Phasing segmented telescopes via deep learning methods: application to a deployable CubeSat

Maxime Dumont,[1,2,3,*] **Carlos M. Correia,**[2,4] **Jean-François Sauvage,**[1,3]
Noah Schwartz,[5] **Morgan Gray,**[3] **and Jaime Cardoso**[2,6]

[1]*DOTA, ONERA, F-13661 Salon Cedex AIR, France*
[2]*Faculdade de Engenharia da Universidade do Porto, Rua Dr. Roberto Frias, s/n, 4200-465 Porto, Portugal*
[3]*Aix-Marseille Université, CNRS, CNES, LAM, Marseille, France*
[4]*Center for Astrophysics and Gravitation, Instituto Superior Técnico, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal*
[5]*UK Astronomy Technology Centre, Edinburgh EH9 3HJ, UK*
[6]*INESCTEC, Porto, Portugal*
*maxime.dumont@lam.fr

Capturing high-resolution imagery of the Earth's surface often calls for a telescope of considerable size, even from low Earth orbits (LEOs). A large aperture often requires large and expensive platforms. For instance, achieving a resolution of 1 m at visible wavelengths from LEO typically requires an aperture diameter of at least 30 cm. Additionally, ensuring high revisit times often prompts the use of multiple satellites. In light of these challenges, a small, segmented, deployable CubeSat telescope was recently proposed creating the additional need of phasing the telescope's mirrors. Phasing methods on compact platforms are constrained by the limited volume and power available, excluding solutions that rely on dedicated hardware or demand substantial computational resources. Neural networks (NNs) are known for their computationally efficient inference and reduced onboard requirements. Therefore, we developed a NN-based method to measure co-phasing errors inherent to a deployable telescope. The proposed technique demonstrates its ability to detect phasing errors at the targeted performance level [typically a wavefront error (WFE) below 15 nm RMS for a visible imager operating at the diffraction limit] using a point source. The robustness of the NN method is verified in presence of high-order aberrations or noise and the results are compared against existing state-of-the-art techniques. The developed NN model ensures its feasibility and provides a realistic pathway towards achieving diffraction-limited images. © 2024 Optica Publishing Group

https://doi.org/10.1364/JOSAA.506182

1. CONTEXT AND MAIN CHALLENGES

The angular resolution of telescopes is fundamentally limited by the size of their aperture for a given wavelength. For example, a telescope with a diameter of 30 cm has a ground sampling of 1 m at a distance of 400 km or 2 m at 800 km at visible wavelengths ($\lambda = 800$ nm). In the fields of Earth observation and astronomy, high-angular-resolution images are crucial for maximizing the scientific output and return on investment, which has driven the demand for larger apertures. However, launching large monolithic pupils into space is expensive, and designing and manufacturing such telescopes pose significant challenges [1]. An effective solution to increase aperture size is to fragment the aperture into smaller segments. By folding the telescope inside a compact volume during launch and deploying it in orbit, both the collecting power and angular resolution can be increased. This volume optimization and gain in mass translate directly into cost savings, enabling the deployment of a constellation with multiple high-angular-resolution platforms and therefore achieving higher revisit rates (i.e., temporal resolution).

Improving both the revisit rate and angular resolution finds applications in Earth climate monitoring, defense, and security for which the highest commercially available spatial resolutions today are given by the satellite Pleiades [with a ground sampling distance (GSD) of 70 cm, at a rate of one image per day] as well as Pleiades-Neo (30 cm GSD color images). For example, flood monitoring (accounting for over 40% of the natural disasters) currently provides information either over large areas or with limited spatial resolutions and a revisit every two days (e.g., Copernicus with a 5 km resolution), or instead with a revisit time at best of a few days yet with a high spatial resolution (e.g., Pléiades, Sentinel-1, and Sentinel-2). Solar system exploration can potentially also benefit from this improvement: although the HRI camera of Deep Impact has an aperture larger than 30 cm, the typical size is 10 to 20 cm (NAC camera on Rosetta, Cassis camera on TGO, New Horizon camera).

The objective of this study is to enable, in ESA's own parlance *very very high-angular-resolution* imaging in the visible spectrum for Earth observation from space at a low cost by

quantifying the performance of a NN-based phasing algorithm during pre-launch, deployment, and exploitation phases. To achieve this, we use AZIMOV [2] as a reference design for our study. AZIMOV is a satellite concept featuring a 30 cm wide, diluted aperture telescope onboard a 6U CubeSat. It is capable of reaching a 1 m GSD in the visible (i.e., $GSD = \lambda/D * z = 800.10^{-9}/0.3 \times 400000 \approx 1$ m) from LEO with a swath of over 4 km. Figure 1 illustrates a simplified concept of the payload. This technology allows the telescope's aperture to surpass the size of the supporting platform, thereby enhancing the telescope's light-gathering power and improving its achievable resolution.

To achieve the theoretical angular resolution for a given segmented pupil, the individual mirror segments must be co-phased with a precision typically on the order of a fraction of the imaging wavelength (usually $\lambda/14$ [3]). Compensation for static and dynamic phase aberrations is required for each segment. High-precision actuators are utilized to control the position of the deployable primary mirror segments in piston and tip-tilt [4] (PTT). For visible wavelength imaging at the metric resolution limit (we select $\lambda = 800$ nm as the study wavelength), the wavefront error (WFE) after correction should be less than 50 nm RMS. However, due to actuator response and dynamics, a portion of the WFE error is independent of sensing. From a dedicated study of this error budget, Sauvage *et al.* [5] allocate $\epsilon_{WFS} < 15$ nm RMS to the sensing and reconstruction contribution. Our study proposes a solution to perform accurate WFS onboard a CubeSat, using directly a single focal plane image and deep learning.

This paper is organized as follows: we first provide an introduction to the study and the motivation behind it. Second we delve into the concepts of image formation in the context of wavefront sensing. The methodology employed in our research is detailed in Section 4. Then, Section 5 presents the results obtained from our experiments, analyzing the performance
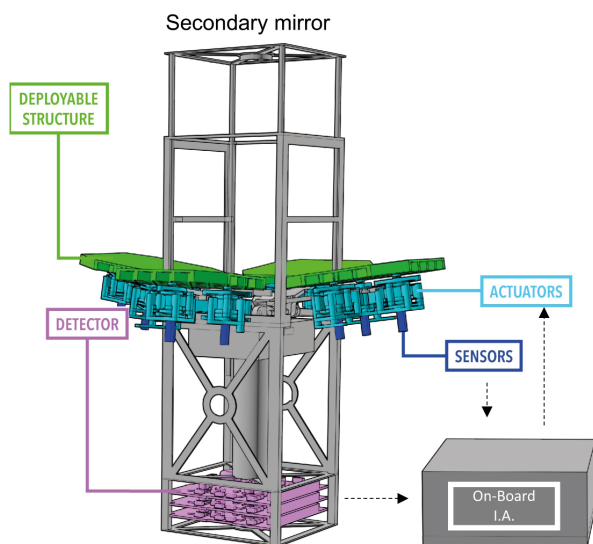
and robustness of the proposed methods. Section 6 discusses the implementation of neural networks in a complete phasing process specifically tailored for the AZIMOV system. Finally, Section 7 concludes the paper.

