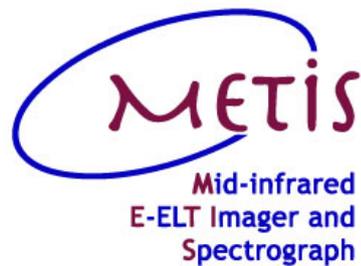




Mid-IR background calibrations for the E-ELT's METIS instrument



THESIS

submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE

in

ASTRONOMY AND INSTRUMENTATION

Author :	Alexander G.M. Pietrow
Student ID :	1054678
Supervisor :	Prof. Bernhard Brandl
2 nd corrector :	Dr. Matthew Kenworthy

Leiden, The Netherlands, July 22, 2016

Mid-IR background calibrations for the E-ELT's METIS instrument

Alexander G.M. Pietrow

Leiden Observatory
P.O. Box 9500, 2300 RA Leiden, The Netherlands

July 22, 2016

Abstract

METIS, one of the E-ELT's first instruments, will not offer classical chopping and nodding. If this is not solved then the residual background on reduced observations will be of the same order or larger than potential science targets. For this reason an investigation has been started to understand and quantify the nature and source of the residuals in order to infer whether or not it is possible to reduce the background and reach the shot noise limit. To answer this question we

have developed a specialised observing plan that allows us to look at the relation between the background residuals and various variables like the chop throw, chop frequency, chop direction, telescope altitude, filter and rotation angle. This plan has been executed on both the VLT and the GTC using the VISIR and CANARICAM IR instruments respectively.

From the analysis of this data we have discovered that contrary to what is believed in the literature the high order residuals are likely to be caused by the telescope and that they are stable over time but not over rotation angle, while the gradient is caused by the atmosphere and is not constant over time. Using this theory we have found a new way of reducing the background without nodding that, for chopping throws below 10", gives nearly identical results when compared to classical chopping and nodding. We believe that with this method we have found a solution for the challenge stated above but conclude that more rigorous follow-up observations are needed.

Contents

1	Introduction	1
1.1	The Mid IR Background	1
1.2	Classical Calibrations	7
1.3	Challenges	11
2	Data	15
2.1	Technical time	15
2.2	Observations	16
3	Methods	21
3.1	Classical Chopping	25
3.2	Chopping without nodding	25
3.2.1	Stability over time	26
3.2.2	Polynomial fitting	28
3.2.3	Principal Component Analysis	29
3.2.4	Multidirectional chopping	31
3.3	Atmosphere VS Telescope	32
3.3.1	Wavelength Dependencies	32
3.3.2	Derotator Test	35
3.3.3	Effects as a function of airmass	35
3.3.4	Closed dome	36
4	Results	37
4.1	Effects as a function of Throw	38
4.2	Residual stability over time	39
4.3	Alternative chop methods	40
4.4	Disentangling Backgrounds	45
4.5	The Derotator test	49

4.6	Effects as a function of airmass	50
4.7	Testing on a source	51
5	Conclusions and Discussion	55

Introduction

1.1 The Mid IR Background

The mid-infrared (MIR), usually defined as the region between 5 and 25 μm , is a relatively underutilized band in the EM spectrum. This is a region that is very important in certain fields of astronomy, especially those where dust plays a big role. Possible science topics that would gain from MIR observations could be the evolution of gas in protoplanetary disks, the study of star and planet evolution, accretion and outflows around protostars and the potential detection of biomarkers in exoplanet atmospheres. Besides this we can also study thermal structure, chemistry and dynamics in the thick atmospheres of our local gas giants and objects at moderate redshift. A more in depth look into these topics can be found in (Brandl et al., 2008; Otarola, 2014; Lacy et al., 2002).

Observations in this regime are done primarily by space telescopes. This is because this wavelength regime suffers heavily from our Earth's atmosphere, which radiates strongly in the IR wavelengths. Influenced by the vertical profiles of temperature, pressure and molecular abundances of, among other things H_2O , CO , CH_4 . The magnitude of all of these factors in turn depend strongly on the telescopes location and altitude above sea level. (Kendrew et al., 2010) The simplest way to understand this is to imagine the atmosphere as a 300K greybody with varying transparency and fully opaque regions. This effect is especially bad in the MIR due to the fact that the atmospheres blackbody peaks in this regime due to Wiens law. In addition to this, the telescope, its dome and its optics emit as a greybody of 15%* that also peaks at this wavelength resulting in a high background and a high photon shot noise (PSN). An estimate of the

*For the E-ELT, this can differ for other telescopes.

brightness of this background as a function of wavelength has been made by Kendrew et al. (2010), where the publicly available Reference Forward Model (RFM)[†] is used in combination with the HITRAN 2004 molecular line database (Rothman et al., 2004) to create an estimate of the sky brightness at two altitudes for a theoretical telescope. (2600m for Paranal and 5000m for an ideal, yet undetermined location.) The results of these estimates can be seen in Fig. 1.1 together with a 10% greybody estimate of the telescope at the local ambient temperature. This shows that the amount of atmosphere between you and the star and the temperature is a big factor, explaining why most people prefer space telescopes. We can see a similar plot done by Jeff Meisner that simulates the

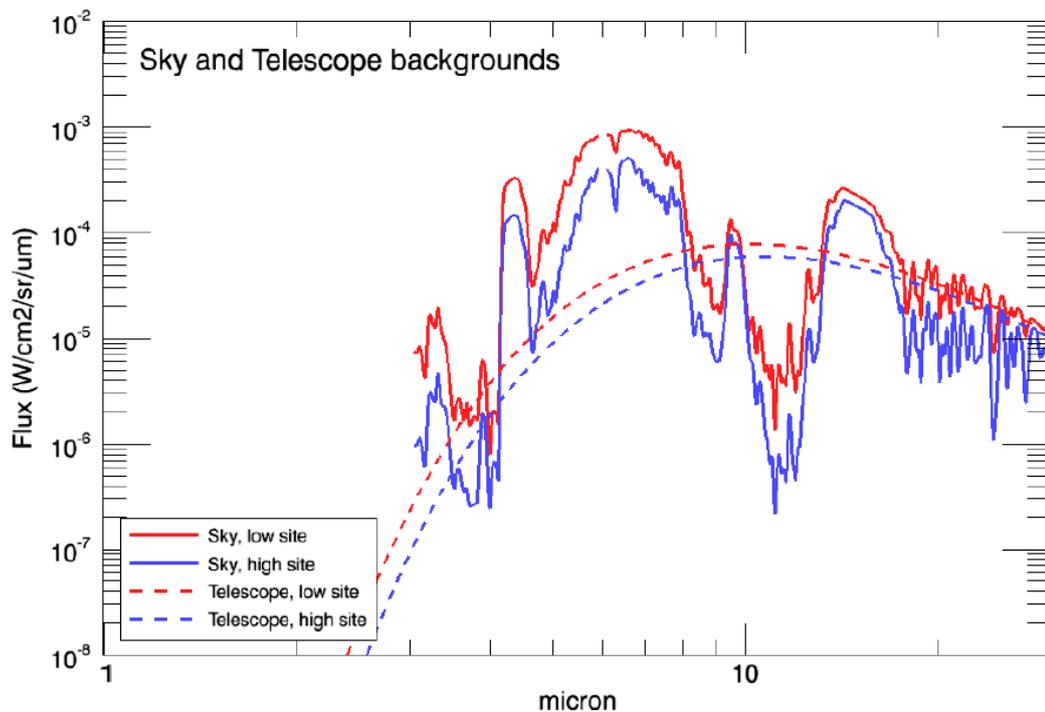


Figure 1.1: The atmosphere from about 2 to 10 micron (solid line) given by the RFM with a resolution of $R=3000$, plotted together with the telescope (dashed line) at 10% emissivity. In Red the conditions for Paranal and in blue a high and dry location. Imagecredit: (Rothman et al., 2004)

E-ELT and its location in Fig. 1.2. To form a better understanding of the background we have added a 0th magnitude star that is spread out over one arcsec² as a comparison. We see that the background is above this value in most of the

[†]<http://www.atm.ox.ac.uk/RFM/atm/>

MIR regime and that it goes up even further with longer wavelengths. If we would compare the MIR to optical bands, the obtained background would be similar to what we see when observing during the day while using a telescope with luminous optics, so not the most ideal setup.

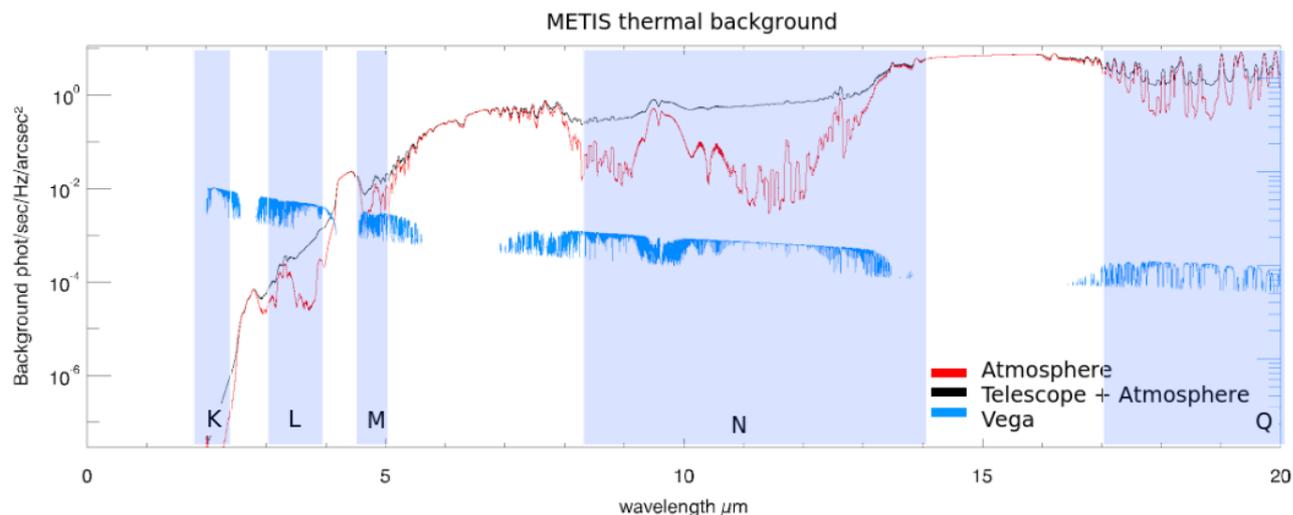


Figure 1.2: The two main background components of the MIR background are given by the atmosphere (Red line), with a varying transmission over the wavelength range and the telescope (The black line is the sum of both) that has been modeled as a gray body of 15%. The combination of the two backgrounds is enough to outshine some of the brightest stars, in this case compared to a 1 arcsecond, 0th magnitude star. In the blue regions one can see the five main observing bands that are used in the MIR. Image was obtained from Jeff Meisner

A bright sky by itself would be fine, if it would be stable. Unfortunately this is not the case in the MIR and we get to deal with varying sky brightness changes with very short (second) time scales over spatially small regions, often much larger than \sqrt{N} , as would be expected from poisson statistics. (Geballe and Mason, 2006) In Fig. 1.4 we can see the three main components of MIR radiation, namely; the telescope, atmosphere and detector. The first two will be looked into now, while the detector will be covered further down.

The biggest effects in the N band are caused by CO_2 , CH_4 , H_2O and O_3 . This means that an increase in water vapour column (WVC) will cause a reduction in transmission and an increase in sky background, as is shown for 1.0 and 3.0 mm H_2O in Fig. 1.5. This means that cold, high and dry places with low emissivity and high cleanliness are the best for MIR observations. Besides this, it is possible to reach the diffraction limit in this regime, but only with good conditions

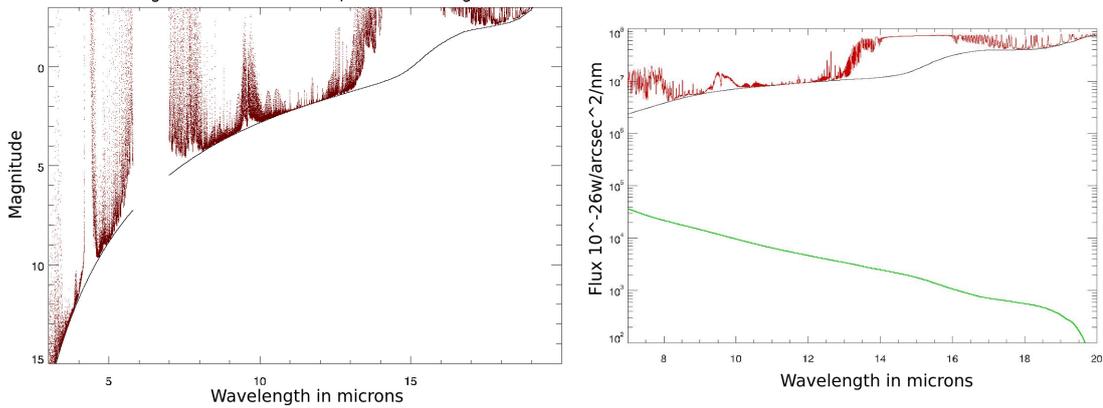


Figure 1.3: **Left:** A diagram showing the brightness of the background in the MIR (red) and the brightness of a diffraction limited star that has the same surface brightness (black). **Right:** A diagram showing the relative difference in brightness between the MIR background (red) and a diffraction limited 10th magnitude star. (green) Both images were obtained from Jeff Meisner.

and guiding.[‡] With the main contributors to background fluctuations being the WVC, unstable weather, thin cloud, wind-borne dust and sometimes bugs and birds.

To get a better feel for the numbers we do some basic S/N calculations to see how bright the background is. We can determine this brightness quite easily by following the method as described at the Gemini website (Geballe and Mason, 2006), but with the numbers for the VLT.

So if we start with an 8.2m class telescope, with a temperature of 300K and an emissivity of 9% (Rupprecht, 2005), we will get a background of 431 Jy/arcsec² at 8.7 micron from the telescope and its optics. We can add about 23% (Meisner, 2016) more to this to account for the atmosphere ending up with 530 Jy/arcsec² of background flux. If we compare this to a 0th magnitude star, which shines at about 46.6 Jy (Cohen et al., 1999) at 8.7 micron in the N band[§] and we focus the star on a 1 arcsecond² region, then we will see that $\frac{530}{35.21+530} = 94\%$ of the light in that region is background, comparable to Fig. 1.2. Of course this number goes down when the star is focused on a smaller region, but we generally also look at much fainter sources. The same star focussed on a diffraction limited 0.3 arcsec² spot would give us a 57% background for example. This is clearly a problem,

[‡]VISIR is diffraction limited when optical seeing is below 0.6" (Tristram, 2016)

[§]This is based on Alpha boo being 633 Jy at -3.14 mag.

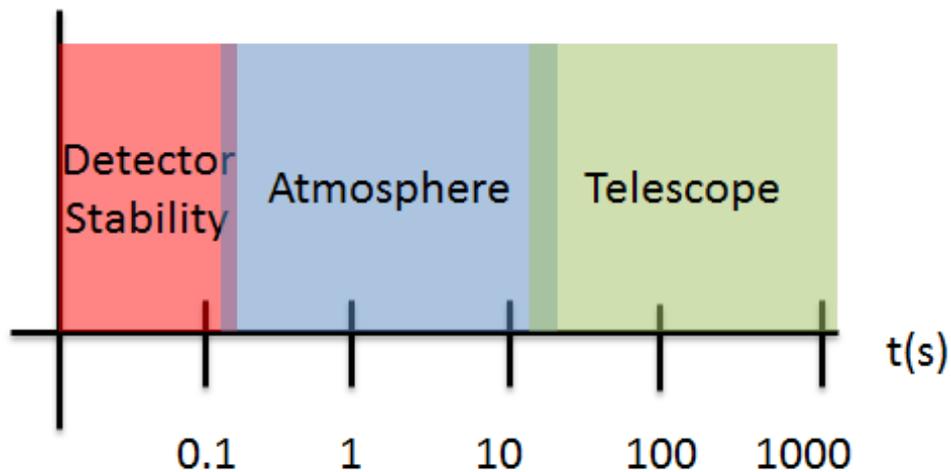


Figure 1.4: The three main components that emit MIR radiation are the atmosphere, the telescope and the detector itself. Looking at the time-scales we can see that the telescope is the most stable with stabilities in the order of hours, followed by the atmosphere in minute time scales and the detector which is stable for subsecond timescales.

as we need a much higher S/N to be able to do any form of science. We can subtract out most of the background, but we eventually we will hit the ultimate limit. This limit is the shot noise that scales with the square root of the counts on the detector. This noise level cannot be surpassed and limits the S/N to our observations. Fortunately there is a lot to gain between the raw background and the shot noise limit, so getting to this point is already a great success.

