

# Alternative Phase-Diverse Phase Retrieval algorithm based on Levenberg-Marquardt nonlinear optimization

Heng Mao\* and Dazun Zhao

Department of Photo-electronic Engineering, Beijing Institute of Technology,  
Beijing 100081, China

\*Corresponding author: [alexmaomao@bit.edu.cn](mailto:alexmaomao@bit.edu.cn)

**Abstract:** A modified Levenberg-Marquardt (MLM) algorithm is proposed to substitute for modified G-S (MGS) algorithm in some situations of phase-diverse phase retrieval wavefront sensing (WFS), such as the obstructed pupil, in which the second derivative information is specifically employed to eliminate the local minimum stagnation. Experiments have been performed to validate MLM algorithm in WFS accuracy (less than  $\lambda/30$  RMS) referring to ZYGO interferometer results and in WFS repeatability (less than  $\lambda/200$  RMS), even the dynamic range is more than  $7\lambda$  PV. Moreover, experiments have shown the MLM algorithm is superior to the MGS algorithm both in WFS accuracy and repeatability.

©2009 Optical Society of America

**OCIS codes:** (010.7350) Wave-front sensing; (100.5070) Phase retrieval; (120.3940) Metrology

---

## References and links

1. R. W. Gerchberg and W. O. Saxton, "A Practical Algorithm for the Determination of Phase from Image and Diffraction Plane Pictures," *OPTIK* **35**, 237-246 (1972).
2. B. H. Dean, D. L. Aronstein, J. S. Smith, R. Shiri, and D. S. Acton, "Phase Retrieval Algorithm for JWST Flight and Testbed Telescope," *Proc. SPIE* **6265**, 626511-1 (2006).
3. R. A. Gonsalves, "Phase Retrieval from Modulus Data," *J. Opt. Soc. Am.* **66**, 961-964 (1976).
4. R. A. Gonsalves, "Phase Retrieval and Diversity in Adaptive Optics," *Opt. Eng.* **21**, 829-832 (1982).
5. C. Roddier and F. Roddier, "Combined Approach to the Hubble Space Telescope Wave-Front Distortion Analysis," *Appl. Opt.* **32**, 2992-3008 (1993).
6. J. R. Fienup, "Phase-retrieval algorithms for a complicated optical system," *Appl. Opt.* **32**, 1737-1746 (1993).
7. D. C. Redding, P. Dumont, and J. Yu, "Hubble Space Telescope prescription retrieval," *Appl. Opt.* **32**, 1728-1736 (1993).
8. D. C. Redding, S. Basinger, A. Lowman, A. Kissil, P. Bely, R. Burg, R. Lyon, G. Mosier, M. Femiano, M. Wilson, G. Schunk, L. Craig, D. Jacobson, J. Rakoczy, and J. Hadaway, "Wavefront Sensing and Control for a Next Generation Space Telescope," *Proc. SPIE* **3356**, 758-772, (1998).
9. D. C. Ghiglia and M. D. Pritt, "Minimum-norm methods," in *Two-Dimensional Phase Unwrapping* (Wiley Press, 1998), pp. 178-278.
10. D. Cohen and D. C. Redding, "NGST High Dynamic Range Unwrapped Phase Estimation," *Proc. SPIE* **4850**, 336-344 (2003).
11. D. S. Acton, P. Atcheson, M. Cermak, L. Kingsbury, F. Shi, and D. C. Redding, "James Webb Space Telescope Wavefront Sensing and Control Algorithms," *Proc. SPIE* **5487**, 887-896 (2004).
12. C. M. Ohara, D. C. Redding, F. Shi, and J. J. Green, "PSF monitoring and in-focus wavefront control for NGST," *Proc. SPIE* **4850**, 416-427 (2003).
13. D. Malacara and S. DeVore, "Interferogram evaluation and wavefront fitting," in *Optical Shop Testing 2<sup>nd</sup> Ed.* (Wiley Press, 1992), pp. 455-487.
14. G. R. Brady and J. R. Fienup, "Nonlinear optimization algorithm for retrieving the full complex pupil function," *Opt. Express* **14**, 474-486 (2006).
15. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, "Nonlinear models," in *Numerical Recipes in C++: The Art of Scientific Computing 2<sup>nd</sup> Ed.* (Cambridge University Press, 2002), pp. 686-694.
16. J. J. Green, D. C. Redding, S. Shaklan, and S. A. Basinger, "Extreme wavefront sensing accuracy for the Eclipse Coronagraphic Space Telescope," *Proc. SPIE* **4860**, 266-276 (2003).
17. Heng Mao, Xiao Wang, and Dazun Zhao, "Application of phase-diverse phase retrieval to wavefront sensing in non-connected complicated pupil optics", *Chinese Opt. Lett.* **5**, 397-399 (2007).

## 1. Introduction

Phase retrieval (PR) is an inverse problem in optics, and generally involves estimating a complex-valued phase distribution from known intensity distributions at some properly selected planes. The original phase retrieval algorithm was proposed by Gerchberg and Saxton in 1972, known as Gerchberg-Saxton (G-S) algorithm [1]. It used the Fourier transform relationship between pupil and in-focus plane to iteratively estimate the pupil phase, and suffered from the non-uniqueness. The algorithm based on the forward and backward Fourier transform is classified as iterative transform algorithm (ITA) [2].

On the other hand, Gonsalves proposed another kind of phase retrieval algorithm [3,4], which was classified as model-based optimization algorithm (MBOA). Therein, the nonlinear optimization theory was introduced to estimate the model solution.

These two kinds of phase retrieval algorithms had been essentially developed in the early 1990s due to the measurement of misalignment aberrations on Hubble space telescope. C. Roddier and F. Roddier proposed a modified Misell algorithm to introduce multiple phase diversities [5]. Fienup derived the parametric model of complicated optical system depending on Fresnel diffraction, and selected gradient-search algorithm to implement MBOA [6]. Redding constructed the parametric model depending on ray tracing, and selected the Levenberg-Marquardt (L-M) algorithm to implement MBOA [7].

In the late 1990s, for seeking a preferred wavefront sensing (WFS) methodology on James Webb Space Telescope (JWST), JPL proposed modified G-S (MGS) algorithm [8], which was developed from modified Misell algorithm and used many defocus PSFs (point spread function) to guarantee the uniqueness. Later, collaborated with proper phase unwrapping algorithms [9], MGS algorithm can be competent for high dynamic aberrations [10]. Besides, JPL used MBOA to measure low-order aberrations of each segment on the primary mirror (PM) of JWST [11,12].

