Statistics of the Exoplanets

Andrew Cumming
University of California, Santa Cruz

SIX NEW PLANETS FROM THE KECK PRECISION VELOCITY SURVEY[1]

— ApJ June 20, 2000

Astronomers Find Evidence for Eight Objects Orbiting Distant Stars

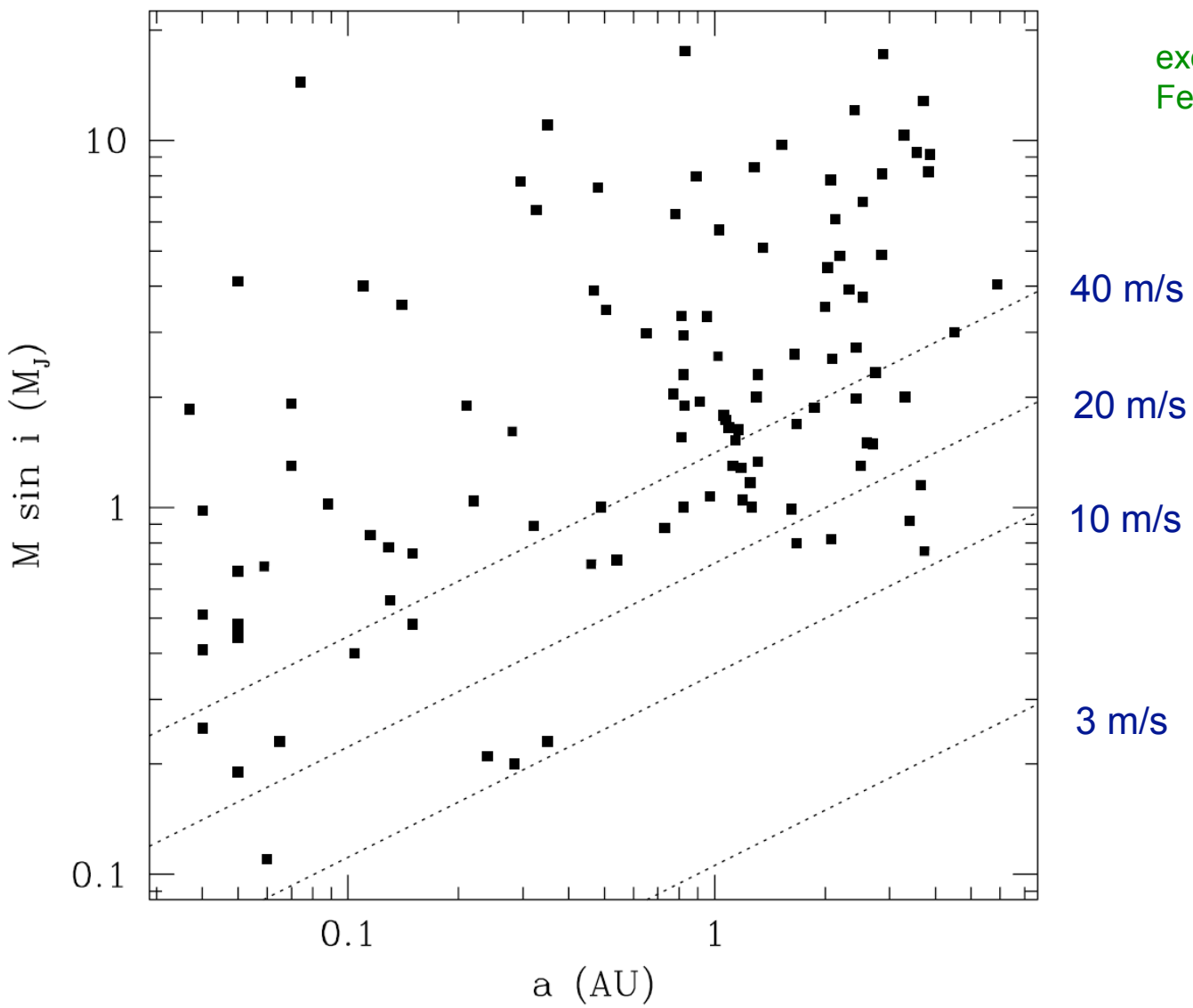
— New York Times
May 10, 2000

SUB-SATURN PLANETARY CANDIDATES OF HD 16141 AND HD 46375[1]

— ApJ June 10, 2000

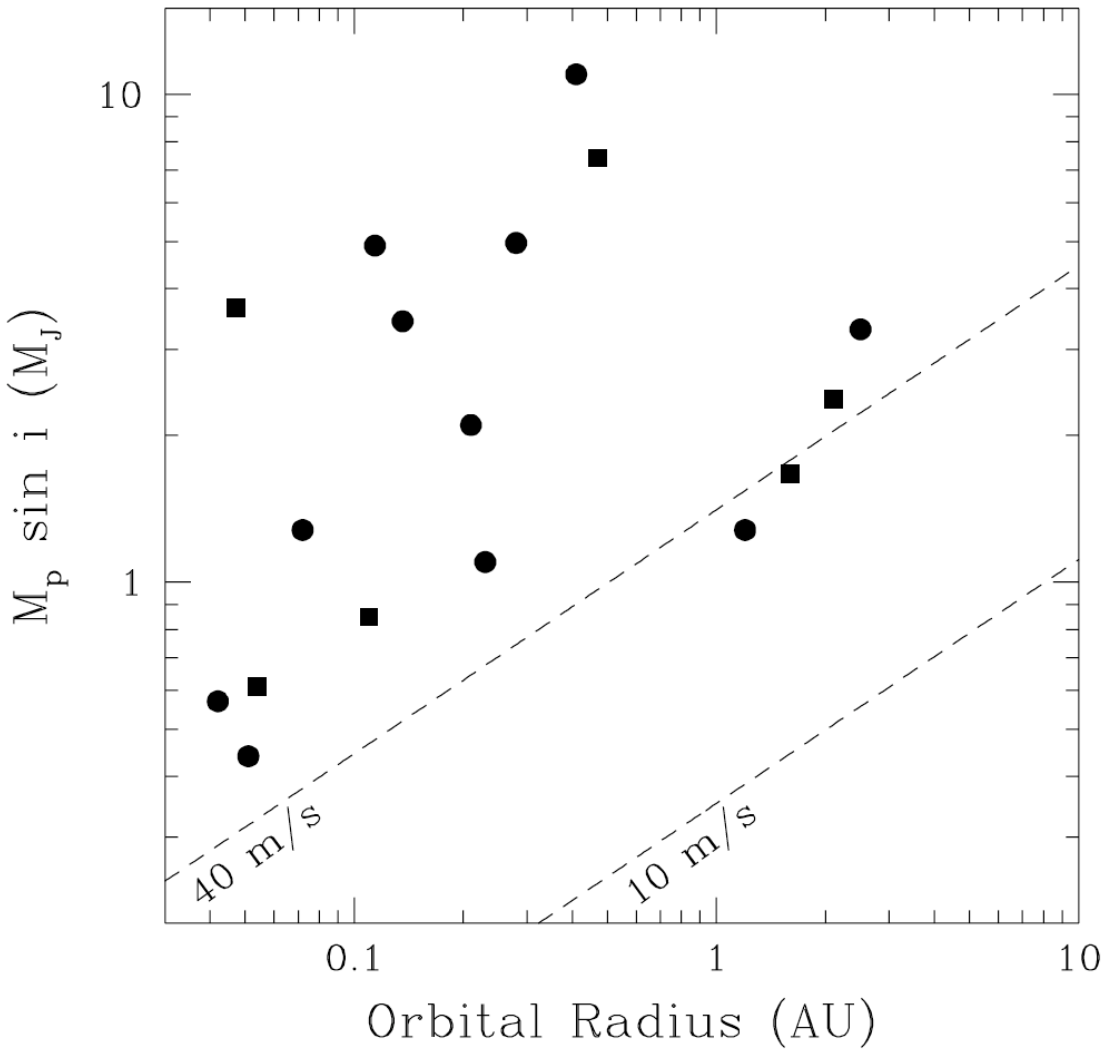
Found: 2 Planetary Systems. Result: Astronomers Stunned.

— New York Times
Jan 10th 2001

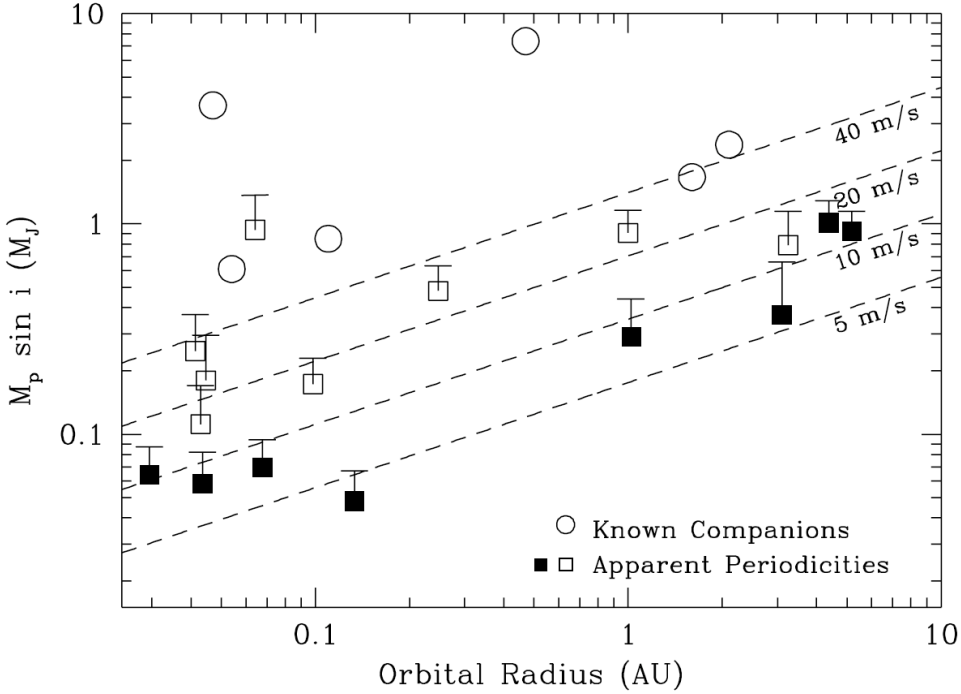


$$K = 28 \text{ m s}^{-1} \left(\frac{M \sin i}{M_J}\right) \left(\frac{a}{1 \text{ AU}}\right)^{-1/2} \left(\frac{M_\star}{M_\odot}\right)^{-2/3}$$

