Forward Modeling the Architectures of Exoplanetary Systems: A Clustered Model using Kepler Data

PennState
Eberly College
of Science

Natural Sciences and Engineering
Research Council of Canada

Matthias Y. He
Eric B. Ford (Advisor)
Pennsylvania State University
Darin Ragozzine
Brigham Young University

ExoPAG 21: January 4, 2020

Kepler's multi-transiting systems are extremely informative to study

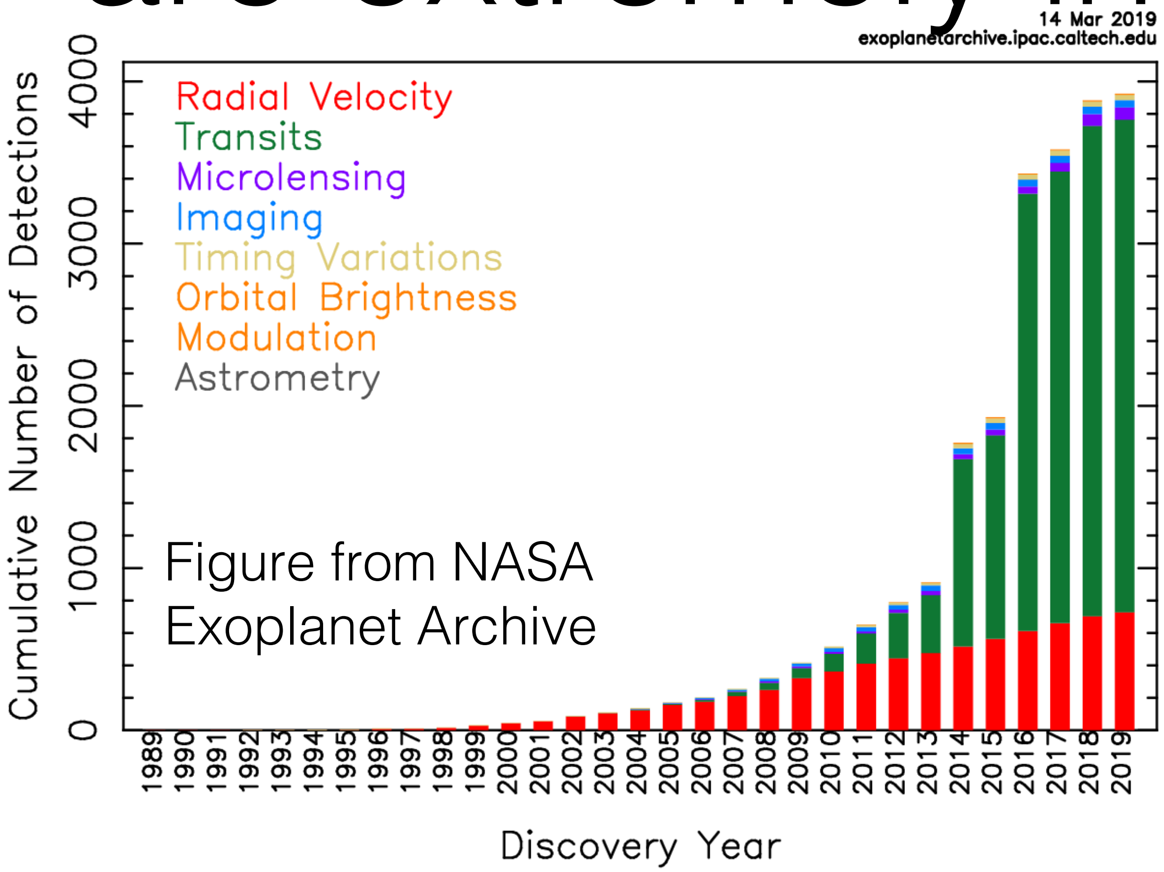
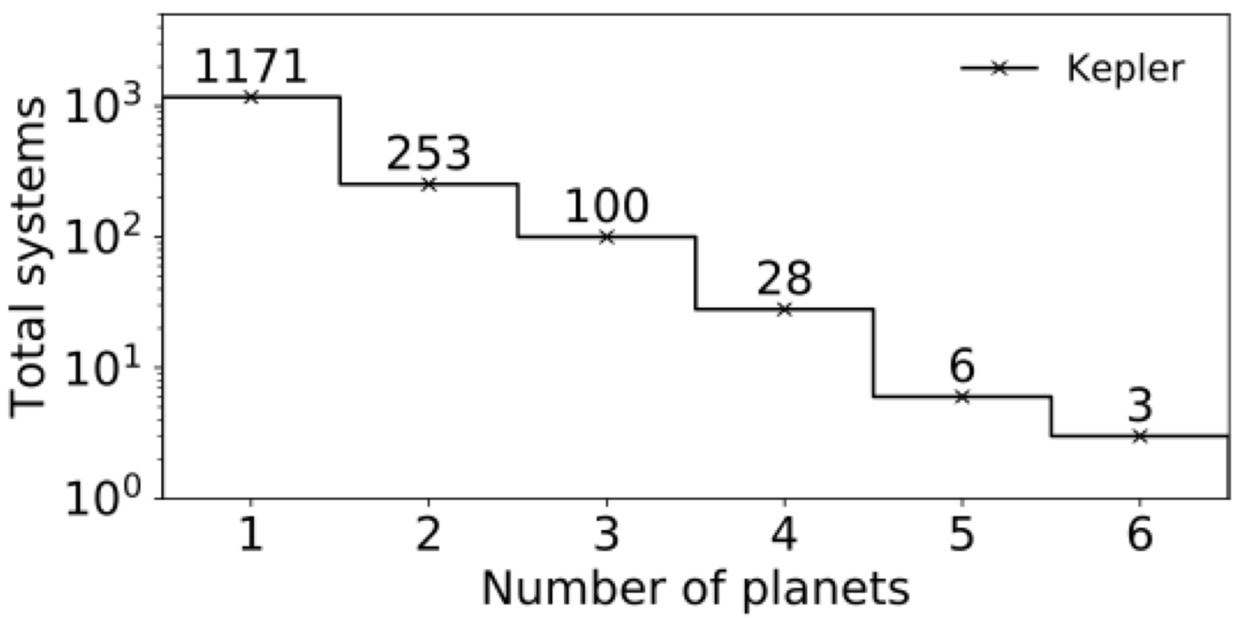
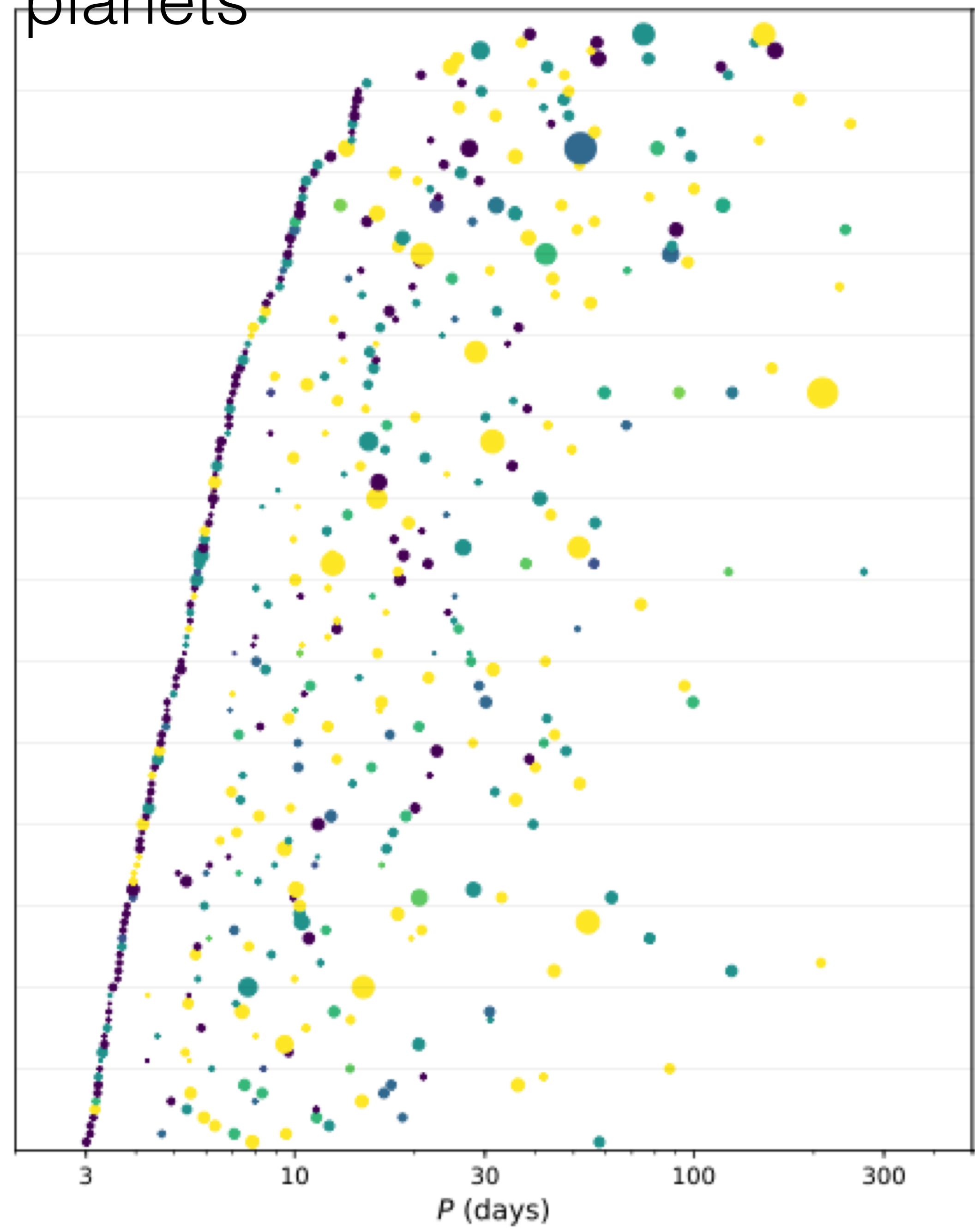


Figure from NASA Exoplanet Archive

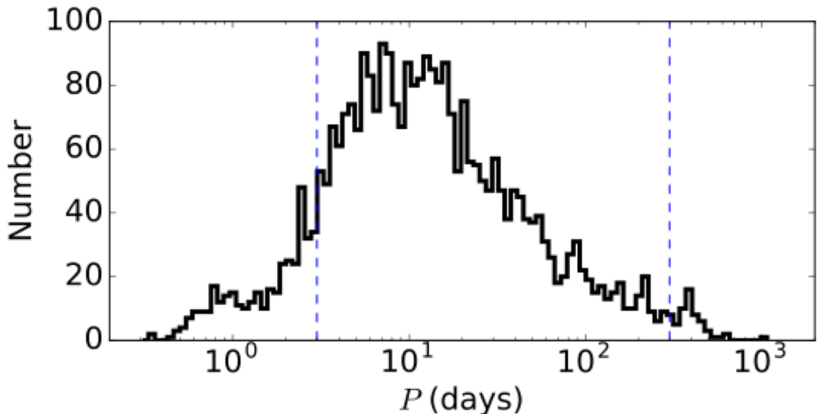


Kepler systems with 3+ planets

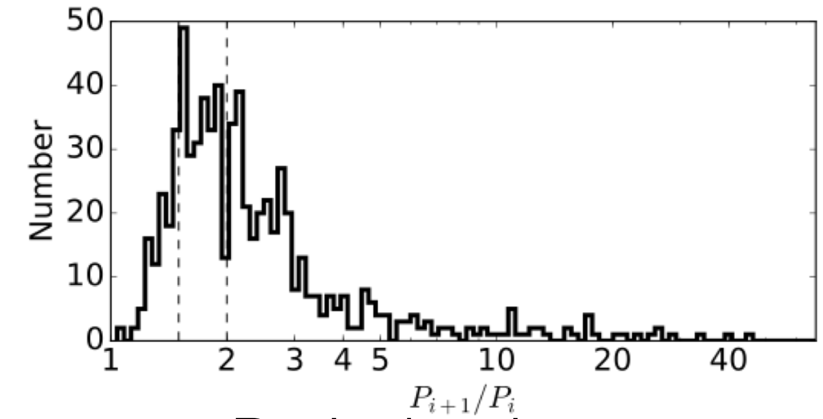


Many of these planets are in multi-transiting systems!

How do we make sense of all these planets?