2. PHASING SEGMENTED MIRROR TELESCOPES

The phasing of segmented mirror telescopes is a critical aspect of their operation, as it directly impacts the quality of their optical performance. Ground-based telescopes such as the segmented, twin Keck telescopes correct for tip-tilt errors using a Shack-Hartmann wavefront sensor (SH-WFS) [6]. However, the SH-WFS is known to be insensitive to global piston error. To overcome this issue, a micro-lenslet array is placed at the junction of two segments on the primary mirror, and allows to identify the deferential piston between two segments [6]. More recently, for the Giant Magellan Telescope, and following roughly the same principle, a differential piston between segments will be identified using a holographic dispersed fringe sensor (HDFS) [7]. It employs a holographic optical element to disperse the incoming light into multiple wavelength channels. By analyzing the interference patterns created by these channels, the HDFS can accurately measure the differential piston. For space-borne telescopes such as the JWST [8], the phasing strategy is divided into three parts: i) segment location-identification, to identify the coarse position of each segment; ii) co-phasing to correct the global phasing error of the 18 segments; iii) finally, wavefront monitoring and maintenance to ensure a correct phasing throughout the life of the mission. The coarse and fine phasing is ensured by a dispersed fringe SH-WFS and weak lens hardware [9]. In this context, this step-by-step approach has inspired our work. However, we propose an innovative approach to co-phase segmented telescopes by using a single focal plane image as the observable, and deep learning to tune our NN-based algorithm. Single-image approaches can greatly simplify the phasing process and enhance the optical performance of future segmented telescopes. Our work focuses specifically on the primary mirror phasing from the point spread functions (PSFs) obtained while imaging a point source object. The use of a point source is a key element in our approach: first to understand the behavior and NN specifics to process focal plane images and estimate wavefront errors from them, and second, to be able to generalize to extended objects. Lastly, the phasing process of AZIMOV is not defined yet. One strategy could be to undertake the coarse phasing imaging of a distant unresolved star and then rotate the telescope towards the Earth to proceed with the fine phasing and capture images for further ends. The use of NN on point sources is therefore central to many applications. Our work aims to employ a computationally efficient phasing method taking into account the typical CubeSat limitations especially in terms of volume and computing power. In particular, we demonstrate the potential of using NN in contrast to more classical optimization methods, focusing on execution time as a proxy to computational complexity. In this context, we use the time required for NN inference to identify a suitable NN model capable of reaching the diffraction limit with minimal computational burden. The AZIMOV payload is currently being designed, and requirements for the

Fig. 1. Simplified illustration of the deployable payload concept: deployable structures (primary and secondary mirrors, baffle), actuators to adjust mirror positions, sensors to measure mirror positions, detector to assess image quality, and onboard artificial intelligence to control and adjust mirror positions to reach diffraction-limited image quality (i.e., active optics). From Schwartz [2].

onboard computer (OBC) are still being finalized. Among other aspects, the OBC will need to run the compute intensive co-phasing algorithm, the data-intensive image processing algorithm (including algorithms such as image registration and super-resolution), and other housekeeping tasks. The satellite being in LEO, it will receive a varying thermal load (approx. every 90 min): multiple updates per orbit will be required to maintain good phasing performance. It is important to minimize the co-phasing algorithm complexity and therefore execution time. The exact execution time will naturally depend on the selected OBC.

A. Classical Wavefront Sensing Approaches

Various methods have been proposed for measuring and estimating wavefront error, broadly classified into two categories: direct wavefront sensing and focal plane methods. Direct wavefront sensing relies on a relatively simple algorithm, at a cost of an optical path specific to the wavefront sensing with dedicated hardware, such as the JWST's dispersed fringe Shack-Hartmann WFS [8].

An alternative method involves indirect measurement, where the wavefront is deduced from the focal plane image captured by the detector. One approach is image sharpening, as proposed in the work of Schwartz [2,10], which entails acquiring a sequence of images and iterating starting from an initial guess of the aberrations. For each of these images, an image quality metric, such as maximum intensity or image contrast, is computed and then used in a non-linear iterative optimization algorithm to estimate the aberration. In contrast, model-based focal plane wavefront sensing methods typically rely on introducing diversity (whether in phase, e.g., phase diversity [11], wavelength, or amplitude) to resolve sign-ambiguity [12] of even modes. A best fit is sought to the focal plane images where from the wavefront aberration parameters are estimated. The algorithms are typically more complex than those used in direct wavefront sensing due to the relationship between the image intensity in the focal plane and the electric field in the pupil plane. Nevertheless, considering the constraints posed by Earth observation and the non-repeatability of scenes, the goal is to conduct focal plane wavefront sensing using only a single image. Data-driven, deep-learning-based methods can also be applied successfully in focal plane wavefront sensing (FPWFS). One of the main advantages lies in the fact that once the network is trained, the model directly outputs the estimated wavefront without requiring iterations. The computation gains time and, consequently, efficiency in calculations. The network's parameters have been learned during the training phase, where all the computational burden is handled before launch. Therefore, once onboard, the process is relatively fast and not contingent on the uncertainty of the number of iterations. The training steps build numerically and iteratively the NN model from the data. Table 1 compares qualitatively the three methods cited above.

B. Motivation for New Phasing Approaches for Space-Borne Segmented Telescopes

Recently, access to space has shown significant changes characterized by the emergence of the "NewSpace" paradigm, which

Table 1.   Feature Comparison of FPWFS Methods

	Deep Learning Method	Image Sharpening	Phase Diversity
Data driven	Yes	No	No
Model-free	Yes	Yes	No
Iterative	No	Very	Yes
Onboard computing	Deterministic	Optical optimization	Numerical optimization
Off-board computing	Stochastic	None	None
Initial guess	No	Yes	Yes

emphasizes cost-effective access to space and introduces novel constraints such as limited volume allocation. Therefore, this necessitates the development of innovative methodologies, such as those shown in this paper. By leveraging a single focal plane image, we estimate and subsequently correct low-order phase aberration piston and tip-tilt using neural networks (NNs) and a space-borne, four-petal deployable CubeSat as the reference design. CubeSats are not only constrained in volume but also in computing power. Therefore, the complexity of the model will be considered in this study, as it plays a crucial role in minimizing optical aberration control for successful implementation on small satellites.

C. Previous Work on NN for Focal Plane Wavefront Sensing

The utilization of deep learning techniques has been proposed in previous studies as a powerful tool for FPWFS, often aiming for performance levels similar to or beyond the state of the art [13–15]. The majority of NN methods for FPWFS employ a single defocused image, as demonstrated in studies such as [13,15–19]. Others have used an optical preconditioner (overexposure, defocus, scatter) to improve the performance of the WFS [20]. Finally, others have used a pair of in- and out-of-focus images to identify Zernike coefficients from a pair of PSFs. [21]. These works have shown excellent performance for wavefront sensing using deep NN architectures, with Xception and Inception v3 being the most widely used. Xception utilizes a depthwise separable convolution approach, which separates spatial and channel-wise convolution operations. These architectures have been, for example, employed by Paine [15], Rajaoberison [17], Andersen [18], Nishizaki [20], and Orban [19]. Other popular architectures such as Resnet 50 and VGG-16, have also been tested by Paine [15] and Orban [19]. Architectures like EfficientNet-B4 (used by Quesnel [21]), Bi-GRU for sequential data (employed by Wang [16]), and Dense NN (used by Rossi [13]) also have demonstrated effectiveness for focal plane wavefront sensing. In these studies, two types of focal plane wavefront sensing have been identified. The majority of them focus on Zernike mode identification for a monolithic pupil imager (Paine [15], Andersen [18], Nishizaki [20], Orban [19], and Quesnel [21]), while others aim at phasing a segmented pupil by identifying low-order modes (Wang [16] focuses on piston and tip-tilt; Rossi [13] and Rajaoberison [17] focus on piston only). The major shortcoming of these studies is that the NN architecture requires hundreds of thousands to hundreds of millions of parameters (e.g., Resnet 50 and VGG-16 require,

respectively, 23 million and 138 million parameters). Our goal, apart from dealing with specific segmentation errors, is to minimize the number of parameters in order to optimize the inference time. For this reason we need to explore new suitable architectures will smaller dimensionalities while keeping a bottleneck architecture such as VGG-net and Resnet.

