Wesleyan University The Honors College

High Precision Photometry of Faint White Dwarf Stars from K2 Data

by

Michael Henderson
Class of 2019

A thesis submitted to the
faculty of Wesleyan University
in partial fulfillment of the requirements for the
Degree of Bachelor of Arts
with Departmental Honors in Astronomy

Middletown, Connecticut April, 2018

If no one comes from the future to stop you from doing it then how bad of a decision can it really be?

–Bill Murray

Twitter

Acknowledgements

I would like to thank the Astronomy department for being supportive and helpful from the first second that I began emailing about getting involved in the department to the second that I graduate. To stop myself from writing an appreciation section of novel length, I have created a short list of people that I would like to thank for helping, specifically, with my thesis. This, in no way shape, or form is meant to downplay the amount of tremendous help and care the whole department has offered me. I appreciate everyone: from Linda, the students, the professors, to the staff all around campus who have helped me outside of academic settings.

To my Advisor, Seth Redfield: Thank you for your patience and guidance through my journey through research and my confusion through it all. Thank you for sharing all of your infectious enthusiasm and genuine joy for research. Thank you for your encouraging words and your willingness to write me a recommendation, sometimes without me even asking yet. I really appreciate all that you have done for me in preparation for my time after Wesleyan.

To Roy: Thank you for all the times you have helped me unbreak my computer and make sense of what ever errors manifested from the ether onto my computer screen. And most importantly thank you for 3D printing not one but two batarangs. Also, I apologize for the many times I have used "sudo pip install" incorrectly, the power went to my head.

To Sophia and Prajwal: Thank you for helping me hone my python coding skills and teaching me the importance of providing thorough comments for my code. It is a lesson I only truly appreciate now.

To Anthony, Ismael, and Jessica: Thank you all for the many distractions imposed on me. Through these exchanges I was not only able to take time and separate myself from my work but also get a chance to know you all for the... interesting people you are. Although the time I have known you all has been fairly short, I enjoyed the many Super Smash Bros. Ultimate matches played as well as the many planned events that never happened.

To McNair: Thank you for the financial and graduate school assistance that made the process more manageable than it otherwise would have been. Thank you for the experiences you were able to give me and the friends I was able to make along the way. I am truly thankful for all that you have done that has allowed me to thrive as a first-generation, low-income student.

And finally, to my loving family: Thank you all for always being in my corner and supporting me through this mysterious journey. Thank you for believing in my brilliance when all I saw was ignorance. Thank you for believing in my adaptability when all I felt was stagnant. Thank you for remembering my sense of humor when all I felt was sadness. Thank you for knowing and understanding me far better than I cared to know myself. It is only because of you all that I can walk away from this journey in one piece and smile. That being said, I expect all you to now address me solely as Dr., in preparation for my next journey. You have only yourselves to blame.

Contents

Chapter 1

Introduction

In 1992, astronomers Alex Wolszczan and Dale Frail made the first confirmation of exoplanets using pulsar timing measurements (Wolszczan & Frail 1992). An exoplanet is a planet that orbits a star beyond our solar system. The planets orbit a pulsar, the highly magnetized remnants of a core collapsed giant star, located in the constellation Virgo. This confirmation was impactful because prior to their confirmation the existence of exoplanets was only hypothesized as possible by astronomers such as Otto Struve. Struve postulated that the absence of rapid rotation in solar-type stars suggest that stars convert angular momentum of rotation into angular momentum of orbital motion of planets as they compress down from enormous, slow rotating gas clouds into small, faster rotating stars (Struve 1952). This first confirmation was also substantial in light of earlier misidentifications of exoplanets, such as the Jupiter sized planets around Barnard's star. Barnard's star is a M dwarf in the Ophiuchus constellation and the second closest system to the Solar System. This system was a part of inaccurate claims of exoplanet detections for nearly 20 years by Peter van de Kamp (1963, 1969a,b, 1975, 1982) who had been claiming an astrometric detection of one or two Jovian planets with orbital periods between 11 and 26 years. Between 1963 to 1973, a number of astronomers accepted that Peter van de Kamp detected a perturbation in the proper motion of Barnard's Star consistent with it having one or two

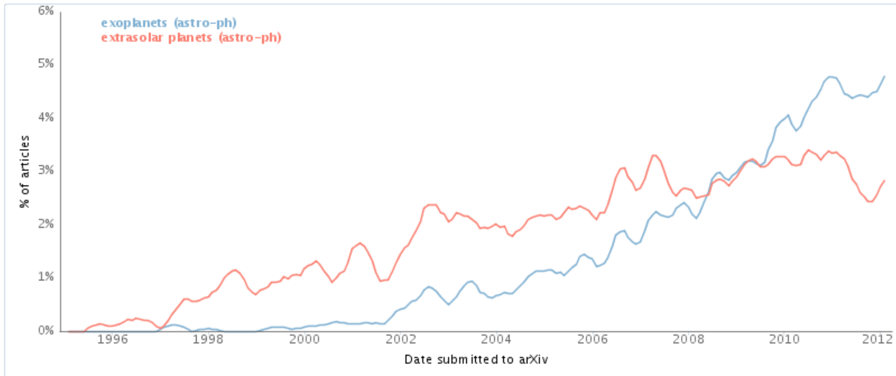

Figure 1.1: Percentage of research papers in astrophysics that mention exoplanets and extrasolar planets (Yaqoob 2012).

Jupiter mass planets. However, due to results that could not be reproduced by multiple astronomers van de Kamp's planetary discovery was considered a result of instrumental errors and thus not considered an actual discovery.

Nearly three decades after the first confirmation of an exoplanet, the field of exoplanet detection has developed into an expansive field whose focus, after nearly 4000 exoplanets confirmed, has shifted from the question of if planets exist outside of the solar system to how frequently planets form around other stars and how much do these systems resemble our solar system. Also, since the first exoplanet detection the field, has seen a dramatic increase in published articles. In Figure 1.1 the popularity of the field can be tracked as the increased percentage of published papers related to exoplanets. Much of the reasoning behind this boost in popularity may be tied to advancements in technology and common knowledge within the field. For example, a clear increase occurs shortly after the launch of the *Kepler Space Telescope* in 2009 which was a telescope dedicated to finding an Earth analog orbiting another star. Therefore it is safe to say that with the advent of new missions, such as *Transiting Exoplanet Survey Satellite* (*TESS*) and *James Webb Space Telescope* (*JWST*), we can expect continued growth of the field.

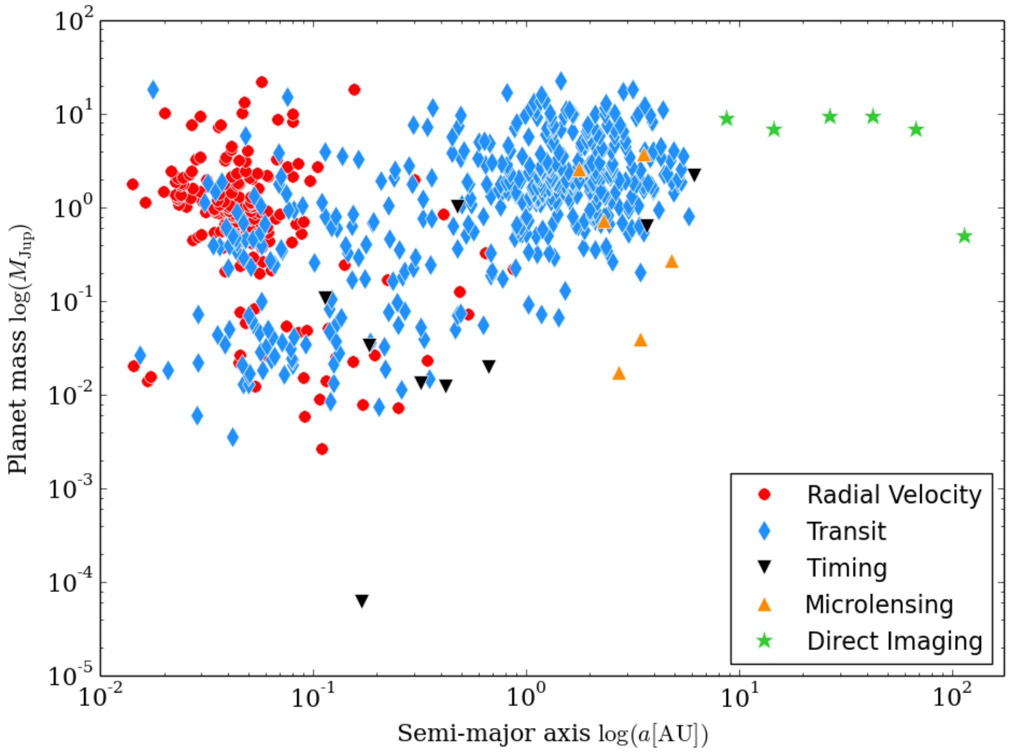

Figure 1.2: Planetary mass as a function of semi-major axis. The shapes represent the various detection techniques; radial velocity (red circles), transits (blue diamonds), timing (black downward triangle), microlensing (orange upward triangle) and direct imaging (green stars).

Since the first confirmation of an exoplanet our repository and understanding of exoplanets has grown and allowed us to see that planets around stars is more commonplace than an exception with estimations saying there is at least one planet around every star in the galaxy[1]. Most exoplanets have been discovered using the transit method, and to a lesser degree the radial velocity method. This can be seen visually in Figure 1.2 which shows the frequency and tendency for each detection method as well as the mass of the exoplanet (y-axis) and its distance from the host star (x-axis). Figure 1.2[2] also conveys our bias for close in hot Juipters due to their large size providing a larger signal for detection.

[1]https://exoplanets.nasa.gov/the-search-for-life/exoplanets-101/
[2]https://www.paulanthonywilson.com/exoplanets/exoplanet-detection-techniques/

The following with be a summary of these two, as well as a few others, methods used to detect planets around other stars.

I. Transit

The transit method is the most widely-used and effective exoplanet detection method to date. The transit method detects exoplanets by measuring the minute dimming of a star due to an orbiting planet that passes directly between an observer and the star. When a planet passes directly between an observer and the star it orbits (i.e. transits), it blocks some of that star's light which briefly dims the star by an amount that is related to the relative size of the two bodies and the planet's distance from the host star. Transits are recurring events, due to the planet being in orbit around the star, which helps validate whether the transit are really due to a planet.

II. Radial velocity

The radial velocity method takes advantage of the fact that a star, when orbited by a planet, has its own small orbit. This is due to the fact that celestial bodies actually orbit the center of mass of the system. This same phenomena happens in our Solar System, however, the center of mass is typically found inside of the Sun so its orbit is not as pronounced as if it were found outside of its radius making it more of a small wobble than an orbit. Although these variations are small they can be measured as a variation in radial velocity by spectrographs. When viewed from a distance these variations lead to Doppler shifts in the star's spectral features. If the star is moving towards the observer then its spectral features will appear shifted towards the blue side of the spectrum; if it is moving away it will appear to be shifted towards the red. These variations are used to detect when a star has planets orbiting it.