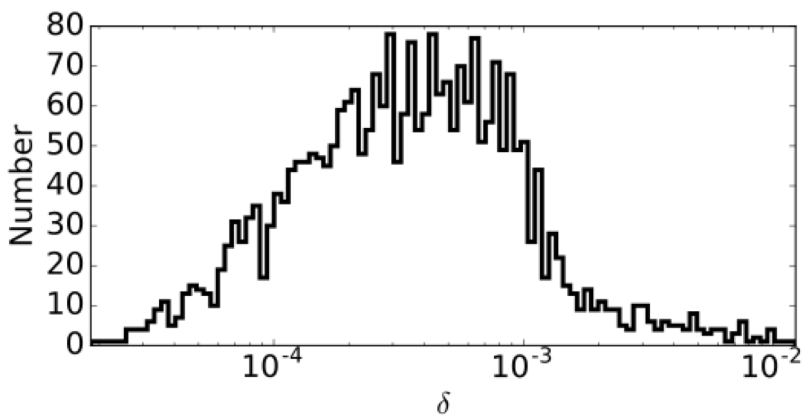


Orbital periods

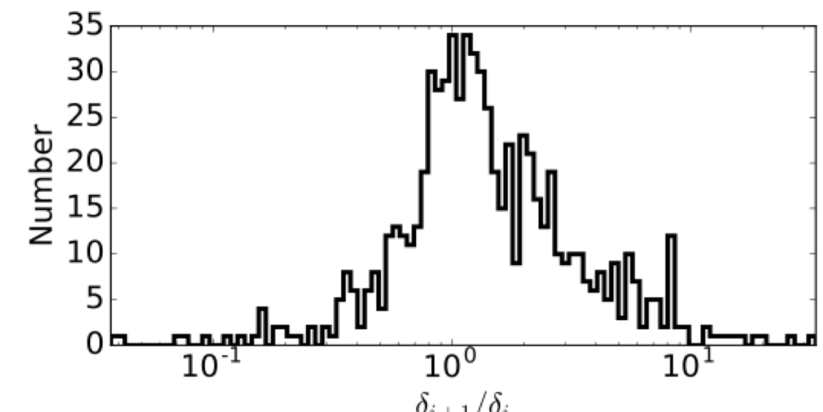


Period ratios

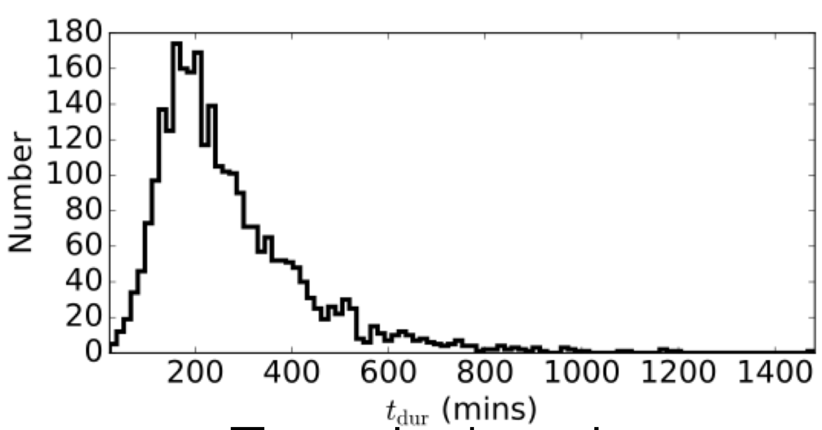
- How do we disentangle **observational biases** from real trends?
- What do these observed planets suggest about the **underlying systems**?
- What are the trends in their **architectures**? What do they suggest about **planet formation processes**?



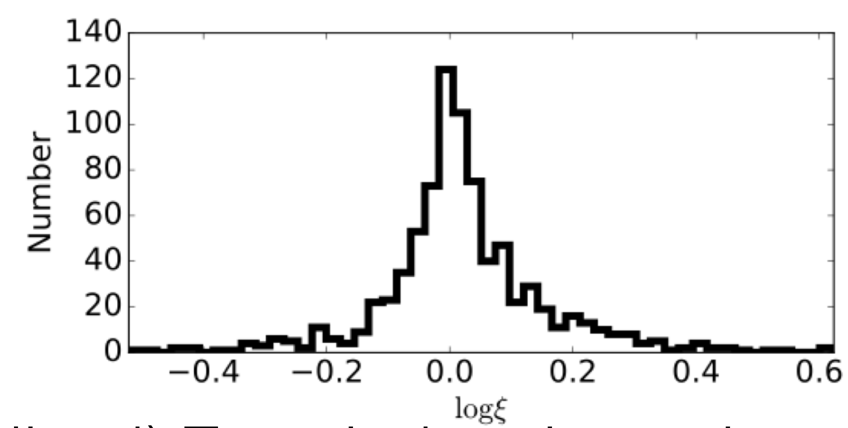
Transit depths



Transit depth ratios



Transit durations



(Period-normalized) Transit duration ratios

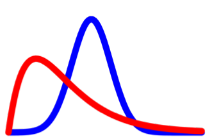
Framework for forward modeling planetary systems and the Kepler mission

$f(P, R)$

Step 1: Define a statistical model for the intrinsic distribution of exoplanetary systems.

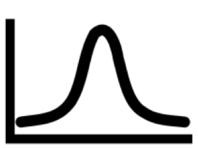
Step 2: Generate an underlying population of exoplanetary systems (*physical catalog*) from a given model.

Step 3: Generate an observed population of exoplanetary systems (*observed catalog*) from the *physical catalog*.



Step 4: Compare the simulated *observed catalog* with the Kepler data using a distance function.

Step 5: Optimize the distance function to find the best-fit model parameters.

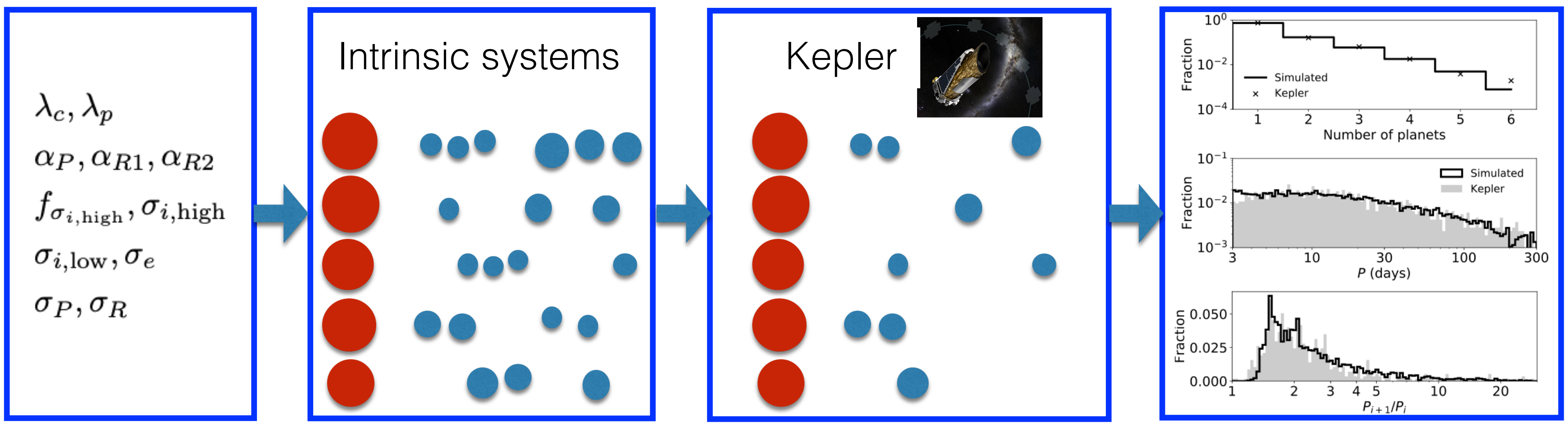


Step 6: Explore the posterior distribution of model parameters using a Gaussian Process (GP) model.

Framework for forward modeling planetary systems and the Kepler mission

We have a **full forward model** for simulating the Kepler mission!

Define model | *Simulate a physical catalog* | *Simulate an observed catalog* | *Compare with Kepler data*



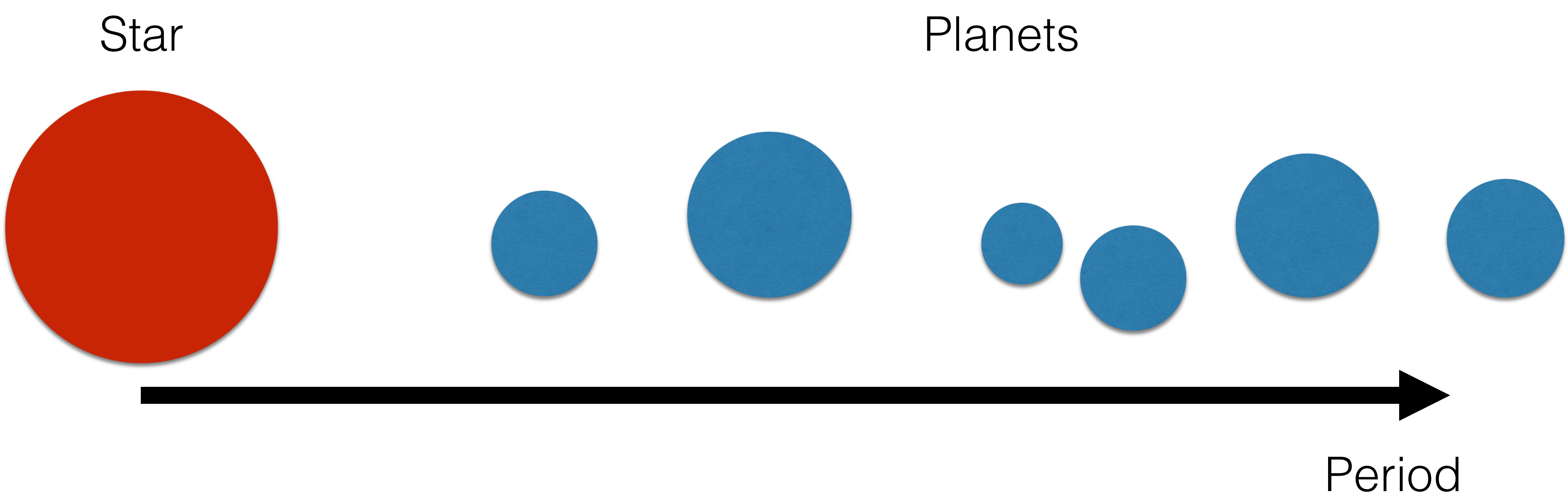
Key differences:

Updated stellar radii from Gaia DR 2 (and b_p-r_p colors)

Detailed Kepler detection efficiency
Hsu et al. (2018, 2019)
He, Ford, & Ragozzine (2019)

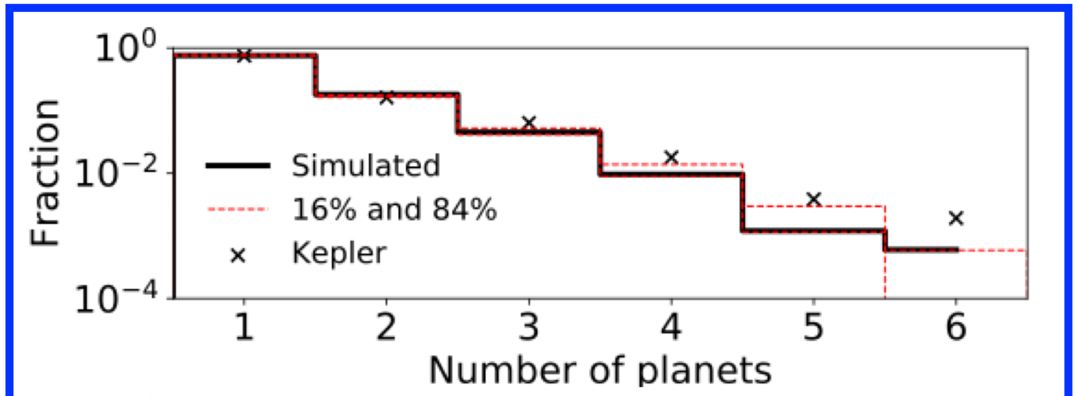
Kepler DR 25 catalog (uniform vetting)
Fit to all observed marginal distributions simultaneously

Many previous studies assume that planets are independent in period and in size

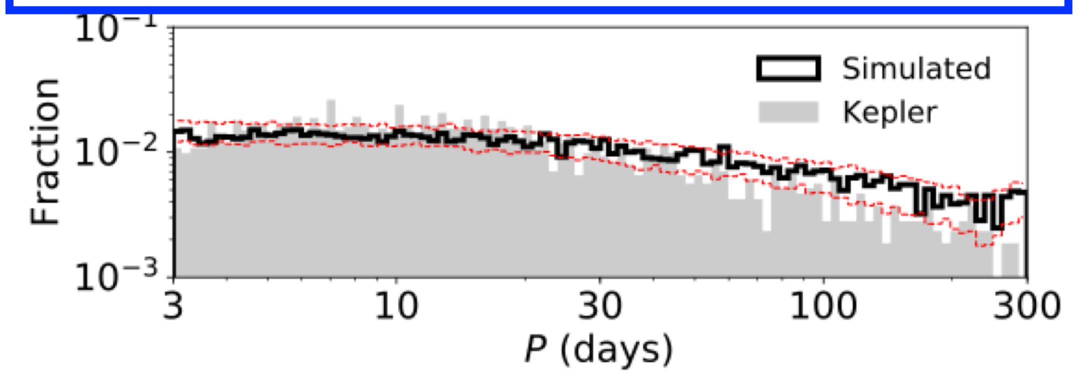


We test a model where the period and radius of each planet are drawn independently

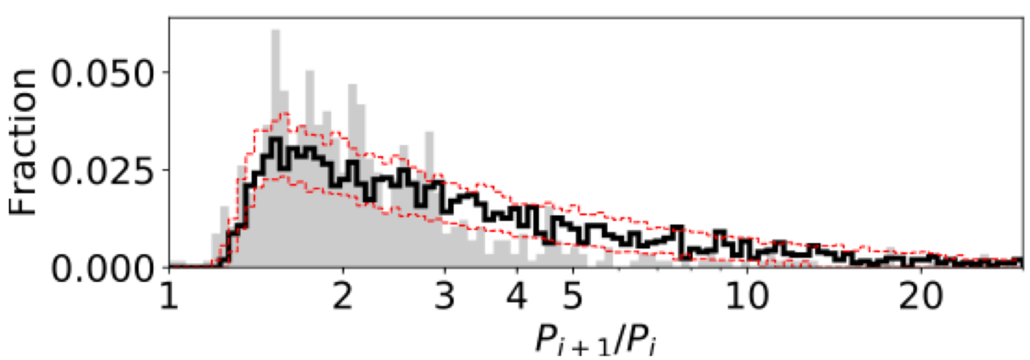
Models assuming independent planets fail to reproduce the observed population



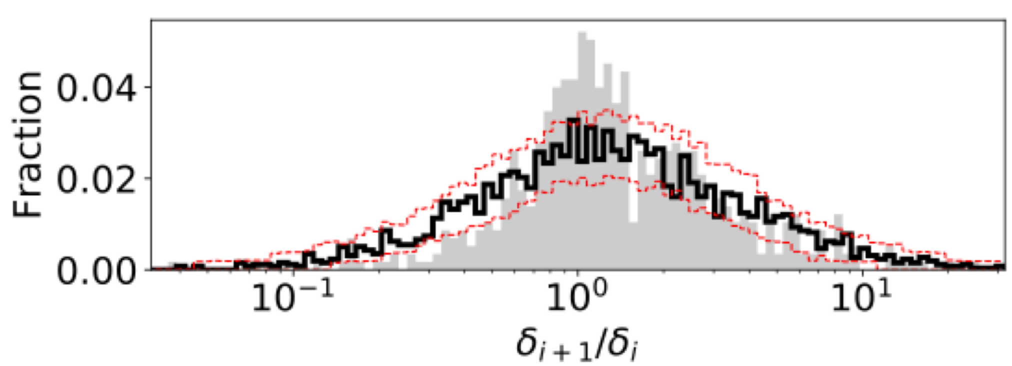
The number of multi-transiting systems is significantly under-produced