Generally, the parametric representation of wanted wavefront in MBOA includes point-by-point type and basis-function type [6,13]. Due to the facts that the basis-function terms are limited, and the convergence of MBOA would probably fall in the local minimum, MBOA possesses higher WFS dynamics and lower WFS accuracy comparing with MGS.

In this paper, based on the phase-diverse phase retrieval (PDPR) model, Zernike basis-functions and L-M optimization, the MBOA has been implemented in Section 2. Section 3 discusses the local minimum stagnation generated in L-M algorithm, and then a modified L-M algorithm for eliminating this issue is proposed. Section 4 presents some proof-of-principle experiments and Section 5 concludes the paper.

## 2. Model-based optimization algorithm

### 2.1 Phase-diverse phase retrieval model

Generally, phase-diverse phase retrieval has two different propagation models from pupil to focal plane: One locates the pupil directly at the lens plane; another locates the pupil at the front focal plane of lens to set up a telecentric system in image space.

As we known in Fourier Optics, the sampling spacing at different defocus planes is variant with the propagation distance in the first model, so some additive calculations are needed to keep invariant sampling spacing among all the defocus images [14]. For the second model, the sampling spacing satisfies the spacing requirement, but the dimension of this WFS system will be expanded (see in Fig. 1). In this paper, we select the second PDPR model to implement the MBOA and corresponding experiments.

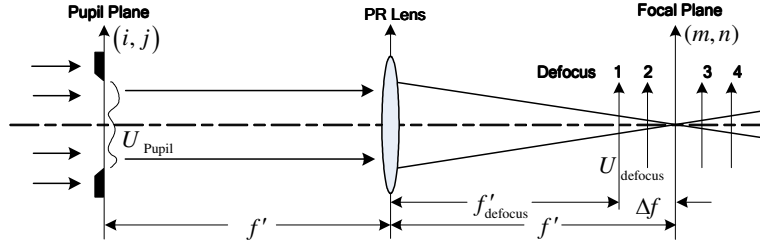


Fig. 1. Second PDPR propagation model.

## 2.2 Objective function of MBOA

As shown in Fig. 1, an optical field (in monochromatic) at the pupil plane,  $U_{\text{pupil}}(i, j)$ , propagates to the focal plane by experiencing two Fresnel diffractions and one phase modulation of PR lens.  $U_{\text{focus}}(m, n)$  is the field at the focal plane,  $U_{\text{defocus}}(m, n)$  is the field at the defocus plane. The first  $S$  terms of Zernike polynomials [13] are adopted to describe the phase distribution of the pupil,  $\phi_{\text{pupil}}(i, j)$ , and the pupil field can be expressed as

$$U_{\text{pupil}}(i, j) = \sqrt{I(i, j)} \cdot \exp[i \cdot \phi_{\text{pupil}}(i, j, \vec{\alpha})] = \sqrt{I(i, j)} \cdot \exp\left[i \cdot \left(\sum_{u=1}^S \alpha_u \cdot Z_u(i, j)\right)\right], \quad (1)$$

where  $i^2 = -1$ ,  $I(i, j)$  is the known intensity of the pupil field,  $Z_u(i, j)$  is the  $u$ -th Zernike polynomial,  $\alpha_u$  is the  $u$ -th Zernike coefficient, and  $\vec{\alpha}$  is a vector constituted of  $S$  coefficients.

According to the phase-diverse theory, the defocus field with a defocus of  $\Delta f$  is

$$U_{\text{defocus}}(m, n, \Delta f, \vec{\alpha}) = \text{FFT}\left\{U_{\text{pupil}}(i, j, \vec{\alpha}) \cdot \exp[i \cdot \phi_{\text{defocus}}(i, j, \Delta f)]\right\}, \quad (2)$$

where  $\phi_{\text{defocus}}(i, j, \Delta f)$  is the phase diversity for  $\Delta f$  defocus, and  $\text{FFT}\{\cdot\}$  denotes the fast fourier transform operator. By abbreviating  $\sqrt{I(i, j)}$ ,  $\phi_{\text{pupil}}(i, j, \vec{\alpha})$  and  $\phi_{\text{defocus}}(i, j, \Delta f)$  to  $W$ ,  $\phi_{\text{pu}}$  and  $\phi_{\text{de}}$ , the intensity at  $\Delta f$  defocus can be calculated as

$$I_{\text{de}}^{\text{cal}}(m, n, \Delta f, \vec{\alpha}) = \text{FFT}\left\{W \cdot \exp[i \cdot (\phi_{\text{pu}} + \phi_{\text{de}})]\right\} \cdot \text{FFT}^*\left\{W \cdot \exp[i \cdot (\phi_{\text{pu}} + \phi_{\text{de}})]\right\}, \quad (3)$$

where the superscript  $*$  denotes the complex conjugate operator.

The objective function of this MBOA is defined as

$$F(\vec{\alpha}) = \sum_{m=1}^M \sum_{n=1}^N h^2(m, n, \Delta f, \vec{\alpha}) = \sum_{m=1}^M \sum_{n=1}^N \left[ \hat{I}_{\text{de}}^{\text{sen}}(m, n, \Delta f) - \hat{I}_{\text{de}}^{\text{cal}}(m, n, \Delta f, \vec{\alpha}) \right]^2, \quad (4)$$

where  $\hat{I}_{\text{de}}^{\text{sen}}$  and  $\hat{I}_{\text{de}}^{\text{cal}}$  are normalized CCD sensed intensity  $I_{\text{de}}^{\text{sen}}$  and  $I_{\text{de}}^{\text{cal}}$ , respectively.

Due to the analytic expression of  $F(\vec{\alpha})$  with respect to  $\vec{\alpha}$  is obtainable, we select the Levenberg-Marquardt algorithm to implement the optimization.

## 2.3 L-M algorithm in MBOA

By rearranging  $h(m, n, \Delta f, \vec{\alpha})$  into an one-dimensional sequence  $h(r, \Delta f, \vec{\alpha})$ , where  $r=1, 2, 3, \dots, M \times N$ , the first derivative of  $h(r, \Delta f, \vec{\alpha})$  with respect to  $\alpha_u$  can be written as

$$\frac{\partial h(r, \Delta f, \vec{\alpha})}{\partial \alpha_u} = - \frac{\left[ \text{FFT}\{U_{\text{total}} \cdot i \cdot Z_u\} \cdot \text{FFT}^*\{U_{\text{total}}\} + \text{FFT}\{U_{\text{total}}\} \cdot \text{FFT}^*\{U_{\text{total}} \cdot i \cdot Z_u\} \right]}{\text{Max}\left[I_{\text{de}}^{\text{cal}}(r, \Delta f, \vec{\alpha})\right]}, \quad (5)$$

where  $U_{\text{total}} = W \cdot \exp\left[i \cdot (\phi_{\text{pu}} + \phi_{\text{de}})\right]$ , and  $\text{Max}\left[I_{\text{de}}^{\text{cal}}(r, \Delta f, \vec{\alpha})\right]$  is the maximum value over all  $I_{\text{de}}^{\text{cal}}(r, \Delta f, \vec{\alpha})$  data.

