Comparison of wavefront control algorithms and first results on the high-contrast imager for complex aperture telescopes (hicat) testbed

L. Leboulleux
M. N'Diaye
J. Mazoyer
L. Pueyo
et al.
COMPARISON OF WAVEFRONT CONTROL ALGORITHMS AND FIRST RESULTS ON THE HIGH-CONTRAST IMAGER FOR COMPLEX ApERTURE TELESCOPES (HICAT) TESTBED

L. Lebouleux¹,²,³, M. N’Diaye¹, J. Mazoyer¹, L. Pueyo¹, M. Perrin¹, S. Egron¹,²,³, E. Choquet¹,², J.-F. Sauvage¹,², T. Fusco¹,², R. Soummer³. Contact: lucie.lebouleux@lam.fr.

¹ ONERA, 29 Avenue de la Division Leclerc, 92320 Châtillon, France. ² Aix Marseille Université, CNRS, LAM (Laboratoire d’Astrophysique de Marseille) UMR 7326, 13388, Marseille, France. ³ Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA. ⁴ Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, MS 169-506, Pasadena, CA 91109, USA. ⁵ Hubble Fellow.

I. INTRODUCTION:

The next generation of space telescopes for direct imaging and spectroscopy of exoplanets includes telescopes with a monolithic mirror, such as the Wide Field Infrared Survey Telescope (WFIRST) [1] and Large Ultra-Violet Optical Infrared (LUVIOIR) telescopes with segmented primary mirror, like ATLAST [2, 3] or HDST [4]. Because of the complexity of their pupils, high-contrast imaging becomes more challenging. Furthermore, space telescopes have huge requirements in terms of contrast stability in the presence of vibrations.

The High-contrast imager for Complex Aperture Telescopes (HiCAT) testbed has been developed to enable studies on different components of high-contrast imaging, meaning starlight suppression, wavefront sensing (WFS), and wavefront control (WFC) for such unfriendly pupils. New coronagraph designs are currently developed in simulation [5, 6, 7] for a next implementation on the testbed. The wavefront control of HiCAT will also consist in two deformable mirrors (DM) pupil-remapping techniques (e.g. Active Control of Aperture Discontinuities [ACAD] [8, 9, 10]), that convert complex pupils into friendly apertures for coronagraphy.

In this communication, in section II we introduce the HiCAT testbed, focusing on its objectives and in particular the studies it will enable and the requirements that were deduced from these goals. These requirements led to a final design and environment, that we also present here, before describing its current status.

In section III, we present different well-known wavefront control methods, in particular their prerequisites.

II. HICAT TESTBED:

The HiCAT testbed is currently being developed at the Space Telescope Science Institute (STScI), more precisely at the Russell B. Makidon Laboratory. This facility is dedicated to the development of technologies for future space missions. In particular, HiCAT is designed to provide an integrated solution for high-contrast imaging for unfriendly aperture geometries in space, such as HDST or ATLAST-like pupils. In the section, we will describe the objectives of this optical bench, its final optical and opto-mechanical design that was deduced from these requirements, the environment constraints, and finally the current status of the project.

A. Goals of the testbed

The HiCAT testbed was designed to develop methods for high-contrast imaging, including a starlight and diffraction suppression system and wavefront sensing and control tools. These techniques have to be applied in complex-aperture case telescopes, which includes segment gaps, spiders and central obstruction.

Its initial contrast goal in air is $10^{-7}$ in a dark hole limited by $3\lambda/D$ and $10\lambda/D$ (where $\lambda$ is the wavelength and $D$ is the aperture diameter) in a 2% bandpass, in the visible, assuming a single Boston Micromachines-deformable mirror (DM), which should be improved to higher contrast after implementation of wavefront control methods.

To reach this contrast ratio, the testbed is designed to minimize the impact of its optical components on its final contrast, with focus on the sources of amplitude-induced errors from the propagation of out-of-pupil surfaces. To limit that effect, known as the Talbot effect, we place a requirement on the contrast contribution of amplitude errors to be one order of magnitude fainter than the total contrast, i.e. $10^{-8}$. The goal is that, by minimizing the amplitude-induced errors due to the Talbot effect, the majority of the amplitude errors comes from the discontinuities in the pupil, such as the segment gaps, the spiders or the central obstruction and will be corrected using wavefront control and wavefront shaping.

