Modelling radial velocities and transit light curves

Rodrigo F. Díaz

Department of Astronomy, University of Geneva

rodrigo.diaz@unige.ch

IVth Azores International Advanced School in Space Sciences
July 25th, 2016

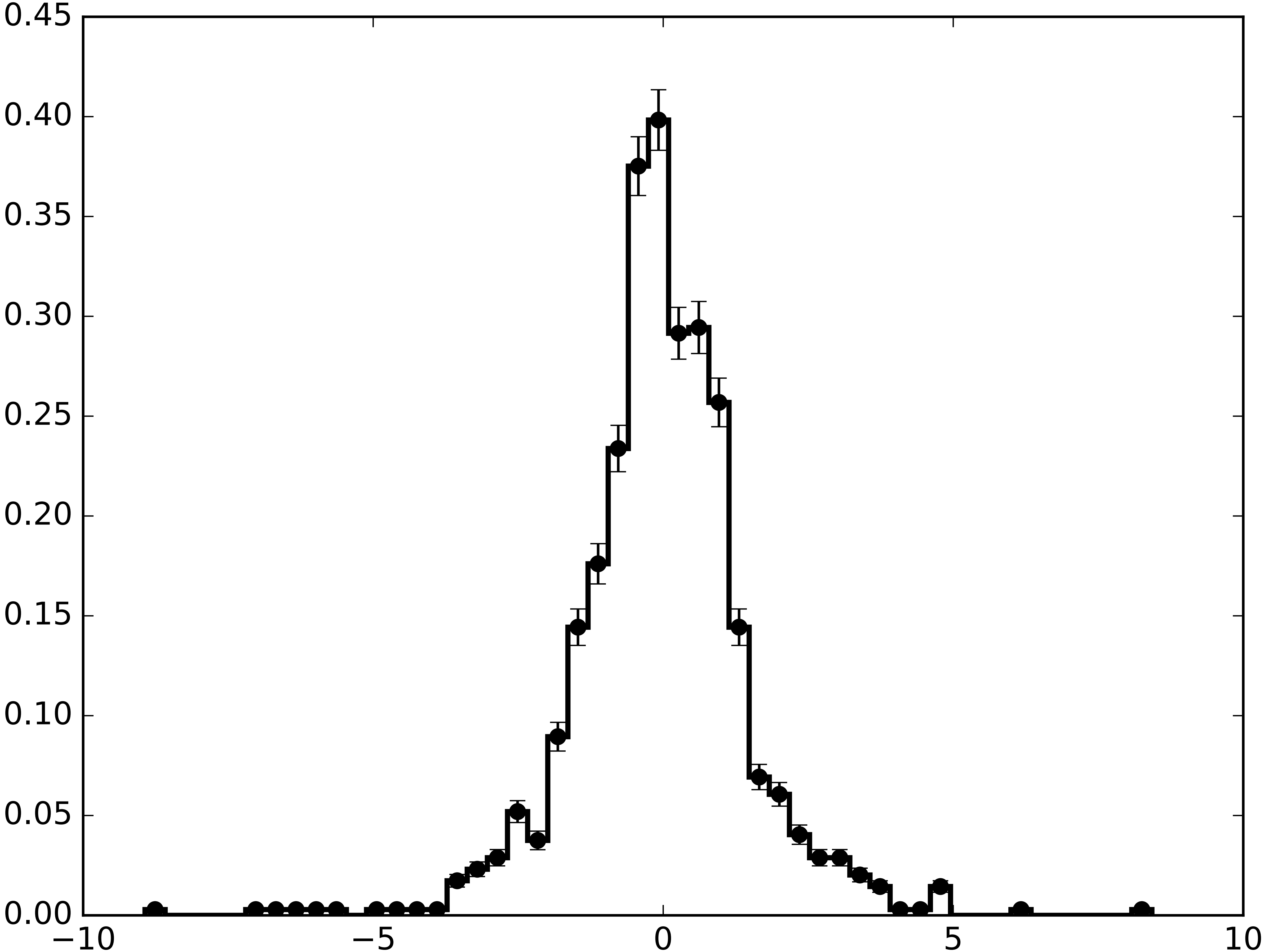
PlanetS
National Centre of Competence in Research

UNIVERSITÉ DE GENÈVE
FACULTÉ DES SCIENCES
Département d'astronomie

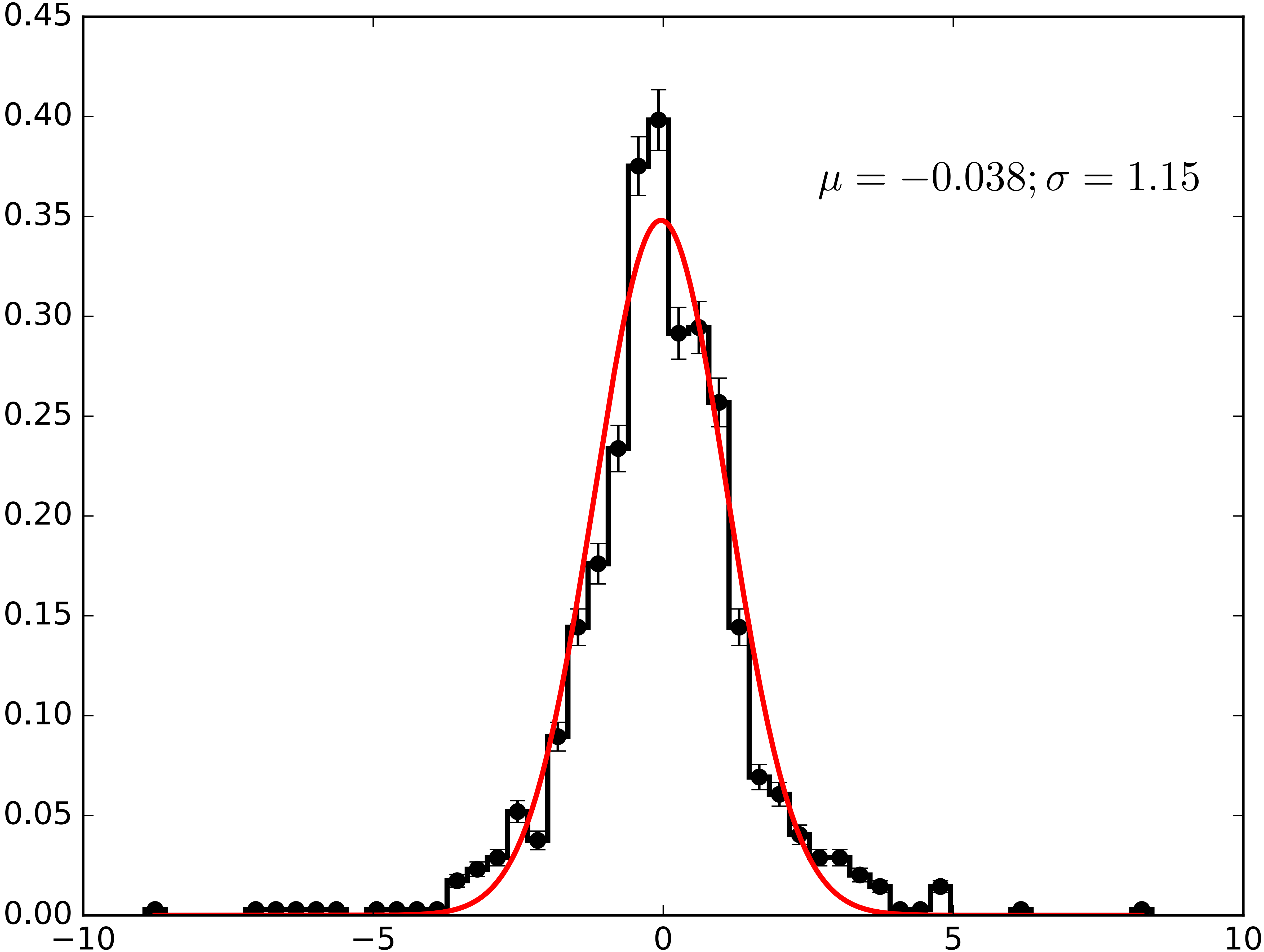
FNSNF
Swiss National Science Foundation

The National Centres of Competence in Research (NCCR) are a research instrument of the Swiss National Science Foundation

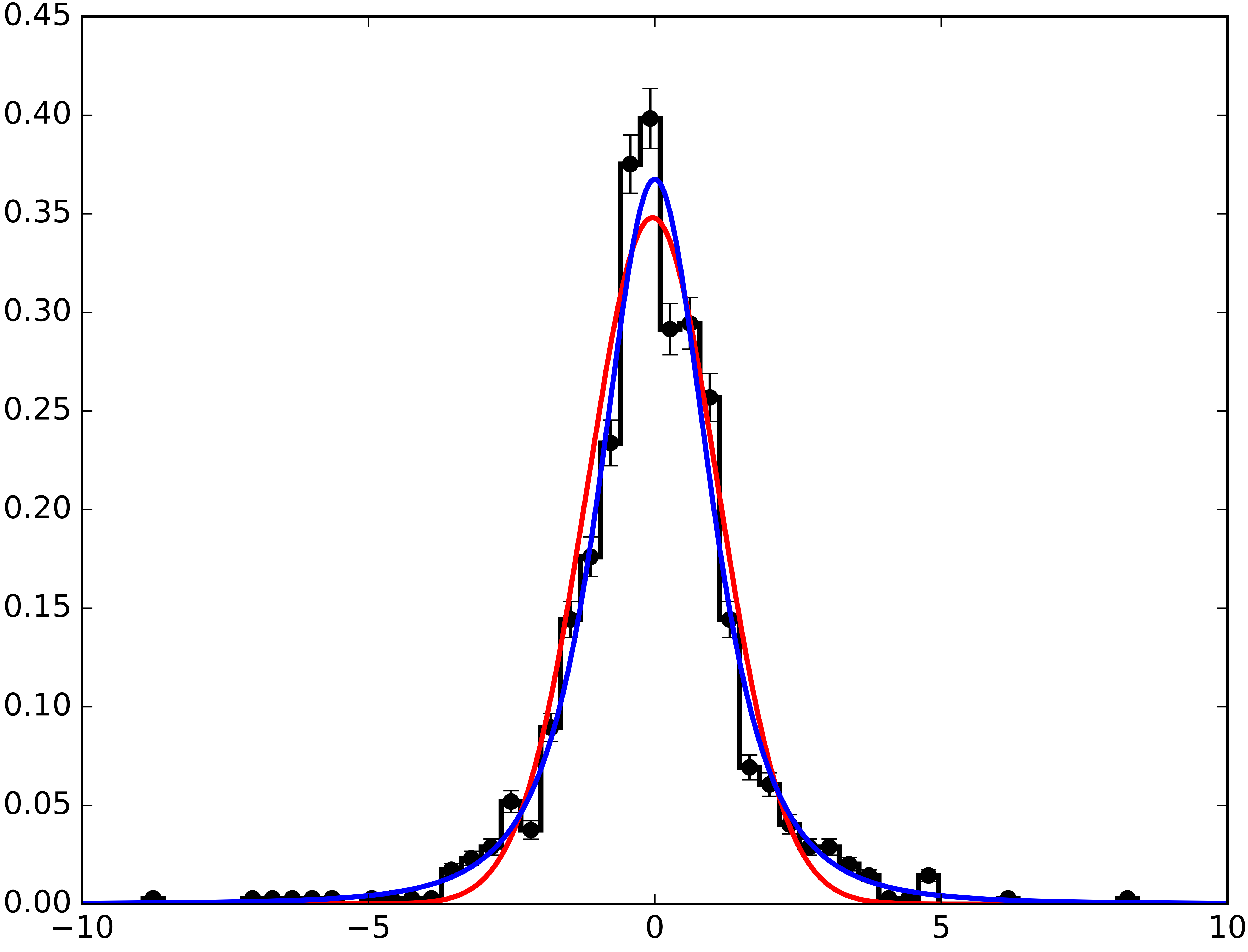
-0.4258, 0.9629, 0.8271, -0.0761, -0.1620, -0.2599, 0.6481, 1.2517, 0.2244, 0.0331, -1.0516, 2.8017, 1.8559, 0.9382, 0.8312, -0.1303, 0.9221, 0.6060, -0.5698,
-0.3927, -1.3152, 1.1827, 1.6165, -0.6727, 0.8533, -0.1789, -1.2074, 0.8020, -0.1568, 0.7057, 0.5656, -0.7147, 1.1954, 0.5416, -0.2113, 1.0522, 1.9738, 0.1474,
0.6597, 0.3949, -1.2146, 0.2744, 0.6703, -0.6913, -0.9657, 0.8498, 0.9167, -0.5060, 0.7082, 0.3312, -0.3092, -0.3044, -1.5772, -0.9272, -0.8858, 1.3993, -0.3382,
-2.8431, 0.9452, -3.0444, -2.0309, 0.1416, 0.1474, -0.0208, -2.9255, 0.3000, -3.8650, -0.0282, 1.1727, 0.5815, -1.1257, 8.4252, -2.5299, 0.0382, -0.6373, 0.2693,
-0.3993, 0.6195, 1.2367, 0.8516, -0.5885, 0.6170, 1.3144, 4.7261, -1.7593, -0.0083, 0.4488, -0.8542, -0.8866, 0.9777, -0.7837, -1.1594, -4.0903, 0.3139, 1.3620,
-0.5746, 1.2583, -0.4483, -0.0271, -0.5161, -1.8327, 3.2739, -0.2937, -0.1833, -1.2682, 0.5631, 2.1498, -0.5458, -0.7951, -0.9052, -0.9792, -1.6048, -1.3072,
-3.5114, -0.1072, -0.5162, -1.2271, -1.2452, -0.3166, -4.9986, 1.4746, -0.4988, -0.0378, 0.0032, 0.2864, -0.4586, -0.1442, 0.4167, 0.1680, 0.4311, 1.2459, 0.4811,
2.0354, -0.1447, 0.0320, 0.0170, 1.1043, 0.0977, -1.0106, 2.1566, -0.3881, 0.8465, 0.7347, -0.6717, -0.7952, 0.0710, 3.6445, -0.0550, -0.3550, 0.6799, -1.1895,
-0.0828, 1.3692, -2.1940, -0.5108, 0.4514, -1.5967, -1.0766, -0.0951, -2.6083, -0.2892, -1.3840, 0.3676, 1.1254, -0.5696, -0.4879, 0.0799, -1.1759, 0.9598,
1.9274, 0.7494, 0.1644, -1.6152, -2.5340, 0.3410, 1.0850, 0.5481, -0.1523, -2.2027, -2.5697, -1.0047, 0.5154, 0.8907, 0.2663, 0.7922, 1.0089, -1.4280, 2.2191,
0.8956, -0.4291, 0.4167, -0.9900, -0.0448, -2.0657, 0.0179, 0.6709, -1.4507, 0.0385, -0.8988, 0.9545, 0.9640, -1.7518, -1.0884, -0.4236, 1.6847, -1.9293, 0.6477,
0.4632, 4.8035, -0.4204, -0.7591, 0.4864, 0.9606, -1.7324, 0.3425, -0.8570, -0.7900, 1.0013, -1.4189, -1.6568, -0.2089, 0.0437, 0.3574, 0.3368, -1.0191, 1.3581,
-1.1470, -1.7700, -0.2553, 0.2416, 1.4234, 0.0679, -2.3670, -0.4121, 0.5269, -0.2469, 0.0158, -0.1713, 1.1389, 1.1288, -0.4167, -0.7257, -0.5442, 0.0650, 0.6510,
-1.5929, 1.0691, 0.3750, -2.2971, 2.9754, 2.9149, -1.3911, 2.2526, 0.1382, -1.7603, -1.0631, 0.5170, -0.1009, 1.1263, -0.0092, -0.0098, -2.7842, 0.6913, -0.9865,
0.2700, 0.9492, -0.0675, 1.0261, 0.0933, -0.9754, -0.9568, -0.2461, -1.5360, 0.4426, 0.2175, 0.7143, -0.5052, -0.8001, 0.5656, 0.7063, 1.4770, -1.3006, -1.5714,
-0.2693, 2.2670, 0.1558, 0.1991, -1.6480, 0.5153, -2.4253, 1.5756, 0.5443, 2.5773, -1.9362, -0.2781, -2.0731, 0.1090, -0.5054, -0.4368, -0.7355, -0.1636, -1.6857,
-1.5232, 1.8666, 0.7096, 1.6585, -0.2773, 1.0332, 1.6389, -1.6783, 0.4412, 0.2219, 0.3065, 0.2976, -0.5809, 2.1519, 0.0975, 1.0190, -0.0502, -2.4786, -1.7093,
1.0806, -3.3706, 1.6150, -0.2482, -0.3966, -0.6100, 0.5156, -0.5557, 0.7086, 1.2471, 0.3134, -1.7028, 0.2730, 1.3062, -2.3565, -0.8197, 1.0612, 1.2029, 0.9465,
0.3207, -0.2665, -0.7594, -1.0070, 2.2674, 0.6885, -1.3172, -0.5294, 0.3230, -0.8870, -1.6809, 0.3694, -0.8466, -0.7692, 0.8474, 1.0270, 0.5491, -1.0072, 2.5180,
-0.6298, -1.5580, 0.7622, 1.3980, 0.3573, -0.2377, -0.1133, -3.4173, -1.2886, -0.6681, -1.9261, 0.5664, 0.0513, -1.0560, -0.0775, -0.0587, -1.5512, 2.4590,
0.4063, 1.4058, 0.0397, 0.1408, -0.3774, 0.3097, -0.7795, -1.3605, -1.3067, -0.7253, -6.3967, 1.2869, -1.3681, 0.3538, -0.0138, 0.9987, -0.0393, -0.0788, -2.8061,
-0.6761, -0.2404, -0.6004, 1.5730, 2.5828, 0.2641, 0.1437, -0.0809, -1.6422, 0.5712, 0.8864, 0.5777, 0.6425, -0.0133, -2.3401, 1.6605, 1.4231, 0.8345, -0.7660,
-1.1168, -0.5035, 0.0766, 2.3952, 1.0476, -1.2858, -0.2306, 1.1802, -3.3299, -0.4132, -0.3664, -1.5994, -0.3148, 4.3496, -2.5871, 1.3305, 0.9727, 0.6839, 0.0276,
0.2045, 1.2147, 2.5365, 1.1441, 3.2629, 0.1362, -0.4081, -0.6274, 0.6130, -0.3459, 0.4389, -0.9244, -2.0190, -1.2441, 1.4978, 0.2328, -0.2747, -0.5876, 0.9303,
0.4465, -0.4086, -0.6834, -1.6257, -3.4097, -0.2872, 0.3217, -0.8903, 0.5751, -2.5417, 0.7618, -0.1990, 0.4749, -0.6886, 0.4808, -0.6274, -0.1490, 1.8698, 0.5414,
0.2706, 0.8269, -0.4423, 0.6014, -0.0533, -0.5480, 0.9465, -0.0274, -0.7399, 0.1950, -3.0122, -0.2117, 1.0063, -1.2950, -0.8230, -0.0239, 0.4498, -0.8517, 3.6613,
1.3914, 1.5189, -2.1787, 0.3822, 0.1255, 0.3432, 1.0617, -0.3204, 1.3348, 0.9355, 0.6301, 3.4833, -0.8114, 0.1575, 3.1708, 0.6257, 0.7065, -1.1157, -2.3580,
-4.6061, 0.8672, 0.9294, 2.9604, 0.1266, -1.9851, -0.8187, -0.1706, 2.7169, 2.7969, 2.1371, 3.0910, -0.8833, 0.2133, -8.9333, -1.2552, -0.8953, -0.0167, 0.5444,
0.4873, -1.1230, -2.2517, -1.1703, 2.3537, 0.7819, 0.6830, 0.2425, -0.5108, 0.0541, 0.3107, 0.0824, -0.0954, -1.9473, -0.4189, 0.5788, -1.7985, 1.1069, -1.0192,
-0.3612, -0.1431, -0.3686, -0.3657, -1.3466, 1.0715, -0.0615, 0.7700, 0.9807, 0.6837, -0.4829, -0.4847, -0.0408, -2.9214, -0.7188, 0.4714, 0.2608, 0.1678,
-0.7726, 2.4223, -1.6223, 1.1990, 0.0309, -3.1811, -0.0955, -1.4879, 0.8964, -0.1063, -0.2639, 1.4454, 0.1860, -0.3108, 1.0299, 0.4298, -0.4844, -0.1514, -0.0100,
-1.2320, 0.9244, 0.3842, -2.0463, -3.5029, 2.2090, -0.0722, -0.3270, -1.6422, -0.6720, 0.1335, 0.7404, -0.0754, -0.9890, 1.7827, 4.0538, -0.0620, -0.5512, 0.8910,
0.6116, -0.9693, 0.2661, -0.4064, -2.1055, 0.4639, 0.3008, -0.1206, 0.1843, 3.2056, -3.3907, 0.8447, -0.6212, 0.2249, -1.9509, 0.5351, -1.2619, 1.8796, -0.7371,
-0.6632, 0.9288, -2.0942, 1.1829, -1.2170, -0.3879, 0.2870, -0.3780, 0.4796, 4.6440, -3.2195, -0.7724, -1.3707, -0.4496, 0.2207, -0.3735, 4.8009, 0.2859, 0.3271,
-0.5053, -3.2197, 0.0770, -3.0705, -2.7360, 0.0243, 1.3920, -1.3508, 1.5471, 0.0783, -2.5703, 1.4966, -1.5306, -2.7655, -1.0365, -0.5016, 0.1065, 0.5727, 0.8461,
-0.1113, -0.2310, 0.9931, -0.3884, 0.7402, -1.5023, -0.6913, -1.0615, -1.9504, 1.1166, 1.6918, -0.1562, -1.4498, 0.0006, 1.3752, -0.3413, -0.7429, 0.3983, 0.4719,
-1.6959, -0.2056, 0.2178, 0.5141, -0.5745, 0.5449, 0.2420, -0.1296, -1.9691, -2.4805, 1.2441, 1.4198, -0.9786, 1.2632, 1.9183, -0.4359, -0.1306, 1.4974, 3.4456,
0.1820, 0.2420, -0.1202, -1.5408, -1.5500, -0.3047, 0.0402, 1.4828, -0.4459, 3.0487, 0.6775, -1.6424, 0.4681, -1.3879, 0.0298, 0.1707, -0.5672, -0.2238, 0.1543,
0.4965, -0.8725, -0.1862, 0.0536, 1.9399, 0.8416, 0.0080, 2.3084, 0.8265, 0.4797, 1.3048, -0.6953, -0.9466, 0.2398, 0.2648, 1.0545, -0.2481, -0.7224, 0.7350,
0.0967, 0.5514, 0.8454, 0.7592, 0.7669, -1.4729, -0.9691, 0.5400, 1.7118, 0.2743, 0.4738, -1.4815, 4.9319, -0.3526, 2.2348, -0.0072, -1.8503, -1.2781, -0.4065,
1.1592, -1.4303, -1.1181, -1.1649, 1.0389, 0.9183, 2.8120, -1.0726, 2.3604, 0.7836, 0.8147, -0.8836, 1.1037, 0.2468, -5.7165, 0.2066, 0.5739, -0.5308, -0.5953,
0.2851, -1.3518, -1.2859, 0.7298, -1.5753, -1.3553, 3.4118, 1.8400, -1.0751, 0.7278, -0.2014, -0.1527, -0.2745, -0.4837, -0.6613, -3.5672, 0.2222, 0.9499, 2.1372,
1.4601, -1.7270, -0.2280, 1.6503, -0.0368, -5.9512, 1.9027, 1.5553, 1.2709, 1.4792, -0.2827, -0.1447, 0.9484, -0.6560, -0.1931, -0.4891, 1.2324, -0.2650, 2.6470,
-1.5580, 1.1036, -2.4841, -0.2640, -0.2440, 0.6009, -0.1926, 1.4688, -0.6251, 3.3118, -0.4526, 1.0283, -0.3820, -3.0154, -1.1445, 0.8195, -0.3249, -0.3703,
3.5886, -0.7405, -1.2940, 1.2646, -0.3361, -1.7244, 0.2509, 2.9173, -0.1308, 0.9136, 0.0706, -0.9477, -2.4055, 0.7462, 0.9538, -0.8179, -0.4440, -1.1852, -0.7789,
-0.5011, -1.8987, -0.3188, 2.0584, 0.6024, -3.1321, -0.6272, -0.5745, 0.4365, 0.9436, -0.6201, 1.3551, -0.4047, 1.5559, 1.0733, -0.3982, -1.4490, -0.0592, 2.8637,
0.9025, -0.7550, 1.5398, -0.5594, 0.4108, -0.5978, 0.0613, -0.9095, -0.4517, -0.9100, -0.3021, -0.3001, -0.2229, 0.9803, 1.8816, 2.9271, 3.3045, -1.2201, -0.9750,
0.0013, -0.9089, -0.2484, 0.0234, -1.1106, -0.8048, 1.0323, -0.5966, 6.1361, -0.1659, -0.0031, 0.0573, -2.0689, -1.6038, 0.6123, 0.4766, -0.1283, -0.3109,
-6.5703, 3.0460, -0.4466, -0.3737, -0.1055, -0.1960, -1.8886, 0.9178, -1.5232, 0.0379, -1.4492, -1.1803, 1.0930, -0.0066, 0.2857, -0.0220, -1.0741, 0.5158,
1.5005, 0.7451, 0.4993, -0.0173, -0.5180, -0.3155, -0.9085, -0.0370, 2.1473, -0.6131, -1.8105, 0.3878, 1.2429, -0.3604, -0.3411, 2.8003, -0.5535, -2.9557, 3.8684,
-2.5195, -0.9888, 1.3046, 2.0976, 0.1146, 1.1547, 0.7065, -0.9618, -0.9320, 0.6342, -0.4095, -0.4565, -1.5873, -0.2245, 2.5148, -6.8740, -0.5089, -0.1434,
-0.4689, 1.6055, -0.2133, -0.4027, 3.5856, -0.0116, -1.7678, 0.0375, 0.0212, 0.0312, -0.8189, 1.3116, 0.9817, 1.8707, 0.0307, -0.6598, -0.5693, 0.2901, -0.1544,
-0.7304, -0.8341, 1.1035, 0.4524, -0.5887, 1.7743, 0.8751, -0.2558, 0.4451, 0.4259, 0.0815, -0.5457, -0.3495, -0.8443, -0.3872, -2.6320, 0.8772, -1.7566



0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
−10
−5
0
5
10



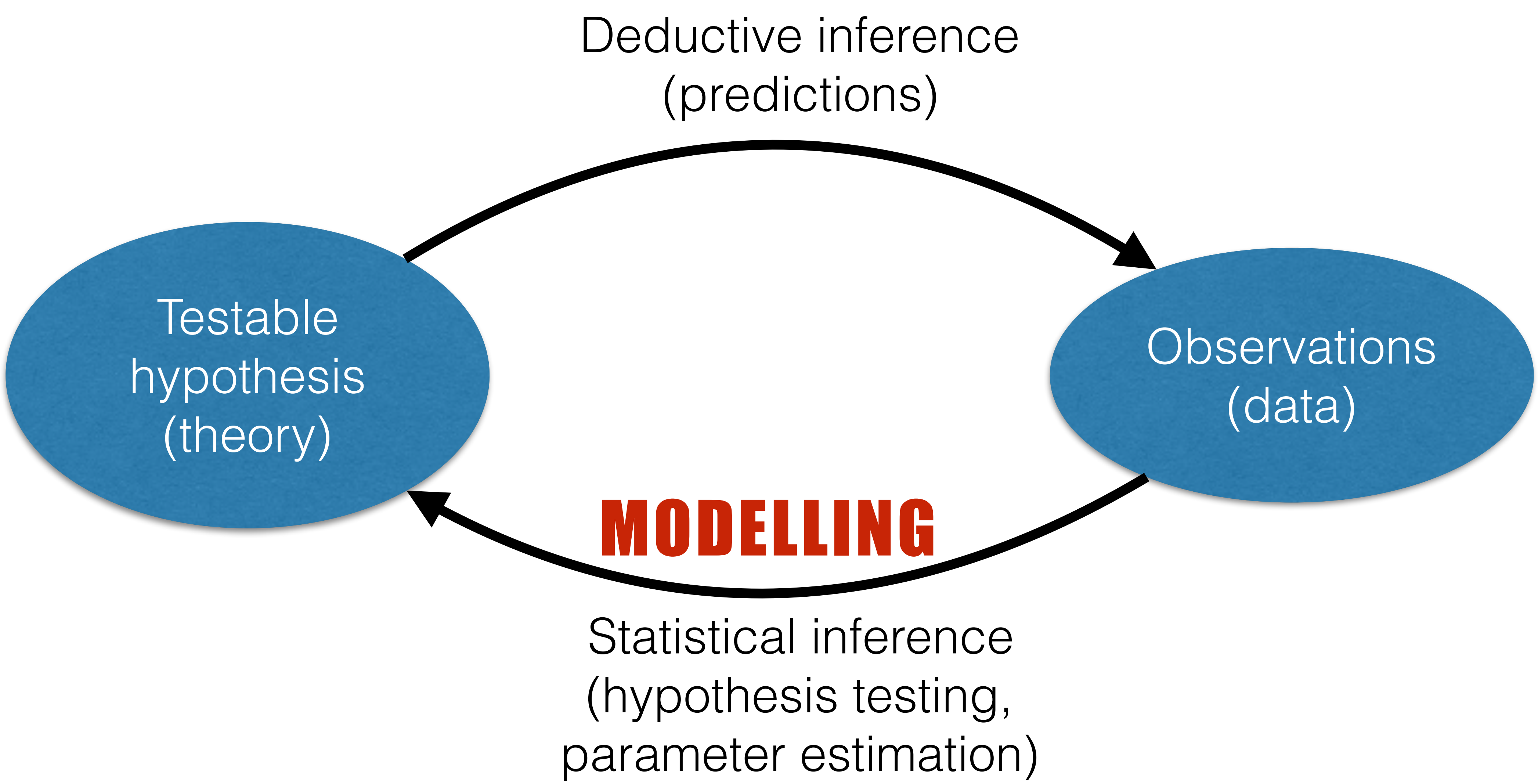
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
−10
−5
0
5
10
$\mu = -0.038; \sigma = 1.15$



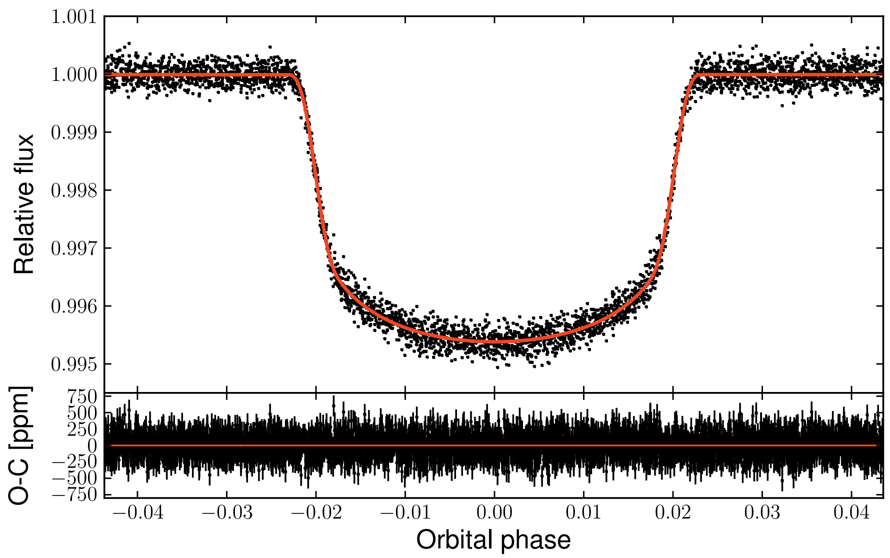
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
−10
−5
0
5
10

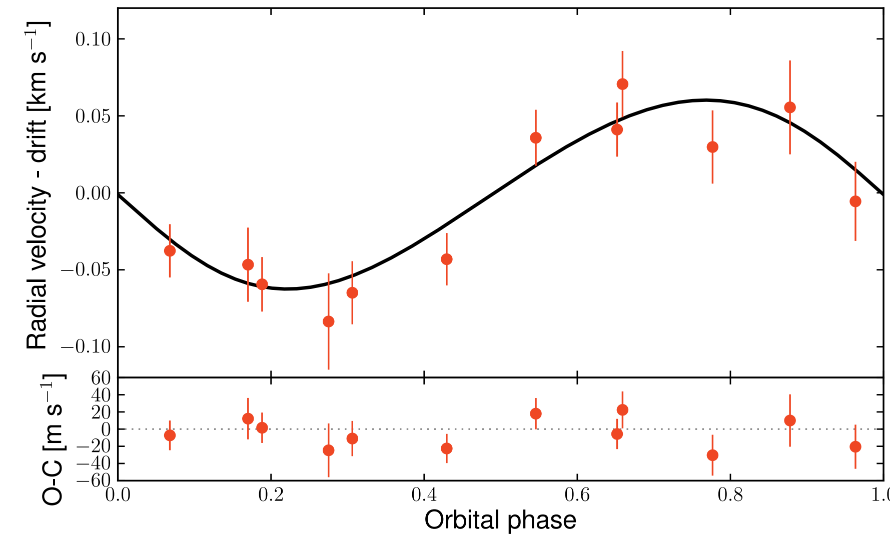
Data modelling in the scientific method

Fig. adapted from Gregory (2005)



Transits + Radial velocity —> density

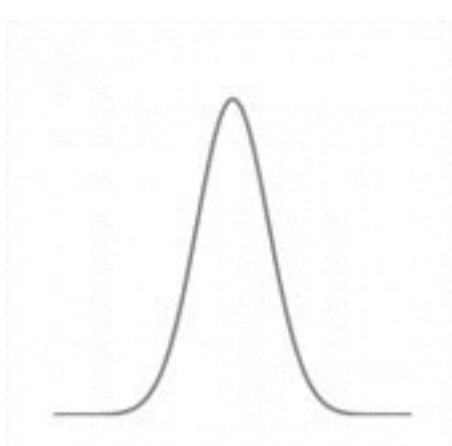




$$\frac{R_p}{R_\star}$$

$$\frac{M_p}{(M_p + M_\star)^{2/3}}$$

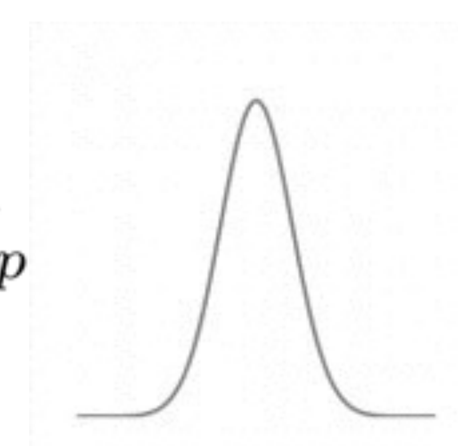
with stellar models

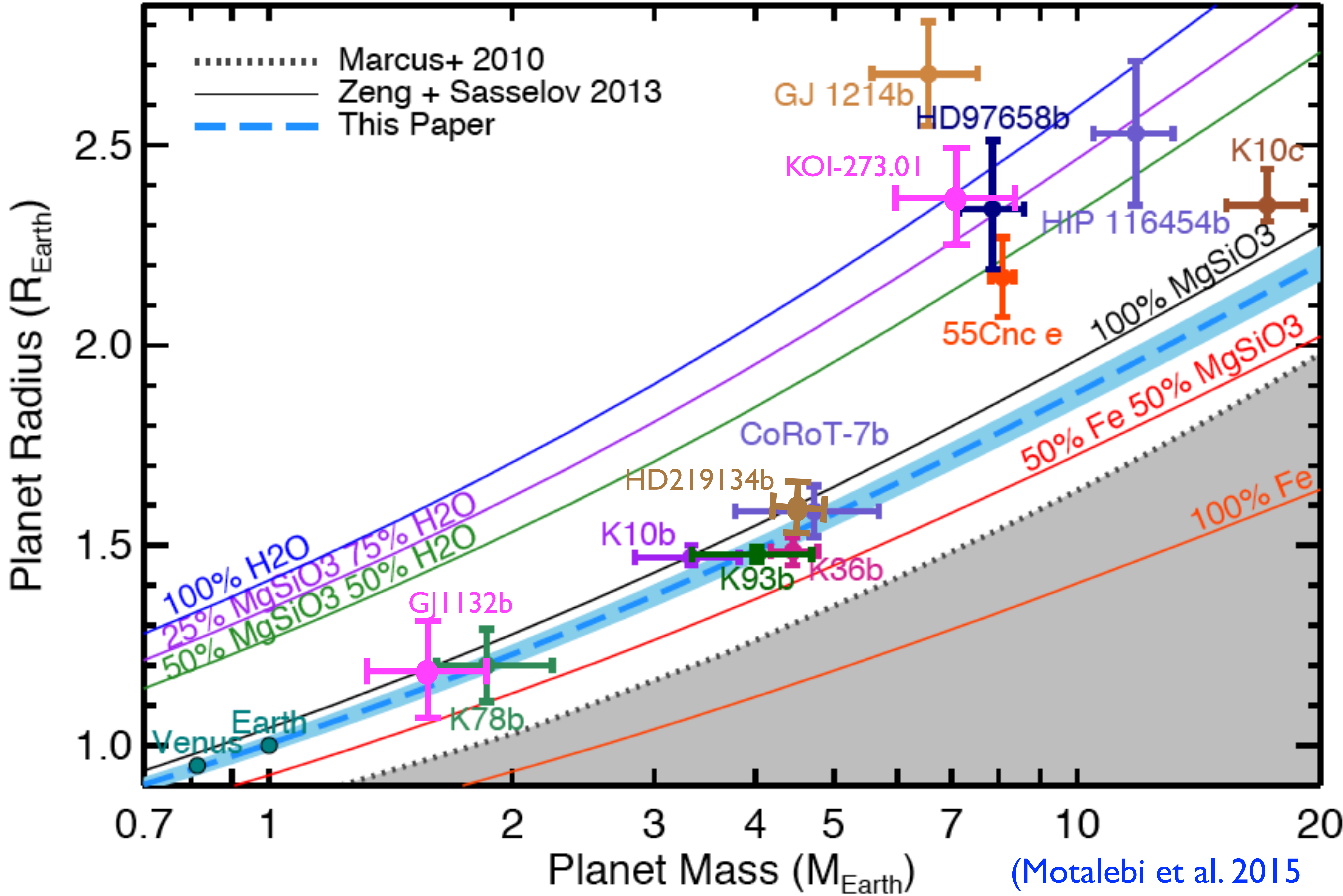


$$R_p$$

$$\rho_p$$

$$M_p$$





(Motalebi et al. 2015 updated from Dressing et al. 2015)

(precision on the mass < 20%)

Physical models

Radial velocity

Transit light curve.

$$d_i = m_i + e_i$$

data model error

Data analysis

Likelihood functions

MCMC

$$p(D|\theta, H, I)$$ → **Priors**

Error model

Beyond the independent normal paradigm

Physical models

Radial velocity

Transit light curve.

$$\underbrace{d_i}_{\text{data}} = \underbrace{m_i}_{\text{model}} + \underbrace{e_i}_{\text{error}}$$

Data analysis

Likelihood functions

MCMC

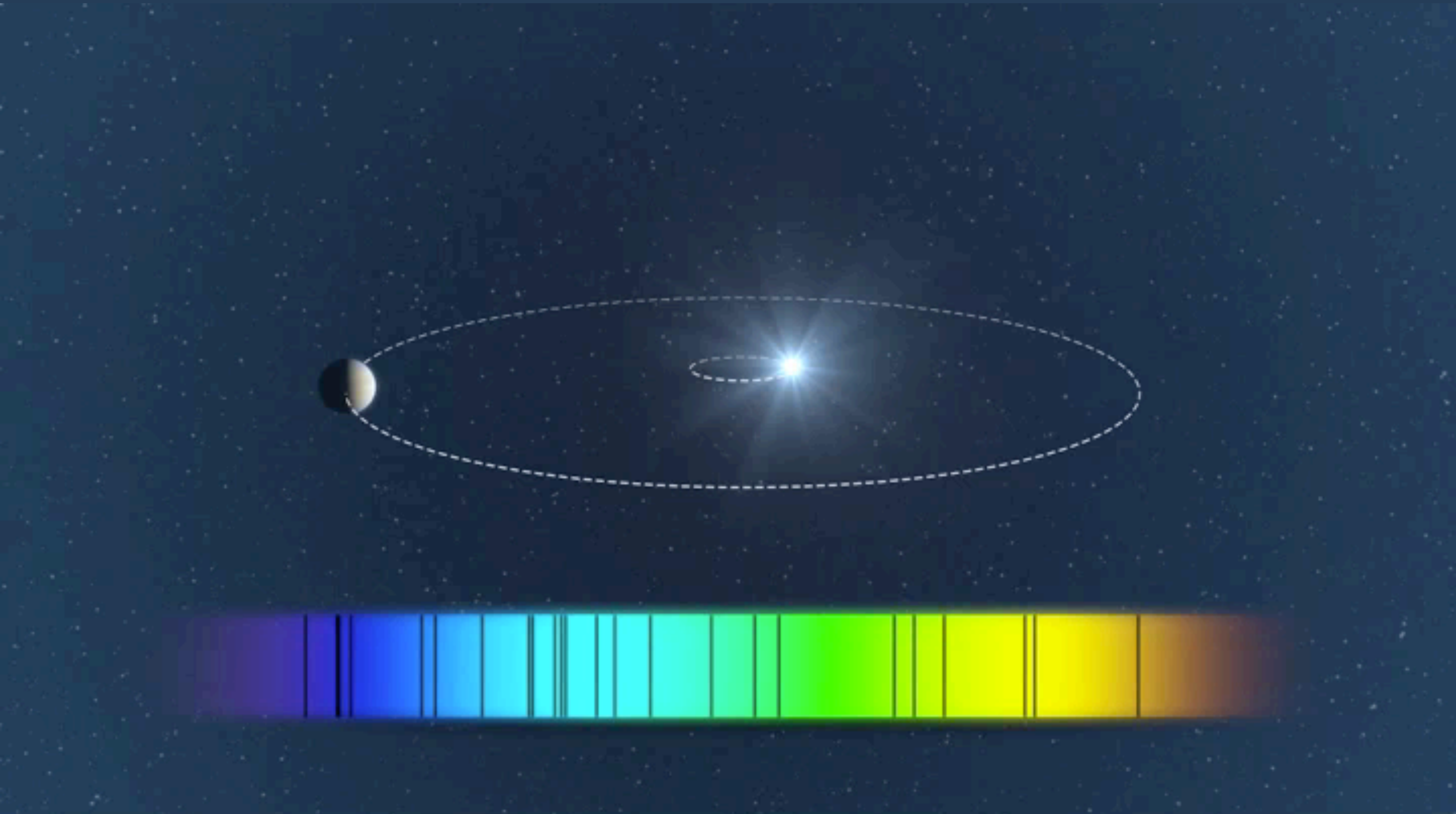
$p(D|\theta, H, I)$

Error model

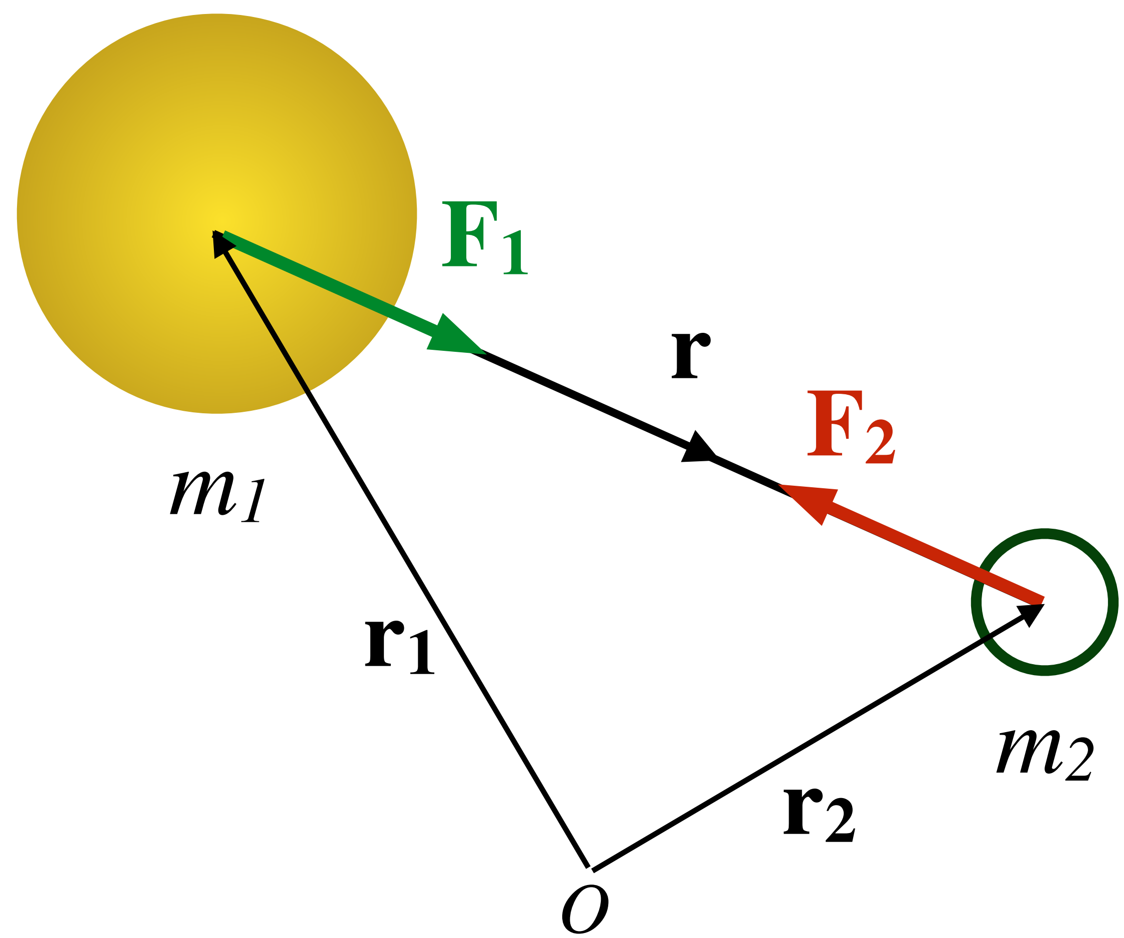
Beyond the independent normal paradigm

Physical Models

Radial Velocities



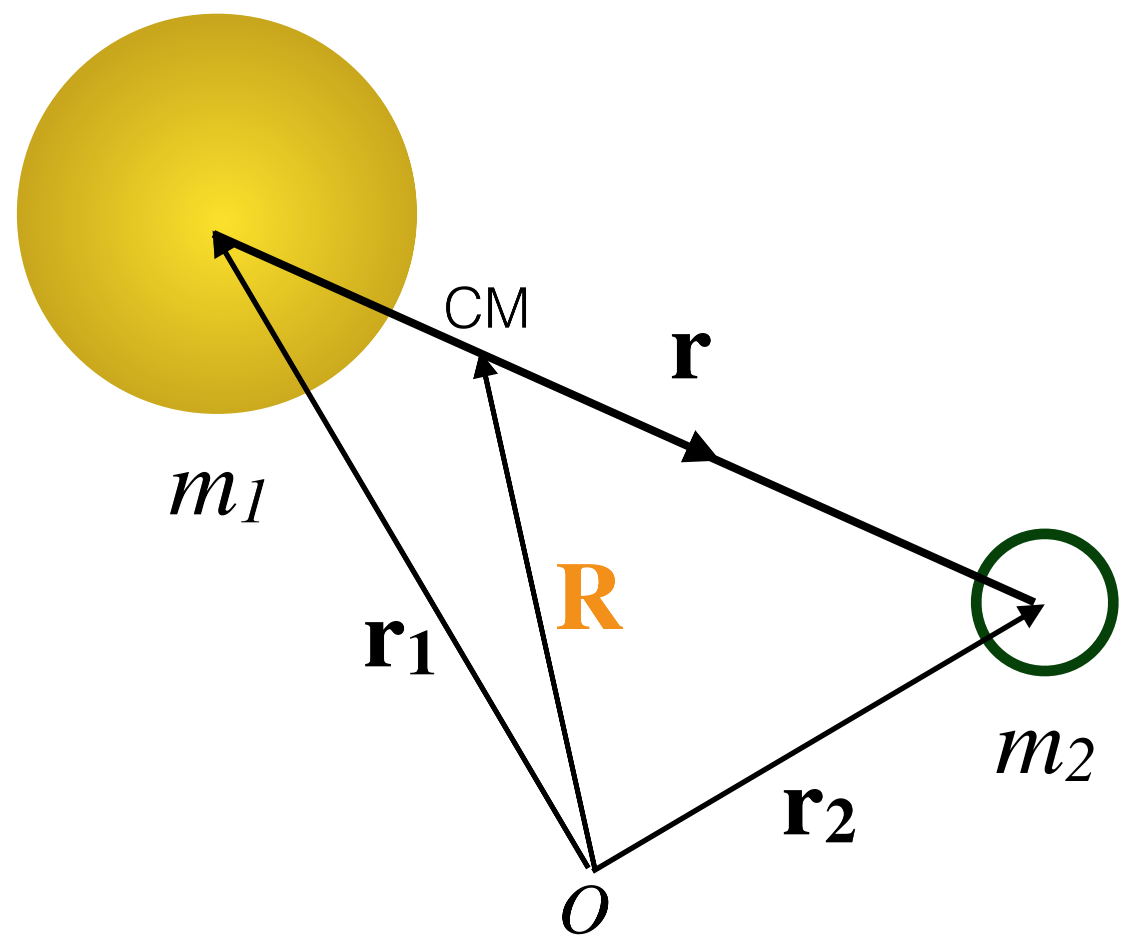
The two-body problem



$$\mathbf{F_1} = +\mathcal{G}\frac{m_1 m_2}{r^3}\mathbf{r} = m_1\ddot{\mathbf{r}}_\mathbf{1} \qquad \mathbf{F_2} = -\mathcal{G}\frac{m_1 m_2}{r^3}\mathbf{r} = m_2\ddot{\mathbf{r}}_\mathbf{2}$$