Depending on Eq. (5), we can optimize the estimate  $\vec{\alpha}$  by L-M iterative formula

$$\vec{\alpha}^{(k+1)} - \vec{\alpha}^{(k)} = - \left[ J^T(\vec{\alpha}^{(k)}) \cdot J(\vec{\alpha}^{(k)}) + \lambda^{(k)} \cdot \mathbf{I} \right]^{-1} \cdot J^T(\vec{\alpha}^{(k)}) \cdot H(\vec{\alpha}^{(k)}), \quad (6)$$

where  $H = \left[ h(r, \Delta f, \vec{\alpha}) \right]_{(M \times N) \times 1}$ ,  $J = \left[ \frac{\partial h(r, \Delta f, \vec{\alpha})}{\partial \alpha_u} \right]_{(M \times N) \times S}$ ,  $\lambda^{(k)}$  is an adjustable factor at the  $k$ -th

iteration,  $\mathbf{I}$  is an identity matrix,  $[\cdot]^T$  and  $[\cdot]^{-1}$  respectively denotes transpose and inverse operation [15].

Hence, by iteratively making a joint estimate from each defocus L-M estimate, this MBOA is finally implemented. However, simulation indicates that L-M algorithm probably suffers from the local minimum stagnation, which stops the convergence approaching to the truth. In Section 3, a modified L-M algorithm will be proposed to eliminate this issue.

### 3. Modified Levenberg-Marquardt algorithm

#### 3.1 Local minimum stagnation issue

According to PDPR model and Eq. (6), the local minimum stagnation in L-M algorithm can be ascribed to the following reasons: the defocus PSF is usually concentrated in paraxial region, and this enhances the sparsity and singularity of the first derivative matrix  $J$ . Besides, high dynamic range of pupil aberrations extends the residual distance between  $\vec{\alpha}^{(k)}$  and the truth. Thus, it's hard to calculate the steepest descent direction at the fore-phase of L-M iteration, when the nonlinearity degree of  $\vec{\alpha}$  probably dominates the objective function  $F(\vec{\alpha})$ .

In nonlinear optimization theory, L-M algorithm is a simplified version of the Newton iterative formula

$$\vec{\alpha}^{(k+1)} - \vec{\alpha}^{(k)} = - \left[ J^T(\vec{\alpha}^{(k)}) \cdot J(\vec{\alpha}^{(k)}) + S(\vec{\alpha}^{(k)}) \right]^{-1} \cdot J^T(\vec{\alpha}^{(k)}) \cdot H(\vec{\alpha}^{(k)}), \quad (7)$$

where  $S = \left[ \sum_{r=1}^{M \times N} h(r, \Delta f, \vec{\alpha}) \cdot \frac{\partial^2 h(r, \Delta f, \vec{\alpha})}{\partial \alpha_u \partial \alpha_v} \right]_{S \times S}$  involves the second derivative of  $h(r, \Delta f, \vec{\alpha})$ .

Due to  $S(\vec{\alpha}^{(k)})$  is quite complicated in expression and calculation, it is discarded in L-M algorithm and substituted for the fudge term  $[\lambda^{(k)} \cdot \mathbf{I}]$ . However, this simplification significantly shrinks the adaptability of L-M algorithm to all nonlinearity dominated cases. Anyway, making use of the second derivative information would be better to eliminate the local minimum stagnation.

Fortunately, the second derivative of  $h(r, \Delta f, \vec{\alpha})$  in PDPR model can be derived as

$$\frac{\partial^2 h_r(\vec{\alpha})}{\partial \alpha_u \partial \alpha_v} = \frac{\text{FFT}\{U_{\text{total}} \cdot Z_u \cdot Z_v\} \cdot \text{FFT}^*\{U_{\text{total}}\} + \text{FFT}\{U_{\text{total}}\} \cdot \text{FFT}^*\{U_{\text{total}} \cdot Z_u \cdot Z_v\}}{\text{Max}[I_{\text{de}}^{\text{cal}}(r, \Delta f, \vec{\alpha})]} - \frac{\text{FFT}\{U_{\text{total}} \cdot i \cdot Z_u\} \cdot \text{FFT}^*\{U_{\text{total}} \cdot i \cdot Z_v\} + \text{FFT}\{U_{\text{total}} \cdot i \cdot Z_v\} \cdot \text{FFT}^*\{U_{\text{total}} \cdot i \cdot Z_u\}}{\text{Max}[I_{\text{de}}^{\text{cal}}(r, \Delta f, \vec{\alpha})]}. \quad (8)$$

### 3.2 Modified L-M algorithm and implementation details

Here, we propose a modified L-M algorithm (MLM) by hybridly using L-M iterative formula Eq. (6) and Newton iterative formula Eq. (7). The algorithm flow (Fig. 2) includes three inner-loops in parallel for three participant defocus PSFs and a global outer-loop for joint estimation. For every round outer-loop, each defocus PSF firstly completes P times L-M iteration, then these estimates  $\vec{\alpha}_{1,2,3}^{(P)}$  individually experience one time Newton iteration to give further estimates  $\vec{\alpha}_{1,2,3}^{(P+1)}$ , then three ultimate estimates  $\vec{\alpha}_{1,2,3}^*$  are selected among  $\vec{\alpha}_{1,2,3}^{(P)}$  and  $\vec{\alpha}_{1,2,3}^{(P+1)}$  depending on multi-metrics, finally  $\vec{\alpha}_{1,2,3}^*$  are combined in distributed weightings to make a joint estimate which is served as initial value  $\vec{\alpha}^{(0)}$  in next round outer-loop. After several rounds, the joint estimate may satisfy the criterion, which presents the solution similarity in last abutted outer-loops.

