# Optimal scheduling of radial velocity follow-up of transiting exoplanets

**Pedro Viana**

Instituto de Astrofísica e Ciências do Espaço (IA)
Faculdade de Ciências da Universidade do Porto

in collaboration with

**Lorenzo Cabona** (INAF), **Marco Landoni** (INAF) and **João Faria** (IA)

Extreme Precision in Radial Velocity
Grindelwald, 18th – 21st March 2019

# Bayesian Decision Theory

The best experiment is the one which maximizes its expected utility

$$EU(\boldsymbol{e}) = \int U(\boldsymbol{D_e}, \boldsymbol{e}) P(\boldsymbol{D_e} | \boldsymbol{D_c}, \boldsymbol{e}) d\boldsymbol{D_e}$$

$$= \int U(\boldsymbol{D_e}, \boldsymbol{e}) \int P(\boldsymbol{D_e} | \boldsymbol{\theta}, \boldsymbol{M}, \boldsymbol{e}) P(\boldsymbol{\theta} | \boldsymbol{M}, \boldsymbol{D_c}) d\boldsymbol{\theta} d\boldsymbol{D_e}$$

where

$e$ **experiment parameters (e.g. spatial or time coordinates)**

$D_e$ ***a priori* possible data values acquired through experiment $e$**

$D_c$ **current data values acquired through previous experiments**

$\theta$ **parameters of the data-generating model ($M$, assumed true *a priori*)**

# Bayesian Decision Theory

**Most useful from a Bayesian point of view are the data $D_e$, obtained under experiment $e$, which maximize the distance between posterior and prior distributions**

$$\mathrm{U}(\boldsymbol{D_e}, \boldsymbol{e}) = \int P(\boldsymbol{\theta}|\boldsymbol{M}, \boldsymbol{D_c}, \boldsymbol{D_e}, \boldsymbol{e}) \ln\left[\frac{P(\boldsymbol{\theta}|\boldsymbol{M}, \boldsymbol{D_c}, \boldsymbol{D_e}, \boldsymbol{e})}{P(\boldsymbol{\theta}|\boldsymbol{M}, \boldsymbol{D_c})}\right] d\boldsymbol{D_e}$$

**Kullback-Leibler divergence**

And if $P(\boldsymbol{D_e}|\boldsymbol{\theta}, \boldsymbol{M}, \boldsymbol{e})$ is independent of $e$ then

$$EU(\boldsymbol{e}) = \text{constant} - \int P(\boldsymbol{D_e}|\boldsymbol{M}, \boldsymbol{D_c}, \boldsymbol{e}) \ln[P(\boldsymbol{D_e}|\boldsymbol{M}, \boldsymbol{D_c}, \boldsymbol{e})] d\boldsymbol{D_e}$$

$$= \text{constant} + E[I(\boldsymbol{D_e}, \boldsymbol{e})]$$

**Shannon entropy**

→ **maximum entropy sampling**

**Ford (2008)**
**Loredo et al. (2004, 2012)**

where $I(\boldsymbol{D_e}, \boldsymbol{e})$ is the information associated with a given set of measurements $D_e$ in nats.

# Problem

**When to schedule radial velocity measurements in order to maximize the information with respect to the masses of exoplanets detected in transit?**

# Generic Model

For each radial velocity measurement $d_i$

$$d_i = \sum_{j=1}^{n_p} K_j \{ e_j \cos \omega_j + \cos[\omega_j + \phi(t_i | P_j, \tau_j, e_j)] \} + v_s(t_i) + \varepsilon_i$$

where $n_p$ is the number of planets, $v_s$ is the radial velocity of the star, and $\varepsilon_i$ represents stochastic instrumental, atmospheric and astrophysical (e.g. stellar activity) model components.

We will assume each $\varepsilon_i$ is **iid Gaussian distributed**

$$P(\{\varepsilon_i\}) \sim \prod_{i=1}^{N} N(0, \sigma_i)$$

# Schedulling Model

For each radial velocity measurement $d_i$

$$d_i = \sum_{l=1}^{n_t} K_l \cos\left[\omega_l + \frac{2\pi(t_i - \tau_l)}{P_l}\right] + v_s + \varepsilon_i$$ → **linear model**

where now $n_t$ is the number of transiting planets, all assumed to have circular orbits, and $v_s$ is the assumed constant radial velocity of the star. The values of $\omega_l$ and $\tau_l$ are completely degenerate, and such that at the time of transit $d_i = 0$.

**For a linear model, and in the presence of iid Gaussian uncertainty, if all $M$ parameter ($K_l$, $v_s$) prior distributions are uniform then**

$$EU(e) = -\frac{1}{2}[M + M\ln(2\pi) - \ln(|G^T E^{-1} G|)] + \text{constant}$$

$G$ – matrix containing the values the basis functions take at $t_i$

$E$ – data variance-covariance matrix (diagonal, with terms $\sigma_i^2$)

# Transiting planets with circular orbits

**If no extra planets were assumed to exist in the system**, then analysis and schedulling models would coincide, and the marginal posterior distributions for any $K_l$ and $v_S$ would be Gaussians. In the particular case of $n_t = 1$,

$$K = \mu_K \pm \sigma_K$$

$$\sigma_K^2 = \left[ \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \sum_{i=1}^{N} \frac{\left(\cos\left[\omega + \frac{2\pi(t_i - \tau)}{P}\right]\right)^2}{\sigma_i^2} - \left( \sum_{i=1}^{N} \frac{\cos\left[\omega + \frac{2\pi(t_i - \tau)}{P}\right]}{\sigma_i^2} \right)^2 \right]^{-1} \sum_{i=1}^{N} \frac{1}{\sigma_i^2}$$

**If the $\{\sigma_i\}$ were all equal, when would $\sigma_K^2$ be minimized?**

**When RV measurements are performed in quadrature!**

$$\sigma_K = \frac{\sigma}{\sqrt{N}}$$ c.f. $$\sigma_K = \frac{\sigma\sqrt{2}}{\sqrt{N}}$$ for uniform sampling in phase (e.g. Cloutier et al. 2018)

