

# Neural networks and PCA coefficients to identify and correct aberrations in adaptive optics

A. Terreri<sup>1,2,3,6</sup> , F. Pedichini<sup>1,6</sup>, D. Del Moro<sup>3</sup>, G. Li Causi<sup>5,6</sup>, M. Stangalini<sup>4,6</sup>, M. Mattioli<sup>1</sup>, and R. Piazzesi<sup>1,6</sup>

<sup>1</sup> Osservatorio Astronomico di Roma, INAF, Via di Frascati 33, Monte Porzio Catone (RM) 00078, Italy  
e-mail: [alessandroterreri88@gmail.com](mailto:alessandroterreri88@gmail.com)

<sup>2</sup> Università degli Studi di Roma “La Sapienza”, Physics Department, Piazzale Aldo Moro 2, Rome 00185, Italy

<sup>3</sup> Università degli studi di Roma Tor Vergata, Via della ricerca scientifica 1, Rome 00133, Italy

<sup>4</sup> Agenzia Spaziale Italiana (ASI), Via del politecnico, Rome 00133, Italy

<sup>5</sup> Istituto di Astrofisica e Planetologia Spaziale (IAPS), Via del fosso del cavaliere 100, Rome 00133, Italy

<sup>6</sup> ADONI – ADaptive Optics National lab in Italy, Italy

Received 10 December 2021 / Accepted 9 May 2022

## ABSTRACT

**Context.** Static and quasi-static aberrations represent a great limit for high-contrast imaging in large telescopes. Among them the most important ones are all the aberrations not corrected by the adaptive optics (AO) system, which are called non-common path aberrations (NCPA). Several techniques have been proposed to mitigate it. The typical approach is to set an offset on the AO system with exactly the opposite sign of the NCPA in order to correct for the aberrations introduced by all the optical components downstream the wave-front sensor (WFS) up to the science camera. An estimate of the NCPA can be obtained with a trial-and-error approach or by more sophisticated techniques of focal-plane wave-front sensing.

**Aims.** In all cases, a fast procedure is desirable to limit the telescope downtime and to repeat, if needed, the correction procedure to cope with the temporal variation of the NCPA. Very recently, new approaches based on neural networks (NNs) have also been proposed as an alternative.

**Methods.** In this work, through simulated images, we test the application of a supervised NN for the mitigation of NCPAs in high-contrast imaging at visible wavelengths and, in particular, we investigate the possibility of applying this method to fast imagers such as SHARK-VIS, the forthcoming visible-band high-contrast imager for the Large Binocular Telescope (LBT).

**Results.** Preliminary results show a measurement accuracy of the NCPA of 2 nm root mean square (RMS) for each sensed Zernike mode in turbulence-free conditions, and 5 nm RMS per mode when the residual turbulence has a wave-front error (WFE) of approximately 42.5 nm RMS, a typical value during LBT AO system calibration. This measurement is sufficient to guarantee that, after correction, NCPA residuals in the system are negligible compared to the typical WFE > 100 nm RMS of the best AO systems at large telescopes.

**Conclusions.** Our simulations show this method is robust even in the presence of turbulence-induced aberrations that are not labelled in the training phase of the NN. The method could thus be used in a real-world setting by offloading a corrective static offset to the AO system of a telescope to mitigate the NCPA.

**Key words.** instrumentation: adaptive optics – methods: data analysis – techniques: image processing – techniques: miscellaneous – telescopes – turbulence

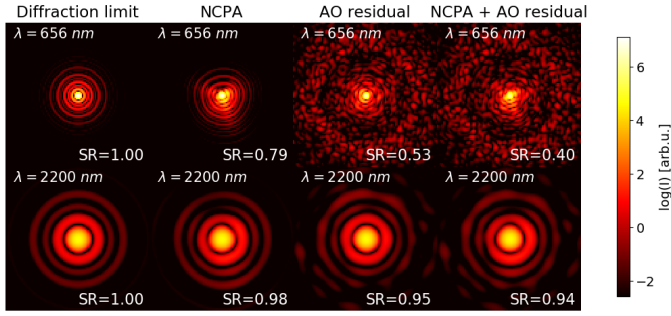
## 1. Introduction

Astronomical high-contrast imaging demands accurate wave-front control to minimise residual atmospheric aberrations that can hamper the detection of very faint objects, for example for extrasolar planet imaging (e.g. Mesa et al. 2019; Lagrange et al. 2009; Marois et al. 2008). A significant limitation to the achievement of the theoretical contrast limit is represented by static or quasi-static aberrations that can result from, for example, optical imperfections, misalignments, or thermal effects. This is the case for both space- (e.g. Hubble Space Telescope Gonsalves & Nisenson 1991) and ground-based observations. In the latter case, particularly with regard to adaptive optics (AO) systems, imaging instruments operated at ground-based telescopes are affected by wave-front errors (WFEs) that result from different contributions: residual atmospheric turbulence not corrected by the AO, telescope vibrations, gravity- or temperature-induced deformations of the optics, and the performance of the optical

elements themselves. Image aberrations originating downstream of the WFS are not sensed by the wave-front sensor (WFS) itself, and therefore are not corrected by the AO system, thus contributing to the increase of the overall WFE, as noted in Esposito et al. (2020).

In this regard, wave-front sensing performed by the science camera would be ideal to effectively mitigate these residual aberrations (Jedrzejewski et al. 1993), which are generally referred to as non-common path aberrations (NCPA).

Achieving a robust focal plane or close-to-focal plane wave-front sensing, based on the analysis of the point spread function (PSF), would also offer a series of advantages in different fields of astronomical instrumentation, including post-facto image deconvolution and analysis. In general, it would also be useful whenever a mitigation of the residual WFE is desired but a dedicated WFS or a secondary focal plane are not available, leading to analysis of extrafocal plane image momentum, as in Terreri et al. (2019b).



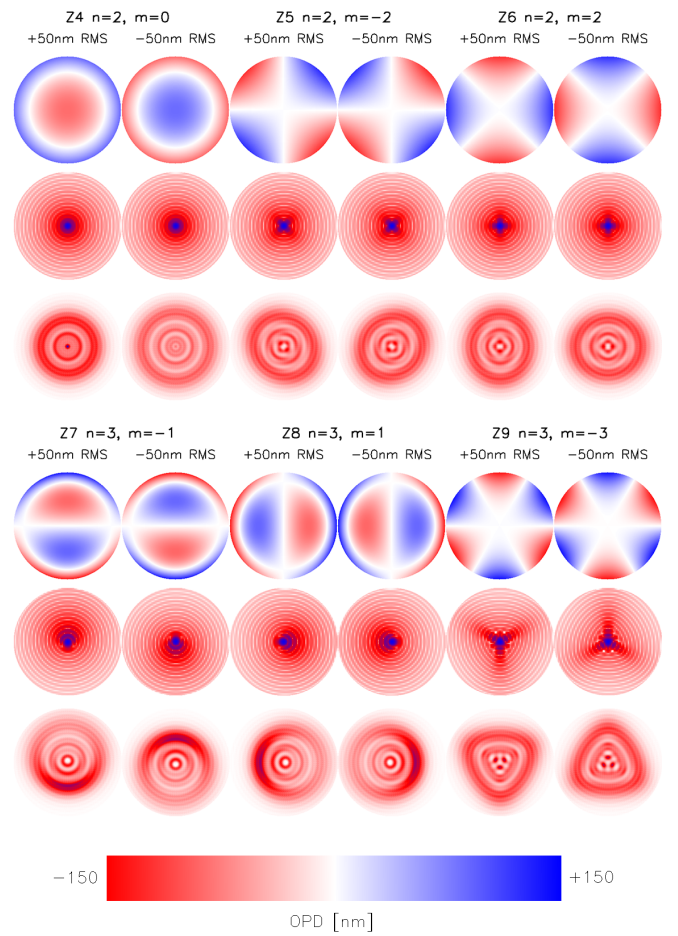
**Fig. 1.** Focal-plane images simulated at two different wavelengths:  $H_\alpha = 656$  nm and  $K$  band = 2200 nm. *From left to right:* diffraction limited PSF; PSF with NCPAs of  $WFE_{CPA} \approx 50$  nm RMS; PSFs with AO residual  $WFE_{AO} \approx 100$  nm RMS; PSFs with AO residuals and NCPAs. The Strehl ratio (SR) is calculated using Szapiel (1983); Ross (2009). The NCPA effect is clearly stronger at the shorter  $H_\alpha$  wavelength.

Even if smaller than the residual amplitude of the seeing aberrations, NCPAs can still represent a severe limitation (see Fig. 1), in particular in applications such as high-contrast imaging of circumstellar discs and exoplanets where extreme wave-front control is required (Potier et al. 2021).

Preliminary on-sky tests (Esposito et al. 2012) conducted on the First Light Adaptive Optics (FLAO) instrument at the Large Binocular Telescope (LBT, Arizona USA; Hill 2010) have shown an overall residual WFE ranging from an 80 nm root mean square (RMS) for a star of  $R_{Mag} = 7.5$  to a 218 nm root mean square RMS for a star of  $R_{Mag} = 14.5$ , as also reported in Pinna et al. (2015, 2016); Agapito & Pinna (2019) and provided by the Strehl Calculator<sup>1</sup>. According to Li Causi et al. (2017); Causi et al. (2018); Hope et al. (2019); Stangalini et al. (2018); Mattioli et al. (2019), these results limit the contrast achievable at visible wavelengths to about  $5 \times 10^{-5}$  at 100 mas separation from the central star under typical observing conditions (seeing ranging from 0.8'' to 1.2'').

