

METIS high-contrast imaging: design and expected performance

Brunella Carlomagno,^a Christian Delacroix,^a Olivier Absil,^{a,*}
Faustine Cantalloube,^b Gilles Orban de Xivry,^a Prashant Pathak,^c
Tibor Agocs,^d Thomas Bertram,^b Bernhard Brandl,^e Leonard Burtcher,^e
Markus Feldt,^b Adrian Glauser,^f Stefan Hippler,^b Matthew Kenworthy,^e
Remko Stuik,^e and Roy van Boekel^b

^aUniversity of Liège, STAR Institute, Sart Tilman, Belgium

^bMax-Planck-Institut für Astronomie, Heidelberg, Germany

^cEuropean Southern Observatory, Garching, Germany

^dASTRON, NOVA Optical Infrared Instrumentation Group, Dwingeloo, The Netherlands

^eLeiden University, Leiden Observatory, RA Leiden, The Netherlands

^fETH Zürich, Institute for Particle Physics and Astrophysics, Zürich, Switzerland

Abstract. With the advent of 30- to 40-m class ground-based telescopes in the mid-2020s, direct imaging of exoplanets is bound to take a new major leap. Among the approved projects, the Mid-infrared Extremely Large Telescope (ELT) Imager and Spectrograph (METIS) instrument for the ELT holds a prominent spot; by observing in the mid-infrared regime, it will be perfectly suited to study a variety of exoplanets and protoplanetary disks around nearby stars. Equipped with two of the most advanced coronagraphs, the vortex coronagraph and the apodizing phase plate, METIS will provide high-contrast imaging (HCI) in L -, M - and N -bands, and a combination of high-resolution spectroscopy and HCI in L - and M -bands. We present the expected HCI performance of the METIS instrument, considering realistic adaptive optics residuals, and investigate the effect of the main instrumental errors. The most important sources of degradation are identified and realistic sensitivity limits in terms of planet/star contrast are derived. © The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JATIS.6.3.035005](https://doi.org/10.1117/1.JATIS.6.3.035005)]

Keywords: exoplanets; high-contrast imaging; coronagraphy; adaptive optics; end-to-end simulations.

Paper 20007 received Jan. 21, 2020; accepted for publication Sep. 1, 2020; published online Sep. 24, 2020.

1 Introduction

Exoplanet science has been making great progress over the last few years, with, e.g., the detection of potentially habitable planets using radial velocity and transit methods^{1,2} and the detection of candidate protoplanets using direct imaging techniques.^{3–5} The future looks even brighter with the upcoming launch of the James Webb Space Telescope and the planned construction of three giant segmented mirror telescopes, among which is the Extremely Large Telescope (ELT). In particular, direct imaging techniques are currently pushing their sensitivity limits toward fainter, closer-in planets due to the development of better adaptive optics systems (e.g., SAXO,⁶ SCExAO,⁷ and SOUL⁸), of more efficient coronagraphic techniques [e.g., vortex coronagraph (VC),⁹ apodizing phase plate (APP),¹⁰ apodized Lyot coronagraph,¹¹ and shaped pupil¹²], and of advanced postprocessing algorithms (e.g., Andromeda,¹³ KLIP-FM,¹⁴ and SODINN¹⁵). These techniques will unleash their full potential on future giant telescopes where they are bound to close the gap with indirect detection methods (transits and radial velocity) in terms of semimajor axis (indirect methods are sensitive to separations typically smaller than 3 AU, while direct methods are sensitive to separations generally larger than 10 AU) and mass of the detected

*Address all correspondence to Olivier Absil, E-mail: olivier.absil@uliege.be

companions. One of the keys to bridge this gap is expected to be the combination of high-contrast imaging (HCI) and high-resolution spectroscopy (HRS), which not only could give access to fainter planets but also holds the promise of characterizing their atmospheres.^{16–18} In this context, the Mid-infrared Extremely Large Telescope (ELT) Imager and Spectrograph (METIS^{19,20}) is a future workhorse instrument for exoplanet science, combining HCI and HRS capabilities in the mid-infrared regime, fed by a high-performance adaptive optics system.²¹

The mid-infrared is widely regarded as a key regime for exoplanet studies in general, and as a sweet spot for exoplanet imaging in particular. The main incentive to observe planetary systems in the mid-infrared is the lower contrast between the (warm to cool) circumstellar material and their (generally hot) host star. This is particularly true not only for protoplanetary disks and exozodiacal disks but also for exoplanets, which generally radiate most of their thermal energy in the mid-infrared domain.²² In addition, the wavefront quality provided by adaptive optics is significantly better in the mid-infrared than in the visible or near-infrared range. This has already resulted in high image quality for first-generation single-conjugate Adaptive Optics (AO) systems on 10-m class telescopes at the *L*-band (Strehl $\sim 80\%$, e.g., Ref. 23), and even more so at the *N*-band (Strehl $>95\%$, see e.g., Refs. 24 and 25). The advantage of observing in the thermal infrared domain is becoming even clearer considering that the majority of candidate protoplanets published to date have been found at the *L*-band,^{5,26–28} and that even fully formed exoplanets are sometimes hard to recover at shorter wavelengths (e.g., HD95086,^{29,30} which had to wait for the advent of extreme-AO HCI instruments to be confirmed in the near-infrared³¹). METIS will take advantage of this optimal spectral range to push the sensitivity limits of direct imaging to unprecedented levels in terms of planetary masses, possibly reaching the realm of terrestrial planets around a few targets.³²

In this paper, we analyze the HCI performance of METIS. The instrument and its main characteristics, including coronagraphic devices, are presented in Sec. 2. The simulation tools used to predict its HCI performance are then presented in Sec. 3. In Sec. 4, we present the speckle noise-limited performance of METIS HCI in the presence of residual turbulence from the AO system and discuss the influence of the AO simulation parameters. In Sec. 5, we present the nominal HCI performance for the four main spectral bands (*L*, *M*, *N1*, and *N2*), taking into account photon noise and thermal background emission in addition to speckle noise. Section 6 is dedicated to the analysis of the HCI performance in the presence of various instrumental errors. In Sec. 7, we conclude by illustrating the effect of a combination of various error sources on the METIS coronagraphic performance.

2 ELT/METIS Instrument

METIS is one of the first-light instruments of the ELT. It consists of two diffraction-limited imagers covering the mid-infrared wavelength range from 3 to 14 μm over a $10'' \times 10''$ field of view featuring HCI and medium-resolution ($R \sim 5000$) long-slit spectroscopy and an integral-field spectrograph providing a spectral resolution $R \sim 100,000$ from 2.9 to 5.3 μm . While perfectly suited to exoplanet science, observing in the mid-infrared regime from the ground comes with the major hindrance of the high thermal background emission originating from the Earth atmosphere and from the warm optical train of the telescope. This background emission significantly affects the sensitivity of ground-based telescopes. To reduce the influence of thermal background as much as possible, METIS will be fully enclosed in a cryostat, thereby limiting the number of warm components to its bare minimum, as already done in other mid-infrared instruments such as VISIR³³ or LBTI.³⁴ METIS will also feature an internal (cold) chopping mirror to enable rapid switching between the target source and an empty sky region. Cryogenic chopping is essential for the 40-m class ELT, because, unlike 8-m class telescopes, it cannot provide classical M2 chopping due to its optical architecture and to the mass/size of the telescope mirrors.

2.1 METIS Coronagraphic Modes

To enable HCI, METIS will implement two main state-of-the-art coronagraphic concepts: a focal-plane (FP) coronagraph, VC, and a pupil plane (PP) coronagraph, APP. These

coronagraphic concepts are complementary in terms of discovery space and sensitivity to instrumental errors. They will be used for different science cases and/or under different environmental conditions.

The VC^{9,35} consists of an FP phase mask (the vortex phase mask), combined with a downstream PP amplitude mask (the Lyot stop). Because the light is completely diffracted outside of the geometric pupil of the telescope by the phase mask, the Lyot stop does not require any undersizing in principle, so that the throughput of a VC is virtually 100%, being limited only by the intrinsic transmission of the phase mask. In our case, a thin synthetic diamond substrate is used to implement the phase mask, which reduces the throughput down to about 90% at the *L*-band (95% at the *N*-band) for a 300- μm -thick substrate.

As with most FP coronagraphs, the VC is affected by a substantial loss in contrast performance for partially obstructed pupils (central obstruction and secondary support structures). Pupil apodization can compensate for the presence of a central obstruction: the ring-apodized vortex coronagraph (RAVC³⁶) combines a VC with a grayscale amplitude ring apodizer in an upstream PP to produce a perfect cancellation of the on-axis light for a centrally obscured aperture in a region of reduced sized in the Lyot stop plane. For the sake of simplicity, without significantly affecting the overall coronagraphic performance, we will consider in the following a circularized version of the ELT pupil, using the largest circle inscribed ELT pupil and a central obstruction of 30% (also known as an all-glass pupil). In this case, the optimized ring apodizer transmission is set to 76% for radii larger than 62%. Although it provides a more modest stellar cancellation performance, a classical vortex coronagraph (CVC) without an apodizer will also be implemented for the cases where throughput is more important. For both the classical vortex and the ring-apodized vortex, the Lyot stop outer diameter is set to 98% of the external pupil diameter to account for pupil stability and blurring, and the Lyot stop spider width to 3% of the pupil diameter. The central obstruction is oversized to 33% of the external pupil diameter in the case of the classical vortex, while it goes up to 65% for the RAVC. Neglecting the effect of the phase mask (which is technology-dependent), the overall throughput of the CVC is about 90%, while it goes down to 34% for the RAVC. Due to the presence of the support structure (spider arms), the total stellar leakage (i.e., the total fraction of stellar light making it through the Lyot stop) is limited to 0.2% in the case of the RAVC (see Fig. 1), while it is 3% for the classical vortex. A CVC will be available at the *L*-, *M*-, and *N*-bands in METIS, while the RAVC will be restricted to the *L*- and *M*-bands, where stellar rejection rate is more critical due to the lower thermal background.

The APP^{10,37} is a pupil-plane phase coronagraph consisting of an antisymmetric phase pattern that redistributes the light in the image plane, creating a dark hole of 180-deg clear space between typically two and several diffraction widths (λ/D). In the grating-vector APP (gvAPP), the phase pattern is produced by self-aligning a multitwist liquid crystal retarder,³⁸ on top of which a polarization grating is added. The result is an image with two complementary APP point-spread functions (PSFs) and an on-axis leakage term. The two coronagraphic PSFs can

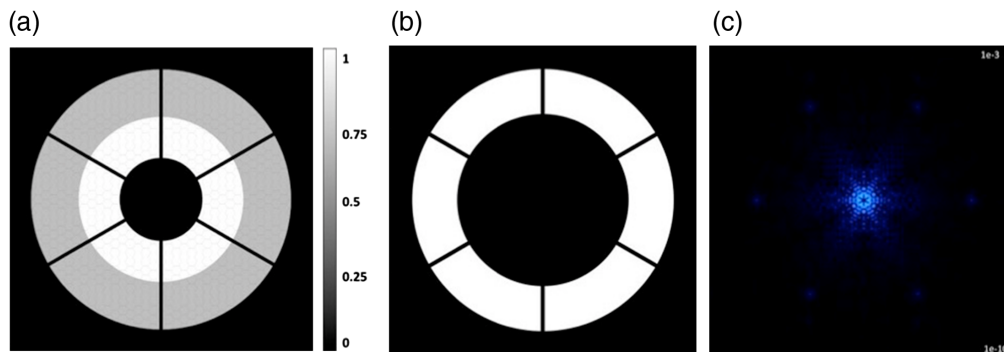


Fig. 1 Illustration of the RAVC design for METIS, showing the input pupil (circularized version of the ELT pupil) with its (a) grayscale apodizer, (b) the Lyot stop, and (c) the associated theoretical PSF for the charge-2 RAVC on a field of view of $100\lambda/D$.

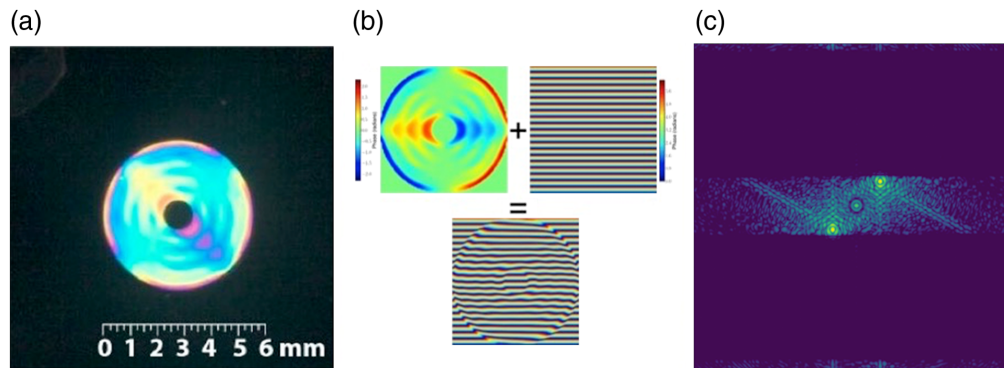


Fig. 2 (a) The vAPP is obtained by superimposing three layers of liquid crystal. (b) The gvAPP concept, where a polarization grating is added on top of the vAPP phase pattern. (c) The final image produced by the METIS gvAPP, with the characteristic three PSFs: the faint leakage PSF in the center, and the two coronagraphic PSFs on either side with their 180-deg suppression zone. In this design, we have also created secondary dark holes to make sure that the two conjugated PSFs do not fill up each other's dark hole.

be combined to obtain an almost 360-deg clear discovery space, while the leakage term can be used as photometric and astrometric reference. Being a pupil-plane coronagraph, the gvAPP can be optimized for complex pupils with central obstruction and secondary support structures. Unlike VC, it is insensitive to tip-tilt errors to the first order. METIS will make use of gvAPPs in the *L*- and *M*-bands, which will feature their own, slightly undersized cold pupil stop to ensure uniform illumination of the APP phase pattern at all times. The structure of the gvAPP PSF designed for the METIS imager is shown in Fig. 2.

2.2 Instrumental Concept

Before entering the METIS cryostat, the light collected by the ELT primary mirror (ELT-M1) is reflected by a series of five additional mirrors, including the ELT-M4 deformable mirror and a tip-tilt mirror (ELT-M5). After bouncing off the final mirror of the ELT optical train (M6), the light then enters the METIS cryostat, which ensures low background emission for all further optical components and includes four main subsystems.

- The common fore-optics (CFO), which formats the beam for injection into the scientific cameras and includes important functionalities, such as atmospheric dispersion compensation (ADC), derotation, pupil stabilization, chopping, and coronagraphy.
- A single-conjugated adaptive optics (SCAO) system, which implements an infrared pyramid wavefront sensor (PyWFS) and controls the shape of the ELT-M4 deformable mirror.
- A pair of diffraction-limited imaging cameras at *L/M*- and *N/Q*-bands with $10'' \times 10''$ field of view (referred to as IMG-LM and IMG-NQ, respectively), including coronagraphy at *L*-, *M*-, and *N*-bands, and medium-resolution ($R \sim 5000$) long-slit spectroscopy (available in *L/M*-bands).
- An integral-field high-resolution ($R \sim 100,000$) spectrograph, referred to as LMS (LM-band Spectrograph), operating at *L/M*-band (2.9 to 5.3 μm) with a $05 \times 1''$ field of view and including coronagraphic capabilities.

