# Bayesian inference of stellar parameters and interstellar extinction using parallaxes and multiband photometry

Coryn A.L. Bailer-Jones @ MPIA, Heidelberg

AG, Bonn, 14 Sept. 2010

# Main points of this talk

AP = astrophysical parameter, e.g. $T_{eff}$, [Fe/H], $A_V$

- Good inference requires
  - a *forward model*, Data = f(APs)
  - a *probabilistic* representation of the results, P(APs | Data) ....
    - ... requiring a *probabilistic model* of uncertainties, P(Data | APs)
- Good inference requires us to
  - use all data/information available
  - achieve self-consistent solutions

# An opportunity and a problem

- *Opportunity:* parallaxes ($\varpi$) potentially allow accurate inference of absolute stellar magnitudes ...

$$\mathbf{q} \equiv \boxed{V + 5\log\varpi} = M_{\rm V} + A_{\rm V} - 5$$

... but only if $A_V$ can be estimated

- *Problem*: $A_V$ is degenerate with $T_{eff}$ when inferred from broad band photometry

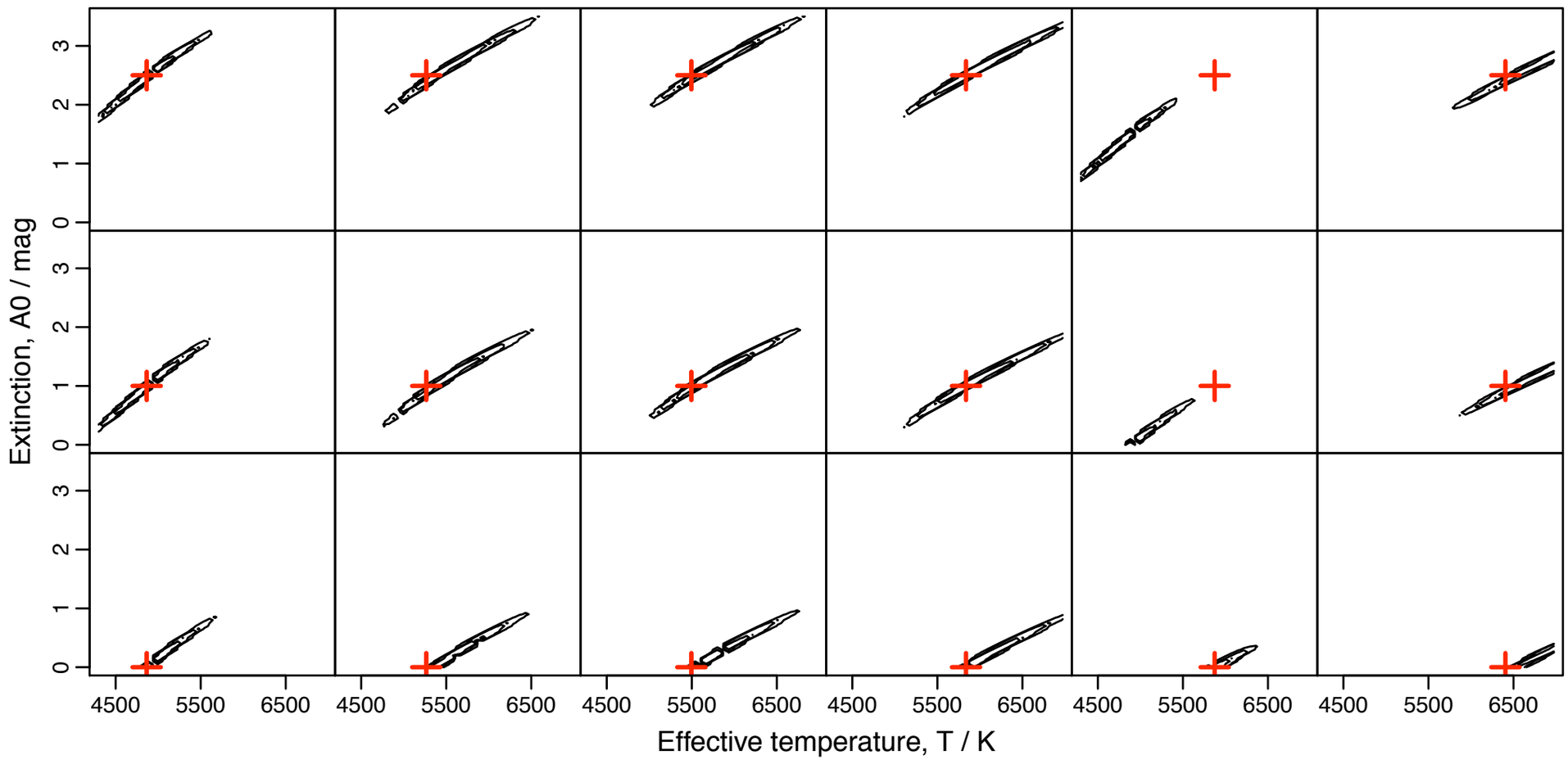
Notation:
**p** = ("photons") normalized SED, e.g. colours
**Φ** = astrophysical parameters (APs), here just $A_V$ and $T_{eff}$