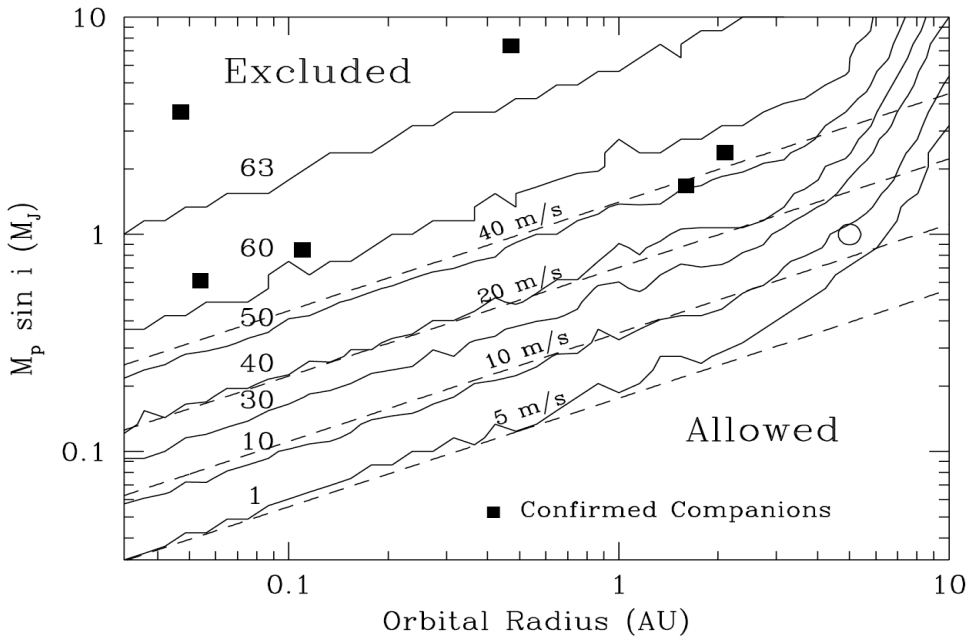
Known planets at the beginning of 1999



(CMB99 - squares indicate Lick planets)



Detections

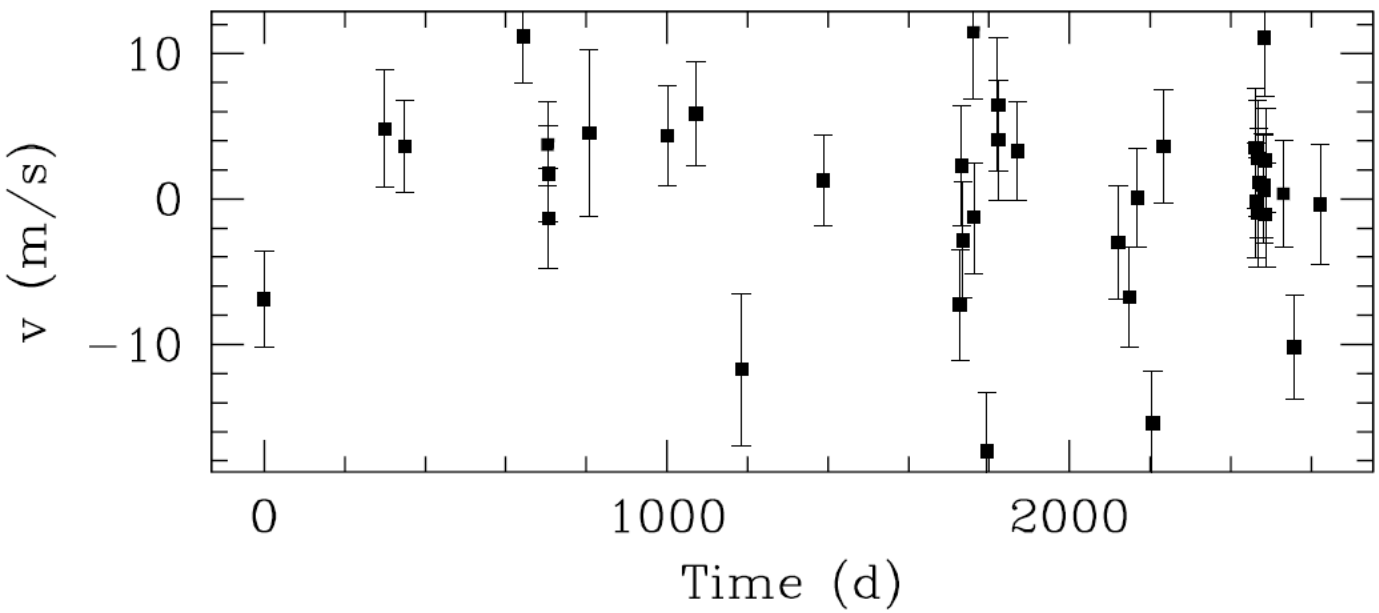


Upper limits

Cumming, Marcy, & Butler (1999)

This talk

- detecting periodicities in unevenly-sampled data
- intrinsic stellar “jitter”
- detections and candidates from the Keck survey
- the mass-period distribution


10
0
−10
v (m/s)
0
1000
2000
Time (d)

Search for Periodicities in Unevenly-sampled Data

basic idea : χ^2 fitting

ask whether χ^2 is improved "enough" when a periodicity is included in the model

$$z = \frac{\Delta\chi^2}{\chi_0^2}$$

Lomb-Scargle periodogram

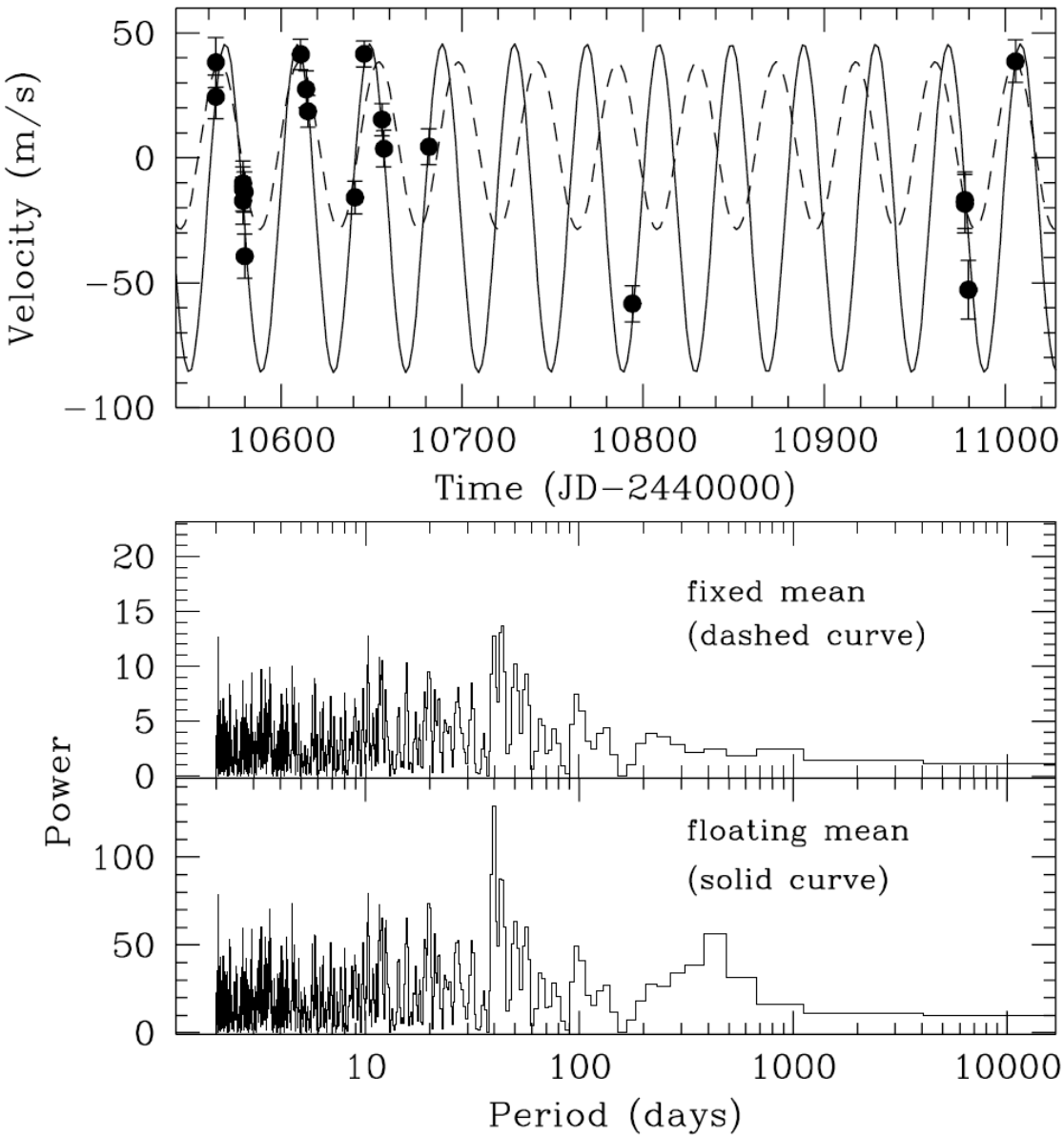
$$f_j(t_j) = A \cos \omega(t_j - \tau) + B \sin \omega(t_j - \tau)$$

$$\hat{z}(\omega) = \frac{[\sum w_j v_j \cos \omega(t_j - \tau)]^2}{\sum w_j \cos^2 \omega(t_j - \tau)} + \frac{[\sum w_j v_j \sin \omega(t_j - \tau)]^2}{\sum w_j \sin^2 \omega(t_j - \tau)}$$

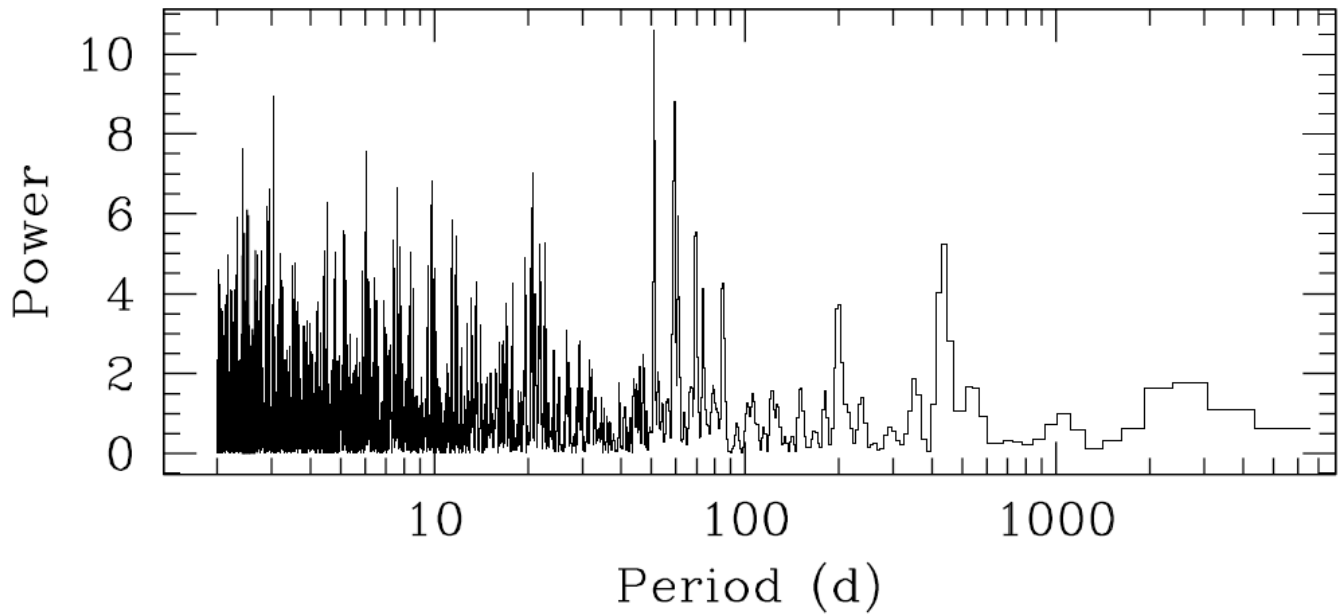
Lomb (1976); Scargle (1982); Horne & Baliunas (1986)

need to include a constant term for sparsely-sampled data
or for periods > duration of observations