III. Pulsar Timing

Pulsar timing is the method used by Alex Wolszczan and Dale Frail to detect the first confirmed exoplanets. We use the pulsar timing method to detect planets around rapidly rotating neutron stars with strong electromagnetic fields, with field strengths between 10^{11} to 10^{13} G (Hooper et al. 2013). Neutron stars are the >1.44 $M_{\odot}$ remnants of violent high mass explosions, such as core collapse supernovae of stars with masses greater than 8 $M_{\odot}$. With a radius of $\sim$ 13 km (Ng et al. 2016) these compact objects are one of the densest objects in the universe. Because of the difference in radius after the supernovae pulsars can complete a rotational period anywhere between 0.01 to 10 seconds (Ng et al. 2016). As pulsars rotate they emit a beam of regular and precise electromagnetic radiation that is detectable from Earth. Slight anomalies in the timing of the pulses indicate that the pulsar has a companion that is perturbing its orbit (Tsapras 2018).

IV. Microlensing

Einsteins Theory of General Relativity tells us that gravity causes an affect on space-time and is able to curve it. This can cause light to distort and bend when in the presence of a large gravitational field. This affect can also briefly act as a lens as a foreground star bends and focuses the light of a background star, making the distant star appear brighter. Gravitational microlensing utilizes this phenomenon and looks for the marginal effects of a planet on the lensing of a star behind it (Beaulieu et al. 2006). Figure 1.3[3] illustrates the aforementioned microlensing phenomenon with a planet companion.

[3] http://www.planetary.org/multimedia/space-images/charts/microlensing.html

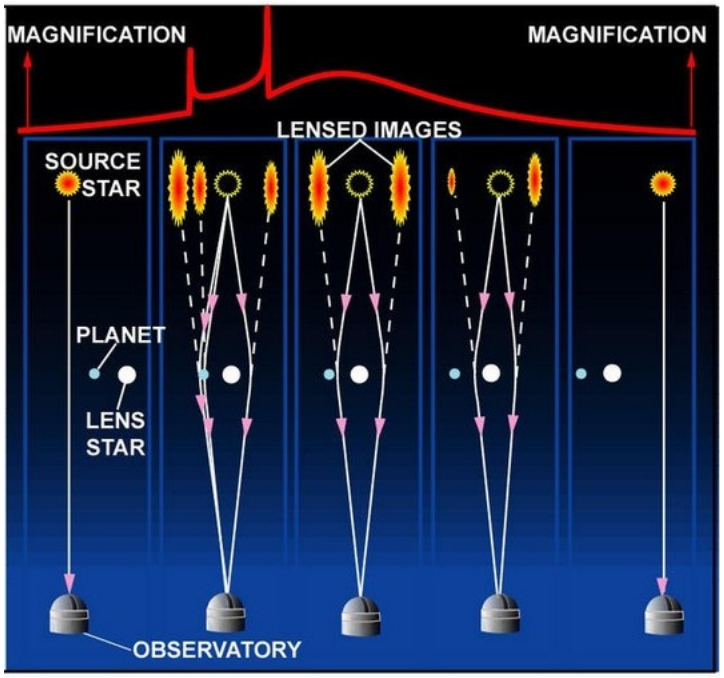

Figure 1.3: The microlensing process in stages, from right to left. In the fourth image from the right the planet has its own observable contribution to the lensing of the star.

V. Direct Imaging

Direct imaging of an exoplanet is the only method aimed at visually isolating the planet from the star. It is also the easiest exoplanet detection method to understand conceptually but is extremely difficult to execute in practice due to the fact that planets are small and faint compared to their large and bright host stars. This constraint means that only planets that orbit far from their host star can be found with this method.

The most influential telescope for planet detection has been the 9 year long mission of the the *Kepler Space Telescope*. The objective of the *Kepler* mission was to explore the structure and diversity of planetary systems using the transit method (Borucki et al. 2010). It is important to note the influence that a bias has on its observations and what the results of an observation means. Looking at the transit method the apparent bias is the magnitude limitation meaning that fainter objects are underrepresented because light falls off at $1/r^2$. This inverse square law tells us that closer, brighter stars are then favored over further ones.

Another bias experienced by faint stars occurs during the reduction and analysis process. When analyzing images from *K2*, the second phase of the *Kepler* mission, the bright stars are usually the ones that are reduced first. This is because they are easiest to reduce and speaks about the culture of how data reduction of *K2* data is performed. Since *K2* releases new sets of data every few months there is not enough time to deeply analyze each star before the next data set is released. However, the end of the *Kepler* mission has come and Campaign 20 will be the last campaign available. This will allow the chance for a much richer analysis of stars and fields of views that we previously did not look into. One such opportunity being crowded field of views (Barentsen et al. 2018). Seldom are faint targets isolated in a field of view and instead are commonly in crowded fields or in the halos of bright stars. These field of views often fail in the traditional code or must be masked which brings uncertainties into play. An alternative to this will be presented in this paper as an algorithm tailored to reduce multiple stars in a field of view with higher precision than previously used.

1.1 Transit Method & Light Curves

The transit method is one of several methods used in the search for exoplanets and, by far, has been the most successful method, detecting nearly 3000 exoplanets (NASA Exoplanet Science Institute 2018). Exoplanets are planets that are found outside of the solar system and can range broadly in radius, mass, density and distance from their host star. These parameters, as will be shown shortly, have profound effects on the results from a transit search and can lead to a null result even though there is a planet in that system.

In astronomy, a transit is the phenomenon where a secondary body, in our case the exoplanet, passes in front of and blocks a portion of a primary body, the star that the exoplanet orbits. When the planet covers a part of the star we observe a decrease in brightness, or flux, because the planet is blocking the light of the star from reaching us. Figure 1.4 illustrates how an exoplanet can affect the light we see from its host star over one orbital period. If I begin analyzing this figure when the planet is directly in front of the star, or when the planet is transiting the host star, there is a corresponding drop in flux called the transit depth (δ). The transit depth is the ratio of the surface area of the star's disk blocked out by the planet's disk (the star and planet are considered disk because they are the 2 dimensional projection of a 3 dimensional object) and is maximized when the planet's disk is completely overlapping with the star's disk. However, when the planet is no longer directly in front of the star the planet now contributes to the flux that we see because the planet's atmosphere now reflects light from the star into our direction. This effect is maximized right before the planet goes behind the host star. Finally, occultation is the event that when the planet is hidden by the host star in which case we only see the flux from the star alone.

With each subsequent transit we observe we increase our confidence that the periodic dip in light is caused by a planet and not another phenomenon, such as star spots, stellar activity, etc. This is the same phenomenon that occurs when the Earth falls in the shadow of the Moon, known as a solar eclipse, and the Sun temporally appears dimmer than it is, depending how much of the Sun is blocked by the Moon from the observer's perspective. A transit, however, does

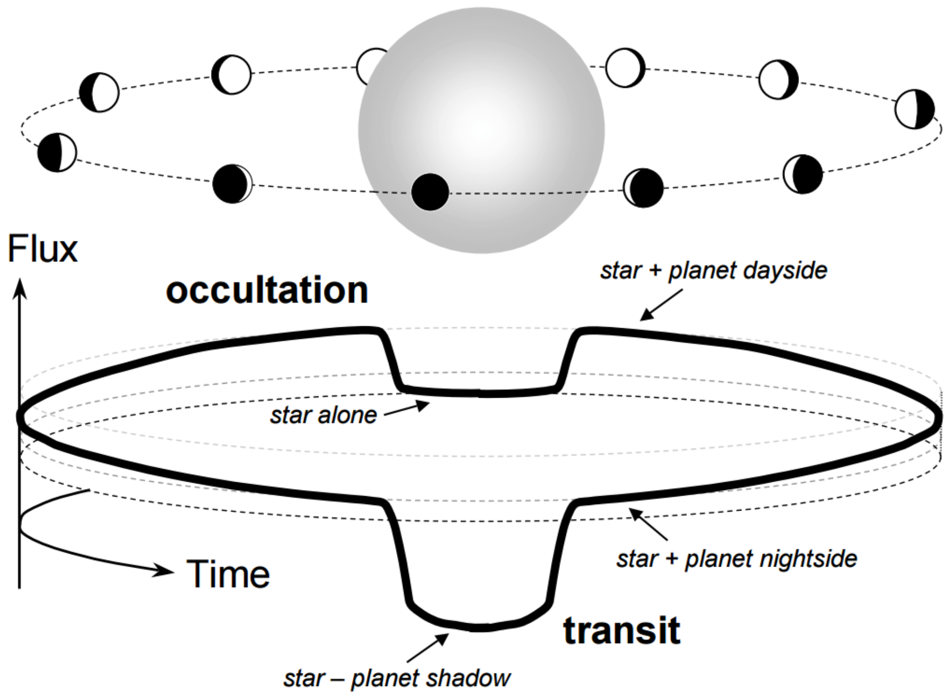

Figure 1.4: Figure showing the occurrence of a transit as well as the effect of the planet on observed brightness. Adopted from (Winn 2010).

not occur in every planet hosting system. A transit requires a special alignment with respect to the observer and has a probability of occurring estimated using the equation but also an occultation, when the secondary body is fully hidden behind the primary body. It is important to note $\frac{R_*}{a}$, where R_* is the radius of the star and a is the semi-major axis.

When able to observe transits we are able to uncover various information about the planet's atmosphere, bulk composition and much more. For instance, from analyzing the transit depth we are able to determine the radius of the planet, if the radius of the star is known, from Equation 1.1

$$\delta = \left(\frac{R_p}{R_*}\right)^2, \tag{1.1}$$

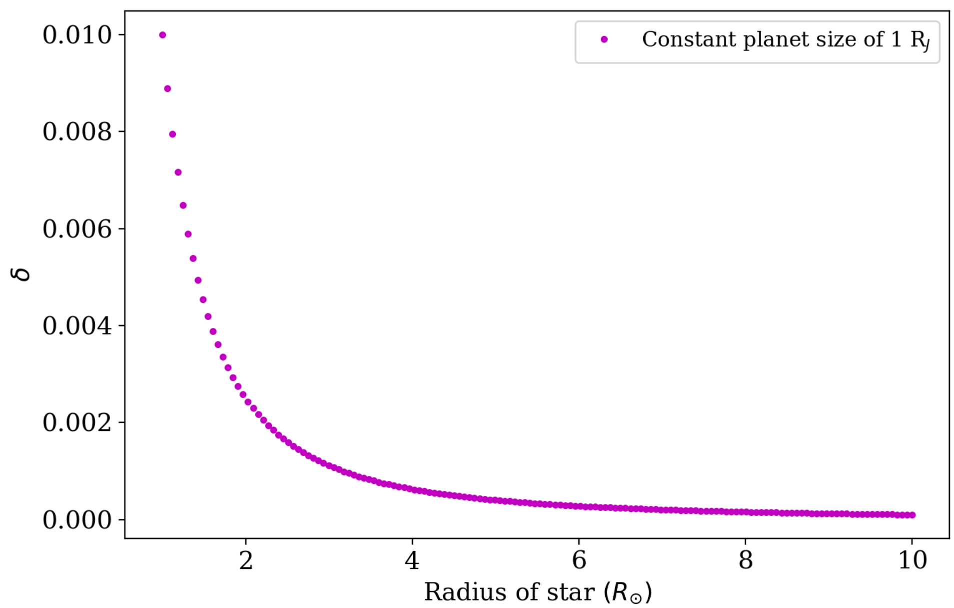

Figure 1.5: Figure illustrating the relationship of the radius of a host star and the planetary transit depth while assuming constant planet radius.

where R_p is the radius of the planet. Herein lies the premise for searching for exoplanets around small stars; keeping the size of the planet constant, a smaller star will allow for a relatively larger transit depth making detection easier, see Figure 1.5. Discoveries, such as WD 1145+017 (Vanderburg et al. 2015), have sparked the interest of the exoplanet community and provided a basis for saying that there are planets around white dwarfs, a stellar core remnant composed mostly of carbon and oxygen and supported by electron-degenerate pressure, and that we can detect them. However, as we focus on the faintest stars there will be more coincident faint targets or target stars in the halo of a brighter star, making them cumbersome to analyze. In order to understand why this traditional pipeline for analyzing stars is not as efficient for faint stars, and especially in an area with a high density of stars, we must look into what the pipeline is doing and why. But first we will describe where the data we reduce is coming from.