# AP estimation from BVJHK colours (only)

$$P(\boldsymbol{p}|\boldsymbol{\phi}) \propto e^{-D^2/2} = \exp\left(-\frac{1}{2}[\boldsymbol{p} - \boldsymbol{f}(\boldsymbol{\phi})]^T \mathbf{C}_p^{-1} [\boldsymbol{p} - \boldsymbol{f}(\boldsymbol{\phi})]\right)$$

forward model

data covariance




$P(\boldsymbol{\phi}|\boldsymbol{p})$

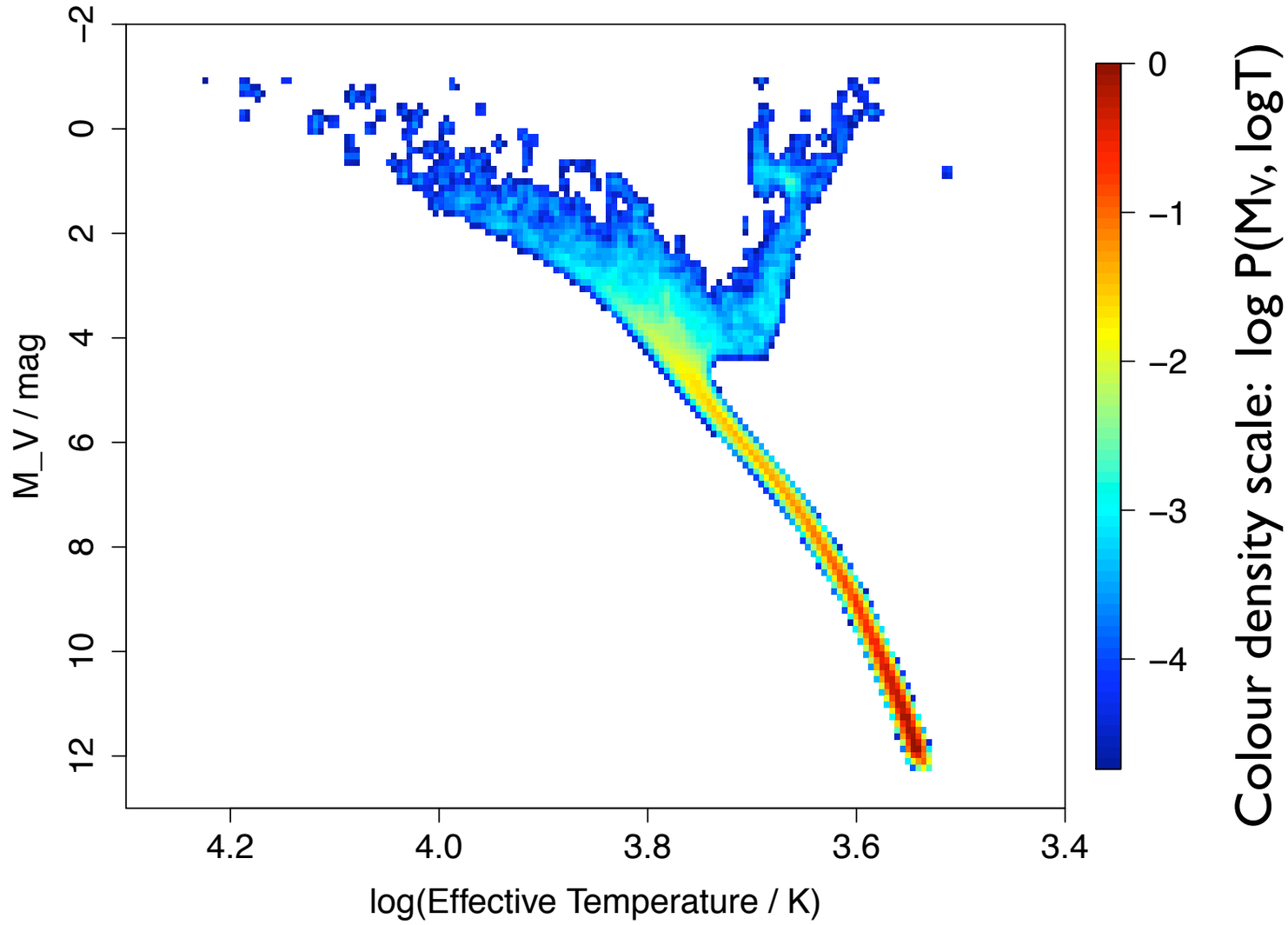
- real data; "true" APs shown as red cross
- "true" $T_{eff}$ from Valenti & Fischer (2005) from high-res. spectroscopy
- artificially reddened colours

