Non-common Path Aberration Compensation Using the NWIWM Method

L.F. Rodríguez-Ramos*[a], Javier López Campos[a], Oscar Tubío Araújo[a], Carlos Colodro Conde[a], Miguel Núñez Cagigal[a], José Marco de la Rosa[a]

[a]Instituto de Astrofísica de Canarias, Calle Vía Láctea, s/n, 38200 La Laguna. Tenerife, SPAIN

ABSTRACT

Accurate measurement of non-common path aberrations (NCPAs) is an important step to be undertaken correctly when operating a real adaptive optics system. NCPAs are defined as the aberrations which are present at the science image but cannot be seen by the wavefront sensor (WFS), basically due to the different placements along the optical path. Compensating these aberrations is required to obtain the best image at the science detector from the resolution point of view.

Obtaining the best resolution image, available by a particular optical system having a deformable mirror in the optical path, can be also considered an interesting problem from an abstract point of view, because it can simultaneously compensate for the NCPA and provide the best initial setup for the DM actuators, independent of the accuracy or calibration of the WFS.

Phase Diversity (PD) is a very commonly used method of measuring the NCPAs, based in the analysis of two images taken at different focus positions, or just one defocused image. There are also Focal plane Sharpening (FPS) methods, which only deal with the science image, trying to minimize the width of the Point Spread Function by blindly actuating on the DM.

The method described in this contribution, Noise Weighted Image Width Minimization (NWIWM), lies in the latter FPS group, and has been developed and tested for the AOLI, EDIFSE and GTCAO projects being developed at Institute of Astrophysics of the Canary Islands (IAC). It is based on the signal to noise analysis of a function describing the width of the PSF with respect to the DM actuations, both zonally and modally, in order to select the actuators or modes to be used during the minimization.

A complete description of the algorithm is included, together with simulation results and practical examples obtained within the above mentioned projects.

Keywords: Non-common Path Aberrations, Deformable mirror, adaptive optics, wavefront sensor, Point Spread Function Optimization

1. INTRODUCTION

Non-Common Path Aberrations can be defined as the wavefront difference between the end of the science path and the one existing at the wavefront sensor. They basically comprise the combined effect of the two different paths travelled by the object light, once separated by a beam splitter or dichroic, and sent to both the wavefront sensor and the science instrument. Given that these path will always be physically different, they will necessarily end up generating aberrations that cannot be measured by the WFS, but could eventually be corrected, or at least minimized, by an adequate setup of the DM.

Figure 1 depicts a typical configuration of a basic AO system in which a calibration facility has been implemented, and is providing the input light for the system. The light is then relayed to the deformable mirror (DM) before being split in the two aforementioned paths: The one towards the wavefront sensor (WFS) and the one to the science instrument. Relaying optics between elements has been represented by small boxes with two lenses, and the real-time computer, in charge of computing the required values of the actuators of the DM, has been included for completeness. The optics directly involved in the NCPA has been signalled, but it should also be noted that the science instrument could include a

huge amount of optics, i.e. a complete spectrograph or polarimeter, and the aberrations introduced by those optics elements cannot be measured by the WFS. There could also be differences between the paths coming out from the calibration subsystem and the real telescope. These differences, normally neglected, will be addressed later.

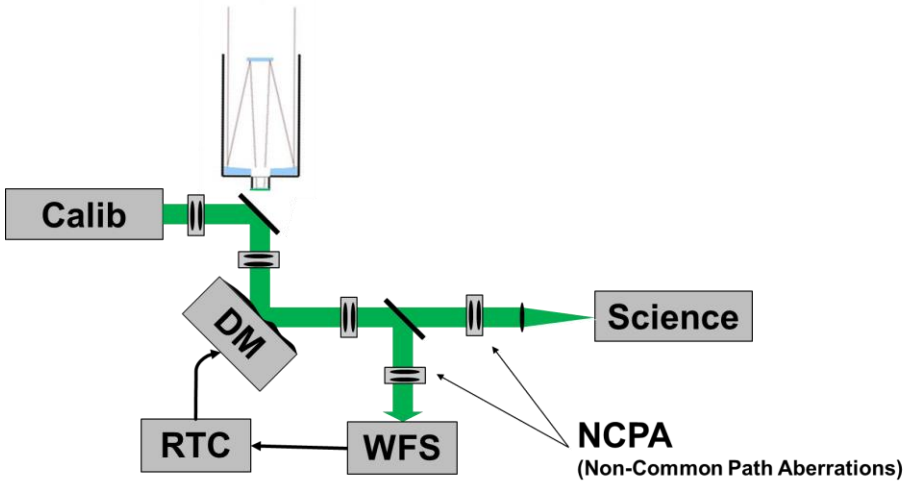

Figure 1. Typical layout for a basic AO closed-loop system. A calibration system, simulating the telescope output, is being used in the figure for generating a reference wavefront. Once after the deformable mirror, the light is divided in the two aforementioned paths, to the WFS and to the science instrument. Relaying optics have been depicted with small boxes with two lenses, and those involved in the NCPA have been identified.

