OWL Instrument Concept Study
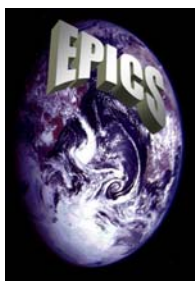
EPICS

Earth-like Planets Imaging Camera Spectrograph

Norbert Hubin (ESO): Project Manager

Markus Kasper (ESO): Instrument Scientist

Christophe Vérinaud (ESO): System Scientist

EUROPEAN SOUTHERN OBSERVATORY

Organisation Européenne pour des Recherches Astronomiques dans l'Hémisphère Austral
Europäische Organisation für astronomische Forschung in der südlichen Hemisphäre

Change record

Issue	Date	Section / Paragraph affected	Reason / Initiation Documents / Remarks		
1.0	20.10.05	All	First Issue of the Document	Author(s)	C. Vérinaud et al.
				Approval	N. Hubin
				Author(s)	
				Approval	
				Author(s)	
				Approval	
				Author(s)	
				Approval	
				Author(s)	
				Approval	

1 List of contributors

1.1 Science case working group

Name	Institute
Markus Kasper (Chair)	ESO
Florian Kerber (Co-chair)	ESO/STECF
Farid Rahoui	ESO
Hans Martin Schmid	ETHZ
Daphne Stam	Astronomical Institute, University of Amsterdam
Raffaele Gratton	INAF, PADOVA
David Mouillet	LAOM
Massimo Turatto	INAF, PADOVA
Franck Selsis	ENS Lyon
Markus Feldt	MPIA
Jean-Luc Beuzit	LAOG
Jimmy Paillet	ENS LYON

1.2 Adaptive Optics working group

Name	Institute
Christophe Verinaud (Chair)	ESO
Norbert Hubin (Co-chair)	ESO
Thierry Fusco	ONERA
Carlos Correia Da Silva	ESO
Enrico Fedrigo	ESO
Miska Le Louarn	ESO

1.3 Coronagraphy working group

Name	Institute
Natalia Yaitskova (Chair)	ESO
Anthony Boccaletti (Co-chair)	LESIA
Remi Soummer	American Museum of Natural History
Pierre Baudoz	LESIA

1.4 Differential Imager working group

Name	Institute
Andy Longmore (Chair)	UKATC
Rainer Lenzen (Co-chair)	MPIA
Farid Rahoui	ESO
David Mouillet	LAOMP
Christophe Verinaud	ESO
Markus Kasper	ESO

1.5 Integral Field Spectroscopy working group

Name	Institute
Jacopo Antichi (Chair)	INAF/ Padova
Kjetil Dohlen (Co-chair)	OAMP/LAM
Etienne Le Coarer	LAOG
Harald Kuntschner	ESO/STECF
Christophe Vérinaud	ESO
Chalabaev Almas	LAOG
Patrick Rabou	LAOG
Turatto Massimo	INAF/ Padova
Raffaele Gratton	INAF/ Padova

1.6 Differential Polarimetry working group

Name	Institute
Hans Martin Schmid (Chair)	ETHZ
Jaap Tinbergen (Co-chair)	ASTRON
Daphne Stam	Astronomical Institute, University of Amsterdam
Ramon Navarro	ASTRON
Christian Thalmann	ETHZ
Achim Gandorfer	MPS

1.7 Acknowledgements

1.8 List of Institutes

- Istituto Nazionale di AstroFisica (INAF), Astronomical Observatory of Padova, Italy

- Swiss Federal Institute of Technology of Zurich (ETHZ) , Switzerland

- Astronomical Institute, University of Amsterdam, Netherlands

- Laboratoire d'Astrophysique de l'Observatoire de Midi-Pyrénées (LAOMP), France

- Ecole Normale Supérieure de Lyon (ENS), France

- Max-Planck Institut fur Astronomie, Heidelberg (MPIA), Germany

- Laboratoire d'AstrOphysique de Grenoble (LAOG), France

- Office National d'Etudes et de Recherches Aérospatiales (ONERA), France

- United Kingdom Astronomy Technology Centre (UKATC), UK

- Observatoire Astronomique de Marseille Provence, Laboratoire d'Astrophysique de Marseille (OAMP/LAM), France

- Observatoire de Paris, Laboratoire d'Etudes Spatiales et d'Instrumentation en Astrophysique (LESIA), France

- American Museum of Natural History, USA

- Netherlands Foundation for Research in Astronomy (ASTRON), Netherlands

- Max-Planck Institut fur Sonnensystemforschung (MPS), Germany

Contents

Index of Figures

Index of Tables

Abbreviations

AD   Applicable Document
AO   Adaptive Optics
ICD   Interface Control Document
FOV   Field of View
GLAO Ground Layer Adaptive Optics
MCAO Multi-Conjugate Adaptive Optics
PSF   Point Spread Function
TBC   to be confirmed
TBD   to be determined
XAO   eXtreme Adaptive Optics
SH   Shack-Hartmann
PSD   Power Spectral Density
PSF   Point Spread Function
SNR   Signal to Noise Ratio
ADC   Atmospheric Dispersion Corrector
SR   Strehl Ratio
BS   Beam Splitter
RMS Root Mean Square
VLT-PF Very Large Telescope Planet Finder
DM   Deformable Mirror
WFS   Wave-Front Sensor
FWHM Full width at Half Maximum
IFS Integral Field Spectrograph
DI   Differential Imager
FOV Field-Of-View
IR Infrared
NIR Near Infrared
VIS Visible
VLT Very Large Telescope
RTC   Real Time Computer
TLR Top-Level Requirement
PTV Peak To Valley
MEMS Micro-Electro-Mechanical System
FPGA Field Programmable Gate Array

Symbols

r_0 atmospheric turbulence coherence length
τ_0 atmospheric turbulence coherence time
L_0 atmospheric turbulence outer scale
C_n^2 integrated atmospheric turbulence profile

2 Scope

The scope of the document is to present the conceptual study of the EPICS instrument including Adaptive Optics, Coronagraphy, Differential Imaging, Integral Field Spectroscopy and Differential Polarimetry.

3 References and applicable documents

3.1 Reference documents

No	Title	Doc. Number	Issue
RD 1	ZIMPOL conceptual design	CHEOPS-TRE-ETH-00039	1
RD 2	CHEOPS Polarimetry: System Aspects and Optional Upgrades	CHEOPS-SRE-EXT-00057	1
RD 3	ZIMPOL detection Simulations	CHEOPS-TRE-ETH-00035	1
RD 4	Testing ZIMPOL for CHEOPS	CHEOPS-TRE-ETH-00030	1
RD 5	VLT-PF CONCEPT FEASIBILITY STUDY, Adaptive Optics Analysis Report	VLT-TRE-ONE-14690-0023	1
RD 6	VLT-PF CONCEPT FEASIBILITY STUDY, System Analysis Report	VLT-TRE-LAM-14690-0008	1
RD 7	VLT-PF CONCEPT FEASIBILITY STUDY, Conceptual design	VLT-TRE-LAM-14690-0005	1
RD 8	VLT-PF CONCEPT FEASIBILITY STUDY, Science Test Cases Report	VLT-TRE-LES-14690-0002	1

3.2 Applicable documents

The following applicable document of the exact issue shown forms a part of the present document to the extent specified herein. In the event of conflict between the document and the content of the present specification, the latter shall be taken as superseding.

No	Title	Doc. Number	Issue
AD 1	OWL Telescope-Instruments Interface Control Documents	OWL-ICD-ESO-00000-0139	1
AD 2	OWL concept design report, phase A design review	OWL-TRE-ESO-0000-0001	1

AD 3	SPARTA for OWL, Standard Platform for Adaptive Optics Real Time Applications Straw-man Design	OWL-CSR-ESO-00000-0178	1

4 Executive Summary

4.1 Introduction

The study of **EPICS** was carried out at ESO with the support of scientists and engineers from different European Institutes that are in parallel involved in the VLT Planet Finder project. It was started in april 2005, after the completion of the VLT Planet Finder phase A . The EPICS concept has been naturally biased but also inspired by the VLT Planet Finder feasibility studies made by the two external European Consortia: the VLT PF led by LAOG and the CHEOPS instrument led by MPIA. Those studies have demonstrated that it is necessary to combine an "extreme" adaptive optics system (hereafter XAO) with other methods (coronagraphy and differential detection) to reach the contrast permitting exoplanets detection. The need to understand the interaction and to control the error sources from the different sub-systems calls for a global system approach in the definition and in the evaluation of the performance of EPICS.

Several working groups (see contributors list in section 1) have been defined in order to develop in parallel the studies in the many areas necessary for exo-planet detection and characterization. The general organization of EPICS is represented in Figure 4-1.
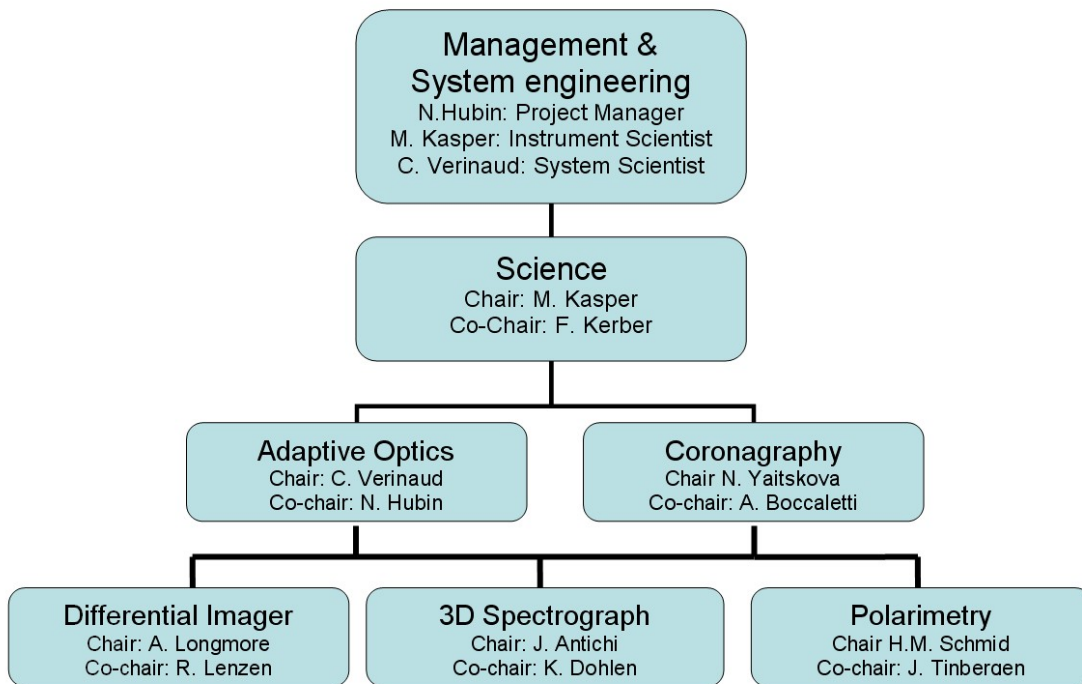
Figure 4-1: EPICS organization.

In the course of this first phase of the project a concept for an Earth-like Planets Imaging Camera Spectrograph (EPICS) has been defined, addressing the main functions required to fulfill the science goals. However, the short amount of time available for this preliminary study did not permit to design an integrated opto-mechanical set-up, which reflects also the fact that many issues need still to be solved in most of the areas before making a final decision on the overall concept. Still, important sub-systems have been deeply investigated by the different groups, involving detailed modeling and sub-system designs. The results of these studies are presented in the various sections of this document.

An SNR analysis of the performance for exo-planets detection with EPICS has been done and demonstrates the exciting potential of a 100-m class telescope for discovering Earth-like planets and possibly bio-markers.

The roadmap of the success of this instrument will need to overcome challenging technological limitations, but since EPICS is expected to be installed on OWL several years after the full capability of the telescope has been reached, important results from Research and Developments is expected in all the areas that are listed in section 4.9.

4.2 Science case

- **Primary science goal: the detection of Earth-like planets**

One of the most ambitious science objectives of OWL is the detection and characterization of extra-solar systems in an advanced evolutionary stage, for a statistically meaningful sample of stars. Rocky planets with possibly Earth-like features is the ultimate and most challenging goal of EPICS. The direct detection of exo-planets is made very difficult by the very high relative flux ratio from the star and planets orbiting it and their small angular separation. Figure 4-2 illustrates the requirements in contrast for different types of planets as a function of angular separation. Ultimately the primary science goal of EPICS requires the detection of faint point sources in proximity of a bright star with an object-star contrast down to about 2.10^{-10} at 0.05 arcsec from the star. Moreover, to observe a planet and to characterize its atmosphere, EPICS must be sensitive at the wavelengths of H_2O, CO_2, CH_4 and O_2 molecular absorption lines.

- **Gas giant planets in a late evolutionary stage**

EPICS will also permit a significant breakthrough in the detection and characterization of cold gas giant planets. The better contrast (the contrast of Jupiter at 5 AU is 10^{-9}) and larger separation, permits an easier detection, and opens the door to high resolution spectroscopy. In particular, radial velocity measurements and the analysis of atmospheric composition and dynamics of close-in giant planets will be possible. The contrast between a Jupiter mass planet at 0.5 AU and its star is around 10^{-7}, so roughly corresponding to the stellar AO residuals. For 10 pc distance from Earth, assuming a G2 star, its magnitude would be around 22.5 and the photon flux at resolution of 50.000 would be about 0.5 photons per second and spectral bin (16% overall quantum efficiency). Therefore, a reasonably high SNR for the high resolution spectroscopy appears feasible in observing times of a couple of hours.

Figure 4-2. Contrast vs. angular separation for different types of planets. (Courtesy O. Lardiere).

- **Targets**

Considering stars visible from a low latitude site in southern hemisphere (~ Paranal) and the zenith angle is restricted to <30 degrees we computed a selection of targets of the most interesting spectral types G, K and M. In order to have access to about 100 stars for each spectral type, one has to observe out to:

- 25 pc for G-stars ($m_V \approx 7$)

- 20 pc for K-stars ($m_V \approx 8.5$)

- 15 pc for M-stars ($m_V \approx 9 – 16$ for M0 to M5). There are about 50 M-stars with $m_V < 10$ and 100 M-stars with $m_I < 10$ observable at low latitudes.

This sets the Top-Level Requirements in terms of contrast/separation and of limiting magnitude for the AO system:

- $2. \, 10^{-10}$ at 40 mas: G star at 25 pc, Mv = 7

- $8. \, 10^{-10}$ at 25 mas: K star at 20 pc, Mv = 8

- $8. \, 10^{-9}$ at 15 mas: M star at 15 pc, Mv = 10

4.3 Concept

The main ingredients of the VLT-PF instrument proposal are present in the EPICS concept. This permits to EPICS to be sensitive to a wide range of exo-planets spectro-polarimetric physical properties from the visible to NIR, thanks to three different instruments based on

34

differential imaging methods. A choice we made at an early stage of the study, is to allow these instruments to work simultaneously for maximizing the scientific return.
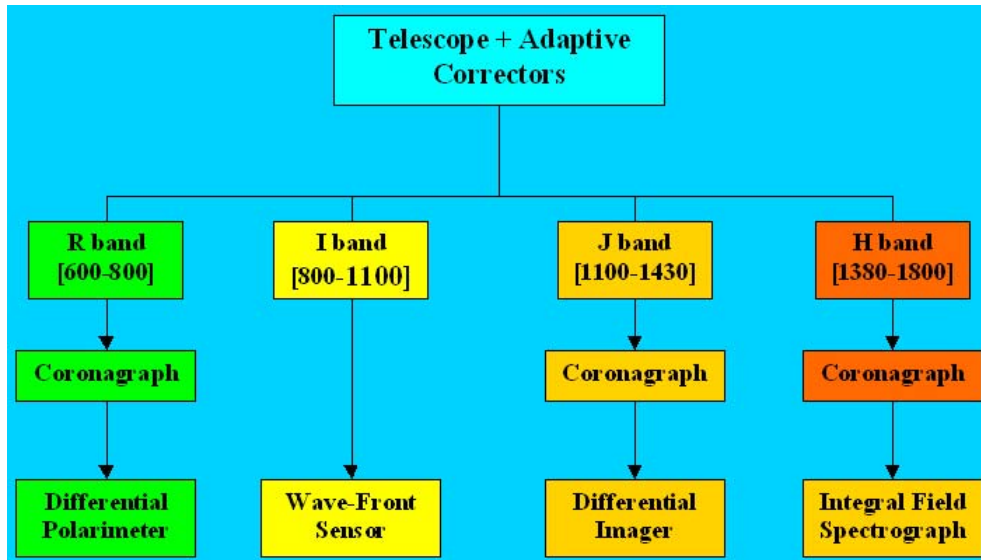
Figure 4-3: EPICS concept. EPICS will be composed of three spectral channels for the scientific instruments and one for wave-front sensing.

A total of four spectral channels (3 for science and one for wave-front sensing) are defined. Each scientific channel will be equipped with its own coronagraph.

- The **R band** is dedicated to the Polarimetric Differential Imager for detecting rocky planets and to the follow-up observations for the detection of O_2.

- The **J band** will be equipped with a differential imager using pairs of filters that will be sensitive to both CH_4 and H_2O absorption bands.

- The **H band** will be equipped with an Integral Field Spectrograph. The main features that can be detected in this band are CH_4 and CO_2.

- The **I band** is reserved for wave-front sensing. This band has been chosen because of the lesser scientific interest for planet detection. Its location, spectrally speaking, between the visible and NIR instruments, is optimal with respect to important atmospheric chromatic limitations for XAO on ELTs. Moreover no light is taken from the scientific channels.

The EPICS concept is summarized in Figure 4-3.

4.4 Adaptive optics

Adaptive optics is one of the major components of EPICS, and the success of the instrument depends closely on the performance of the real-time correction of atmospheric turbulence. For EPICS, the AO system has been optimized, in the NIR, to provide a high efficiency for the rejection of the starlight halo for angular separations ranging from 30 mas to 200 mas, corresponding to the expected locations of Rocky planets in the habitable zone around stars of the three spectral types defined. To achieve this, the proposed concept is based on a double stage: a first system is an extrapolation of the VLT-PF concept with a classical Shack-Hartmann (500x500 sub-apertures) controlling a 1.7×10^5 actuators system at 1 kHz. This system is not enough to provide a sufficient rejection of starlight for the most challenging science goal. A second system in cascade with the first one, is based on $1.5 \ 10^4$ actuators controlled at 3 kHz with a Pyramid sensor (150x150 sub-apertures). Thanks to the high speed and to low noise propagation for low spatial frequencies in the pyramid sensor measurements, more than an order of magnitude of halo rejection is obtained in the central part of the field of view (see Figure 4-4). This translates in a net gain of more than a factor of 10 in integration time for the detection of Rocky planets in the habitable zone with respect to the single stage only.

Note that the main reason for this double stage scheme is a computing power issue. Fast reconstruction algorithms exist for a Shack-Hartmann WFS but not yet for a pyramid sensor. A 500x500 pyramid sensor at 3 KHz would have been ideal in terms of noise propagation, but the only way to reconstruct the command, at least today, through classical matrix-vector multiplication, would have required an increase of more than a factor of 200 of computing power with respect to the 150x150 system for which we already took into account that the available computing power is increased by a factor of about 50 with respect to the SPARTA OWL platform (**AD 3**).
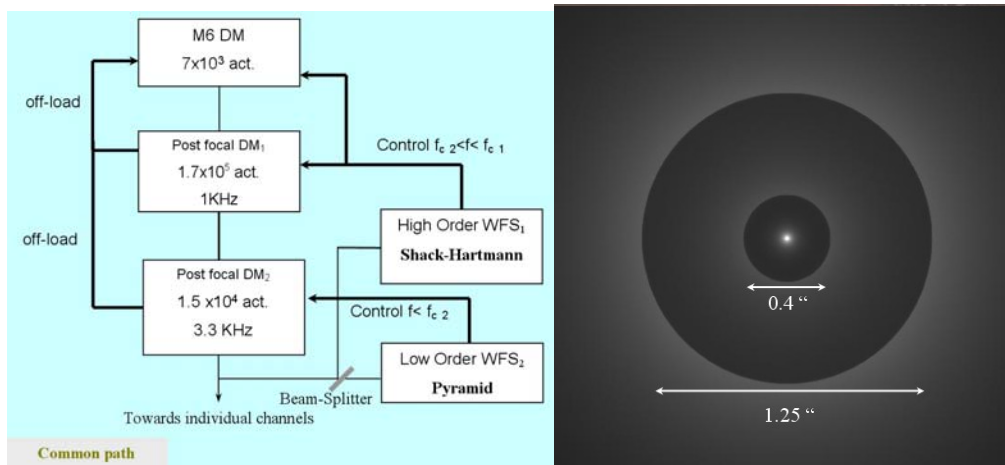
Figure 4-4: EPICS double stage AO system concept (left). Example of simulated coronagraphic image expected with this system (perfect coronagraph) (right). Wave-length: 1220 nm. Seeing: 0.5 arcsec, G2 star at 10 pc (M_v=5.0). The image has been scaled for better rendering. The large outer corrected field (due to 1st stage correction) is 1.25 arcsec in diameter. The inner corrected field (due to 2nd stage correction) is 0.38 arcsec in diameter.

4.5 Coronagraphy and high contrast

Coronagraphy is another major component of EPICS. It is foreseen to have a total of 3 coronagraphs, each optimized for a given scientific channel. Extremely high performance is required to remove the diffraction effects due to OWL segmented pupil. No concept has yet been selected but at least two options are being further studied:

- Prolate Apodized Lyot coronagraph

- Gauss-Lyot reticulated coronagraph

Both are Lyot coronagraphs that are well suited for an ELT, if more than $10\lambda/D$ separation angle is expected. A major complication is the problem of segmentation that needs to be carefully treated if high performance is expected.
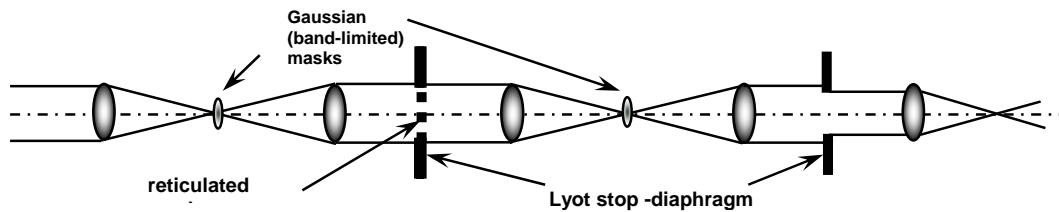
Figure 4-5: Double stage Gauss-Lyot reticulated coronagraph.

A double stage Gauss-Lyot coronagraph with complex reticulated stops (in amplitude and phase) has been studied and shows a theoretically extremely high extinction.
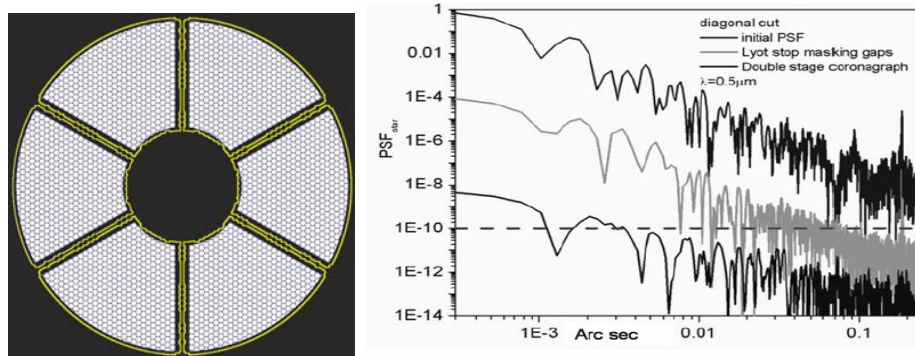
Figure 4-6: Gauss-Lyot reticulated coronagraph. Left: Reticulated Lyot stop corresponding to the primary mirror. Right: performance at 500 nm on OWL pupil without phase errors.

The very fine details of the masks will make them very challenging to manufacture and to align continuously during the observation as the telescope moves. The achromatization of the coronagraph is also an issue that remains to be solved, although several new ideas proposed recently seem very promising.

Important point: There is often a mis-conception that the performance requirement of a coronagraph should be tuned according to the actual contrast that the AO system coupled to

it should deliver. In other words, there would be no need to build an extremely powerful coronagraph, if the AO system doesn't permit to reach the intrinsic coronagraph contrast. **This is actually not true if differential imaging is considered.** A preliminary system analysis has been made (section 7) in order to define the requirements of some key elements in order to attain the goal contrast for Earth-like planets detection. This analysis was made without considering special calibration strategies that could alleviate somehow the requirements. These **requirements** can thus be considered **as a goal** with some possible margin which will depend on the performance of special calibration strategies one can use to further increase the contrast (deconvolution, reference subtractions, rotation of key optical elements for static speckle smoothing, etc.)

For EPICS, the goal requirement for the coronagraphs coupled with OWL pupil (co-phasing residuals excluded because corrected by AO, see section 8.4.3), is to deliver **an instrumental contrast better than 10^{-9} for separations larger than 10 λ/D**. This requirement is essential for being able to attain the 10^{-10} ultimate contrast after dual imaging. Actually, the coronagraphic static residuals (as opposed to the atmospheric and AO residuals that average out), will mix and amplify static chromatic differential aberrations in the instrument and in the common path and limit the ultimate contrast achievable whatever the integration time.

These requirements, which are probably the most stringent for the whole instrument calls for the following more **general system requirements**:

- An intrinsic coronagraph performance of better than **10^{-9}** for the rejection of the telescope diffraction effects (segments and gaps) for **10-20 % spectral bandwidth.**

- To fulfill the first point, one needs a static error in front of the coronagraph less than **0.3 nm rms** (flat PSD) on a limited spatial frequency range: from **10 to 75 cycles/pupil**. The use of an active mirror with 10^4 actuators in front of each coronagraph is mandatory to achieve this precision through off-line calibration (see section 8.3.1.2). Experiments conducted for TPF-C (High Contrast Imaging Testbed HCIT, [35]) show that MEMS mirror with very accurate positioning capability (resolution of <0.10 Å surface/step and position stability of <0.20 Å surface/hour) are already available and permitted to obtain a halo level of 5×10^{-9} at 800 nm for 4 λ/D separation for 40 nm spectral bandpass. These results are very encouraging, and show that with an efficient coronagraph like the Gauss-Lyot reticulated (Figure 4-6), and with an efficient thermal control of only a few key elements in EPICS AO common path (e.g. WFS optics and active corrector to stabilize non common path errors), the goal of 10^{-9} instrumental contrast is reachable.

- **Reflectivity variations on the pupil of less than 0.5% per segment** in primary mirror. An amplitude corrector in front of the coronagraph is probably needed, depending however on the spatial distribution of the dispersion between the segments reflectivity due to ageing and re-coating (TBD).

- **Differential chromatic aberrations**: this complex issue is described in section 7.3. Some important system choices (use of an ADC as up-stream as possible in the optical train), permit to limit **the differential chromatic error in front of the coronagraph to less than 0.1 nm**. To achieve this, one of the most stringent

requirement is a parabolic collimating mirror of 750 mm with 3 nm rms precision (wave) and an ADC of same size (10 nm rms precision), see section 7.3.4 for detailed analysis. These specifications are certainly stringent but not impossible to achieve.

Again, all these specifications can be considered as a goal, the use of smart strategies for calibration could probably alleviate them, but a more in-deep study using end-to-end models is absolutely necessary to address the problem.

4.6 Instruments

Specifications for the three instruments were derived from the Top-Level Requirements and an analysis of the main issues was made. Some preliminary sub-systems design has also been produced in some cases. Note that, except for the larger field available for the differential imager, the latter and the integral field spectrograph could at the end be exchangeable or form a unique instrument. The parallel development of both concepts reflects the situation of VLT-PF studies where special investigations are pursued on the question of differential aberrations in a TIGRE-type IFS. An IFS is theoretically free of differential aberrations but some diffraction effects could severely limit this capability unless extreme over-dimensioning is done or new concepts developed. Most of the answers to this problem should be expected from the VLT-PF studies and will thus directly benefit to EPICS.

4.6.1 Wave-length splitting Differential Imager

A dichroic based differential imager with 4 simultaneous channels and high throughput (> 90 %) has been designed (see Figure 4-7). The choice of the wave-lengths for the filters permits to be sensitive to both CH_4 and H_2O.

Filter	Central wavelength	FWHM	$\lambda/\Delta\lambda$
1 ('on-line')	1140 nm	75 nm	~ 15
2	1220 nm	80 nm	~ 15
3	1300 nm	85 nm	~ 15
4 ('on-line')	1385 nm	90 nm	~ 15
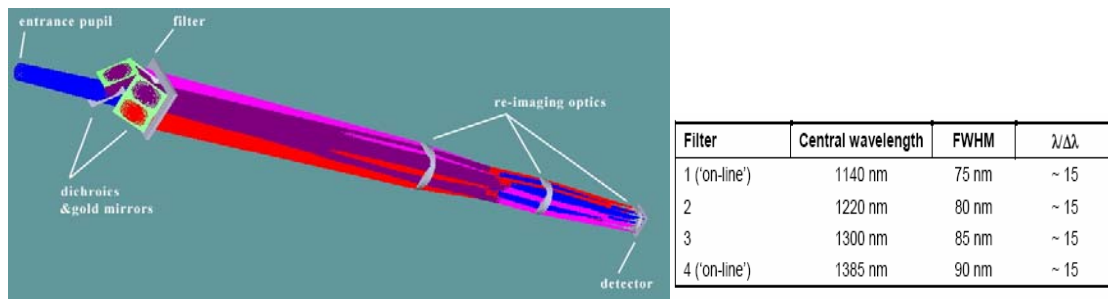
Figure 4-7: Dichroic-based 4 channels differential imager.

Extremely tight specifications are required for the optical quality of the dichroics and of the filters (~ 1nm) for the detection of Rocky planets, and could be somewhat relaxed for the detection of Jupiter-like planets. An 8Kx8K detector is needed to cover the 4 arcsec field-of-view.

4.6.2 Integral Field Spectroscopy

Two concepts are being studied in parallel:

- A TIGRE-type Integral Field Spectrograph

- A Fourier Transform Integral Field Spectrograph
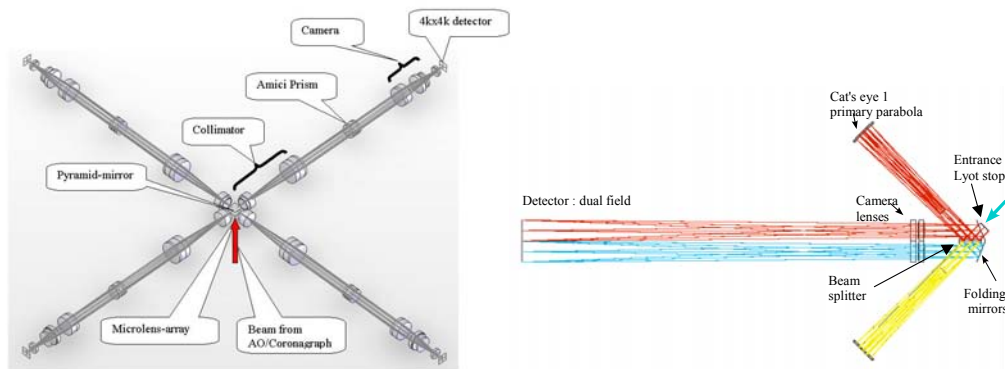
Figure 4-8: The two possible concepts for EPICS integral field Spectrograph. Left: TIGRE-type IFS. Right: Fourier Transform spectrograph.

As said in the preceding section, a TIGRE-type IFS is being considered for VLT-PF. In particular important limitations due to diffraction effects are currently being studied. To overcome some of the effects, significant over-dimensioning of the detectors is needed. In the proposed TIGRE-type concept for EPICS, to cover 2x2 arcsec field in H band with a resolution R=15, 4 detectors with 4K4K pixels are needed (see possible implementation in Figure 4-8).

A Fourier transform spectrograph on the other hand, is much less hungry than a TIGRE-type IFS. 4x4 arcsec could be covered with a single 4Kx4K detector, permitting a very compact design. Moreover, better control of systematic errors is expected since no diffraction effects occur in the FTS. Another nice feature, which still needs to be investigated, is that an FTS is naturally free from aliasing so that for sufficiently large separations in terms of Airy disks, as it will be the case for an ELT, the speckles become unresolved. Temporal variations of speckle intensity can however be an additional issue.

The pros and cons of these two spectrographs need to be carefully analyzed in the future.

4.6.3 Differential polarimeter

A differential polarimeter concept based on the CHEOPS proposal for VLT-PF has been studied for EPICS. Different modes of operation can be used to detect the polarization signal of an exo-planet against the starlight halo. For extremely high contrast application, very precise calibration can be obtained by inserting a polarimetric switch upstream in the optical train.

Figure 4-9: Differential polarimeter concept for EPICS.

The current plan to include a high precision imaging polarimeter in the VLT planet finder instrument is an ideal situation for the implementation of such a device in the EPICS instrument. If this project is realized then much experience will be gained during the development and with real on-sky observations in combination with an extreme AO system.

4.7 Expected performance

Before presenting the results, let us make a simple exercise by comparing VLT-PF and EPICS.

The goal contrast of VLT-PF is 5×10^{-7} at 500 mas. For EPICS, the ultimate goal is 2×10^{-10} at 50 mas. The main elements that can permit to make this quantum jump in performance are listed below:

- The 100-m diameter: the collecting area of OWL is 140 times larger than the VLT 8-m telescopes. Hence, the instantaneous contrast in the halo, that has a direct impact on the integration time needed to overcome photon noise, is also enhanced by a factor of 140. The estimated instantaneous contrast for VLT-PF in good conditions is 5×10^{-5}. With the same AO system OWL would get about 3×10^{-7}. With this level of halo the integration time needed to attain 2×10^{-10} for detecting a planet at 10 pc, is of several tens of hours [33], so not very suitable for a survey for exo-planet detection.

- An additional gain related to the diameter is that the separation at a given position in the field is, in terms of λ/D (Airy disks), 12 times larger for OWL than for VLT-PF. This matter of fact means that it makes sense to compare the performance of EPICS at 50 mas separation to the one of VLT-PF at 500 mas since in the case of VLT-PF, the contrast at less than 100 mas separation is dominated by coronagraphic residuals.

- Thanks to the use of an AO system about 3 times faster and a pyramid based WFS in the second stage, an additional factor of 10 halo rejection can be obtained in the field of view of interest (30 mas – 200 mas in H), permitting integration times of a few hours instead of tens of hours for the detection of an exo-Earth at 10 pc.

- Static and differential aberrations: more than 10 times better is required for EPICS. To get this gain in systematic errors in EPICS, several system choices have been made like the use of an ADC upstream in the optical train, and a precise active corrector in front of each coronagraph.

The instantaneous contrast curves achieved with the double stage AO system at different wave-lengths are represented in Figure 4-10 for the case of a G2 star at 20 pc. Analytical models of the AO system were used (see section 8.2.3). One can see that after correction, the halo at, let's say 50 mas, is still about 1000 times brighter than the potential Earth-like planet we search. Long integration time is thus needed to increase the SNR and differential imaging is needed to overcome the speckle noise.
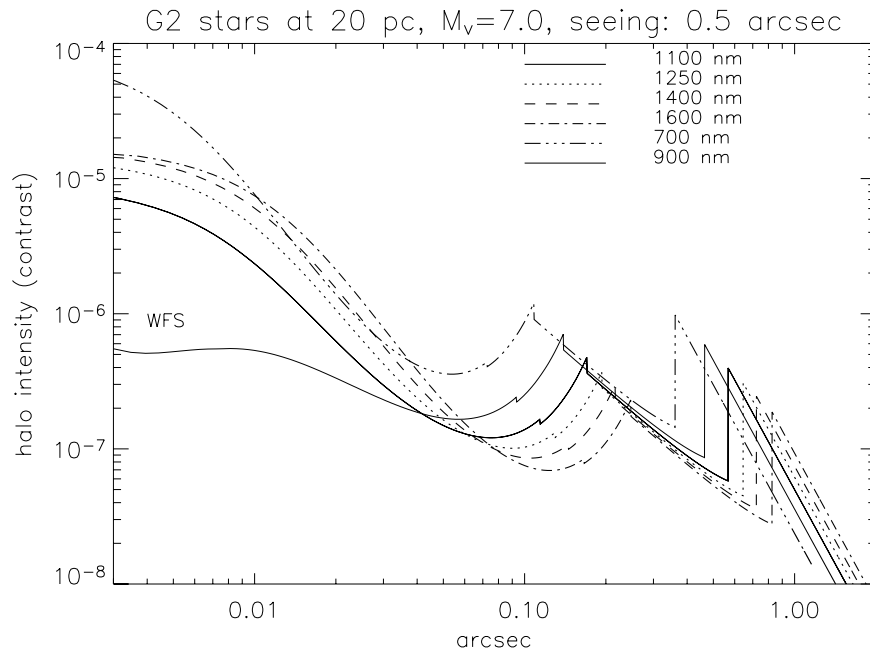
Figure 4-10: Coronagraphic PSFs for a G2 star at 20 pc in function of wave-lengths. Good seeing.

The characteristics of the test planets used in the simulation are similar to Earth-like and Jupiter-like planets at position around their host star such that their temperature is similar to the one in the solar system. (see Table 4-1 and Table 4-2). The system star-planet for each case is displaced to distances from Earth ranging from 10 to 25 pc. The adaptive optics haloes after coronagraph were computed using analytical models and SNR estimations were computed.

Star spectral type	Star-planet distance (AU)	Star-planet contrast in NIR	Angular separation at 20 pc 90 deg phase
G2	1.00	2.21×10^{-10}	50 mas
K2	0.51	8.07×10^{-10}	25 mas
M2	0.16	8.30×10^{-9}	8 mas

Table 4-1: Characteristics of Earth-like planets used in the simulation.

Spectral type	Star-planet distance (AU)	Star-planet contrast in NIR	Angular separation at 20 pc 90 deg phase
G2	5.10	1.40×10^{-9}	250 mas
K2	1.67	5.32×10^{-9}	80 mas
M2	0.83	5.50×10^{-8}	40 mas

Table 4-2: Characteristics of Jupiter-like planets used in the simulation.

Some characteristic performance curves are given here. They show that H_2O and CO_2 (if abundant) can be detected in Rocky planets up to 20 pc in about one or two nights of observation. O_2 is also detectable but requires several tens to hundreds of hours of observation, still possible for follow-up. The complete set of results can be found in section 13.

H_2O: The strong water bands of an Earth-like planet can be detected even if our own atmosphere absorbs similarly. Indeed, thanks to differential imaging, the starlight halo becomes a powerful calibrator of our own atmosphere absorption, so that the contrast of the exo-Earth in the water bands becomes detectable after a long enough integration time. The detection performance is given in Figure 4-11. Note that these results will depend a lot on the type of planet and its albedo, and the real contrast will depend a lot on cloud cover (in case of no clouds or with a lot of high cirrus clouds the contrast will be lower, whereas abundant low altitude clouds can produce a stronger signal.)

Figure 4-11: Time to detect an Earthlike planet orbiting a Main Sequence star at 5σ in water bands. λ=1250 nm, 80nm bandwidth, 0.44 atmospheric transmission, 0.32 contrast. t0=4ms, r0=20cm .

CO_2: The detection of CO_2 if abundant (10 %) is also possible in rocky planets with Earth-like size with similar integration times than for H_2O.

O_2: The detection of O_2 is much more difficult since the feature is much less contrasted, the needed spectral resolution is higher, and furthermore the adaptive optics correction is worse in the visible. The performance given in Figure 4-12 shows however that for close K and M stars, O_2 could be detected in a few tens of hours at 10 pc whereas several hundreds of hours are needed for a G star in this particular case.

Important warning: The case of G star is actually a "bad luck case" because at 10 pc, the planet separation is at 0.1 arcsec, so falling exactly on a maximum of the halo, corresponding to the second stage system (pyramid aliasing). Note that if the phase angle would be 70 deg or so instead of 90 deg the results would be significantly better for G stars by maybe a factor 2 or 3. The lack of time didn't permit to investigate all the possibilities and a better picture will be drawn in the next phase of the project.

With a more powerful AO system, optimized for correction in the visible, one could probably even gain up to a factor of 4 to 5 in integration time. This possibility has not been investigated for O_2. However, earlier results for polarization, with a more optimistic AO system, shows that an order of magnitude can be gained (see section 12.5). The same gain could apply for O_2 detection.

 Even if it is difficult to use the O_2 band to actually detect a planet, the detection of O_2 in follow-up observation of an already detected planet seems feasible. The use of an AO system with better optimization for the visible could also permit a significant gain.

Figure 4-12 Time to detect O_2 in an Earthlike planet orbiting a Main Sequence star at 5σ. λ=0.7μm, 30nm bandwidth, 0.81 atmospheric transmission, 0.79 contrast. τ_0=4ms, r_0=20cm .

CH_4: Jupiter-like planets can be detected with high signal-to noise ratio (50 σ in Figure 4-13) in less than one night, opening the door to high resolution spectroscopy. For rocky planets the signal would be too weak to be used for the detection. Follow-up observation could be foreseen to detect it in Early Earth, but has not been investigated.
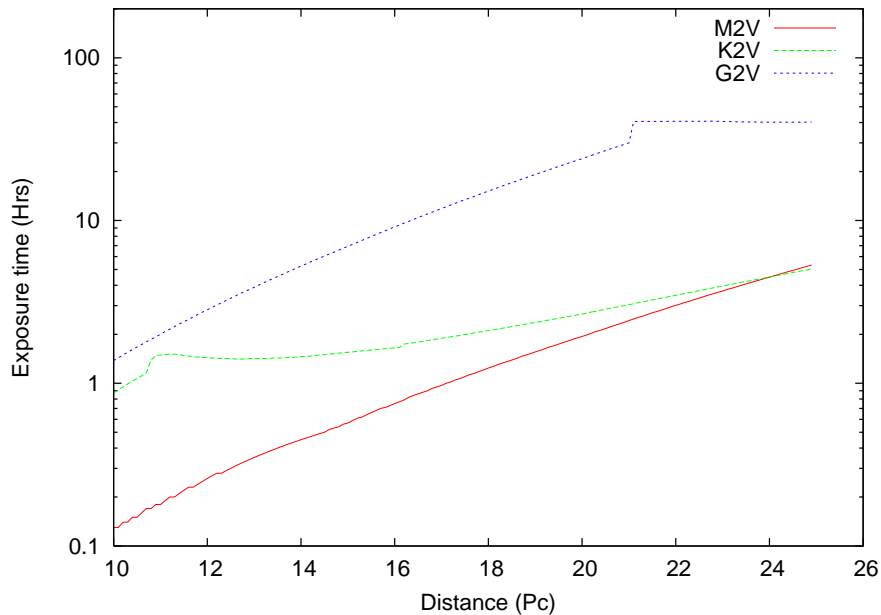
Figure 4-13: Time to detect a Jupiter like planet orbiting a Main Sequence star at 50σ (methane). λ=1.6μm, 55nm bandwidth. τ_0=4ms, r_0=20cm.

45

Polarization: For the same reasons as for O_2, since polarization detection is made in the visible, the starlight halo is much stronger than in the NIR and detection of polarization in planets around G2 star is problematic because of the particular PSF structure in the double stage system (bad luck case…). The situation is more favorable for M and K stars if close.
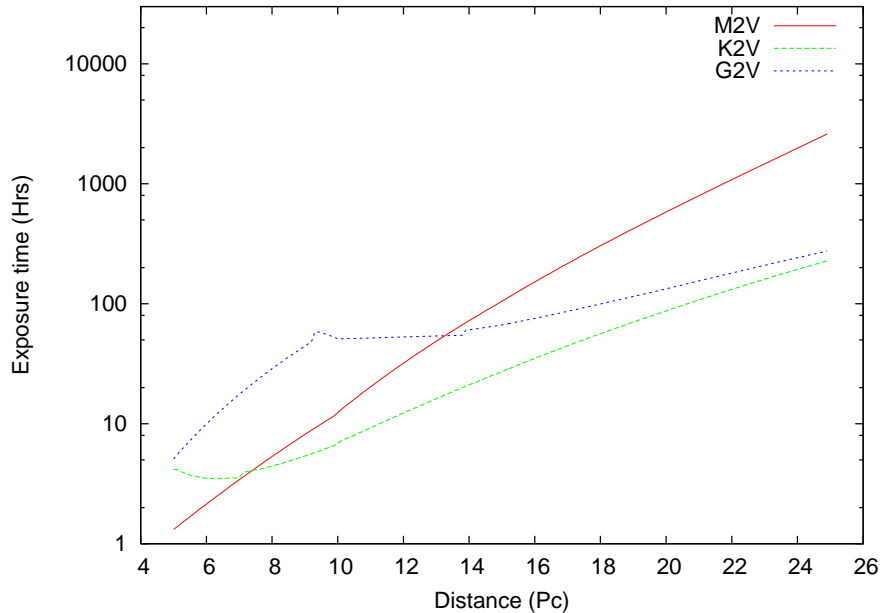
Figure 4-14. Time to detect an Earthlike planet orbiting a Main Sequence star at 5σ (Polarization 15 %). λ=0.7μm, 200nm bandwidth. τ₀=10ms, r₀=20cm (top).

Better results can be obtained with a more powerful AO system, free of the inner corrected area with its halo intensification at 0.1 arcsec. In section 12.5, polarimetric performance estimation based on an earlier AO concept (500×500 pyramid single stage system pending fast reconstruction algorithm for pyramid sensor still to be invented), an order of magnitude can be gained in integration time. The conclusions of this preliminary study are the following:

- At 5 pc an Earth (D = 1 AU) can be detected with S/N ≈ 10 and a Jupiter (D = 5 AU) with S/N ≈ 100 around a G2V star. For a low mass star of spectral type M0V an Earth like planet in the habitable zone (separation of about 0.3 AU) would also be detectable. A Jupiter like planet would be detectable with a S/N = 20 or higher.
- The distance limit for the detection of an Earth (D = 1 AU) around a G2V star in 4 hour integration time is about 10 pc.
- Jupiter sized objects around G2V stars can be detected with S/N ≥ 10 in the entire FOV (2" × 2") offered by the instrument out to distances of 20 pc (physical separations up to 40 AU).

We would like to mention that the results presented here are consistent with earlier results presented by several authors, like R. Angel [60], O. Lardiere et. al. [61], and A. Chelli [62]. In the recent paper of C. Cavarroc et al. [33], taking into account an AO system based on VLT-PF concept, the integration times found are about an order of magnitude larger. The discrepancy between these results and EPICS results relies essentially on the AO system considered. For EPICS, the AO concept was optimized in order to get the highest halo rejection in the field-of view of interest (30 mas – 200 mas). This is obtained thanks to the second stage pyramid-based (see section 8.2.3.3 on noise propagation) AO system and to very low static aberrations. This permits to get a halo residual at 100 mas, down to 5×10^{-8} i.e., 10 times lower than the one assumed in [33].

4.8 Feed-back to OWL and compliance

Here are a few important comments regarding the compatibility of the instrument with OWL current design.

ICD item	Consistency	Comment
D=100 m	Y	Integration times scale as D^{-4} a 60 m telescope would have its efficiency reduced by a factor of ~8.
F/6	Y	Was not really a problem
Segmentation	(N)	- (segment size should be larger by factor 3…) - Probably difficult to avoid but current segment size is unfortunately such that the corresponding spatial frequencies correspond to the field of view of interest for Earth-like planets search in NIR. - A major concern is the segments reflectivity dispersion (PSD to be studied) due to ageing and the capability of the coronagraph to deal with segmentation - Mitigation: amplitude corrector
		Extremely stable instrument

No gravitational stability	N	is required. Complex metrology needed if not gravity stable. If a gravity stable platform could be provided, dedicated ADC insertion should be possible upstream
Mirror coating	Y	High reflectivity absolutely needed
Atmospheric conditions	Y	Need very good seeing and dry site

Table 4-3: Compliance to OWL design.

4.9 Major risk areas

We concentrate here on the main risks areas that could be potential show-stoppers for the most demanding science goal.

Problem / topic	Specification	Prognosis / Mitigation
High density Deformable mirror	$1.7\ 10^5$ actuators	- Need strong interaction with manufacturers. - Could consider smaller DM, at the expense of loss of Strehl, but with optimized halo rejection (to be studied)
Computing power	1.7×10^5 at 1 KHz 10^4 at KHz	- close interation with manufacturers
Coronagraph Complex Lyot stops for segmentation effects rejection	Less than 10^{-9} contrast in J at 10 λ/D (intrinsic performance) for > 10 % spectral bandwidth	- This performance is needed because of systematic errors issue in differential imaging concept. - very accurate calibration can probably alleviate the requirements - New ideas such as focal plane interferometer as science detector are much less sensitive to systematic errors and could alleviate this requirement

Tight error budget	Some key optical elements need nanometric or sub-nanometric precision and/or stability	- Gravity invariant platform - Again requirements would be alleviated with new concepts (plane interferometer) - smart calibration strategies, deconvolution, - rotation of optical elements during observation… - precise thermal control of key opto-mechanical components
ADC Validity of models	ADC needed as upstream as possible to avoid chromatic beams shifts	- Adaptive ADC with image sensors for real-time correction ? - cascaded ADCs ? - will be studied in FP6
Telescope stability Wind effects	Telescope should be atmospheric limited wrt. wind shake effects	Strong damping of vibrations Integrated model strong interaction between XAO system and telescope control
Scattering: dust, turn down edges …	Must not degrade the goal contrast	- not studied, will be studied in next phases
CCD flat fielding	Must not degrade the goal contrast	- not studied, will be studied in next phases - depends on technique (e.g. FTS less sensitive than IFS)
Pyramid WFS	- new sensor no system currently working in the diffraction limited regime - Stability issues - Diffraction effects around central obstruction, interferences between quadrants	- WFS in I band, close to linearity domain, stability should be less an issue, to be studied - large dynamic modulation should be possible (up to ± 10 λ/D on a 100 m) without loss of sensitivity for the correction in the 30 mas-200 mas: diffraction effects should be attenuated (to be studied) - HOT experiments

Focal pane WFS	- new sensor	- not studied - simulations and experiments foreseen
Differential Imager	- dichroic and filters with nanometric precision - problem of coating	- super-polishing R&D of coated optical elements foreseen
IFS- TIGRE	Cross-talk, diffraction effects must not degrade the goal contrast	- studied in frame of VLT-PF
Differential Polarimeter, stability of modulator	Differential errors must not degrade the goal contrast	- studied in frame of VLT-PF

4.10 Future development plan

4.10.1 Adaptive optics

4.10.1.1 Theoretical studies, experiments

EPICS XAO system involves several new concepts that need to be proven through numerical simulations and experiments. The ESO parallel simulation code is in constant evolution and we have several plans to increase the ESO AO simulation cluster computing power and also to port the code to super-computing facilities (collaboration with Australia through European FP6). In the frame of the Joint Research Activity 1 of OPTICON, we develop at ESO an High Order Test Bench for XAO and coronagraphy experiments.

- Pyramid wave-front sensor: even though very promising this concept is very new and only one system is actually working on the sky. The big advantage in sensitivity is at the expense of non-linearities of the measurements and sensitivity to pupil shape through diffraction effects. We will study optimization of the pyramid sensor for XAO application and test them with HOT.

- Multi-stage schemes, Woofer-Tweeter schemes. The HOT bench will be composed of two deformable mirrors, a low order curvature mirror (60 actuators) and a high order MEMS mirror (1000 actuators). We will develop and test different control algorithms involving multiple stages

- New WFSs, focal plane WFS: the extreme contrast needed for exo-planet search calls for new WFS concepts where the correction is optimized by analyzing directly focal plane images. As an extension to HOT we plan to test these new concepts, like for

example the focal plane interferometer [10], especially in the frame of high precision control of systematic errors.

- Research on Fast algorithms for Shack-Hartman and for Pyramid sensors (see section 8.5.3 for details) New algorithms to be developed in frame of FP6 WP 9600.

4.10.1.2 Key components development

EPICS XAO system requires significant hardware developments.

- Adaptive mirrors: an important risk area is the availability in the near future of MEMS adaptive mirrors with an extremely high number of actuators ($> 10^5$). In the frame of OPTICON JRA1, we are developing now a 2K actuators adaptive mirror based on MEMS technology. The next step will be the development of a 10K mirror that would make the second stage of the proposed scheme. Still an order of magnitude in number of actuators is then needed for the final goal.

- Detectors. CCDs: 1KxK (goal 3Kx3K) detectors with fast read-out (3 KHz) and low noise (read-out noise less than one electron) are required. Developments of L3CCD are already part of the OPTICON Joint Research Activity 1.

- Real-time computers (see section 8.5.3 for details), need close interation

    o Input/Output communications: need 50 Gb/s

    o Fast processing elements , FPGA.

    o Faster CPU-CPU busses

    o Faster memory

4.10.2 Coronagraphy, high contrast

- Coronagraphy. The performance of EPICS relies entirely on the best conjugation of XAO with coronagraphy so that these two sub-systems should be tested together as much as possible. This is an important goal of HOT. Several coronagraphy concepts will be studied and realized in collaboration with LESIA (Observatoire de Paris Meudon) in the frame of FP6 program and will be tested on HOT.

- Effect on contrast of scattering due to dust, segments turned down edges etc…

4.10.3 Instruments

- The experience and results of the Planet finder development phase will be extremely valuable. Important feed-back is expected from extreme Adaptive optics developments, coronagraphy differential imaging, polarimetry and integral field spectroscopy.

- Fourier Transform Integral Field Spectrograph. Is not part of VLT-PF project but revealed to have a potential niche for high contrast imaging on ELTs. R&D and prototype development strongly encouraged

- Super-polishing: Optical polishing and coating quality: a number of optical surfaces in the EPICS design need to be of extremely good quality. The effect of coating on super-polished surface is an important aspect of this topic. R&D with close interaction with manufacturers is needed. Realization of a prototype of super-polished coated optics (e.g. beam splitter or filter) strongly encouraged.

- Detectors: specific requirements to be studied during phase B.

- Polarimetry: differential errors induced by modulator, studied for VLT-PF

- IFS: cross talk and diffraction effects, studied for VLT-PF

- Investigation of new detection methods to improve efficiency: speckle elongation, speckle coherence…

4.11 Preliminary cost estimate

CAPITAL INVESTMENT, DIRECT COSTS [M€]

Item	Lowest estimate	Highest estimate	Best estimate	Inte
EPICS	18.850	28.350	21.250	
Technology development	3.250	3.850	3.550	
High contrast demonstrator for EPICS	0.600	0.800	0.700	a
Coronagraphy development	0.150	0.150	0.150	b
Feasibility study	0.150	0.150	0.150	c
1kx1K WFS CCD	1.800	2.200	2.000	d
Integral Field spectrograph study/prototype	0.200	0.200	0.200	e
High dynamic polarimeter study/prototype	0.200	0.200	0.200	f
Differential imager study and prototype	0.150	0.150	0.150	g
Real Time Computer development	0.200			h
Wavefront sensing	0.400	0.600	0.500	i
Visible Wavefront sensors (6 standard units)??	0.000	0.000	0.000	j
Postfocal WFS	0.400	0.600	0.500	k
Corrective elements	4.600	7.000	5.800	l
100kact MDM	4.000	6.000	5.000	m
Four 10kact MDMs (assuming R&D above)	0.600	1.000	0.800	n
Real Time Computer XAO	4.600	9.200	4.600	o
AO Relay opto-mechanics	0.400	0.600	0.500	p
Integral Field spectrograph instrument	2.200	2.800	2.400	q
Differential Polarimeter instrument	2.200	2.600	2.400	r
Differential imager instrument	1.200	1.700	1.500	s
Instruments maintenance facility	2.000	5.000	4.000	

a	Based on typical cost for MAD/HOT
b	Dev. two different coronagraph developments
c	typical VLT instrument feasibility study
d	Typical delta cost for a new CCD development of such complexity (OPTICON-JRA2)
e	typical VLT instrument feasibility study+ prototypes
f	typical VLT instrument feasibility study+ prototypes
g	typical VLT instrument feasibility study+ prototypes (simpler)
h	
i	
j	Focal station not equiped with WFS
k	Two WFSs (Pyramid and SHWFS) + optomechanics
l	
m	Blind guess: Development from 10k to 100kact. + science grade device.
n	Blind guess but recurrent cost only
o	High risk cost estimate (need of new technology: little lisibility)

5 Science Case

5.1 Introduction

Since the epochal discovery of the first extra-solar planet around 51 Peg by Michel Mayor and Didier Queloz in 1995, the study of extra-solar planets has captured an increasing attention among astronomers, as well as the general public. The discovery of other planetary systems has enormously boosted the investigations about the occurrence and formation of planets around other stars. The ultimate goals are the understanding of the mechanisms of formation and evolution of planets, and possibly the discovery of other planets able to host life.

The extreme contrast in mass and luminosity between the planets and the stars around which they orbit make detailed studies of extra-solar planets challenging. In particular, direct observations of extra-solar planets is still beyond capabilities of currently available instrumentation, save for perhaps a few extreme cases of very young and massive planets at large distances from the central star. While progresses in instrumentation might allow significant progresses either with 8-10 m ground-based telescopes, or with some of the satellites currently in the construction phase (like JWST), imaging of extra-solar planets over a wide range of parameters requires extremely large ground-based telescopes like OWL or specially designed space instrumentation (TPF or Darwin).

Imaging extra-solar planets would be of extreme importance for a number of reasons: (i) acquisition of their images would provide immediate confirmation of their same existence; (ii) analysis of their spectra would provide direct information about the presence and composition of their atmospheres, and even about the presence of life on its surface; (iii) analysis of the polarization characteristics can provide clues on the structure of their atmospheres – e.g. the presence of dust; (iv) determination of their positions at different epochs would provide their orbit; (v) perturbations around an elliptical orbit might disclose the presence of other unseen bodies in the system (like it happened for the discovery of Neptune in our own Solar System); (vi) analysis of the light curve might provide the period of rotation of the planet around its own axis, information about the presence of satellites and rings, and finally data about weather or even about the presence of continents covered by vegetation. It is clear then that only through direct imaging, our knowledge of extra-solar planets can really progress into a mature stage of detailed characterization, so that we will begin to answer to basic questions like those about the possible presence of life on such systems.

Detection and characterization of extra-solar planets, possibly down to terrestrial planets, is among the main goals of OWL. **EPICS** is the instrument considered for imaging extra-solar planets with OWL. Follow up observation in the thermal infrared as well as the detection of molecular features such as CO_2 and O_3 (which do not exhibit strong bands in the visible/NIR region) on nearby objects, will be the domain of the thermal infrared instrument for OWL (**T-OWL**). This section will outline the primary science that will be performed with EPICS, although plenty of interesting secondary science in the fields of star formation and debris disks would certainly be supported by the instrument. It will also motivates the Top Level Requirements (TLR).

Two main classes of planets will be considered:

- Giant gaseous planets, likely similar to the giant planets of our own Solar System (Jupiter, Saturn, Uranus and Neptune)
- Smaller rocky planets, similar to the inner planets of the Solar System (in particular Venus, Earth and Mars).

In the solar system, these two classes of planets are clearly distinct according to their mass and location in the system. Rocky Planets have a mass similar to or smaller than the Earth mass, and are all located within the Mars orbit (about 1.5 AU) from the Sun. The Giant Planets are much more massive, with masses ranging from about 15 Earth masses (Uranus and Neptune), up to several hundreds of solar masses (Jupiter). They are all located at large distances from the Sun (from 5 to about 40 AU). These facts are generally thought to be related to the mechanism of formation of planets within the proto-planetary disk (cite). However, observation of extra-solar planetary systems have shown that these characteristics are not at all universal among planetary systems, suggesting that mechanisms not included in the simplest models of planet formation (in particular orbital migration) are important.

As a consequence, we might expect that the simplified picture based on the Solar System is not correct. In particular, the minimum distance from the star where giant gaseous planets occur may clearly be much smaller than in the solar system – and may even be as small as a few hundredths of AU. Furthermore, the value in mass at which the transition from small rocky planets to giant gaseous planets – reflecting the capability of planets to attract a considerable amount of gas from the proto-planetary disk, and then expected to be a quite general property of planets – is uncertain, and possibly significantly variable from system to system. In spite of this considerable uncertainty, we would hereinafter assume that the transition from small rocky to giant gaseous planets occurs somewhere around 10 Earth masses. A better determination of this transition is certainly amongst the basic goals of extra-solar planet science for the next decades.

5.2 Giant Planets: Science milestones

In order to properly characterize the possible contribution of OWL+EPICS to the science of extra-solar planets, it is important to have a broad perspective of such studies. In Table 5-1 we present schematically our view of the most important milestones in the studies of Giant Planets. For each milestone, we indicate the crucial instrument/technique, as well as when we expect the various milestones will be reached (of course, in a few cases such milestones have been already achieved). We use a color code to distinguish those milestones that have been already achieved (green); those that will likely be achieved before OWL+EPICS get into operation (that is, before ~2015); and those that will more likely require a longer time, so that they will be achieved in an epoch where OWL+EPICS would be contributing to science.

– **First detection**
– **Indirect: RV (OHP - 1995), transits (1999), μ-lens (2004), astrometry VLTI-PRIMA? (2007)**
– **Direct: VLT+NACO - 2005 (BD + massive planets at large separation)**
– **Statistics (Mass, orbital parameters, central star parameters, multiple systems)**
– **Hot Jupiters: RV surveys (<2005), Transit Surveys (>2005), Corot (2006), Kepler (2010)**
– **Solar-system like Jupiters: RV surveys (>2005), μ-lens Surveys (>2005), VLTI-PRIMA? (2007), VLT-PF (2010), SIM (>2011), GAIA (>2012), 30m-PF (>2014)**
– **Understanding Giant Planet formation: relation between giant planets and protoplanetary/debris disks**
– **Systems from RV surveys: present (ε Eri), ALMA (2007), nulling interferometers (>2007)**
– **Direct observation of planets and disks: VLT PF (2010), JWST (2012)**
– **Understanding Giant Planet Physics (structure, atmosphere)**
– **Hot Jupiters: HST (<2010), Spitzer (2005), Kepler (2010)**
– **Solar-system like giant planets: VLT-PF (2010), JWST (2012), SIM (>2011), GAIA (>2012)**
– **Rotation, atmospheric motions: 30m-PF (>2014), EPICS (>2017)**
– **Satellites and Rings of Giant Planets: 30m-PF (>2014), EPICS (>2017)**

Table 5-1 Giant Planets Science Milestones

While details of Table 5-1 can be questioned, since they result from the authors' personal view on these issues, even a coarse exam reveals that a large fraction of the most interesting science goals on giant planets will be reached in the 1995-2015 epoch, before OWL will be available. OWL+EPICS may still provide very valuable results about the presence of satellites and rings, that require large telescopes and rather high precision photometry over a significant interval of time: this can only be obtained using extremely large ground based telescopes. However, we expect that for the epoch of operation of OWL+EPICS, science of giant extra-solar planets would have reached a considerable state of maturity, where significant further progresses would require considerable effort.

5.3 Rocky Planets: Science Milestones

We turn then our attention to rocky planets. Table 5-2 is analogous to Table 5-1, but now focuses on rocky planets. The same color code is used.

– **First detection**
– **Indirect: μ-lens Surveys (??), Kepler (>2008) (planets in the habitable zone? No satellite foreseen after Kepler)**
– **Direct: Coronagraphic-TPF (>2014), Darwin (>2015), Interferometric-TPF (>>2014), EPICS (>2017)**
– **Statistics:**
– **Hot planets: μ-lens Surveys (??), Kepler (>2008)**
– **Planets in the habitable zone: μ-lens Surveys (??), Future satellite transit survey (??), Coronagraphic-TPF (>2014), EPICS (>2017)**
– **Understanding the relation between rocky planets and giant planets**
– **Coronagraphic-TPF (>2014), EPICS (>2017)**
– **Understanding Rocky Planets physics (orbital parameters, spectra, polarimetry)**
– **EPICS (>2017)?**
– **Detecting life signatures**
– **Darwin (?) (>2015), Interferometric-TPF (?) (>>2014)**
– **EPICS (>2017)??**
– **Possibly >2020 from space**

Table 5-2 Rocky Planets Sciences Milestones

It should be clear that details of Table 5-2 are even more questionable than those of Table 5-1, since they are based on expected performances of instruments that are still to be constructed. However, a coarse exam of it reveals that very few significant milestones are expected to be reached significantly before the epoch in which OWL+EPICS will become operative.

At present, we expect that first detection of rocky planets will be achieved well before the OWL epoch using transits by the NASA Kepler mission, whose launch is currently scheduled for 2008, even though there is a limited chance that this epochal result could perhaps be obtained even earlier using other instruments (the French satellite Corot expected to be launched in 2007) or techniques (gravitational micro-lens observations or even radial velocities). However, rocky planets discovered by these techniques would be hard to be confirmed as such, mainly so if their orbits have periods longer than a few days. There is at present no other space transit mission foreseen after Corot and Kepler, so that it is well possible that if these missions will not be able to observe rocky planets, such a discovery will need at least a full decade from now. Furthermore, almost all other milestones in rocky planet science require imaging techniques, and then clearly fall in the realm of extremely large telescopes and specialized imaging satellites.

5.4 Detection of Rocky Planets in the habitable zone

In order to further examine imaging of rocky planets with OWL+EPICS in a more general context, we consider in Table 5-3 detection of rocky planets in the habitable zone, that is of

planets similar to our own Earth, with different techniques. The habitable zone is better defined in the next Section. Here, it is generally considered to be of the order of 1 AU for solar type stars. In this Table we compare the accuracy required for detection with that achievable with the various techniques. In the last column of the Table we list the main sources of concern in the various observing techniques. In the Column 2 of Table 5-3, we use a color code: green means that the technique can reach the required accuracy; red means that this is not possible.

Method		Achievable Accuracy	Required Accuracy	Problems
Radial Velocities		>0.1 m/s	0.03 m/s	jitter, pulsations
Astrometry (Space)		10 μarcsec	1 μarcsec	stellar diameter, activity
Transits	Ground	1 mmag	0.1 mmag	scintillation, false alarms, no follow-up, rare
	Space	0.01 mmag		
Gravitational Micro-lenses		1 mmag	1 mmag	unique event, no follow up, very rare
Imaging	VLT	10^8	10^{10} contrast	differential imaging, coronagraphy, nulling interferometry
	30 m	10^9		
	OWL	10^{10}		
	Space	10^{10}		

Table 5-3 Detection of an extra-solar Earth with different techniques

An exam of Table 5-3 reveals that those indirect techniques, where the planet presence is revealed by the reflex motion of the central star, are not well suited for observation of rocky planets in the habitable zone: this is essentially due to problems related to time variations and asymmetries in the flux emitted from stars, and are then difficult to overcome even by significant technological progresses. This limits substantially the available techniques, and give much more weight to the imaging techniques. Also, transits and gravitational micro-lens, while likely sensible enough to detect rocky planets in the near future, are severely limited by the rarity of the alignment conditions required for detection, and by the size of the field used in the searches: in both cases, rocky planets in the habitable zone detected will likely be at distances much larger than 100 pc from the Sun, so that detailed studies are beyond capabilities of even extremely large telescopes and of those space instruments expected to be working in the next 15-20 years. It is then clear that direct imaging will likely play a crucial role in the extra-solar rocky planet science.

The problem cell for the imaging case signals that various special techniques must be used to achieve the maximum contrast. They include differential imaging, coronagraphy and nulling interferometry. Note that some of these techniques may be combined, to further boost instrument performances. This is e.g. the case of differential imaging and coronagraphy. Furthermore, ground based observations assume extreme Adaptive Optics, providing a high

Strehl Ratio at the wavelengths of observations. Combining some of these various techniques in a single instrument concept is the challenge of EPICS.

Summarizing the content of Table 5-2 and Table 5-3 we may conclude that:

- Most interesting science on rocky planets will be done in the 2010-2030 epoch, just when OWL will be available
- A 30m telescope is not large enough for rocky planets
- Early rocky planet science will be mainly made by OWL+EPICS, TPF-Coronagraphic, Darwin

In the following, we will assume that the main goal of OWL+EPICS is detection and characterization of rocky planets in the habitable zone (see below for a definition of the habitable zone). For characterization, we mean observation of a statistically significant sample of rocky planets, to understand the relation between rocky and giant planet, and to understand the main features of rocky planets physics (we will better define these points in the next sections). On the other hand, detection of undisputable life signatures on the spectra of these rocky planets is likely extremely challenging (again, see following discussion): while OWL+EPICS might possibly achieve such goal, we consider it beyond its baseline goals. Specialized space instruments working in the mid-IR like Darwin may do this part of the job.

5.5 Habitable Zone and Optimal Separation

As mentioned above, the main scientific goal of EPICS is the detection and characterization of a statistically significant sample of rocky planets in the habitable zone around a star. The habitable zone is defined as the region around a star where liquid water may be present on the surface of a rocky planet. This definition stems from assumptions about the nature of life (requiring liquid water as solvent), as well as on the stability over time of the conditions needed for its development (this eliminates giant gaseous planets as likely host of life, due to the presence of thick atmospheres characterized by deep convective circulation).

Figure 5-1 Habitable Zone of various star types as function of mass and orbital distance (Kasting, 1997)

For planets without atmospheres, the habitable zone around a star may be determined from the equilibrium conditions between the heat received by a planet from the star, and that dissipated by emission from the planet surface. Habitable zones should then depend on the brightness (mass) of the central star (see Figure 5-1). For cool main sequence stars (M-dwarfs) the habitable zone can be so close to the central star that the planets are expected to be locked in synchronous orbits, so that the same hemisphere always face the star (a condition usually considered to be not favorable to life).

The situation is more complicated for planets with atmospheres due to the greenhouse effect. In the case of the Earth, the greenhouse effect causes a significant warming of about 30 K of the surface, making it adequately warm to host life. The intensity of the greenhouse effect depends on the thickness and composition of the atmosphere. The thick CO_2 dominated atmosphere of Venus produces a strong greenhouse effect that warms its ground temperature to values of ~500 K, while the thin atmosphere of Mars causes an almost negligible greenhouse effect, so that both planets are inhospitable for life. The characteristics of the atmospheres, and then intensity of the greenhouse effect, likely depend on the irradiation from the central star. However, other mechanisms are also relevant: in the Earth, the amount of CO_2 in the atmosphere is controlled by a complex cycle that include addition of CO_2 to the atmosphere through volcanism, removal by solution in the ocean water, deposition of carbonate on ocean bottoms, subduction in the mantle and then recycling again through volcanism as key elements. This implies that active plate tectonic plays a crucial role in maintaining stable the CO_2 level in the atmosphere, and then the Earth habitable. The thick CO_2 atmosphere of Venus is attributed to the fact that plate tectonic is not any more active there. Since it is not clear why plate tectonic is still active in the Earth, but not in Venus, we should consider quite extended habitable zones around other stars.

Figure 5-2 Evolution of habitable zones around stars of different masses

Other mechanisms should be taken into account. The central star evolves becoming brighter. If not compensated by a decreasing efficiency of the greenhouse effect (due e.g. from a decrease in the CO_2 content in O_2 dominated atmospheres), the stronger irradiation should result in increasing temperatures, that would eventually make the planet inhospitable to life. The habitable zones are then expected to evolve with time (see Figure 5-2).

Separation of planets in habitable zones around other stars depend of course not only on the distance from the planet to the star, but also on the distance from the Sun to the star and on the angle of sight. Figure 5-3 gives the maximum separation of solar system planets from the Sun (in arcsec) as seen from various distances (in parsec). From this figure, it is clear that typical separations of planets in the habitable zone are expected to be <0.2 arcsec, save for the very closest stars.
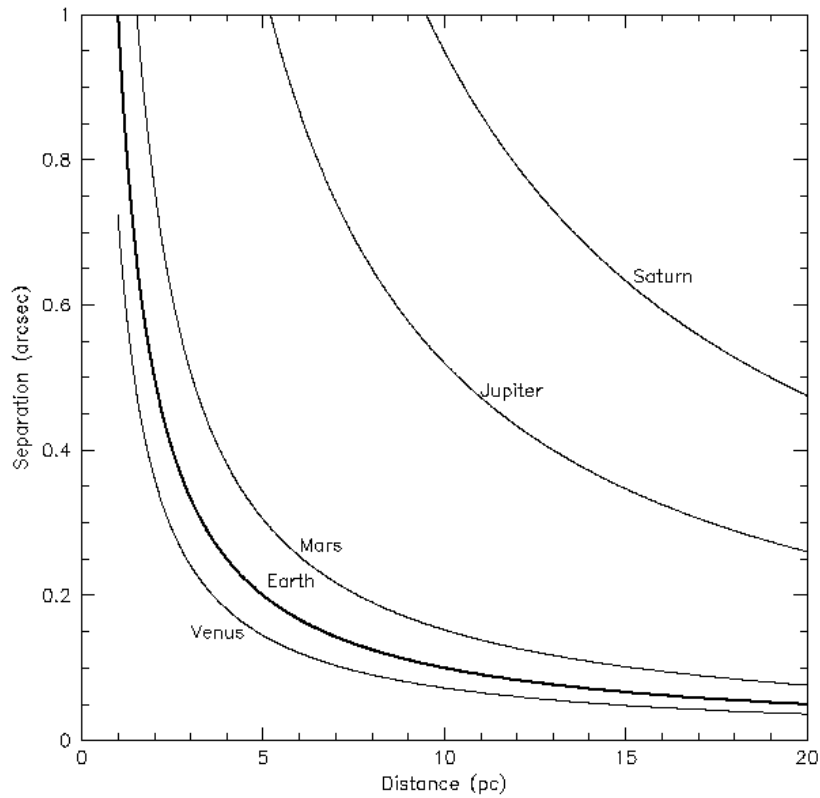
Figure 5-3 Maximum separation of solar system planets from the Sun as seen from various distances

Figure 5-4 shows the maximum separation of the habitable zones around a list of targets selected for Darwin. Since most of the stars considered in this list are M-dwarfs, the habitable zones are even closer to the central stars than expected from Figure 5-3 (which consider a G2V central star). For almost all the Darwin targets considered here the habitable zones are within a few hundredths of arcsec from the central star.

Summarizing the data shown in Figure 5-3 and Figure 5-4, we conclude that the region of interest around the star for EPICS is from a few hundredths to a few tenths of arcsec. The outer limit roughly corresponds to the control radius of high order Adaptive Optics. The inner limit sets the constraints on the coronagraphic performance of EPICS.

Figure 5-4 Habitable zone for Darwin targets

5.6 WAVELENGTH RANGE

Figure 5-5 compares the flux emitted by the Sun (a G2V star) with those coming from the planets of the solar system (Jupiter, Venus, Earth, and Mars). The spectral distribution of the zodiacal light is also shown for comparison. The two peaks in the visible-near IR and in the mid-IR corresponds to the maxima of reflected light and intrinsic emission respectively. From this Figure, it is clear that there are two favorite spectral regions for planet observations: the visible-near IR and the mid-IR. Ground-based observations in the mid-IR are severely hampered by the atmospheric thermal background, that is about 4 orders of magnitude brighter than what can be obtained from space observations.
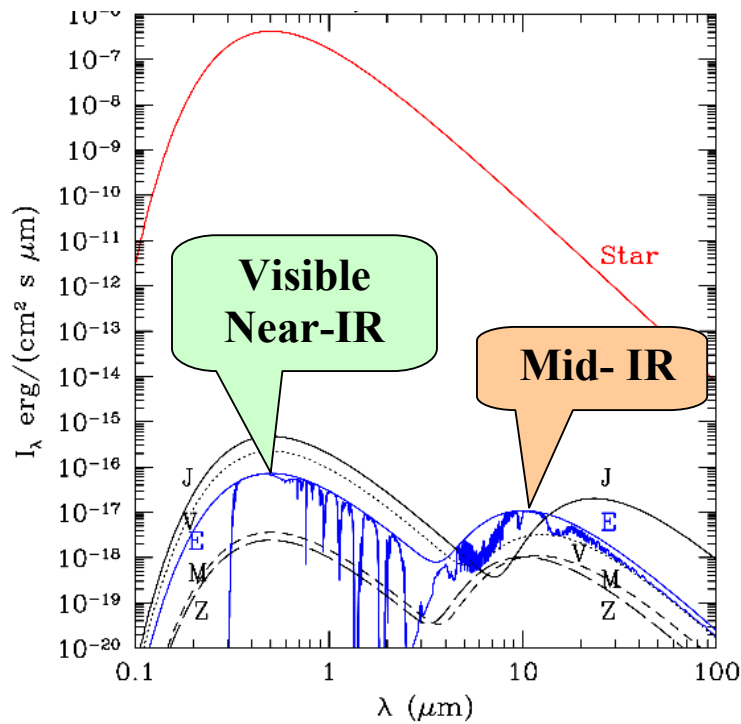
Figure 5-5 Comparison between the flux emitted by the Sun (a G2V star) and those coming from the planets of the solar system (J=Jupiter, V=Venus, E=Earth, M=Mars). Z represents the spectral distribution of the zodiacal light. The two peaks in the visible-near IR corresponds to the maxima of reflected light and intrinsic emission respectively.

Table 5-4 compares the relative pro's and con's of these two spectral regions. The main advantages of the Visible-Near IR are the low background level (both Zodiacal and Atmospheric) and the fact that only small parts of the instrument need to be cryogenic. The main advantage of the Mid-Infrared is in the much smaller contrast between the planet and the star. The high atmospheric background means that space instruments are much favored. **Hereinafter it will then be assumed that EPICS will be optimized for observations in the visible and near-IR.**

	Visible-Near IR	Mid-Infrared
Earth/Sun Contrast	$\sim 10^{-10}$	$\sim 10^{-7}$
Zodiacal background	Low	High
Atmospheric background	Low	High
Spectral features	H_2O, O_2, CH_4, CO_2	H_2O, O_3, CH_4, CO_2, N_2O
Ground-based Instrument	EPICS	T-OWL
Space Instrument	TPF-Coronagraph	Darwin

Table 5-4 Comparison between Visible - Near IR spectral regions

5.7 Molecular Spectra

Based on the variety of atmosphere composition observed in the Solar system, it is expected that also the planets targets of EPICS will possess a wide range of atmospheric properties. Essentially, we consider at least four different classes of planets:

- **Planets without atmosphere**, or with a very thin atmosphere that does not produce significant features, like Mars. These are likely planets of either small mass and/or very close to the central star
- **Planets with atmospheres dominated by methane**, like the giant planets in the Solar System. These are likely either massive and/or cold planets, able to maintain a substantial amount of H in their atmosphere
- **Planets with atmospheres dominated by carbon dioxide**, like Venus
- **Planets with atmospheres dominated by oxygen**, like the Earth

Life might be present in all classes of planets: however planets with atmospheres dominated by oxygen are the most appealing candidates, essentially because it is thought that free oxygen in the atmosphere should be of biological origin, since all primeval oxygen is expected to be locked in rocks as oxides. Most clear diagnostics of the presence of life would be the simultaneous presence of O_2 and O_3 in the planet spectra. Alternatively, a strong coverage by green plants (a secure life signatures) can be provided by evidences of the presence of chlorophyll features in the spectrum.

The three last classes of planets cited above have quite different spectra, both in the Visible and Near-IR, as well as in the Mid-IR. Figure 5-6 illustrates the main features in these spectral regions of the most important molecules present in planetary atmospheres.

Figure 5-6 Main features in the Visible-Near IR (left column) and in the Mid-IR (right column) spectra of the most important molecules expected to be present in planetary atmospheres (from Traub & Jucks, AGU Geophys. Monograph 130, 2002; and Des Marais et al, Astrobiology 2002).

An exam of Figure 5-6 reveals that in the Visible-Near IR spectral range (selected for EPICS), features due to H_2O, O_2, CH_4, and CO_2 are well visible. This should allow appropriate classification of the detected planets into the various classes defined above even from low resolution (survey type) spectra. The narrowest spectral features are those of O_2 (the best band is around 0.76 μm): optimization of observation of this band requires a **spectral resolution of ≥100**. The spectral region to be covered by the spectra ranges from about 0.6 (see next section) up to about 1.7 μm. Addition of the *K*-band would add further information about CO_2, that is however not critical since relevant data can be obtained also at shorter wavelengths.