Coryn Bailer-Jones, MPIA Heidelberg

# Information beyond the spectrum

1) parallax, apparent magnitude (q)

2) HRD




- Use of complementary information
  - **p** constrains $T_{eff}$ and $A_V$
  - q constrains $M_V$ + $A_V$ $\quad q \equiv \boxed{V + 5\log\varpi} = M_V + A_V - 5$
  - HRD prior constrains $M_V$ and $T_{eff}$

# Probabilistic inference

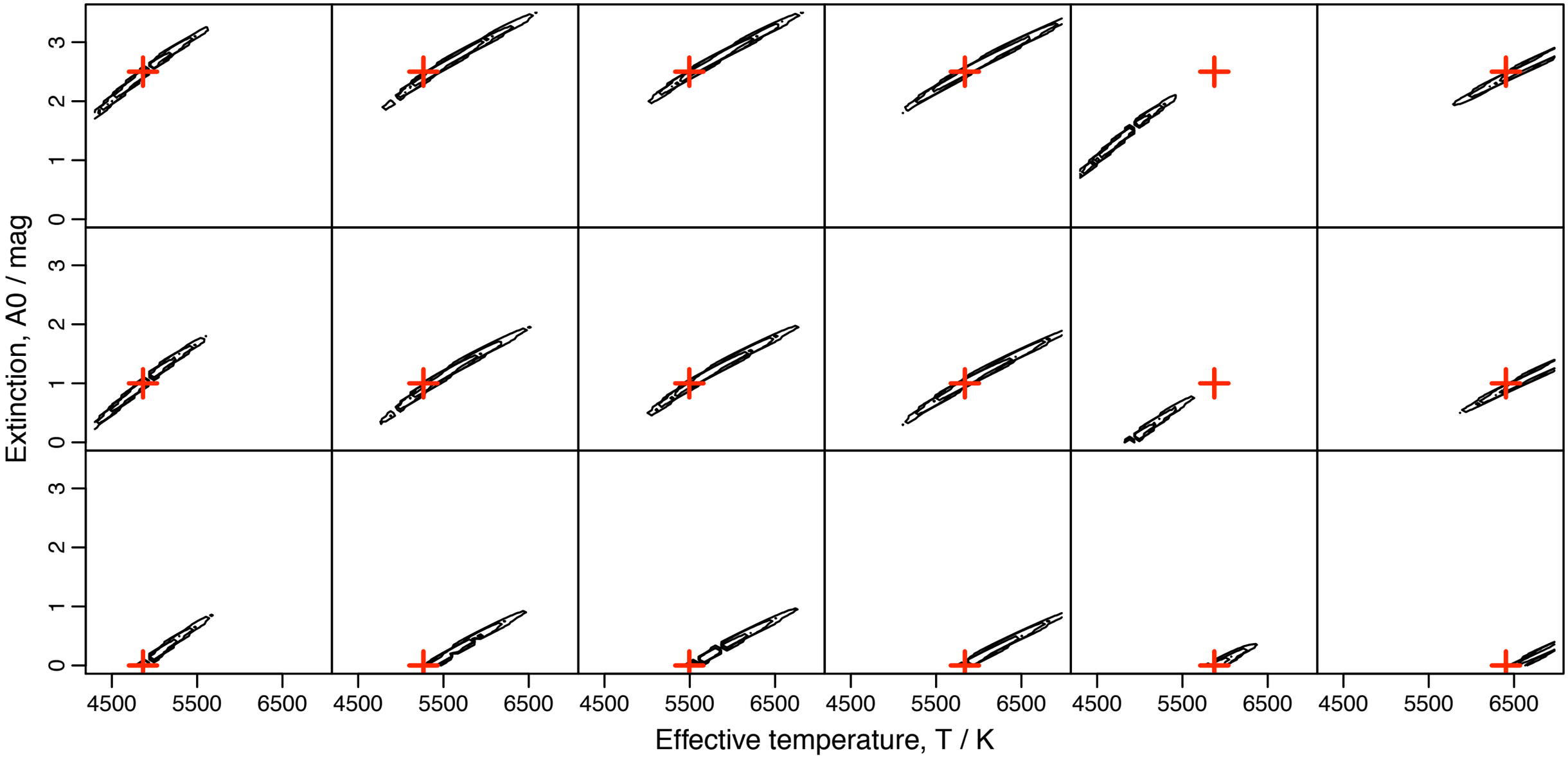
$$P(A_0, T|\boldsymbol{p}, q) = \frac{P(\boldsymbol{p}, q|A_0, T)P(A_0, T)}{P(\boldsymbol{p}, q)}$$

