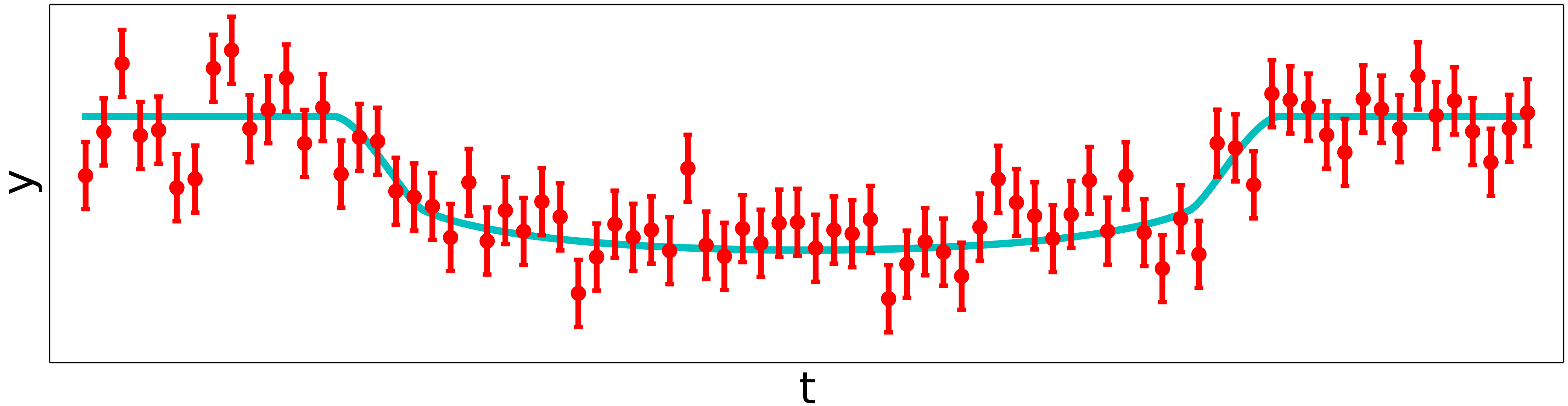
Gaussian process treatment of IRAC lightcurves

Tom Evans
University of Exeter, UK

Tom Evans (Exeter)

What is a Gaussian process (GP)?

A collection of data points, any subset of which has a multivariate normal distribution



i-th data point:

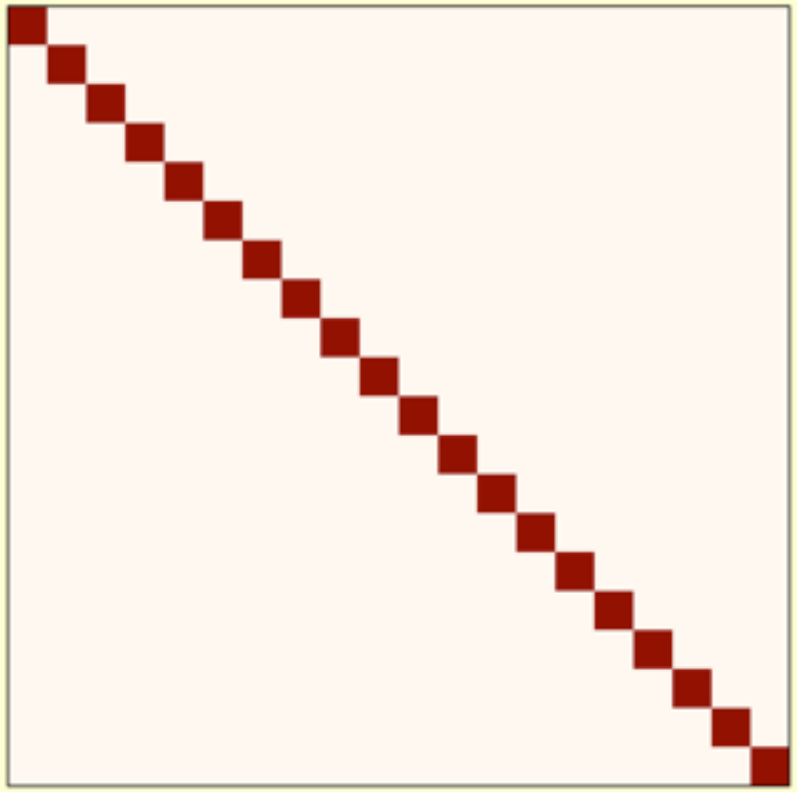
$$p(d_i|\mu_i,\sigma_i) = \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left[-\frac{(d_i-\mu_i)^2}{2\sigma_i^2}\right]$$

full dataset:

$$d = \{d_1,\ldots,d_n\} \quad \mu = \{\mu_1,\ldots,\mu_n\} \quad \sigma = \{\sigma_1,\ldots,\sigma_n\}$$

$$p(d\,|\,\mu,\sigma) = \prod_{i=1}^{N} p(d_i|\mu_i,\sigma_i) \;=\; \mathcal{N}(d\,|\,\mu,\Sigma) \quad \text{where } \Sigma = \text{diag}\left(\sigma^2\right)$$

