PSF modeling using the Gaia stellar catalog for the Dark Energy Survey

Rebecca Chen
Academic Year 2018-2019
Advisor: Prof. Chihway Chang

Contents

1 Project Introduction

This project was originally intended to to help test the performance of Piff, a PSF modeling software package, on real and simulated images for DES and LSST weak lensing. It developed into more of a look into the use of matching the Gaia stellar catalog with PSF solution stars in potentially improving PSF modeling using Piff. Most of the relevant code can be found on my GitHub at: `https://github.com/rebeccachen0/piff_work/`, but one file is on NERSC at global/homes/r/rcchen/ and should be publicly available. There is a ReadMe file that describes the uploaded files, and they will be described as relevant in this document. Hyperlinks and file names will be bolded. If there are any problems accessing anything or any questions, please feel free to email me at rebecca.c.chen@duke.edu.

2 Piff and PSFs

2.1 PSF background

PSF is a point spread function, which is essentially a description of the way that an optical system such as a telescope blurs out objects. As light travels from distant objects to us, it is affected by conditions such as the configuration of the telescope, our atmosphere, seeing, etc, that alter the "original" image. The PSF is a modeling of these effects– the actual image we see is a convolution of the true object and the PSF. Accurate PSF modeling is crucial for image analysis for ground-based surveys such as DES and the upcoming LSST, and in particular for weak lensing analysis, which requires extremely high precision and understanding of systematics. Currently, DES makes use of PSFEx for PSF modeling.

The basic idea behind PSF modeling is that the PSF varies across the focal plane, and we want to know what the PSF looks like at every possible location, since stars and galaxies can fall anywhere. We get this model by using stars, which are point-like sources. We assume that the spatial plane is smooth and use stars as anchor points to interpolate between them. This allows us to construct a function that outputs the PSF given a location on the focal plane. There are different model profiles that can be used for the PSF, such as Gaussian or Moffat, as well as different kinds of interpolation, such as polynomial or basis polynomial.

In order to evaluate the performance of a PSF model, 20% of the stars that could be used for modeling are reserved for testing ("reserve stars"). Several diagnostics have been designed to reflect how PSF modeling errors will affect galaxy shape. Because the PSF solution is interpolated, the error in PSF model for nearby galaxies will be correlated and will propagate as non-random, systematic errors in the two-point shear correlation function. These so called ρ statistics are defined in the DES weak lensing shear catalog papers (**SV** and **Y1**). To examine the goodness of the model we also compare direct measurement of size (T) and shape (ellipticity) of the stars with the corresponding values of the model at their locations. Looking at these residuals either through histogram or as a whisker plot helps give a measure of how well the model is performing.

2.2 Piff details

Piff stands for PSFs in the Full FOV (field of view). It is a newer PSF modeling software that seeks to improve on the performance of PSFEx. There are many contributors, but Mike Jarvis is the go-to for any Piff-related questions that can't be answered elsewhere. Some of the qualities of Piff to be highlighted that are an improvement over PSFEx include: 1) Can model PSF using all CCDs (chips) in an exposure, as is indicated by the name. 2) the PSF is modeled in "sky/real-world" coordinates, instead of pixel coordinates, which means that information about exposures can be used to fix problems before interpolation. 3) Piff is designed modularly, allowing for different models.

2.2.1 Running Piff

First, download Piff off the **GitHub page** or using pip. The easiest way of running Piff is to use a yaml configuration file: you provide an image and catalog file, specify details of the model, output directories, and whether you want other diagnostic plots. Then to run, use the command "piffify example.yaml". You can modify the configuration from the command line by specifying, for example, "input.image_file_name=image" after the piffify command. You can also edit the configuration and run piffify in a script by using:

```
import piff
piff_config = '/piff.yaml'
config = piff.read_config(piff_config)
config['input']['image_file_name'] = img_file
config['input']['cat_file_name'] = piff_cat_file
config['output']['file_name'] = psf_file
config['input']['wcs']['file_name'] = pixmappy
config['input']['wcs']['exp'] = exp
config['input']['wcs']['ccdnum'] = ccdnum
piff.piffify(config, logger)
```

In any case, running piffify will write out a .piff file which contains the PSF solution. It can also produce plots of the residuals and output rho statistics.

