Starshade Exoplanet Data Challenge

Brian Dunne – 4/30/2022

1 Introduction

This report details the methods, results, and conclusions derived from participating in the Starshade Exoplanet Data Challenge. The goals of the data challenge require solutions to multiple subproblems, where each subproblem is interesting and challenging in its own right. For some of the harder subproblems, multiple methods were explored, where some were fully implemented, some partially implemented and some considered in detail but not implemented. This report aims to explain the learnings derived from these various attempts and to make recommendations for future work. The report will also attempt answer some of the original questions posed by the data challenge, but with the caveat that the obtained results are likely far from optimal, and better results are likely possible given more time and effort. While the results may help inform requirement flow down in regards to instrument calibration and noise limits, it is likely that the present results are overly pessimistic relative to what is possible with more time, more training and test data, and more clever approaches. This report intends to be a useful contribution for informing future work while also addressing some of goals set out for the data challenge.

2 Methods

A majority of the time for this effort was spent on developing approaches for analyzing the data in the Release #2 data: Broadband Imaging Simulations for Starshade Rendezvous. This data release consists of 1440 simulated images, representing various imaging scenarios. In general a given simulated image includes some relatively faint signal from an exoplanet, in additional to simulated background signal from various noise sources: residual starlight, solar glint and other stray light sources, exozodiacal light, detector noise, and variability resulting from the starshade's motion and telescope jitter.

While it is possible to formulate the image analysis problem different ways, for the purposes of this report, the approach is broken into three sub-problems:

1. Background Estimation
2. Planet Detection
3. Planet Parameter Estimation

A majority of the effort was spent on the background estimation and removal sub-problem which proved to be the most challenging, and the most critical limiting factor for the ultimate accuracy of detection and estimation results.

2.1 Background Estimation

The background estimation problem proved to be the most challenging subproblem, and significant time was spent trying different approaches. The approaches are divided into two general categories: parametric and non-parametric.

The parametric approach attempts to explicitly model the backgrounds, find best fit parameters, and then subtract the best fit background model. The non-parametric category includes approaches which do not use an explicit astrophysical or otherwise explicit mathematical model of observed backgrounds, and instead attempt to exploit underlying differences in the behavior of planet signals vs background signals when considering a large training set. Specifically, three non-parametric approaches were explored: PCA, ICA, and an Autoencoder. These three approaches all attempt to find a low-dimensional latent representation of the background signals.

2.1.1 Parametric Background Estimation

After inspecting the Release 1 data, which only contained smooth and radially symmetric exozodi, it seemed appropriate to construct a simple parametric model that would be suitable for nonlinear parameter estimation. The idea is to construct a model which produces a background image I_b given some parameter vector p:

$$\mathrm{I}_p = f(p)$$

Once such a model has been designed, background estimation can be performed via nonlinear optimization to find the parameters p which best explain the background observed in a given image, according to some cost function C:

$$\hat{\boldsymbol{p}} = \arg_p \min C(\mathrm{I}, f(p))$$

Because $C(\mathrm{I}, f(p))$ must be minimized with a nonlinear optimizer, ideally $f(p)$ can be designed such that $C(\mathrm{I}, f(p))$, will have the following properties:

- Is fast to compute for a given p
- Has stable numerical derivatives
- Represents a convex and well-conditioned cost function.

As model complexity increases, the difficulty of solving the optimization problem is also expected to increase. Therefore, it is reasonable to start with a simple model, and add increasing astrophysical fidelity incrementally to find an appropriate level of model complexity that is still workable in a nonlinear parameter estimation context.

Following this reasoning, the initial model developed for this purpose consisted of an assumption of an exozodiacal disk which is circular, with some 1D function describing a radially

symmetric intensity profile. Parameters for inclination and orientation relative to the imaging system are also included. This can be considered an 'empirical model' in the sense that it is not derived from astrophysical principles, and is instead intuited from inspection of the image data without any astrophysical reasoning. The model has the following parameters:

- Center position x_c, y_c (pixels)
- Inclination and Orientation i, θ (deg)
- 1D function for symmetric intensity roll-off:
    - $f(r) = se^{(ar^2+br+c)}$, where:
        - r is radial distance from the center
        - s is an intensity scale factor
        - a,b, and c are polynomial coefficients

This gives a total of 8 parameters, denoted by parameter vector $\boldsymbol{p}$**:**

$$\boldsymbol{p} = [\, x_c, y_c, i, \theta, s, a, b, c]$$

Although this model is not derived from astrophysical principles, it has some physical interpretability in the sense that x_c, y_c represent estimated disk center, while i, θ represent disk inclination and orientation, under the rough assumption of an infinitely thin disk with perfect radial symmetry. The 1D intensity roll-off function is any arbitrary function which can explain the observed roll-off. A decaying exponential with a second order polynomial in the exponent was selected for this purpose after some "guess and check" experimentation.

With an initial model selection, the next step is to define the cost function $C(\mathrm{I}_o, \mathrm{I}_p)$. A typical approach would be to simply compute the sum of squared error:

$$C(\mathrm{I}_o, \mathrm{I}_p) = \left\| \mathrm{I}_o - \mathrm{I}_p \right\|_2^2$$

However, consider that the model only attempts to be a first order model of smooth exozodi, and that various unmodelled effects will be present in the image: planet signals, solar glint, starshade transmission, resonant structures, etc. Given that we expect large amounts of unmodelled signal, a robust regression approach should be used. An assumption is made that these unmodelled effects will be somewhat local to particular pixels in the image, whereas the exozodi, being the broadest structure in the image, will have a certain number of pixels which are in fact well explained by the model, and some pixels which are not well explained by the model. If this general assumption is true, then the problem of unmodeled effects in the image can be thought of as an outlier rejection problem. I.e. many pixels in the image will fit the exozodi model, whereas many pixels will also be significantly affected by other unmodelled affects. In order to address this outlier problem, the following cost function is used:

$$C\left(\mathrm{I}\,,\,f(p)\right) = \sum_{i=1}^{N} L(f(p)_i - I_i)\, w_i$$

Where:

- w_i represents a per-pixel weighting
- L is a modified Huber Loss function

The per-pixel weighting is used to down-weight pixels that are not expected to fit the model. For example, pixels within a defined radius of the starshade are weighted as zero since a combination of transmission effects and solar glint effects are expected to always cause poor model fit in that region. The selected loss function L is similar to the Huber loss, except that for the region of the loss above the threshold, a logarithmic loss is used instead of a linear loss. This provides a compromise between the Biweight loss which plateaus at a maximum loss vs Huber loss which continues to grow linearly.

This approach gave reasonable success for smooth disk scenarios. As shown in figure 1, for a smooth disk scenario, the simple disk model and regression approach can yield fairly ideal background subtraction where the residual consists of planet signal plus sensor noise. Note that for this example image content within a radius of 3 pixels from the starshade center is ignored. A less ideal result is shown in Figure 2, where there is more obvious residual due to model error. It has been proposed that this significant model error found with smooth disk data may be due to some combination of forward scattering and inherent thickness of the disk. Even though the SISTER simulation uses a smooth disk model, it also implements forward scattering, and models non-zero thickness of the disk even when imaged at high inclination angles[1]. Therefore, it can be expected that this simple estimator as implemented will still have some significant model error at higher inclination angles.