Since HiCAT was designed to compensate for both amplitude errors due to its complex entrance pupil and phase errors due to surface errors and non-homogeneous reflectivities of the optical components, two Boston-DM are planned to be used. This is why this value of contrast should be then really improved after setting up the second DM in the optical path.
Such a theoretical limit for the contrast is significantly better than the requirement. We are therefore quite confident than the ultimate performance on HiCAT will satisfy the requirement, even with complex apertures (central obstruction, spiders and segments) and in large spectral broadband operations.

But HiCAT also includes a coronagraph for starlight and diffraction-effect suppression, which is designed considering the contrast as a metric to optimize.

B. Optical and opto-mechanical design

The HiCAT testbed is designed to achieve these goals, performing high-contrast imaging in the case of unfriendly apertures. Therefore, it combines studies in coronagraphy, wavefront sensing and wavefront control, plus a simulated telescope with a complex pupil. The final layout is presented in Fig.1 and is explained in details in [7, 11, 12]. It is a purely reflective testbed, except for the last imaging lenses.

![Fig. 1. Optical and mechanical design of the HiCAT testbed, realized with the software Solidworks, the beam is exported from Zemax. The telescope is simulated by a pupil mask, the segmented mirror and off-axis parabolas.](image)

The telescope is simulated using:
- A non-circular entrance pupil mask with central obstruction and spiders to define an aperture shape. Its size is set the 20 mm to enable small details to be represented with good precision (such as the spider) and the use of 1inch optics.
- A 37-segment Iris-AO MEMs deformable mirror with hexagonal segments that can be controlled in tip, tilt, and piston. The gaps between segments are between 10 and 12 µm and the full segmented mirror has an inscribed circle size of 7 mm. This component is conjugated to the entrance pupil mask.

Together, these two components provide a segmented pupil similar to ATLAST. The segmented mirror can also be replaced with a high-quality flat mirror to give the possibility of studying AFTA-like pupils.

The chosen coronagraph is a Apodized Pupil Lyot Coronagraph (APLC), which combines a classical Lyot coronagraph with entrance pupil apodization [8, 9, 10]. This type of coronagraph is currently implemented in the exoplanet direct imagers P1640, GPI, and SPHERE. It is then composed of:
- An apodizer, located in a pupil plane, so conjugated with the two optical components previously presented.
- A reflective focal plane mask (FPM), with a 334 µm diameter central hole. The beam focal ratio at its location is set at F/80.
- A Lyot Stop, with a diameter equal to the entrance pupil. We also have another possible Lyot Stop of 10 mm.

Both the FPM and the Lyot Stop are motorized and can be controlled from a computer.

The final design of the apodizer is still under development, and the testbed is currently equipped with a Lyot coronograph, the apodizer being replaced by a high-quality mirror. For more details about the investigations on coronagraph designs, please see [7]. Furthermore, thanks to the hole in the FPM, part of the beam can be reused and this coronagraph is compatible with a low-order wavefront sensor [13, 14, 15].
Finally, the wavefront sensing and control system includes:
- Two Boston Micromachines deformable mirrors (kilo-DM), named DM1 and DM2, each of them with 952 actuators in a 9.9 mm diameter disk. DM1 is calibrated and set in a pupil plane [7], and DM2, currently replaced by a flat mirror, is located out of pupil. It will enable active correction for aperture discontinuities (ACAD) and both amplitude and phase control.
- A focal plane camera (CamF), with a motorized translation stage along the optical axis. This translation stage will allow phase diversity applications [16, 17].

The testbed also includes a pupil plane camera (CamP), a 4D AccuFiz interferometer for alignment and wavefront measurements, and convergent mirrors.

Combining all these components, the total wavefront error (WFE) in the testbed is 150 nm RMS without the correction from the DMs. This enables the use of λ/20 surface error optics and an alignment tolerance of 100 to 500 µm, depending on the optic.

