Pixel Level Decorrelation for CHEOPS

Pedro Silva

Mestrado Integrado em Engenharia Física

Departamento de Física e Astronomia

2019

Orientador

Inv. Dr. Olivier Demangeon, Centro de Astrofísica da Universidade do Porto

Coorientador

Prof. Dr. Nuno Santos, Faculdade de Ciências da Universidade do Porto

U.PORTO

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, ______/______/________

UNIVERSIDADE DO PORTO

MASTERS THESIS

Pixel Level Decorrelation for CHEOPS

Author:
Pedro SILVA

Supervisor:
Olivier DEMANGEON

A thesis submitted in fulfilment of the requirements for the degree of MSc. Engineering Physics

at the

Faculdade de Ciências da Universidade do Porto
Departamento de Física e Astronomia

December 20, 2019

" As the joke goes in astronomy, the USA actually has several Hubble-class telescopes, it's just most of them are pointing down. "

Reddit comment by Andromeda321

Acknowledgements

Thank you all.

UNIVERSIDADE DO PORTO

Abstract

Faculdade de Ciências da Universidade do Porto
Departamento de Física e Astronomia

MSc. Engineering Physics

Pixel Level Decorrelation for CHEOPS

by Pedro SILVA

Pixel Level Decorrelation (PLD) [1–3] was successfully implemented for the CHaracterising ExOPlanets Satellite (CHEOPS) [4] from ESA. PLD is a fast and powerful reduction tool used to remove the instrumental noise from astrophysical data (transit photometry). To do so it relates fluctuations of the stellar intensity directly to the individual pixels. The high number of pixels in CHEOPS translates to an high number of regressors which were constrained from overfitting through regularization. Gaussian Processes are used to capture the astrophysical signal, thus completing the PLD model of the observed data. The biggest source of instrumental noise in CHEOPS is caused by its rotating ($\sim$ 100 minutes period) field of view that, among others problems, induces a small pointing jitter.

Since the satellite is yet to be launched, two simulated datasets [5] (WASP-43b and WASP-18b) were used to access the PLD for CHEOPS (PfC) ability to detrend CHEOPS data. The PfC was capable of removing most of the instrumental noise in both simulations, with the 100 minutes trends being gone from the data. The detrended light-curves had the proper transit depth and length. For WASP-18b the PfC achieved a photometric precision of 50.86 ppm and 139.4 ppm for the WASP-43b for 15 and 30 minutes of integration time respectively. The highest precision possible is set at the photon noise, calculated through [6]: 47.83 ppm for WASP-18b and 125.32 ppm for WASP-43b, using the integration times previously mentioned.

UNIVERSIDADE DO PORTO

Resumo

Faculdade de Ciências da Universidade do Porto
Departamento de Física e Astronomia

Mestrado Integrado em Engenharia Física

Decorrelação ao Nível do Pixel para o CHEOPS

por Pedro SILVA

Decorrelação ao Nível do Píxel (PLD) [1–3] foi implementado para o satélite de caracterização de exoplanetas (CHEOPS) [4] da Agência Espacial Europeia (ESA). PLD é um método rápido e poderoso de redução de dados, sendo utilizado para a remoção de ruído instrumental de dados astrofísicos, neste caso, fotometría de trânsito. Para o fazer, o PLD relaciona flutuações da intensidade estelar diretamente aos píxeis. O grande número de píxels presentes numa imagem do CHEOPS equivale, portanto, a um enorme número de regressores que são controlados através de regularização. Impede-se assim que estes deem *overfit* aos dados. Processos Gaussianos são utilizados para capturar implicitamente o sinal astrofísico, completando o modelo de PLD que modela os dados observados. A maior fonte de ruído instrumental no CHEOPS é a rotação do seu campo de visão (período de $\sim$ 100 minutos), que entre outros problemas introduz pequenas erros de apontamento.

Como o satélite ainda não foi lançado, os dois conjuntos de dados analisados são simulações [5], uma do WASP-43b e outra do WASP-18b. É através destas que determinamos a habilidade do PfC (PLD para o CHEOPS) de reduzir dados do CHEOPS. PfC foi capaz de remover grande parte do ruído instrumental em ambas as simulações, apagando o sinal de 100 minutos de período que lhe é característico. Em ambas as correções a profundidade e a largura do trânsito foram obtidos. Para o WASP-18b o PfC obteve uma precisão fotométrica de 50.86 ppm e 139.4 ppm para of WASP-43b para tempos de integração de 15 e 30 minutos, respetivamente. A maior precisão alcançável está definida no ruído fotão, calculado através de [6]: 47.83 ppm para o WASP-18b e 125.32 ppm para o WASP-43b, usando os tempos de integração anteriormente mencionados.

Contents

List of Figures

Chapter 1

Introduction

Are there other planetary systems with solar type stars similar to ours? Do planets preserve information of the composition and shape from the protoplanetary disk? How are planets arranged for stars bigger and smaller than ours? Is life a common byproduct of planetary formation and is life, as we know it, feasible on another planet? This and other questions can only be solved looking outwards, towards the stars and characterizing the planets that orbit around them [7]. While the task of finding an extra-solar planet is in itself a daunting one, research teams around the world have taken up the challenge to meet up the above questions. And along the way, exciting new science has been done, incredible new instruments have been built and new computational methods developed. Among those leading this charge is the European Space Agency which has one mission set to launch this year, CHEOPS (Characterizing Exoplanet Satellite) and two other in the drawing board: PLATO (Planetary Transit and Oscillations) set to launch in 2026 and ARIEL (Atmospheric Remote-sensing Exoplanet Large Survey) in 2028.

This dissertation has as its main goal the development of a data reduction pipeline for the small ESA mission CHEOPS that uses Pixel Level Decorrelation (PLD) for the removal of instrumental noise. Standard methods define and remove the correlation that exist between the fluctuations in the intensity of the star and its position in the CCD. PLD relates the fluctuations of the stellar intensity directly to the individual pixels. In doing so, PLD skips the previous two steps (centroid estimation and correlation removal) which are prone to uncertainty. PLD has already been successfully used for reducing data from the Warm Spitzer [1] that exhibits systematics due to intra-pixel sensitive variations [8]. And from Kepler/K2 [2, 3], whose fail in its second reaction wheel rendered him unable of achieving the fine pointing accuracy required for high precision photometry

[9]. This successes in two satellites whose nominal missions had ended speaks volumes as to its potential and versatility. However, CHEOPS, Spitzer and K2 have crucial design differences (e.g. CHEOPS rotating field of view) which might impair the performances of the PLD for CHEOPS (PfC). It is thus important to constrain its performances and understand its limitations when applied to CHEOPS data.

In this chapter the subject of exoplanet search and characterization will be introduced to the reader, as well as the Satellite for which this thesis is done. A brief introduction will be made for the Charge-Coupled Devices and the noise associated to their measurements. The final portion of this section will be devoted to the CHEOPS Data Reduction Pipeline as an insight to currently used methods for the treatment and correction of astrophysical data. Chapter 2 is fully devoted to Pixel Level Decorrelation, with an overview of its first implementation on the Spitzer Space Telescope and the following ones for the K2 mission. The chapter ends with an implementation for CHEOPS (PfC) where some of the methods implemented are also introduced: L-BFGS-B, Gaussian Processes and the Savitzky–Golay filter. Simulated data of WASP-18b and WASP-48b is reduced with PfC and its outputs are analyzed in Chapter 3. Chapter 4 presents the conclusions and the future work.

1.1 Exoplanets

1.1.1 First Discovery

In 1995, a Jupiter-mass companion was found orbiting a solar-type star [10] through Radial Velocity (RV) using a then new, fibre-fed echelle spectrograph ELODIE [11], from the Haute-Provence Observatory in France. This companion, whose minimum and maximum mass was constrained to 0.5 and $2\,\mathrm{M}_J$ (Mass of Jupiter) respectively, was found orbiting at 0.05 AU around 51 Peg. At that distance, the temperature at the companion's surface would be $1300\,K$, hence the target would be incredibly close to the Jeans thermal evaporation limits. Knowing that non-thermal evaporation effects are dominant over thermal ones and considering that binary stars can be arbitrarily close, this companion could very well be a striped down brown dwarf. This possibility was reinforced by contemporaneous models of planetary formation that could not explain the presence of a giant planet so close to the parent star. Thus, the authors restrained from calling this finding an exoplanet throughout the article, opting for the neutral term ”companion”. Still, they left open the possibility of this being a Jovian exoplanet who wandered during its formation closer to its host. Despite this caution, planetary migration models [12] that would make this companion an hot Jovian planet started to pop up. And the mere idea that Jupiter sized planets could finally be found set the world of astronomy ablaze, as a new age of discovery was at hand.

This finding, was a prelude for hundreds of other planets with masses, sizes, densities and orbits that defied knowledge, and display how ill equipped we were to characterize them, since no true frame of reference could be establish from the solar system alone. The past two and a half decades have been spent surveying the existence of these exoplanets. And despite our limitations, the era of physical and chemical characterization of this planets and their systems has already begun.

1.2 A search for Exoplanets

Anyone that ventures into this field of research will soon realize that a planet's light signal is incredibly weak both in relative and absolute terms[7]. When searching for an exoplanets there are two sources of information available, the host star and the exoplanet. Direct techniques use the light from the exoplanet. However, the planet's low luminosity and the small distance to a bright star represent major difficulties in this kind of search. As of now, the two main methods of finding a planet orbiting a distant star are Radial Velocity and Transit Photometry. Both of them are indirect methods, that search for small variations in starlight to infer of this planets and to characterize them. Transit photometry can also be used for direct searches, when the light of the planet is noticeable in the light-curve (phase curve and occultations) as seen in image 1.1

1.2.1 Radial Velocity

Radial velocity (RV), uses the movement of the star caused by the exoplanet orbiting around it. This reflex motion is detected by Doppler shift in the spectrum of the host star, manifesting has a redshift when the star is moving away from us and a blueshift when it is coming toward us. This shift is converted to radial velocity using $v = c\frac{\Delta\lambda}{\lambda}$[*]. Once we have the velocity of the star as a function of time, a Lomb-Scargle periodogram [13] is used as a standard procedure to discover and determine the statistical significance of the various periodic velocity changes. Once the periodicity of the reflex motion of the

[*] c is the speed of light, λ_0 the original value from a spectrum line and $\Delta\lambda_0$ the deviation from the regular value.

star due to the planet is ascertained, the Kepler 3rd law (1.1) can be used to determine the semi-major axis of the planetary orbit (a)*.

$$\tau = \frac{a^{3/2}}{\sqrt{M_\star}} \tag{1.1}$$

$$\frac{M_p V_p^2}{a} = \frac{G M_p M_\star}{a^2} \tag{1.2}$$

$$M_p = V_\star \frac{M_\star}{V_p} \tag{1.3}$$

From this distance and using 1.2, the planetary velocity is gained and with 1.3 the mass of the exoplanet if finally inferred. As the name suggests, the observer will only get the velocity projected to the line of sight, for this reason, equation 1.3 only works when the planet's orbit is in the same plane as the observer. Generally $V_\star$ encapsulates $V_* sin(i)$, where i is the tilt between the planetary orbit and the relative plane of the sky. The reflex motion is proportional to the planet/star mass ratio, hence small planets that orbit big stars are more challenging to detect. Considering a Sun-Jupiter system whose reflex motion if 13 meters per second, a variety of stellar phenomena can produce effects that mimic this motion [7].

RV is used in vast high precision surveys from the ground and claims the responsibility for the identification of hundreds of exoplanets, that, through successive improvements in accuracy, can come from ever greater orbits and be ever smaller.

In the first half of the 20^{th} century, the capture of the reflex motion of a star was either limited to photographic astrometry or photographic spectra, in other words photographic plates and films. With the introduction of solid-state, signal-generating detectors, the study of the stellar spectra, saw its error reduced to the point where it was theoretically possible to find an extra-solar planet. Boosted by this, Gordon Walker and Bruce Campbell began their RV search for an exoplanet, but where ultimately beat to it by Michel Mayor and Didier Queloz who won the 2019 Nobel Prize in Physics, for this achievement.

1.2.2 Transit Photometry

In Transit Photometry, the intensity from stellar light is scanned to see a dip in its value as seen in figure 1.1. This decrease, if periodic, might be caused by a planet that sits between the observer and its host star. Likewise, the observer and the planet's orbit need

*G is the gravitational constant, $M_\star$ is the stellar mass which was previously determined. τ is the periodicity, M_p is the planetary mass, V_p and $V_\star$ denote the exoplanet and stellar velocity respectively

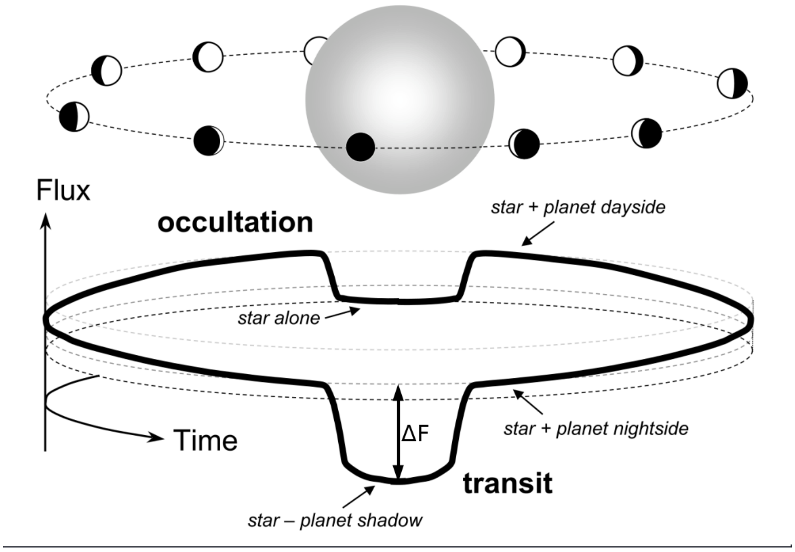


FIGURE 1.1: Here we have a perfect light-curve as a function of the exoplanet's orbit. This light-curve has two dips, one where the planet is in front of the host star and from which the exoplanet's radii can be inferred 1.4. And a second, smaller dip due to the occultation of the planet by the star, where albedo characterization is possible. Image from [14]

to be in the same plane, but this time around the yield of a single observation is smaller than in RV, since a planet, a star and an observer, that are not in the same plane could still achieve a finding and mass estimation for the latter. This yield does not translate in the total number exoplanets discovered by each method. Unlike radial velocity, transit photometry can monitor hundreds of thousands of stars at the same time, making him more prolific.

In a light-curve (plot of stellar intensity over time), the dip caused by the planet is called transit and its main product is the size of the planet relative to the host star as seen in equation 1.4 [15],

$$\Delta F = \left(\frac{R_p}{R_\star}\right)^2 \tag{1.4}$$

where ΔF is the depth of the transit, R_p the radius from the exoplanet and $R_\star$ the stellar radius. Equation 1.4 paints a simple picture: smaller stars will produce deeper and easy to observe transits. In reality there are other factors that will influence the targets desirability. Bright stars with low stellar variability might compensate their bigger frame and produce better light-curves. And while bigger and closer to the parent star exoplanets are more easily and frequently observed, smaller and more distanced ones are more interesting and less known about. The difficulties of transit photometry can be seen considering a Jupiter-Sun systems: when Jupiter transits the Sun, only 0.1 % of the stellar sphere is covered, it lasts 30 hours, happens every 12 years and the solar flux will decrease only 1 % [7].

CoRoT and Kepler/K2 are two spaceborne missions whose goal of discovering new planets using Transit Photometry was largely achieved [4]. This telescopes had to look at a given field full of thousands of targets for large amounts of time. In doing so, long and short orbital periods could be seen and the chance of detection was greatly maximized. Most of their observations where carried out in the range $V\sim$ 13-16 mag and the two found small rocky planets like Corot-7b [16] and Kepler-10b [17]. They also provided reliable statistics as to their numbers, but since they target faint stars only a small number of exoplanets can be characterized using RV and transmission spectroscopy. This follow up study would help obtain their mass and density and allow a more in depth characterization.

Ground-based transit surveys such as the Hungarian Automated Telescope Network (HATNet) and the Wide Angle Search for Planets (WASP) have also been highly fruitful with the discovery of hundreds of Jupiter and Saturn sized objects. The diversity found among the detected planets, paved way for future surveys. Missions like TESS, CHEOPS and PLATO will work towards planets with brighter hosts and smaller planets.

1.2.3 First Detection

The first successful usage of transit photometry traces back to the photometric measurements of the star HD 209458 [18] where a planetary-mass companion in a close orbit had been previously found using RV [19]. The primary objectives where to find if the star was photometrically stable and to search for planetary transits. The chances that those ten nights of observation, in the summer of 1999 would yield the first success in Transit Photometry where not that great, even if the star was stable, the probability of finding this transit was 10% [20].

Still, they obtained $R_p = 1.27 \pm 0.02 R_{\text{Jup}}$ and $i = 87.1^{\text{o}} \pm 0.2^{\text{o}}$. Furthermore, they used radial velocity data to obtain a mean density of $0.38\, g\, cm^{-3}$, which is less dense than Saturn (the least dense gas giant in the solar system), the effective temperature of the planet and the thermal velocity of hydrogen were also computed. The latter being significantly less than the escape velocity of this planet, $v_e \approx 42\,\text{km}\,\text{s}^{-1}$, justifies the possibility of this being a stable planet.

1.3 The Characterization Effort

1.3.1 Planetary Structure

A planet that is discovered through RV, will have a mass value associated with it. This value, however, is not good enough to make a good assertion about its composition, since planets with similar masses are known to have very different compositions. A good example of this are Kepler-10c [21] and Kepler-18c [22, 23], where the first has three hundred and eighteen times the mass of the latter, but is only half its size. For this motive, there is a symbiotic relation to be found, between different missions, be it to improve previous measurements or to complement them with new information. A RV study is positively complemented with a transit search follow up and vice-versa, that could provide the radius and mass of the exoplanet and unlock its mean density. With the density alone, it is possible to build an ensemble of possibilities as to the general composition of the planet. These ensembles are particularly useful for planets whose size do not exist in our solar system, as in the case of super-Earths which could be giant ocean planets, mini-Neptunes, dwarf gas planets or a combination of metal core, rocky mantle and volatile species such as water and gas that we don't have a frame of reference for [4].

1.3.2 Planetary Formation

Planetary formation occurs in dense protoplanetary disks during the pre-main phase of a star's evolution. Despite not being completely understood, it is believed that gas giants and telluric planets formation occur in two different regions of the protoplanetary disc that are separated by the ice condensation line. Estimates of the location of this transition are based on the equilibrium temperature of a dust grain that is illuminated by a host star at a given distance from her: $T(r) = (\frac{L_S}{16\pi\sigma r^2})^{1/4}$. L_S is the stellar luminosity, σ is the Boltzman constant, and r is the orbital radius.With this equation is possible to determine the radius at which water starts to condense into ice grains, r_c.

Jupiter-like planets are believed to form beyond the ice line, because they need a big enough solid core to trigger run away gas accretion which is the gravitational collapse of the surrounding gas to form the gaseous envelope. Without water ice, it is believed that there is not enough solids in the disk to trigger gas accretion. Surprisingly, it is close to r_c the preferential location for the formation of giant planets[24]. Further constrains can be

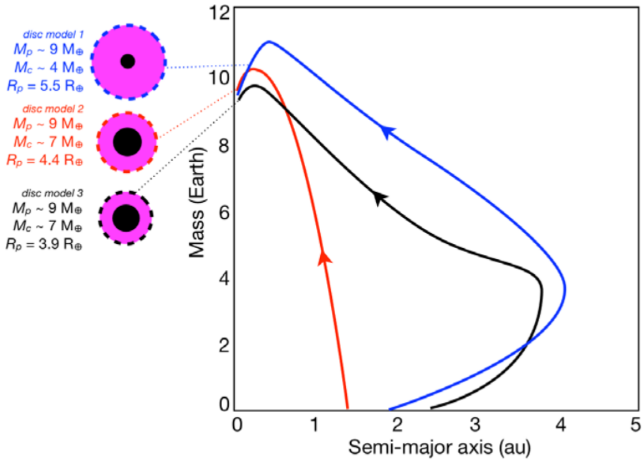


FIGURE 1.2: Three different planets that lie at the same distance from their host star, and have the same mass, but differ in size and in composition. This difference is due, allegedly, to different migration paths. Planets that spend their genesis closer to the star will be telluric like, will those that stray further away will have a more gaseous structure. Image from [4]

placed in order to have a more robust model, but the ice condensation radius will remain the lower bound for the formation of this giant planets.

This model is missing as to explain the origin of Uranus and Neptune, since it predicts bulkier gaseous envelopes for them. It also ignores migration during the formation of a planet, which are prevalent for gas-giants. The timing, the speed, the lengths and the different regions of the disk that are crossed during the planet's formation will significantly impact its final composition, as seen in figure 1.2. These migration paths are difficult to constrain. Only with a precise determination of the mass and radius of hundreds of exoplanets, it will be possible to disentangle all the factors and understand the formation history of a given planet. This is one of the goals of future exoplanets surveys like CHEOPS.

1.3.3 Atmospheres and Envelopes

Exoplanets that have been previously studied and have a precisely measured radii and mass, might be good candidates for the analysis of their atmosphere by future missions. Among these, stands the James Webb Space Telescope (JWST) and the towering European Extremely Large Telescope. This study is done through transmission spectroscopy during the transit or emission spectroscopy during planetary occultations. Despite the breakthrough that will come from this new exciting pieces of technology, this kind of study has been happening in the past decade and a half. The infrared study of an exoplanet can yield a phase curve * capable of displaying the efficiency of heat distribution from the day side to the night side of the extra-solar planet. And while the JWST and other futures missions are poised to work in the infrared, there still is a place for visible phase curves, that have ingrained in their shape the relative size and amount of clouds or hazes in the atmosphere.

1.3.4 Orbital Dynamics

The characterization of a planet is not limited to the planet itself, analyzing what surrounds it is also an important part of the study. In this matter, TTV (transit time variation) is a powerful tool, that uses minor changes in the orbital period to show the observer other celestial bodies that are also bound to the host star and are responsible for perturbing the exoplanet under scrutiny. TTV as shown itself capable of using transit surveys from CoRoT and Kepler to extract other perturbing planets, while RV found the faintness of the star a prohibitive parameter.

1.4 CHEOPS

The CHaracterising ExOPlanet Satellite or CHEOPS for short is a $<$ 250kg satellite using a 33cm diameter Ritchey-Chrétien telescope with a single back-side illuminated Charged-Coupled Device (CCD) located in the focal plane. CHEOPS includes a dedicated field stop and a baffling system to minimize stray light contamination. Its bandpass covers the visible-near-infrared range of $330 - 1100\,nm$.

Despite being a small mission, CHEOPS is designed for the ultra high precision transit photometry study of extra-solar planets. Performing characterization of super-Earths and achieving new insights on Neptunes, through the measurement of their radii comprise its primary goals. Among these ″new insights″ is the limiting planetary mass that enables a massive hydrogen envelope. CHEOPS is also capable of going further and estimate the albedo of hot Jupiters from their phase curves and occultations. It can also search for co-aligned, inner and smaller planets in systems with transiting Neptunes using TTV.

CHEOPS is also suited for the alluring task of characterizing an habitable planet, particularly around an M dwarfs. These stars are cooler and smaller than Sun-like stars and yield larger transit and Doppler signals. They also have their habitable zone at shorter

*flux of the star and planet system as a function of the orbital phase distances which eases the detection. Considering [25] on the maximum radius that still allows habitability, CHEOPS can search for this planets, even if constrained to Super-Earth and Neptune sized planets that have a 50 day orbital period (maximum orbital period for a CHEOPS target) or less.

While ground transit surveys like the NGTS and the WASP have a photometric precision of typically 1000 and 10000 ppm over one hour of integration, CHEOPS, for its dimmest star (12th magnitude) is expected to have a precision of 100 ppm. This improvement is due, in part, to the Earth's atmosphere that plagues ground based missions.

1.4.1 Synergy

The CHEOPS mission is a joint venture between ESA (European Space Agency), Switzerland and other member states, of which Portugal is included. It has a low-budget and a mission implementation time of around four years. CHEOPS will observe one target at a time. Its strength and efficiency comes from knowing where to look by following up on previous missions. Its importance is found in its own goal of creating a collection of ”golden targets” that will be used for in-depth characterization in future missions. Hence CHEOPS is the middle ground between Past and Future missions and instruments and synergy with them is of utmost importance.

CHEOPS target list will come from ground-based RV's (HARPS, HARPS-N, ESPRESSO, etc.) and transit surveys (WASP, K2, TESS, NGTS, etc.). A fifth of CHEOPS available time will be open to the community through regular ESA open calls. CHEOPS finds in TESS (Transiting Exoplanet Survey Satellite) a striking opportunity for follow up research since CHEOPS has around nine times the equivalent collecting area, operates in a similar magnitude range, but allows bigger orbits, where as TESS is bound to observations of 27 days for the targets close to the equator. Any exoplanet from that region with an orbital period bigger than 3 days will see an increase in the transit parameters accuracy with a CHEOPS light-curve.

The new, space-borne, James Webb Space Telescope (JWST), whose unmatched thermal infrared sensitivity will make of it the general purpose infrared space observatory of the next decade. However, its thermal emission detection is highly dependent on the size of the planet, thus a precise radius value from CHEOPS will help prioritize its targets. A precise characterization of Super-Earths by CHEOPS will allow to select the best targets for atmospheric characterization by JWST.