# Transiting planets with non-circular orbits

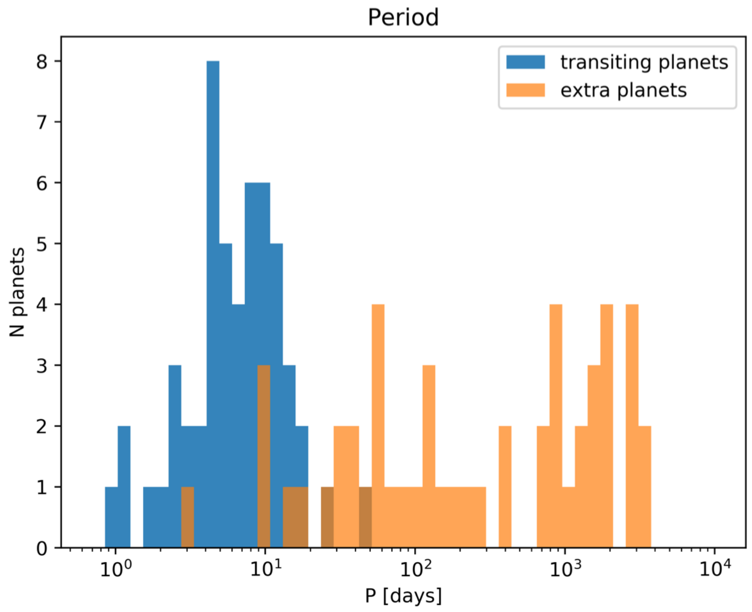
**A Revised Exoplanet Yield from the Transiting Exoplanet Survey Satellite (TESS)**
**Barclay, Pepper and Quintana, 2018, ApJSS, 239**

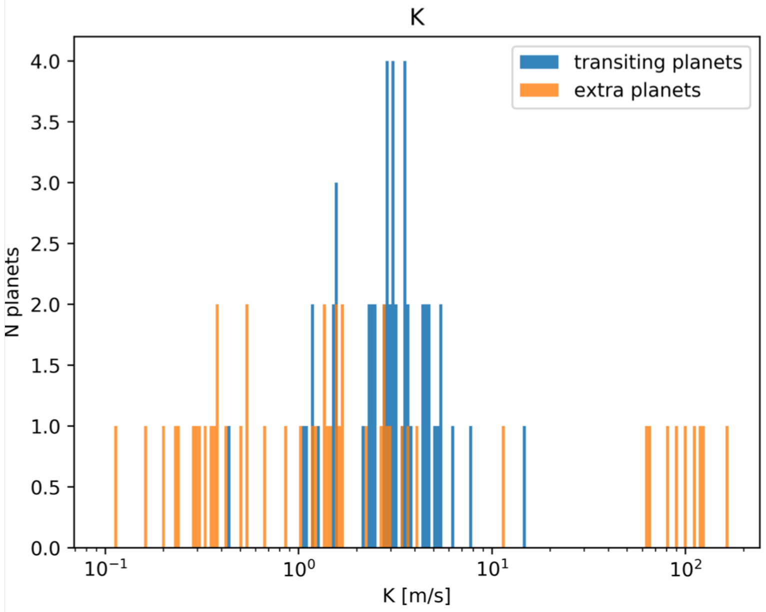
**Among which, 50 stars were selected for ESPRESSO follow-up with:**

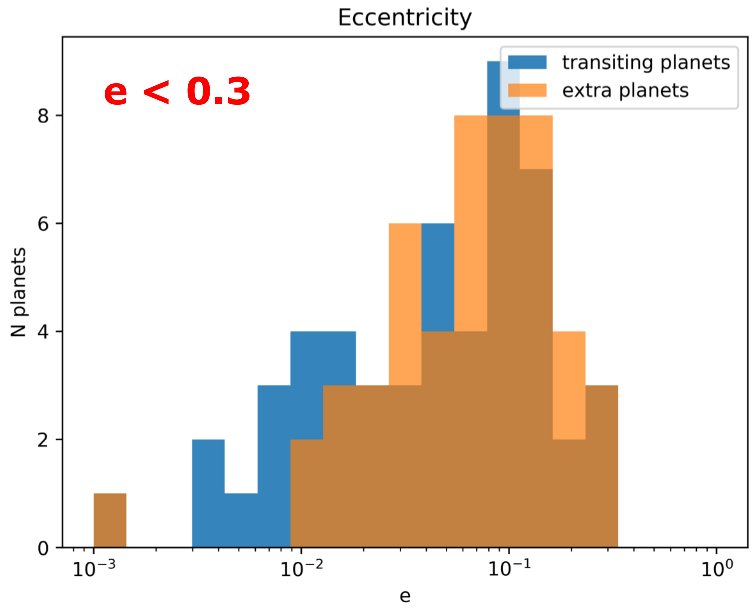
**$-80 < \delta < +30$, $V < 11$, $T_{eff} < 6000$ K, $\log g > 4.0$**

**and at least one orbiting planet with 3 TESS detected transits, S/N<10 and a radius below $4\ R_{\oplus}$**

Extra planets, non-detectable by TESS and with K > 0.1 m/s, were randomly added to each system according to Poisson distributions with means as in Fressin et al. (2013), extrapolated up to 2 years, and with means as in Herman, Hu and Wu (2019), for periods between 2 and 10 years. All planets in any system were assumed co-planar, with eccentricities drawn from the Beta distribution assumed in Barclay et al. (2018).