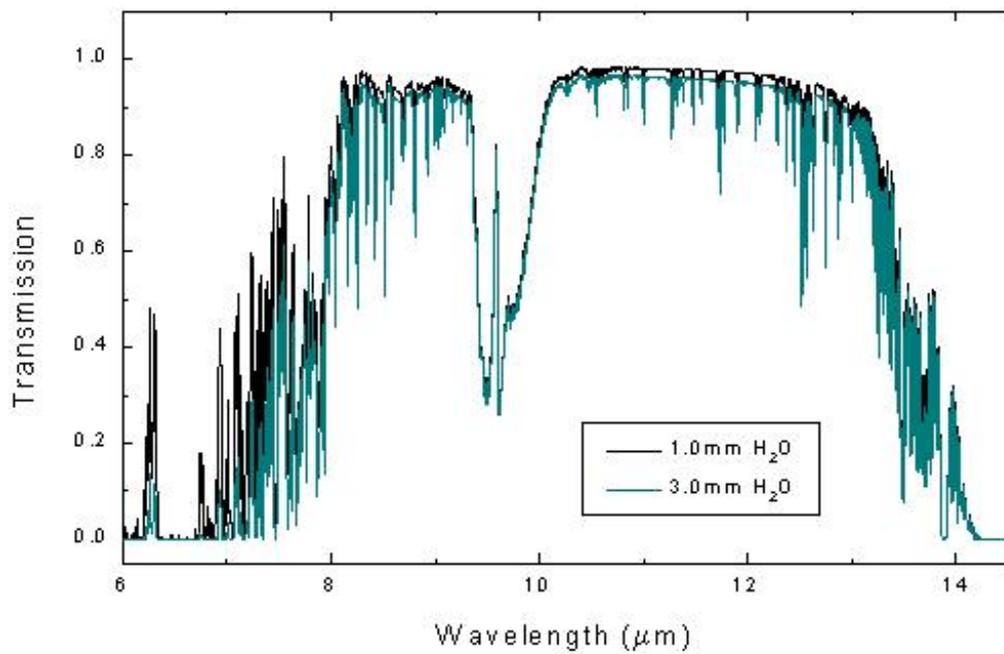


Figure 1.5: The background between 6 and 14 μm showing the difference in background transmission as a function of the vapor column. Lord et al. (1992)

1.2 Classical Calibrations

The classical solution to removing the large MIR background is a combination of two methods, namely chopping and nodding. **Chopping** is the act of moving the secondary mirror of your telescope by a little bit, shifting the image on the detector, creating an offset in the pointing by a few arcseconds from the detectors point of view. This operation is repeated at a frequency of a few Hz with exposures being taken every few ms. This frame rate has to be high enough to prevent over saturation while being as low as possible to avoid excessive readout noise. Typically we use frame rates that read out around every 20 milliseconds.

After averaging these frames and subtracting the two pointings, we are left with a so called '*chop difference frame*' from which all constant components have been removed. Ideally this would only leave the source and remove the entire background, but because we moved the telescope beam across the primary mirror, illuminating a slightly different part of the telescope, we will see slightly different parts of the telescope. Because the telescope emits strongly in this wavelength range, these differences result in residuals and we are left with something in the order of a few percent of the original background. Unfortunately, if we remember the numbers calculated above, this is close to, and usually much higher than, the brightness of the source and therefore not enough to get a proper S/N. The classical solution to these residuals is to move or **nod** the telescope by the same offset and in the same direction as was used for chopping and chop in the same way as before. (So that the science target and empty space switch places.) When repeated about once or twice a minute, this process produces two pairs of images that both contain the same residuals but have inverted signs for the source. We get one chop difference frame for each nod position and when we subtract those we will get a '*nod difference frame*'. This gets rid of most aberrations and gets the noise down to the shot noise limit. This process is illustrated in Fig. 1.7.

In general this method works relatively well, removing almost all of the background. However it is very important to have a stable sky in the region that is being observed, otherwise the difference frames will not cancel each other out properly leaving large residuals on the final image. (Geballe and Mason, 2006) Looking at Fig. 1.6, we can see the four aforementioned pointings, of which two overlap. In Fig. 1.8, we can see how such an image would look on the chip. These nod difference frames have three copies of the source on the image. (If the chopping throw was shorter than the image size, otherwise we would see only

one image.) The central image is positive and has counts for half the exposure time and the other two images are negative and correspond to a quarter of the time each.

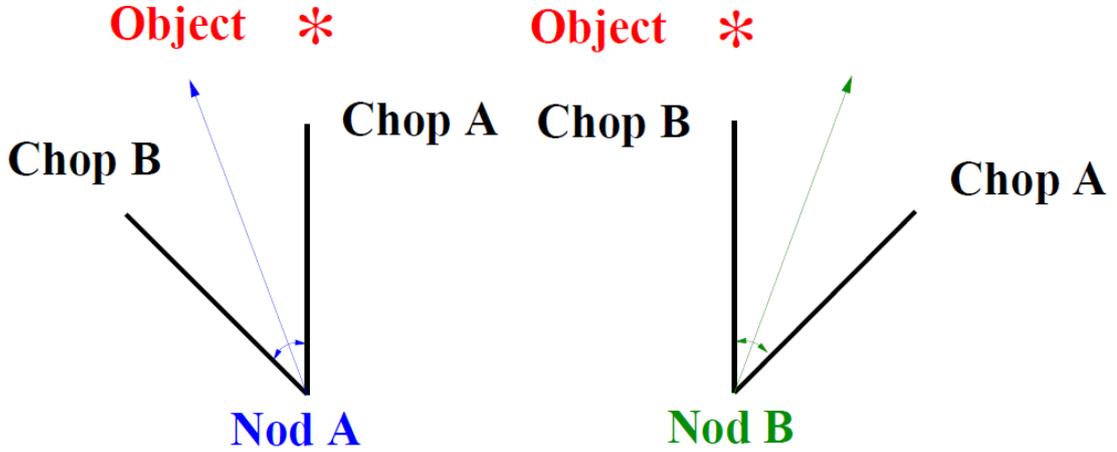


Figure 1.6: A diagram showing the four pointings that are used when chopping classically. In nod position A the telescope is aimed at the source and then offset by chopping away from it. In nod position B, the telescope is moved as a whole to the position that it was chopping to, switching the source with the offset position. Now it chops again. (Volk, 2007)

We can express the above story qualitatively by using the terminology as described in Volk (2007). Using Fig. 1.6 as our reference image we can make a telescope exposure starting with nod position A and define the Flux, F_{nod}^{chop} , background B and telescope background T. With this we can write the flux for Nod A and Chop A as,

$$F_A^A = B_1 + T_A + S. \quad (1.1)$$

Similarly we can express the flux for Nod A, Chop B as,

$$F_A^B = B_2 + T_B. \quad (1.2)$$

These two frames are identical apart from the fact that chop position B has no source in it and that T_B differs from T_A because the beam has moved over the primary mirror. However, this effect is small with $T_B \approx T_A$ and the backgrounds are of similar magnitude too. Next we can look at the flux for Chop A and Chop B for Nod B, which are given by,

$$F_B^A = B_3 + T_A \quad (1.3)$$

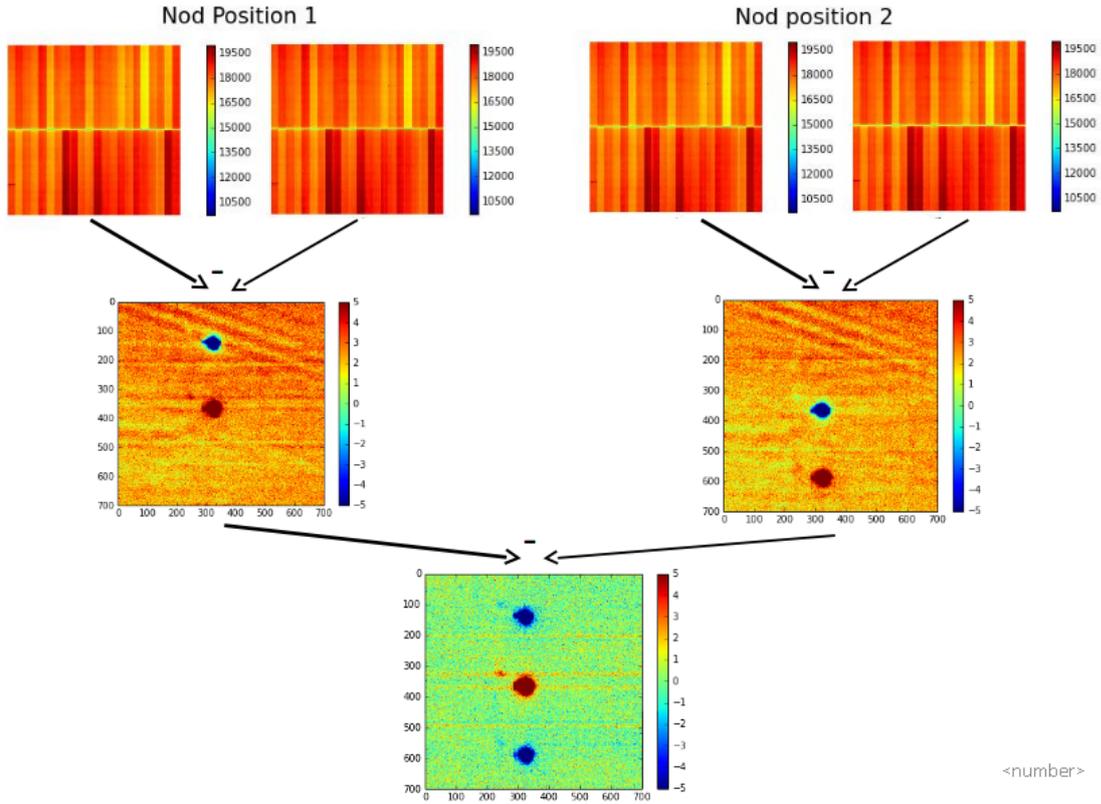


Figure 1.7: A classical chop/nod reduction shown step by step. We start with two pairs of chop exposures which we can subtract in order to remove most of the background. The remaining two images have nearly the same aberrations but have inverted symbols for the source. This means that subtracting them removes the background and doubles the source.

and,

$$F_B^B = B_4 + T_B + S. \quad (1.4)$$

We see that these two equations are nearly identical to the two above, with only the sky backgrounds being different and the source being on the other chop. Subtracting the two chop positions for both nods will leave us with two chop difference frames, one for each nod position. We will call these frames 'nA' and 'nB'.

$$\begin{aligned} F_A^A - F_A^B &= B_1 + T_A + S - B_2 - T_B \\ &= \Delta B_{12} + \Delta T_{AB} + S \\ &= nA \end{aligned} \quad (1.5)$$

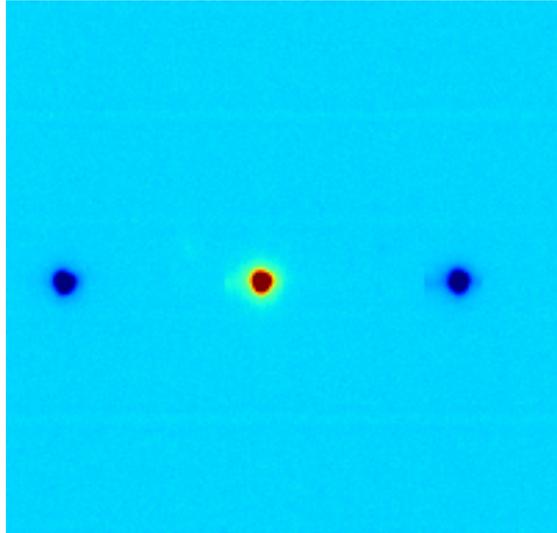


Figure 1.8: The resulting image obtained after executing the plan as described in Fig. 1.6 on observations of the standard star HR 2652.

and,

$$\begin{aligned}
 F_B^A - F_B^B &= B_3 + T_A - B_4 - T_B - S \\
 &= \Delta B_{34} + \Delta T_{AB} - S \\
 &= nB.
 \end{aligned}
 \tag{1.6}$$

Now if the background and telescope terms would have been small enough, then we would have just the source left, but we have discussed above that this is not the case, as can be seen in Fig. 1.7. These terms can still be of the order of magnitude or brighter than the source and we need to find a way to rid ourselves of them. Therefore we want to subtract the two chop difference frames to get

$$\begin{aligned}
 nA - nB &= \Delta B_{12} + \Delta T_{AB} + S - (\Delta B_{34} + \Delta T'_{AB} - S) \\
 &= \Delta B_{12} - \Delta B_{34} + 2S \\
 &\approx 2S.
 \end{aligned}
 \tag{1.7}$$

Here we see that the telescope background is gone and if taken with a short enough time in between exposures, we can assume that the background hardly changes between exposures and thus can say that the two background terms cancel out, leaving us with a flat background and our source.

The noise of this image is composed of shotnoise from the photon flux itself, 1/f noise from background variations, readnoise and the detector. 1/f noise is quite

hard to model, but if we assume that chopping and nodding takes care of it, this allows us to reach background limited performance. We can estimate a value for this noise if we take the example value of 530 Jy/arcsec^2 for the background that was given above and apply the flux identity described in Oke (1974) This will give us a flux of $8.1 \cdot 10^{10} \text{ photons/m}^2/\text{s}/\mu\text{m}/\text{arcsec}^2$ at $8.7 \mu\text{m}$. If we then assume a $0.7 \mu\text{m}$ bandpass, $0.045'' \times 0.045''$ pixels, a throughput of 0.7 and a Quantum efficiency of 70% , we will be left with $8.1 \cdot 10^8 \text{ electrons/s/pixel}$. For a single chop frame the pipeline averages 7 exposures of 87.5ms in total, this gives us an average amount of $10.1 \cdot 10^6 \text{ electrons/pixel}$, giving us a shot noise of $\approx 3.2 \cdot 10^3 \text{ electrons/pixel}$. This means that our image is shot noise limited if the other noise components are below that. Which is true for the $150 \text{ electrons/pixel}$ RMS noise introduced by the detector (Ives et al., 2012), and we assume it to be above the $1/f$ noise introduced by the atmosphere if the chopping takes care of it. So if we know that we are shot noise limited and if we know the counts of all the individual flux components and assume that they are all Poisson distributed, then we can express the noise limit as the square root of the sum of these components. This is true for all background components but not for the source, however its contribution to the counts is very small compared to the other components (like the read noise), so we can neglect it and write the noise for a single frame as.

$$\sigma \approx \sqrt{B_1 + B_2 + B_3 + B_4 + 2T_A + 2T_B} = \sqrt{N}. \quad (1.8)$$

With N the amount of counts in electrons on our images.

1.3 Challenges

METIS, the ‘Mid-infrared ELT Imager and Spectrograph’, is the dedicated MIR instrument, working in between 3 and $19 \mu\text{m}$, that has been selected by ESO as a first generation E-ELT instrument. This instrument will be at the forefront of MIR astronomy, rivalled only by the Spitzer Space Telescope and the upcoming James Webb Space Telescope. However it does have a couple of advantages over space telescopes, the main ones being the higher angular resolution due to the larger aperture, a much higher spectral resolution than is currently possible due to weight limitations for space telescopes and the fact that the instrument is easy to reach and maintain during its operation. (Brandl et al., 2008)

The challenges that come with this telescope are mostly because of the sheer size of the 39m primary and the mirrors that come after it. The secondary mirror is too large (4.2m) and heavy to use as chopping mirror and because of that

it has been decided to make an internal chopper in the pupil plane of the instrument.(Paalvast et al., 2014) METIS would be located at the nasmyth focus (see Fig. 1.9) of the E-ELT and have 6 warm and several cold mirrors before the chopper.[¶] (One of which would be the AO deformable mirror) This will introduce a larger than usual thermal footprint from the telescope, which will be modelled as a 15% greybody compared to 9% in the VLT.

Now that we know the challenges that come with observing in the MIR and the solutions that are used, one might wonder why we bother with this at all. The answer to this question is simple, the classical solutions will not work for the European Extremely Large telescope (E-ELT).(Pantin, 2015) Due to the fact that the E-ELT will be a dynamical structure that continuously reajusts to keep up the alignment and image quality. This makes it impossible to nod properly, as the symmetry cannot be kept, meaning that using classical methods we will not be able to get down further than a few percent of the original background, which is still too high for observing. For this reason we need to find alternative ways of calibrating the images.

Unfortunately we do not yet properly understand the origin and magnitude of the observed systematic effects, which makes it very difficult to construct reliable thermal models of the E-ELT. The currently agreed upon explanation blames the beam motion on the primary mirror for the residuals that can be seen in chop difference frames, which is odd as a surface in the pupil plane should not create such structures. Theories exist about the cryostat window being the cause of the abberations (Brandl, 2015), but this also has never been quantified.

