Exoplanet image postprocessing

Principal Investigator: Graça Rocha (326)

Michael Bottom (383A), Jeff Jewell (398L), Dimitri P. Mawet (326, Caltech), Garreth Rouane (383A)

Jacob Golomb (UMD)

Program: Strategic Initiatives

Project Objective:

The objective of this research work is to help improving the detection and characterization of exoplanets in high-contrast imaging data.

Work proposed: We proposed to develop the first fully Bayesian methodology, based on Powellsnakes, PwS, and a post-processing technique, KLIP.

Powellsnakes - is a Bayesian approach and algorithm co-developed by the PI for detecting compact objects embedded in a diffuse background. This algorithm has been applied successfully to Planck satellite.(Carvalho, Rocha & Hobson 2009; Carvalho, Rocha et al 2012).

KLIP – is the Karhunen-Loeve , KL, Image Processing technique to do stellar Point Spread Function, PSF, subtraction (ie speckle subtraction) in direct imaging of exoplanets. It uses Principal Component Analysis (PCA) to identify the modes of highest variance from reference images of the host star, and subtracts these modes from the target image to remove the stellar PSF and make a potential astrophysical signal (eg planet) more visible. (Soummer, Pueyo, & Larkin 2012).

pyKlip is a python library for direct imaging of exoplanets (Wang et al 2015, 2016). It uses an implementation of KLIP and KLIP-FM to perform point spread function (PSF) subtraction.

KLIP Forward Modeling (**KLIP-FM**) (Pueyo 2016) a companion algorithm to KLIP method, uses an input instrumental PSF at a given location and feeds it into the KLIP algorithm, thus resulting in a model instrumental PSF that can be corrected for biases due to PCA subtraction. However, KLIP-FM allows for accurate astrometry of a potential object, *given the precondition that the point source has been detected by eye in the image.*

We explored one of two ways to extend pyKLIP to Bayesian (blind) detection and characterization: (i) Starts with the coadded frames (from the set of temporal and wavelength frames) and after subtracting the stellar PSF (speckles); (ii) Starts with a joint analysis of all frames in time and wavelength. We completed approach (i) while approach (ii) will be completed in the near future.

Work accomplished:

We extended pyKLIP to a fully Bayesian blind detection step following PwS approach. To this end we developed a new algorithm **PlanetEvidence** and integrated it into pyKLIP. This step involves integration of a nested sampling, Multinest, (Skilling 2004, Feroz & Hobson 2008, 2009, Buchner et al 2014 for PyMultiNest) for detection and characterization of the planet (task completed).

PlanetEvidence is integrated in the pyKLIP library and publicly available here https://pyklip.readthedocs.io/en/latest/

The corresponding paper by Golomb J., Rocha G., Meshkat T., Bottom M., Mawet D., Mennesson B., Vashisht G., and Wang J., 2019, in preparation, will be submitted to the journal soon.

The Next step is a blind detection step, **pyPwS.** The module **pyPWS** will be integrated in the pyKLIP library (task in development).

FY18/19 Results:

Procedure:

Following (Soummer 2012) we estimate the Karhunen-Loeve, KL, modes (up to some pre-set number) and subtract them from the image. Figure 1. shows a PSF-subtracted image of beta Pictoris (left hand side), where, after subtracting 2 KL modes, the planet beta Pictoris b is visible in the image. Next we move onto forward modeling with KLIP, we follow the methodology of Wang et al 2016, with the residuals defined as: $R(x_p, y_p, \alpha) = (D - \alpha F(x_p, y_p))_{\mathcal{F}}$ Where D is the data (value of a pixel) and F is the Flux according to the instrumental PSF model at (x,y). We use α as a scale parameter, to allow it to vary to scale up or down the flux of the PSF and minimize the residuals. The likelihood function is a correlated Gaussian of R, the natural log of which is given by:

$$ln\mathcal{L} = -\frac{1}{2}(R^T C^{-1} R + ln(detC) + N_{pix} ln(2\pi))$$

Where the covariance matrix, C, accounts for the residual correlated noise in the image. Motivated by Czekala et al. 2015 we use the Matern covariance function with $\vartheta = 3/2$. This has the form:

$$C_{ij} = \sigma_i \sigma_j (1 + \frac{\sqrt{3} r_{ij}}{\ell}) e^{\frac{-\sqrt{3} r_{ij}}{\ell}}$$

