Delft University of Technology

Phase retrieval based on the vectorial model of point spread function

Thao, Nguyen Hieu; Soloviev, Oleg; Verhaegen, Michel

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Phase retrieval based on the vectorial model of point spread function

NGUYEN HIEU THAO,[1,2,*,†] **OLEG SOLOVIEV,**[1,3,†] **AND MICHEL VERHAEGEN**[1]

[1]*Delft Center for Systems and Control, Delft University of Technology, 2628CD Delft, The Netherlands*
[2]*Department of Mathematics, School of Education, Can Tho University, Can Tho, Vietnam*
[3]*Flexible Optical B.V., Polakweg 10-11, 2288 GG Rijswijk, The Netherlands*
Corresponding author: h.t.nguyen-3@tudelft.nl

We present an efficient phase retrieval approach for imaging systems with high numerical aperture based on the vectorial model of the point spread function. The algorithm is in the class of alternating minimization methods and can be adjusted for applications with either known or unknown amplitude of the field in the pupil. The algorithm outperforms existing solutions for high-numerical-aperture phase retrieval: (1) the generalization of the method of Hanser *et al.*, based on extension of the scalar diffraction theory by representing the out-of-focus diversity applied to the image by a spherical cap, and (2) the method of Braat *et al.*, which assumes through the use of extended Nijboer–Zernike expansion the phase to be smooth. The former is limited in terms of accuracy due to model deviations, while the latter is of high computational complexity and excludes phase retrieval problems where the phase is discontinuous or sparse. Extensive numerical results demonstrate the efficiency, robustness, and practicability of the proposed algorithm in various practically relevant simulations. © 2019 Optical Society of America

https://doi.org/10.1364/JOSAA.37.000016

1. INTRODUCTION

Phase retrieval is a prominent inverse problem in optics that aims at recovering a complex signal at the pupil plane of an optical system given one or several out-of-focus intensity images measured along the optical axis. Various formulations and modifications of phase retrieval have originated from a wide range of important applications in adaptive optics [1–4], microscopy [5,6], astronomy imaging [7,8], x-ray crystallography [9,10], etc. For optical systems with low numerical aperture (NA), the intensity distribution in the focal plane and the complex signal in the pupil plane are simply related via the Fourier transform [11]. This fundamental relationship has given rise to a wide range of phase retrieval algorithms [12–17] in low-NA settings since the pioneering work of Sayre [18] revealing that the phase of a scattered wave can be recovered from the recorded intensity images at and between Bragg peaks of a diffracted wavefront. Recent overviews on this topic can be found, for example, in the surveys [19–22]. An essential condition for the validity of the Fresnel approximation is that the effect of light polarization on the diffracted images is negligible. For high-NA optical systems, however, the vector nature of light cannot be neglected and point spread functions (PSFs) are formed according to a more involved formulation [23,24], which is called the "vectorial PSF model" in this paper, to be distinguished from the scalar one according to the Fresnel approximation. In contrast to a

vast number of existing phase retrieval algorithms in low-NA settings, only a few solution approaches have been proposed for phase retrieval in high-NA settings.

The first and natural solution approach to high-NA phase retrieval is to adapt the scalar diffraction theory for high-NA imaging systems [25]. The main modification is that the additional defocus term used to calculate the corresponding out-of-focus image in low-NA settings is replaced by an appropriate spherical cap in high-NA settings [25–28] (see Sections 2.A and 2.B.) For higher-NA imaging systems (NA ≥ 0.6), however, the accuracy of this phase retrieval approach is limited due to model deviations. This approach can be enhanced in manner by an ad hoc Gaussian fitting scheme [27]. However, this ad hoc scheme does not give consistent results in terms of phase restoration and cannot be theoretically supported or analyzed.

Another existing approach to deal with high-NA phase retrieval is based on the decomposition of the generalized pupil function (GPF) as a weighted sum of extended Nijboer–Zernike (ENZ) basis functions [29,30] in a vector diffraction theory setting. The main challenge of this approach is to evaluate the complex-valued coefficients of the polarized field components decomposed in terms of Zernike polynomials [31]. The numerical complexity of this approach is high, since the integrals that need to be evaluated have no closed-form solution. Moreover, since the GPF is approximated by a finite number of ENZ basis

functions, it is not valid to phase retrieval applications where the phase is sparse (discontinuous).

In this paper, we present a new phase retrieval approach for high-NA imaging systems, which is directly based on the vectorial PSF model. The algorithm is in the class of alternating minimization methods and can be adjusted for both cases of application with either known or unknown amplitude of the field in the pupil. The proposed algorithm overcomes the highlighted drawbacks of the two existing solution approaches indicated in the previous paragraphs. Numerical experiments demonstrate that for wavefronts with small amplitude (to exclude phase wrapping issues) the algorithm enables to precisely (up to a piston term) retrieve the wavefront from several vectorial PSFs in the noise-free setting.

The effectiveness and robustness of the algorithm numerically shown in Section 5 demonstrate its potential for applications to practical high-NA imaging systems. More importantly, our approach paves the way to extend the class of projection algorithms for phase retrieval in the setting of scalar diffraction theory to the framework of the vectorial one. It is worth recalling that projection methods outperform the phase retrieval algorithms in the other classes in terms of computational complexity, convergence speed, effectiveness and robustness [22]. Laying the groundwork for designing various projection algorithms for phase retrieval in the high-NA settings is another important outcome of this paper. That will be reported in a forthcoming publication.

To avoid possible confusion in terms of terminology, we mention that the term "vector phase retrieval" was also used in a different manner in the literature [32] where the objective is to restore two one-dimensional complex signals from the intensity of their Fourier transforms and the intensity of two structured interference patterns of the transformed signals.

2. POINT SPREAD FUNCTION MODELS

A. Scalar PSF Model for Low NA

For low-NA imaging systems, the PSF $I(\boldsymbol{u})$ can be related to the wavefront aberration $\Phi(\boldsymbol{x})$ as

$$I(\boldsymbol{u}) = \left| \mathcal{F}\left(\chi(\boldsymbol{x}) \cdot e^{j\Phi(\boldsymbol{x})}\right)\right|^2,$$

where $\boldsymbol{x} = (x, y)$, $\mathbf{u} = (u, v) \in \mathbb{R}^2$ are the coordinates in the pupil and focal planes, respectively, $I(\boldsymbol{u})$ is the intensity of the optical field in the focal plane, χ and Φ are, respectively, the amplitude and the phase of the collimated beam in the pupil plane, and $\mathcal{F}$ is the (two-dimensional) Fourier transform.

In this model, the out-of-focus PSF (see Fig. 1) corresponding to some distance z_d from the focal plane is calculated by adding a corresponding phase diversity term, ϕ_d, expressed via the defocus Z_2^0 (the Zernike polynomial of order two and azimuthal frequency zero):

$$\phi_d(\boldsymbol{x}) = \frac{\pi}{2\lambda}\mathrm{NA}^2 z_d Z_2^0(\boldsymbol{x}). \tag{1}$$

From now on, the coordinates $\boldsymbol{x}$ and $\boldsymbol{u}$ respectively corresponding to the pupil and focal planes will be omitted for the sake of brevity. The out-of-focus PSF at a distance z_d from the focal plane is given by the formula

$$I(\phi_d) = \left| \mathcal{F}(\chi \cdot e^{j(\Phi+\phi_d)})\right|^2. \tag{2}$$

