Gaussian processes: the next step in exoplanet data analysis

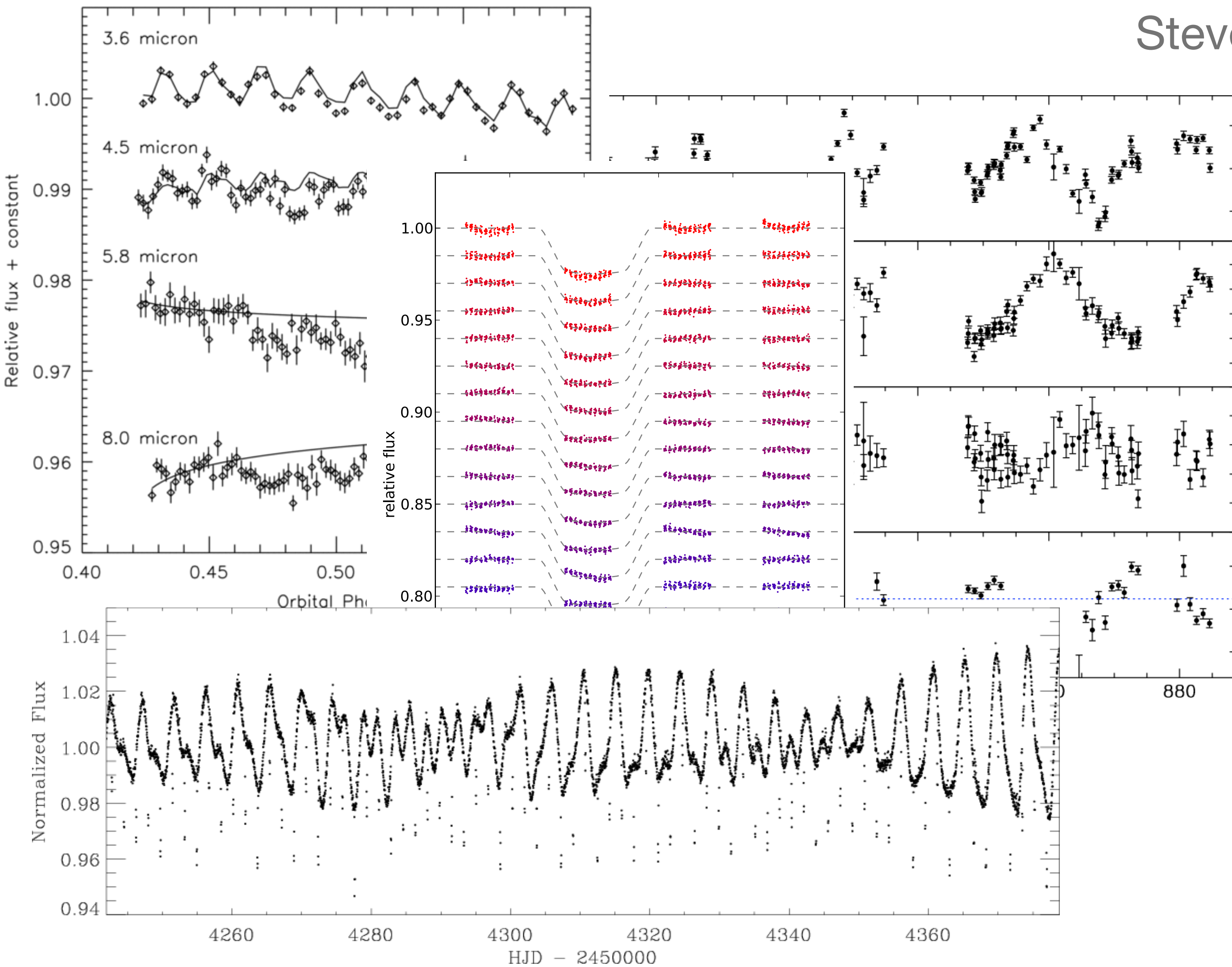
Suzanne Aigrain (University of Oxford)
Neale Gibson, Tom Evans, Amy McQuillan
Steve Roberts, Steve Reece, Mike Osborne

... let the data speak

University of Oxford

Gaussian processes: the next step in exoplanet data analysis

Suzanne Aigrain (University of Oxford)
Neale Gibson, Tom Evans, Amy McQuillan
Steve Roberts, Steve Reece, Mike Osborne



stochastic
high precision
time-series
correlated
disentangle
noise

... let the data speak

UNIVERSITY OF OXFORD

A Gaussian process in a nutshell

$$P\left(\boldsymbol{t} \,|\, \mathbf{X}, \boldsymbol{\theta}, \phi\right) = \mathcal{N}\left[m(\mathbf{X}, \phi), \mathbf{K}\right]$$

A Gaussian process in a nutshell

output vector (e.g. fluxes)   input matrix (e.g. time, pointing, detector temperature...)

likelihood

$$P\left(\boldsymbol{t} \,|\, \mathbf{X}, \boldsymbol{\theta}, \phi\right) = \mathcal{N}\left[m(\mathbf{X}, \phi), \mathbf{K}\right]$$

(hyper-)parameters

A Gaussian process in a nutshell

output vector (e.g. fluxes)

input matrix (e.g. time, pointing, detector temperature...)

likelihood

$$P\left(\boldsymbol{t} \,|\, \mathbf{X}, \boldsymbol{\theta}, \phi\right) = \mathcal{N}\left[m(\mathbf{X}, \phi), \mathbf{K}\right]$$

multivariate Gaussian

(hyper-)parameters

mean function

covariance matrix

A Gaussian process in a nutshell

output vector (e.g. fluxes)   input matrix (e.g. time, pointing, detector temperature...)

likelihood
$$P\left(\boldsymbol{t} \,|\, \mathbf{X}, \boldsymbol{\theta}, \phi\right) = \mathcal{N}\left[m(\mathbf{X}, \phi), \mathbf{K}\right]$$
multivariate Gaussian

(hyper-)parameters

mean function

covariance matrix

$$K_{nm} \equiv \mathrm{cov}\left[\boldsymbol{x}_n, \boldsymbol{x}_m\right] = k(\boldsymbol{x}_n, \boldsymbol{x}_m, \boldsymbol{\theta})$$

kernel function

A Gaussian process in a nutshell

output vector (e.g. fluxes)   input matrix (e.g. time, pointing, detector temperature...)

likelihood

$$P\left(\boldsymbol{t} \,|\, \mathbf{X}, \boldsymbol{\theta}, \phi\right) = \mathcal{N}\left[m(\mathbf{X}, \phi), \mathbf{K}\right]$$

multivariate Gaussian

mean function   mean function parameters

(hyper-)parameters

covariance matrix

$$K_{nm} \equiv \mathrm{cov}\left[\boldsymbol{x}_n, \boldsymbol{x}_m\right] = k(\boldsymbol{x}_n, \boldsymbol{x}_m, \boldsymbol{\theta})$$

kernel function   kernel parameters

Gaussian process regression

Gaussian process regression

a priori knowledge

Gaussian process regression

a priori knowledge

- chose kernel & mean functions
- guess initial hyper-parameters

Gaussian process regression

a priori knowledge → chose kernel & mean functions → prior probability distribution over functions

a priori knowledge → guess initial hyper-parameters → prior probability distribution over functions