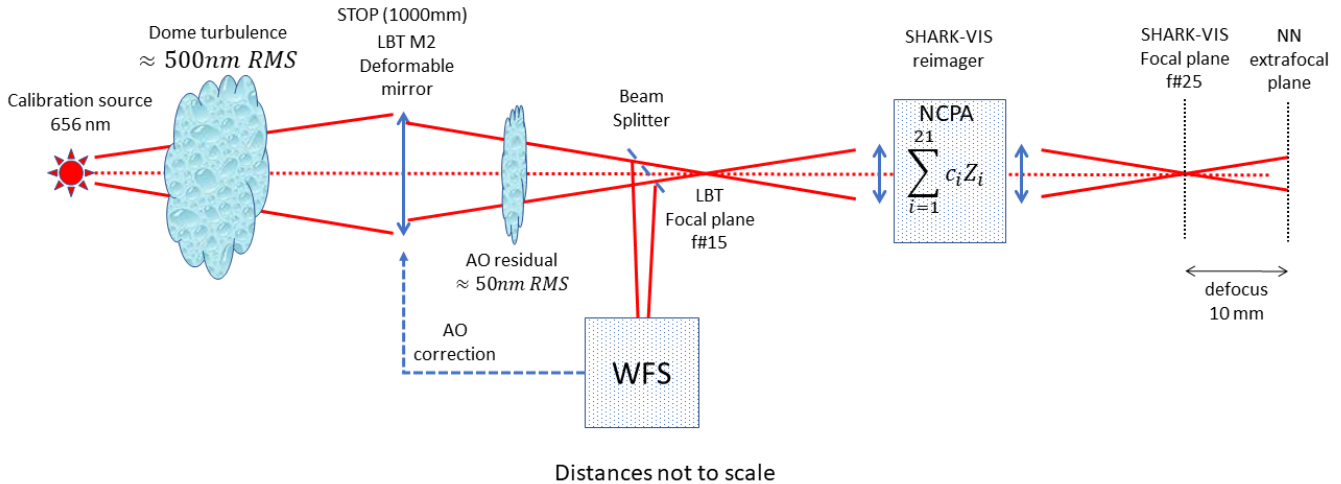
NCPAs contribute to the overall WFE budget and show typical sub-Hz frequencies and amplitudes in the range of a few tens of nanometers RMS, as in Pedichini et al. (2017); Esposito et al. (2020). NCPAs usually affect the low orders of the Zernike components of the wave-front (in this article Zernike components are indexed according to the Noll convention Noll 1976), and the effects are stronger as the wavelength becomes shorter as confirmed by experimental and theoretical results reported in Vassallo et al. (2018); Esposito et al. (2020); Szapiel (1983); Ross (2009). In the case of point-like sources, in the last decade new wave-front sensing technologies (Guyon et al. 2006; Korkiakoski et al. 2014) have been developed based on the analysis of images taken directly from the science camera: these are referred to as focal-plane wave-front sensors (FPWFS). A few examples are phase diversity (PD; Gonsalves 1982), curvature WFSs (Roddiier 1988), the Roddiier Test (Roddiier et al. 1988), the Iterative Fourier Transform method (Roddiier & Roddiier 1991), the Photonic Lantern WFS (Norris et al. 2020), and approaches based on machine learning (Orban de Xivry et al. 2021) (Menduiña Fernández 2021). These methods are useful to measure NCPAs and mitigate their effect. In particular, on AO-assisted telescopes this is done through the addition of a specific static offset to the AO real-time correction, which corresponds to the opposite value of the measured NCPA.

<sup>1</sup> <http://adopt.arcetri.inaf.it/strehl.html>



**Fig. 2.** Detailed view of the degeneracy problem for even radial orders ( $n$ ) at focal plane. *Top panel:* aberrations induced by even radial-order Zernike with  $n = 2$  ( $Z_4, Z_5, Z_6$ ). *Bottom panel:* aberrations induced by odd radial-order Zernike with  $n = 3$  ( $Z_7, Z_8, Z_9$ ). *Top rows:* wave-front OPD at pupil. *Centre rows:* focal-plane images of aberrated PSFs. *Bottom rows:* extra-focal-plane images, obtained with 10 mm of defocus on an  $f/25$  beam. The removal of the degeneracy is evident in each extra focal image (the WFE amplitude was always 50 nm RMS).

However, one of the principal drawbacks of this approach is the degeneracy at the focal plane of the PSF with respect to the sign of some aberration components, that is the fact that for some aberrations, opposite sign amplitudes produce the same PSF at the focal plane; hence, there is a difficulty in disentangling the contribution of these aberration terms. This effect can be seen in Fig. 2, where pupil plane wave fronts and their corresponding PSFs are compared: for even Zernike orders  $n = 2, 4, 6, \dots$ , a positive (+50 nm) or negative (−50 nm) Optical Path Difference (OPD) produces exactly the same PSF, while for odd Zernike orders  $n = 3, 5, 7, \dots$  the PSFs are different. In the extra-focal plane, however, this degeneracy is removed, so the use of extra-focal plane images allows the retrieval of information even for even radial-order Zernike components. In order to solve this ambiguity, most methods exploit images from intra- and extra-focal planes, such as the Roddiier test (Roddiier & Roddiier 1993), which uses both, or the Phase Diversity (PD) technique described in Gonsalves (1982), which uses one image taken in-focus and another taken with a known amount of defocus, as done by SPHERE at the VLT by Sauvage et al. (2011). It is, however, also possible to solve this problem with a single PSF image



**Fig. 3.** Conceptual optical layout of SHARK-VIS at LBT for closed-dome calibration operations.

taken at a defocused plane by using its morphology and intensity distribution; an example of this approach is the Wave-front Reconstruction Software (WRS), which is currently used at the Large Binocular Camera (LBC) at the LBT to preset the shape of the active primary mirror and was used to identify the first low Zernike order static aberration values in Terreri et al. (2019b).

Building on this knowledge base, we introduce the idea of using an Artificial Intelligence (AI) approach to FPWFS, using Principal Component Analysis (PCA) encoded out-of-focus images as input for a Neural Network (NN) to retrieve aberrations in optical systems (Hotelling 1933, Terreri et al. 2019a). Neural Networks are extremely efficient in mapping non-linear relationships between input and output in regression problems and they have almost real-time results once the model is trained. The present work used a Multi-Layer Perceptron (MLP) NN to analyse defocused images and measure their wave-front aberrations. All images reported in this work are simulations generated following an optical prescription corresponding to the optical design of SHARK-VIS (Fig. 3), the new forthcoming high-contrast and high-resolution visible-band imager for the LBT, as described in Pedichini et al. (2016) and Mattioli et al. (2018).

We generated two datasets of simulated images in different conditions and, for each one of them, an NN model was trained to measure the first 21 Zernike polynomials. In the first dataset, only aberrations encoded in the first 21 Zernike polynomials were injected, while the second set also contained high-order aberrations typical of the LBT AO residual turbulence that occurs during closed dome calibrations (Esposito et al. 2012). The range of this is from 40 to 50 nm RMS. The latter situation closely represents the conditions of an instrument such as SHARK-VIS at the LBT, when calibration procedures intended to minimise NCPAs are executed by using the telescope diffraction-limited source of ARGOS, as in Schwab et al. (2010).

## 2. Methods

Our wave-front analysis procedure takes extra-focal-plane images of optical systems, generated with physical optics propagation, and applies a two-stage process to retrieve the Zernike expansion (Noll 1976) of our original wave-front. First, a PCA

(Hotelling 1933) of the image is performed, and then the resulting PCA coefficients are analysed by a properly trained NN using the python code in the scikit-learn (Ross et al. 2008) library package.

Neural Networks (Sect. 2.3) are algorithms designed to recognise patterns in data by establishing correlations between input and desired output data. As described in Van Der Malsburg (1986), a simple toy model can be graphically described by a net of lines connecting a set of nodes called ‘neurons’, representing input numbers, with other nodes representing output, possibly including auxiliary nodes in-between. These neurons are usually structured in three different blocks called ‘layers’: the input layer, the hidden layer(s), and the output layer. The net works by propagating the node values from the input layer to the output layer, processing the values at each node and propagating the result to the next. At each node, the separate input values are each multiplied by a weight and summed. A bias value is then added to the sum. The resulting value is then fed as an argument to an activation function; the output value of the function is the value for that neuron and is fed to the next layer. Activation functions can be of different types (Sigmoid, ReLU, atan,...) and they are the element that makes the whole process non-linear. In our case, the activation function is a fixed ReLU for all neurons and the biases are set to 0. The values in the output layer are thus dependant on the weights. Before the NN can be used, the weights must first be optimised: this is the ‘learning’ process of the net and is called ‘training’ the NN. In this phase, data consisting of known corresponding inputs and outputs are fed to the NN, respectively to the input and output layers, and used to iteratively update connection weights. After the training process is completed, it is possible to apply the NN to new and unknown datasets in a ‘testing’ phase. In this phase (which can be also called the ‘working’ or ‘prediction’ phase), data are fed to the input layer and the NN outputs the result(s).

In our case, the NN is trained to output a measure of the wave-front aberrations from input-simulated images whose aberration is known. Once the training is completed, a different set of images with different known aberrations (the testing dataset) is fed to the NN; the output is then compared with the known input and the accuracy of the prediction is measured. For our purpose, we generated different sets of defocused PSF images for the training and testing phases (Sect. 2.1) simulating different Zernike aberrations by means of an optical propagation software.

**Table 1.** Settings used for the optical simulations and for the development of the NN.

Parameter	Value
F#	25
Defocus [mm]	10
Simulation grid points	$1024 \times 1024$
Detector pixel size [ $\mu\text{m}$ ]	6.5
Entrance Pupil Diameter [mm]	8200
Wave-front sampling [mm]	18.6
Final image size [pixel]	$137 \times 137$
Simulation wavelength [nm]	656
$\lambda \backslash D$ [mas]	16.5
# pixel corresponding to $\lambda \backslash D$	2.52
# sample training	8000
# sample validation	2000
# sample testing	100
# PCA components	500
# hidden layers	1
# hidden layers neurons	10 000
range $c_i$ [nm RMS]	$\pm 50$
# Zernike simulated	21
# Zernike used for NN	18

We then applied a PCA (Sect. 2.2) to the simulated defocused PSF images, which encodes each image into a vector of a few hundred coefficients. These coefficients are used as values for the input layer nodes, and the values of the output layer correspond to the Zernike coefficients for all the simulated PSFs. In practice, the NN maps PCA coefficients to Zernike coefficients, allowing us to directly measure the wave-front aberrations from a defocused image of the PSF. In Sect. 2.4 we describe how atmospheric turbulence residuals were simulated and then introduced together with Zernike coefficients.

### 2.1. Simulations

All the simulations used in this work were made using the `proper_lib` python optical propagation library developed by Krist (2007). Using this library, it is possible to inject known aberrations expressed in terms of Zernike polynomials as

$$W = W_0 + \sum_i^N c_i Z_i \quad (1)$$

where  $W$  is the final wave front that generates our images,  $W_0$  is the original wave front (with no NCPA), and  $\sum_i^N c_i Z_i$  is the NCPA aberration, with  $c_i$  being the amplitude RMS of the  $i$ -th Zernike polynomial, using the Noll index (Noll 1976) on the index  $i$  to identify the Zernike order.

For our simulations, we generated several series of random  $c_i$  with a uniform distribution in the  $\pm 50$  nm RMS range, with one series for each image of each set. These were organised into 2D-arrays ('Zernike dataset'), with row indices corresponding to the image number and column indices corresponding to the  $i$ -th Zernike polynomial. Hereafter, all the measures regarding the WFE are to be considered as RMS.

Running the optical propagation we generated an image for each coefficient set and created a cube of images ('Image Dataset'). Optical simulation parameters were chosen in order to reproduce the working conditions of SHARK-VIS. Table 1 summarises the parameters used to generate the image datasets

for each of the three stages of NN learning and application (i.e. training, validation, and testing; see Sect. 2.3).