B. Scalar PSF Model for High NA

The imaging model presented in this section is based on the scalar diffraction theory for high-NA imaging systems. The main difference from the model in Section 2.A is that due to the presence of large ray angles, the additional phase term cannot be approximated by a scaling of the defocus as in Eq. (1). Instead, it should be modified to a spherical cap [25,26]:

$$\phi_d = \frac{2\pi}{\lambda} z_d k_z, \tag{3}$$

with

$$k_z = \sqrt{1 - k_x^2 - k_y^2},$$

where λ is the wavelength, and (k_x, k_y) are the x and y components of the unit wave vector that satisfy

$$k_x^2 + k_y^2 \leq \mathrm{NA}^2.$$

In addition, an obliquity factor,

$$1/\sqrt{\cos\theta} = k_z^{-1/2}, \tag{4}$$

should be introduced into the amplitude to account for the angle θ between the Poynting vector and the normal to the imaging plane. The PSF with aberration Φ and out-of-focus displacement z_d is calculated according to Eq. (2) with ϕ_d given by Eq. (3) instead of Eq. (1) and the modified amplitude $\chi/\sqrt{\cos\theta}$ in place of χ.
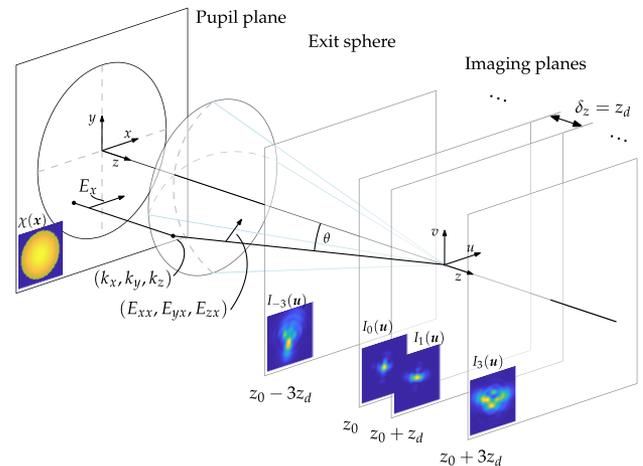
Fig. 1. Schematic diagram depicts the vectorial PSF model and the setup of phase retrieval given several out-of-focus measurements. A collimated beam with (possibly unknown) amplitude χ at the entrance pupil plane is focused by an aplanatic system at plane $z = z_0$, not necessarily on axis. Several out-of-focus PSFs are registered in z planes with known displacements from the focal plane. For high values of NA, the bending of rays introduced by the lens produces a significant z component of the electrical field (here shown on example of x-polarization component E_x), which should be taken into account when calculating the intensity in the imaging planes.

C. Vectorial PSF Model

For high-NA optical systems (e.g., NA ≥ 0.6 according to [29]), the bending of the rays created by a lens introduces a significant z component of the electromagnetic field in the region behind the lens, which introduces discrepancy with the scalar model of Eq. (2). One can model the PSF according to the vector theory of diffraction by considering the x, y, z components of the field right after the lens separately for x and y components (for a collimated beam, the z component is approximately zero) of the electromagnetic field just before the lens [24,26,27,33]. Thus, vectors $E_x(1, 0, 0)$ and $E_y(0, 1, 0)$ of unit length in x and y directions produce field with components denoted (E_{xx}, E_{yx}, E_{zx}) and (E_{xy}, E_{yy}, E_{zy}), respectively (see Fig. 1), and given by {see, e.g., Table 3.1 of [24] or Eq. (3) of [27]}

$$E_{xx} = 1 - \frac{k_x^2}{1 + k_z}, \quad E_{yx} = -\frac{k_x k_y}{1 + k_z}, \quad E_{zx} = -k_x,$$

$$E_{xy} = -\frac{k_y k_x}{1 + k_z}, \quad E_{yy} = 1 - \frac{k_y^2}{1 + k_z}, \quad E_{zy} = -k_y. \quad \textbf{(5)}$$

The six field components can be used to calculate the total electrical field energy at any point after the lens; in particular, they determine the intensity seen by an imaging plane in any orientation. For instance, the z components do not have the Poynting vector along the z direction and do not contribute to the intensity registered by a detector perpendicular to the z axis, and thus they may be discarded when calculating the PSF at z planes. However, a more general approach adopted in this paper is to keep all six field components and account for the angle between the Poynting vector and the imaging plane via the obliquity factor [24,27]. For imaging planes perpendicular to the z axis as depicted in Fig. 1, the obliquity factor is given by Eq. (4). By resetting $\chi := \chi/\sqrt{\cos\theta}$, one can integrate the obliquity factor into the amplitude χ, and, as a consequence, it will no longer be explicitly involved in the subsequent analysis. Then each of the right-hand-side terms in Eq. (5) can be treated as a corresponding amplitude modulation in the entrance pupil for calculation of a PSF with the scalar Fourier method of Eq. (2):

$$p_{cc} = |\mathcal{F}(E_{cc} \cdot \chi \cdot e^{j\Phi})|^2. \quad \textbf{(6)}$$

In the above and elsewhere in this paper, the index cc stands for one of the six pairs of coordinate indices: xx, yx, zx, xy, yy, and zy. One thus obtains six constituent PSFs according to Eq. (6), which can be used to calculate the vectorial PSF corresponding to any linear polarization in the entrance pupil. For an unspecified polarization state, all six PSFs are summed incoherently:

$$I = \sum_{cc} p_{cc}.$$

Therefore, the vectorial PSF model with an additional phase term ϕ_d is given by

$$I(\phi_d) = \sum_{cc} |\mathcal{F}(E_{cc} \cdot \chi \cdot e^{j(\Phi + \phi_d)})|^2. \quad \textbf{(7)}$$

The PSF with out-of-focus displacement z_d is obtained by plugging Eq. (3) into Eq. (7).

    Remark 2.1. *The imaging model in this section can be related to the, introduced by McCutchen, [26,33] three-dimensional Fourier transform of the angular spectrum given by the complex amplitude on the exit sphere of the optical system, which, in turn, for aplanatic systems, focusing a collimated beam is given by the field in the entrance pupil plane. This allows one to use the model without change for off-axis cases as well. For the systems violating the Abbe sine condition, the field amplitude on the exit sphere should be used instead of the pupil plane amplitude.*

    Remark 2.2. *The phase diversity ϕ_d can be more general than a term compensating for out-of-focus displacement given by* Eq. (3) *e.g., it can be introduced by using phase modulator devices.*

3. PROBLEM FORMULATION

A. Phase Retrieval Problem

For an unknown phase aberration $\Phi \in \mathbb{R}^{n \times n}$, let $r_d \in \mathbb{R}_+^{n \times n}$ ($d = 1, \dots, m$) be the measurement of m PSF images generated by Eq. (7) corresponding to different phase diversities ϕ_d. The phase retrieval problem is to restore Φ given r_d and ϕ_d ($d = 1, \dots, m$) and the physical parameters of the optical system. Mathematically, we consider the problem of finding $\Phi \in \mathbb{R}^{n \times n}$, such that

$$r_d = \sum_{cc} |\mathcal{F}(E_{cc} \cdot \chi \cdot e^{j(\Phi + \phi_d)})|^2 + w_d \quad (d = 1, \dots, m),$$

$$\textbf{(8)}$$

where χ is the (possibly unknown) amplitude of the generalized pupil function, and $w_d \in \mathbb{R}^{n \times n}$ ($d = 1, \dots, m$) represent noise.