Gaussian process regression

a priori knowledge

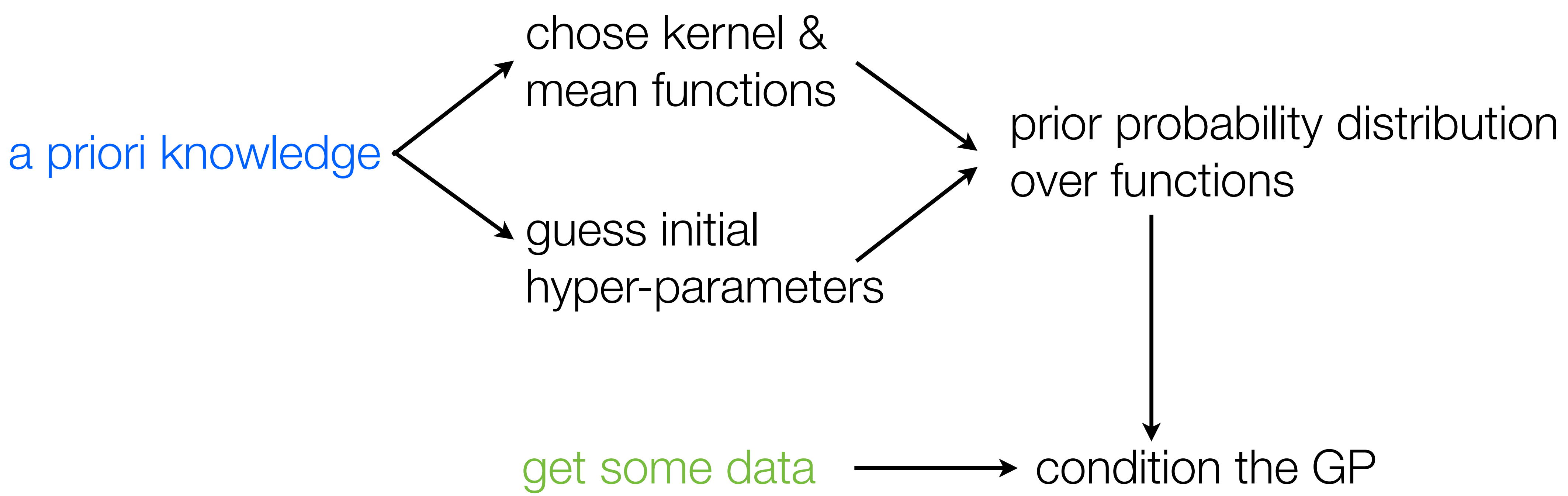
chose kernel & mean functions

guess initial hyper-parameters

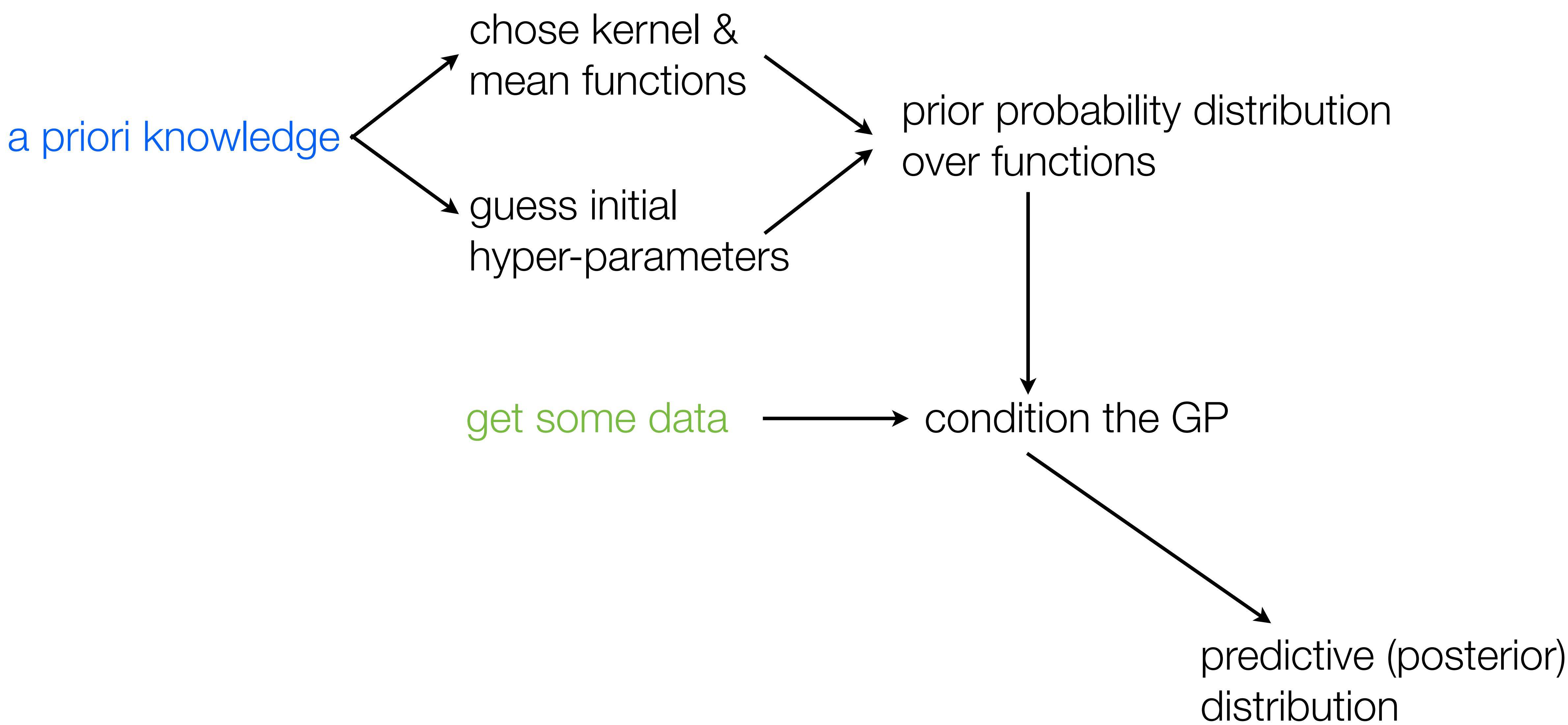
prior probability distribution over functions