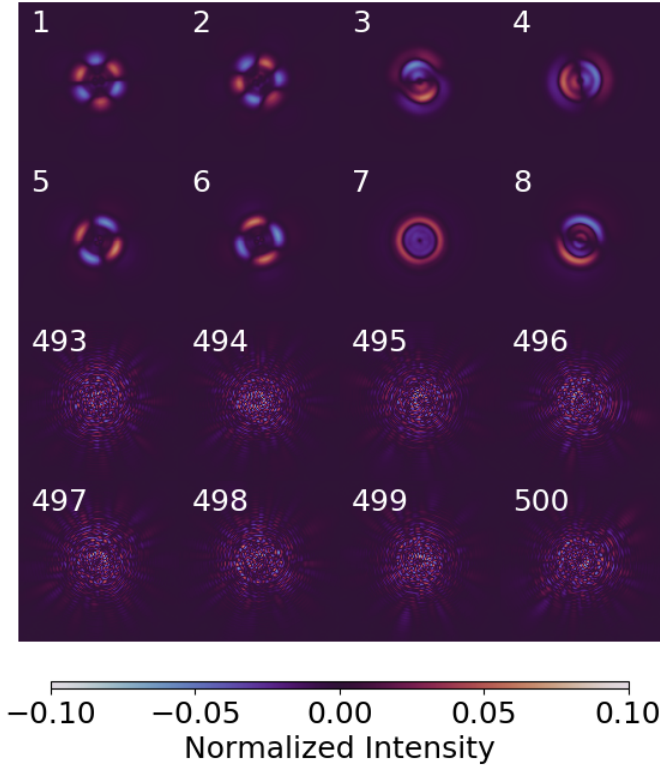
Each set of Zernike coefficients is thus mapped to each image of the 'image dataset'. Out-of-focus images were initially generated on a  $1024 \times 1024$  grid to guarantee super-Nyquist sampling of high-order components and avoid aliasing; these were then re-sampled onto a smaller  $147 \times 147$  grid to reproduce the real sampling of the  $6.5 \mu\text{m}$  pixels of the SHARK-VIS science camera detector (Andor Zyla as described in Andor 2018). Each image was then registered to make the array centre coincide with the position of its photo-centre and cropped to a dimension of  $137 \times 137$  pixels to remove tip and tilt ( $Z_2$  and  $Z_3$ ).

The expected NCPA typical of an imaging instrument such as SHARK-VIS, resulting from Pedichini et al. (2017), is of the order of a few tens of nanometers, with the dominant aberrations described by Zernike polynomials from  $Z_4$  to  $Z_{11}$ . To make our simulation robust and realistic, in Eq. (1) we set  $N = \{1 \dots 21\}$  and the  $c_i$  range of each component to  $\pm 50$  nm RMS in the Noll representation. The piston term  $Z_1$  is neglected because it is a constant phase offset on the optical pupil of the system and does not affect the in-focus or out-of-focus PSF. Tip ( $Z_2$ ) and tilt ( $Z_3$ ) do not modify the PSF shapes on the image plane, only its position; their effects are eliminated by the centring during the dataset pre-processing and are not relevant for the NCPA analysis. Consequently, it is the Zernike from 4 to 21 that are processed by the NN. We set the value of the defocusing of the images to 10 mm after the nominal focal plane in order for it to be compatible with the range of the focus stage of the SHARK-VIS science camera, which is between  $-10$  mm and  $+40$  mm. This value was found to work through preliminary tests: as stated in Sect. 5, we plan to further explore the effect of different defocus positions in future studies. The number of image samples for the NN training phase was set to 10 000.

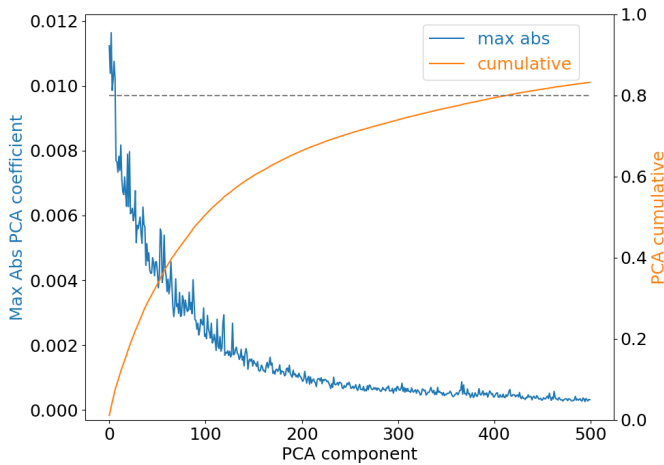
The simulation tool ran on a laptop with an INTEL core i5 3320M processor. The average time for the generation (with or without AO residual turbulence) of one aberrated extra-focal image of  $1024 \times 1024$  pixels is  $\sim 2.5$  s when using  $N = 21$ ; a further  $\sim 0.5$  s are necessary for the post-processing (centring, cropping, and disc storage).

### 2.2. PCA

After its generation, each data set is processed with a PCA that computes a basis of eigen images from an image dataset (Fig. 4). This technique is widely used to reduce data dimensions as in JPEG encoding for image compression, PSF reconstruction for angular differential imaging (ADI) as in Amara & Quanz (2012), and pattern recognition. Every image of the dataset can then be expressed in terms of this basis as a sum of eigen images with a specific set of coefficients. These coefficients are used as input for the NN in place of the actual images, thus requiring only a small number of neurons in the NN input layer. We see that 500 components are enough to keep most of the information present in the image dataset, as can be verified in Fig. 5, where the maximum absolute values of each PCA coefficient over the whole dataset and their cumulative function are plotted, versus the PCA component order. It is evident that after 400 modes there is little added information; the high number of components needed to properly encode our problem reveals the faintness of information present in the NCPA signal we are trying to measure. The datasets are thus encoded on one side by the Zernike coefficients of the injected aberrations and on the other side by the PCA coefficients (Fig. 6). The NN training process creates a robust correlation between these two descriptors.



**Fig. 4.** Example of the first and last 8 eigen images output by the PCA of one extra-focal dataset. The first PCA eigen images encode wide spatial scales in the image, while the last ones contain information on high spatial frequencies.



**Fig. 5.** Contribution of PCA components to image information. Blue: maximum absolute value for each PCA coefficient across all images used to create the basis. Orange: cumulative value of all previous PCA coefficients. The dashed grey line is the 0.8 limit reference reached using more than 400 PCA components.

### 2.3. Artificial neural network and multi-layer perceptron

Artificial Intelligence is a branch of informatics focused on the creation of machines that are able, using specific algorithms, to make decisions autonomously when stimulated by external inputs. Inside this AI family, there are ‘deep-learning’ methods based on artificial NNs. In the biological world, cognitive functions are provided by the presence of ‘neurons’ which are

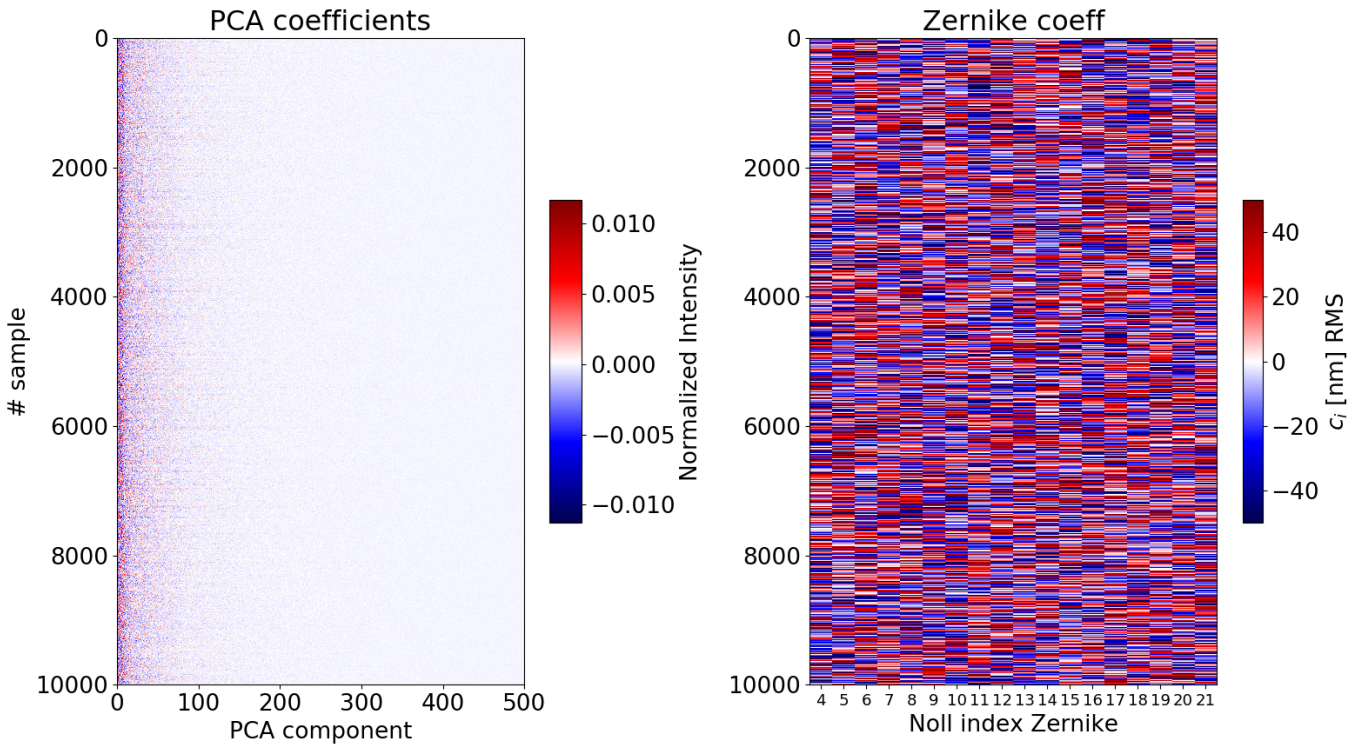
arranged in an interconnected network. The multi-layer perceptron (MLP) is one of the most used NN layouts and is detailed in Van Der Malsburg (1986). This layout (Fig. 7) has a series of artificial neurons organised in layers, where each neuron communicates with all the others placed on the previous and next layers; the first and last layers manage the input and output functions and are called input and output layers, while all the others are referred to as the hidden layers.

In order to tune the connection weights and biases of the neurons, the NN must be trained. In this phase, the input and output are both fed with known linked data: the process is iterative, and at each step, known as an ‘epoch’, the weights and biases are modified, until the RMS error between the NN output and the expected output is below the required precision. There are different optimisation algorithms to train the NN weights. We used ADAM, which is described in Kingma & Ba (2017). Such training processes usually require two phases: the NN is first trained on a large dataset, called a training set; then, after convergence, the NN is applied to a smaller dataset, called a validation set to check if the prediction error is still low on different data than that used for training. If this condition is not met the network parameters are changed until a validated NN model is achieved.