METIS also comprises a warm calibration unit, which provides calibration sources for all observing modes.

Inside METIS, the CFO subsystem provides a series of addressable PPs and FPs to implement coronagraphic capabilities, as shown in Fig. 3. In particular, the CFO-PP1 pupil plane and the CFO-FP2 focal plane will be the hosts of wheel mechanisms that can insert coronagraphic components into the beam. This is also the case of the first PP located inside the imager and spectrograph subsystems, respectively, referred to as IMG-LM-PP1 and IMG-NQ-PP1 for the

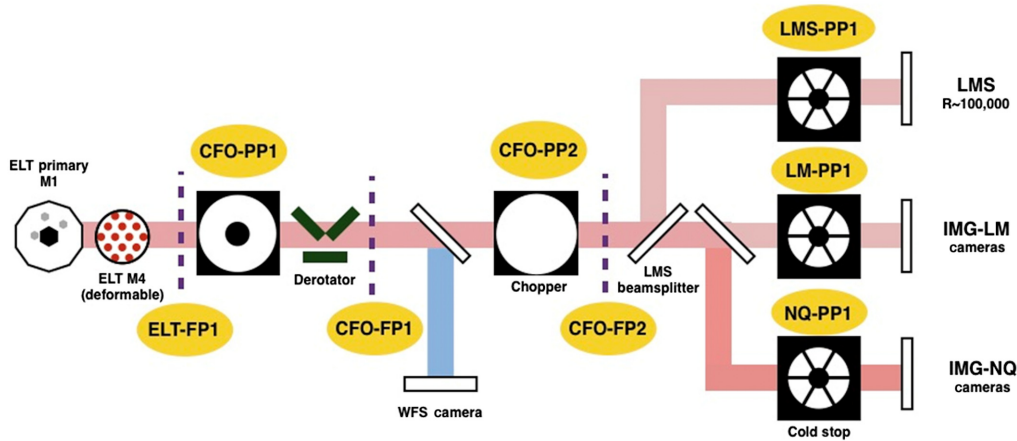


Fig. 3 Schematic layout of the relevant optical components for HCI in METIS.

imager (the two PPs inside the imagers will be referred to as IMG-PP1 for convenience, although they are physically distinct) and LMS-PP1 for the spectrograph. This combination of two PPs with one FP in between enables the implementation of the RAVC, while the gvAPP only requires one PP and will be implemented directly into the IMG-LM and LMS subsystems.

CFO-PP1 is located upstream of any derotation or pupil stabilization mirror. Because the image of the ELT pupil will not be stable there and will be rotating during the observations, this PP can only implement rotationally symmetric masks, such as the RAVC. More advanced apodizer designs taking into account spider arms cannot be implemented there nor the APP, whose design is also tuned to the exact shape of the ELT pupil. The cold stop implemented in CFO-PP1 will be slightly oversized when using the coronagraphic modes to transmit the full image of the ELT-M1 pupil. Because the image of ELT-M1 will drift with respect to the position of CFO-PP1 during observations, the grayscale apodizer itself will also be oversized. Due to a double mechanism, a fixed ADC, optimized for a few specific airmasses, will be inserted slightly upstream of the location of the pupil image. Following CFO-PP1, a K -mirror will be used to derotate the beam. The derotator can be used either to stabilize the field or the pupil. The latter option will be used for the HCI observing modes, to exploit angular differential imaging (ADI) techniques. Close to the first focal plane (CFO-FP1), a pupil stabilization mirror will be used to compensate for residual drifts of the pupil image due to the ELT pupil motion and to the imperfect derotation, and to freeze the position of the ELT-M1 pupil in the downstream optical train with an accuracy of 1% derogator the projected M1 diameter. This pupil stabilization mirror will be fed by measurements from the SCAO PyWFS, which will measure the position of the ELT-M1 pupil in real time.²¹ The SCAO subsystem will pick up the near-infrared part of the beam using a dichroic filter downstream of the pupil stabilization mirror, which reflects the H- and K-band signals toward SCAO and transmits the longer wavelengths to the science cameras. The next PP (CFO-PP2) will be hosting the cold chopping mirror, while CFO-FP2 provides the mechanism to insert both the vortex phase mask and/or the LMS beamsplitter into the beam. The gvAPPs and Lyot stops will be located in IMG-PP1 and LMS-PP1, where the stabilization of the ELT-M1 pupil image will allow efficient use of these masks. The L - and M -band vortex phase masks and the image slicer in the LMS will be centered on the same optical axis to maximize the discovery space around the vortex when used in combination with the LMS. When combined with the gvAPP, the image slicer of the LMS will be fed with one of the two dark holes created by the gvAPP.

2.3 HCI Performance Requirements

The development of the METIS instrument is guided by a set of top-level requirements defined by ESO. Because it specifically aims for HCI, METIS has a formal requirement on the HCI performance. This requirement is expressed as a 5σ sensitivity in terms of contrast that needs to be reached after postprocessing of a 1-h observing sequence on a relatively bright star ($L \leq 6$).

The requirement, defined at the L -band, is to reach a postprocessed 5σ sensitivity of 3×10^{-5} in terms of contrast at an angular separation of $5\lambda/D$ (i.e., $0.1''$ at the L -band). It was derived by the ELT Project Science Team, based on preliminary results from the SPHERE instruments, with the goal to enable the detection of gas giant planets and super-Earths around nearby stars. To verify that this requirement can be fulfilled with the proposed instrumental concept, end-to-end simulations need to be performed.

Before embarking on end-to-end simulations, it is useful to break down this requirement. There are several sources of errors that contribute to the overall HCI performance. The most fundamental limits are photon noise associated with residual stellar light and with thermal background emission, and speckle noise associated with residual turbulence after SCAO closed-loop control. These contributions are detailed in Sec. 3. In addition, several sources of instrumental errors also affect the HCI performance budget, as described in Sec. 6. While these effects are studied one by one, in practice they add up and interfere with each other to create the final HCI performance, as discussed in Sec. 7. Considering the number of effects that need to be taken into account and the fact that they do not simply add up quadratically (they can, in some cases, constructively interfere), we will try to limit the influence of each individual contributor to 1/10th of the performance requirement, i.e., 3×10^{-6} at $0.1''$. This will be referred to as our “goal.”

3 METIS HCI End-to-End Simulator

Obtaining a realistic estimation of the METIS HCI performance requires performing end-to-end simulations, which must include three main steps.

1. Perform AO simulations to provide residual phase screens.
2. Propagate the residual phase screens through the instrument.
3. Produce a mock HCI data set in the pupil-stabilized mode and analyze it to evaluate the associated performance by applying an ADI postprocessing method to the image cube.

These three steps are detailed in the following sections.

3.1 Adaptive Optics Simulations

The quality of the adaptive optics correction provided by the SCAO module inside METIS will set the fundamental limits of its HCI performance. In the context of HCI observations, the SCAO module of METIS will be fed with a bright, on-axis natural guide star. Because HCI is at the heart of the METIS science cases, the SCAO system has been designed to provide the best possible performance for such guide stars. It is based on an infrared PyWFS³⁹ feeding a low-noise, high-speed, infrared electron-avalanche photodiode array. The PyWFS is expected to provide an overall better Strehl ratio than a standard Shack–Hartmann WFS²¹ but also a better sensitivity to low-order modes, which is key for HCI performance at small angular separations.

METIS will use ELT-M4 and M5 to implement the AO correction. Assuming a median seeing of $0.65''$ and a stellar magnitude of $K = 5$, the SCAO system is expected to deliver a high Strehl ratio: above 90% at the L -band and around 99% at the N -band.²¹ The Computing Platform for Adaptive optics System (COMPASS⁴⁰) has been used for this study to perform SCAO simulations. COMPASS is a (partly) open-source end-to-end adaptive optics simulation platform (<https://github.com/ANR-COMPASS/shesha>) built on a graphics processing unit (GPU) architecture to enable long ELT-scale simulations within a reasonable computation time. Using an NVIDIA Tesla K40M GPU, it is possible to simulate 10-min AO sequences with a resolution of 592-pixels square in about 10 h. The simulation parameters are summarized in Table 1. In addition to the parameters listed in that table, we use the tabulated C_N^2 and wind profiles provided by ESO for the Cerro Armazones site based on long-term Multi-Aperture Scintillation Sensor–Differential Image Motion Monitor (MASS-DIMM) site testing (internal communication). For the wavefront reconstruction, COMPASS uses a zonal approach controlling 4128 actuators. The reconstructor matrix is computed via a truncated Singular Value Decomposition (SVD) (where 88 eigenmodes are filtered out) and a simple integrator controller is used in our simulations.

Table 1 SCAO baseline parameters used to run long-simulations with COMPASS.

Parameter	Value
Pupil diameter	37 m
Central obstruction	11.1 m
Spiders	None
Segmentation	None
Target star magnitude	5 (<i>K</i> -band)
Zenith angle	30 deg
Fried parameter at 500 nm	0.157 m
Outer scale	25 m
Atmospheric layers	35
Wind direction	Random for each layer
Coherence time at 500 nm	5.35 ms
Sensing wavelength	2.2 μ m (<i>K</i> -band)
Number of subapertures along diameter	74
PyWFS modulation circle	$4\lambda/D$, sampled at 24 points
Wavefront sensor throughput	0.327
WF slope calculation	Ref. 41
Controller	Intergrator
High Order Deformable Mirror (HODM) ELT/M4 type	Piezo-stack array (Gaussian influence functions)
Tip Tilt Deformable Mirror (TTDM) ELT/M5	Yes
Loop frequency	1000 Hz
Total frame delay	2
Loop gain	0.4
Sampling	592 pixels across aperture (-6.25 cm/pixel)

For the purpose of the HCI performance tests, we ran a series of 10-min AO sequences and a 1-h long AO sequence. For each simulation sequence, we first ran the simulation for 1 s without recording the residual phase to make sure the AO loop is closed before recording the residual phase screens. We typically save one phase screen every 100 ms to feed the HCI simulations. In addition to the atmospheric turbulence, no other (potentially) variable effects, such as vibrations, noncommon path aberrations (NCPA), or observing conditions, have been added to the SCAO simulations. It is worth noting that the wind direction is assumed to be random in all atmospheric layers. This is an optimistic assumption, which reduces the strength of the wind-driven halo in AO-corrected images. Correlation between wind direction in various layers will be explored in a future work. All simulations are performed on a circularized version of the ELT pupil, corresponding to the largest all-glass inscribed circle of ~ 37 -m diameter. Spider arms were not included in the simulations to remove any residual piston effect between the six ELT petals. This effect will be studied separately in Sec. 6.2. A representative phase screen and the evolution of the standard deviation of the residual phase across the pupil for a 1-h simulation are shown in Fig. 4.

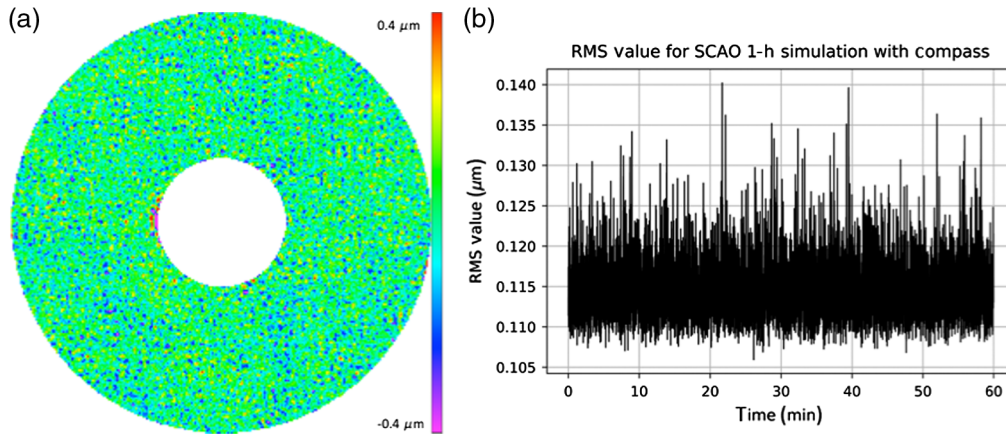


Fig. 4 (a) Typical AO residual phase screen. (b) rms phase across the pupil as a function of time for a 1-h SCAO simulation.

3.2 Optical Propagation

The optical propagation inside the instrument is performed using a custom Python package referred to as the HCI end-to-end performance simulator (HEEPS, <https://github.com/vortex-exoplanet/HEEP>), which relies on the PROPER⁴² optical propagation library. The PROPER library contains routines of optical propagation in the near-field and far-field conditions, using Fourier transform algorithms (Fresnel approximation). The implemented propagation layout is very simple, as it includes only the PPs and FPs where HCI elements are placed inside METIS (see Fig. 5). The input pupil of the coronagraph, defined by the ELT-M1 mirror, is reimaged at the CFO-PP1 pupil plane, where a ring apodizer can be inserted. CFO-FP2 is the FP containing the vortex phase mask. The Lyot stops and APPs are included inside the imager and spectrograph at PPs referred to as IMG-PP1 for the imager and LMS-PP1 for the spectrograph. Finally, the coronagraphic image is formed onto the detector, the second FP included in our simulations. The main simulation parameters are listed in Table 2.

The ELT-M1 contains 798 quasihexagonal segments, 1.45 m from side to side, with 4-mm gaps between them (see Fig. 6). Six 60-cm-wide spider arms divide the pupil into six sectors or petals. The pupil used for the HCI end-to-end simulations is the circular pupil inscribed in the ELT pupil (the all-glass pupil), with an external diameter of 37 m and an internal diameter of 11.1 m (30% central obstruction). Spider arms and pupil segmentation are taken into account.

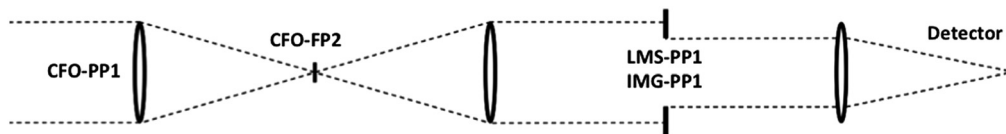


Fig. 5 Simplified scheme of the METIS HCI instrumental set up used in our simulations.

Table 2 Main parameters for the optical propagation simulations.

