Halftoning for High-contrast Imaging: Developments for the SPHERE and EPICS Instruments

Patrice Martinez[1]
Christophe Dorrer[2]
Emmanuel Aller-Carpentier[1]
Markus Kasper[1]
Anthony Boccaletti[3]
Kjetil Dohlen[4]

[1] ESO
[2] Aktiwave, Rochester, New York, USA
[3] LESIA, Observatoire de Paris, France
[4] Laboratoire d'Astrophysique de Marseille, France

Controlling the amplitude of light is crucial for many scientific applications, such as in imaging systems, astronomical instruments or laser physics. We provide a brief overview of recent R&D activities at ESO using halftoning, the process of presenting a continuous image through use of dots. Customised filters with spatially varying transmission are produced using a binary array of metal pixels that offers excellent control of the local transmission. Applications to the production of an apodiser for the VLT SPHERE instrument and a mask for E-ELT EPICS are presented.

High-contrast instruments, such as SPHERE, the Spectro-Polarimetric High-contrast Exoplanet Research instrument (the planet-finder instrument under construction for the ESO Very Large Telescope or VLT), or EPICS (the planet-hunter project for the future European Extremely Large Telescope or E-ELT), will require customised components with spatially varying transmission. A typical example is a stellar coronagraph, which is an optical element that is capable of reducing the contrast between a companion and its parent star. The goal of these subsystems is to control the spatial transmission, either in a pupil plane (pupil apodisation), or in a focal plane of the instrument (occulting mask, i.e. low frequency filter). Reliably producing these components with spatially varying transmission is not trivial, and different techniques have already been investigated for application to astronomy (e.g., metal deposition with spatially varying thickness, or high energy beam-sensitive glass treated by electron beam lithography). We present some results related to the recent development of components with spatially varying transmission using a relatively simple technique analogous to the digital halftoning process used for printing applications.

Halftone dot process principle and applications

Halftoning has been used for hundreds of years in the printing industry as a way of generating continuous tone images with only black or white dots. The so-called halftone image, seen from a distance, should resemble the original continuous tone image as much as possible, based on human visual perception. Following this idea, the continuous transmission of an optical filter can be generated with a specific implementation of opaque and transparent pixels as presented in Figure 1.

The microdot filter is an array of dots (i.e. pixels) that are either opaque (0 % transmission) or transparent (100 % transmission). It is fabricated by lithography of a light-blocking metal layer deposited on a transparent glass substrate. The technique has several advantages: relative ease of manufacture, achromaticity, reproducibility and the ability to generate continuous and rapidly varying transmission functions, without introducing wavefront errors. The technique can be applied in many areas, such as laser beam shaping (Dorrer et al., 2007), intensity modulation in projection and illumination systems, coronagraphy, apodisation (soft-edged apertures) and optical testing. It allows customised filters to be generated with specific and rapidly varying transmission shape, e.g., Gaussian, prolate, sinusoidal-like functions and test zones with calibrated transmissions.

Key parameters

In order to specify a microdot filter, several application-dependent parameters must be carefully defined. Microdot filters are arrays of pixels with binary transmission, and thus light propagating through them has a continuously varying intensity after free space propagation, or after filtering in the far field. Therefore, these filters have a power spectrum different from the power spectrum of an ideal continuous filter. The way the dots are distributed, the size of the dots, and the density of the dots in the pattern directly impact the power spectrum of the filter, and these parameters must be carefully tuned according to the application.

Distribution of the dots

The error diffusion algorithm provides excellent continuous tone rendering and only produces noise at high frequencies in the power spectrum. Unlike many other halftoning algorithms, the error diffusion algorithm has the advantage of only introducing high frequency noise, i.e. the low and middle range of frequencies in the generated transmission have little or no noise due to pixellation and the binary nature of the dot transmission. The error diffusion technique is very simple to describe. The algorithm sets the transmission of a given pixel (either 0 % or 100 %) by comparing the transmission required at this location to a 50 % threshold. The transmission is set to zero if the required transmission is less than 50 % and to one otherwise. The induced transmission error is diffused to adjacent pixels that have not yet been processed by biasing the transmission required at the corresponding locations. This locally cancels the error of the binary optics introduced by the process of translating the required
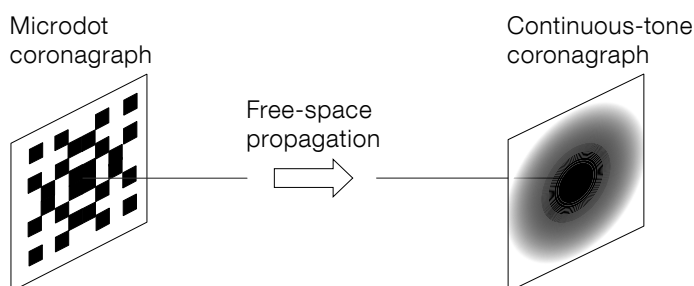
Microdot coronagraph

Continuous-tone coronagraph

Free-space propagation

Figure 1. Principle of the generation of a continuous modified intensity transmission using a microdot filter after Fourier filtering.

continuous tone transmission into binary values. This procedure is used for gray level reproduction with black and white printing techniques, and has been used for laser beam shaping (Dorrer et al., 2007).