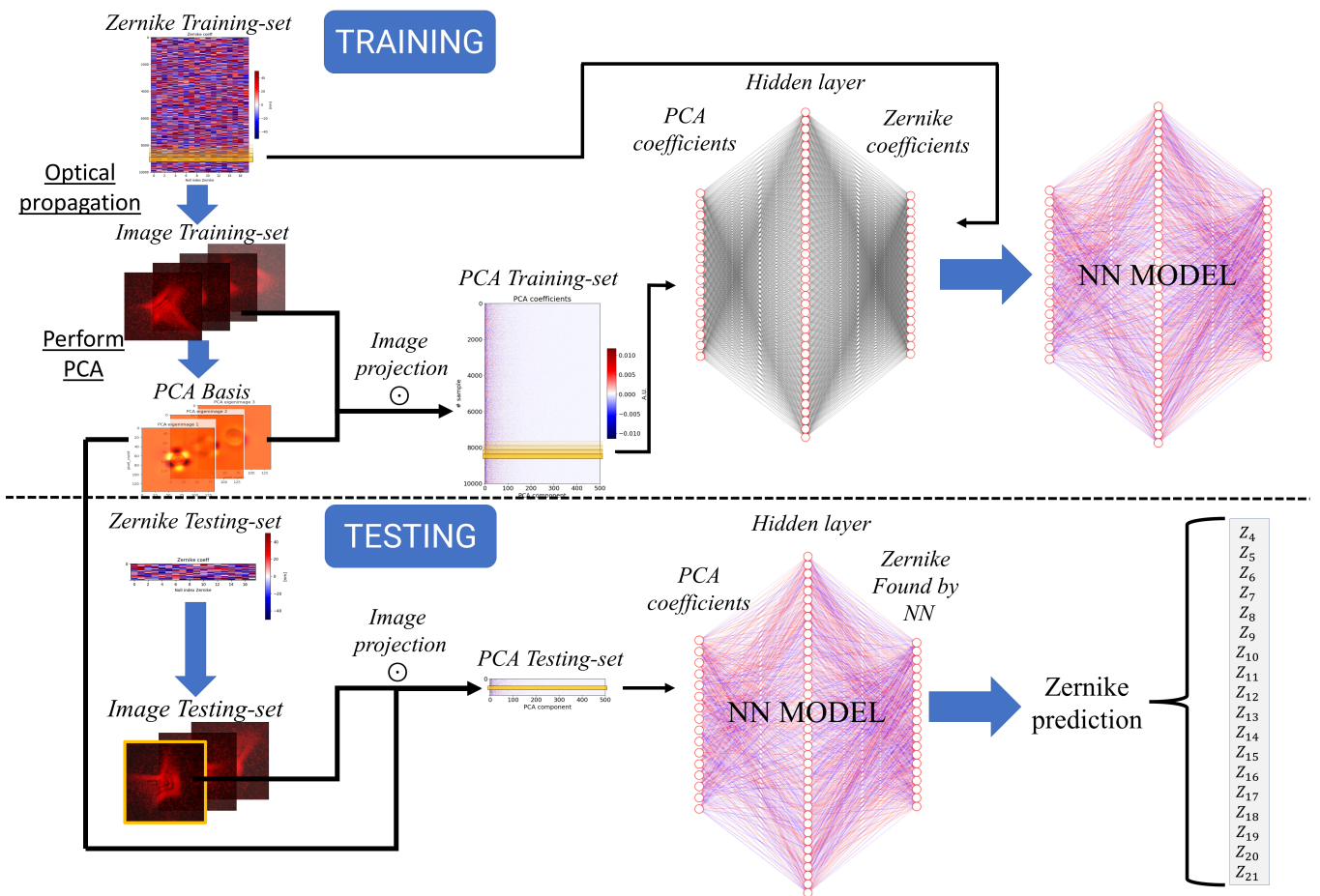
In our case, each set is composed of PCA data and Zernike data. During training, we fed the NN with the PCA training set as input and the Zernike training set as output. At specific regular epoch intervals, a PCA validation set and its corresponding Zernike-validation-set elements are injected into the NN. The validation-set elements, which represent 20% of the entire training dataset, are useful to avoid the NN ending up on local minima. After typically 100 epochs, the accuracy stabilizes just below 4 nm, which is an acceptable level for our kind of application (see Fig. 8). Once a NN training phase has been concluded, it is said that a specific NN model has been generated. In our NN, the input layer has  $N_{\text{input}} = 500$  neurons, corresponding to the number of PCA components, the single hidden layer has  $N_{\text{hidden}} \gg N_{\text{input}}$  and the output layer has 18 neurons corresponding to the number of Zernike components. Our hidden layer is composed of  $N_{\text{hidden}} = 10\,000$  neurons with a ReLU activation function; this number was chosen by trial and error, following the principle expressed in Hornik et al. (1989) whereby a NN with one hidden layer whose number of neurons tends to infinity can approximate any functional relation.

The capture range of the NN is defined by the training process, for instance if we train the NN with Zernike coefficients randomly distributed in the  $\pm 50$  nm range then the NN will be able to recognise aberrations in this range. After the training stage, the testing set is used to test our concept and explore the NN performance on unknown data, which are to those used for training. The testing set is composed of  $N_{\text{test}} = 100$  images. These are decomposed on the same PCA eigen image basis as the training dataset to obtain the PCA testing set. The corresponding Zernike testing set contains the aberrations encoded into each image and the values against which the NN output will be compared to verify its accuracy in working conditions.

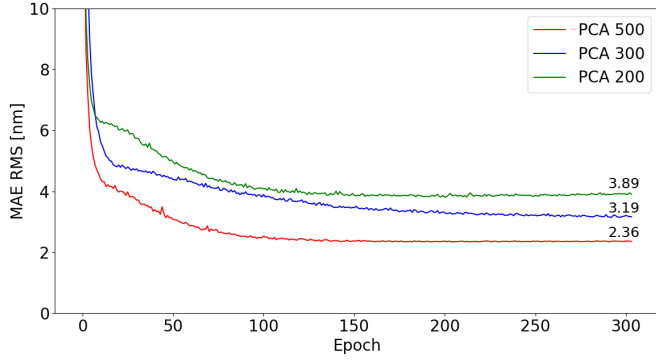
We also tested how fast the NN reached a plateau for different numbers of PCA components [200, 300, 500] and for what mean absolute error (MAE) over all the Zernike components. In Fig. 8, we notice that with 500 components the plateau is reached faster ( $\sim 100$  epochs) and, more importantly, with a lower MAE ( $\sim 2.4$  nm). In the same figure, with 300 components the plateau is reached at 300 epochs and  $\sim 3.2$  nm MAE and with 200 at 150 epochs and  $\sim 3.9$  nm MAE.



**Fig. 6.** *Left* matrix represents the PCA-training-set input of the NN and *right* matrix is the Zernike-training-set output of the NN. Each row of the two matrices represents one of the 10 000 images of the training dataset.



**Fig. 7.** Working flow structure of the NN.



**Fig. 8.** Mean absolute error (MAE) over the Zernike component vs. training epochs of the NN. A larger number of PCA components allows a faster and more accurate training of the NN. The asymptotic values for each PCA curve are reported above the plot final parts.

#### 2.4. AO residual and NCPA

As stated previously, we analysed two scenarios, one where only NCPA aberrations were present and one with both NCPA and AO residuals. AO residuals are high-order aberrations left after the correction of the AO system with a timescale  $\tau$  that corresponds to the atmospheric turbulence timescale and is of the order of a few milliseconds (Stangalini et al. 2017). These residuals combine with other present aberrations (e.g. NCPA), further reducing the SR at the science focal plane. The final wave front can then be expressed as follows:

$$\text{WFE}_{\text{total}} \approx \sqrt{\text{WFE}_{\text{AO}}^2 + \sum_{i=1}^N (c_i Z_i)^2}. \quad (2)$$

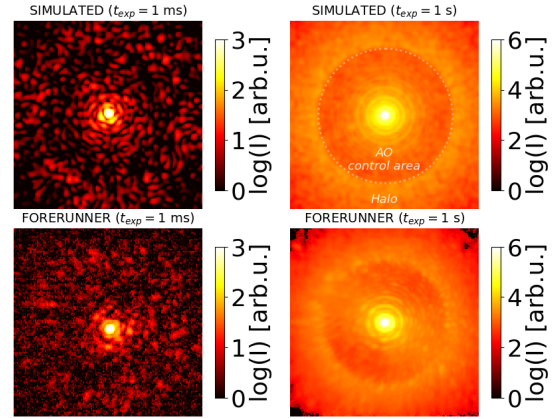
In order to create a NN model capable of identifying quasi-static aberrations within images affected by turbulence, i.e. AO residuals, we had to create a realistic model of the turbulence itself. Creating a realistic model of the turbulence requires comparing the model with a real-case measurement at the telescope that is to be simulated. This in turn implies comparing long-exposure PSFs of the simulated and real cases.

Long- and short-exposure PSFs in AO systems are different. Short-exposure PSFs ( $t_{\text{exp}} < \tau$ ) are dominated by a bright diffraction core and high-frequency spatial structures called ‘speckles’. Long-exposure PSFs ( $t_{\text{exp}} \gg \tau$ ) have a central core with the presence of diffraction rings that are similar to those of the Airy diffraction pattern, while the halo is smooth and resembles that of a seeing-dominated PSF (Fig. 9). The transition point between the two regions is called the control radius and is identified by a bump in the PSF luminosity profile (see Fig. 10). The control radius is a function of the number of corrected modes in the AO system according to the following equation:

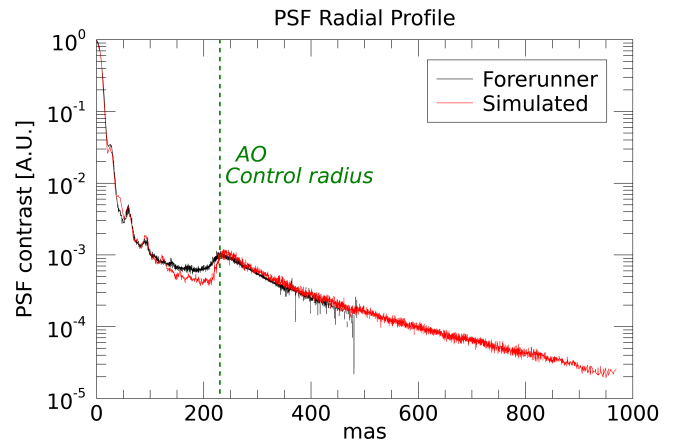
$$R_c [\text{mas}] \approx \frac{180 \cdot 3600}{\pi} \sqrt{\frac{N_{\text{modes}}}{\pi}} \frac{\lambda}{D} = 210 \text{ mas} \quad (3)$$

where  $R_c$  is the control radius,  $N_{\text{modes}}$  the total number of actuators in the AO system,  $\lambda$  is the observing wavelength and  $D$  is the diameter of the telescope pupil. For the LBT forerunner experiment,  $N_{\text{modes}} = 550$ ,  $\lambda = 630 \text{ nm}$ , and  $D = 8.2 \text{ m}$ , resulting in  $R_c = 210 \text{ mas}$  (Pedichini et al. 2017; Tyson 2011).