C. Environment constraints

To limit air turbulence and dust on the optical components, which would degrade the contrast performance, HiCAT is located in a class 1000 clean room with temperature control in a 1°C range and humidity that is maintained under 40%. Furthermore, the testbed is on a floating table, which is on a platform independent from the rest of the building, to remove vibration effects. A box covers all the testbed to protect it from dust and particles.

In addition to these first protections, the deformable mirrors have stronger constraints, in particular about humidity (below 30%), which lead to the installation of temperature and humidity sensors and a complementary dry air system inside the box containing the optical bench.

This air supply may create unwanted turbulence effects in the bench box, that might make the wavefront more unstable and so high-contrast imaging implementation more challenging. This is why we plan to make this supply external by limiting the humidity in the entire room below 30%, which would minimize the turbulence inside the box.

D. Timeline and first results

The HiCAT testbed was fully aligned in Summer 2014, except for the three deformable mirrors (2 Boston-DMs and the Iris-AO segmented mirror) and the apodizer. This alignment resulted in a wavefront errors of 12±3 nm RMS (instead of the 150 nm RMS required in section B) over an 18mm circular pupil, after passing through an optical train of 15 components. Fig. 2 shows the direct and coronagraphic images obtained at the end of the testbed after this alignment. The direct image corresponds to a nice Airy diffraction pattern, with seven visible rings. Furthermore, the coronagraphic image shows a lot of speckles, that result from the residual wavefront errors of the testbed, and was not optimized at that time since the FPM and the Lyot Stop were not optimally centered yet.

In 2015, the first DM was calibrated and integrated into the testbed, and replaced a flat mirror located in a pupil plane. After alignment of the DM, we obtained a wavefront error of 13±3 nm RMS, which makes us hope for very good results after implementation of wavefront control.

Fig. 2. Direct and coronagraphic PSF obtained at the end of the HiCAT testbed, in log scale, but not on the same dynamical range. The camera was not precisely aligned yet. A 18mm circular pupil was used, combined with a 10mm Lyot Stop. The source gives a monochromatic light at λ=640nm.

Unfortunately, the first wavefront control tests could not be achieved, due to an instability issue, that is described in [18]. The resolution of this problem is currently on going, and once it will be solved, the wavefront control implementation will go on, leading to the installation of the second DM to apply also the ACAD method.
Furthermore, the apodizer design studies is on-going and gives very promising results and an apodizer should be added to the optical bench.

III. WAVEFRONT SENSING AND WAVEFRONT CONTROL: THE WAY TO HIGH-CONTRAST:

Since Brown and Burrows set [19] the typical requirements in term of contrast for exoplanet detection, the theoretical feasibility of starlight subtraction in a so-called dark hole thanks to a DM has been proven [20]. Since 1995, several methods have been developed and have proven their efficiency. For a faster correction, most of them separate estimation and control of the wavefront, using complementary sensing methods such as the Self-Coherent Camera (SCC) [21] and the COronagraphic Focal-plane waveFront Estimation for Exoplanet detection (COFFEE) [22]. In this section, we introduce different well-known wavefront control algorithms.

A. Speckle Nulling

The Speckle Nulling (SN) is the first algorithm that was developed and has already several times been experimentally tested and has proven its efficiency and its robustness in broadband light [23, 24].

It is an estimation-free approach, that only uses the image on the science camera. Furthermore, this method typically focuses on the correction of the brightest speckle in the dark hole. This is why it has to be applied many times in a row to correct for speckles in the entire dark hole.

It is based on the relationship between one speckle in a focal plane and a sine error phase in a pupil plane, where the DM is located. This is why a speckle on the science camera plane will be corrected by applying a sine command on the DM surface, such as:

\[ \varphi_{DM} = A \times \sin(\varphi_0 + 2\pi \times (f_x x + f_y y)) \] (1)

where \( A \) is the amplitude of the sine, \( \varphi_0 \) is the origin phase and \( f_x \) and \( f_y \) are the spatial frequencies of the phase and can easily be computed from the position \((\alpha, \beta)\) of the speckle.