Black & Scargle (1989); Walker et al. (1995); CMB99



data for HR 5968
(Noyes et al. 1987)



Simple analytic distribution for Gaussian noise

e.g. if σ is known $\mathrm{Prob}(z > z_0) = \exp(-z_0)$

If you look at M frequencies, the “false alarm probability” is

$$1 - (1 - \exp(-z_0))^M \approx M \exp(-z_0)$$

Uneven sampling => need Monte Carlo simulations to determine M

aliasing problems much less severe because better phase coverage

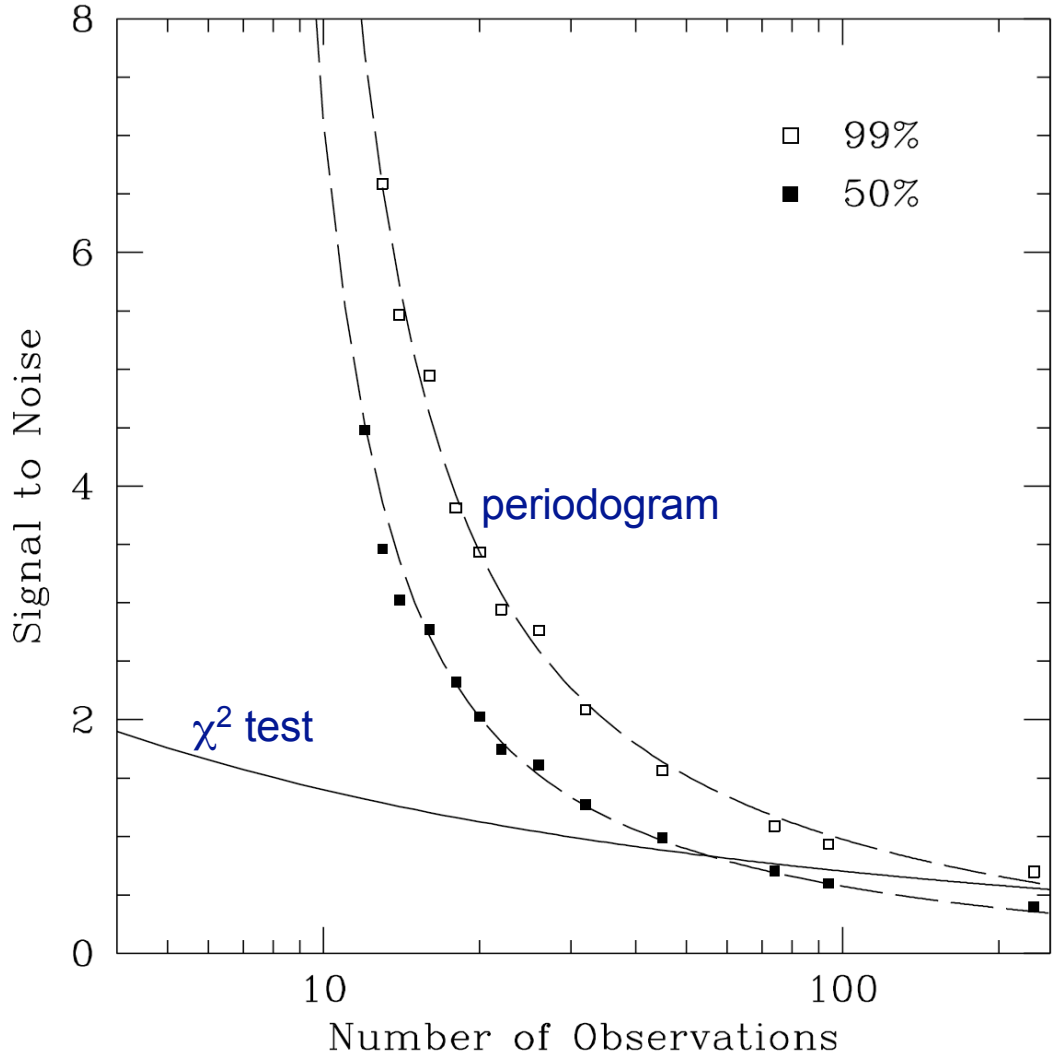
Signal to noise needed to detect a signal 50% of the time

$$\frac{K}{\sqrt{2}\sigma} = \left[\left(\frac{M}{F}\right)^{2/(n-3)} - 1\right]^{1/2}$$

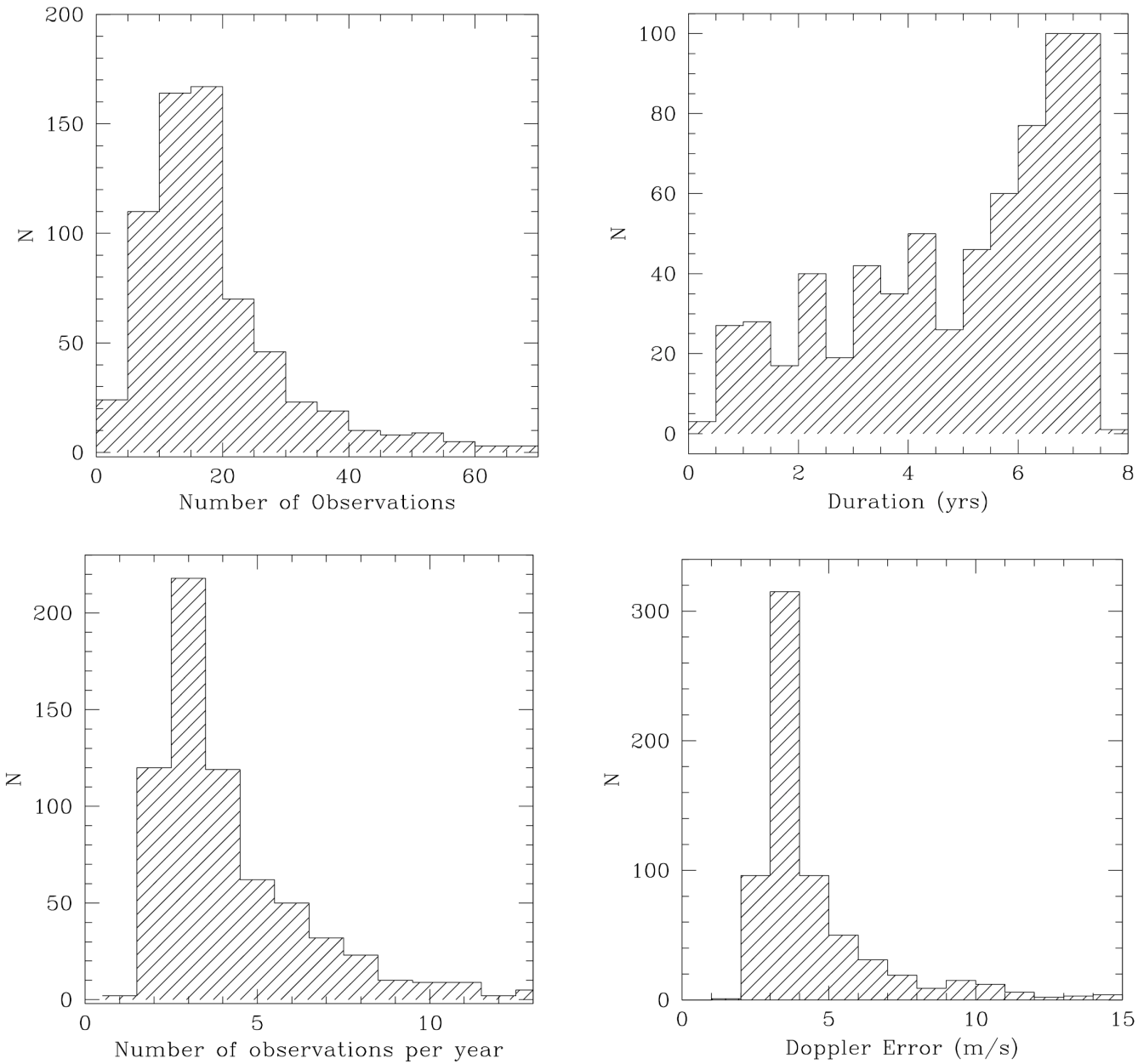
$$\approx \left[\frac{2\ln(M/F)}{n}\right]^{1/2}$$

$$\approx \frac{5}{\sqrt{n}}$$

- K>1-2σ variability detected even for small n
- hard to characterize the orbit for n<10
- detect S/N~2-4 when n=10-20
- require n>50 to get S/N<1

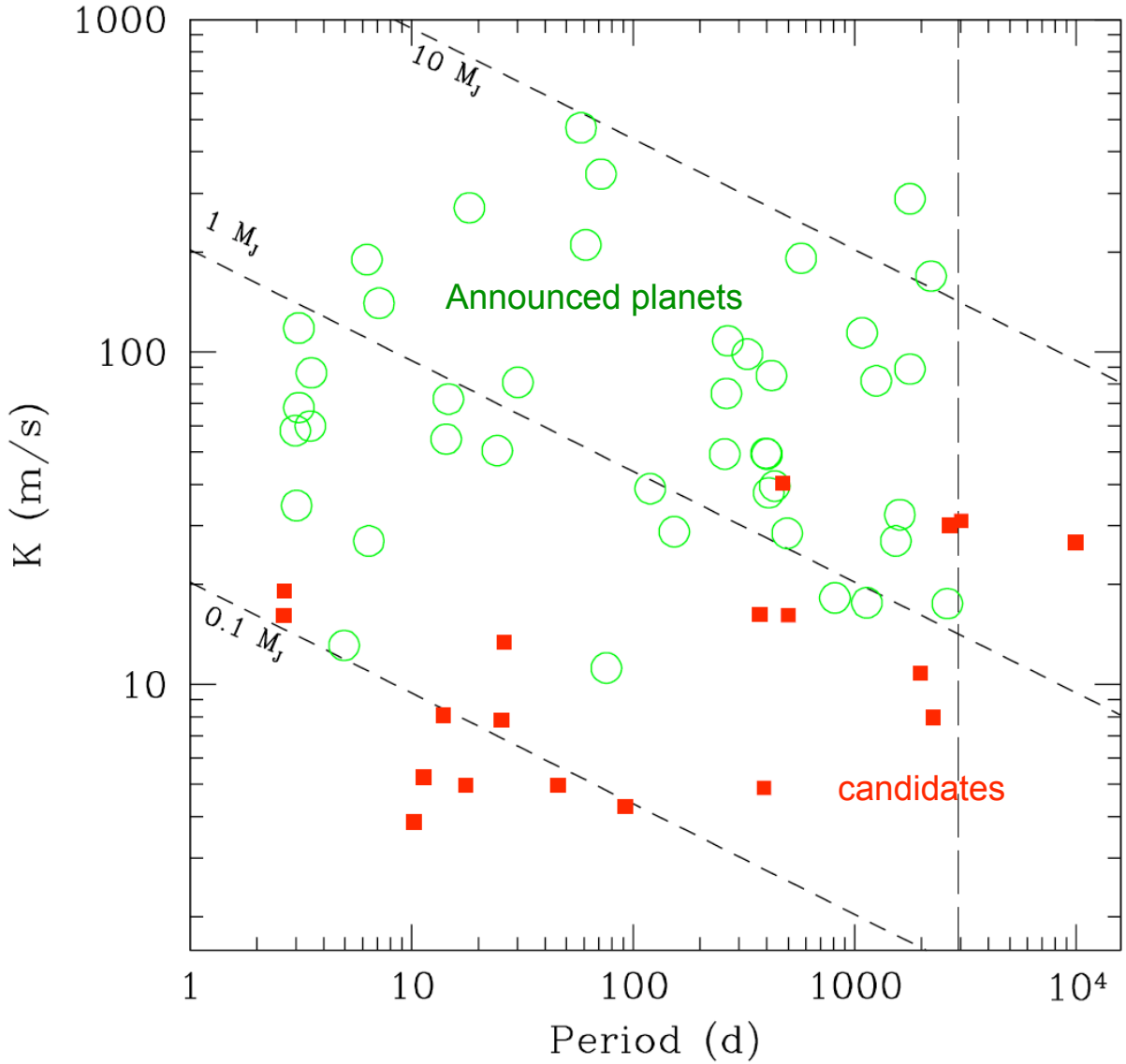


CMBV (2003)

