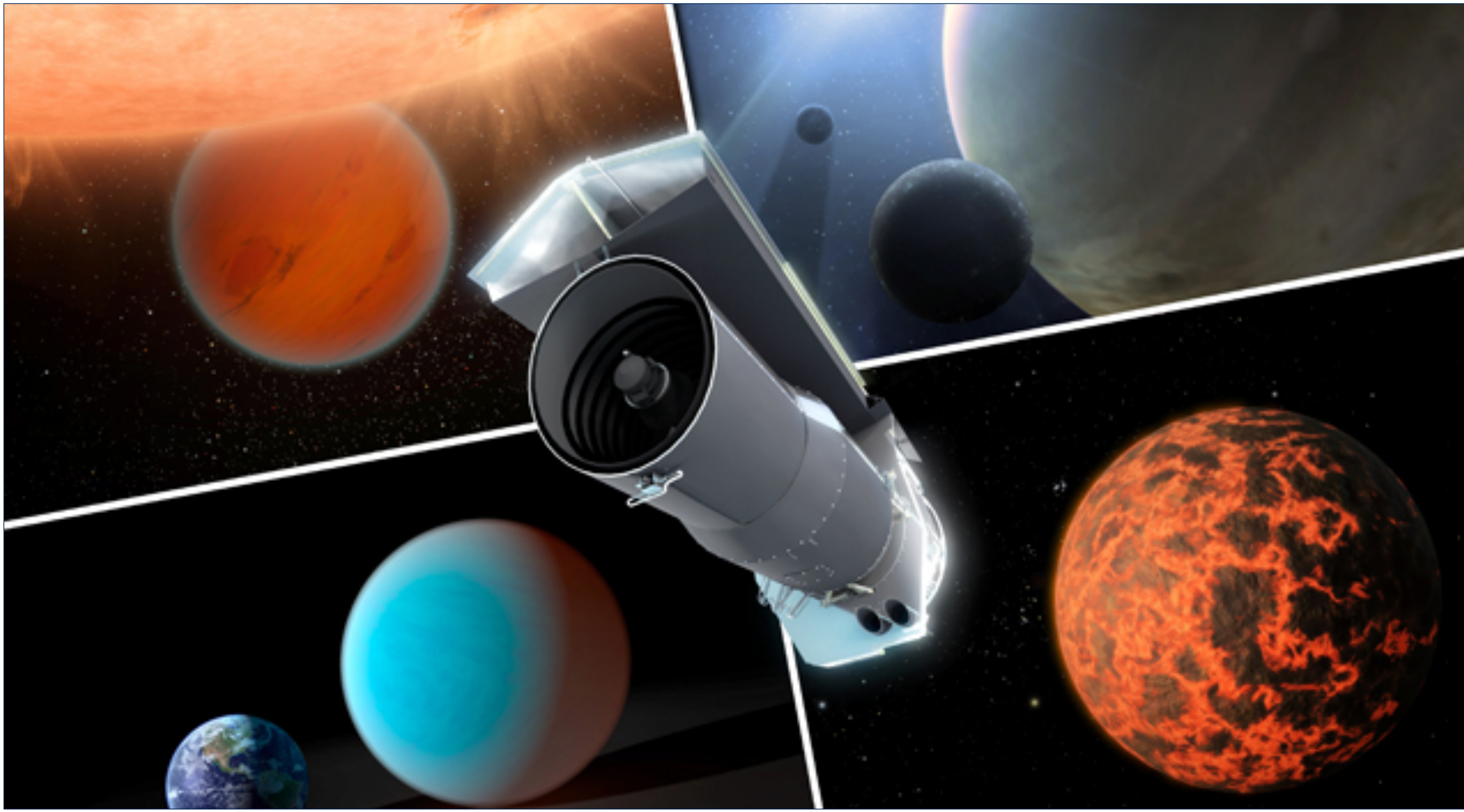
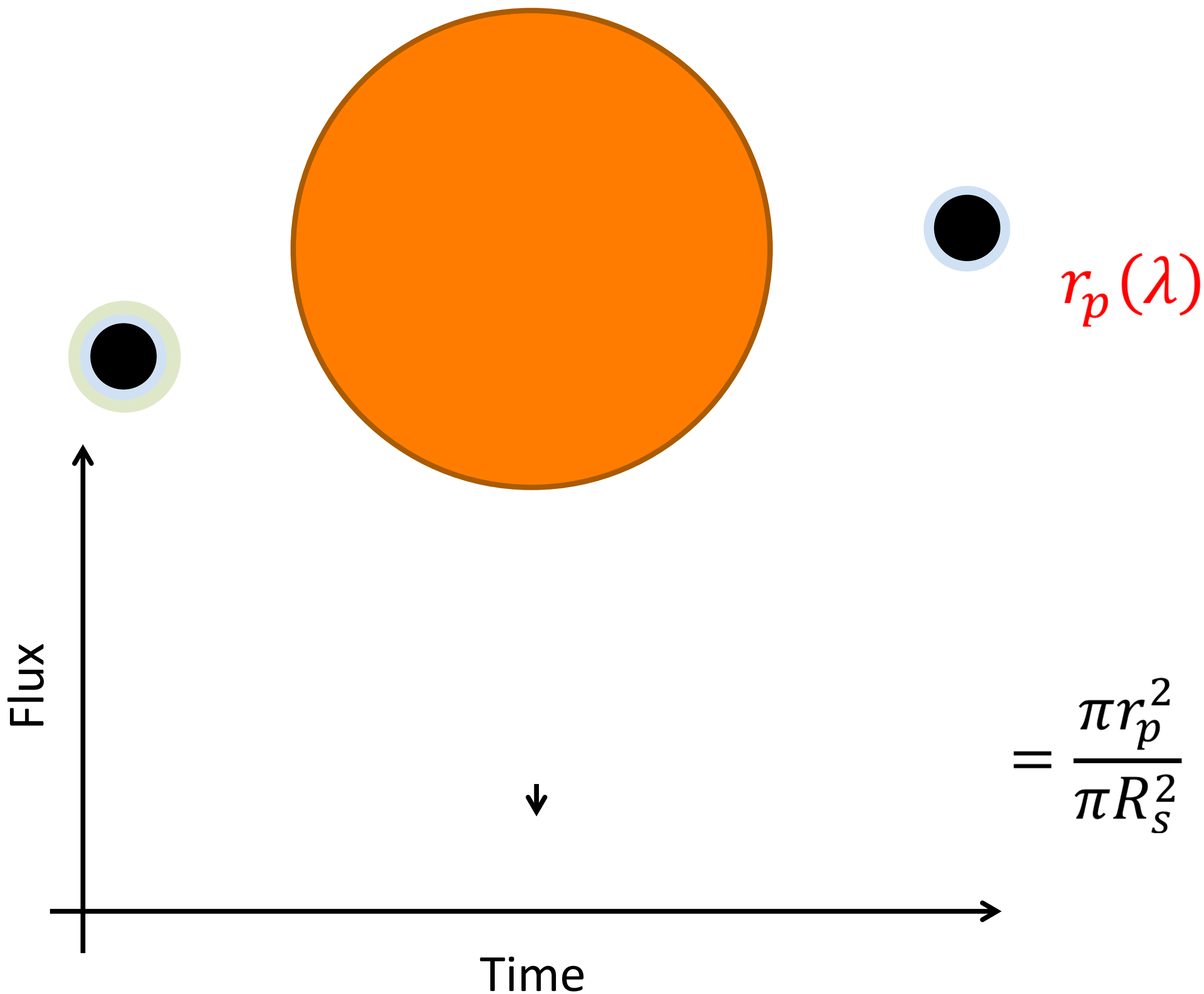
A blind method to detrend instrumental systematics in exoplanetary light-curves

