

Potential of phase-diversity for metrology of active instruments

Visa Korkiakoski^a, Lars Venema^b, Tibor Agocs^b, Christoph U. Keller^a, Niek Doelman^c, Rufus Fraanje^d, Raluca Andrei^d and Michel Verhaegen^d

^aLeiden Observatory, Niels Bohrweg 2, 2333CA Leiden, The Netherlands;

^bASTRON, P.O. Box 2, 7990AA Dwingeloo, The Netherlands;

^cTNO Science and Industry, Stieltjesweg 1, 2628CK Delft, The Netherlands;

^dDelft Center for Systems and Control, Mekelweg 2, 2628CD Delft, The Netherlands

ABSTRACT

We investigate the potential of phase-diversity (PD) and Gerchberg-Saxton (GS) algorithms in the calibration of active instruments. A set of images is recorded with the focal-plane scientific camera, each image having a known and unique defocus. The phase-retrieval algorithms are used, with those images, to estimate the non-common path aberration that needs to be compensated by correct alignment of the instrument. We demonstrate by numerical simulations that such algorithms, in particular GS, are sufficient detection methods to fully correct wavefronts with an rms error at least up to 6 rad — but this requires several iterative correction stages.

Keywords: phase retrieval, phase diversity, active instruments

1. INTRODUCTION

Modern astronomical instruments are becoming more and more complex, and they tend to include an increasing number amount of different optical components. This present challenges for the alignment of the optical components, since an ever increasing amount of manual work is needed both at the integration and at the operating sites to bring the instruments to the working condition.

In addition, even more challenging is the fact that components can be located in cryostatic environments, which prevents direct manual adjustments — and important fine-tuning at the operating temperatures.

Most of the top-class optical telescopes have taken approaches to use active and adaptive optics to compensate optical distortions caused by the changing operation conditions. However, these approaches have not yet been taken to the instrumental level at the same extent, although adaptive components are now being more and more deployed (e.g., X-Shooter VLTI).

We would like to propose a general concept making automatic optical calibrations possible at a wide range of new optical instruments. This would be done in two stages: at first, phase-retrieval techniques^{1,2} would be used to determine the wavefront (WF) shape at different points of the focal plane. This would be done by using the intensity images of the imaging camera, and by introducing a known defocus as a phase-diversity. Then, an inverse model of the optical system would be used to determine the alignments that would optimally compensate the observed distortions. These two steps are iterating until a convergence is reached.

This paper investigates the first part of the metrology concept. We compare two general phase-retrieval techniques to determine the parameter range at which sufficient phase-retrieval is possible. The latter part of the problem is discussed in our another paper.³

Section 2 describes in detail the cases we studied, and Section 3 shows the results. Finally, conclusions are discussed in Section 4.

Further author information: (Send correspondence to V.K.)
V.K.: E-mail: korkiakoski@strw.leidenuniv.nl

2. SELECTED STUDY CASES

We concentrate on two different phase-retrieval algorithm: phase-diversity (PD), as described by,^{2,4} and Gerchberg-Saxton (GS).⁵ The former makes a gradient search to find the unknown wavefront and object that best fit to the measured data. The latter uses iterative Fourier transforms to apply known restrictions in pupil and focal plane.

We simulated typical wavefront distortions and corresponding defocused focal-plane images. Then, we calculated how well the algorithms were able to reconstruct the unknown wavefronts by using the camera images. The common simulation assumptions are enumerated in the following.

- Imaging: the camera images are simulated using the standard FFT-approach. No additional rebinning is done on the arrays simulated by the Fourier transforms. We assume monochromatic point sources are used as objects.
- The wavefront error is simulated by 45 Zernike polynomials, the energy distribution follows a power law of $(n + 1)^{-2}$, where n is the polynomial radial order. We remove tip/tilt before applying the phase-retrieval since it can be determined simply from the position of the spot. For each identical set of parameters, we always computed 12 different random WF realizations to model the statistics. The wavefront is sampled by an array of 128×128 pixels.
- The used phase-diversity implementation is described in.⁶ As a modal base, we use Zernike modes,⁷ and we reconstruct 166 modal coefficients.
- The implementation of GS algorithm is well described in the literature. The implementations of the phase-unwrapping, however, can vary. Due to an ease of implementation, we have done our unwrapping without any quality-checks. Therefore, our GS phase-retrieval can result in a complete failure, although a smarter unwrapping would give acceptable results. This can be improved in the future.
- The defocus shape, acting as a phase diversity, is created by 3–6 different defocus magnitudes. The applied diversities are evenly distributed between the used focal range. The first image is always focused, and the following images are defocused by distances enumerated in Table 1 (a being the maximum applied defocus distance).

Table 1. Applied defocus distances

#image	1	2	3	4	5	6
$N = 3$:	0	$-a$	a			
$N = 4$:	0	$-a$	$-1/3 a$	$1/3 a$		
$N = 5$:	0	$-a$	$-1/3 a$	$1/3 a$	a	
$N = 6$:	0	$-a$	$-3/5 a$	$-2/5 a$	$1/5 a$	$3/5 a$

The phase-retrieval performance was evaluated by what we call an improvement ratio. It is defined as

$$R_i = \frac{\text{rms}(\phi_r - \phi_0)}{\text{rms}(\phi_0)},$$

where ϕ_r is the reconstructed wavefront and ϕ_0 is the unknown aberration (having TT removed). Thus, improvement ratio 0 means a perfect reconstruction, and ratios higher than 1 mean that the phase-retrieval completely fails.

