UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Head Office: Università degli Studi di Padova

DEPARTMENT OF PHYSICS AND ASTRONOMY
Ph.D. course in Astronomy
Series XXX

A virtual coronagraphic test bench for SHARK-NIR, the second-generation high-contrast imager for the Large Binocular Telescope

Coordinator:
Prof. Giampaolo Piotto

Supervisor:
Dr. Jacopo Farinato

Ph.D. student:
Daniele Vassallo

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

SHARK-NIR is the second-generation high-contrast coronagraphic imager for the Large Binocular Telescope (LBT). In my Ph.D. project I have been involved in the conceptual and final design phase of the instrument. In specific, I developed a simulator in IDL language that operated as a virtual test bench to make a comparative study of several coronagraphic techniques identified as suitable candidates for implementation in the instrument. The simulator is based on physical optics propagation and adopts an end-to-end approach to generate images in presence of several sources of optical aberrations, from atmospheric residuals to telescope vibrations and non common path aberrations (NCPA). In particular, a big effort has been devoted to the optimization of the software efficiency through a dedicated parallelization scheme, to modelling of NCPA spatial and temporal properties, to the investigation of the effects of telescope vibrations and of the impact of the forthcoming upgrade of LBT Adaptive Optics system. I explored the coronagraphic performance in a wide range of observing conditions and characterized the coronagraphs sensitivity to aberrations, misalignments of optical components and chromatism. I also helped developing a data reduction pipeline to process simulated data adopting several algorithms. Simulations results have been used to define a final set of coronagrahic solutions that allow to fulfill the top-level scientific requirements.
Finally, I validated with simulations the phase diversity approach as a strategy for on-line sensing of NCPA. Simulations contributed to the final choice of the internal DM for both NCPA and fast tip-tilt correction.

SUMMARY

SHARK-NIR is the second-generation high-contrast coronagraphic imager for the Large Binocular Telescope (LBT). In my Ph.D. project I developed a simulator in IDL language that has been used as a virtual test bench to make a comparative study of several coronagraphic techniques.

This thesis reads as follows:

THE FIRST CHAPTER introduces the instrument. The science cases are presented, together with the opto-mechanical layout and the instrument modes. The chapter ends with the top-level scientific requirements.

THE SECOND CHAPTER describes the Fresnel simulator. The chapter enters into the details of image formation and describes the sources of optical aberrations that are implemented. Finally, the global architecture of the code is illustrated.

THE THIRD CHAPTER describes the data-reduction pipeline. All the steps leading from the raw output of the simulator to the extraction of performance metrics are illustrated.

THE FOURTH CHAPTER illustrates the coronagraphic designs. All the techniques implemented in the simulator are described, and the steps leading to designs optimization are detailed. The chapter ends with a comparison between all the techniques, from which the final selection has been made.

THE FIFTH CHAPTER shows instrument performance in some representative cases.

THE SIXTH CHAPTER reports the results of an investigation of chromatic effects and their impact on the coronagraphic performance.

THE SEVENTH CHAPTER reports the results of a tolerances analysis aimed at investigating the impact on performance of misalignments of optical components.

THE EIGHTH CHAPTER describes the application of the phase diversity approach for on-line sensing of NCPA.

1 THE SHARK INSTRUMENT

SHARK is an instrument proposed for the Large Binocular Telescope (LBT) in the framework of the "2014 Call for Proposals for Instrument Upgrades and New Instruments" (Farinato et al. 2015). It is composed by two separate cameras, a visible and a near infrared channel, to be installed one for each LBT telescope, and it will exploit, in its binocular fashion, unique challenging science from exoplanet to extragalactic topics with simultaneous spectral coverage from B to H band, taking advantage of the outstanding performances of LBT binocular Extreme Adaptive Optics (XAO) capability. The XAO features two Adaptive Secondary Mirrors (ASM) and pyramid wavefront sensors which will allow, especially after the forthcoming SOUL upgrade, unprecedent sensitivity in the faint-end regime.
The visible channel (SHARK-VIS) will operate between 0.5 μm to 1.0 μm, while the near infrared one (SHARK-NIR) will cover the range from 0.96 μm to 1.7 μm.
My Ph.D. project is centered on the NIR channel. In this first chapter I will present the camera: section 1.1 presents the science cases, section 1.2 describes the instrument opto-mechanics, section 1.3 presents the foreseen instrument modes to conclude with the top-level scientific requirements (section 1.4).

1.1 THE SCIENCE CASES

SHARK-NIR aims at addressing several fundamental, but as yet unanswered, topics in modern astrophysics, covering a broad context ranging from the exoplanet detection and characterization, to the study of the jets and circumstellar disks in very young stars, up to distant AGN and QSO. It is noteworthy that the science cases discussed here are only some of the important issues that can be addressed by exploiting SHARK-NIR, but many more topics can be investigated. Most important, the science questions we propose to answer provide us with the utmost synergy of SHARK-NIR with the existing (LMIRCAM) and forthcoming instrumentation at LBT (SHARK-VIS), resulting in a very powerful tool that is not currently available for other facilities in the world in terms of simultaneous spectral coverage from optical to mid-infrared.

1.1.1 Exoplantes: where is SHARK competitive?

The direct detection of extra solar planets is one of the most exciting goals in modern astrophysics. It provides us with access to information on giant gaseous planets outside the Solar system, that is planets similar to ours Jupiter, Saturn, Uranus and Neptune. At present, the atmosphere of giant planets in external orbits can be probed only by direct imaging techniques. In fact, the combination of radial velocity and transit methods, the other way to fully constrain the planetary properties, can only target exoplanets close to their hosting star, while astrometry, although powerful in detecting giant planets at distance between 0.1 and few tens of AU from the star, gives no hints on the composition and structure of the atmosphere of the detected planets.
Due to the very demanding resolution and contrast required on the scientific images, direct detection of extra solar planets is a very challenging task to be achieved. The available database of known planets at distances larger than 1 AU includes about two tens of young sub-stellar objects of mass larger than that of Jupiter, all detected via direct imaging. This number will increase in the near future thanks to the contributions of the ESO SPHERE and Gemini GPI instruments. In the following, I will discuss all the fields that can be covered with SHARK, with particular attention to the ones in which we can give a unique contribution.

Planets in Star-Forming Regions

The identification and the physical properties of planets in their very early stages will allow to obtain decisive inferences on the formation mechanisms (core accretion, disk instability, disk fragmentation), along with studying the interactions between planets and disks. Indeed, several features observed in circumstellar disks, such as spiral arms, gaps, and warps are possibly linked to the gravitational perturbation of planets. A few very interesting cases are starting to emerge (see, e.g. Quanz et al. 2015).
In this context, SHARK may allow detecting planets in formation in the well-studied Taurus-Auriga star forming region. This region is northern and hence badly accessible with SPHERE and GPI. The bulk of the members have ages of about 1-2 Myr, at a distance of about 140 pc. About 350 members were identified, 130 of which brighter than $R = 15$. Recently a survey has been performed by Bonavita et al. 2014 using Gemini-N, yielding the identification of a planet candidate. A detection in the gap of the circumstellar disk around LkCa 15 has also been proposed by Kraus and Ireland 2012 as a possible planet caught in the formation phase. The search for planets in star-forming regions represents a program capable of fully exploiting the potential of SHARK@LBT as proposed. The NIR channel will be used to reveal planet thermal emission. The visible channel could be used for narrow band imaging in Hα to reveal signatures of accretion from their circumplanetary disks. The simultaneous observations uniquely allowed by LBT would allow

to highly mitigating the issue of variability. Solar-type stars in Taurus-Auriga star-forming region represents targets for which LBT is uniquely competitive (faint magnitudes, northern declination). Considering its distance (140 pc), this science case motivates the instrument design with an IWA as small as 100 mas.

Giant planets around low-mass stars

This is a special niche for LBT AO at faint magnitude, especially with AO upgrade SOUL. Several members of young moving groups (age $\sim 10-100$ Myr) were recently identified (see e.g., Schlieder et al. 2012). There are several tens of potential targets, depending on the exact magnitude limit of the instrument, accessible for a deep search for planets in wide orbits, thus complementing results by SPHERE and GPI, which will be mostly limited to solar-type and early-type stars. In particular, with the current limit at $R< 10.5$ our sample comprises 33 targets (basically FGK-type stars), whereas adopting $R = 12.5$ as a magnitude limit we would gain more than a factor of three in sample size (that is 108 objects).

Brown dwarfs in young Open Clusters

Brown dwarfs (BDs) were originally proposed as a distinguishable class of astrophysical objects, with intermediate mass between stars and planets. During the last decade there was an increase in our knowledge of the bottom of the stellar main-sequence and of the low-mass stellar and sub-stellar (BDs) population of the solar neighborhood. Two new classes have been added to the spectral type sequence following M: L and T (Kirkpatrick 2005). T dwarfs with effective temperatures as low as $\sim$ 500 K are now known and we are finding objects that provide the link between the low-mass stars and giant planets.
The spectra of these objects are replete with atomic and molecular absorptions. Most dominant among these are lines of neutral alkali metals and bands of metal oxides, metal hydrides, water, methane, and ammonia. For L dwarfs, the spectral shapes cannot be matched unless dust grains, produced by the condensation of atomic and molecular species, are suspended in the photosphere. On the other hand, those of mid and late T dwarfs cannot be properly produced unless this dust is assumed to have vanished from the atmosphere. Recently (for a review see Kirkpatrick 2013) colder BDs were discovered by WISE (NASA's Wide-field Infrared Survey Explorer) giving origin to a new spectral class: the Y dwarfs. The boundary between T dwarfs and Y dwarfs roughly coincides with the location where the J-H colors of BDs, as predicted by models, turn back to the red at effective temperatures below $\sim$ 400 K.
In the field of star formation, these coldest brown dwarfs contain an historical record of the formation process at very low masses and at epochs many

billions of years before the stellar birthplaces we see today. Recently a lot of infrared surveys and high contrast observation have unambiguously revealed planetary masses objects isolated in the field or as wide companions to stars. Their existence confirms that the formation mechanisms proposed to form stars can actually form objects down to the planetary mass regime. Theoretically, there have been two main competing hypotheses regarding the formation of gas-giant planets: core accretion and disk instability. In the core accretion model, relatively small giant planets, such as Jupiter and Saturn, are thought to form at about 10 AU or less from a solar-type host star in several Myr. On the other hand, in the disk-instability model, planets of a few to 10 Jupiter masses can be created within a few 10 to 100 AU from the central star on a dynamical timescale of several thousand years. These formation models therefore predict two populations of giant planets segregated by orbital distance, with the closer planets formed by core accretion and the outer ones by disk instability, showing that stellar and planetary mechanisms overlap in the substellar regime. They can both lead to the formation of planetary mass objects including companions to stars and BDs. Distinct statistical properties such the occurrence, the mass, and the main orbital parameters, should help to identify the dominant mechanism for forming substellar companions. In these imaging studies, objects belonging to moving groups and local associations could be preferred targets in order to clarify this topics. First of all because these associations are nearby (20–100pc) and young (several to several hundred Myr) so their substellar objects (planets and BDs) are relatively bright, and thus easier to detect by direct imaging. Furthermore, stars in the same cluster have similar ages and distances from Earth, which statistically improves the accuracy of the age and luminosity estimates, and hence the derivation of the mass. Pleiades is the richest of these clusters with its about 800 known members. It is $\sim 133 \pm 1.2$ pc distant form Sun and it is 125 ± 8 Myr old (e.g., Ann et al. 2007).

Direct imaging offers the possibility to detect self-luminous companions placed on wide orbits (> 1 AU), to characterize their atmosphere and, for some cases, constrain their orbital parameters. Nearby ($d < 100$ pc) and young (age < 150 Myr) BDs represent observational niches for the direct imaging search of planetary mass companions. At these ages, companions are still hot and large. They can therefore be detected at favorable contrasts in the near-infrared. Given the characteristics of LBT AO system with respect to those of both SPHERE and GPI, we expect $1-2$ magnitude fainter targets to be reachable, and this will allow to observe many more nearby small mass stars and solar type stars in star-forming regions at close distance. SHARK instrument will allow to study with LBT the earliest evolution of sub-stellar objects and their link with the circumstellar disks, then providing specific clues to the formation mechanism and in the particular case of brown dwarfs also in the understanding the L-T transition.

Astrometry

Direct imaging provides both spectrum and position of sub-stellar companions. Astrometric follow-up allows to constrain the total dynamical mass for short-period systems ($\leqslant$ 10 AU typically) and, if combined with radial velocity data, the individual masses of the components of a system. Most of the directly-imaged planets and brown dwarfs have long orbital periods, so that their mass is usually inferred by comparing the measured luminosity to atmospheric and evolutionary models, assuming that the age of the object is equal to the age of the host star. However, these models are poorly constrained for young and low-mass objects (e.g., Burrows et al. 1997) because of the lack of calibrators. Thus, it is critical to detect objects for which the mass can be derived independently for testing and calibrating the models. These objects can be used to discriminate between different formation scenarios («hot-start», «warm-start», and «cold-start» models), by constraining the initial entropy at the epoch of their formation (Marleau and Cumming 2014). Short-period systems are especially relevant for this purpose.
There are two kinds of short-period systems relevant for astrometric monitoring. Low-mass binaries refer to binary systems with components of similar masses, like brown-dwarf binaries. In the last years, dynamical masses have been measured for a growing number of these systems (see e.g. Michael C. Liu et al. 2008). It has been shown evidence for luminosity problems for several systems, suggesting missing physical mechanisms in the atmospheric and evolutionary models of cloudy brown dwarfs with dust-rich atmospheres. For instance, over luminosity problems by factors of about two were reported for the L-dwarf systems HD130948BC and Gl 417BC (Dupuy et al. 2009; Dupuy et al. 2014). An explanation could be that their atmosphere is composed of patchy clouds instead of uniform cloud layers. The second type of systems is composed of a young primary star and a brown-dwarf or giant planet companion. M. C. Liu et al. 2002 reported the discovery of an L-dwarf companion to the active solar analog HR7672. Astrometric follow-up combined with radial velocity measurements allowed to constrain the orbital elements and the dynamical mass of the brown dwarf companion to $65.6 - 71.1$ Jupiter masses at 68.2% confidence level. A giant planet candidate was discovered around the young star β Pictoris ($\sim$ 20 Myr) in 2009 (Lagrange et al. 2009). Hot-start models suggest a mass of $\sim$ 12 Jupiter masses, while «cold-start» models indicate a mass in the brown-dwarf regime. Astrometric monitoring and radial velocity follow-up allowed to constrain the mass of the planet candidate to below 20 Jupiter masses with $\geqslant$ 96% probability (Bonnefoy et al. 2014). Its bolometric luminosity and the constraints on its dynamical mass are only reproduced by warm-start and hot-start models. The contribution of SHARK-NIR astrometry to the measurement of dynamical masses of directly imaged companions around bright young stars in the Northern hemisphere is best seen in connection to the Gaia mission (e.g., Sozzetti et al. 2016). Its presently envisaged instrumental configuration should warrant a per-measurement astrometric performance comparable to that of SPHERE (3-5 mas). SHARK-NIR will monitor directly imaged young

massive planets and brown dwarfs with the objective of detecting the orbital motion of the companions, and the combination with Gaia astrometry of the primaries will allow for very tight constraints to be placed on the actual masses of the imaged objects. Alternatively, Gaia discoveries of long-period massive companions to bright young stars in the Northern hemisphere (e.g. in Taurus-Auriga) will be followed-up with SHARK-NIR to obtain direct images of the companions and thus infer with good statistical confidence a relatively narrow range of likely masses (or tight upper limits in case of no detection).

Spectroscopic characterization of known planets

The physical characterization of the known planets is becoming increasingly important as planet discoveries cumulate. One way to distinguish between young and old brown dwarfs is to look for gravity-sensitive spectral lines. The radius of field brown dwarfs varies only slightly with mass and age, and therefore the surface gravity ($g \sim GM/R^2$) is determined by the mass ($\log g \sim 5$). The radius of the young objects, on the other hand, can be as much as three times greater than their eventual equilibrium state. As a result, young objects can exhibit significantly lower surface gravities (10–100 times) than the more massive evolved dwarfs of the same spectral type. Gorlova et al. 2003 showed that the neutral potassium (K I) lines in the J-band are very sensitive to surface gravity. These features might be investigated taking medium-resolution spectra for planets orbiting HR8799 using SHARK-NIR LSS mode. Furthermore, LSS observations would allow to shed light on L-T transition and on the characteristics of brown dwarfs and giant planets, which are expected to somewhat overlap but also significantly differ in terms of chemistry of the atmospheres and mechanisms of clouds formation. Kirkpatrick 2005 noted that optical spectral type and effective temperature (T_{eff}) show a tight correlation throughout the range of L dwarfs, but the correlation is broken at early T. On the other hand, in the NIR the behavior is different: T_{eff} and spectral type are well correlated only from early to mid-L, then we have an almost flat trend at $T_{eff} \approx 1400$ K (although the scatter is quite large); finally, cooler T dwarfs show again the correlation. A discrepancy between optical and infrared spectral types for L dwarfs was already highlighted by Stephens 2003. The author in this case suggested that whereas optical type was primarily a proxy for temperature, NIR is more influenced by clouds (and possibly gravity). The presence of condensate clouds is one of the most unique features of the ultra-cool atmosphere of directly imaged planets and BDs. The re-emergence of the 0.99 μm FeH feature in early- to mid-T spectral type has been suggested as evidence for cloud disruption where flux from deep, hot regions below the Fe cloud deck can emerge. The same mechanism could account for color changes at the L/T transition and photometric variability. This molecular feature along with CH4 bands (the hallmarks of T spectral types, see figure 1) will be covered from SHARK-NIR LSS modes.

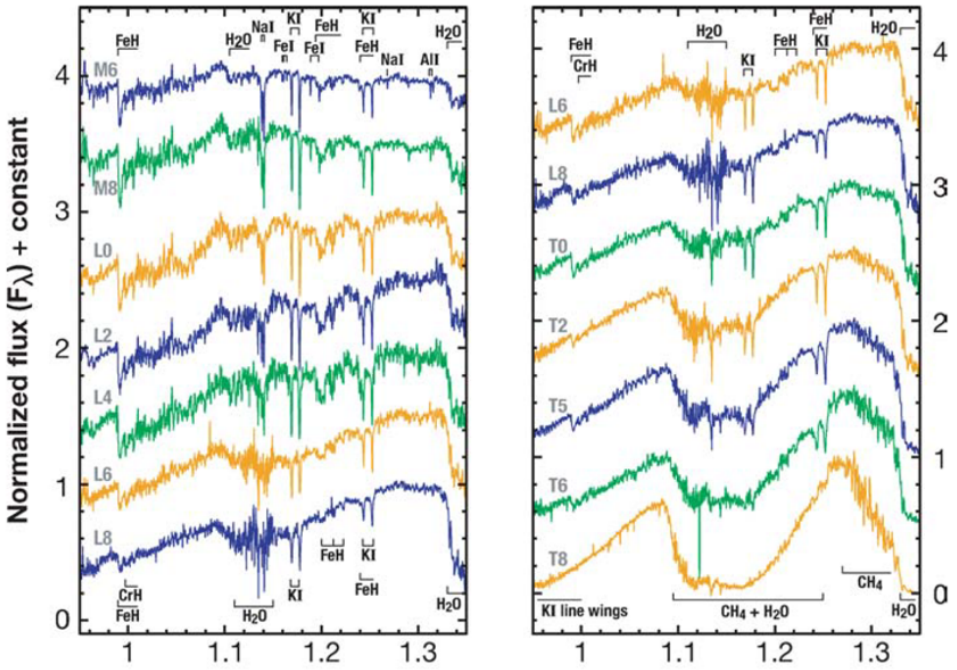


Figure 1: Example of spectra for M, L, and T objects. The spectral diagnostics, such as FeH, H2O and CH4 are labelled (from Kirkpatrick 2005).

Detecting very cool sub-stellar objects via dual-band imaging techniques

Simultaneous dual imaging (SDI) was first proposed by Racine et al. 1999 for the detection of faint companions and then extensively studied. This technique relies on the fact that planetary objects have strong molecular features in their spectra, whereas the host star has a relatively flat spectrum. By simultaneously acquiring two images of a system at two close wavelengths located around one of these sharp features and subtracting them, the star contribution can be partially eliminated, and the planet signal revealed. SDI is most effective when used for detecting cool companions that show deep molecular absorption bands caused by e.g., H2O, CH4 and NH3 at low T_{eff}, according to state-of-the-art atmosphere models of planetary-mass objects. With carefully selected filter pairs, a contrast of several magnitudes on the planet flux between the two filters can be obtained. The SDI method is straightforward: the images taken at λ_1 need to be spatially rescaled to account for the spectral dependence of the PSF and subtracted from the images at λ_0, with a possible amplitude correction factor to minimize the residual speckle noise. The implementation in SHARK-NIR of a dual band imager with purposely designed filter pairs (e.g., H2 and H3 centred at 1593 and 1667 nm, respectively) will provide us a very powerful tool to reveal very cool targets, such as e.g., T type sub-stellar objects that are characterised by strong absorption in the CH4 band.

1.1.2 Disks and jets around young stars

The circumstellar environment of a young star where planets are believed to form consists of an accretion disk and the often associated stellar jets. The study of proto-planetary disks is fundamental to comprehend the formation of our own Solar System as well as of extra-solar planetary systems. On the other hand, the study of the jet launch mechanism allows us to investigate the processes related to the transfer of angular momentum in the system and the feedback on the inner AUs. SHARK in the NIR and/or in combination with the VIS channel will offer a unique opportunity to investigate these phenomena, allowing:

- **High-contrast coronagraphic imaging of disks**. Optical and NIR images with an angular resolution of 20-40 mas and a pixel scale of 5-20 mas/pixel, combined with a coronagraph giving contrast ratios of $10^{-4} - 10^{-5}$, will allow to probe the small (μm-size) dust grains in the disk through scattered light imaging down to a few AU from the star. This is crucial to probe the disk structure and planetary dynamics in the planet-forming region (<30 AU). For example, the detection of asymmetries in the spatial distribution of dust, such as warps, gaps, spirals, and rings (see e.g. figure 2, left-hand panel) provides evidence for unseen co-orbital planetary-mass companions through their dynamical interactions with the disk grains (e.g., Garufi et al. 2013). Moreover, the comparison with millimeter interferometric images which probe the large (mm-size) grains can highlight the presence of dust traps which allow dust grains to grow beyond the so-called "radial drift and fragmentation barrier" (~ 1 m at 1 AU from the star) and to form planetesimals and hence planets (figure 2, right-hand panel).

- **Coronagraphic imaging of stellar jets**. Stellar jets emit in atomic and molecular lines, originated in the internal shocks, with almost no continuum. AO observations of the base of jets disclose a clear picture of the jet very close to the launching point and of the instabilities and shocks in the flow (figure 3). Conversely, measurements of jet opening angle and collimation scales can give stringent constraints to the launch mechanism and the consequent removal of angular momentum from the disk. A Coronagraph with an IWA of $150 - 200$ mas and modest contrast in J and H bands would constitute a big leap in the observational modes available today: it would allow, in principle, to sample the jet as close as 10 AU from the source in nearby systems like Lupus and Taurus (10^{-2}–10^{-3} contrast is sufficient to image the bright jet nebulae). Atomic jets in the IR are mainly seen in [Fe II] lines: narrow band imaging in the bright 1.64 μm and 1.25 μm lines can be used to efficiently explore the reddened inner jet regions. Finally, the molecular counterpart of the jet, which is particularly bright in the H2 line at 2.12 μm, will be accessible at LBT after LMIRCAM will be upgraded to provide K-band coronagraphic imaging.

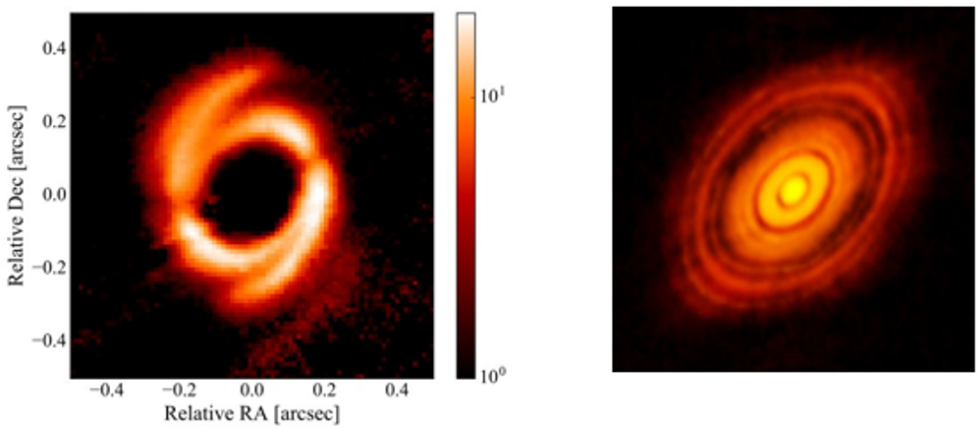


Figure 2: *Left-hand panel*: Spiral patterns in the disk around HD 100453 in J-band polarized emission obtained with the dual-band imager IRDIS on VLT/-SPHERE (Wagner et al. 2015). Right-hand panel: ALMA image of the disk around HL Tau, presenting dark rings attributed to the presence of unseen planets (ALMA partnership et al. 2015).

- **Long-slit spectroscopy of jets**. A fundamental complementary investigation with SHARK is to observe the jets with the planned LSS mode in the NIR combined possibly with a spectrograph in the VIS channel. The interesting region of the base of the jets from sources located in Taurus has typical size of 0.7x1.5″, thus well within the instrument FoV. The planned spectral resolution R = 700, although not allowing to resolve internal kinematics, is enough to separate the different emission lines and study diagnostic line ratios.

1.1.3 Extragalactic science

Also in the extra-galactic field the unique capabilities of SHARK in terms of spatial resolution and contrast enhancement may be successfully applied. Both VIS and NIR channel are required, as well as both imaging and coronagraphic modes.
In particular, SHARK will allow to study in great detail:

- **AGN feedback on nearby host galaxies**. Near-infrared sub-arcsec resolution images will allow us to map dust lanes on scales down to one hundred pc, and to investigate whether outflows are dusty or rather the AGN feedback has already swept the ISM. Moreover, color maps and IFU follow-up of star-forming regions in the galaxy nucleus and disk will allow to constrain the galaxy SFR, age, and metallicity. Figure 4 shows the state of the art LBT AO-assisted imaging of extragalactic sources and demonstrates the feasibility of AO observations guided by AGN with the current LBT capabilities.

- **Dumped Ly-α systems (DLAs)**. These systems are identified in the spectra of far background quasars and have a column density of neu-

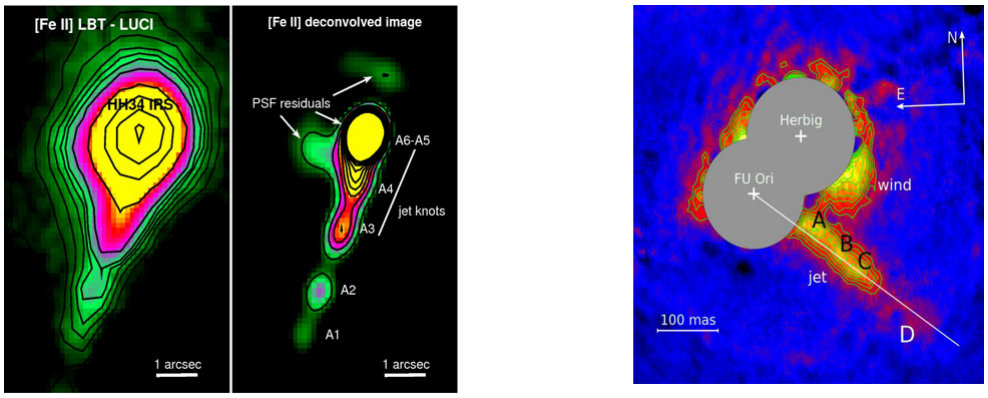


Figure 3: *Left-hand panel*: seeing-limited LBT-LUCI observations of HH34 jet in a narrow band [Fe II] 1.64 μm line filter, before and after PSF deconvolution (Antoniucci et al. 2014). *Right-hand panel*: SPHERE-ZIMPOL observations of the jets emanating from the binary system Z CMa in the [OI] 6300 Å narrow-band filter (Antoniucci et al. 2016). The position of the two young stars of the system (an Herbig star and a FU Ori object) are marked with a cross. Areas heavily corrupted by artefacts have been masked.

tral hydrogen in excess of $2 \cdot 10^{20} \mathrm{cm}^{-2}$. Hence, they are dense clouds most likely to be found within the optical extent of galaxies, within a few kpc ($< 1-2''$) from the centre, e.g. as part of their disk or as dense in-falling clouds. Direct rest-frame optical imaging of the galaxies associated with the DLAs will allow us to put in place another piece of the puzzle that represents the complex interplay between dense cold gas (possibly still in-falling from the cosmic web) and the different stages of galaxy evolution at high z.

- **Quantum space-time degradation at high redshift**. Indeed, quantum space-time scenarios predict a degradation of the diffraction images of distant sources. The best limits on this fuzziness have been obtained so far by HST in the optical band. The limit on the phase variation is of the order of a few 10^{-7} radians. Obtaining SR$\sim 40-50\%$ in the I-band with an 8m class telescope can improve such limit by a factor of 2-3.

1.2 SHARK-NIR OPTO-MECHANICAL LAYOUT

SHARK-NIR will image a field of view of 18x18 arcsec onto a 2048x2048 pixels H2RG Teledyne detector, with a plate scale of 14.5 mas/px. The Airy disk is two pixels in radius (Nyquist sampled) at the wavelength of 0.95 μm. The instrument can be divided in three main sub-systems:

1. The science channel, composed by the relay optics that makes the intermediate images of the pupil/sky, the coronagraphic masks, the filters and the scientific camera.

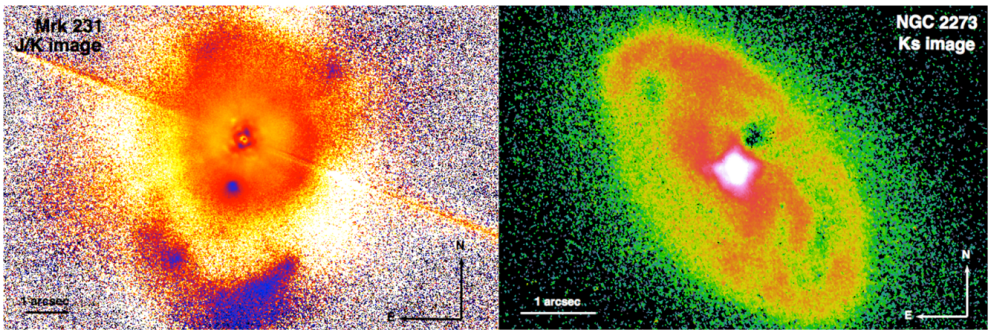


Figure 4: AO-assisted images obtained with PISCES at LBT. *Left*: J-K color image of Mrk 231. *Right*: Ks unsharped image of NGC 2273. The achieved SR was, respectively, 30% and 20% with seeing worse than 1" and R+I $13.7 + 12.9$, $15.5 + 15.0$.

2. The tip-tilt WFS that picks up part of the scientific light, by means of the beam splitter of figure 8, to measure and correct the movement of the star during an exposure.

3. The calibration unit used for flat field and wavelength calibration of the instrument.

1.2.1 The science channel

The relay optics used to make the intermediate images of the pupil and focal planes are Off-Axis Parabolic (OAP) mirrors. There are four parabolic mirrors in total: the first to make the first image of the pupil, the second to make the intermediate focal image, the third to make the second pupil image and the fourth to produce the scientific image on the detector. The relation between the off-axis angles of the parabolae and their magnification factor was optimized as described in Ghedina and Ragazzoni 1997. This approach allows to achieve a very good optical quality but imposes some restrictions on the direction of the light beam. In particular, for each couple of parabolae, the outgoing beam is bent on the same side of the incoming beam as depicted in (Figure 5). For this reason, in order to avoid superposition of the optical components and keep the instrument in the very small volume available, three additional flat mirrors are used to fold the system.

Figure 6 shows a CAD view of the optical bench, while the optical design of the instrument is shown in Figure 7. Following the light path from the left side of Figure 7, the light beam from the telescope is picked-up before the Nasmyth focus by a dichroic mirror which directs the light between $\lambda = 0.96$ µm and $\lambda = 1.70$ µm inside SHARK-NIR and transmits the visible light to the WFS of LBTI. Then, the light is directed inside the instrument by a folding mirror supplied with two very precise DC motors (IN-TT in Figure 7). The motorized axes will allow the mirror to perform TT al low frequency to center the pupil image on the coronagraphic masks with high precision (better than 1/100th of the pupil diameter). Dichroic and TT mirror

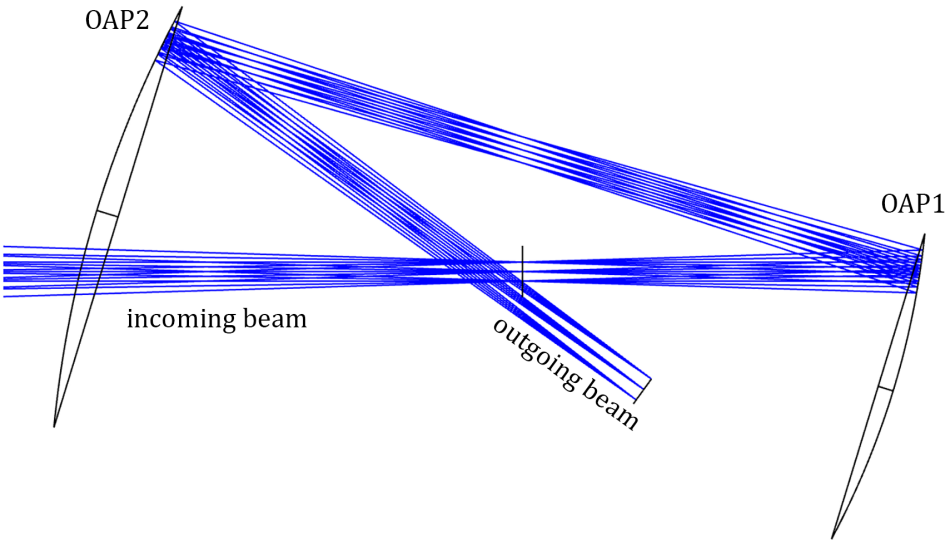


Figure 5: Optimized configuration of two consecutive OAP mirrors.

are both part of a single sub-system, which is mounted on a deployable arm not connected to the rotating part of the instrument.
Light then enters through a flat optical window into the carter in which the optical bench is enclosed (to prevent dust contamination of the optics). The entrance window has a wedge of about 15 arcmin in order to move the ghost image 5 arcsec far from the central field. The science channel is composed as following:

- An off-axis parabola (OAP1) is creating a pupil of diameter 11.2 mm. In this plane (PP1) is located the internal ALPAO DM 97-15, that will be used both for fast tip-tilt correction and for Non-Common Path Aberrations (NCPA) compensation.
- A filter wheel is positioned 50 mm after the pupil plane and will host the apodizing masks. This displacement may represent a limitation for those techniques relying on pupil apodization and has been taken into account in the choice and design of the coronagraphs.
- Immediately after, in the collimated beam, the ADC is placed.
- Between the ADC and the second off-axis parabola (OAP2), a beam splitter will send a small portion of the light ($\sim$ 5%) to the internal tip-tilt sensor.
- OAP2 is refocusing the beam on an intermediate focal plane (FP-CORO) with an $F/\# = 22$, where a filter wheel can select between different occulting masks. The same wheel will accommodate the slit for the spectroscopic mode, for spectral characterization of the science targets. The clear aperture required for the filters in this focal plane, in order to not vignet the FoV, is about one inch in diameter.

- After a folding mirror (FM1), a third off-axis parabola is creating the second re-imaged pupil plane (PP2), where a filter wheel can select between different Lyot stops in order to minimize diffraction effects. The pupil diameter is 14.9 mm. In the same wheel, the dispersive elements for the spectroscopic mode will be placed.

- On the collimated beam, two additional filter wheels (positioned between the pupil plane and a folding mirror, FM3) will allow the insertion of the scientific filters.

- After a folding mirror, the fourth off-axis parabola (OAP4) is creating the final image onto the detector (FP-SCI). A filter wheel just before the dewar window will accomodate dual-band filters. Additionally, a deployable small optical group (an achromatic doublet), not shown in Figure 7, can be inserted between OAP4 and the dewar window with the purpose to create an image of the pupil onto the detector, which will be used to align the coronagraphic pupil masks before exposure.

The entrance window of the dewar is kept at a distance of 200 mm from the detector in order to limit the solid angle through which the detector sees the external environment and thus the thermal background. Moreover, a gold-coated annular Narcissus mirror is placed 40 mm before the dewar window for the same purpose. The mirror has a central hole ($\emptyset = 43.4$ mm) to transmit the scientific FoV and the surface looking at the detector has spherical shape with the center of curvature placed on the detector (R=250 mm). The solid angle subtended by the hole on the narcissus mirror, as seen from the detector center, is $\omega = 0.023$ sr.

1.2.2 The tip-tilt WFS

The layout of the TT-WFS arm is shown in Figure 8. The arm is oriented in the vertical direction with respect to the plane of the drawing in Figure 7 (Y-axis in the figure) and is composed of an imaging lens and a detector.
The sensor was originally concieved to drive a slow TT loop, aimed at monitoring the long-term drifts of the star during the exposure in order to keep it aligned to the coronagraphic focal plane mask. After a dedicated study which followed the FDR, we decided to move towards a fast TT loop (> 1 kHz), with the idea to compensate also for high-frequency drifts induced by differential vibrations between the instrument and the WFS. The original commercial camera will be thus replaced by a First Light C RED-2 camera, which is equipped with a 640x512 InGaAs sensor operating in the SWIR (Short Wave Infrared), from 0.9 to 1.7 μm, with a very good Quantum Efficiency (over 70% in the whole wavelength range). This camera would allow to perform TT sensing down to magnitude $12 - 13$, keeping the centroid error below 3 mas rms.
The TT corrector will be the same DM located in the first intermediate pupil plane and used to compensate for NCPA. The model we selected features a

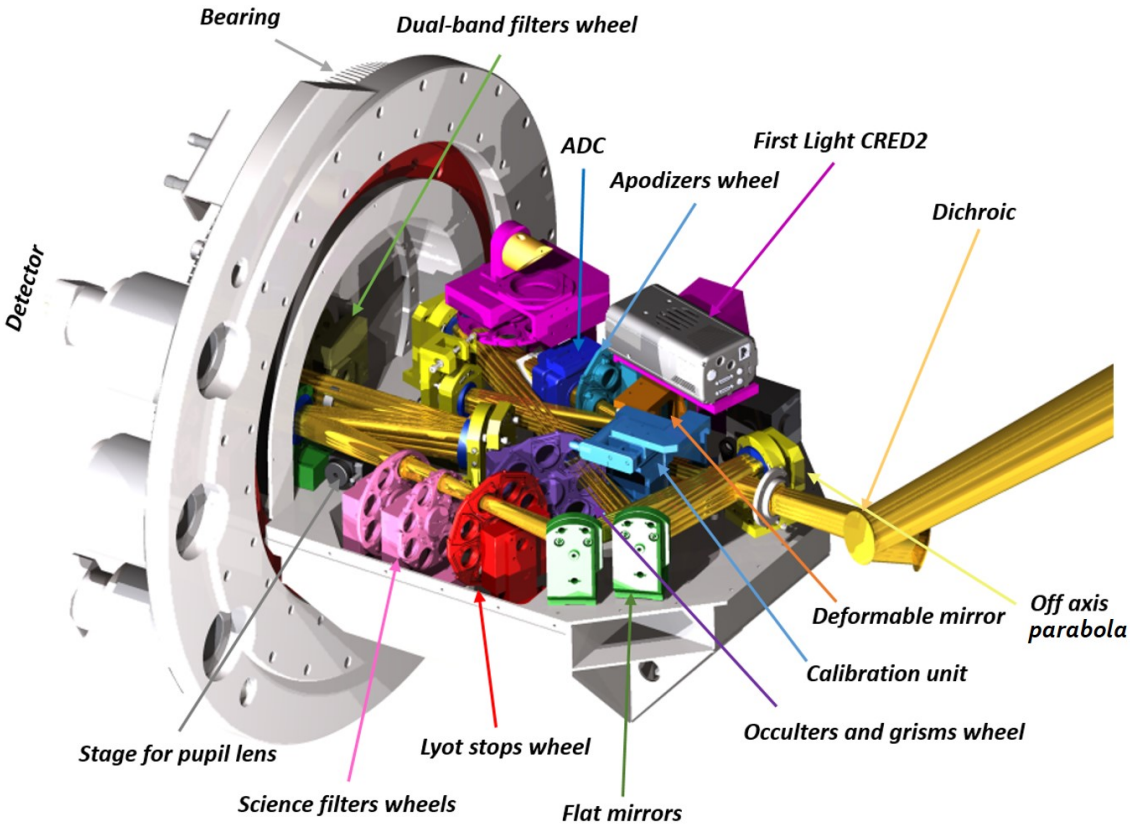


Figure 6: The opto-mechnical concept of SHARK-NIR bench.