A major difficulty in the observations of oxygen dominated atmospheres, similar to the Earth one, is that similar features are present in the Earth atmosphere too. Features due to the extra-solar planet can be distinguished by the radial velocity difference, that would shift the extra-solar planet absorption out of the telluric feature: even at the low resolution likely considered for EPICS, this would result in stronger overall absorption, at least for those spectral bands that are not heavily saturated. However, it is clear that **detection of atmospheric signatures of O_2 and H_2O dominated atmospheres require special (extremely dry) observing conditions**. Insofar detection of planet possibly hosting life is among the major science objectives of OWL, this should be considered in the selection of telescope site. Detection of the most critical life diagnostics requires a more extended discussion, that will be given in the next subsection.

5.7.1   Life diagnostics

The molecules that indicate beyond doubt the presence of life on a planet (biomarkers) are ozone and chlorophyll. The presence of O_2, while possibly connected with life, might be interpreted in alternative ways.

The presence of water indicated by the strong water vapor absorption bands separating the observing windows in the NIR is regarded as a prerequisite for habitability but not as a proof for the existence of life (definition of the habitable zone). However, the presence of O_2 in combination with the presence of H_2O seems to indicate the presence of carbon based life and would be observable in the Vis/NIR wavelength range.

Ozone has transitions both in the visible (at about 0.6 μm) and in the mid-IR (see Figure 5-6). In the integrated Earth spectrum (see Figure 5-7 and Figure 5-8), the feature at 0.6 μm is quite broad (being resolved at R>10) and weak (central absorption of about 5%), and it can only be observed in spectra of rather high S/N. Observation with EPICS will be extremely challenging, both for the short wavelength and weakness. Much more intense is the mid-IR transition at about 10 μm (see Figures Figure 5-9 and Figure 5-10), that should be easily visible even in spectra of rather coarse resolution and S/N.

Chlorophyll causes the so-called vegetation red edge in the integrated solar spectrum. This shows up as a sudden decrease of flux below about 0.7 μm in the Earth integrated spectrum.

The exact wavelength of the vegetation red edge depends on the kind of green plants which dominate the spectrum, and it is then expected to be variable from planet-to-planet. However, the visibility of the vegetation red edge even in the Earth spectrum is questionable: for instance, very little if any evidence of this red edge is observed in the spectrum of Figure 5-8; even in the spectrum of Figure 5-7 the evidence for a vegetation red edge is unclear. This lack of evidence for a vegetation red edge is likely related to the fact that only a fraction of the Earth surface is covered by green plants, and that only part of this can be seen from outside due to the presence of clouds. The vegetation red edge is then expected to be strongly variable not only from planet to planet, but also with time for a given planet. On the whole, it is expected to be weak and difficult to observe, unless high S/N observations are available.
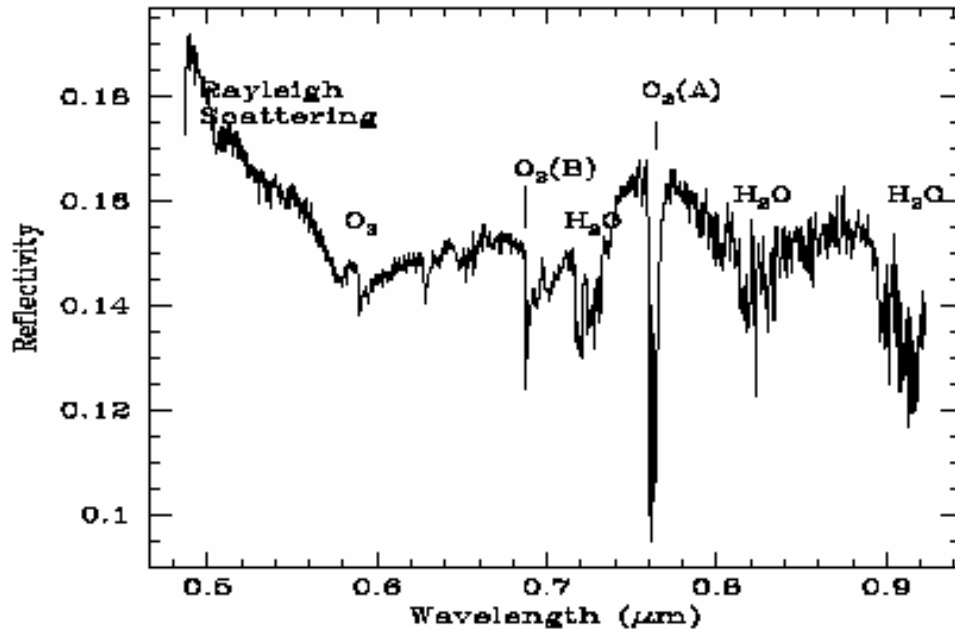
Figure 5-7 Integrated earth light reflected by Moon dark side in the visible: chlorophyll, O_2, O_3, H_2O (from Woolf, Smith, Traub, & Jucks, ApJ 574 2002, astro-ph/0203465)

Figure 5-8 Integrated earth light reflected by Moon dark side in the visible: chlorophyll, O_2, O_3, H_2O (from Montanes Rodriguez et al., ApJ in press 2005, astro-ph/0505084)

Figure 5-9 Mid-IR integrated Earth light observed by TES (Thermal Emission Spectrometer) flying toward MARS (1996): CO_2 (15 μm), O_3 (9.6 μm), H_2O.

In summary, biomarkers are difficult targets for EPICS, planned to work in the visible and near-IR. Complementary and compelling evidence for the presence of life on a planet is expected from instruments optimized for the mid-IR observation of atmospheric features in an extrasolar terrestrial atmosphere, like Darwin (see Figure 5-10).
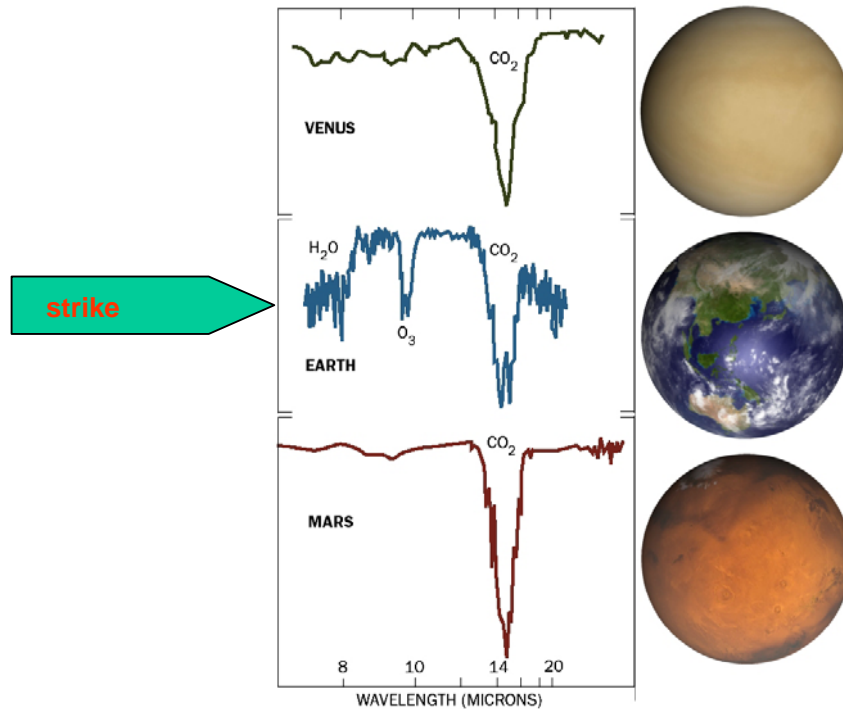
Figure 5-10 Mid-IR integrated spectra of the rocky planets of the Solar System: Venus, Earth and Mars. Note the presence of ozone and water vapor in the Earth spectrum, but not those of dry planets like Venus and Mars. The quite strong ozone feature at about 10 μm is considered the best biomarkers.

5.7.2 Species and their properties

Table 5-5 summarizes the capabilities of EPICS, Coronagraphic TPF and Darwin for analyzing atmospheres of extra-solar planets. Working in the Visible and near-IR (from 0.6 to 1.7 μm), EPICS will be able to observe most relevant species and properly characterize the planetary atmospheres. Detection of clear biomarkers is very challenging, and it can be considered as a follow-up observation dedicated to only a few targets selected from an extended survey.

	Visible/Near IR EPICS - TPF	Mid-IR DARWIN
CO_2	Yes (if abundant)	yes

H₂O	**Yes**	**Yes**
O₂	**Yes**	**No**
O₃	**Yes**	**Yes**
CH₄	**Yes (if abundant)**	**Yes (if abundant)**
N₂O	**No**	**Yes (if abundant)**
T	**Yes (derived, surface)**	**Yes (direct, stratosphere)**
pressure	**Yes**	**No**
radius, mass	**Yes (inferred)**	**Yes**
chlorophyll	**Yes?**	**No**

Table 5-5 Summary of capabilities of EPICS, Coronagraphic TPF and Darwin for analyzing planetary atmospheres.

5.8 Scaling from VLT to OWL

In order to understand the main properties of EPICS for imaging planets, it is useful to start from scaling the expected performances of 8 meter imagers (studied in detail in the Planet Finder studies for VLT) to a planet imager for OWL. The main scaling relations are as follows:

- The telescope area scales as d^2 ➔ everything else constant planet signal and background per unit sky area raise as d^2
- The area of the Airy disk scales as d^{-2} ➔ everything else constant the background per Airy disk is constant
- If the background (star or sky) dominates, the S/N scale as d^2
- VLT-PF limiting contrast is ~10^8 at 1 arcsec
- The limiting contrast for OWL should then be ~10^{10} at 1 arcsec
- For comparison, the Sun/Earth contrast is about 2×10^{10}. However the separation would appear smaller than 1 arcsec for all targets
- Sampling should be done at the Airy Disk limit (10^{-3} arcsec) ➔ relatively large detectors are required to explore the useful region of sky

It is also useful to consider the AO requirements for OWL. In fact, the scaling laws considered above assume:

- High Strehl ratios (>0.5 in *I*, >0.7 in *J*).
- This requires to sample the pupil at about r_0 (20 cm) ➔ large number of actuators (~$2 \ 10^5$!)
- Note that this imposes limiting magnitudes similar to those of CHEOPS (I~10) if the central star is used as the reference star!
- Note also that MS stars within 5 pc are resolved at this resolution.

5.9   Expected Output of OWL+EPICS

To simulate planet detection capabilities with EPICS we performed a Monte Carlo simulation using the code described in Desidera et al. 2004, Doc. CHEOPS-SRE-OPD-00065. 10000 planets were generated following the mass function of Solar System planets, the period and eccentricity distributions of known extra-solar planets. Only reflected luminosity was considered (for terrestrial planets intrinsic luminosity in *I* band would be significant only on very limited timescales). The projected separation and the phase factor were calculated self-consistently from the orbital parameters. The planet luminosity was assumed to scale with mass using the mass-radius relation of Solar System planets. The same albedo was assumed for all masses. The detection threshold shown in Fig. 12 was adopted to evaluate the detectability of each generated planet. The simulation was performed for the 650 targets considered in Desidera et al. 2004, Doc. CHEOPS-SRE-OPD-00065: this sample includes all the stars within 20 pc in Hipparcos catalogue with declination <+20. The actual stellar properties (mass, distance, *I* magnitude) were considered in deriving the planet parameters and the star/planet contrast.

Binarity was taken into account excluding from detectability the planets outside the stability limits.

We included planets down to the mass of Mars, considering that lower mass planets are not expected to have atmospheric spectral features dominating their spectra, making more challenging their detection using differential imaging or differential polarimetric techniques. The frequency of terrestrial planets is currently unconstrained by the observations. In the simulations we adopted as normalization that each star have 3 planet with mass <1 M_{Earth} within 2 AU. The absolute number of expected detections scales with the adopted normalization while other results (minimum mass of detectable planets, mass distribution of detected planets etc) do not. For this normalization, it results that the number of expected detection of planets within 2 AU with masses smaller than 10 Earth masses is more than 500.

The results are shown in Figures Figure 5-11 to Figure 5-14 14-17. Figures Figure 5-11 and Figure 5-12   the minimum mass of detected planets as a function of distance and *I* magnitude. Mars-size planets result detectable around a large fraction of the targets, while a minimum detectable mass above 1 M_{Earth} is found for a few objects with distance larger than 10 pc and *I* magnitude fainter than 6 (the few close/brighter objects with larger minimum mass are members of binaries for which stable orbits are not possible in the region with the largest detectability, 0.05-0.10 arcsec).

Figure 5-11 Maximum detectable mass as a function of stellar distance
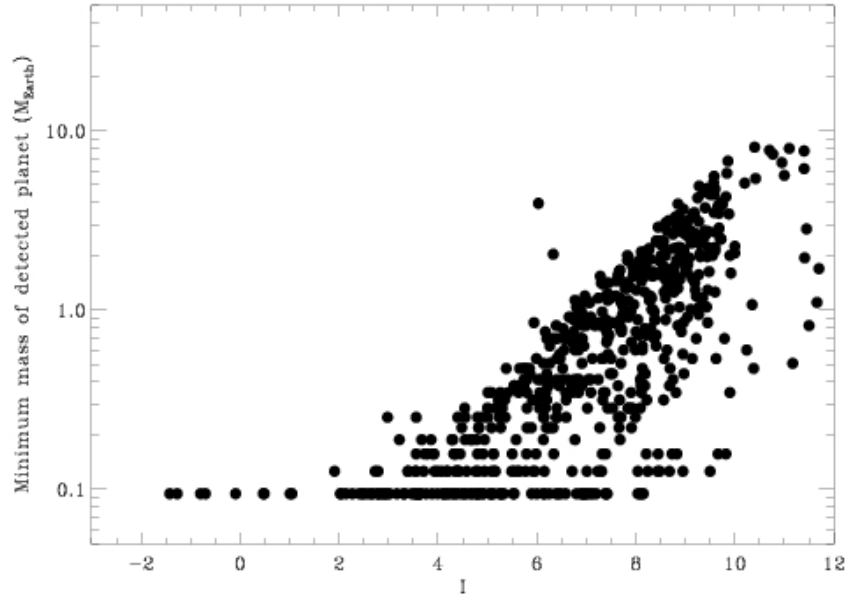
Figure 5-12 Minimum detectable mass as a function of stellar magnitude

Figure 5-13   Distribution of masses of extra-solar terrestrial planets expected to be discovered by an appropriate survey with OWL + EPICS

Overall, it results that Earth mass planets are detectable (at least in a limited separation range) around a few hundredths objects in our sample, but the inclusion of farther, relatively bright stars will certainly increase this number. Figure 5-13 shows the mass histogram of the detected planets, strongly peaked toward lower mass planets, more numerous according to the input distribution.

Figure 5-14 shows the histogram of the star/planet contrast of the detected planets, compared with the results obtained for our simulation of the VLT Planet Finder. EPICS will allow an improvement of more than two orders of magnitude.

Since these simulations were carried out assuming the spectral properties of methane dominated giant planet spectra, they are not necessarily representative for a population of terrestrial planets in the habitable zone and only give an indication of the discovery space to be explored by EPICS.

Figure 5-14 **Histogram of star/planet contrast of the planets that results detected in our simulation (continuous line), compared with derived for the VLT Planet Finder.**

5.10 Photometry with EPICS

Accurate photometric time series for the detected extra-solar planets discloses the possibility of studying several interesting aspects of planetary science.

(a) Rotation

Planet rotational period can be obtained by periodic variations in the light curve. Periods are expected to range from a few hours (giant planets of the solar system) up to hundreds of days (synchronous orbits, Venus). The rotational signal is due to variations in the planet reflectivity due to not-uniform surface features. These are expected to be strong in planets having atmospheres with incomplete cloud coverage, like the Earth *[give values]*. In other cases the signal variations might be more subtle. However, photometry accurate to a few per cent should be adequate to study planet rotation if long enough time series are obtained.

(b) Weather and Phase Variations.

Short and long term variations in the planet reflectivity can occur due to weather and to seasonal variations along the orbit. These last can be strong for planets on eccentric orbits, like a large fraction of the extra-solar planets discovered from radial velocity surveys. An important complication in the observation of seasonal reflectivity changes is due to the variation of the stellar background related to the different projected separation along the orbit. This will likely limit the photometric accuracy. Rings around planets also cause phase variations. For a detailed study of the impact of rings on observation of extra-solar planets, see Arnold & Schneider (A&A, 420, 1153, 2003).

(c) Observations of satellites (transits and eclipses).

In the case of the Earth, the presence of the Moon, with its stabilizing effect on the Earth rotation axis and then on the climate, is thought to favour the presence of life. Observations of satellites around other terrestrial planets would then be important to better understand the mechanisms of life formation on planets. In the solar system, satellites are very common: in particular, a numerous satellite family seems a natural outcome of the formation mechanism of giant planets, possibly indicating the presence of circum-planetary rings in early phases of formation of gas-rich planets. The mass ratios of planet-satellite systems range from 81 for the Earth-Moon case, which appears to be exceptional, up to ~10^3 for Titan and Triton, and 10^4-10^5 for the other major satellites of Jupiter and Saturn. Even larger values are found for the very numerous groups of small satellites, which however appear in general to be objects formed elsewhere and then captured by the planets, being often in highly inclined or even retrograde orbits. Diameter ratios range from a factor of ~4 (Earth-Moon case) up to a few tens (major satellites of Jupiter and Saturn). Some of the satellites have atmospheres. Satellites with atmospheres around giant planets are as good candidates as terrestrial planets to host life if they are in the habitable zone.

Based on solar system data, planet-satellite separations are expected to be of the order of 10^6 km, i.e. about 6 mas at a distance of 10 pc.. In most cases satellite images will then not appear resolved from the planet ones (but there might be interesting exceptions). However, satellites might still be visible by considering their impact on the combined light of the planet-satellite system. Furthermore, satellite orbital period and transit duration would allow direct estimates of the planetary mass. Transits of satellites would be a rare phenomenon (less than 1/10 of the detected planets would have transiting satellites, the fraction being likely more closer to this value for giant planets). The signal is in this case of the order of the square of the diameter ratios, which may be of the order of per cent in most favorable cases. Eclipses of the satellites by the planet or of the planet by the satellite would be much more common, since it is expected that orbital planes of the satellites around the planet, and of the planet around the central star, would be roughly aligned. In these cases (especially the second one) the signal is smaller, due to the finite distance of the planet from the star. Again, circumstances of the eclipse would provide crucial information about the satellite orbit and planet mass.

5.11 Astrometry with EPICS

With a diffraction image of the order of 1 mas, EPICS will have a large potential for astrometry (relative to the central star) on the planetary images, insofar atmospheric and instrument effects can be appropriately removed. These observations may provide a wealth of information, depending on its accuracy. Possible goals are:

(a) Of particular interest are the determinations of the periods and eccentricities. Also of interest is the determination of relative orbital inclination relative to the axis of rotation of the central star, to other planet orbits in the case of multiple planet systems, or finally to the stellar companion orbits in stellar binary systems. Orbital inclination is in fact a basic parameter in modeling the formation and evolution of planetary system. Also, statistical studies of the orbital parameters for planetary systems of different ages allow a discussion of the secular evolution of the orbital parameters and of dynamical stability. The accuracy

required for orbit determination depends on the star-planet separation and on planet periods. For short period planet (like those in the habitable zone), one or more orbits can be followed within the EPICS lifetime, so that astrometric accuracy is less critical. Since the orbits of planets detectable with EPICS have semi-axis of the order of 0.1 arcsec, measures with accuracies of the order of 1mas will already provide very strong constraints.

(b) Long enough time series might allow to show evidence for orbital perturbations due to other planets in the system, even if not observed. In the most favorable cases, this might allow determinations of planet masses.

(c) The presence of satellites around planets can also at least in principle be deduced from (short term – baseline of days or months) perturbations of the astrometric orbits. For unresolved images, measurements of shifts of photocentres appear more viable than those due of the barycentre owing to the more favorable dependence on the mass ratio: in fact photocentre shifts are expected to roughly scale with the mass ratio at 2/3 power (assuming comparable densities and albedos for the planet and the satellite), rather than the linear scaling expected for the motion of the barycentre. For the Earth-Moon system seen from a distance of 10 pc, the barycentre shift is about 2.5 μas, while the photocentre shift is about ten times larger (assuming a similar photometric signal from the planet and the satellite). Given the low SNR that will be obtained on terrestrial planets, this method appears promising for giant planets orbited by giant moons.

5.12 High Resolution Spectra of Giant Planets

Giant Planets in close orbits (so that their images fall within the AO correction radius), those far enough that stellar background becomes low, or young and then intrinsically bright, may be observed at high spectral resolution with acceptable *S/N* ratios using EPICS. *S/N* ratios of the spectra are expected to scale roughly inversely to the square root of resolution, essentially because of the smaller pixel size in wavelength units. For an old giant planet (radius 10 times the Earth radius) around a G2V star at 10 pc from the Sun, orbiting at about 2 AU from the star and observed when in quadrature (separation of 0.2 arcsec), spectra over a 12 hr exposure time are expected to have $S/N \sim 1000/\sqrt{R}$. Spectra at a resolution of $R=4\times10^4$ can then be obtained with a *S/N*~5. Situation will be even more favorable for young planets, like e.g. that orbiting around ε Eridani, that are intrinsically much brighter in the *J* and *H* bands. In these cases spectra with *S/N* ratios well in excess of 100 can be obtained at virtually all separations. Even spectra with *S/N*~5 may provide radial velocities with errors of a few tenths of km/s, depending on the spectrum width (the giant planet spectra are expected to be very rich of molecular lines, and are then well suited for precise radial velocity measurements). They may be used to better constrain the orbit, to look for large satellites (in the case of old systems, satellites having a mass larger than 5% the mass of the planet would be detectable, if in close orbits like e.g. Io around Jupiter; for young planets, satellites with even 1/1000 of the planet mass are detectable), to study planet rotation (Jupiter equatorial velocity is about 10 km/s) and velocity fields in the planet atmosphere, or to study in detail the planet atmospheric composition, including trace contributors.

5.13 Polarimetry with EPICS

5.13.1 Introduction

Stellar light that is reflected by a planet is generally polarized. The degree of polarization, P, of the reflected light depends on (1) the composition and structure of the planetary atmosphere and that of an underlying surface, (2) the wavelength, λ, of the light, and (3) the illumination and observation geometries, i.e. the angles of incidence and reflection. For spatially unresolved extrasolar planets, the illumination and observation geometries can directly be translated into the planetary phase angle α (the angle between the star and the observer as observed from the center of the planet). In general, P of a spatially unresolved planet will be largest when its phase angle α is around 90° [53], because light that has been singly scattered by gaseous molecules has its highest degree of polarization around 90°.

Polarimetry promises to be a valuable tool for the direct detection of extrasolar planets, because the direct, i.e. unscattered, stellar light can be considered to be unpolarized (integrated over the stellar disk) [ref], while the stellar light that is reflected by the planet is generally polarized. Polarimetry thus enables one to distinguish the direct stellar light from the planetary light. Making this distinction would otherwise be very difficult because of (1) the small angular distance between a star and its planet, and (2) the weakness of the planetary flux as compared to the stellar flux. Besides being useful for *detecting* extrasolar planets, polarimetry is also useful for *characterizing* them, because, as also described above, the degree of polarization of the reflected light is very sensitive to a planet's physical characteristics. This latter application of polarimetry has proven its strengths in studies of solar system planets, such as Venus, Earth, Mars, and the giant planets.
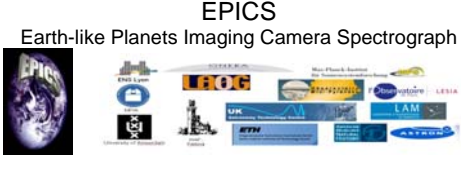
Here, we address polarimetry with EPICS as a tool for detecting and characterizing (terrestrial) extrasolar planets. First, we shortly introduce the flux vectors that are used to describe polarization. Then, we present polarization observations of planets in our own solar system, and discuss what these observations teach us about what to expect for observations of extrasolar planets. Next, we show results of numerical simulations of the polarization of extrasolar planets, both for terrestrial extrasolar planets and for Jupiter-like extrasolar planets, and discuss the wavelength dependence of the polarized signal as well as the dependence on the planetary phase angle.

5.13.2 Describing reflected light and polarization

The flux and state of polarization of planetary radiation with wavelength λ is fully described by a flux (column) vector $\vec{F}$

$$\pi\vec{F}(\lambda,\alpha) = \pi\big[F(\lambda,\alpha),Q(\lambda,\alpha),U(\lambda,\alpha),V(\lambda,\alpha)\big]. \tag{1}$$

Here, α is the planetary phase angle, i.e. the angle between the star and the observer as seen from the center of the planet, scalar F describes the total flux, Q and U the linearly polarized flux, and V the circularly polarized flux (all divided by π) (all with dimension W m^{-2} m^{-1}). In the following, we will neglect V, because it will be extremely small for extrasolar planets, and because there are no plans to measure it.

We define the fluxes Q and U with respect to a reference plane. Our reference plane is the *planetary scattering plane*, i.e. the plane through the centers of star, planet, and observer. Defining F_x as the flux that is measured through a polarization filter oriented perpendicular to the direction of propagation of the light and with its optical axis making an angle x with the reference plane, Q and U can be obtained through the following measurements [50].

$$Q = F_0 - F_{90} \tag{2}$$

$$U = F_{45} - F_{135} \tag{3}$$

The (total) flux F is simply the sum $F_0 + F_{90}$ (or $F_{45} + F_{135}$). Flux vectors can be transformed from one reference plane to another, e.g. the optical plane of a telescope instrument, by multiplying them with a rotation matrix [51]. In EPICS, the fluxes F, Q, and U will be measured not by using polarization filters, but with the sophisticated Zürich IMaging POLarimeter (ZIMPOL).

Neglecting flux V, the degree of polarization of the reflected light, P, is defined as the ratio of the linearly polarized flux to the total flux, thus

$$P(\lambda,\alpha) = \frac{\sqrt{Q(\lambda,\alpha)^2 + U(\lambda,\alpha)^2}}{F(\lambda,\alpha)}. \tag{4}$$

In the following, we assume that (1) a planet is symmetric with respect to the planetary scattering plane and (2) the incoming stellar light is unpolarized, flux U of the reflected stellar light will equal zero, and we redefine P as

$$P_s(\lambda,\alpha) = -\frac{Q(\lambda,\alpha)}{F(\lambda,\alpha)}. \tag{5}$$

Here, we introduced a minus sign to add information about the direction of polarization, i.e. if $P_s > 0$ ($P_s < 0$), the reflected light is polarized perpendicular (parallel) to the reference plane. The subscript s simply refers to the word "sign." Note that the assumption that a planet is symmetric with respect to the planetary scattering plane is reasonable considering the long integration times that will be required for the direct detection of extrasolar planets. Detailed numerical simulations for planets with significant asymmetries with respect to the planetary scattering plane will be the subject of later studies.

The flux vector of stellar light that has been reflected by a planet and that arrives at an Earth-based telescope located at a distance d is given by

$$\pi\vec{F}(\lambda,\alpha) = \exp(-\tau(\lambda))\frac{r^2}{d^2}\frac{R^2}{D^2}\frac{1}{4}\ddot{S}(\lambda,\alpha)\pi\vec{B}_0(\lambda). \tag{6}$$

Here, r is the radius of the planet, R the radius of the star, D the distance between the star and the planet, S the 4x4 planetary scattering matrix [53], and πB_0 is the flux vector describing the stellar surface flux. The disk-integrated degree of polarization of a solar-type star can be assumed to be zero [Kemp et al.,], thus vector B_0 is given by B_0 [1,0,0,0]. We assume that the starlight that is incident on the planet is unidirectional ($D \gg R$). Furthermore in Eq.(6), the exponent describes the transmission of the reflected starlight through the Earth's atmosphere, with τ the atmospheric extinction *along the line of sight*. Note that in Eq.(6), we have not included the response of the observing instrument, i.e. the telescope and

the instrument behind it. Furthermore, because the integration times for direct observations of extrasolar planets will be very long, the time-dependence of e.g. matrix S (upon rotation of the planet), the atmospheric extinction, τ, and the instrumental response function should be taken into account when actually using Eq.(6) for the interpretation of direct observations of extrasolar planets.

A planetary scattering matrix S not only depends on the wavelength and the planetary phase angle, but also on the composition and structure of the planetary atmosphere and an underlying planetary surface. Using the planetary scattering plane as reference plane and assuming the planet is symmetric with respect to this reference plane, matrix S can be described by

$$\bar{S}(\lambda,\alpha) = \begin{bmatrix} a1(\lambda,\alpha) & b1(\lambda,\alpha) & 0 & 0 \\ b1(\lambda,\alpha) & a2(\lambda,\alpha) & 0 & 0 \\ 0 & 0 & a3(\lambda,\alpha) & b2(\lambda,\alpha) \\ 0 & 0 & -b2(\lambda,\alpha) & a4(\lambda,\alpha) \end{bmatrix}. \tag{7}$$

Matrix S is normalized so that the average of the planetary phase function, which is represented by matrix element $a1$, over all directions, equals the planet's Bond (or spherical) albedo. The geometric albedo A_G of a planet is given by

$$A_G(\lambda) = \frac{1}{4} a1(\lambda, 0^o). \tag{8}$$

In case the incident stellar light is unpolarized, the degree of polarization P_s equals

$$P_s(\lambda,\alpha) = -\frac{b1(\lambda,\alpha)}{a1(\lambda,\alpha)}. \tag{9}$$

The degree of polarization of the reflected stellar light that has been reflected by a planet and that arrives at an Earth-based telescope thus only depends on elements $a1$ and $b1$ of the planetary scattering matrix S, and not on e.g. the distances between the star and the planet and between the star and the Earth. P_s is also independent of the extinction of the Earth's atmosphere, which is very convenient when looking for terrestrial-type extrasolar planets.

Matrix elements $a1$ and $b1$ can be numerically calculated for a given model planet as described in [53]. Results of such calculations for terrestrial-type extrasolar planets will be shown later. First, we will show some polarization observations of solar system planets.

5.13.3 Polarization observations of Solar System planets

Polarimetry has long been used in observations of solar system planets, mainly to characterize their atmospheric properties. A famous example that proved the strength of polarimetry was the analysis of observations of the disk-integrated degree of polarization of Venus as a function of its planetary phase angle that not only correctly predicted the sizes of the particles constituting the Venusian clouds, but also that they should be composed of 75% sulfuric acid solution [49]. Besides Venus, the planets Earth, Mars, Jupiter, Saturn, Uranus, and Neptune have been studied with polarimetry. The polarization observations of Solar

System planets can be used to estimate the polarization properties of extrasolar planets, and to test numerical algorithms that predict the degree of polarization of extrasolar planets.

To illustrate the wavelength and phase angle dependence of the degree of polarization of light that is scattering within an Earth-like planetary atmosphere, Figure 5-15 shows the flux of sunlight that is reflected by the Earth itself and Figure 5-16 shows the degree of polarization of sunlight that has been scattered within the Earth's atmosphere. Both observations have been performed by Global Ozone Monitoring Experiment (GOME) instruments of ESA: the reflected flux has been observed from space, and the degree of polarization (Figure 5-16) from the ground (it thus pertains to scattered light that is transmitted through the atmosphere). Note that neither of these observations represents a disk-integrated signal (what one will obtain for an extrasolar planet). The observations do, however, show the variation that can be expected for observations of Earth-like extrasolar planets.
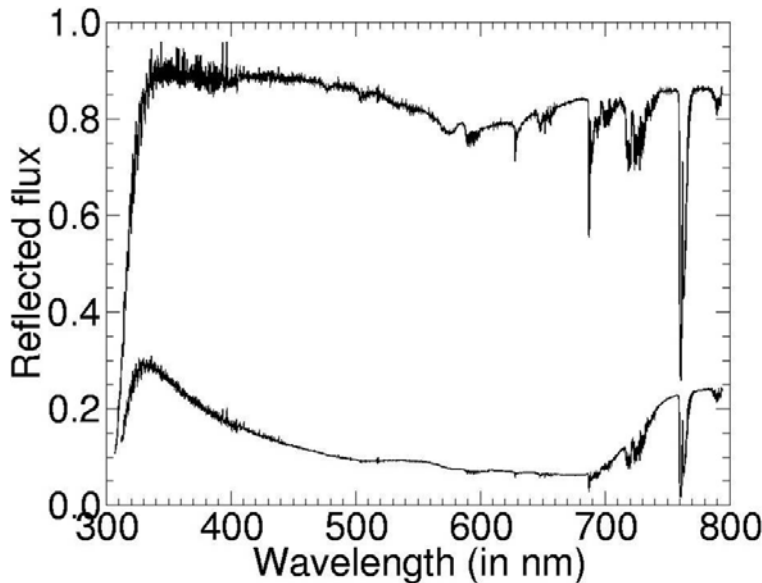
Figure 5-15. Flux reflected by a completely cloudy pixel (upper curve) and a cloud-free pixel (lower line) over vegetation on Earth as observed by GOME onboard ESA's ERS-2 satellite. The pixel size is 40x80 km2. In both cases, the solar zenith angle is about 34° and the viewing angle is 0° (nadir view). The fluxes are normalized to the incoming solar flux.

The curves in Figure 5-15 and Figure 5-16 can be thought of to be composed of a continuum with superimposed high-spectral resolution features. The continua in Figure 5-15 are mainly determined by the (almost wavelength independent) reflective properties of the clouds in case of the cloudy pixel and the surface albedo in case of the cloud-free pixel. The surface was mainly covered by vegetation: the small bump between 500 and 600 nm reflects the green color of the vegetation, and the high reflectivity above 700 nm is associated with reflection within plant cells. This high reflectivity is usually referred to as the "red edge" of vegetation. The high reflectivity at the shortest wavelengths is due to light scattering by gaseous molecules ("Rayleigh scattering"). Below about 330 nm, both curves show the strong so-called Huggin's absorption bands of ozone. The Chappuis absorption band of ozone is visible in the cloudy pixel case as the broad, shallow depression between 500 and

700 nm. Both curves show abundant high-spectral resolution features, most of which are due to absorption by atmospheric gases. For example: absorption bands of oxygen are found near 630 nm, 690 nm, and 760 nm. Absorption bands of water vapor can be seen around 590 nm, 650 nm, 740 nm, and 790 nm. The noise-like, narrow peaks that are in particular visible below 400 nm, coincide with Fraunhofer lines and are due to rotational Raman scattering by molecules, an inelastic scattering process.
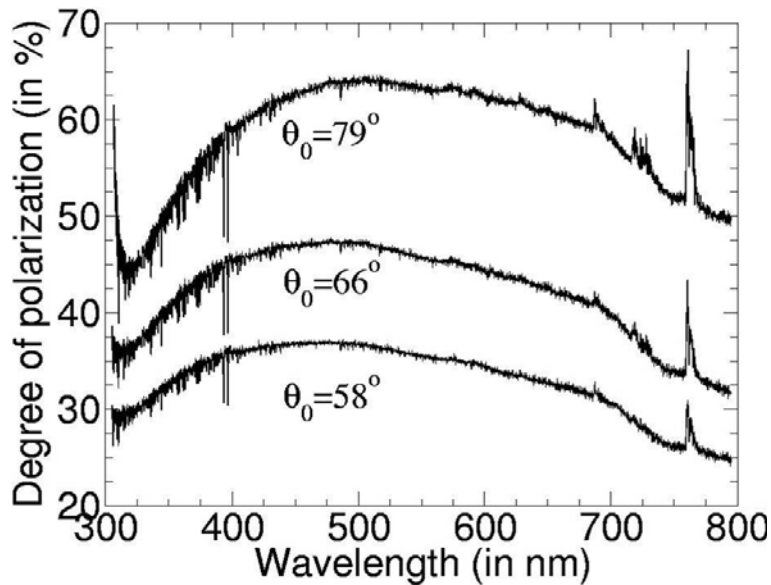
Figure 5-16. The degree of polarization of the cloud-free zenith sky, for solar zenith angles equal to 79°, 66°, and 58°. These geometries are comparable to planetary phase angles α of, respectively, 101°, 114°, and 122°.

The continua of the degree of polarization P (Figure 5-16) show a strong dependence on the solar zenith angle: the larger the solar zenith angle, the larger the maximum value of P. This solar zenith angle dependence is due to the scattering angle dependence of the degree of polarization of light that is singly scattered by gaseous molecules: this single scattering P equals zero at scattering angles equal to 0° and 180°, and it is maximum, close to 1.0 or 100%, at a scattering angle of 90°. In Figure 5-16, a solar zenith angle of 79° corresponds to a single scattering angle of 79°, which implies a relatively high degree of polarization for the light that is singly scattered by the gas, whereas the solar zenith angle of 58° corresponds to a single scattering angle of 58°, and thus to a much lower P for the light that is singly scattered by the gas. Note that multiple scattering generally lowers the degree of polarization.

Each curve in Figure 5-16 shows that at the shortest wavelengths, P is relatively low. This is due to multiple scattering of light between gaseous molecules (the strong increase of P with decreasing wavelength below 310 nm is due to the strong absorption by ozone, which significantly reduces the multiple scattering of light). With increasing wavelength, the scattering cross-section of the molecules decreases (as λ^{-4}), and as a result, the continuum P

increases. With decreasing molecular scattering cross-section, however, the scattering optical thickness of the atmosphere decreases, too, and as a result, more light is reflected by the surface below the atmosphere. Because the surface reflection generally depolarizes light, the continuum *P* decreases with increasing wavelength at the longer wavelengths. Note that above 700 nm, *P* appears to decrease even stronger than at shorter wavelengths. This is due to the strong surface reflection associated with the "red edge" of vegetation. Apart from scattering by molecules and reflection by the surface, the continuum *P* is also shaped by scattering by atmospheric aerosol. This type of scattering, however, does not show a clear spectral feature or behavior, and is thus difficult to point out in the figure. The aerosol contribution becomes clear when one tries to fit the observed continua with only (single and multiple) molecular scattering and surface reflection, which is not sufficient.

The curves in Figure 5-16 show abundant high-spectral resolution features that correspond with the features in the reflected flux (cf. Figure 5-15). In particular, in Figure 5-16, the gaseous absorption bands of ozone (the Huggin's band below 310 nm), oxygen, and water vapor show up. These bands show up mainly because absorption of light reduces multiple scattering and hence increases *P* (light scattering by different types of aerosol at different altitudes in the atmosphere also influences *P* in the absorption bands, see [52] .

Rotational Raman scattering shows up in *P*, too, just like in the reflected flux spectra (see Figure 5-15). In *P*, the solar Fraunhofer lines appear as sharp dips in the continuum, because light that is inelastically scattered in a rotational Raman scattering process is less polarized than the elastically scattered light [47].

5.13.4 Polarization simulations of terrestrial-type extrasolar planets

For our numerical simulations of light that is reflected by terrestrial-type extrasolar planets, we assume planets that are symmetric with respect to the planetary scattering plane and unpolarized incoming starlight (hence, *U*=0). The terrestrial-type planets have homogeneous, layered atmospheres that are bounded below by a homogeneous, flat surface. The atmospheric layers contain gaseous molecules, and, optionally, aerosol and/or cloud particles. We use an adding-doubling radiative transfer algorithm [54][48] in combination with a disk-integration algorithm to calculate the flux vector of the reflected and observable light for a given model planet and planetary phase angle. Numerical simulations of light reflected by giant, gaseous extrasolar planets can be found in [53].

Figure 5-17 shows the degree of polarization *P* for planets with black surfaces and atmospheres with various (scattering) optical thicknesses as functions of wavelength. The atmospheres contain only non-absorbing gaseous molecules. The planetary phase angle α is 90°. The model atmosphere with an optical thickness of 10 (at 0.55 μm) is representative for the Venus atmosphere, that with an optical thickness of 0.1 (at 0.55 μm) for the Earth atmosphere, and that with an optical thickness of 0.01 (at 0.55 μm) for the Mars atmosphere. As can be seen in the figure, for all atmospheres, *P* is lowest at the shortest wavelengths. This is due to the large amount of multiple scattering that takes place at these wavelengths. *P* increases with wavelength for all atmospheres, because the atmospheric optical thickness decreases with wavelength, and the multiple scattering thus decreases, too. For the thinnest atmospheres, *P* increases with wavelength up to its single scattering value of almost 0.96. For the thickest atmospheres, even at the longest wavelengths of our spectral interval, the

multiple scattering is still significant. Note 1) although multiple scattering can strongly reduce *P*, the minimum value of *P* is still about 0.28 (28 %) at this phase angle, 2) the degree of polarization is independent of the number of photons that is received: a high value of *P* is generally related with a low number of photons!
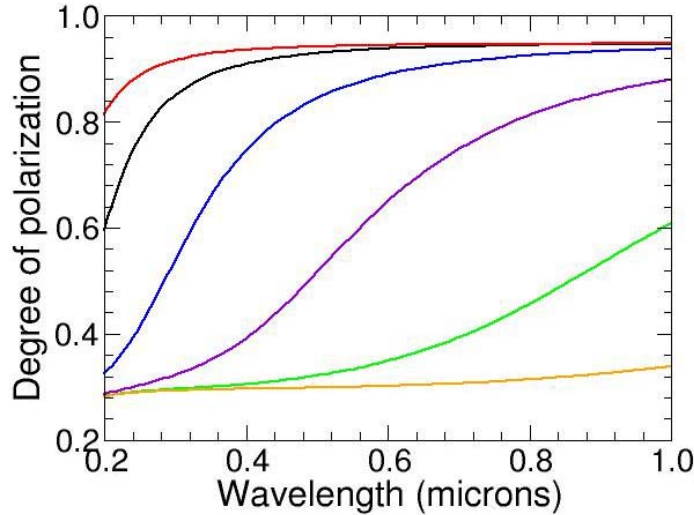
Figure 5-17.Polarization of black planets with atmospheres of various optical thicknesses (at 0.55 microns): 100 (orange), 10 (green; Venus), 1 (purple), 0.1 (blue; Earth), 0.01 (black), 0.001 (red; Mars). The planetary phase angle α is 90°.

Figure 5-18 shows the degree of polarization of extrasolar planets with Earth-like atmospheres, i.e. the ambient pressure and temperature within the 16-layer model atmospheres are values as measured for the Earth. The model atmospheres contain gaseous molecules, including ozone, oxygen, and water. Absorption bands of these 3 gases are taken into account, using typical, Earth-like, ozone and water vapor mixing ratios (that vary with altitude). Oxygen is well-mixed within our atmosphere, and its mixing ratio is assumed to equal 21%. One of the model atmospheres contains a thick, low altitude, water cloud with an optical thickness of 10 at 0.55 μm and consisting of Mie-scattering cloud particles. The model atmospheres are bounded below by reflecting surfaces. We adopt two surface types: ocean (a flat, Fresnel reflecting surface with a black ocean body below) and vegetation (a flat, completely depolarizing surface with a wavelength dependent reflectivity as measured for mixed forest). We thus perform our numerical calculations for 4 types of homogeneous planets: an ocean planet (an ocean surface and no clouds), a vegetation planet (a vegetation surface and no clouds), two cloudy planets (one with an ocean surface and complete cloud cover and one with a vegetation surface and complete cloud cover).

Figure 5-18 The calculated polarization at a phase angle of 90 degrees for a 100% ocean planet (black), a 100% vegetation planet (green), a 100% cloudy planet (blue), a 70% ocean + 30% vegetation planet (orange), and the latter planet with 40% cloud cover (red).

All curves in Figure 5-18 can be thought of to consist of a continuum with superimposed high-spectral resolution features (see also Figure 5-16). The continuum polarization is determined mainly by the (wavelength dependent) atmospheric scattering optical thickness, and the surface reflectivity, while the high-spectral resolution features are due to gaseous absorption bands. Note that our numerical simulations do not include rotational Raman scattering, our spectra thus do not show the narrow dips at the shortest wavelengths.

At the shortest wavelengths, P is high for all model planets, because absorption by ozone strongly reduces the multiple scattering of light. With increasing wavelength up till about 0.32 microns, P decreases, because the ozone absorption decreases, and hence the multiple scattering increases. For the ocean planet, the continuum P then starts to increase with wavelength, because the atmospheric optical thickness decreases and because there is almost no contribution of unpolarized light from the surface (which is black). For the vegetation planet, the continuum P first increases with wavelength, just like for the ocean planet, but at wavelengths larger than about 0.35 microns, P shows a general decrease with wavelength, because at these wavelengths, unpolarized light reflected by the surface starts to contribute to the total reflected light. The spectral signatures of vegetation (e.g. the red edge) clearly show up in the spectrum of P. For the completely cloudy planet, P simply decreases with increasing wavelength because the light reflected by the thick cloud is practically unpolarized. Note that the continuum P will depend on the type (size, composition) of the cloud particles, on the optical thickness of the cloud as well as on its cloud top altitude (for a cloud as thick as ours, P is virtually independent of the surface reflectivity).

In reality, terrestrial-type extrasolar planets are not expected to be homogeneous. To simulate heterogeneous planets, we simply take a weighted sum of the reflected fluxes F and Q, and then recalculate P. Figure 5-18 thus also shows P for a planet with 70% of its surface covered by ocean and 30% by vegetation (similar as on Earth). Clearly, P of this mixed planet still shows spectral signs of the vegetation. When we include a coverage of 30% clouds (evenly distributed over the ocean and the vegetation), only a slightly steeper curve at the lower wavelength end of the red edge indicates the presence of vegetation. Note that the thickness of our cloud is not representative for the average cloud thickness on Earth; with most clouds, some spectral signature of the surface will show through.

All curves in Figure 5-18 show absorption bands by atmospheric gases. As explained earlier, these bands show up in (high-spectral resolution) polarization spectra, mainly because the absorption decreases the multiple scattering and hence increases P.

Figure 5-19 shows the degree of polarization of the 3 homogeneous and the 2 heterogeneous planets in Figure 5-18 as functions of the planetary phase angle α. For this figure, the spectra have been integrated across the I-band, thus from 0.64 to 0.9 μm. The I-band is a typical band the ZIMPOL instrument will perform observations in. Figure 5-19 clearly shows that P varies strongly with α. In particular, P equals zero (for all planets) at $\alpha=0°$ (the dayside of the planet is in full view) and at $\alpha=180°$ (the night side of the planet is in full view), simply because of symmetry. These phase angles are anyway not interesting for direct imaging of extrasolar planets, because there the planet is too close to its star.

For the cloud-free, ocean planet, P is highest around $\alpha=90°$, because there the degree of polarization of light singly scattered by gaseous molecules is highest. For the other (homogeneous and heterogeneous) planets, the surface reflectivity and/or the clouds lead to a low P across all phase angles. It should, however, be remembered that degrees of polarization as small as 0.001 (0.1%) can be measured. Degrees of polarization of a few percent, as shown in Figure 5-19 , will thus pose no problem for polarimetry. The curves indeed show that with polarimetry a distinction between the various types of planets could be made.

The phase angles in Figure 5-19 range from 0° to 180°. The actual range of phase angles a given extrasolar planet will be observable at, depends strongly on the orbital inclination angle i. In case the orbit is viewed edge-on ($i=90°$), the phase angle ranges from 0° to 180°, whereas in case the orbit is viewed face-on ($i=0°$), the planet can only be seen at a phase angle of 90°. In general, with an orbital inclination angle i, the phase angle ranges from 90°-i to 90°+ i. Thus, independent of the orbital inclination angle, a planet will be at a phase angle of 90° at least twice every orbit. This is very advantageous for polarimetry.

Figure 5-19 The degree of polarization of the 5 terrestrial-type planets of Figure 1-4 as functions of the planetary phase angle α (spectrally integrated across the I-band, i.e. 0.64 to 0.9 microns).

5.13.5 Summary

Observations of solar system planets and (disk-integrated) numerical simulations show that the degree of polarization of light reflected by extrasolar planets can be several (tens) of percents. The degree of polarization of the reflected light depends strongly on the planetary phase angle (with the highest values generally occurring at intermediate phase angles, around 90°), as well as on the composition and structure of the planetary atmosphere and the reflection properties of the underlying surface, and thus on the wavelength. Polarimetry can thus be used both to detect an extrasolar planet (the polarized signal can be distinguished form the direct, unpolarized starlight) and to characterize it.

5.14 Data calibration and analysis

Instruments for OWL will be physically larger and heavier by almost an order of magnitude than instruments for the VLT. The scientific goals for OWL require that its instruments achieve levels of performance (contrast, stray light suppression, spatial and spectral resolution, stability) substantially beyond what is achievable today. As a result of these requirements the instruments will be very complex and technically demanding. In order to achieve the science goals it will be absolutely essential that excellent calibration is achieved for all scientific observations during standard operations. This will only be possible by having full understanding and full control over the physical configuration of the instrument. The mode of operation for OWL observations has not been fully developed. It may differ substantially from current operations that are characterized by a roughly equal split between

visitor and service mode. Some science goals of OWL may be best served by allocating the telescopes for extended periods to a few projects, possibly only a single one. Such a methodology might be similar to the approach used for big experiments in laboratory and particle physics. Regardless of the exact operational approach adopted it is safe to assume that all scientific goals will be best served by making calibration an integral part of the instrument design and operations, already in the planning phase.

Baseline assumption: The instrument should not be considered primarily an observational tool but a physical experiment for which we need to strive to eliminate or at least properly describe and understand all sources of relevant systematic and random errors.

5.14.1 Use of instrument physical modeling techniques:

The high complexity of OWL instrumentation and the need to fully control and describe the instruments physical properties over extended periods of time make it appear unlikely that the required performance can be achieved by applying standard empirical calibration methodology.

Instead the approach of choice is to use instrument physical modeling techniques. In this the engineering information used to build the instrument is being used to describe its actual physical configuration and from that to assess its actual performance.

ESO's Instrumentation division (INS) has recently enhanced its capability to develop models for its instruments. The effort will build on the ample experience gained at ECF (Giacconi, 2005, Ann. Rev. Astr & Astroph, 2005, 43, 1) for both space-based (FOS, STIS) and ground-based (UVES, CRIRES) spectrographs (e.g. Ballester & Rosa, 1997, A&A 126, 563). In order to achieve the full capability an integrative approach is required.

As an example, the current package of the Instrument Physical Modeling Team (IPMT) consists of:

- Optical ray trace of instrument
- Imaging module to produce simulated 2D spectra (order and line shapes) taking into account the slit width and detector geometrical properties
- Dedicated software package to compare model with actual calibration data
- Optimizer to establish the exact instrument configuration under varying environmental conditions

An integral part of the approach is the provision of high quality input data; items of particular relevance are: a proper characterization of the material properties of the refractive elements as well as a detailed understanding of the output of the calibration sources and the relevant calibration standards.

Note that the physical instrument model should begin with the design phase. It will then accompany the development of the instrument through its phases of testing, verification and operations. An extensive interaction between the modeling and the laboratory verification is essential in order to achieve the full capabilities of the approach and to ensure that that excellent calibration can be achieved early on. It will be a key ingredient to maximize the performance of EPICS and to minimize the effort required for routine operations.

5.14.2 Calibration sources:

For routine operations we advocate the use of calibrations sources avoiding astrophysical calibrators whenever possible. The main advantages are:

- more reliable and repeatable
- independent of environmental conditions (atmosphere)
- less burden on observing time
- if time dependent then in a measurable manner (e.g. lamp aging)
- selection of calibration sources can be optimized for the scientific (e.g. line density, blending – spectral resolution)

The practical points above are relevant for any instrument but gain special significance when one considers an instrument such as CODEX for which an accuracy of few 10 cm/s and the corresponding stability over more than decade is envisaged. For planet finding and characterization, the field that EPICS is going to excel in, another aspect of calibration is particularly important and somewhat unusual in astronomy.

5.14.3 Calibration for intrinsically time dependent phenomena

The main scientific goal of EPICS is finding and characterization of extra-solar planets. Excellent calibration of each observations will be vital not only because we will be studying features at low S/N ratios in many cases but also because the study of extra-solar planets involves intrinsically time dependent effects such as:

- Orbital motion
- orbital phase,
- rotation,
- seasons,
- atmospheric flows, weather, clouds, dust,
- variable surfaces, oceans, continents, vegetation

Exotic chemistry of the atmospheres may result in unexpected properties and spectral features. High quality calibration will again be key to ensure proper identification of these features and to avoid pitfalls e.g. in identification of biomarkers.

Therefore the standard procedure of verifying critical observations by repeating them may or may no be useful since the relevant properties can be time dependent.

In some cases changes as a function of time will be at the heart of the scientific result: e.g. common proper motion as result of physical connection or changes in relative position as result of orbital motion. In such cases calibration of each individual observations will be essential in order to derive a quantitative result.

5.14.4 Calibration of the quasi-static speckle pattern

The challenge of planet detection close to bright stars is the discrimination from quasi-static speckles remaining after differential imaging techniques. These speckles mainly originate

from the combined effects of optical imperfections of telescope, instrument common path and very critically of differential aberration after splitting the beams for differential imaging techniques. Fortunately, speckles have different properties than real objects that can be exploited for the discrimination and should be supported by the instrument design and the data calibration strategy:

- Chromaticity [43]. The position of a speckle scales with wavelength while a planet would be at the same position (not speaking of atmospheric dispersion) at all wavelengths. The effect significantly elongates speckles which are located at angular distances greater than $\lambda^2 / (\Delta\lambda * D)$ from the star.
- Symmetry [55]. High Strehl ratio speckle patterns show a significant amount of symmetry, which is apparent from Figure 5-20.
- Shape [56]. Figure 5-20 also demonstrates that speckles do not necessarily have the point-like shape of a planet. In this context it is interesting to note that many near-by stars will be resolved by a 100-m telescope observing at ~ 1 micron with an angular resolution of about 2 mas FWHM. For example, the angular diameter of the sun at a distance of 10 pc is about 1 mas, while Vega has an angular diameter of 3 mas. Therefore, speckles from the star would be larger than the planet.
- Time dependency. Depending on the location of the aberration origin, quasi static speckles change their position and intensity during the observations on rather short timescales (~minutes). If aberrations originate from the telescope, they will tend to rotate with the telescope pupil. If they originate from inside the instrument, they will tend to be fixed on the detector independent of the field orientation. This behavior is demonstrated by Figure 5-21. A planet would rotate with the field.
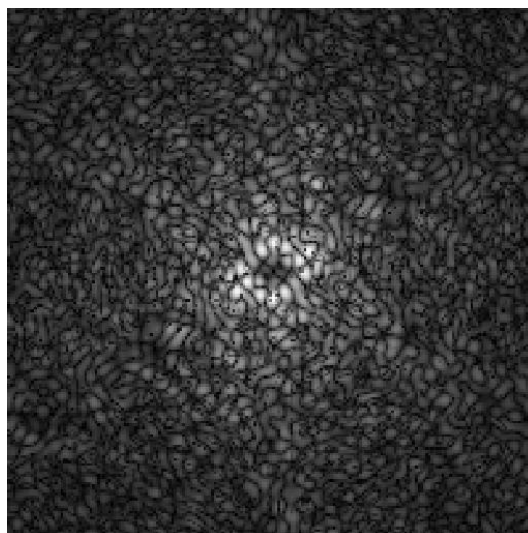
Figure 5-20. Residual static speckle pattern after a perfect coronagraph and differential imaging with small differential aberrations. Note symmetry and shape of speckles.

Figure 5-21. Quasi-static speckle pattern. The images show the high spatial frequency content of long exposures taken with NACO at the VLT separated by about one hour.

5.15 Virtual Observatory Aspects

Extra-solar planet research currently is still a very young field of astronomy. It will be in a much more mature state by the time EPICS will become operational. While EPICS holds great potential for discovery, it will also make important contribution towards the characterization of extra-solar planets. Research with EPICS will make use of data provided by missions such as HARPS, NACO, COROT, GAIA etc. Other large facilities operating at the same time will be e.g. VLT 2nd generation instruments, VLTI, JWST, Darwin, TPF-C, SIM. In this context it is important to note that our understanding of the physical properties of extra-solar planets and the environments in which they form and evolve, will benefit greatly by combining data sets from different sources and across a wide range of the electromagnetic spectrum. Progress in the field will require capabilities which are mostly very similar to the ones currently being developed for the Virtual Observatory (VO).

Baseline assumption: The design of EPICS observing modes and data taking as well as OWL modes of operations should comply with standards of the VO in order to facilitate the exchange of data between various archives. To this end it will be essential to closely follow the evolution of these standards as guided by the International Virtual Observatory Alliance (IVOA). For issues specific to extra-solar planet research attention should also be given to the recommendations of the joint ESA-NASA Planet Finding Data Archive Working Group (PFDAWG).

5.15.1 Virtual Observatory Alliance (IVOA):

The VO is an international effort by the astronomical community, one in which ESO has played a leading role from the beginning. Its main goal is to allow global electronic access to the available astronomical data archives of space and ground-based observatories and sky survey databases. It also aims to enable data analysis techniques through a coordinating entity. In 2002 the existing VO projects formed this coordinating entity, namely the International Virtual Observatory Alliance IVOA (http://www.ivoa.net). It sets standards - after discussion in expert working groups - for data exchange and procedures that will be presented to the International Astronomical Union (IAU) for endorsement. Crucial standards

are being defined at this point and it will be important to closely monitor the evolution to make sure that EPICS and its data products will be developed towards these standards. This will be best achieved by interfacing with ESO's Virtual Observatory Systems Department.

5.15.2 The Planet Finding Data Archive Working Group (PFDAWG):

ESA and NASA have set up a working group dedicated to establishing common formats and procedures for databases related to extra-solar planet research. While many aspects of this work will be similar to the issues addressed by the VO, there is a significant number of issues specific to the field or to individual missions such as Darwin and TPF.

The working group consists of about 8 persons from both sides of the Atlantic, including one representative from ESO, and it will have its first meeting in November 2005. It will identify the relevant needs and will make recommendations to both space agencies. For EPICS it will be important to be aware of these developments for space missions planned to be operational contemporaneously.

6 Top Level Requirements

The TLR are focused on the detection of extra-solar systems in a later evolutionary stage and in particular Earth-like planets for a large enough statistically meaningful sample of stars. The following Figure 6-1 illustrates the requirements in contrast for different types of planets as a function of angular separation.
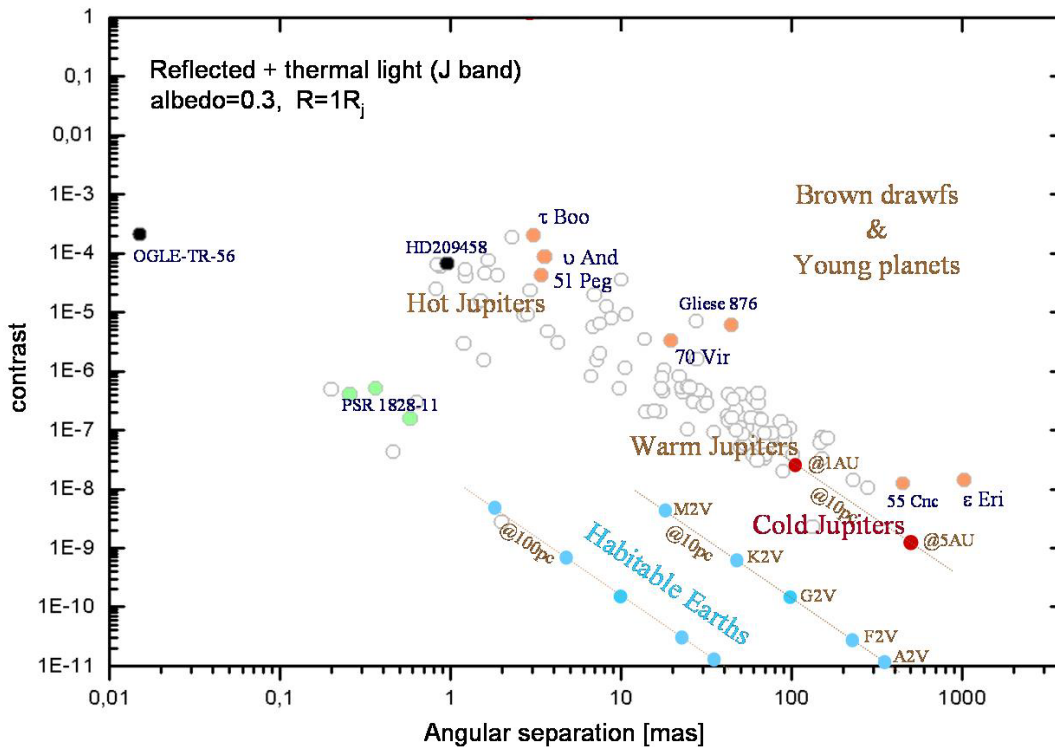
Figure 6-1. Contrast versus angular separation for different types of planets. By courtesy of Olivier Lardiere.

The requirements are subdivided into requirements on the EPICS instrument, observational and the telescope as well as on a proper calibration procedure. The instrument TLR are to be used by the system group to develop an instrument design that fulfills the TLRs assuming a certain set of observing conditions. In this sense the TLR on site and telescope are a result of the EPICS study and therefore cannot be provided yet.

6.1 Instrument, general

- **TLR_1: The instrument covers the wavelength range 0.6 – 1.75 micron**

For the detection of terrestrial planets the wavelength range 700 nm – 900 nm is most interesting with its O_2 band. Gas giant spectra are dominated by the methane features in J- and H-band. Wavelengths shorter than 700 nm might be interesting because of a higher degree of polarization.

- **TLR_2: The total field of view in all observing modes is at least 2″ in diameter at visible wavelengths and 4″ in diameter in the NIR.**

A field of view of 2″ in diameter corresponds to 1 AU at 1 pc, large enough to cover terrestrial planets at all distances. The bigger field in the NIR accounts for the larger separation at which giant planets are searched. The 4″ field covers the solar system (apart from Neptune and Pluto) at distances larger than 10 pc.

- **TLR_3: The inner working angle in all observing modes working at visible wavelengths is smaller than 30 mas (goal 15 mas).**

This inner working angle corresponds to 0.3 AU at 10 pc, small enough to cover the solar system at 10 pc or to resolve Earth out to 25 pc. The atmospheres of close-in and therefore bright giant planets could be studied at very small angular separations.

- **TLR_4: The spatial sampling will fulfill at least the Nyquist criterion (precise value TBD) at all working wavelengths**

The re-scaling and subpixel-shifting of the differential techniques require image interpolation. The interpolation leaves residuals that depend on the sampling and limits the speckle reduction. Therefore, the pixels-scale has to be taken into account for the error budget and a trade-off has to be found.

6.2 Main observing modes and performance

- **TLR_5: There will be a low resolution differential spectroscopic mode covering at least the following lines**

- O2 at 760nm,   R = 150 (see Figure 6-2)

- CH4   in J- and H-band,   R > 15

- H2O   between the bands, R > 15

- CO2   in H-band,   R > 15

Figure 6-2. O₂ A-band at high spectral resolution.

- **TLR_6: There will be a broad band differential polarimetric mode. The absolute polarimetric accuracy will be 1%**

- **TLR_7: The relative astrometric precision in all main observing modes is better than ~100 μarcsec (goal 10 μarcsec)**

The instrument should support relative astrometry (if permitted by the SNR) at this level to allow for the determination of the main orbital parameters of the planets and the detection of unseen satellites by photocentre-wobble. However, the photocentre wobble would only be detectable in the most extreme case for almost equal mass planet and satellite.

- **TLR_8: The photometric (absolute/relative) precision in all main observing modes is better than 1%**

The instrument should support absolute photometry (if permitted by the SNR) at this level to allow for observations of planetary rotation, weather and phase variations, satellites (transit and eclipses)

- **TLR_9: Earth up to 20 pc at a phase angle of 90° is detected in all main observing modes at SNR > 5 in one night of observation**

Properties of Earth at 20 pc: Contrast 2e-10, m_V = 30.6, angular separation 50 mas.

- **TLR_10: Jupiter up to 20 pc at a phase angle of 90° is detected in spectroscopic mode at SNR > 50 in one night of observation**

Properties of Jupiter at 20 pc: Contrast 1e-9, m_V = 28.8, angular separation 250 mas.

6.3 Adaptive Optics related

- **TLR_11: The AO control radius is larger than 0.4″ at 800 nm (goal 0.8″)**

This control radius corresponds to about 1 AU at 2.5 pc, and ensures that – besides for the Alpha Centauri system – the prime targets are inside the control radius. Note that the control radius is given by the $\lambda/(2d)$, where λ is the observation wavelength and d is the actuator pitch of the deformable mirror. This TLR corresponds to an actuator pitch of ~0.2 m (goal 0.1 m).

- **TLR_12: AO limiting magnitude for achievement of TLR: Good for > 100 stars of the spectral types G, K and M.**

The following plots show the number of stars as a function of distance from Earth listed in the NSTAR database (http://nstars.arc.nasa.gov). As one can see on the right plot a telescope at low latitudes (here Paranal) will be able to observe ~50% of the stars when the zenith angle is restricted to 30 degrees. In order to have access to about 100 stars of the spectral types G, K and M, one has to observe out to

- 25 pc for G-stars ($m_V \approx 7$)
- 20 pc for K-stars ($m_V \approx 8.5$)
- 15 pc for M-stars ($m_V \approx 9 - 16$ for M0 to M5). There are about 50 M-stars with $m_V<10$ and 100 M-stars with $m_I<10$ observable at low latitudes.
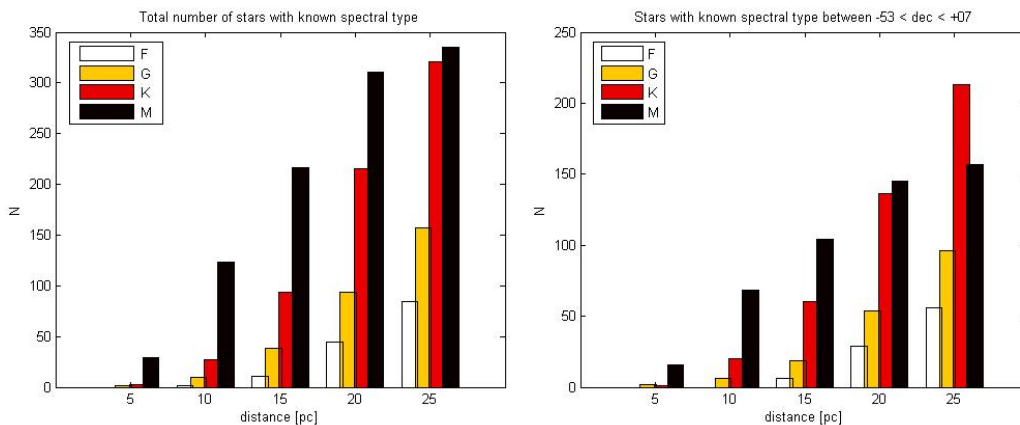
Figure 6-3. Number of stars versus distance from Earth for different spectral types

6.4 Secondary (goal) requirements

- **TLR_13: EPICS data will comply with Virtual Observatory (VO) format**

- **TLR_14: Simultaneous search for terrestrial planets and gas giants will be possible**

- **TLR_15: There will be an NIR high spectral resolution (R ~ 50.000) mode for characterization of giant planets**

This mode would be well suited for radial velocity measurements and the analysis of atmospheric composition and dynamics of close-in giant planets. The contrast between a

Jupiter mass planet at 0.5 AU and its star is around 10^-7 roughly corresponding to the stellar residual. For 10 pc distance from Earth assuming a G2 star, its magnitude would be around 22.5 and the photon flux at resolution 50.000 would be about 0.5 photons per second and spectral bin (16% optics transmission). Therefore, a reasonably high SNR for the high resolution spectroscopy appears feasible in observing times of a couple of hours.

6.5 Calibration

- **TLR_16: Physical model of instrument (Internal calibration sources, functions for calibration of time-dependent effects) required**

7 EPICS Concept and preliminary system analysis

In this section, we present a notional concept of the EPICS instrument. Due to lack of time, no integrated opto-mechanical design could be produced. Moreover due to the numerous issues still to be solved in all the areas of the instrument, adaptive optics, coronagraphy, polarimetry, differential imaging, integral field spectroscopy, a complete opto-mechanical design would be very premature. Nevertheless, some sub-system designs have been proposed in the various chapters of this document describing the instrument parts. In this section we concentrate on the common path optics that shall include the AO system correctors. We developed a preliminary optical design for the purpose of having a starting framework in order to investigate some crucial issues of ground-based ELTs for high contrast imaging applications.

EPICS concept is based on the experience of the VLT Planet Finder Phase A study. The core of the instrument is based on an "eXtreme"Adaptive Optics system of very high performance that feeds three main instruments, each one equipped with a coronagraph:

- a wave-length splitting Differential Imager
- an Integral Field Spectrograph
- a Differential Polarimeter.

This choice reflects the concepts that have been developed for VLT-Planet Finder. Due to science goals that are significantly different from VLT-Planet Finder, the EPICS concept is based on somewhat different choices on the system level that are described in section 7.1 .

7.1 Overall instrument concept

The EPICS concept should be compatible with the detection of both gas giants and rocky planets. Due to different locations of the spectral and polarimetric features of these two groups of planets, different channels split in the spectral domain are needed. Each scientific channel will be equipped with its own coronagraph.

- The **R band** is dedicated to the Polarimetric Differential Imager and to the follow-up observations for the detection of O_2.

- The **J band** will be equipped with a differential imager using pairs of filters that will be sensitive to both CH_4 and H_2O absorption bands.

- The **H band** will be equipped with an Integral Field Spectrograph. The main features that can be detected in this band are CH_4 and CO_2 and H_2O.

- The **I band** is reserved for wave-front sensing. This band has been chosen because of the lesser scientific interest for planet detection. Moreover its location, spectrally speaking, between the visible and NIR instruments, is optimal with respect to important atmospheric chromatic limitations for XAO on ELTs.

The EPICS concept is summarized in Figure 7-1. The baseline concept for the XAO system is described in details in section 8.3.

Possible concepts for the coronagraph are studied in section 9. The three different instruments developments are presented in sections 10 (Differential Imager), 11 (Integral Field Spectrograph) and 12 (Differential Polarimeter). The performances of EPICS in terms of Signal-to-Noise Ratio for exo-planet detection are analyzed in section 13.

In the sub-sequent sections we will concentrate more on the common path optical implementation with the goal of studying the impact of several parameters on the problem of systematic errors that are a crucial problem for Earth-like planets detection. We limited the study to the problem of chromatic differential errors. The problem of systematic polarimetric errors is introduced in section 12.
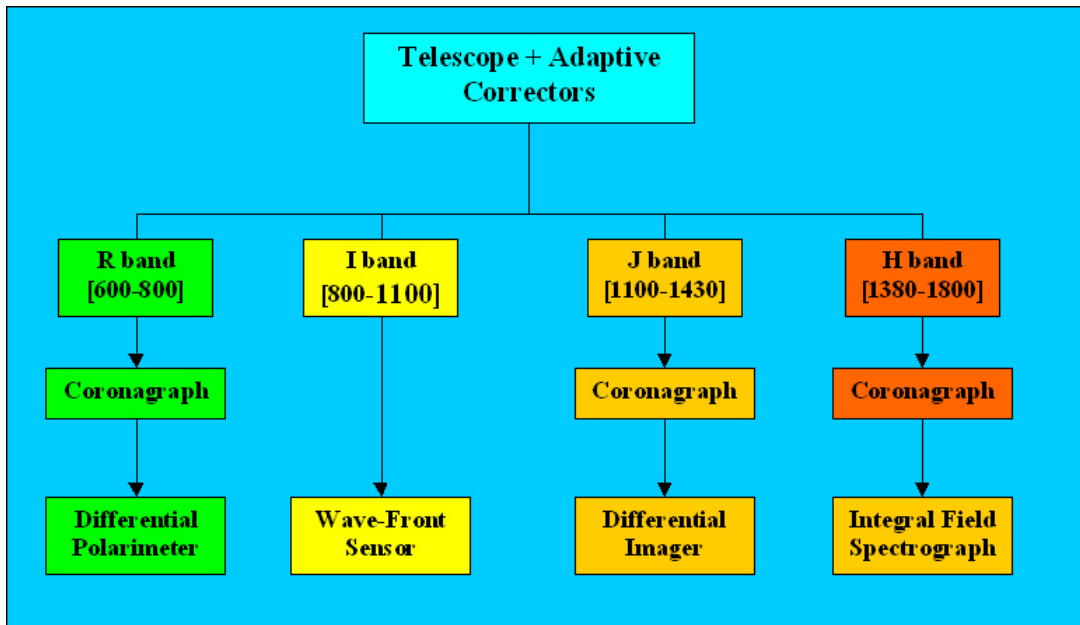
Figure 7-1: EPICS will be composed of three spectral channels for the scientific instruments and one for wave-front sensing.

7.2 Preliminary optical design of the common path

A preliminary optical design of the common path (XAO system only) has been studied with ZEMAX in order to address some critical aspects that the system should fulfill. A baseline optical design and some possible options are given in section 7.2.1. Atmospheric dispersion correction is addressed in section 7.2.3. Section 7.3 explicates some of the concept choices and analyses its performance in terms of differential chromatic aberrations and ultimate contrast achievable by wave-length splitting differential imaging. Some optical specifications are derived from this analysis.

7.2.1 Common path optics

The common path optics includes essentially an ADC and the XAO correction devices. We recall the baseline concept for the XAO that is based on a double stage (see Figure 7-2 and section 8.3 for detailed description).
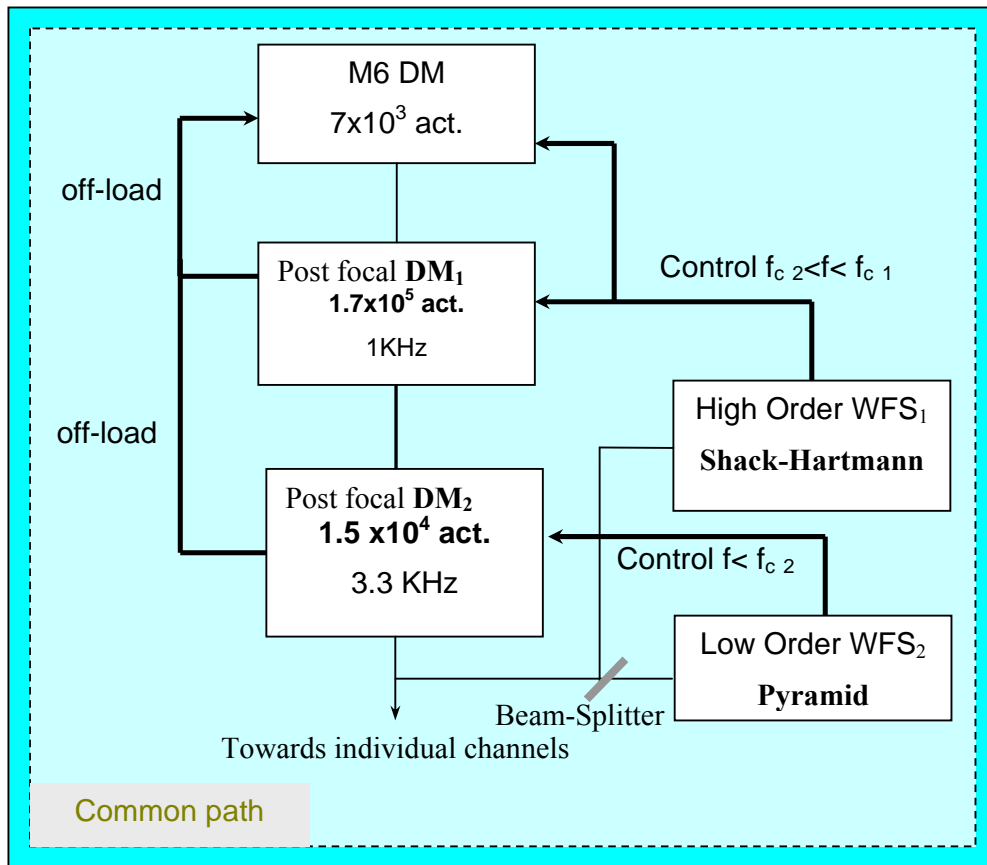
Figure 7-2: Common path XAO concept for EPICS.

The M6 DM is part of OWL telescope. The role of the common path optical system is to include the two post-focal adaptive DMs DM_1 and DM_2 after the F/6 focus. These two DMs differ by the number of actuators. For the sake of the design we assume MEMS technology with a fixed physical inter-actuator spacing of 0.5 mm for both mirrors. The sizes of the deformable mirrors are:

- DM_1: $1.7 10^5$ actuators, (500x500) : $\varnothing$ **250 mm**.
- DM_2 : $1.5 10^4$ actuators, (150x150): $\varnothing$ **75 mm**.

The main constraints on the optical set-up are:

- The DMs should be located in a pupil plane.

- The ADC should be as upstream as possible for the reason explained in section 7.3.

- The optical quality should be better than SR>90 % for 2 arcsec field (goal 4 arcsec)

The general view of EPICS common path optics included in OWL telescope is shown in Figure 7-3, and detailed description is given in Figure 7-4 and Figure 7-5.
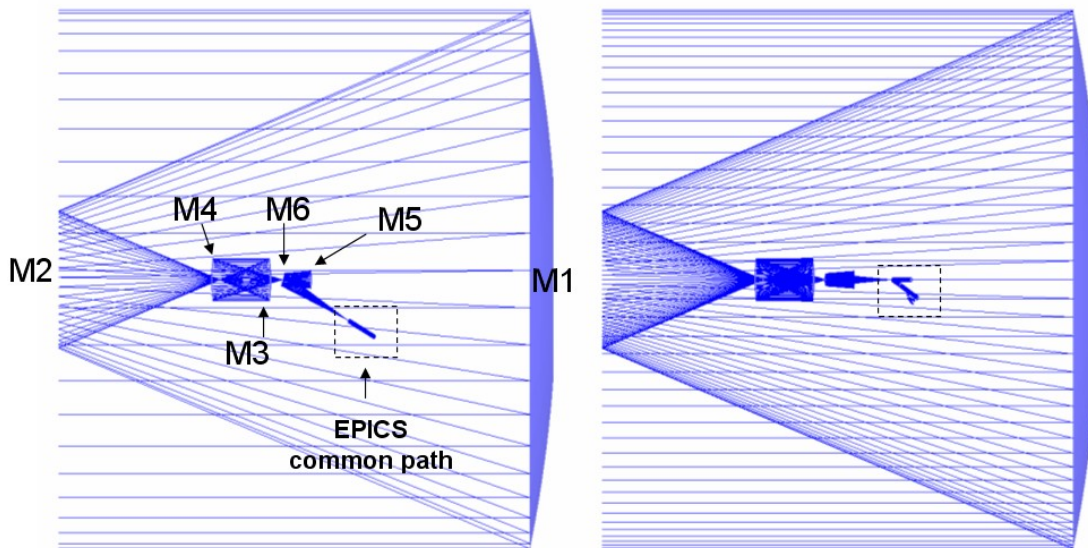
Figure 7-3: EPICS common path location in OWL optical set-up. The right image is 90 deg rotated along the telescope optical axis with respect to the left image.

The common path optics is placed directly at the F/6 OWL focus in the instrument cage. The first optical element is a 750 mm parabolic collimator (M7) forming a pupil plane on the ADC after reflection by a folding mirror (M8). This folding mirror is with a central hole taking advantage of the large central obstruction avoiding the use of an off-axis parabola at this stage. The size of these optical elements is justified in section 7.3. Other options with smaller optics are also briefly described in section 7.2.2.

A second parabolic mirror (M9), identical to M7 in order to cancel field aberrations induced by M7, refocuses the beam and a folding mirror reflects the beam on the side towards the optical system for the inclusion of the DMs. This optical system is based on off-axis parabolic mirrors (2 for each DM) with reflection angles lying in two perpendicular planes in order to minimize polarization.