3. SIMULATION RESULTS

3.1 Simulations studying the optimal diversity magnitude

At first, we studied how well the phase-retrieval methods can reconstruct an unknown wavefront by a single iteration.

We are studying five different WF rms: 0.5, 1, 2, 4 and 6 rad. Since the tip/tilt is removed before applying PD, the actual values are about 20% smaller.

Next, we will study the optimal maximum applied diversity, a (shown in Table 1).

Fig. 1 shows how the improvement ratio changes as a function of a . We show the plots for two different CCD samplings, when phase-diversity retrieval is applied. PSF width (λ/D) is 2 or 3 pixels.

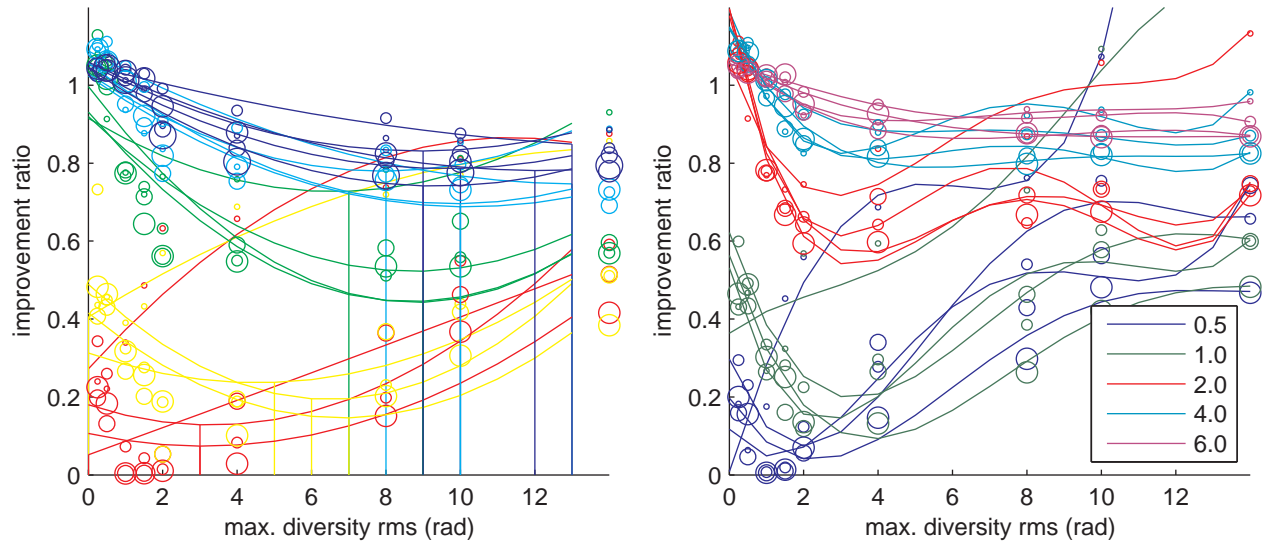


Figure 1. Improvement ratio for phase-diversity as a function of maximum applied diversity. The colors show the original WF rms 0.5–6 rad (red: 0.5, yellow: 1.0, green: 2.0, light blue: 4.0, dark blue: 6.0). The bubble sizes show the number of used images (3–6), smaller bubbles denote less images. Lines are polynomial fits to the data. Left: $\lambda/D = 2$ pix. Right: $\lambda/D = 3$ pix.

The two different PSF resolutions do not show fundamentally different behavior. The less aberrations in the original WF, the better the PD performs: the optimal improvement ratios are around 0, 0.2 0.6 and 0.8 for WF rms of 0.5, 1.0, 2.0 and 4.0 respectively. As long as more than 3 images are used, the optimal diversity magnitude does not depend on the number of used images. However, it is necessary to use larger diversity magnitudes when the WFs have larger aberrations: the optimums are about 2, 4, 6 and 12 rad for the studied cases.

Regarding the performance of GS algorithms, we noted that the improvement ratio not strongly depend on the applied diversity nor the magnitude of the aberrations in the original WF. Therefore, the results are not plotted here.

Next, we show, in other plots in Fig. 2, the optimal improvement ratio as a function of required images. Fig. 3 shows the optimal improvement ratio as a function of initial WF aberrations.

It is seen that the number of images has very little impact on the optimal reconstruction, when GS is used. With PD, the optimal reconstruction improves 10–20%, when the number of images is increased from 3 to 4–5. Using more than 5 images has a negligible impact.

The improvement ratios of GS are significantly better compared to PD, in particular at WF aberrations larger than 2 rad rms. At such large aberrations, PD is correcting only 10–30% of the initial WF, but GS can still correctly detect more than 60%.

We also see that the improvement ratio as a function of initial WF rms behaves very differently for PD and GS. With PD, the phase-retrieval gets relatively worse and worse, when the WF aberrations are increased. Below 1 rad, 80–100% of the aberrations are correctly detected, but only 10–50%, when the aberrations are 2–6 rad rms. GS, on the other hand, seems to work relatively best, when the initial WF aberrations have a 2–4 rad rms error: then they correct 80–90% of the WF. However, at WF rms errors of 0.5 or 5–6 rad rms, GS corrects only 70%.