The Piff runs found on **Redmine** are made using **this script**. This script works by reading iterating over a list of exposures, storing exposure information in a pandas DataFrame, and then iterating over the rows, each of which represents a single ccd. For each ccd, it downloads and later deletes the files needed (image, background, catalog), runs SExtractor and findstars (algorithm for selecting PSF stars), removes bad stars, measures star shapes and sizes, runs piff, and measures piff shapes. This skips a few steps and details, but this is the general outline. Most of the files required to run this are located at `https://github.com/rmjarvis/DESWL/tree/master/psf` and/or on NERSC at /global/project/projectdirs/des/wl/desdata/users/mjarvis and the astro subdirectory.

2.3 Measuring shapes and sizes

To measure the shapes and sizes as mentioned previously, we use either hsm, which requires GalSim, or ngmix, which is found **here**. Code for this is found at both **y3_gaia_plotting_clean.ipynb** and **psf_plot.ipynb**. To measure for actual stars, we use GalSim to select a small region around each star and and then measure using hsm or ngmix code. For the PSF model, we need to read in the piff solution, then draw the psf at a location and measure it. This can be done using:

```
psf = piff.read('solution.piff')
temp = psf.draw(x_position, y_position)
```

For further help, run "dir(psf)", "help(psf.draw)", etc.

2.4 Color dependence

Using just the reserved catalog, I made a few plots that examine the color dependence of the residuals. Mike noted it's a bit surprising that the g band looks alright here, since it is the one with the worst rho statistics. The code for this is at **color_plot.ipynb**.

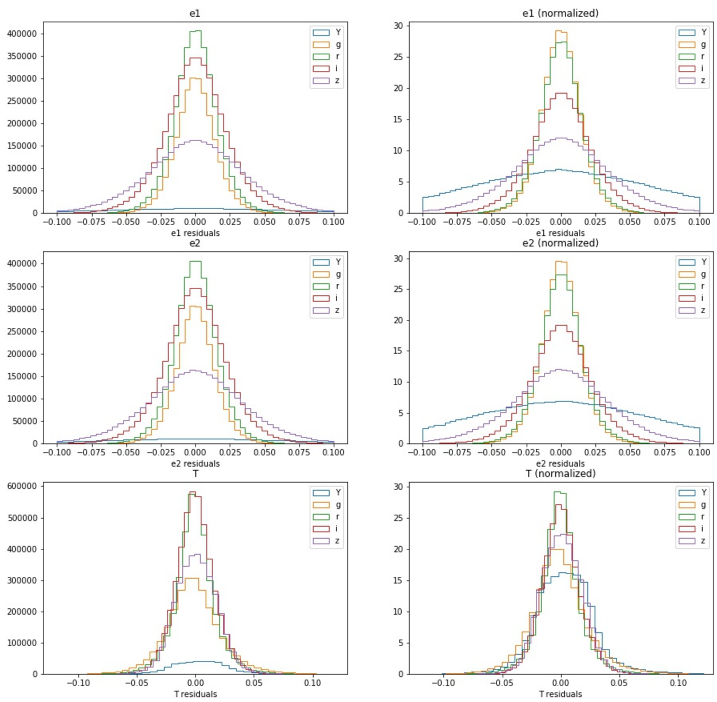

Figure 1: Color dependence of residuals

3 Gaia Stellar Catalog

Gaia is a space telescope run by ESA, designed specifically to do good astrometry and map out stars in the Milky Way. Kuang Wei has previously done work to match the Gaia stellar catalog to DES Y3 data in order to create a ultrahigh purity stellar catalog for DES. Documentation for this and the catalog download can be found **here**. If possible, download the catalog directly, as we ran into some problems using the easyaccess downloads. Using this catalog, we can further match to PSF stars used for Piff to examine possible galaxy contamination. We looked into the feasibility of using this overlap to either help further constrain the selection of PSF stars and whether this information might be usable for improving PSF modeling with Piff generally.