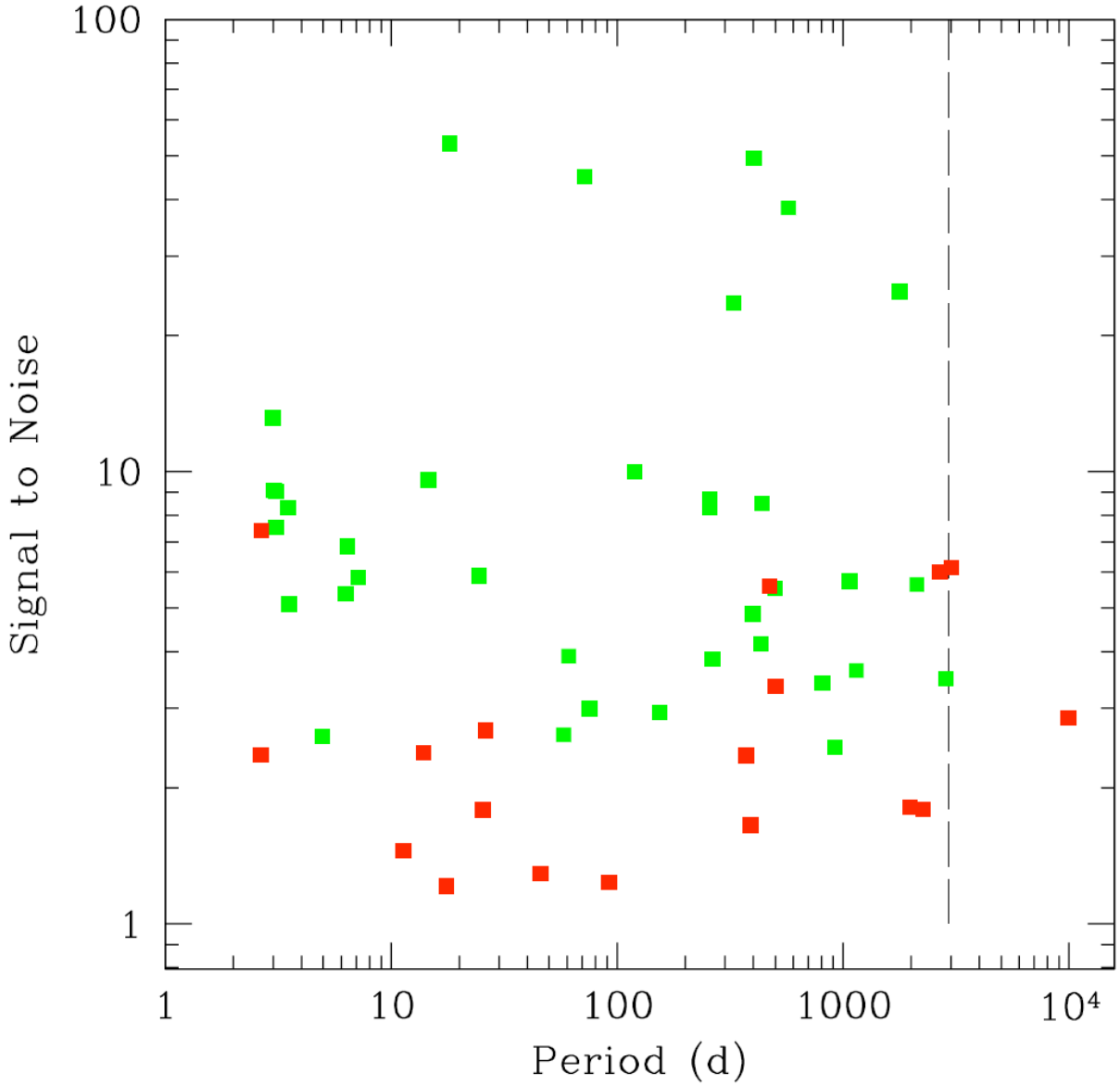

N
Number of Observations
Duration (yrs)
Number of observations per year
Doppler Error (m/s)

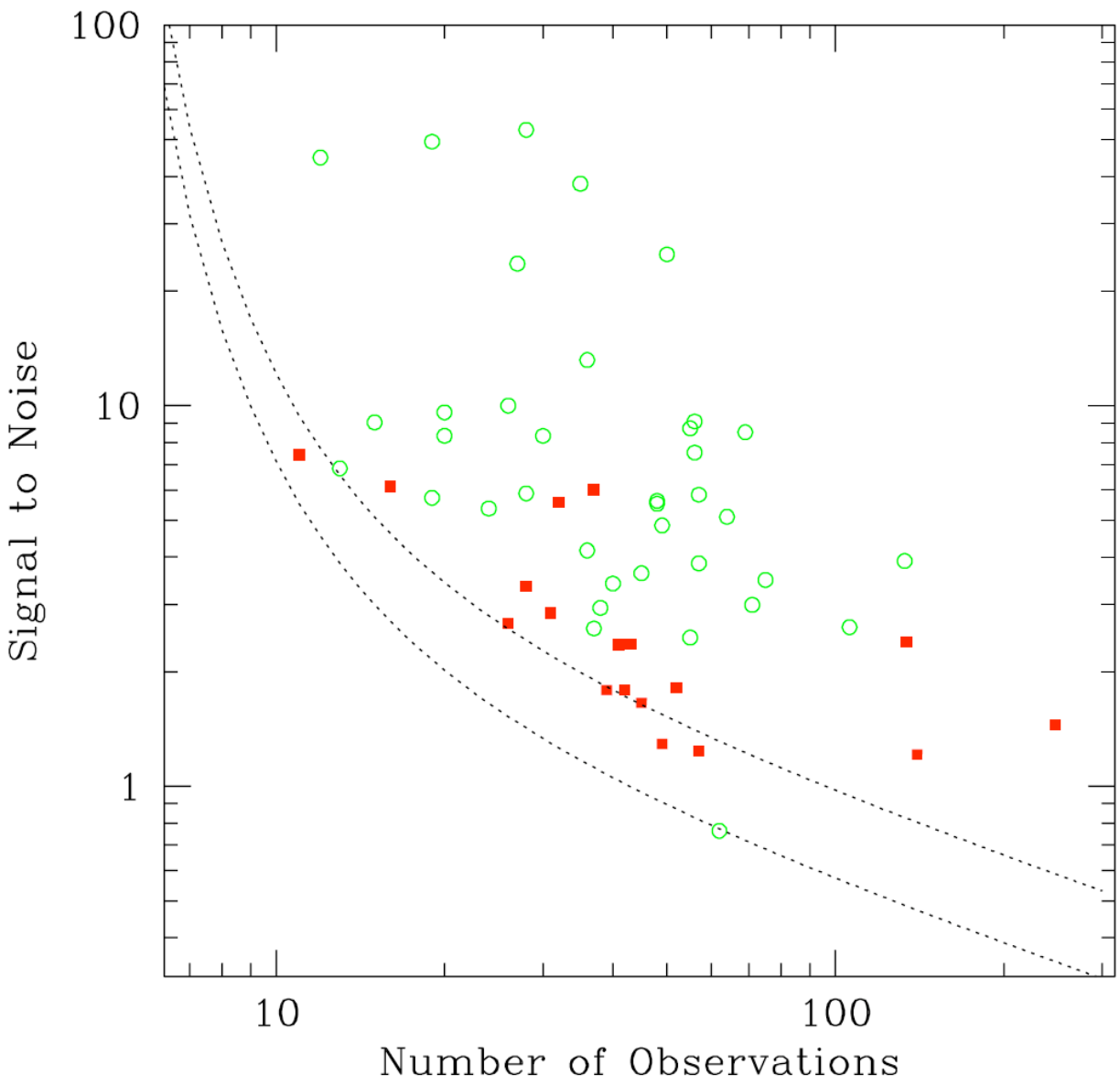
Recipe

- for each star, use a periodogram to get an initial guess for (P,K)
- use this as input for a Keplerian fit
- calculate the F-statistic
- find the FAP with Monte Carlo


1000
100
10
K (m/s)
1
10
100
1000
10^4
Period (d)
10 M_J
1 M_J
0.1 M_J
Announced planets
candidates

CMBV04


100
10
1
Signal to Noise
1
10
100
1000
10⁴
Period (d)


100
10
1
Signal to Noise
10
100
Number of Observations

“Jitter”

Saar & Donahue (1997)

- rotational modulation of spots

$$\mathrm{v}\sin i = 5\ \mathrm{km\ s^{-1}} \left(\frac{P_{\mathrm{rot}}}{10\ \mathrm{days}}\right)^{-1}$$

$$A_s \approx f_s\ \mathrm{v}\sin i \approx 5\ \mathrm{m\ s^{-1}} \left(\frac{f_s}{10^{-3}}\right) \left(\frac{P_{\mathrm{rot}}}{10\ \mathrm{days}}\right)^{-1}$$

