SIAM J. MATH. DATA SCI.
Vol. 2, No. 1, pp. 246–263

Phase Retrieval with Sparse Phase Constraint*

Nguyen Hieu Thao†, David Russell Luke‡, Oleg Soloviev§, and Michel Verhaegen¶

Abstract. In this paper, we propose and investigate the phase retrieval problem with the a priori constraint that the phase is sparse (SPR), which encompasses a number of practical applications, for instance, in characterizing phase-only objects such as microlenses, in phase-contrast microscopy, in optical path difference microscopy, and in Fourier ptychography, where the phase object occupies a small portion of the whole field. The considered problem is strictly more general than the sparse signal recovery problem, which assumes the sparsity of the signal because the sparsity of the signal trivially implies the sparsity of the phase, but the converse is not true. As a result, existing solution algorithms in the literature of sparse signal recovery cannot be applied to SPR and there is an appeal for developing new solution methods for it. In this paper, we propose a new regularization scheme which efficiently captures the sparsity constraint of SPR. The idea behind the proposed approach is to perform a metric projection of the current estimated signal onto the set of all the signals whose phase satisfies the sparsity constraint. The main challenge here is that the latter set is not convex and its associated projector in general does not admit a closed form. One novelty of our analysis is to establish an explicit form of that projector when restricted to those points which are relevant to the solutions of SPR. Note that this result is fundamentally different from the widely known calculation form for projections onto intensity constraint sets. Based on this new result, we propose an efficient solution method, named the sparsity regularization on phase (SROP) algorithm, for the SPR problem in the challenging setting where only one point-spread-function image is given, and we analyze its convergence. The algorithm is the combination of the Gerchberg–Saxton (GS) algorithm with the projection step described above. In view of the GS algorithm being equivalent to the alternating projection for an associated two-set feasibility, the SROP algorithm is shown to be the cyclic projection for an associated three-set feasibility, one of the sets being analyzed in this paper for the first time. Analyzing regularity properties of the involved sets, we obtain convergence results for the SROP algorithm based on our recent convergence theory for the cyclic projection method. Numerical results show clear effectiveness and efficiency of the proposed solution approach for the SPR problem.

Key words. phase retrieval, sparsity constraint, nonconvex optimization, prox-regularity, projection algorithm, linear convergence

*Received by the editors June 6, 2019; accepted for publication (in revised form) January 10, 2020; published electronically March 25, 2020.

https://doi.org/10.1137/19M1266800

Funding: This work was supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013)/ERC grant agreement N° 339681. The second author's research was supported in part by German Israeli Foundation grant G-1253-304.6 and Deutsche Forschungsgemeinschaft Collaborative Research Center SFB755.

†Delft Center for Systems and Control, Delft University of Technology, 2628 CD Delft, The Netherlands, and Department of Mathematics, School of Education, Can Tho University, Can Tho, Vietnam (h.t.nguyen-3@tudelft.nl, nhthao@ctu.edu.vn).

‡Institut für Numerische und Angewandte Mathematik, Universität Göttingen, 37083 Göttingen, Germany (r.luke@math.uni-goettingen.de).

§Delft Center for Systems and Control, Delft University of Technology, 2628 CD Delft, The Netherlands, and Flexible Optical B.V., 2288 GG Rijswijk, The Netherlands (o.a.soloviev@tudelft.nl).

¶Delft Center for Systems and Control, Delft University of Technology, 2628 CD Delft, The Netherlands (m.verhaegen@tudelft.nl).

AMS subject classifications. 65K15, 65Z05, 65T50, 78A45, 78A46, 90C26

DOI. 10.1137/19M1266800

1. Introduction.

Phase retrieval is an inverse problem of recovering a complex signal from one or several measured intensity images. It appears in many scientific and engineering fields, including astronomy, crystallography, microscopy, optical manufacturing, and adaptive optics [5, 6, 15, 24]. An important application of phase retrieval is to quantify the properties of an optical system via its generalized pupil function (GPF). The fundamental advantage of this approach compared to those using intensity point spread functions (PSFs) or intensity optical transfer functions is that it is modifiable and automatically includes specific characterizations of the imaging system under investigation.

Since PSF images capture the complex-valued GPF via the squared amplitude of its Fourier transform, several PSF images are, on the one hand, usually needed for retrieving the phase of the GPF. On the other hand, simultaneously measuring multiple PSF images is not a simple task in many practical applications, especially for real time imaging systems. A reason is that the use of double exposure procedures for such a measurement process would inevitably generate unwanted noise to the obtained data. Fortunately, under certain circumstances the phase can be estimated from only one PSF image; see, for example, [3, 5, 7]. In this paper, we investigate the phase retrieval problem, given a single PSF image, with a priori information that the phase is sparse. More specifically, we study the *phase retrieval problem with sparse phase constraint* (SPR) as follows:

$$\begin{aligned} &\text{finding} \quad x = \chi \cdot \mathrm{e}^{\mathrm{j}\Phi} \in \mathbb{C}^{n\times n} \\ &\text{such that} \quad \left|\mathcal{F}\left(x \cdot \mathrm{e}^{\mathrm{j}\Psi}\right)\right|^2 = b + w \quad \text{and} \quad \|\Phi\|_0 \le s, \end{aligned} \tag{1.1}$$

where $\chi \in \mathbb{R}_+^{n\times n}$ is the known amplitude of the GPF, $\Phi \in (-\pi,\pi]^{n\times n}$ is the unknown phase variable, $\mathcal{F}$ is the (discrete) 2-dimensional Fourier transform, $\Psi \in (-\pi,\pi]^{n\times n}$ is a known phase diversity term,[1] $b \in \mathbb{R}_+^{n\times n}$ is the measured PSF image, $w \in \mathbb{R}^{n\times n}$ is unknown noise, and s is an a priori upper bound of the sparsity level of Φ. In (1.1) and elsewhere in the paper, the L0-norm $\|\cdot\|_0$ of a complex- or real-valued matrix is the number of its nonzero entries. The equality, the multiplication, the exponential, the modulus, and the square operations are understood elementwise. The imaginary unit is denoted by j. To capture the case that the GPF is nonzero on the entire aperture of the imaging system, we assume that χ is nonzero everywhere in the aperture.