	L- and M-bands	N-band
Wavelengths	3.8 and 4.8 μm	8.7 and 11.5 μm
Grid size	1024 \times 1024 px	1024 \times 1024 px
Input pupil sampling	253 \times 253 px	253 \times 253 px
IMG pixel size	5.21 mas	10.78 mas

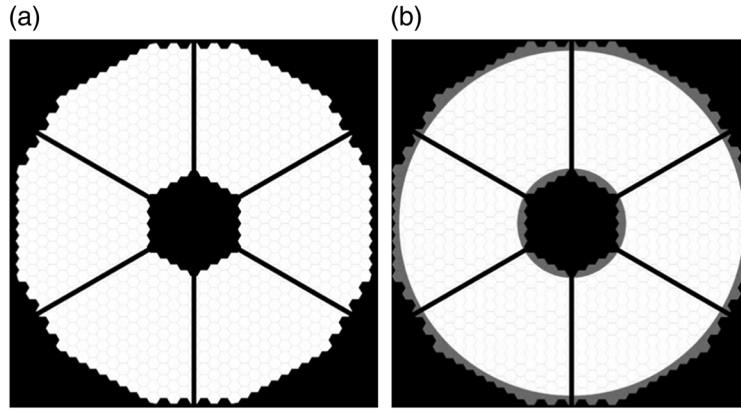


Fig. 6 (a) Geometrical description of ELT-M1, including spider arms. (b) Inscribed circular (all-glass) pupil used for our performance analysis.

Optical propagation of the residual SCAO phase screens produces noiseless coronagraphic PSFs, leading to a cube of 36,000 coronagraphic PSFs for a 1-h SCAO sequence sampled every 100 ms. In Sec. 6, we add instrumental errors in the optical propagation prescription, such as pointing errors, pupil alignment errors, or differential piston between the ELT pupil petals. All our simulations are performed at a single wavelength, but will be considered as representative of broadband observations. This inherently neglects all chromatic errors, and in particular the effect of atmospheric dispersion in the Earth atmosphere. Our simulations are, therefore, more representative of narrowband observations, even though we assume here that such a performance can be preserved in the broadband METIS filters defined in Table 3.

3.3 Mock Observations and Postprocessing

To turn the noiseless coronagraphic PSFs into a realistic observing sequence, we first need to define the main observing parameters. We will assume a standard HCI observing sequence with a bright, early-type star ($K = 5$, $L = 5$, and $M = 5$) observed for 1 h around meridian crossing, and culminating at a minimal zenithal distance of 20 deg. This observing sequence provides a total rotation of the parallactic angle of about 35 deg and would correspond more or less to the case of 51 Eri observed from the ELT site. It is worth noting that the variation of airmass during the observation is not captured in the AO simulation. This is not expected to be a strong contributor to the HCI performance, as the zenith angle varies only by 1 deg during the whole observing sequence. At the N -band, for the sake of illustration, we will assume a star as bright as Alpha Centauri ($N = -1.6$), while retaining the rest of the observation parameters.

Once the observing sequence is defined, producing a realistic mock observation requires to scale the noiseless coronagraphic PSFs produced above to the actual stellar flux in photoelectrons and to add the emission from the thermal infrared background. The METIS radiometric budget takes into account a model for the sky transmission (ESO SkyCalc) using standard

Table 3 Filter definition, total stellar flux for $L = M = N = 5$, and background flux per pixel (source: METIS simulator, SimMETIS, version 0.2, metis.strw.leidenuniv.nl/simmetis).

	Filter	Wavelength (μm)	50% width (μm)	Stellar flux (e^-s)	Background flux ($e^-/s/px$)
L-band	HCI-L long	3.82	0.27	9.00×10^8	8.88×10^4
M-band	CO ref	4.80	0.22	2.45×10^8	6.71×10^5
N-band	N1	8.65	1.17	2.98×10^8	9.63×10^7
N-band	N2	11.25	2.34	2.82×10^8	2.14×10^8

atmospheric conditions (2.4 mm of precipitable water vapor) and an airmass of 1.3, an emissivity of 23% for the telescope with an ambient temperature of 9°C, an emissivity of ~2% for the METIS entrance window, and an overall system throughput of ~36% for the METIS cold optics and detector. This leads to the star and background fluxes listed in Table 3 for the selected filters. It is worth noting that we do not use the full L - or M -band for HCI observations to reduce the effect of atmospheric dispersion, as well as the amount of speckle smearing in the innermost regions. Once the PSFs have been scaled to the appropriate flux and the thermal background added, we finally generate random shot noise independently on each pixel. Classical ADI processing⁴³ is used to process the mock data cubes using the open-source package VIP.⁴⁴ Classical ADI processing consists of removing the temporal median of the cube from each frame, derotating all frames to a common north, and averaging them to produce a final, postprocessed image. The main metric used in this document to evaluate the METIS HCI performance is the postprocessed 5σ sensitivity in terms of planet/star contrast for a 1-h observing sequence, also known as a “contrast curve.” The noise level in the final image is computed at a given angular separation from the star as the standard deviation of the aperture fluxes measured in as many independent resolution elements as can be defined at that separation. To evaluate the sensitivity limit of the instrument in terms of contrast, this noise level must then be compared to the non-coronagraphic signal of the host star, taking into account the throughput of the postprocessing algorithm. This throughput is evaluated by injecting fake companions at various positions into the raw data and by applying the same postprocessing algorithm to measure the amount of signal preserved in the final postprocessed image. The formula defining the 5σ contrast is the following:

$$c = \frac{5 N}{\tau S}, \quad (1)$$

where N is the noise, S is the noncoronagraphic stellar signal measured in an aperture of size equal to one resolution element (λ/D), and τ is the algorithm throughput. It is worth noting that the factor 5 is associated with a fixed false alarm probability (3×10^{-7}) for purely Gaussian noise. The small number of samples at short angular separations modifies the relation between false alarm probability and noise level.⁴⁵ To keep the same false alarm probability at all separations, this requires modifying the definition of contrast, taking into account a t -Student distribution instead of a standard Gaussian distribution. The penalty associated with small sample statistics is taken into account in all contrast curves presented in this document.

A second contrast-related metric useful to evaluate the HCI performances is the raw contrast, which is defined, at any given separation, as the ratio between the intensity of the considered PSF (be it coronagraphic or not) and the peak of the off-axis PSF with the same METIS configuration (Lyot stop and apodizer). It indicates the mean level of residual starlight at a given separation. Raw contrast can readily be assessed by plotting the PSF profile of interest in a figure where the direct PSF profile culminates at 1 on the optical axis. This metric does not tell what is the faintest source that can be detected as it is not a direct indicator of noise level. Yet, raw contrast can be useful to assess a coronagraph’s capability to cancel stellar light. In the case of the APP, this estimation is done on the clean half of the field of view.

4 SCAO Simulation Setup

Before presenting the nominal HCI performance, we need to study how the duration and time sampling of the AO residual phase screen sequence affects the estimated performance to select a baseline set of AO simulation parameters that provide robust performance estimation within a reasonable computation time. All simulations performed in this section are done for a RAVC operating at the L -band, because it is supposed to be the best-performing HCI mode for which the influence of the AO sequence can most readily be assessed. To reduce computation time, we have generally cropped the field of view to a radius of about $0.6''$, which produces individual frames of about 240×240 pixels based on the sampling of 5.21 mas/pixel for the LM -band images detector. We do not include background noise or photon noise in this section to highlight the effect of pure atmospheric residual speckle noise. Background noise will be included and

discussed in Sec. 5, where we show that it creates a performance plateau at large angular separations, drowning the actual effect of SCAO residuals at these separations.

Ideally, we would like to use a 1-h SCAO simulation to match the typical duration of an ADI observing sequence, with a sampling down to a few tens of ms to properly capture the evolution of atmospheric turbulence in the thermal infrared where the atmospheric coherence time τ_0 ranges from a few tens of ms to a few hundred of ms depending on weather conditions and wavelength. A 10-ms sampling would, however, result in 360,000 phase screens for each individual end-to-end simulation, which is not acceptable in terms of computation time. To perform a parametric study of the SCAO performance and of the instrumental noise sources within a reasonable time, we need to reduce the computation time associated with one single end-to-end simulation. We, therefore, investigate the influence of the duration and sampling of SCAO simulations on the ADI contrast curves.

4.1 Duration of the SCAO Sequence

Assuming a 1-h-long simulation as a baseline, the best sampling that we can use while maintaining a reasonable CPU time and a reasonable RAM usage is about 100 ms, which gives a total of 36,000 phase screens. The computation time is then around one day for the SCAO simulation on our GPU server, about 10 h for the optical propagation on a multicore CPU server, and a couple of hours for the contrast curve computation. The RAM usage for the optical propagation and contrast curve computation can be kept below the 128 Gb available on our CPU server.

To investigate the effect of the SCAO sequence duration, we produce two other sequences, respectively, 1- and 10-min long, with a similar number of phase screens. For that, we save phase screens every 2 ms in the first case and every 20 ms in the second case, which results in 30,000 phase screens in both cases. This largely reduces the time required to produce the SCAO sequence, while the CPU time required for optical propagation does not change significantly due to the similar number of phase screens. To simulate a 1-h-long ADI sequence based on these two shorter SCAO sequences, we artificially stretch them in time so that the time between samples becomes 120 ms instead of 2 or 20 ms. All three sequences are then postprocessed with the same classical ADI algorithm to produce contrast curves. This is shown in Fig. 7, where we note that using the 10-min sequence results in a degradation of the contrast by less than a factor 2. Conversely, going to a 1-min sequence results in a very significant degradation of the contrast, by a factor around 5. Stretching a SCAO sequence does not preserve the timescale of correlation between successive phase screens (or equivalently, the shape of the temporal PSD), and so also between the associated coronagraphic PSFs, degrading the contrast as the SCAO sequence is shortened. Using longer time sequences reduces the overall amount of correlation in the phase screens.

The conclusion of this comparison is that while 1-h-long sequences should ideally be used to assess HCI performance, using 10-min sequences can also be tolerated in cases where the

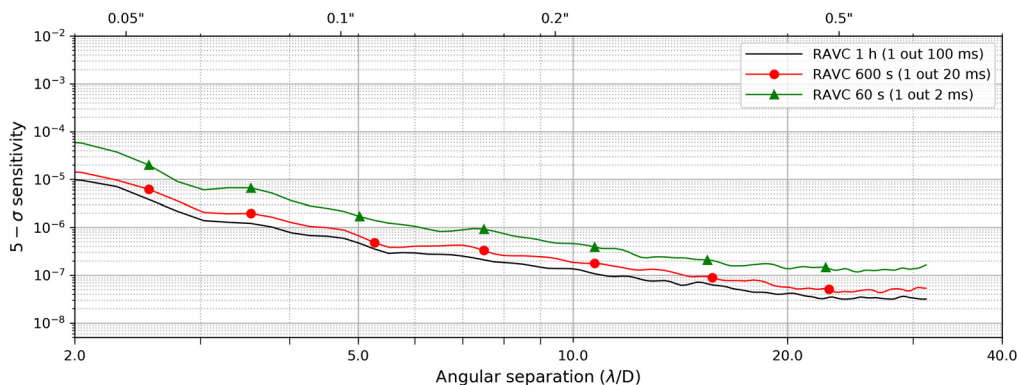


Fig. 7 Comparison of contrast curves obtained for the RAVC at the L -band with three SCAO sequences of different duration and sampling: 1 h with 100-ms sampling (black), 10 min with 20-ms sampling (red), and 1 min with 2-ms sampling (green).

computation time for the SCAO simulation becomes prohibitively large. In particular, testing various SCAO configurations/parameters with 1-h-long sequences is impractical and will generally be restricted to 10-min-long sequences.

4.2 Sampling of the SCAO Sequence

We now study the effect of sampling on the final postprocessed contrast. While SCAO simulations are always performed with a 1-ms sampling, one can choose to save the phase screens to be propagated in our end-to-end simulations at a slower pace. Reducing the number of phase screens improves the computation time for optical propagation and contrast curve calculation. To illustrate the effect of sampling, we use our two longest SCAO simulations, respectively, spanning 10 min and 1 h. For the former, we originally saved one phase screen every 20 ms to produce a cube of 30,000 phase screens, while for the latter, we originally saved one phase screen every 100 ms to produce a cube of 36,000 phase screens. In Fig. 8, we explore how the final postprocessed contrast is affected when downsampling these series of phase screens in the time domain. From the 10-min simulation [Fig. 8(a)], we conclude that downsampling to one phase screen every 100 ms still produces an acceptable result in terms of contrast level (degradation by less than a factor 2). This suggests that reducing the number of phase screens by a factor of 5 does not significantly reduce the amount of diversity in the produced coronagraphic PSFs, which seems to be in line with a coherence time of up to about 100 ms at the *L*-band. The 1-h simulation [Fig. 8(b)] confirms that a coarser sampling than 100-ms starts to take a significant toll on the estimated performance, although we note that the 300-ms sampling on the 1-h sequence provides a final contrast similar to the 10-min sequence with 20-ms sampling due to the improvement related to extending the duration of the sequence from 10 min to 1 h. The degradation in performance associated with a downsampling from 100 to 300 ms is still generally not larger than a factor of 2, which we deem acceptable.

In conclusion, to reduce the required computation time for the following study, we will use either a 1-h-long sequence with 300-ms sampling (12,000 phase screens, ~ 4 h for the optical propagation and a couple of hours for the contrast curve calculation) or a 10-min-long sequence with 100-ms sampling (6000 phase screens, ~ 1 h for the optical propagation and half an hour for the contrast curve calculation). The former will only be used to study instrumental effects

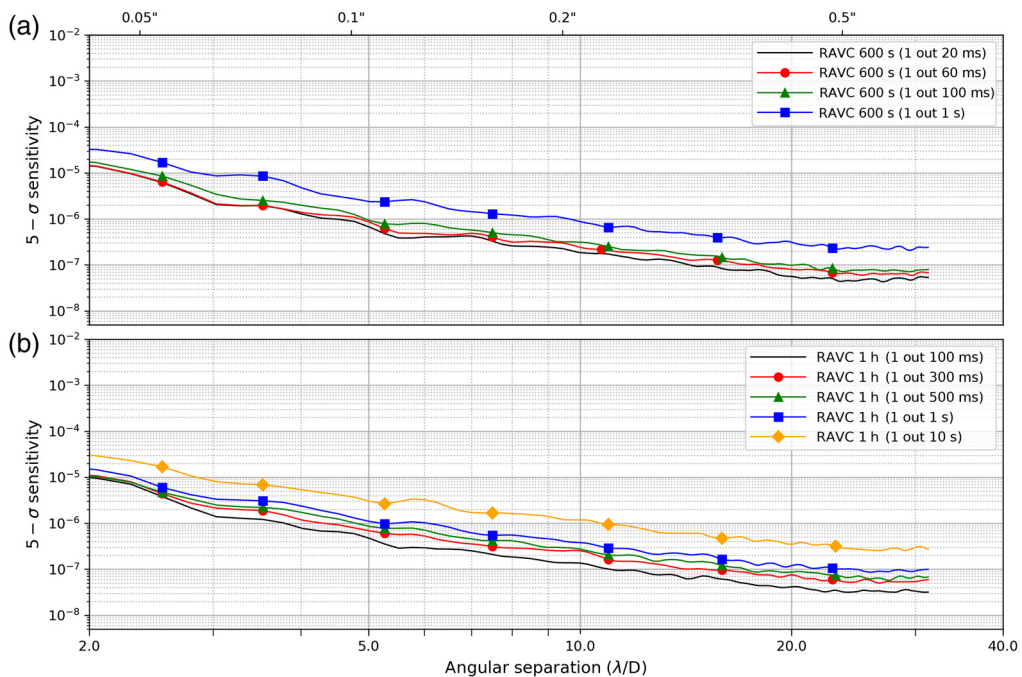


Fig. 8 Comparison of contrast curves obtained with the RAVC at the *L*-band with various sampling times for two SCAO sequences: (a) 10-min long and (b) 1-h long.

that do not require to produce new SCAO simulations (i.e., effects associated with the coronagraphs), while the latter will be used to explore the influence of SCAO performance (the computation time goes from about one day for the 1-h sequence to about 10 h for the 10-min one).

