THE EFFECT OF THE RANGE OF A MODULATING PHASE MASK ON THE RETRIEVAL OF A COMPLEX OBJECT FROM INTENSITY MEASUREMENTS

V. Karitans[1,2]*, M. Ozolinsh[1,2], A. Lapins[3], S. Fomins[1,2]

[1]Institute of Solid State Physics, University of Latvia
8 Kengaraga Str., Riga, LV-1063, LATVIA
[2]Department of Optometry and Vision Science, University of Latvia
1 Jelgavas Str., Riga, LV-1004, LATVIA
[3]Laser Centre, University of Latvia
3 Jelgavas Str., Riga, LV-1004, LATVIA
*e-mail: Varis.Karitans@cfi.lu.lv

In many fields of science, it is often impossible to preserve the information about the phase of the electromagnetic field, and only the information about the magnitude is available. This is known as the phase problem. Various algorithms have been proposed to recover the information about phase from intensity measurements. Nowadays, iterative algorithms of phase retrieval have become popular. Many of these algorithms are based on modulating the object under study with several masks and retrieving the missing information about the phase of an object by applying mathematical optimization methods. Several of these algorithms are able to retrieve not only the phase but also the magnitude of the object under study. In this study, we investigate the effect of the range of modulation of a mask on the accuracy of the retrieved magnitude and phase map. We conclude that there is a sharp boundary of the range of modulation separating the successfully retrieved magnitude and phase maps from those retrieved unsuccessfully. A decrease in the range of modulation affects the accuracy of the retrieved magnitude and phase map differently.

Keywords: *Coherent diffractive imaging, magnitude, phase retrieval, phase problem.*

1. INTRODUCTION

Due to the extremely rapid oscillations of the electromagnetic field [1], it is often impossible to preserve the information about the phase of the field, and only the information about the magnitude is available [2]. This is widely known as the phase problem and is encountered in many fields of science. The information about the phase of the diffracted field, however, is very important to retrieve the structure of the diffracting object [3]. In order to overcome the phase problem, various algorithms of phase retrieval have been proposed. Historically, the first algorithm proposed to retrieve the phase of an object from diffraction patterns was the Gerchberg-Saxton algorithm based on alternating projections [4]. Later, Fienup modified the Gerchberg-Saxton algorithm by replacing some of the constraints in the real domain with other constraints in accordance with the measured Fourier magnitude [5]. The Fienup algorithm is itself divided into two different subversions [6]. The Gerchberg-Saxton and Fienup algorithms were followed by the transport of intensity equation (TIE) [7]. In TIE, two intensity measurements at adjacent planes are sufficient to retrieve the phase of an object. TIE has also several versions like standard TIE, higher-order TIE and Gaussian process TIE.

Today, many iterative phase retrieval algorithms have been developed. An iterative algorithm PhaseLift requires modulating the object under study by a very limited number of masks and recovers the structure of the object from coded diffraction patterns under very noisy conditions [8], [9]. The original solution to PhaseLift is based on convex optimization; however, it was later optimized using low-rank Riemannian optimization methods [10]. PhaseLift formulates phase retrieval as finding a rank-one matrix. An algorithm PhaseCut [11] may be seen as a modification of PhaseLift. PhaseCut separates phase and magnitude variables and has been shown to be more stable than PhaseLift in the presence of noise. An algorithm SR-SPAR using sparsity in a complex object domain allows achieving superresolution about a quarter wavelength [12], while an algorithm GESPAR is able to recover a sparse object from intensity measurements and repeatedly updates the support of the object to avoid getting stuck [13]. While the aforementioned algorithms have been primarily developed to solve the phase problem, as the name suggests, several of them, e.g., PhaseLift, SPAR and GESPAR are also able to retrieve the complex field, i.e., the magnitude of the object along with the phase. As already mentioned, the algorithm PhaseCut separates the phase and magnitude of the object to be recovered. Generally, the magnitude retrieved by algorithms of phase retrieval has been studied minimally [14], [15].