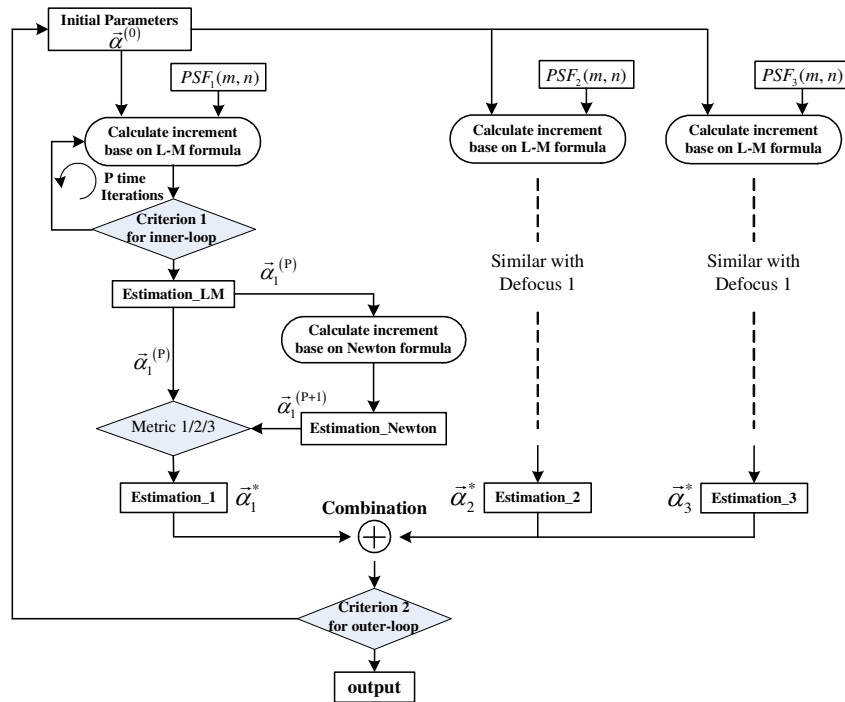


Fig. 2. Flowchart of modified L-M algorithm.

Meanwhile, some modification details are carried out:

(1) As discussed in Section 3.1, defocus PSF is over-concentrated and the data in periphery are susceptible to the CCD noise and stray light. So, the rearrangement of  $h(m, n, \Delta f, \vec{\alpha})$  follows the ranking of  $I_{\text{de}}^{\text{sen}}(m, n, \Delta f)$ , and a gray threshold is set to discard all

dark points participating in formula calculation. In this way, the size of the sequence  $h(r, \Delta f, \vec{\alpha})$  is much less than  $M \times N$ .

(2) Although the calculation by Newton iterative formula can effectively pull the convergence out of the local minimum and accelerate the convergent speed, the one time Newton estimation  $\vec{\alpha}^{(P+1)}$  is not always better than  $\vec{\alpha}^{(P)}$ . Therefore, multi-metrics are set as follows for optimal selection among  $\vec{\alpha}^{(P)}$  and  $\vec{\alpha}^{(P+1)}$ :

Metric 1 is the objective function value about each defocus optimization, and other two metrics are established on a reference defocus PSF. This reference defocus is different from any of the above three defoci, and its PSF should be known. Metric 2 is the objective function value at reference defocus. Metric 3 calculates the normalized cross correlation matrix in  $5 \times 5$  central window between  $\hat{I}_{de}^{sen}(m, n, \Delta f_{refer})$  and  $\hat{I}_{de}^{cal}(m, n, \Delta f_{refer})$  at reference defocus and then sums up the matrix. In this way, the centering errors (sub-pixel part) could be ignored.

As evaluating metrics, the SSE (sum of the squares of errors, such as metric 1 and metric 2) and Correlation (metric 3) respectively represent the similarity between estimated PSF and actual PSF in different respects. As long as the multi-metrics are fully satisfied,  $\vec{\alpha}^{(P+1)}$  is treated as the better estimate.

(3) It is considerate to use two defocus pairs involved. Therein, three of them are used for parametric optimization, and the rest one is used as reference defocus. Moreover, distributing weightings to each estimate is crucial for eliminating the local minimum stagnation.

At fore-times of outer-loop, considering the convergence at each defocus optimization is probably not well, the similarity in PSF configuration (metric 3) is the most reasonable criterion to distribute weightings. And in the former step (2) the values of metric 3 with respect to  $\vec{\alpha}_{1,2,3}^*$  have been already calculated. Thus, using an adjustable positive exponential  $\gamma$ , these three weightings can be determined as

$$Weighting(i) = [Metric\ 3(\vec{\alpha}_i^*)]^\gamma / \sum_{i=1}^3 [Metric\ 3(\vec{\alpha}_i^*)]^\gamma, \quad (9)$$

where  $i$  is the sequential number among three defocus planes.

### 3.3 Simulations

In simulation, by using the first 21 terms Zernike polynomials (excepting piston mode), the actual phase distribution at the pupil plane could be assumed. Then, the corresponding actual defocus PSFs were calculated and digitized, and the maximum gray-level of each PSF was shrunk to 3200. Random Gaussian noise within 8 gray-level standard deviation was added into each PSF.

For every different trial, the WFS errors (WFSE) of MLM algorithm (estimating 20 Zernike coefficients from 2<sup>nd</sup> to 21<sup>st</sup> terms) are remarkably smaller than the WFSE of MGS algorithm. For example, as PV and RMS of the actual phase assumption in one trial are 37 radians ( $5.88\lambda$ ) and 7.3 radians ( $1.16\lambda$ ), the WFSE of MLM and MGS are  $\lambda/1290$  and  $\lambda/46$ , respectively. It has shown that MLM algorithm is less susceptible to the CCD noise than MGS algorithm when its representation of basis-function about the actual phase distribution is full.

## 4. Experimental setup and results

The experimental setup for MLM algorithm, as well as MGS algorithm, based on the second PDPR model is shown in Fig. 3. The light source was a monochromatic LD (@OZ, 670nm) which involves CW mode and TTL mode. The point source was collimated by lens  $L_1$ , and the light went through the phase plate, which is mounted on a pupil module (see in Fig. 4), and lens  $L_2$  to image on CCD chip. This scientific grade CCD camera (@ANDOR, back-illuminated, 16 bit,  $13\mu m$  pixel size, readout noise  $2.5 \sim 10e^- @ 50kHz \sim 2.5MHz$ ) was

mounted upon the motorized translator (@SIGMA, 85mm) to complete the defocus movements, and it also accurately modulated the LD fire under its timing. Moreover, for recording the actual pupil intensity, lens  $L_3$  was inserted to re-image pupil near the focus of  $L_2$ , and it can be accurately repositioned by another motorized translator (same as the former one).