Where r_{ij} is the distance between the i^{th} and j^{th} pixels and l is the correlation length scale, the expected size of the spatial correlations in the image. We treat l as a parameter, but the expectation for l is $\frac{\lambda}{D} \approx 3$, in our case. In addition we adopt non-informative priors for the model parameters probability distributions. For *detection* and *characterization* in the image we consider *α, l, x*, and *y*, as the parameters of interest. We sample the model parameter values with MCMC and Multinest. The latter has several advantages including overall speed of computation as well as giving the overall "evidence" (average of the likelihood over the prior) values for models being tested.

We run PlanetEvidence in Gemini Planetary Imager (GPI) data of beta Pictoris b. With the GPI data the instrumental PSFs are modeled empirically. Previously we had performed consistency checks running MCMC and obtained results consistent with Pueyo et al 2016. To assess whether the detection is more likely a point source than background, we consider two models: H_1: a planet is present in the image, and H_0 (null hypothesis): only noise is present in the image. For H_1 model, we simply use the Likelihood defined above. For H_0 model we set α = 0 and get the distributions for the remaining 3 parameters. With Multinest we compute the evidence of each model (Z) and the evidence ratios (Z_1/Z_0) which will tell us whether the data favor H_1 or H_0 . Figure 1. (2nd panel) shows the results from PlanetEvidence runs on the location of beta Pictoris b for both models. Next we inject planets at various flux levels into the image, and see how confidently we can decide they are point sources rather than background noise, as shown in the 3rd panel of Figure 1 and Figures 2. and 3.

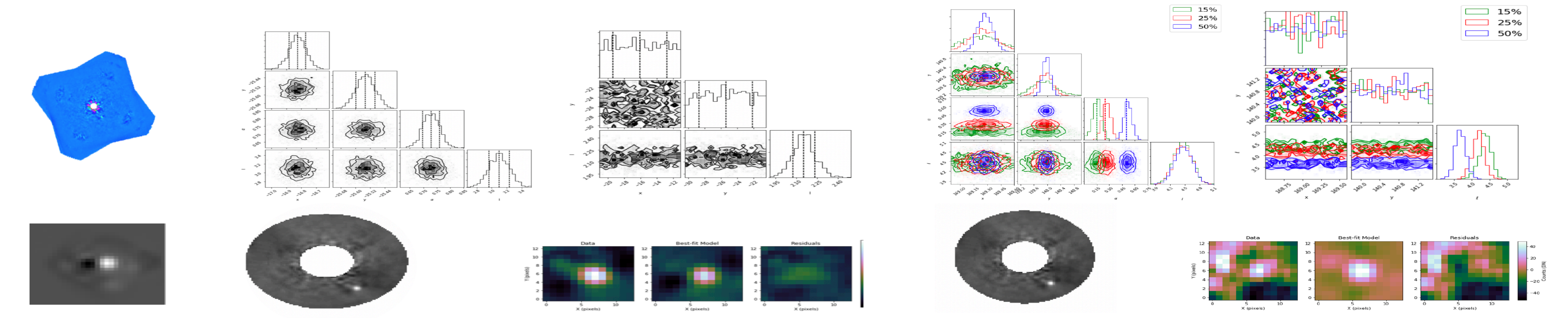

Figure 1. From left to right: (Top) A 2 KL-mode subtracted image of beta Pictoris (beta Pictoris b is located to the bottom right of the coronagraph). (Bottom) A GPI instrumental PSF forward modeled in the location of beta Pictoris b. 2nd panel: (Top) PlanetEvidence runs for both H_1 and H_0 models, the corner plot of (x, y, α, l) and (x, y, l) for H_1 and H_0 respectively (from left to right) on the location of beta Pictoris b, (Bottom left) the KLIP-subtracted image of beta Pictoris b, (Bottom right) shows from left to right: the data, the best fit model and residuals; with $\frac{Z_1}{Z_0} \approx 2 * 10^{76}$ i.e. very strong evidence for H_1 over the null hypothesis (as expected). 3rd panel: Same as 2nd panel but this time for for a planet injected at the 3 o'clock position with 50% (blue), 25% (red) and 15% (green) of the Forward Model flux (FMF).