B. Maximum-Likelihood Formulation

Let us define

$$I_d[\chi, \Phi] := \sum_{cc} \left| \mathcal{F}\left(E_{cc} \cdot \chi \cdot e^{j(\Phi + \phi_d)}\right) \right|^2 \quad (d = 1, \dots, m)$$

$$\textbf{(9)}$$

and denote their concatenation by $I[\chi, \Phi]$. Let r and w be the concatenations of r_d and w_d ($d = 1, \dots, m$), respectively. We assume that the measurement noise along pixels is independent and identically distributed with the conditional probability density function denoted by $\mathcal{P}$. Then the co-log-likelihood function of the noise distribution for each of the pixels is given by

$$\mathcal{L}(w) := -\log \mathcal{P}(I[\chi, \Phi]|r). \quad \textbf{(10)}$$

Using Eq. (10), we can associate the problem (8) with a maximum-likelihood formulation (minimum-co-log-likelihood) given by [34]

$$\min_{\chi \in \mathbb{R}_+^{n \times n}, \Phi \in \mathbb{R}^{n \times n}} f(\chi, \Phi) := \sum_{\text{pixels}} \mathcal{L}(r - I[\chi, \Phi]). \quad \textbf{(11)}$$

When the noise in Eq. (8) is of Gaussian model, the objective function f reduces to (a scaling of) the energy kernel, and Eq. (11) becomes a nonlinear least squares problem:

$$\min_{\chi \in \mathbb{R}_+^{n \times n}, \, \Phi \in \mathbb{R}^{n \times n}} f(\chi, \Phi) := \sum_{d=1}^{m} \| r_d - I_d[\chi, \Phi] \|_F^2 . \quad \textbf{(12)}$$

Remark 3.1. *When the amplitude χ of the GPF is known, one can always reset $E_{cc} := E_{cc} \cdot \chi$ without loss of generality. In that case, χ can be omitted in this section and elsewhere in the paper, particularly in expressions (8)–(12).*

Remark 3.2. *In this paper, we investigate the problem where the intensity images are registered with a sufficiently large number of photon counts, and hence Poisson noise can be approximated by a Gaussian distribution in view of the central limit theorem. Therefore, the assumption on Gaussian model of noise in order to derive Eq. (12) has no restriction for practical applications involving Poisson noise, provided that a sufficiently large number of photons are counted. For completeness, a specification of Eq. (11) for the case of Poisson noise is also available but more involved in terms of numerical complexity; see, e.g., [35,36].*

4. ALGORITHMS

Based on the vectorial PSF model (7), we propose a new solution approach for the nonlinear least squares problem (12) in cases of both known and unknown amplitude χ. In the sequel, the index d stands for $1, 2, \ldots, m$, where $m \geq 2$ is the number of input PSF images.

In step 4 of Algorithm 1, we use the following arithmetic convention: $\frac{0}{0} = 1/\sqrt{6}$. This is simply for ensuring the uniqueness of that step and has no substantial effect on the performance as well as the analysis of the algorithms.

Remark 4. 1. *The idea of step 8 of Algorithm 2 can be traced back at least to the work of Gonsalves [37].*

Remark 4. 2. *In the setting of out-of-focus images, i.e., the phase diversities ϕ_d are given by Eq. (3), the displacement values z_d and*

Algorithm 1.   Vectorial PSF model-based alternating minimization with amplitude constraint (VAM$_+$)

Input:

   $r_d \in \mathbb{R}_+^{n \times n}$—$m$ PSF images
   $\phi_d \in \mathbb{R}^{n \times n}$—$m$ phase diversities
   χ—amplitude of the GPF
   N—number of iterations, and τ—tolerance threshold
   Φ^0—initial guess for Φ.

Iteration procedure: given Φ^k

   1. $x_{cc}^k = E_{cc} \cdot \chi \cdot e^{j\Phi^k}$

   2. $X_{cc,d}^k = \mathcal{F}(x_{cc}^k \cdot e^{j\phi_d})$

   3. $I_d^k = \sum_{cc} |X_{cc,d}^k|^2$

   4. $Y_{cc,d}^k = \frac{X_{cc,d}^k}{\sqrt{I_d^k}} \cdot \sqrt{r_d}$

   5. $y_{cc,d}^k = e^{-j\phi_d} \cdot \mathcal{F}^{-1}(Y_{cc,d}^k)$

   6. $\bar{y}_{cc}^k = \left(\sum_{d=1}^{m} y_{cc,d}^k \right)/m$

   7. $\Phi^{k+1} = \arg\left(\sum_{cc} (E_{cc} \cdot \bar{y}_{cc}^k) \right)$.

Stopping criteria: $k > N$ or $\sum_{d=1}^{m} \| I_d^k - I_d^{k+1} \|_F < \tau$.
Output: $\hat{\Phi} = \Phi^{end}$—the estimated phase.

Algorithm 2.   Vectorial PSF model-based alternating minimization (VAM)

Input: the same input as Algorithm 1, except
       χ^0—initial guess for χ.
Iteration procedure: given χ^k and Φ^k
   1. $x_{cc}^k = E_{cc} \cdot \chi^k \cdot e^{j\Phi^k}$
   2–7. Steps 2–7 of Algorithm 1
   8. $\chi^{k+1} = \frac{|\sum_{cc} E_{cc} \cdot \bar{y}_{cc}^k|}{\sum_{cc} |E_{cc}|^2}$.
Stopping criteria: the same as Algorithm 1.
Output: $\hat{\Phi} = \Phi^{end}$—the estimated phase.
       $\hat{\chi} = \chi^{end}$—the restored amplitude.

---

the number m of images typically used for scalar phase retrieval algorithms are expected to be relevant for VAM and VAM$_+$.

Remark 4. 3 (stopping criteria). *Since running additional iterations of VAM and VAM$_+$ does not decrease the quality of phase retrieval in the sense that the PSFs reconstructed from the temporally estimated wavefront will not diverge any further from the data PSFs, the maximum number of iterations should be set sufficiently large whenever time consumption is not a major factor of concern. It is predetermined according to the number of data images, e.g., N can be tens of thousands for VAM with two input images (respectively, VAM$_+$ with one image), while it can be hundreds for experiments given seven or more images. In the presence of noise, error-reduction in terms of PSF restoration can be too small after some number (smaller than N) of iterations. One may then opt to terminate the algorithm using a predetermined threshold value τ, which is ideally the machine precision. We note that for offline applications without restriction on computing time, this stopping criterion is not really needed and can be dropped.*

The proposed algorithms belong to the class of Fourier transform-based methods and can be viewed as descendants of the classical Gerchberg–Saxton algorithm [12]. Algorithm 1 is (referred to as VAM$_+$) designed for phase retrieval given multiple intensity images and the amplitude constraint of the GPF. Algorithm 2 (referred to as VAM) is the adjusted version of Algorithm 1 tuned for phase retrieval applications where the amplitude of the GPF is not available. It can be viewed

---

Algorithm 3.   Scalar PSF model-based alternating minimization (SAM)

Input:

   r_d, ϕ_d, N, and τ—the same as Algorithm 1
   x^0—initial guess for $\chi \cdot e^{j\Phi}$.
Iteration procedure: given x^k

   1. $X_d^k = \mathcal{F}(x^k \cdot e^{j\phi_d})$

   2. $Y_d^k = \frac{X_d^k}{|X_d^k|} \cdot \sqrt{r_d}$

   3. $y_d^k = e^{-j\phi_d} \cdot \mathcal{F}^{-1}(Y_d^k)$

   4. $x^{k+1} = \left(\sum_{d=1}^{m} y_d^k \right)/m$.

Stopping criteria: $k > N$ or $\| x^k - x^{k+1} \|_F < \tau$.
Output: $\hat{\Phi} = \arg(x^{end})$—the estimated phase.

as the vectorial PSF model-based extension of the scalar PSF model-based alternating minimization proposed by Hanser *et al.* [25] for phase retrieval in high-NA settings. In this paper, the latter is referred to as the SAM algorithm and will serve as the basic for demonstrating the advancement of VAM and VAM$_+$ algorithms. For convenience of the reader, it is also recalled in Algorithm 3.