Note that the two-off axis parabola M11 and M13 for the inclusion of DM_1 (M12) are used with reflection angles such that they compensate the field aberrations introduced by a single off-axis parabola. This special configuration is needed here because the field of view in entrance of the first parabola is large when considering the long focal ratio of OWL.

A similar set-up is defined for DM_2 (see Figure 7-5). A total of 11 reflexions compose the common path. The details about the components can be found in Table 7-1.

Some more sophisticated configurations may exist, involving for example powered mirror also for the DMs but have not been investigated. Some alternatives are described in section 7.2.2.
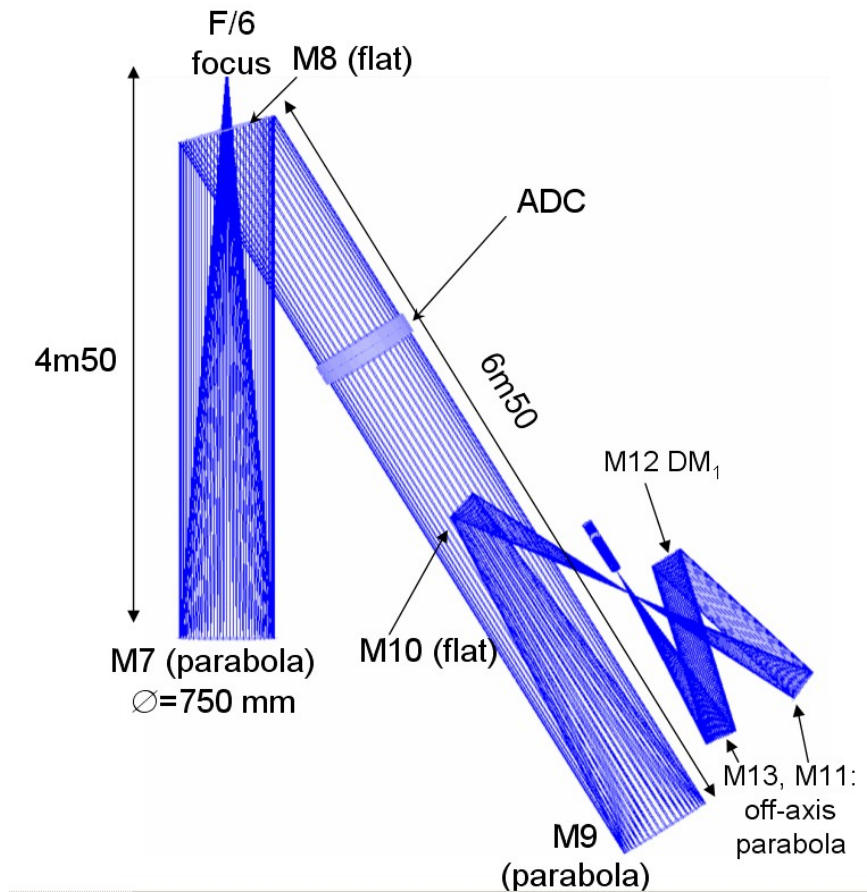
Figure 7-4: Zoom on EPICS common path at F/6 focus.
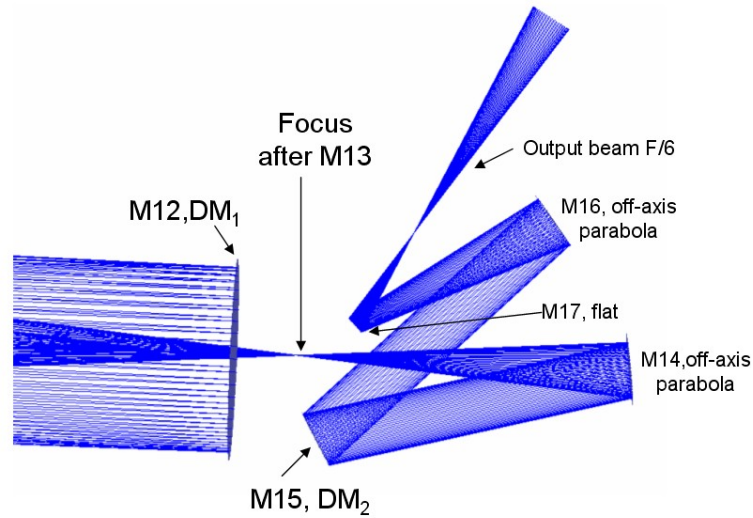
Figure 7-5: Zoom of optics for DM₂ implementation. The view is as Figure 7-4 seen from the right.

Figure 7-6: Spot diagrams for a 4 arcsec field. Wave-length is 1 micron. The field curvature is about 200 mm. Field points are 1 in the center (top left) , then from left to right 4 points at corners at 0.5 , at 1 arcsec separations and 3 points at 2 arcsec separations.

Figure 7-7: Spot diagrams for a 4 arcsec field with correction of field curvature. Wave-length is 1 micron.

The spot diagrams at the output focus for a 4 arcsec field of view are shown in Figure 7-6. The images are diffraction limited up to 1.5 arcsec separation (86 % Strehl) and deteriorate at the border of the field (2 arcsec separations, 68 % Strehl).

Thus to be able to work up to 4 arcsec, some field distortion correction should be provided after the coronagraph for instance (applies essentially to the Differential Imager).

The spot diagrams taking into account field distortion correction (by the instruments) is given in Figure 7-7.

Item	Radius of curvature (mm)	Diameter (mm)	Off-axis distance (mm)
M7	9000	750	-
M8	∞	750	-
ADC	∞	750	-
M9	9000	750	-
M10	∞	250	-
M11	3000	260	388.2
M12 (DM$_1$)	∞	250	-
M13	3000	260	388.2

M14	900	80	116.5
M15 (DM$_2$)	∞	75	-
M16	900	80	116.5
M17	∞	30	-

Table 7-1: Optical elements parameters. All mirrors are off-axis parabola except flats and DMs

7.2.2   Alternative optical designs

Some variants of the baseline design (Figure 7-4) have been produced. The main difference is in the size of the collimator and subsequent ADC, which can be reduced if needed. For example, M7 can be reduced to 250 mm as well as the ADC and the second parabola. However, in this case the relative chromatic beam shift (see section 7.3.2) is larger and the optical quality needs to be improved.

A detailed sensitivity analysis is needed to fix the optimum sizes of M7, M8 and M9 and the ADC.

Two other designs have been produced and permit to save 3 reflections:
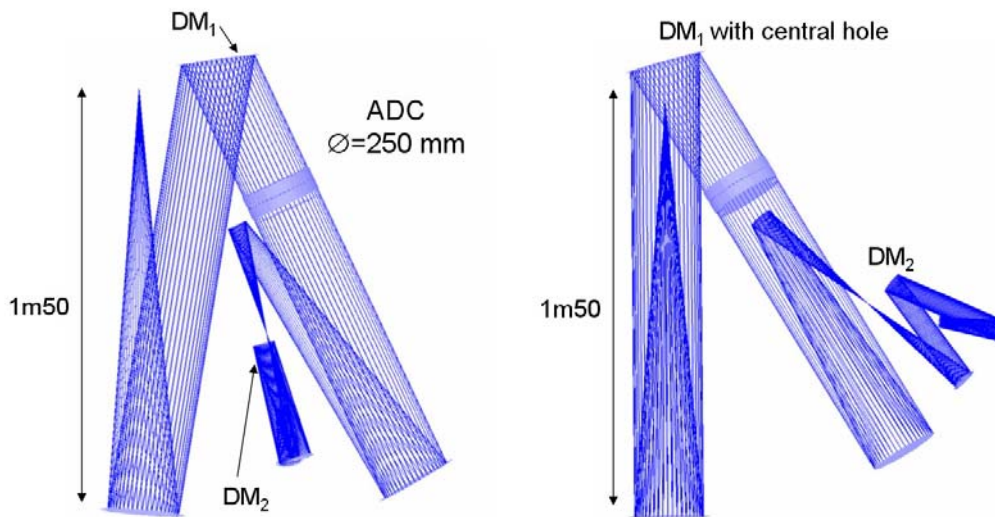
Figure 7-8: Alternative solutions. Solution 1 (left) is based on an off-axis first collimator. Solution 2 (right) is based on an on-axis collimator and assumes an adaptive mirror with a central hole.

The first one (left in Figure 7-8) is based on a first collimator that is an off-axis parabola. The problem with this set-up, is that the field aberrations are quite large so that, because of atmospheric dispersion, the images after correction by the ADC are affected by important wave-length dependent aberrations (see Figure 7-9). An on-axis (solution 2) case of this design permits to solve this chromatic error problem but involves here a adaptive mirror with

a central hole. The possibility of producing such a MEMS device is questionable, but could permit an easy implementation.
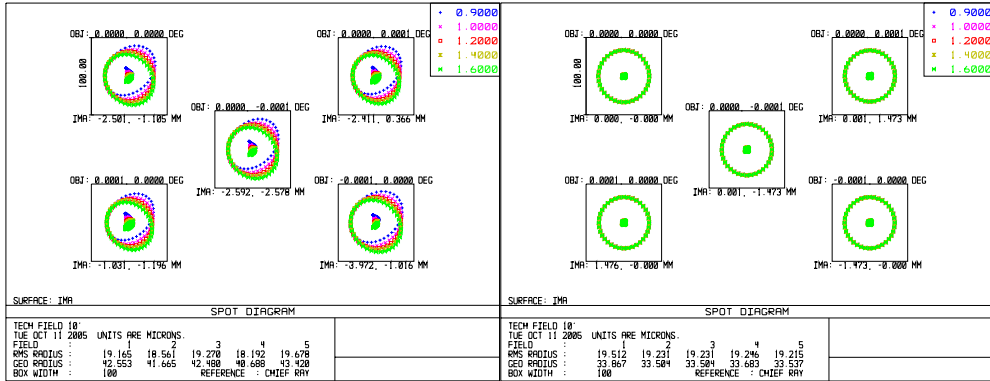
Figure 7-9: Spot diagrams for the two alternative solutions. Zenith angle Z=50 deg.

7.2.3 Atmospheric Dispersion Correction

Atmospheric dispersion correction for ELTs is a crucial issue. For the moment it is not clear if the existing models are still valid at the milli-arcsec level. It could be that an adaptive instead of a passive component could be needed. This issue will be studied within the ELT design study FP6 framework. Here we assume the classical laws apply and just produced an ADC for the purpose of this preliminary optical design.

The baseline solution is optimized for the wave-length range covering the WFS central wave-length (900 nm) to the H band (1750 nm). This permits to reject most of the chromatic errors due to beam shifts as shown in section 7.3.2. Each component of the ADC is 100 mm thick (50 mm per glass) and 750 mm diameter.

The two glasses are LASF32 and LAKN12.

The angles of each surface are respectively, 0 deg, 1.44 deg, 0.25 deg.

The transmission of this ADC is 93 % in I, 85 % in J and 80 % in H.

```
3D LAYOUT
LENS HAS NO TITLE.
FRI OCT 14 2005
```

Figure 7-10: The atmospheric dispersion corrector based on two glasses: LASF32 and LAKN12.


```
OBJ: 0.0000, 0.0000 DEG

+   0.9000
×   1.0000
□   1.2000
✳   1.4000
✳   1.6000
⊕   1.7500

200.00

IMA: 0.000, 0.014 MM
SURFACE: IMA
                    SPOT DIAGRAM
LENS HAS NO TITLE.
FRI OCT 14 2005  UNITS ARE MICRONS.
FIELD      :        1
RMS RADIUS :    16.933
GEO RADIUS :    30.581
AIRY DISK  :    70.78        REFERENCE  : CHIEF RAY
```
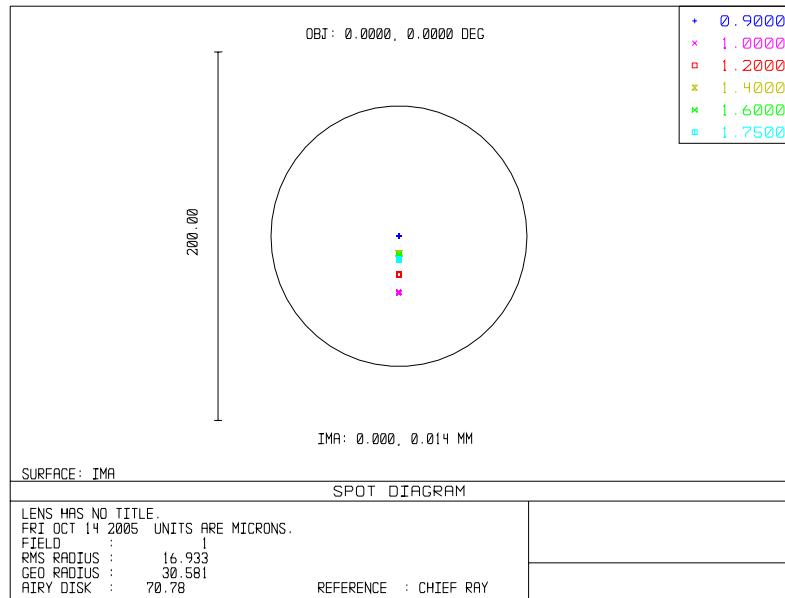
Figure 7-11: Spot diagram after ADC compensation for Z=50 zenith angle.

A solution with 3 glasses for ADC correction from visible to NIR has also been studied but led to a quite thick component with low transmission in the NIR. Further developments are needed to see if a wide wavelength range correction is feasible. Hence for the moment, the ADC for the visible is supposed to be downstream and dedicated to visible only. Since the differential polarimeter is insensitive to chromatic aberrations this is not an issue. However for follow-up observation for O_2 detection for example with a dedicated IFS or DI, the problem of chromatic differential aberrations (see section 7.3) needs to be solved and an

ADC upstream optimized in the visible needs to be implemented. No simultaneous observation in visible and NIR would then be possible, but this is not an issue in the case of a follow-up observation.

7.2.4 Transmission budgets

Some preliminary transmission budgets have been estimated.

The WFS budget that has been used for the simulations is detailed in table 6-2.

An error budget for the scientific channel transmission is given in Table 7-3.

Item	Transmission	Intermediate total
OWL 6 mirrors at 0.985	0.913	
AO common path 11 reflections 0.98	0.800	0.73
ADC	0.93	0.679
BS	0.980	
2 Lenses (2 ref) + pyr double (4 ref) + pupil modulator 1 = 0.98^11	0.800	0.532
CCD optimized for I 90%	0.900	0.479
Excess noise LLLCCD	0.500	0.240
Contingency 30 surfaces 1%	0.740	0.177
TOTAL	0.177 (excess noise included)	

Table 7-2: Transmission budget for the WFS path.

Item	Transmission
Atmosphere 0.95	0.95
OWL +AO+ADC	0.45
Coronagraph apodization	0.6
Additional optics (instrument)	0.9
Detector QE	0.7
TOTAL	0.16

Table 7-3: Provisional transmission budget for the scientific channels

7.3 Analysis of systematic and differential chromatic errors.

The analysis presented in this section results from numerous discussions inside the differential imaging group. We made the choice to put this analysis in the concept part since it affects the system as a whole and in particular it has an impact on the final performance of both the Differential Imager and the 3D Spectrograph.

The purpose of differential imaging is to attenuate the speckle noise in the long exposure images taking advantage of differences in the planet flux in two adjacent wave-lengths and thus to attain the required contrast. By splitting the beam in two channels (either through two dichroic filters in the Differential Imager or through two channels in the 3D Spectrograph), one introduces inevitably differential aberrations that translate into speckles in the final images that don't disappear after the subtraction process. More problematic is the static or quasi static nature of these speckles that remain whatever the integration time.

The problem is complex and for a complete assessment, the overall calibration strategies and data processing should be defined and properly taken into account.

We make here an analysis of the main error sources and derive specifications in order to get the required ultimate contrast by dual imaging without further processing or calibration. The requirements derived in this section can thus be considered as a goal with some possible margin which will depend on the performance of special calibration strategies one can use to further increase the contrast (deconvolution, reference subtractions, rotation of key optical elements for static speckle smoothing, etc.)

In sections 7.3.1 and 7.3.2 we analyze the effect of chromatic errors occurring before the coronagraph (related to atmospheric differential refraction) and those occurring after the coronagraph, in the wave-length splitting process. Section 7.3.3 introduces the problem of amplitude errors that has an important impact on the ultimate contrast achievable with a coronagraph. In section 7.3.4, we try to combine the effects in an error budget and derive some first system specifications of some key elements for the fulfillment of the Top-Level-Requirements.

We analyze now the two main error sources in differential chromatic errors and show how they mix with the static part of the wave-front provided by the AO system. A schematic view of the problem is given in Figure 7-12.

(All values on optical quality and aberrations are expressed in terms of wave and not mechanical surface. Perfect coronagraph is assumed.)

Figure 7-12: Schematic representation of the issue of differential chromatic aberrations. φ dif, shift is the differential aberration caused by chromatic beam shift due to differential refraction. φ dif, filter is the differential aberration caused by wave-length splitting in the instrument (DI or 3D spectro). φstat is the static uncorrected part of the wave-front (due mainly to non common path between the WFS and the coronagraphic optics) at the entrance of the coronagraph.

All the results presented in this section are related to the case of observation in J band in the water band. Results for other wave-lengths will differ (requirement in H band is slightly relaxed) In the visible, the requirements may be slightly more stringent (on the one hand differential refraction is more important but on the other hand spectral bands are narrower) and have not been analyzed in detail.

7.3.1 Differential aberrations occurring after the coronagraph

The effect of differential aberrations occurring after the coronagraph and their amplification by the static part of the wave-front has been studied in great details in ref. [33]. We recall here the main results.

Let's call ϕ_{stat} the static aberration after AO before the coronagraph, ϕ the AO residuals with zero mean, and $\phi_{dif.,filter}$ the differential aberrations between two channels (due to imperfections of the dichroics for example).

$$A_1 = \Pi\left(\sqrt{E_c} - \exp(i(\phi + \phi_{stat}))\right)$$

$$A_2 = \Pi\left(\sqrt{E_c} - \exp(i(\phi + \phi_{stat}))\right) \cdot \exp(i\phi_{dif.,filter}) \qquad \text{eqs.1}$$

$$I_1 - I_2 = \left\langle \left[\hat{A}_1\right]^2 - \left[\hat{A}_2\right]^2 \right\rangle = 2\,\text{Im}\left[\left(\hat{\Pi} \otimes \hat{\phi}_{stat}\right) \cdot \left(\hat{\Pi} \otimes \hat{\phi}_{stat} \otimes \hat{\phi}_{dif.,filter}\right)^*\right]$$

where A_1 and A_2 are the complex amplitudes corresponding to the two beams propagated in each channel of dual imager. Π is the pupil function. The symbols $\wedge$ and $\otimes$ represent respectively the Fourier Transform and the convolution product. E_c is the coherent energy.

In the equations above, the complex amplitudes and the intensities have been developed and the assumption of zero mean for ϕ has been used. The remaining term in the last line of the above equations involves only ϕ_{stat} and $\phi_{dif.,filter}$ and shows how these two terms mix together to influence the final contrast. Note that the static part of the wave-front after the coronagraph has not been taken into account. Its effect should in principle only be a global redistribution of light. First estimation show that about 20-30 nm rms precision for the re-imaging optics before the scientific CCD should however be required (TBC).

We can see that the final effect on the contrast, given by the intensity difference distribution I_1-I_2 shows a quadratic dependence with ϕ_{stat} and a linear dependence with $\phi_{dif.,filter}$. This illustrates the huge importance of the static part of the wave-front. In ref. [33], the authors use assumptions for the static aberrations (about 20 nm rms on the low and mid-spatial frequencies) that are similar to the requirements for VLT-PF instrument. Their conclusions are that this imposes requirements on the differential errors after the coronagraph that are extremely tight (10 pm).

It is clear that for EPICS, with the ultimate goal of Earth-like planets detection, the static part of the wave-front after AO needs to be lower by more than one order of magnitude than the assumption in ref. [33].

For the AO system defined in section 8.3.1.1, it makes no sense to allow a static error as big as several tens of nanometers since for this system, optimized for high halo rejection in the region of interest, the halo in the center of the field -of-view (20 mas to 200 mas corrected by the second stage) would just be completely dominated in terms of intensity by the 20 nm static errors speckles. For the same reasons *one wouldn't equip his race car with bicycle wheels*, we chose a completely different strategy to take the best advantage of the AO system, and to be able to correct for the static errors to a very low level compatible with the primary science goal. Hence, in EPICS (see section 8.3.1.2), each scientific channel is equipped with an active mirror with a focal plane WFS in order to calibrate the non common path errors on a reference source and bring them down as much as possible. It is expected from our strategy a much better performance than the usual way of reference measurements subtraction form the AO WFS data.

As shown in the following, to fulfill the primary science goal, the specifications of some key optical elements and of the active corrector are very tight, but not impossible to achieve, and are discussed in section 7.3.4.

For the modeling we assume that the PSD of the static error in front of the coronagraph is flat, as it can be expected from a wave-front corrected by a focal plane WFS (similar noise

propagation as a phase sensor). The PSD of the differential chromatic errors after the coronagraph are supposed to be in f^{-2}, typical of polishing errors. This PSD could probably be different for high precision optics, but in our simulation we didn't find a strong dependence on the PSD.

In order to make an analytical model of the effect of chromatic errors after the coronagraph we found a quite useful approximate expression, under the hypothesis of a flat PSD for the static aberration, to the eqs.1:

$$\left| I_1 - I_2 \right| \approx \left| \hat{\phi}_{stat} \right|^2 \cdot \sigma_{dif} \qquad \text{Eq.2}$$

where σ_{dif} is the standard deviation of the phase (in radians) at the observation wave-length corresponding to the differential aberration after the coronagraph. Eq. 2 simply expresses that the level of the contrast after dual imaging is equal to the product of the halo that would be produced by the static aberration alone, $(\left| \hat{\phi}_{stat} \right|^2)$ by the standard deviation in radians of phase of the differential aberration after the coronagraph. A Fourier-based simulation was used to confirm this approximation (see Figure 7-13).
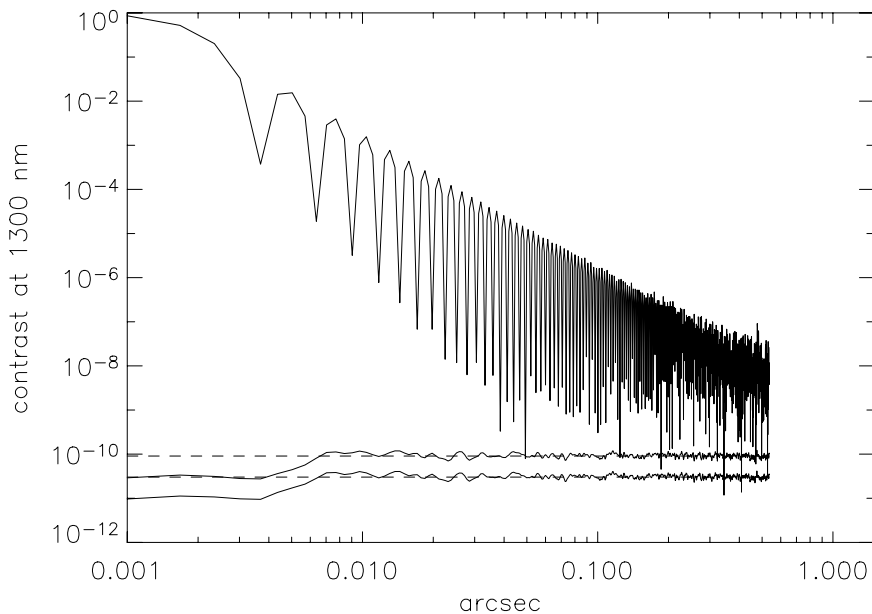
Figure 7-13: Illustration of the effect of differential aberrations after the coronagraph and comparison to theoretical value: 1 and 3 nm rms differential error, f^{-2} PSD. The static error in front of coronagraph is 0.3 nm (10 c/p < f < 75 c/p, flat spectrum) corresponding to a static halo level of 1.25×10^{-9} at 1300 nm. The upper curve is the PSF without coronagraph. Solid line is simulated contrast after dual imaging. Dashed line is theoretical contrast after dual imaging.

7.3.2 Differential aberrations occurring before the coronagraph

The well-known effect of atmospheric differential refraction occurring when observations are made away from the zenith creates also chromatic differential aberrations, which revealed to be even more effective than the one described in section 7.3.1.

Atmospheric differential refraction causes the star to appear at shifted positions depending on the wave-length (see Figure 7-14). In the optical train, the effect is that the beams for different wave-length follow different optical paths. Optics located in the pupil plane see all the beams hitting the surface at the same position, so that all wave-length are affected by the same optical surface defects. For optics not located in a pupil plane, especially for those close to an image plane, the beams are shifted with respect to each other proportionally to the differential refraction angle (see Figure 7-12 on top). In case of a small beam shift, which will almost always be the case, the differential error is equal to (for a beam shift in the x direction):

$$\Delta\phi = \phi(x) - \phi(x + \delta p) \approx \frac{\partial\phi}{\partial x}\delta p$$

The beam shift is proportional to the derivative of the phase and to the beam shift δp. This matter of fact has a quite strong implication. There is no 'blurring effect' when considering a continuous range of wave-lengths. Only the brightness of the speckles will change. The same remark stands for the variation of the beam shift with zenith angle, except that the field will rotate. To what extend these effects can be calibrated still needs to be studied.

For the observation case in J band, the differential refraction at Z=30 deg, between two typical wave-lengths of the dichroic Differential Imager (λ_1=1300 nm and λ_2=1385) is 14 mas. This corresponds to a beam shift between the two wave-lengths of 55 microns on the mirror M7 which is the main contributor to the differential chromatic error (when ADC is in function). The rms value of the wave-front difference assuming a 3 nm optical quality for M7, is 68 pm. Even this very low value is sufficient to create a halo of speckles that don't disappear after dual imaging (see Figure 7-15) and are just 5 times lower than the contrast of an Earth-like planet (allowing theoretically a 5 σ detection).

Figure 7-14: Refraction angle for z=60 deg. zenith angle. Temperature T_0=273.15 K, pressure P_0=1013.25×10^2 Pa.
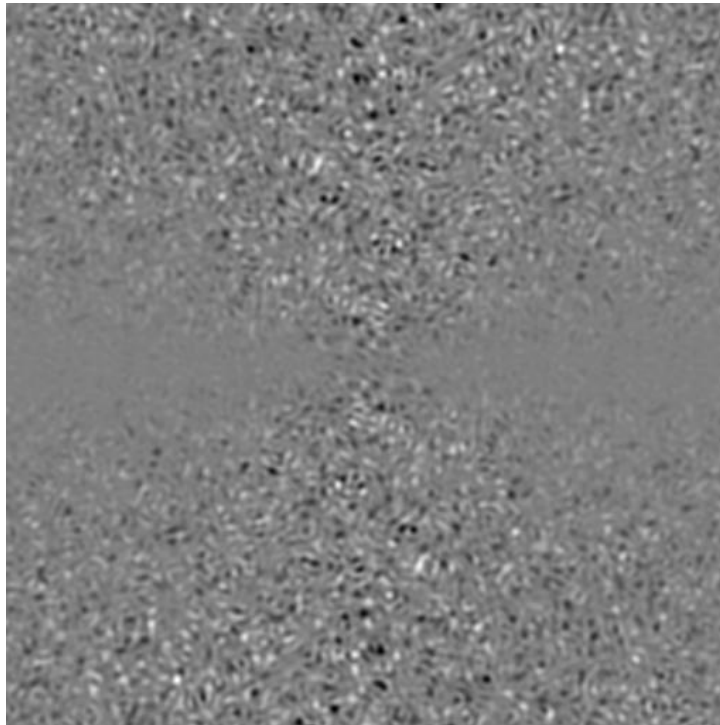
Figure 7-15: Simulation of Differential image in presence of 72 pm (flat PSD f < 250 c/p) differential aberrations before the coronagraph (corresponding to a beam shift on M7 (3 nm optical precision) of 55 microns between λ_1=1300 nm and λ_2=1385 nm for z=30 deg. (14 mas)) . The static error in front of

coronagraph is 0.3 nm (10 c/p < f < 75 c/p, flat spectrum) corresponding to a static halo level of 1.25×10^{-9} at 1300 nm. 5 σ contrast level is 1.5×10^{-10}.

The spectral theory of chromatic beam shift has been described for VLT-PF phase A study in **RD 6**. An important result is that the PSD of the differential error is flat when considering optics phase defects in f^{-2} PSD for typical polishing.

The PSD for polishing errors can be written:

$$PPSD(f) = PPSD_0 f^{-2} \quad \text{Eq.3}$$

where PPSD$_0$ is the value of the phase PSD normalized to the telescope pupil size for a spatial frequency f = 1 cycle/pupil (noted c/p here after).

The differential error power spectral density ΔPPSD for a relative beam shift δp/d where d is the size of the beam on the optics and δp the shift, is given by:

$$\Delta PPSD(f) = (2\pi\delta p / d)^2 PPSD_0 \quad \text{Eq. 4}$$

The PSF structure in high Strehl images has been described in ref. [34]. The main contributor is the so-called halo speckles which intensity can be approximated by the phase PSD. A development similar than in ref. [33] can be done to evaluate the effect of differential aberrations before the coronagraph. Again, atmospheric speckles average out and the remaining terms after dual imaging depends on both the differential error $\phi_{dif., shift}$, and the static error ϕ_{stat}. After development and some approximations one obtains:

$$\left| I_1 - I_2 \right| \approx \left| \hat{\phi}_{stat} \right| \left| \hat{\phi}_{dif.,shift} \right| \quad \text{Eq. 5}$$

Eq. 5 simply expresses that the level of the contrast after dual imaging is equal to the square-root of the product of the halo that would be produced by the static aberration alone, multiplied by the halo corresponding to the differential aberration PSD. Numerical simulations agree with this approximation. The fact that the ultimate contrast is here proportional to the phase standard deviation (and not to the square as in section 7.3.1) implies that the effect is more severe.

7.3.3 Amplitude errors

All the developments done in this section suppose a perfect coronagraph. This assumption means also that no amplitude errors are present in the incoming pupil in front of the coronagraph. In reality, important intensity variations will exist on the scales of the segments size, and variations of more than 1 percent of individual segment reflectivity is expected. The way the segments are replaced plays an important role on the spatial distribution of amplitude errors in the pupil. To evaluate the effect we suppose here a worse case, i.e. completely random distribution of amplitude errors with a flat spectrum, with a cut-off at the spatial frequency 0.6 m^{-1} corresponding to the size of the segments in the primary mirror. This spatial frequency is inside the critical spatial frequency range that affects the separations between 30 mas and 200 mas in NIR, and its effect needs absolutely to be taken into account. Unlike scintillation, the amplitude errors caused by optics are static and are thus contributing directly to the ultimate contrast budget.

We made a simple simulation by generating a random distribution of amplitude errors (see Figure 7-16) and studied the effect on the halo contrast of different levels of reflectivity dispersion amongst the segments. Figure 7-17 displays the results. Those agree with an analytical approximation given by the following formula, based on the translation of amplitude errors in scattered light and assuming a uniform distribution over the field affected by amplitude errors:

$$C_{\delta I} = \frac{\left(\arcsin\left(1 - \sqrt{1 + \delta I}\right)\right)^2}{1.22^2 (2D/d)^2}$$

where δI are the intensity variations standard deviation, D is the size of the telescope and d is the size of the segments in the primary. The theoretical curve corresponding to this formula is displayed in Figure 7-18.

$C_{\delta I}$ will directly contribute to the contrast budget as the static error. In section 7.3.4, the contrast corresponding to the static is specified to be 1.25×10^{-9} in J. $C_{\delta I}$ will directly contribute to this value and this needs to be significantly lower to not constrain even more the requirements on the static wave-front.

From these results we can derive a specification: for example for a contribution to the **contrast of less than 5.10^{-10} the reflectivity variations per segment needs to be less than 0.5 % rms.** This is probably significantly lower than what will permit the ageing. However depending on the sequence with which the segments are replaced, the PSD of the variations could be more gentle and be then acceptable. More detailed studies is needed about the ageing and the effect of the procedure to re-coat the segments to decide whether an amplitude corrector is needed or not. If needed it could be implemented in front of the coronagraph.

Effect of intensity variations due to phase defects on optics out of pupil is another source of errors that needs to be investigated at a later stage.

Figure 7-16: Intensity variation map used in the simulation.
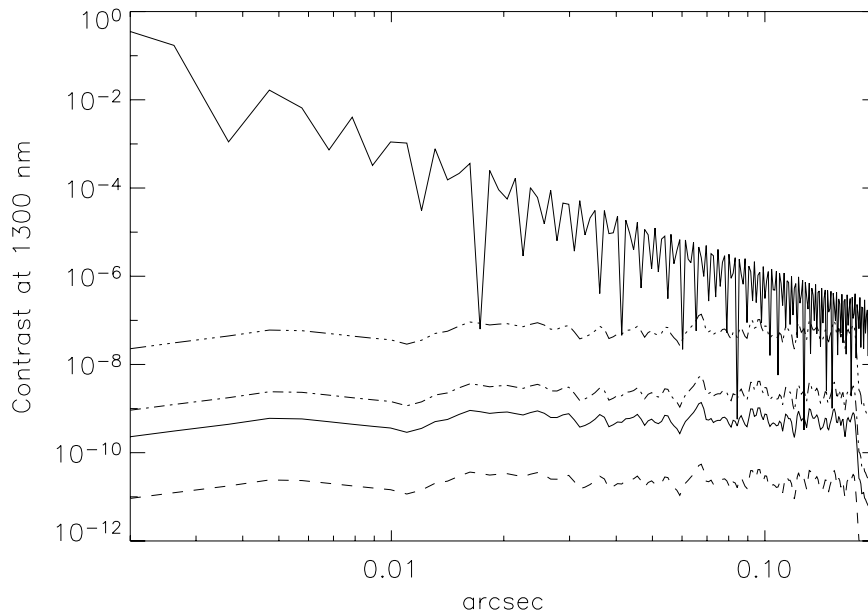
Figure 7-17: Effect of amplitude errors on the contrast. The upper curve is the PSF without coronagraph. Three dotted dashed line: 5% rms intensity dispersion among segments. Dotted dashed line: 1% rms. Solid line: 0.5 % rms. Dashed line: 0.1%.

Figure 7-18: Theoretical contrast in function of intensity dispersion among segments.

7.3.4   Error budget and preliminary specifications

Using the results of sections 7.3.1 and 7.3.2, it is possible to make an error budgeting combining the two independent sources of errors. It was checked by numerical simulations that the two error sources are indeed independent so that the contrast achieved is simply given by:

$$C = \sqrt{C_{before}^{\;2} + C_{after}^{\;2}}$$

where C_{before} and C_{after} are respectively the contrast achievable considering differential aberrations before and after the coronagraph.

7.3.4.1   Chromatic beam shift error budget

Using Eq. 4 we can compute the chromatic beam shift error PSD induced by each optical element of the telescope and the instrument (here limited to the AO system).

Using Zemax, the relative shift sensitivity of the beam for each optical element has been derived and the differential aberration standard deviation corresponding to 14 mas differential refraction angle (Z=30 deg., λ_1=1300 nm and λ_2=1385 nm) has been computed for each mirror (see Table 7-4). **Here we suppose there is no ADC at this stage** (for the purpose of showing its actual utility)**.** In this example, the OWL mirrors are supposed to have a precision of 20 nm rms (similar to a VLT primary mirror) and other mirrors are of 10 nm rms. Note that here the spatial frequency range considered for calculating the rms values

(of optical surfaces and differential errors) is 1 to 250 c/p. Table 7-4 shows that, in the hypothesis of an ADC placed downstream in the optical train (this is for example the case in the LAOG VLT-PF proposal (**RD 7**)), the total chromatic errors is several nanometers. Thus it is larger by almost two orders of magnitude than the requirement of a few tens of picometers (example of Figure 7-15).

In this case the 5 σ contrast level would be not better than 1.5×10^{-8}!

A few solutions exist to this problem. One can imagine for example to use optics with much better quality.

Optical element	Beam Size (m)	Beam shift (mm.arcsec^{-1})	Relative beam shift (arcsec^{-1})	Diff. Error for 14 mas angle (pm)
M1	100	0	0	0
M2	34	0.452	1.3e-5	1.1
M3	8	0.645	8.06e-5	7.1
M4	8	0.078	9.75e-6	0.8
M5	3.2	1.136	3.55e-4	31.2
M6	2.4	0	0	0
M7	0.75	3.896	0.0052	228.7
M8	0.75	1.277	0.0017	74.8
M9	0.75	2.916	0.00388	170.6
M10	0.26	2.937	0.0112	492.6
M11	0.25	2.957	0.0118	520.3
M12	0.25	0.114	0.00045	20
M13	0.25	2.736	0.0109	481.3
M14	0.075	3.01	0.0401	1763
M15	0.075	0.1	0.00133	58.5
M16	0.075	3.310	0.04413	1941
M17	0.032	2.966	0.09268	4076
TOTAL rms (pm) (for f < 250 c/p)				4932

Table 7-4: Beam shifts sensitivity in OWL and AO optical set-up. The last column represents, for each mirror, the differential error for 14 mas differential refraction (Z=30 deg). No ADC is supposed to be implemented at this stage for the purpose of showing its actual utility.

The only way to really tackle this problem is to implement an ADC as up-stream as possible in the optical train and to ensure that the mirrors before the ADC are not contributing too much to the differential errors. It is the solution we consider.

In the baseline design of Figure 7-4, all the mirrors after the ADC will no longer see the chromatic beam shift. The OWL mirrors contribute by a very little amount because of their big size. The only mirrors contributing significantly are M7 and a little bit M8. The specification of the optical quality of M7 and marginally of M8, needs to be such that the ultimate contrast achievable is compatible with the TLRs. In the present case, 3 nm rms optical quality for M7 (and M8) is required to attain $1.5 \ 10^{-10}$ contrast at 5 σ (Figure 7-15).

7.3.4.2 Total error budget and preliminary specifications

In the following we derive some specifications for some key optical elements that are required for the goal contrast. We restrict this preliminary study to J band observation in water bands for Earth-like planets. The specifications shouldn't be very different for H band. In visible, the situation is different and a study is needed, especially there is probably the need of a specific ADC for visible.

For Earth-like planets, (ref. section features Markus), it was shown that the contrast in water bands is about 0.2. We adopt for an Earth-like planet around a G2 star a contrast of $2.2 \ 10^{-10}$ (see 13.4). In this case the searched contrast in the water band is reduced by a factor 0.8.

$$C_{earth} = 1.8 \times 10^{-10}$$

Note that this is the extreme case, the contrast being more favorable for exo-Earths around K and M stars.

We consider a Z=30 deg. zenith angle for the computation of the chromatic beam shifts effects on M7 and M8. The three main parameters we can specify and tune are the following:

- The static aberration output from AO in front of the coronagraph: as shown in section 8.3.1.2, the static aberrations due non-common path between the WFS and the coronagraph paths will be corrected by an adaptive mirror before the observation. The specification will consist on the precision that need to attain the actuation of this mirror and the ability to maintain it. The main limitation of the calibration process is the potential contribution of chromatic beam shift to the static error. Since this one depends on the telescope pointing, it is difficult to calibrate it, even though it is possible. We estimated that the error is however small: the differential refraction between 800 nm (the lowest WFS wave-length) and the NIR (J and H) is about 0.2 arcsec (to be compared to 14 mas in Table 7-4) which translates for the EPICS optical set-up to a chromatic error of 1 nm (Z=30 deg, f < 250 c/p). On the critical frequency range (10 c/p < f < 75 c/p), the error is 0.1 nm, well below the specification defined hereafter. The static aberration requirement can also be expressed in terms of instrumental contrast to be attained during the calibration, which gives a good indication of the level of difficulty.

- The differential aberrations after the coronagraph: essentially the precision of the dichroics and mirrors in the Differential Imager, or the equivalent differential aberrations in a 3D spectrograph (for the latter an extremely low differential aberration (null in theory) could be obtained but important diffraction effects that could seriously limit this capability are still under investigation for VLT-PF also).

- The differential aberrations before the coronagraph: the main contributor is here M7, and so the specification will be the quality of the mirror .

- Note that the ADC will also produce some differential aberrations before the coronagraph due to the optical defects and the glass dispersion. With a precision of 10 nm on the outer surfaces, the contribution is 25 pm.

Combining the whole errors sources, we derived a set of specifications that fulfills the contrast requirements of **1.8×10^{-10} at 5 σ.** The detailed specifications of the most critical optical elements and their contribution to the differential errors are listed in Table 7-5. Note that the segments reflectivity dispersion is not included and is supposed to be corrected.

Error budget item		Optical quality (rms reflected wave)	Contribution to diff. error
Static errors, adaptive mirror precision in scientific channel (PSD f^0 10 <f < 75 c/p)		0.3 nm	(before corono. only)
Differential chromatic errors before the coronagraph	M7 (parabolic) (PSD f^2 f < 250 c/p)	3 nm	69 pm
	M8 (flat) (PSD f^2, f < 250 c/p)	3 nm	22 pm
	OWL 6 mirrors (PSD f^2, f < 250 c/p)	20 nm each	32 pm
	ADC	10 nm	25 pm
Total Contribution to diff. error (before corono. only)			83 pm
Differential aberrations after coronagraph (PSD f^2, f < 250 c/p)		1 nm	

Table 7-5: Example of specification of critical elements and differential aberrations error budget for Earth-like planet detection goal (J band). The 5 σ level contrast with these specifications is 1.8×10^{-10}.

The specifications of Table 7-5 will need to be tuned according to what will technologically be feasible in 10 to 15 years. In order to have an idea of the sensitivity of these specifications, several curves have been produced where one parameter is varied and the others kept constant.

- The sensitivity to the static error before the coronagraph is displayed in Figure 7-19 and in Figure 7-20. Those two figures are essentially the same except that one is expressed in terms of rms error (Figure 7-20) and the other one in terms of equivalent contrast to be attained in the calibration procedure (see section 8.3.1.2). In the case of 0.3 nm (10 c/p f < 75 c/p), this corresponds to a halo level of 1.25×10^{-9} at 1300 nm at separations larger than 10 λ/D. This can be compared to what has already been achieved in visible light (875 nm) in the laboratory experiment HCIT for the TPF-C space mission project [35]: 10^{-9} in monochromatic light at 4λ/D and 5×10^{-9} with 40 nm bandwidth. The real challenge for EPICS is to be able to attain this instrumental contrast inside the OWL focal station. Note that the instrumental contrast must include the coronagraph intrinsic residuals with segmentation effects and amplitude errors but not co-phasing errors which are corrected by AO (see section 8.4.3).

- Figure 7-19 shows that if the instrumental contrast requirement is relaxed to 10^{-8} (control of static errors to 1 nm instead of 0.3 nm for 10 c/p f < 75 c/p) the ultimate contrast achievable is 10^{-9} compatible with Jupiter-like planets.

- Figure 7-21 and Figure 7-22 show the sensitivity to differential aberrations after the coronagraph. It confirms the fact that those can be significantly relaxed and are much less important than the one occurring before the coronagraph (provided the static aberration is sufficiently low, remember the quadratic dependence on this term).

It is important to note that all the specifications derived in this section are considering no further data processing or calibration procedure to increase further the contrast. Taking into account the latter, and using a very precise modeling of the instrument, the most stringent specifications may be somewhat relaxed. For example for the VLT-PF project it has been shown that calibration using reference data from other observations can significantly increase the achievable contrast after data processing (RD 8). The same strategy could be implemented for EPICS. However special care needs to be taken for the treatment of differential aberrations before the coronagraph. Those create speckles with fixed position (the differential error is merely proportional to the wave-front derivative), but their brightness in the field will depend on the pointing direction of the telescope. Very precise instrument modeling and a very stable critical elements in the instrument will then be needed.

All the specifications have been done for J band for water band observations. In H band the chromatic beam shift is relaxed by about a factor of 2. In visible, since for O_2 the separation between two adjacent channels would be of about 10 nm, the chromatic beam shift is of about 9 mas at z=30, so slightly less than in J. However the requirement would be more severe because the feature contrast is not so favorable as for water. The requirements for this

particular case has not been studied, but since only follow-up observations are foreseen in O_2, the knowledge of the position of the planet will help a lot the calibration.
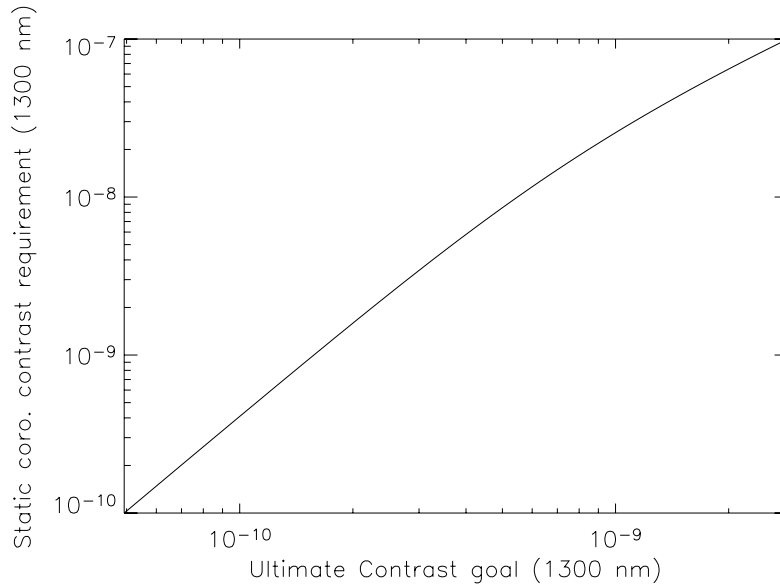
Figure 7-19: Effect of instrumental coronagraphic contrast in J band for the following conditions: Differential aberrations after coronagraph: 1 nm. Differential aberrations before coronagraph: 83 pm (total f< 250 c/p) at z=30 deg., equivalent to M7 (750 mm diameter) with 3 nm rms optical surface precision (10 c/p < f < 75 c/p). Contrast goal is for 5 σ.
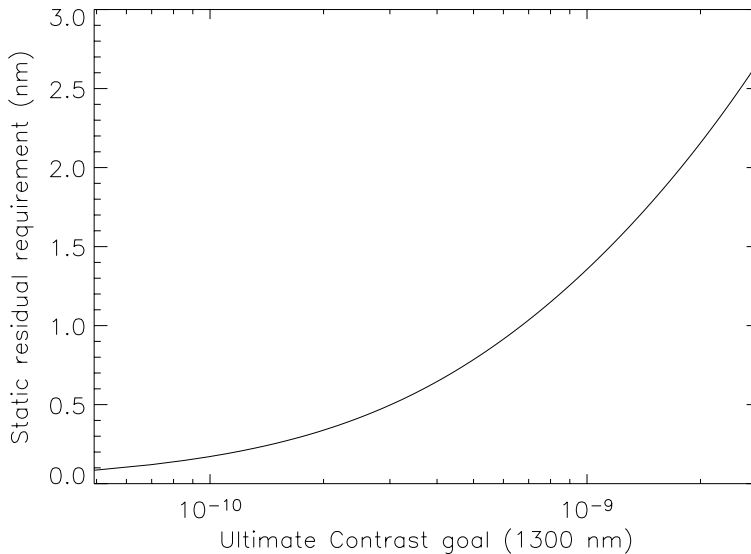
Figure 7-20: Effect of static aberration rms error (10 c/p < f < 75 c/p) in front of coronagraph (supposed perfect) for the following conditions: Differential aberrations after coronagraph: 1 nm.

Differential aberrations before coronagraph: 83 pm (total f< 250 c/p) at z=30 deg., equivalent to M7 (750 mm diameter) with 3 nm rms optical surface precision (10 c/p < f < 75 c/p). Contrast goal is for 5 σ.

Figure 7-21: Effect of differential aberrations after the coronagraph for the following conditions: Instrumental coronagraphic contrast: 1.25×10^{-9} (at 1300 nm), equivalent to 0.3 nm rms static error (10c/p < f < 75 c/p). Differential aberrations before coronagraph: 83 pm (total f< 250 c/p) at z=30 deg., equivalent to M7 (750 mm diameter) with 3 nm rms optical surface precision (10 c/p < f < 75 c/p). Contrast goal is for 5 σ.
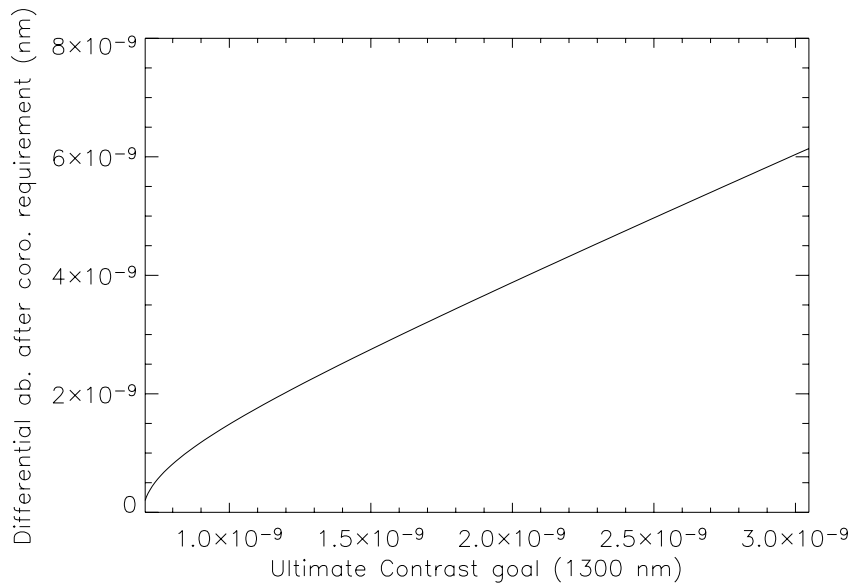
Figure 7-22: Effect of differential aberrations after the coronagraph for the following conditions: Instrumental coronagraphic contrast: 2×10^{-8} (1300 nm) equivalent to 1.2 nm rms static error (10 c/p < f

< 75 c/p). Differential aberrations before coronagraph: 83 pm (total f< 250 c/p) at z=30 deg., equivalent to M7 (750 mm diameter) with 3 nm rms optical surface precision (10 c/p < f < 75 c/p). Contrast goal is for 5 σ.

Figure 7-23: Effect of optical quality of M7 for the following conditions: Instrumental coronagraphic contrast: 1.25×10^{-9} (at 1300 nm), equivalent to 0.3 nm rms static error (10 c/p < f < 75 c/p). Differential aberrations after coronagraph: 1 nm. Contrast goal is for 5 σ.

8 Extreme Adaptive Optics

Single conjugate Adaptive Optics is now a mature technology in Astronomy. The next decade will see a new evolution towards more complicated systems based on the advanced concepts we study now. Among them Ground Layer, Multi-Conjugate, Multi-Objects Adaptive Optics aim at increasing both the sky coverage and the corrected field-of-view. We are here interested in another concept with increasing interest: Extreme Adaptive Optics. At a first glance such an AO system is extreme just in the number of actuators of the deformable mirror one needs to use to get a high Strehl Ratio in corrected images. However, with the increasing interest of high contrast imaging for exo-planets detection, the main goal is to provide not only a high concentration of energy in a diffraction limited core but also to minimize the halo surrounding the star image, in order to be able to get the deepest starlight rejection with a coronagraph. The concept that is proposed in this chapter tries to fulfill these needs. With the concern of designing a system that is not to prohibitive in real-time computing power and WFS detectors requirements, we ended up with a double stage concept where the first stage controls the needed high number of actuators at moderate speed to get a high Strehl Ratio in NIR, and the second stage, significantly faster is optimized to provide an additional halo suppression in the central region of the PSF where the most challenging targets are searched. This chapter is structured in the following way: After a description of the requirements (section 8.1) and a first order choice for the implementation of the AO system (8.2.1), we make a list of the different error sources that guided us in making the trade-off analysis that led to the concept (8.2.3). The concept is detailed in section 8.3 and the performance is evaluated in section 8.4. The section 8.5 gives a list of the different items to be worked out in the mid-term development plan.

8.1 Requirements

8.1.1 General

The Top-Level Requirements that are the most relevant for the design of the EPICS XAO system are recalled here:

TLR_1: The instrument covers the wavelength range 0.6 – 1.75 microns

The instruments are covering a large spectral region from visible to near infrared. **TLR_5** and **TLR_6** call for two classes of instruments for exo-planet detection:

- one in the near IR from 1.1 to 1.7 microns split in two instruments , a Differential Imager and an Integral field Spectrograph.

- one in the visible for the polarimetric mode.

TLR_14: Simultaneous search for terrestrial planets and gas giants will be possible

Even though both classes of instruments in Visible and Near infrared will be sensitive to gas giants and terrestrial planets, the visible range was initially preferred for rocky planet since

those are expected to exhibit a significant polarization. The differential imager having a larger FOV is better suited for gas giants. However it also contains the interesting H_2O feature. Whatever the detection goal, for the highest detection efficiency of the system, the whole three instruments should preferably be able to work simultaneously.

TLR_9 and TLR_10: Contrast requirements.

TLR_9: Earth up to 20 pc at a phase angle of 90° is detected in all main observing modes at SNR > 5 in one night of observation

TLR_10: Jupiter up to 20 pc at a phase angle of 90° is detected in spectroscopic mode at SNR > 50 in one night of observation

TLR_9 which is the most challenging translates, in case of an Earth-like planet in an Earth-like orbit around a G2 star, into a scattered starlight rejection capability at 50 mas permitting to attain a contrast of 2.10^{-10} in one night of observation.

The detection of Earth-like planets by differential methods with wave-length splitting (the two NIR instruments) impose that the system is compliant with low chromatic differential aberrations. Polarimetric methods impose a low telescope and AO system polarization with high stability. These requirements have an important impact on the optical design of the common path AO system (see section 7.2.1).

TLR_11:The AO control radius is larger than 0.4″ at 800 nm (goal 0.8″)

The actuator pitch projected on the pupil needs to be 20 cm (goal 10 cm) to comply with **TLR_11**. This translates into 500x500 (or 1kx1k) sub-aperture across the pupil. The number of degrees of freedom of the corrector assuming a 100-m telescope with 35 % central obstruction is $1.7\ 10^5$ ($6.8\ 10^5$).

TLR_12: AO limiting magnitude for achievement of TLR: Good for > 100 stars of the spectral types G, K and M.

According to section 6.3, G stars are considered up to 25 pc (G2 at 25 pc, M_v=7.0), K stars up to 20 pc (K2 at 20 pc, M_v=8.5) and M stars up to 15 pc (K2 at 15 pc, M_v=9.5).

8.1.2 Performance requirements

The EPICS ultimate contrast requirement is 4-5 orders of magnitude higher than the VLT-Planet Finder science goal of about $10^{-5} – 10^{-6}$ contrast at 0.1 arcsec. When scaling from a 10-m to a 100-m class telescope, the contrast naturally improves by a factor of 100 for a given rms value of the wave-front error. This means that the XAO system for EPICS should provide a 2 or 3 orders of magnitude better starlight halo rejection than a simply scaled version of the VLT Planet Finder system. This matter of fact calls for system specifications that are tremendously more stringent.

- a significantly higher AO system frame rate (up to 3-4 KHz) to reach high rejection in the central part of the field-of-view (for separations less than 0.1 arcsec for the Earth-like planets detection goal).

- the <u>systematic errors must be kept at a very low level</u> on the low and mid spatial frequencies (f < 2.5 cycles/m in the entrance pupil frame). For VLT Planet finder, on these spatial frequency range, the static errors contributes by about 40-50 nm. A gain of more than an order of magnitude is needed.

- the wave-front sensing measurements error propagation on low and mid-spatial frequencies must be very low: the use of <u>phase-type sensor</u> instead of a slope sensor is needed at least for the correction of the halo at separations less than 0.1 arcsec.

The role of an XAO system for a planet finder is two-fold:

- **Condition 1**: to deliver <u>a high Strehl Ratio</u> at the science wave-lengths in order to concentrate most of the candidate exoplanet's light in a diffraction core.

The top-level requirements of the XAO system as given by **AD 2** are summarized in Table 8-1:

Star mag	Seeing (arcsec)	Strehl Ratio			
		V	R	J	H
< 10	< 0.6	0.40	0.64	0.87	0.92

Table 8-1 Top-level requirements for XAO in terms of Strehl Ratio.

- **Condition 2**: to provide, in combination with a coronagraph, <u>the level of rejection of scattered starlight (typically halo intensity lower than 10^{-7} at 0.1 arcsec),</u> that permits planet detection and characterization in a reasonable amount of time.

  - <u>the part of the halo that averages out</u> defines mainly the level of photon noise against which the planet needs to be detected. It directly impacts the total integration time needed.

  - <u>the part of the halo that doesn't average out</u> (mainly quasi-static speckles), is the most critical part and defines the ultimate level of contrast one can reach.

These guidelines are very important in the definition of the XAO post-focal system.

8.1.3 Computing power requirements

Given the huge number of degrees of freedom needed, at least $1.7 \ 10^5$, the computing power requirements are a very important aspect of the system and this issue needs to be considered at an early stage of the project and definitely needs to be part of the trade-off analysis. A straw-man design of a standard platform for OWL first generation AO real time applications is described in AD 3 (OWL/SPARTA). The technological improvements with respect to OWL SPARTA needed for implementing a Shack-Hartmann based XAO system with the baseline (and goal) number of degrees of freedom has been derived and is recalled in Table 8-2. Four areas that need significant improvement in terms of time needed to achieve a specific task have been identified. The first two are hardware related whereas the last two concern the choice of the algorithms to be implemented. The developments needed to achieve these gains are discussed in section 8.5.3 on the mid-term developments plan for the Real-Time systems architecture.

Technology	500x500	1k x 1k
Input/output communications. Today it is based on a 2.5 Gb/s serial communication. The 10Gb/s is becoming available and it has been used as the baseline for the other designs. Here we need 50 Gb/s for the first case, 100Gb/s for the second	5	10
Faster processing elements, faster CPU-to-CPU busses, faster memory.	10	30
Integer arithmetic. FPGAs perform faster if integer arithmetic is used. By observing that input data (pixels) are integers and output data (control voltages) are as well, one could think of arranging the computation in integer arithmetics. Study is required here. However, performance gain is already known.	2	3
The sparseness of the interaction matrix of an XAO system is very high. However a control matrix (the inverse) is not. Smart algorithms will be able to take advantage of the sparseness of the IM and require less processing power.	5	10

Table 8-2. Four technologies to achieve the XAO requirements. The numbers given here a the gain factors needed with respect to the SPARTA architectures for the 1st generation AO systems for OWL,

8.2 Trade-off analysis

8.2.1 First order choice

Optimizing an AO system taking into account the scientific requirements is a very complex task. Especially for EPICS, the wide range of wave-lengths involved complicates a lot the problem. Ideally, an AO system should provide the best possible correction, while using some part of the target flux for wave-front sensing, <u>preferably outside the scientific spectral bandwidth</u> for the obvious reason of maximizing the SNR for the scientific instrument. Up to now, most of the AO systems were aimed at providing a correction in the near IR by using the visible photons. The scientific wave-lengths of EPICS as shown in section 7 cover the visible and Near IR at the exception of the I band. In XAO, one has to deal with at least two additional error sources which are usually overlooked and that are directly linked to the difference between the scientific wave-length and the wave-front sensing wave-length and that come from the dispersion of the air index of refraction. These error sources are known as chromatic seeing (section 8.2.3.7) and anisoplanicity error due do atmospheric differential refraction (section 8.2.3.6) and the only way to completely avoid them would be to make the wave-front sensing at the same wave-length as the scientific one.

Different possible scenarios have been considered at the beginning of the project, and only one has been selected for evaluation. For reasons of lack of time, this choice was based only on qualitative arguments and the analysis in section 8.4 shows that the performance of the selected concept is rather close to the top-level requirements, even if not completely satisfying them, in particular for the faintest star of M spectral type. The different possible scenarios and some of their advantages and dis-advantages are given in Table 8-3.

Scenario	Description	Advantages	Disadvantages
1	2 independent AO systems, and thus two correction paths, one in the visible and one in the NIR, with WFS at scientific wave-length	- no air chromaticity related errors - visible and NIR instruments can work simultaneously	- takes out photons from scientific beams - important complications: two complete AO systems needed, two correction paths, two WFS, one in visible and one in NIR
2	1 AO system only (either for visible or for NIR) and WFS at scientific wavelength	- no air chromaticity related errors - simpler system: only one correction path	- takes out photons from scientific beam - visible and NIR instruments cannot work simultaneously - needs two WFS one in visible and one in NIR
3	1 AO system only (either for visible or for NIR) and WFS at non scientific wave-length	- no photons taken out of scientific beams - large number of photons for wave-front sensing - simpler system: only one correction path	- potentially high air chromaticity related errors - visible and NIR instruments cannot work simultaneously - needs two WFS one in visible and one in NIR
4	**1 common path corrector for both visible and NIR and WFS in I band**	**- simple system: only one AO system and WFS in available spectral band** **- no photons taken out of scientific beams** **- air chromaticity related errors not null but attenuated** **- visible and NIR instruments can work simultaneously**	**- less number of photons for wave-front sensing (but use of phase-type sensor will reduce a lot this disadvantage and the fainter stars are also the reddish one which favors to go to longer wave-length)**

Table 8-3: Different possible scenarios for EPICS XAO implementation and their advantages and disadvantages.

The scenario number 4 has been chosen, at a first place because it is the one that is the simplest on the system point of view and it mitigates the effect of chromaticity. One possible drawback is that less photons are available for wave-front sensing. However, as it is shown in section 8.2.3.3, by using a phase-type sensor one can increase the sensitivity by more than one order of magnitude, which makes the impact of this drawback much less important. The location of the WFS in I band has also important advantages with respect important limitation due to chromaticity of the air index of refraction.

8.2.2 Method: analysis of the phase errors Power Spectral Densities as an approximation of the scattered starlight halo.

This first order evaluation of adaptive optics error sources and trade-off is based on a Power Spectral Density based analytical model [15]. All PSD formula given in this section are derived from reference [15], unless otherwise stated. It has been shown that the phase PSD is a very good estimation of the halo contribution in a diffraction limited PSF [16]. The PSD is thus a good estimate of the final infinitely long exposure image after a perfect coronagraph, showing the effect of the phase defects only.

The reason why the trade-off analysis is done on the residual halo is because it determines directly the SNR and thus the integration time needed for a given observation. In high ground-based contrast imaging, the detection will always be in a background limited regime, where the background is here the starlight scattered halo. In this case, the needed integration time to achieve a given SNR is directly proportional to the halo intensity level, thus every factor gained in reducing this halo translates in the same gain in observing time.

8.2.3 Main Adaptive Optics error sources

This section describes the different error sources affecting the AO performance. The right balance between this error sources and the unavoidable computing power limitations (section 8.2.4) has been used to define the concept presented in section 8.3.

The atmospheric turbulence model adopted in this study is the Van Karman model characterized by the Fried parameter r_0 and the outer scale L_0.

When a vertical turbulence profile was needed (mainly for propagation effects and differential refraction anisoplanetism), a three turbulence layers model with different weights has been adopted (Table 8-4). This profile is based on recent MASS-DIM measurements campaign and represent a pessimistic case.

Atmospheric Layer	C_N^2 (%)	Altitude (km)
#1	20	0
#2	60	1
#3	20	15

Table 8-4: Turbulence profile model

8.2.3.1 Fitting error

The fitting error is simply represented by the high spatial frequency part of the atmospheric phase power spectrum that cannot be corrected by an AO system. If d is the sub-aperture size of the AO system, then the highest spatial frequency controllable by the system is $f_c=1/(2d)$, also called the AO spatial cut-off frequency. In the Fourier domain, the fitting

error is simply represented by the Kolmogorov PSD, $0.0229\, r_0^{-5/3}\, f^{-11/3}$ where r_0 is the Fried parameter, for a spatial frequency f larger than f_c and is null everywhere else.
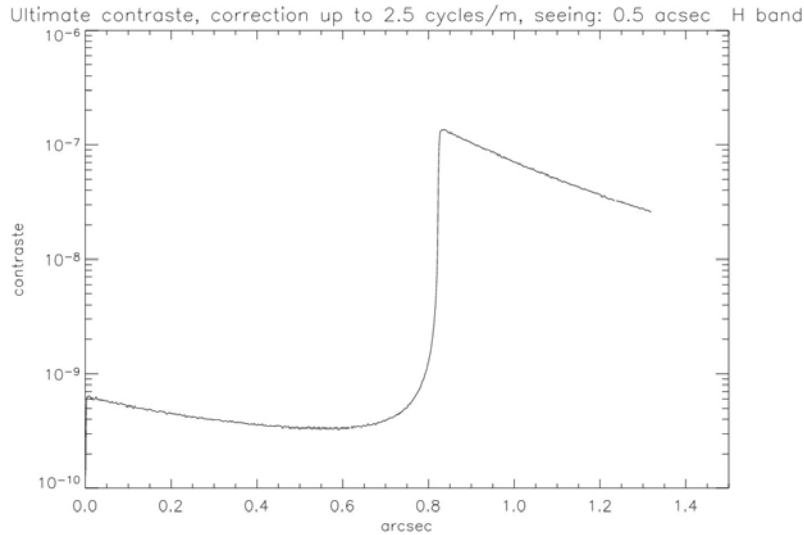
Figure 8-1: Halo for a perfect AO correction (Fourier semi-analytical model) up to 2.5 cycles/m (inter-actuator distance: 0.2 m), seeing:0.5 arcsec, wavelength: 1.6 microns.

Figure 8-1 shows the halo contribution produced by the fitting error using a semi-analytical model where phase screens following the fitting error statistics have been produced and a coronagraphic image produced. The cut-off frequency is $f_c=2.5$ m^{-1}, which corresponds to a AO control radius in H band (λ=1600 nm) of r=λf_c = 0.8 arcsec. One can see that the main contribution is for separations larger than this control radius, and the analytical model only describes this contribution. Inside the control radius, the contribution is generally negligible with respect to others and is not included in model. However it represents the ultimate contrast one can achieve for a perfect correction of all spatial frequencies up to the cut-off frequency f_c.

8.2.3.2 Temporal error

The temporal error depends mainly on the wind speed, on the pure delay t_d in the AO loop and on the integration time t_i during which the measured aberration is averaged. We use the formula in reference [15] and a correction factor for closed loop effect (gain = 0.5) [17].

The halo contribution of the temporal error is plotted in Figure 8-2 for two values of the AO system frame rate. One can notice that increasing the frame rate by about a factor of 3 permits a halo rejection better by an order of magnitude, hence the importance of a fast AO system.

Figure 8-2: Temporal error or servo-lag contribution to the halo (λ=1600 nm). Dashed line: frame rate 1 KHz, t_i=1 ms, t_d=1 ms. Dotted-dashed line: 3.3 KHz t_i=0.3 ms, t_d=0.3 ms. R_0=12.1 cm, wind speed: 12.5 m/s.

An important characteristic of the residual temporal error is that it increases significantly towards small separations. The reason for that is simply the predominance in the atmospheric turbulence of large amplitude low spatial frequency aberrations. This complicates of course severely the search of exo-planets and explains why a fast AO system is required especially to correct the halo at small separations.

8.2.3.3   Photon noise error propagation

The precision of the wave-front sensor measurements are set by the flux of the guide star used to derive the wave-front correction to be applied. The type of wave-front sensor influences dramatically the shape of the contribution of the photon noise error to the scattered light [2]. The comparison between slope sensors and "phase-type" sensors (for example the pyramid sensor), which are the two WFS that were proposed for the VLT-PF, have been extensively studied in reference [1] and [3]. The advantage of a pyramid sensor over a Shack-Hartmann becomes more and more obvious for ELTs when looking at small separations. The noise propagation of a pyramid sensor affects almost equally all the spatial frequencies so that the contribution to the scattered light in the image is fairly uniform. This is not the case for a Shack-Hartmann sensor: the smaller the spatial frequency, the higher the propagated noise, so that the contribution to scattered light dominates at small angular separations. Figure 8-3 compares the halo contributions of a Pyramid sensor with small modulation with the one of an optimized Shack-Hartmann [8]. The difference is quite

impressive. In the example of Figure 8-3, at 100 mas for λ=1400 nm, the gain provided by a Pyramid sensor translates already in 15 times less scattered light. At 30 mas, i.e. at the inner working angle defined by the Top-Level Requirements, the pyramid permits, at this angular separation, 200 times less scattered light due to photon noise propagation. Since there is an inverse proportionality factor between scattered light and number of photons for wave-front sensing, this can be translated in a gain in limiting magnitude for the pyramid varying from 3 to almost 6 on the range 30-100 mas at 1400 nm (note that this wave-length is actually the one used for H_2O detection). This region corresponds well with the area of interest for the search of Rocky Planets in the habitable zone (see section 5.5 and Figure 5-4).
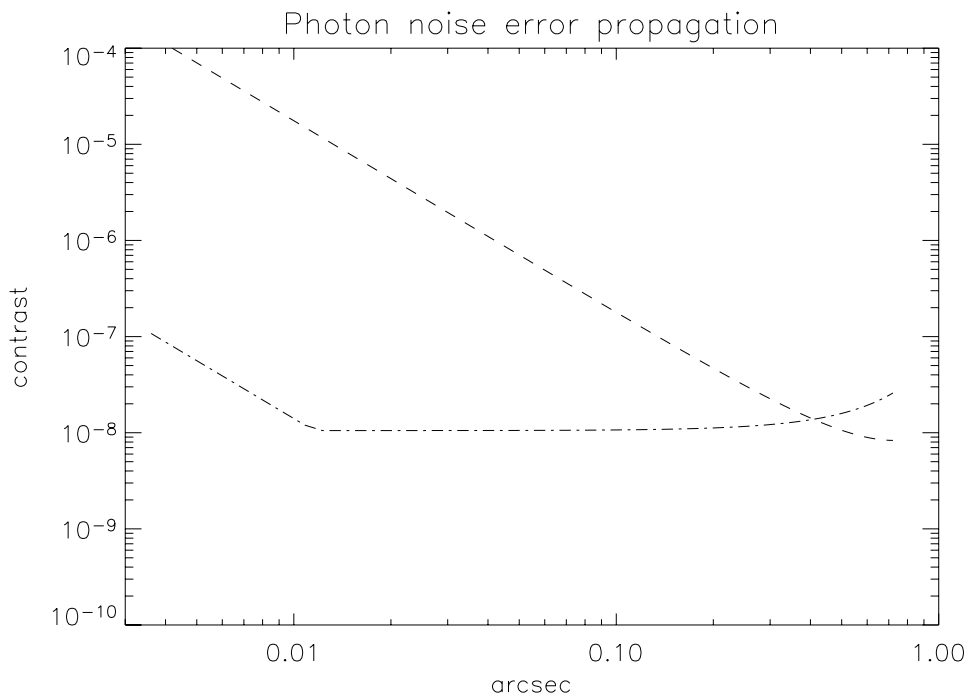
Figure 8-3: Photon noise error contribution to scattered starlight halo in H band (λ=1400 nm). Sub-aperture size: d=20 cm. Number of photons per sub-aperture: 100. Dashed line: Optimized Shack-Hartmann wave-front sensor. Dotted-dashed line: Pyramid wave-front sensor, dynamic modulation: ±2λ/D.

One can notice also that the Pyramid gain is retained only when considering scattered light rejection at separations less than 0.4 arcsec at this particular case. Beyond this point an optimized Shack-Hartmann works better than a Pyramid.

As a general statement, we can say that the Pyramid sensor is very well adapted for halo rejection at small separations whereas the Shack-Hartmann is better for correcting the halo at large separations.

These outstanding properties of the phase-type and slope sensors played an important role in the choice of the XAO concept presented in section 8.3.1.1.

8.2.3.4 Aliasing error

The aliasing error is a consequence of the finite sampling of the wave-front sensor. It can be corrected for a Shack-Hartmann sensor by employing a spatial filter [18] (see Figure 8-4). For the pyramid, the aliasing is mainly limited to large spatial frequencies and becomes negligible at low spatial frequencies [1] (Figure 8-5). A spatial filter is not very useful.
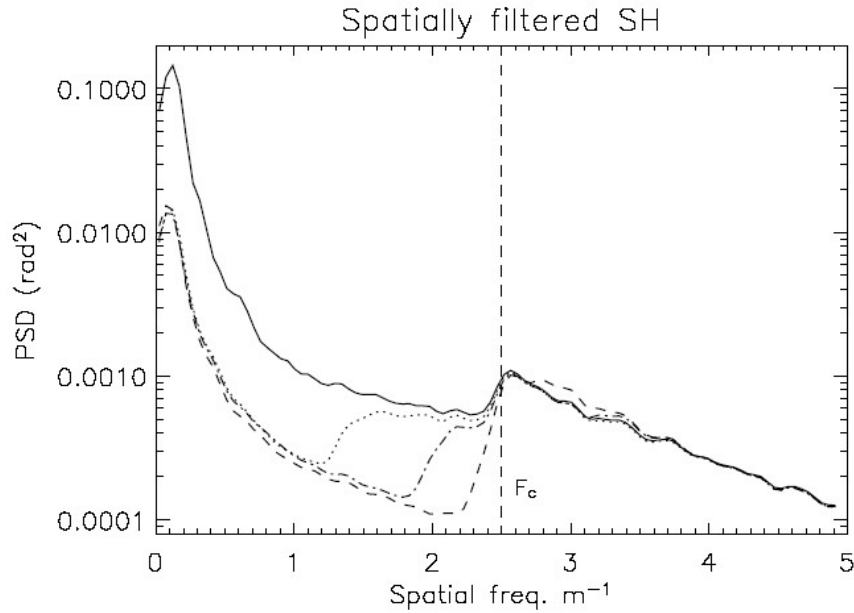
Figure 8-4: Aliasing error (PSD) for a Shack-Hartman wave-front sensor for different size of the spatial filter: dotted line:1.5 λ/D, dotted-dashed line: 1.25 λ/D, dashed line: 1.1 λ/D, where λ is the wave-front sensing wave-length. Solid line: without spatial filter. F_c: AO system spatial cut-off frequency.

Figure 8-5: Aliasing of Pyramid sensor (PSD): SHS (solid line), PWS operated at modulation path of 5% (dashed line), 20% (dotted dashed line), 50%(three dotted-dashed line) and 100% (long dashed line) of an equivalent SHS spot (sub-aperture size d =0.25 m).

In the case of the double stage XAO system presented in section 8.3, a special development has been done to evaluate the aliasing error of the Pyramid sensor after a first stage system based on a Shack-Hartmann (see section 8.4.2).

8.2.3.5 Propagation effects: scintillation

The main physical process addressed in this section is the Fresnel diffraction induced by high altitude turbulence layers. Propagation effects have been studied in the frame of the VLT-PF phase A study (see RD 5). Here only the impact on focal plane images is shown. The results in this sub-section are obtained using a propagation code (PILOT) and a Fourier-based simulation of AO system both developed at ONERA.
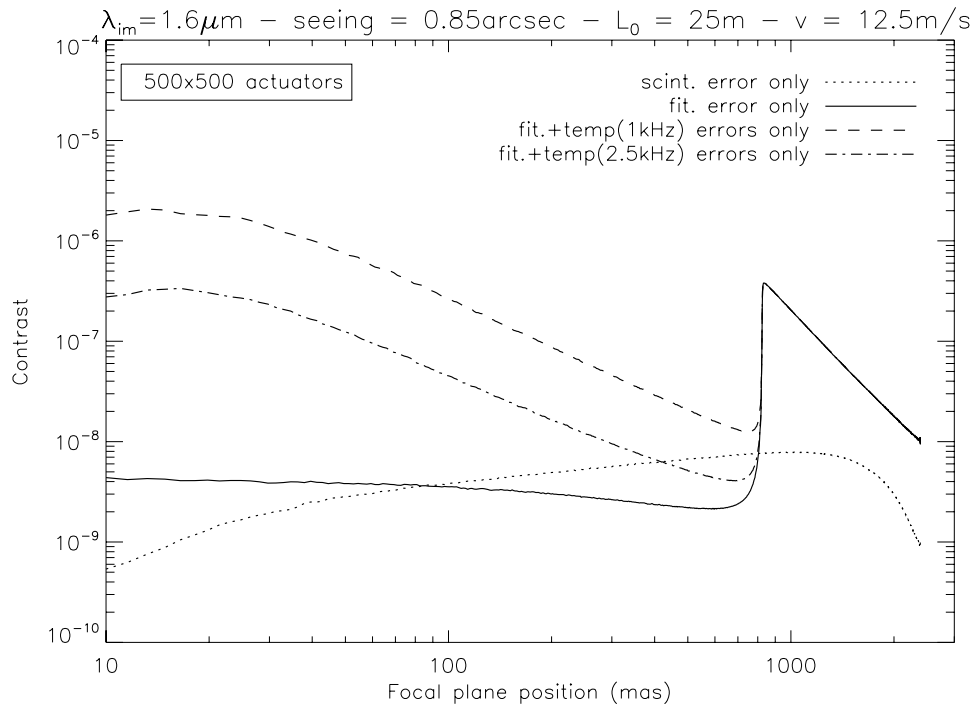
Figure 8-6 Comparison of contrasts with a perfect coronagraph. [dotted line] fitting + ref. diffraction (λ_{wfs}=1.25 µm – λ_{im}=1.65µm), [dashed line] fitting + ref. diffraction (λ_{wfs}=0.65 µm – λ_{im}=1.65µm). [Solid line] fitting (@zenith) + temporal effects (2.5 kHz)..

Figure 8-6 shows the effect of scintillation on the contrast when compared with another component, the temporal error. Scintillation effects (at 1.65 microns) are negligible as soon

as temporal errors are considered. Even in the perfect case of a AO with fitting error only, the scintillation effects only dominate in the 100-800 mas range (and only by a factor 1.5 to 2 in comparison to the plateau produced by high spatial frequencies (uncorrected part of the phase) spread over the whole focal plane by the diffraction pattern and not cancelled out by the coronagraph (even a perfect one).

When comparing the scintillation PSD at various wavelength it is interesting to note that if their integral over all the spatial frequencies (that is the log amplitude variance) evolves in $\lambda^{7/6}$, this dependency is essentially due to the modification to high spatial frequency behavior (due to the modification of the characteristic size of the scintillation pattern -in $\lambda^{1/2}$ -). In the low – medium frequency domain (lower than the AO cut-off frequency), the PSD shape barely depends on the wavelength (see Figure 8-7). It is mainly affected by the L_0 value.
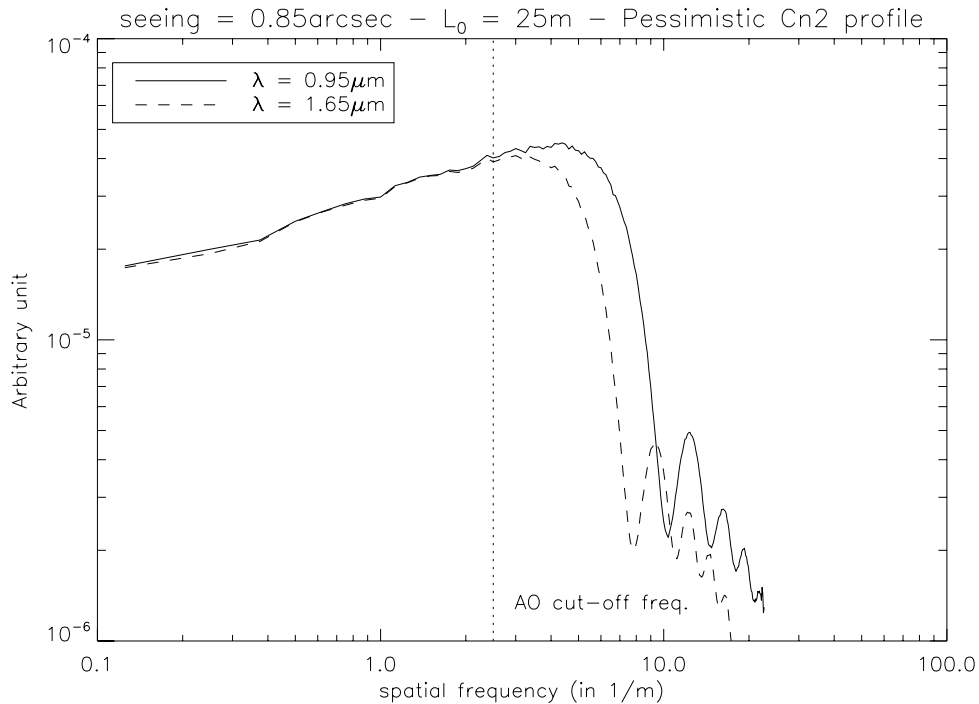
Figure 8-7: Comparison of log-amplitude PSD (circular average) for two wavelengths (0.95 and 1.65 μm). If the log-amplitude variance evolves in $\lambda^{7/6}$ as expected by the Rytov theory, it is interesting to note that the PSD shapes are close to be identical in the low – medium frequency domain (where OA correction is efficient for the phase perturbation)

The effect of scintillation on the wave-front sensing has not been evaluated. It has been studied in RD 5, concluding on a loss of Strehl of about 0.5% in H band. A thorough investigation of this effect will be needed at a later stage of the project and cannot be done in the time of this initial phase since intensive end-to-end simulations are needed.

