It's just a phase: high-contrast imaging with patterned liquid-crystal phase plates to facilitate characterization of exoplanets

Doelman, D.S.

6 Design and Performance Analysis of a prototype Holographic Aperture Mask for the Keck OSIRIS Imager

Adapted from
D. S. Doelman , J. P. Wardenier, P. B. Tuthill, M. P. Fitzgerald, J. Lyke, S. Sallum, B. Norris, N. Z. Warriner, P. Wizinowich, C.U. Keller, M. J. Escuti, F. Snik
Accepted by Astronomy and Astrophysics

Context As an interferometric technique, sparse aperture masking (SAM) is capable of imaging beyond the diffraction limit of single telescopes. This makes SAM an important technique to study processes like planet formation at solar-system scales. However, it comes at the cost of a reduction in throughput, typically by 80-90%.
Aims We report on the design, construction and commissioning of a prototype aperture masking technology implemented at the Keck OSIRIS Imager: the holographic aperture mask. Holographic aperture masking (HAM) aims at increasing the throughput of sparse aperture masking by selectively combining all subapertures across a telescope pupil in multiple interferograms using a phase mask, in addition to adding low-resolution spectroscopic capabilities.
Methods Using liquid-crystal geometric phase patterns we manufacture a HAM mask that uses an 11-hole SAM design as central component and a holographic component that compromises 19 different subapertures. Thanks to a multi-layer liquid-crystal implementation, the mask has a diffraction efficiency higher than 96% from 1.1 to 2.5 micron. We create a pipeline that extracts monochromatic closure phases from the central component as well as multi-wavelength closure phases from the holographic component. We test the performance of the HAM mask in the lab and on-sky.
Results The holographic component yields 26 closure phases with a spectral resolutions between R$\sim$6.5 to R$\sim$15, depending on the interferogram positions. On April 19 2019, we observed the binary star HDS 1507 in the H_{bb} filter ($\lambda_0 = 1638$ nm and $\Delta\lambda = 330$ nm) and retrieve a constant separation of 120.9 $\pm$0.5 mas for the independent wavelength bins, which is excellent agreement with literature values. For both the lab measurements and the observations of unresolved reference stars we record non-zero closure phases; a potential source of systematic error that we traced to polarization leakage of the HAM optic. We propose a future upgrade that improves the performance, reducing this effect to an acceptable level.
Conclusions Holographic aperture masking is a simple upgrade of SAM with increased throughput and a new capability of simultaneous low-resolution spectroscopy which provides new differential observables (e.g. differential phase with wavelength).

6.1 Introduction

Many of the most critical aspects of stellar physics play out in a theatre at solar-system scales. These include star and planet formation, mass loss and debris disks to name only a few. High-fidelity imaging of circumstellar environments can provide key insights in these processes. High-contrast imaging instruments with adaptive optics (AO) provide high resolution imagery with great sensitivity, resolving many protoplanetary disks and sub-stellar companions (Chilcote et al., 2018; Beuzit, J.-L. et al., 2019). However, the performance of high-contrast imaging systems is still limited by residual phase and non-common path aberrations, reducing the sensitivity (Macintosh et al., 2019; Beuzit, J.-L. et al., 2019). Nonetheless, extreme AO facilities have been able to reach high contrasts (>14 magnitudes) down to ∼200 mas: a few times the diffraction limit in the near-infrared (Vigan et al., 2015; Nielsen et al., 2019). For nearby populous star forming regions such as Taurus, 200 mas corresponds to ~ 30 AU (larger than the orbit of Jupiter and Saturn), which leaves a blind spot for critical scales of disk evolution and planet formation. An additional technique called sparse aperture masking (SAM), often used in concert with AO, has been able to resolve finer structures beyond the diffraction limit, e.g. 20 mas at 1.65 μm (Tuthill et al., 1999).

SAM works by turning a telescope aperture into an interferometric array using an opaque mask with small holes (Haniff et al., 1987; Tuthill et al., 2000). For most applications, the holes are placed in a non-redundant fashion, which means that each baseline (the vector that connects two apertures) appears only once. Imaging with such a mask results in an interferogram that contains many fringes with unique spatial frequencies in the image plane. The first null of these fringes is at $0.5\lambda/B$ instead of $1.22\lambda/D$, where λ is the wavelength, B the longest baseline and D the telescope aperture diameter. A second profound advantage is the rejection of phase noise. Non-redundancy acts to remove noise in both visibility amplitudes and phase measurements, and in particular robust observables known as closure phases have been exploited with great success. Closure phases are formed by taking the sum of phases around baselines forming a closed triangle of subapertures in the pupil. Even before adaptive optics (AO) became well established, the robust observables delivered by SAM allowed for imaging the regions closest to stars (Tuthill et al., 1999).

Sparse aperture masking is even more powerful when used in concert with AO, providing extremely stable closure phases that resulted in a leap in sensitivity and contrast (Tuthill et al., 2006). The superior calibration particularly of closure phases makes SAM more sensitive than coronagraphs for probing the smallest separations ($< 1 - 2\lambda/D$), e.g. Gauchet et al. (2016); Cheetham et al. (2016); Samland, M. et al. (2017). Furthermore, SAM has established productivity in directly resolving stellar environments, recovering dust shells (Haubois, X. et al., 2019) and structures in protoplanetary disks (Kraus et al., 2008; Willson et al., 2019) at solar system scales.

Current SAMs, such as those in Keck/NIRC2 and VLT/SPHERE, block $80 - 90\%$ of the incident light, improving resolution at the expense of throughput and Fourier

coverage. Several different approaches to improve this throughput have been proposed. One approach is the kernel phase method introduced by Martinache (2010). Here, the concept of closure phase is generalised to the so-called kernel phase, which are also phase-noise resistant observable quantities. The advantage of this formalism over traditional closure phases is the fact that it can also be applied to redundant pupils (placing it within the lineage of development from Speckle Interferometry), provided that phase aberrations are small. In the high-Strehl regime, the Kernel phase method presents a huge increase in throughput compared to sparse aperture masking, although the achieved contrast is comparable to non-redundant masking (Pope et al., 2015; Sallum & Skemer, 2019). Another advantage of Kernel phase is that a full aperture yields more independent Kernel phases, allowing for a better characterization of the target. Lastly, it easy to implement in any extreme-AO instrument, as it requires only regular imaging and a different post-processing algorithm.

A different approach that aims to increase the throughput of aperture masks is segment tilting interferometry (Monnier et al., 2009). With this approach subsets of mirrors are imaged onto separate image plane locations, resulting in a few completely independent non-redundant fringe patterns. While an on-sky survey using this approach with the Keck 1 telescope proved successful, determining the sizes of many young-stellar objects, the experiment was stopped shortly after with the decommissioning of the Long Wavelength Spectrometer instrument. The simple concept is appealing, it has a huge increase in throughput, superior uv-coverage and the possibility of a photometric channel. However, the technical aspect is challenging with strict requirements on keeping the tip-tilt alignment and piston offset of individual mirror segments (Tuthill, 2012). Even so, there are many opportunities to revive this approach, as many future telescopes having a segmented primary mirror, e.g. JWST (Gardner et al., 2006), the TMT (Sanders, 2013), and the ELT (Gilmozzi & Spyromilio, 2007).

A third approach is pupil remapping interferometry, which uses photonic techniques to remap the pupil onto a non-redundant output pupil, such that the full pupil can be used (Perrin et al., 2006). Different photonic techniques have been implemented to remap the pupil, e.g. single-mode fibers (Huby et al., 2012, 2013) and integrated photonic chips (Jovanovic et al., 2012). The photonic chips are especially powerful, as a single integrated chip can contain multiple components that split single subaperture waveguides into multiple channels, control the phase delay of every channel and recombine all subaperture waveguides, interfering every subaperture with all others (Cvetojevic et al., 2018; Norris et al., 2019). Both photonic techniques perform spatial filtering, are compact and their monolithic design promises superior stability. Moreover, pupil remapping lends itself for spectroscopy, as the output can be arranged as 1D arrays of subapertures, which are easily dispersed. Overall, pupil remapping has some unique advantages compared to sparse aperture masking at the cost of increased complexity, as pupil remapping requires a separate instrument.

Sparse aperture masking is an elegant technique due to its simplicity, i.e. a single mask in the pupil plane adds the capability of imaging beyond the diffraction limit.

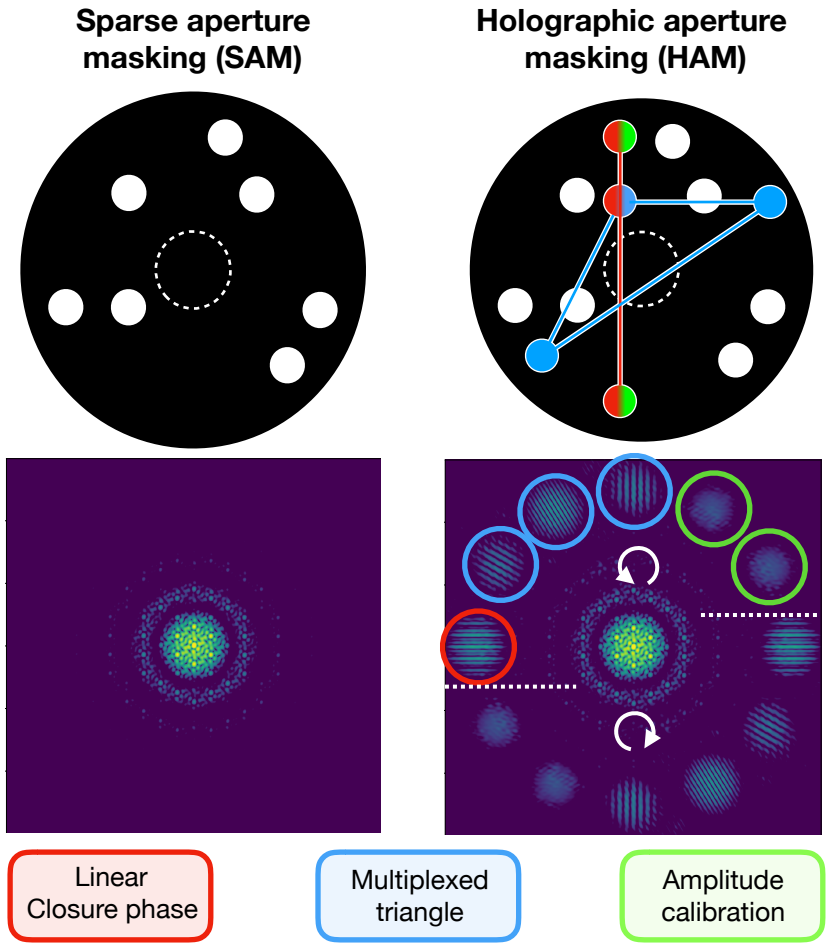

Figure 6.1: Cartoon of sparse aperture masking (left) and holographic aperture masking (right). The masks are shown on top with their resulting PSFs at the bottom. Holographic aperture masking (HAM) combines an aperture mask with a phase plate. HAM adds baselines by imaging otherwise redundant subapertures at separate locations in the focal plane. The SAM mask design is incorporated as the central component and the off-axis interferograms are part of the holographic component. Subapertures can be imaged onto multiple PSF locations to make different combinations. The different colors shown for HAM indicate different ways of combining subapertures and correspond to the same color coded off-axis interferograms. Multiple subapertures in one dimension can be combined at the same location (red) or in two dimensions by imaging single baselines at different locations (blue). It is also possible to make a non-interferometric PSF of a single subaperture for amplitude monitoring (green). Two copies of each PSF are created with opposite circular polarization state.

Yet, the approaches that have an increased throughput compared to SAM present demanding system requirements, requiring high Strehl (Kernel phase), a complete overhaul in the primary mirror alignment (segment tilting) or even a completely new instrument (pupil remapping).
In this paper we present the implementation of a novel approach, holographic aperture masking (HAM), that improves the throughput, number of baselines and closure phases, adds the capability of low-resolution spectroscopy, while also maintaining the simplicity of a single pupil-plane optic. The fundamental idea of HAM is equivalent to segment tilting, where interfering different subapertures on separate locations in the pupil allows for increasing the throughput without creating redundant baselines (Doelman et al., 2018). Instead of tilted mirrors, the phase mask introduces achromatic phase tilts to shift the location where subapertures are imaged onto the detector to form an interferogram. As developed in this paper, HAM is implemented as an addition to a non-redundant SAM mask, with the distinction between the two components of the hybrid experiment discussed as the SAM *central component*, and the off-axis HAM interferograms, the *holographic component*. Both components provide complementary information and can be used independently of each other. The difference between SAM and HAM is shown in Fig. 6.1. One unique advantage of the phase masks is that single subapertures can be imaged onto multiple locations. This is done by holographically combining these phase tilts through multiplexing. For example, the three blue subapertures in Fig. 6.1 are all imaged onto two holograms, such that all baselines are measured and a closure phase can be extracted. Moreover, the achromatic phase tilts make the HAM mask diffractive, i.e. the diffraction angle depends on wavelength. When the one dimensional subaperture combinations are deflected to a point in the image plane orthogonal to their baseline, the wavelength smearing does not act to blur out the fringes. Therefore, the holographic component can be designed to operate as a low-resolution spectrograph without an additional dispersing element.
We address the implementation and design of a HAM mask in Sect. 6.2, the influence of polarization leakage in Sect. 6.3, the HAM mask design for the Keck OSIRIS Imager in Sect. 6.4, the data reduction pipeline in Sect. 6.5, the manufacturing of two HAM masks for the Keck OSIRIS Imager in Sect. 6.6, the lab verification of one mask in Sect. 6.7, and the on-sky verification with a binary in Sect. 6.8.

6.2 HAM design

The HAM concept shows promise to improve upon SAM mask designs adding throughput, Fourier coverage and wavelength diversity. Here we explore the implementation of the HAM phase mask, and how the properties of the phase mask influence the design. A critical property of the phase mask is that it needs to be able to image subapertures in off-axis interferograms. The off-axis interferograms are rather large with size scaling with λ/D_{sub}, where λ is the wavelength and

D_{sub} is the diameter of the subaperture. Therefore, imaging multiple interferograms onto separate locations (so as to avoid overlap) on the detector requires large phase tilts. This makes it difficult to manufacture classical phase implementations of a HAM phase mask for transmissive pupil planes. A solution is offered by liquid-crystal geometric phase holograms (GPHs), as they have unbounded continuous phase over a large wavelength range. Next, we will explore the properties of these GPHs.

6.2.1 Geometric phase holograms

A geometric phase hologram (Escuti et al., 2016) is a phase plate that applies geometric phase (or Pancharatnam-Berry phase), rather than classical phase that arises from optical path differences. Geometric phase arises when light travels through an anisotropic medium, and only depends on the geometry of the light path through the medium. One type of GPHs are half-wave retarders with varying fast-axis orientation, and they apply geometric phase to circularly polarized light. The geometric phase, $\phi(x, y)$, of such a device only depends on the geometry of the fast-axis orientation $\theta(x, y)$, i.e.

$$\phi(x, y) = \pm 2\theta(x, y), \tag{6.1}$$

where the sign is determined by the handedness of the circular polarization state of the incoming light and x, y indicate the pupil-plane coordinates (Escuti et al., 2016). Unpolarized light is defined as having no preferred state of polarization and contains on average equal amounts of left and right circular polarization. The fraction of light that acquires geometric phase depends on the retardance, where half-wave retardance yields close to 100% diffraction efficiency. The fraction of light that does not acquire geometric phase is called the leakage term and is apart from a global piston term unaffected by the optic. The impact of this leakage term for HAM is explored in Section 6.3.

A GPH phase mask for HAM requires control of the fast-axis orientation and tuning of the retardance for optimal efficiency over the desired bandwidth. Both of these properties can be controlled to a very high degree with liquid-crystal technology. With a direct-write method, almost arbitrary phase patterns can be written in a photo-alignment layer (PAL) (Miskiewicz & Escuti, 2014). Birefringent liquid-crystal layers deposited on the PAL keep this orientation pattern due to spontaneous self-alignment. Changing the retardance is possible by stacking these layers with different thickness and twist into a monolithic film (Komanduri et al., 2013). By tuning these parameters, multi-twist retarders are capable of achieving high diffraction effiencies over large bandwidths. The layers are cured with UV radiation, and the liquid-crystal film therefore constitutes a static phase pattern. Moreover, the optic is completely flat and can easily be combined with an amplitude mask.