The period distribution appears well modeled with a single power-law (between 3 and 300 days)

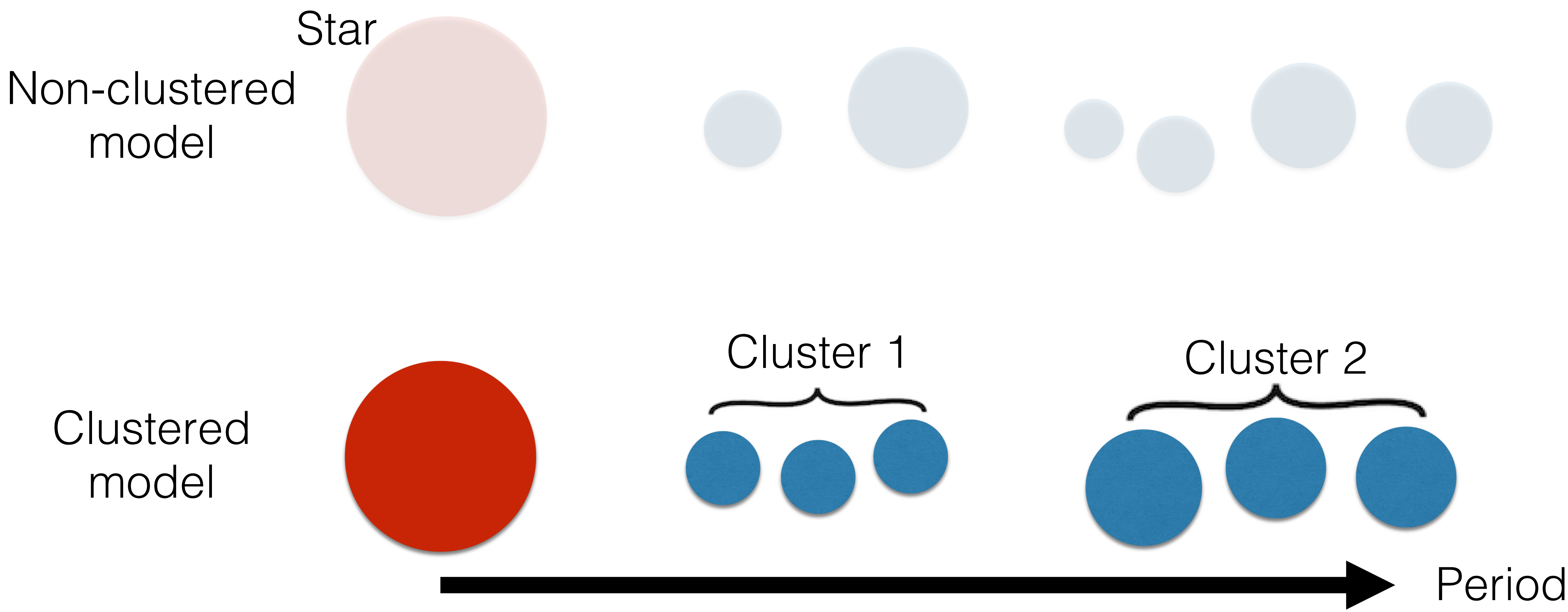


The period ratio distribution is poorly modeled



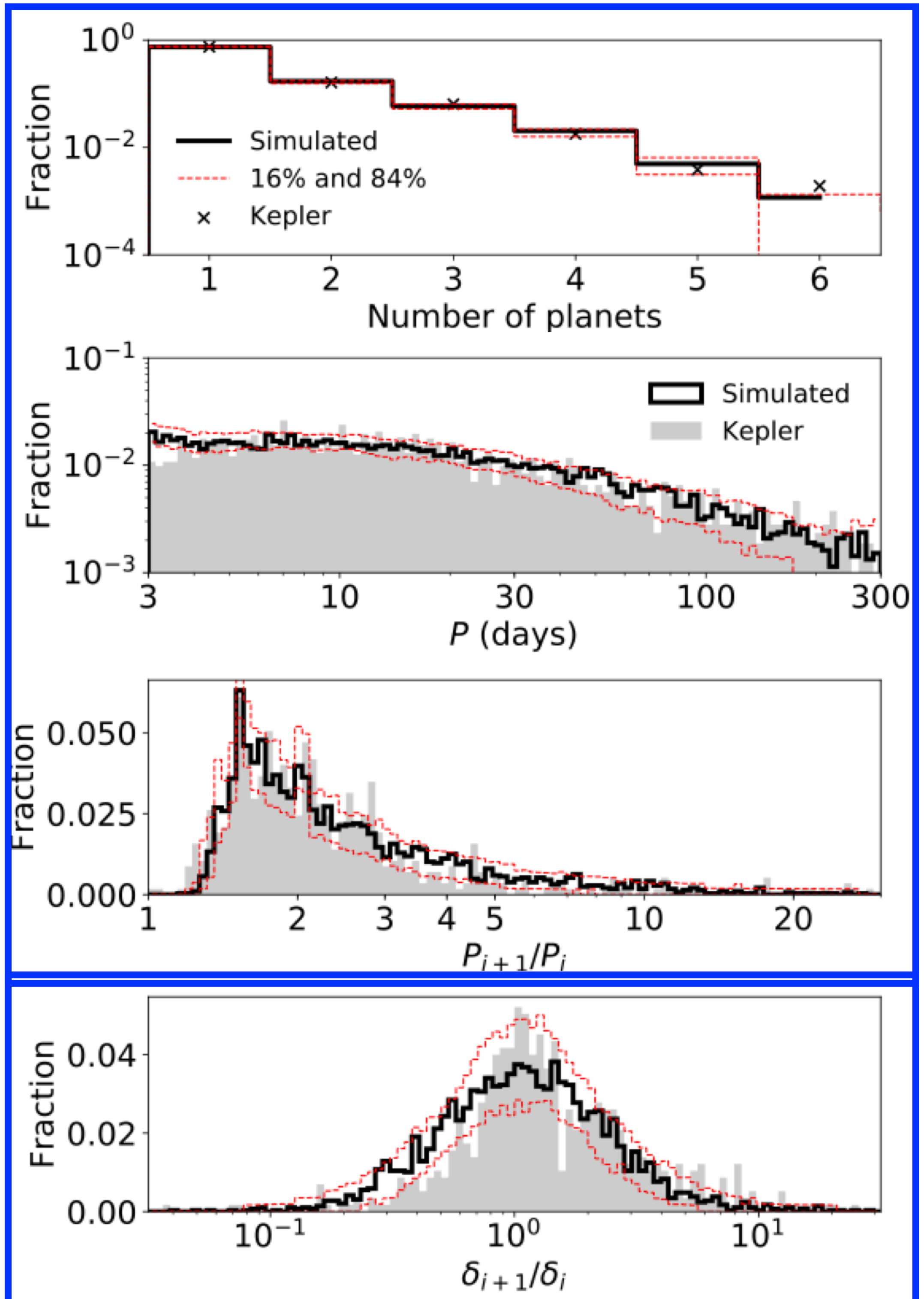
The transit depth ratios are not as peaked

There are significant intra-system correlations: planets are clustered in periods and sizes



Planets are drawn as a clustered point process, where each cluster has a period scale and radius scale

Our clustered model provides a significantly improved description of planetary systems!

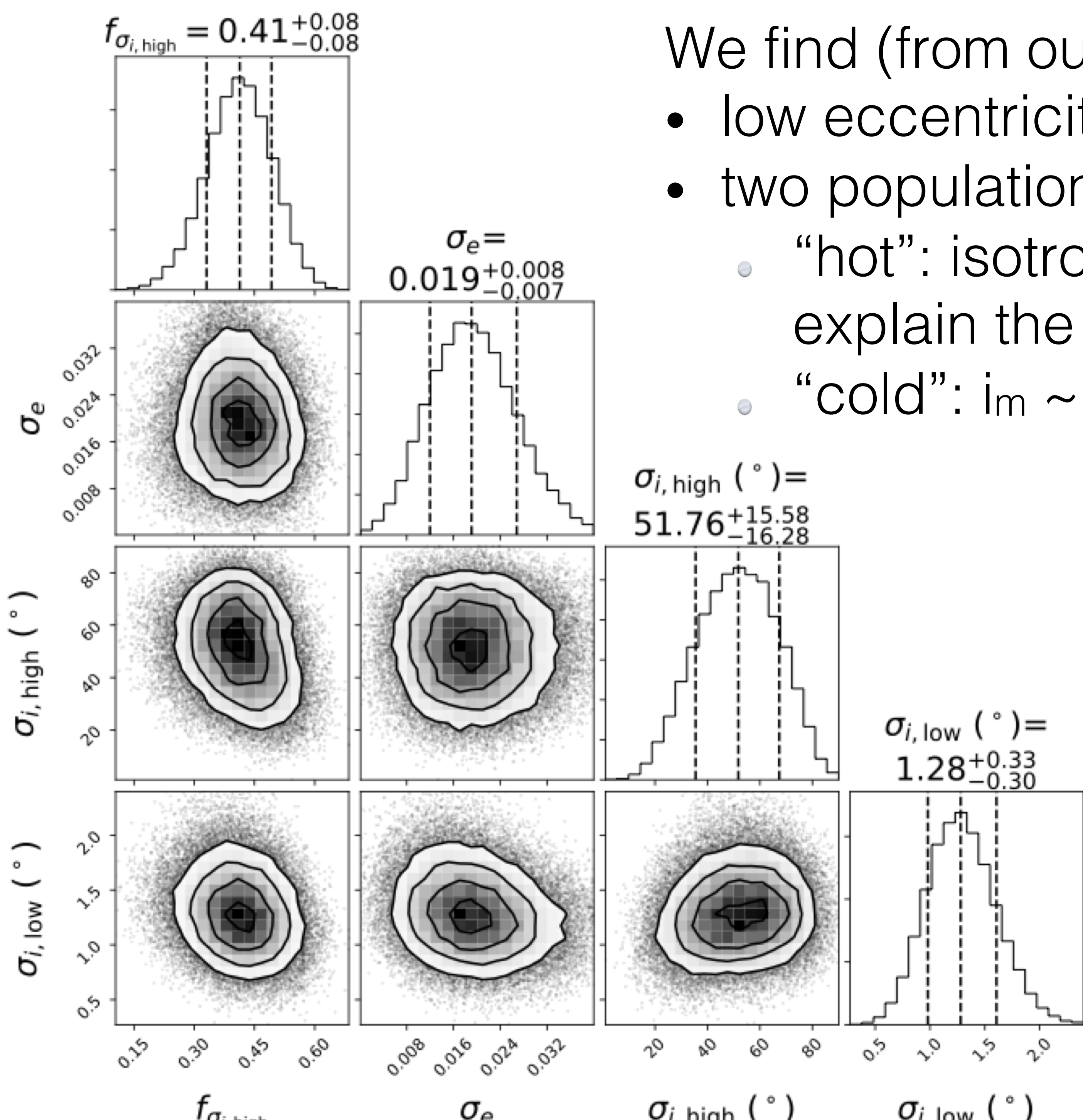


Observed multiplicities are fit extremely well

Both the period and period ratio distributions are well reproduced

Fit to transit depth ratios appear better, but distances not improved

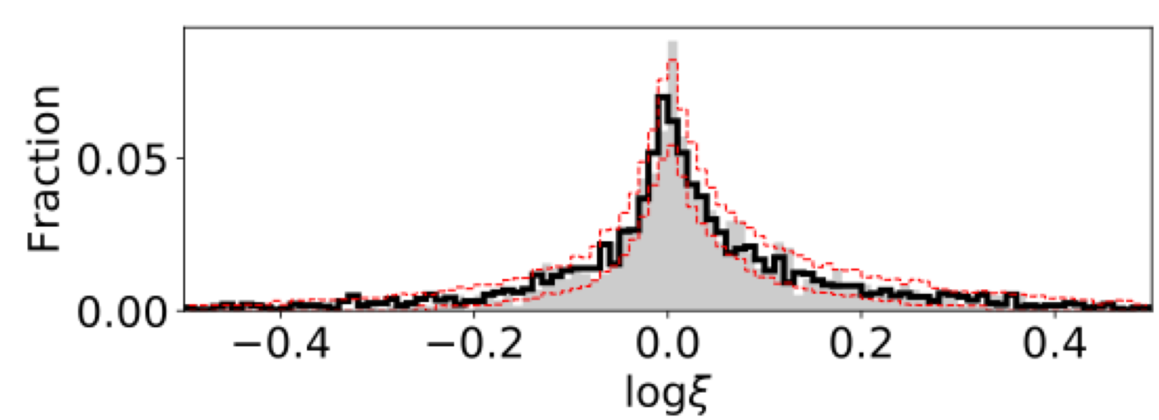
Planetary systems have low eccentricities and consist of two populations of mutual inclinations