Murray & Dermott "Solar System Dynamics"

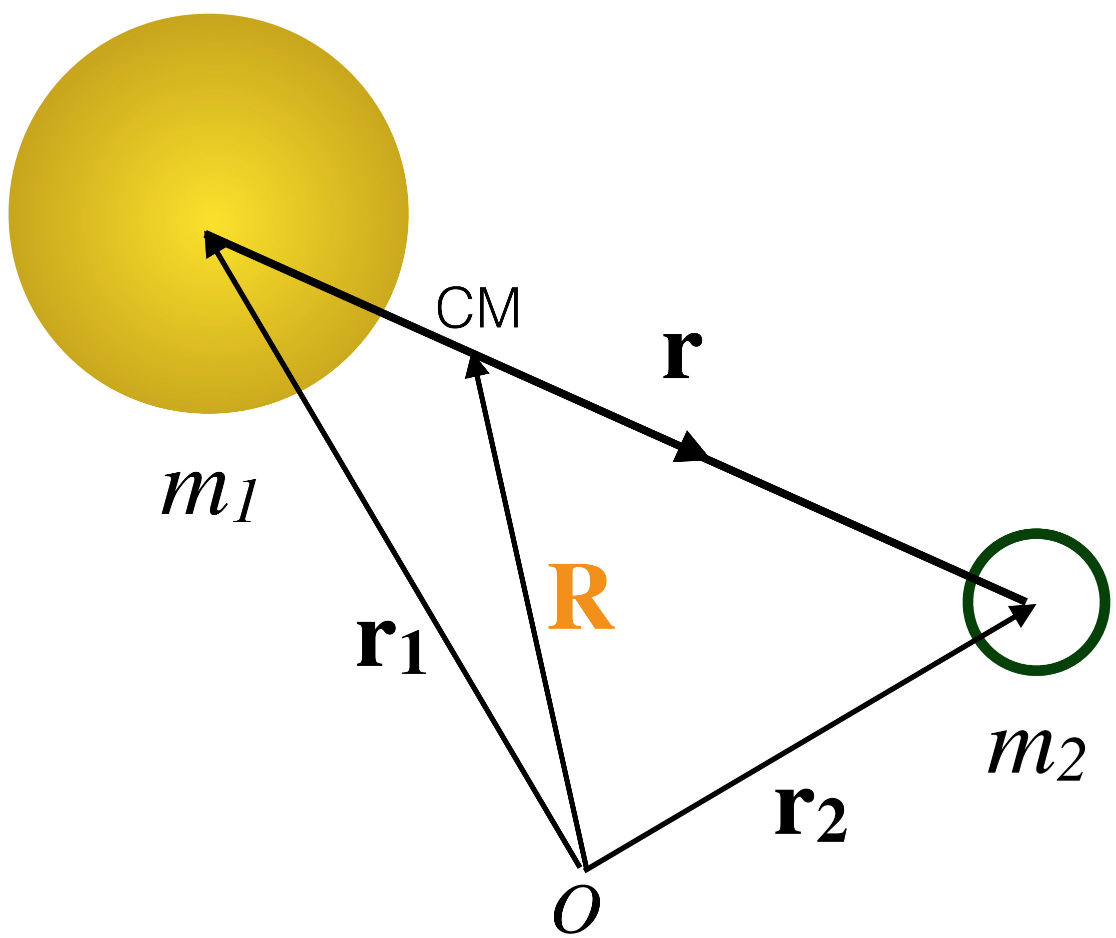
The two-body problem



$$m_1\ddot{\mathbf{r}}_\mathbf{1} + m_2\ddot{\mathbf{r}}_\mathbf{2} = 0$$

Murray & Dermott “Solar System Dynamics”

The two-body problem



$$m_1\ddot{\mathbf{r}}_1 + m_2\ddot{\mathbf{r}}_2 = 0$$

integrate

$$m_1\mathbf{r_1} + m_2\mathbf{r_2} = \alpha t + \beta$$

define

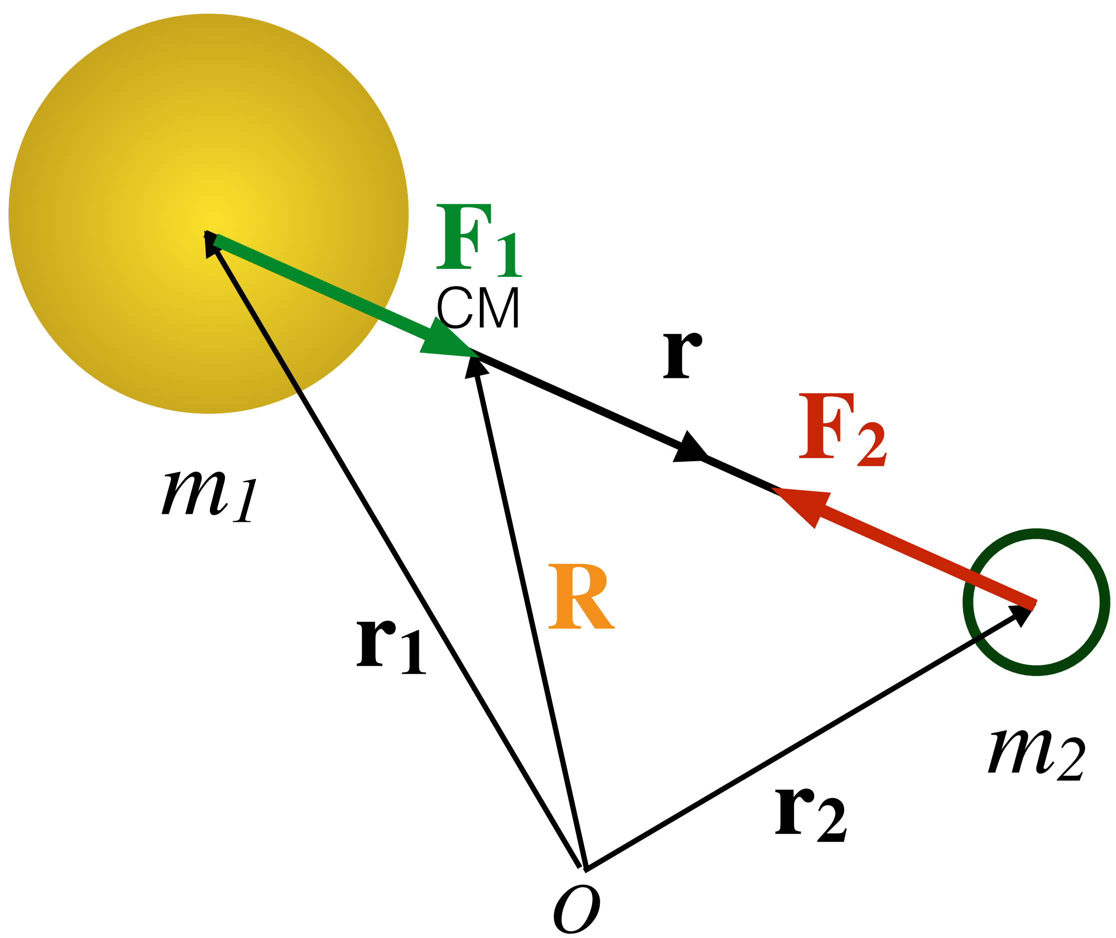
$$\mathbf{R} = \frac{m_1\mathbf{r_1} + m_2\mathbf{r_2}}{m_1 + m_2}$$

$$\boxed{\mathbf{R} = \frac{\alpha t + \beta}{m_1 + m_2}}$$

The centre-of-mass is stationary or moving in a straight line with constant velocity.

Murray & Dermott "Solar System Dynamics"

The two-body problem - Relative orbit



Now consider the motion of m_2 relative to m_1

$$\mathbf{r} = \mathbf{r_2} - \mathbf{r_1}$$

$$\ddot{\mathbf{r}} = \ddot{\mathbf{r_2}} - \ddot{\mathbf{r_1}}$$

$$\mathbf{F_1} = +\mathcal{G}\frac{m_1 m_2}{r^3}\mathbf{r} = m_1\ddot{\mathbf{r_1}}$$

$$\mathbf{F_2} = -\mathcal{G}\frac{m_1 m_2}{r^3}\mathbf{r} = m_2\ddot{\mathbf{r_2}}$$

Divide through: $\mathbf{F_1}$ by m_1, and $\mathbf{F_2}$ by m_2 and combine

Equation of relative motion

$$\ddot{\mathbf{r}} + \mathcal{G}(m_1 + m_2)\frac{\mathbf{r}}{r^3} = 0$$

Murray & Dermott "Solar System Dynamics"

The two-body problem - Relative orbit

$$\ddot{\mathbf{r}} + \mathcal{G}(m_1 + m_2)\frac{\mathbf{r}}{r^3} = 0$$

Equation of relative motion

Take vector product with $\mathbf{r}$

$$\mathbf{r} \times \ddot{\mathbf{r}} = 0 \longrightarrow \mathbf{r} \times \dot{\mathbf{r}} = \mathbf{h} = \text{constant}$$

Polar coordinates

$$\mathbf{r} = r\hat{\mathbf{r}}$$

$$\dot{\mathbf{r}} = \dot{r}\hat{\mathbf{r}} + r\dot{\theta}\hat{\theta}$$

$$\boxed{\mathbf{h} = r^2\dot{\theta}\,\hat{\mathbf{z}}}$$

- $\mathbf{h}$ is a sort of angular momentum.
- The movement of m_2 with respect to m_1 lies **in a plane**.

Murray & Dermott "Solar System Dynamics"

The two-body problem - Relative orbit

$$\ddot{\mathbf{r}} + \mathcal{G}(m_1 + m_2)\frac{\mathbf{r}}{r^3} = 0$$

Equation of relative motion

Scalar equations obtained

$$\hat{\mathbf{r}} \longrightarrow \quad \ddot{r} - r\dot{\theta}^2 = -\frac{\mu}{r^2}$$

with $\mu = \mathcal{G}\left(m_1 + m_2\right)$

$$\hat{\theta} \longrightarrow \quad \frac{\mathrm{d}}{\mathrm{d}t}\left(r^2\dot{\theta}\right) = 0$$

conservation of angular momentum.

$$r = \frac{p}{1 + e\,cos(\theta - \varpi)}$$

where ϖ and e are integration constants and $p = h^2/\mu$

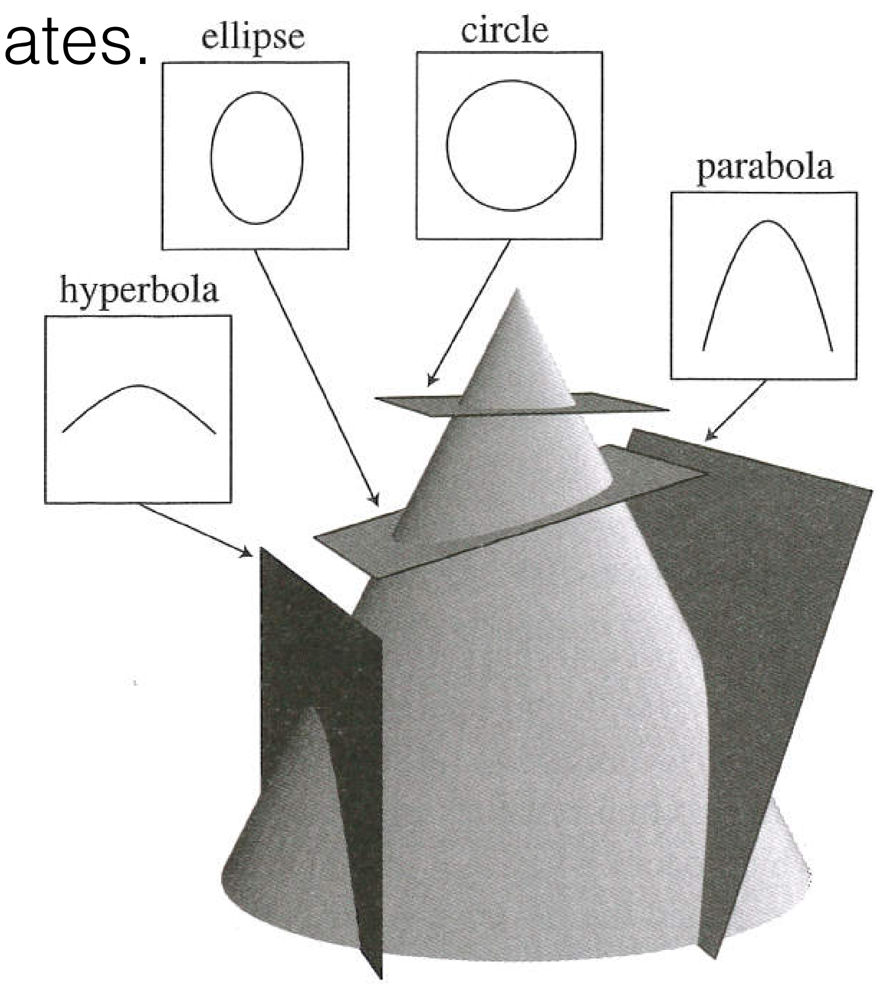
Murray & Dermott “Solar System Dynamics”

The two-body problem - Relative orbit

$$r = \frac{p}{1 + e\,cos(\theta - \varpi)}$$

where ϖ and e are integration constants and $p = h^2/\mu$

General equation of a conic in polar coordinates.



circle $e = 0; \quad p = a$

ellipse $0 < e < 1; \quad p = a(1 - e^2)$

parabola $e = 1; \quad p = 2q$

hyperbola $e > 1; \quad p = a(e^2 - 1)$

Murray & Dermott “Solar System Dynamics”

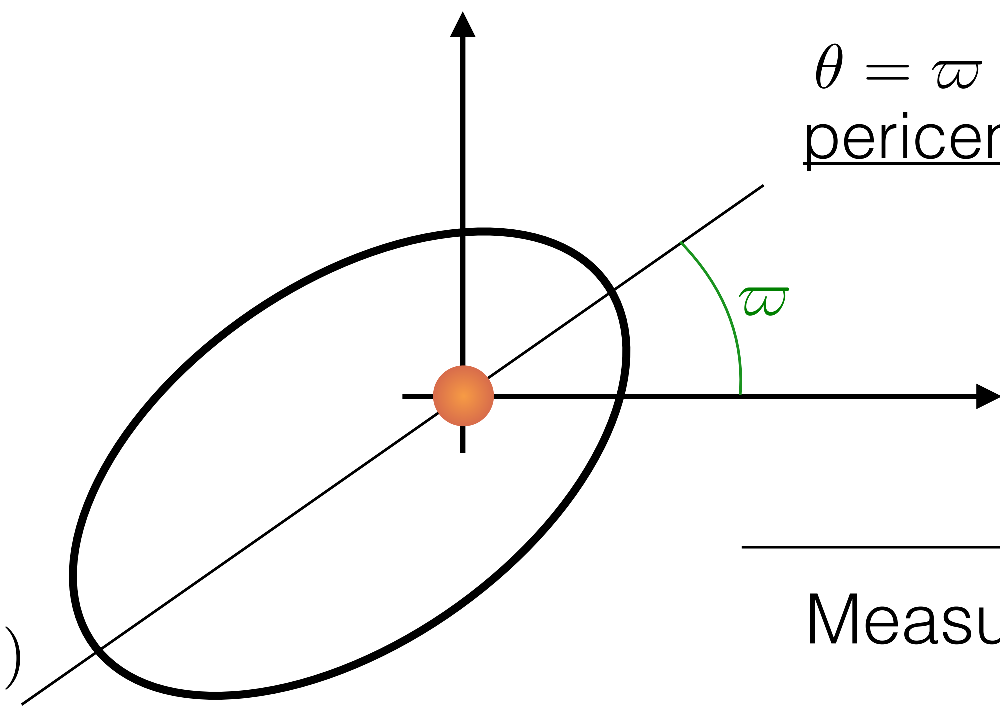
The two-body problem - Elliptical orbits

$$r = \frac{a(1-e^2)}{1 + e\,cos(\theta - \varpi)}$$

Maximum and minimum distances:

$\theta = \varpi$

<u>pericentre</u> $\quad r_{\rm peri} = a(1-e)$



$\theta = \varpi + \pi$

<u>apocentre</u>

$r_{\rm apo} = a(1+e)$

Measure angles from pericentre.

Define true anomaly: $\nu = \theta - \varpi$

$$r = \frac{a(1-e^2)}{1 + e\,cos(\nu)}$$

Murray & Dermott "Solar System Dynamics"

The two-body problem - Kepler's laws

$$h = r^2\dot{\theta} = \sqrt{\mu a(1-e^2)} = \text{constant}$$

It can be shown that the area swept by m_2 around m_1 per unit time is proportional to h.

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{1}{2}h$$ **Second Law of Kepler**

The area swept over a period P is the area of the ellipse $A = \pi ab$

$$\frac{\pi ab}{P} = \frac{1}{2}h \longrightarrow \frac{\pi^2 a^4(1-e^2)}{P^2} = \frac{1}{4}\mu a(1-e^2)$$

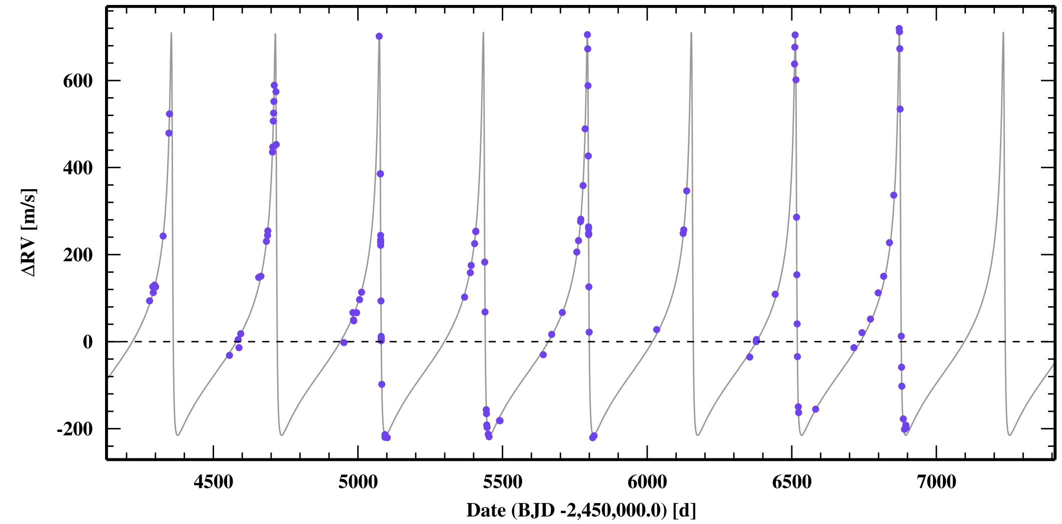
$$P^2 = \frac{4\pi^2}{\mu}a^3$$ **Third Law of Kepler**

Murray & Dermott "Solar System Dynamics"

Where is time?

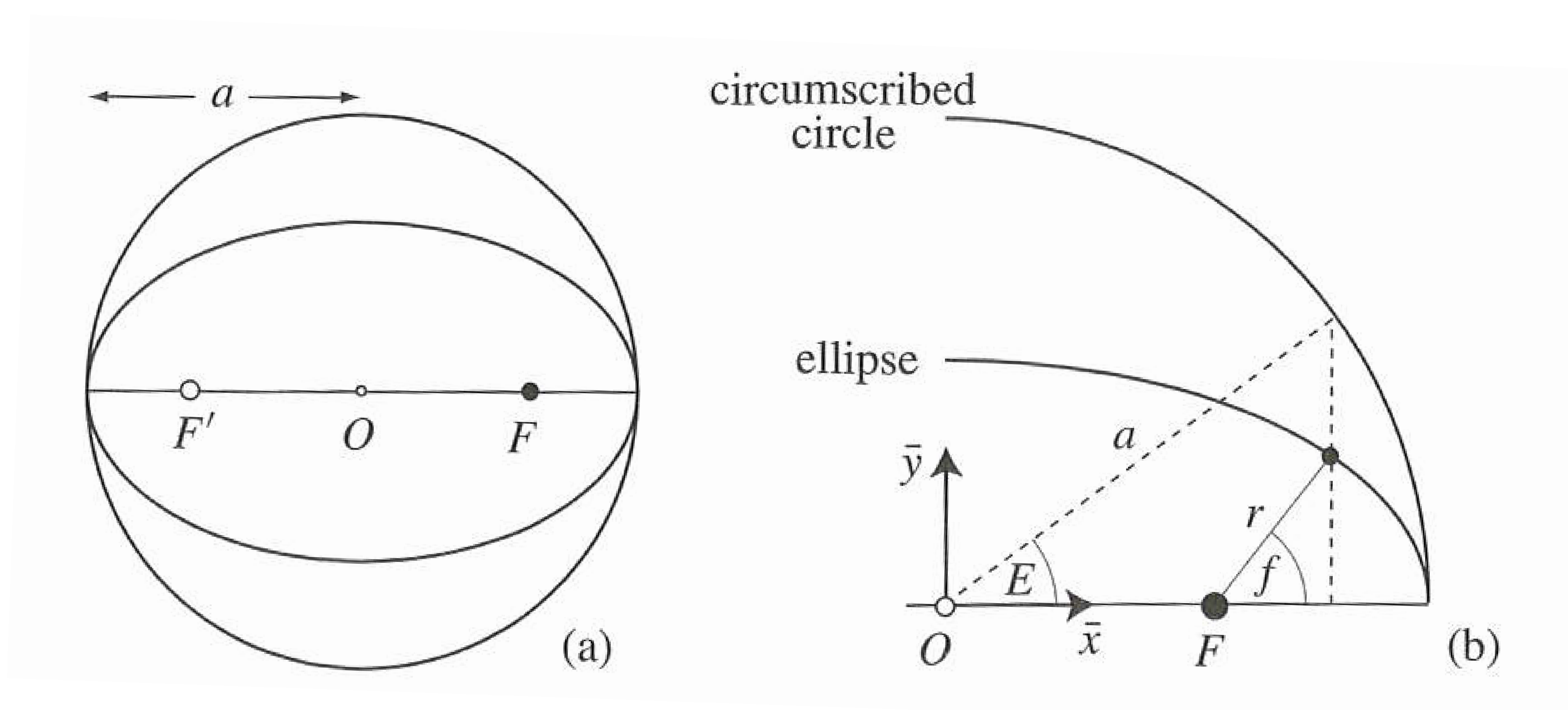
$$r = \frac{a(1-e^2)}{1+e\,cos(\nu)}$$

???



Mean anomaly $M = \frac{2\pi}{P}(t-\tau)$

τ Time of periastron passage



Murray & Dermott “Solar System Dynamics”

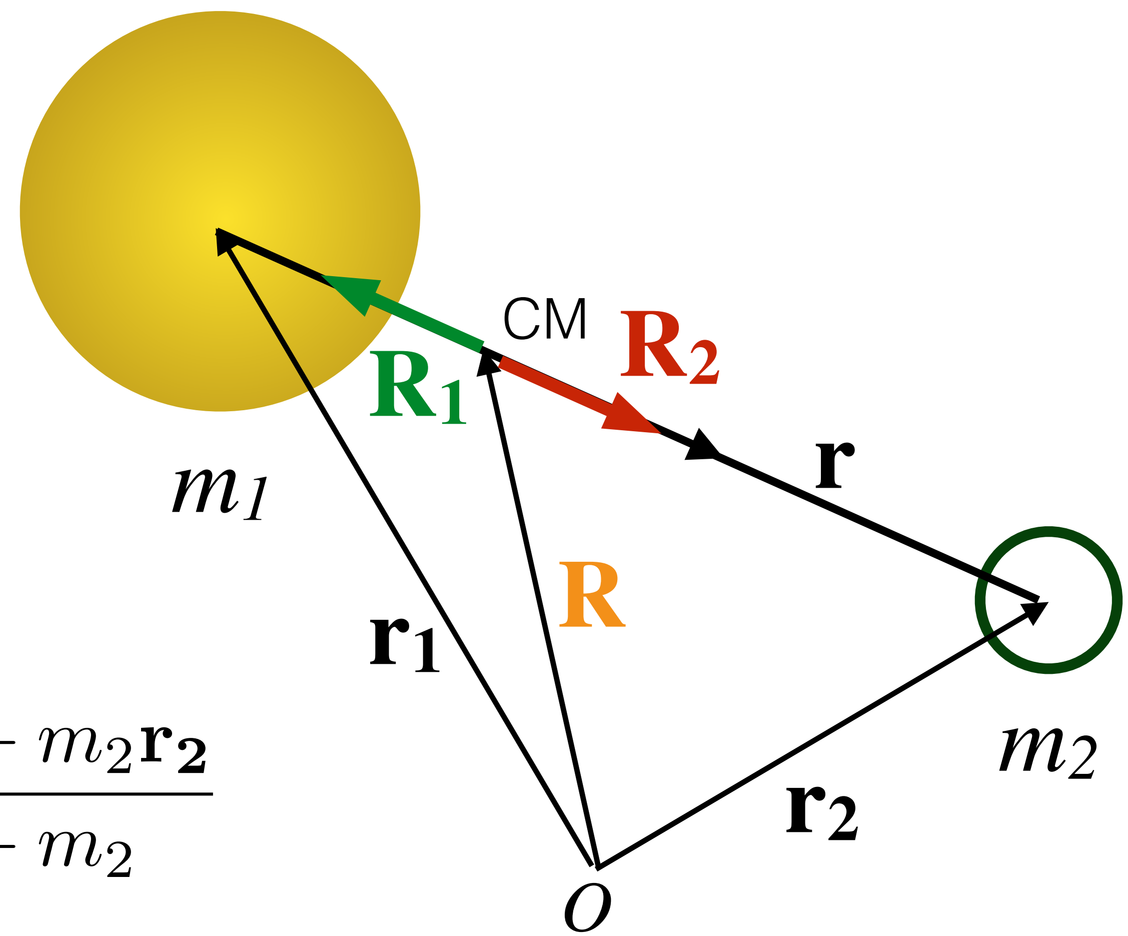
Where is time?

Mean anomaly	$M = \frac{2\pi}{P}(t - \tau)$	τ Time of periastron passage
Kepler's equation	$M = E - e\sin(E)$	Transcendental; requiere iterative methods to solve.
Eccentric to true anomaly	$\tan\left(\frac{\nu}{2}\right) = \sqrt{\frac{1+e}{1-e}}\tan\left(\frac{E}{2}\right)$	

$$\boxed{r = \frac{a(1-e^2)}{1+e\,cos(\nu)}}$$

Murray & Dermott "Solar System Dynamics"

The two-body problem - Barycentric orbit



remember

$$\mathbf{R} = \frac{m_1\mathbf{r_1} + m_2\mathbf{r_2}}{m_1 + m_2}$$

new vectors of interest

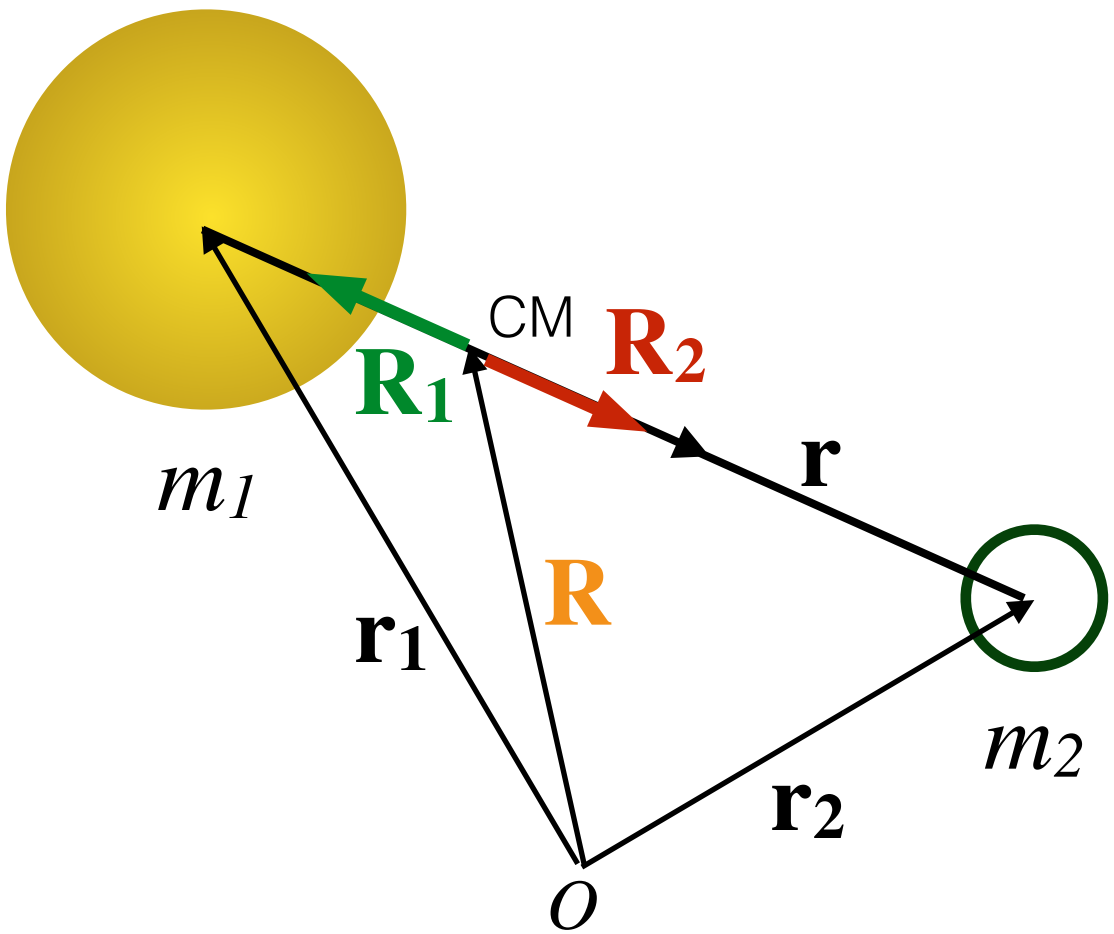
$$\mathbf{R_1} = \mathbf{r_1} - \mathbf{R} \qquad \mathbf{R_2} = \mathbf{r_2} - \mathbf{R}$$

$$m_1\mathbf{R_1} + m_2\mathbf{R_2} = 0$$

vectors lie in the same line, with opposite directions.

Murray & Dermott "Solar System Dynamics"

The two-body problem - Barycentric orbit



$$\mathbf{R_1} = \mathbf{r_1} - \mathbf{R}$$

$$\mathbf{R_2} = \mathbf{r_2} - \mathbf{R}$$

$$m_1\mathbf{R_1} + m_2\mathbf{R_2} = 0$$

we can write *scalar* equations

$$R_1 = \frac{m_2}{m_1 + m_2} r \qquad R_2 = \frac{m_1}{m_1 + m_2} r$$

Question: what would be the barycentric motion of m_1 and m_2 look like?

Murray & Dermott "Solar System Dynamics"

The two-body problem - Barycentric orbit

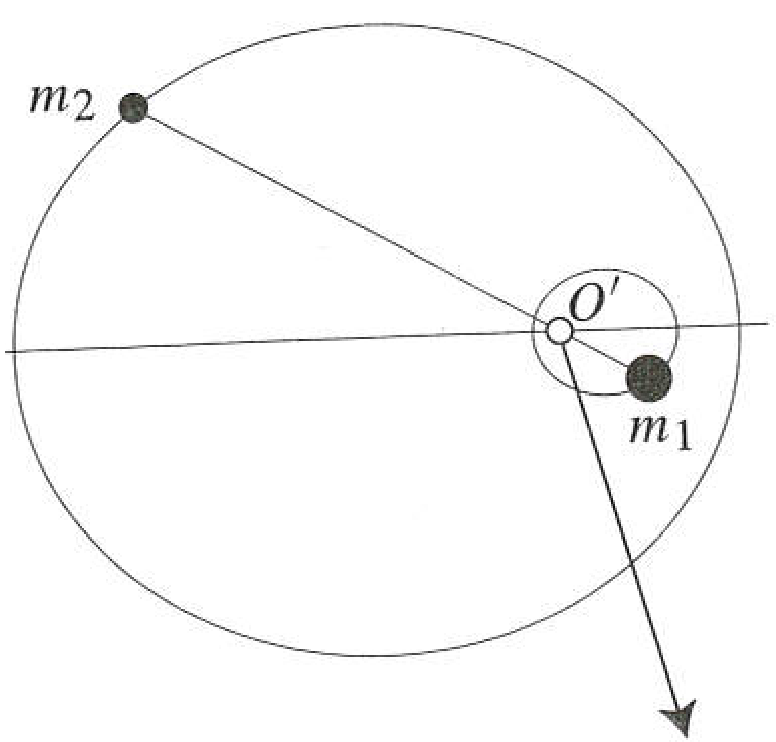
Question: what would be the barycentric motion of m_1 and m_2 look like?

$$R_1 = \frac{m_2}{m_1 + m_2} r \qquad R_2 = \frac{m_1}{m_1 + m_2} r$$

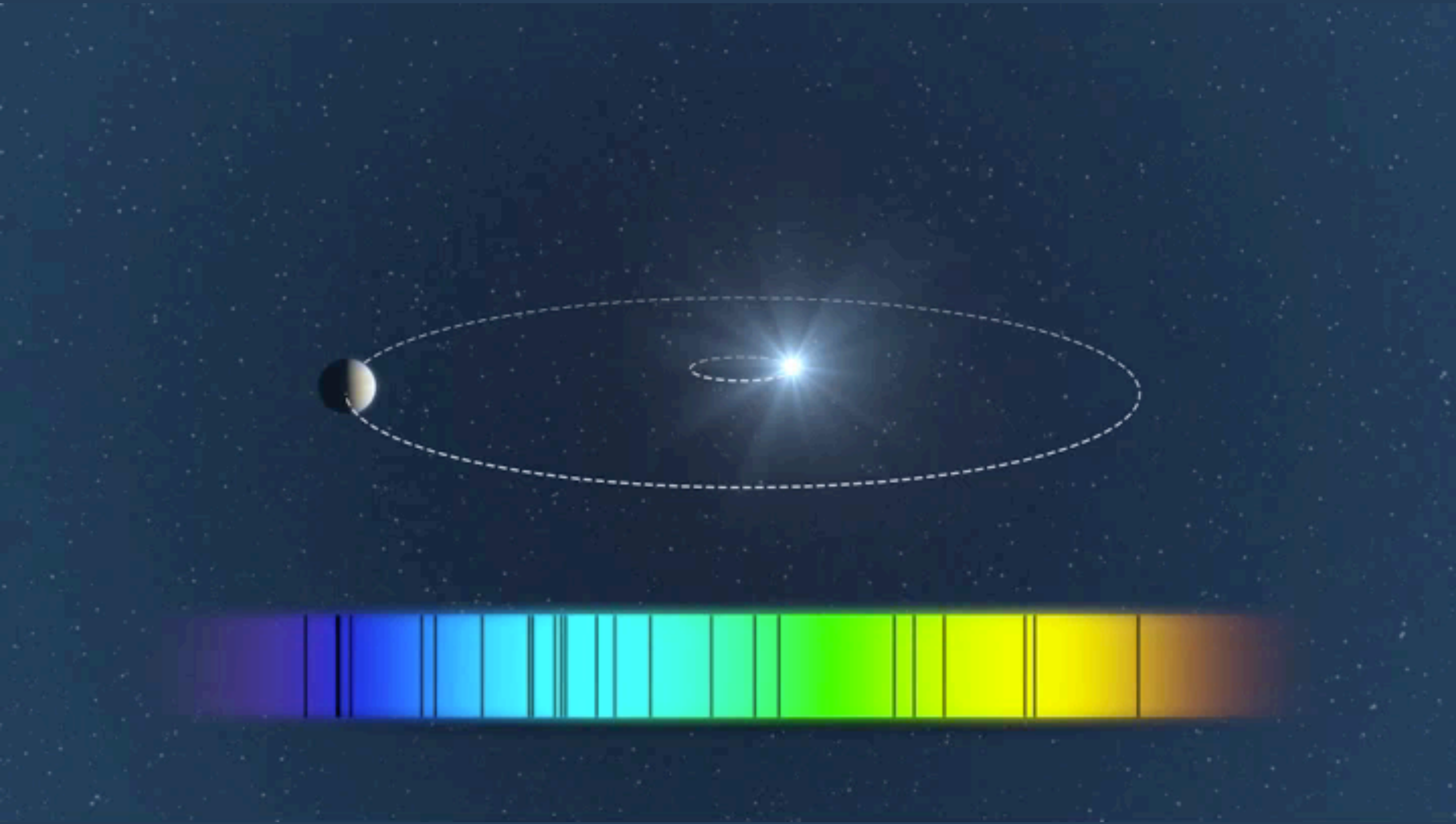
Answer: the orbits are a scaled-down version of the relative orbit.

$$a_1 = \frac{m_2}{m_1 + m_2} a \qquad a_2 = \frac{m_1}{m_1 + m_2} a$$

the orbital period and eccentricities are identical, but the arguments of pericentre differ by 180 degrees.

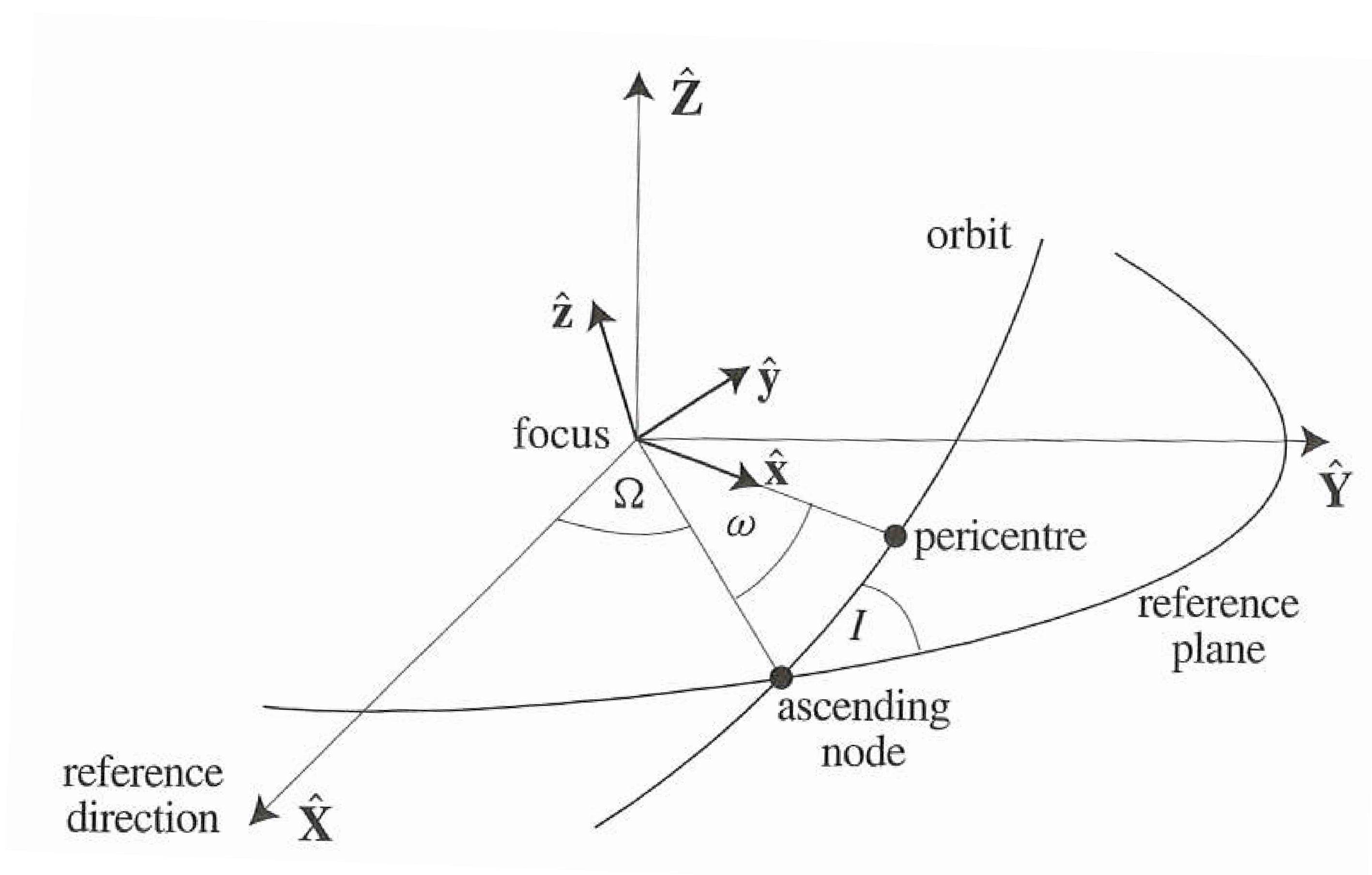


Murray & Dermott "Solar System Dynamics"



The two-body problem - 3D orbits

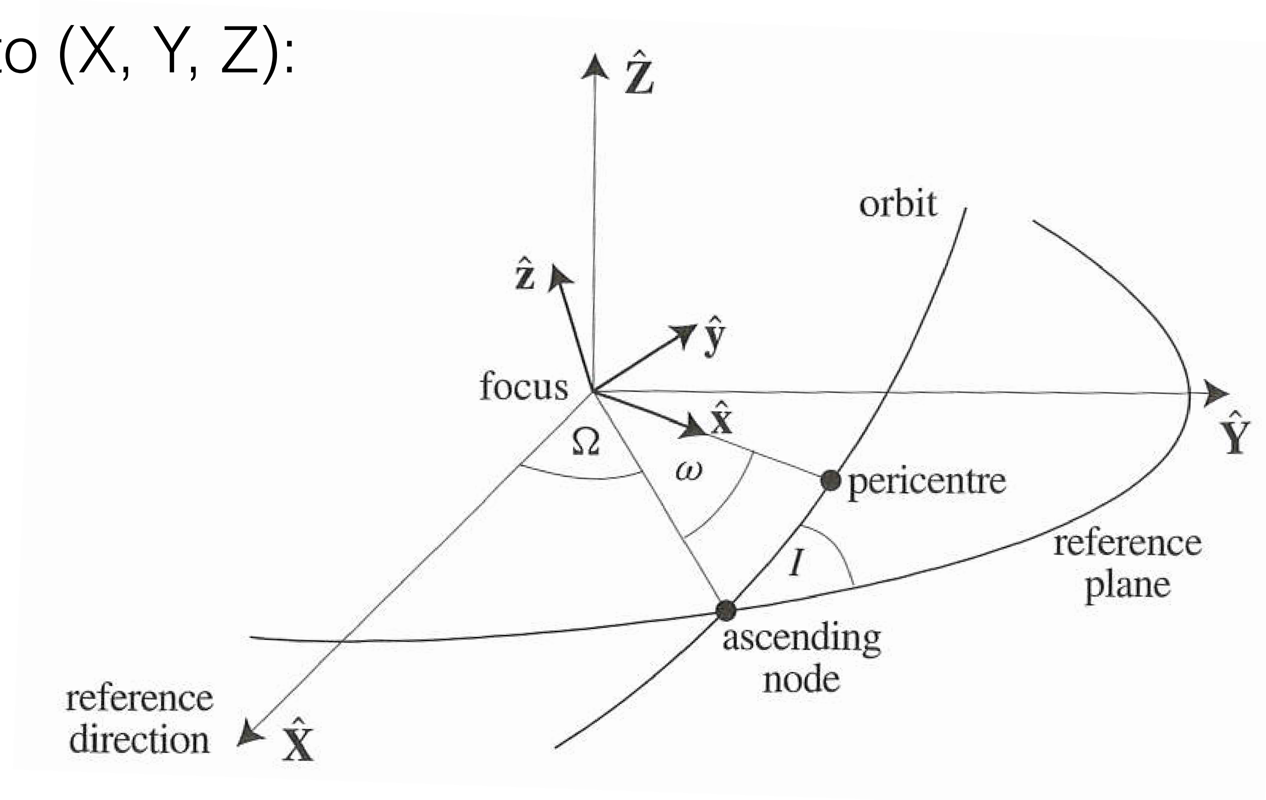
- The motion of the bodies is confined to a plane ($z = 0$).
- Want to transform to reference system (X, Y, Z).
- Three rotations to go from (x, y, z) to (X, Y, Z):
  - ✳ A rotation around z.
  - ✳ A rotation around x.
  - ✳ A rotation around z.