An example of a GPH is a polarization grating (PG). It has a continuously rotating fast-axis orientation, which can also be interpreted as a phase ramp in accordance

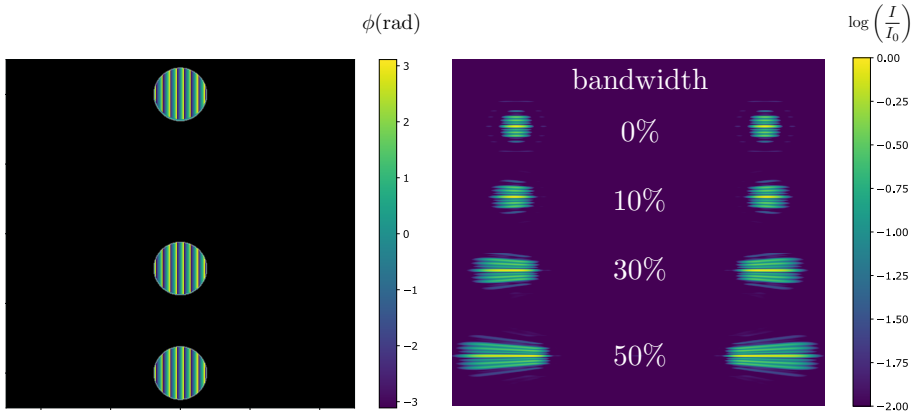

Figure 6.2: A linear arrangement of subapertures diffracted by a phase ramp or polarization grating. Such a design can be used over a large simultaneous bandwidth because the fringes are orthogonal to the bandwidth smearing.

to Eq. 6.1. A simple version of a HAM phase mask could therefore be manufactured as a collection of PGs for different subapertures. Yet, it is important to consider the diffractive nature of a GPH, which rules out a direct copy of mask designs for segment tilting. As geometric phase is independent of wavelength, the diffraction angle does change with wavelength. Thus, a continuous phase ramp will disperse the light, hence the term polarization grating. Large phase tilts required for HAM smear out fringes in the direction of the phase ramp for any non-zero bandwidth, while a major advantage of liquid-crystal technology is the high efficiency over large bandwidths. However, like segment tilting, there is total freedom to choose where baselines are imaged off-axis. If the fringes are orthogonal to the phase ramp, no amount of wavelength dispersion decreases the fringe visibility. We demonstrate this is Fig. 6.2. In this case, the broadband capabilities of the liquid-crystal technology can be fully exploited for broadband observations. A second advantage of using liquid-crystal GPHs, is that it is possible to write any phase pattern. We exploit this by combining polarization gratings into a single phase pattern that images a single subaperture onto multiple locations in the focal plane. The mathematical description of multiplexing can be found in Doelman et al. (2018). An example of multiplexing blazed gratings is shown in Fig. 6.3. Multiplexing creates great freedom to image any subaperture to multiple locations on the focal plane with high efficiency. However, multiplexing also gives rise to higher-order diffraction effects, that could potentially create interference between unwanted subapertures, see Fig. 6.3.

To summarize, the holographic component of HAM is created using geometric phase holograms, and the properties that will affect the HAM design are as follows.

1. Subapertures can be imaged onto any location in the focal plane using phase ramps.

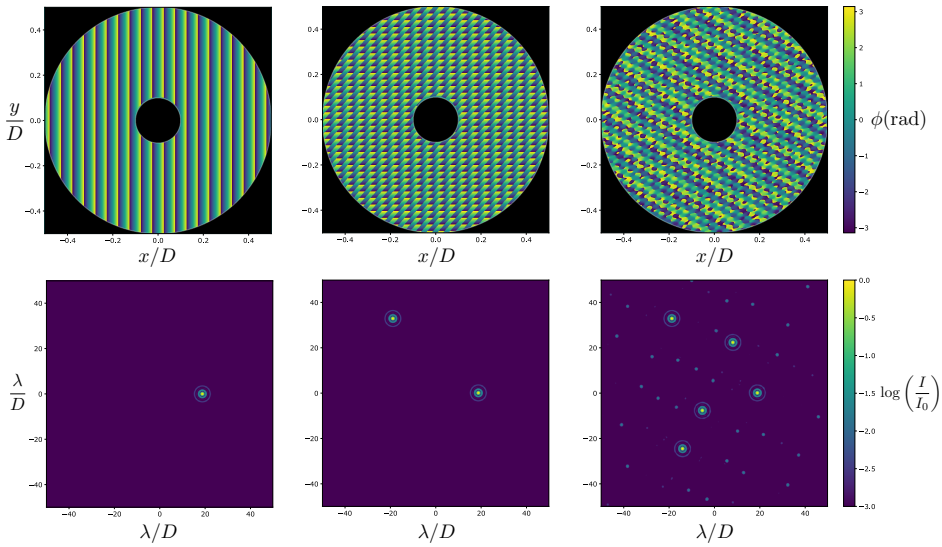

Figure 6.3: Simulations of combining multiple blazed gratings in a single hologram with multiplexing. The aperture multiplexes one, two and five holograms. Multiplexing multiple gratings increases crosstalk, i.e. the fainter PSFs seen easily in the 3rd panel.

2. Phase ramps can be combined for a single aperture with multiplexing.

3. A GPH is a diffractive phase element and the location of the PSF will change with wavelength.

4. A phase ramp will produce two PSFs at opposite location, one for each circular polarization state.

6.2.2 General considerations for a HAM design

Now that we have established how the liquid-crystal technology can be used to create off-axis interferograms, we can explore the design space it offers. The design of a HAM mask can be separated into several steps.

1. *Central component.* Select a SAM design that becomes the central component.

2. *Holographic component.* Select combinations depending on the number of closure phases versus the number of holograms and S/N per baseline.

3. *Focal plane.* Select the distances of holograms from the central component, which determines the spectral resolution of each baseline.

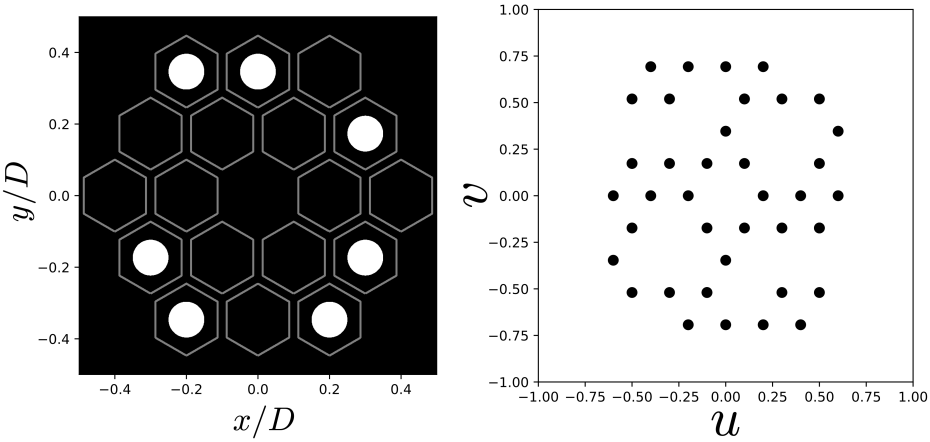

Figure 6.4: Example of a semi-redundant SAM design and the *uv*-coverage.

4. *Dynamic range.* Change piston phases of subapertures in the central component to reduce peak values.

We will elaborate on each of these points, adding design considerations.

The Central component

The goal of HAM is to increase the number of closure phases, throughput, simultaneous bandwidth and spectral resolution compared to a sparse aperture mask. A defining feature of aperture masking is that the detector information density, i.e. the number of baselines and closure phases, rapidly increase with the number of non-redundant subapertures that are imaged onto the same location. The diffractive nature of a liquid-crystal HAM mask limits the number of apertures that can be combined into a single hologram. As such, the highest information density on the detector can be reached by including a SAM mask that is imaged on-axis. Therefore, the design of a HAM mask starts with optimizing a SAM mask. Details on this optimization can be found in Carlotti & Groff (2010) and Tuthill (2018). We select a 7-aperture semi-redundant SAM mask for the demonstration of HAM design, as shown in Fig. 6.4.

The holographic component

With the central component defined, there is still $\sim 80\% - 90\%$ of the mask that can be used for the holographic component. Combined with the possibility to multiplex each subaperture, there are an extremely large number of combinations possible. Here we will explore the parameter space and list design rules and some of the trade-offs that can be made.
The design rules provide handles for a HAM design, yet they do not limit the

design freedom to a single solution, as this is also dependent on the science case. The design rules come from the properties of the geometric phase holograms listed in Section 6.2.1. From the diffractive nature (GPH property 3), we derive the first design rule, which also has the largest implication on the design.

1. Make 1D combinations of subapertures and image them onto a single interferogram that is placed orthogonally to their baselines.

If the baselines of subapertures are not coaligned, the fringes will smear out due to the dispersion not being orthogonal to the fringes. Only a limited number of subapertures can be arranged in a non-redundant way in a single dimension. Therefore, the number of closure triangles per interferogram is inherently restricted. Adding more independent closure triangles therefore requires GPH property 2: multiplexing. Multiplexing provides great design freedom but also has drawbacks, which is why design rule 2 states:

2. Minimize multiplexing.

The first two design rules that are in tension with each other. First we will explore the reasons multiplexing should be minimized before we look at the trade-offs that can be made.
Multiplexing splits up the light of a single aperture and directs it to multiple focal-plane locations. Adding one baseline, or multiplexing one additional grating, reduces the light in all other baselines of this subaperture. To demonstrate this, we look at a holographic component with n subapertures where every baseline has a separate interferogram. The number of holograms grows as $n(n-1)/2$, every subaperture is multiplexed $(n-1)$ times, and the number of closure phases grows as $(n-1)(n-2)/2$. So the number of holograms divided by the number of closure phases, i.e. the detector density, goes as $n/(n-2)$. However, the intensity of the interferograms decreases with $(n-1)$ due to the multiplexing. Overall, multiplexing does decrease the signal to noise ratio of all interferograms rather quickly. This is unlike the central component, where adding more subapertures does not decrease the visibility amplitude of baselines. This fundamental principle of the holographic component demonstrates why creating many closure triangles comes at a large cost.
A way to optimize the holographic component is to maximize the number of closure phases per holographic interferogram divided by the number of gratings multiplexed in a subaperture. That way we maximize the intensity of the holograms in combination with the number of baselines and closure phases, i.e the information density. This information density is highest when all supabertures combinations that contain coaligned baselines are imaged onto a single focal-plane location. For example, three subapertures imaged onto one location reduces the number of interferograms from 3 to 1. In conclusion, in order to minimize multiplexing while keeping a sufficient amount of closure triangles, it is key that a subset of subapertures is chosen that can be combined in multiple one-dimensional arrays. We define a subset as a collection of subapertures that are interfered with all other

Table 6.1: Specification of the holographic components of the three designs presented in Fig. 6.5. The central component (CC) is added as reference. $N_{subapertures}$ is the total amount of subapertures that are used in the holographic component, $N_{baselines}$ is the number of unique baselines, and $N_{holograms}$ is the total number of holograms. Moreover, N_{CLPH} is the number of closure phases, and $N_{Multi,max}$ and $N_{Multi,mean}$ show the maximum and mean amount of gratings multiplexed in the phases of the subapertures in the holographic component, where g and r indicate the green and red subset.

Design	1	2	3	CC
$N_{subapertures}$	9	10	11	7
$N_{baselines}$	9	15	19	19
$N_{holograms}$	6	28	46	-
N_{CLPH}	3	12	18	15
$N_{Multi,max}$	1	4	g: 6, r: 3	-
$N_{Multi,mean}$	1	3.1	g: 5.14, r: 3	-

subapertures in that collection and no other subapertures. A subset can have multiple interferograms, or only one like the central component.

3. Combine subapertures in multiple subsets.

Combining many subapertures leads to many baselines. The number of holograms that can be imaged onto the detector depend on the detector size, the subaperture diameter and the bandwidth of the filter. Different subsets of subapertures can be created to add baselines and closure triangles when an increase of the number of holograms for a single subset goes beyond the amount that would fit on the detector. Another advantage is that making multiple subsets reduces the amount of gratings that are multiplexed in single apertures. Again, this shows the trade-off between adding one subaperture that gives many new baselines and closure triangles but reduces the SNR of all holograms, versus adding a few new subapertures with higher SNR but less unique closure traingles. We will explore this trade-off in more detail using the designs in Fig. 6.5.

4. A subset is non-redundant if all interferograms, not baselines, are non-redundant.

Rule 4 adds more freedom for designing subsets, which becomes more useful when most of the mask is already filled with subapertures. A limit is set by the available detector space in the direction orthogonal to the redundant baselines, in accordance to rule 1, which automatically leads to a more homogeneous spread of baselines in all directions for square detectors.
With these design rules in mind, we can explore some of the possibilities of HAM designs. This is a similar trade-off that is present for SAM design, trading subaperture size with the number of non-redundant baselines that fit onto a mask (= SNR vs. *uv*-coverage). A mask with large subapertures and limited *uv*-coverage

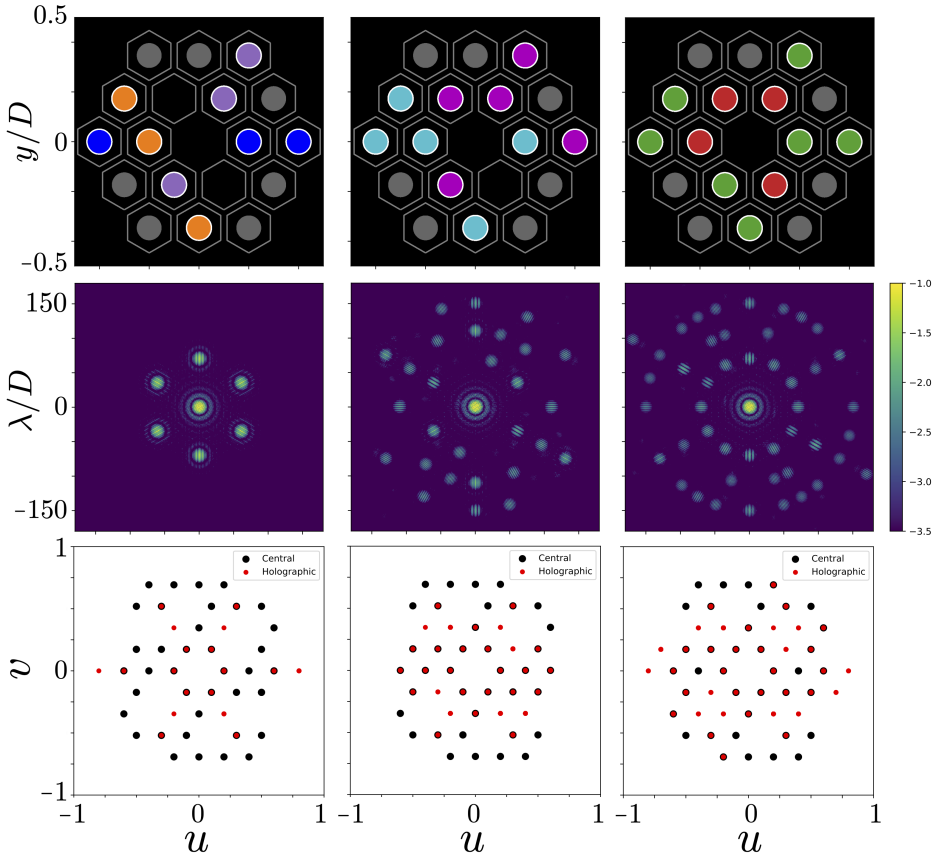

Figure 6.5: Three different designs of the holographic component. *Top:* pupil designs, where the colors indicate the subapertures that are combined in single or multiple holographic interferograms. *Middle:* focal plane image of the mask on logarithmic scale. *Bottom: uv*-coverage of the masks, with the central component in black and the holographic component in red.

provides requires less observation time to characterize targets with low complexity, e.g. a binary system, than a mask with smaller holes and fine sampling in the *uv*-plane. However, the finer sampling is key to retrieve complex structures of a source. For HAM the trade-off is also influenced by the available detector space and available bandpass.
We explore three designs, each with a different SNR and *uv*-coverage. The designs, their PSF and the *uv*-coverage are presented in Fig. 6.5, and the specifications are shown in Table 6.1. The first design optimizes the SNR, consists of three subsets of 1D combinations of subapertures without multiplexing. With only nine unique baselines and three closure phases, the holographic component is not optimal for operating in stand-alone mode. However, it is a good design for extreme broadband observations, e.g. $\Delta\lambda/\lambda_0 \sim 100\%$. Moreover, with enough sky-rotation it can still provide sufficient *uv*-coverage. Design 3 is the complete opposite of the first design, with extreme *uv*-coverage but low SNR. The number of unique baselines is equal to the SAM mask, and the number of closure phases even exceeds the SAM mask design. Each subaperture is multiplexed 4.3 times on average, some even six times. The focal plane is crowded, filling most of three rings with holograms that have less than 1% peak intensity compared to the central component peak intensity. Together, both components almost fill all possible locations in the *uv*-plane that correspond to the hexagonal grid of the mirror segments. On average the holograms are quite close together, limiting the simultaneous spectral bandwidth. A compromise of these two extremes is presented in design 2. Two components of five subapertures create 12 closure phases and 15 baselines, which comes close to the central component design. This design only adds four new baseline with respect to the design of the central component. Compared to the green subapertures in design 3 the holograms have increased a factor 1.7 in brightness. This trade-off makes the second design suitable for stand-alone mode for a more diverse range of targets.
It is worth pointing out that this specific example assumes a hexagonal grid of subapertures and an equal size of the subapertures in both components. For example, a good design will more look like design 1 if the subapertures are small in comparison to the full aperture. Not only are there more non-redundant subapertures available in a line, the off-axis holograms are also larger. As such, it is difficult to generalize the example. A complete review on design trade-offs is outside the scope of the paper.