8.2.3.6 Anisoplanetism due to differential refraction

This effect has been extensively described in RD 5. We just recall briefly the effect: atmospheric differential refraction makes the beams corresponding to different wavelengths appear at shifted locations in the sky. This chromatic differential tilt is compensated by the ADC. However, the different beams have already traveled though different parts of the high altitude layers so that a chromatic error similar to anisoplanetism occurs. This error is important only if the wave-front sensing wave-length is different from the scientific one.

Figure 8-8 shows the difference between sensing in the visible (650 nm) and in I band (900 nm), when observing in J band (1400 nm). The first order choice of the wave-front sensing wave-length in I band (section 8.2.1) permits to gain almost an order of magnitude in scattered light with respect to the usual scheme of sensing in the visible and observing in NIR. The evolution of the error with zenith angle is given in Figure 8-9.
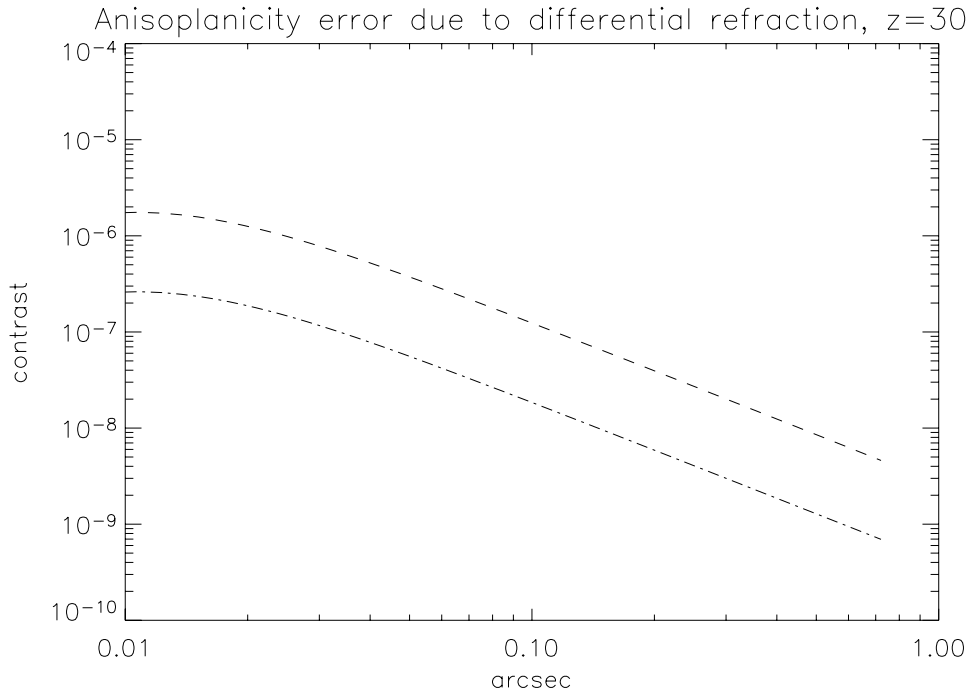
Figure 8-8: Anisoplanatic error (contribution to scattered light) due to atmospheric differential refraction. Observing wavelength: λ_{obs} = 1400 nm. Dashed line: wave-front sensing wave-length: λ_{wfss}= 650 nm. Dotted-dashed line: λ_{wfss}= 900 nm. Seeing: 0.85 arcsec, L_0=25 m.

Figure 8-9: Anisoplanatic error (contribution to scattered light) due to atmospheric differential refraction. Observing wavelength: λ_{obs} = 1400 nm. Wave-front sensing wave-length: λ_{wfss}= 900 nm. Zenith angles: Solid line: z=10 deg, dashed line: z=20 deg, dotted-dashed line: z=30 deg, Three dotted dashed line: z=40 deg. Seeing: 0.85 arcsec, L_0=25 m.

8.2.3.7 Chromatic seeing

Another effect directly linked to the air index of refraction chromaticity, is the chromatic seeing and still affects the performance when sensing and observation wave-lengths differ. This effect occurs whatever the zenith angle and is present and important even at zenith. Since the index of refraction depends on wave-length, the wave-front error when expressed in meters (so not the phase errors), should be the same if the turbulence perturbations were an achromatic process. This is not the case because phase errors due to turbulence are due to spatial variations of the difference of indices of refraction. As a result, the phase errors differ for different wave-lengths by an error which is directly proportional to the turbulence Van Karman PSD and attenuated by a scaling factor depending on the refractivities. The chromatic error is given by [19]:

$$PSD_{obs}(f) = \left(\frac{n_{obs} - n_{wfs}}{n_{wfs} - 1}\right)^2 \left(f^2 + \frac{1}{L_0^2}\right)^{-11/6} \times 0.0229 \left[r_0(\lambda_{obs})\right]^{-5/3}$$

Figure 8-10 shows the chromatic seeing error for the two cases already considered for the wave-front sensing wave-length. A factor of 5 less scattered light is obtained when the sensing is done in I band instead of R band for observation at 1400 nm. It is important to

notice that the PSD of this error is directly proportional to the uncorrected turbulence PSD, so that the very lowest spatial frequencies are overrepresented. For this reason, the chromatic seeing is also strongly dependent on the turbulence outer scale L_0 (see Figure 8-11).
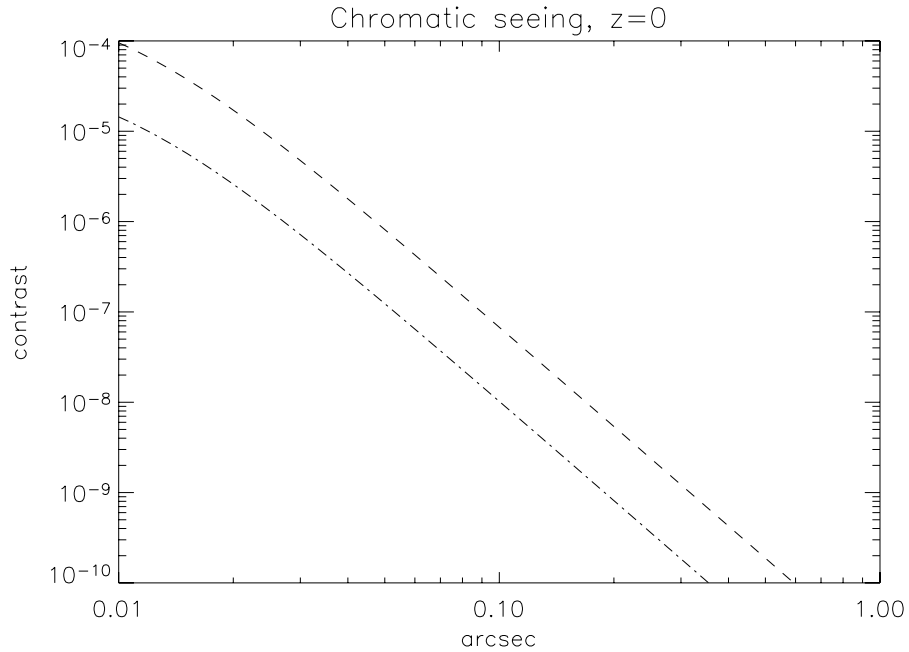
Figure 8-10: Chromatic seeing error (contribution to scattered light) due to chromaticity of air index of refraction. Observing wavelength: λ_{obs} = 1400 nm. Seeing: 0.85 arcsec. L_0=25 m. Dashed line : Wavefront sensing wave-length: λ_{wfss}= 650 nm. Dotted-dashed line: λ_{wfss}= 900 nm.

Figure 8-11: Chromatic seeing error (contribution to scattered light) due to chromaticity of air index of refraction. Observing wavelength: λ_{obs} = 1400 nm. λ_{wfss}= 900 nm. R_0=0.121. Effect of different L0 values: solid line: 10 m, dashed line: 20 m, dotted-dashed line: 50 m, three-dotted dashed line: 200 m.

8.2.4   Shack-Hartmann and Pyramid interaction matrix sparseness

As shown in section 8.2.3.3 a pyramid sensor would be a good choice for an XAO system with a high number of degrees of freedom on an ELT. However, an important aspect needs to be considered: the required computation time. Smart techniques exist for the SH thanks to the possible sparseness of the command matrices. This section shows that it is unfortunately not the case for the pyramid sensor.

The sparseness attained for a given configuration depends on whether zonal or modal interaction matrices are used. Whatever the basis employed is, the inversion and storage of a full-matrix, irrespective of the basis, can be computationally very complex to accomplish. A first-approach technique consists of truncating the control matrix so a considerable factor of sparseness can be introduced directly, allowing considerable reduction of fill-in with limited degradation of performance. Recent developments, however, use zonal reconstructors with penalty matrices: noise and curvature for the case of Least-Squares Reconstructor (LSR) and noise and phase for the case of Minimum-Variance Reconstructor (MVR). The use of such techniques reduces the overall complexity while exploiting sparse and parallel techniques.

The sparseness of SH-based systems can be easily assessed as a function of the radius of influence, which is normally kept small and localized. The zonal interaction matrix of one such system has a fill-in given by:

$$fill \approx \frac{r^2 N^2}{N^4} = \frac{r^2}{N^2}$$

The nnz (number of non-zeros) per line is $\approx r^2$; r is the radius of influence over an actuator, or the radius of the region with significant slope measured around a single actuator when this actuator is activated. N^2 stands for the total number of lines whereas N^4 is the total number of elements in the interaction matrix. Typical fill-in for a SH system is of the order of 1% (see Figure 8-12).

Figure 8-12: Interaction matrix fill-in pattern for a 40x40 SH WFS. The observed fill-in of 1% .

Pyramid-based systems, however, do not have the same sparseness characteristics. Although the diagonally dominance (see Figure 8-13) might suppose a threshold could be applied to filter out some elements and thus increase the sparseness, the distribution of the elements depicted in Figure 8-14 shows that the main concentration is in the interval $[10^{-3}:10^{-4}]$. Even truncating all the elements below 10^{-4} the resultant fill in is 78.3%.

Figure 8-13: Pattern of a pyramid-based system interaction matrix
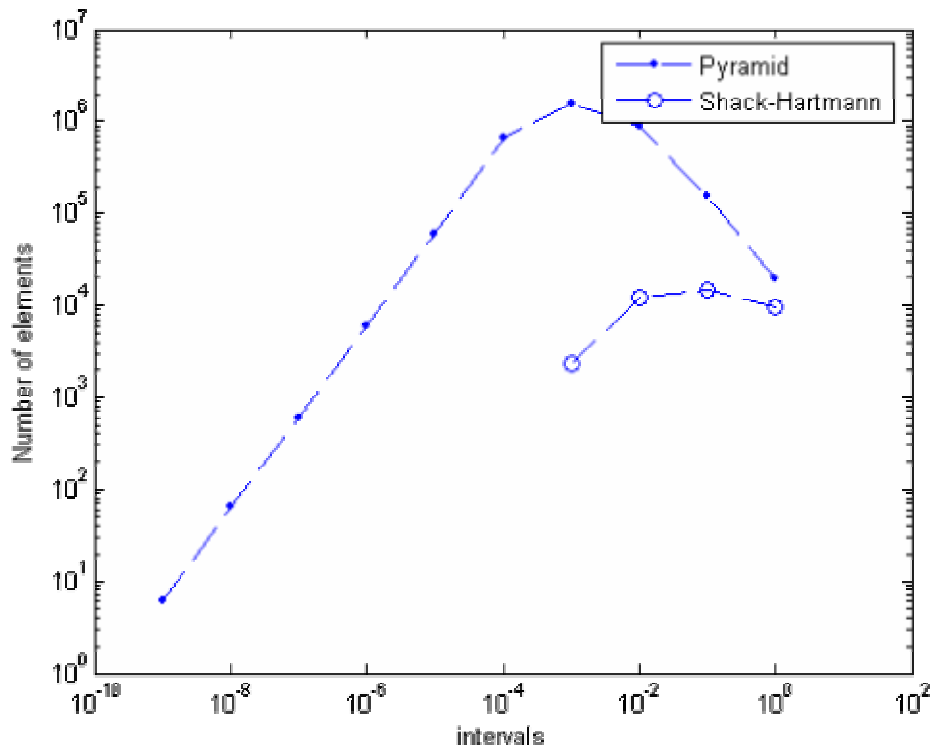
Figure 8-14: Distribution of matrix elements for the SH and Pyramid system. Pyramid has many more and wide-spread distributed elements.

One could think instead, of truncating directly the control matrix (the inverse of the interaction matrix). The results obtained using singular-values decomposition with truncation of 21 values for the Pyramid and 78 in the SH case (correspond to 95% of the Frobenius norm) for the distribution are depicted in Figure 8-15.
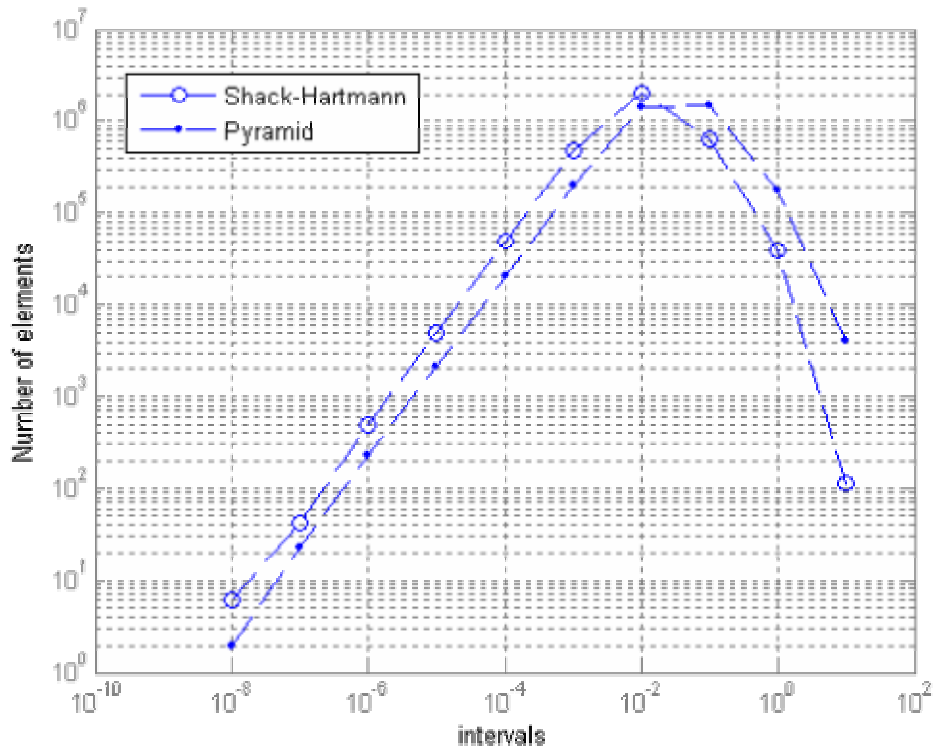
Figure 8-15: Distribution of matrix elements for the control matrices obtained by truncated singular-value decomposition (ensuring 95% of the Froebenius norm). In this case the distributions are similar.

The effect of truncation can be evaluated by measuring the relative error norm after reconstruction - Figure 8-16. For a threshold of 10^{-2}, over which the relative error norm starts to increase, the attained fill-in is 50% for the pyramid and 20% for the SH.

As a conclusion, a pyramid sensor interaction or control matrices cannot be made significantly sparse without impairing the quality of the reconstruction.

Figure 8-16 : Top: variation of fill-in (percentage) with the threshold level, i.e. all matrix entries whose absolute value is below the threshold level are eliminated. Bottom: Relative error norm for the reconstruction.

8.2.5 Wave-front sensing and computing power requirements trade-off

The wave-front sensor in an XAO system is a very important component, and its noise propagation properties must be carefully taken into account. The ultimate science goal of EPICS calls for a very efficient scattered star light rejection very close to the center of the field-of-view. Section 8.2.3.3 has shown the important advantage brought by a Pyramid sensor in terms of sensitivity with respect to photon noise for the reconstruction of low to mid-spatial frequencies.

Our first choice for the EPICS post-focal XAO system, was to couple a 500x500 pyramid sensor with a single high density MEMS mirror with 1.7×10^5 actuators and to control this system at 3 KHz. This is certainly a good choice in terms of noise propagation but revealed to be very risky in terms of required computing power and CCD read-out time requirements. Indeed, the signal provided by a pyramid sensor is a complex function of the entrance phase error, and is characterized by a very non-sparse interaction matrix (see section 8.2.4). As today, the only way of deriving the correction commands from a pyramid sensor is to use a full rank matrix-vector multiplication. Some possibilities combining different approaches will be studied but are still in a very preliminary state.

Even taking into account a significant increase in computing power over the next 10-15 years (see section 8.1.3) , the control of a 1.7 10^5 degrees of freedom or more AO system at 3 KHz using a full rank matrix-vector multiplication seems extremely difficult, if not impossible, to achieve. Indeed when considering Table 8-2, to achieve the computing power of a 500x500 shack-Hartmann based system, an improvement of up to a factor of 5 to 10 (last row of tables) is needed by using smarter algorithms. This cannot apply here, basically because of the non- sparseness of the pyramid sensor.

For Shack-Hartmann-based systems the situation is different; at least two new methods for fast reconstruction exist : one using the sparseness of the zonal interaction matrices [4], and another one using Fourier methods permitting also a modal control [5][6]. These elements led to propose a double stage concept that is detailed in the next section.

8.3 Implementation Concept

8.3.1   Global overview

8.3.1.1   Common path system:

The common path AO system for EPICS is composed of M6 and a post-focal XAO system. The control of these two systems (M6 + XAO post-focal system) will be based on a Woofer - Tweeter scheme, where M6 is dedicated to the correction of the large PTV low spatial frequency aberrations whereas the post-focal system ensures the correction of the fast evolving aberrations and of the high spatial frequencies. To fulfill the requirements on the AO control radius, i.e. a 20 cm inter-actuator separation as projected on the 100-m pupil, the post focal corrector needs to be composed of 1.7x10^5 degrees of freedom**.

Two-stage post-focal XAO system.

The solution we propose is to split the wave-front sensing and the correction in two stages. This permits to alleviate a lot the requirements in terms of computing power as well as of WFSs CCDs read-out. Here is a brief description of this concept (see also Figure 8-18):

* the post focal corrector will be composed by two post-focal DMs:

    * post-focal DM$_1$ (with an equivalent d_1 = 0.2 m actuator pitch): 1.7x10^5 actuators controlled at 1KHz or so, providing the required 0.4 arcsec control radius at 0.8µm. The cut-off spatial frequency of this system is f_{c1} = 1/d_1 = 2.5 cycles/m. This system alone already permits to get a high Strehl greater than 90% in J and H band (it fulfils condition 1), but is unable to provide an acceptable rejection in the central part of the field-of-view.

    * post-focal DM$_2$ (with an equivalent d_2 = 0.67 m actuator pitch): 1.5x10^4 actuators controlled at 3 KHz. The cut-off spatial frequency of this system is f_{c2} = $(2d_2)^{-1}$ = 2.5 cycles/m. This system will permit to increase by an order of magnitude the rejection of the scattered light for separations less than 0.1 arcsec (it fulfils condition 2).

- After reflection by the two DMs, the beam is split between two wave-front sensors with different pupil sampling. This solution, known as hierarchical wave-front sensing, has been proposed [7] as a way to increase the sensitivity of the Shack-Hartmann sensor. We use this concept with a different scope, i.e. to optimize the system in terms of correction bandwidths at the expense however of some moderate loss in terms of sensitivity. Moreover we propose to use two different types of wave-front sensors:

  - WFS$_1$ (Shack-Hartmann WFS): very High Order WFS (500x500 sub-apertures) to control the post-focal DM$_1$ at a **1 KHz** frame rate. A possible control algorithm for this stage is a Fourier reconstructor using optimized modal control in the Fourier domain [6] for spatial frequencies f such as $f_{c2}<f< f_{c1}$. The gain for spatial frequencies $f < f_{c2}$ are essentially put to 0 (or to very low values) because the measurements of WFS$_1$ for this range would be much more noisy than the one provided by WFS$_2$. The Shack-Hartmann detector size is **1Kx1K** (quad-cell SH).

  - WFS$_2$ (pyramid WFS): medium order WFS (150x150 sub-apertures) to control the post-focal DM$_2$ at a **3 KHz** frame rate. A full rank matrix-vector multiplication is used. For the pyramid sensor, 4 independent detectors of at least **150x150** can be used for each quadrant. This way the interference effects can also be mitigated.
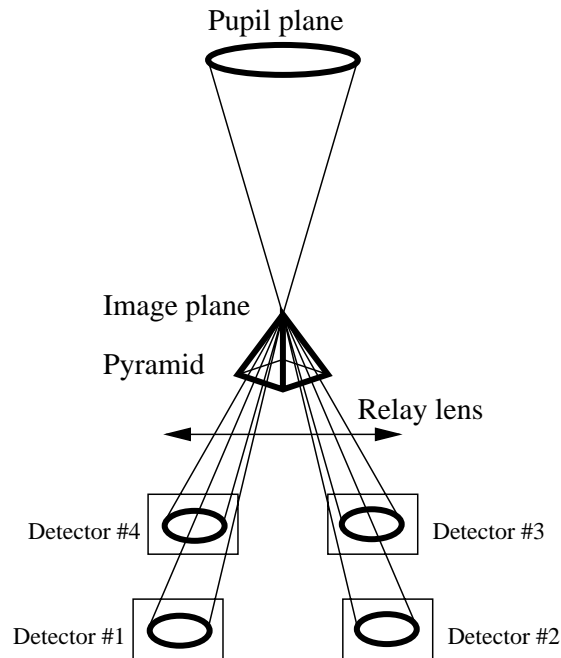
Figure 8-17: Pyramid WFS detector concept

Even though challenging, the detector sizes and frame rates associated in this double stage are not overwhelming. The main reason of this double stage is however the computing power needed to implement a matrix-vector multiplication for a pyramid sensor, whereas the Shack-Hartmann can benefit from much faster algorithms. The development of fast algorithms for a pyramid sensor, if they exist, could permit to consider a single stage with a pyramid sensor only that would provide a better performance and a simpler optical implementation, with however a more demanding detector.

For detection of both gas giants and rocky planets, the light splitting (50/50 for example) is adjusted to provide a level of the halo more or less balanced in the field-of view. For follow-up observations the light splitting and temporal control bandwidths can be adjusted in function of the location of the target. More light can be sent for example to WFS_1 to ensure a higher halo rejection in case of a follow-up observation of rocky planet near the center of the field of view.
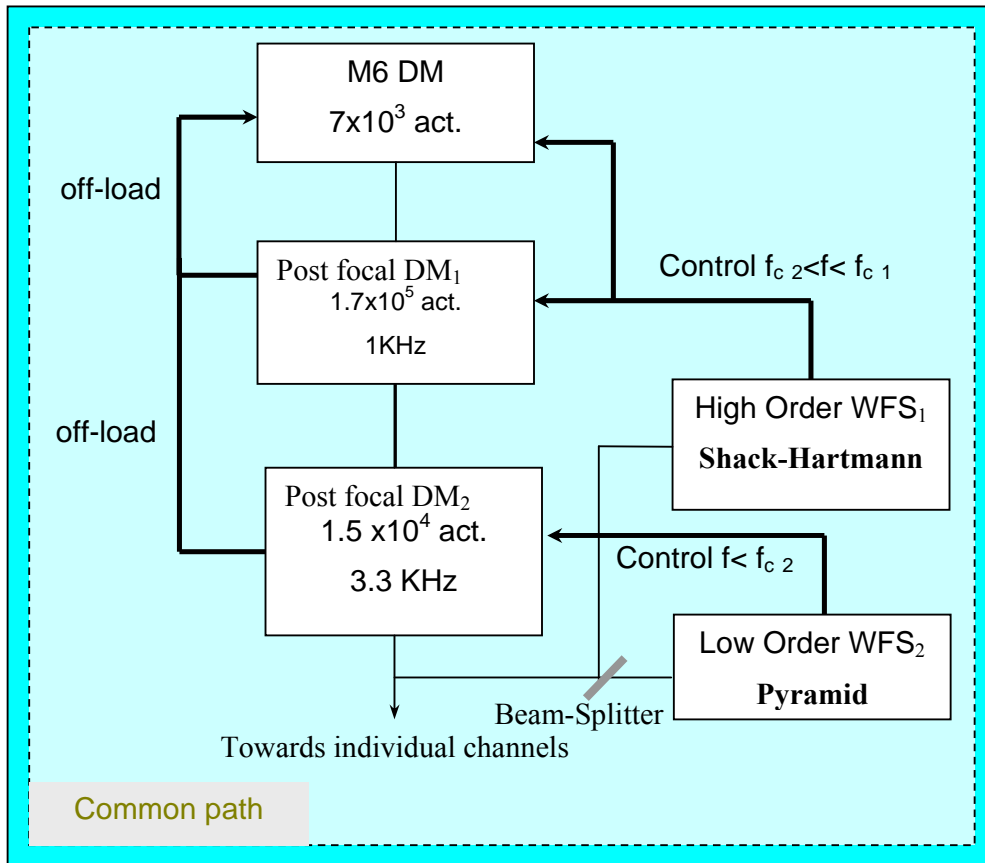
Figure 8-18: Common path XAO concept for EPICS.

8.3.1.2 Individual Channel paths

The very demanding requirements on the static errors before the coronagraph (see section 7.3) imposes to follow a new strategy for correcting them. Static errors come in general from non common path errors: the main contributors are the WFS optics and the optics feeding the coronagraph. The usual way is to use the AO deformable mirror to correct them (using phase

diversity as in the VLT-PF project see **RD 5**) and to send off-sets in the AO control loop. In this case the system drives the correction to a non-null measurements of the wave-front sensor. In the VLT-PF study (**RD 5**), the precision that can be attained has been estimated to several tens of nanometer (40-50 nm) and confirmed by experiments. In EPICS case we consider in the actual baseline three coronagraphs. So using only one deformable mirror we cannot correct for the aberrations before the coronagraphs and moreover this strategy could not permit to attain the sub-nanometric precision needed.

Thus, each scientific channel will be equipped with an active mirror of about 10^4 actuators (it could actually be a copy of DM_2 of the second stage XAO system) for the correction of the residual static error in the common path AO system before the coronagraph. A focal plane sensor (focal plane interferometer [10] or other sensor (TBD)) is used to measure and compensate the static errors using an artificial source for calibration (see Figure 8-19). In section 7.3.3 it has been shown that a control of the amplitude errors is also needed, especially the transmission variations of OWL M1 and M2 segments due to ageing. The individual channel path thus includes also an amplitude corrector. This one could be a spatial light modulator or another DM in a plane conjugated far from the pupil plane. But note in the latter case, the correction would be chromatic.

The calibration strategy will be the following:

- - close the double stage AO loops using the artificial reference source. The WFS will be driven to a null measurement.

- - Close the loop on the active mirror and amplitude corrector and the focal plane WFS to attain the highest possible extinction. At this stage, the residual errors should be such to obtain at least a 10^{-9} extinction in J band at $10\lambda/D$ separation in order to fulfill the requirement for the primary science goal.

Once the system is calibrated, the reference source can be switched off and the observation can start. A full understanding of the stability of this static error correction is needed to decide how often the calibration needs to be done (once a night, every hours or even less…). Moreover, the active mirror must be free of any creep effect, or thermal effect that would induce a modification of its shape during the observation. An additional monitor could be needed to record the active mirror deformation and to maintain it during the observation. The whole AO system, and especially the WFS should be located in a thermally controlled environment.

The calibration strategy described above is based on the assumption that the static error in output of the AO system is essentially due to non common path errors between the WFS path and the coronagraph. In reality there are several sources of static errors in the AO output:

- Non common path errors between the WFS path and the coronagraph: the error the calibration strategy is able to correct

- Chromatic beam shift between the beams at the wave-front sensing wave-length and the beams at the observation wave-length. This error varies with the telescope pointing and is thus difficult to calibrate. It is one of the main errors source in the VLT-PF system proposal (see **RD 7**) where ADCs are located downstream and WFS

wave-length and observation wave-lengths are very far one from each other. About 20 nm rms of static errors are due to chromatic beam shift. In EPICS, since the WFS wave-lengths (I band) is closer to the observation wave-length, and most of the chromatic beam shift is corrected by the ADC up-stream, the error we evaluated is only 1 nm which corresponds to 0.1 nm on the critical spatial frequencies (10 < f < 75 c/p), so much less than the specification of 0.3 nm (see section 7.3).

- Aliasing of static errors (mis-figure of optics in high spatial frequencies). Thanks to the very high order of correction, this errors is small, and the use of a moderate spatial filter in the SH and in the pyramid sensor can filter out this effect.

- Static error arising in the AO system itself: in the pyramid sensor, diffraction effects can create some off-sets in the measurement that need to be calibrated. However this errors are very low spatial frequency thus affecting the halo only very near the star (in any case inside the inner working angle).

Hence, the only significant contributor to the static error will be the non common path errors between the WFS path and the coronagraph .
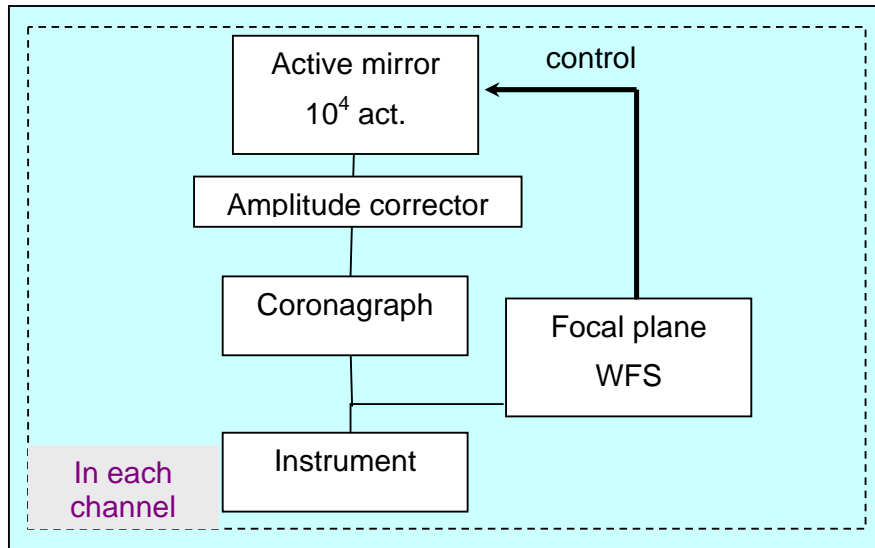
Figure 8-19: Individual scientific channel path.

8.4 Predicted performance

8.4.1 AO Error budget

Using the models described in section 8.2.3, a breakdown of the error budget has been produced and is described in Table 8-5. Two seeing cases are considered, a good one (0.5 arcsec) and a mean one (0.85 arcsec). For the computation of chromatic errors the zenith angle is supposed to be z=30 deg, and the observing wave-length 1600 nm (the WFS wave-

length being 900 nm). This numbers should change slightly when considering other wave-length. The breakdown for other wave-length has not been made but the effect is properly taken into account for the computation of the contrast curves in section 8.4.2.

The error budget has been divided into 3 spatial frequency ranges.

- Low: On the range 1 to 5 c/p (cycles per pupil), the errors affect only the very center of the image, inside the inner working angle.

- Mid: On the range 5 to 75 c/p, the errors affect directly the main area of interest for the search of rocky planets (15 to 200 mas at 1400 nm, 7.5 to 100 mas at 700 nm). This range is controlled by the pyramid sensor (2nd stage).

- High: On the range 75 to 250 c/p, the errors affect the separations larger than 200 mas at 1400 nm, where mainly gas-giant planets are expected. This range is controlled by the Shack-Hartmann sensor (1st stage).

Note that the fitting error, affecting the outermost region that cannot be corrected by the AO system is here included in the high spatial frequency range.

For the computation of the noise on the WFSs, two stellar magnitudes are considered: I=6 for a G2 star at 25 pc, and I=9 for a M2 star at 15 pc, representing the limiting magnitudes to be expected for the two extreme spectral types (for K2 star the performance was not found to be very different from G stars). For the low and mid-spatial frequencies only photon noise propagated in the pyramid sensor is considered. Indeed it is assumed the SH is controlled with a Fourier filter avoiding noise propagation in the low and mid spatial frequencies as defined in section 8.3.1.1. Light is equally split between the two WFSs.

The SH is supposed to be quad-cell based and equipped with a spatial filter [31] (no aliasing). The pyramid aliasing is taken into account since no real gain has been observed by using a spatial filter [32]. Both the high spatial content of the non corrected part and the 1st stage AO residuals (SH) are aliased in the mid-spatial frequency range.

Seeing		0.5 arcsec			0.85 arcsec		
Spatial frequencies		Low 1-5 c/p	Mid 5-75 c/p	High > 75 c/p	Low 1-5 c/p	Mid 5-75 c/p	High > 75 c/p
Error sources							
Fitting (> 250 c/p)		0	0	41	0	0	63
Temporal		<1	3	6	1	5	9
Noise* on WFS	I=6	0	5.5	6		5.5	6
	I=9	1.5	20.	21.5	1.5	20.	21.5

Pyramid* aliasing	I=6	0	7	0	0	9	0
	I=9	0	18	0	0	19.5	0
Diff. Refract. (z=30 deg)		0.5	3	1	1	4	1.5
Chromaticity (z=30 deg)		5	4.5	0	7.5	7.	0
Sub-Total	I=6	5.	10.8	41.9	7.6	14.2	63.9
	I=9	5.2	27.6	46.7	7.8	29.5	67.2
Total	I=6	43.6			65.9		
	I=9	54.5			73.8		

Table 8-5: Adaptive optics error budget. All values are in nanometers. WFS wave-length is 900 nm and observing wave-length is 1600 nm.

The Strehl ratios corresponding to the total errors budgets are displayed in Table 8-6.

STREHL	Band	V	R	I	J	H
Seeing	Mag.					
0.5 arcsec	I=6	0.74	0.86	0.91	0.95	0.97
	I=9	0.625	0.79	0.865	0.93	0.955
0.85 arcsec	I=6	0.505	0.705	0.81	0.895	0.935
	I=9	0.42	0.645	0.765	0.87	0.92

Table 8-6: Strehl ratios.

8.4.2   Residual halo with analytical models

The Strehl ratios given by table Table 8-6 are useful the compute the signal from the test planets for the estimations of the SNRs. This parameter is however not enough to compute the noise coming from the stellar residuals against which the planet has to be detected. In order to compute the stellar noise, PSF shapes of the coronagraphic images needs to be estimated using the errors sources definition of section 8.2.3. In all this section, the

153

coronagraph is supposed to be perfect, which means that the diffraction residuals of the coronagraph are negligible, in terms of intensity with respect to AO phase residuals.

As an example, the performance at all wave-length for G2 stars at 20 pc and M2 stars at 15 pc are shown in Figure 8-20 and Figure 8-21.

For the G2 (Figure 8-20) stars the peak performance in terms of instantaneous contrast is found to be better than 10^{-7} (down to 5×10^{-8}) from 50 mas to 200 mas separation in J and H bands. The performance remains better than 10^{-6} in R band.

For the M2 stars (Figure 8-21), the magnitude is larger and photon noise error propagation dominates much more the halo. Some optimization of the correction bandwidth may improve the performance but has not been investigated here.

In super-position to Figure 8-20 and Figure 8-21, the contrast curve corresponding to the WFS wave-length (900 nm) has also been represented. One can notice that the low frequency part is much lower for this wave-length and demonstrates that the chromatic effects (differential refraction and chromatic seeing) dominate essentially the very center of the image but still affects the halo at the inner working angle.

Figure 8-22, Figure 8-23, Figure 8-24 and Figure 8-25 display the performance for G2 and K2 stars in function of the distance of the star for 700 nm and 1400 nm observation wave-lengths and illustrates the effect of photon noise propagation in the WFS.
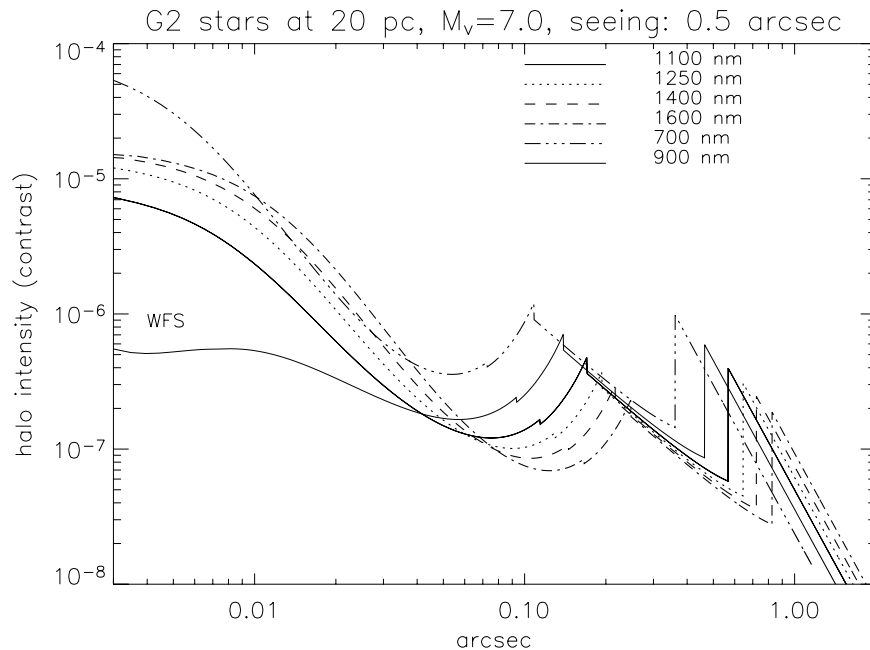
Figure 8-20: Coronagraphic PSFs for a G2 star at 20 pc in function of wave-lengths. Good seeing.
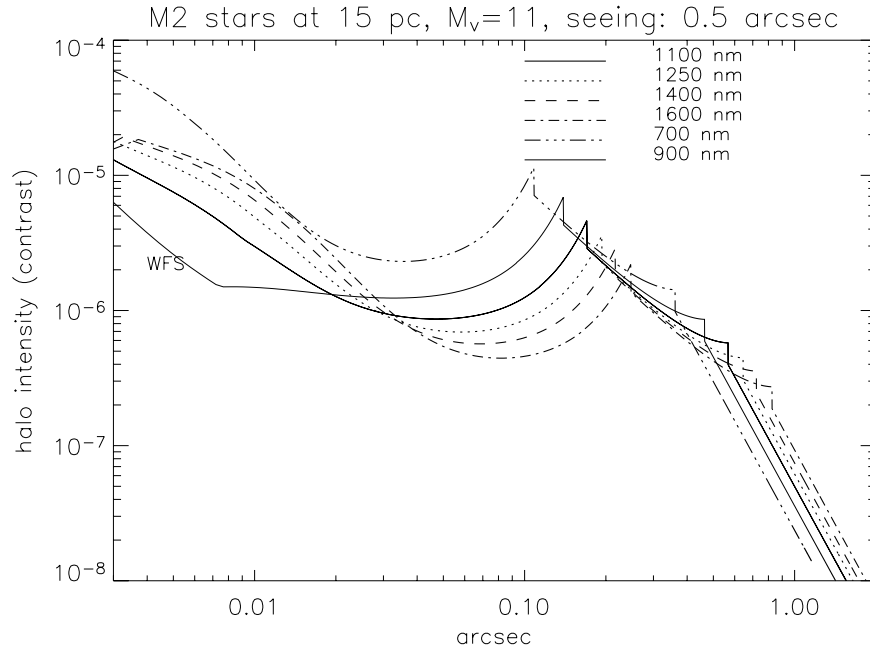
Figure 8-21: Coronagraphic PSFs for a M2 star at 15 pc in function of wave-lengths. Good seeing.
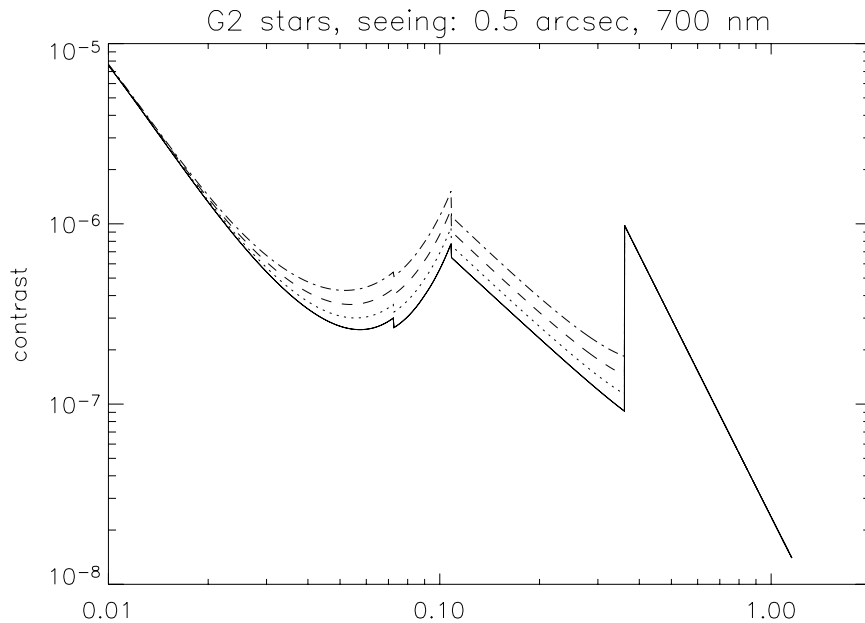
Figure 8-22: Coronagraphic PSFs in R band for G2 stars. Good seeing. Solid line: at 10 pc , dotted line: at 15 pc, dashed line: at 20 pc, dotted-dashed line: at 25 pc.
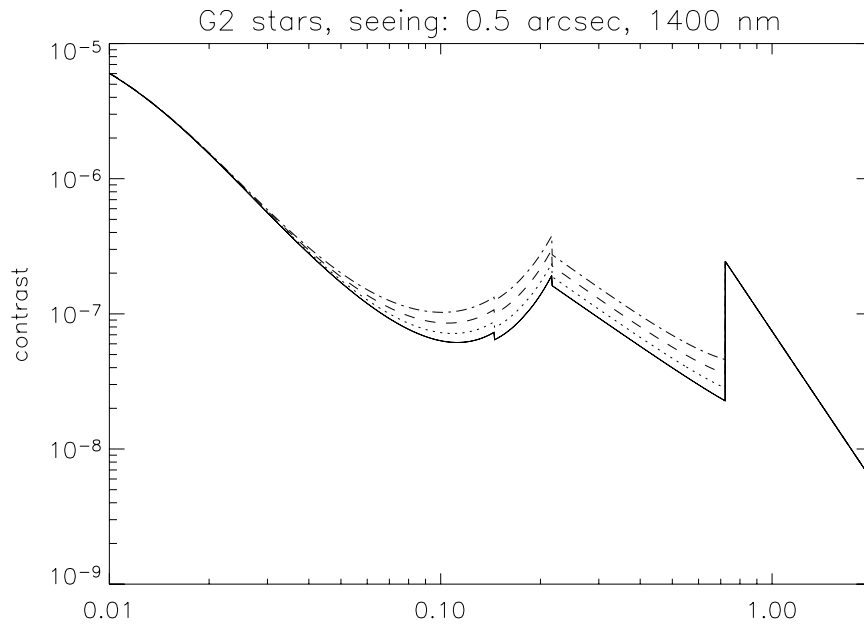
Figure 8-23: Coronagraphic PSFs in J band for G2 stars. Good seeing. Solid line: at 10 pc , dotted line: at 15 pc, dashed line: at 20 pc, dotted-dashed line: at 25 pc.
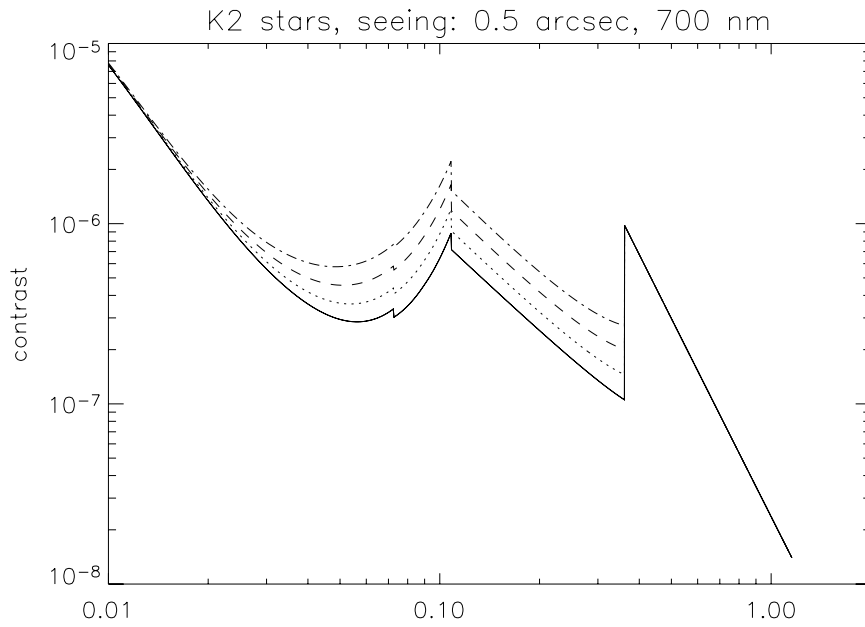
Figure 8-24: Coronagraphic PSFs in R band for K2 stars. Good seeing. Solid line: at 10 pc , dotted line: at 15 pc, dashed line: at 20 pc, dotted-dashed line: at 25 pc.
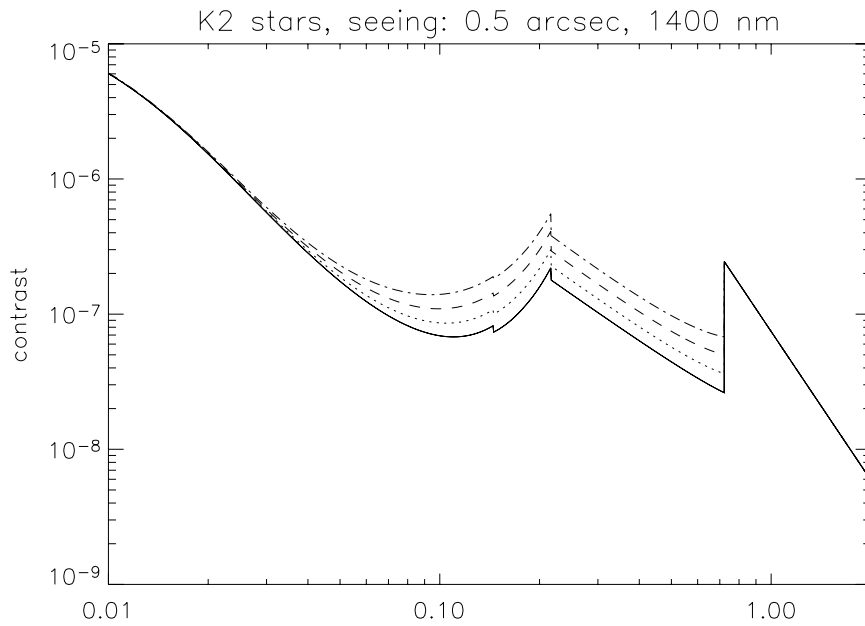
Figure 8-25: Coronagraphic PSFs in J band for K2 stars. Good seeing. Solid line: at 10 pc , dotted line: at 15 pc, dashed line: at 20 pc, dotted-dashed line: at 25 pc.

An example of coronagraphic image is given in Figure 8-26. It clearly shows the two corrected area corresponding to the two stages correction. This image has been obtained by computing with Fourier methods 800 phase screens following the PSD of a typical correction.
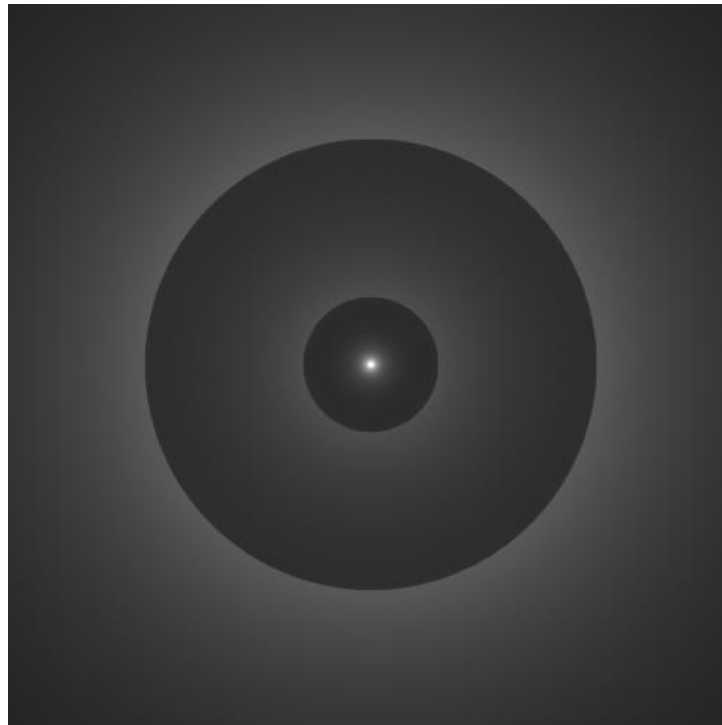
Figure 8-26: Coronagraphic image, 800 iterations. Wave-length: 1220 nm. Seeing: 0.5 arcsec, G2 star at 10 pc (M_v=5.0). The image has been scaled for better rendering. The large outer corrected field (due to 1st stage correction) is 1.25 arcsec in diameter. The inner corrected field (due to 2nd stage correction) is 0.38 arcsec in diameter.

8.4.3   Correction of co-phasing residuals

Co-phasing residuals on an ELT is often presented as an important show-stopper for planet finding. Lardiere et al. [9] estimate that the co-phasing rms residuals should be at the level of 1 nm so that its contribution to the halo remains negligible with respect to the AO residuals. Their main argument is that piston errors cannot be corrected with usual continuous deformable mirrors, so that dedicated fast piston correctors should be implemented. We demonstrate that a continuous mirror can actually correct for piston errors, or more specifically it can correct for the Fourier components of the wave-front that affect the field-of-view of interest. For this we considered a typical co-phasing errors figure from M_1 and M_2 with 20 nm rms total error and fitted, using a Fourier method, a phase function which spectral content is limited to spatial frequencies less than fc$_2$ the cut-off frequency

corresponding to the second stage DM$_2$ inter-actuator separation (67 cm). The initial wave-front map and the DM$_2$ fitted correction wave-front are represented in Figure 8-27 and the residual wave-front (difference of the two) is represented in Figure 8-28.
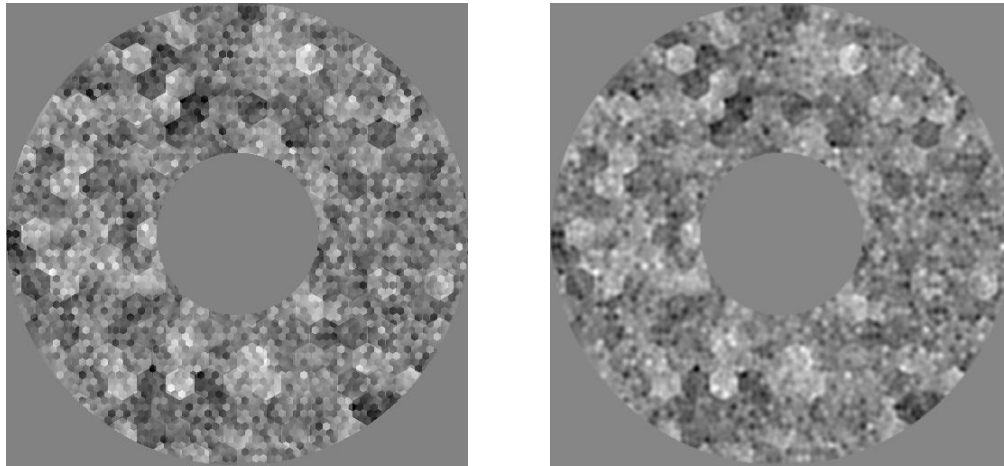
Figure 8-27: AO correction of co-phasing residuals. Left: initial co-phasing errors (20 nm rms of tip-tilt and pisotn). Right: best fit with DM$_2$ (0.67 m actuator separation).
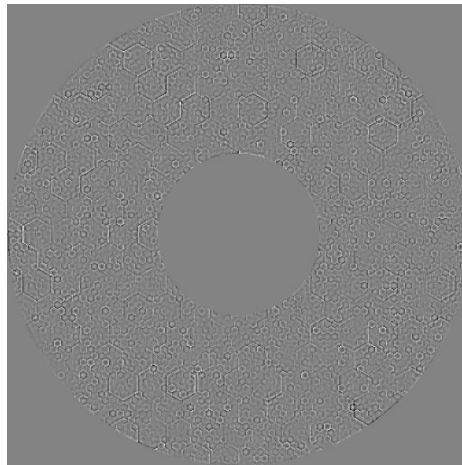
Figure 8-28: Residual left by best fit of DM$_2$ (6 nm rms)

The best fit obtained with DM$_2$ leaves <u>a residual of 6 nm rms</u>. A cut in the three wave-front maps is represented in Figure 8-30. One can verify that indeed, DM$_2$ is unable to correct for the 'jumps' of the wave-front at the segment edges, and those appear still in the residuals. These jumps can even be as high as several tens of nanometers. However, what counts for imaging is the spatial content of the phase. Since we are interested to obtain a very high contrast in the centre of the field of view, the most important is to correct the low spatial Fourier components. It is exactly the case of the residual phase of Figure 8-28.

To convince ourselves that such a residual error figure can be acceptable, we simulated the coronagraphic images corresponding to the wave-fronts with co-phasing residuals and co-

phasing residuals plus correction by best fit of DM_2. The results are displayed in Figure 8-29. One can see that indeed if the co-phasing residuals only are present, the halo is very bright and the raw contrast is larger than 10^{-6} at 50 mas. After correction by DM_2 the residual is only about 10^{-9} in the centre of the field of view, so negligible with respect to AO residuals.

It is important also to notice that the pyramid sensor is sensitive to piston errors (see ref. [1] and Figure 8-31) contrary to the Shack-Hartmann, so that it will be able to measure them. This is another reason of why a phase-type sensor should be used.
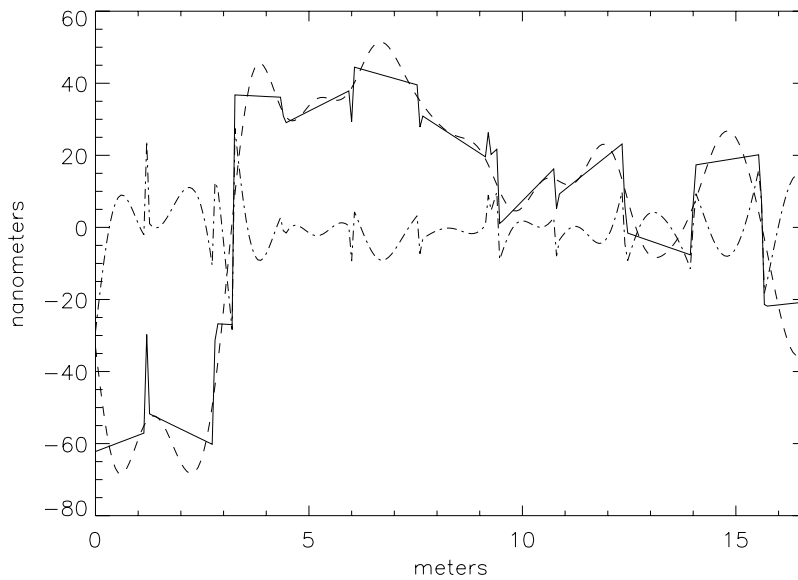
Figure 8-29: Cut in wave-front maps of Figure 8-27 and in Figure 8-28. Solid line: initial co-phasing errors (piston and tip-tilt, 20 nm rms). Dashed line: DM_2 fit. Dotted-dashed-line: residual error figure after correction by DM_2 (6 nm rms error residual of high spatial frequencies).

Figure 8-30: Effect of co-phasing on coronagraphic image at 1600 nm (Circular average of 2D image). Solid line: PSF without coronagraph. Dashed line: Coronagraphic image with initial co-phasing residuals of 20 nm rms. Dotted-dashed line: Coronagraphic image of residuals after AO correction (6 nm rms).
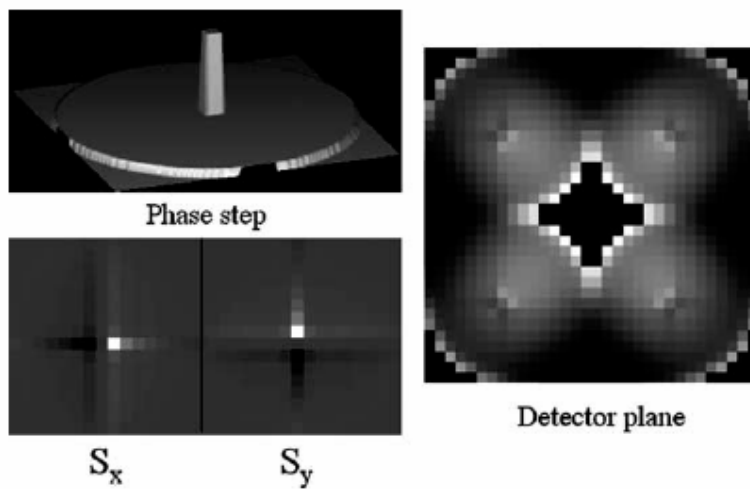
Figure 8-31: Response of the PWS to a pure phase step of 200 nm over one single sub-aperture (sensing wavelength is 700 nm, modulation is 5% of equivalent SHS spot). The signals are represented in two dimensions in grey scale.

8.5 Mid-term development plan

8.5.1 Theory and experimentation

EPICS XAO system involves several new concepts that need to be proven through numerical simulations and experiments. The ESO parallel simulation code is in constant evolution and we have several plans to increase the ESO AO simulation cluster computing power and also to port the code to super-computing facilities (collaboration with Australia though European FP6). In the frame of the Joint Research Activity 1 of OPTICON, we develop at ESO an High Order Test Bench for XAO and coronagraphy experiments.

- Pyramid wave-front sensor: even though very promising this concept is very new and only one system is actually working on the sky. The big advantage in sensitivity is at the expense of non-linearities of the measurements and sensitivity to pupil shape through diffraction effects. We will study optimization of the correction for XAO application and test them with HOT.

- Multi-stage schemes, Woofer-Tweeter schemes. The HOT bench will be composed of two deformable mirrors, a low order curvature mirror (60 actuators) and a high order MEMS mirror (100 actuators). We will develop and test different control algorithms involving multiple stages

- New WFSs, focal plane WFS: the extreme contrast needed for exo-planet search calls for new WFS concepts where the correction is optimized by analyzing directly focal plane images. As an extension to HOT we plan to test these new concepts, like for example the focal plane interferometer, especially in the frame of high precision control of systematic errors.

- Coronagraphy. The performance of EPICS relies entirely on the best conjugation of XAO with coronagraphy so that these two sub-systems should be tested together as much as possible. This is an important goal of HOT. Several coronagraphy concepts will be studied and realized in collaboration with LESIA (Observatoire de Paris Meudon) in the frame of FP6 program and will be tested on HOT.

8.5.2 Key components development

EPICS XAO system requires significant hardware developments.

- Adaptive mirrors: an important risk area is the availability in the next decade of MEMS adaptive mirrors with an extremely high number of actuators ($> 10^5$). In the frame of OPTICON JRA1, we are developing now a 2K actuators adaptive mirror based on MEMS technology. The next step will be the development of a 10K mirror that would make the second stage of the proposed scheme. Still an order of magnitude in number of actuators is then needed for the final goal.

- Detectors. CCDs: 1KxK (goal 3Kx3K) detectors with fast read-out (3 KHz) and low noise (read-out noise less than one electron) are required. Developments of L3CCD are already part of the OPTICON Joint Research Activity 1.

- Super-polishing: Optical polishing and coating quality: a number of optical surfaces in the EPICS design need to be of extremely good quality. The effect of coating on super-polished surface is an important aspect of this topic.

- The experience and results of the Planet finder development phase will be extremely valuable. Important feed-back is expected from extreme Adaptive optics developments.

- Coronagraphy. High precision and complex coronagraphic masks (in phase and in amplitude) are needed. Strong interaction with industry is needed.

- Real-time computers. See next section

8.5.3 Real-Time-Computers developments

Many developments in the field of Real-Time Computers are required to reach the goals of this ambitious system. To make it achievable we need to improve 4 technologies that are the foundations of the main design of all the other AO systems for OWL (see AD 2). Table 8-2 illustrates these four technologies and the improvement that is required in order to implement XAO in two specific cases 500x500 sub-apertures and 1kx1k.

Since this system is not supposed to be built before 2015, we can benefit from the technological advances that will happen in the next 10 year, and the corresponding efforts that can be spent in studying better algorithms. Famous Moore's law predict an increase of computing power in 10 year by a factor 100, so the highest value we used, 30, is not so aggressive.

With the assumptions listed above, the 500x500 system would be made of 3 crates with 20 boards each, each carrying 2 high performance next-generation FPGA chips. Each crate will actually host 3 identical sub-systems, for a total of 9 identical sub-systems each processing one slice of the detector data. Final control values will be merged and exchanged with 6 back-end stages that will control the deformable mirror through multiple parallel lines.

The 1000x1000 system will have a similar complexity, but it requires more aggressive technological improvements and a higher parallelism.

In conclusion, the XAO system at 500x500 (more for the 1000x1000 case) is not achievable with today's available technology. However, we identified 4 critical technologies to improve and the required factor of improvement is within reach for 2015.

Hardware Development

To get to the final performance of the last AO system, XAO, we need to strengthen the relationship with the industrial partner(s) even more than in the previous cases in order to aggregate in a super-performing board the latest technology in FPGAs, CPUs, busses and memory. This, of course, can only be done if the first level of co-operation had been successful. Unfortunately there not many similar applications around and industry could

consider our as a small niche, so convincing them to develop the products we need might be difficult

Algorithms

The XAO system is too big to be implemented using plain matrix-vector multiply thus smart algorithms must be used. Fortunately there are several options.

Generic Algorithm Improvements

- The control matrix can be reduced in precision and compressed. Pixels are normally 16 bit values and mirror controls are 14 or 16-bit values. A loss-less compression, if possible, will allow us to reduce the storage size of the matrix, thus its loading time and the overall performance of the RTC.

- Fixed-point arithmetics. By observing that input data (pixels) are integers and output data (control voltages) are as well, one could think of arranging the computation in integer arithmetics.

- Multi-rate control: the RTC could send more than one command within a single frame time. Multiple commands can be generated by a sequence of approximated commands where the last is the final correct value. This technique is useful if the dynamic of the mirror is not particularly fast. In this way the mirror can be pushed toward the final value, even if approximated, very soon, and while it is reaching the position another refined value will be sent for the final adjustment.

Architecture Specific

- Local reconstruction algorithms can be used to maximize the parallelism of the reconstructor and the controller.

- Fourier domain techniques can be used to reduce the complexity of the reconstructor from n2 to nlog(n). However it has been developed so far only for Shack-Hartmann systems

System Specific

- SCAO and XAO systems are characterized by a large sparsity factor, i.e.: the interaction matrix is mainly made of zeros since the size of the influence functions is small. This characteristic can be exploited to design special algorithms that do not explicitly invert the interaction matrix.

9 Coronagraphy

9.1 Introduction

The term "high contrast imaging" (HCI) is referred to a large number of optical techniques combined by one common aim. This aim is detection and imaging of faint objects (planets, planetary disks, companions etc.) on the intensive background formed by scattered and diffracted light from the parent star. Generally speaking there are two characteristic inputs of this task: first is that the background is much brighter than the object (typically the difference is 4-5 orders of magnitude), and second, that the faint object is located relatively close to the star (as close as 20 mas angular separation). The purpose of HCI is to suppress the star light in the area of the expected planet location with minimal suppression of the planet light.

All HCI techniques are divided into two main groups according to the optical principle of light suppression. In nulling interferometry the light is collected in several pupils, the phase offset between which equals π for the on-axis beam (from the star). The efficiency of nulling is defined by number of pupils and the accuracy of the phase offset. The nulling interferometry is most efficient for the long wavelength (mid infrared, 6-18μm). The complementary technique for the shorter wavelength (visible and near infrared) is coronagraphy. In coronagraphy star light collected in single pupil is suppressed in the image plane. The latter can be achieved by light absorption (Lyot coronagraphy), light destructive interference (phase mask coronagraphy or interferometric coronagraph), or virtually – changing the shape of the star PSF by reshaping the complex field in pupil (pupil apodized coronagraphy). The combination of the techniques as well as optical way to perform the light suppression result into a long list of coronagraphs, which is still growing.

9.2 General principle

The principle scheme (Figure 9-1) includes the apodizer placed in a first pupil plane, the mask placed in the focal plane to block the core image of the star and a diaphragm (Lyot stop) in the next pupil plane to cut the remaining light diffracted outside the pupil.
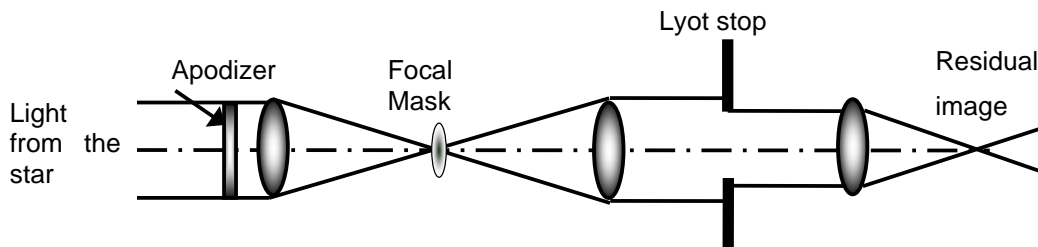
Figure 9-1 Principle scheme of coronagraph.

Depending on the type of the coronagraph and the telescope pupil shape, the apodizer, mask and the stop can be:

Masks:

- Occulting top-hat

- Occulting apodized

- Phase mask with a round shape introducing π-shift

- Phase mask, dividing the field in 4 quadrants with π –phase shift in 2 of them

Entrance pupil apodizer:

- Gaussian

- Sonine

- Prolate

Stop

- Undersized pupil

- Pupil replication

- Reticulated

- Apodized

For some type of coronagraphs the scheme of Figure 9-1 can be used in cascade.

9.3 World of coronagraphic systems

In the reference [20] there is a quite representative list of the coronagraphic concepts organized in "The Coronagraphic Tree of Life" (shown in Figure 9-2), together with the list of the most important characteristics of the coronagraphic designs. The latter include, among others, the compatibility with segmented pupil and compatibility with the on-axis telescope design. This Table may serve as a starting point for the study of the coronagraphs for OWL. Nevertheless, not only the disadvantage (segmentation, huge central obscuration, secondary support obscuration) we have to keep in mind doing a study, but also the obvious advantages of OWL size: narrow size of Airy disk (2.6mas for 0.5micron wavelength) and large collecting area ($\sim7*10^3$ m^2). Not making the full use of the last factors would mean not use of a full system capability, which for the HCI is unaffordable. (The high telescope resolution is also disadvantage for some coronagraphic systems: many stars are not point source anymore, but the disk of a limited angular size.)

The Coronagraphic Tree of Life
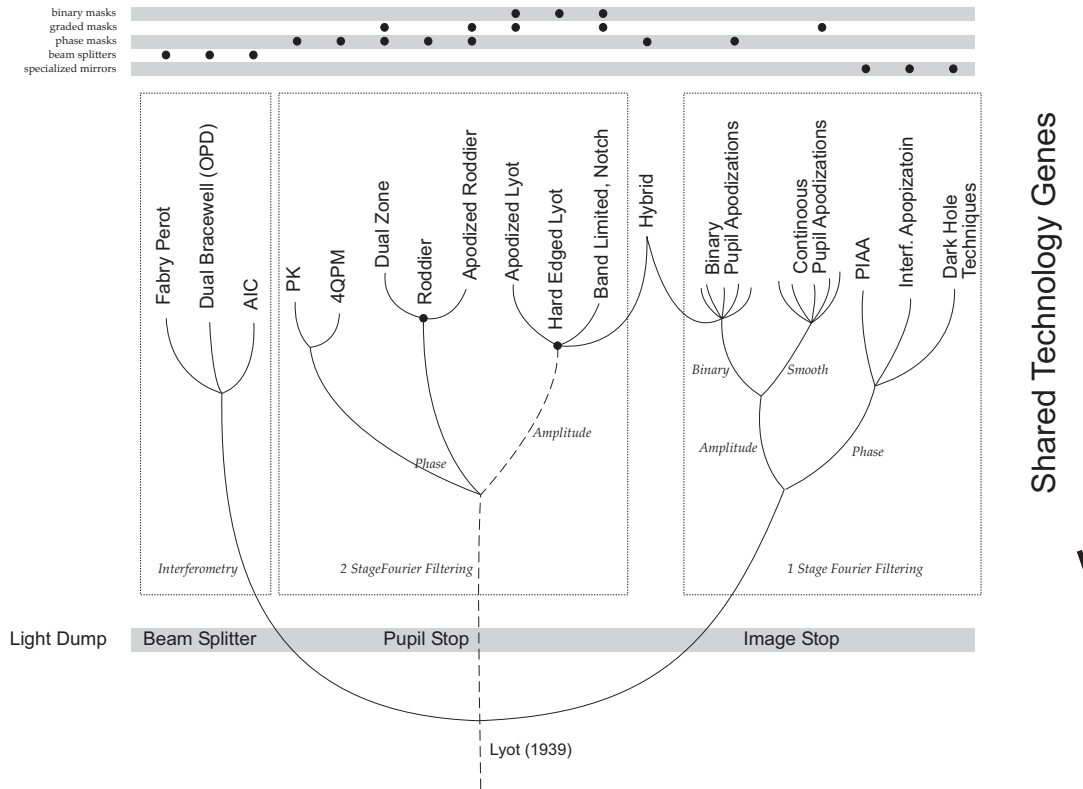
Figure 9-2: Classification of coronagraphic concepts.

The choice and design of the coronagraph strictly depends on the telescope design. In case of OWL we deal with a single pupil, the choice of coronagraphy rather than nulling interferometry is automatic. (Although we can imagine doing nulling with the pupil segments, that idea requires, first, very precise segments control and, second, the temporal shut down of other astronomical programs running at that time on the telescope) The prime input for coronagraph choice: giant pupil of the diameter 100m (truncated by circular mask it is 97m) with the huge central obscuration (the ratio 0.36) and the variable gaps between segments. The segmentation on the secondary mirror is also needed to be taken into account.

In the following we'll discuss several systems which are now being studied and tested in the world, applying them virtually on OWL pupil.

9.4 Classical Lyot coronagraph

Classical Lyot coronagraph, hardly applicable for the telescope of diameter from 2 to 8m might get a new life for the 100m telescope. This scheme includes the top-hat mask placed in

the image plane to block the core image of the star and the diaphragm in the next pupil plane – to cut the remaining light diffracted outside the pupil. The efficiency of the system depends on the size of the image mask measured in the number of Airy disk. On the other hand, the mask cannot be very large in order not to block the planet itself. Clearly, the larger telescope diameter the more number of Airy rings is covered by mask of the given size measures in arc seconds, the more efficient the coronagraph. Figure 9-3 shows the location of the planet with respect of the Airy pattern from the star. The angular separation star – planet is taken 50mas, the wavelength 0.5 microns.
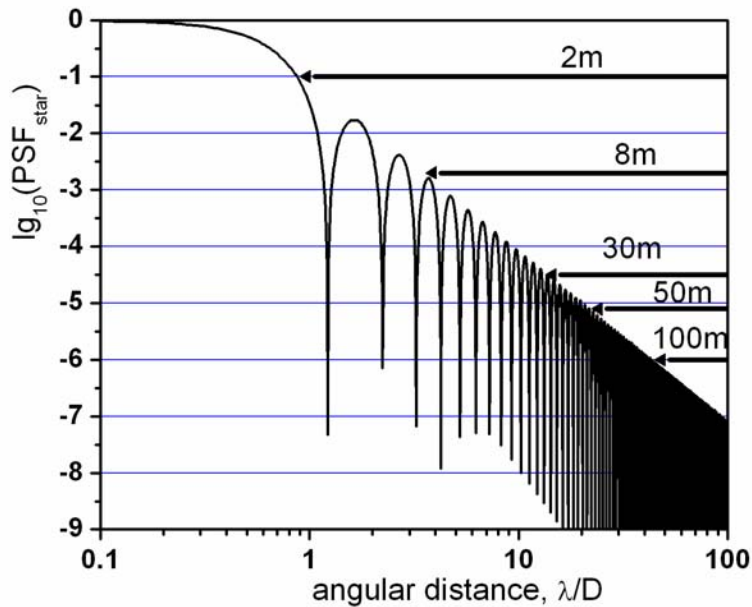
Figure 9-3. Number of Airy rings over a radius of 0.05 arc seconds as a function of telescope diameter. Wavelength is 0.5 μm.

The benefits in increasing the size of the mask are also that the system becomes less sensitive to the tip-tilt jitter and that the limited size of the star is less affecting the performance. The drawback of Lyot coronagraphy is that the theoretical full star extinction is not achievable. Besides, the segmentation gaps play non-negligible role. Nevertheless, both problems can be overcome by a system modification. The theoretical full extinction can be achieved using the band-limited masks; while the influence of the gaps – by implying complex Lyot stop and a second coronagraphic stage. About the second modification we'll speak in more details later.

9.5 Apodized Pupil Lyot Coronagraph (APLC)

The APLC [21][22][23] is an improvement over the classical Lyot coronagraph (LC) [24]. An APLC is basically a classical Lyot coronagraph with a hard-edged Focal Plane Mask (FPM) and a hard-edged Lyot Stop, but using an upstream apodized pupil.

To understand the APLC, and why apodization improves performance, consider what a coronagraph does to the central star's wave front amplitude in the final Lyot stop plane. The complex amplitude in the Lyot plane (before setting the blocking diaphragm) is the difference between the complex amplitude in the input plane (U_1) and its convolution with the Fourier image of the mask's absorption function:

$$U_C = U_1 - U_1 * M .$$ **EQ 1**

To achieve a full extinction of the star light the function Uc must equal zero. For the monolithic telescope in diffraction limit regime it can be achieved by the choice of the mask profile such that $U_1 \approx U_1 * M$ within the pupil. The best coronagraphic efficiency is obtained when these two amplitude distributions are as closely matched as possible. In Figure 9-4 we show these two functions for both unapodized and apodized first pupil plane in the simple example of a monolithic telescope with central obstruction. The amplitude mask with a gradual transmission function is applied in the pupil plan in a way to minimize the residual illumination in the next after coronagraph pupil. Mathematically this condition can be written as the integral equation on the real function $\Psi(\xi)$ representing an apodizer:

$$U_1(\xi)\Psi(\xi) \approx \frac{1}{\lambda^2} \int U_1(\xi')\Psi(\xi')M(\xi - \xi')\, d^2\xi' ,$$ **EQ2**

where $U_1(\xi)$ is the pupil function, $M(\xi)$ is the Fourier image of the mask.
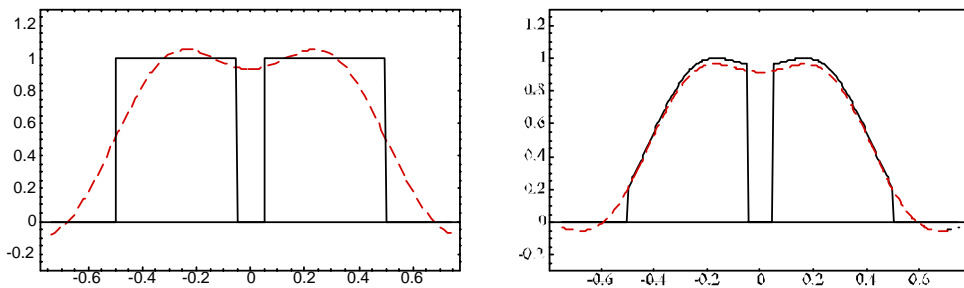
Figure 9-4. One dimensional plots of the normalized amplitude of the wave fronts that are differenced optically by the coronagraph. On the left is a clear aperture with a 10% central obscuration and on the right is an apodized system. The resultant actual amplitude is the solid curve minus the dashed curve. This subtraction is much better in the case of an apodized pupil (right), where the two curves match one another much better. This is the primary motivation for making an APLC.

The formal problem involves finding the combinations of focal mask dimensions and apodizers that maximize total throughput while making these two wave amplitude distribution components as closely matched as possible. A property of APLC is that they do not require a downsized Lyot Stop, improving the throughput of the system. This not only

improves throughput in comparison to LCs, but there is also another great advantage: the angular resolution of the final PSF. With a LC the angular resolution is often limited by the Lyot stop. With a centrally obscured telescope, it is common to loose almost half of the angular resolution due to the Lyot stop dimensions, which downsize the primary and increase the size of the central obscuration [24]. APLCs loose some angular resolution because of the entrance aperture apodization, but the effect is far smaller (about a 10% increase in FWHM of the PSF).

The new apodization functions ([22], Appendix 4.20) are general enough that APLCs can now be optimized for arbitrary apertures, including structure such as central obscurations, support "spider" vanes, or segments.

For a given telescope geometry and a FPM size, a unique apodization function exists. Therefore the throughput and performance of the APLC varies with the chosen mask size. The apodizer throughput and static PSF intensity at a radius of 5 λ/D from the center is given in Figure 9-5 in the simple case of a circular aperture with central obstruction (neglecting segments and spiders). Figure 9-6 shows examples of an apodization function optimized for OWL

Apodizer Troughput as a function of mask size

Figure 9-5: **Illustration of APLC performance for a VLT/Gemini geometry. Top: throughput of the optimal apodizer as a function of the mask size. This number is the overall coronagraph throughput since there is no Lyot stop reduction. Bottom: PSF intensity, normalized to unity. Masks sizes over 4.7 λ/D a high throughput and therefore relatively higher angular resolution as well.**
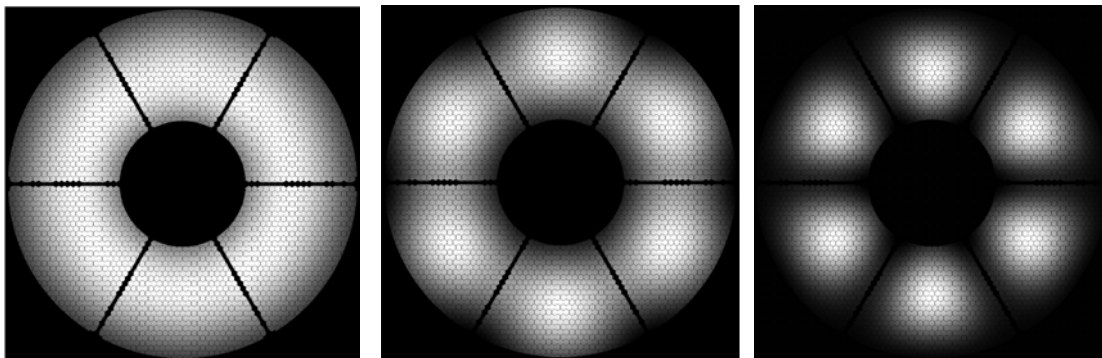
Figure 9-6: Example of apodizer transmission for the OWL telescope geometry, including secondary mirror support structures and segmentation, for 3 FPM sizes of 5 λ/D, 7 λ/D, 10 λ/D. The coronagraph throughput for the 5 λ/D can reach 54%; the larger the FPM, the more aggressive the apodizer, the lower the throughput and the higher the rejection.