1.2 Kepler/K2

The *Kepler Space Telescope* was launched in March of 2009 with its original mission geared towards constantly observing a single field of view, of around 150,000 stars, located in the constellations of Cygnus and Lyra (Borucki et al. 2010). *Kepler* provided high precision photometric measurements for four years until the loss of two reaction wheels. This marked the end of the original mission for *Kepler* and in 2014 the second phase, *K2*, was initiated in order to take advantage of the remaining functionality of the telescope. The *K2* mission was designed to use the transfer of momentum from photons coming from the Sun to solar panels on the telescope to stabilize the spacecraft's pointing and drift, allowing it to point near the ecliptic, sequentially observing fields, campaigns, as it orbits the Sun. Each campaign can be continuously observed for approximately 75 days. This strategy observes a different set of 10,000 targets in each campaign (Howell et al. 2014) and offers a photometric precision approaching that of the original *Kepler* mission.

While the two remaining reaction wheels control two axes of movement the radiation pressure from the Sun, along with an approximately 6 hour periodic thruster firing events to mitigate radiation pressure, control the third axis and allows for free movement of the telescope. Although the *K2* mission has successfully re-purposed *Kepler* its photometric precision is not as high as before and the drifting and thruster events create a problem that must be accounted for during the reduction process. This is what calls for a new and different pipeline than what was needed during the original mission. More on the *K2* data reduction and the detrending process will be discussed in §1.3.

1.3 Traditional Code

Although the *K2* mission is an innovative and resourceful re-purpose of *Kepler* it has introduced errors, such as errors induced from the telescope drifting as it observes and the periodic thruster events to correct for this drifting, that must be corrected during the reduction process to recover photometric precision. The pipeline that we use to accomplish these corrections is a publicly available pipeline[4] by Van Eylen et al. (2016). The pipeline can be broken into several main sections that details what the procedure of the code:

I. Aperture photometry:

The data used is downloaded from the Barbara A. Mikulski Archive for Space Telescopes (MAST) archive[5] in *K2* pixel mask files. The code takes the files and takes the sum of the flux per pixel over the full time series of the observation. Subsequently, the median flux over the different pixels is calculated and is used as an estimate for the background flux. This is an acceptable assumption to make as long as the field is not excessively crowded, this will be discussed further in Chapter 2. The target star in the field is then found by grouping together spatially adjacent pixels that exceed the background flux by some amount (usually 1.05 $\times$ median). If multiple spatially disjointed groups are detected, the group with the most pixels is selected while the others are ignored. The total flux in the aperture is then calculated over the full time series and the flux centroid position is calculated based on the flux-weighted mean x and y coordinates of the group of pixels. Because there is a thruster event approximately every 6 hours there

[4]https://github.com/vincentvaneylen/k2photometry
[5]https://archive.stsci.edu/k2/data search/search.php

are flux outliers. To remove these the differences between subsequent center-of-fluxes are calculated (Van Sluijs & Van Eylen 2018) and used to iteratively clip 3σ outliers. Afterwards, the background flux, which is estimated as the median of the pixels, is subtracted.

II. Light curve detrending:

To detrend the light curve, decorrelate the movement of the spacecraft from the field of view, the code uses a linear fit to centriod positions. What this does is truncate the light curve into lengths (typically 300 data points) for which a polynomial function of centroid position and time is fitted to the flux in each chunk. To separate out long-term astrophysical variations from *Kepler's* drift every approximately 6 hours, a process called self-flat-fielding is performed. This is the process of taking each chunk and dividing the light curve by the model.

III. Transit search:

In order to detect periodic events in the light curves a box least square algorithm[6], developed by Ruth Angus and Dan Foreman-Mackey, is implented. This algorithm is able to detect and extract transit-like events to allow us to view and validate if they are planetary transits and if there are any remaining transits in the system.

Although this code functions well it does have restrictions placed on it by an underlying assumption. The assumption being that the median flux is used as an estimate for the background flux, which restricts ones use of this code to fields with relatively few stars. To circumvent this one may mask other stars in the field, however, this does not eliminate the contribution of other stars surrounding the target star. This is an especially significant problem with faint stars. The process we use to accurately analyze a crowded *K2* field is to model all of the stars in

[6]https://github.com/dfm/python-bls

the field and determine parameters to reconstruct, and therefore, analyze them individually. The steps used to model, find best fit parameters and reconstruct a target star will be outlined in Chapter 2.

Chapter 2

Kepler Data

2.1 *Kepler/K2* Mission

The *Kepler Space Telescope* was launched in March of 2009 and lasted until October 2018 despite encountering mechanical challenges along the way. Over *Kepler's* lifespan it has provided unrivaled precision in the visible spectrum. One of the most influential challenges faced by the *Kepler* team was the failure of a second reaction wheel, leaving two functional wheels remaining, on the spacecraft in 2013, forcing the transition to a second phase of the mission to enable continued scientific observations.

Beginning in 2014 the data for the second phase, *K2*, was released in 20 campaigns each lasting approximately 80 days, barring campaign 20 which was stopped short because of exhausted fuel, due to constraints set by the angle the solar panels of the spacecraft make with the Sun. Figure 2.1 shows the importance of the angle the spacecraft's solar panels makes with the Sun to avoid sunlight from entering the telescope, which would saturate the pixels on the CCDs (charged coupled devices). Figure 2.1 also shows a cartoon of how *Kepler* balances the pressure from solar photons to ensure stable pointing for high photometric precision. Note that all of the patches of sky *Kepler* has observed are in the ecliptic plane, the apparent plane of motion of the Sun as seen from the Earth during orbit.

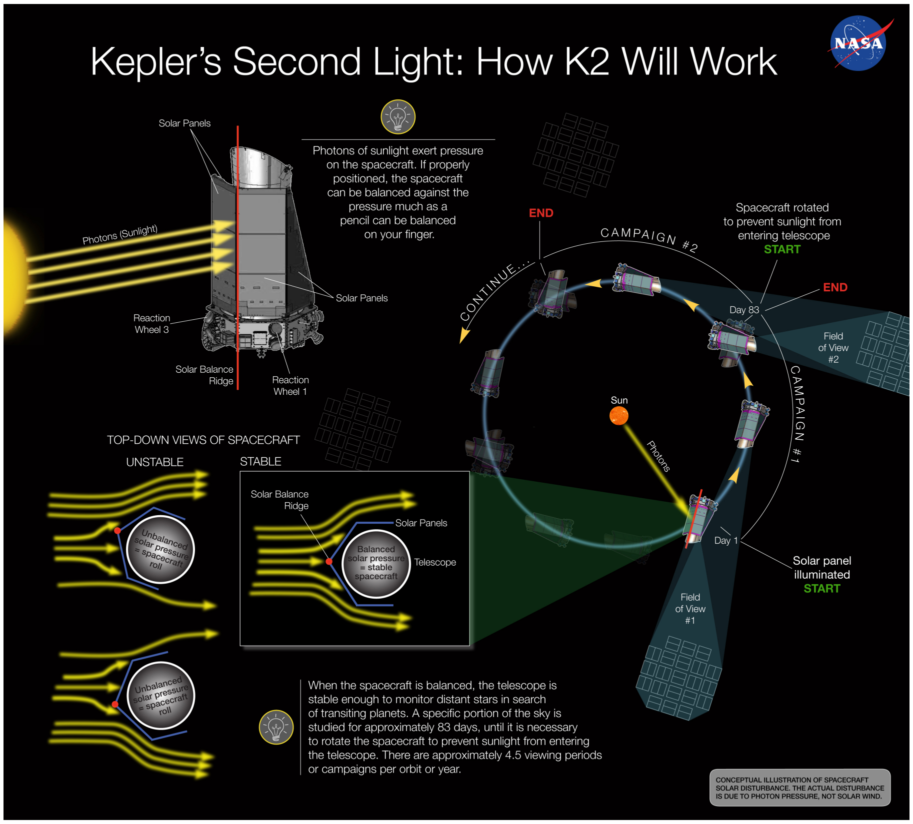

Figure 2.1: Infographic showing the *Kepler Space Telescope* during its second phase and describing how it utilizes solar photons to achieve stability for precision photometry. Credit: NASA Ames/W Stenzel

Observing in the ecliptic plane allows for the spacecraft to be controlled through a combination of thrusters and the two remaining reaction wheels to maximize stability through use of photons (Ramsay & Doyle 2014). Figure 2.2 shows a projection of the observed *K2* campaign fields. The placement of the campaign field were chosen based on community input as part of the Guest Observer program[1] while considering the spacecraft's constraints and fuel usage. Fields that overlap have fewer new targets but benefit from a longer baseline of observation for stellar activity, asteroseismology, and AGN studies, as well as revisiting *K2*

[1]The full list of targets and programs can be found at https://keplerscience.arc.nasa.gov/k2-approved-programs.html#all-k2-campaigns

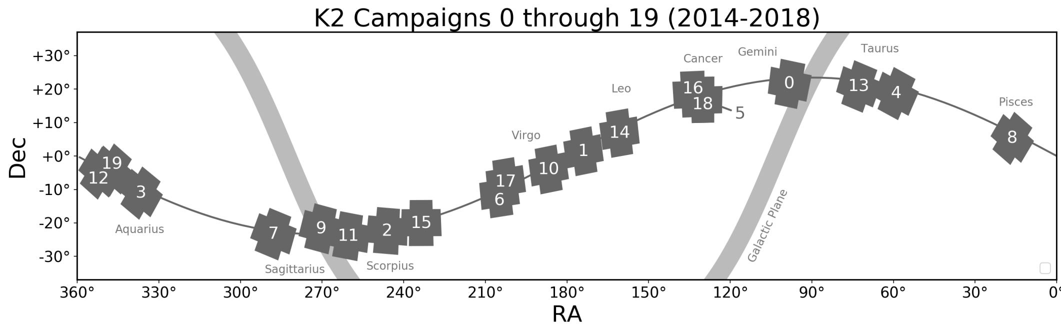

Figure 2.2: Image showing placement of campaigns 0 through 19 with the Galactic plane placed for reference. Credit: NASA Ames

planet candidates (Mayo et al. 2018).

The *Kepler* photometer was engineered while considering the original mission objective, to continuously observe over 100,000 stars in the Cygnus and Lyra constellations for three and a half years. Given this objective the photometer was crafted with one array of 42 (CCDs) and a 12 degree diameter field-of-view (FOV), approximately the size of two 2 side by side dips from the Big Dipper (credit: NASA Kepler Mission/Dana Berry). *Kepler's* FOV is large, compared to other space telescopes, so that it may observed as many stars as possible to increase the chance of an observed transit, since transit detection is rare. For comparison, the *Hubble Space Telescope's* Wide Field Camera 3 has a FOV of 7 square arc minutes while *Kepler* has a FOV of 410,000 square arc minutes. Figures 2.3 and 2.4 are full frame image (FFI) displays taken over nine years apart. The thin black lines in each image shows adjacent pairs of CCDs in a module while the thicker black lines that cross through the image are from the structures holding all the modules together which also were purposely oriented to block out the very brightest stars in *Kepler's* FOV (NASA Ames). The four black corners that are apparent in both images are where the guidance sensors are placed.