In this study, we simulated the retrieval of a complex object with the algorithm PhaseLift and studied its accuracy depending on the range of modulation of pure phase masks. We chose the algorithm PhaseLift for its several advantages compared to other algorithms. As already mentioned, it requires a smaller number of modulating masks compared to other algorithms. Candes and Huang [8]–[10] showed that six modulating masks were sufficient for successful phase retrieval by PhaseLift. We have shown previously that four amplitude masks generated from a single mask are sufficient for reasonable phase retrieval [16], while other algorithms require all masks being different. PhaseLift is also able to retrieve the phase of an object from

very sparse data suggesting that imaging cameras of low bit depth are sufficient for capturing diffraction patterns. As far as the authors know, the performance of Phase-Lift has been studied for different types and number of modulating masks; however, the effect of the range of modulation on its performance has never been studied before.

2. METHODS

2.1. Phase Retrieval

A two-dimensional complex object to be retrieved is given by

$$x_0 \in \mathbb{C}^{n_1 x n_2}, \tag{1}$$

while its vectorized form is given by

$$x_0 \in \mathbb{C}^n, \tag{2}$$

where $n = n_1 \cdot n_2$.

The discrete Fourier transform (DFT) of x_0 is given by

$$\Im(x_0) = T \cdot x_0, \tag{3}$$

where the matrix T is the Kronecker product of the DFT matrix W, i.e.,

$$T = W \otimes W. \tag{4}$$

The size of the matrix W is n_1 by n_2. Let us also suppose we have l complex random modulating masks

$$d_i \in \mathbb{C}^{n_1 x n_2}, \tag{5}$$

where $i = 1, 2, ..., l$.

Let us denote the vectorized form of d_i by $d_i \in \mathbb{C}^n$ and point-wise multiplication by $\cdot$. The DFT of the modulated object $d_i \cdot x_0$ is given by

$$\Im(d_i \cdot x_0) = Z_i \cdot x_0, \tag{6}$$

where the matrix Z_i is given by an equation

$$Z_i = T \cdot \text{Diag}(d_i), \tag{7}$$

where $\text{Diag}(d_i)$ is an n by n matrix with the modulating waveform d_i on its main diagonal all other elements being zeros.

For the i-th mask, we have the modulus of the DFT of the modulated object in a vector b_{i_k} given by an equation

$$b_{i_k} = \left| \left\langle z_{i_k}, x_0 \right\rangle \right|, \tag{8}$$

where z_{i_k} is the k-th row of the matrix Z_i and the index of measurement $k = 1, 2, ..., n$.

Let us define a matrix Z for all l masks:

$$Z = \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_l \end{pmatrix}. \tag{9}$$

We can write

$$b_p = \left| \left\langle z_p, x_0 \right\rangle \right|, \tag{10}$$

where z_p is the p-th row of the matrix Z and the index of measurement $p = 1, 2, ..., m$. Here, $m = l \cdot n = l \cdot n_1 \cdot n_2$.

5

If we denote a vector containing all the moduli by b then we can write

$$b = \left| Z \cdot x_0 \right|. \tag{11}$$

As only intensity, i.e., the squared modulus of the DFT or Fraunhofer diffraction pattern, is detected, we have quadratic measurements of form

$$b_\bullet^2 = \text{diag}(Z \cdot x_0 \cdot x_0^* \cdot Z^*), \tag{12}$$

where $b_\bullet^2$ denotes the vector containing squared elements of the vector b, $\text{diag}(Z \cdot x_0 \cdot x_0^* \cdot Z^*)$ denotes the vectorial form of the matrix $Z \cdot x_0 \cdot x_0^* \cdot Z^*$ while the asterisk * denotes the conjugate transpose.

If we denote $x_0 \cdot x_0^*$ by X_0, the phase retrieval by PhaseLift can be formulated as finding a rank-one matrix [8]–[11]:

find

$$X_0$$

subject to

$$b_\bullet^2 = \text{diag}(Z \cdot x_0 \cdot x_0^* \cdot Z^*)$$
$$X_0 \succ= 0$$
$$\text{rank}(X_0) = 1$$

Here, X_0 is the rank-one solution and factorizing it in the form $x_0 \cdot x_0^*$ yields the solution to the phase retrieval problem. The rank minimization problem, however, is NP hard and in convex programming can be relaxed to trace minimization:

min

$$\text{trace}(X_0)$$

subject to

$$b_\bullet^2 = \text{diag}(Z \cdot x_0 \cdot x_0^* \cdot Z^*)$$
$$X_0 \succ= 0$$