Bayes' theorem

$$= \underbrace{P(\boldsymbol{p}|A_0, T)}_{\text{likelihood}} \underbrace{\frac{P(A_0)}{P(\boldsymbol{p}, q)}}_{\text{priors}} \underbrace{\int_{M_\mathrm{V}} \underbrace{P(q|M_\mathrm{V}, A_0, T)}_{q\text{ constraint}} \underbrace{P(M_\mathrm{V}, T)}_{\text{HRD prior}} dM_\mathrm{V}}_{\text{HRD/q factor}}$$

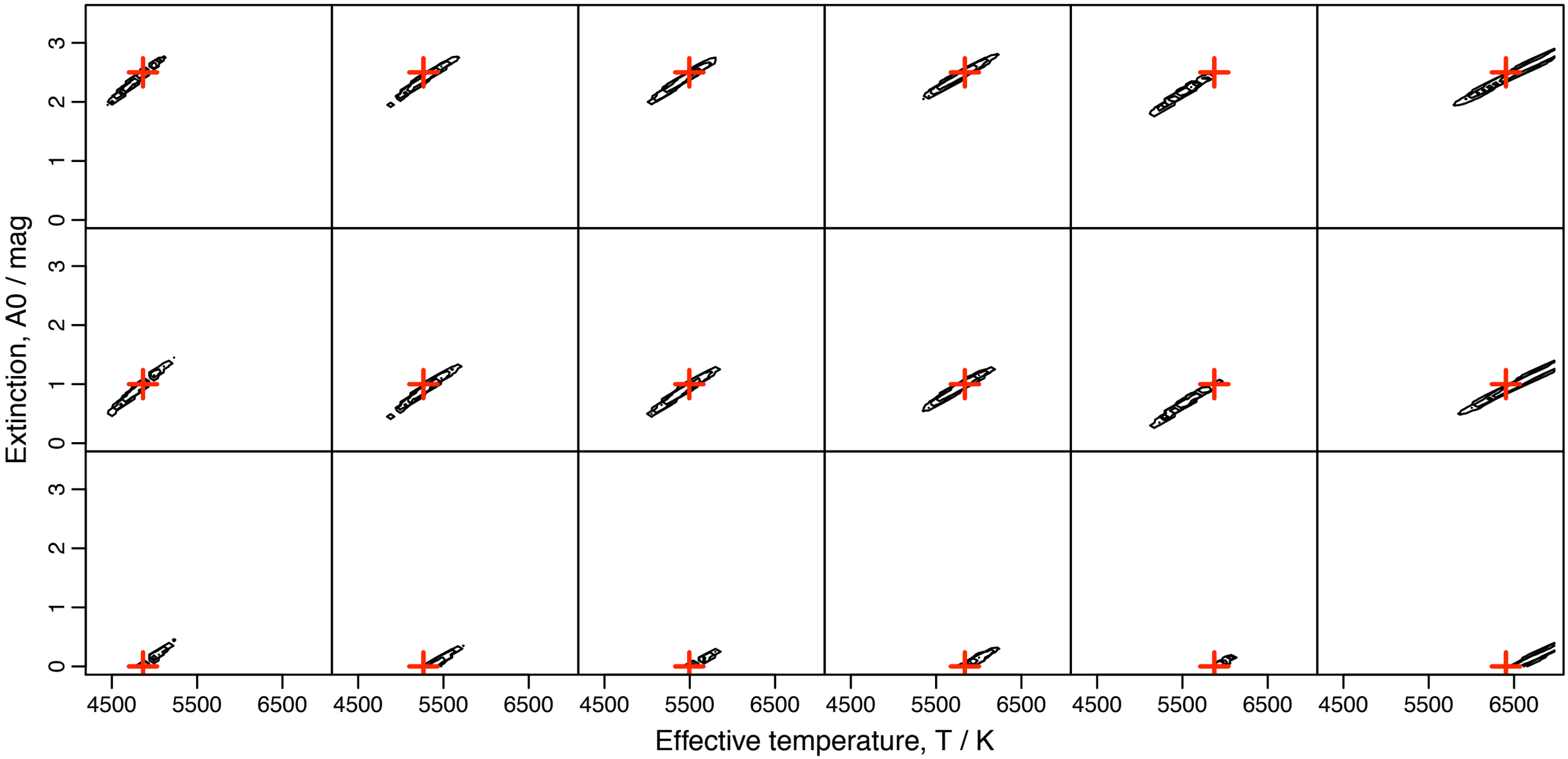
colour-only estimation

# AP estimation from BVJHK colours




Coryn Bailer-Jones, MPIA Heidelberg

# AP estimation from BVJHK colours + q, HRD




Coryn Bailer-Jones, MPIA Heidelberg

# Numerical improvement by use of q and HRD

5280 stars with $T_{eff}$ = 4700–6600 K, $A_V$ = 0–2.5 mag

Precision: 90% confidence interval

- $\log T_{eff}$ improvement: 0.055 dex to 0.038 dex
- $A_V$ improvement: 0.63 mag to 0.43 mag