Because of this we want to gain a better quantitative understanding of the thermal background and its uniformity as a function of various variables and calibration techniques. On top of this we want to develop, implement and analyse new background reduction strategies that do not require nodding.

To do this we will request technical time at two large telescopes capable of observing in the MIR, the VLT and the Grantecan (GTC). These observations will be specially developed to test various aspects of the background, as can be read in Chapter 2. The obtained data will be subjected to various tests and reduction methods that are described in Chapter 3 and their outcome in Chapter 4. Finally conclusions will be drawn in Chapter 5.

[¶]Classically there would only be one, namely M1.

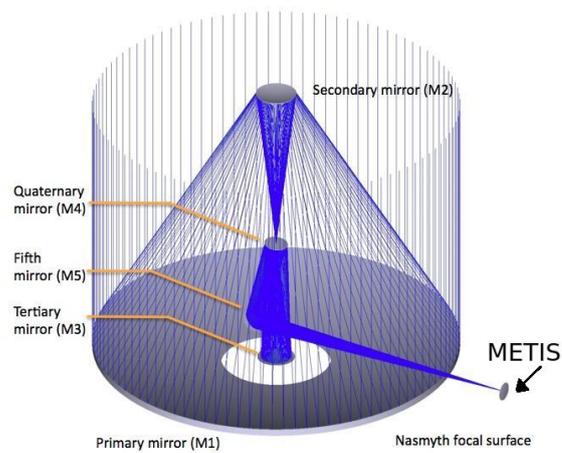


Figure 1.9: A raytrace of the E-ELT with a beam of light going to the Nasmyth platform. The METIS chopper will be inside of the instrument, meaning that the light needs to reflect off 6 warm mirrors before hitting the chopper. This is significantly more than in classical setups.

Data

2.1 Technical time

To better understand the origin of the various contributions of the thermal background we have requested observing time on both the VLT and the GTC. With the former we would want to observe while using the VISIR instrument, which is located in UT3 (Melipal) and similarly for the GTC we want to observe with their IR instrument, the CANARICAM. The aim of the proposal is to obtain a coherent dataset, tailored especially for this purpose, to better understand the thermal background and its uniformity under different observing schemes.

For these observations we propose to take a set of 'cheap' exposures that should take no more than two hours in total and can be done during twilight and does not require good seeing conditions. This would make our tests non invasive in the existing schedule and should be doable in the time frame given for a master thesis.

While it is desirable to have an IR point source in the images, we also wish to have a stable telescope with no movement other than that of the chopper. Tracking and derotation should be turned off, which makes it hard to look at a source. For this reason we have chosen to look at empty sky with a given Alt and Az=0.

We know that most of these observations can be found in the archive, but we think that it is essential to have a coherent data set with the same boundary conditions and instrument settings, otherwise we will have too many variables and never be able to pinpoint the exact reason. This dataset taken on both telescopes should be a good first step in identifying the origins of the thermal background. Throughout the observing procedure we have tried to keep doing the same set

Chop Amplitude	2,5,10,15,30 arcseconds
Chop Frequency	0.5,1,2 Hz
Chop Direction	N-S, E-W
Filters	J8.9, J7.9, SIV
Altitude	25, 30 ,45, 90 degrees ALT
Derotator angle	0, 90 degrees

Table 2.1: All variables that we want to vary during our tests with the values that we wish to vary.

of observations and control what we vary as strictly as possible. For this reason we have chosen to use a certain nodding pattern that shall henceforth be referred to as the 'nodding cross'. This is a set of eight science observations as displayed in Fig. 2.1 that is designed to allow us to make different nodding and chopping patterns in different directions, allowing us to get the maximum amount of information from a minimal amount of observations. The cross is defined as a collection of standard chop/nod operations where we start chopping at pointing A in direction B, then nod to B and continue nodding in the same direction. Then we repeat the block but nod to position C in the direction of C, etc. Most of the observations in this observing run will follow this 'nodding cross' template. During these tests we will keep the directions the same but vary chopping throw and frequency.

Furthermore we want to add some observations in other filters and with different rotation positions on the derotator to see if the residuals respond to that.

2.2 Observations

The technical time observations for both the VLT as the GRANTECAN have resulted in multiple datasets with different variations. The original observing plan proposed to make cross shaped nodding patterns with three different chop throws (5", 10", 30") and two different frequencies (1Hz, 4Hz). All of these tests were done while using the J8.9 filter for the VLT and the Si2-8.7 filter for the GTC, which both are similar and have relatively broad filter on the lower end of the N-band. These filters were chosen because they were the closest filters that have significant contributions from both the telescope and the atmosphere. These tests were taken during twilight with reasonable conditions and were also partially repeated with the dome closed, as a comparison without atmosphere.

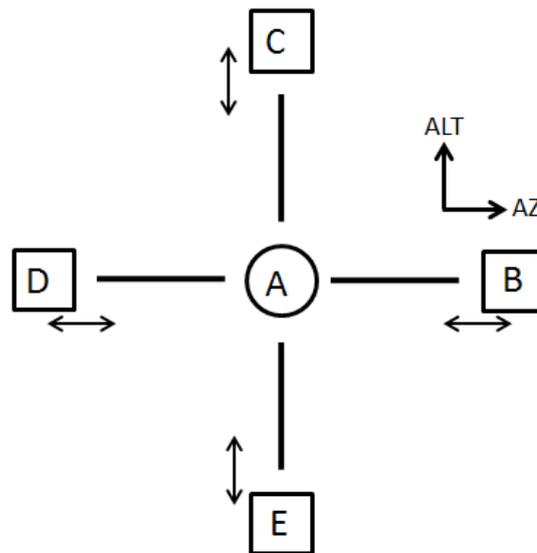


Figure 2.1: The so called 'nodding cross', showing how the nods will be set. The chopping is always parallel to the nodding direction. The cross starts with chopping from A to B, then nodding to B and chopping in the same direction. We call this the AB arm. After this we nod back to the A position and begin chopping in the AC direction. In this way we work through all four arms.

Our field is 38.0×38.0 arcsec², but will be windowed down to a region of interest of the central 31.5×31.5 arcsec². In all cases we will chop parallel to the nodding direction, have no random jitter, use only 1 nodding cycle and have a 60s total integration time in high gain mode.* The only things that will change are the chopping position angle, the frequency and the chopping amplitude. In addition we have turned off the telescope tracking and observe an empty patch of sky during most observations.

If everything would have gone well we would have four sets of data, two from each telescope. This would be the entire observing plan taken with an open and closed dome. However it is not as straightforward to obtain and retain observing time as it is with science observations. For both the VLT and the GTC the observing date has been pushed back many times due to a myriad of reasons including weather and priority going to more important technical tests. For the VLT we eventually set an observing date at the end of January where three hours were allotted to this project, however when the observing began we got the message that the time was cut down to 1 hour. For this reason we had to make on the spot decisions on what tests to do and what tests not to do. In the

*This means $20 \text{ e}^- / \text{ADU}$

end we split the time into two sets, one with a low (88.9 degrees ALT) airmass and one with high (25 degrees ALT) airmass. The first set consisted of three nodding crosses with a throw of 5,10,30 arcsec at 4Hz, one nodding cross with a 30 arcsec throw at 1Hz and a filter test. The latter consisted of two nodding crosses with a throw of 5 and 30 arcsec at 4Hz and one cross with a 30 arcsec throw at 1Hz. We also did a filter and derotator test at this ALT. Unfortunately the exposure times of the observations were not altered, so all the high airmass observations have passed their maximum saturation level and are therefore not usable. Because the derotator test was an important test for us, we have requested for it to be done once more and it has been. We have received the data plus a half nodding cross on the 16th of June. (Taken with the same settings and position.) We also got a nodding cross taken at two airmasses of 30 and 45 degrees ALT with a random pointing and rotation. These images are not ideal, but better than nothing. Finally we also got a nodding cross centred on a source, namely the star HR 2652. This was done with tracking but at a high altitude.

With the Grantecan we had even less luck. We got a closed dome dataset with three crosses with throws of 5,10,30 arcsec at 3.6 Hz[†]. We got the same data for 1Hz, a derotator test and a time stability test. The rest of the observations were scheduled to be taken later that month, but due to faults in the instrumentation the CANARICAM needed to be removed from the telescope for at least a year, making further observations impossible within the time frame of this project.

A summary of all of these observations can be found in table 2.2.

[†]The GCT choppers maximum frequency.

Summary of observations made for the METIS calibration

Date	Telescope	Type	Amplitude ["]	Freq [Hz]	Filter	t_{int}	ALT [Deg]	Burst
3-11-'15	VLT [C]	Cross	8.0	3.0	NeII	0.0278	88.9	N
28-12-'15	VLT [O]	Cross	5.0	4.0	J8.9	0.0125	45	Y
29-12-'15	VLT [O]	Cross	5.0	4.0	J8.9	0.0125	30	Y
23-01-'16	GTC [C]	Cross	5.0	3.6	Si2-8.7	0.07	80.0	N
23-01-'16	GTC [C]	Cross	10.0	3.6	Si2-8.7	0.07	80.0	N
23-01-'16	GTC [C]	Cross	30.0	3.6	Si2-8.7	0.07	80.0	N
23-01-'16	GTC [C]	Cross	5.0	1.0	Si2-8.7	0.07	80.0	N
23-01-'16	GTC [C]	Cross	10.0	1.0	Si2-8.7	0.07	80.0	N
23-01-'16	GTC [C]	Cross	30.0	1.0	Si2-8.7	0.07	80.0	N
23-01-'16	GTC [C]	Cross	5.0	3.6	Si2-8.7	0.07	80.0	N
23-01-'16	GTC [C]	Derot	30	3.6	Si2-8.7	0.07	80.0	N
22-03-'16	VLT [O]	Cross	5.0	4.0	J8.9	0.0125	88.9	Y
22-03-'16	VLT [O]	Cross	10.0	4.0	J8.9	0.0125	88.9	Y
22-03-'16	VLT [O]	Cross	30.0	4.0	J8.9	0.0125	88.9	Y
22-03-'16	VLT [O]	Cross	30.0	1.0	J8.9	0.0125	88.9	Y
22-03-'16	VLT [O]	Derot	30.0	4.0	J8.9	0.0125	25	Y
22-03-'16	VLT [O]	Filter	30.0	4.0	J7.9, SIV	0.0125	25	Y
22-03-'16	VLT [O]	Cross	5.0	4.0	J8.9	0.0125	25	Y
22-03-'16	VLT [O]	Cross	30.0	4.0	J8.9	0.0125	25	Y
22-03-'16	VLT [O]	Filter	30.0	4.0	J7.9, SIV	0.0125	88.9	Y
24-03-'16	VLT [O]	Star	10.0	4.0	J8.9	0.0125		Y
16-06-'16	VLT [O]	Derot	10.0	4.0	J8.9	0.0125	88.9	Y
16-06-'16	VLT [O]	Time	10.0	4.0	J8.9	0.0125	88.9	Y

Table 2.2: A table showing all obtained observations for both the VLT and the GTC. All files are grouped together into relevant blocks to avoid clutter, these blocks can be seen in the 'Type' column. With 'Cross' being the standard nodding cross, 'Derot' the derotator test, 'Filter' the filter test, Time a set of identical chopping patterns with set times in between and 'Star' being a standard chopping cross executed on a source. In the 'telescope' column, we can see the telescope that has been used and whether the dome is [O]pen or [C]losed. Observations with striked out dates are oversaturated and not usable.

Methods

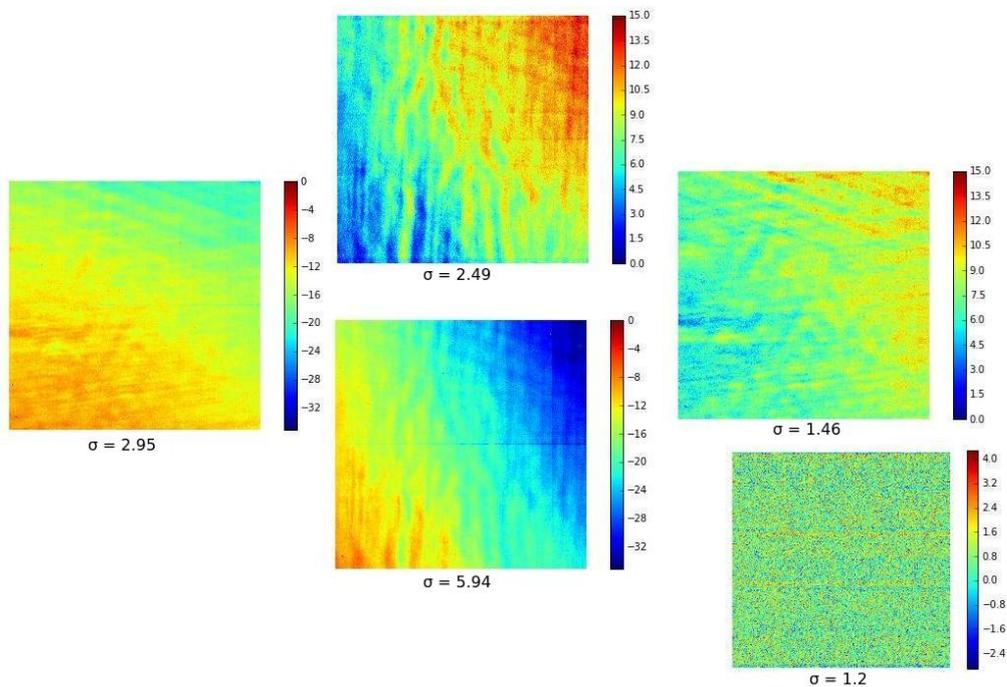


Figure 3.1: In this illustration we see four chop difference frames corresponding to the nodding cross of Fig. 2.1. The background left after the chop subtraction is still significant compared to average sources and very high compared to a nod subtracted frame. Such a frame can be seen in the lower right corner of the image.

Since we will not be able to nod with the E-ELT, we will have to find ways to deal with the chopped difference residuals that are left after a chop subtraction. These residuals are often in the order of a few percent of the original back-

ground, but since the source is also of this magnitude and often less, we cannot just ignore it. In Fig. 3.1 we can see the residuals of the chopped difference frames for the four directions of the nodding cross. Here it is clear that the background residuals are far above the shot noise limit and that these images cannot be used without further reduction due to the gradients and small scale structure inside them. For comparison we have added a nod difference frame in the lower right corner, showing what we are lacking without proper nodding and what we want to achieve with our alternative methods.

In this chapter we introduce different background reduction techniques and compare them to classical chopping and nodding. We begin by studying the residuals as a function of four variables. Namely the chop throw, direction, frequency and telescope elevation. Afterwards we test the residual stability over time and then we choose a subset of these variables and explore new background reduction methods.

After obtaining this data we have to define methods to quantify and display background quality of the reduced frames after different background reduction techniques. Primarily we want to see if the background is flat and limited by shot noise. This means that we want the mean to be zero, without any gradients or higher order structures and that the noise should have a standard deviation of around a theoretical minimum of

$$\sigma = \sqrt{\text{electrons}} \quad (3.1)$$

But since we have our counts in ADU, we should convert them by using the conversion factor found in the image header. This value is given to be $20 \text{ e}^- / \text{ADU}$.

$$\sigma = \frac{\sqrt{2}\sqrt{20 \cdot \text{ADU}}}{20\sqrt{n}} \quad (3.2)$$

Because we subtract two frames the noise rises with $\sqrt{2}$ and because we average over n chop pairs, the noise goes down with that number. Filling in the ADUs from the images and the amount of frames we get $\sigma = 2.16 \text{ ADU}$.

However when comparing these numbers to our calculations from before, we find that there we would expect to have $1262 \text{ e}^- / \text{ADU}$ rather than 20. Repeating the calculation with this number gives us $\sigma = 0.27 \text{ ADU}$. Neither of these numbers seem to be correct, as our chop and nod subtracted frames have a STD that is lower than 2.16 and the e^- / ADU rate from our calculations is too high. Therefore we assume that the images are shotnoise limited and that the shot-noise limit is 1.1 ADU.

1 μ_1 σ_1	2 μ_2 σ_2
3 μ_3 σ_3	4 μ_4 σ_4

Figure 3.2: With the Quadrant test we test the flatness of the reduced images. To do this we cut them into four smaller squares and calculate the mean and STD of every quadrant. The variation in these values tells us about the flatness on both large and small spatial scales.

We attempt to quantify the gradient and higher order structures by cutting the images into four quadrants and calculating the mean and standard deviation of every quadrant. (See Fig. 3.2.) Large scale variations (like a gradient) will be shown by a large spread in the four means while small scale structures will be described by a larger standard deviation. To illustrate the method we prepared two test images, one with a large overall gradient and one with smaller higher order structures. Both images can be seen in Fig. 3.3 and in Table 3.1 as the gradient test and the structure test.