get some data

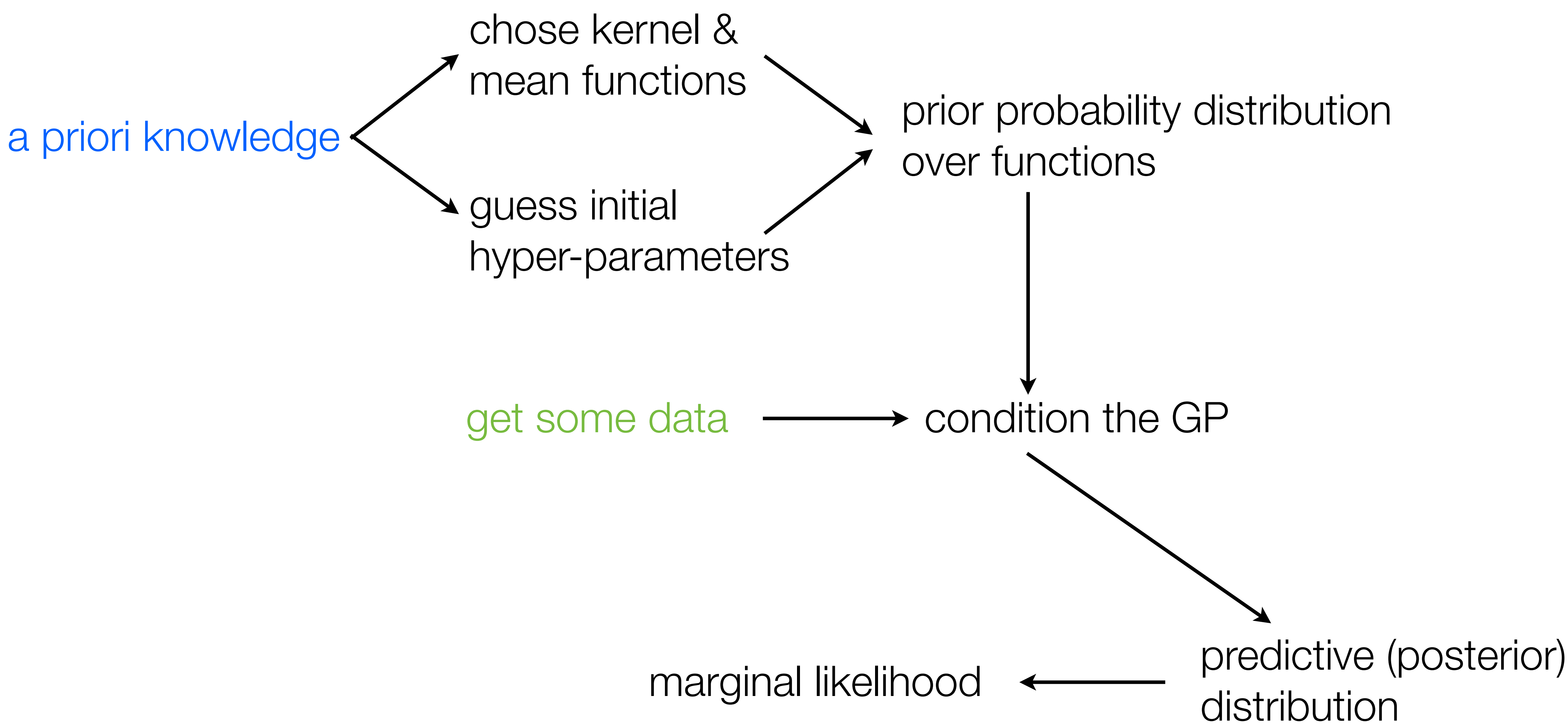
Gaussian process regression



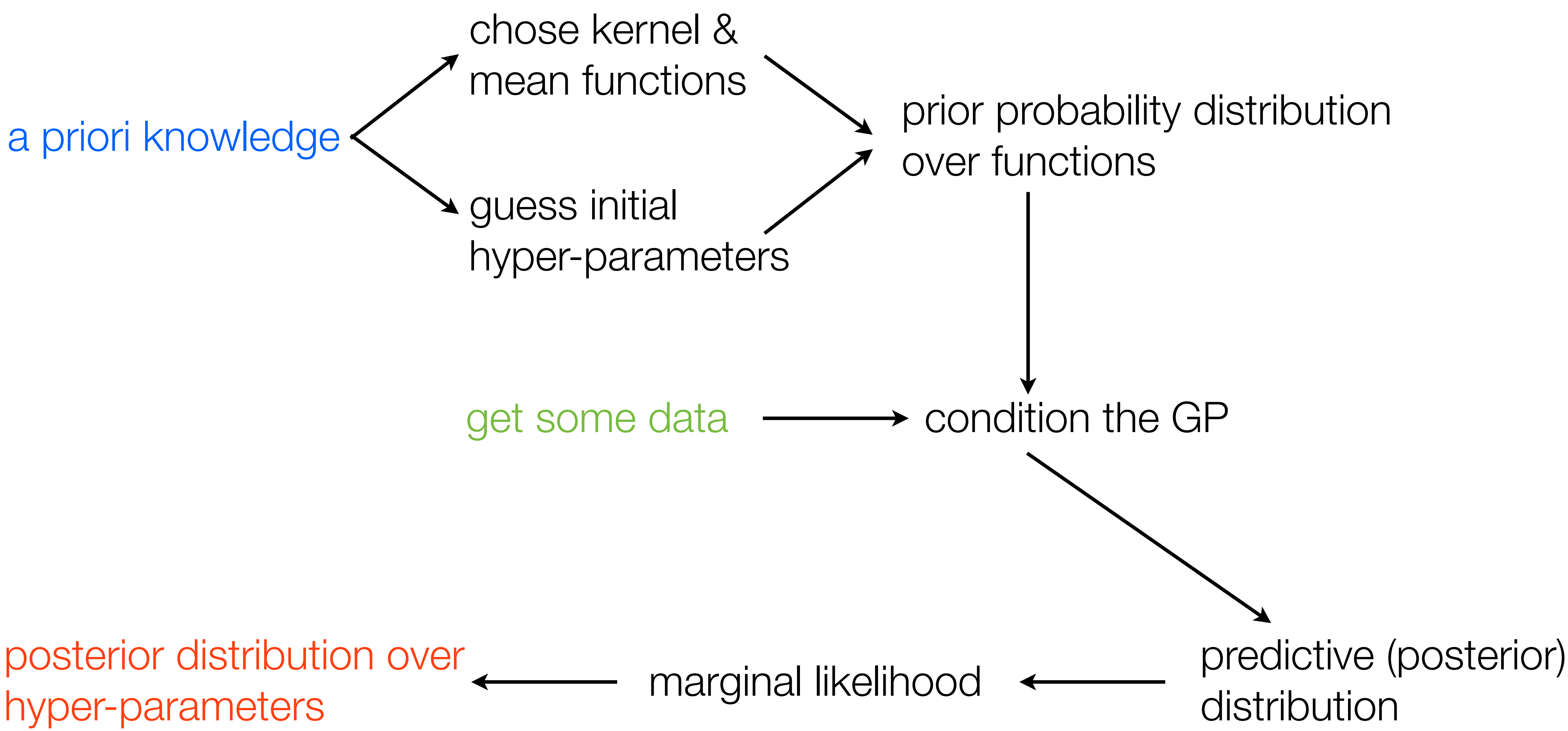
Gaussian process regression