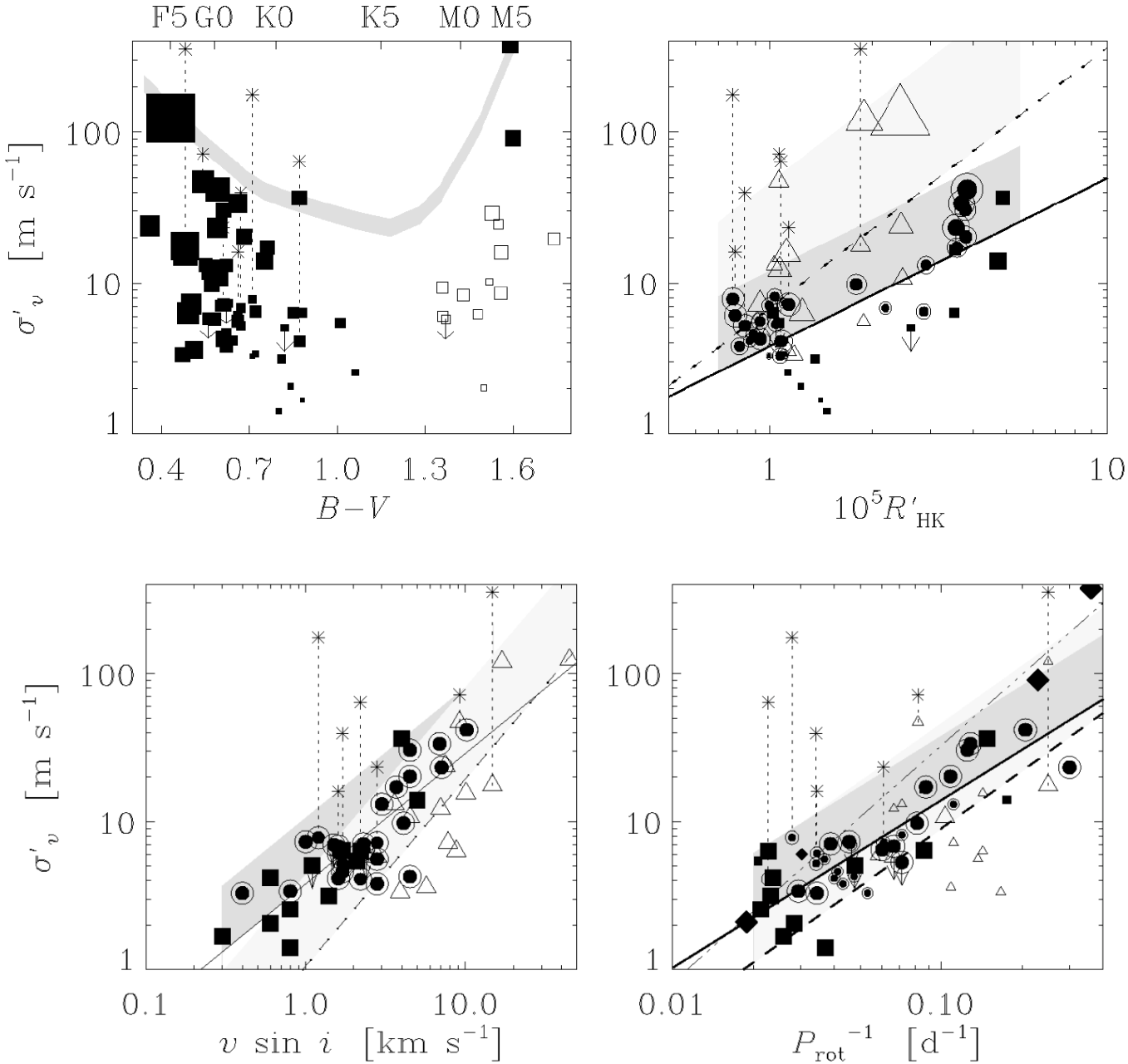
dominates in active G,K

- convective inhomogeneities

$$A_c \approx f_P\ \left(\mathrm{v}_{\mathrm{conv}} + \mathrm{v}\ \sin i\right)$$

dominates in F, quiet G

extra source of jitter needed for M dwarfs?

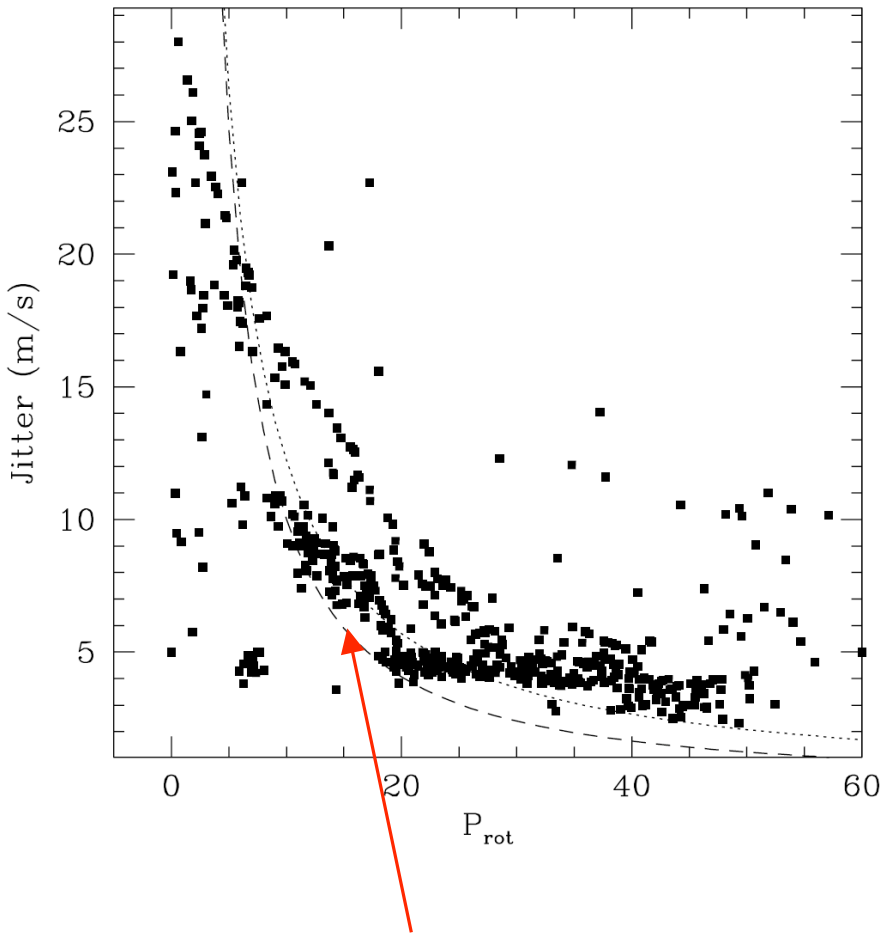


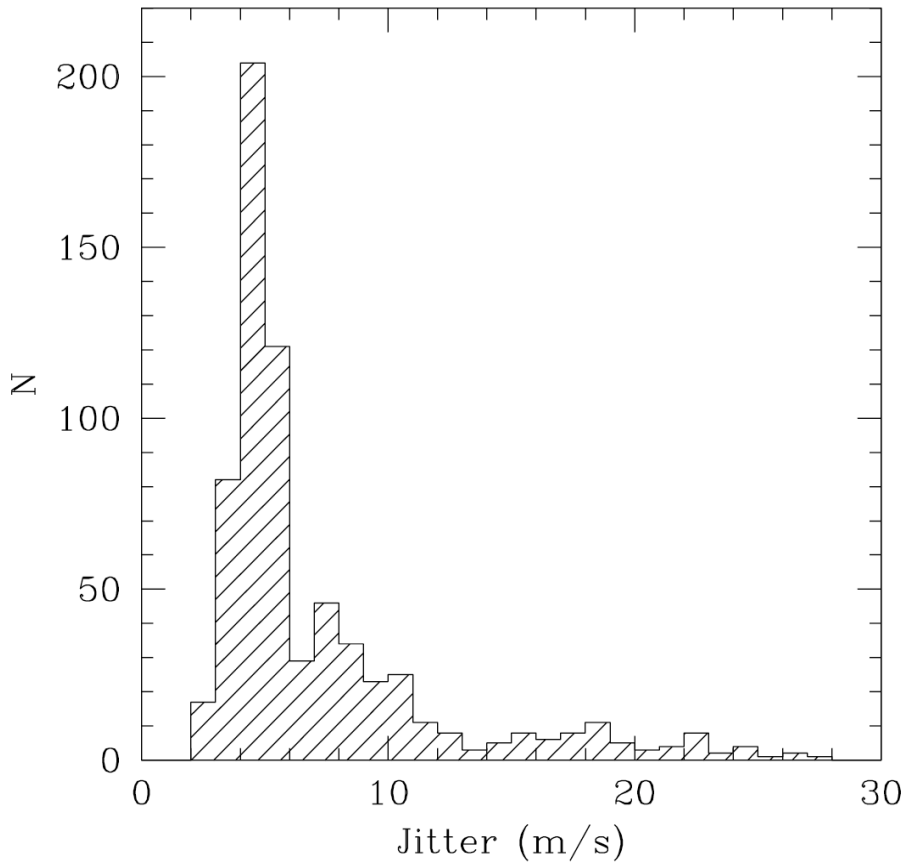
a rough fit is

$$\sigma_v \approx 10 \text{ m s}^{-1} \left(\frac{P_{\rm rot}}{12 \text{ days}} \right)^{-1.1}$$

Wright & Marcy (2004)

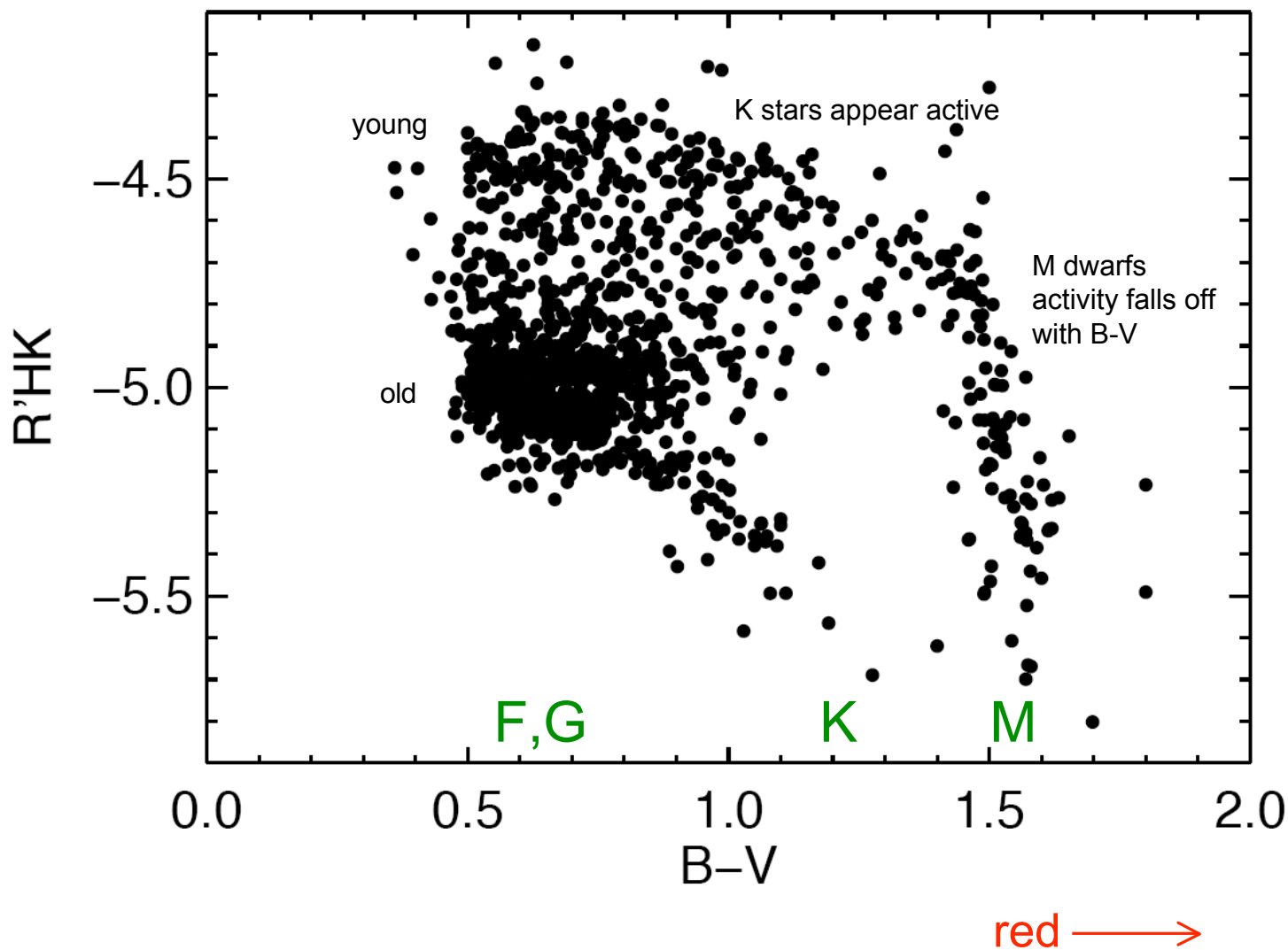
empirical calibration of jitter as a function of R'HK, MV, and B-V