1.4.2 Scientific Requirements

According to CHEOPS Definition Study Report [4], it must be able to detect an Earth-size planet transiting a G5 dwarf star (0.9$R_\odot$) with a V-band magnitude in the range of $6 \leq V \leq 9$ and characterize a Neptune-size planet transiting a K-type dwarf star (0.7$R_\odot$) with a V-band magnitude as faint as 12 with the goal of going as low as 13. This two achievements are unlocked if the flat field is measured down to a pixel-to-pixel precision of $\sigma_{ff} = 0.1\%$, is stable over two days, and if the pointing accuracy is better than 8 arcsec rms.

CHEOPS will have a 800 km sun-synchronous orbit of 101 minutes, and because of this low altitude, 40 % of the data will be lost to the South Atlantic Anomaly (SAA)[26].

The satellite will be nadir locked and will roll around its line of sight, this is done to keep the radiators pointing away from the Earth (source of radiation) and increase the thermal stability. While the main target star will appear undisturbed in the middle of the image, the stars in the background will rotate around the target. Thus, between 1 min exposures the image will have rotated 3.6^o, and is now subject to a different pointing jitter and this will be the source of some headaches that will be discussed further down.

1.5 CCDs

Originally envisioned to be memory devices by Bell Labs in 1969, Charge Coupled Devices (CCDs) soon saw their light sensitive properties exploited [27]. CCDs are, not only, responsible for an improvement of two orders of magnitude in the overall sensitivity of astronomical instruments, but are also fast and efficient in digitizing data [28]. They have spawn a revolution in astronomy and are now indispensable to this field.

1.5.1 Working Principle

CCDs are semiconductor based devices, when hit by a photon with more energy than its band gap, an electron jumps from its valence band to its conduction band, and a hole will be placed in his stead. This pair would defuse in the semiconductor lattice and after sometime, they would recombine and the effect created by the photon would no longer be noticeable.

Being capable of linearly generating a large number of electrons from an incident flux of photons, maintaining them stored, transferring and quantifying them is what makes

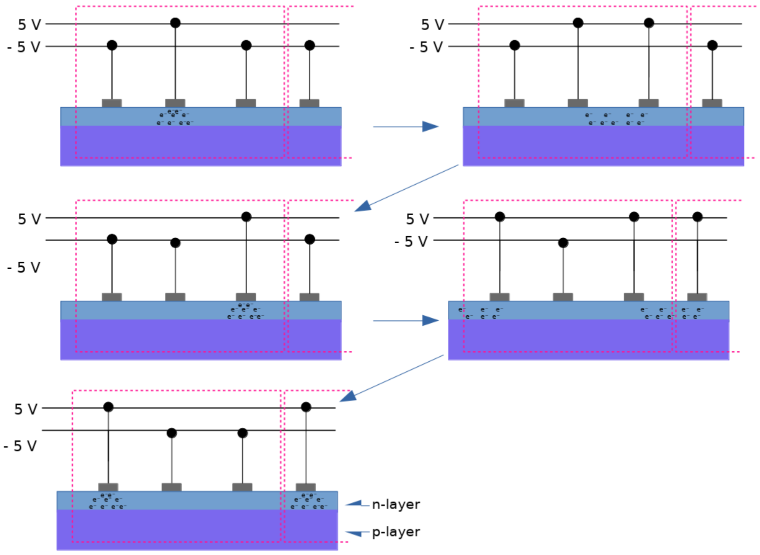


FIGURE 1.3: The transfer method of the electrons captured in each pixel is displayed sequentially until they reach a new pixel. The pixels are delimited by the pink dash and each has 3 electrodes. By applying 5 V to the middle electrode, and maintaining the other two at $-5\,V$ the electrons are confined to the area under the highest positive voltage. Giving the same high voltage to the electrode at the right and then lowering the voltage in the middle electrode will make the electrons move towards the right side of the pixel. This is done until the charges get to the middle electrode from their neighboring right pixel.

a CCD useful. Much like in a diode, each pixel will have a n-layer and a p-layer, each filled respectively with electrons and holes. The diffusion of the electrons and holes to the opposite layers creates a charge depleted area and induces an electric field on top of it. It is this depleted region that is sensitive to photons. When a photon hits this region with an energy close to the Gap energy, it creates a free pair of an electron and a hole. The electron and the hole will remain trapped close to the n-layer zone and the p-layer zones that are under the electric field respectively. This way they don't recombine immediately and can be stored in the pixel for the duration of the exposure. A photon that hits outside of the charge depleted area and manages to create a free pair will see it recombine since there is nothing to keep them apart.

The electrode system of figure 1.3 prevents electrons from moving freely vertically. Horizontally this is guarantied with channel stops that define different columns and physically separate the n-layers. During the electrons' transport, each row is transferred simultaneously until it reaches the Serial Register. In the Serial Register, each pixel of a given row is transferred one at a time to an amplifier, where voltage is induced from the present charge with a given gain (V/e). This voltage is ultimately converted to a discrete digital number in Analog-to-Digital Units (ADUs), but not before a bias voltage is applied that ensures that the read value is positive. This allows the usage of an extra bit that would otherwise be used for the sign of the voltage.

1.5.2 Noise and Performance

CCDs are the bridge between the stellar light and the observer, understanding their performance and the noise that is associated with them is critical. The first mechanism of noise that is associated with the CCD is not caused by him, but by the act of measuring light: The photon noise is created by the uncertainty of the arrival of a given photon to the CCD, hence it scales with the square root of the signal. Since there are many sources of light in the same field of view of the target star (e.g. other stars, thermal emission from the telescope, exozodiacal light, stay light etc.) the total flux will have entangled information from the target star and the background noise.

The light that enters the telescope is guided to the focal plane where the CCD lies with a given optical throughput T. When photons reach the CCD, it is not guaranteed that electron will be knocked into the potential well, some might even pass through the CCD untouched, since the latter could be invisible to them depending on their wavelength. Some might even be knocked, but will not reach the well due to impurities. All these factors contribute to a Quantum Efficiency rate which depends on the wavelength. The overall number of electrons in the CCD is also influenced by electrons that are thermally knocked to the conducting band. This generated signal is called dark current, since it exists in the absence of a light source.

No two pixels are alike, and the way they differ from one another is a source of noise in the image. Moreover there are bad pixel who have an abnormally high (hot) or low (dead) flux response or might have stochastic changes of their sensitivity in an otherwise normal behavior (telegraphic pixels).

The transfer of the electrons is not perfect and can be measured through the Charge Transfer Efficiency (CTE) which measures the percentage of electrons that is transferred in each pixel transfer.

The act of reading the charge in a given pixel, is prone to uncertainty and thus noise. Typically, a good average value is obtained by the amplifier, but with some random scatter called Readout noise. This noise is the minimum noise value for a CCD measurement, and is most relevant when dim stars are being observed. Another problem that affects dim stars and very bright ones that are being measured is the CCD's response deviation from

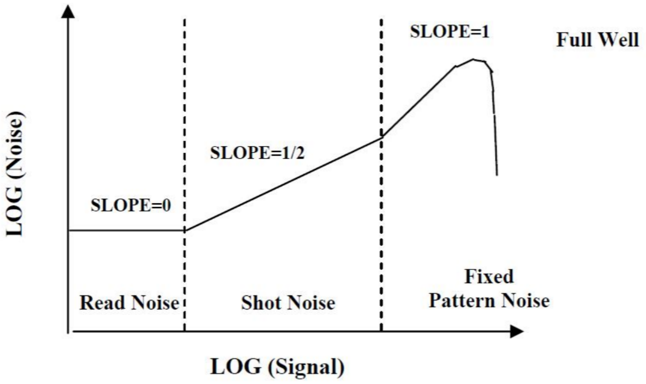


FIGURE 1.4: Logarithm of the noise over the logarithm of the signal plot divided in the three main regions that rule the noise behavior: Read Noise, Shot Noise and Fixed Pattern Noise. Image from [29].

its linear behavior (where electrons are directly proportional to photons). If the number of electrons starts to reach the full well capacity some electrons will be lost to inactive areas of the semiconductor and other might bleed to neighboring pixels in the same column.

The amount of noise in data, depends in the amount of signal read by the CCD, as well as other factors that have been mentioned above. The three different regions that set the main component of the noise are displayed in figure 1.4. For dim signals the noise is controlled by the readout noise. If the signal increases sufficiently, the photon noise region is entered, whose 1/2 slope comes from the square root of the signal a property of Poisson noises. If it gets even bigger the final region might be achieved, where the noise is set by the differences that exist between each pixel (Fixed Pattern Noise).

$$\frac{S}{N} = \frac{N_*}{\sqrt{N_* + n_{pix}(1 + \frac{n_{pix}}{n_B})(N_S + N_D + N_R^2 + G^2\sigma_f^2)}} \tag{1.5}$$

The Signal-to-noise ratio of a CCD follows equation 1.5 [30], where N_* is the total number of collected photons, n_{pix} the number of pixels, n_b the number of pixels in the background estimate. N_S is the number of background photons per pixel, N_D the number of dark current electrons per pixel, N_R the number of electrons per pixel that come from the read noise. $G^2\sigma_f^2$ is the digitization noise, where G is the gain and σ is the noise of a single ADU step. If the number of photons that are collected is large enough, one can reduce the formula to $S/N = N_*/\sqrt{N_*}$.

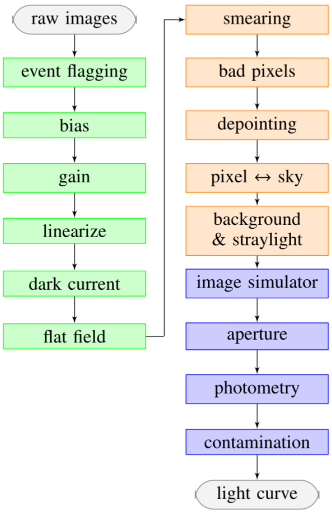


FIGURE 1.5: The green steps comprise the calibration, the orange ones the corrections and the blue the photometry. Image from [31]

1.6 CHEOPS Data Reduction Pipeline

CHEOPS' Data Reduction Pipeline (DRP) is responsible for the on-ground data processing of CHEOPS images. The DRP happens automatically without any external agent at play and its final outputs are ready for scientific analysis. The DRP has been built with the help of simulated data from the CheopSim [5]. For this reason, the DRP will be subject to further optimization during the commissioning phase, when CHEOPS data starts to be available.

The DRP, as explained in [31], can be divided in three main sections: Calibration, responsible of correcting instrumental noise, correction, accountable for the corrections of environmental effects and, finally, photometry, tasked with outputting a final light-curve, from the previous corrected and calibrated images. The different sections and each step that comprises them can be seen in image 1.5. Beside these three sections, this pipeline has a Report module that automatically outputs a document with the evolution of the signal across the successive DRP steps at the end of the run.

1.6.1 Calibration

Calibration begins with raw images from CHEOPS in ADU. This step uses known characteristics from the satellite to correct the instrumental response. As the satellite is still on the ground, all that is known comes from lab calibrations, further improvements can be done during the commissioning phase done when the sattelite is in orbit. At this stage, the bias is removed, the units are converted from ADU to photo-electrons, dark current is corrected and flat fielding is done.

Firstly, the Event Flagging is done to remove bad exposures whose raw frames are beyond being reduced to something meaningful and might even compromise the detrending of the regular frames. This is followed by the removal of the bias voltage. The bias is sensible to slight variations due to fluctuations of voltages and temperature. For this reason, it needs to be adjusted for each exposure. In CHEOPS, an image is the sum of up to 60 shorter exposures. The bias measurement is done with prescan pixels, who are virtual pixels devoid of photon noise and dark current. Obtaining the bias of a single image and removing it, would not only be bandwidth expensive, but would also increase the readout noise, since there are a lot less prescan pixels than aperture ones. For this reason the bias correction is done removing a constant component estimated over the whole visit, leaving a time varying component left to be done with the background correction.

Gain correction is done dividing the images by the gain function, converting the units of the image from ADU to photo-electrons. Linearization is done with the help of the imagettes, smaller frame image of each exposure. It is in them, where a correction law determined from laboratory measurements will be applied.

Dark current starts to accumulate from the beginning of the exposure and will continue to do so until the last pixel in the CCD reaches the amplifier. In this manner, it is possible to map the amount of time each pixel will wait until they are read, as seen in equation 1.6,

$$M(x,y) = t_{exp} + n(t_z + y \cdot t_y + x \cdot t_x) \tag{1.6}$$

where t_{exp} is the exposure time, n is the number of stacked images, t_y and t_x is the time needed to shift a row or a line respectively and t_z is the transfer time of the imagette to the covered storage zone. There are sixteen dark columns of pixels which are only affected by the dark current. The dark current estimate will be separated into a constant value that is determined by averaging the dark current over the entire visit (observation period) and a variable term that will be mitigated during the background correction step.

Pixel Response Non Uniformity (PRNU) are corrected dividing the image by the flatflied. PRNU is highly dependent on the wavelength, hence the flat-field image that is used in this pipeline is a linear combination of several monochromatic exposures weighted according to the spectrum of the target star.

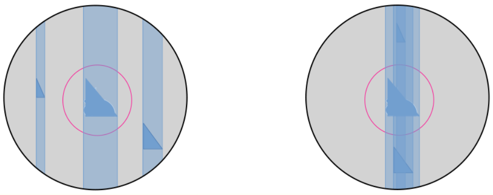


FIGURE 1.6: The self smear is always present in the aperture defined by the pink circle, but the same does not happen for neighboring star's smear. They will only be notice by the viewer when they are directly above or bellow the aperture, as seen in the image from the right.

1.6.2 Correction

Ideally, the CCD would only receive light when every pixel is in place and would stop receiving it once the exposure is over. Unfortunately this is not the case for CHEOPS due to the absence of a shutter. For him, light will also be gathered when new rows start to replenish the CCD at the end of the previous exposure and during the movement to the concealed area at the end of the current one. The light received before and after the exposure is responsible for the smear flux $f_k(y)$ as described in equation 1.7

$$f_k(y) = \sum_{i=y+1}^{N} s_{k-1}(i) + \sum_{i=1}^{y-1} s_k(i) \tag{1.7}$$

where $s_k(i)$ is the flux collected during the kth readout of row i. The smear estimation is done by scanning six extra rows (overscan). The overscan will make the same travel as the other pixels, but without the exposure time, thus accurately representing the smear effect. The smear is, in large part, caused by the target (self smear) and its neighboring stars in the CCD. By itself the self smear is innocuous and more noise would be introduced correcting it than leaving it untreated. This is not the case when the target and the neighboring stars are aligned, as seen in image 1.6. Simulations (see section 1.6.3) will determine if it is worth correcting the smear considering the neighbors position in the CCD.

Next is the detection of bad pixels which tend to increase as the mission progresses over the years.It begins by reducing the jitter in the images by shifting them in the opposite direction of the depointing. Then, a map of the relative variations of the individual pixels compared to their neighbors is made:

$$r = \log(\frac{f}{f * k}) \tag{1.8}$$

where f is the image and k is an unitary smoothing kernel of 10×10 pixels. The r images have an unique footprint for each type of bad pixels. Hot and Dead pixels are spatial outliers. For them an average of the residuals (r) is done and a threshold is set to 30σ, avoiding the peaks of the point spread function (PSF). Telegraphic pixels are searched for individually through their intensity over time. Since regular pixels don't vary as much as them, they are easily distinguishable with a 7σ threshold. Cosmic Rays are temporal outliers which are difficult to detect when they strike the point spread function. For this reason, their detection is done with the imagettes, whose short exposure allows for a better contrast. The residuals of each pixel is plotted over time and a 6σ detection threshold is used. Overall, this module will output the corrected images till this point as well as a 3D map of all the bad pixels in the data.

Now, all the images are centered, pixel coordinates are converted into physical sky coordinates and, lastly, the background and straylight noise is removed. Due to the rotation of the images the DRP removes the photometric aperture and uses the remainder of the image for the background estimation. An histogram is made from the remainder pixel intensities, its upper bound will be excluded to prevent the contamination of the target star in the background and to reduce the contributions of the neighboring stars. A Gaussian is fitted to the histogram and its mode will be the background value to be subtracted to the image.

1.6.3 Photometry

Photometry starts by making two simulations of CHEOPS field of view. They are made to gauge the contamination of nearby stars in the aperture, hence one of the simulations will have only the target star and the other will have the neighboring ones. From the brightness of the target and the contamination that comes from around her, the aperture size will be determined.

The aperture is circular and avoids sharp edges by weighting its bordering pixels with the pixel fractions that are under the aperture as seen in image 1.7. The aperture's shape will not change, but will be displaced according to the centroid position. The DRP will use four apertures: a default one with 33 pixels radius that covers 97.5% of the PSF, one that covers 80% and another that covers 120% and an optimal one. The optimal aperture is calculated by minimizing the noise to signal ratio:

$$\mathrm{NSR} = \frac{\sqrt{f + c + \sigma_c^2 + \sigma_{ron}^2}}{f} \tag{1.9}$$

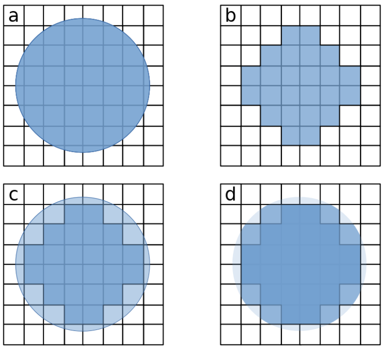


FIGURE 1.7: a) has the aperture on top of the pixels, b) displays a binary aperture, where only the pixels that are completely covered by the aperture are used. c) uses only a fix percentage for the partially covered pixels and d) has a weighted border, where pixels with more aperture coverage have an higher influence than the others.

where f, c and σ_c come directly from the two simulations and are, respectively, the shot noise from the target star and the contaminants and the noise from the ingress-egress in the aperture from the contaminants. σ^2_{ron} is the readout noise that was previously calculated with the help of the prescan pixels. Once the optimal is chosen, each aperture will, through the sum of their pixels, produce four different light-curves which are the main output of the DRP.

Chapter 2

Pixel Level Decorrelation

2.1 Spitzer

Pixel Level Decorrelation (PLD) traces its origins to [1] where the study of a rather odd transiting exoplanet took place, HAT-P-20b. Its quirkiness came from its high density despite its gigantic size of $0.87R_J$. This fact coupled with a parent star with a high metallicity strongly suggests that this is a metal-rich planet. This observations were done with the Spitzer Space Telescope, who claimed the first ever instance of a direct search of an exoplanet through secondary eclipses [32]. However, at the time of this study, Spitzer was past its nominal mission due to the depletion of the liquid helium (coolant). Unable of cooling its IRAC (Infrared Array Camera), its temperature rose from -271^oC to -242^oC, increasing the spacecraft thermal emission and rendering the IRAC's longer bands unusable. The shorter ones (3.6 and $4.5\mu m$) still worked, but suffer from enhanced intra-pixel variations, which can translate into 8% variations in the stellar intensity due to regular pointing wooble [8].

To have surveyed HAT-P-20b with the ”warm” Spitzer it was necessary to overcome its systematics first. Standard options define and remove the correlation between the fluctuations of the stellar intensity and the position of the stellar image. These methods, while successful, have some shortcomings as seen for the $3.6\mu m$ band's red noise where intra-pixel effect is most aggressive [1]. They are handicapped by the assumptions made about the nature of this correlations and by estimating the centroids, which are a secondary products from photometry prone to uncertainty.

Pixel Level Decorrelation was built not only as an alternative to this methods, but to go beyond them. PLD encapsulates the idea that the best way to correct the noise due to the motion of the stellar image is to operate on the pixel intensities directly. Fitting for the stellar position and solving for correlation are bypassed, but the position of the star will remain present implicitly through the relative intensity of the pixels.

2.1.1 Brightness Variation Model

The brightness of the star, S^t, at time t depends on the intensity of every pixel in the image, as described in equation 2.1.

$$S^t = F(P_1^t, P_2^t, ..., P_N^t) \tag{2.1}$$

Here, F is a general function and P_N^t is the intensity of pixel N at time t . Since the point spread function of the telescope is broader than any individual pixel, the variations in S^t due to the image motion will be smooth. This makes F continuous and differentiable and a Taylor series expansion can be applied for small variations. Only the linear terms of the Taylor expansion is used for the Spitzer. They will represent small fluctuations caused by the combination of image motion and spatial inhomogeneities on the CCD depicted in equation 2.2.

$$\delta S^t = \sum_{i=1}^{N} \frac{\partial F}{\partial P_i^t} \delta P_i^t \tag{2.2}$$

The total brightness variations of S^t are not only dependent on the noise from the image motion, but also on the temporal variations of the detector's sensitivity and on the exoplanet's eclipse. By representing the eclipse with $DE(t)$, where D is the eclipse's depth and $E(t)$ the eclipse's shape normalized to a unit amplitude. With temporal variations being represented by a quadratic function of time, $ft + gt^2$. The complete model for the brightness variations is written in equation 2.3.

$$\Delta S^t = \sum_{i=1}^{N} c_i \delta P_i^t + DE(t) + ft + gt^2 + h \tag{2.3}$$

With h being an added offset constant and the c_i's representing the partial derivatives from the Taylor expansion. A final modification is done to equation 2.3 with the δP_i^t being replaced by the normalized pixels $\hat{P}_i^t$:

$$\hat{P}_i^t = \frac{P_i^t}{\sum_{i=1}^{N} P_i^t} \tag{2.4}$$

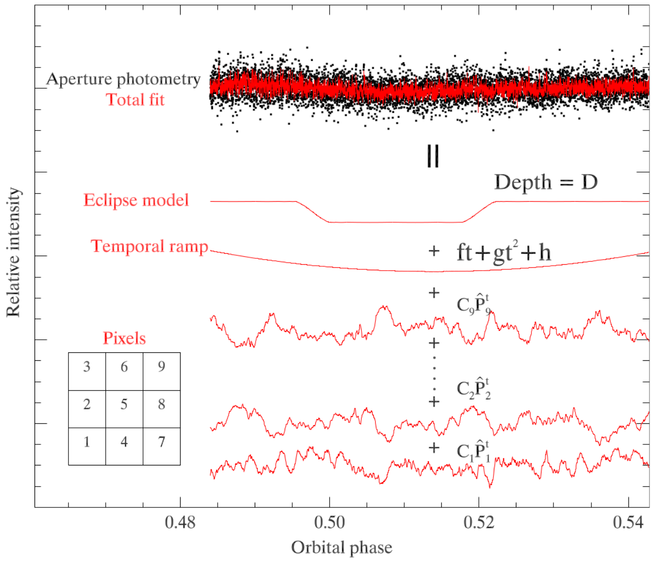


FIGURE 2.1: All the components in a light-curve as seen from the total variation of brightness used for the PLD. The complete light-curve is divided in the eclipse model, the temporal ramp and the contributions from each normalized pixels which through their coefficients represent the instrumental noise. An array of three by three pixels was used here. Image from [1]

The normalized pixels will not contain the eclipse of the planet, since all the astropyshical signal is removed through the normalization. All the components in the light-curve that are modeled through equation 2.3 can be seen in image 2.1.

2.1.2 Least Squares and Data Binning

The coefficient from the normalized pixels, the temporal variations, transit depth and the offset are obtained through Least Squares. This is a standard method of regression analysis where the best model minimizes the sum of the squares of the residuals between the observed data and the model employed. The brightness variation model over time without the coefficients will be stored in matrix X of size $M \times N$, equation 2.5. M is the number of observations and N the sum of the total number of normalized pixels used ($\hat{P}_i^t$), the temporal ramp inputs (t t^2 1), the phase of the transit ($E(t)$) and the offset (1).

$$X = \begin{pmatrix} \hat{P}_0^0 & \hat{P}_1^0 & \dots & \hat{P}_{N-4}^0 & 0 & 0^2 & E(0) & 1 \\ \hat{P}_0^1 & \hat{P}_1^1 & \dots & \hat{P}_{N-4}^1 & 1 & 1^2 & E(1) & 1 \\ \hat{P}_0^2 & \hat{P}_1^2 & \dots & \hat{P}_{N-4}^2 & 2 & 2^2 & E(2) & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \hat{P}_0^M & \hat{P}_1^M & ... & \hat{P}_{N-4}^M & M & M^2 & E(M) & 1 \end{pmatrix} \tag{2.5}$$

$$\chi^2 = (y - Xc)'(y - Xc) \tag{2.6}$$

Equation 2.6 portrays the squared residuals, where y is the SAP (simple aperture photometry) over time, c is a column of coefficients that are yet to determine and $'$ the transpose symbol. Equation 2.3 is used to model the light-curve that is being observed, with each of its elements playing an equally important rule. From within it we want the values that model the instrumental systematics and long term variations, equation 2.7.

$$m(t) = \sum_{i=0}^{N-4} c_i \hat{P}_i^t + ft + gt^2 + h \tag{2.7}$$

When $m(t)$ is removed from the SAP, a detrended light-curve is obtained, where the best depiction of the secondary eclipse is present. Having y and X from photometry, we will set the derivative of χ^2 to 0 to find the coefficients associated with its minimum.

$$\frac{\partial \chi^2}{\partial c} = 2\chi \frac{\partial \chi}{\partial c} = 0$$

$$X'(y - Xc) = 0$$

$$X'Xc = X'y$$

$$c = (X'X)^{-1}X'y$$

$$\Delta S = X(X'X)^{-1}X'y \tag{2.8}$$

This equations can be extended to the case where data has correlated errors [33].

$$\chi^2 = (y - Xc)'K^{-1}(y - Xc)$$

$$c = (X'K^{-1}X)^{-1}X'K^{-1}y$$

$$\Delta S = X(X'K^{-1}X)^{-1}X'K^{-1}y \tag{2.9}$$

K being the covariance matrix of errors.

