A Fast, Two-dimensional Gaussian Process Method Based on Celerite: Applications to Transiting Exoplanet Discovery and Characterization

T.Gordon, E. Agol, D. Foreman-Mackey (2020)

Mateusz Mróz
Statistics Journal Club
OAUW 24.01.2023

Noise in transit event

- white noise:

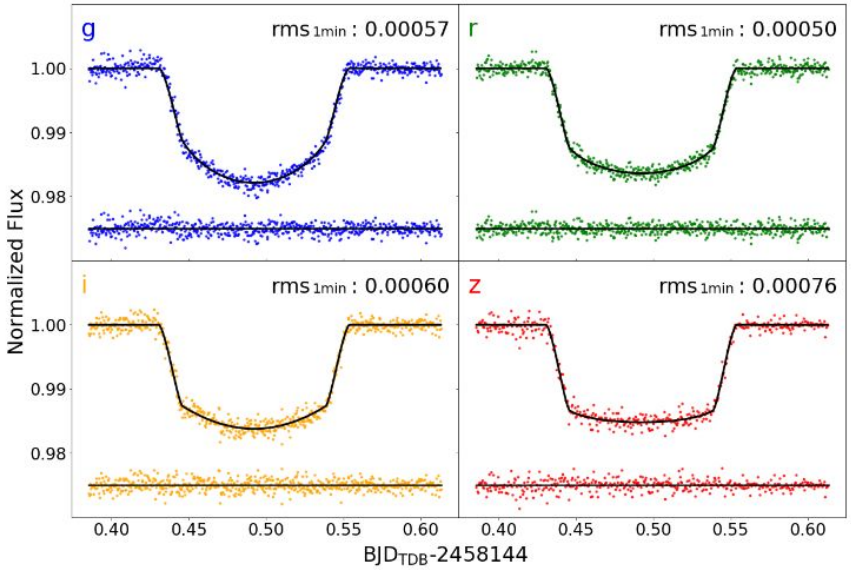
  instrumental,

  depend on photon counts

- correlated noise:

  astrophysical, e.g. stellar activity

  wavelength dependent



WASP12 N.Narita et al. (2019)

Gaussian Process

ordered collection of random variable along one or more axes

for exoplanets transits : random variables model a series of observations of the star's flux taken at discrete time

relation between random variables, we model N' observations with N'-dimensional Gaussian distribution.

characterized by :

- covariance matrix define by kernel function $k(x_i, x_j)$

$$K_{i,j} = k(x_i, x_j) + \delta_{i,j}\sigma_i^2, \qquad (1)$$

σ_i - white noise component

- mean function $\mu(t)$

deterministic component, transits model

GP likelihood function

$$\ln \mathcal{L} = -\frac{1}{2}(\boldsymbol{y} - \boldsymbol{\mu})^{\mathrm{T}} K^{-1}(\boldsymbol{y} - \boldsymbol{\mu}) - \frac{1}{2}\ln \det(K) - \frac{N'}{2}\ln(2\pi) \qquad (2)$$

$\boldsymbol{y}$ - observations

$\boldsymbol{\mu}_i = \mu(x_i)$ - model

Multiwavelength Noise Model

model of noise correlated across both time and wavelength

- Time-correlated Variability Model

- Wavelength Dependence of Variability

Time-correlated Variability Model

Anderson & Jefferies (1990)

> stellar oscillations from stochastic excitations damped by convection and turbulent excitation in the star

$$\frac{1}{\omega_0^2}\frac{d^2}{dt^2}y(t) + \frac{1}{\omega_0 Q}\frac{d}{dt}y(t) + y(t) = \epsilon(t) \qquad (3)$$

- ω_0 - characteristic frequency of the oscillations
- Q - quality factor of the oscillator
- $y(t)$ - amplitude of the oscillations
- $\epsilon(t)$ - stochastic driving force

Time-correlated Variability Model

if $\epsilon(t)$Gaussian distributed solution is GP with power spectrum :

$$S(\omega) = \sqrt{\frac{2}{\pi}} \frac{S_0 \omega_0^4}{(\omega^2 - \omega_0^2)^2 + \omega_0^2 \omega^2 / Q^2}. \qquad (4)$$

setting $Q = 1/\sqrt{2}$

$$S(\omega) = \sqrt{\frac{2}{\pi}} \frac{S_0}{(\omega/\omega_0)^4 + 1}. \qquad (5)$$

giving kernel function:

$$k(\tau) = S_0 \omega_0 e^{-\omega_0 \tau/\sqrt{2}} \cos\left(\frac{\omega_0 \tau}{\sqrt{2}} - \frac{\pi}{4}\right), \qquad (6) \qquad \tau = |t_i - t_j|$$

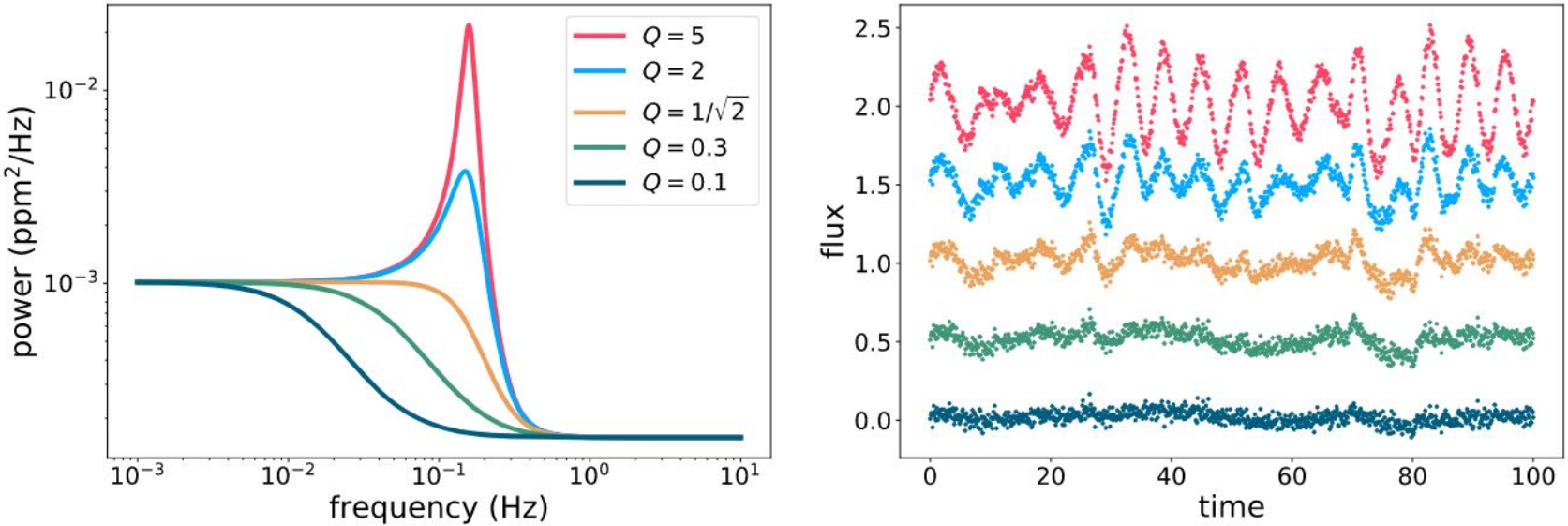


Figure 1. Left: power spectrum of the simple harmonic oscillator kernel for several values of the quality factor Q. For $Q < 1/\sqrt{2}$ the system is overdamped. For $Q > 1/\sqrt{2}$ the system is underdamped and the Gaussian process (GP) shows oscillations at the characteristic frequency. For our simulations we set $Q = 1/\sqrt{2}$, in which case the system is critically damped. Right: noise realizations for each power spectrum on the left. Note the decreasing coherency of the oscillations as we move from high to low values of Q. The decreasing noise amplitudes from top to bottom are a result of the fact that the GPs with larger Q values have more total power at constant S_0.