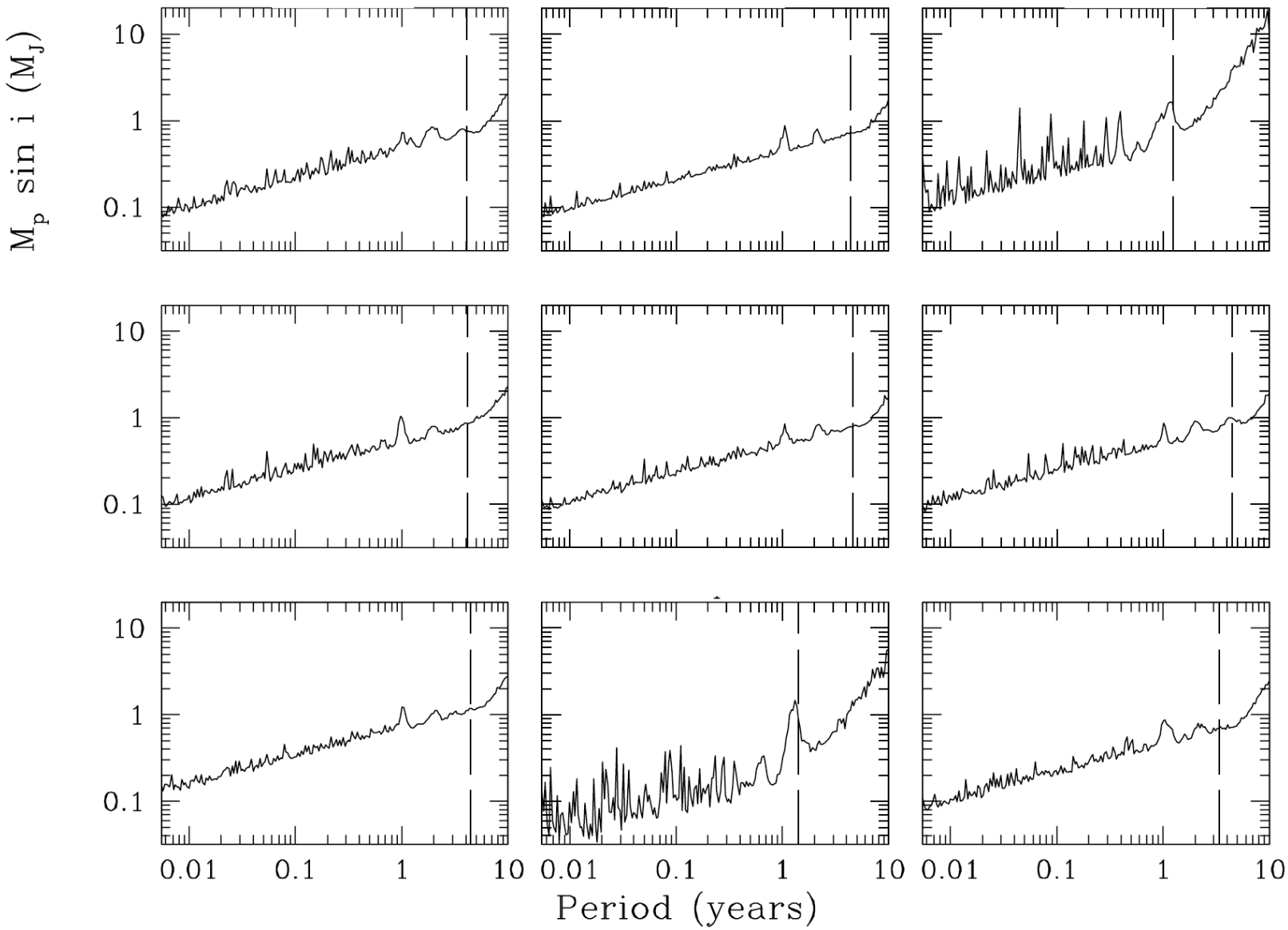




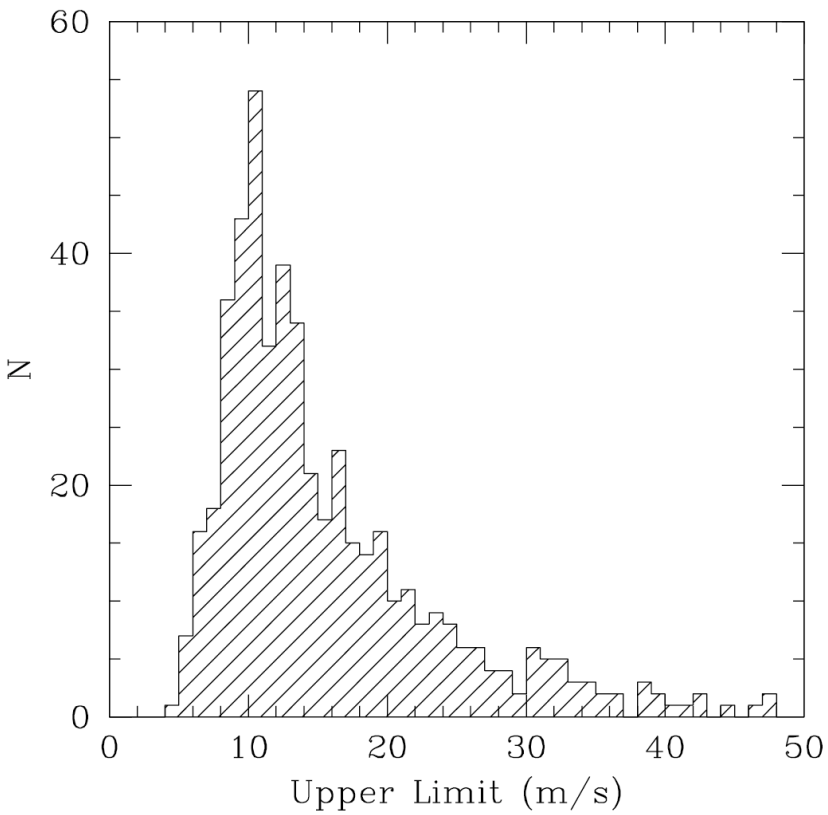
roughly jitter ~ 1/Prot

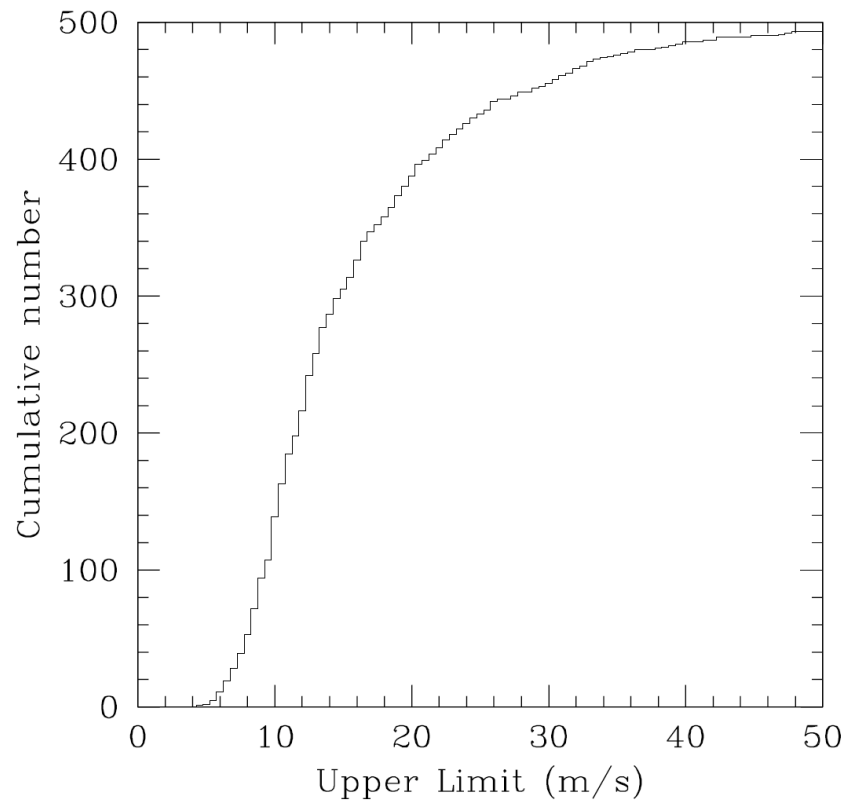
see also Santos et al (2000) (CORALIE)


young
K stars appear active
M dwarfs
activity falls off
with B-V
old
F,G
K
M
R'HK
−4.5
−5.0
−5.5
0.0
0.5
1.0
1.5
2.0
B−V
red