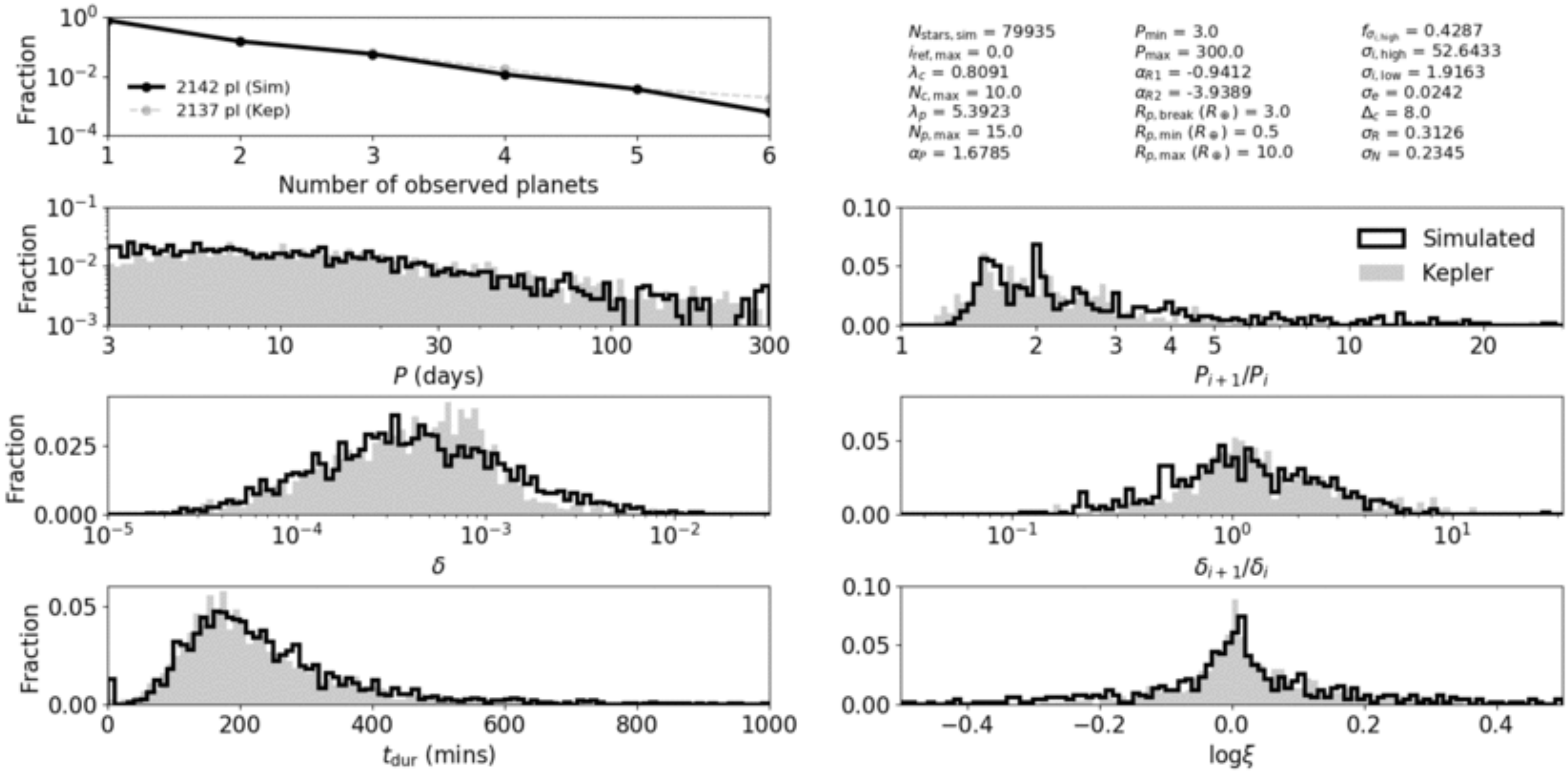
We find (from our clustered models):

- low eccentricities (e ~ 0.02)
- two populations of mutual inclinations
  - “hot”: isotropic i_m for ~40% of systems to explain the excess of single-transit systems
  - “cold”: $i_m \sim 1.3°$ for remaining ~60% of systems

The transit duration ratio distribution is very well matched



Our clustered models provide a great fit to the Kepler data!

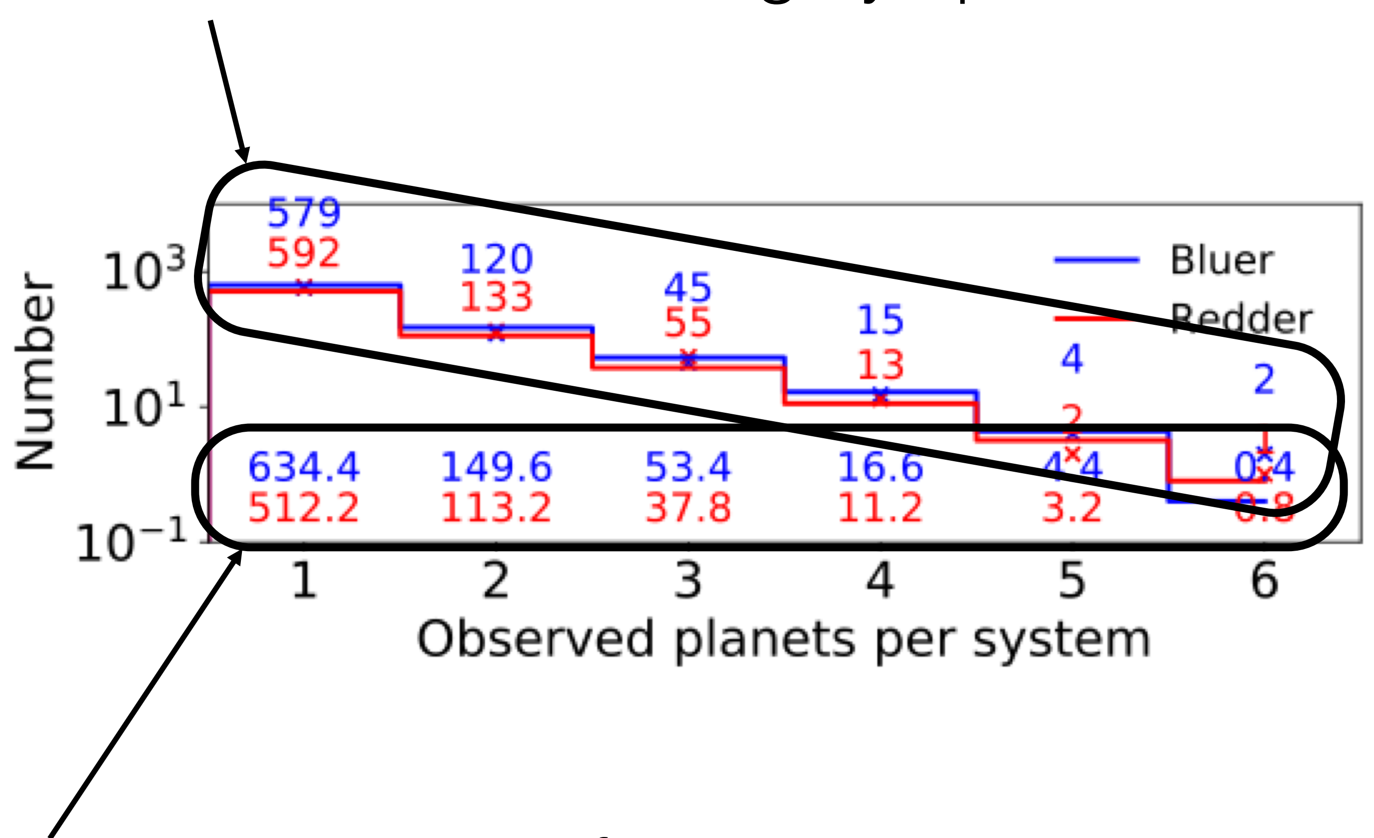


Access our model catalogs (or simulate your own) at:

https://github.com/ExoJulia/SysSimExClusters

What about correlations between planetary systems and their host stars?

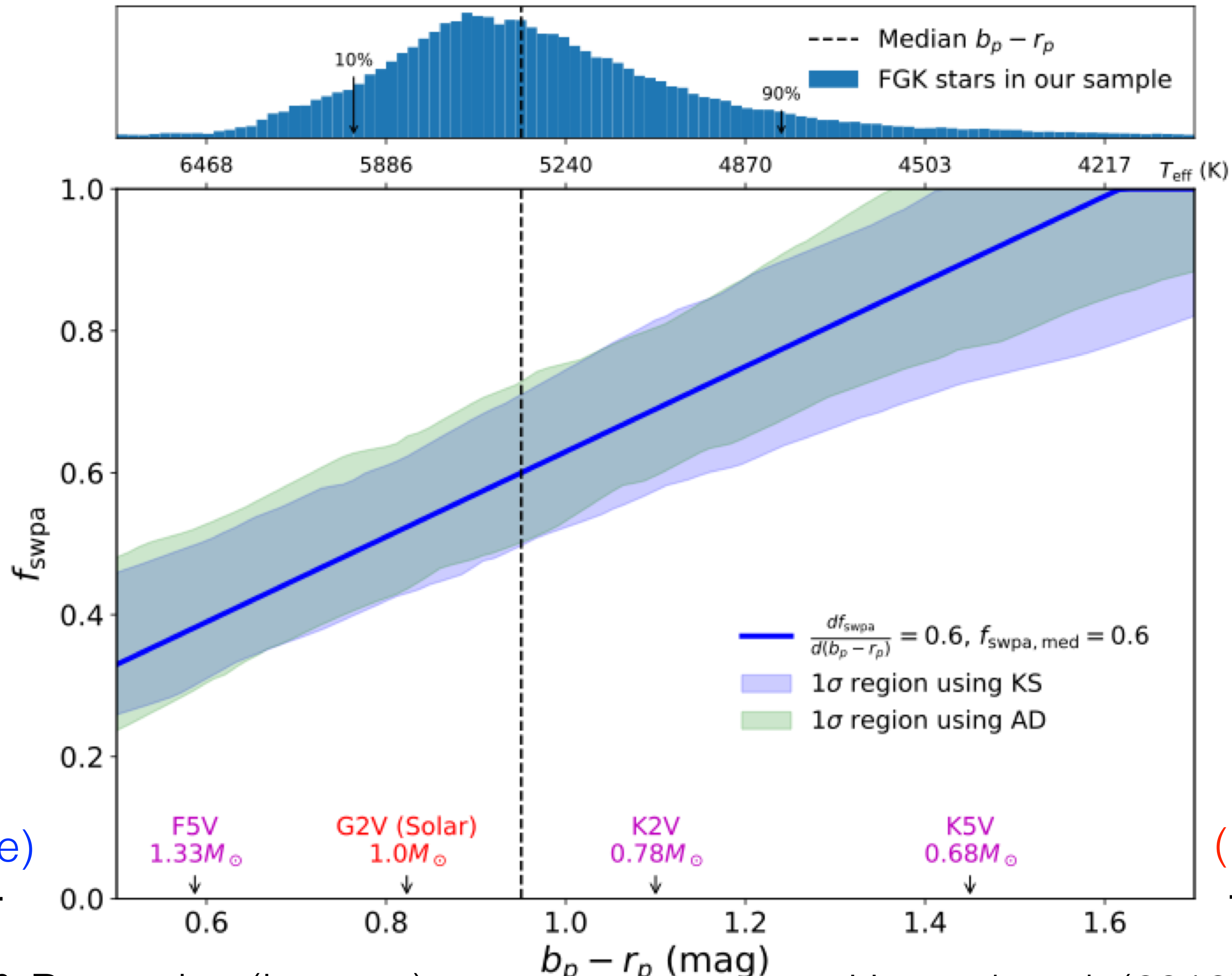
We split the stellar sample (79935 FGK stars) in half by Gaia b_p-r_p color
—> The exoplanet counts are also roughly split in half



Assuming the same distribution of planets around all stars produces more detected planets around the bluer stars!

Planetary systems are more common around late type stars than early type stars

Fraction of stars with planets



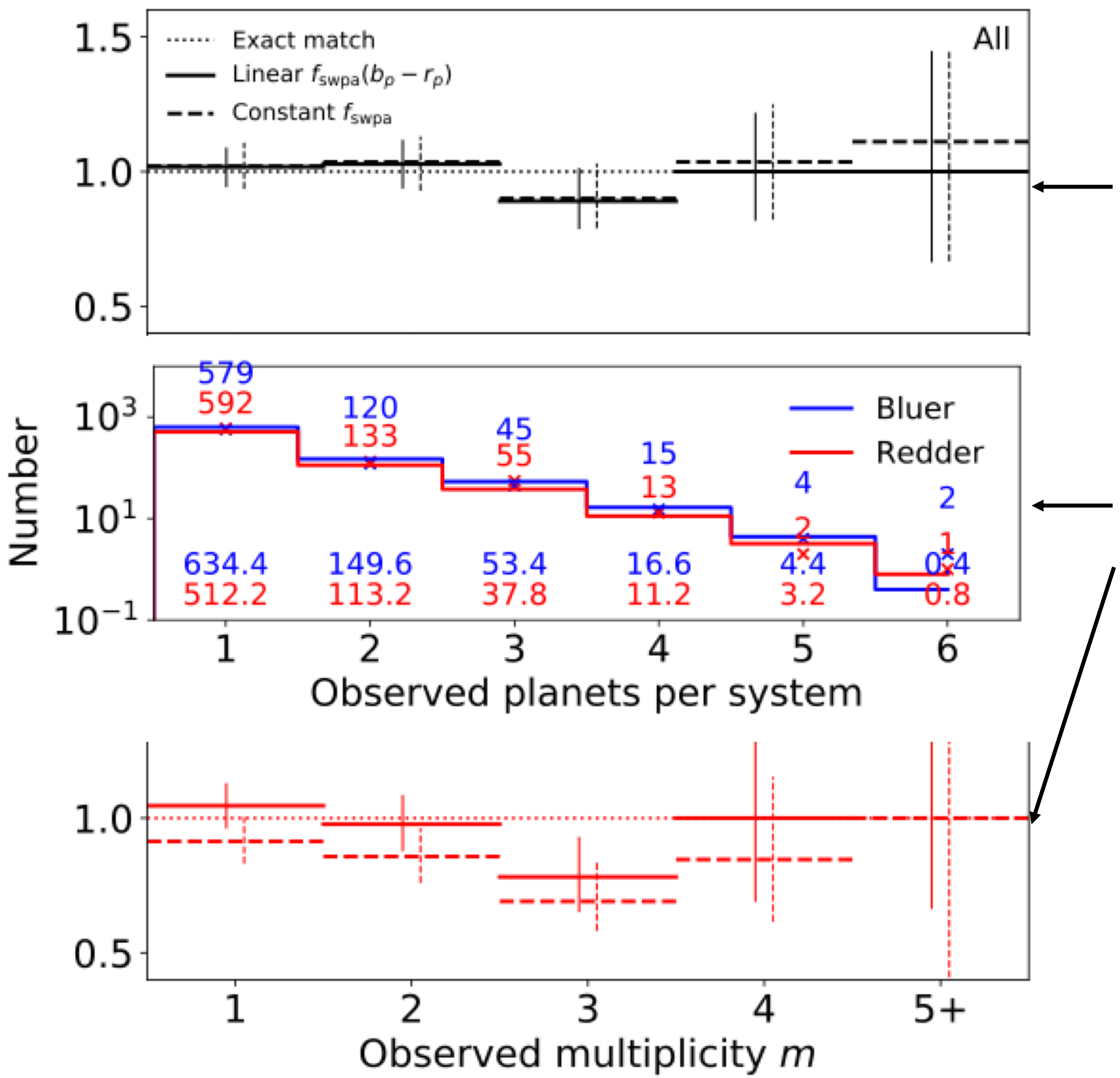
bluer (earlier type) ←

→ redder (later type)

He, Ford, & Ragozzine (in prep.)