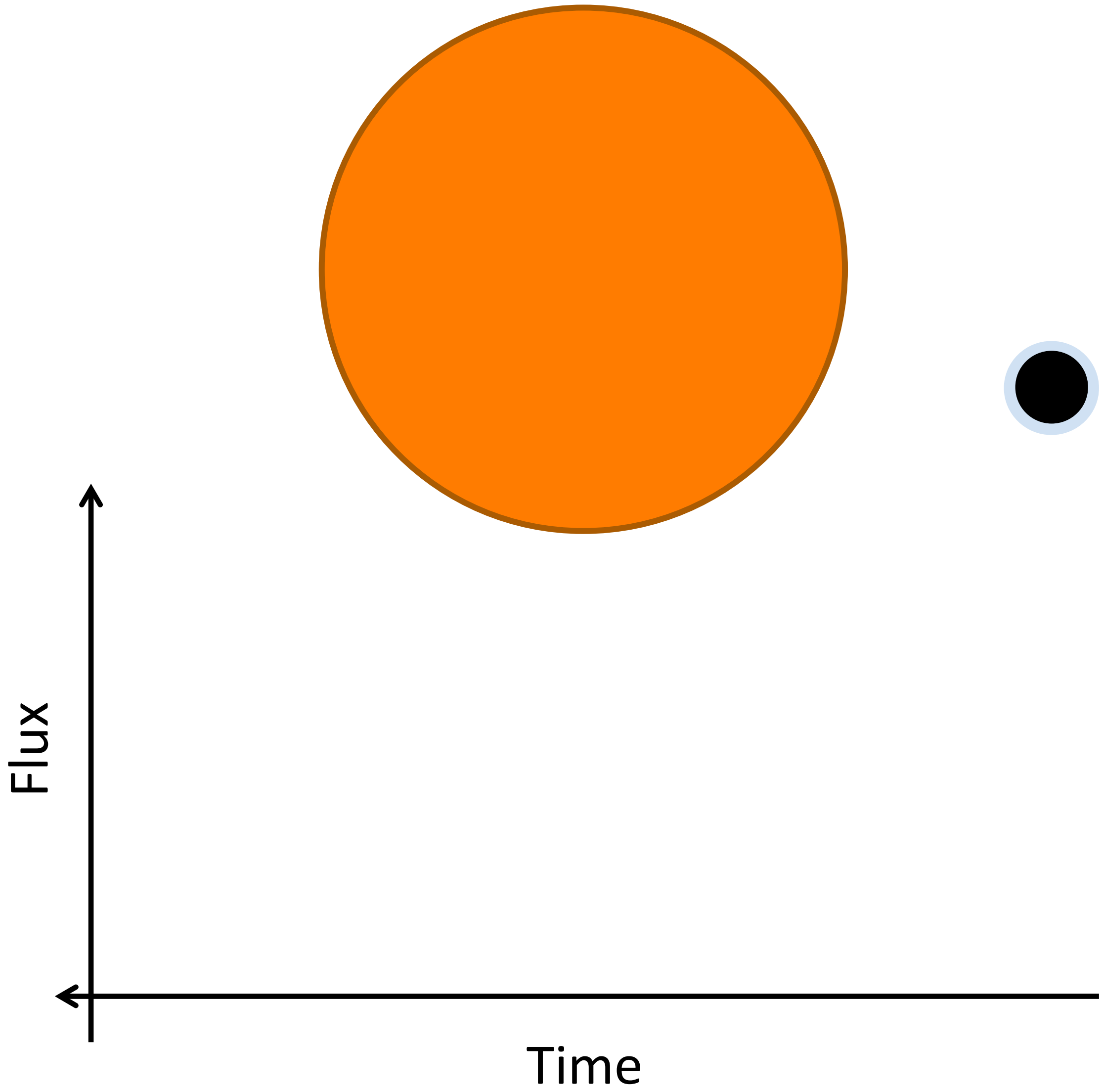


Giuseppe Morello, UCL

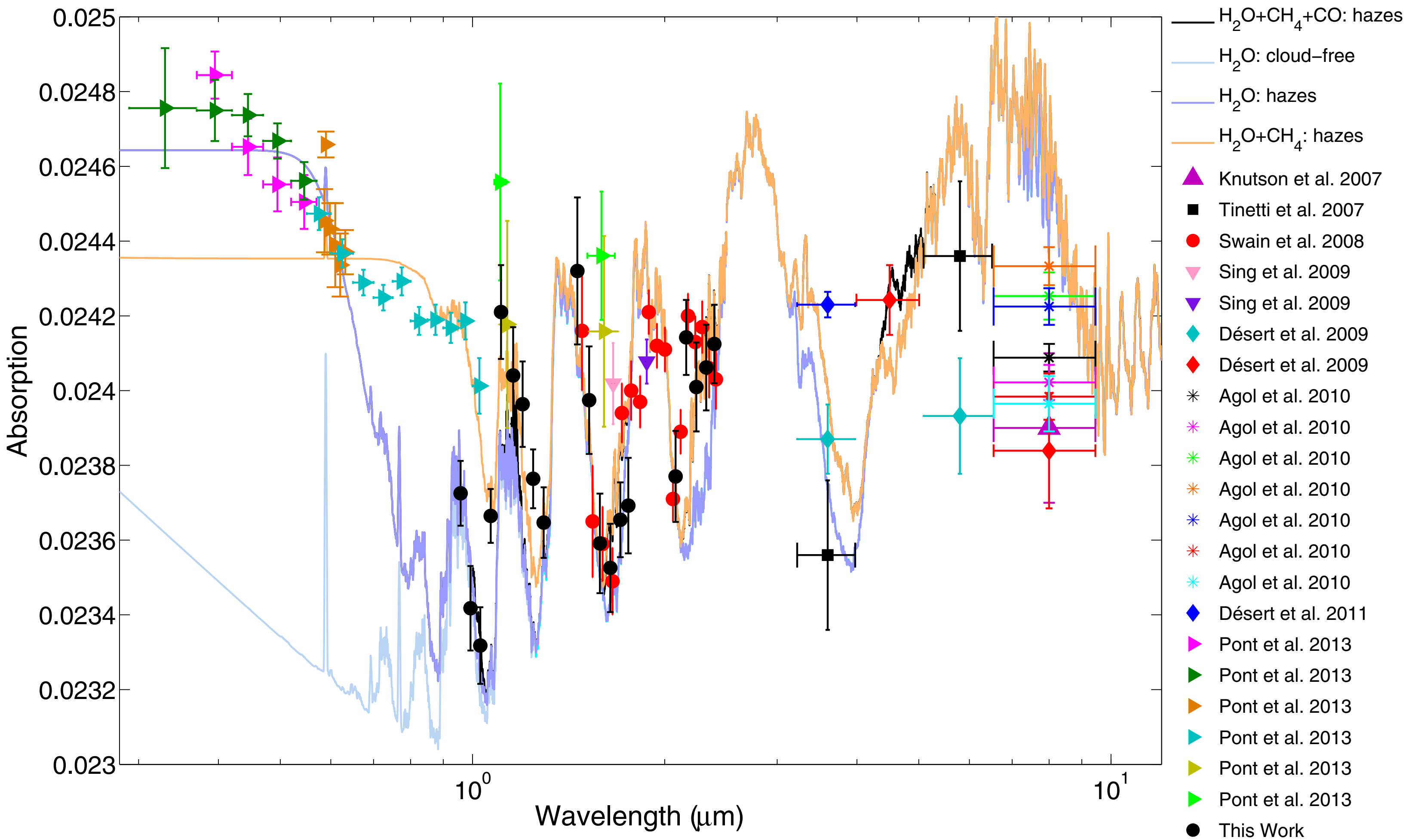
Transit of an exoplanet with atmosphere



Secondary eclipse of an exoplanet



HD189733b transmission spectrum



Danielski et al. 2014

Primary transit

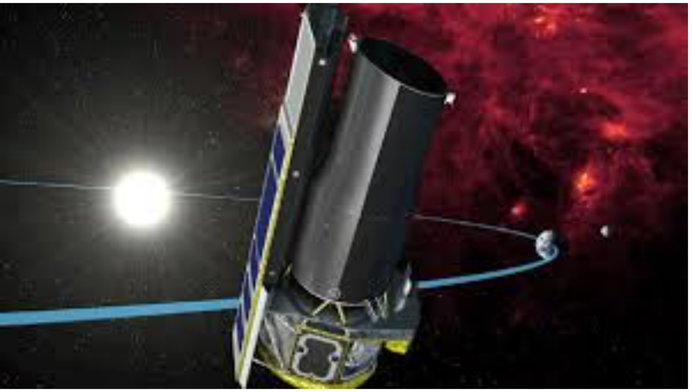
- Transmission spectroscopy
- Molecular absorbers, clouds
- Transit depth < 3%
- Photometric precision $\sim 10^{-4}$

Secondary eclipse

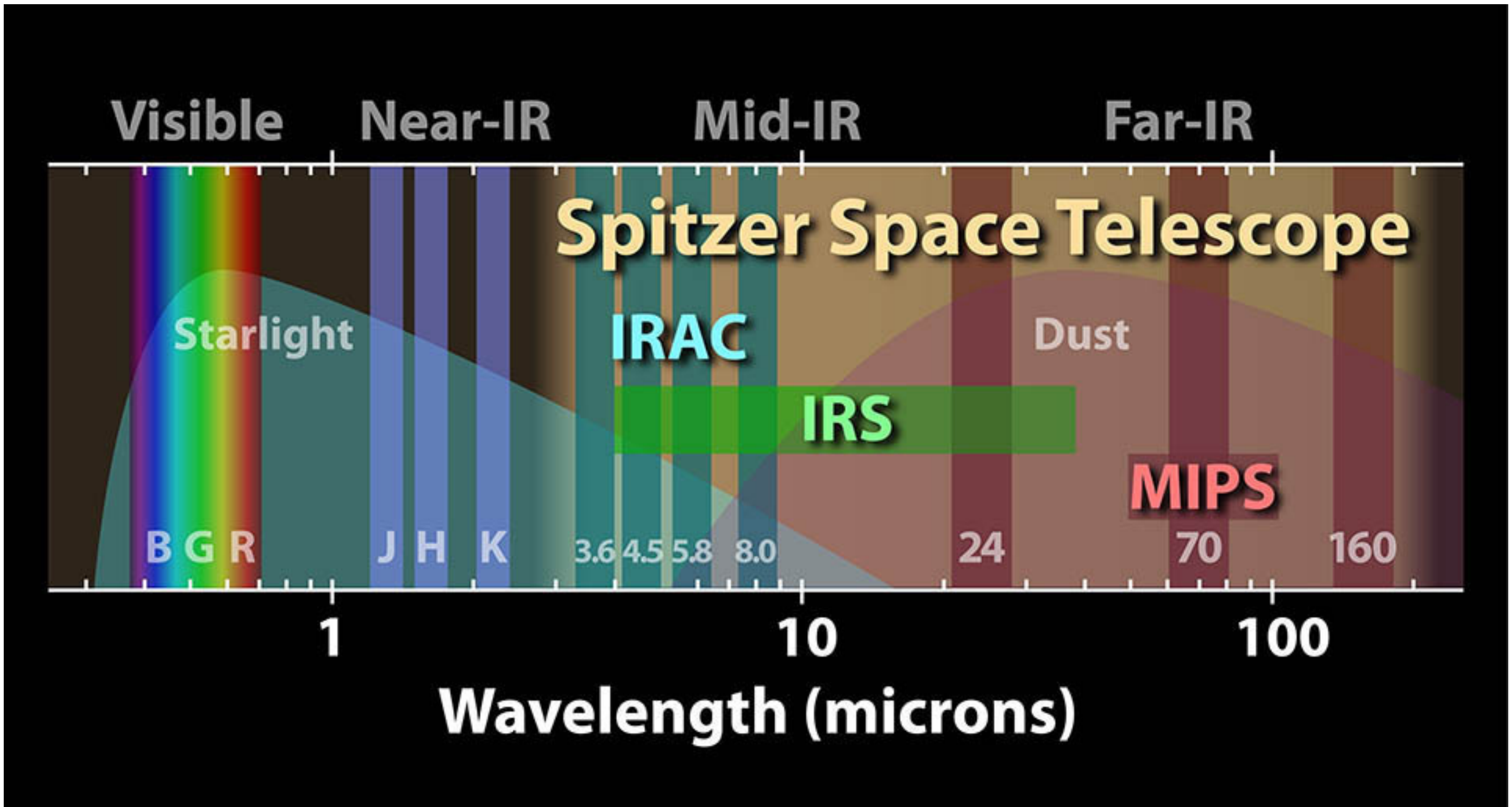
- Emission spectroscopy
- Thermal radiation, albedo
- Eclipse depth < 0.3%
- Photometric precision $\sim 10^{-4}$

Beyond the native precision of current instruments

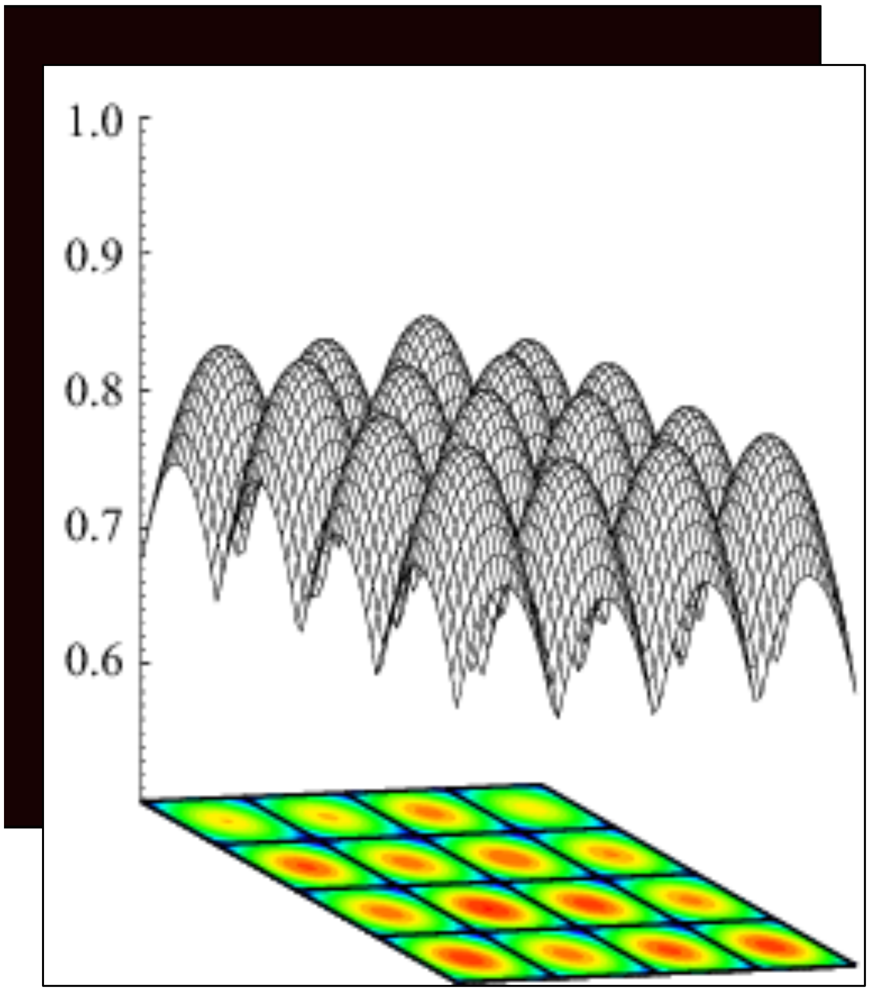
Data detrending is needed to reduce instrumental systematics

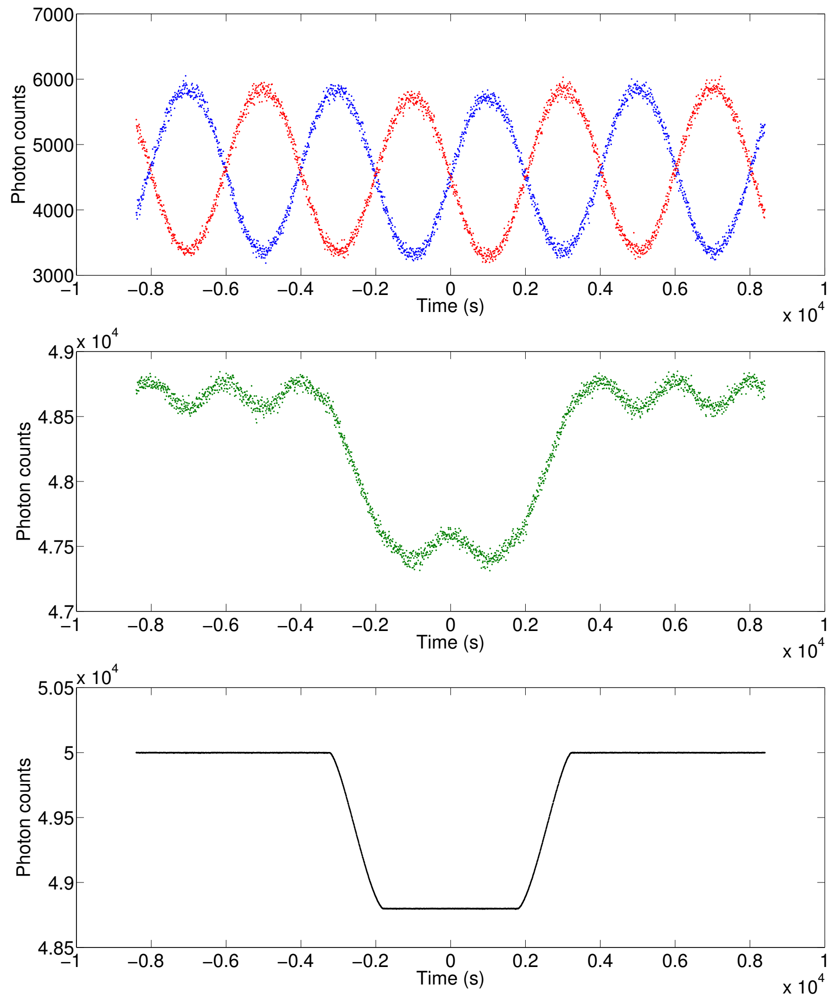


Spitzer Space Telescope



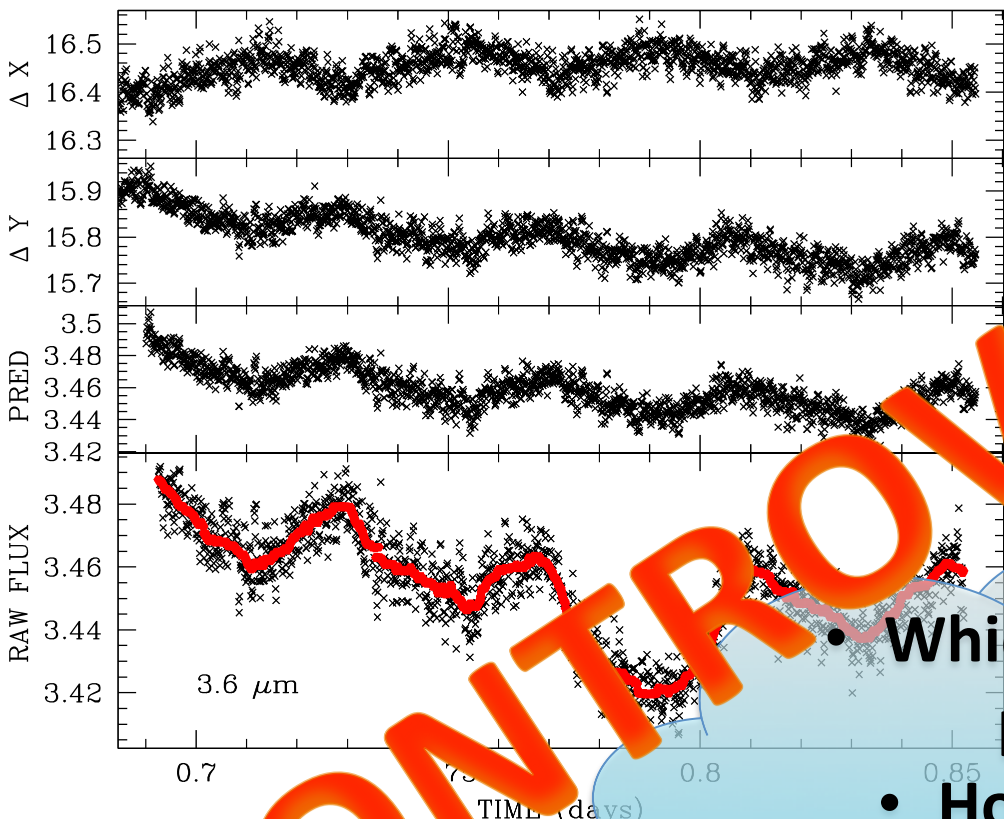
Spitzer pixel-phase effect





Morello 2015, ApJ, 808, 56

Parametric detrending



- Measured flux is correlated with centroid coordinates;
- Detrending by division for a polynomial function of centroid coordinates.