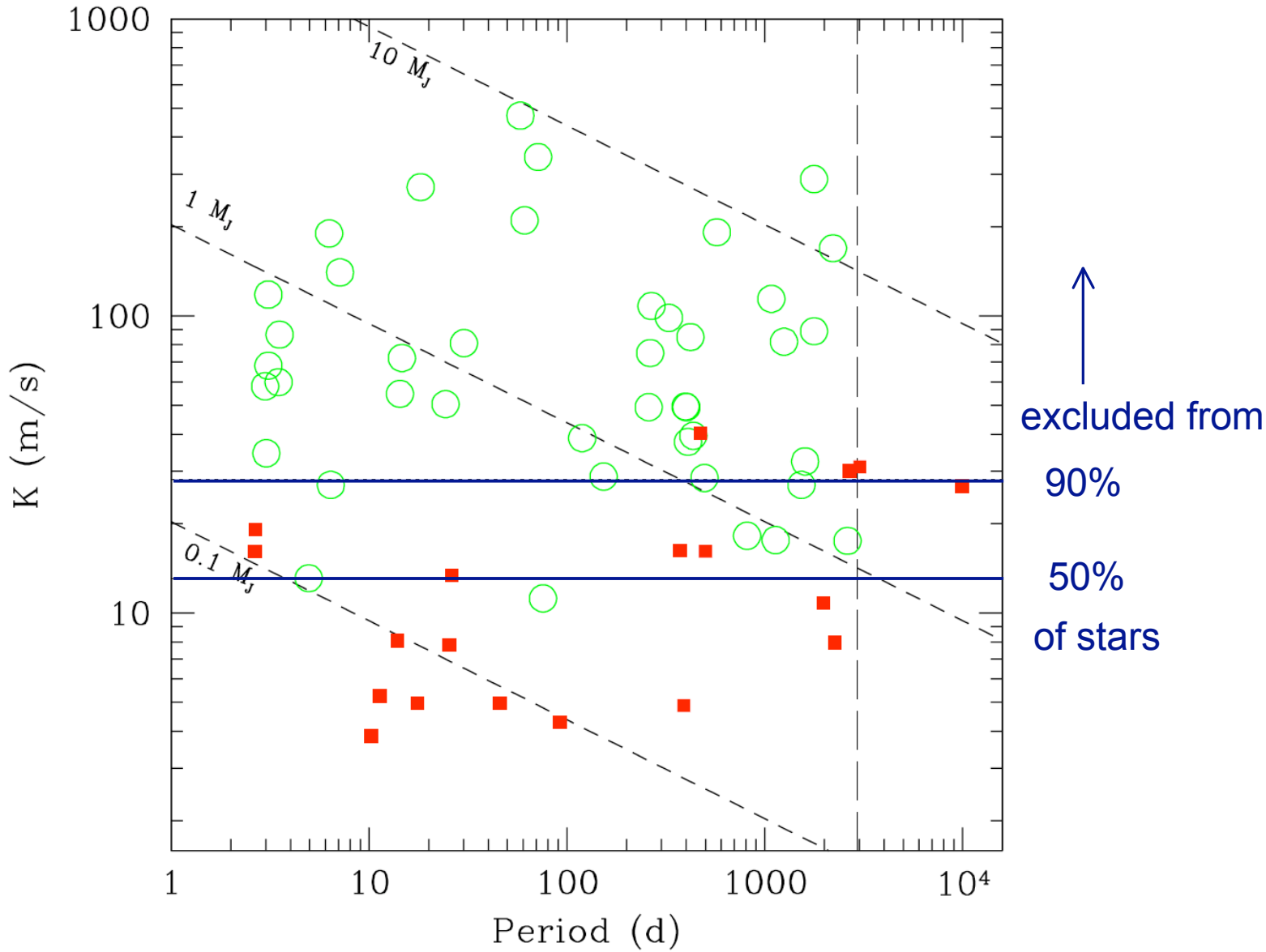

M_p sin i (M_J)
10
1
0.1
0.01
0.1
1
10
Period (years)

Upper limits





CMBV04


1000
100
10
K (m/s)
1
10
100
1000
10^4
Period (d)
10 M_J
1 M_J
0.1 M_J
excluded from
90%
50%
of stars

A Bayesian approach

basic probability

$$\mathrm{Prob}(M, P, \phi) = \prod_{k=1}^{N} \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left(-\frac{[v_k - f_k(M \sin i, P, \phi)]^2}{2\sigma_k^2}\right)$$

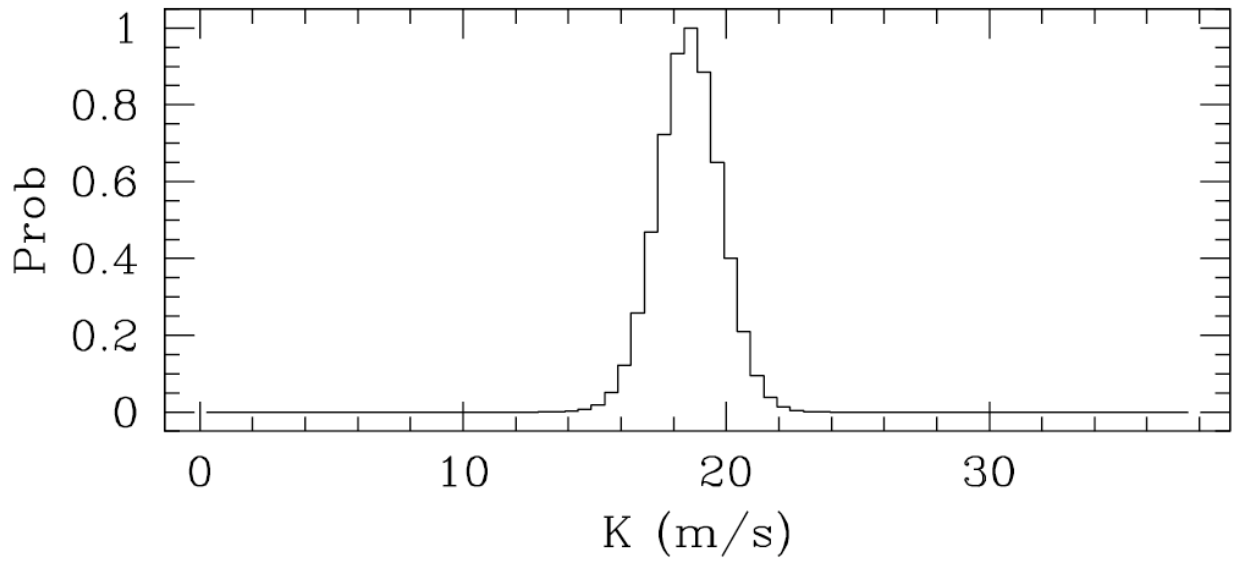
“integrate out” uninteresting parameters

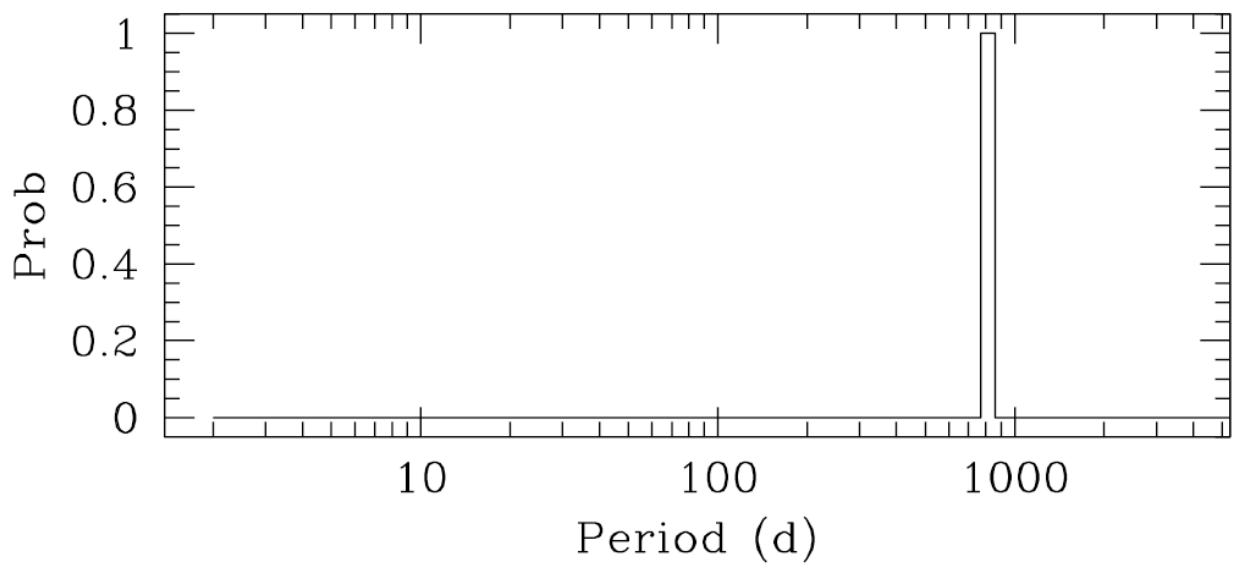
$$\longrightarrow \quad \begin{array}{l} \mathrm{Prob}(K) \\ \mathrm{Prob}(P) \\ \mathrm{Prob}(M, P) \end{array}$$