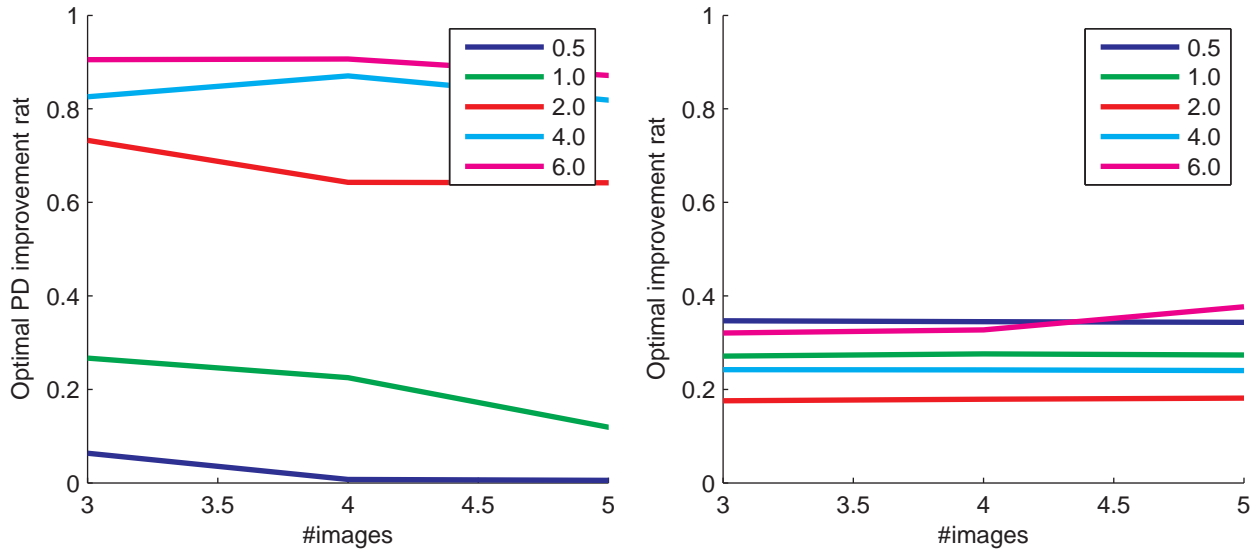


Figure 2. Optimal improvement ratio as a function of used images. Used PSF resolution: $\lambda/D = 3$ pix. The colors show the amount of unknown WF aberrations. Left: PD. Right: GS.

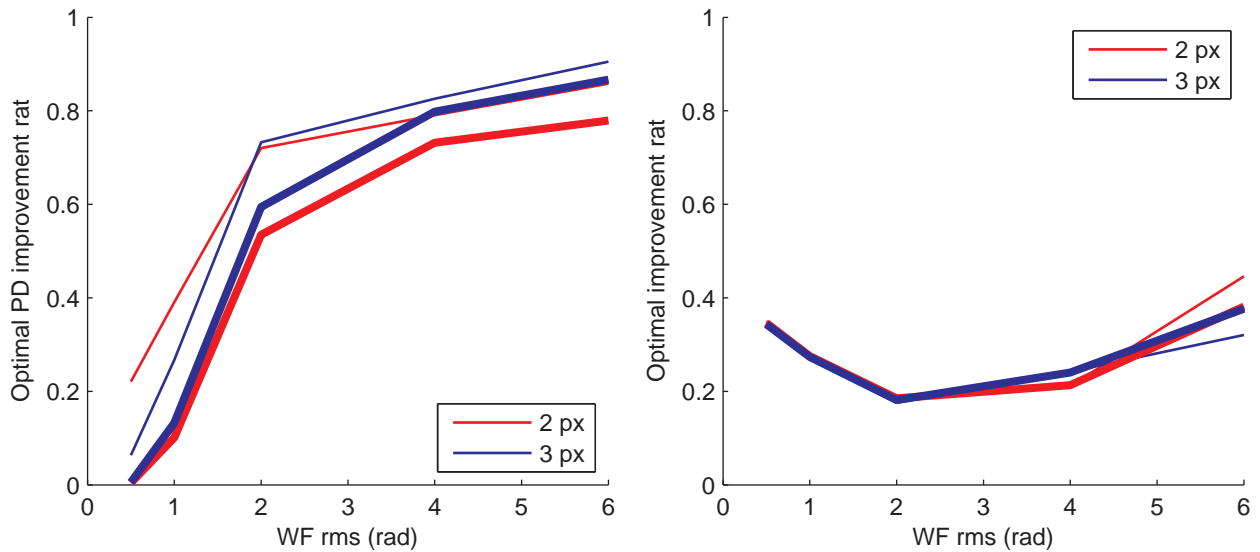


Figure 3. Optimal improvement ratio as a function of initial WF aberrations. Colors show the PSF resolution. Line widths show the number of used images (thin lines: 3 images, thick lines: 6 images). Left: PD. Right: GS.

When comparing the PSF resolutions of $\lambda/D = 2$ and 3 pixels, we see that the impact on optimal improvement ratio is not dramatic. At the largest WF aberrations, PD appears to be $\sim 5\%$ better with the coarser sampling of 2 pixels. This will be discussed in more detail in the following sections.

3.2 Failure rate of phase retrieval

The previous section illustrated the phase-retrieval performance by an averaged improvement ratio, where the average is over random WF realizations. Another way to evaluate the retrieval is to calculate the percentage of the cases where the reconstruction is successful.

We define a case as a failure, when the improvement ratio is more than 0.9–1.0. Thus, we can compute a failure rate by comparing the number of failed reconstructions to the total number of used WF random realizations.

Figs. 4 and 5 show the failure rates of the WF reconstructions for PD and GS, respectively. We study PSF resolutions of $\lambda/D = 2$ pix and $\lambda/D = 3$ pix, and 3–5 diversity images are used.

It is seen that, again, PD and GS have different characteristics for the failure rates.