5 HCI Reference Performance in Presence of SCAO Residuals

In this section, we describe the expected HCI performance of METIS in the presence of photon noise associated with stellar residuals and to thermal background for typical stellar magnitudes, taking into account SCAO residual phase screens as the only perturbation to the system, assuming the instrument to be otherwise perfect. The 5σ sensitivity limits in terms of contrast produced here will be used as a reference for the exploration of instrumental and environmental parameters in the next sections. They will also be used in the future to assess the improvements that can be brought by advanced AO control strategies (e.g., Kalman filtering and predictive control), which are not considered in this work. We remind that the performance presented here pertains to a 1-h observing sequence and to a given stellar brightness. Longer observing sequences and/or brighter stars would decrease the contribution of photon noise and would bring the HCI performance closer to the SCAO-limited (i.e., speckle noise-limited) performance at all separations.

5.1 L-Band Performance

Using our 1-h SCAO sequence with 300-ms sampling, we perform a first set of simulations at the *L*-band ($3.8 \mu\text{m}$) for the three main coronagraphic concepts considered for METIS: the CVC, the RAVC, and the APP. In Fig. 9, we show representative instantaneous PSFs for the three coronagraphic concepts, as well as their long-integration PSF, averaged over the full 1-h sequence. This figure clearly highlights the well-corrected region corresponding to the spatial frequencies that can be controlled by the ELT deformable mirror (about $80\lambda/D$ across). It also highlights the

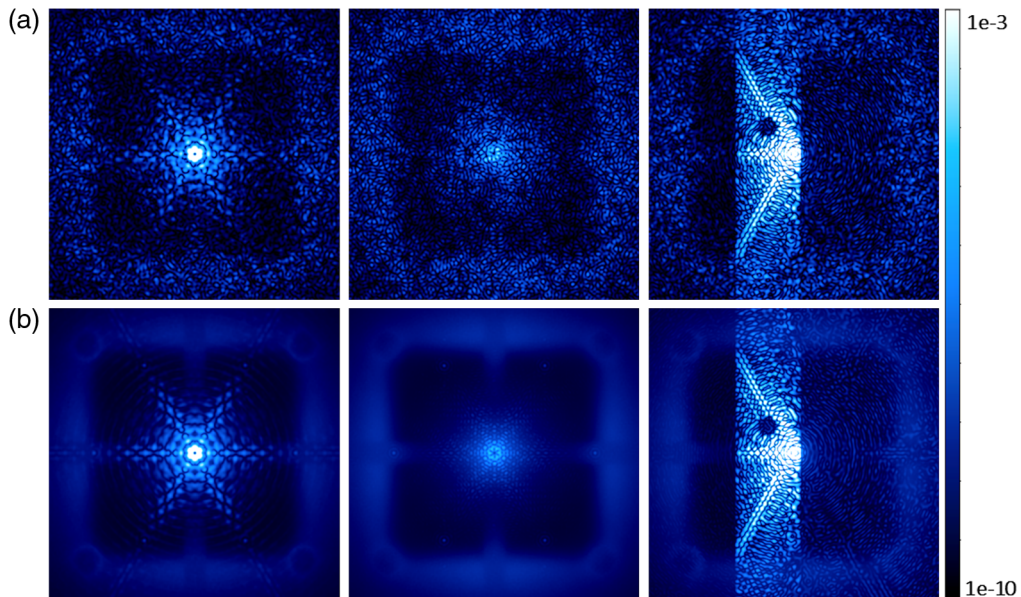


Fig. 9 (a) Representative instantaneous PSFs at the *L*-band, using a single SCAO residual phase screen obtained in the cases of CVC (left), RAVC (center), and APP (right). (b) Same for long-exposure PSFs, averaged over the 1-h SCAO sequence. All images have been cropped to a field of view of $100\lambda/D$, i.e., $2.1''$ at the *L*-band. All PSFs have been normalized by the maximum intensity in the noncoronagraphic PSF of the corresponding HCI mode. The dark spot in the APP PSF was purposely created on the optical axis to enable high-SNR measurements of the so-called leakage term (not included here), a scaled-down version of the stellar PSF due to imperfection in the phase shift induced by liquid crystals, which can be used to monitor the photometry and astrometry of the central star without saturating the detector.

amount of starlight remaining in the science FP, ranging from full stellar light in the case of the APP to very low starlight for the RAVC. The mean total stellar leakages for the CVC and RAVC in presence of AO residuals are around 8% and 3%, respectively. We remind that these figures are due to a combination of diffraction effects in the telescope with aberrations due to residual atmospheric turbulence. In the absence of any perturbation, the total stellar leakage for these two configurations, due exclusively to diffraction effects associated with the shape of the input pupil (and to the Lyot stop design), would amount to 3% (CVC) and 0.2% (RAVC). The results presented here are valid for a sufficiently narrow spectral filter (bandwidth typically smaller than 10%), and for sufficiently compact stars (angular diameter smaller than 4 mas), in which case the effect of atmospheric dispersion and of partly resolved stellar surfaces on coronagraphic performance is reduced below the level of stellar leakage associated with pure atmospheric turbulence.

The raw PSF profiles (aka raw contrast) and the 5σ sensitivity limits in terms of flux ratio after ADI postprocessing (aka contrast curves) are shown in Fig. 10. In the case of the APP, the raw and postprocessed contrasts were computed by combining the dark holes associated with the two coronagraphic PSFs of the gvAPP. This leaves, in principle, a bright stripe of speckles in between the two dark holes. Here, we have artificially dimmed this bright stripe to facilitate the contrast computation. The contrast curves for the APP are thus somewhat optimistic, especially at the smallest angular separations, where the bright stripe has the largest relative influence. A more advanced way to compute contrast curves for the APP will be implemented in the future. We note that the RAVC, CVC, and APP meet our goal of 3×10^{-6} contrast at $5\lambda/D$, while the noncoronagraphic mode does not meet that requirement.

While the raw contrast provided by the various coronagraphs is significantly enhanced compared to noncoronagraphic imaging, the improvement in terms of ADI contrast curve may seem small, especially in the presence of thermal background and photon noise (dotted lines). Only for the smallest angular separations do coronagraphs (and the RAVC in particular) provide a large gain. This situation is bound to change in the presence of instrumental errors. While random atmospheric speckles efficiently average out over the 1-h observing sequence, instrumental

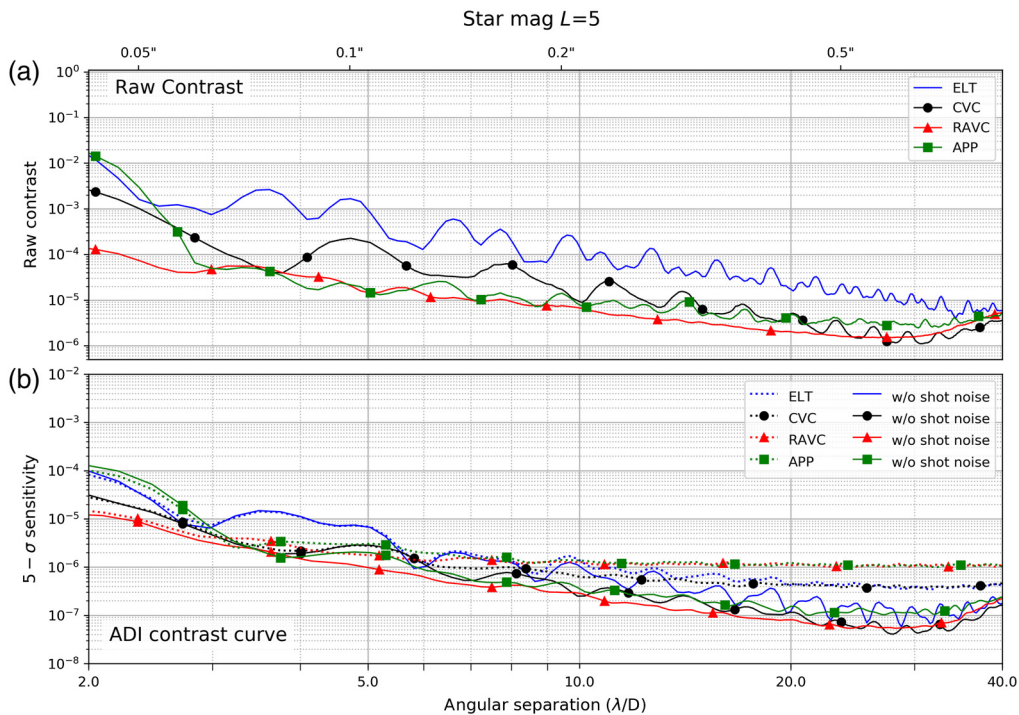


Fig. 10 (a) Raw contrast and (b) ADI contrast curve for noncoronagraphic (ELT), CVC, RAVC, and APP observing modes at the L -band. In (b), the solid curves include only the contribution of speckle noise, while the dashed curves include the effect of the thermal background (shot noise) for a star of magnitude $L = 5$.

errors are expected to create long-lived speckles, whose intensity is modulated by (or “pinned” to Ref. 46) the stellar diffraction pattern. The coronagraphic modes, which produce a much lower level of diffracted stellar light (see Fig. 9), are thus expected to provide significantly improved performance with respect to noncoronagraphic modes when these additional effects are taken into account. It must also be noted that the contrast curves do not take into account detector saturation effects, which would significantly limit the discovery space at small angular separations from bright stars in the noncoronagraphic case. Since the noncoronagraphic contrast curve is a purely hypothetical performance that can never be reached, it will be omitted for the rest of this work.

It is noteworthy that the sensitivity in terms of contrast becomes dominated by thermal background noise beyond about $0.2''$, and for the stellar magnitude $L = 5$ considered here. Further simulations show that for $L = 8$, the transition between the speckle-dominated and the background-dominated regimes shifts to about $0.1''$ (i.e., $5\lambda/D$). Using a coronagraph for fainter targets becomes less and less useful. We note that the APP contrast takes into account the fact that the planetary signal is divided into two conjugated PSFs by the phase plate and that its Strehl ratio is reduced by the redistribution of the light required to create a dark hole. The background-limited contrast of the APP is, therefore, similar to the RAVC’s despite its intrinsically higher throughput. The CVC provides the best contrast in the background-limited regime, thanks to its high throughput.

5.2 M-Band Performance

The same analysis as in the L -band is carried out at the M -band in Fig. 11. The raw contrast shows a significant improvement compared to the L -band, because all phase errors have a smaller effect at longer wavelengths. The ADI contrast curves computed in the absence of thermal background and photon noise improve accordingly. However, in the presence of thermal background and photon noise (dotted lines), the performance is largely decreased due to the higher background level at the M -band. This also means that the RAVC becomes less useful at this wavelength and will only be used for stars significantly brighter than $M = 5$.

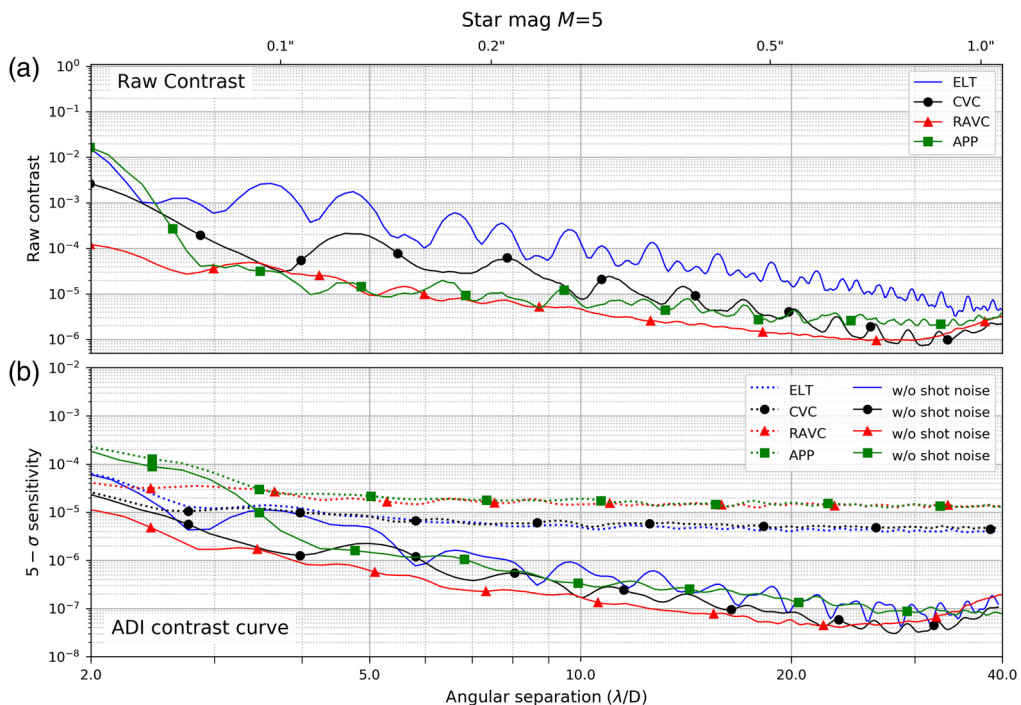


Fig. 11 Same as Fig. 10 for the M -band in the case of a star of magnitude $M = 5$.

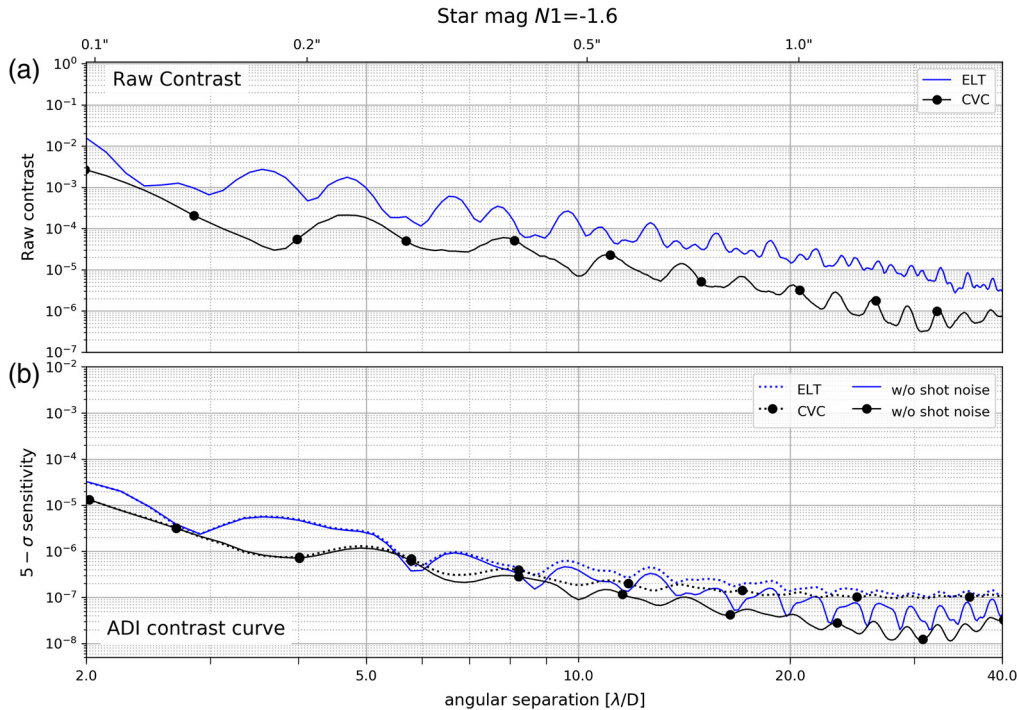


Fig. 12 Same as Fig. 10 for the $N1$ -band (around $8.7 \mu\text{m}$) in the case of a star of magnitude $N = -1.6$.