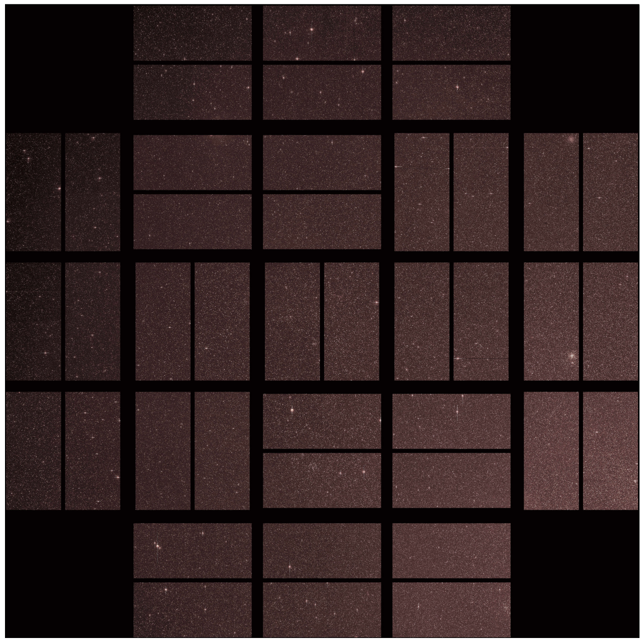

Figure 2.3: Image showing the full FOV of *Kepler* taken in April of 2009. There appears to be a gradient of brightness in the image because the lower right of the image is closer to the plane of our Galaxy, which has an abundance of stars, while the upper left is further from the Galactic plane where stars are relatively more sparse. The image is color-coded such that brighter stars are white and fainter stars are red. Credit: NASA Ames

2.2 Standard Photometric Procedure

The images that are used for photometric analysis are only a small fraction of *Kepler's* FFI, much less than 1%. The data are available for download as a flexible image transport system (fits) file by target name, region, system parameters, etc. The targets available are based on the target list generated by submitted proposals by the community. The fits files are composed of several hundred pixels of which the target star falls on some portion of those pixels. The files also contain information pertaining to the quality of the images, the name, magnitude, and location the on the sky of the target.

Figure 2.5 is a time-slice, one picture that *Kepler* took, from a fits file. This

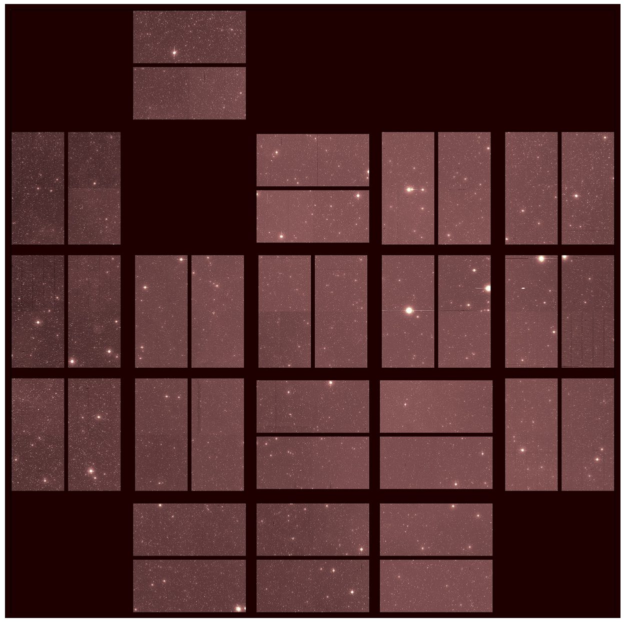

Figure 2.4: Image showing the full FOV of *Kepler* taken in September of 2018. The blackened gaps in the center and along the top of the image, which are not present in 2.3, are results of random part failures in the camera. The image is color-coded such that brighter stars are white and fainter stars are red. Credit: NASA Ames

image is less than 64 pixels which allows for a quicker download time as well as minimize file size. When we perform simple aperture photometry we only include the pixels inside the aperture, represented by the red box in Figure 2.5, which is centered on the centroid of the star, represented by the green dot. To produce a light curve we integrate over this region and across all the exposures.

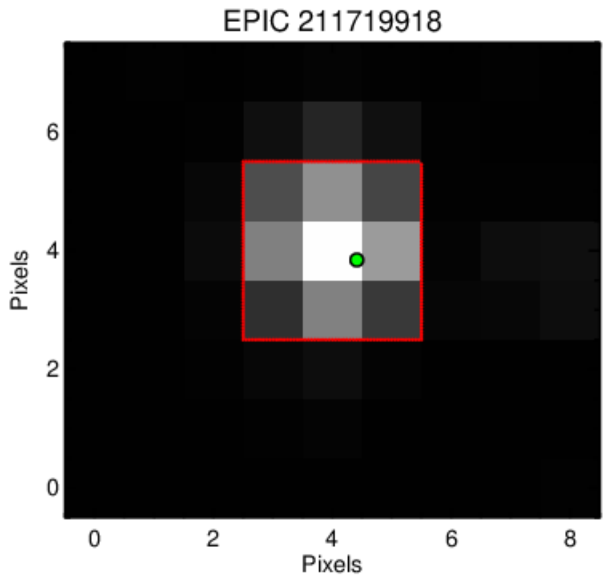

Figure 2.5: Aperture used to perform aperture photometry. The red line represents the aperture and the green dot is where the target star's centroid is.

Chapter 3

Data Reduction

The primary focus of this work is to address the problem commonly found when analyzing faint stars in the K2 field. When focusing on faint stars there are usually more coincident faint targets in view which creates the problem of blended light. Blending is a consequence of $Kepler's$ relatively large pixels ($\sim$4 arc-seconds $\approx 10^{-3}$ degrees on the sky) which allows for flux from multiple stellar objects to fall in the same pixel, mixing the light. When this event occurs with a target star we intend to perform aperture photometry on, the light from neighboring stars must be disentangled and removed to increase precision and strengthen signals present in the data. To combat these occurrences I have developed the Gaussian fitting algorithm to model faint stars in crowded fields and eliminate blending affects. The following chapter will be dedicated to exploring the notable characteristics of the algorithm.

3.1 Alternative Photometric Techniques

When the *Kepler Space Telescope* was repurposed for the *K2* mission it was necessary to balance the solar panels of the spacecraft against the solar radiation pressure to achieve photometric precision rivaling the original *Kepler*, despite operating on only two reaction wheels (Howell et al. 2014). The solar photons, however, cause the *K2* FOV to drift over time which is then corrected by thruster firing events in the opposite direction of the drift every 6-8 hours. For the *K2* mission to approach precision levels of the original *Kepler* mission the drift and thruster firing events require a more rigorous pipeline to process the data. As a solution to the issue of a drifting FOV within *K2* data there have been several open source *K2* pipelines made available.

A handful of the pipelines that have been optimized for *K2* include **K2SFF** (Vanderburg & Johnson 2014), **K2Phot** (Van Eylen et al. 2016), and **Everest** (EPIC Variability Extraction and Removal for Exoplanet Science Targets; Luger et al. (2016)). There are several other pipelines available but a complete comparison of all pipelines is outside the scope of this work. Rather, I intend to explore the aforementioned pipelines and give a general understanding of the goal of these data processing methods.

3.1.1 Everest

Everest utilizes a variant of pixel level decorrelation (PLD), a method developed by (Deming et al. 2015) for *Spitzer*. **Everest** uses a higher order of PLD given by

$$\begin{aligned} m_i = &\sum_l a_l \frac{p_{il}}{\sum\limits_k p_{ik}} + \\ &\sum_l \sum_m b_{lm} \frac{p_{il} p_{im}}{\left(\sum\limits_k p_{ik}\right)^2} + \\ &\sum_l \sum_m \sum_n c_{lmn} \frac{p_{il} p_{im} p_{in}}{\left(\sum\limits_k p_{ik}\right)^3} \; + \; \alpha \; + \; \beta t_i \; + \; \gamma t_i^2 \; , \end{aligned}$$

where m_i denotes the noise model at time t_i, p_{il} denotes the flux in the l^{th} pixel at time t_i, and a_l is the linear PLD coefficient for the l^{th} pixel. PLD is limited from higher orders to avoid overfitting. A Gaussian Process (GP) is also used to fit the stellar PSF and remove instrumental noise while preserving astrophysical signals. PLD combined with GP allows **Everest** to aid numerous photometric studies including exoplanet detection and characterization, although Luger et al. (2016) cautions its use in the presence of outliers due to the GP assuming a Gaussian noise. Deviations from Gaussianity can lead to poor fits in data if outliers are not removed before the GP.

3.1.2 K2Phot

K2Phot performs simple aperture photometry when analyzing *K2* images. **K2Phot** uses a linear fit to the multi-dimensional centroid positions X_c and Y_c (calculated relative to the mean centroid position), time T, and flux F, and fit the model M:

$$\begin{aligned} M = t_0 \ &+\ t_1 T \ +\ t_2 T^2 \ +\ t_3 T^3 \\ &+\ x_1 X_c \ +\ x_2 X_c^2 \ +\ x_3 X_c^3 \\ &+\ y_1 Y_c \ +\ y_2 Y_c^2 \ +\ y_3 Y_c^3 \\ &+\ z_1 X_c Y_c \end{aligned}$$

where t_i, x_i, y_i, and z_1 are fitting parameters determined by fitting the data with least square fit methods.

3.1.3 K2SFF

K2SFF estimates the image centroid of a drifting star and then calculate the covariance matrix between horizontal and vertical centroid positions and its eigenvectors. The image centroid is then rotated in the direction of the eigenvector with the largest eigenvalue, direction of maximum change. A fifth-order polynomial is then fit between the drift distance, while calculating the arclength along the curve, as a function of locally normalized photometric variation. A self-flat-fielding (SFF) process is then used to remove photometric variability.

All of these methods are adequate for a wide range of *K2* targets but require tweaking for specific targets due to underlying assumptions. One of which are faint stars and, to a lesser extent, crowded fields. This is the motivation for developing a PSF fitting algorithm. This algorithm will assist in the analysis of stars that will normally fail or have complications when analyzed using the aforementioned methods.

3.2 PSF Modeling

When a star, assumed to be a point source described by a delta function, is observed through a telescope with a circular aperture the resulting image is not the same as the source. The resulting image is a disk surrounded by a number of faint rings, which are produced by Fraunhofer diffraction of the light due to the optics. The diffraction of the light spreads (blurs) out any point-like object to a certain minimal size and shape. The point spread function (PSF) is the function that defines the two dimensional spread, or the diffusion, a point source undergoes when it is far smaller than the maximum resolution of the imaging device and can be approximated as a Gaussian function and characterized by the full width at half maximum (FWHM). In an ideal world, the PSF would be a delta function and would not spread out the point source. However, as we do not live in an ideal world there are factors to consider that will contribute to the PSF. Figure 3.1 illustrates the spreading a point source undergoes as it is affected by a PSF. Because *Kepler* is in space we do not have to consider factors associated with Earth's atmosphere, such as turbulence and refraction. However, we still have to consider the effects of the optics within the telescope, such as aberrations from the shape of the mirror and how the image is focused, to understand how the PSF is affected. It is important to note that because of *Kepler's* emphasis on photometry the focal plane geometry was not optimized for compact PSFs but intentionally defocused to spread light out across more pixels to provide excellent photometric images. The goal of this work is to efficiently model the PSF of faint stars which has been shown to be more effective for faint stars in crowded *K2* fields (Montet et al. 2017; Libralato et al. 2016).