It is worth noting that this simple model and estimation technique was developed for release 1 data, which contained only smooth exozodi images. Once the resonant structures were observed in the release 2 data, it became clear that the initial model would be limited for that scenario. Some results for resonant disks are shown in Figures 3 and 4. The resonant disks show some degree of radial symmetry, so it may be possible to improve the fit of the original model by updating the arbitrary 1D intensity function.

However, with the growing complexity of the scene, the best use of time would be developing a model with some more astrophysical fidelity, especially since these effects have already been studied and modeled in SISTER.[1] Using the models from SISTER would present a dilemma because while the results of solving an estimation problem would be interesting, the results would be overly optimistic when using the same model for estimation as is used in the simulation. In the real world, we expect more unmodelled effects, i.e. we never expect to have a perfect model. So while using the models in SISTER to solve estimation problems for images generated by SISTER may represent non-trivial estimation problem it would not necessarily serve the goals of the data challenge. However, it would be reasonable to demonstrate some

degree of robustness to unmodelled effects by turning of or changing some effects in the test set such that the models used for testing do not exactly match the models used for estimation. Another tool which was explored for possible conversion into an implementation for $f(p)$ is the ZODIPIC package developed by Marc Kuchner.[2] The tool was downloaded and run successfully in IDL but was not pursued further.

Developing an $f(p)$ with both sufficient fidelity to capture all the background effects, and also the right properties for an inverse estimation problem seemed like too large of an effort to complete within the scope of this project, but would be of great interest for future work.

A decision was made to switch focus to non-parametric background estimation approaches, since there is reason to believe the data can be decomposed into background and foreground components due to underlying differences in their behavior across a large dataset.

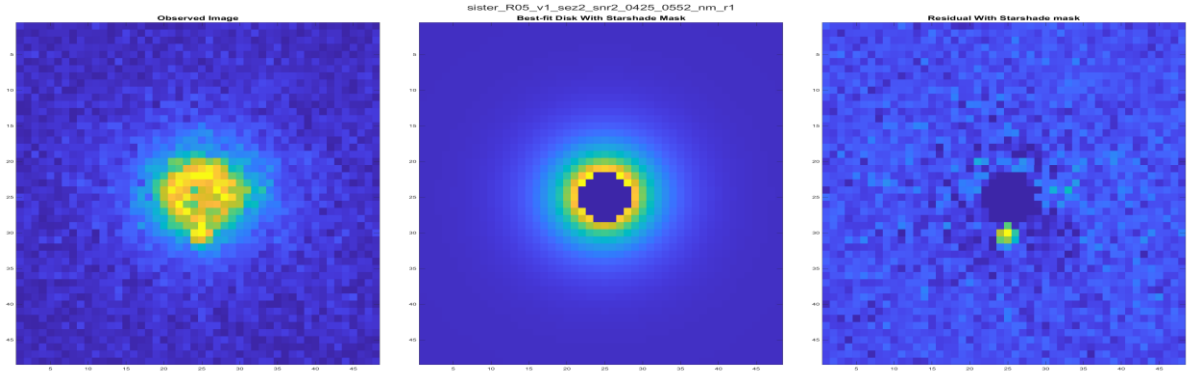

Figure 1: Example background subtraction results using a simple symmetric disk model with robust regression. From left to right: a. the observed image for R05_v1_sez2_snr2_0425_0552_nm_r1, b. The estimated disk model, c. the residual.

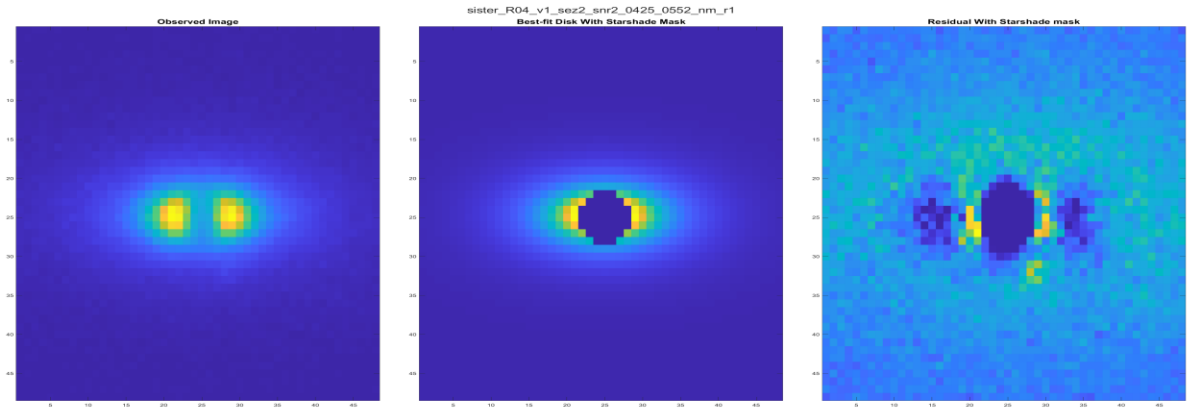

*Figure 2 Example background subtraction results using a simple symmetric disk model with robust regression. From left to right: **a.** the observed image for R04_v1_sez2_snr2_0425_0552_nm_r1, **b.** The estimated disk model, **c.** the residual. In this case the residual contains more obvious model error which is greater in magnitude than the planet signal to the bottom right.*

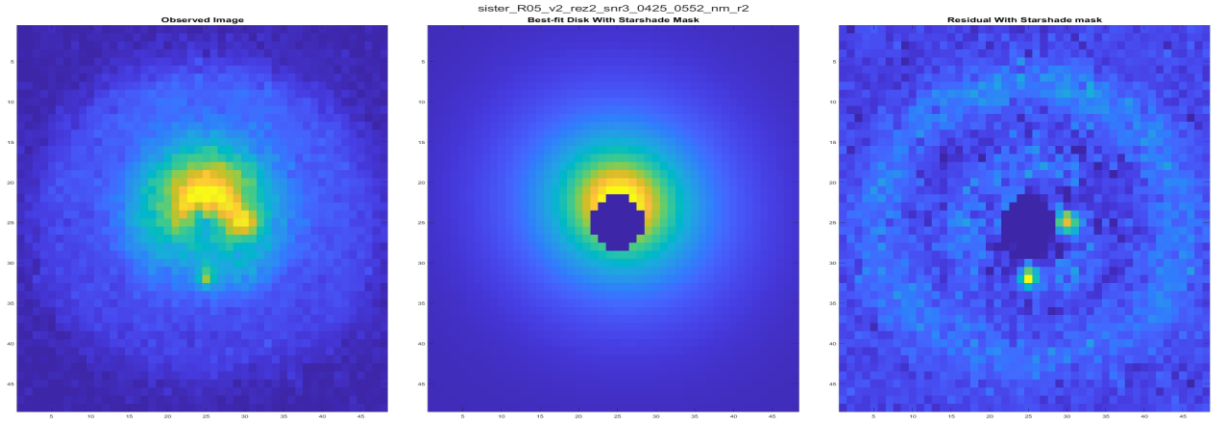

Figure 3 Example background subtraction results using a simple symmetric disk model with robust regression. From left to right: ***a.*** *the observed image for R05_v2_rez2_snr3_0425_0552_nm_r2,* ***b.*** *The estimated disk model,* ***c.*** *the residual. In this case the residual contains some significant model error due to resonant structures. This residual becomes dominant when the relative magnitude of the exozodi is greater.*

