High-contrast imaging of protoplanetary disks

Boer, J. de

Chapter 6

Characterizing the polarimetric imaging mode of SPHERE/IRDIS

Based on

J. de Boer, M. Langlois, R. van Holstein, J. H. Girard, J.-L. Beuzit, M. Kasper, C. U. Keller, D. Mouillet, H. M. Schmid, F. Snik, D. Stam, A. Vigan [1]

and

R. van Holstein, J. de Boer, J. H. Girard, J.-L. Beuzit, M. Kasper, C. U. Keller, M. Langlois, D. Mouillet, H. M. Schmid, F. Snik, D. Stam, A. Vigan [2]

Abstract

Polarimetric imaging is among the most promising techniques for high-contrast imaging and characterization of protoplanetary disks, and can become instrumental in the characterization of exoplanets. The new Spectro-Polarimetric High-contrast Exoplanet REsearch (SPHERE) instrument of the VLT contains a InfraRed Dual-band Imager and Spectrograph (IRDIS) with a dual-beam polarimetric imaging (DPI) mode. We aim to characterize VLT/SPHERE/IRDIS/DPI and validate our results with on-sky observations of the protoplanetary disk

[1] About to be submitted to Astronomy & Astrophysics.
[2] About to be submitted to Astronomy & Astrophysics.

surrounding TW Hydrae. We analyze the polarimetric efficiency and measured polarization angle offset for observations of TW Hydrae taken with the B_H filter. We calibrate the instrument for all broad-band filters with the internal light source in combination with the calibration polarizer and with unpolarized standard stars. These calibrations are used to create a Mueller matrix model of the instrument which we compare with, and use to correct the data of TW Hydrae. The calibrations show a strong dependency of the efficiency and polarization angle offset on the specific instrument cofiguration, especially the derotator angle in the filters B_Y, B_H and B_K_s. However, this dependency is small in B_J. The instrumental effects seen in the observations of TW Hydrae can be well explained by and corrected for with our instrument model. IRDIS/DPI is among the most powerful high-contrast polarimetric imagers currently available. However, the instrument performance is strongly dependent on the specific instrument configuration. We suggest adjustments to the observing strategy to minimize efficiency loss and application of the instrument model to derive the true polarization state of the incident light by correcting for the effects of instrument and telescope polarization and crosstalk.

6.1 Introduction

The direct detection of planets and protoplanetary disks in the visible and near-infrared (NIR) requires us to bridge the large brightness contrast between the bright star and its faint surroundings. Polarimetry has proven to be a powerful tool in high-contrast imaging (e.g. with HST/NICMOS Perrin et al. (2009), Subaru/HiCIAO Mayama et al. (2012) and VLT/NACO Quanz et al. (2013)). Starlight scattered by circumstellar material becomes polarized and can therefore be distinguished from the polarized stellar speckle halo. With the aid of extreme adaptive optics, polarimetric imaging has been successful in detecting circumstellar disks down to very small separations ($\sim 0.1''$, e.g. Garufi et al. 2016, Quanz et al. 2013). Compared to alternative high-contrast techniques such as Angular Differential Imaging (ADI, Marois et al. 2006), Polarimetric Differential Imaging is especially well suited to image disks seen at a low-inclination (i) angle (where $i = 0$ represents a face-on orientation), such as TW Hydrae (Krist et al. 2000, Rapson et al. 2015, van Boekel et al. 2017, $i \approx 7°$,). While ADI suffers from self-subtraction of signal from a disk with a low inclination, PDI remains sensitive to its scattered light.

Young, hot and therefore self-luminous extra-solar planets (or exoplanets) are expected to emit polarized light, typically up to a few percent of the total intensity (de Kok et al. 2011, Sengupta & Marley 2010). This polarization traces clouds or hazes in the planets' atmospheres, rotational flattening and the direc-

tion of the planets' spin axes. Compared to their younger counterparts, cooled down older ($\gtrsim 1\,\mathrm{Gyr}$) planets will emit at a much lower brightness in total intensity relative to their parent star (i.e. larger contrast). However, star light scattered off these planets is expected to be strongly polarized in the visible and NIR ($> 10\,\%$ for phase angles $\approx 90°$, Seager et al. 2000, Stam et al. 2004). These high degrees of polarization make polarimetric imaging a promising technique for the detection of older exoplanets.

In 2014, the Spectro Polarimetric High-contrast Exoplanet REsearch (SPHERE, Beuzit et al. 2008) instrument was commissioned at the Very Large Telescope (VLT) at Cerro Paranal, Chile, Unit Telescope 3 (UT3). The instrument's three scientific sub-components, assisted by the extreme Adaptive Optics (AO) system SAXO (SPHERE AO for eXoplanet Observation, Fusco et al. 2014) are: the (visible light) Zurich IMaging POLarimeter (ZIMPOL, Thalmann et al. 2008); the (NIR) Integral Field Spectrograph (IFS, Claudi et al. 2008); and the (Near) InfraRed Dual-band Imager and Spectrograph (IRDIS, Dohlen et al. 2008).

IRDIS is primarilly designed to detect planets in imaging modes combined with pupil tracking, which allows the removal of the stellar speckle halo by performing ADI. Two optical components at the end of the light path (just before IRDIS' detector) are a beam splitter and a filter wheel (FW2) with which we can introduce two different filters for the separate beams. Observations in two different color filters allows the characterization of planets (e.g. by observing methane emission in their atmosphere) with Dual-Band Imaging (DBI, Vigan et al. 2010). We can also choose to do Classical Imaging (CI) by inserting one single filter in an aditional filter wheel (FW1) just upstream from the beam splitter within IRDIS.

The inclusion of two sets of orthogonal polarizing filters (polarizers) in FW2 makes IRDIS a polarimeter. This Dual-beam Polarimetric Imaging (herafter DPI or IRDIS/DPI) mode is especially well suited for high-contrast imaging of circumstellar disks (e.g. Benisty et al. 2015).

Due to the complexity of the instrument and its many reflecting surfaces, the polarimetric efficiency of SPHERE/IRDIS is strongly dependent on the specific instrumental setup. This study describes the instrumental setup and characterizes the polarimetric performance of SPHERE/IRDIS. We describe the optical components encountered by the light beam in section 6.2; we use the observational data of TW Hydrae (van Boekel et al. 2017) as a case study where we explore the performance of the system in Section 6.3; a theoretical framework for telescope and instrumental polarization and polarimetric crosstalk is presented in Section 6.4; The instrumental effects described in Section 6.3 will be explained in Section 6.5 by performing additional calibrations and modeling the optical components described in Section 6.2. We apply this model to correct for the polarization effects along the optical path in the telescope and instrument in

Section 6.5.4, and end with our conclusions and recommendations in Section 6.6.

6.2 Design of the polarimetric mode IRDIS/DPI

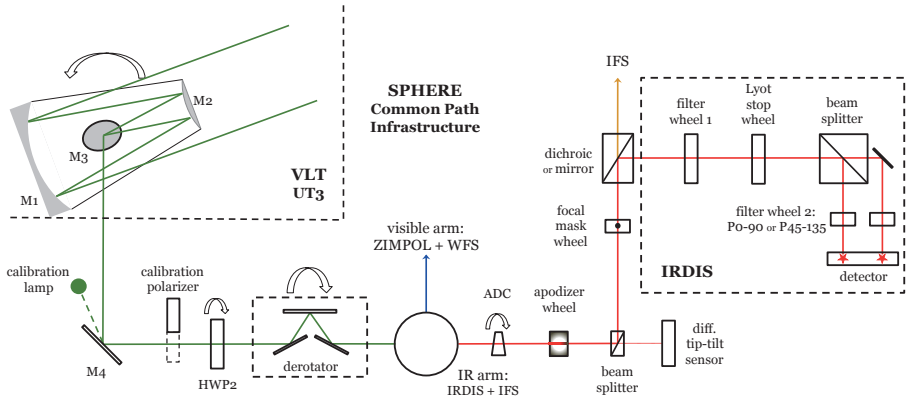

Figure 6.1: Schematic overview of VLT/SPHERE/IRDIS/DPI. The red arrows indicate which components rotate during an observation block. Reflections at angles of incidence ≥ 45° in the instrument are represented with similarly large incidence reflections in the figure. The green beam shows the starlight before color filters are applied, blue represents visible light, red and orange represent NIR light (with the orange beam towards the IFS showing the shorter wavelengths).

6.2.1 Unit telescope 3 and SPHERE common path

SPHERE is on the Nasmyth platform of VLT/UT3. To direct the light from the telescope towards the Nasmyth focus, the third mirror (M3) of UT3 is used, which introduces the first reflection at a non-zero (45°) angle of incidence. The incidence angles are important since larger angles are more prone to introduce modifications to the incident polarization signal, such as instrumental polarization and polarimetric crosstalk. As summarized in Figure 6.1, we describe below the optical components of SPHERE's Common Path and Infrastructure (CPI) and IRDIS, which create the instrumental effects on the polarization signal, are useful for calibrations or can be changed to alter the observational strategy.