Focal plane design

We switch to the focal plane to explore the effects of hologram placement in more detail. If the subsets of apertures are chosen, the focal plane has a set of holograms that need to be given a location. Each hologram can be imaged onto a line, where the separation with respect to the central component is a design freedom. This impacts spectral resolution, spectral bandwidth, and interference between higher-order terms. First, we will explore the spectral resolution of holograms.
The spectral resolution of the holographic component is actually defined by the

maximum of two terms. Both terms are independent and are defined in different planes, i.e the focal plane and the uv-plane, and are derived in Appendix 6.10. The first term has to do with the diffraction by the gratings of the HAM phase mask. For two subapertures with a subaperture diameter D_{sub}, a phase grating with period P, the spectral resolution in the focal plane is given by

$$R_{fp} = \frac{\lambda}{\Delta\lambda} = \frac{D_{sub}}{1.22P}. \tag{6.2}$$

The second term is a fundamental property of interferometry, where the uv-points, or baselines, are wavelength dependent due to diffraction. If the two subapertures form a baseline $\mathbf{b}$, their spectral resolution in the uv-plane is given by

$$R_{uv} = \frac{\lambda}{\Delta\lambda} = \frac{|\mathbf{b}|}{D_{sub}}. \tag{6.3}$$

With different methods we are able to fit fringes directly in the focal plane or retrieve complex visibilities from the uv-plane independently. We can choose which method to apply to obtain the highest spectral resolution. For most holograms $\frac{D_{sub}}{1.22P} > \frac{|\mathbf{b}|}{D_{sub}}$, yet it can be worthwhile to fill the image plane closest to the central component with the longest baselines. A maximum spectral resolution is obtained when the holographic interferograms are placed near the edges of the detector. Doing so introduces a larger sensitivity to the effects of non-perfect optics. This is particularly problematic for closure triangles where the baselines are imaged at different parts of the detector. For example, image distortion can change the shape and location of the holographic interferograms. Variable image quality across the focal plane can add different phase offsets to each baseline, resulting in a non-zero closure phase. A full analysis of these effects is outside of the scope of the paper. Another consideration is the crosstalk between multiplexed gratings, as shown in Fig. 6.3. By minimizing multiplexing this effect is already reduced, however, it is also possible to reduce the impact by switching circular polarization states of interferograms. Changing the sign of the grating frequency of gratings changes the crosstalk locations. Moreover, changing the sign of the grating frequency changes the circular polarization state, potentially making the interferogram incoherent with the crosstalk term. This is useful if an interferogram is near a crosstalk PSF, or if two interferograms have similar baseline directions.

Dynamic range

A last design step is to change the peak fluxes of the central component to reduce the brightness ratio to the holographic components. Due to multiplexing, the light from any HAM subaperture is redirected to multiple image plane locations. HAM interferogram spots therefore have a fraction of the intensity compared to the central component. For example, a central component can be the interferogram of nine holes, compared to a holographic interferogram that may only combine one third of the intensity of two holes: an issue made still more pronounced when

off-axis light is spread by bandwidth smearing. This can lead to more than an order of magnitude difference between the peak intensities of the holographic and central components.
All subapertures of the central component constructively interfere on the optical axis, leading to a highly peaked PSF. Nulling these peaks of the central component reduces the required dynamic range significantly. The nulls are generated by changing the piston phase offsets of individual subapertures of the central component. It is not necessary to reach a perfect null, the goal is only to homogenize the intensity of the central component. To this end, the piston phase offsets could be random, as long as the piston was corrected for in the closure phase calculations. However, liquid-crystal technology limits us to 0 or $\pm\pi$ phase offsets due to applying opposite phase to the opposite circular polarization state, i.e. Eq. 6.1. Other piston phases would make the baseline appear as redundant with two piston offsets. As a phase piston of 0 or $\pm\pi$ is the optimal way of nulling, this is not a problem. Another advantage of geometric phase is that the phase offsets are achromatic, such that the interferogram does not change with wavelength. Selecting which subapertures apply the phase offset depends on the central component design. To null the central peak, half of the subapertures should acquire the phase piston. This is not possible for designs with an odd number of subapertures. When the central component design contains doubly redundant baselines, there is an additional constraint. The identical baselines need to have equal piston offsets or they are nulled. An example of the nulling of the central component is presented in Fig. 6.6. The peak brightness in the three designs, normalized on the first and from left to right, are [1,0.41, 0.53]. So nulling reduces the peak intensity by a factor two.

6.3 The influence of polarization leakage

Both the central component and a polarization leakage term are imaged onto the optical axis of the system. Here we wile the influence of polarization leakage on the central component, and we will explain its effects on HAM observations. Polarization leakage emerges when the retardance of a geometric phase hologram is not exactly half-wave. As mentioned before, the geometric phase hologram is a half-wave retarder with a spatially varying fast axis. The space-variant Jones matrix of such a retarder in the circular polarization basis is given by

$$\mathbf{M} = c_V \begin{bmatrix} 0 & \mathrm{e}^{i2\chi(x,y)} \\ \mathrm{e}^{-i2\chi(x,y)} & 0 \end{bmatrix} + c_L \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{6.4}$$

Here $\chi(x, y)$ is the spatially varying fast-axis orientation, and both c_V and c_L are parameters that depend on the retardance $\Delta\phi$ (Mawet et al., 2009; Ruane et al., 2019) and are given by

$$c_V = \sin\frac{\Delta\phi}{2},\ c_L = -i\cos\frac{\Delta\phi}{2}. \tag{6.5}$$

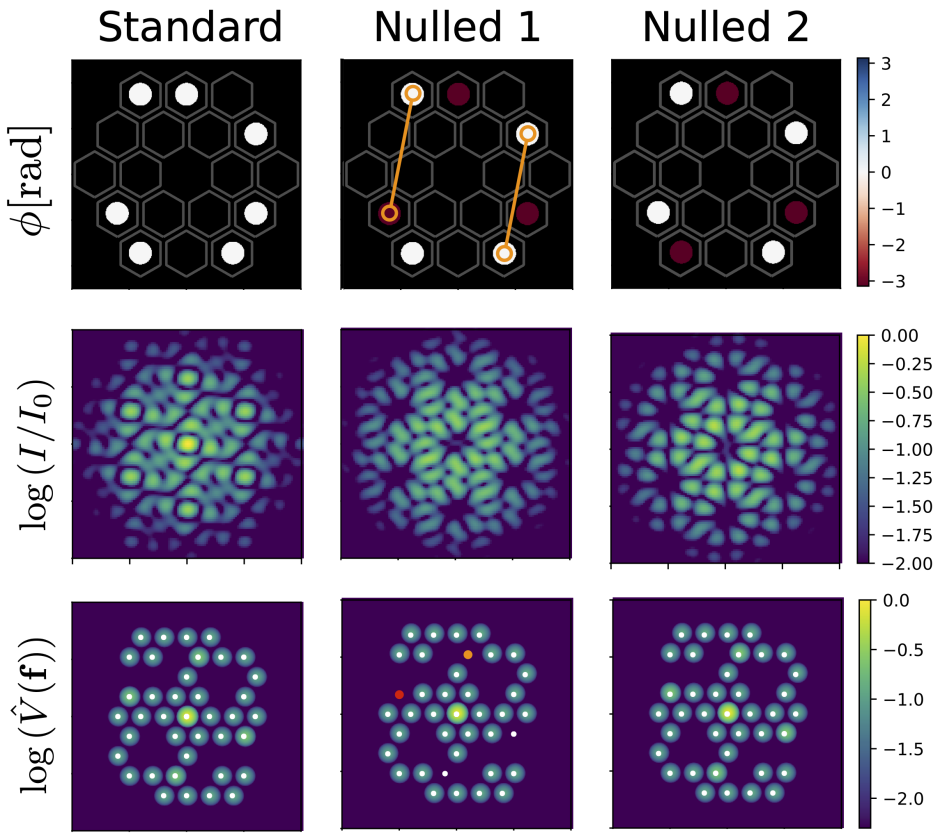

Figure 6.6: Phase, PSF and the visibilities of the central component with and without nulling. Adding π phase offsets to different holes shapes the PSF and reduces peak intensity. If doubly redundant baselines (orange) have different phase offsets, the baseline is nulled achromatically.

The first term in Eq. 6.4 describes that a fraction C_V of the light acquires a geometric phase of

$$\Phi(x, y) = \pm 2\chi(x, y), \tag{6.6}$$

where the sign of the phase depends on the handedness of the incoming circular polarization. The second term describes the polarization leakage beam, and is unaffected by the fast-axis orientation pattern. When the retardance is perfectly half-wave, i.e. $c_V = 1$ and $c_L = 0$, the right-handed circular polarization states are converted to

$$\mathbf{RC_{out}} = \mathbf{MRC_{in}} = \mathbf{M} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \mathrm{e}^{-i\Phi(x,y)} \end{bmatrix}, \tag{6.7}$$

With non-perfect retardance, the output electric field for incoming right-circular polarization is given by

$$\mathbf{RC_{out}} = \mathbf{MRC_{in}} = \mathbf{M} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} c_L \\ c_V \mathrm{e}^{-i\Phi(x,y)} \end{bmatrix}. \tag{6.8}$$

The leakage term and main beam have an orthogonal polarization state and are therefore incoherent, assuming no polarization cross-talk due to the optical system. In addition, unpolarized light contains on average equal amounts of left- and right-circular polarization and these states are incoherent, see Hecht & Zajac (1974). Therefore, we can describe the impact of leakage, typically on the order of 1%, for 1 circular polarization state without loss of generality. While objects do not appear fully unpolarized due to instrumental polarization or interstellar polarization, we will first explore this simplification because it demonstrates how closure phases and complex visibilities are less resistant against wavefront aberrations.
From Eq. 6.8, it is clear that the point-spread function (PSF) is actually an incoherent sum of the HAM PSF and the leakage PSF, where the leakage PSF is the unaltered PSF from the HAM amplitude mask. Moreover, the Fourier transform is a linear operator. Therefore, we can calculate the visibilities of the HAM PSF and the leakage PSF separately and co-add them. The combined measured visibility is then given by

$$V(\mathbf{f}) = c_V^2 V_{SAM}(\mathbf{f}) + c_L^2 V_{holo}(\mathbf{f}). \tag{6.9}$$

Here, $V_{SAM}(\mathbf{f})$ is different from $V_{holo}(\mathbf{f})$, as their amplitude masks are different. The leakage term contains many redundant baselines for realistic HAM designs, as it is the combination of the subapertures of the central and holographic component. Therefore, $V_{holo}(\mathbf{f}$ is a sum over all redundant baselines, each with a different phase. While $C_V^2 \gg C_L^2$ for a retardance close to half-wave, $c_L^2 V_{holo}(\mathbf{f})$ can be a significant fraction of $V(\mathbf{f})$ if their phases are non-random. Any phase aberration has a non-zero sum of phase for the redundant baselines in the HAM mask, and adds an offset to the visibility phase that is unique per baseline. This offset is affects both the closure phase retrieval as well as the squared visibilities. The impact depends on aberration strength and time dependent wavefront errors change the closure phases in a way that can not be calibrated by observing a calibrator.

For any linearly polarized light fraction it is not correct to assume that the leakage PSF is incoherent. Linearly polarized light can also be written as the sum of the two circular polarization states, however, they are still coherent. That means that the polarization leakage of left-circular polarization interferes with the main beam of the right-circular polarization state that becomes left-circularly polarized after going through the half-wave retarder, i.e.

$$\mathbf{RC_{out}} = \mathbf{MLP_{in}} = \frac{1}{\sqrt{2}}\mathbf{M}\begin{bmatrix}1\\-i\end{bmatrix} = \begin{bmatrix}c_L - ic_V \mathrm{e}^{i\Phi(x,y)}\\ c_V \mathrm{e}^{-i\Phi(x,y)} - ic_L\end{bmatrix}. \tag{6.10}$$

The PSF will contain an interference term with a relative intensity $\sim C_V C_L$, which is smaller than C_V^2 when the retardance is close to half-wave, but much larger than C_L^2. The effect of this interference depends on many factors, including the retardance of the HAM mask, the linear polarization fraction, the wavefront aberrations and the nulling of baselines using $\pm\pi$ offsets. This effect is strongest for 100% linearly polarized light coming in on the HAM mask, but can be significant even for small fractions of linearly polarized light at the level of 1%. Moreover, the equations are the same for polarization crosstalk due to instrumental polarization. This means that time-variable polarization states, e.g. induced by a derotator at different orientations (van Holstein, R. G. et al., 2020), cannot be calibrated with a single calibrator at the beginning of the observing sequence. Similarly, differences in linear polarization state between a calibrator and a science target induces a non-zero signal in closure phase that can also not be calibrated. Simulations of a mask design in Appendix **??** show that the maximum deviation of closure phases is linear with the degree of linear polarization. We find that a linear polarization fraction of 1% corresponds to a 1 degree offset for 2.5% percent leakage. In the same fashion, the $\pm\pi$ phase offsets of some subapertures are also impacting the retrieved complex visibilities when leakage is present. A unique baseline from the central component with this phase offset is nulled by multiple non-shifted baselines from the polarization leakage PSF. This leads to a reduction of squared visibility of phase-shifted baselines. Wavefront aberrations further complicate the picture, and calibrating the squared visibilities is not possible with a calibrator. Overall, extreme caution is warranted when the linear polarization fraction is non-zero and the polarization leakage is above 1%.

There are two solutions that will limit the effect of polarization leakage. First, one can filter circular polarization states before and after the HAM optic, similarly to what is done for the vector-vortex coronagraph (Foo et al., 2005; Mawet et al., 2009). For an existing instrument this is usually not possible to implement, although a HAM optic could be sandwiched between circular polarizers to make it into a single optic. The downside of such a solution is the 50% throughput reduction. The second solution is a double-grating HAM optic (Doelman et al., 2020). Adding a polarization grating pattern to the full phase pattern, followed by a second identical polarization grating, reduces the polarization leakage by at least one order of magnitude. This solution is explored further in Section 6.9.

Table 6.2: Specifications of the full HAM mask (right), decomposed into contributions from the central component (CC, left) and the holographic component (HC, middle).