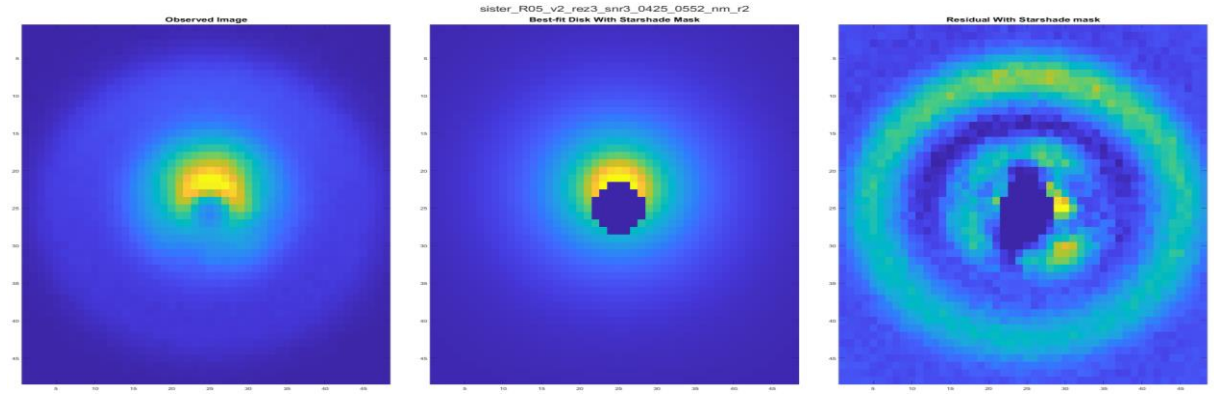

Figure 4: Example background subtraction results using a simple symmetric disk model with robust regression. From left to right: ***a.*** *the observed image for R05_v2_rez3_snr3_0425_0552_nm_r2,* ***b.*** *The estimated disk model,* ***c.*** *the residual. This shows the same scenario as Figure 4, but with exozodi intensity level 3. In this case the residual due to model error is on the order of planet signal.*

2.1.2 Non-Parametric Background Estimation

There are a few reasons to believe that planet signal can be separated from background signal without an explicit background model:

1. Planets move between multi-epoch observations, whereas backgrounds do not.
2. The shape and scale of the disk structure in the image changes significantly as a function of wavelength, whereas planet positions do not. This is relevant when combining images from the same observation into a multi-channel image considered as a single sample.
3. Disk structures have underlying similarity between different scenarios and dominate the signal energy.
4. Planet signals are point sources with low signal energy, appearing in non-repeatable positions.

A general idea which may exploit these underlying effects to separate planets from background is to use dimensionality reduction techniques to find a latent model which models backgrounds but not planets. A summary of the idea is that a low dimensionality latent model will have a denoising effect, where planet signal can be thought of as a noise signal on top of the broad background structure. Three approaches were explored: PCA, ICA, and an Autoencoder.

For all three approaches, the image data is collected into a matrix $\mathbf{X}$, where the rows represent samples, and the columns represent variables, which in this case are the image pixels flattened into a row. The data matrix $\mathbf{X}$ is prepared as follows:

1. Each pair of 425-522nm and 615-800nm images taken in the same visit number for the same scenario are consider a single two-channel image representing a single observation.
2. The subset of images with exozodi intensity levels 2 or 3 are selected to reduce modelling of planet signal.
3. A 41 x 41 pixel ROI is imposed on each image to further reduce the starting number of pixels and focus on the interesting region of the scene. This reduces the number of variables in a row of $\mathbf{X}$ by 60%.
4. Each 41 x 41 x 2 image is duplicated with a combination of flips and rotations. (data augmentation)
5. Each two-channel 41 x 41 pixel has its channels concatenated, and then the image is flattened to a 3362-element row. The rows are combined to form data matrix $\mathbf{X}$.

Given this $\mathbf{X}$, the general idea is to find some transform $\mathbf{T}$ which can be applied to latent scores $\mathbf{S}$ to construct $\widehat{\mathbf{X}}$ such that some measure of the error between $\mathbf{X}$ and $\widehat{\mathbf{X}}$ is minimized. The scores in $\mathbf{S}$ represent the reduced dimensionality representation of $\mathbf{X}$. Rows of $\mathbf{X}$ are observed samples in the input space $\mathbf{R}^{\mathrm{N}}$, and rows of $\mathbf{S}$ are those same samples represented in the latent space $\mathbf{R}^{\mathrm{M}}$, where M < N. In this case N = 3362, and the dimensionality of the latent space M must be tuned per model.

While each of these methods are complicated areas of study in their own right, the three approaches can generally be thought of in these terms with different levels of generality on the transform $\mathbf{T}$:

PCA: Principal Component Analysis

$\widehat{\mathbf{X}} = \mathbf{TS}, \mathbf{T}'\mathbf{T} = \mathbf{I}$ ($\mathbf{T}$ is orthonormal)

Note: It is also required that the rows of $\mathbf{T}$, which represent the "principal axes", must be in order of greatest variance of $\mathbf{X}$ to least variance explained.[3]

ICA: Independent Component Analysis

$\widehat{\mathbf{X}} = \mathbf{TS}$ ($\mathbf{T}$ is a linear transform, but not constrained to be orthonormal)[3]

Autoencoder

$\widehat{\mathbf{X}} = \mathbf{T}(\mathbf{S})$ ($\mathbf{T}$ is a nonlinear decoder function implemented by a neural network)[4]

In all cases $\mathbf{S}$ represents the samples of $\mathbf{X}$ transformed into latent space, where the transform $\mathbf{T}$ reconstructs samples from the latent space to original feature space represented by $\mathbf{X}$**.** In this case $\mathbf{T}$ transforms row vectors in the latent representation $\mathbf{S}$ back into the 3362-dimensional pixel space.

Once a model is obtained, the process of obtaining a background estimate for a given observed image $\mathbf{I_o}$ is as follows:

1. Flatten the image $\mathbf{I_o}$ to a row vector $\boldsymbol{x}$ in the input pixel space $\mathbf{R}^{\mathrm{N}}$.
2. Use $\mathbf{T}$ to project $\boldsymbol{x}$ onto the latent space $\mathbf{R}^{\mathrm{M}}$, producing score vector $\boldsymbol{s}$
3. Apply $\mathbf{T}$ to $\boldsymbol{s}$ to obtain reconstructed image $\mathbf{I}_{\boldsymbol{r}}$
4. Treat $\mathbf{I}_{\boldsymbol{r}}$ as the background estimate since the model is tuned for backgrounds.

Assuming that the model preferentially represents background signals, $\mathbf{I}_{\boldsymbol{r}}$ is used as the background signal to subtract for the purpose of this data challenge.

Of the three methods discussed, PCA is the oldest, most well-known, and easiest to compute. ICA can be considered a generalization of PCA where the orthonormality constraint on T is removed.[3] In turn autoencoders can be considered a generalization of ICA where the encoder and decoder are implemented by a neural network which can represent an arbitrary nonlinear function.[4] Autoencoders are more complicated to work with due to long training times and hyperparameter tuning. Due to the higher number of free parameters, more training data is required to constrain the model.