At focal plane 1, a calibration light source, such as an integrating sphere, can be inserted. The first reflection in the light path is the pupil tip-tilt mirror (PTTM or M4, with a 45° incidence angle), which is the only mirror in SPHERE coated with aluminum.[1] The remaining mirrors of SPHERE are coated with protected silver for its higher reflectivity. A calibration polarizer with a fixed polarization angle can be inserted in the light path, just before the beam encounters HWP2, the only HWP available for IRDIS/DPI. The purpose of HWP2 rotations is twofold: The first is to account for field rotation to keep the incident polarization angle fixed on the detector; the second function of HWP2 is to switch between 4 angles (switch angles[2] $\theta^{s}_{\rm hwp} = 0°, 45°, 22.5°, 67.5°$) to measure $\pm Q$ and $\pm U$ with both beams in IRDIS independently. The next optical component is the derotator: three mirrors shaped like a 'K' (K-mirror) which rotates around the optical path to stabilize either the field or the pupil on the detector. Currently, IRDIS/DPI is only offered in field-stabilized mode. Multiple reflective surfaces with small angles of incidence follow in the AO common path, including the image tip-tilt mirror (ITTM), the 41×41 actuator deformable mirror (DM) and three toric mirrors (Hugot et al. 2012).

At focal plane 2, a dichroic beam splitter separates the light into a visible and a NIR arm. The visible light is reflected by the dichroic beam splitter and sent to SAXO's wave front sensor (WFS) and when required also to ZIMPOL (not offered simulteneously with IRDIS or IFS). The NIR beam is transmitted by the dichroic beam splitter, and is corrected for atmospheric dispersion, determined by the airmass during the observations. The Atmospheric Dispersion Corrector (ADC) is assumed not to affect the incident polarization in our analysis below, because the angle of incidence for the transmission through the ADC prisms is only $\sim 2.2°$. The beam then passes the apodizer wheel (which allows the apodization of the pupil in combination with Lyot coronagraphs) and is sent to a grey beamsplitter, which transmits 2% to a differential tip-tilt sensor (DTTS) and reflects the remaining light at a 45° angle of incidence. The reflected beam passes the NIR coronagraph (focal) masks (Boccaletti et al. 2008, Martinez et al. 2009) wheel in focal plane 3, and the Neutral Density (ND) filter wheel before reaching the final 45° angle reflection directing the beam towards IRDIS. For this reflection, we choose a mirror when only IRDIS is used, which is the option offered for DPI. We use a dichroic beam splitter when we use IRDIS in concert

[1] This coating gives M4 similar reflective properties as M3 of UT3, which is only useful for ZIMPOL. SPHERE contains a visible light HWP (HWP1) inbetween M3 and M4, which can keep the angles of the polarization induced by M3 crossed with M4, effectively canceling both their contributions. For the NIR, there is no HWP1 available.

[2] To avoid confusion with the true angle of HWP2 with respect to the Nasmyth platform, which also depends on the field tracking law, we use the superscript 's' in $\theta^{s}_{\rm hwp}$ for switch angles.

with IFS (IRDIFS is only offered in combination with DBI and CI, Claudi et al. 2008).

6.2.2 SPHERE/IRDIS

IRDIS is described in detail by Dohlen et al. (2008). Below, we summarize the optical components for a better understanding of the polarimetric performance of the system and for reference later in this text. The first optical component inside IRDIS is a common filter wheel (FW1). The filters of FW1 are the only color filters we can insert in DPI mode (since FW2 must contain the analyzer/polarizer pair). Among several narrow-band and spectroscopy filters, FW1 contains four broad-band filters which are offered for DPI (see Table 6.1).

filter	λ_c (nm)	$\Delta\lambda$ (nm)
B_Y	1042.5	139
B_J	1257.5	197
B_H	1625.5	291
B_K_s	2181.3	313.5

Table 6.1: IRDIS broad-band filters available in FW1, as described by the SPHERE manual.

Next, the beam encounters a Lyot stop wheel, just before the beam is split by the non-polarizing beam splitter (NBS) The beam transmitted by the NBS is reflected by an extra mirror in the direction parallel with the beam reflected by the NBS (and therefore with the same angle of incidence as the reflected beam: 45°). Each beam is finally reflected by their own spherical camera mirror which focuses the beam to the detector (not shown in Figure 6.1). The second filter wheel (FW2) is located inbetween the camera mirrors and the detector.

FW2 contains two pairs of polarizers: P0-90 and P45-135. The P0-90 analyzer set filters the light with polarization angles perpendicular to and aligned with the plane of the Nasmyth platform: the reflection plane for all reflections downstream from the derotator, while the P45-135 set polarizes at angles of 45° and 135° with respect to this plane. We expect the P45-135 to be highly sensitive to crosstalk introduced by all reflections in this plane. Therefore, we limit this study to the analysis of the telescope and instrumental modification of the polarization measured by IRDIS using the P0-90 analyzer pair, while using HWP2 to switch between $Q^{\pm}_{\rm meas}$ and $U^{\pm}_{\rm meas}$ measurements. The Hawaii-2RG detector is mounted on a dither stage and has 2048×2048 pixels with 18 μm pitch. Two

quadrants (2048 × 1024 pixels) are used, each of which records one beam with a field of view (FOV) of ~ 11″.

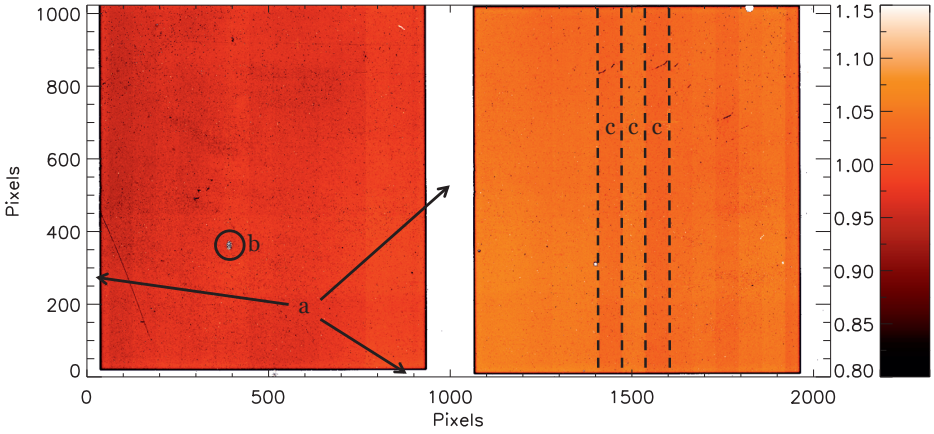

Figure 6.2: Normalized IRDIS master-flat image in *B_H*-band. The illuminated detector halves are surrounded by the field-mask {**a**}. The masked regions (and all pixels with value < 0.1) are set to 1000 and shown as white in this image. This enhances several clusters of dead pixels (e.g. {**b**} around [x,y] = [400,400]). A time varying read-out signature of the detector is visible as columns of 64 pixels wide {**c**}.

6.3 Case study: TW Hydrae

We observed TW Hydrae during the night of March 31, 2015, with the polarimetric mode of SPHERE/IRDIS in *B_H*-band (see Table 6.1), while an apodized pupil Lyot coronagraph with a focal plane mask with radius of 93 mas was used. The observations were taken with a detector integration time (DIT) of 16 s per frame, 4 frames per file, during 25 polarimetric cycles (also called 'HWP cycles', containing the switch angles $\theta^{s}_{\mathrm{hwp}}$ = 0°; 45°; 22.5°; and 67.5°). This adds up to a total exposure time of 106.7 min. The data and their analysis are discussed in van Boekel et al. (2017). In this study we give a more detailed description of the data reduction. We limit our analysis of the reduction to the polarimetric performance of the instrument, and consider the astrophysical analysis of the data to lie outside the scope of this study.

6.3.1 Data reduction

For the reduction and post-processing of the data, we did not use the official Data Reduction and Handling (DRH) pipeline[1], but our own custom routines described below.

Raw IRDIS frames are of 2048×1024 pixels in the x, y direction. The two beams separated by the beam splitter (see Section 6.2.2) are centered roughly on the left and right detector halves, while a field mask avoids leakage of signal from one half to the other.

We median combine all dark observations taken with DIT = 16 s to create a master-dark image . From internal light source (flat-field) measurements, we create a (dark-subtracted) master-flat image: we take median over two regions of 800×800 centered on the left half ($[x, y] = [512, 512]$) and the right half ($[x, y] = [1536,512]$) and use this value to normalize. To avoid emphasizing dead pixels or pixels masked by the field mask, we change all pixel values < 0.1 into 1000. The final master-flat is shown in Figure 6.2.