The three algorithms above fall in the class of alternating optimization methods; in particular, they can be proven to be descent methods for minimizing an associated objective function. The descent property of VAM and VAM$_+$ algorithms for minimizing the objective function given by Eq. (12) is clearly demonstrated in the numerical experiments in Section 5.C. Note that the descent property of the SAM algorithm is with respect to its own objective function. In optimization, an optimal solution is more desirable than the optimal value of the objective function. On the one hand, a direct measure of the distance from the temporal estimate to an optimal solution, which is not known in advance, is not obtainable in practice. On the other hand, the temporal objective value can be useful for establishing an error bound of the temporal approximate solution provided that the objective function satisfies a corresponding *error bound property* [38]. The latter is a property of the objective function itself and does not relate to a particular solution method. For the objective function given by Eq. (12) associated with VAM and VAM$_+$, the error bound property is not trivial and further analysis in this direction is quite involved and requires *a priori* mathematical assumptions; see, e.g., Ref. [39].

In terms of computational complexity, the evaluation of the Fourier transform constitutes the major part of their complexity. Since one iteration of a fast two-dimensional Fourier transform algorithm costs a scaling of $n^2 \log(n)$ flop counts, the complexity of the above Fourier-transform-based algorithms is $O(n^2 \log(n))$.

5. NUMERICAL EXPERIMENTS

Throughout this section, the vectorial PSF model (7) is taken as the forward imaging model for generating the PSF images. We simulate an imaging system having circular aperture with NA value of 0.95, except for the analysis of NA in Section 5.A. The amplitude χ of the GPF is (a scaling of) the two-dimensional Gaussian distribution truncated at 0.5 on the boundary of the aperture. The wavelength of the illuminating light is $\lambda = 300$ nm, the image size is 128×128 pixels, and the pixel size is 60 nm. For each analysis, we perform 75 phase retrieval experiments for 75 wavefront realizations taking values in one wavelength, i.e., $[-\pi, \pi]$. If not otherwise specified, data of each experiment consist of a set of seven out-of-focus PSF images uniformly separated by one depth of field/focus (DOF) along the optical axis (with $z_d = \text{DOF}$ in notations in Fig. 1). It is important to emphasize that the solution approach proposed in Section 4 is applicable not only to phase retrieval given out-of-focus images as depicted in Fig. 1, but any phase-diverse phase retrieval with known phase diversities (cf. Remark 2.2). White Gaussian noise is used for all experiments with noisy data. Except for the robustness analysis on various levels of noise in Section 5.B, the intensity images after being normalized to unity

energy are corrupted by white Gaussian noise at the signal-to-noise ratio of 40 dB (decibel). Recall that the signal-to-noise ratio expressed in decibels is defined by $\text{SNR} = 10 \ln(P/P_0)$, where P and P_0 are the powers of the signal and noise, respectively. The quality of phase retrieval is measured by the relative root mean square (RMS) error: $\|\hat{\Phi} - \Phi\|_F / \|\Phi\|_F$, where Φ and $\hat{\Phi}$ are the data and the estimated wavefronts, respectively. Since phase retrieval is ambiguous up to (at least) a piston term (the first Zernike mode), in this paper, the Frobenius norm of a phase object is always computed after the removal of its piston term. The algorithms considered in this section are of error-reduction type. Hence, running additional iterations would not decrease the quality of phase retrieval in terms of the reconstructed PSFs. For offline applications simulated below, the stopping criterion based on a tolerance threshold τ described in Section 4 is not so relevant and will be dropped. Each experiment consists of 150 iterations, except those for analyzing convergence properties of the algorithms in Sections 5.C and 5.D, where a few more hundreds of iterations are needed.

A. Analysis of Numerical Aperture

This section demonstrates the advanced features of the vectorial-PSF-model-based approach compared to the scalar one proposed by Hanser *et al.* [25]. Figure 2 summarizes the performance of Algorithms 1–3 on 75 experiments for six different NA values ranging from 0.7 to 0.95 in the presence of noise. The RMS errors of the restored wavefronts relative to the correct solution are shown. Suffering model deviations, the SAM algorithm (black) is much less accurate than VAM (red) and VAM$_+$ (blue). VAM and VAM$_+$ also have smaller error variance than SAM as indicated by the black boxes, which are much taller than the corresponding red and blue ones. This demonstrates the consistency of the phase retrieval approach based on the vectorial PSF model proposed in this paper. With additional
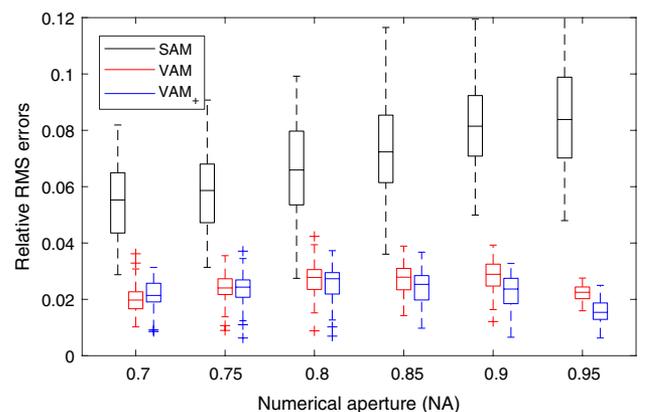
Fig. 2. Experiments with noise show the advantages of VAM and VAM$_+$ over SAM for various NA values. The relative RMS errors of the restored wavefronts compared to the correct solution are presented for 75 wavefront realizations and six different NA values ranging from 0.7 to 0.95. The performance of VAM (red) and VAM$_+$ (blue) is consistent for all experiments. Suffering model deviations, SAM (black) is significantly outperformed by the others in terms of both accuracy and consistency. VAM$_+$ slightly outperforms VAM thanks to the additional information of χ.
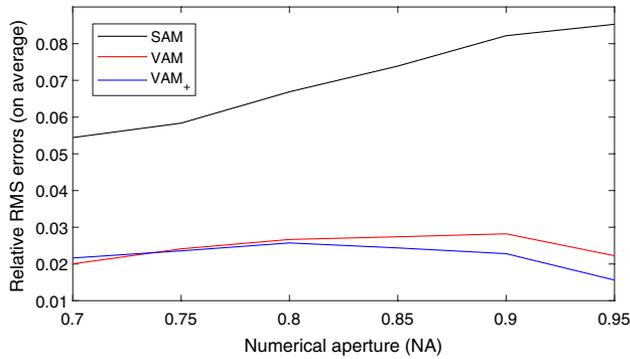
Fig. 3. Errors of phase retrieval over 75 experiments are presented for six different NA values. The relative errors of VAM (red curve) and VAM$_+$ (blue curve) are consistent at about 2%, while that of SAM (black curve) increases from 5.5% for NA = 0.7 up to 8.5% for NA = 0.95. The error of phase retrieval by SAM is approximately proportional to the NA value as indicated by the upward black curve.

Fig. 4. Experiments demonstrate the robustness of VAM and VAM$_+$ against noise. The two algorithms are consistent and reliable for SNR from 35 dB as demonstrated by small variation ranges of errors. The performance of SAM (black) is almost unaffected by noise with SNR from 30 dB; however, due to model deviations, it is clearly outperformed by VAM (red) and VAM$_+$ (blue).

Fig. 5. Average of the relative RMS errors of phase retrieval over 75 experiments is presented for seven different levels of Gaussian noise. VAM$_+$ (blue curve) is clearly superior to VAM (red curve) and SAM (black curve). The quality of phase retrieval by VAM and SAM for SNR 25 dB is at the same level, but the difference becomes more substantial for higher SNR. The relative RMS error of VAM$_+$ and VAM sharply decreases for higher SNR, while that of SAM remains high (above 8.5%) for all SNR due to model deviations.

information of the amplitude χ, the overall performance of VAM$_+$ is slightly better than that of VAM.