1.25 kHz bandwidth with a TT stroke of ±30 µm, corresponding to ±1.2″ on sky.

1.3 INSTRUMENT MODES

The instrument will offer to the community the possibility to choose between the following observing modes: *Direct Imaging* (DI), *Coronagraphic Imaging* (CI), *Dual Band Imaging* (DBI) and *Long Slit Spectroscopy* (LSS).

1.3.1 Direct imaging

Direct imaging assumes no coronagraph nor slit are inserted along the optical path. All the scientific filters combinations are available to the observer in this mode. The DI field of view is 18″ × 18″ (square).

1.3.2 Coronagraphic imaging

The presence of wheels in the coronagraphic focal plane and in the two intermediate pupil planes will offer to the user the possibility to choose

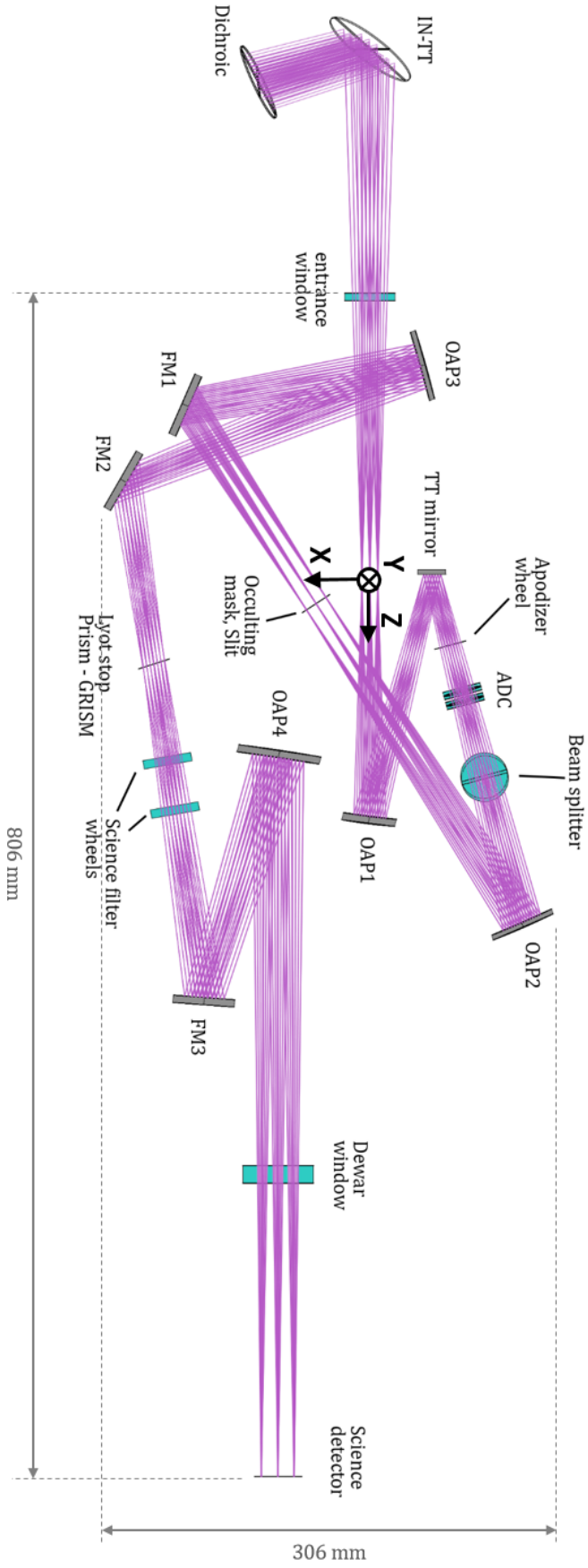


Figure 7: Optical layout of SHARK-NIR (top view of the optical bench). The light beam enters the instrument on top of the figure.

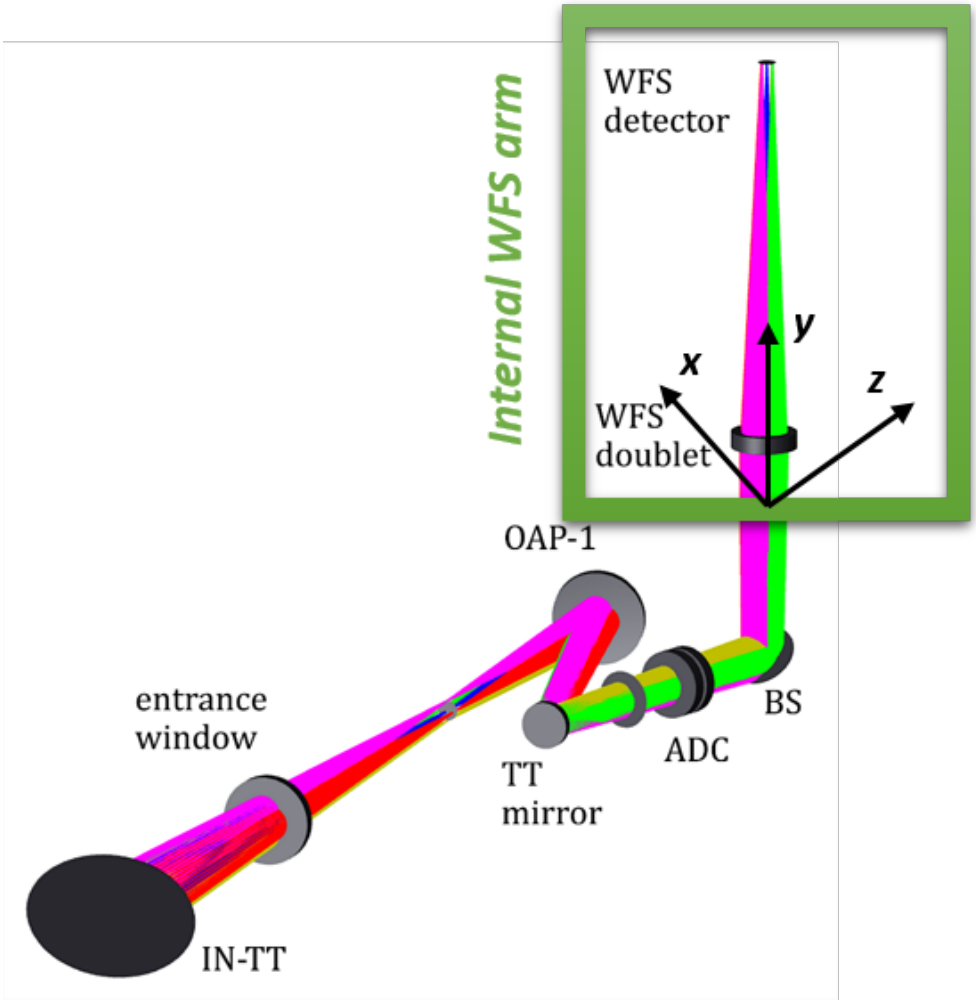


Figure 8: Layout of the internal Tip-Tilt WFS arm. The optics up to the beam splitter (BS) are in common with the scientific channel.

between a suite of coronagraphic options, depending on the need in terms of contrast, proximity to the star and imaged field of view.

1.3.3 Dual Band Imaging

The presence of a Wollaston prism in PP2 also allows to perform Dual Band Imaging (DBI). The prism is designed to achieve a beam deviation of $\delta = 2.15°$, which corresponds to 14″ on-sky. The material of the birefringent crystal is YVO4, selected for its low chromatic elongation in the wavelength range of SHARK-NIR. The DBI field of view is restricted to 2″ in diameter so to allow the two beams to be completely spatially separated before the Dewar window and thus the dual-band filters wheel to be positioned outside the cryostat. Dual band imaging can be operated in coronagraphic mode, with the restriction to those options not critically dependend on a Lyot stop (Wollaston and Lyot stops are on the same wheel).

1.3.4 Long-slit spectroscopy

The LSS mode can assume two different configurations, namely a low resolution (R ~ 100) and a medium resolution (R ~ 700) one. In LSS mode, the light focused on the FP-CORO is then selectively masked by a 0.1" wide slit, equipped with an occulter, whose size can be 0.1" or 0.2", to be selected by the observer, according to the scientific target to be observed. On the second Pupil Plane (PP2), the dispersive element can be inserted, choosing between an Amici prism, which allows a resolution at the level of the detector of ~ 100, and a more complex solution which could include a single grism or 2 prisms and two gratings combined in multiplexing mode, to push the resolution to a medium value of ~ 700.

1.4 THE SCIENTIFIC REQUIREMENTS

Table 1 summarizes the top-level scientific requirements to be satisfied in order to address the primary SHARK-NIR scientific cases discussed in this chapter. These requirements are the fundamental drivers in the design of the coronagraphs.

Table 1: Requirements for the primary SHARK-NIR science cases.

Target	Mode	Requirements
Exoplanets	CI, DBI	IWA: maximum 300 mas, goal: $100-200$ mas
		Contrast: $10^{-5}-5\text{x}10^{-6}$ for separations <300 mas; $10^{-5}-10^{-6}$ in the range $300-500$ mas;
		Rmag: $5-10$
		FoV: $<1\times1''$
	LSS	Resolving power: R $\sim$ 100 and R $\sim$ 700
		SNR: >5
Disks	CI	IWA: $150-200$ mas
		Contrast: $10^{-4}-10^{-5}$ at separations of $200-300$ mas
		Rmag: $6-12$
		FoV: $<1\times1''$
Stellar jets	LSS	IWA: $150-200$ mas
		Contrast: 10^{-3} at separations of $200-300$ mas
		Rmag: $6-12$
		FoV: $<1\times1''$
		Resolving Power: R $\sim$ 100
AGN	DI, CI	IWA: $200-600$ mas
		Contrast: $10^{-5}-5\text{x}10^{-6}$ at 2000 mas for $z=0$; 10^{-4} at 700 mas for $z=0.1\div1.1$;
		Rmag: $9-14$
		FoV: $5\times5''$ and $15\times15''$ for, respectively, DLAs and AGN

2 THE SIMULATION CODE

SHARK-NIR is a coronagraphic camera. Since coronagraphs deal with diffraction of light, it is necessary to operate in the framework of wave-propagation physics: given the mathematical complexity of this theory, the most common approach is to make use of numerical simulations. For the purpose, I developed a Fresnel simulator in IDL language. Starting from the optical design of SHARK-NIR, I developed a simplified 1-D model of the telescope+camera system in the form of an IDL module whose backbone is a sequence of calls to procedures and functions belonging to IDL library PROPER (Krist 2007). This module, hereafter called the *prescription*, is the very functional core of the simulator (section 2.1): it takes an electric field disturbance and numerically propagates it through the optical train according to scalar theory of diffraction, from the entrance pupil of the system until the detector plane. Intensity in this final plane is then recorded and used to generate the synthetic image (section 2.2). Initial conditions for the disturbance are here represented by the different sources of optical aberrations: residuals from AO correction, non common path aberrations and telescope jitter (sections 2.3 to 5.3).

2.1 THE PRESCRIPTION

The light path from the entrance pupil of the telescope to the SHARK-NIR detector is unfolded into a linear sequence of planes (table 2). PROPER makes no distinction between mirrors and lenses: both are treated as ideal phase transformers and are generally named 'lenses'. LBT M1 assumes the correct effective aperture and focal ratio to account for M2 without introducing it as a separate surface. For the off-axis parabolas, focal lengths are given by the optical design. Together with lenses, there are also four'simple' planes. The last corresponds to the detector, where the final image is formed, while the other three are the *coronagraphic planes*: two pupil planes (PP1 and PP2) and one focal plane (FP_CORO). These planes are reached by the wavefront only if a specific optical element is foreseen by the coronagraphic technique, otherwise they are skipped so to speed up the simulation.
There are several optical components in SHARK-NIR not included in this model (entrance window, ADC, science filters and dewar window). However, considering all of these elements would result very demanding from the computational point of view. The idea behind this model is to lay the basis for operating in a more advanced framework with respect to a simple Fraunhofer approximation, but keeping only a reasonable number of sur-

Table 2: The sequence of surfaces defining the simplified 1-D optical model used for propagation. Surfaces can be either lenses or simple planes. In order: 'f' is the focal length, Δz is the intra-surface distance travelled by the wavefront, while 'z' is the absolute position along the propagation axis.

#	**Surface**	**Type**	fl [mm]	**Δz [mm]**	**z [m]**
1	M1	Lens	123300	-	0
2	OAP1	Lens	165	123465	123.465
3	Apodizers wheel	Plane	-	215	123.68
4	OAP2	Lens	240	302	123.982
5	FP_CORO	Plane	-	240	124.222
6	OAP3	Lens	325	325	124.547
7	PP2	Plane	-	120	124.667
8	OAP4	Lens	460	410	125.077
9	Detector	Plane	-	460	125.537

faces. In such a framework, a step forward in the complexity of the model for any future need can easily be implemented.

For propagation, PROPER adopts the technique of the pilot beam. This technique consists in propagating a Gaussian beam analytically in parallel to the wavefront: in this way, the method used to propagate the wavefront from surface i to surface $i+1$ is selected by looking at the relative position of the surfaces themselves with respect to the pilot beam waist. Both near and far field propagation are implemented, using, respectively, the angular spectrum or the standard Fresnel method.
In the following, I schematically report the sequence of calls to PROPER routines as they appear in the prescription, with a short description for each of them. The reported arguments of the routines (focal lengths and propagation distances) refer directly to table 2.

- **prop_begin**. This routine starts the propagation: it initializes the pilot beam and creates the complex-valued wavefront array. This array is organized into an IDL structure, together with several other variables carrying out informations about the state of the wavefront itself. The computational grid is 1024×1024 pixels, while the ratio of the pupil size to the grid one (the zero-padding factor) is chosen to be 4. When using Fast Fourier Transform algorithm (FFT), the zero-padding defines the focal planes sampling. A value of four means that the pupil is 256 pixels in size and that $1\lambda/D$ in focal planes is sampled with four pixels, instead of the traditional two (Nyquist).

- **prop_circular_aperture, prop_circular_obscuration, prop_rectangular_obscuration**. These routines are used to create the system entrance aperture (Figure 9). The aperture is antialiased by means of gray pixels at its edge.

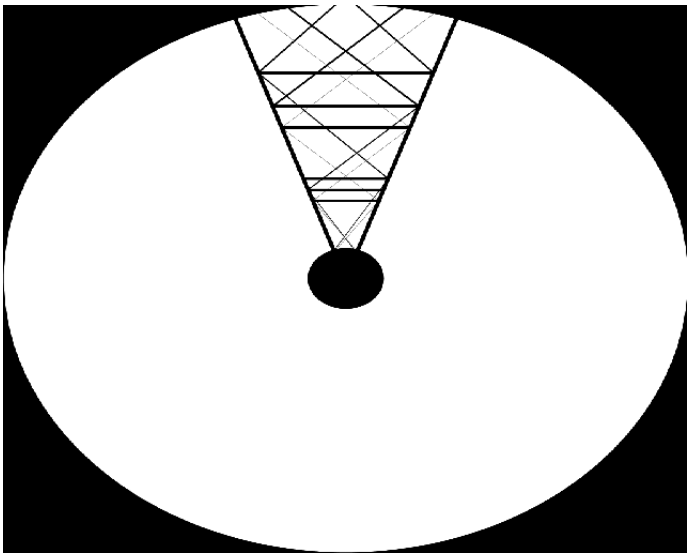


Figure 9: LBT pupil used in simulation built with PROPER.

- **prop_define_entrance**. The call to this routine tells PROPER to consider the current wavefront array as the wavefront at the entrance pupil of the system, so it is renormalized to have total intensity equal to 1 over the pupil.
- **prop_lens**, *f_M*1. This routine applies the phase transformation to the wavefront. The effective focal ratio of M1 is $F/\# = 15$. At the same time, the routine computes the new position of the pilot beam waist as the beam crosses the lens.
- **prop_propagate**, Δz_2, *'OAP*1*'*. Wavefront is propagated to the first parabola, that collimates the beam. The telescope $F/\# = 15$ focal plane is made coincident with the front one of the parabola, hence the distance travelled by the wavefront is the sum of M1 and OAP1 focal lengths. Using the position of the first parabola as a reference, all the others parabolas are positioned along the propagation axis accordingly to their real reciprocal by-design distances.
- **prop_lens**, *fl_OAP*1.
- **prop_propagate**, Δz_3, *'Apodizer'*. Wavefront is propagated to the position of the apodizer wheel. In the paraxial approximation, the first intermediate pupil (PP1) forms at a distance from OAP1 equal to its focal length. The apodizer wheel is then positioned 50mm after it.
- **prop_multiply**, *Apodizer*. The apodizer is multiplied to the wavefront.
- **prop_propagate**, Δz_4, *'OAP*2*'*. The wavefront reaches the second parabola.
- **prop_lens**, *fl_OAP*2.
- **prop_propagate**, Δz_5, *'FP_CORO'*. Wavefront is propagated to the coronagraphic focal plane. This plane is positioned in the focus of OAP2.

- **prop_multiply**, *FP_mask*. The occulting mask is applied to the wavefront. Several options are available, depending on the selected coronagraph.
- **prop_propagate**, Δz_6, *'OAP3'*. The wavefront reaches the third parabola, where the beam is re-collimated. The distance travelled is equal to the focal length of the parabola.
- **prop_lens**, *fl_OAP3*.
- **prop_propagate**, Δz_7, *'PP2'*. The wavefront reaches the second intermediate pupil plane. Again, the position of the pupil is computed according to the paraxial approximation.
- **prop_multiply**, *Lyot stop*. Wavefront is multiplied for the Lyot stop.
- **prop_propagate**, Δz_8, *'OAP4'*. The last parabola is reached.
- **prop_lens**, *fl_OAP4*.
- **prop_propagate**, Δz_9, *'Detector'*. The wavefront finally reaches the detector plane. The distance travelled corresponds to the focal length of OAP4.
- **prop_end**. These routine ends the prescription. The intensity (square module of the complex field) is computed and returned to the caller.

2.2 IMAGE FORMATION

In the framework of diffraction theory it is common and convenient to work with phasors, rather than electric fields directly. A phasor is a complex function of the position, defined by a phase and an amplitude, mathematically related to the real physical disturbance. It can be shown that, following its definition, the squared module of the phasor, named intensity, is directly proportional to the power density of the optical wavefield, which is its directly measurable physical attribute.
The raw output of the propagation prescription is thus an intensity map on the detector plane, namely the wavefield energy (or equivalently the number of photons) per unit time and area. How to generate an image integrated over a given exposure time? The total energy per unit area after a time T is simply the integral of intensity:

$$E = \int_0^T I(t)dt \tag{1}$$

This integral translates into a discrete sum after decomposing interval T in a finite number of points N:

$$E = dt \cdot \sum_{i=1}^{N} I_t = (Ndt) \cdot \frac{1}{N} \sum_{i=1}^{N} I_t = T \cdot \langle I \rangle \tag{2}$$

The total number of photons reaching the detector after a time T is then proportional to the average wavefield intensity. To generate the image then the simulator runs the prescription N times, generating N instantaneous PSFs, which are then averaged. The number of sampling points in interval T depends on the time step dt of the simulation:

$$N = \frac{T}{dt} \tag{3}$$

Since atmospheric residuals are generated at a precise cadence, depending on the AO loop framerate (see section 2.3), the most natural choice is to use the same framerate for temporal integration. In this way, at each atmospheric screen will corresponds an instantaneous PSF, obtained running the prescription for that initial condition. The AO frame rate depends on the brightness of the target: it can range from a minimum of 0.67 ms (bright target) up to 4 ms in the very faint-end regime explored. This means 250 up to 1500 instantaneous PSF per second of integration.

The effect of integration time on a closed-loop image is clearly visible in figure 10. The simulation includes a coronagraph, while no other noise source is considered except for atmospheric speckles. Seeing is 0.4″ and star R magnitude is 8. One second of integration corresponds to 8000 phase screen employed (AO works at 1 kHz in these conditions). Bright speckles are distributed in an external annulus located at the AO control radius (fitting error), but also in an horizontal pattern inside the corrected region (lag due to the wind). By visual inspection, these speckles tend to gradually form a smooth structure as integration time increases. Looking at the standard deviation of intensity as a function of time (figure 11), it can be seen that the curves flatten rapidly. This is expected, since speckle noise in the long-exposure regime tends to a constant value which is ultimately dependent on the power spectrum of the aberration itself. Based on these curves, in all simulations reported in this thesis a reference integration time of one second is assumed. This value is found to be a good comprimise between the computational time required to generate the images and minimization of the atmospheric speckle noise.
As explained in chapter 3, simulated images are converted into counts units before being processed. The amount of photons in the images is directly proportional to the integration time and impacts on the SNR. With this choice of the integration time, we can in principle assume in the conversion to counts an arbitrary higher integration time than the nominal one used in the simulation. In this way, the impact of high DIT on performance can be expolored without having to run highly computationally expensive end-to-end simulations. To remark the concept, this relies on the assumption that the effect of integrating for more than one second is simply that of an increment in the number of photons in the image, with no effect on their spatial distribution.

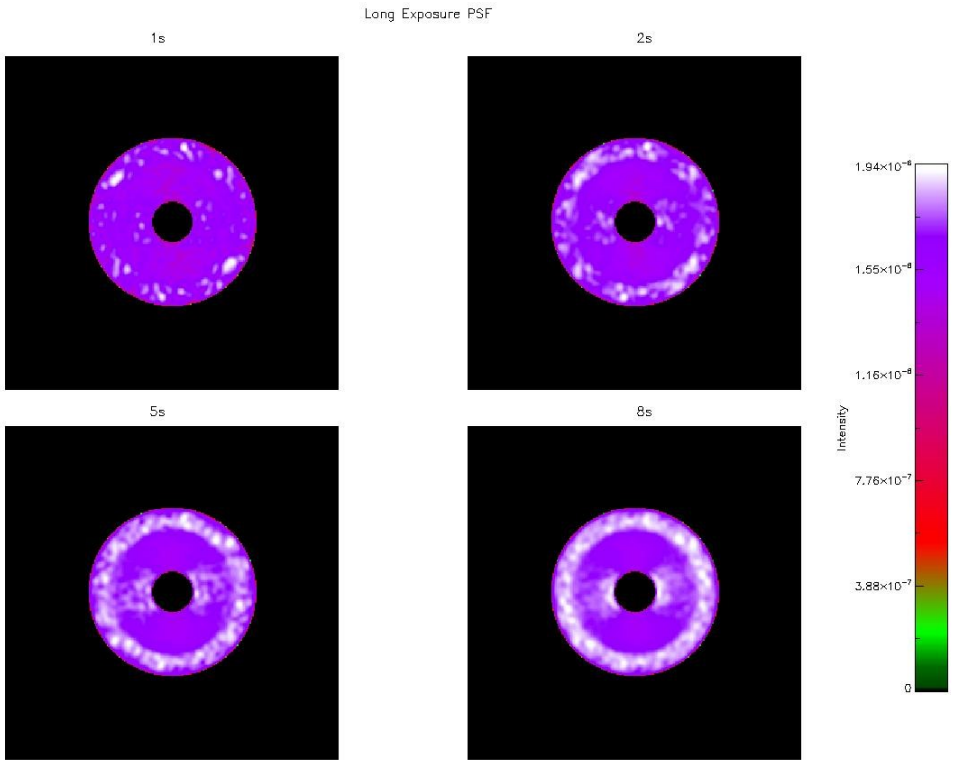


Figure 10: Coronagraphic PSFs showing the distribution of atmospheric speckles after 1, 2, 5 and 8 seconds of integration.

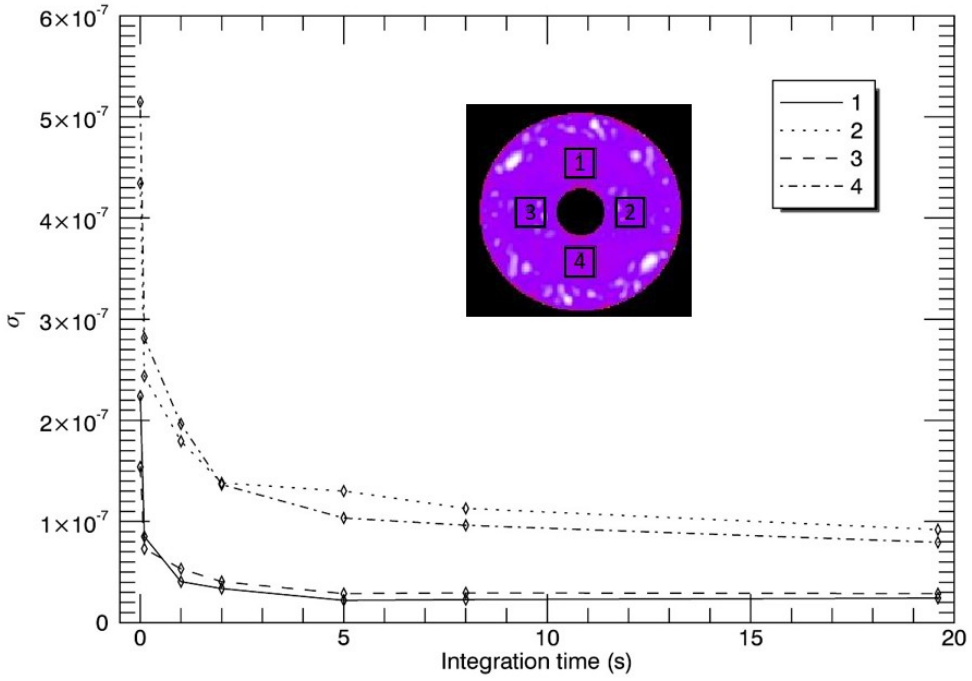


Figure 11: Standard deviation of intensity in the four quadrants displayed in the window as function of integration time. Intensity is normalized to have an input power equal to one.

2.3 ADAPTIVE OPTICS

The correction delivered by the AO system is one of the key aspects to consider in the analysis of the coronagraphic performance. SHARK-NIR will share with LBTI the WFS module of the central bent-gregorian station of LBT. The module is one of four identical and forms, together with the Adaptive Secondary Mirror (ASM), the LBT natural guide star (NGS) single-conjugate AO system FLAO (First-Light Adaptive Optics, see e.g. Esposito et al. 2011).
In order to implement realistically the expected AO correction, we generated phase screens of atmospheric residuals with the official LBT AO end-to-end simulator PASSATA (Agapito et al. 2016), developed in Arcetri (Florence). Phase screens are organized in data cubes, each composed of 30 up to 100 seconds of closed-loop operation. We asked for data cubes in several configurations of atmospheric seeing and guide star magnitude, in order to explore a wide range of conditions. Furthermore, at the time SHARK-NIR will be at the telescope, all FLAO modules will have undergone an important upgrade. The Single conjugated adaptive Optics Upgrade for LBT (SOUL, Pinna et al. 2016) has been first proposed, together with SHARK, in answer to the call for instrument upgrades and new instrumention issued by LBTO in 2014. It consists in the installation of a new generation visible detector with sub-electron RON, together with other hardware/software upgrades and a new camera lens. This modifications will lead to an improvement in the system sampling capability in both spatial and temporal domain. For this reason, we asked for data cubes also in the new configuration. Hereafter, I will use the term FLAO to refer exclusively to the current system, while SOUL will denote the future upgraded configuration.

The PASSATA simulator

The AO simulator assumes a two-layer atmosphere under the Taylor frozen flow hypothesis. The two layers flow in opposite directions, with speeds of 15 m/s at the ground layer and 18 m/s at the higher layer (located at an altitude of 6 km). C_n^2 fractions are, respectively, 0.6 and 0.4. Adopted wind speeds are deliberately high, resulting in a conservatively small wavefront coherence time: for example, the median velocity at Mt. Graham site at 6 km is around 14 m/s (extrapolated from figure 2 in Hagelin et al. 2010). Seeing can be either assumed to be constant or to fluctuate. The baseline for my simulations is to assume a constant seeing. Finally, the simulator considers influence matrices directly measured on the ASM of LBT and the real characteristics of WFS detector, optics and AO Real Time Computer.

Using a semi-analytical AO error budget model, the Arcetri group explored the system parameters in order to find the best configuration in terms of residual wavefront error as a function of guide star brightness. The pa-

rameters considered for the optimization are *pupil sampling*, *loop framerate*, *number of corrected modes* and *integrator gain*. Table 3 reports the optimal values for the first three in the explored magnitude ranges for both FLAO and SOUL configurations. Integrator gain is also a function of seeing, hence it is not reported in the table to keep it readable. FLAO works at the best possible spatial and temporal sampling at magnitudes $\leqslant 8$, while for SOUL this limit is shifted to magnitude 10. At magnitudes brighter than these, AO loop parameters are of course unchanhged. The corresponding Strehl ratios in H-band in all the explored combinations of seeing and NGS magnitude are reported in table 4. These numbers are retrieved directly from the synthetic phase screens using the exponential approximation.

Table 3: Optimal number of sub-apertures, framerate and number of corrected modes in the magnitude range explored in simulations, for both FLAO (white) and SOUL (gray) systems.

R	Pupil sampling	Framerate [kHz]	# modes
8	30×30	1	400
	40×40	1.5	630
10	15×15	1	153
	40×40	1.5	630
12	15×15	0.7	153
	40×40	0.5	630
14	-	-	-
	20×20	0.4	170
15	-	-	-
	13×13	0.3	90

As an example, I report in the plot in Figure 12 the screen by screen rms wavefront error (WFE) calculated on 30 seconds of closed-loop operation simulated with PASSATA. The target is bright ($R = 8$): 30 seconds correspond in this case to thirty thousands screens (FLAO configuration). Seeing is good ($0.6''$). The statistical dispersion of residual WFE around the mean value is of only ~ 4 nm, consistently reproducing a constant seeing observation. The choice of constant seeing might look obvious considering 100 seconds of closed-loop operation at most. However, the real motivation behind this assumption will appear clear in the following chapter when discussing how a simulated observation is strucured.

Table 4: Strehl Ratios @1.6μm delivered by the AO simulator PASSATA for both FLAO and SOUL configurations as a function of seeing and guide star magnitude.

R	seeing ['']	FLAO	SOUL
8	0.4	96%	97%
	0.6	92%	96%
	0.8	85%	-
	1.0	75%	-
10	0.4	88%	94%
	0.6	77%	91%
	0.8	62%	87%
12	0.4	78%	91%
	0.6	69%	85%
	0.8	53%	77%
14	0.4	-	79%
	0.6	-	68%
	0.8	-	52%
15	0.4	-	69%
	0.6	-	49%
	0.8	-	27%

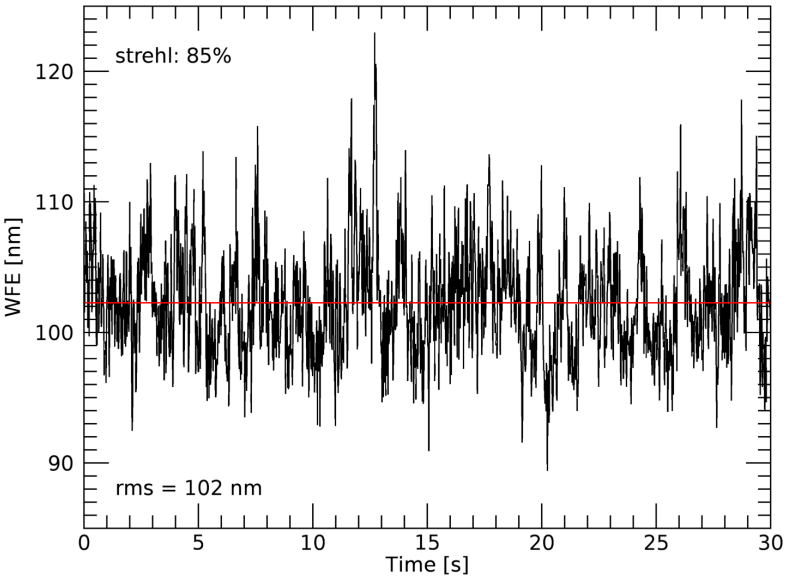


Figure 12: Screen by screen rms wavefront error for 30 seconds of closed-loop operation simulated with PASSATA in the case of a bright target (R = 8) and good seeing (0.6″).

2.4 NON COMMON PATH ABERRATIONS

Speckles originating because of Non Common Path Aberrations (NCPA) represent a significant source of noise in high-contrast imaging because of their long life and slow temporal evolution. Figure 14 shows the temporal power spectrum of speckle intensity obtained in a 20 minutes observing run of star Gliese 777 with SHARK-VIS experimental imager (the Forerunner, Pedichini et al. 2017). The instrument was installed at the right bent Gregorian focus of LBT in 2014. Its simple optical layout is shown in figure 13. During the experiment, more than one million frames were acquired in closed-loop at 1 ms cadence. A significant rise in speckle intensity at frequencies lower than 0.1 Hz was clearly highlighted in the data: this is likely due to slowly-evolving NCPA. The implementation of this phenomenon in simulations is crucial for a realistic noise estimation after post-processing.

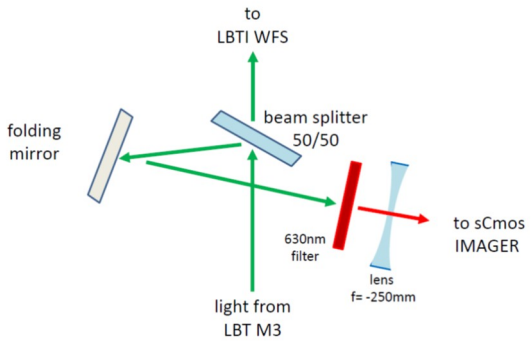


Figure 13: The SHARK-VIS Forerunner optical layout. The instrument simply consisted of two refractive surfaces, a narrow-band filter centered at 630 nanometers and a divergent lens to form the image of the star.

In my simplified optical model, NCPA are introduced at the entrance pupil of the system. Modelling NCPA spatial and temporal behaviour is not easy. It is important to stress that, from this point on, what I will call NCPA are actually the residual NCPA after a generic off-line compensation procedure is actuated. This is the situation the instrument will face at the telescope. Throughout the instrument design phase, the baseline approach to the problem was to make simplified assumptions on the residual NCPA directly, without caring about the full budget and related compensation strategy from which these residuals would have originated. In this sense, an end-to-end approach has been then investigated after the FDR, leading to a detailed modeling of the whole input NCPA budget and simulations of the compensation capabilities using the Phase Diversity algorithm. I stress that in all detection limits shown in this thesis NCPA are included with the assumptions reported in this section. The post-FDR analyisis is detailed in chapter 8.

Three basic assumptions are made: power spectrum, total integrated power and temporal evolution scheme. Regarding the power spectrum, I assumed

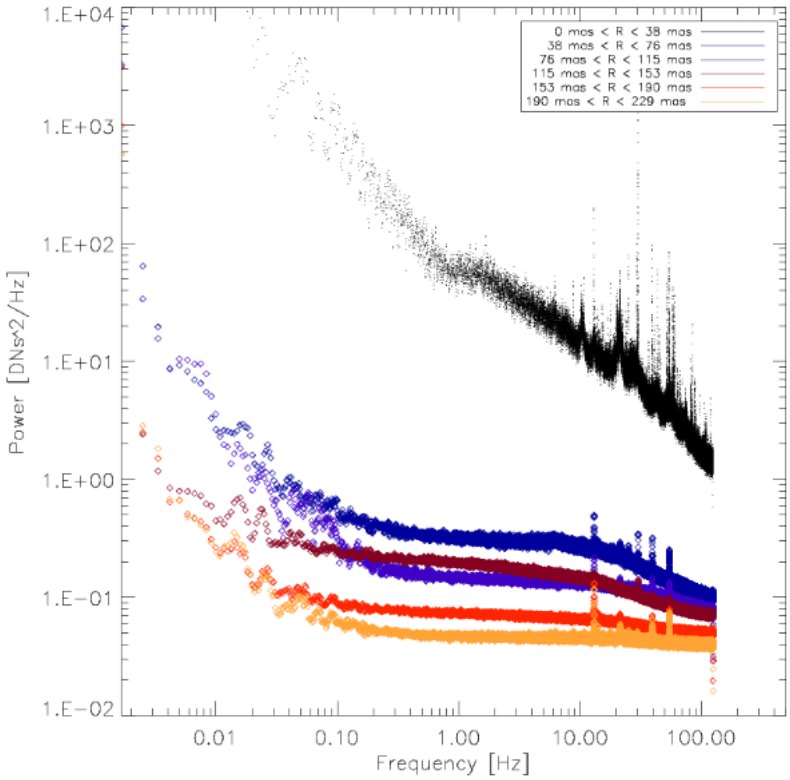


Figure 14: Temporal power spectra of intensity fluctuations at different angular separations. Data were acquired at LBT during a 20 minutes observing run of Gliese 777 in June 2015.

a flat power spectral density function (PSD) until the AO cut-off frequency, followed by a f^{-2} decline (figure 15). This model is the same used in VLT/-SPHERE simulations and relies on the assumption that the compensation is performed with the ASM. However, the assumption on the power spectrum is not critical: simulations show that that the shape of the spectrum does not critically influence detection limits. Total power is 30 nanometers rms. This value is routinely achieved at the telescope with the LUCI camera, so we considered it as a reasonable estimate. The evolution law is much more critical. Temporal evolution becomes of course important in the search for exoplanets, where few hours of observation on a single object are usually required to apply differential imaging (e.g. ADI). Under these conditions, NCPA usually become a significant source of noise. Martinez et al. 2013 applied a statistical model for noise variance in high-contrast imaging to data obtained with the SPHERE instrument to derive an empirical law for the temporal evolution of quasi-static speckles wavefront error. This error is found to be in the regime 0.1 to 1 Å and to increase linearly. The approach I followed consists in scrolling a fixed NCPA map along the pupil every time a new image of the sequence has to be generated. In this way, the pupil always "sees" a slightly different input aberration, resulting in a modification of the speckle pattern whose entity depends on how fast the NCPA map

crosses the pupil. The choice to apply a shift is not really motivated by any physical argument, but it is only a way to introduce a source of noise that mimics NCPA evolution. Although in this way the change in the speckle pattern is not due to an increasing wavefront error, simulations show that the net effect is similar if the amount of shift is properly tuned.

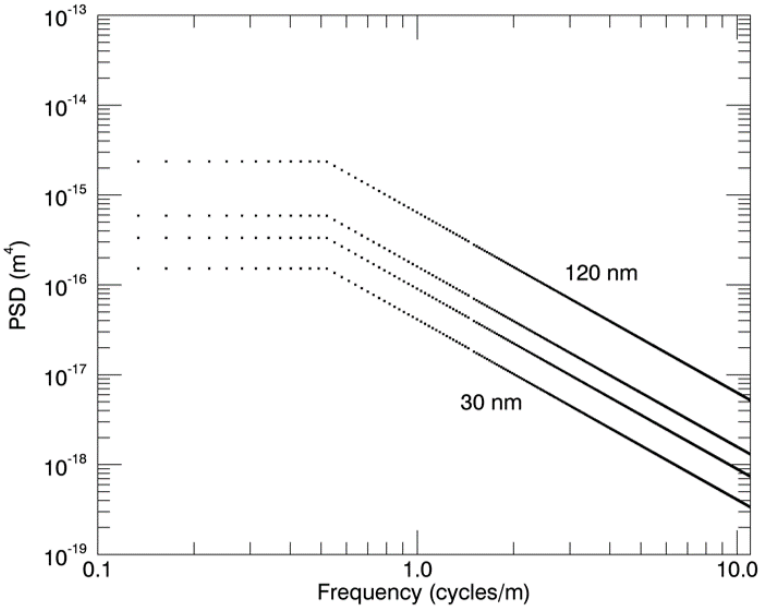


Figure 15: Power spectral density function of residual NCPA for different values of the total injected power. Spectra are flat until the AO control radius, then they decline following a f^{-2} law.

2.5 TELESCOPE VIBRATIONS

Vibrations arise because of resonant modes in the structure of the telescope, in particular the swing arm supporting the ASM. These modes are excited by wind shacking and/or telescope tracking and may introduce differential tip-tilt aberrations between the AO system and the camera. Vibrations at LBT have been characterized during FLAO commissioning run: the median value is around 6 mas rms, with most of the power concentrated at a frequency of $\sim$ 10 Hz. These frequencies are much smaller than FLAO/-SOUL framerates even in the faint-end regime, hence the effects of vibrations can be very well sampled in the temporal domain. For the purpose of modeling the phenomenon, I developed three different schemes:

- **Sinusoidal jitter**
  In this simple scheme, vibration spectrum is approximated with its dominant Fourier component: jitter is modeled as a tilt-only aberration varying sinusoidally with frequency $\nu_s = 13$Hz:

$$\phi(r,\theta,t) = C(t)\cdot Z_2(r,\theta) \qquad C(t) = A\sin(2\pi\nu_s t) \tag{4}$$

The amplitude A of the sinusoid is a free input parameter: the rms is simply $A/\sqrt{2}$.