To correct for detector dark current, we subtract the master-dark from each science frame recorded of TW Hydrae, after which we divide the science frames by the master-flat to correct for flat-field errors. Next, we crop the frames to separate the left ($1 \leq x \leq 1024$) from the right ($1025 \leq x \leq 2048$) detector half. We center the star by cross-correlation with a 2-dimensional Moffat function with a central gap of six pixels (to account for the coronagraph mask) for both the left and the right frame.

We determine I_{meas} according to Equation 1.6, and median combine the images of all HWP cycles to create a reference image. To enhance our centering, we repeat the cross-correlation for all frames, this time using the reference image. Due to irregular diffraction paterns close to the edge of the coronagraph mask, we cannot guarantee with the two centering steps described above that the star-center lies at the center of the image. However, it does place the star close to the center of the image, and at the same location for each frame, which is crucial to perform the single and double difference subtractions. A more advanced centering is performed at a later stage. The single difference images (Q^+_{meas}; Q^-_{meas}; U^+_{meas}; and U^-_{meas}) are determined frame by frame for $\theta^s_{\text{hwp}} = 0°; 45°; 22.5°;$ and $67.5°$, respectively, and median combined per file. For each HWP cycle, Q_{meas} and U_{meas} are computed with the double difference method of Equations 1.10 and 1.11. The total intensity image I_{meas} is computed per HWP angle with Equation 1.6 and subsequently averaged over the four HWP angles.

We perform a first order correction of instrumental polarization created upstream from HWP2 (most likely by M3 and M4) on the Q_{meas} and U_{meas} images of each polarimetric cycle. This correction method (as described by Canovas et al.

[1] http://www.eso.org/sci/software/pipelines/

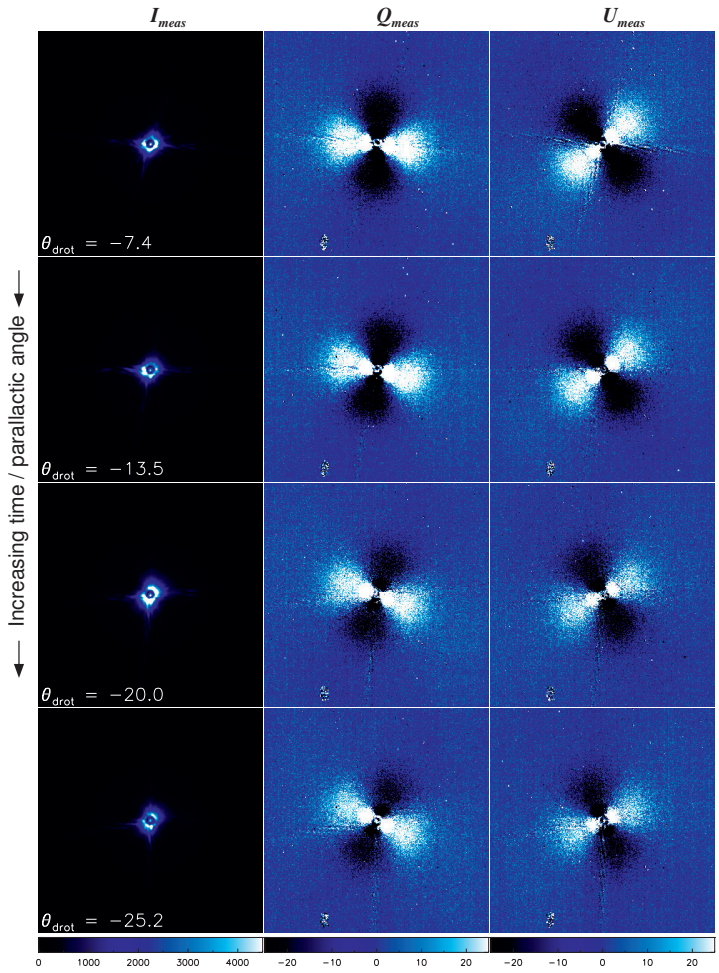

Figure 6.3: Each row shows the $I_{\rm meas}$, $Q_{\rm meas}$ and $U_{\rm meas}$ images (from left to right) for polarimetric cycles with increasing parallactic angle from top to bottom. Although the orientation is the same for all panels (north is up, east is left), the measured polarization angle (Equation 1.5) is changing and the polarimetric efficiency ($Q_{\rm meas}$ and $U_{\rm meas}$ signal) decreases with derotator angle.

2011) is based on the assumption that the direct stellar light is unpolarized. We

take the median of the $Q_{\rm meas}/I_{\rm meas}$ signal over an annulus centered around the star (excluding the coronagraph mask) to obtain the scalar $C^q_{\rm annulus}$ (likewise, we determine $C^u_{\rm annulus}$ with $U_{\rm meas}/I_{\rm meas}$); multiply this scalar with $I_{\rm meas}$; and subtract this from the $Q_{\rm meas}$ image. The instrumental polarization corrected linear stokes components are:

$$Q^{\rm ipc}_{\rm meas} = Q_{\rm meas} - C^q_{\rm annulus} \times I_{\rm meas}. \quad (6.1)$$

$$U^{\rm ipc}_{\rm meas} = U_{\rm meas} - C^u_{\rm annulus} \times I_{\rm meas}. \quad (6.2)$$

We remove five HWP cycles where the AO was not correcting optimally, which leaves us with 20 pairs of $Q^{\rm ipc}_{\rm meas}$ and $U^{\rm ipc}_{\rm meas}$ images.

A residual of the read-out columns (see feature **c** in Figure 6.2) remains visible in the double difference images. Similar to how Avenhaus et al. (2014) removed noise across detector rows from NACO images, we remove these artefacts from the double difference images by taking the median over the top and bottom 20 pixels (to avoid including signal from the star) on the image per pixel column (not the 64 pixel wide read-out column), and subtract this median value from the entire pixel column.

Figure 6.3 shows the $I_{\rm meas}$, $Q^{\rm ipc}_{\rm meas}$ and $U^{\rm ipc}_{\rm meas}$ images for four polarimetric cycles observed with increasing parallactic angle for each subsequent panel row. While each $Q^{\rm ipc}_{\rm meas}$ and $U^{\rm ipc}_{\rm meas}$ panel displays the typical 'butterfly' signal of an approximately face-on and axi-symmetric disk, strong variations occur between the polarimetric cycles: the butterflies appear to rotate in clockwise direction and the $Q^{\rm ipc}_{\rm meas}$ and $U^{\rm ipc}_{\rm meas}$ signal decreases with time. Since the observations were taken in field tracking mode, the image of the disk itself does not rotate on the detector, rather the polarization angle $\theta_{\rm p}$ (Equation 1.5) is varying between the polarimetric cycles.

Because of this variation between the polarimetric cycles, we cannot simply stack all $Q^{\rm ipc}_{\rm meas}$ and $U^{\rm ipc}_{\rm meas}$ images. To create the final polarization image we have two choices: **1)** We can compute the polarized intensity $PI_{\rm L}$ according to Equation 1.3 for each HWP cycle and median combine these to create a final $PI_{\rm L}$ image (with a higher signal-to-noise then the $PI_{\rm L}$ images of individual cycles). The problem with this method is that the squares taken in Equation 1.3 boost the noise, which is no longer averaged around 0. The squared noise then becomes harder to disentangle from the $PI_{\rm L}$ signal. For example, artefacts seen as a bright (positive or negative) feature detected at a point in the $Q^{\rm ipc}_{\rm meas}$ image where the signal should be ~ 0, (on the diagonal 'null' lines separating the positive from the negative $Q^{\rm ipc}_{\rm meas}$ signal) or a strong positive signal in a region where $Q^{\rm ipc}_{\rm meas}$ ought to be negative, will be indistinguishable from true disk signal in the $PI_{\rm L}$ image. This is actually a general problem we encounter when computing $PI_{\rm L}$,