Likelihood functions



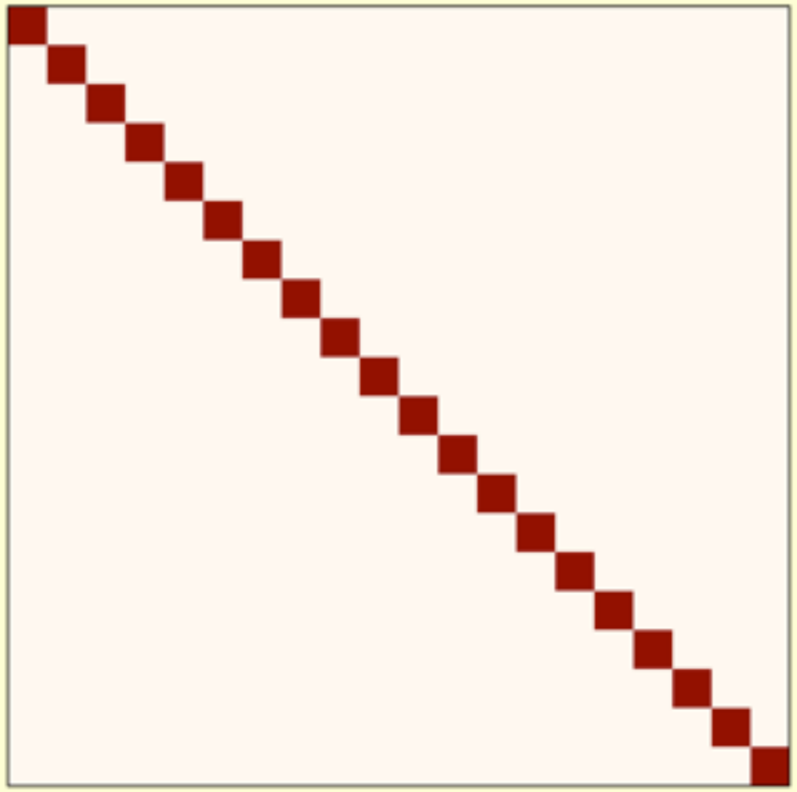
$= \Sigma$

'traditional' method

$$p(d\,|\,\mu,\Sigma) = \mathcal{N}(\mu,\Sigma)$$

$$\ln p(d\,|\,\mu,\Sigma) = -\frac{1}{2}\chi^2 - \frac{1}{2}\ln|\Sigma| - \frac{1}{2}\ln 2\pi$$