The above model encompasses a number of important applications, for instance, in characterizing phase-only objects such as microlenses, in phase-contrast microscopy, in optical path difference microscopy, and in Fourier ptychography, where the phase object (for example, blood cells) occupies less than 10% of the whole optical field. Let us clarify the difference between the SPR problem (1.1) and the *sparse signal recovery* (SSR) problem, which is widely

[1] This term plays an important role in application, but it brings no difference to (1.1) in terms of mathematics.

known and has been well investigated in the literature; see, for example, [19, 20, 23]:

$$\text{(1.2)} \qquad \begin{aligned} &\text{finding} \quad x = \chi \cdot \mathrm{e}^{\mathrm{j}\Phi} \in \mathbb{C}^{n\times n} \\ &\qquad\quad \text{such that} \quad \left|\mathcal{F}\left(x\cdot \mathrm{e}^{\mathrm{j}\Psi}\right)\right|^2 = b + w \quad \text{and} \quad \|x\|_0 \le s. \end{aligned}$$

The difference between the two problem models is that (1.1) captures the sparsity property of the phase Φ while (1.2) captures the sparsity property of the signal x. On the one hand, the condition $\|x\|_0 \le s$ implies the condition $\|\Phi\|_0 \le s$ since Φ is the (elementwise) argument[2] of x; in other words, the sparsity constraint in (1.2) is stronger than the one in (1.1). Thus, the problem (1.1) covers the problem (1.2) as a special case. On the other hand, the sparsity of the phase Φ does not imply any sparsity property of the signal x, for instance, when the amplitude χ is nonzero everywhere in the aperture, as is the case considered in this paper. This means that the problem (1.1) is in general not covered by the problem (1.2), and hence the former one is strictly more general than the latter one. As a consequence, existing solution algorithms for sparse signal recovery, for example, [19, 20, 23], are not applicable to the SPR problem (1.1). This fact has motivated the development of new efficient schemes for addressing the new problem (1.1), and the analysis of this paper is devoted to this task. To avoid possible confusion in terms of terminology, we mention that the SSR problem (1.2) is also referred to as sparse phase retrieval elsewhere in the literature.

Inspired by the success of the sparsity regularization scheme incorporated into projection algorithms for solving the SSR problem [19] and the *affine sparse feasibility* problem [2, 9, 25], in this paper we develop a new efficient regularization scheme for capturing the sparsity property of the phase in the setting of the SPR problem (1.1). The idea is to project the temporal estimated signal onto the set of all the signals which satisfy both the amplitude modulation and the sparse phase constraint. The main challenge is that the latter set is not convex and its associated projector in general does not admit a closed form, and one novelty of this paper is an explicit form of that projector when restricted to those points which are relevant to the solutions of the SPR problem (1.1); see Lemma 2.1. It is worth noting that this result is fundamentally different from the widely known formula for projections onto intensity or amplitude constraint sets; see, for example, [1, 15].

The remainder of this paper is organized as follows. Other mathematical notation will be introduced in the rest of this section. In section 2, we present a new regularization scheme for capturing the sparsity property of the phase in the setting of the SPR problem (1.1). In section 3, we introduce a natural and efficient solution method for (1.1), named the sparsity regularization on phase (SROP) algorithm, based on the classical Gerchberg–Saxton (GS) algorithm and the regularization scheme proposed in section 2. In section 4, the SROP algorithm is proved to be equivalent to the cyclic projection algorithm for solving an associated feasibility problem involving three sets—two amplitude constraint sets and the set describing both the amplitude modulation and the sparsity of the phase defined in section 2. Based on the analysis of the cyclic projection for nonconvex feasibility recently established in [17], we obtain convergence results for the SROP algorithm by analyzing regularity properties of the three component sets. In section 5, we numerically demonstrate the effectiveness and efficiency of

[2]Without loss of generality, zero can always be taken as the argument of itself.

the proposed regularization scheme by showing substantial advantages of the SROP algorithm over the corresponding algorithm but without the additional regularization step. The latter algorithm is nothing else but the GS algorithm [7]. Since the SPR problem (1.1) is proposed and investigated in this paper for the first time, existing phase retrieval algorithms are not relevant to it. The goal of section 5 is to demonstrate the effectiveness and efficiency of the new regularization scheme rather than to compare the SROP algorithm with the GS method.

Mathematical notation. The underlying space in this paper is the Hilbert space $\mathbb{C}^{n\times n}$. The Frobenius norm and the maximum norm are denoted by $\|\cdot\|$ and $\|\cdot\|_\infty$, respectively. Mathematical operations including the multiplication, the division, the modulus, the argument,[3] the exponential, the square, the square root, and the equality are understood elementwise in this paper. The distance function associated to a set $\Omega \subset \mathbb{C}^{n\times n}$ is defined by

$$\mathrm{dist}(\cdot,\Omega)\colon \mathbb{C}^{n\times n} \to \mathbb{R}_+ \colon x \mapsto \inf_{w\in\Omega} \|x-w\|,$$

and the set-valued mapping

$$P_\Omega : \mathbb{C}^{n\times n} \rightrightarrows \Omega : x \mapsto \{w \in \Omega \mid \|x-w\| = \mathrm{dist}(x,\Omega)\}$$

is the corresponding *projector*. A selection $w \in P_\Omega(x)$ is called a *projection* of x on Ω. Since only projections on compact sets are involved in the analysis of this paper, the existence of projections is always guaranteed. The fixed point set of a possibly set-valued mapping $T : \mathbb{C}^{n\times n} \rightrightarrows \mathbb{C}^{n\times n}$ is defined by $\mathrm{Fix}\, T \equiv \{x \in \mathbb{C}^{n\times n} \mid x \in T(x)\}$; see, for example, [17, Definition 2.1]. An iterative sequence $x_{k+1} \in T(x_k)$ generated by T is said to *converge linearly* to a point x with rate $c \in [0,1)$ if there exists a constant $\gamma > 0$ such that

$$\|x_k - x\| \le \gamma c^k \quad \forall k \in \mathbb{N}.$$

Dealing with matrices of size $n \times n$, we frequently use the set of all indices

$$\mathcal{J} \equiv \{\xi = (\xi_1,\xi_2) \in \mathbb{N}^2 \mid 1 \le \xi_1, \xi_2 \le n\} \tag{1.3}$$

and the set of all s-element subsets of $\mathcal{J}$ given by

$$\mathcal{J}_s \equiv \{J \subset \mathcal{J} \mid J \text{ has } s \text{ elements}\}.$$

For a real- or complex-valued matrix $\mathcal{O}$ of size $n \times n$, we define the set (see [2, equation (33)] or [9, equation (17)])

$$C_s(\mathcal{O}) \equiv \left\{ J \in \mathcal{J}_s \mid \min_{\xi\in J} |\mathcal{O}_\xi| \ge \max_{\xi\in\mathcal{J}\setminus J} |\mathcal{O}_\xi| \right\}. \tag{1.4}$$

In (1.4) and elsewhere in the paper, the entry at index ξ of a matrix $\mathcal{O}$ is denoted by $\mathcal{O}_\xi$. Our other basic notation is standard; cf. [4, 18, 22]. The open ball in $\mathbb{C}^{n\times n}$ with radius $\delta > 0$ and center x is denoted by $\mathbb{B}_\delta(x)$.

[3]For a complex number $z = re^{\mathrm{j}\varphi}$, where $r \in \mathbb{R}_+$ and $\varphi \in (-\pi,\pi]$ (using Euler's formula), its modulus and argument are $|z| = r$ and $\arg(z) = \varphi$, respectively. By convention, $\arg(z) = 0$ for $z = 0$.

For simplicity in terms of terminology, when $x = \chi \cdot \mathrm{e}^{\mathrm{j}\Phi}$ is a solution to (1.1), Φ is accordingly called a *phase solution* to that problem. Without loss of generality, we can assume that the amplitude χ is strictly positive everywhere.[4]

2. Regularization for sparse phase constraint.

This section presents the novel idea of this paper. In the literature of sparse signal recovery and sparse optimization, for example, [9, 19, 20], the sparsity constraint on the signal has been shown to be useful in reconstructing the sparse signal. In this section, we propose a regularization scheme which can effectively capture the sparsity property of the phase in the setting of the SPR problem (1.1). The idea is to project the temporal estimated signal onto the set of all the signals satisfying both the amplitude modulation and the sparse phase constraint. As a result, the proposed scheme can be incorporated into the framework of projection methods for solving the SPR problem (1.1) as detailed in the subsequent sections. The analysis developed in this section is new.