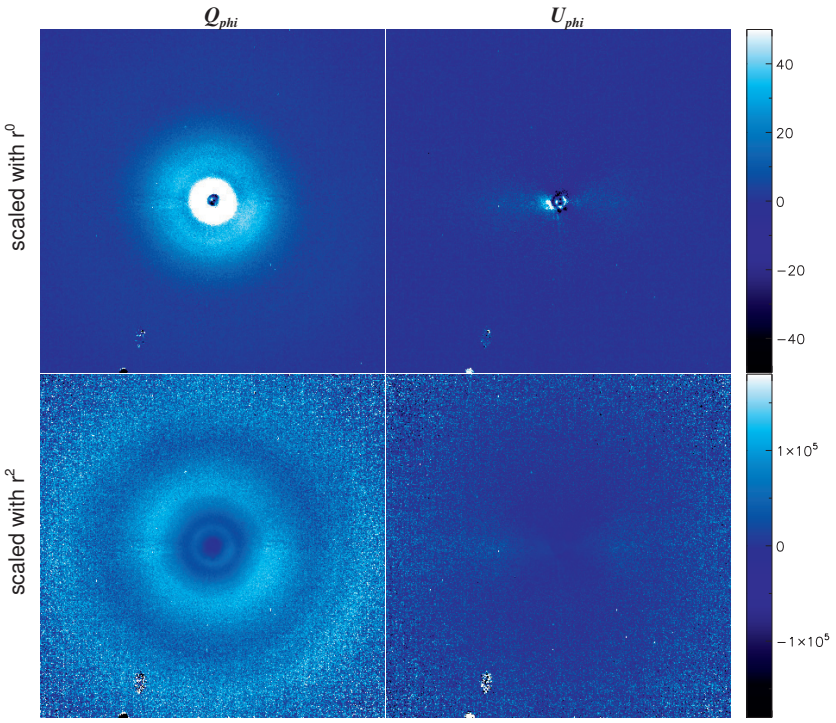

Figure 6.4: Final Q_ϕ (left) and U_ϕ (right) images of TW Hydrae. Both Q_ϕ and U_ϕ are displayed with identical linear scale, and either unscaled (top) or scaled with r^2 (bottom). All four panels are shown with North up and East to the left for the same FOV: 4.9″ × 4.9″ or up to a separation of r = 2.45″ from the star in both RA ($-x$-axis) and Dec ($+y$-axis).

but even more so when we are dealing with images of short integration times, such as resulting from individual HWP cycles. **2)** Alternatively, we can compute the azimuthal Stokes vectors (Avenhaus et al. 2014, Schmid et al. 2006):

$$Q_\phi = Q^{\rm ipc}_{\rm meas} \times \cos 2\phi + U^{\rm ipc}_{\rm meas} \times \sin 2\phi, \quad (6.3)$$

$$U_\phi = Q^{\rm ipc}_{\rm meas} \times \sin 2\phi - U^{\rm ipc}_{\rm meas} \times \cos 2\phi, \quad (6.4)$$

where ϕ describes the azimuth angle, which can be computed for each pixel (or

x, y coordinate) as

$$\phi = \arctan \frac{x - x_{\text{star}}}{y - y_{\text{star}}} + \theta_{\text{pa}}. \tag{6.5}$$

In Equation 6.5, x, y_{star} are the coordinates of the star in the image, and θ_{pa} gives an offset of the azimuth angle, which allows us to describe the temporal variation of θ_{p} (which itself varies accross the image) with one single value per HWP cycle. Polarization oriented in azimuthal direction (with respect to the position of the star) will be measured as a positive Q_ϕ signal; radial polarization will show up as a negative Q_ϕ; while polarization angles oriented at ±45° with respect to azimuthal will result in $\pm U_\phi$ signal. Disks which have high inclinations or where multiple scattering is expected to produce a significant part of the scattered light might be expected to have a significant signal in U_ϕ (Canovas et al. 2015). Fortunately, for low inclination disks, such as the TW Hydrae system, we can expect all starlight to be scattered in azimuthal direction, which means that Q_ϕ will de facto show us PI_{L}, with the the benefit that we do not take the square over the noise, resulting in cleaner images. The U_ϕ image should ideally show no signal at all, which makes it an ideal metric for the merrit of our reduction.

6.3.2 Refining the reduction by minimizing U_ϕ

To account for the variation of the measured θ_{p} between the cycles, we determine the correct value for θ_{pa} for each cycle separately, based on the assumption that the polarization is oriented in azimuthal direction, and therefore U_ϕ should be 0. We achieve this by measuring the (absolute) signal over a centered annulus in the U_ϕ image computed for a range of θ_{pa} values: $C^{u_\phi}_{\text{annulus}}(\theta_{\text{pa}})$. We select the θ_{pa} value which yields the lowest $C^{u_\phi}_{\text{annulus}}(\theta_{\text{pa}})$. Next, we improve our centering by shifting $Q^{\text{ipc}}_{\text{meas}}$ and $U^{\text{ipc}}_{\text{meas}}$ with a range of x and y steps to find the minimum $C^{u_\phi}_{\text{annulus}}(x, y)$ value. Because the improved centering will affect the minimization process with which we found θ_{pa}, we repeat the minimization of $C^{u_\phi}_{\text{annulus}}(\theta_{\text{pa}})$ on the centered data, and find θ_{pa} with increasing values between $6° \leq \theta_{\text{pa}} \leq 25°$ for the 20 HWP cycles. A final U_ϕ minimization is performed to enhance our instrumental polarization correction: we find the minimum of $C^{u_\phi}_{\text{annulus}}$ by subtracting a grid of constants C^q and C^u multiplied with I_{meas} from Q_{meas} and U_{meas}, respectively (see Equations 6.1 & 6.2).

After these U_ϕ minimizations are performed for the i^{th} cycle, $Q_{\phi,i}$ & $U_{\phi,i}$ are computed. We can derive the relative efficieny by measuring the (absolute) signal over an annulus in the Q_ϕ image for each cycle, and dividing these values by that of the highest (first) HWP cycle. We find that compared to the first cycle (set to 100%), the efficiency has declined to 62%. The final Q_ϕ & U_ϕ

images shown in Figure 6.4 are created by median combining the 20 $Q_{\phi,i}$ and $U_{\phi,i}$ images, respectively. We choose not to correct for the loss of efficiency, because that would require giving low signal to noise images a larger weight, which will enhance the noise in the final images.

6.3.3 Temporal variations in efficiency and $\theta_{\rm pa}$

Figure 6.3 shows the Stokes $I_{\rm meas}$, $Q^{\rm ipc}_{\rm meas}$ and $U^{\rm ipc}_{\rm meas}$ images for four separate derotator angles ($\theta_{\rm drot}$). We do not expect either the incident ('true') Degree of Linear Polarization ($P_{\rm L}$, see Equation 1.4) to vary or the true $\theta_{\rm pa}$ to change with respect to the meridian (north-south is fixed top-bottom on the detector). Therefore, the measured variations $\theta_{\rm pa}$ and $PI_{\rm L}$ (notice that $I_{\rm meas}$ remains roughly constant in Figure 6.3) must come from instrumental modifications of the true polarization signal, often called instrumental crosstalk. During the observations of the 20 HWP cycles, the derotator (see Figure 6.1) rotates from $\theta_{\rm drot}= -6.8°$ to $\theta_{\rm drot}= -27.4°$ ($\Delta\theta_{\rm drot}= -20.6°$). To explore whether we can explain the measured variations in $\theta_{\rm pa}$ and $PI_{\rm L}$, we calibrate the polarimetric performance with the internal light source and unpolarized standard stars in Section 6.5.

6.4 Description of instrumental effects

In Section 6.3 we have demonstrated that both the polarization angle and polarimetric efficiency measured by IRDIS are strongly dependent on the specific instrument configuration. We can express the measured Stokes vector $\mathbf{S}_{\rm meas}$ in terms of the incident Stokes vector $\mathbf{S}_{\rm in}$ as

$$\mathbf{S}_{\rm meas} = \mathbf{M}\,\mathbf{S}_{\rm in}, \tag{6.6}$$

where $\mathbf{M}$ is the 4×4 Mueller matrix describing the polarization modification of the telescope and instrument combined. The effect of the Mueller matrix is summarized by Snik & Keller (2013) as

$$\mathbf{M} = \begin{pmatrix} I \rightarrow I & Q \rightarrow I & U \rightarrow I & V \rightarrow I \\ I \rightarrow Q & Q \rightarrow Q & U \rightarrow Q & V \rightarrow Q \\ I \rightarrow U & Q \rightarrow U & U \rightarrow U & V \rightarrow U \\ I \rightarrow V & Q \rightarrow V & U \rightarrow V & V \rightarrow V \end{pmatrix}, \tag{6.7}$$

where $i \rightarrow j$ describes the contribution of the incident i into the measured j Stokes vector component. The bottom three components of the first column ($I \rightarrow [Q, U, V]$) show the conversion of unpolarized into partially polarized

light, commonly called instrumental polarization. The instrumental polarization created downstream from the HWP (except $I \rightarrow V$) is negated by the double difference (see Section 1.3.3.2). The bottom three rows of the last three columns ($[Q, U, V] \rightarrow [Q, U, V]$, without the components on the diagonal) show the crosstalk between different polarization components. Crosstalk is not necessarily negated by the double difference (for details see de Boer et al. 2014). When an optical system is comprised of n optical components and the instrumental effects on polarization depends on the specific configuration of these components, we rewrite Equation 6.6 into (see Tinbergen 1996):

$$\mathbf{S}_{\text{meas}} = \mathbf{M}_n \, \mathbf{M}_{n-1} \, \ldots \, \mathbf{M}_2 \, \mathbf{M}_1 \, \mathbf{S}_{\text{in}}, . \tag{6.8}$$

In Equation 6.8, we do not have to include every separate mirror or component independently. We are allowed to combine components which share a fixed reference frame, such as the 3 mirrors of the derotator (see Section 6.2). This allows us to create a model with Mueller matrices for only five component groups:

1. $\mathbf{M}_{\text{ut3}}$: the three mirrors of UT3,
2. $\mathbf{M}_{\text{m4}}$: M4,
3. $\mathbf{M}_{\text{hwp}}$: HWP2,
4. $\mathbf{M}_{\text{drot}}$: the derotator,
5. $\mathbf{M}_{\text{ird}}$: CPI downstream from the derotator + IRDIS.