Since the relationship between the electric fields in DM plane and in science camera plane is linear, the links are easy between position of the speckle and frequencies of the sine function and between intensity in the image and amplitude of the sine function. This second step also implies to acquire a flux reference prior to the SN to get the DM response. This algorithm is illustrated in Fig. 3 and explained in details in [18]. The main issue is to find the phase offset \( \varphi_0 \). This is done by testing different phase offsets and selecting thanks to an interpolation the one that gives the best results on the corrected image in term of speckle extinction.

Fig.3. SN algorithm principle. The coronagraphic PSF is used as an input. In the desired dark hole, the brightest speckle is selected. Thanks to its position, the frequency of the optimal correction sine function is computed. The amplitude of the correction function is obtained thanks to a calibration step and the intensity of the speckle. The optimal phase offset is deduced from test commands that are sent to the DM before selection of the most efficient one. After correction of the speckle, the second brightest speckle is selected in the dark hole, and the algorithm is applied again.

This algorithm is interesting since it does not need any prerequisite, except for a fast calibration of the deformable mirror and a reference flux calibration. However, this method has some drawbacks, such as the absence of physical constraints on the DM: by adding as many sine commands as speckles in the dark hole, huge actuator strokes can be produced, which is not realistic. Furthermore, the correction is very slow, since one speckle is corrected at each step of the control loop. Many hundreds of thousands of correction iterations are necessary to correct for the electric field in the entire dark hole, which is not compatible with space missions. This is why, after this first test, we want to focus on other methods, which are directly correcting for the wavefront in the entire dark hole.
B. Speckle Field Nulling

The Speckle Field Nulling (SFN) algorithm [25] brings an improvement to the SN as it proposes an enhanced convergence speed. This method is a generalization of the previous one and is based on the minimization of the speckle energy in all over the dark hole.

The electric field $E$ in the pupil plane can be expressed as:

$$\forall (x,y) \in Pup, E(x,y) = P(x,y) \times e^{A(x,y)+i\varphi(x,y)+i\varphi_{DM}(x,y)}$$  \hspace{1cm} (2)$$

where $P$ is the pupil function, $A$ is the amplitude aberration, $\varphi$ is the phase aberration, and $\varphi_{DM}$ is the phase correction, equivalent to the DM surface equation. If we consider the amplitude aberrations as negligible and the correction efficient enough so that $\varphi + \varphi_{DM}$ is small, then we have the following linearization:

$$\forall (x,y) \in Pup, E(x,y) \approx P(x,y) \times (1 + i\varphi(x,y) + i\varphi_{DM}(x,y))$$ \hspace{1cm} (3)$$

After passing through the coronagraph, modeled here as a linear function $C$, the electric field in the detector plane $E_f$ can be expressed as following:

$$\forall (\alpha,\beta) \in DH, E_f(\alpha,\beta) \approx C\{P(x,y)\} + iC\{P(x,y)\varphi(x,y)\} + iC\{P(x,y)\varphi_{DM}(x,y)\}$$ \hspace{1cm} (4)$$

where $C\{P\}$ is the image of the star, $iC\{P\varphi\}$ corresponds to the field of speckles, and $iC\{P\varphi_{DM}\}$ is the correction brought by the DM as seen in the detector plane. Since the image of the star is cancelled by the coronagraph, this expression becomes:

$$\forall (\alpha,\beta) \in DH, E_f(\alpha,\beta) \approx iC\{P(x,y)\varphi(x,y)\} + iC\{P(x,y)\varphi_{DM}(x,y)\}$$ \hspace{1cm} (5)$$

Furthermore, the DM surface can be expressed as the sum of the contribution of its different actuators, meaning:

$$\forall (x,y) \in Pup, \varphi_{DM}(x,y) = \sum_{k=1}^{N} \sum_{l=1}^{N} a_{kl} f_{kl}(x,y)$$ \hspace{1cm} (6)$$

where $(a_{kl})_{(k,l)\in[1,N]^2}$ are the actuator strokes and $(f_{kl})_{(k,l)\in[1,N]^2}$ are the influence functions of the DM, obtained by calibration.