Beaulieu et al. 2011

CONTROVERSIES

- **Which degree of the polynomial?**
- **How to estimate centroid?**
- **Is this the only effect?**

Newer detrending techniques for Spitzer

- Spatial weighting functions (e.g. Ballard et al. 2010, Cowan et al. 2012, Lewis et al. 2013)
- Bliss mapping (Stevenson et al. 2012, b)
- Independent Component Analysis (Morello et al. 2014, 2015, Morello 2015)
- Pixel-level decorrelation method (Deming et al. 2014)
- Gaussian Processes (Gibson et al. 2012, Evans et al. 2015)

Independent Component Analysis

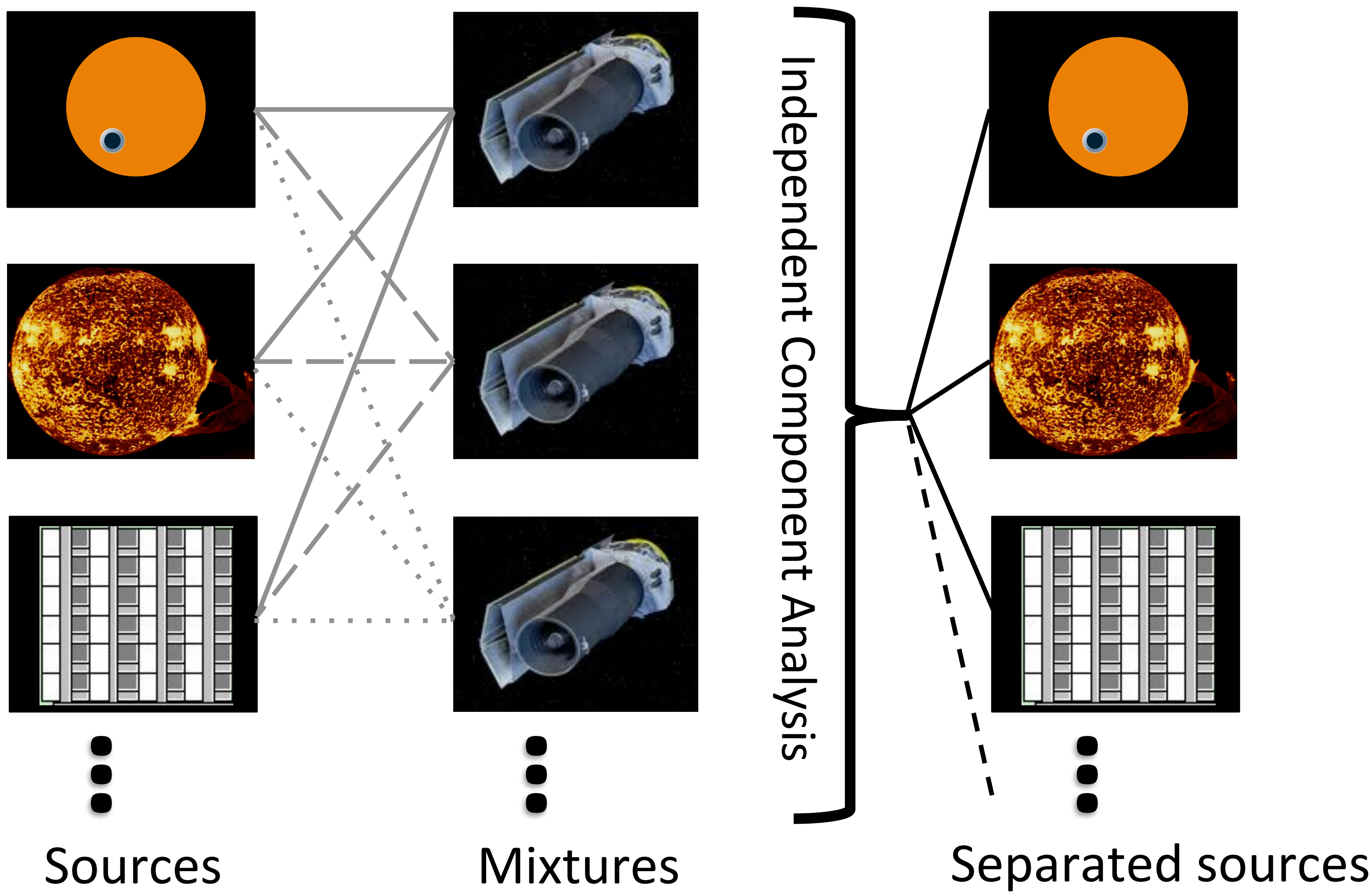
- **Blind Source Separation** technique, i.e. no prior knowledge of the instrument systematics
- Applicable in a general context, not just IRAC light-curves

ICA in astrophysics

- ICA has been used to separate the cosmic microwave background or signatures from distant galaxies from their galactic foregrounds and instrumental noise (e.g. Stivoli et al. 2006, Maino et al. 2002, 2007, Aumont & Macías-Pérez 2007, Wang et al. 2010).

- ICA has been used to detrend exoplanetary light-curves taken with different instruments (Waldmann et al. 2013, Waldmann 2012, 2014).

Independent Component Analysis



ICA: mathematical model

$$x_k = a_{k,1} s_a + a_{k,2} s_{wn} + \sum_{l=3}^{N_{sn}} a_{k,l} s_{sn}$$

observed time series

astrophysical source

white noise

systematic noise

observations

signals

$$X = AS \longrightarrow S = A^{-1} X$$

mixing matrix

UNKNOWN

ICA: statistics (1)

$$H(\mathbf{y}) = -\sum_k p(\mathbf{y}_k) \log p(\mathbf{y}_k)$$ Shannon entropy

It is the statistical measure of uncertainty associated with a random variable.

$$I(y_1, y_2, ..., y_n) = \sum_{i=1}^{n} H(y_i) - H(\mathbf{y})$$ mutual information

maximum independence = minimum mutual information

$$I(s_1, s_2, ..., s_n) = \sum_i H(s_i) - H(\mathbf{x}) - log|det(\mathbf{W})|$$

ICA: statistics (2)

Among all the distributions with fixed mean and covariance, the gaussian distribution has the maximum entropy.

$$J(\mathbf{y}) = H(\mathbf{y}_{gauss}) - H(\mathbf{y})$$ negentropy

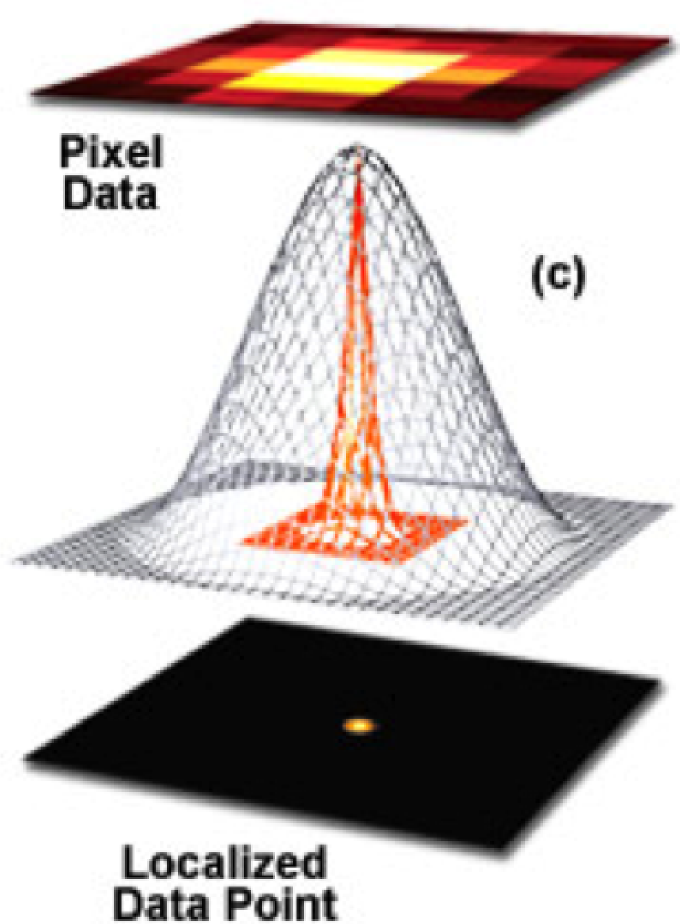
- Mutual information and negentropy are hard computing.
- Alternatively, we can maximize non-gaussianity of the source signals, through different estimators.

$$kurt(y) = E(y^4) - 3E(y^2)^2$$ kurtosis

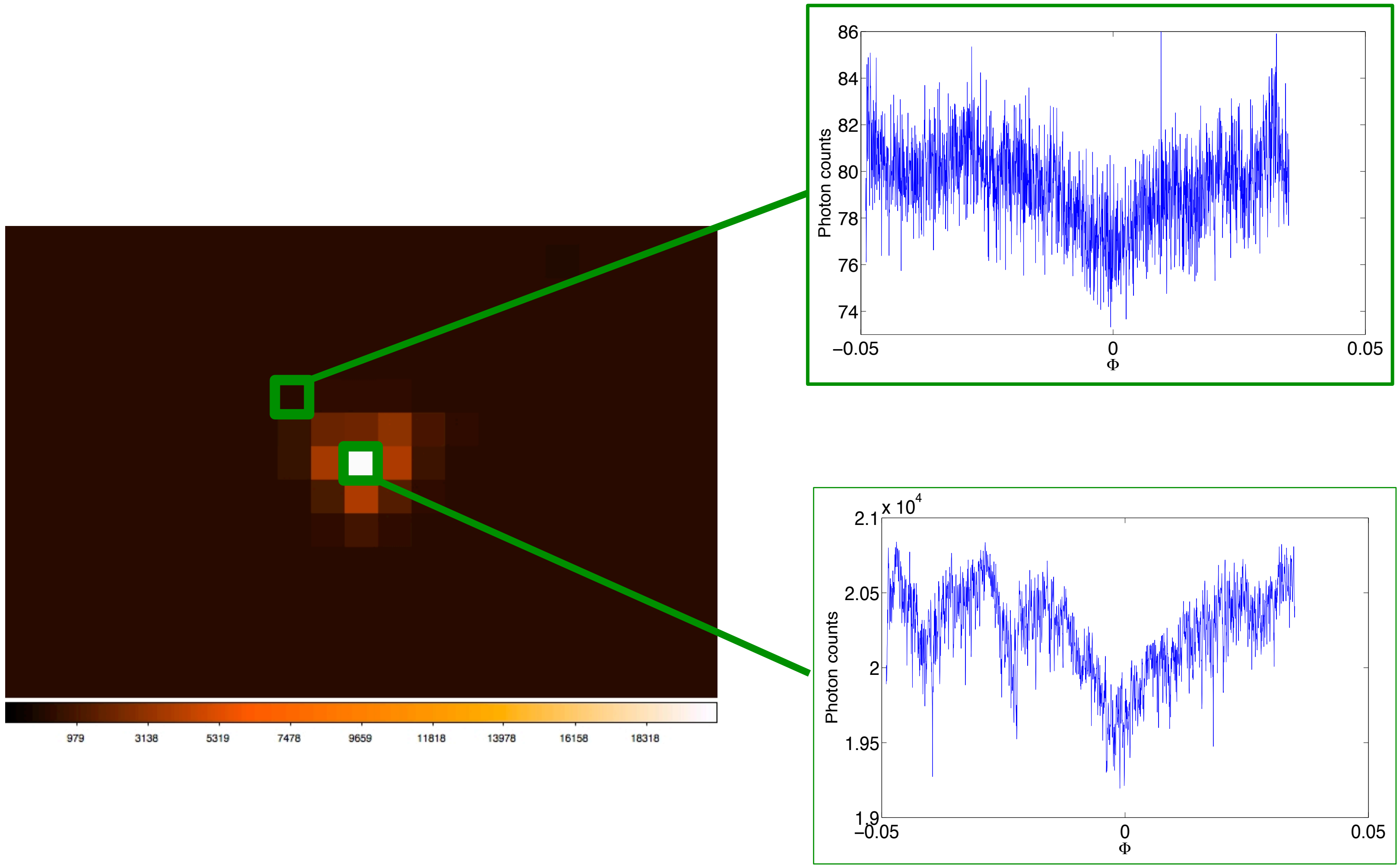
$$J(y) \approx \sum_{i=1}^{p} k_i [E\{G_i(y)\} - E\{G_i(\nu)\}]^2$$ approximated negentropy