Given that the WFS is by definition blind to the NCPAs, the science image is then the only measurement system available for dealing with them, and at the same time is the final result of the whole instrument. The problem of compensating the NCPAs can also be formulated as an optimization problem, where the Point Spread Function (PSF) of the complete instrument, even including the telescope, should be made as close as possible to the Fourier Transform of the overall pupil shape, which is normally well known.

The degrees of freedom available for compensating NCPAs, understanding that proper optical alignment has been accomplished during commissioning phases, are directly related with the number of actuators of the DM. There might be other optical elements capable of modifying the shape of the wavefront, like focusing elements, and they could be eventually incorporated to the optimization algorithm, but they will not be included in our study because they are normally very specific of a particular instrument.

As our DM provide us with a wavefront control mechanism, it is normally needed some way to extract the wavefront shape exclusively from the science image, or from a number of them, taken under specific setups, to compensate NCPA. NCPAs are supposed to be time-invariant, at least for a particular telescope+instrument configuration, so there is normally plenty of time to perform the required calculations, which are done off-line from astronomical observations. Iterative algorithms are used for this task[1, 2], where all available information about the optical system is included on top of the science image(s) itself.

There is a family of algorithms which do not rely on the knowledge of wavefront shape for trying its correction. Instead, they just use the shape of the science image itself and tries to optimize it using the a priori knowledge of how it should be in an ideal situation. These algorithms are named as “focal plane sharpening” following the nature of the used approach. As it will be shown, the proposed algorithm lies in this field.

2. THE NWIWM ALGORITHM

The ideal shape for a science image, when using a point object as source and a circular aperture, is the very well-known Airy disc. Other more complex apertures may have slightly different optimal imaging results, but they always present a central core where most energy should be concentrated. The very first rationale of the NWIWM algorithm is the maximization of the encircled energy within this central core, which in turn will provide the desirable minimum width of the point spread function (PSF) of the overall system.

The second key idea behind the NWIWM is related with the practicalities of the measurement of the encircled energy, and the role of any specific actuator in its maximization. The signal to noise ratio of the gradient measurement is understood to be a combined piece of information which includes many relevant factors, like the position of the actuator within the aperture, or the amount of differential actuation required in each iteration, or the reliability of the measurement itself. This approach allows a completely blind optimization from the point of view of the nature and shape of the deformable mirror, or the optical setup, and thus absolutely no use of the mirror shape, orientation or actuator arrangement is required, because all relevant effects will be automatically extracted from the signal to noise results of the gradient measurements.

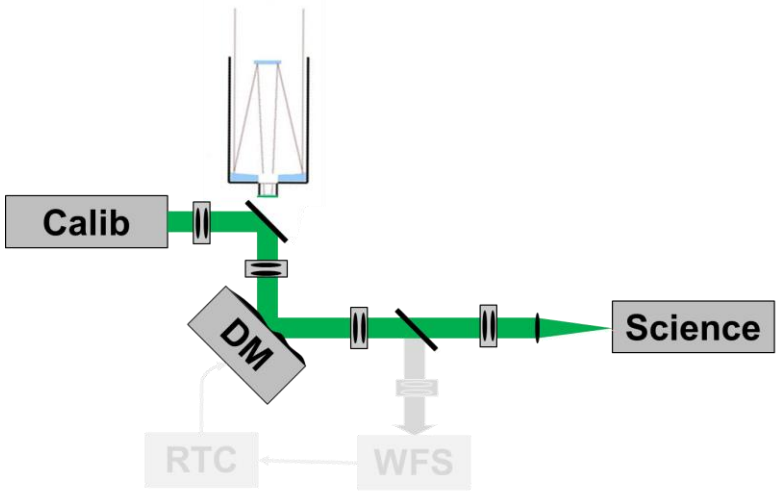

Figure 2. Fraction of the typical layout of a classic AO system which is used for the NCPA correction with the NWIWM algorithm. The wavefront sensor is not used at all, and only the part of the real-time computer devoted to DM actuator driving is used, with no need of any knowledge of the shape or arrangement of the actuators in the DM.

2.1 Encircled energy as a function of actuator position

The Encircled energy (EE) is defined within the NWIWM algorithm as a scalar function of all values of the deformable mirror actuators:

$$EE\ (a_1,a_2,\ldots a_N) = \sum_r p_i \qquad (1)$$

Where p_i are the values of the science image pixels and the sum is made on a region defined by a radius (r). This is by far the simplest function available and the only one tested to date, but it could be sophisticated by using distance weighted functions trying to exploit the a priori knowledge of the expected Airy shape. In order to implement the "keep it simple" approach, we have chosen to use the central core of the diffraction shape as the region used to compute the encircled energy, but we have also found that smaller regions can be used in a second or third round of the minimization algorithm to obtain further improvements of the Strehl ratio.