All three methods were implemented, and an explanation of results for each method is given in the following subsections. It should be noted that the judgement of performance for each method is somewhat subjective. Since no ground truth information is available, there is no obvious way to directly quantify background subtraction performance. In all cases visual inspection and human judgement are used to understand background subtraction performance.

PCA Results

Various combinations of data augmentation and preprocessing for generating $\mathbf{X}$ were attempted, and a PCA model was generated for each case. The model that appeared to give best results was a 17-component model with mean centering and some degree of data augmentation. The components of this model are shown below in figure 5. These components may be combined in any linear combination to approximate a given image test sample. The first few components have some degree of interpretability: Component 1 represents a typical smooth disk shape. (happens to be negated) Component 2 represents a term to impose some degree of top to bottom asymmetry. Component 3, introduces some higher order asymmetry. The higher order components are rarely interpretable, but in some cases they can appear to contain planet signals which is useful to see for diagnostic purposes.

Using this 17-component model, the background subtraction results achieved with PCA are reasonable in some cases, but poor in other cases. While it is difficult to summarize the varied performance across the test samples, an example of PCA based background subtraction for one particular case is shown in Figure 6. Note the clear planet signals in the residual, but also the presence of some model error in the residual. (i.e. structured residual content that is neither planet signal nor instrument noise) The general finding with PCA was that as you increase the number of principal components used (the dimensionality of $\mathbf{S}$), at some point above 17 principal components, the components start to include planet signal.

The performance is limited to some maximum number of principal components before you start to lose planet signal along with background signal as you increase the number of principal components. At this level there is still a lot of structured error in the residual for certain cases. While some further tuning is possible in terms of how $\mathbf{X}$ is generated and preprocessed, the method is generally fairly rigid by definition, so a dramatic improvement in performance is not expected with further tuning.

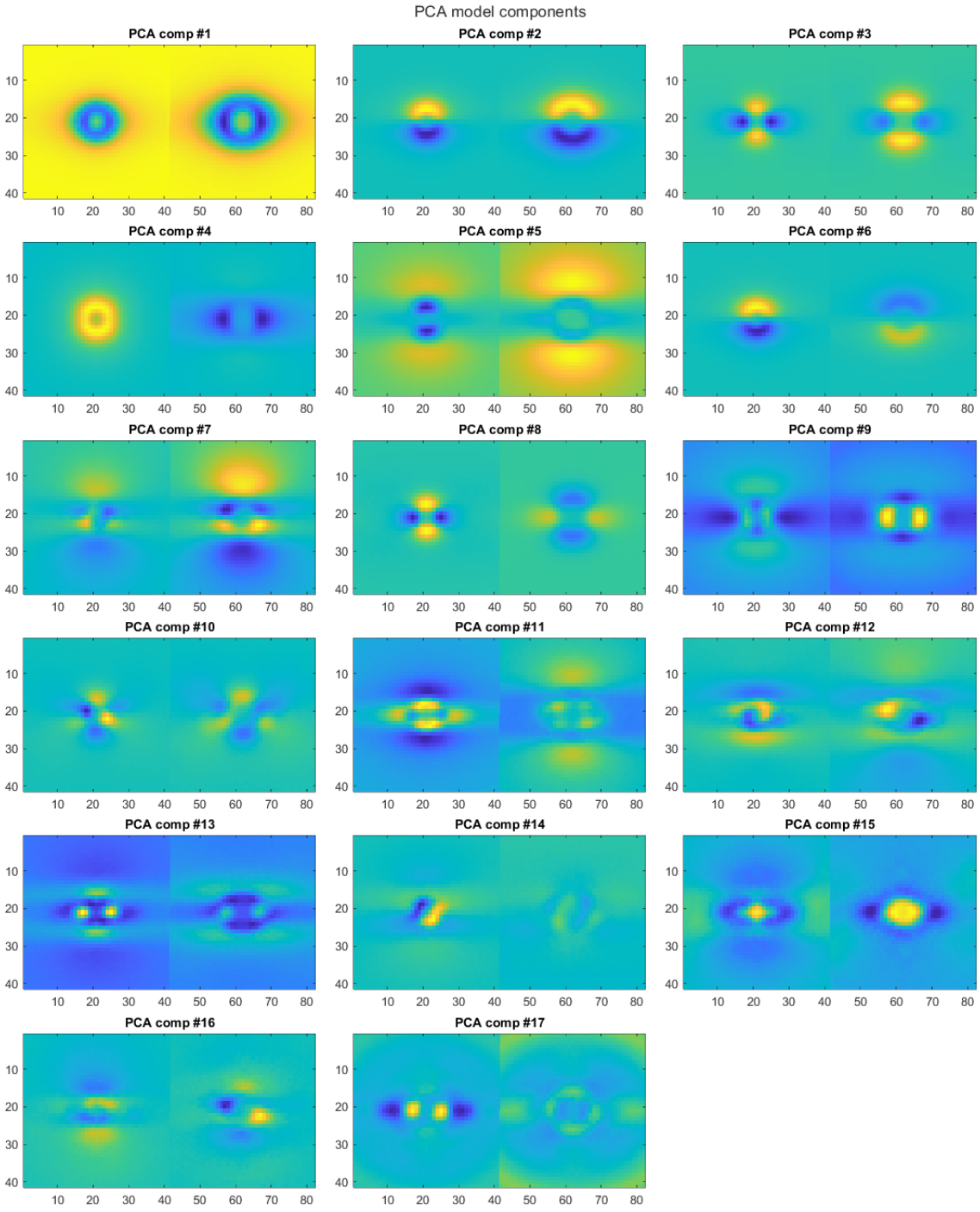

Figure 5: PCA model components up to PC # 17. Components are shown as 41x82 images where in each image the left 41x41 section corresponds to the [425-0552] nm passpand, and the right 41x41 section corresponds to the [615-800] nm passband.