- **On-sky data**
  To model vibrations more realistically, another way could be to use real on-sky measurements. I thus took the centroids measured on Forerunner images (section 2.4) and converted them into a tip-tilt aberration to reproduce in simulations the same star 'wandering'. In this way, a complete vibration temporal spectrum is considered. Although the sample refers of course to a single case of vibrations, data can be in principle arbitrarily rescaled to change the rms. The spectral information is preserved in this operation.

- **AO**
  Associating vibrations to image motion is still an over-simplification. The true impact of vibrations on image quality in closed-loop operations is difficult to assess because it critically depends on the AO system response and filtering. The most realistic approach in this sense would be that of accounting for vibrations directly in the AO simulator when generating phase screens. This approach has been already used by the Arcetri group to characterize SOUL performance: the basic idea is to model open-loop vibrations and then introduce them in input directly to the AO simulator. Phase screens generated under these conditions then automatically incorporate residual vibrations. In this way, the total open-loop vibration power becomes the true free parameter, while the residual is exactly the one we should expect in the simultated conditions (seeing and NGS magnitude). Furthermore, loop parameters themselves can also be re-optimized to mitigate the impact of vibrations: this is typically achieved increasing the AO framerate.

Each approach has its pros and cons. Sinusoidal jitter was the baseline scheme at the beginning for its simplicity. However, it has several limitations. The polar plot in figure 16 shows the resulting PSF centroid position using this approach compared to positions extracted from a subsample of Forerunner data. The rms is 10 mas in both cases. The difference between the two approaches is evident. The Forerunner time-series clearly shows that focal plane coverage is uniform (at least close to the center of the field) and that the star may exhibit significant PtV displacements during an exposure: this may cause an important fraction of star light to leak out from the coronagraphic mask. However, this effect could also result not so dramatic in the end: its impact depends of course on the amount of time spent by the star at high displacements compared to the total integration time. Another advantage of using the Forerunner time-series is that, as already said, although data refer to a single case of vibrations, they may be easily rescaled to any desired rms value. Finally, the third scheme is probabily the most comprehensive one, but it is difficult to apply in practice: in fact, being vibrations accounted for in the generation of phase screens means that every time one

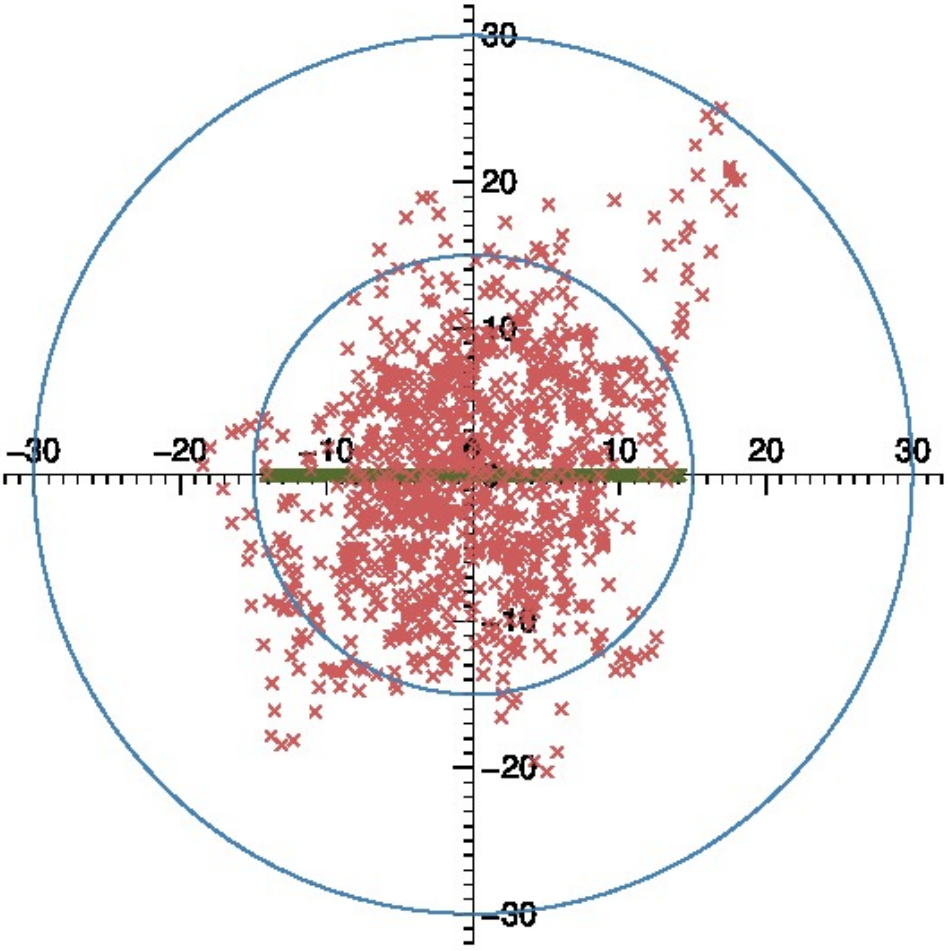


Figure 16: Polar plot showing the PSF centroid position in case of pure tilt (*green crosses*) and using a subsample of Forerunner data (*red crosses*). The rms is 10 mas in both cases. Data span 1 second at 1 ms cadence. Numbers on both axes are in units of milliarcseconds and the two blue circles have a radius of 15 and 30 mas.

wants to change the amount of vibration that a new data cube is required. This approach is thus suitable for investigation on a reduced scale, namely on a few case studies.
Further details on the schemes will be provided in chapter 5, in the section dedicated to the analysis of vibrations impact on performance based on complete simulations.

.

2.6 BROADBAND SIMULATIONS

The prescription delivers a purely monochromatic result. To generate a broadband image is thus necessary to define a given number of wavelengths inside the band, run the prescription for each of them and then properly

combine the resulting images. The simulator also offers this possibility. The final data cube in a polychromatic simulation has an additional dimension of size equal to the number of simulated wavelengths. Being the focal plane sampling constant in λ/D units with FFTs, images at different wavelengths are differently sampled. Instantaneous PSFs are thus saved only after the complex field is properly rescaled using damped sinc interpolation to conform all of them to a user-defined sampling. Finally, during pre-processing images are coadded and divided for the number of wavelengths.
All considered sources of optical aberrations in my simulations are introduced in the system entrance pupil plane, hence phase-amplitude mixing phenomena (Talbot effect) are not taken into account. However, each coronagraph is sensitive to chromatism somehow. This functionality of the code has been used to explore these effects in order to quantify their impact on instrument performance (chapter 6).

2.7 GLOBAL SOFTWARE ARCHITECTURE

As explained in Section 2.2, a single image is indeed the average of several hundreds of frames, each of which results from the propagation through the optical train described in 2.1. Depending on the employed number of coronagraphic planes, propagation may take from six up to nine Fourier transforms on a big computational matrix (1024×1024 pixels). In order to generate sequences of tents of images in a reasonable time, a large effort has been put into optimizing the software to fully exploit the computational resources of the available machines.
The i-th frame composing the image is uniquely determined by its corresponding input aberration ϕ_i. Since initial conditions for each frame are known from the beginning, frames can be generated in any order in principle. The flow non-sequentiality opens the possibility of parallelizing the code. Parallelization in IDL is not as straightforward as in other programming languages: the easiest way to accomplish it is by opening multple IDL sessions, each running as a separate process (child process). Following this approach, I developed and implemented in the simulator the parallelization scheme illustrated in figure 17. $P(\lambda_i, \phi_i)$ in yellow boxes stands for 'prescription' run for the i-th wavelength and the i-th input phase aberration. For each image of the sequence, the N calls to the prescription are split among four IDL sessions running in parallel. For displaying purposes, $N = 8$ is assumed. In order to minimize memory allocation, the outputs of each execution of the prescription are not stored, but co-added directly internally to each session. When all the sessions have completed their tasks, they are closed and their outputs passed to the main level. The final image is then the sum of the four outputs divided by N. If the simulation is polychromatic, the process is repeated for each wavelength before passing to the successive image. The whole process takes roughly 12 minutes per second of integra-

tion per wavelength at 1 kHz rate in the worst case where all coronagraphic planes are included.

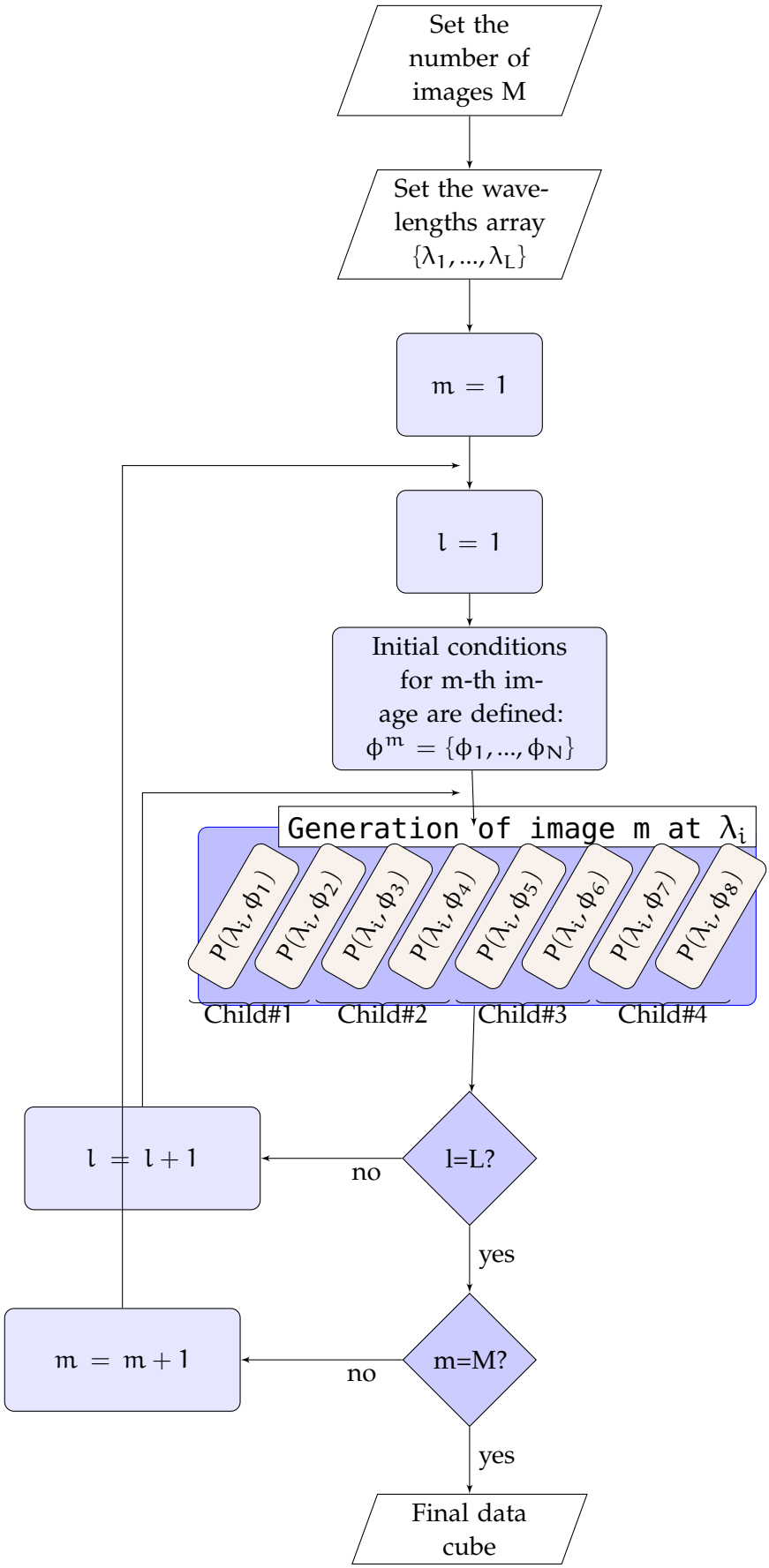


Figure 17: The parallelization scheme adopted by the simulator. For each image of the sequence and for each wavelength, the N calls to the prescription are split among four IDL sessions running in parallel.

3 DATA REDUCTION PIPELINE

In order to derive instrument performance from simulated images, I developed a data-processing pipeline which complements the physical optics propagation code. In this chapter I am describing the pipeline, detailing all the steps transforming the 'raw' outputs of the Fresnel simulator into a final product suitable for applying the desired post-processing algorithm. A structured pipeline is mandatory when dealing with post-processing algorithms and noise calculations: for results to be considered realistic, it is fundamental to include all those processes playing a role in shaping real images (photon and read-out noises for example).
The pipeline is conceived mainly for reproducing Angular Differential Imaging (ADI) processing. ADI is the algorithm we identified as the baseline to derive SHARK-NIR detection limits of exoplanets. All steps in this direction are detailed in sections 3.1 to 7.1.
Finally, no post-processing has been tested so far for the non-exoplanetray cases: the reduction and derivation of a performance metric, much simpler in these cases, is described in section 3.4.

3.1 RAW DATA

The raw input to the pipeline is a sequence of images (a data cube). The number of simulated images is limited by the total number of phase screens and by the single-frame DIT. Using a DIT of one second, sequences up to 100 frames can be generated by the simulator. However, because of computational time, the baseline length assumed throughout the design phase is 30 images.
Together with the scientific sequence, the pipeline also needs a PSF out-of-coronagraph (or non-coronagraphic PSF) under the same observing conditions. As explained is section 3.3.3, this image gives the star signal reduced for the coronagraph throughput and is necessary to compute the detectability curve.

3.2 PRE-PROCESSING

The raw image generated by the Fresnel simulator contains the time-averaged wavefield intensity per unit incident power at different locations on the detector plane. Pre-reduction consists in converting these raw data into a real CCD output. This is accomplished in five steps:

1. The image values are first converted into units of power density by multiplication for the incident power P_i in photons/s from the star. The calculation accounts for LBT and SHARK-NIR throughputs and is given by the formula:

$$P_i = 1.51 \cdot 10^{-3} \cdot A \cdot \eta_{LBT} \cdot \eta_{SHARK} \cdot \frac{\Delta\lambda}{\lambda} \cdot 10^{-0.4 \cdot (M - M_0)} \tag{5}$$

   where:

   - A is the telescope effective collecting area.
   - η_{LBT} is the throughput of LBT M1, M2 and M3 mirrors, given by the product of their respective reflectivities.
   - η_{SHARK} is SHARK-NIR throughput (coronagraph excluded). This number is estimated via a dedicated budget including all possible internal light-loss sources.
   - M is the magnitude of the star in the scientific filter.
   - M_0 is the zero-point magnitude in the scientific filter.
   - $\frac{\Delta\lambda}{\lambda}$ is the filter fractional bandwith.

   Values for these constants are reported in table 5. It is important to point out that, while the Fresnel simulator requires in input the magnitude in the visible (R band) to set the AO correction, the one in the scientific channel M comes into play only at this stage in the pre-reduction of the simulator output. The relation between the two depends of course on the spectral type of the target. Throughout the design phase, I assumed as a reference a late-type star in the Taurus-Auriga star forming region. For this kind of targets, the following approximate relations hold:

$$H = R - 2 \tag{6}$$

$$J = H + 0.2 \tag{7}$$

2. The image is multiplied for the Detector Integration Time (DIT). According to equation 2 in chapter 2, this operation delivers the total number of stellar photons per unit surface reaching the detector plane over the exposure.

3. Counts are integrated over the H2RG Teledyne 18 µm pixels. The result is the number of stellar photons striking each detector pixel. Because of the limited size of the computational matrix, only a portion of the detector surface can be covered.

4. Sky background is added to the image. I used Mount Graham sky brightness measurements, namely 15.82 and 14.29 mag/arcsec^2 in J and H band, respectively (Pedani 2014). Thermal background is also implemented. However, its effect is negligible according to our estimates.

5. Image are converted into photo-electrons (phe-): this is accomplished by adding Poissonian fluctuations on the stellar photons.

6. Detector read-out noise is added. I assumed a stationary white Gaussian noise model, with a standard deviation of $15e^{-}$.

7. Images are finally converted into Data Numbers (DN) by specifying a gain factor. The gain has been usually assumed to be 1.

Table 5: Parameters used in the computation of the total number of photons/s from the star reaching SHARK-NIR detector.

η_{LBT}	0.70
η_{SHARK}	0.50
$\Delta\lambda/\lambda$	$0.159(J) - 0.125$ (H)
M_0	$0.03(J) - 0.03(H)$

3.3 POST-PROCESSING

3.3.1 The observation

For exoplanets detection, the baseline data-processing algorithm is Angular Differential Imaging (ADI). ADI is a technique routinely used on direct imaging data to suppress the quasi-static structure present in the PSF of the star (Marois et al. 2006). The acquisition of a set of images (usually up to a few hundreds) is performed with the instrument rotator turned off. In this way, an hypothetical off-axis source is made to rotate with respect to the star speckle floor, which is in turn static. Image subtraction then strongly attenuates starlight contamination while highly preserving the planet flux.
Because of the very small temporal sampling dictated by the AO framerate, simulating more that one hundred seconds of closed-loop operation is very demanding from the computational point of view. ADI observations typically span up to a few hours, in order to maximize field rotation and, as consequence, minimize planet self-cancellation. If ADI is applied based on the real timeline of the simulation, then planet cancellation would be very close to 100% even at very high angular separations from the host star. A solution to the problem is find putting ourselves in the following situation: *images in a simulated sequence are not acquired consecutively, but rather they are distributed to cover a lead time sufficiently long for ADI to be applied*. Of course, this is quite an unrealistic situation for two main reasons. First, images are acquired at very slow temporal cadence: 100 seconds ideally distributed over one hour means $\sim$ 95% overhead time). Second, the need to dilate the timeline necessarily forces us to assume the atmosphere as highly stable for long times. This is because seeing fluctuations in atmospheric data cubes are very small (see section 2.3). However, these two unrealisic aspects somehow go

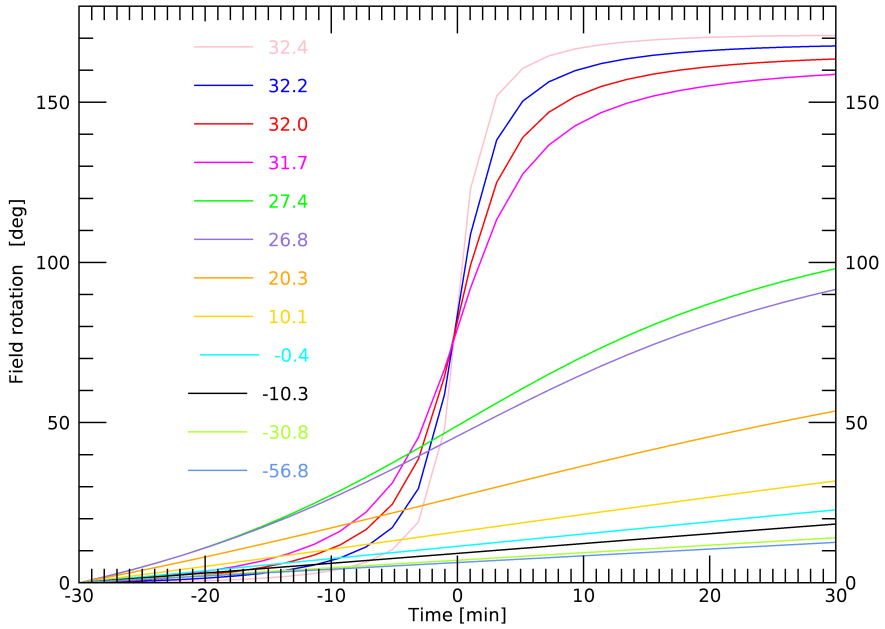


Figure 18: Field rotation at LBT as a function of time for different values of the target declination.

in the opposite direction: in real observations, the additional noise coming from changes in the atmosphere might be partially compensated by the acquisition of a much larger number of frames. In this sense, the performance obtained assuming a short and 'diluted' sequence may not be dramatically different from the one obtained in a more realistic case.
According to the new timeline, I associate an hour angle to each frame in the sequence assuming that culmination (HR $= 0$) happens at the midpoint of the observation. For example, if the sequence is made to span one hour, then hour angles are equally distributed between -0.5 and $+0.5$ hours. Assuming a declination for the target, parallattic angles can be finally computed from hour angles, to be used to de-rotate images. By changing the declination it is possible to reproduce any desired field rotation (figure 18). Targets in Taurus-Auriga range in declination from $\sim 18°$ to $\sim 30°$. At the highest of these values, targets are close to zenithal culmination at LBT latitude. This is a very suitable condition for ADI observations because of the large FoV rotation. To be conservative, however, in detection limit curves showed in this thesis I will assume the smallest possible declination of 18°, corresponding to $\sim 50°$ total field rotation.

3.3.2 The ADI algorithm

The ADI algorithm is structured in three steps (figure 19):

I. **Subtraction of a reference PSF from each image in the datacube**
The reference PSF is extracted directly from the data. The pipeline implements three different variants of the algorithm, differing in the way this operation is performed:

- *Single-median*: the reference PSF is the same for each image and it is simply the median of the whole sequence (classical ADI).
- *Multiple-medians*: the sequence is divided into two or more subsets and the single-median subtraction is applied separately for each of them. In this way, reference PSFs are closer in time (and hence more correlated) to the images they are subtracted from. The drawback is the enhanced planet cancellation (section 3.3.4)
- *Principal Component Analysis (PCA)*: PCA algorithm is based on a statistical representation of each frame as a linear combination of its principal orthogonal components. These components are estimated by diagonalization of the covariance matrix associated to the signal. Its application has been reported to be very effective particularly by manipulation of the number of principal components to maximize the signal from the planet near its host star [RD11]. The number of modes used for the PCA analysis is associated to the variance of the corresponding principal component.

II. **De-rotation of image differences**
Images are de-rotated to align the fields of view in preparation of the final step. De-rotation is performed according to images parallattic angles. FoV rotation is critical in defining the amount of planet self-cancellation (see section 3.3.4).

III. **Median-combination of de-rotated images**
By median-combining the image differences, noise is further attenuated at best by a factor $\sqrt{N}$, where N is the number of images in the sequence.

The baseline for generating the detection limit curves that I will show in the following chapters is classical ADI with 30 images and 50° of FoV rotation. However, the developement of a detailed pipeline to reduce simulated data offers a unique opportunity to carefully investigate the parameter space related to post-processing. For example, Carolo et al. 2016 compared different ADI variants based on simulated data, while in Carolo and Vassallo in prep we discuss how to optimally process ADI data in different observing conditions to maximize the SNR at any given angular separation.

3.3.3 The performance metric

The metric adopted is the 5-σ detection limit: namely, this quantity is defined as the planet-to-star flux ratio $\mathcal{F}$ yielding a SNR of five after post-

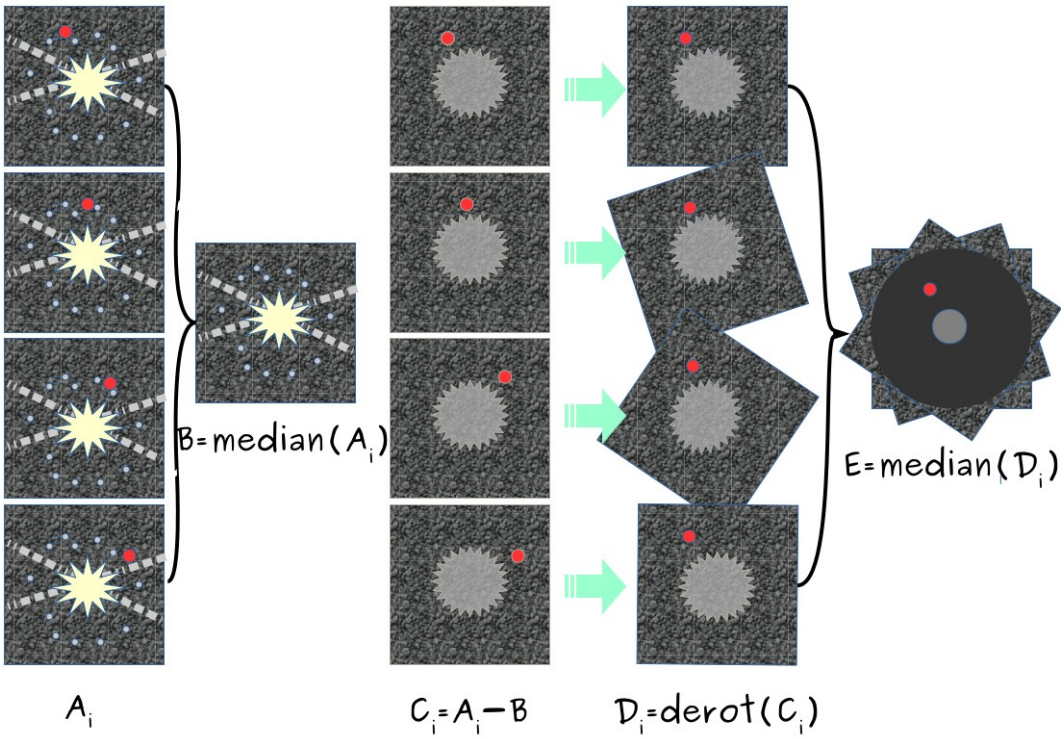


Figure 19: Graphical sketch of the ADI algorithm.

processing. There exists at least two approaches to compute it. One way is to introduce into the simulated sequence some fake planets at different angular separations and flux ratios and calculate the final SNR at their (known) positions after ADI. This method thus computes the SNR starting from the flux-ratio $\mathcal{F}$. This approach is probably the most rigorous one, but it is also time consuming since the value of $\mathcal{F}$ of the planets has to be tuned iteratively until a SNR of exactly five is reached. Alternatively, one can face the problem from the opposite point of view, namely starting form the planet SNR. Its value at a given position r in the coronagraphic FoV can be expressed as:

$$\mathrm{SNR}(r) = \frac{\mathcal{P}^*}{\mathrm{Noise}} = \frac{\mathcal{P} \cdot \mathcal{T}(r) \cdot \mathcal{B}(r)}{\mathrm{Noise}(r)} \tag{8}$$

where:
$\mathcal{P}$ is the original planet flux
$\mathcal{P}^*$ is the residual planet flux in the final ADI-processed frame
$\mathcal{T}(r)$ is the coronagraphic throughput
$\mathcal{B}(r)$ is the photometric bias or self-cancellation factor.

Given a targeted SNR, then the associated planet-to-star flux ratio $\mathcal{F}$ is simply given by inversion of (4):

$$\mathcal{F}(r) = \frac{\mathcal{P}}{\mathcal{S}} = \frac{\mathrm{SNR} \cdot \mathrm{Noise}(r)}{\mathcal{S} \cdot \mathcal{T}(r) \cdot \mathcal{B}(r)} \tag{9}$$

In the formula:

- The term $\mathcal{S} \cdot \mathcal{T}(r)$ is calculated from the out-of-coronagraph PSF. This PSF, in fact, is simply a stellar image through the coronagraph without focal plane elements. Its flux is then the one of the star times the coronagraph throughput. Specifically, I decided to adopt for the metric the *peak* of this PSF, instead for example of the integrated flux over a given area. In those cases in which the throughput changes with angular separation, a correction a-posteriori is applied to this peak.

- The noise as a function of angular separation is calculated after applying ADI to the simulated sequence. No fake planet is injected. For each pixel in the final ADI processed frame, the associated noise is computed as the standard deviation of counts in two close-by regions, as showed in figure 20. This is the same method used for example in Mesa et al. 2011 to simulate planet detection with SPHERE-IFS. It might seems strange that such a method is used when no planet is truly injected. However, I tested different approaches (e.g. simply taking the standard deviation on concentrinc rings) and all leads more or less to the same results. This approach has the advantage to produce a bi-dimensional map, allowing to spot eventual non homogeneities in noise distribution along the azimuthal coordinate. In the simplest cases, a 1-D radial profile is simply extracted via azimuthal averaging.

- The self-cancellation factor profile $\mathcal{B}(r)$ is computed separately using the method illustrated in section 3.3.4.

This approach is of course less straightforward than directly injecting fake planets into the simulated sequence. However, it has the advantage of being much faster. For this reason, it is the one I adopted for computing detection-limit curves shown in this thesis.
I conclude by pointing out that when the planet falls very close to the edge of the focal plane mask then the metric is no more completely reliable. In such a situation, an additional term at the denominator of equation 9 should be added to account for the signal loss due to partial obscuration of the planet disk. Moreover, in this case also the noise calculation is difficult because the two regions of computation are inevitably contaminated by pixels behind the mask. Because of the difficulty to model this effect, I decided to plot all detection curves starting half a PSF core away from the IWA of the coronagraph (e.g. 20 mas in H band) when a hard edge mask is present.

3.3.4 The self-cancellation factor

Using small subsets to generate the reference PSF can help attenuating the speckle noise, but it also results in a growing risk of planet removal if not

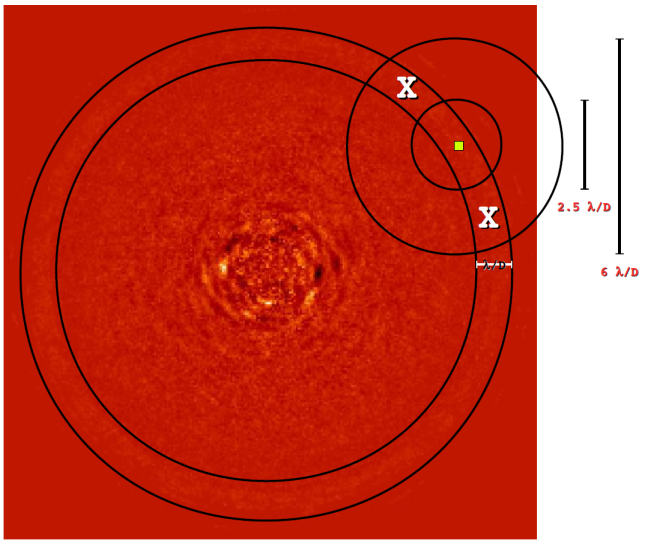


Figure 20: Geometry of noise calculation in ADI-processed frames. Noise associated to the green pixel is calculated taking the standard deviation of counts inside the two areas marked with an 'X' in the image. These areas are chosen to be at same distance from the star as the pixel itself, but separated in azimuth in order to avoid possible contamination from a planet near the pixel position.

enough field rotation has occurred in the subset itself. In general, self-subtraction depends on planet separation from the host star: the drop of its light is bigger at small separations and decreases at large separations. This is due to the fact that the same FoV rotation corresponds to a slower linear motion of the planet at small radial separations from the star with respect to larger separations. This effect results in a subtraction of both speckles and the planet signal near to the star.

The amount of self-cancellation is calculated empirically by using 'probe' planets: basically, I apply ADI to a sequence of images containing only planets at different separations from the star. In these images, planets are properly rotated according to parallattic angles. Their residual (post-ADI) signal is then compared to the initial one to generate a cancellation profile. As already pointed out, the amount of cancellation depends on planet angular separation and, at fixed separation, on the ADI variant used and on the overall FoV rotation. Additionally, there is also a somewhat weaker dependence on the number of images in the ADI sequence. Figure 21 reports the self-cancellation profile using a classical ADI processing based on 30 images in the case of a FoV rotation of 50°. Residual planet flux is stable to $\sim 95\%$ until ~ 200 mas, where it starts decreasing very rapidly down to $\sim 30\%$ at 80 mas.

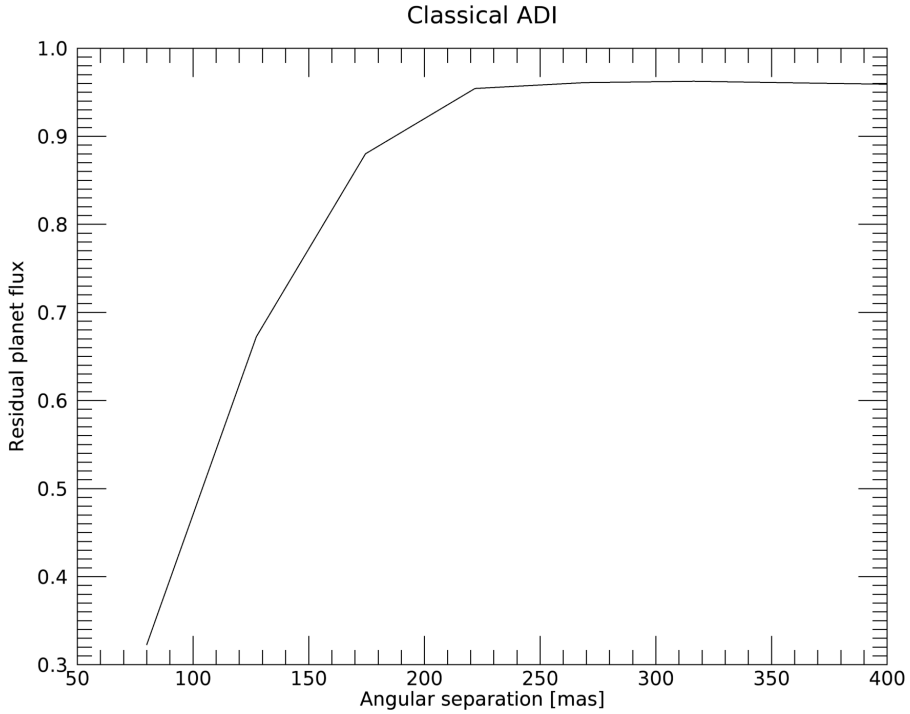


Figure 21: Planet self-cancellation as a function of angular separation applying classical ADI on a 30-images sequence with a FoV rotation of 50°. Cancellation starts being significant below ~ 200 mas, but a small flux loss of ~ 5% is experienced even at large separations.

3.4 FIELD-STABILIZED OBSERVATIONS

The non-planetary cases will require field de-rotation. In addition, sources are usually extended. No post-processing specifically devoted to this kind of observations has been investigated so far. The metric I adopt in these cases is thus the raw contrast. I generate with the simulator a single scientific frame, together with its corresponding out-of-coronagraph PSF. The intensity as a function of position in the scientific image is then normalized to the peak of this PSF. In doing this simple intensity ratio I am not considering any noise contribution and, in general, any factor contributing to the instrument limit magnitude. Hence, especially at very high angular separations, the contrast may be only theoretical. However, this metric is still commonly used for estimating coronagraphic performance and can be considered a good first-order tool to assess whether the scientific requirements can be fulfilled or not.

4 CORONAGRAPHIC DESIGNS

Coronagraphy is a very wide field: several techniques have been proposed through years, with a huge effort both theoretical and experimental. A relevant reference in this context is e.g. Guyon et al. 2006. A few different solutions will be finally implemented in SHARK-NIR, thanks to the presence of wheels in all of the three coronagraphich planes. The goal of this study is to converge on a final suite. The final selection is based on the simulated performance. Several factors might influence the performance, with each technique generally behaving differently in response to each of them:

1. Geometry of LBT pupil
2. Vibration environment at LBT
3. Upstream aberrations (AO residual and NCPA)
4. 50 millimeters separation between the position of the first intermediate pupil plane and the apodizers wheel
5. Finite tolerance in the alignment of the apodizer with respect to the telescope pupil
6. PSF stability in the coronagraphic focal plane
7. Chromatism

As a result of a preliminary study, we first narrowed the focus to a small subsample of techniques. There are three planes in SHARK-NIR that can be used for coronagraphy: two pupil planes and one focal plane in-between. All techniques that are not compatible with the opto-mechanics of the instrument are of course immediately discarded. The Lyot concept and its improvements with both amplitude and phase masks are the most suitable options. In this chapter I will introduce these techniques and then detail the design steps for each of them. Before starting, I will introduce some fundamental concepts in the framework of coronagraphic design:

- **Contrast**
  The efficiency of a coronagraph resides in its capability to suppress the incoming light selectively with the position of the source in the field: it has to block on-axis light while transmitting off-axis radiation. Contrast is the ratio of this two components. In a formula:

$$C(r) = \frac{I(r)}{\widehat{I}(0) \cdot |M(r)|^2}$$

$\widehat{I}(0)$ is the peak intensity of the off-axis PSF, which is computed removing the focal plane mask from the optical train. This peak is multiplied a-posteriori by an intensity transmission term (namley, the square of the focal plane mask amplitude transmission M(r)). For those coronagraphs acting on phase rather than amplitude in the focal plane, this term is neglected since off-axis transmission can be usually assumed to be one at all angular separations above a few λ/D. Finally, I(r) is the residual intensity in the coronagraphic (on-axis) PSF.
In the ideal case of no aberrations, I will use the expression 'Nominal contrast'. Nominal contrast is quick to compute and represents the metric commonly used when designing a coronagraph. Although it only accounts for diffractive effects, it should be stressed that the amount of diffracted light in the focal plane regulates the speckle noise once aberrations are introduced: in this sense, this metric is a good first-order estimate of the performance we shall expect after running complete simulations.

- **Throughput**
  This parameter quantifies the attenuation experienced by an off-axis source because of the coronagraph. One way to fairly compare off-axis attenuation when different coronagraphic concepts are compared is to use the PSF throughput, namely the encircled energy within 1 λ/D radius for the off-axis PSF compared to the non-coronagraphic one. In general, this parameter is a function of Strehl and position in the focal plane. Unless otherwise specified, I will always refer to the maximum coronagraphic throughput, namely I am assuming an unaberrated system and a source located far away from the center of the field.

- **Discovery Space**
  Also named coronagraphic FoV, it is the the region of the image plane where high-contrast is generated. The inner boundary, named Inner Working Angle (IWA), is arbitrarily defined as the separation $\tilde{r}$ from the star at which the coronagraph throughput has dropped to 50% of its maximum.
  $$|M(\tilde{r})|^2 = 0.5$$
  In some coronagraphs, the high contrast region has a finite extent: the outer boundary is named Outer Working Angle (OWA). Its definition is the same as for the IWA.
  In some special cases, also the azimuthal extent shall be specified: the coronagraph may not generate high contrast all around the star (360° extent), but only in specific areas of the FP (non-annular discovery space). This is for example the case of some special Shaped Pupil coronagraphs, as I will detail later on in this chapter.

- **Tolerance to pupil/focal plane components misalignments**
  This parameter quantifies the sensitivity of the coronagraph to displacements of some of its components from their nominal position.

- **Bandwidth**
  This parameter defines the maximum wavelength range over which the performance does not differ from the nominal monochromatic one more than a given reasonable amount.

4.1 CLASSICAL LYOT

The principle

This coronagraph was first proposed by Lyot back in the thirties to observe the solar corona outside eclipses. It consists of a hard-edge occulting disk in the focal plane, usually a few λ/D in radius, and a stop in a downstream pupil plane (the Lyot stop) to mask the light diffracted by the occulter near the edge of the pupil. The mask amplitude transmission is thus simply 0 inside the occulter and 1 outside.

The design

In a Classical Lyot coronagraph, the parameters to be specified are the occulter and Lyot stop dimensions. The presence of the occulter in the focal plane makes the dark zones of LBT pupil to appear bright in the downstream pupil plane. The ideal Lyot stop should then trace the shape of the secodary mirror and its supporting arm, resulting then non circularly symmetric. The symmetry of optical components is crucial in all coronagraphic designs, since it is directly related to the operational mode in which the coronagraph can be used. In fact, since the telescope pupil is not rotating inside the instrument, any non circularly-symmetric component in the pupil plane shall not rotate as well during the exposure. This means that the coronagraph can only work if the instrument de-rotator is switched off, namely in pupil-stabilized mode. This mode will be used in SHARK-NIR only for exoplanet research, thanks to the possibility to exploit ADI. Operating a classical Lyot coronagraph in pupil-stabilized mode is of little utility for exoplanet research because of the modest contrasts achievable with this technique. For this reason, I designed the Lyot stop as a classic circularly symmetric element, to be thus used in field-stabilized observations. In addition to pupil undersizing, the central region of the pupil is also masked in correspondence of the secondary mirror shadow. The central suppressor, as I will call it, is thus the third free parameter, together with focal plane occulter and pupil stop sizes.

The IWA corresponds to the radius of the focal plane occulter. The choice for this parameter is driven by the science. The goal is 150-200 mas (see 1.4). A good choice to match the requirement is 3 λ/D at 1.6 μm, corresponding to $\sim$ 120 mas.
Lyot stop and central suppressor are then optimized performing a run of simulations: the two parameters are varied at a step of 5% and, for each

combination, I generate both a coronagraphic and an out-of-coronagraph image. The performance metric is the average contrast inside the AO corrected field. Contrast is calculated by normalizing intensity in the coronagraphic image to the peak of the out-of-coronagraph one. Finally, the bidimensional contrast plane is filled in-between the discrete 'nodes' using the cubic convolution interpolation method. Figure 22 shows the generated contrast map.
Simulations are at 1.6 μm and include atmospheric residuals and 30 nm of NCPA. The GS magnitude is 8, integration time is 1 second. The red region in the map identifies the parameter combination that maximizes the contrast. There is of course some dispersion around the optimal solution. In this case, the optimal Lyot stop has a diameter which is 84% of the pupil one, while the central suppression is 11% (as the secondary mirror). Nominal contrast in the optimized configuration is plotted in figure 23. This technique does not allow to reach high starlight rejection at small angular separations.
An IWA of 120 mas corresponds to 4 λ/D in J-band, meaning that the light distribution in the Lyot plane is different from the H-band case. However, in presence of aberrations, simulations show that the optimized parameters does not differ significantly in the two cases. I can thus conclude that I do not expect chromatism to be a concern and that such a coronagraph can operate both in J and H bands.
It is also important to point out here than this optimization depends on the Strehl Ratio. In the case of the Classical Lyot coronagraph, this dependence has not been further explored since this technique was soon abandoned in favour of its variant with Gaussian amplitude modulation in the focal plane (Gaussian Lyot).