The size in pixels of this central core is normally known in advance by simple analysis of the optical design of the science imaging system, and the working wavelength. In the event that the science image (or spectrograph) is not designed to provide enough sampling of the diffraction PSF, it should be considered to temporally install a higher resolution camera just for NCPA compensation.

The objective of the NWIWM algorithm is then the iterative optimization (maximization) of the encircled energy as defined in (1), as function of the actuations values. This optimization does not require the use of any wavefront sensor nor any knowledge of the deformable mirror, except the way in which its actuators can be moved.

2.2 Science image pre-processing

Common pre-processing techniques are be applied to science images before the calculation of the EE is performed. After bias and flat corrections, some thresholding may be needed in order to improve the sensitivity of the EE to the actuator

movements, if the image contain any illumination pedestal. However, the key processing is the averaging of a number of images, which will provide both a better estimation of the EE and the most important piece of information, its measurement noise. This measurement noise plays a very important role in the algorithm, as stated before, and will allow us to adapt the measurement step in order to make the best possible gradient measurement.

2.3 The optimization loop

The rationale under the optimization loop may be stated as a very simple "Steepest Ascent to the Encircled Energy Peak". Having defined the EE as a scalar function of all actuators, there would be a peak in the N-dimension space which will provide the best actuator configuration from the point of view of the encircled energy. In order to keep on moving towards this peak, we measure the gradients with respect to each of the dimensions, compute the steepest direction as the one combining all gradients and then we seek the maximum along this direction to obtain the new position. This approach is depicted in Figure 3.

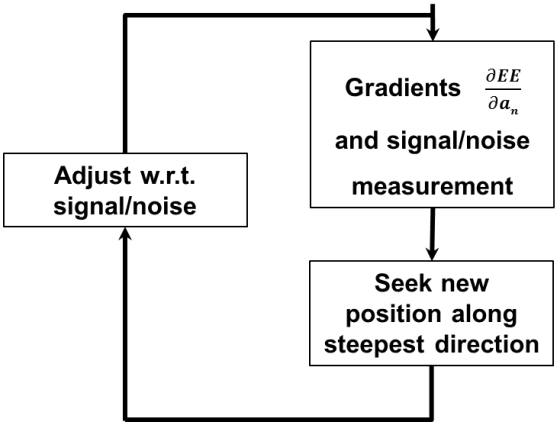

Figure 3. Schematic drawing of the NWIWM iteration concept. Steepest ascent to the EE peak is made by measuring gradients and moving along the steepest direction till a new maximum if found, which will be considered the next position.

Each gradient measurement is done using both a forward and a backward steps, whose size is adjusted at every iteration following the signal to noise results, in a way to be described later. EE is repeatedly measured at each backward and forward position and both average and standard deviation are recorded, in order to compute the gradient with respect to each actuator and its signal to noise value. The amount of EE samples to be averaged is one of the key parameters of the NWIWM algorithm, and depends on the expected noise behaviour of the system. Values from 4 to 50 have been used in our experiments and simulations, and no special dependence have been noticed except for the logical increase in measurement time. In the event that both measurements (backward and forward) give an EE lower than the central one, the gradient is understood to be zero and no movement will be made with this actuator. Gradients with a signal to noise below some level are also discarded, i.e. considered as zero, being this level also a parameter. Simulations and laboratory tests suggest that even values as low as 0.5 can be successfully used, due to the iterative nature of the algorithm.

The next position is sought along the direction constructed as a combination of the valid measured gradients, by computing encircled energies along a fixed number of steps (20-40) of a size which could be adapted to the estimated Strehl ratio, but in practice it can be maintained in a fixed value in a range of a few tens of nanometres in the DM actuation. The position along this line with the maximum EE is selected as the new one, and if eventually this positon is located at the last sample, a second run is done after doubling the step size, and so on.

Finally the last step of the iteration is to adjust the size of the actuator step to be used when measuring the gradients. This step is set at the beginning at a fairly high value, in the order of a micron or two, and changed individually for each actuator in order to end up obtaining a desired signal to noise figure in the next iteration, which turns out to be, as stated before, one of the key parameters of the NWIWM algorithm. This adjustment can be either to increase the step size if the signal to noise is below the desired level, or to reduce the step if the SNR is higher.

As it can be easily confirmed, there is no need of any previous knowledge of the system setup at all, because the algorithm probes the system systematically and blindly extracts the required information for the EE maximization. Of course, the system is considered to be time-invariant during the measurements.