Multiple observations

- Spectroscopically resolved light-curve, i.e. simultaneous observations in different wavele… …n et al. 2013, Waldmann 2012, 20…
- Multiple photometric observati… Waldmann 2012)
- Individual pixel-times series (e.g… Morello 2015)

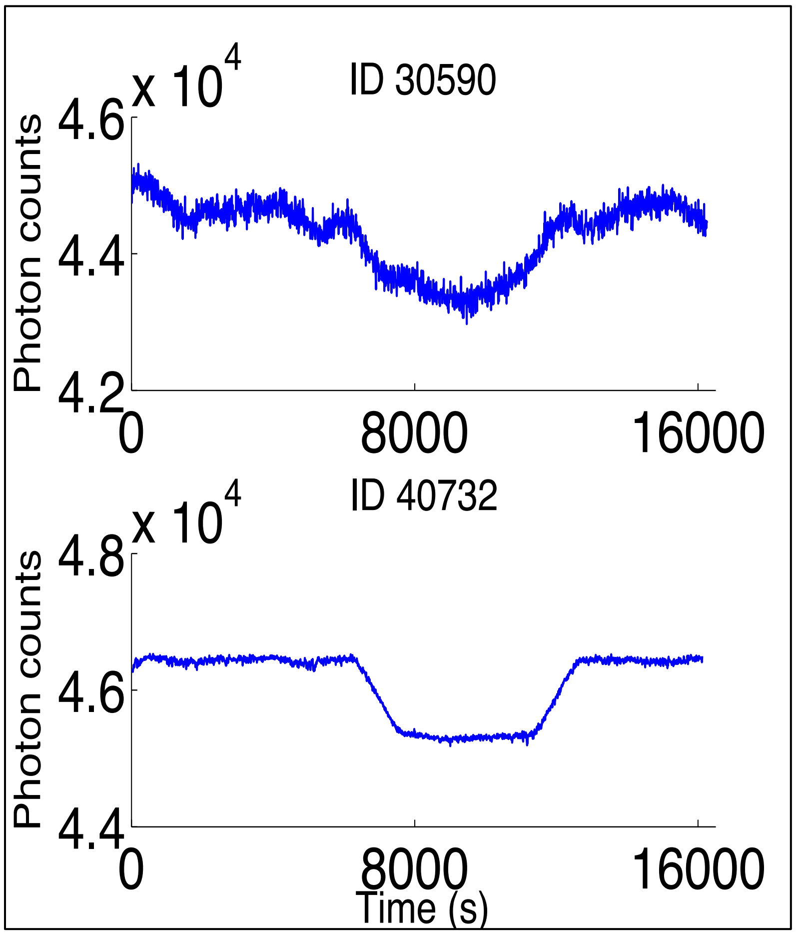


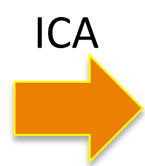
Pixel-lightcurves

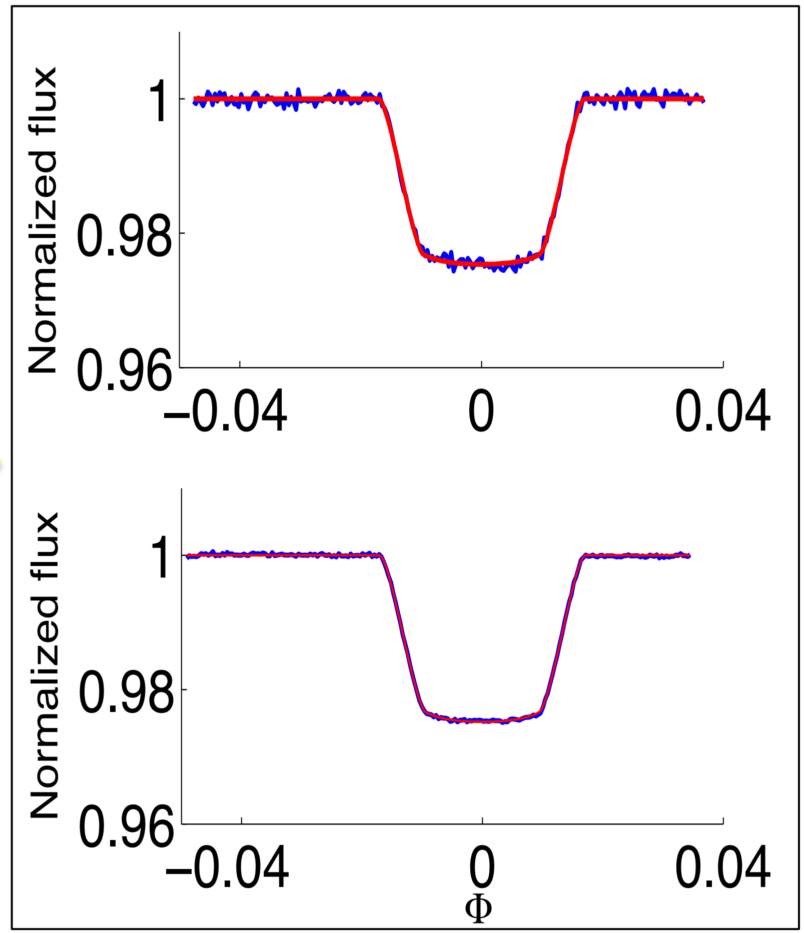


Spitzer/IRAC observations at 3.6 μm of HD189733b

Raw lightcurves | **Detrended lightcurves + models**

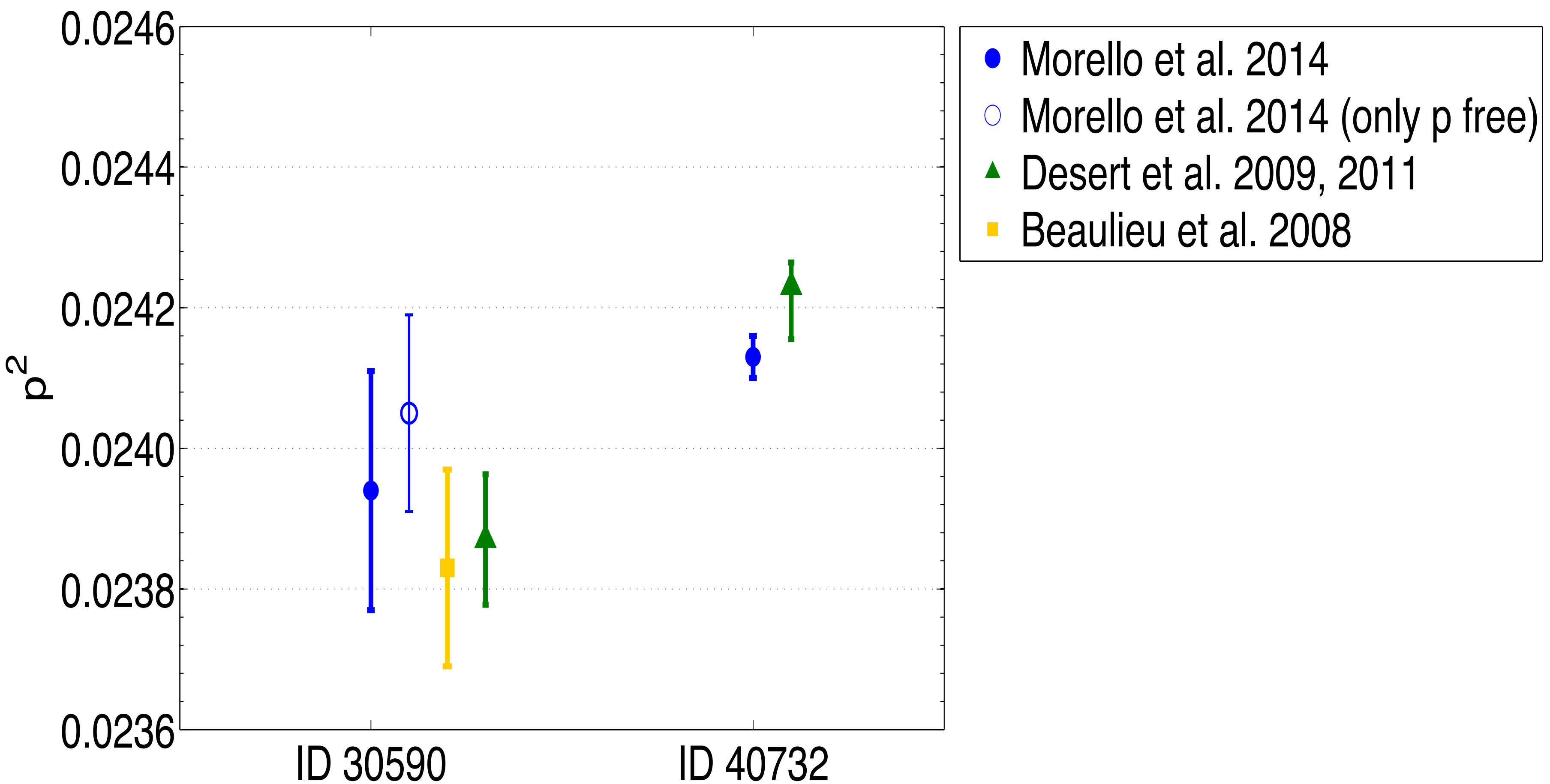






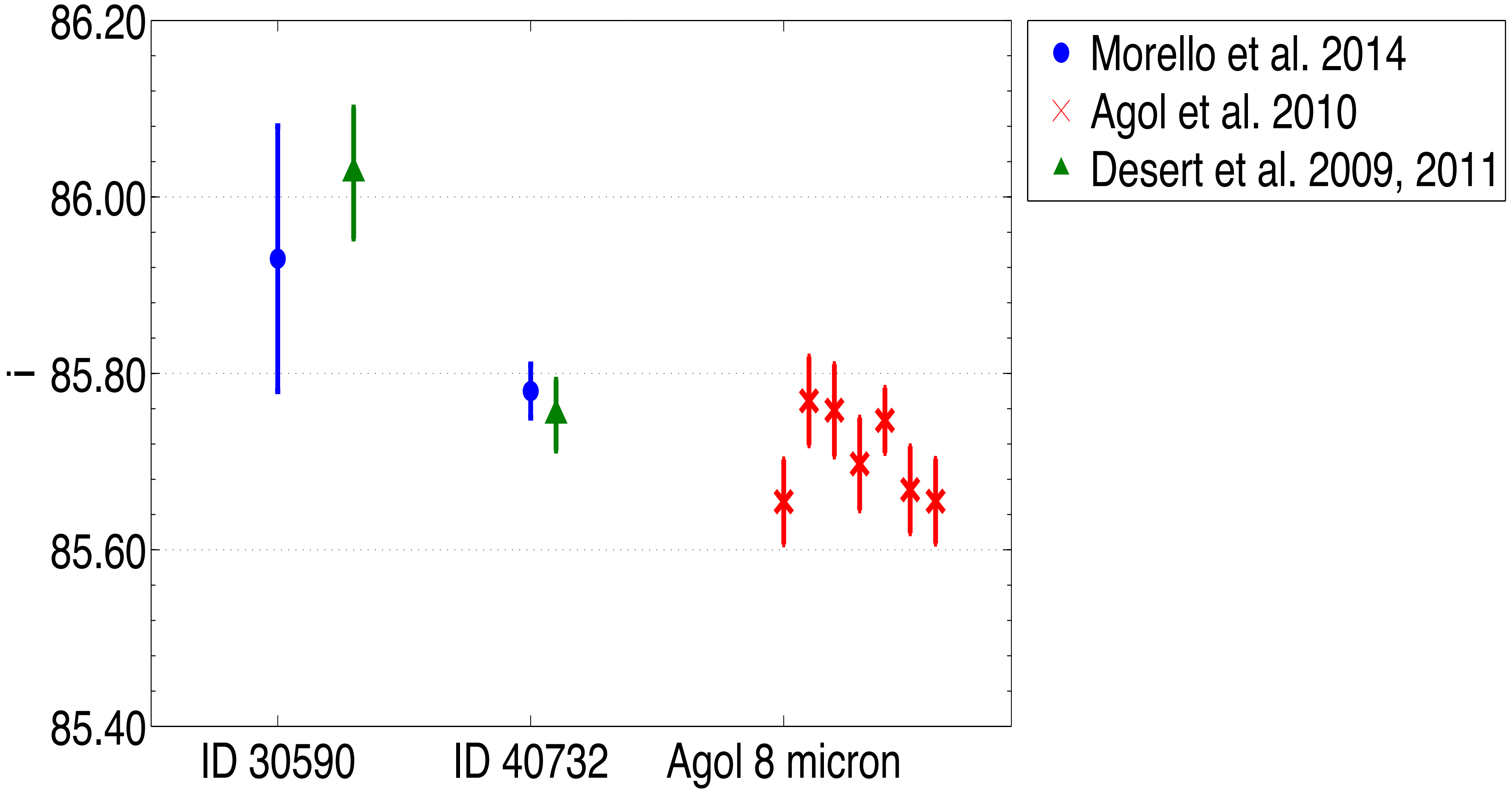
Morello et al. 2014, ApJ, 786, 22

Spitzer/IRAC observations at 3.6 μm of HD189733b - Results



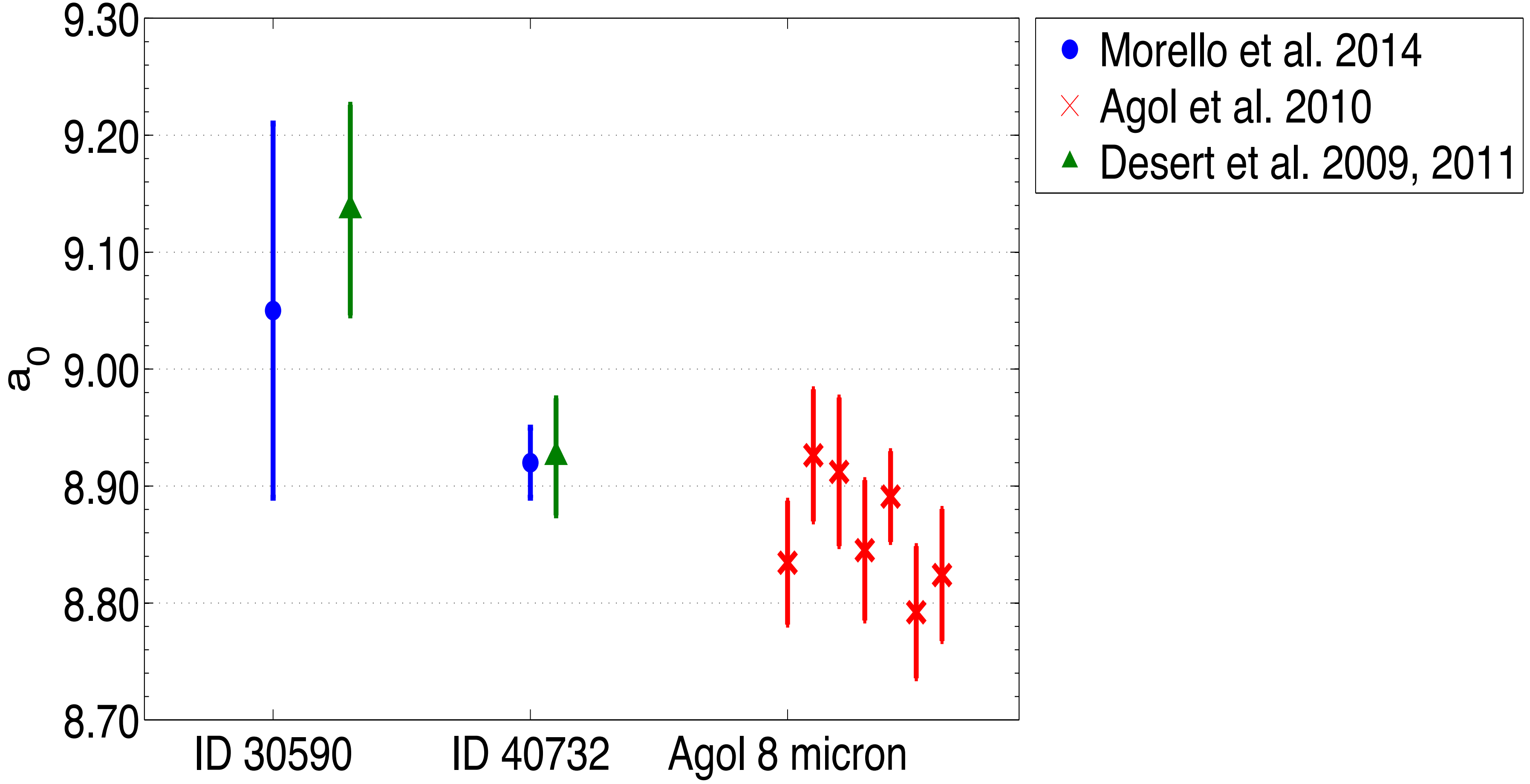
Morello et al. 2014, ApJ, 786, 22

Spitzer/IRAC observations at 3.6 μm of HD189733b - Results



Morello et al. 2014, ApJ, 786, 22

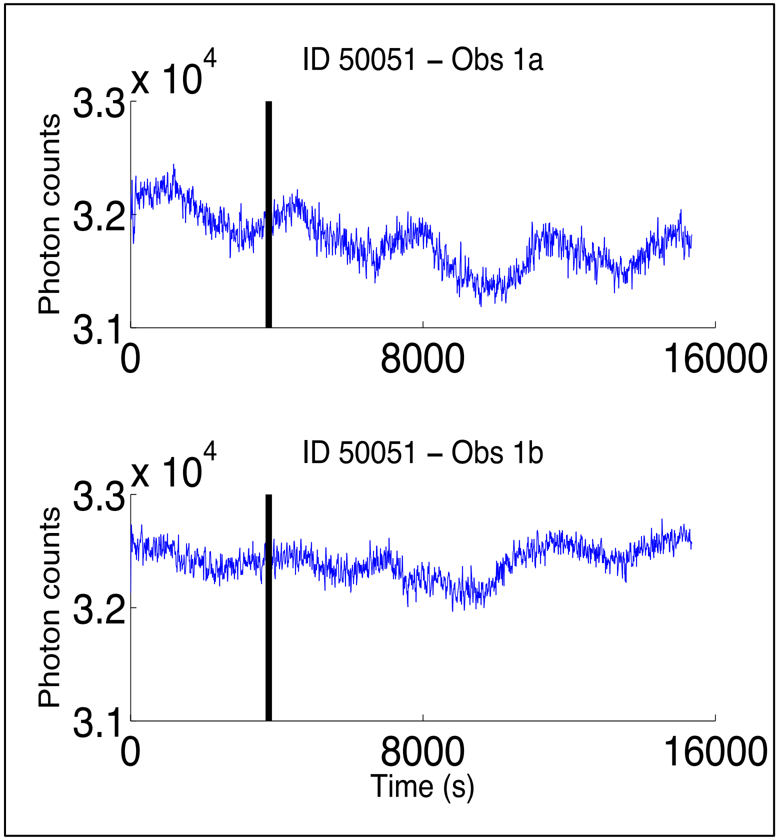
Spitzer/IRAC observations at 3.6 μm of HD189733b - Results



