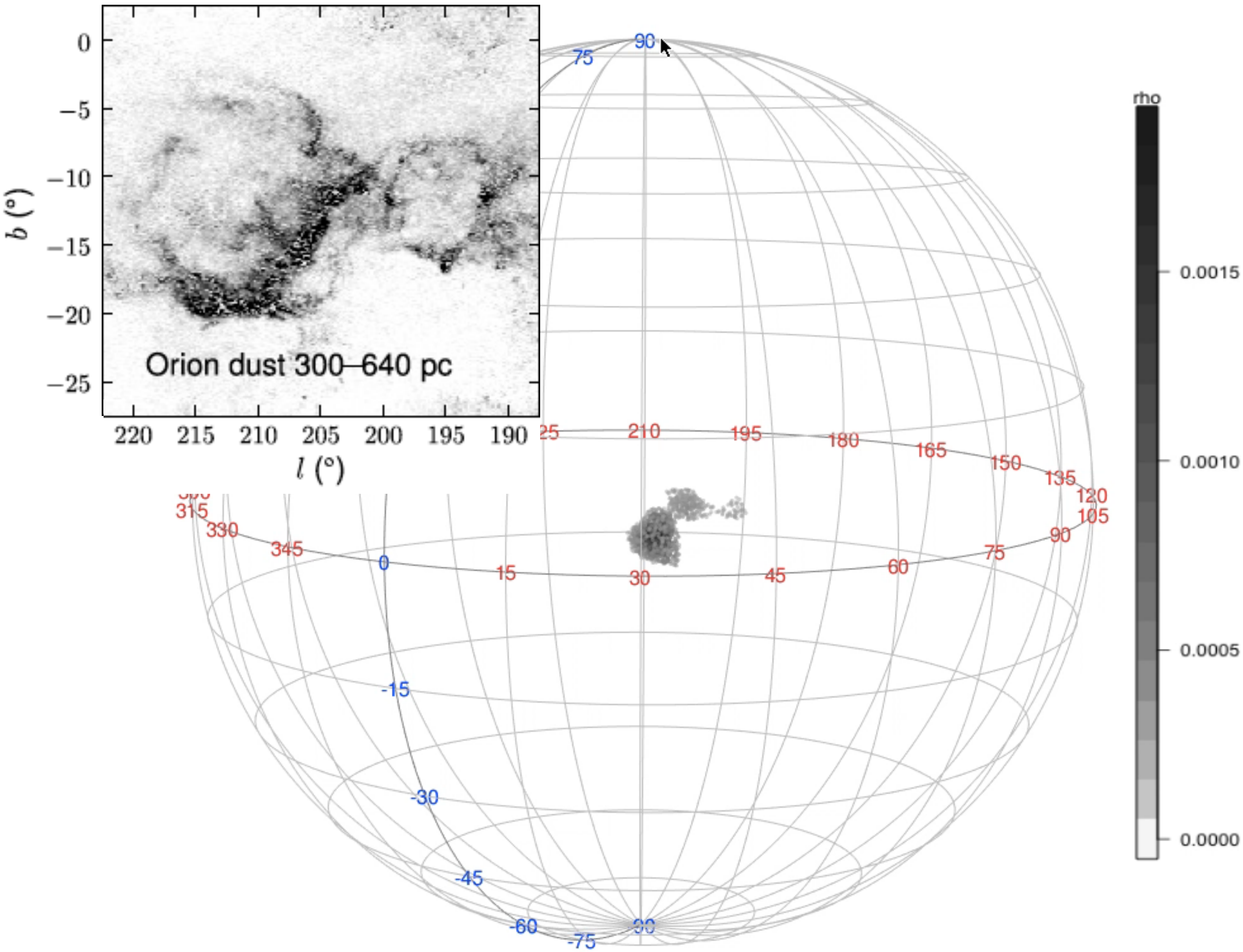


Rezaei Kh. et al. 2017

Coryn Bailer-Jones, MPIA

Dust clouds in/towards Orion from TGAS

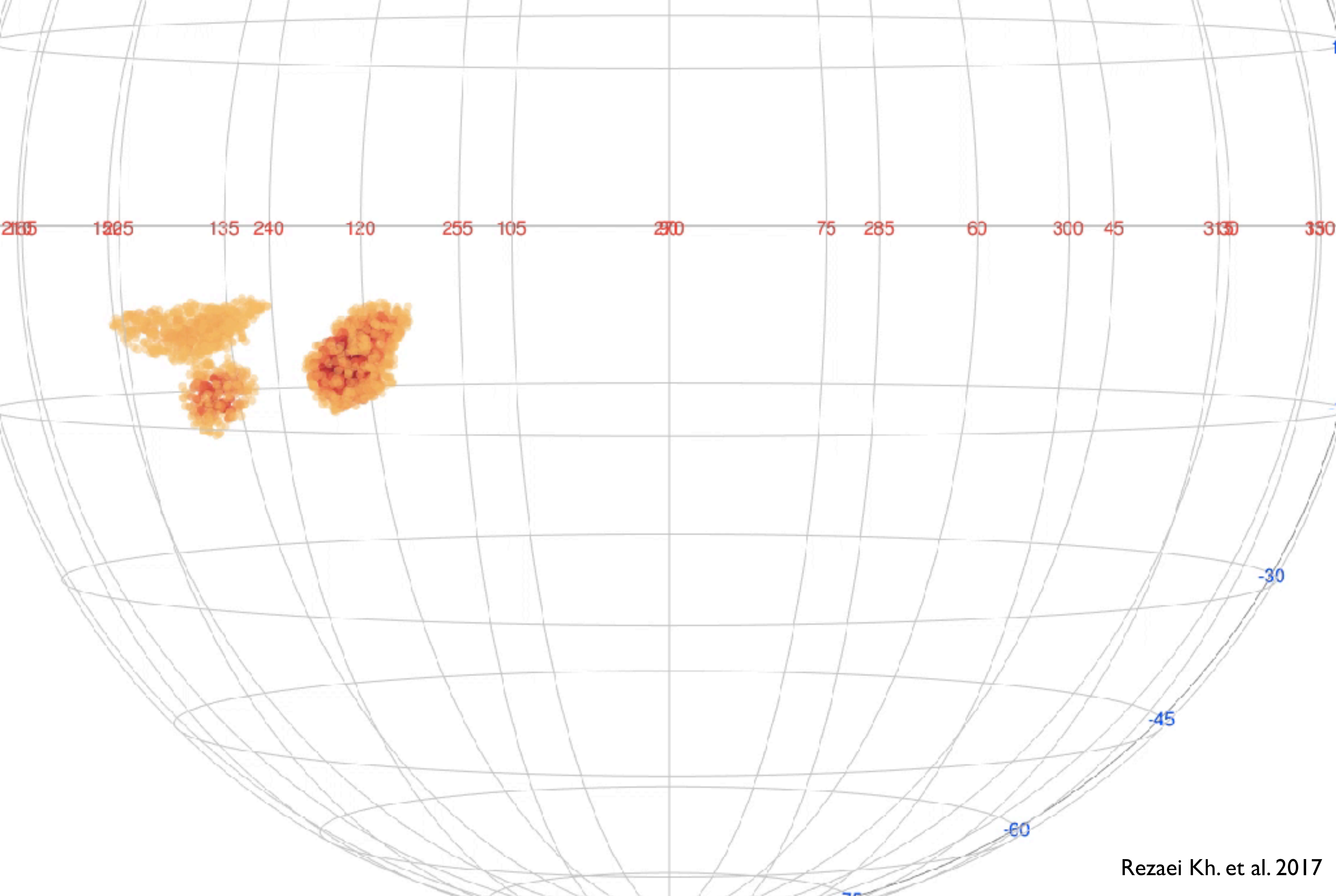
Schlafly et al. 2015



Dust density
mag/pc

Rezaei Kh. et al. 2017

Dust clouds in/towards Orion from TGAS



Rezaei Kh. et al. 2017

Conclusions

- "Bayes" is an approach to learning about models from data
  - remains consistent in limit of poor data
- Priors
  - incorporate other knowledge you have
  - flexible: theoretical, empirical, (non)-parametric, …
  - are just one choice of many we must make when analysing data
- Many applications, esp. data when noisy/incomplete
  - distances, kinematics, stellar parameters, dust mapping, TGAS, …