To begin, let us define on $\mathbb{C}^{n\times n}$ the three constraint sets in accordance with the SPR problem (1.1) as follows:

$$
\begin{aligned}
\Omega_1 &\equiv \left\{x \in \mathbb{C}^{n\times n} \mid \left|\mathcal{F}\left(x \cdot \mathrm{e}^{\mathrm{j}\Psi}\right)\right|^2 = b\right\},\\
\Omega_2 &\equiv \left\{x \in \mathbb{C}^{n\times n} \mid |x| = \chi\right\},\\
\Omega_3 &\equiv \left\{x \in \Omega_2 \mid \|\arg(x)\|_0 \le s\right\}.
\end{aligned} \tag{2.1}
$$

The set Ω_3 describes the sparsity of the phase, and a natural way to incorporate it into a solution method for (1.1) is to perform an additional projection step onto this set. The next lemma provides an explicit form of P_{Ω_3} when restricted on the points relevant to the solutions of (1.1). Note that this result is different from the widely known formula for projections onto intensity or amplitude constraint sets.

Lemma 2.1 (projector P_{Ω_3}). *For any $x = \chi \cdot \mathrm{e}^{\mathrm{j}\Phi} \in \Omega_2$, let us define the matrix*[5]

$$
\mathcal{O} \equiv \chi^2 \cdot \left|\mathrm{e}^{\mathrm{j}\Phi} - \mathbf{1}_n\right|^2, \tag{2.2}
$$

where the squared amplitude χ^2 is understood elementwise, and $\mathbf{1}_n$ is the all-ones matrix of size $n \times n$. Then it holds that

$$
P_{\Omega_3}(x) = \left\{\chi \cdot \mathrm{e}^{\mathrm{j}\varphi} \mid \exists J \in C_s(\mathcal{O}) \textit{ such that } \varphi \textit{ equals } \Phi \text{ on } J \textit{ and vanishes on } \mathcal{J} \setminus J\right\},
$$

where the sets $\mathcal{J}$ and $C_s(\mathcal{O})$ are given by (1.3) and (1.4), respectively.

Proof. We take an arbitrary $J \in C_s(\mathcal{O})$ and $\varphi \in \mathbb{R}^{n\times n}$ determined by

$$
\varphi_\xi = \begin{cases} \Phi_\xi & \text{if } \xi \in J,\\ 0 & \text{if } \xi \in \mathcal{J} \setminus J \end{cases} \tag{2.3}
$$

and show that $z \equiv \chi \cdot \mathrm{e}^{\mathrm{j}\varphi}$ is a projection of x on Ω_3. That is, we need to verify that

$$
\|x - z\| \le \|x - w\| \quad \forall w \in \Omega_3. \tag{2.4}
$$

[4] Since the phase being sparse relative to the aperture size is the objective of our analysis, all the results of this paper are also valid for the case where the aperture has a smaller size than $n \times n$.

[5] $\mathcal{O}$ depends only on Φ since χ is known.

Thanks to $|x| = |z| = |w| = \chi$ (elementwise) and the definition of Ω_3, condition (2.4) amounts to

$$\sum_{\xi\in\mathcal{J}} \chi_\xi^2 \left|\mathrm{e}^{\mathrm{j}\Phi_\xi} - \mathrm{e}^{\mathrm{j}\varphi_\xi}\right|^2 \le \sum_{\xi\in\mathcal{J}} \chi_\xi^2 \left|\mathrm{e}^{\mathrm{j}\Phi_\xi} - \mathrm{e}^{\mathrm{j}\phi_\xi}\right|^2 \quad \forall\phi \in \mathbb{R}^{n\times n} \text{ with } \|\phi\|_0 \le s. \tag{2.5}$$

From the definition (2.3) of φ, we have that

$$\begin{aligned}\sum_{\xi\in\mathcal{J}} \chi_\xi^2 \left|\mathrm{e}^{\mathrm{j}\Phi_\xi} - \mathrm{e}^{\mathrm{j}\varphi_\xi}\right|^2 &= \sum_{\xi\in J} \chi_\xi^2 \left|\mathrm{e}^{\mathrm{j}\Phi_\xi} - \mathrm{e}^{\mathrm{j}\varphi_\xi}\right|^2 + \sum_{\xi^c\in\mathcal{J}\setminus J} \chi_{\xi^c}^2 \left|\mathrm{e}^{\mathrm{j}\Phi_{\xi^c}} - \mathrm{e}^{\mathrm{j}\varphi_{\xi^c}}\right|^2\\ &= \sum_{\xi\in J} \chi_\xi^2 \left|\mathrm{e}^{\mathrm{j}\Phi_\xi} - \mathrm{e}^{\mathrm{j}\Phi_\xi}\right|^2 + \sum_{\xi^c\in\mathcal{J}\setminus J} \chi_{\xi^c}^2 \left|\mathrm{e}^{\mathrm{j}\Phi_{\xi^c}} - 1\right|^2\\ &= \sum_{\xi^c\in\mathcal{J}\setminus J} \chi_{\xi^c}^2 \left|\mathrm{e}^{\mathrm{j}\Phi_{\xi^c}} - 1\right|^2 = \sum_{\xi^c\in\mathcal{J}\setminus J} \mathcal{O}_{\xi^c}.\end{aligned} \tag{2.6}$$

Since $J \in C_s(\mathcal{O})$ and $\mathcal{O}$ is nonnegative by its definition (2.2), we have that

$$\mathcal{O}_{\xi^c} = |\mathcal{O}_{\xi^c}| \le |\mathcal{O}_\xi| = \mathcal{O}_\xi \quad \forall\xi \in J, \forall\xi^c \in \mathcal{J} \setminus J. \tag{2.7}$$

Note that $\|\phi\|_0 \le s$ since $w = \chi \cdot \mathrm{e}^{\mathrm{j}\phi} \in \Omega_3$ and hence

$$\operatorname{card}\left(\{\xi \mid \phi_\xi = 0\}\right) \ge n^2 - s = \operatorname{card}\left(\mathcal{J} \setminus J\right), \tag{2.8}$$

where card$(\cdot)$ denotes the cardinality of the set in the brackets. The combination of (2.7) and (2.8) yields that

$$\sum_{\xi^c\in\mathcal{J}\setminus J} \mathcal{O}_{\xi^c} \le \sum_{\xi\in\{\xi\,|\,\phi_\xi=0\}} \mathcal{O}_\xi. \tag{2.9}$$