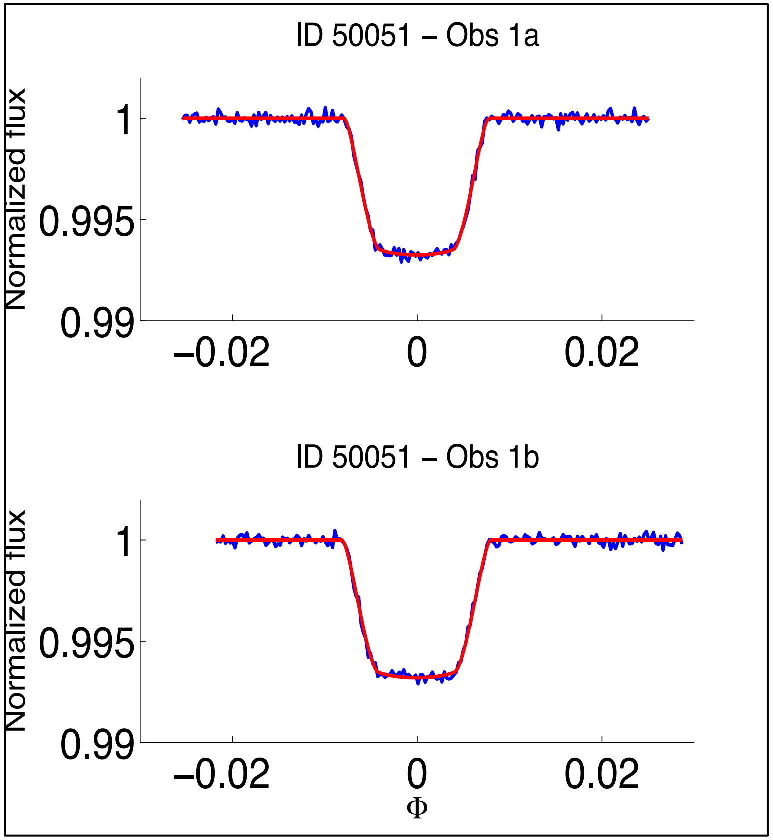
Morello et al. 2014, ApJ, 786, 22

Spitzer/IRAC observations at 3.6 μm of GJ436b

Raw lightcurves | **Detrended lightcurves + models**



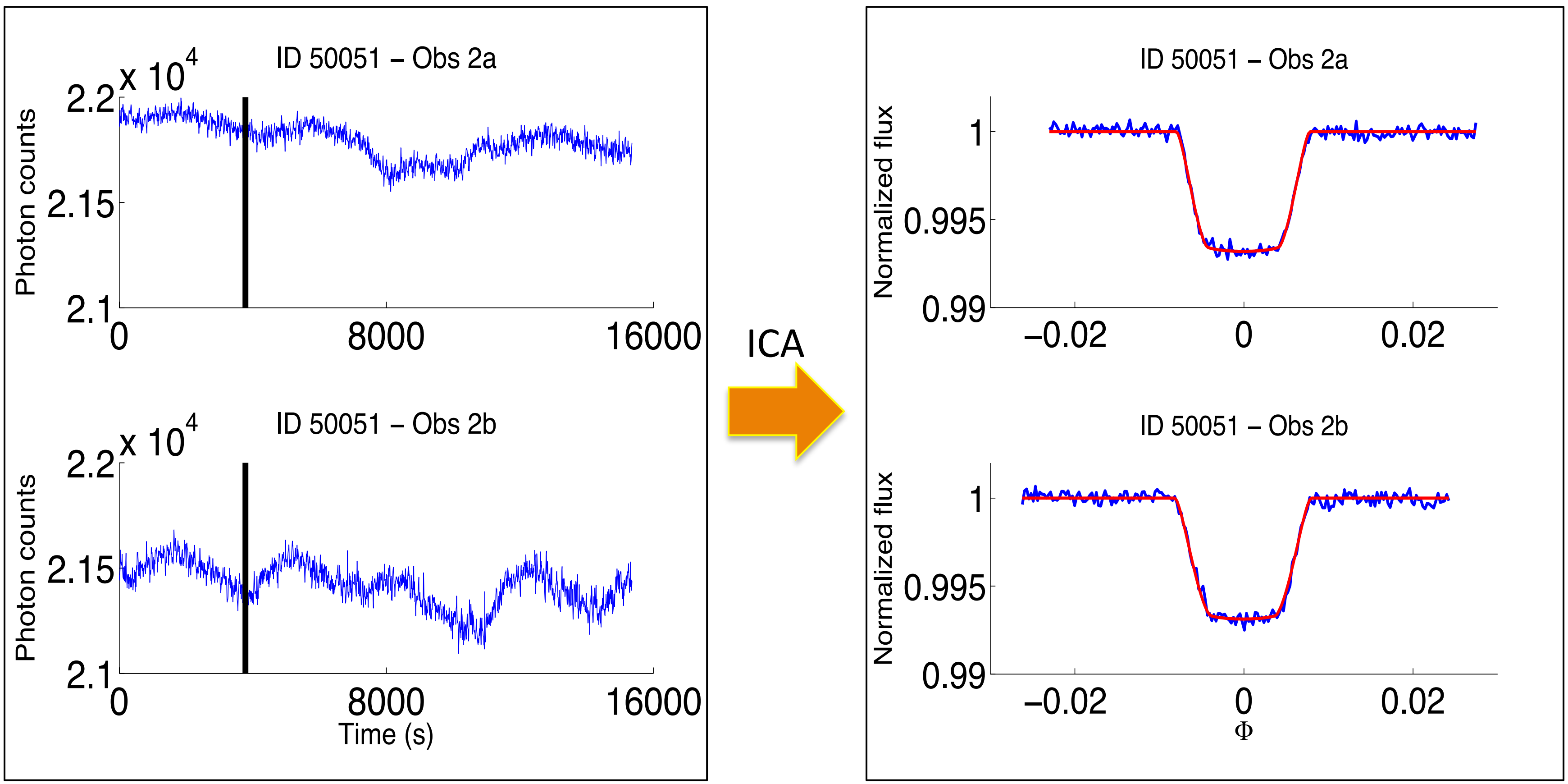
ICA



Morello et al. 2015, ApJ, 802, 117

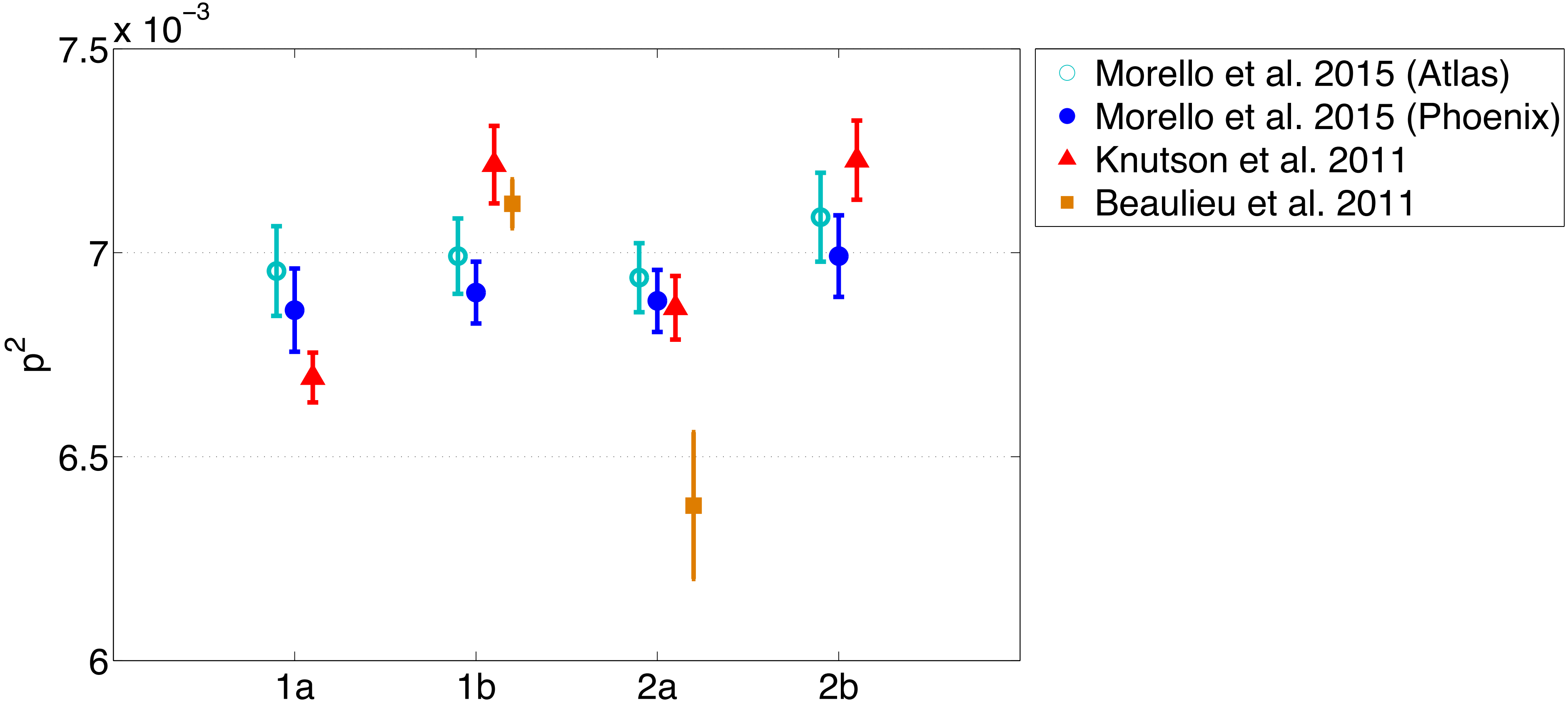
Spitzer/IRAC observations at 4.5 μm of GJ436b