Regarding the PD, as long as at least 5 images are used, all the less-strict fail-rates drop to zero when optimal diversity magnitude is applied, even at largest WF aberrations of 6 rad. The reconstructions appear to become significantly more robust, when 5 diversity images are used instead of 3–4; and also a coarser PSF resolution of $\lambda/D = 2$ pix gives better results. It is seen that at the optical case (5 images, 6–8 rad rms defocus diversity and $\lambda/D = 2$ pix), at least 5–10% of the WF aberrations can be correctly detected if the initial WF rms error is 4–6 rad.

Regarding the GS, it is seen that the failure ratios are remarkably lower. The PD reconstructions are successful virtually always when the initial WF error is lower than 4 rad rms. When the aberrations are increased to 6 radians, we see that with most parameters, the reconstructions fail in 10–20% of the cases. However, by selecting the parameters optimally — enough diversity images and sufficiently large image resolution — this can be avoided. Unlike PD, GS appears to work better with a higher PSF resolution of $\lambda/D = 3$ pix. We also see that the optimal diversity magnitude is about 50% lower with GS compared to PD (2–3 rad rms).

3.3 Convergence issues of phase-diversity

As observed in the previous section, PD is able to correct only about 20–40% of the unknown WF, if its rms error is 2–4 rad. This happens even if we use a large number of images to obtain a sufficient knowledge for the inversion problem.

When PD is applied for a point-source, the error metric that is optimized should be very close to zero at the optimal wavefront reconstruction. However, there is no guarantee that the gradient search will converge to the global optimum.

We illustrated this issue in Fig. 6. It shows the improvement ratio as a function of relative error metric: $L_m(\phi)/L_m(0)$, where $L_m(\phi)$ is the error metric at the convergence and $L_m(0)$ is the error metric for a flat wavefront. The points come from individual wavefront reconstructions, and the colors illustrate the number of used diversity images.

It is seen, in the left plot, that the $L_m(\phi)/L_m(0)$ ratios tend to appear in bands. For the WFs having aberrations of 0.5–1 rad rms, we see one band close to origin, where both L_m and improvement ratios are small. This is the optimal point where the PD algorithm should converge to. Then, we see another band that forms a sharp triangle, its tip at the origin and base pointing towards the top-right corner of the plot. It is interesting that the areas where L_m would be small and improvement ratio large, or L_m large and improvement ratio small, have no data points. This means that the PD algorithm has a region of convergence where the phase reconstruction differs from the aberrations causing the measured images — the convergence gets stuck to local minima.

When we increase the aberration up to 4 rad WF rms, we see that the PD no longer converges to such neat bands. Instead, in particular with aberrations of 4 rad rms, we find many cases where L_m is very small, but also the improvement ratio is small. We conclude that the phase wrapping ($\phi > 2\pi$) further increases the issues with the converge-stuck at the local minima.