Murray & Dermott "Solar System Dynamics"

The two-body problem - 3D orbits

- Three rotations to go from (*x, y, z*) to (X, Y, Z):
  - ✳ A rotation around z.
  - ✳ A rotation around the new axis x.
  - ✳ A rotation around the new axis z.



$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \mathbf{P_3 P_2 P_1} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$\mathbf{P_1} = \begin{pmatrix} \cos\omega & -\sin\omega & 0 \\ \sin\omega & \cos\omega & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{P_2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos I & -\sin I \\ 0 & \sin I & \cos I \end{pmatrix} \quad \mathbf{P_3} = \begin{pmatrix} \cos\Omega & -\sin\Omega & 0 \\ \sin\Omega & \cos\Omega & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Murray & Dermott “Solar System Dynamics”

The two-body problem - 3D orbits

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \mathbf{P_3 P_2 P_1} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$\mathbf{P_1} = \begin{pmatrix} \cos\omega & -\sin\omega & 0 \\ \sin\omega & \cos\omega & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{P_2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos I & -\sin I \\ 0 & \sin I & \cos I \end{pmatrix} \quad \mathbf{P_3} = \begin{pmatrix} \cos\Omega & -\sin\Omega & 0 \\ \sin\Omega & \cos\Omega & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$r \begin{pmatrix} \cos(\nu+\omega)\cos\Omega - \sin\Omega\sin(\nu+\omega)\cos I \\ \cos(\nu+\omega)\sin\Omega + \cos\Omega\sin(\nu+\omega)\cos I \\ \sin(\nu+\omega)\sin I \end{pmatrix} = \mathbf{P_3 P_2 P_1} \begin{pmatrix} r\cos\nu \\ r\sin\nu \\ 0 \end{pmatrix}$$

Murray & Dermott “Solar System Dynamics”

The two-body problem - 3D orbits

None of the methods are sensitive to the orientation of the orbit in the plane of the sky.

We can therefore simplify by considering the X axis is coincident with the line of nodes ($\Omega = 180$ deg).

$$r \begin{pmatrix} \cos(\nu+\omega)\cos\Omega - \sin\Omega\sin(\nu+\omega)\cos I \\ \cos(\nu+\omega)\sin\Omega + \cos\Omega\sin(\nu+\omega)\cos I \\ \sin(\nu+\omega)\sin I \end{pmatrix} = \mathbf{P_3 P_2 P_1} \begin{pmatrix} \boldsymbol{r\cos\nu} \\ \boldsymbol{r\sin\nu} \\ \mathbf{0} \end{pmatrix}$$

$$\begin{aligned} X &= -r\cos(\nu+\omega) \\ Y &= -r\sin(\nu+\omega)\cos I \\ Z &= r\sin(\nu+\omega)\sin I \end{aligned}$$

$$\boxed{r = \frac{a(1-e^2)}{1+e\,cos(\nu)}}$$

Murray & Dermott "Solar System Dynamics"

The two-body problem - RV amplitude

Now, we are interested in the motion of the star with respect to the centre of mass of the system. Substituting a by a_1

$$Z_\star = \left(\frac{m_2}{m_1 + m_2} a \right) \frac{(1 - e^2)}{1 + e \cos \nu} \sin(\nu + \omega) \sin I$$

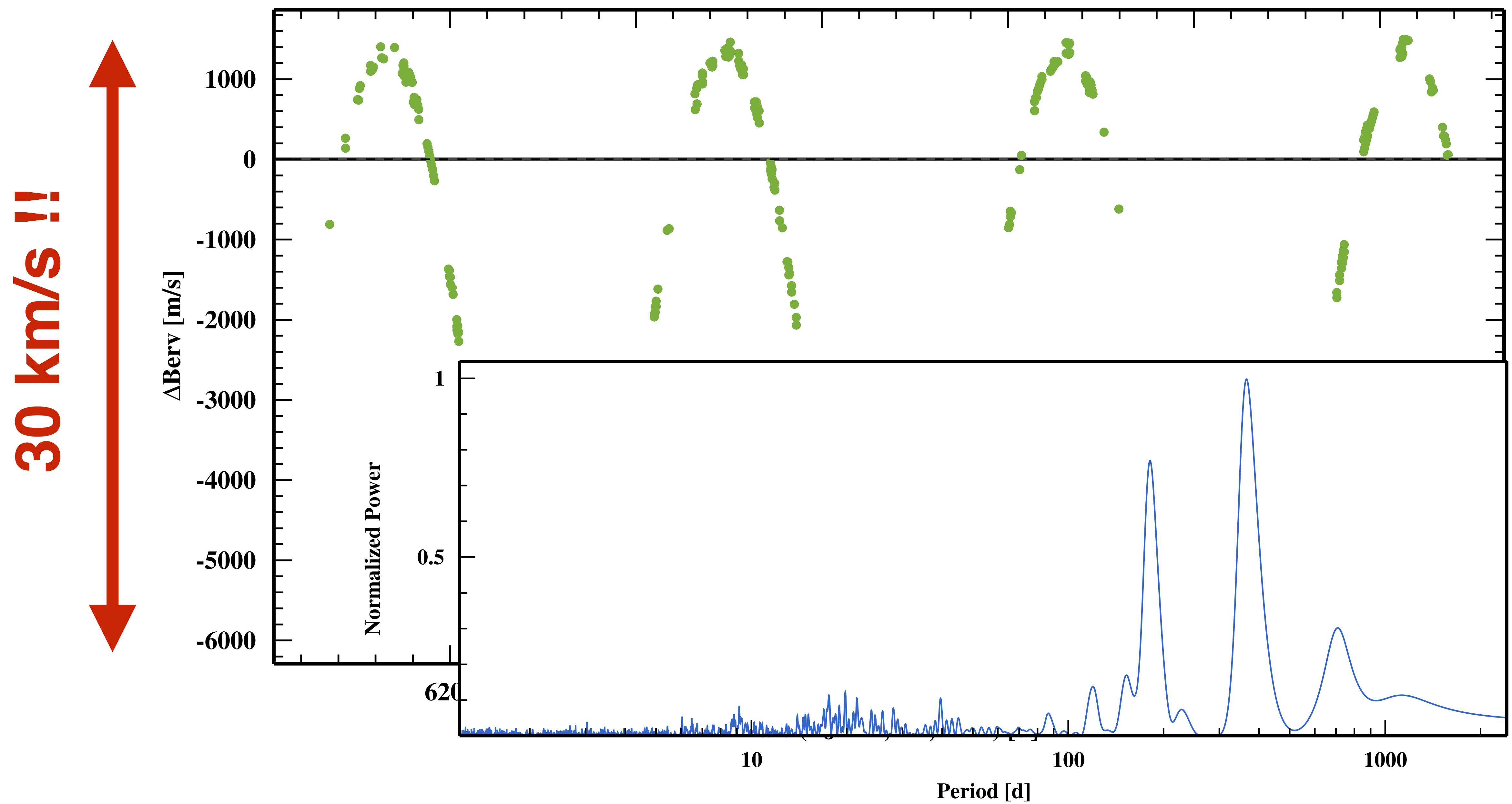
Deriving and using the fact that $\dot{\nu} = \dot{\theta} = h/r^2 = \frac{2\pi}{P} \frac{a^2 \sqrt{1 - e^2}}{r^2}$

$$V_\star = V_0 + K_\star \left[\cos(\nu + \omega) + e \cos \omega \right]$$

$$K_\star = \left(\frac{2\pi G}{P} \right)^{1/3} \frac{1}{\sqrt{1 - e^2}} \frac{m_2 \sin I}{(m_1 + m_2)^{2/3}}$$

Real life nuisance - the Earth moves

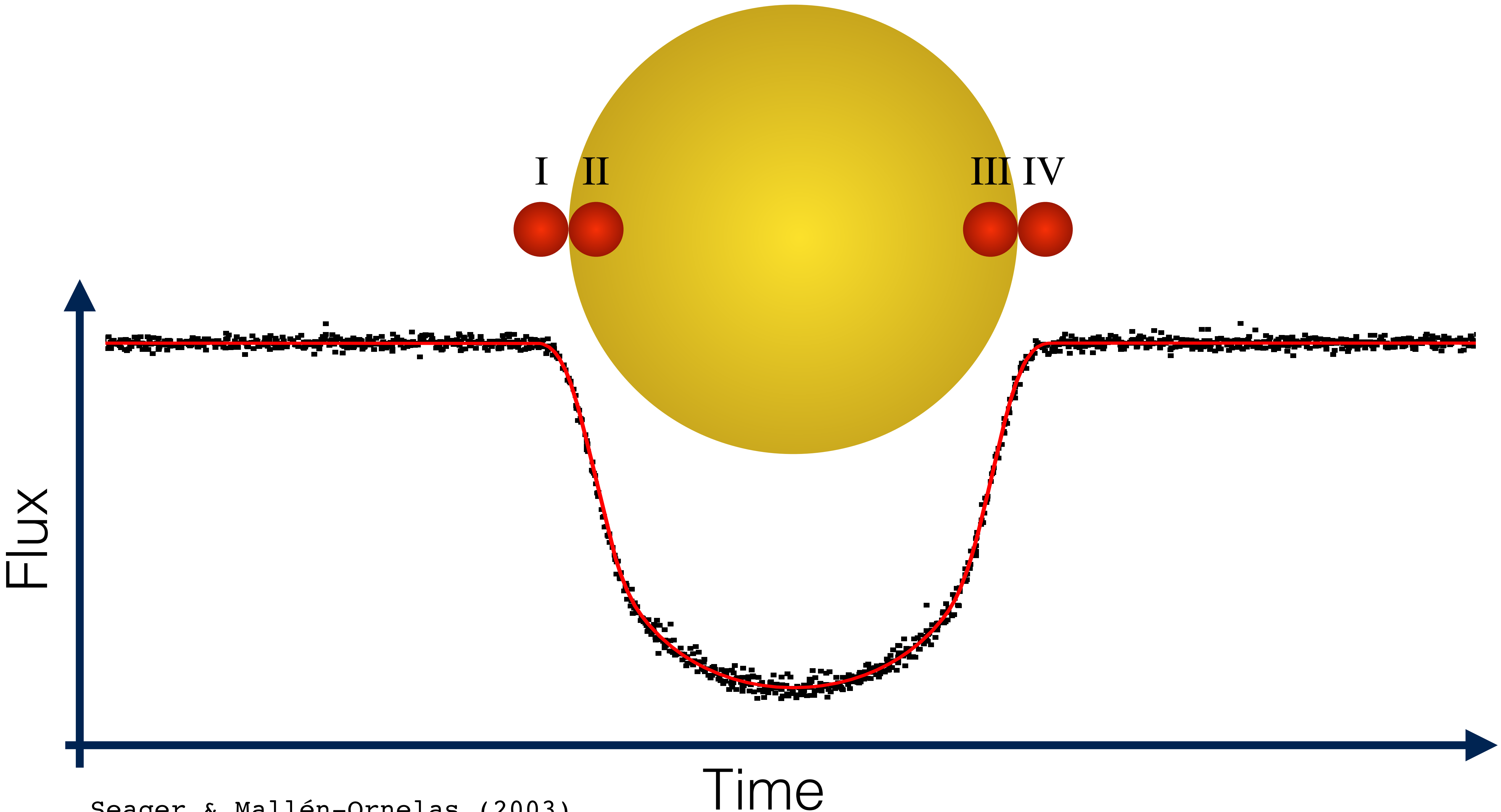
One of the most stable HARPS stars.



Physical Models

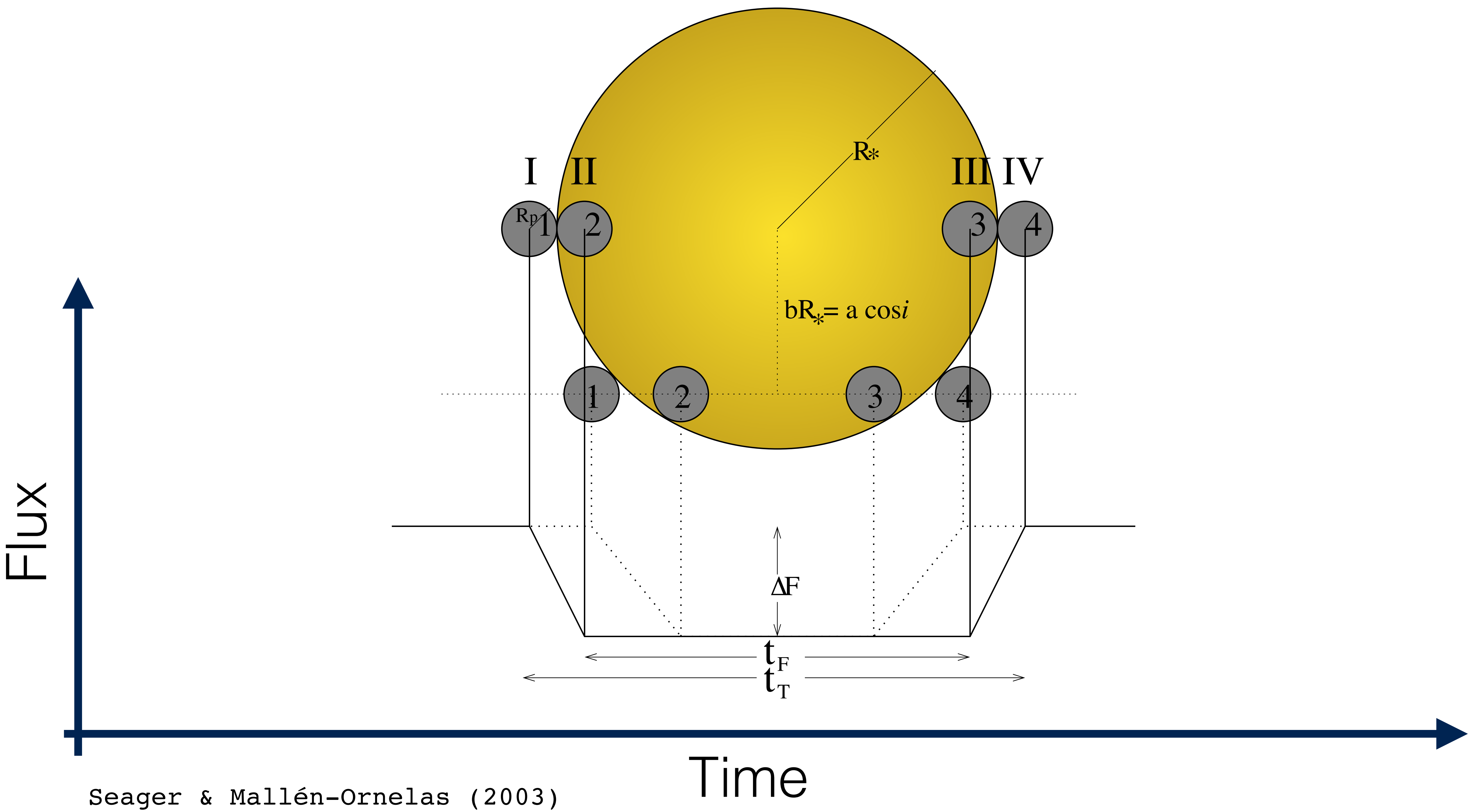
Transits

Transiting planets



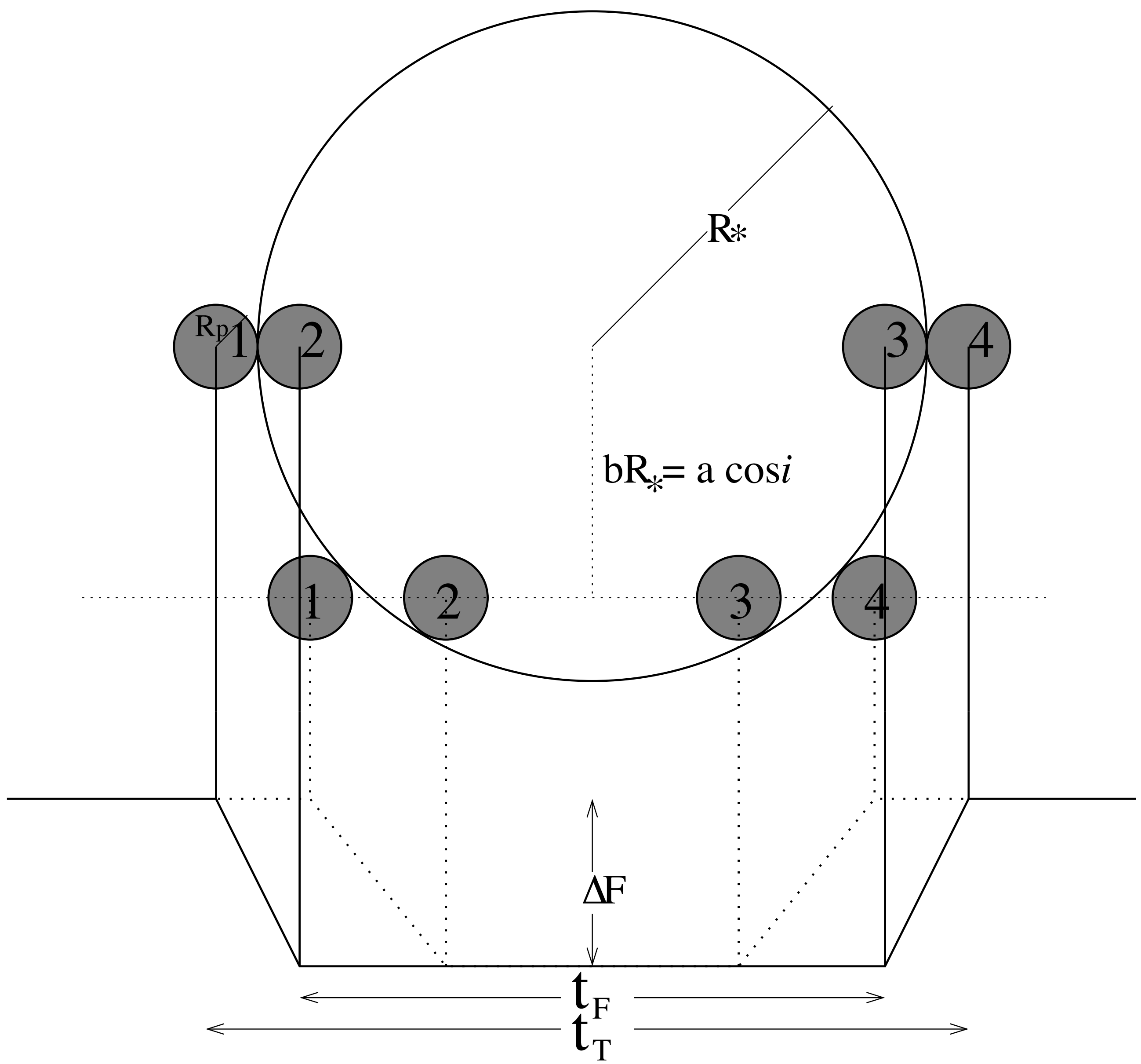
Seager & Mallén-Ornelas (2003)
Mandel & Agol (2002)

Transiting planets - transit geometry



Seager & Mallén-Ornelas (2003)
Mandel & Agol (2002)

Transiting planets - transit geometry



Seager & Mallén-Ornelas (2003)
Mandel & Agol (2002)

A **four**-parameter model

$$P, \Delta\mathrm{F}, t_{\mathrm{T}}, t_{\mathrm{F}}$$

In principle, **seven** physical parameters

$$M_\star, R_\star$$
$$M_p, R_p$$
$$i, e, \varpi$$

Only some **combinations** of these are obtainable from the light curve alone (degeneracy).

Seager & Mallén-Ornelas (2003)

Transiting planets - transit geometry

In the circular case, these **combinations** are:

$$\frac{R_p}{R_\star} = \sqrt{\Delta F}$$

$$b = \frac{a}{R_\star}\cos i = \left\{ \frac{\left(1-\sqrt{\Delta F}\right)^2 - \sin^2(t_F\pi/P)/\sin^2(t_T\pi/P)\left(1+\sqrt{\Delta F}\right)}{1-\sin^2(t_F\pi/P)/\sin^2(t_T\pi/P)} \right\}^{1/2}$$

$$\frac{a}{R_\star} = \left\{ \frac{\left(1+\sqrt{\Delta F}\right)^2 - b^2\left[1-\sin^2(t_T\pi/P)\right]}{\sin^2\left(t_T\pi/P\right)} \right\}$$

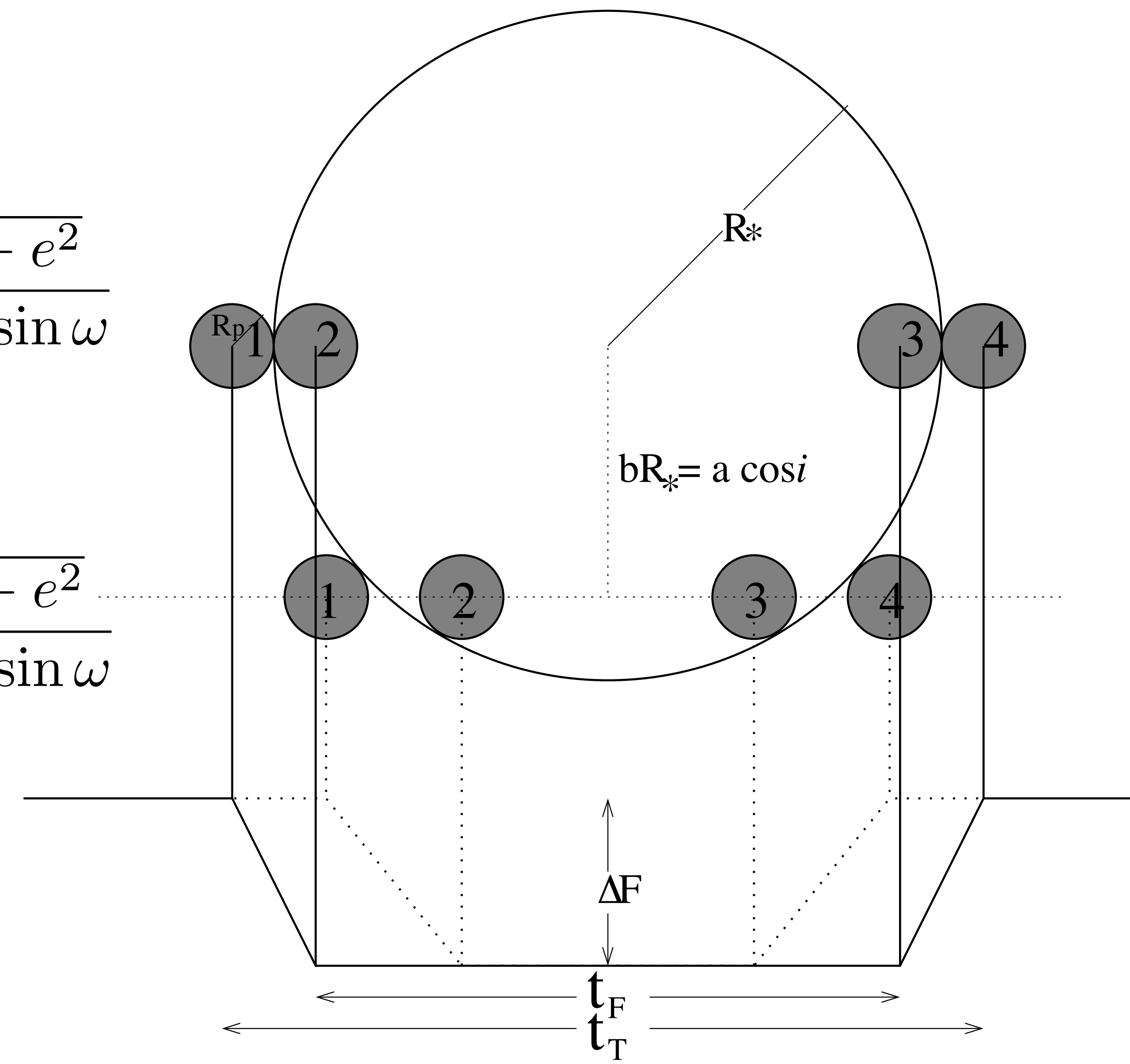
Seager & Mallén-Ornelas (2003)

Transiting planets - eccentricity effects

Complicated algebra arises when eccentricity is included. Inversion is no longer easy, but the forward equations are given, approximately, by Winn (2008)

$$t_T = \frac{P}{\pi}\sin^{-1}\left[\frac{R_\star}{a}\frac{\sqrt{(1+R_p/R_\star)^2-b^2}}{\sin i}\right]\frac{\sqrt{1-e^2}}{1+e\sin\omega}$$

$$t_F = \frac{P}{\pi}\sin^{-1}\left[\frac{R_\star}{a}\frac{\sqrt{(1-R_p/R_\star)^2-b^2}}{\sin i}\right]\frac{\sqrt{1-e^2}}{1+e\sin\omega}$$



Winn (2008)

Transiting planets - eccentricity effects

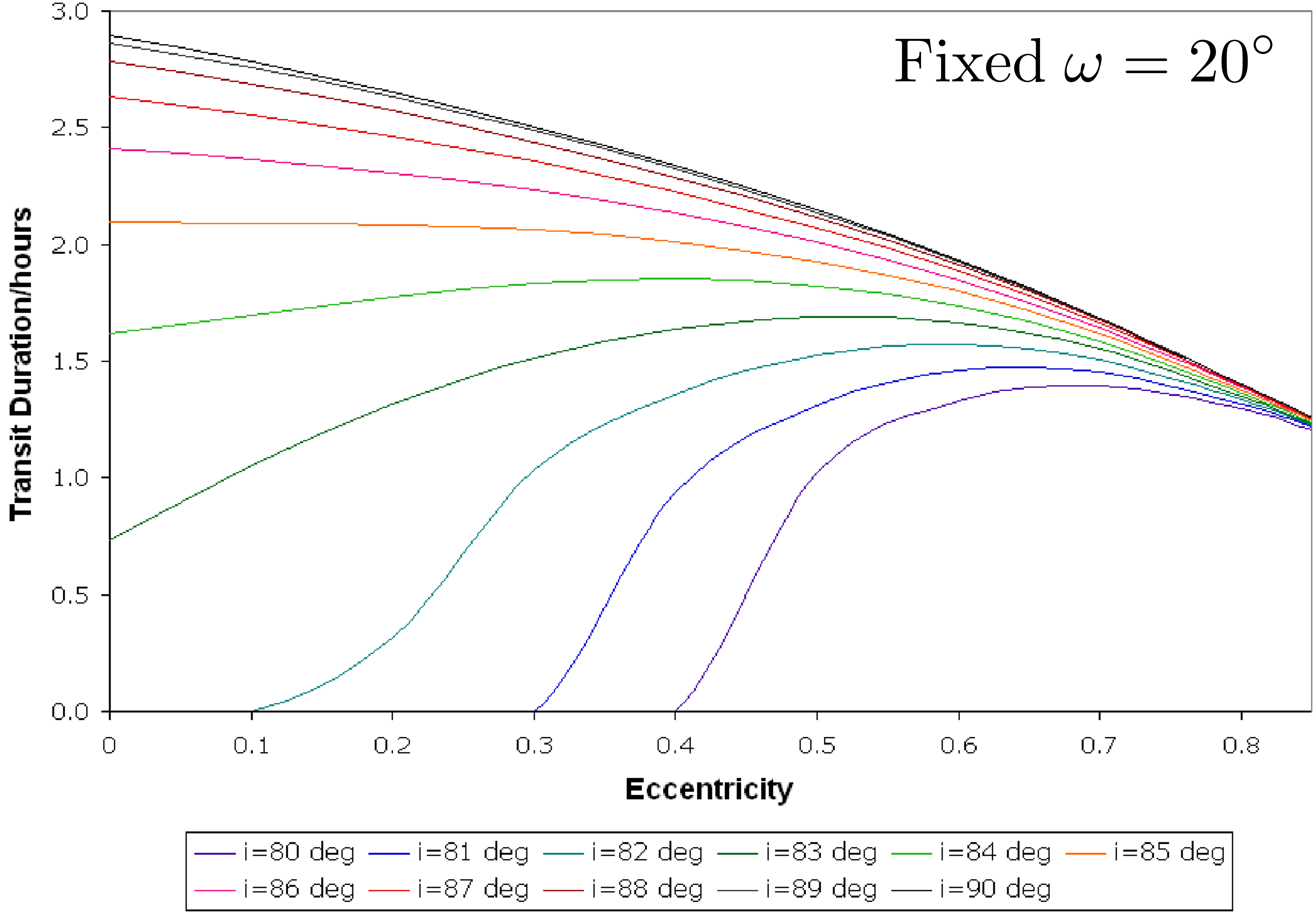
Inversion is possible for non-grazing transits under the condition $R_p \ll R_\star \ll a$

$$b^2 = \frac{\left(1-\sqrt{\Delta F}\right)^2 - (t_F/t_T)^2\left(1+\sqrt{\Delta F}\right)^2}{1-(t_F/t_T)^2}$$

$$\frac{R_\star}{a} = \frac{\pi}{2\Delta F^{1/4}}\frac{\sqrt{t_T^2 - t_F^2}}{P}\left(\frac{1+e\sin\omega}{\sqrt{1-e^2}}\right)$$

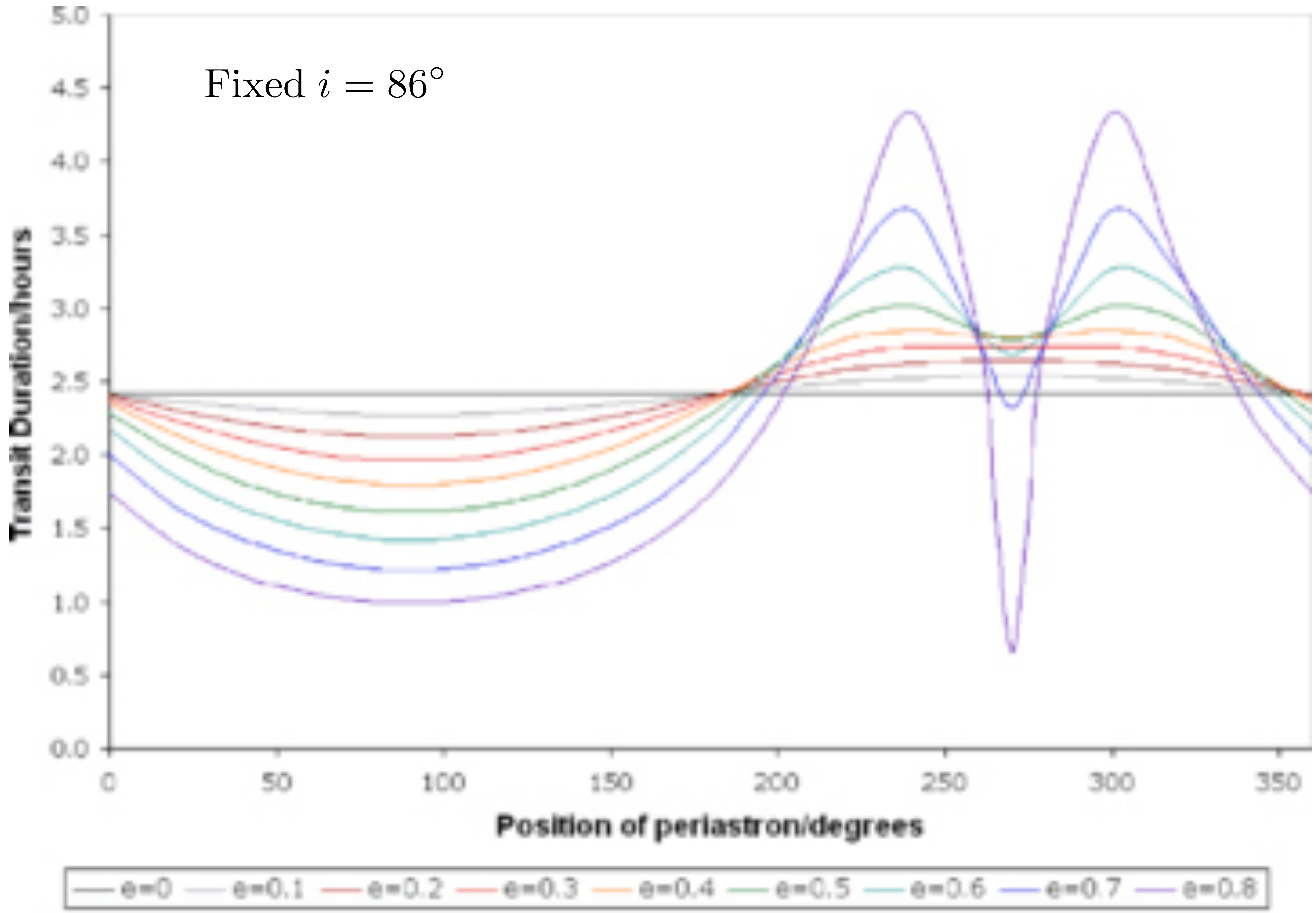
Winn (2008)

Transiting planets - eccentricity effects



Kipping (2008)

Transiting planets - eccentricity effects



Kipping (2008)

Transiting planets - stellar bulk density

Directly from Kepler's third law.

$$\rho_* + \left(\frac{R_p}{R_*}\right)^3 \rho_p = \frac{3\pi}{\mathrm{G}P^2}\left(\frac{a}{R_*}\right)^3$$

Importantly, in the common case that $R_p/R_* << 1$:

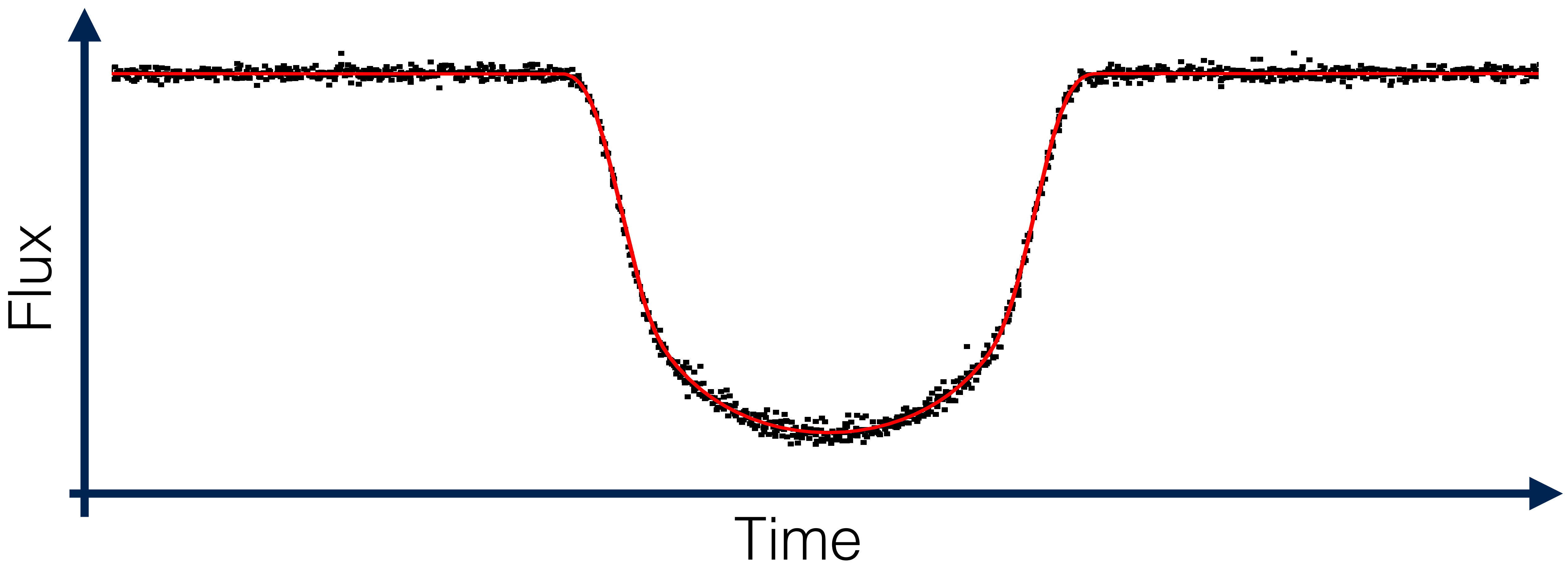
$$\rho_\star \sim \frac{3\pi}{GP^2}\left(\frac{a}{R_\star}\right)^3$$

ρ_* obtained from transit geometry.

Caveat: **orbital eccentricity.**

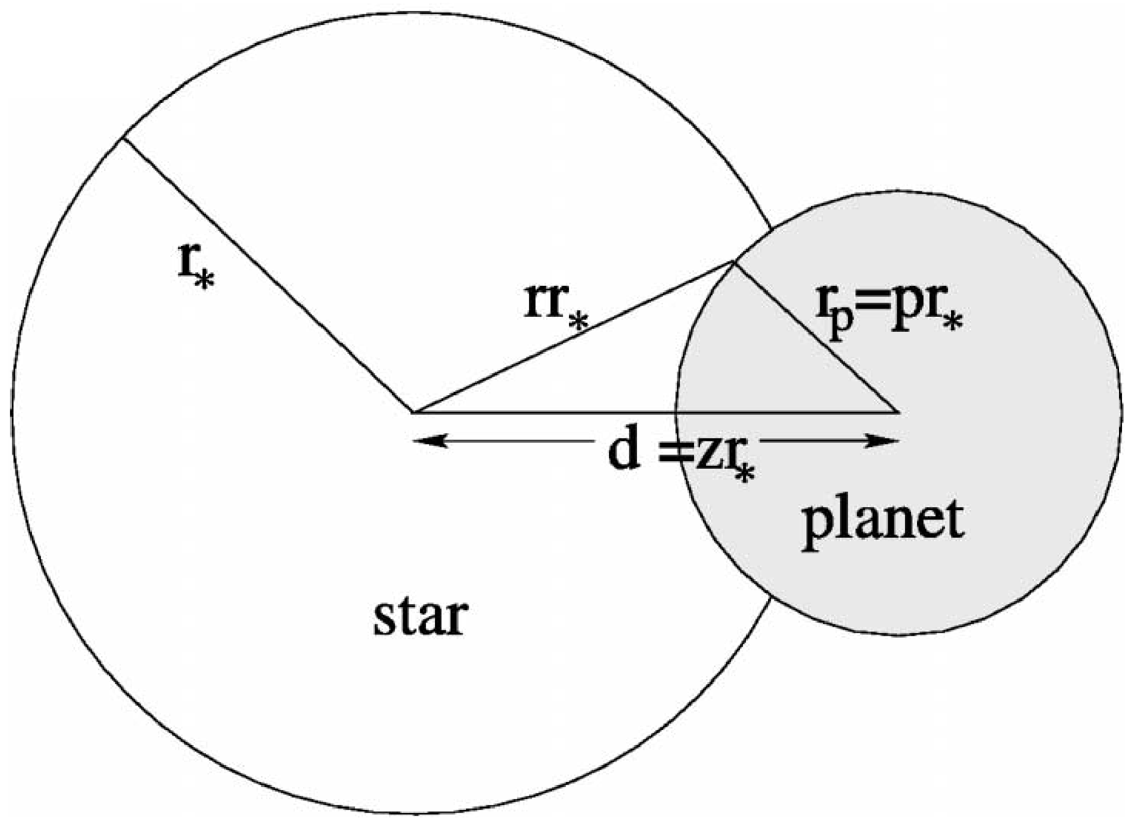
Seager & Mallén-Ornelas (2003)

Analytical light curve expression



$$flux = f(\text{time})$$

Analytical light curve expression

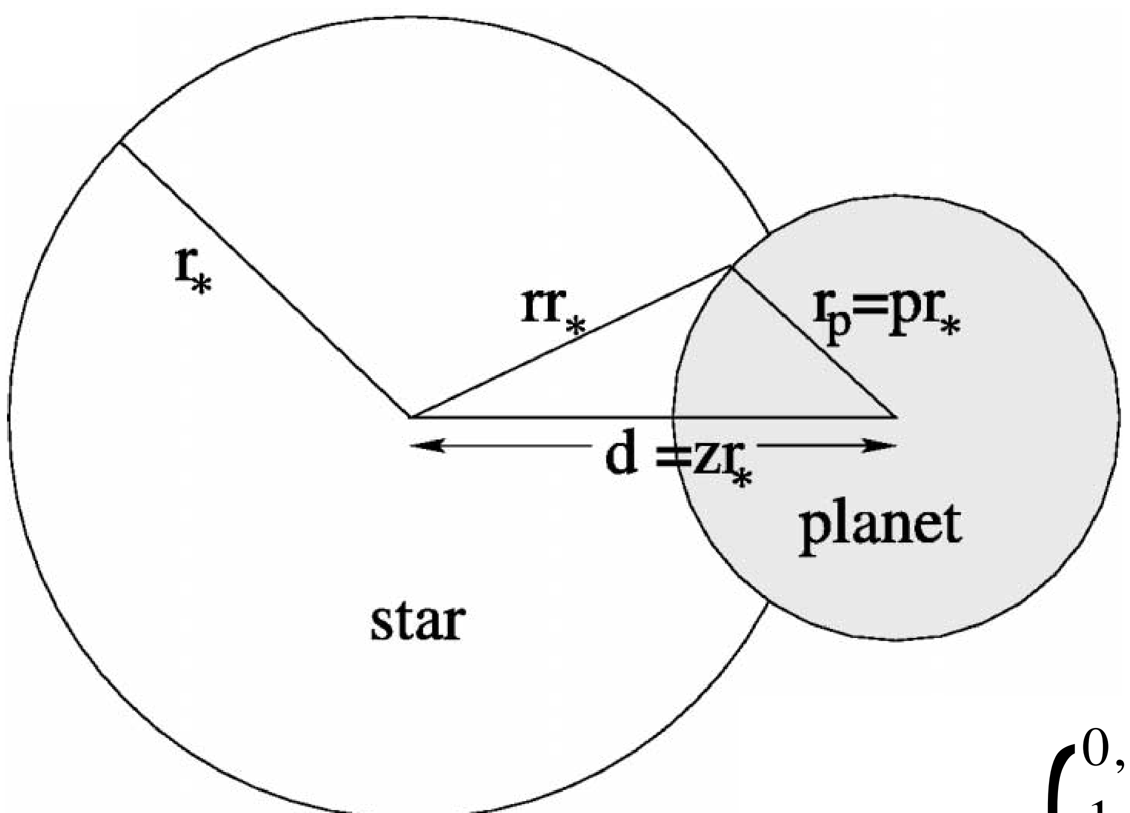


Mandel & Agol (2002)

$$flux = f(z)$$
$$z = d/R_\star$$
$$d = g(t)$$

Analytical expressions for the light loss as a function of z and $p = R_p/R_s$ are given by **Mandel & Agol (2002)**, for a uniform source and for nonlinear limb darkening laws (see also **Giménez 2006**).

Analytical light curve expression



Uniform source

$$F(p,z) = 1 - \lambda(p,z)$$

$$\lambda^e(p,z) = \begin{cases} 0, & 1+p<z, \\ \frac{1}{\pi}\left[p^2\kappa_0 + \kappa_1 - \sqrt{\frac{4z^2-(1+z^2-p^2)^2}{4}}\right], & |1-p|<z\leq 1+p, \\ p^2, & z\leq 1-p, \\ 1, & z\leq p-1, \end{cases}$$

$$\kappa_0 = cos^{-1}\left[\left(p^2+z^2-1\right)/2\,pz\right]$$
$$\kappa_1 = cos^{-1}\left[\left(1-p^2+z^2\right)/2z\right]$$

Mandel & Agol (2002)

Computation of d as a function of time.