3. FORWARD MODEL

A. Image Formation

The image of a point source, or PSF, can be described as

$$\text{PSF} = |\text{FT}(Pe^{j\phi})|^2 = \left|\text{FT}\left(Pe^{j(\phi_{\text{PTT}}+\phi_{\text{HO}})}\right)\right|^2, \quad \textbf{(1)}$$

where FT is the Fourier transform, and P is the pupil transmission function, with one inside the pupil boundaries and zero outside. ϕ is the phase function, which can be decomposed as $\phi = \phi_{\text{PTT}} + \phi_{\text{HO}}$ corresponding, respectively, to the PTT phase and the higher-order (HO) phase. Also, $\phi = \frac{2\pi\Delta}{\lambda}$, where Δ is the corresponding optical path difference (OPD) and the wavelength $\lambda = 800$ nm.

The AZIMOV prototype (Fig. 1) is composed of a pupil with four deployable petals. The precise design of the AZIMOV pupil is not finalized today. The global shape is four segments with an approximately square shape, but the final design might evolve due to the reduced volume constraints. In particular the four segments might not be symmetric in the final design. In this paper, we propose a simplified pupil composed of four squared segments where two segments are 10% cropped on the edge in order to break the pupil centro-symmetry as shown in Fig. 2. The asymmetric pupil Fourier wavefront sensor (APF-WFS) in Martinache's work shows great potential for small aberrations on an asymmetric pupil. Pope [22] shows an interest in using Martimache's work for segmented pupil, retrieving a piston value for an interval of $\frac{\lambda}{3.5}$. However, it has not been studied for large tip-tilt errors, a subject discussed in this paper. This method appears to be faster than our approach (only requiring a FFT) but lacks the ability to estimate large aberrations and only works on a point source—essential considerations for AZIMOV. In our case, cropping 10% of the segments is chosen as a trade-off between the loss of the collecting area and the correct wavefront estimation: the more asymmetric the pupil is, the better the even parts of the phase will be identified [12]. However, a limit is shown in the ratio of even part identification over the percentage of the cropped area. For instance, doubling the cropped area (from 10% to 20% of two segments) improves the estimation by 25%. For our space-borne application, the pupil design and the need to use the images from the science detector directly compel us to use instead an amplitude diversity as shown in Fig. 2.

To address the specifics of a segmented pupil, we define our modal basis with three modes $B_{i=1\text{-}3}^{k=1\text{-}4}$ on each of the four segments, i.e.,

$$\phi = \sum_{i=1}^{3} \sum_{k=1}^{4} c_i^k B_i^k, \quad \textbf{(2)}$$
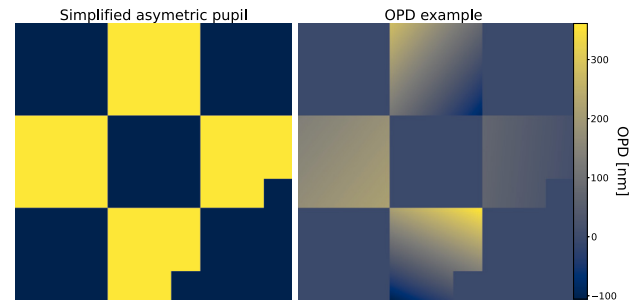
where the modes are

Fig. 2. Example of the pupil and OPD map.

$$\begin{aligned} B_1^k(x, y) &= 4, \\ B_2^k(x, y) &= 4\sqrt{3}x, \\ B_3^k(x, y) &= 4\sqrt{3}y, \end{aligned} \quad \textbf{(3)}$$

representing, respectively, piston, tip, and tilt on segment k. Clearly these modes are orthogonal across different segments and their inner product:

$$\langle B_i, B_j \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}. \quad \textbf{(4)}$$

The expansion coefficients in Eq. (3) are easily found when establishing an orthonormalized basis with unitary root-mean-square (RMS) over the whole, un-cropped, pupil in Fig. 2:

$$\frac{1}{4}\sqrt{\frac{1}{4}\int_{-1}^{1}\int_{-1}^{1} M_i M_j \mathrm{d}x\mathrm{d}y} = 1 \quad \text{if } i = j, \quad \textbf{(5)}$$

where the 1/4 in front of the square root represents one-fourth of the pupil, i.e., the relative area of one segment, and the integral is computed on normalized units; and M_i represents the phase mode of one squared pupil. For instance M_1 corresponds to a piston on a squared segment, when the three others are set to zero. Equation (5) allows to identify normalization coefficients in Eq. (3).

Here we allow ourselves a slight lack of mathematical rigor in that the integration bounds, on account of the removed portion shown in Fig. 2, would not be exactly 1 and -1. Since the amount of amplitude diversity needed to lift the even mode sign ambiguity is an open optimization parameter, we gloss over this detail and assume the same normalization for each and all segments.

The WFE definition for segmented pupils differs from the definition over monolithic pupils as shown in Eq. (6):

$$\text{WFE}(\{c_i^k\}) = \sum_{k=1}^{4} \sqrt{\sum_{i=1}^{N} c_i^{k\,2}}. \quad \textbf{(6)}$$

This WFE is described by a quadratic sum of the phase coefficients over each segment. Provided the segments are non-overlapping, the total RMS is the sum of individual RMS values and not their quadratic sum.

An example of the PSFs studied is shown Fig. 3. These PSFs, paired with their Zernike coefficients, are directly used for the NN training.
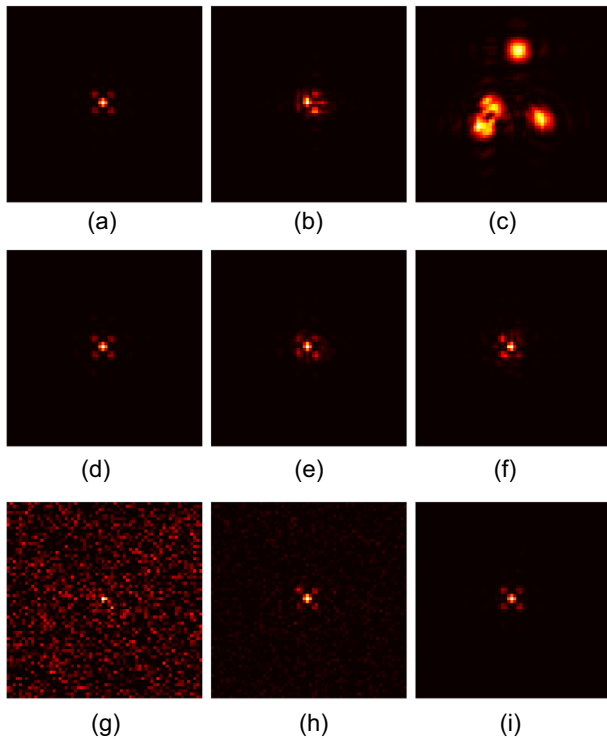
Fig. 3. Illustration of AZIMOV PSFs obtained in different configurations of WFE and signal to noise ratio (SNR). First row is PSF disturbed by PTT aberrations only. (a) PSF diffraction limit, (b) PSF $\text{WFE}_{\text{PTT}} = 97$ nm, (c) PSF $\text{WFE}_{\text{PTT}} = 454$ nm. Second row is PSF disturbed both by PTT and HO aberrations. (d) PSF diffraction limit + 30nm HO, (e) PSF $\text{WFE}_{\text{PTT}} = 96$ nm $\text{WFE}_{\text{HO}} = 30$ nm, (f) PSF $\text{WFE}_{\text{PTT}} = 97$ nm $\text{WFE}_{\text{HO}} = 60$ nm. Last row is the noise influence depending on the SNR. (g) PSF diffraction limit, SNR = 10, (h) PSF diffraction limit, SNR = 30, (i) PSF diffraction limit, SNR = 100.