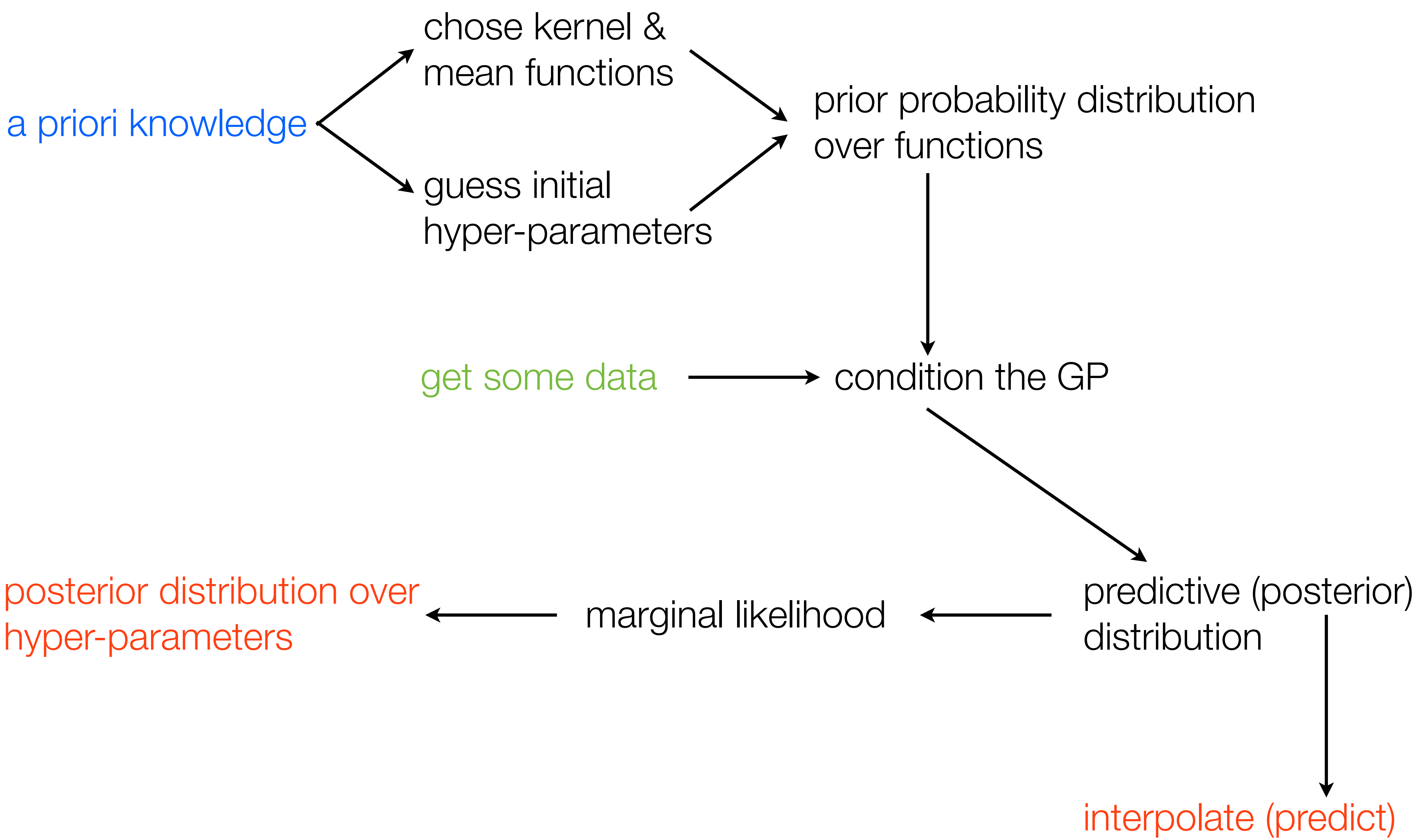
Gaussian process regression



Gaussian process regression

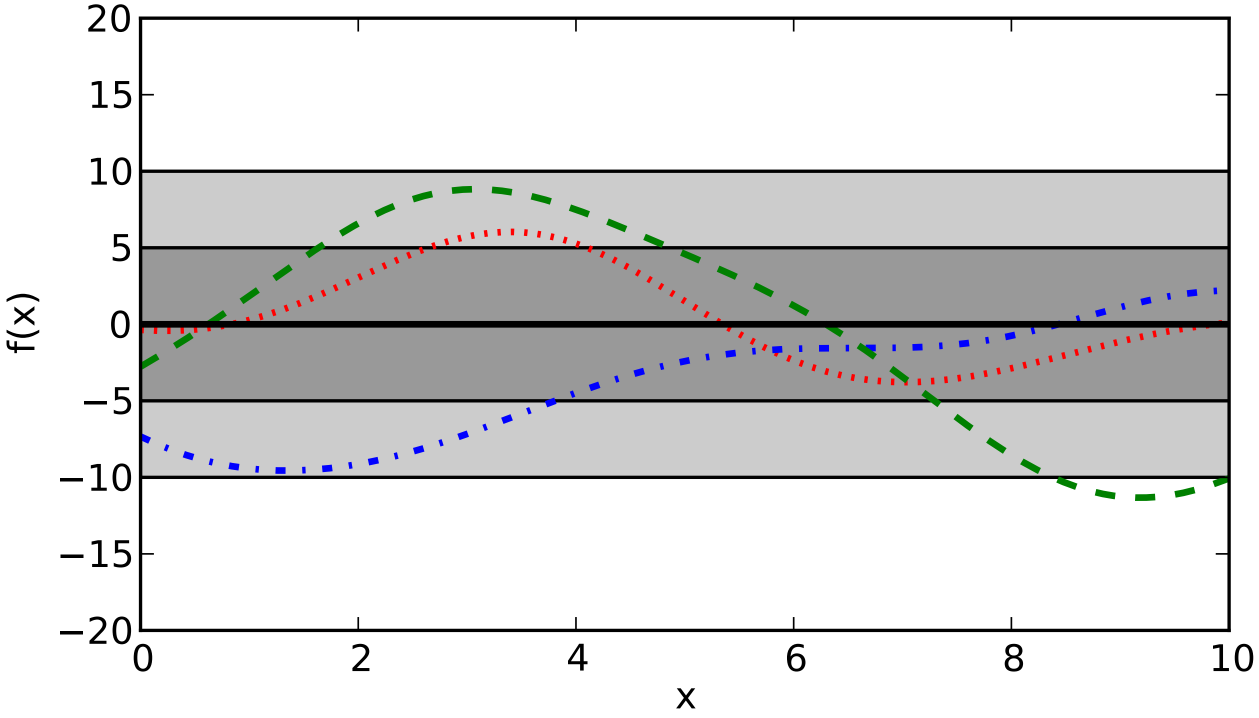


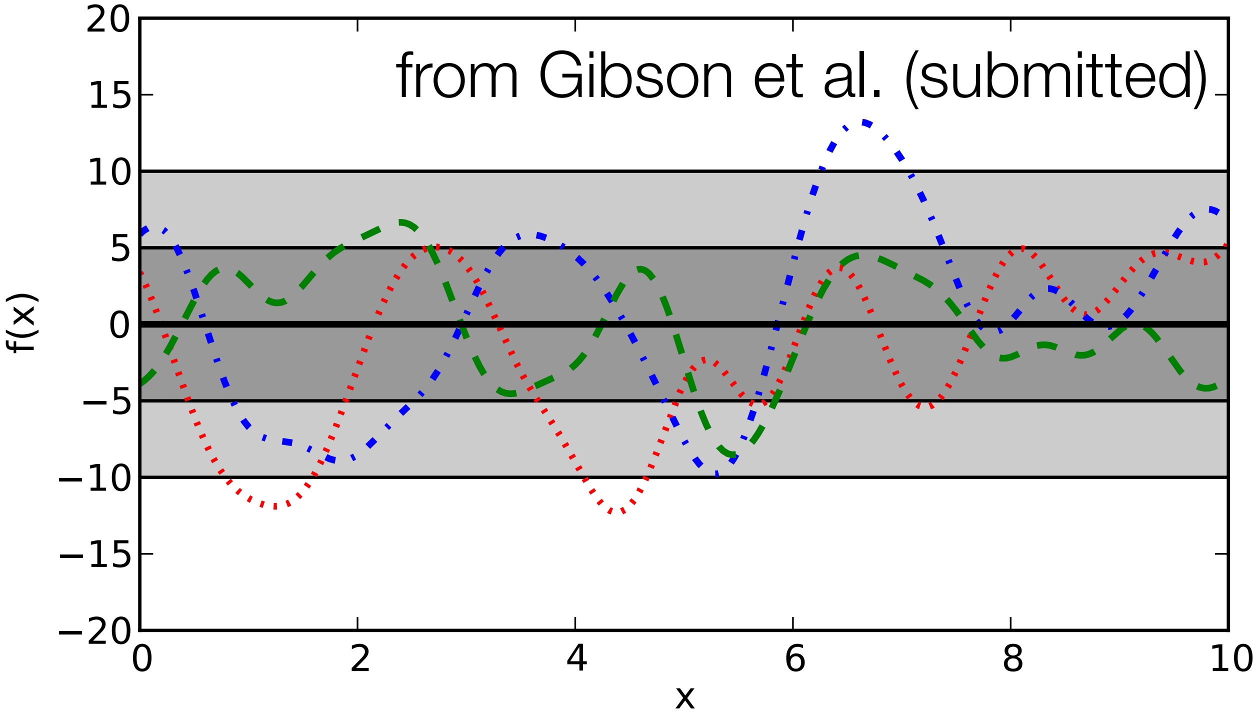