Howard et al. (2012)
Mulders, Pascucci, & Apai (2015a)

A linear dependence between the fraction of stars with planets (FSWP) and the host star color fits the multiplicity distribution



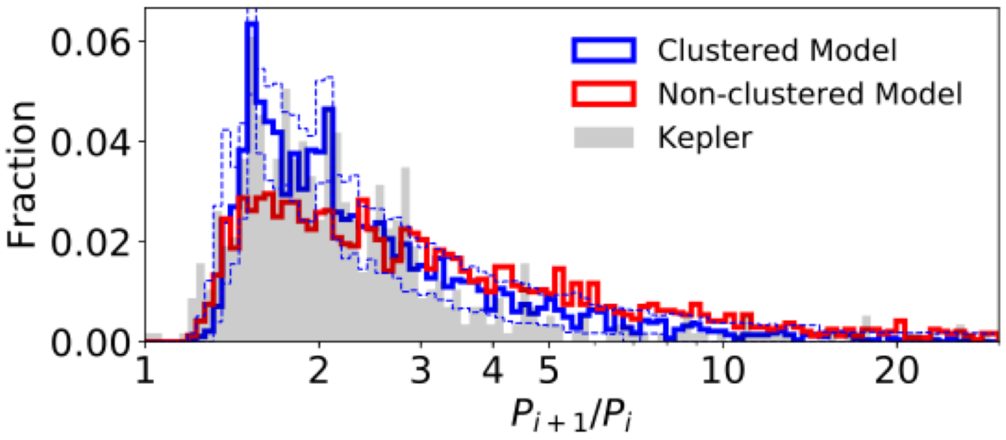
Both clustered models fit the overall multiplicity distribution well, but…

The linear function significantly improves the fit to the "bluer" and "redder" halves

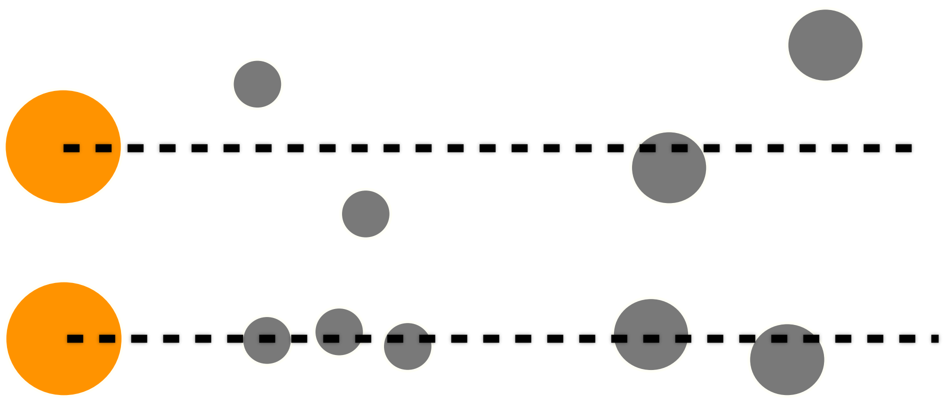
The architectures of planetary systems across FGK stars are similar, aside from the FSWP

Summary

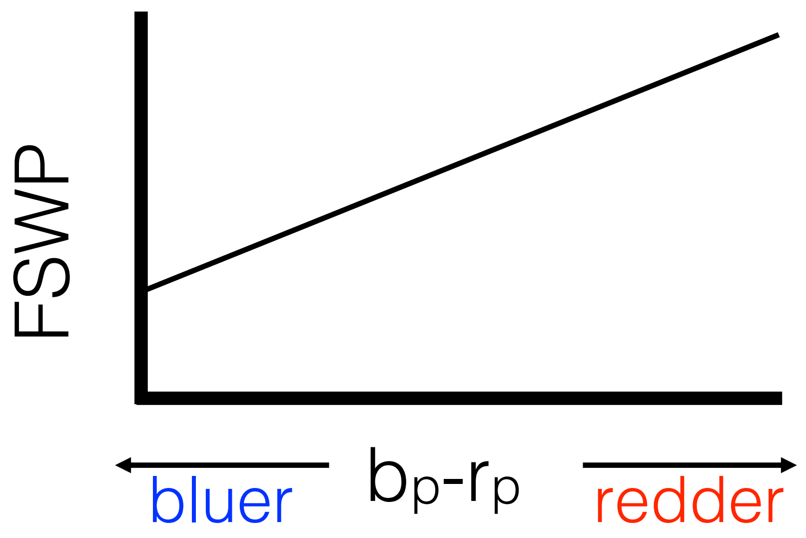
- **Multi-planet systems are clustered in periods and planet sizes**
  - The non-clustered model cannot fit the multiplicities, period ratios, and radius ratios
  - A clustered model better reproduces the multiplicities and period ratios



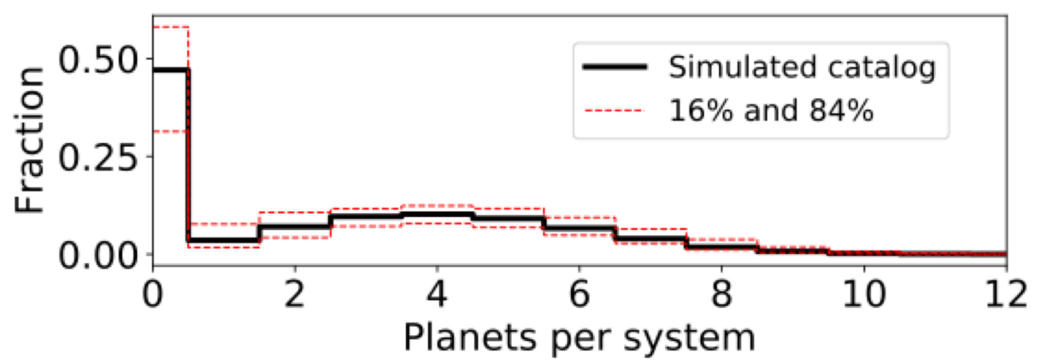
- **There are two populations of orbital architectures: low and high inclinations**
  - e ~ 0.02, i_m ~ 1.3° for 60% of systems
  - i_m > 10° for 40% of systems (Kepler dichotomy)



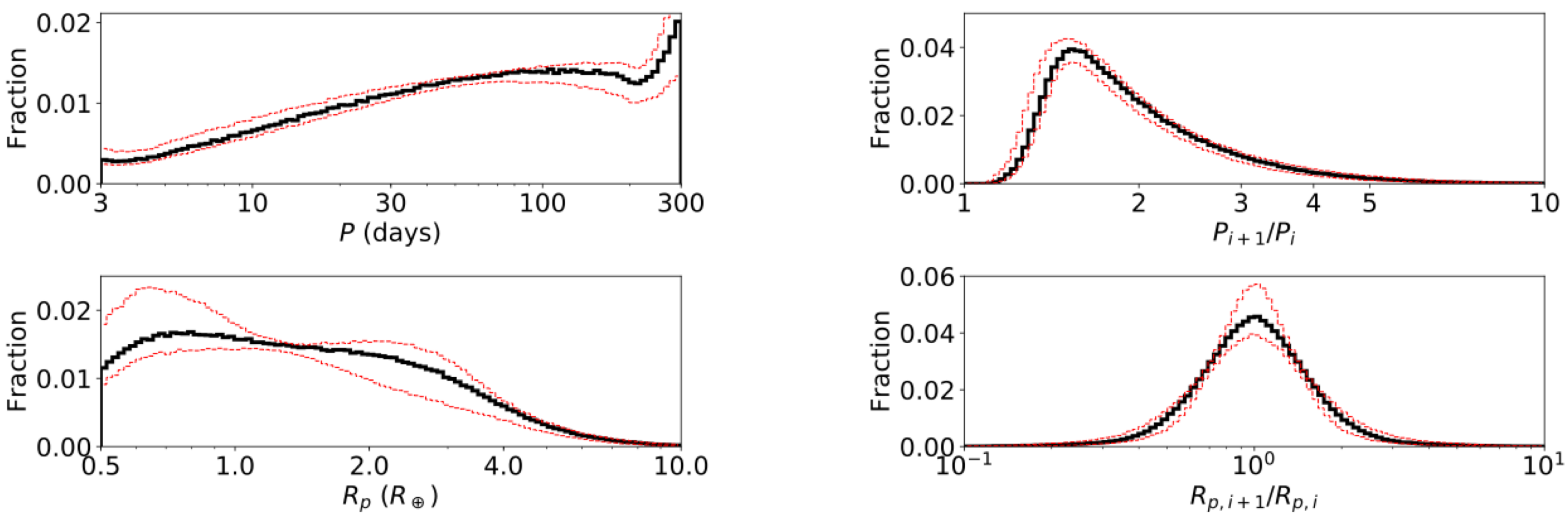
- **Planetary systems are more common around late type (cooler) stars**
  - The *overall* fraction of FGK stars with planets (FSWP) between 3 and 300 days is ~60%
  - FWSP increases from ~30% (early F) to ~100% (late K)



Feel free to use our models and simulated catalogs!



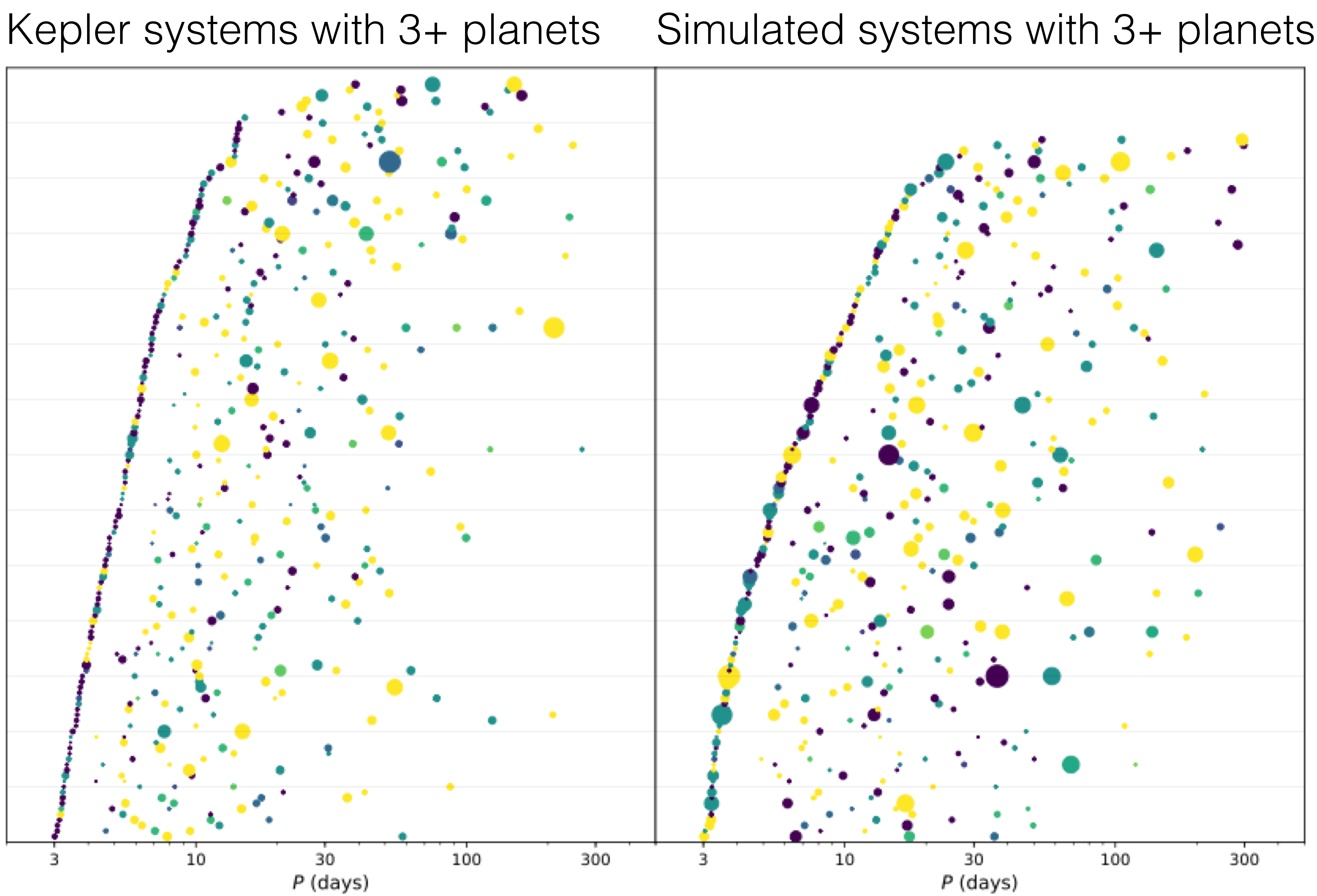
- Serves as a point of comparison for planet formation simulations
- Can be used to predict additional planets given already detected planets



Hsu et al. (2018, 2019)
He, Ford, & Ragozzine (2019)
He, Ford, & Ragozzine (in prep.)

Download or simulate model catalogs:
https://github.com/ExoJulia/SysSimExClusters

Questions?