B. Data Generation

The training of NNs necessitates a substantial volume of data to effectively learn from diverse realizations. Unfortunately such an amount of real data is not available yet; therefore, data are simulated. All the generated imagery data are Nyquist sampled with a field of view (FoV) of $64 * 64$ pixels at $\lambda = 800$ nm. We selected this field of view (FoV) based on the anticipated amplitude of the wavefront error (WFE) to be measured. The PSF's global intensity is normalized to one. Figure 4 presents the two studied distributions, one for large WFE errors corresponding to the initial mirror segments at the deployment (i.e., coarse phasing), and the other for smaller WFE errors for the final fine phasing step. The distribution of the 12 piston-tip-tilt coefficients follows a Gaussian distribution centered on 0 nm with $\sigma_{\text{fine}} = 10$ nm or $\sigma_{\text{coarse}} = 70$ nm. However, identification of pistons, under a single-image FPWFS setting, is constrained to a range between $-\frac{\lambda}{4}$ and $\frac{\lambda}{4}$ or their estimate can suffer from the so-called lambda ambiguity. We constrained piston values on a $\frac{\lambda}{2}$ width interval so that no piston can be identified at $+\frac{\lambda}{2}$ while another would be identified as $-\frac{\lambda}{2}$ since it will lead to the same PSF. Therefore, for the coarse phasing, the piston values are randomly drawn following the same distribution as for the fine phasing ($\sigma_{\text{coarse}}^{\text{piston}} = \sigma_{\text{fine}}^{\text{piston}}$) so that the same NN model studied in the following section can be used also in Section 6.
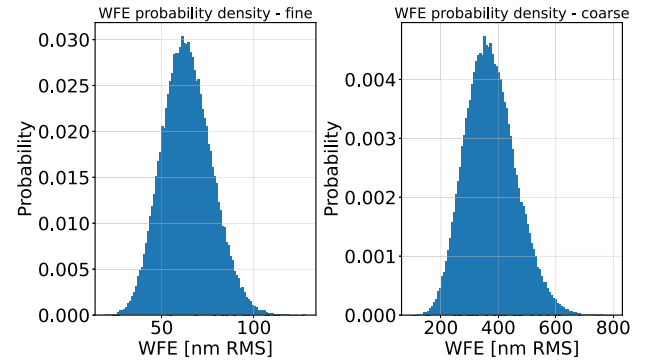

Fig. 4. Probability density of WFE over the full pupil for the training set calculated over 10^6 samples using Eq. (6). The figure on the left corresponds to the fine phasing. On the right, the distribution of the coarse phasing, studied in Section 6.

The capture range of the depicted distribution enables a range of minimum and maximum values within approximately $[-50, 50]$ nm RMS, corresponding to $[-200, 200]$ nm peak to valley (PV) for the piston and $[-350, 350]$ nm PV for tip-tilt. This amplitude is reached after an initial phasing step using the ELASTIC method [23], or our embedded onboard capacitive displacement sensors [2].

The determination of an optimal dataset size, depending on the complexity of the problem and the complexity of the learning algorithm (e.g., nonlinear relationships between input and output features), is controversial as there is currently no definitive methodology for its precise determination. We generated a dataset comprising 10^5 samples randomly split: 90% of the samples are allocated to the training set, while the remaining 10% are assigned to the validation. An additional 10^4 samples are generated exclusively for the testset so that the NN model can be tested with unseen data. All the datasets follow the same coefficient distribution.

C. Disturbance Sources

1. Noise

In a low-flux regime, the image mostly suffers from readout noise, estimated at 2.1 electrons (e.g., using a Sony Pregius-S IMX530 CMOS detector with 5320×4600 pixels) whereas photon noise dominates in high-flux cases. These are the two noises considered in the study, while other sources of noise (background, gain, drift) have been chosen to be ignored. Therefore, we define the signal to noise ratio (SNR) as

$$\text{SNR} = \frac{f}{\sqrt{f + \sigma_{\text{ron}}^2}}, \tag{7}$$

where f is the star flux in numbers of photons and $\sigma_{\text{ron}} = 2.1\text{e}^-$ the readout noise. The photon noise is described by a Poisson distribution $\mathcal{P}(f)$ for each pixel. The readout noise is a Gaussian noise drawn at $\mu = 0$, $\sigma = \sigma_{\text{ron}}$ added to the pixel value. The corresponding noisy images can be found in Fig. 3.

2. Higher-Order Aberrations

Higher-order aberrations introduce additional distortions to the wavefront on top of the PTT. However, their impact is different

from the PTT: although they degrade the PSF and therefore the delivered resolution, they cannot be compensated for. We carry out an analysis of their impact in the system: how they degrade the PTT estimation, and if it is needed to include them in the training set for PTT. These aberrations are added in addition to the PTT in the wavefront. They follow a $1/f^2$ power spectrum, typically encountered with good quality optics [24] and are orthogonal to PTT aberrations. The magnitude of HO aberrations is approximately 30 nm RMS over the full pupil, or in other words $\sum 1/f^2 = 30$ nm RMS, consistent with current estimates for AZIMOV.

4. METHODS

A. Deep Learning

Deep learning is a type of machine learning method able to extract representations or model directly from data with a complex architecture of learning successive layers with a high level of abstraction (increasingly meaningful representations). These architectures leverage a combination of diverse non-linear transformations to accomplish their objectives. The NN architecture has a lot of influence on the output accuracy and is composed of different layers, of which the two primary types are noted below.

• *Fully connected layers* [25]: layers composed of neurons. A neuron computes a weighted sum of its input, plus a bias and then applies a non-linear activation function. The model parameters are therefore the neurons' weights and biases.
• The *convolutional layer* [25], particularly adapted for image processing, computes the convolution of an image with one or several kernels. In this case, the neural network optimizes all the convolution kernels and biases.

The model architecture is created by adding layers one to another. Choosing the right architecture depends on the more important criterion of the project: it can be either the accuracy of the estimation, the model size, or the inference speed. In our work we aim to find the right balance between performance (the precision of WFE estimation) and model size (the size of the NN algorithm that needs to be embedded onboard the CubeSat). The number of model parameters depends on the choice of NN architecture. Decreasing it avoids overfitting and makes the model size smaller. Since our input data consists of 2D images, convolutional NNs have fewer parameters compared to fully connected layers and can capture pixel neighborhood information effectively. The simplest NN architecture to infer coefficients from an image is the VGG-net [26] or the Resnet [27] architecture, which are optimal in terms of number of parameters. Historically, the Resnet architecture is known for the "*skip connection*," allowing the loss gradient to float easily through the deepest layer of the model during backpropagation. To achieve the best trade-off between performance and time complexity, we consider these two highly suitable architectures: VGG-net [26] and Resnet [27].

We therefore propose a NN baseline composed of five convolutional layer stacks as VGG-net architecture (two convolutional layers followed by one max-pooling layer) and three fully connected layers with a ReLu activation function. In the case of the Resnet, two skip connections are added to the
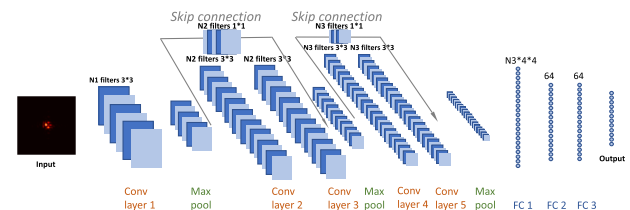

Fig. 5.   Adopted Resnet architecture.

architecture (each skip connection is a convolutional layer with a 1×1 filter): the first one starts after the first layer to merge with the output of the third layer, and the second one after the third layer to merge after the fifth layer. The architecture is presented in Fig. 5 and has been chosen in order to find the smallest architecture suitable for reaching the diffraction limit requirement. Indeed such architecture goes up to 10^6 number of parameters whereas the smallest Resnet found in the literature, the Resnet-18, reaches 10^7 parameters, 10 times more than our heavier network.

N1, N2, and N3 correspond to the number of filters of the convolutional layers. This number of filters impacts the internal parameters and changes the size of the NN model (see Table 2); all model parameters are stored in a double precision float64 format as it is the default format in the Pytorch library. The number of parameters also impacts the computation time and influences the model's performance. Increasing the number of parameters can improve performance until the point of overfitting occurs, leading to limited generalization on new data.