Raw lightcurves | **Detrended lightcurves + models**



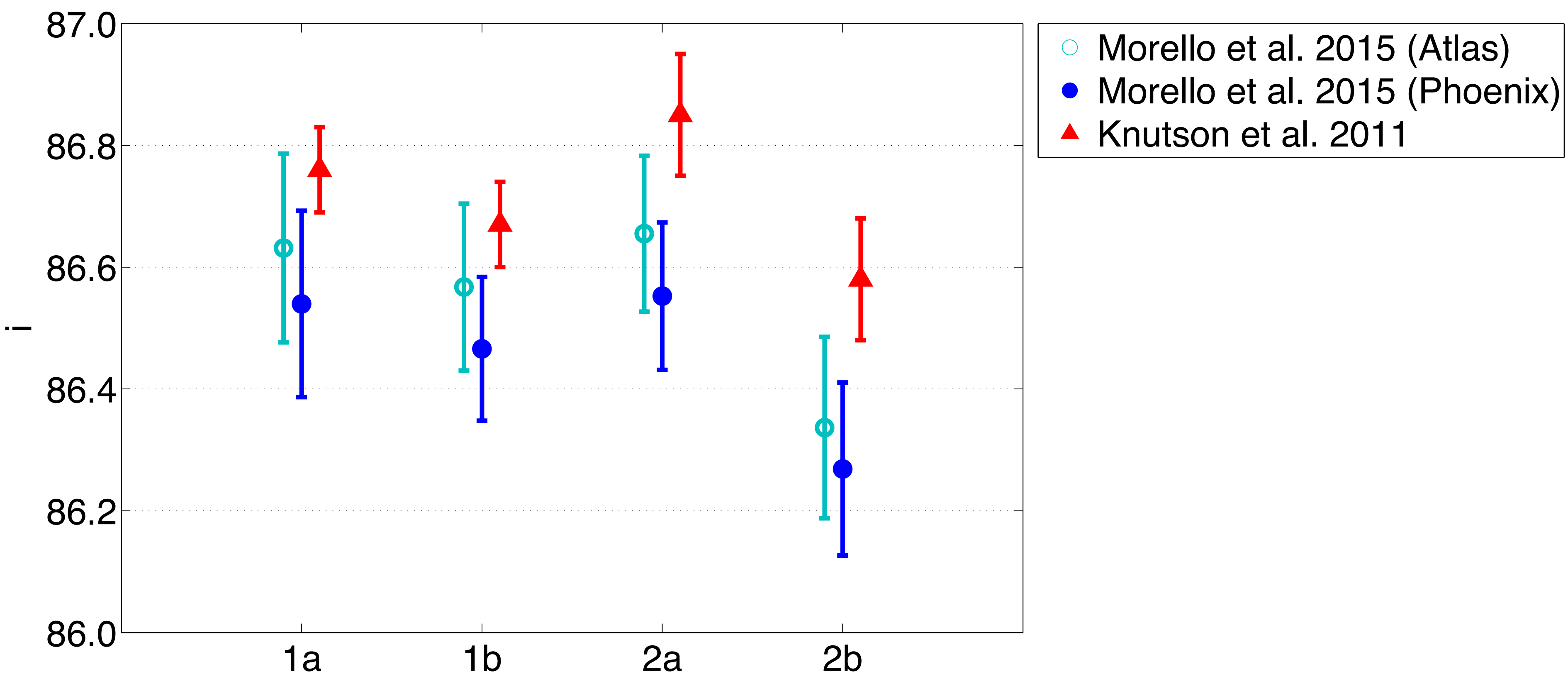
Morello et al. 2015, ApJ, 802, 117

Spitzer/IRAC observations of GJ436b - Results



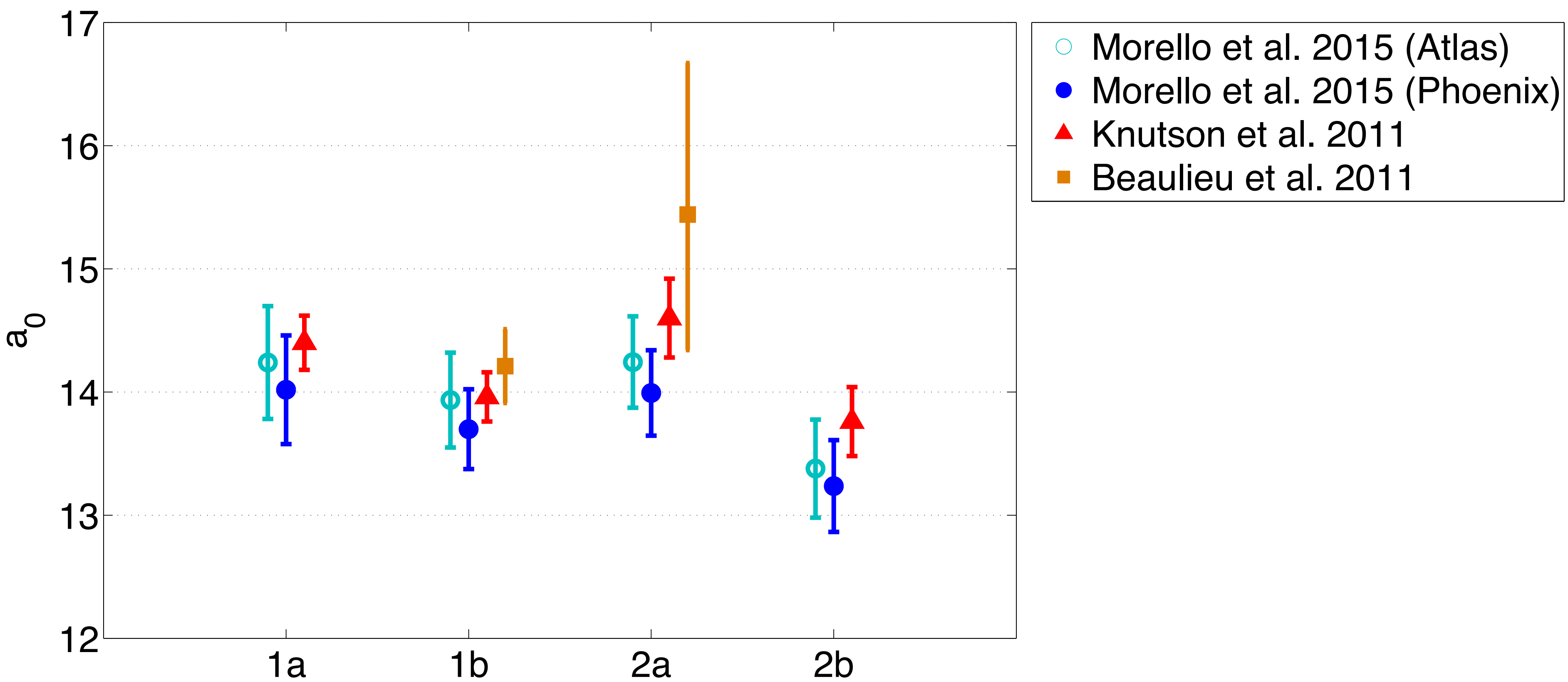
Morello et al. 2015, ApJ, 802, 117

Spitzer/IRAC observations of GJ436b - Results



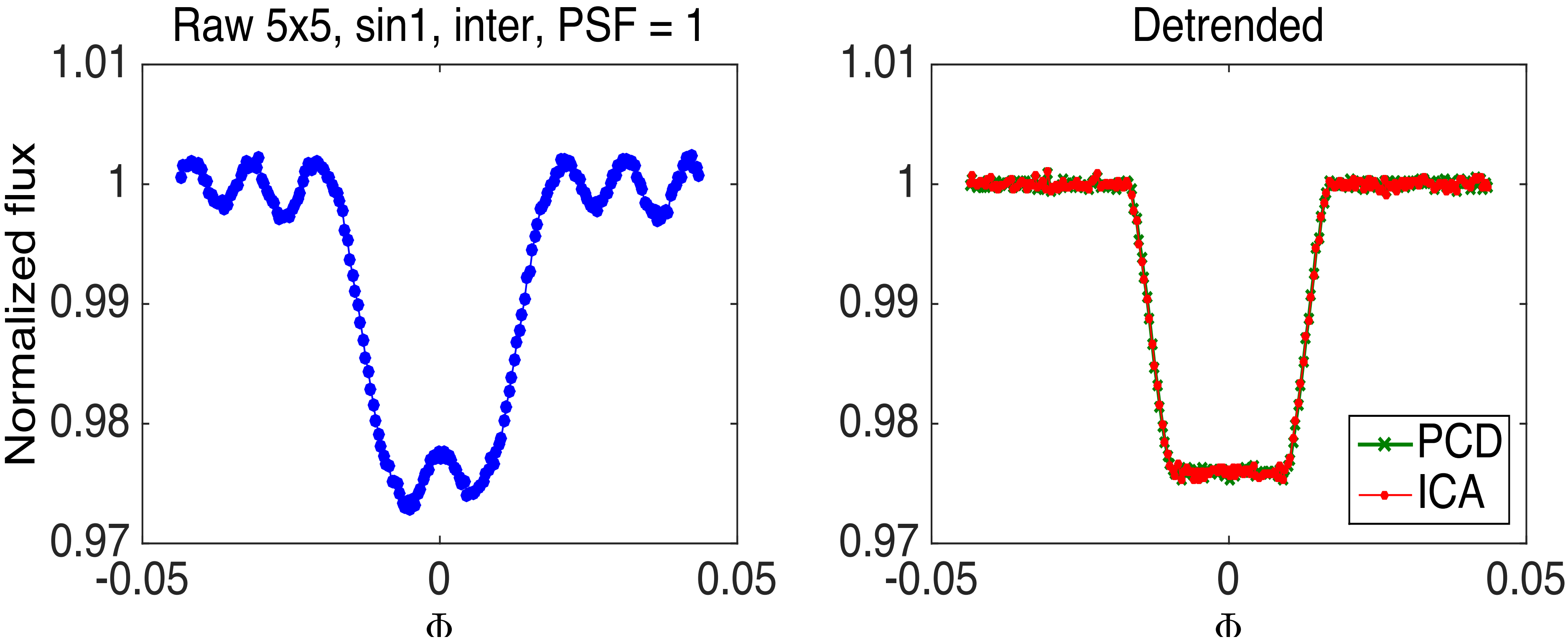
Morello et al. 2015, ApJ, 802, 117

Spitzer/IRAC observations of GJ436b - Results