Characteristics of the microdot filters

On account of the pixellation and binary transmission of the pixels, high frequency components are generated in the power spectrum of a microdot filter. The smaller the dots are, the higher the spatial frequencies at which the light is scattered. With smaller pixels, the agreement between the required transmission and the obtained transmission improves (Martinez et al., 2009a; 2009b). This issue is formally equivalent to a sampling problem. Better results are obtained when increasing the sampling rate (i.e. more binary dots are used to reproduce the continuously varying transmission). In the specific case of coronagraphs, theoretical derivations of the impact of the dot size in the science image for pupil apodiser-type components have been developed (Martinez et al., 2009a), and verified experimentally (Martinez et al., 2009b), as well as for focal plane masks (Martinez et al., 2009c).

One must generally avoid dots with sizes comparable to the wavelength of the light with which they are designed to operate. In such a regime, it is difficult to predict how the field will react to such filters, and the transmission might be strongly dependent on the wavelength. Typical pixel sizes that are useful for astronomical applications range from a few tens of microns to a few microns.

Coronagraphs fabricated for pupil plane apodisation do not require a high opacity. However, the opacity of the metal layer (the optical density [OD], which is wavelength dependent) is an important parameter for focal plane masks, whether they are fabricated with microdot filters or other technologies. High OD specification is required for focal plane mask coronagraphs, at least OD > 6 (transmission $< 10^{-6}$) in order to avoid starlight leaking on to the detector.

Application of halftoning to SPHERE

The Apodised Pupil Lyot Coronagraph (APLC) is one of the three coronagraphic concepts for the SPHERE IRDIS, (Infra-Red Dual beam Image and Spectrograph) instrument, which will enable direct images, and spectra to be obtained for warm, self-luminous Jovian planets. A recent study also points out its potential application to E-ELT instruments (Martinez et al., 2008), such as EPICS, the extreme adaptive optics planet imager. The APLC combines an apodiser in the entrance aperture with a small opaque mask in the focal plane. Manufacturing an accurate apodiser is critical for such a coronagraph, as for any other concept using pupil apodisation (e.g., a dual zone coronagraph, or for conventional pupil apodisation concepts).

The definition of the prototype profile has been optimised (Martinez et al., 2007) for its use on HOT, the High Order Test bench at ESO (Aller Carpentier et al., 2008), and is roughly similar to the one defined for SPHERE. The main characteristics of the microdot version of the APLC apodiser are summarised in Table 1. The main difference with SPHERE is the size of the pupil to apodise (18-mm diameter for SPHERE, while in HOT it is only 3 mm, increasing the difficulty of producing the apodiser because smaller pixels are required). The size of the square chrome dots of our prototype is 4.5 µm; see Figure 2, where metrology inspection of the microdot apodiser is illustrated. The dot spatial distribution across the pupil was analysed using a shadowgraph (contour profile projector for quality control) and compared with a simulation map, demonstrating extremely good agreement. The principle of the shadowgraph inspection consists in placing the specimen on a glass stage that is illuminated from below, and the resulting image is picked up by a microscope objective and

projected onto a large projection screen, where only the details of the contour and profile are seen. The shadowgraph is a powerful tool for controlling dot size and spatial distribution, which can be easily compared to specification maps.

Figure 3 shows the good agreement between the measured and required profiles. The transmission error is approximately 3 %. The achromaticity of the transmission profile is also demonstrated: the profile error increases by only about 2 % between the narrow H-filter and the broadband J-filter. Figure 4 shows a laboratory experiment where we have analysed the effect of the dot size in the coronagraphic image. Different microdot filters were fabricated to approximate the same coronagraph with dot size specified as 15, 30, 60, 120 and 240 µm, corresponding to masks 2 to 6 respectively.

Coronagraphic images exhibit speckles with intensity and position in the field that were confirmed to be a function of the dot size, and the agreement with the theoretical derivation is very good (Martinez et al., 2009b). When the dots are sufficiently small, a usable field of view free of speckles appears and reveals the residual diffraction from the VLT-like pupil (spider vane diffraction spikes). Overall the prototype has excellent quality and meets the specifications.

Application of halftoning to EPICS

Even though the APLC is considered as well, for EPICS (an instrument that is planned to enable direct images and spectra for both young and old Jupiter-mass planets in the infrared), several

Table 1. Summary of the main characteristics of the prototype halftone components for SPHERE and EPICS.