Huang suggested using an alternate cost function due to the complexity associated with convex programming [10]. It is known that if the rank of X_0 is p, then there is $Y_p \in \mathbb{C}^{nxp}$ satisfying $Y_p \cdot Y_p^* = X_0$. Then the following cost function can be used:

$$f_r\left([Y_r]\right) = \frac{\left\| b_\bullet^2 - \text{diag}(Z \cdot Y_r \cdot Y_r^* \cdot Z^*) \right\|_2^2}{\left\| b_\bullet^2 \right\|_2^2} \tag{13}$$

PhaseLift optimizes the cost function in each iteration using one of several low-rank Riemannian optimization methods [10]. The algorithm was stopped when the norm of the gradient of the cost function fell below 10^{-5}.

2.2. Simulations

Complex objects of unity large magnitude maps and distorted phase maps were used as test objects in the simulations (see Fig. 1 left). The phase maps of the objects were scaled versions of each other. One of the objects was selected as the original object, while the phase maps of other objects were obtained by multiplying the phase map of the original object with a scaling factor s taking values 0.5, 1.5, 2.0., and 2.5. In order to generate the phase map of the original object, random Zernike coefficients in the range from -1 μm to 1 μm up to the 6th order were generated. The units of phase distortions were radians. The size of the objects were 256 by 256 pixels.
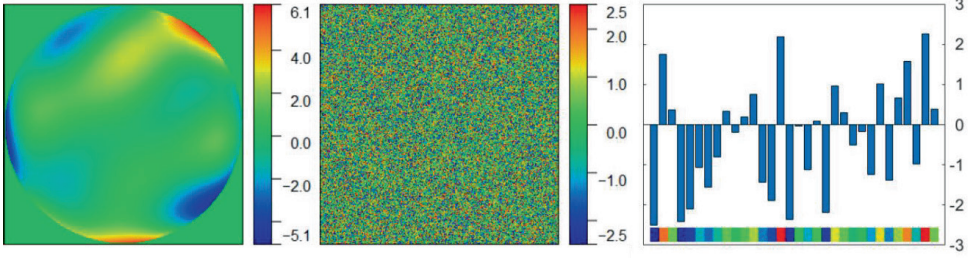
Fig. 1. Left – one of the objects used in the simulations. The scale on the right side is given in radians. Middle – one of the phase masks with the range of modulation = 0.8. The scale on the right side is given in radians. Right – the surface profile of a cut-out of the phase mask. The length of the cut-out is 16 pixels.

The root-mean-square error of the phase map of an object $RMSE_{po}$ was calculated according to an equation

$$RMSE_{po} = \sqrt{\frac{\sum\limits_{i=1}^{n}(\overline{\varphi_0} - \varphi_{0i})^2}{n}} , \qquad (14)$$

where $\overline{\varphi_0}$ is the average value of all data points of a phase map φ_0, φ_{0i} is the value of the i-th data point, and n is the number of data points.

In order to assess the quality of the retrieved magnitude map A_r, the root-mean-square error of the residual magnitude map of an object $RMSE_{mres}$ was calculated according to an equation

$$RMSE_{mres} = \sqrt{\frac{\sum\limits_{i=1}^{n}\left(\left(\overline{A_0 - A_r}\right) - \left(A_0 - A_r\right)_i\right)^2}{n}} , \qquad (15)$$

where $\overline{A_0 - A_r}$ is the average value of all data points of the residual magnitude map $A_0 - A_r$, $\left(A_0 - A_r\right)_i$ is the value of the i-th data point, and n is the number of data points.

In order to assess the quality of the retrieved phase map, the root-mean-square error of the residual phase map of an object $RMSE_{pres}$ was calculated according to an equation

$$RMSE_{pres} = \sqrt{\frac{\sum\limits_{i=1}^{n}\left(\left(\overline{\varphi_0 - \varphi_r}\right) - \left(\varphi_0 - \varphi_r\right)_i\right)^2}{n}} , \qquad (16)$$

where $\overline{\varphi_0 - \varphi_r}$ is the average value of all data points of the residual phase map $\varphi_0 - \varphi_r$, $\left(\varphi_0 - \varphi_r\right)_i$ is the value of the i-th data point, and n is the number of data points.