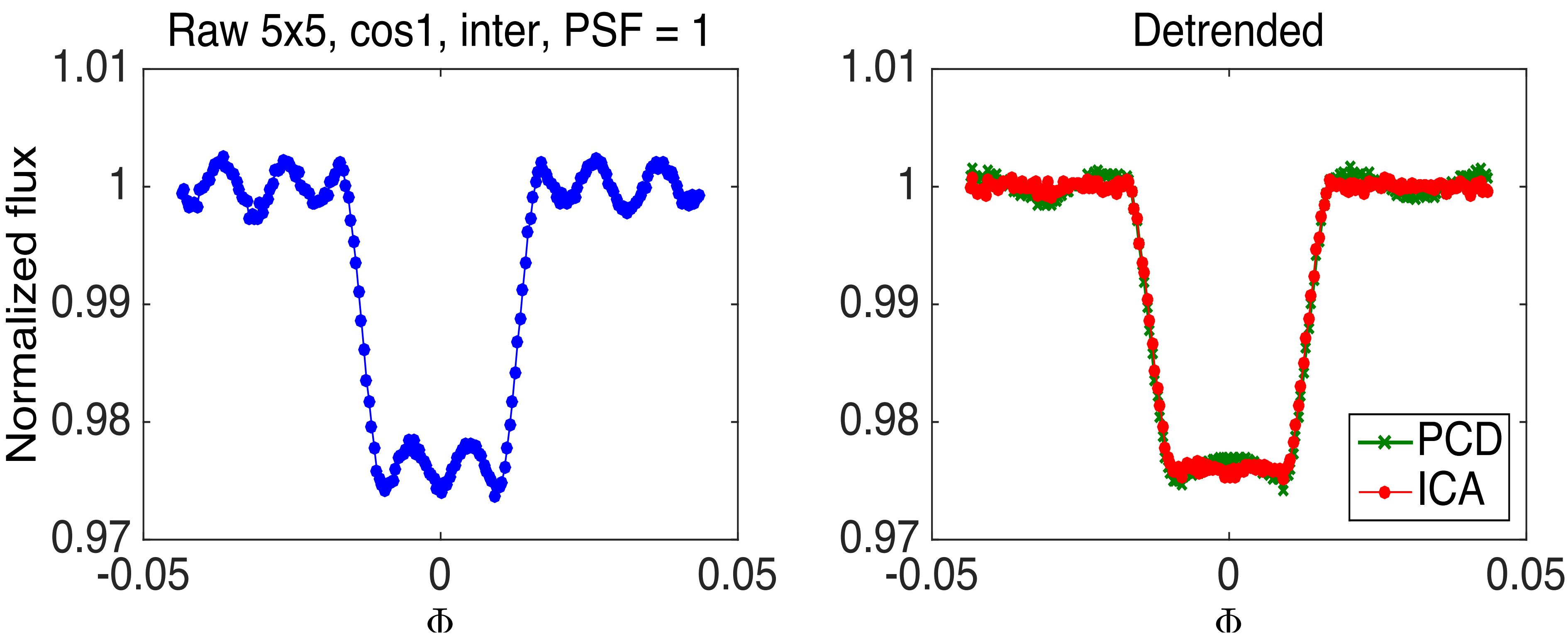
Morello et al. 2015, ApJ, 802, 117

Simulated datasets (example 1)



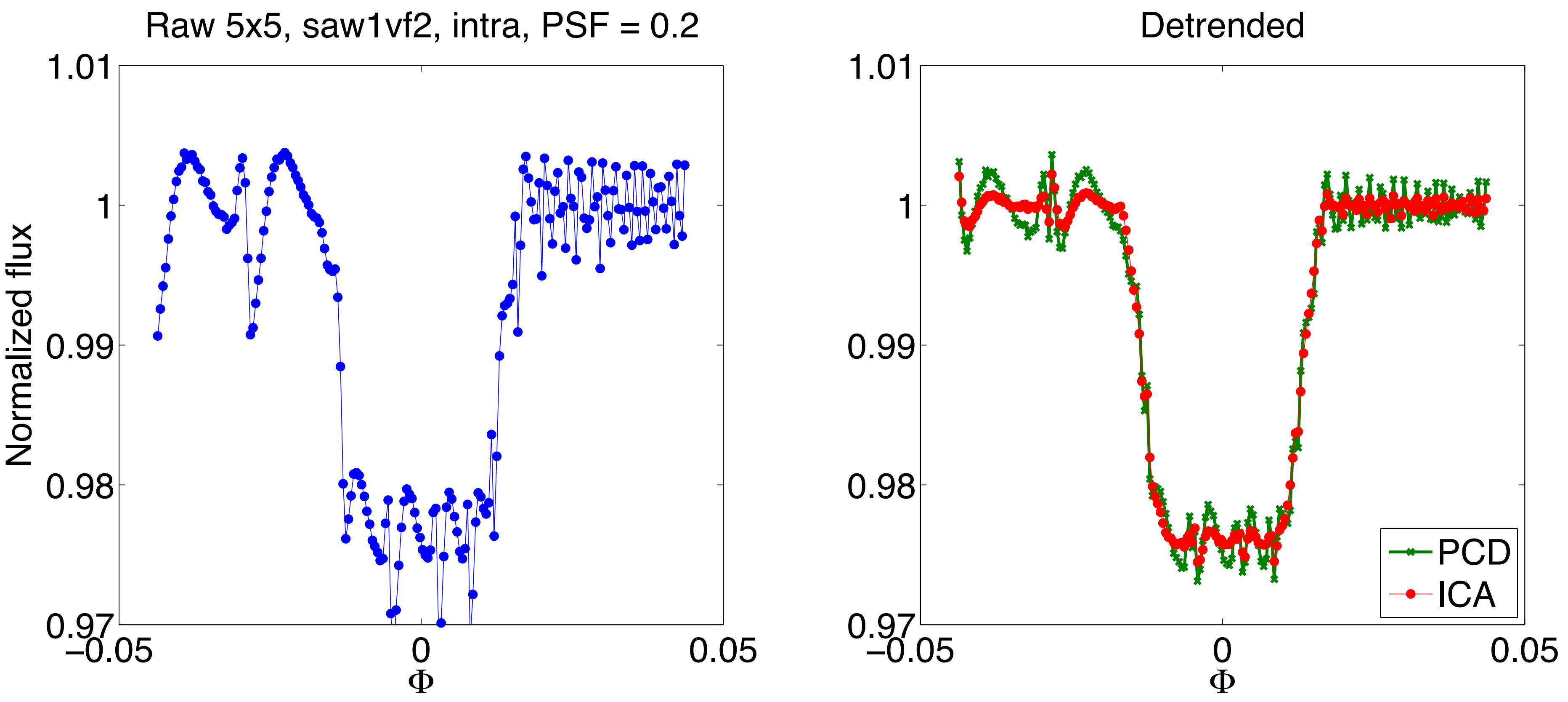
Morello 2015, ApJ, 808, 56

Simulated datasets (example 2)



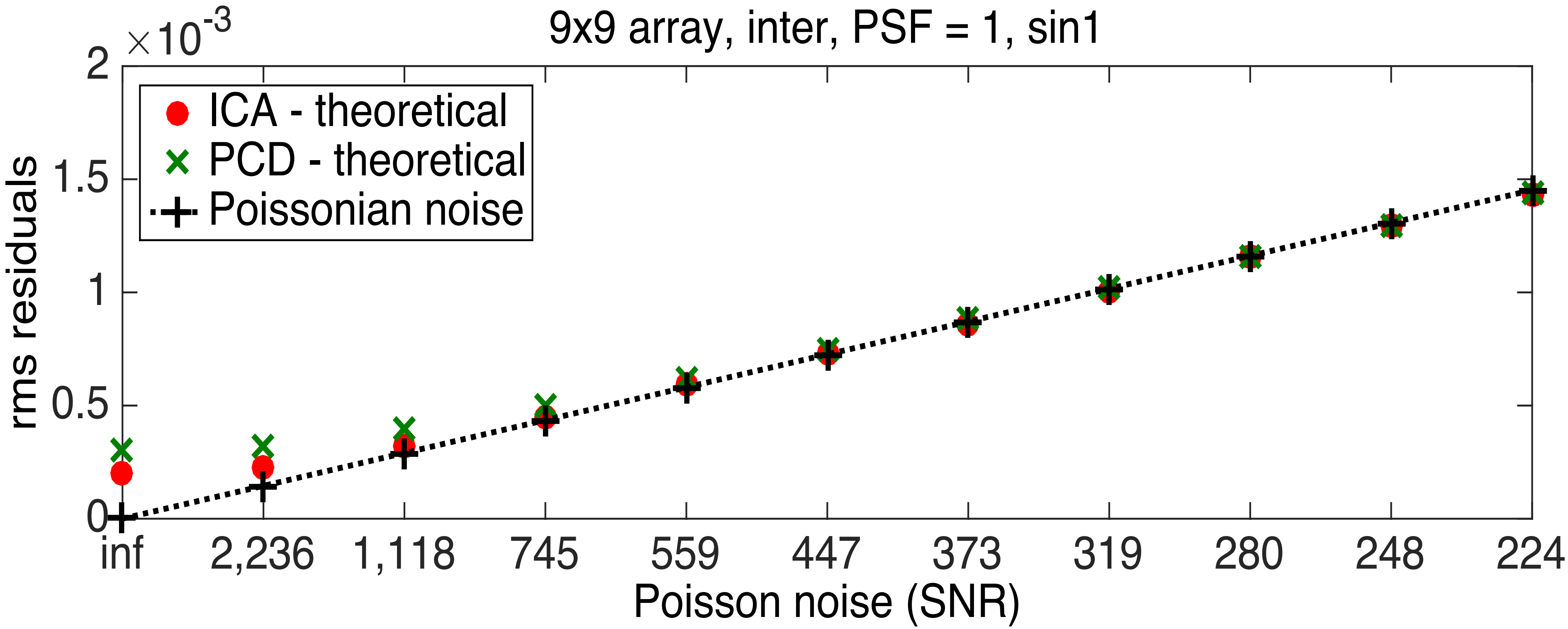
Morello 2015, ApJ, 808, 56

Simulated datasets (example 3)



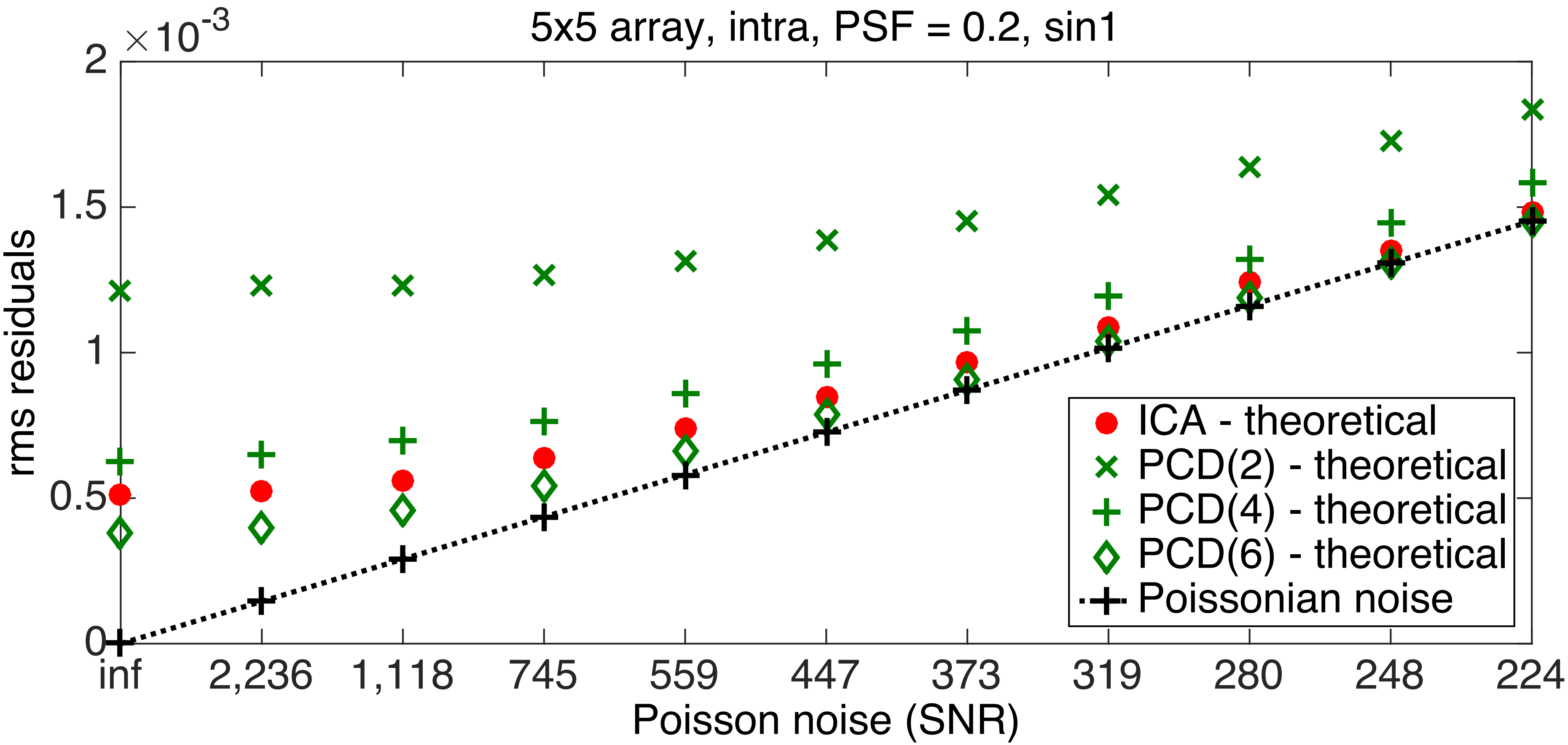
Morello 2015, ApJ, 808, 56

Effect of Poisson noise (1)