Gaussian process regression



A very simple example

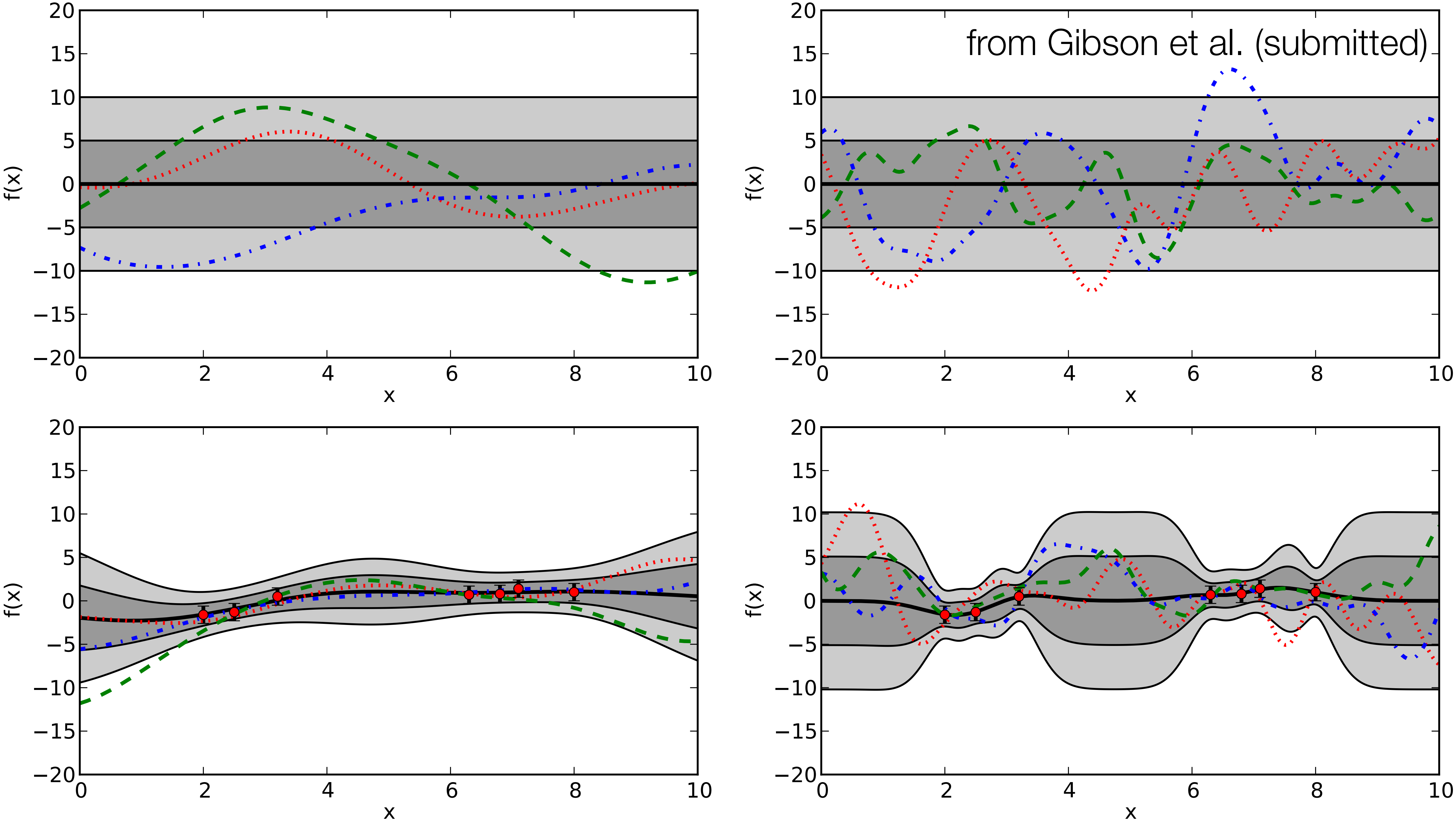
$$k_{\mathrm{SE}}(t, t') = A^2 \exp\left(-\frac{(t-t')^2}{2l^2}\right)$$