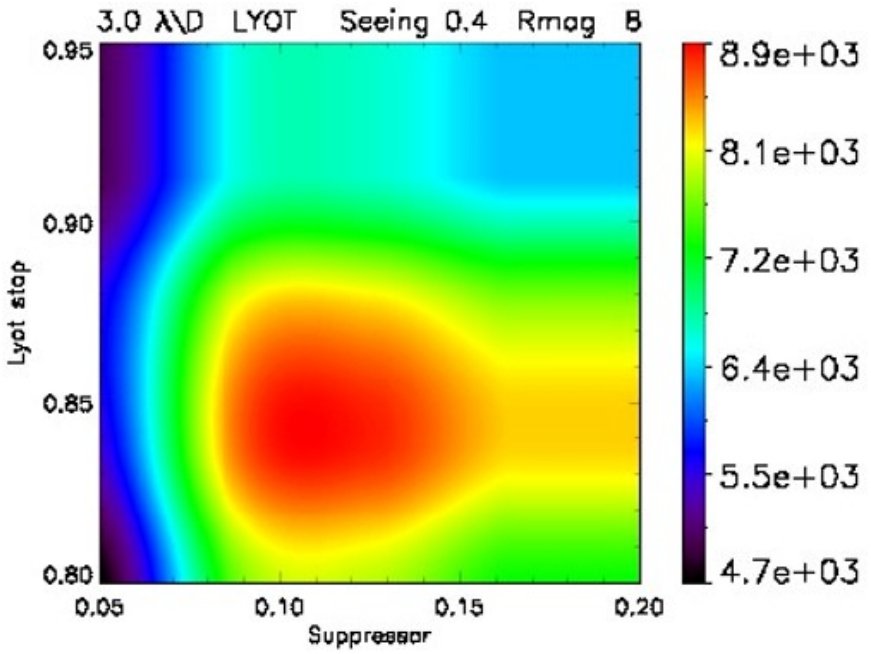


Figure 22: Contrast map for the optimization of Lyot stop and central suppressor for the Classical Lyot coronagraph. The red region in the map identifies the parameter combination that maximizes contrast inside the AO corrected field.

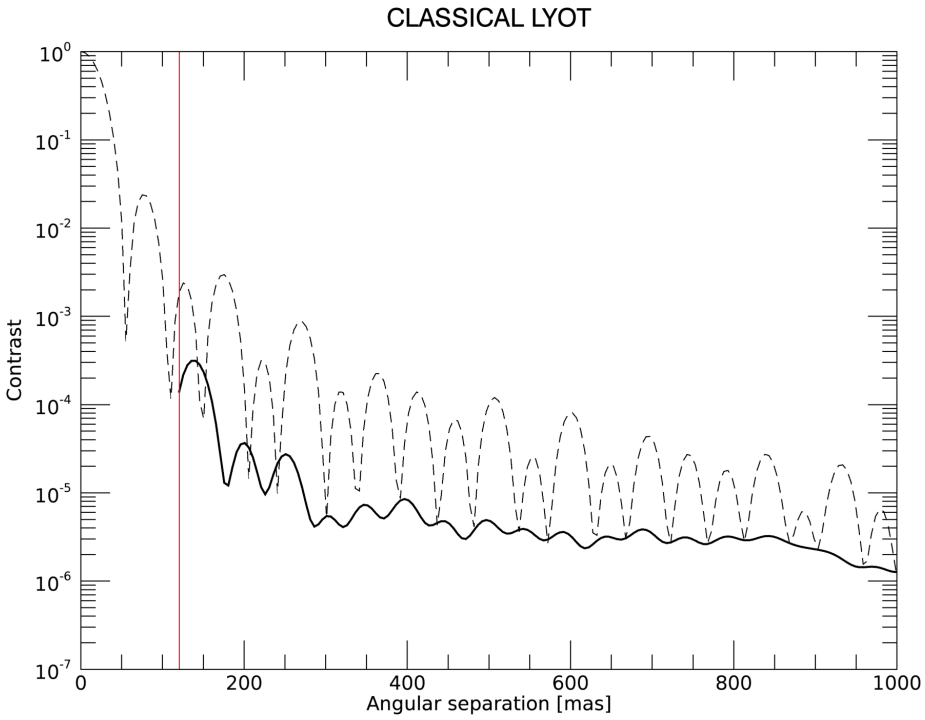


Figure 23: Nominal contrast obtained with the Classical Lyot coronagraph (continuous line). Dashed line is the off-axis PSF, while the red vertical line marks the position of the IWA (120 mas in H-band).

4.2 GAUSSIAN LYOT

The principle

In the classical Lyot configuration, amplitude transmission is a simple boxcar function. Several other transmission profiles have been proposed through years to control diffraction. In a Gaussian Lyot coronagraph, the electric field amplitude in the image plane is spatially modulated according to a Gaussian transmission profile. The mask amplitude transmission is given by the formula:

$$M(r) = 1 - e^{-(\frac{r}{\alpha})^2}$$

Where r is the angular separation and α is a free parameter. It can be shown that, following the definition, the IWA is given by:

$$IWA = \alpha \cdot \sqrt{\ln(2+\sqrt{2})}$$

This equation is used at reverse to calculate the parameter α given the desired IWA. An example of transmission profile, both amplitude and intensity ($M(r)^2$), is shown in Figure 24 for a 120″ mas IWA. This corresponds, in SHARK-NIR coronagraphic focal plane, to ~ 100 μm. The IWA definition is somehow less intuitive in this case with respect to a hard-edge mask: in principle, a Gaussian mask can work even inside its IWA, since transmission is zero only at its very center.

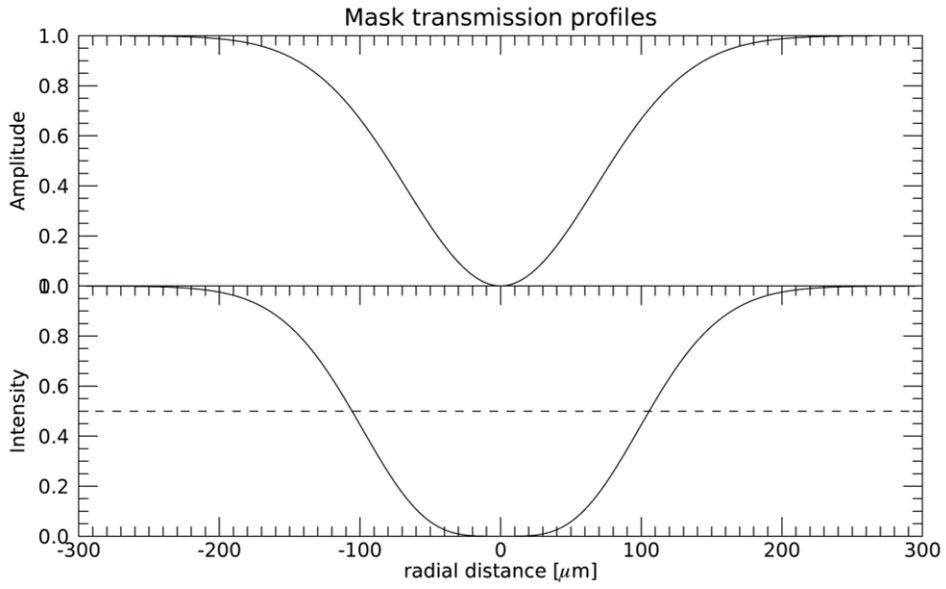


Figure 24: 1-D amplitude and intensity transmission profiles of a Gaussian mask designed for a 120 mas IWA. The IWA location is identified by the horizontal line in correspondence of 50% intensity transmission and corresponds to ~ 100 μm considering SHARK-NIR plate scale in the coronagraphic focal plane.

The design

All considerations done for the the Classical Lyot coronagraph shall be repeated here. The IWA is still 3 λ/D and the coronagraph can operate in both J and H bands. A Lyot stop in the downstream pupil plane is still mandatory to block the residual diffracted light. Its size has been optimized using contrast maps. As anticipated in the previous section, it is not possible to find the optimal configuration in an absolute sense. Table 6 reports the optimal parameters in a few different cases. I let the SR to vary by changing the target brightness and/or the atmospheric seeing. The corresponding maps are shown in Figure 25. It can be seen that, as SR get worse, the optimal configuration moves towards the top of the maps, namely in the direction of a small pupil undersizing (and thus high throughput). On the other hand, the secondary suppressor size remains constant on a value close to the secondary obscuration. In our intentions, such a coronagraph is not supposed to operate in very high SR regime because of the moderate achievable contrast compared to other techniques. As a consequence, we opted for a compromise between the three invesigated case: $85\% - 12\%$. A slight oversizing of the secondary suppressor is considered so to make the coronagraph more robust to misalignments.

Contrast in this configuration is shown in figure 26. In the first ~ 1 λ/D after the IWA, this coronagraph outperforms Classical Lyot. Moreover, also the region inside the IWA is theoretically accessible.

Table 6: Optimal values for Lyot stop and central suppressor of the designed Gaussian Lyot coronagraph. The optimization is performed in different conditions in order to explore the impact of SR.

R	seeing	SR	Lyot stop	Suppressor
8	$0.4''$	84%	79%	10%
8	$0.8''$	75%	86%	10%
10	$0.6''$	70%	87%	10%

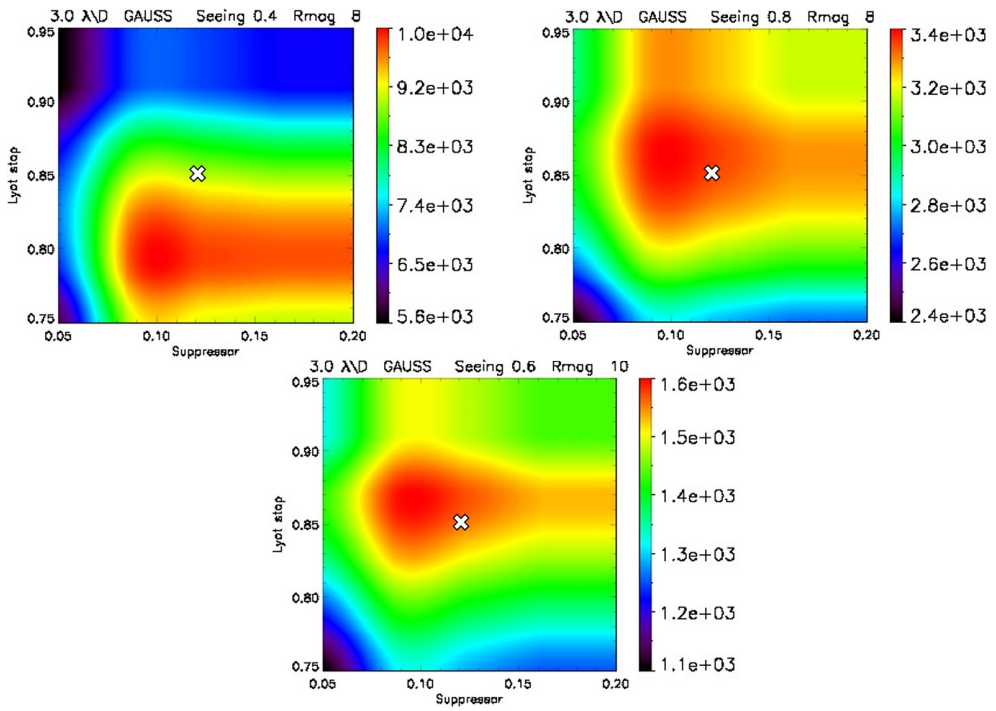


Figure 25: Contrast maps for the optimization of Lyot stop and central suppressor for the Gaussian Lyot coronagraph. The white cross marks the position of the finally chosen values.

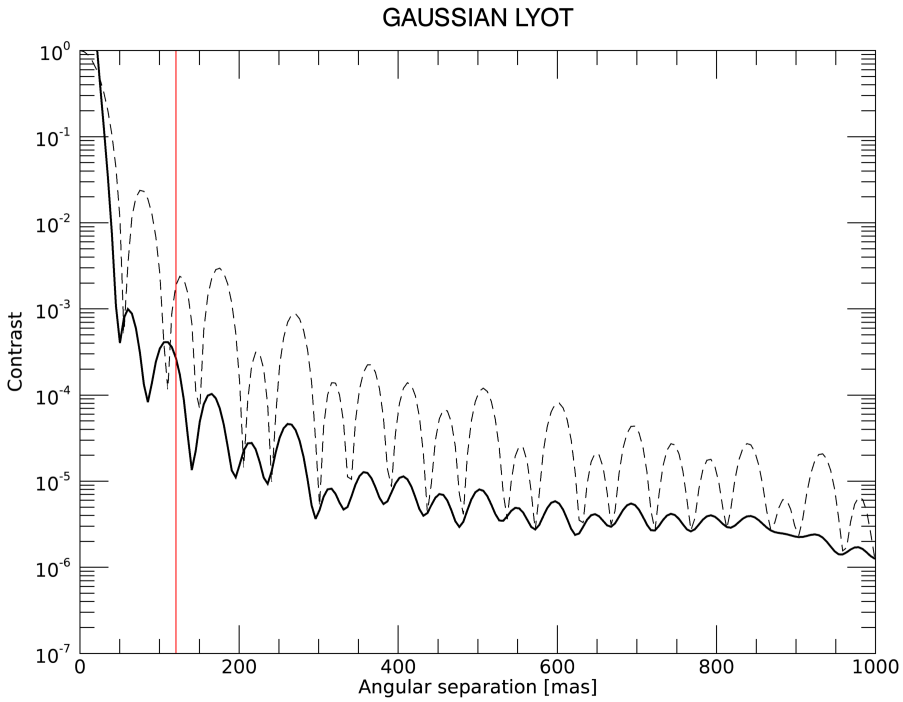


Figure 26: Contrast obtained with the Gaussian Lyot coronagraph (continuous line). Dashed line is the off-axis PSF, while the red vertical line marks the position of the IWA (120 mas in H-band).

4.3 SHAPED PUPIL

The principle

As suggested by its name, the Shaped Pupil (SP) technique consists in reshaping the telescope pupil so to generate dark areas in the focal plane via distructive interference. Stricly speaking, ''shaping" here means to constrain transmission to be either 0 or 1 across the pupil (for a review of apodization techniques see e.g. Kasdin et al. 2003). With currently available algorithms, masks can be designed for any arbitrary pupil geometry to generate regions of high contrast in the image plane tailored to any desired shape or size (Carlotti et al. 2011). These algorithms compute the pupil shape which fulfills the specifications at a minimum cost in terms of blocked light. From the mathematical point of view, throughput maximization can be formulated as a linear problem: the apodizing function A(x,y) is build so that the on-axis field $E(0,0)$ in the focal plane is maximized, subject to the constraints:

$$0 < A(x,y) < 1 \tag{10}$$

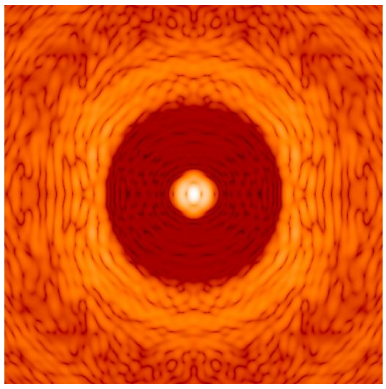
$$-10^{-\frac{c}{2}} < \frac{E(u,v)}{E(0,0)} < 10^{-\frac{c}{2}} \tag{11}$$

where 'c' is the targeted contrast. The second constraint is of course satisfied only inside the desired discovery space. It has been shown that, using sufficiently big arrays for optimization, transmission values tend to cluster around either 0 or 1. The apodizers computed in this way are thus quasi-binary: practically, pixels are artificially rounded as a final step to make them perfectly binary.
Figure 27 shows an example of apodizer designed for LBT pupil geometry, together with the generated 2-D and 1-D (azimuthal average) PSF profiles. The apodizer is designed for a contrast of 2×10^{-7} from 4 to 15 λ/D. According to equation 11, the contrast to be specified to the algorithm is the ratio between residual intensity in the high-contrast region and the PSF peak. As it will become clear in a moment, this quantity is not necessarily the coronagraph nominal contrast (as defined in the previous section). In the discussion that follows, where using the term contrast I will always refer to the former since it is the one inherent to the apodizer design.

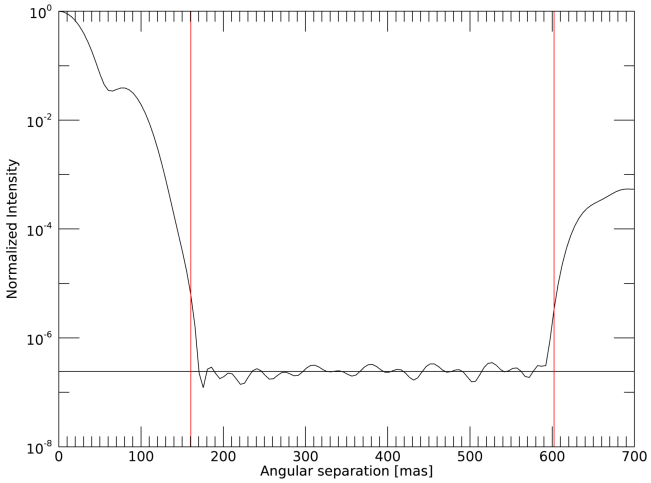
The design

In principle, the only binary mask is sufficient to create high contrast. However, there are two main drawbacks: first, the shaped PSF features a bright central core, which may easily saturate the detector. To give an idea, with an apodizer having 25% throughput, I estimated more than one milion photons per second of integration striking the central pixel for a $H = 6$ target in high-Strehl regime. The second important drawback is that NCPA downstream of the coronagraph fully contribute quasi-static speckles, since the stellar light

(a) Apodizer.



(b) Fourier Transform of re-shaped pupil.



(c) Radial intensity profile in the focal plane.

Figure 27: Example of Shaped Pupil design. Contrast of the plateau is 2×10^{-7}, while the discovery space ranges from 160 to 600 mas in H-band (4 to 15 λ/D).

is only re-distributed but not damped. In order to mitigate these effects, we can exploit the focal plane downstream the apodizer to put a hard edge occulter that masks the PSF core. A Lyot stop in the second intermediate pupil plane then blocks the residual light diffracted by the occulter, as in the classical Lyot configuration. The three-planes configuration also allows to go deeper in contrast with respect to the apodizer-only configuration.

The design of a Shaped Pupil depends on several parameters: in order to explore the interplay between them, our collaborators at IPAG provided us with several apodizers. First of all, the optimization algorithm can easily account for LBT pupil geometry and apodizer alignment tolerance constraints. They both translates in a re-definition of the support A(x,y) for optimization: the desired tolerance is achieved by simply thickening the spiders and central obscuration. In this way, the tolerance is obtained at the cost of a reduced throughput. According to a detailed error budget, the estimated precision in centering the apodizer with respect to the telescope pupil is of

0.23% of the pupil diameter. Applying a factor $\times 2$ as a safety margin, this translates into a specification for all the designs of 0.5%.

An apodizer is designed once an IWA and OWA are specified as delimiters of the focal plane region in which the contrast is constrained. The focal plane mask will be then dimensioned on them so to block all the light falling outside of this region. This mask is thus composed of an occulter of size equal to the IWA and a field stop of size equal to the OWA. Since these boundaries are λ-dependent, a choice has to be made on the band the coronagraph will operate. Alternatively, we might consider more than one focal plane mask for one apodizer. Since the second option is not doable because of the limited number of slots in the focal plane wheel, we decided to operate the coronagraph in H-band. This band is the one designated for exoplanet research. In fact, this technique allows to achieve very good contrasts at quite small angular separations: hence, it is a good candidate for this kind of science. Moreover, since apodizers are not circularly symmetric (because of the asymmetry of of LBT pupil, see figure 27), this technique requires pupil stabilization. This observing mode is exactly the one commonly used in exoplanet research, because of the possibility to exploit ADI.

If we want to address this kind of science with this coronagraph then the goal is to reach a ~100 mas IWA. Pushing towards these very small IWAs necessarily forces us to accept compromises: either the contrast requirement shall be relaxed or the discovery space shall be reduced. This situation is clearly exemplified in table 7, where I compare three different apodizers designed for the same contrast, but at increasing IWA from 2 to 3 λ/D. If we ask for a given contrast, then obtaining it at 2 λ/D is definitely more demanding than at 3 λ/D. The result is that the obtained throughput is smaller in the first case, as expected. The 2 λ/D design is actually so demanding that the tolerance requirement has been even relaxed to 0.5% and OWA reduced to 4 λ/D in order to reach a throughput of at least 10%.
How to constrain then the IWA? The performance metric is better suited for the purpose. The plot in figure 28 compares the performance of the three different designs. Going from 2.5 up to 3 λ/D IWA has small impact on detection limit, while if we push in the opposite direction then the performance is strongly affected. However, I stress that a role in this loss is also played by planet self-cancellation, which starts being significant below ~200 mas. Despite of this factor, which is not coronagraph-dependent of course, the level of tradeoffs that characterizes this technique makes it probably not the most suitable one for pushing towards extremely small IWAs.

One way to obtain high contrast at small IWA without significantly affecting the throughput is to constrain high-contrast on reduced areas in the focal plane. Figure 29 shows one example of apodizer creating high contrast only on two symmetric 110°-wide regions. 1-D contrast profile is calculated averaging intensity only in these two regions. This design is able to reach 5×10^{-6} at $\sim$ 130 mas with a throughput (25%) much higher than its analog

Table 7: Specifications of three SP designed to investigate the small IWA regime. The contrast by design is 2×10^{-5} for all the three.

Specification	SPa	SPb	SPc
IWA [λ/D]	2	2.6	3
OWA [λ/D]	8	8	8
Throughput	12%	20%	29%
Tolerance to pupil shifts	0.25%	0.5%	0.5%

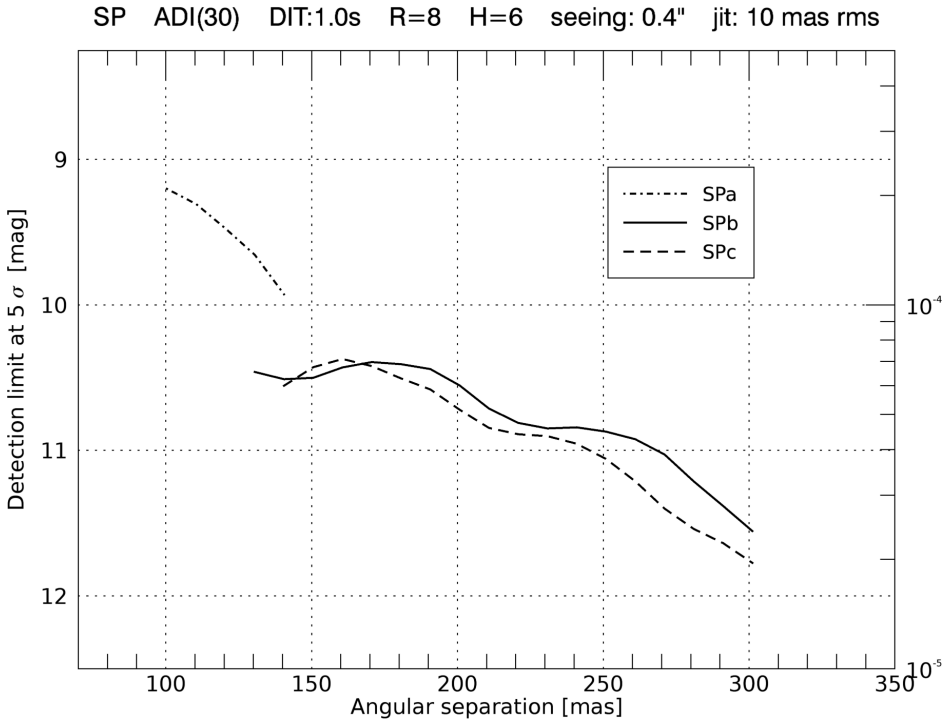
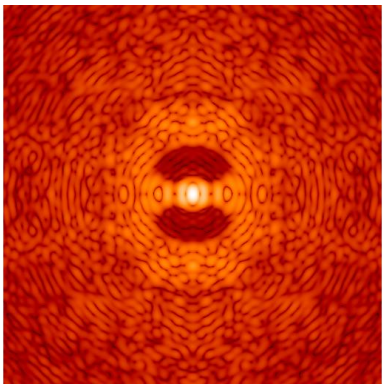


Figure 28: Detection limits obtained with the three shaped pupil designs having small IWA specified in table 7.

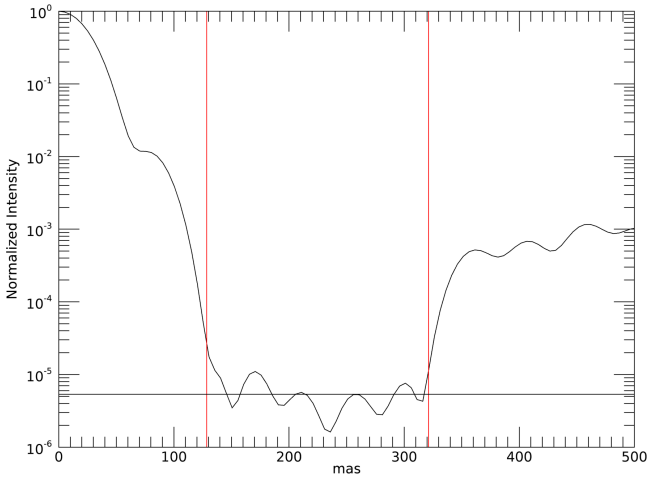
360° design. Of course, such a design is not optimal to search for planets because of the reduced discovery space. This limitation, however, has to be thought in the context of a pupil-stabilized observation: a planet falling behind the focal plane mask at the beginning of the observation may not be completely lost since the FoV rotates with respect to the mask. Viceversa, for the same reason a planet which is visible for some time during the observation may disappear after some time.

A preliminary investigation led us to conclude that detection probability will increase if the focal plane mask could be flipped to the perpendicular orientation at some point during the observation. To make this work, of course, the region of high contrast shall be re-oriented accordingly: to do it, the apodizer has to be rotated as well. This is not feasible, however. In fact, the apodizer must be aligned (in rotation) to the telescope pupil, which can not be rotated. There is only one possible solution to this problem: designing

(a) Apodizer.



(b) Fourier Transform of re-shaped pupil.



(c) Radial intensity profile in the focal plane.

Figure 29: Example of Shaped Pupil design having a non circular-symmetric discovery space. The two high-contrast regions are 110°-wide and range from 128 to 320 mas in H-band (3.2 to 8 λ/D). Contrast is 5×10^{-6}.

two different apodizers that, aligned to the same pupil orientation, generate 'complementary' discovery spaces. After a dedicated study, our collaborators in Grenoble managed to optimize their algorithm so that for a given non-360° apodizer its 'twin' can be generated to satisfy our requirement. The only drawback of this procedure is that a perfect twin can not actually be generated. This problem can be overcome if we accept to have slightly different IWAs for the two apodizers.

To summarize, it is hard to outline a SP design with simple theoretical arguments starting from scientific requirements. The contrast by-design of the apodizer tells us about its potential capability to pin down the speckle noise, but throughput is crucial as well in contributing to the final SNR. As shown in this section, the feedback coming from complete simulations was thus crucial to progressively converge towards a final design. In addition, this process was also inevitably influenced by the comparison of SP performance with other coronagraphic techniques: I stress again that SHARK-NIR

is conceived to offer several options. In this sense, it is important to differentiate the solutions as much as possible so to have an optimal choice for each situation.
In the end, the viable options we identified are three: specifications are reported in table 8. SP1 is designed for moderate contrast, but the smallest possible IWA, SP2 for high contrast slightly further away from the star and SP3 for extreme contrast at a large IWA. SP3 also has a more extended discovery space, with an OWA of 12 λ/D.
Nominal contrasts for the three designs are plotted in figure 30. It is clearly visible how these contrasts are higher than the ones delivered by the apodizers alone. This is due to the add of focal plane mask and Lyot stop. The Lyot stop has been optimized using contrast maps and resulted in an optimal configuration of $98\% - 12\%$, regardless of the design.

Table 8: Specifications of three possible Shaped Pupil designs for SHARK-NIR.

Specification	**SP1**	**SP2**	**SP3**
IWA [λ/D]	2.6	3.5	5.3
OWA [λ/D]	8	8	12
Contrast	2×10^{-5}	1×10^{-7}	2×10^{-7}
Throughput	22%	26%	38%
Discovery space	360°	220°	360°
Tolerance to pupil shifts	0.5%	0.5%	0.5%

SP2 is a non circular-symmetric design. As already discussed, two different apodizers are thus needed to look all around the star. These masks constrain the contrast in two 110°-wide regions each that are oriented perpendicularly (figure 31). The angle is specifically chosen so that high-contrast regions partially overlap and that at the IWA it is possible to fit a 1 λ/D wide object (figure 32). Hence, if a planet falls near the lateral edge of the occulter, it can be spotted with both masks. As already pointed out, the two apodizers have slightly different IWAs. The value reported in table 8 is the biggest of the two. The samller is 3.3 λ/D.

In trying to choose the optimal design looking at performance, we cannot ignore that the apodizer plane is located 50 mm away from the nominal pupil plane. This may induce a contrast degradation. In fact, Fresnel effects alter phase and amplitude of the wavefront seen by the apodizer, undermining the distructive interference (figure 33). The impact on contrast depends on the characteristics of the design. Apodizers designed for high contrast are more penalized, but also the IWA plays a role. SP3, despite of its high contrast, suffers from a marginal degradation since its IWA is big. SP1 has a small IWA but moderate contrast by design, hence the effect is smaller also in this case. The design most affected is indeed SP2. Of course, the 50 mm displacement is accounted for in all simulations.

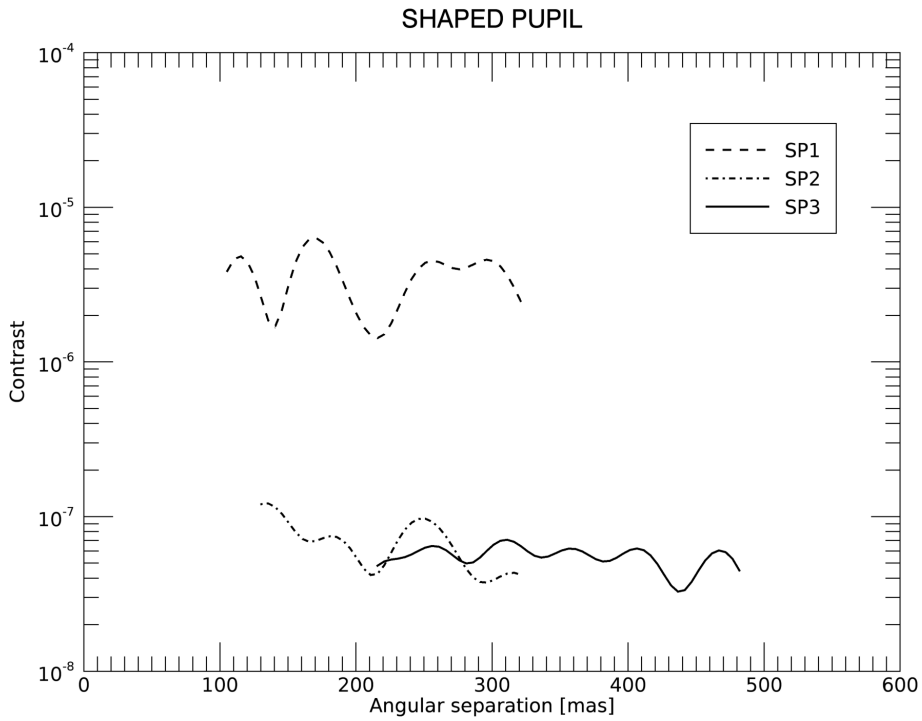


Figure 30: Nominal contrast obtained with the three shaped pupil design specified in table 8. The curves range from IWA to OWA of the respective design.

After a study specifically dedicated to this problem, our collaborators at IPAG managed to optimize the algorithm in order to mitigate this effect. Basically, the step forward consisted in applying the optimization algorithm directly on a Fresnel-perturbated field, instead of starting from a flat and uniformly illuminated field. Of course, propagation is a wavelength-dependent phenomenon: the new designs are then expected to show enhanced chromatism. Dedicated polychromatic simulations, however, showed that chromatism does not impact significantly the performance. As already pointed out, only SP2 has been optimized using the new algorithm, being negligible the impact of the 50 mm on the other two.

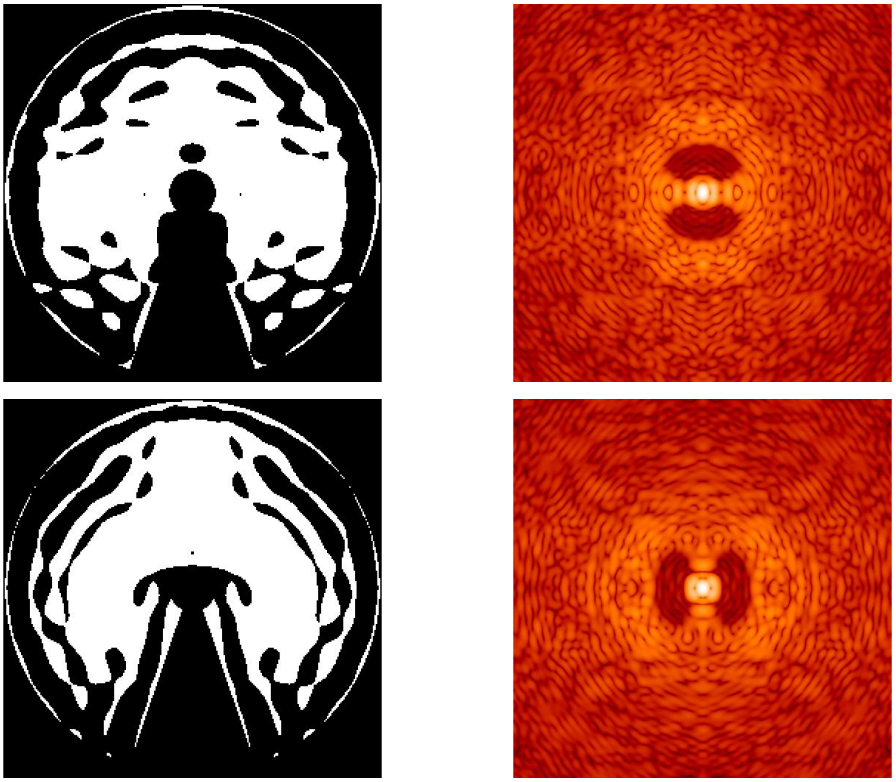


Figure 31: The design of SP2. Two apodizers are needed to cover the all 360° around the star.

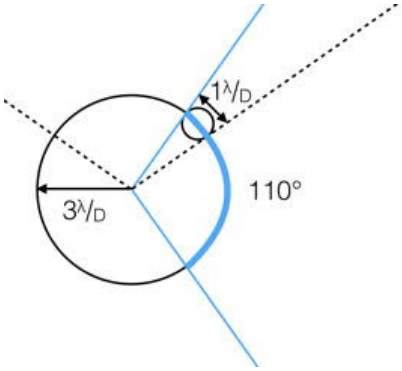


Figure 32: The high-contrast regions generated by the two apodizers SP2a and SP2b are complementary, except for a total of four triangular overlap regions. One of them is shown here as the plane angle in which the planet is inscribed. The angle is selected so that a planet falling in one of these regions is always spotted down to an angular separation equal to the IWA of the coronagraph.

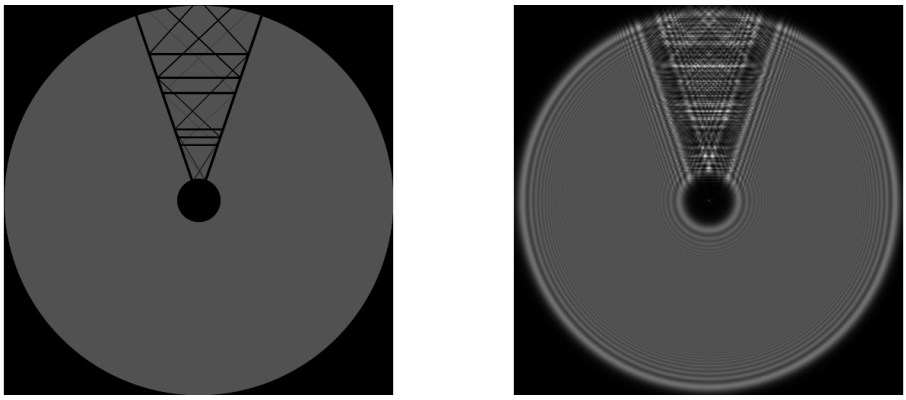


Figure 33: Pupil image in PP1 (*Left*) compared to the same image in the apodizer plane (*Right*), located 50 mm after PP1. Images are displayed in the same linear scale. Fresnel effects clearly re-shape the image, thus degradating the contrast achievable with the apodizer.

4.4 APODIZED PUPIL LYOT CORONAGRAPH

The principle

While Shaped Pupil acts on the pupil transmission function (0 or 1 transmission), APLC (or simply ALC) makes use of pupil apodization in the strict sense of the term, namely transmission varies continuously across the pupil. Moreover, the apodizing function is not conceived as a stand-alone element (as for a SP), but rather it clears out the pupil from starlight when combined with a properly sized focal plane mask (Aime et al. 2002). APLC is the coronagraph adopted in top-level high contrast imagers such as SPHERE (Guerri et al. 2011) and GPI (Savransky et al. 2014). This technique relies on the joint optimization of the two and allows to reach very good contrasts on large discovery spaces (10-20 λ/D).

The design

As for the SP, this coronagraph has been designed at IPAG and then tested in my simulator. In the optimization process, the very critical issue is to deal with diffraction by LBT spiders. Several attemps has been made to optimize the algorithm, none of them providing a clear advantage when run in the test bench. At the end, we obtained very interesting results using a double pupil apodization (figure 34). Design specifications are listed in table 9. The coronagraphic PSF is displayed in figure 35, while contrast is plotted in figure 36. Unlike SP, bandwidth here plays an important role: the optimal choice is to match the one of the scientific filter the coronagraph will be optimized for. The design presented here has 15% bandwith to match the one of SHARK-NIR J-band filter. However, H-band filter is only slightly narrower (12.5%), hence the coronagraph can be used in both bands. However, still the occulting mask has different size in the two bands in terms of λ/D, hence

Table 9: Specifications of the Apodized Pupil Lyot Coronagraph designed for SHARK-NIR.

Specification	Value
IWA [λ/D]	3
OWA [λ/D]	17
Contrast	1×10^{-6}
Throughput	20%
Discovery space	360°

one mask for each band would be required in order not to let chromatic effects to reduce the contrast.

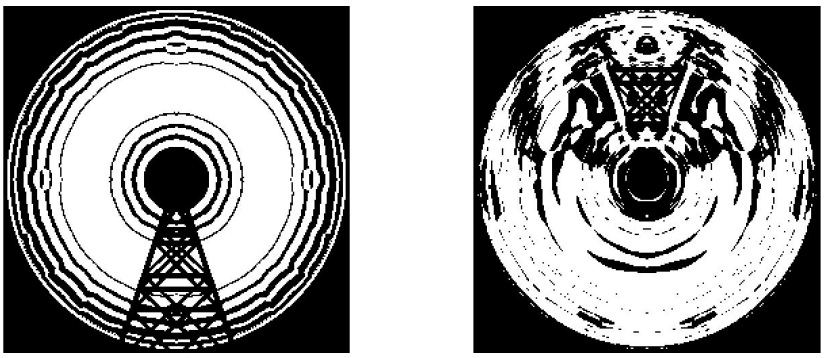


Figure 34: The APLC designed by our collaborators at IPAG. *Left*: apodizer in the first pupil plane. *Right*: apodizer in the second pupil plane. A 3 λ/D occulter is placed in the focal plane in-between.

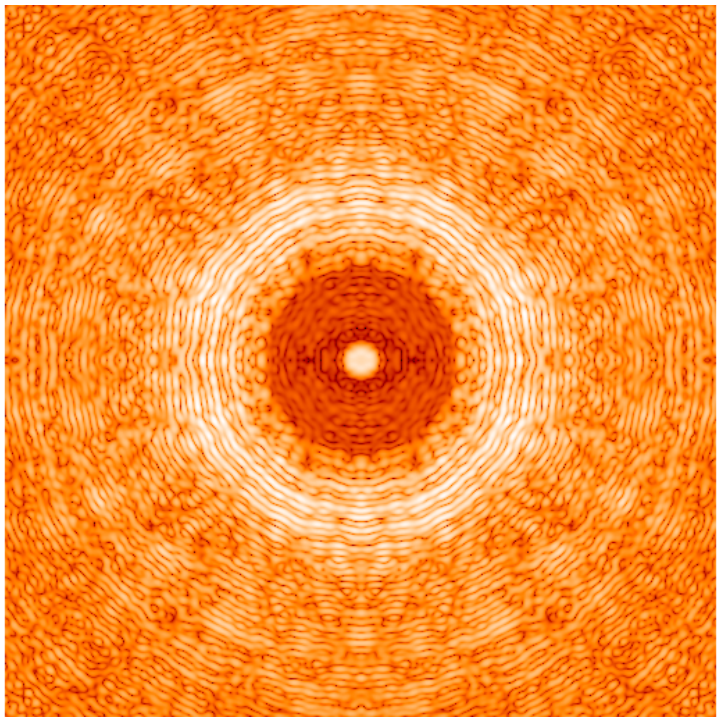


Figure 35: Coronagraphic PSF with the APLC designed for SHARK-NIR.