3. SIMULATION RESULTS

The overall behaviour of the NWIWM algorithm has been simulated using the OOMAO software package[3]. This simulation package is MATLAB based, object oriented, very user friendly, and it provides resources to simulate most of the features required by an AO system. A 1 m telescope has been simulated using 60x60 pupil samples, with a science camera having a 20 electrons readout noise. The calibration system has been simulated using a time-invariant atmosphere with a very smooth turbulence layer and a 6^{th} magnitude star measured at I-band. The deformable mirror has been simulated to have 11x11 actuators (97 within the entrance pupil) and a monotonic influence function with 20% coupling between adjacent actuators.

NCPAs have been simulated using random contents of Zernike polynomials ranging from 3 to 15 in Noll's numbering scheme[4], as suggested by previous literature[5]. Piston (for obvious reasons), tip and tilt are then explicitly excluded, counting on the capability of obtaining a correct position alignment during AIV phase. NWIWM may eventually be also in charge of doing corrections of tip and tilt misalignments, if so decided. These random Zernike were adjusted to generate NCPAs ranging from 300-700 nanometers RMS. An example of the NCPA starting image and its NWIWM correction can be found in figure 4, and the user interface describing the intermediate stages of the simulation is shown in figure 5.

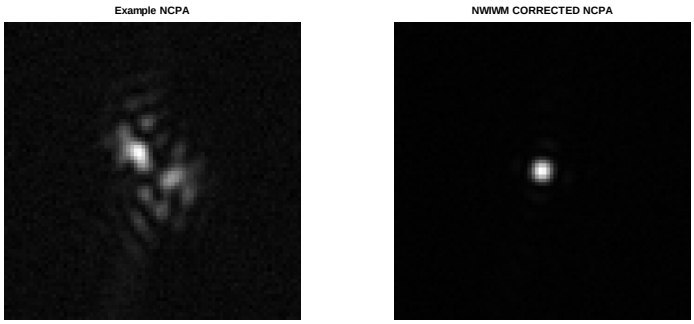

Figure 4. Example of NCPA correction using the NWIWM algorithm. On the left, starting image with NCPA of 270 nm RMS. On the right, the corrected image after 40 iterations.

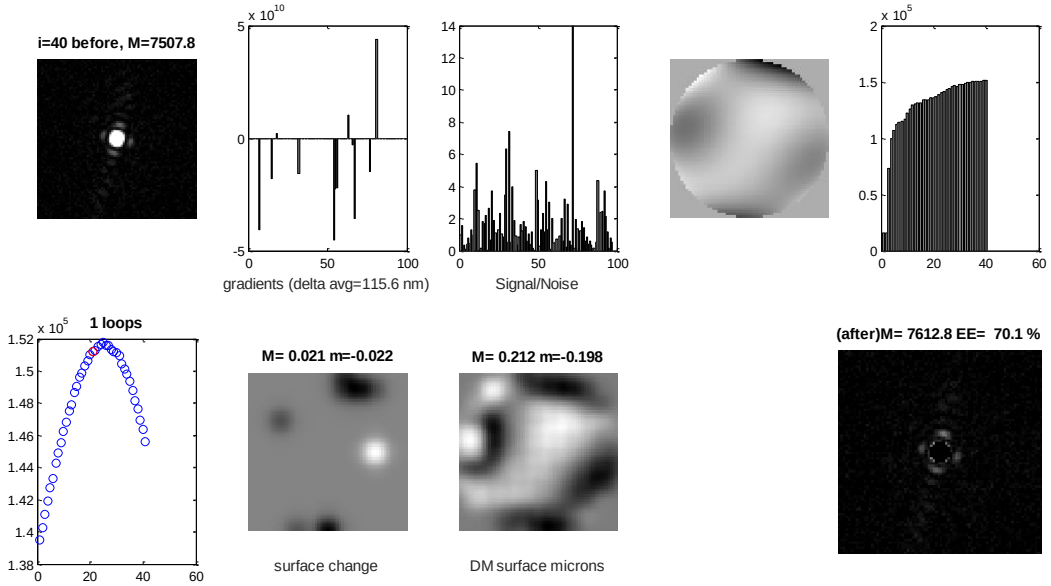

Figure 5. Example of the NWIWM user interface, taken after the 40th iteration of the NCPA depicted in previous figure. From left to right, in the upper row it can be found the starting image (after the previous iteration), the result of the gradient measurement, their SNR, the NCPA shape, the graph of the encircled energies found in all iterations. In the lower row, the line search performed (red circle is the starting position), the computed surface change, the resulting DM surface, and the residual image having extracted the core pixels used to determine the encircled energy. Both first and last images has been thresholded and amplified ten times to better notice the residual power being left out of the central core.

3.1 NCPA correction simulation results

Figure 6 shows the results of 120 random NCPA compensations using NWIWM, regarding the encircled energy. We have considered the diffraction image at the working wavelength as the overall reference for evaluating the EE, identifying the sum of the pixels within its central core with the absolute maximum EE achievable, and assigned as 100%. As far as within the simulation the NCPA value is known, a best square fit, computed using the set of influence functions, provides the best correction achievable in each case, known in the literature as the fitting error. The EE corresponding to this best possible correction is drawn for reference as blue asterisks, and they range between 85%-95% of the diffraction EE, with a clear dependence on the NCPA value.