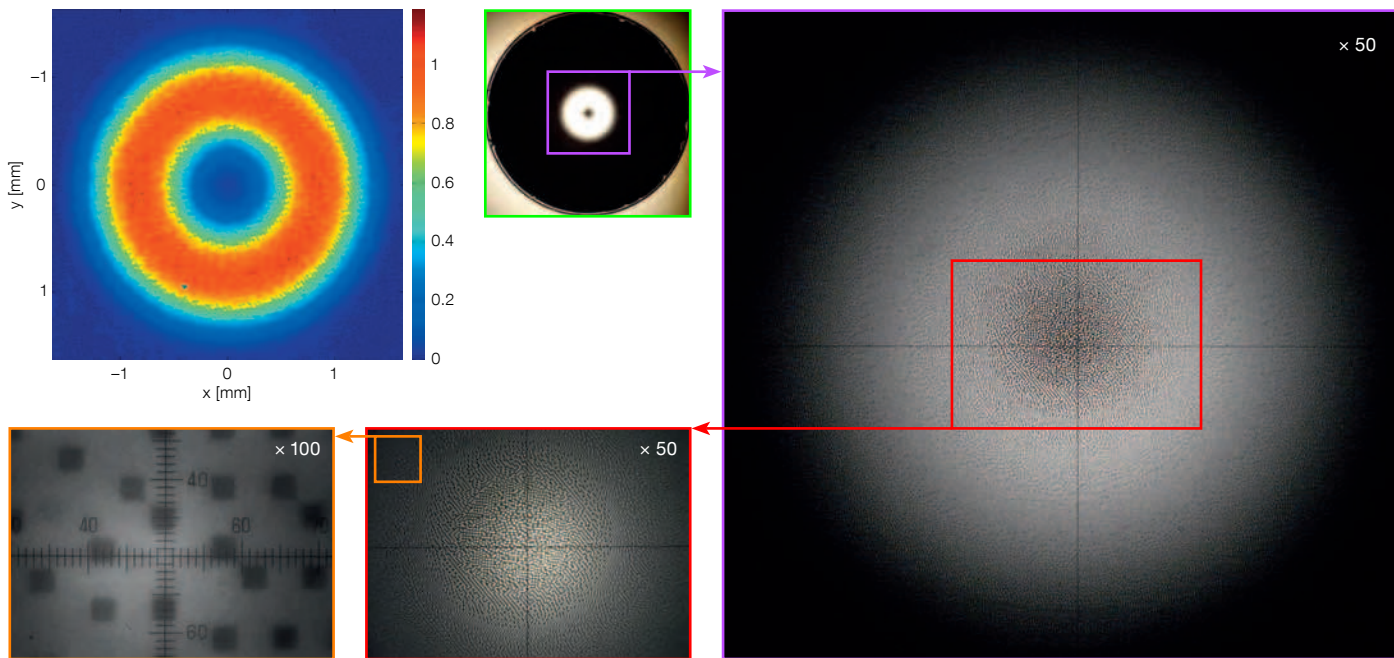| | Prototype | |
	Pupil apodiser (APLC)	Focal plane mask (BLC)
Function	Prolate-like function	(1-sinc)
Metal layer	Chrome	Aluminium
OD (dot opacity)	4+ (Visible)	8+ (near-infrared)
Dot size	4.5 µm	5/10 µm
Diameter of the pattern	3 mm	12.0 mm
Diameter of the substrate (BK7)	12.7 mm	12.7 mm
Accuracy of the profile	3 %	5 %