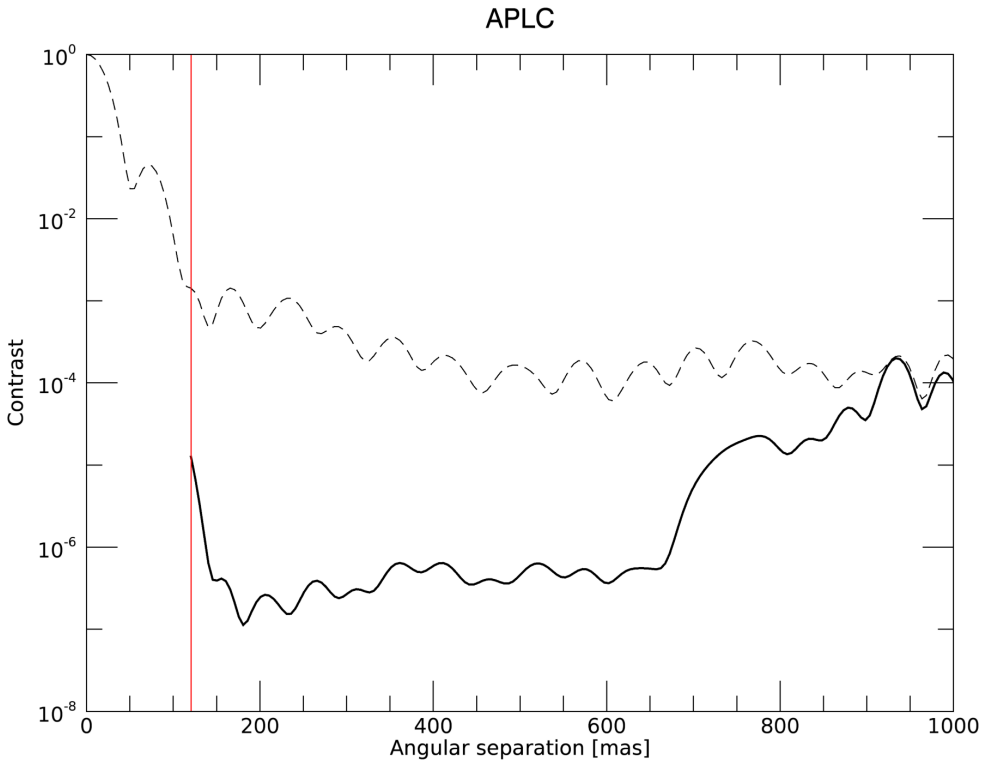


Figure 36: Contrast obtained with the APLC (continuous line). Dashed line is the off-axis PSF, while the red vertical line marks the position of the IWA (120 mas in H-band).

This design allows to achieve excellent performance in ideal conditions. However, it suffer from three main problems:

1. **Out-of-pupil position**
   This is the same problem as for SP, namely the apodizer (the first of the two in this case) is 50 mm after the nominal pupil position. Figure 37 shows the result of a simulation in which the in-pupil and out-of-pupil configurations are compared. The loss in performance is huge, of order of 1 mag from the IWA up to 400 mas.

2. **Pupil chromatic shear**
   Since the ADC is not exactly in the pupil plane, it generates lateral chromatic aberration in the downstream pupil plane (PP2), where the second apodizer is placed. This effect is not problematic for SP, since the Lyot stop is a rather 'passive' component (the contrast is created by the apodizer, the Lyot stop only slightly improve it). For APLC, on the contrary, the second apodizer is an active component. In addition, its fine structures make it much more sensitive to misalignments with respect to a classical stop. In fact, we discovered that pupil shear has a notable impact on contrast, expecially close to the IWA.

3. **Alignment**
   Of course, every optical element needs to be aligned. The second apodizer has quite a complex morphology, that makes it difficult to determine with precision its geometrical center.

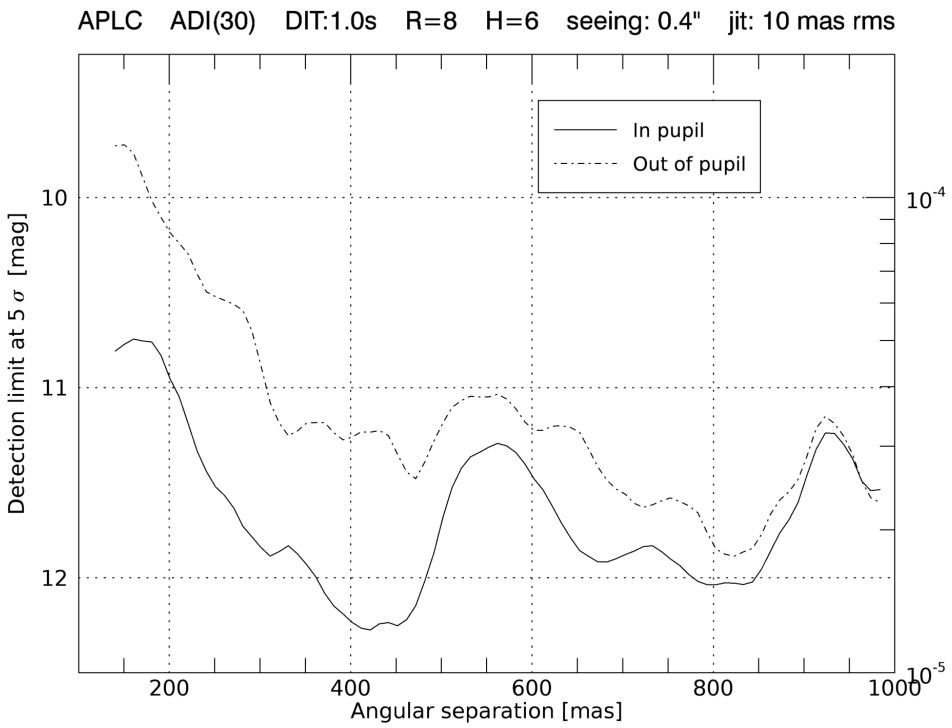


Figure 37: Impact of apodizer displacement for the APLC. The loss in detection limit is of the order of one magnitude in the AO-corrected region.

4.5 APODIZED PHASE PLATE CORONAGRAPH

The principle

An Apodized Phase Plate (APP) is a transmissive optic that modifies the phase of the incident wavefront at a pupil plane. This mask modifies the PSF of all objects in the field by clearing out the light over a 180° 'D'-shaped region on one side of the PSF core (Kenworthy et al. 2007). This concept presents several advantages: a single optics is required, the IWA is typically small and throughput is high and constant with angular separation. Furthermore, the coronagraph is insensitive to tip-tilt aberrations. On the other hand, the discovery space is restricted to only one side of the PSF and the bandwith is typically small. Recently, Snik et al. 2012 showed that device vectorization is a very promising way to solve both problems at the same time.

The design

The mask design has been carried out at IPAG. Table 10 summarizes the specifications. The phase apodizer and the corresponding PSF are shown in figure 38. The apodizer takes into account LBT pupil geometry. Contrast is plotted in figure 39. Since there is no focal plane mask, the PSF stays the same over the field of view. In principle, also in this case a focal plane mask could be added, in combination with a Lyot stop. However, this option has

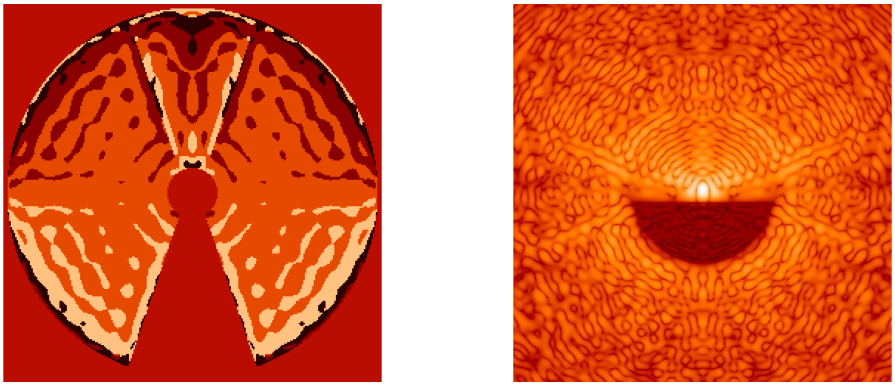


Figure 38: The APP coronagraph designed for SHARK-NIR. *Left*: the phase apodizer. *Right*: the PSF with the apodizer, featuring the characteristic D-shaped discovery space.

not been explored so far for the APP. Assuming to use the phase apodizer alone, then it could be in principle positioned in PP2 (the second intermediate pupil plane), eliminating thus the problem of the 50 mm displacement that affects the first pupil plane.

Table 10: Specifications of the Apodized Phase Plate Coronagraph designed for SHARK-NIR.

Specification	Value
IWA [λ/D]	2.1
OWA [λ/D]	12
Contrast	3×10^{-7}
Throughput	30%
Discovery space	$\sim 180°$

4.6 FOUR-QUADRANT PHASE MASK

The principle

This coronagraph suppresses on-axis starlight by means of a phase mask in the focal plane. The mask divides the FP into four quadrants and induces a π phase shift on two of them on one diagonal. Provided that the image of the star is perfectly centered on the common vertex of the quadrants, then the four outcoming beams combine destructively at infinity and the stellar light in the downstream pupil plane is totally rejected outside of the pupil area (Rouan et al. 2000). This light is then easily blocked by means of a Lyot stop.

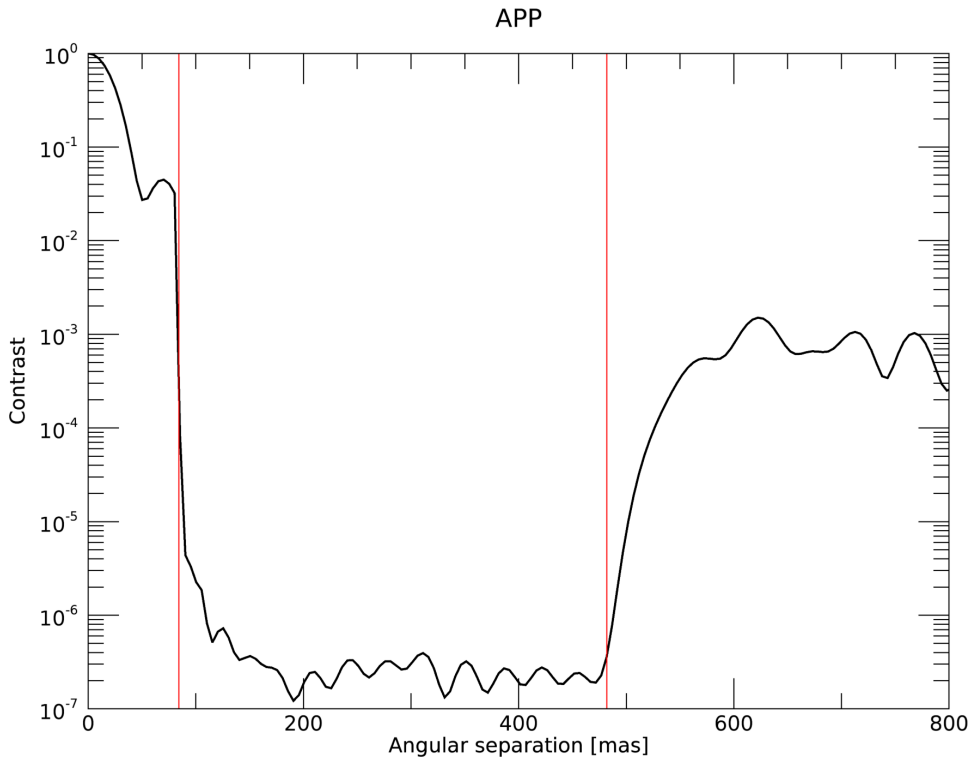


Figure 39: Nominal contrast obtained with the Apodized Phase Plate designed for SHARK-NIR. Red vertical lines mark the position of IWA and OWA in H-band, respectively 84 and 480 mas. Mean contrast in this region is $\sim 3 \times 10^{-7}$.

This concept has been tested with near-infrared light both in the lab (Riaud et al. 2003) and on sky (Boccaletti et al. 2004). The two major drawbacks of this technique are its high sensitivity to aperture geometry and tip-tilt. The former stems from the fact that the phase mask diffracts light to the outer regions of pupil intensity discontinuities, as central obscurations or support structures. Thus, light is not completely rejected outside the pupil image in presence of these elements, resulting in a contrast degradation proportional to the obscured area. Tip-tilt aberrations cause instead starlight to unfairly breakdown into the four beams, undermining their destructive interference. Despite of this limitations, simulations showed that FQPM represents an intriguing solution for SHARK-NIR, ensuring a way to further increase the detection capability of the instrument at very small angular separations in the high-strehl regime.

The design

The design foresees the focal plane mask and a downstream Lyot stop. The IWA is fixed and is of order of 1 λ/D. The only component to be optimized is thus the Lyot stop. For the purpose, I used the same approach as for the Classical and Gaussian Lyot coronagraphs. Table 11 reports the optimal parameters in different cases. Because of the high sensitivity of this

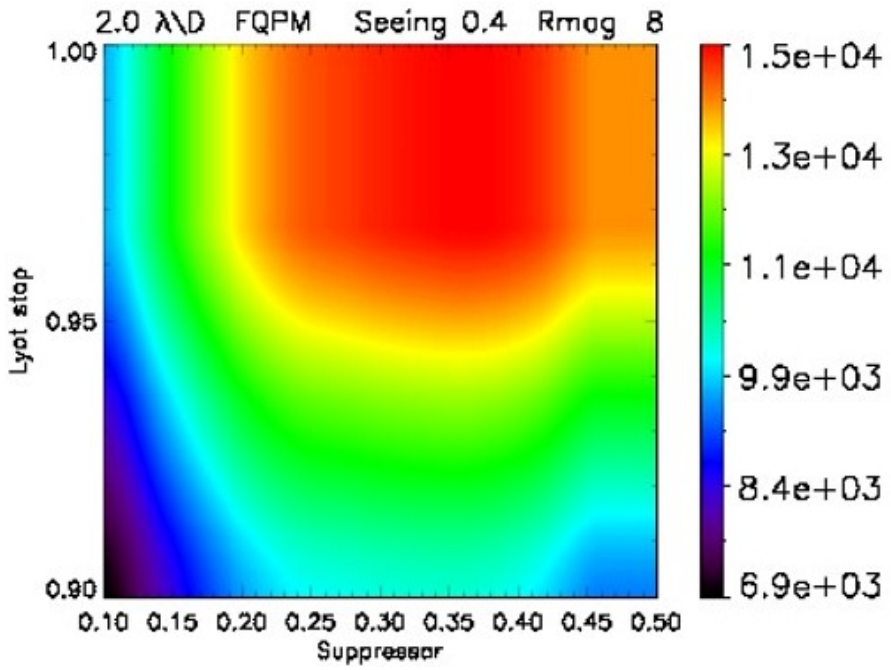


Figure 40: Contrast maps for the optimization of Lyot stop and central suppressor for the FQPM coronagraph in the case of 95% SR.

coronagraph to tip-tilt, I also explored the impact of residual jitter on the optimization. Simulations show that for SR approaching 100% the maximum contrast is achieved with a significant central obscuration (~ 40%). However, there is quite large dispersion around this value (figure 40). There is no difference in this case between 84% and 75% SR. In conclusion, since 95% SR (without residual jitter) is somewhat extreme, we opted for a more conservative 20% suppressor.

Contrast in this configuration is shown in figure 41. Diffraction by LBT spiders severly limit the nulling capability of this coronagraph. However, this coronograph has the advantage of attenuating the starlight very close to the star without zeroing or strongly reducing the throughput in the area.

Table 11: Optimal values for Lyot stop and central suppressor of the designed four-quadrant phase-mask coronagraph. The optimization is performed in different conditions in order to explore the impact of SR on optimization.

R	seeing	jitter [mas rms]	SR	Lyot stop	Suppressor
8	0.4″	0	95%	98%	36%
8	0.4″	10	84%	98%	17%
8	0.8″	10	75%	98%	16%

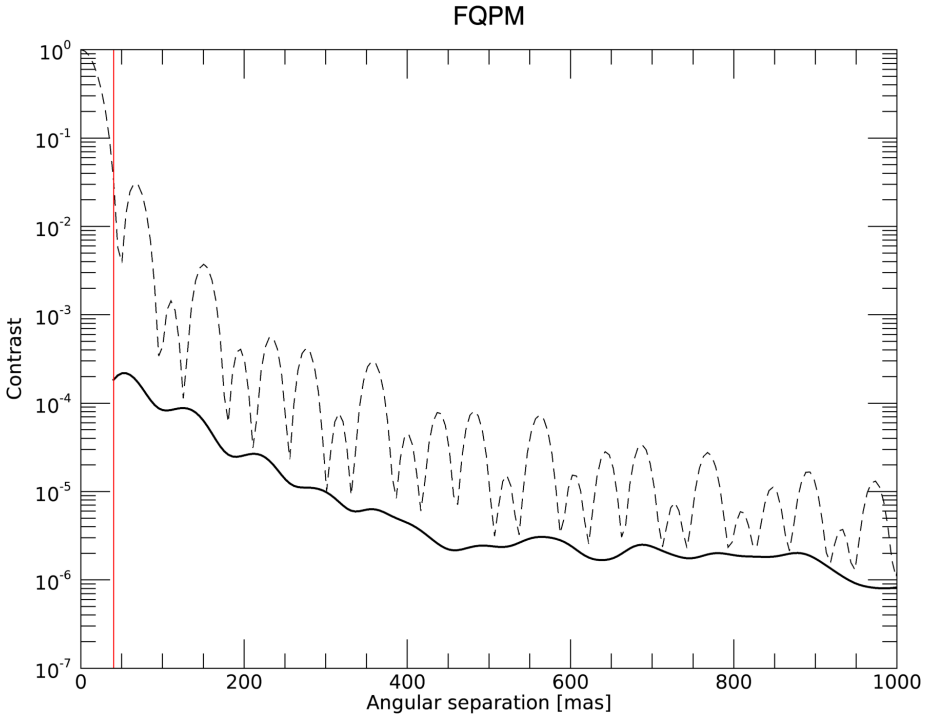


Figure 41: Contrast obtained with the FQPM (continuous line). Dashed line is the off-axis PSF, while the red vertical line marks the position of the IWA, which is around 40 mas in H-band.

4.7 VORTEX CORONAGRAPH

The principle

Like the FQPM, the vortex acts on the phase of the wavefront in the focal plane. In this case, an optical element is used whose complex transmittance depends in a linear fashion on the azimuth angle. The so called 'phase ramp' completely cancels a diffraction-limited image by diffracting all the light outside of the geomerical area of the pupil, where it can be easily blocked with a Lyot stop (Foo et al. 2005, Mawet et al. 2005). This optical system has been proven, both on laboratory and on-sky, to achieve very high stellar suppression (Serabyn et al. 2017). The limitations of this technique are the same as for the FQPM, namely the high sensitivity to pupil geometry and tip-tilt aberrations.

The final design

The topological charge is set equal to 2, in order to obtain the same IWA as the FQPM. The Lyot stop has been optimized in the usual way. Even in this case, the sensitivity to residual jitter is investigated. The results in this case are exactly identical to FQPM in all conditions. The Lyot stop is thus 98% of the pupil size, with a central obscuration of 20%.

4.8 THE FINAL SELECTION

The plot in figure 42 shows detection limits in high Strehl regime with all the coronagraphs presented in this chapter. The target is bright ($R = 8$) and seeing excellent ($0.4''$). Figure 43 reports the same comparison with a lower Strehl: $R = 10$ and seeing is $0.6''$. Jitter is 10 mas rms in both cases. Quantitatively, SR are, respectively, 85% and 70%. The tested coronagraphs span quite a large range in detection limit at all angular separations. The performance of classical and Gaussian Lyot are very similar. The latter, despite of some bumps in correspondence of diffraction rings, is preferable because it allows to go closer to the star, exploiting the fact that even inside the IWA transmission is higher than zero. I also stress that the Gaussian Lyot design is not optimized for high Strehl observation (see section 4.2). Even FQPM and Vortex are very similar and, despite of vibrations, they deliver the best performance either at 100 mas and at separations bigger than ~300 mas in high-Strehl regime. With moderate Strehl, they outperform all other coronagraphs at every separation, with Vortex slightly better than FQPM. The only coronagraph delivering a performance better than Vortex and FQPM in limited ranges in high-Strehl is the Shaped Pupil. SP1 is intriguing since it is the best one at 130 mas, with half a magnitude gain with respect to phase masks. SP2 is the best coronagraph between 160 and 200 mas from the star, while SP3 from 250 to 300 mas. Finally, APLC and APP are definitely not optimal at any angular separation.

In conclusion, looking at detection curves we selected Gaussian Lyot, FQPM, SP1 and SP2. The reason why FQPM has been preferred to Vortex is for its smaller sensitivity to misalignments, as it will be shown in chapter 7. Table 12 summarizes the performance indicators for the selected coronagraphs.

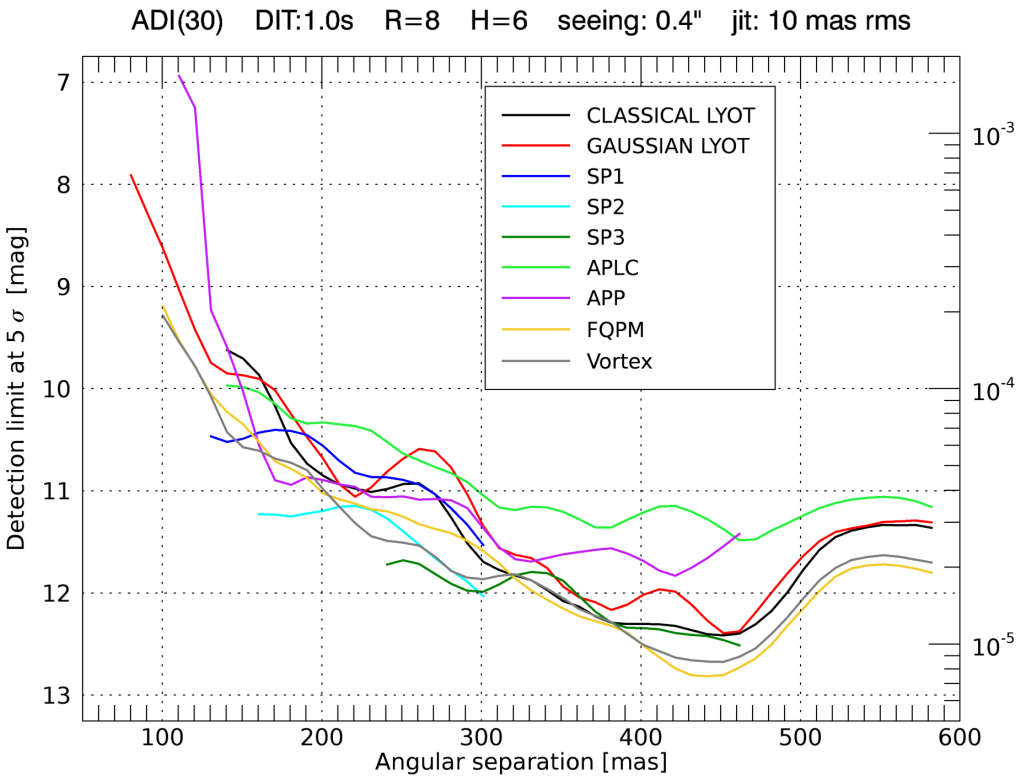


Figure 42: Detection limit for all the coronagraphs described in this chapter in case of high Strehl.

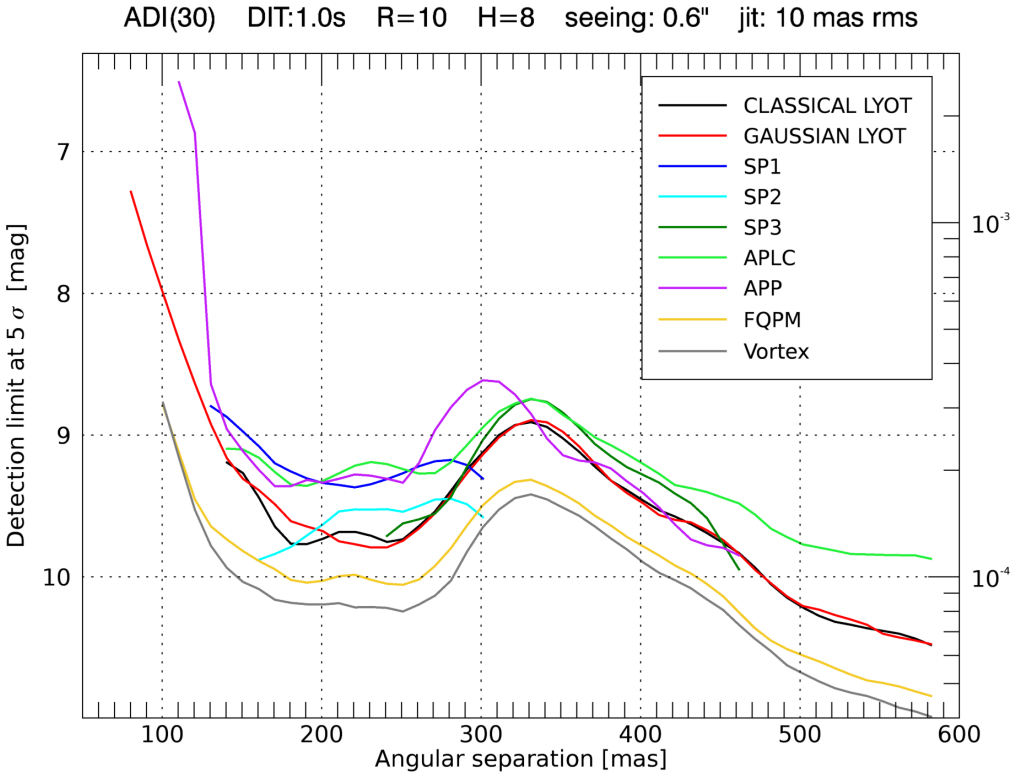


Figure 43: Detection limit with for the coronagraphs described in this chapter in case of moderate Strehl.

Table 12: Properties of the four coronagraphic designs selected for implementation in SHARK-NIR.
*Maximum throughput. Throughput decreases with increasing proximity to the star (at IWA, it is 28%).

Coronagraph	**IWA**	**OWA**		**Contrast**	
	[mas]	**[mas]**	@100 mas	@200 mas	@400 mas
Gaussian Lyot	120	–	5×10^{-4}	1×10^{-5}	1×10^{-5}
SP1	100	320	2×10^{-5}	2×10^{-5}	–
SP2	100	320	1×10^{-7}	1×10^{-7}	–
FQPM	40	–	1×10^{-4}	2×10^{-5}	5×10^{-6}

	Throughput	**Discovery space**	**band**	**Science**
Gaussian Lyot	56%*	360°	J,H	Disks/Jets/AGN
SP1	22%	360°	H	Exoplanets
SP2	26%	220°	H	Exoplanets
FQPM	95%	360°	H	Exoplanets

5 INSTRUMENT PERFORMANCE

Detection limit curves in several combinations of guide star magnitude, atmospheric seeing and residual jitters have been generated during the design phase. I will not report in this chapter the results of the whole simulation campaign, but I will only focus on a few cases. Contrast curves will be shown only for those coronagraphs that have been selected for implementation (see section 4.8). I will first report some results using ADI (section 5.1), then a few raw contrast curves (section 5.2). Section 5.3 is finally dedicated to the investigation of the impact of residual jitter on coronagraphic performance.

5.1 ADI

In this section I report instrument performance in few different observing conditions using differential imaging. The processing algorithm is classical ADI. The baseline for these simulations is 30 images sequences and FoV rotation of $\sim 50°$ (see section **??**). I introduced 30 nm of residual NCPA and a residual jitter of 10 mas rms derived from Forerunner data (section 5.3). All simulations are monochromatic at the central wavelength of the H-band (1.6 µm).
Figure 44 and 45 show detection limits in high-Strehl regimes with FLAO and SOUL, respectively. Figure 46 and 47 reports other examples at lower Strehl ratios with FLAO.
A contrast of 2×10^{-5} can be reached at 300 mas in high Strehl conditions. It should be stressed that performance can be improved by optimizing post-processing. Figure 48 and 49 show two examples of the application of the optimized data reduction algorithm described in Carolo and Vassallo in prep. The algorithm basically optimize the reference PSF subtraction as a function of angular separation in order to maximize the SNR. The gain in performance is significant, both in high and low Strehl regime. Using this optimized algorithm, it is possible to reach a 10^{-5} contrast at 200 mas and less than 5×10^{-6} at 400 mas, matching thus the scientific requirements for exoplanets.

5.2 RAW CONTRAST

For those science cases requiring field stabilization, the simple raw contrast metric is used (see chapter 3). Since the instrument will rotate, the

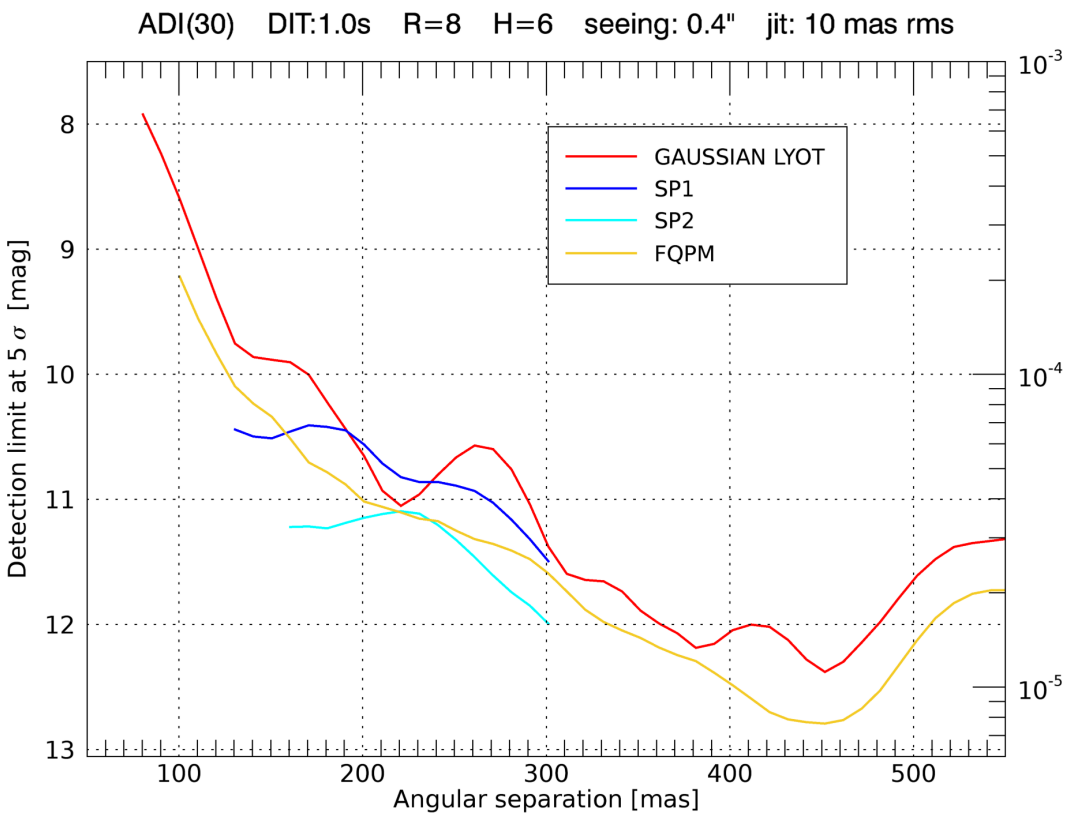


Figure 44: Detection limits for a R = 8 guide star and 0.4″ seeing with FLAO.

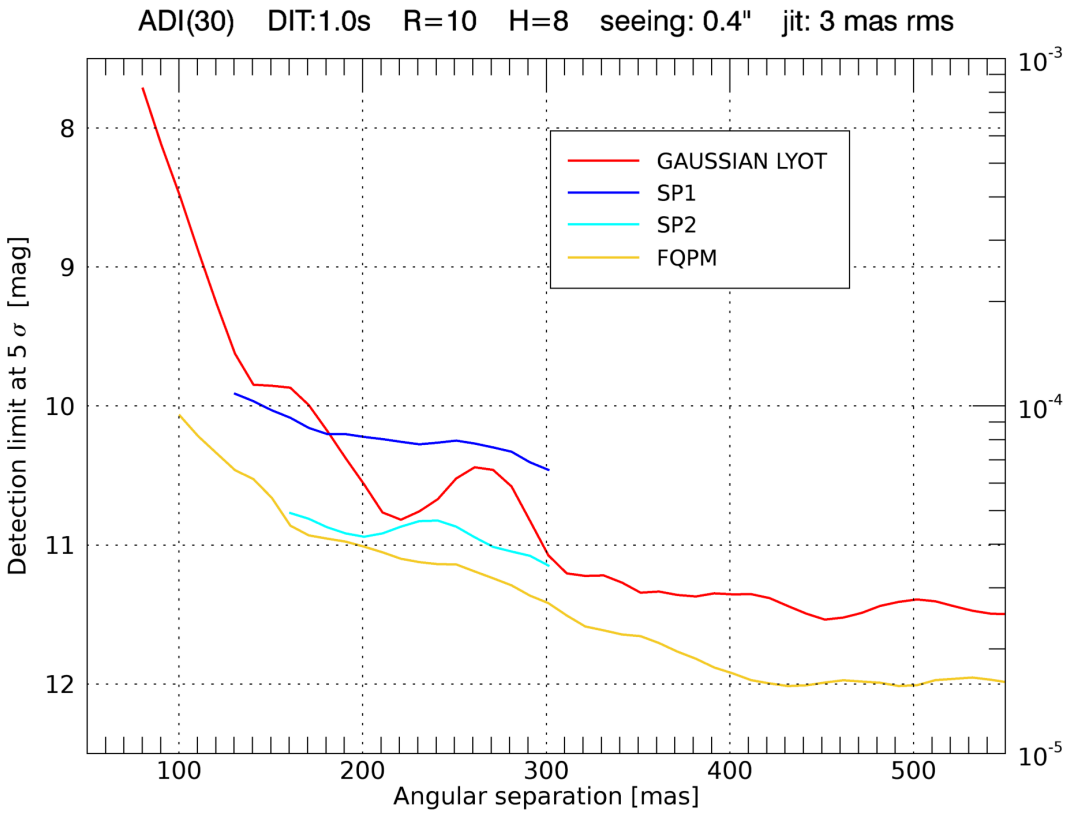


Figure 45: Detection limits for a R = 10 guide star and 0.4″ seeing with SOUL.

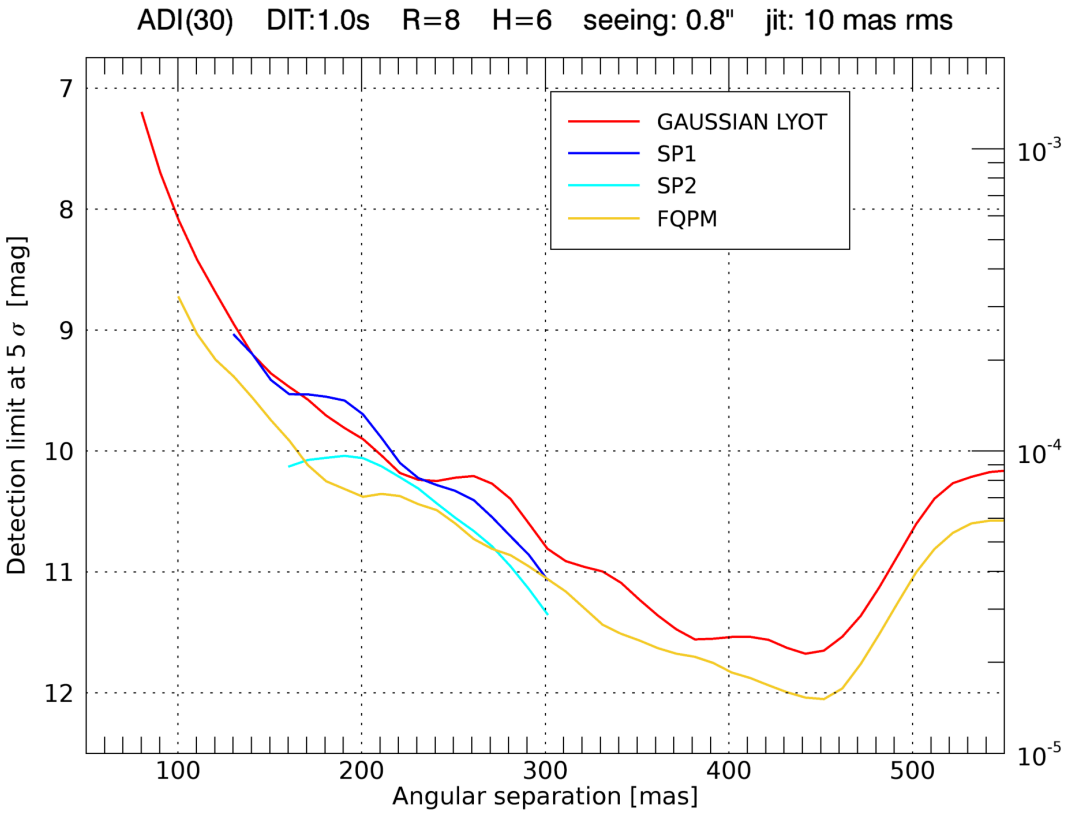


Figure 46: Detection limits for a R = 8 guide star and 0.8″ seeing with FLAO.

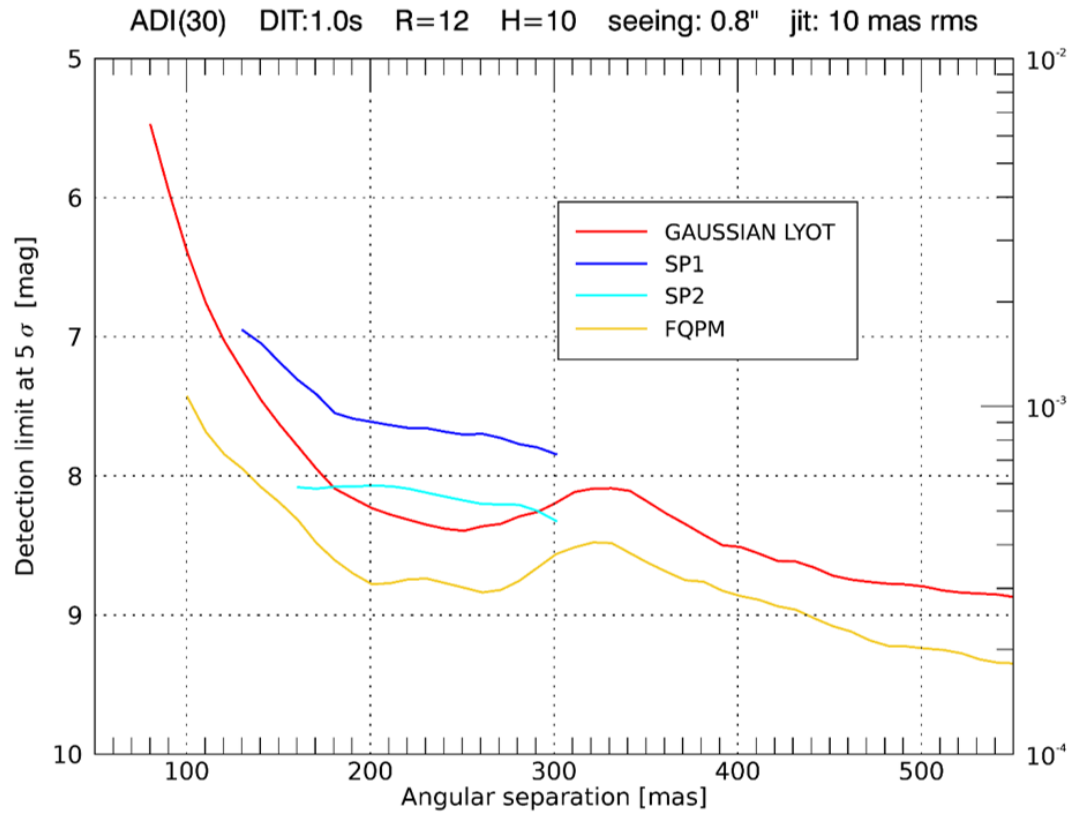


Figure 47: Detection limits for a R = 12 guide star and 0.8″ seeing with FLAO.