Likelihood functions

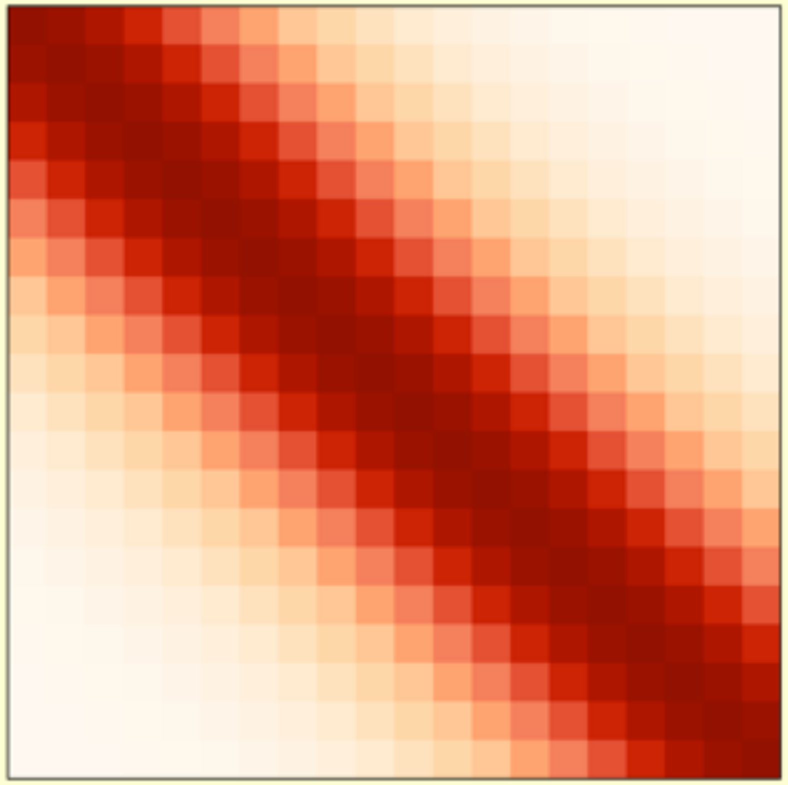


$= \Sigma$

'traditional' method

$$p(d \,|\, \mu, \Sigma) = \mathcal{N}(\mu, \Sigma)$$

$$\ln p(d \,|\, \mu, \Sigma) = -\frac{1}{2}\chi^2 - \frac{1}{2}\ln|\Sigma| - \frac{1}{2}\ln 2\pi$$



$= K + \Sigma$

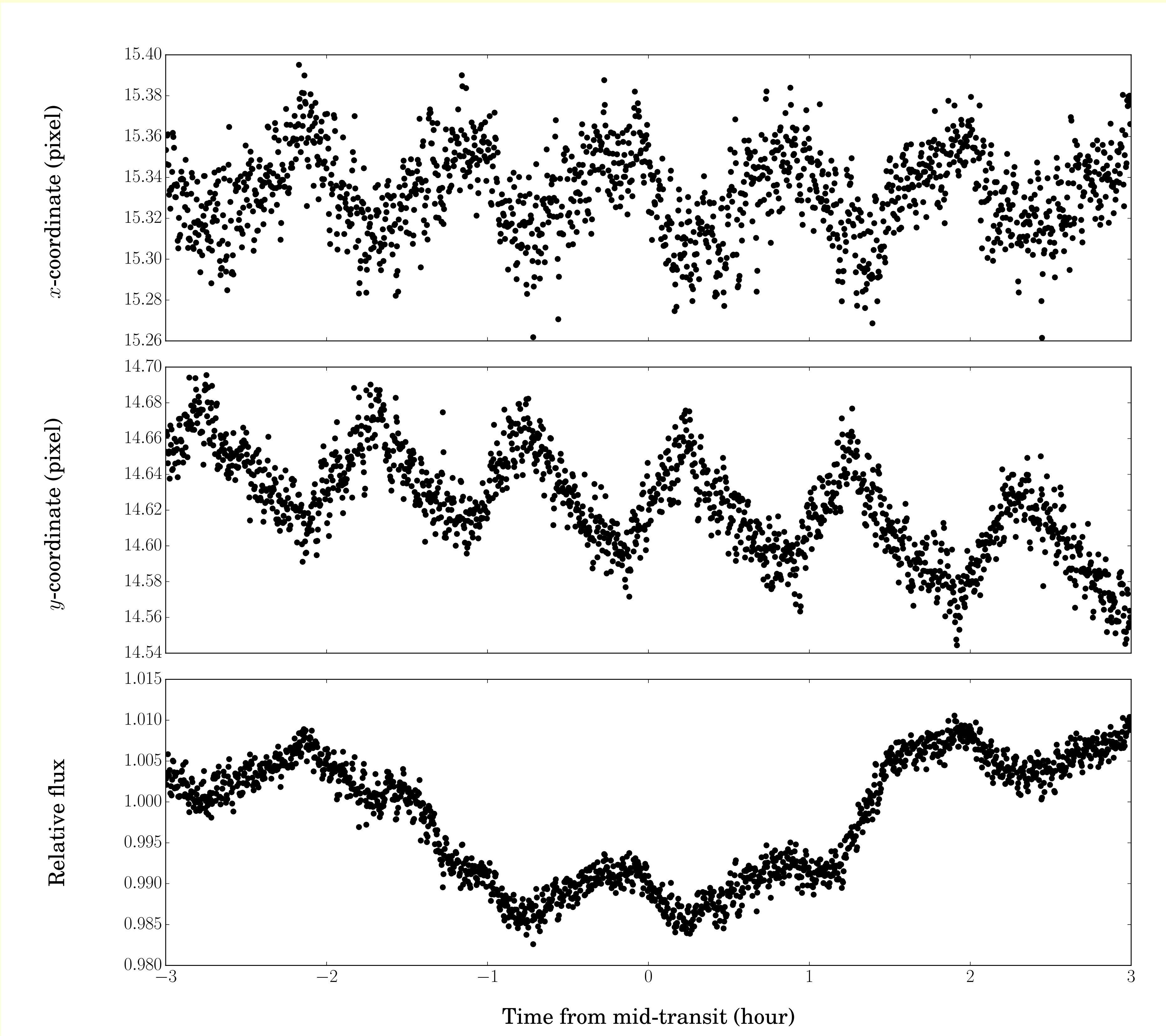
general Gaussian process

$$p(d \,|\, \mu, K+\Sigma) = \mathcal{N}(\mu, K+\Sigma)$$

$$\ln p(d \,|\, \mu, K+\Sigma) = -\frac{1}{2}(d-\mu)^{\mathrm{T}}(K+\Sigma)^{-1}(d-\mu) - \frac{1}{2}\ln|K+\Sigma| - \frac{1}{2}\ln 2\pi$$