Back to Cartesian description of relative orbit

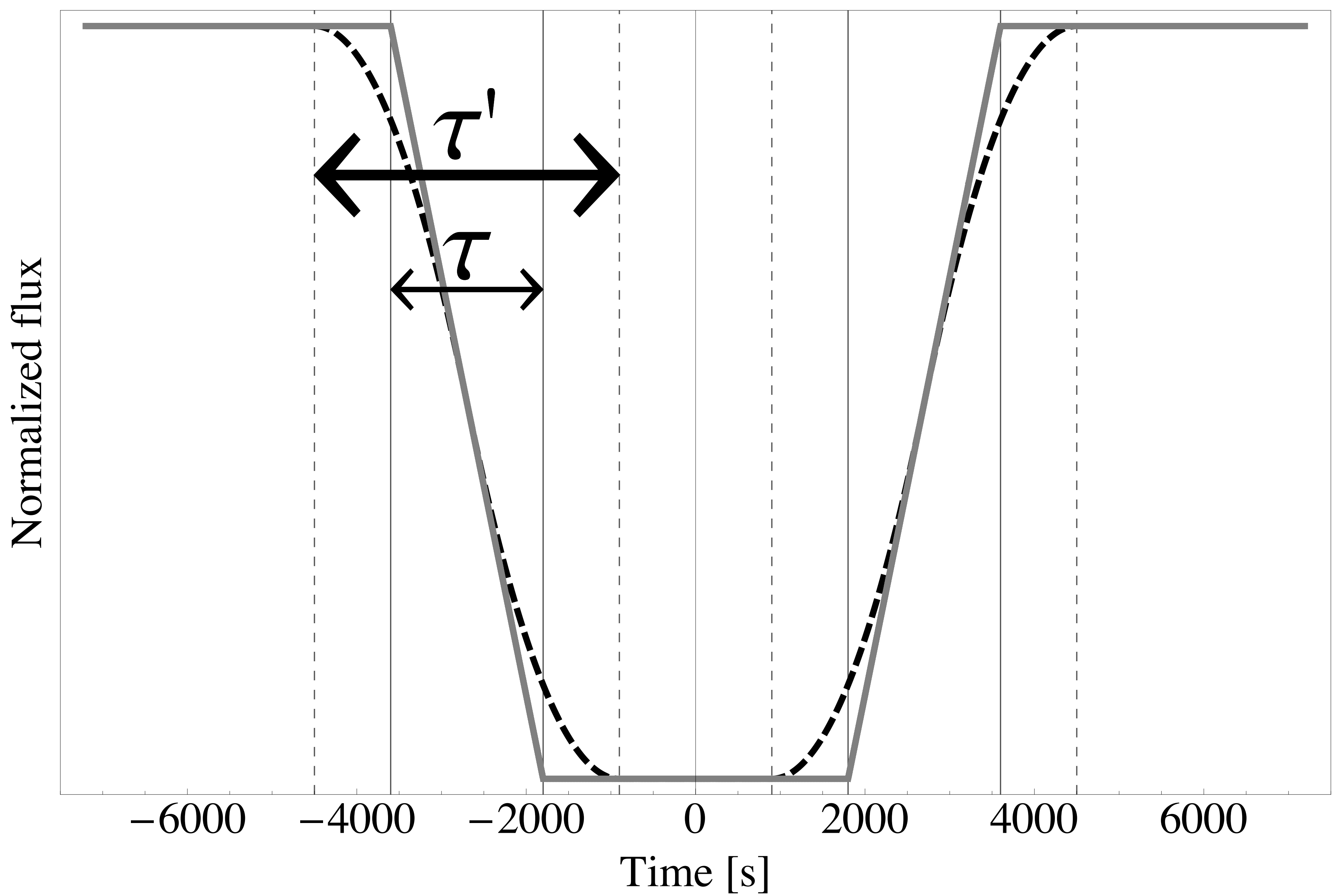
$$
\begin{aligned}
X &= -r\cos(\nu+\omega) \\
Y &= -r\sin(\nu+\omega)\cos I \\
Z &= r\sin(\nu+\omega)\sin I
\end{aligned}
$$

$$r = \frac{a(1-e^2)}{1+e\,cos(\nu)}$$

Planet-to-star centre distance is easily written as

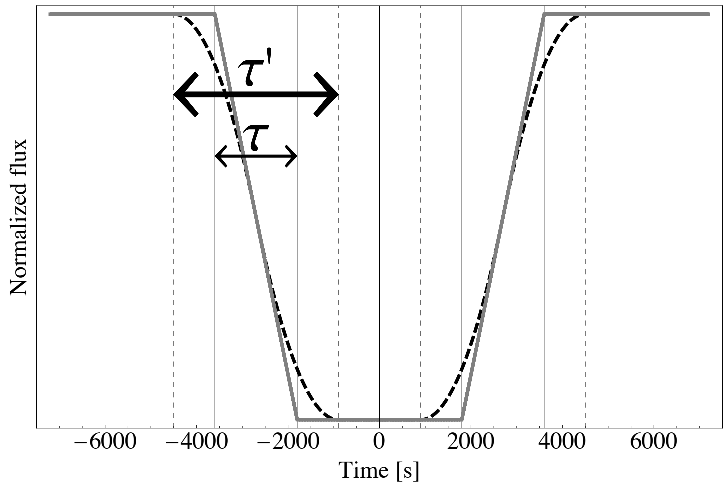
$$d = \sqrt{X^2+Y^2} = \frac{a\left(1-e^2\right)}{1+e\cos\nu}\sqrt{1-sin^2(\nu+\omega)\sin^2 I}$$

Real life nuisance - integration smearing

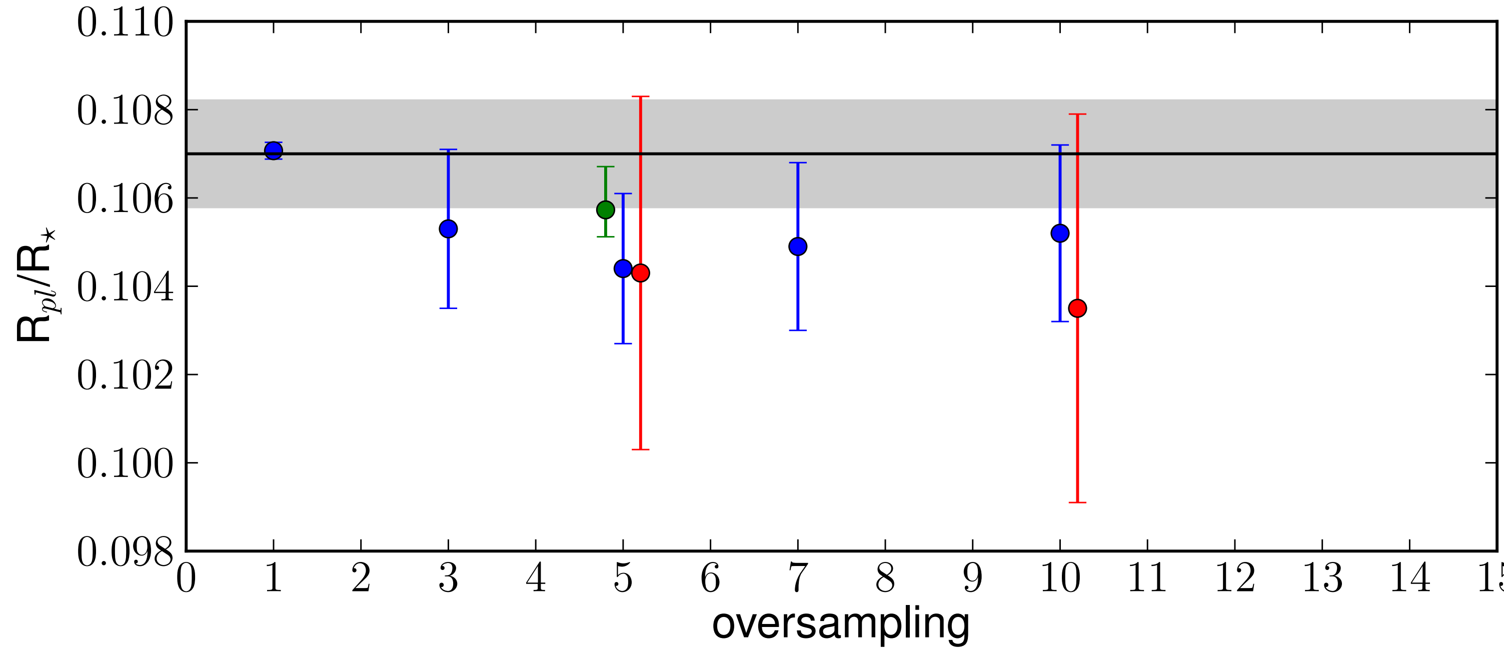


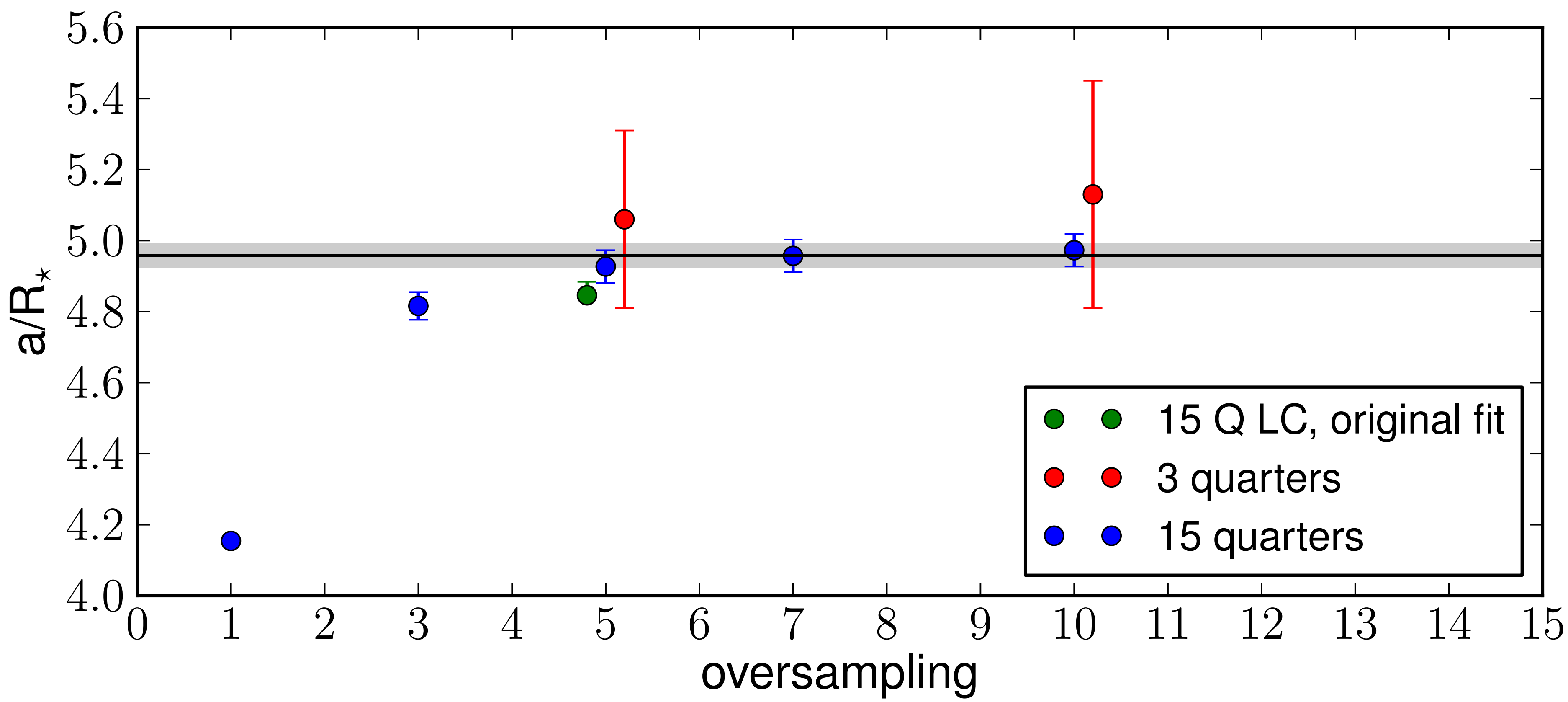
Kipping et al. (2010)

Real life nuisance - integration smearing



Kipping et al. (2010)





Credit: J. Almenara

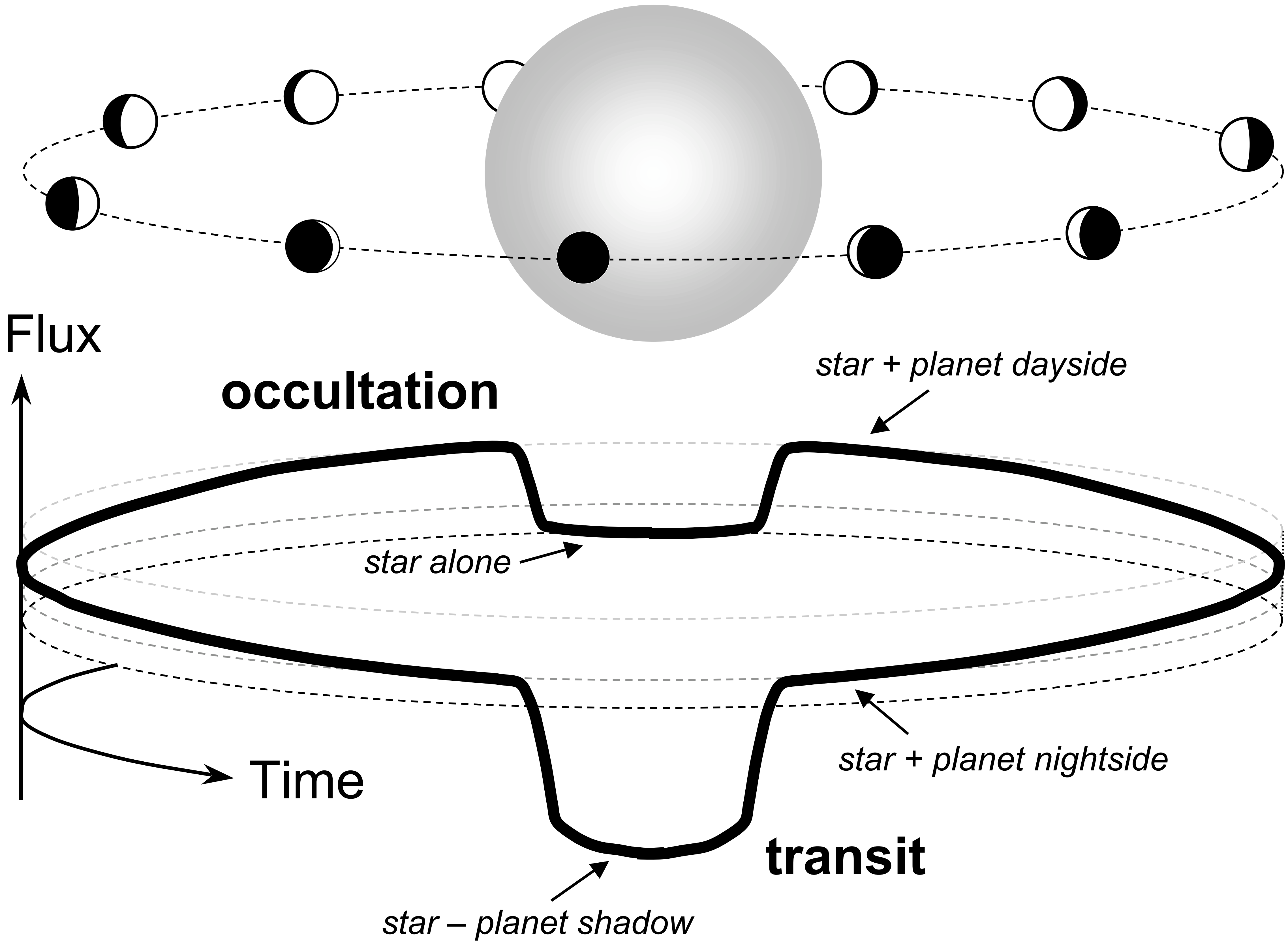
Physical Models

Beyond the basics

Improving model realism.

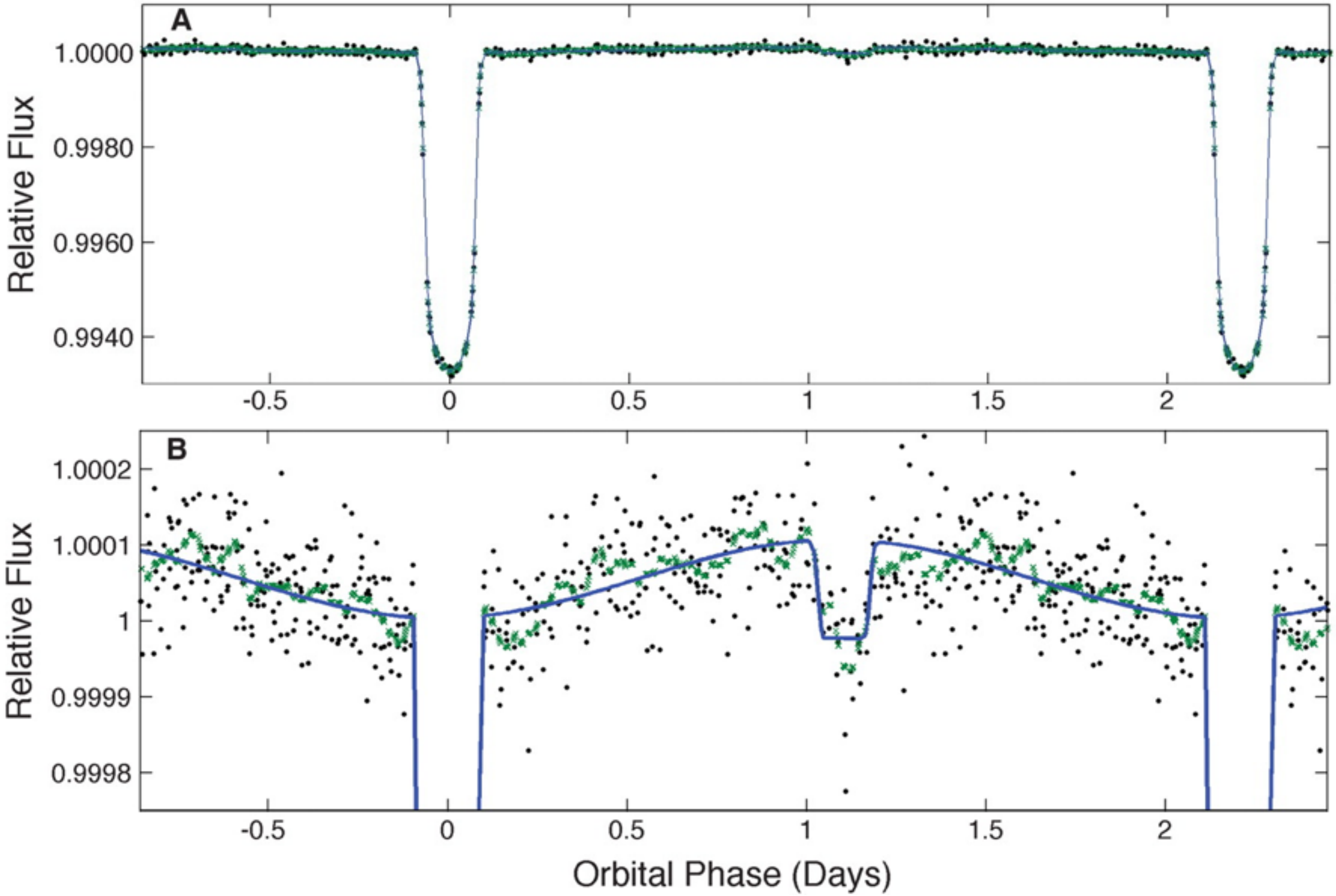
> "All models are wrong,
> but some are useful."
>
> George Box

- Phase curves, secondary eclipses.
- Rossiter-McLaughlin effect.
- Transit timing variations.



Credit: J. Winn

Optical phase curves

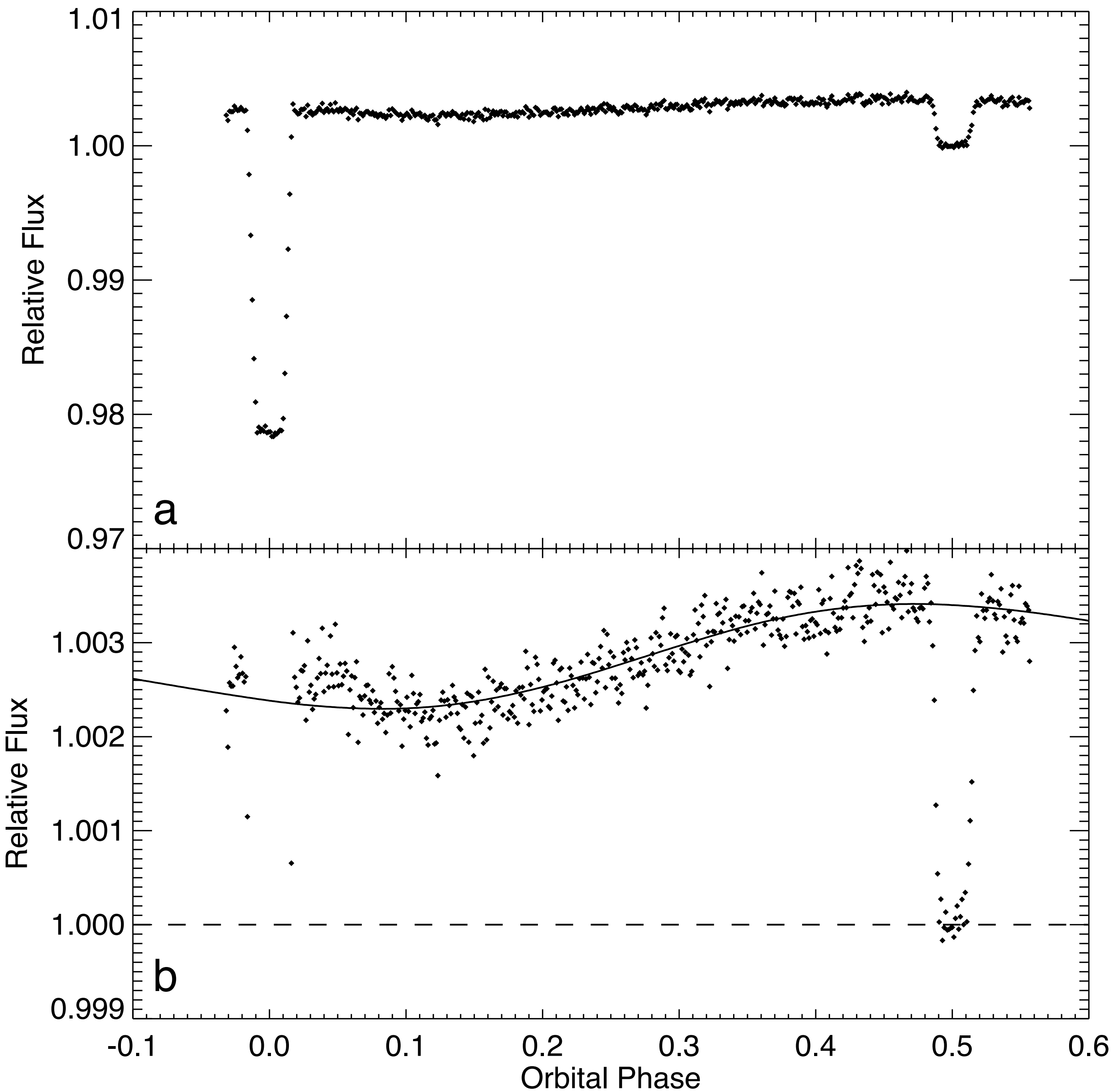


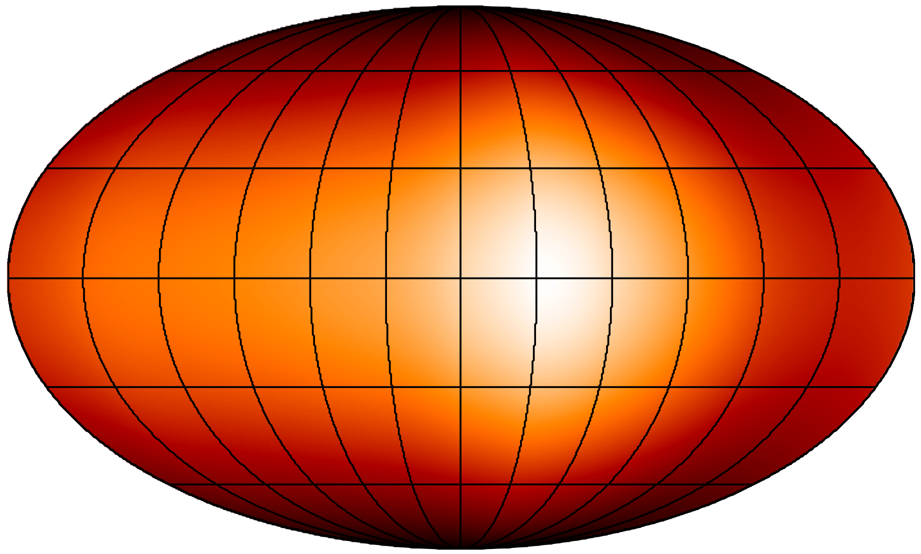
Borucki et al. (2009)

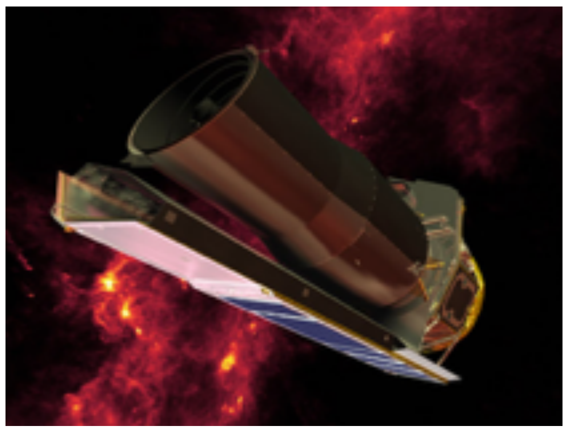
IR phase curves

HD189733 b, a Hot Jupiter

Knutson et al. (2007)

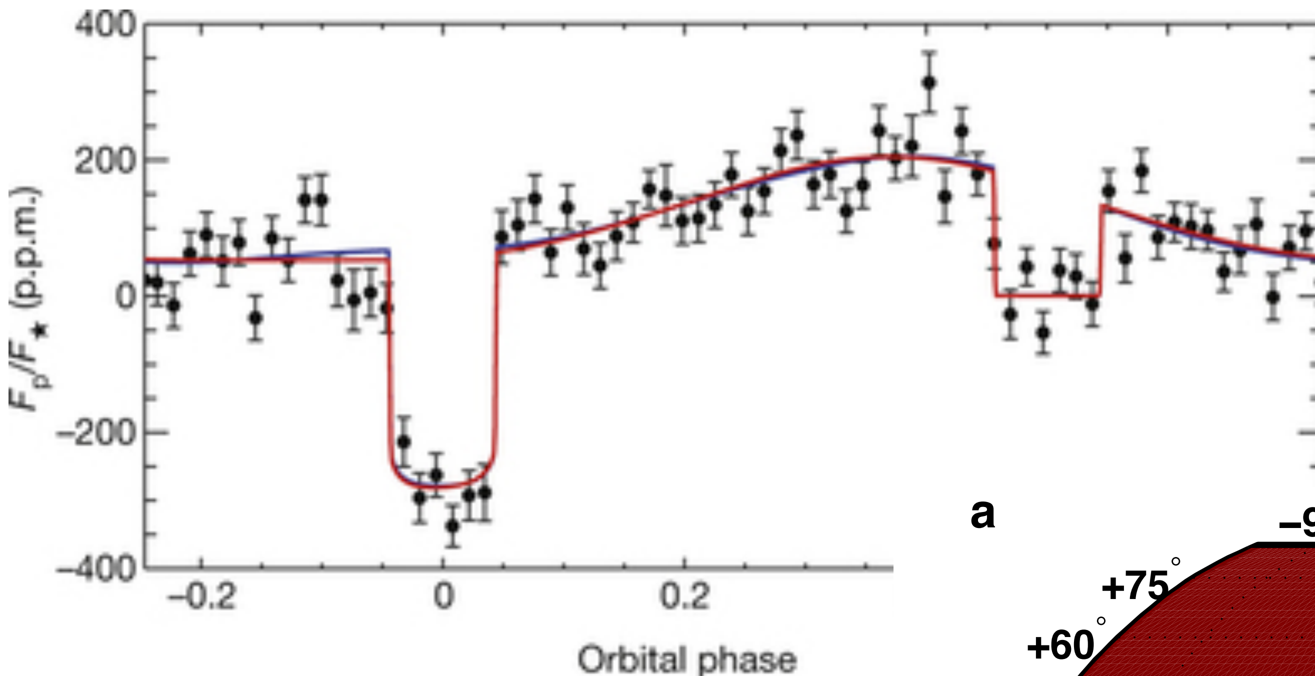


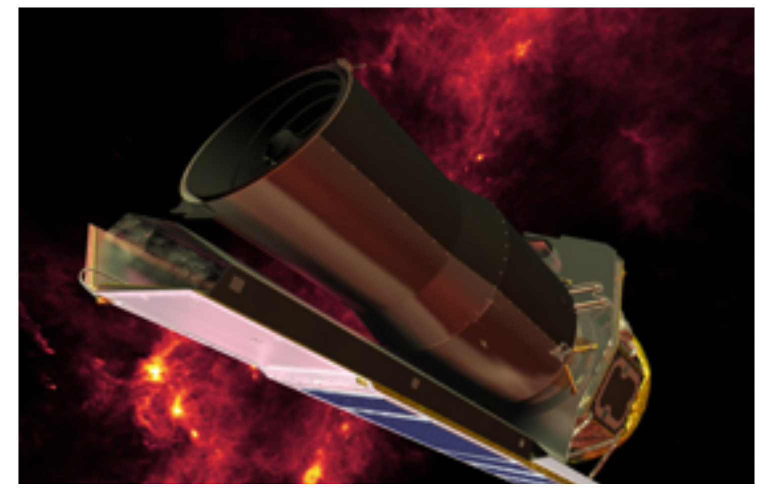


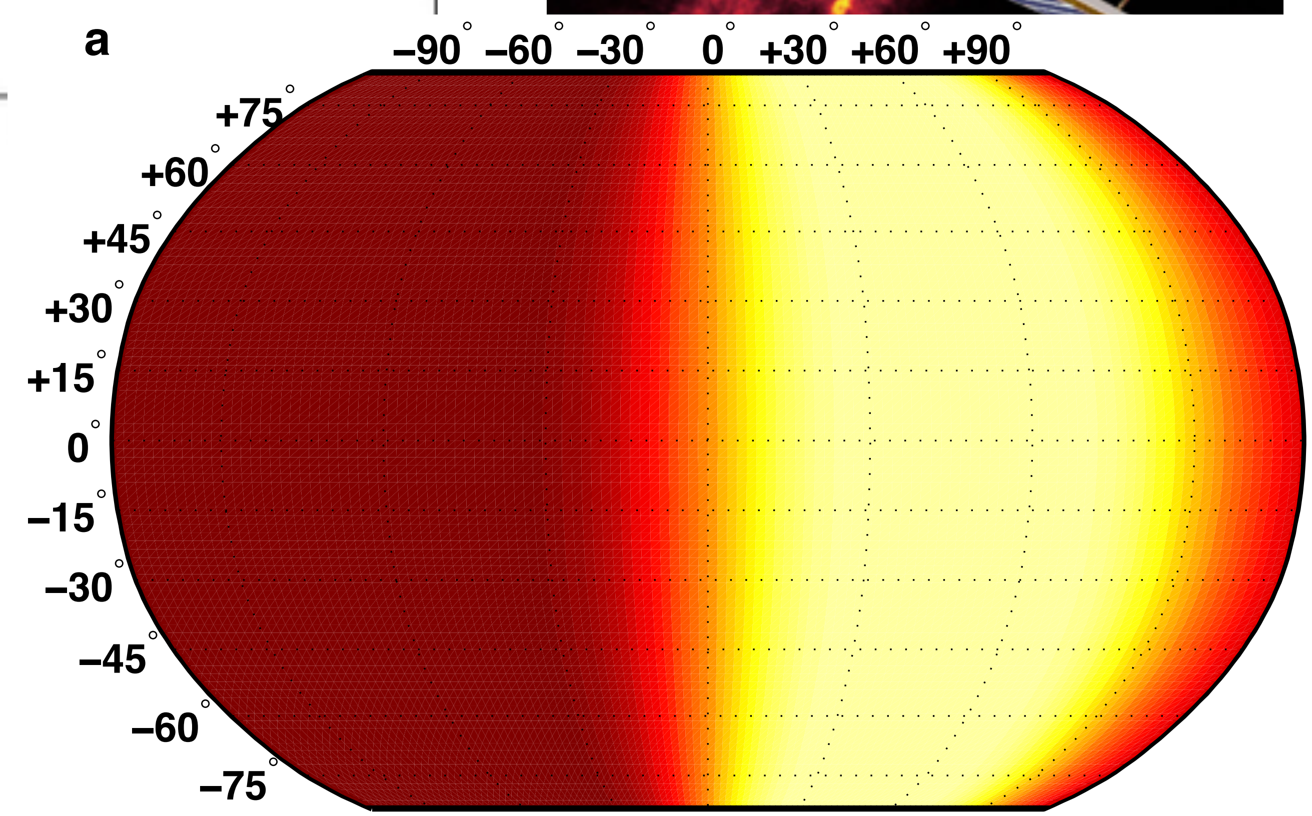


Optical and IR phase curves

55 Cnc e, a Super-Earth on a P=0.74-day orbit

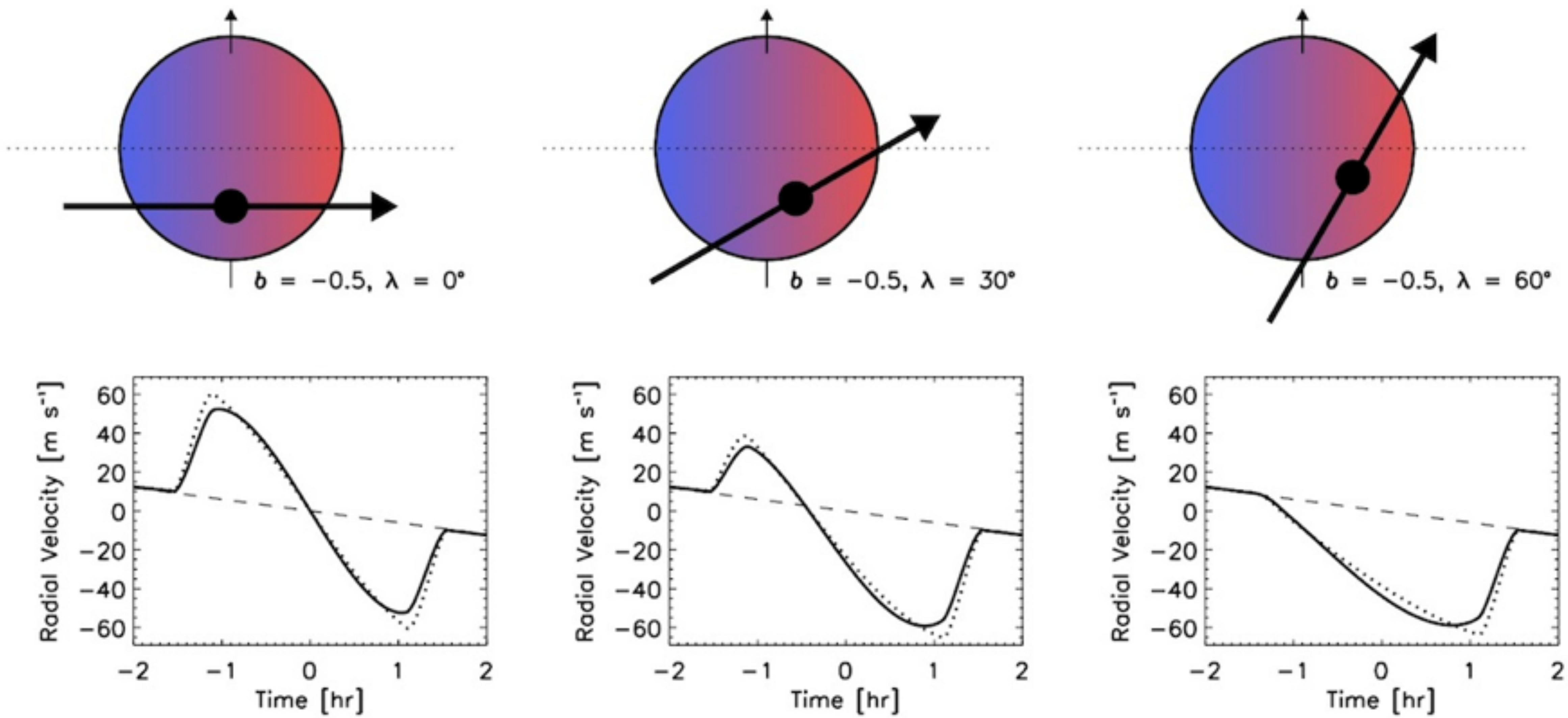






Demory et al. (2016)