The binning of data in time is an integral part of the PLD fitting method used for the Spitzer. Spitzer exhibits pointing jitter on a wide range of timescales: 40-minute oscillations due to the battery heater and short-term fluctuations as high as 10 seconds [8]. Binning out the short-term effects will allow for the PLD model to focus on removing the long-term variations. In essence we are trading-off more noise on shorter timescales for less in longer ones. Binning data also helps with the normalized pixels located at the edge of the point spread function, whose intensity and consequently their significance are increased in the basis vector of the PLD. The bin widths shouldn't produce distortion of the eclipse curve. Nor should they reduce the number of data points comparable to the number of the coefficients. PLD is, after all, based on regression analysis whose approximate solutions are obtained for overdetermined systems.

2.1.3 Implementation Overview

The PLD implementation for Spitzer begins by performing aperture photometry with different radius and storing their values. Following up, equation 2.8 is solved for different transit centers with the one that as the best fit being choosen. With the eclipse model set, different bin sizes, different apertures and two centroiding methods will be used to minimize χ^2 . To determine the best fit, the coefficients of each combination will be used in the unbinned data and the residuals (data minus fit) are calculated, (r_i). The aim is to evaluate the photometric precision of each PLD model across different timescales. Thus, each r_i will be binned in different even sizes smaller than 18 and their standard deviation will be calculated. A second χ^2 is used, this time between a line of slope -0.5 that passes through the standard deviation of the unbinned residuals and line made of the different binned residuals standard deviations. The best χ^2 will have its bin size, centroid method and aperture size chosen as the PLD regression solution.

2.1.4 Advantages

PLD claims several advantages over standard methods. These standard methods have difficulties modeling pixels with mis-calibrated sensitivity (errors in the flat-field) that receive a significant portion of the stellar flux. This due to the integrated intensity being affected in a manner poorly represented by a low degree polynomial of the centroid position. PLD is imune to this, because every pixel will have a coefficient attached to it, hence a natural byproduct of PLD is an efficient flat-fielding.

PLD is computationally fast, having its most time consuming section, the estimation of the eclipse model, that is also found in all other methods.

2.2 K2

K2 is a follow up of the Kepler mission, the latter being responsible for discovering nearly five thousand exoplanets. Some of whom were small planets in the habitable zone of their host stars [34]. This mission came to an end after the failure of a second reaction wheel in

2013 which dictated the loss of its high pointing accuracy. K2 relies on a clever observing strategy: by pointing the satellite in the ecliptic plane, the depointing force produced by solar pressure can be mitigated with only two reaction wheels. While Kepler delivered fifteen hundred thousand observations, K2 is expected to deliver ten thousand during its nominal mission. Moreover, the delivered data from K2 will be less precise and will display stronger instrumental artifacts. Among them, a six hour trend that endangers its ability to detect small transits.

Many powerful methods have been developed with the aim of bringing the precision of K2 to the levels of its predecessor [35–37]. Despite not regaining the initial precision, the improvement of data obtained from the raw aperture photometry have allowed this second mission to fulfill its goals. As it was for Spitzer, and despite their differences, most of the pipelines employed for K2 depend on numerical methods that determine and remove the correlations between the stellar position and intensity fluctuations. Even when non-parametric techniques are used such as Gaussian Processes, there still are assumptions being made about the correlation from the spacecraft motion and the instrumental noise. The second instance of PLD usage [2] is implemented in the EVEREST pipeline, used to detrend data from the *K*2 mission.

2.2.1 EVEREST

The jitter in the Spitzer telescope was small enough to constrains the PLD to its first order, this is not the case for K2. Despite the solutions found, there still is large pointing variations on short timescales ($\sim$ 6h) in K2. For Spitzer, solutions where drawn to focus PLD in longer timescale variations, such as the binning of data and the parameter optimization having in mid the effect in different timescales Now, higher orders of PLD will be used to battle short timescale variations, thus the noise model for image i is

$$\begin{aligned} m_i = &\sum_l a_l \frac{p_{il}}{\sum_k p_{ik}} + \\ &+ \sum_l \sum_m b_{lm} \frac{p_{il} p_{im}}{(\sum_k p_{ik})^2} + \\ &+ \sum_l \sum_m \sum_n c_{lmn} \frac{p_{il} p_{im} p_{in}}{(\sum_k p_{ik})^3} + \alpha + \beta t_i + \gamma t_i^2. \end{aligned} \tag{2.10}$$

Equation 2.10 maintains the same variables as before, but with the additional second and third order terms. The increase of the PLD orders will augment the pool of pixels from which the instrumental noise can be computed. This is particularly important for K2 due to the small point spread functions of the stars in its CCD. Despite the added complexity, this remains a linear model which can be solved as before. The detrending power increases steeply with each order of PLD that is being used. However, one should be cautious with the increase of regressors. The PLD will become more and more computationally demanding. A K2 star, with an aperture of 23 pixels, will have the equivalent of $C_3^{23} + C_2^{23} + 23 = 2047$ pixels when using the third order PLD and 10902 for the fourth order. Another draw back of using a higher-order PLD is the higher likelihood of overfitting, leading to an artificially low scatter in the detrended light-curve. This low value will come from the removal of both the white noise and the astrophysical signal which includes the transit, our principal object of study

For K2, a new design matrix will be used, one that does not include the temporal polynomial term. The thermal stability of K2 is still very good, thus not requiring the polynomial term:

$$X = \begin{pmatrix} \hat{P}_0^0 & \hat{P}_1^0 & \dots & \hat{P}_N^0 & 1 \\ \hat{P}_0^1 & \hat{P}_1^1 & \dots & \hat{P}_N^1 & 1 \\ \hat{P}_0^2 & \hat{P}_1^2 & \dots & \hat{P}_N^2 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \hat{P}_0^M & \hat{P}_1^M & ... & \hat{P}_N^M & 1 \end{pmatrix}. \tag{2.11}$$

N is the number of pixels used including their combinations for PLD of orders higher than one. Due to the missing critical components such as the transit and the polynomial representing the temporal variability, this matrix alone is not enough to accurately model the raw flux. However the previous design matrix (used for Spitzer) was not without fault, being incapable of properly modeling a variable star over long timescales. If a sinusoidal signal where to be injected to a raw light-curve, the "Spitzer" PLD would find difficult modeling it even with a tenth order polynomial has its temporal term, figure 2.2 c). Without the polynomial term the PLD will more or less maintain the same shape of instrumental noise as see in figure 2.2 a) and b). But in the absence of a model for the astrophysical variability the PLD will trade red noise for white noise, compromising the precision of the final light-curve.

Using a polynomial term as the astrophysical model in the design matrix and by computing its coefficient through least squares we are treating the stellar signal explicitly. In doing so, the PLD model will now, not only model the instrumental noise, but also the stellar variability, figure 2.2 c). Straying away from its purpose, while better than 2.2 b)

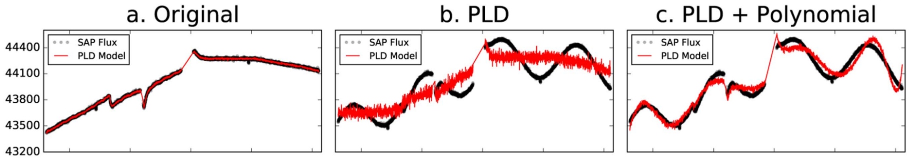


FIGURE 2.2: a) In black we have the SAP flux of the Kepler-803 star from the original *Kepler* mission, in red the PLD model of the instrumental noise. SAP to which a sinusoidal signal with a 25 day period was added and the PLD attempt to model the instrumental noise without any temporal term b). c) is the same as before, but with a tenth order polynomial to model the temporal variability. Image from [2]

doesn't really make the cut. The solution found to make PLD more robust to different types of variable stellar signals is treating these same signals non-parametrically using a GP (Gaussian Process) regression, image 2.3.

2.2.1.1 GPs

Data used in exoplanet science is often difficult to work with, with many of its hardships being transverse to different tools, different methods and objectives. Observations are both costly and time sensitive and the sources of noise that plague them are plentiful and complex. In addition, the sought after targets lay at the edges of current capabilities. Probabilistic inference has shown to be a principled framework that brings a robust, stable and computationally practical approach to the challenges at hand [38]. Bayesian inference aims to make probability statements of unobserved quantities from the observed data [39]. For model fitting this means estimating the probability distribution of the model parameters given the data and prior information. The parameters are not the common coefficients of a given polynomial (parametric model). They are less specific values such as the amplitude, characteristic timescale and smoothness that have been constrained from prior knowledge from the system.

The likelihood function is the probability of a set of observations following a given model, $P(D|H)$, with D being the dataset and H the hypothesis [39]. Prior distributions encapsulates the current knowledge from previous measurements, which also molds our parameter expectations before a new set of observations takes place. With this information, and through the Bayes Theorem we can recover the intended aim of Bayesian inference $P(H|D)$, equation 2.12.

$$P(H|D) = \frac{P(H)P(D|H)}{P(D)} \tag{2.12}$$

In the standard case of modeling N observed outputs (y) from a N inputs of size M (X) one uses equation 2.13 where ϵ represents independent and identically distributed Gaussian noise. For transit photometry the inputs contain only time, while the outputs represent the measured flux.

$$y = m(X, \phi) + \epsilon \tag{2.13}$$

m is the model and ϕ the model parameters. Given that the noise follows a normal distribution $\mathcal{N}$, than its probability distribution can be calculated through:

$$p(\epsilon) = \mathcal{N}(0, \sigma^2 I) \tag{2.14}$$

where $\sigma^2 I$ represents covariance matrix diagonally populated with the variance, σ^2. If the model is added as the mean function in $\mathcal{N}$, it will result in the probability distribution of y (likelihood function)

$$p(y|X,\, \phi,\, \sigma) = \mathcal{N}(m(X, \phi), \sigma^2 I) \tag{2.15}$$

Noise in a dataset is rarely only white, with a vast majority of its source signals being aggregated into time-correlated noise (red). If Red noise is stochastic in time and follows a normal distribution, it may be modeled as a Gaussian process [39]. GPs are the go to method in machine learning for Bayesian inference in non-parametric problems and classification ones [40]. They are defined as a collection of random variables, any finite number of which have a joint Gaussian distribution [41]. A GP model is defined by its mean and its covariance functions:

$$y \approx \mathcal{GP}(m(X, \phi), K(X, \theta)) \tag{2.16}$$

with θ being the parameters of the covariance function. ϕ and θ are generally referred as hyperparameters. The mean function should represent the overall behavior of the observations and will for this reason dominate the forecast of the GP model in regions without data or far from it. Usually this function is set to zero, mirroring the lack of knowledge of the data from the observer. Despite not being ideal, not having a mean function is not an impediment on GPs. It will mean the structures captured and the extrapolated data by the GP model are completely determined by the kernel [42]. The likelihood function of the data given the model, with zero mean, is

$$P(y|\theta) = \frac{1}{\sqrt{2\pi}^n \sqrt{|K|}} \exp(-\frac{r' K^{-1} r}{2}) \tag{2.17}$$

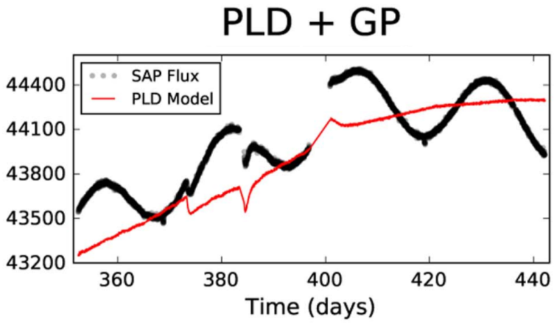


FIGURE 2.3: The design matrix of the PLD has only the pixel from the aperture as well as their combinations and an offset. The covariance matrix used in the computations of the coefficients is obtained through GP regression to represent the astrophysical signal. Image from [2]

with its log-likelihood written as

$$\ln P(y|\theta) = -\frac{1}{2}(n\ln 2\pi + ln|K| + r'K^{-1}r) \tag{2.18}$$

with $r = y - m(X, \phi)$. From equation 2.18 its is possible to infer the model parameters, an important tool that comes with an heavy price: It requires the inversion of the covariance matrix whose time complexity is $\mathcal{O}(N^3)$. However, in some cases the matrix inversion can be sped up or bypassed, [43] and [44].

2.2.1.2 GPs in Everest

In short, GPs are used to estimate the covariance matrix K, seen in equation 2.9. If K is reasonably close to the true covariance of the stellar signal the PLD coefficients will be an accurate representation of the instrumental noise as seen in figure 2.3.

Using the GP regression one would think that the kernels used would have to be hand picked as well as the bounds that constrain its hyperparameters. However, [2] proved that in practice, in the case of K2, the PLD model is relatively insensitive to the kernel function and its hyperparameters. This is a major advantage of using the GP regression, because the stellar signal is unknown a priori. This is showcased in image 2.3, where the 25 day period sinusoidal is modeled by a radial Matérn-3/2 kernel with a timescale of 20 days without the white noise being inflated.

Also missing from the design matrix is the transit model $DE(t)$, and looking at image 2.4 it is possible to see that the transits are shallow and localized when compared with the rest of the light-curve. While possible, modeling the transits of potentially several planets

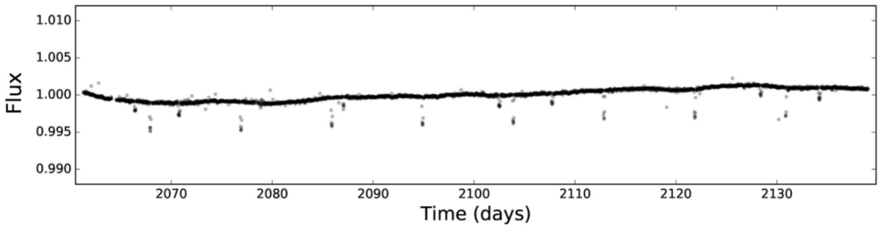


FIGURE 2.4: Light-curve from EPIC 205071984, detrended by EVEREST pipeline which employs pixel level decorrelation. The vertical stream of dots represent the transits of three different planets that orbit the host star. Image from [2]

would be an hassle. Since they account for so little in the overall light-curve, the transits and eclipses that might be in it are treated as outliers. For PLD, outliers in the light-curve are identified and removed from its least-square fitting. Leaving the outliers and the transit in the least square fitting would result in the red noise - white noise trade already mentioned earlier in this section. When computing the final detrended light-curve, the outliers are reintroduced and the transits are thus generally preserved.

2.2.2 Implementation

The first step of Pixel Level Decorrelation for the K2 mission is choosing the aperture. It should be big enough to have enough pixels to generate a good basis set for PLD, all the while preventing contamination from neighboring stars. Then, the background is removed from the pixels in the aperture. This background is computed to be the median of the pixel that lay outside of the aperture for each frame.

Now, through iterative σ-clipping all the features that are not modeled by the PLD nor the GP (outliers and transits) are hopefully removed. To achieve this, the light-curve is divided in five chunks with each one being detrended by the first order PLD with the covariance being computed with a 2 day Matérn-3/2 kernel with the median standard deviation of all 2 days segments in the chunk as the amplitude. A median absolute deviation cut at 5σ is then done to identify the outliers in the residuals of the SAP light-curve by this first order detrending. This first set of outliers are masked and the PLD model will be re-computed for the entire light-curve and a new set of additional outliers are added to the mask. This is repeated until the mask remains the same between iterations or until the maximum number of iterations is reached.

$$K_{ij} = k_w(t_i, t_t) + k_t(t_i, t_j). \tag{2.19}$$

In the first iteration of the EVEREST the best combinations of different kernels would be used to calculate the covariance matrix. Following, principal component analysis (PCA) would be done through cross-validation, in order to control the PLD model from overfitting [2, 45]. In the most update version, a single optimized Matérn-3/2 kernel is used and Regularization takes the place of PCA. Since overfitting occurs when the PLD coefficients become too large, the latter consists in imposing a prior that penalizes the likelihood that increases with an increasing coefficient value. A more in depth explanation is given in section 2.3.1 followed by its implementation for CHEOPS

Once detrended, the combined differential photometric precision (CDPP), is used to evaluate the quality of the light-curve. This is the formal photometric noise metric developed by the *kepler* team [46], but an easier to calculate version is used for EVEREST [47]. The CDPP of a light-curve computed for a given duration outputs the depth of a transit of that duration that would have a signal-to-noise ratio of 1 and is thus very convenient to asses the scientific quality of the light-curve. The CDPP begins by removing the correlated stellar noise using a 2 day quadratic Savitsky-Golay high-pass filter to detrend the flux. This prevents the white noise to be inflated by the red one in the precision metric. Following this, outliers laying outside the 5σ area are removed. For a z hours CDPP, the light-curve is divided in chunks of z hours. The precision of each chunk is then estimated by computing the standard deviation in the chunk and dividing by the square root of the number of samples. Finally the CDDP over z hours is computed as the median of the precision of all the chunks of z hours in the light-curve.

2.2.3 Limitations

PLD traces two of its limitations to the assumption that a normalized pixel is devoid of any astrophysical signal which is not always the case for saturated stars nor for crowded apertures. Saturated pixels that bleed electrons to their neighbors in the same column will globally preserve the overall shape of the light-curve in simple aperture photometry. However, the overflowed pixels contain no transit information to begin with, since their full-well capacity is reached even for the in-transit flux. Normalizing the saturated pixels will over-correct them, since they don't contain the entire signal they receive. This leads to an inverted transit shape in the time series of those pixels. PLD is not able to properly detrend this targets as it removes both the transit and the instrumental noise. Collapsing the pixels that compose the saturated columns would solve this problem.

Crowed apertures are also tricky for PLD, since the assumption that the transit is removed through normalization, equation 2.20, is no longer true.

$$\frac{p_{il}}{\sum_k p_{ik}} = \frac{a_{il}\tau_i}{\sum_k a_{ik}\tau_i} = \frac{a_{il}}{\sum_k a_{ik}}, \tag{2.20}$$

where a_{il} is the stellar signal and τ_{il} the transit signal. This time around a different stellar signal, b_{il}, will be added to equation 2.20 resulting in

$$\frac{p_{il}}{\sum_k p_{ik}} = \frac{a_{il}\tau_{il} + b_{il}}{\sum_k a_{ik}\tau_i + b_{ik}} \tag{2.21}$$

from which the transit cannot be completely erased. Considering that the perturbing star is dim compared to the target one it is possible to use the difference between the crowding in a give pixel and the crowding in the entire aperture, $\Delta = \frac{b_{il}}{a_{il}} - \frac{\sum_k b_{ik}}{\sum_k a_{ik}}$ and re-write equation 2.20 (appendix A):

$$\frac{p_{il}}{\sum_k pi_{ik}} \approx \frac{a_{il}}{\sum_k a_{ik}} \left(1 + \frac{\Delta}{\tau_i}\right). \tag{2.22}$$

From equation 2.22, it is seen that the bigger the Δ compared with τ_i, the worse of the PLD will be. This is the case for bright contamination sources that hoover around the edge of the aperture. Or when the transit or stellar variations are shallow, for which the contamination will change greatly across the aperture. However if the contaminant star is co-located with the transiting planet host the quantity of b_{il}/a_{il} is constant across the detector leading to a $\Delta = 0$ and a PLD regression that will work as expected. PLD was found to overfit for contaminants that are separated by more than one pixel and that are brighter or within two orders of magnitude of the target star. This conclusions are specific to K2 due to the small size of the PSF.

The final limitation comes from the usage of the GP to model the astrophysical signal. If the stars being studied are highly variable, having a high amplitude and high frequency signal, the GP will capture both the astrophysical and the instrumental noise. The high frequency from the astrophysical signal is thus a gateway for the GP to overpower the PLD in the short timescales. If the instrumental noise is modeled in the covariance matrix, the coefficients of the PLD will not produce a model that removes the instrumental noise from the light-curve. Moreover, additional artifacts will be introduced by the tension between GP and PLD models in the short timescales.

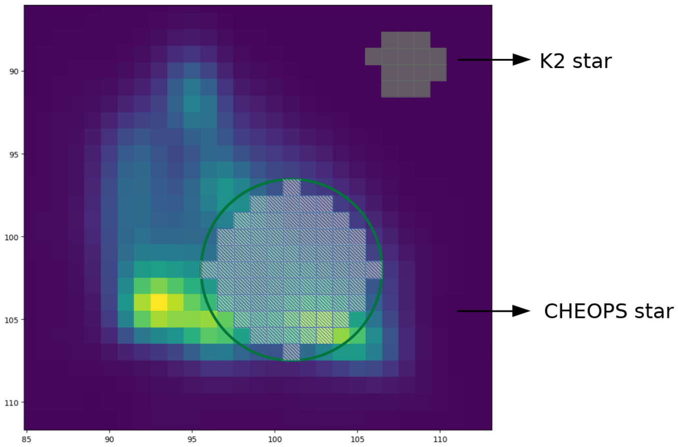


FIGURE 2.5: A CHEOPS simulated image is displayed with a 75 pix aperture inside,this being the maximum number of pixels allowed to have a third order PLD. The green aperture of 5.5 pix radius, is not able to encapsulate the majority of the flux captured by the CCD. Displayed in grey is the aperture of the EPIC 201505350 (K2 target), which is much smaller than the minimum aperture needed to detrend a CHEOPS target.

2.3 PLD for Cheops

CHEOPS has two major characteristics that must be addressed before the PLD regression is used: 1) The high number of pixels inside the PSF and 2) its rotating field of view.

K2 saw its maximum number of pixels constrained to 75, preemptively stopping the code if this number was reached. This is done to avoid memory errors since it equates to a design matrix of 2GB for the third order PLD. This memory usage comes from the 3840 data points of each pixel from K2's 80 days light-curves with a 30 minutes cadence. 75 pixels corresponds to a circular aperture of 5.5 pixels of radius, where only the completely covered pixels are used. Even considering the generally smaller light-curves from CHEOPS, $\sim$ 1440 data points*, that might allow bigger apertures, this constitutes a problem given the 99% of the energy in its PSF spreads in a radius of $\sim$ 30 pixels. While the aperture could be reduced to a radius of 14-15 pix and yield good results, a 5.5 pix radius accounts for less than 20% of the PSF and is well inside the image of the target, has portrayed in image 2.5.

PLD is inherently prone to overfitting due to the high number of coefficients that it computes, one for each pixel. Fortunately there are three ways of controlling this problem: 1) Keeping more data points than coefficients, as it was done for Spitzer where the binning of data points was restricted by the number of coefficients. 2) Reducing the number of coefficients through PCA as it was done in the first iteration of EVEREST or through pixel

*1 day light-curves with a 1 minute cadence binning. 3) Regularization as it was used in the latest version of EVEREST and in the Casual Pixel Model [48], which is another instance of pixel level detrending.

Keeping the number of regressors down and using only the first order PLD might not be a viable option for CHEOPS. This, due to its complex jitter from its rotating field of view, mixed with the usual noise introduced by the guidance system. Using PCA goes against one of CHEOPS biggest attributes: its big PSF and would lead to the loss of information. Thus, the only option left for CHEOPS is regularization.

2.3.1 Regularization

The minimization of χ^2, equation 2.9, is the equivalent of maximizing the log-likelihood, equation 2.15, [49]. However, for ill-posed problems where the number of data points is not big enough when compared to the number of coefficients, neither will work. If we have prior knowledge on the problem at hand, a previously insufficient number of data points, might now provide valuable information. Instead of minimizing χ^2, we will focus on maximizing the posterior [50].

$$\begin{aligned}
\hat{\theta} &= \arg\max_{\theta} P(\theta|y) \\
&= \arg\max_{\theta} \frac{P(y|\theta)P(\theta)}{P(y)} \\
&= \arg\max_{\theta} P(y|\theta)P(\theta) \\
&= \arg\max_{\theta} \log P(y|\theta) + \log P(\theta)
\end{aligned} \tag{2.23}$$

$\hat{\theta}$ are the kernel hyperparameters that maximize the posterior. In $\log P(\theta)$ we have our regularization terms that will allow the PLD to run on ill-posed problems, such as CHEOPS datasets.

2.3.1.1 rPLD

In the latest version of EVEREST, L2 regularization is used in the dentrending of K2 lightcurves (rPLD). This type of regularization uses a normal distribution with a zero mean as the prior. The PLD coefficients are now obtained through maximizing 2.23:

$$\begin{aligned}
\log\mathcal{L} = &-\frac{1}{2}(y - Xc)' K^{-1}(y - Xc) - \frac{1}{2}\log|K| - \frac{N_{dat}}{2}\log 2\pi \\
&-\frac{1}{2}c'\Lambda^{-1}c - \frac{1}{2}\log|\Lambda|
\end{aligned} \tag{2.24}$$

$$\Lambda_{m,n} = \lambda_n^2 \delta_{m,n} \tag{2.25}$$

with λ_n^2 being the regularization term of all the pixels of the nth PLD order. The model for the rPLD after the minimization of the $\log\mathcal{L}$ is:

$$m = X \cdot (X' \cdot K^{-1} \cdot X + \Lambda^{-1})^{-1} \cdot X' \cdot K^{-1} \cdot y \tag{2.26}$$

and is found to have an higher predictive power than the PCA model from EVEREST 1.0.5. The λ values will be obtained through cross-validation: a model is calculated with the training section of the light-curve with a given λ and is then used to detrend the validation section. The validations sets that yield the best scatter will have its regularization term save to be used in PLD detrend.