Kepler systems with 3+ planets
Simulated systems with 3+ planets
3
10
30
100
300
P (days)
3
10
30
100
300
P (days)

Extra slides

Framework for forward modeling planetary systems and the Kepler mission

$f(P, R)$ **Step 1:** Define a statistical model for the intrinsic distribution of exoplanetary systems.

Step 2: Generate an underlying population of exoplanetary systems (*physical catalog*) from a given model.

Draw stars from Kepler catalog ← Kepler DR 25 (uniform vetting) Gaia DR 2 (updated radii, and b_p-r_p colors)

Populate each system with planets ← Draw periods, radii, masses

Assign an orbit to each planet ← Allow for two populations: Dynamically “hot” and “cold”

Framework for forward modeling planetary systems and the Kepler mission

$f(P, R)$ **Step 1:** Define a statistical model for the intrinsic distribution of exoplanetary systems.

Step 2: Generate an underlying population of exoplanetary systems (*physical catalog*) from a given model.

Step 3: Generate an observed population of exoplanetary systems (*observed catalog*) from the *physical catalog*.

Calculate which planets transit

Add a transit noise and detection efficiency model ← Christiansen (2017), Hsu et al. (2018, 2019)

Compute observed properties of planets ← Multiplicity, periods, depths, durations

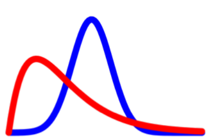
Framework for forward modeling planetary systems and the Kepler mission

$f(P, R)$

Step 1: Define a statistical model for the intrinsic distribution of exoplanetary systems.

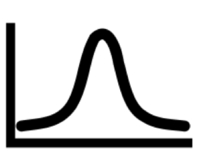
Step 2: Generate an underlying population of exoplanetary systems (*physical catalog*) from a given model.

Step 3: Generate an observed population of exoplanetary systems (*observed catalog*) from the *physical catalog*.



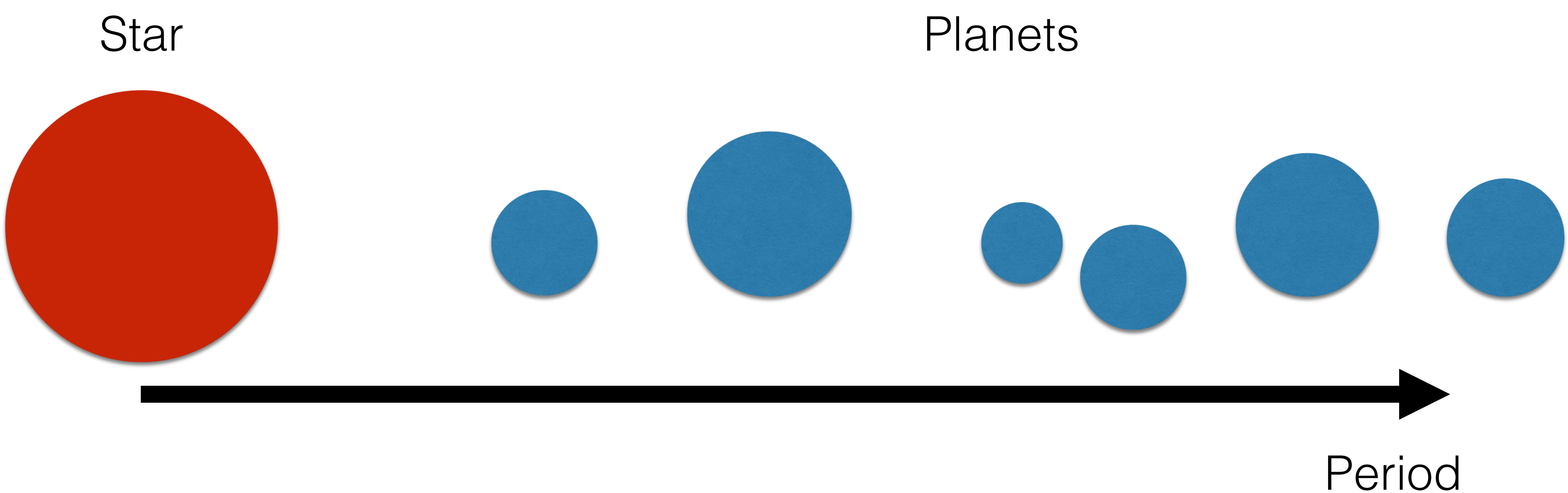
Step 4: Compare the simulated *observed catalog* with the Kepler data using a distance function.

Step 5: Optimize the distance function to find the best-fit model parameters.



Step 6: Explore the posterior distribution of model parameters using a Gaussian Process (GP) model.

Model 1: planets are drawn independently in period and size



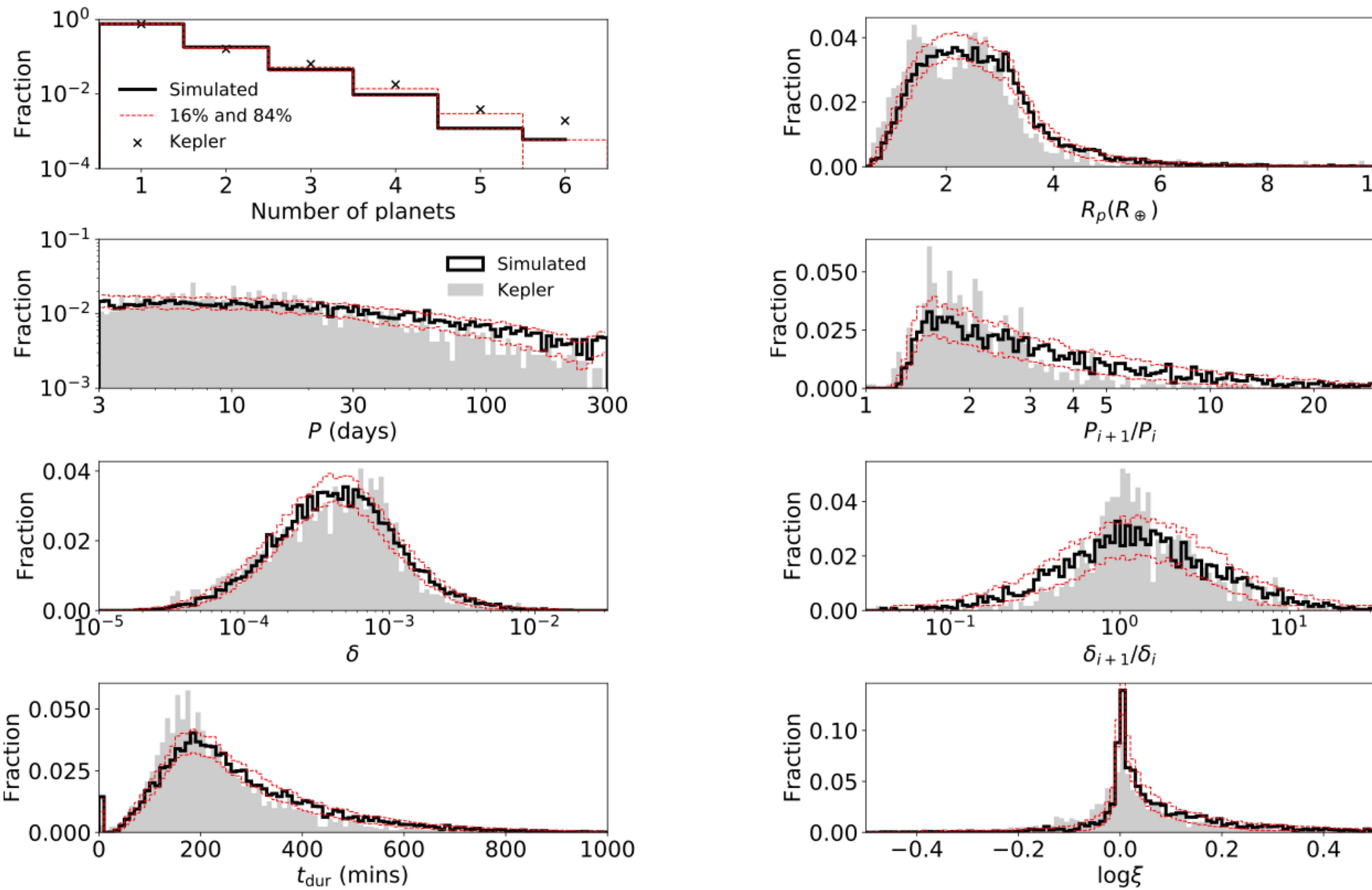
$$N_p \sim \text{Poisson}(\lambda_p)$$

Planets are drawn as a Poisson point process with independent periods and sizes

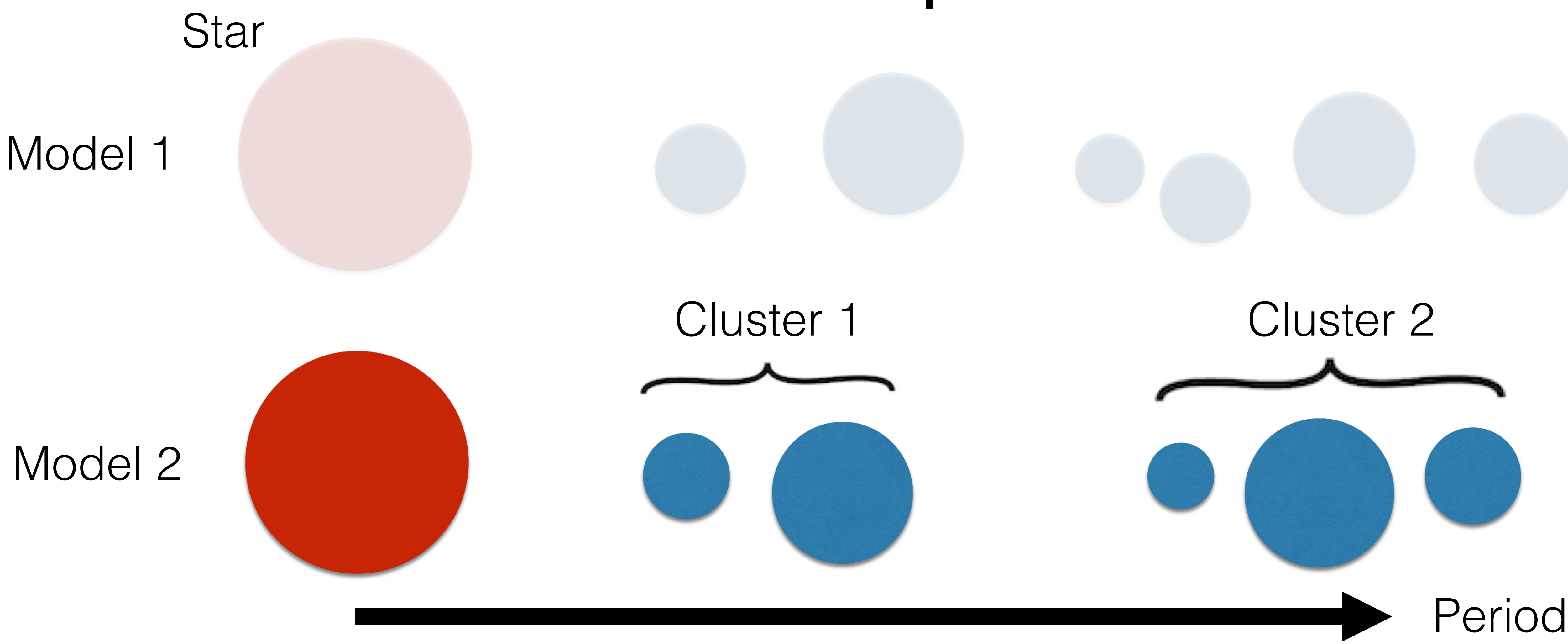
$$P \sim f(P) = C_1 P^{\alpha_P}$$

$$R_p \sim f(R_p) \propto \begin{cases} {R_p}^{\alpha_{R1}}, & R_{p,\mathrm{min}} \leq R_p \leq R_{p,\mathrm{break}} \\ {R_p}^{\alpha_{R2}}, & R_{p,\mathrm{break}} < R_p \leq R_{p,\mathrm{max}} \end{cases}$$

Models assuming independent planets fail to reproduce the observed population



Model 2: planets are clustered in periods

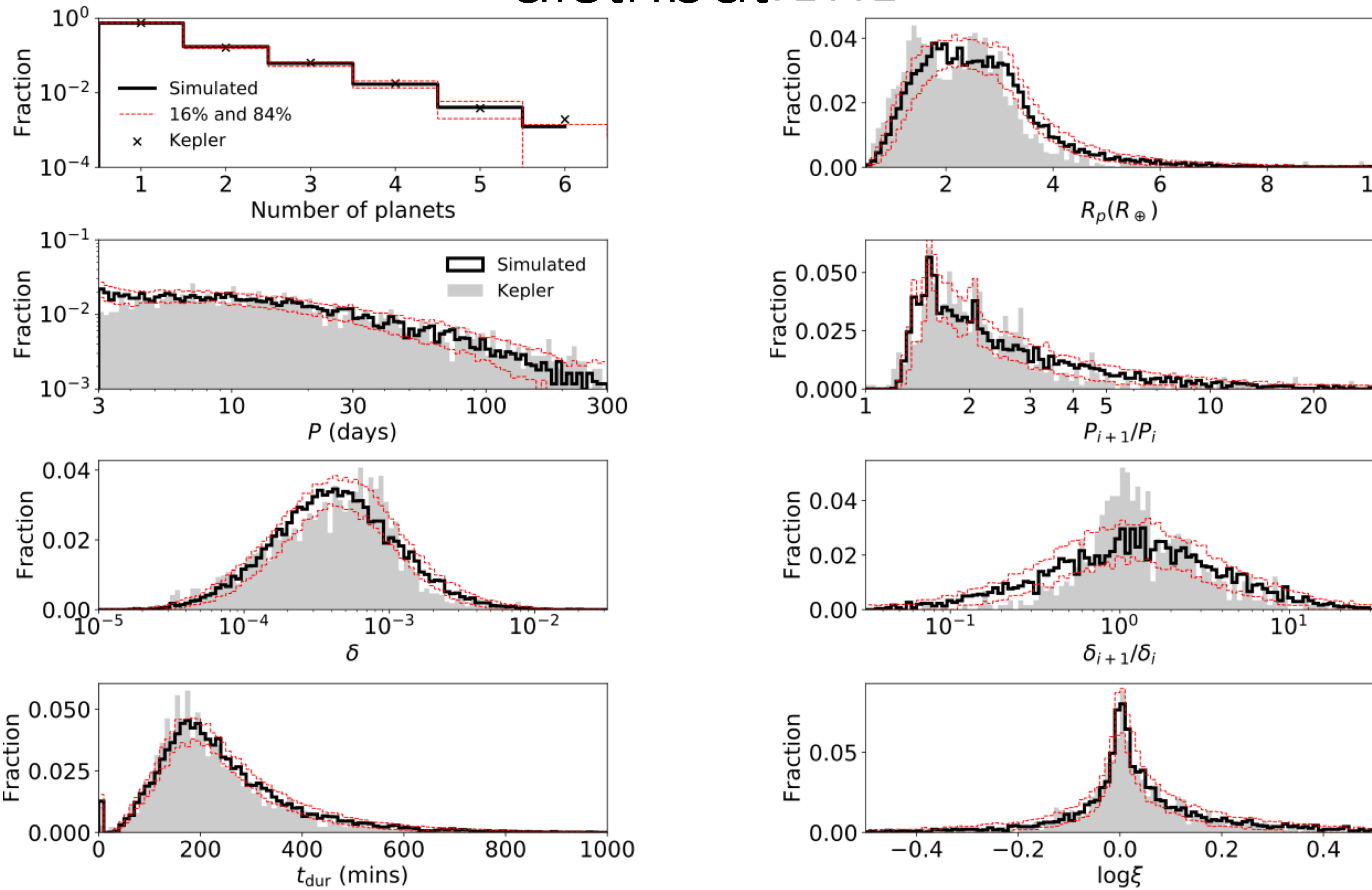


Planets are drawn as a clustered point process better reproduce the observed period ratios