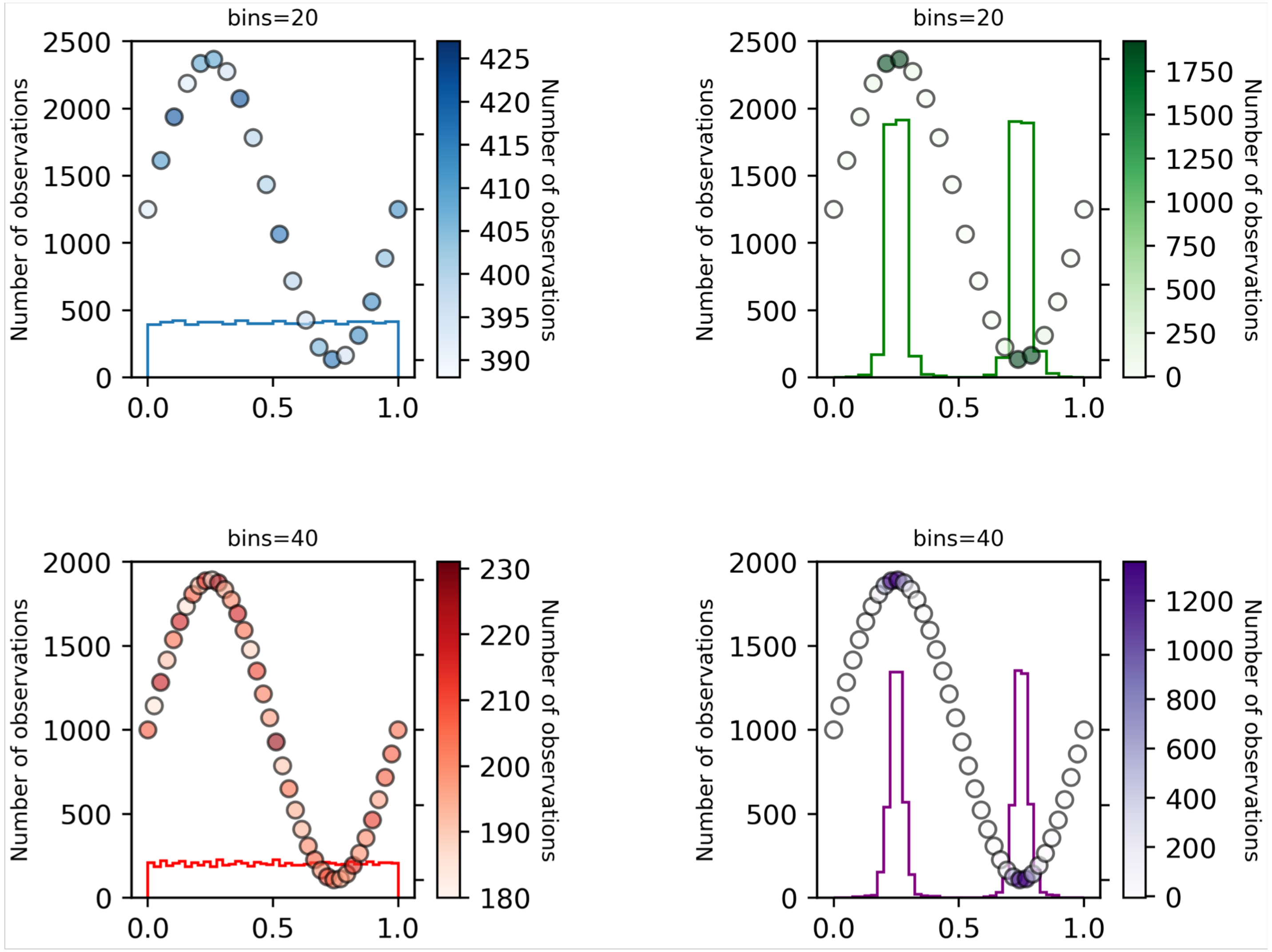



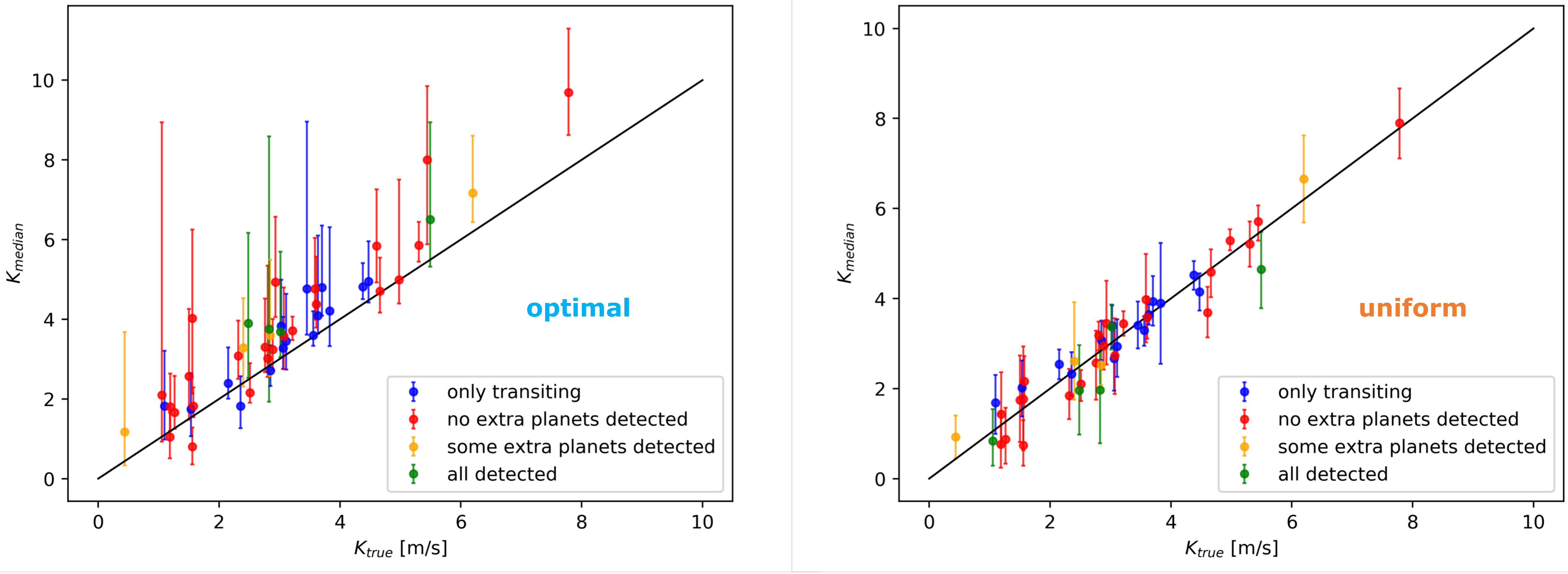
# Transiting planets with non-circular orbits

**ESPRESSO GTO-type schedules for transit follow-up** were simulated 7 times, assuming that **30% of the ESPRESSO guaranteed time for exoplanet science** was available during **3 years**, which amounts to **1102 observational slots of (15+9) minutes**, i.e. about **61.4 effective nights**.

**Optimal (for $e$ =0)** and **uniform-in-phase** schedulers were then applied. The later evenly distributes the available observational slots among all 50 stars, while ensuring each star is observable with airmass below 2 and the phase space is sampled as uniformly as possible (uses a maximin space-filling design based on $L^2$ relaxation, e.g. Pronzato 2016). **For each possible time, $t_i$, the value of $\sigma_i$ was estimated using the ESPRESSO ETC, and then a time-independent contribution from stellar activity, specific to each star, was added in quadrature (Cegla et al. 2014).**