The model of equation 2.26, has a matrix of ($N_{reg} \times N_{reg}$) to be inverted and for a third order PLD, it is not computationally tractable. As a solution, the Woodbury equation identity [51] is applied. Now the matrix to be inverted is of size $N_{dat} \times N_{dat}$ and the new model is displayed in equation 2.27.

$$m = X \cdot \Lambda \cdot X' \cdot (X \cdot \Lambda \cdot X' + K)^{-1} \cdot y \tag{2.27}$$

$$X \cdot \Lambda \cdot X' = \sum_n \lambda_n^2 X_n \cdot X_n' = \sum_n \lambda_n^2 X_n^2 \tag{2.28}$$

with n being the PLD order and a third order design matrix being represented by

$$X = (X_1 \; X_2 \; X_3) \tag{2.29}$$

Equation 2.27 greatly increases the computation speed of the PLD since with $N_{reg} \approx 10N_{dat}$ for K2. For CHEOPS this transformation is fundamental, with the difference between N_{dat} and N_{reg} being even greater.

2.3.2 Implementation Overview

The input data of PLD for CHEOPS (PfC) will be Raw data from satellite that has been previously treated by DRP's Calibration. There is no data reduction during the first three steps (blue boxes) of image 2.6. They are used to set the behavior of the PfC (PLD parameters) such as the maximum PLD order, the kernel to be used and the number of iterations in other sections of the code. To guess the initial kernel parameters (Initial Kernel Guess) and to assess the initial photometric precision (CDPP). The ensuing steps will

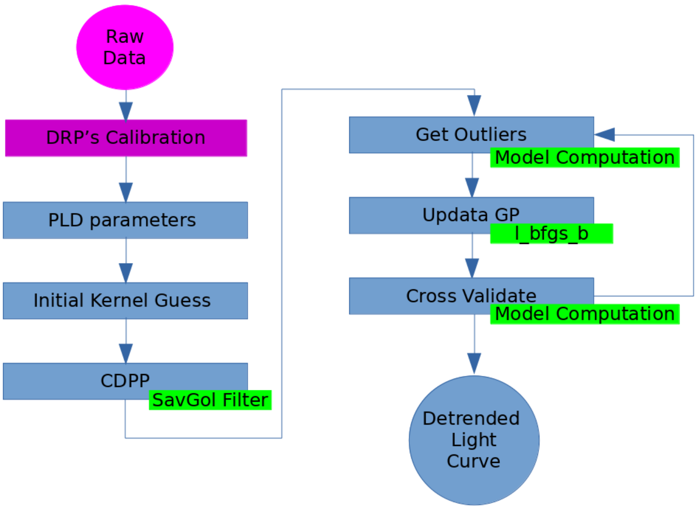


FIGURE 2.6: PfC is divided in its different stages (blue boxes). The purple box highlights the source of PfC input data and the green ones represent modules used in the blues boxes. CDPP is used through the code, but, for simplicity, only the first instance is recorded in this scheme.

begin to carve the desired light-curve from its initial state, with the flagging and masking of outliers (Get Outliers). The optimization of the kernel hyperparameters needed for the covariance matrix (Update GP), the estimation of the best regularization value for the PLD model and its computation (Cross Validate).

Each PLD order is computed separately, but with the help of the preceding orders. The loop that links the Cross Validate and the Get Outliers is only triggered if the PLD order desired has not yet been reached. For example, with a second order PLD:

1. The first order PLD model is first computed in the Get Outliers, using a covariance matrix made from the hyperparameters of the Initial Kernel Guess and a standard value for λ_1 given in the PLD Parameters.

2. The kernel hyperparameters are updated (Update GP) using the light-curve reduced with the model calculated in 1).

3. λ_1, used to regularize the first order is updated in the Cross Validate and the final model for the first order PLD is calculated.

4. A first order PLD model is computed in the Get Outliers with the previously found hyperparameters and regularization value.

5. Update GP updates the hyperparameters using the light-curve reduced with the model calculated in 4).

6. Design matrix has the second order pixels added to it, $X = (X_1\,X_2)$. Keeping λ_1 fixed, the Cross Validate seeks to find λ_2. Once found, the second order PLD model is calculated and the PfC output is obtained.

PLD Parameters also sets the number of breakpoints in the light-curve. Breakpoints are divisions in the light-curve, whose resulting chunks are independently solved by the PfC. This reliefs the strain in the PLD model that might try to model instrumental noise that radically changes from one section to another. In the PfC this values are not used, because the observations done with CHEOPS are much shorter (a few hours or days), when compared to the 80 days of $K2$.

2.3.2.1 Initial Kernel Guess

The PfC begins by determining an initial guess of the hyperparameters for a given kernel. This guess will be responsible for creating a covariance matrix, needed for the PLD model. This is a kick-start for the PLD model used in the masking of outliers and will be the seed in the hunt for better parameter values (GP Update). For example, a Matérn-3/2 kernel with an added white noise term:

$$K_{ij} = \alpha \left(1 + \frac{\sqrt{3(t_i - t_j)^2}}{\tau}\right) e^{-\frac{\sqrt{3(t_i-t_j)^2}}{\tau}} + \sigma^2 \delta_{ij} \tag{2.30}$$

will need a guess for its white noise (σ), red noise amplitude (α) and red noise timescale (τ).

A first order PLD model (simple PLD) will be applied to the data to remove from it some of its instrumental noise, helping to better estimate the hyperparameters. Only the fifteen pixels with the highest signal to noise ratio will be used in the design matrix and the covariance matrix will be diagonally populated with the SAP uncertainties, Δy^t. Despite its simplicity, this model proves itself capable of fulfilling its purpose even when the PfC is applied directly to CHEOPS data, image 2.7. There are no values for the individual pixel uncertainties in the CHEOPS data. For this reason their uncertainties will be set to the photon noise, $\Delta P_i^t = \sqrt{P_i^t}$, with the covariance matrix being populated with

$$\Delta y^t = \sqrt{\sum_{i=0}^{N} P_i^t}. \tag{2.31}$$

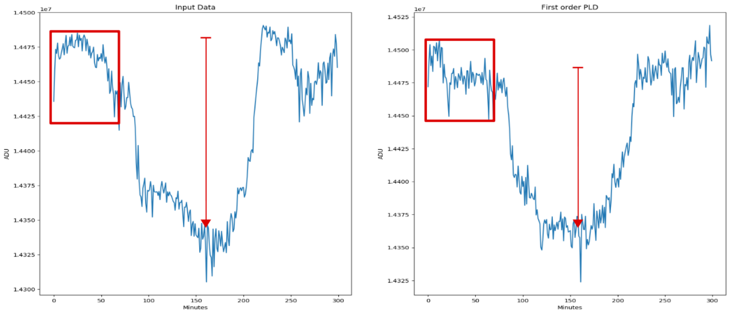


FIGURE 2.7: Input data is showed in the left and the detrended data from the first order PLD with 15 regressors in the right. The red arrows showcase the biggest temporal trend to be modeled by the GPs. The red boxes show the windows from which the white noise will be obtained.

The S/N for each pixel is obtained dividing the median pixel intensity by the median absolute deviation (MAD). The design matrix (X) is made from the chosen normalized pixels and the SAP will be the y of equation 2.9. To estimate the white noise, the light-curve is divided in several chunks and the median of their standard deviations is the guess. The red noise amplitude is the standard-deviation of the light-curve times a coefficient. τ is not calculated being set to the width of the transit. The improvement between the input and the reduced data is clear figure 2.7, with some of the red noise in the first being corrected by the simple PLD. The white noise is more or less the same, which is a positive sign. The data is not being fully transformed by this simple PLD model. And the majority of the dentrending is being left to be done by the proper regularized PLD model that accounts for the astrophysical data with its covariance matrix.

2.3.2.2 CDPP

The mechanics behind the CDPP used in PfC are the same as the ones used in K2's (described in section 2.2.2), but some details of the implementation are different. The CDPP begins by removing all the long term trends in the light-curve and while for the K2 this meant removing the red noise, for the CHEOPS it will also mean removing the transit. Ideally we want the noise to be displayed horizontally before we compute the photometric precision. To this end, the long timescale trends are modeled by a Savitzky-Golay (SavGol) filter with its output being removed from the data, thus flattening it. Next, outliers are detected and removed, thus mitigating their impact in the following photometric precision estimation.

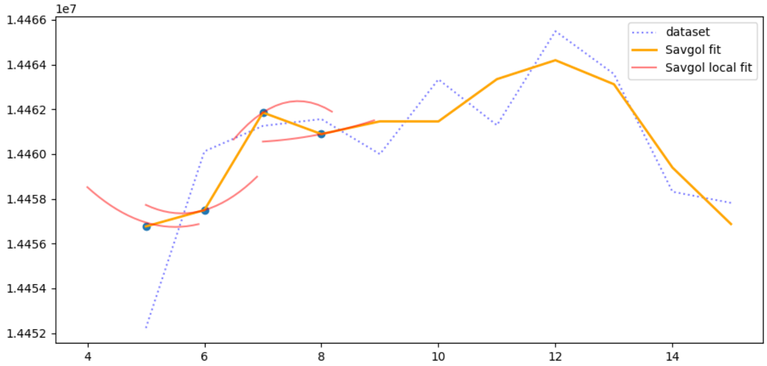


FIGURE 2.8: A small section of a simulated light-curve from CHEOPS is used to show the local fits of each point as well has the SavGol filter for the light-curve. A window of 5 points was used for the filtering of each point. And the local fit is done through a second order polynomial. This is a plot of evenly spaced data in ADUs

To better understand how the CDPP works and when it doesn't for CHEOPS, first we must look the the SavGol filter. The SavGol fits a set of local samples with a polynomial, using least squares with the central value of the fit being used as the filter value for that point. If the SavGol is using a quadratic function for a window of five points, then equation $c_0 + c_1 * x_i + c_2 * x_i^2 = y_i$ must be solved for each point. x is the relative distance to the center point, hence for a five point window $x = \{-2,\ -1,\ 0,\ 1,\ 2\}$ and y_i are the sampled points. This is displayed in the form $Ac = y$ in equation 2.32. c_0, c_1 and c_2 are the desired coefficients.

$$\begin{pmatrix} -2^0 & -2^1 & -2^2 \\ -1^0 & -1^1 & -1^2 \\ 0^0 & 0^1 & 0^2 \\ 1^0 & 1^1 & 1^2 \\ 2^0 & 2^1 & 2^2 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} \tag{2.32}$$

As it is for the PLD, the coefficients for the SavGol are obtained through $c = (A'A)^{-1}A'y$. Matrix A from equation 2.32 is only valid for datasets whose x's are evenly spaced, computed over five point windows and with a second order polynomial. As $(A'A)^{-1}A'$ is always the same,

$$\begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \frac{1}{35} \begin{pmatrix} -3 & 12 & 17 & 12 & -3 \\ -7 & -4 & 0 & 4 & 7 \\ 5 & -3 & -5 & -3 & 5 \end{pmatrix} \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} \tag{2.33}$$

we need only change the window of y's to compute the c_0s. The c_0s being the values of the SavGol filter, since for the central point: $y_2 = c_0 + c_1 * 0 + c_2 * 0^2$.

The SavGol filter of the Scipy module, as well as in others modules of different programming languages, exploit the fact that for an evenly spaced dataset you don't have to continuously compute the $(A'A)^{-1}A'$. You compute this matrix once for a given window size and polynomial order and compute the c_0 multiplying the first row of $(A'A)^{-1}A'$ by the local set (y_i's). For unevenly space data such as a detrended light-curve from the PfC, due to the SAA and outlier removal, this simplified version might not work. Considering this data as even, would amount to virtually shifting each point to equally spaced locations. This shift is responsible for an increase in the data noise which could be too great a compromise. A rough criterion [52] to see if it is worth using the evenly method is: If the difference in intensity between the first and last point in the window is less than $\sqrt{N/2}$* times the measurement noise of a single point, then we can use the simple method. For datasets that have an evenly enough spaced dataset, the rough criterion is not triggered and the simple SavGol is used.

For the above mentioned reasons we implemented for this pipeline a SavGol filter that works with unevenly spaced data and that takes into account the noise of each measurement. Matrix A is changed to

$$A = \begin{pmatrix} (x_{-2}-x_0)^0 & (x_{-2}-x_0)^1 & (x_{-2}-x_0)^2 \\ (x_{-1}-x_0)^0 & (x_{-1}-x_0)^1 & (x_{-1}-x_0)^2 \\ (x_0-x_0)^0 & (x_0-x_0)^1 & (x_0-x_0)^2 \\ (x_0-x_1)^0 & (x_0-x_1)^1 & (x_0-x_1)^2 \\ (x_0-x_2)^0 & (x_0-x_2)^1 & (x_0-x_2)^2 \end{pmatrix} \tag{2.34}$$

*N being the window size

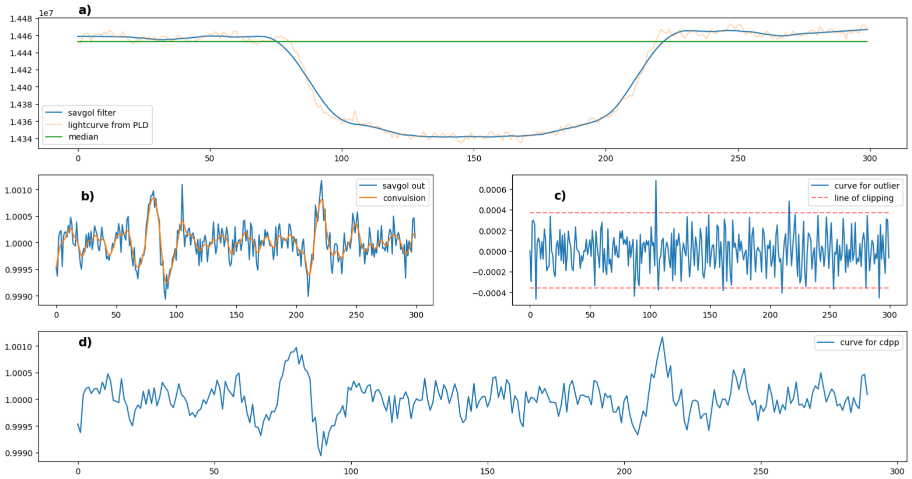


FIGURE 2.9: The different stages of the CDPP are shown here. a) depicts the SavGol filter stacked against a detrended light-curve of PfC. b), c) and d) show the outlier removal and the final plot of the residual where the photometric precision is taken from.

with x_i's representing the actual temporal values of a five points window. The equation that yields the coefficients is now

$$c = (A'K^{-1}A)^{-1}A'K^{-1}y \tag{2.35}$$

with K being a diagonal covariance matrix populated with the uncertainties of each light-curve point.

Figure 2.9 b) shows in blue the residuals from the SavGol filter removal from the light-curve. Outlier detection begins with a copy of the residuals that goes through a convolution with a hanning window. The convolution result is then removed from the previous residuals. This second set of residuals goes through a single σ-clipping iteration: 1) calculate the MAD and the median (m), 2) remove all point bigger or smaller than $m \pm \alpha$MAD[*], as seen in image 2.9 c). The photometric precision estimation of the residual and the outlierless light-curve is the same as in section 2.2.2.

For K2, the typical duration of the light-curve is 80 days. The CDPP can thus be computed for time scales of 6 hours and the width of the SavGol filter window is set to around 1.5 days. The SavGol filter will thus pass through any unmasked transit, leaving them to influence the CDPP. This is not a big problem since only shallow transits escape the outlier detection, hence the end result will be close to the ideal. In CHEOPS, the typical duration of the light-curve is 10 hours. The CDPP cannot be computed for timescales of 6 hours, because there aren't enough chunks. Thus the photometric precision is evaluated over

[*] α is a coefficient chosen by the observer

1 hour of less. The SavGol filter must have a window size bigger than the photometric precision timescale to ensure that it is not filtering out any noise (ten times bigger). This is a problem for CHEOPS, because a SavGol windows with more than 41 points begins to introduce noise when filtering the planetary ingress and egress discontinuities, figure 2.9 a) and d).

The user has to ensure that a 41 points SavGol window is not eroding the light-curve noise. If it is, a larger window must be chosen. For these windows, the peaks in the residuals introduced by the rough light-curve filtering need to be removed. This removal is done by an added option in the CDPP that takes the derivative of the light-curve and through σ-clipping removes the ingress and egress location from the photometric precision estimation. The option does not need to be used for every light-curve that yields a transit, since this peaks are not always as noticeable as in image 2.9 d). The CDPP of the PfC is further analyzed in section 3.3.

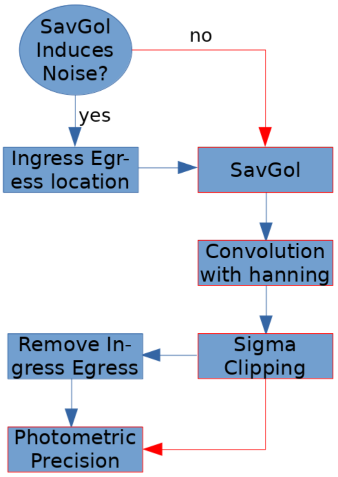


FIGURE 2.10: CDPP scheme, with the option of detecting and removing the egress and ingress of a transit in the light-curve (blue) or not (red).

2.3.2.3 Get Outliers

The data reduction effectively begins with the Get Outliers. The outliers are saved in a global variable and are copied to a local list inside the Get Outliers module. The process is similar to the one described in the CDPP section, but this time the σ-clipping happens directly in the residuals of the SavGol filter. The PLD model is computed before the SavGol filter and uses the most up to date set of outliers also stored in the local list. This process is repeated until the new set of outliers are equal to the previously found ones or a maximum number of iterations is reach, image 2.11.

The PLD model requires the design matrix, the covariance matrix and the regularization one. The regularization weights are only computed during the Cross Validation section, hence the weights used in the first order PLD are set to 10^5. Despite having the lowest number of regressors, the first order PLD might still have more regressors than data points. Hence, without this guess for λ_1 the first order PLD will remove the entire signal. In the first order, this will also mean using the hyperparameters guesses from the Initial Kernel Guess.

2.3.2.4 GP Update

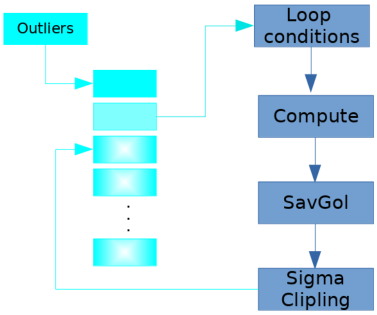


FIGURE 2.11: Scheme of the Get Outliers with the blue boxes. The light blue boxes are the arrays of outliers, with the gradient ones being the ones that are still empty. The "outliers" box is the global outliers array that is updated by this function when the loop condition is meet.

There are three kernels available to model the astrophysical signal: 1) Matérn 3/2 kernel from equation 2.30, 2) the Squared Exponential and 3) the Exponential Sine Squared (Quasi-Periodic) the latter two being displayed respectively in 2.36.

$$k(r^2) = \exp(-\tfrac{r^2}{2})$$

$$k(X_i, X_j) = \exp(-\Gamma) sin^2 \left[\tfrac{\pi}{P}|x_i - x_j|\right]) \tag{2.36}$$

Here $r^2 = (x_i - x_j)' C^{-1} (x_i - x_j)$ with C being the covariance matrix of the input coordinates x_i and x_j. Γ is the scale of the correlations and P the period of oscillations.

GP Update begins by masking all the previous discovered outliers from the data, followed by a single iteration σ-clipping has the one used in the Get Outliers. At this stage the user can choose to keep the hyperparameters that reach GP Update or guess new ones. Each hyperparameter has its respective bounds defined. For the case of the Matérn-3/2 kernel, the white noise lower bound is set to ten times smaller and the upper one to ten times bigger than the guess. The Red noise timescale will depend on the target, but as a rule of thumb it must engulf the transit duration and other known variability from the host star. The Red noise amplitude ought to be comparable to the transit size ($\sim$ 1%) and its bounds are dependent on its guess, to the hundredth for the lower and hundred times for the upper. Just before the initial guesses goes through the optimization process they are slightly randomized by $(1 + 0.1 * rdv) * guess$. With rdv being a random value from a Gaussian distribution of mean 0 and variance 1.

The optimization of the hyperparameters is done through the `fmin_l_bfgs_b` function from the `scipy.optimize` module. As the name states, it implements the L-BFGS-B, a quasi-Newton method to estimate the minimum of a given function f, with a given set of bounds on its parameters. The L-BFGS-B is less memory demanding than a standard

Newton method, because instead of the Hessian matrix, it uses a smaller matrix (B_n) for the optimization of equation f [53, 54].

f will take four set parameters which won't change during the optimization: time, stellar intensity and its uncertainties as well as the kernel to be used. The guesses will be the values to optimize. f begins by building the desired kernel using the `GP` function from the `George` module. Time and errors are used to `compute` the kernel. f will output the likelihood and the gradient of the light-curve being currently modeled by the kernel through `log_likelihood` and `grad_log_likelihood` respectively.

```python
def f(guess,time,flux,errors,kernel):
  w = guess[0] #white noise
  a = guess[1] #red noise amplitude
  t = guess[2] #red noise timescale

  gp = george.GP(a ** 2 * Matern32Kernel(t ** 2),
                          white_noise=np.log(w ** 2),
                          fit_white_noise=True)

  gp.compute(time,errors)
  nll = -gp.log_likelihood(flux)
  ngr = -2 * gp.grad_log_likelihood(flux) / \
          np.sqrt(np.exp(gp.get_parameter_vector()))

  return nll, ngr
```

If everything goes as expected, the optimization will terminate by one of two ways: 1) the change in the likelihood is not significant compared to its current or previous values or 2) the normal of the projected gradient becomes sufficiently small. This two are respectively described in equation 2.37.

$$\begin{gathered} \frac{(f_n - f_{n+1})}{\max(|f_{n+1}|, |f_n|, 1)} \leq \mathrm{factr} \cdot \mathrm{epsmch} \\ ||\mathrm{proj}\, g|| \leq \mathrm{pgtol} \end{gathered} \tag{2.37}$$

Where pgtol and factr are user controlled parameters and epsmch is the machine precision.

2.3.2.5 Cross Validation

Cross Validation is responsible for the regularization values that prevent the PLD from overfitting. It begins by using the masked light-curve and updating the covariance matrix with the previously optimized kernel hyperparameters The light-curve is then divided in three chunks, with each (validation set) being used to gauge the scatter of a model computed with the other two (training sets).

This model is computed for different values of λ that logarithically range from 1 to 3.16×10^{21}. It would be computationally expensive to compute the PLD model 46 times. For this reason, the computation is divided in two stages: One that is done once for each validation/training set (PreCompute) and another done for each regularization value (LamCompute).

The PreCompute will compute six design matrices, three without the outliers ($X1_n$) and the other three without the outilers and the validation set ($X2_n$), one for each order (n). Flux and covariance matrix are also masked from outliers and validation The LambCompute will then calculate the model of equation 2.27:

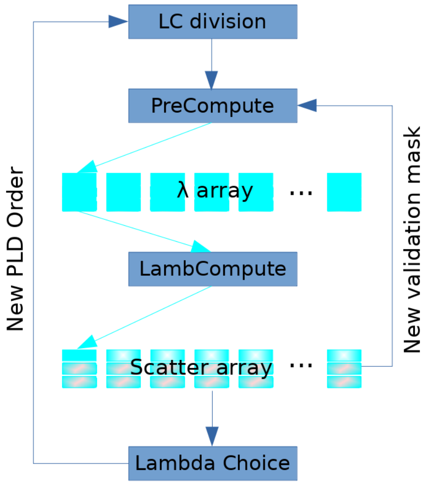


FIGURE 2.12: Cross validation scheme, with the blue boxes being the described stages. The light blue boxes representing the λ array of values (big) and the scatter array (small).

$$m = \sum_n \lambda_n^2 X1_n \cdot X2'_n \cdot (\sum_n \lambda_n^2 X2_n \cdot X2'_n + K)^{-1} \cdot y. \tag{2.38}$$

The model produced by the LambCompute is applied to the entire light-curve, but we are only concerned with the variability of the masked chunk of data (validation set). The SavGol filter that is used in the CDPP to remove long term variations in the light-curve, is substituted by a GP model. The median absolute deviation (MAD) is used to determine the variability of the residuals. Each regularization value has three MAD values from each of the validations sets. The biggest λ that yields the lowest scatter is stored as well as the the lowest λ whose scatter is within 5% of the previous. The smallest λ of the two is saved and used as the regularization value of the PLD order it was computed for. At the end of this section the final PLD model is calculated. This computation is equivalent to joining both the PreCompute and the LamCompute, but with $X1_n$ having outliers and $X2_n$ having them masked.