5.3 N -Band Performance

Because of the large thermal background, coronagraphy at N -band will only be useful for the brightest stars. Here, we focus on the case of α Cen A, with magnitudes $K = -1.5$ and $N = -1.6$. For the sake of simplicity, we keep the same set of SCAO residual phase screens as described in Sec. 4, recognizing that photon noise is not dominant for stellar magnitudes $K \leq 5$ so that we do not expect a significant improvement in terms of SCAO residuals compared to the case described in Sec. 4. We also keep the same trajectory in the sky as in the previous section, although we recognize that a 1-h observing sequence on α Cen would actually lead to a parallactic angle variation of only about 23 deg (instead of 35 deg).

The contrast at 8.7 (aka $N1$ -band) and $11.5 \mu\text{m}$ (aka $N2$ -band) is displayed for α Cen A in Figs. 12 and 13, respectively. These figures show that, even for one of the brightest stars in the sky, the performance is background limited for angular separations larger than about $0.5''$. Because the RAVC largely decreases the total throughput, this mode is not considered at the N -band. Indeed, additional simulations show that, even for a bright star such as α Cen, the RAVC would improve HCI performance only within about $0.1''$. A simple CVC is preferred at the N -band to provide the highest possible throughput while still reducing the total amount of stellar light from the host star by a factor ~ 30 , and the peak intensity in any pixel by a factor slightly larger than 100. It is worth noticing that, at longer wavelengths, the difference between direct imaging and coronagraphy is somewhat reduced, but still significant at the small angular separations where it matters the most. Additionally, coronagraphy will still help avoid saturation of the detector for bright stars such as α Cen.

6 Influence of Instrumental and Environmental Effects

In this section, we investigate the influence of the most prominent contributors to the HCI performance budget. This includes (i) instrumental and environmental effects appearing within the SCAO, such as the influence of turbulence conditions and of noncorrected atmospheric effects (Sec. 6.1), residual petal piston (Sec. 6.2), or residual pointing jitter (Sec. 6.3), (ii) instrumental effects appearing within the METIS instrument itself, such as pointing drifts between the SCAO

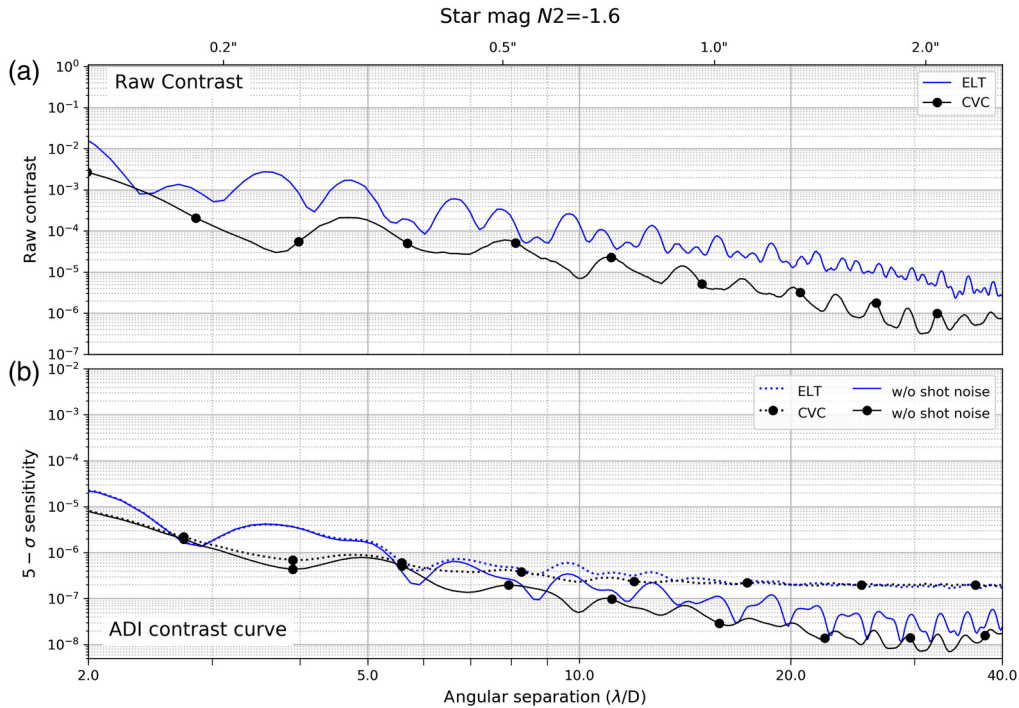


Fig. 13 Same as Fig. 10 for the N2-band (around 11.5 μm) in the case of a star of magnitude $N = -1.6$.

and coronagraphic arms (Sec. 6.3) or higher-order NCPA (Sec. 6.4), and (iii) instrumental effects appearing within the telescope, such as the shape of the input pupil (Sec. 6.5) or drifts in the input pupil position (Sec. 6.6). We note that beyond seeing quality, some other effects have an influence on SCAO performance, and hence on HCI performance. These effects are described in Ref. 21, but will not be detailed here for the sake of conciseness. This includes the effects of mirror segmentation, cophasing, and amplitude aberrations in the M1 pupil, which are mostly constant over the duration of an ADI observation, and will, therefore, cancel out during post-processing. Some additional instrumental effects, such as (residual) atmospheric dispersion or defects in the coronagraphic devices themselves, will not be discussed here as they can be mitigated through careful design and manufacturing.

Here, and for the rest of this study, we consider L -band simulations at 3.8 μm as a baseline to explore the influence of instrumental and environmental error sources and quantify their effect on HCI performance. The reason for using the L -band is that on one hand, the contrast requirement for METIS is defined at the L -band, and on the other hand, instrumental effects are supposed to have the highest relative influence at the shortest wavelengths. We will focus on the RAVC mode, which shows the best performance in the atmospheric residuals speckle-dominated regime, and is also, therefore, expected to show the strongest (relative) dependence on instrumental errors. All contrast curves from now on do not include the influence of the thermal background, as we are interested in the speckle-limited contrast here. Adding the effect of thermal background would only conceal the effects that we aim to evaluate. When discussing the performance of HCI modes, we should, however, bear in mind that a large part of the search region will actually be background-dominated for most stars. The need for reaching very deep contrasts in these regions will be weighted accordingly.

6.1 Atmospheric Effects

One of the main parameters influencing the SCAO performance is the strength of atmospheric turbulence. Here, we study the influence of the Fried parameter r_0 , or equivalently of the seeing. In Fig. 14, we show the ADI postprocessed contrast for a series of 10-min COMPASS simulations obtained under various seeing conditions ranging from excellent (0.45") to poor (0.95").

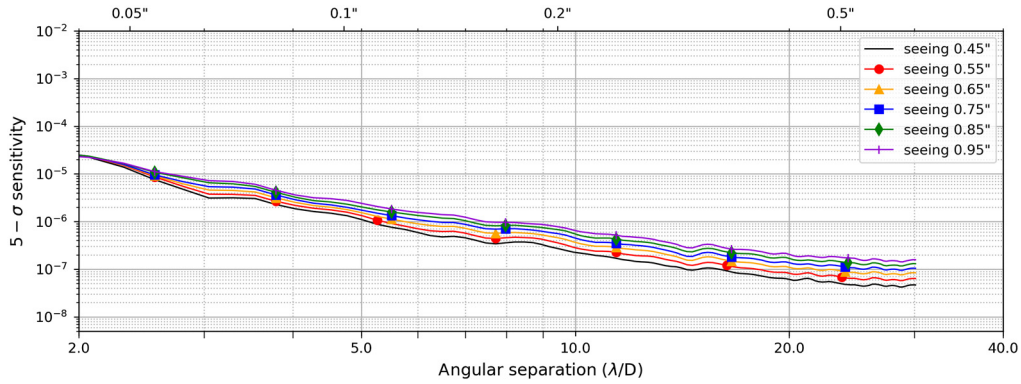


Fig. 14 ADI postprocessed contrast for the RAVC at the L -band for various seeing conditions, considering only the contribution of AO residuals speckle noise (no photon noise considered here and in the next following figures).

While seeing has a significant influence on HCI performance, we note that performance at small separations is weakly affected. Even the worst seeing explored here does not break our goal of 3×10^{-6} at $5\lambda/D$. This suggests that, provided that the SCAO loop remains robust under such conditions (which seems to be the case based on our COMPASS simulations), HCI observations could, in principle, be carried out most of the time with METIS at the ELT.

In addition to pure SCAO residuals, atmospheric turbulence has a series of other consequences on HCI performance, as discussed in Ref. 47. We briefly review these effects below and argue that they can generally be neglected in our HCI performance study.

- Scintillation is due to phase aberrations turning into amplitude aberrations due to Fresnel propagation through the Earth’s atmosphere. Reference 47 gives an expression for the raw coronagraphic contrast associated with a perfect coronagraph in the presence of atmospheric scintillation [C_1 term defined in their Eq. (17)]. Plugging in the METIS parameters and Cerro Armazones standard atmospheric conditions in this expression, we get a raw contrast ranging from 2×10^{-6} to 6×10^{-6} for angular separations from 1 to $40\lambda/D$. Comparing with the SCAO-limited raw contrast (Figs. 10–13), we conclude that scintillation is negligible at the angular separation where the contrast requirement is set ($5\lambda/D$). It becomes non-negligible only for angular separations larger than $10\lambda/D$.
- Wavefront chromaticity refers to the fact that Fresnel propagation is chromatic, leading to wavelength-dependence in wavefront aberrations. When perfectly corrected at one wavelength (the WFS wavelength), the wavefront at the science wavelength will show a small residual that limits the achievable contrast, as described by the C_4 term in Eq. (24) of Ref. 47. Plugging the METIS and Cerro Armazones parameters into this expression, it can be shown that the contribution of wavefront chromaticity remains below 10^{-6} in terms of raw contrast inside the AO control radius of $\sim 40\lambda/D$. It can, therefore, be neglected.
- Refraction index chromaticity, and more specifically the effect of water vapor on the refractive index chromaticity, is a relatively poorly constrained atmospheric effect. Reference 24 discusses its influence on the performance of future mid-infrared ELT instrumentation. While the degradation in wavefront quality is expected to be significant under standard atmospheric conditions, with an estimated WFE around 40-nm rms at L -band and 300-nm rms at N -band, a thorough investigation of water vapor seeing is beyond the scope of this paper. Here, we will assume that N -band observations will only be performed during episodes of very low precipitable water vapor (<1 mm), which typically happen 12% of the time each year at Cerro Paranal.⁴⁸

In addition to atmospheric turbulence, another contribution of the Earth’s atmosphere is refraction, which is chromatic and, therefore, leads to dispersion in the resulting broadband PSF. For instance, Ref. 49 shows that a Strehl loss of about 6% can be expected at the N -band on an ELT-class telescope for an airmass of 1.5—an airmass above which HCI observations will generally not be attempted with METIS. The Strehl loss would be even larger at the L -band,

which justifies our decision to add an ADC in METIS for that wavelength range. While the Strehl loss associated with atmospheric dispersion may look critical for HCI observations, the variations in PSF shape during a typical 1-h ADI sequence are actually quite small. We verified through dedicated end-to-end simulations that even in extreme cases such as a 3-h ADI sequence on α Cen (airmass around 1.25), the contribution of atmospheric dispersion will be negligible in the final ADI-processed contrast curves compared to the other effects described below for the standard HCI filters of METIS (Table 3).

6.2 Petal Piston

As for most ground-based telescopes, the ELT pupil is divided into a number of separated regions (referred to as petals) by the support structure of the secondary mirror (spider arms). The presence of spiders has recently been shown to be at the origin of the “low wind effect,” which creates additional aberrations at low spatial frequencies that are not well detected by standard Shack–Hartmann PyWFSs.⁵⁰ Even worse, the discontinuity of the pupil can lead to an artificial buildup of piston error between petals in closed-loop AO operations. Fortunately, the near-infrared PyWFS used in METIS is more robust to this petal piston mode, and specific reconstruction and control strategies are being implemented to mitigate this effect.⁵¹

To evaluate the effect of petal piston on the coronagraphic performance, we introduced differential pistons between petals in the SCAO phase screens with various amplitudes (10, 20, 50, and 100 nm). We first introduced a static petal piston, whose main effect is to increase the intensity level of the diffraction pattern, especially at small angular separations. However, this additional contribution is not significantly larger than the starlight level. The resulting effect on the contrast curves is negligible, because the ADI postprocessing captures the static part of the stellar diffraction pattern so that we are just left with a small enhancement on the speckle intensity in the post-ADI images. We, therefore, only consider the case of variable petal pistons, assuming a sinusoidal behavior for the mean phase in all petals. To avoid coherent effects between petals, the phase of the sinusoid is chosen to be different for the pistons in all petals, and the amplitude is slightly varied from one petal to the other. Two different timescales were chosen to inject petal pistons into our 1-h SCAO residual phase screen sequence, with periods of ~ 1 s [Fig. 15(a)] and

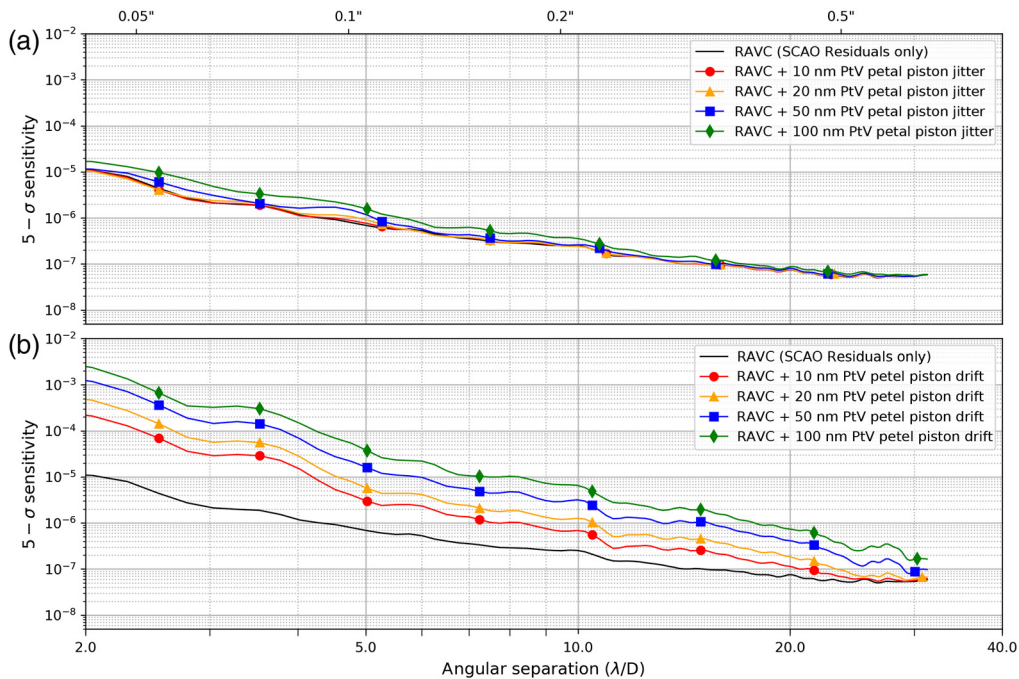


Fig. 15 ADI contrast curve for HCI performance in the presence of variable pistons between the six ELT petals, (a) for the case of piston jitter at 1 Hz, and (b) petal piston drift with a drift timescale of 1 h, considering only the contribution of AO residuals speckle noise.