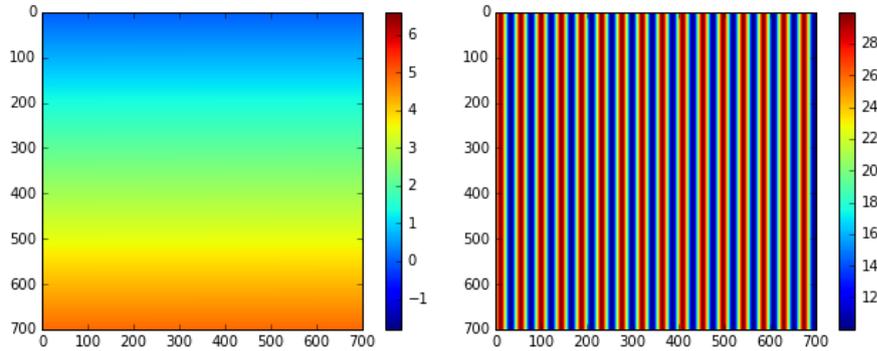


Figure 3.3: Two test images designed to illustrate the working of the quadrant test. The first image shows a gradient moving smoothly from top to bottom, while the second image shows a high order repeating pattern. The gradient will be picked up by the mean while the STD is sensitive to small scale variations.

Data	$\overline{\text{counts}}$	μ_1	μ_2	μ_3	μ_4	$\overline{\mu} \pm \sigma_\mu$	$\overline{\sigma}$
Gradient Test	2.7	1.36	1.36	4.08	4.08	2.72 ± 1.36	0.79
Structure Test	20.02	20.01	20.01	20.01	20.02	20.02 ± 0.007	7.08

Table 3.1: Results of the Quadrant test for the two test datasets that are illustrated in Fig. 3.3. This shows the strength and weakness of looking at the mean and STD of the quadrants and how using both can help as a first look. The first column shows the average amount of counts in ADU of the image. The next four columns show the mean value of the four quadrants, followed by the mean of these four means \pm the spread on the values. The final value is the mean STD of the four quadrants.

For the gradient test we can see that the means are different for the top two and bottom two quadrants, suggesting a vertical gradient in the data. Because the gradient goes up very smoothly the standard deviation will be small and not vary across the four quadrants. In the structure test we can see that the opposite is true, the means are very close to each other, suggesting a flat image, while the standard deviation is very big. This test is not perfect, but it can indicate complete outliers and methods that won't work at all. Once a sufficiently flat image has been produced we can compare the standard deviation of the entire image with the theoretical value that we derived above. Besides this it is also important to do a visual inspection of the resulting image and to repeat the process on a binned down image.

3.1 Classical Chopping

To get a good feeling for the numbers and a baseline as reference, we want to calculate the means and STDs for a large array of classical nodding patterns. Trying to extract the maximum amount of data from our observations we compare classical nod patterns with three different chop throws, four directions, two frequencies and two airmasses. With this we hope to find a pattern in the brightness of the background residuals and with that an optimal chopping strategy and reduction method. We have taken the nod difference frames of observations with one of these variables being varied every time and carried out the quadrant test on these frames. The results can be seen in Table 3.2.

3.2 Chopping without nodding

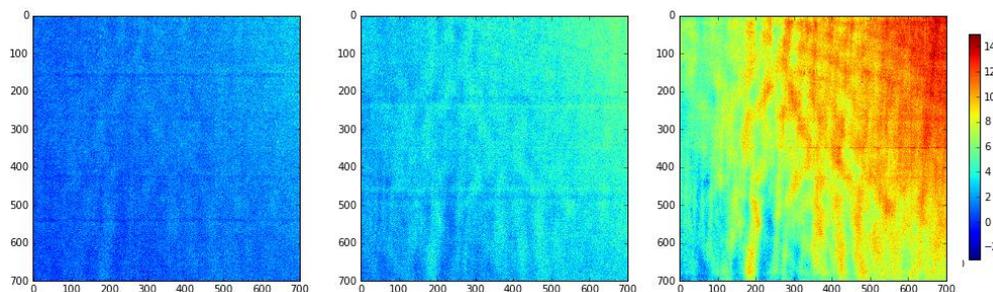


Figure 3.4: Chop difference frames for the AC nodding direction for 5'', 10'' and 30'' chop throws with the same scaling.

Since we cannot nod and have to suppress the background by testing new reduction strategies, we should first observe the individual difference frames to understand them better. In Fig. 3.1 we can see the chop difference frames for all of the four nodding directions. It is clear that the pattern, both high and low order, is different in every direction, but similarities can be seen in opposite facing chops. This raises the question of whether or not we can remove the background by using these symmetries. When looking at a single chop direction as a function of chop throw. (Fig. 3.4) We can see that both the high and low order aberrations seem to stay the same, but change in amplitude. This brings up the question of stability. Do the aberrations that gain in brightness as a function of throw grow linearly and stay identical? We can easily find the scale factors between the frames by comparing two regions where the gradient is largest and plotting the difference in counts between the two regions. This has been done in Fig. 3.5.

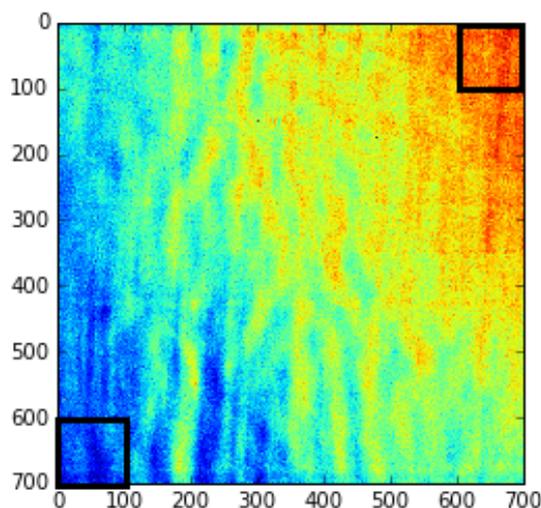


Figure 3.5: To test the linearity of the background we will compare the difference in counts in the two black boxes for three chop throws of 5", 10", 30".

3.2.1 Stability over time

This brings up the question of stability of time, as there is time between the frames. We know that the sky changes in the order of minutes and the telescope stays stable for a longer time. There is a time difference of about 15 minutes between each of the three exposures and therefore 30 minutes between the first and last exposure. If the images scale linearly this either means that both backgrounds did not change in this timespan or perhaps that several effects cancel out.

On top of that stability of the residuals over time is very important if we want to stack multiple exposures and lower the noise that way. In Pantin (2015) we see that while the S/N of nod subtracted frames improves with longer integrations, the S/N of chop subtracted frames stays roughly the same for integrations over 1 minute. It is also mentioned that the aberrations introduced by the telescope don't change rapidly for about 2-3 hours, potentially, allowing us to make some sort of flat that we can take out of our observations. However, when put to the test, we can see in Fig. 3.6 that the background changes very strongly over a period of two hours, going from lines at about 30° to nearly horizontal patterns.

This would suggest that our frames which have a Δt of 30 minutes at most should be fine, but it can never hurt to repeat this test for ourselves. This can be done by using observations with identical chop throw, frequency and direction.

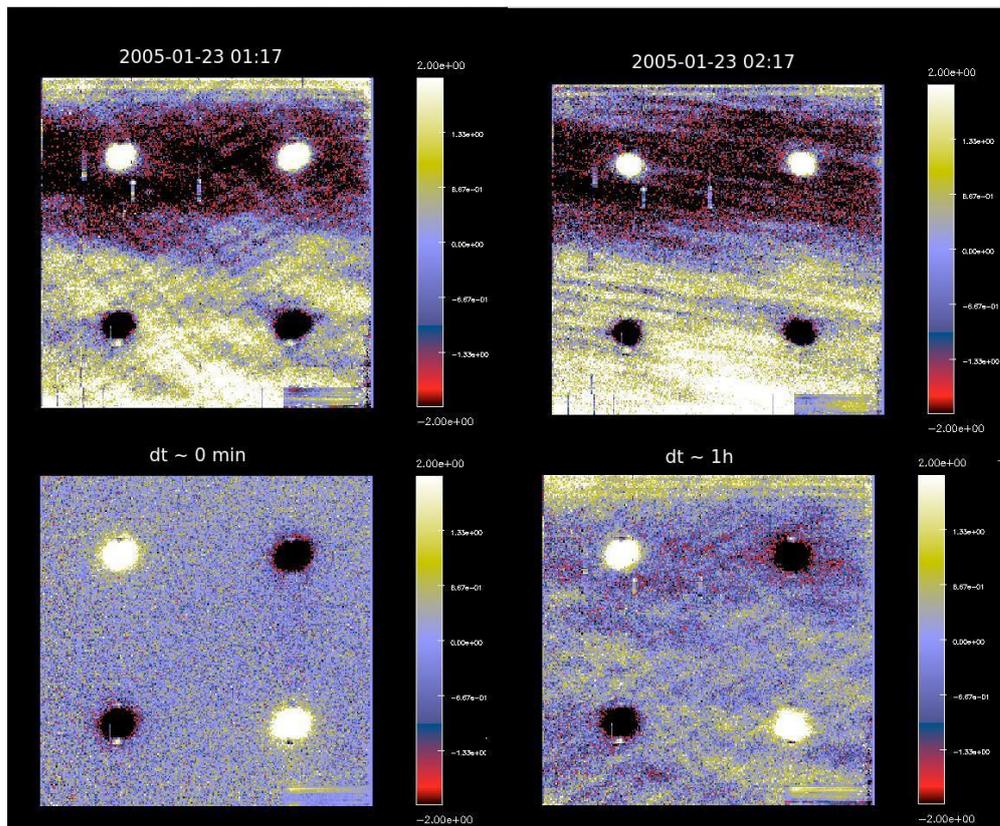


Figure 3.6: The two top images show a chop difference frame for a VISIR standard star. The two images are taken one hour apart, illustrating the strong changes in background residuals. The lower two images are nod difference images. The left one is taken with minimal time between the two nod frames and the second nod difference frame has 1 hour between the two nod frames. A clear difference is visible in effectiveness of the subtraction. Pantin (2015).

From our data we can find three identical datasets with 6 minutes between each set and one that has been taken about 3 months later. Therefore we can explore the stability of the patterns over time within the range covered by Pantin and far beyond it. This will be done in the same way as before, namely by use of the quadrant test on the nod subtracted frames. However instead of subtracting the datasets corresponding nod frame, we take the corresponding nod frame of the dataset that was taken at a later time. (see Fig. 3.7a.) Following the theory of Pantin (2015), this should give us four flat images at $dt=0,6,12,18$ min and one completely mismatched frame at $dt=3$ months.

A different way to test this linearity is to use asymmetrical nodding patterns

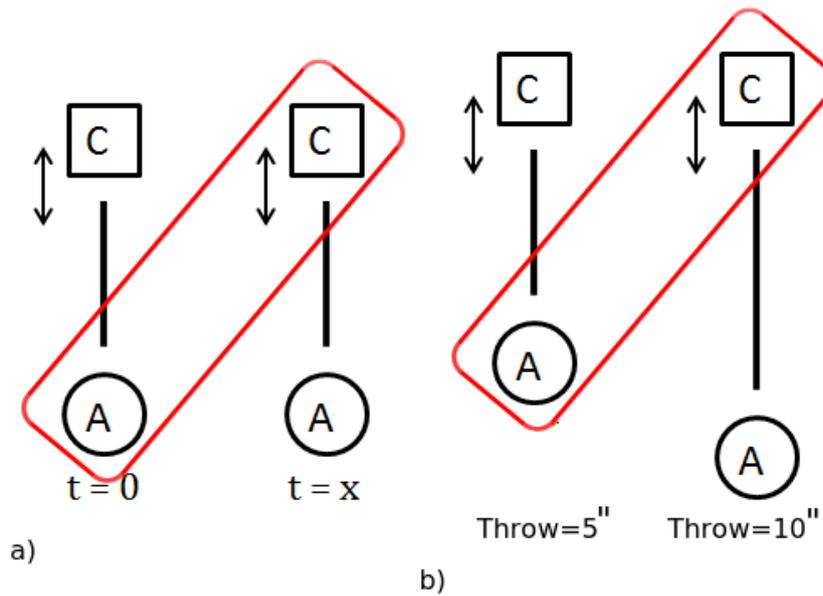


Figure 3.7: a) An illustration of classical chop/nod reductions with a certain time in between observations. We simulate this by matching two different pairs of observations with the same chop throw and chop direction. b) The same process as seen in (a) but here we match frames with a different chop throw. We attempt to correct for the residual amplitude by multiplying the first term by a scalefactor.

where we pair two nod frames with different chop throws. So for example we take nod position A in the AC direction with a 5'' throw and subtract the linearly scaled nod position B with a 10'' throw in the AC direction. (see Fig. 3.7b.) These frames are taken 15 to 30 minutes apart from one another so temporal effects, if negligible, will not play a role here. The results of both tests can be found in table 3.1.

3.2.2 Polynomial fitting

Looking at the structure of the aberrations as seen in Fig. 3.1 and Fig. 3.4, one could imagine that a simple polynomial fit might be enough to remove all large structures from the background, leaving only the higher order aberration. This would not be a good method for images with large structures, but could perhaps be useful for observations with small chop amplitudes. A first attempt would be to fit a 1st order polynomial to take care of the gradient and leave the sources intact. Therefore it is interesting to see if the fringes on the images are still dominant over the poisson noise. To test this theory we have fitted a first order polynomial to three chop difference frames with a chop throw of 5, 10

and 30". These images have then been subjected to the Quadrant test and the obtained results can be seen in 3.2. This method by itself will probably not be the best method of background subtraction, but could prove useful in combination with other methods.

3.2.3 Principal Component Analysis

Principal component analysis (PCA) is a statistical procedure that orthogonalises a dataset of observations into a set of linearly uncorrelated variables that we call principal components. We always get the same number of variables as the number of datasets that are put in, with the first few components being dominant and the rest negligible. The transformation also sorts the component on their magnitude, meaning that the component that introduces the biggest variability to the data will be first, the biggest contribution after that will be second and so on. Using this method we end up with a vector of uncorrelated orthogonal components which are eigenvectors of the symmetric covariance matrix. This method is sensitive to the scaling of the original variables. (Pearson, 1901)

This method is not useful for reducing the background, as it will subtract out the source, but we can plug in all 24 frames from the three nod crosses of 5", 10" and 30" throw to see if the images share certain structures.

Data	counts	μ_1	μ_2	μ_3	μ_4	$\bar{\mu} \pm \sigma\mu$	$\bar{\sigma}$
Theoretical	8000	0	0	0	0	1.1 ± 0	0
Classical 5'' AC	7990	0.5	0.5	0.60	0.60	0.60 ± 0.04	1.19
Classical 10'' AC	8076	0.2	0.16	0.17	0.17	0.18 ± 0.02	1.21
Classical 30'' AC	8160	0.71	0.72	0.68	0.62	0.68 ± 0.04	1.21
Classical 30'' AD	8189	0.32	0.30	0.30	0.38	0.33 ± 0.03	1.20
Classical 30'' AE	8033	0.31	0.26	0.19	0.15	0.23 ± 0.06	1.19
Classical 30'' AC 1Hz	7173	4.12	4.02	4.24	4.14	4.13 ± 0.08	1.64
Classical 30'' Airmass	7305	1.32	1.39	1.41	1.39	1.38 ± 0.03	1.13
Classical 30'' dt=0	8160	0.71	0.72	0.68	0.62	0.68 ± 0.04	1.21
Classical 30'' dt=6 min	6766	0.17	0.14	0.18	0.08	0.14 ± 0.04	1.10
Classical 30'' dt=15 min	6670	0.33	0.27	0.29	0.18	0.27 ± 0.05	1.10
Classical 30'' dt=84 days	6670	0.15	0.38	0.	0.28	0.2 ± 0.14	1.17
Asymmetric 5-10'' AC	7996	1.56	2.04	1.33	1.78	1.68 ± 0.26	1.22
Asymmetric 10-30'' AC	7992	3.65	5.71	2.29	4.61	4.07 ± 1.26	1.5
Asymmetric 5-30'' AC	7998	5.41	7.91	3.80	6.56	5.92 ± 1.51	1.64
Polynomial fit 5''	7990	0.01	0.02	0.02	0.01	0.01 ± 0.01	0.85
Polynomial fit 10''	8123	0.03	0.05	0.02	0.05	$0. \pm 0.04$	0.88
Polynomial fit 30''	8076	0.12	0.06	0.15	0.09	$0. \pm 0.11$	1.17
ICA 30'' AC-AE	8160	-9.38	-15.82	-7.47	-12.74	-11.35 ± 3.2	2.5
ICA 30'' AC-AE - polyfit	8160	0.02	0.03	0.07	-0.08	0.01 ± 0.05	1.48
ICA 10'' AC-AE - polyfit	8076	0.01	-0.	0.02	0.03	0.02 ± 0.01	1.2
ICA 5'' AC-AE -polyfit	7990	0.01	0.	0.01	0.02	0.01 ± 0.01	1.19

Table 3.2: In the first section we see the results of the Quadrant test for varying chop throw, direction, frequency and airmass. In the second part we show the Quadrant test for different times in between nods and asymmetrical chopping/nodding. In the last part we can see the Quadrant test results for the other reduction methods. The first column shows the average amount of counts in ADU of the image. The next four columns show the mean value of the four quadrants, followed by the mean of these four means \pm the spread on the values. The final value is the mean STD of the four quadrants.