	CC	**HC**	**HAM**
Subapertures	11	19	30
Throughput	10.1%	21.9%	32.0%
Baselines	55	30	85
Unique baselines	50	23	51
Single baselines	45	17	27
Repeated baselines	10	13	58
• Pairs	5	5	16
• Triplets	0	1	7
• Quadruplets	0	0	0
• Quintuplets	0	0	1
Closure Triangles			
• w/ repeated SAM baselines	165	26	191
• w/o repeated SAM baselines	88	26	114
Unique Closure Triangles			
• w/ repeated SAM baselines	165	26	190
• w/o repeated SAM baselines	88	26	114
***uv*-points**			
• w/ repeated SAM baselines	100	46	102
• w/o repeated SAM baselines	90	46	98

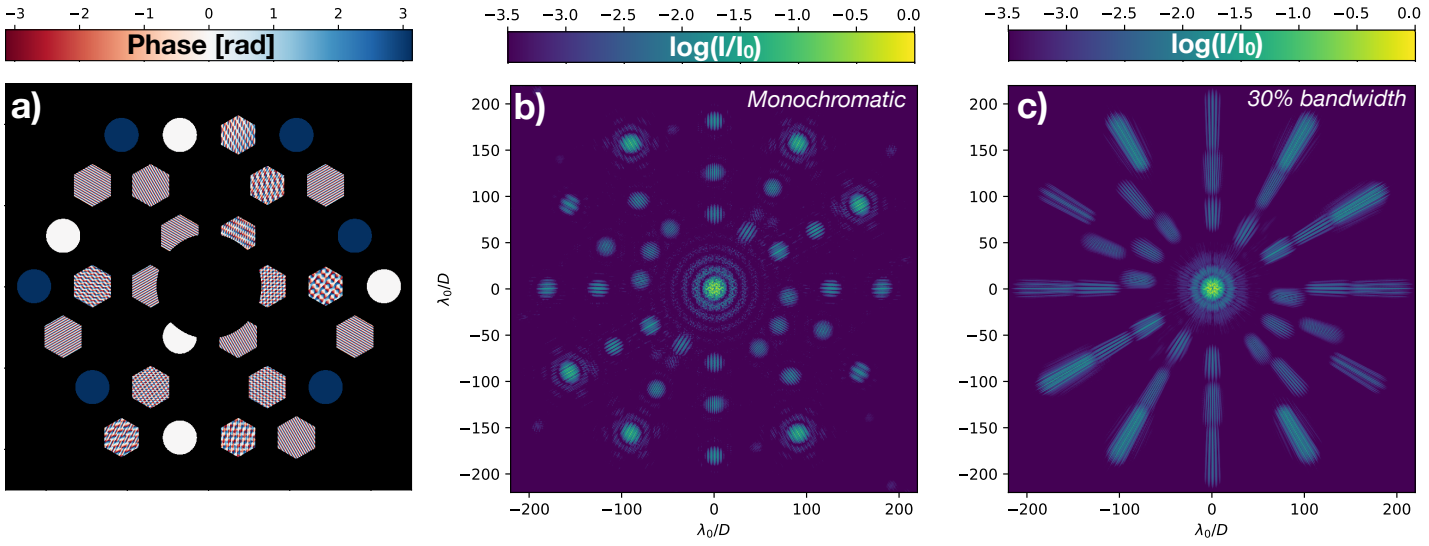

Figure 6.7: The design of the HAM mask for OSIRIS. (*a*) Phase pattern of the HAM optic, masked by the amplitude mask. (*b*) The simulated monochromatic PSF. (*c*) The simulated PSF with 30% bandwidth.

6.4 Design of a prototype mask for Keck OSIRIS

In this section, we discuss the design of a prototype HAM mask for the OH-Suppressing Infrared Integral Field Spectrograph (OSIRIS) (Larkin et al., 2006). The design follows the considerations mentioned in Section 6.2.2. The OSIRIS detector, a Teledyne Hawaii-2RG HgCdTe detector with a size of 2048×2048 pixels, provides a unique opportunity. With a plate scale of 10 mas/pixel, the detector is large enough to add many baselines to the holographic component. We place the subapertures in alignment with the Keck primary mirror segmentation, such that every subaperture is centered on one segment. The segments that are crossed by a spider are not used in the design. Moreover, the subapertures of the holographic component are hexagonal to increase their throughput, and we increase their diameter if the smallest baselines are larger than one the distance between neighboring subapertures. The subapertures of the central component are circular as this keeps the central component PSF circularly symmetric and the outer Airy rings have equal strength for all holograms at equal radius. The full design and both a monochromatic and broadband PSF are shown in Fig. 6.7.

The central component consists of an 11-hole SAM mask. It is optimized to maximize throughput, the number of baselines, and the number of closure triangles at the cost of a few redundant baselines. Tuthill (2018) showed that adding redundancy allows for a boost the S/N ratio with respect to non-redundant masking by $\gtrsim$50%. In addition, we offset five holes of the central component by π to reduce the required dynamic range of the detector and the influence of polarization leakage, as explained in Sect. 6.2.2. This specific combination of phase-shifted subapertures minimized the PSF of the central component at the location of the polarization leakage PSF maxima and does not null redundant baselines.

The holographic component uses the remaining apertures, for a total of 19. Nine subapertures are used for three linear combinations of three holes each, yielding nine baselines and three closure phases for three holograms. Because these holograms are not multiplexed, they yield the interferograms with the highest signal-to-noise ratio. The other ten subapertures are devided into a highly-multiplexed combination of six holes and a combination of four holes. These designs are optimized to maximize the number of baselines and closure triangles per hologram. Overall, the holographic components consists of 18 holographic interferograms per polarization, yielding 26 closure triangles.

The subaperture locations and their mapping to the focal plane can be found in Appendix **??**. The full overview of HAM properties is provided in Table 6.2. Additionally, it makes a distinction between the individual contributions from the central component and holographic component. Several numbers are worth pointing out. Apart from transmission and diffraction efficiency, the total throughput of the mask is $> 30\%$, which is a factor of 3 higher than the throughput of the central component. Whereas the central component features 10 repeated baselines (i.e. five pairs of "redundant" baselines), 13 out of the 30 HAM baselines have at least one identical counterpart. However, because the light is mapped onto different spots in the focal plane, their complex visibility can be computed independently.

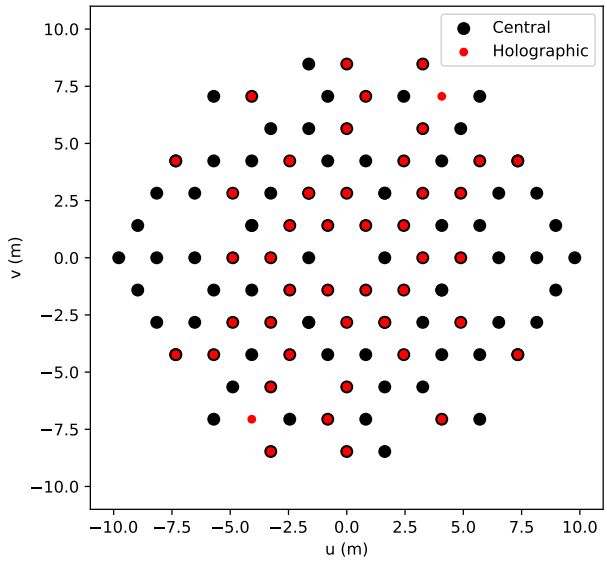

Figure 6.8: The *uv*-coverage of both the central and holographic component.

The HAM mask has a total of 85 baselines, 51 of which are unique. The unique baselines can be divided into two groups: (i) *single* baselines, which occur only once, and (ii) *repeated* baselines, which occur either 2, 3 or 5 times (see table 6.2). With the 50 unique baselines accommodated by the central component, the *uv*-plane is uniformly sampled and provides great coverage. The 23 unique baselines of the holographic component provide sufficient *uv*-coverage that allow it to be used on its own. Moreover, the baselines present in both components can be used to improve calibration of the data.

6.5 Data reduction pipeline

In this section, we will briefly go over the new pipeline that was developed for HAM data reduction.

6.5.1 Pipeline overview

Generally, pipelines for aperture masking consist of the same building blocks.

1. *Pre-reduction.* This includes flat-field corrections, bad-pixel corrections and background subtraction;

2. *Complex visibility.* The phase and amplitude of the complex visibility are measured on each baseline. To this end, there exist two approaches: (i) fringe-fitting methods, which rely on measuring the fringe phases and amplitudes directly from the focal plane, e.g. Lacour et al. (2011), Greenbaum et al. (2014), and (ii) Fourier methods, which extract information from the Fourier Transform of the PSF, as shown in Figure 1.4, e.g. Tuthill et al. (1999), Sallum & Eisner (2017);

3. *Closure phases.* The visibility phases are combined to yield robust observables, such as closure phases and squared visibilities;

4. *Calibration.* Closure phases of a point source are subtracted from the closure phases of the science target, to correct for systematic effects;

5. *Model fitting.* By fitting the data with an analytical source model, one can estimate the source parameters.

The HAM pipeline is written in Python using the `HCIPy` package (Por et al., 2018). All operations are carried out separately for the central and the holographic component. The pipeline uses both methods to retrieve the complex visibilities. For the fringe-fitting method, we apodize the PSF of the central component using a power 2 super-Gaussian window function with a FWHM of 110 λ/D in both axes. This separates the central component from the holographic component, in addition to suppressing high-frequency noise in the uv-plane. Similarly, for the holographic component we apodize the PSFs with a power 2 super-Gaussian with a FWHM of 33 λ/D, placed at the location of the interferogram. This location is pre-computed according to the plate scale of the data. The pipeline builds a fringe-library for every interferogram at their own location, such that a direct fit can be made without shifting any interferograms. The fit is a least-squares optimization with a model matrix that contains the flattened versions of all fringes in the fringe-library.
The Fourier method uses the same super-Gaussian masks for both components. Masking causes information from neighbouring uv-pixels to blend, such that the central pixel value is representative for the entire splodge. Visibilities are then extracted at the central locations of splodges in the uv-plane. For the holographic component we mask individual interferograms before doing the Fourier transform. It is not necessary to center on the interferograms. While this does introduce a large phase slope in the uv-plane, we know the location of the PSF and can subtract a pre-calculated phase slope. Moreover, if we assume that there is no distortion or only symmetric distortion in the image plane, we can average the visibilities of the interferograms with opposite circular polarization. As they are on exactly opposite sides of the image center, the phase slopes cancel.
The closure phases are calculated for each closure triangle from the individual visibility phases. After subrracting the closure phases of a calibration target, the pipeline can fit an analytical model to the residual closure phases. For now the only model is a binary object, and the fit is performed using least-squares minimization.

6.5.2 Spectroscopic closure phase extraction

The spectroscopic capabilities of HAM are also exploited in the pipeline, which is capable of extracting multi-wavelength closure phases from a single image. For the Fourier method this is an almost trivial addition, as the baseline length in the uv-plane can be scaled according to the wavelength. The sampling in the uv-plane can be changed accordingly with the use of a matrix Fourier transform (Soummer et al., 2007). In addition, we change the mask for the interferogram to be an elliptical super-Gaussian, with the major axis in the smearing direction. Extracting the fringe visibilities and closure phases is done at a wavelength sampling that is higher than the intrinsic spectral resolution. This spectral resolution is given by the dispersion in the uv-plane which depends on the wavelength and baseline length, see Eq. 6.3. Therefore, the spectral resolution between baselines that are combined in a single closure triangle might differ. We oversample the uv-plane compared to the highest spectral resolution and recover the wavelength-dependent closure phases using this sampling . As such, the recovered closure phases are not completely independent as function of wavelength.
Spectroscopic extraction of closure phases is different for the fringe-fitting method. Using the fact that a broadband PSF is the incoherent sum of multiple monochromatic PSFs, it is possible to fit a fringe to multiple wavelength-scaled locations of the holograpic interferograms. The spectral resolution is given by Eq. 6.2. If the monochromatic PSFs have a separation of $1.22\lambda/D_{sub}$, they can be used as individual fringes in the fringe-library and fitted simultaneously. However, extracting the fringes at a higher spectral resolution with overlapping PSFs requires regularization of the design matrix, as the columns are no longer linearly independent. We tried two methods of regularization, i.e. Tikhonov regularization and regularization using polynomial coefficients. For Tikhonov regularization, we construct the Tikhohnov matrix Γ such that it acts as a difference operator. As a result, the regularisation punishes large values of the finite differences of the fringe phase between two adjacent wavelength bins. While this enforces a smooth transition with wavelength, it could affect the calculated closure phases. In addition, Tikhonov regularization requires tuning which is computationally expensive and might change between targets. As this is not ideal, a different regularization method was explored. For simple systems like a binary system, the fringe phases behave smoothly with wavelength. A low-order polynomial could be a good estimate of such behavior, and, can therefore be combined with the fringe-fitting method as regularization. Instead of directly fitting fringe phases, we fit polynomial coefficients to the wavelength dependency of the fringe phases. The full equations are shown in Appendix **??**.
We simulate a binary with a separation of 1.08 λ/D, a brightness ratio of r = 0.5 and use all three methods described in this section to retrieve the broadband closure phases of the holographic component.The results for two closure triangles are shown in Fig. 6.9. We explore a different sampling in wavelength for every method, i.e. uv-plane sampling (Fourier), number of wavelength bins (Tikhonov) and polynomial order (Polynomial). For the Fourier method, the solutions barely

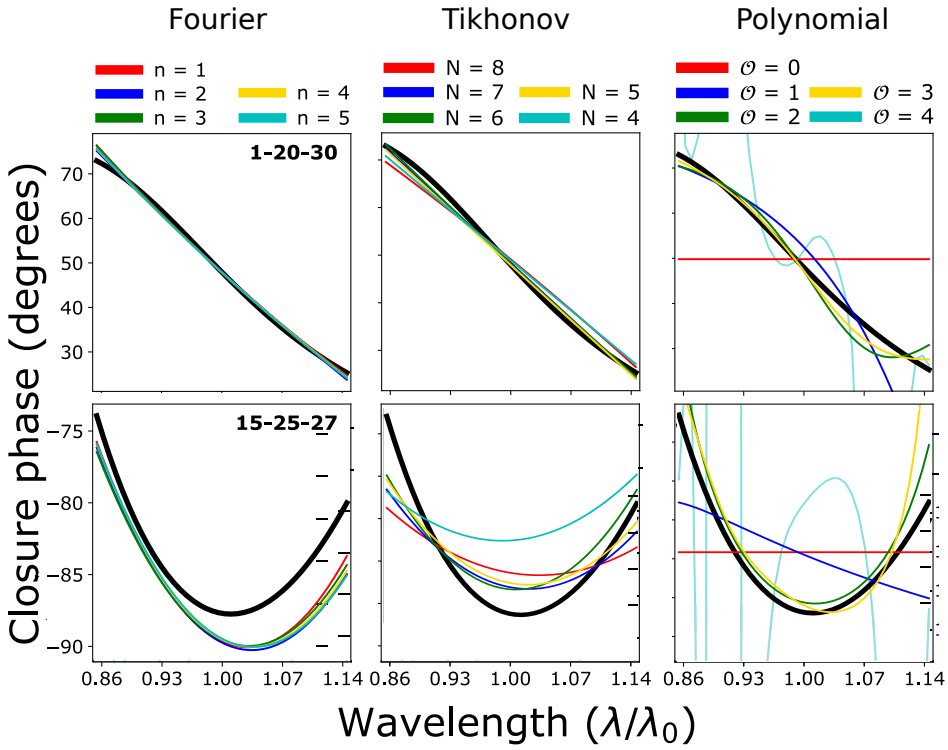

Figure 6.9: Analytical closure phases (black) and extracted closure phases (color) of a binary system applying different closure phase extraction algorithms to the holographic component. Two representative closure triangles were selected. The binary has a separation of 1.08 λ/D, a brightness ratio of r = 0.5. The simulated bandwidth is 30% (i.e. $\Delta\lambda/\lambda_0 = 0.3$). For the Fourier method we change the uv-plane sampling from n = 1 to 5, meaning $2D/1000$ to $2D/5000$. Closure phases extracted using fringe-fitting with Tikhonov regularization is carried out for a different number of wavelength bins N, and we vary the polynomial order $\mathcal{O}$.

depend on the uv-plane sampling, as long as it is high enough. The polynomial fit best captures the behavior of the analytical model at order two and three. The extreme behavior of the fourth order polynomial suggests over-fitting, implicating that the spectral resolution is not high enough to constrain more than four components. The Tikhonov extraction has similar performance with $N > 4$, and has decreased performance toward the edge of the band. This could be a direct result of the choice of regularization.
To summarize, all three methods have similar performance for the optimal sampling of the bandwidth. They are capable of reproducing the change of closure phases with wavelength within a five degrees over 30% bandwidth. The limited spectral resolution of the holographic spots is just enough to capture the change of baseline units with wavelength for this binary system. More rapidly changing closure phases, e.g. from a more structured object, are not sampled well and will be recovered as a smoothed function. This limits the detail that can be added to the model.