The EE directly related to the simulated NCPA is depicted using green dots, normally below 10% of the diffraction EE. This value is the starting point of the NWIWM algorithm, and the EE values after 80 iterations have been depicted using red circles. The reference value used for the SNR of the gradient measurement was 1 in this case, and only gradients with SNR better than 0.5 were considered in each iteration. The improvement in EE is clear in all cases, but specially for low NCPA, where the final point can reach the 80% EE.

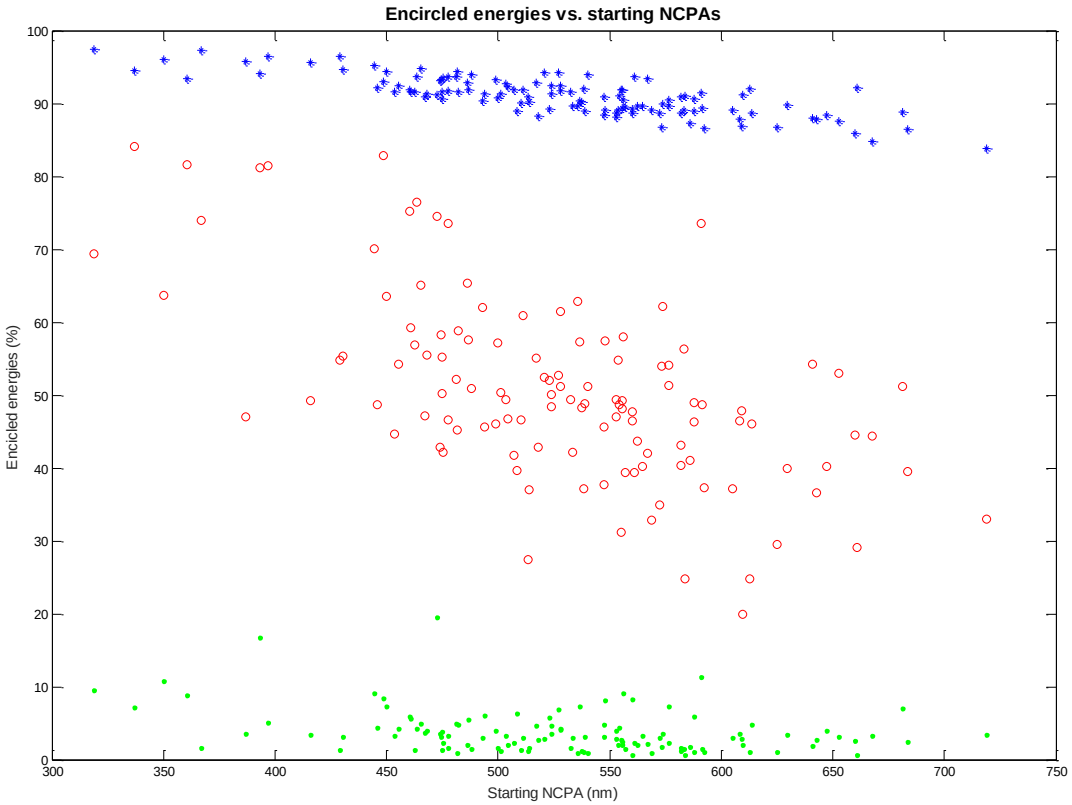

Figure 6. Summary of results of 120 simulations of the NWIWM algorithm. Considering the diffraction EE as 100%, blue asterisks represent the best possible correction with the simulated mirror (fitting error), green dots represent the EE low level caused by the simulated NCPA, measured before the start of the algorithm, and red circles represent the EE values after 80 iterations of NWIWM.

3.2 The random nature of optimization

There are a number of consequences of the random nature of the algorithm which can be found in the behaviour of NWIWM and its results. Figure 7 shows the simulation results of a ten-fold repetition of ten fixed NCPA values **¡Error! Marcador no definido.**optimization, selected along the line of linear regression of the fitting error results of previous figure. The simulations were done starting exactly from the same NCPA, and they ended up in a certain range of optimal positions, with some spread in the encircled energies obtained in each case. This spreading is understood to be generated by the reach of local minima and seem to be bigger for bigger NCPAs.