Since $\{\xi \mid \phi_\xi = 0\} \subset \mathcal{J}$, it is clear that

$$\begin{aligned}\sum_{\xi\in\{\xi\,|\,\phi_\xi=0\}} \mathcal{O}_\xi &= \sum_{\xi\in\{\xi\,|\,\phi_\xi=0\}} \chi_\xi^2|\mathrm{e}^{\mathrm{j}\Phi_\xi} - 1|^2 = \sum_{\xi\in\{\xi\,|\,\phi_\xi=0\}} \chi_\xi^2|\mathrm{e}^{\mathrm{j}\Phi_\xi} - \mathrm{e}^{\mathrm{j}\phi_\xi}|^2\\ &\le \sum_{\xi\in\{\xi\,|\,\phi_\xi=0\}} \chi_\xi^2|\mathrm{e}^{\mathrm{j}\Phi_\xi} - \mathrm{e}^{\mathrm{j}\phi_\xi}|^2 + \sum_{\xi\in\{\xi\,|\,\phi_\xi\neq0\}} \chi_\xi^2|\mathrm{e}^{\mathrm{j}\Phi_\xi} - \mathrm{e}^{\mathrm{j}\phi_\xi}|^2\\ &= \sum_{\xi\in\mathcal{J}} \chi_\xi^2 \left|\mathrm{e}^{\mathrm{j}\Phi_\xi} - \mathrm{e}^{\mathrm{j}\phi_\xi}\right|^2 \quad \forall\phi \in \mathbb{R}^{n\times n} \text{ with } \|\phi\|_0 \le s.\end{aligned} \tag{2.10}$$

Then for any $\phi \in \mathbb{R}^{n\times n}$ with $\|\phi\|_0 \le s$, using (2.6), (2.9), and (2.10) successively we have that

$$\sum_{\xi\in\mathcal{J}} \chi_\xi^2 \left|\mathrm{e}^{\mathrm{j}\Phi_\xi} - \mathrm{e}^{\mathrm{j}\varphi_\xi}\right|^2 = \sum_{\xi^c\in\mathcal{J}\setminus J} \mathcal{O}_{\xi^c} \le \sum_{\xi\in\{\xi\,|\,\phi_\xi=0\}} \mathcal{O}_\xi \le \sum_{\xi\in\mathcal{J}} \chi_\xi^2 \left|\mathrm{e}^{\mathrm{j}\Phi_\xi} - \mathrm{e}^{\mathrm{j}\phi_\xi}\right|^2.$$

Hence, we have proved (2.5); that is, $z = \chi \cdot \mathrm{e}^{\mathrm{j}\varphi}$ is a projection of x on Ω_3.

Conversely, let us take an arbitrary $z = \chi \cdot \mathrm{e}^{\mathrm{j}\varphi} \in P_{\Omega_3}(x)$ and show that there exists a $J \in C_s(\mathcal{O})$ such that φ equals Φ on J and vanishes on $\mathcal{J} \setminus J$; that is, φ satisfies (2.3). Note that if $x \in \Omega_3$, then $P_{\Omega_3}(x) = \{x\}$ and $z = x$, and (2.3) holds trivially. Hence, it suffices to consider $x \notin \Omega_3$, that is, $\|\Phi\|_0 > s$ (recall that $x = \chi \cdot \mathrm{e}^{\mathrm{j}\Phi} \in \Omega_2$). We claim that $\|\varphi\|_0 = s$ (recall that $\|\varphi\|_0 \leq s$ as $z \in \Omega_3$). Indeed, suppose to the contrary that $\|\varphi\|_0 \leq s - 1$. Since $\|\Phi\|_0 > s$, there exists an index ξ^c such that $\varphi_{\xi^c} = 0$ and $\Phi_{\xi^c} \neq 0$. Define $\phi \in \mathbb{C}^{n\times n}$ by $\phi_{\xi^c} = \Phi_{\xi^c}$ and $\phi_\xi = \varphi_\xi$ for all $\xi \in \mathcal{J} \setminus \{\xi^c\}$, and set $w \equiv \chi \cdot \mathrm{e}^{\mathrm{j}\phi}$. Then $w \in \Omega_3$ since $\|\phi\|_0 = \|\varphi\|_0 + 1 \leq s$, and by the construction of ϕ and $\varphi_{\xi^c} = 0$, we have that

$$\|x - z\|^2 - \|x - w\|^2 = \chi_{\xi^c}^2 \left|\mathrm{e}^{\mathrm{j}\Phi_{\xi^c}} - \mathrm{e}^{\mathrm{j}\varphi_{\xi^c}}\right|^2 = \chi_{\xi^c}^2 \left|\mathrm{e}^{\mathrm{j}\Phi_{\xi^c}} - 1\right|^2 > 0,$$

where the last strict inequality is due to $\Phi_{\xi^c} \neq 0$. This implies that

$$\|x - z\| > \|x - w\| \geq \operatorname{dist}(x, \Omega_3),$$

which contradicts $z \in P_{\Omega_3}(x)$; that is, we have shown $\|\varphi\|_0 = s$.

On the one hand, let $K_\varphi \subset \mathcal{J}$ be the set of indices of nonzero entries of φ. By the definition of the projection and $z \in P_{\Omega_3}(x)$, it holds that $\|x - z\| = \operatorname{dist}(x, \Omega_3)$. Then

$$\begin{aligned}
&\sum_{\xi^c \in \mathcal{J}\setminus K_\varphi} \mathcal{O}_{\xi^c} + \sum_{\xi \in K_\varphi} \chi_\xi^2 \left|\mathrm{e}^{\mathrm{j}\Phi_\xi} - \mathrm{e}^{\mathrm{j}\varphi_\xi}\right|^2 \\
&= \sum_{\xi^c \in \mathcal{J}\setminus K_\varphi} \chi_{\xi^c}^2 \left|\mathrm{e}^{\mathrm{j}\Phi_{\xi^c}} - 1\right|^2 + \sum_{\xi \in K_\varphi} \chi_\xi^2 \left|\mathrm{e}^{\mathrm{j}\Phi_\xi} - \mathrm{e}^{\mathrm{j}\varphi_\xi}\right|^2 \\
&= \sum_{\xi \in \mathcal{J}} \chi_\xi^2 \left|\mathrm{e}^{\mathrm{j}\Phi_\xi} - \mathrm{e}^{\mathrm{j}\varphi_\xi}\right|^2 = \|x - z\|^2 = \operatorname{dist}^2(x, \Omega_3).
\end{aligned} \tag{2.11}$$

On the other hand, it follows from the definition (1.4) of $C_s(\mathcal{O})$ that

$$\sum_{\xi \in J} \mathcal{O}_\xi \quad \text{is constant over all } J \in C_s(\mathcal{O}), \tag{2.12}$$

$$\sum_{\xi \in K} \mathcal{O}_\xi < \sum_{\xi \in J} \mathcal{O}_\xi \quad \forall K \subset \mathcal{J}, \operatorname{card}(K) = s, K \notin C_s(\mathcal{O}), \forall J \in C_s(\mathcal{O}). \tag{2.13}$$

In view of (2.12), the distance $\operatorname{dist}(x, \Omega_3)$ which was calculated in the first part of the proof (see (2.5) and (2.6)) is independent of the choice of J in $C_s(\mathcal{O})$, that is,

$$\sum_{\xi^c \in \mathcal{J}\setminus J} \mathcal{O}_{\xi^c} = \operatorname{dist}^2(x, \Omega_3) \quad \forall J \in C_s(\mathcal{O}). \tag{2.14}$$