An interesting property of APLC is that the Lyot Stop amplitude is simply proportional to the entrance pupil amplitude (Figure 9-7). Thus, it is possible to add a second identical coronagraphic stage in a cascaded mode, without the need of another apodizer. An apodizer is to be used in an upstream pupil plane, followed by two stages of classical Lyot coronagraphy. In this case, the coronagraphic rejection is squared [26] (Figure 9-8).

Figure 9-7: With an APLC, the OWL entrance aperture apodization (left) is simply proportional to the Lyot stop amplitude inside the Lyot Stop (right). The center image shows the amplitude in the whole Lyot stop plane, before the application of the Lyot Stop. Note that the central obstruction and spider structures appear bright is this plane.
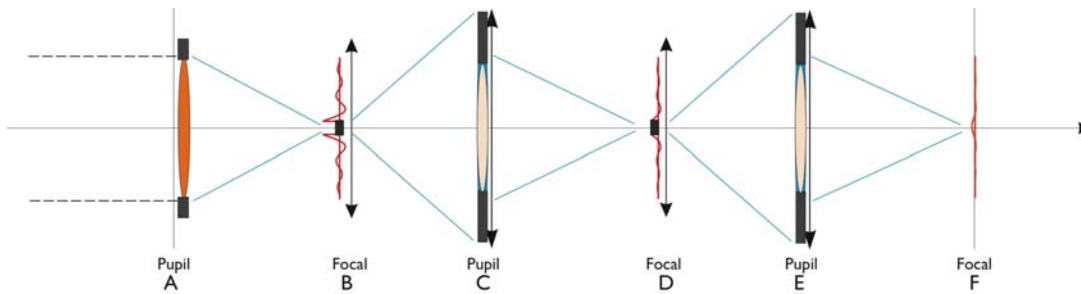
Figure 9-8 Scheme of a multiple stage APLC. A unique apodizer is required in the upstream pupil plane (A); The Lyot stop amplitude (C) is proportional to the original apodized pupil (Figure 9-7). A second stage is possible without a need of a second apodizer. The rejection is multiplied at each stage.

The PSF suppression for large bandpass observations is worse than for the monochromatic case since the apodizer and FPM must be chosen for a single wavelength and therefore, the FPM is too small on the red edge of the band pass and too large on the blue end. If the apodizer could be manufactured to affect the beam with a slightly different apodization as a function of wavelength, we could greatly improve the efficiency of the APLC in practice and it would approach the monochromatic performance, as shown by [23]; this effect can be produced by interferometry. Another approach is to use an absorbing material for which the transmission chromaticity is such as to produce this effect. Absorbing material such as HEBS glass have the property that the absorption in the infrared decreases with wavelength. This naturally well-behaved chromaticity of the apodizer already improves the polychromatic coronagraphic PSF by a factor of 5, based on the simulation of the performance using the infrared spectrum of the material. The material can be further optimized to approach the monochromatic performance [27].

Depending on the mask size, the apodizer can present more or less structures (Figure 9-5). This increases the complexity with azimuthal features in the apodizer. Regardless of

whether they are apodized, the spiders will appear bright in the Lyot Stop plane, and will have to be blocked out (Figure 9-7). The tradeoff between increased complexity (rotation needed) vs. performance has to be studied.

Radial profiles of the APLC PSF are shown in Figure 9-9.
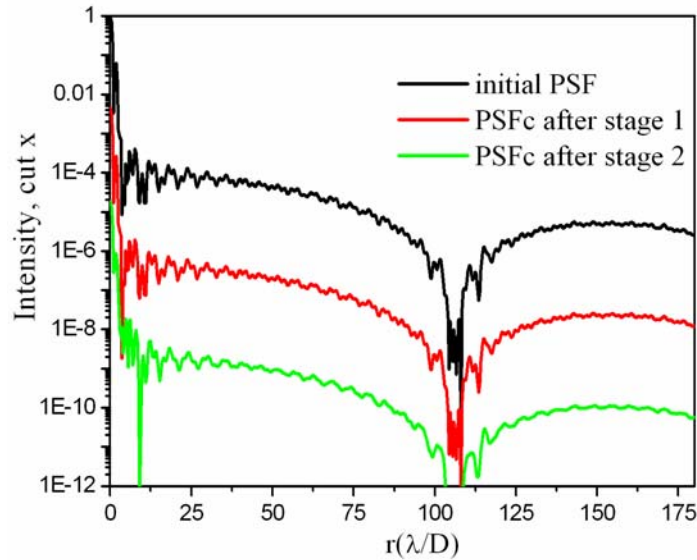
Figure 9-9 Radial profiles of the PSFs corresponding to the left apodizer in Figure 2.5.3. This corresponding to a 5 λ/D mask in diameter (22marcsec for a K-band mask), for a total throughput of coronagraph of 54%. The red curves shows the monochromatic (or achromatized) performance for a single stage, and the green curve for a two-stage APLC. A theoretical level of 10^-9 can be reached around 0.05 arcsec and 10^-15 at 0.15 arcsec. Higher performance can be obtained with more aggressive apodizers, or combining this mild apodization with larger masks, which it totally affordable here given the angular resolution of the instrument.

Figure 9-10 shows the PSF when the same coronagraph being applied on the segmented pupil including the gaps, while the Lyot stop does not masks out the gaps. As is can be seen from the picture no light suppression happens for the diffraction peaks. The effect of the diffraction from the gaps is not particular for APLC only, but to any type of coronagraph. The special care to mask the gaps in a Lyot plane must be done for any system. This is described in more details in a next chapter.

Figure 9-10 Star images before and after the coronagraph with pupil apodization, apodizer applied to a segmented pupil. Averaged size of the gap 14mm. The residual PSF after the second stage is also shown. Mask 5λ/D.

In actual implementation, the APLC is identical to a LC except for the entrance pupil apodizer. The back end LC has been proven both in lab and on the sky and is therefore a low risk part. The construction of the apodizer is critical and poses moderate risk, even considering partial achromatization given the options already available, and the near future results to be expected (implementation of an APLC at the Lyot Project coronagraph). Full achromatization using interferometric apodizers or other techniques is more challenging.

9.6 Non-apodized one stage Lyot coronagraph with a Gaussian focal mask

The mask with the Gaussian profile is the intermediate case between top-hat mask and band-limited masks. The performance of the top-hat mask is not sufficient. The band-limited masks is the subject of a future report. The background for the following calculations can be found in [28].

9.6.1 Mask size

The mask with the Gaussian profile has following absorption function

$$\mathrm{m}(w) = \exp\left[-\left(2\sqrt{\ln 2}\,w/a\right)^2\right] \approx \exp\left[-\left(w/0.6a\right)^2\right], \qquad \textbf{EQ 3}$$

174

with a being a full width half maximum of the mask, w - angular distance from the center. We assume the approximate values of the coefficient $1/2\sqrt{\ln 2} \approx 0.6$. If a is expressed in fraction s of Airy disks then:

$$a = s\lambda/D \qquad \textbf{EQ 4}$$

Parameter s is a critical parameter in Lyot coronagraphy and should be set as large as possible. We define the maximal size of the mask from the condition that the extinction of the planet located at <u>inner working angle</u> specified in section 6 as 0.03″, is less than 1% and 10%. For two cases correspondingly:

$$m(w = 0.03'') = \exp\left[-(w/0.6a)^2\right] \le 0.01 \quad \Rightarrow \quad a_{max} = 0.023''$$

$$m(w = 0.03'') = \exp\left[-(w/0.6a)^2\right] \le 0.1 \quad \Rightarrow \quad a_{max} = 0.033'' \qquad \textbf{EQ 5}$$

The maximal s for D=100m is shown in Figure 9-11 for the range of wavelength defined in [28]: 0.6μm ÷ 1.7μm. The values for 4 filters, R, J and H, belonging to this region are presented in Table 9-1. (*I* from 0.8μm to 1μm band is reserves for the wave-front sensing)

Table 9-1 Maximal size of the focal mask

Band	R	J	H
s_{max} / 1% lost	17	9	7
s_{max} / 10% lost	25	13	10

The small size of the focal mask leads to the large lost of throughput at the pupil stop. Figure 9-12 helps to estimate which combination of mask and stop is required to achieve a given suppression level.
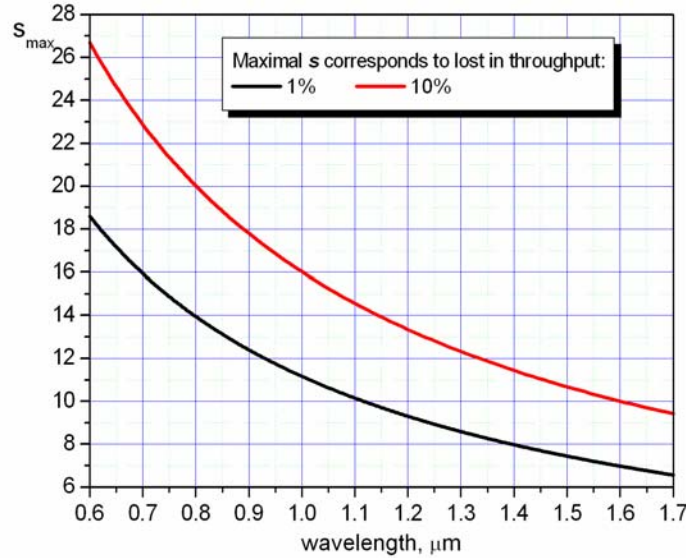
Figure 9-11 Maximal size of the mask measured in numbers of Airy rings for D=100m.
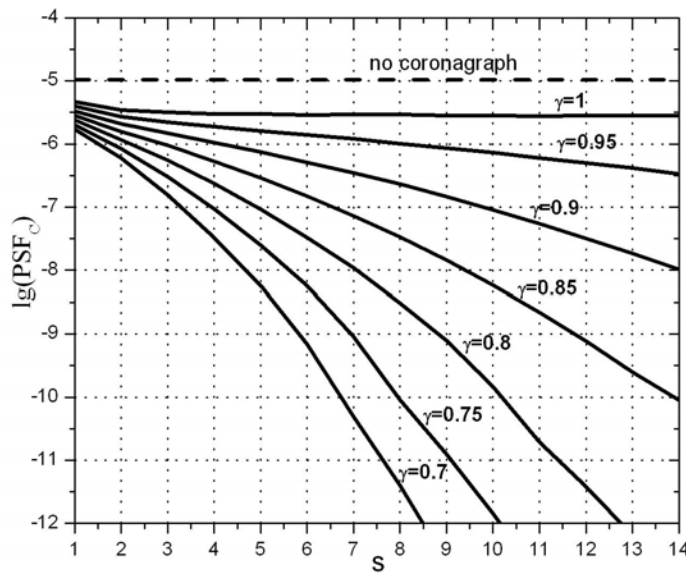
Figure 9-12 Contrast (maximum of residual PSF) at angular distance 20λ/D for circular telescope without central obscuration as a function of mask size for different sizes of Lyot stop. Here *s* is the FWHM of the mask, γ is the linear reduction of pupil diameter by Lyot stop.

The main role of any coronagraph is to suppress the light from the star, diffracted by the telescope pupil. It is not capable to reduce the speckle noise and other speckle cancellation technique serves that purpose. On the other hand, the performance of the differential imaging

176

depends on the level of speckles, and the latter is defined by residual diffraction limited star PSF (pinned speckles).

9.6.2 Segmented telescope: complex amplitude in Lyot plane and the stop

The performance of the coronagraph is easier to define considering the complex amplitude in Lyot plane before applying a diaphragm (Figure 9-1). The simulated U_c is shown in Figure 9-13. The bright areas to be covered by Lyot stop are: gaps on primary and secondary mirrors, mirror borders and spider and the area around the spider. The stop must cover all bright areas. Besides, as during the observation there is a lateral shift of the secondary mirror with respect to the primary (18mm with declination from 0 to 60º), there are must be 2 stops corresponding to 2 segmented patterns. The stop for the secondary mirror must be movable. Two stops are shown in Figure 9-14 and Figure 9-15 .

Figure 9-13 Complex amplitude (modulus) in Lyot plane for OWL pupil, s=17, R-band.
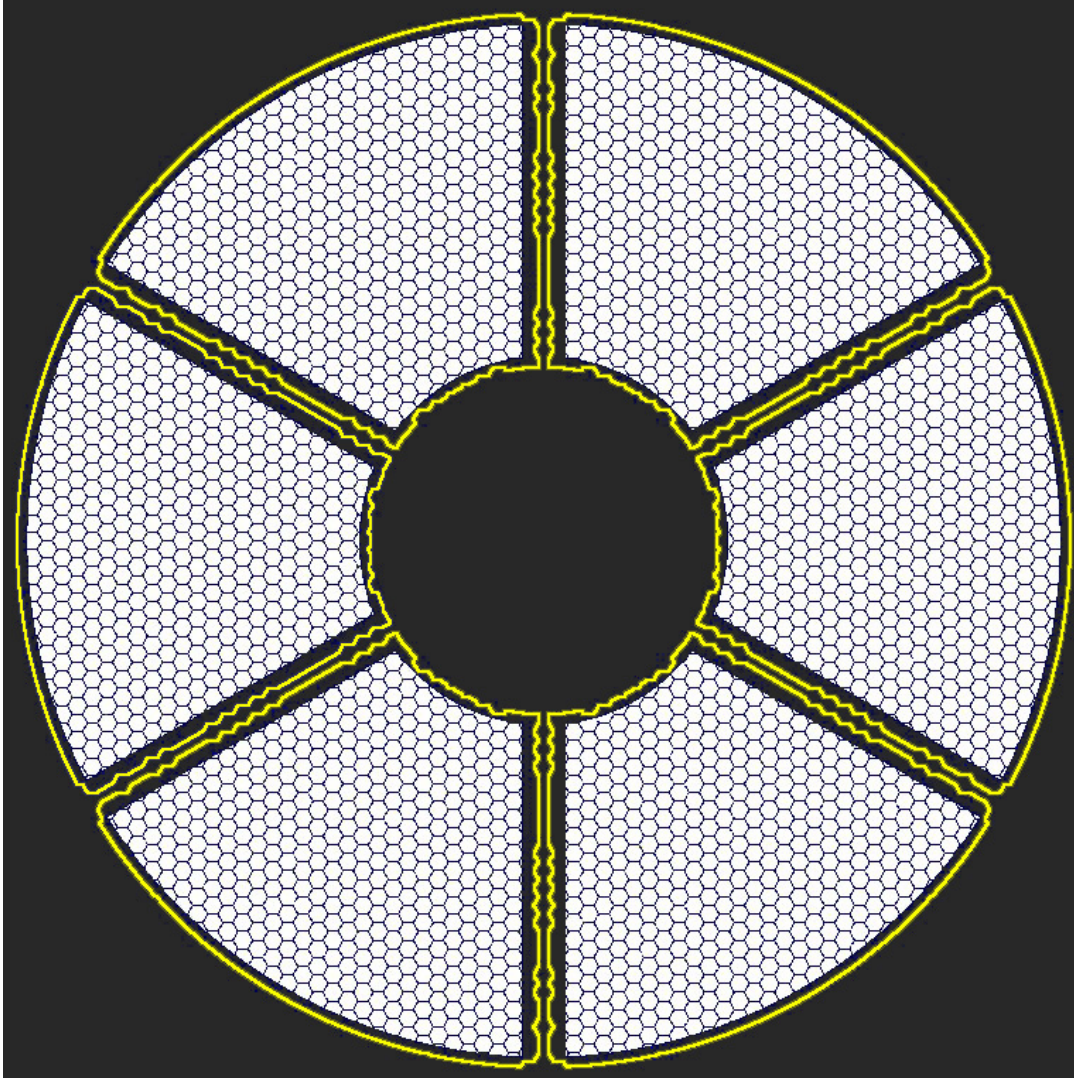
Figure 9-14 Lyot stop for the primary mirror. Yellow line shows the initial border of the mirror

On the stop for the primary mirror the width of the "gaps" corresponds to the maximal size of the gap on M1, i.e. 14mm. For the pupil image reduced from 97m to 1m and 0.5m that corresponds to 140μm and 70μm respectively.

The "gaps" on the secondary mirror stop with the same reduction factor have size 310μm and 156μm. The full shift of the secondary mirror stop is 180μm (1m pupil) and 90μm (0.5m pupil).

The diameter of the mirror is reduced by a factor γ: new outer diameter is γ·97m; new inner diameter is 32m+97m· (1-γ). "Spider" is removed by a new "spider" with the arms each being of δm width. The throughput Σ can be calculated by approximate formula:

$$\Sigma = 3.07 \cdot \gamma - 2.04 - \delta \cdot (0.09 \cdot \gamma - 0.06)$$

EQ 6
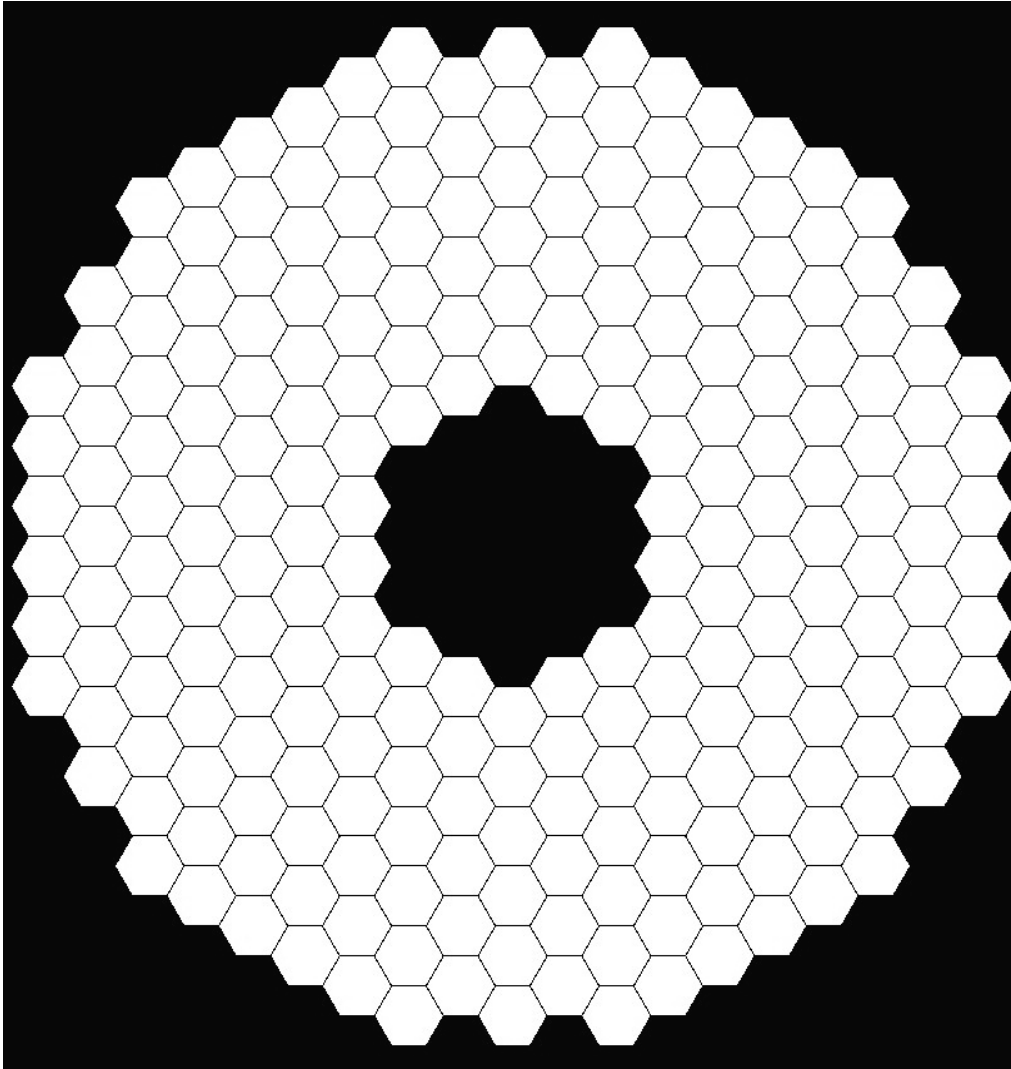
Figure 9-15 Lyot stop for the secondary mirror. An active control of lateral position in a range of 18mm/60º during the observations is required.

9.6.3 Simulation

The results are calculated numerically, assuming a perfect alignment of the stops. The size of the grid used for FFT 4096 x 4096. Mirror sampling grid 0.0651m, which corresponds 1492 pixel across 97m. Wavelength 0.65µm. Mask size 0.023" (s=17). Stop parameters γ=0.92 and δ=5m. The solid strait line define the maximal level of contrast define in [28].

179
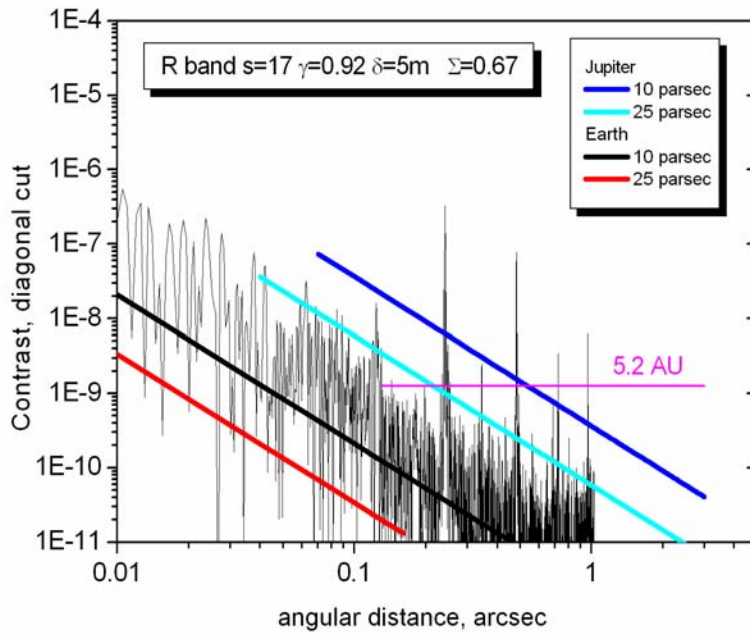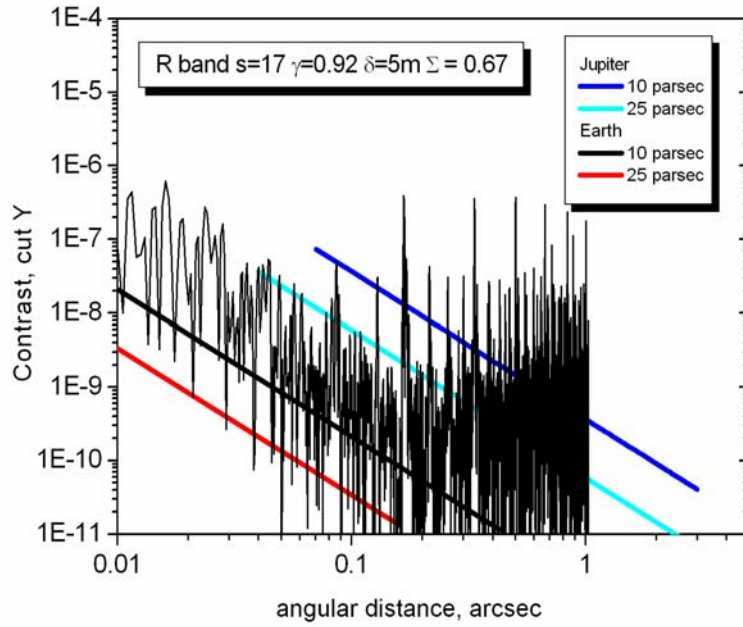
Figure 9-16 Residual PSF for R band in Y-direction (perpendicular to the direction of the diffraction from "spider") and at 45°.

Figure 9-17 Residual PSF for H band at 45º

9.6.4 Conclusions and concerns

- Two segmented patterns require two reticulated stops copying the segmentation geometry for primary and secondary mirrors. Difficulties associated with this approach: fabrication and alignment of the stops and also the lateral shift of two stops with respect to each other to compensate for the shift between the M1 and M2 during the observation.

- The performance of the coronagraph is far from sufficient. The first reason is that the Gaussian mask is not band-limited. The second and the most critical reason is that the diffraction from the gaps forms the high order peaks which are not removed by coronagraph. Also the gaps create a floor of residual intensity in the Lyot plane. (for more details see [28]).

- For R and H bands the location of Jupiter ad 5.2AU coincides with the location of the peaks formed by diffraction on the gaps of M1.

- Huge lost in throughput in Lyot plane, especially for a long wavelength.

- Simulations contain an artifact: because of non sufficient sampling the gaps are simulated with the use of a gray pixel approximation. Therefore the reticulated stops

are also calculated using this approach. That provokes an error in the residual PSF calculation, which is difficult to evaluate.

9.7 Gauss-Lyot coronagraph with reticulated Lyot stop

9.7.1 Coping with segmentation

As we have seen (Figure 9-12) for the Lyot coronagraph there exits the combination of focal mask and Lyot stop which can provide a required extinction. From the study with the segmented pupil we observed that the segmentation changes a situation dramatically. For the segmented telescope with the gaps (assuming diffraction limit regime and a perfect phasing) the complex amplitude U_1 is a comb-like function, which equals unit within the segment and zero in the gap. If the width of the function M is larger than the segment size, the complex amplitude in a next after the Gaussian mask pupil plane becomes:

$$\mathrm{U}_c(r) = \mathrm{U}_1(r) - (1 - g/d)^2 \frac{s}{\sqrt{\pi}} \int_{-1}^{1} dt \exp\left[-s^2\left(t - \frac{2r}{D} \right)^2 \right] \Phi\left(s\sqrt{1-t^2} \right) \qquad \textbf{EQ 7}$$

where s is the mask size measured in a number of Airy disks, g is an averaged size of the gap, d is the size of the segment, Φ is the probability integral. If we assume that the size of the mask is more than six Airy disks we can neglect the pupil boundary (which is truncated by the diaphragm anyway) and study the complex amplitude in the Lyot plane (Uc) within the pupil. The result of the convolution is $(U_1 * M)_{gaps} \approx 1 - 2\omega, \quad \omega = g/d$. Therefore Uc equals $2\omega \ll 1$ within the segment area and $-(1 - 2\omega) \sim -1$ in the gap area (Figure 9-18).

Figure 9-18 Amplitude in the Lyot plane for the telescope with (solid) and without (dash) gaps

The gaps in the Lyot plane are bright and have negative complex amplitude. The reticulated Lyot stop with absorbing grid was studied above. We have seen that even if this type of the stop can be accurately manufactured and precisely aligned, the second type of coronagraph does not give a full extinction due to the remaining "floor" 2ω.

The reduction factor of the initial PSF after such a stop is $(2\omega)^2$. One can think to use the coronagraph in cascade. The PSF reduction factor after the second stage is $(2\omega)^4$. For the segment d=1.6m and gaps g=14mm the first stage coronagraph give attenuation $3\cdot10^{-4}$ and the second 10^{-7}. The intensity of the PSF from OWL pupil at the distance 30mas is about 10^{-5} (for λ=0.5µm). The first stage coronagraph will produce the contrast at this distance 10^{-9} and the second stage coronagraph 10^{-12}. The last case can be considered already as a "perfect coronagraph" totally illuminating the star light.

The last approach requires the alignment of four narrow grids (2 for each segmentation pattern) with the micrometric precision and control of the lateral shift of the two pairs of grids. That increases the risk of having huge error due to the possible misalignment. Systems with only one complex element are more preferable. The alternative is the apodized Lyot stop placed in the first stage. As before this stop has unit transmission within the segments. Within the gaps it introduces the π shift and intensity attenuation equal to $(2\omega)^2/(1-2\omega)^2$. The idea is to produce the uniform illumination in the Lyot plane of $(2\omega)^2$ everywhere and by doing so, obtain a "monolithic" telescope. As the amplitude within the gaps is negative, the π shift is used.

Placing a diaphragm in the first Lyot plane we decrease the telescope diameter and hence the throughput. Following the same principle we may suggest to modify U_c on the pupil border

183

and even outside to have a uniform illumination everywhere within the initial pupil boundary.

The complex amplitude U_{CL} does not carry anymore the segmentation; and the second stage coronagraph with a simple Lyot stop ideally provides a full extinction. The second stage does not require the reticulated stop, only a diaphragm. The principle scheme is shown in Figure 9-19
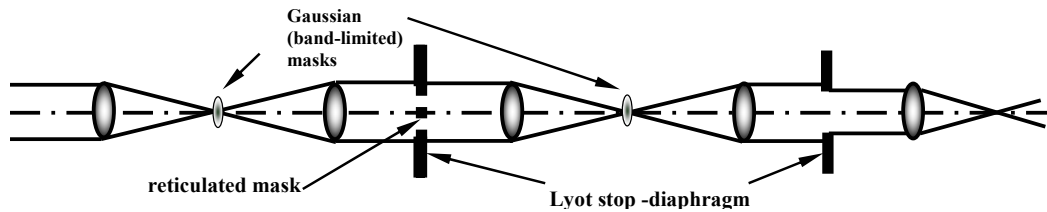
Figure 9-19 Principle of a double stage Lyot coronagraph: the reticulated mask inverses the phase and attenuate the intensity; second stage acts like on a monolithic pupil

The subsequent Lyot coronagraph with the same mask (or slightly bigger) and the diaphragm should act as the coronagraph applied to the monolithic telescope. We study this concept on OWL pupil assuming averaged gaps of 14mm. The residual after two stages star PSF in comparison with initial PSF (Figure 9-20). In the same graph we show residual star PSF after one stage Gaussian Lyot coronagraph with the reticulated stop totally absorbing within the gaps area. To block the diffracted light around the borders (including very strong diffraction around the spider) the stop was placed to cover the central obscuration, outer border of the mirror and the spider in both configurations. This leads to the loss of flux for the planet, which can be calculated as 66% throughput.

Figure 9-20 Initial and residual PSF from OWL pupil (primary mirror only) after applying the one stage coronagraph with stop making the gaps and after the second stage in the system with the reticulated apodized stop.

9.7.2 Chromatic effect

For the one stage scheme considered in Section 9.6, the only sensitive to the chromaticity effect is the size of the mask *s*. Nevertheless as we consider the case of large s, the chromatic effect is negligible. On the contrary, for the double stage coronagraph the π-shift in the first Lyot plane makes a new set-up wavelength dependant. If the mask is designed for wavelength λ_0, than for $\lambda=\lambda_0+\Delta\lambda$ the complex amplitude in a first Lyot plane after the stop becomes

$$U_{CL}(\xi) \approx \begin{cases} 2\omega, & in\ the\ segment\ aperture \\ 2\omega\exp(-i\varphi_\lambda), & in\ the\ gap \\ 0, & \xi > D\gamma_1/2 \ \ or \ \ \ \xi < D(\Delta+1-\gamma_1)/2 \end{cases} \qquad \textbf{EQ 8}$$

where $\varphi_\lambda = \pi\Delta\lambda/\lambda_0$ and we assume that $\Delta\lambda << \lambda_0$. In other words, there is again a segmentation effect. Consequently, the chromatic effect reveals in loss of contrast and appearance of the diffractive peaks determined by a second "segmented" component. The detail analysis for contrast we plan for a future work. Nevertheless now we can make estimation of the peaks intensity. As the first type of the Lyot stop does not alter the peaks, their intensity can be defined directly. For example for A-peaks

$$\begin{aligned} I_{A1,cor} &= 0.68\omega^2(2\omega)^2\gamma_1^4 2[1 - cos(\varphi_\lambda)], \\ I_{A2,cor} &= 0.171\omega^2(2\omega)^2\gamma_1^4 2[1 - cos(\varphi_\lambda)], \\ I_{A3,cor} &= \omega^4(2\omega)^2\gamma_1^4 2[1 - cos(\varphi_\lambda)], \end{aligned} \qquad \textbf{EQ 9}$$

Figure 9-21 shows the contrast in the location of the first diffraction peaks (A$_1$- peaks) from M1. For comparison on the same graph we plotted the intensity of the A$_1$ peaks after the one stage coronagraph with Lyot stop of the second type. For the large bandwidth the performance of the double stage coronagraph is comparable with the performance of its simplified ancestor.

Figure 9-21 Contrast at the location of the first diffractive peaks A_1 as a function of relative bandwidth for different gap size, ω. $\gamma_1=0.9$, $\gamma_2=0.72$. Dash lines – intensity of the peaks after one stage coronagraph with the second type of the stop.

9.8 Interferometric Coronagraph

Another type of coronagraph which is currently being studied for OWL is Hybrid Interfero Coronagraph (HIC) [30]. It is based on the interferometric combination between a wave-front and a spatially filtered π-dephased wave-front (Figure 9-22). When recombining the two arms, a destructive interference occurs between the complex amplitude of an on-axis object and a spatially filtered part of this complex amplitude. For an off-axis object, the complex amplitude is removed on one arm by the spatial filter, and the nulling process does not take place. A convenient way to create a spatially filtered π-dephased wave-front is to place a pinhole or a monomode fiber in a focal plane on one arm of the interferometer. Spatial filtering by a finite pinhole or a fiber creates a non-uniform light distribution in the exit pupil plane. Therefore an apodization is mandatory to reach large rejection rate.

Crude lab tests have shown a reduction of the PSF central peak by a factor of 200 through the coronagraph (Figure 9-22). The use of a Lyot stop is mandatory for the HIC but the light is diffracted only outside the geometric aperture. The spatial filtering reduces the bandwidth usable with the HIC, but a contrast of 10^6 can be expected for reasonable spectral resolution (R=10 to 20). An interesting feature of the HIC is the fact that the pupil obstruction improves the achromaticity of the coronagraph. The main issues here will be manufacturing of the apodization function (chromaticity, phase) and, as for all interferometric coronagraph, the OPD stability.

Figure 9-22 Left: Principle of the HIC in terms of complex amplitudes of the main star (grey) and of the companion (black). Right: Comparison between average radial profiles obtained with a laboratory source with HIC (red line) and without HIC (green line). Both data recorded with a Lyot stop of 73%. Working wavelength: $\lambda_0 = 630$nm, $\Delta\lambda = 100$nm

9.9 Other coronagraphs

9.9.1 Phase mask coronagraph

In the π-mask coronagraph (the modification of which is the double zone phase mask serves to decrease the chromatic effect) the absorbing mask is replaced by the transparent mask introducing the phase shift. The optimal size of the mask is found to be $1.06\lambda/D$, which makes a fabrication and alignment of such a mask for OWL difficult.

9.9.2 Four quadrant coronagraph

In 4Q coronagraph the complex amplitude in image plane is split in 4 quadrants in two of which the phase shift of π is introduced. This method is very efficient for the monolithic off-axis telescope, but its efficiency drops very fast in the presence of the central obscuration. The large central obscuration ratio and again the gaps are the serious obstacles for application 4Q (as any image plane phase mask coronagraph) on OWL pupil.

9.9.3   Phase introduced amplitude apodization

The concept of PIAA coronagraph attracts the high interest for the designer of a planet finder on a monolithic off-axis telescope. The deliberately introduced phase aberration in pupil produces the PSF with a intensive narrow core and the very low diffraction wings, so that the desired contrast difference is achieved not by blocking the light, but by the re-shaping the star image. The big advantage of this system is 100% throughput. The calculation of the required pupil phase for the telescope with large central obscuration and segments is pessimistically considered as an impossible task; nevertheless the proving of that impossibility has not yet been demonstrated.

9.10 Effect of resolved star

Along with segmentation another particularity of a giant telescope is that a number of stars are spatially resolved. The evaluation of the sensitivity to the angular diameter of a solar-type star is analogous to the one to tip-tilt jitter. Obviously, the coronagraphs that are sensitive to tip-tilt errors will give worse results than the coronagraphs with less sensitivity to tip-tilt. But it does not mean that they have to be ruled out. To estimate the real impact of the stellar diameter on the coronagraph detection capabilities, the curves obtained for each coronagraph and diameter size must be compared to the limiting coronagraphic profiles that take into account the residual errors of the Adaptive Optics. The careful evaluation of these residuals is not finished so, as a first attempt to estimate the sensitivity of several coronagraphs to stellar diameter, we assumed no phase residual errors.

9.10.1 Hypothesis and definitions

The results are presented for the following parameters. Two sizes of telescope are considered: 60m and 100m. We assume a pupil perfectly round without obstruction and with no phase errors. The working wavelength is 1.65 μm and only 1 wavelength is simulated.

Coronagraphs :

- Achromatic Interfero Coronagraph (AIC)
- Four Quadrant Phase Mask coronagraph (FQPM)
- Apodized Hybrid Interfero Coronagraph (AHIC)
- Apodized Phase Mask Coronagraph (APMC)
- Apodized Pupil Lyot Coronagraph (APLC)

The coronagraphs are supposed perfect and achromatic. The size of the mask of the AHIC, the APMC and APLC is the same: 3 λ/D.

The same apodization function is used for these three coronagraphs. The energy transmission of the apodization function is 25% while the ratio between the PSF with and without apodization is 18%.

We assume a solar system located at 10 parsec. In that case, the angular diameter of the star is 0.93 mas. The same stellar diameter expressed in terms of telescope resolution (λ/D) gives: 0.16 and 0.27 λ/D for telescope diameter of 60m and 100m respectively. No limb darkening is taken into account in this simulation.

Peak-to-peak rejection rate: the ratio of the brightest intensity with and without coronagraph.

Contrast @ a given distance: the ratio of the intensity at a given distance with coronagraph to the brightest peak intensity of the PSF without coronagraph.

9.10.2 Diameter: 100 m

When the source is point-like and located on-axis of the coronagraph, the simulation is limited by numerical errors (10^{15}) for all coronagraphs except for APLC, for which the nulling is not complete. For a 100 m telescope and a solar system located at 10 parsecs, the angular diameter of the star is not negligible (0.27 λ/D). When observing such star with a coronagraph, the contributions of the off-axis sources decrease the performance of the coronagraphs. Both AIC and FQPM are more sensitive to stellar diameter than the AHIC and APMC (Figure 9-23). APLC is the best coronagraph in terms of decreasing the effect of the stellar size. This effect seems logical since AIC and FQPM are sensitive to details that can be as near to the central star as the resolution of the telescope while the apodized coronagraphs cannot detect such details. The APMC and the AHIC are very similar because we supposed here that both are perfectly achromatic. With this hypothesis, they are theoretically almost identical.
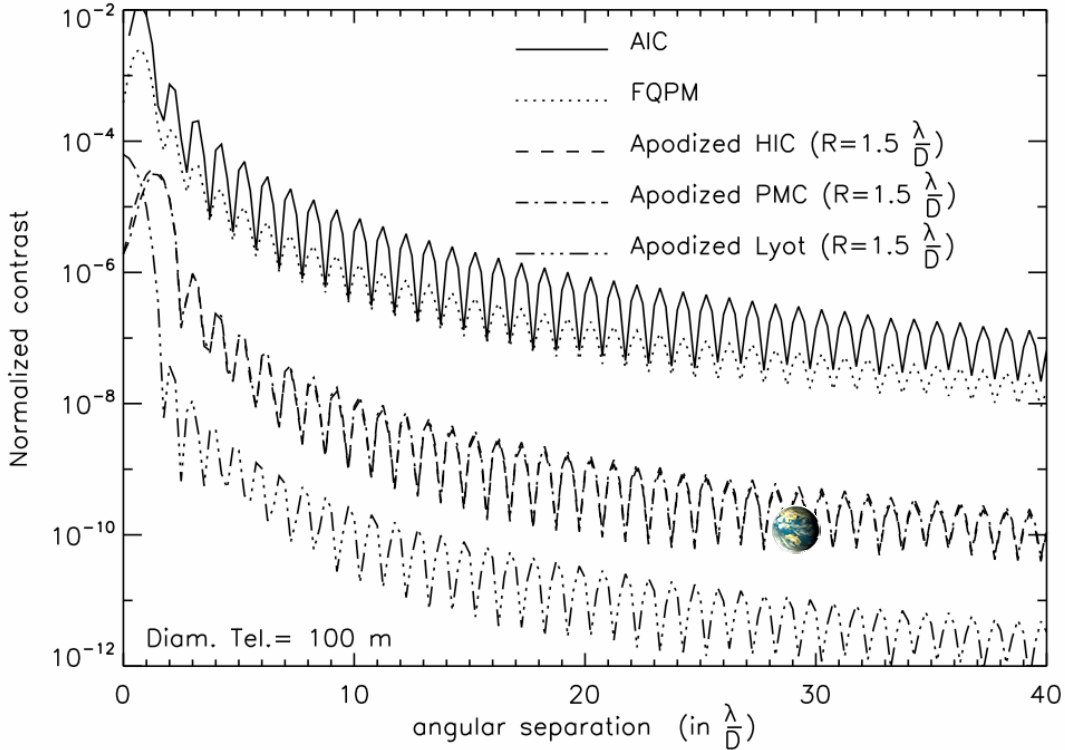
Figure 9-23 Normalized profiles for several coronagraphs on a 100 m telescope assuming a solar system located @ 10 pc (perfect coronagraphs).

If the residual phase errors of the AO are lower than 10^{-8} at 0.1'' then the AIC and the FQPM might be ruled out of the selected interesting coronagraph because they would limit the detection capability of exoplanets on ELT. A strong limitation comes also from the peak-to-peak rejection rate (Table 9-2) that does remove enough photons near the star (10^{-2}-10^{-3} for these coronagraphs). Compared APLC or APMC/AHIC, the exposure time must be 10^3 times shorter for AIC and FQPM (without saturation of the detector). Introducing AO and post-AO phase errors for all the coronagraphs will certainly smooth this difference but it has to be studied.

	AIC	FQPM	AHIC	APMC	APLC
Peak-to-peak rejection rate	$1.2 \ 10^{-2}$	$2.6 \ 10^{-3}$	$3.8 \ 10^{-5}$	$3.1 \ 10^{-5}$	$6.3 \ 10^{-5}$
Contrast @ 0.1 ''	$1.7 \ 10^{-7}$	$4.0 \ 10^{-8}$	$2.7 \ 10^{-10}$	$3.1 \ 10^{-10}$	$5.8 \ 10^{-12}$

Table 9-2 Peak-to-peak rejection rate and contrast for at the distance of an Earth-like planet for a 100 m telescope and assuming a solar system located @ 10 pc (perfect coronagraphs).

9.10.3 Diameter: 60 m

With a 60 m telescope, the sensitivity to the stellar diameter decreases as D^2, but the contrast at the angular separation of the planet varies faster, as D^3. So the contrast at the same angular distance on the sky is a bit worse than with 100 m (Tab. 2). This is true for all coronagraphs except for the APLC for which the variation is stronger. In fact, the curves of the APLC do not change with decreasing diameter because it is fully limited by the uncomplete rejection of APLC and is not sensitive to such stellar diameter. So the contrast only varies with diffraction (D^3 if no apodization were used). The peak-to-peak rejection rate varies as expected (D^2) except for the APLC for the same reason than above, it is limited by the uncomplete rejection.
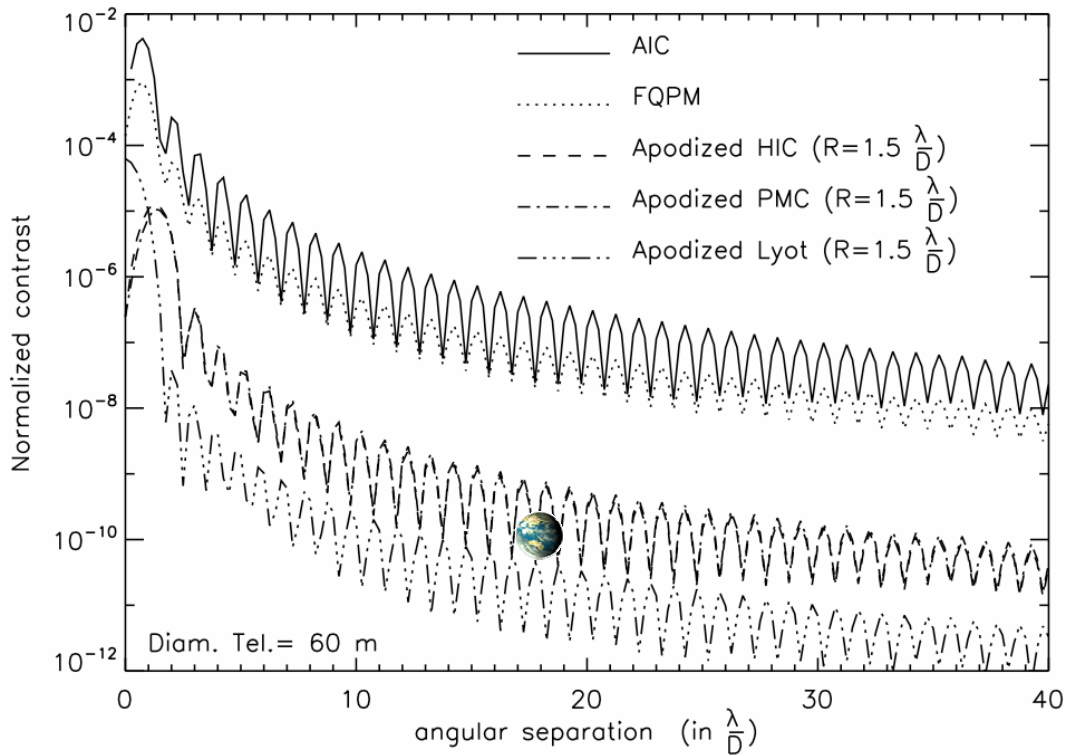
Figure 9-24: Normalized profiles for several coronagraphs on a 60 m telescope assuming a solar system located @ 10 pc (perfect coronagraphs)

	AIC	FQPM	AHIC	APMC	APLC
Peak-to-peak rejection rate	$4.2\ 10^{-3}$	$9.2\ 10^{-4}$	$1.3\ 10^{-5}$	$1.1\ 10^{-5}$	$6.3\ 10^{-5}$
Contrast @ 0.1 ''	$2.3\ 10^{-7}$	$5.5\ 10^{-8}$	$3.5\ 10^{-10}$	$4.1\ 10^{-10}$	$3.5\ 10^{-11}$

Table 9-3 Peak-to-peak rejection rate and contrast for at the distance of an Earth-like planet for a 60 m telescope and assuming a solar system located @ 10 pc (perfect coronagraphs).

9.10.4 Conclusion

The sensitivity of the coronagraph to the stellar diameter does not strongly depend on the telescope diameter if we study the contrast reached at a distance of an Earth-like planet (Figure 9-25). But the peak-to-peak rejection rate that measures the overall efficiency of the coronagraph decreases with larger diameter and could be a limit for coronagraph with close-sensing capabilities like AIC or FQPM. A study taking into account residual phase defects from the AO must be done to verify that the resolving power of the telescope is not limiting the exoplanet detection. To choose the best coronagraph for exoplanet detection we need to take into account defects, limitations and manufacturing issues of each coronagraph.
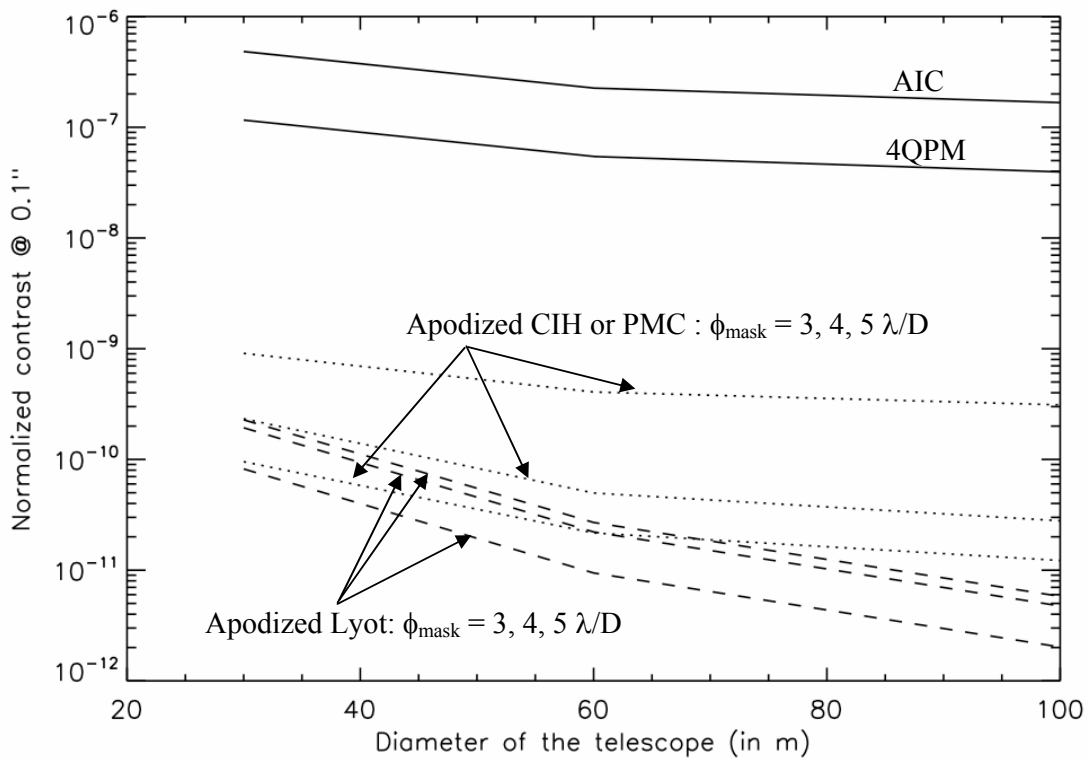
Figure 9-25 Normalized contrast at 0.1'' for several coronagraphs. Different Apodized focal mask coronagraphs are calculated (with diameter of the focal mask that varies from 3 to 5 λ/D; the higher the curve, the smaller the mask).

10 Differential Imager

10.1 Introduction

10.1.1 Differential imaging concepts

The contrast within the seeing patch of a diffraction limited 100m-telescope is limited by residual speckle noise, which can be dramatically improved by Differential Imaging (DI) methods. The DI concept to extract the signal of a very faint source against a bright halo relies on some *well understood* observational property of the faint source being very different from that of the (host-star generated) background. The main candidates for this property are the spectrum and polarization. The remit of Differential Imaging Working Group (DIWG) was to investigate an instrument mode that was based on a small number of narrow-band filter imaging channels. Differential imaging modes like polarimetry and integrated field spectroscopy will be considered elsewhere (IFU-group, Polarization-group).

Initial reduction of the halo and maintaining its stability are the function of systems up-stream of the Differential Imager itself. However it is required of the DI that it does not add significantly to the background, so very low scattering surfaces and good optical baffling must be considerations in the final design. The main advantage of DI over an Integral Field Unit spectrometer (IFU) is its simplicity and the fact that it needs to maintain very small optical path differences across only a small number (probably 2 – 4) of channels. DI's main disadvantage is the very low level of spectral multiplexing, so that it can only be applied to one or two target source (i.e. planet) spectral properties that must be understood and selected in advance.

The key up-stream systems for halo control are the AO system and the coronagraph. Note, though, that the EPICS TLR task each instrument module to have an independent coronagraph so that individual instrument performances can be optimized for their wavelength range and individual science priorities.

By consideration of the overall top-level EPICS requirements, in which the roles of the various instrument modes are defined to optimize the overall science capability of EPICS, the role of the Differential Imager has been assigned the near-IR J waveband (nominally 1100 – 1430nm). DI selection using polarimetry is best at wavelengths shorter than the J-band and is the subject of section 12 in this document, so using spectral properties is the method selected for the DI concept. The challenge is then to create two or more optical paths that can pass different (usually neighboring) wavelengths with no or minimal introduction of optical path difference across any part of the image (or pupil). The limits to allowable path differences are determined by modeling, ideally as part of an end-to-end system model in which a realistic wave-front input to the DI instrument + coronagraph allows self-consistent iteration to an overall system error budget.

10.1.2 Status of design work and analysis

The first meeting of the DIWG was a brainstorming session to help define the scope of what was necessarily to be a strictly limited (by time and amount of effort available) study. Many ideas were proposed that may be realizable through advanced technology and some of these are mentioned in the 'Mid-term Development Plan section' as being worthy of further study. However it was agreed that for a limited study it was appropriate just to try and identify and quantify the basic requirements for a simple DI design.

The final outcome is a simple optical design option using dichroic mirrors that serves to illustrate the main features required for differential imaging, and an analysis defining the main optical wave-front specifications that must be met to detect an Earth-like planet in accordance with the science requirements (see section 7.3.4). Because the main target of DI is Jupiter-like planets, important DI science would be achieved with significantly relaxed requirements. Preliminary results of calculations on image (over)sampling are also presented, relevant to TLR_4 (see section 13.2).

10.2 Requirements

10.2.1 Introduction and prior assumptions

We list two sets of requirements: those reflected or derived directly from Science and Top Level Requirements and some secondary requirements deduced for the DI mode from analysis and other specific considerations. Some of the secondary requirements are illustrative in nature to show where further design consideration or modeling must be applied.

10.2.2 DI instrument requirements

The following requirements reflect or confirm EPICS Science and TLRs that drive DI considerations.

DIR_1: The DI mode shall operate over the wavelength range from 1100 to 1430nm.

DIR_2: The DI mode shall provide a capability for the detection of a Jovian planet (requirement) and an Earth-like planet (goal), in accordance with discussion in the Science Case.

The goal requirement means achieving image contrasts of close to 10^{-10} to 5-sigma in one night of observation. The contrast requirements may be met by any combination of design feature and subsequent calibration provided these also meet overall operating efficiency requirements.

DIR_3: The DI module shall have its own Coronagraph optimized for the specified DI mode wavelength range and IWA priorities.

The coronagraph itself will not be the responsibility of the DI module. This requirement is noted so that the need for a specific interface to this individual coronagraph is properly recorded.

DIR_4: The FOV of the DI mode shall extend from less than the IWA of TLR_3 to ≥ 4 arcsec in accordance with TLR_2.

Implications for DI mode are that the separate images must not overlap in the focal plane. For (e.g.) 1 mas sampling and a 4-channel DI instrument, four 4K x 4K pixel science detectors or one 8K x 8K pixel detector would be required.

DIR_5: The image shall be at least Nyquist sampled, in accordance with TLR_4.

The need to re-sample, re-scale and subtract images, in the presence of noise and without (significant) loss of S/N ratio, is an essential requirement of DI mode. Some level of over-sampling (TBD) is almost certain to be required to achieve this. See section 10.4.

DIR_6: It shall be possible to calibrate the DI mode for image quality (focus, minimization of common-path and non-common path wave-front errors, alignment errors etc), science detector flat-field and photometric responsivity in accordance with TLR_26.

DI mode, in conjunction with other system elements, must allow for the test and optimization of system alignment and related quality control parameters.

The following instrument requirements are those derived from modeling results, the nominal design and general considerations.

DIR_7: The DI mode optical design shall meet the DI WFE specifications required to meet the science goals in DIR_2 indicated by the modeling in the DI See section 10.4.

Iteration on apportionment of error budgets across the telescope, instrument common-path optics and DI module would be an on-going process throughout the EPICS project.

DIR_8: The optical transmission of the DI mode shall be > 90% (TBC) in each wavelength channel, excluding detector QE.

Needs to be part of overall system throughput error budget. (Total EPICS throughput is currently assumed to be 16%).

DIR_9: The DI mode shall have a pupil stop or stops optimized together for the specific coronagraph design and required science throughput.

There are also other implications for the pupil stop in terms of the possible (likely?) need for rotation to match input pupils.

DIR_10: Ghosts: optical designs shall take into account the requirement to maintain a very low level of ghost images. Analysis must be provided of the level of ghosting from in or close to focus images and of the surface brightness of out-of-focus ghosts. The analysis should indicate the fraction of potential science field lost or significantly compromised because of ghosts, as a function of angular separation from the star.

DIR_11: Science detector stability and flat-field properties, after calibration, shall be such as to not degrade the detectivity of a potential planet detection by >10% (TBC).

DIR_12: The science detector R/O noise shall not degrade the detectivity of a potential planet by >5% (TBC) at any separation from the host star.

This requirement will be driven by the shortest integration times consistent with any operational requirements regarding observing efficiency, field and or pupil rotation considerations, calibration requirements and any additional steps to reduce systematic errors such as flat field noise and instrument speckle noise (see below).

DIR_13: The science detector non-linearity shall be <1% peak-valley (TBC) up to 50% full well capacity and < 3% p-v (TBC) up to 80% full-well capacity. Non-linearity shall be repeatable at given flux input and operating conditions.

These sample numbers are illustrations from the VLT-PF requirements. They may need to be more stringent for the greater contrast required for EPICS DI mode, but this has not yet been modeled.

DIR_14: The science detector shall have image persistence lower than a level that degrades DI mode planet detectivity by less than (TBD) under the required operating and calibration conditions.

'The required operating conditions' are those of integration time, exposure time, field/pupil rotation, detector non-linearity limitations etc that enable EPICS DI mode to meet the science goals for the full range of (J-band) brightness of host star and potential planet.

10.2.3 Further notes on DI requirements

Field rotation. Field rotation rates could introduce >0.2 milli-arcsec/sec of image elongation at 2 arcsec radius, or about one diffraction-limit image size every few seconds for a 100m telescope. Therefore unless system designs provide both a pupil de-rotator and a field de-rotator, the observing mode for DI will be to stabilize the field: the pupil will rotate. This has implications for the coronagraph and its apodizer and for calibration and removal of instrument speckles. For planet detections taking many hours of observation, a large angle of pupil rotation could help to smooth out instrumental effects but could also complicate calibration. The rate will not remain uniform over a wide range of hour angles; this would need to be taken into account in any plan to smooth out pupil effects.

Other operational questions that would need to be answered in a detailed design include the following (not a complete list).

Would calibration requirements lead to a need to take the coronagraph out of the optical path? If so, a very strict tolerance on repeatability of replacement would be set.

Would calibration requirements necessitate a change of filters, analogous to that suggested in the VLT-PF concept to reduce certain DI systematic errors?

Science Requirements that may not be met even if all current technical specifications are met:

TLR_8 (1% absolute photometry if SNR allows). Don't have calibration procedure that allows this. May not be possible after a coronagraph anyway, because of variable and field-position dependent throughput, especially with a complex coronagraph mask.

10.3 Implementation concept

10.3.1 Introduction

To search for exo-planets by direct imaging methods highest spatial resolution and highest available contrast is required. We concentrate here on describing simultaneous (or semi-simultaneous) narrow band imaging methods to achieve this: the full pupil is split into several (we propose four) nearly identical images, using the full pupil (and thus the full spatial resolution) of the telescope for all sub-images. The four imaging beams are observed through narrow band filters of slightly differing central wavelengths for the four images, e.g. around the absorption edge at 1.6µm (methane) (see ref. [59]).

To optimize the contrast enhancement efficiency, the different images should provide very similar wave-front behaviors: differential wave-front errors of less than 1nm are required (see section 10.4). An optical accuracy of better than 1nm RMS (633nm) in general is hard or even impossible to achieve, so a differential imager that meets the differential static aberration specs should use the same optical components for the different beams as far as possible.

In the following, three general concepts are considered, presenting different techniques of splitting the full beam into fairly identical images of the full resolution at slightly different wavelengths.

1. The pupil is split by a double Wollaston prism, similar to the solution built into NACO. Just in front of the detector, the four images are passing through a quadrant filter defining the different central wavelengths.
2. The pupil is split by means of a combination of dichroic and solid mirrors, the filters are part of this mirror system. The re-imaging optics is common to all beams.
3. The image is split in time: Within the collimated beam near the pupil image position the beam passes through a low resolution Fabry-Perot filter (FPI). The central wavelength is chopped synchronously with detector readout and charge-shuffling.

In the following advantages and disadvantages of the different concepts are discussed.

10.3.2 The double Wollaston concept

10.3.2.1 Description

This double Wollaston prism concept is illustrated in Figure 10-1. Within the collimated beam, near the pupil image to minimize differential static aberrations, a double Wollaston device is introduced. The collimated beam is required to minimize astigmatic effects of the thick Wollaston cubes. The first Wollaston prism is of similar designed to the standard

Wollaston concept, however, the crystal axis´ are rotated by 45deg relative to the splitting direction. Thus, there are two beams leaving the first Wollaston prism, which are inclined against each other by means of ±e, where (approximation):

$$\sin(e) = (n_o - n_e)\tan(\varphi)$$

and which are linearly polarized in 45 deg direction relative to the cube sides. The second Wollaston prism is a regular one, with a splitting angle perpendicular to the first one, which splits again into two beams of orthogonal direction. As the polarization direction is 45deg against the splitting direction of the second Wollaston, all four resulting beams are of identical intensity, which is about ¼ of the incoming intensity.

Just in front of the detector, as near as possible to reduce static aberration effects, a quadrant filter is inserted that defines the central wavelengths of the four sub-images.
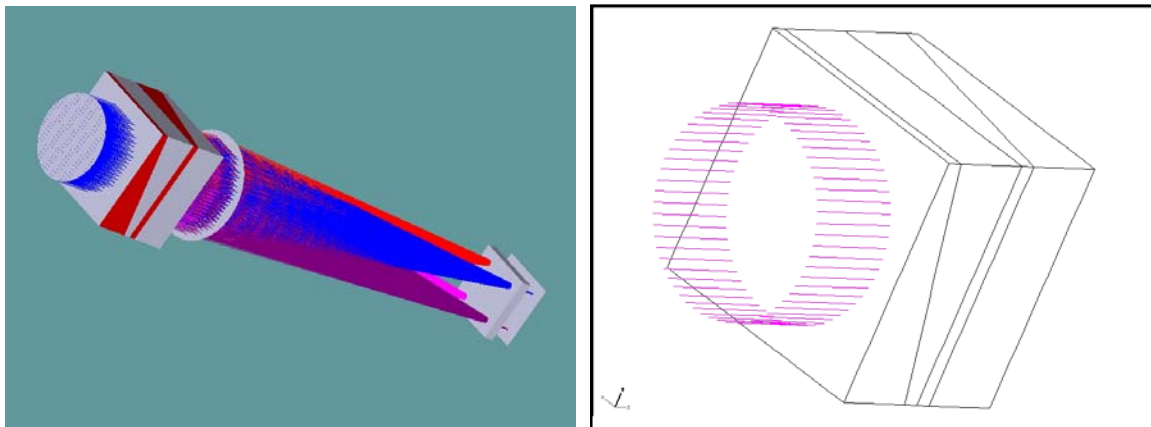
Figure 10-1: 3D overview of DI concept using crossed Wollaston prisms. Crystal axis are parallel to the wedge edge of the first Wollaston prism, and rotated by 45deg for the second Wollaston prism.

10.3.2.2 <u>Birefringent materials</u>

In Figure 10-2 all birefringent material transparent between 1.0µm up to 1.4µm are plotted with their birefringency at 1.2µm and their chromatic error within the given wavelength range. Minimum chromatic errors are expected for $AgGaS_4$, LiO_3 and YVO_4. But even for these materials of minimum chromatic effects, the variation of birefringence over a 0.1µm FWHM-filter is of the order of 1.6pm. Assuming a FOV of 4arcsec this gives a chromatic error of 5.6mas, significantly larger than the diffraction limit of a 100m-telescope at 1µm (2.06mas).
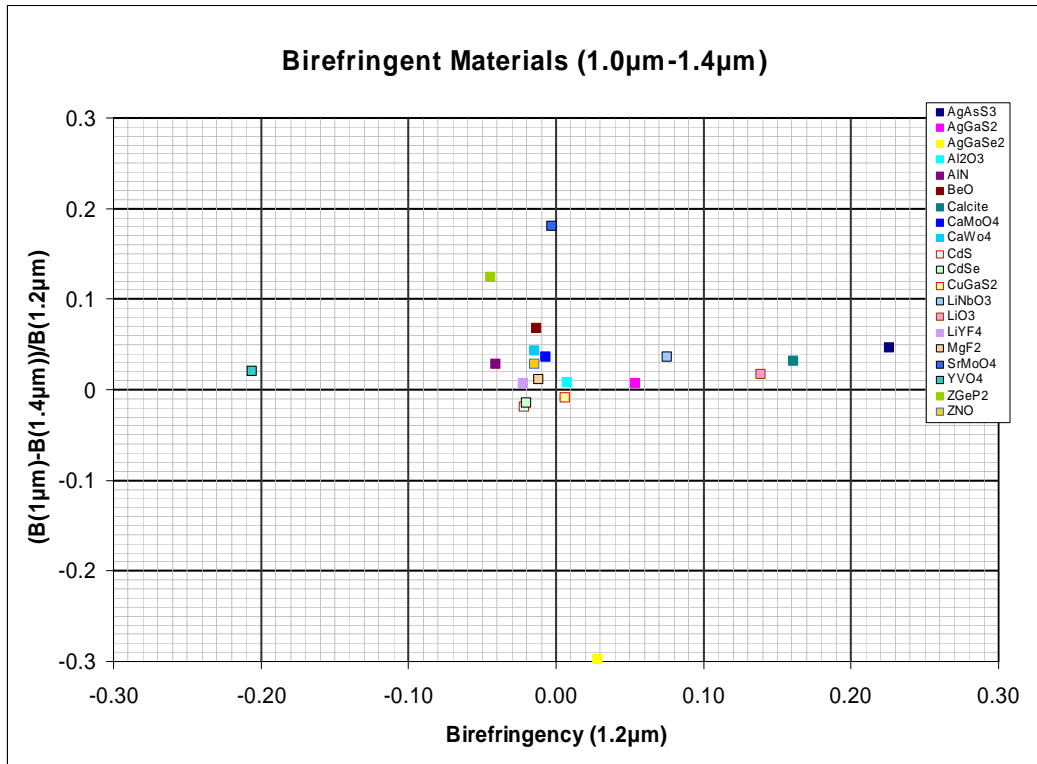
Figure 10-2: All birefringent materials transparent between 1.0µm and 1.4µm

10.3.2.3 Advantages

a) Splitting unit (Double Wollaston Prism) nearly identical for all four sub-images.
b) Re-imaging optics is small, nearly identical for all four images.
c) Different filter directly in front of the detector, resulting in small differential static aberrations.

10.3.2.4 Disadvantages

a) Wollaston spitting is wavelength dependent, resulting in different chromatic aberrations for the sub-images.
b) Throughput reduced by a factor 4 by polarizer.
c) Not useful for highly polarized sources.
d) Filter in convergent beam, additional aberration, large f-ratio required.

10.3.2.5 Tolerancing

A first tolerancing analysis has been performed for each of the three DI concepts. Only tolerances on surface irregularities using standard Zernikes (TEZI) are considered, which are the main source for differential wave-front errors. The maximum tolerance value gives the exact standard RMS value in mm (the minimum tolerance value is automatically set to the negative of the max value). The given criteria is the resulting RMS wave-front error in 633nm waves (to be compared to the nominal criteria).   Thus, for example, an RMS surface irregularity of 10nm produces a wave-front RMS error of 0.01145 waves (633nm).

```
Analysis of Tolerances

File : G:\OwnData\LensData\EPICS\Wollaston_LZ_Tol1.zmx
Title: Wollaston birefringent sample
Date : WED JUN 22 10:2:19 2005


Units are Millimeters.

Paraxial Focus compensation is on. In this mode, all
compensators are ignored, except paraxial back focus change.


Criteria          : RMS Wave-front Error in waves
Mode              : Sensitivities
Sampling          : 5
Nominal Criteria  : 0.00041131
Test Wavelength   : 0.6327


Sensitivity Analysis:

                 |------------ Minimum ------------|  |------------ Maximum ------------|
Type                    Value    Criteria     Change        Value    Criteria      Change
TEZI Wollaston in  -0.000010    0.002047    0.001635     0.000010    0.002099    0.001688
 Change in Focus                 :       0.000000                                 0.000000
TEZI requires surface 3 to be a Standard or Even Asphere surface!
TEZI Wollaston out -0.000010    0.001600    0.001188     0.000010    0.001549    0.001138
 Change in Focus                 :       0.000000                                 0.000000
TEZI Wollaston in  -0.000010    0.002053    0.001642     0.000010    0.002104    0.001693
 Change in Focus                 :       0.000000                                 0.000000
TEZI requires surface 6 to be a Standard or Even Asphere surface!
TEZI Wollaston out -0.000010    0.001628    0.001217     0.000010    0.001570    0.001159
 Change in Focus                 :       0.000000                                 0.000000
TEZI Filter in     -0.000010    0.002062    0.001651     0.000010    0.002315    0.001903
 Change in Focus                 :       0.000000                                 0.000000
TEZI Filter out    -0.000010    0.002311    0.001900     0.000010    0.002057    0.001646
 Change in Focus                 :       0.000000                                 0.000000
```

End of Run.

Used science wavelength is 1μm. Unfortunately, ZEMAX does not allow to calculate Zernike irregularities for tilted surfaces. However, the required specs for the surfaces between ordinary and extraordinary Wollaston part are expected to be $2(n_o - n_e)/ (n_o + n_e)$ times less strict than for the outer surfaces (if cemented).

10.3.3 The dichroic mirror concept

10.3.3.1 Description:

Within the collimated beam a system of four mirrors is inserted, all having a tilt angle near 45deg against the optical axis. Two of these mirrors are solid gold coated mirrors, two are dichroic mirrors. The gold mirrors are slightly tilted against the 45deg angle to create an overlapping area of all four beams where the common imaging optics can be placed. The quadrant filter is placed directly following the splitting device.

The re-imaging system consists of three components, two BAK50/FK51 cemented achromats and one single fused silica lens just in front of the detector. As an example, a three layer AR coating on fused silica is shown which demonstrates that for the given wavelength region effective AR coatings can be found which provide less than 1pm reflection losses. The reflection losses at the interface between BAK50 and FK51 would be less than 7pm.
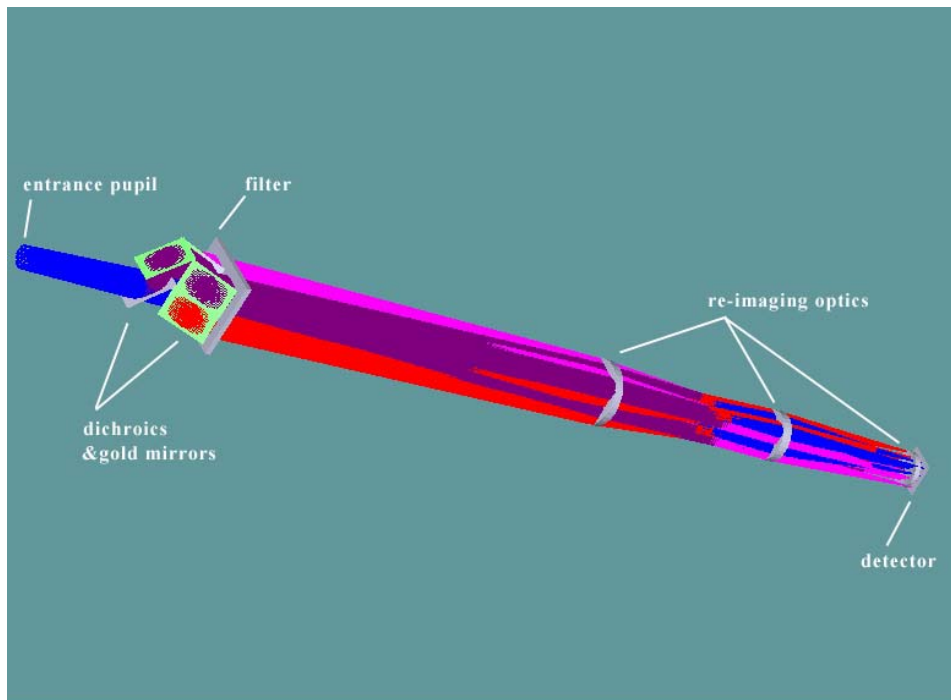
Figure 10-3: 3D overview of dichroic beam slitting design

Figure 10-4: Expanded view of dichroic mirror arrangement + quadrant filter

The resulting imaging quality is extremely good for all four channels all over the FOV and over the full wavelength range. Theoretical Strehl ratios are always better than 97% (as an example see Figure 10-5). The collimator in front of the dichroic system is not yet designed, it has to be adapted to the final f-ratio and exit pupil position of the coronagraphic module.

10.3.3.2 Advantages

a) Strictly achromatic image splitting
b) High efficiency
c) Filter within collimated beam, no additional aberrations

10.3.3.3 Disadvantages.

a) 2 to 4 individual optical components (mirrors) of very tight specs required for each beam.
b) Differential static aberration of the Quadrant filter fully applies to the overall budget.
c) Long overall design (about 1m from dichroics to detector)
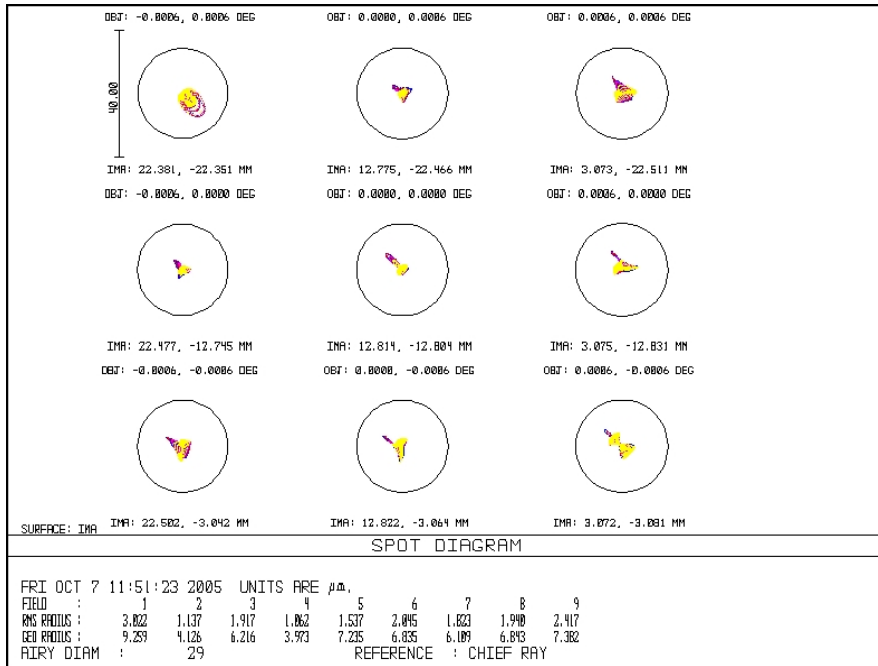
Figure 10-5: Spot diagram of one channel for a 4x4arcsec FOV. For comparison, the airy disk at 1.3µm is plotted.

10.3.3.4 Tolerancing

```
Analysis of Tolerances

File : G:\OwnData\LensData\EPICS\Beamsplitter_DI_Tol2.ZMX
Title:
Date : WED JUN 22 10:0:26 2005


Units are Millimeters.

Paraxial Focus compensation is on. In this mode, all
compensators are ignored, except paraxial back focus change.

WARNING: Boundary constraints on compensators are ignored when
         using fast mode or user-defined merit functions.

Criteria            : RMS Wavefront Error in waves
Mode                : Sensitivities
Sampling            : 5

Nominal Criteria    : 0.00000000
Test Wavelength     : 0.6328
```

```
Fields: User Defined Object height in Millimeters
 #     X-Field      Y-Field      Weight     VDX    VDY    VCX    VCY
 1   0.000E+000   0.000E+000   1.000E+000  0.000  0.000  0.000  0.000

Sensitivity Analysis:

                |------------ Minimum -----------|  |------------ Maximum -----------|
Type                  Value    Criteria     Change       Value    Criteria     Change
TEZI Mirror 1     -0.000010    0.011450    0.011450    0.000010    0.011285    0.011285
 Change in Focus                 :       0.000000                              0.000000
TEZI Mirror 2     -0.000010    0.011578    0.011578    0.000010    0.011782    0.011782
 Change in Focus                 :       0.000000                              0.000000
TEZI  Filter in   -0.000010    0.000989    0.000989    0.000010    0.000989    0.000989
 Change in Focus                 :       0.000000                              0.000000
TEZI  Filter out  -0.000010    0.000992    0.000992    0.000010    0.000992    0.000992
 Change in Focus                 :       0.000000                              0.000000


Estimated Performance Changes based upon Root-Sum-Square method:
Nominal RMS Wavefront   :   0.000000
Estimated change        :   0.016359
Estimated RMS Wavefront :   0.016359
End of Run
```

Used science beam wavelength is 1µm. Main contributors to the wave front deterioration are the dichroic and gold mirrors, which are fully contributing to the differential static aberration budget. The influence of the filter is by a factor of 10 smaller.

10.3.4 The frequency-switched Fabry Perot

10.3.4.1 Description

In this case the optical system is just the same for the different wavelength channels, the wavelength is chopped in time using a tuneable Fabry-Perot-Device within the common collimated beam. To freeze-out the speckle pattern, a chopping frequency well beyond the typical speckle correlation time has to be applied. Synchronously with the FPI-chopping, the image is shuffled from one memory to another. Even though in general this technique allows for several chopping positions, because of timing problems a two position mode will be preferred.

10.3.4.2 Advantages

   a) Optical components really identical for all sub-images
   b) Compact design

10.3.4.3 Disadvantages

   a) No real simultaneous imaging
   b) High read-out noise
   c) Deterioration of image quality by charge shuffling

10.3.4.4 Tolerancing

No differential static aberrations are expected due to the identical optical path.

10.3.5 Illustration of filter choice in J waveband

This section illustrates a set of four filters that can be used either for CH_4 or for H_2O spectral differential imaging in the J-band using multiple wavelengths. In principle this allows detection of either a "Jovian" or a terrestrial planet (or both!) in a single observation, although the effects of the Earth's atmospheric absorption water vapor spectrum on the observations must be taken into account very carefully (see section 13.3.2).

Figure 10-6 shows a "Jovian" planet spectrum with a methane feature and the relevant terrestrial atmospheric windows to be representative of a water feature. The CH_4 bands are centered around 1500nm and 1400nm.

Figure 10-6: Emergent spectrum of a class I ("Jovian"; ~5 AU) EGP from Sudarksy et al, ApJ 588, 2003. The temperature in the outer atmosphere is cold enough (≤ 150 K) that ammonia condenses into ice, which provides significant reflection.

Figure 10-7 displays the atmospheric transmission produced using the program IRTRANS4 with the following parameters: Altitude: 4200m, Airmass: 1.0, H_2O column: 1.2mm, Resolving power 3000. H_2O bands are centered around 1130nm and 1380nm.

Figure 10-7: The atmospheric transmission versus wavelength for Mauna Kea (see http://www.jach.hawaii.edu/UKIRT/astronomy/utils/atmos-index.html)

Filter	Central wavelength	FWHM	$\lambda/\Delta\lambda$
1 ('on-line')	1140 nm	75 nm	~ 15
2	1220 nm	80 nm	~ 15
3	1300 nm	85 nm	~ 15
4 ('on-line')	1385 nm	90 nm	~ 15

Table 10-1: Proposed filter set.

Filters 1 and 4 would be 'on-line', i.e., a planet with either CH_4 or H_2O would be significantly dimmed by them, while it would appear bright in the other filters 2 and 3. Three-wavelength imaging to further reduce chromatic speckle residuals would also be possible. The spectral resolution is kept nearly identical to avoid relative speckle elongation.