Wavelength Dependence of Variability

Stellar correlated variability model:

- two-component host's photosphere
- "hot" and "cold" spectral $S_{\rm h}(\lambda)$ and $S_{\rm c}(\lambda)$
- converting fractions $x_{\rm h}$ and $x_{\rm c} = 1 - x_{\rm h}$

$$F_{B_{\rm l}} = \frac{\pi R_*^2}{d^2} \int (x_{\rm c} S_{\rm c}(\lambda) + x_{\rm h} S_{\rm h}(\lambda)) \mathcal{R}_{B_{\rm l}}(\lambda) d\lambda \qquad (7)$$

$\mathcal{R}_{B_{\rm l}}(\lambda)$ - response curve for the filter

$$F_{B_1} = \frac{\pi R_*^2}{d^2}\left(\int S_{\rm h}(\lambda)\mathcal{R}_{B_1}(\lambda)d\lambda\right) - \frac{\pi R_*^2}{d^2}x_{\rm c}\left(\int (S_{\rm h}(\lambda) - S_{\rm c}(\lambda))\mathcal{R}_{B_1}(\lambda)d\lambda\right). \quad (8)$$

$$F_{\rm B_1,hot} = \frac{\pi R_*^2}{d^2}\int S_{\rm h}(\lambda)\mathcal{R}_{B_1}(\lambda)d\lambda \quad (9)$$

$$\alpha_1 = \frac{\pi R_*^2}{d^2}\sigma_c \int (S_{\rm h}(\lambda) - S_{\rm c}(\lambda))\mathcal{R}_{B_1}(\lambda)d\lambda, \quad (10) \qquad \sigma_c^2 = \mathrm{var}(x_{\rm c})$$

$$F_{B_1} = F_{\rm B_1,hot} - \frac{x_{\rm c}}{\sigma_c}\alpha_1. \qquad F_{B_2} = F_{B_2,\rm hot} - \frac{x_{\rm c}}{\sigma_c}\alpha_2.$$

$$\begin{aligned}\mathrm{cov}(F_{B_1}, F_{B_2}) &= \sigma_c^{-2}\mathrm{cov}(x_{\rm c}\alpha_1, x_{\rm c}\alpha_2)\\ &= \alpha_1\alpha_2\mathrm{corr}(x_{\rm c}, x_{\rm c}). \quad (13)\end{aligned}$$

$T_{i,j} = \mathrm{corr}(x_c(t_i), x_c(t_j))$

x_c -> $x_c(t)$

drawn from one-dimensional GP at times t_i for i= 1, ... , N

$$K = \begin{bmatrix} \Sigma_1 + T_{1,1}R & T_{1,2}R & \dots & T_{1,N}R \\ T_{2,1}R & \ddots & & \\ \vdots & & & \\ T_{N,1}R & & & \Sigma_N + T_{N,N}R \end{bmatrix}, \qquad (14)$$

$$\Sigma_i = \begin{pmatrix} \sigma_{i,1}^2 & 0 \\ 0 & \sigma_{i,2}^2 \end{pmatrix}$$

white noise

$$T_{i,j} = \mathrm{corr}(x_c(t_i), x_c(t_j))$$

time covariance

$$R = \begin{pmatrix} \alpha_1^2 & \alpha_1\alpha_2 \\ \alpha_2\alpha_1 & \alpha_2^2 \end{pmatrix}.$$

bands covariance

For M bands $B_1, B_2, \ldots, B_M$ with amplitudes $\alpha_1, \alpha_2, \ldots \alpha_M$

$$R = \begin{pmatrix} \alpha_1^2 & \alpha_1\alpha_2 & \ldots & \alpha_1\alpha_M \\ \alpha_2\alpha_1 & \ddots & & \alpha_2\alpha_M \\ \vdots & & & \vdots \\ \alpha_M\alpha_1 & \alpha_M\alpha_2 & \ldots & \alpha_M^2 \end{pmatrix} = \boldsymbol{\alpha}\boldsymbol{\alpha}^T \qquad (17)$$

$$\Sigma = \begin{bmatrix} \Sigma_1 & \ldots & 0 \\ 0 & \ddots & \\ \vdots & & \Sigma_N \end{bmatrix},$$

$$K = \Sigma + T \otimes R,$$

Validation of multiwavelength stellar variability

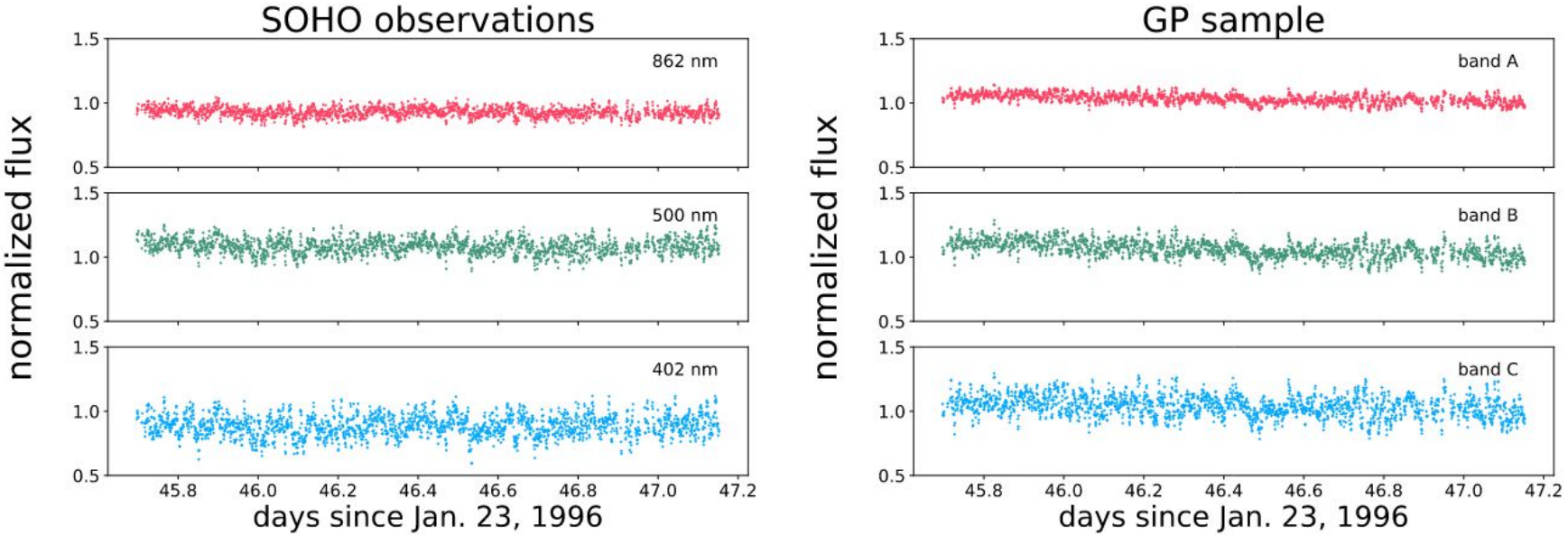


Figure 2. Left: SOHO three-channel Sun photometer time series of the Sun. Right: a three-band light curve simulated from a GP with a kernel consisting of three Kronecker-product terms (see Equation (59)), each term having the covariance described by Equation (17). The GP hyperparameters were obtained by optimizing the GP likelihood with respect to the data in the left panel.