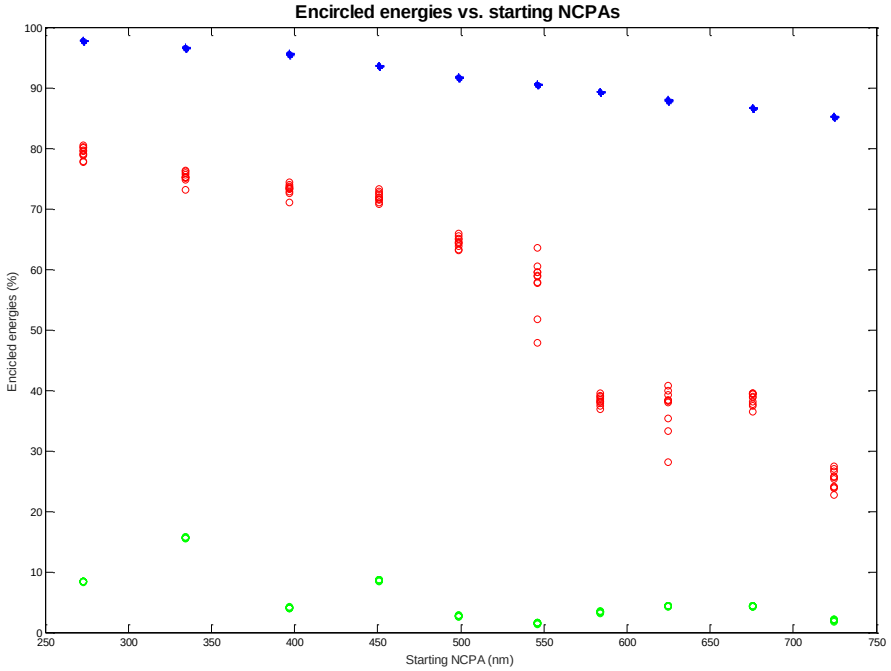

Figure 7. Results of the ten-fold repetition of ten selected NCPAs. Considering the diffraction EE as 100%, blue asterisks represent the best possible correction with the simulated mirror (fitting error), green dots represent the EE low level caused by the simulated NCPA, measured before the start of the algorithm, and red circles represent the EE values after 40 iterations of NWIWM in this case. The random nature of the algorithm can be noticed as the spread of the outputs, even with the same input, which is bigger for greater NCPAs.

4. LABORATORY RESULTS

The NWIWM algorithm has been used to date in two different adaptive optics projects at IAC: AOLI and EDIFISE. Both projects are demonstrating prototypes of specific techniques, like lucky imaging or equalized fiber-optic based integral field units.

4.1 NWIWM in AOLI

The Adaptive Optics Lucky Imager (AOLI)[6] instrument is designed to combine the techniques of AO and Lucky Imaging into a single instrument. It consists of a geometric curvature WFS and a low order adaptive optics wavefront corrector using a deformable mirror in conjunction with a wide-field, array detector Lucky Imaging camera. A calibration subsystem is also present in order to provide a reference flat wavefront and also to simulate atmospheric-like wavefront distortions for day-time system tuning.

AOLI is equipped with an ALPAO 241 actuators mirror which provides an excellent stroke capability, and for NWIWM tests a general purpose laboratory GigE camera was used, PHOTONIS Nocturn-GV, with 1280x1024 pixels of 9.7 x 9.7 microns. Results of the first iterations of the NWIWM algorithm are shown in figure 8, where a clear improvement on the mirror shape can be appreciated after just a few iterations.

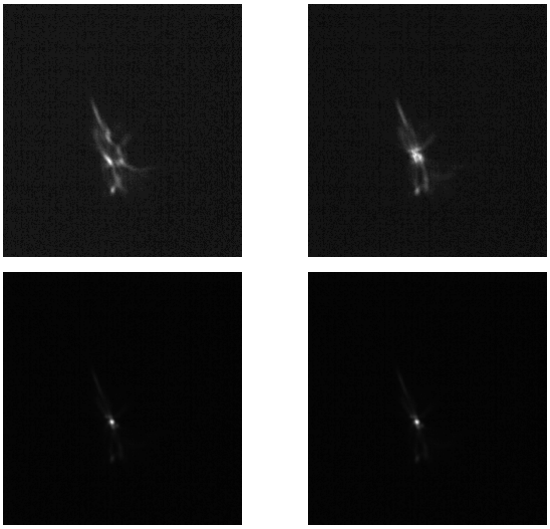

Figure 8. Results of only three iterations of the NWIWM algorithm in order to improve the shape of the ALPAO 241-actuators DM. Top left is a region of the GigE test camera used, with the mirror driver switched on and all actuators at zero current level. Top right is the image after the first iteration. Bottom left, after the second iteration. Bottom right, after the third iteration.

4.2 NWIWM in EDiFiSE

EDiFiSE (Equalized and Diffraction-limited Field Spectrograph Experiment) is a prototype featuring an adaptive optics subsystem, with both high and low order corrections, and a spectrograph with equalized integral field unit (EIFU)[7]. It is intended to be a test platform to verify the technology and its viability in bigger projects for large telescopes, where available information regarding the real-time statistics of the atmosphere is used for optimizing the configuration of the system. Both subsystems, AO and EIFU should then be designed concurrently to fulfil this purpose. The main scientific case aiming the technique is related to compact objects with high intensity contrast. The resolved detection of the spatial components will be used for improving both spatial and spectral resolution.