Figure 2 (above). Metrology inspection of the final prototype using a camera (green box), a shadow-graph (purple and red boxes), a microscope (orange box), and the corresponding spatially resolved transmission (top-left).
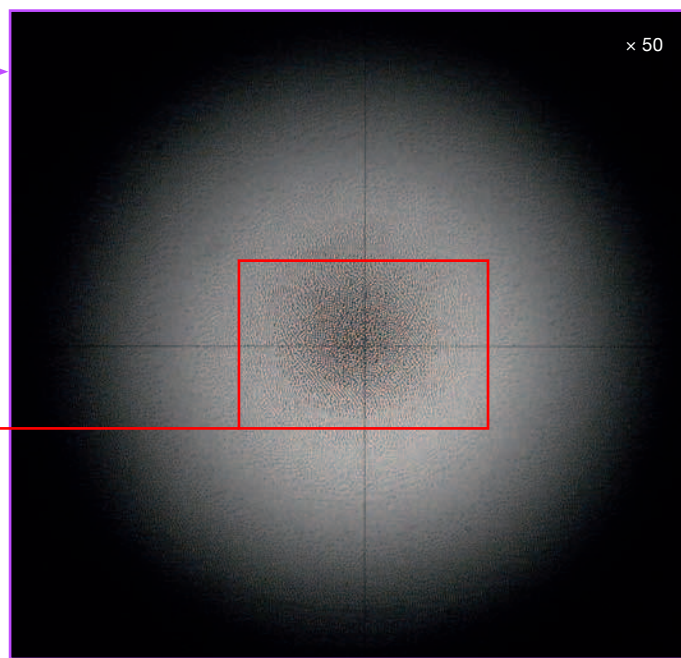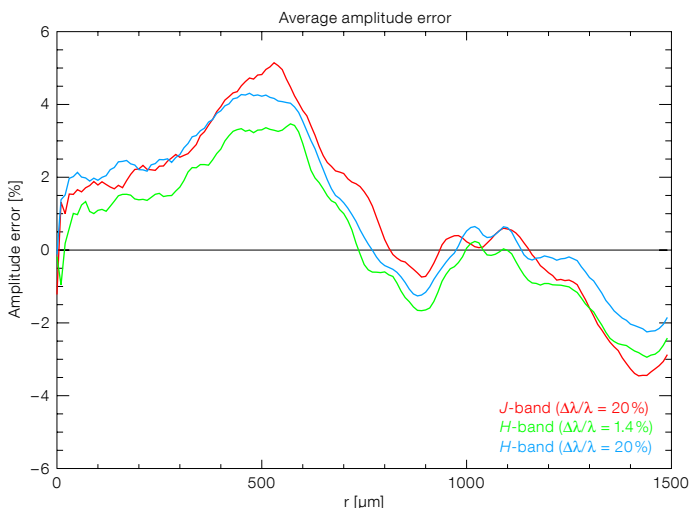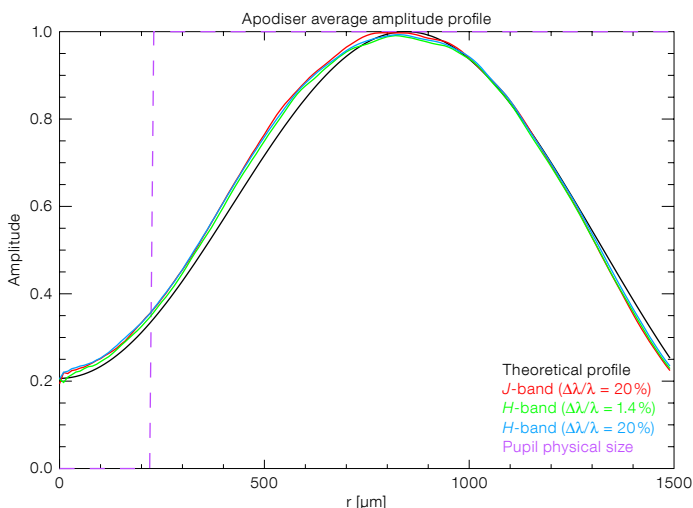
Figure 3 (below). Left: Apodiser azimuthally averaged from centre to edge using different filters (*J*- and *H*-band) with width indicated compared to the specification (black curve). Right: Corresponding average amplitude error as a function of radius.

other coronagraphic concepts are under investigation. Some of them require the microdot technique for pupil apodisation (conventional pupil apodisation corona-graph or dual zone coronagraph), while the band-limited coronagraph (BLC) requires an accurate filter with a spatially varying transmission (e.g., a sinusoidal function such as 1-sinc) in the focal plane of the instrument. A theoretical analysis

of the design issues (different from the APLC) is detailed in Martinez et al. (2009c).

We produced several prototypes (Figure 5) based on the same function, but with different dot sizes (indicated as a and b in Figure 5, where a corresponds to a 5-μm dot size and b to a 10-μm dot size) and function bandwidth (BL5