Chapter 3

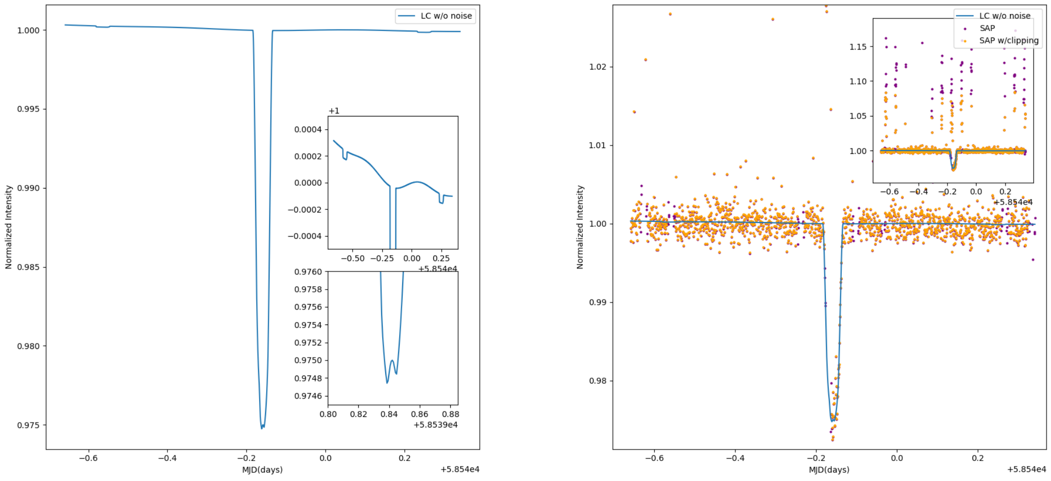
Data Analysis

The PfC was tested for two simulated datasets from CHEOPSim [5], one of the WASP-43b and another of WASP-18b. In this chapters, the outputs from the PfC and the behavior of its different sections will be analyzed for sets of data that haven't yet been treated by the DRP. For this reason, the smear is removed with the overscan pixels as well as the constant component of the dark current through the dark rows. The outer most rim of pixels in the aperture is used to estimate the background noise and remove it. Flat-fielding is performed even if the latter is not mandatory when using a PLD model. Memory and runtime tests will be carried out to access if the PLD remains a powerful and fast detrending tool for CHEOPS data. These are done in a laptop with a quad-core Intel i7-7500U at 3.5 GHz and 7844 MB of RAM.

3.1 WASP-43b

Found in 2011 [55] through the ground-based transit search WASP [56], WASP-43b is a ultra-short period Jovian planet orbiting a K7V type star of 12.4m_V. Further characterization [57] unveiled, among other values, its mass, 2.034M_J and its radius, 1.036R_J, as well as WASP-43's mass 0.717$M_\odot$ and radius 0.667$R_\odot$.

CHEOPSim's light-curve of this exoplanet without any source of noise is displayed in image 3.1 a) and the raw flux from a 17 pix aperture is displayed in b). This simulation represents one visit of 24 hours with a 1 minute cadence. CHEOPSim follows a table that relates the effective temperature, the radius, the mass, etc. to the spectral type [58]. Since this star is bigger and has more mass than a usual K7V, it is simulated with a spectral type



(A) (B)

FIGURE 3.1: These plots comprise the starting point b) and the desired planetary signal a) that CHEOPS must obtain. In a) a Light-curve containing only the desired astrophysical features, 2 occultations, 1 transit and a phase curve and a bump. The latter could be interpreted as a solar dark spot that is being covered by the planet [14]. SAP with and without an initial 10σ-clipping is displayed against the desired astrophysical features of the light-curve, b).

of K5V. Using CHEOPS Estimation Time Calculator [6] there is: 108.59 ppm of instrumental noise (N_{ins}), 125.32 ppm of photon noise (N_{pho}). Together, they account for 179.52 ppm ($\sqrt{N_{ins}^2 + N_{pho}^2}$) over 30 minutes of integration time. Applying the CDPP function of the PfC to the raw flux that have been previously 10σ-clipped yields 343 ppm. The objective for the PfC pipeline is to reduce the photometric precision to a value well inside the 179.52 ppm, thus proving the effectiveness of PLD in removing instrumental noise. Going bellow the photon noise, would mean that the PfC is overfitting and could be bleaching the simulated astrophysical data. This dataset has 1440 data points and even if outlier detection and removal brought this value down to 1000, the design matrix for a third order PLD (17 pix aperture radius) would have 1.24×10^{11} elements or 992 GB. Hence, a second order PLD will be used for this dataset, which for the same aperture equates to 3.3 GB.

3.1.1 Optimal Aperture and PLD order

This dataset is tested for different apertures for the first and second order PLD, figure 3.2 a). For small aperture ($<$ 15 pixels), a significant fraction of the stellar flux goes in and out of the photometric aperture with the movement of the star on the CCD. This creates an highly variable light-curve (figure 3.2 b)) which is one the limitations of the PLD (see

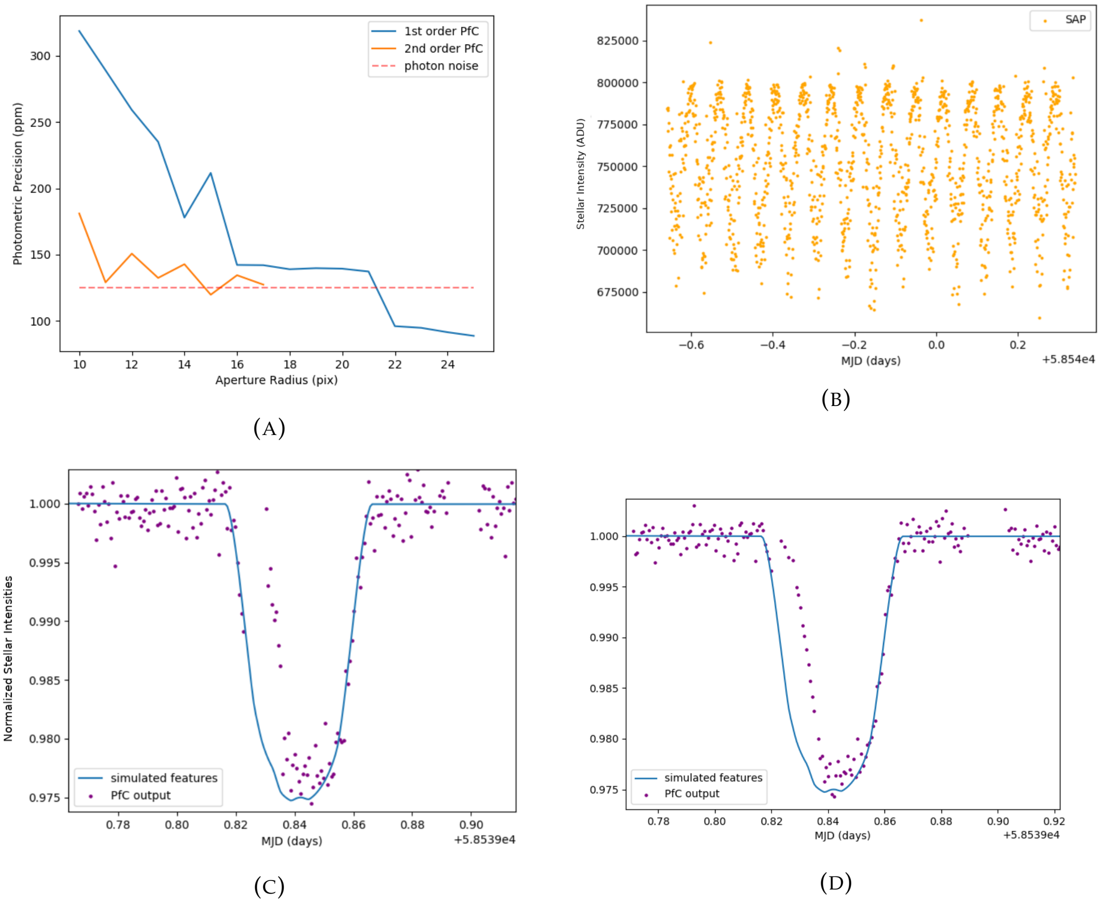
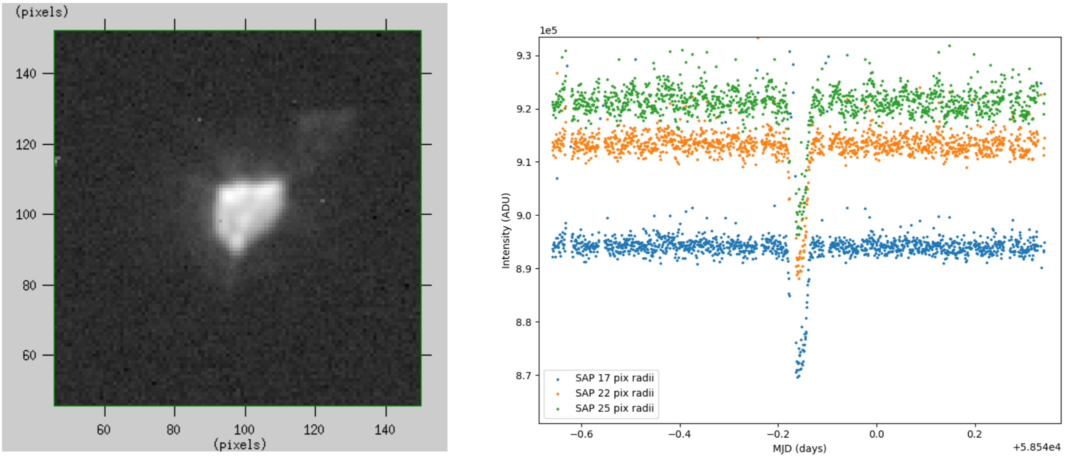


FIGURE 3.2: a) Photometric precision of both PLD order reductions for different radii apertures with 30 minutes of integration time. The second order is limited to 17 pixels radius due to memory constrains. SAP made with an aperture for a 10 pix radius, b). PfC output for a 10 pix radius aperture made with a first order PLD model, with a close up on the main transit, c). PfC output for 10 pix radius aperture made with a second order PLD, with a close up on the main transit, d). The 10 pix radius aperture is chosen to showcase the PLD's struggle with smaller apertures and the effect of the SAA in the ingress of the transit.

section 2.2.3). This could lead to a case of the GP overpowering the PLD model by fitting both the instrumental and the astrophysical data. However, the photometric precision of the first three apertures greatly increases in the second order. This is indicative that first order PLD isn't capable of detrending the light-curve, because the instrumental noise is not well represented by a linear combination of the normalized pixel values. The ingress and egress of the star in the aperture also makes the SAA less prominent in the SAP, figure 3.2 b). This hampers the σ-clipping from removing outliers altogether, including the SAA itself. This corrupted data will have repercussions in the detrending as seen in figure 3.2 c) and d). Here, the SAA data is detrended into a shorter transit and influences the overall depth of transit. For smaller apertures, the places that are not affected by the SAA



(A)

(B)

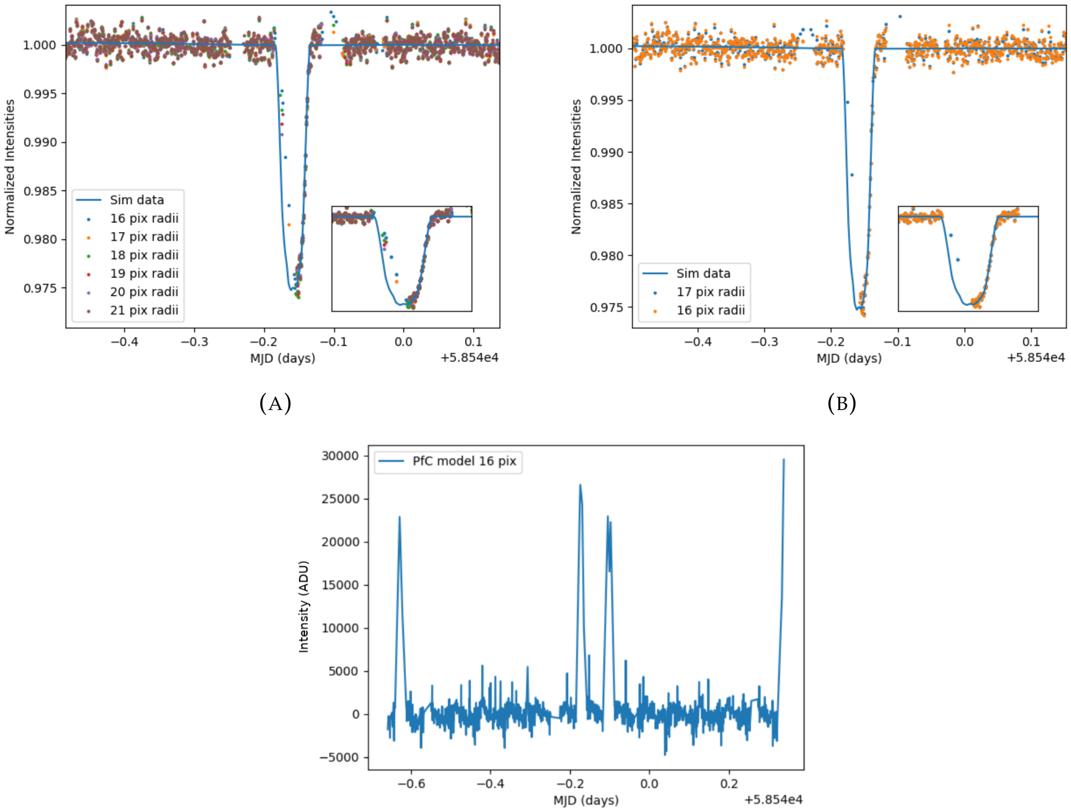
FIGURE 3.3: Neighboring star, next to the target, 23 pixels away from the center of the image, a). SAP for an aperture of 17, 22 and 25 pix radius, displaying an increasing sinusoidal signal, b).

are detrended remarkably well, considering the SAP of figure 3.2 b). For this light-curve, the covariance matrix of the PfC might not be modeling the astrophysical signal, but the motion of the star in the aperture as if it where the stellar variability. This possibility is left open due to the loose bounds imposed in the red noise timescale. If the red noise timescale were properly bounded and the optimized hyperparameters where easily translated to the light-curve parameters we could ensure that this wasn't happening. The Matérn-3/2 kernel does not allow such maneuverability in this dataset (see section 3.1.5)

Bigger apertures are also the source of some headaches, as seen in the 22-25 pix radius. This SAPs will start to have a sinusoidal component in the signal, figure 3.3 b). This is due to the interference of a faint neightboring star close to the target. As explained in section 2.2.3, the PLD doesn't fare well when a neighboring star is interfering in the aperture. Moreover, the GPs are, again, open to model that interference as astrophysical signal from a star/planet system, which is not true. This reasons lead the PfC to overfitting with its output values having less 40 ppm[*] than the photon noise.

One can also note the descending trend in the photometric precision between the first and the second order. The fact that we get so close to the photometric precision might suggest that CHEOPS does not need the third order PLD for the analysis of this dataset.

[*] This was estimated with the CDPP of the PfC. It was found in 3.3 that for this dataset the CDPP is removing some of the noise in the photometric precision. Still, even accounting for the removed noise, this values remain well bellow the photon noise value.



(A)

(B)

(C)

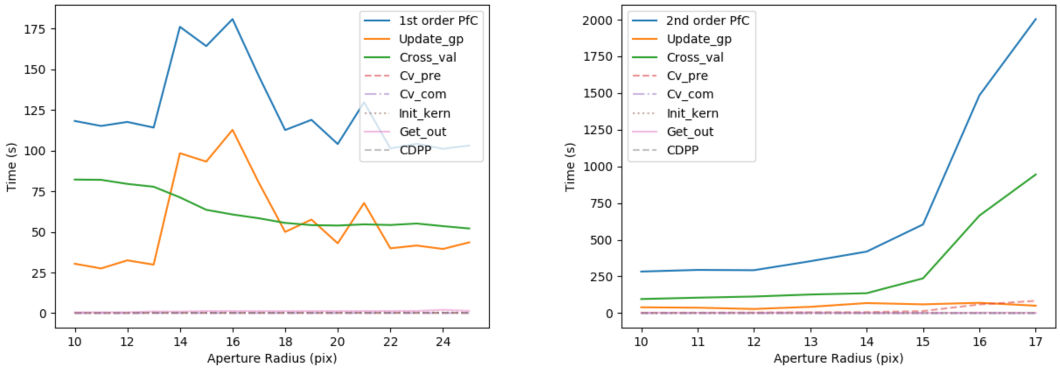
FIGURE 3.4: a) has the light-curves for the 1st order PfC of radius 16 through 21 for the WASP-43b simulation. b) has the detrended light-curves for the 2nd order with a 16 and a 17 pix radius aperture. c) has the PfC model that detrends the 16 pix radius aperture

However the amount of memory needed for the third order PLD is too high and the confirmation is unfeasible.

Apertures of radius 16-21 pix comprise the nominal instances of PfC behavior for this dataset. The closeness of the photometric precision to the photon-noise will be evaluated in the CDPP section of this chapter.

3.1.2 PLD Model

The PfC, for apertures of radius 16 through 21 pix, yields, for the first order PLD, similar light-curves as seen in figure 3.4 a) which is also in accordance with the photometric precision plateau found for this apertures in figure 3.2. For the first order, all apertures have remnants of the SAA, with the detrended transit having points that don't match the simulated signal. This problem is not as prominent as before with each aperture having four or less points that are badly detrended. Besides, all detrends seem to achieve the



(A)

(B)

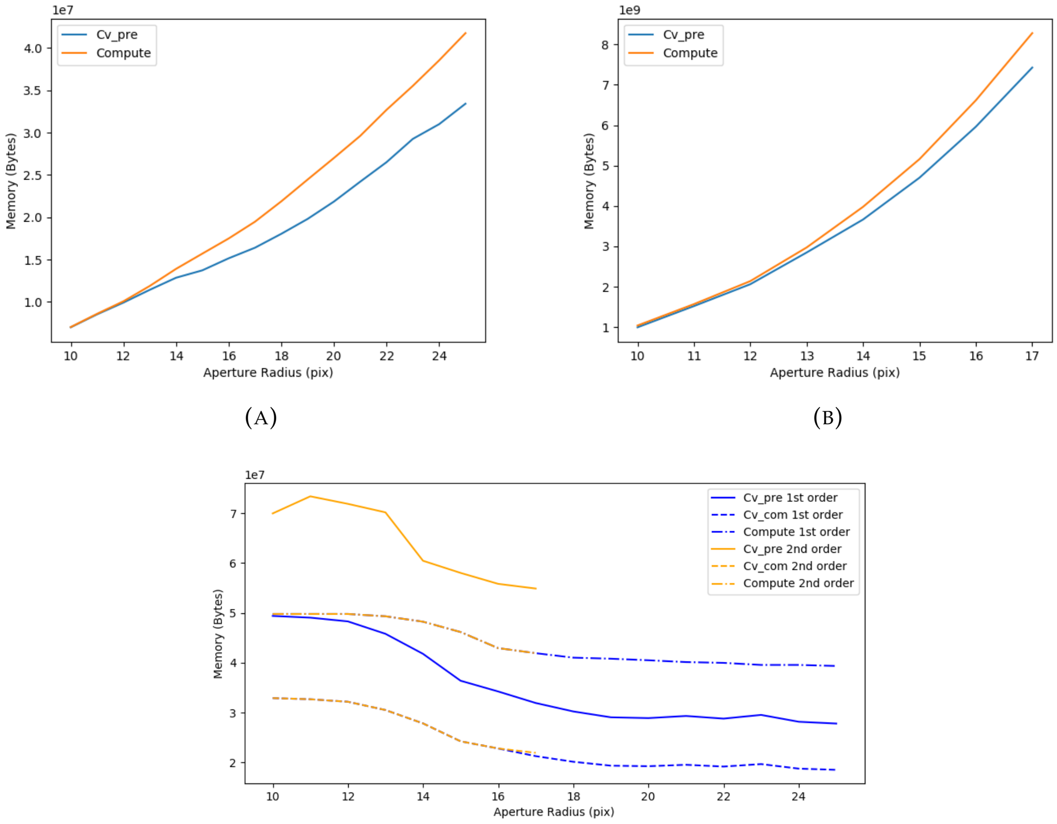
FIGURE 3.5: Runtimes as a function of the radius of the photometric aperture for first a) and second b) PLD orders. The blue curves represent the runtime of the whole PfC reduction, while the other plots display the run times of its different components.

proper transit depth. For the 16 radius aperture this problem is solved with the second order PfC, figure 3.4 b), managing to properly identify the points taken inside the SAA and masking them as outliers. The photometric precision achieved for this aperture is 134.4 ppm. And while the 17 aperture radius achieves a value closer to the photon-noise (127.5 ppm), it still wasn't fully capable of clipping the SAA, having two of its points*. The PLD model used in the detrending of this apertures is more or less the same to the one in figure 3.4 c). And in it, a trend with a period of 100 minutes is displayed. Since one of the major sources of instrumental noise is the satellite's rotation, we can say that the PfC was able, for this aperture sizes, of achieving its goal. This increased photometric precision is however not sufficient to properly detect and characterize the phase curve and secondary eclipse (who are partially affected by the SAA) whose amplitude of 100 ppm is lower than the photo noise for these observations.

3.1.3 Memory and Runtime

Runtimes for the first order PLD are not obvious. With an increase in the size of the aperture, a increase in the runtime should follow. This is not the case (figure 3.5 a)), because, as previously mentioned, the SAA is sequentially removed, through σ-clipping, with the increase of the aperture size. The successive improvements in the clipping of outliers are more impactful on the design matrix size than the aperture size for the first

*Both this values have less 5 ppm then supposed to, see section 3.3.



(A)

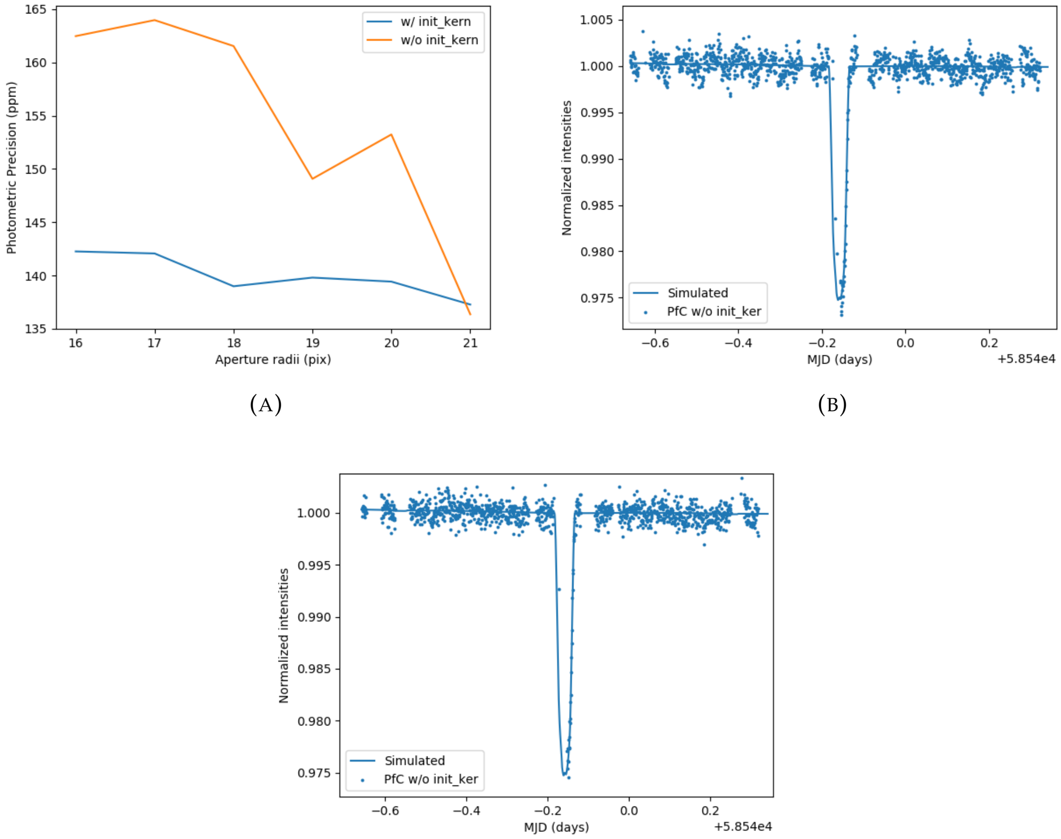
(B)

(C)

FIGURE 3.6: a) and b) have respectively the memory requirements for the first and second order PfC functions Compute and Pre-compute of the cross-validation. c) has the different memory requirements once the design matrices have been multiplied.

order PLD. It is also noteworthy that for this order, the GP optimization through L-BFGS-B is often the most time demanding task.