The training step relies on a minimization algorithm of a cost function commonly known as the loss function. Conceptually, a loss function is a way of prioritizing which error to fix from our training samples, so that NN parameter updates result in adjustment to the output, decreasing the loss. The minimization algorithm optimizes the NN parameters (weights, biases, and filters) to minimize the loss function. Usually, the root mean square error is used, but in our study, the WFE over the pupil is computed as the loss function:

$$\text{Loss}_{\text{WFE}} = \text{WFE}\left(c_i^k - \hat{c}_i^k\right), \tag{8}$$

where c_i^k is the true coefficient for mode $i \in [1 \dots 3]$ and segment k, and $\hat{c}_i^k$ is the estimated coefficient. This loss, as mentioned above, allows the model to adjust accurately the model parameters according to the optical criterion on the pupil.

Table 2.   Resnet Model Sizes Depending on the Number of Convolutional Filters[a]

Label	N1	N2	N3	# Parameters	Model Size [MB]
NN1	4	8	16	13.10^3	0.134
NN2	8	16	32	31.10^3	0.278
NN3	16	32	64	94.10^3	0.94
NN4	32	64	128	326.10^3	2.6
NN5	64	128	256	$1.2.10^6$	11

[a]This table approximates VGG-net sizes as well, since Resnet architecture has only slightly more internal parameters due to the skip connection that needs to link the number of filters from the input layer to its output layer.

B. Classical Phase Diversity and Image Sharpening

Other more classical FPWFS methods have also been proposed for the phasing of segmented telescopes, namely, phase diversity and image sharpening. The whole problem of phase diversity (PD) lies in the estimation from the focal plane image and the defocused image of the unknowns, which are the phase and the object [28]. However, in our case, the object is not to be estimated as it is a point source. Since our system does not include any additional optical components and aims to maintain the integrity of the focal plane image without any defocus, it is not possible to acquire a defocused image. Therefore, we propose a single-image phase diversity that minimizes the distance between the real PSF and the fitted one by adjusting the 12 phase coefficients. A criterion based on a quadratic pixel-to-pixel difference between the real PSF and the fitted one is minimized iteratively using the conjugate-gradient method [29]. A Tikhonov regularization is added to the criterion, in order to avoid the divergence of the phase coefficient, particularly at low SNR. The conjugate-gradient has shown the best performance in time, number of iteration, and WFE among others minimization methods.

Another method is image sharpening (IS). By taking focal plane images, the algorithm iteratively maximizes an image maximum intensity centered window by adjusting iteratively the PTT. Classical metrics are the maximum intensity in the center of the image, contrast, or the sum of the pixel squared. Alternative optimization methods are also possible and may lead to faster and improved estimates. Powell optimization methods [30] have shown the best efficiency in terms of noise robustness and computing time. To perform IS, we consider an exposure time of 0.1 ms. This exposure time is taking into account in the method calculus time as an image has to be acquired at each iteration of the algorithm. The computation time of generating the PSFs is deduced from the total computing time.

5. SAMPLE NUMERICAL RESULTS

A. Model Architecture

First, both types of model architectures (VGG-net and Resnet) are trained using identical hyperparameters (*batch size*, *learning rate*, *epochs*) and their inference performances are compared. The performances of the Resnet architecture are observed to systematically outperform the VGG-net even after fine-tuning both model hyperparameters (i.e., other models are trained) enhancing performance with respect to the architecture and number of parameters. The optimal set of hyperparameters for the Resnet includes a progressive scheduled learning rate (LR) depending on the epoch e:

$$\text{LR}(e) = \begin{cases} 10^{-4} & \text{if } e < 70 \\ 10^{-6} + (10^{-4} - 10^{-6})\exp^{-(e-70)*0.0125} & \text{otherwise.} \end{cases}$$

Specifically, a batch size of 32 PSFs is used during 300 learning epochs.

Figure 6 presents the performance comparison between the two architectures mentioned above for various model sizes. Performance is compared using the same test dataset and optimal sets of hyper-parameters, in terms of *residual WFE*, assuming a perfect correction based on the algorithm estimates.
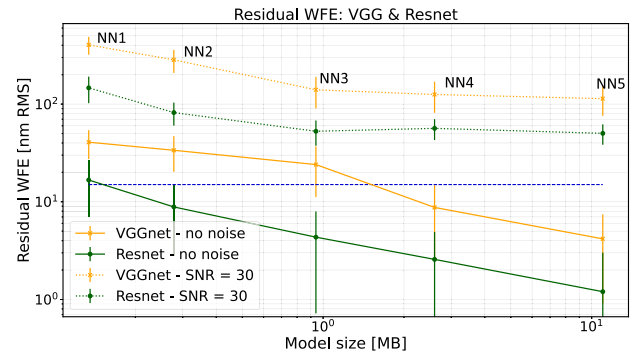

Fig. 6. Performance comparison of estimation between the VGG-net architecture (yellow line) and the Resnet architecture (green line) as a function of the model size. Dotted-line corresponds to model estimation for data at SNR = 30. The dashed blue line represents the residual target to reach the diffraction limit at 15 nm RMS. VGG-net size is rounded at the same value as Resnet for the graphic's readability.

Both NN models are trained with noiseless data and tested with either noiseless data or data at SNR = 30 (the noise study will be developed in Subsection 5.B). As expected, larger models lead to better performance, but at the cost of increased computational complexity. The best performance reaches about 1 nm of WFE, far below the residual target at 15 nm. However, when facing a low SNR, both NN models show weak performance where the best performance achieved is 50 nm of WFE for the heaviest Resnet. Improvement on noisy data is shown in Subsection 5.B. Whatever the model size, the Resnet outperforms the VGG-net. Additionally, Resnet has demonstrated faster convergence and lower loss function values compared to VGG-net. The skip connection in the Resnet architecture enables more effective parameter updates in the top layers and accelerates the convergence of the loss function. Based on the NN performance, the Resnet architecture is preferable. The requirement of 15 nm RMS is fulfilled by nearly all Resnet models regardless of the model size. However, in presence of noisy data, the methodology should be improved to reach the diffraction limit.

Subsequently, only the Resnet architecture will be examined in the rest of the paper. It is important to note that the simulated data used for evaluation lacks noise diversity and higher-order aberrations. Therefore, it is crucial to assess the robustness of the models by introducing noise and HO aberrations into the samples.

B. Model Robustness to Noisy Data

As mentioned above, images suffer from detector and photon noise. For the training step, the noise is added "on-the-fly" meaning that during each epoch, the noise is added to each noiseless PSF so that the noise realization is always different at each epoch for each PSF. Two scenarios are identified.

• Scenario 1: each NN model is trained with a noiseless dataset and is tested with a noisy dataset. Each single NN model is then tested over different SNRs. Several model sizes are tested. The aim is to evaluate the error made on noisy datasets by a model that has been trained with a noiseless dataset.
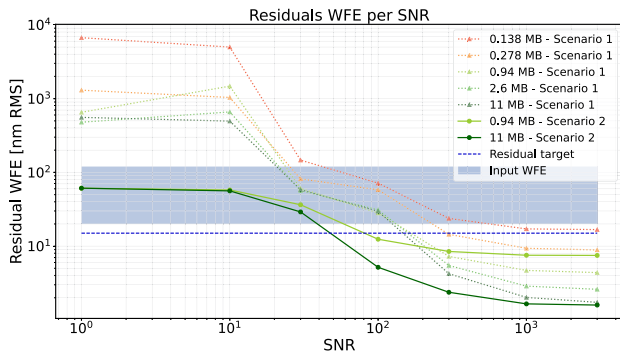
Fig. 7. Resnet robustness to noise: scenario 1 (dotted lines) and scenario 2 (solid lines). For the scenario 2 curves, each point represents one NN model, trained and tested on the corresponding SNR.

Fig. 8. Resnet robustness to SNR. Dark green curve illustrates the estimation quality of the NN5; light green shows NN3 performances.