As shown in Fig. 3, the error of phase retrieval by SAM (black curve) is approximately proportional to the NA value since the occurrence of light polarization becomes more substantial for higher NA values. It increases from 5.5% (on average) for NA = 0.7 and up to 8.5% for NA = 0.95 as indicated by the upward black curve, while the relative errors of VAM (red curve) and VAM$_+$ (blue) are around 2%. The performance of VAM and VAM$_+$ depends very little on the NA value, since the latter information has been consistently incorporated into the problem setting and the input data. One may notice slightly better phase retrieval for NA = 0.95 compared to the smaller values, but this comes mainly from the approximation of presenting the results rather than from the quality of phase retrieval. More specifically, since the image resolution is fixed at 128 × 128 pixels for all experiments, a larger NA value corresponds to a finer grid on the circular aperture, which in turn makes the conversion from a phase screen to Zernike polynomials and vice versa more accurate. Figure 3 also reflects the level of deviation between the scalar and vectorial PSF models for different NA values in the sense that more model deviation leads to higher error of phase retrieval.

B. Noise Analysis

The influence of noise on the performance of VAM and VAM$_+$ is analyzed in this section. Different levels of Gaussian noise ranging from 25 dB to 55 dB are respectively introduced to 75 simulation data sets. The relative RMS errors of phase retrieval by VAM, VAM$_+$, and SAM are presented in Fig. 4. Due to the additional information of the amplitude χ, which is not corrupted by noise in these experiments, VAM$_+$ (blue) outperforms VAM (red). As expected, such a difference is not notable for high SNR (from 50 dB), where they both retrieve almost exact solutions. The influence of noise on the algorithms is also reflected by the variance of the retrieval errors, which appears to be inversely proportional to SNR. For SNR from 35 dB, the two algorithms are consistent and reliable. The complication of the vectorial PSF model compared to the scalar one explains why

VAM and VAM$_+$ are more sensitive to noise than SAM (black) whose performance is almost unaffected by noise with SNR from 30 dB. However, for SNR from 25 dB, the above drawback of VAM and VAM$_+$ compared to SAM is well compensated for by the usage of the correct PSF model. As shown in Fig. 5, the quality of phase retrieval by VAM (red curve) and SAM (black curve) for SNR 25 dB is at the same level, but the difference becomes more substantial for higher SNR. In particular, the relative RMS error of VAM$_+$ (resp., VAM) sharply drops to 5.3% (resp., 4.2%) at SNR 30 dB (resp., 35 dB), while that of SAM is above 8.5% for all SNR.

C. Convergence Properties

In this section, several convergence properties of VAM and VAM$_+$ including the descent property with respect to the objective function of (12) are numerically demonstrated. Figure 6 presents the objective value versus the number of iterations
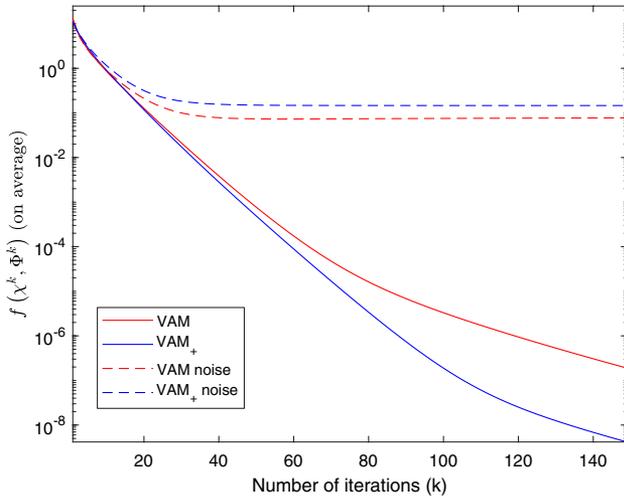
Fig. 6. Experiments demonstrate *descent property* of VAM and VAM$_+$. In the noise-free setting (solid curves), the objective value converges to zero, and with the additional information of χ, VAM$_+$ (blue) reaches the optimal value faster than VAM (red). In the presence of noise (dashed curves), the objective value reaches an *objective gap* in about 50 iterations, and we note that VAM optimizing over both χ and Φ yields a solution better fitting to the noisy data (smaller objective gap) than VAM$_+$ optimizing over only Φ.

Fig. 7. Experiments show the feasibility gap versus the number of iterations of VAM and VAM$_+$ compared to SAM. Regardless of the presence of noise, the performance of VAM (red curves) and VAM$_+$ (blue curves) is consistent, and the latter one is favorable due to the additional information of χ. Model deviations make SAM (black curves) much less accurate than VAM and VAM$_+$.

Fig. 8. Experiments demonstrate the convergence of VAM and VAM$_+$ in terms of the change in the estimated phase in iteration. In both scenarios of noise, the iterative change of the estimate decreases to zero in a stable and consistent manner. Due to the additional knowledge of χ, VAM$_+$ (blue curves) converges slightly faster than VAM (red curves).

for 75 phase retrieval experiments in the two settings with and without noise. The descent property is observed for all the experiments. For noiseless experiments (solid curves), the objective value $f(\chi^k, \Phi^k)$ decreases to zero (up to the machine precision) in a few hundreds of iterations. Recall that since the amplitude χ is fixed (assumedly known) for VAM$_+$, its objective function depends only on the sequence Φ^k. For experiments with noise (dashed curves), the sequence of objective values $f(\chi^k, \Phi^k)$ reaches an "objective gap" in about 50 iterations. The latter term is meant to indicate the gap in terms of the objective function, which is different from the "feasibility gap," i.e., the distance from the current estimate to an optimal solution. The objective gap empirically reflects the inexactness level of the input data in the sense that lower SNR yields a larger objective gap. It is worth noting that such an objective gap is not mathematically ensured to be unique, since the objective function f is far away from convexity, and hence it is not guaranteed to be the optimal value of problem (12).

In this section and Section 5.D, to observe convergence properties of the algorithms, we run a few more hundreds of iterations compared to 150 as in the other sections. However, the tailing parts of the curves are not plotted in Figs. 6–8, and 12 because otherwise different features of the algorithms at early iterations are hardly visible.

Let us point out an interesting influence of noise on the objective gaps shown in Fig. 6. For noiseless experiments (solid curves), the objective value of VAM$_+$ (blue) clearly decreases to zero faster than that of VAM (red) due to the additional knowledge of the amplitude χ, whereas in the presence of noise (dashed curves), VAM appears to achieve a smaller objective gap than VAM$_+$. This is because the optimization over both χ and Φ of VAM in general yields a better fit to the noisy data than the optimization over only Φ of VAM$_+$.

In the simulation framework, it is possible to monitor the feasibility gap in iteration, i.e., $\|\Phi^k - \Phi\|_F$, as plotted in Fig. 7. For the experiments without noise (solid curves), the feasibility gap of VAM and VAM$_+$ decreases to zero. This demonstrates the fundamental feature of the phase retrieval approach proposed in this paper. Roughly speaking, the two algorithms are capable of inverting the vectorial PSF model (7).

Similar to the case of the objective gap analyzed above, the feasibility gap of VAM$_+$ decreases faster than that of VAM due to the additional information of χ. For the experiments with

noise (dashed curves), the performance of VAM and VAM$_+$ is also consistent. It takes about 60 iterations to reach an approximate solution, i.e., an estimate corresponding to the feasibility gap that is almost no longer reduced by increasing the number of iterations. In contrast to the analysis of the objective gap shown earlier in Fig. 6 (dashed curves), VAM$_+$ outperforms VAM in terms of the feasibility gap, and hence it is favorable in these experiments. The influence of noise on the performance of SAM (black) is not observable due to model deviations that make the algorithm much less accurate than VAM and VAM$_+$.