6.6 Manufacturing of prototypes HAM v1 and HAM v1.5

Two prototype HAM devices were manfactured by ImagineOptix in August 2018, labelled part A and part B. The first HAM device, part A, was manufactured using a 1 inch flat CaF2 substrate with a thickness of 5 mm, while the second HAM device, part B, was fabricated using a 1 inch wedged CaF2 substrate with a thickness of 1 mm. The front-sides have the same 3-layered liquid-crystal multi-twist retarder film, aimed at minimizing polarization leakage between 1 and 2.5μm. Both devices have an anti-reflection coating for the this bandpass on the backside of the substrates. The phase pattern with a diameter of 25.4 mm was generated with 5 micron pixels. The polarization leakage measured by the manufacturer is less than 3% between 1 and 2.5μm, see Fig. 6.14. Additional alignment markings have been added to the pattern, outside of the pupil diameter of 13.5 mm. An image of the optic between polarizers is shown in Fig. 6.10a. In addition, the optic was inspected under a microscope between polarizers. The four microscope images presented in Fig. 6.10b show the high quality of the manufacturing process. Amplitude masks were laser cut in 100um brass and/or 304 stainless steel foils with a diameter of 20.83 mm, using a picosecond laser machining facility (OptoFab node of ANFF, Macquarie University, Sydney). It is screwed in place in a holder in the filter wheel assembly. The first version (HAM v1) was installed in the imaging arm of OSIRIS at the Keck I telescope in September 2018. Each position of its first filter wheel contains separate pupil mask and filter mount assemblies. The amplitude mask was installed in the pupil assembly facing the incoming beam, while the HAM optic was installed in the opposite 1-inch filter side of the wheel assembly. Therefore, a gap of several mm was present between the amplitude mask and the HAM optic. This version has been tested with an internal source in OSIRIS in April 2019. The results are presented in section 6.8.

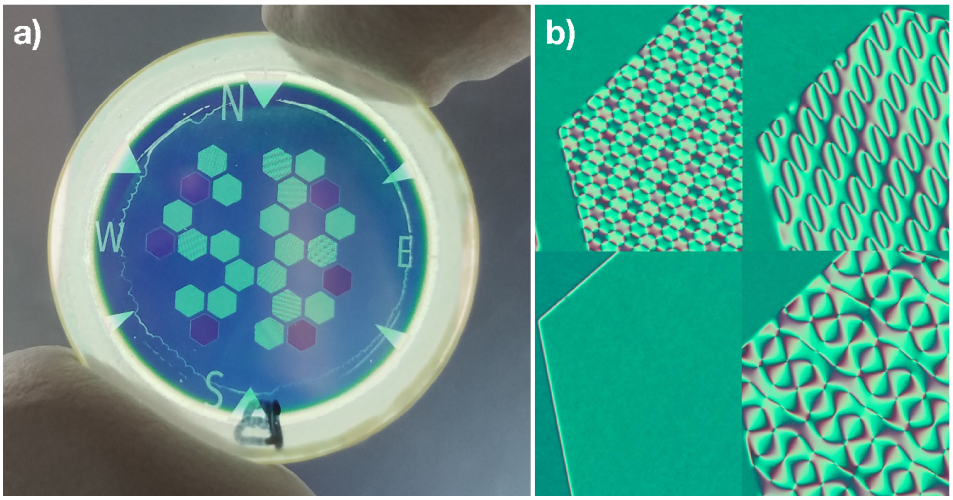

Figure 6.10: Images of the HAM optic between polarizers. Image credit: ImagineOptix

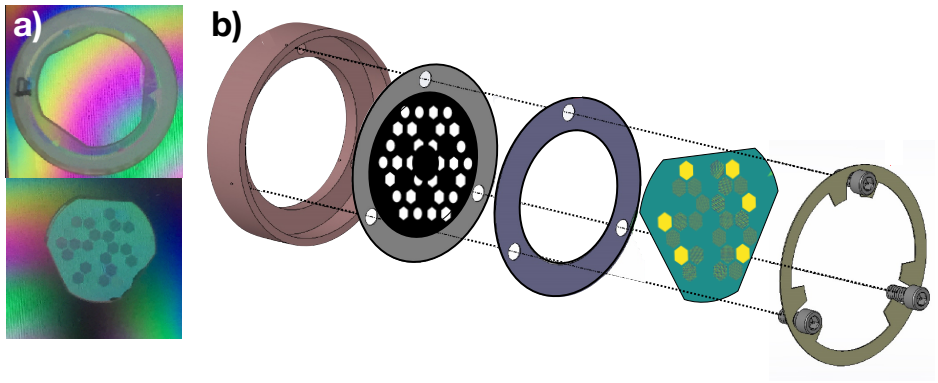

Figure 6.11: Manufacturing of HAM v1.5. (*a*) Images of the diced HAM v1.5 phase optic. (*b*) The assembly of HAM v1.5 in the OSIRIS pupil mount.

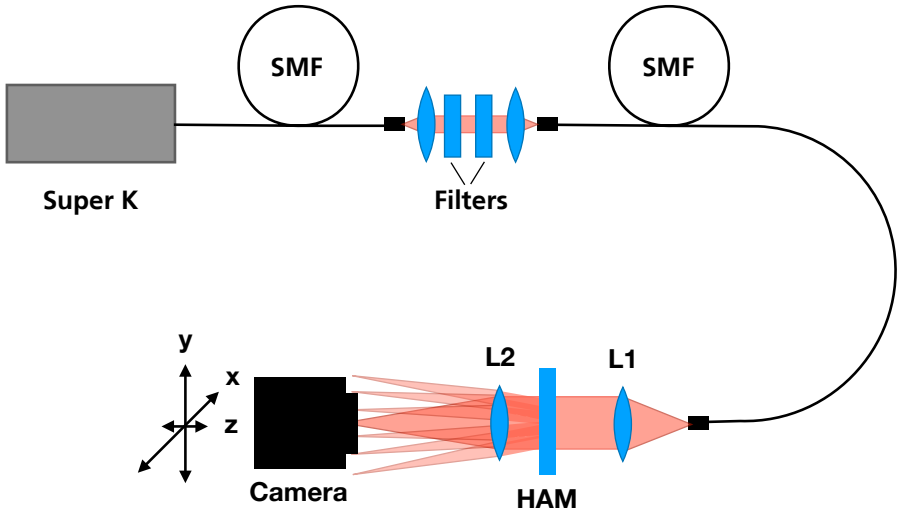

Figure 6.12: Lab setup used to characterize the HAM v1.5 optic.

An updated version (HAM v1.5) has been created to allow the HAM optic and the amplitude mask to be installed in the same pupil mask holder, reducing the separation between the optics to almost zero. The HAM optic of HAM v1.5 is a cut-out version of the spare HAM v1 phase mask (part B). The cut-out mask and assembly of the optic in the mount are shown in Fig. 6.11. Initial lab tests were conducted in Sydney in the Fall of 2019. HAM v1.5 has been installed in OSIRIS in February 2020, replacing HAM v1 in OSIRIS. We present the results of the lab tests in the next section.

6.7 Lab tests

We tested the HAM v1.5 optic in the lab using the setup described in Fig. 6.12. The light source is a SuperK COMPACT from NKT photonics and is connected to a custom reimaging system through a single-mode fiber (SMF). The reimaging system allows us to insert spectral filters and neutral density filters in a collimated beam before injection into the optical setup with a second SMF. We used filters from the Thorlabs IR Bandpass Filter Kit from 1000 nm to 1600 nm. Light from the second SMF is collimated with a Thorlabs 1 inch doublet with a focal length of 150 mm (AC254-150-C-ML). A second 300 mm (AC254-300-C-ML) doublet is placed close to the HAM optic to prevent vignetting of individual HAM apertures. The camera is a CRED2 and is mounted on motorized X,Y,Z translation stages to fully capture the HAM PSF. We used this motorized camera because a 2K pixel science grade detector, like a Teledyne Hawaii-2RG HgCdTe detector, did not fit

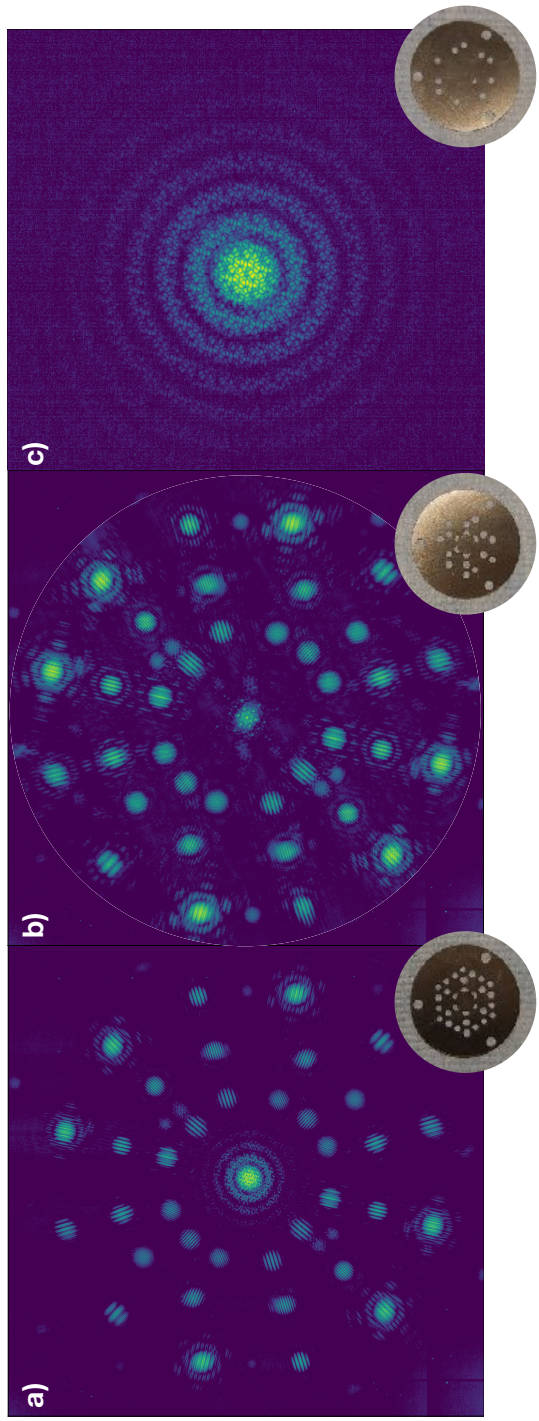

Figure 6.13: Images of the normalized HAM PSF with different amplitude masks at 1400 nm. (*a*) Full HAM PSF showing that all holograms are present and contain interferograms. (*b*) HAM PSF of the holographic component where the central term is zero order diffraction term of the holographic gratings. (*c*) Close-up of the central component PSF.

in the budget of the lab experiment. The full PSF is captured by recording images in a 5x5 grid. With this sampling neighbouring images now overlap significantly, and this overlap is use for more precise image registration between images. Both the CRED2 and the translation stages are controlled by a Matlab script that can capture images of the PSF. The background is estimated from the median of 100 images where the source is turned off. The background is captured before the image sequence, where for every position 10 images are averaged, and background subtracted.

6.7.1 The Point-spread function

The translation between neighbouring images are extracted using the scikit-learn `feature.register_translation` function on a masked PSF. Individual images are shifted accordingly and stored in a 3D array, where the first axis corresponds to the number of images and the other two to the x and y positions in the combined image. The final mosaic is the median along the first axis, which can be the single pixel value if a PSF region is imaged only once, or the median of multiple values when there is overlap. We remark that this method heavily relies on PSF stability, especially as fringes are the features used to align images with respect to each other. Any change to the fringe phase could lead to minor misalignment of images with respect to each other. We assume this effect is small as the overlapping region between neighbouring images contains multiple interferograms with different orientations. Any shift of the fringe phase in a single spot would not throw off the alignment. A passive setup can only generate changes in many closure phases simultaneously when optics move, which does not occur on timescales of two consecutive images. However, it demonstrates the sensitivity of fringe phases to the alignment of images with respect to each other. In addition, the PSFs are Nyquist sampled for the shortest wavelengths, meaning that even small subpixel shifts lead to large phase offsets. This is only a limitation of the lab setup, and will not affect the performance of HAM in the OSIRIS instrument where the PSF is fully captured by the camera.
We recorded images of the HAM PSF using three different amplitude masks. he different amplitude masks are the full HAM amplitude mask and two masks to isolate only the central and holographic components of the PSF. This allows us to analyze their individual PSFs and characterize the zero-order leakage of the HAM phase optic. The three masks and their corresponding PSFs at 1400 nm are presented in Fig. 6.13. The HAM PSFs closely resemble the simulated PSFs, and they confirm that the HAM optic was not damaged during the dicing process. The holographic component directly shows the zero-order leakage. Using `HCIPy`, we forward-model the HAM PSF with variable retardance. Changing the retardance changes the intensity of the central leakage term with respect to the holographic interferograms. By minimizing the normalized difference of the modelled and measured PSFs, we extract the retardance at different wavelengths. The fitting results are shown in Fig. 6.14 in addition to measurements of the efficiency of a polarization grating with the same liquid-crystal recipe. There is an offset of

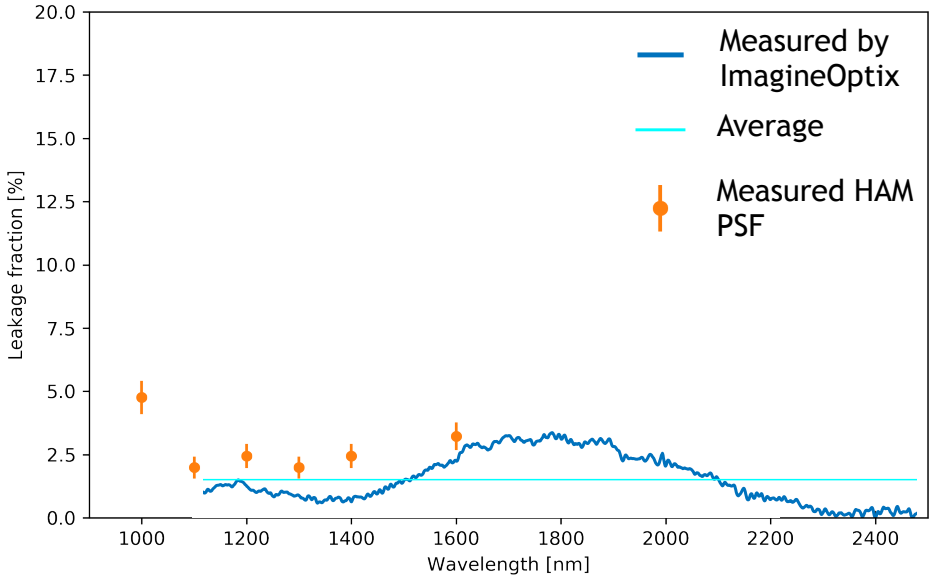

Figure 6.14: Measured zero-order leakage fraction of the HAM v1.5 optic.

roughly 1% between these two methods of measuring the zero-order leakage. A reduced diffraction efficiency for multiplexed gratings as compared to single gratings could explain this difference in zero-order leakage intensity. Multiplexing gratings with high frequencies leads to local phase patterns with extreme phase gradients, which might not be fully captured by the direct-write method. If that is the case, the diffraction efficiency is reduced. However, the presented measurements are unable to distinguish a change in diffraction efficiency due to multiplexing from a change in retardance. A consequence of the increased zero-order leakage is an increased sensitivity of the closure phases to polarized light, as discussed in Sect. 6.3.

6.7.2 Closure phases

We extract closure phases of both the central component and the holographic component of the full HAM PSF using our pipeline. In Fig. 6.15 we show the closure phases of the central component for 1400 nm, some with large deviations from zero. As explained in section 6.3 and in appendix **??**, these closure phases are non-zero due to the sensitivity to wavefront aberrations and the polarization state of the incoming light. Therefore, we fit a simple model that includes some low-order aberrations, a linear polarization fraction and polarization leakage to the closure phases. The low-order aberrations are represented with 10 Zernike modes starting with defocus. A monochromatic PSF is calculated using `HCIPy` with matching plate scale to the lab PSF, and our pipeline extracts the closure phases from the

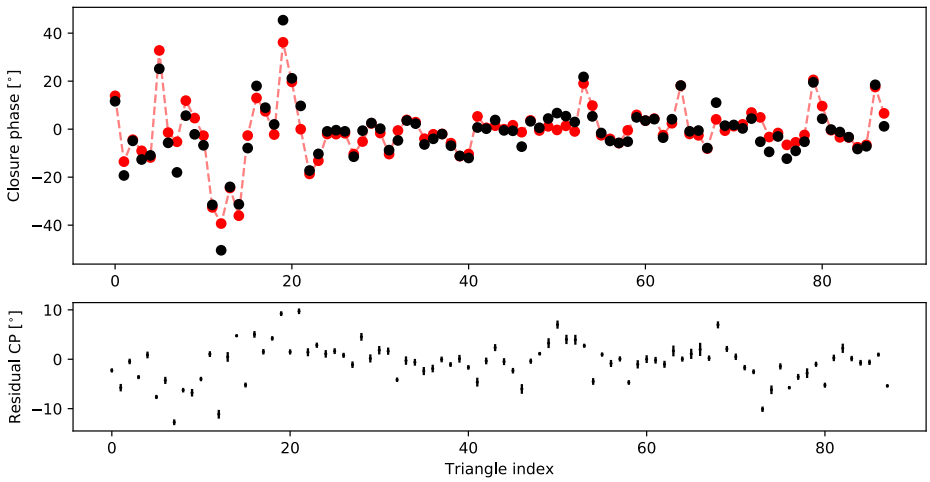

Figure 6.15: *Top:* closure phase of each closure triangle of the central component of the full HAM PSF at 1400nm (black), in addition to closure phases of a forward model (red). The forward model includes a linear polarization fraction of the source, polarization leakage and low-order aberrations. *Bottom:* The residuals between the model and observed closure phases. The error bars indicate the standard deviation of the residual closure phases during 22 hours of measurements at 15 minute intervals.