In view of (2.14), subtracting (2.13) by $\sum_{\xi \in \mathcal{J}} \mathcal{O}_\xi$, we obtain that

$$\sum_{\xi^c \in \mathcal{J}\setminus K} \mathcal{O}_{\xi^c} > \operatorname{dist}^2(x, \Omega_3) \quad \forall K \subset \mathcal{J},\ \operatorname{card}(K) = s,\ K \notin C_s(\mathcal{O}). \tag{2.15}$$

Then, in view of (2.15) and noting that $\mathrm{card}(K_\varphi) = \|\varphi\|_0 = s$, we infer from the equality (2.11) that

$$\sum_{\xi^c \in \mathcal{J}\setminus K_\varphi} \mathcal{O}_{\xi^c} = \mathrm{dist}^2\left(x, \Omega_3\right), \tag{2.16}$$

$$\sum_{\xi \in K_\varphi} \chi_\xi^2 \left| \mathrm{e}^{\mathrm{j}\Phi_\xi} - \mathrm{e}^{\mathrm{j}\varphi_\xi} \right|^2 = 0. \tag{2.17}$$

The combination of (2.15) and (2.16) implies that $K_\varphi \in C_s(\mathcal{O})$, and (2.17) ensures that φ equals Φ on K_φ. Obviously, φ vanishes on $\mathcal{J} \setminus K_\varphi$ by the definition of K_φ.

The proof is complete. ∎

Note that Lemma 2.1 does not provide the formula of $P_{\Omega_3}(x)$ for all $x \in \mathbb{C}^{n\times n}$ but is restricted to those points in Ω_2 which are the only ones relevant for the purpose of capturing the sparse phase constraint in the SPR problem (1.1).

Remark 2.2. Lemma 2.1 is crucial for the analysis of this paper and rather unexpected since the sparsity regularization applied to the phase turns out to be a metric projection on the underlying space $\mathbb{C}^{n\times n}$. In the special case that the amplitude χ is constant over its entries, P_{Ω_3} reduces to the hard thresholding/L0-norm regularization but is applied to the phase. It is worth mentioning that even the latter special case has not been analyzed in the literature.

3. SROP algorithm.

Based on the regularization scheme for the sparse phase constraint proposed in section 2, this section presents an efficient solution algorithm for the sparse phase retrieval problem (1.1).

Algorithm 3.1 (SROP algorithm for SPR).

Input:

χ — *amplitude modulation*
Ψ — *phase diversity*
$b \in \mathbb{R}_+^{n\times n}$ — *measured PSF image*
s — *sparsity parameter*
τ — *tolerance threshold*
N — *maximal number of iterations*
Φ_0 — *initial guess for* Φ.

Iteration process: *given* Φ_k

(i) $x_k = \chi \cdot \mathrm{e}^{\mathrm{j}\Phi_k}$
(ii) $X_k = \mathcal{F}\left(x_k \cdot \mathrm{e}^{\mathrm{j}\Psi}\right)$ — *phase diversity and Fourier transform*
(iii) $Y_k = \sqrt{b} \cdot \frac{X_k}{|X_k|}$ — *amplitude constraint*
(iv) $y_k = \mathrm{e}^{-\mathrm{j}\Psi} \cdot \mathcal{F}^{-1}(Y_k)$ — *inverse Fourier transform and phase diversity removal*
(v) $z_k = \chi \cdot \mathrm{e}^{\mathrm{j}\varphi_k}$ *with* $\varphi_k = \arg(y_k)$ — *amplitude modulation constraint*
(vi) $\mathcal{O}_k = \chi^2 \cdot \left|\mathrm{e}^{\mathrm{j}\varphi_k} - \mathbf{1}_n\right|^2$
(vii) *find* $J_k \in C_s(\mathcal{O}_k)$ *and set* Φ_{k+1} *equal* φ_k *on* J_k *and vanish on* $\mathcal{J} \setminus J_k$.

Stopping criterion: $\|\Phi_k - \Phi_{k+1}\| < \tau$ *or* $k \geq N$.
Output: $\widehat{\Phi} = \Phi_{end}$ — *the estimated phase.*

In step (iii) of Algorithm 3.1 and elsewhere in this paper, we use the arithmetic convention $\frac{0}{0} = 1$. This is simply for ensuring the uniqueness of that step and has no effect on the performance as well as the analysis of the algorithm.

Algorithm 3.1 is the combination of the GS method (steps (i)–(v)) and the regularization scheme for the sparse phase constraint analyzed in section 2 (steps (vi)–(vii)). When applied to SPR (1.1), the additional regularization step brings substantial advantages of the SROP algorithm over the GS method; see section 5.

4. Convergence analysis.

In this section, we will establish a local linear convergence criterion for the SROP algorithm. We first show that the algorithm is equivalent to the cyclic projection for solving the feasibility problem consisting of the three sets defined in (2.1). Then the analysis of the cyclic projection for nonconvex feasibility recently developed in [17] can be applied to the SROP algorithm.

The next lemma shows that the SROP algorithm is nothing else but the cyclic projection method $P_{\Omega_3}P_{\Omega_2}P_{\Omega_1}$.

Lemma 4.1 (SROP as cyclic projection). *The following statements hold true:*

(i) *The combination of steps (ii), (iii), and (iv) of Algorithm 3.1 amounts to $y_k \in P_{\Omega_1}(x_k)$.*

(ii) *Step (v) of Algorithm 3.1 amounts to $z_k \in P_{\Omega_2}(y_k)$.*

(iii) *The combination of steps (vi), (vii), and (i) of Algorithm 3.1 amounts to $x_{k+1} \in P_{\Omega_3}(z_k)$.*

In other words, the SROP algorithm is characterized by the fixed point operator $T : \mathbb{C}^{n\times n} \rightrightarrows \mathbb{C}^{n\times n}$ given by

$$x \mapsto T(x) \equiv P_{\Omega_3}P_{\Omega_2}P_{\Omega_1}(x) \quad \forall x \in \mathbb{C}^{n\times n}. \tag{4.1}$$

Proof. (i) From the definition of Ω_1 in (2.1), we can write

$$\Omega_1 = \left\{x \in \mathbb{C}^{n\times n} \mid \mathcal{F}\left(x \cdot \mathrm{e}^{\mathrm{j}\Psi}\right) \in C\right\}, \tag{4.2}$$

where the set C is given by

$$C \equiv \left\{X \in \mathbb{C}^{n\times n} \mid |X|^2 = b\right\}. \tag{4.3}$$

The following projection on the set C is straightforward from the simple structure of C defined in (4.3) (recall that the arithmetic convention $\frac{0}{0} = 1$ is used in this paper):

$$\sqrt{b} \cdot \frac{X}{|X|} \in P_C(X) \quad \forall X \in \mathbb{C}^{n\times n}. \tag{4.4}$$

Let us define the linear operator $\mathcal{G} : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ by

$$x \mapsto \mathcal{G}(x) \equiv x \cdot \mathrm{e}^{\mathrm{j}\Psi} \quad \forall x \in \mathbb{C}^{n\times n}. \tag{4.5}$$