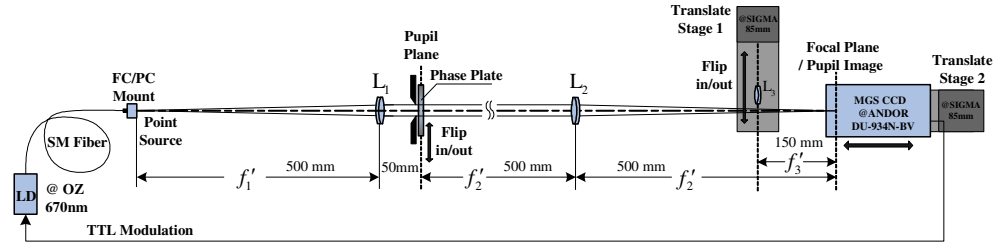


Fig. 3. Layout of experimental setup.

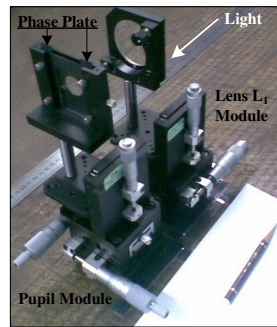


Fig. 4. Pupil module and Lens  $L_1$  module.

#### 4.1 Experimental steps and settings

At first, the systematic errors of this optics were calibrated as the initial aberrations at pupil plane by using MLM or MGS algorithm. Then, one phase-plate sample was inserted and the whole aberrations (including initial part and sample part) can be retrieved by using MLM or MGS algorithm. Finally, we took the pupil module off and tested the sample in ZYGO interferometer.

There are two pupil apertures discussed in this experiment, one was the circular aperture and another was within obstructions. For each case, we completed both MLM and MGS WFS, and respectively calculated nominal WFSE by comparing with the ZYGO result (as reference). Therein, all the phasemaps participated in comparison didn't include the piston and tip/tilt mode.

In this experiment, nominal focal length was 500mm, pupil diameter was 20mm, and sampling matrix was set as  $512 \times 512$  for MGS and  $256 \times 256$  for MLM, so the actual sampling points of pupil are respectively 398 and 199 in MGS and MLM algorithm. Two defocus pairs were employed both in MGS and MLM WFS, and the Zernike polynomials estimated in MLM were the first 21 terms excepting piston mode.

#### 4.2 Systematic errors calibration

As shown in Fig. 5, a WFS GUI drove the calibration of systematic errors in MGS WFS. The GUI was compatible for both MGS and MLM algorithm in many different aperture cases. In detail, in the bottom panel of GUI, there were four subfigures and two group data on the right.

From left to right, these subfigures presented systematic errors, actual phasemap, retrieved phasemap and WFS errors (all in radians). There are many different representations of actual phasemap to compare with the retrieved result, such as alternative algorithm estimates, ZYGO results and even same algorithm estimates from different measurement sequence. Those two group data displayed the individual PV and RMS value of actual phasemap and WFS errors, and nearly 20 points of pupil edge were cut off in calculating WFS errors.

Fig. 5 indicates the PV and RMS of MGS estimate is  $0.7041\lambda$  and  $0.1171\lambda$ . Then by 20 terms Zernike decomposition, as listed in Table 1, we know the systematic errors are dominated in 5<sup>th</sup> mode (defocus), 4<sup>th</sup> and 6<sup>th</sup> modes (3<sup>rd</sup> order astigmatism), 8<sup>th</sup> and 9<sup>th</sup> modes (3<sup>rd</sup> order coma) and other higher modes.

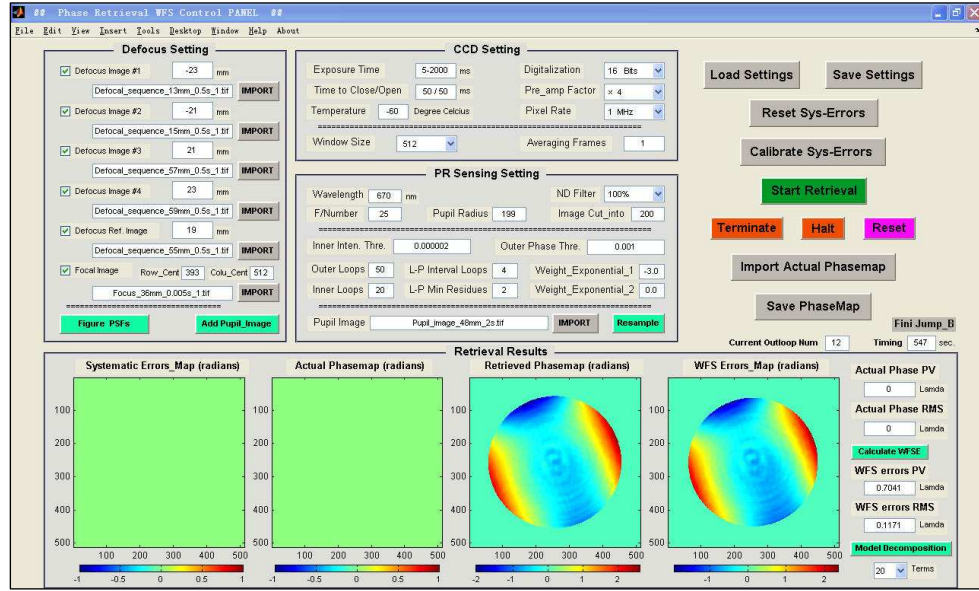


Fig. 5. GUI in systematic errors calibration.

Table 1. Zernikes decomposed from MLM estimate and MGS estimate (in radians).

Zernike Modes	1	2	3	4	5	6	7
MLM result	0	removed	removed	<b>1.4840</b>	<b>0.5570</b>	<b>-1.4513</b>	0.0752
MGS result	0	removed	removed	<b>0.9245</b>	<b>0.4515</b>	<b>-2.1944</b>	0.0902
Zernike Modes	8	9	10	11	12	13	14
MLM result	<b>-0.0394</b>	<b>0.0617</b>	-0.0687	0.0835	<b>-0.0153</b>	-0.0624	<b>0.0024</b>
MGS result	<b>0.1810</b>	<b>-0.1779</b>	-0.1134	0.1087	<b>0.1070</b>	-0.0446	<b>-0.5835</b>
Zernike Modes	15	16	17	18	19	20	21
MLM result	0.0400	0.0119	0.0146	0.0158	-0.0188	0.0162	-0.0038
MGS result	0.0403	0.0138	-0.0059	-0.0207	-0.0009	0.0428	0.0077

The 4<sup>th</sup>, 5<sup>th</sup>, 6<sup>th</sup>, 8<sup>th</sup>, 9<sup>th</sup>, 12<sup>th</sup>, 14<sup>th</sup> modes are bolded as their differences are larger than 0.1 radians.