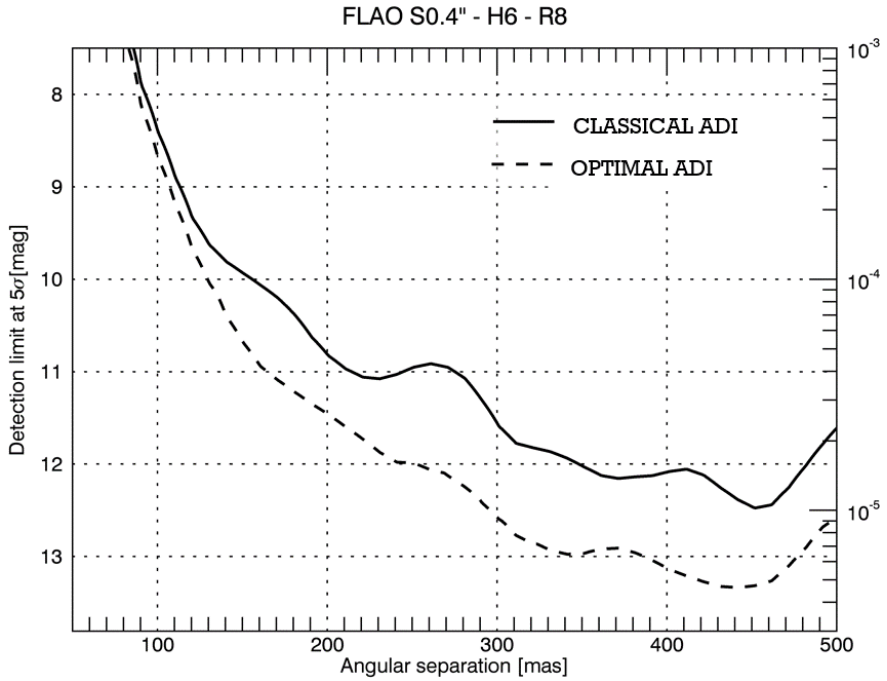


Figure 48: Detection limit in high Strehl regime with the Gaussian Lyot obtained with classical ADI compared to the optimized data processing algorithm described in Carolo and Vassallo in prep.

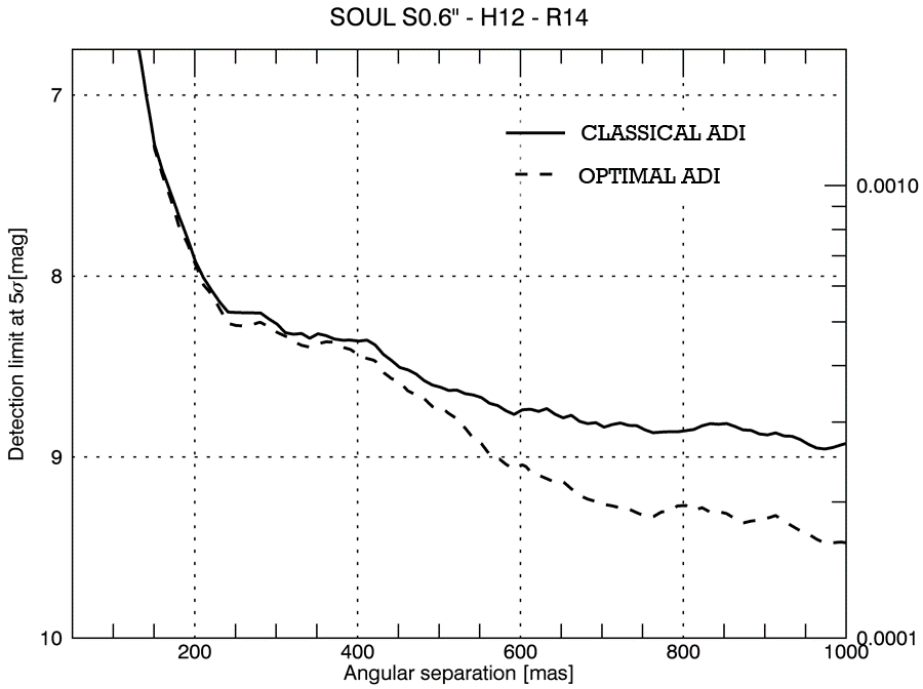


Figure 49: Detection limit in low Strehl regime with the Gaussian Lyot obtained with classical ADI compared to the optimized data processing algorithm described in Carolo and Vassallo in prep.

Gaussian Lyot is the only coronagraph that can work in this mode, being both the focal plane mask and the Lyot stop circularly symmetric. The possibility to also exploit the FQPM will be investigated via simulations.
Figure 50 shows contrasts achieved with both FLAO and SOUL systems for a R = 12 target. The smaller AO cut-off frequency of the FLAO system in the faint regime is clearly visible in the bump at 300 mas. With SOUL it will be possible to achieve a contrast close to 10^{-4} at 300 mas directly in raw coronagraphic images even in the faint tail of the magnitude distribution for disks and jets that SHARK-NIR will target.
Figure 51 shows a wide-field contrast curve without coronagraph for a R = 10 target. The contrast at 2 arcsec is $\sim 2 \cdot 10^{-6}$, thus fulfilling the requirement for AGN in the local universe (table 1 in chapter 1).

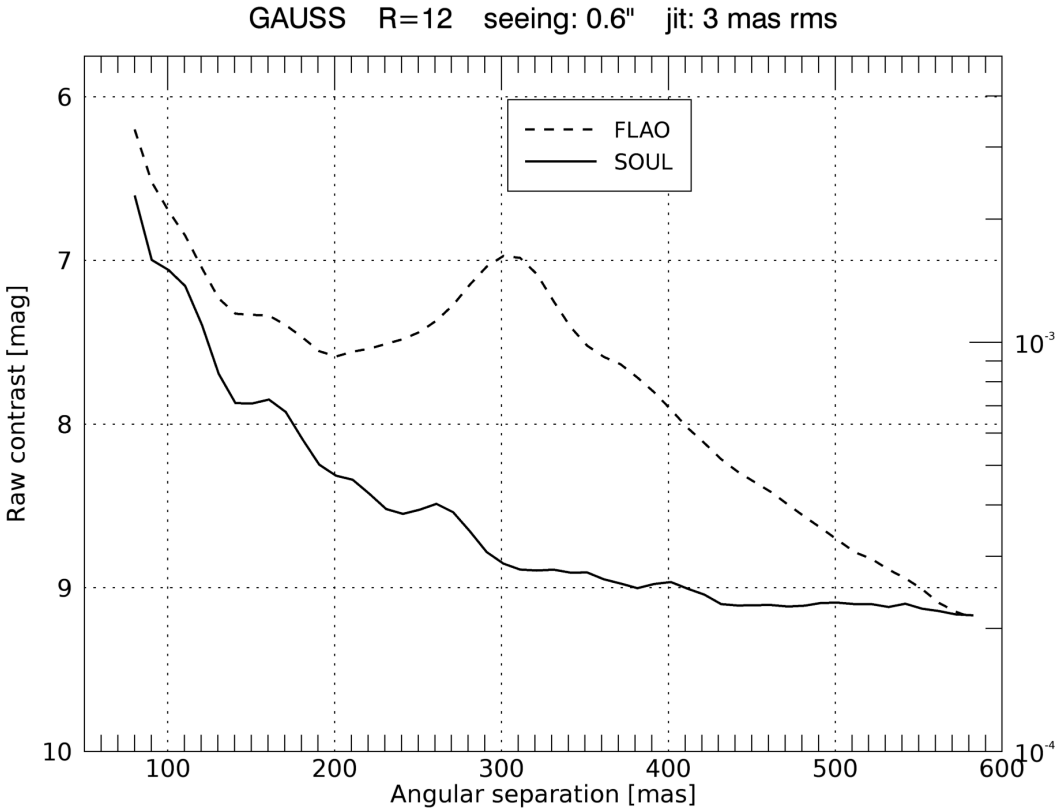


Figure 50: Raw contrast obtained with the Classical Lyot coronagraph for a R = 12 target and 0.6″ seeing with both FLAO and SOUL. Contrast with SOUL is close to 10^{-4} at 300 mas, which fulfills the requirement for disks and jets science.

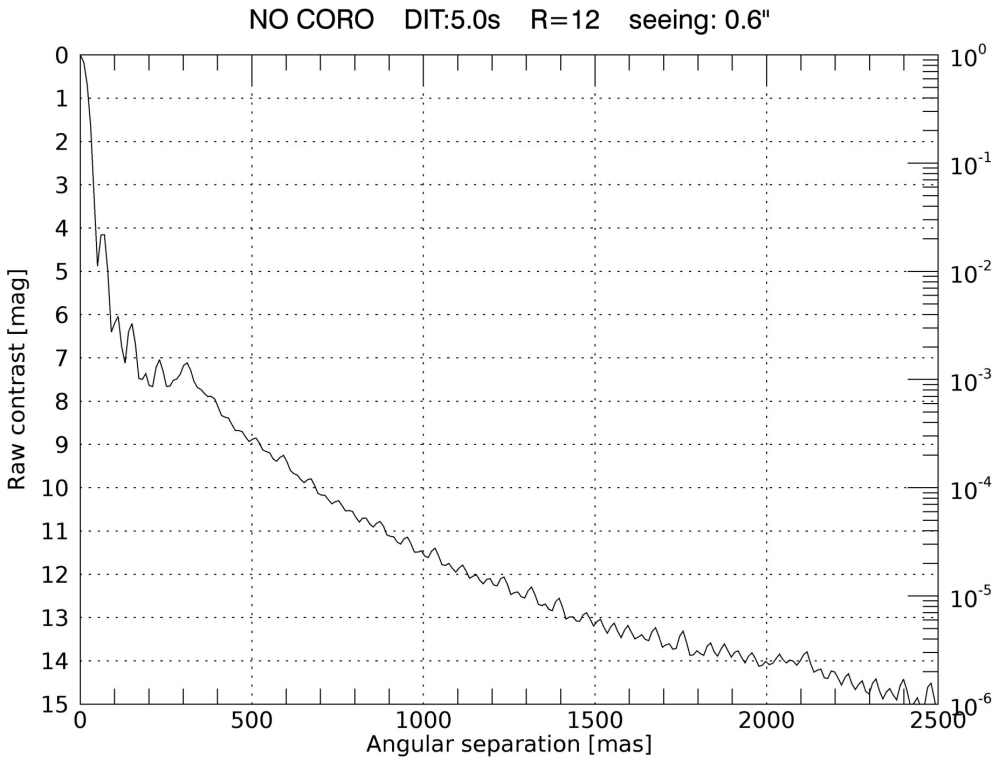


Figure 51: Wide-field raw contrast without coronagraph. R magnitude of the target is 12, seeing is 0.6″. Contrast at 2 arcsec is $\sim 2 \cdot 10^{-6}$, which is in line with the requirement for AGN science in the local universe.

5.3 THE IMPACT OF JITTER ON CORONAGRAPHIC PERFORMANCE

After the FDR, we performed a study specifically dedicated to the issue of telescope vibrations and their impact on coronagraphic performance, which led to a progressive refinement of jitter implementation in the simulator. In this section I report the main results of this assessment study.

Sinusoidal jitter

As detailed in section 5.3, modelling vibrations as a tilt-only aberration at a fixed frequency has the limitation of not introducing significant PtV displacements of the star during an exposure that may cause an important fraction of star light to leak out from the coronagraphic mask. This effect can be accounted for by using the Forerunner time-series. But how much does it affect detectability? Figure 52 and figure 53 compare detection limits using the two approaches for the Gaussian Lyot and the FQPM, respectively. Jitter is 10 mas rms and SR is 85%. For the Gaussian Lyot there is almost no difference between the two approaches, while in the case of the FQPM the implementation based on Forerunner data causes a performance loss below

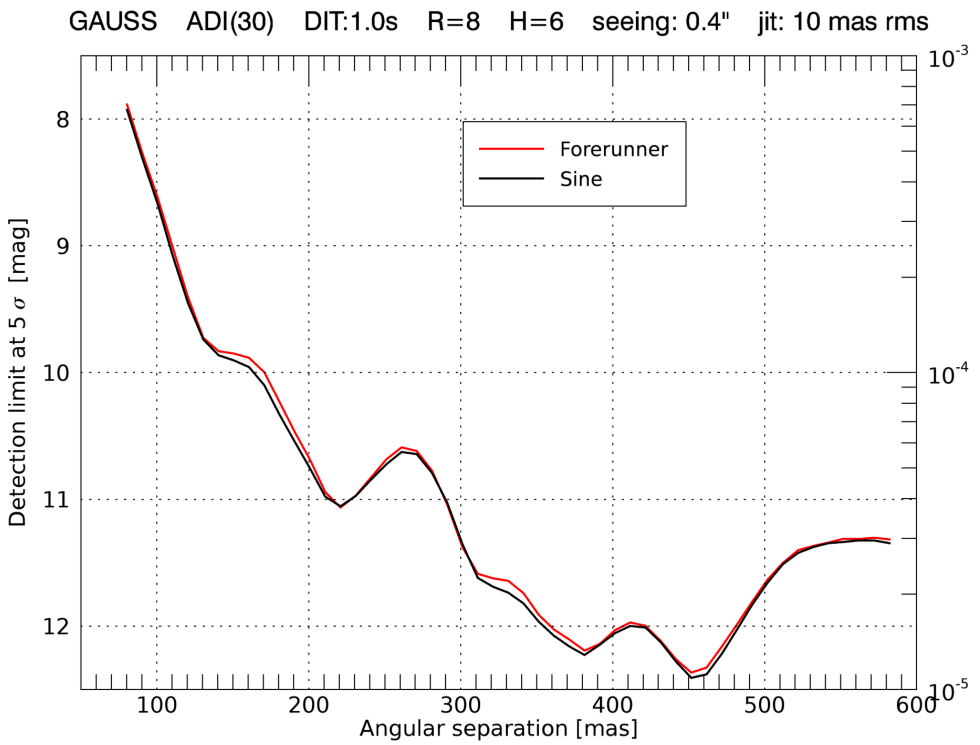


Figure 52: Detection limits with the Gaussian Lyot in presence of 10 mas rms jitter modeled as a single Fourier component (*sine*) or using the Forerunner time series (*Forerunner*).

200 mas. We can thus conclude that PtV displacements during the exposure might cause a performance degradation. For this reason, the Forerunner implementation has been adopted as the baseline in simulations.

Sensitivity to vibrations

I then investigated the sensitivity of coronagraphs to jitter. Again, the analysis is performed in the two limiting cases of a very robust technique (Gaussian Lyot) and a very sensitive one (FQPM). Results are shown in figure 54 and 55, respectively. I compared the case of no vibrations with 3 mas and 10 mas rms. While there is almost no difference between 0 and 3 mas, the FQPM experiences a loss of 1 magnitude at 100 mas in case of 10 mas rms jitter. The Gaussian Lyot, on the other hand, is not at all affected by the jitter increase.

Vibrations on the WFS

Vibrations can be introduced directly at the level of WFS, so that SR degradation not only results from image motion (tip-tilt), but also the impact on the other modes is accounted for. Having these data cubes at hand, it is pos-

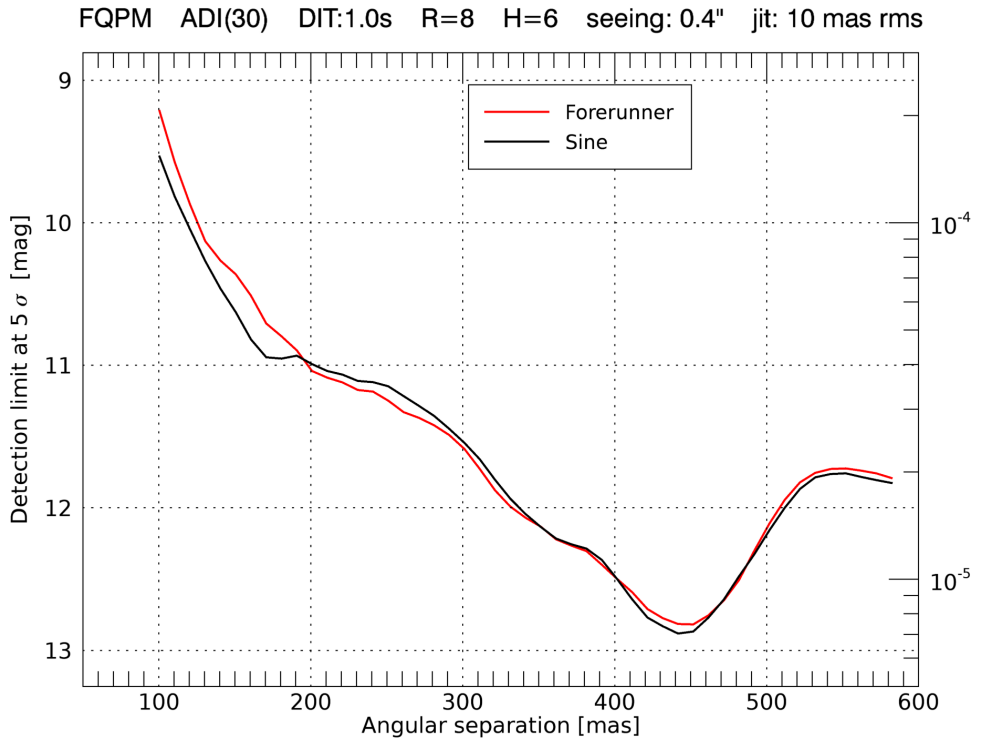


Figure 53: Detection limits with the FQPM in presence of 10 mas rms jitter modeled as a single Fourier component (*sine*) or using the Forerunner time-series (*Forerunner*).

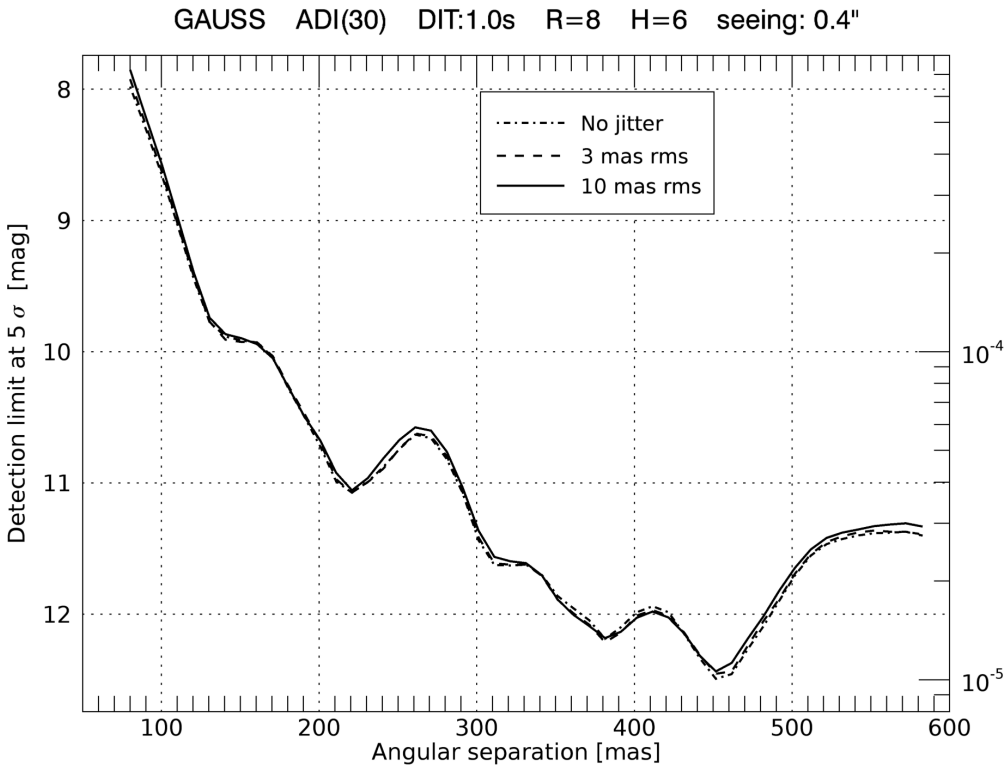


Figure 54: Detection limits with the Gaussian Lyot for three different values of residual jitter.

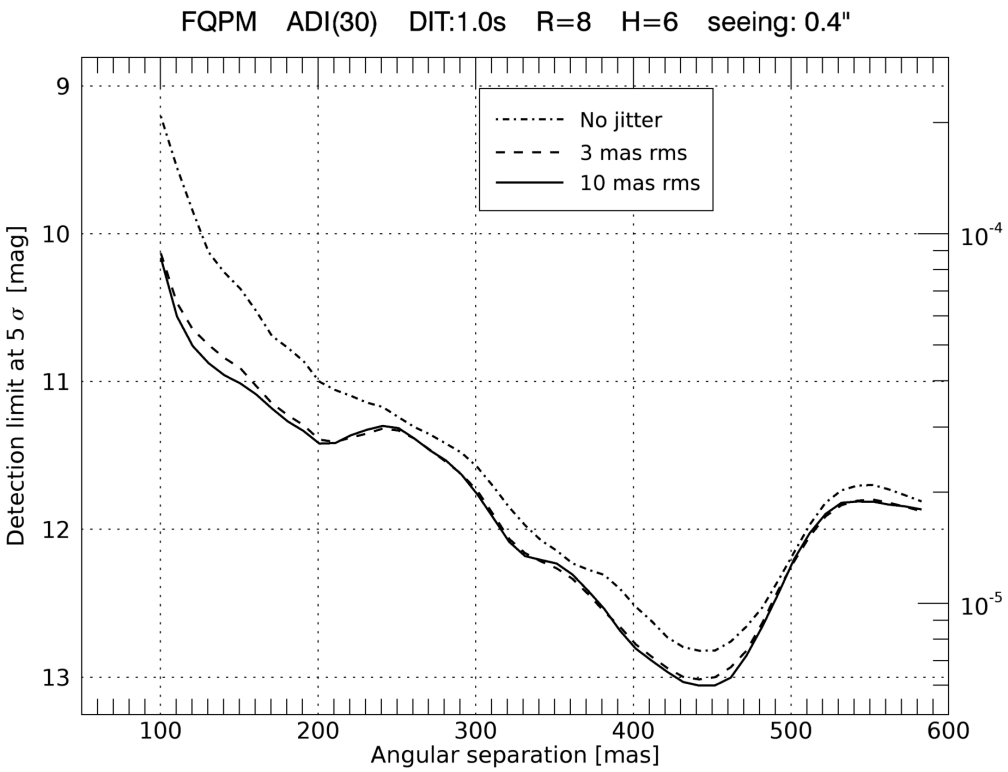


Figure 55: Detection limits with the FQPM for three different values of residual jitter.

sible to simulate the advantage that a fast tip-tilt correction can deliver. Data cubes with vibrations all assume a quite bad open-loop vibration spectrum of ~70 mas rms. When closing the loop, the expected residual of course depends on the parameters of the loop itself, expecially the framerate, which is in turn dependent on the guide star brightness.

I simulated the observation of a faint target (R = 12) with FLAO. Seeing is good (0.6″). As anticipated in section 5.3, the effect of vibrations might be mitigated by increasing the AO loop framerate. This is the situation: in this case it has been demonstrated that the maximum Strehl ratio is achieved going from 700 Hz to 1 kHz, despite of the smaller SNR on the four pupils of the WFS. The information about residual vibrations in this configuration resides inside the data cube itself. It can be extracted for example by simply taking the Fourier transform of a sufficiently large number of screens and retrieving its centroid. The so calculated residual is of 23 mas rms. SR is 40%. What happens then if we apply a tip-tilt correction? To test it, I retrieved the tip-tilt coefficients for all phase screens in the data cube and repeated the simulation applying them (using the opposite sign) in input. Of course, if using exactly the same coefficients then residual tip-tilt is completely removed. In order to make the correction 'imperfect', I corrupted the tip-tilt coefficients adding some random noise. Residual vibrations have been reduced in this way to ~5 mas. Figure 56 and 57 compare detectability curves using the Gaussian Lyot and the FQPM, respectively.

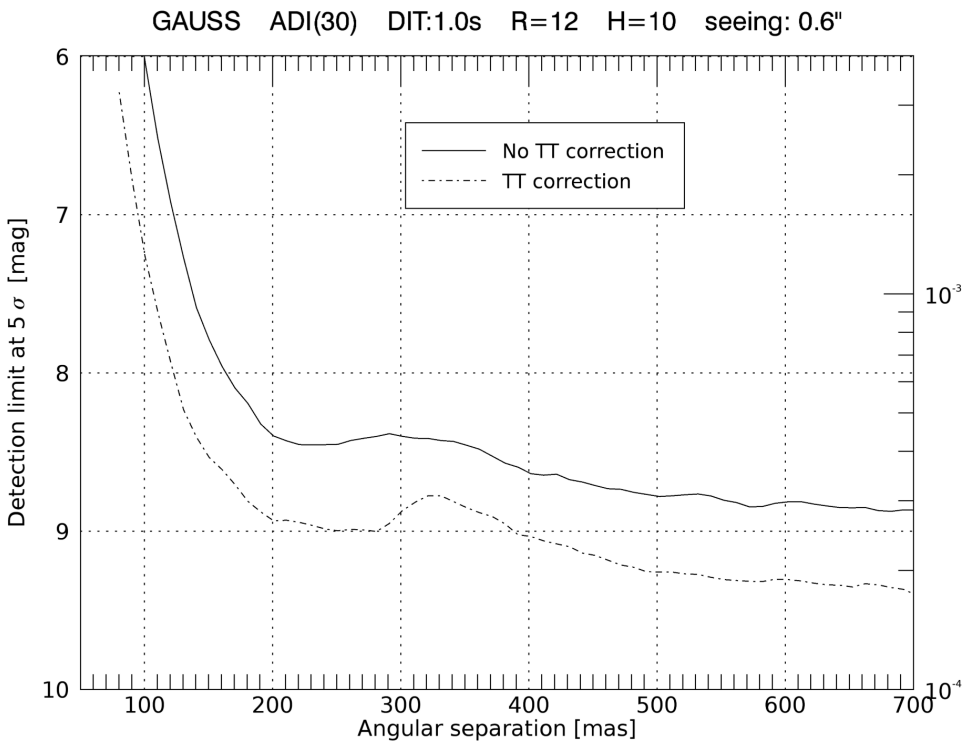


Figure 56: Detection limits with the Gaussian Lyot in presence of 70 mas open-loop vibrations with and without tip-tilt correction.

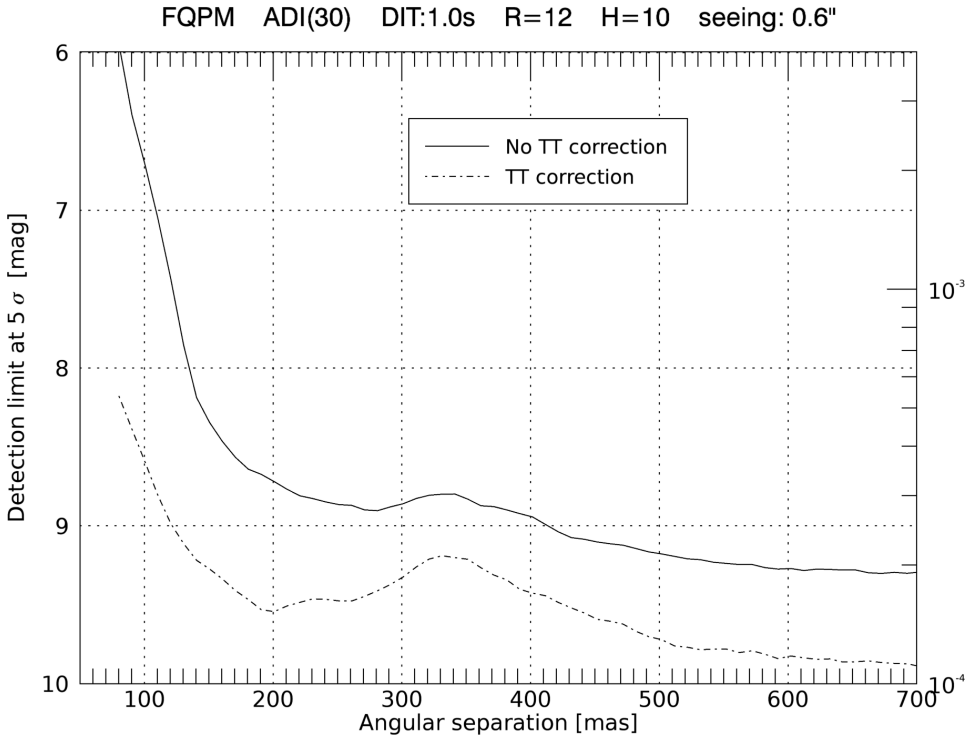


Figure 57: Detection limits with the FQPM in presence of 70 mas open-loop vibrations with and without tip-tilt correction.

SR with tip-tilt correction grew up to 70%. The gain in performance when the correction is applied is huge for both coronagraphs. In the case of the FQPM, it reaches more than two magnitudes at separations smaller than 100 mas.

6 THE EFFECT OF CHROMATISM ON CORONAGRAPHIC PERFORMANCE

Results shown so far are monochromatic. Although one single wavelength is usually sufficient for many purposes, there are still some second-order effects that can only be investigated by using a multi-wavelength approach. In fact, each coronagraph has its peculiar concerns when operated in broadband. In this chapter I will detail all the issues due to chromatism that came up during the design phase. The formation of a polychromatic image has already been discussed (see chapter 2). I will thus start directly by detailing the tests I performed on the coronagraphs case by case.

Figure 58 shows a comparison between a monochromatic and a polychromatic image generated with the simulator. By eye inspection, the latter appears somewhat smoother, with radial structures that are easily visible at the AO control radius. In fact, the combination of multiple wavelengths causes an attenuation of high-frequency noise in the image. For this reason, when comparing mono to polychromatic sequences we should always expect a small inherent gain in performance in the polychromatic case.

Gaussian Lyot

The selected IWA is 3 λ/D. The by-design proportion between the Gaussian filter FWHM and the Airy disk is actually respected only for the central wavelength of the passband. In other words, it is like each wavelength sees a different IWA in its reference frame: it results bigger for wavelengths smaller than λ_0 and viceversa. Figure 59 shows this effect clearly: it reports

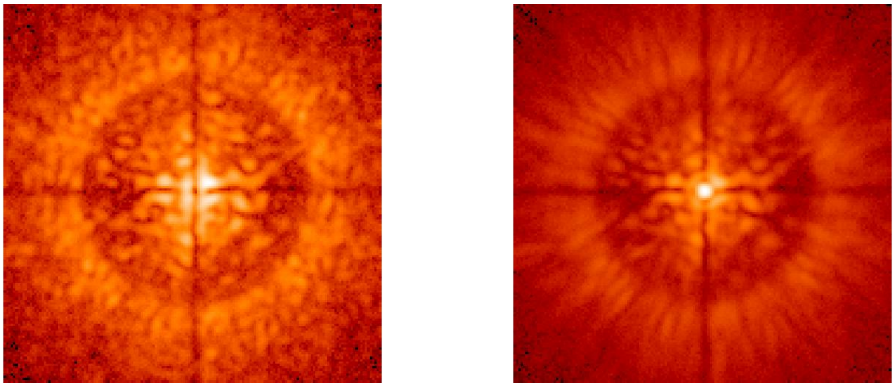


Figure 58: *Left*: monochromatic image obtained with the FQPM. *Right*: same image in broadband. The scale is logarithmic.

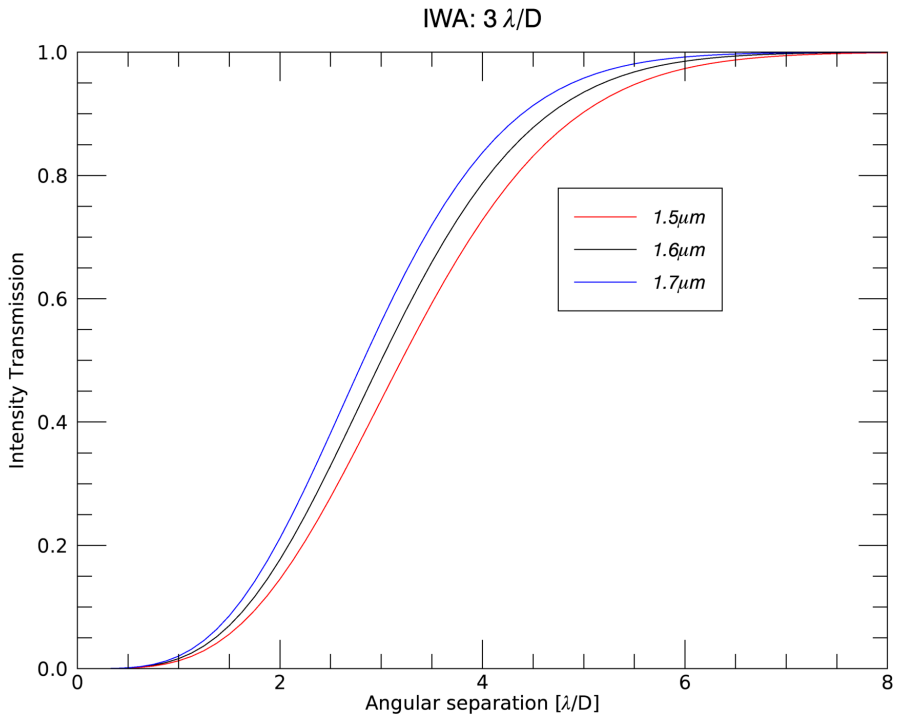


Figure 59: Intensity transmission of the Gaussian filter for the initial, central and final wavelength of the H-band. The filter is designed so that transmission at 3λ/D for the central wavelength is 50%.

how Gaussian filter transmission changes at the extremes of the H-band for a 3 λ/D IWA. Transmisson at 3 λ/D for the central wavelength is 50% by definition, while it is only 43% at 1.5 μm and rises to 56% at 1.7 μm.
In order to assess whether this effect causes a contrast degradation or not, I generated an ADI sequence of 20 images in high Strehl regime (R=8). Since the differences in transmission are very small near the center of the field, chromatism is expexted to be negligible for a perfectly centered star. For this reason, in order to amplify the effect I introduced 14 mas rms of jitter. In figure 60 polychromatic detection limits for different numbers of sampling wavelengths are compared with the monochromatic case (central wavelength). Simulations show better detection limits in the polychromatic case, with a weak trend towards the monochromatic curve as the number of sampling wavelengths increases. We can thus conclude that there is not any major concern for this coronagraph to work in broadband.

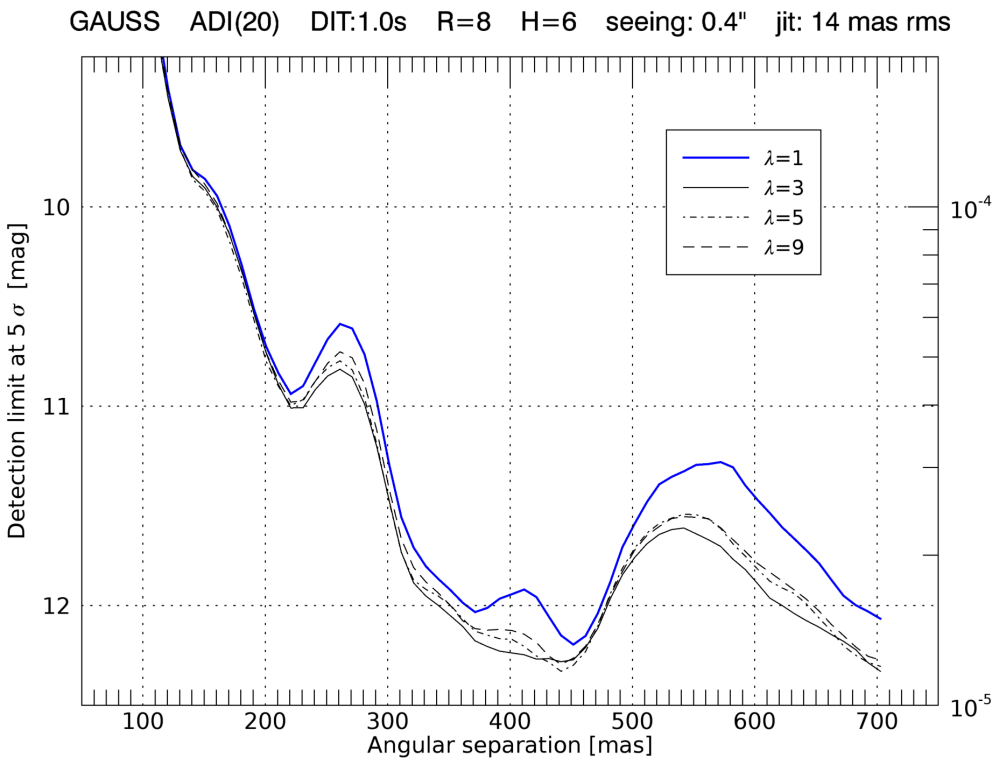


Figure 60: Detection limits in H-band with the Gaussian Lyot coronagraph. The blue curve corresponds to the case of a single wavelength (1.6 μm), while black curves correspond to polychromatic simulations with increasing number of sampling wavelengths λ across the band. Sampling wavelengths are chosen to be equally spaced from 1.5 to 1.7 nm.

Shaped Pupil

The Shaped Pupil coronagraph used in broadband suffers from a problem in some sense similar to the Gaussian Lyot: the boundaries of the high-contrast region created by the apodizer are wavelength dependent. Then how to dimension the focal plane mask? A new 'effective' discovery space can be defined as the intersection of all the spaces as the wavelength moves across the passband. The new effective IWA then becomes the one of the final wavelength of the band while, viceversa, the effective OWA becomes the one of the initial wavelength. The former thus moves outward, while the latter moves inward, causing a shrinking of the discovery space. The variation in angular units with respect to the IWA at the central wavelength can be quantified as:

$$\Delta IWA = IWA_{\lambda_f} - IWA_{\lambda_c} = IWA_{\lambda_c}\left(\frac{\lambda_f}{\lambda_c} - 1\right) = IWA_{\lambda_c} \cdot \frac{\Delta\lambda}{2\lambda_c}$$

The increment is then by half the fractional bandwidth in percentage. It can be easily shown that the same happens, in the opposite direction, for the OWA. If we define the linear extent of the nominal discovery space as the

difference $S = OWA - IWA$, then the new space when observing through a filter of fractional bandwith $\frac{\Delta\lambda}{\lambda_c}$ shrinks by the fractional amount:

$$\frac{\Delta S}{S} = -\frac{\Delta\lambda}{2\lambda_c} \cdot \frac{OWA + IWA}{OWA - IWA}$$

Table 13 reports the effective IWA and OWA for the two selected Shaped Pupil designs, together with the corresponding discovery space fractional shrink. SP2a and SP2b are shown separately because their IWAs are slightly different (see section 4.3).
Is is then necessary to shrink the discovery space by dimensioning the focal plane mask using the new calculated effective IWA and OWA? If we preserve the discovery space by dimensioning focal planes masks at the central wavelength λ_0 of H-band, then we should expect some leackage at both the inner and the outer border of the mask. Simulations show that performance degradation near IWA and OWA is indeed observed (figure 61). However, increasing the number of wavelengths the degradation seems to vanish progressively. Again, I simulated a high Strehl observation with significant residual jitter in order to enhance the investigated effect.
In conclusion, simulations show that it is not strictly necessary to shrink the discovery space by oversizing the dimension of the occulter and downsizing the one of the field stop.

Table 13: Effective IWA and OWA and linear discovery space shrinking for each selected Shaped Pupil design.

Design	IWA_e[mas]	OWA_e[mas]	$\Delta S/S$ **[%]**
SP1	111	301	12
SP2b	141	301	15
SP2a	149	301	16

FQPM

From the manufacturing point of view, the easiest way to provide the π-shift is by deposition of a metallic material on a glass substrate. Layer thickness e must satisfy:

$$2(n-1)e = \lambda$$

where n is the refractive index of the material. The desired shift can thus be obtained only for one single wavelength in this way. Achromatization of the FQPM has been subject of intensive research in the last ten years. Mawet et al. 2006 successfully tested in laboratory a new technology that exploits the birefringent properties of halfwave plates, the same then used to manufacture the FQPM for SPHERE. Hou et al. 2014 proposed and validated via simulations a phase mask dividing the focal plane in six instead of four regions.