Tom Evans (Exeter)

IRAC systematics



$$p\left(d \,|\, \mu, K + \Sigma\right)$$

random noise

correlated noise

deterministic model

GPs + other methods

GPs: comparison to pixel mapping methods

Gaussian kernel regression

$$W(x_i, y_i) = \frac{\sum_{j\neq i}^{n} k_{ij} f_j}{\sum_{j\neq i}^{n} k_{ij}}$$

Ballard et al (2010)

$$k_{ij} = \exp\left[-\frac{1}{2}\left(\frac{x_i - x_j}{\sigma_x}\right)^2 - \frac{1}{2}\left(\frac{y_i - y_j}{\sigma_y}\right)^2\right]$$

Knutson et al (2012), Lewis et al (2013)

$$k_{ij} = \exp\left[-\frac{1}{2}\left(\frac{x_i - x_j}{\sigma_x}\right)^2 - \frac{1}{2}\left(\frac{y_i - y_j}{\sigma_y}\right)^2 - \frac{1}{2}\left(\frac{\sqrt{\beta_i} - \sqrt{\beta_j}}{\sigma_\beta}\right)^2\right]$$

Gaussian process:

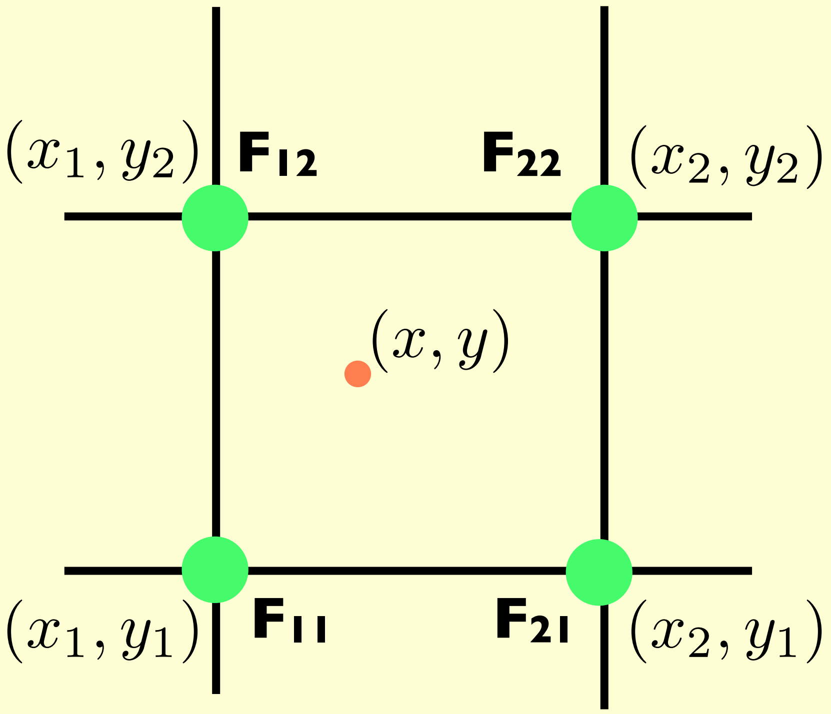
$$p(d \,|\, \mu, K + \Sigma) = \mathcal{N}\,(d \,|\, \mu, K + \Sigma)$$

$$K_{ij} = A^2 \exp\left[-\left(\frac{x_i - x_j}{L_x}\right)^2 - \left(\frac{y_i - y_j}{L_y}\right)^2\right]$$

GPs: comparison to pixel mapping methods

Bilinear interpolation (BLISS) Stevenson et al (2012)

$$\begin{aligned} W(x,y) &= F_{11}\left(\frac{x_2-x}{x_2-x_1}\right)\left(\frac{y_2-y}{y_2-y_1}\right) \\ &+ F_{21}\left(\frac{x-x_1}{x_2-x_1}\right)\left(\frac{y_2-y}{y_2-y_1}\right) \\ &+ F_{12}\left(\frac{x_2-x}{x_2-x_1}\right)\left(\frac{y-y_1}{y_2-y_1}\right) \\ &+ F_{22}\left(\frac{x-x_1}{x_2-x_1}\right)\left(\frac{y-y_1}{y_2-y_1}\right) \end{aligned}$$