3.2.4 Multidirectional chopping



Figure 3.8: Instead of subtracting the chop difference frames obtained from $Nod_{AB}Chop_A$ from $Nod_{AB}Chop_B$ we now add the chop difference frames obtained from $Nod_{AB}Chop_A$ to $Nod_{AD}Chop_A$.

The final method requires us to chop in different directions to simulate nodding without actually moving the telescope. This method uses the fact that opposing chop directions seem to be eachothers inverse and can remove close to all low order aberrations by simply adding the two difference frames. So we would be adding the chop difference frame obtained from Nod position A in the AB direction to the chop difference frame obtained from nod position A in the AD direction. (see Fig. 3.8) This is not just a hunch based on symmetries seen by eye, as mathematically this also makes sense, as shown below.

We start in the much same way as we did with Eq. 1.1 and define two chop observations A and B for our two nod directions AB and AD . We change the naming convention slightly to preserve clarity of which chop and nod frame we are working with. Therefore the flux F will have a subscript with a certain nod n followed by the chop position in the superscript. Besides this we introduce an offset term O that is defined as the difference between the telescope background of two chops. Now we can define the following four pointings.

$$\begin{aligned}
 F_{AB}^A &= B_1 + T_A + S \\
 F_{AB}^B &= B_2 + T_A + O_A \\
 F_{AC}^A &= B_3 + T_B + S \\
 F_{AC}^C &= B_4 + T_B + O_B
 \end{aligned} \tag{3.3}$$

Here we can assume $O_B \approx -O_A$ because we expect them to be roughly equal and that the reversed chopping direction also inverts the sign of the aberrations. Knowing this we can combine the four chops into two chopped difference frames, which we name P1 and P2.

$$\begin{aligned}
 F_{AB}^A - F_{AB}^B &= B_1 + T_A + S - B_2 - T_A - O_A = P1 \\
 F_{AC}^A - F_{AC}^C &= B_3 + T_A + S - B_4 - T_A - O_B = P2
 \end{aligned} \tag{3.4}$$

Next instead of subtracting them, we have to add these two frames in order to counter the minus sign introduced by reversing the chop direction. This gives

us our source and six background terms that can be paired and subtracted for each other, leaving the difference flux.

$$P1 + P2 = \Delta B_{12} + \Delta B_{34} - \Delta O_{AB} + 2S \quad (3.5)$$

The noise is given in the same way as for the classical chopping, namely by the square root of the sum of all the components, again with the source being neglected due to its amplitude.

$$\sigma = \sqrt{4T_A + B_1 + B_2 + B_3 + B_4 + O_A + O_B} \quad (3.6)$$

Just like with the offset term we can assume $\Delta B_{34} \approx -\Delta B_{12}$, meaning that we end up with a smaller sky and telescope background than in the separate chop difference frames, but a higher one than when applying classical chopping and nodding.

In table 3.2 we have placed the results of the quadrant test completed on a frame resulting from inverse chop addition and three frames resulting from a combination of this method and a polynomial fit.

3.3 Atmosphere VS Telescope

3.3.1 Wavelength Dependencies

To find out which components are contributed by the sky and which by the telescope we need to find a way to vary the amount of light given by either of the two components. While this is not possible with our currently used data, we could remedy this by observing in different filters. Therefore we have made observations in two more filters to see if this can be confirmed. (See Fig. 3.10.) The first filter was the *J7.9* filter which was chosen because most of the background comes from the atmosphere, while the *SIV* filter is chosen for the opposite reason of having a background dominated by the telescope. With this test and the closed dome observations we hope to be able to discover what effects are caused by the atmosphere and what effects are caused by the telescope. If the difference is big enough, it should even be possible to see the variations by eye.

If this is not the case, then we can use a more theoretical approach and write down these two filters as the sum of the two backgrounds. Now assuming that one of the backgrounds dominates completely in these images, we can say that,

$$\frac{F_1}{F_2} = \frac{B1_{Sky} + B1_{Telescope}}{B2_{Sky} + B2_{Telescope}} \approx \frac{B1_{Sky}}{B2_{Telescope}}. \quad (3.7)$$

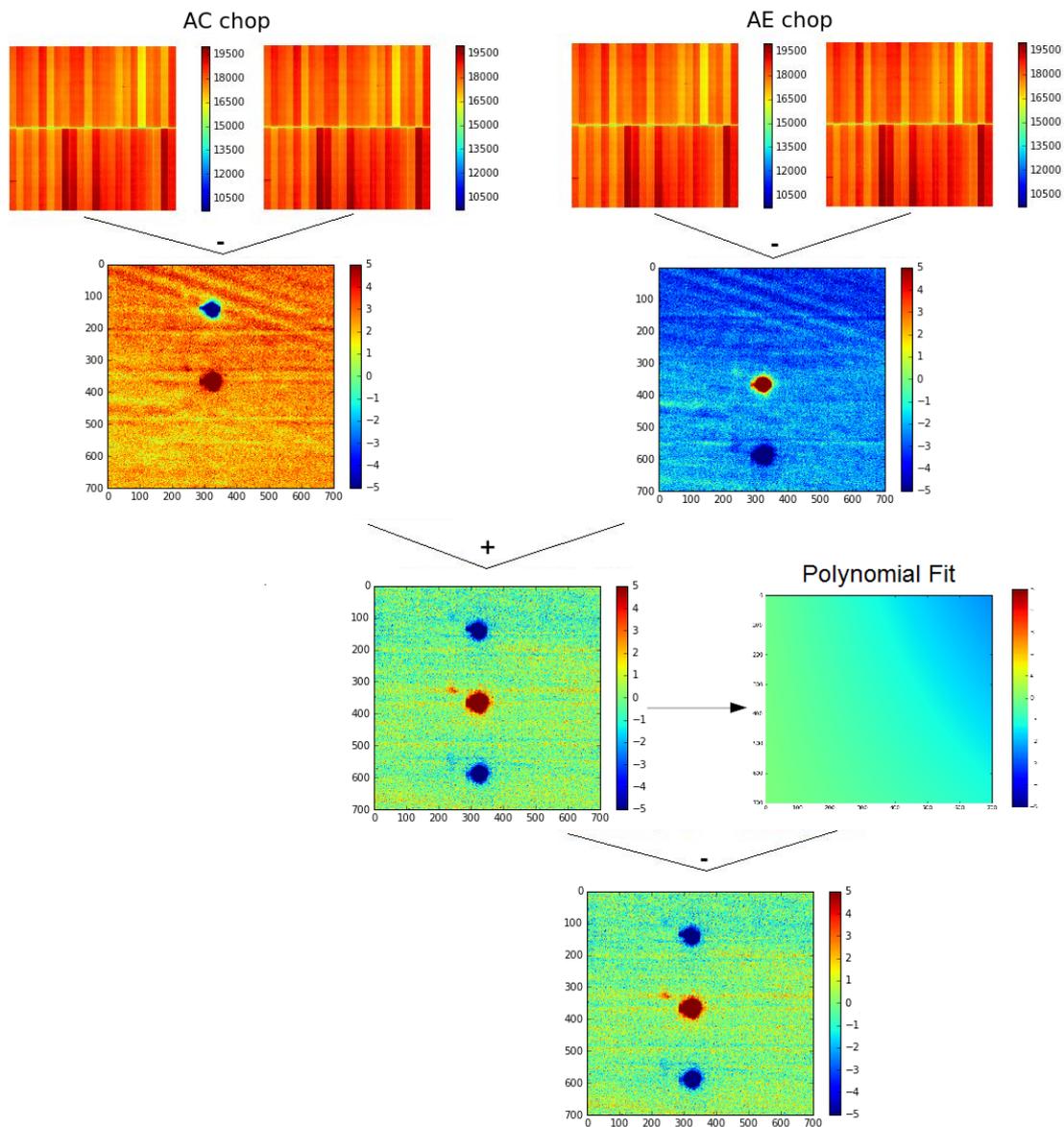


Figure 3.9: Inverse chop addition starts similarly to a classical chop/nod reduction by chopping on and off from the source in a given direction. Then instead of nodding the telescope, we repeat the same observations but with the chopping angle $+180^\circ$. This gives us two frames that apart from the source are each others mirror image. Adding these two images removes the higher order background residuals and leaves us with a twice as bright source and a gradient. After fitting and subtracting this gradient we are left with a flat image comparable to a classically reduced chop/nod image.

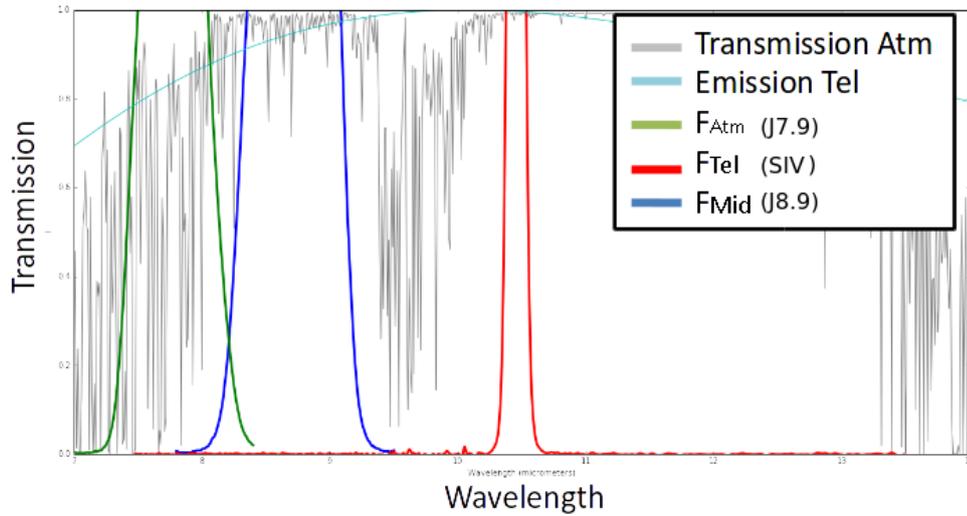


Figure 3.10: Three filters plotted over the sky transparency and a blackbody curve at 288K. The filters are chosen in such a way that the first (F_{Atm}) sees mostly the atmosphere and the third (F_{Tel}) sees mostly the telescope radiation while the second is in between. (F_{Mid}) These filter are named J7.9, SIV, J8.9 respectively.

Here the first filter is dominated by sky background and is approximated by assuming $B_{1Sky} \gg B_{1Telescope}$ while the second filter is quite the opposite and it is assumed that $B_{2Sky} \ll B_{2Telescope}$. With these ideal filters we would see just a gradient on one of the two filters and just the higher order pattern on the other. Dividing these two filters will result in an image that has certain characteristics. If we divide the pattern by the gradient we will see the pattern with an overlaid gradient but if we flip the fraction we will divide the gradient by the pattern and see a gradient inside of the pattern with a smooth background. This is illustrated in Fig. 4.8, where we took an arbitrary gradient and a higher order structure in the form of three bars of uniform brightness. Looking at both of these fractions we should be able to distinguish whether or not the relation is true for the two filters where $B_{2Sky} \ll B_{2Telescope}$ and $B_{1Sky} \gg B_{1Telescope}$

Unfortunately, we do not possess these ideal filters and all exposures have non negligible contributions of both backgrounds. However the amplitude of both components varies, so we should still be able to distinguish a pattern as shown in Fig. 4.8.

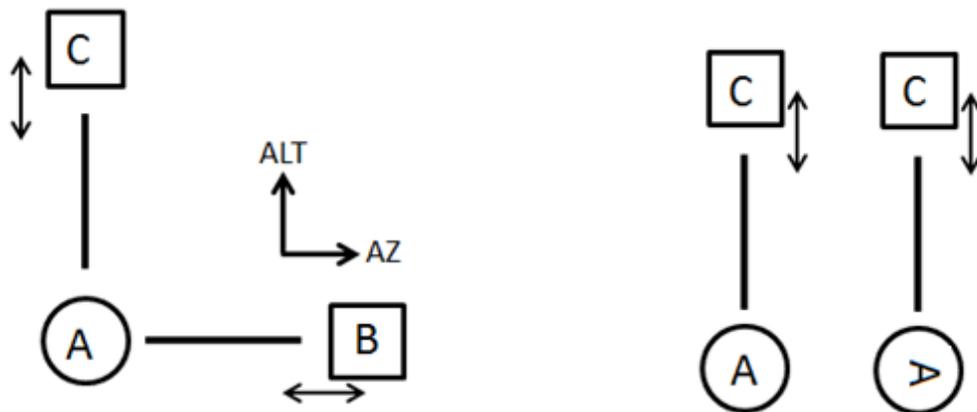


Figure 3.11: The derotator test does a two direction nod and then two nods in the same direction but with a 90° turn in the derotator. On the detector this will seem the same as doing the same motion twice.

3.3.2 Derotator Test

The second test is the so called derotator test, where we try to distinguish between the effects caused by the telescope and everything that happens inside of the instrument. This is achieved by making two standard chop observations in an L shaped pattern, as seen in Fig. 3.11, followed by a similar observations in which we chop up, turn the instrument by 90 degrees and chop up again. From the detectors point of view these two manoeuvres were both the same, namely an L shaped movement. This means that if the cryostat window or anything inside of the instrument is not rotationally symmetrical, we should see a difference between the two observations. Any instrument related asymmetries should show up as variations between the AB and \triangleright B frames, while telescope related variations would be the same in all observations taken in the same chop direction.

3.3.3 Effects as a function of airmass

Once we know that our residuals are stable over time and assume that the higher order aberrations are caused by the telescope angle, the question might come up of how this scales when the telescope moves and looks through different airmass. First we would assume that as there is more atmosphere to look through, the gradient will become stronger with respect to the higher order structures, but it would not be strange to assume that the higher order patterns stay the same. Due to the fact that most of our low airmass data is useless we

are forced to make do with one dataset taken at 45 and 30 degrees ALT with a random telescope rotation angle of 164 degrees. Due to this the patterns will probably differ from the ones in the other tests, but we should be able to compare the frames within this dataset to see the stability as a function of airmass.

3.3.4 Closed dome

Observing the inside of a closed dome should help us disentangle the backgrounds even further, as we will not see the atmosphere and only the telescope. Therefore doing these tests should give us more insight in these components.

Results

In this chapter we will discuss the perceived effects that occur when changing the variables discussed above, analyse the outcome of the quadrant test and other methods to find the optimal chopping strategies and pinpoint potential causes of aberrations.

We start by looking at table 3.2 to weed out all reduction methods that clearly do not pass the test. The first candidate for this is the classical chop/nod reduction at 1Hz, this frame has a significantly higher noise than its 4Hz counterpart and suggests that any other tests at lower chopping frequencies would only raise the noise floor, but also might mean that a higher frequency could lower it further.

Second is the polynomial fit, these reductions have taken out the overall gradient but left the higher order structure, leaving an image with a low STD but lacking in flatness for larger chops, as can be seen in Fig. 4.1. This could be a viable solution for chops with a throw of 5'' or less.

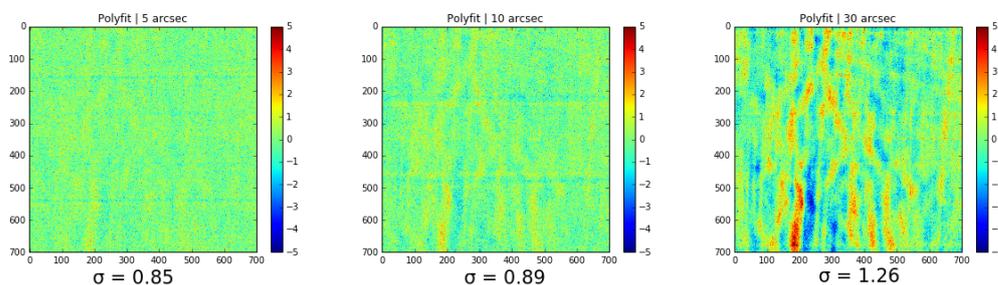


Figure 4.1: Background reduction done by a simple first order polynomial fit, removing only the gradient and leaving the rest.