The feasibility gap is a desirable quantity for any solution process, but it is not achievable in practice since no exact solution is known in advance. Instead, one is often interested in how a certain temporal estimate changes in iteration. For example, the change in the reconstructed PSFs (resp., the GPF) can be used as a stopping criterion for VAM and VAM$_+$ (resp., SAM), as presented in Section 4. When linear convergence appears to be the case, such an iterative change can be useful for estimating the rate of convergence, provided that the generated sequence remains in the convergence area. Our experiments with VAM and VAM$_+$ show the vanishing of the change in terms of temporally estimated wavefront, i.e., $\|\Phi^k - \Phi^{k+1}\|_F$, as presented in Fig. 8 (for VAM, $\|\chi^k - \chi^{k+1}\|_F$ is also vanishing in the same manner). Note that this implies the vanishing of the change in terms of reconstructed PSFs, since the imaging model (7) is continuous in Φ and χ. In both scenarios of noise, the iterative change in VAM and VAM$_+$ decreases to zero in a stable and consistent manner. In the noiseless case, the vanishing of the feasibility gap shown in Fig. 7 already confirmed the retrieval of the exact solution of Eq. (12). In the presence of noise, the steady decrease in the feasibility gap in Fig. 7 and the vanishing of the iterative change in Fig. 8 show that the algorithms find a local solution to that problem. For the same reason as in the analysis of the objective gap, the uniqueness of such a local solution is not guaranteed, and hence a global solution to Eq. (12) is not theoretically guaranteed to be obtained, although numerical experiments indicate so. Due to the additional information of the amplitude χ, VAM$_+$ converges slightly faster than VAM.

In summary, VAM and VAM$_+$ consistently exhibit desirable convergence properties in terms of the objective gap, feasibility gap, and iterative change for all the experiments conducted. In the noiseless setting, they are able to precisely restore the phase as indicated by the zero feasibility gap in Fig. 7 (solid curves). In the presence of noise, the two algorithms also clearly outperform SAM in terms of both convergence and phase retrieval quality. A particular example is presented in Fig. 9, where phase retrieval by VAM (bottom-left), VAM$_+$ (bottom-right), and SAM (top-right) are shown. The first two algorithms are about six times more accurate than the last one with relative RMS errors 2.4% and 2.2% compared to 15.2%, respectively. For completeness, four noiseless vectorial PSF images corresponding to the phase shown in Fig. 9 are presented in Fig. 10.

D. Number of Input Images

In this section, the effectiveness of the proposed algorithms with respect to the number of PSF images (m) is analyzed. Phase retrieval experiments with two, four, and six noisy PSF images are respectively conducted for 75 wavefront realizations.
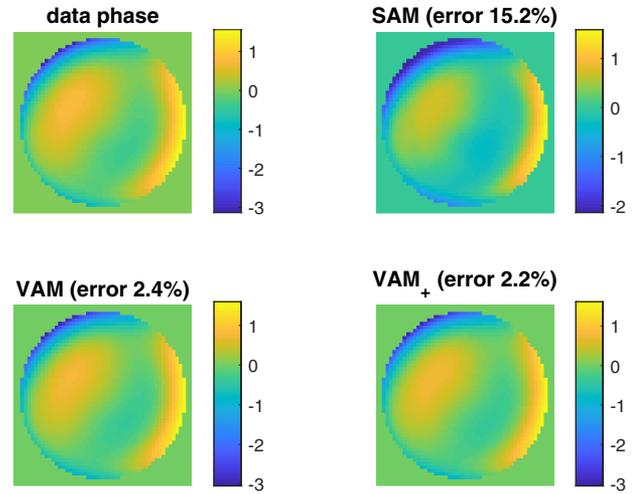
Fig. 9. Realization of phase retrieval by VAM (bottom-left), VAM$_+$ (bottom-right) and SAM (top-right).
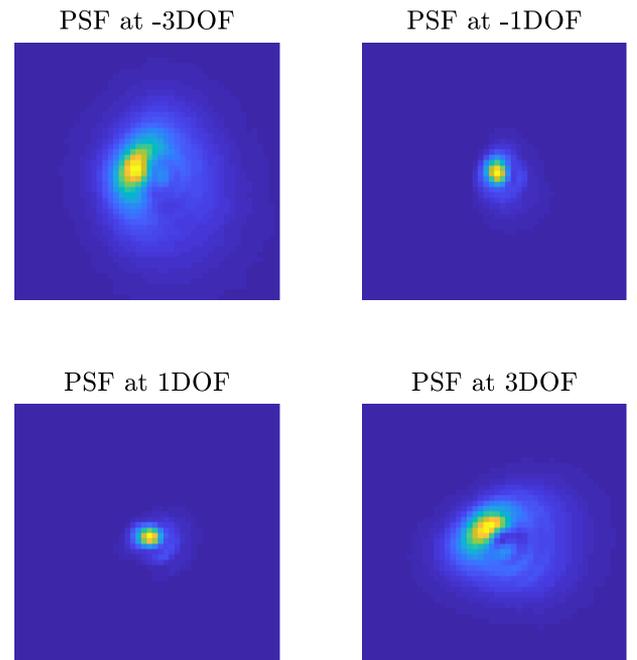
Fig. 10. Four noiseless out-of-focus PSFs corresponding to the phase presented in Fig. 9: the central part of each image with size 48×48 pixels is shown.

Figure 11 plots the relative RMS errors in iteration of VAM and VAM$_+$. The overall comparison between the two algorithms with the same number of input images is the same as in the previous sections. Figure 11 shows that phase retrieval is more accurate if more input images are used. For example, the relative RMS errors of VAM$_+$ (blue) given two (bold curve), four (dashed curve), and six (thin curve) images are, respectively, 8.3%, 1.6%, and 0.5% (on average). This is well explained by the widely known fact that additional measurements can be useful for suppressing noise. Note that the complexity of these Fourier-transform-based algorithms is approximately linearly proportional to the number of input images; hence, there is a
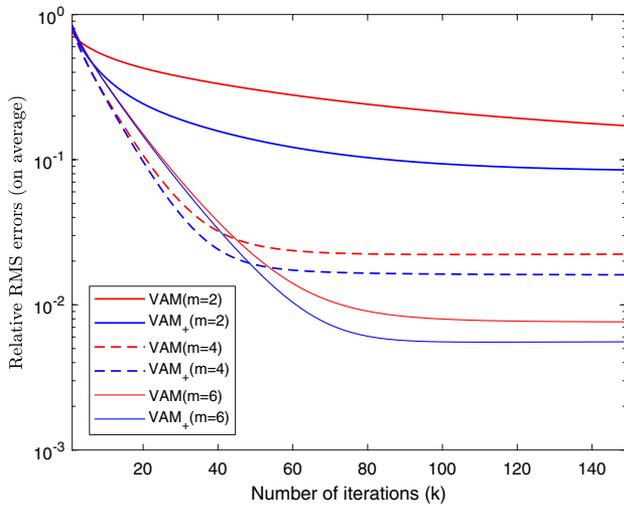
Fig. 11. Experiments show that more input images lead to more accuracy of phase retrieval by VAM and VAM$_+$.

Fig. 13. Experiment with noise demonstrates the solvability of VAM and VAM$_+$ for phase retrieval with discontinuous phase. VAM$_+$ (bottom-right) outperforms VAM (bottom-left) due to the knowledge of χ, while SAM (top-right) is more erroneous than the others due to model deviations with relative RMS error 9.4% compared to 5% and 4.3% of VAM and VAM$_+$.

trade-off between the computing time versus noise suppression. Moreover, accurately registering multiple PSF images is often not a trivial task in many applications.