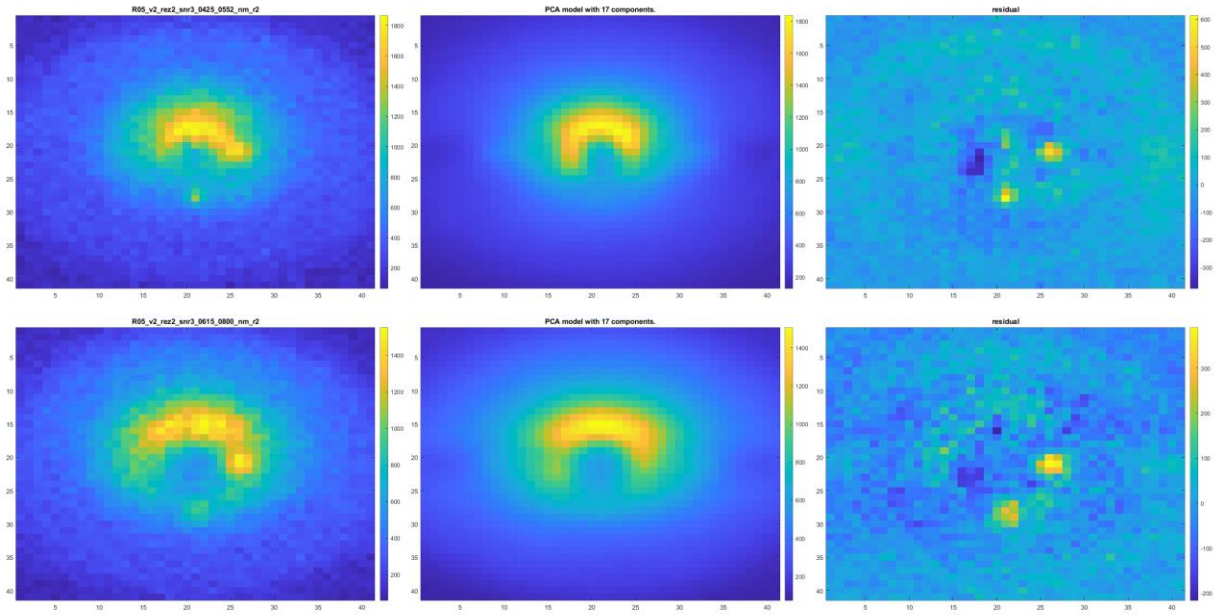

Figure 6: Example of PCA based background subtraction results for "R05_v2_rez2_snr3_1em10" scenario. A 2x3 set of subplots is shown where rows 1,2 correspond to passbands [425-522] and [615-800] respectively. Columns 1,2,3 of the plot correspond to: original image, estimated background, residual.

ICA Results

After hitting diminishing returns while tuning PCA, a natural next step is to try ICA, which can be thought of as a generalization of PCA which removes the orthogonality constraint on $\mathbf{T}$. The latent representation still uses a linear transform, but any arbitrary image components can be found such that taking linear combinations of those components model the content of $\mathbf{X}$. The removal of the orthogonality constraint when going from PCA to ICA makes the practical problem of generating a good model more difficult. Whereas a PCA model can be computed in closed form with a singular value decomposition, computing ICA components requires an iterative optimization approach with more configurations to tune. It is also the case that whereas a single PCA model can be used for any number of components, an ICA model must be recomputed for each proposed dimensionality of **S**. A single ICA model was selected after computing multiple ICA models with different configurations for: data augmentation, preprocessing, and optimization settings. The selection criteria for the model was visual inspection since no ground truth is available. The components for the selected model are shown in Figure 7. (A 17-component model) Note that in contrast to the PCA model, the components are not ordered by percent variance explained. Visual inspection of the model is important to rule out the presence of planet like content or measurement noise content. The

general heuristic used to arrive at this model was to keep increasing the number of components until the model starts to visibly include planet signal or measurement noise signal.

Inspecting the background subtraction results using this model yielded similar results to PCA in many cases, but appeared to work better in some of the edge cases where PCA had previously performed poorly. Based on visual inspection of several samples, it was concluded that this ICA model had somewhat better performance on average. An example result is shown in figure 8. This example was selected as a resonant disk example with exozodi intensity level 2. This is overall a “medium difficulty” sample in the sense that there are resonant structures, but the exozodi intensity relative to planet intensity is only at level 2. It is clear that the ICA model (and PCA model as well) does a reasonable job of accounting for the broad resonant structures without an explicit model.

The ICA model background subtraction was ultimately selected for moving forward to produce results for the data challenge. One further improvement was made, which was to use a robust regression with Huber loss to solve for the scores vector $\boldsymbol{s}$ for a given test image. This additional tweak adds robustness against outlier pixels such as relative bright planet signals which may otherwise bias the background estimate for a particular image.

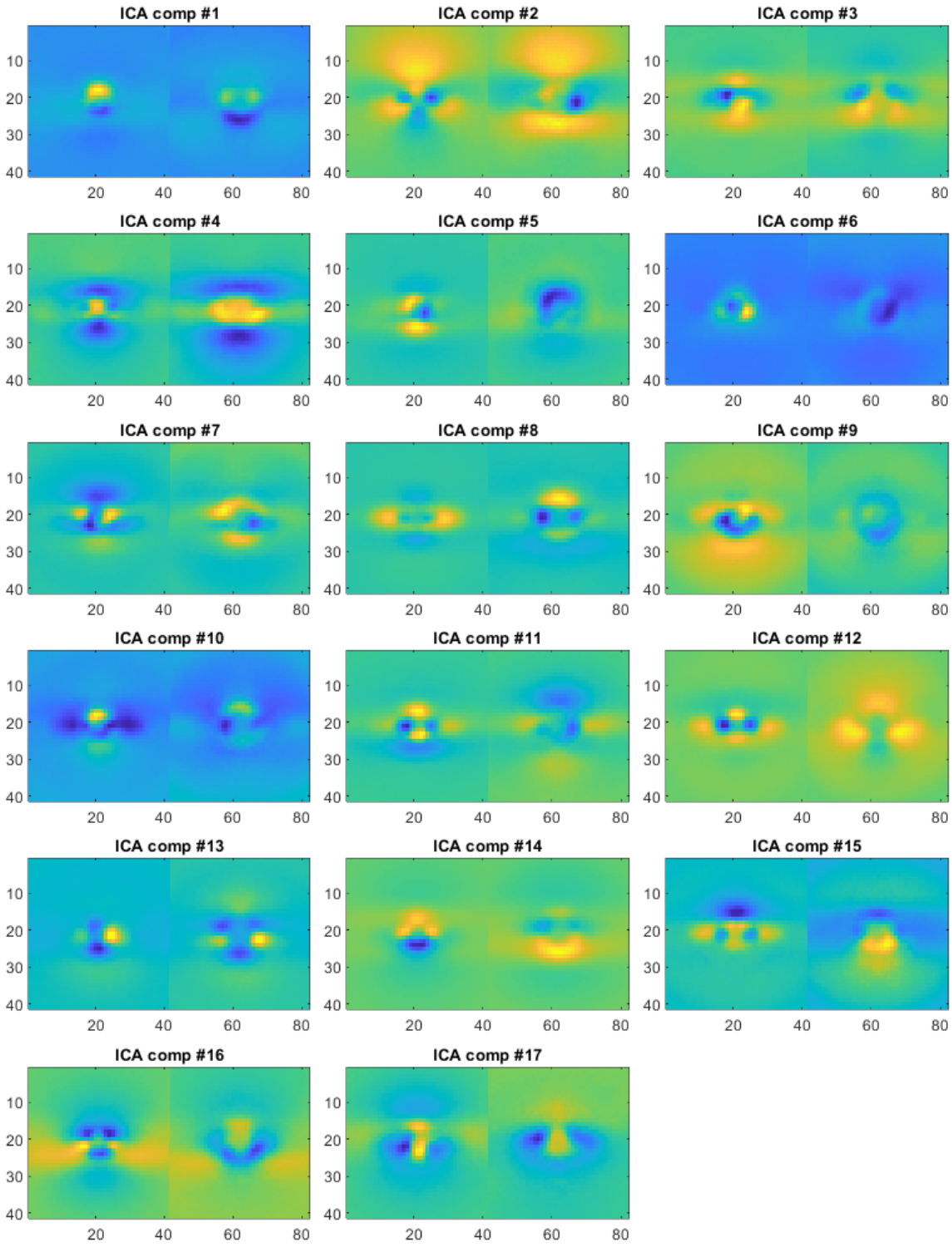

Figure 7: ICA model components for the selected 17 component model. Components are shown as 41x82 images where in each image the left 41x41 section corresponds to the [425-0552] nm passpand, and the right 41x41 section corresponds to the [615-800] nm passband.