Modeling the *Kepler* PSF is challenging due to a poorly understood flat field (Montet et al. 2017). To model faint stars in the field I assume a normal distribution described by a Gaussian function:

$$f(x,y) = Ae^{-\left(\frac{(x-x_o)^2}{2\sigma_x^2}+\frac{(y-y_o)^2}{2\sigma_y^2}\right)},$$

where A is the amplitude, $x_o,\ y_o$ is the center and $\sigma_x,\ \sigma_y$ are the x and y spreads of the distribution. To determine initial conditions for the function I utilize information within the fits files to produced averaged values of the image. To ensure quality is preserved the process of modeling the star field is performed again for each exposure, accounting for the jitter induced due to the two-wheel era of *Kepler*. Modeling the PSF is useful because it allows us to increase our sensitivity and analyze stars individually. Figure 3.2 presents an example of a simple Gaussian distribution centered at (0, 0). The flux of a star is equivalent to the amplitude

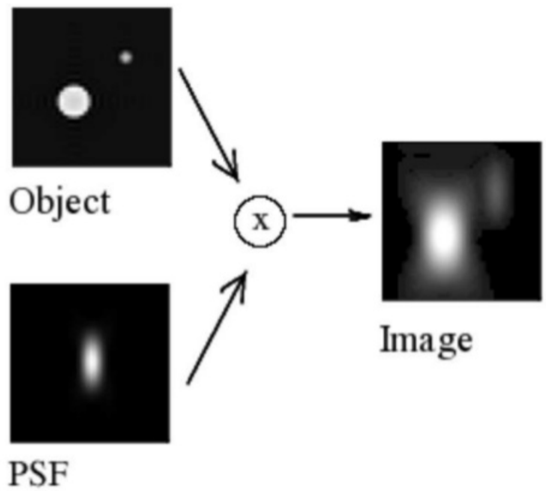

Figure 3.1: Infographic detailing the process of convolution, symbolized as $\otimes$. Convolution is the mathematical operation of two functions to produce a third function that expresses how the shape of one is modified by the other. Credit: Scientific Volume Imaging

where the centroid of the star, its center, corresponds to the centroid of the Gaussian which correctly suggest that the flux must decreases as we move from the center of the star.

When I model the stars I subtract the background level, assumed to be the median of the image (Van Eylen et al. 2016), from the image and model the system as such. This prevents the algorithm from fitting the background level as well as the Gaussian, which in most cases results in over/under approximations in other areas of the curve. Subtracting the background does not decrease our effectiveness because what matters are the flux counts within the aperture set around the star, which is where the Gaussian distribution will be. The process of subtracting the background before fitting the data with a Gaussian will be discussed in further detail in Chapter 5.

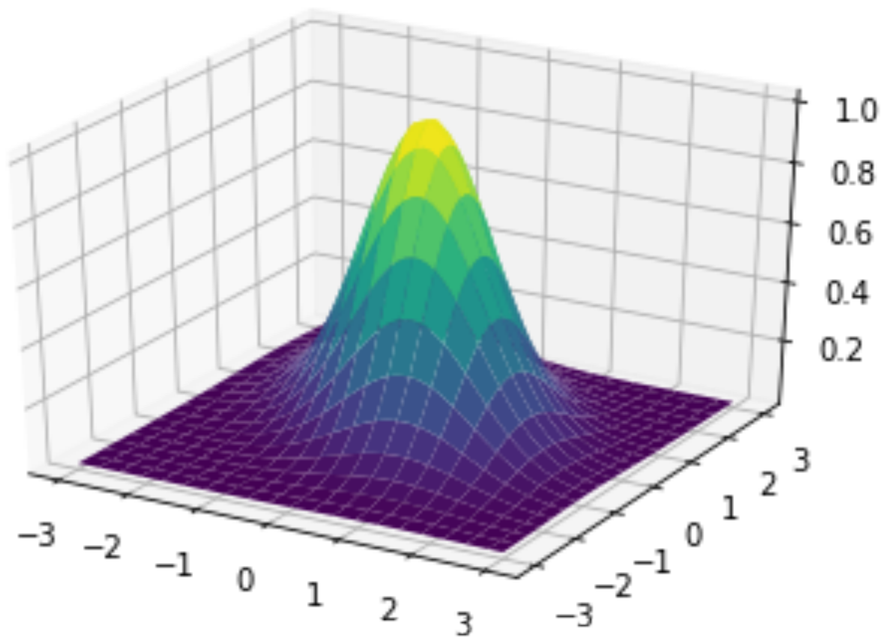

Figure 3.2: Figure illustrating a simple Gaussian distribution centered at (0, 0). The vertical axis represents amplitude which, in our case, has a maximum value of 1.

3.2.1 Raw image analysis

Included in the target pixel files (TPFs), 30-min or 1-min cadence images around targets used to perform photometry on the raw or calibrated data, are records of instances when certain spacecraft events occurred, or when an interesting phenomenon was flagged and are saved under a data column called "QUALITY." Most quality flags only provide information about the event based on flag number and do not notify if the data is not usable. To maintain a stream lined and automated process all images that raise a flag or not included in the analysis.

To determine the number of stars there are in an image the algorithm utilizes the module "IRAFStarFinder" which searches for local density maxima within the constraints of defined bounds and thresholds. These thresholds include constraints on the PSF full-width at half-maximum, background level thresholds, etc. Therefore, our sensitivity to the number of stars in the field is tied to the capabilities of IRAFStarFinder.

Chapter 4

Sample Analysis

The sample of white dwarfs (WD) used in this work are from Campaigns 15, 17, and 18. Campaigns 16 and 19 were not analyzed due to a time constraint but in the future we will incorporate these campaigns. The motivation for analyzing campaigns 15, 17 and 18 comes from a similar analysis that was performed on the first 14 campaigns by Van Sluijs & Van Eylen (2018). From this data we focused on the long cadence (LC) data, defined as having a nearly 30 minute exposure. The data used in this work is publicly available for download from MAST where I searched investigation IDs from the guest observation (GO) program. The main principal investigators (PI) that submit for WD programs are Seth Redfield (Wesleyan University), James J. Hermes (University of North Carolina, Chapel Hill), Matt R. Burleigh (University of Leicester), and John Arban Lewis (Harvard-Smithsonian Center for Astrophysics).

The data downloaded from MAST is then modeled using the Gaussian PSF fitting algorithm where the output is a fits file of a two-dimensional Gaussian characterized by the parameters found during the fitting process. The fits file is then processed through the K2SFF and K2Phot detrending pipelines.

While the purpose of this algorithm is primarily driven to increase photometric precision for faint stars, particularly WDs, it may also be applied to increasing the precision within crowded fields. As explained in Chapter 3.2, I am able to

Campaign	Investigation ID	Number of Targets
15	GO15016, GO15046, GO15062, GO15903	270
17	GO17016, GO17029, GO17037	260
18	GO18016, GO18029, GO18037	961

Table 4.1: Table displaying relevant campaigns as well as the investigation IDs associated with WD observation programs

model the star field and determine the parameters that allow me to recreate the stars with a Gaussian function. This process works with multiple stars in the field as well as providing enough parameters to model each star individually. The occurrence of multiple star fields, including WDs, is a non-trivial concern and occurs with varying levels of severity. Figures 4.1 and 4.2 are provided to illustrate the frequency and severity associated with crowded fields containing white dwarfs. With this secondary purpose we did not limit ourselves to isolated fields, however, there was a restriction on the brightness we are focusing on ($K_P > 12$).

There are a number of interesting phenomenon that can appear in the light curve of a white dwarf. For instance, it may have a companion, such as a M or K dwarf star, that flares periodically. The flaring events of the companion will reveal itself as periodic increases in brightness in the light curve of the white dwarf. Another common phenomenon is rotation modulation which resembles beats within the light curve. Rotation modulation in WD can be caused by near surface activity in a convection zone can produce magnetic spots, which could reveal rotation through photometric variation (Kawaler 2015). Tables 4.2, 4.3, 4.4, and 4.5 list stars that display characteristics similar to those described above as well as other phenomenon these in their light curves. Chapter 5 will go into

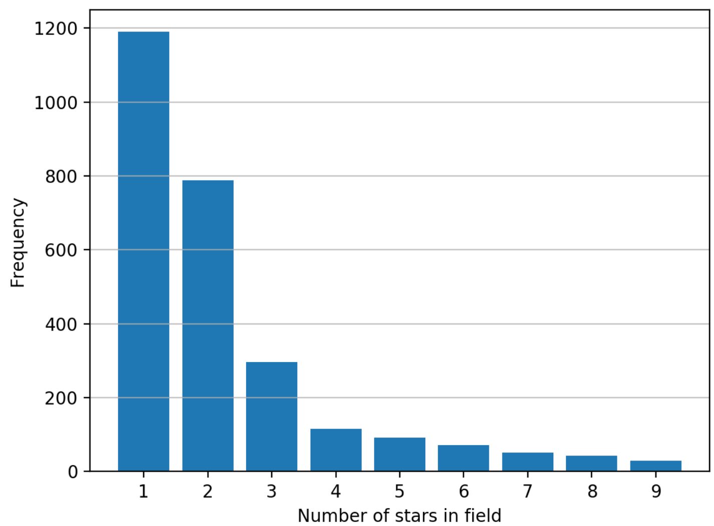

Figure 4.1: Histogram of the number of stars in the field from campaigns 15 through 19.

depth about what the light curves for these stars look like.

It is important to note that all of the stars listed are not known with their EPIC ID in SIMBAD astronomical database. To circumvent this issue the tables below include RA and Dec so that is is possible to query these objects, if you so choose.

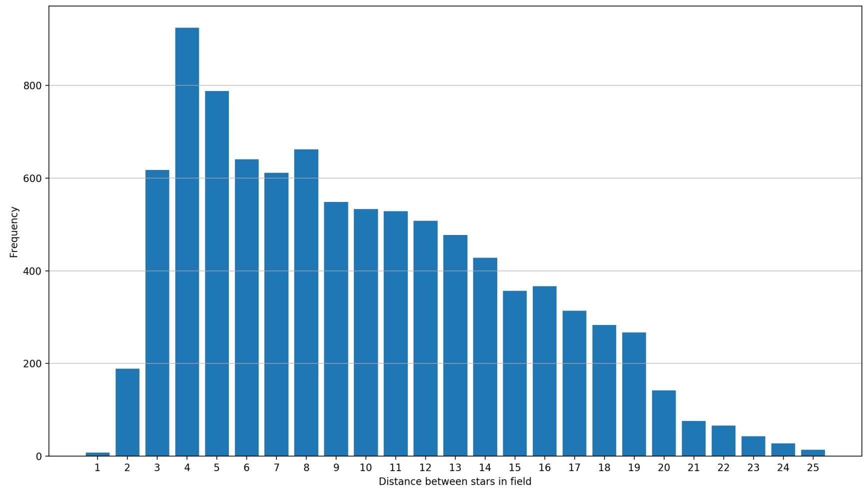

Figure 4.2: Histogram of the distribution of distances between stars in the field from campaigns 15 through 19.