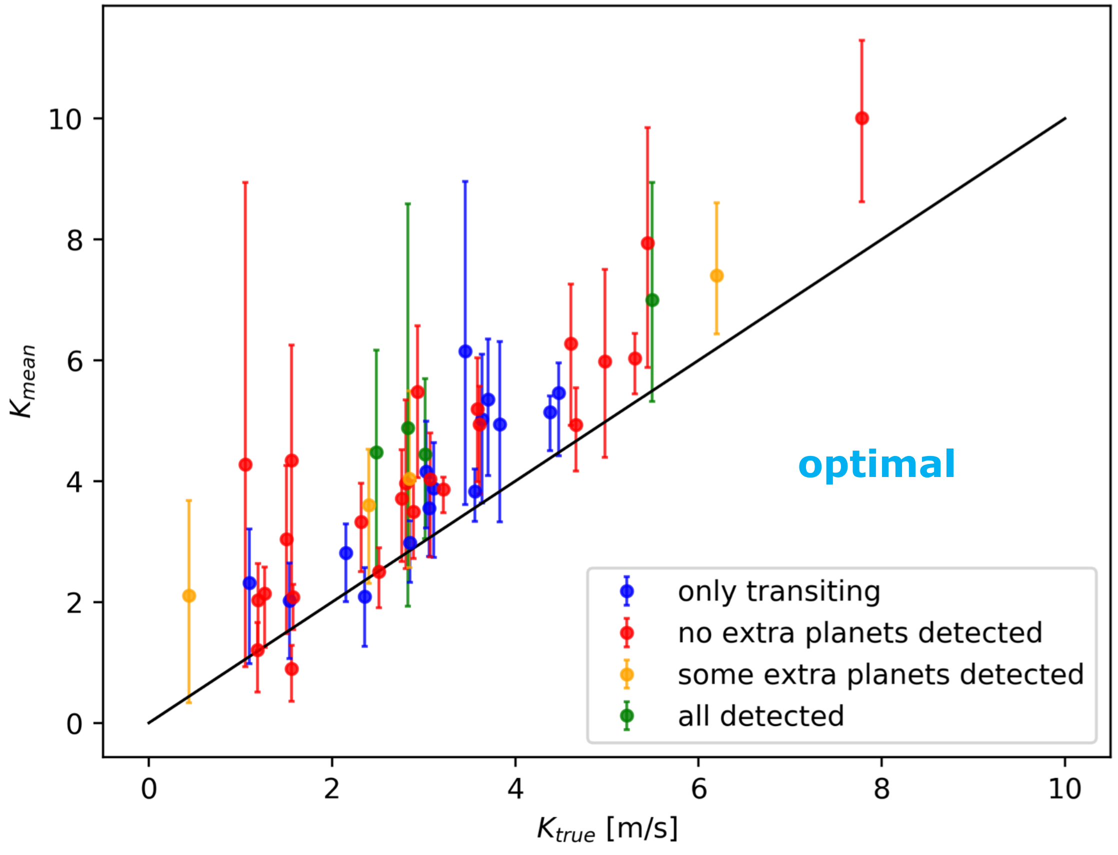


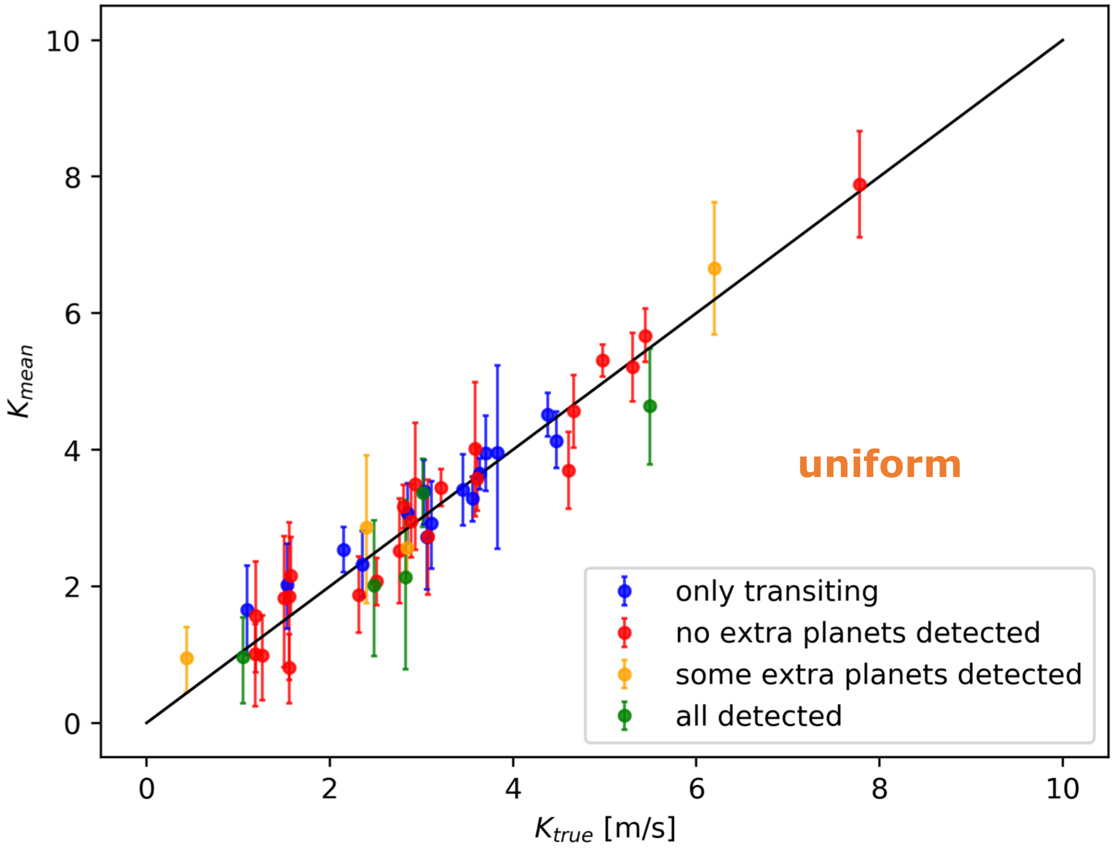
# Transiting planets with non-circular orbits




**Medians** plus symmetric credible intervals centred on median containing 0.68 of the probability, for one realization of each type of schedule