1 h [Fig. 15(b)]. While the effects of high-frequency variations of petal piston are largely averaged out by the ADI postprocessing, slow drifts have a very detrimental effect on the final contrast. Our simulations indicate that a high-frequency jitter of up to 100-nm PtV is acceptable, while low-frequency drifts should be limited to about 10-nm PtV. Preliminary simulations of advanced wavefront reconstruction with the PyWFS suggest that a drift of around 20-nm PtV is within reach.⁵²

6.3 Pointing Errors

FP phase mask coronagraphs, such as the VC, are known to be particularly sensitive to low-order aberrations, and to tip-tilt errors in particular.⁵³ Tip-tilt errors create additional leakage of stellar light through the Lyot stop, thereby reducing the achievable contrast.

Pointing errors can be divided into two main parts: high-frequency pointing jitter (typically >0.1 Hz) and low-frequency pointing drifts (typically <0.1 Hz). In the former case, random deviations from perfect alignment of the star with the center of the vortex phase mask create rapid fluctuations of the stellar leakage, which increases the starlight-related photon noise. When telescope vibrations are not taken into account, SCAO simulations only contain a very small amount of pointing jitter (around 0.1 mas). Taking into account telescope vibrations could increase this jitter beyond 1 mas rms.²¹ To explore the influence of pointing jitter, we have added an additional high-frequency random jitter to the SCAO phase screens, with four different levels: 0.5, 1, 2, and 5 mas rms, using the 1-h sequence sampled at 300 ms. Figure 16(a) shows that the influence of high-frequency pointing jitter appears mostly at short angular separations and averages out in the ADI postprocessing to produce an insignificant performance reduction.

While the influence of high-frequency pointing jitter is minor on the postprocessed contrast performance, slow pointing drifts are expected to have a very large impact because they are not easily removed by ADI postprocessing. Slow pointing drifts are not expected to be due to SCAO itself, as it will provide excellent stability of the set point, but rather due to the noncommon path between SCAO and the coronagraph, including the effect of differential atmospheric refraction between the K -band (where SCAO operates) and longer wavelengths (where coronagraphic observations are performed). Quasistatic pointing errors can even lead to systematic biases

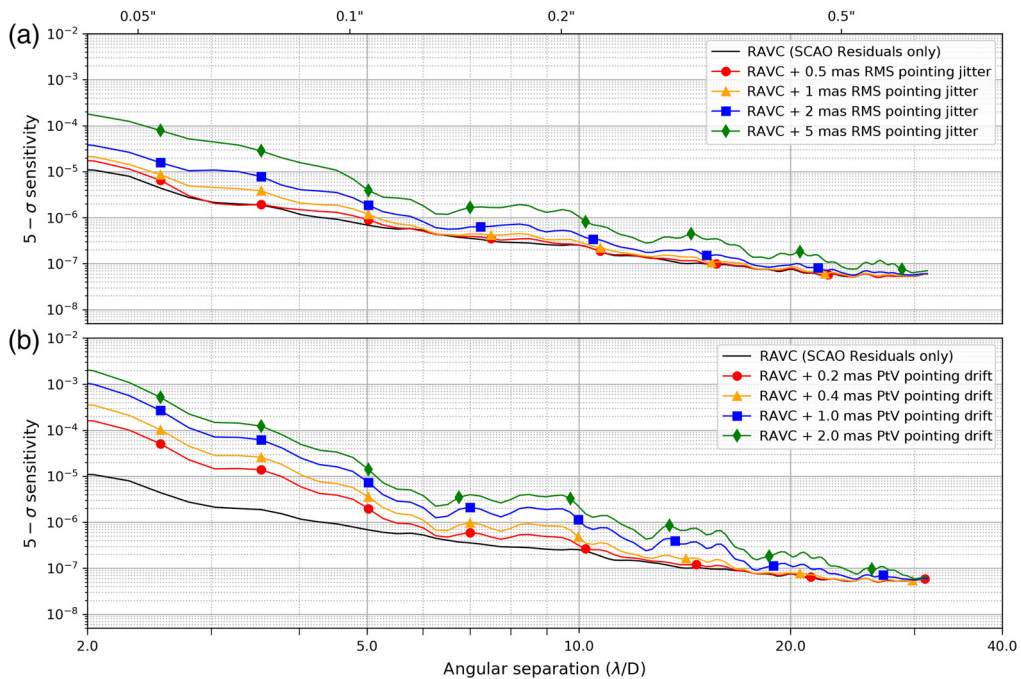


Fig. 16 (a) ADI contrast curve for HCI performance for the RAVC at the L -band in the presence of pointing jitter and (b) various PtV amplitudes of low-frequency pointing drift considering only the contribution of AO residuals speckle noise.

in the final postprocessed images, e.g., by making one side of the image brighter than the other. Such errors can be associated with mechanical drifts in the coronagraphic components (e.g., vortex phase mask) or to imperfect predictions of the effect of atmospheric refraction on the PSF position at the vortex phase mask location. This is the reason why common-path pointing control will be implemented in METIS using the QACITS algorithm.⁵⁴ This algorithm derives the centering error by analyzing the shape of the coronagraphic PSF on the science camera. To investigate the influence of residual noncommon path pointing errors after QACITS closed-loop control, we have added low-frequency pointing drifts to the SCAO phase screens with a timescale equal to the full observing sequence and with PtV amplitudes ranging from 0.2 to 2 mas. The star was set to move along the horizontal axis during the AO sequence. This represents a worst-case scenario, because the timescale of the offset perfectly matches the timescale of the ADI sequence. We implicitly assume here that our pointing control algorithm (QACITS) does not produce any significant bias on the centering of the star, which should be the case for a good alignment between the input pupil and the Lyot stop.⁵⁵

Low-frequency pointing drifts have a huge impact on the final contrast [see Fig. 16(b)]: a PtV amplitude of 0.4 mas would already marginally break our goal of 3×10^{-6} at $5\lambda/D$. This corresponds to $0.02\lambda/D$ PtV at the L -band. Our previous experience with QACITS at Keck/NIRC2⁵⁵ and our first results obtained in the framework of the NEAR experiment on VLT/VISIR (Maire et al.)⁵⁶ show that the closed-loop pointing stability using feedback from the science camera can go down to about $0.01\lambda/D$ rms (i.e., $\sim 0.025\lambda/D$ PtV) using a simple integrator with a control frequency around 0.1 Hz. While this is slightly larger than the requirement to keep the METIS performance compliant with our goal, it must be noted that speckles associated with temporal frequencies higher than 0.001 Hz will average out to some level during the ADI post-processing, and that the contribution of frequencies below 0.001 Hz to the total rms pointing error should be significantly smaller than $0.01\lambda/D$ due to the integral control.

Performing a detailed study of the frequency content of residual pointing errors and of the QACITS control loop is beyond the scope of this paper. Confirming that QACITS is capable of reaching 0.4-mas PtV at L -band on a timescale of an hour will need to be investigated with more advanced simulations in the future. Here, we recognize that pointing drifts could be one of the main contributors to the final performance budget, but that, based on our best knowledge of the QACITS performance, pointing drifts alone should not jeopardize the fulfillment of the contrast requirement.

6.4 Noncommon Path Aberrations

NCPAs are known to be a major contributor to HCI performance budgets and are supposed to be the limiting factor to the performance of many state-of-the-art instruments under good observing conditions and AO correction. In this section, we focus on aberrations of higher order than tip and tilt, which were discussed in Sec. 6.3. To mitigate the effect of NCPA on the METIS HCI performance, we plan to use a common-path wavefront sensing/correction technique, referred to as phase sorting interferometry (PSI⁵⁷). With this algorithm, NCPAs are sensed in real time using the data from the IMG camera and the telemetry from the PyWFS, and fed back to the SCAO system for correction through WFS slope modifications. In this situation, the NCPAs measured by PSI are transferred to the SCAO, which then operates around a nonzero wavefront error. This may lead to a decrease in wavefront control performance for the SCAO. Our first SCAO simulations in presence of NCPA suggest that this is not the case under the expected level of NCPA, and we, therefore, focus on the direct consequences of NCPA on the coronagraphic performance. To simulate the influence of NCPA, we performed a series of Monte Carlo simulations of the wavefront aberrations between the PyWFS inside SCAO and the IMG-PP1 pupil plane, where the Lyot stops and APP are placed inside the IMG camera. These simulations take into account our current knowledge of the METIS optical train at preliminary design review. An example of such an NCPA phase map is shown in Fig. 17, and will be used in the rest of this section.

In closed-loop operations, PSI will typically be able to measure and correct for spatial frequencies up to 10 cycles per pupil, although this figure depends on the fine-tuning of the algorithm and of the filter bandwidth, which is beyond the scope of this paper. This means that static NCPA appearing at spatial frequencies beyond 10 cycles/pupil will not be corrected.

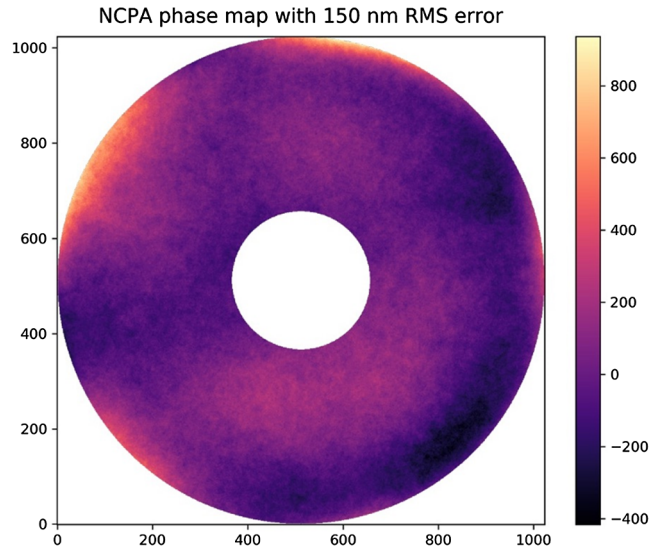


Fig. 17 Representative NCPA phase map between the SCAO PYR and IMG-PP1, where the Lyot stops and APP are located.

In Fig. 18(a), we investigate the influence of static NCPA after filtering the low spatial frequencies out of the NCPA map shown in Fig. 17. We also show the case where low spatial frequencies have not been filtered out, which would correspond to a case where the PSI correction would fail. To simulate various amounts of static NCPA, our NCPA map was rescaled to various rms wavefront errors, ranging from 100 to 200 nm (the latter being significantly beyond what we expect for METIS). This Fig. 18(a) shows that, even though they create a purely static contribution to the coronagraphic PSF, static NCPA would have a significant influence on the HCI performance

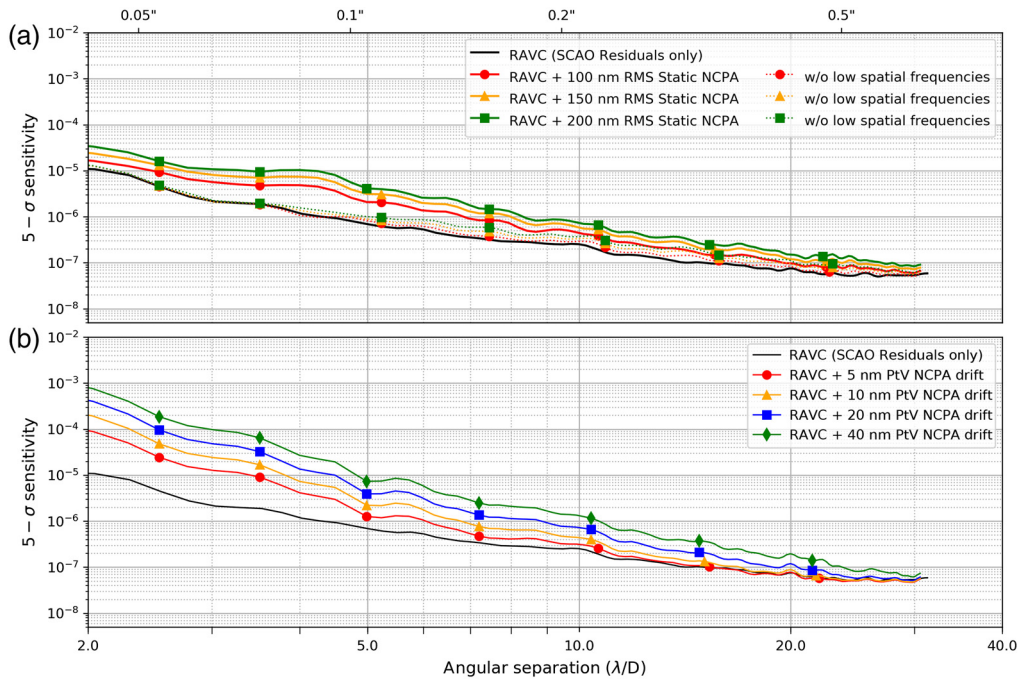


Fig. 18 ADI contrast curve for HCI performance in the presence of various levels of static NCPA on the HCI side (a) the dashed curves correspond to the same NCPA maps as the solid curves from which we have removed the low-order modes up to 10 cycles per pupil. (b) NCPA drifts with a timescale equal to the duration of the ADI sequence considering only the contribution of AO residuals speckle noise.

if they were not corrected by a common-path wavefront sensing method such as PSI. To provide a postprocessed contrast compliant with our goal, the static NCPA level would need to be smaller than 150-nm rms. This is within reach based on the METIS preliminary design, but operating with such a level of static aberrations is not desirable. Once the static low-order modes are removed by PSI, the contribution of medium- to high-order NCPA to the HCI performance becomes completely negligible at the angular separation where the contrast requirement is defined.