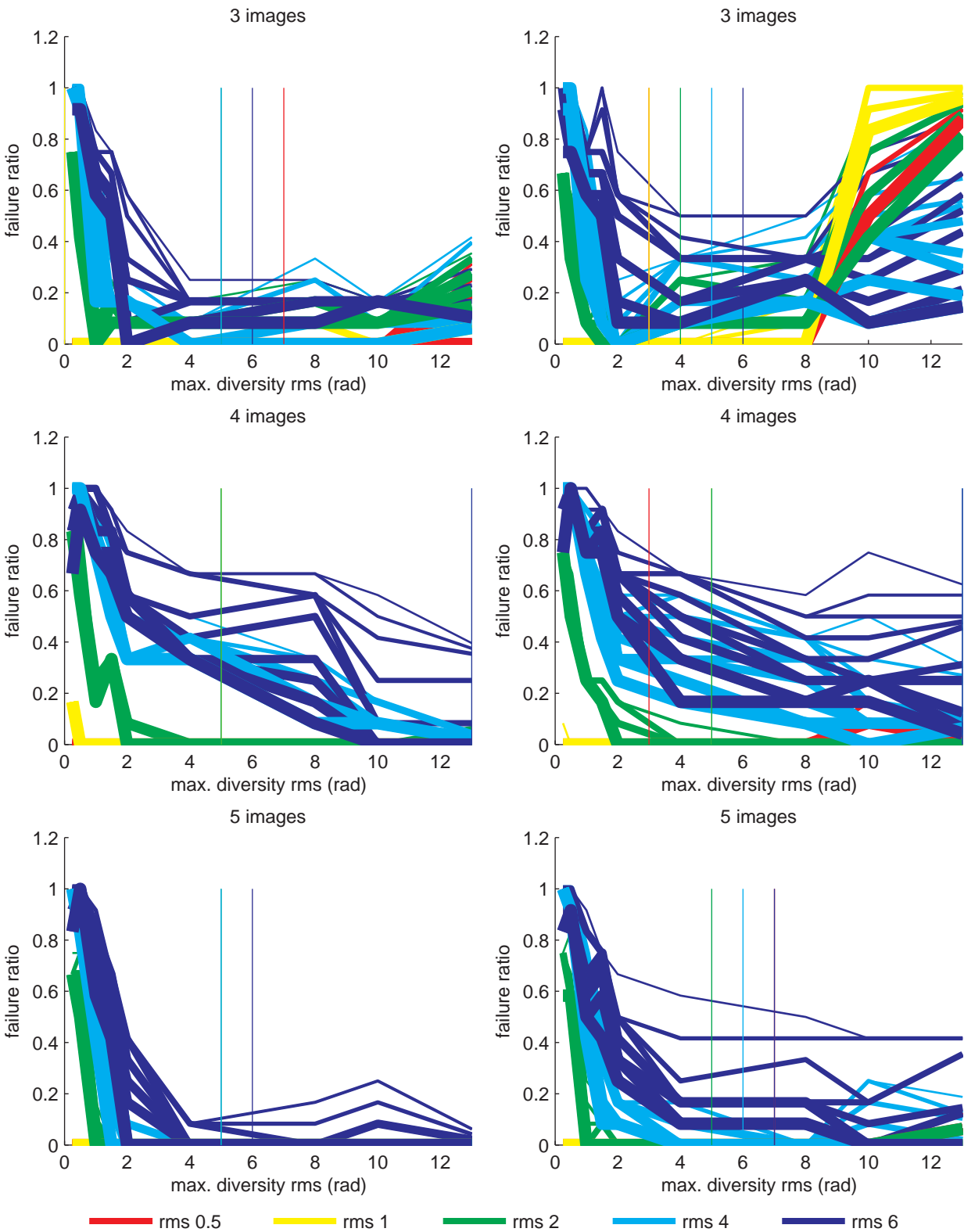


Figure 4. PD success rates. Line thickness indicate the success definition strictness (thinnest lines: success if $R_i < 0.9$, thickest lines: success if $R_i < 1.0$). Colors indicate initial WF rms. Left: $\lambda/D = 2$ pix. Right: $\lambda/D = 3$ pix.

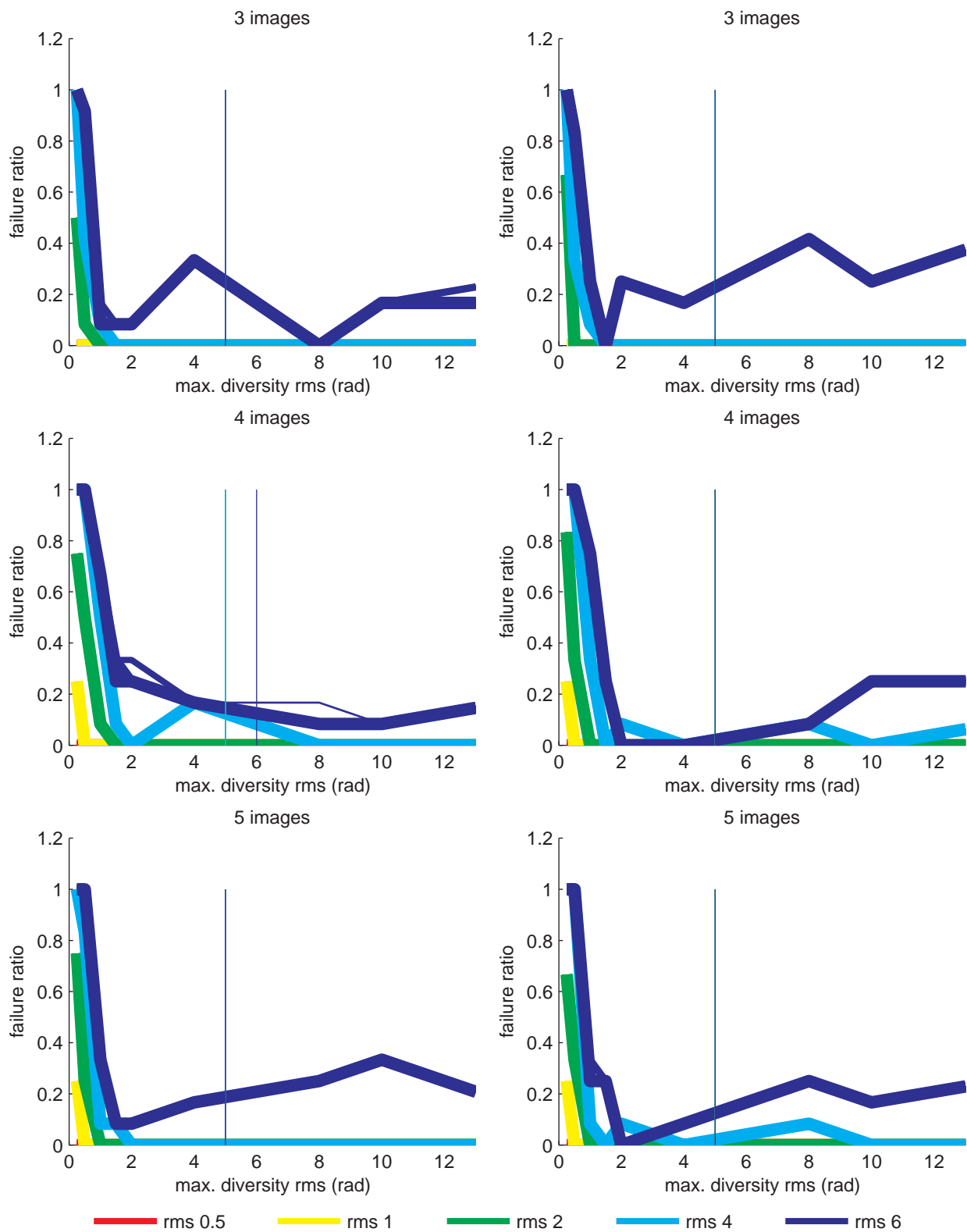


Figure 5. GS success rates. See details in Fig. 4.