In order to compare the $RMSE_{pres}$ among objects with different $RMSE_{po}$, the $RMSE_{pres}$ was divided by the respective $RMSE_{po}$.

The masks used to modulate the object were of the same size as the object. The masks were random pure phase masks modulating the phase of the object at each point by a random value within a certain range (see Fig. 1 right). The inset shown on the right side of Fig. 1 is a region of the phase mask contained within the white square. Here, we introduce the range of modulation of a mask denoted by M and defined as

$$M = \frac{\varphi_{\max} - \varphi_{\min}}{2 \cdot \pi} , \qquad (17)$$

where $\varphi_{\max}$ and $\varphi_{\min}$ are the maximum and minimum value of the phase mask, respectively, within the range from -π to +π.

The range of modulation M varied between 0.1 and 0.8 in steps of 0.1. The selected range of modulation was based on

a spatial light modulator (SLM) available in our laboratory guaranteeing that the range of phase modulation M as the results of simulations indicate the achievable accuracy of the retrieved magnitude and phase maps of an object. The size of a pixel of the SLM is 8 μm so that an object and mask of size 256 by 256 pixels correspond to the lin-ear size 2048 by 2048 μm. The maximum phase shift that can be introduced by the SLM is 2.9·π radians at the wavelength of a He-Ne laser (λ = 0.6328 μm corresponding to 2·π radians). The peak-to-valley value of the original object was about 11.2 radians corresponding to about 1.1 μm at the wavelength of a He-Ne laser.

3. RESULTS

Figure 2 shows the retrieved magnitude and phase maps of the object with the scaling factor s = 2 for all ranges of modulation M. The range of modulation M increases from left to right and from top to bottom. It can be easily noticed that the range of modulation M of at least 60 % of a full wave is required for a success-ful retrieval of the magnitude map of the selected object. A sharp boundary separating successfully retrieved magnitude maps from those retrieved unsuccessfully was characteristic of all objects. The values of the magnitude maps are shown in arbitrary units and have been normalized so that the maximum value equals unity.
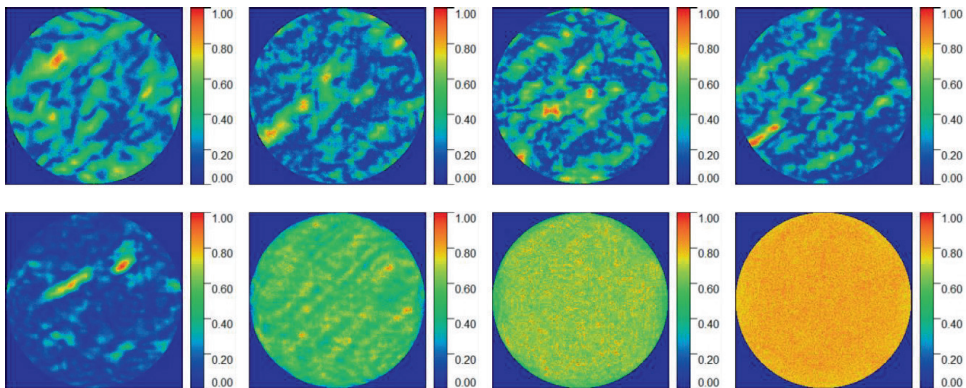

Fig. 2. The retrieved magnitude maps of the object with the scaling factor s = 2 for all ranges of modulation M changing from 0.1 to 0.8 in steps of 0.1. The range of modulation M increases from left to right and from top to bottom. One can notice the sharp boundary separating successfully retrieved magnitude maps from those retrieved unsuccessfully.