and BL10). The main characteristics of the prototypes are detailed in Table 1. Qualitative and quantitative descriptions of the resulting prototypes are shown in Figures 5 and 6. The accuracy is about 5% in the high transmission part (recently improved to an accuracy of 2% with new prototypes). Accuracy in the low transmission part (centre of the mask) is extremely good.
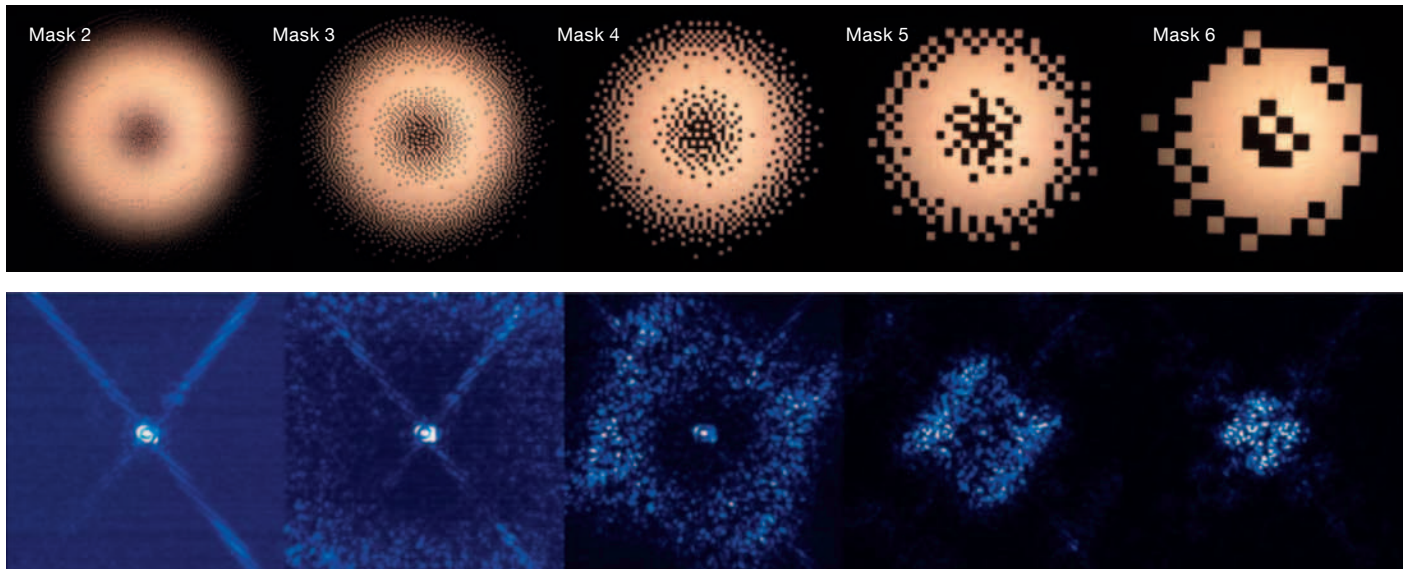
Figure 4 (above). Top row: Shadowgraph inspection (× 50) of different masks based on the same profile, but differing dot size. Bottom row: The coronagraphic images (with narrowband illumination, $\Delta\lambda/\lambda = 1.4\%$) corresponding to the different masks.
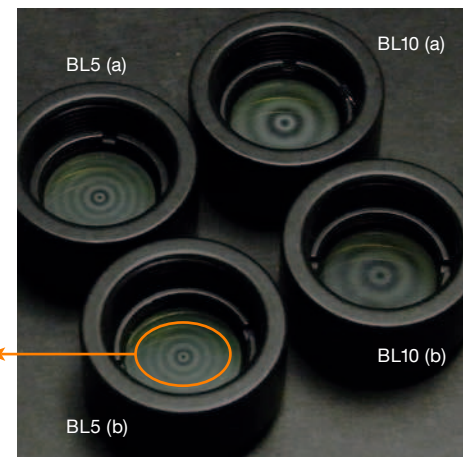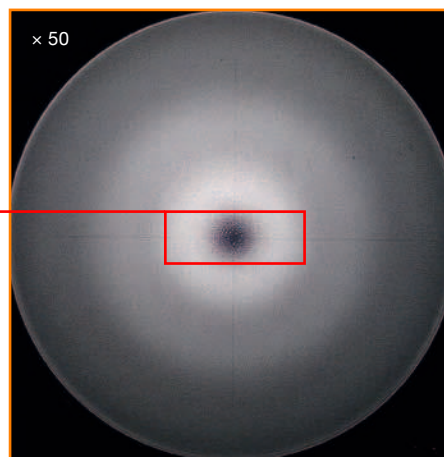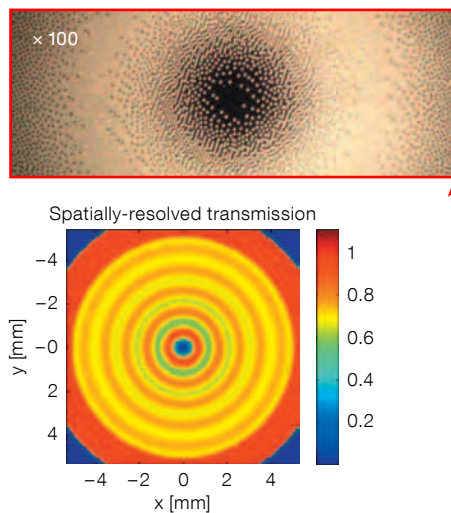
Figure 5 (below). Metrology inspection of the band-limited coronograph prototypes. From right to left: The four prototypes in their integration mounts, shadowgraph image (× 50), shadowgraph image (× 100) and spatially resolved transmission of BL5 (b), with 5 μm dot size.