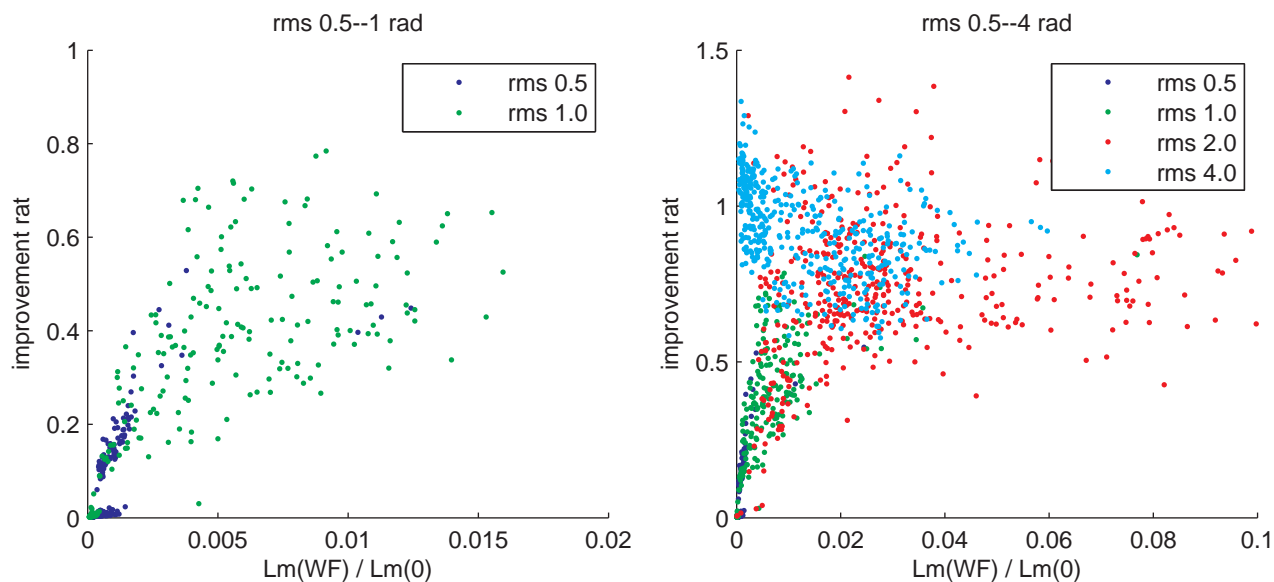


Figure 6. Improvement ratio as a function of $L_m(\phi)/L_m(0)$.

3.4 Simulations with multiple iterations steps

Next, we show preliminary simulations to illustrate how an iterative phase correction would work with phase-retrieval. We simulate a system that would detect the phase aberrations at an imaging camera, using the phase-diversity. Then, the detected WF error would be corrected by the best approximation of Zernike modes that fit to the reconstructed wavefront.

Since the GS algorithm does the phase retrieval in a more reliable way, we know that these simulations show the pessimistic case. At large WF aberrations, PD can correctly detect only 10–30% of the aberrations, and we will next study, if that is sufficient.

We start with an aberration phases similar to the previous section (initial wavefront rms error 2.0 rad), and then compute the PD reconstruction of 166 modes. Then, the PD reconstruction is removed from the residual phase, and the obtained residual WF is fed to the PD again. We study the convergence during 30 iterations.

We apply a constant diversity magnitude of 1 rad, which is the optimum for small WF aberrations of 0.5 rad, but less what is optimal for 2.0 rad.

Fig. 7 shows, in the left plot, the residual WF rms errors as a function iteration. We show the the largest, mean and smallest rms values over the 12 random WF realizations.

It is seen that when 3,4 or 6 images are used, the iteration either diverges or converges to non-zero values, depending on the initial random WF realization. However, when 5 or 7 images are used, 5–10 iterations are sufficient to obtain almost perfect wavefront correction.

To evaluate the robustness of the algorithm, we also made a similar test, but introduced a small error on the applied diversities (defocus distances). This was modeled as a randomly distributed noise, added on all the defocus distances.

Fig. 7 shows, in the right plot, the best obtained WF correction as a function of diversity error. It is seen that the algorithm seems remarkably robust, and diversity errors of 0.5 rad (in WF rms) cause an rms error in corrected WF only of 0.2–0.4. Using more images are effective in compensating the diversity errors: using 7 instead of 5 images can more than halve the diversity-based uncertainty error in the corrected WF.

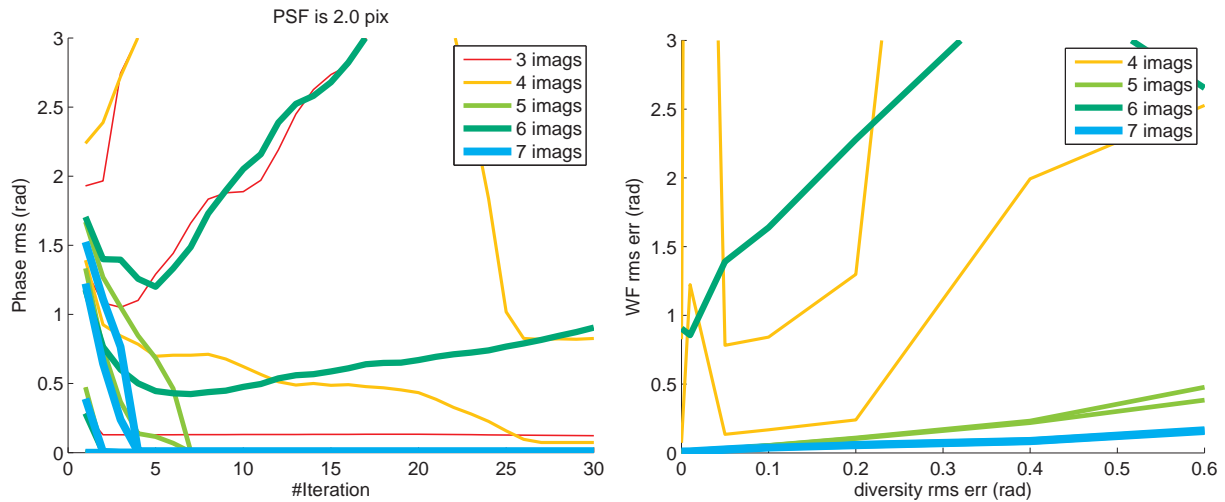


Figure 7. Left: residual WF energy after consecutive PD iterations. The colors indicate number of images (3–7). Upper lines show the maximum over 12 random realizations, the middle lines show the means and lower lines the minima. Right: Residual WF energy, after 20 consecutive PD iterations, as a function of diversity error.