Dynamic (aka quasistatic) NCPA is more worrisome than static NCPA and is expected to largely affect HCI performance. While the METIS design should minimize NCPA variations due to the gravity-invariant, temperature-controlled, and mostly nonmoving noncommon path optical train (except for the chopper and the SCAO field selector/modulator), we can still expect NCPA to vary to some level during the observations, e.g., because of chromatic variations in the footprint of the beam on the optics. In Fig. 18(b), we show the simulated HCI performance in the presence of NCPA variations during the entire ADI sequence. To create these NCPA variations, we have rescaled the input NCPA map with a sinusoidally varying factor that makes the rms wavefront error in the rescaled map go from zero up to the value quoted in the legend, assuming a period of 1 h for the sinusoid (worst-case scenario). The mean PtV amplitude over time for the wavefront error measured at each individual pixel of the pupil is a factor of $\sqrt{2}$ smaller than the values quoted in the legend. We have assumed here that PSI does a good job cancelling the static, low spatial frequency part of the NCPA map, and that higher spatial frequencies have a negligible effect on the HCI performance as described above so that we can perform our variable NCPA simulations around a zero-mean NCPA map. To preserve our goal contrast requirement of 3×10^{-6} at $5\lambda/D$, dynamic NCPA should have an amplitude smaller than about 15-nm PtV for the worst-case scenario simulated here (1-h-period sinusoid). A more comprehensive simulation of NCPA variations, obtained by generating random NCPA fluctuations from a white power spectrum covering timescales of 100 s to 1 h, would significantly relax this requirement to about 20-nm rms, based on additional simulations not shown here. We also note that the overall contrast requirement (3×10^{-5} at $0.1''$) would still be met with NCPA variations significantly larger than the largest NCPA variability simulated here (40-nm PtV), which leaves some room for a possible relaxation of the NCPA stability requirement on the instrument. This could turn out to be a challenging specification to meet anyway, be it by design or by usage of PSI closed-loop control. Although it is too early to assess whether PSI on METIS will achieve this correction accuracy, preliminary simulations suggest that this is within reach, especially for the long timescale considered here.

In addition to exploring NCPA variations occurring on the timescale of the ADI sequence, we also explored more rapid NCPA fluctuations. For instance, using a sinusoidal NCPA variation with a period 10 times smaller than the duration of the ADI sequence, we improve the resulting ADI postprocessed contrast by a factor of around 2. More rapid fluctuations would have an even smaller effect. This is another demonstration that the worst-case scenario for noise source is slow variations on the timescale of the ADI sequence. This means that the control bandwidth of PSI can be relatively low, down to about 0.01 Hz if needed, which should allow a very high signal-to-noise ratio to be built up for the NCPA measurements.

6.5 Pupil Shape (Misaligned Segments)

The ELT pupil is not coronagraphy-friendly due to the presence of a big central obstruction, secondary support structures, and segments. All these features have a significant impact on HCI performance, all the more because the image of the ELT pupil is not stabilized in position and rotation at CFO-PP1, where the effects of diffraction could have been handled by advanced apodizer designs. Both APP and RAVC have been optimized to mitigate the effect of the 30% central obstruction, but the spiders still represent an important source of spurious diffracted light. This is particularly true for the VC, because the ring apodizer is located ahead of the derotator so that it cannot be optimized to handle starlight diffracted by spider arms. Conversely, the APP has been optimized to take into account the effect of the spiders, because it is positioned after the derotator in a PP where the image of the ELT pupil is stabilized. The spiders have two effects: on one hand, they lead to additional diffracted stellar light inside the Lyot stop, and on the other

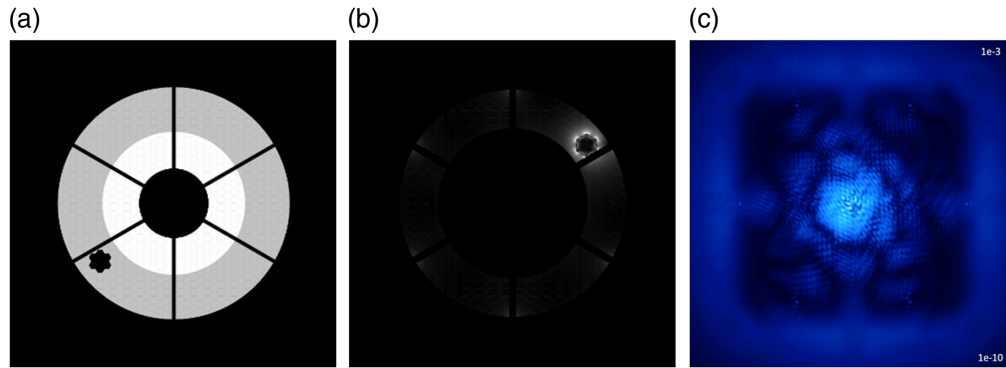


Fig. 19 (a) Input pupil with seven misaligned segments in flower configuration. (b) Corresponding distribution of light in the Lyot stop plane for the RAVC. (c) Averaged noiseless coronagraphic PSF in the presence of atmospheric turbulence.

hand, their own thermal emission increases the background emission if not properly masked by the cold Lyot stop. Due to the stabilization of the ELT pupil inside METIS, the diffracted light by the spiders can be masked at the Lyot stop/APP level with a reasonable level of oversizing: for 60-cm spiders in the ELT PP (1.6% of the input pupil diameter), the spider width in the Lyot stop/APP plane amounts to about 3% of the pupil diameter (i.e., about 110 cm in the ELT pupil), also taking into account the effects of pupil blurring and shear.

In addition to spiders, mirror segmentation can also lead to performance degradation. Segmentation itself creates spurious replicas of the ELT PSF within the coronagraphic images. These aliases, however, appear at spatial frequencies higher than the frequencies that can be corrected by the M4 deformable mirror. They will, therefore, be located in a region where HCI performance is relatively poor, and are thus not considered as a critical issue here. The real problem with mirror segmentation is that during night-time operations, several mirror segments could be misaligned due to maintenance. The average number of misaligned segments will not be larger than seven over the ELT lifetime. The most probable spatial distribution of missing segments is a flower pattern [i.e., seven neighboring segments, see Fig. 19(a)] due to a failure within a segment concentrator. Because these misaligned segments may change from night to night, it is impossible to optimize the coronagraphs to compensate for them. Assuming that the misaligned segments will be pointed toward an empty sky region, their net effect is to create a “hole” in the starlight intensity distribution in the input pupil. This hole will create additional diffracted starlight inside the Lyot stop, leading to additional stellar leakage in the coronagraphic PSF (see Fig. 19).

Figure 20 shows the contrast curves in presence of one or two sets of seven misaligned segments arranged in a flower pattern. In this configuration, the effect of misaligned segments starts

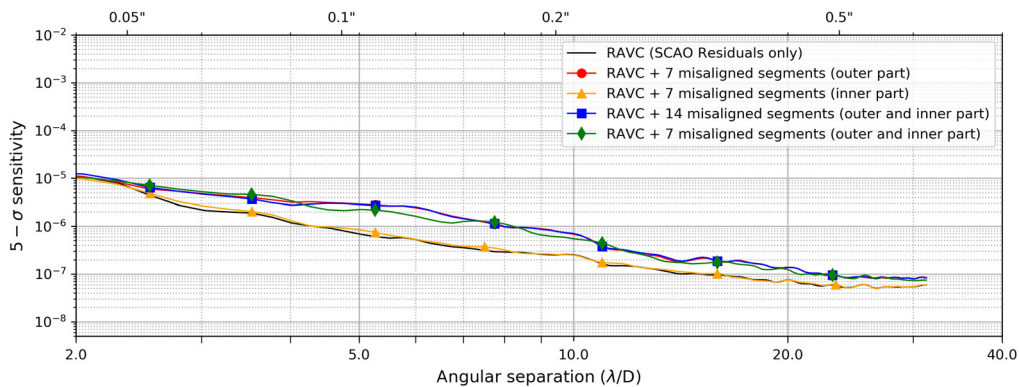


Fig. 20 ADI contrast curves for the RAVC at the *L*-band in the presence of residual atmospheric turbulence for various configurations of misaligned segments in the ELT pupil, considering only the contribution of AO residuals speckle noise.

to show at angular separations larger than about $3\lambda/D$. At $5\lambda/D$, where the contrast requirement is defined, the ADI contrast is degraded by a factor of around 5 in the case where one set of misaligned segments is located in the outer part of the pupil defined by the Lyot stop of the RAVC. While the ADI contrast performance is still compliant with our goal (3×10^{-6}), it would clearly be preferable to operate the METIS HCI modes in the absence of misaligned segments. We note that the presence of misaligned segments in the inner part of the pupil (blocked by the RAVC Lyot stop) does not significantly impact the RAVC performance.

6.6 Pupil Stability

Due to the pupil stabilization mechanism, the ELT pupil will be stabilized inside METIS with a relative accuracy of about 1%. This stabilization, however, occurs downstream of CFO-PP1, where the ring apodizer will be installed. This means that the ring apodizer will be seen as drifting on top of the image of the ELT pupil, with an expected PtV amplitude of about 2.5%. Here, we investigate the effect of this drift on the METIS HCI performance. The residual drift of the ELT pupil image downstream of the pupil stabilization mirror will be compensated by oversizing/undersizing the Lyot stop appropriately, and will, therefore, not contribute significantly to the stellar leakage level (but will reduce the overall throughput).

The effect of a misalignment of the ring apodizer with respect to the input pupil and Lyot stop is shown in Fig. 21. The additional amount of stellar leakage in the FP remains relatively low as long as the footprint of the transition between the two zones in the apodizer remains hidden behind the central obscuration of the Lyot stop, which we have purposely oversized by 3% to mitigate this effect. To simulate the effect of drifts in the ELT pupil position during an HCI observing sequence, linear motions of the ring apodizer with PtV amplitudes ranging from 1% to 5% of the M1 projected diameter have been injected in the optical propagation of the phase screens with a timescale equal to the duration of the ADI sequence. By assuming that the drift happens on the same timescale as the ADI observing sequence, we are investigating a worst-case scenario, as the variations in coronagraphic PSF shape associated to the changing intensity distribution in the input pupil will not average out sufficiently well on typically observing sequences using classical ADI processing. Under this hypothesis, a drift of 1% PtV would already be detrimental to the HCI performance at the shortest angular separations, as shown in Fig. 22. However, we can tolerate a drift of the ELT pupil of up to 3% PtV with respect to the CFO-PP1 pupil inside METIS before reaching our goal in terms of postprocessed contrast. This is compatible with the expected stability of the ELT pupil at the level of CFO-PP1. Any pupil drift occurring at higher temporal frequencies, including pupil jitter due to vibrations, have a smaller influence on the HCI performance than the low-frequency drift simulated here, because they partly average out in ADI postprocessing. The amplitude of rapid pupil jitter is expected to be smaller than 1% and will, therefore, not contribute significantly to the HCI performance budget.

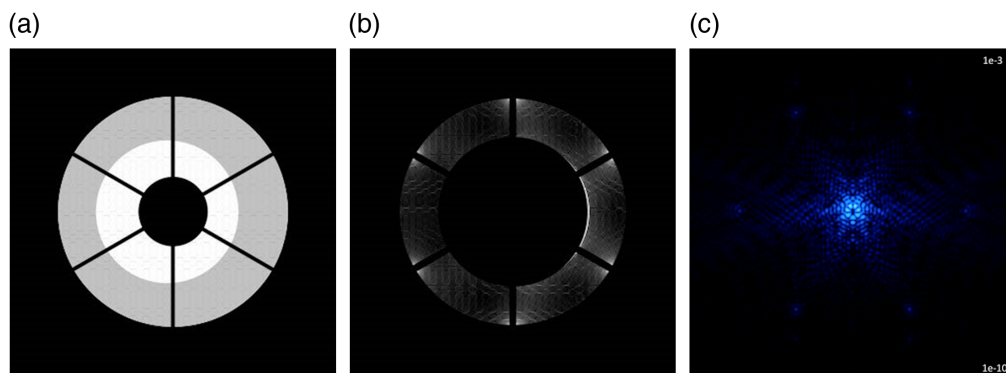


Fig. 21 Effect of the ring apodizer misalignment on the RAVC behavior. The misalignment has been exaggerated to 5% here for the sake of illustration. Illustration of (a) the apodizer, (b) intensity distribution at the Lyot plane just after the Lyot stop, and (c) associated coronagraphic PSF.

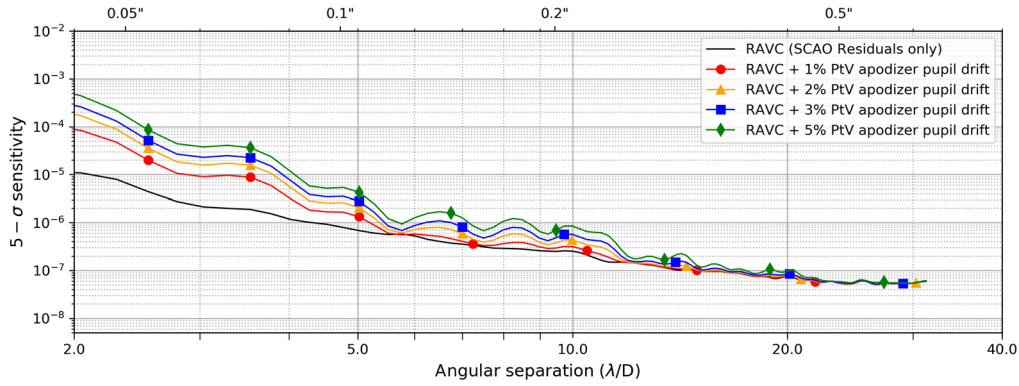


Fig. 22 ADI contrast curve for HCI performance for the RAVC at the L -band in the presence of various PtV amplitudes of low-frequency drift in the alignment of the ring apodizer compared to the ELT pupil, considering only the contribution of AO residuals speckle noise.

It must be noted that, whatever their amplitude, drifts of the ELT pupil with respect to METIS are an effect on which we have almost no control, and which we will need to live with. In case the ELT pupil drifts turn out to be larger than anticipated, we could resort to more advanced post-processing techniques to (partly) mitigate this effect. For instance, we have tested ADI post-processing algorithms based on principal component analysis,^{58,59} and checked that we could gain up to a factor of 2 in terms of postprocessed contrast at $5\lambda/D$, which would potentially allow us to work with pupil drifts larger than 5% PtV.

7 Conclusions

In this paper, we have presented a performance study of the HCI modes of the ELT/METIS instrument. Several sources of errors have been analyzed, including residual wavefront errors after SCAO control, petal piston, pointing jitter, low-frequency pointing drifts, NCPA, misaligned segments in the ELT-M1, and pupil drifts. Table 4 summarizes the results of this analysis by showing the level of each effect that would lead to an ADI contrast 10 times below the top-level contrast requirement of 3×10^{-5} at $5\lambda/D$. In the last column of the table, we give our current best estimate of the level of each of these effects in METIS based on system-level analysis simulations or on experience with similar projects on other telescopes.

Table 4 Tolerable values for noise sources in L -band RAVC observations to reach our goal contrast requirement. The last column gives our best estimate on the actual value for each effect based on experience on previous projects and/or simulations.