EPIC ID	Campaign	RA	Dec	KEP mag
249620955	15	15 13 58.721	-20 14 45.99	14.155
200233450	18	08 51 09.661	+11 41 45.61	-
228682348	18	08 32 53.659	+23 31 21.25	19.140

Table 4.2: List of flaring targets

EPIC ID	Campaign	RA	Dec	KEP mag
249546378	15	15 21 09.220	-21 09 30.42	17.611
200233343	18	08 51 38.692	+11 51 12.11	-
211719918	18	08 56 18.949	+16 11 03.77	15.733
211830319	18	08 42 25.265	+17 44 50.74	16.501
200233385	18	08 51 08.226	+11 47 15.16	-

Table 4.3: List of targets that have rotation modulation features

EPIC ID	Campaign	RA	Dec	KEP mag
211934173	18	08 38 45.866	+19 14 16.80	18.229

Table 4.4: List of binary targets

EPIC ID	Campaign	RA	Dec	KEP mag
200233259	18	08 51 19.027	+11 58 11.03	-
200233287	18	08 50 57.082	+11 55 15.85	-
200233454	18	08 51 32.560	+11 42 04.68	-
200233420	18	08 51 21.301	+11 44 44.32	-

Table 4.5: List of pulsar targets

Chapter 5

Discussion

In this chapter I will present case studies for common phenomenon found in campaigns 14, 17, and 18 as well as provide a comparison between a reduction with and without the PSF fitting algorithm. This chapter will not focus on the differences between the detrending pipelines, as that lies outside the scope of this work, but instead will focus on the sensitivity achieved using the PSF algorithm as opposed to simple aperture photometry.

5.1 Binaries

A binary system is a system of two celestial bodies that are bound gravitationally to each other and orbit around a barycenter. As explained in Section 1.1, when one body passes in front of another, within our line of sight, we call this phenomenon a transit or an eclipse. When we observe this phenomena in other systems and create a light curve, a graph of the output of light of the star over time, we can see the dips and raises in total light as we observe over time. While we have discussed this and how it relates to exoplanets it also has applications with binary stars where a fainter star passes in front of a brighter star, decreasing the total amount of light we see. The transits of a binary star system are almost always deeper than exoplanet transits because in a binary star system the transiting star will cover more of the host star than if a smaller planet were to orbit the same star, much like how a larger planet creates a deeper transit.

EPIC 211934173 is a spectroscopically confirmed binary system consisting of a WD and a M dwarf (Rebassa-Mansergas et al. 2010; Heller et al. 2009). Figures 5.1 and 5.2 show the light curve that is generated by the **K2SFF** and **K2Phot** detrending methods. Figure 5.2 shows the light curve with the PSF treatment while in Figure 5.1 file was treated with simple aperture photometry. There are minute differences between the light curves but each method detects the transit events.

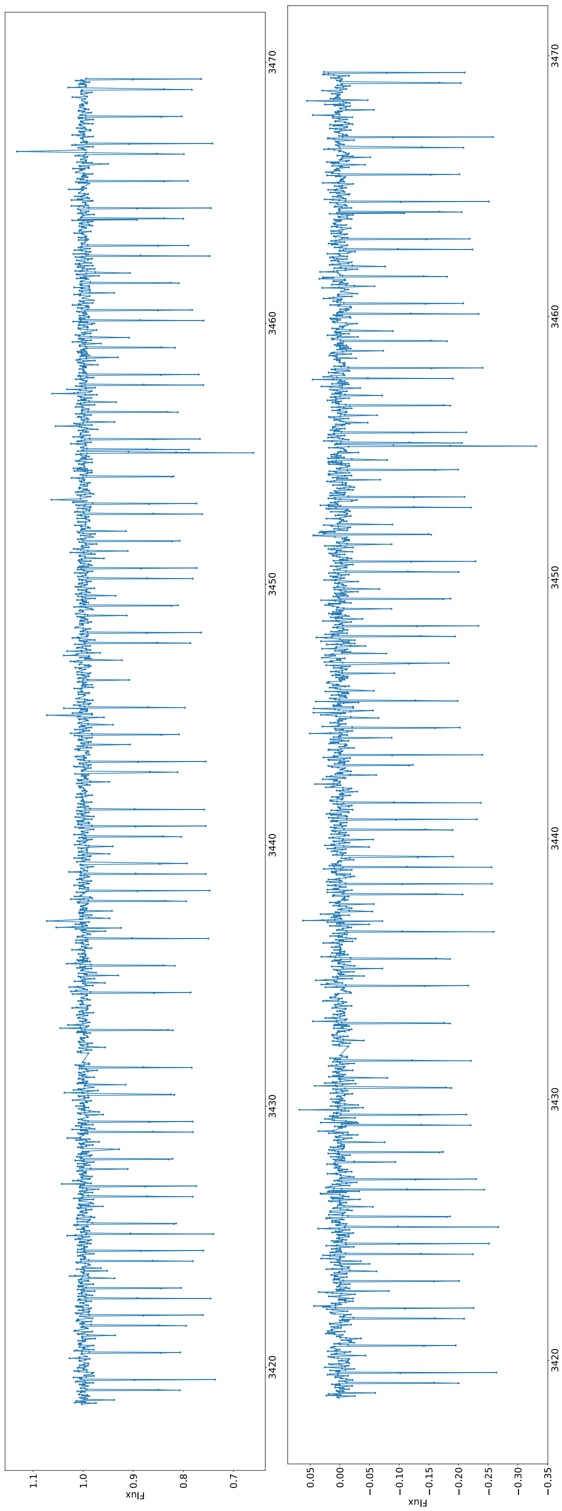

Figure 5.1: Light curve of EPIC 211934173 produced without using PSF modelling algorithm.

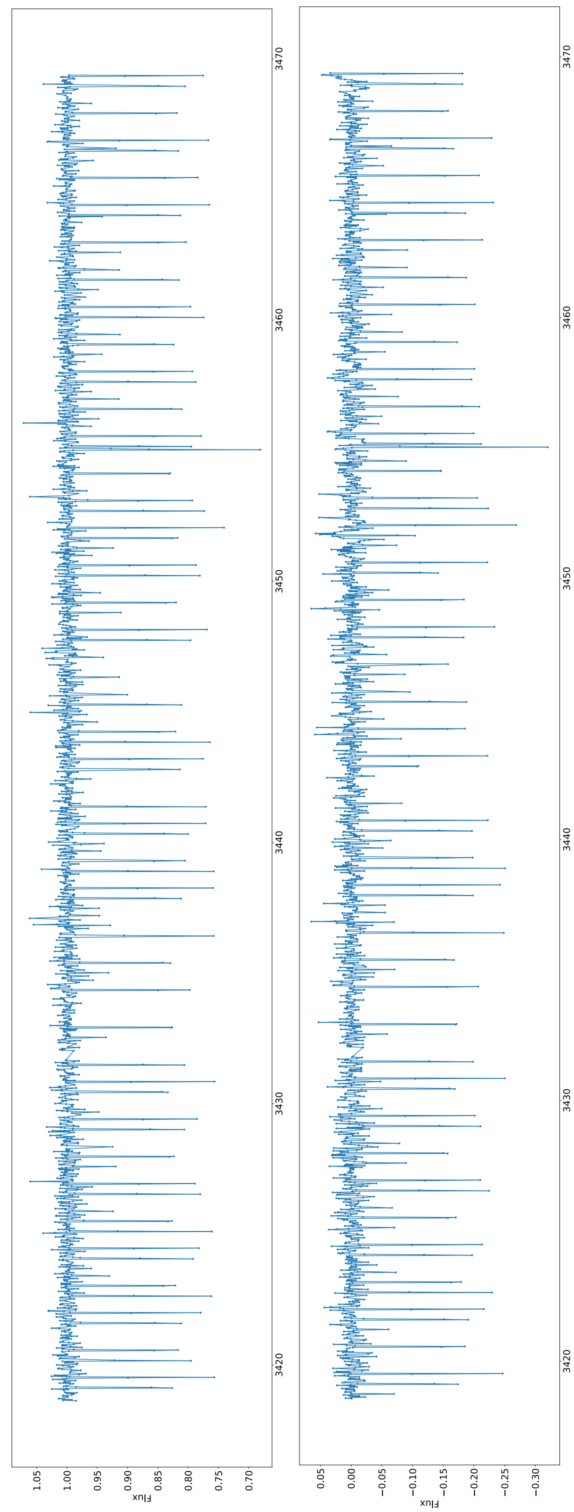

Figure 5.2: Light curve of EPIC 211934173 produced after PSF modelling.

5.1.1 Background subtraction example

Figure 5.3 shows an image from a target pixel file (TPF). The TPF contains a time-stamped sequence of uncalibrated and calibrated postage stamp pixel images of a Kepler target. The images are on a logarithmic scale to aid in identifying differences after background subtraction. In this case the median flux value was calculated to be a negative number, which is why there are pixels appearing in the right-hand image and not in the left-hand image.

Figures 5.4 and 5.5 illustrate the codes ability to fit the data with and without the background. Note that these slices were taken from the predicted centroid of the target star. Unpacking the graphs we can determine that when the function has to fit the background as well as the peak the resulting fit is less accurate. In Figure 5.5 graphs b) and d) the values are concentrated near zero rather than in the equivalent graphs in Figure 5.4. This is an important aspect of the code because it allows for higher precision by minimizing the amount of light that is not from the star.

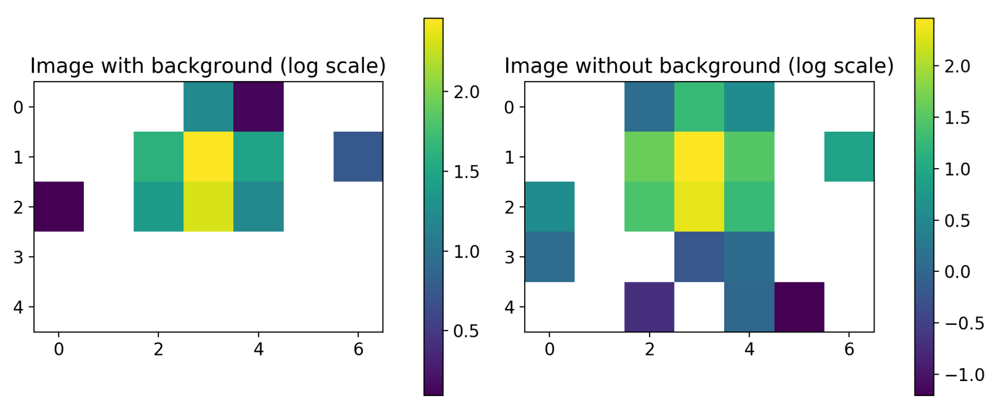

Figure 5.3: Figure comparing an exposure of EPIC 211934173 with and without the background. The image is in log scales which results in blank (white) pixels where a negative value was recorded.

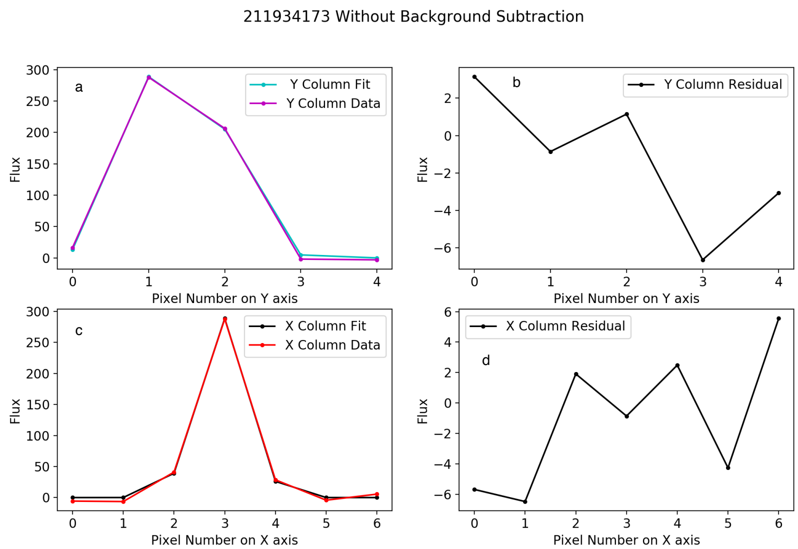

Figure 5.4: Figure comparing how well the function fits the data in a two-dimensional space. Graph a) (top left) shows the flux per pixel when fixing the x value and selecting all the pixels in that column. Graph b) (top right) shows the residual when the difference in flux between the data and the fit is taken at each pixel. Graph c) (bottom left) shows the flux per pixel when fixing y and selecting all the pixels in that row. Graph d) (bottom right) is again the residual.