Transit model

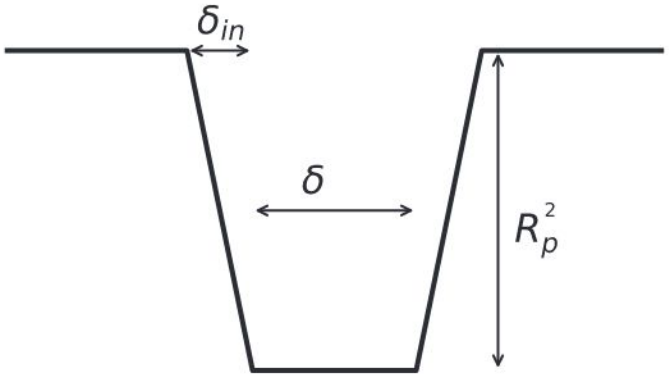


Figure 3. Schematic of the trapezoidal transit model. The center of transit t_0 is the midpoint of the transit.

$$\mu_{\text{trap}}(t, \boldsymbol{\theta})$$

$$\boldsymbol{\theta} = (R_p, t_0, \delta, \delta_{\text{in}})$$

Simulations

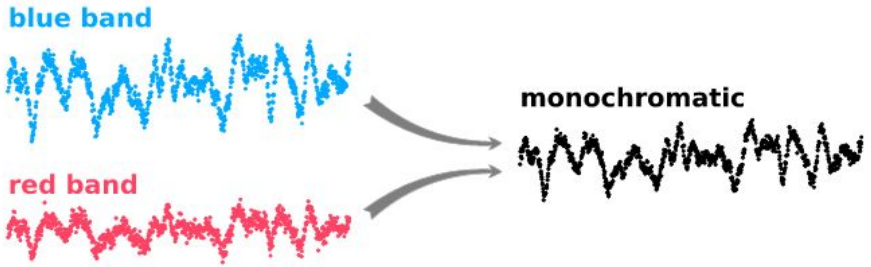


Figure 4. Two bands from a multiband simulation combined to simulate a monochromatic light curve with the same noise realization. Note that the white noise amplitude is smaller in the monochromatic light curve than for either individual band, while the amplitude of the correlated noise is the photon-weighted mean of the amplitude in the two bands. Here the blue band has a correlated noise amplitude twice that of the red band.

Three regimes

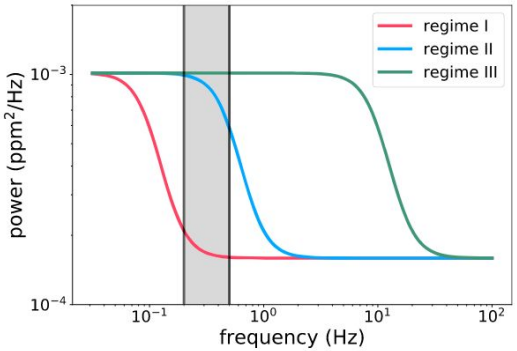


Figure 6. Power spectral densities for the three regimes. The shaded region spans from the inverse transit duration on the left to the inverse ingress/egress duration on the right. Note that the densities plotted here are only meant to be illustrative, and do not correspond to the power spectra of the light curves in Figure 5.

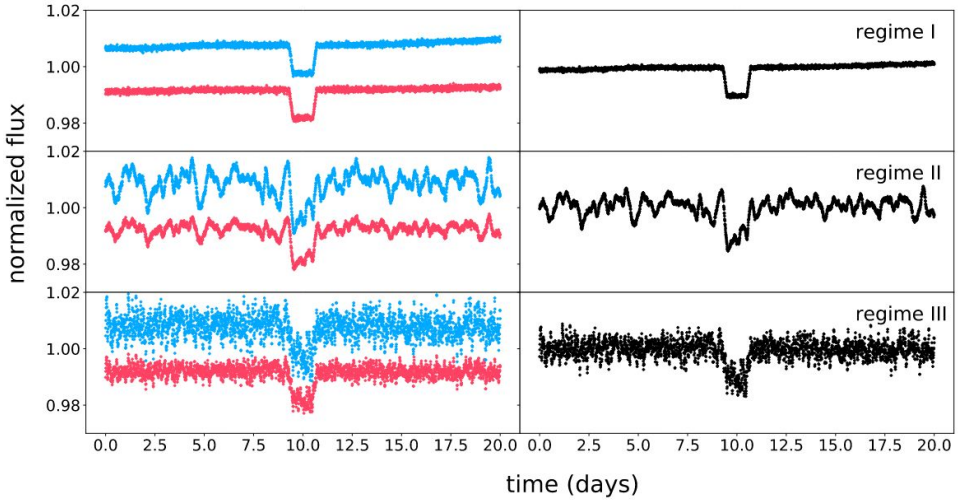


Figure 5. Representative light curves for the three noise regimes. The left panels show the two bands separately and the right panels show the monochromatic light curve resulting from the summation of the two bands. Top: in regime I the variability timescale is much longer than the transit duration. Middle: in regime II the variability timescale is between the transit duration and ingress/egress duration. Bottom: in regime II the variability timescale is shorter than the ingress/egress duration. Figure 6 shows power spectra corresponding to each of these regimes (but not to the light curves pictured here).

1. regime I: $1/f_0 > \delta$
2. regime II: $\delta_{\rm in} < 1/f_0 < \delta$
3. regime III: $1/f_0 < \delta_{\rm in}$

$$f_0 = \omega_0/(2\pi)$$

$$\bar{\alpha}^2 + \sigma^2 = \text{const}$$

$$\alpha_2 = 2\alpha_1$$

Information matrix

model made up of mean functions $\boldsymbol{\mu}_{\boldsymbol{\theta}}$ with N_θ parameters $\theta_1, \theta_2, \ldots \theta_{N_\theta}$

$$[\mathcal{I}_{\boldsymbol{\theta}}]_{i,j} = \left(\frac{d\boldsymbol{\mu}}{d\theta_i}\right)^{\mathrm{T}} K^{-1} \left(\frac{d\boldsymbol{\mu}}{d\theta_j}\right). \qquad (21)$$

$$[\mathcal{I}_{\boldsymbol{\theta}}^{-1}]_{i,j} \approx \mathrm{cov}(\theta_i, \theta_j).$$

Analytical Estimates for Parameter Uncertainties -R_p

- limb-darkening - ignored
- $\delta_{\rm in} \approx 0$
- no other uncertainty
- out-of-transit flux measured with high precision ->

$$\frac{\partial \boldsymbol{\mu}_{\rm trap}}{\partial R_p^2} = \begin{cases} 0 & {\rm out-of-transit} \\ -1 & {\rm in-transit} \end{cases}, \qquad (23)$$

Two-band

$$K = \begin{pmatrix} \sigma_1^2 + \alpha_1^2 & \alpha_1\alpha_2 \\ \alpha_1\alpha_2 & \sigma_2^2 + \alpha_2^2 \end{pmatrix},$$

$$\sigma^2_{R_p^2,\text{poly}} = \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)^{-1} \left(\frac{1 + \left(\frac{\alpha_1}{\sigma_1}\right)^2 + \left(\frac{\alpha_2}{\sigma_2}\right)^2}{1 + \frac{(\alpha_1 - \alpha_2)^2}{\sigma_1^2 + \sigma_2^2}} \right).$$

Monochromatic

Poisson noise

$$\sigma^2_{R_p^2,\text{mono}} = \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)^{-1} + \bar{\alpha}^2,$$

$$\bar{\alpha} = \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)^{-1} \left(\frac{\alpha_1}{\sigma_1^2} + \frac{\alpha_2}{\sigma_2^2}\right).$$