Figure 3 shows the retrieved phase maps of the same object with the scaling factor s = 2 for all ranges of modulation M. The retrieved phase maps were wrapped within the range from -π to +π and had to be unwrapped using an algorithm freely available on the Internet. The retrieved phase map is also comparable to the original one starting from the same range of modulation M, i.e., 60 %. Below this level, there is no similarity between the retrieved phase map and the object. Generally, the minimum values of the range of modulation M for successful retrieval of the magnitude and phase map are always equal given one and the same object. The units of phase maps are given in radians.
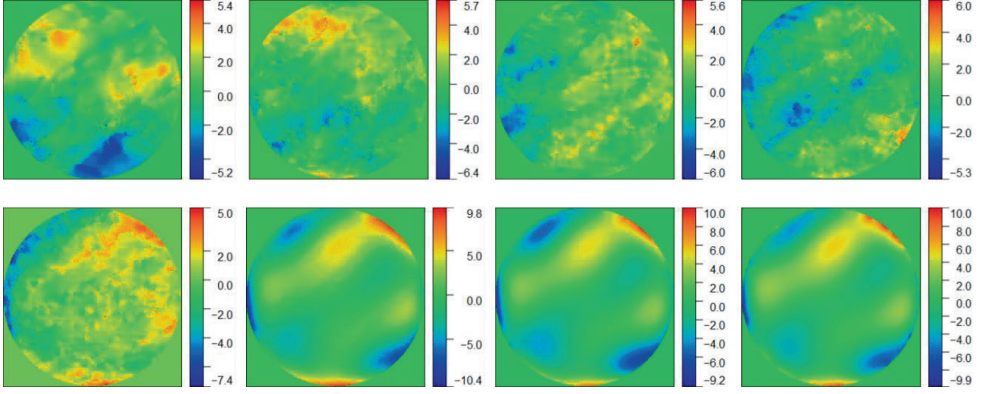
Fig. 3. The retrieved phase maps of the object with the scaling factor $s = 2$ for all ranges of modulation M changing from 0.1 to 0.8 in steps of 0.1. The range of modulation M increases from left to right and from top to bottom. One can notice the sharp boundary separating successfully retrieved phase maps from those retrieved unsuccessfully.

Figure 4 shows the $RMSE_{mres}$ (left) and the normalized $RMSE_{pres}$ (right) for all ranges of modulation M. The error bars are standard errors calculated from the results of five simulations. The labels top right show the $RMSE_{po}$ of the objects. The $RMSE_{pres}$ was normalized by dividing it with the corresponding $RMSE_{po}$. It can be easily noticed that the $RMSE_{mres}$ grows for all objects as the range of modulation M is reduced while three distinct regions of the $RMSE_{pres}$ can be noted. For the largest range of modulation M, i.e., 0.8 of a full wave, the $RMSE_{pres}$ of all objects is low except the object having the smallest $RMSE_{po}$. Starting from a certain value of

the range of modulation M, the $RMSE_{pres}$ starts to increase rapidly as the range of modulation M is decreased. The exact range of modulation M at which the $RMSE_{pres}$ starts to grow rapidly depends on the $RMSE_{po}$. It should also be noted that the highest $RMSE_{pres}$ also depends on the $RMSE_{po}$, i.e., the $RMSE_{pres}$ was higher for objects having small $RMSE_{po}$. As the range of modulation M is further reduced, the $RMSE_{pres}$ of objects having small $RMSE_{po}$ remained constant while the $RMSE_{pres}$ of objects having large $RMSE_{po}$ started to fall reaching the level of objects having small $RMSE_{po}$.
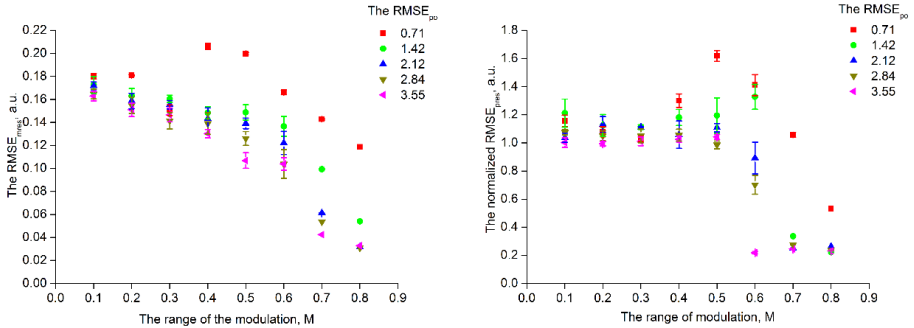

Fig. 4. Dependence of the $RMSE_{mres}$ (left) and $RMSE_{pres}$ (right) on the range of modulation M. Generally, the growth of the $RMSE_{mres}$ is similar for all objects except the object with the smallest $RMSE_{po}$. The $RMSE_{pres}$ is always low for large ranges of modulation M irrespective of $RMSE_{po}$. For medium ranges of modulation M, the $RMSE_{pres}$ strongly depends on the $RMSE_{po}$ while for low small ranges of modulation M the $RMSE_{pres}$ is approximately constant and does not vary considerably with the $RMSE_{po}$.