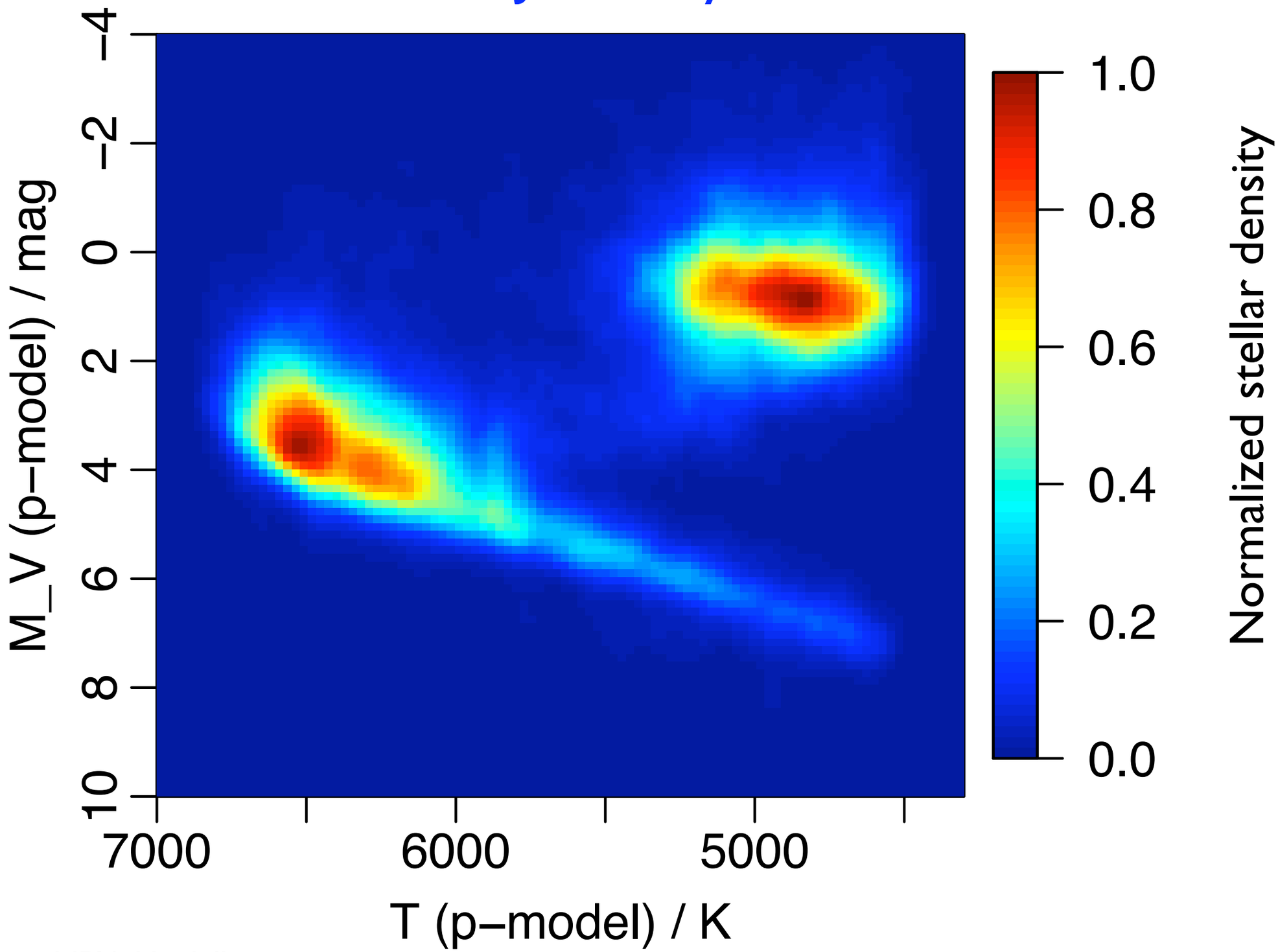
Accuracy: mean absolute residual (estimated minus true)

- $\log T_{eff}$ improvement: 0.025 dex to 0.015 dex
- $A_V$ improvement: 0.29 mag to 0.19 mag