Results obtained with APLC and BLC

We tested our prototypes first in a near-infrared coronagraphic transmission bench developed at ESO (turbulence-free and AO-free, operating in *H*-band). APLC demonstrates the ability to reduce the on-axis starlight by a factor of 700 on the peak, and to reach a contrast of 5×10^{-5} and 10^{-6} at 3 and 20 λ/D offset (0.12 and 0.8 arcseconds for an 8-metre telescope) respectively (Figure 7, and Martinez et al. [2009a] for more details). BL5 demon-strates a peak reduction of 2500, and a contrast of 3×10^{-5} and 2.7×10^{-8} at 5 (inner working angle), i.e. offset angle below which a companion cannot be resolved and 20 λ/D respectively (i.e. 0.2 and 0.8 arcseconds), while BL10 yields a peak reduction of 83 000 and a contrast of 1.3×10^{-7} and 1.3×10^{-8} at 10 (inner-working angle) and 20 λ/D respectively (Martinez et al., 2009c). The theoretical limits of the coronagraph have not been reached. Limitations originate from align-ment issues, and from the static aberra-tions of the optical bench. Contrast levels reached are similar (APLC) or better (BLC) than the quasi-static speckle halo level expected after adaptive optics (AO) cor-rection in a real instrument (10^{-5}–10^{-6}).

We further investigated the APLC per-formance on HOT, which is a high-con-trast imaging eXtreme Adaptive Optics (XAO) bench developed at ESO. Its objective is to test and optimise differ-ent techniques and technologies (e.g., wavefront sensors, coronagraphs,
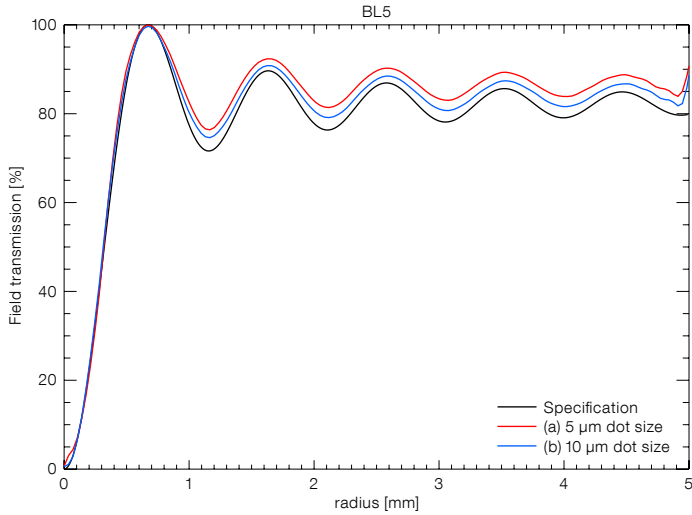
Figure 6 (above). Azimuthally averaged profiles (from centre to edge) of the band-limited coronagraphs BL5 (left) and BL10 (right) with 5 and 10 µm dot sizes.
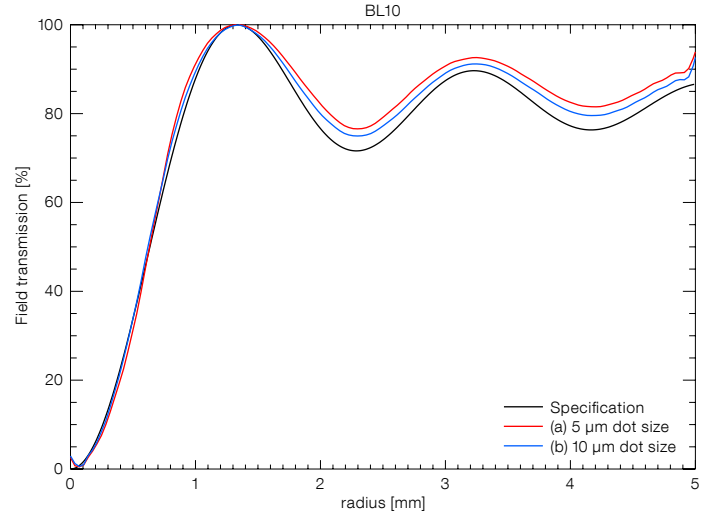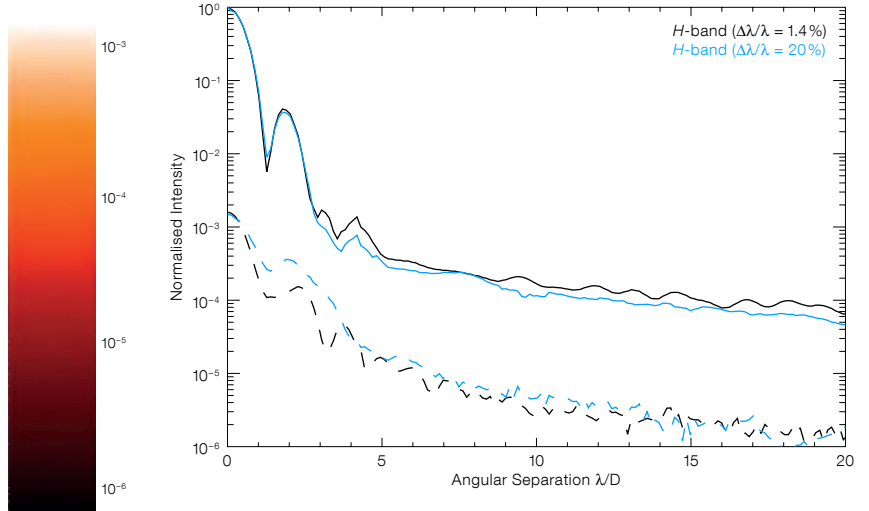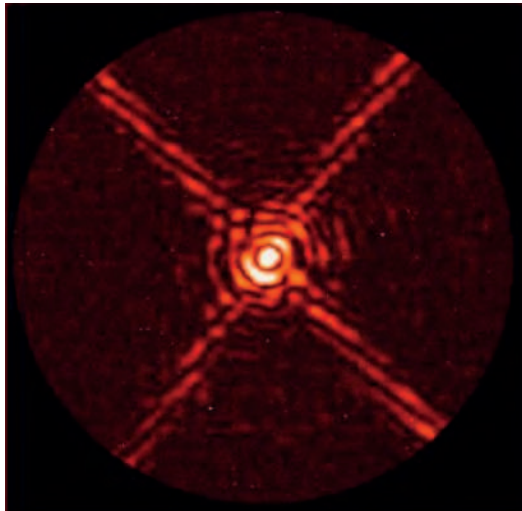
Figure 7 (below). Apodised pupil Lyot coronagraph (APLC) results obtained in *H*-band (94 % Strehl optical bench assuming the VLT-pupil without turbu-