To obtain the long-exposure PSF profiles, we added random phases to our optical simulation of the wave front at the



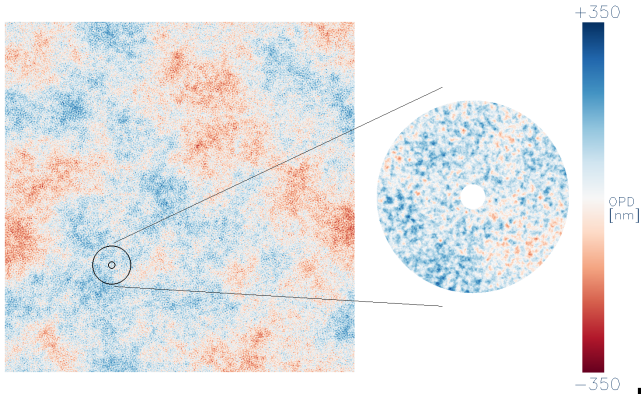
**Fig. 9.** Comparison between simulated and on-sky images of forerunner PSFs at LBT. *Top row:* PSF-simulated images. *Bottom row:* forerunner at LBT PSF images at  $\lambda = 630 \text{ nm}$ ,  $f_{\#} = 44$ . *Left column:* Short 1 ms exposures. *Right column:* long 1 s exposures. In the real on-sky image, it is possible to see the slow wind effect left uncorrected by AO and causing an asymmetric scintillation halo structure, as explained in Madurowicz et al. (2019).



**Fig. 10.** Long exposure PSF (1 s) radial profile comparison between real (Forerunner) and simulated data.

pupil plane. The phase screen representing the random phases was sampled every 18.6 mm across the LBT aperture of 8.2 m diameter; its PSD was derived from a low-pass filtered Kolmogorov spectrum similar to the response of the LBT AO system with a WFE value of approximately  $42.5 \pm 2 \text{ nm RMS}$ . This Optical Path Difference (OPD) phase screen, which simulates typical AO residual turbulence during closed dome calibrations, has been verified experimentally at LBT in the same conditions (Quirós-Pacheco et al. 2010). Data were saved in a large fits file of  $4096 \times 4096$  pixel (see Fig. 11).

We then took random  $440 \times 440$  pixel portions of this phase screen and injected them in simulations generating a series of images equivalent to an exposure time of 1 ms. We summed 1000 of these short-exposure images in order to create the PSF of an equivalent long-exposure of 1 s and normalised it to its peak to compare its luminosity profile with the real one obtained from the SHARK-VIS forerunner experiment PSF at 1 s exposure (Pedichini et al. 2017); see Figs. 10 and 9. The simulated PSF has a control radius that matches the real one well, and the core and wing profiles also match. The phase screen of AO residuals is also useful to simulate short integration times. As



**Fig. 11.** OPD layer phase screen ( $4096 \times 4096$  pixel) used to emulate LBT AO residual turbulence with WFE  $\approx 42.5$  nm RMS. The annulus shows a 440 pixel diameter disc corresponding to the real pupil of the telescope.

the refreshed timescale of speckles is of the order of milliseconds, as expected according to Stangalini et al. (2017), using a high-speed acquisition frame rate ( $>1$  KHz) it is possible to freeze the atmospheric turbulence and acquire images similar to our AO residual simulation (see Fig. 9). This condition is still valid in every image plane, including the extra-focal one; furthermore, the defocusing performs a smoothing and broadening effect on high spatial frequency structures, which is helpful for the PCA encoding. We can exploit the fast imaging approach and the corresponding atmospheric frozen flow condition to infer the NCPA as an average of the NN prediction results. The testing set with AO residuals was obtained from a set of constant Zernike coefficients, representing a fixed NCPA and random AO residuals added as different OPDs for each of the simulated images of the dataset. This is a good representation of a short burst of 1 ms exposure images acquired during a calibration run when the NCPA is static and the AO residual varies on milliseconds scales. The AO residual can be considered as noise for the NN and thus is not expected to influence the training process.

### 2.5. NCPA mitigation concept with AO residuals

To test the NN efficiency, we first generated different sets of images: each image of one set had the same fixed NCPA. For each set, we let the NN measure the NCPA (Zernike coefficients) of every image. These values were then averaged over the whole set to estimate the total NCPA for that set. We then measured the MAE between the injected and measured values and implemented an iterative process where successive NCPA estimates are removed and the corrected images are then fed to the NN for a new estimate. In a real-world operation, this is equivalent to measuring the NCPA, sending the corresponding correction to the AO system for removal and repeating the measure until the NCPA value is below a certain threshold.

This process was simulated with the `proper_lib` and the 42.5 nm phase screen. The NCPA measured by the NN in one set is multiplied by a gain factor and then subtracted from the initial NCPA (which had been injected in that set): the new NCPA is then injected in the simulation tool to create a new set of images, as per the following equation:

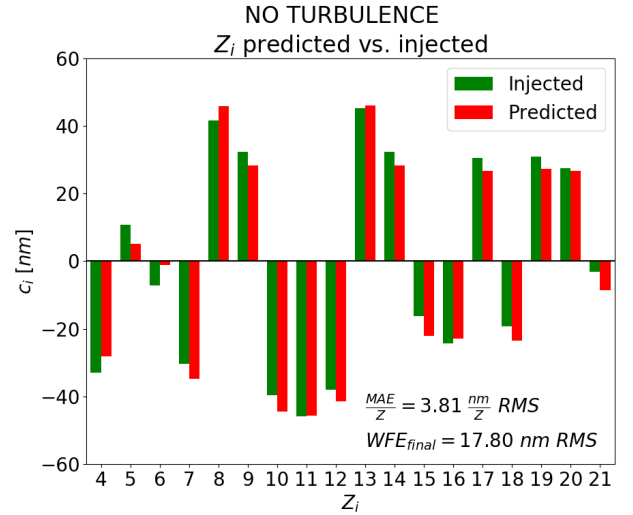
$$\text{NCPA}(i+1) = \text{NCPA}(i) - g \cdot \text{NN}(i), \quad (4)$$

where  $i$  is the iteration step,  $\text{NCPA}(i)$  is the NCPA at step  $i$ ,  $\text{NN}(i)$  is the NN output at step  $i$ , and  $g$  is the gain factor.

**Table 2.** Summary of the results obtained for the WFE [nm] RMS by the NN prediction averaged for every dataset.

Dataset ( $\lambda = 656$ nm)	Initial WFE [nm]	Recovered WFE [nm]	MAE [nm]
No turbulence + NCPA	140	26	2.36
AO res.+ NCPA	133	58	5.36
AO res. NCPA mitigation	150	44	1.84

**Notes.** AO residual RMS WFE is 42.5 nm.



**Fig. 12.** No turbulence case. Comparison of predicted (red) and injected (green) Zernike values of a single image of the testing set (exact values in Table 3).

We let the simulation run for 15 iterations to verify if oscillations were triggered for certain gain values. Results are presented in Sect. 3.4.

## 3. Results

The values from NN models with and without AO residual turbulence are summarised in Table 2. The recovered RMS WFEs are the final residual RMS after applying the NCPA calculated by the NN as a correction to the input wave front. The results are the average over the whole set.

### 3.1. Ideal condition: No turbulence

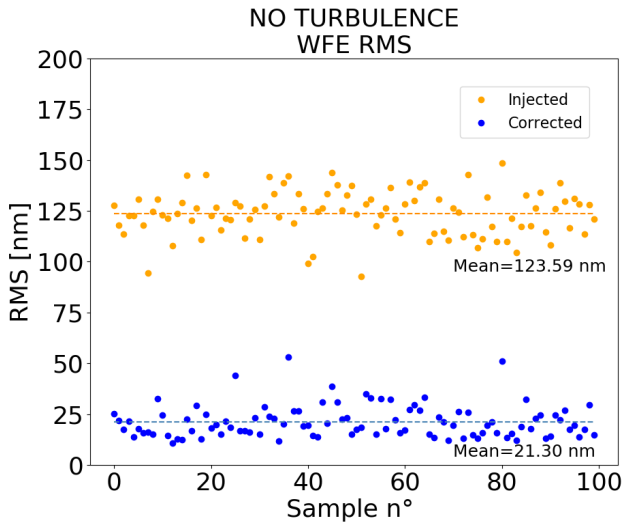
The median MAE achieved from our predictions in this case is 3.67 nm, and in Fig. 12 we compare the magnitudes of injected and retrieved values versus Zernike order for a specific case, with values reported in Table 3. We further reduced the WFE and flattened the NCPA components by applying recursive predictions and corrections, as described in Sect. 2.5.

In Fig. 13, we plot the RMS WFE values before and after the prediction and correction phase. As shown, the NN reduces the RMS WFE by a factor of almost 6, taking the average on the test set from  $\text{WFE}_{\text{initial}} \sim 124$  nm to  $\text{WFE}_{\text{final}} \sim 21$  nm. In Fig. 14, we show the effects on the PSF after the correction of NCPA by the NN prediction in one of the simulated cases. The NN prediction, however, can then be re-applied after the first step of correction and so on with the aim of improving the final results, as in Fig. 15, where we find that using a process gain of 0.8 makes it possible to flatten the wave front to a final MAE of  $\approx 10$  nm RMS



**Table 3.** Injected Zernike coefficients and the corresponding values found by the NN in two images of the testsets for the two different cases of no turbulence and AO residual 42.5 nm turbulence.