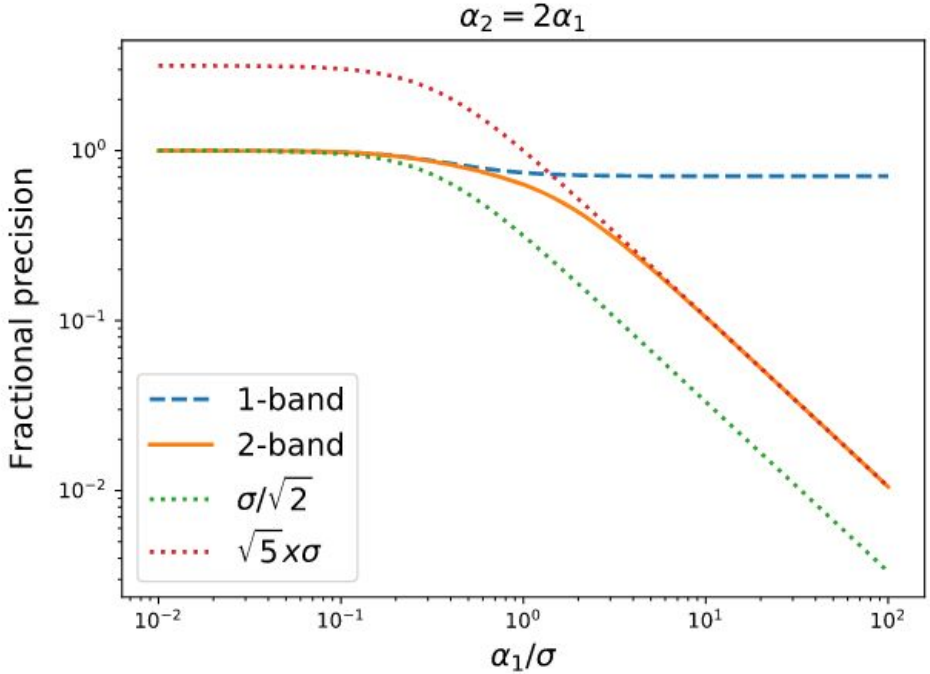


Figure 7. Analytic approximation for the fractional uncertainty on depth for two bands vs. the ratio of the correlated noise to white noise in first band, α_1/σ, in the limit of a constant amplitude of the sum of correlated and white noise (so that the white noise declines as the correlated noise increases). The ratio of the correlated noise in the two bands is two, i.e., $\alpha_2 = 2\alpha_1$. Plotted are the single-band case (blue dashed), two-band case (orange solid), and the white noise in each band, σ, times $1/\sqrt{2}$ and $\sqrt{5}$ (dotted). The fractional precision is normalized to the case $\alpha_1 = 0$.

M - bands generalization

$$\frac{\sigma^2_{R_p^2,M,\mathrm{poly}}}{\sigma^2} = \frac{1+\sigma^{-2}\sum_{i=1}^{M}\alpha_i^2}{M\left(1+\sigma^{-2}\sum_{i=1}^{M}\alpha_i^2\right)-\left(\sigma^{-1}\sum_{i=1}^{M}\alpha_i\right)^2}, \quad (28)$$

for the M-band case, and

$$\frac{\sigma^2_{R_p^2,M,\mathrm{mono}}}{\sigma^2} = \frac{1}{M}+\left(\frac{1}{M\sigma}\sum_{i=1}^{M}\alpha_i\right)^2. \quad (29)$$

$$\lim_{\sigma\ll\alpha_{1,2}} \frac{\sigma^2_{R_p^2,M,\mathrm{poly}}}{\sigma^2} = \frac{\sum_{i=1}^{M}\alpha_i^2}{M\sum_{i=1}^{M}\alpha_i^2-\left(\sum_{i=1}^{M}\alpha_i\right)^2}. \quad (30)$$

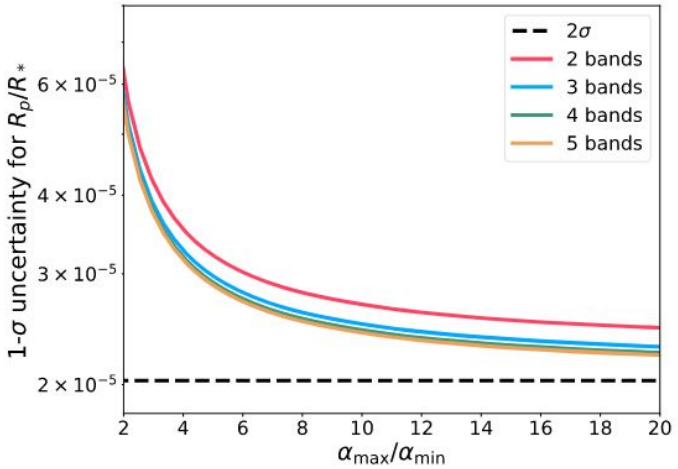


Figure 8. Information uncertainty curves for the planet–star radius ratio as a function of contrast ratio for a spectrum that increases linearly with photon flux from α_{min} to α_{max}. We plot the Information uncertainty for different values of M, the number of bands into which the spectrum is binned for modeling. The dashed line is the minimum uncertainty achievable as the contrast ratio becomes infinite which, for the two-band case, is equal to 2σ where σ is the Information uncertainty in the absence of correlated noise.

1D celerite*

GP for N flux measurements $\boldsymbol{y} = (y_1, y_2, \ldots y_N)$ taken at times $\boldsymbol{t} = (t_1, t_2, \ldots t_N)$

$$\ln \mathcal{L} = -\frac{1}{2}(\boldsymbol{y} - \boldsymbol{\mu})^{\mathrm{T}} K^{-1} (\boldsymbol{y} - \boldsymbol{\mu}) - \frac{1}{2} \ln \det(K) - \frac{N}{2} \ln(2\pi) \qquad (31)$$

$$\boldsymbol{\mu} = (\mu(t_1), \mu(t_2), \ldots \mu(t_N))$$

computing covariance matrix K requires $\mathcal{O}(N^3)$ operations

*celerity <- celeritas latin - quickness

celerite kernel, stationary GP

$$k(x_i, x_j) = k(|x_i - x_j|).$$

$$k_{\boldsymbol{\beta}}(t_n, t_m) = \sigma_n^2 \delta_{nm} + \sum_{j=1}^{J} \frac{1}{2}[(a_j + ib_j)e^{-(c_j+id_j)\tau_{nm}} + (a_j - ib_j)e^{(c_j-id_j)\tau_{nm}}] \quad (33)$$

$\boldsymbol{\beta} = (a_1 \ldots a_J, b_1 \ldots b_J, c_1 \ldots c_J, d_1 \ldots d_J)$