Now that we know what data to not pay attention too, we can focus on the data that is left. The first question that is in want of an answer is the linearity of the background as a function of throw.

4.1 Effects as a function of Throw

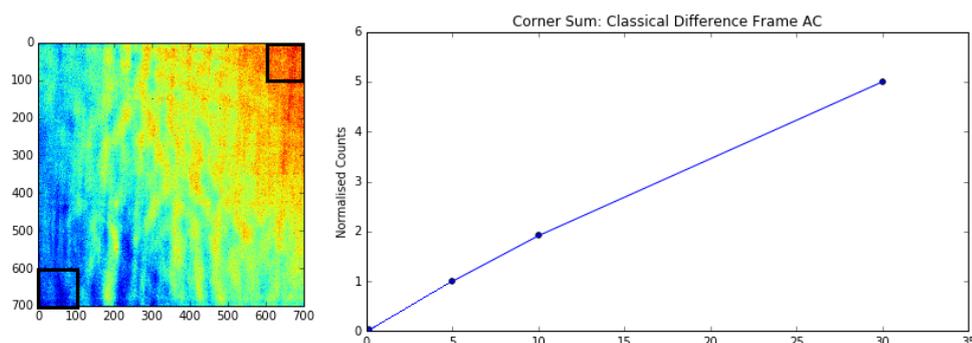


Figure 4.2: We compare two corners of a chop difference frame by summing the counts inside of the black squares and subtracting these two values to see if the patterns scale linearly. In the left image we see a chop difference frame in the AC direction with the two regions marked and in the right image we see the obtained values plotted as a function of throw.

In the last chapter we plotted the difference in counts between the maximum and minimum values of the gradient. When looking at Fig. 4.2, we see that the gradient seems to be stable over both time and throw. (As there is 15 minutes between each of the four frames.) However, we cannot assume that the increase is linear without also looking at the chop difference frames taken in other directions. Using the corner test on these data that we plotted in Fig. 4.3. The errors are given as the STD of the two used regions with a slightly larger samplebox, however they are so small compared to the data points that we cannot see them, meaning that there is no linear relation between chop throw and or time.

To learn more about which of the two are the main contributor to the non-linearity we can look into the asymmetrical chop reduction. In Tab. 3.2 we can see that these frames had by far the highest STD and the biggest variation in mean quadrant value. We still cannot say if this is due to time variations or not.

One of the ways of seeing this is by using PCA on all 8 frames of a nodding cross and the combined frames of the three crosses that we have for 5", 10", 30". Un-

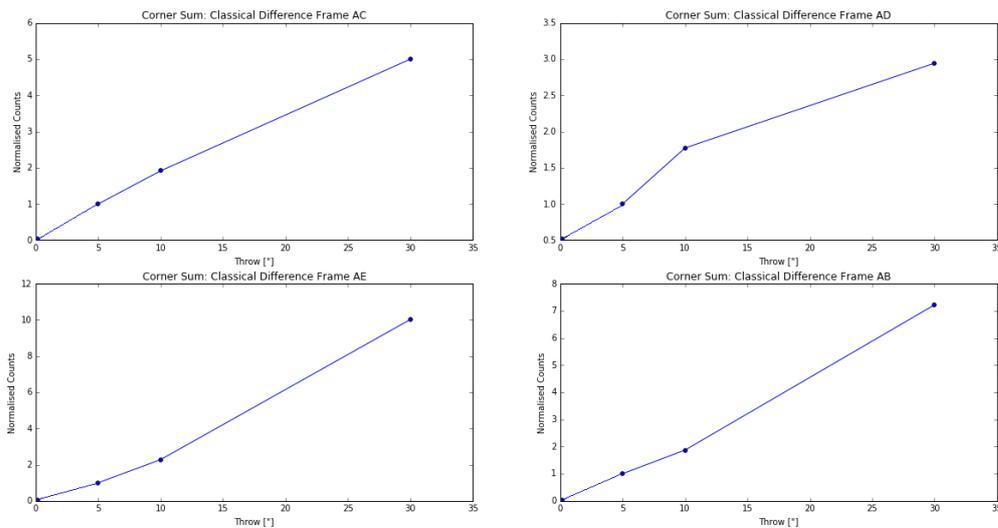


Figure 4.3: A test of linearity for all four directions of the nodding cross as a function of chop throw. The top left image shows the AC direction and the top right images shows the AD direction. The lower left and right images correspond to the AE and AB directions respectively

fortunately we need more than 5 principal components to reconstruct an image to within 10% for both cases and therefore we do not think that we cannot add anything useful by using this method with the limited amount of data that we have. For this reason we need to look for another route and make the residual stability as a function of time our next subject of interest.

4.2 Residual stability over time

Stability of the residuals over time is very important to understand properly if we wish to successfully reduce the background. We discussed in the chapter above how the theory from Pantin (2015) suggested that the background is stable for only 2 hours before it changes dramatically. This seems to be confirmed by our short dt data, where the background hardly changes at all in 15 minutes. But when comparing a frame to one taken nearly 3 months later, we see that this has also hardly changed at all. (See Fig. 4.4) This suggests that the change in high order aberrations is not dominated by the temperature fluctuations but something more systematic and is a big discrepancy compared to what is suggested in Pantin (2015).

It is hard to compare the data from Pantin with our own as the detector used for

this study has recently been replaced by the new Aquarius detector.(Ives et al., 2012) Therefore we will not be able to exactly replicate the effects seen in that paper. In spite of this we can still compare the images and find that the most striking difference is the fact that both the tracking and de-rotator have been turned off in our dataset.

The two observations in the Pantin data are one hour apart, which would translate to a 15° rotation on both the telescope and derotator. Comparing the difference in the angle of the pattern, we see that they compare to the 0 and 90 degree exposures displayed in Fig. 3.1. If we look at these images then we see that the high order aberrations have a stronger magnitude when chopping in the 0 degrees direction than when chopping in the 90 degrees direction. This means that if we subtract these two frames, we will be left with mainly the patterns on the 0 degree frame with a bit of the 90 degree patterns, as shown in Fig. 4.5 and also visible in the Pantin data on Fig. 3.6. Therefore we think that despite the difference in detector, the datasets are comparable and the effects causing the aberrations are the same. This suggests that the background is virtually stable with respect to the time when not moving the telescope or detector. On top of this it also suggests that the chopping residuals do not scale linearly as a function of throw as was theorised in the last section.

4.3 Alternative chop methods

Now that we have learned that the chop residuals do not scale with throw but are stable over time, we can investigate the effectiveness of our new chop reduction method. If we compare the values from the quadrant test of the classical method and those of the inverse chop addition with polynomial subtraction that are given in table 3.2, we will see that the quadrant test gives values that are identical within the given error apart from the largest throw of $30''$ where we have an offset of 18 times the STD. This suggests that the method does not work well for very large chop throws, but works well for smaller ones. In Fig. 4.6 we have plotted the fully reduced difference frames for both classical chopping and nodding and the inverse chop addition method. All the images have the overall STD given as a means of comparison of the overall flatness. From both visual inspection and this metric we can see that the methods seem to give the same result for $5''$ and $10''$ throws but break up for 30 arcsec. Given the non-linearity that we have seen in the background as a function of chop throw, this might mean that the balance between the two backgrounds breaks up or that the remaining contribution of the high order aberrations is no longer negligible. This effect would be more apparent in this method than classical chopping due

to the fact that the reduced terms still have a telescope offset term. (See Eq. 3.5.)

If we take a look at the inverse chop addition without subtracting a polynomial we will see that we get an image with just a gradient and no higher order aberrations. This means that the higher order structures and the gradient have a different source, which according to Eq. 3.5 can only be the atmosphere or telescope contributions. We may not precisely know which component contributes what, but we now do have a means of splitting the higher and lower order components. The gradient can be removed by a polynomial fit of first order and the other structure by the inverse chop addition. Combining these methods gives us a flat image comparable to classical chopping, as can be seen in Fig. 4.7.

Now that we have looked different methods of reducing the background in the chop difference frames, we should also realize that using the inverse chop method, even though we do not know which component introduces which effect, we could find out if we can vary the amount of power that is put into either of the two contributors.

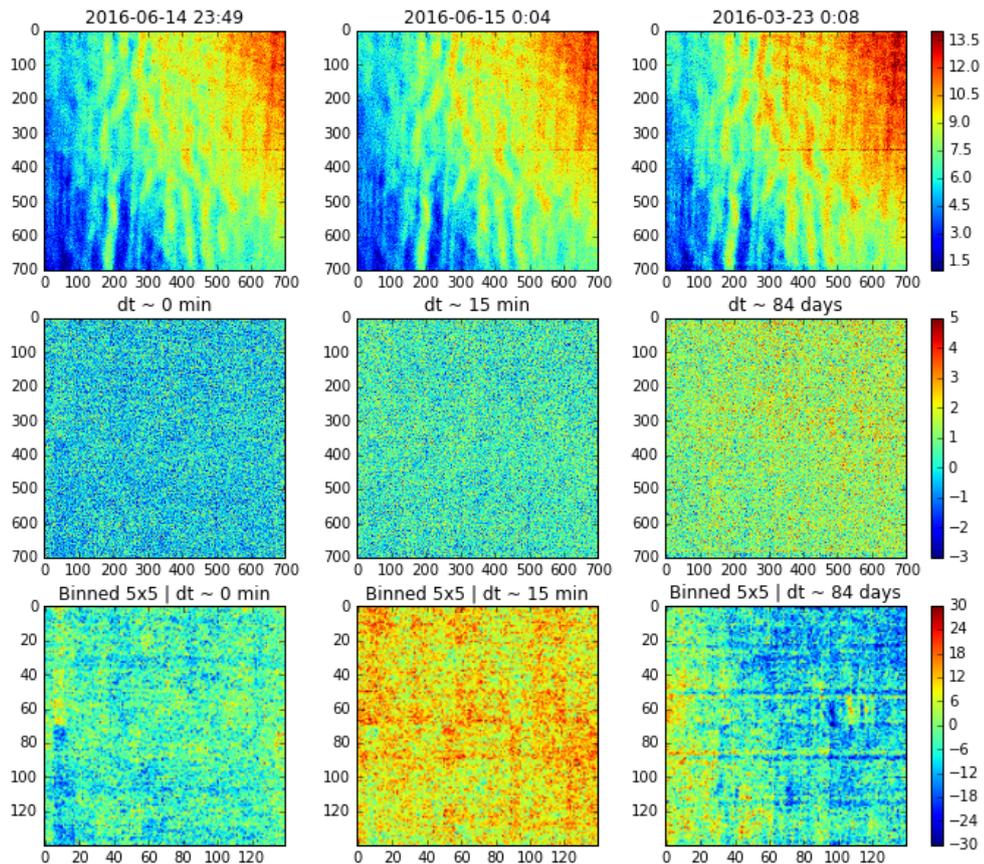


Figure 4.4: To test the stability over time we have taken the same observation three times at different times. Two frames are 15 minutes apart and the third is almost 3 months earlier. **Top Row:** The chopped difference frame for the three exposures. **Middle Row:** We see three classically chop/nod reduced images but with three different times in between the nods. **Bottom Row:** The same images as in the middle row but binned down 5x5 times.

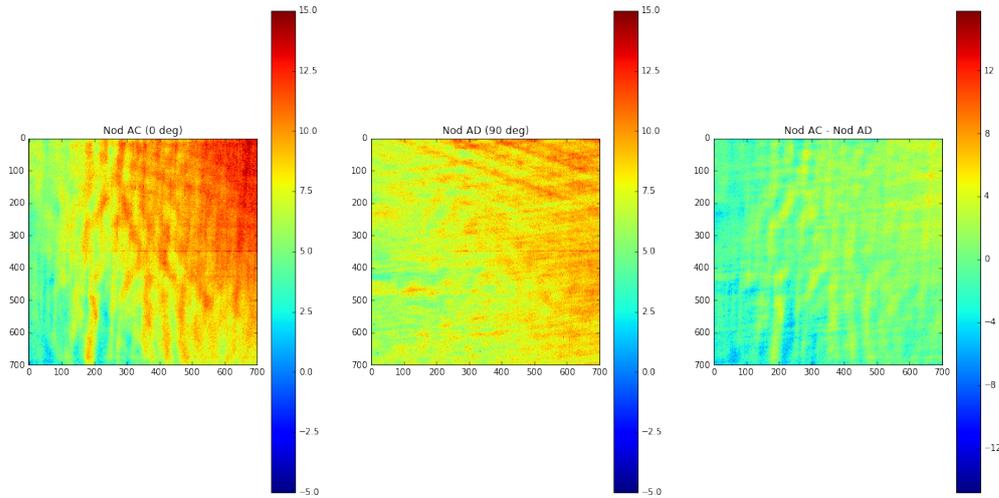


Figure 4.5: Left and Center: An illustration of the background obtained from chop difference frames at the 0 and 90 degree arm (AC and AB) and the result that is obtained when these two are used for classical chop/nod reduction. Right: The resulting nod difference frame is comparable to the one obtained by Pantin in Fig. 3.6.

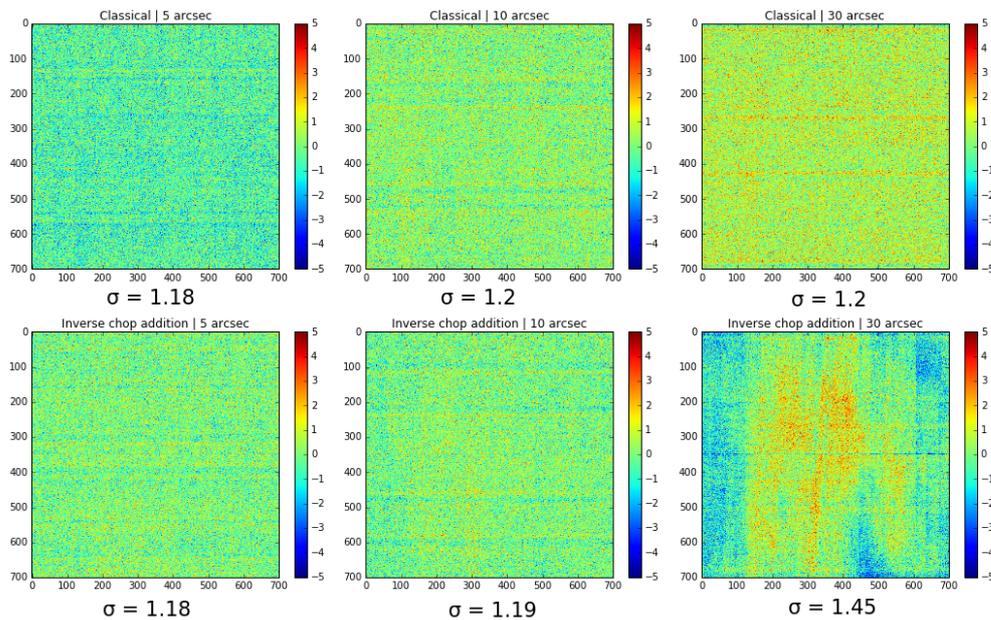


Figure 4.6: We have tested the inverse chop addition with polynomial subtraction method on three different throws and compared the results to classical chop/nod reductions. Top row: Classically reduced chop/nod images taken in the AC direction with a throw of 5", 10", 30" respectively. Bottom row: Inverse chop addition with polynomial subtraction on the same three frames as in the top row.

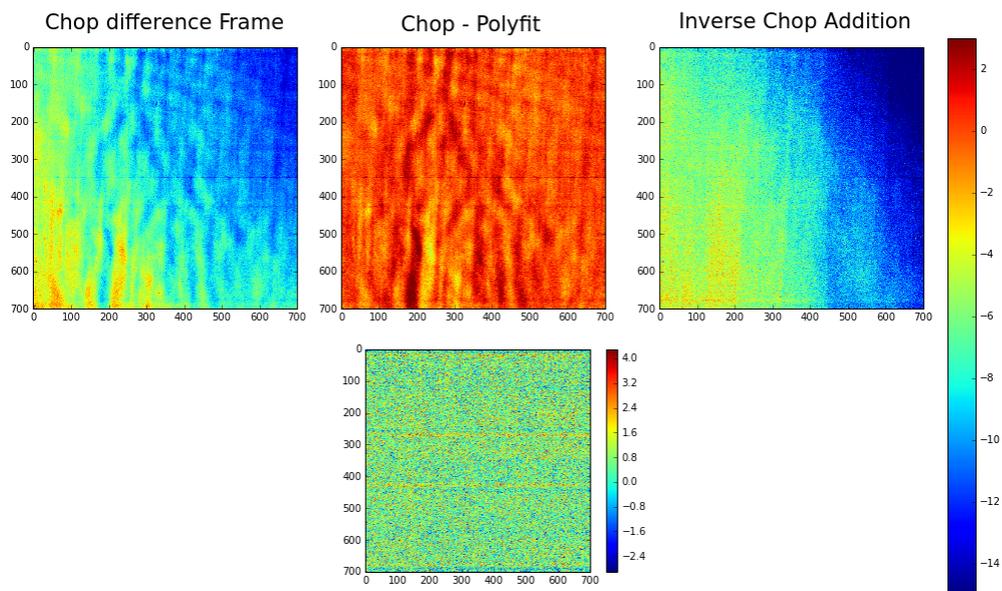


Figure 4.7: **Top Left:** We can see a strong gradient and higher order aberrations in the chop difference frame. (To be able to use the same colorbar, we have changed the sign of this frame.) **Top Center:** We can easily remove the gradient by fitting a first order polynomial, thus leaving the higher order structure. **Top Right:** We can also use the inverted chopping technique and add two difference frames to get rid of all the higher order structure but retain the gradient. **Bottom Center:** Combining both methods leaves us with a relatively flat image, comparable to that of classical chopping.