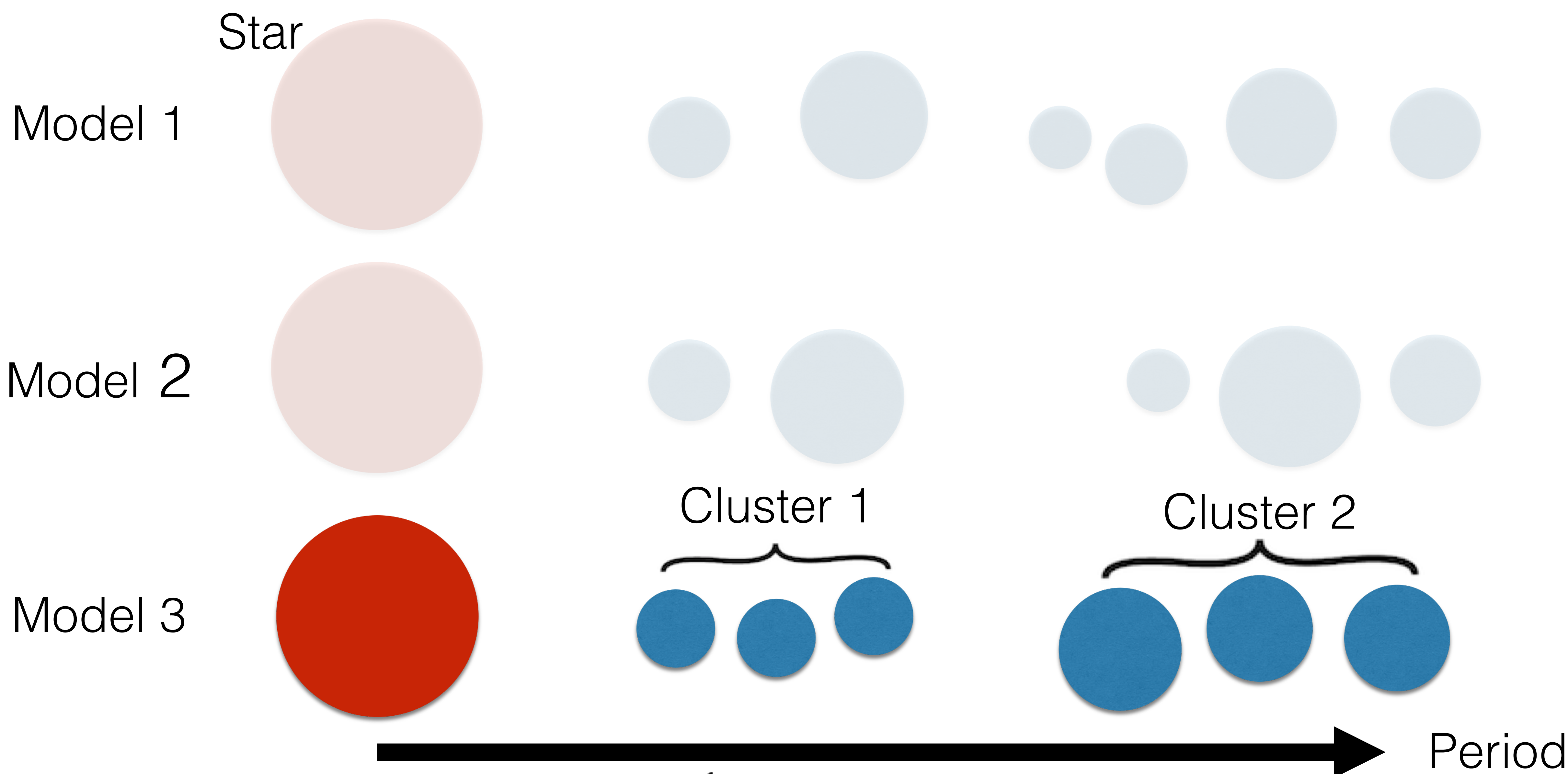
$$N_c \sim \mathrm{Poisson}(\lambda_c)$$
$$N_p \sim \mathrm{ZTP}(\lambda_p)$$

$$P_c \sim f(P_c) = C_1 P_c^{\alpha_P}$$
$$P' \sim \mathrm{Lognormal}(0, N_p \sigma_p)$$
$$P = P' P_c$$

The clustered periods model fits the multiplicity, period, and period ratio distributions



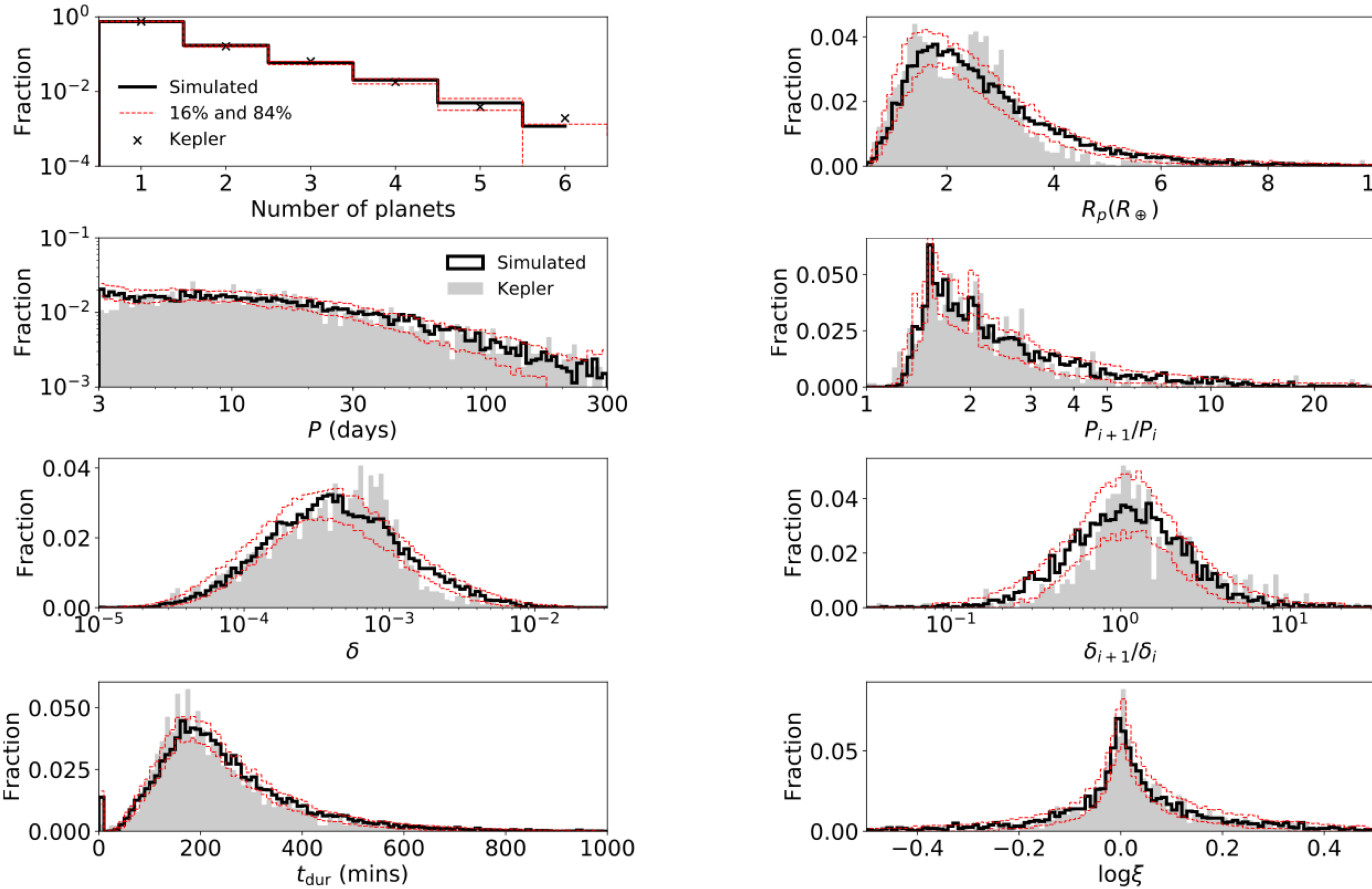
Model 3: planets are clustered in periods and sizes



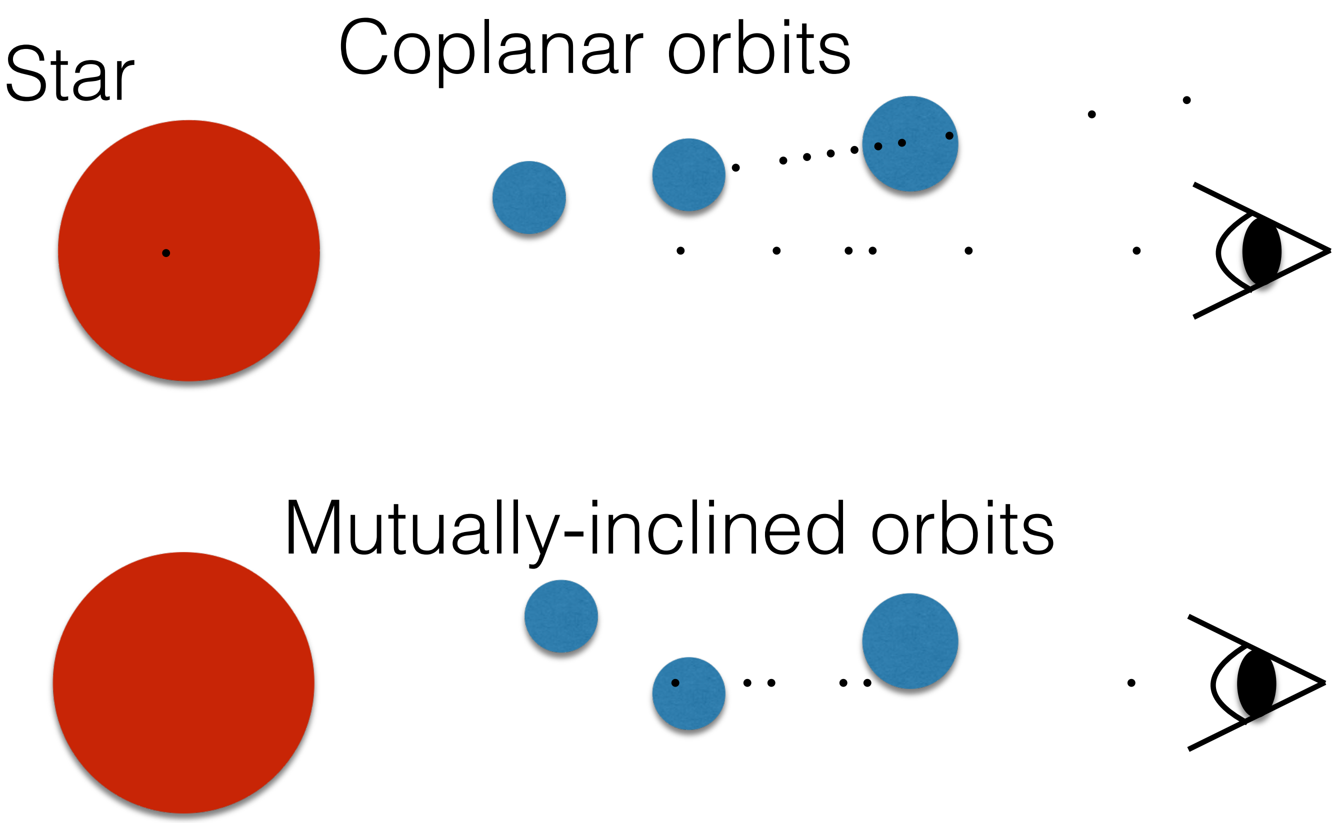
$$R_{p,c} \sim f(R_{p,c}) = \begin{cases} C_3 {R_{p,c}}^{\alpha_{R1}}, R_{p,c} \leq R_{p,\mathrm{break}} \\ C_4 {R_{p,c}}^{\alpha_{R2}}, R_{p,c} > R_{p,\mathrm{break}} \end{cases}$$

$$R_p \sim \mathrm{Lognormal}(R_{p,c}, \sigma_R)$$

Our fully clustered model provides a significantly improved description of planetary systems!