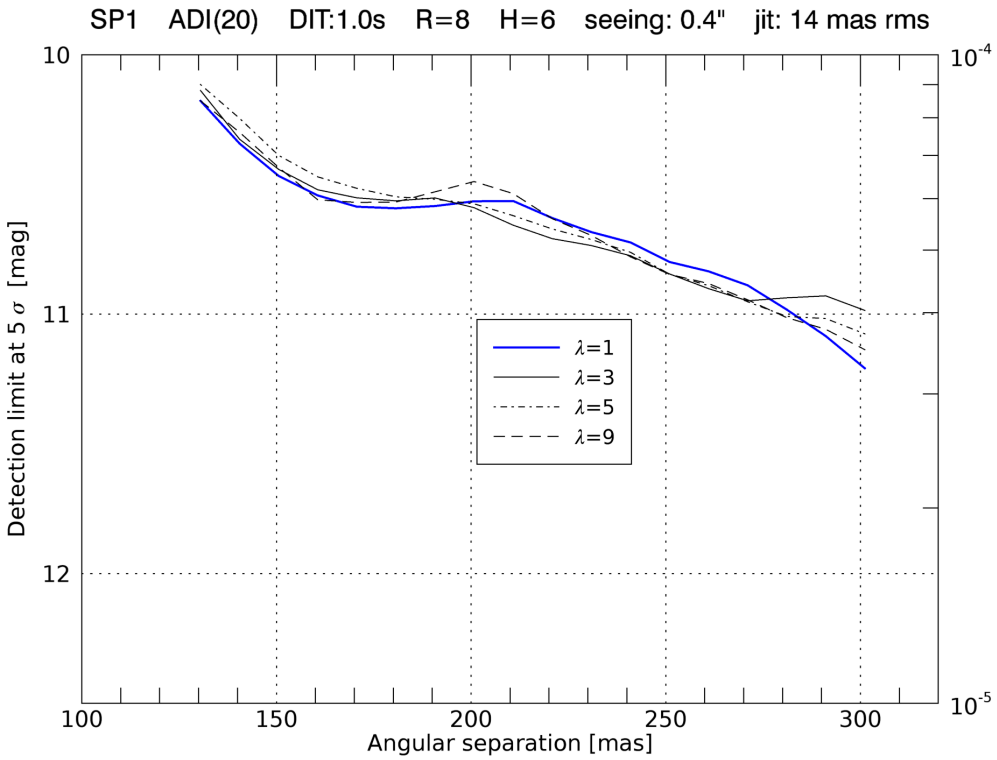


Figure 61: Detection limit in H-band with SP1 using a focal plane mask dimensioned at 1.6 µm. The blue curve corresponds to to the case of a single wavelength (1.6 µm), while black curves correspond to polychromatic simulations with increasing number of sampling wavelengths λ across the band. Sampling wavelengths are chosen to be equally spaced from 1.5 to 1.7 nm.

In the case of SHARK-NIR, the idea is to optimize the design for observation in H-band. Working on a limited wavelength range, the coronagraph may not be severly limited by chromatism even if decide for a standard technology. In order to confirm this hypothesis, I run a polychromatic simulation in H-band to compare a real case in which the phase shift changes with the wavelength with an ideal case in which the shift is π throughout the whole band. In order to maximize the impact of chromatism, I simulated a high-Strehl observation ($\approx$ 95%) without residual jitter. The band is sampled with five equally spaced wavelengths. The result is shown in figure 62. Except for a small region around 150 mas, where a gain of a quarter of magnitude is observable, the difference between the two cases is negligible. This confirms that the FQPM is not limited by chromatism if operated with the H-band filter foreseen for SHARK-NIR.

Finally, I also simulated a six-level phase mask, which is an instance of the six-region phase mask coronagraph proposed by Hou et al. 2014. For the comparison with the FQPM, I used the same Lyot stop. Of course, since we are not limited by chromatism, there should be no real gain in performance, which is indeed confirmed by simulations (see figure 63).

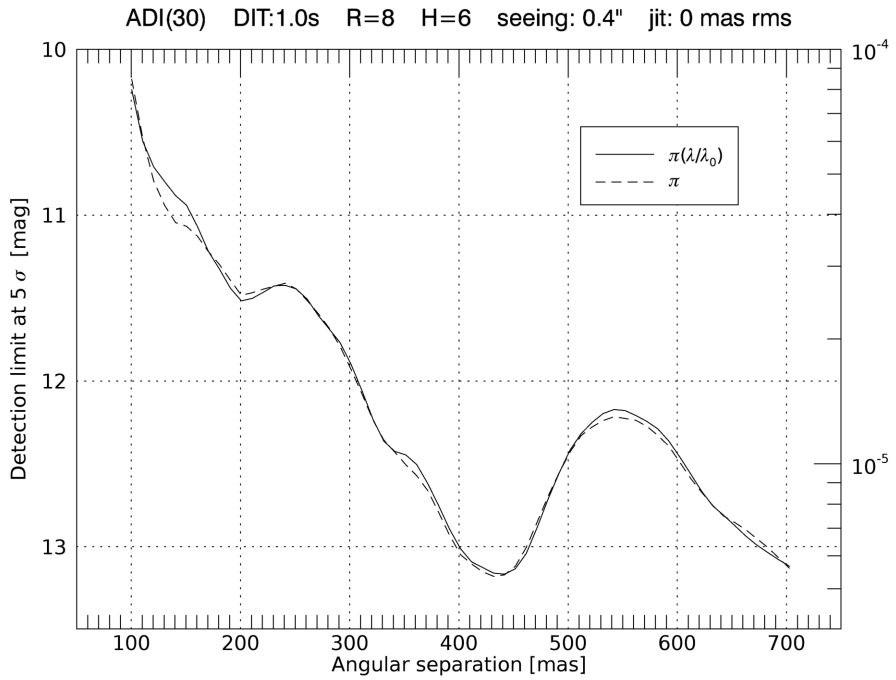


Figure 62: Detection limit in H-band with a FQPM in the ideal case of a π-phase shift at all wavelengths (dashed line) compared to a real case in which the shift scales with the wavelength (solid line).

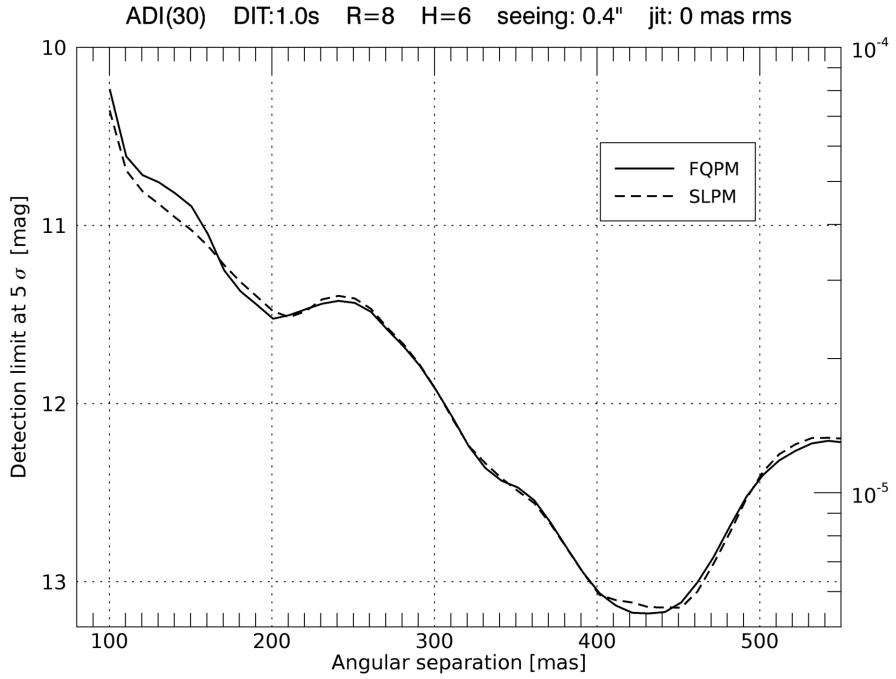


Figure 63: Comparison between detection limit in broadband using a four-quadrant (continuos line) and a six-level (dashed line) phase mask coronagraph. Although the latter is inherently less chromatic, the difference between the two is small, meaning that we are not limited by chromatism even in very high Strehl regime.

In conclusion, simulations show that there is not really any particular concern when operating the designed coronagraphs in broadband. On the contrary, detection limits computed in the monochromatic case might even slightly underestimate the true instrument performance.

7 TOLERANCES ANALYSIS

All the results shown in previous chapters assume a perfectly aligned system. In this chapter I investigate the impact of misalignments on coronagraphic performance. In reporting the results, I will distinguish between misalignments of pupil plane components (section 7.1) from focal plane ones (section 7.2). For both, I will first discuss the stability error budget and then illustrate how these effects impact on detectability.

7.1 MISALIGNMENTS IN PUPIL PLANE

7.1.1 Error budget

Apodizers and Lyot stops are the two kinds of coronagraphic optics in SHARK-NIR intermediate pupil planes. These optics are aligned to the telescope pupil by means of the pupil re-imager (PRI) deployable doublet. Inserting it in the optical path, an image of the apodizer on the scientific focal plane is formed and its position retrieved. The image of LBT pupil is then aligned to this reference before exposure using SHARK-NIR motorized entrance folding mirror (IN_TT hereafter). The reference retrieval can be performed either in a dedicated characterization campaign or directly before pupil adjustment.
The error budget of the alignment procedure is detailed in table 14. The following contributors are considered:

- **LBT pupil stability during exposure**. We here assume that the position stability of LBT pupil during a single exposure (~ 1 hour) is below 0.05% of the overall pupil diameter. This value is very pessimistic, because it corresponds to worst case scenario experienced during ASM commissioning and characterization campaign.

- **PP1/FP-SCI mechanical stability**. The stability of the position of the pupil plane relative to the scientific camera during a single exposure. We assume a relative stability lower than 5 µm, corresponding to ~ 0.05% of the pupil diameter on PP1.

- **PP1/FP-SCI thermal stability**. The position stability of the pupil plane relative to the scientific camera during a single exposure due to thermal variations. From a dedicated study, we retrieved for a conservative ΔT variation of 2° a relative shift of about 1.2 µm, corresponding to ~ 0.012% of the pupil diameter.

- **Apodizer position retrieval**. We estimate to achieve a precision higher than 1 pixel. Being the pupil image sampled with ~ 660 pixel on its diameter, this translates in 0.15% of the pupil size.
- **Pupil position retrieval**. From preliminary tests, an optimized software routine could allow to define the center of the pupil, from the image, with a precision better than 0.2 pixel, corresponding to 0.03% of the pupil size.
- **Repeatability of the apodizing mask wheel positioning**. It drives the actual position of the apodizing mask itself with respect to the pupil. Basically, this term is driven by the requirement set on the wheel repeatability, which is 1 μm at the edge of the rotating device, which translates into 4 arcsec and corresponds to 0.01% of the pupil.
- **PRI positioning repeatability**. This has a (almost) linear effect on the positon of the re-imaged pupil onto the detector. The PRI positioning linear stage unidirectional repeatability requirement is 1 μm, translating into 0.01% of the pupil.
- **Out-of-focus apodizer position**. The image of the apodizer onto the scientific camera is blurred because of its 50 mm displaced along the optical axis with respect to the actual pupil plane. This blur is expected to be small (comparable with the camera pixelsize).
- **IN_TT minimum incremental step**. The sensitivity of the pupil position on the tilting angle of the IN_TT has been computed. In particular, 1/100 of the diameter of the pupil corresponds to a tilt of IN_TT of about 200 arcsec. Since the required minimum incremental step of the IN_TT adjustment is 1 arcsec (actual expected value is about 0.3 arcsec by design), we can assume a pupil displacement better than 0.005%.
- **Rotation alignment**. SHARK-NIR de-rotator can be used to adjust the rotation of the pupil image with respect to the masks, looking at the pupil image on the camera, once the deployable doublet is inserted. This would imply a calibration of the masks rotation angles with respect to the camera. Being the bearing minimum incremental step requirement 20 arcsec, the pupil displacement (at the edge) shall be about 0.006%.

7.1.2 Simulations

The effect of a misalignment between LBT pupil and apodizers is reproduced in simulations by simply displacing the pupil itself while keeping the apodizers centered on the optical axis. The minimum testable shift in simulations is of order of 1 pixel, corresponding to 0.4% of the pupil diameter on a 256 pixels pupil. The pupil displacement is kept constant during the observation. We expect pupil stability to have small impact on coronagraphic

Table 14: SHARK-NIR error budget for centering the apodizer to LBT pupil

CONTRIBUTOR	UNCERTAINTY [PUPIL DIAM %]
LBT PUPIL STABILITY DURING EXPOSURE	0.05
PP1/SCICAM MECHANICAL STABILITY	0.05
PP1/SCICAM THERMAL STABILITY	0.006
APODIZER POSITION RETRIEVAL	0.15
PUPIL POSITION RETRIEVAL	0.03
APODIZERS WHEEL REPEATABILITY	0.01
PRI REPEATABILITY	0.01
OUT-OF-FOCUS	0.15
IN_TT MIN STEP	0.005
FIELD DEROTATOR ACCURACY	0.006
TOTAL	0.23 (rms)

performance: in fact, for those techniques relying on pupil apodization (i.e. SP and APLC), the tolerance to misalignments is specifically accounted for in the design. In particular, all solutions selected for SHARK-NIR are designed to tolerate twice the maximum misalignment estimated from the complete error budget (see chapter 4).
Simulations indeed confirm that there is no loss in performance for SP and APLC for a pupil misalignment equal to the tolerance by-design.
Concerning Lyot stops, they are usually the less critical components in coronagraphs. This can be deduced from the dispersion around the optimal configuration in presence of atmospheric residuals and NCPA, which is clearly visible in the contrast maps used for the designs. Figures 64 compare detection limits when the system is perfectly aligned to a pessimistic case of 1% pupil misalignment with a FQPM. I performed this tests in high Strehl regime, where the sensitivity to pupil movements is expected to be maximized. As expected, no appreciable loss in performance is found. The same holds for the Gaussian Lyot.

7.2 MISALIGNMENTS IN FOCAL PLANE

7.2.1 Error budget

Every simulated coronagraph (except for APP) features a focal plane mask. Before any exposure, the local DM will be moved in tip-tilt to center the PSF on the occulting mask. This will be accomplished by measuring, on the scientific focal plane, the position of the image of the source (without the occulting mask) and then comparing it with the position of the image of

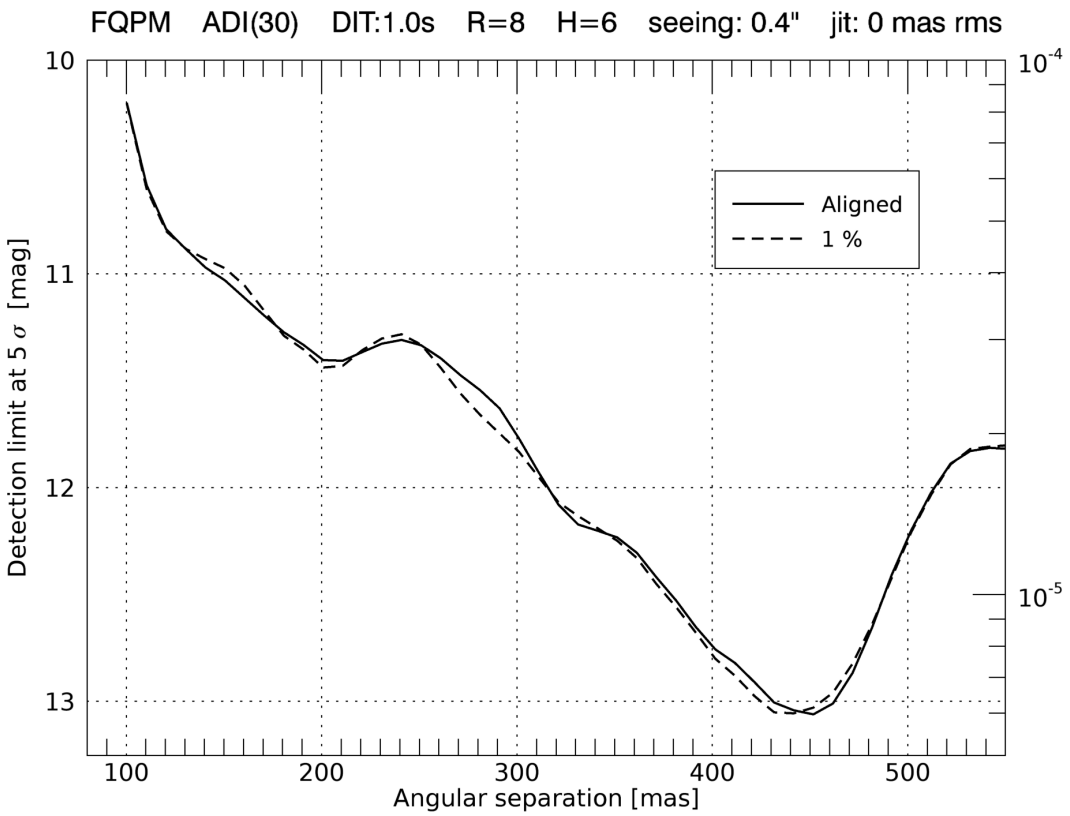


Figure 64: Detection limits with the FQPM in case of a perfectly aligned system (*continuous line*) compared to a case where 1% pupil misalignment is introduced (*dashed line*).

the occulting mask (which can be illuminated with the flat field calibration source). The occulter center is measured first and used as a reference for the alignment of the source using the DM.
The error budget of this procedure is detailed in table 15. The following contributors are considered:

- **DM minimum incremental step**. The minimum incremental step adjustment of the ALPAO DM 97-15 has been computed considering the PtV of the best flat that the mirror can achieve, resulting in ~ 70 nm. Being the mirror diameter 13.5 mm, the corresponding tilt of the mirror is ~ 1 arcsec.

- **Occulting mask wheel positioning repeatability**. It drives the actual position of the occulting mask. If we want to keep this term below 1μm (computed on the edge of the rotating device – conservative case), we need to have a repeatability better than 4 arcsec.

- **PSF center retrieval**. From a preliminary test, a possible PSF position characterization routine could allow a centering better than 0.1 pixel, with a PSF size in the range 4-7 pixel of diameter (this corresponds to ~ 1.26 μm on FP-CORO).

- **Occulting mask center retrieval**. For the occulting masks, we assume, as a first guess, the same position characterization routine tested for PSF centering.

Table 15: SHARK-NIR error budget for centering the star PSF onto the focal plane occulter. The uncertainty is expressed in microns on FP-CORO.

CONTRIBUTOR	UNCERTAINTY [μm]
TTM MINIMUM INCREMENTAL STEP	1.4
OCCULTER WHEEL REPEATABILITY	1
PSF CENTER RETRIEVAL	1.26
OCCULTER CENTER RETRIEVAL	1.26
TOTAL	2.48 (rms)

7.2.2 Simulations

According to the error budget, focal plane masks can be aligned no better than 2.5 μm, corresponding to ~ 3.5 mas. However, with the adopted resolution of 4 pixels per λ/D, the linear sampling of the simulation grid in the focal plane is 9 μm. This number also corresponds to the minimum testable misalignment. It is possible to increase the zero-padding factor by a factor 2 to reach 4 μm, at the cost of halving the pupil sampling. This possibility has been explored for the FQPM and Vortex.
The presence of hard-edge occulter and field stop makes Shaped Pupils rather insensitive to focal plane misalignments. The most critical coronagraphs are FQPM and Vortex, while the Gaussian Lyot is also expected to be robust. As for pupil misalignments, also in the simulations that follow displacements are assumed to be fixed during the observation.

Gaussian Lyot

Figure 65 shows the result of a 9 μm misalignment of the focal plane mask for the Gaussian Lyot coronagraph. In addition to the misalignment, a 10 mas rms jitter is also added. The overall impact on performance is very small, confirming the robustness of this coronagraph.

FQPM

We expect a misalignment of the focal plane mask to impact on the performance in this case. Figure 66 shows detection limits when 4 μm and 9 μm of misalignment are introduced. Again, the comparison is made in a 'bad' case of high SR. Jitter is small (3 mas rms). With a 4 μm misalignment, the loss in performance can be quantified in a quarter of magnitude at separations smaller than 200 mas. If the misalignment is 9 μm, than the loss is of half

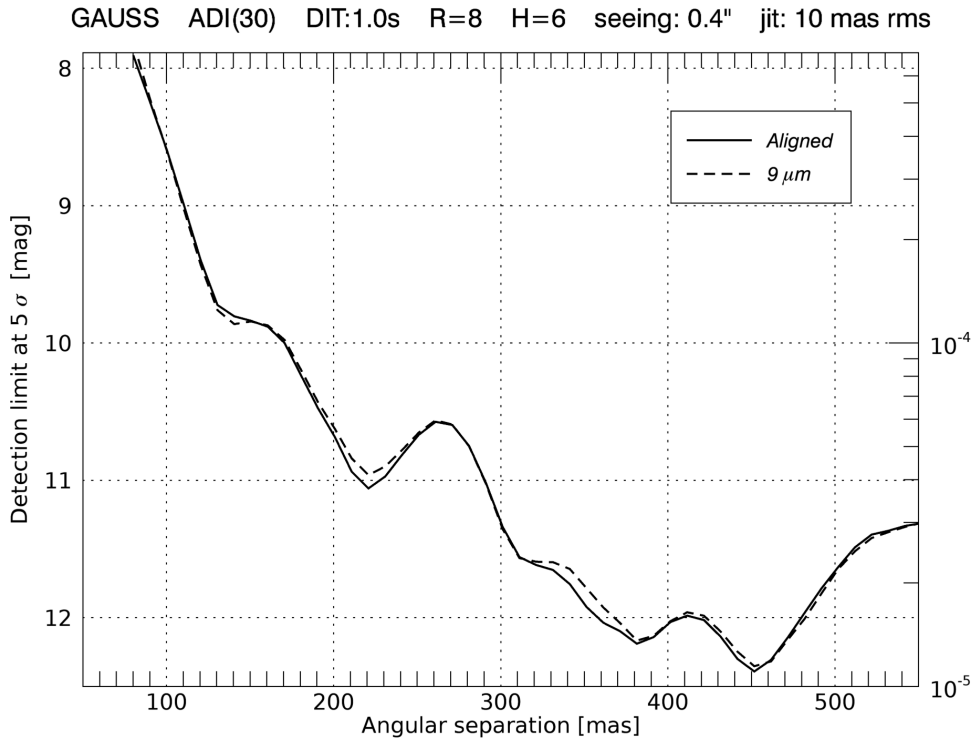


Figure 65: Detection limit with the Gaussian Lyot coronagraph in case of a perfectly aligned system (*continuous line*) compared to a case where the focal plane mask is misaligned by 9 μm (*dashed line*). The impact of misalignment on performance is very small.

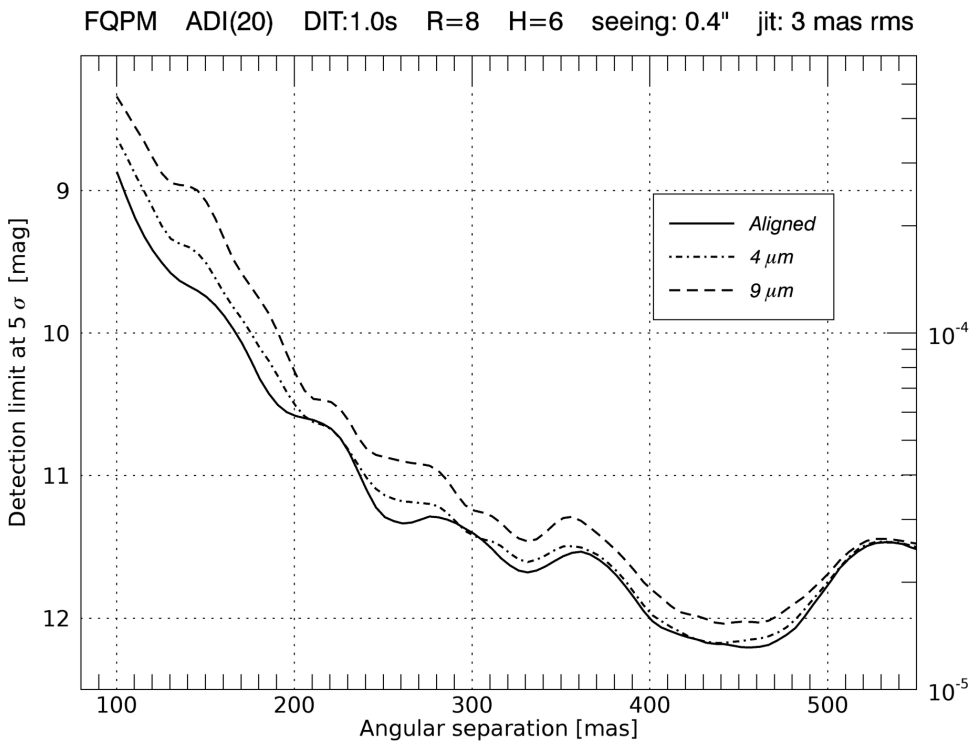


Figure 66: Detection limits with the FQPM coronagraph in case of a perfectly aligned system compared to 4 and 9 μm of misalignment of the focal plane mask (*dashed line*).

a magnitude or more below 300 mas. If residual jitter increases, than the impact of misalignments is reduced (figure 67).
Extrapolating from this plots, if centering precision is really kept below 2.5 μm, then we could expect very small impact on the performance of FQPM.

Vortex

Figure 68 shows detection limits when 4 μm and 9 μm of misalignment are introduced. It can be seen how the Vortex is much more sensitive to this effect: around 200 mas it outperforms the FQPM when perfectly aligned, but it then experiences a loss of order of 1 mag and more in case of a 9 μm displacement, while for FQPM this loss is much smaller. Even at 100 mas, the performance loss is at least twice the FQPM. Even in this case, of course, if centering precision is kept below 2.5 μm, then the impact on performance is not very significant.

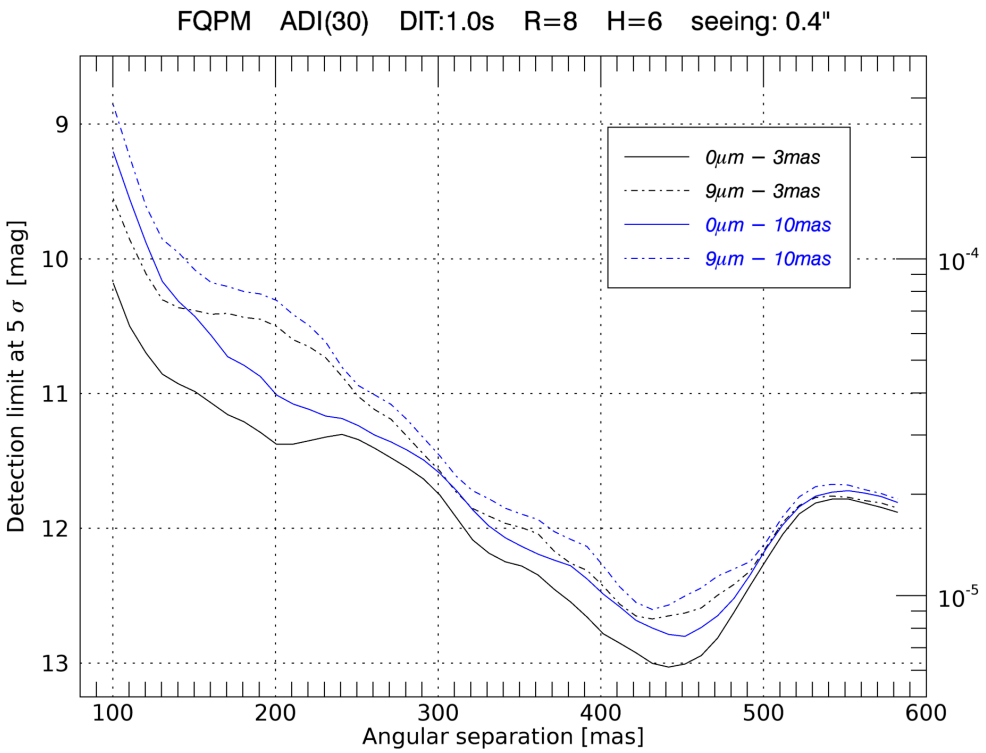


Figure 67: Detection limits with the FQPM coronagraph in case of a perfectly aligned system compared to 9 μm of misalignment of the focal plane mask for two different values of residual jitter.

To summarize, Gaussian Lyot and Shaped Pupil are the most robust coronagraphs to both pupil and focal plane misalignments. FQPM and Vortex are very robust to pupil misalignments, but they show a higher sensitivity to focal plane ones. However, if focal plane mask centering precision will really match the one estimated in the error budget, then none of the two should suffer a critical performance degradation. FQPM, however, is less sensitive than Vortex. The same behaviour is found when looking at the sensitivity to residual jitter when both masks are perfectly aligned. Logically, residual jitter and misalignments both have the effect of displacing the star from the center of the focal plane mask. For this reson, the FQPM has been selected for implementation in SHARK-NIR.

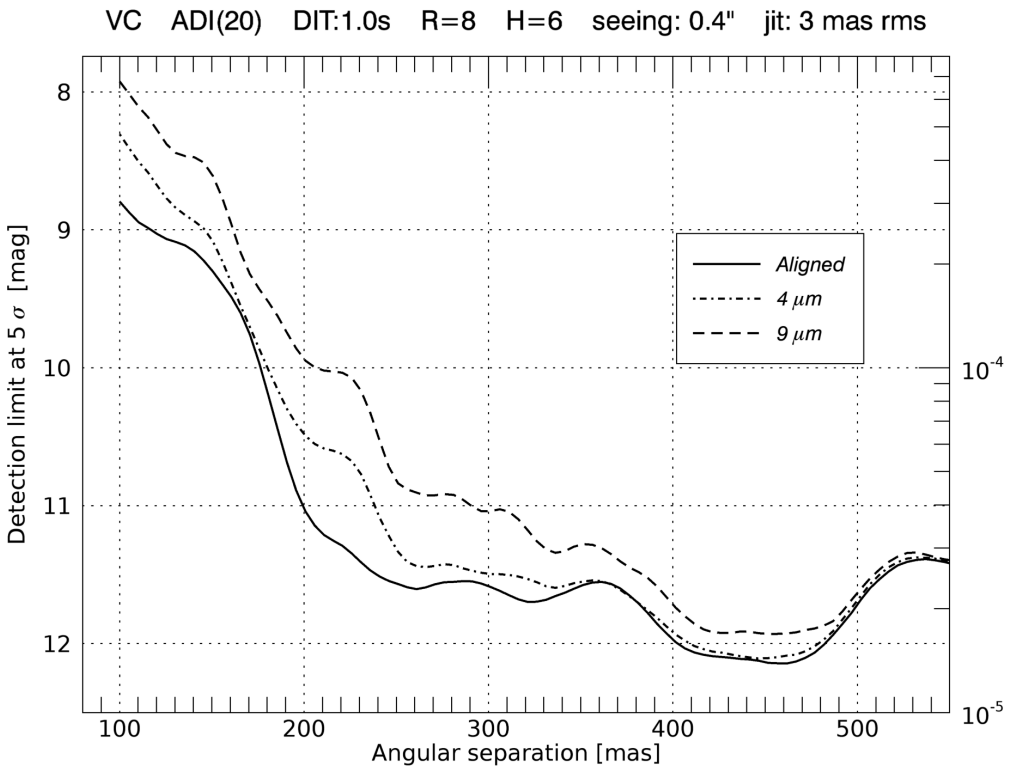


Figure 68: Detection limits with the Vortex coronagraph in case of a perfectly aligned system compared to 4 and 9 μm of misalignment of the focal plane mask.

8 THE PHASE DIVERSITY APPROACH FOR NCPA COMPENSATION

Iin defining a compensation strategy, the key aspect to take into account is how stable we expect NCPA to be. In fact, NCPA typically change slowly in time because of thermal and/or mechanical flectures of the bench. Even thought we expect these effects to be small for SHARK-NIR, a full characterization will be performed only at the telescope during the commissioning phase. If NCPA will not prove to be as stable as we expect, we might need to sense them routinely before observation. A suitable option in such a situation would be to use a phase diversity approach (PD hereafter, see e.g. Mugnier et al. 2008). Phase diversity is a focal plane wavefront sensing technique which is able to retrieve the phase aberration introduced by a camera starting from two images of whatever object, one of which (the diverse image) is intentionally corrupted by a know aberration.
In order to investigate the potential of this method, I applied the algorithm on images generated with the Fresnel simulator. The input NCPA map to be retrieved has been modeled on a realistic error budget of the instrument (section 8.2). Section 8.1 illustrates the theoretical framework, while the results of the analysis are reported in section 8.3.

8.1 THE PRINCIPLE

In incoherent monochromatic imaging, the image I is given by the convolution of the object O with the instrument PSF S:

$$I(r) = O(r) * S(r)$$

The PSF is a function of the phase aberration ϕ in the pupil of system:

$$S(r) = \left| \int_{-\infty}^{+\infty} A(x) exp(i \frac{2\pi}{\lambda F} r \cdot x)\, dx \right|^2 \tag{12}$$

$$A(x) = P(x) exp(i\phi(x)) \tag{13}$$

The relation between the image and the aberrated phase is non-linear: no analytical solution exists to derive the latter from the former. The phase diversity approach basically consists in using a maximum a-posteriori (MAP) approach to estimate the aberrations that are most likely given a set of recorded images and our prior information on the aberration. To apply this method, it is required to record a reference image I_0 of any object 'o' plus one or more diverse images I_k of the same object, differing from the reference by known aberrations ϕ_k. The MAP approach then basically consists in maximizing

the likelihood of the data once a theoretical model for the data themselves is chosen. In a formula, the criterion looks as:

$$J(\phi) = \frac{||I_0(r) - o(\phi, r) * S_0(\phi, r)||^2}{\sigma_0^2} + \sum_k \frac{||I_k(r) - o(\phi, r) * S_k(\phi, \phi_k, r)||^2}{\sigma_k^2}$$

where $||\ ||^2$ denotes the sum of squared pixel values of the argument. It is important to remark that the criterion is a function of the unknown phase ϕ alone, since the solution for the object given ϕ is analytical (that is why ϕ appears as an argument for the object). The model for the data relies on the convolution principle. Sauvage et al. 2012 extended the basic convolution model to include also a coronagraph. This approach, although intriguing, has not been tested so far for SHARK-NIR, but it might be investigated in the future. The adopted model includes detector noise, which is assumed to be stationary white Gaussian noise. Furthermore, its variance σ^2 is assumed to be constant for all images.
Numerical minimization of the criterion is performed via the fast conjugate-gradient method. The phase reconstruction is pixel-wise.

8.2 SHARK-NIR NCPA BUDGET AND MODELING

A dichroic is used to split the beam coming from the telescope in two: visible light is *transmitted* to the WFS, while IR light is *reflected* to the instrument. NCPA then include the contributions of the dichroic itself, the optics in the WFS arm and the optics in SHARK-NIR arm. The full budget of NCPA expected for SHARK-NIR is composed of the following contributions:

WFS ARM

- **Dichroic transmitted wavefront error**: the presence of a refractive element inside the optical train to the AO WFS introduces a WFE. The aberrations introduced by this element are mainly chromatic dispersion ($\sim$ 66%) and astigmatism ($\sim$ 33%). The first has been estimated in $\sim$ 30 nm, which have negligible effect on the WFS (private communication with the AO group in Arcetri). Only astigmatism is thus considered.

- **Dichroic manufacturing tolerance**: in transmission, the total WFE due to manufacturing is the sum in quadrature of both front and back dichroic surfaces. We assumed for the surfaces a $\lambda/50$ rms at HeNe (632.8 nm) over the beam footprint (33 mm). Aberration introduced by manufacturing errors, in the high fractal regime, are characterized by a power spectral density (PSD) $\sim f^{-2}$, where f is the spatial frequency (Dohlen et al. 2011).

- **Wavefront sensor optics**: the contribution from optics internal to the wavefront sensor has been estimated to be negligible.

SHARK-NIR ARM

- **Optical design**: residual aberration from the optimization of the optical quality through SHARK-NIR field.
- **Dichroic manufacturing tolerance**: In reflection, only the front surface contributes to NCPA (with a factor $2\times$ because of the reflection).
- **Manufacturing tolerances (dichroic excluded)**: WFE associated to manufacturing tolerances has been computed considering a bad case (98° percentile) determination of a 10000 Montecarlo analysis, including all the optics manufacturing tolerances, combined together. Again, the aberration introduced by manufacturing, in the high fractal regime, goes as f^{-2}.
- **Alignment tolerances (dichroic excluded)**: an alignment tolerance analysis has been carried on, with respect to the nominal design. A 5000 Montecarlo analysis has been carried on to translate the acceptable misalignments into expected wavefronts. The analysis reveals that most of the residual aberration is astigmatism.
- **ADC**: the design has been optimized in order to minimize the residual spread in waves on the scientific camera, for each considered waveband. The worst case is the H-band, for which the ADC effect results in a residual chromatism on the detector on 1.1 μm. In the budget, we considered the conservative case of the ZA=50°. It shall be noticed that the contribution for lower zenithal angles is more and more negligible.
- **Thermal effect (defocus)**: image plane shift along the optical axis results in an overall enlargement due to defocus at the level of the scientific camera of 2.8 μm.
- **Flexures**: the contribution of flexures has been estimated, and resulted negligible.
- **FP-CORO / FP-SCI differential WFE**: the NCPA that we would ideally correct for are the ones upstream of the coronagraph. The proposed approach for NCPA measurement, however, implies the use of the scientific camera as a sensor, meaning that the contribution of the differential WFE between this plane and the coronagraphic focal plane shall be included in the budget as a contribution which will not be compensated with the ASM. A 5000 Montecarlo Zemax simulation has been performed to quantify this contribution, including both manufacturing and alignment tolerances for all the optical elements, and optimizing the focal position of the scientific camera. The 98° percentile results to be 17nm rms.

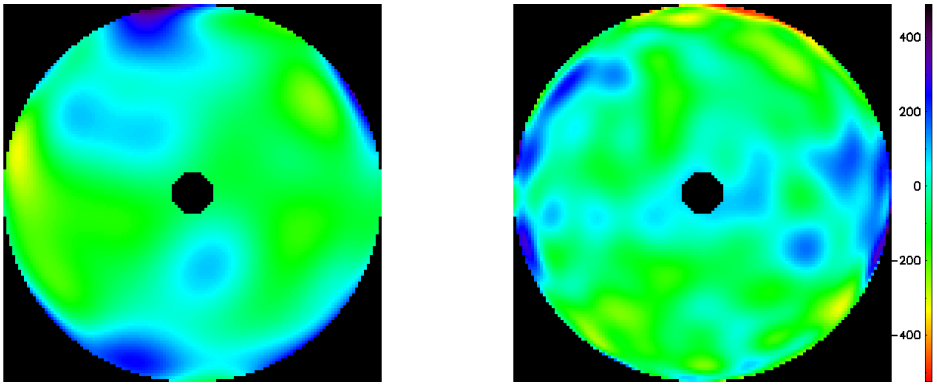


Figure 69: Aberration maps generated according to SHARK-NIR NCPA error budget. *Left*: Map with manufacturing errors distributed on 10 Zernike orders. *Right*: Map with manufacturing errors distributed on 20 Zernike orders. The overall rms is 106 nm in both cases. The maps are displayed on the same linear scale. Units are nanometers.

- **FP-CORO / FP-SCI defocus alignment tolerance**: since the Montecarlo analysis for manufacturing and alignment considers detector refocusing as a compensator, the tolerance in the positioning the camera along the optical axis during alignment shall be considered as a separated contribution for the overall budget. This contribution cannot be compensated, since it can't be disentangled from the measurement, taken at the level of the camera itself. A 250 μm residual displacement has been assumed.

Table 16 reports all budgeted NCPA contributions in terms of WFE and type of aberration. The overall expected WFE is of ~ 100 nm.
The first step in the analysis consists in generating an NCPA map starting from this budget. For the purpose, I chose a modal approach based on Zernike polynomials. Some contributions are easy to deal with since they correspond to given polynomials. On the other hand, manufacturing errors are only specified in terms of a power spectrum. To model them, I generated several realizations of the f^{-2} spectrum using random phases, fitted them with Zernike polynomials and then averaged the coefficients mode by mode. All the other contributions in the budget are then added in quadrature according to the type of aberration to obtain the final NCPA map. In fitting the random realizations, the maximum number of radial orders is not theoretically constraint. For this reason, I generated both a map assuming only ten radial orders (66 polynomials) and a more pessimistic case of twenty radial orders (231 polynomials). The maps are displayed in figure 69.

Table 16: SHARK-NIR NCPA error budget.

Contributor	rms WFE [nm]	Aberration type/spectrum
WFS ARM		
Dichroic transmitted WFE	13	Astigmatism
Dichroic manufacturing tolerance	28	f^{-2}
SHARK-NIR ARM		
Optical design	16	-
Dichroic manufacturing tolerance	23	f^{-2}
Manufacturing tolerances (dichroic excluded)	90	f^{-2}
Alignment tolerances (dichroic excluded)	35	29 nm Astigmatism 2 nm Coma 3 nm Spherical aberration
ADC	8.7	Chromatic dispersion
Thermal effect (defocus)	3	Defocus
FP-CORO / FP-SCI differential WFE	17	15 nm f^{-2} 9 nm Astigmatism
FP-CORO / FP-SCI defocus alignment tolerance	10	Defocus

8.3 SIMULATIONS

The aberration commonly chosen for phase diversity applications is defocus. The biggest advantage of this choice resides in the small practical effort required to introduce and calibrate it (see Jolissaint et al. 2012 for an example of phase diversity experiment). In my simulations, I used defocus as the diverse aberration as well. I did not consider any other diverse image in the analysis: the algorithm is then applied to two images: a reference (focused, hereafter) and a defocused one. I remark that both images could be generated simultaneously on the science camera, exploiting SHARK-NIR dual band imaging mode. An optic composed of a flat filter next to a low-power plano-convex lens will then introduce the desired defocus on one of the beams.