Otherwise, MLM WFS was done for sensing the systematic errors, and the 20 estimated coefficients are also listed in Table 1.

As discussed in [16], the PSF centering errors in practice translate into the WFSE of MGS. Concretely, the mismatch between the true optical axis and the PDPR model causes the MGS algorithm to develop false astigmatism, coma and other higher order errors in its estimate. This explains the obvious differences between MLM and MGS estimate at certain modes in Table 1. The differences between MLM and MGS estimate (as actual phasemap) are shown in Fig. 6, while the PV and RMS of differences are  $0.2759\lambda$  and  $0.0575\lambda$ .

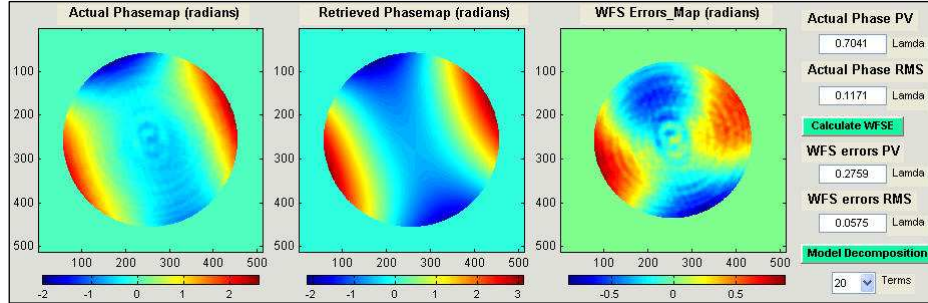


Fig. 6. Systematic errors calibration: MLM estimate vs MGS estimate (actual map).

#### 4.3 Circular aperture case

After inserting sample 1, all WFS experiments were done and repeated 10 times at different hours lasting two days. The comparison between one MLM estimate and one MGS estimate (as actual phasemap) for sensing whole aberrations is recorded in Fig. 7. Here, the PV of MLM estimate and MGS estimate are both nearly  $7.0\lambda$  (including tip/tilt mode), the PV of differences between two estimates is  $0.1561\lambda$ , and the RMS of differences is  $0.0256\lambda$ .

Later, statistic calculation is given to demonstrate the repeatability of MLM over 10 estimates, and the RMS of differences between each estimate and mean estimate is individually less than  $0.005\lambda$ .

Otherwise, the raw ZYGO result experienced the realignment processing, including centering, interpolation, wavelength conversion and rotation compensation. Here, by using edge mask, the rotation angle is calibrated as  $-1.05$  degree. Thus, the nominal WFSE between one MLM estimate and ZYGO result (as actual phasemap) is recorded in Fig. 8, while the PV and RMS of differences are  $0.1377\lambda$  and  $0.024\lambda$ . The nominal WFSE between one MGS estimate and ZYGO result (as actual phasemap) is recorded in Fig. 9, while the PV and RMS of differences are  $0.2313\lambda$  and  $0.0461\lambda$ . Meanwhile, all above WFS results are summarized in Table 2.

Being consistent with results from other samples, it has shown that the WFSE created by mismatch errors in MGS is reduced (see in Fig. 7) when the sample phase distribution is larger (in dynamic range) and higher (in spatial frequency) than the systematic errors at pupil plane.

Moreover, the WFSE of MLM presents many local errors and few artifacts (see in Fig. 8), which are suspected to be the residual modes and certain offspring of ZYGO sensing. The WFSE of MGS is quite larger than the WFSE of MLM (see in Fig. 9), and the most WFSE belong to the 3<sup>rd</sup> order astigmatism and 3<sup>rd</sup> order coma modes which are related to the estimate of systematic errors in MGS WFS.

In all above MLM iterations, the consumed times of outer-loop are all less than 10 times, and in most of outer-loop rounds, at least one estimate from  $\vec{\alpha}_{1,2,3}^{(P+1)}$  is selected. Besides, it has demonstrated that adoption of using PSFs on two pairs of defocus planes can effectively prevent the local minimum stagnation at the fore-phase of outer-loop and ultimately enhance the guarantee of the uniqueness of the retrieved phase distribution.

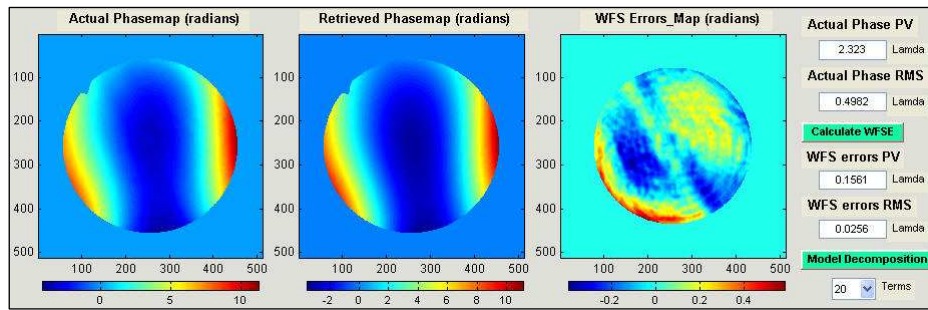


Fig. 7. Sample 1: MLM estimate vs MGS estimate (actual map) for whole aberrations.

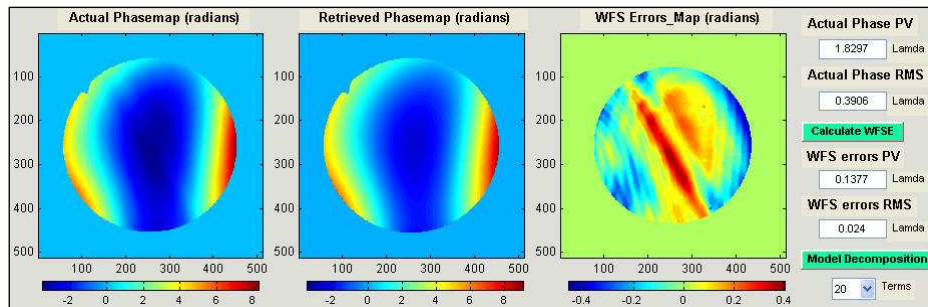


Fig. 8. Sample 1: MLM estimate vs ZYGO result (actual map).