- Scenario 2: each NN model is trained with noisy samples at a given SNR. Each NN model is then tested on noisy data at the same SNR as for the training. Two NN model sizes are tested. The aim is also to evaluate the error made on noisy datasets when the NN models have been trained with noisy datasets.

We conducted a performance comparison between the two scenarios based on residual WFE after correction. Figure 7 illustrates the residual wavefront error depending on the SNR. The dotted lines represent NN models trained without noise, while the solid lines represent NN models trained with noisy PSFs (each point on the solid line corresponds to a NN model trained for a specific SNR). The performance of the models unsurprisingly improves as their size increases. The diffraction limit, defined by WFE < *Residual target*, is not achieved by the smallest models (dashed red line) even at high SNR. For low SNR, where the detector noise prevails over the photon noise, scenario 1 outputs an estimation error larger than the initial WFE: this would lead to an increase of the WFE if used in a closed loop. On the other hand, the scenario 2 network's estimation stays in the mean of the input WFE distribution: during training, the models minimize their loss function by setting all the phase coefficients to zero as the PSF geometry cannot be properly identified in the noise. At high SNR, in both scenarios, the largest model (NN5) converges towards the same WFE value. However, the NN3 scenario 2 model exhibits a slight limitation, achieving a WFE slightly below 10 nm RMS, whereas the same model trained without noise performs better.

An alternative approach is to train models using a range of SNRs drawn uniformly in the interval [30,100] to assess their performance.

Figure 8 illustrates a comparison of wavefront estimation quality between models trained on a fixed SNR and those trained on varying SNRs. Remarkably, the models trained on various SNRs show great performance, reaching the diffraction limit more rapidly at SNR = 50 for the NN5 and SNR = 70 for the NN3. Surprisingly, even at high SNRs, the models trained on varying SNRs demonstrate comparable performance to the models trained at fixed SNR. However, in a very-low-flux regime, performance is worsening the wavefront. Globally speaking, a SNR of at least 50 is required to reach the diffraction limit with a relatively small NN model. This is somehow similar to phase diversity.
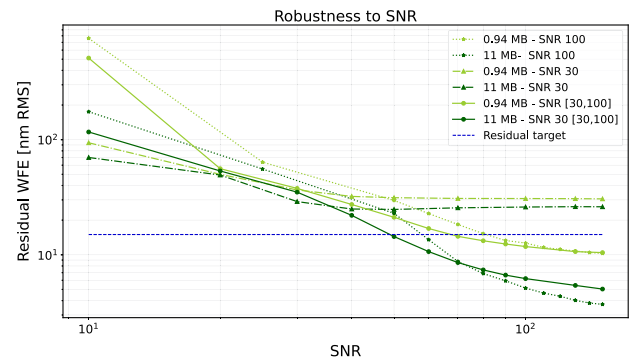
C. Models Robustness to Higher-Order Aberrations

To validate the performance of the neural network, the model is also trained considering the presence of higher-order aberrations on the primary mirror. In the same way as for the noise robustness discussed in Subsection 5.B, we distinguish between two scenarios.

- Scenario 3: each NN model is trained with a dataset where higher-order aberrations are not considered, and is tested on PSFs generated with PTT and higher-order aberrations.
- Scenario 4: each NN model is trained with samples generated from PTT and higher-order aberration phase maps. The model is then tested over data that suffers from the same higher-order amplitude.

For these two scenarios the 12 PTT coefficients follow the same distribution. They are the only coefficients inferred by the NN model since the HO aberrations cannot be compensated for. The *residual WFE* refers here to the PPT part of the phase (i.e., excluding HO aberrations).

Figure 9 shows that even in the presence of unexpected HO aberrations, the inference step achieves a performance better than the diffraction limit requirement. The difference between scenario 3 and scenario 4 is only of a few nanometers, indicating that all model sizes exhibit robustness against small changes in the focal plane image due to HO aberrations. Surprisingly, the most robust model is for a size of 1MB, where the difference between scenario 3 and scenario 4 is the smallest, approximately 0.3 nm RMS. In contrast, the heaviest model begins to overfit the data: it infers with the best accuracy the raw data but when new data is introduced for testing, the performance slightly degrades by a few nanometers, while still reaching the diffraction limit. The inference accuracy worsens when stronger HO aberrations are added to the phase map, as shown by the red curve Fig. 9, which represents the inference performance when 60 nm RMS of WFE is added on top of PTT. However, the performance continues fulfilling the accuracy requirement and adding 60 nm RMS on the phase map introduces only a few nanometers of estimation errors.

D. Comparison to Classical Minimization Methods

Having explored the robustness of different model sizes in term of estimation accuracy, we now turn our attention to comparing
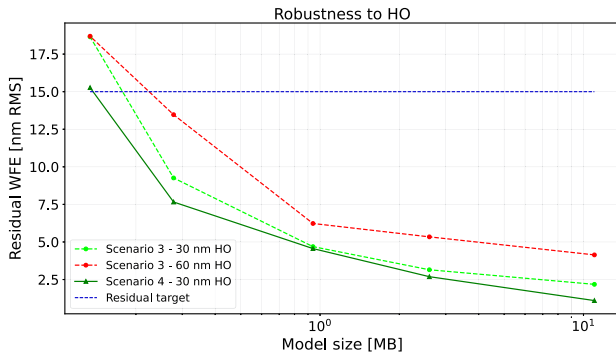
Fig. 9. Resnet performance as a function of model size for images including higher-order aberrations. The dashed light green line corresponds to scenario 3 (i.e., no HO aberrations included in the training) for 30 nm RMS of HO. The red line corresponds to scenario 3 for 60 nm RMS of HO aberrations. Solid green line, scenario 4 results for a NN trained with PSF aberrated with PTT and HO.
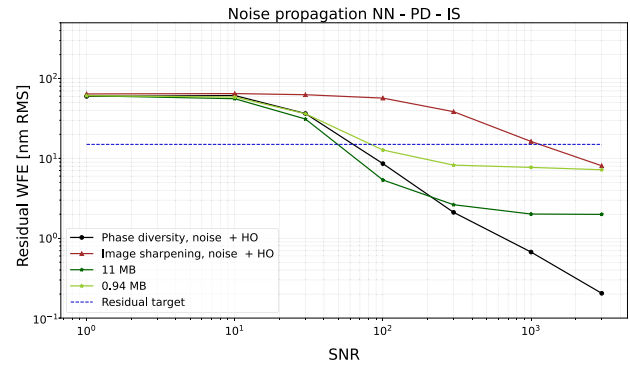

Fig. 10. PTT estimation of the NN (green line), phase diversity method at $\epsilon_3^{PD} = 10^{-15}$ (dark line), and image sharpening (brown line) facing the same dataset of PSF generated with 30 nm RMS of higher-order aberrations at several SNRs regardless of the computing time.

Fig. 11. Estimation time and precision of estimation for NN, phase diversity (PD), and image sharpening (IS). Two types of PD are studied, one fast but less accurate, one slow and more accurate. Test made over the same 1000 samples from the NN testset with HO aberrations for SNR = 100. Algorithm computed on a 2,3 GHz Quad-Core Intel Core i7.

our method with phase diversity and image sharpening (IS) in terms of computation time and performance. Phase diversity relies on a numerical iterative minimization of a quadratic criterion over the image pixels, involving a computing time depending on the number of iterations of the algorithm. Each iteration requires several Fourier transforms as the model PSF is fitted to the measured data (see Subsection 4.B). Three tolerance criteria are chosen: one fast but less accurate $\epsilon_1^{PD} = 10^{-4}$, one finer but slower $\epsilon_2^{PD} = 10^{-7}$, and a very slow and precise $\epsilon_3^{PD} = 10^{-15}$. The principle is also iterative for IS: the metric is maximized until a tolerance criterion is reached; here, a tolerance criterion of $\epsilon^{IS} = 2.10^{-4}$ is chosen. In our case, the metric employed is the maximum pixel intensity of a 2×2 window centered on the image. At each step, PTT corrections are applied (i.e., mirrors are moved), another image is acquired, and the maximum intensity is subsequently calculated.