Then $\mathcal{G}$ is a unitary transform since $\|x\| = \|\mathcal{G}(x)\|$ for all $x \in \mathbb{C}^{n\times n}$. Its inverse is also unitary and given by $\mathcal{G}^{-1}(x) = \mathrm{e}^{-\mathrm{j}\Psi} \cdot x$ for all $x \in \mathbb{C}^{n\times n}$. Now using (4.2), we obtain that

$$P_{\Omega_1}(x) = \mathcal{G}^{-1}\mathcal{F}^{-1}\left(P_C\left(\mathcal{F}\mathcal{G}(x)\right)\right) \quad \forall x \in \mathbb{C}^{n\times n}. \tag{4.6}$$

Combining steps (ii), (iii), and (iv) of Algorithm 3.1 in view of (4.4) and (4.6), we have that

$$y_k = \mathrm{e}^{-\mathrm{j}\Psi} \cdot \mathcal{F}^{-1}\left(\sqrt{b} \cdot \frac{X_k}{|X_k|}\right) = \mathcal{G}^{-1}\mathcal{F}^{-1}\left(\sqrt{b} \cdot \frac{\mathcal{F}\left(x_k \cdot \mathrm{e}^{\mathrm{j}\Psi}\right)}{|\mathcal{F}(x_k \cdot \mathrm{e}^{\mathrm{j}\Psi})|}\right)$$

$$= \mathcal{G}^{-1}\mathcal{F}^{-1}\left(\sqrt{b} \cdot \frac{\mathcal{F}\mathcal{G}(x_k)}{|\mathcal{F}\mathcal{G}(x_k)|}\right) \in \mathcal{G}^{-1}\mathcal{F}^{-1}\left(P_C\left(\mathcal{F}\mathcal{G}(x_k)\right)\right) = P_{\Omega_1}(x_k) \quad \forall k \in \mathbb{N}.$$

The proof of part (i) is complete.

(ii) This part follows by comparing step (v) of Algorithm 3.1 with the definition of Ω_2 in (2.1).

(iii) This part was already proved in Lemma 2.1. ∎

To this end, the SROP algorithm is the cyclic projection algorithm for solving the associated feasibility problem of

$$\text{finding} \quad x \in \Omega_1 \cap \Omega_2 \cap \Omega_3, \tag{4.7}$$

which can be used to address the SPR problem (1.1). Note that in the noise-free setting, the two problems are equivalent since x is a solution to (1.1) if and only if it is a solution to (4.7).

As a result, one can fully apply the convergence theory for the cyclic projection algorithm developed in [17, section 3.1] to the SROP algorithm. It is worth mentioning that the above mentioned theory also encompasses inconsistent feasibility, which in the setting of this paper corresponds to the SROP algorithm applied to SPR (1.1) with noise. To avoid quite formidable technical details inherently arising for inconsistent feasibility, this application paper presents convergence results in the noise-free setting.

We next recall an amount of mathematical material which is sufficient for formulating a local linear convergence criterion in a concise and memorable statement. Since phase retrieval is a typical nonconvex problem and our approach requires no convex relaxations, we can only prove local convergence of the SROP algorithm as iterations generated by the fixed point operator T given by (4.1).

Recall that a set Ω is called *prox-regular* at a point $x \in \Omega$ if the projector P_Ω is single-valued around x; see, for example, [21, 22]. Ω is called a *prox-regular set* if it is prox-regular at every point of it. The analysis regarding prox-regularity in the context of the phase retrieval problem was first given in [14]. The following lemma provides the geometry of the feasibility problem (4.7).

Lemma 4.2 (prox-regularity of Ω_i). *The following statements regarding the sets defined in* (2.1) *hold true:*

(i) *Ω_1 is a prox-regular set.*

(ii) *Ω_2 is a prox-regular set.*

(iii) *If s is the sparsity of the solutions with the sparsest phase to the problem* (4.7)*, then the set Ω_3 is prox-regular at every point in $\Omega_1 \cap \Omega_2 \cap \Omega_3$.*

As a consequence, in the setting of part (iii)*, the sets Ω_i ($i = 1, 2, 3$) are prox-regular at every solution to the problem* (4.7)*.*

Proof. (i) The proof follows the idea of [14, section 3.1]. We first note that the set C defined in (4.3) is prox-regular at every point of it since C is the product of a number of circles which are typical examples of prox-regular sets. It is clear from the definition of Ω_1 that $\Omega_1 = \mathcal{F}^{-1}\mathcal{G}^{-1}(C)$, where $\mathcal{G}$ is given by (4.5). Being a unitary transform, $\mathcal{F}^{-1}\mathcal{G}^{-1}$ preserves the geometrical structure of C; that is, Ω_1 is also prox-regular at every point of it and the proof is complete.

(ii) The argument for part (i) also encompasses part (ii).

(iii) Suppose that s is the sparsity of the solutions with the sparsest phase to (4.7). By its definition, Ω_3 is the union of $\binom{n^2}{s}$ sets, each of which we call a component of Ω_3, and is the product of s circles and $n^2 - s$ singletons. Here, $\binom{n^2}{s}$ denotes the binomial coefficient indexed by the pair of integers n^2 and s. In view of part (i), each component of Ω_3 is a prox-regular set. Taking an arbitrary point $x \in \Omega_1 \cap \Omega_2 \cap \Omega_3$, we will show that Ω_3 is prox-regular at x. By the definition of Ω_3 and $x \in \Omega_3$, we can write $x = \chi \cdot \mathrm{e}^{\mathrm{j}\Phi}$, where $\Phi \in (-\pi, \pi]^{n\times n}$ satisfying $\|\Phi\|_0 \leq s$. Let $J \subset \mathcal{J}$ be the set of indices of the nonzero entries of Φ. Since x is a solution to (4.7) and s is the sparsity of the solutions with the sparsest phase to this problem, it holds that $\|\Phi\|_0 \geq s$. Hence, we have that $\|\Phi\|_0 = s$. This implies that there is the unique component of Ω_3 containing x, denoted by L. Since the component L is a prox-regular set, it suffices to show the existence of a neighborhood U of x such that

$$\Omega_3 \cap U = L \cap U. \tag{4.8}$$

Let us define

$$\theta \equiv \min\left\{|\Phi_\xi| \mid \xi \in J\right\}, \tag{4.9}$$

$$r \equiv \min\left\{\chi_\xi \mid \xi \in J\right\}, \tag{4.10}$$

$$\delta \equiv r|\mathrm{e}^{\mathrm{j}\theta} - 1|. \tag{4.11}$$

It is clear that $\theta \in (0, \pi]$, $r > 0$, $\delta > 0$, and

$$|\mathrm{e}^{\mathrm{j}\theta} - 1| \leq |\mathrm{e}^{\mathrm{j}\Phi_\xi} - 1| \quad \forall \xi \in J. \tag{4.12}$$