σ_n^2 variance of Gaussian - distributed white noise

$\tau_{nm} = |t_n - t_m| \quad n, m \in 1, \ldots, N$

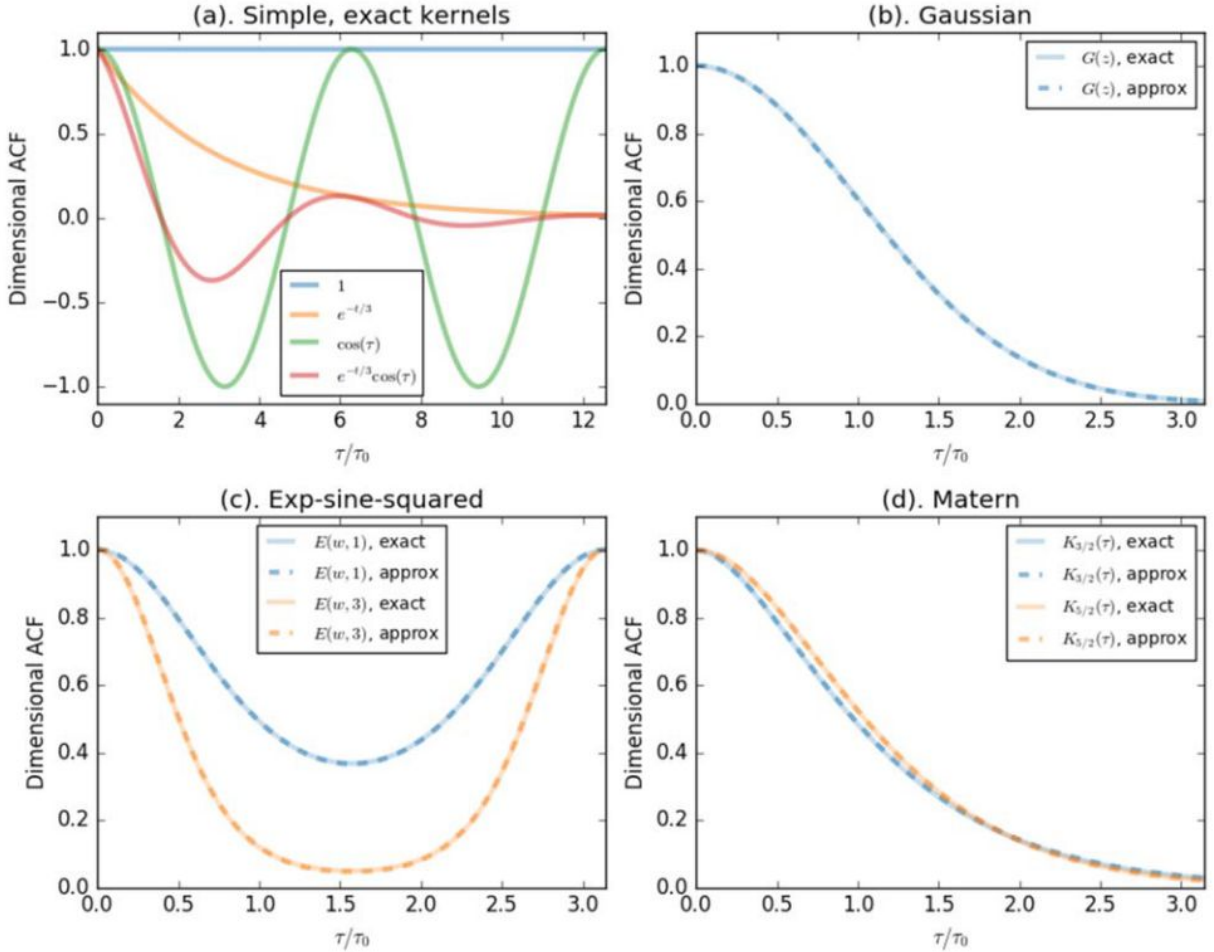


Figure 9. Approximation to various commonly used GP kernels. (a) Simple kernels with an exact celerite representation: cosine, or exponential times cosine. (b) Approximation of (a) referred to as "exponential-squared" to distinguish it from sine-squared kernels, indicated in (c). (d) Matern kernels.

..for such kernel covariance matrix is a symmetric and semiseparable

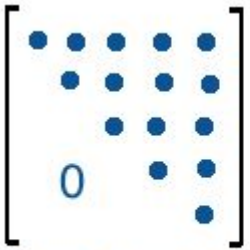
$$K = A + \text{tril}(UV^{\mathrm{T}}) + \text{triu}(VU^{\mathrm{T}}), \qquad (34)$$

U and V, both of size $(N \times P)$

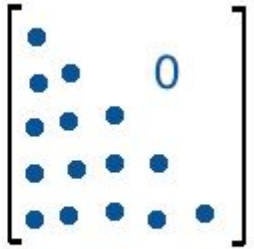
$$P = 2J$$

$$\begin{aligned}
U_{n,2j-1} &= a_j e^{-c_j t_n} \cos(d_j t_n) + b_j e^{-c_j t_n} \sin(d_j t_n), \\
U_{n,2j} &= a_j e^{-c_j t_n} \sin(d_j t_n) - b_j e^{-c_j t_n} \cos(d_j t_n), \\
V_{m,2j-1} &= e^{c_j t_m} \cos(d_j t_m), \\
V_{m,2j} &= e^{c_j t_m} \sin(d_j t_m),
\end{aligned} \qquad (35)$$

$$A_{n,n} = \sigma_n^2 + \sum_{j=1}^{J} a_j.$$



Upper Triangular Matrix



Lower Triangular Matrix

Cholesky factorization

$\mathcal{O}(NJ^2)$

$$K = LDL^{\mathrm{T}}$$

L is the lower-triangular Cholesky factor and D is a diagonal matrix

ansatz: $L = I + \mathrm{tril}(UW^{\mathrm{T}}).$

$$S_{n,j,k} = S_{n-1,j,k} + D_{n-1,n-1}W_{n-1,j}W_{n-1,k}$$

$$D_{n,n} = A_{n,n} - \sum_{j=1}^{P}\sum_{k=1}^{P} U_{n,j}S_{n,j,k}U_{n,k}$$

$$W_{n,j} = \frac{1}{D_{n,n}}\left[V_{n,j} - \sum_{k=1}^{P} U_{n,k}S_{n,j,k}\right],$$

$S_{j,k,1}$ is a matrix of zeros and P is both the rank of the semiseparable covariance matrix and the number of columns in U and V, here equal to 2J.

but exp in (35) can be numerically unstable -> additional matrix ϕ with same dimension as U and V

$$\tilde{U}_{n,2j-1} = a_j \cos(d_j t_n) + b_j \sin(d_j t_n)$$
$$\tilde{U}_{n,2j} = a_j \sin(d_j t_n) - b_j \cos(d_j t_n)$$
$$\tilde{V}_{m,2j-1} = \cos(d_j t_m)$$
$$\tilde{V}_{m,2j} = \sin(d_j t_m).$$

$$\tilde{W}_{n,2j-1} = e^{-c_j t_n} W_{n,2j-1}$$
$$\tilde{W}_{n,2j} = e^{-c_j t_n} W_{n,2j}.$$

$$\phi_{n,2j-1} = \phi_{n,2j} = e^{-c_j(t_n - t_{n-1})}.$$

$$S_{n,j,k} = \phi_{n,j}\phi_{n,k} \left[S_{n-1,j,k} + D_{n-1,n-1}\tilde{W}_{n-1,j}\tilde{W}_{n-1,k}\right]$$

$$D_{n,n} = A_{n,n} - \sum_{j=1}^{P}\sum_{k=1}^{P} \tilde{U}_{n,j} S_{n,j,k} \tilde{U}_{n,k}$$

$$\tilde{W}_{n,j} = \frac{1}{D_{n,n}} \left[\tilde{V}_{n,j} - \sum_{k=1}^{P} \tilde{U}_{n,k} S_{n,j,k}\right].$$

Two-dimensional GP

- first large dimension (of size N)
- second small dimension (of size M)

Cholesky decomposition in case when covariance of second dimension can be written as the outer product of a vector with itself:

$$R = \boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}},$$

possible when the correlated noise has the same shape along large dimension and varies proportionally in amplitude along the second small dimension

Covariance matrix

$$K = \Sigma + T \otimes R,$$

white noise diagonal matrix

covariance matrix for first dimension define by celerite kernel

covariance matrix in second dimension

$$R = \boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}},$$

$$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_M)^{\mathrm{T}}$$

$$
\begin{aligned}
K &= \Sigma + [A_0 + \mathrm{tril}(UV^{\mathrm{T}}) + \mathrm{triu}(VU^{\mathrm{T}})] \otimes R \\
&= \Sigma + \mathrm{diag}(A_0 \otimes R) \\
&\quad + \mathrm{tril}(UV^{\mathrm{T}} \otimes R) + \mathrm{triu}(VU^{\mathrm{T}} \otimes R), \qquad (48)
\end{aligned}
$$

$$
\begin{aligned}
K = \Sigma \; &+ \mathrm{diag}(A_0 \otimes R) \\
&+ \mathrm{tril}(UV^{\mathrm{T}} \otimes \boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}) \\
&+ \mathrm{triu}(VU^{\mathrm{T}} \otimes \boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}).
\end{aligned}
$$

$$(AB) \otimes (CD) = (A \otimes C)(B \otimes D),$$

$$
\begin{aligned}
K = \Sigma \; &+ \mathrm{diag}(A_0 \otimes R) \\
&+ \mathrm{tril}((U \otimes \boldsymbol{\alpha})(V \otimes \boldsymbol{\alpha})^{\mathrm{T}}) \\
&+ \mathrm{triu}((V \otimes \boldsymbol{\alpha})(U \otimes \boldsymbol{\alpha})^{\mathrm{T}}).
\end{aligned}
$$

structure as in 1D case

$$
\begin{aligned}
A' &= \Sigma + \mathrm{diag}(A_0 \otimes R) \\
U' &= U \otimes \boldsymbol{\alpha} \\
V' &= V \otimes \boldsymbol{\alpha}
\end{aligned}
$$

$$\tilde{U}'_{M(n-1)+p,2j-1} = \alpha_p(a_j \cos(d_j t_n) + b_j \sin(d_j t_n))$$

$$\tilde{U}'_{M(n-1)+p,2j} = \alpha_p(a_j \sin(d_j t_n) - b_j \cos(d_j t_n))$$

$$\tilde{V}'_{M(m-1)+p,2j-1} = \alpha_p \cos(d_j t_m)$$

$$\tilde{V}'_{M(m-1)+p,2j} = \alpha_p \sin(d_j t_m),$$

$$\phi'_{M(n-1)+p,:} = \begin{cases} e^{-c_j(t_n - t_{n-1})} & p = 1 \\ 1 & p > 1 \end{cases},$$

$$S_{n,j,k} = \phi'_{n,j}\phi'_{n,k}\,[S_{n-1,j,k} + D_{n-1,n-1}\tilde{W}_{n-1,j}\tilde{W}_{n-1,k}]$$