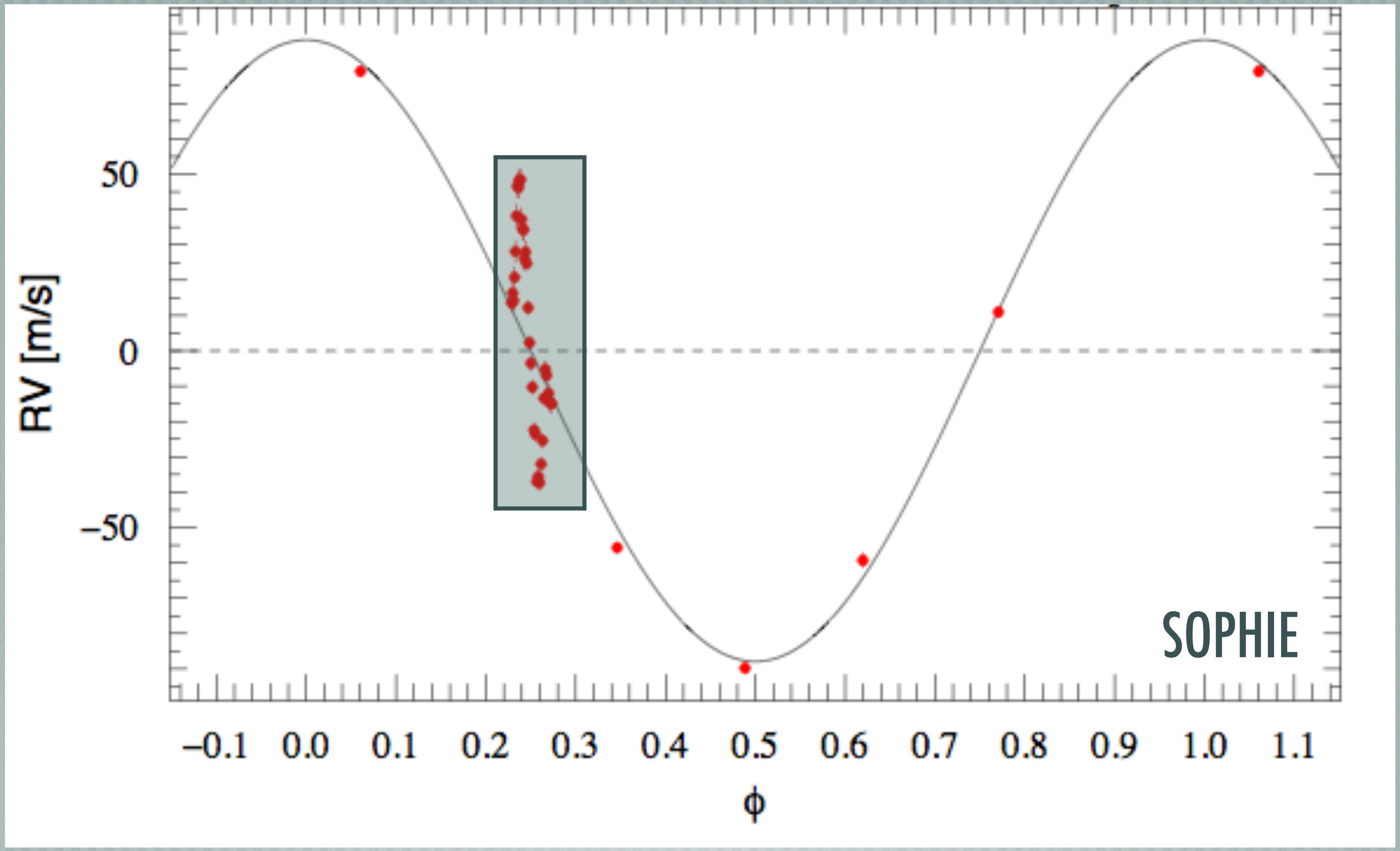
Rossiter-McLaughlin effect



Credit: A. Triaud

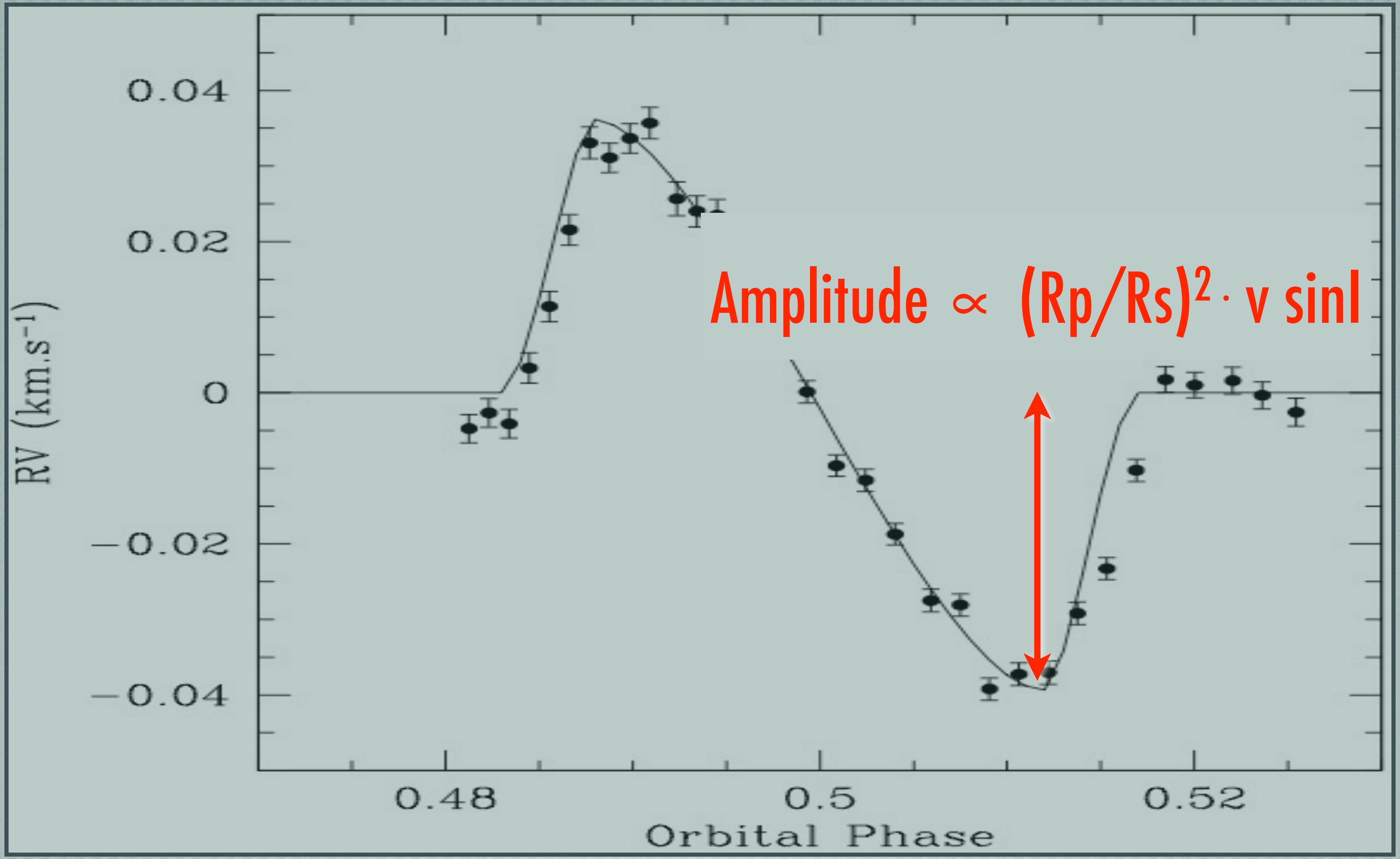
Rossiter-McLaughlin effect

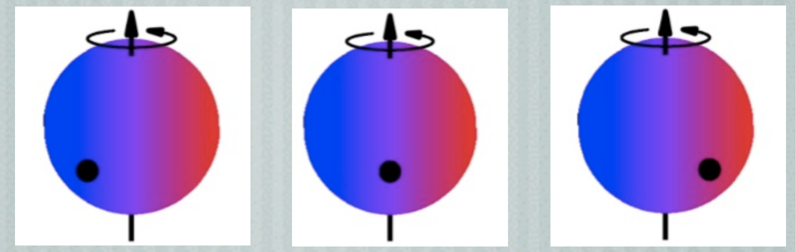
HD209458 b



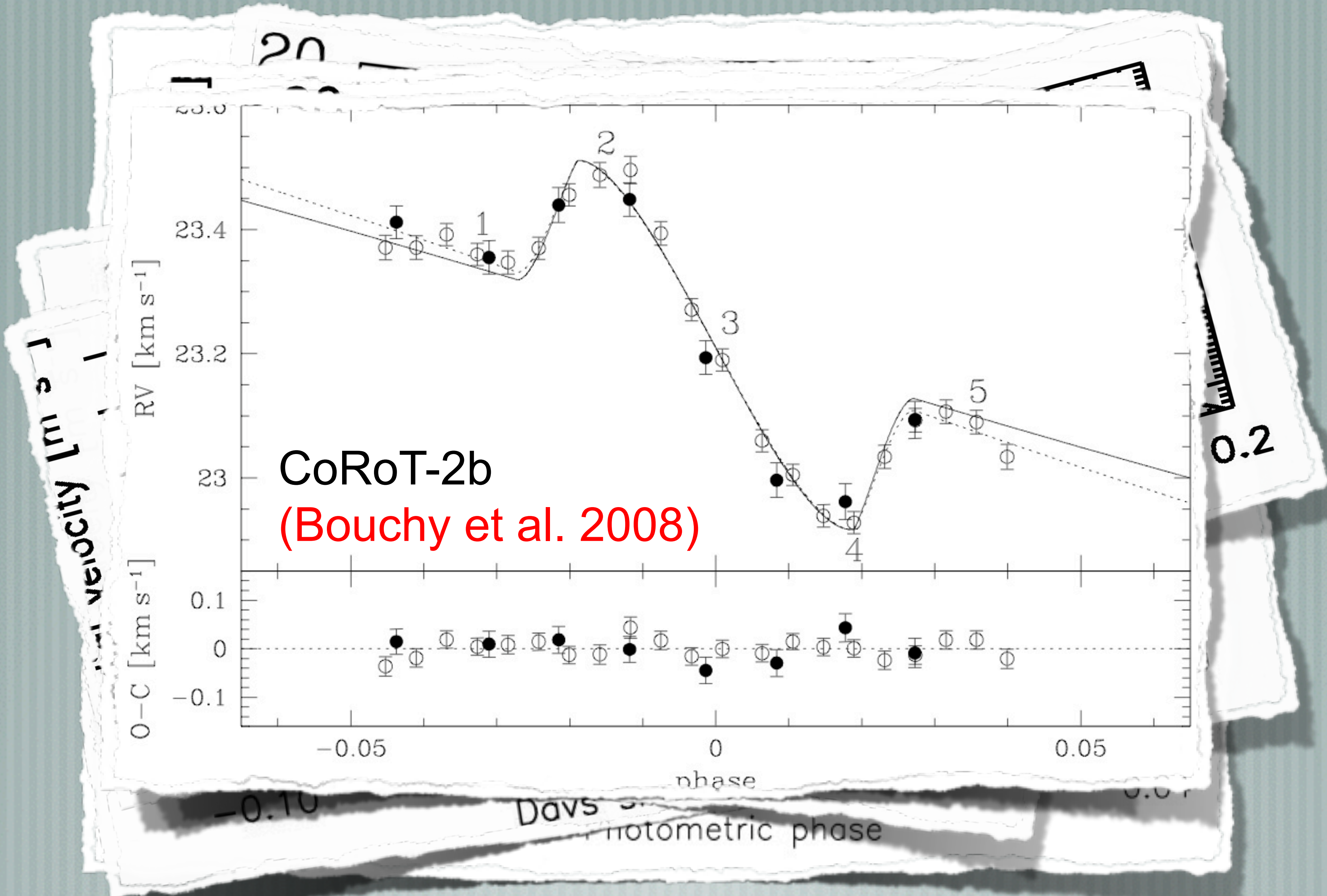
Rossiter-McLaughlin effect

HD209458 b

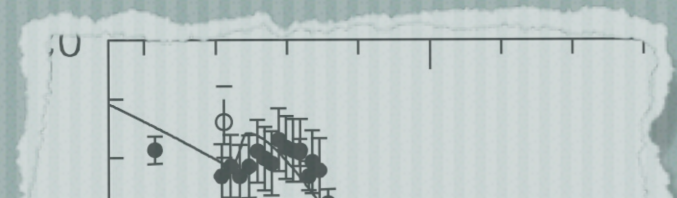


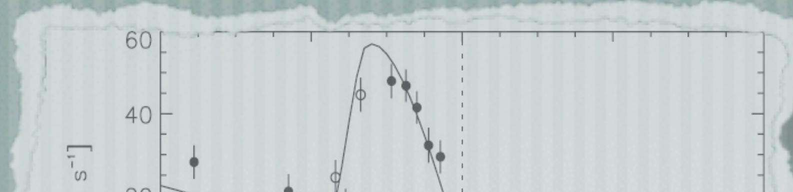


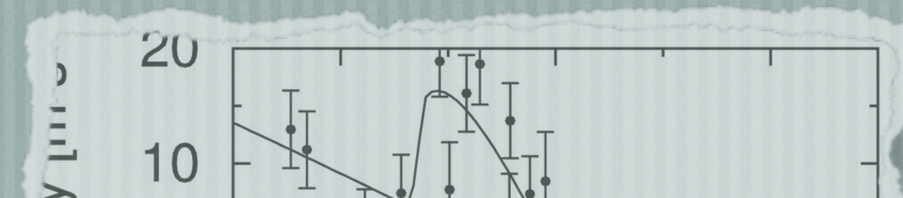
First measurements: alignment galore



First measurements: alignment galore

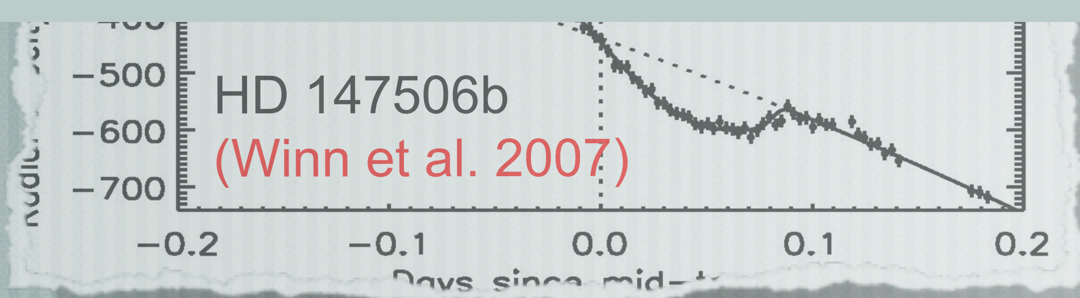




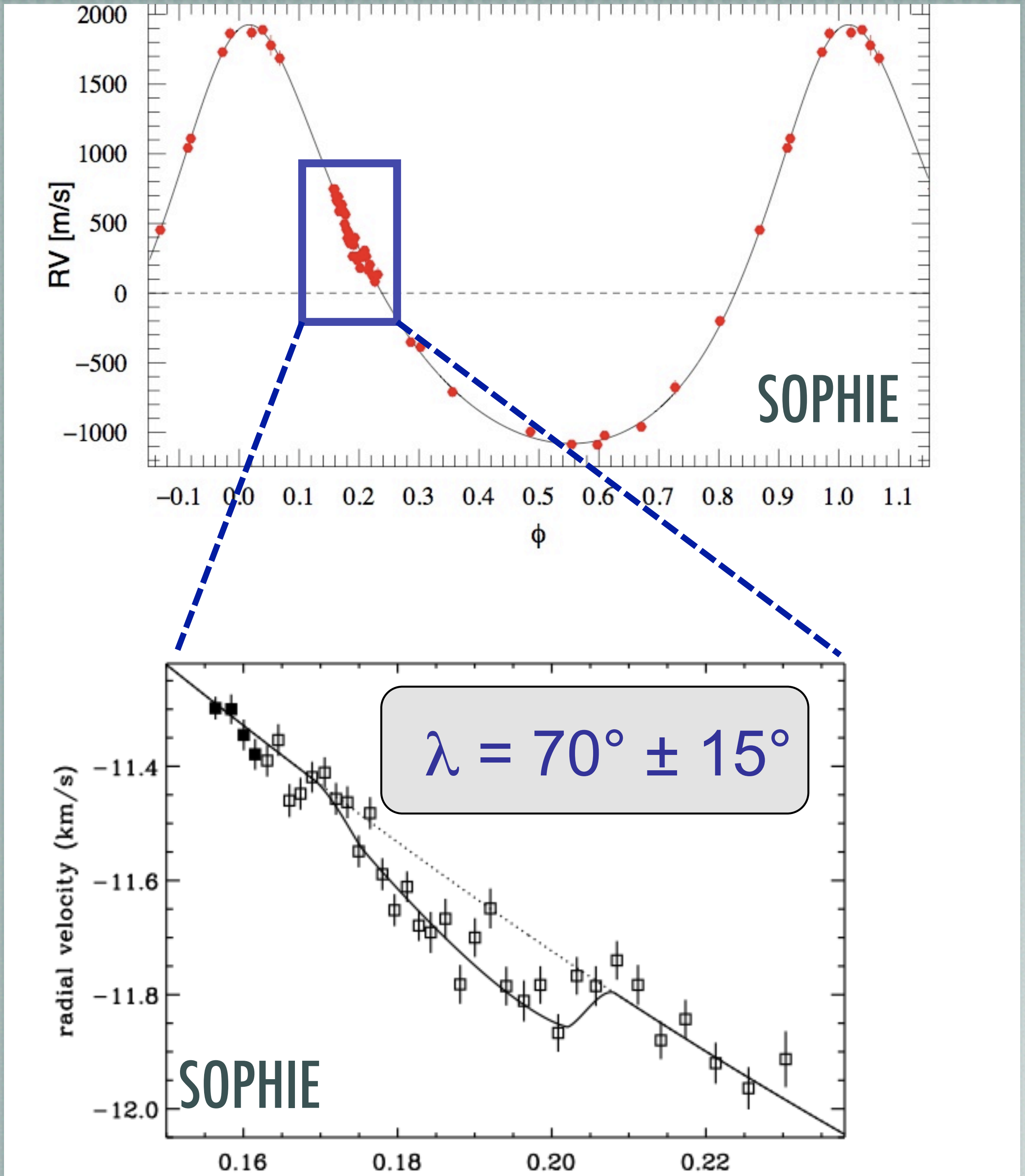


By 2008, the obliquity of about ⅕ of the known transiting planets had been measured.

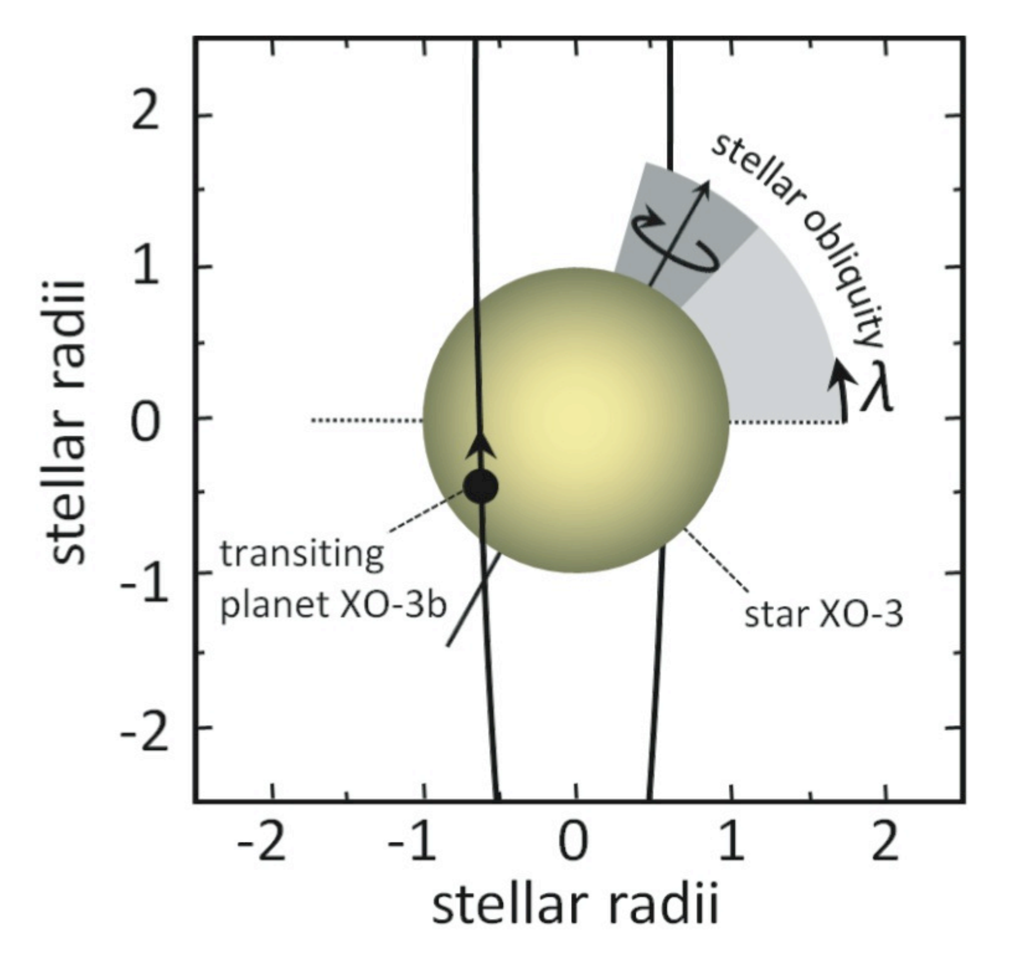
All orbits were aligned and prograde, in agreement with the expected result of planetary formation and migration in a protoplanetary disk



The first case of a misaligned system, XO-3 b

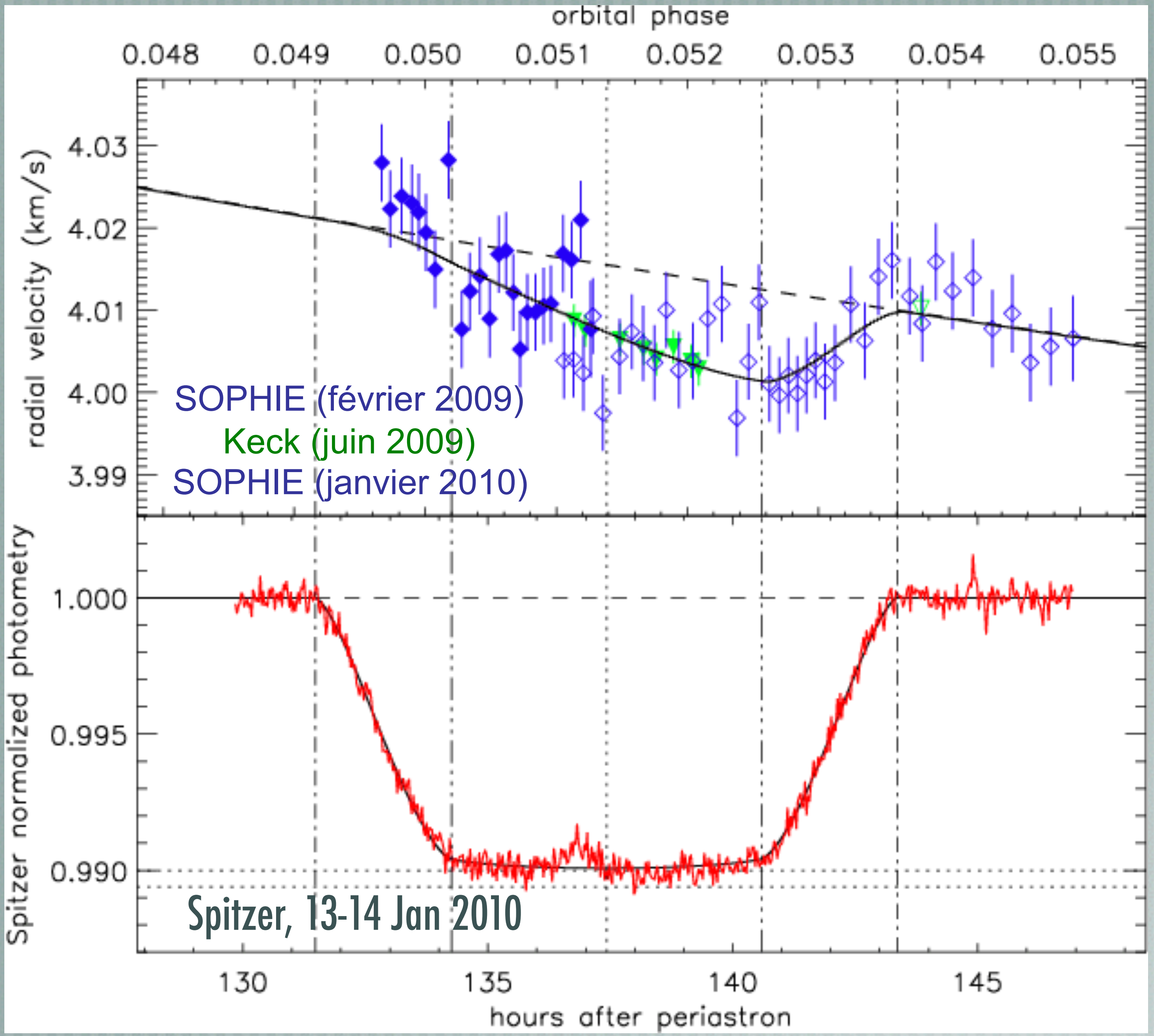


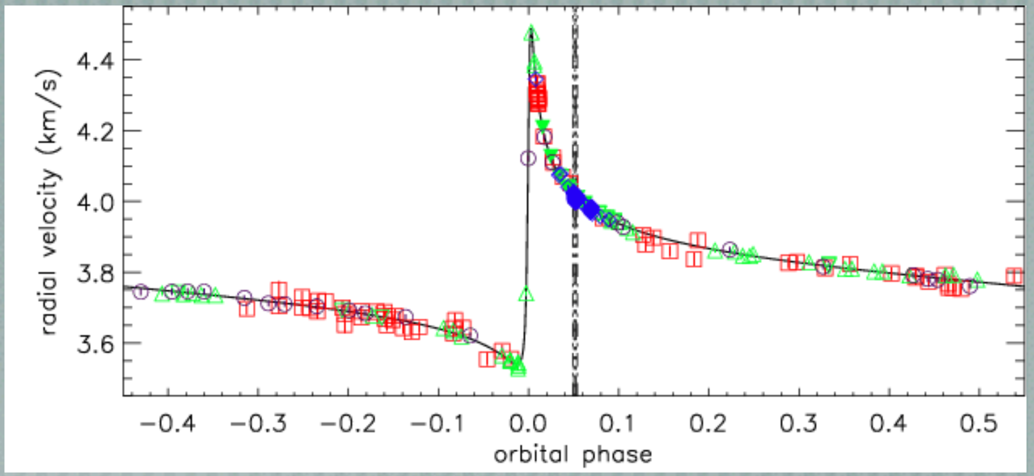
Hébrard, Bouchy, Pont et al. (2008)



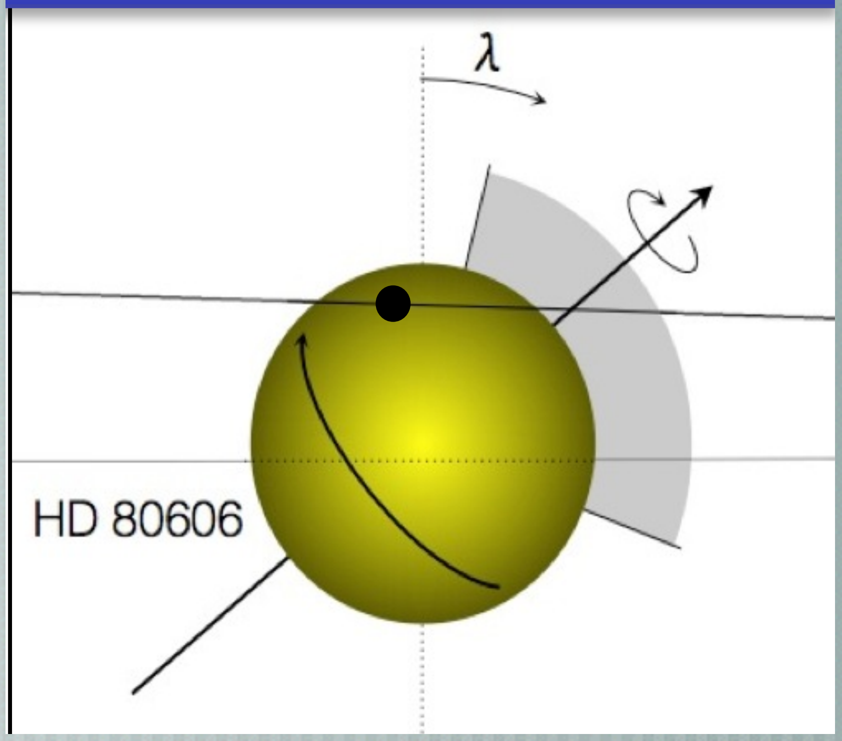
Confirmed in 2009 by Winn et al.

A long-period transiting planet in a very eccentric misaligned orbit





$\lambda = 42° \text{ +/- } 8°$

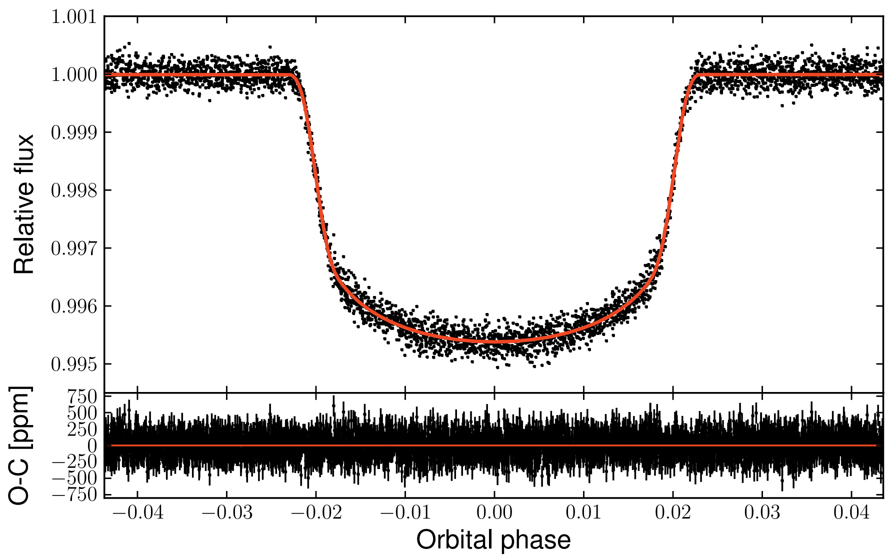


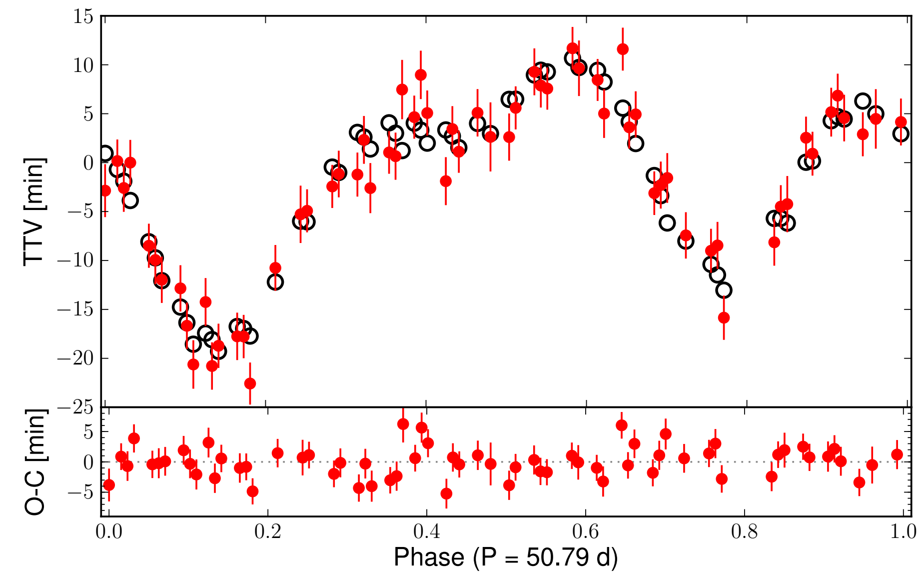
Hébrard, Désert, Díaz, et al. (2010)

Physical Models

Transit timing variations

Transits + TTV —> density

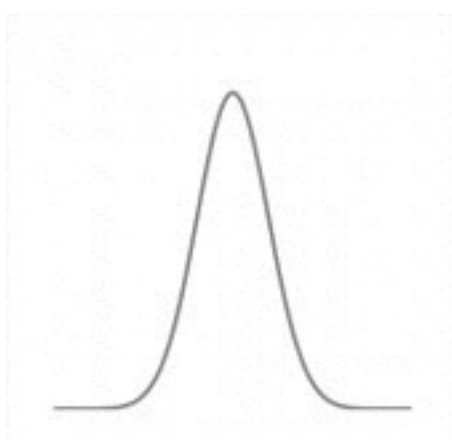




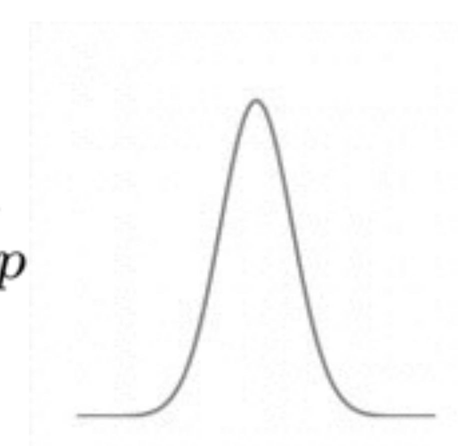
$$\frac{R_p}{R_\star}$$

$$\frac{M_p}{M_*}$$

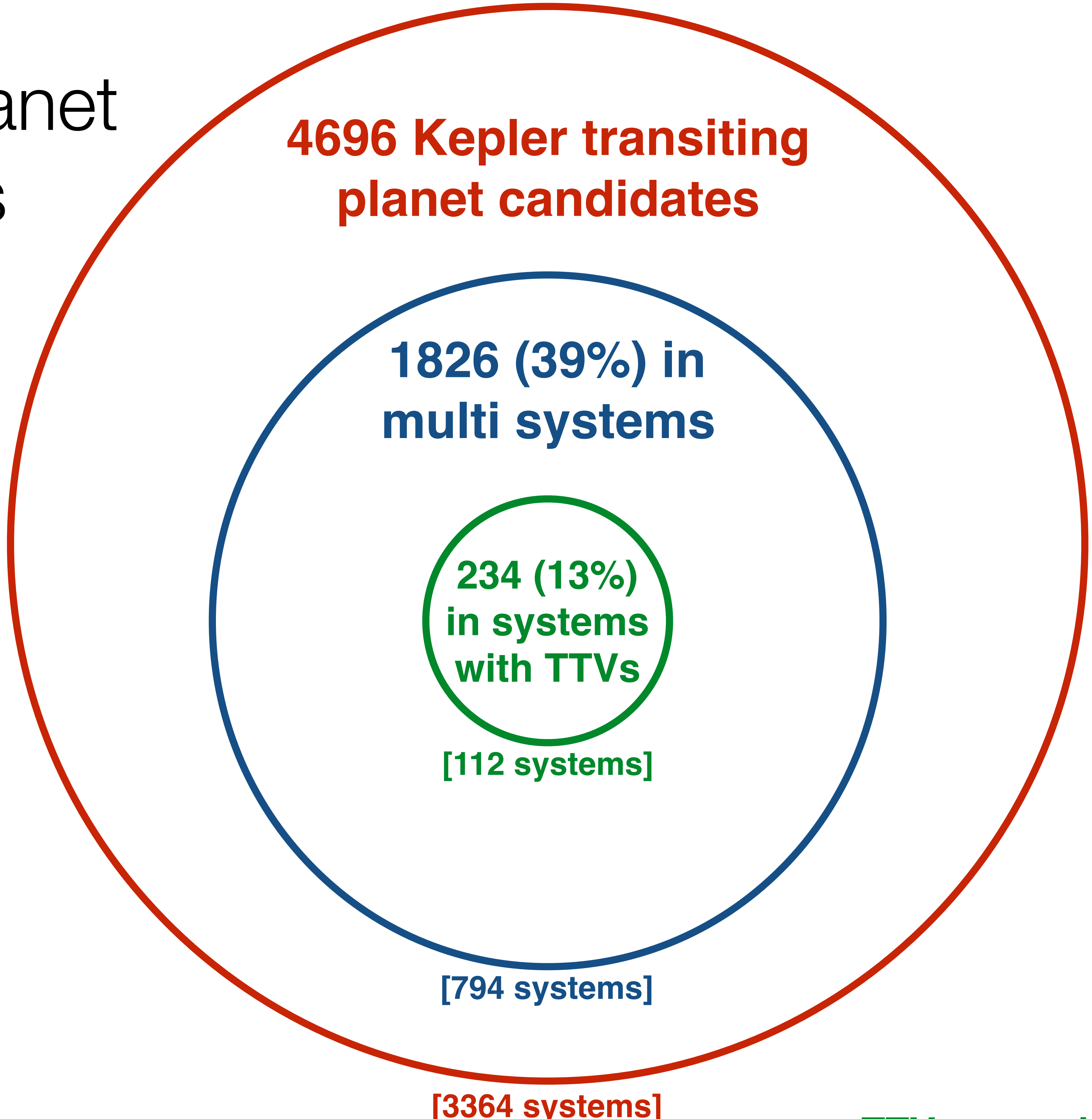
with stellar models



$$R_p \qquad \rho_p \qquad M_p$$

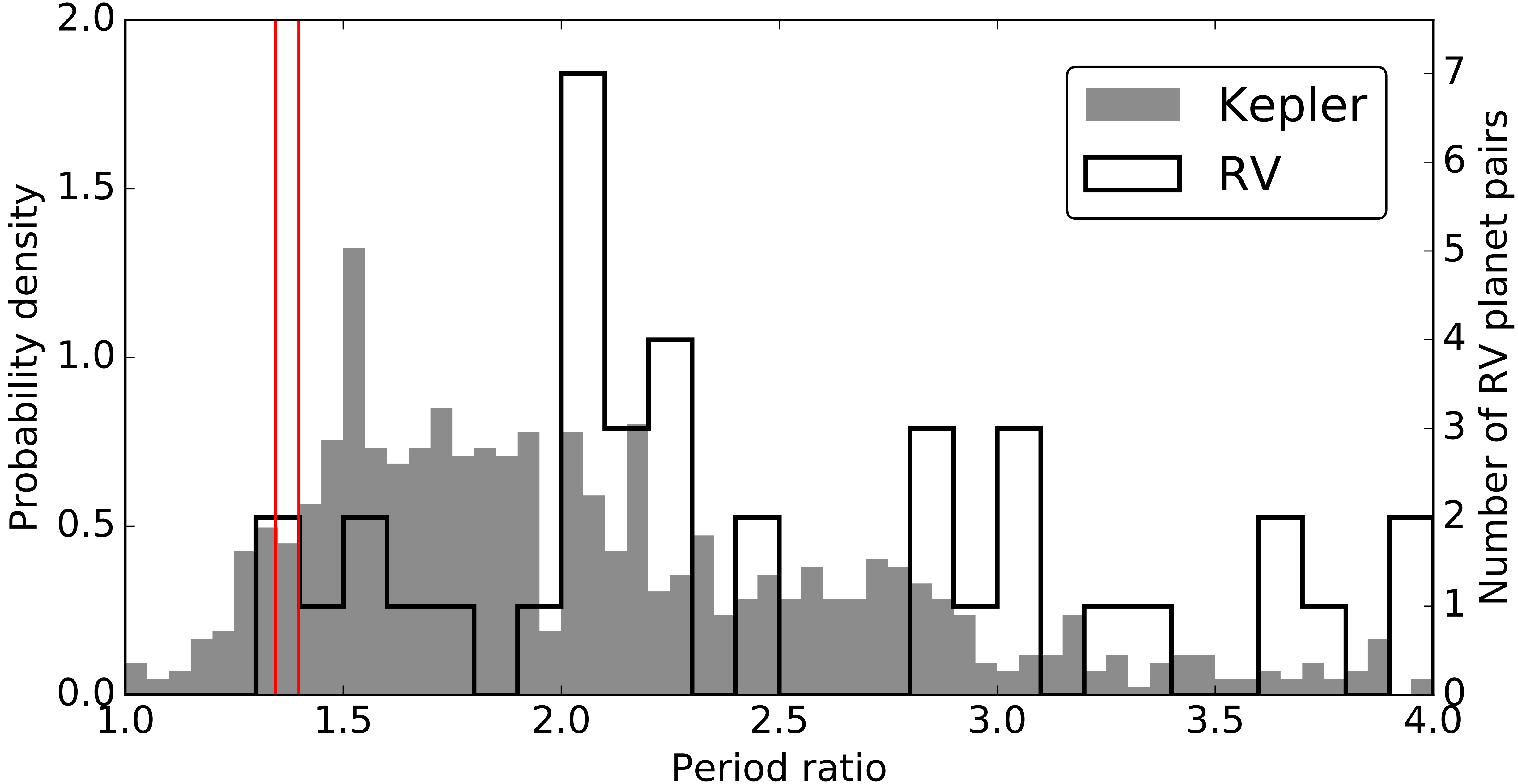


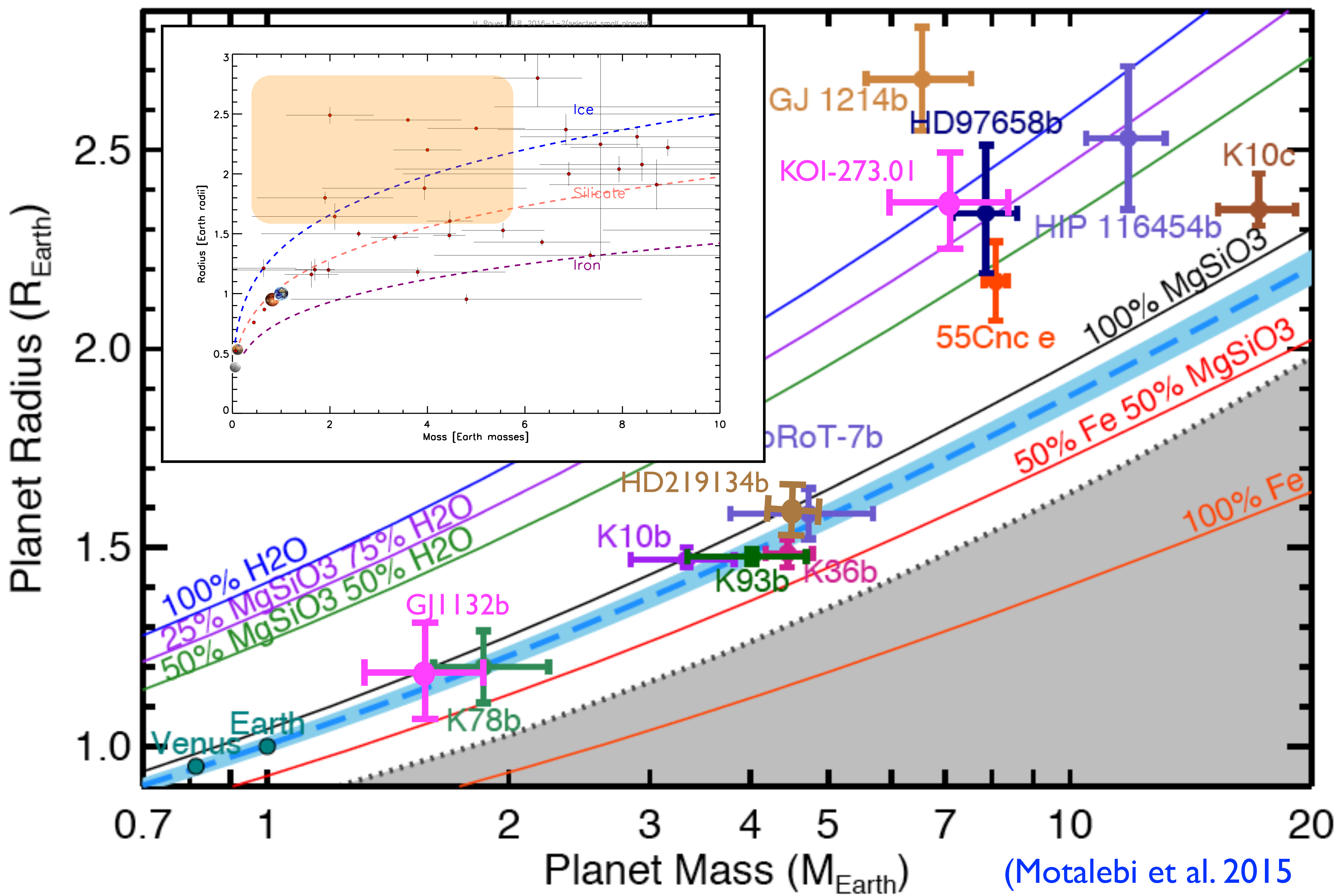
Kepler multi-planet systems



TTVs: transit timing variations

Missing compact planetary systems.





(precision on the mass < 20%)

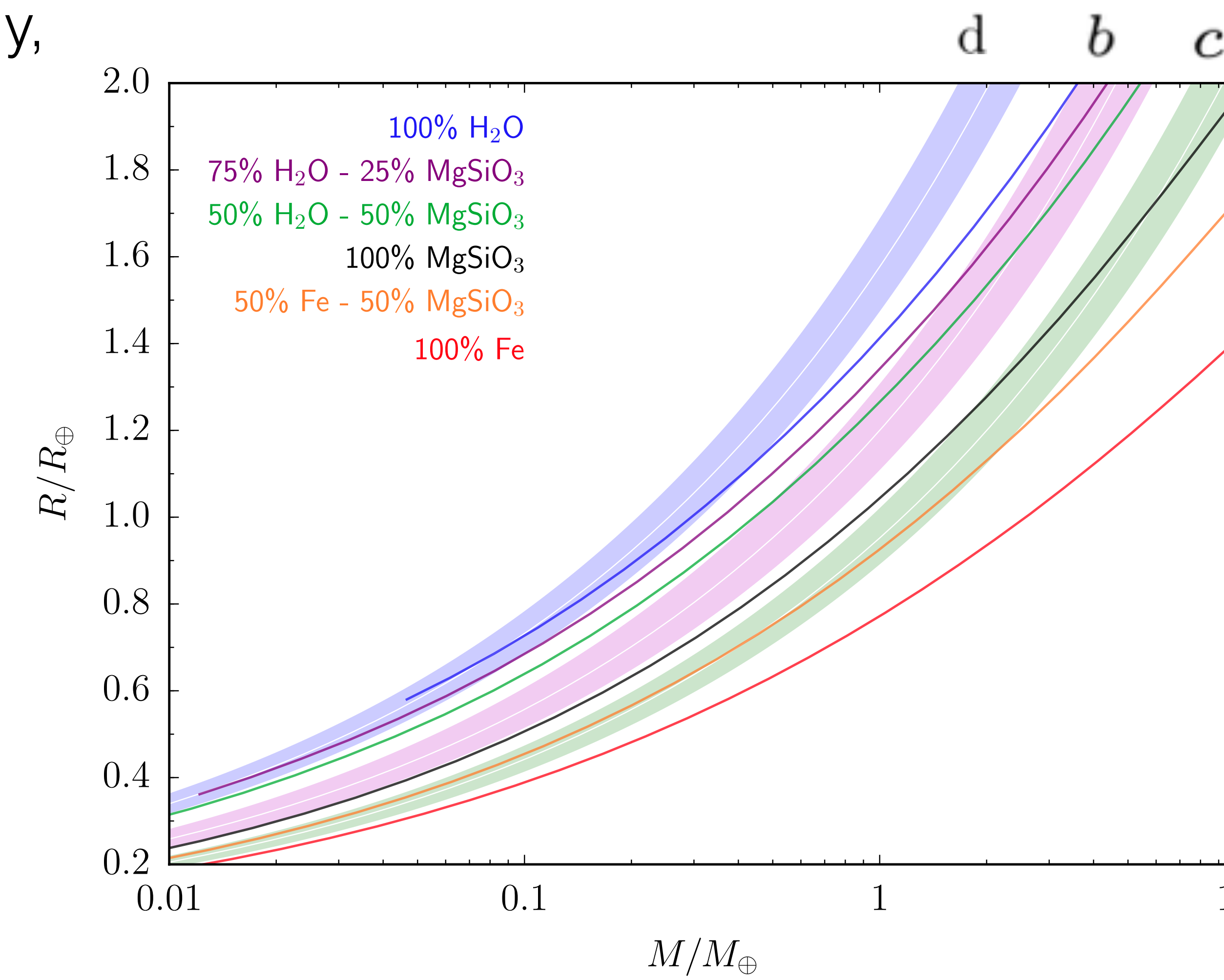
(Motalebi et al. 2015 updated from Dressing et al. 2015)

Planet density from light curve data

The planetary density can be obtained directly, without going through M_* or R_*

$$\frac{\rho_p}{\rho_\star} = \frac{\frac{M_p}{\frac{4}{3}\pi R_p^3}}{\frac{M_\star}{\frac{4}{3}\pi R_\star^3}}$$

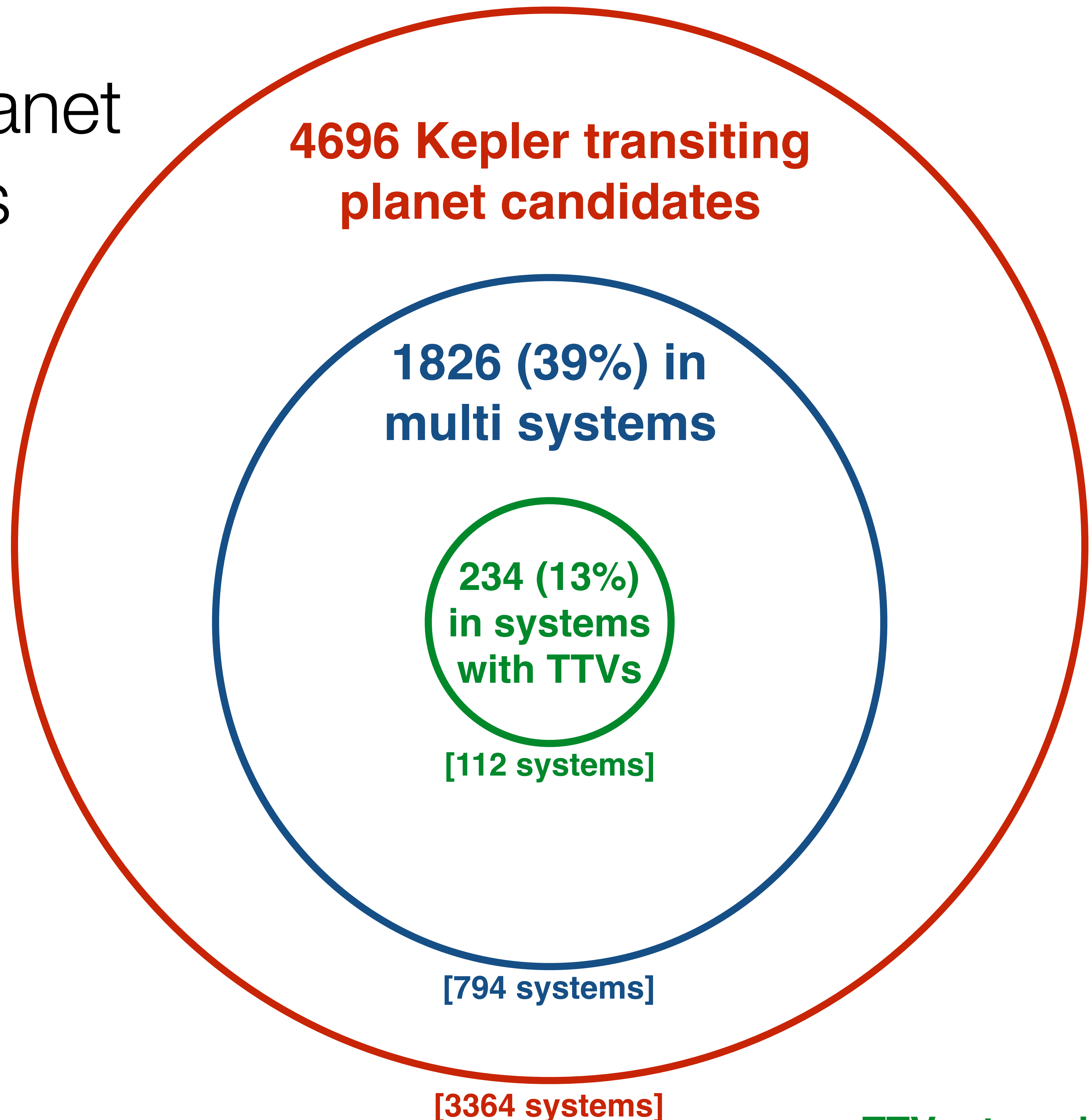
$$\rho_p = \rho_\star \left(\frac{M_p}{M_\star}\right)\left(\frac{R_p}{R_\star}\right)^{-3}$$



The masses and radii **cannot** be obtained independently.
Gravity is scale invariant, i.e. under the transformation:

$$M' = \alpha^3 M \qquad d' = \alpha d \qquad [\mathrm{G}] = \left(\frac{\mathrm{kg}}{\mathrm{m}^3}\right)^{-1} s^{-2}$$

Kepler multi-planet systems

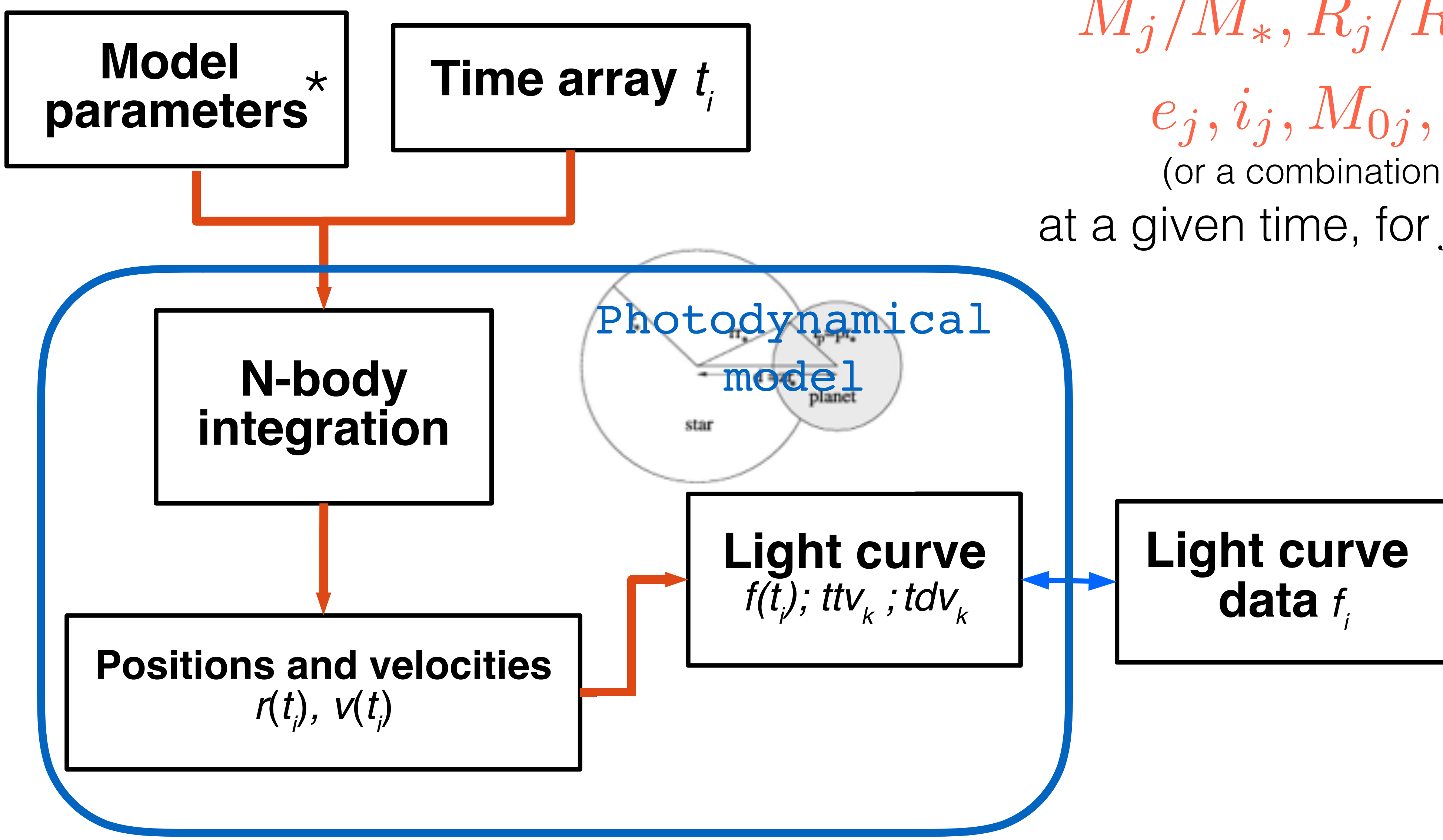


TTVs: transit timing variations

Photo-dynamical modelling

ρ_*

$*$ $M_j/M_*, R_j/R_*, a_j/R_*,$
$e_j, i_j, M_{0j}, \varpi_j, \Omega_j$
(or a combination thereof),
at a given time, for $j = 1, \ldots, N_{\mathrm{pla}}$



Kepler-117, two planets far from resonance

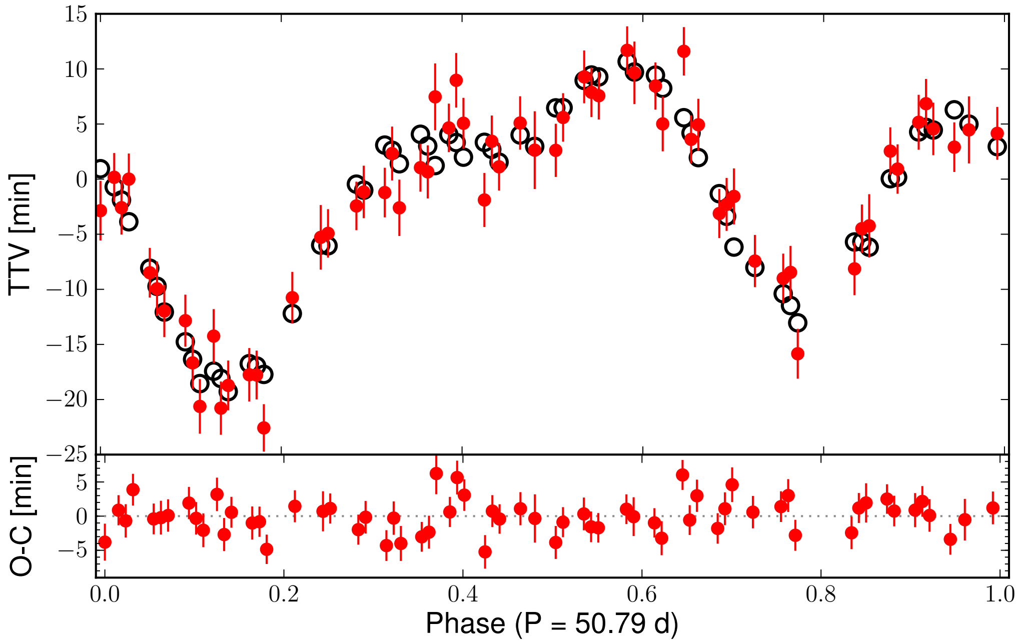
Almenara, Díaz, Mardling, et al. (MNRAS, 2015)

$K_p = 13$

$P_b = 18.8\,\mathrm{d}$
$P_c = 50.8\,\mathrm{d}$

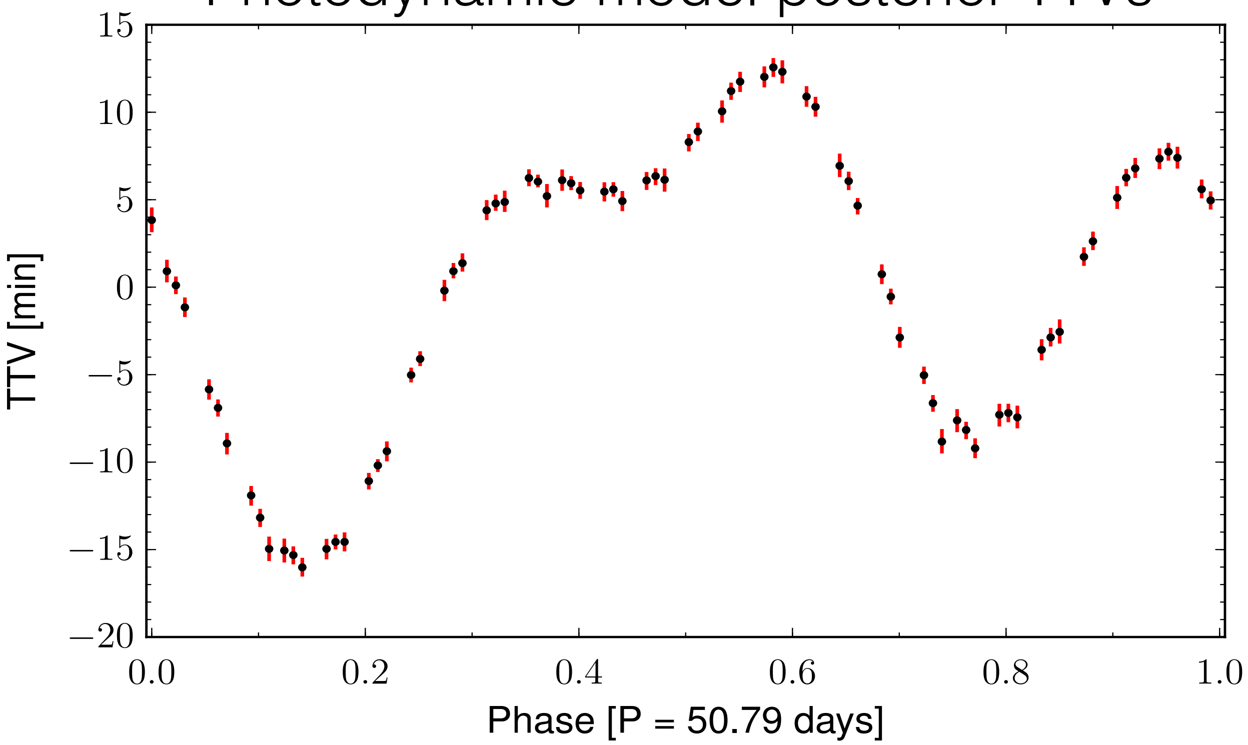
$\frac{P_c}{P_b} = 2.7$

Individual TTV meas.