For the second order (figure 3.5 b)), the behavior is as expected and can be explained by the memory increase presented in images 3.6. The values of this image represent the two massive matrices used to obtain the X_2^2s of equation 2.28 that are $N_{dat} \times N_{dat}$. This single matrix multiplication is responsible for the massive runtime increase. This increase is so drastic, because, for apertures bigger than 15 pix of radius, each of the two matrices starts to be bigger than 3GB, hence the SWAP memory, starts to be used. The Woodbury equation identity is responsible for keeping the rest of the PfC a low memory usage process. Due to him, the matrices used have the same size in the first and second order as seen in figure 3.6 c). The exception is the output matrices of the cross-validation PreCompute, with the matrices X_1^2 and X_2^2 separated waiting to be combined with their respective



(A)

(B)

(C)

FIGURE 3.7: The photometric precision of the first order PfC with its kernel guess against the one used in EVEREST a). The regular products of the EVEREST kernel guess appears in b) and a case where it achieves the proper detrend appears in c).

lambdas and summed together in the cross-validation LambCompute, $\sum_n \lambda_n^2 X_n^2$ of equation 2.28. In Appendix B a plot of the effective memory usage over time is showed for a second order PfC with a 10 pix radius aperture.

3.1.4 Initial Kernel Guess and First order

In the PfC, the light-curve is always improved upon its previous values: The second order PLD uses the product of the first order. And the first uses the product of a simple computation with a fixed value for regularization and a covariance matrix whose kernel hyperparameters are initially guessed. The Initial Kernel Guess is thus the foundation for the entire PfC. And while its estimation of kernel parameters seems crude its importance is easily checked. In figure 3.7 a) the photometric precision for the first order PLD are almost always better when the implemented Initial Kernel Guess is used in the PfC. Using the kernel guess employed in the latest version of the EVEREST pipeline, which applies

Aperture radii (pix)	White (ADU)	Red Amp (ADU)	Red Time (Days)
16	1.170e+03	7.080e+05	9.971e-01
17	1.061e+03	8.686e+05	1.567e+00
18	1.082e+03	1.326e+06	2.972e+00
19	1.127e+03	6.223e+05	1.052e+00
20	1.217e+03	1.021e+05	1.430e-01
21	1.293e+03	1.032e+06	1.772e+00

TABLE 3.1: Update GP optimized hyperparameters for the first order PfC, for a Matérn-3/2 kernel with the bound specified in 3.1.

an unregularized first order PLD to the entire aperture, would set back the following Update GP. Since the influence of a badly seeded GP Update is shown in the first order PfC as seen in figure 3.7 b). In this figure there is a periodic signal which is not part of the simulated astrophysical signal. This might be due to the GP wrongly fitting the instrumental noise signal of 100 min period. The aperture with 21 pix radius is an exception to this (figure 3.7 c)), where the Update GP is capable of going through the bad start and produce a light-curve as good as the one made with the kernel guess tailored for CHEOPS.

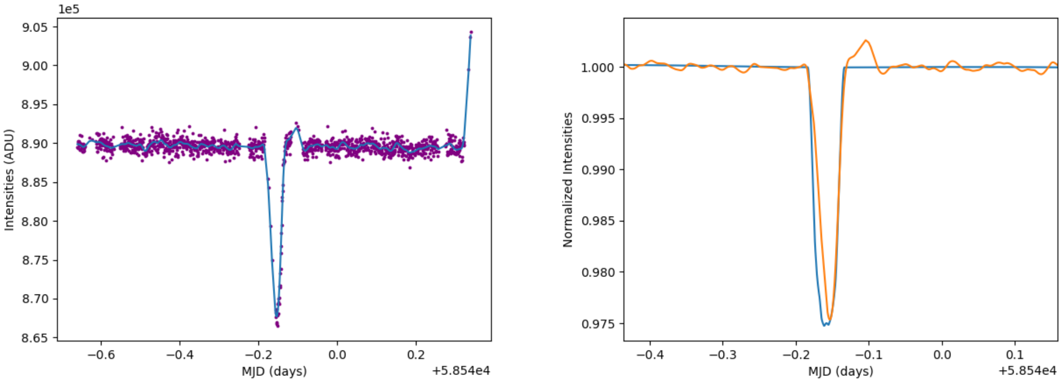
One shouldn't forget that, the Initial Kernel Guess is responsible for the covariance matrix in the Get Outliers of the first order PLD. Hence, the photometric precision improvement of figure 3.7 is also due to the more precise SAA removal that it allows.

3.1.5 Update GP

Figure 3.7, not only showcases the importance of the initial guess, but also the importance of the kernel and its hyperparameters. This dataset was ran with a Matérn-3/2 kernel with very broad bounds in order to understand how L-BFGS-B performs. Different aperture sizes, combined with the pointing motion will produce completely different light-curves, whose bound constriction would have to be done individually. With bounds

$$\begin{gathered} 0.1 * w \leq w \leq 10 * w \\ 0.01 * a \leq a \leq 10000 * a \\ 0.0069 \leq t \leq 100 \end{gathered} \tag{3.1}$$

where w, a and t are respectively the white noise, the red noise amplitude and red noise timescale guesses as explained in section 2.3.2.1. The kernel hyperparameters for the first order PfC are presented in table 3.1 with the complete table displayed in Appendix C. We can see that the white-noise value chosen by the optimization method is more or less in agreement with the white-noise amplitude in the light-curve, being slightly smaller than



(A) (B)

FIGURE 3.8: The GP fit from the hyperparameters of an aperture with 16 pix of radis stacked against the PfC output for this same aperture a) and the simulated data without noise b)

the latter. The same cannot be said for the red noise amplitude which corresponds to 10% of the light-curve median level, while the transit is only 1% in depth. Nor for the red noise timescale, where only two of its values are inside the visit length (1 day). We can see in figure 3.8, the GP is not accurately representing the transit, with its biggest shortcomings being due to the missing data. Still, this proved enough for the PfC, since the PLD only requires a covariance functions that sufficiently represents the true covariance matrix of the astrophysical signal. Trying to restrain the values to the expected red noise amplitude and red noise timescale will make the L-BFGS-B too constricted and the GP unable to follow the transit pattern with every run ending at the boundary condition. We can however constrain the GP hyperparameters with:

$$\begin{gathered} 0.1 * w \leq w \leq 10 * w \\ 0.01 * a \leq a \leq 100 * a \\ 0.0069 \leq t \leq 5 \end{gathered} \tag{3.2}$$

and get values that come close to the 1% for the red noise amplitude and timescales inside the 1 day length as seen in table 3.2 (Appendix C). This apparent improvement is not meet with significant changes to the GP fit seen in figure 3.8, nor in the photometric precision as seen in figure 3.9 where the two sets of bounds are compared. The stricter bounds yeild results equal or worse than the previous ones. For looser bounds, the L-BFGS-B might lose one of its runs in a set of hyperparameters that are incredibly far from the expected ones, but still have a likelihood on par with the other sets. Still, three iterations of Update GP have proved efficient enough to keep those away from the PfC model.

Aperture radii (pix)	White (ADU)	Red Amp (ADU)	Red Time (Days)
16	1.171e+03	2.467e+05	2.511e-01
17	1.088e+03	7.439e+04	1.228e-01
18	1.082e+03	2.401e+05	2.950e-01
19	1.139e+03	2.225e+05	2.824e-01
20	1.216e+03	1.026e+05	1.433e-01
21	1.291e+03	2.308e+05	2.582e-01

TABLE 3.2: Update GP optimized hyperparameters for the first order PfC, for a Matérn-3/2 kernel with the bound specified in 3.2.

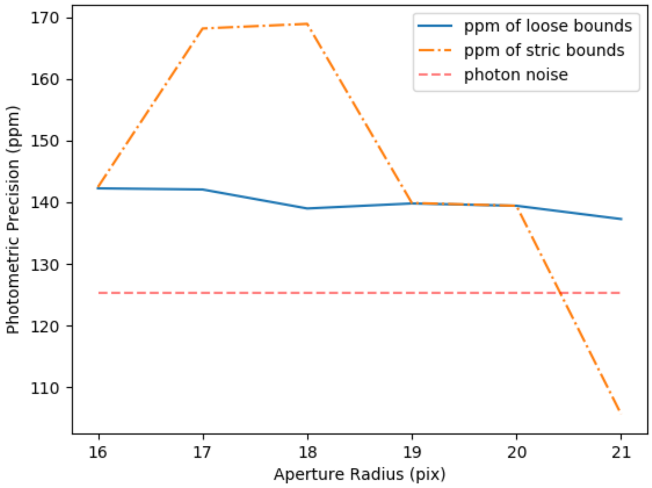
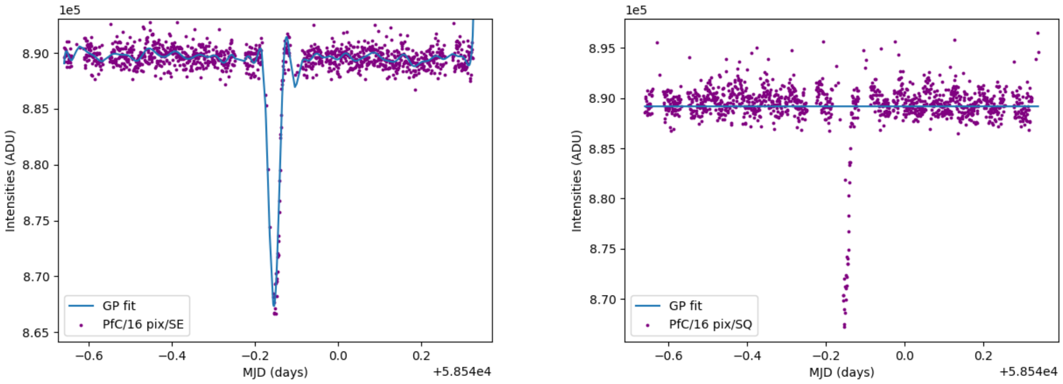


FIGURE 3.9: Photometric precision achieved with loose and strict bounds versus photometric aperture radius.

The difficulty in constraining the hyperparameters for the Matérn-3/2 kernel and their optimization results that don't translate accurately to the light-curve showcase that this is not the appropriate kernel to model a light-curve with a transit. The kernel must be pliable and effective in its purpose, and the Matérn-3/2 is neither. Its shortcomings don't compromise this dataset, since it does not model the instrumental noise even with its broad bounds. Moreover it is capable of roughly following the transit.



(A)

(B)

FIGURE 3.10: a) has a light-curve detrended using a squared exponential kernel (163.4 ppm) and the GP fit. b) has the PfC output (213.6 ppm) for a quasi-periodic kernel as well as the GP fit

3.1.5.1 Kernel options

We wanted to see if another kernel would be able to better reproduce the transit signal. We tried the Quasi-Periodic and the Squared Exponential. With their bounds set to

$$\begin{array}{cc} 1 \leq a_{SE} \leq 1 \times 10^{6} * a & 0.1 * w \leq w_{SQ} \leq 10 * w \\ 6 \times 10^{-4} \leq t_{SE} \leq 10 & 1 \leq a_{SQ} \leq 10000 * a \\ & 1 \times 10^{-5} \leq \tau_{SQ} \leq 100 \\ & 0.02 \leq \Gamma_{SQ} \leq 10 \end{array} \tag{3.3}$$

the WASP-43b simulation was detrended for a 16 pix radius aperture yielding the plots in figure 3.10. The final output of the PfC using the Squared Exponential, has a periodic systematic that is not simulated. This is a sign that the squared exponential is actively modeling the instrumental noise as an astrophysical signal. Changing the bounds for the Squared Exponential will either yield a kernel incapable of following the transit or a kernel similar to the current one, neither will improve figure 3.10 a). Furthermore, using the second order with the Squared Exponential leads to overfitting. As for the Quasi-Periodic kernel, it tries to follow a long timescale trend, thus it appears to be flat to the PLD. The end result is a light-curve barely reduced by the PfC.

3.2 WASP-18b

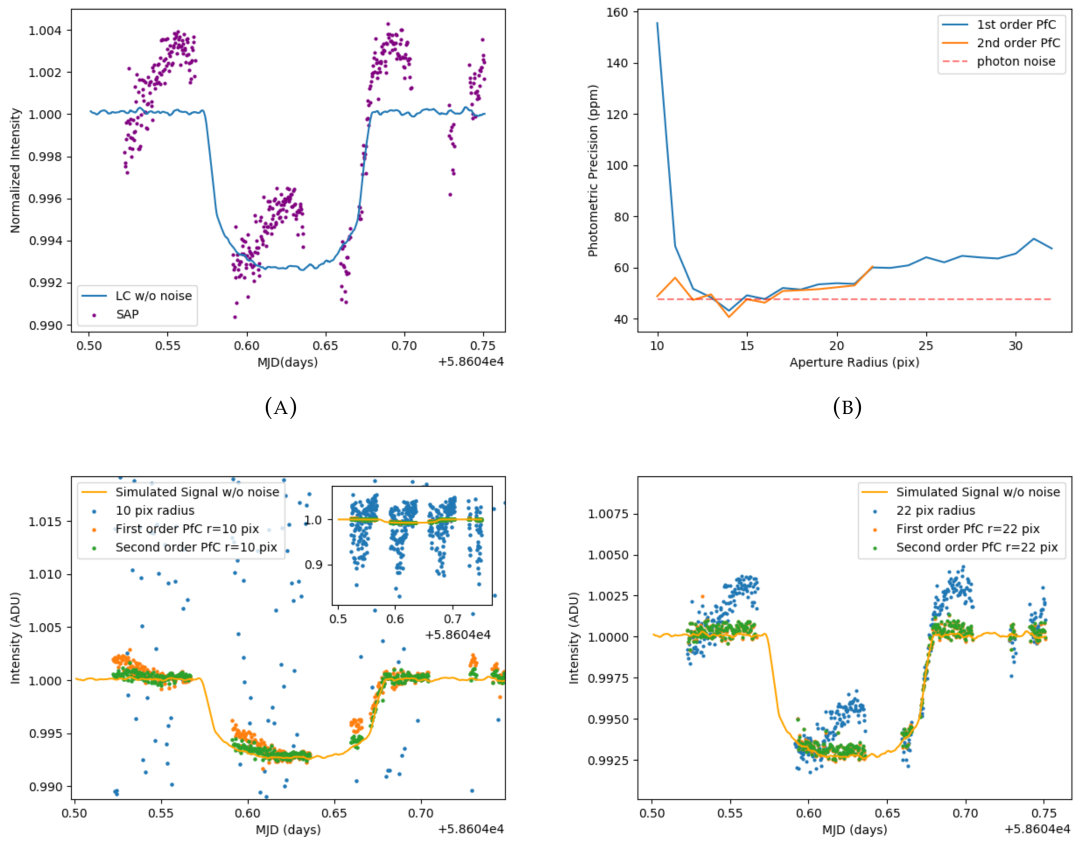
WASP-18b [59] is a Jovian planet orbiting a F6V type star of 9.3 m_v with $1.25 \pm 0.04\, M_{\odot}$ and $1.23 \pm 0.045\, R_{\odot}$ [60]. It has 10 times the mass of Jupiter, but only 1.165 R_J [61]. It has one of the shortest-periods found, 0.94 d, which allied to its large mass makes him an important target for the study of tidal dissipation of a planet/star system [62].

CHEOPSim is again used to simulate the dataset. Its SAP for a 20 pix aperture is displayed with a pure astrophysical signal with photon noise, figure 3.11 a). The light-curve displays several differences when compared to WASP-43b among them: its shorter lenght with 416 data points (208 minutes) obtained through 30 second exposures, compared to the day long light-curve of WASP-43b with a 60 seconds cadence. Data contaminated with the SAA is not sent to the observer, easing the job of clipping the outlier. The transit comprises a larger chunk of the light-curve easing the SavGol filtering and the GP model. The biggest challenge for the PfC will be the frequent and large interruptions present in the light-curve, due to the SAA, where data is not available. Those gaps might complexify the GP Update, since during the gaps the GP will be unconstrained.

The simulations were made using the spectral type F5V for the parent star, instead of the F6V. Through the CHEOPS Estimation Time Calculater we can see that for this simulation, with a 15 minute integration time: the instrumental noise is 86.13 ppm, the photon noise is 47.83 ppm and the combined noise estimate is set at 103.79 ppm, accounting for interruptions.

3.2.1 Optimal Aperture and PLD order

Different aperture sizes were used to detrend WASP-18b's light-curve, figure 3.11 b). As before, the smaller apertures (10-13) are not having a good detrend from the first order PfC. We maintain that the most likely cause is the first order PLD not being capable of modeling the instrumental noise with linear combinations of the pixel values due to the ingress and egress motion of the star in the aperture. The second order PLD is capable of mitigating the instrumental noise and of following closely the transit shape (figure 3.11 c)) for this apertures. This improvement did not happen for WASP-43b in this aperture size, because the SAA was not properly removed. Larger apertures, in this dataset, are not prone to overfitting due to neighboring stars, since none is in the vicinity of the target. In fact, greater apertures improve the SAP considerably, hence in theory they should yield better values. In reality the photometric precision reaches a minimum for a 14 pix radius aperture and starts to steadily increase for greater aperture sizes. This is caused by the background noise and the read out noise that rise for bigger apertures.



(A)

(B)

(C)

(D)

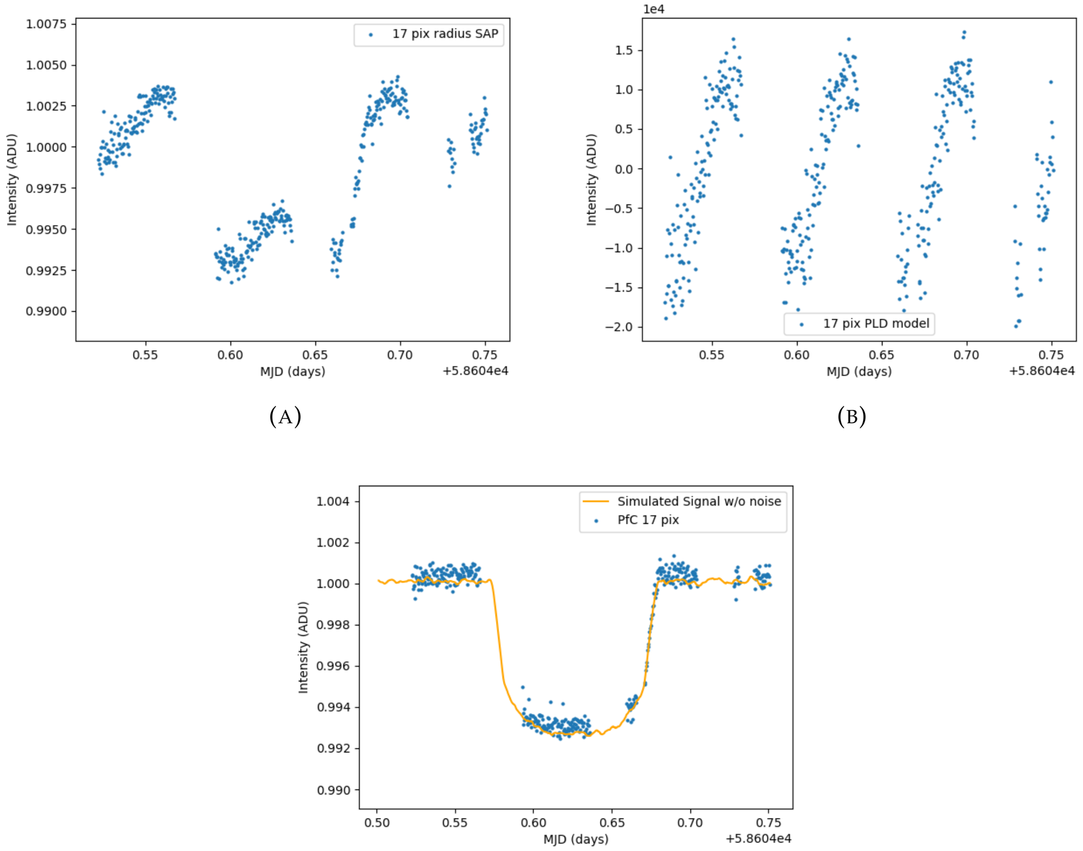
FIGURE 3.11: The simulated astrophysical signal (blue), stacked against the SAP for a 20 pix radius aperture (purple) a). b) describes the photometric precision for different aperture sizes. c) has inscribed the progression from the SAP to the second order PfC for a 10 pix radius aperture. d) has the same progression for a 22 pix radius aperture.

Bigger apertures don't seem to require the second order PLD, with the light-curves having both similar features and photometric precision to the detrends of the first order PLD.

Aperture 12-16 seem to be beyond the photon noise for this target. It is however difficult to access if they actually are. Their light-curves follow the simulated astrophysical signal incredibly well. They are not smooth due to the bleaching of white-noise, nor do they exibit any other sign of overfitting.

3.2.2 PLD Model

The 17 pix radius aperture is chosen to represent the PfC output of this dataset, figures 3.12. It achieves 50.86 ppm and the light-curve accuratly represents the simulated astrophysical signal as seen in figure 3.12 c). The $\sim$ 100 minute trend in figure 3.12 a) is all but



(A)

(B)

(C)

FIGURE 3.12: SAP from a 17 pix radius aperture, displaying correlated noise with the same period of the satellite rotation a). A second order PfC b) with the same periodicity of CHEOPS rotation. Output of a second order PfC c).

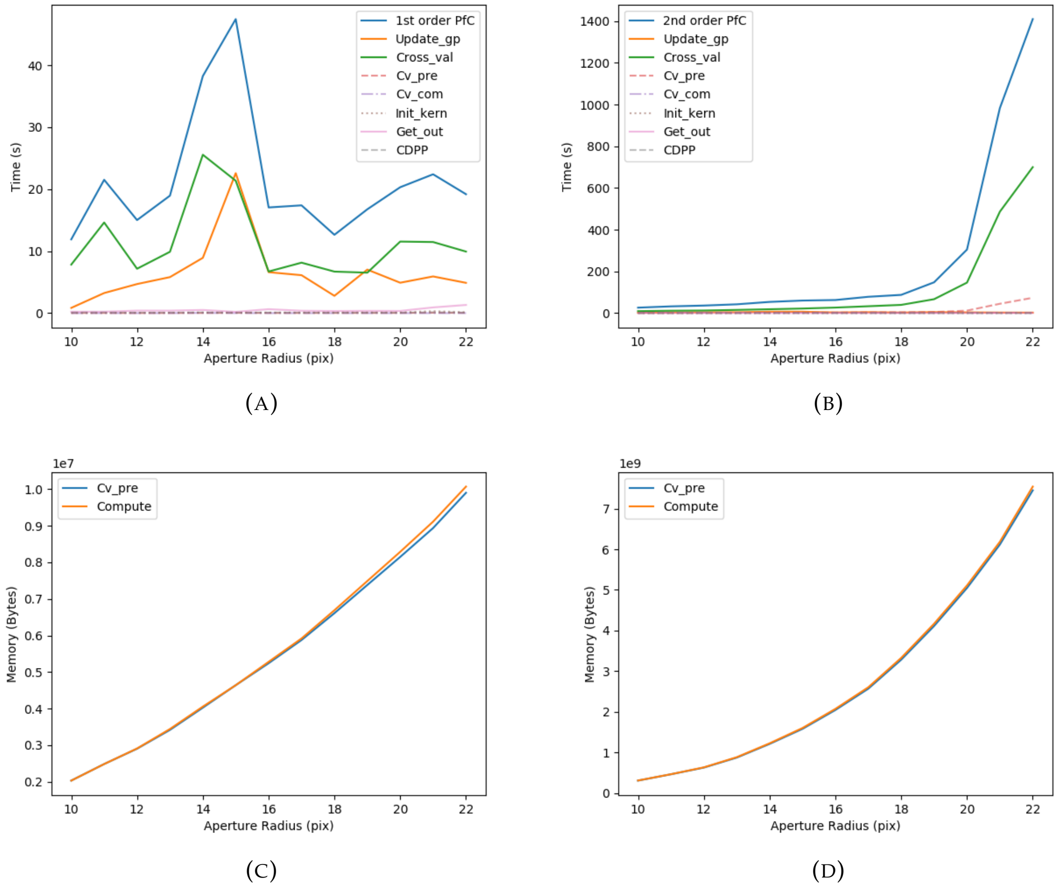
gone, proving that the PLD is capable of eliminating the instrumental noise for a CHEOPS dataset.

3.2.2.1 Memory and Runtime

The memory requirements and the runtime plots remain very similar to the ones of WASP-43b, figure 3.13. In figure 3.13 a) there is an unforeseen jitter in the Cross validation runs. It might be attributed to other processes running during this method. Due to the small number of datapoints (416), the second order PfC can be extended to a 22 pix radius aperture.