These five groups rotate with respect to each other during an observation to account for variations in parallactic angle θ_{par} and altitude θ_{alt}. We take the rotation of the i^{th} component θ_i (as a function of θ_{par} and θ_{alt}) into account by using a rotation matrix $\mathbf{T}(\theta_i)$ according to (Tinbergen 1996):

$$\mathbf{M}_i = \mathbf{T}(-\theta_i)\mathbf{M}_{i,r}\mathbf{T}(\theta_i), \tag{6.9}$$

where $\mathbf{M}_{i,r}$ is the mueller matrix of component i within its own (fixed) frame of reference. $\mathbf{T}(\theta_i)$ rotates the Stokes vector incident on the i^{th} component toward the reference frame of this component, while $\mathbf{T}(-\theta_i)$ returns the output stokes vector into the original reference frame. The rotation matrix is described as:

$$\mathbf{T}(\theta_i) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\theta_i & \sin 2\theta_i & 0 \\ 0 & -\sin 2\theta_i & \cos 2\theta_i & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{6.10}$$

When we include the rotation matrices according to Equation 6.9 for all the subcomponents into Equation 6.8, we obtain a complete description of the VLT/SPHERE/IRDIS system:

$$\begin{aligned}\mathbf{S}_{\rm meas} =& \mathbf{M}_{\rm ird}\,\mathbf{T}(-\theta_{\rm drot})\,\mathbf{M}_{\rm drot}\,\mathbf{T}(\theta_{\rm drot})\,\mathbf{T}(-\theta_{\rm hwp})\,\mathbf{M}_{\rm hwp}\,\mathbf{T}(\theta_{\rm hwp})\\ & \mathbf{M}_{\rm m4}\,\mathbf{T}(\theta_{\rm alt})\,\mathbf{M}_{\rm ut3}\,\mathbf{T}(\theta_{\rm par})\,\mathbf{S}_{\rm in},\end{aligned} \tag{6.11}$$

which is defined in the reference frame of the P0-90 analyzer pair, which can be changed into the 'sky' reference frame by starting the sequence on the righthand side of Equation 6.11 with $\mathbf{T}(-\theta_{\rm par})\mathbf{T}(-\theta_{\rm alt})$.

6.5 Calibrating the instrumental effects

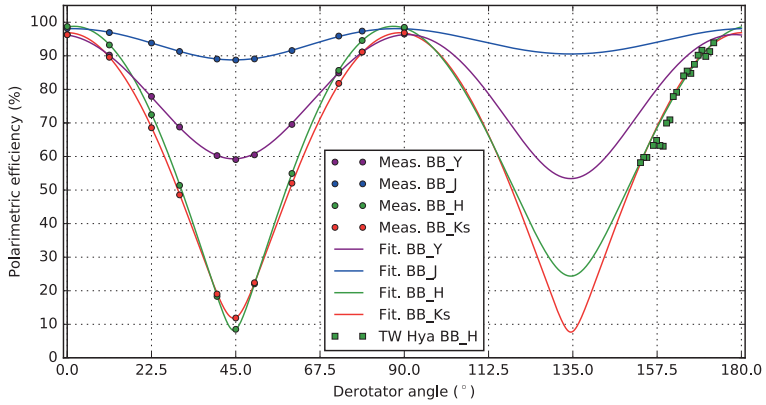

Figure 6.5: Polarimetric efficiency dependence on derotator angle $\theta_{\rm drot}$ for all broad-band filters listed in Table 6.1. The circles represent calibration measurements, the solid line is a fit to the calibration data. The green squares show the efficiency measured for TW Hydrae (Section 6.3.2) after scaling all data points such that the first HWP cycle (highest efficiency) has the same efficiency as the fit of the calibration data for the same $\theta_{\rm drot}$.

The calibrations of various instrumental effects on polarization and their analyses towards a complete Mueller matrix model of the instrument are described in detail in the Master thesis of Rob van Holstein[1]. In this section we briefly summarize the calibrations and present the major results, to the point

[1] "Accurate high-contrast imaging polarimetry of exoplanets with SPHERE/IRDIS", http://resolver.tudelft.nl/uuid:56dd5ce4-c68f-4145-a328-0d8054591ae1

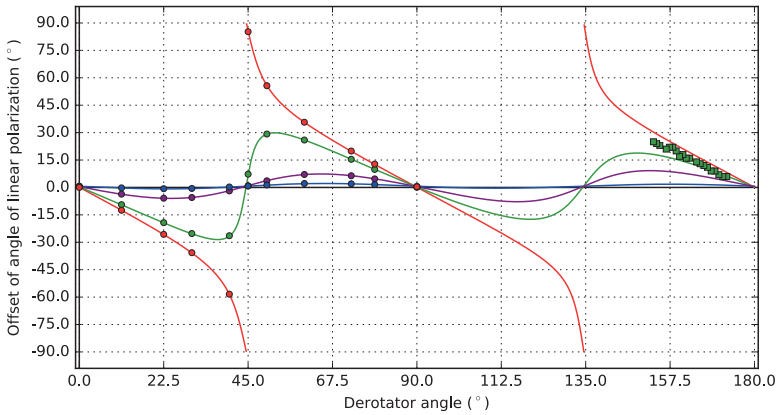

Figure 6.6: The polarization angle offset plotted against θ_{drot} for the same filters (shown with the same colors) as Figure 6.5. In the ideal case, the polarization angle offset should remain 0 for any θ_{drot}. However, there is a clear dependancy on θ_{drot} for the B_H and B_K_s filters, and to a lesser extend for B_Y, while the B_J filter remains close to ideal. The angle offset measured as θ_{pa} in the TW Hydrae dataset (B_H) is shown in green squares. The TW Hydrae (B_H) angle offset (measured as θ_{pa}) does not seem to follow the B_H calibration curve very closely. This shows that θ_{drot} cannot fully explain the polarization angle offset. Other instrumental components are likely to play an important role (e.g. θ_{hwp} is not the same for the TW Hydrae dataset as for the calibration data).

where we can explain the observed instrumental polarization effects described in Section 6.3.2 and formulate recommendations for future observations with IRDIS/DPI. Since the calibration of each component in Equation 6.11 requires knowledge of the components downstream of this component, we start with the final component $\mathbf{M}_{\text{ird}}$ and work our way up to $\mathbf{M}_{\text{ut3}}$.

For this study, it suffices to assume $\mathbf{M}_{\text{ird}}$ to produce $I \leftrightarrow Q$ instrumental (de)polarization, and only $U \leftrightarrow V$ crosstalk, because all reflection planes described by $\mathbf{M}_{\text{ird}}$ lie in the same plane as / are perpendicular to the polarization axes of the P0-90 analyzer pair. For future calibrations including the polarization effects of the ADC and/or the use of the P45-135 analyzer pair, it will be largely the $\mathbf{M}_{\text{ird}}$ component that requires a more detailed analysis, which lies outside the scope of this study. Since the P0-90 analyzers do not filter U and V polarized light any different than unpolarized light, the instrumental (de)polarization is the only relevant effect. The double difference negates this effect of $\mathbf{M}_{\text{ird}}$. Therefore, we can replace $\mathbf{M}_{\text{ird}}$ with the identity matrix.

6.5.1 The derotator and HWP2

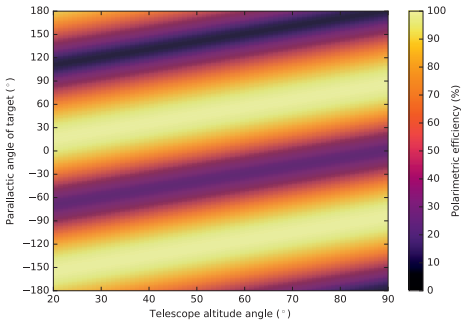

Figure 6.7: The behavior of the efficiency with $\theta_{\rm par}$ and $\theta_{\rm alt}$, when the derotator is correcting for field rotation (i.e. is in "field tracking" mode) without an additional derotator offset.