Bruno, et al. (A&A, 2014)

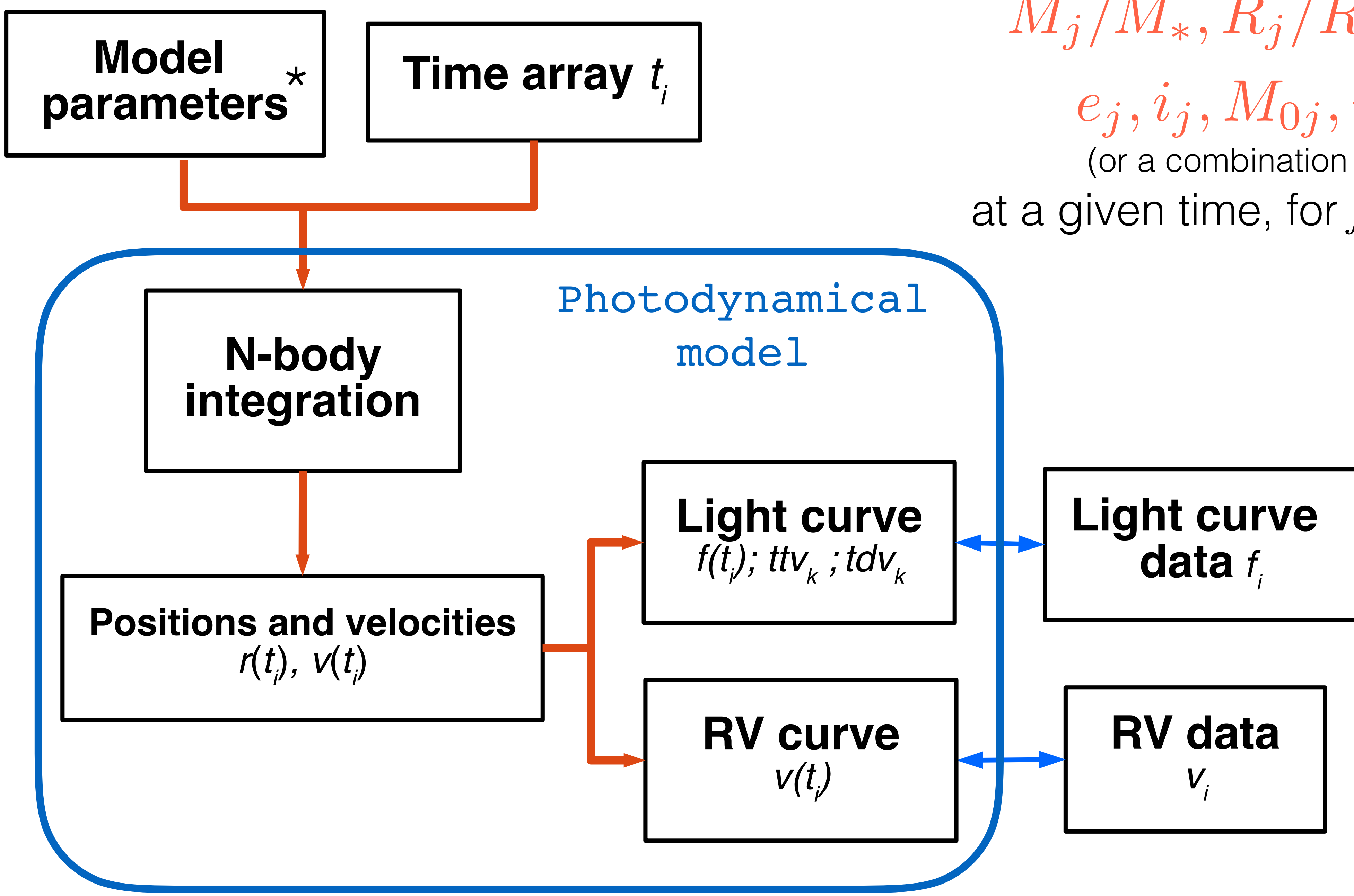
Photodynamic model posterior TTVs



Stellar and planetary densities obtained with 2% - 5% precision, **without RV**, and **independently of stellar models**.

$(M_b+M_c)/M_* \sim 0.4\%$

Photo-dynamical modelling



* ρ_*

$M_j/M_*, R_j/R_*, a_j/R_*,$

$e_j, i_j, M_{0j}, \varpi_j, \Omega_j$

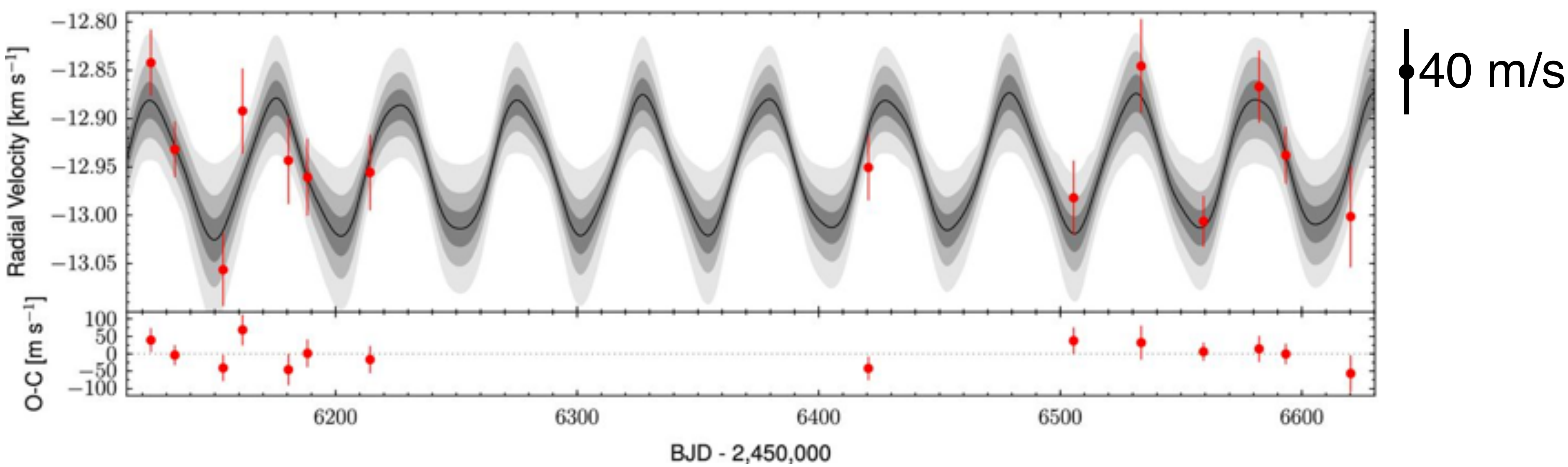
(or a combination thereof),

at a given time, for $j = 1, \ldots, N_{\rm pla}$

Kepler-117, two planets far from resonance

Almenara, Díaz, Mardling, et al. (MNRAS, 2015)

As a bonus, SOPHIE **RVs** break the Newtonian degeneracy.



Masses and radii are imprecise due to poor **RV** errors, but **independent of stellar models**.

Simulations of **RV** data with 1-m/s precision lead to masses and radii at a few percent precision.

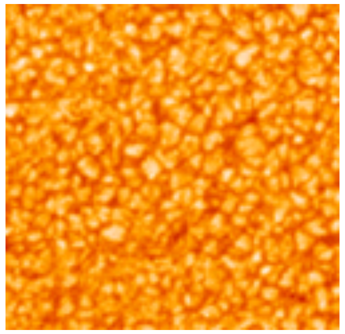
Limitations

- Computing time. → Analytical TTVs analysis
- Complex parameter space (usually multi-modal). → Bayesian model comparison
- Additional potential planets in the system. → Bayesian model comparison
- Incomplete model:
  - Light-time effects. → Usually negligible
  - Post-Newtonian / GR. → Usually negligible
  - Stellar activity. → Gaussian process regression
  - Tides.

Stellar activity.

All models are incomplete.

<u>Stellar jitter</u>



<u>Stellar activity</u>

- spots
- cycles

<u>Instrumental effects</u>

- outliers
- false positives

$$d_i = m_i + e_i$$

Physical models

Radial velocity

Transit light curve.

$$\underset{\text{data}}{d_i} = \underset{\text{model}}{m_i} + \underset{\text{error}}{e_i}$$

Data analysis

Likelihood functions

MCMC

$p(D|\theta, H, I)$

Error model

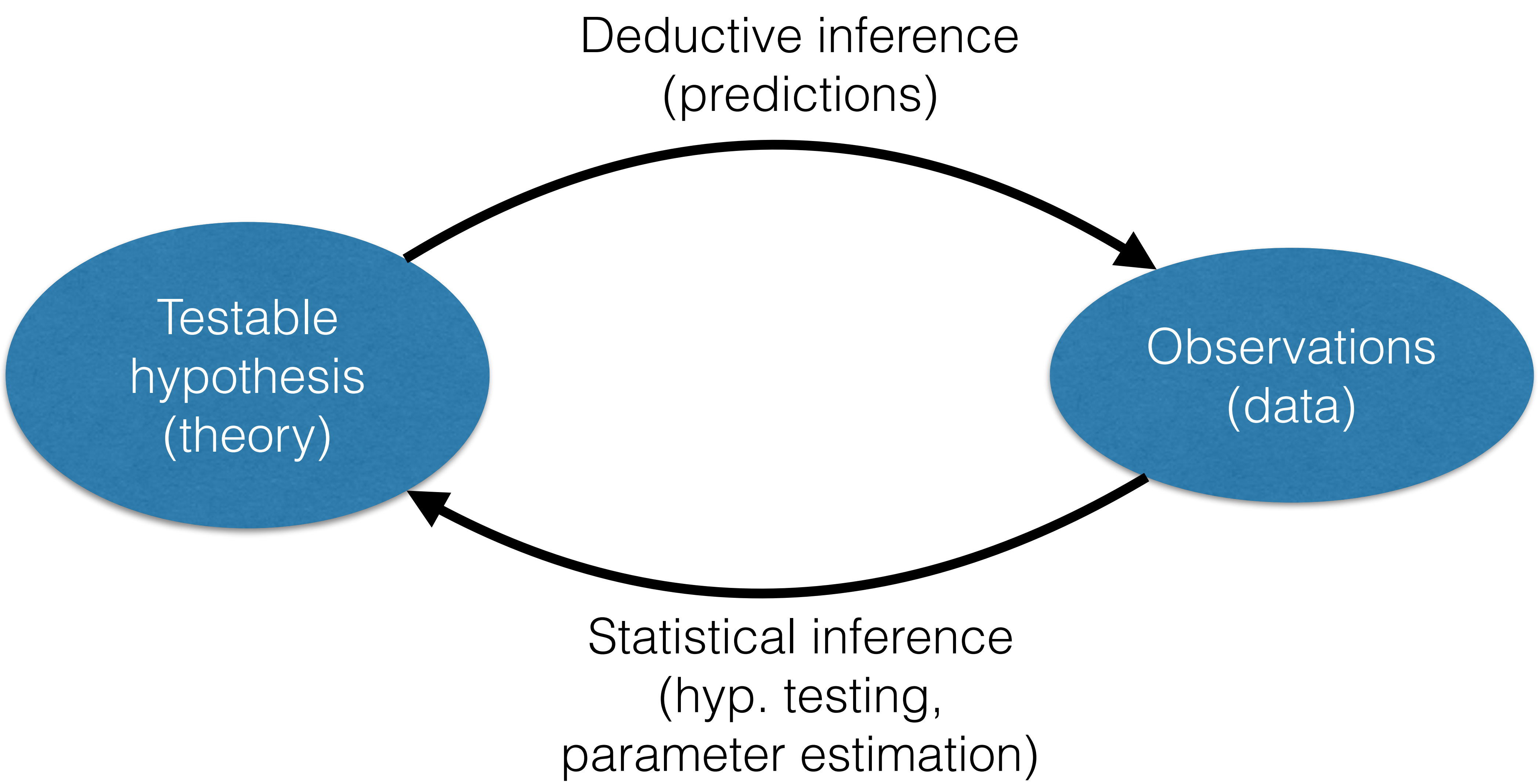
Beyond the independent normal paradigm

Data analysis

The Bayesian revolution

Statistical inference

Fig. adapted from Gregory (2005)



Statistical
inference
requires a
probability theory

Frequentist

Bayesian

Thomas Bayes (1701 – 1761)

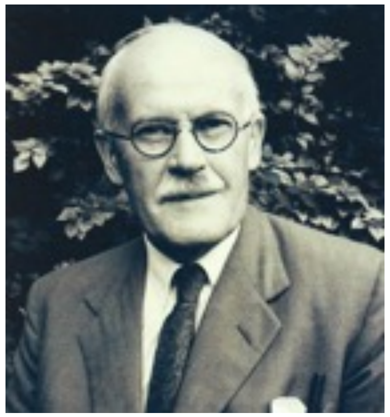
First appearance of the **product rule** (the base for the Bayes' theorem; *An Essay towards solving a Problem in the Doctrine of Chances*).

$$p(H_i|I,D) = \frac{p(D|H_i,I)}{p(D|I)} \cdot p(H_i|I)$$

Pierre-Simon Laplace (1749 – 1827)

Wide application of the **Bayes' rule.** Principle of insufficient reason (non-informative priors). Primitive version of the Bernstein–von Mises theorem.

Laplace's "inverse probability" is largely rejected for ~100 years. The reign of frequentist probability. Fischer, Pearson, etc.



Harold Jeffreys (1891 – 1989)

Objective Bayesian probability revived.
Jeffreys rule for priors.

(1940s - 1960s)
R. T. Cox
George Pólya
E. T. Jaynes

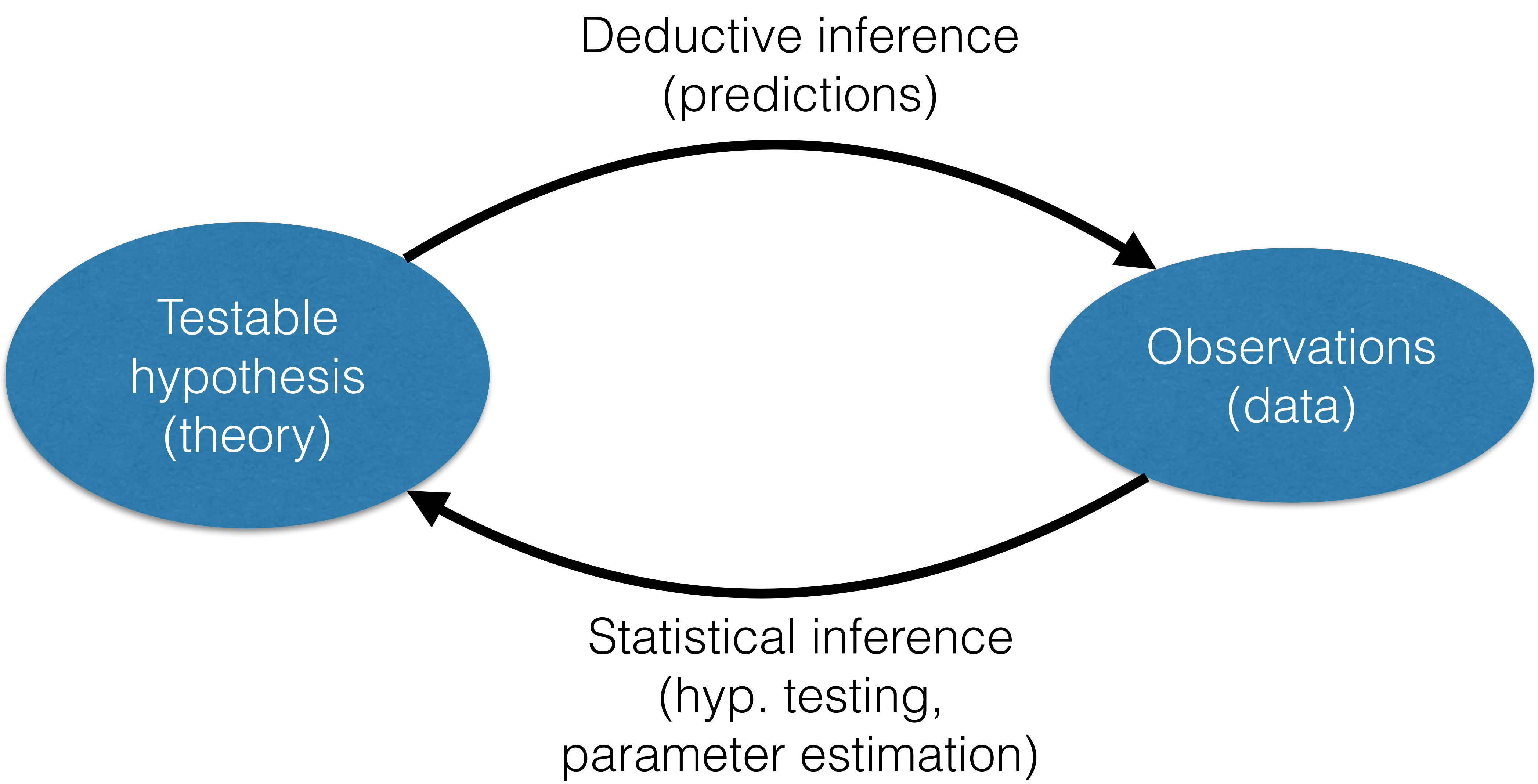
Plausible reasoning. Reasoning with uncertainty.
Probability theory as an extension of Aristotelian logic.
The product and sum rules deduced for basic principles.
MAXENT priors.

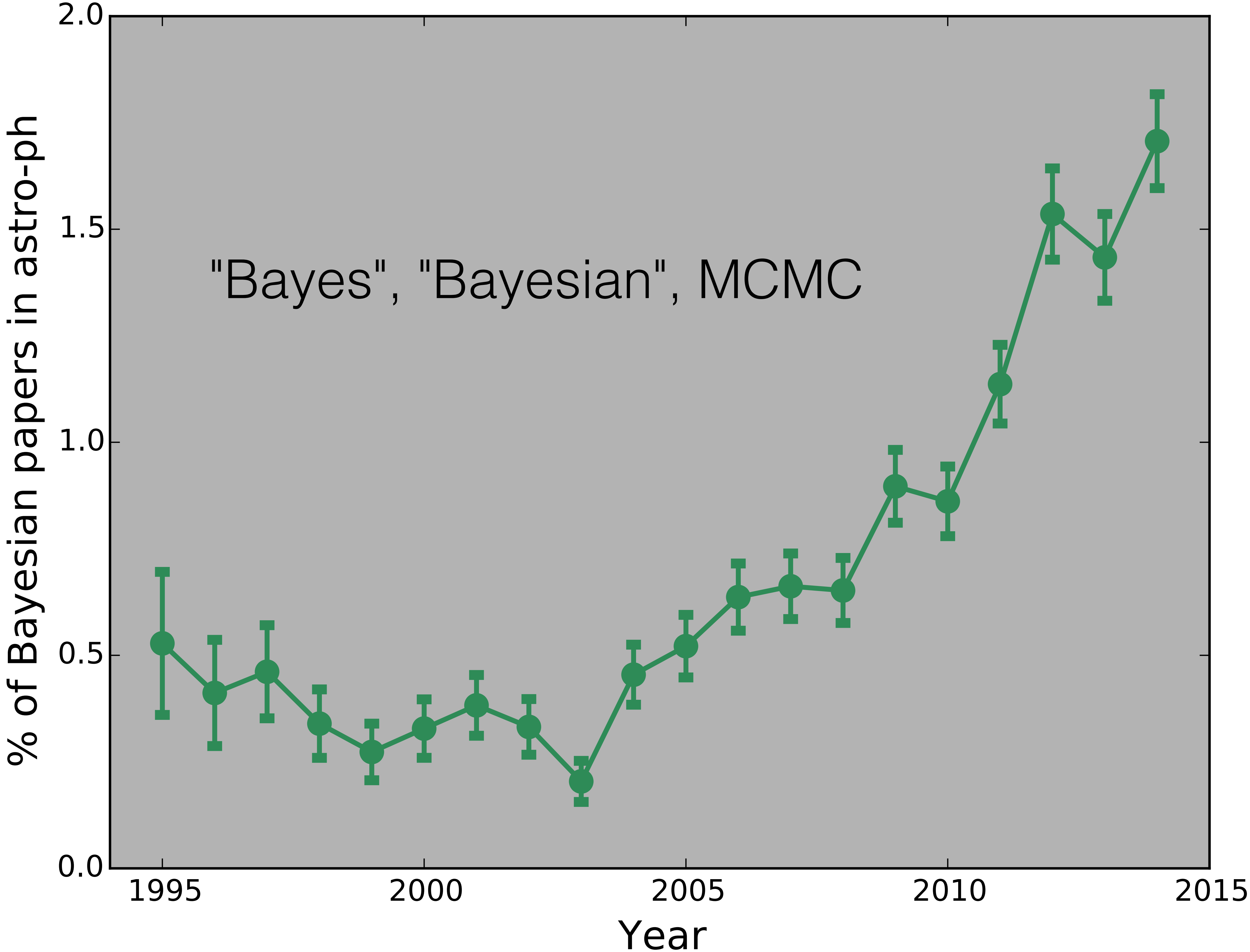
See E.T Jaynes. Probability Theory: The Logic of Science.

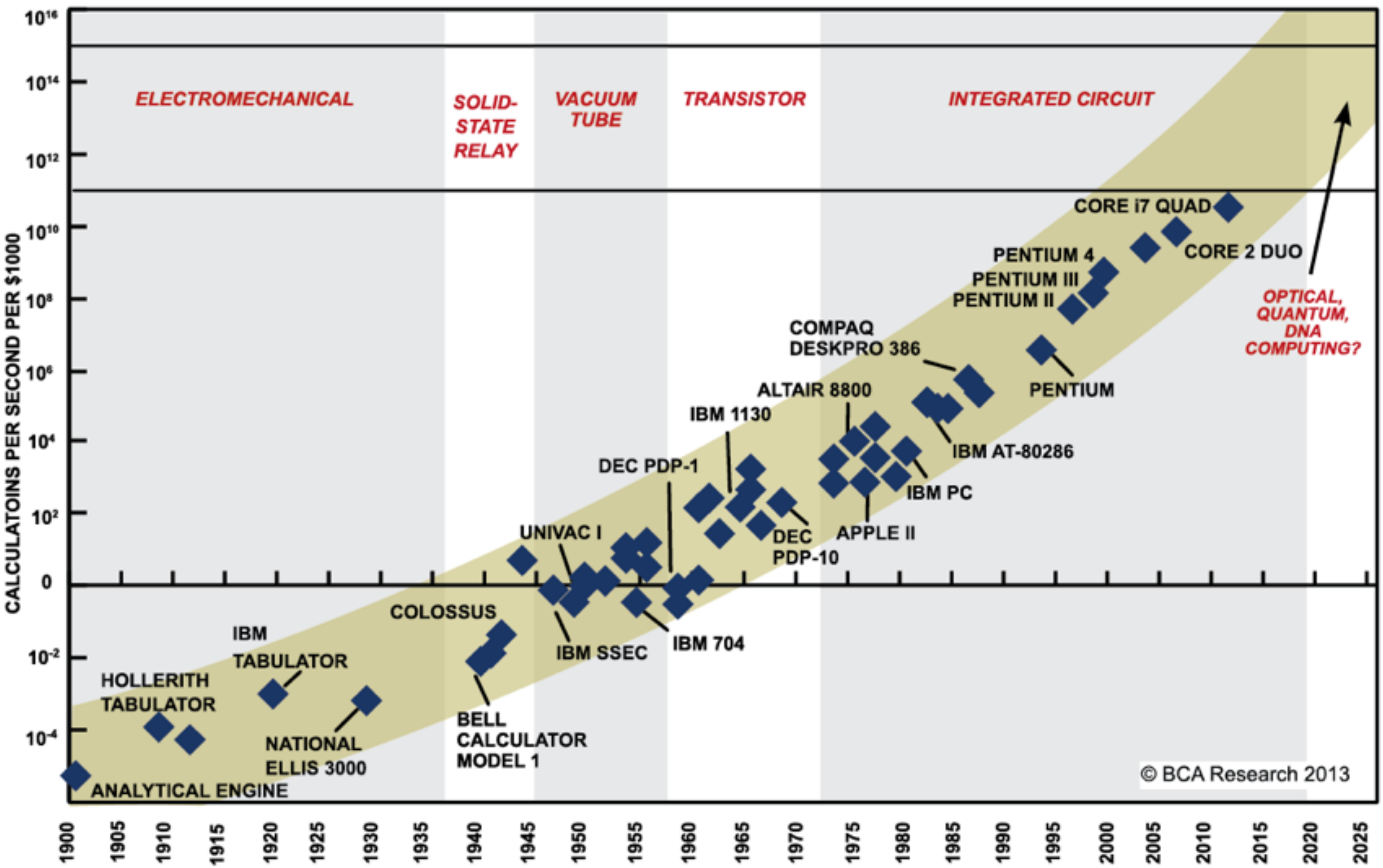
http://www-biba.inrialpes.fr/Jaynes/prob.html

Statistical inference

Fig. adapted from Gregory (2005)




"Bayes", "Bayesian", MCMC
% of Bayesian papers in astro-ph
Year
0.0
0.5
1.0
1.5
2.0
1995
2000
2005
2010
2015



SOURCE: RAY KURZWEIL, "THE SINGULARITY IS NEAR: WHEN HUMANS TRANSCEND BIOLOGY", P.67, *THE VIKING PRESS*, 2006. DATAPOINTS BETWEEN 2000 AND 2012 REPRESENT BCA ESTIMATES.

Two basic tasks of statistical inference

Learning process

(parameter estimation)

Decision making

(model comparison)

Learning process

Bayesian **probability** represents a state of

knowledge

$$p(\bar{\theta}|H_i, I) \longrightarrow p(\bar{\theta}|D, H_i, I)$$

Prior Posterior

$\bar{\theta}$: parameter vector
H_i: hypothesis
I: information
D: data

Discrete space
(hypothesis space)

$$p(H_i|I) \longrightarrow p(H_i|I, D)$$

$\bar{\theta}$: parameter vector
H$_i$: hypothesis
I: information
D: data

Learning process

Enter the **likelihood** function

$$p(\bar{\theta}|H_i, I, D) = \frac{p(D|\bar{\theta}, H_i, I)}{p(D|H_i, I)} \cdot p(\bar{\theta}|H_i, I)$$

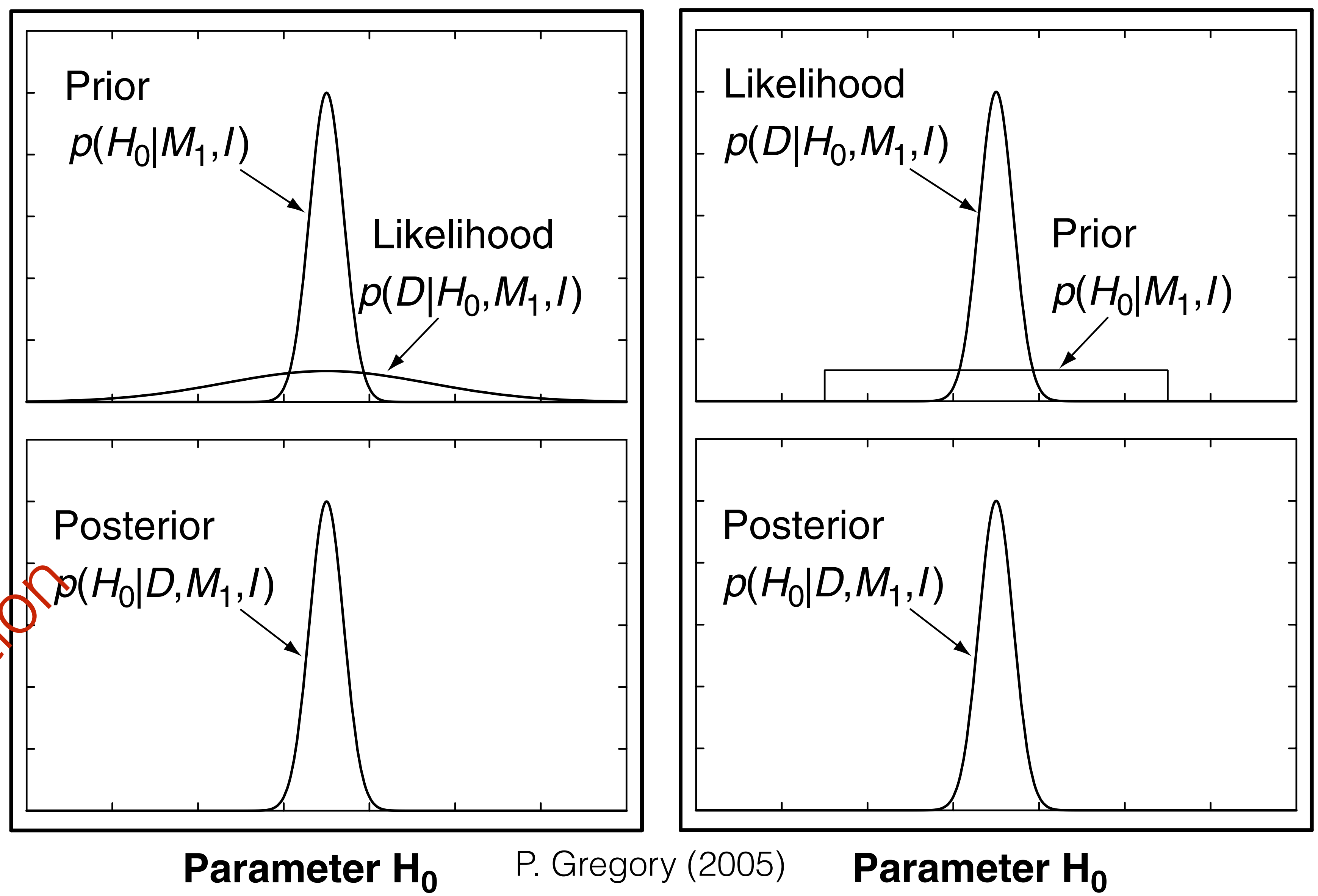
Posterior Prior

$$\boxed{p(\bar{\theta}|D, H_i, I) \propto \mathcal{L}_{\theta}(H_i) \cdot p(\bar{\theta}|H_i, I)}$$

The proportionality constant has many names: marginal likelihood, global likelihood, model evidence, prior predictive. Hard to compute. **Important**.

Optimising the learning process

- The likelihood needs to be selective for the learning process to be effective.



Dominated by prior information

Dominated by data

P. Gregory (2005)

Data analysis

Likelihood functions

$\bar{\theta}$: parameter vector
H$_i$: hypothesis
I: information
D: data

The likelihood function

$$p(\bar{\theta}|H_i, I, D) = \frac{p(D|\bar{\theta}, H_i, I)}{p(D|H_i, I)} \cdot p(\bar{\theta}|H_i, I)$$

Posterior Prior

$$\boxed{p(\bar{\theta}|D, H_i, I) \propto \mathcal{L}_\theta(H_i) \cdot p(\bar{\theta}|H_i, I)}$$

The likelihood is the probability of obtaining data D, for a given prior information I and a set of parameters θ.

Remember, likelihood is *not* a probability for parameter vector θ

(for that you need the prior)

$\bar{\theta}$: parameter vector
H$_i$: hypothesis
I: information
D: data

Likelihood function

Ingredients

Physical model

- Analytic model
- Simulations
- …

Statistical (non-deterministic) model

- Unknown errors (jitter)
- Instrument systematics
- Complex physics (activity, …)
- …

Error statistics

- Covariances
- Non-Gaussianity
- …

$$p(D|\bar{\theta}, H_i, I) = \mathcal{L}_\theta(H_i) \overset{indep.,gauss.}{\propto} \exp -\frac{\chi^2_\theta}{2}$$

Constructing the likelihood

The data:

$$D = D_1\, D_2 \ldots D_n = \{D_i\}$$

D_i: the *i-th* measurement is in the infinitesimal range y_i to $y_i + \mathrm{d}y_i$

The errors:

E_i: the *i-th* error is in the infinitesimal range e_i to $e_i + \mathrm{d}e_i$

$p(E_i|\theta, H, I) = f_E(e_i)$ The probability distribution of statement E_i

Most used f_E

$$f_E(e_i) = N(0, \sigma_i^2)$$

The model:

M_i: the *i-th* error is in the infinitesimal range m_i to $m_i + \mathrm{d}m_i$

$p(M_i|\theta, H, I) = f_M(m_i)$ The probability distribution of statement M_i

Constructing the likelihood

The data:

$$D = D_1\, D_2 \dots D_n = \{D_i\}$$

We want to build the probability distribution:

$$p(D|\theta, H, I) = p(D_1, D_2, \dots, D_n|\theta, H, I)$$

Remember: $y_i = m_i + e_i$

It can be shown that:

$$\boxed{p(D_i|\theta, H, I) = \int \mathrm{d}m_i\, f_M(m_i)\, f_E(y_i - m_i)}$$

Convolution equation

Constructing the likelihood

$$\boxed{p(D_i|\theta, H, I) = \int \mathrm{d}m_i \, f_M(m_i) \, f_E(y_i - m_i)}$$

But for a **deterministic model**, m_i is obtained from a (usually analytically) function f without any uncertainty (say, a Keplerian curve for RV measurements)

$$m_i = f(x_i|\theta)$$
$$f_M(m_i) = \delta(m_i - f(x_i|\theta))$$

Then:

$$p(D_i|\theta, H, I) = \int \mathrm{d}m_i \, \delta(m_i - f(x_i|\theta)) \, f_E(y_i - m_i)$$
$$= f_E(y_i - f(x_i|\theta)) = p(E_i|\theta, H, I)$$

$$\boxed{p(D|\theta, H, I) = p(D_1, D_2, ..., D_n|\theta, H, I) = p(E_1, E_2, ..., E_n|\theta, H, I)}$$

Constructing the likelihood

$$\boxed{p(D|\theta, H, I) = p(D_1, D_2, ..., D_n|\theta, H, I) = p(E_1, E_2, ..., E_n|\theta, H, I)}$$

Gaussian likelihood

$$\ln \mathcal{L} = -\frac{1}{2}\left[n \ln(2\pi) + \ln|K| + \mathbf{r} \cdot K^{-1} \cdot \mathbf{r}^{\intercal}\right]$$

K: covariance of the data; n: number of data points; $\boldsymbol{r} = \boldsymbol{y} - \boldsymbol{m}$: residuals vector

Constructing the likelihood

$$\boxed{p(D|\theta, H, I) = p(D_1, D_2, ..., D_n|\theta, H, I) = p(E_1, E_2, ..., E_n|\theta, H, I)}$$

Now, for **independent errors** (K is diagonal)

$$\begin{aligned} p(D|\theta, H, I) &= p(E_1, E_2, ..., E_n|\theta, H, I) \\ &= p(E_1|\theta, H, I)...p(E_n|\theta, H, I) \\ &= \prod_{i=1}^{n} p(E_i|\theta, H, I) \end{aligned}$$

$$\ln \mathcal{L} = -\frac{1}{2}\left[n\ln(2\pi) + \sum_{i=1}^{n} \ln \sigma_i^2 + \chi^2\right] \propto \chi^2$$

Constructing the likelihood

Back to the **convolution equation**

$$\boxed{p(D_i|\theta, H, I) = \int \mathrm{d}m_i\, f_M(m_i)\, f_E(y_i - m_i)}$$

For a **non-deterministic model**, M_i is distributed:

M_i: the *i-th* error is in the infinitesimal range m_i to $m_i + \mathrm{d}m_i$

$p(M_i|\theta, H, I) = f_M(m_i)$ The probability distribution of statement M_i

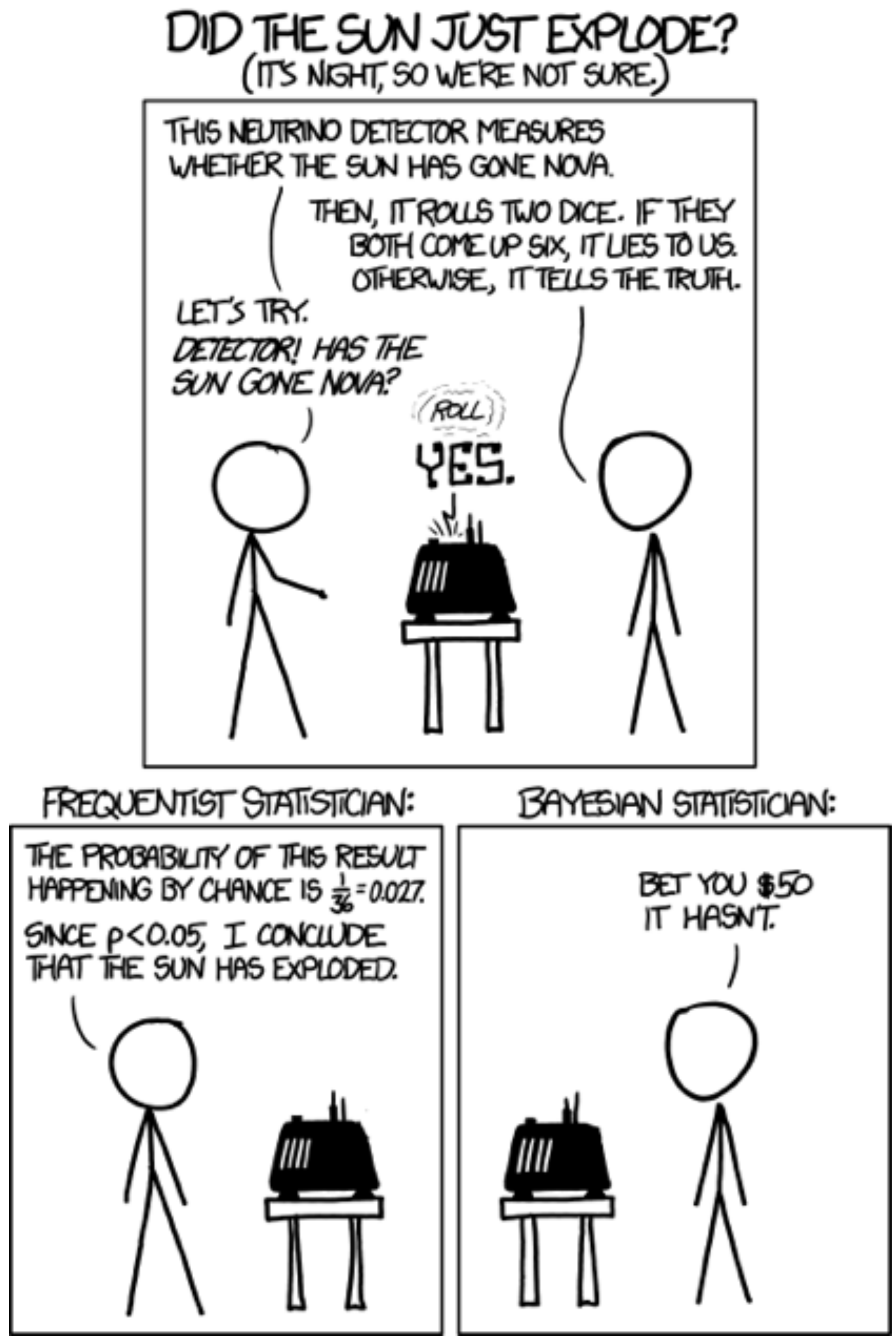
E.g. adding instrumental error / resolution:

$$f_M(m_i) = N(f(x_i|\theta), \sigma^2_{\mathrm{inst}})$$

Data analysis

Priors

Ignorance priors


DID THE SUN JUST EXPLODE?
(IT'S NIGHT, SO WE'RE NOT SURE.)
THIS NEUTRINO DETECTOR MEASURES WHETHER THE SUN HAS GONE NOVA.
THEN, IT ROLLS TWO DICE. IF THEY BOTH COME UP SIX, IT LIES TO US. OTHERWISE, IT TELLS THE TRUTH.
LET'S TRY.
DETECTOR! HAS THE SUN GONE NOVA?
ROLL
YES.
FREQUENTIST STATISTICIAN:
THE PROBABILITY OF THIS RESULT HAPPENING BY CHANCE IS $\frac{1}{36}$=0.027.
SINCE $p<0.05$, I CONCLUDE THAT THE SUN HAS EXPLODED.
BAYESIAN STATISTICIAN:
BET YOU $50 IT HASN'T.

xkcd.com

Learning process

$\bar{\theta}$: parameter vector
H$_i$: hypothesis
I: information
D: data

The role of the **prior** distribution

$$p(\bar{\theta}|H_i, I, D) = \frac{p(D|\bar{\theta}, H_i, I)}{p(D|H_i, I)} \cdot p(\bar{\theta}|H_i, I)$$

Posterior

Prior

$$\boxed{p(\bar{\theta}|D, H_i, I) \propto \mathcal{L}_\theta(H_i) \cdot p(\bar{\theta}|H_i, I)}$$

Prior probabilities

$$p(H_i|I)$$

H_i: hypothesis (can be continuous).
I: information

- Prior information I is always present:
    - The term "prior" does not necessarily mean "earlier in time".
- Philosophical controversy on how to assign priors.
    - Subjective vs. objective views.
    - No single universal rule, but a few accepted methods.
- **Informative** priors. Usually based on the output from previous observations. (What was the prior of the first analysis?).
- **Ignorance** priors. Required as a starting point for the theory.