Data used for the comparison is from the same dataset so that the same data is compared between the three methods.

We aim to compare the three methods in terms of wavefront estimation performance under the presence of noise and higher-order aberrations. NN3 and NN5 are the two neural networks used to compare our three methods, trained on at the corresponding SNR with HO aberrated data. NN coefficients estimation yields better phase estimations for both the NN5 and NN3 across all SNRs levels. Plus these models demonstrate a significant improvement in terms of computational time compared to PD and IS. NN models with their relatively lightweight architecture are much faster, with the majority of computational time occurring during training (which is obviously not taken into account for the inference). Figures 10 and 11 illustrate the performances of both methods relative to their execution time. The $\epsilon_3^{PD} = 10^{-15}$ criterion and $\epsilon_2^{PD} = 10^{-7}$ achieve nearly the same performance at SNR = 100 but with a different computing time. In Fig. 11 only ϵ_2^{PD} is used to provide a fairer comparison between the methods. Compared to the NN inference, which is stable in terms of computing time, PD shows limitations in execution time. First, PD takes longer than NN inferences, and second, the standard deviation of the computing time is larger. In addition, two distinct clusters are identified computing phase diversity with a tolerance criterion
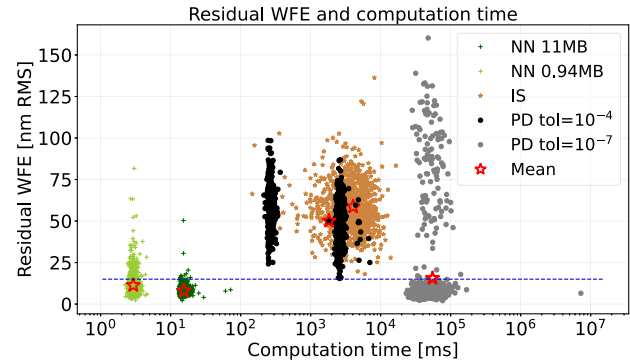
of 10^{-4}, indicating that the number of iterations to reach convergence depends on the criterion shape. Better estimations can be achieved with a convergence criterion of 10^{-7}. However, it requires a longer computing time. Additionally, while PD reaches the diffraction limit requirement on average (red star), it does exhibit approximately 100 outliers out of 1000 samples. While IS is widely spread in terms of computing time and is not reliable, only PD follows NN performance in terms of WFE. However, the estimation time is at least three orders of magnitude slower compared to NN estimation. Therefore, the NN method appears to be more reliable to accurately estimate the wavefront compared to a mono-image phase diversity and image sharpening.

In this section, our primary focus has been examining NN methods and comparing them to two alternative approaches. Both phase diversity and image sharpening can naturally be improved upon. PD methods, for instance, could potentially benefit from linearized analytic phase diversity [31] to improve computation time and load. For IS, one could potentially investigate different image sharpness metrics that provide better

performance, but also look at faster optimization algorithms correcting one mode at a time [32].

6. TOWARDS A COMPLETE PHASING SEQUENCE

Once launched, the telescope will deploy the four petals with a mechanical accuracy of a few micrometers [5]. The initial stage of the phasing strategy aims to mitigate the significant piston and tip-tilt errors introduced during deployment. After initial segment identification and phasing procedure recurring, for instance, to the ELASTIC method [2,23], the expected WFE reaches a sub-λ amplitude over the entire pupil. We have not managed to train a suitable single NN able to retrieve high performance on large aberrations. However, we propose a solution with two NNs in cascade (a first NN for the initial large tip-tilt sensing ignoring the piston, and a second NN for the piston and the remaining PTT). The AZIMOV phasing strategy relies on two main steps. First, after deployment as the TT errors are large (as well as the piston errors), the spots do not superimpose and the piston errors do not affect (or only marginally) the resulting image. The primary mirror needs to be first corrected in tip-tilt in order to recenter each of the four PSFs generated by the individual segments. This allows the four PSFs to be superimposed on top of each other, resulting in a single PSF in the focal plane image. After this superposition, the piston error can be estimated. Finally, the remaining piston and residual tip-tilt errors can be estimated using the NN presented in Section 5. Our correction can be applied iteratively in closed-loop at SNR = 100, to process real-time wavefront estimation. The proposed phasing strategy can be found Fig. 12.

The first iteration corrects for the coarse tip-tilt errors. Figure 13(b) shows this while the piston errors remain unchanged. Then after the second iteration, the NN responsible for fine PTT phasing corrects for the remaining aberrations eventually converging towards about a nanometer of piston and residual tip-tilt resulting in <10 nm WFE over the full pupil. The diffraction limit is reached after two or three iterations for the 11 MB model, and slightly after four or five iterations for the 0.94 MB model, Fig. 13(a). The presence of noise at SNR = 100 increases the standard deviation of the estimation error. However, the estimation remains well under the requirement and shows a great stability during 100 iterations for both models. In addition, it is important to notice that we limited the tip-tilt amplitude due to the amount expected in the deployment, yet larger tip-tilt could be easily added and learned by our method as long as the PSFs remain in the field of view of the detector.
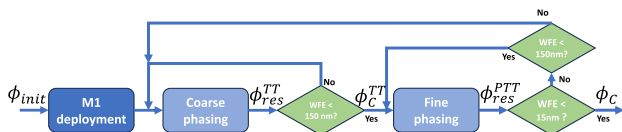

Fig. 12. Diagram of the AZIMOV phasing strategy, where ϕ_{init} is the initial aberrated phase, ϕ_{res}^{TT} the residual tip-tilt phase with uncorrected piston, ϕ_C^{TT} the phase with pre-corrected tip-tilt, ϕ_{res}^{PTT} the residual phase after PTT correction, and ϕ_C the phase corresponding to the diffraction limit.
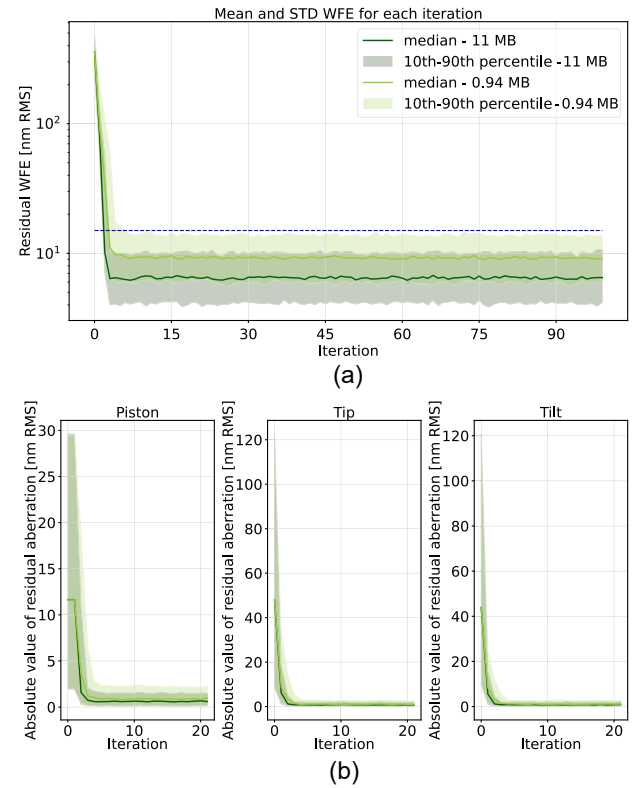

Fig. 13. Closed-loop residuals for AZIMOV deployment regime; mean and STD results over the same 1000 PSFs of the testset at SNR = 100. Both figures contain the same legend. (a) Average residual WFE for 100 iterations in a closed-loop over 1000 PSFs of the testset. The green shaded area represents the bounds of the 10th and 90th percentiles. (b) Average PTT residuals for 15 iterations in a closed-loop. The green shaded area represents the bounds of the 10th and 90th percentiles. Iteration 0 corresponds to the initial WFE.