The number of input images also affects the convergence speed of the algorithms, as shown in Fig. 12 where the *change* in the restored phase in iteration is plotted. Obviously, more input images lead to faster convergence of the algorithms. The additional knowledge of χ makes VAM$_+$ converge faster than VAM, e.g., VAM$_+$ with $m = 4$ (blue dashed curve) converges faster than VAM with $m = 6$ (red thin curve). It is important to keep in mind that faster convergence does not imply higher quality of restoration, e.g., VAM$_+$ with $m = 4$ is less accurate than VAM with $m = 6$, as shown in Fig. 11.

E. Phase Retrieval with Discontinuous Phase

This section demonstrates that VAM and VAM$_+$ are also efficient for phase retrieval with discontinuous phase, e.g., phase

Fig. 12. Experiments show that more input images lead to faster convergence of VAM and VAM$_+$.

retrieval with sparse phase constraint for applications in characterizing phase-only objects such as microlenses, phase-contrast microscopy, optical path difference microscopy, and in Fourier ptychography, where the phase object occupies less than 10% of the whole field [40]. As mentioned in the introduction, it is not feasible to accurately approximate a discontinuous wavefront or its associated GPF with a weighted sum of Zernike modes because the continuity property is invariant with respect to linear combination. More precisely, such an approximation would require a very high order of Zernike polynomials that are fast oscillating and impede the numerical calculations. As a consequence, solution methods based on modal formulations of phase retrieval (see, e.g., [41] for a comparison to the zonal formulation) are no longer relevant. In particular, the solution approach based on the use of ENZ expansion in [29] is not applicable to the class of problems analyzed in this section.

The discontinuous phase shown in Fig. 13 is obtained by adding a binary phase to a defocus term, and thus it is not sparse, i.e., its sampling is not sparse with respect to the pixel basis. The sparse phase shown in Fig. 14 is (a scaling of) a truncation of the above discontinuous phase. For experiments with these phases, VAM and VAM$_+$ precisely restore the exact solution in the noise-free setting, while SAM suffers model deviations. In the presence of noise, Figs. 13 and 14 show the phase retrieval results for the discontinuous and the sparse phases, respectively. The overall comparison among the three algorithms is the same as for the case of continuous phases analyzed in the previous sections. VAM$_+$ (bottom-right) outperforms VAM (bottom-left) due to the additional knowledge of χ, while SAM (top-right) is more erroneous than the others due to model deviations. This demonstrates the solvability of VAM and VAM$_+$ for this class of phase retrieval problems. These results also prove the advantage of our phase retrieval approach compared to the one in [29], as we claimed in the introduction. For discontinuous phases,
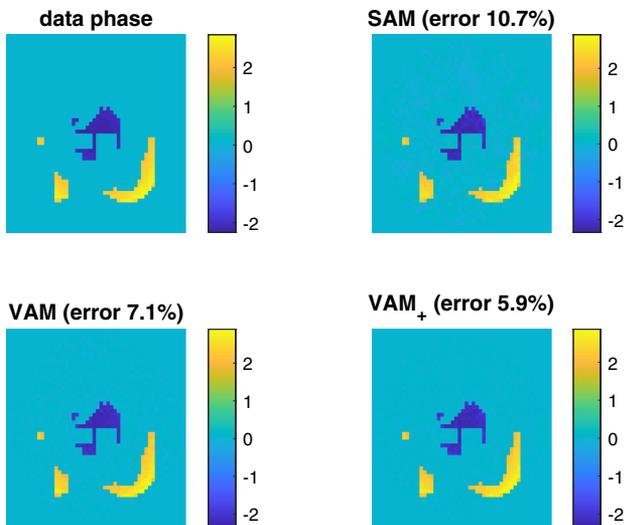
Fig. 14. Experiment with noise demonstrates the solvability of VAM and VAM$_+$ for phase retrieval with sparse phase constraint. The main features observed here resemble those in Fig. 13.

correction schemes after restoration (e.g., smoothness using Zernike polynomials) cannot be applied for noise suppression. Due to this disadvantage, the errors of VAM and VAM$_+$ in Figs. 13 (5% and 4.3%) and 14 (7.1% and 5.9%) are higher than the ones in Fig. 10 (2.4% and 2.2%).

F. Amplitude Restoration

The dominance of VAM$_+$ over VAM as observed so far demonstrates the positive influence of the additional information of χ on phase retrieval. This observation shows merely the effectiveness of χ whenever it is available; it does not imply any conclusion of comparing the two algorithms. They are two versions of the same phase retrieval approach reflexively adjusted for two scenarios of application, i.e., with known and unknown amplitude of the GPF. This section briefly analyzes the effectiveness of VAM in restoring the amplitude.

In the noise-free setting, the algorithm precisely restores the amplitude for all experiments regardless of the continuity of the phase. In the presence of noise, VAM enables to restore the amplitude with the same level of accuracy as for retrieving the phase. For example, Figs. 15, 16, and 17 show the amplitude restoration by VAM for the experiments presented in Figs. 10, 13, and 14, respectively. The restoration is highly accurate;
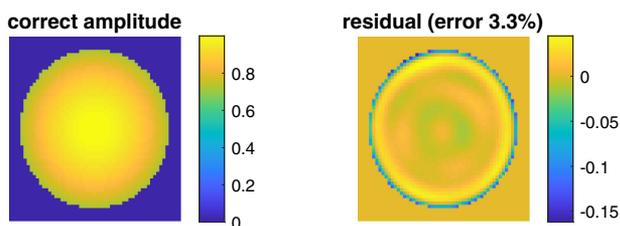
Fig. 15. Amplitude restoration by VAM for the experiment with continuous phase in Fig. 10: correct amplitude (left) and its residual (right) relative to the restoration (error 3.3%).
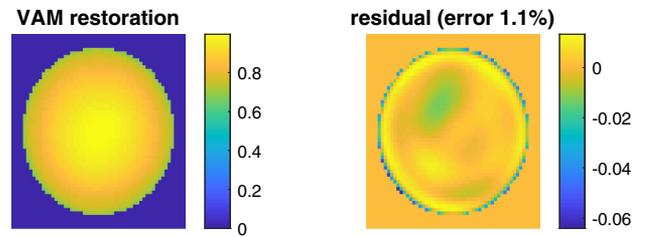
Fig. 16. Amplitude restoration by VAM for the experiment with discontinuous phase in Fig. 13: restoration (left) and its residual (right) relative to the data (error 1.1%).
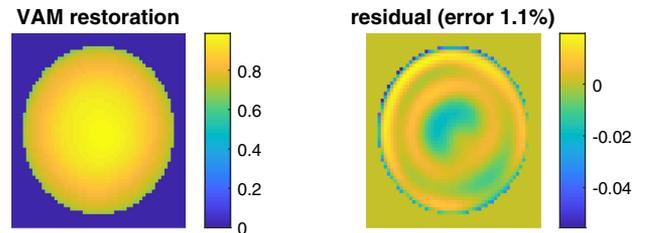
Fig. 17. Amplitude restoration by VAM for the experiment with sparse phase shown in Fig. 14: restoration (left) and its residual (right) relative to the data (error 1.1%).

the relative RMS errors are 3.3%, 1.1%, and 1.1%, respectively. It appears that the discontinuity of the wavefront has no substantial influence on the quality of amplitude restoration.