A very simple example

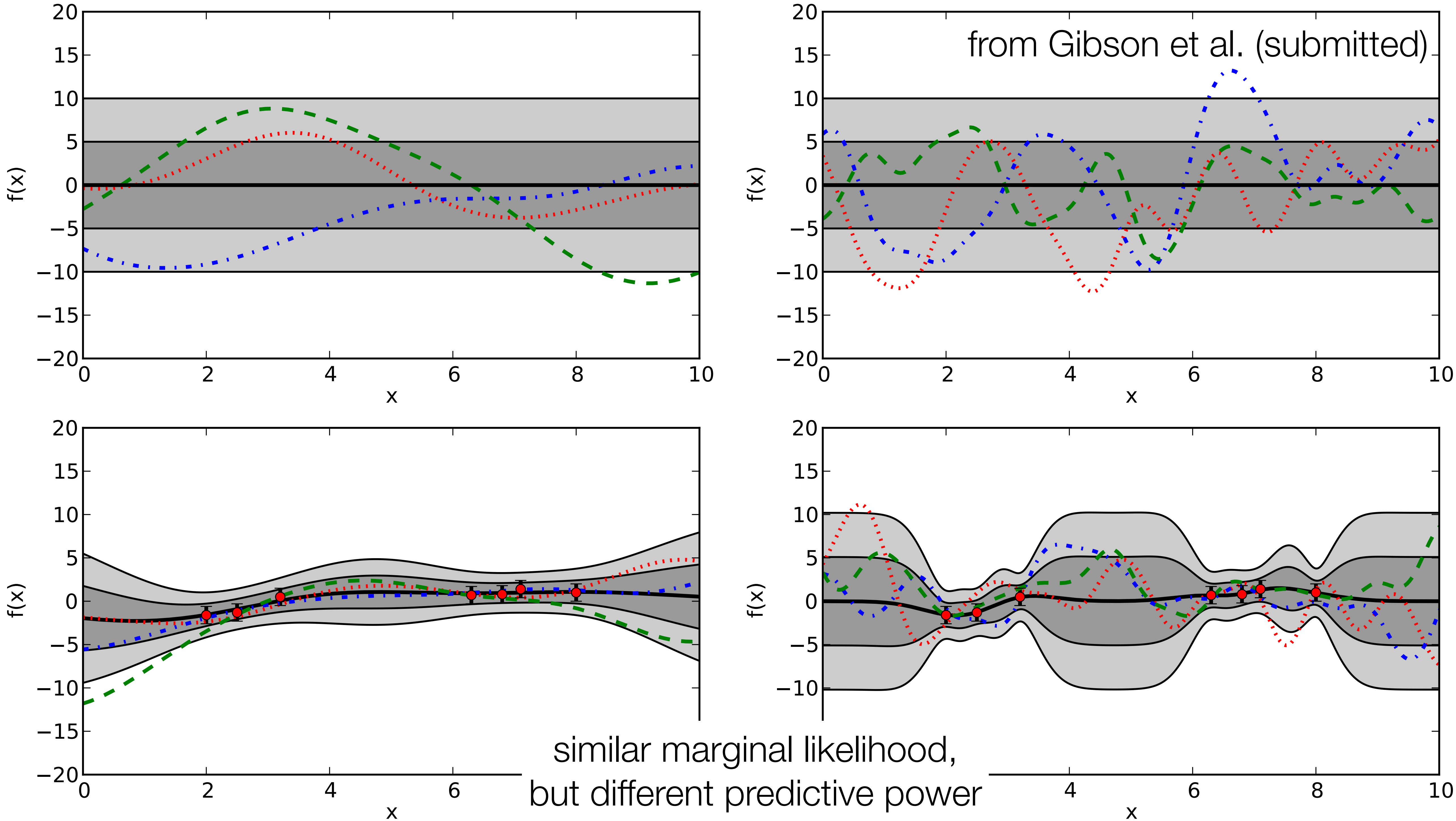
$$k_{\mathrm{SE}}(t,t') = A^2 \exp\left(-\frac{(t-t')^2}{2l^2}\right)$$



A very simple example

$$k_{\mathrm{SE}}(t, t') = A^2 \exp\left(-\frac{(t-t')^2}{2l^2}\right)$$

from Gibson et al. (submitted)



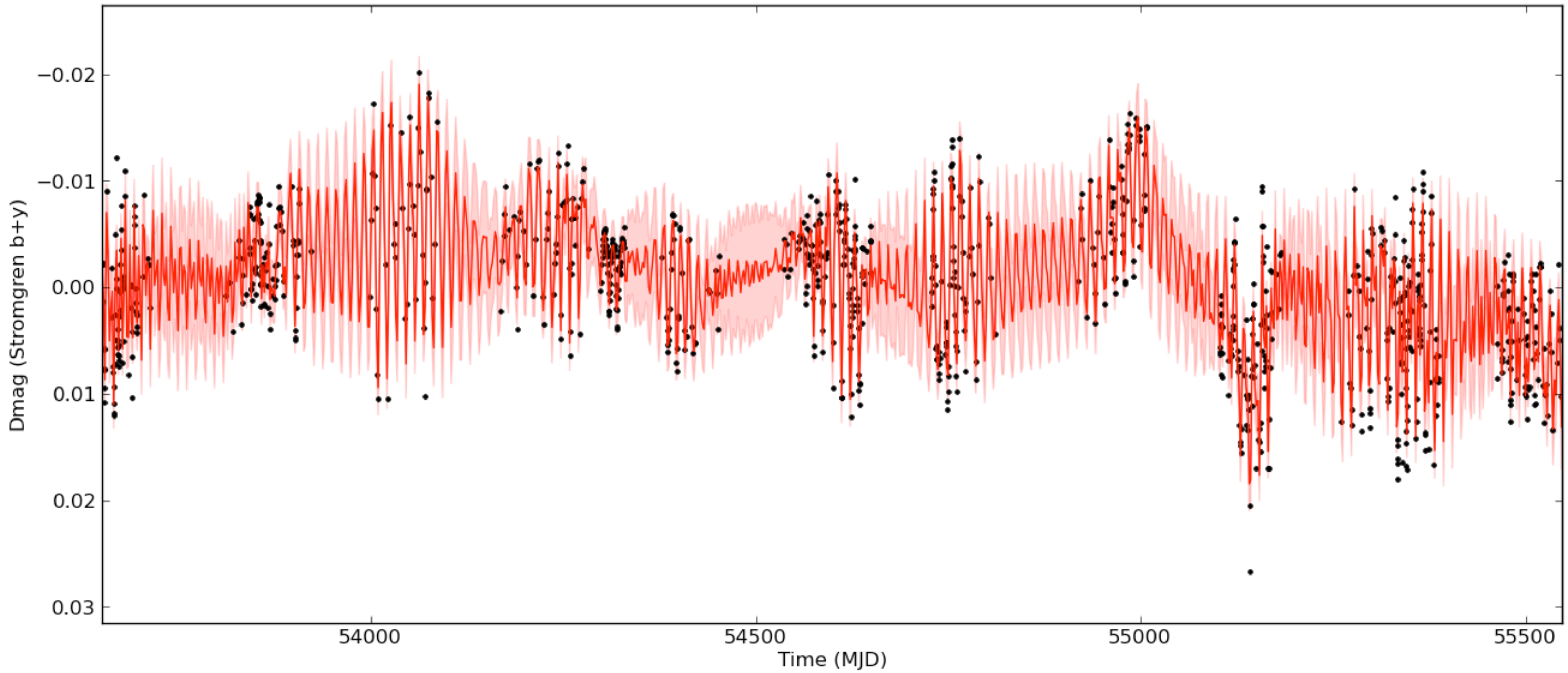
similar marginal likelihood,
but different predictive power