10.3.6 Ghost images

Special care has to be taken to avoid ghost images within the differential imager (the same is true for the whole optics of EPICS). The main components that could produce ghost images are the collimator, the slitting unit (which is the system of dichroics for the favorite concept), the band pass filters and the re-imaging camera system.

1. Imaging optics has to be AR-coated to a very high percentage. HEAR better than 99.9% per surface are feasible for the given small wavelength region. Thus, total ghost image energy will be below 10^{-6} of the transmitted star energy (which is already reduced by a large factor by the coronagraphic device).
2. Bandpass filters should be tilted to avoid reflection between detector and filter.
3. Plane-parallel optics (dichroics and filters) should be slightly wedge-shaped or plane-parallel to a very high accuracy.
4. The imaging optics (collimator as well as camera optics) should be optimized concerning minimum ghost effects. No ghost focus should be found near the focal plane.
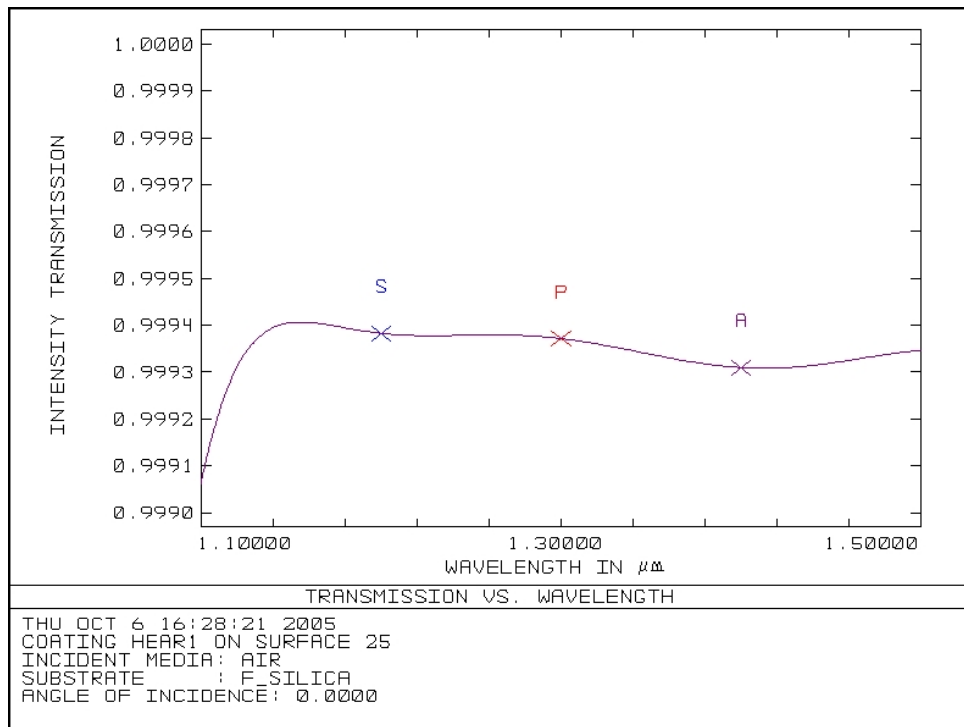
Figure 10-8: High efficiency anti-reflection coating on fused silica for the wavelength region 1.1µm to 1.5µm.

10.3.7 Elements in the Final Design

The following sub-units will be or are likely to be elements of the final design for the DI mode.

*   System Configuration Software

*   Mechanisms: Coronagraph assembly (minimum capability will be fine adjustment for alignment, possible requirement for removal from optical path for calibration purposes); Stop Wheel Assembly

(likely requirement for rotation); possible requirement for exchanging filters/dichroics to reduce WFE systematics; detector focus mechanism; possible detector X,Y translation mechanism.

- Detector Sub Assembly (including above-mentioned mechanisms)

- Detector

- Data Analysis Software

- Infrastructure: vacuum vessel, radiation shield, mechanical support, power, LCU, instrument workstation, handling equipment.

Note: The vacuum vessel and radiation shield are needed at least to house and cool the detector. Because the DI mode is defined for the J-band only, the dominant (sky + telescope + instrument) background will be OH airglow. Thermal emission from sky, telescope and instrument surfaces will be negligible through the nominated band-pass **BUT** if the detector used is sensitive past a wavelength of ~ 1.6μm a cold, broad-band J filter and/or a cold stop/baffle will be needed to limit the total background seen by the detector. Other than this, optics will not need to be at cryogenic temperatures. The location and quality of any such filter will need to be carefully considered to minimize ghosting and additional aberrations.

10.3.8 Summary

To design a differential imager for a 100m-telescope using a FOV of at least 4x4 arcsec, three concepts have been presented. The most powerful solution especially for mid- to broad-band imaging is the dichroic approach. No chromatic imaging errors have to be taken into account, the sub-images are strictly achromatic. In addition, for this solution the throughput is optimum - no splitting by polarization is applied.

However, this solution requires maximum care regarding optical quality especially of the dichroic filters and bandpass filters. Fortunately, there is no need to cool down to cryogenic temperatures (but see section 10.3.7 with regard to blocking broad band radiation if the detector is sensitive to wavelengths longer than ~ 1.6μm). Optical quality insurance for coated filters under cryogenic environment is quite hard to achieve if values better than λ/5 are required. A separate study should clarify if the optical requirements for dichroic and bandpass filters can be met.

10.4 Predicted Performance

The final performance of the differential imager depends on a huge amount of parameters.

First of all the performance of adaptive optics sets the intensity of the halo against which the planets have to be detected by differential imaging. Thus the main error sources from the AO output is the photon noise (speckle noise is most of the cases negligible after image subtraction, see section 13.4.2) from the star scattered light and has a direct impact on the needed integration time to attain the goal SNR.

The analysis in section 13 deals with the performance obtained considering AO residuals derived from the models of section 8.4.2. No hypothesis was made on the systematic errors

since those don't have an impact on integration time but set the ultimate performance one can achieve.

The problem of systematic errors in differential imaging is complex and involves the following main error sources:

- static part of the wave-front in entrance of the coronagraph

- intrinsic performance of the coronaragraph (level of diffraction residuals)

- amplitude errors in the pupil (mainly from segments ageing)

- differential chromatic aberrations occurring before and after the coronagraph, with a significant difference in sensitivity between those two

Since the systematic errors depend on the concept design as a whole and in particular on the common path optics, they have been analyzed in section 7.3.

10.5 Mid-term development Plan

- The development plan for DI Mode itself needs the following elements.

- Adoption and inclusion of a coronagraph design from the Coronagraphy analyses that is optimized for DI application.

- Decision on appropriate pupil stop(s), apodizer(s) and the need to rotate them.

- Assumption made about detector long wavelength cut-off, so that design can include appropriate stop, baffle or filter(s) to prevent broad-band, 'warm' radiation reaching the detector.

- Definition of internal calibration requirements

- Placement of the resulting optical design in a mechanical framework.

It must be done in the context (a) of an iteration of the various sub-systems modeling to gain a first set of consistent error budgets for optical surface quality and alignment stability and (b) an overall system calibration and operational model

Iterate these to thermal and operational stability requirements of mechanical mountings.

In addition, identify and initiate appropriate further technology development to eliminate or mitigate risk areas in terms of error budget items that cannot be met at present. The following table illustrates some of these areas.

Problem / topic	Impact	Specification	Where are we?	Prognosis / Mitigation
WFE before coro mask (e.g. lateral chromatism)	Major	80 pm	Analysis adequate to indicate risk.	Difficult. ADC early in optical path; additional ADC for each mode?
DI path WFE	Major	1 nm for Earth-like to 3-4 nm for Jupiter-like	Analysis adequate to indicate risk. ~ x10 beyond current optics?	Difficult but is focus of technology dev on optical surface quality. Also consider options that remove multiple optical paths: frequency switching or 3D detectors?
Instrument / telescope speckles	Major	10^{-9} contrast at J at 10 λ/D 10-20 % bandwidth	HCIT TPF, $5 10^{-9}$ at 800 nm10% bandwidth	Focal plane WFS. Should part of future analysis / desi development.
Throughput	Medium	16%		Continued development of AR coatings. Simplification of design to reduce # of surfaces.
Photometric accuracy	Minor?	1%	Not analyzed	May not be possible in DI mode
Metrology: initial alignment to achieve WFE budget; mechanical design / calibration to maintain it.	Major	As per WFE specifications above	Not analyzed	Method must be developed along with system design.
Coatings (filters and AR)	Major		Similar to optics specification generally: ~ x10 from requirements	Promote further progress by optics companies through national and EU technology development funding

11 Integral Field Spectroscopy

Integral Field Spectroscopy (IFS), coupled with Extreme Adaptive Optics (XAO) and Coronagraphy is a powerful technique for detection and characterization of extrasolar planets, including extrasolar earths which is the prime science goal of EPICS.

The purpose of IFS for planet finding is to exploit the spectral features of the sub-stellar companions in order to increase the contrast on the stellar background and to use Simultaneous Differential Imaging (SDI) techniques to remove the speckle noise. This is the dominating noise source in the region between the AO control radius and the Coronagraph inner working angle and makes direct detection of exo-planets extremely difficult.

Before choosing an IFS concept, optical analysis based on deep understanding of how to handle the residual speckles pattern is required. Its behavior has to be analyzed in terms of spectral and spatial power spectrum via software simulations and laboratory experiments. The basic technical requirement for this instrument is given in terms of its capacity to overcome this speckle noise. Other requirements depend directly on the specificities of each optical concept.

Two IFS concepts are described in the following, each in principle able to aliment an SDI-type analysis. The first is a Tigre-type IFS and the second in a Fourier Transform Integral Field Spectrograph (FT-IFS).

11.1 Options and Requirements

11.1.1 Global Requirements and design considerations

The formal top-level requirements defined for EPICS states that all observing modes shall have 2" (Vis) to 4"(NIR) FOV, with at least Nyquist sampling, corresponding to 4k x 4k image pixels in both bands. This is extremely demanding for 3D spectroscopy, in particular when spectral and spatial information is recorded simultaneously, such as in Slicer and Tigre-type instruments. The Slicer option (Figure 11-1 a) is more efficient in terms of pixel usage, since no separation between spectra from adjacent pixels is required. The number of detector pixels (N_{Det}^2) required for a given number of spatial pixels (N_{Im}^2) and spectral samples (N_{Spec}) is given by:

$$N_{Det} \sim N_{Im} \sqrt{N_{Spec}}$$

In the Tigre-type option (Figure 11-1 b), where spectra are interleaved in a more intricate manner, a separation between spectra is necessary. For a separation of N_{Sep} pixels, the required number of detector pixels becomes:

$$N_{Det} \sim N_{Im} \sqrt{N_{Spec} \cdot N_{Sep}}$$

For a typical separation of 5-10 pixels, depending upon detailed design and cross-talk requirements, the Tigre-type is up to 3 times hungrier in terms of linear detector dimension than the Slicer-type option.

(a)	(b)	(c)

Figure 11-1. Schematic comparison of three Integral Field Spectrograph concepts considered for Planet Finding instruments

In the Slicer option (a), the image is cut into a series of pixel-wide slices which are organized to form a virtual long slit at the entrance of the spectrograph. In the Tigre-type option (b), The image is chopped into pixel-sized squares using a micro-lens array. The flux within each lens is compressed into small dots (micro slits) at the entrance of the spectrograph, leaving space for spectra to be dispersed in the intervening space. In the Fourier Transform option (c) the full image is projected onto the detector array after passage through an interferometer. The spectral information is collected over time in the form of an interferogram for every pixel of the image by scanning the optical path difference of the interferometer.

In spite of this obvious advantage, the Slicer option has been eliminated from the EPICS study based on arguments developed in the PF phase A study. The two main arguments concern differential aberrations and instrument size and complexity. A 180x180 pixel Slicer instrument was designed to a great level of detail [42]. Although within reasonable limits in terms of volume and complexity, extension of this design to 4kx4k pixels would be extremely cumbersome, even in ELT standards. The design was finally abandoned as a VLT-PF option due to an argument of differential aberrations: The goal of speckle elimination is to compare and subtract speckle formed at different wavelengths. But since the speckle pattern is chromatic (see Figure 11-2), corresponding speckles in different wavelengths do not fall at the same spatial pixel, hence not on the same slice. The speckles

216

therefore follow different routes through the following optics, and will be affected, in the along-slice dimension, by different aberrations.

(a)   (b)
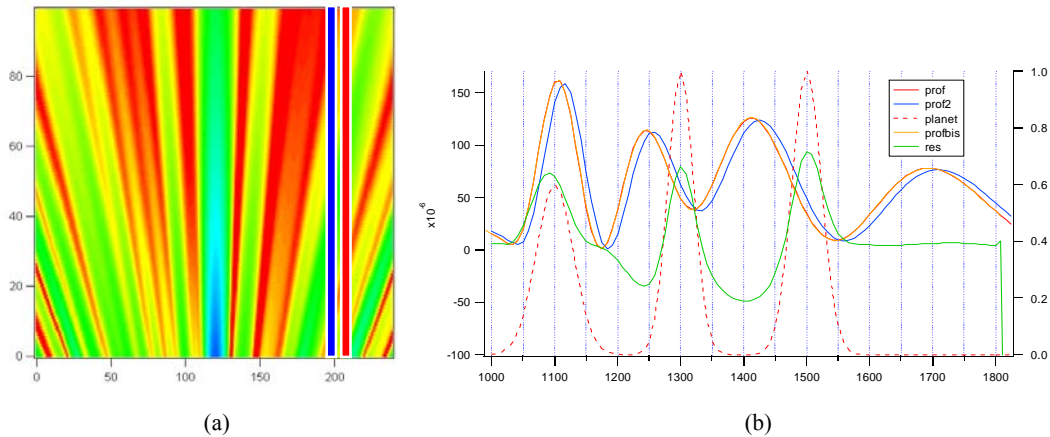
Figure 11-2. Slice through a 3D data cube corresponding to a speckle pattern obtained with an "ideal" coronagraph (a). The horizontal axis represents spatial position and the vertical axis represents the spectral dimension, covering an octave, linear in wavelength. Speckles are seen as streaks ("spaghetti") running through the data cube obliquely, hence creating speckles in the spectral dimension. The spectra for two radially neighboring spatial pixels are extracted in (b), the red and blue curves corresponding to the red and blue lines in (a).

Clearly, the spectra are identical, apart from a slight shift and scale difference. A planetary spectrum (broken line) has been added to one of the pixels at a level of 1e-6 with respect to the peak stellar flux. The green line shows this spectrum recovered (x100) after subtraction of scaled and shifted speckle spectra, illustrating an IFS-based Simultaneous Dual Imaging technique. The demonstration ignores any problems related to spectral aliasing, cross-talk, or differential aberrations.

The Tigre option avoids this problem because the image is sampled in both dimensions at the entrance of the spectrograph. The advantage is real only insofar as the cross-talk between spectra is controlled, however, hence the need for a sufficiently large N_{Sep}. The spectral separation must be considered an advantage, rather than a disadvantage in favour of the Tigre option, and the difference between Eqs 1 and 2 reflects the cost of solving the cross-talk problem.

For the FTS (Figure 11-1 c) the situation is of course different. Here, the spectral information is obtained by scanning the optical path difference (OPD) between the arms of a Mach-Zehnder interferometer. The detector is therefore entirely dedicated to the spatial dimensions, so that:

$N_{Det} = N_{Im}$

This constitutes an enormous advantage for the FTS, both in terms of detector real-estate (the detector cost is a major part of present-day instruments) and instrumental dimensions and complexity, which is strongly related to the size and resolution of the final image, i.e.

the number of detector pixels, N_{Det}. The interferometric setup and cryogenic scanning mechanism certainly adds complexity to the system, but the entire 4k x 4k instrument fits well within a trash-can sized cryostat.

Other advantages of the FTS include its flexibility in terms of spectral band and resolution. The same instrument can be used throughout the efficiency band of its detector, possibly extending a standard "NIR" instrument both into the visible and out to K and even further.

In terms of signal-to-noise (SNR) performance, a study (Sec. 11.3.4.1) concentrating on low spectral resolution as required for EPICS, shows that the FTS has an advantage in systems limited by detector and background noise. Its preferred domain from that point of view is therefore the infrared where detectors (still) have considerable read-out noise and where a cooled FTS will have considerably less background than a Tigre-type instrument, especially when the latter is only partially cooled.

The main drawback for the FTS is the variability of speckles in time, and we shall discuss ways to limit, if not eliminate this drawback in Sec. 11.3.1. It seems even that in certain observing modes, the speckle variability can be used to our advantage. However, all this is fairly speculative and needs intensive further study in order to be fully explained and exploited.

In view of all this, both a Tigre and an FTS instrumental concepts are described. The Tigre is specified for a 2"x2" FOV with a goal of 4"x4", covering the H band where main spectral features due to CH_4, H_2O, CO_2, characteristic of expected exo-planets, are located (see Figure 11-3). Taking full advantage of its reduced size and efficient use of detector pixels, the FTS is specified for a 4"x4" FOV, covering at least the Y, J, H bands with an extension into the visible and/or K band as goal. In particular, it would seem interesting to include the O_2 forest at 760nm as well as CH_4 and CO_2 features in the K-band.
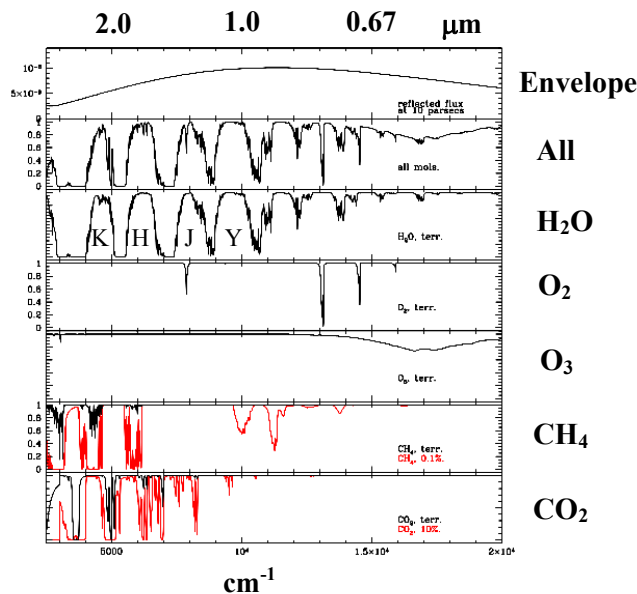
Figure 11-3. Decomposition of the earth's atmospheric absorption spectrum.

11.1.2 TIGRE IFS Requirements

11.1.2.1 Spectral domain

TLR_1: The TIGRE IFS covers the wavelength range 1.38 – 1.75 microns (i.e. the H-band). The main argument for this choice is that the H-band gets all the features useful for detection and characterization of earth-like planets: CO_2, CH_4, H_2O.

11.1.2.2 Field of View

TLR_2: The present TIGRE IFS concept covers the minimum requested field of view : 2″ diameter. No conceptual show stoppers exist for its extension up to 4″ diameter, but cost and complexity would be much larger.

11.1.2.3 Strehl Ratio

TLR_3: The lower limit considered in the present TIGRE IFS concept is SR = 0.95 (H-band).

11.1.2.4 Spectral Resolution

TLR_4: The present TIGRE IFS concept is developed to cover the low spectral resolution mode in the H-band: $R_{2pixels} > 15$.

11.1.2.5 Operation

TLR_5: The operation of TIGRE-IFS should follow the operation standards defined for OWL. While these are not yet fully defined, we expect these will include e.g. remote operability, operation through predefined Observing Blocks's, definition of templates and observing modes, and well defined safety rules.

TLR_6: All hardware needed for the calibration procedures of TIGRE-IFS should be implemented: these will include at least detector and lenslet flat fields, detector dark current and biases determinations, wavelength calibration, cross talk determination.

11.1.3 FT IFS Requirements

11.1.3.1 Particularities and specifications of the optics

The Planet Finder has to provide the highest possible contrast and Strehl ratio on the detector. The detectivity of a planet is further increased if the signal can be analyzed simultaneously in more than one spectral band, so that one can discriminate and reduce the "speckle" noise. Spectral analysis is also required for characterization of newly found objects

11.1.3.2 Spectral domain

The FT-IFS should be designed for optimal performance within the NIR covering at least the Y, J, H atmospheric bands. Extending this to the 0.8 (0.75?) to 2.5 micron range covered by current-technology detectors appears feasible in terms optical efficiency (coatings etc) and should be considered as a goal. This goal could be further extended in function of future technological advances.

11.1.3.3 Field-of-View

A FT-IFS based on a 4kx4k detector and Nyquist sampling at the shortest design wavelength covers a FOV of 4" for the basis version (1µm) and 3.2" for the goal version (0.8µm).

11.1.3.4 Spectral resolution

A resolving power of R~15 as required for observation of the H-band spectral features requires a scan length of only 10 µm. The extension to R=100, requiring less than 0.1mm scans can be considered as trivial. A scan length of 5-10mm (R~10 000) should be seen as an upper limit of the proposed concept.

11.1.3.5 Operation

Various observational modes should be considered, allowing optimization in terms of integration time per OPD step, integration time per scan, and total integration time. While very fast operation time allowing scanning faster than the atmospheric coherence time may be optimal, this is clearly not realistic in terms of detector performance. Step-and-integrate operation is therefore assumed, probably with individual integration times of the order of seconds. The possibilities offered by the non-destructive read process must be studied and optimally used.

TLR_1: The instrument covers the wavelength range 0.6 – 1.75 micron

11.2 TIGRE IFS Implementation Concept

This Section describes the TIGRE IFS concept for extreme-AO planet finding instruments. Contributors: Jacopo Antichi, Raffaele Gratton, Massimo Turatto, Stefano Andolfo, Alessandro Berton, and Harald Kuntschner.

The main scope of an IFS within a high contrast imaging system aimed to detect planets is to fully exploits the potentials of spectral differential imaging. With respect to simple differential imaging obtained through use of filters, the IFS offers a multiplex advantage over the number of usable spectral bands: this is of paramount importance for detection of terrestrial planets, since their spectra may present a variety of bands, that cannot be predicted in advance. As mentioned in Sec. 11.2, the TIGRE type design possibly reduces some of the main concerns of differential imaging, in particular those related to non common path aberrations, because essentially the same optical path is followed by rays of different wavelength through the IFS (neglecting the small variations due to chromaticity of the optics and the small effect due to the dispersing element). On the other hand, the speckles themselves are chromatic (see Figure 11-1). This is potentially a severe problem in TIGRE type design, whose impact should be carefully examined. Construction of monochromatic images extracted from the spectral data cube (needed within the differential imaging procedure) requires interpolations between the data provided by individual pixels. Spectral undersampling and inhomogeneity of illumination at sub-pixel levels are a potential source of errors in such interpolations, possibly severely reducing the accuracy of differential imaging, and then reducing the sensitivity of the IFS. This is in particular a serious concern for the TIGRE IFS, because in this system micro-pupils generated by the lenslet array are re-imaged onto the detector. The shape of these micro-pupils is set by diffraction at the individual lenses, and it is not uniform, depending on the illumination of the individual lenses. For this reason, it should be expected that adequate sampling of the micro-pupil images is required in order to properly take into account these variations. This require significant magnification by the IFS optics, and huge detector space. We call this solution Super-TIGRE, to underline the requirement for very large detectors. While technically possible, this solution is extremely expensive. On the other hand, the OTF of the IFS itself, as well as the box-car sampling due to the detector, act in the sense of cutting the high frequencies, possibly alleviating the concern related to pupil undersampling. How much speckle chromatism really degrades the performances of differential imaging based on the TIGRE IFS concept with undersampled micro-pupils is currently under study for the VLT-PF, and a final answer is still pending.

The problem of pupil undersampling can be also solved in alternative ways. One possible practical solution is to re-image on the detectors not the micro-pupils, but rather images of the micro-lenses. This can be achieved by coupling each-microlens of the TIGRE design (sampling a small portion of the FOV), with a second suitable microlens. The focal length of this second microlens should be quite shorter than the first one, in order to avoid confusion between the spectra. Due to the presence of two microlens arrays, we call this second solution BIGRE. The basic equations ruling the BIGRE concept have been developed within the VLT-PF study, as a possible back-up solution in case the TIGRE IFS concept should

reveal inadequate to properly treat the speckle chromatism. In that case, optical solutions within spec's seem possible.

 Hereinafter, we will present a normal TIGRE concept for the EPICS-IFS, with a few comments about a Super-TIGRE concept too. The results of the VLT-PF studies will show how well the normal TIGRE solution may match the EPICS-IFS requirements. In case of negative answers, it will be likely possible to develop a BIGRE design with quite similar characteristics.

11.2.1 Optical Concept

The number of lenses shaping the lenses array (i.e. the spatial sampler device) of a TIGRE type IFS has to be equal to the number of angular resolution elements contained in the field of view. This is because the instrument works at the diffraction limit and the Nyquist sampling of the Telescope LSF is necessary in order to fully exploits the potentials of Simultaneous Differential Imaging.

Considering the minimum field of view ($2''$ diameter) and the telescope angular resolution element ($\lambda/2D_{TEL} = 1.42 \times 10^{-3}''$, $\lambda = 1.38$ micron, $D_{TEL} = 100$ meters), the numerical size of the lenses array is 1000 x 1000.
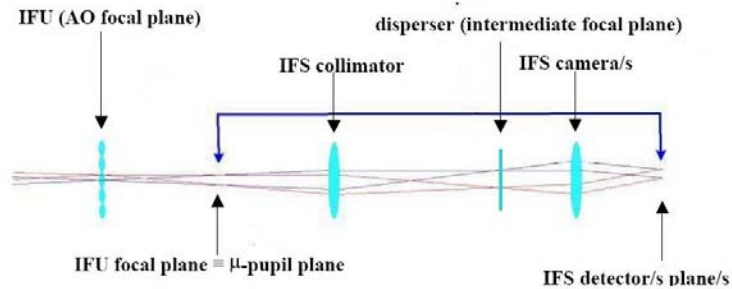
Figure 11-4 : Optical concept of aTIGRE type IFS for planet finding: the lenses array samples at the Nyquist limit the Telescope LSF.

Figure 11-5 : Optical implementation of the H-band CHEOPS IFS. Note that the disperser is placed on an image of the Telescope focal plane. This happens because the IFS object plane is a micro-pupils array which is imaged and dispersed on the detector pixels

11.2.1.1 Fore-optics focal ratio

For the determination of the fore-optics focal ratio (FR_{fore}), we have to impose the Nyquist sampling of the telescope LSF on the lens-array plane which is an image of the Telescope focal plane (see Figure 11-4). The adopted Nyquist spatial sampling is fixed as follows.

The maximum (i.e. evaluated at the minimum wavelength) spatial frequency band of the Telescope LSF, projected on the lens-array plane is:

$$B_{LSF-Tel} = \frac{1}{\lambda_{min} \cdot FR_{fore}}$$

The Nyquist criterion implies:

$$D_L = \frac{\lambda_{min} \cdot FR_{fore}}{2}$$

where D_L is the pitch of the lens-array. The requested fore-optics focal ratio then is:

$$FR_{fore} = \frac{2 \cdot D_L}{\lambda_{min}}$$

11.2.1.2 Length of the spectra on the detector/s

In the TIGRE concept the entrance slits of the IFS are micro-image of the exit pupil of the telescope and they form a micro-pupils-array. The micro-pupil-array is obtained by the lens-

array which is placed on an image of the Telescope focal plane, this latter working as a field-stop-array. The telecentricity of the input beam guarantees that the micro-pupils-array is focused on the focal plane of the lens-array.

The spectral purity of the system is:

$$R_{2\,pixels} = \frac{\lambda_C}{\Gamma \cdot W}$$

where λ_c is the wavelength at the center of the range, W is the spectral resolution and Γ is the maximum between the Nyquist sampling limit on the detector/s plane and the FWHM of the spatial shape of the slit evaluated at the central wavelength ($S_{slit}(\lambda_C)$, see Sec. 11.2.1.3), even on the detector/s plane:

$$\Gamma = MAX\{2, FWHM[S_{SLIT}(\lambda_c)]\}$$

Then the spectral resolution in microns per pixel is:

$$W = \frac{\lambda_c}{\Gamma \cdot R_{2\,pixels}}$$

and the spectrum length in pixel units (l_S) is:

$$l_S = \left(\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_c}\right) \cdot R_{2\,pixels} \cdot \Gamma$$

11.2.1.3 Spatial shape of the monochromatic entrance slit

The monochromatic slit is a micro-pupil i.e. a micro-image of the exit pupil of the telescope.

The Nyquist sampling of the diffraction limited Telescope LSF made by the lens-array implies, for dualism principles between pupil planes and focal planes, that any microlens undersamples the exit pupil of the telescope. This fact makes the monochromatic signal on the micro-pupil plane equal to the single-lens OTF (OTF_L). Moreover, due to the position of the single lens in the optical train (i.e. a field-stop position) the diffraction limited shape of its monochromatic OTF is a perfect monochromatic Airy-pattern (if the microlens illumination is uniform), meaning that the spatial shape of the monochromatic entrance slit of the TIGRE is a perfect monochromatic Airy-pattern (see details in Sec. 11.2.1.3.1):

$$S_{SLIT}(u,\lambda) = SINC^2\left(\frac{u}{\lambda \cdot FR_L}\right)$$

11.2.1.3.1 *The formalism*

In the optical train the single lens acts as a field stop this means that its monochromatic Linear Spread Function (LSF_L) is as follows:

$$LSF_L(x,\lambda) = \begin{cases} \dfrac{2}{\pi}\cdot\left(\arccos(x) - x\cdot\sqrt{1-x^2}\right), if\;|x|\leq \dfrac{1}{\lambda\cdot FR_L} \\[2ex] 0, if\;|x| > \dfrac{1}{\lambda\cdot FR_L} \end{cases}$$

The variable x represents the assigned coordinate on the lens-array plane and its physical dimension is millimeters. Then, the single-lens OTF is as follows:

$$OTF_L(u,\lambda) = \left[\dfrac{\sin(\dfrac{\pi\cdot u}{\lambda\cdot FR_L})}{\dfrac{\pi\cdot u}{\lambda\cdot FR_L}}\right]^2 \equiv SINC^2\left(\dfrac{u}{\lambda\cdot FR_L}\right)$$

The variable u represents the assigned coordinate on the micro-pupils plane and its physical dimension is millimiters^{-1}. FR_L is the single-lens focal ratio.

11.2.1.4 Format of the spectra on the detector/s

The format of the spectra must allow the best use of the detector/s area in such a way that the transversal distance (δ_S) among various spectra is constant and large enough to reduce cross talk contamination at acceptable values.

Clearly the allowed spectra layout on the IFS detector focal plane depends on the shape of the lens-array. To this aim, in the present study we have analyzed the hexagonal and square shapes which represent the best choices in terms of filling factor.

As shown in Table 11-1, we have taken into account several geometrical configurations giving different values for the length-to-distance ratio R_G, defined as follows:

$$R_S = \frac{L_S}{\delta_S}$$

(L_s is the allowed length of the spectrum and δ_S works as a separation from its nearest neighborhood) and different values for the position angle α formed by the reference direction with the direction of the dispersion.

Square	R_S	α	Hexagonal	R_S	α

case A	1	0°	case A	1.15	30°
case B	2	45°	case B	3.46	0°
case C	5	26.56°	case C	8.08	10.89°
case D	10	18.43°	case D	15.01	16.10°
case E	13	33.69°	case E	21.94	6.56°

Table 11-1: Configurations for the spectra alignment in the case of hexagonal and square shape for the lens-array.

The various configurations are displayed in Figure 11-6 and Figure 11-7.

Figure 11-6: Possible configurations for a square lens-array. The reference axis is aligned to a side of the square. Note that the symmetries of a square lens-array allow to analyze only configurations having position angles $\alpha \leq \pi/4$.

Figure 11-7: Possible configurations for a hexagonal lens-array. The reference axis is aligned to a side of the hexagon. Note that the symmetries of a hexagonal lens-array allow to analyze only configurations having position angles $\alpha \leq \pi/6$.

As explained in Sec. 11.2.1.3, the spatial shape of the monochromatic slit projected onto the detector/s plane is a perfect monochromatic Airy-pattern in the case the lenses of the array are uniformly illuminated. This estimation do not consider the instrumental OTF of the IFS-optics, from the entrance slit to the detector/s plane because this latter should involve smaller variations to the perfect Airy-pattern of the exit slit, when an IFS optical design will be optimized. Thus, the crossing factor between two adjacent spectra, evaluated in the direction orthogonal to the dispersion direction (i.e. the cross talk level, see Sec. 11.2.1.5), has to be evaluated considering two perfect monochromatic Airy-patterns, one shifted with respect to the other one of a distance equal to the requested separation level (δ_S).

The requested value of δ_S depends on the cross talk upper limit which we set to be $< 10^{-4}$ (see Table 11-3) and it changes varying the configuration. Finally, L_S depends only on the configuration one chooses to adopt for the layout of the spectra and it must verify the following inequality:

$$L_S - \delta_S \geq l_S$$

This latter guarantees that the cross talk of two adjacent spectra with respect to the dispersion direction do not exceed the specified upper limit.

11.2.1.5 Cross talk contamination of the spectra on the detector/s

The cross talk contamination of one image onto the adjacent ones is defined as the value of the LSF evaluated at the positions of the centers of the others.

In general, the signal (S) of the spectrum containing the planet signal at two different wavelengths will be:

$$S(\lambda_1) = s(\lambda_1) + p(\lambda_1) + \int_{\lambda_{i\,min}}^{\lambda_{max}} c_1(\lambda') \cdot S(\lambda') d\lambda'$$

$$S(\lambda_2) = s(\lambda_2) + p(\lambda_2) + \int_{\lambda_{min}}^{\lambda_{max}} c_2(\lambda'') \cdot S(\lambda'') d\lambda''$$

where $s(\lambda)$ is the monochromatic signal of the parent star, $p(\lambda)$ is the monochromatic signal of the planet, $c_1(\lambda)$ and $c_2(\lambda)$ are the monochromatic cross talks of the nearest neighbor spectrum onto the reference one at the position corresponding to the wavelength λ_1 and λ_2, respectively.

Basic assumptions of Simultaneous Differential Imaging are: $p(\lambda_2) = 0$ and $s(\lambda_1) = s(\lambda_2)$ (possibly after suitable normalization).

Then, the observable is:

$$O \equiv S(\lambda_1) - S(\lambda_2) \approx p(\lambda_1) + \int_{\lambda_{min}}^{\lambda_{max}} c_1(\lambda') \cdot s(\lambda') d\lambda' - \int_{\lambda_{min}}^{\lambda_{max}} c_2(\lambda'') \cdot s(\lambda'') d\lambda''$$

and the noise is:

$$N = \sqrt{err^2[s(\lambda_1)] + err^2[s(\lambda_2)] + err^2[p(\lambda_1)] + \int_{\lambda_{min}}^{\lambda_{s\,max}} err^2[c_1(\lambda') \cdot s(\lambda')] d\lambda' + \int_{\lambda_{min}}^{\lambda_{max}} err^2[c_2(\lambda'') \cdot s(\lambda'')] d\lambda''}$$

Let us assume that any single error term is a photon-noise source in a poissonian regime, then:

$$N = \sqrt{s(\lambda_1) + s(\lambda_2) + p(\lambda_1) + \int_{\lambda_{min}}^{\lambda_{max}} c_1(\lambda') \cdot s(\lambda') d\lambda' + \int_{\lambda_{min}}^{\lambda_{max}} c_2(\lambda'') \cdot s(\lambda'') d\lambda''}$$

We now express the ratio between parent star and planet observed fluxes in terms of the SR level of the global signal, the mutual star-planet contrast (I_{star}/I_{planet}) and the ratio between the number of resolution elements on which, respectively, the planet signal and the parent star signal are focused :

$$\frac{s(\lambda_1)}{p(\lambda_1)} = \left(\frac{1 - SR(\lambda_1)}{SR(\lambda_1)}\right) \cdot \frac{I_{star}}{I_{planet}} \cdot \left(\frac{n_L^{planet}}{n_L^{star}}\right)^2$$

Adopting a reference value of SR = 0.95 for the TIGRE IFS wavelength range, a contrast reference value $I_{star}/I_{planet} = 2 \times 10^{10}$, the Nyquist sampling of the planet-LSF on the lens-array plane (i.e. $n_L^{planet} = 2$) and finally, adopting the numerical size of the lenses array as the number of resolution elements devoted to collect the parent star signal (i.e. $n_L^{star} = 1000$, see Sec. 11.2.1), we obtain:

$$s(\lambda_1) \approx 4.2x10^3 \cdot p(\lambda_1)$$

This result and the further condition $c_1(\lambda') \approx c_2(\lambda'') << 1$ imply:

$$N \approx \sqrt{s(\lambda_1) + s(\lambda_2)}$$

The SNR will then be:

$$\frac{O}{N} = \frac{p(\lambda_1) + \int_{\lambda_{mon}}^{\lambda_{max}} c_1(\lambda') \cdot s(\lambda')d\lambda' - \int_{\lambda_{min}}^{\lambda_{max}} c_2(\lambda'') \cdot s(\lambda'')d\lambda''}{\sqrt{s(\lambda_1) + s(\lambda_2)}}$$

From the above equation the planet would be detectable if the cross talk satisfies the following relation:

$$p(\lambda_1) > \sqrt{err^2[\int_{\lambda_{min}}^{\lambda_{max}} c_1(\lambda') \cdot s(\lambda')d\lambda'] + err^2[\int_{\lambda_{min}}^{\lambda_{max}} c_2(\lambda'') \cdot s(\lambda'')d\lambda'']}$$

Since $s(\lambda') \approx s(\lambda'') \approx s(\lambda_1)$, and in the upper limit condition $c_1(\lambda') \approx c_2(\lambda'') \approx c$, we obtain that the above inequality is broadly verified if the planet monochromatic signal verifies the following inequality:

$$p(\lambda_1) > \sqrt{l_S} \cdot c \cdot s(\lambda_1)$$

or, equivalently:

$$c < \frac{p(\lambda_1)}{s(\lambda_1) \cdot \sqrt{l_S}} = \frac{1}{4.2x10^3} \cdot \frac{\sqrt{l_S}}{l_S}$$

Then fixing the length of the spectrum on the detector/s plane we can fix the cross talk upper limit for the detection of planet up to contrast of order of 2×10^{10}.

As described in Sec. 11.2.1.2, the spectrum length depends on TLRs (spectral resolution), on sub-system TLRs (wavelength range of the TIGRE IFS) and finally on the FWHM of the spatial profile of the exit slit.

11.2.1.6 <u>Aliasing budget on the monochromatic exit slit</u>

The stellar background is scattered into each pixel through the wave-front error function, giving rise to speckles. These appear both in the image and, due to the speckle scaling effect, in the spectrum making important the quantification of the aliasing the instrument commits on the monochromatic exit slit.

As explained in Sec. 11.2.1.3, the spatial shape of the monochromatic slit projected onto the detector/s plane is a perfect monochromatic Airy-pattern in the case the lenses of the array are uniformly illuminated (i.e. the phase variations of the speckle pattern in perfectly removed on the lens-array plane). Thus, the spatial frequencies band of the slit is bounded and its maximum band (i.e. evaluated at the minimum wavelength) projected on the detector/s plane and normalized to the pixel size is:

$$B_{SLIT} \equiv B_{MTF-L} = \frac{d_{pixel}}{\lambda_{min} \cdot FR_L \cdot m_{IFS}}$$

where d_{pixel} is the pixel size and m_{IFS} represents the optical magnification between collimator and camera optics for a TIGRE-type concept design.

Starting from the relation above and according to the Shannon theorem, we can define the aliasing budget of the monochromatic slit sampling as follows:

$$\Delta_{ALIASING} \equiv \frac{|2 - B_{SLIT}^{-1}|}{2} \cdot 100$$

The aliasing budget is a key parameter quantifying the capacity of an IFS concept to remove the speckle chromatism exploiting the correct sampling of the monochromatic signal. However, in normal TIGRE $\Delta_{ALIASING} > 0$.

11.2.2 IFS Optical Design

In this Section we describe a schematic design of the TIGRE IFS matching the 3D-Spectroscopy TLRs (see Sec. 11.2.1) and the requirement for the cross talk upper limit (see Sec. 11.2.1.5).

The cross talk upper limit fixes the best lens-array layout. It turns that the best layout, i.e. that one allowing to match the cross talk upper limit requirement in both directions (spatial and spectral on the detector/s plane) is the hexagonal case-B (see Figure 11-8). We notice here that this same configuration turned out to be optimal also for the VLT-PF TIGRE IFS concept study.
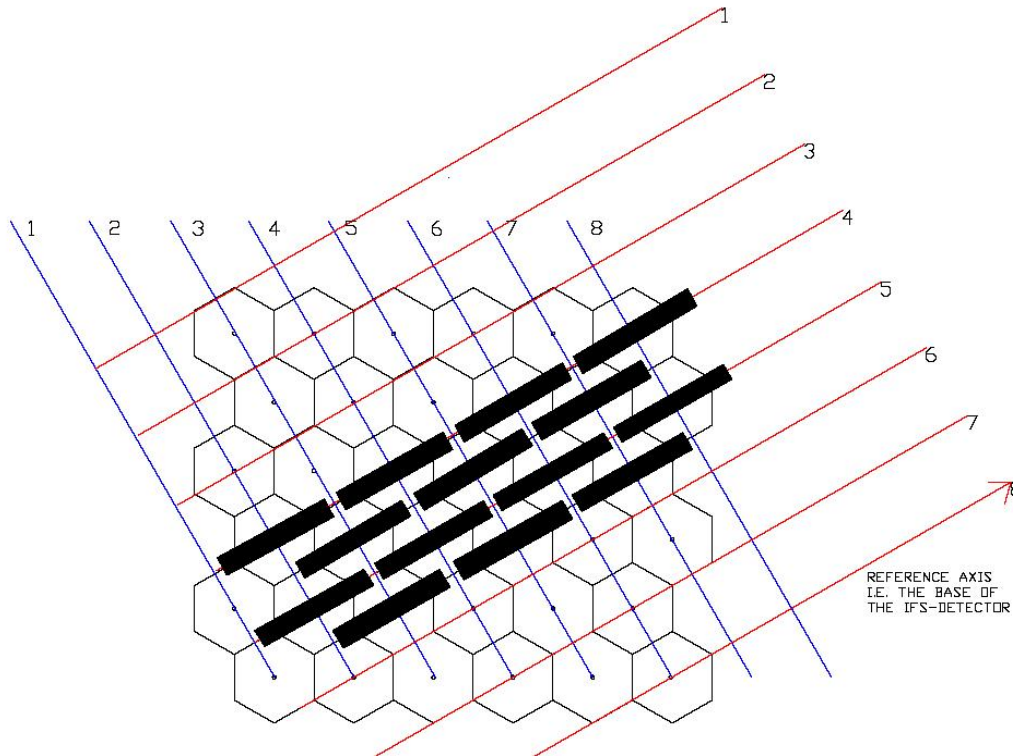
Figure 11-8: Thin hexagons represent the lens-array. Black filled rectangles represent the spectra on the detector/s plane corresponding to the hexagonal case-B layout. Oblique rows and columns represent a reference matrix orientated in the same way as the IFS detector/s.

The spectrum length and cross talk upper limit together with the sub-system TLRs and optical specifications (detector/s pixel size, magnification factor on the detector/s plane) set the TIGRE IFS main parameters (see Table 11-2).

As shown in Table 11-3, this TIGRE layout solution verifies the 3D-Spec. sub-system TLRs and the cross talk upper limit (in the case of a diffraction limited - i.e. free from static aberrations - IFS design which is a reasonable assumption in this context); notice that this latter also determines the maximum spectrum length allowed in this solution.

The aliasing budget is large: the exit slit FWHM is undersampled (0.31 pixel @ λ_{min} and 0.39 @ λ_{max}), producing an aliasing budget of order of (84/80 %) at the minimum/ maximum wavelength respectively. In order to avoid the aliasing but saving the TIGRE IFS concept, one might introduce a magnification factor (m_{IFS}) equal to 6.6; however this Super-TIGRE layout solution would require an enormous detector size (~ 176 detectors [4k x 4k]).

The layout solution described in Table 11-2 allows to conceive a design for the TIGRE IFS optics following the same strategy of the TIGRE IFS proposed in the CHEOPS phase-A study for the VLT-PF (cfr. CHEOPS-TRE-OPD-00038), even if with opto-mechanical variations due to the strong difference between the size of the lens-array needed for EPICS and VLT-PF respectively (1000 x 1000 vs 254 x 254, cfr. CHEOPS-SPE-OPD-00023).

Table 11-4 provides approximate values for the sizes of the optics needed.

TIGRE IFS main parameters	TIGRE IFS other parameters
Fore optics focal ratio = F/217	Lens-array refraction index = 1.5
Lens-array layout = hexagolal-B	Lenses-mask diam. = 0.95*Lens-array pitch
Lens-array pitch = 150 microns	Lens-array focal length = 0.6 mm
Lens-array focal ratio = F/4	Lens-array curvature radius = 300 microns
Lenses number = 1000 x 1000	Lens-array sag = 9.5 microns
Wavelength dispersion = 0.05 micron/pixel	Detector/s pixel size = 18 microns
Spectrum length = 10 pixels	IFS optical magnification = 1
Spectrum area = 60 pixels	Exit slit FWHM (@ λ_{min}) = 0.31 pixels
Detector size = 4 [4k x 4k units]	Exit slit FWHM (@ λ_{min}) = 0.39 pixels

Table 11-2: The TIGRE IFS layout solution for EPICS.

3D-SPEC. TLRs	TIGRE IFS layout solution
Min. wavelength = 1.38 microns	1.38 microns
Max. wavelength = 1.75 microns	1.75 microns
Wavelength purity ≥ 15 (2 pixels)	21 (2 pixels)
Field of View = 2 arcsecs [diagonal]	2 arcsec [diagonal]
Cross talk upper limit < 7.5 x10^{-5}	7.0 x 10^{-5}

Table 11-3: Confront among the 3D-Spectroscopy TLRs and the values obtained in the final layout.

1-QUADRANT TIGRE IFS	4-QUADRANT TIGRE IFS
Diagonal size of lens-array ≤ 213 mm	Diagonal size of lens-array ≤ 106.5 mm
Collimator focal length before opt. = 600 mm	Collimator focal length before opt. = 600 mm
Max.size of the collimator optics ≤ 380 mm	Max.size of the collimator optics ≤ 274 mm
Max.size of the collimator optics ≤ 380 mm	Max.size of the collimator optics ≤ 274 mm
Collimator corrected field ~20 deg.	Collimator corrected field ~10 deg
Camera corrected field ~ 20 deg	Camera corrected field ~ 10 deg
IFS length (from the entrance slit)= 2400 mm	IFS length (from the entrance slit)= 2400 mm
Number of collimator+camera arms = 1	Number of collimator+camera arms = 4
Single detector size = 4 [4k x 4k units]	Single detector size = 1 [4k x 4k units]

Table 11-4: Global sketch of the TIGRE IFS optics. Two different opto-mechanical configurations are resumed one with no-division/4Q-division of the FOV respectively.

11.2.3 Mechanical Design

In the figures of this Section (

Figure 11-9 and Figure 11-10) we show a possible arrangement for the EPICS 4Q-TIGRE IFS.

Figure 11-9: 3D-sketch of EPICS 4Q-TIGRE IFS showing the four arms. The light from the AO/Coronagraphic-module comes from bottom. Each arm is fed by a pyramid-mirror. Each arm includes: collimator, an Amici prism disperser, a camera, and a 4k x 4k detector. The detectors should be within cryostats not shown in this figure. The length of each arm is ~ 2.4 meters. Suitable folding-flats could be inserted into each arm to modify the geometry of the IFS.

Figure 11-10: Side and top views of the EPICS 4Q-TIGRE IFS.

11.2.4 IFS Operation

11.2.4.1 Observing Procedure

In order to optimize the detection rates a detailed understanding of the instrument and in particular the detector is required. Typically total long integration times (from 1 to a few hours) are needed to achieve the required contrast for the planet detection. Therefore we will need to combine a large number of individual exposures. Individual exposures times will be adjusted such that the optimal compromise between field rotation, flat field noise reduction through dithering, detector saturation and good sampling of the atmospheric changes is achieved. The availability of an automatic and close to real time data-reduction system at the telescope is of paramount importance for the operation of the IFS.

11.2.4.2 Instrument Calibration

Various calibrations are required for the IFS. These include:

Basic detector characteristics: Dark current, gain, etc.

Fixed pattern noise: (a) Channel-to-channel transmission variations in the IFS. This may be due to variations of optical performances (transmittance, focal length, optical quality) of the lenslet array/stray light suppression mask system, and it is expected to be at the percent level. It is expected to be constant with time. (b) Wavelength dependent pixel-to-pixel efficiency variations on the detector. These are expected to be quite large and to be time dependent. Fixed pattern noise is very dangerous, since it may create artifacts similar to planet images. Reduction of fixed pattern noise will be achieved by two strategies: (1) dithering of images by exploiting field rotation or X-Y translations of the detector/s. (2) detailed understanding of IFS and detector flat-fields. In order to allow for an efficient on-line data reduction, the flat-fields will be stored in reference files; for example the detector flat-field can be stored in a cube where the wavelength dependent pattern comprises the third axis.

Wavelength calibration: This is required for all individual IFS spectra in order to provide an accurate reconstruction of monochromatic images from the combination of the spectra. Wavelength calibration will be achieved by means of images obtained through simultaneous homogenous illumination of the IFS by a few selected monochromatic sources.

Spectral trace information: A very accurate knowledge of the spectral position of each input lenslet is required. The software will need to be able to treat very accurately sampling effects and thus will need to rely on instrument modeling of the spectral PSF. Furthermore a regular monitoring of the cross talk levels is extremely important.

Outline of a typical image reconstruction:
- acquisition of an individual image

- bias and dark subtraction
- removal of cosmic ray hits
- removal of the detector fixed pattern (using a detector flat field)
- extraction of individual spectra (using a previously determined map of the positions of the spectra)
- possibly iterative removal of the cross talk
- wavelength calibration of the individual spectra
- rebinning to constant wavelength step
- extraction of monochromatic images
- removal of the IFS fixed pattern (using an IFS flat field)
- examination of image quality; if not passed image is rejected, else then:
- if required, reentering of the monochromatic images to a common reference system
- rotation of the monochromatic images to the common reference system
- if required, scaling of the monochromatic images to a common reference system
- sum of the monochromatic images
- scaling of diffraction pattern as a function of wavelength for appropriate speckle subtraction

computations of differences/ratios of the summed monochromatic images.

11.2.5 Results of simulations

Some simulations have been conducted based on the TIGER IFS layout proposed in Sec. 11.2.2. The aim of this exercise was not to simulate a detection of an extrasolar planet (and more specifically an earth-like one) but to test the capability of the TIGER IFS simulator adopted in the present VLT-PF study (and based on the CHEOPS IFS simulator, cfr. CHEOPS-TRE-OPD-00034) to be modified and fine-tuned to the OWL-EPICS science case.

The two figures below show a different scales the pattern of a star on the detector/s plane (no division of the FOV has been requested in this simulation and no coronagraph has been inserted in the optical path of the star light insofar).

Figure 11-11: Image of the TIGRE IFS spectra of the central star as obtained with the proposed layout covering a 1k x 1k portion of the detector. Intensities are in log-scale.

Figure 11-12: Image of the TIGRE IFS spectra of the central star as obtained with the proposed layout covering a 0.5k x 0.5k portion of the detector. Intensities are in log-scale.

11.2.6 TIGRE IFS Mid-term development plan

The most critical components of the proposed TIGRE IFS design (microlens array, optical components, detector) are based on currently available technology. Prototypes with characteristics similar to those required for the EPICS IFS will be developed within the VLT-PF IFS.

11.3 FT IFS Implementation Concept

This Section describes an FTS concept for extreme-AO planet finding instruments. Contributors: Etienne le Coarer, Kjetil Dohlen, Almas Chalabaev, Patrick Rabou.

11.3.1 General discussion and modes of operation

The problem of detection of extra-solar planets is mainly one of eliminating the speckles created by the passage of the coherent stellar light through the atmosphere and the instrument optics. Highly performing AO systems minimize wave-front errors, and coronagraphs eliminate pupil diffraction. Still, planets will be buried deep below a residual speckle pattern, and so the first role of any planet imager is to characterize the speckle halo in order to eliminate it. Dual or multiple band differential imaging is one approach. 3D spectroscopy is another, related approach [43], which, in theory at least, would appear superior because it multiplies the number of measurements of the speckle pattern, hence potentially increasing the precision of its elimination.

In practice, a number of practical problems arise in trying to measure the 3D spectral cube of speckles, as was pointed out during the VLT-PF Phase A study. As can be seen in the slice through a cube shown in Figure 11-2 (a), each speckle runs through the cube obliquely, a bit like a fan of spaghetti. The spectrum of a given pixel, found by extracting a column as in Figure 11-2 (b), is therefore itself speckled. This spectral speckle has a period which depends upon the distance from the centre of the image; its resolution is approximately equal to the distance expressed in λ/D units. For a spectrograph of resolving power R, the spectral speckles will therefore tend to be unresolved for FOV>$2R\lambda/D$, beyond which spectral aliasing will occur unless the optical transfer function of the spectrograph eliminates all spectral features with higher resolution. Eliminating aliasing noise is a critical and difficult problem in Tigre-type systems.

In FTS systems, since spectral information is obtained via an interferogram, spectral aliasing cannot occur. The finite length of the interferogram sets a precise limit to the spectral resolving power, acting as an anti-aliasing filter, and no higher resolution components of the spectrum can be detected. Also, this filtering effect is expected to eliminate speckle to a great extent for FOV>$2R\lambda/D$.

Still, the effects of temporal variations in speckles may be important. While the single-output FTS (Michelson interferometer) is known to be very sensitive to source intensity variations, this sensitivity is efficiently eliminated in the dual-output version (Mach-Zehnder interferometer) since the sum of the two outputs is equal to the instantaneous source flux. Speckle variations, however, induce random variations in the source spectrum, and this cannot be calibrated photometrically. To some extent, the above-mentioned filtering process will reduce the time-variable part of the speckles. Further improvements are expected by application of various algorithms such as blind deconvolution. Further study of this and other options must be studied.

But the FTS may be used in different modes where speckle variations can possibly be ignored or even used as an observational advantage. We describe here briefly two such modes, stressing that again, deep further studies are required to determine the validity and real value of such modes.

11.3.1.1 Harmonic features

Certain spectral features in atmospheric spectra consist of a forest of fine spectral lines, more or less evenly spaced. The O2 band at 760nm, Figure 11-13 (a), is a good example of such a feature. A perfectly harmonic spectral feature will give rise to a strong signal in the interferogram occurring at an OPD corresponding to the period of the feature, Figure 11-13 (b). It should therefore be possible to study the characteristics of this feature simply by scanning a range of OPDs corresponding to that signal. While not providing a high-resolution spectrum in the classical sense, this method could give access to information about such high-resolution features in a very efficient way. This method is similar to the concept of a Fourier transform stellar seismograph proposed by Maillard (1996)[44].
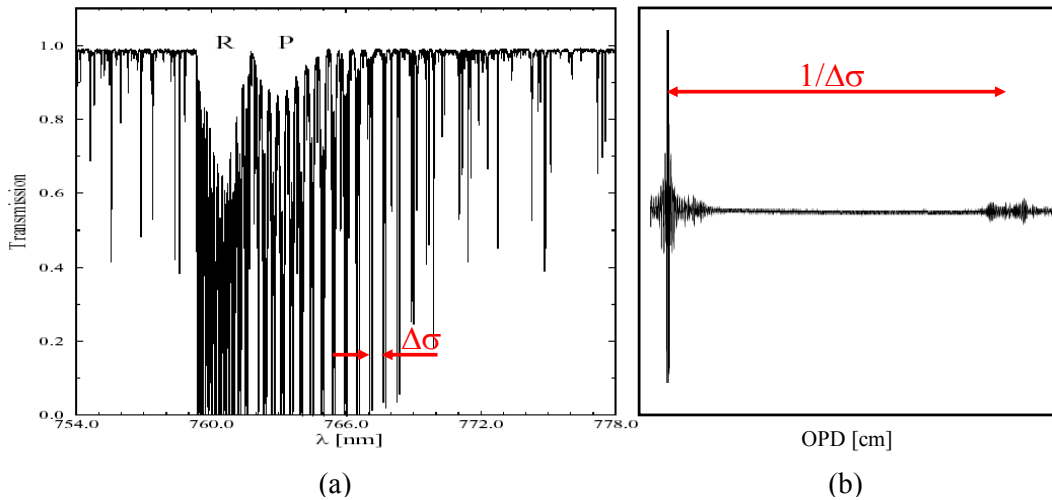
Figure 11-13. The O_2 A-band at high spectral resolution (a) indicating the period of a quasi-harmonic structure in the spectrum. An indication of what the interferogram of this spectrum would look like is shown in (b), showing the presence of strong features at an OPD corresponding to the harmonic period of the spectrum.

The problems of speckle noise will in this case be reduced if the frequency of the harmonic feature is beyond that of the spectral speckle. Since, for a 4k x 4k detector, the FOV extends to a radius of 1000 λ/D, this mode is therefore expected to be useful for features whose resolution is R>1000.

Difficulties include the fact that our own atmosphere, containing much the same molecular constituents as those we seek to discover in exo-atmospheres, will also create an interferogram signal at the same OPD. The interaction between these signatures and their disentangling, which should be possible thanks to the finite redshift of the observed exo-worlds, again needs further study.

11.3.1.2 <u>Speckle demodulation</u>

If the optical bandwidth ($\Delta\lambda$) is reduced sufficiently to be smaller than the spectral speckle period, then the temporal speckle variation can be considered as purely photometric. This is true within a limited field of view of FOV < 2 $\lambda/\Delta\lambda$ λ/D. ie, FOV < 100λ/D for $\lambda/\Delta\lambda$ = 50. The speckle noise is therefore efficiently eliminated by the photometric interferogram correction provided the two outputs (I_1 and I_2) of a Mach-Zehnder interferometer are used:

$$I_{Corr} = (I_1 - I_2)/(I_1 + I_2).$$

After Fourier transformation of I_{Corr}, a high-fidelity, non-speckled but small-bandwidth spectrum of the star will be found in every pixel of the detector. Creating a fiducial stellar spectrum by taking the average of all the spectra and subtracting the average from every spectrum should yield the planet spectrum.

However, the light from the planet is not affected by the apparent photometric variation for which the stellar interferogram was corrected. This correction has therefore corrupted the planetary spectrum, spreading it as noise across the entire, mostly empty, spectral range. After subtraction of the fiducial stellar spectrum, the residual must therefore be transformed back into the interferogram domain where the photometric correction is undone (demodulated) by re-multiplication with the factor ($I_1 + I_2$). Transforming the demodulated interferogram into the spectral domain again finally recovers the planetary spectrum. If the band is centered on a sharp spectral feature of the planet spectrum, for example a low-resolution version of the O2 forest, detection of this feature will help identifying the planet.

Note that this detection mode will have an important noise advantage because of its reduced bandwidth, hence low spectral fill factor. Residuals of the stellar signal will be spread as noise into the unoccupied parts of the spectrum and hence eliminated, but so will also most of the stellar photon noise.

It would therefore seem that this mode, if confirmed by further studies, could be an important tool for planet search in the inner regions of the stellar halo.

11.3.2 Optical implementation

This section describes the optical implementation of IRFTS for Planet Finder. The size of the telescope has no impact on the size of the instrument because the instrument is located after an adaptive optics which permits instrument to sample the diffraction limit of images. For a 100m diameter telescope, a 4kx4k detector covers a 4"x4" FOV as required. A Rockwell Hawaii detector covering the spectral range from 0.8 to 2.5 micron with a quantum efficiency exceeding 50% is foreseen. It is possible to have a beam splitter which covers this spectral range with the same coating.

Figure 11-14: Proposed optical design for the EPICS FT-IFS. The blue arrow at the right indicates the incoming collimated beam (D=10 mm) from the PF adaptive optics. A Lyot mask is placed at the entrance of the spectrometer. After the first wave-front division by the beam splitter, the light goes through the cat's eyes which image the pupil again just after the second part of the beam splitter and separates the incoming and outgoing beams vertically. A beamsplitter of double length is used for the entrance and the output beams. The camera optics is an air-spaced doublet, as proposed for the IRDIS dual imaging camera for VLT-PF. Adding a pupil imaging capability may be useful for image quality calibration purposes.

11.3.3 Mechanical implementation

Figure 11-15 illustrates the size of the instrument, estimated to match a 50cmx50cmx100cm cryogenic box. Moveable parts include the OPD scanning mechanism, an adjustable tip/tilt in one of the arms to tune the interferometer alignment, and filter wheels.

The limitation of scanning range to a maximum of ~10mm is mainly set by the wish to avoid use of adjustable curvature mirrors in the cat's eyes. The scanning mechanism must ensure parallel displacement with precision of 0.1 micron. Such mechanisms have been developed both for ground-based applications (CFHT) and cryogenic space operation (IASI, SPIRE, see Figure 11-16).

Dispersive Integral Field
Spectrograph

Integral Field Fourier
Transform Spectrograph

Figure 11-15 :The proposed Fourier Transform Spectrometer (small red box) compared with the VLT-PF instrument installed at the Nasmyth platform (left) and a Slicer spectrograph design under consideration for VLT-PF (upper right).

(a)	(b)

Figure 11-16 :Existing scan mechanisms for two cryogenic space-borne FTS systems, IASI on METOP (a) and SPIRE on Herschel. The former is built by CSEM (Suisse) for a 20mm stroke, the latter by LAM for 30mm stroke.

11.3.4 FT-IFS predicted performance

We studied the FT-IFS option in terms of a comparison with integral field dispersive spectrographs.

Dispersive spectrographs coupled to mosaic detectors have the advantage of recording the field simultaneously in all used spectral channels. While costly in terms of detector real estate, this is advantageous in terms of signal-to-noise ratio, in particular for high resolving powers and low detector noise. However, in the case of EPICS, this advantage is reduced by the need for additional optics which can introduce a loss of contrast and transmission.

The Fourier Transform Spectrometer (FTS hereafter) turns out to fit the EPICS specifications on the contrast, Strehl ratio and the optical path differences between colors in a satisfying way. Indeed, the FTS makes use of only simple optics: beam splitter and mirrors; no dispersive elements, only little loss in transmission and contrast. Furthermore, since the EPICS specifications require a large sky field (~4arcsec) and a spectral resolution >1000 for some science cases, the gain in size of the optics of the FTS compared with that of any dispersive option is enormous: cat's eyes and the beam splitter in the FTS do not exceed 50 mm (see Figure 11-14). The optics of such a size can be ordered with the specifications of surface precision down to few nm rms. The maximum path difference to be achieved for a spectral resolutions up to 100 is 0.2 mm (at $\lambda = 2$ µm), and the mechanical travel is only 0.1 mm. With no additional investment, commercially available high precision equipment can provide mechanical travel up to 1 mm even in cooled environment, i.e. spectral resolution up to 1000. This is achieved without laser metrology, using only magneto- or capacitive sensors providing 10nm precision. Finally, thanks to its compactness, the FTS can be implemented rather easily.

We shall consider in more details other issues of the FTS such as the signal-to-noise ratio, the color-to-color differential aberrations, and so on. We will summarize the analysis before giving the designed implementation in the PF.

11.3.4.1 Signal-to-noise ratio calculations

A detailed analysis of this important and subtle issue is required. Here, we will present only main lines leading to conservative estimates of the FTS performance relative to other possible techniques.

Let us write the general expression of the signal-to-noise ratio S/N in the case of the spectrographs. Let F_λ denote the spectral flux density of the target as measured by the detector pixel at the output. Then we can write:

$$\frac{S}{N} = \frac{\eta_{det} \cdot T_F \cdot F_\lambda \cdot t}{\left[\delta_{det}^2 + \delta_{phot}^2 + \delta_{resp}^2 \right]^{1/2}}$$

where η_{det} is the detector QE, T_F is the optical transmission, t is the integration time, and δ is the standard deviation of the noise fluctuations respectively for the detector read-out noise, the photon noise and the pixel response fluctuations ("flat-field" noise). To develop this expression further, three instrument concepts are considered :

FTS : (Fourier Transform Spectrograph) which is sufficiently small to be entirely cooled

IFSC : (Integral Field Spectrograph Cold) an entirely cooled Tigre spectrograph, requiring a 2 cubic meter vacuum vessel

IFSW : (Integral Field Spectrograph Warm) Tigre spectrograph, only the detector is cooled.

The development of each noise term is pursued in Sec. 11.5

11.3.4.2 <u>Signal to noise simulation for IR</u>

A simulation tool has been developed to include above formalism and to estimate the relative performance of IFSC, IFSW, IFTS.

We have considered (see the inputs parameters panel of Figure 11-17):

- infrared detectors with 15 electrons readout noise (actually Rockwell guarantee 18electrons).

- a few seconds exposures ,

- the sky/instrument background 1 ph/sec

- 1% of uncertainty on FF, these simulation show that this factor is negligible

- transmission of IFS is 70% which is an optimistic value. OASIS, SAURON have 25% including detector and filters and gratings rather than prisms.

- transmission of FTS is 90% including unbalance of beam splitter this numbers include only parts which are not shared by IFS and FTS :

Results of the simulations are shown in Figure 11-18. The signal to noise is plotted versus spectral resolution (number of channels) in order to define the range where FTS concept are better or worse than IFS.

Global slope of FTS is greater than the IFS, which is consequence of the intrinsic role of noise in FTS, FTS noise is sensitive to number of spectral channels.

The FTS has an advantage over the cooled Tigre up to R~15, and over the warm Tigre up to R~75 due to better transmission and less sensitivity to detector noise. The cold grating spectrograph is seen to be an ideal spectrograph for higher resolving powers. The difference between warm and cold IFS below K band is essentially due to a loss of efficiency of the observation strategy dark-sky/object.

Note that the reduced FTS performance at high resolution can be regained by reducing the transmitted bandwidth so that the number of spectral channels (k) is much smaller than R.

Note also that the IFS has been considered with prism dispersive unit to justify the 70% transmission. If we push the dispersion of Tigre beyond R=100 resolution, gratings must be used, reducing the transmission to typically 50%. This assumption permits FTS to be competitive still R=100-200.

Figure 11-17: Simulation tool panel. The exposure time is hidden by the Background level or the number of electron of the signal.
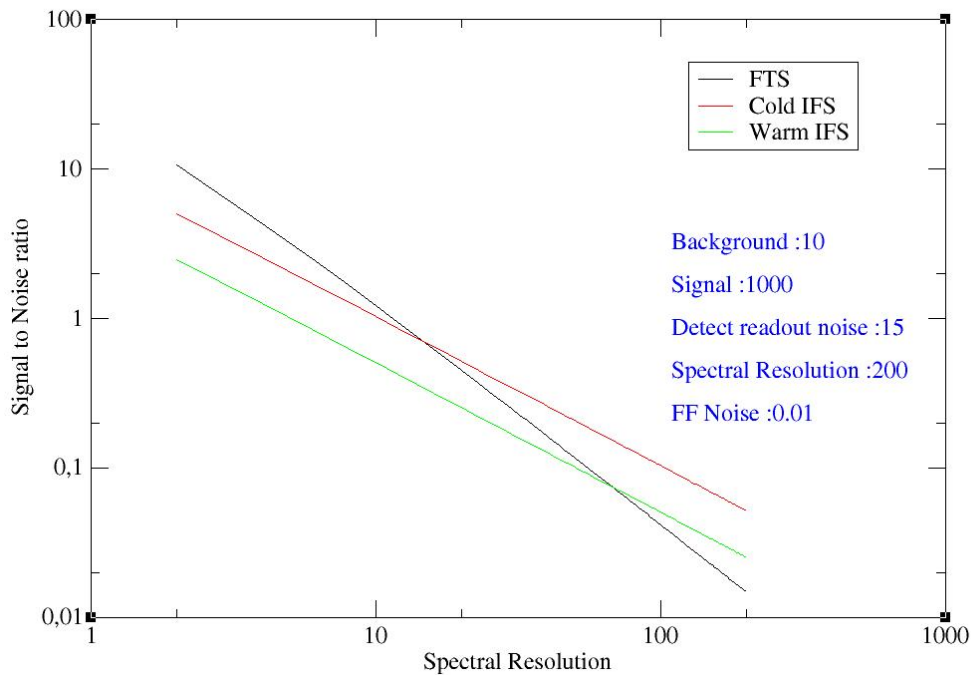
Figure 11-18: Plot of signal to noise ratio variation versus resolution.

11.3.4.3 Signal to noise simulation for Visible

We have considered :

- Visible detectors have 2 electrons readout noise.
- a few seconds exposures ,

- the sky/instrument background 0 ph/sec

- 0% of uncertainty on FF

- transmission of prism IFS is 70%

- Transmission if FTS is 90% including unbalance of beamsplitter this numbers include only parts which are not shared by IFS and FTS :

In the visible there is no clear advantage for FTS in terms of signal to noise except to have access to a large field of view. Other advantages remain and should be discussed.
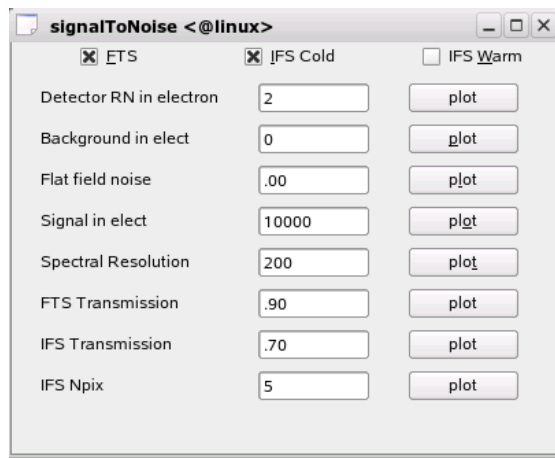
Figure 11-19: Signal to noise calculator panel

Figure 11-20: Result for signal-to-noise simulations in the visible.

11.3.4.4 <u>Speckle noise</u>

Speckles are moving with three different time scales, a) faster than *t*, the elementary frame time; b) between *t* and *T*, the total exposure time, and c) longer than *T*.

Case b) is clearly the worst for FTS. The FTS measures a sequence of white images which can give some additional information to discriminate between speckle and planet and the exploitation of this and other specific features must be carefully studied. An important analysis and simulation effort has therefore to be invested to make this study complete by the description of specific data reduction procedure and optimal performance.

11.3.4.5 <u>Contrast</u>

- There is no grating and therefore no diffuse light added in the FTS, so that the contrast is preserved.
- Full-pupil optics of the FTS avoids the diffraction and the diffusion due to pupil fragmentation.

11.3.4.6 <u>Color-to-color optical path differences</u>

In the FTS, the light of all wavelengths passes through the same optics. Misbalance may still appear between the two outputs, which of course has to be reduced to a minimum, but this misbalance is achromatic.

11.3.4.7 <u>Polarization</u>

An alternative FTS design, based on the Martin-Puplett (Martin 1982) concept can introduce a polarization separation which permits the FTS to observe simultaneously the two polarizations.

11.3.4.8 <u>Miscellaneous</u>

- The 1/f noise which can strongly disadvantage the FTS due to the necessity of several read-outs to produce the spectrum is reduced by the technique of partial detector shielding successfully used in the AMBER instrument.
- All interferogram values for a given pixel are recorded by same detector pixel in each output, reducing the "flat-field" noise.
- The concept opens the possibility to implement a Shearing Interferometer (to be studied) which can further improve the performance of the instrument close to the star [46].
- Spectral resolution flexibility: The optimum resolution is chosen without changing the optics.

- Simplicity and compactness of the opto-mechanical setup, easy implementation.

11.3.5 FT IFS Mid-term development plan

As we have seen in the preceding Sections, the FT-IFS option for EPICS is, in spite of the maturity of the general FTS concept, immature and requires large R&D efforts. Still, we have tried to put forward the advantages of the option, mainly in terms of opto-mechanical and cost considerations, but also in terms of signal-to-noise performance. We have also indicated avenues of investigation in terms of operational modes, leading - we are quite convinced - to a viable and extremely interesting instrumental option for planet search and characterization.

A major obstacle in terms of development plan for this instrument, most certainly shared by many other aspects of EPICS as well as the entire ELT concept, is the need for continued study in parallel with the development of current-generation instruments, VLT-PF in particular. In the near term, most of the competence in terms of staff researchers and technicians is therefore available only at reduced levels of implication. This does not appear to be a major problem for EPICS, since, in any realistic scenario, an XAO capability for an ELT will not be available before several years after the telescope's first light. We therefore propose to delay any serious instrument development activities until VLT-PF is commissioned.

Still, issues related to fundamental aspects of the FT-IFS concept should ideally be addressed in the intervening time, possibly through pre-doc or post-doc projects. We propose here four such subjects, the exact definition of which could vary according to the person or institute which takes on the work:

- Theoretical work/simulation: speckle noise reduction by blind deconvolution-type algorithms, etc
- Theoretical work/simulation: harmonic features, speckle demodulation, etc
- Performance simulation, end-to-end modeling
- Lab work, setting up a demonstrator FTS and realistic speckle simulation

The French FTS community is spread over several labs, including LAOG (Grenoble); CRAL (Lyon); OAMP-LAM (Marseille); Onera, IAP, IAS (Paris); etc. Although representatives of only a small part of this community have participated in the present report, several others have expressed interest in the subject, and they should be included in the discussions concerning a possible future consortium. But the FTS world does not stop at the French border, and there is of course no reason not to include other European colleagues in the design and construction of such an important and promising instrumental concept.

11.4 Conclusions, summary of IFS options

In this concept study we have put forward two very different instrumental options for the EPICS Integral Field Spectrograph, a Tigre-type option and a Fourier transform option. Both concepts are classical concepts developed since many years for various branches of science, including astronomy. However, as illustrated by the discussions going on within the world-wide planet imaging community and in the VLT-PF community in particular, the specific problems related to this subject pose completely new questions, like diffraction induced slit effects, spectral aliasing, speckle-induced spectro-photometric source variations, etc. Also, specifications in terms of inter-channel cross talk, flat-field precision and stability, etc turn out to be extremely severe. New ways of thinking about and characterizing these instruments must therefore be invented, and this document only pretends to give embryonic answers to these questions.

Much further work is therefore required. It is probable that the VLT-PF instrument, currently in its definition phase, will include a Tigre-type instrument. This is of course an enormous advantage for the evolution of its EPICS counterpart, since most of the questions posed are the same in each case. In order for the FTS to keep up the pace, it is therefore necessary to invite the community to initiate fundamental research projects in order to address at least some of the issues that we have raised.

The FTS option has not been selected for VLT-PF, for which the Tiger-type option appears to fit more snugly to the specific requirements in terms of FOV and spectral resolution without raising many of the questions posed in terms of temporal scanning of the FTS. It appears clear, however, that the increased spatial resolution offered by an ELT and the enlarged spectral range considered for the EPICS science case (earth-like planets) point towards the improved flexibility and economic use of detector pixels offered by an FTS. Also, the present report indicates that some aspects specific to an ELT, where the distance to observed objects in λ/D units is up to 10 times larger than in a VLT, can efficiently reduce its sensitivity to temporally varying speckle noise.

On the other hand, the Tigre Section of this document indicates a design avenue satisfying most of the critical issues involved, in particular the pixel-to-pixel cross talk. Further work is required to answer certain outstanding questions related to diffraction effects and spectral aliasing, but these questions are already addressed within the VLT-PF study, and design options overcoming them (Super-Tigre, Bigre, ...) are being studied. The realized FOV and spectral coverage, including many scientifically important spectral features, are acceptable while using a reasonable amount of detector real estate.

Eventually, the advantages of both options must be considered in the framework of finite coronagraph and detector performances, and compared in terms of final efficiency in terms of planet detection and characterization capacity. We are clearly not ready to perform a concept selection at this point, but, in view of any realistic development scenario for an ELT and its associated XAO system, this does not appear to be a pressing issue.

Finally, we stress that although much future work is clearly required in terms of fundamental issues for both concepts, these issues are related to the specific problems related to extremely high dynamic range imaging. No technological show stoppers have been identified; these instruments could essentially be built tomorrow. We also indicate that the problems involved, related essentially to elimination of static and quasi-static aberrations,

are not specific to ground-based systems, but also shared with any similar space-based instrument.

11.5 Appendix: Development of noise terms

This Appendix develops the formalism upon which the signal-to-noise comparison between FTS and dispersive spectrographs referred to in Sec. 11.3.4.1 is based.

For the following tables, we consider that the resolution R and the number of FTS channels k are the same value.

Detector readout	δ^2_{det}	Comments
FTS	$2k \cdot Ron^2$	2: simple approach to include that a step is a mixture of two detector subarray reads k: for each wave slice, the readout noise of whole scan is included R_{on}: standard readout noise deviation
IFSW	$2 \cdot 2 \cdot Npix \cdot Ron^2$	2: the final spectral element is the result of subtraction of dark measurement and object measurement N_{pix} is the distance between two spectra, extraction of spectral channel is the measurement of profile surrounding dark parts separating two spectra 2: this extraction use two times spectra separation
IFSC	$2 \cdot 2 \cdot Npix \cdot Ron^2$	id

Table 11-5 detector noise

Photon Noise	δ^2_{phot}	Comments
FTS	$\eta_{\det} \cdot T_{FTS}(F_\lambda + 2F_B) \cdot t$	T_{FTS} : transmission of FTS $\eta_{\det}$: detector efficiency F_λ: object flux F_B: background including detector dark . For wavelength between 1.5μm and 1.8μm the background is 1 photon per second. It is twice because the FTS has two entries, the second one can be set on sky then the background is automatically subtracted but the photon noise remains t: whole exposure time
IFSW	$\eta_{\det} t(T_{IFSW}(F_\lambda + 2F_B) / R + N_{PIX} \cdot 2 \cdot F_B)$	T_{IFSW} : transmission of IFSW $\eta_{\det}$: detector efficiency F_λ: object flux N_{pix}: interspectra distance, in warm instrument , whole detector receive background contribution, it is twice because flux is extracted according the two side of spectra. this background comes from instrument parts close to dispersive units, it illuminates directly detector and don't depend of instrument transmission R: Number of channels the signal coming from sky is dispersed by grating, in warm instrument the instrument background is not dispersed then not depend of number of channels F_B: Flux spectral density of the background including detector dark . For wavelength between 1.5 and 1.8 the background is 1photon per second. It is twice because of dark exposure. t: exposure time
IFSC	$\eta_{\det} \cdot T_{IFSC}(F_\lambda + 2F_B) / R \cdot t$	T_{IFSC} : transmission of IFSC $\eta_{\det}$: detector efficiency F_λ: object flux R: Number of channels the signal coming from sky is dispersed by prism Fb: Flux spectral density of the background including detector dark . For wavelength between 1.5 and 1.8 the background is 1photon per second. It is twice because of dark exposure. t: exposure time

Table 11-6 photon noise

Responsivity Noise	δ^2_{resp}	Comments
FTS	$(F_\lambda + F_B) \cdot FF_{\det} \cdot \sqrt{2}$	an error on the flat-field determination can be considered, but it is applied on the signal + background as a fraction of it. This uncertainty is close to percent and is negligible in front of photon noise itself here only two pixels are considered in the flux evaluation,
IFSW	$(F_\lambda + F_B) \cdot FF_{\det} \cdot N_{PIX}$	Id
IFSC	$(F_\lambda + F_B) \cdot FF_{\det} \cdot N_{PIX}$	Id

Table 11-7 Response noise. There is not a dominating source of noise, this translates into an advantage for the FTS since the reconstruction of the spectrum is obtained using measures of temporal evolution of flux within one pixel, while the reconstruction of the spectrum in a dispersive system requires the precise knowledge of the wavelength response of each pixel. This factor is a delicate factor for infrared arrays.