The process for matching uses astropy; it loops through each exposure, checks that the exposure wasn't blacklisted for some reason, and then saves the indices for matched stars. The code for matching and generating plots is located on NERSC at: global/homes/r/rcchen/gaia_match_plots.ipynb This includes code for producing plots of: 1) magnitude ranges for various matched/unmatched subpopulations 2) fraction of stars matched as a function of magnitude 3) histogram of fraction of stars matched by exposure. This also writes out a file in_gaia_info.fits that is a BinTableHDU with a row for every star (running through each exposure with the condition if(all(f==0 for f in flag)) in the order listed in exp_nums.txt and a single column — 1 if it is in Gaia and 0 if it is not.

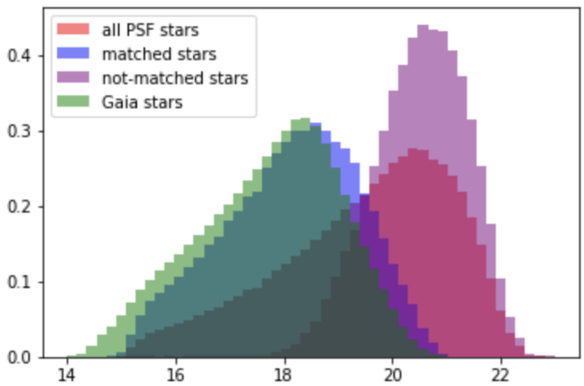

Figure 2: Individually normalized to show magnitude ranges

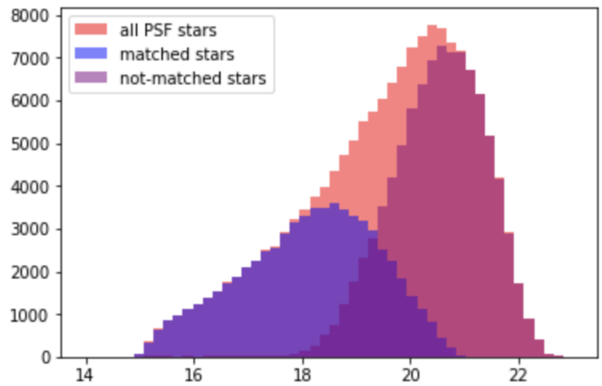

Figure 3: Not normalized to show actual distribution numbers

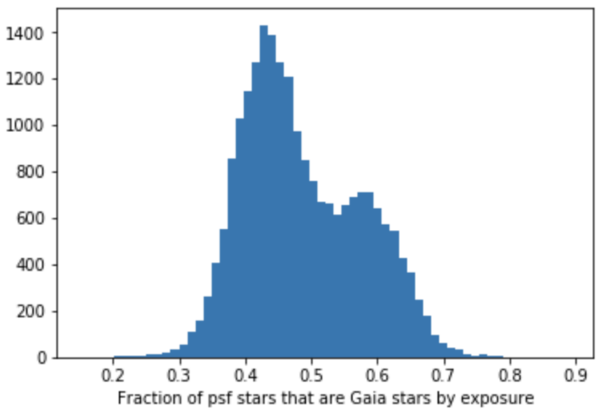

Figure 4: For each exposure, the fraction of PSF stars that are matched to Gaia is computed. These values are histogrammed.