4.4 Disentangling Backgrounds

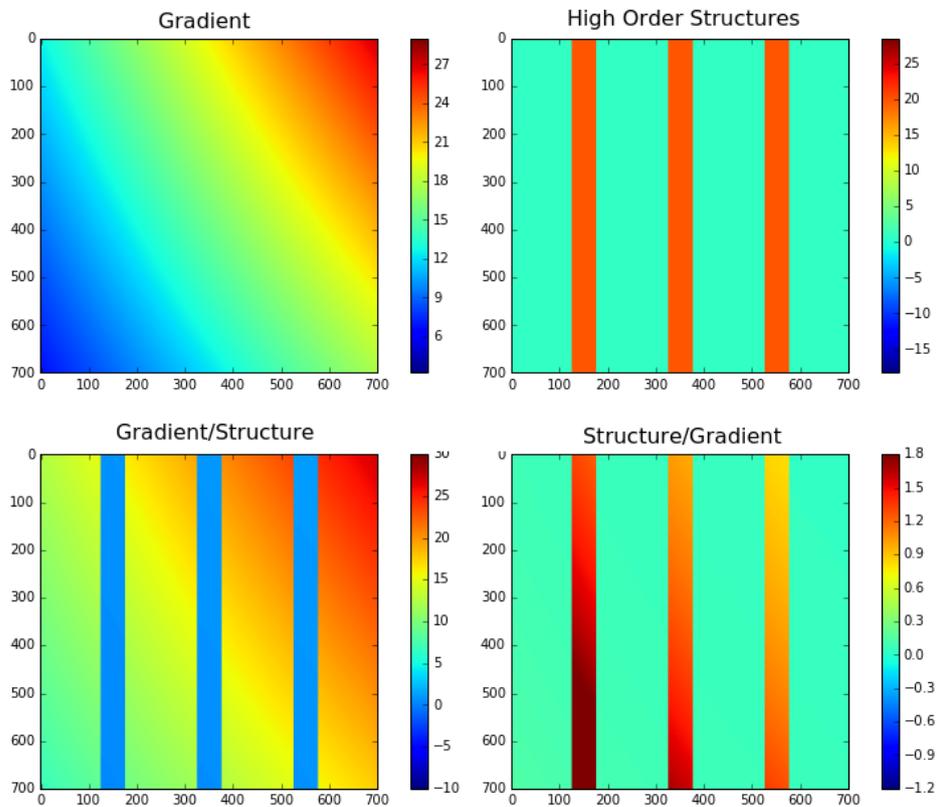


Figure 4.8: If we have a gradient (top left) and higher order structure (top right) we can divide the two to get the structures as seen in the lower row. Gradient/structure (bottom left) gives an overall gradient with the structure being dampened while structure/gradient (bottom right) will give us a flat image with the structure overlaid but with an internal gradient.

In the last section we reasoned that it should be possible to disentangle the two background components if we could vary the amount of power in either of the two contributors. The most straightforward way of doing this is by observing through different filters. In Fig. 4.9 we see a plot of the three filters that have been observed during our technical run at 0 and 90 degrees. Namely J7.9, SIV and the standard J8.9, which will be called the 'Sky', 'Telescope' and 'Middle' filters respectively for sake of simplicity. Looking by eye at the top row it is possible to see a difference in amplitude of the two components. the Sky filter seems to have less ripples and a stable gradient, while the Telescope filter has much stronger features. The bottom row seems to have less of a gradient alto-

gether, but the amount of higher order structure varies a lot. However since the filters also have different bandwidths and throughputs while having the same exposure time, the differences could be caused by that. Therefore we should take a look at the divided result of the two and see what comes out of that, since divisions are less sensitive to scalar differences.

We have done just that by taking the fractions between the Telescope, Sky and Mid filters. However due to the fact that the Sky filter was underexposed compared to the other two we could not get any useful results out of fractions containing this filter. The remaining two filters were much more similar in the amount of ADUs and therefore comparable. In Fig. 4.10 we have show the fractions between the Mid and Telescope filters for both directions.

Looking at the images obtained from the 0 degrees chop direction we have a difficult time saying anything based on the pattern as it seems to disappear on its own in the upper right corner, which is where we would like to see the effects of the division. Fortunately this is not the case for the images taken in the 90 degrees chop direction, where we can the strong features that go across the entire image get dampened strongly in the upper right corner and being brighter in the lower left corner. When compared to its counterpart in Fig.4.9, we have reason to believe that the gradient in the Mid filter image is stronger than the gradient in the Telescope filter image and therefore we see the structure/gradient effect when dividing telescope/Mid. This Suggests that the gradient is caused by the atmosphere and the structure by the telescope.

We could compare this data to the closed dome observations, but the chop difference frames are all perfectly flat with no residuals or gradients to be seen. (See Fig. 4.11) We think that this is due to the fact that the dome is significantly warmer than the atmosphere and completely out of focus and therefore smooth. If the dome background is much brighter than the telescope background then the latter would disappear in the shotnoise of the former, which seems to be confirmed by the high STD. This would give us identical images when chopping and thus the chop difference frames would already be flat.

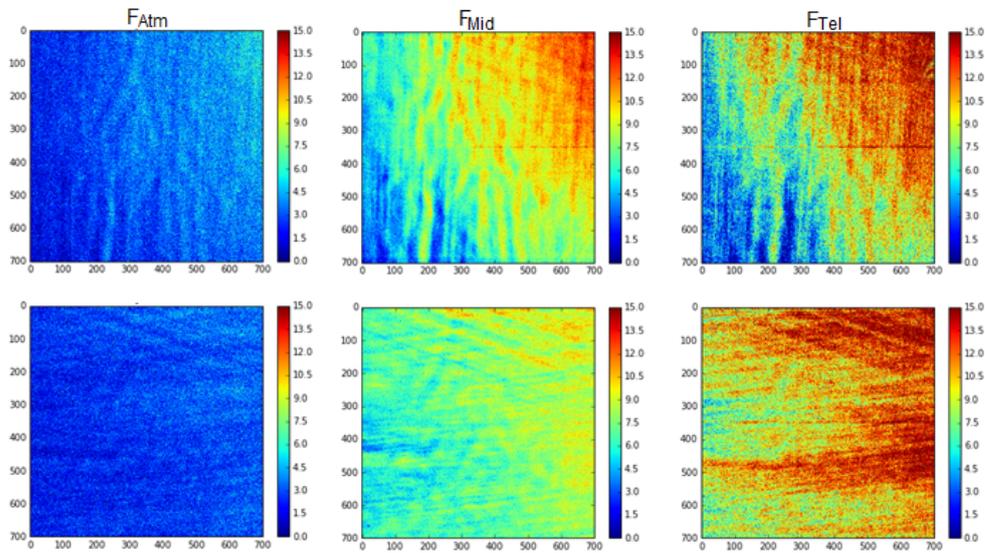


Figure 4.9: We make observations in three filters, namely the J7.9 as a filter that is dominated by the sky background, SIV as a filter that is dominated by the telescope background and J8.9 as something in between. **Top Row:** We see the the 0 degrees, AC chop arm in the three respective filters. **Bottom Row:** We see the same but for the 90 degrees, AB arm.

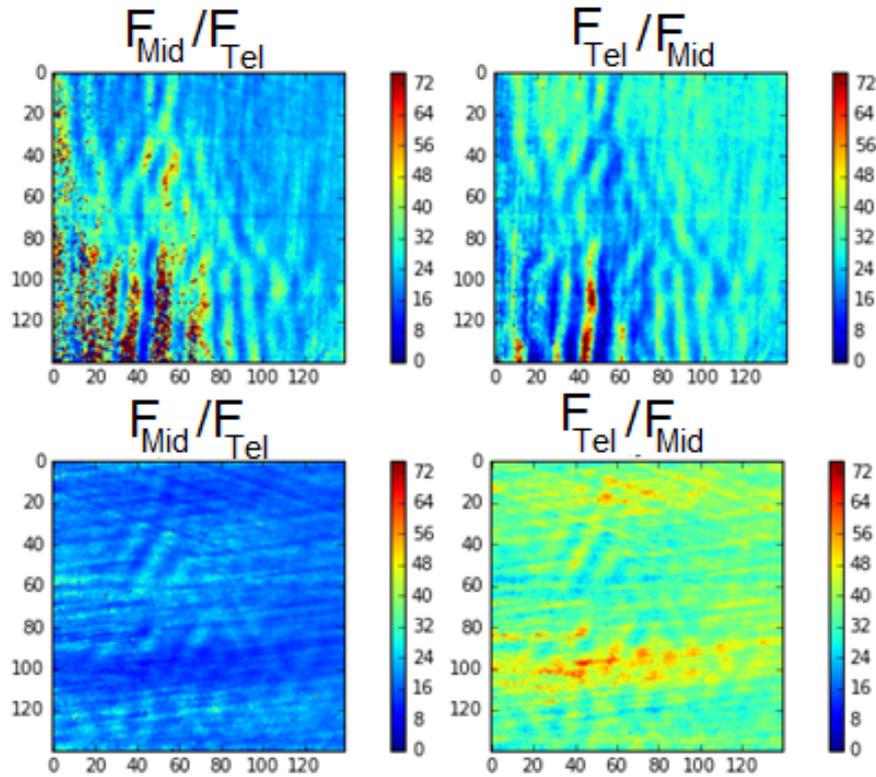


Figure 4.10: Results of the filter test for the two chop arms. All six possible permutations for divisions have been done to attempt to see the predicted patterns, but only two gave an usable result. **Top Row:** The filter test used on the AC arm with the Middle filter and the Telescope filter. Left we see Middle/Telescope and right we see Telescope/Middle. **Bottom Row:** The same as in the top row but for the AB arm.

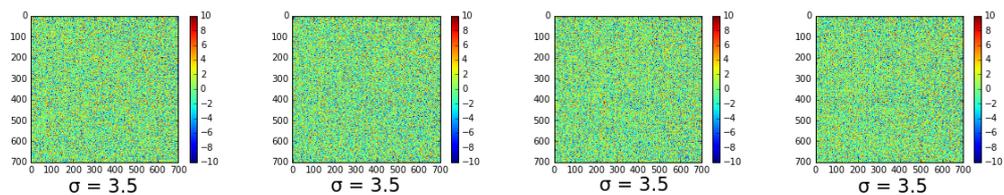


Figure 4.11: The chop difference frames for all four directions of the nodding cross taken with a closed dome. All images are identical with a STD that is about three times higher than the open dome images. No higher order residuals can be seen on these chop only images.

4.5 The Derotator test

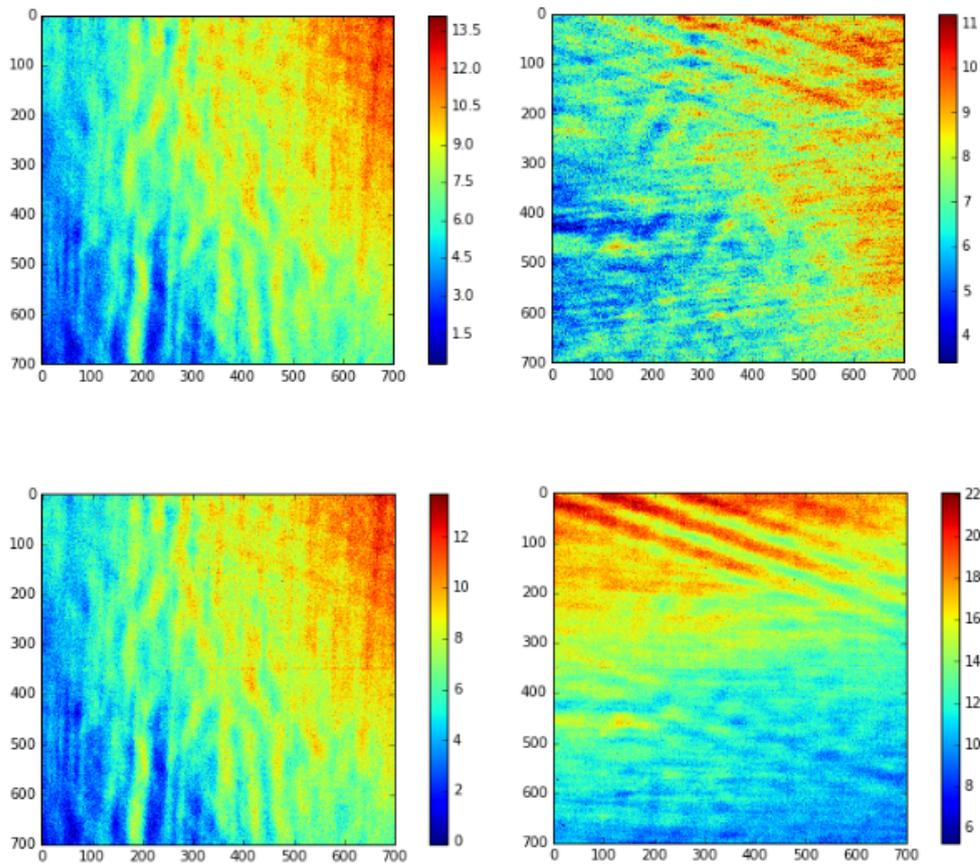


Figure 4.12: The derotator test with a 30'' chop throw in the AC and AB direction. **Top Row:** The control image where we have chopped in the AC and AB direction for the two images respectively. **Bottom Row:** The derotator test images with both images chopping in the AC direction and with the detector rotating 90° in the second image, simulating an AB chop from the detectors point of view. See Fig. 3.11 for a diagram of the test.

Our data suggests that the higher order residuals come from the telescope. We could take it a step further by trying to determine which part of the telescope is the main contributor of these residuals. Using our data from the derotator test can do just that and see whether the main factor is the telescope or the instrument. In Fig. 4.12 we see the results of the test with a 30'' throw. The pattern in the lower right panel is clearly more similar to the upper right panel than the left one. The higher order structure changed implying that this is caused by the instrument while the gradient direction flipped by 90 degrees, suggest-

ing that its direction is determined purely by the telescope chopping direction. This suggests that we might be able to remove the higher order structures by removing the defect in the instrument that causes them.

4.6 Effects as a function of airmass

It might also be interesting to see if the gradient and higher order structures change as a function of airmass. To test this we have plotted the four arms of the nodding cross at the two altitudes of 45 and 30 degrees ALT in Fig. 4.13. Here we can see that the patterns are indeed not the same as in Fig. 3.1 due to the rotation, but we can see some similarities. If we subtract the two frames we see that the background disappears quite well and we are left with flat images that have a similar spread to classically chopped and nodded images. It is not sure how this scales with airmasses with a larger difference than 22.5%

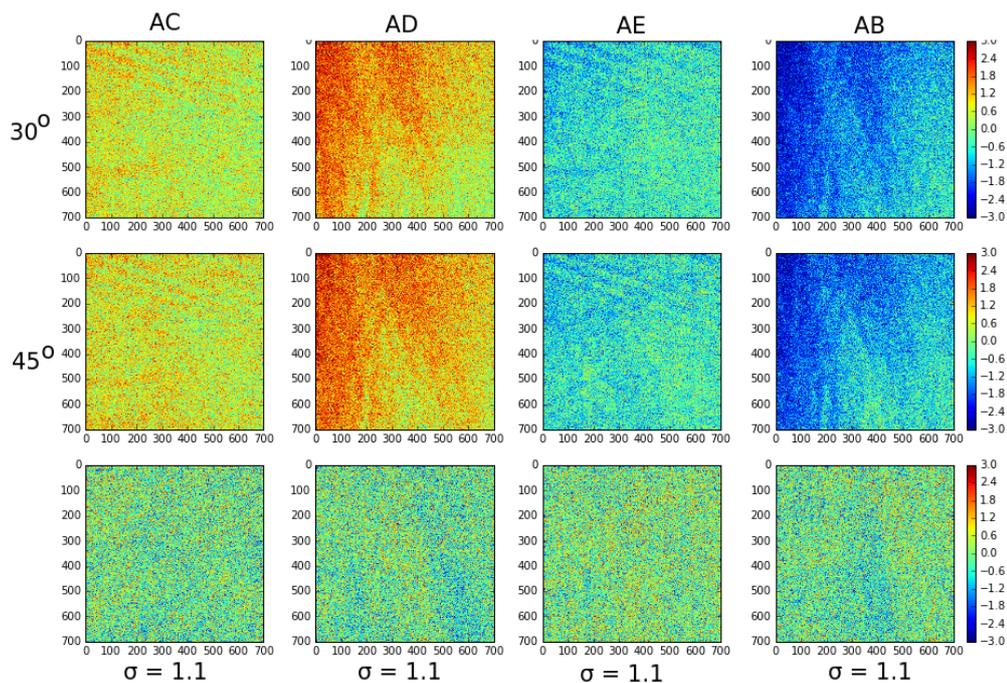


Figure 4.13: **Top Row:** A standard nodding cross executed at an elevation of 30° . **Center Row:** A standard nodding cross executed at an elevation of 45° . **Bottom Row:** The result obtained when the two frames above frames are subtracted from one another. The flatness of the image, comparable to classical chopping and nodding, implies that the patterns are not dependant on airmass.