Zernike n°	Values from Fig. 12 [nm]		Values from Fig. 16 [nm]	
	Injected	Predicted	Injected	Predicted
Z <sub>4</sub>	-32.96	-28.18	-12.20	-7.19
Z <sub>5</sub>	10.84	5.21	0.48	-2.77
Z <sub>6</sub>	-7.25	-1.05	39.03	38.53
Z <sub>7</sub>	-30.41	-34.89	7.73	9.91
Z <sub>8</sub>	41.55	45.89	-10.33	-15.54
Z <sub>9</sub>	32.40	28.27	-10.60	-6.57
Z <sub>10</sub>	-39.65	-44.58	29.52	18.29
Z <sub>11</sub>	-45.94	-45.68	-25.53	-30.26
Z <sub>12</sub>	-38.05	-41.37	28.65	20.17
Z <sub>13</sub>	45.17	46.04	3.66	11.02
Z <sub>14</sub>	32.42	28.23	46.86	34.53
Z <sub>15</sub>	-16.27	-22.20	-24.22	-24.71
Z <sub>16</sub>	-24.38	-22.98	-7.34	-6.87
Z <sub>17</sub>	30.49	26.77	-46.07	-54.12
Z <sub>18</sub>	-19.29	-23.58	-21.41	-27.45
Z <sub>19</sub>	30.90	27.31	-12.94	-11.59
Z <sub>20</sub>	27.54	26.6	34.53	34.54
Z <sub>21</sub>	-3.06	-8.65	-49.8	-47.18

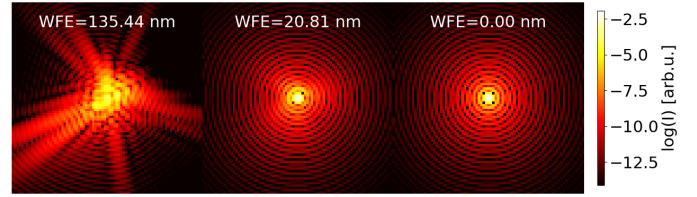


**Fig. 13.** NCPA correction in the case of no turbulence. Dashed lines represent the mean, over the entire testing set, of initial and corrected values. Each sample represents a ‘single iteration’ of the NN prediction and simulated NCPA correction.

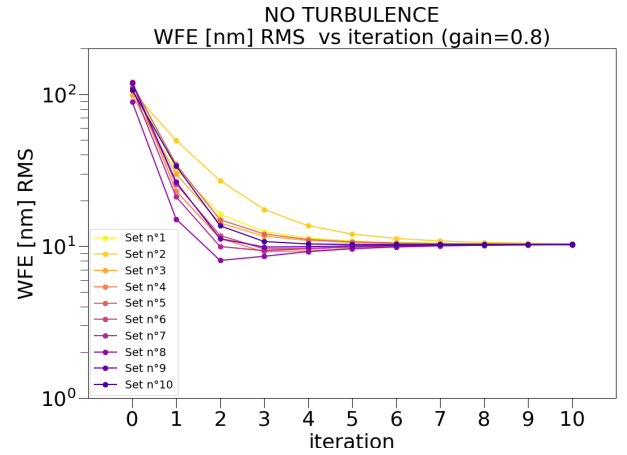
in a few steps despite the initial NCPA components. This result is also mathematically supported by the quadratic sum of the MAE of the 18 Zernike components (each one with a MAE of 2.36 nm RMS) giving a final value for the WFE of  $\approx 10$  nm RMS.

### 3.2. Addition of AO residual turbulence: NN behaviour

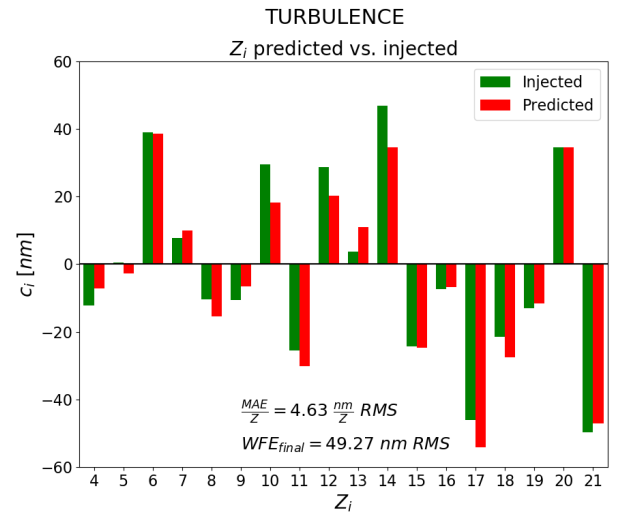
In this section, we cover the behaviour of the NN when typical AO residual aberrations are present and when using short exposure images as NN input (see Sect. 2.4). In this test set, every image has both different AO residuals and different NCPA aberrations. We increased the number of training samples from



**Fig. 14.** PSF of one case from Fig. 13 in log scaling. *From left to right:* NCPA aberrated; corrected for NCPA with only one NN prediction; diffraction-limited PSF (for comparison).

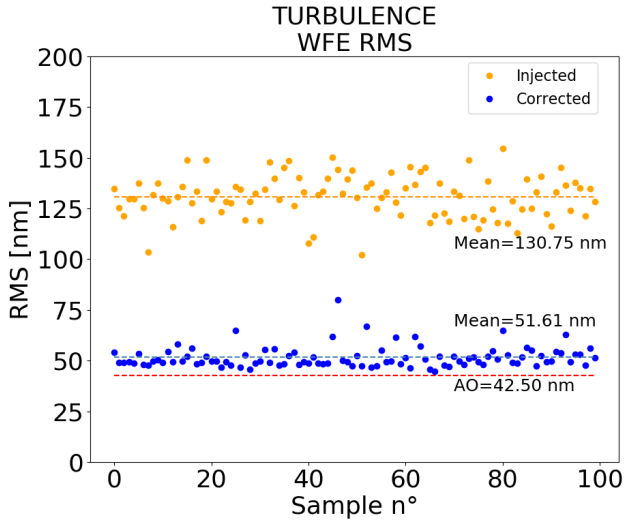


**Fig. 15.** Effects of iterative application of the NCPA prediction and correction by NN on ten different cases of initial NCPA using a gain of 0.8. The MAE is always reduced to only 10 nm RMS in a few steps.



**Fig. 16.** Turbulence case. Comparison of predicted (red) and injected (green) Zernike values of a single short exposure image of the testing set (exact values in Table 3).

10 000 to 15 000 in order to improve the NN prediction capability, reaching a median value of MAE = 5.17 nm. Figure 16 shows a specific case to compare the magnitudes of injected and retrieved values versus Zernike order. Exact numerical values are reported in Table 3. The WFE after prediction by the NN and correction is less than half of the initial value and close to the AO correction without NCPA (Fig. 17).



**Fig. 17.** NCPA correction in case of  $\approx 42.5$  nm turbulence based on our NN model predictions. Dashed lines represent the mean, over the entire testing-set, of initial and corrected values. The red dashed line is the AO system limit without NCPA. In this test set every image has both different AO residuals and different NCPA aberrations.

This test shows that the NN is effective in identifying aberrations with sufficient accuracy even if high-order aberrations from AO residuals are present. This includes aberrations that were not labelled during the training phase, i.e. the NN has not been 'told' what they are.

### 3.3. Estimation of static NCPA with AO residuals

In a sequence of short exposure images, we have different WFEs, and, consequently PSFs with different sharpnesses, depending on the combination between the NCPA and the instantaneous AO residual; thus, it is not possible to estimate quasi-static aberrations and minimise the average WFE with a NN prediction on a single image. This is why several images arranged in a time series are needed for a correct evaluation of the NCPA. To verify this, we fixed a static NCPA value for a dataset of 100 images and let the NN infer the NCPA value. The predictions of the injected NCPA are shown for each Zernike component in Fig. 18. The average of the NN predictions for each Zernike is assumed as the NCPA estimate for that component. For most of the NCPA Zernike components, the typical estimate error is less than 10 nm.

### 3.4. NCPA mitigation process

As described in Sect. 2.5, we simulated NCPA correction in a real-world system by iteratively estimating it from a set of images, correcting the same set and feeding the corrected images to the NN for a new estimate. In Fig. 19 we present the results of 15 steps and of the values of the single Zernike coefficients after each successive step in the iteration.

We then ran three 15-step simulations, each with different gain settings. The results are presented in Fig. 20 with the resulting WFE at every step of simulated correction; a gain,  $g = 1$  is still stable and allows the NCPA to be removed after only three simulation steps. This means that a few hundred images are sufficient to correct the NCPA of the system. In a fast imaging instrument, such as SHARK-VIS, this corresponds to about a few seconds of exposure, assuming the time for NN predictions

is negligible with respect to the exposure time. The RMS WFE goes from the initial value of WFE  $\approx 185$  nm to WFE  $\approx 43$  nm, which is the typical value of the residual turbulence left by AO in closed dome conditions.

We also verified the NN mitigation process on 20 sets of data with different NCPA values to see if the method is robust in different conditions. We used a gain value of  $g = 1$ , 42.5 nm of residual AO turbulence and only three iteration steps.

Figure 21 shows the results, that NCPA are almost completely removed (see Sect. 3.2). Figure 22 shows the effect of the NN NCPA mitigation process on the sharpness of the PSF in one of the simulated cases. It is worth noting that, in our specific case, there is no temporal correlation between the images of the 100 image testing set because the AO residual is sampled randomly. This allows us to use a smaller number of samples in the testing set to give a solution with sufficient accuracy.

In a real-world setting, the states of atmospheric turbulence between consecutive frames during acquisitions will not be independent, and this decoupling effect will not be valid. Thus, to achieve sufficient accuracy at each step of the mitigation process, images should be acquired with short exposures ( $t_{\text{exp}}$  ranging from 1 to 2 ms) and sampled more slowly than the typical turbulence de-correlation time, which is between 10 ms and 100 ms, as expected by Stangalini et al. (2017).

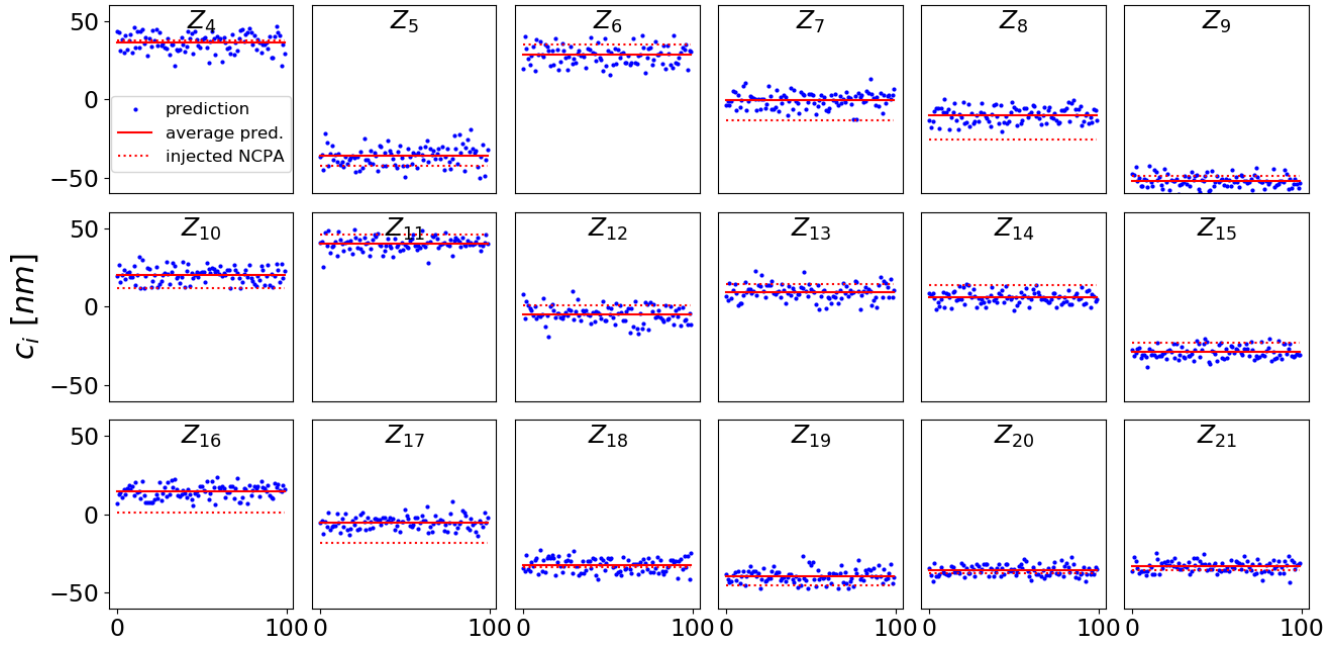
## 4. Discussion

Comparing the mitigation processes without turbulence (Fig. 15) to that with AO residuals (Fig. 20), it is possible to notice that, counter-intuitively, the measurements of the NCPA is more accurate when working with turbulence than without it, passing from a value of approximately 10 nm to almost 7 nm RMS (Fig. 19). We also have to take into account that the NN trained on turbulence has almost double the MAE for each Zernike, so the expected value for the mitigation process (given by the squared sum of the MAE for each Zernike) is  $5.17 * \sqrt{18} \approx 22$  nm. The reason is in the averaging of the NN predictions, which provides the system with a way to overcome the NN limits, making it perform better in the presence of turbulence on instruments capable of fast imaging.