	Requirement	Expected
Seeing (arc sec)	0.95''	0.65''
Pointing jitter (mas rms)	2	<2
Pointing drift (mas PtV)	0.4	~0.4
Static NCPA (nm rms)	150 (total)	0 (low orders)
Quasistatic NCPA (nm PtV)	15	20
Quasistatic petal piston (nm PtV)	10	20
Misaligned segments (N/A)	7	0 or 7
Input pupil stability (PtV)	3%	2%

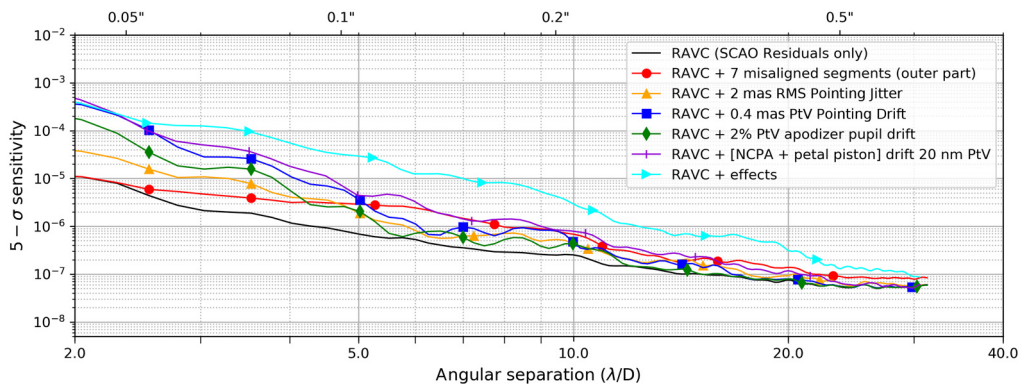


Fig. 23 ADI contrast curve for HCI performance in the presence of a combination of effects: seven missing segments, 2-mas rms pointing jitter, 0.4-mas PtV pointing drift, 2% PtV pupil drift, and a combination of NCPA and petal piston drift with 20-nm PtV amplitude, considering only the contribution of speckle noise.

To conclude this study, we give a preliminary estimation of how the contrast will be impacted by a combination of all these effects, using our best estimate for their nominal value. We assume for these simulations a pupil with seven misaligned segments under a nominal seeing of $0.65''$, with 2 mas rms pointing jitter, 0.4 mas PtV of slow pointing drifts, 20-nm PtV of slow NCPA and petal piston drifts, and 2% PtV pupil drift. Note that because the petal piston is mostly associated to low spatial frequencies, it will also be sensed by PSI, and is, therefore, counted together with the NCPA drift budget. In our simulations, we have introduced an equivalent amount of petal piston and NCPA drift (in terms of amplitude) and rescaled the resulting series of wavefront maps so that the temporal, pixel-wise PtV variation is 20 nm on average across the pupil. In this simulation, all drifts are assumed to occur at the worst possible timescale, i.e., similar to the duration of the ADI sequence. The contrast performance associated with this simulation is shown in Fig. 23. These simulations show that the combination of all these effects leads to a final postprocessed contrast that is still compatible with the top-level requirement, although the margin is small. This margin would become much more significant in the absence of misaligned segments, which is expected to be the standard case for HCI operations. We will nevertheless strive to reduce some of the main contributors to the HCI performance budget in the next phases of the METIS development (e.g., NCPA).

We remind the reader that some effects are still missing in this analysis, such as water vapor turbulence, or some SCAO-specific effects. Whether adding these additional contributions keeps the postprocessed contrast compliant with the requirement (in the absence of misaligned segments) will be explored in the future. On a more positive note, we also remind the reader that all contrast curves were computed here using a classical ADI postprocessing algorithm. During this study, we tried more advanced algorithms (such as PCA-based algorithms) on a few occasions and noted that significant gains (up to a factor two) can be obtained on the postprocessed contrast. Systematically exploring the parameters of more advanced postprocessing algorithms is, however, beyond the scope of this paper and is left for future works.

Based on the estimated performance presented in this work, we can conclude that METIS is bound to provide a dramatic increase in terms of sensitivity to faint companions in the thermal infrared range. Compared to the on-sky results obtained at the L -band with VCs installed at VLT, Keck, and LBT,⁶⁰ we estimate a gain ranging from a factor 100 to 1000 in terms of contrast at $0.1''$, which will give access to planetary-mass objects around the snow line for stars up to 30 pc. And with achievable contrasts below 10^{-6} at the N -band, METIS is bound to reach the rocky-planet regime around our nearest neighbors.³²

Acknowledgments

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Program (ERC Grant Agreement No. 337569)

and under the European Union’s Horizon 2020 research and innovation program (ERC Grant Agreement No. 819155). The research was also supported by the Wallonia-Brussels Federation (grant for Concerted Research Actions).

References

1. M. Gillon et al., “Temperate earth-sized planets transiting a nearby ultracool dwarf star,” *Nature* **533**, 221 (2016).
2. G. Anglada-Escudé et al., “A terrestrial planet candidate in a temperate orbit around Proxima Centauri,” *Nature* **536**, 437–440 (2016).
3. S. Sallum et al., “Accreting protoplanets in the LkCa 15 transition disk,” *Nature* **527**, 342–344 (2015).
4. M. Keppler et al., “Discovery of a planetary-mass companion within the gap of the transition disk around PDS 70,” *Astron. Astrophys.* **617**, A44 (2018).
5. M. Reggiani et al., “Discovery of a point-like source and a third spiral arm in the transition disk around the Herbig Ae star MWC 758,” *Astron. Astrophys.* **611**, A74 (2018).
6. T. Fusco et al., “SAXO, the SPHERE extreme AO system: on-sky final performance and future improvements,” *Proc. SPIE* **9909**, 99090U (2016).
7. O. Guyon et al., “The Subaru Coronagraphic Extreme AO (SCEXAO) system: wavefront control and detection of exoplanets with coherent light modulation in the focal plane,” *Proc. SPIE* **7736**, 773624 (2010).
8. E. Pinna et al., “SOUL: the single conjugated adaptive optics upgrade for LBT,” *Proc. SPIE* **9909**, 99093V (2016).
9. D. Mawet et al., “Annular groove phase mask coronagraph,” *Astrophys. J.* **633**, 1191–1200 (2005).
10. M. Kenworthy et al., “First on-sky high-contrast imaging with an apodizing phase plate,” *Astrophys. J.* **660**, 762–769 (2007).
11. R. Soummer, “Apodized pupil Lyot coronagraphs for arbitrary telescope apertures,” *Astrophys. J.* **618**, L161–L164 (2005).
12. N. J. Kasdin et al., “Extrasolar planet finding via optimal apodized-pupil and shaped-pupil coronagraphs,” *Astrophys. J.* **582**, 1147–1161 (2003).
13. F. Cantalloube et al., “Direct exoplanet detection and characterization using the ANDROMEDA method: performance on VLT/NaCo data,” *Astron. Astrophys.* **582**, A89 (2015).
14. L. Pueyo, “Detection and characterization of exoplanets using projections on Karhunen Loeve Eigenimages: forward modeling,” *Astrophys. J.* **824**, 117 (2016).
15. C. A. G. Gonzalez, O. Absil, and M. Van Droogenbroeck, “Supervised detection of exoplanets in high-contrast imaging sequences,” *Astron. Astrophys.* **613**, A71 (2018).
16. I. Snellen et al., “Combining high-dispersion spectroscopy with high contrast imaging: probing rocky planets around our nearest neighbors,” *Astron. Astrophys.* **576**, A59 (2015).
17. J. Wang et al., “Observing exoplanets with high dispersion coronagraphy. I. The scientific potential of current and next-generation large ground and space telescopes,” *Astron. J.* **153**, 183 (2017).
18. H. J. Hoeijmakers et al., “Medium-resolution integral-field spectroscopy for high-contrast exoplanet imaging: molecule maps of the β Pictoris system with SINFONI,” *Astron. Astrophys.* **617**, A144 (2018).
19. B. R. Brandl et al., “METIS: the mid-infrared E-ELT imager and spectrograph,” *Proc. SPIE* **7014**, 70141N (2008).
20. B. R. Brandl et al., “Status of the mid-IR ELT imager and spectrograph (METIS),” *Proc. SPIE* **10702**, 107021U (2018).
21. S. Hippler et al., “Single conjugate adaptive optics for the ELT instrument METIS,” *Exp. Astron.* **47**, 65–105 (2019).
22. A. Burrows, D. Sudarsky, and I. Hubeny, “Spectra and diagnostics for the direct detection of wide-separation extrasolar giant planets,” *Astrophys. J.* **609**, 407–416 (2004).

23. D. Mawet et al., “ L' -band AGPM vector vortex coronagraph’s first light on VLT/NACO. Discovery of a late-type companion at two beamwidths from an FOV star,” *Astron. Astrophys.* **552**, L13 (2013).
24. S. Kendrew et al., “Atmospheric refractivity effects on mid-infrared ELT adaptive optics,” *Proc. SPIE* **7015**, 70155T (2008).
25. M. Kasper et al., “NEAR: low-mass planets in α Cen with VISIR,” *Messenger* **169**, 16–20 (2017).
26. S. P. Quanz et al., “A young protoplanet candidate embedded in the circumstellar disk of HD 100546,” *Astrophys. J. Lett.* **766**, L1 (2013).
27. B. A. Biller et al., “An enigmatic point-like feature within the HD 169142 transitional disk,” *Astrophys. J. Lett.* **792**, L22 (2014).
28. M. Reggiani et al., “Discovery of a companion candidate in the HD 169142 transition disk and the possibility of multiple planet formation,” *Astrophys. J. Lett.* **792**, L23 (2014).
29. J. Rameau et al., “Confirmation of the planet around HD 95086 by direct imaging,” *Astrophys. J.* **779**, L26 (2013).
30. T. Meshkat et al., “Further evidence of the planetary nature of HD 95086 b from Gemini/NICI H-band data,” *Astrophys. J.* **775**, L40 (2013).
31. R. Galicher et al., “Near-infrared detection and characterization of the exoplanet HD 95086 b with the Gemini Planet Imager,” *Astron. Astrophys.* **565**, L4 (2014).
32. S. P. Quanz et al., “Direct detection of exoplanets in the 3 – 10 μ m range with E-ELT/METIS,” *Int. J. Astrobiol.* **14**, 279–289 (2015).
33. P.-O. Lagage et al., “Final design of VISIR: the mid-infrared imager and spectrometer for the VLT,” *Proc. SPIE* **4008**, 1120–1131 (2000).
34. P. M. Hinz et al., “Overview of LBTI: a multipurpose facility for high spatial resolution observations,” *Proc. SPIE* **9907**, 990704 (2016).
35. G. Foo, D. M. Palacios, and G. A. Swartzlander, “Optical vortex coronagraph,” *Opt. Lett.* **30**, 3308–3310 (2005).
36. D. Mawet et al., “Ring-apodized vortex coronagraphs for obscured telescopes. I. Transmissive ring apodizers,” *Astrophys. J. Suppl. Ser.* **209**, 7 (2013).
37. F. Snik et al., “The vector-APP: a broadband apodizing phase plate that yields complementary PSFs,” *Proc. SPIE* **8450**, 84500M (2012).
38. M. J. Escuti, J. Kim, and M. W. Kudenov, “Controlling light with geometric-phase holograms,” *Opt. Photonics News* **27**, 22 (2016).
39. R. Ragazzoni, “Pupil plane wavefront sensing with an oscillating prism,” *J. Mod. Opt.* **43**, 289–293 (1996).
40. D. Gratadour et al., “COMPASS: an efficient, scalable and versatile numerical platform for the development of ELT AO systems,” *Proc. SPIE* **9148**, 91486O (2014).
41. C. V erinaud, “On the nature of the measurements provided by a pyramid wave-front sensor,” *Opt. Commun.* **233**, 27–38 (2004).
42. J. E. Krist, “PROPER: an optical propagation library for IDL,” *Proc. SPIE* **6675**, 66750P (2007).
43. C. Marois et al., “Angular differential imaging: a powerful high-contrast imaging technique,” *Astrophys. J.* **641**, 556–564 (2006).
44. C. G. Gonzalez et al., “VIP: vortex image processing package for high-contrast direct imaging,” *Astron. J.* **154**(1), 7 (2017).
45. D. Mawet et al., “Fundamental limitation of high-contrast imaging at small angles set by small sample statistics,” *Astrophys. J.* **792**, 97 (2014).
46. E. E. Bloemhof et al., “Behavior of remnant speckles in an adaptively corrected imaging system,” *Astrophys. J. Lett.* **558**, L71–L74 (2001).
47. O. Guyon, “Limits of adaptive optics for high-contrast imaging,” *Astrophys. J.* **629**, 592–614 (2005).
48. A. Turchi et al., “Forecasting water vapour above the sites of ESO’s Very Large Telescope (VLT) and the Large Binocular Telescope (LBT),” *Mon. Not. R. Astron. Soc.* **482**, 206–218 (2019).
49. A. J. Skemer et al., “A direct measurement of atmospheric dispersion in N-band spectra: implications for mid-IR systems on ELTs,” *Publ. Astron. Soc. Pac.* **121**, 897 (2009).

50. J. Milli et al., “Low wind effect on VLT/SPHERE: impact, mitigation strategy, and results,” *Proc. SPIE* **10703**, 107032A (2018).
51. V. Hutterer et al., “Wavefront reconstruction for ELT-sized telescopes with pyramid wavefront sensors,” *Proc. SPIE* **10703**, 1070344 (2018).
52. V. Hutterer et al., “Advanced wavefront reconstruction methods for segmented extremely large telescope pupils using pyramid sensors,” *J. Astron. Telesc. Instrum. Syst.* **4**(4), 049005 (2018).
53. D. Mawet et al., “The vector vortex coronagraph: sensitivity to central obscuration, low-order aberrations, chromaticism, and polarization,” *Proc. SPIE* **7739**, 773914 (2010).
54. E. Huby et al., “Post-coronagraphic tip-tilt sensing for vortex phase masks: the QACITS technique,” *Astron. Astrophys.* **584**, A74 (2015).
55. E. Huby et al., “On-sky performance of the QACITS pointing control technique with the Keck/NIRC2 vortex coronagraph,” *Astron. Astrophys.* **600**, A46 (2017).
56. A. L. Maire et al., “Design, pointing control, and on-sky performance of the mid-infrared vortex coronagraph for the VLT/NEAR experiment,” *J. Astron. Telesc. Instrum. Syst.* **6**(3), 035003 (2020).
57. J. L. Codona and M. Kenworthy, “Focal plane wavefront sensing using residual adaptive optics speckles,” *Astrophys. J.* **767**, 100 (2013).
58. R. Soummer, L. Pueyo, and J. Larkin, “Detection and characterization of exoplanets and disks using projections on Karhunen–Loève Eigenimages,” *Astrophys. J.* **755**, L28 (2012).
59. A. Amara and S. P. Quanz, “PYNPOINT: an image processing package for finding exoplanets,” *Mon. Not. R. Astron. Soc.* **427**, 948–955 (2012).
60. O. Absil et al., “Three years of harvest with the vector vortex coronagraph in the thermal infrared,” *Proc. SPIE* **9908**, 99080Q (2016).

Biographies of the authors are not available.