3.5 Computational issues

The gradient-search based phase-diversity algorithms are computationally extremely demanding. Here, we illustrate the current status of our codes on a typical computing machine.

We wrote our codes in C, and they are able to compute the gradients in parallel. In the examples illustrated in the following, we used 4 CPUs in a single processor (Quad-core Xeon E5504, 2.0 GHz, 800 MHz, 4 MB cache), to make the computations.

Fig. 8 illustrates the computational demands. We show, as a function of used images, the total computation time needed for a single WF reconstruction, the number of gradient step iterations and the time per a single iteration.

It is seen that the total computation time per one WF reconstruction increases roughly quadratically from 3–7 min (with 3 images) to 15–60 min (9 images). The total computation time depends strongly on the amount of aberrations in the unknown wavefront: it is possible to reconstruct a WF with an rms of 0.5 rad in about 3–15 min (depending on the amount of used images), but a WF with an rms of 4.0 rad requires 2–3.5 times more computations. This happens because the gradient search converges more slowly in the presence of larger aberrations.

The upper right plot in Fig. 8 shows the number of required iterations as a function of used images. It is seen that a WF with an rms 0.5 rad requires only 5–10 iterations, but more than 15 are needed when the WF aberrations have an rms of 2–4 rad.

It is also interesting how the number of required iterations changes, when more images are used. When we use 7–9 images instead of 3, it is necessary to make 1.5–2.5 times less iterations for WFs with an rms of 0.5–1.0. However, the WFs with an rms of 2–4 rad require 0–10% more iterations.

It needs to be mentioned, though, that we limited the maximum number of iterations to 20, and therefore the data is slightly misleading, in particular with larger WF aberrations.

Fig. 8 also shows that a single iteration takes 20–200 s to complete, and that depends only on the number of images (the slight deviations in the plots are due to numerical timing errors).

Based on the previous sections, it seems beneficial to use up to 6 images in the reconstruction. Therefore, it can be concluded that, with the current codes and machines, it will take 8–25 min to reconstruct a single WF, and a complete iterative WF correction would take 5–10 times longer. Taking into account the potential of code optimizations (for instance, using more CPUs for gradient computations, reconstructing less modal coefficients,

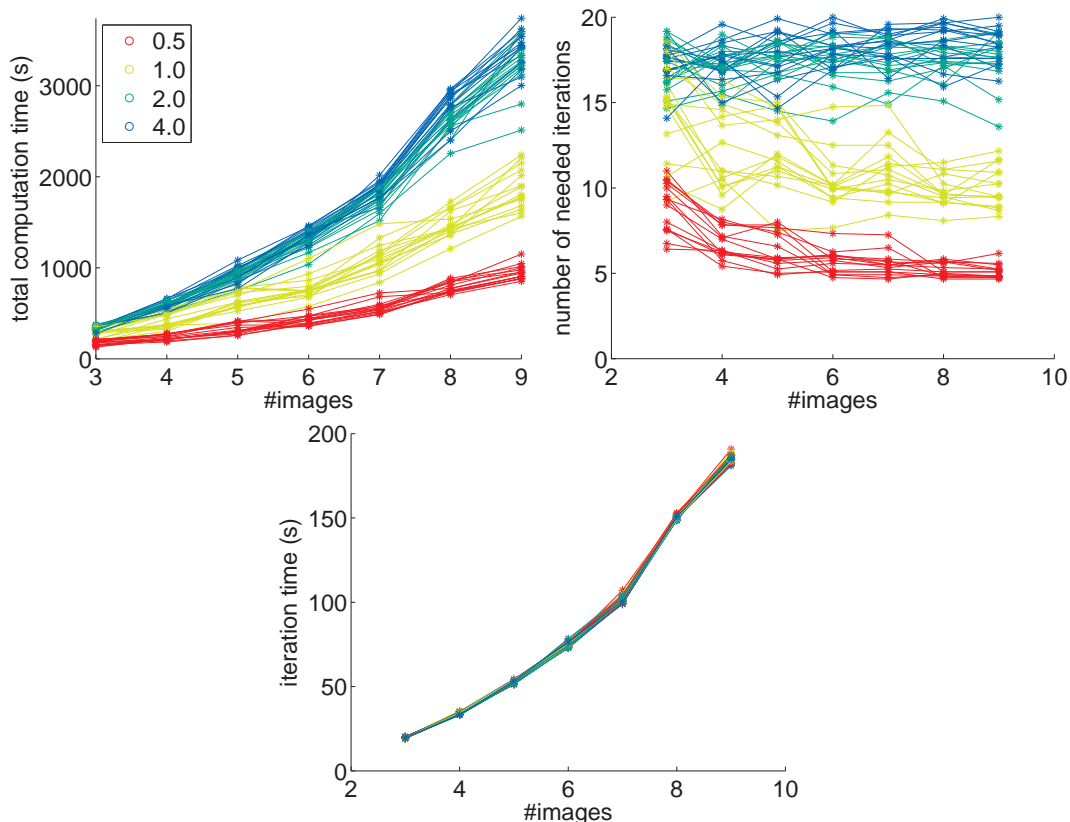


Figure 8. Illustration of computational demands for the simulations illustrated upper left in Fig. 1. Colors indicate the amount of aberrations in the original WF. The plots are averages over 12 random WF realizations, and each line shows the results for a different diversity magnitude.

numerically more efficient gradient computation⁸), we conclude that the computational issues are not a major problem.