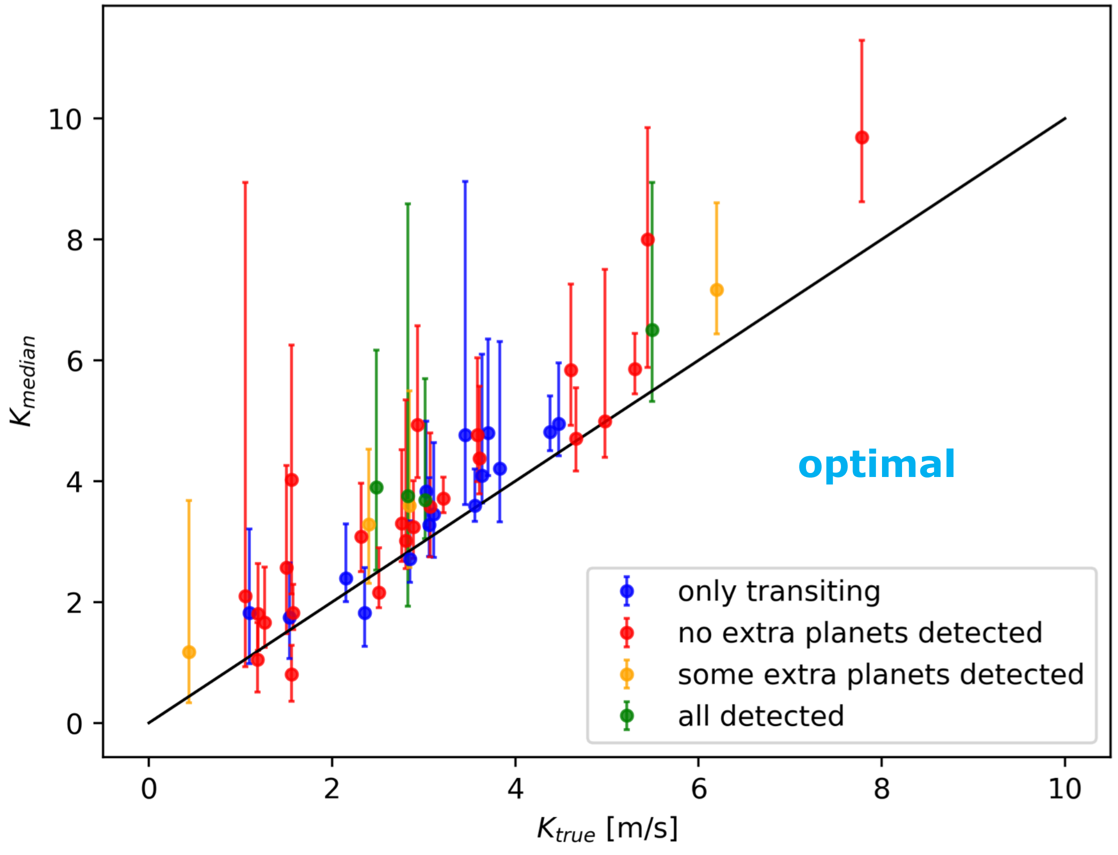
# Transiting planets with non-circular orbits

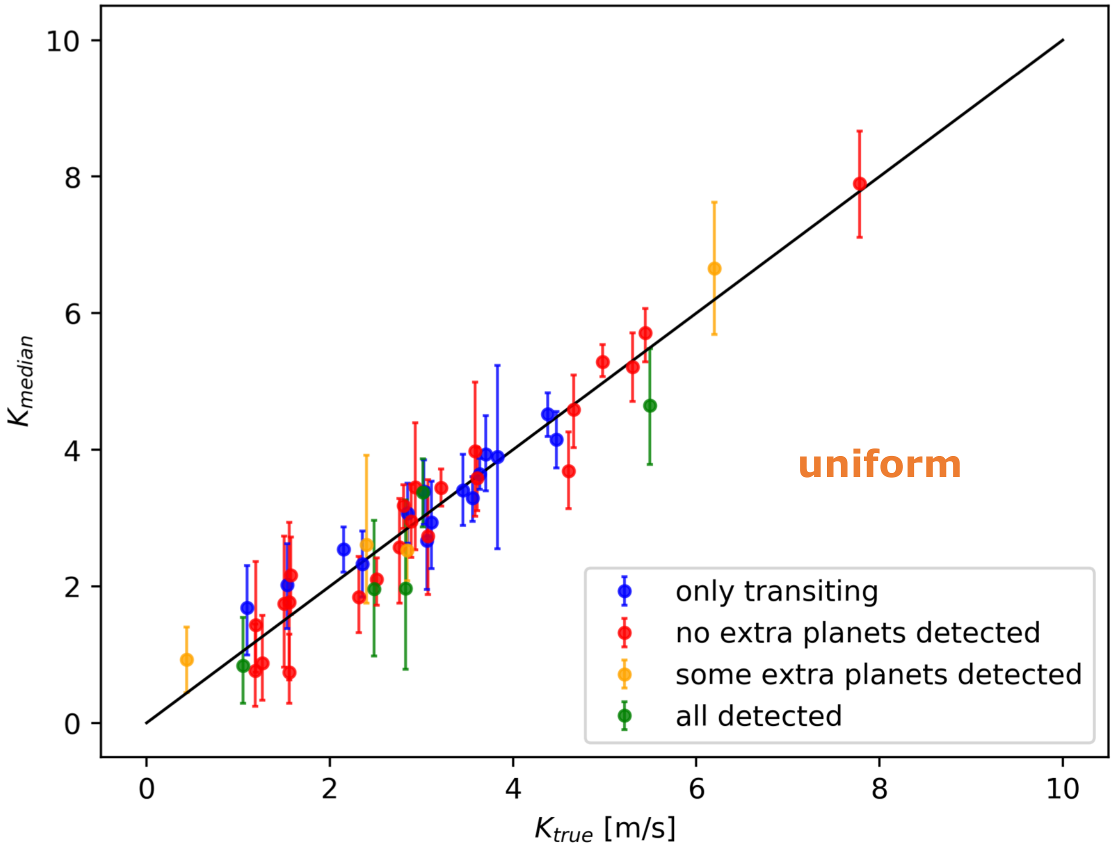



**Means plus symmetric credible intervals centred on median containing 0.68 of the probability, for one realization of each type of schedule**

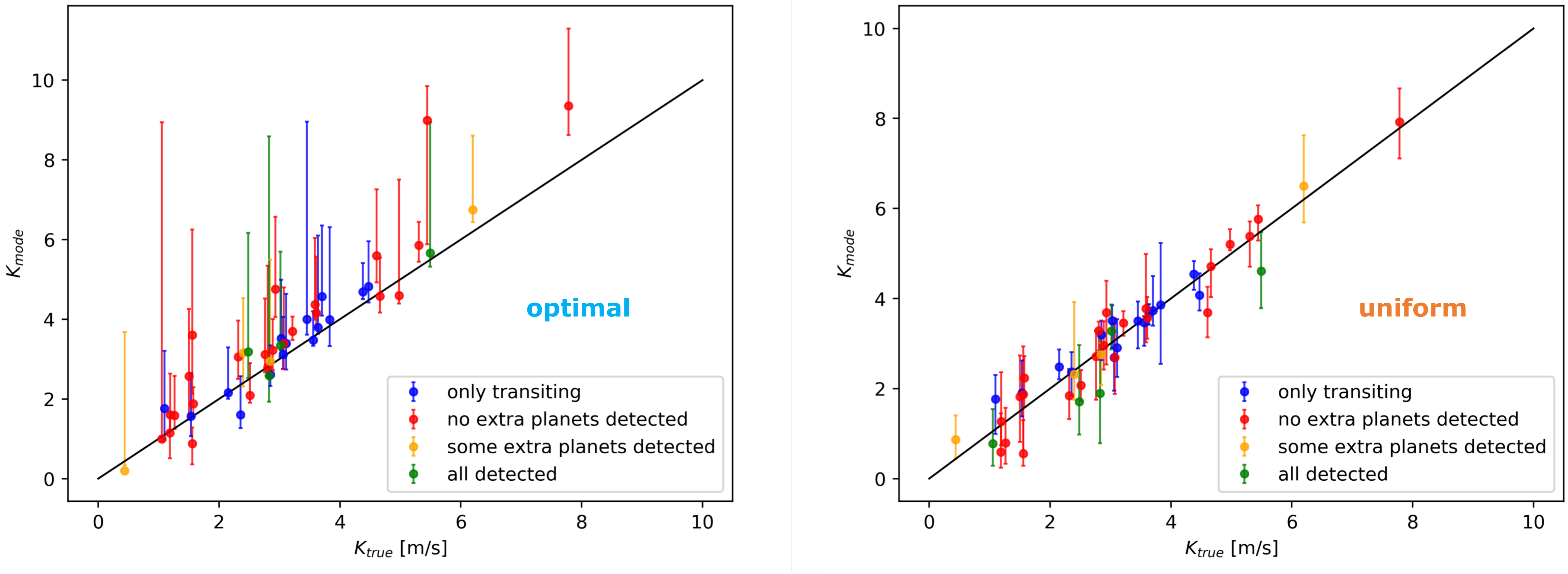
# Transiting planets with non-circular orbits