In other words, we have quantified the confidence with which we accept the point-source detection and reject the null hypothesis. As the flux of the injected planet decreases (in a fixed location), so does the confidence with which the decision can be made. Dim sources with fluxes of the order SNR_{an} ~2σ can be detected and characterized as true point sources rather than background noise. This is very encouraging as it allows us to lower the detectability threshold to ~ 2σ.

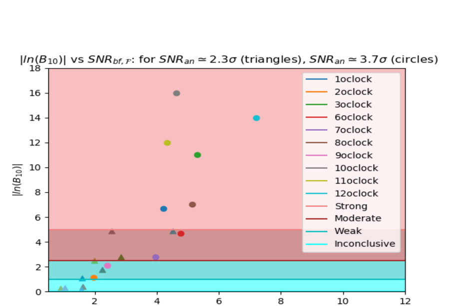

Figure 2. Evidence vs SNR: SNR estimated in the annulus, SNR_{an}, after masking beta Pictoris b and the injected planet; and SNR estimated in the fitting area, $SNR_{bf,F}$ with noise estimated after subtracting the best fit model from the fitting area. The horizontal lines represent the threshold values which are empirically set, and they occur for values of the logarithm of the Bayes factor of |lnB10| = 1.0, 2.5 and 5.0. Shaded areas represent the different levels of evidence above these thresholds Trotta (2008).

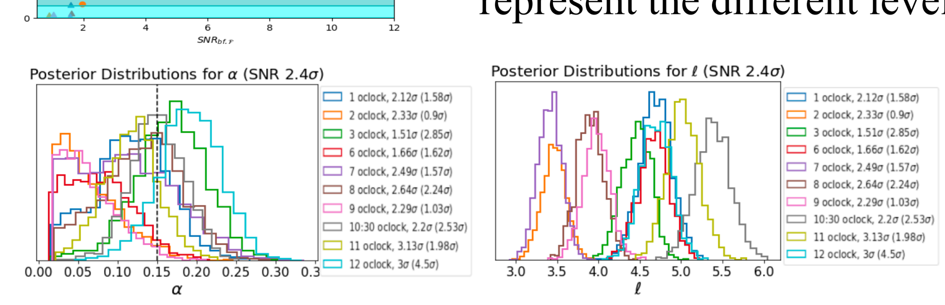

Figure 3. Left: Posterior distributions of parameter α ; Right: Posterior distributions of parameter l, for all injected sources (ie for the several positions on the image) for $SNR_{in,an}$ ~2.4 sigma; 15% of the input FMF.

Finally we tested this approach by running PlanetEvidence on "noise", that is, when no synthetic planet is injected into the data. We find that seven out of the ten cases considered hold B_{10} ~ 0.3, using Harold Jefreys scale interpretation for B_{10} this indicates that Evidence supports the null hypothesis, H_0 (Harold Jeffreys 1961). This is a reassuring result.

Discussion and Conclusions:

These results demonstrate that applying a Bayesian evidence-based method along with the post-processing technique, KLIP, as implemented in our algorithm PlanetEvidence, allows the detectability of dim sources (barely or not visible by eye after speckle subtraction). Furthermore our method allows to quantify the confidence with which we reject the null hypothesis. However to fully reach such conclusion we need to prove the false positive rate is low enough to confidently say they are planets. This is the matter of future work using noise only simulations. PlanetEvidence in conjunction with the blind detection step (in development) will prove to be a powerful tool for detection and characterization of planets in direct imaging data. This novel method will greatly help the (automated) data-analysis of future experiments with heavy JPL involvement.

Benefits to NASA and JPL (or significance of results):

This novel method developed in this study will greatly help the (automated) data-analysis of future experiments with heavy JPL involvement such as WFIRST, HabEX, JWST, LUVOIR, etc. It will ultimately help enhancing the science rendered by these missions.

Publications:

Golomb J., Rocha G., Meshkat T., Bottom M., Mawet D., Mennesson B., Vashisht G., and Wang J., 'Planet-Evidence: is it a planet or just noise?', 2019, in preparation.

Graca Rocha/R.18.183.023

Contact Information: 818-393-0095, Graca.M.Rocha@jpl.nasa.gov

Poster No. RPC-042

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

www.nasa.gov