lence or AO correction). Left: Coronagraphic image. Right: APLC contrast profiles (dashed lines) compared to the PSF contrasts, i.e. without the coronagraph (full lines), for two bands.

speckle calibration methods, image post-processing). HOT mimics the conditions occurring at a real telescope (e.g., the VLT), including turbulence generation, XAO, and various near-infrared (near-IR) coronagraph concepts. The experiment was performed under 0.5-arcsecond seeing, with real-time AO-correction using a 31 × 31 DM (15 λ/D AO cut-off frequency, i.e. 0.6 arcseconds in *H*-band) and a Shack–Hartman wavefront sensor. Recently, we have also experimented with differential imaging (DI) tests, using sequential coronagraphic image subtraction in between closely spaced narrowband filters in *H*-band (1.56 µm and 1.60 µm, *H1* and *H2* respectively).

Figure 8 qualitatively describes the different steps of the experiment.
1. The AO-corrected (apodised) point spread function (PSF) reveals the diffraction pattern of the VLT pupil by achieving a high (90 %) Strehl ratio. The apodisation of the PSF helps in reducing the intensity level of the PSF wings.
2. The coronagraphic images recorded in *H1* and *H2* both confirm the effect of the APLC, where images exhibit atmospheric speckles with lower intensity in the AO correction domain (from the second or third wing of the PSF to the rise of the AO cut-off frequency clearly seen in the images), while the central core of the coronagraphic PSF being

brighter, appears dominated by diffraction residuals and pinned-speckles.
3. The differential image, where the speckle contrast has been improved by removing the atmospheric speckle halo and a part of the instrumental speckle contribution. The image has been high-pass filtered to remove smooth structures, leaving the small-scale higher frequency objects unaffected (e.g., faint companion).