Assigning ignorance priors

1. Principle of indifference.

Given n mutually exclusive, exhaustive hypothesis, $\{H_i\}$, with $i = 1, \ldots, n$, the PoI states:

$$p(H_i|I) = 1/n$$

Assigning ignorance priors

2. Transformation groups. Location and scale parameters.

For a certain type of parameters (location and scale), "total ignorance" can we represented as invariance under certain (group of) transformation.

Location: "position of highest tree along a river."
Problem must be invariant under a translation.

$$X' = X + c$$

$$p(X|I)\,\mathrm{d}X = p(X'|I)\,\mathrm{d}X' =$$
$$p(X'|I)\mathrm{d}(X+c) = p(X'|I)\mathrm{d}X$$

$$\boxed{p(X|I) = \text{constant}}$$

Uniform prior.

Assigning ignorance priors

2. Transformation groups. Location and scale parameters.

For a certain type of parameters (location and scale), "total ignorance" can we represented as invariance under certain (group of) transformation.

Scale: "life time of a new bacteria" or "Poisson rate"
Problem must be invariant under scaling.

$$X' = aX$$

$$p(X|I)\,\mathrm{d}X = p(X'|I)\,\mathrm{d}X' =$$
$$p(X'|I)\mathrm{d}(aX) = ap(X'|I)\mathrm{d}X$$

$$\boxed{p(X|I) = \frac{\mathrm{constant}}{x}}$$

"Jeffreys" prior.

Assigning ignorance priors

3. Jeffreys rule.

Besides location and scale parameters, little more can be said using transformation invariance.

Jeffreys priors use the Fisher information; parameterisation invariant, but strange behaviour in many dimensions.

Observed Fischer information: $$I_D = -\frac{\mathrm{d}^2 \log \mathcal{L}_D}{\mathrm{d}\theta^2}$$

But D is not known when we have to define a prior. Use expectation value over D.

$$I(\theta) = -\mathrm{E}_D\left[\frac{\mathrm{d}^2 \log \mathcal{L}_D}{\mathrm{d}\theta^2}\right]$$

Assigning ignorance priors

3. Jeffreys rule says:

$$\boxed{p(\theta|I) \propto \sqrt{I(\theta)}}$$

Examples:

- Mean of Normal distribution (μ) with known variance σ^2.

$$p(\mu|\sigma^2, I) \propto \text{constant}$$

- Rate λ of Poisson distribution.

$$p(\lambda|I) \propto 1/\sqrt{\lambda}$$

- Exercise: Scale of Normal with known mean value?

Assigning ignorance priors

3. Jeffreys rule:

$$\boxed{p(\theta|I) \propto \sqrt{I(\theta)}}$$

$$I(\theta) = -\mathrm{E}_D\left[\frac{\mathrm{d}^2 \log \mathcal{L}_D}{\mathrm{d}\theta^2}\right]$$

- Favours parts of parameter space where data provides more information.
- Is invariant under reparametrisation.
- Works fine only in one dimension…

See more examples here: en.wikipedia.org/wiki/Jeffreys_prior

Data analysis

Sampling the posterior (MCMC)

Sampling from the posterior

$$p(\bar{\theta}|D, H_i, I) \propto \mathcal{L}_\theta(H_i) \cdot p(\bar{\theta}|H_i, I)$$

- The posterior distribution is proportional to the likelihood times the prior.
- The normalising constant (called model evidence, marginal likelihood, etc.) is of importance when comparing different models.
- The posterior contains all the information on a given model a Bayesian statistician can get for a given set of priors and data.
- Posterior is only analytical in few cases:
  - Conjugate priors.
- Other methods needed to sample from posterior.

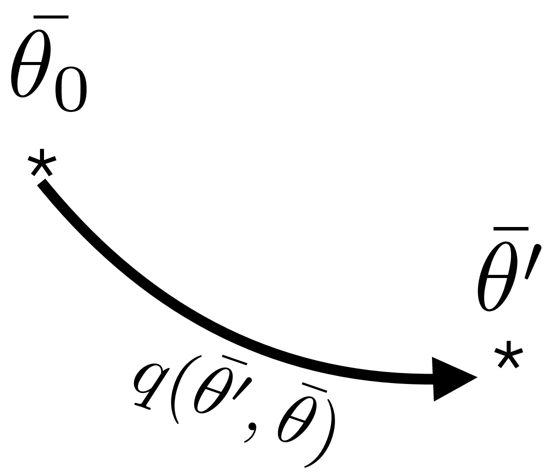
Most Bayesian computations can be reduced to expectation values with respect to the posterior.

$\bar{\theta}$: parameter vector
H$_i$: hypothesis
I: information
D: data

Markov Chain Monte Carlo

$$p(\bar{\theta}|D, H_i, I) \propto \mathcal{L}_\theta(H_i) \cdot p(\bar{\theta}|H_i, I)$$

Metropolis-Hastings



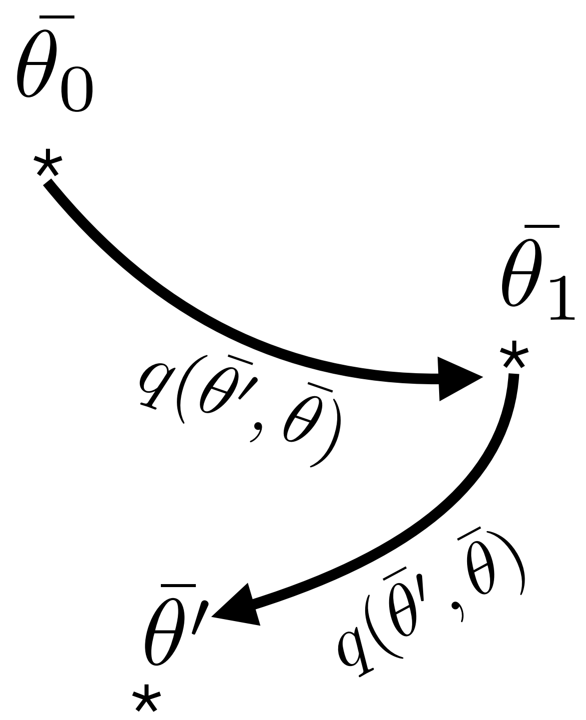
1. $\mathcal{L}_{\theta_0} \cdot p(\bar{\theta}_0|I)$
2. Create proposal point.
3. $\mathcal{L}_{\theta'} \cdot p(\bar{\theta}'|I)$
4. $r = \dfrac{\mathcal{L}_{\theta'} \cdot p(\bar{\theta}'|I)}{\mathcal{L}_{\theta_0} \cdot p(\bar{\theta}_0|I)}$

Markov Chain Monte Carlo

$\bar{\theta}$: parameter vector
$\mathbf{H_i}$: hypothesis
$\mathbf{I}$: information
$\mathbf{D}$: data

$$p(\bar{\theta}|D, H_i, I) \propto \mathcal{L}_\theta(H_i) \cdot p(\bar{\theta}|H_i, I)$$

Metropolis-Hastings



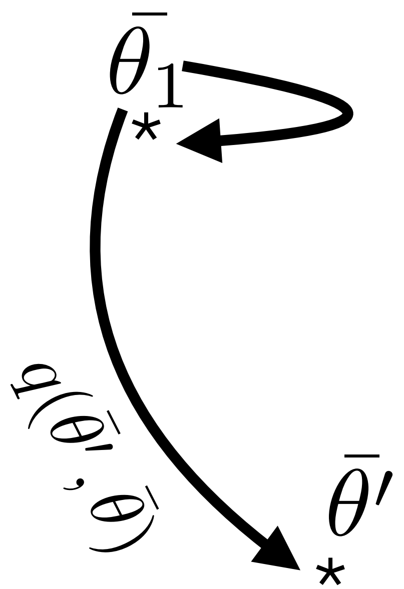
1. $\mathcal{L}_{\theta_0} \cdot p(\bar{\theta}_0|I)$
2. Create proposal point.
3. $\mathcal{L}_{\theta'} \cdot p(\bar{\theta}'|I)$
4. $r = \frac{\mathcal{L}_{\theta'} \cdot p(\bar{\theta}'|I)}{\mathcal{L}_{\theta_0} \cdot p(\bar{\theta}_0|I)}$
5. Accept proposal with probability $\min(1, r)$

Markov Chain Monte Carlo

$\bar{\theta}$: parameter vector
H$_i$: hypothesis
I: information
D: data

$$p(\bar{\theta}|D, H_i, I) \propto \mathcal{L}_\theta(H_i) \cdot p(\bar{\theta}|H_i, I)$$

Metropolis-Hastings



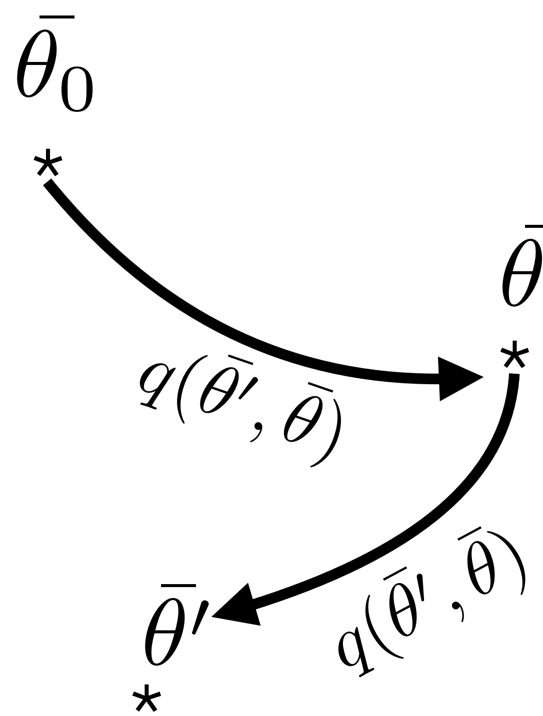
1. $\mathcal{L}_{\theta_0} \cdot p(\bar{\theta}_0|I)$
2. Create proposal point.
3. $\mathcal{L}_{\theta'} \cdot p(\bar{\theta}'|I)$
4. $r = \frac{\mathcal{L}_{\theta'} \cdot p(\bar{\theta}'|I)}{\mathcal{L}_{\theta_0} \cdot p(\bar{\theta}_0|I)}$
5. Accept proposal with probability $\min(1, r)$

Markov Chain Monte Carlo

$\bar{\theta}$: parameter vector
H$_i$: hypothesis
I: information
D: data

$$p(\bar{\theta}|D, H_i, I) \propto \mathcal{L}_\theta(H_i) \cdot p(\bar{\theta}|H_i, I)$$

Metropolis-Hastings



Algorithms

Metropolis-Hastings
Gibbs sampling
Slice sampling
...
Hybrid Monte Carlo

Codes

pymc
emcee
kombine
cobmcmc
...

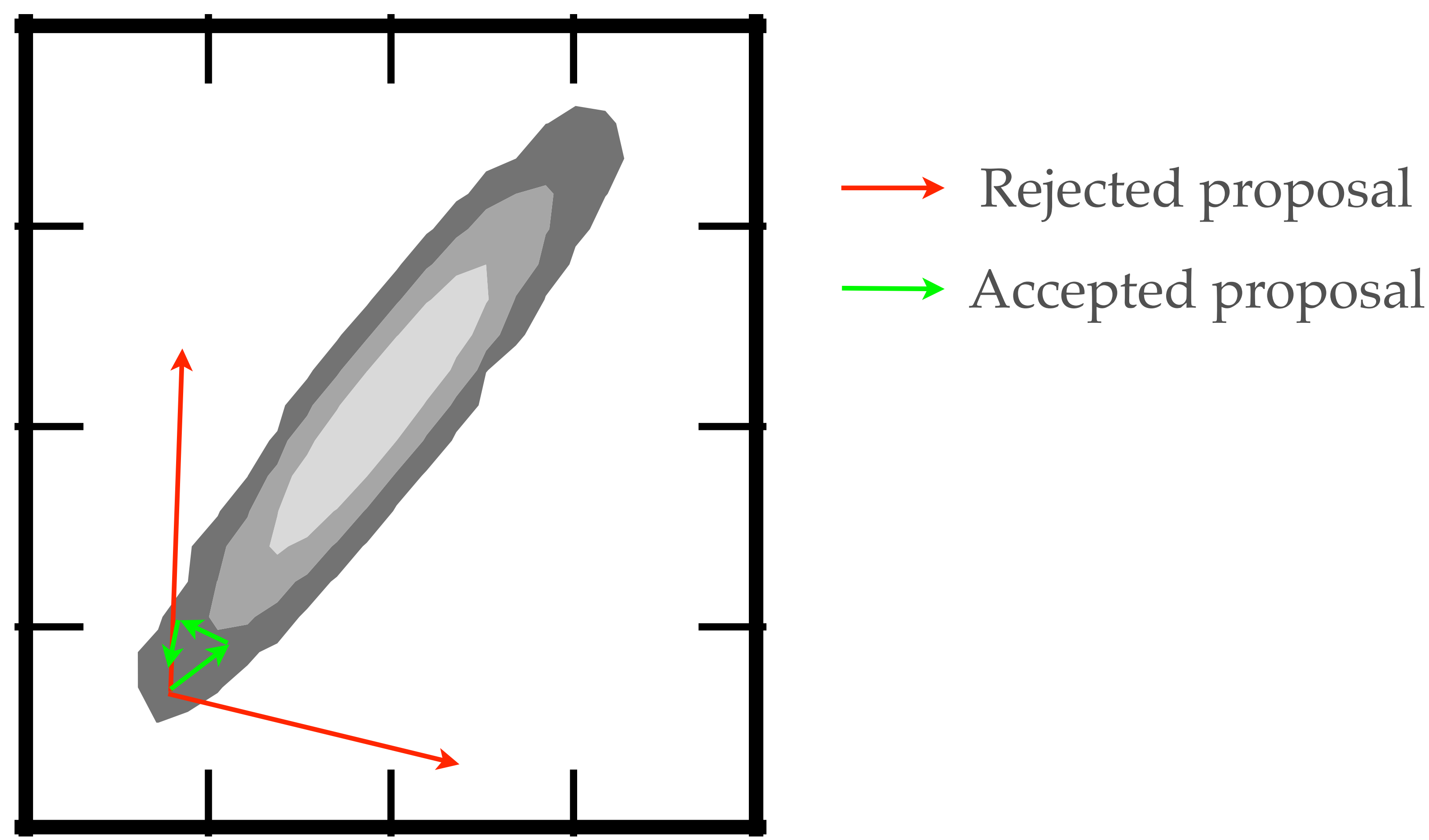

5 minutes
to LEARN
a lifetime
to MASTER

But beware. Non-convergence, bad mixing. The dark side of MCMC are they.

Problems with correlations and multi-modal distributions.

The problem with Correlations

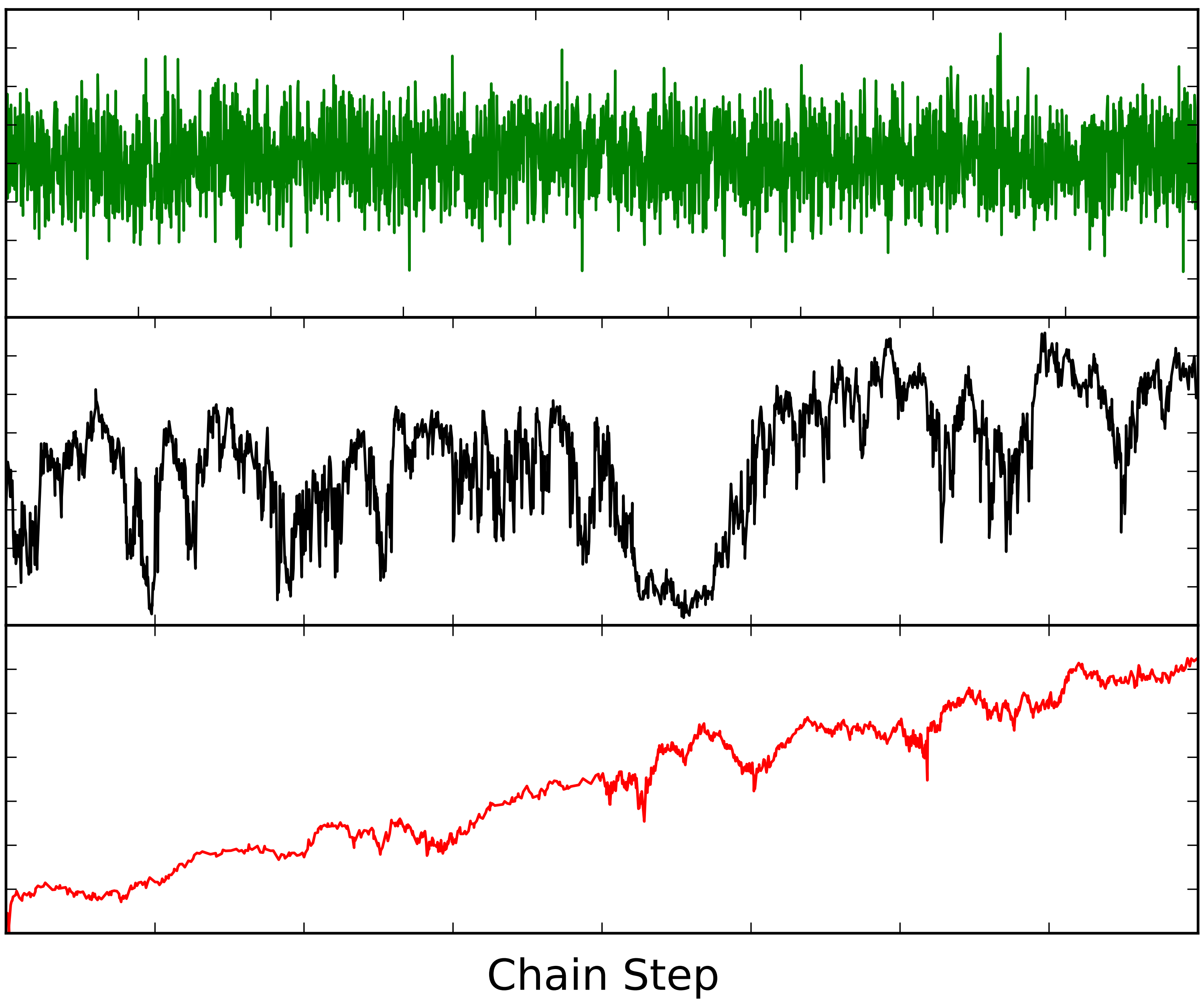
If parameters exhibit correlations, then step size must be small to reach the demanded fraction of accepted jumps.



Need a very long chain to explore the entire posterior. Or, more relevant, the *entire posterior will not be explored thoroughly* (i.e. reduced error bars!)

MCMC: the Good, the Bad, and the Ugly

Visual inspection of traces.



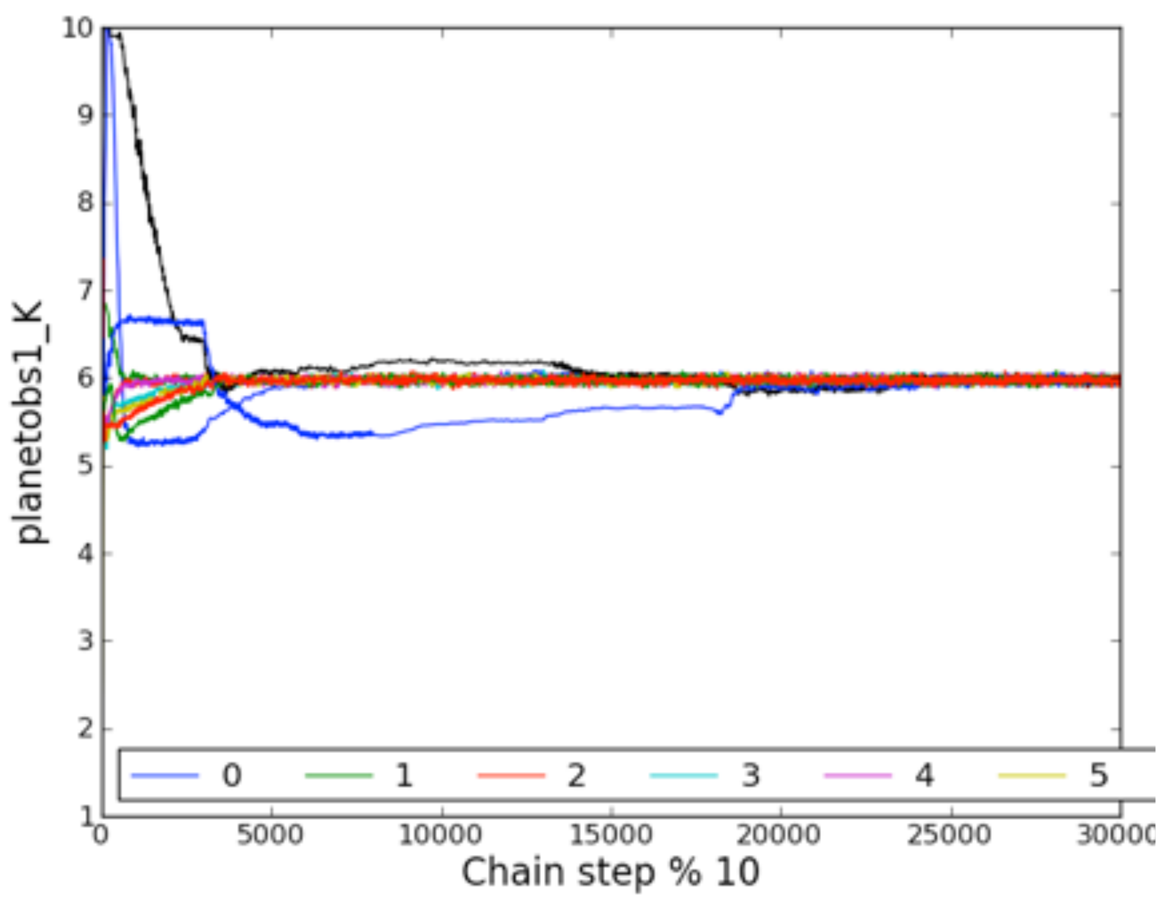
Good

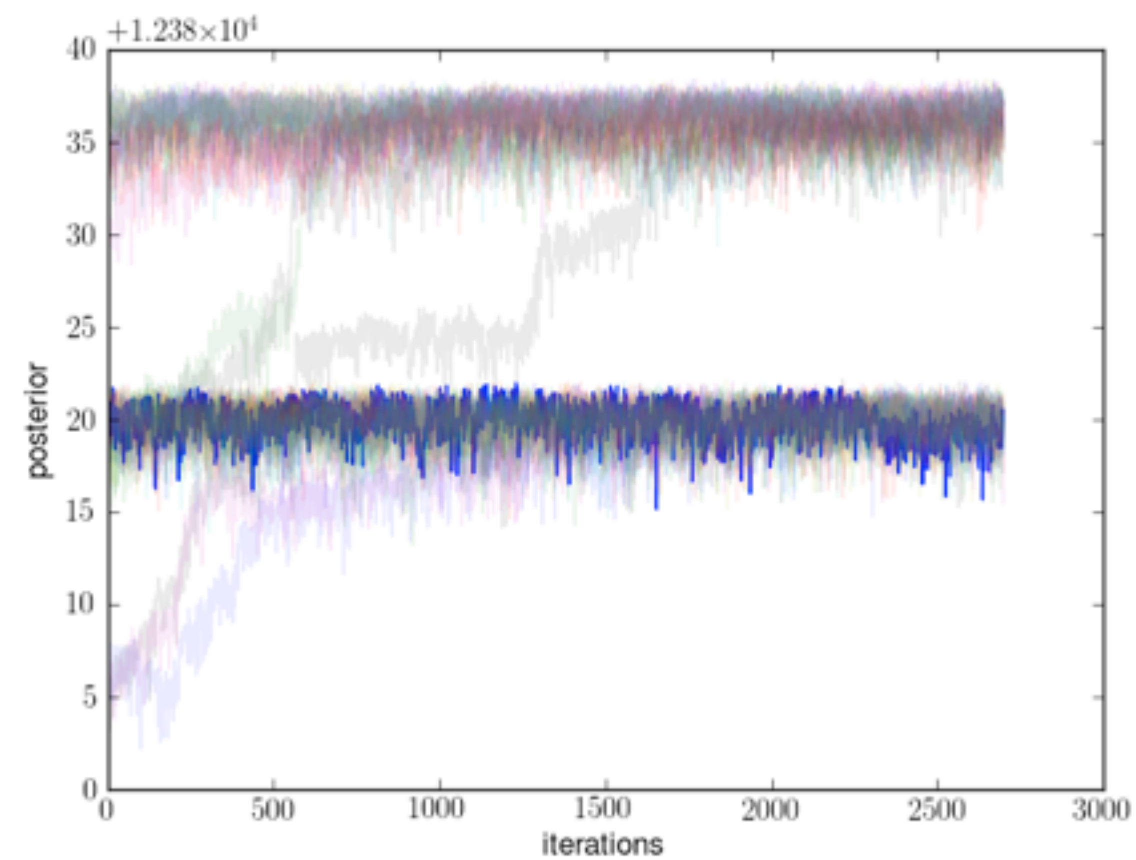
Marginal mixing

No convergence

Multi-modal posteriors.

Run as many chains as possible starting from significantly different places in the prior space.





Be paranoid! You can always be missing modes.

Physical models

Radial velocity

Transit light curve.

$$\underset{\text{data}}{d_i} = \underset{\text{model}}{m_i} + \underset{\text{error}}{e_i}$$

Data analysis

Likelihood functions

MCMC

$p(D|\theta, H, I)$

Error model

Beyond the independent normal paradigm

Error models

Gaussian process regression

The Gaussian likelihood

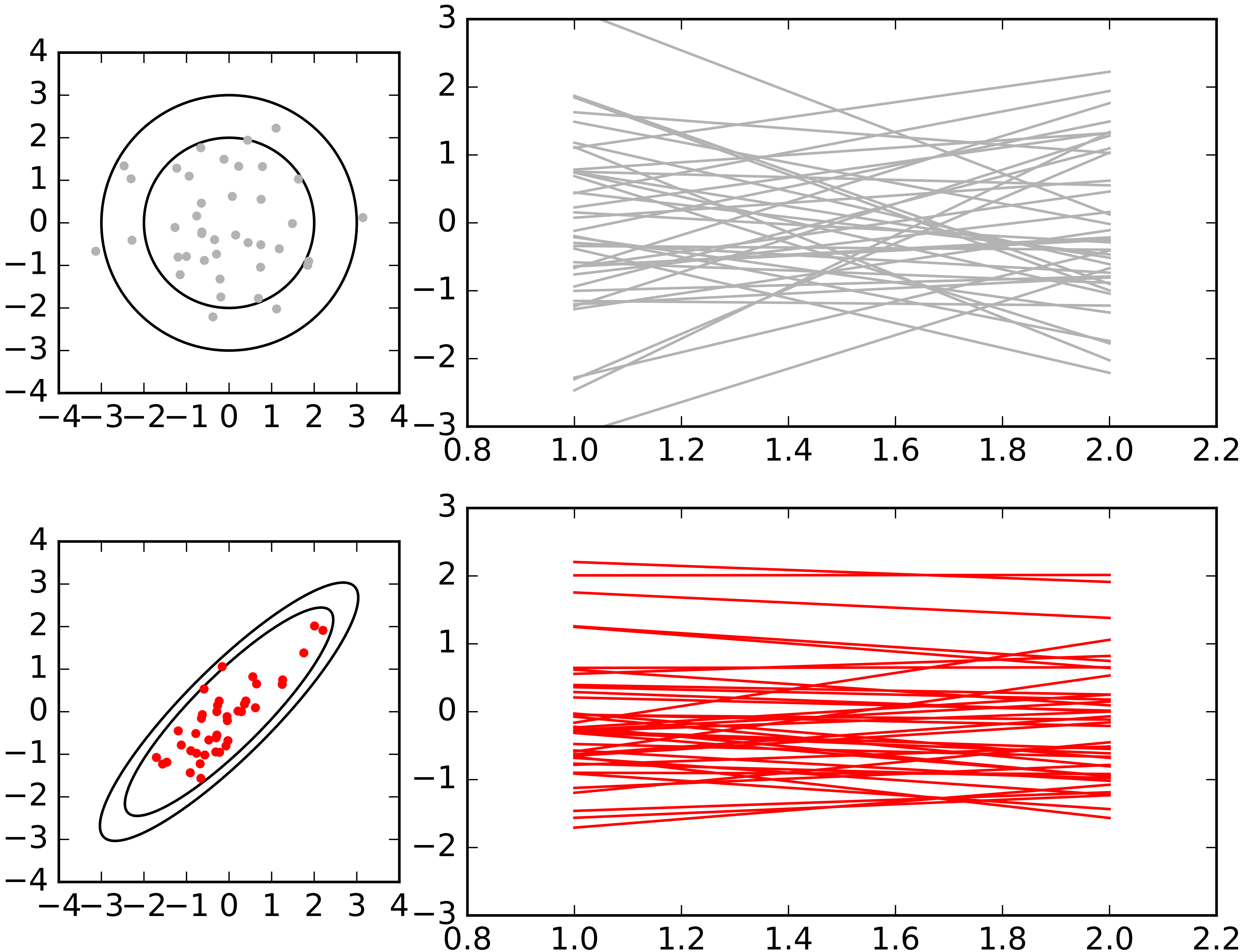
$$\ln \mathcal{L} = -\frac{1}{2}\left[n\ln(2\pi) + \ln|K| + \mathbf{r}\cdot K^{-1}\cdot \mathbf{r}^{\mathsf{T}}\right]$$

K: covariance of the data; n: number of data points; $\boldsymbol{r} = \boldsymbol{y} - \boldsymbol{m}$: residuals vector

If K is diagonal (i.e. uncorrelated errors σ_i)

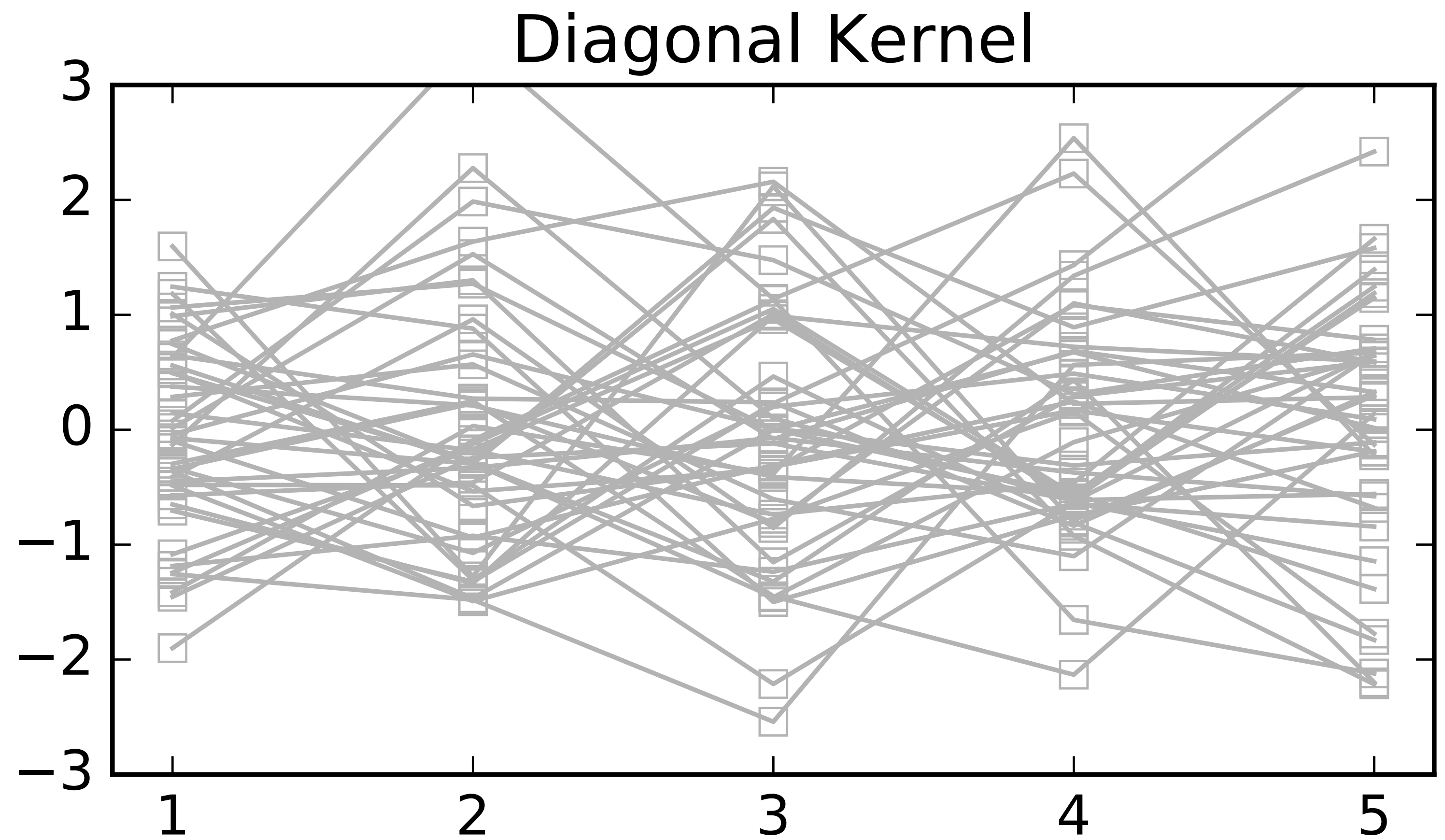
$$\ln \mathcal{L} = -\frac{1}{2}\left[n\ln(2\pi) + \sum_{i=1}^{n}\ln\sigma_i^2 + \chi^2\right] \propto \chi^2$$

Alternatively, K can be generated by a kernel function $k(x_i, x_j)$

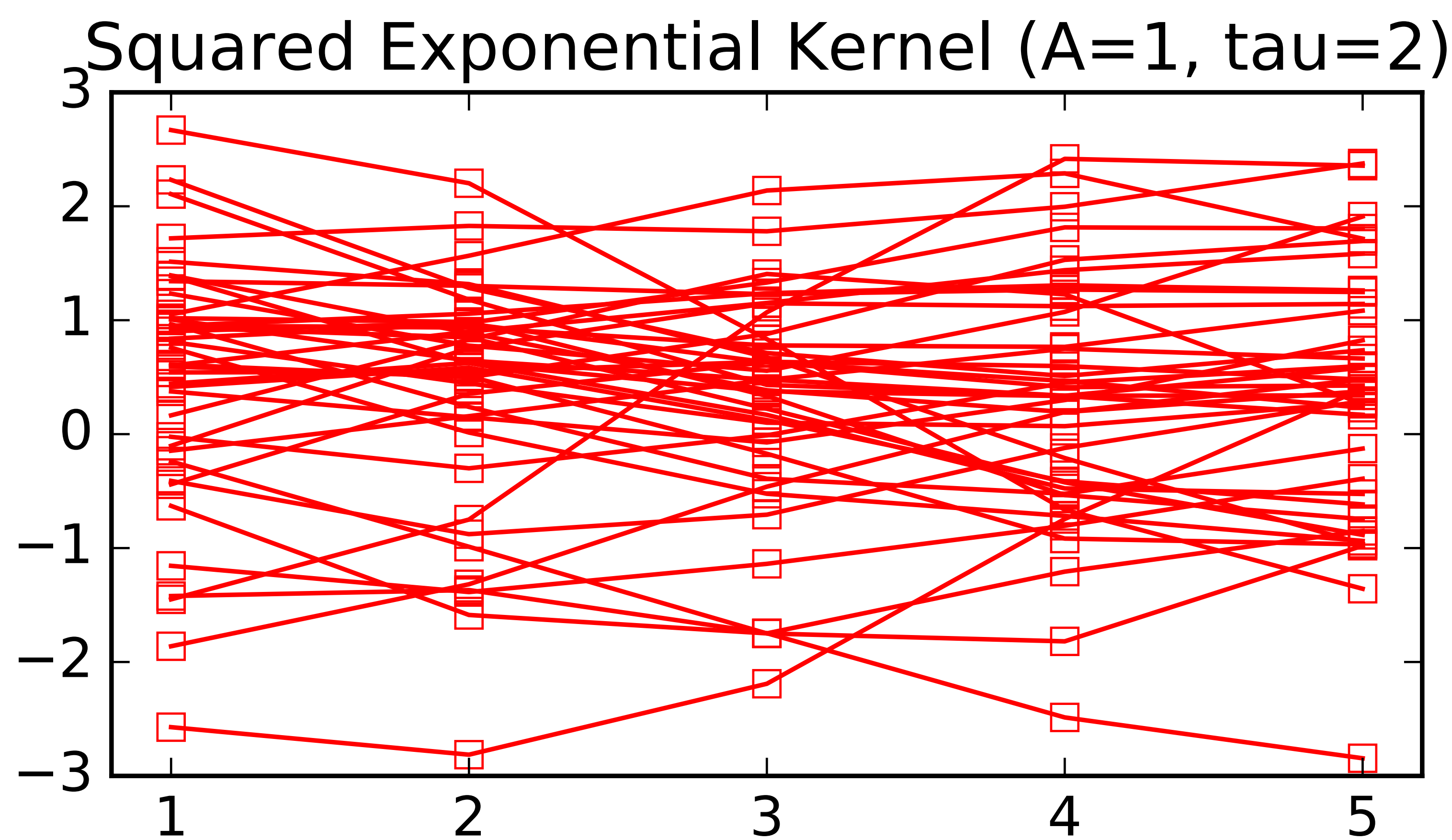


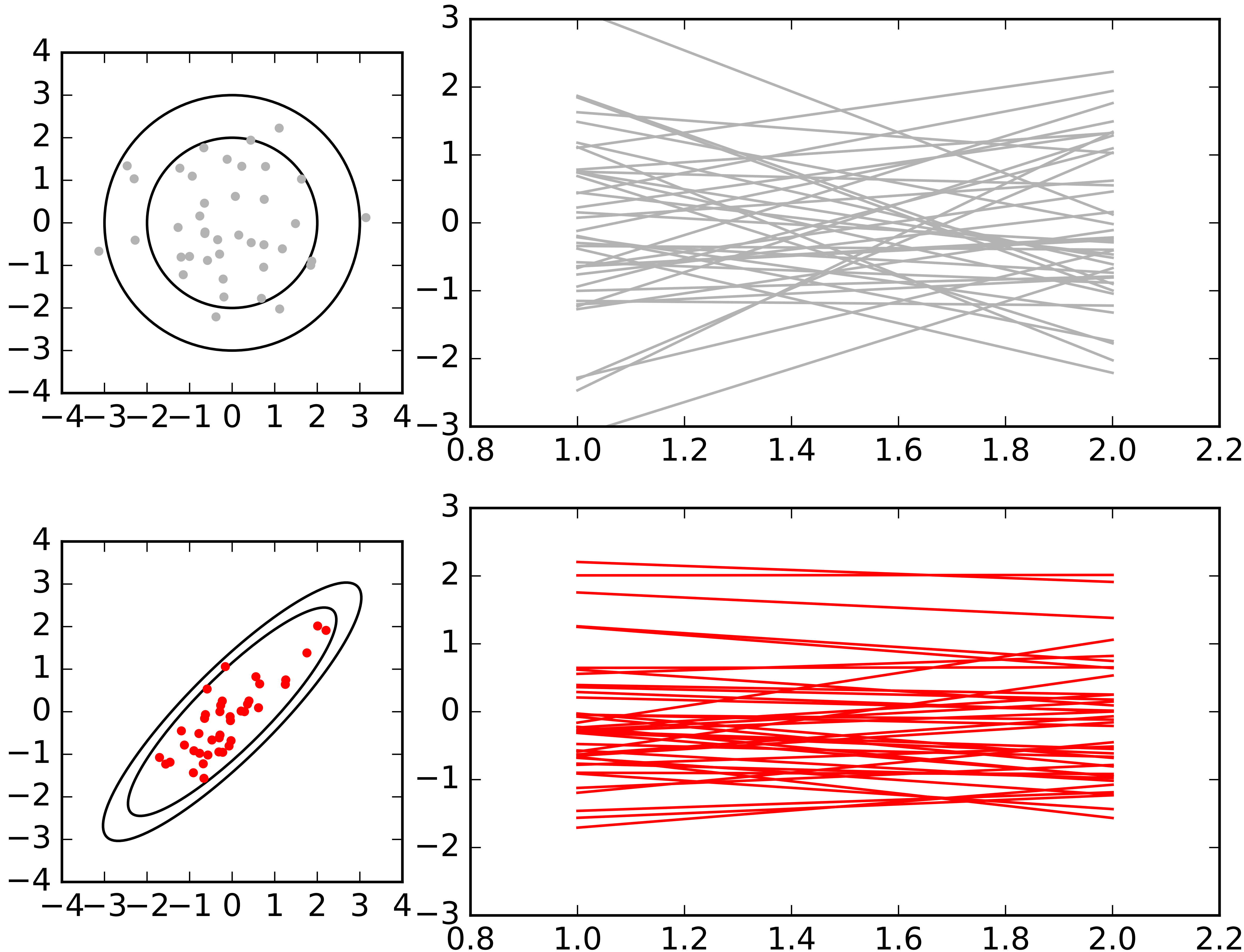
4
3
2
1
0
−1
−2
−3
−4
−4 −3 −2 −1 0 1 2 3 4
0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2

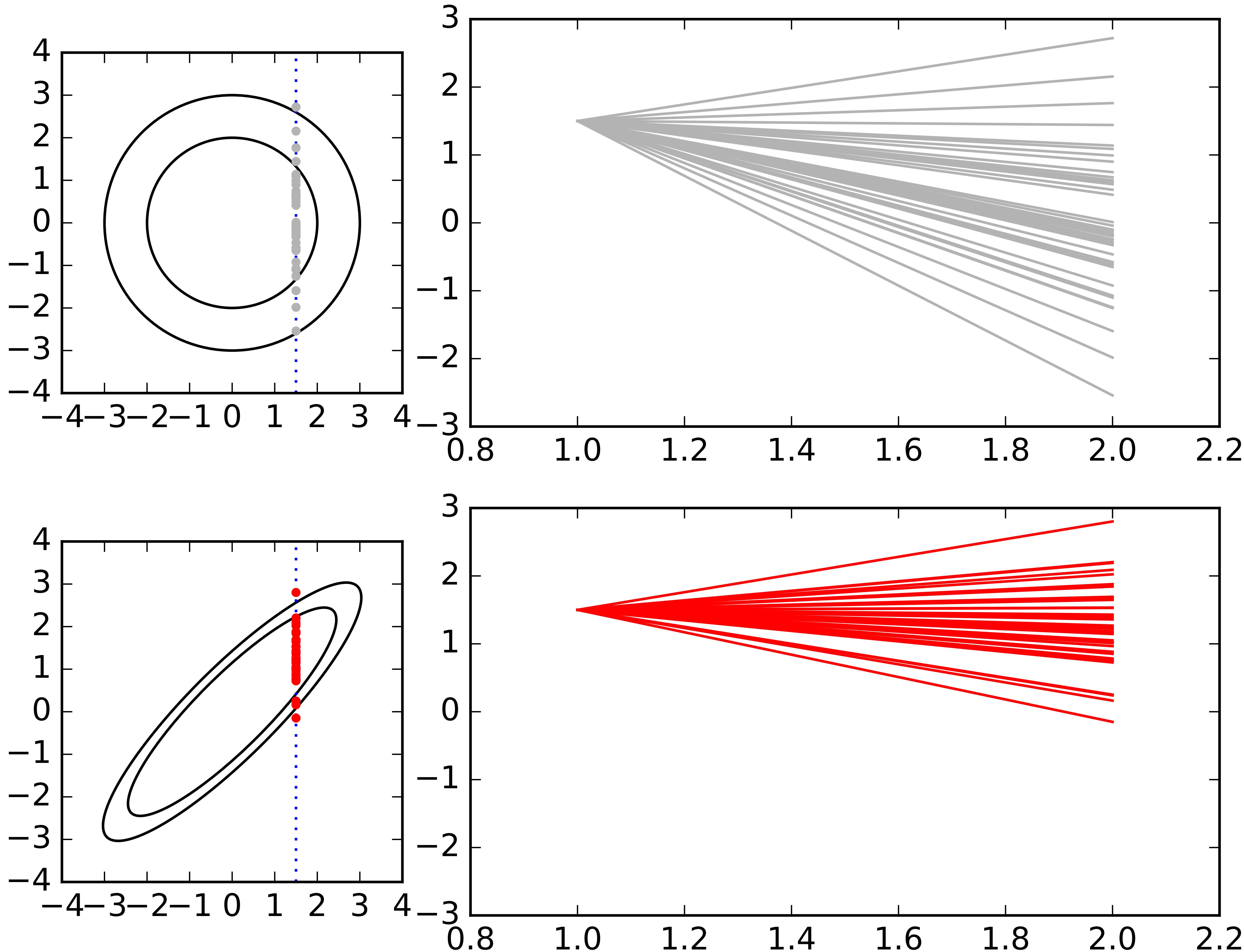
1.,	0.,	0.,	0.,	0.
0.,	1.,	0.,	0.,	0.
0.,	0.,	1.,	0.,	0.
0.,	0.,	0.,	1.,	0.
0.,	0.,	0.,	0.,	1.