3.2.3 Initial Kernel Guess

For this dataset, the PfC's Initial Kernel Guess proved to be better than using the guess from EVEREST. The Update GP continues to be restrained by the bad seed of EVEREST's



(A)

(B)

(C)

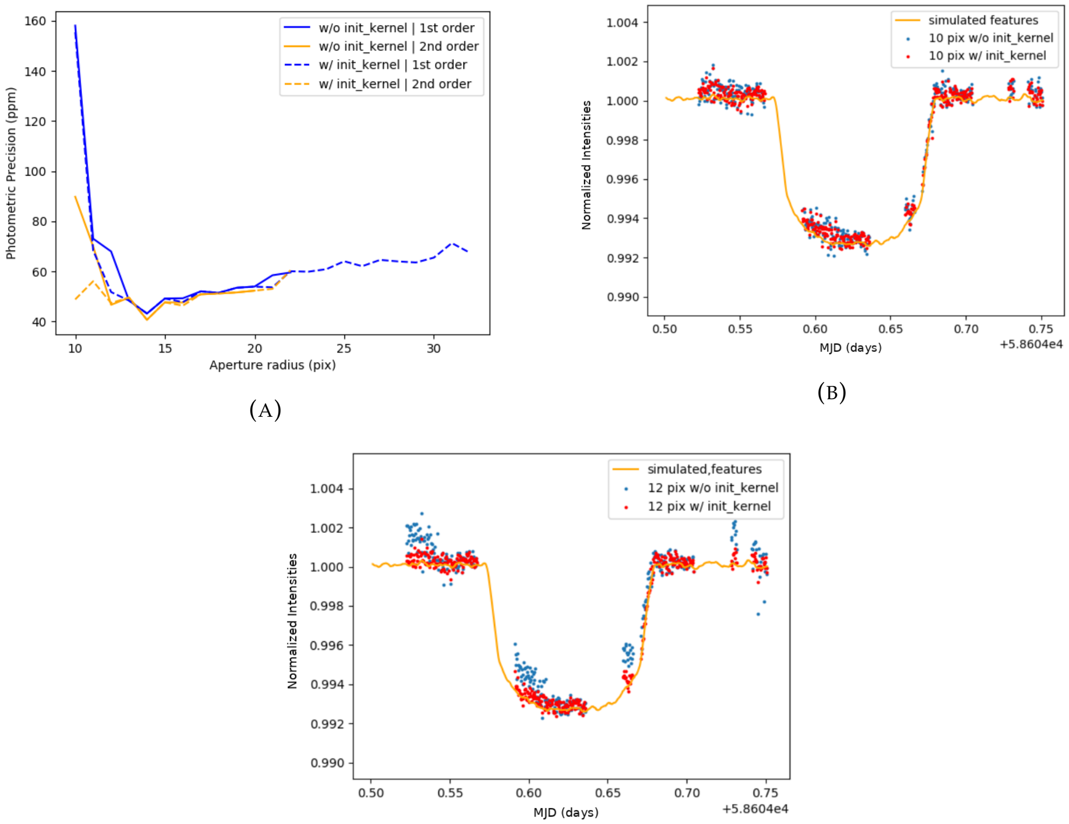
(D)

FIGURE 3.13: Runtimes for the first a) and second order PfC b), as well as the memory usage, c) and d) for each.

guess, with it exiting the hyperparameter search with bound proximity in nearly every run. This time the photometric precision remains closer between both. The 12 (figure 3.14 c)) and 21 pixels radius aperture yield worse photometric precision for the first order PLD. Apertures 10 (figure 3.14 b)) and 11 are their counterparts in the second order. While the absence of the Initial Kernel Guess was felt in the Update GP, the latter's constriction was not generally felt by the PLD.

3.2.4 Update GP

Keeping the constrains of 3.1 in the Matérn-3/2 kernel hyperparameters, the Update GP outputs for both orders are specified in table 3.3 (Appendix C). The hyper parameters chosen are not in sync with the data, with the white noise in the first order being the only exception. The red noise amplitude has the same order of magnitude of the light-curve, when in fact it should only be around 1%. The red noise timescale is beyond the visit



(A)

(B)

(C)

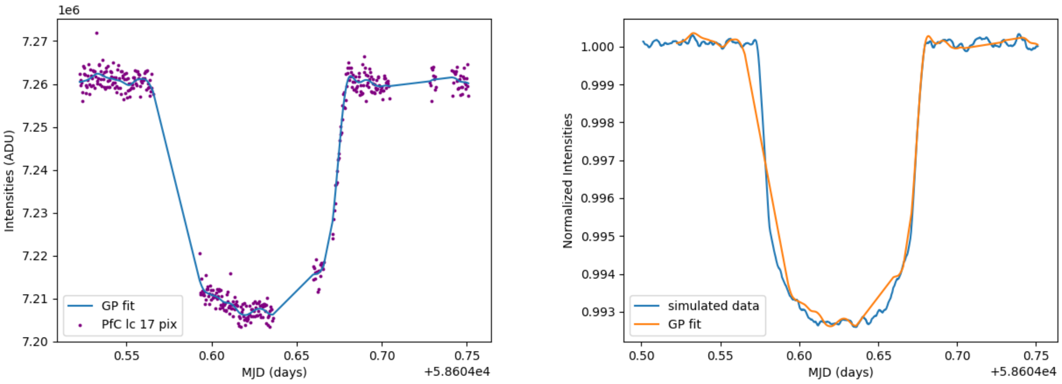
FIGURE 3.14: Photometric precision for the PfC with its Initial Kernel Guess (dashed) and using EVEREST's (solid) a). First order c) and second order b) examples where the PfC is affected by EVEREST's hyperparameter guess.

PLD order	White (ADU)	Red Amp (ADU)	Red Time (Days)
1	8.195e+03	6.396e+06	1.016e+00
2	2.117e+02	6.242e+06	2.190e+00

TABLE 3.3: Update GP optimized hyperparameters for the first and second order PLD, for a Matérn-3/2 kernel with the bound specified in 3.1.

length of 3.5 hours, having more than a day. For the second order the results remain disproportional for the red noise (amplitude and timescale), but now they also are for the white noise being one order of magnitude less than supposed. In figure 3.15 b) we can see that the hyperparameters for the second order give a pretty good fit to the simulated astrophysical signal. The GP fit finds its weakest moments during the gaps from the SAA and among its best during the egress of the transit. This problem could be solved setting a mean function in the GP model that accurately depicts the astrophysical data.

Constraining the bounds to equations 3.2 did improve the hyperparameters values without sacrificing the photometric precision, yielding the red noise time scale under a



(A)

(B)

FIGURE 3.15: GP fit for the second order PLD using the hyperparameterso of table 3.3 a). And how the fit stacks against the astrophysical signal it should model b).

PLD order	White (ADU)	Red Amp (ADU)	Red Time (Days)
1	8.371e+03	3.389e+05	6.758e-02
2	1.971e+02	2.435e+05	1.427e-01

TABLE 3.4: Update GP optimized hyperparameters for the first and second order PLD, for a Matérn-3/2 kernel with the bound specified in 3.4.

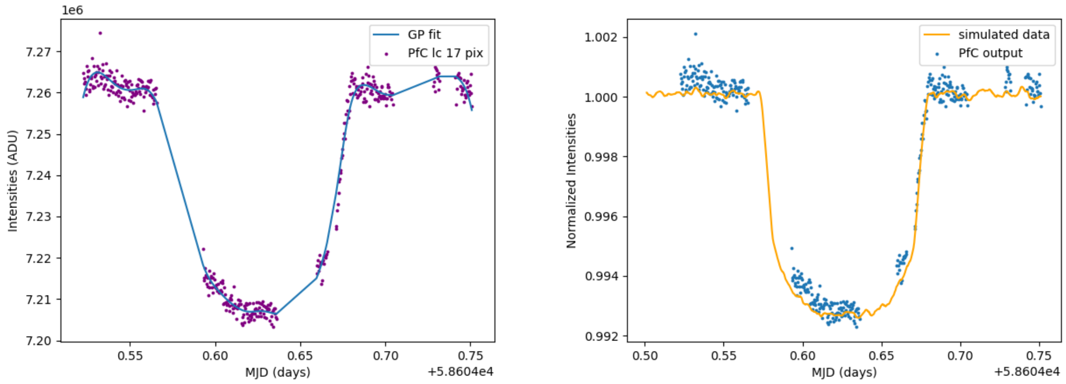
day. For that reason we went further and constrained the values to equations 3.4.

$$\begin{aligned} 0.1 * w \leq w \leq 10 * w \\ 0.01 * a \leq a \leq 10 * a \\ 0.0069 \leq t \leq 0.8 \end{aligned} \tag{3.4}$$

The output hyperparameters are displayed in table 3.4. Now the only value that does not respect the light-curve characteristics is the white noise value of the second order that is one order of magnitude smaller than expected. And while this values are met with an improvement in the photometric precision of the output, the light-curve will be deformed, not following the simulated data, figure 3.16 b).

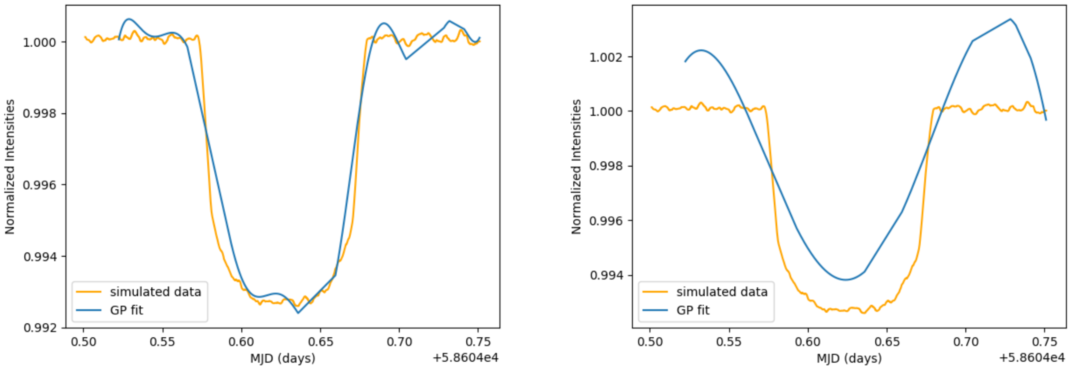
3.2.4.1 Kernel options

The same bounds of 3.3 were used for the Quasi-Periodic and the Squared Exponential kernels. Neither was able to emulate, the simulated data as the Matérn-3/2 kernel was able to, with the Squared Exponential having a much better attempt than the Quasi-Periodic, figure 3.17.



(A) (B)

FIGURE 3.16: a) has the GP fit with the 3.4 bounds to the data and b) yields the reduced data next to the simulated signal.

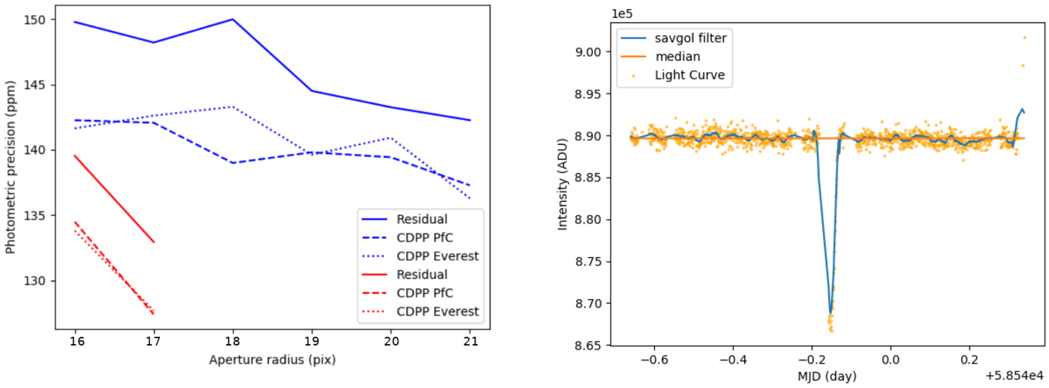


(A) (B)

FIGURE 3.17: a) Squared Exponential fit of the light-curve compared to the simulated astrophysical signal. b) has the same plot for the Quasi-Periodic.

3.3 CDPP

You can recall that the CDPP is a numerical method used to estimate the photometric precision over a given timescale. The CDPP routine (described in section 2.3.2.2) involved the filtering (done with the SavGol filter) of the astrophysical signal to access both the instrumental and the photon noise precision of the light-curve. In this section we want to check if this filtering doesn't also remove part of the instrumental and photon noise, thus artificially increasing the quality of the light-curve. With that objective we are going to compute, for the light-curve of WASP-43b and WASP-18b, the CDPP implemented in the EVEREST pipeline, the modified version that we implemented for PfC and the CDPP of PfC without the filtering and outliers removal made on the residuals light-curve (light-curve to which we have removed the artificial signal that is provided with these simulated



(A)

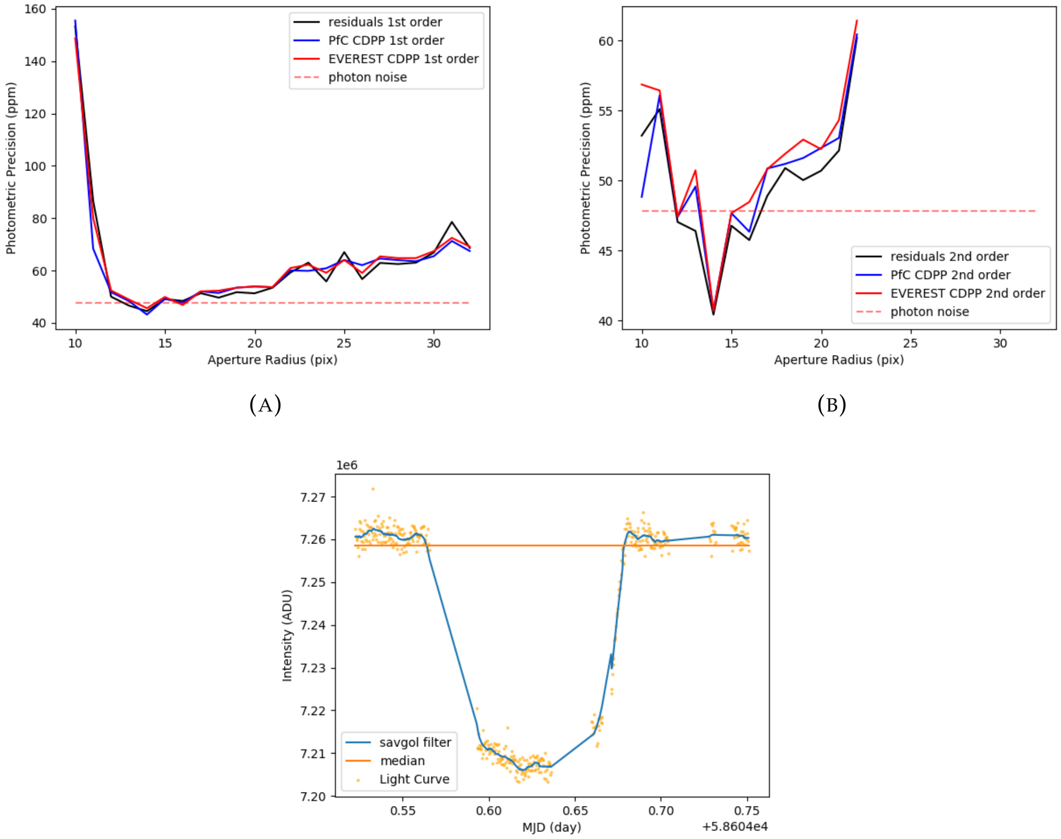
(B)

FIGURE 3.18: The photometric precision from both CDPPs are compared with the photometric precision obtained through the residual of the simulated data a). This are the results for WASP-43b. The blue data comes from the first order PfC and the second order is in red. The SavGol filter with a window of 41 points is displayed in b)

datasets). Both CDPP's use a SavGol window of 41 points (41/20.5 minutes)[*] and the derivation is turned off (see section 2.3.2.2). The photometric precision is calculated for an integration time of 30/15 minutes[†]. Figure 3.18 a) displays the results for WASP-43b and both versions of the CDPP are removing some of the instrumental noise, having their photometric precision under the residuals' for both the first and second order. For the CDPP of the PfC the difference to the residual was between 4-11 ppm in the first order and 5 ppm for the second. For EVEREST's it was between 2-8 ppm for the first order. Apertures 22-25 remain well bellow the photon-noise ($\sim$ 100) and their light-curves show signs of overfittings: white noise following an unsimulated a periodic signal.

The goal of the SavGol Filter is to model the long-term features of the light-curve and to also model the transit. This is only possible if the SavGol window is at most 41 points. From image 3.18 b), we can see that the goal of following the transit even with the removed outliers was largely achieved. Unfortunately 41 points is small enough to filter some of the noise present in the detrended light-curve. Moreover, both CDPPs remove outliers that still might exist in the light-curve widening the gap to the residuals. EVEREST's CDPP is closer to the values from the residuals, because the gaps that exist in the light-curve are removed in its SavGol filter. Mismatched data that is joined together will be badly filtered by the SavGol filter, thus introducing noise. The biggest perpetrator is the gap present during the transit. This does not happen for the PfC's CDPP, because the

[*] 41 minutes for the WASP-43b and 20.5 minutes for the WASP-18b
[†] 30 minutes for the WASP-43b and 15 minutes for the WASP-18b



(A)

(B)

(C)

FIGURE 3.19: Residuals are compared with both instances of the CDPP in the first a) and second order b), to determine how well they achieve the photometric precision. c) has the output of the SavGol filter implemented for the PfC from the reduced dataset.

beginning and the end of the gaps remain distanced by their temporal values in the SavGol filter implemented. Both CDPPs can be made to have the same ppm has the residuals by increasing the SavGol window to 121 points (121 minutes). However this agreement is artificial since it's due to a bad filtering of the transit and not to a correct assessment of the instrumental noise.

For WASP-18b, both instances of the CDPP (PfC and EVEREST) fared reasonably well, with neither having a clear gap in the first order to the residuals as seen for WASP-43b. For the second order both CDPPs yield lower photometric precision than the residuals, with the PfC's being consistently closer to them, 3.19 b). The SavGol filter behaves as intended mitigating the noise added to the photometric precision due to the discontinuities. Still, it erodes a small portion of the noise of the light-curve, figure 3.19 c).

Chapter 4

Conclusion and Future Work

Pixel Level Decorrelation was successfully applied to simulated data of CHEOPS. The PfC was capable of reducing the instrumental noise to a bare minimum, coming close to the photon noise limited precision for both sets of data (WASP-18b and WASP-43b). Both reductions were able to accurately preserve the simulated data, with the depths and lengths of the transits being the correct ones.

As expected, crowded apertures remain a difficult problem for the PLD to solve with their instance in the WASP-43b dataset leading to overfit. Memory constrains set the second order PLD as the highest order available for CHEOPS due to its high number of pixels. From the datasets studied, the first order PLD was capable of reducing bigger apertures ($>$ 16 pixel radius), but broke down for the smaller ones ($<$ 13 pixel radius). This is due to the variability introduced with the in and out motion of the star in the photometric aperture. The linear approach of the first order PLD is not sufficient to model this type of instrumental noise. The second order PLD was capable of removing most of the instrumental noise in all the studied apertures. With the exception of apertures with $>$ 20 pix radius who have their photometric precision lowered by the increase in both the photon and background noise.

The Initial Kernel Guess that employs a first order PLD with 15 regressors also proved to work well. The hyperparameter guesses that come from it facilitate the job of the Update GP and are fundamental in the removal of outliers in the first order PLD.

The Matérn-3/2 kernel proved sufficient for the modeling of the overall stellar signal. Its hyperparameters from the Update GP did not meet the expected values seen in the light-curve features, partially, because this kernel is not the best to model a light-curve with a transit.

The SavGol filter implemented for the PfC is a step in the right direction for an automated and accurate determination of the photometric precision of a light-curve with the CDPP. It is capable of precisely filtering datasets that are unevenly spaced. However, for the datasets studied here, there is no clear difference between the EVEREST's and the PfC's CDPPs.

During the last stage of my thesis, I identify several avenues to further characterize and improve the performance of the PfC:

- The Matérn-3/2 kernel showed signs of strain while modeling the transits. This is not a problem for the EVEREST pipeline since it actively masks them. A possible solution, would be implementing a transit model similar to the one used for the Spitzer (see section 2.1). Or implementing a mean function in the GP that models the stellar signal. This would also help constricting the GP model during the SAA gaps.

- Evaluating the viability of the neighboring Pixel Level Decorrelation (nPLD) from EVEREST [3] in CHEOPS data. nPLD includes in the design matrix pixels from neighboring stars to better identify and correct instrumental noise (like temperature variations) which would affect the whole CCD and thus other stars observed at the same time. Unlike for K2, obtaining the light-curves from the neighboring stars is not an easy task for CHEOPS data. The rotating field of view means that these stars need to be precisely tracked, which goes against the zen of PLD. Fortunately this is already done by the ARCHI pipeline [63]. Joining the photometry outputs of ARCHI to the design matrix of the PfC might yield a quick but interesting study as to the viability of this method.

- Study the PfC behaviour for different stars, with different magnitudes and different amounts of crowding. See the performance changes in the PfC when its inputs come from the CHEOPs Data Reduction Pipeline. They would allow a more comprehensive characterization of the performance of the PfC.

Appendix A

Saturated Pixel Demonstration

The deduction from equation 2.21 to 2.22 is:

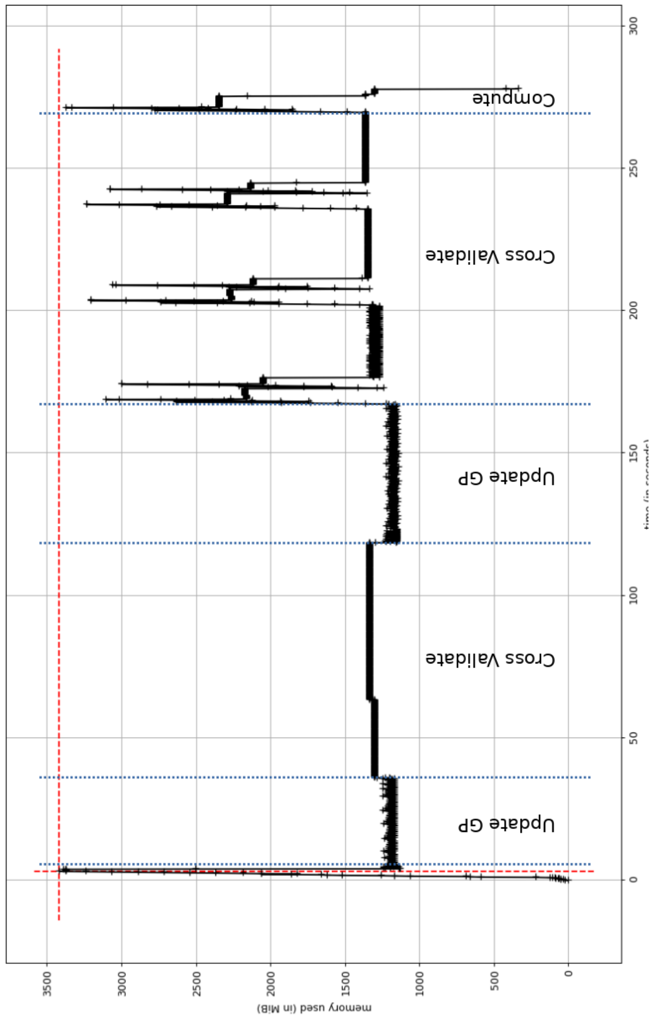
$$\begin{aligned}
\frac{p_{il}}{\sum_k p_{ik}} &= \frac{a_{il}\tau_i + bil}{\sum_k a_{ik}\tau_i + b_{ik}} \\
&= \frac{a_{al}\tau_i(1 + \frac{b_{il}}{a_{il}\tau_i})}{\sum_k a_{ik}\tau_i(1 + \frac{\sum_k b_{ik}}{\sum_k a_{ik}\tau_i})} \\
&= \frac{a_{il}}{\sum_k a_{ik}} \left(\frac{1 + \frac{\sum_k b_{ik}}{\sum_k a_{ik}\tau_i} - \frac{\sum_k b_{ik}}{\sum_k a_{ik}\tau_i} + \frac{b_{il}}{a_{il}\tau_i}}{1 + \frac{\sum_k b_{ik}}{\sum_k a_{ik}\tau_i}} \right) \\
&= \frac{a_{il}}{\sum_k a_{ik}} \left(1 + \frac{(\frac{b_{il}}{a_{il}} - \frac{\sum_k b_{ik}}{\sum_k a_{ik}}) \times \frac{1}{\tau_i}}{1 + \frac{\sum_k b_{ki}}{\sum_k a_{ik}\tau_i}} \right) \\
&= \frac{a_{il}}{\sum_k a_{ik}} \left(1 + \frac{\frac{\Delta}{\tau_i}}{1 + \frac{\sum_k b_{ik}}{\sum_k a_{ik}\tau_i}} \right)
\end{aligned}$$

if $1 >> \frac{\sum_k b_{ik}}{\sum_k a_{ik}\tau_i}$ then:

$$\frac{p_{il}}{\sum_k p_{ik}} = \frac{a_{il}}{\sum_k a_{ik}} \left(1 + \frac{\Delta}{\tau_i} \right)$$

Appendix B

PfC memory usage over time



Appendix C

L-BFGS-B tables

The full tables of the GP Update made with the L-BFGS-B for WASP-43b are:

Matérn-3/2 kernel - bounds 3.1				
Aperture	White noise	Red noise Amplitude	Red noise Timescale	Log Likelihood
16	1.170e+03	7.080e+05	9.971e-01	-1.021e+04
	1.186e+03	1.746e+06	1.699e+00	-1.023e+04
	1.164e+03	6.926e+05	9.112e-01	-1.021e+04
17	1.057e+03	1.558e+07	7.936e+01	-9.855e+03
	1.061e+03	8.686e+05	1.567e+00	-9.851e+03
	1.050e+03	5.873e+05	8.471e-01	-9.852e+03
18	1.082e+03	1.326e+06	2.972e+00	-9.604e+03
	2.143e+03	8.364e+04	2.615e+00	-1.006e+04
	1.893e+03	8.316e+04	3.592e+01	-1.016e+04
19	1.164e+03	9.078e+04	1.383e-01	-9.554e+03
	1.127e+03	6.223e+05	1.052e+00	-9.449e+03
	1.163e+03	9.287e+04	1.391e-01	-9.548e+03
20	1.218e+03	9.979e+04	1.419e-01	-9.530e+03
	2.231e+03	1.026e+05	8.203e+00	-9.808e+03
	1.217e+03	1.021e+05	1.430e-01	-9.526e+03
21	2.085e+03	1.103e+05	9.659e+01	-9.964e+03
	1.274e+03	6.764e+05	9.143e-01	-9.609e+03
	1.293e+03	1.032e+06	1.772e+00	-9.609e+03