7. CONCLUSION

This study explored the performance of neural networks for phasing a four-segment deployable space telescope with piston, tip, and tilt aberrations in the range of (20 − 120 nm RMS). The existing state-of-the-art algorithms are deemed too complex because of their iterative nature and large computational burden for space applications. Alternatively, simplified NN architectures are explored, which exhibit promising results for accurately estimating the piston-tip-tilt phasing errors. A five-convolutional-layer-based Resnet has been implemented for demonstrating robustness even in the presence of additional high-order aberrations and noise, by reaching the diffraction limit in different cases at SNR = 50 and for 30 nm RMS and 60 nm RMS of HO aberrations. Additionally, it outperforms concurrent methods such as mono-image phase diversity and image sharpening in terms of computation time by at least a factor of 100 with comparable noise propagation performance. Our implementation shows great performance and stability once applied in a closed loop, for both coarse and fine phasing, using a relatively simple architecture allowing for fast inference times of about a few milliseconds per step.

However, using a single in-focus image for FPWF sensing appears to limit the piston capture range between $-\frac{\lambda}{4}$ and $\frac{\lambda}{4}$. Nevertheless, this limitation might be overcome by using several

input images. For instance, capturing two images at different wavelengths or alternatively using an in-focus and an out-of-focus image could overcome the lambda ambiguity and extend the capture range. Ultimately, the NN approach provides a more generic solution, both in terms of aberration capture range and potentially in terms of scene (point source versus extended).

In our case of simulated data, NNs are driven by the image formation model. Once on-sky data will be available, fine-tuning our method can enhance the performance to fit better to realistic cases.

Finally, our work presents results employing a point source. Finding such a source may be possible during the initial calibration stage. However, in the context of Earth observation, obtaining such a point source may be difficult in practice as objects of interest are typically more complex and extended. Further work will be directed towards the study of extended objects in regions of adequate contrast to convey sufficient phase information.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

REFERENCES

1. P. A. Sabelhaus and J. E. Decker, "An overview of the James Webb space telescope (JWST) project," Proc. SPIE **5487**, 550–563 (2004).
2. N. Schwartz, W. Brzozowski, M. Milanova, *et al.*, "High-resolution deployable CubeSat prototype," Proc. SPIE **11443**, 1144331 (2020).
3. M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (Elsevier, 2013).
4. N. Schwartz, W. Brzozowski, Z. Ali, *et al.*, "6U CubeSat deployable telescope for optical earth observation and astronomical optical imaging," Proc. SPIE **12180**, 1073–1084 (2022).
5. J.-F. Sauvage, N. Schwartz, S. Vievard, *et al.*, "First error budget of a deployable CubeSat telescope," Proc. SPIE **11443**, 1144330 (2020).
6. G. A. Chanan, J. E. Nelson, and T. S. Mast, "Segment alignment for the Keck telescope primary mirror," Proc. SPIE **628**, 466–471 (1986).
7. S. Y. Haffert, L. M. Close, A. D. Hedglen, *et al.*, "Phasing the Giant Magellan Telescope with the holographic dispersed fringe sensor," J. Astron. Telesc. Instrum. Syst. **8**, 021513 (2022).
8. D. S. Acton, S. Knight, M. Carrasquilla, *et al.*, "Phasing the Webb telescope," Proc. SPIE **12180**, 121800U (2022).
9. M. D. Perrin, D. S. Acton, C.-P. Lajoie, *et al.*, "Preparing for JWST wavefront sensing and control operations," Proc. SPIE **9904**, 142–160 (2016).
10. M. P. Lamb, C. Correia, J.-F. Sauvage, *et al.*, "Quantifying telescope phase discontinuities external to adaptive optics systems by use of phase diversity and focal plane sharpening," J. Astron. Telesc. Instrum. Syst. **3**, 039001 (2017).
11. L. M. Mugnier, A. Blanc, and J. Idier, *Advances in Imaging & Electron Physics* (2006), Vol. **141**, p. 1.
12. F. Martinache, "The asymmetric pupil Fourier wavefront sensor," Publ. Astron. Soc. Pac. **125**, 422–430 (2013).
13. F. Rossi, C. Plantet, A.-L. Cheffot, *et al.*, "Machine learning techniques for piston sensing," Proc. SPIE **12185**, 1669–1674 (2022).
14. J. Herbel, T. Kacprzak, A. Amara, *et al.*, "Fast point spread function modeling with deep learning," J. Cosmol. Astropart. Phys. **2018**, 054 (2018).
15. S. W. Paine and J. R. Fienup, "Machine learning for improved image-based wavefront sensing," Opt. Lett. **43**, 1235–1238 (2018).
16. Y. Wang, F. Jiang, G. Ju, *et al.*, "Deep learning wavefront sensing for fine phasing of segmented mirrors," Opt. Express **29**, 25960–25978 (2021).
17. H. F. Rajaoberison, J. S. Tang, and J. R. Fienup, "Machine learning wavefront sensing for the James Webb space telescope," Proc. SPIE **12180**, 2210–2216 (2022).
18. T. Andersen, M. Owner-Petersen, and A. Enmark, "Image-based wavefront sensing for astronomy using neural networks," J. Astron. Telesc. Instrum. Syst. **6**, 034002 (2020).
19. G. Orban De Xivry, M. Quesnel, P. Vanberg, *et al.*, "Focal plane wavefront sensing using machine learning: performance of convolutional neural networks compared to fundamental limits," Mon. Not. R. Astron. Soc. **505**, 5702–5713 (2021).
20. Y. Nishizaki, M. Valdivia, R. Horisaki, *et al.*, "Deep learning wavefront sensing," Opt. Express **27**, 240–251 (2019).
21. M. Quesnel, G. Orban de Xivry, G. Louppe, *et al.*, "A deep learning approach for focal-plane wavefront sensing using vortex phase diversity," arXiv, arXiv:2210.00632 (2022).
22. B. Pope, N. Cvetojevic, A. Cheetham, *et al.*, "A demonstration of wavefront sensing and mirror phasing from the image domain," Mon. Not. R. Astron. Soc. **440**, 125–133 (2014).
23. S. Vievard, F. Cassaing, and L. Mugnier, "Large amplitude tip/tilt estimation by geometric diversity for multiple-aperture telescopes," J. Opt. Soc. Am. A **34**, 1272–1284 (2017).
24. K. Dohlen, F. Wildi, J. Beuzit, *et al.*, "SPHERE: confronting in-lab performance with system analysis predictions," in *Second International Conference on Adaptive Optics for Extremely Large Telescopes* (LESIA, 2011).
25. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT, 2016).
26. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv, arXiv:1409.1556 (2014).
27. K. He, X. Zhang, S. Ren, *et al.*, "Deep residual learning for image recognition," in *IEEE Conference on Computer Vision and Pattern Recognition* (2016), pp. 770–778.
28. A. Blanc, "Identification de réponse impulsionnelle et restauration d'images: apport de la diversité de phase," Ph.D. thesis (Paris 11, 2002).
29. B. T. Polyak, "The conjugate gradient method in extremal problems," USSR Comput. Math. Math. Phys. **9**, 94–112 (1969).
30. M. J. Powell, *A Direct Search Optimization Method That Models the Objective and Constraint Functions by Linear Interpolation* (Springer, 1994).
31. I. Mocoeur, "Analyse de front d'onde en plan focal: développement d'algorithmes temps-réel et application au cophasage de télescopes multipupilles imageurs," Ph.D. thesis (Université Paris Sud-Paris XI, 2008).
32. D. Débarre, E. J. Botcherby, M. J. Booth, *et al.*, "Adaptive optics for structured illumination microscopy," Opt. Express **16**, 9290–9305 (2008).