Signal	Detected signal Signal	Comments
FTS	$F_\lambda \cdot \eta_{\det} \cdot T_{FTS}$	T_{FTS} : transmission of FTS includes also the unbalance introduced in beam separator ; this transmission can be estimated to 0.9
IFSW	$F_\lambda \cdot \eta_{\det} \cdot T_{IFSW} / 2$	T_{IFSW} : transmission of warm IFS which includes also some masks along the path to limit the diffusion introduced by micro lenses, dispersive units ; it is estimated to 70% 2: because of dark/sky exposure, the field of view is small, it is not possible to simply chop object in two opposite corners
IFSC	$F_\lambda \cdot \eta_{\det} \cdot T_{IFSC}$	T_{IFSW} : transmission of warm IFS which includes also some masks along the path to limit the diffusion introduced by micro lenses, dispersive units ; it is estimated to 70% A cooled environment avoid the Sky/Object exposure

Table 11-8 - Detected Signal

257

12 Differential polarimetry

12.1 Scope

This section describes a possible concept for a polarimetric observational mode of EPICS, the planetary camera-spectrograph for OWL. Basic instrumental techniques are discussed in [36].

12.1.1 Why polarimetry?

Light reflected from a planet is polarized, mainly due to Rayleigh scattering (see section 5). This raises the possibility of detecting extra-solar planets by using its polarization signal. For this reason a polarimetric imaging mode is a desirable technique for the EPICS project.

Even with very good adaptive optics, success in detecting Earth-like planets will depend on distinguishing the signal of a faint planet in the presence of a very much stronger background (residual seeing disk). Even more daunting than mere detection is the requirement for characterizing the planets one has detected. Simple photometric imaging fails by several powers of 10 and mere improvement of photometric techniques is unlikely to solve this problem. We need a differential technique which discriminates against starlight and favors planetary light.

Imaging polarimetry is such a technique. Since planets are seen (in the visible) by scattered light only, their light will be polarized to some extent. The expected polarization is 5-50% and covers a broad range of terrestrial and Jupiter-like planets, as shown by models and by observations within our own solar system (see section 5). The parent star (if closer than 30-50 pc and of not too extreme a spectral type, see ref. [11]) generally has a polarization of less than about 10^{-4}, so that the task of detecting the polarized planet against the virtually unpolarized stellar residual seeing disk is made 500 to 5000 times easier by doing it via polarimetry rather than photometry. Help is also provided by the fact that the planet is a point source (the residual seeing disk is, when averaged over time, extended; the speckles constitute – mainly unpolarized – noise), which moreover may be expected to move around the central star (as opposed to a faint background star with interstellar polarization). And finally, the "degree-of-polarization spectrum" of planets differs from that of the parent star. All this means that one can use the light-gathering power of ELTs and long integration times, as long as one can suppress sufficiently the systematic errors of the polarimetric process. The process is differential in 3 cascaded ways: optical polarimetry is itself differential photometry, a point source of polarization is distinguished from a polarized background distribution by spatially differential polarimetry and on top of this one may use temporal differencing to establish that the source moves as expected for an extra-solar planet, or spectral differencing to establish that the polarization spectrum contains planet-like features. System design of EPICS should reflect these cascaded means of discriminating against spurious detections.

12.1.2 Top level requirements for the polarimetric mode of EPICS

The top level requirements (TLRs) for EPICS are defined in section 5.1. Most relevant for the polarimetric mode of EPICS are the following points:

- **TLR_1: The instrument covers the wavelength range 0.6 – 1.75 micron**

This point includes in the explanation the explicit statement that the wavelengths shorter than 700 nm might be interesting because of a higher degree of polarization.

- **TLR_14: Simultaneous search for terrestrial planets and gas giants will be possible**

- **TLR_2: The total field of view in all observing modes is at least 2″ in diameter at visible wavelengths and 4″ in diameter in the NIR.**

- **TLR_3: The inner working angle in all observing modes working at visible wavelengths is smaller than 30 mas (goal 15 mas).**

- **TLR_4: The spatial sampling will fulfill at least the Nyquist criterion (precise value TBD) at all working wavelengths**

- **TLR_6: There will be a broad band differential polarimetric mode. The absolute polarimetric accuracy will be 1%**

The polarimetric mode measures the fractional linear polarization of a point-like source with a photon-noise-limited precision down to a level of 3×10^{-5} with respect to the surrounding background (star's AO-processed PSF).

- **TLR_9: Earth up to 20 pc at a phase angle of 90° is detected in all main observing modes at SNR > 5 in one night of observation**

- **TLR_10: Jupiter up to 20 pc at a phase angle of 90° is detected in spectroscopic mode at SNR > 50 in one night of observation**

- **The operational efficiency of all modes, including acquisition and observation of the target and any required calibrations is better than 50%.**

- **Polarimetric stability better than 1% over 4 hours.**

The instrumental polarization induced by the telescope up to the telescope final focus (after M6, see **AD 1**) is lower than 1% and stable on a level better than $\Delta Q, \Delta U < 0.1\%$ during several consecutive nights.

- **Physical model of instrument (Internal calibration sources (TBD), functions for calibration of time-dependent effects)**

These TLRs are quite specific for the polarimetric mode of EPICS and there remains not much freedom for the instrument requirements.

Wavelength range: There is some freedom in the exact choice of the wavelength domain. However, the TLRs already favor wavelengths in the domain ~600 nm. According to the

science case the instrument should be optimized for the polarized light reflected from molecules and aerosols. This favors strongly short wavelength for polarimetry, at least down to 600nm, if possible down to 500 nm. Wavelengths longer than 1 micron are not very attractive because the intrinsic polarization of planets will be low. Further, detector technology changes around 1 micron, which might add significant complexity for an instrument working above and below 1 micron.

Thus the wavelength range to be considered with highest priority is **600 nm – 850 nm.** If possible the wavelength range should be extended to 500nm-950nm.

Achievable polarimetric precision: The achievable polarimetric precision must be high enough to reach the photon noise limit for the measurement of the polarization signal of a planet in the AO-corrected, coronagraphic PSF of a bright star observed with an integration time of at least 4 hours. Estimate based on a preliminary PSF for EPICS/OWL indicate that a polarimetric precision of about 3×10^{-6} must be achieved to meet this goal.

The absolute polarimetric accuracy of the polarimetric mode should reach 10^{-4} for most parts of the field of view. This would allow to obtain reliable measurement for extended objects, such as circumstellar disks and shell, which are important secondary science cases.

12.2 Basic principles for high precision polarimetry

12.2.1 Fast modulation

The basic principle of high-precision polarization measurements includes a fast polarization modulator with a modulation frequency in the kHz range, combined with an imaging photometer which demodulates the intensity signal in synchronism with the polarization modulation. The polarization modulator and associated polarizer convert the degree-of-polarization signal into a fractional modulation of the intensity signal, which is then measured in a demodulating detector system by a differential intensity measurement between the two modulator states. Each active pixel measures both the high and the low states of the intensity modulation and dividing the differential signal by the average signal eliminates all gain changes, notably changes of atmospheric transparency or electronic gain drifts. The next paragraph develops this general principle for a particular implementation.

Figure 12-1 The measuring principle for a fast modulation imaging polarimeter.

Figure 12-1 illustrates the basic measuring principle employed in the ZIMPOL (Zurich Imaging Polarimeter) system. In ZIMPOL a CCD is used as detector where every second row of the CCD is masked. Charge packages created in the unmasked row during one half of the modulation cycle are shifted for the second half of the cycle to the next masked row, which is used as temporary buffer storage (the CCD can be equipped with cylindrical micro-lenses which focus all the light onto the open CCD rows). After many thousands of modulation periods the CCD is read out within less than 1 second, thus providing two image planes measured with exactly the same (timeshared) detector array. The sum of the two images is proportional to the intensity,, the normalized difference represents the degree of polarization (of the Stokes component for which the polarimeter optics are adjusted). An important point is that the degree-of-polarization image is the basic result of the measurement; it will be loosely referred to as 'the polarization image'.

Other detector systems, such as arrays of avalanche diodes or special CMOS arrays may be considered instead of a masked CCD (see section 11.4).

Because the measurement of degree of polarization is fully differential and rationing, systematic error sources are reduced to a very low level. Key advantages of this technique are:

- images for the two opposite polarization modes are created practically **simultaneously** (the modulation is faster than the seeing variations),

- both images are recorded with the same detector elements (**same pixels),**
- there are **no differential aberrations** between the two images corresponding to opposite polarization due to the atmosphere or telescope/instrument
- the differential (polarization) signal is **not affected by chromatic effects** due to telescope diffraction or speckle chromatism

ZIMPOL has proved to be an extremely precise technique for polarimetric imaging. It is probably the most precise differential imaging technique with array detectors available today. In solar applications ZIMPOL has routinely achieved a precision of better than 10^{-5} in (long slit) spectro-polarimetric mode. The same performance has been demonstrated in a lab experiment for imaging polarimetry for the CHEOPS planet finder phase A study (RD 4). This is more than an order of magnitude better than other instruments.

In Section 12.4 the properties of possible polarization modulators and demodulating detectors are described in more detail.

12.2.2 Instrumental polarization

12.2.2.1 Why is instrumental polarization an issue

Instrumental polarization can introduce an offset to the measured signal which makes an absolute measurement of the polarization signal difficult. In the extra-solar planet case this problem is for many targets significantly relaxed, as the polarization in the PSF from the central star can be used as zero point level for the polarization on top of which a polarization signal from the planet is measured. This approach assumes that the intrinsic polarization of the central star is very low, lower than 10^{-4}. This is a reasonable assumption as it has been shown that the integrated linear polarization of the sun is less than 10^{-6}. A small contribution from interstellar polarization of the order 10^{-3} (the expected interstellar polarization for nearby stars is 10^{-4} or less: see ref. [11]) affects both central star and planet in the same way and can be treated like a contribution to the instrument polarization.

The instrument polarization can strongly reduce the relative precision of the polarization measurement due to non-linearity effects of the demodulating detector. For a high instrument polarization (say of the order 5%) the intensity level is different for the two opposite polarization modes (e.g. I_0 and I_{90}), which poses stringent constraints on the linearity of the pixel efficiency for measuring a target signal which is more than 3 order of magnitudes smaller (e.g. 10^{-5}). It has been demonstrated for the demodulating CCDs in ZIMPOL that an instrument polarization of 1% still allows to measure a weak polarization signal of the order of 10^{-5}. For a higher level of instrument polarization this precision was not achieved with the same measuring technique.

12.2.2.2 Impact of the requirement for a low instrument polarization on EPICS concept

All optical elements may introduce instrumental polarization effects to some level. For this reason it is extremely important to take the instrumental polarization in EPICS carefully into account.

The polarization measurement yields the polarization of the light at the position of the fast polarization modulator. This consists of a combination of the polarization signal from the bright central star, the polarization from the target (e.g. the extra-solar planet), the polarization from the sky (negligible in the bright PSF halo), and the (telescope +) instrument polarization.

The position of the polarization modulator has the following implications on the instrument polarization:

- all components upstream of the polarization modulator may introduce instrumental polarization which must be under control,
- all components downstream of the beam splitting polarizer produce no instrumental polarization, if their time-constants are much longer than the modulation period.

Thus critical components which introduce significant instrument polarization, such as image rotators, should be placed after the modulator/beamsplitter package.

There are various ways to control the instrument polarization:

- The best way is to have an instrument concept which avoids instrumental polarization as far as possible.
- The instrument concept should consider with high priority a hardware compensation of unavoidable instrument polarization effects.
- Very important is that the instrument concept incorporates strategies for calibrating the instrument polarization on a monitoring basis.

There are two other important aspects which can be used for disentangling the signal from the sky targets and the instrument polarization:

- The instrumental effects are locked to the telescope/instrument orientation. Contrary to this the target signal is different from object to object and moreover it will rotate in the instrument frame due to the field rotation in the alt-azimuth OWL telescope.
- Extra-solar planets, the main science case for EPICS, are point-like sources, while at least most of the potential instrumental effects are expected to cause a smooth signal over the FOV. Thus, in the (degree-of-)polarization image, an extra-solar planet is expected to be a point-like source on top of a (probably smooth) "background" signal from the instrument.

12.2.2.3 Polarization switch and the calibration of the instrument polarization

In principle the polarization modulator package should be placed as early as possible in the beam in order to reduce the instrumental polarization. However, for EPICS, a polarization modulator would strongly interfere with the following AO system and other subsystems of the instrument. Due to this, the polarization modulator has to be placed after many optical components of the EPICS instrument. This disadvantage (from the point of view of polarimetry) can be overcome with a polarization switch placed early in the beam.

Figure 12-2 Schematic illustration of the effect introduced by the polarization switch. For a given pixel on the polarization image p_Q there are contributions to the measured polarization from the instrument (from polarization switch to modulator package) and from the sky and telescope. The switch can now reverse the sign of the contribution from the sky and the telescope. This allows disentangling instrumental effects from real signals from the sky (as transmitted by the telescope; separation of telescope and sky contributions is achieved by using the relative rotation in the course of a night).

The polarization switch is important to disentangle the polarization from the sky and the telescope up to the switch position and the polarization from the instrument from the switch position to the polarization analysis (= modulator package). The switch must be at a location where the instrument polarization is low and stable (no moving components in front of it). During long series of short science observation (e.g. 5000 integrations with t_{exp}=1 sec) the switch would rotate the linear polarization position angle by 90 degrees and the resulting linear polarization signal would behave as illustrated in Figure 12-2. For OWL the switch might be located after telescope mirror M6 (near the final 'scientific' telescope focus, see **AD 1**) . An alternative location ('intermediate' focus) for such a switch is discussed in appendix 12.9; both options have advantages and disadvantages, it may be best to include both.

The method with the switch will only be powerful in identifying a real sky signal if the polarization contribution from all the optical components in front of the switch can be eliminated with calibrations. This requirement is significantly relaxed if only point-like sources are searched, because then only the spurious instrumental effects producing point-like polarization features have to be eliminated.

12.3 A polarimetric concept for EPICS

12.3.1 Basic considerations for the EPICS case

12.3.1.1 Science requirements

The main science case for EPICS is the detection and characterization of Earth-like and Jupiter-like gas planets around nearby stars. The polarimetric mode of EPICS should investigate the reflected light from these planets with high precision polarimetry.

Thus the prime task for the polarimetric mode of EPICS is the measurement of a weak, linearly polarized, point-like polarization signal in the residual AO/coronagraphic PSF of a bright star. Measurements in different wavelength bands should be possible for the characterization of the planets.

Of course, the polarimetric mode assumes that the AO system and the stellar coronagraph of EPICS will allow observations with a very high contrast as specified in the TLRs.

Thus the most important properties of the polarimetric mode of EPICS for achieving the highest possible efficiency must be:

- photon noise limited polarimetric precision
- a high photon detection efficiency
- no spurious effects which could mimic a signal from a planet

The primary science case does not require a very high absolute accuracy in polarimetry. This implies that the light in the residual PSF can be used as zero point reference for the polarization on top of which the polarization signal from the planet is measured. This approach seems to be warranted for most nearby stars, which can be considered to be intrinsically unpolarized in integrated light [11]. Thus, the instrument polarization must be controlled at least to a level where it does not affect the polarimetric precision.

For secondary science cases it would be advantageous to achieve also a high absolute accuracy (scale, vector zero point, angle) for the polarization measurements. This would allow the investigations of extended circumstellar structures of bright stars in polarized light with high accuracy because the overall polarization offset introduced by the instrument can be determined accurately.

12.3.1.2 Instrumental restrictions

The overall concept of the polarimetric mode of EPICS must as far as possible be optimized for high precision polarimetric imaging. On the other hand the concept for polarimetry must be adapted in order to avoid or minimize any impact on the performance of the AO-system, of the coronagraphic system, and of the other instrument arms.

Telescope: In the OWL/EPICS configuration we have a telescope which produces a relatively low telescope polarization. The main effect in telescope polarization will be due to

the tilted telescope mirror M6. In addition one should consider the polarization pattern introduced in the PSF by the telescope diffraction effects, in particular due to the segmented primary and secondary mirror. OWL is an alt-azimuth telescope which has the advantage that the telescope polarization can be accurately determined with observation of the same polarization standard star taken at different telescope positions (e.g. two observations of the same star taken during the same night, symmetrically before and after meridian passage and differing by about 90° in parallactic angle).

EPICS: EPICS will be located in one of the focal stations of the OWL telescope. The focal station moves with the telescope, so that the instrument is fixed with respect to the telescope.

It is in principle an advantage for the determination and correction of the instrument polarization if the whole instrument is fixed in the telescope frame, because no time dependent changes in the instrument polarization are introduced. **An image de-rotator in the common path of EPICS would destroy this advantage** and should therefore be avoided. In any case a de-rotator has strongly tilted surfaces which are very problematic for a polarimetric observing mode. From the polarimetric point of view, the correct location for the image rotator is after the final polarizing beamsplitter, in both beams of the polarimeter.

AO-system: The AO-system is the critical key component for very high contrast applications. An optimum performance of the AO-system is a pre-requisite for the success of the EPICS instrument. Therefore, all components which might affect the AO-performance must be carefully considered. Introducing polarimetric components in front of the AO-system is difficult as they must meet not only polarimetric requirements but also requirements set by the AO system. For this reason we do not consider a concept with the fast polarization modulator located in front of the AO system. Up to now, no detailed study has been made about the feasibility of such a system but it is usually assumed that a modulator package in front of an AO system would introduce different types of significant additional difficulties (such as 2 fully polarized beams).

With the modulator downstream of the AO, the concept of the AO-system should be optimized to avoid introducing appreciable instrument polarization.

Other focal-plane instruments in EPICS: EPICS consists of different focal plane instruments. From the point of view of system efficiency it may be **advantageous if different instrument modes can work simultaneously** by using beam splitters. This sets again constraints on the optical component in the common beam section, which must fulfill requirements of all simultaneously working sub-systems. On the other side, contrariwise, polarimetrically necessary components may prohibit some simultaneous modes.

12.3.2 A "simultaneous" polarimetric mode for EPICS

A simple solution for a polarimetric mode for EPICS would consist of a polarimetric arm where all polarimetric components are located after the AO-system except for insertable

calibration components (e.g. polarization switch, polarizer, etc.) which are used for non-science measurements only (see block diagram in Figure 12-3).

Advantages:

- Having no polarimetric components in the common path makes it relatively easy to perform observations with different EPICS subsystems simultaneously (e.g. 3-D spectroscopy in the near-IR, polarimetry in the visual).
- This polarimetric mode is very simple to implement in the EPICS concept.

Disadvantages:

- There are only limited means to control and calibrate the instrument polarization. Therefore there exists the potential danger that it may be difficult to disentangle spurious instrumental effects from a real polarization signal of a planet.
- The performance for measuring absolute polarization values will be clearly less than optimal and the overhead due to on-sky-calibration of the system may be increased.

12.3.3 A high performance polarimetric mode for EPICS

A polarization switch may be introduced in front of the AO system for the science observations in order to enhance the polarimetric performance of EPICS beyond the simple solution presented above. In particular the polarization switch could be decisive for the success of the polarimetric mode as it allows disentangling the real signals from extra-solar planets from spurious signals introduced by the instrument. The polarization switch would be in the beam for science observations, at least for cases where polarimetry has the highest priority.

Advantages:

- Spurious polarization effects due to the instrument (AO system, coronagraph) can be better recognized, calibrated and taken into account in the data reduction process. In particular it may be possible to distinguish between localized polarization signals from the sky and spurious polarization from the instrument
- A significantly higher accuracy can be reached in the measurement of the absolute polarization level.
- Other EPICS subsystems which are basically photometric (e.g. near-IR spectrometer) would be able to function as simple polarimeters (e.g. spectropolarimeter) by using the polarization switch with a downstream polarizer as a very slow modulator.

Disadvantages

- The polarization switch requires the consideration of chromatic effects, defocus and internal reflections which might affect the performance of the AO system and the coronagraphic system if not taken into account.

- Simultaneous observations with the polarimetric arm and another EPICS subsystem becomes difficult to implement because various constraints are introduced by the polarization switch (limitations set on the wavelength range by coatings, chromatic aberrations, defocus, internal reflections, etc.).
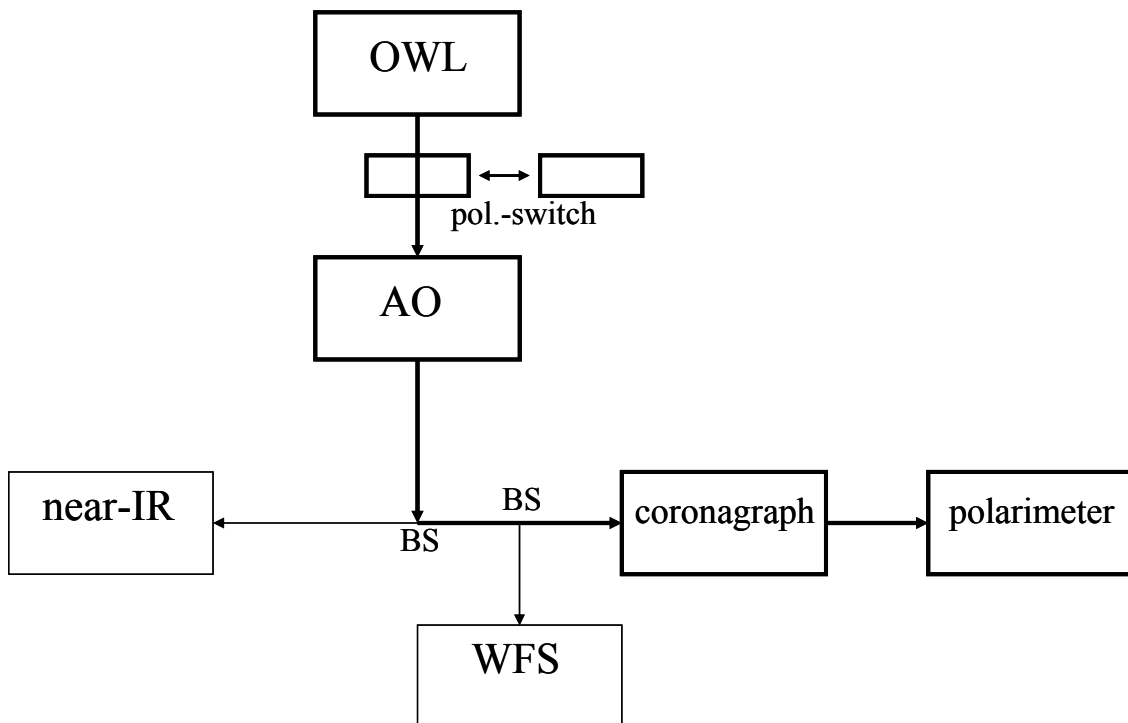
Figure 12-3 Block-diagram for the EPICS polarimetric mode concept.

12.3.4 Recommended polarimetric concept for EPICS

A very attractive solution is to combine the advantages of both observing modes with and without polarization switch (see block diagram in Figure 12-3). In this way it would be possible to have a versatile polarimetric instrument which can be used for survey work but also for the detailed characterization of targets.

Without polarization switch:

- Polarimetry can be operated in parallel with other EPICS subsystems with little restrictions. This may be particular important for surveys as the efficiency of the instrument is high.

- If instrumental effects are well under control (e.g. after the data analysis of many nights) then even this mode may provide a maximum efficiency for the *detection* of point like polarization signals from an extra-solar planet.

With polarization switch:

- The instrument effects can be much better controlled and calibrated. This allows identifying planets with high confidence down to the photon noise limit even for long integrations of bright targets.
- The polarization can be *measured* with a high absolute polarimetric *accuracy*. Extended polarization signals from e.g. circumstellar disks or shells can be disentangled well from instrumental effects.
- Other EPICS subsystems can perform limited polarimetry.

The requirements set for the two modes by the science are different and they have to be defined in detail for the planning of the polarization switch properties. It is clear that the mode without switch must be optimized for survey observations. Contrary to this, the mode with switch must be optimized for the detailed characterization of key targets. For the characterization of planets it seems to be acceptable that other focal-plane instrument of EPICS can not be operated in parallel (or only with restrictions) with the imaging polarimeter. Thus the polarization switch can be tuned to the requirements for a high performance polarimetric measurement with EPICS; as the science case stands at present, extending polarimetry to other EPICS subsystems is of secondary importance.

Issues related to the polarization switch:

There are many engineering issues which are related to the polarization switch:

- focus change if switch is moved in and out
- chromatic effects (aberrations) of the polarization switch (transmitting component)
- internal reflections

In the Appendix 12.9 one possible concept for such a polarization switch is presented and technical issues are addressed. It should be considered that the switch described in the Appendix 12.9 is considered for the intermediate focus (between telescope mirror M4 and M5) where the focal ratio is $F/\# = 2$. From a polarimetric point of view it is advantageous to have this switch as early as possible in the beam. However, other solutions are possible which would be less demanding for the properties of the polarization switch.

We list possible locations for the polarization switch. A possibility not to be ruled out is to include insertable switches at more than one of these locations, for different science programs. The possibilities are:

- Near the intermediate focus after M4: From the polarimetric point of view an attractive solution, because all the components in front of the switch are symmetric (at least hexagonal symmetry) with respect to the beam and the induced instrumental polarization will be very low. Problematic for a transmitting polarization switch is

the fast beam with an F-ratio of F/# = 2; nothing like this has been done before (see Appendix 12.9 for details).

- Near the final telescope focus (after M6): Mirror M6 introduces due to its tilt with respect to the telescope axis an instrumental polarization of a few tenths of a percent. Thus, there will be instrument polarization at the position of the switch, which can and has to be taken into account. The beam is at this position considerably slower with an F-ratio of about F/# = 6; a precision polarimeter at F/6.6 has been built (C.U. Keller, private communication).

- From the optical-design point of view, the ideal location for an insertable component would be a collimated beam section (near the final telescope focus) just after mirror M6. In a collimated beam section components such as a polarization switch can be inserted without introducing strong defocus or chromatic effects. This possibility has to be investigated in connection with the EPICS AO concept.

The concept of an insertable polarization switch is currently foreseen and being investigated in detail for the ESO VLT planet finder instrument.

12.4 High precision imaging polarimeter

The high precision imaging polarimeter measures the polarization of the incoming light at the position of the fast polarization modulator package. According to the recommended concept for EPICS (Figure 12-4) the high precision imaging polarimeter is located after the AO system, the dichroic beam splitter and the WFS beam splitter. Because the coronagraphic concept of EPICS is currently not defined it is not clear whether the modulator package should be located in front of or after the coronagraphic system. The pro and cons are discussed in section 12.4.2. In the following concept we assume that the coronagraphic system is located in front of the modulator package.

12.4.1 A concept for the high precision imaging polarimeter

It is assumed that the beam comes from the coronagraphic focal plane and passes first through a collimating lens and a coronagraphic (Lyot) pupil mask before it enters the high precision polarimeter.

In the high precision polarimeter the beam passes first through a rotatable half wave plate which selects between a Stokes Q or U measurement. Then follows the polarization modulator package, including a modulator and a polarization beam splitter producing two beams (see Figure 12-4). Both beams will be recorded with their own detector systems in order to collect all light from the telescope. In both beams the same polarization signal is encoded as intensity modulation. Thus, double beam mode observations provide two full polarimetric observations which can be reduced and analyzed independently of each other. The results from both beams can then be combined at the end, in order to achieve maximum efficiency and to eliminate common-mode noise. This is the basic concept of all polarimeters based on fast polarization modulation.

Different filters can be introduced in the two beams so that polarimetry and differential imaging in two pass band filters can be taken simultaneously. Thus the polarimeter can also

be used as a "simple" differential imager or differential polarimeter, with 50% efficiency in each beam. As half wave plate and modulator are not used for differential non-polarimetric imaging, it may be advantageous to have the possibility to retract them, in order to reduce stray light and ghosting effects and to enhance efficiency. Retractable transmission components should be placed in a parallel beam section in order to avoid focus shifts and chromatic effects.
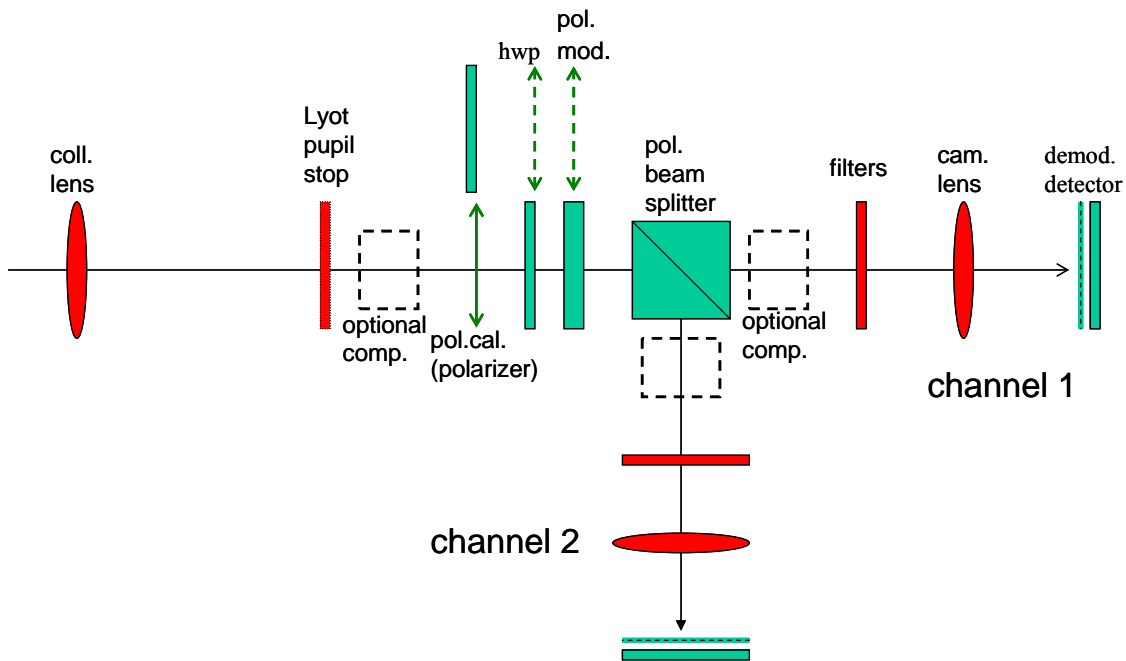
Figure 12-4: Optical scheme for the high precision polarimeter in EPICS. The important elements for the polarization are colored in green. The optional components in the input beam could be (for instance) birefringent filters for differential imaging, those in the 2 polarimeter arms would be image rotators, grisms, etc.

Also required are polarimetric components for the calibration of polarization cross talk effects and the zero point of the position angle for the high precision polarimeter. This requires an insertable and rotatable polarizer plate (and quarter wave retarder for circular polarization) which can be introduced in the beam for calibration measurements with an internal calibration lamp from EPICS; the beam from the calibration optics should be identical (pupil and image) to that from the telescope.

Various other components may be introduced in the (collimated) beam section. Of much interest would be birefringent filters. Such components would allow a fast modulation/demodulation of the spectral signal similar to the polarimetric modulation/demodulation. This type of filter would provide essentially aberration free differential imaging capabilities (see Section 12.4.4).

12.4.2 Location of fast modulator with respect to the coronagraph

It is not clear yet whether the modulator package should be located in front of or after the coronagraphic system. This certainly depends on the concept for the coronagraphic system which is not defined yet. The following points should be considered for the choice of the final concept.

Advantages of a concept with the fast modulator in front of coronagraphic system:

- The coronagraphic system introduces no polarization effects in the final PSF.

Advantages of a concept with a fast modulator after the coronagraphic system:

- The central peak of the PSF is blocked by the coronagraphic system and does not pass through the transmitting polarization components. Thus there is much less danger that point-like ghosts are produced by unwanted reflection from surfaces which might look like a localized polarization signal as expected from a planet.
- Only one coronagraphic system has to be built (if modulator package is in front then two systems, one in each arm, are necessary).

Currently hardly any experience has been gained about the polarization effects which could be produced by a coronagraphic system. We assume (but it is exactly that: an assumption) that the polarization effects from the coronagraph can be calibrated so that the corresponding spurious effects can be taken into account in the data reduction. Therefore we adopt in the following a concept where the coronagraphic system is in front of the modulator package, but this may be changed later in the light of experience with actual coronagraph systems.

12.4.3 Optical concept for the imaging polarimeter

The optical setup of the imaging polarimeter has to re-image the coronagraphic focal plane onto the focal plane of the demodulating detector with a pixel scale of about 0.5 mas/pixel. Different components must be placed in this beam section:

- coronagraphic pupil mask(s)
- half wave plate for selecting Stokes Q or U
- polarization modulator
- polarizing beam splitter
- insertable calibration components for the polarimeter
- exchangeable pass band filters
- optional, exchangeable or retractable components, such as birefringent filters, image rotators or grisms

The optical concept. We propose for EPICS a concept (based on the ZIMPOL set up foreseen for the VLT planet finder) which includes a parallel beam section. Thus the concept consists of a diverging beam from the coronagraphic focus, a collimated (ie parallel) beam section including all the analyzing optics (filters, polarization components), and a converging beam with an F-ratio of the order F/# ≈ 100 (depending on the pixel size of the detector).

This is certainly not the only possible solution for an imaging polarimeter in EPICS, but the study for ZIMPOL/PF supports this choice (**RD 1** to **RD 4**).

Parallel beam. The parallel beam section is ideal for thick transmitting, exchangeable, or insertable components, because chromatic effects for transmitting components and focus positions changes due to the inserting or exchanging of components are for many types of optical components negligible. Retracting or inserting analyzing optical components in the parallel beam allows switching between different observing modes, enhancing significantly the versatility of the instrument.

The surface quality of the components in the parallel beam can be relaxed with a small pupil diameter in the beam. Due to the narrow FOV of EPICS (about 2"× 2") it is possible to have a small pupil diameter without introducing too large field angles in the parallel beam. A small pupil is also ideal for having a compact polarimeter concept.

The small beam requires on the other hand, that the cosmetic surface quality and the internal purity of the components are very high in order to minimize stray light caused by inhomogeneities. Ghosting and the effect of internal reflections must be reduced with very good anti-reflection coatings. In addition, the residual ghosting has to be suppressed by tilting (by a few degrees) as many components as possible.

The mechanical requirements for insertable plane parallel plates (filters, retarder plates, etc) in the parallel beam are not high. Only the coronagraphic pupil mask has to be positioned with a high precision (tolerance a few microns?) in order to achieve the expected performance.

Requirements on non common path aberrations. When a measurement requires combined data from both beams, the requirements on the optical components in the imaging polarimeter will be similar to the requirements for the differential imaging instruments. In particular, non-common path aberrations are an important issue. The fast modulation polarization imaging has the huge advantage that the beams for both opposite polarization modes pass through the same components and are registered with essentially the same detector element. Nonetheless, the birefringence effects in the components, especially in the polarization modulator, may introduce differences between the two polarization modes which have to be investigated.

The imaging polarimeter has two arms after the polarization beam splitter. In principle the imaging polarimeter can also be used as differential imager by inserting different filters in the two arms and removing the modulator. For a very high performance in such differential imaging the differential aberrations in the two arms must fulfill at least the requirements as the differential imager in the IR. The science case must define whether such a high performance differential imaging mode in the visual-red spectral region is desirable.

12.4.4 Upgrade options

12.4.4.1 Differential imaging with modulating filters

A major drawback of the "classical" differential imaging approach is that the two images with different pass-band filters are registered either (quasi-)simultaneously or with the same detector pixels, but rarely or never both in one and the same instrument mode. This implies that the pixel efficiency or non-common-path aberrations have to be calibrated in order to reach a photon-noise-limited precision.

The EPICS imaging polarimeter provides a technique which employs the same detector pixels for the quasi-simultaneous observations of two images by rapidly switching between two different observing modes. In polarimetry this is achieved with a switching between opposite polarization modes using a fast switchable polarization modulator.

A similar modulation/demodulation mode can be achieved for differential imaging using Lyot or Solc filters, or birefringent Fabry-Perot filters in combination with the polarization modulator and demodulating detector of the imaging polarimeter [37]. Birefringent Fabry-Perot filters produce double pass-bands which are shifted in wavelength for the polarization direction 0 and 90 degree. Let us assume that the pass-band for 0 degree is centered on the Hα line (656 nm) and for 90 degree on the adjacent continuum. With a polarization modulator (and the polarization beam splitter) the intensity difference between these two pass bands can be transformed into a fast intensity modulation which can be demodulated by the detector of the imaging polarimeter. With Lyot and (modified) Solc filters the same effect can be achieved; the choice between these depends on the science to be performed, ease of manufacture, etc. A series of papers [41] from the 1960s needs careful study in this respect. In these papers, various synthesis techniques and convincing experiments are described using a design similar to the Solc filter. The applications are for narrow-band filters (< 0.1 nm), but bandwidth is an inverse function of crystal thickness; pass band shape is an input to the procedure, it is quite possible that it can be applied to the exoplanet search case; this would then influence the specifications (e.g. wavelength range) of some of the polarimetric components. Solc filters are reported to require strict tolerances (C.U. Keller, priv. comm.), but that should not deter us in this case.

From differential imaging with modulating filters it can be expected that the **errors due to flat-fielding and differential instrumental aberrations can be substantially reduced**. Such a modulated differential imager could be particularly useful for the Hα line, which is one of the strongest circumstellar emission features seen in a large fraction of bright stars.

The disadvantage of these polarization-derived differential-imaging methods is that the final signal on the detector is one half or even one quarter of that for some of the 'classical' methods (see section 10). If, with the chosen bandwidth and astronomical target, detector operation is readout-noise-limited, then this would affect the performance w.r.t. the classical methods. The conclusion might be that both options should be made available within EPICS.

12.4.4.2 <u>Spectroscopy (Spectropolarimetry)</u>

Low resolution exoplanet spectropolarimetry would require a focal plane aperture mask or lenslet array (in the coronagraphic focus) plus an insertable grism in the parallel beam. It is probably only of interest after a planet has been found.

12.4.5 Modulators

There are different types of modulators available on the market. We provide here only a short overview about the advantages and disadvantages of the different modulator types.

Piezoelastic modulator. Piezoelastic modulators produce oscillating birefringence at a fixed frequency in the low frequency ultrasound range (20 kHz to 100 kHz). In a fused silica slab, which is optically isotropic under normal conditions, a standing wave is excited with a piezoelectric inducer. The oscillating mechanical stress produces optical anisotropy and thus birefringence. The optical retardation depends on the stress amplitude but the orientation of the optical axis is fixed.

Advantages:

- widely used device with which a very high polarimetric precision has been reached in the past (Kemp polarimeter, ZIMPOL)

- good optical quality of fused silica slab

Disadvantages:

- modulation frequency is relatively high; (if a demodulation with charge shuffling is required one may run into problems with the charge transfer efficiency)

- the oscillating retardation is sine-like and therefore not as efficient as for a square-type modulation; further the device has a chromatic birefringence effect, which reduces further the efficiency for broad band work.

- the induced retardance is position dependent and only homogenous in the central region of the modulator (the useful size of the device is limited)

Piezoelastic modulators should certainly be considered for a imaging polarimeter, if not the ultimate photon (polarization) efficiency is required and where the demodulating detector is able to cope with a fast modulation.

Ferroelectric Liquid Crystal retarders (FLCs): FLCs are bistable, electrically switchable wave plates. A thin layer (approximately 2 μm) of a ferroelectric uniaxial, birefringent liquid crystal is sandwiched between two glass plates. By selecting the appropriate thickness, the FLC layer functions as a half wave plate at a distinct wavelength. The optical axis of the liquid crystal layer has two preferred orientations which differ from each other by approximately 45 degrees. The orientation of the optical axis is controlled by a drive voltage applied across the FLC layer. Thus, one of the two states can be selected by applying an

appropriate voltage. Thus FLCs have a fixed retardation but change the orientation of their optical axis with time.

Advantages:

- often used in optics for fast shutters (in combination with a polarizer)

- can be used with a high efficiency for broad band polarimetry when combined with a rotatable zero order half wave plate

- first test in solar physics and astronomy are very promising; this type of modulator is foreseen for the VLT planet finder.

Disadvantages:

- Switch time is typically of the order 10-100 μsec limiting an efficient modulation to a frequency of about 1 kHz (a high modulation frequency requires an enhanced working temperature ~25 C for the device)

- not much experience in demanding astronomical applications yet

FLCs are certainly very efficient and versatile polarization modulators for imaging polarimetry as long as the modulation frequency does not have to be significantly higher than 1 kHz.

Alternative modulators: Pockels cells or stress induced modulators are probably viable alternatives for future modulators. In particular, Pockels cells have already been successfully used for astronomical applications in the past.

12.4.6 Demodulating detectors

A demodulating imaging detector has to register the polarization information encoded as intensity modulation (by the polarization modulation package) so that this signal can be reconverted into a polarization value. For this the detector has to separate the photons arriving during one half of the modulation cycle from the photons registered during the other half of the modulation cycle. The normalized difference between the two half cycles is the resulting polarization signal.

For EPICS the demodulating imaging detector / detector mosaic must have an area of about 4000×4000 pixels. It must be capable of registering a photon flux per pixel of the order of 10^3-10^5 per second without significant overhead (<10% of the integration time) and all instrumental noise effects (e.g. RON) should be smaller than the photon noise under all circumstances.

In principle there exist various possibilities to achieve this technically, for example:

- continuous, synchronous detection of photo-electrons (lock-in detector)
- read out of photo-electrons after each half cycle of the modulation,

- storage of the photo-electrons registered during the first or second half of the modulation cycle in two different buffers respectively, and readout after many modulation cycles.

The lock-in technique is well known for high precision aperture (single channel) polarimetry with photo-multipliers or avalanche diodes. However, for imaging polarimetry large array detectors must be used. In future it may be possible to use arrays of avalanche diodes.

Read-out of large array detectors after every half cycle with a frequency of about 1 kHz or faster requires that the read-out noise for the short readout intervals is smaller than the photon noise. For high precision polarimetry the modulation frequency must be faster than the seeing variations (changes in the speckle pattern) in order to achieve the photon noise limit in the polarimetric precision.

The third method can be realized with demodulating CCDs, which are considered for ZIMPOL in the VLT planet finder project. In future also CMOS detectors may become available which are suited as demodulating array detector.

Key properties of the demodulating array detector for the EPICS imaging polarimeter

- imaging array / array mosaic with 4000×4000 pixels or more
- same detector pixels are used for both image planes (for opposite polarization modes), so that the pixel efficiency can be rationed out in the reduction process
- modulation / demodulation frequency should preferably be faster than seeing variations ($\geq$ 1 kHz)
- must be implemented in a system without differential optical aberrations using an electro-optical or stress-birefringence polarization modulator (not a rotating plate)
- read-out noise lower than photon shot noise for primary science targets (for a full well capacity of 10^5 photoelectrons the read-out noise should be RON $\leq$ 10 e$^-$).
- high quantum efficiency of array detectors ($\geq$ 70%)
- high demodulation efficiency ($\geq$ 95%)
- no significant overhead to register photon fluxes of the order 10^5 photons per sec and pixel

CCD arrays (ZIMPOL-detectors). CCDs can be equipped with a stripe mask which covers every second row of the detector. The photo-electric charges registered during the first half of the modulation cycle in the open rows are shifted during the second half of the cycle to the adjacent covered row which serves as buffer storage. Then the charges are again shifted back to the open rows for the first half of the next cycle and additional photon events are registered. The photo-electric charges from the second half of the modulation cycle are treated in the same way with a phase shift of half a cycle. In this way the CCD registers alternately and with the same detector pixel two images corresponding to opposite polarization modes, e.g. I_0 and I_{90}. The CCD can be read out after many modulation cycles, after having accumulated many (thousands) photons per pixel. Thus the photon shot noise will be larger than the read-out noise.

Due to the stripe mask the CCD has only a filling factor of less than 50 % for the active (open) pixels. In order to retain the system efficiency the filling factor can be restored with cylindrical micro-lenses which focus the light onto the open rows.

This type of detector is foreseen for ZIMPOL in the VLT Planet finder. All previous ZIMPOL (Zurich IMaging POLarimeter) systems are based on demodulating CCD arrays.

CMOS detectors. CMOS detectors consists of a photosensitive layer (Si or HgCdTe) on top of a CMOS multiplexer array for storage and read-out. A typical CMOS multiplexer contains two capacities, one of which is used for integrating the charges of the current exposure, while the other holds the charges from the previous exposure that is concurrently read out. Instead of only 2 capacities, a CMOS multiplexer could be designed and produced with 4 capacities within every pixel to hold 2 separate image states.

Thus, instead of shifting charges back and forth as for the masked CCD a CMOS array would just distribute the charges collected during the first and second half of the modulation cycle into different integrating capacities. In this way the two images from the opposite polarization modes can be detected and integrated in two image planes in the CMOS imaging array.

Avalanche diode arrays and other photon counting detectors: Any type of photon counting detector arrays would be a very interesting device for an imaging polarimeter. In such a device it would be possible to use each pixel of the array as a lock-in detector. Based on this principle it was possible in the past (by Kemp) to achieve a polarimetric precision of almost 10^{-7} using a Piezoelastic modulator and a photomultiplier.

12.5 High precision Imaging Polarimeter detection capability

This section explores with simple simulations the performance of the polarimetric mode of EPICS/OWL for the detection and measurement of a point-like polarization signal in the PSF of a bright star. For this we investigate the dependencies of the expected S/N ratio on a range of both astronomical parameters and observing conditions.

Some parameters in these simulations were adopted from similar simulations carried out for the CHEOPS planet finder phase A study (RD 3). However, we have carefully re-considered for OWL/EPICS all those parameters which have a substantial impact on the final results.

12.5.1 Model Inputs

The models calculate the photon flux from parent star and planet and the resulting S/N ratios for various combinations of input parameters as detailed below. We adopt only parameters for the I (800nm) and R (650nm) bands.

Stars: Relevant parameters for our calculations are:

- Spectral types: G2V, M0V (only for 5 pc)

- Distances from sun: d = 5 pc, 10 pc, 20 pc

In addition we adopt an atmospheric extinction of 4 % in the I-band which corresponds to an air mass of 1.15 or a sky position of 30 degree (zenith angle).

Planets: For our simulations we adopt only representative values for the planet properties in the I and R band. The dominant uncertainty in the planet parameters is introduced by the polarization parameter, which is not well known for Earth-like and Jupiter-like planets. The adopted value of p(90) = 0.15 is a reasonable guess for the R-band (in the I-band the polarization might be a bit lower, perhaps p(90) = 0.10).

- Size: Earth radius (R_E) and Jupiter radius (R_J)
- Phase angle (angle: Star – Planet – Observer) of planet: 90 degrees
- Apparent separation: δ = 0.02" to 0.55" → physical sep.: D [AU] = δ ["] / d [pc]
- Reflectivity for phase angle 90 degrees: A(90) = 0.13 (for Jupiter and Earth)
- Polarization for phase angle 90 degrees: p(90) = 0.15 (for Jupiter and Earth)

The photon flux of the planet can be expressed as contrast C relative to the photon flux from the central star:

$$C = F_{planet} / F_{star} = A(\alpha) \ R^2/D^2$$

R is the radius of the planet and D the physical distance to the central star. The factor R^2/D^2 is $8.4 \cdot 10^{-9}$ for Jupiter and $1.8 \cdot 10^{-9}$ for Earth. The polarimetric imager of EPICS will measure the polarization signal from the planet. Thus the contrast between the polarized flux from the planet and the total flux of the star is:

$$C_{pol} = F_{planet} \ p(\alpha) / F_{star} = A(\alpha) \ p(\alpha) \ R^2/D^2$$

Instrument: The following parameters were used for OWL/EPICS.

- Effective telescope area: 7150 m² (π (50² – 15²))
- Effective spectral pass band width: 280 nm (650 – 930 nm)
- Net exposure time: 4 hours
- Telescope/instrument transmission: T = 0.2. Thus 20 % of the light in the wavelength pass band incident on the telescope area ends up being detected.
- Polarization efficiency: ε_{pol} = 0.9. This quantity describes the factor by which a polarization signal (polarized flux) from a planet is degraded by Q,U →V cross talks by the telescope and EPICS and imperfect modulation and demodulation in the imaging polarimeter.

AO performance: This is the most critical component in the overall performance calculations. We have adopted the resulting PSF from the semi-analytic AO simulations for OWL/EPICS from Christophe Verinaud for a set of conditions named A1 (seeing; 0.85, tau0=3ms, 20cm sub-aperture size (500x500 pyramid sensor)[1], frame rate 3.3 kHz, OWL pupil with spiders, no segmentation).

The corresponding PSF profile for 700nm is shown in the following Figure 12-5.
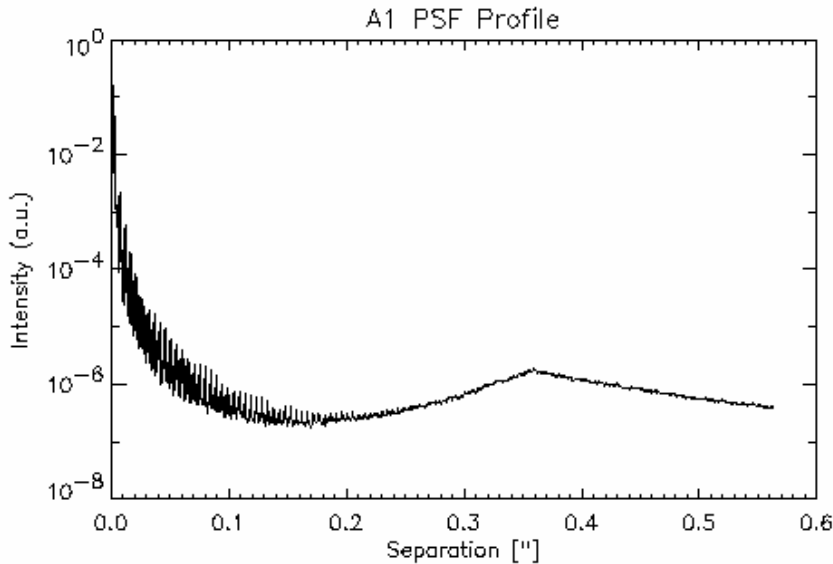
Figure 12-5 Adopted PSF profile for the simulations. It was assumed that the PSF peak can be adopted for the off-axis planet signal. The PSF peak of the central star is blocked by perfect coronagraph introducing no diffraction residuals.

Speckle noise: Speckle noise is neglected in the simulations. It is assumed (based on simple analytic estimates) that fast modulation polarimetry is not affected by speckle noise, because the modulation/demodulation is faster than the seeing variations and any remaining speckle noise may be removed by combining results from both channels of the polarimeter. Thus, the observations are assumed to be limited by photon noise. Due to the high photon flux the read out noise can be neglected for the anticipated demodulating detectors (RON < 10 e).

12.5.2 Model output

Figures Figure 12-6 to Figure 12-9 show the planet signal, residual stellar noise and the resulting S/N ratio as function of apparent separation for Jupiter- and Earth-sized planets. Marked are the positions where the physical separation for the Earth-sized planet is 1 AU, and for the Jupiter-sized planet 5.2 AU as in the solar system. For all simulations a net exposure time of 4 hours was used. The polarization of the planet is always assumed to be p(90)=0.15.

Also indicated in the signal plots are the photon noise limit $(I_{PSF}(\delta))^{1/2}$ from the PSF of the central stars.

Planets at 5 pc around G2V star

Figure 12-6 shows that an Earth around G2V star at 5 pc can be clearly detected with a about S/N=10. A Jupiter-type planet could be detected with about a S/N=100 or higher within the simulated region. A Jupiter with a separation of 5 AU (as in the solar system) would be located at a separation of 1" outside the simulated region. Nonetheless we can assume that an instrument with such a FOV would be capable to see such a Jupiter with S/N=100, because the S/N curves are flat outside the 0.35" (control radius of AO system).
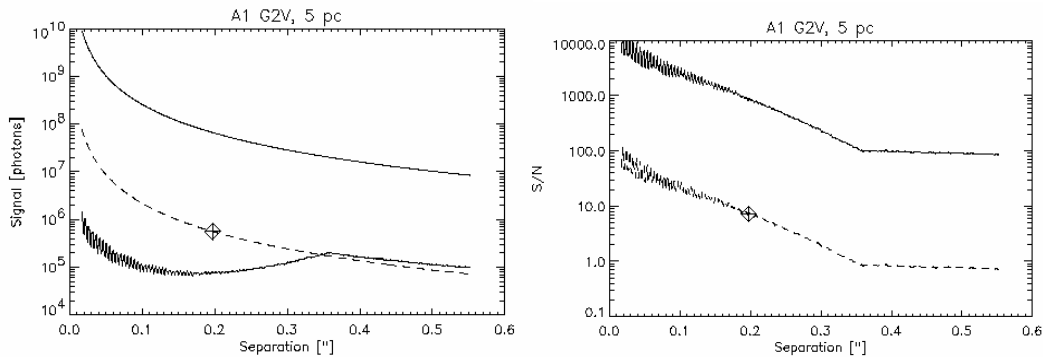
Figure 12-6 : Planet signal and PSF noise (left) and detection S/N (right) for planets around G2V star at 5 pc. The thin solid line is for a Jupiter-sized planet and the dashed line for an Earth-sized planet, both with p(90)=0.15 and A(90)=0.13. The separation corresponding to 1AU is marked for the Earth-sized planet. The thick solid line in the left panel indicates the photon noise limit due to the PSF of the bright central star.

Planets at 5 pc around M0V star

There are many low mass stars in the solar neighborhood. The habitable zone for these stars is located closer to the star with a typical separation of the order 0.3 AU (0.07" for d=5pc). An Earth-like planet with a separation <0.1" would be detectable S/N~5 in 4 hr but it would be hard to see an Earth-like planet at 0.2" (see Figure 12-7). Jupiters around M0V stars at 5pc would be detectable with S/N ≈ 20 or more.
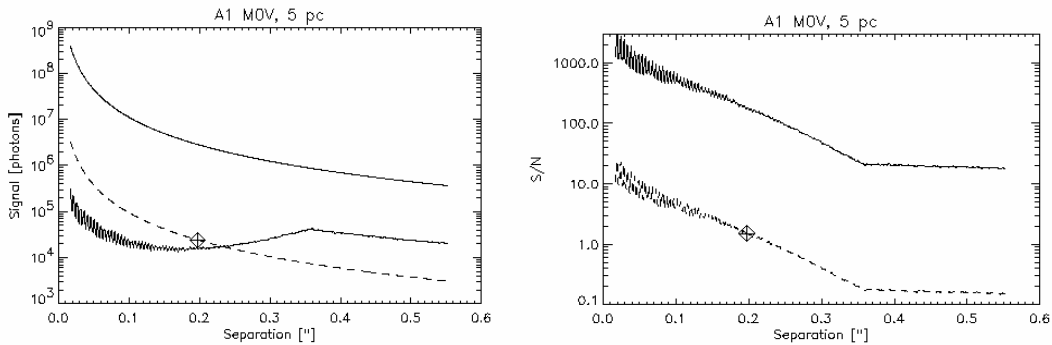
Figure 12-7: Same as Figure 12-6 but for planets around a M0V star at 5 pc.

Planets at 10 pc around G2V star

Figure 12-8 shows that 10 pc is close to the limit where an Earth in a solar system analog can still be detected in 4 hours of integration time.
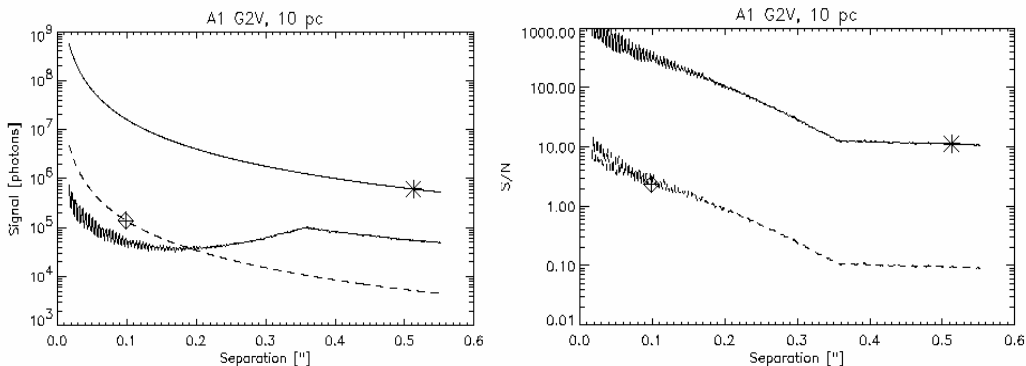
Figure 12-8: Same as Figure 12-6 but for planets around a G2V at 10pc. The separation corresponding to 1 AU for an Earth-sized planet and 5.2AU for a Jupiter-sized planet are marked.

Planets at 20 pc around G2V star

For this distance the Earth-like planet cannot be detected with an integration time of 4 hours. Jupiter-type planet are still detectable with a S/N=10.
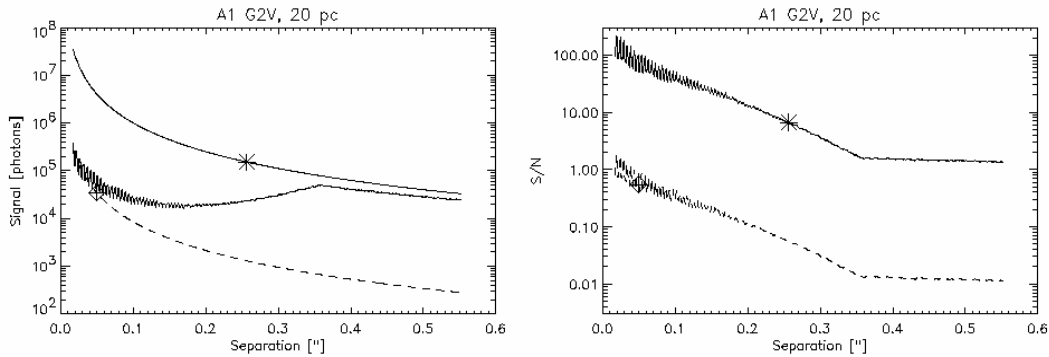
Figure 12-9: Same as Figure 12-6 but for planets around a G2V at 20 pc. The separation corresponding to 1 AU for an Earth-sized planet and 5.2AU for a Jupiter-sized planet are marked.

12.5.3 Conclusions for the EPICS Polarimetric mode performance analysis

The simulations give a first overview what can be achieved in 4 hour exposure time with the proposed polarimetric mode for EPICS / OWL. However one should note that these simulations are not based on the latest AO PSF curves and the latest coronagraphic performance described in earlier Chapters of this document. The estimates made in the previous sections should be redone with updated instrument properties, which was not possible for this version of the document due to the very short time schedule. [1]

Nonetheless we summarize here the preliminary results described in the previous section:

- At 5 pc an Earth (D = 1 AU) can be detected with S/N ≈ 10 and a Jupiter (D = 5 AU) with S/N ≈ 100 around a G2V star. For a low mass star of spectral type M0V an

---

[1] The simulation in the present section are based on an earlier AO concept, involving a single stage 500x500 pyramid sensor. This concept has been temporarily ruled out for important computation time issues due to the complexity of the pyramid sensor signals. However new algorithms to be developed, but still hypothetic, could make this system more realistic. New estimations based on the double stage baseline AO concept have been done in section 13, however this baseline concept is less adapted to visible correction.

Earth like planet in the habitable zone (separation of about 0.3 AU) would also be detectable. A Jupiter like planet would be detectable with a S/N = 20 or higher.

- The distance limit for the detection of an Earth (D = 1 AU) around a G2V star in 4 hour integration time is about 10 pc.
- Jupiter sized objects around G2V stars can be detected with S/N ≥ 10 in the entire FOV (2" × 2") offered by the instrument out to distances of 20 pc (physical separations up to 40 AU).

These results were derived for "normal" system parameters. There might exist a significant population of extra-solar planets with more favorable properties than Jupiter and Earth. For example, a much higher polarization, of the order p=0.3 may be frequent.

It is important to note that the polarimetric mode of a planet finder instrument will only be successful if it achieves a very high polarimetric precision because a very faint polarization signal must be found in a very bright residual PSF halo. This is the case because:

- The Strehl ratio provided by the AO system will be "only" of the order 0.5 in the visual-red spectral region, leaving a lot of photons in the PSF halo.

- Polarimetry will only find a fraction of the reflected light from a planet. Although there might be highly polarized planets out there, the amount of reflected light is very small for all old planets similar to the objects in the solar system. There will be no bright, self-luminous, young planets with a high polarization, because the thermal light from a planet will show only a low polarization level (<1% typically).

- The best targets to find a polarization signal from a planet are very bright (brighter than 5 mag) nearby stars. For these targets it will be necessary to achieve a polarimetric precision of 10^{-6} in order to reach the photon noise limit in a one night integration and to reach the goal of the detection of an Earth like planet. It makes no sense to construct an instrument with much lower precision (less than 10^{-5}), because then it is useless to integrate longer than a few tens of minutes. With such an instrument it would only be possible to conduct a relatively shallow survey for planets of the best targets (bright nearby stars) which would be finished after a few nights of exposure time. Most likely it would only be possible to detect giant planets with a "low"-precision instrument.

12.6   Mid-term development plan

The current plan to include a high precision imaging polarimeter in the VLT planet finder instrument is an ideal situation for the implementation of such a device in the EPICS instrument. If this project is realized then much experience will be gained during the development and with real on-sky observations in combination with an extreme AO system.

Important issues which will be clarified with such an instrument at the VLT will be:

- How efficient is the speckle noise reduction with fast modulation polarimetry under different seeing conditions?

- How much calibration and data reduction efforts are required to achieve a polarimetric precision of 10^{-5}.

- How can a precision of 10^{-6} be achieved and which systematic noise sources are important in this regime.

- Which type of coronagraphy is best suited for high precision polarimetry?

There will be also issues which are not clarified by developing and using a polarimetric system in the VLT planet finder instrument. These issues are:

- How can the optical quality of the polarization components (in particular the polarization modulator) be further improved to achieve the even higher requirements for the EPICS/OWL case?

- Will it be possible to have a "demodulating detector" or "demodulating detector mosaic" with in total 4000×4000 pixels?

Both these points seem currently to be non-critical considering the expected technological advance in the coming decade. Liquid crystal technology is rapidly evolving and also the need for "photon counting" array detectors in science and technology will push the development in the desired direction. However, beside this optimism one should also take into account that polarimetry is only a "niche technology" for which the number of available component types and providers will be limited and may require expensive developments for a sophisticated instrument.

Appendices to polarimetry section

12.7 Appendix: Important points to be considered in connection with the instrument polarization

12.7.1 Ways to avoid instrument polarization

- The position of the modulator package (including the polarizer) should be as early as possible in the beam, because all following components introduce no instrument polarization (impact on instrument: modulator introduces phase shifts and polarizer selects only half the light for the instrument arm of the transmitted beam while the other half is either lost or requires a second instrument arm.)
- The instrument polarization cancels (apart form diffraction effects) for optical components which are rotationally symmetric with respect to the beam. For OWL the telescope mirrors M2, M3, M4 and M5 fulfill this condition. For M1 the geometry of a partially filled mirror is critical for the instrument.
- For all optical components the angle of incidence of the beam with respect to the surface normal should be as small as possible. Inclined surfaces introduce a polarization which depends on the inclination angle α roughly like: $p \sim \alpha^2$??. They also introduce differential phase shifts, causing polarization conversions. Particular attention is required for beam splitters.
- Surface coatings should be optimized for polarimetric purposes. The introduced polarization (instrument polarization and retardance) should be as small as possible. This applies particularly to OWL mirror coatings. It will be necessary to measure the properties of all coatings being considered, both freshly deposited and after several periods of aging under realistic atmospheric conditions; ESO has some such samples (Gilmozzi, private communication).
- Components with non-normal surfaces should be fixed, in order to not introduce a strongly changing instrument polarization. In particular, image rotators should be avoided in front of the modulator package (the effect of deformable mirrors or tip-tilt mirrors is small and most likely negligible).
- Transmitting components (lenses, Atmospheric Dispersion Compensator) must be checked for (stress-) birefringence effects.

12.7.2 Ways to compensate instrumental polarization

- The instrument polarization may be reduced with an inclined glass plate which deflects predominantly light with one polarization direction. A chromatic polarization component will remain.
- The polarization introduced by an inclined mirror can be compensated with a crossed mirror having the same surface coating (including identical aging).

12.7.3 Effects which may cause harmful instrument polarization

- Non-normal reflection from unsuitable or dirty (fingerprints) surfaces can produce strong and harmful polarization effects
- Coronagraphic phase masks can destroy completely the polarization information. If the polarization modulator or switch precedes the coronagraph and the coronagraph optics are stable (for periods much longer than the modulation/switching period), this should not apply (the polarization information has been transformed into a fractional intensity modulation; see section 12.2.1).
- Ghosting may produce harmful polarization signals, especially if the ghost light passes in an unforeseen way a polarimetric (polarization inducing) component.

12.7.4 Not well investigated effects which should be considered in future studies

- Diffraction effects by the telescope optics cause some polarization effects in the diffraction pattern which have so far not be investigated in sufficient detail.
- Coronagraphic focal plane and pupil masks may produce also polarization effects in the resulting diffraction pattern which should be included in the coronagraphic studies. Again, if the coronagraph is downstream of the modulator or switch, no harm should be done.

12.8 Appendix: Basic analysis of telescope polarization

Raytracing software such as ZEMAX can give full numerical information about what happens to the polarization of a celestial signal as it passes through the telescope and the rest of the optical system. This is a comparatively recent development, but some astronomical experience with it does exist. The end product of such an analysis is a prediction of what a backend full-Stokes polarimeter will measure, as a function of the polarization (4 Stokes parameters) of the wave-front entering the telescope.

Such analysis is time-consuming and will be carried out at a later stage, when a likely instrument configuration has emerged. At this point, one must content oneself with approximate results, strictly valid only for the paraxial approximation, idealized coatings and clean surfaces. For now, and indeed for much of a phase A study, that is sufficient.

12.8.1 Introduction

The OWL telescope introduces some instrument polarization which has to be taken into account. In principle, all reflections from an inclined surface produce some instrumental polarization. The incidence angle for the mirrors M1 – M6 is given in Table 12-1.

However, no instrumental polarization is introduced if a telescope system is rotationally symmetric. But, even a rotationally symmetric telescope may cause some depolarization (reduction of the degree-of-polarization signal).

For the introduced telescope polarization of OWL we must consider the following effects:

- M1 and M2 hexagonal structure
- M6, net polarization and field effects
- M1, M2, M3, M4, M5 depolarization
- M1, instrument polarization due to partially filled primary mirror

Mirror	F/# beam in	F/# beam out	Formula α(r) [rad]	Min α [deg] (r=0.3)	Max α [deg] (r=1)
M1	∞	1.4 (=F1)	r / (4·F1)	3.1	10.2
M2 flat	1.4	1.4	2·r / (4·F1)	6.1	20.5
M3	1.4	∞	r / (4·F1)	3.1	10.2
M4	∞	1.8 (=F4)	r / (4·F4)	2.4	8.0
M5	1.8	6.7 (=F5)	r / (4·F4) – r / (4·F5)	1.7	5.8
M6 flat	6.7	6.7	0.279* ± 2·r / (4·F5)	14.7 – 17.3	11.8 – 20.2

*: corresponds to a tilt angle of 16 degrees

Table 12-1: Estimated angle of incidence α(r) as function of the fractional radial position r on the mirror for the light rays falling onto the OWL telescope mirrors M1 – M6. All values are estimated using the paraxial approximation.

12.8.2 Mirror coating

According to the current 6-mirror OWL concept, we adopt for all calculations the expected polarization introduced by mirrors coated by silver. We adopt the refractive indices from ref. [13]. The exact refractive indices and the corresponding polarization effects depend on the protective layers properties on the actual mirror coating and the mirror aging due to dirt and tarnishing. Thus, the calculated effects for bulk silver yield only indicative values for the expected effects of a real OWL telescope.

Wavelength[μm]	0.3	0.4	0.5	0.6	0.7	1.0	2.0	5.0	10.0
Polarization [%]	11.7	1.07	0.37	0.22	0.17	0.12	0.11	0.09	0.11
Reflectivity	0.163	0.866	0.947	0.967	0.974	0.982	0.983	0.985	0.983

Table 12-2 Induced polarization and reflectivity vs. wavelength for a mirror with a silver coating for unpolarized incoming light and angle of incidence of 20 degrees (about the maximum incidence angle on M2).

12.8.3   Mirror M6

12.8.3.1 Net polarization

Mirror M6 produces a beam which is tilted by about 32 degrees from the telescope main optical axis. The angle of incidence is 16 degrees and the introduced instrumental polarization by mirror M6 is about 2/3 of the values given in Table 12-2 for 20 degrees inclination angle.

12.8.3.2 Field effects

The field dependence of the angle of incidence for mirror M6 can be neglected. Only objects close to the optical axis are imaged because the FoV is only of the order 10'' or less for a planet imaging camera.

12.8.4 Depolarization

A polarized signal from a source is slightly depolarized by the reflections from the rotational symmetric telescope mirrors M1, M2, M3, M4 and M5. The effect is however very small

only of the order 0.001 at a wavelength of 600 nm. This means that for a polarization signal of 1% from the target still 0.999% is in principle measurable after the mirrors M1 to M5.

12.8.5 Partially filled mirror M1

Different scenarios for filling the primary mirror can be considered, and also the procedure for re-coating of mirror segments is currently not defined. We consider here the impact of a partially filled mirror M1 on the instrument polarization (purely the geometric effect, neglecting diffraction)

First, let's assume the most extreme case, which is only one mirror segment of the M1 mirror at a location r=50m from the optical axis. Light collected from this single mirror would be affected by an instrument polarization of p = 0.39 % for a wavelength of 600 nm taking into account all reflections from mirrors M1 to M5.

At the same time we might also consider that the M1 mirror is symmetrically filled by about 3000 mirror segments except one which is missing at the outer rim of the M1 mirror. The effect on the total instrument polarization would be p = 0.39% / 3000.

Thus we can conclude that the expected instrument polarization due to the telescope mirrors M1 to M5 is much smaller than p=0.1% for all except very asymmetric primary mirror configuration. Typical values for the instrument polarization with a more or less rotational symmetric configuration due to all 5 mirrors M1 to M5 will be an order of magnitudes less (p << 0.01 %) than the instrument polarization introduced by mirror M6.

12.8.6 Summary and Conclusions

OWL is according to the current concept a telescope which will cause only little telescope polarization. The expected instrument polarization is essentially due to mirror M6 and amounts to about 0.24% at 500nm, 0.14% at 600nm and less for longer wavelengths. If a polarization switch can be implemented at the intermediate focus (see section 12.9), the influence of M6 can be largely eliminated.

Important for this low instrument polarization of OWL are two points:

- the very good polarimetric properties of the anticipated silver coatings to be used for all telescope mirrors,
- the small angle of incidence (or small tilt angle) of mirror M6.

From the polarimetric point of view it is important to retain or even improve these key properties of the OWL concept during the future evolution of the telescope design.

12.9 Appendix: A polarization switch at OWL secondary focus ?

Jaap Tinbergen and Ramon Navarro

In this section we explore the requirements that must be met by particular implementations of a moderately achromatic polarization switch for the intermediate (F/2) focus of the 100-m OWL design. Such an implementation would either be a stress-birefringence Pancharatnam combination or a rotating birefringent-crystal Pancharatnam combination. Both would consist of zero-order waveplates (usually halfwaves) in view of the extremely fast focal ratio and both would have to rotate to some extent; each implementation has its advantages and both would allow switching the sky polarization, which at that point has only been corrupted by the on-axis part of the telescope.
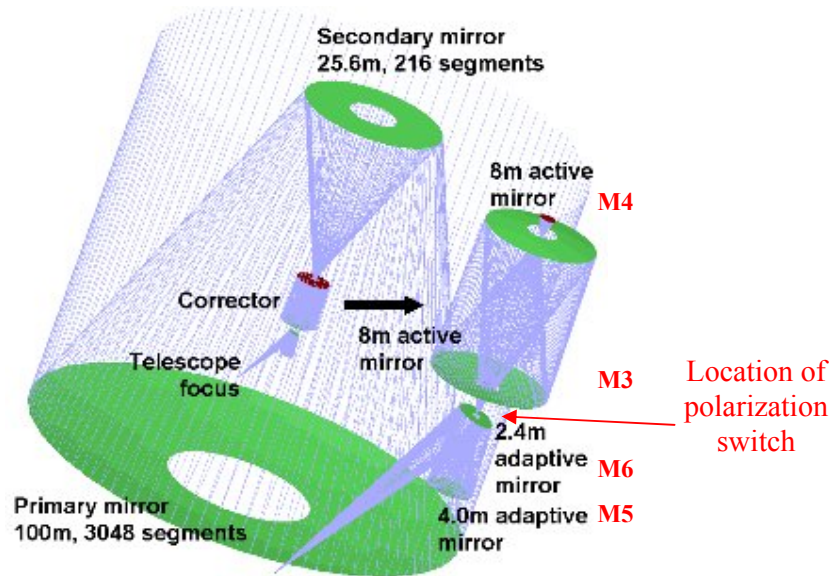
Figure 12-10 The main optics of the OWL telescope. The system consisting of M3 to M6 is enlarged. The proposed polarization switch is at the secondary ('intermediate', F/2) focus, located just behind, or even inside the central aperture of M6. The switch unit will be attached to the M6 support structure.
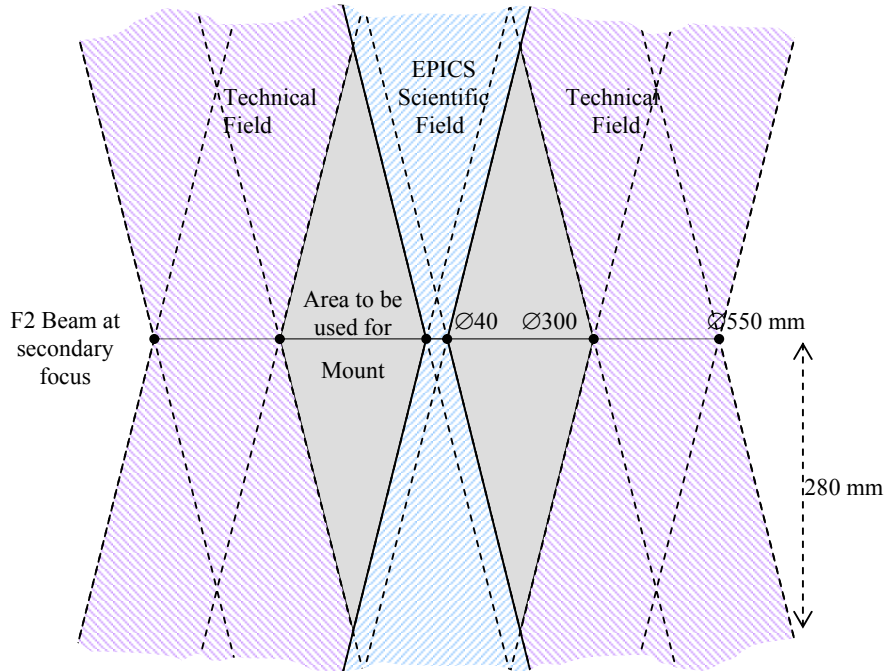
Figure 12-11: **'Vertical' cross section of the light paths at secondary focus. The beams for the scientific and the technical fields are indicated, including diameter dimensions. Vignetting is not allowed in the scientific field (blue), nor in the technical field (violet), restricting the region that can be used for mounting polarization optics to the grey volume.**

Figure 12-12: Focal-plane geometry at OWL intermediate focus. Vignetting is not allowed in the EPICS scientific field (blue), nor in the technical field (violet), restricting the space for supporting the polarization optics to the grey area. However, also the technical field should not be vignetted, so that the final area for the mount is limited to locations where the guide star pickup arms cannot reach. This does not appear to be impossible (Gilmozzi, private communication).

12.9.1 The stress-birefringence Pancharatnam configuration

We first consider the stress-birefringence version of the proposed switch, because it is potentially the more powerful and versatile. It is also the more complicated option, needing much more careful design and interfacing.

The stress-birefringence Pancharatnam configuration of 3 on/off halfwave retarders is based on the mechanism of stressing a slab of isotropic optical material, usually fused silica. By rotating the entire assembly one may switch linear polarization, while switching the retarders on and off will switch the circular polarization. The Pancharatnam configuration is achromatic over a certain band (Figure 12-16). Stress-birefringence retarders are operated as zero-order waveplates, so that they are likely to function better under the fast beam conditions than would achromatic 2-material combinations. By adjusting the orientation of the central retarder w.r.t. the outer 2, one may trade off achromatic quality against achromatic bandwidth, while one may tune the central wavelength by suitably adjusting the retardance values. If this kind of unit can be designed and constructed, it will confer unique properties on the EPICS instrument in its high-precision polarimetry mode: the astronomical polarization signal, corrupted only by the rotationally-symmetric part of the OWL telescope, may thus be labeled, to distinguish it from all polarization effects introduced by tilted mirrors, AO, pupil or image rotators, atmospheric-dispersion corrector, dichroics and other instrumental modules further downstream.

Figure 12-13: Plan and elevation (sketches) of a Pancharatnam stress-birefringence switch: schematic optics only; steel force frame and actuators are assumed and are as yet undefined. Each slab of fused silica can be stressed to provide halfwave retardation (at the passband central wavelength), under servo control (retardance sensor). Pancharatnam's 3-retarder configuration is achromatic over a certain bandwidth. By changing the orientation of the central slab slightly (55°-60°), one may trade achromatic range for achromatic quality (Figure 12-16). The small prisms for the servo beam act as quarterwave retarders at position angles of ±45° with respect to the birefringence axis of the slab; the servo is explained in the text.

Figure 12-14: The stress-birefringence Pancharatnam combination in elevation, including field lenses and the piezo actuators but excluding the force frames. Stress direction is indicated by the red arrows. Servo control requires the birefringence to be measured, individually for each plate. The total thickness of this stack is > 3 cm, having a significant impact on the chromatic effects in the scientific focus (

Table 12-3**). The field lenses could be convex or concave, in this figure they are 'placeholders' for as yet unspecified components to counteract chromatic effects in the scientific focus, and/or manipulate beam focal ratio through the birefringent slabs. The small prisms are quarterwave retarders at position angles of ±45o with respect to the birefringent stress direction.**

Plate 1

Plate 2

Plate 3

Rotated

55° - 60°

Figure 12-15: Top view of a stress birefringence Pancharatnam combination including field lenses and the piezo actuators, but excluding the force frame. The rotation of the middle plate can be adjusted slightly in order to select a suitable wavelength-bandwidth combination. Servo control requires the birefringence to be measured, individually for each plate. The whole combination needs to rotate over angles 0° 22.5° 90° and 112.5° in order to measure all Stokes parameters; more detail in the text. A steel outer frame is needed to apply the necessary force to each fused-silica plate. The ratio of E-modulus of steel (200·109 N/m2) to that of fused silica (70·109 N/m2) is about a factor of three. This indicates that, when a force is applied, a very modest steel frame with the same cross section of 5 cm2 per plate, expands about one third in respect to the compression of the silica. Taking into account the high strength and fracture resistivity of steel versus silica, designing a force frame is not an issue.

Wavelength (um)	Defocus (mm)	Spot size (um) at 50.35 mm defocus
0.5	51.25	160
1.0	50.35	70
1.8	49.60	180

Table 12-3: Defocus and spot size at scientific focus due to a 30 mm thick fused silica plate at secondary focus.