Furthermore, since the objective of this method is to obtain:

$$\forall (\alpha,\beta) \in DH, E_f(\alpha,\beta) = 0$$ \hspace{1cm} (7)$$

by combining this equation with (4) and (5), we finally get:

$$\sum_{k=1}^{N} \sum_{l=1}^{N} a_{kl} C\{P f_{kl}\} = -C\{P \varphi\}$$ \hspace{1cm} (8)$$

This corresponds to a linear system in $(a_{kl})$ and can easily be solved using Singular Value Decomposition (SVD) or Fourier Expansion.

This algorithm proposes a good improvement in term of computing speed from the previous method, since the electric field is corrected in the entire dark hole at once. However, it requires a preliminary pupil plane wavefront sensing step and is based on a model for the coronagraph, that needs to be realistic, and a complex calibration of the deformable mirror. Furthermore, once again, there is no constraint on the deformable mirror, which means big strokes can be produced.

C. Electric Field Conjugation

The Electric Field Conjugation algorithm (EFC) is well described in [26, 27]. This algorithm needs an estimation of the amplitude of the electric field in the focal plane. As in the previous method, the objective is to satisfy the requirement of (6).

This approach is also based on the expression of the electric field in the pupil plane, given in (2), but the linearization is different, since we consider that a small correction phase is applied:

$$\forall (x,y) \in Pup, E(x,y) \approx P(x,y) \times e^{A(x,y)+i\varphi(x,y)} (1 + i \varphi_{DM}(x,y))$$ \hspace{1cm} (9)$$
After passing through the coronagraph, the electric field becomes:

\[ \forall (\alpha, \beta) \in DH, E_f(\alpha, \beta) \approx C\{Pe^{A+i\psi}\} + iC\{\varphi_{DM}\} \tag{10} \]

A preliminary calibration step furnishes the relation between the voltages \( v \) applied on the DM and the electric field on the detector plane. This relationship is modeled thanks to a so-called interaction matrix \( G \), defined as:

\[ C\{\varphi_{DM}\} = Gv \tag{11} \]

By combining this equation, (10) and the criterion (6), we then get:

\[ C\{Pe^{A+i\psi}\} + iGv = 0 \tag{12} \]

which is equivalent to:

\[ v = G^\dagger (iC\{Pe^{A+i\psi}\}) \tag{13} \]

where the generalized inverse matrix of \( G, G^\dagger \), can be obtained by SVD. Finally, since \( v \) is necessary real, we have:

\[ v = \begin{bmatrix} \text{Re}\{G\} \\ \text{Im}\{G\} \end{bmatrix}^\dagger \begin{bmatrix} \text{Re}\{iC\{Pe^{A+i\psi}\}\} \\ \text{Im}\{iC\{Pe^{A+i\psi}\}\} \end{bmatrix} \tag{14} \]

where \( C\{Pe^{A+i\psi}\} \) is the electric field in the image plane, obtained thanks to wavefront sensing.

This algorithm, such as the previous one, needs to be complemented with a wavefront estimator, and is based on a calibration of the deformable mirror. There is again no constraint on the DM strokes, and the linearization of the correction phase implies that it has to remain small. But, in opposition to the SFN method, it does not require any model for the coronagraph.

**D. Stroke Minimization**

The problem of these last three methods is that they do not take into account the physical constraints of the deformable mirror, in particular the limited strokes of the actuators. The last algorithm we introduce here, named Stroke Minimization (SM), answers to this issue. The idea here is to minimize the quadratic sum of the actuator strokes, \( \frac{1}{2} \sum_{k=1}^{N} \sum_{l=1}^{N} a_{kl}^2 \), regardless a constraint in contrast:

\[ \varepsilon \leq 10^{-C} \tag{15} \]

where \( \varepsilon \) corresponds to the total energy in the dark hole and is defined as:

\[ \varepsilon \equiv \langle E_f, E_f \rangle = \int_{DH} |E_f(\alpha, \beta)|^2 \delta \alpha \delta \beta \tag{16} \]

To combine these two requirements, we define a criterion, named \( \varepsilon_M \) and defined as:

\[ \varepsilon_M = \frac{1}{2} \sum_{k=1}^{N} a_k^2 + \mu \times (\varepsilon - 10^{-C}) \tag{17} \]

where \( \mu \) is a weighting parameter that needs to be optimized.