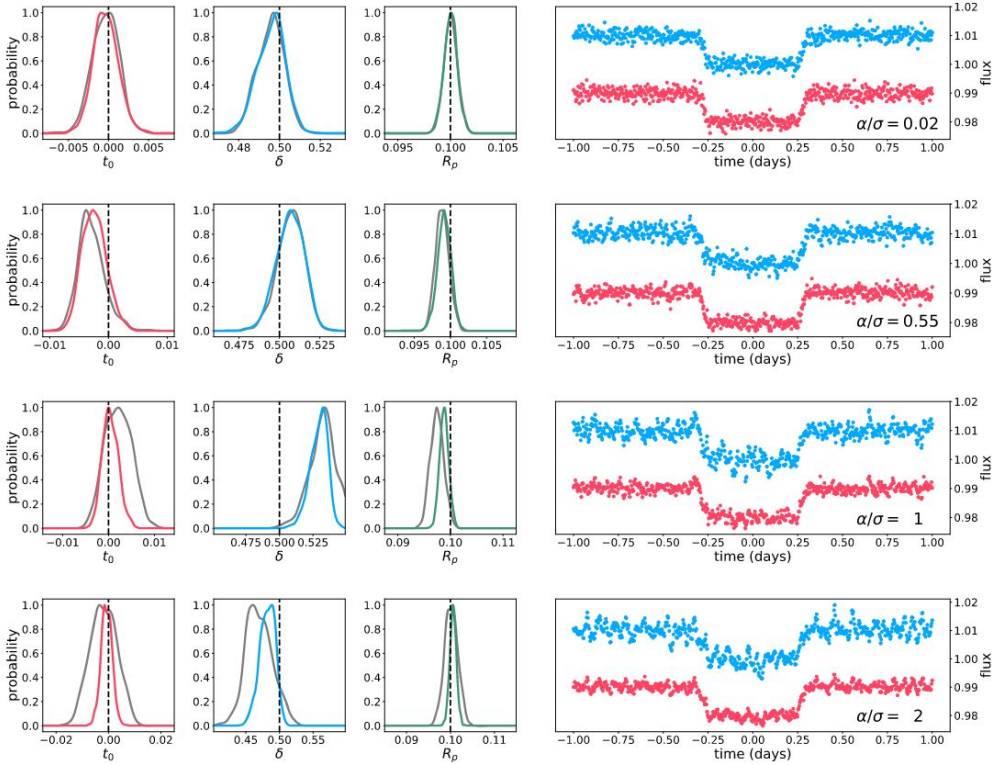
$$D_{n,n} = A'_{n,n} - \sum_{j=1}^{P}\sum_{k=1}^{P} \tilde{U}'_{n,j} S_{n,j,k} \tilde{U}'_{n,k},$$

$$\tilde{W}_{n,j} = \frac{1}{D_{n,n}}\left[\tilde{V}'_{n,j} - \sum_{k=1}^{P} \tilde{U}'_{n,k} S_{n,j,k}\right],$$

$n, m \in 1, \ldots, N,\ p \in 1, \ldots, M,$:

$$N' = NM \qquad \mathcal{O}(NMJ^2)$$

Results


probability
t_0
δ
R_p
flux
time (days)
$\alpha/\sigma = 0.02$
$\alpha/\sigma = 0.55$
$\alpha/\sigma = 1$
$\alpha/\sigma = 2$

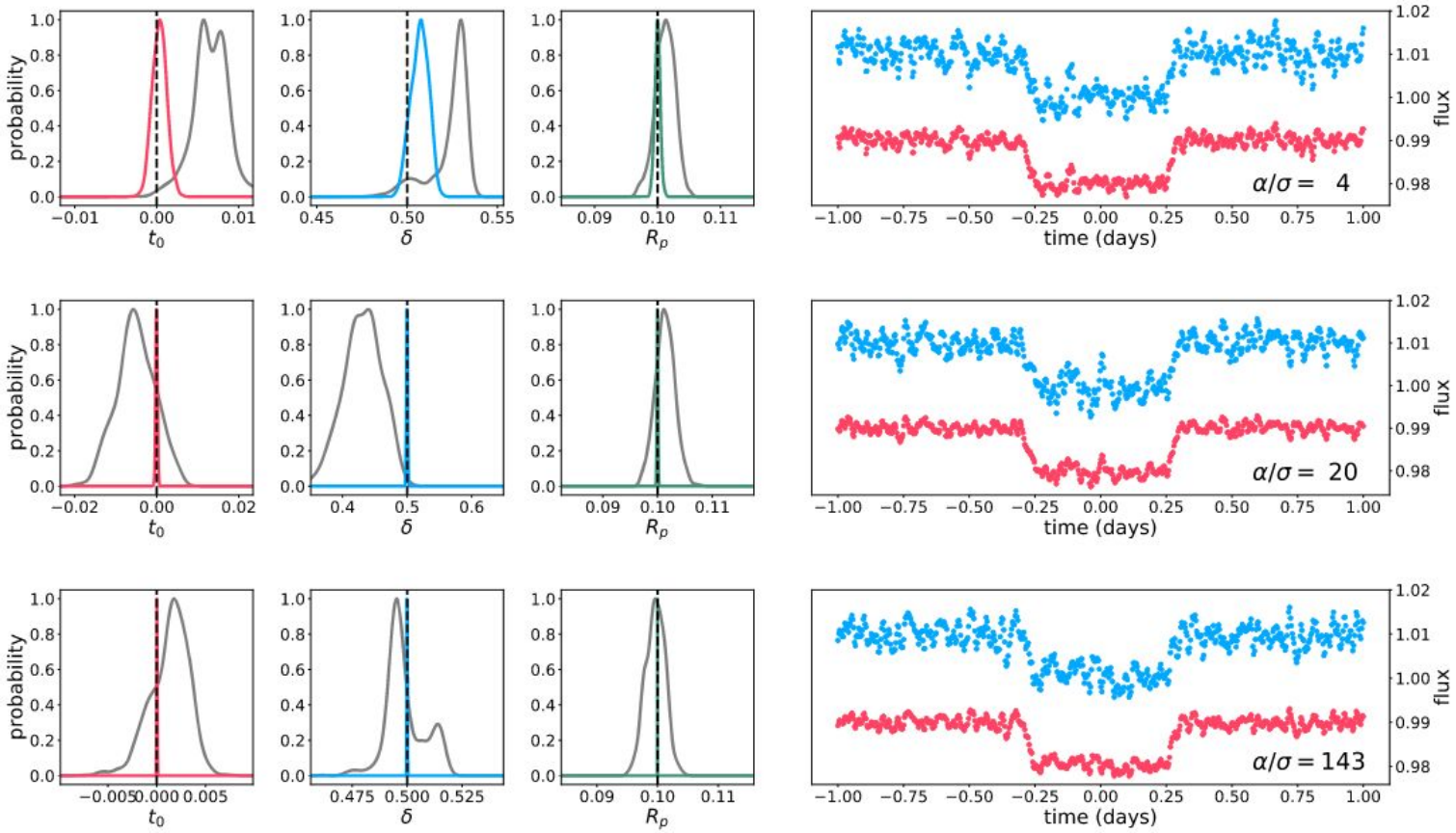
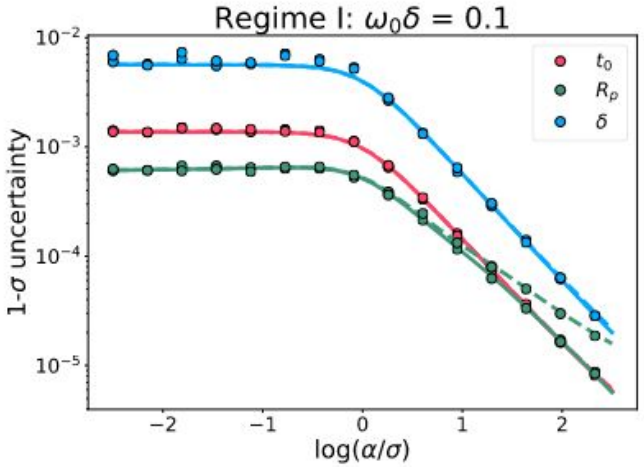
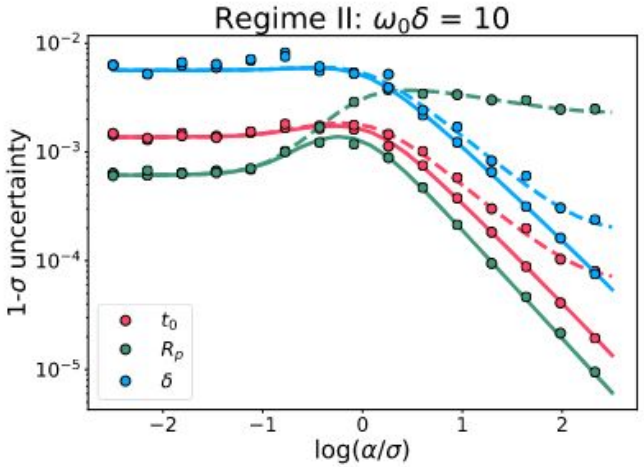