Coryn Bailer-Jones, MPIA Heidelberg

# Application to 85 000 Hipparcos/2MASS stars

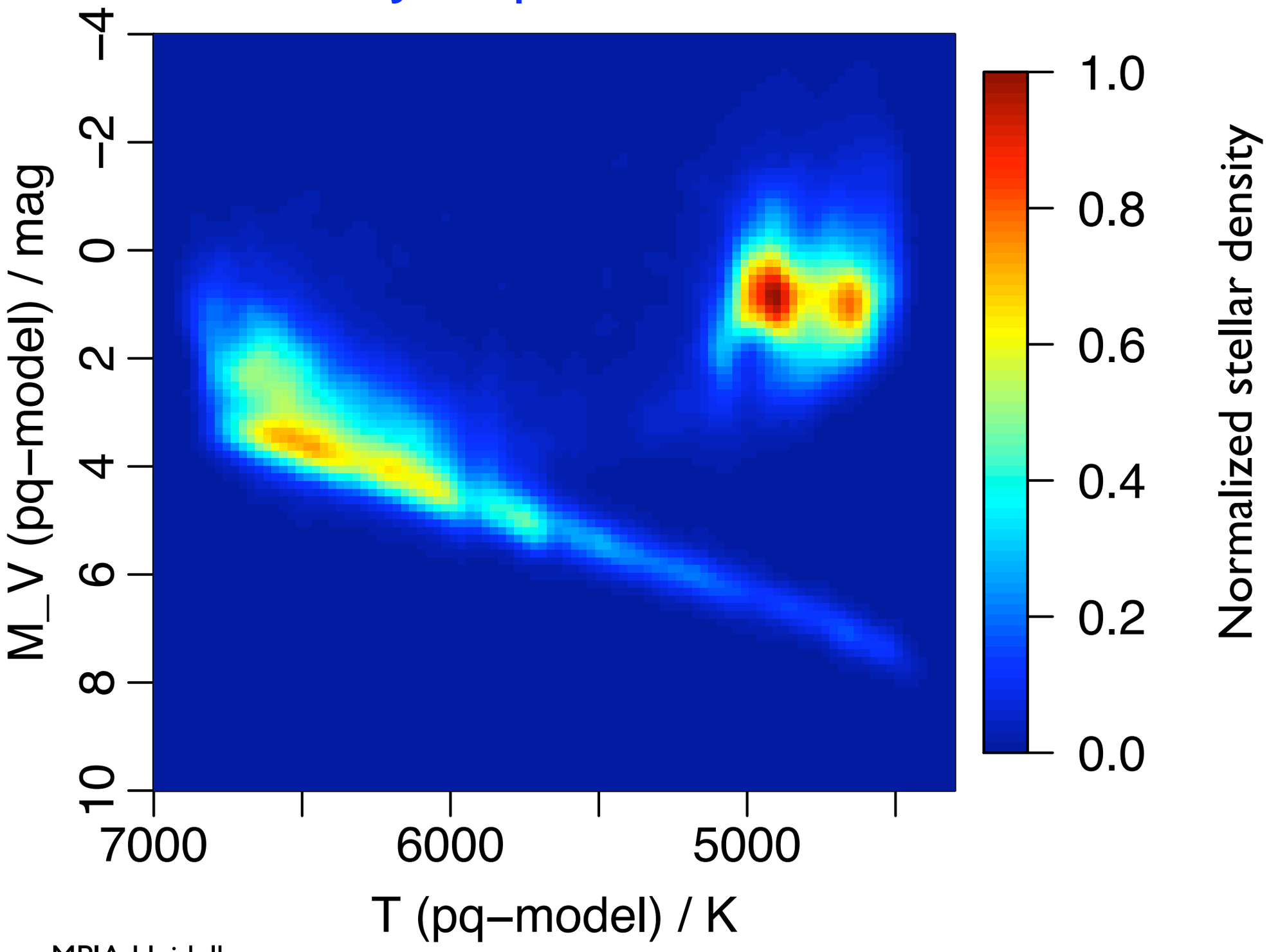
BVJHK only




Coryn Bailer-Jones, MPIA Heidelberg

# Application to 85 000 Hipparcos/2MASS stars

BVJHK, parallax, HRD




Coryn Bailer-Jones, MPIA Heidelberg

# Drawbacks of conventional approaches

Why modelling of y(AP | Data) (e.g. ANN, SVM) is a bad idea

- it's an inverse problem
  - may not be unique
  - cumbersome to fit and inflexible with heterogeneous inputs
- methods are invariably non-probabilistic
  - cannot recognize degeneracies, multiple solutions or give errors bars in any sensible way
- cannot naturally/explicitly incorporate domain knowledge or prior information ⇒ maybe inconsistent, unphysical solutions

# Summary and Conclusions

- Good inference needs a *forward* model and a *probabilistic* model of the uncertainties
- You *do* have information beyond the SED, so use it!
- Use of parallax, HRD improves $T_{eff}$, $A_V$ accuracy by 35% on Hipparcos/2MASS data (BVJHK)
- Method gives full PDF over solutions
- AP estimation by conventional (inverse) approach (e.g. with ANN, SVM etc.) has significant disadvantages
- Plan to use in Gaia data processing
- `http://tinyurl.com/qmethod` (MNRAS, accepted)