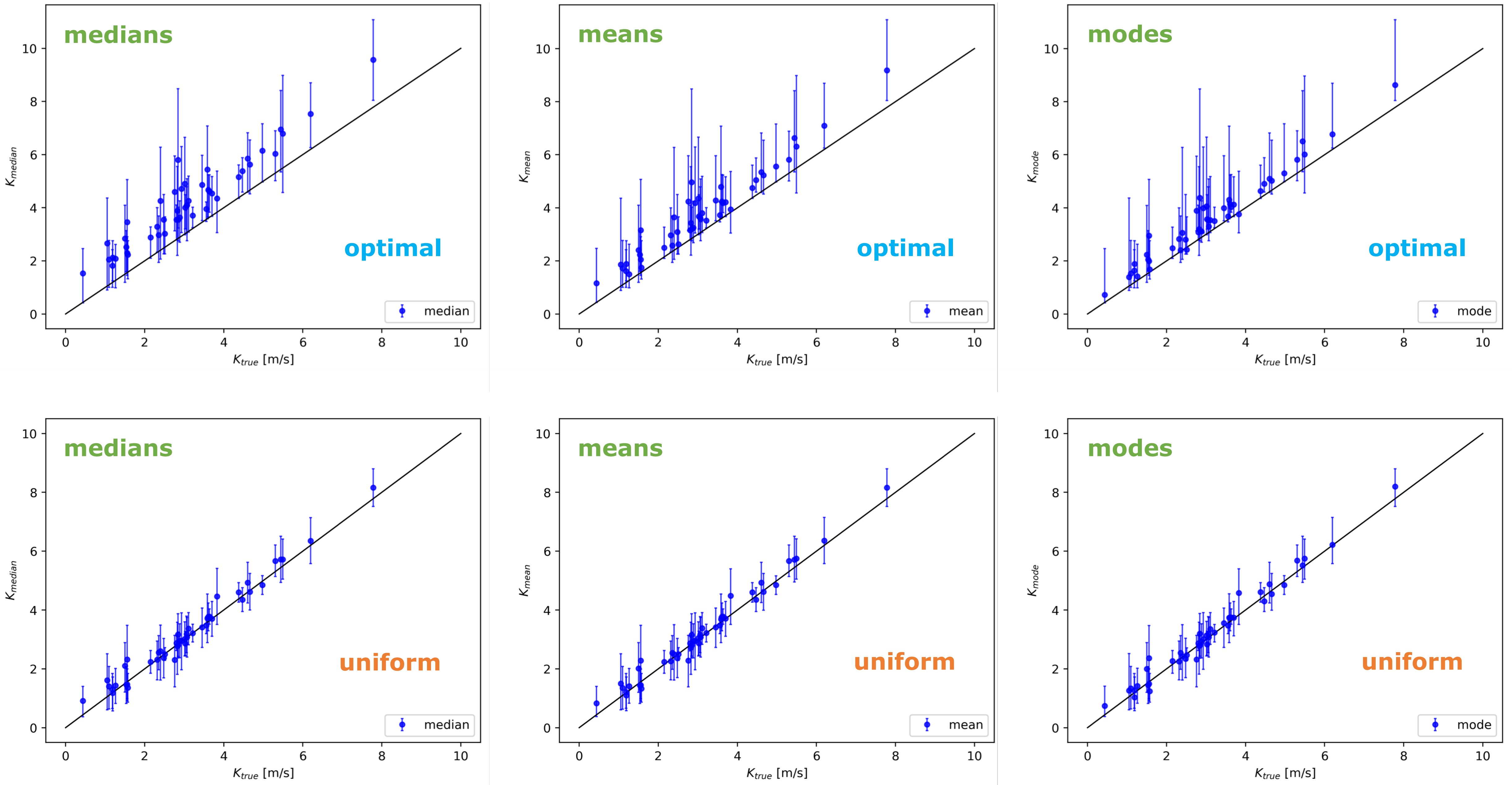

**Medians plus symmetric credible intervals centred on median containing 0.68 of the probability, for one realization of each type of schedule**