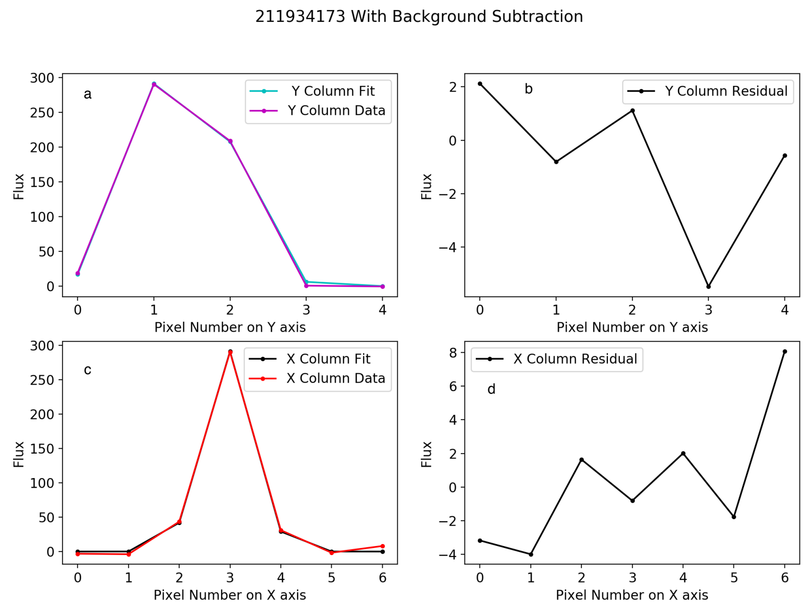

Figure 5.5: The layout of the graphs and what they portray is exactly the same as in figure 5.4 with the only change here is that a background subtraction was applied using the median of the image.

5.2 Modulation

EPIC 211719918 is a known weakly magnetic, variable DBA white dwarf (Putney 1997). However, the period of variability is not well constrained with a possible period between 2 and 24 hours (Brinkworth et al. 2013; Dennihy et al. 2017). Some earlier studies have estimated the temperature to be 21,000-28,000 K which would place it on the DB instability strip. Figure 5.6 presents a graph of the locations of the instability strip as the colored ovals (Córsico 2018). Because this graph is in log base 10 scale the range of temperature puts EPIC 211719918 between 4.3 and 4.45 on the $x-$ axis, along with V777 Her. V777 Her is also variable white dwarf star of the DBV type. On this strip there would be pulsations due

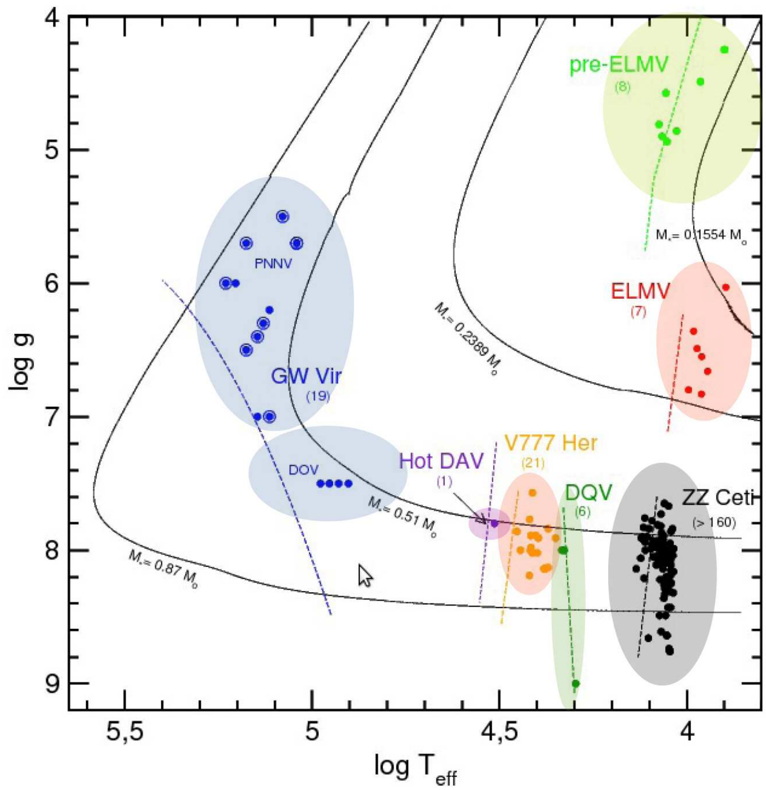

Figure 5.6: The location of the known classes of pulsating WD and pre-WD stars (dots of different colors) in the log Teff log g plane. In parenthesis we include the number of known members of each class. Dashed lines are the theoretical blue edges of the different instability domains (Córsico 2018).

to a helium partial-ionization zone (Wesemael et al. 2001). However, other work has found an effective temperature 34,520 K (Kleinman et al. 2013).

Using the PSF fitting method I was able to obtain a similar period for the modulations, between 3 and 6 hours. Figure 5.7 and 5.8 shows light curves produced where we suspect some form of modulation from the shape of the graph, which appears to have some sort of structure rather than just random noise. There are methods to show that the variability of a stars might be due to rotational modulation (Balona et al. 2019), however, that is not the primary focus of this work.

Figure 5.7: Light curve of EPIC 211719918 produced without using PSF modelling algorithm.

Figure 5.8: Light curve of EPIC 211719918 produced after PSF modelling.

5.3 Pulsations

EPIC 200233287 is a pulsating white dwarf which is a WD whose brightness varies due to gravity wave pulsations and thermal processes that generate pulsations within itself (Córsico 2018). Generally, pulsating WDs have multiple modes of pulsation and therefore have more than one period. In some cases, the different frequencies lead to constructive and destructive interference, such as in the light curve of EPIC 200233287 shown in Figures 5.11 and 5.12. The pulsations are envelopes in the data where there appears to be a defined structure and a periodic changing of brightness lasting between a timescale of hours to a few days.

The detrending pipelines produces two other ways to see the variations and pulsations from a star, one of which is by looking at phase folded data. Figure 5.9 has been phase folded on four different periods chosen by a box least-squares algorithm[1]. An example of a pulse can be seen in the graph second from the top (blue) where there is an underlying structure to the data.

Figure 5.10 shows two power spectrums of EPIC 200233287. A power spectrum is a graph of frequency/power data, converted from time/flux data by Fourier Transform, where modes appear as sharp peaks. White noise, for example, contains all frequencies at the same power so its power spectrum would be flat. However, the power spectrum for a key on a piano will have have a high value at the frequency corresponding to its note, but low values elsewhere. This translates to a power spectrum for a pulsating star having high values for times where it is pulsating and lower values during other times. Figure 5.10 illustrates this behavior and suggests that there is some pulsating occurring in the star.

[1] https://github.com/dfm/python-bls

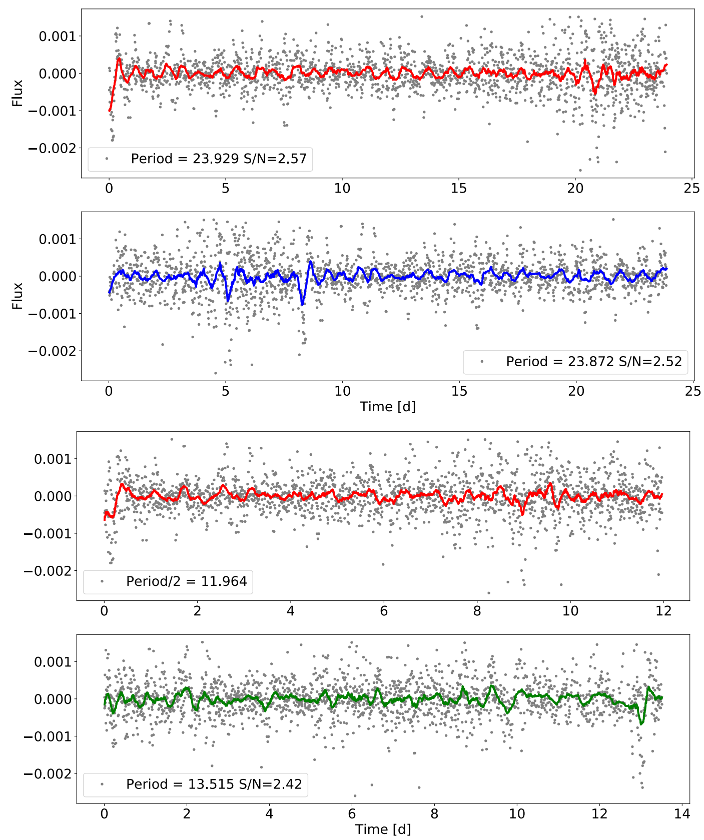

Figure 5.9: Figure showing data after it has been phase folded to the best fit periods.

A crucial aspect from these results is that our precision is tied to the choice of our detrending method, as seen from the included light curves. Future work will be dedicated to understanding why there is a dependence, more so with the **K2SFF** pipeline, on a chosen pipeline and its potential impact on data analysis.

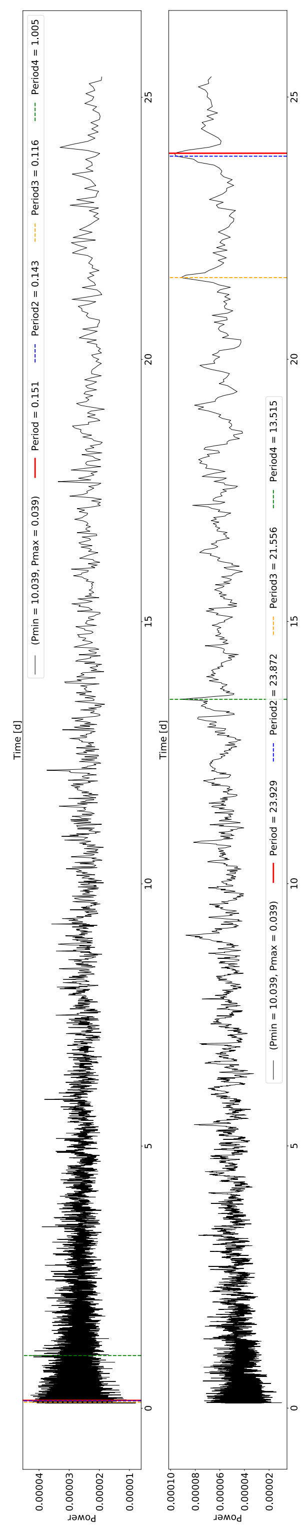

Figure 5.10: Power spectrum of EPIC 200233287 produced without PSF modeling (top) and with PSF modelling (bottom).