EDiFiSE is designed to have a separated tip-tilt correction stage, which is not being used by NWIWM, and relies on a 97-actuators ALPAO mirror for high order correction of the wavefront. The RTC control system is 100% FPGA based, in charge of processing 500 frames/second of 16x16 Shack-Hartmann wavefront sensor with minimum latency, but the DM is driven from a dedicated computer which allows direct commanding of the DM, and this capability is used for running the NWIWM algorithm. Science image is obtained, for these laboratory tests, from a JAI/PULNIX TM-6740CL, with 640x480 pixels of 7.4 x 7.4 microns each. A narrow-band I filter was used during the tests.

Figures 9 and 10 summarizes shows the test results of NWIWM. They are snapshots of the user interface at the first and the 25th iteration of the algorithm, and starting NCPA image is shown at the upper right image of figure 7. A magnification of the central part of this image depicted at the upper left corner, with the pixels being used for EE computation shown with a superimposed green grid. Gradients, signal to noise and EE are depicted in the upper row, and line search results, gradient measuring step, actuator values and images after iteration are shown at the bottom row. The improvement of the science image is clear, so it is the EE value achieved.

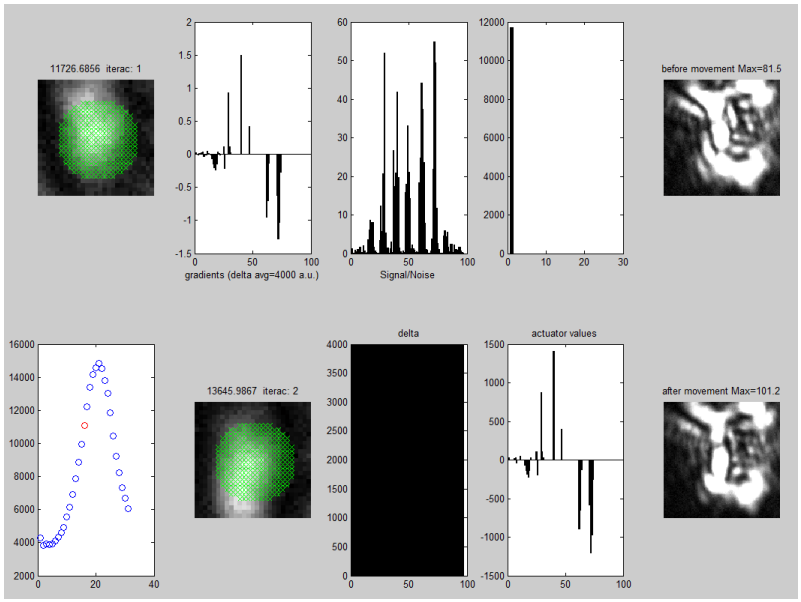

Figure 9. Snapshot of the user interface for laboratory tests of NWIWM in EDiFiSE. From left to right, at the upper row, the central part of the science image with indication in green of the pixels being used for EE computation, the measured gradients, its signal to noise ratio, the EE per iteration, and the (thresholded + amplified) science image. At the bottom row, the results of the line search (red circle shows the present EE), the central image after this iteration, the step to be used in the next gradient measurement, the actuator values after the iteration and the science image, also after the iteration.

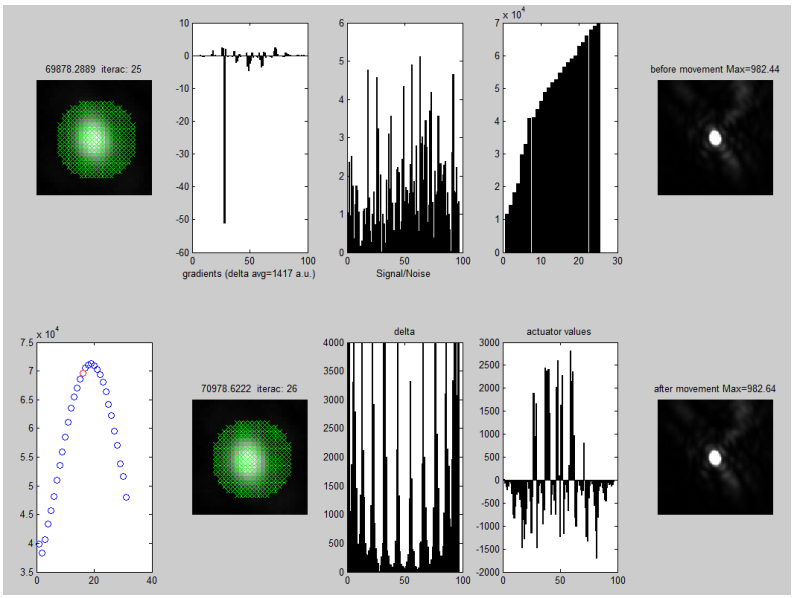

Figure 10. Snapshot of the user interface after 25 iterations of NWIWM in EDiFiSE. The improvement of the image is clear, and the EE has been increased by a factor of ~7. It can be noted that the steps used for gradient measurement, labelled as "delta", area now different for each the actuator, depending on the measured SNR.