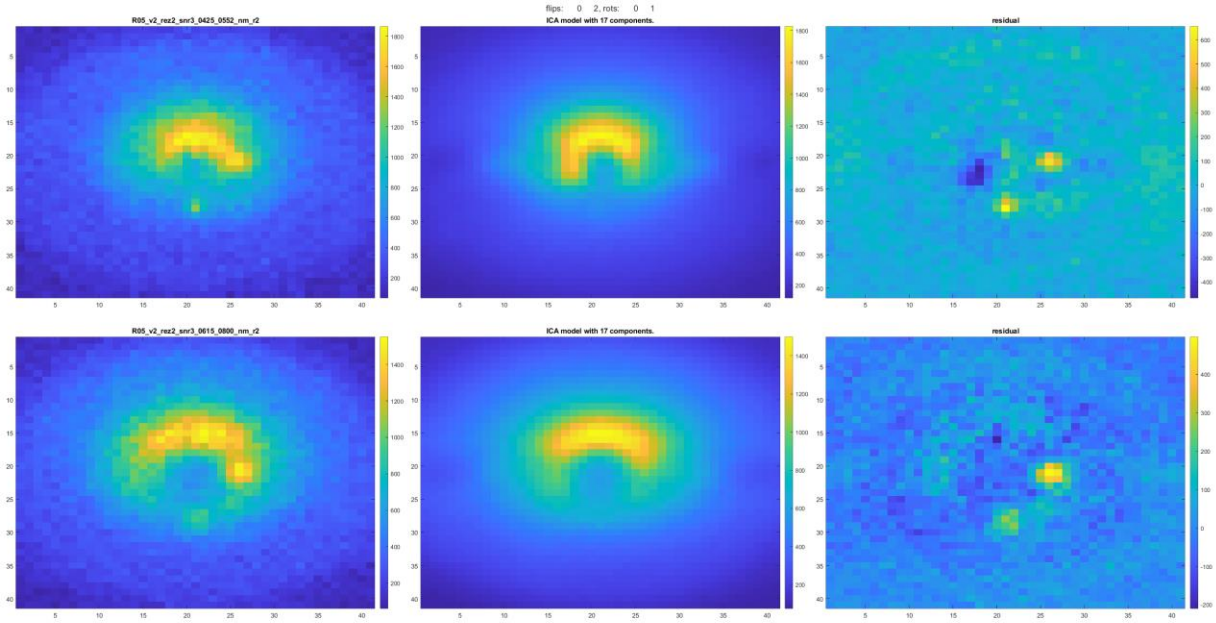

Figure 8: Example of ICA based background subtraction results for "R05_v2_rez2_snr3_1em10" scenario. A 2x3 set of subplots is shown where rows 1,2 correspond to passbands [425-522] and [615-800] respectively. Columns 1,2,3 of the plot correspond to: original image, estimated background, residual.

Autoencoder Results

The autoencoder approach is more general than the PCA and ICA approaches, and in theory it should be able to find a better latent model, since the PCA and ICA models are special cases of the possible models which can be represented by an autoencoder. In practice whether a better latent model is found with an autoencoder depends on various factors such as: size of training set, network design, hyperparameter tuning, training method, and training time. It should also be noted that there are multiple types of autoencoders which could be applied here. In this case a variational autoencoder was used. A convolutional autoencoder approach would also be of interest for future work.

The increase in generality going from ICA to an autoencoder also increases complexity of the implementation and variations that can exist in implementation and parameter tuning. While a detailed discussion of autoencoders is outside the scope of this report, a summary of the approach is as follows: a variational autoencoder was used where loss function used in the training process includes an L_2 regularization term and a sparsity regularization term. Due to the very long training times, tuning of hyperparameters such as training settings and dimensionality of the latent space was not feasible. A hidden layer size of 30 was selected, which is larger than the dimensionality found for the ICA and PCA approach. The hope is that a

larger dimensionality will allow background content to be well explained, but that regularization will encourage sparsity of the representation.

Initial Autoencoder results looked poor compared to PCA and ICA. A much better result should be possible with further effort, but progress is limited by training time on the order of 24 hours, which creates a long iteration loop.

2.1.3 Background Estimation Summary

A simple parametric background model was demonstrated with capability of estimating and subtracting backgrounds for simple smooth disks, excluding the center region of the image. This parametric model is useful for estimating some physically interpretable parameters such as disk orientation and inclination and center in x,y. In order to estimate more complex backgrounds including effects like resonant structures and forward scattering, further work is required to develop a parametric model which has both astrophysical fidelity and is feasible to use in the context of parameter estimation using a nonlinear optimizer.

Three non-parametric background estimation approaches were tried: PCA, ICA and an Autoencoder. All three methods attempt to find a low dimensional latent representation which is expected to separate background from planet signals by exploiting underlying differences in the variance across a large dataset. PCA and ICA gave similar results based on visual inspection of many cases, with ICA performing better for some samples that appeared give poor results with PCA.

Initial Autoencoder results did not look very good, but it is reasonable to expect that with more time and effort background subtraction based on a denoising Autoencoder can ultimately outperform PCA and ICA. One caveat to this prediction is that more training data may be required to properly constrain the large number of free parameters in the network. Currently only 920 training samples are used including data augmentation, which is less than the input dimensionality of 3362. Ideally the number of training samples would be greater than the number of input dimensions, but this is not necessarily required.

The ICA method is selected as a reasonable placeholder for the background estimation problem, in order to move on to other subproblems and address other goals of the data challenge.

2.2 Planet Detection

The planet detection step starts with the background subtracted image, which will be denoted as $\mathrm{I_f}$ (foreground image). In the current workflow $\mathrm{I_f} = \mathrm{I_o} - \mathrm{I_r}$, where $\mathrm{I_o}$ is the observed image, and $\mathrm{I_r}$ is the image obtained by using the ICA model to project $\mathrm{I_o}$ onto the latent space, followed by reconstruction back into image space. For example, the residual images in Figure 8 are considered to be foreground images for the purpose of planet detection.

The pipeline implemented for the purposes of the data challenge includes the following steps:

1. Matched filtering: The foreground images are filtered with the appropriate PSF as a function of wavelength and instrument scenario.
2. Local maxima detection: local maxima detection is performed on the matched filter output to find regions of the image which are well matched to the PSF, indicating possible presence of a point source.
3. Outlier detection: given the distribution of local maxima intensities found in 2. univariate outlier detection is performed to detect which local maxima do not match the main distribution. The Grubbs test is used to detect local maxima which are outliers in terms of matched filter output intensity.
4. Keep the N most intense positive outlier detections and consider them as planet detections.

This detection approach differs from the original proposal to use Bayesian inference for the planet detection step. The Bayesian inference approach would have required characterizing multivariate probability distributions for $p(v|H1)$, and $p(v|H2)$, where v is a vector containing a feature descriptor plus pixel location, and H1 and H2 are respectively the hypotheses that a pixel corresponds to a planet, or does not correspond to a planet. Given that $p(v|H1)$ and $p(v|H2)$ are not known and no obvious method for characterizing them is available, an approach requiring far less prior knowledge was implemented.