6. CONCLUDING REMARKS

We have presented and numerically analyzed a new phase retrieval approach for high-NA imaging systems based on the vectorial PSF model. The solution scheme can be adjusted for both cases of application with either known or unknown amplitude of the GPF. Making use of the correct imaging model, the proposed algorithms enable to restore phase aberrations at a high level of accuracy. They retrieve the correct solution in the noise-free setting and perform consistently with error around 5% for input data with SNR from 30 dB. Numerical results clearly demonstrate the efficiency and practicability of the proposed approach in various practically relevant simulations. Our approach exhibits several fundamental advantages over existing solution methods for phase retrieval in high-NA settings including [25], which is limited in terms of accuracy due to model deviations, and [31], which is expensive in terms of computational complexity and not applicable to problems with discontinuous wavefronts due to its modal-based approach.

Theoretical support for the phase retrieval approach presented in this paper is of great importance; however, we think it does not fit the scope of this paper and will be reported in a forthcoming mathematical-oriented publication.

REFERENCES

1. L. M. Mugnier, A. Blanc, and J. Idier, "Phase diversity: a technique for wave-front sensing and for diffraction-limited imaging," in *Advances in Imaging and Electron Physics*, P. Hawkes, ed. (Elsevier, 2006), Vol. **141**, pp. 1–76.
2. J. Antonello and M. Verhaegen, "Modal-based phase retrieval for adaptive optics," J. Opt. Soc. Am. A **32**, 1160–1170 (2015).
3. C. C. de Visser and M. Verhaegen, "Wavefront reconstruction in adaptive optics systems using nonlinear multivariate splines," J. Opt. Soc. Am. A **30**, 82–95 (2013).
4. C. C. de Visser, E. Brunner, and M. Verhaegen, "On distributed wave-front reconstruction for large-scale adaptive optics systems," J. Opt. Soc. Am. A **33**, 817–831 (2016).
5. S. R. Arridge, "Optical tomography in medical imaging," Inverse Probl. **15**, R41–R93 (1999).
6. T. Kim, R. Zhou, L. L. Goddard, and G. Popescu, "Solving inverse scattering problems in biological samples by quantitative phase imaging," Laser Photon. Rev. **10**, 13–39 (2016).
7. J. C. Dainty and J. R. Fienup, "Phase retrieval and image reconstruction for astronomy," Image Recovery: Theory Appl. **13**, 231–275 (1987).
8. J. W. Hardy and L. Thompson, "Adaptive optics for astronomical telescopes," Phys. Today **53**(4), 69 (2000).
9. R. P. Millane, "Phase retrieval in crystallography and optics," J. Opt. Soc. Am. A **7**, 394–411 (1990).
10. R. W. Harrison, "Phase problem in crystallography," J. Opt. Soc. Am. A **10**, 1046–1055 (1993).
11. J. W. Goodman, *Introduction to Fourier Optics* (Roberts & Company Publishers, 2005).
12. 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).
13. J. R. Fienup, "Phase retrieval algorithms: a comparison," Appl. Opt. **21**, 2758–2769 (1982).
14. H. Hauptman, "The direct methods of x-ray crystallography," Science **233**, 178–183 (1986).
15. H. H. Bauschke, P. L. Combettes, and D. R. Luke, "Phase retrieval, error reduction algorithm, and Fienup variants: a view from convex optimization," J. Opt. Soc. Am. A **19**, 1334–1345 (2002).
16. D. R. Luke, "Relaxed averaged alternating reflections for diffraction imaging," Inverse Probl. **21**, 37–50 (2005).
17. E. J. Candès, Y. C. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," SIAM J. Imaging Sci. **6**, 199–225 (2013).
18. D. Sayre, "Some implications of a theorem due to Shannon," Acta Crystallogr. Online **5**, 843 (1952).
19. J. R. Fienup, "Phase retrieval algorithms: a personal tour," Appl. Opt. **52**, 45–56 (2013).
20. Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: a contemporary overview," IEEE Signal Process. Mag. **32**, 87–109 (2015).
21. D. R. Luke, "Phase retrieval, what's new?" SIAG/OPT Views News **25**, 1–6 (2017).
22. R. Luke, S. Sabach, and M. Teboulle, "Optimization on spheres: models and proximal algorithms with computational performance comparisons," SIAM J. Math. Data Sci. **1**, 408–445 (2019).
23. B. Richards and E. Wolf, "Electromagnetic diffraction in optical systems. II. Structure of the image field in an aplanatic system," Proc. R. Soc. A **253**, 358–379 (1959).
24. M. Mansuripur, *Classical Optics and Its Applications* (Cambridge University, 2009).
25. B. M. Hanser, M. G. L. Gustafsson, D. A. Agard, and J. W. Sedat, "Phase retrieval for high-numerical-aperture optical systems," Opt. Lett. **28**, 801–803 (2003).
26. C. W. McCutchen, "Generalized aperture and the three-dimensional diffraction image," J. Opt. Soc. Am. **54**, 240–244 (1964).
27. B. M. Hanser, M. G. L. Gustafsson, D. A. Agard, and J. W. Sedat, "Phase-retrieved pupil functions in wide-field fluorescence microscopy," J. Microsc. **216**, 32–48 (2004).
28. N. Nakajima, "Phase retrieval from a high-numerical-aperture intensity distribution by use of an aperture-array filter," J. Opt. Soc. Am. A **26**, 2172–2180 (2009).
29. J. J. Braat, P. Dirksen, A. J. Janssen, S. van Haver, and A. S. van de Nes, "Extended Nijboer-Zernike approach to aberration and birefringence retrieval in a high-numerical-aperture optical system," J. Opt. Soc. Am. A **22**, 2635–2650 (2005).
30. S. van Haver, J. Braat, D. Peter, and A. Janssen, "High-NA aberration retrieval with the extended Nijboer-Zernike vector diffraction theory," J. Eur. Math. Soc. **1**, 06004 (2006).
31. J. J. M. Braat, P. Dirksen, A. J. E. M. Janssen, and A. S. van de Nes, "Extended Nijboer-Zernike representation of the vector field in the focal region of an aberrated high-aperture optical system," J. Opt. Soc. Am. A **20**, 2281–2292 (2003).
32. O. Raz, N. Dudovich, and B. Nadler, "Vectorial phase retrieval of 1D signals," IEEE Trans. Signal Process. **61**, 1632–1643 (2013).
33. C. W. McCutchen, "Generalized aperture and the three-dimensional diffraction image: erratum," J. Opt. Soc. Am. A **19**, 1721 (2002).
34. W. H. Southwell, "Wave-front analyzer using a maximum likelihood algorithm," J. Opt. Soc. Am. **67**, 396–399 (1977).
35. D. R. Luke, "Local linear convergence of approximate projections onto regularized sets," Nonlinear Anal. **75**, 1531–1546 (2012).
36. F. Soulez, É. Thiébaut, A. Schutz, A. Ferrari, F. Courbin, and M. Unser, "Proximity operators for phase retrieval," Appl. Opt. **55**, 7412–7421 (2016).
37. R. A. Gonsalves, "Phase retrieval and diversity in adaptive optics," Opt. Eng. **21**, 829–832 (1982).
38. J.-S. Pang, "Error bounds in mathematical programming," Math. Programming, Ser. B **79**, 299–332 (1997). Lectures on Mathematical Programming (ISMP97) (Lausanne, 1997).
39. D. R. Luke, N. H. Thao, and M. K. Tam, "Quantitative convergence analysis of iterated expansive, set-valued mappings," Math. Oper. Res. **43**, 1143–1176 (2018).
40. N. H. Thao, R. Luke, O. Soloviev, and M. Verhaegen, "Phase retrieval with sparse phase constraint," arXiv:1804.01878v2 (2018).
41. R. Doelman, N. H. Thao, and M. Verhaegen, "Solving large-scale general phase retrieval problems via a sequence of convex relaxations," J. Opt. Soc. Am. A **35**, 1410–1419 (2018).