Figure 10. Left: posteriors for three transit parameters estimated by MCMC analysis on the two-band (colored) and single-band (gray) data. Posteriors are smoothed using Gaussian kernel density estimation for $\omega_0\delta = 100$ (corresponding to the final panel of Figure 11). From left to right: the center of transit t_0, transit duration δ, and radius ratio R_p/R_*. For $\alpha/\sigma = 20$ and $\alpha/\sigma = 143$ the posterior distributions for the two-band case are too sharply peaked to be visible. Right: representative light curves for each value of the noise amplitude ratio α/σ zoomed in on the transit signal (the input light curves have a duration of 10 days).





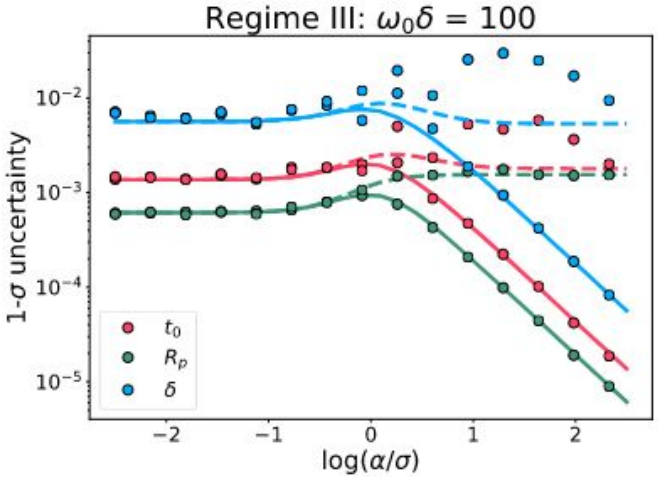


Figure 11. Information uncertainty curves overlaid with MCMC uncertainty estimates for trapezoidal transit parameters. Dashed lines show results for the monochromatic noise model and solid lines show results for the two-band noise model. Circles represent the MCMC uncertainty for distinct realizations of the noise and transit.

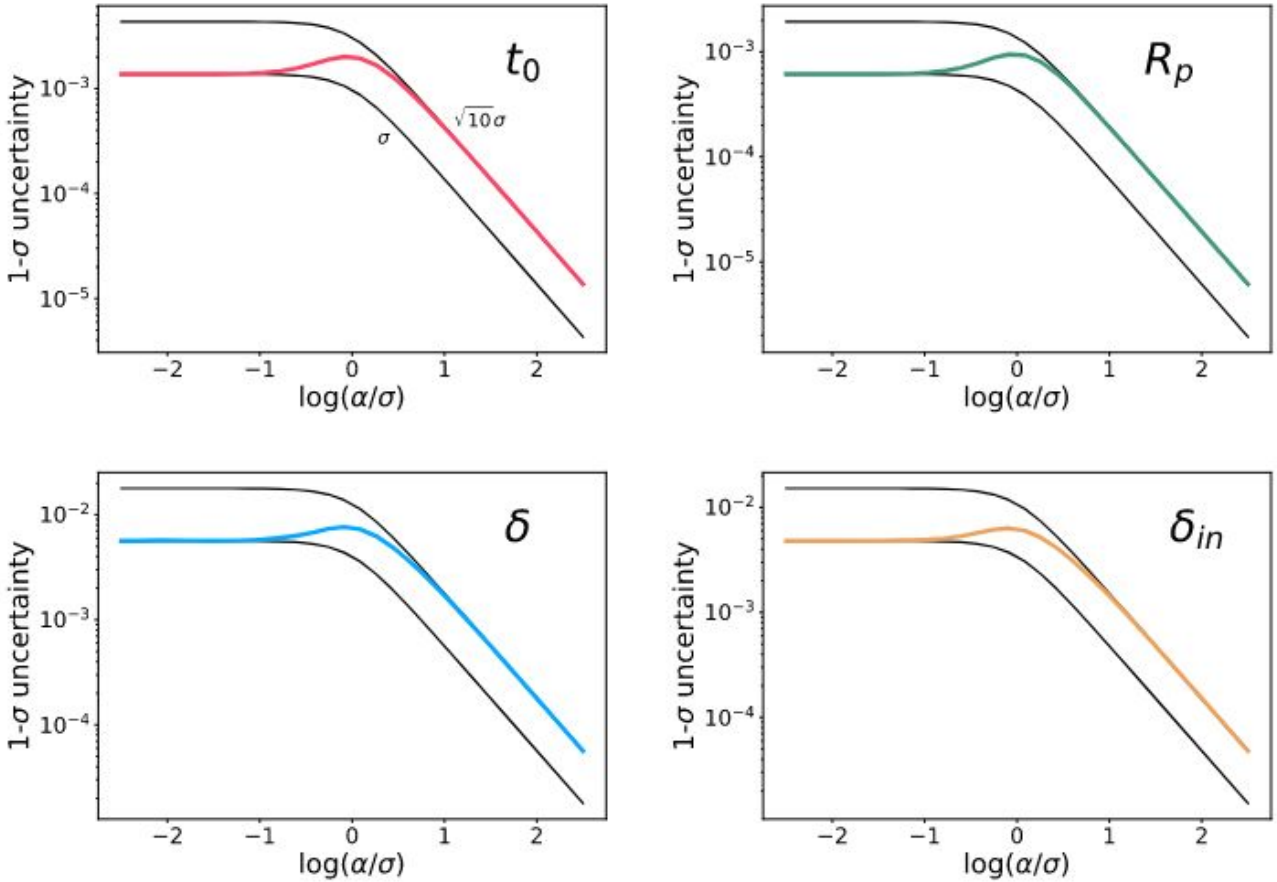


Figure 12. Information uncertainty curves (colored lines) for the two-band model compared to the white noise-only versions of the corresponding monochromatic noise model (black lines) in regime III. For the white noise-only models we set the correlated noise amplitude to zero and leave all other parameters the same as the monochromatic model. As we transition from the white noise-dominated to the correlated noise-dominated regimes the Information uncertainty curves for the two-band model transition from following the white noise model with $\sigma' = \sigma$ to the white noise model with $\sigma' = \sqrt{10}\sigma$. In effect perfect knowledge of the two-band correlated noise hyperparameters allows us to recover transit parameters at the same precision as if the correlated noise were simply white noise with a $\sqrt{10}$ larger amplitude.