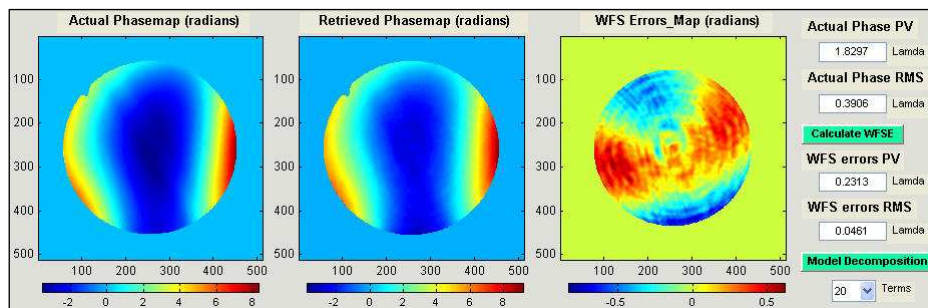


Fig. 9. Sample 1: MGS estimate vs ZYGO result (actual map).

#### 4.4 Obstructed aperture case

WFS in non-connected pupil, especially for high dynamic aberrations, the complexity of MGS algorithm extremely proliferates in phase-unwrapping processing [17]. As known, the non-connected pupil doesn't mean any trouble to MBOA. However, the WFS accuracy of most MBOA generally doesn't achieve the level of MGS because of local minimum stagnation. In this paper, the MLM algorithm is proposed to substitute for MGS in obstructed aperture case.

In this WFS situation, replacing with sample 2, we first completed 10 times WFS experiments in circular aperture, then fixed a pupil mask and completed 10 times WFS experiments in obstructed aperture. These corresponding WFS results are summarized in Table 2.

In one of obstructed experiments, all the related images in MLM WFS are presented in Fig. 10. Subfigure (a) is the raw pupil image; (e) is the resampling pupil window (0-1)

function; (b), (c), (d) and (i) are four sensed defocus PSFs (in  $\pm 21\text{mm}$  and  $\pm 25\text{mm}$ ): (b)~(d) are used for optimization, (i) is used as reference defocus PSF; (f), (g) and (h) are the ultimate PSF estimations of defocus (b), (c) and (d); (j), (k) and (l) are the ultimate PSF estimations of reference defocus respectively from three defocus ultimate estimates.

Regarding all the WFS results of this experiment, the comparison between MLM estimate and MGS estimate (as actual phasemap) for sensing whole aberrations is recorded in Fig. 11. Here, the PV of MLM estimate and MGS estimate are both nearly  $7.2\lambda$  (including tip/tilt modes), and the PV and RMS of differences are  $0.1675\lambda$  and  $0.0269\lambda$ . And the statistic calculation showed the uniform repeatability of MLM as calculated in Section 4.3.

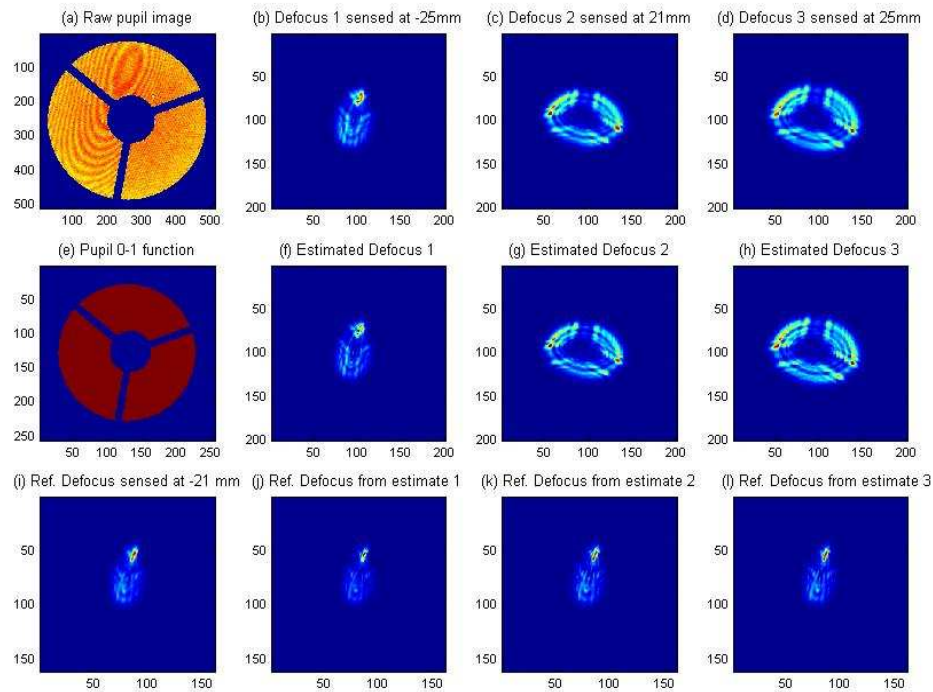


Fig. 10. Sample 2: all related images in MLM WFS.

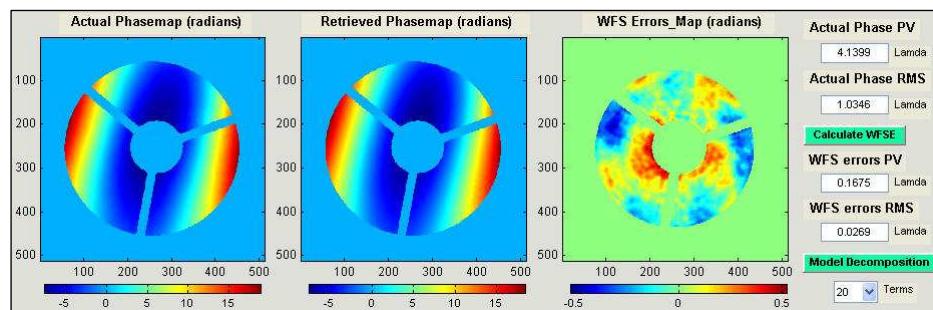


Fig. 11. Sample 2: MLM estimate vs MGS estimate (actual map) for whole aberrations.

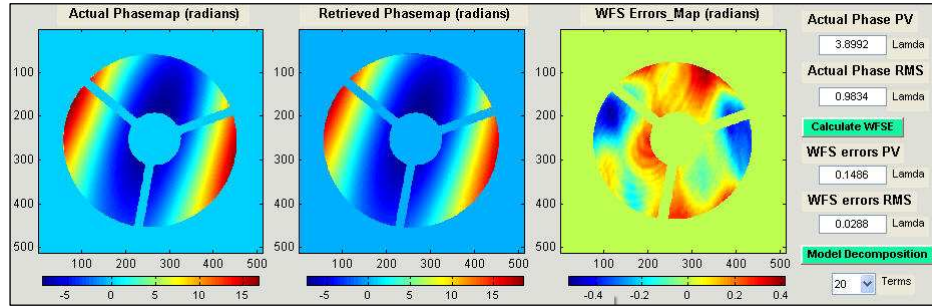


Fig. 12. Sample 2: MLM estimate vs ZYGO result (actual map).