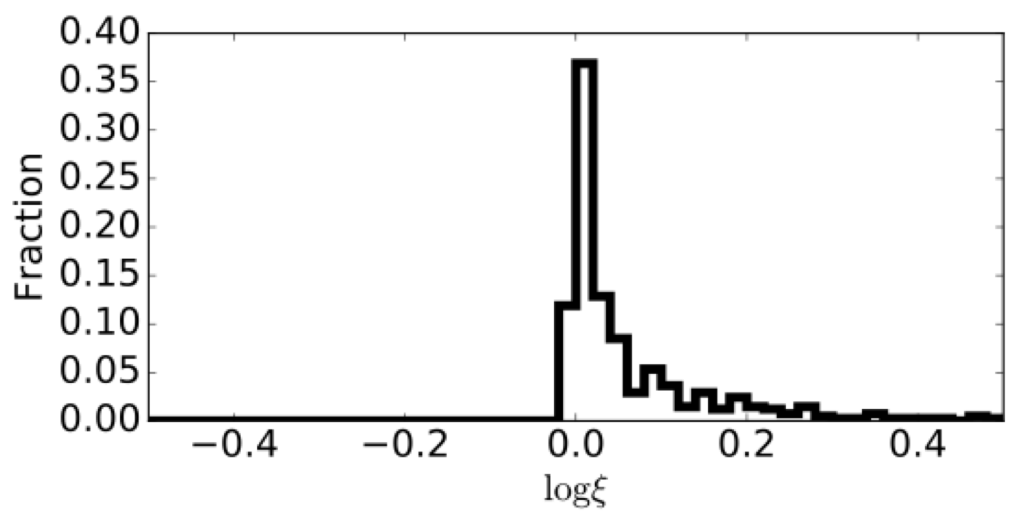
The transit duration ratio distribution encodes population orbital properties



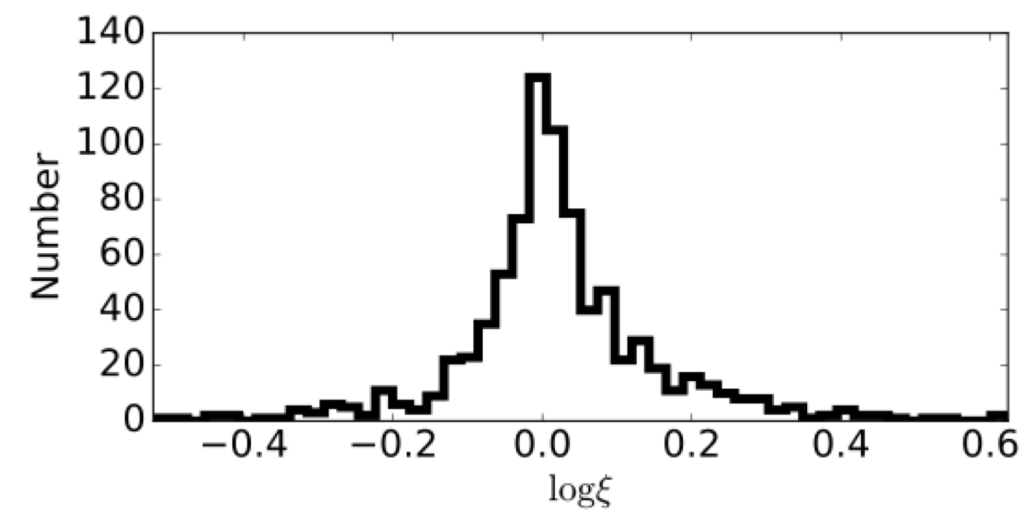
$$\xi = \left(\frac{t_{\mathrm{dur,in}}}{t_{\mathrm{dur,out}}}\right)\left(\frac{P_{\mathrm{out}}}{P_{\mathrm{in}}}\right)^{1/3}$$

- Mutual inclinations randomize impact parameters
- Orbital eccentricities randomize velocity during transit

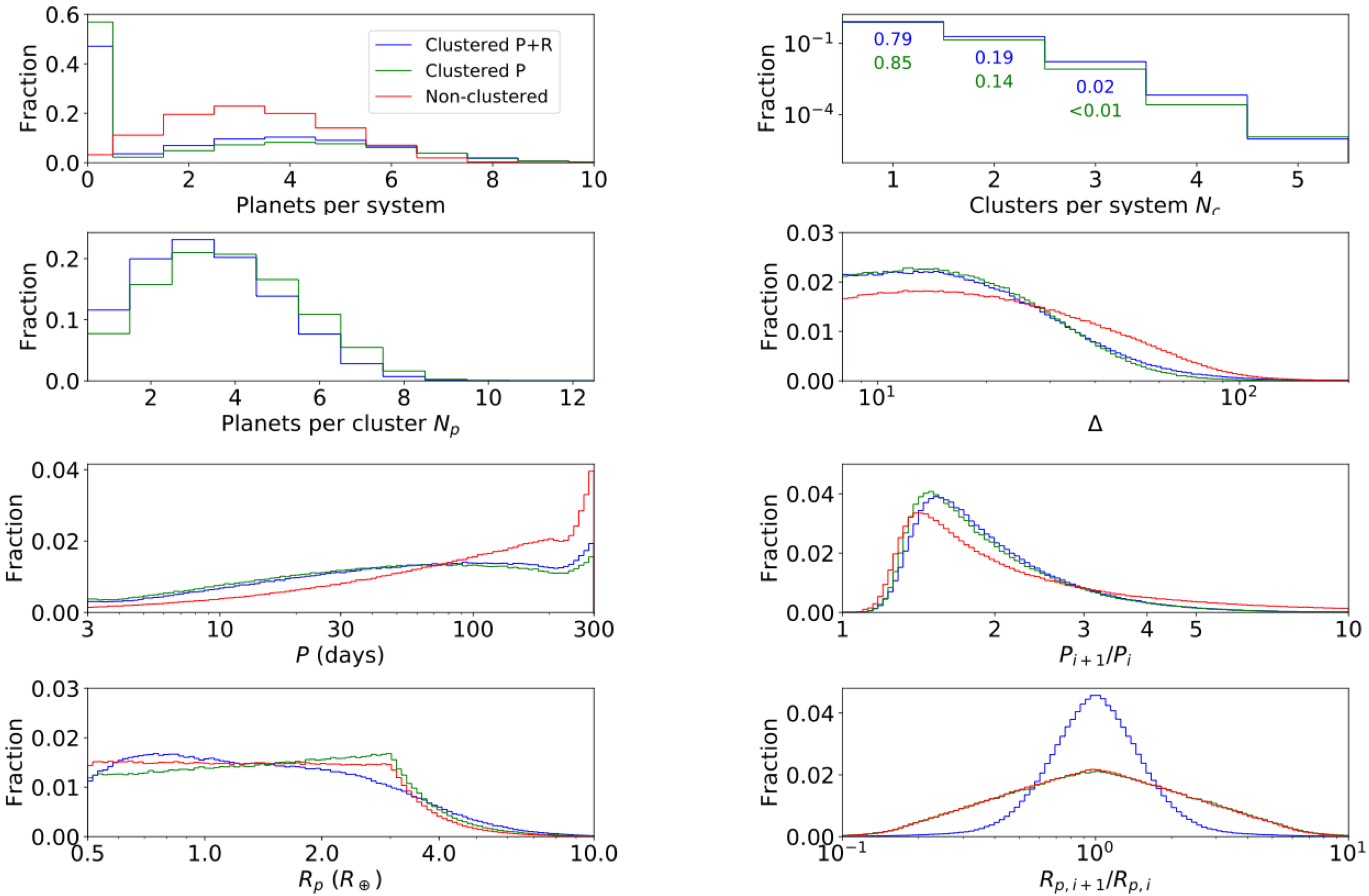
$i_m = 0$, $e = 0$ (coplanar + circular)



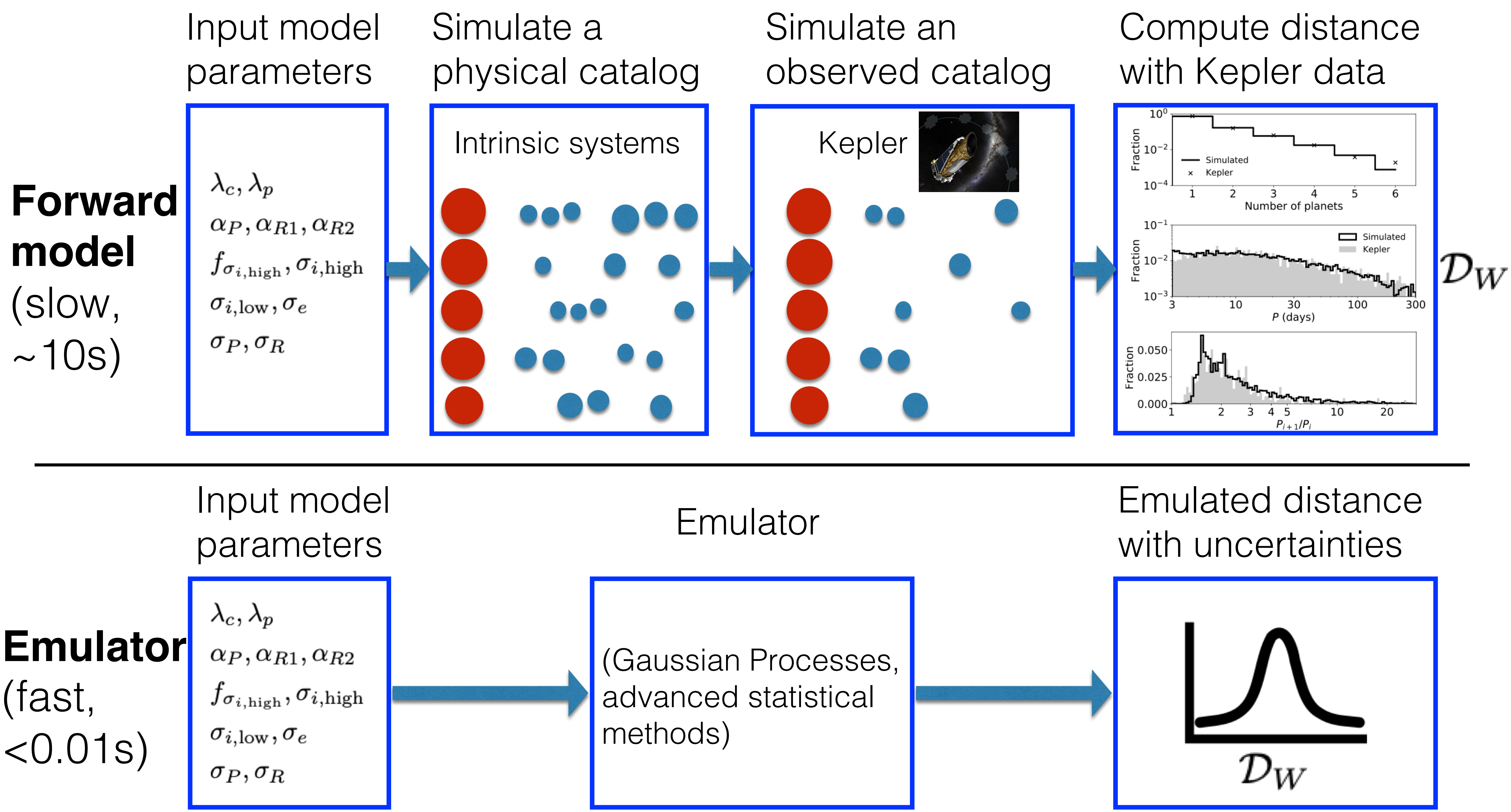
Kepler distribution



Differences between our models in their predictions for the underlying populations

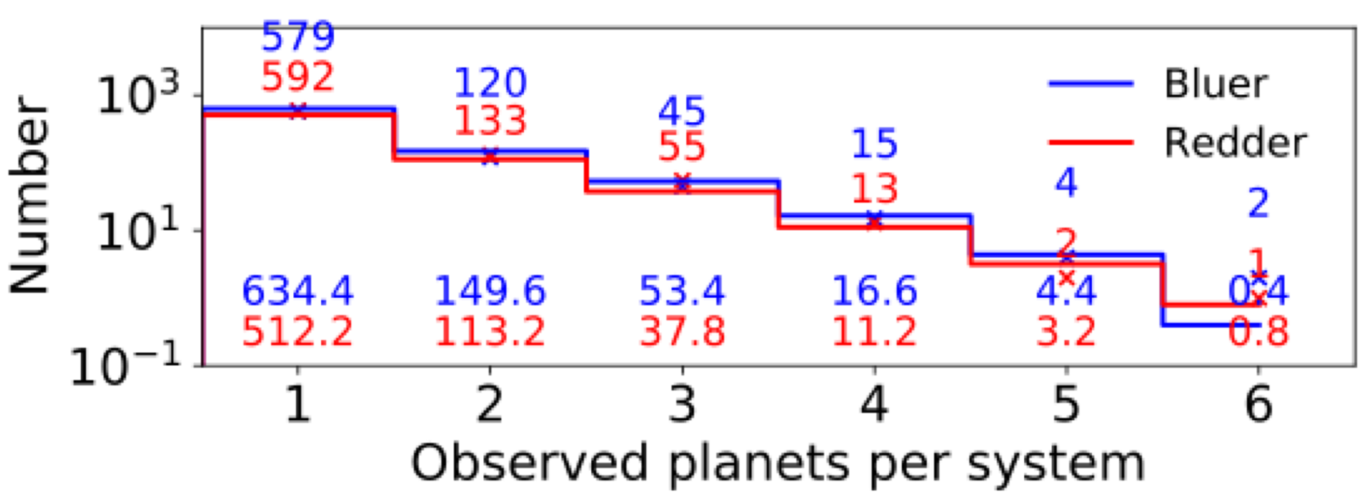


We train a fast emulator for our models to quantify model uncertainties



What about correlations between planetary systems and their host stars?

We split the stellar sample (79935 FGK stars) in half by Gaia b_p-r_p color
—> The exoplanet counts are also roughly split in half



Assuming the same distribution of planets around all stars produces more detected planets around the bluer stars!