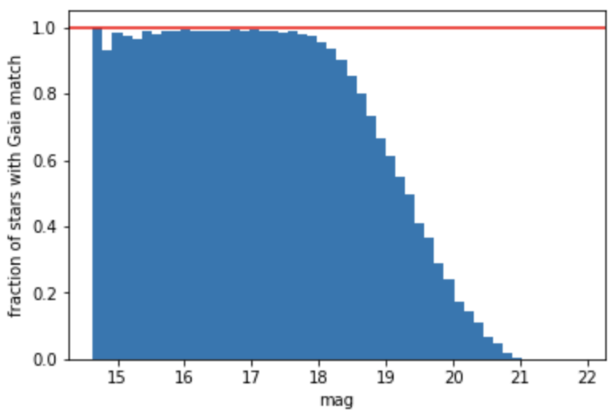

Figure 5: Fraction of PSF stars matched to Gaia as a function of magnitude. Few percent galaxy contamination can be seen in the magnitude less than 19 range.

3.1 Piff with matched catalogs

The next step is to do Piff runs with the matched catalogs to see if they produce reasonable results, as on average the number of stars is reduced to about half. One way to do this is to run Mike's run_piff script locally, inserting a step in/after findstars that removes non-Gaia stars. Instead of this, I looked at directly running piffify (**run_exposure.py**) but was not able to finish looking at the results. There are several problems with my current work that I was not able to fully diagnose; the biggest problem is that the resulting PSF solutions in running piffify directly on the same files used in the run_piff script does not produce the same results as given in the **Y3 Piff catalog** and generally seem to be off/unreasonable. This could be either because my method of running Piff is incorrect or my method of making shape measurements is incorrect. Something to consider is I was not using pixmappy, which doesn't account for the completely unreasonable results, but is something that will need to be done for the future. The comparisons and diagnostics can be found on the **y3_gaia_plotting_clean.ipynb**.

4 Data, Files, and Resources

4.1 Catalogs, etc

- Y3 Piff Catalog, also found on NERSC. Contains one file per exposure in format "exp_psf_cat_.fits" `https://cdcvs.fnal.gov/redmine/projects/deswlwg/wiki/PSF_Modeling`
- Y1 Test Data. Contains chip info for exposure 241238. Includes multiple cuts, catalog, bkg, findstars, and more. Folder "y1_test"
- Y3 Images and Catalogs, found on NCSA. Example url for image file where exposure is 516836 and c01 indicates the chip number. `https://desar2.cosmology.illinois.edu/DESFiles/desarchive/OPS/finalcut/Y2A1/Y3-2380/20160212/D00516836/p01/red/immask/D00516836_z_c01_r2380p01_immasked.fits.fz` Files are downloaded from here in the run_piff script using wget.

4.2 Resources and Links

General Information:

- PSF Modeling Redmine page `https://cdcvs.fnal.gov/redmine/projects/deswlwg/wiki/PSF_Modeling`
- LSST PSF Task Force page. Includes slides and recordings with PSF updates and background. `https://confluence.slac.stanford.edu/pages/viewpage.action?spaceKey=LSSTDESC&title=PSF+Task+Force+Meetings+and+Resources`
- Kuang's Gaia stellar catalog Redmine page `https://cdcvs.fnal.gov/redmine/projects/des-y3/wiki/Gaia_Stellar_Catalog_vs_DES_Y3`
- LSST Dark Energy School, lots of good background. `https://lsstdesc.org/pages/DESchool`
- DC2 Tutorials. `https://github.com/LSSTDESC/DC2-analysis/tree/master/tutorials#dc2-tutorials`

Piff:

- Piff GitHub.https://github.com/rmjarvis/Piff
- Documentation on Piff. http://rmjarvis.github.io/Piff/html/psf.html
- run_piff script. https://github.com/rmjarvis/DESWL/blob/master/psf/run_piff.py

Papers:

- Alex Amon's weak lensing KiDS paper. Similar Gaia comparison in Appendix C. https://arxiv.org/pdf/1707.04105.pdf
- DES Y1 weak lensing shape catalogues. https://arxiv.org/pdf/1708.01533.pdf
- DES SV shear catalogues. https://arxiv.org/pdf/1507.05603.pdf