To test the algorithm, the simulator has been used to generate several couples of focus-defocus images under different conditions. An hypothetical acquisition can not be performed using our internal calibration source, otherwise we would be blind to part of the non-common path (the dichroic for example). The only possibility is to use directly the light of a star. In order to reduce as much as possible the contamination from the atmosphere, then the observation should be carried out in closed-loop. To form the image, I then added to NCPA aberrations the fast evolving AO residuals, like all 'standard' simulations discussed in previous chapters. Telescope jitter is also included. Pre-processing of raw images is also performed in the usual way. All noise sources are included (see section 3.2). In addition, a 'virtual' neutral density filter is eventually applied to keep the flux below the CCD saturation level. This is particularly important for phase diversity, since in presence of noise the stellar flux obviously plays a role in the reconstruction. In particular, being the images acquired simultaneously, if a neutral density filter has to be applied then it will also attenuate the defocused beam. This effect is accounted for in pre-processing of images. Figure 70 shows an example of the two images generated with the simulator. Simulations are monochromatic. The wavelength is 1.558 μm, corresponding to the central one of the H-band narrow-band filter foreseen for SHARK-NIR. The images are then given in input to the phase-reconstruction algorithm. Figure 71 compares the reconstructed map with the original one. The algorithm takes one to two minutes to converge.

A simple subtraction of the two maps tells us about the theoretical reconstruction capability of the algorithm. This operation gives the smallest possible error we can get. In practice, there is also a fitting error due to how well the DM is able to reproduce the reconstructed shape. At a first stage, I will account for this contribution by fitting Zernike polynomials to the reconstructed map. This approach introduces a new free parameter in the analysis, namely the number of polynomials (modes) used for the fit. For this reason, reconstruction error in the following will always be given

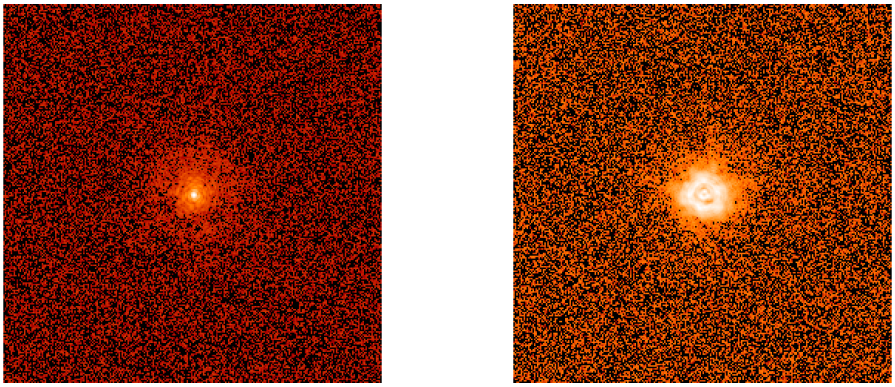


Figure 70: Example of a focused (*Left*) and defocused (*Right*) image generated with the simulator. The scale is logaritmic, but is not the same in the two images.

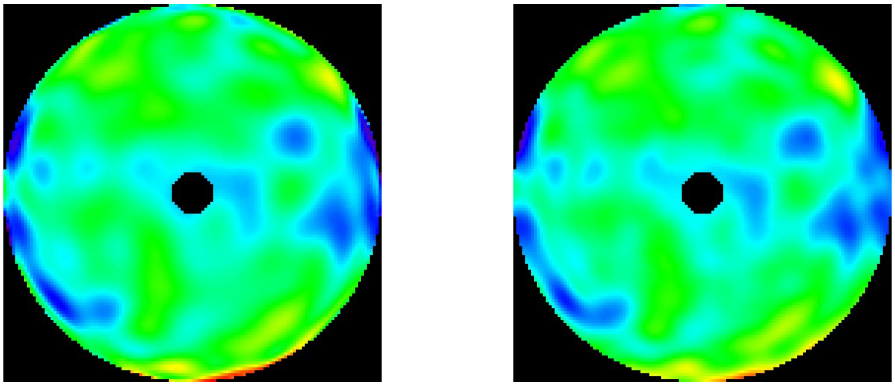


Figure 71: *Left*: the original NCPA aberration map. *Right*: aberration map reconstructed with the phase diversity algorithm. OPDs in the two images are displayed on the same scale.

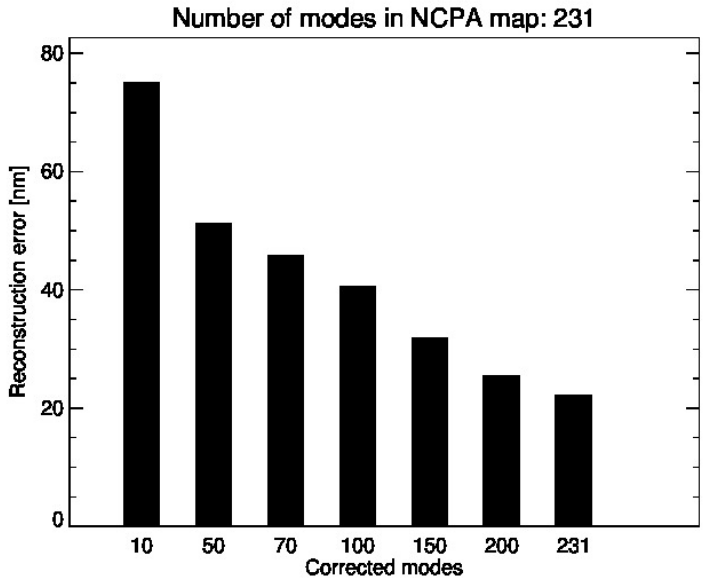


Figure 72: Reconstruction error using phase diversity as a function of the number of modes used to simulate DM fitting error.

as a function of the number of modes corrected by the hypothetical DM. Figure 72 shows the results for the example discussed in this section. Reconstruction error ranges from ~ 75 nm correcting only 10 modes down to ~ 20 nm when correcting 231 modes, which correspond in this case to the same number of modes used to generate the NCPA map.

Of course, using a simple Zernike fit to estimate the DM fitting error is somewhat simplistic. However, this approach has been useful to have an estimate at hand of a reasonable number of modes to be corrected. This quantity direcly translates into a requirement on the number of actuators on the DM. After the DM was selected, ALPAO provided us with influence functions measured directly on the mirror. As discussed at the end the following section, this allowed me to accurately introduce the fitting error in the analysis and compare the results with the previous simpler modelling.

8.4 RESULTS

In this section I report the main results of the simulation campaign I run to constrain the parameters at play in PD reconstruction. In particular, I investigated the sensitivity of the algorithm to the assumption on NCPA spectrum, to the amount of defocus, images integration time and telescope vibrations.

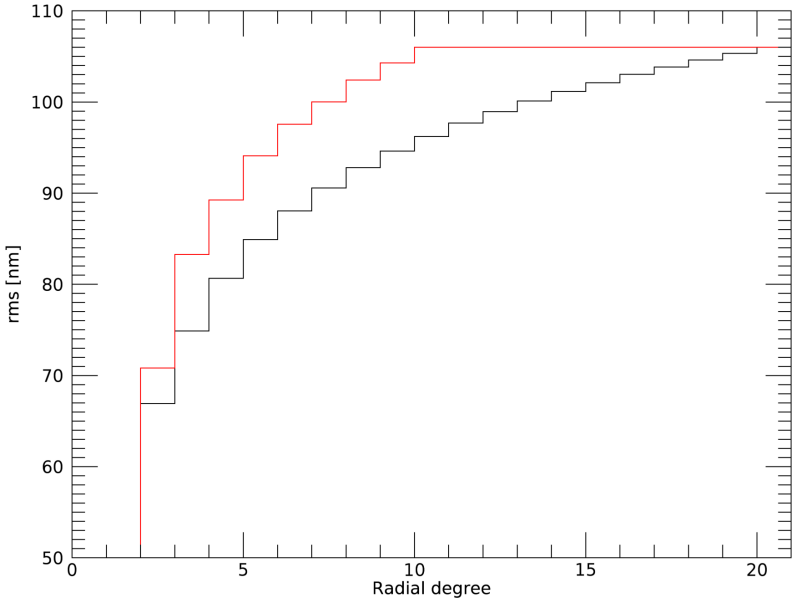


Figure 73: Cumulative power spectra of NCPA. Red line refers to the low-order assumption, namely manufacturing errors are distributed inside the first 10 Zernike degrees. Black line refers to the high-order case, where power is distributed up to the radial degree 20.

8.4.1 NCPA spectrum

As detailed in section 8.2, the fact that manufacturing errors are modeled from their PSD introduces an important degree of freedom in the analysis, namely the number of modes over which these errors are distributed. In order to investigate the sensitivity of the algorithm to this parameter, I tested both a low-order (10 radial degrees) and a high-order (20 radial degrees) aberration spectrum. Figure 73 shows the cumulative power spectra in the two cases. I report in figure 74 reconstruction errors Err_{10}, Err_{20} as a function of the number of corrected modes. It is clearly visible how critical the assumption on the NCPA spectrum is. An almost linear trend is observed in the difference $\mathrm{Err}_{10} - \mathrm{Err}_{20}$ as the number of corrected modes increases. While 60 modes are sufficient to obtain a residual of 20 nm in the low-order case, in presence of high orders the residual is as high as 50 nm.
This test confirms that PD has limited sensitivity to high order aberrations. Moreover, NCPA power spectrum is definitely the dominant source of uncertainty in the analysis. Unless otherwise specified, results showed in the following will assume the high-order hypothesis.

8.4.2 Defocus

The problem of which is the optical defocus for PD is argumented in several papers. Since there is no clear agreement on the topic, I decided to

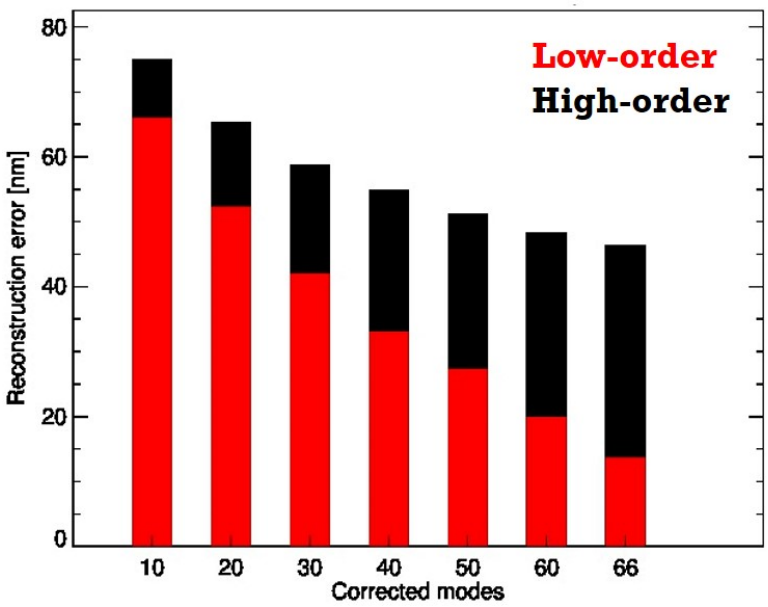


Figure 74: Reconstruction error using PD as a function of the number of corrected modes for the two NCPA power spectra illustrated in figure 73.

investigate the impact of this parameter on the reconstruction. The quantity I used to characterize the defocus is the amplitude of the aberration (half the ptv), which I express in units of λ. Figure 75 shows the reconstruction error as a function of defocus amplitude. Different curves refer different numbers of corrected modes. Strehl ratio is high ($R = 8$, seeing is $0.4''$). Reconstruction error shows a weak dependence on the amplitude of the defocus. As the number of corrected modes increases, then then a minimum around 1 becomes more and more visible. In this test, the 231-modes NCPA map is of course used.

8.4.3 Integration time

Integration time is another important parameter, since it directly affects the amount of time required for NCPA sensing. In order to constrain it, I computed the reconstruction error integrating the two images from 1 up to 30 seconds. These errors result very similar one to each other: fluctuations are comparable with the inherent uncertainty due to random contributions to image formation (basically photon and read-out noise, table 17). The important conclusion of this analysis is that, in closed-loop, one second of integration is sufficient to apply PD. I repeated the analysis also in conditions of lower Strehl ratios and the results are very similiar. We can thus conclude that the bottleneck of this procedure is not in the integration time,

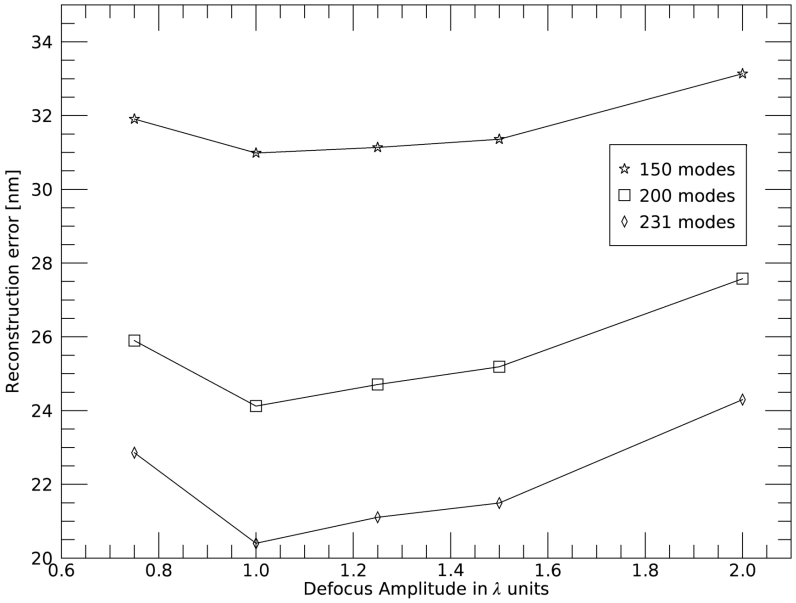


Figure 75: Reconstruction error using PD as a function of the amount of introduced defocus and of the number of corrected modes. A minimum around 1 λ defocus amplitude becomes more and more visible as the number of corrected modes increases.

but rather in the algorithm convergence.

Table 17: Fluctuations in PD reconstruction error as DIT increases from 1 up to 30 seconds as a function of the number of corrected modes. These fluctuations are small and consistent with the inherent uncertainty on reconstruction due to random noises in the processed images.

# of modes	σ_{DIT} **[nm]**	σ_{Rand} **[nm]**
50	0.11	0.11
70	0.17	0.16
100	0.17	0.22
150	0.26	0.21
200	0.37	0.36
231	0.28	0.47

8.4.4 Vibrations

To test the robustness of the algorithm to vibrations, I used the of 70 mas rms open-loop spectrum and simulated a faint star (R = 12). Residual vibrations in this regime are as high as 20 mas rms. I assumed here the low-order hypothesis for the NCPA spectrum. Reconstruction error is shown in figure 76). The error correcting 66 modes only increases from 24

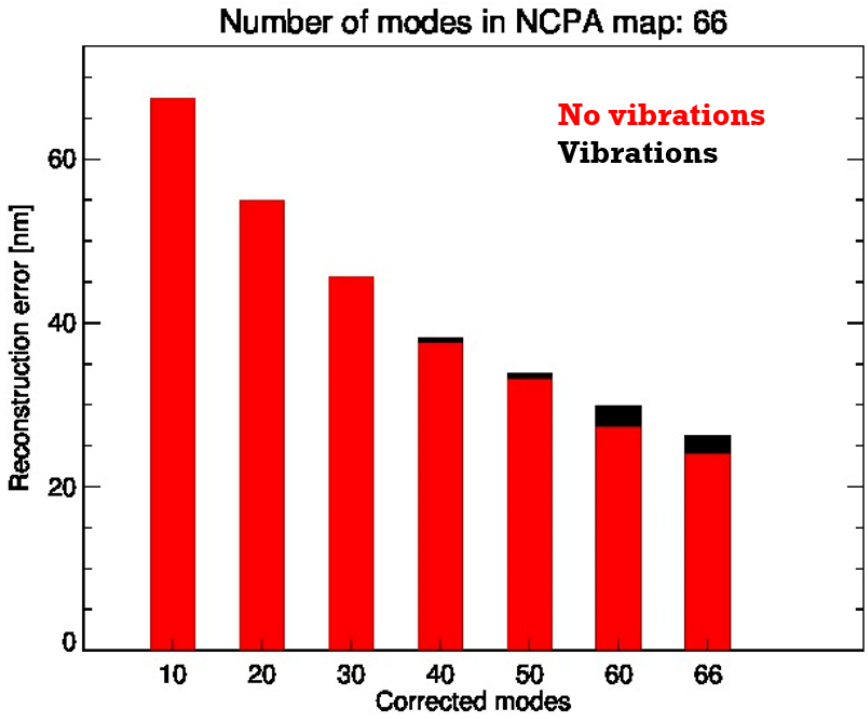


Figure 76: Comparison between reconstruction error using PD in a case of no vibrations (*red bars*) and a case of very strong residuals (20 mas, *black bars*. The difference is small, meaning the the algorithm is rather insensitive to vibrations.

to 26 nanometers in presence of this strong residual. The time required to converge, however, increases by 50%, from 1.8 to 2.6 minutes. We can thus conclude that PD is very robust to vibrations, whose only effect is the one of slowing down the algorithm convergence.

8.4.5 The ALPAO DM 97-15

Looking for example at figure 72, it is clear how in the hypothesis of high-orders the ideal condition would that of correcting as much modes as possible. However, going to 200 modes means at least 200 actuators, with a strong impact on opto-mechanics and costs. Finally we thus opted for a mirror having only 100 actuators. Correcting 100 modes, we expect a reconstruction error of order of 40 nm, which is slightly higher than the assumption made throughout the design phase (30 nm). Under the low-orders hypothesis, however, the residual will be less than 20 nm. The number used in the design phase is thus in-between the two results.
ALPAO provided us with the measured influence matrix of the mirror. This allowed to repeat the analysis by introducing the real fitting error of the mirror. This is accomplished by taking the pseudo-inverse of the influence

matrix and appying it to the PD-reconstructed map in order to derive mirror commands. This commands were then multiplied again for the influence matrix to obtain the best-fit shape. Doing this experiment, I obtained a residual of 37 nm (with the high-orders assumption), which is perfectly consistent with a 100 modes correction.

In conclusion, the reconstruction capability of phase diversity is strongly dependent on the assumption on NCPA power spectrum. The residual is expect to be in the range 20-40 nm. Simulations showed that the defocus shall be around 1 λ in amplitude, while integration time in the two images can be as small as 1 second. Finally, the reconstruction is very robust to vibrations.

CONCLUSIONS

SHARK is an instrument proposed for the Large Binocular Telescope (LBT) in the framework of the "2014 Call for Proposals for Instrument Upgrades and New Instruments". Thanks to a visibile and a near infrared channel, to be installed one for each LBT telescope, and exploiting the synergy with already installed instruments, SHARK will provide LBT with simultaneous spectral coverage from optical to mid-infrared for high-contrast imaging. The outstanding performances of LBT binocular Extreme Adaptive Optics (XAO) system, especially in terms of sensitivity in the faint-end regime thanks to the pyramid WFS, will allow SHARK to adress several fundamental topics in modern astrophysics, from exoplanets discovery and characterization to imaging of jets and disks around young stars up to distant AGN.
In my Ph.D. I have been involved in the design of the near infrared channel. In particular, I developed a code in IDL language to simulate the process of image formation using an end-to-end approach based on scalar theory of diffraction. The code has been used as a test bench of several coronagraphic techniques, thus contributing to the overall instrument error budget finalized at assessing SHARK-NIR final expected performance. Some of the investigated techniques have been designed by me, while some others by collaborators at IPAG-Grenoble. This collaboration also foresees tests in the lab once the coronagraphs will be assembled, offering the opportunity of joint publications in the near future.
After extensive simulation campaigns, the work presented in this thesis allowed to finally select the coronagraphic techniques to be implemented in the instrument and, moreover, allowed to show that the scientific requirements can be fulfilled. In summary, the Gaussian Lyot coronagraph is the option to serve all those science cases requiring field-stabilization and moderate contrast. Observations in pupil-stabilized mode to search for exoplanets can take advantage of two Shaped Pupils (SP) and a four-quadrant phase mask (FQPM) coronagraph. The SP are designed for high contrast on a small field close to the star and are robust to image and pupil jitter, while the FQPM allows to access the entire scientific FoV and delivers excellent performance in ideal conditions (high Strehl ratios), but performance is still good, both close and further away from the star, even at lower Strehl and with moderate vibrations.
Besides of coronagraphs design, I also investigated via simulations two critical issues for the instrument, namely residual vibrations and NCPA. Simulations contributed in making important design choices in this sense (local fast Tip-Tilt correction and introduction of an internal DM). In particular, the phase diversity approach has been investigated and proven to be a valuable option for SHARK-NIR. Considering the expected amount of aberrations ($\approx$ 100 nm), the estimated residual is between 20-40 nm when correcting

$\approx$ 100 modes, depending on NCPA power spectrum.
SHARK-NIR successfully passed the Final Design Review in January 2017 and the instrument construction started with the procurement phase in July 2017.

BIBLIOGRAPHY

Agapito, G., A. Puglisi, and S. Esposito

2016 "PASSATA: object oriented numerical simulation software for adaptive optics", *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 9909, p. 9.

Aime, C., R. Soummer, and A. Ferrari

2002 "Total coronagraphic extinction of rectangular apertures using linear prolate apodizations", *Astronomy and Astrophysics*, 389, pp. 334-344, DOI: 10.1051/0004-6361:20020419.

ALMA partnership, C. L. Brogan, L. M. Perez, et al.

2015 "First Results from High Angular Resolution ALMA Observations Toward the HL Tau Region", *The Astrophysical Journal Letters*, 808, L3, DOI: 10.1088/2041-8205/808/1/L3.

Ann, D., D. Terndrup, M. Pinsonneault, D. Paulson, R. Hanson, and J. Stauffer

2007 "The Distances to Open Clusters from Main-Sequence Fitting. III. Improved Accuracy with Empirically Calibrated Isochrones", *The Astrophysical Journal*, 655, pp. 233-260, DOI: 10.1086/509653.

Antoniucci, S., A. La Camera, B. Nisini, T. Giannini, D. Lorenzetti, D. Paris, and E. Sani

2014 "The HH34 outflow as seen in [Fe ii] 1.64 μm by LBT-LUCI", *Astronomy & Astrophysics*, 566, p. 5, DOI: 10.1051/0004-6361/201423944.

Antoniucci, S., L. Podio, B. Nisini, F. Bacciotti, E. Lagadec, E. Sissa, A. La Camera, T. Giannini, H. M. Schmid, R. Gratton, M. Turatto, S. Desidera, M. Bonnefoy, G. Chauvin, C. Dougados, A. Bazzon, C. Thalmann, and M. Langlois

2016 "Sub-0.1" optical imaging of the Z CMa jets with SPHERE/ZIMPOL", *Astronomy & Astrophysics*, 593, p. 6, DOI: 10.1051/0004-6361/201628968.

Boccaletti, A., P. Riaud, P. Baudoz, J. Baudrand, D. Rouan, D. Gratadour, F. Lacombe, and A.-M. Lagrange

2004 "The Four-Quadrant Phase-Mask Coronagraph. IV. First Light at the Very Large Telescope", *The Publications of the Astronomical Society of the Pacific*, 116, pp. 1061-1071, DOI: 10.1086/425735.

Bonavita, M., S. Daemgen, S Desidera, R. Jayawardhana, M. Janson, and D. Lafreniere

2014 "A new Sub-stellar Companion around the Young Star HD 284149", *The Astrophysical Journal Letters*, 791, L40, DOI: 10.1088/2041-8205/791/2/L40.

Bonnefoy, M., G.-D. Marleau, R. Galicher, H. Beust, A.-M. Lagrange, J.-L. Baudino, G. Chauvin, S. Borgniet, N. Meunier, J. Rameau, A. Boccaletti, A. Cumming, C. Helling, D. Homeier, F. Allard, and P. Delorme

2014 "Physical and orbital properties of β Pictoris b", *Astronomy and Astrophysics*, 567, L9, DOI: 10.1051/0004-6361:201424041.

Burrows, A., M. Marley, W. B. Hubbard, J. I. Lunine, T. Guillot, D. Saumon, R. Freedman, D. Sudarsky, and C. Sharp

1997 "A Nongray Theory of Extrasolar Giant Planets and Brown Dwarfs", *The Astrophysical Journal*, 491, pp. 856-875, DOI: 10.1086/305002.

Carlotti, A., R. Vanderbei, and N. J. Kasdin

2011 "Optimal pupil apodizations of arbitrary apertures for high-contrast imaging", *Optics Express*, 19, pp. 26796-26809, DOI: 10.1364/OE.19.026796.

Carolo, E., D. Vassallo, et al.

in prep .

Carolo, E., D. Vassallo, J. Farinato, M. Bergomi, M. Bonavita, A. Carlotti, V. D'Orazi, D. Greggio, D. Magrin, D. Mesa, E. Pinna, A. Puglisi, M. Stangalini, C. Verinaud, and V. Viotto

2016 "A comparison between different coronagraphic data reduction techniques", *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 9909, DOI: 10.1117/12.2232792.

Dohlen, K., F. P. Wildi, P. Puget, D. Mouillet, and J.-L. Beuzit

2011 "SPHERE: Confronting in-lab performance with system analysis predictions", *Second International Conference on Adaptive Optics for Extremely Large Telescopes*.

Dupuy, Trent J., Michael C. Liu, and Michael J. Ireland

2009 "Dynamical Mass of the Substellar Benchmark Binary HD 130948BC", *The Astrophysical Journal*, 692, pp. 729-752, DOI: 10.1088/0004-637X/692/1/729.

2014 "New Evidence for a Substellar Luminosity Problem: Dynamical Mass for the Brown Dwarf Binary Gl 417BC", *The Astrophysical Journal*, 790, pp. 133-146, DOI: 10.1088/0004-637X/790/2/133.

Esposito, S., A. Riccardi, L. Fini, E. Pinna, A. Puglisi, F. Quiros, M. Xompero, R. Briguglio, L. Busoni, P. Stefanini, C. Arcidiacono, G. Brusa, and D. Miller

2011 "LBT AO on-sky results", *Second International Conference on Adaptive Optics for Extremely Large Telescopes*.

Farinato, J., C. Baffa, A. Baruffolo, M. Bergomi, L. Carbonaro, A. Carlotti, M. Centrone, J. Codona, M. Dima, S. Esposito, D. Fantinel, G. Farisato, W. Gaessler, E. Giallongo, D. Greggio, P. Hinz, F. Lisi, D. Magrin, L. Marafatto, F. Pedichini, E. Pinna, A. Puglisi, R. Ragazzoni, B. Salasnich, M. Stangalini, C. Verinaud, and V. Viotto

2015 "The NIR arm of SHARK: System for coronagraphy with High-order Adaptive optics from R to K bands", *International Journal of Astrobiology*, 14, pp. 365-373, DOI: 10.1017/S1473550414000494.

Foo, G., D. M. Palacios, and G. A. Swartzlander Jr.

2005 "Optical vortex coronagraph", *Optics Letters*, 30, pp. 3308-3310, DOI: 10.1364/OL.30.003308.

Garufi, A., S. P. Quanz, H. Avenhaus, E. Buenzli, C. Dominik, F. Meru, M. R. Meyer, P. Pinilla, H. M. Schmid, and S. Wolf

2013 "Small vs. large dust grains in transitional disks: do different cavity sizes indicate a planet? SAO 206462 (HD 135344B) in polarized light with VLT/NACO", *Astronomy and Astrophysics*, 560, A105, DOI: 10.1051/0004-6361/201322429.

Ghedina, A. and R. Ragazzoni

1997 "Optimum configurations for two off-axis parabolae used to make an optical relay", *Journal of Modern Optics*, 44, 7, pp. 1259-1267, DOI: 10.1080/09500349708230735.

Gorlova, H. I., M. R. Meyer, G. H. Rieke, and J. Liebert

2003 "Gravity Indicators in the Near-Infrared Spectra of Brown Dwarfs", *The Astrophysical Journal*, 593, pp. 1074-1092, DOI: 10.1086/376730.

Guerri, G., J.-B. Daban, S. Robbe-Dubois, R. Douet, L. Abe, J. Baudrand, M. Carbillet, A. Boccaletti, P. Bendjoya, C. Gouvret, and F. Vakili

2011 "Apodized Lyot coronagraph for SPHERE/VLT: II. Laboratory tests and performance", *Experimental Astronomy*, 30, pp. 59-81, DOI: 10.1007/s10686-011-9220-y.

Guyon, O., E. A. Pluzhnik, M. J. Kuchner, B. Collins, and S. T. Ridgway

2006 "Theoretical Limits on Extrasolar Terrestrial Planet Detection with Coronagraphs", *The Astrophysical Journal Supplement Series*, 167, pp. 81-99, DOI: 10.1086/507630.

Hagelin, S., E. Masciadri, and F. Lascaux

2010 "Wind speed vertical distribution at Mt Graham", *Monthly Notices of the Royal Astronomical Society*, 407, pp. 2230-2240, DOI: 10.1111/j.1365-2966.2010.17102.x.

Hou, F., Q. Cao, M. Zhu, and O. Ma

2014 "Wide-band six-region phase mask coronagraph", *Optics Express*, 22, 1, DOI: 10.1364/OE.22.001884.

Jolissaint, L., L. M. Mugnier, C. Neyman, J. Christou, and P. Wizinowich

2012 "Retrieving the telescope and instrument static wavefront aberration with a phase diversity procedure using on-sky adaptive optics corrected images", *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 8447, 844716, DOI: 10.1117/12.926637.

Kasdin, N. J., R. Vanderbei, D. Spergel, and M. Littman

2003 "Extrasolar Planet Finding via Optimal Apodized-Pupil and Shaped-Pupil Coronagraphs", *The Astrophysical Journal*, 582, pp. 1147-1161, DOI: 10.1086/344751.

Kenworthy, M., J. Codona, P. Hinz, J. Angel, A. Heinze, and S. Sivanandam

2007 "First On-Sky High-Contrast Imaging with an Apodizing Phase Plate", *The Astronomical Journal*, 660, pp. 762-769, DOI: 10.1086/513596.

Kirkpatrick, J. D.

2005 "New Spectral Types L and T", *Annual Review of Astronomy & Astrophysics*, 43, pp. 195-245, DOI: 10.1146/annurev.astro.42.053102.134017.

2013 "Cold brown dwarfs with WISE: Y dwarfs and the field mass function", *Astronomische Nachrichten*, 334, pp. 26-31, DOI: 10.1002/asna.201211759.

Kraus, A. L. and M. J. Ireland

2012 "LkCa 15: A Young Exoplanet Caught at Formation?", *The Astrophysical Journal*, 745, 5, DOI: 10.1088/0004-637X/745/1/5.

Krist, J. E.

2007 "PROPER: an optical propagation library for IDL", *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 6675, 66750P, DOI: 10.1117/12.731179.

Lagrange, A.-M., D. Gratadour, G. Chauvin, T. Fusco, D. Ehrenreich, D. Mouillet, G. Rousset, D. Rouan, F. Allard, E. Gendron, J. Charton, L. Mugnier, P. Rabou, J. Montri, and F. Lacombe

2009 "A probable giant planet imaged in the β Pictoris disk. VLT/NaCo deep L'-band imaging", *Astronomy and Astrophysics*, 493, pp. L21-L25, DOI: 10.1051/0004-6361:200811325.

Liu, M. C., D. A. Fisher, J. R. Graham, J. P. Lloyd, G. W. Marcy, and R. P. Butler

2002 "Crossing the Brown Dwarf Desert Using Adaptive Optics: A Very Close L Dwarf Companion to the Nearby Solar Analog HR 7672", *The Astrophysical Journal*, 571, pp. 519-527, DOI: 10.1086/339845.

Liu, Michael C., Trent J. Dupuy, and Michael J. Ireland

2008 "Keck Laser Guide Star Adaptive Optics Monitoring of 2MASS J15344984-2952274AB: First Dynamical Mass Determination of a Binary T Dwarf", *The Astrophysical Journal*, 689, pp. 436-460, DOI: 10.1086/591837.

Marleau, G.-D. and A. Cumming

2014 "Constraining the initial entropy of directly detected exoplanets", *Monthly Notices of the Royal Astronomical Society*, 437, pp. 1378-1399, DOI: 10.1093/mnras/stt1967.

Marois, C., D. Lafreniere, R. Doyon, B. Macintosh, and D. Nadeau

2006 "Angular Differential Imaging: A Powerful High-Contrast Imaging Technique", *The Astrophysical Journal*, 641, pp. 556-564, DOI: 10.1086/500401.

Martinez, P., M. Kasper, A. Costille, J. F. Sauvage, K. Dohlen, P. Puget, and J. L. Beuzit

2013 "Speckle temporal stability in XAO coronagraphic images. II. Refine models for quasi-static speckle temporal evolution for VLT/-SPHERE", *Astronomy & Astrophysics*, 544, A41, DOI: 10.1051/0004-6361/201220820.

Mawet, D., P. Riaud, O. Absil, and J. Surdei

2005 "Annular Groove Phase Mask Coronagraph", *The Astrophysical Journal*, 633, pp. 1191-1200, DOI: 10.1086/462409.

Mawet, D., P. Riaud, J. Baudrand, P. Baudoz, A. Boccaletti, O. Dupuis, and D. Rouan

2006 "The four-quadrant phase-mask coronagraph: white light laboratory results with an achromatic device", *Astronomy and Astrophysics*, 448, pp. 801-808, DOI: 10.1051/0004-6361:20054158.

Mesa, D., R. Gratton, A. Berton, J. Antichi, C. Verinaud, A. Boccaletti, M. Kasper, R. U. Claudi, S. Desidera, E. Giro, J.-L. Beuzit, K. Dohlen, M. Feldt, and D. Mouillet

2011 "Simulation of planet detection with the SPHERE IFS", *eprint arXiv:1103.4766*.

Mugnier, L. M., J.-F. Sauvage, T. Fusco, A. Cornia, and S. Dandy

2008 "On-Line Long-Exposure Phase Diversity: a Powerful Tool for Sensing Quasi-Static Aberrations of Extreme Adaptive Optics Imaging Systems", *Optics Express*, 16, p. 18406, DOI: 10.1364/OE.16.018406.

Pedani, M.

2014 "Sky surface brightness at Mount Graham II. First JHKs science observations with Large Binocular Telescope", *New Astronomy*, 28, pp. 63-69, DOI: 10.1016/j.newast.2013.10.005.

Pedichini, F., M. Stangalini, A. Ambrosino, A. Puglisi, E. Pinna, V. Bailey, L. Carbonaro, M. Centrone, J. Christou, S. Esposito, J. Farinato, F. Fiore, E. Giallongo, J. M. Hill, P. M. Hinz, and L. Sabatini

2017 "High Contrast Imaging in the Visible: First Experimental Results at the Large Binocular Telescope", *The Astronomical Journal*, 154, 2, p. 5, DOI: 10.3847/1538-3881/aa7ff3.

Pinna, E., S. Esposito, P. Hinz, G. Agapito, M. Bonaglia, A. Puglisi, M. Xompero, A. Riccardi, R. Briguglio, C. Arcidiacono, L. Carbonaro, L. Fini, M. Montoya, and O. Durney

2016 "SOUL: the Single conjugated adaptive Optics Upgrade for LBT", *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 9909, 99093V, DOI: 10.1117/12.2234444.

Quanz, S., A. Amara, M. Meyer, J. Girard, M. Kenworthy, and M. Kasper

2015 "Confirmation and Characterization of the Protoplanet HD 100546 b-Direct Evidence for Gas Giant Planet Formation at 50 AU", *The Astrophysical Journal*, 807, 64, DOI: 10.1088/0004-637X/807/1/64.

Racine, R., Gordon A. H. Walker, D. Nadeau, R. Doyon, and C. Marois

1999 "Speckle Noise and the Detection of Faint Companions", *The Publications of the Astronomical Society of the Pacific*, 111, pp. 587-594, DOI: 10.1086/316367.

Riaud, P., A. Boccaletti, J. Baudrand, and D. Rouan

2003 "The Four-Quadrant Phase-Mask Coronagraph. III. Laboratory Performance", *The Publications of the Astronomical Society of the Pacific*, 115, pp. 712-719, DOI: 10.1086/375385.

Rouan, D., P. Riaud, A. Boccaletti, Y. Clnet, and A. Labeyrie

2000 "The Four-Quadrant Phase-Mask Coronagraph. I. Principle", *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 112, pp. 1479-1486, DOI: 10.1086/317707.

Sauvage, J.-F., L. M. Mugnier, B. Paul, and R. Villecroze

2012 "Coronagraphic phase diversity: a simple focal plane sensor for high-contrast imaging", *Optics Letters*, 37, pp. 4808-4810, DOI: 10.1364/OE.16.018406.

Savransky, D., S. J. Thomas, L. A. Poyneer, J. Dunn, B. A. Macintosh, N. Sadanuki, D. Dillon, S. J. Goodsell, M. Hartung, P. Hibon, F. Rantakyro, A. Cardwell, and A. Serio

2014 "Automated alignment and on-sky performance of the Gemini planet imager coronagraph", *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 9147, 914740, DOI: 10.1117/12.2055752.

Schlieder, J. E., S. Lepine, and M. Simon

2012 "Likely Members of the β Pictoris and AB Doradus Moving Groups in the North", *The Astronomical Journal*, 144, 109, DOI: 10.1088/0004-6256/144/4/109.

Serabyn, E., E. Huby, K. Matthews, D. Mawet, O. Absil, B. Femenia, P. Wizinowich, M. Karlsson, M. Bottom, R. Campbell, B. Carlomagno, D. Defrre, C. Delacroix, P. Forsberg, C. Gomez Gonzalez, S. Habraken, A. Jolivet, K. Liewer, S. Lilley, P. Piron, M. Reggiani, J. Surdej, H. Tran, E. Vargas Cataln, and O. Wertz

2017 "The W. M. Keck Observatory Infrared Vortex Coronagraph and a First Image of HIP 79124 B", *The Astronomical Journal*, 153, DOI: 10.3847/1538-3881/153/1/43.

Snik, F., G. Otten, M. Kenworthy, M. Miskiewicz, M. Escuti, C. Packham, and J. Codona

2012 "The vector-APP: a broadband apodizing phase plate that yields complementary PSFs", *The Publications of the Astronomical Society of the Pacific*, 8450, 84500M, p. 11, DOI: 10.1117/12.926222.

Sozzetti, A., M. Bonavita, S. Desidera, R. Gratton, and M. G. Lattanzi

2016 "Gaia: The Astrometry Revolution", *Young Stars & Planets Near the Sun, Proceedings of the International Astronomical Union*, 314, pp. 264-269, DOI: 10.1017/S1743921315006535.

Stephens, D. C.

2003 "The Classification of L Dwarfs: Is It Based on Clouds or Temperature?", *Brown Dwarfs, Proceedings of IAU Symposium # 211. Edited by Eduardo Martin, Astronomical Society of the Pacific, 2003, p.355*.

Wagner, K., D. Apai, M. Kasper, and M. Robberto

2015 "Discovery of a Two-armed Spiral Structure in the Gapped Disk around Herbig Ae Star HD 100453", *The Astrophysical Journal Letters*, 813, L2, DOI: 10.1088/2041-8205/813/1/L2.

ACKNOWLEDGMENTS

Firstly, I would like to express my enormous gratitude to my supervisor Dr. Jacopo Farinato. I thank you for having intoduced me in the field of astronomical instrumentation (totally new for me) with patience and enthusiasm, for your continuous support during all my Ph.D. and for having helped me day by day to acquire the knowledge and instruments necessary to give my (small) personal contribution to the project.

Very special thanks of course to everyone in my group. Thanks for welcoming me in such a fantastic reality and for driving me with teachings, support and expertise. Thanks for the unique opportunity I had to travel with you around the world and to live outstanding experiences. I could not imagine to find better people, both professionally than humanly.

Many thanks to all the people from other institutions I have been in touch during these three years. In particular, I am grateful to Marco Stangalini in Rome for enlightening me at very beginning of my Ph.D. on how to start writing my simulation code and for his constant support to my work and to the project through years. I am particularly grateful to Alexis Carlotti and Christophe Verinaud from IPAG-Grenoble for sharing with me their great knowledge and expertise in the coronagraphic field.

I thank of course the reviewers of this work, Markus Kasper and Damien Gratadour, for their time and insightful comments that helped improving the quality of the manuscript. It is a matter of pride to me that astronomers of such a level appreciated my work.

I thank the Ph.D. school coordinator, prof. Giampaolo Piotto. The activities offered by the school allowed me to gain knowledge in many fields and to live very beautiful experiences.

Last but not the least, I would like to thank my family, my friends and my girlfriend. You have been fundamental in helping me living this experience far from home with serenity and positive state of mind, always offering me moments to relax and have fun.

Padova, ottobre 2017