Example application 1:
instrumental systematics in transmission spectra

See Neale Gibson's talk

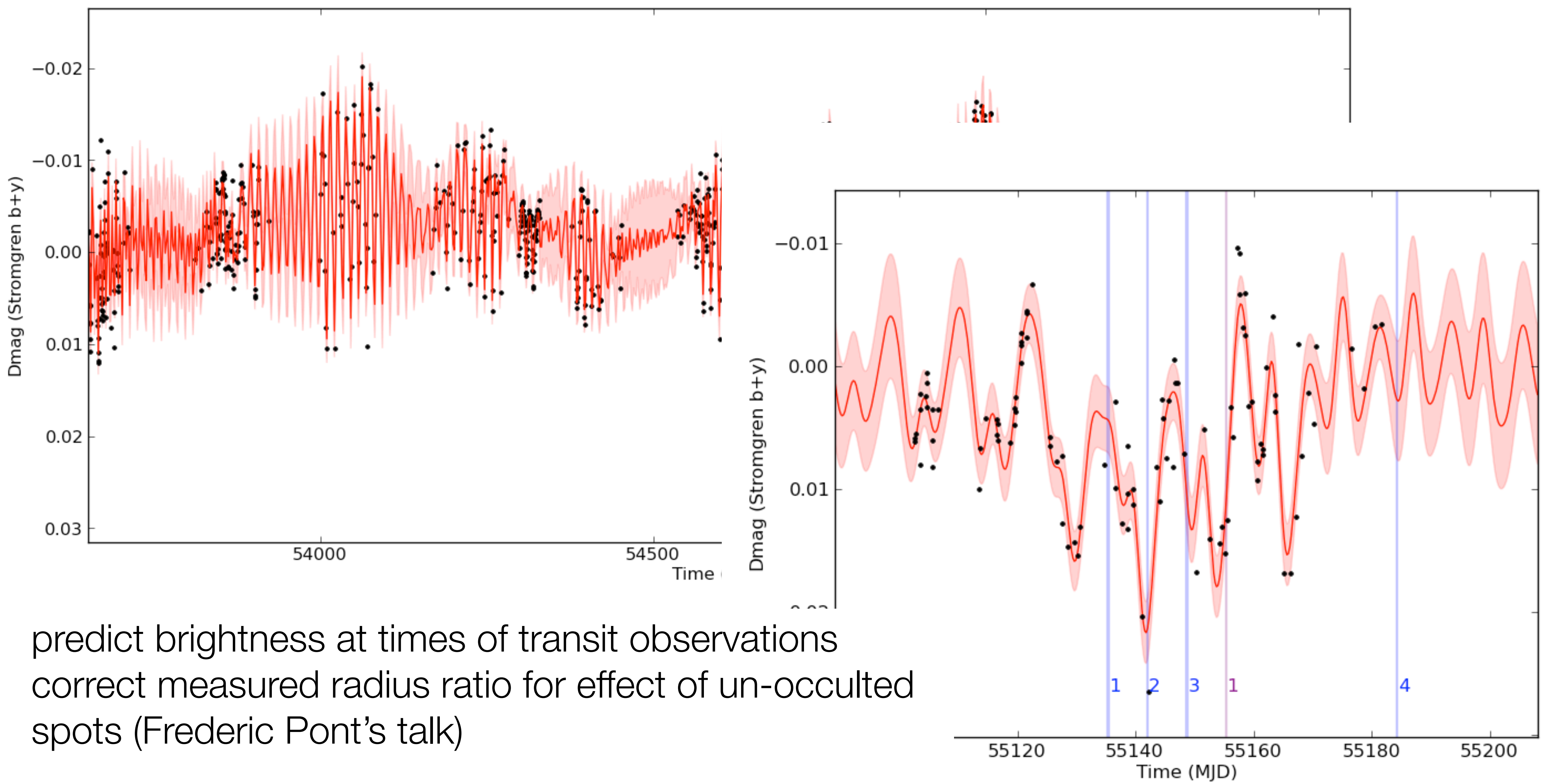
Example application 2: modelling HD 189733b's OOT light curve

$$k_{\mathrm{QP,mixed}}(t,t') = A^2 \exp\left(-\frac{\sin^2[\pi(t-t')/P]}{2L^2}\right) \times \left(1 + \frac{(t-t')^2}{2\alpha l^2}\right)^{-\alpha} + \sigma^2 \mathbf{I}.$$



Example application 2:
modelling HD 189733b's OOT light curve

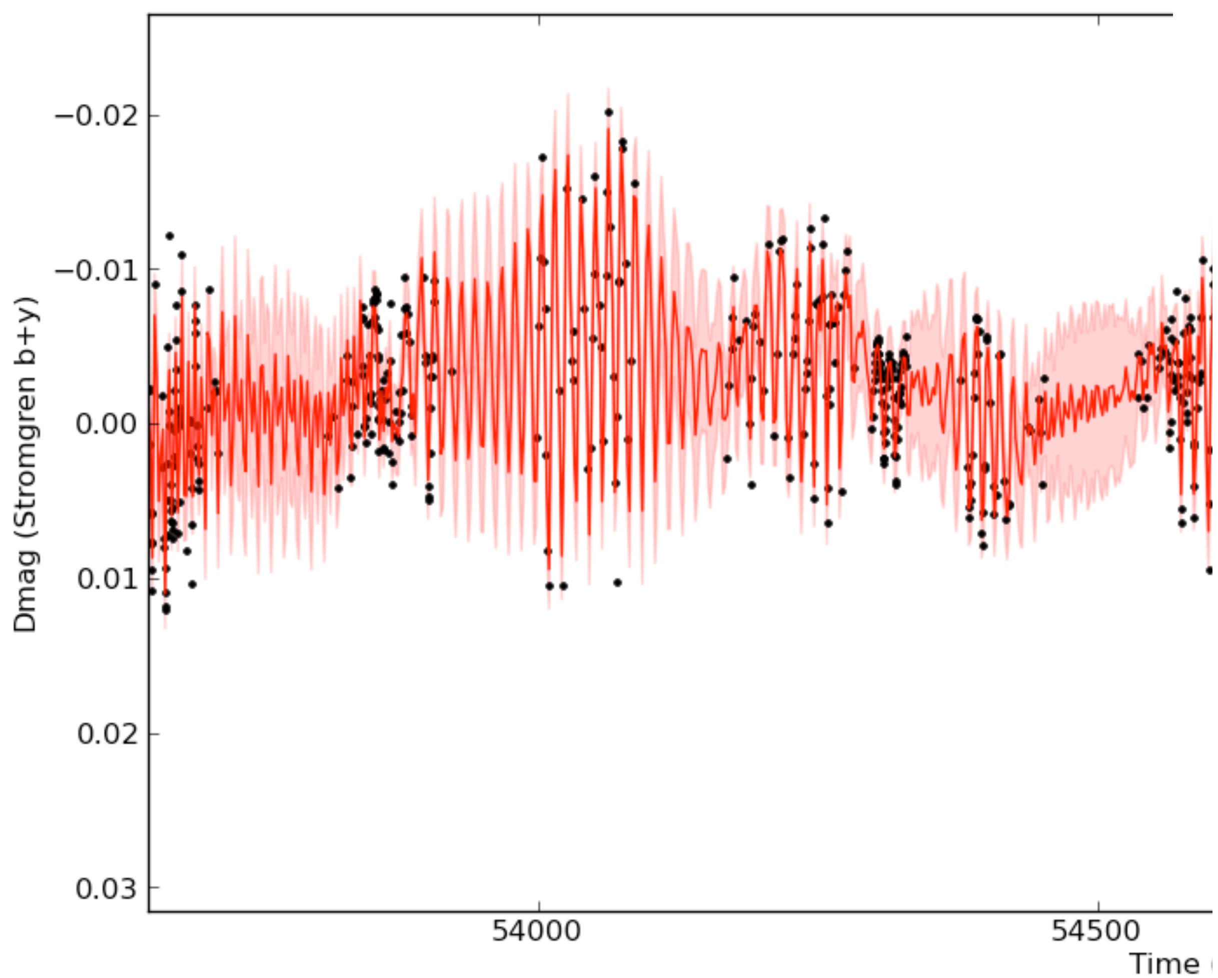
$$k_{\mathrm{QP,mixed}}(t,t') = A^2 \exp\left(-\frac{\sin^2[\pi(t-t')/P]}{2L^2}\right) \times \left(1 + \frac{(t-t')^2}{2\alpha l^2}\right)^{-\alpha} + \sigma^2 \mathbf{I}.$$