While the implemented outlier rejection approach does not compute a likelihood ratio, it is statistically motivated and some comparison is warranted. The outlier rejection approach implicitly characterizes $p(v|H2)$, the probability of observing feature v under the hypothesis that a local maximum does not correspond to a planet. The probability $p(v|H2)$ may also be referred to as the probability that the null hypothesis is true. In this case the v is the intensity of a local maxima in the matched filter output, and only local maxima are considered as opposed to all pixels as originally proposed. The implicit characterization of $p(v|H2)$ is based on the histogram of local maxima in a single image. The outlier rejection step finds local maxima which have very low probability under the null hypothesis. The selected Grubbs test implementation uses an iterative outlier rejection where for each iteration a threshold is selected using the Student's t inverse cumulative distribution function. While it is possible to compute the probability of the null hypothesis for a given detection, these probability values were not computed explicitly. The Grubbs outlier test can be considered as a useful heuristic for setting a detection threshold on matched filter output, even if the associated distributional assumptions are not rigorously demonstrated.

A visualization of detection results for the “R05_v2_rez2_snr3_1em10” scenario is shown in Figure 9. It is apparent in Figure 9 That while the automated detection is successful for two planets, imperfect background subtraction is affecting the background level, creating valleys with no local maxima. For many images, imperfect background subtraction appears to be a limiting factor. Figure 10 shows an example of detection performance when background

residual is small. In this scenario it is apparent that the approach performs well for detection with low signal to noise ratio when background residuals in the foreground image are low.

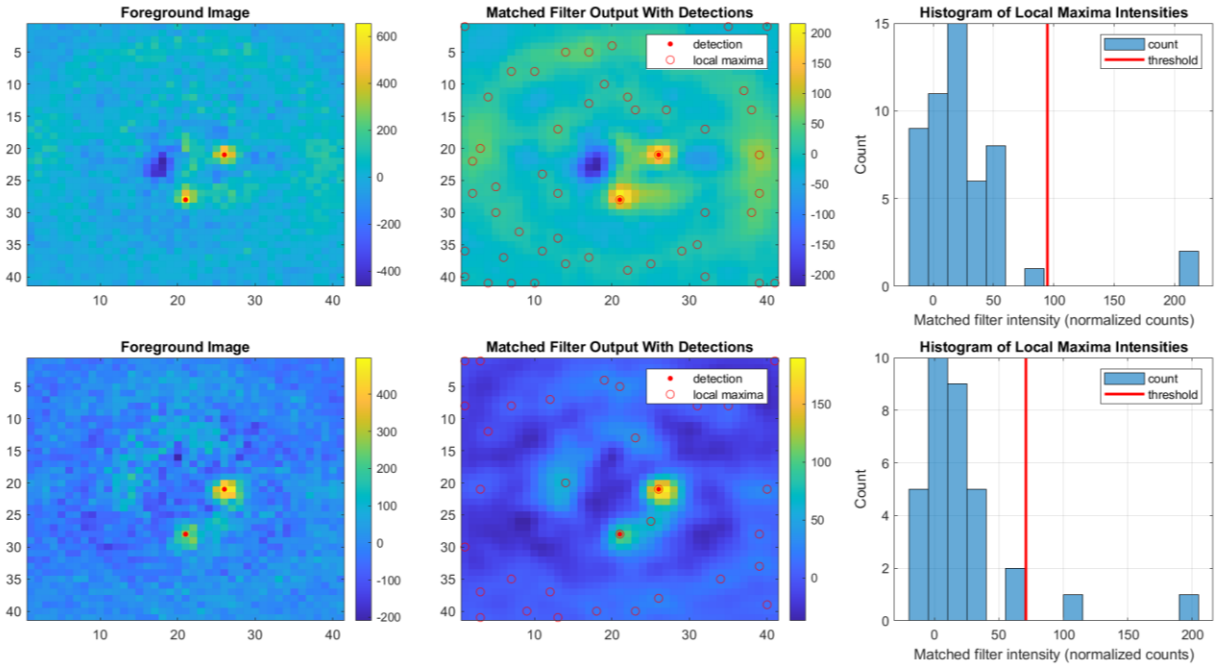

Figure 9: Example planet detection results for the "R05_v2_rez2_snr3_1em10" scenario. A 2x3 set of subplots is shown where rows 1,2 correspond to passbands [425-522] and [615-800] respectively. Columns 1,2,3 of the plot correspond to: foreground images with detections, matched filter images with detections and local maxima, and a histogram of local maxima intensities with outlier/detection threshold.

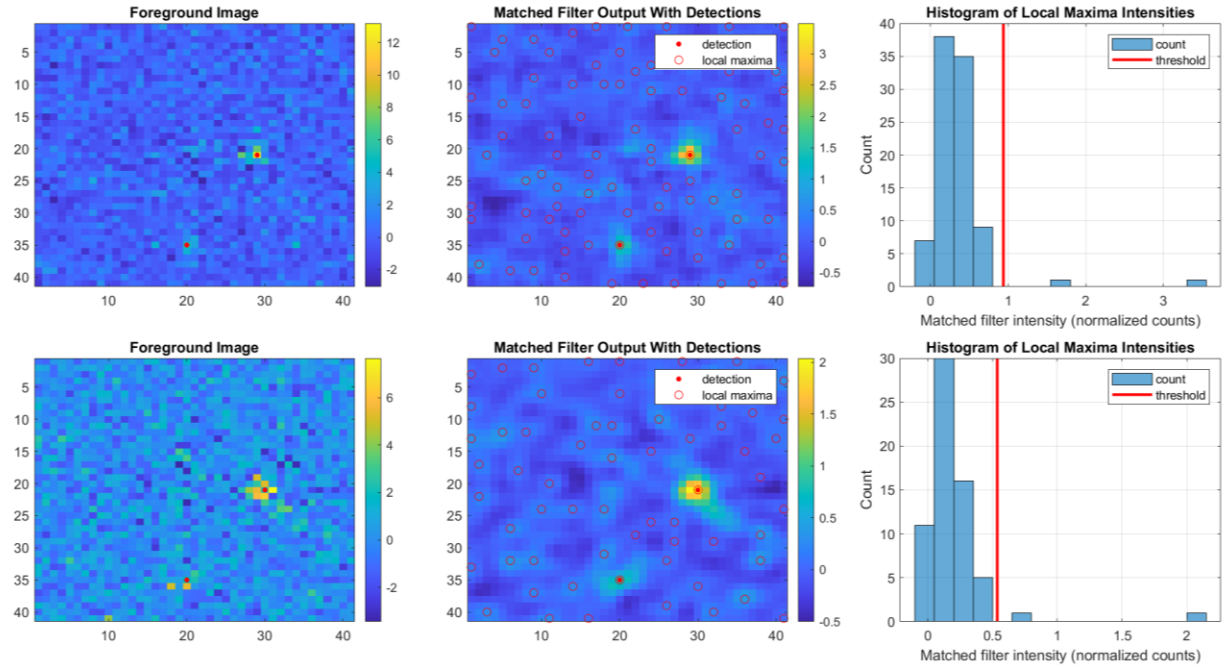

Figure 10: Example planet detection results for the "R02_v2_rez1_snr1_1em10" scenario. A 2x3 set of subplots is shown where rows 1,2 correspond to passbands [425-522] and [615-800] respectively. Columns 1,2,3 of the plot

correspond to: foreground images with detections, matched filter images with detections and local maxima, and a histogram of local maxima intensities with outlier/detection threshold.

In addition to planet detections found using the described approach, human labelled detections are also included for instances where the automated detection approach left obvious true positives undetected. Many presumed false positives were not removed to give a representative idea of false positive rate resulting from the automated detection method.