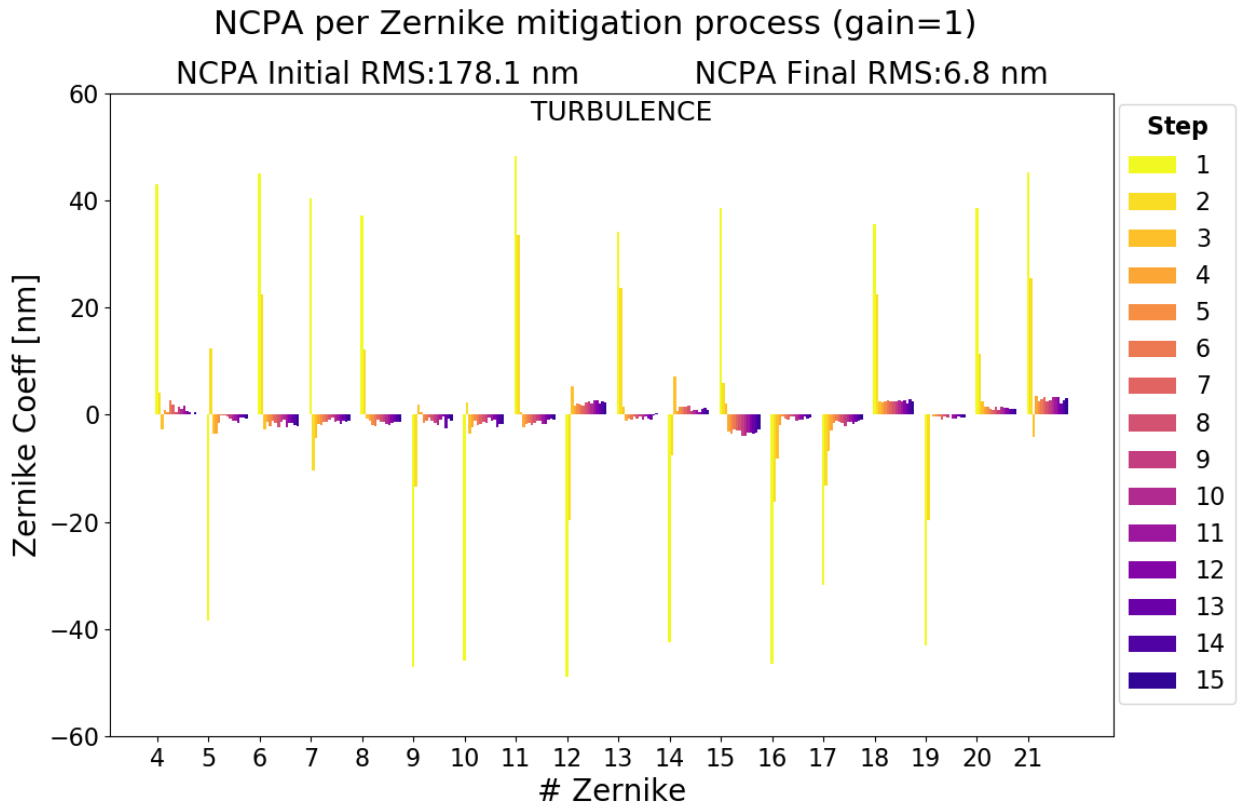
Preliminary results (which will be the subject of future articles) show that the method is robust even in the presence of photon and detector noise, and also in more generic cases. Such cases would be, for example, spiders at the pupil plane, wide-band filter images instead of monochromatic light, coronagraphic images, different  $f\#$ , and defocused offsets.

In Fig. 23, we present just a few of these aspects: the presence of spiders, the effect of photon and read-out noise (RON), and the use of a wide-band filter. We performed these tests both with and without turbulence (WFE = 42.5 nm RMS), basing them only on a single prediction by the NN in order to quickly estimate the potential of the method even in these scenarios. Furthermore, the NN used for these preliminary tests was trained on monochromatic images at  $\lambda = 656$  nm without any spider or image noise and not on similar images, as should be the case.

We added in the optical propagation a mask corresponding to the LBT spiders in order to have a more realistic simulation and applied the mitigation process. We followed the same procedure considering a more realistic source dominated by Poisson noise (PN). We simulated a defocused guide star with a flux of  $5 \times 10^4$  photons  $s^{-1}$  spread over a few thousand pixels, which was affected by photon noise and acquired with a detector read-out noise  $\text{RON}[\text{RMS}] = 1e - /\text{pixel}$ . Wide-band filter images were



**Fig. 18.** Comparison of injected (red dotted line) and averaged prediction of NCPA component (red line) for each  $Z_i$  in the simulation when NCPA is fixed over the whole dataset of 100 images while the turbulence is evolving. Blue dots are the single NN predictions for NCPA.

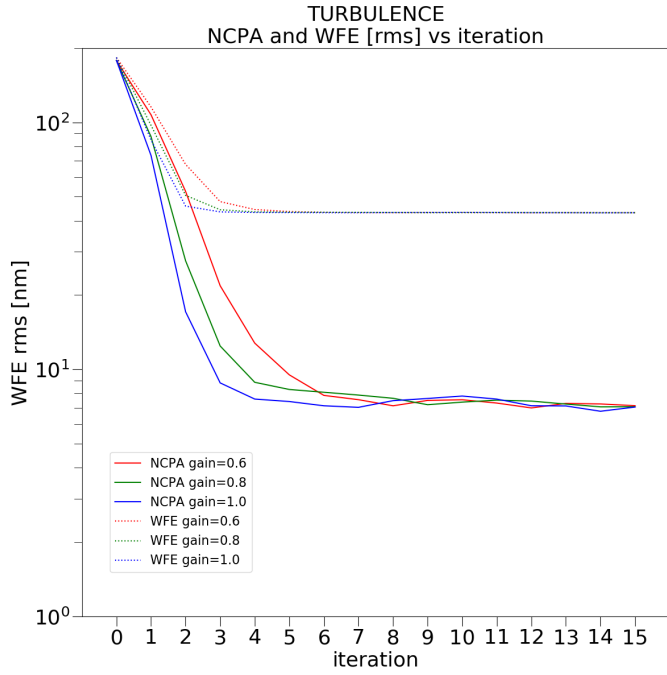


**Fig. 19.** Mitigation iterative process. Barplots of the NCPA Zernike coefficients after each iteration. For each Zernike, the bars, from yellow to blue, show the value of the components almost decreasing at each iteration step.

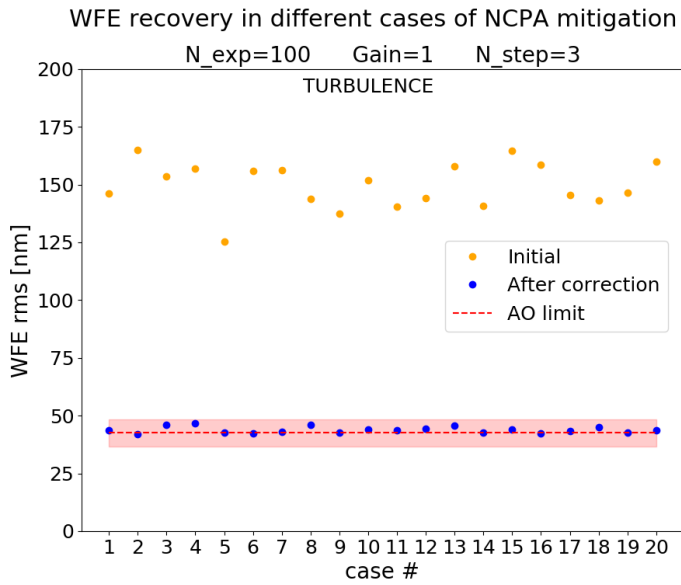
generated with a spectral bandwidth of  $FWHM = 60$  nm centred at  $\lambda = 656$  nm coadding 12 monochromatic PSFs within the spectral bandwidth equally spaced out by 5 nm. As can be seen from the plots in Fig. 23, results are almost identical to the ones obtained by using monochromatic light in ideal conditions both in the presence and absence of turbulence. In all these

described scenarios, we fixed the same amount of initial NCPA values and then performed the mitigation process as explained in Sect. 2.5.

We note that once verified on real hardware (telescope or optical bench), the imaging-detector-with-NN approach may become a cheap substitute for WFSs in all cases in which these

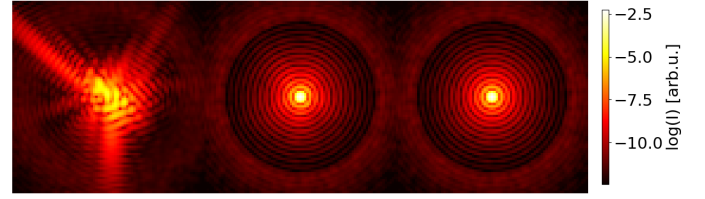


**Fig. 20.** Temporal evolution of NCPA mitigation process for one AO residual condition. NCPA aberrations' RMS WFE (solid lines) and total RMS WFE (dashed lines) versus iterations for different gain values. The NCPA is well mitigated after a few iterations, independently of the loop gain value.