Matérn-3/2 kernel - bounds 3.2

Aperture	White noise	Red noise Amplitude	Red noise Timescale	Log Likelihood
16	1.212e+03	7.302e+04	1.003e-01	-1.039e+04
	1.171e+03	2.467e+05	2.511e-01	-1.022e+04
	7.188e+02	1.801e+05	1.446e+00	-1.069e+04
17	1.090e+03	7.231e+04	1.224e-01	-1.002e+04
	1.088e+03	7.439e+04	1.228e-01	-1.001e+04
	1.089e+03	7.396e+04	1.227e-01	-1.001e+04
18	1.110e+03	8.130e+04	1.227e-01	-9.740e+03
	1.082e+03	2.401e+05	2.950e-01	-9.612e+03
	1.109e+03	8.317e+04	1.232e-01	-9.733e+03
19	1.164e+03	9.076e+04	1.382e-01	-9.554e+03
	1.139e+03	2.225e+05	2.824e-01	-9.459e+03
	1.163e+03	9.285e+04	1.390e-01	-9.548e+03
20	1.218e+03	9.978e+04	1.419e-01	9.530e+03
	1.216e+03	1.026e+05	1.433e-01	-9.525e+03
	1.217e+03	1.0208e+05	1.430e-01	-9.526e+03
21	1.312e+03	1.102e+05	1.374e-01	-9.678e+03
	1.291e+03	2.308e+05	2.582e-01	-9.619e+03
	1.310e+03	1.128e+05	1.386e-01	-9.675e+03

Squared Exponential kernel - bounds 3.3

Aperture	Amplitude	Timescale	Log Likelihood
16	8.635e+04	8.479e-04	-1.238e+04,0
	6.653e+04	9.037e-04	-1.291e+04,0
	7.265e+04	8.794e-04	-1.270e+04,0

Quasi-Periodic kernel - bounds 3.3

Aperture	White Noise	Red Amplitude	Red Timescale	Γ	Log Likelihood
16	2.386e+03	6.486e+05	8.551e+01	1.000e+01	-1.076e+04
	2.577e+03	7.207e+04	4.578e+01	7.863e+00	-1.084e+04
	2.387e+03	6.777e+05	3.125e+01	1.000e+01	-1.076e+04

For WASP-18b the tables are:

Matérn-3/2 kernel - bounds 3.1

PLD Order	White noise	Red noise Amplitude	Red noise Timescale	Log Likelihood
1	8.196e+03	6.396e+06	1.015e+00	-4.375e+03
	2.433e+04	5.298e+05	1.863e-02	-4.771e+03
	1.524e+04	1.105e+05	2.804e-01	-7.268e+03
2	2.117e+02	6.242e+06	2.190e+00	-3.669e+03
	2.117e+02	6.802e+06	2.456e+00	-3.669e+03
	2.117e+02	6.643e+06	2.380e+00	-3.669e+03

Matérn-3/2 kernel - bounds 3.2

PLD Order	White noise	Red noise Amplitude	Red noise Timescale	Log Likelihood
1	8.167e+03	3.389e+06	4.719e-01	-4.376e+03
	1.555e+04	1.145e+05	2.723e-01	-7.089e+03
	1.524e+04	1.105e+05	2.804e-01	-7.268e+03
2	8.105e+03	4.478e+05	2.688e-01	-4.156e+03
	2.116e+02	1.030e+06	2.154e-01	-3.696e+03
	2.116e+02	2.262e+06	5.707e-01	-3.673e+03

Matérn-3/2 kernel - bounds 3.4

PLD Order	White noise	Red noise Amplitude	Red noise Timescale	Log Likelihood
1	7.703e+03	3.385e+05	6.649e-02	-4.706e+03
	8.371e+03	3.389e+05	6.758e-02	-4.703e+03
	1.524e+04	1.105e+05	2.804e-01	-7.268e+03
2	1.971e+02	2.435e+05	1.428e-01	-4.259e+03
	6.839e+03	1.029e+05	5.102e-01	-6.889e+03
	8.005e+03	2.262e+05	4.022e-01	-4.630e+03

Quasi-Periodic Kernel - bounds 3.2					
PLD Order	White noise	Red Amplitude	Red Timescale	Γ	Log Likelihood
1	8.553e+03	4.679e+06	4.701e+01	4.210e-01	-4.401e+03
	2.216e+04	5.565e+06	1.264e+00	2.009e-02	-4.740e+03
	2.215e+04	5.754e+06	1.258e+00	2.536e-02	-4.740e+03
2	4.521e+03	4.767e+06	5.157e+01	1.424e+00	-4.023e+03
	1.170e+03	4.172e+06	4.446e+01	3.388e-01	-3.815e+03
	9.313e+03	5.680e+06	5.796e+01	5.873e-01	-4.117e+03

Squared Exponential Kernel - bounds 3.2			
PLD Order	Amplitude	Timescale	Log Likelihood
1	1.361e+05	1.406e-02	-2.256e+04
	1.049e+05	1.449e-02	-2.447e+04
	1.145e+05	1.944e-03	-1.314e+04
2	1.297e+05	3.145e-02	-7.296e+03
	1.138e+05	3.176e-02	-8.064e+03
	1.117e+05	3.179e-02	-8.188e+03

Bibliography

[1] D. Deming, H. Knutson, J. Kammer, B. J. Fulton, J. Ingalls, S. Carey, A. Burrows, J. J. Fortney, K. Todorov, E. Agol, N. Cowan, J.-M. Desert, J. Fraine, J. Langton, C. Morley, and A. P. Showman, "SPITZERSECONDARY ECLIPSES OF THE DENSE, MODESTLY-IRRADIATED, GIANT EXOPLANET HAT-p-$20{\rm b}$ USING PIXEL-LEVEL DECORRELATION," *The Astrophysical Journal*, vol. 805, no. 2, p. 132, May 2015. [Online]. Available: https://doi.org/10.1088/0004-637x/805/2/132

[2] R. Luger, E. Agol, E. Kruse, R. Barnes, A. Becker, D. Foreman-Mackey, and D. Deming, "EVEREST: PIXEL LEVEL DECORRELATION OFK2 LIGHT CURVES," *The Astronomical Journal*, vol. 152, no. 4, p. 100, Oct. 2016. [Online]. Available: https://doi.org/10.3847/0004-6256/152/4/100

[3] R. Luger, E. Kruse, D. Foreman-Mackey, E. Agol, and N. Saunders, "An update to the EVEREST k2 pipeline: Short cadence, saturated stars, and kepler-like photometry down to kp = 15," *The Astronomical Journal*, vol. 156, no. 3, p. 99, Aug. 2018. [Online]. Available: https://doi.org/10.3847/1538-3881/aad230

[4] T. C. S. Team, *CHEOPS Definition Study Report*. European Space Agency, 2013.

[5] D. Futyan, "Cheopsim user manuel," 2019.

[6] "Cheops exposure time calculator." [Online]. Available: https://cheops.unige.ch/pht2/exposure-time-calculator

[7] T. Donahue, K. Trivers, and D. Abramson, *Planetary Sciences American and Soviet Research*. National Academy Press, 1989.

[8] J. G. Ingalls, J. E. Krick, S. J. Carey, S. Laine, J. A. Surace, W. J. Glaccum, C. C. Grillmair, and P. J. Lowrance, "Intra-pixel gain variations and high-precision photometry with the infrared array camera (IRAC)," in *Space Telescopes and*

Instrumentation 2012: Optical, Infrared, and Millimeter Wave, M. C. Clampin, G. G. Fazio, H. A. MacEwen, and J. M. Oschmann, Eds. SPIE, Aug. 2012. [Online]. Available: https://doi.org/10.1117/12.926947

[9] S. B. Howell, C. Sobeck, M. Haas, M. Still, T. Barclay, F. Mullally, J. Troeltzsch, S. Aigrain, S. T. Bryson, D. Caldwell, W. J. Chaplin, W. D. Cochran, D. Huber, G. W. Marcy, A. Miglio, J. R. Najita, M. Smith, J. D. Twicken, and J. J. Fortney, "The k2 mission: Characterization and early results," *Publications of the Astronomical Society of the Pacific*, vol. 126, no. 938, pp. 398–408, apr 2014. [Online]. Available: https://doi.org/10.1086%2F676406

[10] M. Mayor and D. Queloz, "A jupiter-mass companion to a solar-type star," *Nature*, vol. 378, no. 6555, pp. 355–359, Nov. 1995. [Online]. Available: https://doi.org/10.1038/378355a0

[11] A. Baranne, D. Queloz, M. Mayor, G. Adrianzyk, G. Knispel, D. Kohler, D. Lacroix, J. Meunier, G. Rimbaud, and A. Vin, "ELODIE: A spectrograph for accurate radial velocity measurements." *aaps*, vol. 119, pp. 373–390, Oct. 1996.

[12] D. N. C. Lin, P. Bodenheimer, and D. C. Richardson, "Orbital migration of the planetary companion of 51 pegasi to its present location," *Nature*, vol. 380, no. 6575, pp. 606–607, Apr. 1996. [Online]. Available: https://doi.org/10.1038/380606a0

[13] J. D. Scargle, "Studies in astronomical time series analysis. II - statistical aspects of spectral analysis of unevenly spaced data," *The Astrophysical Journal*, vol. 263, p. 835, Dec. 1982. [Online]. Available: https://doi.org/10.1086/160554

[14] J. N. Winn, "Transits and occultations," 2010.

[15] H. J. Deeg and R. Alonso, *Transit Photometry as an Exoplanet Discovery Method*. Cham: Springer International Publishing, 2018, pp. 1–25. [Online]. Available: https://doi.org/10.1007/978-3-319-30648-3_117-1

[16] A. Lger, D. Rouan, J. Schneider, P. Barge, and M. Fridlund, "Transiting exoplanets from the corot space mission," *Astronomy and Astrophysics*, vol. 506, no. 1, pp. 287–302, Aug. 2009. [Online]. Available: https://doi.org/10.1051/0004-6361/200911933

[17] F. W. Wagner, N. Tosi, F. Sohl, H. Rauer, and T. Spohn, "Rocky super-earth interiors," *Astronomy and Astrophysics*, vol. 541, p. A103, May 2012. [Online]. Available: https://doi.org/10.1051/0004-6361/201118441

[18] D. Charbonneau, T. M. Brown, D. W. Latham, and M. Mayor, "Detection of planetary transits across a sun-like star," *The Astrophysical Journal*, vol. 529, no. 1, pp. L45–L48, Jan. 2000. [Online]. Available: https://doi.org/10.1086/312457

[19] G. Henry, G. Marcy, R. Butler, and S. Vogt, "Hd 209458," vol. 7307, no. 1, Nov. 1999.

[20] P. D. Sackett, "Searching for unseen planets via occultation and microlensing," 1998.

[21] X. Dumusque, A. S. Bonomo, R. D. Haywood, L. Malavolta, D. Ségransan, L. A. Buchhave, A. C. Cameron, D. W. Latham, E. Molinari, F. Pepe, and et al., "The kepler-10 planetary system revisited by harps-n: A hot rocky world and a solid neptune-mass planet," *The Astrophysical Journal*, vol. 789, no. 2, p. 154, Jun 2014. [Online]. Available: http://dx.doi.org/10.1088/0004-637X/789/2/154

[22] S. Hadden and Y. Lithwick, "DENSITIES AND ECCENTRICITIES OF 139keplerplanets FROM TRANSIT TIME VARIATIONS," *The Astrophysical Journal*, vol. 787, no. 1, p. 80, May 2014. [Online]. Available: https://doi.org/10.1088/0004-637x/787/1/80

[23] E. D. Lopez and J. J. Fortney, "UNDERSTANDING THE MASS-RADIUS RELATION FOR SUB-NEPTUNES: RADIUS AS a PROXY FOR COMPOSITION," *The Astrophysical Journal*, vol. 792, no. 1, p. 1, Aug. 2014. [Online]. Available: https://doi.org/10.1088/0004-637x/792/1/1

[24] A. P. Boss, "Proximity of jupiter-like planets to low-mass stars," *Science*, vol. 267, no. 5196, pp. 360–362, 1995. [Online]. Available: http://www.jstor.org/stable/2886240

[25] Y. Alibert, "On the radius of habitable planets," *Astronomy and Astrophysics*, vol. 561, Dec 2013. [Online]. Available: http://dx.doi.org/10.1051/0004-6361/201322293

[26] "The CoRoT satellite in flight: description and performance," *Astronomy & Astrophysics*, vol. 506, no. 1, pp. 411–424, Mar. 2009. [Online]. Available: https://doi.org/10.1051/0004-6361/200810860

[27] E. T. Hamden, F. Greer, M. E. Hoenk, J. Blacksberg, M. R. Dickie, S. Nikzad, D. C. Martin, and D. Schiminovich, "Ultraviolet antireflection coatings for use in silicon

detector design," *Applied Optics*, vol. 50, no. 21, p. 4180, Jul 2011. [Online]. Available: http://dx.doi.org/10.1364/AO.50.004180

[28] M. Richmond, "Notes from observational astronomy (physics 445) classes," 2010. [Online]. Available: http://spiff.rit.edu/classes/phys445/phys445.html

[29] D. Gardner, "Characterizing digital cameras with the photon transfer curve," 2002.

[30] S. B. Howell, *Handbook of CCD Astronomy*, 2nd ed., ser. Cambridge Observing Handbooks for Research Astronomers. Cambridge University Press, 2006.

[31] S. Hoyer, P. Guterman, O. Demangeon, S. Sousa, M. Deleuil, and J. Meunier, "Data reduction pipeline of cheops mission," 2019.

[32] B. Keng, "Spitzer." [Online]. Available: https://www.nasa.gov/mission_pages/spitzer/news/spitzer-20090506.html

[33] A. C. Aitken, "IV.—on least squares and linear combination of observations," *Proceedings of the Royal Society of Edinburgh*, vol. 55, pp. 42–48, 1936. [Online]. Available: https://doi.org/10.1017/s0370164600014346

[34] W. J. Borucki, E. Agol, F. Fressin, L. Kaltenegger, J. Rowe, H. Isaacson, D. Fischer, N. Batalha, J. J. Lissauer, G. W. Marcy, D. Fabrycky, J.-M. Desert, S. T. Bryson, T. Barclay, F. Bastien, A. Boss, E. Brugamyer, L. A. Buchhave, C. Burke, D. A. Caldwell, J. Carter, D. Charbonneau, J. R. Crepp, J. Christensen-Dalsgaard, J. L. Christiansen, D. Ciardi, W. D. Cochran, E. DeVore, L. Doyle, A. K. Dupree, M. Endl, M. E. Everett, E. B. Ford, J. Fortney, T. N. Gautier, J. C. Geary, A. Gould, M. Haas, C. Henze, A. W. Howard, S. B. Howell, D. Huber, J. M. Jenkins, H. Kjeldsen, R. Kolbl, J. Kolodziejczak, D. W. Latham, B. L. Lee, E. Lopez, F. Mullally, J. A. Orosz, A. Prsa, E. V. Quintana, R. Sanchis-Ojeda, D. Sasselov, S. Seader, A. Shporer, J. H. Steffen, M. Still, P. Tenenbaum, S. E. Thompson, G. Torres, J. D. Twicken, W. F. Welsh, and J. N. Winn, "Kepler-62: A five-planet system with planets of 1.4 and 1.6 earth radii in the habitable zone," *Science*, vol. 340, no. 6132, pp. 587–590, Apr. 2013. [Online]. Available: https://doi.org/10.1126/science.1234702

[35] A. Vanderburg and J. A. Johnson, "A technique for extracting highly precise photometry for the two-wheeledkeplermission," *Publications of the Astronomical Society of the Pacific*, vol. 126, no. 944, p. 948–958, Oct 2014. [Online]. Available: http://dx.doi.org/10.1086/678764

[36] D. J. Armstrong, J. Kirk, K. W. F. Lam, J. McCormac, S. R. Walker, D. J. A. Brown, H. P. Osborn, D. L. Pollacco, and J. Spake, "K2 variable catalogue: Variable stars and eclipsing binaries in k2 campaigns 1 and 0," *Astronomy & Astrophysics*, vol. 579, p. A19, Jun. 2015. [Online]. Available: https://doi.org/10.1051/0004-6361/201525889

[37] M. N. Lund, R. Handberg, G. R. Davies, W. J. Chaplin, and C. D. Jones, "K2p2—a photometry pipeline for the k2 mission," *The Astrophysical Journal*, vol. 806, no. 1, p. 30, Jun 2015. [Online]. Available: http://dx.doi.org/10.1088/0004-637X/806/1/30

[38] S. Roberts, M. Osborne, M. Ebden, S. Reece, N. Gibson, and S. Aigrain, "Gaussian processes for time-series modelling," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 371, no. 1984, p. 20110550, Feb. 2013. [Online]. Available: https://doi.org/10.1098/rsta.2011.0550

[39] H. Parviainen, "Bayesian methods for exoplanet science," in *Handbook of Exoplanets*. Springer International Publishing, 2018, pp. 1567–1590. [Online]. Available: https://doi.org/10.1007/978-3-319-55333-7_149

[40] N. P. Gibson, S. Aigrain, S. Roberts, T. M. Evans, M. Osborne, and F. Pont, "A gaussian process framework for modelling instrumental systematics: application to transmission spectroscopy," *Monthly Notices of the Royal Astronomical Society*, vol. 419, no. 3, pp. 2683–2694, Nov. 2011. [Online]. Available: https://doi.org/10.1111/j.1365-2966.2011.19915.x

[41] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[42] D. K. Duvenaud, "Automatic model construction with gaussian processes," Ph.D. dissertation, University of Cambridge, 6 2014.

[43] D. Foreman-Mackey, E. Agol, S. Ambikasaran, and R. Angus, "Fast and scalable gaussian process modeling with applications to astronomical time series," *The Astronomical Journal*, vol. 154, no. 6, p. 220, Nov 2017. [Online]. Available: http://dx.doi.org/10.3847/1538-3881/aa9332

[44] J. A. Carter and J. N. Winn, "PARAMETER ESTIMATION FROM TIME-SERIES DATA WITH CORRELATED ERRORS: A WAVELET-BASED METHOD

AND ITS APPLICATION TO TRANSIT LIGHT CURVES," *The Astrophysical Journal*, vol. 704, no. 1, pp. 51–67, sep 2009. [Online]. Available: https://doi.org/10.1088%2F0004-637x%2F704%2F1%2F51

[45] R. Luger, "Everest 1.0.5 source code," 2016. [Online]. Available: https://github.com/rodluger/everest/tree/1.0.5

[46] J. L. Christiansen, J. M. Jenkins, D. A. Caldwell, C. J. Burke, P. Tenenbaum, S. Seader, S. E. Thompson, T. S. Barclay, B. D. Clarke, J. Li, and et al., "The derivation, properties, and value of kepler's combined differential photometric precision," *Publications of the Astronomical Society of the Pacific*, vol. 124, no. 922, p. 1279–1287, Dec 2012. [Online]. Available: http://dx.doi.org/10.1086/668847

[47] R. L. Gilliland, W. J. Chaplin, E. W. Dunham, V. S. Argabright, W. J. Borucki, G. Basri, S. T. Bryson, D. L. Buzasi, D. A. Caldwell, Y. P. Elsworth, J. M. Jenkins, D. G. Koch, J. Kolodziejczak, A. Miglio, J. van Cleve, L. M. Walkowicz, and W. F. Welsh, "KEPLERMISSION STELLAR AND INSTRUMENT NOISE PROPERTIES," *The Astrophysical Journal Supplement Series*, vol. 197, no. 1, p. 6, Oct. 2011. [Online]. Available: https://doi.org/10.1088/0067-0049/197/1/6

[48] D. Wang, D. W. Hogg, D. Foreman-Mackey, and B. Schölkopf, "A causal, data-driven approach to modeling theKeplerData," *Publications of the Astronomical Society of the Pacific*, vol. 128, no. 967, p. 094503, Jun. 2016. [Online]. Available: https://doi.org/10.1088/1538-3873/128/967/094503

[49] "Chi square or likelihood." [Online]. Available: https://www-cdf.fnal.gov/physics/statistics/recommendations/modeling.html

[50] B. Keng, "A probabilistic interpretation of regularization." [Online]. Available: http://bjlkeng.github.io/posts/probabilistic-interpretation-of-regularization/

[51] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. The Johns Hopkins University Press, 1996.

[52] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, 2nd ed. Cambridge, USA: Cambridge University Press, 1992.

[53] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A limited memory algorithm for bound constrained optimization," *SIAM Journal on Scientific Computing*, vol. 16, no. 5, pp. 1190–1208, Sep. 1995. [Online]. Available: https://doi.org/10.1137/0916069

[54] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal, "Algorithm 778: L-BFGS-b: Fortran subroutines for large-scale bound-constrained optimization," *ACM Transactions on Mathematical Software*, vol. 23, no. 4, pp. 550–560, Dec. 1997. [Online]. Available: https://doi.org/10.1145/279232.279236

[55] C. Hellier, D. R. Anderson, A. C. Cameron, M. Gillon, E. Jehin, M. Lendl, P. F. L. Maxted, F. Pepe, D. Pollacco, D. Queloz, D. Ségransan, B. Smalley, A. M. S. Smith, J. Southworth, A. H. M. J. Triaud, S. Udry, and R. G. West, "WASP-43b: the closest-orbiting hot jupiter," *Astronomy & Astrophysics*, vol. 535, p. L7, Nov. 2011. [Online]. Available: https://doi.org/10.1051/0004-6361/201117081

[56] D. Pollacco, I. Skillen, A. Cameron, D. Christian, C. Hellier, J. Irwin, T. Lister, R. Street, R. West, D. Anderson, W. Clarkson, H. Deeg, B. Enoch, A. Evans, A. Fitzsimmons, C. Haswell, S. Hodgkin, K. Horne, S. Kane, F. Keenan, P. Maxted, A. Norton, J. Osborne, N. Parley, R. Ryans, B. Smalley, P. Wheatley, and D. Wilson, "The wasp project and the superwasp cameras," *Publications of the Astronomical Society of the Pacific*, vol. 118, no. 848, pp. 1407–1418, 2006. [Online]. Available: http://www.jstor.org/stable/10.1086/508556

[57] M. Gillon, A. H. M. J. Triaud, J. J. Fortney, B.-O. Demory, E. Jehin, M. Lendl, P. Magain, P. Kabath, D. Queloz, R. Alonso, D. R. Anderson, A. C. Cameron, A. Fumel, L. Hebb, C. Hellier, A. Lanotte, P. F. L. Maxted, N. Mowlavi, and B. Smalley, "The TRAPPIST survey of southern transiting planets," *Astronomy & Astrophysics*, vol. 542, p. A4, May 2012. [Online]. Available: https://doi.org/10.1051/0004-6361/201218817

[58] E. Mamajek, "A modern mean dwarf stellar color and effective temperature sequence," June 2016. [Online]. Available: http://www.pas.rochester.edu/~emamajek/EEM_dwarf_UBVIJHK_colors_Teff.txt

[59] C. Hellier, D. R. Anderson, A. C. Cameron, M. Gillon, L. Hebb, P. F. L. Maxted, D. Queloz, B. Smalley, A. H. M. J. Triaud, R. G. West, D. M. Wilson, S. J. Bentley, B. Enoch, K. Horne, J. Irwin, T. A. Lister, M. Mayor, N. Parley, F. Pepe, D. L. Pollacco, D. Segransan, S. Udry, and P. J. Wheatley, "An orbital period of 0.94 days for the

hot-jupiter planet WASP-18b," *Nature*, vol. 460, no. 7259, pp. 1098–1100, Aug. 2009. [Online]. Available: https://doi.org/10.1038/nature08245

[60] "Exoplanet catalog." [Online]. Available: http://exoplanet.eu/catalog/wasp-18_b/

[61] J. Southworth, T. C. Hinse, M. Dominik, M. Glitrup, U. G. Jorgensen, C. Liebig, M. Mathiasen, D. R. Anderson, V. Bozza, P. Browne, M. Burgdorf, S. C. Novati, S. Dreizler, F. Finet, K. Harpsoe, F. Hessman, M. Hundertmark, G. Maier, L. Mancini, P. F. L. Maxted, S. Rahvar, D. Ricci, G. Scarpetta, J. Skottfelt, C. Snodgrass, J. Surdej, and F. Zimmer, "Physical properties of the 0.94-day period transiting planetary system wasp-18," 2009.

[62] P. Goldreich and S. Soter, "Q in the solar system," *icarus*, vol. 5, pp. 375–389, 1966. [Online]. Available: https://ui.adsabs.harvard.edu/abs/1966Icar....5..375G

[63] A. Silva, "An expansion to the cheops mission official pipeline," Ph.D. dissertation, Faculdade de Ciências da Universidade do Porto, 9 2019.