4. DISCUSSION AND CONCLUSIONS

The results suggest that the retrieval of the magnitude and phase of an object is compromised differently as the range of modulation is reduced. Differences between the changes in the $RMSE_{mres}$ and $RMSE_{pres}$ along with the changes in the range of modulation M may be explained by separating the magnitude and phase of an object and analysing the spatial spectrum of both components.

The magnitude map being completely flat contains predominantly low spatial frequencies, and hence the intensity of high spatial frequencies strongly depends on the range of modulation M. Due to zeroing the high spatial frequencies because of the low bit depth, the amount of information carried by the high spatial frequencies decreases and the $RMSE_{mres}$ increases as the range of modulation M decreases. For all objects, except the object with the smallest $RMSE_{po}$, the effect of the $RMSE_{po}$ on the $RMSE_{mres}$ seems negligible.

The analysis of the effect of the range of modulation M on the $RMSE_{pres}$ is more complicated. In the region, where the range of modulation M is low, high spatial frequencies are of low intensity irrespective of the $RMSE_{po}$ and are forced to zero due to the low bit depth. The bright central part occupies about the same region of the diffraction pattern irrespective of the $RMSE_{po}$, and therefore the normalized $RMSE_{pres}$ is approximately constant. In the region, where the range of modulation M is large, high spatial frequencies are intense for all objects despite the low bit depth ensuring favourable conditions for the algorithm to converge to a solution.

One can also note that smaller ranges of modulation M are required for objects of large $RMSE_{po}$ compared to objects of smaller $RMSE_{po}$. The middle region is a transition between the two extremes where zeroing the high spatial frequencies depends both on the M and the $RMSE_{po}$. If the $RMSE_{po}$ is small, the high spatial frequencies are still forced to zero because of the low bit depth despite that the range of modulation M is moderate; however, for objects with higher $RMSE_{po}$ high spatial frequencies are non-zero. Particularly large $RMSE_{pres}$ of the object with the smallest $RMSE_{po}$ suggests that the probability of unsuccessful phase retrieval from very sparse intensity measurements increases as the range of modulation M increases given that the sparsity of the intensity measurements remains the same. To summarise, the effect of the range of modulation M dominates when reaching the extreme values while at moderate levels of the range of modulation M the contents of spatial frequencies of the object dictate the $RMSE_{pres}$.

Finally, we conclude that increasing the bit depth of intensity measurements may lower the minimum the range of modulation M sufficient for successful retrieval of a complex object. It can also be expected that the probability of retrieving an object with small $RMSE_{po}$ increases as the bit depth is increased given that the range of modulation M remains the same. Objects having particularly large $RMSE_{po}$ may even be retrieved from diffraction patterns of very low bit depth.

Ability to retrieve phase maps having large $RMSE_{po}$ from very sparse intensity measurements is very important in astronomy where imaging the celestial objects is compromised by very turbulent atmospheric layers [17]. Studying the effect of bit depth on the quality of phase and mag-

nitude retrieval constitutes a wide scope of research, and we plan to address this question in future. It can also be expected that oversampling the diffraction patterns may lower the requirements of the minimum necessary range of modulation M.

ACKNOWLEDGEMENTS

REFERENCES

1. Shechtman, Y., Eldar, Y.C., Cohen, O., Chapman, H. N., Miao, J., & Segev, M. (2015). Phase Retrieval with Application to Optical Imaging. *IEEE Signal Processing Magazine*, *32* (3), 87–109. DOI: 10.1109/MSP.2014.2352673

2. Zuo, C., Li, J., Sun, J., Fan, Y., Zhang, J., Lu, L., ... & Chen, Q. (2020). Transport of Intensity Equation: A Tutorial. *Optics and Lasers in Engineering*, *135*, 106187. DOI: 10.1016/j.optlaseng.2020.106187

3. Tao, S., He. C., Hao, X., Kuang, C., & Liu, X. (2021). Principles of Different X-ray Phase-Contrast Imaging: A Review. *Applied Sciences (Switzerland)*, *11* (7), 2971. DOI: 10.3390/app11072971

4. Gerchberg, R. W., & Saxton, W.O. (1972). Practical Algorithm for the Determination of Phase from Image and Diffraction Plane Pictures. *Optik (Stuttgart)*, *35* (2), 237–250.