PSF. The best fit is shown in red in Fig. 6.15, and is capable of explaining almost all features of the data. The model has a linear polarization fraction of 7%, a polarization leakage of 3% and 47 nm RMS wavefront error, consisting mostly of defocus (38 nm RMS), astigmatism (16 nm RMS) and coma (20 nm RMS). Significant residuals remain, which could be explained by effects not present in our simple model, such as higher-order aberrations, inaccurate sampling of the model PSF or fast-axis deviations from $\pi/2$ in the phase-shifted subapertures.
The non-zero closure phases again showcase the sensitivity of HAM to aberrations and the polarization state of the incomming light. When the polarization state and wavefront aberrations are stable in time, the closure phases should stay constant. We address the stability in the lab by imaging the PSF 90 times during 22 hours, roughly 15 minutes apart. The camera will not return to exactly the same location in the 5x5 grid for every mosaic. Shifts in these positions bias the stability measurement significantly, inflating the error bars. To be less dependent on the registration of translation between images in the mosaic, we first register the translation between all images of a single camera position. We minimize the difference between these 90 images by aligning them all to the first frame in time. After all images of a single camera position are aligned with respect to each other, we repeat the process for all other camera positions. Then, we calculate the translation of the 25 camera positions with the first frames of the aligned images. Because all images are aligned, this offset is the same for all 90 mosaics. We generate the 90 mosaics by translating the aligned images with this general offset, and stitch them together like before.
From the PSF in each mosaic we extract the closure phases of the central component and the standard deviation of each closure phase is indicated by the error bars in the bottom plot of Fig. 6.15. Almost all closure phases have a standard deviation of less than one degree, indicating that under laboratory conditions they are stable. This is also true for the underlying cause of the offsets, i.e. the polarization fraction and the wavefront aberrations. Additionally, we extract the closure phases of the holographic component, and calculate the time variability. The results are presented in Fig. 6.16, showing two interesting features. First, there is a deviation from zero for these closure phases as well, and it is independent of the number of interferograms per closure phase. These closure phases are not affected by polarization leakage, as the holographic component is imaged far away from the central component. Second, there is a difference in time variability, e.g. the largest variation is 6 degrees compared to less than a degree for the first three closure triangles. The most stable closure triangles are the ones imaged in a single interferogram, i.e. closure triangle 0, 1, 2, 10, 14 and 18 in Fig. 6.16. This suggests that the time variability does depend on the number of interferograms per closure phase. Moreover, it suggests that the cause of the closure phase offset is different than the cause of the variability differences. Most likely, the difference in variability is not caused by effects of the optical setup as this would have impacted the closure phases of the central component as well. A possible explanation for the variability difference could be that there are still remaining translation registration errors for different interferograms in time.

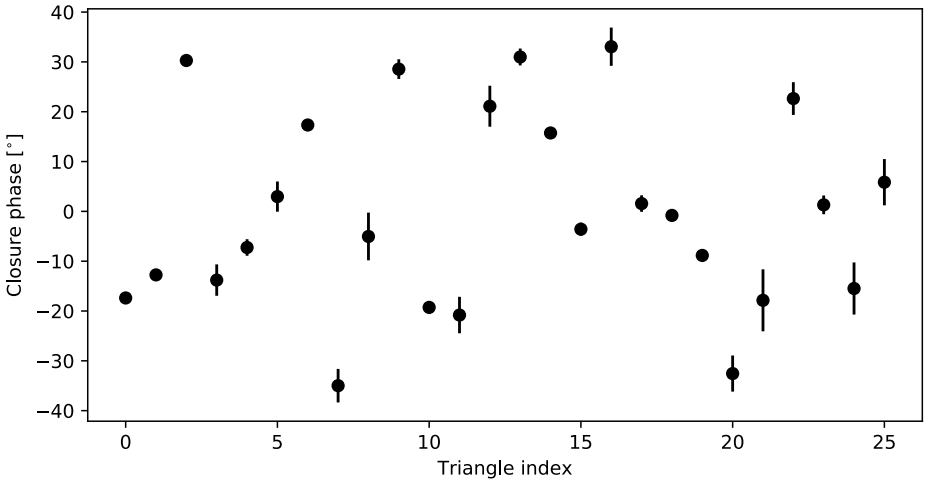

Figure 6.16: Measured average closure phases of the holographic component during 22 hours of measurements with 15 minute intervals.

On the other hand, the offsets from zero for the closure phases of the holographic component could be generated by the optical system. Light from each subaperture has a different optical path between the HAM optic and the focal plane. When two subapertures are imaged onto the same location in the focal plane, differences in optical path length introduces shifts in the fringe phase. These aberrations can differ for each baseline, such that combining the phases into a closure phase does not add to zero. Yet, as the optical setup itself is static, they are stable in time.

6.8 On-Sky Verification

We observed the binary HD 90823 (also known as HDS 1507 or WDS 10294+1211) and an unresolved reference star in two different filters. The primary goal is to verify whether the correct system parameters of HD 90823 (i.e. contrast ratio and separation) can be inferred from the data. Moreover, we aim to assess HAM's broadband performance by extracting wavelength-dependent closure phases from the holographic spots in the focal plane. In the simple case of a binary, the apparent angular separation between the two stars should not change as a function of wavelength.

6.8.1 Binary HD 90823

HD 90823 is an ideal verification target because the binary has a relatively low contrast $\Delta m \approx 1.2$ in the V and I bands, making the companion easy to detect, and, two sets of orbital elements have been published in the literature (Cvetković et al. 2016, Tokovinin 2017), allowing us to compute the predicted on-sky separation

vector at any point in time. In 2016, Cvetković et al. published a result that was based on four measurements spread over more than a decade. The authors found an orbital period of 23 years, yet they warn that their result is *"highly-tentative"* and that *"new observations are very desirable"*. Based on three new data points, Tokovinin *"radically revised"* the orbital elements in 2017 and lowered the binary's period to just over 15 years. Table 6.3 provides a further overview of the relevant HD 90823 parameters that are reported in both papers.

Table 6.3: An overview of different HD 90823 parameters as reported in Cvetković et al. (2016) and Tokovinin (2017). A and B refer to the binary components. For further details, please refer to the tables in the papers.

	Cvetković et al. (2016)	Tokovinin (2017)
Period (years)	23.361	15.59 ± 0.11
M_A ($M_\odot$)	1.66	1.53
M_B ($M_\odot$)	1.30	1.18
Contrast (Δm)	1.19 ± 0.17 (V)	1.21 ± 0.15 (I)
Spectral type	F0 (A) + F7 (B)	F2

6.8.2 Observations

The on-sky verification of the prototype HAM v1.0 test took place in the evening of 16 April 2019 (Hawaiian time). Observations were carried out in H band and time was divided equally among the binary HD 90823 and the calibrator source HD 90700. According to the SIMBAD database, the difference between the apparent magnitudes of both objects in H band is small: $m_H = 6.2$ for HD 90823 versus $m_H = 5.6$ for HD 90700. Both objects were observed in a narrowband filter (henceforth Hn5) and a broadband filter with a 20% bandwidth (henceforth Hbb). An observing log and the details of the filters are provided in Tables 6.4 and 6.5, respectively.

The OSIRIS Imager comprises a Teledyne Hawaii-2RG HgCdTe detector with a size of 2048×2048 pixels and a minimum integration time of 1.476 seconds (see Arriaga et al. 2018 for further specifications). Because this number exceeds the typical time associated with atmospheric seeing and the subapertures are larger than r_0 at the filter bandwidths, aperture masks in OSIRIS can only operate in conjunction with adaptive optics. Tests with an internal source showed that the imaging quality changed considerably as function of the input source location in the field of view. The separation of the HAM phase and amplitude mask of HAM v1.0 was considered as the source of this changing imaging quality. Light that passes through the holes of the amplitude mask at an angle to the optic axis then intercepts the phase mask off-axis, possibly illuminating the phase mask beyond the edges of the phase pattern of individual holes. If present, such a leakage term would affect the central component PSF. However, images of some sections the

[1] `https://www2.keck.hawaii.edu/inst/osiris/scale_filter.html`

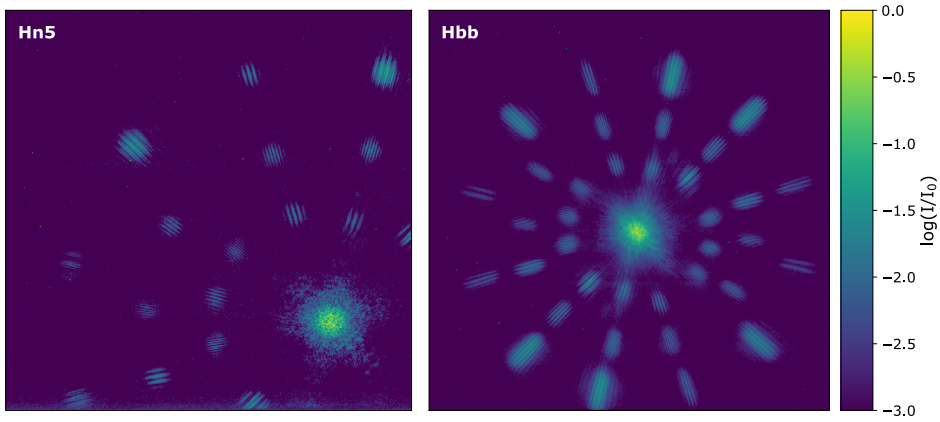

Figure 6.17: The star HD 90700 imaged with HAM in the Hn5 filter (left panel) and the Hbb filter (right panel), respectively. These figures display only a sub-window of the full sensor area. Data recovered for the Hn5 PSF occupied the bottom-right corner of the detector, so not all holographic spots are visible.

Table 6.4: A summary of the observations taken with HAM at Keck during the evening of 16 April 2019 (Hawaiian time).

Start (UTC)	End (UTC)	Target	Filter	N_f	t_{exp} (s)
06:28:52	06:36:30	90700	Hn5	50	1.476
06:39:05	06:46:42	90823	Hn5	49[a]	1.476
06:49:48	06:57:26	90823	Hbb	48[b]	1.476
06:58:56	07:06:47	90700	Hbb	50	1.476

[a] One corrupted frame was discarded.
[b] Two frames with a much lower signal-to-noise were discarded.

Table 6.5: The two filters that were used during the observations. Values were copied from the OSIRIS filter table on the Keck website [1].

Filter	PSF position	λ_0 (nm)	$\Delta\lambda/\lambda_0$ (%)
Hn5	Bottom-right	1765	4.9%
Hbb	Center	1638	20.1%

HAM mask pupil did not show partial illumination of holes. As the whole optic could not be imaged in pupil viewing mode, and we can not be certain that all holes were fully illuminated. Another explanation could be a differential focus between the OSIRIS imager and the OSIRIS spectrograph, which was present at the time of observation. Reconstruction of the visibility amplitudes also showed a gradient in the pupil illumination. The OSIRIS imager has been realigned and HAM v1.5 has been installed after the presented observations. New internal source

measurements with the HAM v1.5 mask show little variation of the PSF quality as function position in the FOV.
Because the image quality of the central component looked highest in the bottom-right corner of the detector, we decided to locate it there during the Hn5 observations (sacrificing the majority of holographic spots – left panel of Figure 6.17). However, with the holographic spots being of primary interest in broadband, the PSF was shifted to the middle of the detector for the Hbb observations (right panel of Figure 6.17). This allowed us to compute HAM's full set of closure phases, albeit with lower image quality.

6.8.3 Data Reduction

We apply a dark correction using 100 dark frames and calculate the power spectrum of the holographic component and the central component separately on a *uv*-grid of 3000×3000 pixels. For the Hn5 filter we mask everything but the central component, while for the Hbb filter we apply the same data reduction as mentioned in Sect. 6.5. Before the fringe phases can be extracted, we need to determine exactly *where* in the *uv*-plane the visibility has to be sampled. This requires two pieces of information: (i) the rotation angle of the PSF, which is set by the mask's orientation with respect to the detector, and (ii) the radial scaling of the PSF, which depends on the wavelength and the magnification of the instrument. To find the optimal parameter values that describe the scaling and orientation of the power spectrum (and thus of the PSF), we use a model of the *uv*-plane, that is cross-correlated with the observed power spectra as a function of radial scaling and rotation angle. We find that the angular orientations of power spectra of the holographic and central components differ by $\sim 2.1°$. This suggests that there is an angular offset between the amplitude mask and the phase mask in both filter wheels. The orientation of the central component's power spectrum is determined by the holes in the amplitude mask, while the orientation of the holographic component's power spectrum is determined by the slopes on the phase mask, which has to be taken into account when fitting a model to the observed closure phases.
We extract closure phases from the observed PSFs using the Fourier methods for both monochromatic and broadband observations.

Closure Phases

Once all visibility phases are sampled in the *uv*-plane, we compute the closure phases associated with each triangle on the mask. This gives rise to 197 sets of closure phases, each of which corresponds to one of the science frames listed in Table 6.4. In addition, we average the closure phases of the calibrator HD 90700 per filter, and we consider the standard deviation as measure of the stability of a closure triangle over time. We then calibrate the binary closure phases of each frame by subtracting the average closure phases of the calibrator.
Some representative results are displayed in Figures 6.18 and 6.19. Fig. 6.18 shows

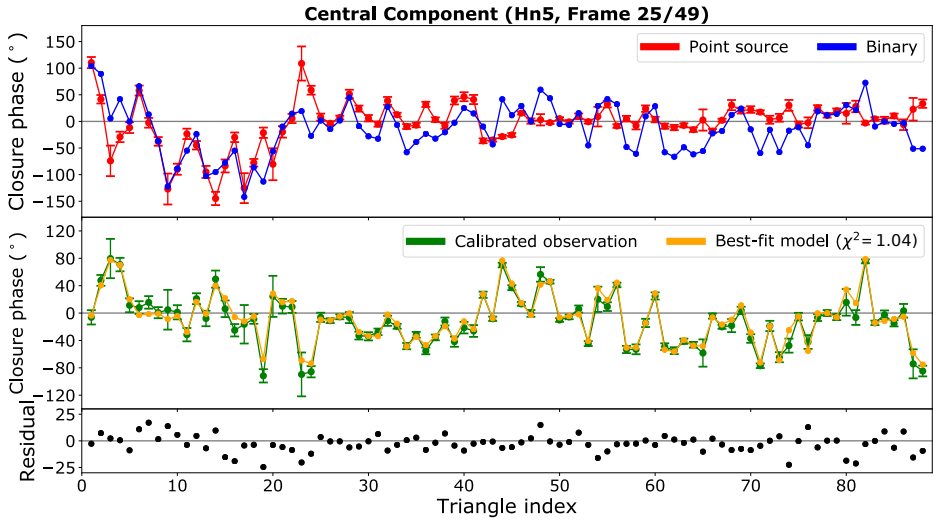

Figure 6.18: *Top:* closure phases of HD 90823 (binary) inferred from the central component in frame 25/49 in the Hn5 filter. The closure phases of HD 90700 (point source) are averaged over all frames. *Middle:* the calibrated observation, obtained by subtracting the closure phases of the point source from the closure phases of the binary. The best-fit model that was found is overplotted, with the corresponding χ^2 value reported in the legend. *Bottom:* residuals left after subtracting the model from the observation.