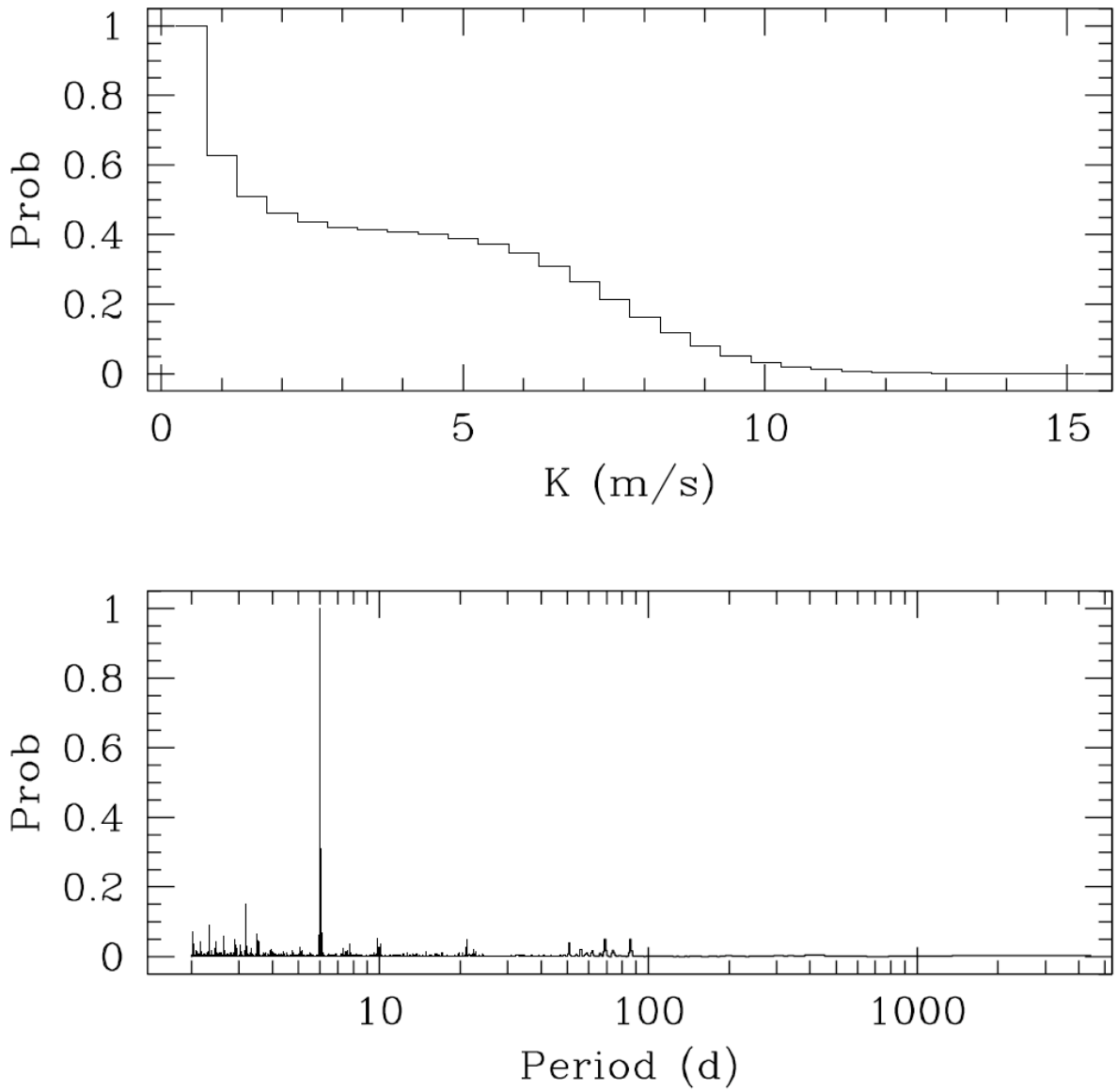
ask questions about the distribution

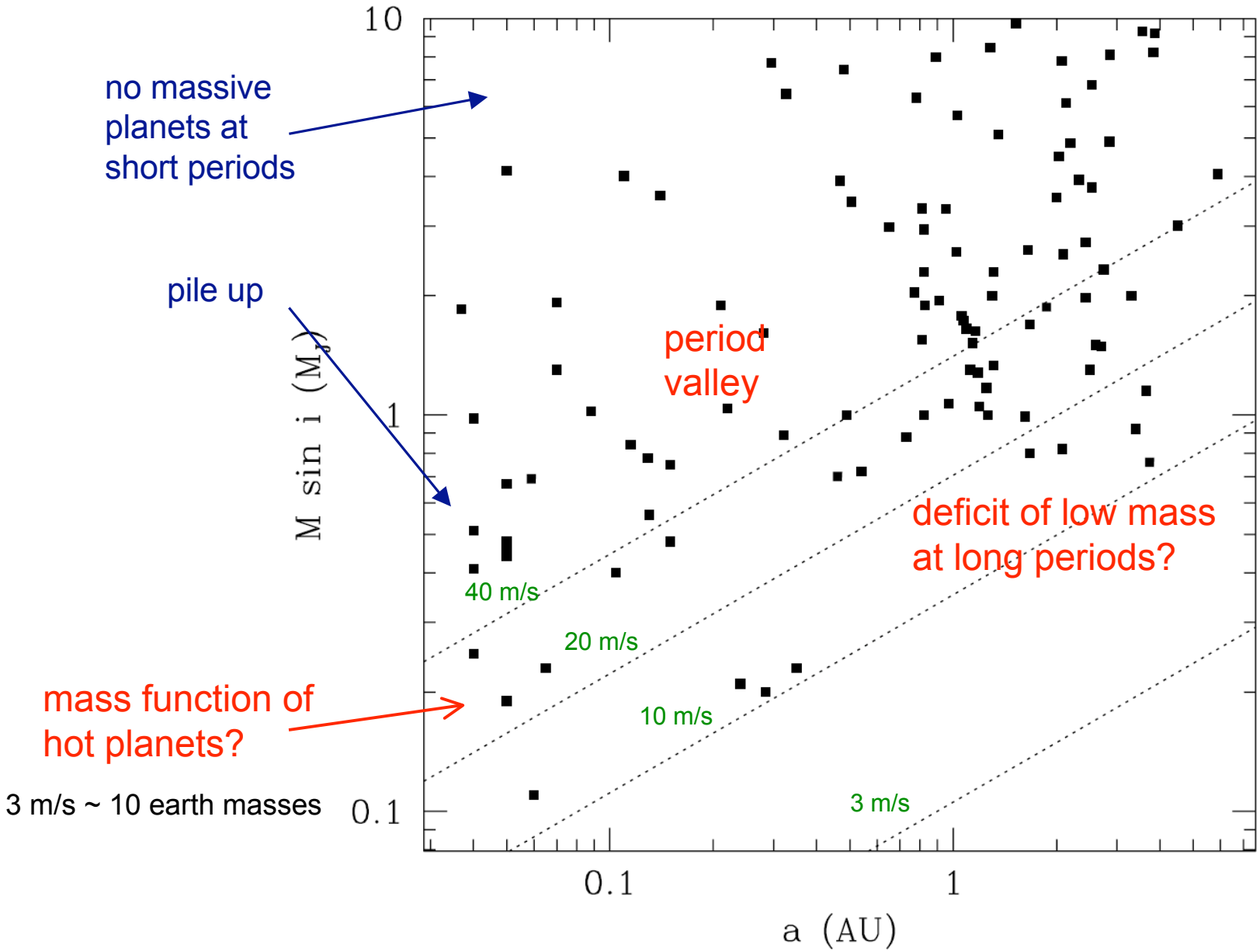
$$L_j = (1 - f) q_j + f \int dM \; dP \; n(M, P) \; p_j(M, P).$$

don’t need to decide what is and isn’t a “detection”
include both possibilities, weighted by probabilities

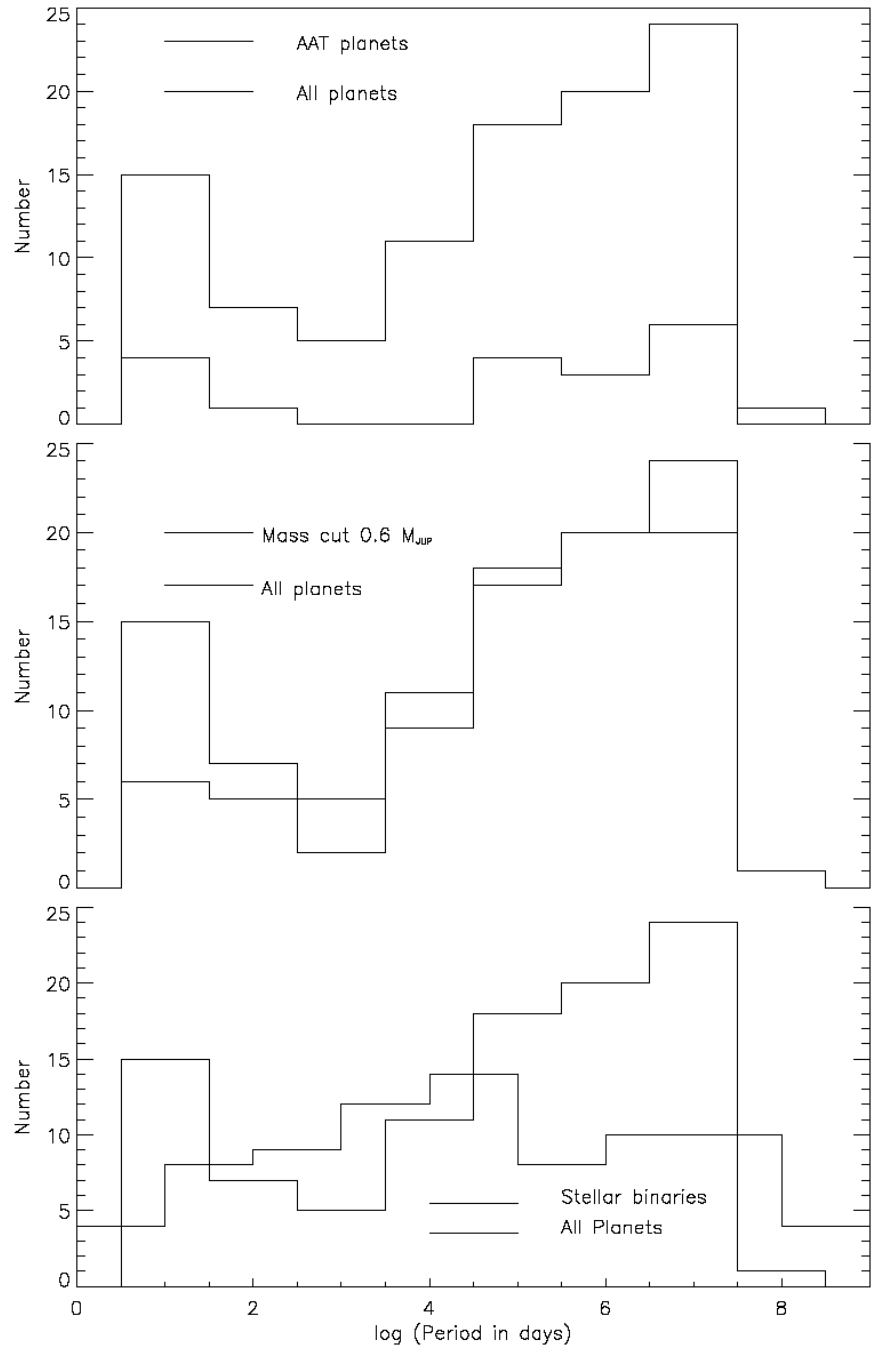

Prob
1
0.8
0.6
0.4
0.2
0
0
10
20
30
K (m/s)


Prob
1
0.8
0.6
0.4
0.2
0
10
100
1000
Period (d)

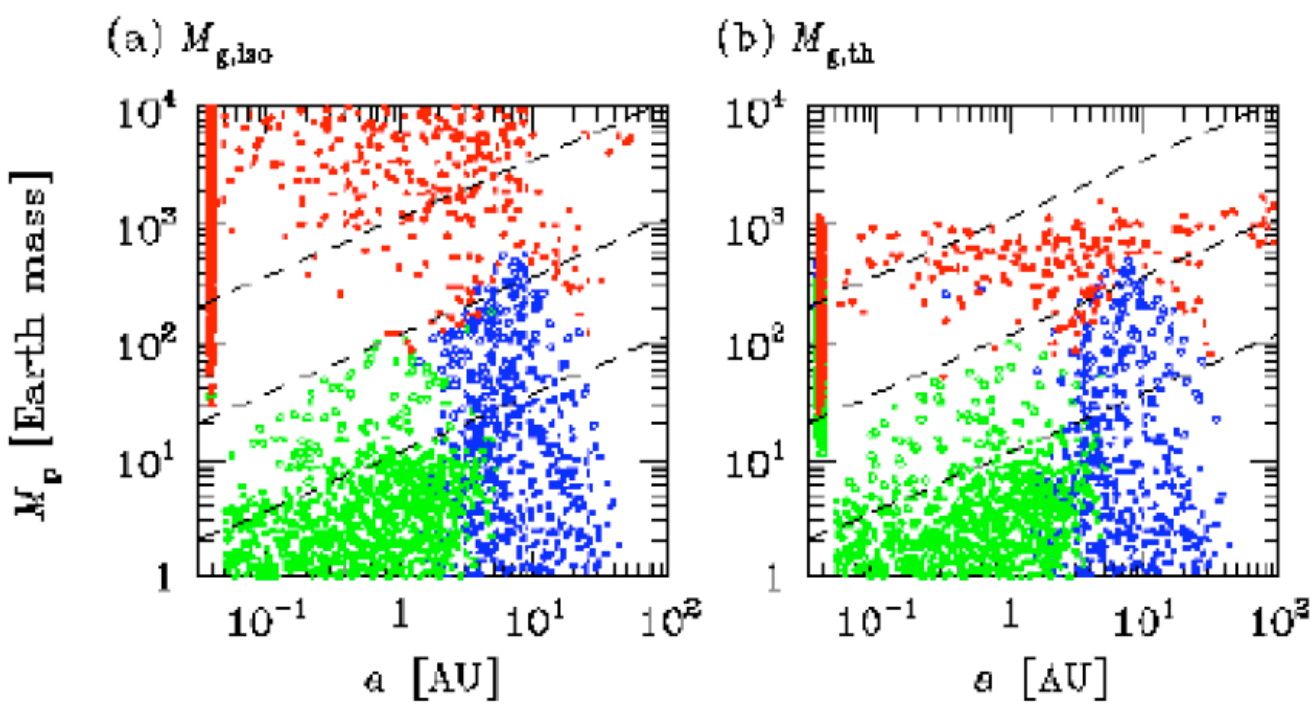

Prob
1
0.8
0.6
0.4
0.2
0
0
5
10
15
K (m/s)
Prob
1
0.8
0.6
0.4
0.2
0
10
100
1000
Period (d)


no massive planets at short periods
pile up
period valley
deficit of low mass at long periods?
mass function of hot planets?
3 m/s ~ 10 earth masses
40 m/s
20 m/s
10 m/s
3 m/s
10
1
0.1
0.1
1
M sin i (M_J)
a (AU)

Jones et al. 2003;
Udry et al. 2003; CMB99



Udry et al. (2003) suggest a deficit of low mass (<0.75 Mjup) planets at periods > 100 days

rapid migration for planets in this range

Ida & Lin (2003) astro-ph/0312144 core accretion + migration