5. Fienup, J. R. (1982). Phase Retrieval Algorithms: A Comparison. *Applied Optics*, *21* (15), 2758–2768. DOI: 10.1364/AO.21.002758

6. Latychevskaia, T. (2018). Iterative Phase Retrieval in Coherent Diffractive Imaging: Practical Issues. *Applied Optics*, *57* (25), 7187–7197. DOI: 10.1364/AO.57.007187

7. Huang, L., Zuo, C., Idir, M., Qu, W., & Asundi, A. (2015). Phase retrieval in arbitrarily shaped aperture with the transport-of-intensity equation. In *SPIE/IS&T Electronic Imaging* (94010B), 8–12 February 2015, San Francisco, California, United States: SPIE Digital Library. DOI: 10.1117/12.2087754

8. Candès, E. J., Eldar, Y. C., Strohmer, T., & Voroninski, V. (2013). Phase Retrieval via Matrix Completion. *SIAM Journal on Imaging Sciences*, *6* (1), 199–225. DOI: 10.1137/110848074

9. Candès, E. J., Li, X., & Soltanolkotabi, M. (2015). Phase Retrieval from Coded Diffraction Patterns. *Applied and Computational Harmonic Analysis*, *39* (2), 277–299. DOI: 10.1016/j.acha.2014.09.004

10. Huang, W., Gallivan, K. A., & Zhang, X. (2017). Solving PhaseLift by Low-rank Riemannian Optimization Methods for Complex Semidefinite Constraints. *SIAM Journal on Scientific Computing*, *39* (5), B840–B859. DOI: 10.1137/16M1072838

11. Waldspurger, I., d'Aspremont, A., & Mallat, S. (2015) Phase Recovery, MaxCut and Complex Semidefinite Programming. *Mathematical Programming*, *149* (1–2), 47–81. DOI: 10.1007/s10107-013-0738-9

12. Katkovnik, V., & Egiazarian, K. (2017). Sparse Superresolution Phase Retrieval from Phase-coded Noisy Intensity Patterns. *Optical Engineering*, *56* (9), 094103. DOI: 10.1117/1.OE.56.9.094103

13. Shechtman, Y., Beck, A., & Eldar, Y. C. (2014). GESPAR: Efficient Phase Retrieval of Sparse Signals. *IEEE Transactions on Signal Processing*, *62* (4), 928–938. DOI: 10.1109/TSP.2013.2297687

14. Kumar Singh, R., Vinu, R. V., & Sharma M., A. (2014). Recovery of Complex Valued Objects from Two-point Intensity Correlation Measurement. *Applied Physics Letters*, *104* (11), 111108. DOI: 10.1063/1.4869123

15. Knapp, J., Paulus, A., Komprobst, J., Siart, U., & Eibert, T. F. (2021). Multifrequency Phase Retrieval for Antenna Measurements. *IEEE Transactions on Antennas and Propagation*, *69* (1), 488-501. DOI: 10.1109/TAP.2020.3008648

16. Karitans, V., Nitiss, E., Tokmakovs, A., Ozolinsh, M., & Logina, S. (2019). Optical Phase Retrieval Using Four Rotated Versions of a Single Binary Amplitude Modulating Mask. *Journal of Astronomical Telescopes, Instruments, and Systems*, *5* (3), 039004. DOI: 10.1117/1.JATIS.5.3.039004

17. Burke, D., Devaney, N., Christof, J., & Hartung, M. (2010). Application of wavelength diversity for astronomical adaptive optics imaging. In *SPIE Astronomical Telescopes + Instrumentation*, (77365U), 27 June–2 July 2010, San Diego, California, United States: SPIE Digital Library. DOI: 10.1117/12.856953