5. NWIWM LIMITATIONS

It is fair to discuss here the main drawbacks found in the NWIWM algorithm. Probably the bigger one is related to the amount time required for doing all gradient measurements, and the risk that it could rise beyond practical limits. To this end, Table 1 summarizes the estimation of the overall time required for the use of the NWIWM algorithm, at a number of representative telescopes of different sizes.

The estimation is based in the need of 40 iterations, which has been found to be a reasonable amount of steps in most cases. Each gradient measurement has been supposed to be the result of averaging 9 samples, and 3 exposures has been selected to allow for complete DM stabilization after movements. A readout rate of 500 images per second of the science camera has been estimated, which might require the use of fast readout modes of small windows within the IR array detector. The results show that the execution time is well within practical limits for telescopes up to 10 m, and even at an ELT could eventually be used provided that the required couple of hours are available in daytime.

Table 1. Estimated execution time for NWIWM at different telescopes. See assumptions in text.

	OGS	WHT	GTC	ELT
size (m)	1	4,2	10	39
Estimated Number of actuators	97	241	373	5000
Exposures per iteration	2733	6189	9357	120405
Total (secs)	219	495	749	9632
Total (mins)	4	8	12	161

A second drawback of the algorithm has been already addressed in paragraph 3.2, where its random nature is shown to potentially allow the algorithm to get trapped in local minima at a certain distance of the achievable optimum. There are a number of solutions for this problem, like successive executions of the algorithm with a decreasing radius of the EE calculation region, together with the adaptation of the gradient measurement step to the expected distance to the optimum.

6. CONCLUSSIONS AND FUTURE WORK

The NWIWM algorithm has been demonstrated to be a simple tool for compensating the NCPA in an adaptive optics system, without the use of any sort of wavefront sensing. It could also be used for flattening mirrors in a blind way, based only on the overall resulting image. The behaviour of the algorithm has been thoroughly simulated and also has been tested in laboratory for two different adaptive optics instruments.

In the future, even more simulation work is likely to be required in order to verify the behaviour with a higher number of actuators, and to compare the results with other well stablished methods like Phase Diversity. Other more complicated stages could be added to the algorithm, like conjugate gradients, distance weighted energy, correlation, and many other possible improvements which, on the other hand, would supress the simplicity that underlies the present NWIWM.

A possible future use of NWIWM could also be the direct use of bright stars to evaluate the overall PSF of the telescope+instrument optical, making extensive use of lucky imaging techniques. The algorithm could be useful helping in the co-phasing of the primary, or compensating for polishing defects of big monolithic telescopes, the 4.2m WHT.

REFERENCES

[1] Gonsalves, R. A., “Phase retrieval and diversity in adaptive optics," Optical Engineering 21(5), 215829-215829 (1982).

[2] Sauvage, J.-F., Fusco, T., Rousset, G., and Petit, C., “Calibration and precompensation of noncommon path aberrations for extreme adaptive optics," Journal of the Optical Society of America A 24, 2334-2346 (Aug. 2007).
[3] R. Conan, C. Correia, "Object-oriented Matlab adaptive optics toolbox", Proc. SPIE 9148, Adaptive Optics Systems IV, 91486C (7 August 2014); doi: 10.1117/12.2054470; http://dx.doi.org/10.1117/12.2054470
[4] Noll, RJ. Zernike polynomials and atmospheric turbulence. J Opt Soc Am 66; 207-211 (1976).
[5] Clélia Robert, Thierry Fusco, Jean-François Sauvage, Laurent Mugnier, "Improvement of phase diversity algorithm for non-common path calibration in extreme AO context", Proc. SPIE 7015, Adaptive Optics Systems, 70156A (15 July 2008); doi: 10.1117/12.787870; http://dx.doi.org/10.1117/12.787870
[6] Craig Mackay et al. " AOLI: Adaptive Optics Lucky Imager: diffraction limited imaging in the visible on large ground-based telescopes ", Proc. SPIE 8446, Ground-based and Airborne Instrumentation for Astronomy IV, 844621 (September 24, 2012); doi:10.1117/12.925618.
[7] García-Lorenzo, B, et al. "EDiFiSE: equalized and diffraction-limited field spectrograph experiment", Proc. SPIE 7014, Ground-based and Airborne Instrumentation for Astronomy II, 70144B (July 11, 2008); doi:10.1117/12.787848;