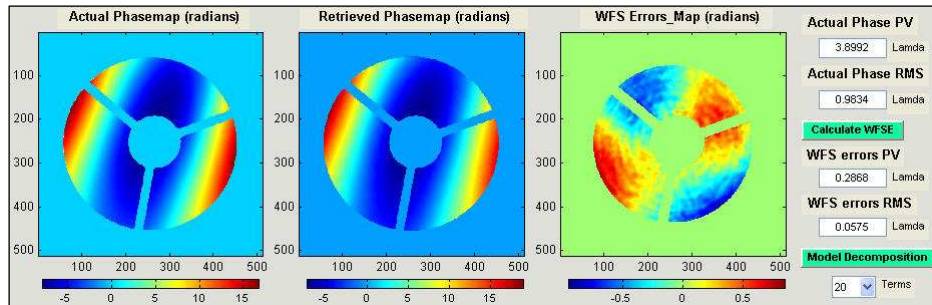


Fig. 13. Sample 2: MGS estimate vs ZYGO result (actual map).

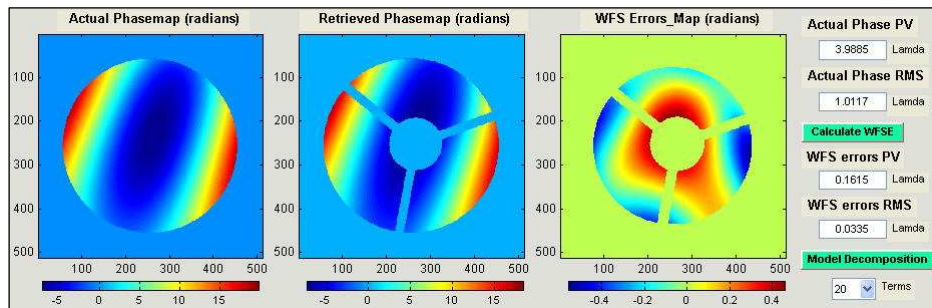


Fig. 14. Sample 2: obstructed MLM estimate vs circular MLM estimate (actual map).

Otherwise, after  $-1.50$  degree rotation compensation, the nominal WFSE between one MLM estimate and ZYGO result is recorded in Fig. 12, while the PV and RMS of differences are  $0.1486\lambda$  and  $0.0288\lambda$ . And the nominal WFSE between one MGS estimate and ZYGO result is recorded in Fig. 13, while the PV and RMS of differences are  $0.2868\lambda$  and  $0.0575\lambda$ .

Besides, the comparison between one MLM estimate in obstructed pupil and one MLM estimate in circular pupil (as actual phasemap) is recorded in Fig. 14, while the PV and RMS of differences are  $0.1615\lambda$  and  $0.0335\lambda$ , which illustrate the effects of pupil obstruction in MLM WFS.

Same experiments were done for other 4 samples. Due to the dynamic range (PV and RMS) of those aberrations are lower than the dynamic range of sample 2, the WFS accuracy and repeatability of those MLM estimates are totally superior to the WFS accuracy and repeatability of sample 2.

#### 4.5. Discussion

Some issues concerning MLM and MGS algorithm from experimental results are discussed here.



(1) As analyzed in Section 3, using Newton iterative formula is a trade off. The validity of MLM algorithm has been demonstrated in simulations and proof-of-principle experiments, which provide a good example to reasonably adopt the Newton iterative formula.

(2) Cross correlation between estimated PSF and actual PSF at the reference defocus, has been demonstrated as the sensitive metric to evaluate estimates and distribute weightings, especially at the fore-phase of outer-loop.

(3) It has shown that the MGS WFS suffers from the undesired influence of PSF centering errors, intensively in calibration of the systematic errors at pupil plane. For resolving this issue, an additional processing named as Image Centering Preprocessing is embedded in the flow of MGS iteration [2,16].

However, the MLM WFS has been demonstrated to be not involved in this issue. And comparing with interferometric result, the WFSE of MLM mainly includes any higher order modes and local errors which are absolutely out of its model representation.

Table 2. Summarization of all WFS results in sample 1 and 2 (in waves).

WFS Results ( $\lambda=670\text{nm}$ )	Sample 1	Sample 2	Sample 2
	circular	circular	obstructed
Estimate PV (with tip/tilt)	6.88 <sup>a</sup>	7.02 <sup>a</sup>	7.26 <sup>a</sup>
Estimate RMS (with tip/tilt)	1.67 <sup>a</sup>	1.65 <sup>a</sup>	1.75 <sup>a</sup>
ZYGO data rotation	-1.05 degree	-1.50 degree	-1.50 degree
MLM - MGS (PV/RMS)	0.1561 / 0.0256	0.1831 / 0.0327	0.1675 / 0.0269
MLM WFSE (PV/RMS)	0.1377 / 0.024	0.1583 / 0.0303	0.1486 / 0.0288
MLM repeatability (RMS)	< 0.005	< 0.005	< 0.005
MGS WFSE (PV/RMS)	0.2313 / 0.0461	0.3012 / 0.0597	0.2868 / 0.0575
MGS repeatability (RMS)	< 0.008	< 0.008	< 0.008
Circular MLM– Obstructed MLM (PV/RMS)	0.1615 / 0.0335		

<sup>a</sup>Mean value between MLM estimate and MGS estimate.

## 5. Conclusion

Here, we propose a modified L-M algorithm to substitute for MGS algorithm being an alternative PDPR algorithm for WFS the obstructed pupil. Proof-of-principle experiments in Section 4 have validated this proposed algorithm in WFS accuracy (less than  $\lambda/30$  RMS including residual modes) referring to the ZYGO interferometric results and in WFS repeatability (less than  $\lambda/200$  RMS), even for sensing high dynamic aberrations (up to  $7\lambda$  PV including tip/tilt mode). Moreover, these experiments have shown that MLM algorithm is superior to MGS algorithm both in WFS accuracy and repeatability, as illustrated in Table 2.

Considering the limited spatial frequency components of phase samples used in above experiments, further work is planning to validate MLM algorithm for WFS some different samples, which are dominated in moderate and/or even higher order modes, both by simulation and experiment. Besides, the calibration of the WFS accuracy of PDPR algorithms has been considered as a complicated and under-resolved problem and we need to more exactly extract the portion which truly belongs to the WFS results from their estimates [18].