4.7 Testing on a source

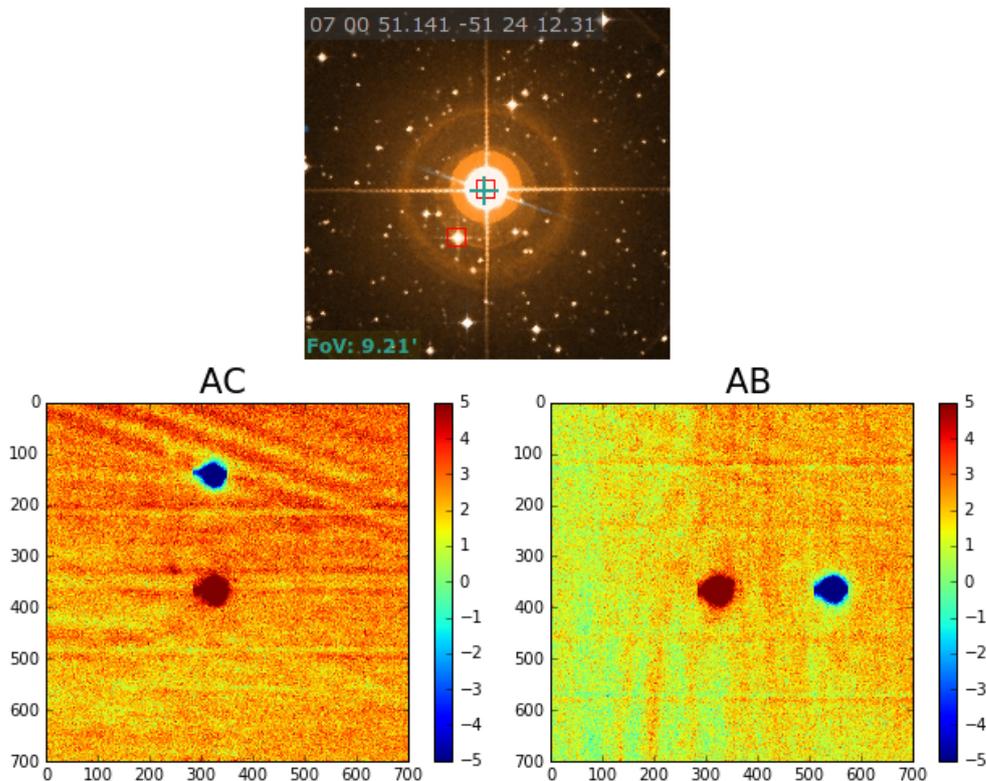


Figure 4.14: **Top:** HR 2652 as seen in the DSS survey. **Bottom:** The same object as seen by VISIR with a 0 (AC) and 90 (AB) degree chopping direction.

Now that we have tried and tested several methods on an empty patch of sky, it would also be interesting to see how these methods behave when looking at a source. For this reason we have made an observation of the bright standard star HR 2652 which has a relatively dim companion named TYC 8130-1560-1. This star was observed at 63 degrees ALT with a rotation angle of nearly 270 degrees and a chop throw of 10". In Fig. 4.14 we can see an image from the Simbad database and the AC and AB arm of the nodding cross, with patterns that correspond to the patterns that we can see in Fig. 3.1 but with a 90 degree rotation. Now we can test the effectiveness of the reduction method by comparing the end result with that of classical chopping and nodding, especially by whether or not we can see TYC 8130-1560-1.

In Fig. 4.15 we can see four sets of reduction methods applied on the two difference frames that were given in Fig. 4.14. We start with classical chopping,

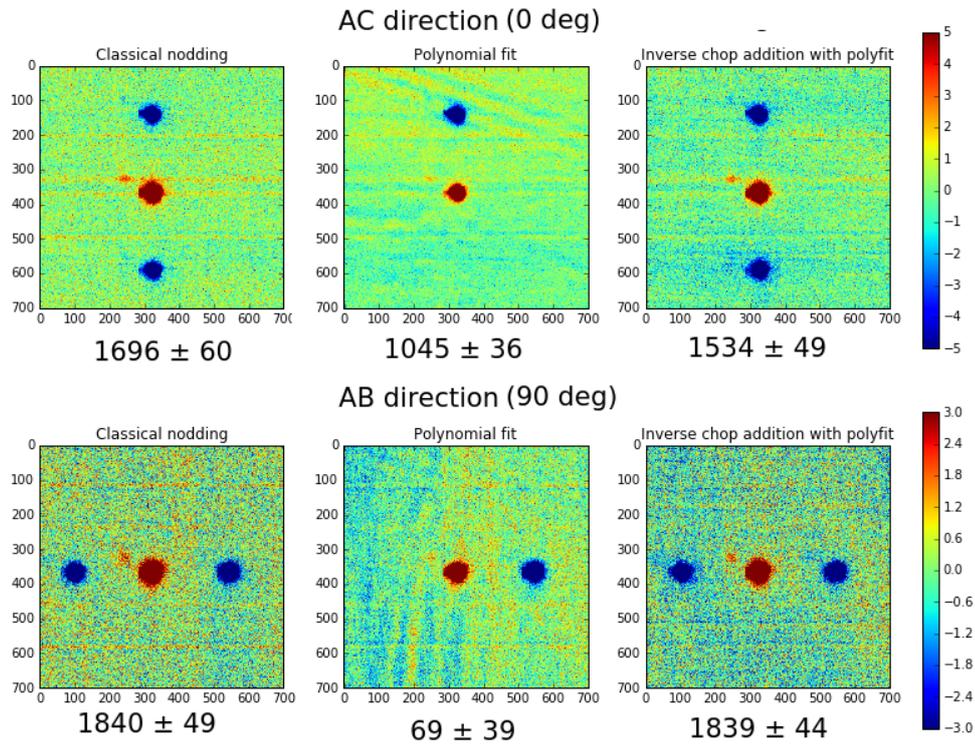


Figure 4.15: A comparison of three different background reduction methods. A simple photometric count has been made on the small companion of the central star and is displayed below each individual frame. **Left Column:** Classical chopping for both AC and AB directions that were shown in Fig. 4.14. **Middle Column:** The same sources after a simple reduction of only a first order polynomial subtraction. **Right Column:** The same sources reduced by use of inverse chop addition with polynomial fitting.

where we see the familiar tripple pattern with two negative stars and one positive star, and for both directions we can see a small point source near the upper left of the central star. We have done some basic photometry on this object by using the *aper* function from the Leiden ‘Sterrenkunde Practicum’. This gave us the background subtracted counts that are found in a region around this object. These numbers and a basic visual inspection will be our baseline for evaluation of the methods.

In the second column we can see the polynomial fit method, in which we only fit a polynomial to the data as a means of reduction. From both the counts and the image it is clear that this method will not work in practice, as the source is hardly visible at all and suffers from over subtraction.

The third column is the most new promising method, namely the inverse chop addition, where we add two opposing chop directions and then remove the resulting gradient. As can be seen this method gives us a very comparable result both visually and when looking at the numbers.

Conclusions and Discussion

We have observed an empty patch of sky for a long time and under different conditions. The Derotator test has given us a number of insights about the nature of the MIR background, its stability and cause. From this research we have drawn 5 conclusions which are listed below.

1. We have looked at the stability of the background as a function of throw and found that the background does not scale linearly and that the reduced images deteriorate in quality with throw. Therefore we suggest using the smallest possible throw for future observations. This will not pose a problem for point sources but means lesser performance for extended objects.
2. There is an alternative to getting a background comparable to classical chopping and nodding. The method is provided in the form of inverse chop subtraction combined with a polynomial fit. This is because the inverse chop subtraction rids us of the higher order structures and leaves a simple gradient that we can remove with a first order polynomial fit. We believe that due to the non-linearity of the background as a function of chop throw, this method is not stable for large ($> 10''$) throws. This is a good alternative for classical chopping and nodding.
3. It is shown that the amplitude of the gradient can vary heavily for different chop directions and over time while the higher order patterns stay the same if all other variables (e.g. AZ, derotator angle) stay the same. This implies that the gradient depends on something with a time-scale longer than 15 minutes while the high order aberrations depend on the chop direction and filter and have no time dependency.

4. Using observations from three different filters we have found that the background contributions from the telescope and sky are different. We have found that the overall gradient is due to the sky and the higher order structures are due to the telescope. Taken together with the last conclusion this implies that the higher order aberrations are introduced by the telescope and depend on the angle of the detector with respect to the ALT direction. Chopping in the ALT direction gives much stronger residuals than the two perpendicular directions.
5. From the derotator test we have learned that the lower order structures are caused by something inside of the instrument. This is either the window, the detector or some kind of interaction between the two. Finding the exact source of this effect could give us a method of solving the higher order aberrations, making it a lot easier to reduce the MIR background. However this is not doable with the data that we have and follow-up tests are needed.

While doing this project we have come up with some new methods of subtracting the background from MIR observations where nodding is not a possibility. We believe that when having a small enough chop throw it will most certainly be possible to properly reduce the image without nodding, however we have only uncovered the tip of the iceberg here. Most of our tests were very superficial and based on a handful of data points taken during one hour of observing at the VLT. To confirm these results we believe that it is crucial to make more observations where we can focus on the conclusions made above. For these observations we suggest the following:

1. Take a couple of nodding crosses in various filters with the proper exposure time to better understand the scaling of the gradient and the higher order structure. We suggest to use a chop throw of 10" to maximize the amplitude of both backgrounds while also staying in the regime where the images become truly flat after reduction. Hopefully this will shed more light on our theory of the atmosphere introducing the gradient and the telescope the other structures.
2. Similarly we need to test the background as a function of airmass. We have lost many good data points due to overexposure and the remaining dataset was far from perfect, although still indicating that the background does not depend on altitude. We need a few extra measurements with the same rotation and large airmass variations to confirm this.

3. To get definitive proof of whether the instability after two hours described by Pantin is truly due to the tracking and derotator, we have to do a similar test where we make a nodding cross with the tracking and derotator running with observations after 1, 2 and 3 hours. We can then compare the background data after three hours with the 90° arm of the nodding cross.
4. Once we have a better understanding of the above points, it is important to try inverse chop addition with polynomial fit methods out on point sources and extended sources. Comparing this to classical chopping should give a definitive answer to the question of whether they work or not.
5. Because we did not get open dome data for the GTC we have not worked with the data. Once the CANARICAM is operational again it would be interesting to finish the observations and redo the analysis that has been done in this thesis. If this proves to be impossible, a different telescope should be considered.

Acknowledgements

After spending a year on this research project I can honestly say that I have learned a lot about preparing an observing plan and working with MIR data. I very much enjoyed the balance between theory and instrumentation and the chance to work with my own data obtained from the VLT. Because of this I would like to express my deep gratitude to my supervisor, Professor Bernhard Brandl, for his guidance, help with ESO and constructive criticism. I would also like to thank Jeff Meisner, Roy van Boekel, Rudolf le Poole and Eric Pantin for their advice and suggestions concerning the project, as well as Konrad Tristram, Daniel Asmus and Julien Girard from ESO and Antonio Cabrera Lavers from the ICA for helping me obtain my data at the VLT and GTC. My grateful thanks are also extended to Steve Hammond, Dominique Petit dit de la Roche for extensively proofreading my work to remove all the typos.

Finally I would like to thank my parents and Dominique for their continuing support and encouragement throughout my study.

Bibliography

- B. Brandl. VLT Technical Time Proposal, A Study to develop Chopping/Nodding Strategies for E-ELT/METIS. 2015.
- B. R. Brandl, R. Lenzen, E. Pantin, A. Glasse, J. Blommaert, L. Venema, F. Molster, R. Siebenmorgen, H. Boehnhardt, E. van Dishoeck, P. van der Werf, T. Henning, W. Brandner, P.-O. Lagage, T. J. T. Moore, M. Baes, C. Waelkens, C. Wright, H. U. Käufl, S. Kendrew, R. Stuik, and L. Jolissaint. METIS: the Mid-infrared E-ELT Imager and Spectrograph. In *Ground-based and Airborne Instrumentation for Astronomy II*, volume 7014, page 70141N, July 2008. doi: 10.1117/12.789241.
- M. Cohen, R. G. Walker, B. Carter, P. Hammersley, M. Kidger, and K. Noguchi. Spectral Irradiance Calibration in the Infrared. X. A Self-Consistent Radiometric All-Sky Network of Absolutely Calibrated Stellar Spectra. 117:1864–1889, April 1999. doi: 10.1086/300813.
- T. Geballe and R. Mason. Introduction to Ground-based Mid-IR Observing. <https://www.gemini.edu/sciops/instruments/mid-ir-resources/mid-ir-observing>, 2006. [Online; accessed 2016-06-15].
- D. Ives, G. Finger, G. Jakob, S. Eschbaumer, L. Mehrgan, M. Meyer, and J. Steigmeier. AQUARIUS, the next generation mid-IR detector for ground-based astronomy. In *High Energy, Optical, and Infrared Detectors for Astronomy V*, volume 8453, page 845312, July 2012. doi: 10.1117/12.925844.
- S. Kendrew, L. Jolissaint, B. Brandl, R. Lenzen, E. Pantin, A. Glasse, J. Blommaert, L. Venema, R. Siebenmorgen, and F. Molster. Mid-infrared astronomy with the E-ELT: performance of METIS. In *Ground-based and Airborne Instrumentation for Astronomy III*, volume 7735, page 77355F, July 2010. doi: 10.1117/12.857043.

- J. H. Lacy, M. J. Richter, T. K. Greathouse, D. T. Jaffe, and Q. Zhu. TEXES: A Sensitive High-Resolution Grating Spectrograph for the Mid-Infrared. 114: 153–168, February 2002. doi: 10.1086/338730.
- S. D. Lord, D. J. Hollenbach, S. W. J. Colgan, M. R. Haas, R. H. Rubin, E. F. Erickson, P. Carral, and P. Maloney. A Far-Infrared Spectral Line Survey of 23 Disk Galaxies. In *American Astronomical Society Meeting Abstracts*, volume 24 of *Bulletin of the American Astronomical Society*, page 1182, December 1992.
- Jeff Meisner. personal communication, May 2016.
- J. B. Oke. Absolute Spectral Energy Distributions for White Dwarfs. 27:21, February 1974. doi: 10.1086/190287.
- Richter M Packham C Chun M Otarola, A. Atmospheric transmission and thermal background emission in the mid-infrared at Mauna Kea. *Adapting to the Atmosphere Conference 2014*, 2014.
- Sander Paalvast, Robert Huisman, Bernhard Brandl, Huub Jassen, Bayu Jayawardhana, Frank Molster, Maurice Teuwen, and Lars Venema. Development and characterization of a 2d precision cryogenic chopper for metis, 2014. URL <http://dx.doi.org/10.1117/12.2054874>.
- E. Pantin. Study of chopping only performances. 2015.
- K Pearson. Liii. on lines and planes of closest fit to systems of points in space. *Philosophical Magazine Series 6*, 2(11):559–572, 1901. doi: 10.1080/14786440109462720. URL <http://dx.doi.org/10.1080/14786440109462720>.
- G. S. Rothman, T. E. Gill, and C. Chang. Saharan Air Layer Interaction with Hurricane Claudette (2003). *AGU Fall Meeting Abstracts*, December 2004.
- G Rupprecht. Very Large Telescope, Requirements for Sceintific Instruments on the VLT Unit Telescopes. *ESO Internal*, March 2005.
- K. Tristram. Very Large Telescope: VISIR User Manual v98.1. 2016.
- K Volk. Chopping and Nodding for Mid-Infrared Astronomy. 2007.