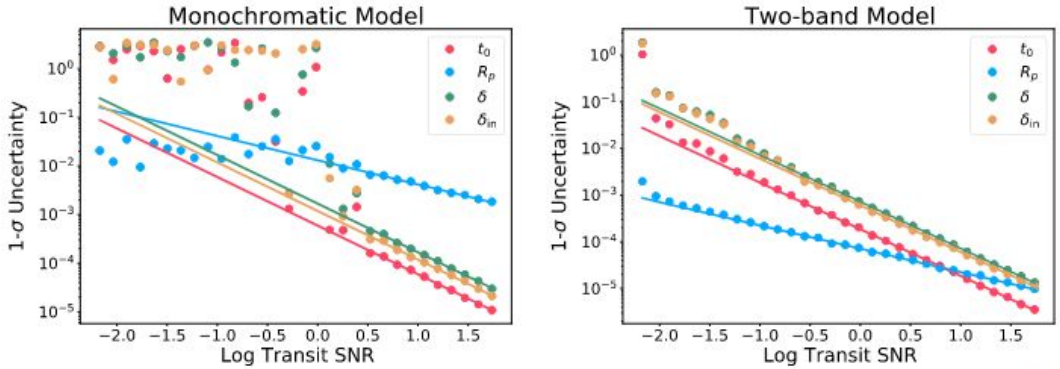


Figure 13. MCMC uncertainties (dots) and Information matrix uncertainties (lines) for monochromatic and two-band noise models as a function of the transit signal-to-noise ratio (S/N) with $\alpha_2 = 2\alpha_1$ for the two-band simulations. For these simulations the correlated noise is held constant at 150 times the amplitude of the white noise component and the total noise, defined to be the sum in quadrature of the white noise and correlated noise amplitudes, is conserved. The variability timescale $1/\omega_0 = \delta/10$, placing these simulations in regime II. For the monochromatic model, the Information and MCMC uncertainties correspond down to an S/N of about 10, which is the point at which the MCMC simulations no longer converge to the correct transit solution, as evidenced by the scatter in MCMC uncertainties at lower S/N. For the two-band simulations the Information and MCMC uncertainties correspond down to an S/N of 1/100.

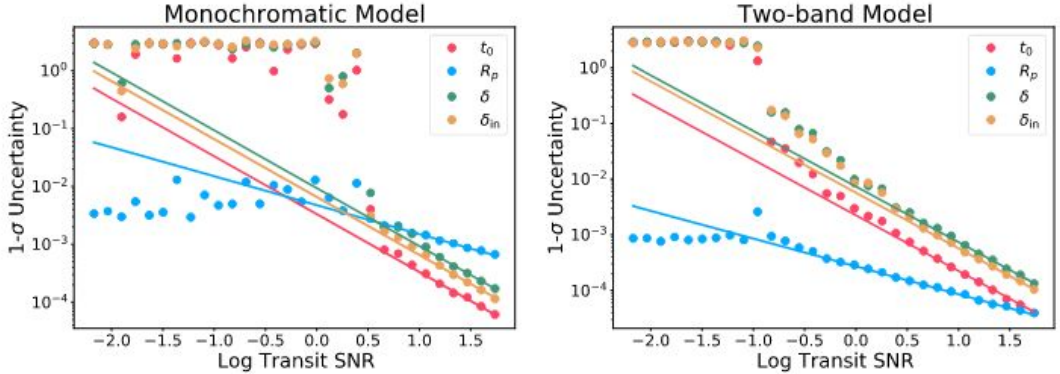
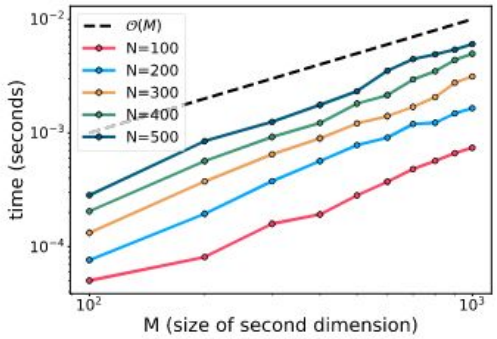
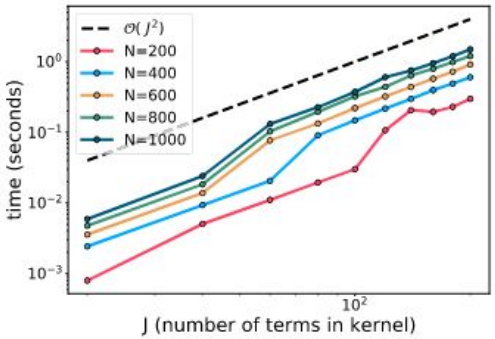
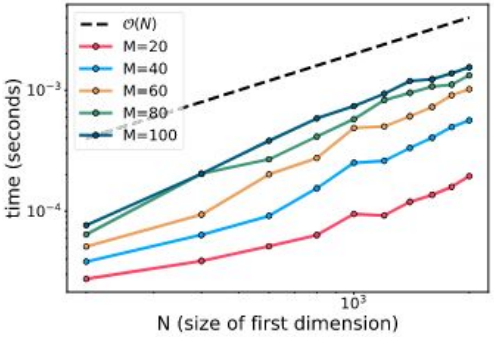
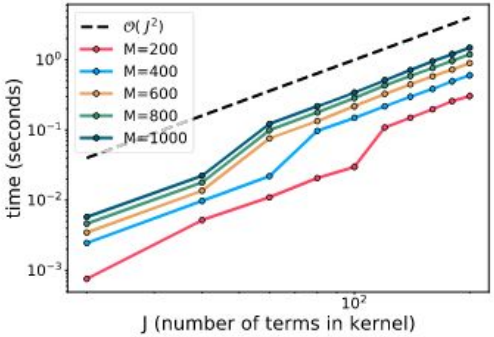


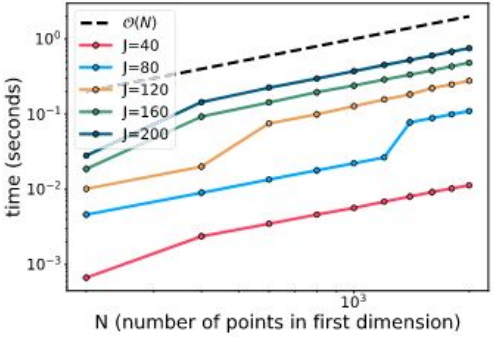
Figure 14. MCMC uncertainties (dots) and Information uncertainties (lines) for monochromatic and two-band noise models as a function of the transit S/N with wavelength dependence specified $\alpha_2 = 2\alpha_1$ for the two-band simulations. For these simulations the correlated noise is held constant at 10 times the amplitude of the white noise component and the total noise, defined to be the sum in quadrature of the white noise and correlated noise amplitudes, is conserved. The variability timescale $1/\omega = \delta/10$, placing these simulations in regime II. The larger white noise component compared to Figure 13 pushes the S/N limit below which the MCMC and Information uncertainties diverge to higher S/N. As before, there is an abrupt transition at this limiting S/N where the MCMC suddenly fails to converge to the correct transit solution.

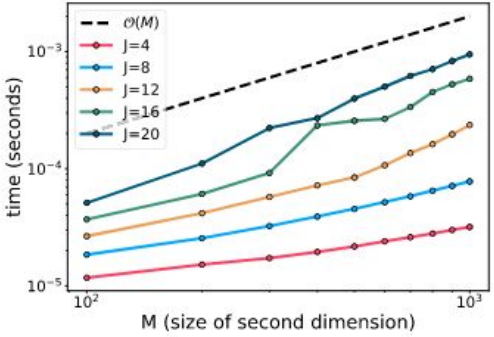

$\mathcal{O}(M)$
N=100
N=200
N=300
N=400
N=500
time (seconds)
M (size of second dimension)


$\mathcal{O}(J^2)$
N=200
N=400
N=600
N=800
N=1000
time (seconds)
J (number of terms in kernel)


$\mathcal{O}(N)$
M=20
M=40
M=60
M=80
M=100
time (seconds)
N (size of first dimension)


$\mathcal{O}(J^2)$
M=200
M=400
M=600
M=800
M=1000
time (seconds)
J (number of terms in kernel)


$\mathcal{O}(N)$
J=40
J=80
J=120
J=160
J=200
time (seconds)
N (number of points in first dimension)


$\mathcal{O}(M)$
J=4
J=8
J=12
J=16
J=20
time (seconds)
M (size of second dimension)

Limitations

- only short scale variability (stationary kernel)
- GP is bad with outliers (pre-processing)
- noise have be stationary, no stellar flares, outbursts…
- when time delay betten bands
- small-amplitude temperature variation, behaving as area variations
- every time of observation contains data in every band