With the internal light source and calibration polarizer (herafter calibrator) inserted, we record $I_{\rm meas}$, $Q_{\rm meas}$ and $U_{\rm meas}$ for derotator angles ranging from $0° \leq \theta_{\rm drot} \leq 90°$ in the *B_Y*, *B_J*, *B_H* and *B_K*$_{\rm s}$ filters. The calibrator has its polarization axis fixed vertically with respect to the (reference) plane of the Nasmyth platform, which in turn is aligned with/perpendicular to the polarization plane of the P0-90 analyzer pair. Figure 6.6 shows how $\theta_{\rm p}$ oscillates around 0 (which should be the value in the ideal case of no crosstalk). While this oscillation is only marginally visible for *B_J*, with a maximum deviation from ideal < 3°; it is < 7° in *B_Y*, and can reach up to 30° in *B_H*. For *B_K*$_{\rm s}$, $\theta_{\rm p}$ does not even return to the ideal angle offset (= 0°) and continues rotating beyond ±90° (where a rotation of +90° is indistinguishable from −90°).

When we use the ∼ 100% polarized light of the calibrator, we determine the polarimetric efficiency curves by computing the Degree of Linear Polarization $P_{\rm L}(\theta_{\rm drot})$. The efficiency curves are shown for the four broad-band filters of IRDIS in Figure 6.5. A dramatic decrease in efficiency is seen for $\theta_{\rm drot} \sim 45°$ (e.g. ∼ 10% in *B_H*) and 135° (e.g. ∼ 20% in *B_H*). The asymmetry between these two derotator angles is caused by a non-ideal behavior of HWP2, i.e. the retardance $\neq \lambda/2$. We calibrated the wavelength dependence of the retardance of HWP2 by observing the $\theta_{\rm drot}$ dependence of the efficiency for different offsets to the HWP2 switch angles $\theta^{\rm s}_{\rm hwp}$ (see the thesis of van Holstein). Also the *B_K*$_{\rm s}$ filter (efficiency ≥ 12%) shows a strongly varying performance like *B_H*, while *B_Y* (≥ 59%) and especially the *B_J* filter (≥ 90%) show a much better polarimetric performance at $\delta_{\rm drot} = 45°$ and 135°. The low efficiencies in the *B_H* and *B_K*$_{\rm s}$ filters are caused by the retardance of the derotator, which is close to that of a quarter-wave ($\lambda/4$)

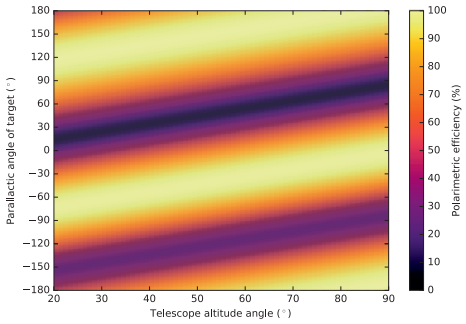

Figure 6.8: Polarimetric efficiency plotted against $\theta_{\rm par}$ and $\theta_{\rm alt}$ for the derotator in field tracking mode with a derotator offset of 45° (INS.CPRT.POSANG = 90°).

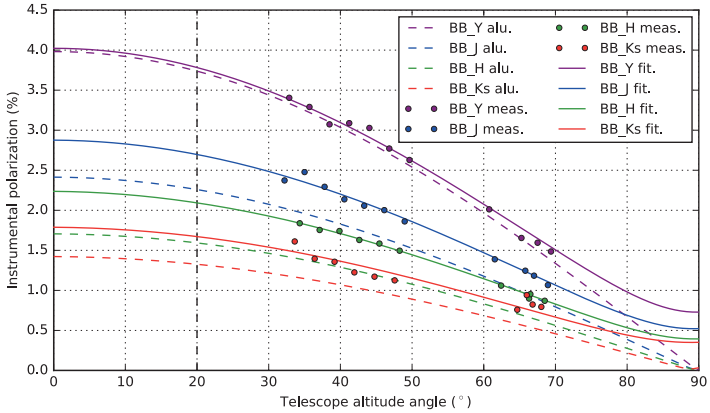

Figure 6.9

plate at these wavelengths. With these retardances, the derotator causes a strong $[Q, U] \rightarrow V$ crosstalk in B_H and B_K_s.

An efficiency $\geq$ 70% is achieved for all filters as long as the derotator stays within $-22.5° < \delta_{\rm drot} < 22.5° + n \times 90°$, with $n \in \mathbb{Z}$. This constraint on $\theta_{\rm drot}$ can be achieved with two simple steps: **1**) split the total observation block into templates where the difference between $\theta_{\rm par}$ and $\theta_{\rm alt}$ does not vary by more than $4 \times 22.5 = 90°$ (note: $\theta_{\rm drot} \propto (\theta_{\rm alt} - \theta_{\rm par})/2$); **2**) within ESO's Phase 2 Proposal Preparation (P2PP), provide each template of the Observation Block (OB) with a derotator offset (keyword: INS.CPRT.POSANG, which gives the offset in position angles, which is $2 \times \Delta\theta_{\rm drot}$) equal to the mean difference

between the values of $\theta_{\rm par}$ and $\theta_{\rm alt}$:

$$\text{INS.CPRT.POSANG} = < \theta_{\rm par} - \theta_{\rm alt} > + n \times 180^\circ, \tag{6.12}$$

with $n \in \mathbb{Z}$. When the previous two steps are performed correctly, the mean $\theta_{\rm drot}$ lies at either 0° or 90°, which orients the derotator horizontal or vertical with respect to the Nasmyth platform.

Figure 6.7 shows the efficiency plotted against $\theta_{\rm par}$ and $\theta_{\rm alt}$ without the application of an offset to $\theta_{\rm drot}$. Observing a star at for example $< \theta_{\rm par} >= 160, < \theta_{\rm alt} >= 70$ will give us an efficiency of only $\sim 10\%$. When we provide a derotator position angle offset of INS.CPRT.POSANG $= 160 - 70 = 90^\circ$, we get $\theta_{\rm drot} = 45^\circ$. The effect of a 45° derotator offset is shown in Figure 6.8: efficiency $\approx 100\%$.

6.5.2 M4 and UT3

Now we know how a Stokes vector incident on HWP2 will be modified downstream by the instrument, we can calibrate the two final components: $\mathbf{M}_{\rm m4}$ and $\mathbf{M}_{\rm ut3}$. When we take the azimuthal symmetry of M1 and M2 into account, we can simplify $\mathbf{M}_{\rm ut3}$ into $\mathbf{M}_{\rm m3}$, the Mueller matrix description of the 45° incidence reflection of the fold mirror M3.

We observe an unpolarized standard star for a range of altitude angles. After correcting the data with the model for the components downstream from M4, we determine the instrumental polarization created by the combination of M3 and M4 (which cannot be removed with the double difference). We fit the instrumental polarization contribution of each mirror to the calibrations and compare this with the curve expected for a perfect aluminum mirror (without dust, which is clearly not true) in Figure 6.9. As expected, the instrumental polarization is maximal when the reflection planes of both M3 and M4 are aligned (at $\theta_{\rm alt} = 0^\circ$). Due to the similarity of the coating, we would ideally expect (as we see for the dashed lines) the instrumental polarization to be fully canceled in a crossed configuration (at $\theta_{\rm alt} = 90^\circ$). That the two mirrors do not cancel each other at $\theta_{\rm alt} = 90^\circ$ indicates a signifficant difference between the instrumental polarization induced by either mirror, possibly caused by a larger dust layer on one of the mirrors.

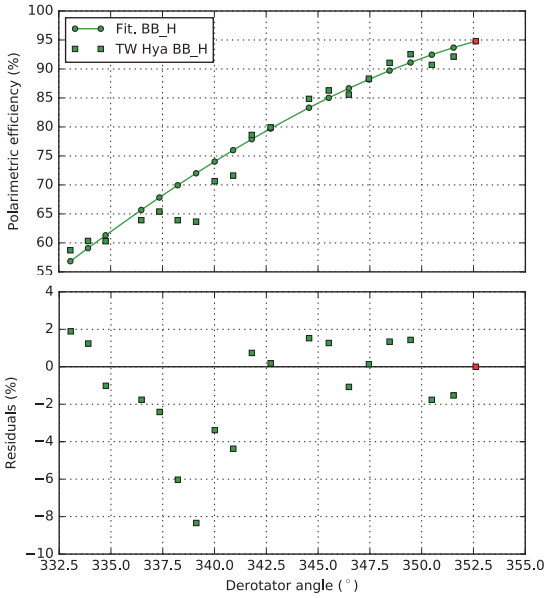

Figure 6.10: Top: Polarimetric efficiency modeled (solid line) for the same derotator and HWP2 angles as used during the observation of TW Hydrae (squares). Note that the efficiencies measured for the HWP cycles of TW Hydrae are only determined relative to the other HWP cycles. We therefore scaled all data points such that the first cycle (red square, with the highest efficiency) matches the value of the model. **Bottom:** Residuals between the model and the efficiencies obtained for TW Hydrae.