Gaussian process:

$$p(d\,|\,\mu, K+\Sigma) = \mathcal{N}(d\,|\,\mu, K+\Sigma)$$

$$K_{ij} = A^2 \exp\left[-\left(\frac{x_i - x_j}{L_x}\right)^2 - \left(\frac{y_i - y_j}{L_y}\right)^2\right]$$

GPs: comparison to pixel-level decorrelation (PLD)

Deming et al (2015)

Measured flux

Unknown function of pixel intensities

1st-order Taylor expansion

$$d^t = \mathcal{F}\left(P_1^t, P_2^t, \ldots, P_n^t\right) \approx \sum_{i=1}^{n} \frac{\partial \mathcal{F}}{\partial P_i^t} P_i^t$$

$$\text{systematics}(t) = a_0 + b_1 t + b_2 t^2 + \sum_{i=1}^{n} c_i \hat{P}_i^t$$

$$\text{where} \quad \hat{P}_i^t = \frac{P_i^t}{\sum_{i=1}^{n} P_i^t}$$

Normalisation removes time-varying astrophysical signals

Tom Evans (Exeter)

GPs: comparison to pixel-level decorrelation (PLD)

Deming et al (2015)

Measured flux

Unknown function of pixel intensities

1st-order Taylor expansion

$$d^t = \mathcal{F}\left(P_1^t, P_2^t, \ldots, P_n^t\right) \approx \sum_{i=1}^{n} \frac{\partial \mathcal{F}}{\partial P_i^t} P_i^t$$

$$\text{systematics}(t) = a_0 + b_1 t + b_2 t^2 + \sum_{i=1}^{n} c_i \hat{P}_i^t$$

$$\text{where} \quad \hat{P}_i^t = \frac{P_i^t}{\sum_{i=1}^{n} P_i^t}$$

Normalisation removes time-varying astrophysical signals

Higher-dimensional input space than xy

Information rich!

Avoids centroiding issues

Works best for small pointing variations

Could be enhanced using GPs?

GPs: comparison to pixel-level decorrelation (PLD)

Could PLD be enhanced using GPs?

GP likelihood: $$p(d\,|\,\mu, K+\Sigma) = \mathcal{N}\left(d\,|\,\mu, K+\Sigma\right)$$

Use the centroid (x,y) coordinates as inputs:

$$K_{ij} = A^2 \exp\left[-\left(\frac{x_i - x_j}{L_x}\right)^2 - \left(\frac{y_i - y_j}{L_y}\right)^2\right]$$

Use the pixel time series as inputs:

$$K_{ij} = A^2 \exp\left[-\left(\frac{P_1^i - P_1^j}{L_1}\right)^2 - \left(\frac{P_2^i - P_2^j}{L_2}\right)^2 - \ldots - \left(\frac{P_n^i - P_n^j}{L_n}\right)^2\right]$$

GPs: comparison to independent component analysis (ICA)

Morello et al (2014, 2015a, 2015b)

Complementary to GPs

Advantages

- No assumptions made about the form of the systematics (!)
- Signals do not need to be stationary
- Computationally fast (?)

Restrictions (?)

- Requires multiple independent, simultaneous measurements
- Requires linear combination of signals to be valid

GPs: comparison to independent component analysis (ICA)

Morello et al (2014, 2015a, 2015b)

Complementary to GPs

Advantages

- No assumptions made about the form of the systematics (!)
- Signals do not need to be stationary
- Computationally fast (?)

Restrictions (?)

- Requires multiple independent, simultaneous measurements
- Requires linear combination of signals to be valid

Use in addition to GPs

Tom Evans (Exeter)

Summary: GP treatment of IRAC lightcurves

- Parameterise statistical covariance between data points, rather than enforce a deterministic systematics model
- Optimise/marginalise resulting likelihood function
- With (x,y) as inputs: Pixel mapping formalised within a Bayesian probabilistic framework
- With pixel time series as inputs: Enhanced PLD?
- Complementary to ICA

Strengths:	Weaknesses:
Flexible with few tunable parameters	Computationally intensive
Built-in Occam's razor	Limited to N∼1000 data points
Relatively weak assumptions, reduces subjectivity	Assumes 'stationary' noise properties

More details: **Evans et al (2015), MNRAS, 451, 680**

Python code: `github.com/tomevans`