In practice, \( \mu \) is set at a small value. \( \varepsilon_M \) is minimized by setting its derivative to 0 and we obtain a value for the actuator strokes \( \langle a_{kl} \rangle_{(k,l) \in [1,N]^2} \). Then, if the contrast constraint is respected, the algorithm is finished, if not, the value of \( \mu \) is increased, which is equivalent to increasing the importance of the contrast constraint in the criterion \( \varepsilon_M \), and the algorithm is applied again, until the requirements are satisfied.

To have a complete description of this algorithm, its formalism, and results, please refer to [28].

Even if this algorithm is more complex than the previous ones and requires a calibration step and a wavefront sensing, it is very efficient and if the wavefront estimation is well done, it can be also very fast. It also takes into account the physical limitations of the DM, but its main advantage is that it can include multiple DMs to enable symmetric correction in the dark hole, by correcting for both amplitude and phase aberrations, which is not possible with one single DM.
E. Non-linear Dark Hole

In opposition to these methods that are all using a linearization of the electric field, another algorithm was developed, the so-called Non-Linear Dark Hole (NLDH) [29]. The idea of this method is the minimization of the energy $\varepsilon$ in the dark hole, as defined in (16). Without any linearization, this energy is equal to:

$$\varepsilon = \left\| \mathcal{C}(P \times e^{A+\phi+\phi_{DM}}) \right\|^2$$

(18)

where:

$$\phi_{DM} = F \nu$$

(19)

$F$ contains the influence functions $(f_k)_{(k,l) \in [1,N]^2}$.

This energy is minimized to obtain an optimal voltage vector $\nu$ thanks to a numerical minimization based on the Variable Metric with Limited Memory and Bounds (VMLM-B) method [29]. The key point to optimize the minimization is to compute the analytical expression of the gradient of the coronagraphic PSF with respect to the aberrations upstream of the coronagraphic mask. The computation of this term $\delta e / \delta \phi$ is explicitly detailed in the thesis of B. Paul, appendix B.0.3.

IV. CONCLUSIONS AND PERSPECTIVES:

The HiCAT testbed will enable high-contrast imaging studies for telescopes with segmented apertures, spiders and/or central obstruction. It is designed as a coronagraph completed with two deformable mirrors to perform wavefront sensing and wavefront control and address both phase and amplitude aberrations.

In parallel to this experimental study, we plan to focus on simulation of several well-known wavefront control algorithms, that we introduced in the second part of this paper. Even if SN is the only one that does not need any wavefront estimation and limits the number of prerequisites (no need of a model for the coronagraph, and a fast single calibration step), it is the longest to converge for a correction on the entire dark hole. This is why, after this first test, it appears necessary to focus on other methods, such as SFN, EFC, SM, and NLDH. As indicated in Fig. 4, these methods require wavefront estimation, a calibration of the deformable mirror actuators with the knowledge of the influence functions or the interaction matrix, and a realistic model for the coronagraph, in the cases of SFN and NLDH.

The objective now is comparing the performance in contrast of these different methods, but also their requirements and their robustness to realistic space-like environment conditions, such as jitter. After this study, a method should be selected to be implemented on the HiCAT testbed for further studies.

Fig. 4. Prerequisites of the different wavefront control algorithms that have been introduced in this paper. In opposition to the last four algorithms, the SN does not need any estimation, but this is the slowest method. The five options require a calibration of the deformable mirror actuators with the knowledge of the influence functions or the interaction matrix, and also a realistic model for the coronagraph, in the case of SFN or NLDH.

REFERENCES


[27] A. Give' on, B. D. Kern, and S. Shalayan, “Pair-wise, deformable mirror, image plane-based diversity electric field estimation for high contrast coronagraphy”, *SPIE* 8151, 815110, October 2011.