4. CONCLUSIONS

We have studied, by numerical simulations, the potential of phase-diversity and Gerchberg-Saxton algorithms in phase-retrieval. We concentrated in a case where the imaged object is a point source, and the diversity is created by several dynamically adjustable defocus distances. The simulations schemas were designed to be descriptive for systems with a capability to change camera focus and correct low-order aberrations caused by errors in the optical alignment.

The simulations demonstrate that the phase-retrieval can be successfully done, but the magnitude of the unknown WF aberrations limit the accuracy of individual phase reconstructions. In particular, with phase-diversity, the issue is severe: wavefronts with an rms error less than 0.5 rad can be reconstructed almost perfectly, but an rms error of 4–6 rad means that only about 10–20% of the unknown WF energy can be successfully detected.

This loss of accuracy in PD is caused by the fact it is using a gradient search prone to getting stuck in local minima. The GS algorithm does not suffer from the effect at the same extent, because it is inherently more efficient to model monochromatic sources and using the imposed constraints in pupil and focal-plane. Therefore, it is not surprising that we saw the GS being able to retrieve the phase with an accuracy of 50% even at 6 rad rms. We believe this is sufficient for our applications.

In addition, we demonstrated that the perfect reconstruction of large wavefronts is not absolutely necessary: starting from initial WF error of at least 2 rad rms, the phase-diversity can be applied iteratively to reach almost

an optimal performance. Since GS was shown to retrieve the phase with more than comparable accuracies at 6 rad rms, we conclude that the iterative WF correction should work in practice up to those initial aberration magnitudes. However, in the iterative approach, it is still necessary to have a rough estimate of the magnitudes of the unknown wavefronts to optimize the applied diversity magnitude. We believe this is possible, for instance, by creating a loop-up-table based on the observed maximum intensities at the detector.

Regarding the computational demands, we studied the WF reconstruction times on current easily available computers. Our phase-diversity implementation, which is computationally much more demanding, should be able to complete a full calibration in less than an hour. Gerchberg-Saxton is computationally less demanding since it does not need to estimate the object parameters. Therefore, we conclude neither the computational issues should be the problem.

The next step will be combining the phase-retrieval approaches presented in this paper to the tomography solvers discussed in.³ To ultimately validate the concept, we need to demonstrate that the iterative optical alignment converges to an optimal wavefront quality at the detector. This means that the inverse problem solving the tomography should be sufficiently tolerant to the retrieval errors we will make at the large WF aberrations. The issue will be addressed in our future work: we will implement detailed end-to-end numerical simulations leading to reliable performance estimates of selected actively calibrated instruments, like the METIS spectrograph.

In the future, it is also possible to include deformable mirrors at the instrumental level to deal with the optical alignment errors. This would make it possible to use arbitrary diversity shapes in the phase retrieval. We have demonstrated by laboratory experiments that such sequential phase-diversity algorithms work well to calibrate the non-common path aberrations in a conventional adaptive optics system.⁹

Acknowledgments

This research, part of the OPTICON program, is supported by the European Commission's FP7 Capacities programme (Grant number 226604). We acknowledge also the STW Smart Optics project.

REFERENCES

- [1] Fienup, J. R., "Phase retrieval algorithms: a comparison," *Applied Optics* **21**, 2758–2769 (Aug. 1982).
- [2] Paxman, R. G., Schulz, T. J., and Fienup, J. R., "Joint estimation of object and aberrations by using phase diversity," *Journal of the Optical Society of America A* **9**, 1072–1085 (July 1992).
- [3] Agocs, T., Venema, L., Korkiakoski, V., and Kroes, G., "Metrology for Optimizing Optical Systems with Active Components," in [*Modern Technologies in Space- and Ground-based Telescopes and Instrumentation II*], *Proc. SPIE*, This conference (2012).
- [4] Löfdahl, M. G. and Scharmer, G. B., "Wavefront sensing and image restoration from focused and defocused solar images," *Astronomy and Astrophysics Supplement* **107**, 243–264 (Oct. 1994).
- [5] Gerchberg, R. W. and Saxton, W. O., "A practical algorithm for the determination of the phase from image and diffraction plane pictures," *Optik* **35**, 237 (1972).
- [6] Korkiakoski, V., Keller, C. U., Doelman, N., Fraanje, R., and Verhaegen, M., "Joint optimization of phase diversity and adaptive optics: demonstration of potential," *Applied Optics* **51**, 102 (Jan. 2012).
- [7] Noll, R. J., "Zernike polynomials and atmospheric turbulence.," *Journal of the Optical Society of America (1917-1983)* **66**, 207–211 (1976).
- [8] Vogel, C. R., Chan, T. F., and Plemmons, R. J., "Fast algorithms for phase-diversity-based blind deconvolution," in [*Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*], D. Bonaccini & R. K. Tyson, ed., *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **3353**, 994–1005 (Sept. 1998).
- [9] Korkiakoski, V., Keller, C. U., Doelman, N., Fraanje, R., Andrei, R., and Verhaegen, M., "Experimental validation of optimization concepts for focal-plane image processing with adaptive optics," in [*Adaptive Optics Systems III*], *Proc. SPIE*, This conference (2012).