6.5.3 Final model

The Mueller matrix of Equation 6.7 for an individual optical component can be written as a function of the component's diattenuation (ϵ) and retardance (Δ):

$$\mathbf{M} = \frac{1}{2}\begin{pmatrix} 1+\epsilon & 1-\epsilon & 0 & 0 \\ 1-\epsilon & 1+\epsilon & 0 & 0 \\ 0 & 0 & 2\sqrt{\epsilon}\cos\Delta & 2\sqrt{\epsilon}\sin\Delta \\ 0 & 0 & -2\sqrt{\epsilon}\sin\Delta & 2\sqrt{\epsilon}\cos\Delta \end{pmatrix}. \tag{6.13}$$

Diattenuation creates instrumental polarization in the $+Q$ direction when $\epsilon > 1$, in the $-Q$ direction when $\epsilon < 1$ and no instrumental polarization when $\epsilon = 1$. Note that the final Mueller matrix of the complete system has to be divided by I_{meas} to normalize it.

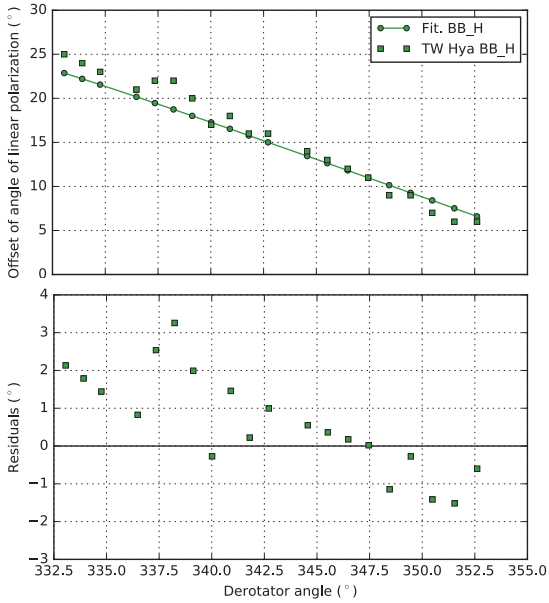

Figure 6.11: Top: Polarimetric angle offsets of the same model (solid line) as used in Figure 6.10; and the offset angles ($\theta_{\rm pa}$ of Equation 6.5) of TW Hydrae (squares) with respect to azimuthal polarization angles. **Bottom:** Residuals between the model offset angles and $\theta_{\rm pa}$ retrieved for TW Hydrae.

As described in detail in the Master thesis of Rob van Holstein, we computed the model by performing a non-linear least squares fit to the calibration datasets. The final model parameters are listed in Table 6.2. The retardances of M3 and M4 and not calibrated but computed using the Fresnel equations with the wavelength dependent complex refractive indices for aluminum (Rakic et al. 1998).

6.5.3.1 Explaining the TW Hydrae data

Figures 6.5 and 6.6 show that although the polarimetric efficiency is rather well explained with the B_H calibration curve, the polarization angle offset $\theta_{\rm pa}$ deviates from the calibration curve. This shows that the derotator is not the only component which affects the polarization angle offset.

For the same $\theta_{\rm drot}$ and $\theta_{\rm hwp}$ as the instrument configuration used for TW Hydrae, we determine the model including the retardances of the derotator

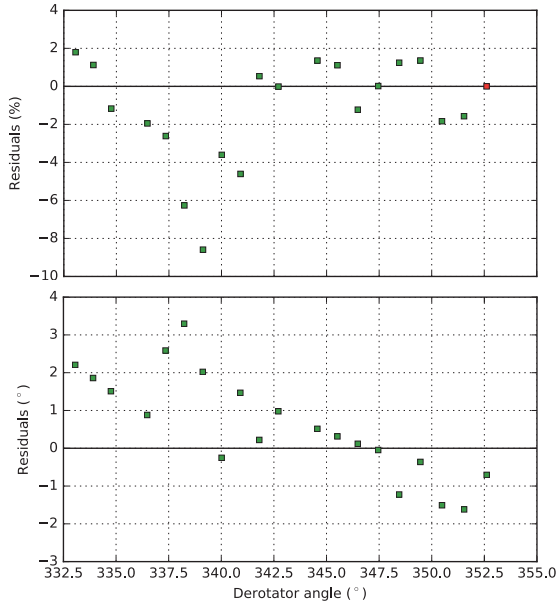

Figure 6.12: Top: Residual between the model and TW Hydrae polarimetric efficiencies, including the retardances of M3 and M4. **Bottom:** Residuals between the offset angles of the same model as the top panel, and θ_{pa} retrieved for TW Hydrae.

and HWP2. Because we cannot detemine the absolute degree of polarization in the TW Hydrae images, we do not include the diattenuation in the model (which is computed for $\sim 100\%$ incident polarization). We compare the model efficiencies and offset angles with on-sky observation efficiencies in Figure 6.10 and angle offsets in Figure 6.11. Since we do not know the absolute values of the polarimetric efficiencies for the TW Hydrae measurements, we scale the efficiencies for TW Hydrae to have the maximum value (first HWP cycle, indicated with a red square in Figure 6.10) match the modeled efficiencies in B_H.

We supplement the model by including the retardances of M3 and M4 and plot the residuals for the efficiencies and offset angles in Figure 6.12. This inclusion of M3 and M4 only marginally improved the residuals between model and on-sky observations. We have some clear outliers in both angle and efficiency at $\theta_{\mathrm{drot}} \approx 139°$, which are most likely caused by a poor fit of θ_{pa} for these HWP cycles. More interestingly, we can detect a clear trend in the residuals of the angle offsets decreasing with θ_{drot}, which does not seem to appear in the

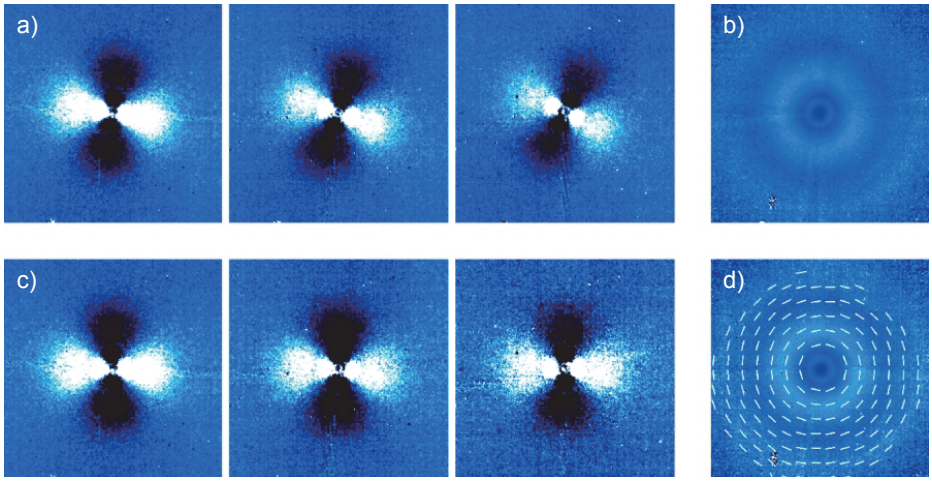

Figure 6.13: a shows $Q_{\rm meas}$ for three HWP cycles, with $\theta_{\rm drot} = 349.7°, 342.0°$, and $334.1°$. We can clearly see the buttefly patterns rotate in clockwise direction and the signal decrease from left to right. When we use the assumption of azimuthal polarization, we can find the correct $\theta_{\rm pa}$, for each HWP cycle and compute a final Q_ϕ image, shown in **b**. Because we have used the assumption that we know the polarization angle at each point, we cannot claim to have determined $\theta_{\rm p}$ in this image. When we apply the inverse of our Mueller matrix model to the $Q_{\rm meas}$ (**c**) and $U_{\rm meas}$ images, we retrieve the ideal orientation of the butterflies and can correct for the reduced efficiency. However, the latter does not improve the signal to noise (see right-most panel of **c**). The model corrected Q_ϕ image is shown in panel **d**. Because we no longer need the assumption that we know $\theta_{\rm p}$ a priori, we can use $Q_{\rm meas}$ and $U_{\rm meas}$ to determine $\theta_{\rm p}$ in our final reduction (the angle is indicated by the white lines in panel d).

efficiency residuals. Because we exclude the diattenuation values from these models, we have only taken the instrumental crosstalk into account and ignored the telescope- and instrumental polarization. Instrumental polarization is likely to create an additional offset in the polarization angle, while having only a minor effect on the relative efficiencies.