The raw contrasts achieved after the APLC (Figure 9, red and green curves) are in good agreement with the SPHERE simulations. When removing the atmospheric speckle halo with a high-pass filter
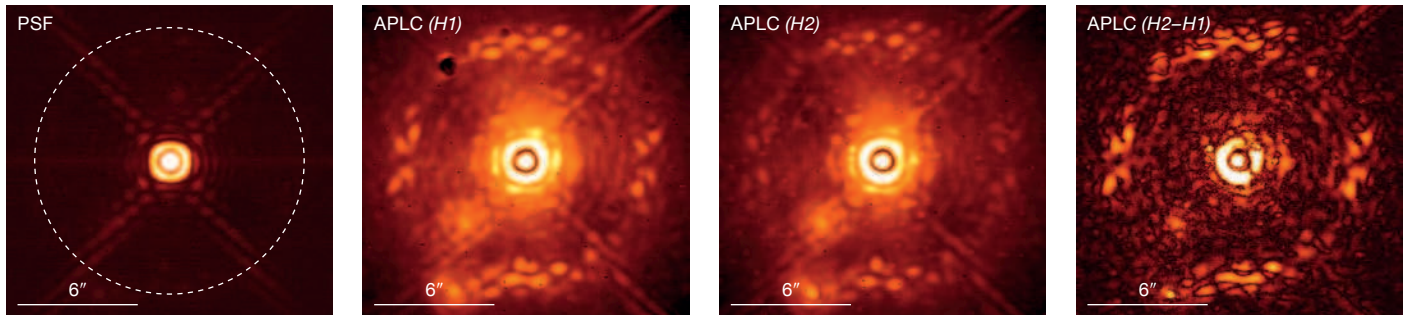
Figure 8. From left to right: AO-corrected and apodised PSF is shown (white-dashed circle indentifies the AO cut-off frequency in the field); APLC coronagraphic images recorded with *H1* and *H2* near-IR filters (middle images); and (right) subtraction of the two coronagraphic images, after high-pass filtering.
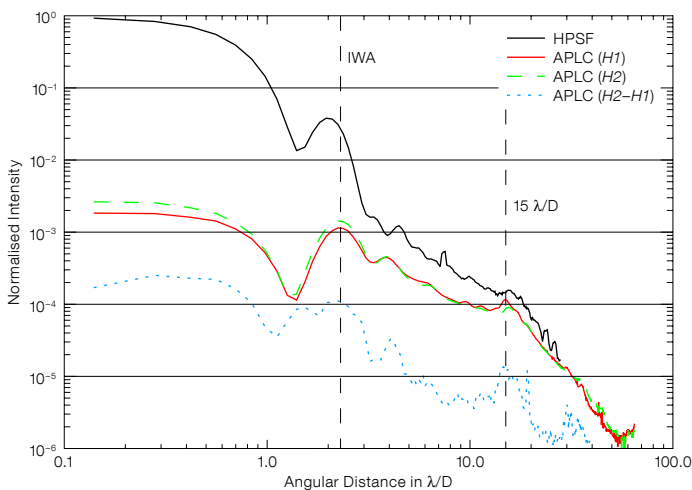

Figure 9. Azimuthally averaged profiles of the AO-corrected PSF (black curve), APLC raw coronagraphic images in filters *H1* and *H2* (red and green curves), and 1σ contrast after differential imaging and low frequency content suppression (blue curve). IWA is the inner working angle.

Table 2. Some of the next instruments that can benefit from the use of the halftoning process for producing (at least) their coronagraphs (in blue). FQPM and PIAA refer to Four Quadrant Phase Mask and Phase Induced Amplitude Apodisation respectively.

Instrument	In service by	Coronagraph type
SPHERE/VLT	2011	FQPM/APLC/Lyot
GPI/Gemini	2011	APLC
HCIAO +	2010	Lyot/APLC/PIAA
JWST	2013	Band-limited/FQPM
EPICS/E-ELT	> 2018	CPA (baseline)
PFI/TMT	> 2020	APLC

applied either on *H1* or *H2* images (not plotted in Figure 8), the contrast curve nicely fulfilled the SPHERE expectation as well. High-pass filtered APLC images demonstrate a 5σ detectability of 2.5×10^{-4} and 2.2×10^{-5} at 0.1- and 0.5-arcsecond offsets respectively, without Spectral Differential Imaging (SDI) or Angular Differential Imaging (ADI) techniques, i.e. a factor of 25 and 45 at 0.1 and 0.5 arcseconds respectively from the goal contrast of SPHERE (after SDI and/or ADI). The 1σ contrast provided after differential imaging (blue dashed curve) demonstrates atmospheric speckle suppression (1 order of magnitude gain), while slightly improving the speckle noise. Improvement of these results is foreseen with a simultaneous implementation of the DI tests (implemented in a sequential way so far). These results provide confidence in the expected performance of SPHERE.

The halftoning process presented has been validated for pupil apodisers (APLC for SPHERE and for the Gemini Planet Imager, GPI), and is defined as the baseline technique for EPICS (for producing a conventional pupil apodisation coronagraph [CPAC], the concept baseline). Space-based telescopes can also benefit from the halftoning process, such as the JWST NIRCam coronagraphic masks for which microdot band-limited coronagraphs are being manufactured. Table 2 summarises some of the future instruments that can take advantage of the halftoning technique at least for coronagraphs. Its interest has been recently extended in the context of the manufacture of the apodisation mask for the BIGRE Integral Field Spectrograph for SPHERE (Antichi et al., 2009).

References

Aller Carpentier, E. et al. 2008, Proc. SPIE, 7015, 108
Antichi, J. et al. 2009, ApJ, 695, 1042
Dorrer, C. & Zeugel, J. D. 2007, J. Opt. Soc. Am. B, 24, 1268
Martinez, P. et al. 2007, A&A, 474, 671
Martinez, P. et al. 2008, A&A, 492, 289
Martinez, P. et al. 2009a, A&A, 495, 363
Martinez, P. et al. 2009b, A&A, 500, 1281
Martinez, P. et al. 2009c, ApJ, accepted