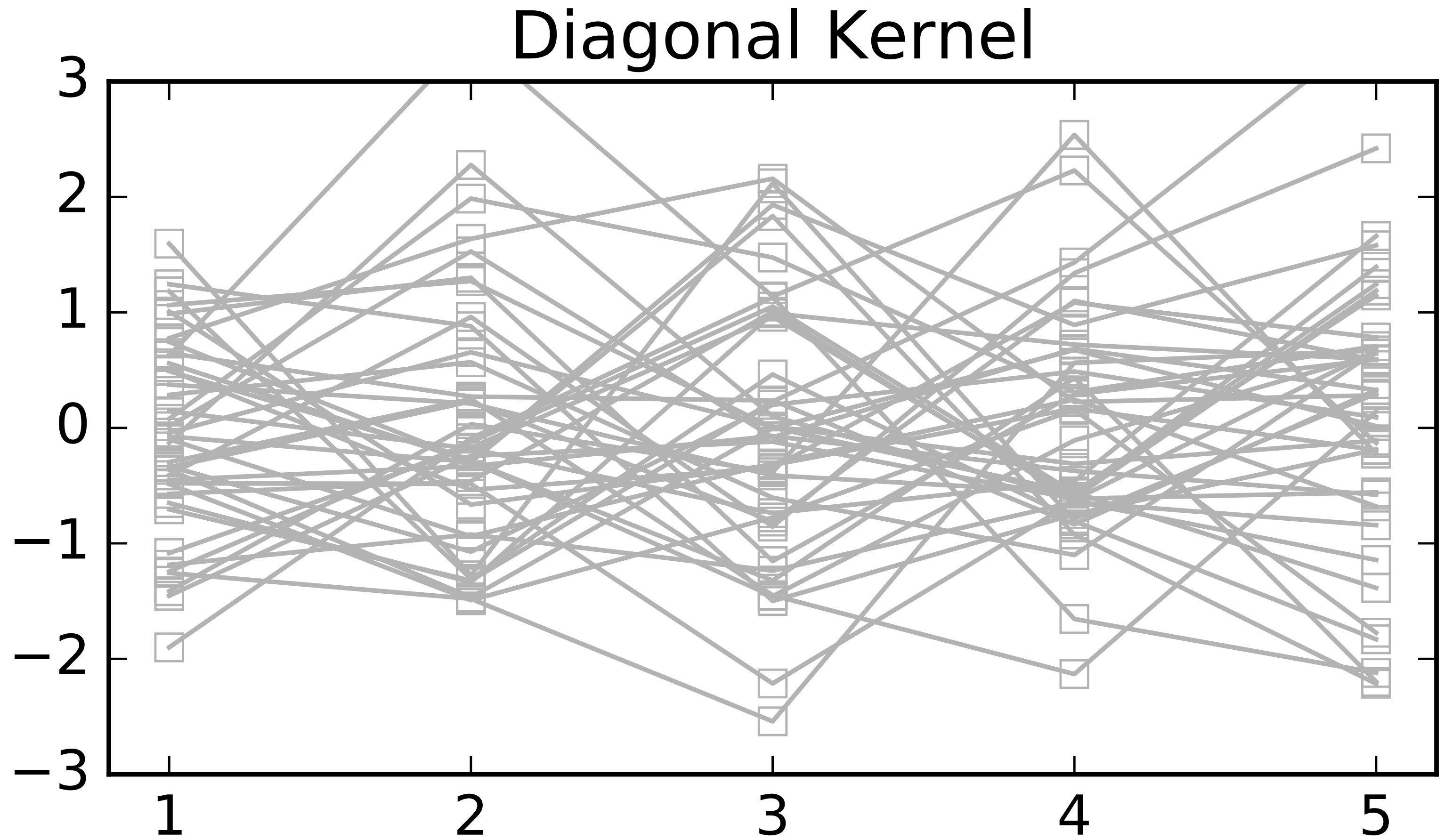
1. ,	0.88,	0.61,	0.32,	0.14
0.88,	1. ,	0.88,	0.61,	0.32
0.61,	0.88,	1. ,	0.88,	0.61
0.32,	0.61,	0.88,	1. ,	0.88
0.14,	0.32,	0.61,	0.88,	1.



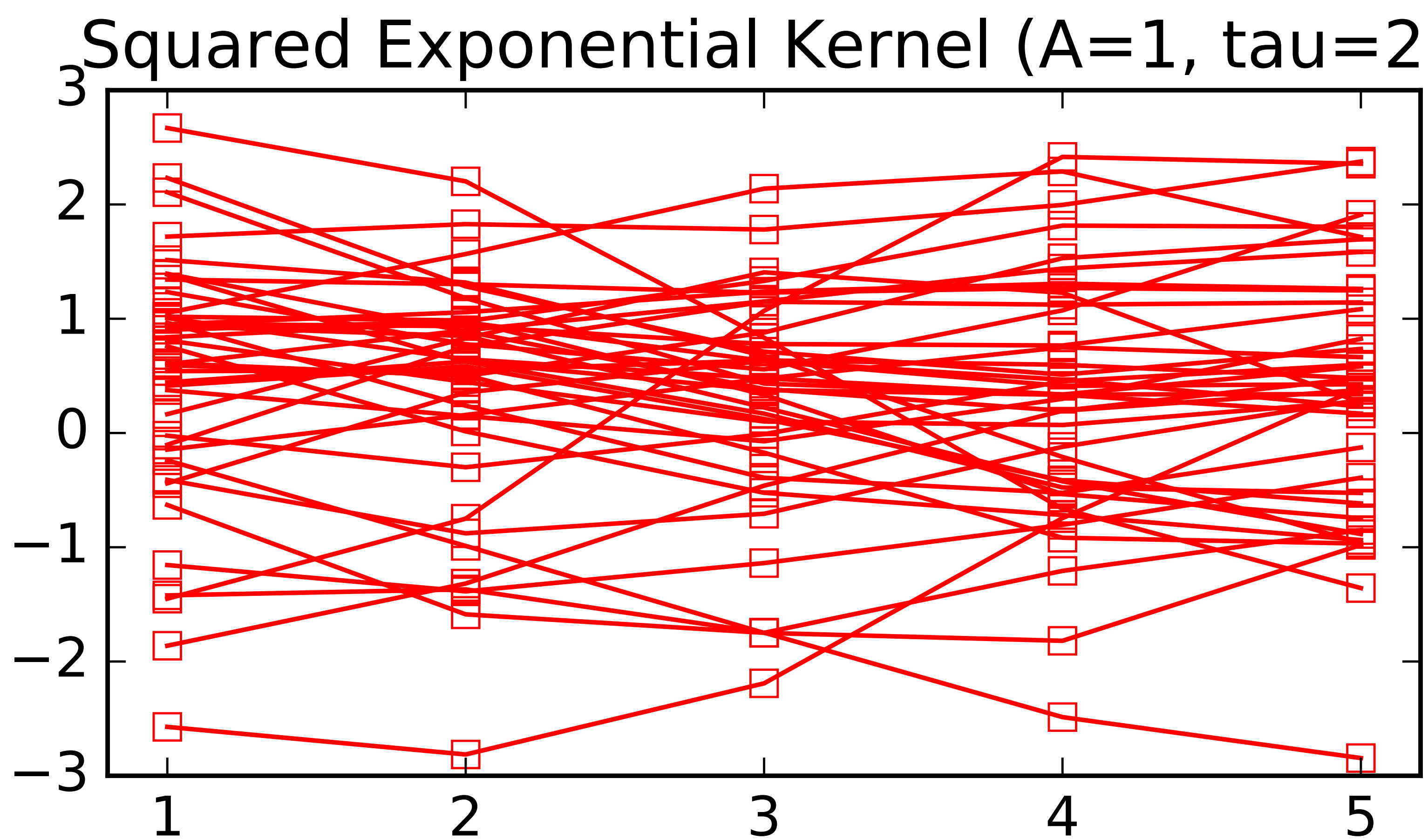

4
3
2
1
0
−1
−2
−3
−4
−4 −3 −2 −1 0 1 2 3 4
3
2
1
0
−1
−2
−3
0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2

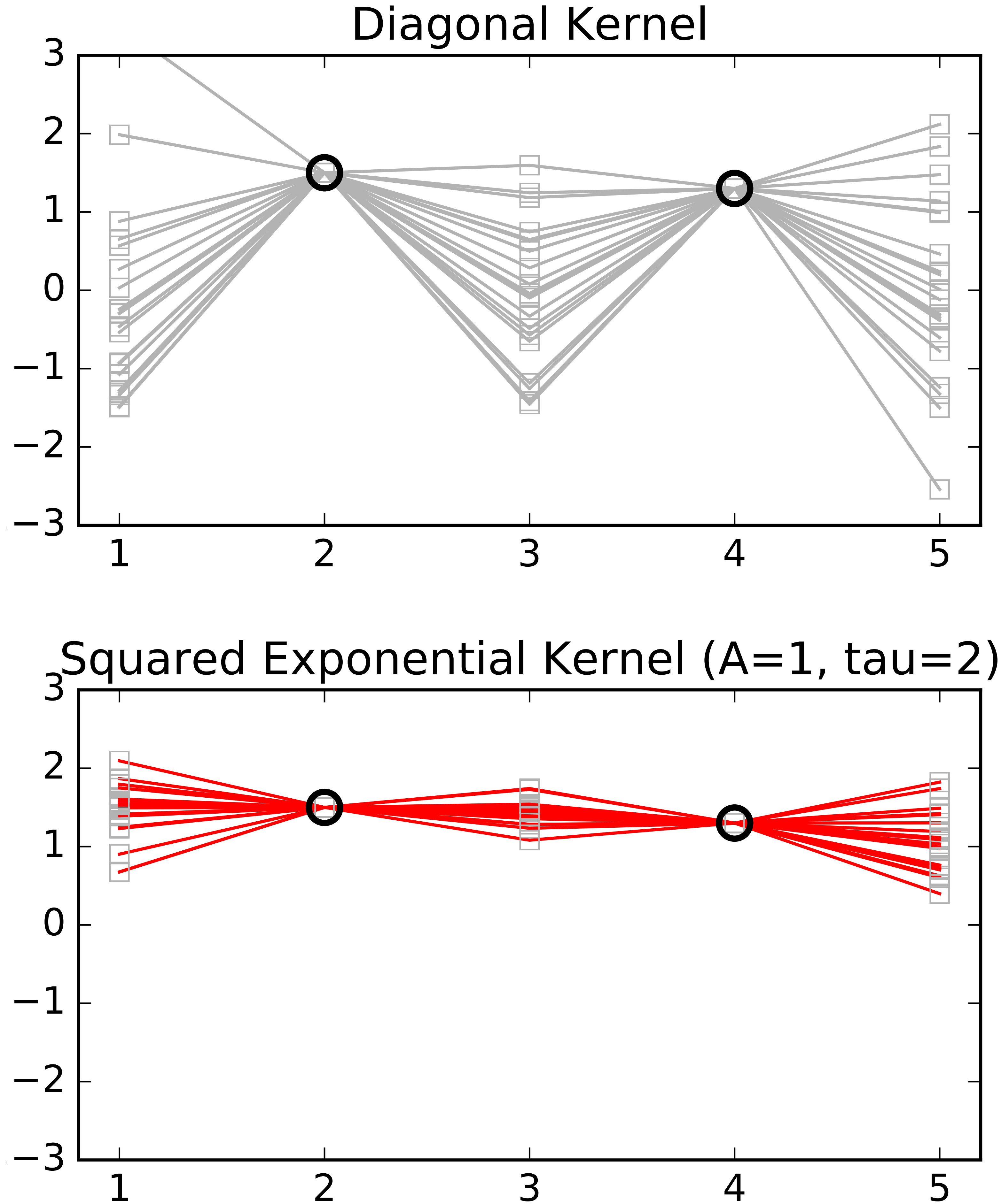

4
3
2
1
0
−1
−2
−3
−4
−4 −3 −2 −1 0 1 2 3 4
3
2
1
0
−1
−2
−3
0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2

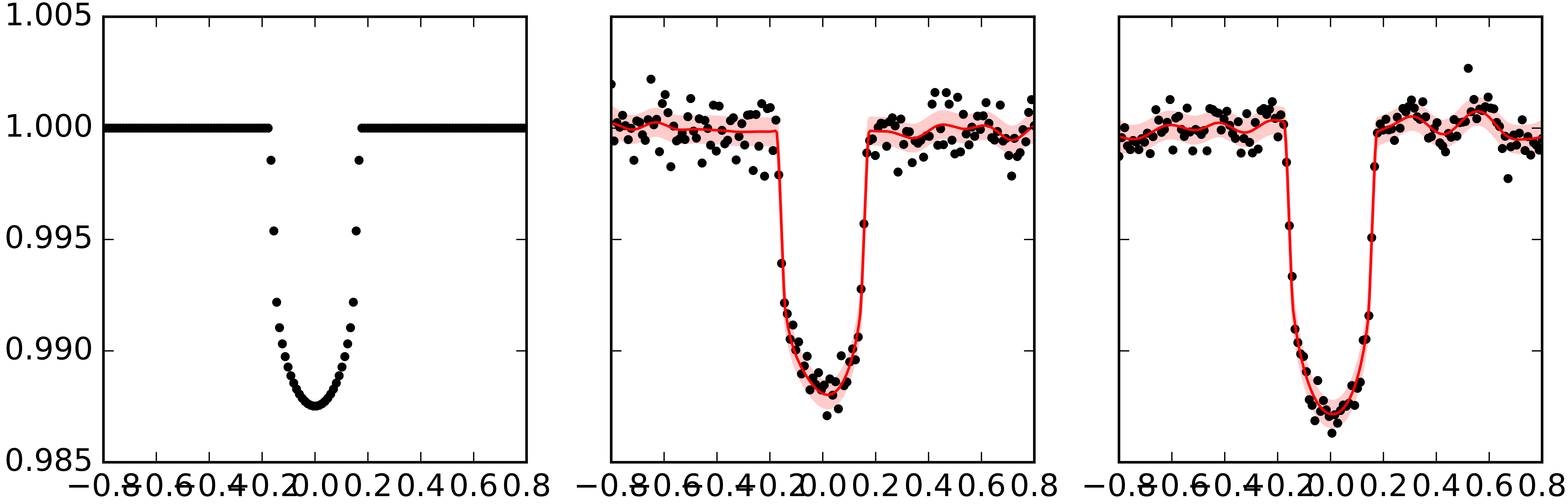
1.,	0.,	0.,	0.,	0.
0.,	1.,	0.,	0.,	0.
0.,	0.,	1.,	0.,	0.
0.,	0.,	0.,	1.,	0.
0.,	0.,	0.,	0.,	1.



1. ,	0.88,	0.61,	0.32,	0.14
0.88,	1. ,	0.88,	0.61,	0.32
0.61,	0.88,	1. ,	0.88,	0.61
0.32,	0.61,	0.88,	1. ,	0.88
0.14,	0.32,	0.61,	0.88,	1.

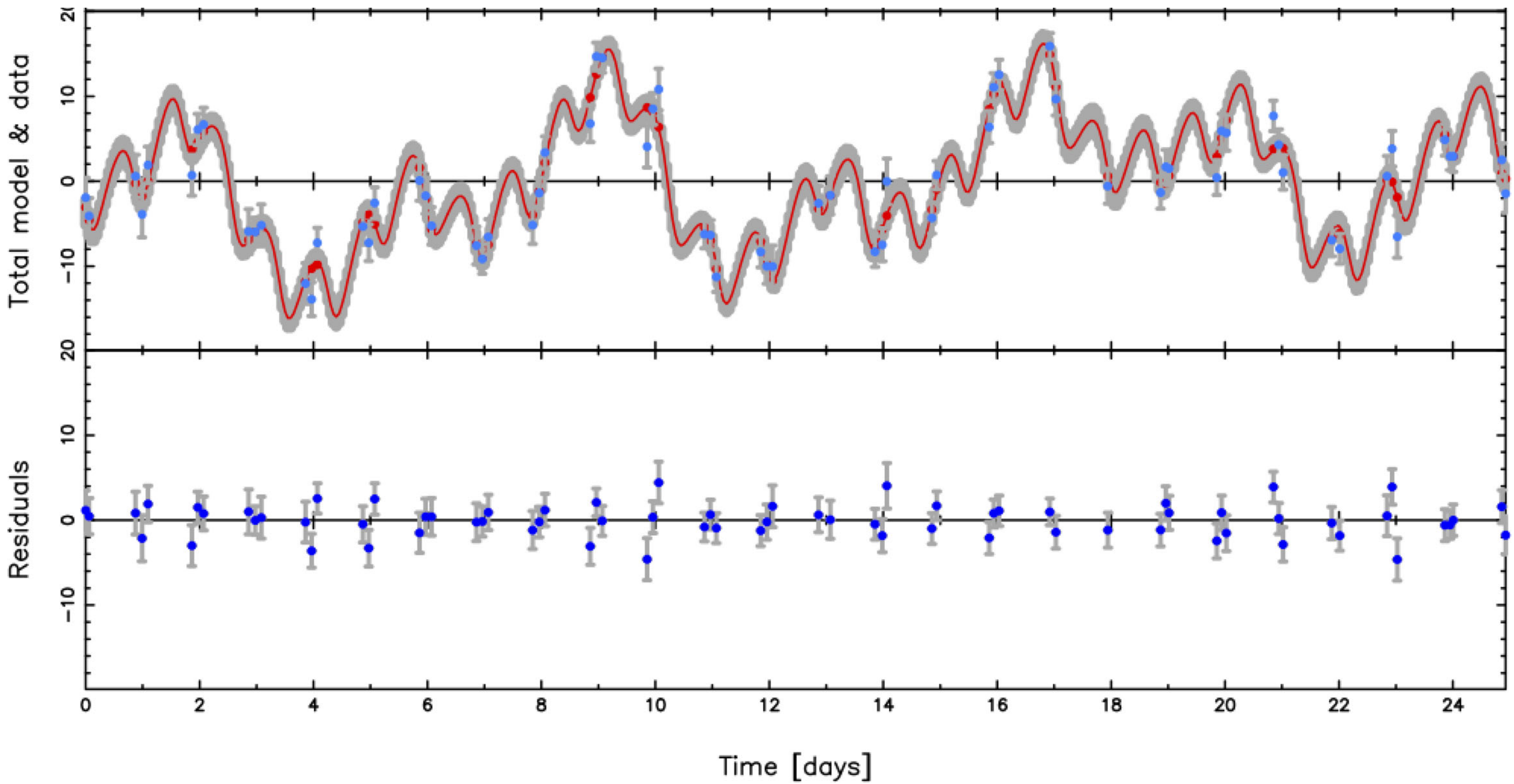



Diagonal Kernel
3
2
1
0
−1
−2
−3
1
2
3
4
5
Squared Exponential Kernel (A=1, tau=2)
3
2
1
0
−1
−2
−3
1
2
3
4
5


1.005
1.000
0.995
0.990
0.985
−0.8 −0.6 −0.4 −0.2 0.0 0.2 0.4 0.6 0.8
−0.8 −0.6 −0.4 −0.2 0.0 0.2 0.4 0.6 0.8
−0.8 −0.6 −0.4 −0.2 0.0 0.2 0.4 0.6 0.8

Using GP regression for activity

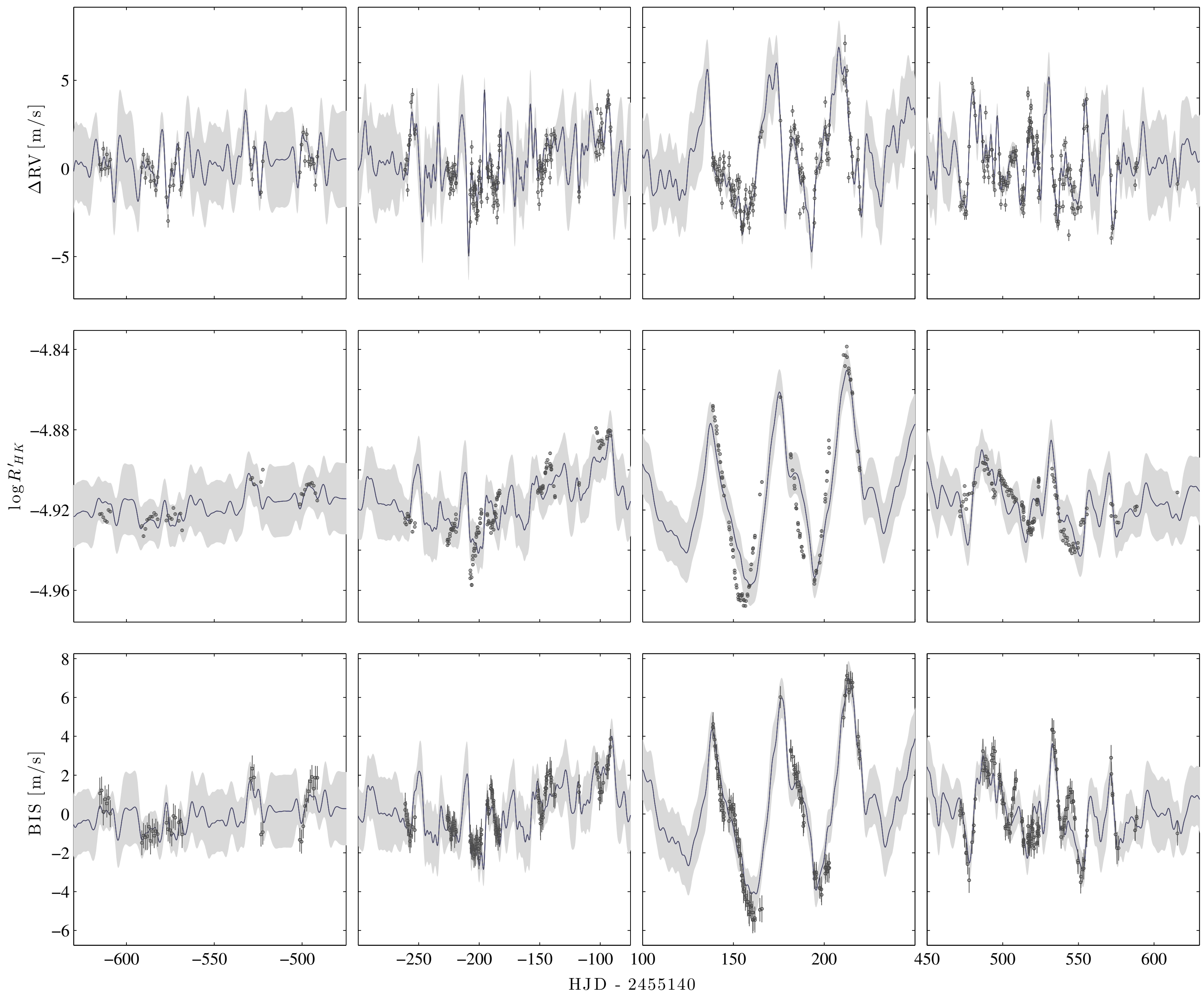
CoRoT-7



Haywood et al. (2014)

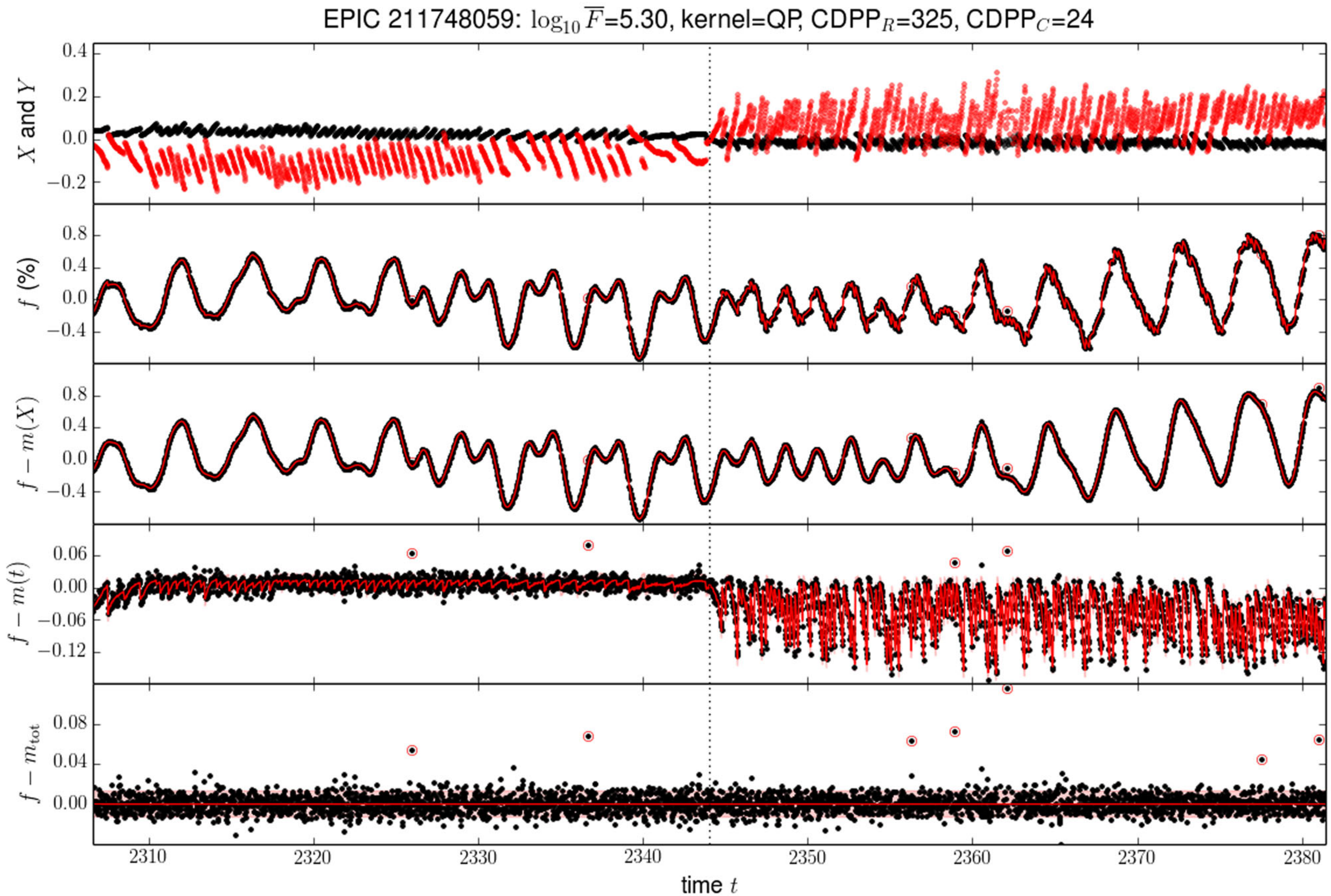
Using GP regression for activity

Alpha Cen B



Rajpaul et al. (2015)

GP regression to correct K2 LC



Aigrain et al. (2016)

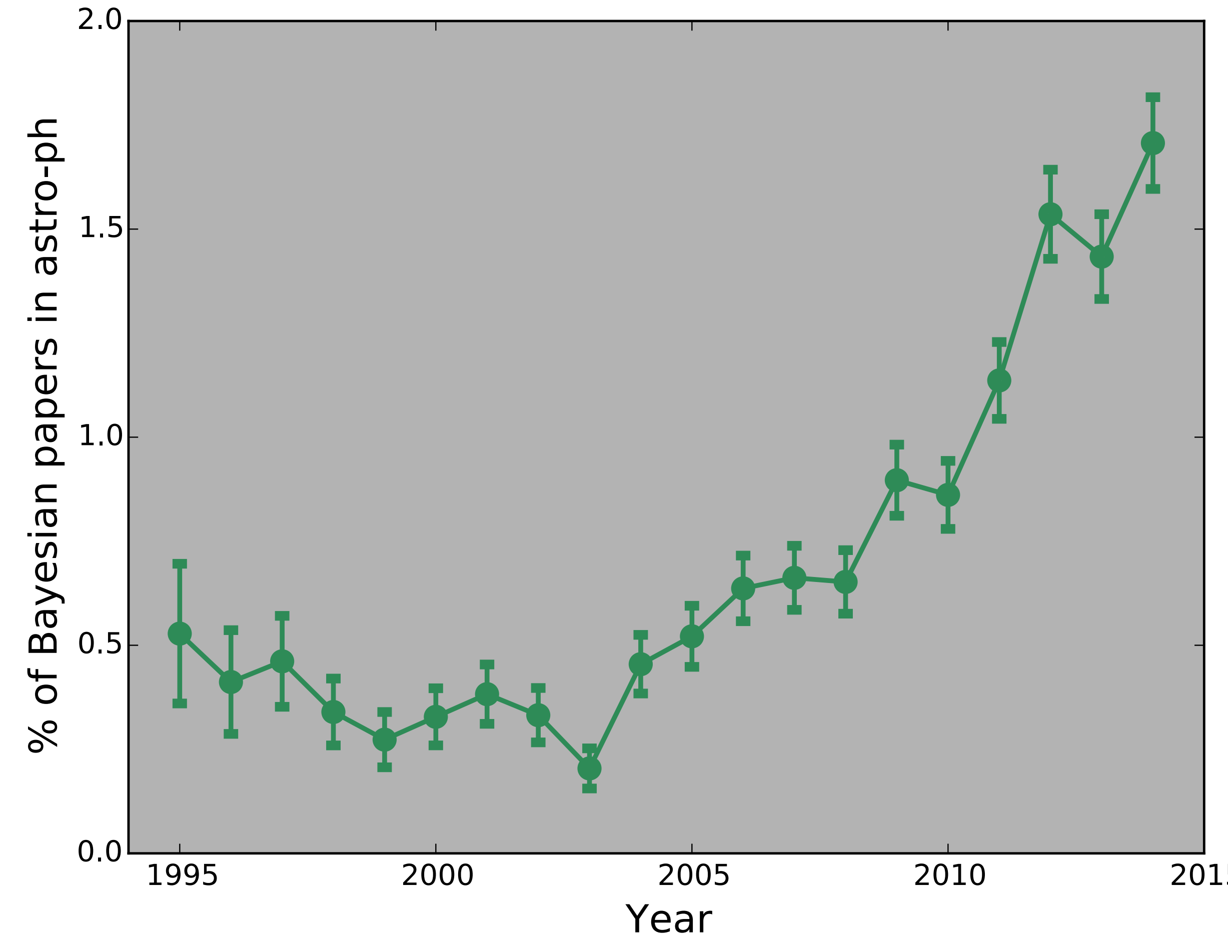
Epilogue

Recap' and conclusions

- Data is silent without a model. It can be treacherous with the wrong one.
- "All models are wrong, but some are useful."
- Physical models and error models are two faces of the same coin.
- As data precision improves, so must our ability to model and analyse them (thank computers).
- There is no shame in using other people's code, but there's nothing like writing (and debugging) your own.

Recap' and conclusions

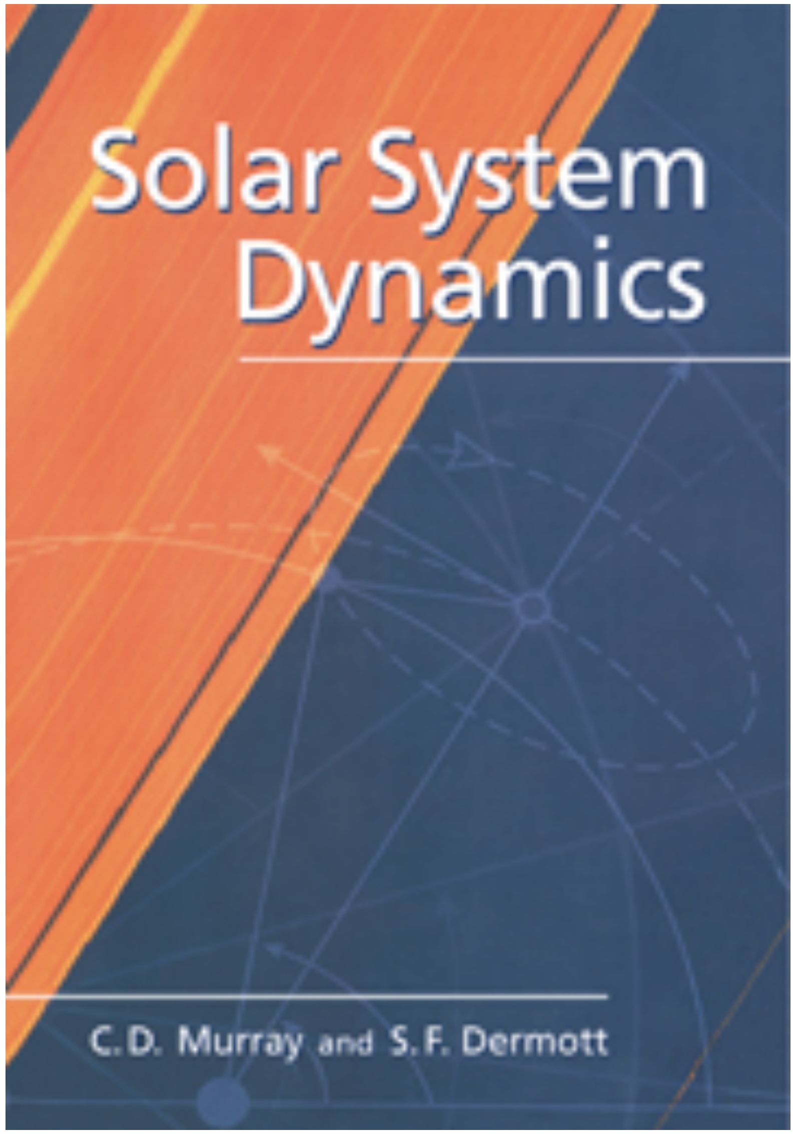
- The Bayesian approach is becoming the industry standard.
- Powerful and flexible way of thinking about all things around data.
- "If you're doing chi2 minimization, you could be doing something better."



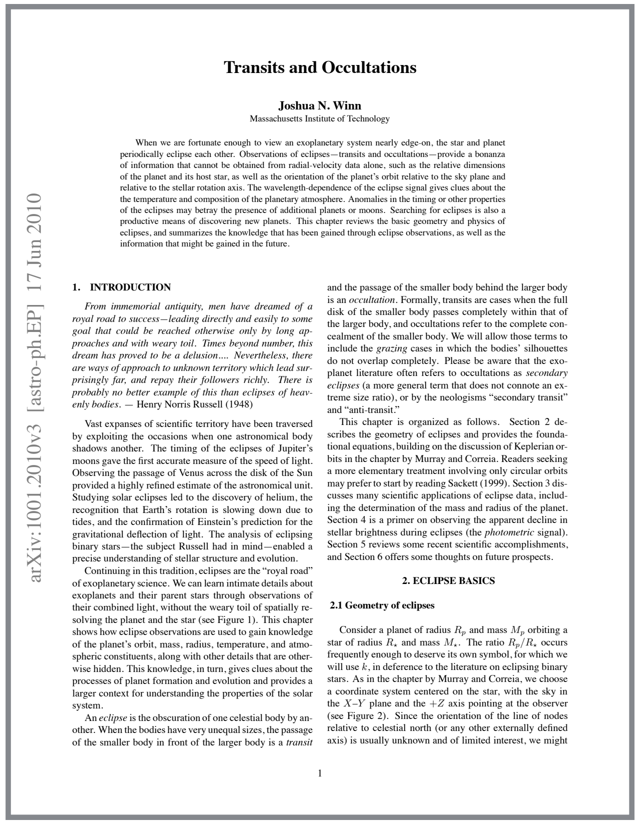
Recap’ and conclusions

- Stay tuned for:
  - effects of stellar activity.
  - planetary atmospheres.
  - more!

Main references

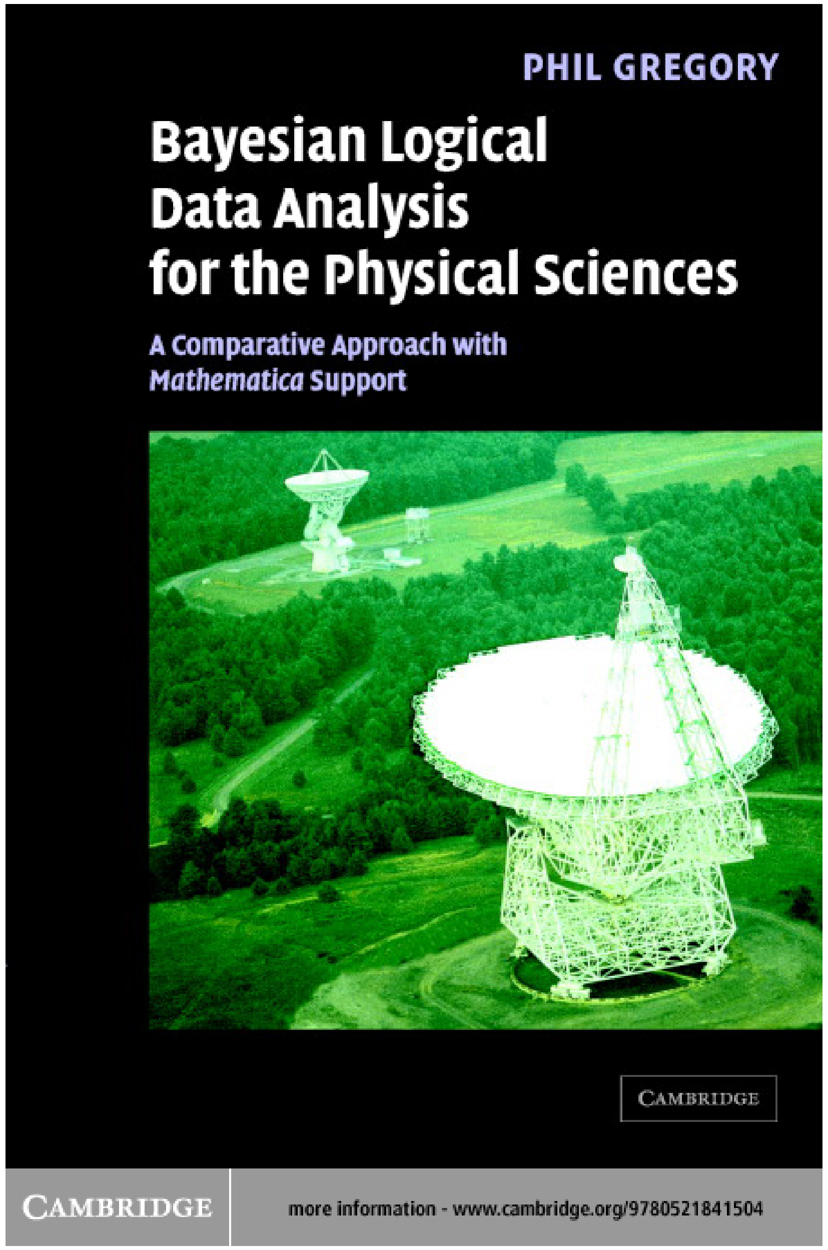


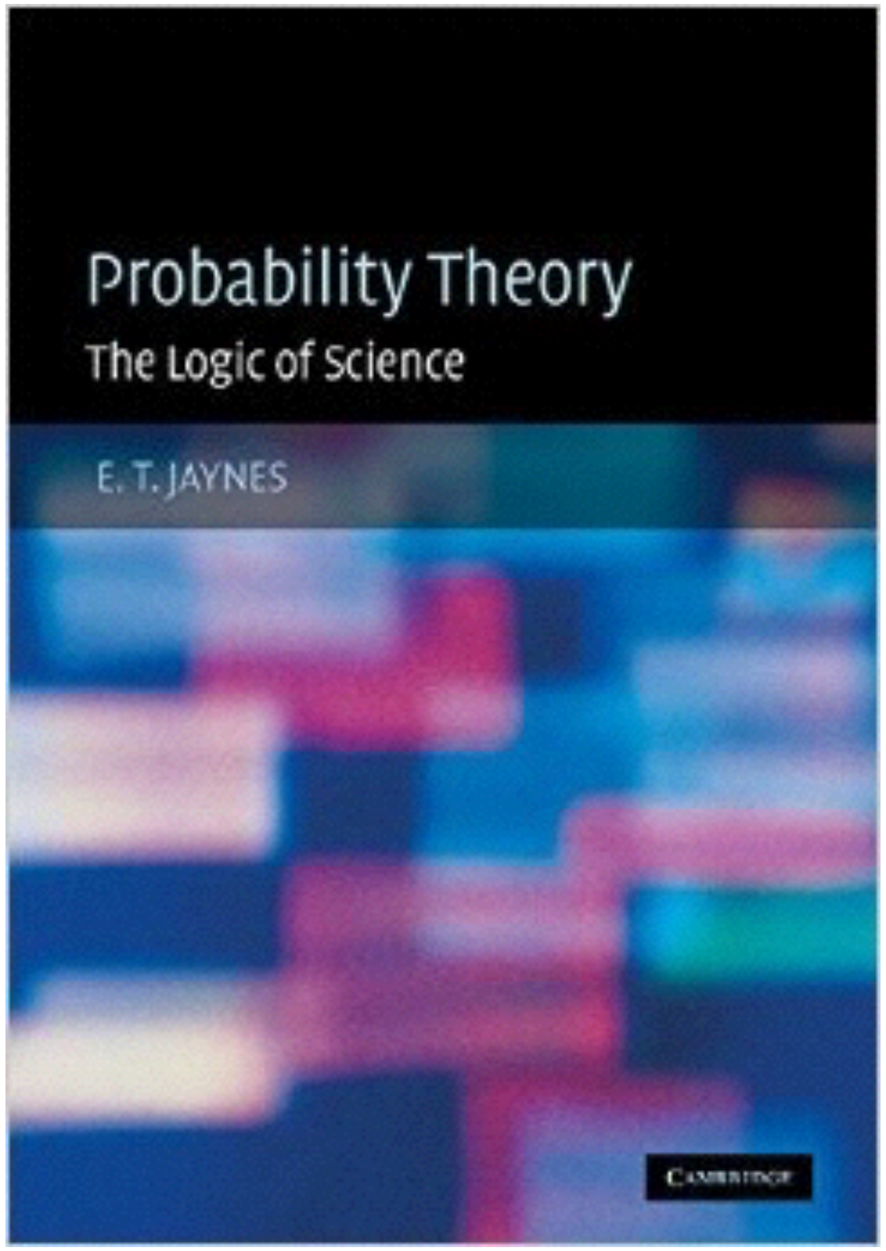
Seager & Mallén-Ornelas (2003)



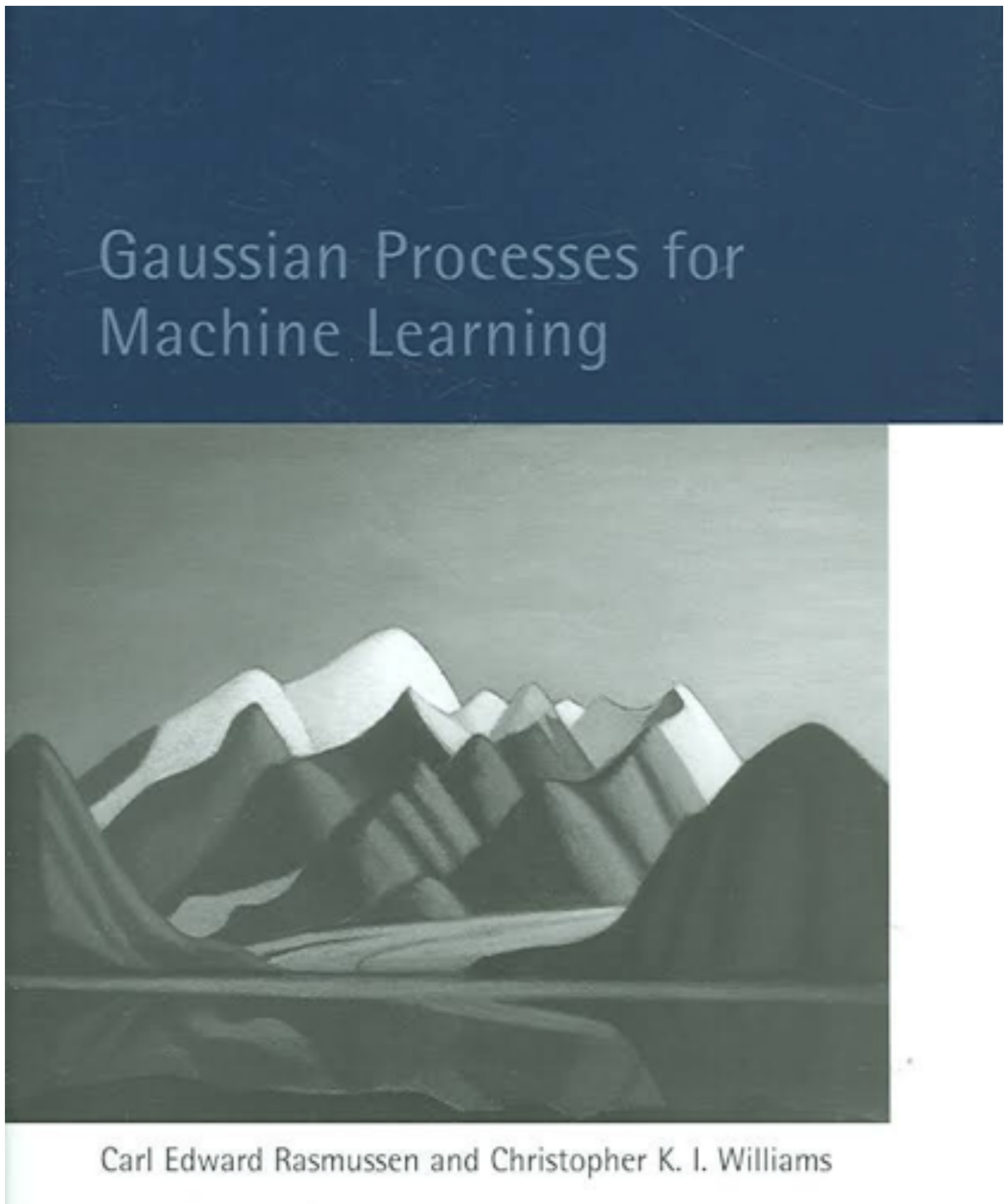
Winn (2010) arXiv:1001.2010

Main references - Bayesian





Main references - Gaussian processes



Available online

http://www.gaussianprocess.org/gpml/

Code

Bayesian evidence computation

`github.com/exord/bayev`

Gaussian process regression

`github.com/exord/gp`

`pypi.python.org/pypi/pygpr`

Spot LC simulator

`github.com/exord/lcspotter`

`pypi.python.org/pypi/lcspotter`