6.5.4 Correcting observations with the instrument model

We can largely explain the observed instrumental effects on the incident polarization by including only the retardances of the derotator and HWP2, but are

Parameter		*B_Y*	*B_J*	*B_H*	*B_K$_s$*
$\epsilon_{\rm hwp}$		$1.00046 \pm 3 \cdot 10^{-5}$	$1.000851 \pm 9 \cdot 10^{-6}$	$1.00055 \pm 1 \cdot 10^{-5}$	$1.00082 \pm 2 \cdot 10^{-5}$
$\Delta_{\rm hwp}$	(°)	184.2 ± 0.2	182.5 ± 0.2	170.5 ± 0.1	177.5 ± 0.1
$\delta_{\rm hwp}$	(°)	$0.18589 \pm 5 \cdot 10^{-5}$	$0.18589 \pm 5 \cdot 10^{-5}$	$0.18589 \pm 5 \cdot 10^{-5}$	$0.18589 \pm 5 \cdot 10^{-5}$
$\epsilon_{\rm drot}$		$1.00182 \pm 4 \cdot 10^{-5}$	$1.01658 \pm 1 \cdot 10^{-5}$	$1.00453 \pm 1 \cdot 10^{-5}$	$0.99302 \pm 1 \cdot 10^{-5}$
$\Delta_{\rm drot}$	(°)	126.1 ± 0.1	203.9 ± 0.1	99.39 ± 0.06	84.17 ± 0.05
$\delta_{\rm drot}$	(°)	$0.53088 \pm 5 \cdot 10^{-5}$	$0.53088 \pm 5 \cdot 10^{-5}$	$0.53088 \pm 5 \cdot 10^{-5}$	$0.53088 \pm 5 \cdot 10^{-5}$
$\epsilon_{\rm m3}$		$0.9526 \pm 4 \cdot 10^{-4}$	$0.9662 \pm 3 \cdot 10^{-4}$	$0.9738 \pm 2 \cdot 10^{-4}$	$0.9785 \pm 6 \cdot 10^{-4}$
$\Delta_{\rm m3}$	(°)	188.1	186.6	185.0	183.7
$\epsilon_{\rm m4}$		$0.9666 \pm 4 \cdot 10^{-4}$	$0.9761 \pm 3 \cdot 10^{-4}$	$0.9813 \pm 2 \cdot 10^{-4}$	$0.9851 \pm 6 \cdot 10^{-4}$
$\Delta_{\rm m4}$	(°)	188.1	186.6	185.0	183.7

Table 6.2: Model parameters diattenuations (ϵ), retardance (Δ), and zero-angle offset (δ) for the four broad-band filters of IRDIS. Note that the retardances for M3 and M4 are based on the chromatic values for the complex refractive index of aluminum (from refractiveindex.info). The retardances are expressed in degrees with $180° = \lambda/2$.

left with lower order effects caused by instrumental polarization. However, the most useful application of our model will be the correction (rather than explanation) of observational data. Contrary to the explanation given in Section 6.5.3.1, in our correction we can include the diattenuation and therefore account for the telescope and instrumental polarization. However, as a first test of the method, we have used the instrument model including only the retardances of all optical components. How the correction method is derived from the model is described in detail in the thesis of van Holstein, and summarized below.

In our correction we use the reference frame of the analyzer pair (not the 'sky'), and therefore consider each double difference measurement (Q and U in 'sky') to be a $Q_{\rm meas}$ measurement in this reference frame, irrespective of $\theta_{\rm hwp}$. We choose this reference frame, because the properties (rotation, retardance and diattenuation) of HWP2 (which creates the switch between $Q_{\rm meas}$ and $U_{\rm meas}$ in the sky reference frame) are either variables or parameters of our model, while the analyzer pair is fixed to transmit $(I_{\rm meas} \pm Q_{\rm meas})/2$. Using this convention, observing with n polarimetric cycles, we obtain $2n$ pairs of $I_{\rm meas}$ and $Q_{\rm meas}$ measurements. We list these measurement for each new pair i in the array

$$\mathbf{Y} = [I_{\rm meas,1}, Q_{\rm meas,1}, \ldots I_{\rm meas,\mathit{i}}, Q_{\rm meas,\mathit{i}}, \ldots I_{\rm meas,4\mathit{n}}, Q_{\rm meas,4\mathit{n}}]^{\rm T},$$

a $4n$ column vector. These observations are related to the true Stokes vector $\mathbf{S}_{\rm in}$ as:

$$\mathbf{Y} = \mathbf{X}\mathbf{S}_{\rm in}, \tag{6.14}$$

where $\mathbf{X}$ is a $4 \times 4n$ matrix describing the instrumental effects on the incident polarization, as defined by the model parameters listed in Table 6.2 and the specific rotation angles of the optical components.

To derive the corrected stokes vector $\hat{\mathbf{S}}_{\rm in}$ given $\mathbf{X}$ and $\mathbf{Y}$ we invert Equation 6.14 with the linear least-squares solution:

$$\hat{\mathbf{S}}_{\rm in} = (\mathbf{X}^{\rm T}\mathbf{X})^{-1}\mathbf{X}^{\rm T}\mathbf{Y}. \tag{6.15}$$

We have applied Equation 6.15 to the dataset of TW Hydrae. For illustrative purposes, we have also applied Equation 6.15 to three individual polarimetric cycles with $\theta_{\rm drot} = 349.7°, 342.0°$, and $334.1°$ of this dataset. Figure 6.13 shows the original images and the result of these corrections. The $Q_{\rm meas}$ images (in the sky reference frame) of the three HWP cycles are shown in panel a; the final Q_ϕ image of the reduction described in Section 6.3.2 is shown in panel b; the corrected $Q_{\rm meas}$ of panel a are shown as $\hat{Q}_{\rm in}$ images in panel c; and the final corrected Q_ϕ is shown in panel d.

While the original $Q_{\rm meas}$ of Figure 6.13a show the rotation ($\theta_{\rm pa}$) of the butterfly caused by crosstalk; the corrected $\hat{Q}_{\rm in}$ are clearly oriented such that $\theta_{\rm pa} \sim 0°$. Although the $\hat{Q}_{\rm in}$ display a surface brightness which is approximately the same for all three $\theta_{\rm drot}$, the decrease in efficiency between the three $Q_{\rm meas}$ images (going from left to right in panel a) is still visible as a decrease of the signal to noise ratio (SNR). The azimuthal direction of the true polarization angle was used as an assumption in our reduction of Q_ϕ in Section 6.3.2. Therefore, we cannot claim to have derived $\theta_{\rm p}$ in panel b. Since we do not need to assume a-priori knowledge about $\theta_{\rm p}$ to compute the final Q_ϕ image of panel d, we can confidently claim to have determined $\theta_{\rm p}$, which is therefore indicated with white bars in the image.

6.6 Conclusions and recommendations

SPHERE/IRDIS/DPI is currently among the most powerful instruments (modes) to perform polarimetric high-contrast imaging. Due to design choices, its performance is strongly dependent on the observation strategy, as we illustrated with the observations of TW Hydrae. We characterize the instrument and find that polarimetric crosstalk can cause the efficiency to drop towards $\sim$ 10% in the B_H and B_K_s filters; efficiency remains above $\sim$ 60% in B_Y and well above 90% in B_J. Crosstalk also causes an offset in the measured polarization angle up to 25° for the B_H and more in B_K_s. We have demonstrated that our Mueller matrix model can be used to explain and correct for the variations in efficiency and polarization angle offset due to crosstalk observed in the TW Hydrae data.

Optimal results can be optained from IRDIS/DPI observations when two important considerations are taken into account: 1) adjust the observation strategy beforehand to minimize a loss in efficiency; 2) apply the Mueller matrix to correct the data for instrumental (and telescope) polarization and crosstalk. The observation strategy should be adjusted accordingly:

- When no strict wavelength requirements are present, use IRDIS/DPI in J-band to achieve a $> 90\%$ polarimetric efficiency which is nearly independent of the remaining instrumental setup.
- If the previous recommendation cannot be met, and the filters B_Y, B_H, or B_K_s are used, avoid the use of derotator angles $22.5 \lesssim |\theta_{\rm drot}| \lesssim 67.5^\circ + n \times 90^\circ$, with $n \in \mathbb{Z}$.
- For short observing blocks with little field rotation, this $\theta_{\rm drot}$ constraint can be achieved by determining the average $\theta_{\rm par}$ and $\theta_{\rm alt}$ and including a derotator (posang) offset of:

  $$\text{INS.CPRT.POSANG} = < \theta_{\rm par} - \theta_{\rm alt} > + n \times 180^\circ,$$

  with $n \in \mathbb{Z}$.
- When the observing block is so long that $\Delta\theta_{\rm drot} > 45^\circ$, we recommend to split up the observing block in multiple templates, and compute the derotator offset for each template separately. The latter can most accurately be performed during observing runs in Visitor Mode at the VLT, since the required offset will be strongly dependent on the exact start of the observation template.
- Alternatively, a constraint on the Local Siderial Time (LST) around a time when the parallactic angle does not change too fast (i.e. far from the meridian) will enable the user to predict the required derotator angle offset more accurately (depending on the length of the observing block), which allows for (remote) Service Mode observations.