# Transiting planets with non-circular orbits

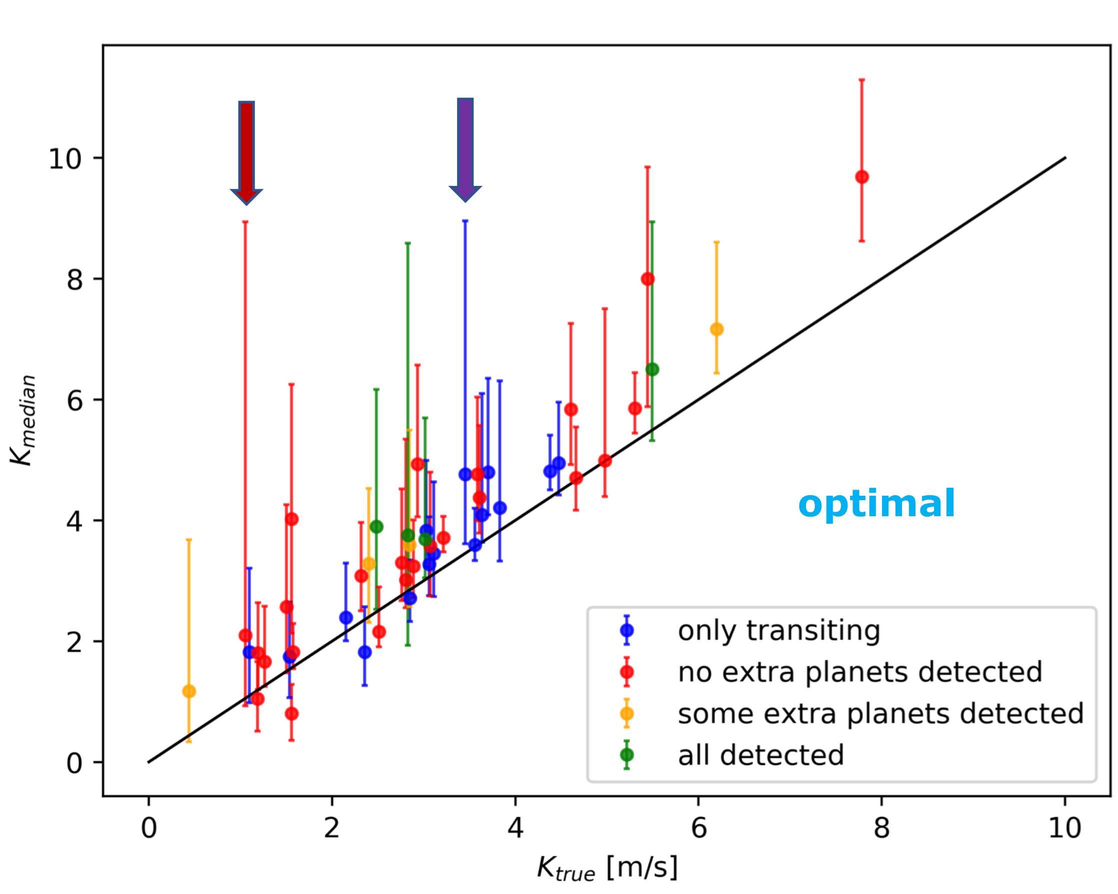


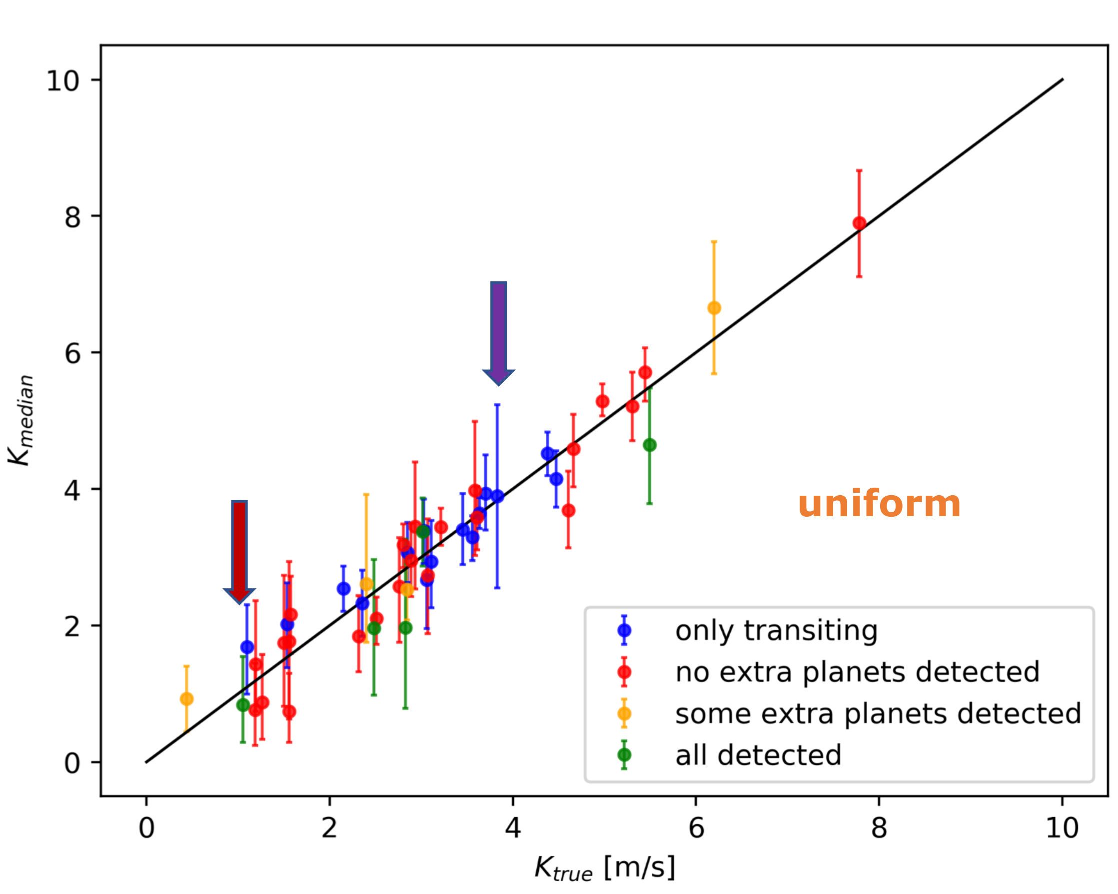

Modes plus symmetric credible intervals centred on median containing 0.68 of the probability, for one realization of each type of schedule

**Average over 7 realizations of each type of schedule, of medians, means and modes, as well as of the limits of the symmetric credible intervals centred on median containing 0.68 of the probability**