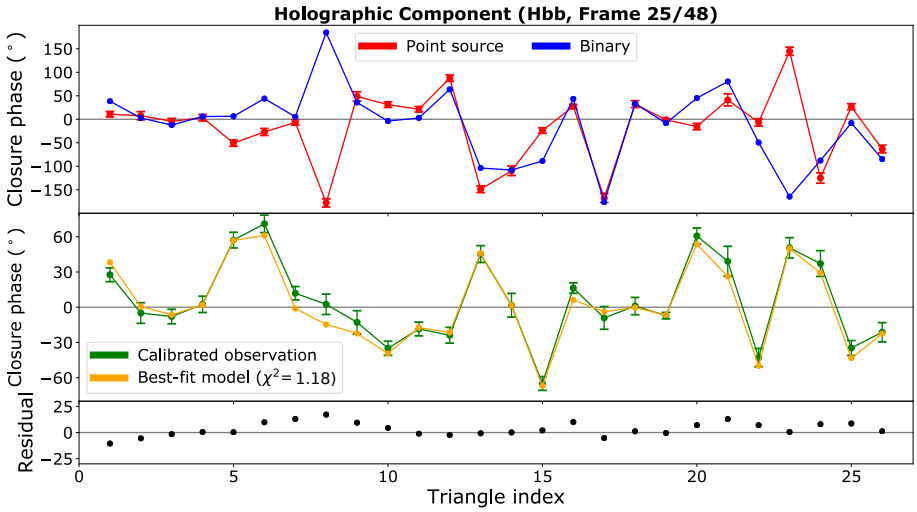

Figure 6.19: The same as Figure 6.18, but now for the **Holographic spots** in frame 25/48 in the Hbb filter. The *uv*-plane was sampled at 1638 nm.

the both the non-calibrated and calibrated closure phases of HD 90823 obtained from the PSF core in frame 25/49 in the Hn5 filter, as well as the best-fit model that was found. The average closure phases of the point source are plotted in the background. From Fig. 6.18 it is clear that the closure phases of HD 90700 (point source) exhibit a remarkably strong deviation from zero. The deviations are much stronger for HAM v1.0 on-sky than the offsets measured in-lab for HAM v1.5, as presented in the previous section. It is unclear if this is purely from the differences between HAM v1.0 and HAM v1.5. However, the overall structure of the closure phases agree with the lab measurements, with the first 24 closure phases (with triangle indices ≤ 24) of the central mask component having the largest deviations. Moreover, Fig. 6.18 suggests that the errors on these 24 triangles are also greater than the other errors, as they are more sensitive to changes in defocus. Even the PSF at the corner of the detector in the Hn5 filter has a poor quality, indicating that the PSF is strongly aberrated. Even with small fractions of linearly polarized light present, it is expected that the closure phase offsets are much larger under these conditions. This implies that calibration is essential to extract physical information from the data.

Fig. 6.19 shows the measured closure phases for frame 25/48 in the Hbb filter for the holographic spots at the central wavelength λ_0. Again, the closure phases for the single star deviate strongly from zero, and again much more than the lab measurements. We did not find a good explanation for this offset, but we can calibrate it for the binary system using the calibrator.

Table 6.6: The estimated separation and contrast ratio for HD 90823. A different set of closure phases applies to each row. The closure phases in the broadband Hbb filter were sampled at the central wavelength λ_0.

Subset	N_f	Separation ρ (mas)	Contrast ratio r
Central (Hn5)	38	121.9 ± 0.5	0.45 ± 0.01
Central (Hbb)	47	121.1 ± 0.8	0.45 ± 0.02
Holographic(Hbb)	46	120.9 ± 0.5	0.44 ± 0.02
Full HAM (Hbb)	93	121.0 ± 0.7	0.44 ± 0.02

Parameter Estimation

Given the measured closure phases and calibration of the zero points, the separation ($\boldsymbol{\rho} = (\rho_x, \rho_y)$) and the contrast ratio of HD 90823 can be estimated. We do this by finding the parameter combination $(r, \boldsymbol{\rho})$ that minimizes the chi-squared difference χ^2 between the observation and an analytical model of a binary system:

$$\chi^2(r, \boldsymbol{\rho}) = \frac{1}{N-m} \sum_{i=1}^{N} \left(\frac{[\phi_{i,\mathrm{obs}} - \phi_{i,\mathrm{mod}}(r, \boldsymbol{\rho}) + \pi]\%(2\pi) - \pi}{\sigma_i} \right)^2. \tag{6.11}$$

Here, N is the number of closure phases, m the number of free parameters, $\phi_{i,\mathrm{obs}}$ is the i-th observed closure phase (in radians, with error σ_i) and $\phi_{i,\mathrm{mod}}$ is the i-th modelled closure phase (in radians). Adding π, applying the modulo operator % and subtracting π makes sure that all differences are mapped onto the interval $[-\pi, +\pi]$. We perform a total of 145 fits: 49 in the Hn5 filter and 2×48 in the Hbb filter, whereby the closure phases obtained from the PSF core and the holographic spots are treated separately. Table 6.6 provides an overview of the HD 90823 parameters that were found after evaluating Equation 6.11 on a high-resolution grid. The reported values for $\rho = \sqrt{\rho_x^2 + \rho_y^2}$ and r result from averaging the best-fit parameters over different subsets of frames. Only fits for which $\chi^2 < 2N_{CP}$ (with N_{CP} the number of closure phases) are included to reduce the effect of outliers. We find the same contrast ratio $r = 0.45 \pm 0.02$ for each of the different subsets. The value inferred for the separation ρ is ca. 1 mas larger in the Hn5 filter as compared to the Hbb filter, which can be explained by random errors only.
As far as the performance of the mask's holographic component is concerned, it is reassuring that the retrieved parameter values in the Hbb filter are consistent with each other. The observables obtained from the central and holographic components of the mask are within 1σ.

6.8.4 Spectroscopic parameter retrieval

Holographic aperture masking has the unique capability to extract low-resolution spectroscopic closure phases using the holographic component. Here, we will look

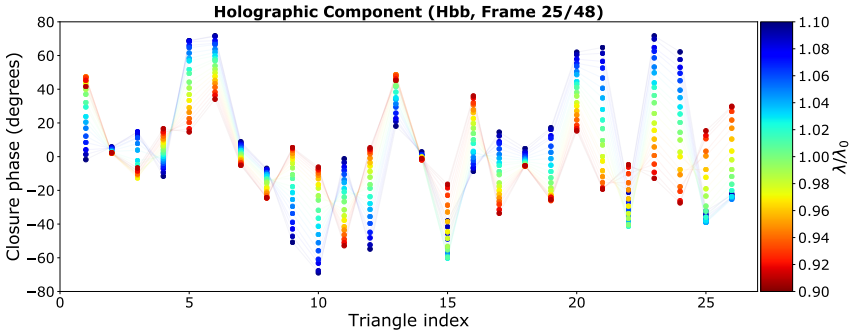

Figure 6.20: Calibrated closure phases obtained from the **Holographic spots** in frame 25/48 in the Hbb filter, plotted for different wavelengths within the 20% bandwidth (see colour bar).

at the multi-wavelength extraction of the closure phases, which allows us to extract the separation and magnitude from multiple wavelength channels in the Hbb band, spanning from 1473 nm to 1803 nm. We illustrate the wavelength-dependence of the measured closure phases in the Hbb filter in Fig. 6.20. The closure phases from the holographic spots are sampled at different coordinates in the uv-plane (corresponding to different wavelengths). The number of sampling points far exceeds the number of *independent* measurements. The latter, the spectral resolution, will be 2 or 3, depending on the triangle index, according to Eq. 6.2 and Eq. 6.3. The closure phases behave well and there seems to be a smooth transition from one wavelength to another.

We extract the binary parameter values by minimizing the χ^2 (Equation 6.11) for closure phases sampled at different points in the uv-plane. Fig. 6.21 and Fig. 6.22 display the parameters of HD 90823 that were inferred from the holographic spots in the Hbb filter as a function of wavelength. Fig. 6.21 shows the separation ρ between the binary components. Expressed in units of λ/D, the separation exhibits a $1/\lambda$ drop-off, which implies that ρ must be constant throughout the bandwidth. This is the expected result for a binary system, and is a powerful method to distinguish astronomical observables from instrument artefacts. As mentioned in Table 6.6, we find a separation of 121−122 mas based on the closure phases at λ_0. This is also the value that follows from averaging over all wavelengths in the bandwidth, as shown by the horizontal line. Some points deviate significantly from the average. The error bars are determined using a Jackknife method (Roff & Preziosi, 1994), which does not take into account systematic errors, e.g. wavelength dependent errors in closure phase retrieval as shown in Fig. 6.9. The measured brightness ratio r as a function of wavelength is plotted in the right panel of Fig. 6.22.

The spectral types of the HD 90823 components are uncertain. According to the SIMBAD database, both stars are of type F2, but the corresponding quality labels suggest little reliability. On the other hand, Cvetković et al. (2016) state that the

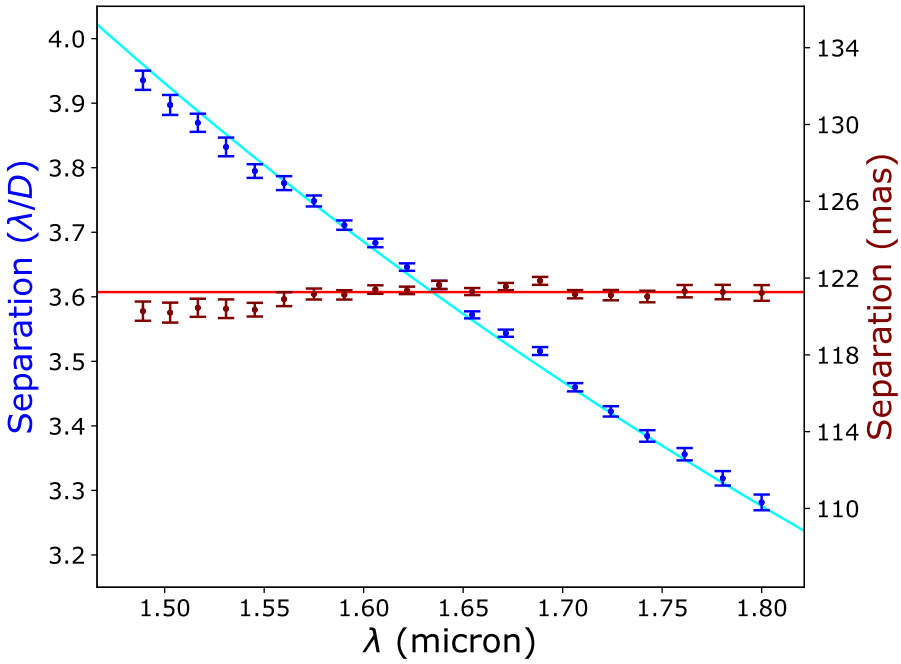

Figure 6.21: The separation between the components of HD 90823 inferred with HAM as a function of wavelength, expressed in units of λ/D (blue points) and expressed in milli-arcseconds (red points). The blue line is proportional to $1/\lambda$. The red line is a weighted average of the measured separations in mas.

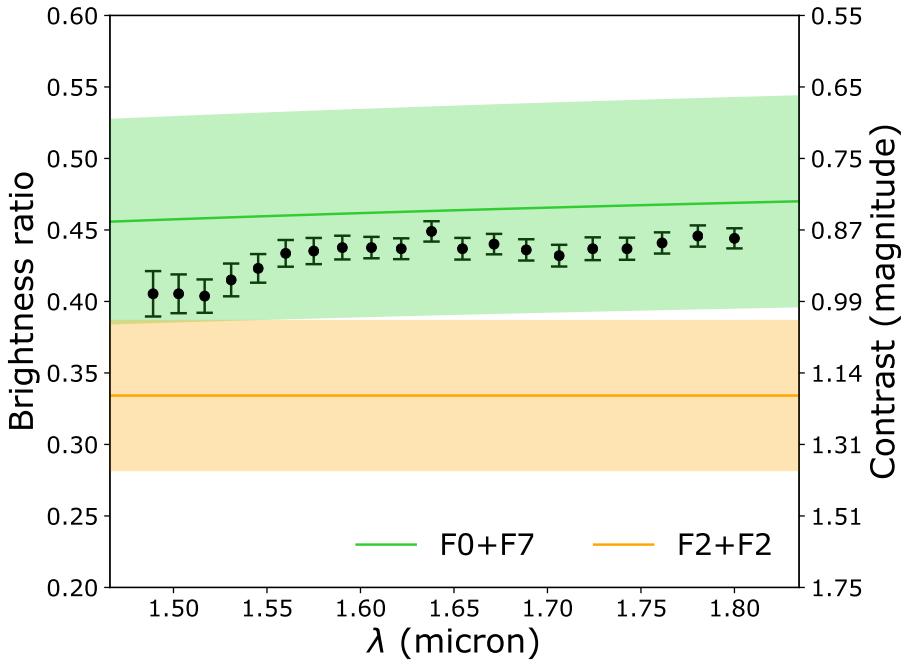

Figure 6.22: The recovered brightness ratio of the binary as a function of wavelength. The green line shows the expected brightness ratio for the spectral types (F0 and F7) reported by Cvetković et al. (2016). The orange line is the expected brightness ratio when both components are of type F2. The envelopes represent the 1σ error from the V-band and I-band measurements presented in table 6.3. The number of data points is much larger than the number of independent measurements.

bright component is hotter (type F0) than the faint one (type F7) with the difference in surface temperatures being roughly 1000 K. In order to find out which claim is most likely based on the HAM data, we computed the Planck spectra of the components in both scenarios, scaled them according to the reported contrast in V band (see Table 6.3) and divided them in the Hbb filter. Fig. 6.22 illustrates that the actual measurements lie in between the theoretical curves, but fully within the 1σ envelope of the F0+F7 scenario. Moreover, the gradient as function of wavelength is consistent with the F0+F7 scenario. This means that our estimate of the brightness ratio also suggests a difference between the components' temperatures and spectral types, in line with Cvetković et al. (2016).

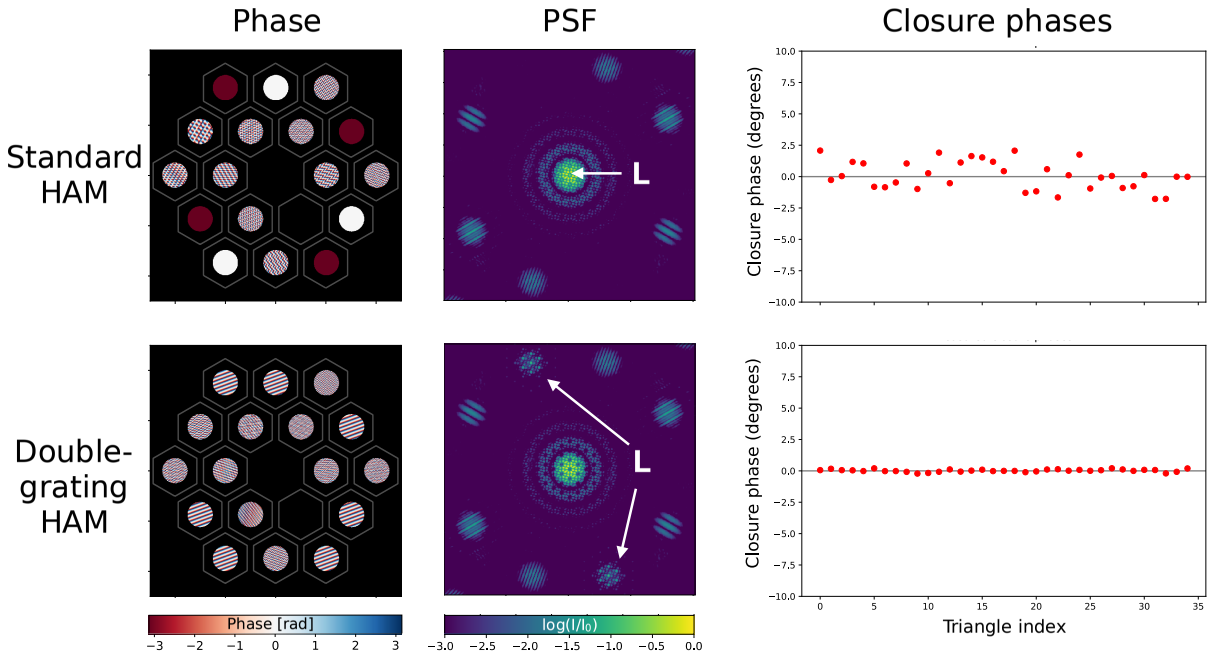

Figure 6.23: The difference between a standard HAM and a double-grating HAM. The first element of the double-grating is the standard HAM pattern with an added polarization grating pattern (= phase ramp) with 70 periods across the full pupil. The second grating is not shown. The resulting PSFs show that the polarization leakage is directed from the center into two off-axis PSFs. Both elements have 2% leakage and we assume a linear polarization fraction of 10%. Closure phase offsets due to polarization leakage are greatly reduced with the double-grating method.