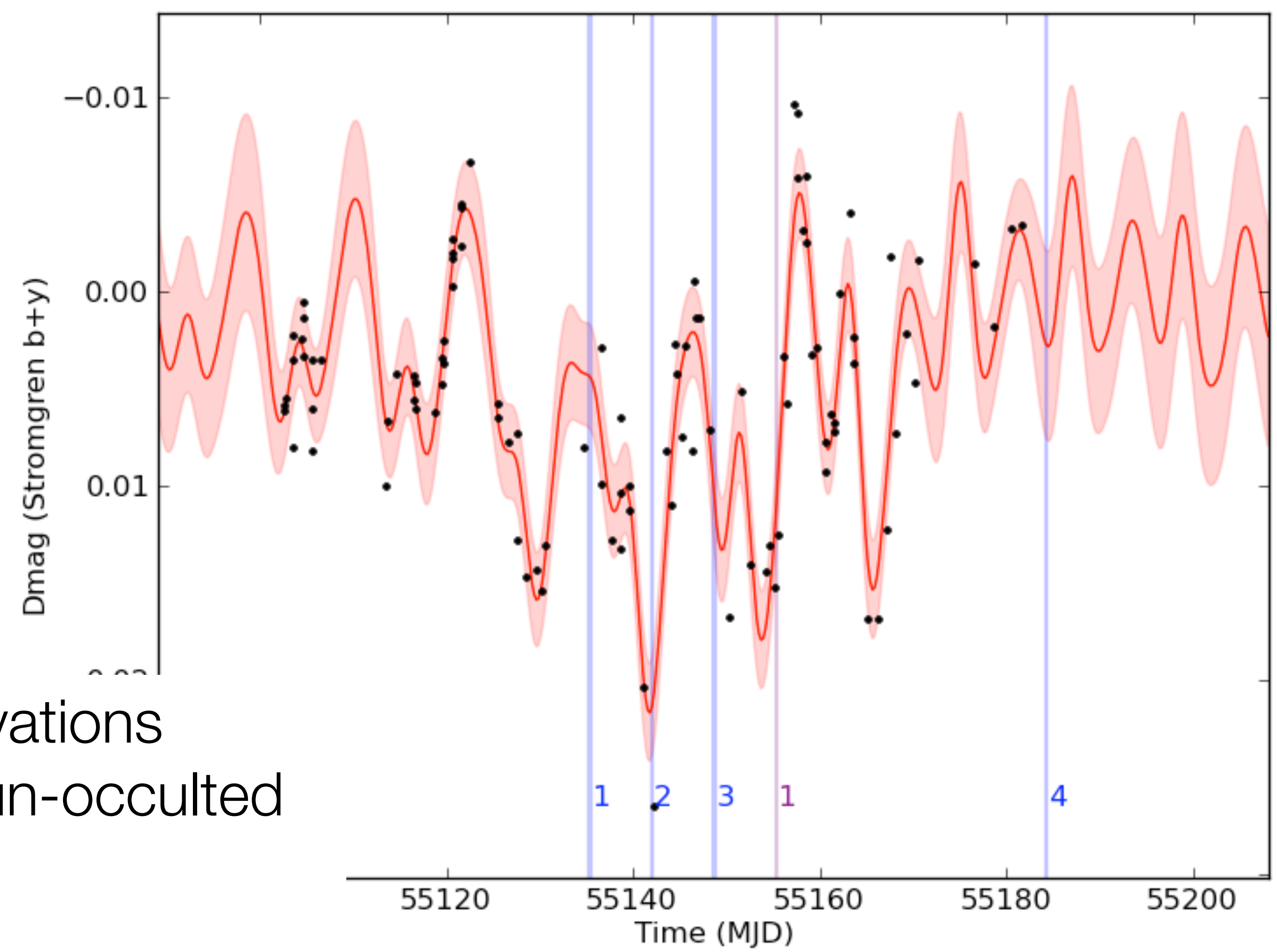
predict brightness at times of transit observations
correct measured radius ratio for effect of un-occulted spots (Frederic Pont's talk)

Example application 2: modelling HD 189733b's OOT light curve

$$k_{\mathrm{QP,mixed}}(t,t') = A^2 \exp\left(-\frac{\sin^2[\pi(t-t')/P]}{2L^2}\right) \times \left(1 + \frac{(t-t')^2}{2\alpha l^2}\right)^{-\alpha} + \sigma^2 \mathbf{I}.$$



hyper-parameters contain information about rotation rate, spot distribution and evolution...



predict brightness at times of transit observations

correct measured radius ratio for effect of un-occulted spots (Frederic Pont's talk)

Pros ... ... and cons

- Rigorous error propagation
- Extremely versatile
- Built-in Ockam's razor
- Joint modelling of arbitrary number of inputs (and outputs)
- Easy to combine with other techniques e.g. MCMC

- Computationally intensive: $O(N^3)$
  - ok up to N~1000
  - alternative: Variational Bayes (see Tom Evans' poster)

Pros ...

- Rigorous error propagation
- Extremely versatile
- Built-in Ockam's razor
- Joint modelling of arbitrary number of inputs (and outputs)
- Easy to combine with other techniques e.g. MCMC

... and cons

- Computationally intensive: $O(N^3)$
  - ok up to N~1000
  - alternative: Variational Bayes (see Tom Evans' poster)

Want to try?
Python GP module under development