Fig. 5. Performance of Pancharatnam achromatic half-wave and quarter-wave retarders. The abscissa Φ is relative retardation. The ordinates are retardation (2Δ) and position angle of the fast axis (c'_1), of the combination. The numbers denote Pancharatnam f-values. The zero of position angle is in each case defined as the position angle for relative retardation $(1 \pm f)$. "D" denotes Dollfus' (1963) half-wave combination

Figure 12-16: The Pancharatnam achromatic waveplate. Left: the halfwave retarder. All 3 components of the Pancharatnam halfwave are themselves halfwave at the wavelength(s) of mean relative

retardation. **The halfwave is a special case in that the curves are symmetrical, which means that the halfwave is a particularly good polarization component. From the different curves one may see how achromatic range may be traded for quality of achromatization.** T**he abscissa is relative retardation (normalized to the mean value of the range in use), it** can best be thought of as normalized inverse wavelength. **Figure from ref. [38].**

Operation of a stress-birefringence Pancharatnam switch

The suggested mode of operation is to switch the 3 stressplates on/off in synchronism at 0.1-0.01 Hz and to record the "AC" degree-of-polarization signal for 4 orientations of the complete assembly: combining 2 of these will yield q=Q/I, the other pair will yield u=U/I. This eliminates (as "DC", or constant polarization) all effects due to misalignment, vignetting etc, downstream of the switch; "AC", the switched polarization, is the result of true (sky, as modified by the telescope) polarization only. A similar scheme yields v=V/I. This is worked out in somewhat more detail in the rest of this paragraph.

A halfwave retarder (the "on" state) at orientation θ will have the Mueller matrix:

$$M_{halfwave}\left(\theta\right)=\left[\mathbf{R}\left(-\theta\right)\right]\times\left[M_{halfwave}\left(0\right)\right]\times\left[\mathbf{R}\left(\theta\right)\right]$$

$$=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \left(c^2-s^2\right).h_{22}-2cs.h_{23} & \left(c^2-s^2\right).h_{23}-2cs.h_{22} & -\left(c.h_{42}-s.h_{43}\right) \\ 0 & \left(c^2-s^2\right).h_{23}-2cs.h_{22} & -\left(c^2-s^2\right).h_{22}+2cs.h_{23} & -\left(s.h_{42}+c.h_{43}\right) \\ 0 & +\left(c.h_{42}-s.h_{43}\right) & +\left(s.h_{42}+c.h_{43}\right) & h_{44} \end{bmatrix}$$

where $c=\cos 2\theta, s=\sin 2\theta, c^2-s^2=\cos 4\theta, 2cs=\sin 4\theta$ and h_{lm} are the elements of the Mueller matrix of a halfwave retarder in principal orientation. $\mathbf{R}\left(\theta\right)$ represents rotation of the coordinate system through an angle θ.

Subtract the unit Mueller matrix (for the "off" state) to obtain the "AC" Mueller matrix (representing the action of the switch):

$$M_{switch} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & h_{22}\cos 4\theta - h_{23}\sin 4\theta - 1 & h_{23}\cos 4\theta - h_{22}\sin 4\theta & -(ch_{42} - sh_{43}) \\ 0 & h_{23}\cos 4\theta - h_{22}\sin 4\theta & -h_{22}\cos 4\theta + h_{23}\sin 4\theta - 1 & -(sh_{42} + ch_{43}) \\ 0 & ch_{42} - sh_{43} & sh_{42} + ch_{43} & h_{44} - 1 \end{bmatrix}$$

Note that m_{44} is constant ($\approx$ -2), m_{42}, m_{43}, m_{24} and m_{34} are periodic in $2\theta,$ whereas $m_{22}, m_{33},$ m_{23} and m_{32} are periodic in $4\theta.$ This fact allows us to disentangle circular and linear polarization and to reject error terms due to the halfwave retarder being imperfect:

- Use the (AO train +) polarimeter to measure $(q,u)_{out}$ at θ = x and x+90°, and at x+22.5^o and x+112.5^o. This yields q_{in} and u_{in} (4θ -periodic) and eliminates v_{in} (2θ - periodic).
- Similarly measure v_{out}. Average over θ = x and x+90°.This yields v_{in} (constant) and eliminates q_{in} and u_{in} (2θ -periodic).

In practice the retardation will only be exactly halfwave at 3 spot wavelengths and only for certain ray directions; the deviations from halfwave will have to be evaluated (partly by ZEMAX) and averaged over the passband and F/2 beam (taking account of partial M1 filling) to obtain the effective Mueller matrix of the polarization switch. If from nothing else, it is clear from this example that full-polarization ZEMAX or equivalent is essential for EPICS.

Construction requirements for a stress-birefringence Pancharatnam switch

To recapitulate: in constructing a polarization switch of this kind at the OWL intermediate focus, something like the following constraints will apply; at this point, we don't guarantee that a solution is possible:

- Vignetting in the technical field to be minimal. Therefore the scientific field must be small (it is !) and the entire unit must be suspended on a structure in parts of the field not going to be used by star trackers.
- Outer diameter of unit limited by the vignetting condition.
- Entire unit to be inserted from the rear of M6 and removed from the beam when not needed. If astronomical *circular* polarization is to be detected, this action should not take more than about 5% of the length of the switching cycle.
- Entire optical unit rotatable, at least to positions 0°, 22.5^o, 90° and 112.5^o. This action should not take more than about 5% of the length of the switching cycle.
- Outer 2 slabs coaligned, inner slab at relative orientation ('azimuth') 55°-60°, adjustable (for trade-off of bandwidth vs achromatic quality).
- 3 identical units, each consisting of:

1) Outer frame, to allow squeezing the optical slab
2) Optical slab
3) (Piezo?) actuator within force frame
4) Auxiliary optics (with light-source, detector and remotely-rotatable polaroid) to operate the retardance servo; see next section
5) Optical aperture (unvignetted and with uniform stress): a few arcseconds plus imperfect focus by OWL $\equiv$ 4 cm approximately

- Possibly specialized lenses, to condition the beam through the slabs assembly and preserve scientific-focus achromatism for a small on-axis field and restricted wavelength range (an achromatic focus shift is allowed in principle, could be compensated by moving a mirror elsewhere).

A servo-stabilized switched halfwave retarder

A switched halfwave retarder (of which a Pancharatnam switch has 3, operated in synchronism) has 2 states, zero and halfwave. Both of these must be stable under changing external conditions (temperature, gravity). A servo is required and this servo must stabilize (therefore measure) the retardance itself rather than, for instance, the applied mechanical stress. The present section outlines such a servo, for the half-wave state. **The zero-retardance state may perhaps be obtained by releasing all mechanical stresses. However, practical implementation of this may be difficult and it may be necessary to use a servo-stabilized low-retardance state (rather than precisely zero) crossed with a low-retardance fixed compensating retarder. That is a relatively minor detail, but may have a bearing on achromatization.**

To our knowledge, there is no 'retardance-meter' simple enough to incorporate into an optical component. We therefore propose to use a theorem stated by Jones in one of his papers on the polarization calculus that bears his name (see [39], the theorem is discussed in the last section before the final Conclusion): *"...... it is possible to produce an optical system which behaves rigorously as a (polarization) rotator .... by inserting a retardation plate between two crossed quarter-wave plates in such a position that its axes are at 45° with respect to the axes of the quarter-wave plates; ..... the angle of the rotation is one-half of the phase retardation of the inserted plate..... According to Billet, ... (this) was known to Fresnel.".* By multiplying out the matrices of Jones' equation 28, we may confirm the identity in that equation. However, since in astronomy we are dealing with partially polarized radiation, we prefer to present the analogous Mueller matrices for a retarder (of retardance γ) at orientation 45° between quarter-waves at orientations 0° and 90°.

$$[M] = \left[Ret(\lambda/4) \right] . \left[Rot(+45^0) \right] . \left[Ret(\gamma) \right] . \left[Rot(-45^0) \right] . \left[Ret(-\lambda/4) \right]$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\gamma & \sin\gamma \\ 0 & 0 & -\sin\gamma & \cos\gamma \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\gamma & \sin\gamma & 0 \\ 0 & -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \equiv \left[Rot(\gamma/2) \right]$$

This theorem tells us that, to measure the retardation of the stressed slab of fused silica, we have to place it between crossed quarter-waves and measure the rotation of the plane of polarization, which is something we can do by relatively simple means. If the quarter-waves are implemented as right-angled prisms [40], the auxiliary light path can enter and leave from the side, an almost certain advantage in construction and in eliminating stray light. We can use any wavelength that is convenient, preferably outside the instrumental bandwidth; from the known wavelength dependence of the slab's birefringence we can compute what we need to obtain halfwave retardation at our target wavelength

To determine the retardance, we have to measure the rotation of the plane of polarization and compare this with the desired value, to yield the error signal for the servo . There are various ways of doing this using 2 polaroids, such as:

1) rotating a polaroid at the input or output and determining the phase of the sinusoidal component of the output intensity; this uses a single detector, but requires continuous rotation of a polaroid

2) nulling the output intensity by rotating the input polaroid; a cosine-like null is not a good way to operate a servo, we have to convert to a sine-like passage through zero (this may be done by using 2 output channels, with polaroids offset either way from the desired position and subtracting; however, this makes us dependent on detector gain ratio)

3) switching between 2 input polarizations by a simple modulator and looking for the sine-like passage through zero of the AC component of the output signal. This is the most sophisticated option, probably the best of the 3; it uses a single detector and a single polarization of the light striking it.

To be able to program the retardance value, we have to be able to control the rotation, therefore the orientation of either the input or output polaroid; this can be a simple

mechanism, since the value stays constant during an observation (unless switching passbands between, for example, red and IR as part of the observing scheme)

12.9.2 Crystal-plate Pancharatnam polarization switch

If the chromatic effects of the triple silica stressplate turn out to be prohibitive and cannot be compensated, we shall have to fall back on the alternative of zero-order crystal waveplates (e.g. quartz or sapphire), which are usually cemented components and cannot be tuned. For different applications, we shall have to have different versions in a magazine, with a robot picking one or the other and placing it in the rotation mechanism.

We do not elaborate on this alternative, which we consider less attractive but more or less straightforward. It is illustrated in Figure 12-17 and Figure 12-18.
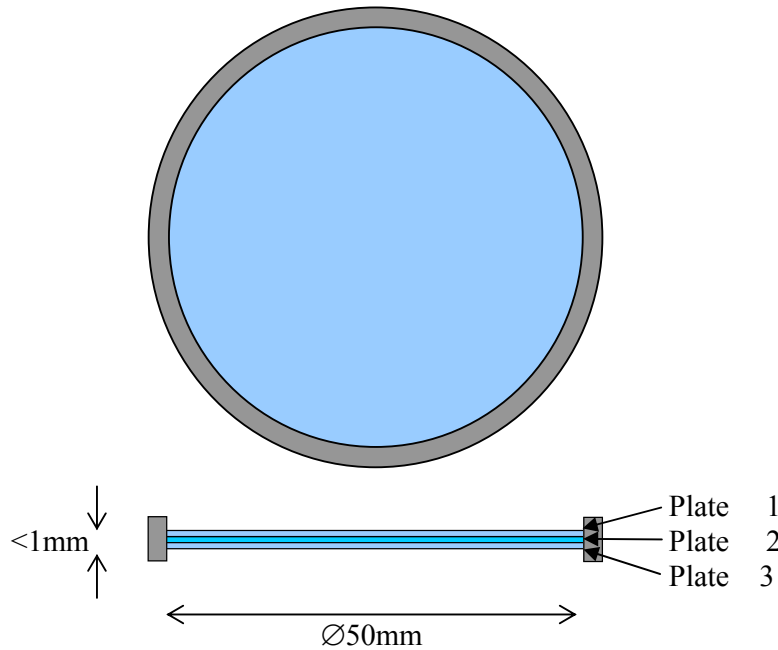
Figure 12-17: As an alternative to the stress-birefringence switch, a zero-order birefringent-crystal Pancharatnam halfwave combination is very thin, limiting the chromatic effects in the scientific focus.

Figure 12-18: Very schematic: a robot for exchange of a selection of birefringent-crystal zero-order Pancharatnam combinations and rotating them inside the scientific field. The mount should not vignette the technical field in the areas that can be reached by the guide star pickup arms.

12.9.3 Conclusion

If the chromatic path length of 3 slabs of fused silica can be compensated in the design somewhere, the stress-birefringence Pancharatnam halfwave polarization switch would be a flexible facility to have at OWL's intermediate focus. It would relax considerably the requirements on other parts of the polarimetric system and it would enable simple (imaging, spectro-)polarimetry by other photometric backend subsystems (imagers, spectrometers). If the chromatic path length difference is a showstopper, the crystal-plate version of the switch offers a viable alternative.

13 EPICS performance analysis

This section presents an analysis of the performance of EPICS in detecting extrasolar terrestrial planets and characterize them by their main prominent features. The AO corrected PSF after correction by a "perfect" coronagraph (see section 8.4.2) feeds a differential spectral or differential polarimetric imager that reduces speckle noise. In the first subsection, the pixels-size required for efficient wavelength re-scaling and image re-centering is determined. Then, various spectral features of the Earth are analyzed with respect to their detectability from the ground when present in an exo-Earth. Hereby, a focus has been put on the exploitation of the Doppler-shift of molecular bands of the exo-Earth moving at some tens of kilometers per second with respect to Earth. Finally, exposure times, required to detect a spectral or polarimetric signature of an Earth-like planet in the habitable zone around G2V, K2V, and M2V stars at distances between 10 pc and 25 pc and an SNR of 5, are plotted in section 13.3.

13.1 Assumptions

Although sounding overly optimistic, a **"perfect" coronagraph** isn't actually such a strong assumption. Since EPICS' discovery space starts at a relatively large separation in terms of λ/D (50 mas correspond to about 20 λ/D at 1.2 µm wavelength), and coronagraphs can be tailored to deliver nearly perfect performance over a restricted range of angular separation with not so small inner working angle, the assumption of a "perfect" coronagraph is reasonable.

As shown in section 8.4, the **AO PSF** contrast in the corrected area is limited mainly by temporal errors, errors introduced by chromaticity of the turbulent wave fronts and WFS noise. It is possible to correct for the chromatic errors by either operating the WFS at the same wavelength as the science camera, or by modifying the AO control law in a way that the chromatism of the wave fronts is taken into account. The latter option is preferable since it does not lead to a loss of photons for the science instrument. The control law has been derived and the concept will be tested by lab experiment. It is worth noting that the other limitations (temporal and noise) are fundamentally set by guide star brightness and inherent to all kind of WFS including focal plane interferometers.

Sky and detector noise has been neglected in this analysis. Given that EPICS will only observe bright (mag < 10) stars, noise will always be dominated by residual stellar light, so sky and detector noise can safely be neglected.

Systematic errors such as static aberrations (common and differential, before and after the coronagraph) leading to slowly evolving semi-static speckles have been neglected. Our approach has been to very carefully analyze the systematic errors and to develop a concept (see section 7) that reduces them to a negligible level. Therefore, the reported numbers for the detectability of spectral features are representative for spectral differential imaging either with an SDI type device or with an IFS.

Speckle noise is removed by either spectral differential or polarimetric differential imaging. Hereby, the relatively low number of photons per speckle at the angular distances of interest (~ few tens of photons per speckle for a 6 magnitude star assuming a speckle lifetime of 10

ms) ensures that simple subtraction of two wavelengths provide sufficient speckle noise reduction for the spectral differential imaging. If the number of photons per speckle would be much larger than this, three-wavelength subtraction would be required to reduce chromatic residuals.

13.2 Pixel scale and interpolation error

A speckle image was simulated by using a high order AO corrected phase screen and imaging it through a 4 quadrant phase mask coronagraph at a high (8 pixels per FWHM) resolution. Subimages were cut out of this large image with shifts of a couple of pixels. Therefore these subimages were identical but displaced by some pixels. Then, the subimages were binned to the desired resolution (2x2 for 4 pixels / FWHM, 3x3 for 8/3=2.67 pixels / FWHM, 4x4 for 2 pixels / FWHM and 8x8 for 1 pixels / FWHM). Finally the final subimages were subpixel shifted with respect to each other using a spline interpolation algorithm provided by MATLAB and subtracted from each other.

Results

Figure 13-1 and Figure 13-2 display the result of this experiment. Essentially the speckles are suppressed by the factors listed in Table 13-1. Although the spline interpolation might not be optimum, other interpolation algorithms (linear, cubic) performed worse.

The results indicate that just Nyquist sampled imaging would hardly provide the required speckle subtraction factors. A sampling of 2.5-3 pixel per FWHM appears necessary in order to achieve adequate interpolation performance for two-wavelength differential imaging with our wavelength separation (see section 10).

Table 13-1. Speckle suppression limited by interpolation.

Pixelscale	Imax raw image	Imax diff. image	Suppression
1 / FWHM	~250000	~90000	~2.8
2 / FWHM	~80000	~4000	~20
2.67 / FWHM	~45000	~400	~112
4 / FWHM	~22000	~35	~630

Figure 13-1. Images at 4 pixels / FWHM (left column) and 2.67 (not 3 as indicated in the pictures) pixels / FWHM (right column). The images in the top and in the middle row are identical but slightly displaced with respect to each other on the detector grid. The bottom row shows the difference between the two after proper subpixel shifting using a MATLAB spline interpolation algorithm.

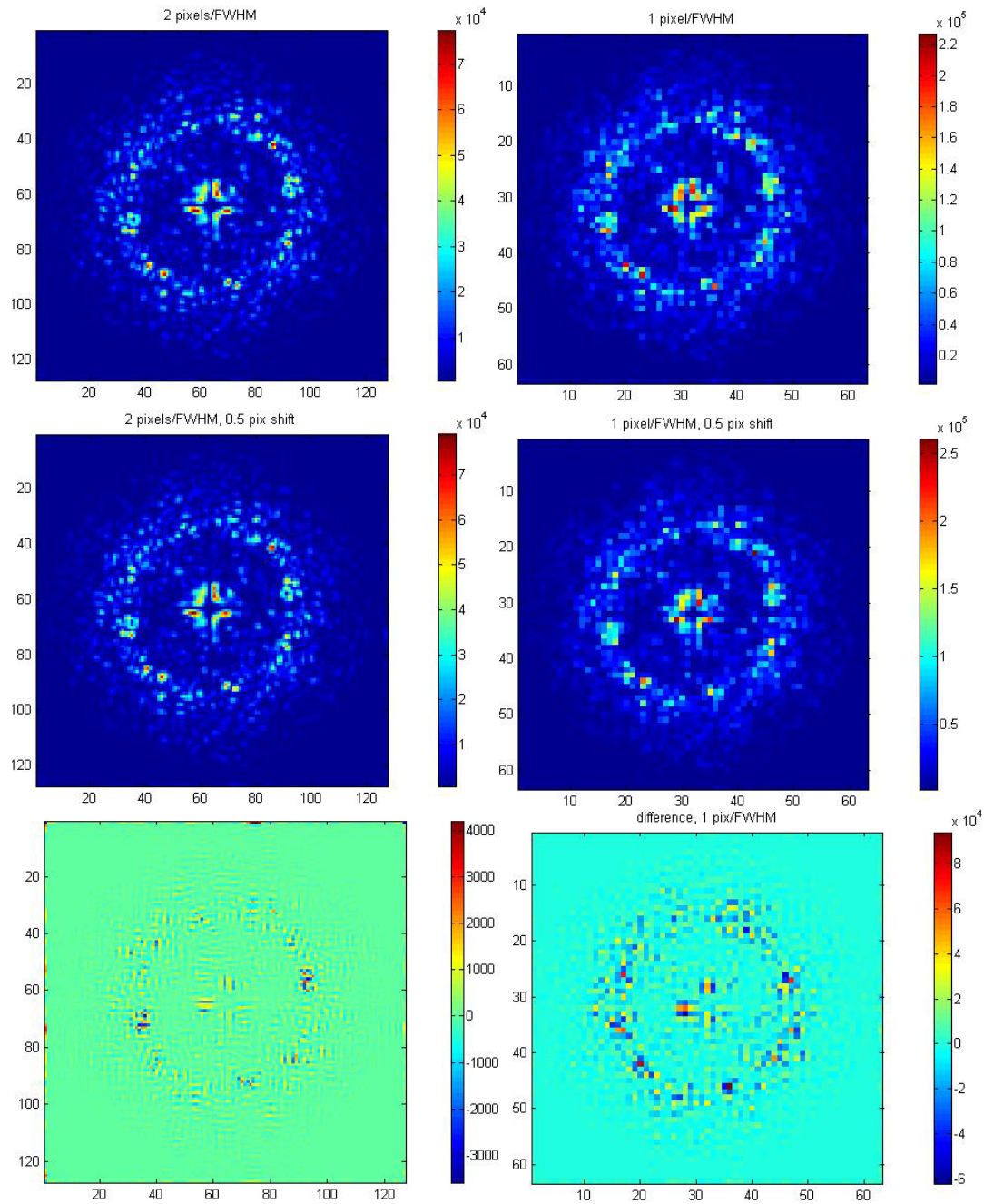**Figure 13-2. Images at 2 pixels / FWHM (left column) and 1 pixels / FWHM (right column). The images in the top and in the middle row are identical but slightly displaced with respect to each other on the detector grid. The bottom row shows the difference between the two after proper subpixel shifting using a MATLAB spline interpolation algorithm.**

13.3 Detection of exo-Earths in O_2, CO_2, CH_4 and H_2O bands

This section reports on the analysis of the dual-imaging contrast of an Exo-Earth's O_2 and H_2O achievable while observing through our own atmosphere.

The Exo-Earth is assumed to have an atmosphere of the same composition as ours and is moving at speeds between -50 km/s and +50 km/s relative to us along the line of sight corresponding to typical orbital motion (± ~40km/s) and proper motion (± ~10 km/s) for neighboring stars. The stellar spectrum is assumed to be featureless. Atmospheric transmissions were calculated with the HITRAN software.

Cloud coverage would only reduce the contrast in the spectral significantly in case of high Cirrus clouds where most of the light would be reflected from above the atmosphere and would not show the atmospheric spectral signatures.

13.3.1 O_2

The following Figure 13-3 shows a simulated atmospheric transmission for assuming and observatory in 4 km altitude and observation at zenith. The assumed atmosphere is a mid-latitude winter-model one. The resolution of the simulated spectrum is R ≈ 500.000.

The spectrum of the Exo-Earth was calculated by raising this transmission spectrum to the 4th power, because the O_2 column density is about 4 times higher when traveling all the way down and up through the atmosphere than just down to 4000 m.
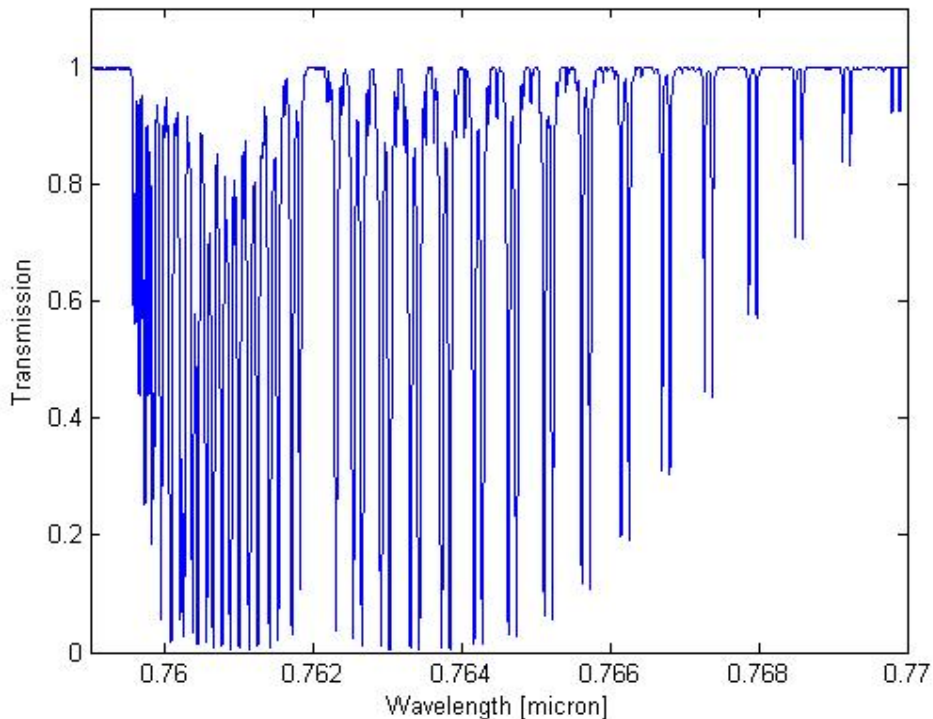

Figure 13-3. Simulated atmospheric transmission in the A-band of O_2.

In order to evaluate the effect of Earth's atmospheric transmission on the achievable contrast of differential imaging using Earth-like atmospheric features, the following contrast ratio was calculated

$$c = \sum_{\lambda=759nm}^{770nm} F_{ee,\lambda} \cdot A_{\lambda} \bigg/ \sum_{\lambda=759nm}^{770nm} F_{s,\lambda} \cdot A_{\lambda} \ ,$$

with $F_{ee,\lambda}$ being the spectrum of the exo-Earth, $F_{s,\lambda}$ being the spectrum of the star (here = 1) and A_{λ} being the atmospheric transmission.

The following Figure 13-4 plots c as a function of the relative radial velocity between Earth and exo-Earth. In the worst case of $\Delta v = 0$ km/s, the O_2 band of Earth and exo-Earth are superimposed and the contrast is 0.79. When the band is Doppler shifted the situation becomes more favorable. The ultimate contrast for $A_{\lambda} = 1$ (in space) is $c = 0.67$.

The average atmospheric transmission over the O_2 band is about 0.81.
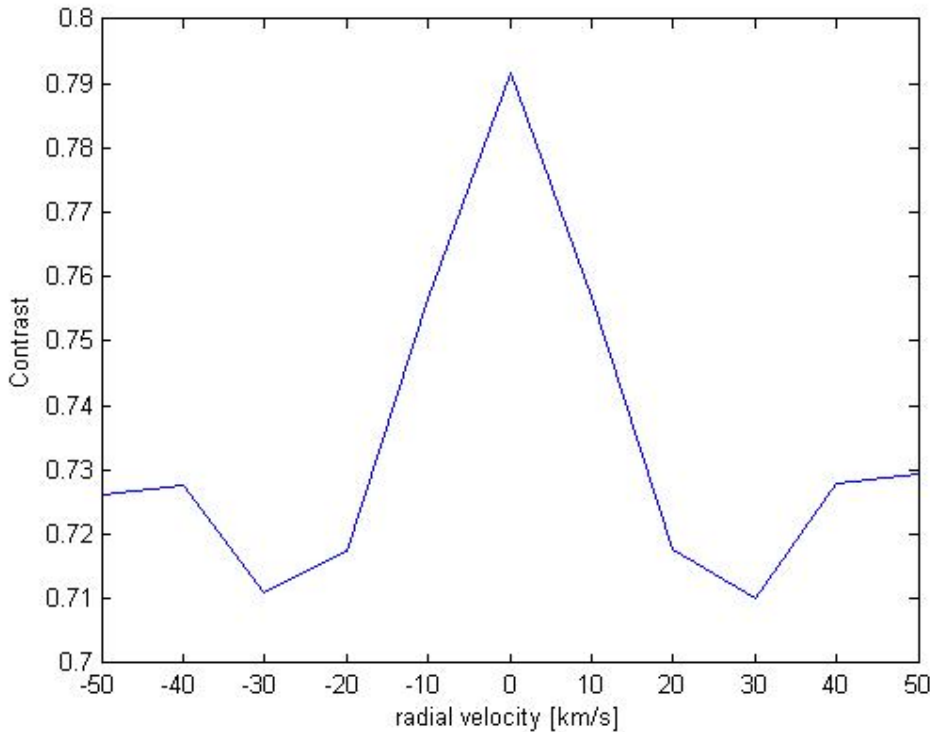

Figure 13-4. O_2 contrast (normalized planet brightness / normalized star brightness) as a function of the relative radial velocity between Earth and exo-Earth. The optimum in case of no atmospheric O_2 absorption would be 0.67.

13.3.2 H₂O

The following Figure 13-3 shows a simulated atmospheric transmission for assuming an observatory in 4 km altitude and observation at zenith. The assumed atmosphere is a mid-latitude winter-model one. The resolution of the simulated spectrum is $R \approx 500.000$.

The spectrum of the Exo-Earth was calculated by raising this transmission spectrum to the 4th power, because the H_2O column density is about 4 times higher when traveling all the way down and up through the atmosphere than just down to 4000 m. The resulting spectrum, smoothed to R=150, was compared with the atmospheric transmission on Mauna Kea (http://www.jach.hawaii.edu/UKIRT/astronomy/utils/index1.html) and found to be very similar.
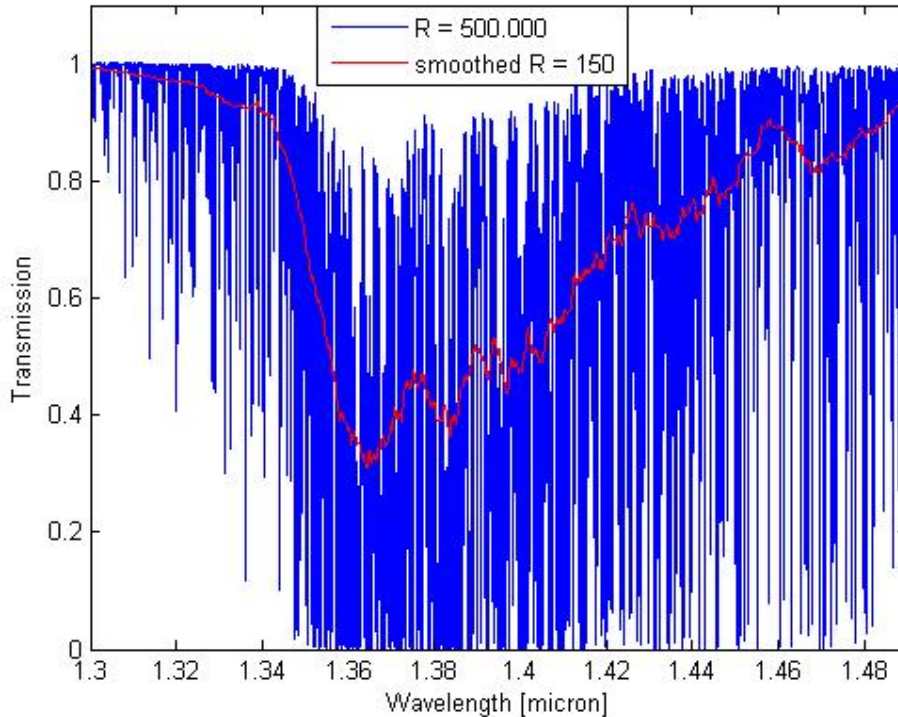
Figure 13-5. *Simulated atmospheric transmission of the H₂O separating the J and H astronomical bands*

In order to evaluate the effect of Earth's atmospheric transmission on the achievable contrast of differential imaging using Earth-like atmospheric features, the following contrast ratio was calculated

$$c = \sum_{\lambda=1350nm}^{1400nm} F_{ee,\lambda} \cdot A_\lambda \Bigg/ \sum_{\lambda=1350nm}^{1400nm} F_{s,\lambda} \cdot A_\lambda \,,$$

with $F_{ee,\lambda}$ being the spectrum of the exo-Earth, $F_{s,\lambda}$ being the spectrum of the star (here = 1) and A_{λ} being the atmospheric transmission.

The following Figure 13-6 plots c as a function of the relative radial velocity between Earth and exo-Earth. In the worst case of $\Delta v = 0$ km/s, the H_2O band of Earth and exo-Earth are superimposed and the contrast is 0.32. When the band is Doppler shifted the situation becomes more favorable. The ultimate contrast for $A_{\lambda} = 1$ (in space) is $c = 0.17$.

The average atmospheric transmission over the H_2O band is about 0.44.
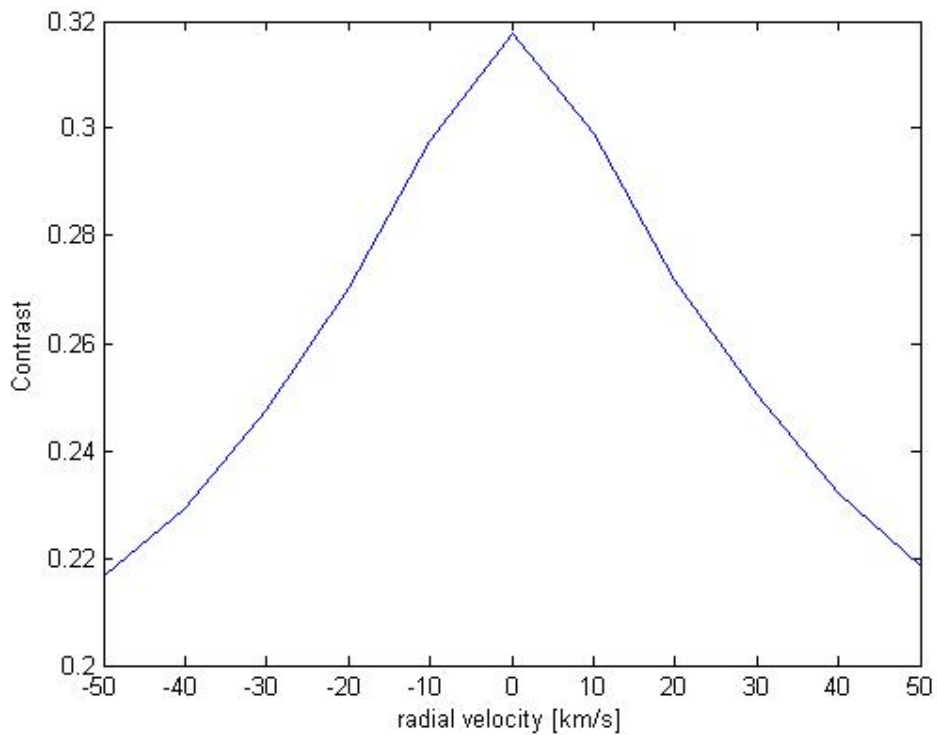
Figure 13-6. H_2O contrast (normalized planet brightness / normalized star brightness) as a function of the relative radial velocity between Earth and exo-Earth. The optimum in case of no atmospheric H_2O absorption would be 0.17.

13.3.3 CH_4

The following Figure 13-7 shows a simulated atmospheric spectrum of a Methane-rich (1000 ppmv) Exo-Earth calculated by raising a calculated transmission spectrum (4000 m height) to the 4th power. The resolution of the simulated spectrum is $R \approx 500.000$. Since the Methane concentration on Earth nowadays is just about 2 ppmv, atmospheric transmission and stellar spectrum were assumed to be featureless.

Hence, the contrast is given by

$$c = \sum_{\lambda=1630nm}^{1700nm} F_{ee,\lambda} \cdot A_\lambda \Bigg/ \sum_{\lambda=1630nm}^{1700nm} F_{s,\lambda} \cdot A_\lambda$$

with $F_{ee,\lambda}$ being the spectrum of the exo-Earth (Figure 13-7), $F_{s,\lambda}$ being the spectrum of the star (here = 1) and A_λ being the atmospheric transmission (here = 1).

The achievable contrast is 0,78. Already in the transmission spectrum at 4000 m height, the CH_4 lines are mostly saturated. Therefore raising it to the 4th power did not greatly improve the achievable contrast. Therefore, even higher concentrations of CH_4 would not greatly improve the situation.
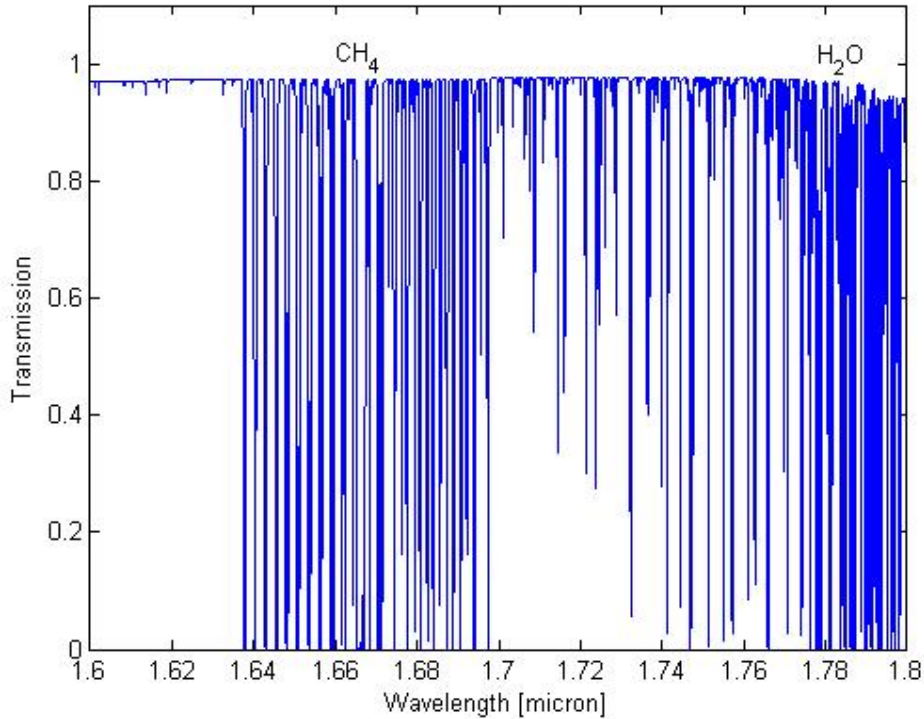
Figure 13-7. *Simulated spectrum of a Methane-rich (1000 ppmv) Exo-Earth $_4$ in H-band*

13.3.4 CO_2

The following Figure 13-8 shows a simulated atmospheric spectrum of a CO_2-rich (10%) Exo-Earth calculated by raising a calculated transmission spectrum (4000 m height) to the 4th power. The resolution of the simulated spectrum is R ≈ 500.000. Since the CO_2 on Earth nowadays is just about 370 ppmv, atmospheric transmission and stellar spectrum were assumed to be featureless.

Hence, the contrast is given by

$$c = \sum_{\lambda=1560nm}^{1660nm} F_{ee,\lambda} \cdot A_\lambda \Bigg/ \sum_{\lambda=1560nm}^{1660nm} F_{s,\lambda} \cdot A_\lambda$$

with $F_{ee,\lambda}$ being the spectrum of the exo-Earth (Figure 13-7), $F_{s,\lambda}$ being the spectrum of the star (here = 1) and A_λ being the atmospheric transmission (here = 1).

The achievable contrast is 0.43 (10% concentration) and 0.30 (50% concentration).
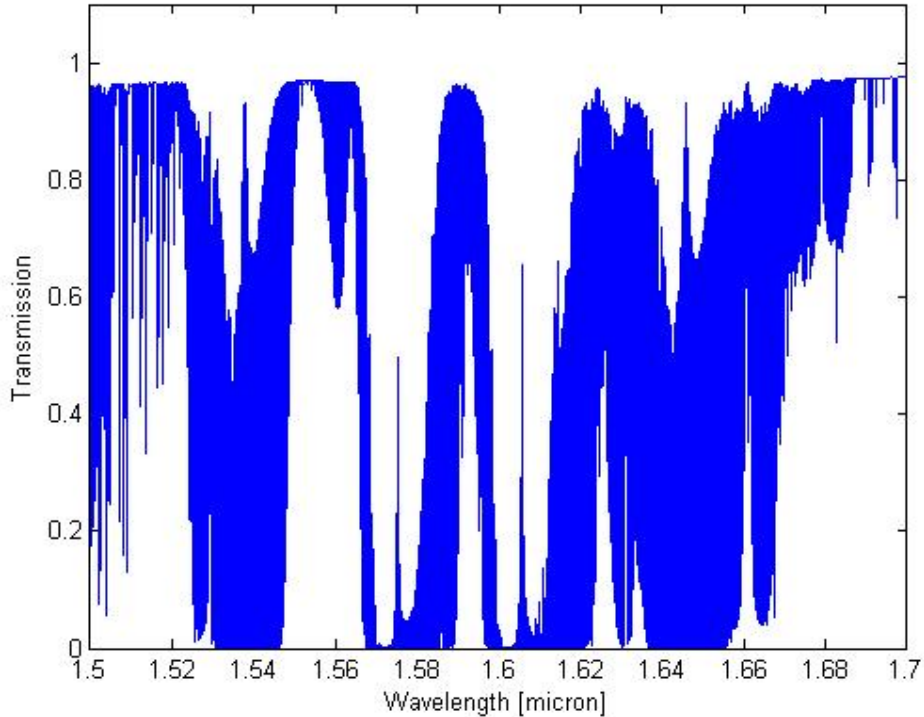
Figure 13-8. Simulated spectrum of a CO_2-rich (10 percent) Exo-Earth in H-band

13.3.5 Conclusions

The main conclusion of this sections is that **differential imaging and detection of atmospheric features of an exo-Earth through our atmosphere is possible**. Especially the detection of water vapor appears very promising. The Doppler shift of the molecular bands of Earth and exo-Earth due to their differential motion improves the contrast but not by a great deal.

13.4 Detectability and SNR

This section reports the exposure times required to detect various spectral features of an Exo-Earth under the assumptions stated in section 13.1.

13.4.1 Signal and noise

The star

In the blackbody approximation, the signal N_* (in photo-electrons) we receive from a star with a radius R_* at a distance D_*, observing at a wavelength λ, integrating during a time t in a Δ_λ bandwidth with a telescope having a diameter d is:

$$N_* = \Gamma \left(\frac{R_*}{D_*}\right)^2 \frac{\pi c}{2\lambda^4} d^2 \Delta_\lambda t \frac{1}{e^{\frac{hc}{\lambda k_b T_*}} - 1},$$

where Γ is the telescope's total transmission and T_* the star's effective temperature.

Moreover, the star is supposed to be on the main Sequence which means that we have a relation between its radius and its temperature:

$$R_* = \left(\frac{T_*}{5790K}\right)^{\frac{3.2}{1.9}}$$

in solar radius unit.

The planet
The planet is supposed to reflect the light it receives from the star. Moreover, we have to take into account it phase angle φ as well as its orbit obliquity i (when $i = 90°$ the planet is edge-on). So, the number of photo-electrons we receive from the planet is:

$$N_p(\varphi, i) = A_G \left(\frac{R_p}{a}\right)^2 \left[\frac{\sin(\alpha) + (\pi - \alpha)\cos(\alpha)}{\pi}\right] N_*,$$

where A_G is the planet's geometric albedo, supposed achromatic (0.367 for an Earthlike planet and 0.52 for a Jupiter-like planet), R_p is the planet's radius, a the orbit's semi-major axis and α the angle such as:

$$\cos(\alpha) = -\sin(i)\cos(\varphi).$$

Moreover, planets are supposed to be in thermodynamical equilibrium with the star so:

$$a = a_{ss} \left(\frac{T_*}{5790K}\right)^{\frac{7}{1.9}} \quad \text{in AU,}$$

where a_{ss} is the Sun/Planet distance in our own Solar System (1AU for the Earth and 5.2AU for Jupiter).

As an example, the Star/Earthlike contrast (for $\varphi = 90°$) in J band in function of the star's temperature is displayed in Figure 13-9

Figure 13-9. Star/Earthlike contrast in J band.

Measurement SNR

If $I(r,\theta)$ is the halo, then the photon noise is:

$$\sigma_p = \sqrt{N_* I(r,\theta)},$$

and speckle noise is[58]:

$$\sigma_s = (1 - S_r) N_* \sqrt{\frac{1.7}{\pi}\frac{\tau_0}{t}} I(r,\theta),$$

where S_r is the Strehl ratio (see section on AO) and τ_0, the speckles' lifetime assumed to be 10 ms. Moreover, the planet's signal is measured in the core of the PSF so, it is:

$$F_p = 0.85 \times S_r \times N_p(\varphi, i)$$

13.4.2 Two-wavelength Differential Imaging

Spectral DI

SDI allows to lower the speckle noise level under the photon noise one by subtracting 2 images taken at the same time at 2 close wavelengths, λ_1 and λ_2, λ_2 being one of the Extrasolar planet's absorption features. After that subtraction, the residual speckle noise is [57]:

$$\sigma_{ss} = 2(1 - S_r)\frac{\Delta\lambda}{\lambda}\sigma_s$$

315

Of course, the photon noise is going to increase due to this subtraction but we have to consider 2 different cases depending on whether or not our Earth has the same absorption feature (see previous section).

No similar absorption by Earth atmosphere:

$$\sigma_{ps} = \sqrt{2}\sigma_p .$$

Similar absorption by Earth atmosphere:

If β is our atmosphere's transmission at λ_2, we have to divide the second image by this factor to be sure that we obtain the same star's flux on both images, which involves the following photon noise:

$$\sigma_{ps} = \sqrt{\frac{1+\beta}{\beta}}\sigma_p$$

Moreover, if C is the contrast between our Earth's feature and the extrasolar planet's one, then the received signal after subtraction is:

$$F_{ps} = (1-C)F_p$$

13.4.3 Study cases

The characteristics of the planets used in the simulation are similar to Earth-like and Jupiter-like planets at position around their host star such their temperature is similar to the one in the solar system (around a G2 star) (see Table 13-2 and Table 13-3). The system star-planet for each cases is displaced to distances from Earth ranging from 10 to 25 pc.

Spectral type	Star-planet distance (AU)	Star-planet contrast in NIR	Angular separation at 20 pc 90 deg phase
G2	1.00	2.21×10^{-10}	50 mas
K2	0.51	8.07×10^{-10}	25 mas
M2	0.16	8.30×10^{-9}	8 mas

Table 13-2: Characteristics of Earth-like planets used in the simulation.

Spectral type	Star-planet distance (AU)	Star-planet contrast in NIR	Angular separation at 20 pc 90 deg phase
G2	5.10	1.40×10^{-9}	250 mas
K2	1.67	5.32×10^{-9}	80 mas
M2	0.83	5.50×10^{-8}	40 mas

Table 13-3: Characteristics of Jupiter-like planets used in the simulation.

13.4.4 H_2O, Earth-like planet (no cirrus)
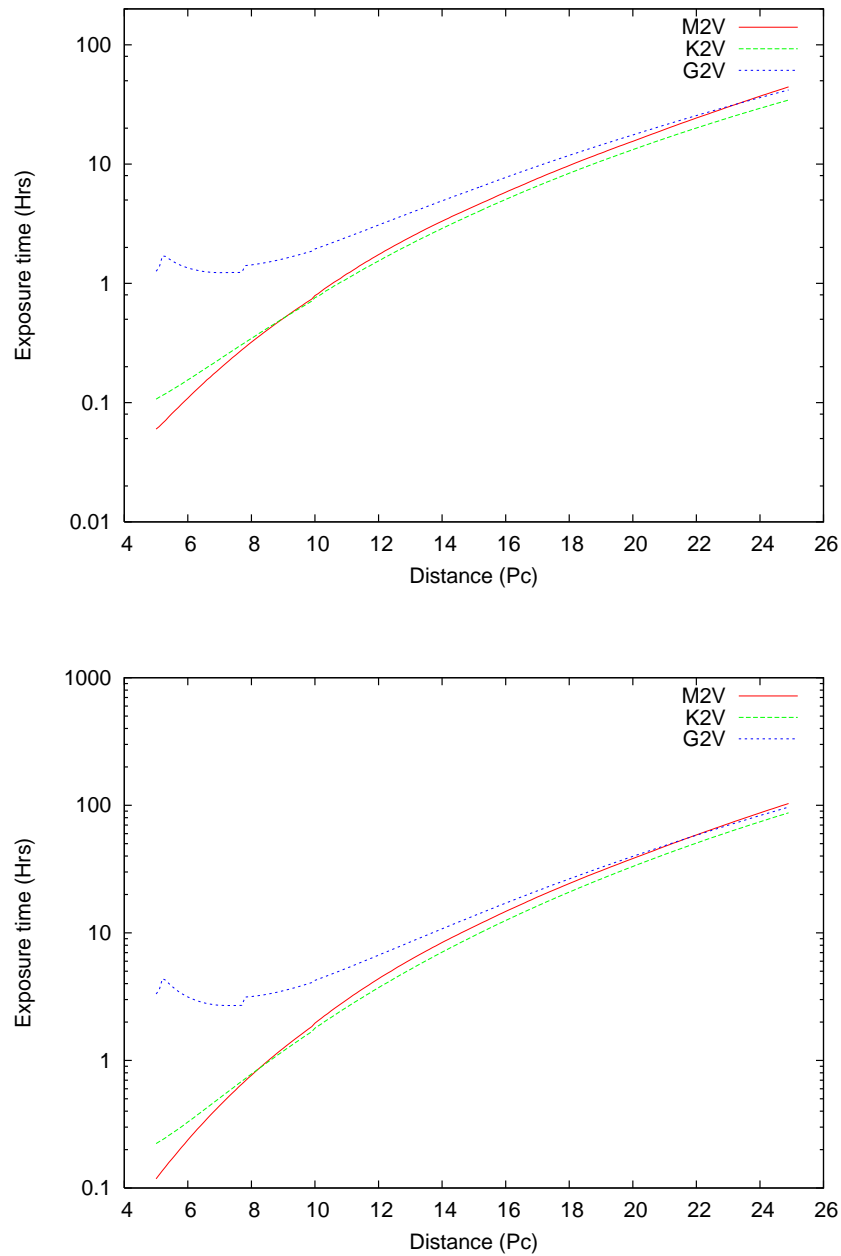
Figure 13-10. Time to detect an Earthlike planet orbiting a Main Sequence star at 5σ (water bands). λ=1300 nm, 80nm bandwidth, 0.44 atmospheric transmission, 0.32 contrast. t0=4ms, r0=20cm (top) and r0=12.1cm (bottom).

13.4.5 CO₂, 10% concentration, Earth-like planet (no cirrus)
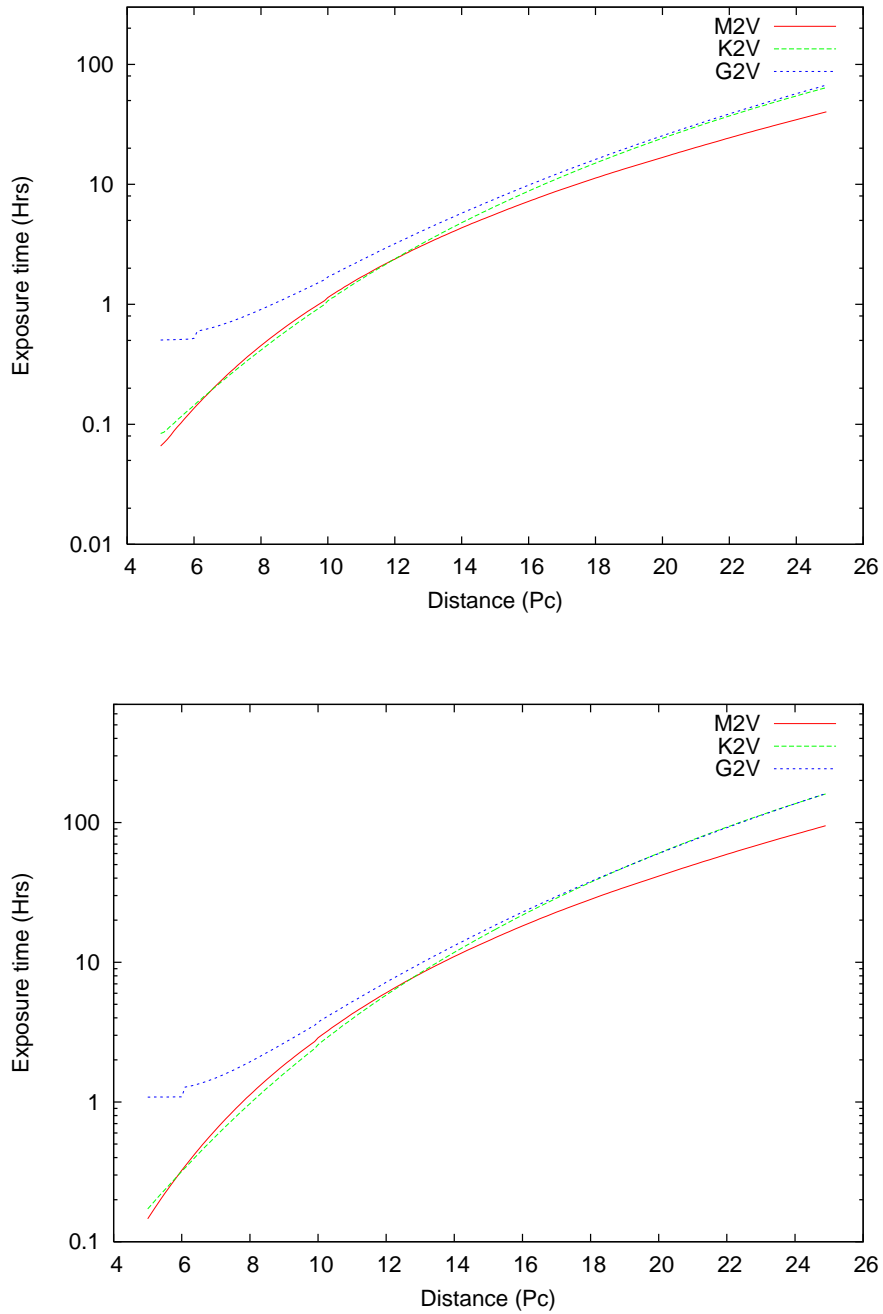
Figure 13-11. Time to detect an Earthlike planet orbiting a Main Sequence star at 5σ (CO₂ band). λ=1.60μm, 100nm bandwidth, 1 atmospheric transmission, 0.43 contrast. τ_0=4ms, r_0=20cm (top) and r_0=12.1cm (bottom).

13.4.6 CO_2, 50% concentration, Earth-like planet (no cirrus)
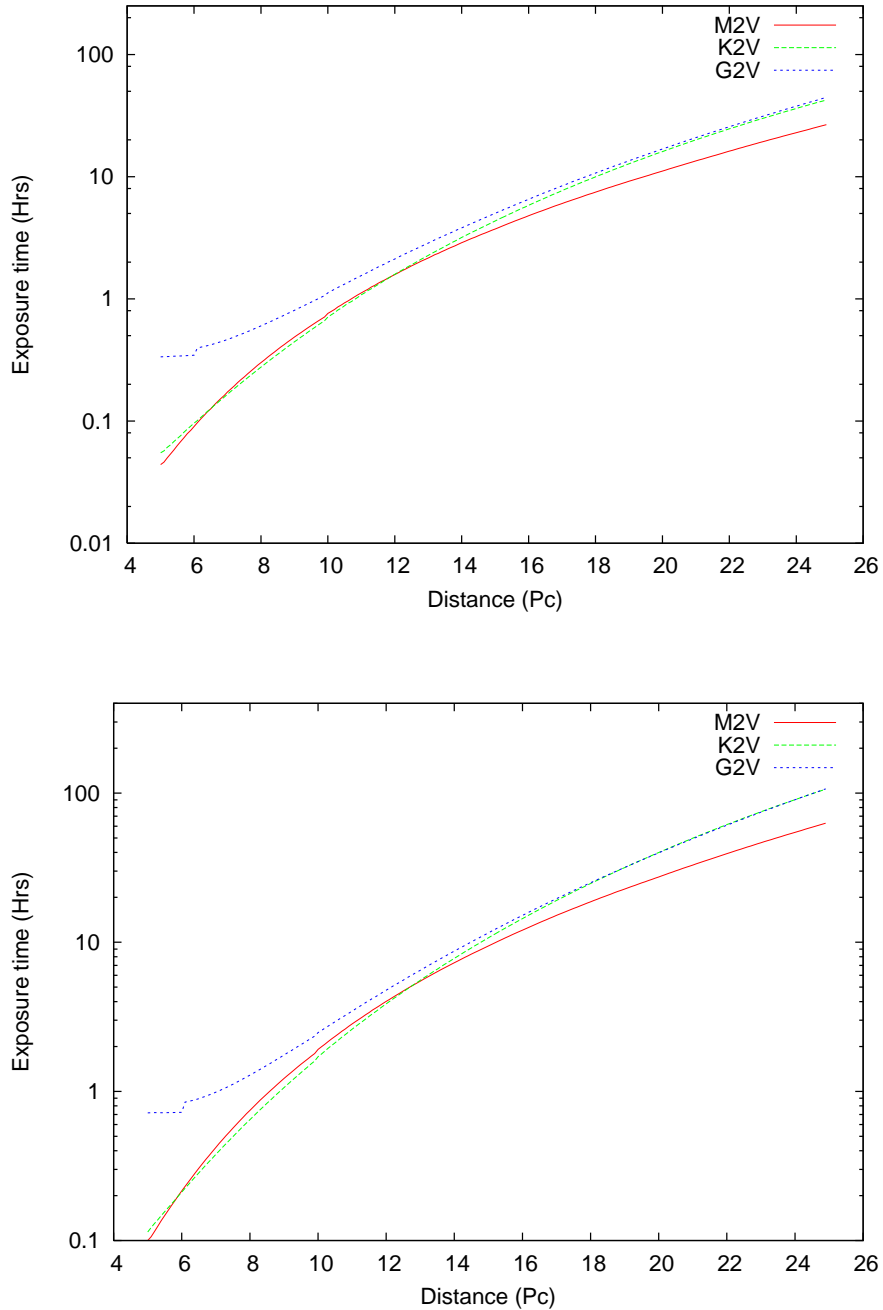
Figure 13-12. Time to detect an Earthlike planet orbiting a Main Sequence star at 5σ (CO_2 band). λ=1.60μm, 100nm bandwidth, 1 atmospheric transmission, 0.3 contrast. τ_0=4ms, r_0=20cm (top) and r_0=12.1cm (bottom).

13.4.7 O_2, Earth-like planet (no cirrus)

Figure 13-13. Time to detect O_2 in an Earthlike planet orbiting a Main Sequence star at 5σ. λ=0.7μm, 30nm bandwidth, 0.81 atmospheric transmission, 0.79 contrast. τ_0=4ms, r_0=20cm (top) and r_0=12.1cm (bottom).

13.4.8 CH₄, Jupiter-like planet
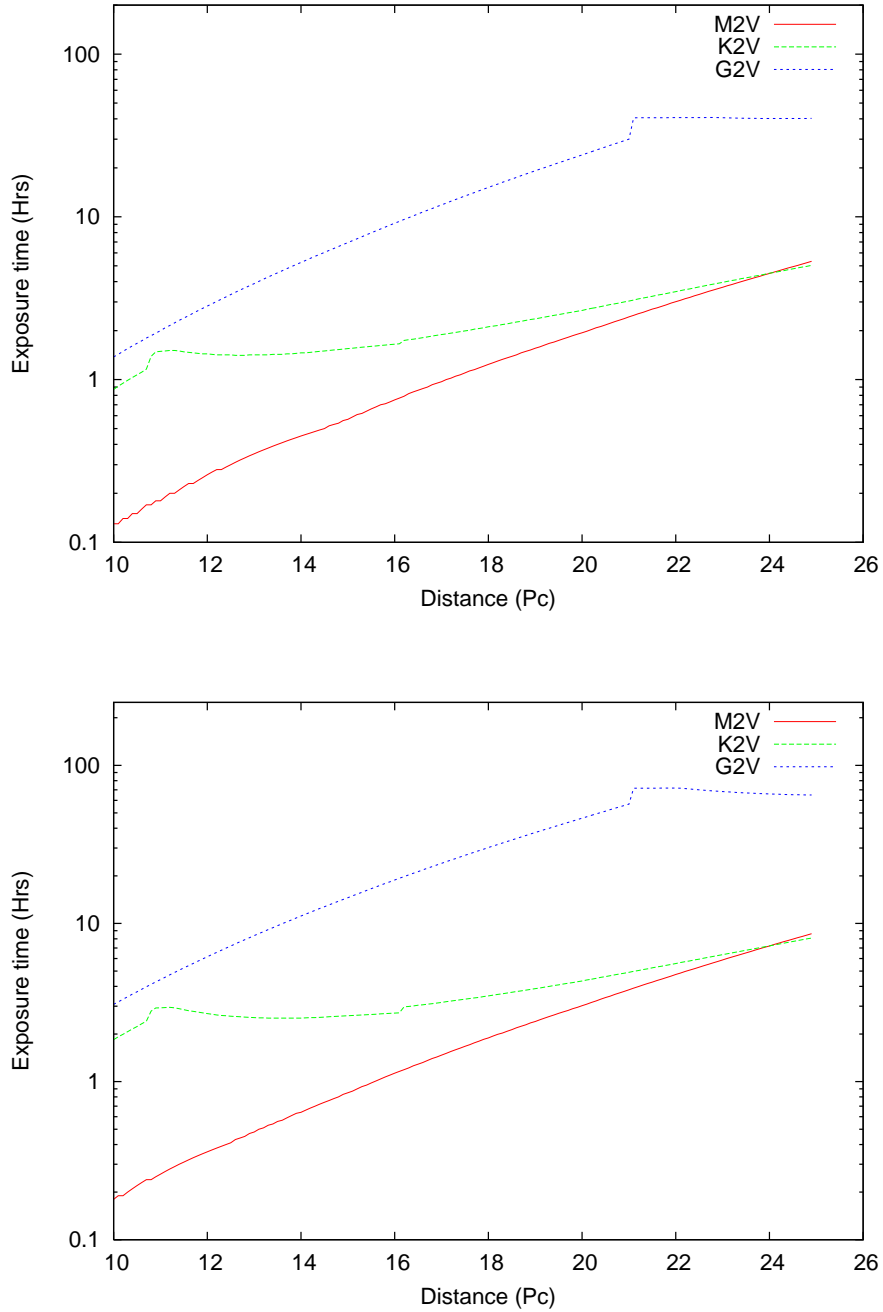
Figure 13-14. Time to detect an Jupiter-like planet orbiting a Main Sequence star at 50σ. λ=1.6µm, 55nm bandwidth. τ₀=4ms, r₀=20cm (top) and r₀=12.1cm (bottom).

13.4.9 Polarimetric signal, Earth-like planet (with clouds and oceans, 15% polarization)

In this case, we considered that speckle noise has been removed by the temporal modulation of a ZIMPOL-like device.

If p is the total extrasolar planet's polarization, then the received signal after subtraction is:
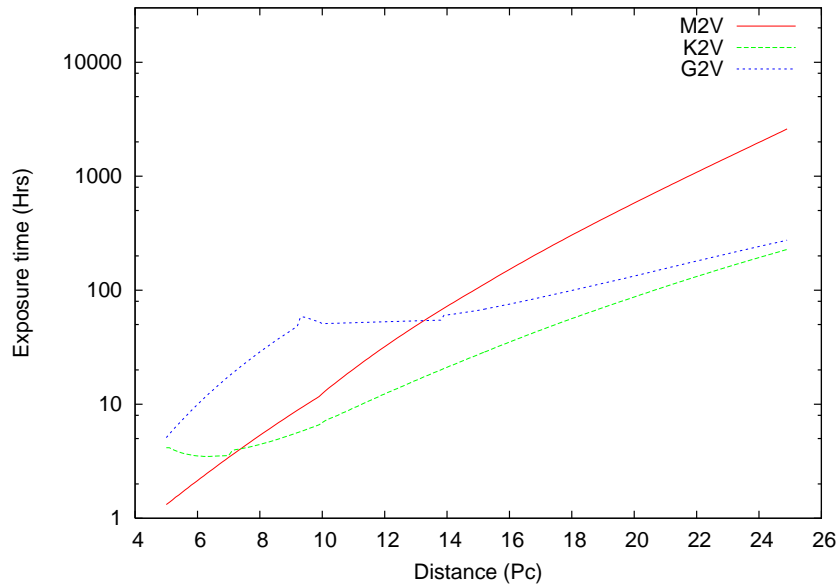
$$F_{pp} \equiv pF_p .$$

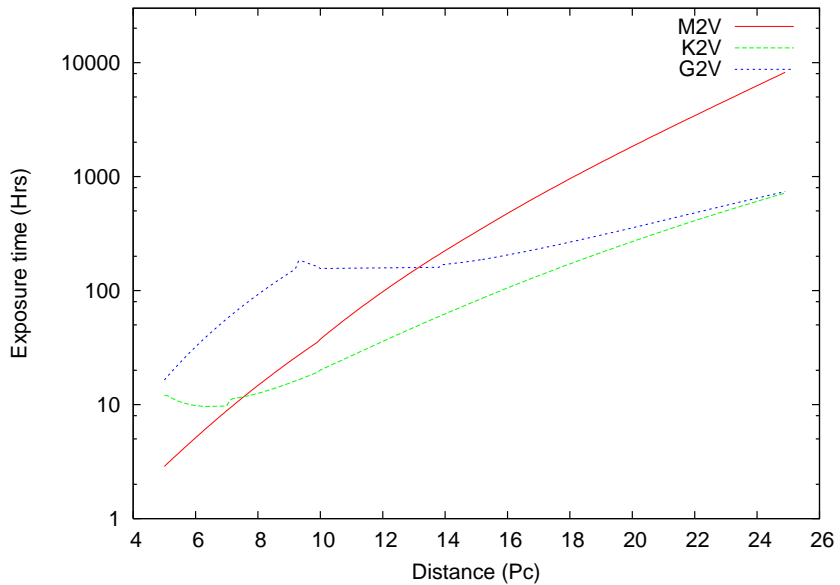**Figure 13-15. Time to detect an Earthlike planet orbiting a Main Sequence star at 5σ. λ=0.7μm, 200nm bandwidth. τ_0=4ms, r_0=20cm (top) and r_0=12.1cm (bottom).**

323

The figures on the previous pages demonstrate that the discovery of exo-Earths and the detection of their most prominent spectral features are possible with a 100-m telescope.

Figure 13-10 shows that the discovery of an exo-Earth by **differential imaging in H_2O is possible** in a couple of hours even in regular observing conditions and fulfills the TLRs (1-2 nights per target, ~ 1hr for closest objects in good conditions). In the same time the filter pair of the differential imager would also be **sensitive to methane**, and Figure 13-14 shows that an exo-Jupiter could be discovered in the same time at an SNR of 50. It is worth noting that contrasts are quite similar for the different types of stars, so increased brightness compensates for reduced angular distance in the case of later spectral types.

Figure 13-12 shows that SNR similar to H_2O can be reached in **CO_2 if abundant**. This could be rather important given that two (Mars, Venus) of the three major rocky planets in the solar system have CO_2 atmospheres. Also, CO_2 is better observed in H-band with the IFS which can be operated simultaneously to the differential imager. In this way, EPICS would be sensitive to all major types of planets in the solar system at the same time, allowing for an efficient search program for extra-terrestrial planets.

Follow up observations on discoveries in O_2 differential imaging and polarimetry is possible in good conditions, but a considerable amount of observing time (10-100 hrs) has to be spent to achieve high enough SNR. The long exposure times in the visible are primarily attributed to the more shallow features compared to the NIR leading to small photon return. For example, the bandwidth of the O_2 A-band at 760 nm is about 10 nm, and the photon flux delivered by an exo-Earth at 10 pc observed with a 100-m telescope is just ~2500 per hour (assuming 16% transmission) considering differential imaging of the O_2 feature with a contrast of ~0.7 (see section 13.3.1). These photons have to be detected against the stellar residuals which are almost 10000 times higher (cf. Figure 8-22). The discontinuities of the contrast curve representing the G2V star in Figure 13-13 and Figure 13-15 result from the PSF contrast delivered by the two stage AO system (Figure 8-22). The control radius of the fast second stage corresponds to about 100 mas at visible wavelengths.

Note that the effect of seeing is quite important: A factor 2 to 3 in integration times is found between observations with good seeing (0.5 arcsec) and mean seeing (0.85 arcsec).

The considerations above lead to a proposed strategy for an **EPICS survey.** Each of the about 300 target stars would have to be observed for 1-2 nights at two different periods or more. Multiple visits are required in order to limit misses because of unfavorable orbit configurations. Hence, the survey would require roughly 1000 nights of observing time. After that, follow up observations to confirm common proper motion and to do additional bands and polarimetry at visible wavelengths and in the thermal IR (T-OWL) would be required.

14 Conclusions

This document describes a preliminary study of EPICS a Planet finder instrument for OWL. It is based on the experience of the phase A study for VLT- Planet Finder. The main concept for planets detection is differential imaging by spectroscopy and polarimetry. An XAO system concept and several coronagraph concepts have been studied. Three instrument concepts have been developed.

The performances show that, thanks to a very efficient XAO system, a high rejection of the starlight halo in J and H band can be obtained permitting to detect H_2O and CO_2 (if abundant > 10%) in Rocky planets in about one or two observation nights up to 20 pc.

Detection by polarimetry will also be possible, with however the difficulty that the observation takes place in the visible where extreme AO correction is harder.

O_2 can also be detected but requires several tens to several hundreds of hours for detection in an Earth-like planet.

Detection of Jupiter-like planets in CH_4 bands in J and H band can be done in one night of observation with high SNR (50 σ) opening the door to high resolution spectroscopy of gas giant planets.

The target sample is statistically compatible with the one of DARWIN space mission so that a very good complementarity of EPICS in visible and NIR and DARWIN in thermal IR can be expected.

There are numerous risk areas (see section 4) in technology developments or in more fundamental issues related to the Earth atmosphere or to ultimate optical precision (systematic differential aberrations issue), but we are already on the roadmap to solve these issues.

The preliminary system analysis permitted to define some specifications that the instrument should fulfill. The exercise we made consisted in deriving the specifications of some key elements of the design, in terms of systematic errors, permitting to attain the goal contrast with the instrument without further calibration.

The outcome is that the requirements in terms of stability and optical quality for some key elements are quite demanding, but not impossible to achieve. Such requirements will affect any instrument dedicated to the extreme contrast required for exo-planets detection especially for Earth-like planets (TPF-C for example). The use of elaborated calibration strategies, that we haven't studied yet, will give even more confidence that we will be able to actually achieve the goal contrast. Moreover recent new ideas go into the direction of solving the problems of systematic errors [60], [63].

15 Bibliographic references

[1]   C. Verinaud, Optics Communications, Volume 233, Issue 1-3, p. 27-38 (2004).

[2]   O. Guyon, The Astrophysical Journal, Volume 629, Issue 1, pp. 592-614 (2005).

[3]   Vérinaud, C.; Le Louarn, M.; Korkiakoski, V.; Carbillet, M., Monthly Notices of the Royal Astronomical Society: Letters, Volume 357, Issue 1, pp. L26-L30 (2005)

[4]   B. Ellerbroek, Journal of the Optical Society of America A: Optics, Image Science, and Vision, vol. 19, iss. no. 9, p. 1803-1816 (2002)

[5]   L. Poyneer, Gavel D.T., Brase J.M., Journal of the Optical Society of America A, vol. 19, Issue 10, pp.2100-2111 (2002)

[6]   L. Poyneer, J.P Veran, accepted in Journal of the Optical Society of America A, (2005).

[7]   Le Roux, B.; Coyne, J., Ragazzoni, R., Applied Optics IP, vol. 44, Issue 2, pp.171-177 (2005).

[8]   Nicolle, M.; Fusco, T.; Rousset, G.; Michau, V., Optics Letters, Volume 29, Issue 23, pp. 2743-2745 (2004).

[9]   Lardiere, O., Carbillet, M, Riccardi, A., Salinari, P. Proceedings of the SPIE, Volume 5490, pp. 516-526 (2004).

[10]   Codona, J.L., Angel, R., The Astrophysical Journal, Volume 604, Issue 2, pp. L117-L120 (2004).

[11]   J. Tinbergen, *Interstellar polarization in the immediate solar neighbourhood,* Astron. Astrophys., **105**, 53-64, 1982

[12]   I'll look uo which is the most significant

[13]   Lynch D.W., Hunter W.R., 1985, Handbook of optical constants of solids, E.D. Palik (ed.), Academic press

[14]   Schmid H.M., Appenzeller I., Stenflo J.O., Kaufer A., 2002, in: Scientific drivers for ESO future VLT/VLTI instrumentation, J. Bergeron & G. Monnet (eds), ESO Astrophysics Symposia, p. 231

[15]   F. Rigaut, J.P. Veran, O. Lai, Proc. SPIE 3353 (1998) 1038.

[16]   Jolissaint, L., Veran, J.P., ESO conf. Proc 58 Beyond conventional Adaptive Optics, 2002

[17]   R. Conan, in preparation

[18]   Poyneer, Lisa A.; Macintosh, Bruce, Journal of the Optical Society of America A, vol. 21, Issue 5, pp.810-819, 2004.

[19]   Wallner, E. P, Journal of the Optical Society of America, vol. 67, issue 3, p.407, 1977

[20]   Coronagraphic methods for the detection of Terrestrial planets", Ed. A. Quirrenbach, ESA Publication Division, 2005

[21]   C. Aime, R. Soummer, A. Ferrari, "Total coronagraphic extinction of rectangular apertures using linear prolate apodization", A&A 389, 334, 2002

[22]   R.Soummer, "Apodized Pupil Lyot Coronagraphs for Arbitrary Telescope Apertures" ApJ. Letters, 618, L161, 2005

[23]   C. Aime "Proposal of an Achromatic Prolate Apodized Lyot Coronagraph", A&A, 431, 1105 2005

[24]   B. Lyot, "A study of the solar corona and prominences without eclipses" Mon. Not. R. Astron. Soc. 99, 579-595, 1939

[25]   A. Sivaramakrishnan, C. Koresko, R. Makidon, T. Berkefeld, M, Kuchner, "Ground- based coronagraphy with high-oredr adapive optics", The Astrophysical Journal 552:397 − 408, 2001

[26]   C. Aime and R. Soummer "Cascading properties of Prolate Apodized Lyot Coronagraphs", EAS Publications Series, **12** 281-286, 2004

[27]   Soummer etal. 2005 in preparation

[28]   N.Yaitskova, "Double stage Lyot coronagraph with the apodized reticulated Lyot stop for ELT", SPIE 5905-35.

[29]   M. Kasper, "EPICS Top-Level Requirements", EPICS report

[30]   Baudoz et al. 2005, *Stellar Coronagraphy: Study and Test of a Hybrid Interfero-Coronagraph,* PASP, in press (available from PASP electronic early-release papers)

[31]   Poyneer, Lisa A.; Macintosh, Bruce, Journal of the Optical Society of America A, vol. 21, Issue 5, pp.810-819, 2004.

[32]   Verinaud, Christophe; Le Louarn, Miska; Korkiakoski, Visa; Braud, Jeremy, Advancements in Adaptive Optics. Edited by Domenico B. Calia, Brent L. Ellerbroek, and Roberto Ragazzoni. Proceedings of the SPIE, Volume 5490, pp. 1177-1188 (2004).

[33]   C. Cavarroc, A. Boccaletti, P. Baudoz, T. Fusco, D. Rouan, 2005, A&A, accepted.

[34]   Perrin, Marshall D.; Sivaramakrishnan, Anand; Makidon, Russell B.; Oppenheimer, Ben R.; Graham, James R, The Astrophysical Journal, Volume 596, Issue 1, pp. 702-712.(2003)

[35]   http://planetquest.jpl.nasa.gov/TPF/TPF-CTechPlan.pdf

[36]   Basic references for optical polarization techniques:
       R.M.A. Azzam and N.M. Bashara, *Ellipsometry and polarized light,* ISBN 0444870164

       C.U. Keller, 2001: chapter 8 of J.Trujillo-Bueno, F.Moreno-Insertis and F.Sanchez (eds), *Astrophysical spectropolarimetry,* ISBN: 052180983.

       J.O. Stenflo, *Solar Magnetic Fields,* ISBN 0792327934, 1994.

       J.Tinbergen, 1996, *Astronomical Polarimetry,* ISBN 0521475317

[37]   C.U. Keller and R.N. Smartt, Solar Physics, vol 166, pp 311-315 (1996)

[38]   J.Tinbergen, *Precision Spectropolarimetry of Starlight: Development of a Wide-band Version of the Dollfus Polarization Modulator,* Astron. & Astrophys., vol 23**,** pp 25-48 (1973).

[39]   R. Clark Jones, *A New Calculus for the Treatment of Optical Systems III,* Journal of the Optical Society of America, vol 31, pp 500-5033 (1941) (reprinted as SPIE Milestone Series: MS23 Selected Papers on Polarization, 336-339).

[40]   E. Cojocaru et al, *Achromatic thin-film totally reflecting quarterwave retarders,* Applied Optics, vol 28**,** pp 211-212 (1989)

[41]   S.E. Harris, E.O. Ammann and I.C. Chang, JOSA, vol 54, pp 1267-1279 (1964); more in: Ammann and Chang, JOSA, vol 55, pp 835-841 (1965); Ammann, JOSA, vol 56, pp 943-951, (1966); Ammann, JOSA, vol56, pp 952-955 (1966), Ammann and J.M. Yarborough, JOSA, vol 56, pp 1746-1754 (1966) (lossless networks=Solc-like filters), Ammann and Yarborough, JOSA, vol 57, pp 349-353 (1967); Yarborough and Ammann, JOSA, vol 58, pp 776-783 (1968) (experiments, passband examples); Ammann, JOSA, vol 55, p 412-417 (1965) (applications to arbitrarily polarized light)

[42]   Prieto, Eric; Saisse, Michel; Hibon, Pascale; Beuzit, Jean-Luc; Lagrange, Anne-Marie 2004, "High-contrast differential-aberration-free 3D spectrometer for the Planet Finder instrument," SPIE 5492, 271.

[43]   Sparks, W. B. and Ford, H. C. 2002, "Imaging spectroscopy for extrasolar planet detection," ApJ 578, 5436-564.

[44]   Maillard 1996, "Seismology with a Fourier-transform spectrometer: applications to giant planets and stars," Applied Optics 35, 2734.

[45]   Martin, D. H. 1982 "Polarizing (Martin-Puplett) Interferometric Spectrometers for the Near- and Submillimeter Spectra." Ch. 2 in Infrared and Millimeter Waves 6, 65-148. New York: Academic Press.

[46]   Roddier & Roddier 1989, "Pupil-Plane Interferometry", Proceedings of the NATO Advanced Study Institute, held in Cargese, 1988, ASI Series C, Volume 274, p.221.

[47]   Aben, I., F. Helderman, D.M. Stam, and P. Stammes, Geophys. Res. Lett., Volume 26, pp. 591-594 (1999)

[48]   De Haan, J.F., P.B. Bosma, and J.W. Hovenier, A&A, Volume 183, issue 2, pp. 371-391 (1987)

[49]   Hansen, J.E., and J.W. Hovenier, Journal of Atmospheric Sciences, Volume 31, issue 4, pp. 1137-1160 (1974)

[50]   Hansen, J.E., and L.D. Travis, Space Science Reviews, Volume 16, pp. 527-610 (1974)

[51]   Hovenier, J.W., C. Van der Mee, and H. Domke, Transfer of Polarized Light in Planetary Atmospheres, Basic concepts and practical methods, Kluwer Academic Publishers (2004)

[52]   Stam, D.M., J.F. De Haan, J.W. Hovenier, and P. Stammes, Journal of Geophysical Res. D., Volume 104, pp. 16843-16858 (1999)

[53]   Stam, D.M., J.W. Hovenier, and L.B.F.M. Waters, A&A, Volume 428, pp. 663-672 (2004)

[54]   Van de Hulst, H.C., *Multiple Light Scattering. Tables, formulas, and application.* New York, Academic Press (1980)

[55]   Perrin et al, the structure of high Strehl ratio PSF, ApJ 596, 2003

[56]   Masciadri & Raga, exoplanet detection using a wavelet analysis technique, ApJ 611, 2004

[57]   Marois, C., Doyon, R., Racine, R. and Nadeau, D., Efficient Speckle Noise Attenuation in Faint Companion Imaging, PASP 112, 91-96 (2000)

[58]   Racine, R., Walker, G.A.H., Nadeau, D., Doyon, R. and Marois, C., Speckle Noise and the Detection of Faint Companions, PASP 111, 587-594 (1999)

[59]   Lenzen, Rainer; Close, Laird; Brandner, Wolfgang; Biller, Beth; Hartung, Markus, Ground-based Instrumentation for Astronomy. Edited by Alan F. M. Moorwood and Iye Masanori. Proceedings of the SPIE, Volume 5492, pp. 970-977 (2004).

[60]   Angel, R., 2003, in Towards Other Earths: DARWIN/TPF and the search for Extrasolar Planets, ESA SP-539, 221

[61]   Lardiere, O., Salinari, P., Jolissait, L. et al. 2004, in Proc. SPIE 5382,550.

[62]   Chelli, A., Astronomy and Astrophysics, Volume 441, Issue 3, October III 2005, pp.1205-1210, 2005.

[63]   Guyon, O., *The Astrophysical Journal*, volume 615, part 1 (2004), pages 562–572.