2.3 Planet Parameter Estimation

While the detection problem described in section 2.2 can be summarized as detecting pixels in the image which correspond to planets, the problem of estimating parameters for each detected planet can be considered as a fairly independent subproblem. The following approach was taken for estimating parameters per planet:

1. For each planet detection in a given image, select an NxN pixel ROI centered on the detection. In this case the select ROI width was 7 pixels.
2. Fit a parameterized model of a PSF to the sampled ROI region, using a standard nonlinear least squares approach.
3. Use the center location and magnitude from the best-fit PSF to ultimately compute: planet location and uncertainty, photon counts, and SNR.

Since the PSFs are known and provided, ideally a PSF with known form can be used for step 2 above, such that the true PSF shape is known and the position and amplitude can be estimated for a given planet. Because the PSFs provided for the instrument are at single pixel instrument resolution rather than some sub-pixel resolution, it is difficult to use them effectively for estimation of sub-pixel planet locations. One idea would be to take the PSFs provided at native resolution, and parameterize subpixel shifts via some smooth 2D interpolation. This may not correctly account for binning effects of pixel sampling, and therefore may not be a great representation of the true sub-pixel PSF. A different approach was taken to achieve sub-pixel parameterization. A 2D gaussian function was used, but the relationship between the Gaussian model and the provided PSFs was characterized in advance such that the biases introduced by using a Gaussian PSF instead of the true instrument PSF may be corrected. The process for accounting for the biases introduced by using a Gaussian PSF are as follows:

1. The known PSFs at single pixel resolution are loaded and a 2D gaussian model is fit to each known PSF within NxN ROI around the PSF center. For each best fit Gaussian, the sigma is saved along with the ratio of the sum of the counts of Gaussian PSF intensity, divided by the sum of the counts of the true PSF. The sigma and counts ratio are then saved in a lookup table vs image position, passband, and instrument scenario.
2. When performing estimation for detected planets, the lookup table generated in step 1. is queried to obtain correct sigma and counts ratio for the specific detection being performed.

3. A least-squares fit is performed by fitting a 2D gaussian to the observed planet in the NxN ROI. The sigma is constrained to equal the value from the lookup table, and three free parameters are estimated, x,y location and amplitude.
4. The x,y location is taken as-is to be the sub-pixel estimate of planet position, whereas planet photon count is estimated as the sum of the best fit gaussian PSF, divided by the counts ratio correction factor which was previously generated relative to the known instrument PSF.

This is a somewhat ad hoc approach to leverage the known instrument PSFs for sub-pixel estimation despite the PSFs being available only in single pixel resolution.

A more ideal approach would be to find an appropriate sub-pixel parameterization of the provided PSFs using an interpolation scheme, and then solve a weighted least squares problem where weights are used to reduce the influence of pixels in the 'tails' of the PSF with poor SNR. In the case of the Gaussian proxy model there is a quick intensity roll off from the peak, such that using weighted least squares to account for per-pixel SNR may be less critical. Repeating some of this estimation with the benefit of ground truth values some subset of the data would be informative for guiding development and ranking the accuracy of approaches. For the purpose of the data challenge, the results from the proxy Gaussian PSF approach are used.

It should also be noted that in some cases, structured non-random residual from imperfect background subtraction may bias the PSF fit substantially. The problem of optimal estimation for individual planet location is likely limited by background subtraction quality rather than particulars of some PSF estimation approach assuming only the presence of random measurement noise.

The SNR estimate for a given planet was calculated using a variation on equation 6.6 from the *Wide Field and Planetary Camera 2 Instrument Handbook*.[5] The general equation was modified to use available input:

$$SNR = \frac{S_{object}}{\sqrt{S_{object} + \ noise\_per\_pixel^2/sharpness}}$$

Where:

- S_{object} [e^-] is the total PSF counts estimated for the detection.
- $noise\_per\_pixel$ [e^-] is estimated as a local standard deviation of pixels around the outer perimeter of a detected PSF.
- $sharpness$ [unitless] is calculated as specified in equation 6.5 of *Wide Field and Planetary Camera 2 Instrument Handbook.*

This is a significantly simplified version of equation 6.6 of *Wide Field and Planetary Camera 2 Instrument Handbook* where $noise\_per\_pixel$ is a value derived empirically from the image. The $noise\_per\_pixel$ is estimated by taking the standard deviation of a local set of pixels

around a perimeter of given PSF for a detected planet. This should give a reasonable local estimate of all noise sources affecting the detection, other than the Poisson noise which increases near the center of the PSF.

3 Results

All estimated parameters and detections are compiled in the following files which are attached and delivered with this report:

- level_1_results_rev4_1em10.json: All generated results for the 1em10 scenario
- level_1_results_rev4_1em9.json: All generated results for the 1em9 scenario

4 Implications

This section provides opinions on some of the original questions posed by the StarShade data challenge:

Can backgrounds be calibrated to the photon-noise limit?

For some scenarios, based on visual inspection, it appears that magnitude of residual after background subtraction is dominated by random instrument noise. These ideal background residuals tend to be either from scenarios starting with low background intensity, or medium background intensity scenarios where the ICA model happened to fit very well. For these scenarios with low background residual, matched filtering detection should yield point source detection approaching optimal detection performance given random noise sources.[6] In many other scenarios the residuals from background subtraction included large structured model errors. In general backgrounds were not successfully calibrated to the photon-noise limit.

Are planets and exozodiacal dust disks or clouds separable from a single image?

Yes, both the parametric approach and non-parametric approaches demonstrate that it is possible to develop a background model which can separate background signal from planet signal in a single broadband image. However it may require many images to develop good background models, but once a background model is developed, it can be applied to a single broadband image to provide reasonable separation. Multi epoch and multi spectral images of the same system will certainly help and probably give better results vs the single broadband image scenario.

How sensitive is the planet detection capability to the instrument performance?

The main limiting factor tended to be background subtraction quality rather than instrument performance. For the special cases where background subtraction works well, the instrument

performance becomes limiting. The dominant background source tends to be exozodi, whose intensity are more a function of the planetary system being imaged rather than the instrument. As background subtraction methods improve, instrument performance may tend to become limiting.

5 References

[1] Rafels, S. H., Shaklan, S. B., Turnbull, M. C., & Cady, E. J. (2019). SISTER: Starshade Imaging Simulation Toolkit for Exoplanet Reconnaissance (Conference Presentation). *Techniques and Instrumentation for Detection of Exoplanets IX*. doi:10.1117/12.2528332

[2] Marc Kuchner, *ZODIPIC*, available: https://asd.gsfc.nasa.gov/Marc.Kuchner/home.html

[3] Hastie, T., Friedman, J., & Tisbshirani, R. (2017). The Elements of statistical learning: Data mining, inference, and prediction. New York, New York: Springer.

[4] Lyle Ungar, Unsupervised Neural Nets: Autoencoders and ICA, available: https://www.seas.upenn.edu/~cis520/lectures/unsupervised_networks.pdf

[5] Heyer, Biretta, et al. (2004), *WFPC2 Instrument Handbook, Version 9.0*

[6] Vio, R., Andreani, P., & Wamsteker, W. (2004). Some good reasons to use matched filters for the detection of point sources in CMB maps. *Astronomy & Astrophysics, 414(1), 17-21. doi:10.1051/0004-6361:20031632*