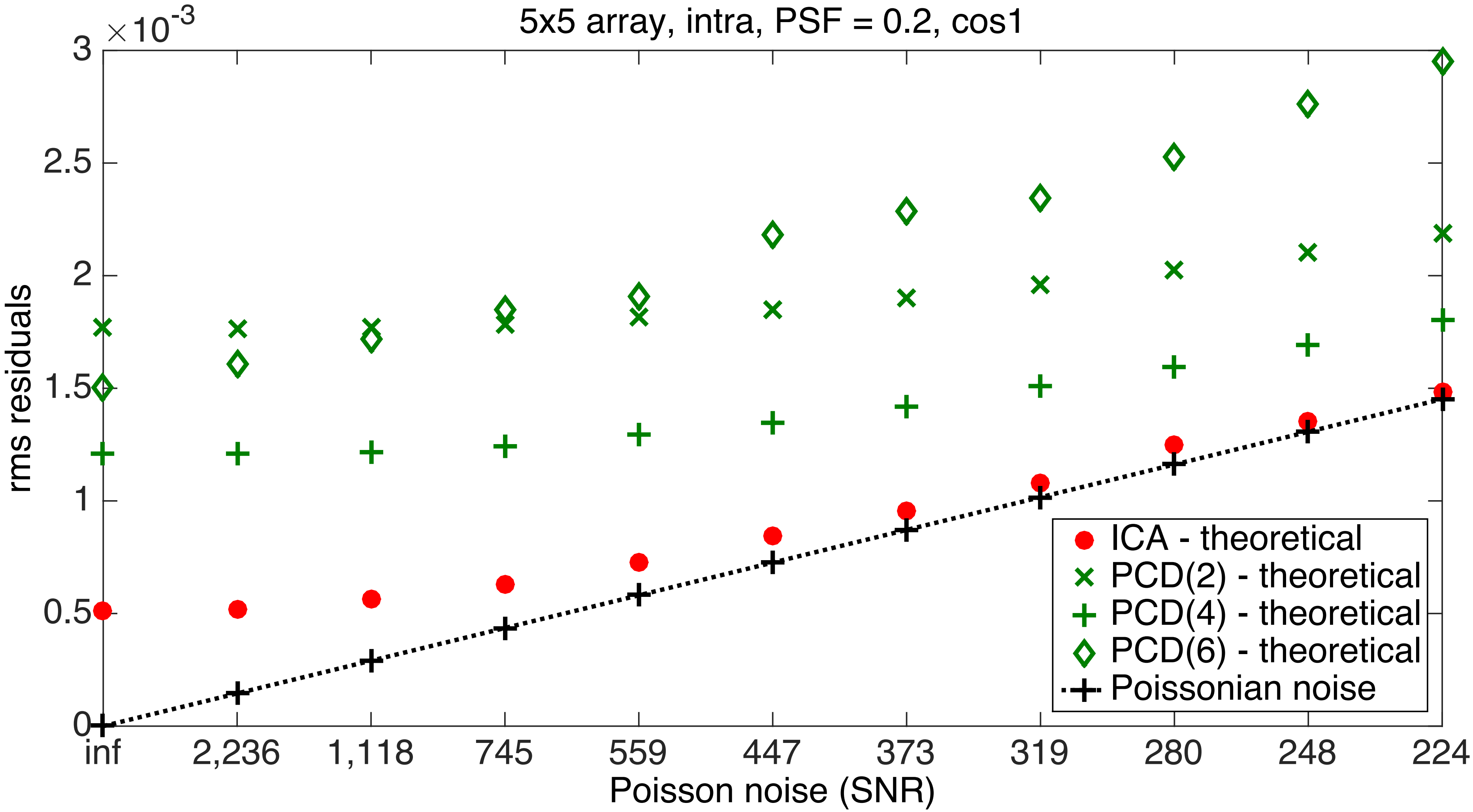
Morello 2015, ApJ, 808, 56

Effect of Poisson noise (2)



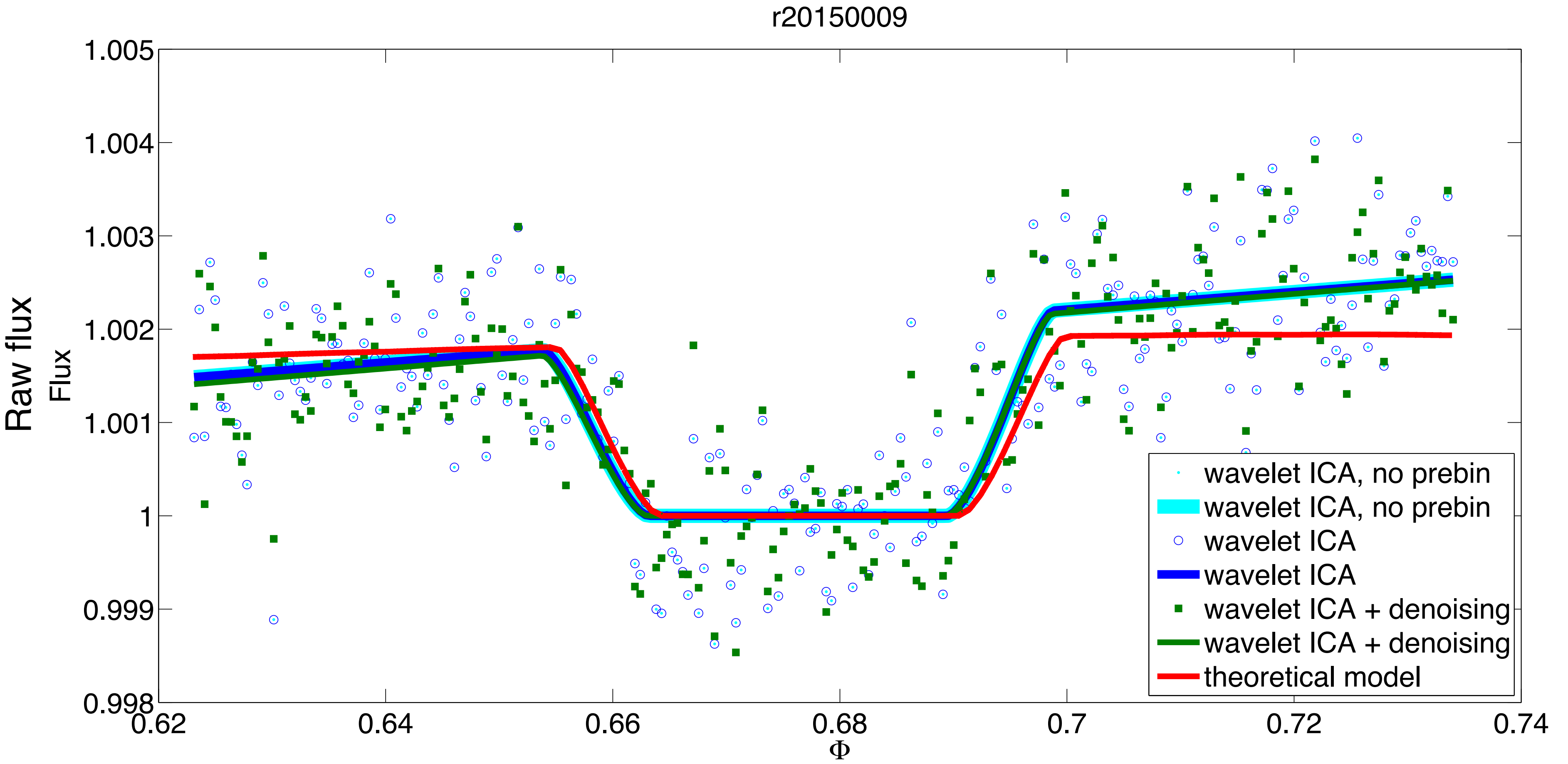
Morello 2015, ApJ, 808, 56

Effect of Poisson noise (3)



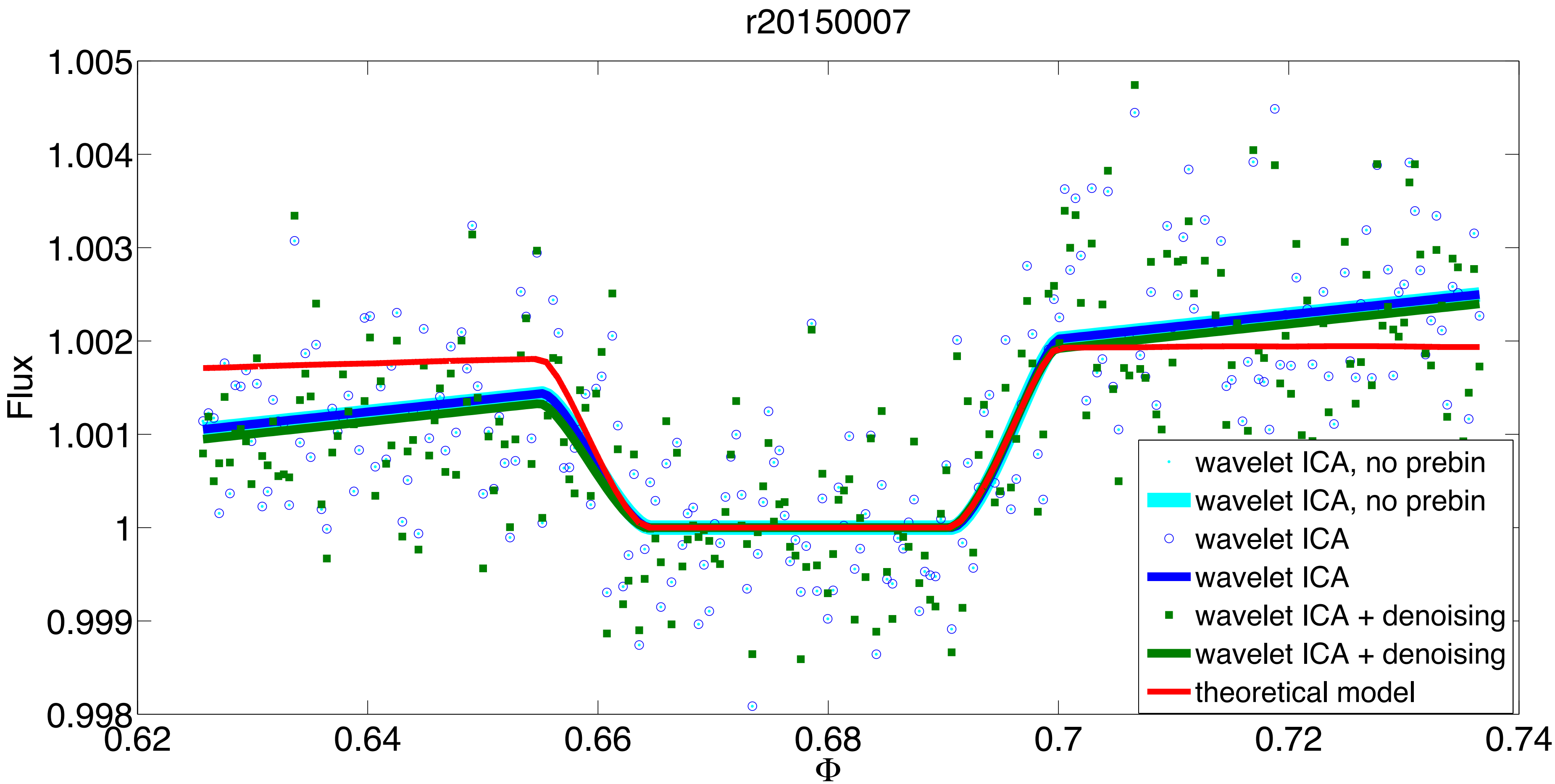
Morello 2015, ApJ, 808, 56

IRAC data challenge (1)



Measured eclipse depth: 0.00194 +/- 0.00012 True value: 0.00188

IRAC data challenge (2)



Measured eclipse depth: 0.00171 +/- 0.00020 True value: 0.00188

Conclusions

- Transit/eclipse spectroscopy can be used to understand exoplanets (composition, climate, history);
- Data detrending methods are crucial to achieve the target precision, i.e. $\sim 10^{-4}$;
- **Pixel-ICA** is a **blind signal-source separation** method to detrend systematics from a **single photometric observation**;
- Consistent transit parameter results from multiple observations;
- High performance over simulated datasets of primary transits with instrument systematics and noise.

Future projects

- Reanalysis of archive datasets;
- Improving detrending techniques through instrument simulations;
- Optimizing the method for the case of secondary eclipses (low signal-to-noise, amplitude of the astrophysical signal smaller than instrument systematics);
- Data analysis from different instruments, e.g. Spitzer/IRS, other Spitzer/IRAC passbands.

Interference-to-Signal Ratio

- If the source signals and the true mixing matrix are known, it is possible to test the goodness of the separation:

Normalized gain matrix

$$\tilde{\mathbf{G}} = \hat{\mathbf{W}}\mathbf{A}\mathbf{D}^{\frac{1}{2}}$$

Diagonal matrix of the variances of the estimated source signals

Estimated inverse mixing matrix

True mixing matrix

- In case of perfect demixing, the normalized gain matrix is the identity.

$$\mathbf{ISR}_{ij} = \frac{\tilde{\mathbf{G}}_{ij}^{2}}{\tilde{\mathbf{G}}_{ii}^{2}} \approx \tilde{\mathbf{G}}_{ij}^{2}$$

- For certain algorithms, it is possible to calculate asymptotical expressions for the **ISR** matrix, which are independent on the mixing matrix.

Error bars

$$\sigma_{par} = \sigma_{par,0}\sqrt{\frac{\sigma_0^2 + \sigma_{ICA}^2}{\sigma_0^2}}$$

$$\sigma_{ICA}^2 = f^2\left(\sum_j o_j^2 \mathbf{ISR}_j + \sigma_{ntc-fit}^2\right)$$

MULTICOMBI:

$$\mathbf{ISR} = \frac{\mathbf{ISR}^{EF} + \mathbf{ISR}^{WA}}{2}$$

$$\boxed{\mathbf{ISR}_{i,j} = min\left(\mathbf{ISR}_{i,j}^{EF}, \mathbf{ISR}_{i,j}^{WA}\right)}$$