We will show that the neighborhood

$$U \equiv \left\{z \in \mathbb{C}^{n\times n} \mid \|z - x\|_\infty < \delta\right\} \tag{4.13}$$

of x satisfies condition (4.8). Since the inclusion $(\Omega_3 \cap U) \supset (L \cap U)$ is trivial, we only need to prove the converse one, that is, $(\Omega_3 \cap U) \subset (L \cap U)$. Let us take an arbitrary $z \in \Omega_3 \cap U$. Note that $|z| = \chi$ (elementwise) since $z \in \Omega_3 \subset \Omega_2$. Also by $z \in \Omega_3$, we can write $z = \chi \cdot \mathrm{e}^{\mathrm{j}\varphi}$, where $\varphi \in (-\pi, \pi]^{n\times n}$ satisfying $\|\varphi\|_0 \leq s$. Then for every index $\xi \in J$, using the triangle inequality, (4.13), (4.10), (4.12), and (4.11) successively, we have that

$$\begin{aligned}
|z_\xi - \chi_\xi| &\geq |x_\xi - \chi_\xi| - |x_\xi - z_\xi| \\
&= \chi_\xi|\mathrm{e}^{\mathrm{j}\Phi_\xi} - 1| - |x_\xi - z_\xi| \\
&> \chi_\xi|\mathrm{e}^{\mathrm{j}\Phi_\xi} - 1| - \delta \\
&\geq r|\mathrm{e}^{\mathrm{j}\theta} - 1| - \delta = \delta - \delta = 0.
\end{aligned}$$

This in particular implies that $\varphi_\xi \neq 0$ for all $\xi \in J$ and thus $\|\varphi\|_0 = s$. Hence, J is also the set of indices of nonzero entries of φ. That is, $z \in L$ and the proof is complete. ■

We are now ready to formulate a local linear convergence criterion for the SROP algorithm.

Theorem 4.3 (linear convergence of SROP). *Let Φ be a solution with the sparsest phase to the SPR problem (1.1), and let $\|\Phi\|_0 = s$. Suppose that the set-valued mapping $\Theta \equiv \mathrm{Id} - T$, where* Id *is the identity mapping and T is defined in (4.1), is* metrically subregular *at $x \equiv \chi \cdot \mathrm{e}^{\mathrm{j}\Phi}$ for $0 \in \mathbb{C}^{n\times n}$; that is, there are constants $\kappa > 0$ and $\delta > 0$ such that*

$$\kappa \,\mathrm{dist}(z, \Theta^{-1}(0)) \leq \mathrm{dist}(0, \Theta(z)) \quad \forall z \in \mathbb{B}_\delta(x).$$

Then every iterative sequence generated by T converges linearly to a point in $\mathrm{Fix}\, T$, *provided that the initial point is sufficiently close to x.*

Proof. Since $\|\Phi\|_0 = s$ and s is the sparsity of the sparsest phase solutions to the SPR problem (1.1), Φ is a sparsest phase solution to (1.1) if and only if $x = \chi \cdot \mathrm{e}^{\mathrm{j}\Phi}$ is a solution to (4.7). By Lemma 4.2, the sets Ω_i ($i = 1, 2, 3$) defined in (2.1) are prox-regular at x. By [8, Theorem 2.14(ii)], the projectors P_{Ω_i} ($i = 1, 2, 3$) are *almost firmly nonexpansive* [17, Definition 2.2] with a violation that can be made arbitrarily small by shrinking the effective neighborhood of x if necessary. Then by [17, Proposition 2.4(iii)], there is a neighborhood U of x such that the mapping T is almost averaged on U with averaging constant $3/4$ and violation ε, which can be made arbitrarily small. Hence, taking also the metric subregularity of $\Theta = \mathrm{Id} - T$ at x for 0 into account, we can apply [17, Corollary 2.3] to obtain the local linear convergence of T as claimed. ■

The above proof is mostly built on the results proved in [17], which require much more background, for example, the foundations of pointwise almost averaged operators. In this application paper, we choose to use the theory established in our recent paper [17], not to rederive it, since a self-contained proof would be neither instructive nor useful for the emphasis on the novelty of this work. In the remainder of this section, we discuss several frequently asked questions around the convergence result formulated in Theorem 4.3.

Remark 4.4 (convergence to a fixed point). The limit of an iterative sequence generated by a general mapping T is not necessarily a point of $\mathrm{Fix}\, T$, which is the only case of interest when analyzing T as an iterative algorithm. For example, any sequence generated by $T : \mathbb{R} \to \mathbb{R}$ given by

$$T(t) \equiv \begin{cases} t/2 & \text{if } t \neq 0, \\ 1 & \text{if } t = 0 \end{cases}$$

converges linearly with rate $1/2$ to 0, but $0 \notin \mathrm{Fix}\, T = \emptyset$. In Theorem 4.3, the limit is guaranteed to be a point of $\mathrm{Fix}\, T$ thanks to the almost averagedness of T on the effective neighborhood of x. The latter property is derived from the physical structure of the SPR problem (1.1); see Lemma 4.2.

Remark 4.5 (necessity of metric subregularity). There are two types of regularity properties often required to establish a convergence criterion in the nonconvex optimization literature. The physical structure of the SPR problem automatically satisfies one type of regularity (Lemma 4.2). It is not a secret that the second type of regularity, named *metric subregularity* in Theorem 4.3, is difficult to verify, but as has been recently shown in [16], this condition is not only *sufficient* but also *necessary* for local linear convergence. In other words, there is no way around it. Several concrete examples in [17] show that this property is satisfied almost always for phase retrieval, but a systematic study is way beyond the scope of the present article. In the setting of set feasibility, this property is closely related to the mutual arrangement of the sets at the reference point [8, 10, 11, 12, 13]. A deeper look at this type of regularity for the collection of sets $\{\Omega_1, \Omega_2, \Omega_3\}$ defined in (2.1) is obviously of importance but again beyond the scope of this paper.

Remark 4.6 (global convergence of cyclic projection for nonconvex feasibility). With very few exceptions, there is no global convergence guarantee for nonconvex problems. A notable exception is sparse affine feasibility: under an *asymmetric restricted isometry* assumption, cyclic projection was shown to be globally linearly convergent to a global solution [9]. However, even in that situation, verifying the restricted isometry condition is as expensive as solving the original problem.

5. Numerical simulations. This section demonstrates that the new regularization scheme for the sparse phase constraint is efficient for the SPR problem (1.1). Since existing solution algorithms in the literature of sparse signal recovery, for example, [19, 20, 23], are not applicable to SPR as explained in section 1, only the comparison between the SROP algorithm and the corresponding algorithm, but without the regularization step, is presented in this section. The latter one is the GS algorithm [7].

The common parameters for all experiments in this section are below:

- image size: 128×128 pixels;
- circular aperture: diameter 64 pixels;
- amplitude modulation χ: constant across the aperture with unity value;
- phase diversity: $\Psi = 4Z_2^0$, where $Z_2^0(\rho) = 2\rho^2 - 1$ is the Zernike polynomial of order 2 and azimuthal frequency zero;
- PSF image: $b = \left|\mathcal{F}\left(\chi \cdot \mathrm{e}^{\mathrm{j}(\Phi+\Psi)}\right)\right|^2$, where Φ is the simulation phase;
- initial phase guess: zero phase everywhere;
- noise: Poisson noise introduced to the PSF b by using the MATLAB function *imnoise*; and
- number of iterations: 1200.

The other input data and parameters, such as the simulation phase Φ, its sparsity level $\|\Phi\|_0$, and the sparsity parameter s, will be specified for each of the experiments.