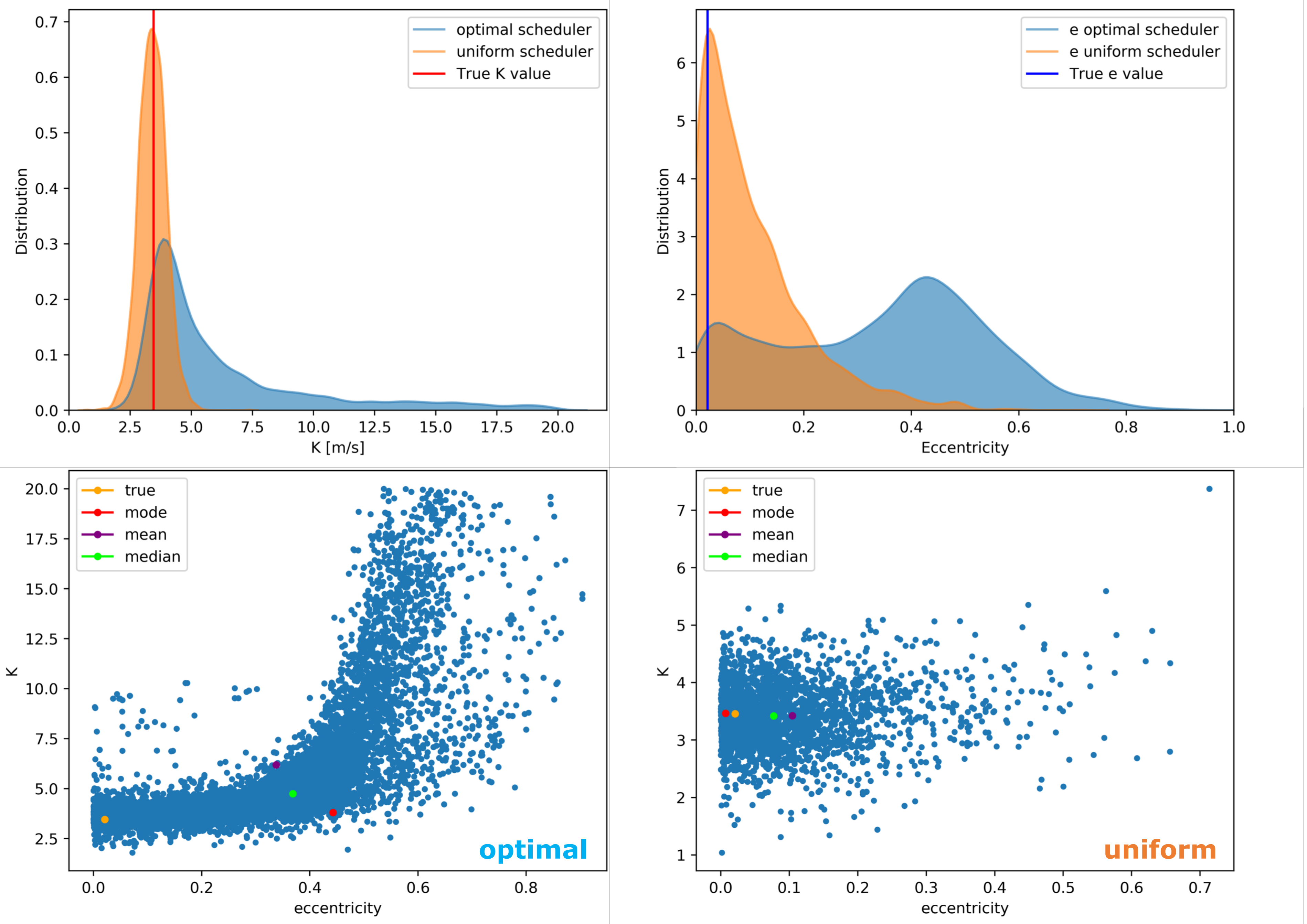
# Transiting planets with non-circular orbits



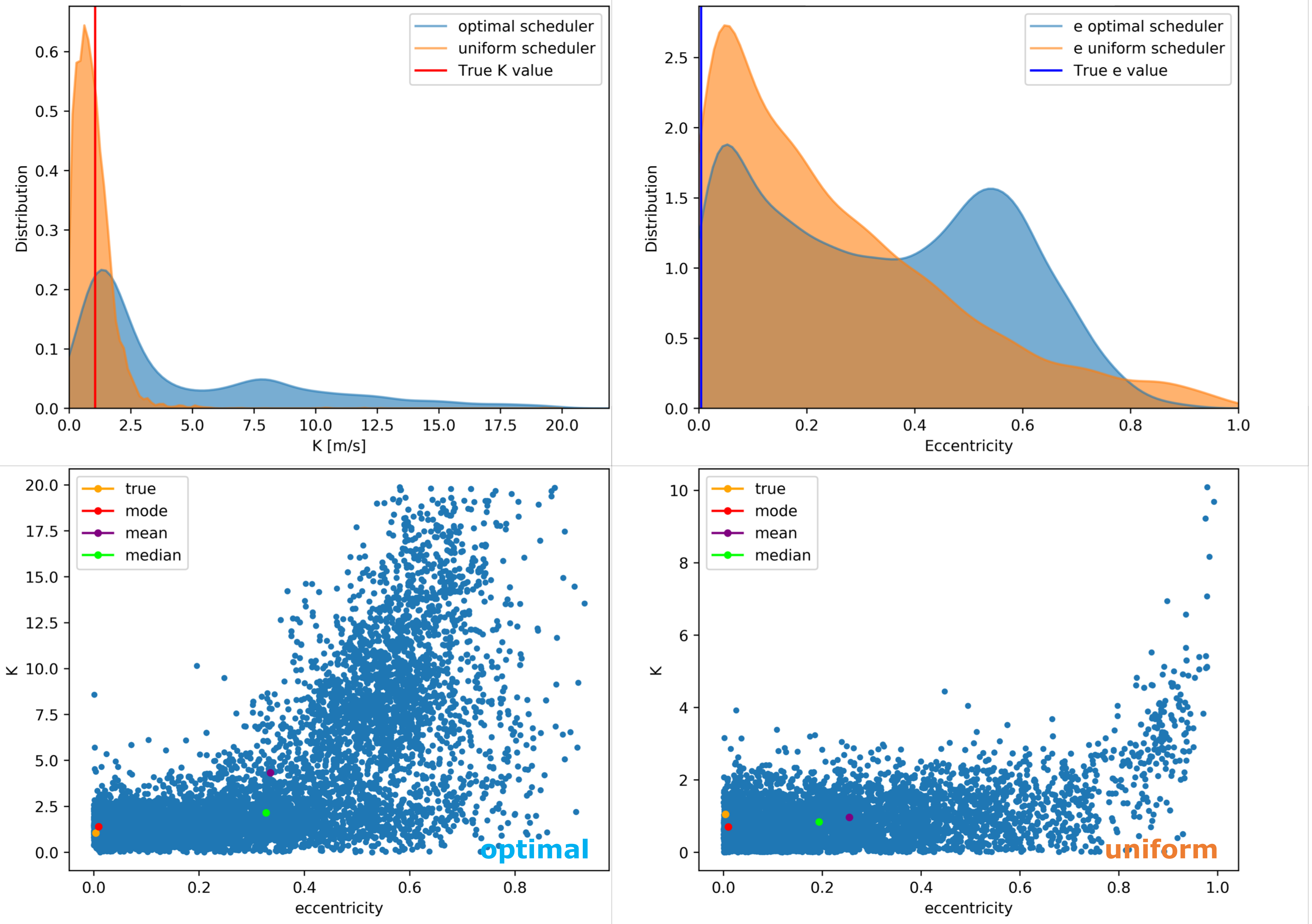




**Medians plus symmetric credible intervals centred on median containing 0.68 of the probability, for one realization of each type of schedule**

# One transiting planet, no extra planet, $\sigma_{act}$ = 2.0 m/s



**One transiting planet, one extra planet (K=11.6 m/s, e=0.057, P=806.1 days), detected under uniform sampling, not-detected under optimal sampling, $\sigma_{act}$ = 2.1 m/s**



## Posterior probability distributions for K, centered on its true value, for 15 systems with only one (transiting) planet, assuming *e* = 0 and no other planets exist