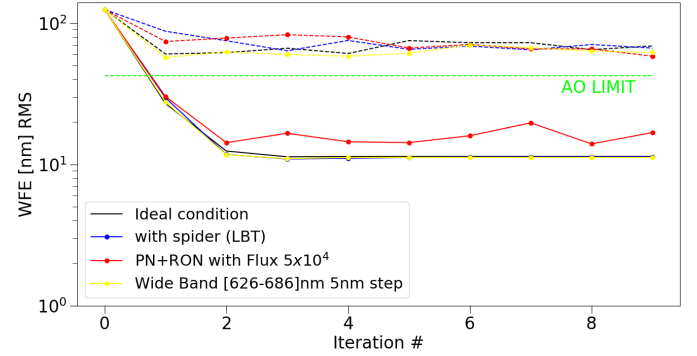


**Fig. 21.** WFE expressed in RMS [nm] versus a different set of NCPA. The applied NCPA correction is estimated by the average of NN predictions over 100 images. The red gradient bar represents the  $\pm 3\sigma$  error bar for the AO residual.

are not retro-fittable in already developed and installed instrumentation, provided a focusing system is present. Despite its imperfections and limited accuracy with respect to a dedicated WFS or an optical interferometer, simulations show the NN is sufficiently precise to mitigate NCPA in AO systems where a final WFE less than 80–100 nm RMS is not expected. In the future, the team will investigate the opportunity to increase the number of used components to map the aberrations and the use



**Fig. 22.** Long-exposure (100 ms) PSF of one case of Fig. 21 in log scaling. *From left to right:* NCPA aberrated with a WFE  $\approx 185$  nm RMS; NN corrected for NCPA; AO limit PSF (for comparison).



**Fig. 23.** Comparative results of iterative mitigation process based only on a single NN prediction in case of presence of spider, a noise-dominated faint source of  $5 \times 10^4$  photons, and a wide-band filter. Solid lines represent the case without turbulence while dashed lines represent the case with an additional WFE = 42.5 nm RMS of turbulence.

of different NNs tuned to different amplitudes of aberrations in a cascaded system where the first NN is trained on large aberration ranges (e.g. 250 nm RMS) and the successive ones are instead tuned and optimised to smaller ranges. This would act as a multiple-step process in which the larger aberrations are measured and removed first, and smaller scale aberrations are then measured and removed by increasingly more finely tuned systems.

## 5. Conclusions

In this work, thanks to optical propagation simulations based on a simplified layout where we used monochromatic light, ideal optical elements without obstructions (such as spiders or occulters) and no photon noise, we demonstrated that a simple multi-layer perceptron NN is able to establish a robust relation between defocused PSF images and the corresponding pupil phase aberrations. This relation also holds in the presence of limited AO residual turbulence, such as the one present during closed-dome calibrations. This is possible thanks to a pre-processing step in which the dimensionality of the data is reduced by projecting images onto an orthonormal base of PCA eigen images and the pupil wave front phase onto Zernike polynomials. In the training phase, we feed the NN with the PCA components of a set of images and the Zernike coefficients from which they were generated. In the predicting phase, after being fed with the images to be analysed through their PCA components, the NN outputs the Zernike coefficients that describe the NCPA aberrations.

Typical WFE measurements, retrieved during our simulations, showed errors of a few nanometers per Zernike mode with a MAE ranging from 2 to 3 nm RMS that encouraged us to apply

this method to a more realistic case. Here, the AO residual turbulence typical of LBT closed-dome calibrations was added to the static pupil aberration. Here, by using a dedicated NN and starting from an initial RMS WFE of 180 nm, the final RMS WFE was drastically reduced to less than a few nm above the AO residual limit value of 42.5 nm RMS with only three steps of measurement and correction of the NCPA component.

Further studies in this field are planned to improve the MAE so that the method works even in the case of AO full turbulence residual, photon, and detector noise. This will be done by searching for an optimal defocus position, for the optimal ratio between the number of Zernike polynomials and the number of samples in the training stage, for different kind of basis to be used instead of PCA and Zernike, and for different NN layouts. The method presented in this work has two important features. The first is the possibility to train the NN to recognise static aberrations hidden by the larger and dynamic aberrations of atmospheric turbulence. The second is that the method is computationally fast, both in the training and application stages, because we do not use the whole image but its compact representation through PCA coefficients. In a fast imaging approach, this can be used to measure NCPA as the average of the NN prediction, taking care that the average is taken over a period at least one order of magnitude longer than the typical turbulence timescale. We expect that this method can be used directly on sky, once the NN is trained on a realistic model of the telescope optics or by injecting aberrations with its AO system, because it does not need additional hardware to be added to the instrumentation or telescope. We are currently working to confirm the results of these simulations on a dedicated optical bench in the ADONI laboratory at INAF (Rome) and at the ESO (Garching) GHOST bench Kasper et al. (2021).

*Acknowledgements.* The authors would like to thank Dr. Michele Buzzicotti, researcher at University of Tor Vergata, for his suggestions during the final review of this paper. We also thank Dr. Markus Kasper who has agreed to give us the opportunity to test the method on the ESO GHOST testbench.

## References

- Agapito, G., & Pinna, E. 2019, *J. Astron. Teles. Instrum. Syst.*, 5, 1
- Amara, A., & Quanz, S. P. 2012, *MNRAS*, 427, 948
- Andor 2018, in <https://andor.oxinst.com/products/scmos-camera-series/zyla-4-2-scmos#product-information-tabs>
- Causi, G. L., Stangalini, M., Antonucci, S., et al. 2018, *SPIE*, 10703, 892
- Esposito, S., Riccardi, A., Pinna, E., et al. 2012, *SPIE*, 8447, 312
- Esposito, S., Puglisi, A., Pinna, E., et al. 2020, *A&A*, 636, A88
- Gonsalves, R. A. 1982, *Opt. Eng.*, 21, 829
- Gonsalves, R. A., & Nisenson, P. 1991, *SPIE*, 1567, 294
- Guyon, O., Gallet, B., Pluzhnik, E. A., Takami, H., & Tamura, M. 2006, *SPIE*, 6272, 1081
- Hill, J. M. 2010, *Appl. Opt.*, 49, D115
- Hope, D., Jefferies, S., Li Causi, G., et al. 2019, in *Advanced Maui Optical and Space Surveillance Technologies Conference*, ed. S. Ryan, 3
- Hornik, K., Stinchcombe, M., & White, H. 1989, *Neural Netw.*, 2, 359
- Hotelling, H. 1933, *J. Educational Psychol.*, 24, 417
- Jedrzejewski, R. I., Hartig, G. F., Nota, A., et al. 1993, *SPIE*, 1945, 244
- Kasper, M., Urra, N., Pathak, P., et al. 2021, *The Messenger*, 182, 38
- Kingma, D. P., & Ba, J. 2017, arXiv e-prints [arXiv:1412.6980]
- Korkiakoski, V., Keller, C. U., Doelman, N., et al. 2014, *SPIE*, 9148, 1746
- Krist, J. E. 2007, *SPIE*, 6675, 250
- Lagrange, A.-M., Gratadour, D., Chauvin, G., et al. 2009, *A&A*, 493, L21
- Li Causi, G., Stangalini, M., Antonucci, S., et al. 2017, *ApJ*, 849, 85
- Madurowicz, A., Macintosh, B., Chilcote, J., et al. 2019, *J. Astron. Teles. Instrum. Syst.*, 5, 1
- Marois, C., Macintosh, B., Barman, T., et al. 2008, *Science*, 322, 1348
- Mattioli, M., Pedichini, F., Antonucci, S., et al. 2018, *SPIE*, 10702, 1332
- Mattioli, M., Pedichini, F., Antonucci, S., et al. 2019, *Res. Notes Am. Astron. Soc.*, 3, 20
- Mendiña Fernández, A. 2021, PhD thesis, University of Oxford, UK
- Mesa, D., Keppler, M., Cantalloube, F., et al. 2019, *A&A*, 632, A25
- Noll, R. J. 1976, *J. Opt. Soc. Am.*, 66, 207
- Norris, B. R. M., Wei, J., Betters, C. H., Wong, A., & Leon-Saval, S. G. 2020, *Nat. Commun.*, 11, 5335
- Orban de Xivry, G., Quesnel, M., Vanberg, P.-O., Absil, O., & Louppe, G. 2021, *MNRAS*, 505, 5702
- Pedichini, F., Ambrosino, F., Centrone, M., et al. 2016, *SPIE*, 9908, 964
- Pedichini, F., Stangalini, M., Ambrosino, F., et al. 2017, *AJ*, 154, 74
- Pinna, E., Pedichini, F., Esposito, S., et al. 2015, in *Adaptive Optics for Extremely Large Telescopes 4*, Conf. Proc. (University of California Center for Adaptive Optics), Conf. Proc., AO4ELT 2015
- Pinna, E., Esposito, S., Hinz, P., et al. 2016, *SPIE*, 9909, 1183
- Potier, A., Ruane, G., Chen, P., et al. 2021, *SPIE*, 11823, 118231L
- Quirós-Pacheco, F., Busoni, L., Agapito, G., et al. 2010, *SPIE*, 7736, 1245
- Roddier, F. 1988, *Appl. Opt.*, 27, 1223
- Roddier, F., & Roddier, C. 1991, *Appl. Opt.*, 30, 1325
- Roddier, C., & Roddier, F. 1993, *J. Opt. Soc. Am. A*, 10, 2277
- Roddier, F., Roddier, C., & Roddier, N. 1988, *SPIE*, 0976, 203
- Ross, T. S. 2009, *Appl. Opt.*, 48, 1812
- Ross, D., Lim, J., Lin, R.-S., & Yang, M.-H. 2008, *Int. J. Comput. Vision*, 77, 125
- Sauvage, J.-F., Fusco, T., LeMignant, D., et al. 2011, *Second International Conference on Adaptive Optics for Extremely Large Telescopes*, 48,
- Schwab, C., Peter, D., & Aigner, S. 2010, *SPIE*, 7736, 1081
- Stangalini, M., Pedichini, F., Pinna, E., et al. 2017, *J. Astron. Teles. Instrum. Syst.*, 3, 1
- Stangalini, M., Li Causi, G., Pedichini, F., et al. 2018, *ApJ*, 868, 6
- Szapiel, S. 1983, *Opt. Lett.*, 8, 327
- Terreri, A., Pedichini, F., Li Causi, G., Mattioli, M., & Stangali 2019a, in <https://doi.org/10.36116/WFS2019.23>
- Terreri, A., Stangalini, M., Pedichini, F., & De Petris, M. 2019b, PhD thesis, Università di Roma “La Sapienza”, Italy
- Tyson, R. L. & Frazier, B. W. 2011, *Principles of Adaptive Optics (USA)*: CRC Press, Taylor & Francis Group, LLC
- Van Der Malsburg, C. 1986, in *Brain Theory*, eds. G. Palm, & A. Aertsen (Berlin, Heidelberg: Springer Berlin Heidelberg), 245
- Vassallo, D., Farinato, J., Sauvage, J.-F., et al. 2018, *SPIE*, 10705, 433