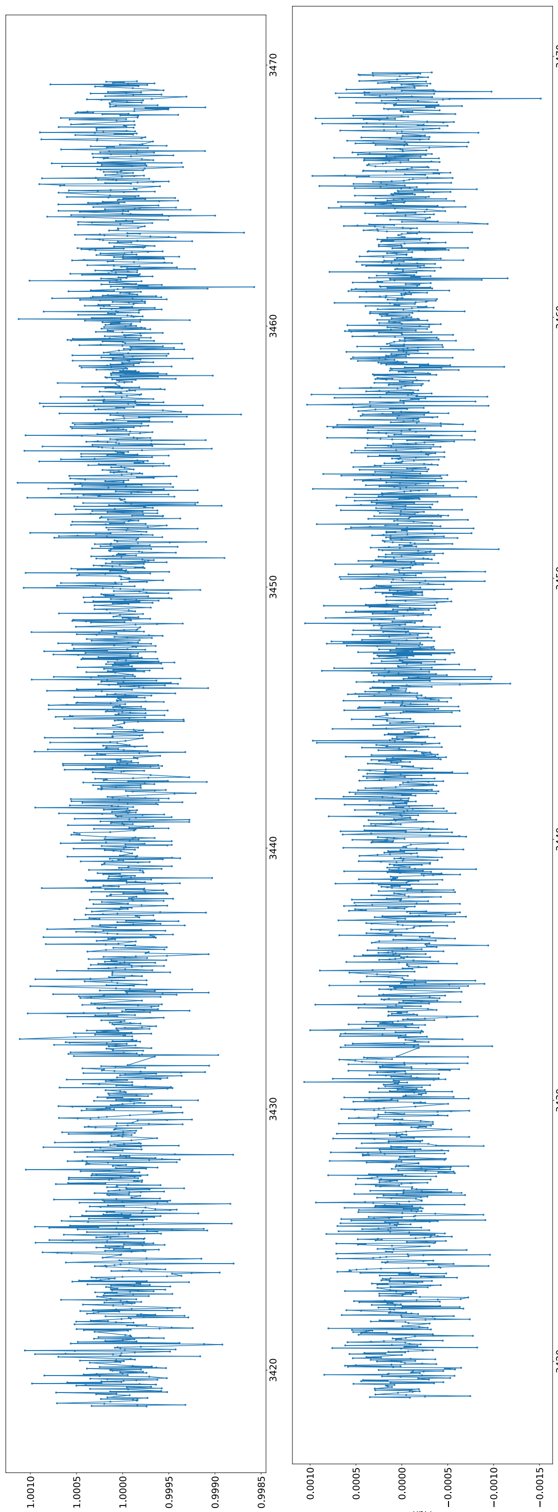

Figure 5.11: Light curve of EPIC 200233287 produced without using PSF modelling algorithm.

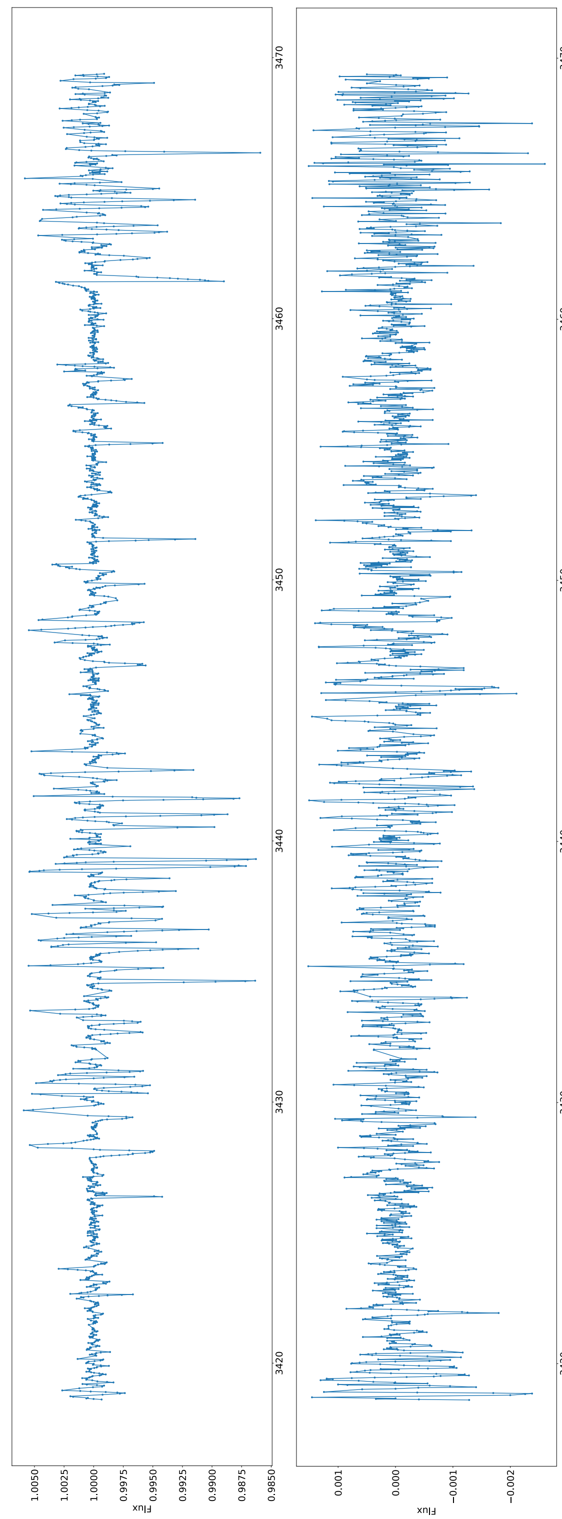

Figure 5.12: Light curve of EPIC 200233287 produced after PSF modelling.

Chapter 6

Conclusion

This paper explored an alternative to traditional aperture photometry methods used to analyze faint stars in crowded K2 fields. This method was used a Gaussian fitting algorithm on single star fields as well as multi-star fields with known transiting exoplanets. We have shown the capabilities of the algorithm through various case studies that are to be expected when analyzing a sample of WDs. We have also discussed the alternative applications for the algorithm as well as the direction of future work may go.

The transit method has proven to be a very accomplished planet detection method, greatly due to the innovations of the *Kepler Space Telescope*. It has provided an abundance of pristine photometric images over a long period of time and because of this not all of the data has been fully exploited. It has very much been akin to drinking water from a fire-hose due to data releases happening nearly three months apart. This high volume of new data lead to surface level analysis the majority of the time, leaving more difficult and time expensive targets to be left for another time. Now that the *K2* mission has come to an end we finally have the time to look back on the wealth of data amassed over nearly nine and a half years and begin to interrupt every aspect of this phenomenal data. The algorithm introduced in this work aims to aid in this process and take full advantage of the data available.

As it stands the code to perform the PSF fitting is functional and performs well with single star fields. It has also been tested with multi-star fields, which were not WDs, and it still performed well in separating and fitting the PSF of each star. Although the fitting algorithm is functional there are areas that can be approved on. To begin we can make the process of background subtraction a more sophisticated and rigorous process by accounting for all the stars. This would allow for a more accurate estimate of the background noise, leading to a higher signal to noise. Another area for improvement is to iteratively perform the Gaussian fitting process to ensure convergence on the optimal parameters.

Future work will focus on utilizing short cadence data. As of now not all WD are taken with short cadence exposures, however, for the targets that do have it available it would be advantageous to analyze. Short cadence WD observations are advantageous because their transit depths would be less diluted due to their relative transit times, on the order of minutes. If a WD system, with an exoplanet transit time of $\sim$2 minutes, is observed with a long cadence exposure then the system has at most 28 minutes of out of transit observing. What this does is increase the overall light collected and makes it more difficult to detect the decrease of light caused by a transit. A short cadence observation would not have this problem and would allow us to be more sensitive to systems like this.

There will also be work put into testing the limitations of the algorithm in crowded fields and when there is a negligible benefit due to large pixel separations. Determining the limitations on the effectiveness in a crowded field will allow us to understand the practicality of the code and potential applications outside of WD surveys. Although we do expect a higher photometric precision for faint stars in isolated fields, as opposed to simple aperture photometry, placing a limit on when the code is necessary will cut down on computation time.

Bibliography

Balona, L. A., et al. 2019, , 485, 3457

Barentsen, G., Hedges, C., Saunders, N., Cody, A. M., Gully-Santiago, M., Bryson, S., & Dotson, J. L. 2018, arXiv e-prints, arXiv:1810.12554

Beaulieu, J. P., et al. 2006, Nature, 439, 437

Borucki, W. J., et al. 2010, Science, 327, 977

Brinkworth, C. S., Burleigh, M. R., Lawrie, K., Marsh, T. R., & Knigge, C. 2013, The Astrophysical Journal, 773, 47

Córsico, A. H. 2018, in Third Conference On Stellar Astrophysics, 13–22

Deming, D., et al. 2015, , 805, 132

Dennihy, E., Fuchs, J. T., Hermes, J. J., Gnsicke, B. T., Hollands, M. A., Gentile Fusillo, N. P., Raddi, R., & Redfield, S. 2017, Monthly Notices of the Royal Astronomical Society, 468, 1946

Heller, R., Homeier, D., Dreizler, S., & Østensen, R. 2009, , 496, 191

Hooper, D., Cholis, I., Linden, T., Siegal-Gaskins, J. M., & Slatyer, T. R. 2013, , 88, 083009

Howell, S. B., et al. 2014, PASP, 126, 398

Kawaler, S. D. 2015, in Astronomical Society of the Pacific Conference Series, Vol. 493, 19th European Workshop on White Dwarfs, ed. P. Dufour, P. Bergeron, & G. Fontaine, 65

Kleinman, S. J., et al. 2013, , 204, 5

Libralato, M., Bedin, L. R., Nardiello, D., & Piotto, G. 2016, , 456, 1137

Luger, R., Agol, E., Kruse, E., Barnes, R., Becker, A., Foreman-Mackey, D., & Deming, D. 2016, , 152, 100

Mayo, A. W., et al. 2018, , 155, 136

Montet, B. T., Tovar, G., & Foreman-Mackey, D. 2017, , 851, 116

NASA Exoplanet Science Institute. 2018, NASA Exoplanet Archive

Ng, C. W., Takata, J., & Cheng, K. S. 2016, , 825, 18

Putney, A. 1997, The Astrophysical Journal Supplement Series, 112, 527

Ramsay, G., & Doyle, J. G. 2014, , 442, 2926

Rebassa-Mansergas, A., Gänsicke, B. T., Schreiber, M. R., Koester, D., & Rodríguez-Gil, P. 2010, , 402, 620

Struve, O. 1952, The Observatory, 72, 199

Tsapras, Y. 2018, Geosciences, 8, 365

Van Eylen, V., et al. 2016, The Astrophysical Journal, 820, 8

Van Sluijs, L., & Van Eylen, V. 2018, Monthly Notices of the Royal Astronomical Society, 474, 4603

Vanderburg, A., & Johnson, J. A. 2014, Publications of the Astronomical Society of the Pacific, 126, 948

Vanderburg, A., et al. 2015, Nature, 526, 546

Wesemael, F., Liebert, J., Schmidt, G. D., Beauchamp, A., Bergeron, P., & Fontaine, G. 2001, , 554, 1118

Winn, J. N. 2010, arXiv e-prints, arXiv:1001.2010

Wolszczan, A., & Frail, D. 1992, Nature, 335, 145

Yaqoob, T. 2012, Exoplanet Research Paper Statistics and Sociology (New Earth Labs (Education and Outreach))