6.9 Conclusions and Outlook

In the paper we show that holographic aperture mask show promise to empower a new generation of sparse aperture masking experiments: retaining the instrumental simplicity while adding spectroscopic capabilities, higher throughput and added Fourier coverage. We discuss the trade-offs in mask design, and how this was implemented for a prototype HAM mask in OSIRIS. A first version of this mask was installed at Keck OSIRIS in 2018. By observing the binary HD90823 with this mask, we obtain low resolution spectra of closure phases, confirming the broadband capabilities of the HAM mask. We investigated two limitations of this version. The first is the spatial separation between the phase and amplitude mask, resulting in spatially varying PSF quality. This is solved with an upgraded version, HAM v1.5, with a diced version of the same phase mask combined in the same holder as the amplitude mask, installed in OSIRIS in early 2020. Secondly, we find non-zero closure phases for a single star. We show in simulation that polarization leakage can produce these non-zero offsets in closure phases, and confirm this with lab demonstrations.
A future upgrade of the HAM mask can remove the effects of polarization leakage by several orders of magnitude suppression of the unwanted light using the double-grating method (Doelman et al., 2020). The double-grating method adds a phase ramp to the phase pattern, so that any polarization leakage travels in a different direction than the main beams. A second grating with the same phase ramp (= polarization grating) is installed directly after the first phase pattern, folding the two main beams back on axis. The polarization leakage term of the first grating, on the other hand, is diffracted away by this second grating, reducing the on-axis leakage by at least an order of magnitude. The location of the polarization leakage can be controlled with the phase ramp slope and direction, similar to the holographic interferograms. Placing the polarization leakage on an empty spot on the detector reduces the phase slope. Adapting the focal plane design of a HAM device to leave room for this leakage term would be beneficial. We simulate the performance of a HAM device with the second design presented in Fig. 6.7, assuming 2% leakage for both GPHs and a linear polarization fraction of 10%. The results are shown in Fig. 6.23. Using the double-grating method reduces the standard deviation of the closure phases from 1.1 degree to 0.1 degree. The residual closure phase pattern of the double grating HAM is not correlated with the standard HAM, which suggests that the deviation from zero is caused by different effect, e.g. inaccuracies in the data reduction. These simulations prove that a double-grating version of HAM would greatly reduce the impact of polarization leakage on the performance of HAM.
The potential of a double-grating HAM is exciting.

We outline a few scientific prospects that are enabled only by HAM. Studies of thermal emission from protoplanets in protoplanetary disks is complicated by disk features that can emulate exoplanet signals, e.g. light scattered by dust, that these systems can display (Kraus & Ireland, 2012; Sallum et al., 2015; Currie et al., 2019). However, spectral information can help with discriminating between disk

and planet signal (Haffert et al., 2019). HAM's simultaneous spectral and spatial measurements help constrain infrared spectral slopes, distinguishing between scattered light and thermal signals. Other opportunities are enabled by the higher throughput of a HAM mask compared to a SAM mask. Follow-up on the brightest *TESS* targets can help with ruling out background or binary contaminators. Monitoring brown dwarf binaries with HAM increases efficiency on determining orbits and dynamical masses, directly testing predictions for lithium burning, the stellar/substellar mass boundary, and substellar cooling rates (e.g., Dupuy & Liu (2017)). Suppressing the polarization leakage will be critical for improving closure phase stability, resulting in better contrast.

Acknowledgements

The research of David Doelman and Frans Snik leading to these results has received funding from the European Research Council under ERC Starting Grant agreement 678194 (FALCONER). Laser cutting of aperture masks and planar optics was carried out with the assistance of the OptoFab node of the Australian National Fabrication Facility, utilising NCRIS and NSW Gov. funding. We specifically thank Benjamin Johnston from OptoFab for the quick turn-around of the many amplitude masks and for laser cutting the HAM v1.5 mask. Some of the data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain. We thank John Canfield and Peter Wizinowich for the technical support. We also thank Michael Liu for the fruitful discussions, which helped improve the results presented in this work.

6.10 Appendix: Spectral resolution of the holographic interferograms

Here we will derive the spectral resolution of hologrpahic interferograms. We start by looking at a single baseline, $\mathbf{b}$, between two subapertures with a phase ramp with a period P and a direction $\hat{\mathbf{a}}$. Assuming no piston phase offset between the subapertures, the electric field is given by

$$M(\mathbf{r}) = \Pi(\mathbf{r}) \otimes \left[(\delta(\mathbf{r} - \mathbf{b}/2) + \delta(\mathbf{r} + \mathbf{b}/2))\, \mathrm{e}^{2\pi i \mathbf{a}\mathbf{r}} \right], \tag{6.12}$$

where $\otimes$ is the convolution operator and $\delta(\mathbf{r})$ is the Dirac delta function, $\mathbf{a} = P/D_{sub}\hat{\mathbf{a}}$, and $\Pi(\mathbf{r})$ defines the subaperture. For a circular subaperture with di-

ameter D_{sub},

$$\Pi(\mathbf{x}) = \begin{cases} 1 & \text{if } |\mathbf{x}| \leq \frac{D_{sub}}{2} \\ 0 & \text{otherwise.} \end{cases} \tag{6.13}$$

The PSF is then described by

$$p(\theta) = P(\boldsymbol{\theta}, \lambda) \left[2 + 2\cos\left(\frac{2\pi \mathbf{b}\boldsymbol{\theta}}{\lambda} \right) \right] \otimes \delta(\boldsymbol{\theta} - \lambda \mathbf{a}). \tag{6.14}$$

Here $\boldsymbol{\theta}$ and $\frac{\mathbf{r}}{\lambda}$ are the Fourier plane coordinates, and $P(\boldsymbol{\theta}, \lambda)$ is the PSF of the aperture function

$$P(\boldsymbol{\theta}, \lambda) = \text{Airy}\left(\frac{\pi D_{sub} |\boldsymbol{\theta}|}{\lambda} \right) \tag{6.15}$$

As the PSF location is directly proportional to $\lambda\mathbf{a}$ due to the grating, we have an independent measurement of a baseline phase when the shift is $1.22\lambda/D_{sub}$, i.e. the Raleigh criterion. Therefore we can define the spectral resolution as

$$R_{fp} = \frac{\lambda}{\Delta\lambda} = \frac{D_{sub}}{1.22P}. \tag{6.16}$$

Increasing the subaperture size and grating frequency yields a higher spectral resolution. For a subaperture with $D_{sub} = \frac{1}{10}D$ imaged at $100\lambda/D$, the period is $D_{sub}/10$ and $R \sim 8$.

The second spectral resolution of the holographic component is defined in the uv-plane. We calculate the uv-plane distribution $\hat{V}(\mathbf{f})$ with the Fourier transform of the PSF,

$$\hat{V}(\mathbf{f}) = (\Pi(\mathbf{r}) \star \Pi(\mathbf{r}))^{\frac{1}{2}} \otimes [\delta(\mathbf{f} - \mathbf{b}/\lambda) + \delta(\mathbf{f} + 2\delta(\mathbf{f}) + \mathbf{b}/\lambda)]\, \mathrm{e}^{2\pi i \mathbf{af}}. \tag{6.17}$$

We define the cross-correlation between the subaperture function $Pi(\mathbf{r})$, shifted by $\mathbf{b}/\lambda$ as a splodge. The location of the splodges change with wavelength and the shift depends on the length of the baseline. The phases of two splogdes of different wavelengths for the same baseline can be uniquely extracted when they are separated by D_{sub}. In principle this is possible for every SAM mask, however, to increase throughput many of these masks have large subaperture diameters and many baselines, such that even for a small bandwidth the splodges start to overlap. In the case of HAM, the holographic interferograms only contain a limited amount of baselines and the effect of overlapping splodges can be reduced by design. The spectral resolution in the uv-plane is given by

$$R_{uv} = \frac{\lambda}{\Delta\lambda} = \frac{|\mathbf{b}|}{D_{sub}}. \tag{6.18}$$

Bibliography

Arriaga, P., Fitzgerald, M. P., Johnson, C., Weiss, J. L., & Lyke, J. E. 2018, in Ground-based and Airborne Instrumentation for Astronomy VII, ed. H. Takami, C. J. Evans, & L. Simard, Vol. 10702 (SPIE), 103, doi: 10.1117/12.2313101

Beuzit, J.-L., Vigan, A., Mouillet, D., et al. 2019, A&A, 631, A155, doi: 10.1051/0004-6361/201935251

Carlotti, A., & Groff, T. 2010, in Ground-based and Airborne Telescopes III, Vol. 7733, International Society for Optics and Photonics, 77333D

Cheetham, A. C., Girard, J., Lacour, S., et al. 2016, in Optical and Infrared Interferometry and Imaging V, Vol. 9907, International Society for Optics and Photonics, 99072T

Chilcote, J. K., Bailey, V. P., De Rosa, R., et al. 2018, in Ground-based and Airborne Instrumentation for Astronomy VII, Vol. 10702, International Society for Optics and Photonics, 1070244

Currie, T., Marois, C., Cieza, L., et al. 2019, The Astrophysical Journal Letters, 877, L3

Cvetković, Z., Pavlović, R., & Ninković, S. 2016, The Astronomical Journal, 151, 83, doi: 10.3847/0004-6256/151/3/83

Cvetojevic, N., Huby, E., Martin, G., et al. 2018, in Optical and Infrared Interferometry and Imaging VI, ed. M. J. Creech-Eakman, P. G. Tuthill, & A. MÃ©rand, Vol. 10701, International Society for Optics and Photonics (SPIE), doi: 10.1117/12.2313262

Doelman, D. S., Por, E. H., Ruane, G., Escuti, M. J., & Snik, F. 2020, Publications of the Astronomical Society of the Pacific, 132, 045002

Doelman, D. S., Tuthill, P., Norris, B., et al. 2018, in Optical and Infrared Interferometry and Imaging VI, ed. A. Mérand, M. J. Creech-Eakman, & P. G. Tuthill, Vol. 10701 (SPIE), 12, doi: 10.1117/12.2313547

Dupuy, T. J., & Liu, M. C. 2017, The Astrophysical Journal Supplement Series, 231, 15

Escuti, M. J., Kim, J., & Kudenov, M. W. 2016, Optics and Photonics News, 27, 22, doi: 10.1364/OPN.27.2.000022

Foo, G., Palacios, D. M., & Swartzlander, G. A. 2005, Opt. Lett., 30, 3308, doi: 10.1364/OL.30.003308

Gardner, J. P., Mather, J. C., Clampin, M., et al. 2006, Space Science Reviews, 123, 485

Gauchet, L., Lacour, S., Lagrange, A.-M., et al. 2016, Astronomy & Astrophysics, 595, A31, doi: 10.1051/0004-6361/201526404

Gilmozzi, R., & Spyromilio, J. 2007, The Messenger, 127, 11

Greenbaum, A. Z., Pueyo, L., Sivaramakrishnan, A., & Lacour, S. 2014, The Astrophysical Journal, 798, 68, doi: 10.1088/0004-637X/798/2/68

Haffert, S., Bohn, A., de Boer, J., et al. 2019, Nature Astronomy, 3, 749

Haniff, C. A., Mackay, C. D., Titterington, D. J., et al. 1987, Nature, 328, 694, doi: 10.1038/328694a0

Haubois, X., Norris, B., Tuthill, P. G., et al. 2019, A&A, 628, A101, doi: 10.1051/0004-6361/201833258

Hecht, E., & Zajac, A. 1974, Reading, Mass, 19872, 350

Huby, E., Perrin, G., Marchis, F., et al. 2012, A&A, 541, A55, doi: 10.1051/0004-6361/201118517

Huby, E., Duchêne, G., Marchis, F., et al. 2013, A&A, 560, A113, doi: 10.1051/0004-6361/201321894

Jovanovic, N., Tuthill, P. G., Norris, B., et al. 2012, Monthly Notices of the Royal Astronomical Society, 427, 806, doi: 10.1111/j.1365-2966.2012.21997.x

Komanduri, R. K., Lawler, K. F., & Escuti, M. J. 2013, Optics Express, 21, 404

Kraus, A. L., & Ireland, M. J. 2012, The Astrophysical Journal, 745, 5, doi: 10.1088/0004-637X/745/1/5

Kraus, A. L., Ireland, M. J., Martinache, F., & Lloyd, J. P. 2008, ApJ, 679, 762, doi: 10.1086/587435

Lacour, S., Tuthill, P., Amico, P., et al. 2011, Astronomy & Astrophysics, 532, A72, doi: 10.1051/0004-6361/201116712
Larkin, J., Barczys, M., Krabbe, A., et al. 2006, New Astronomy Reviews, 50, 362
Macintosh, B., Nielsen, E., & De Rosa, R. 2019, AAS, 233, 104
Martinache, F. 2010, The Astrophysical Journal, 724, 464, doi: 10.1088/0004-637X/724/1/464
Mawet, D., Serabyn, E., Liewer, K., et al. 2009, Optics Express, 17, 1902
Miskiewicz, M. N., & Escuti, M. J. 2014, Optics Express, 22, 12691, doi: 10.1364/OE.22.012691
Monnier, J., Tuthill, P. G., Ireland, M., et al. 2009, The Astrophysical Journal, 700, 491
Nielsen, E. L., De Rosa, R. J., Macintosh, B., et al. 2019, The Astronomical Journal, 158, 13
Norris, B. R. M., Cvetojevic, N., Lagadec, T., et al. 2019, Monthly Notices of the Royal Astronomical Society, 491, 4180, doi: 10.1093/mnras/stz3277
Perrin, G., Lacour, S., Woillez, J., & Thiébaut, E. 2006, Monthly Notices of the Royal Astronomical Society, 373, 747, doi: 10.1111/j.1365-2966.2006.11063.x
Pope, B., Tuthill, P., Hinkley, S., et al. 2015, Monthly Notices of the Royal Astronomical Society, 455, 1647, doi: 10.1093/mnras/stv2442
Por, E. H., Haffert, S. Y., Radhakrishnan, V. M., et al. 2018, in Adaptive Optics Systems VI, ed. D. Schmidt, L. Schreiber, & L. M. Close, Vol. 10703 (SPIE), 152, doi: 10.1117/12.2314407
Roff, D. A., & Preziosi, R. 1994, Heredity, 73, 544
Ruane, G., Mawet, D., Riggs, A. E., & Serabyn, E. 2019, in Techniques and Instrumentation for Detection of Exoplanets IX, Vol. 11117, International Society for Optics and Photonics, 111171F
Sallum, S., & Eisner, J. 2017, The Astrophysical Journal Supplement Series, 233, 9, doi: 10.3847/1538-4365/aa90bb
Sallum, S., & Skemer, A. 2019, Journal of Astronomical Telescopes, Instruments, and Systems, 5, 1, doi: 10.1117/1.JATIS.5.1.018001
Sallum, S., Follette, K. B., Eisner, J. A., et al. 2015, Nature, 527, 342, doi: 10.1038/nature15761
Samland, M., Mollière, P., Bonnefoy, M., et al. 2017, A&A, 603, A57, doi: 10.1051/0004-6361/201629767
Sanders, G. H. 2013, Journal of Astrophysics and Astronomy, 34, 81
Soummer, R., Pueyo, L., Sivaramakrishnan, A., & Vanderbei, R. J. 2007, Opt. Express, 15, 15935, doi: 10.1364/OE.15.015935
Tokovinin, A. 2017, The Astronomical Journal, 154, 110, doi: 10.3847/1538-3881/aa8459
Tuthill, P., Monnier, J., Danchi, W., Wishnow, E., & Haniff, C. 2000, Publications of the Astronomical Society of the Pacific, 112, 555
Tuthill, P., Lloyd, J., Ireland, M., et al. 2006, in Advances in Adaptive Optics II, ed. B. L. Ellerbroek & D. B. Calia, Vol. 6272, International Society for Optics and Photonics (SPIE), 1064 – 1073, doi: 10.1117/12.672342
Tuthill, P. G. 2012, in Optical and Infrared Interferometry III, ed. F. Delplancke, J. K. Rajagopal, & F. Malbet, Vol. 8445 (International Society for Optics and Photonics), 844502, doi: 10.1117/12.925794
Tuthill, P. G. 2018, in Optical and Infrared Interferometry and Imaging VI, ed. A. Mérand, M. J. Creech-Eakman, & P. G. Tuthill (SPIE), 11, doi: 10.1117/12.2313145

Tuthill, P. G., Monnier, J. D., & Danchi, W. C. 1999, Nature, 398, 487, doi: 10.1038/19033

van Holstein, R. G., Girard, J. H., de Boer, J., et al. 2020, A&A, 633, A64, doi: 10.1051/0004-6361/201834996

Vigan, A., Gry, C., Salter, G., et al. 2015, Monthly Notices of the Royal Astronomical Society, 454, 129, doi: 10.1093/mnras/stv1928

Willson, M., Kraus, S., Kluska, J., et al. 2019, A&A, 621, A7, doi: 10.1051/0004-6361/201630215