Numerical performance is measured by two output quantities: (1) the *change* of the distance between two consecutive iterations $\|\Phi_k - \Phi_{k+1}\|$, which is of interest for understanding convergence properties of the algorithms, and (2) the *feasibility gap* $\|\Phi_k - \Phi\|$, which indicates the quality of phase retrieval in iteration.

5.1. Effectiveness of sparsity regularization on phase. This section numerically demonstrates the effectiveness of the proposed regularization scheme for capturing the sparse phase constraint of the SPR problem (1.1). This is done by comparing SROP with the GS algorithm; the former one is the combination of the latter one and the new regularization scheme.

Experiment setup:

- simulation phase: Φ randomly generated with values in $(-\pi, \pi]$,
- sparsity level: $\|\Phi\|_0 = 319$ (the pixel totality of the aperture is 3168), and

- sparsity parameter: $s = 335$ (about 105% of the sparsity level).

The performance of the SROP algorithm for solving the SPR problem (1.1) compared to the GS algorithm in the two settings without and with noise is summarized in Figures 1 and 2, respectively. Thanks to the additional regularization step, the SROP algorithm (the blue curves) exhibits faster convergence (the left figures) and more accurate restoration (the right figures) compared to the GS method (the red curves). In the noise-free setting, the SROP algorithm restores the exact solution, as indicated by the zero feasibility gap in Figure 1 (right), while the GS algorithm does not. We mention that the features observed from Figures 1 and 2 are consistent for experiments with different realizations of Φ. These numerical results show substantial advantages of the SROP algorithm over the GS method thanks to the additional regularization step.

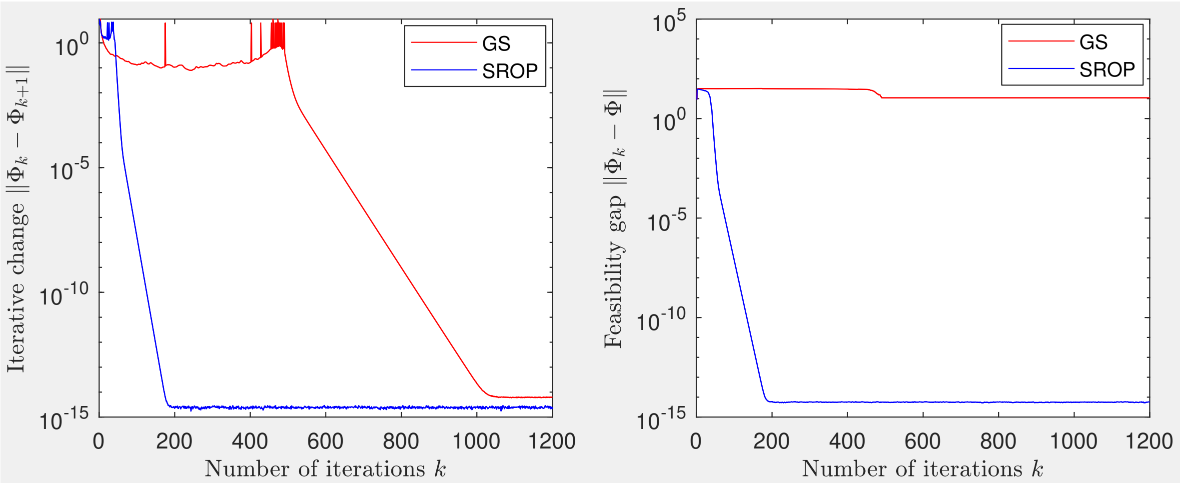

Figure 1. *Experiment shows the effectiveness of the proposed regularization scheme for capturing the sparse phase constraint of SPR in the noise-free setting. Thanks to the additional regularization step, the SROP algorithm (the blue curves) clearly exhibits faster convergence (the left figure) and more accurate restoration (the right figure) compared to the GS method (the red curves).*

5.2. Solvability versus sparsity level. This section numerically addresses the following question: to which sparsity level of the phase Φ is the SROP algorithm effective for the SPR problem?

Experiment setup:

- simulation phase: 120 realizations of Φ with 120 different sparsity levels randomly generated with values in $(-\pi, \pi]$,
- sparsity level: 120 different values of $\|\Phi\|_0 = 4k$ for $k = 1, 2, \ldots, 120$, and
- sparsity parameter: $s \simeq 1.05\|\Phi\|_0$.

The performance of SROP compared to the GS algorithm for 120 experiments in the presence of noise is summarized in Figure 3. The right figure shows that the SROP algorithm (the blue curve) is highly accurate for SPR with sparsity level up to 10% (330 over the totality of 3168 pixels). Within that range of sparsity level, the SROP algorithm is convergent, as shown in the left figure. The SROP algorithm clearly outperforms the GS method (the red

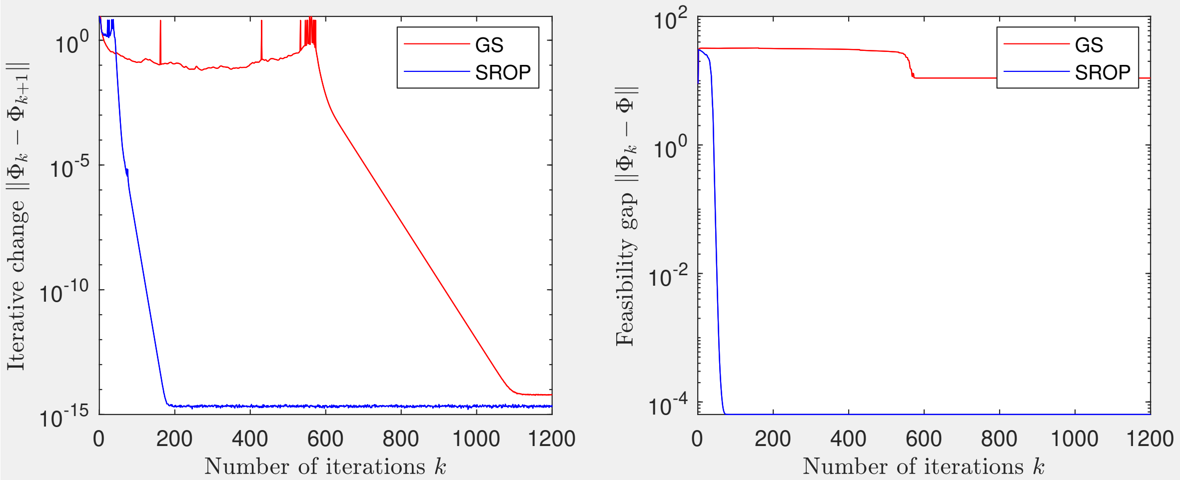

Figure 2. *Experiment shows the effectiveness of the proposed regularization scheme for capturing the sparse phase constraint of SPR in the presence of noise. The overall features are similar to the ones in the noise-free setting shown in Figure 1, except that only approximate solutions can be restored as expected (the right figure).*

curves) in both convergence speed and accuracy thanks to the additional regularization step already observed in section 5.1. As mentioned in section 1, the analysis of this paper can be applied to characterizing phase-only objects in a number of applications in optical science where the phase object occupies less than 10% of the whole optical field. The results of this section show that the SROP algorithm is efficient for potential applications with phase objects of that level of sparsity.

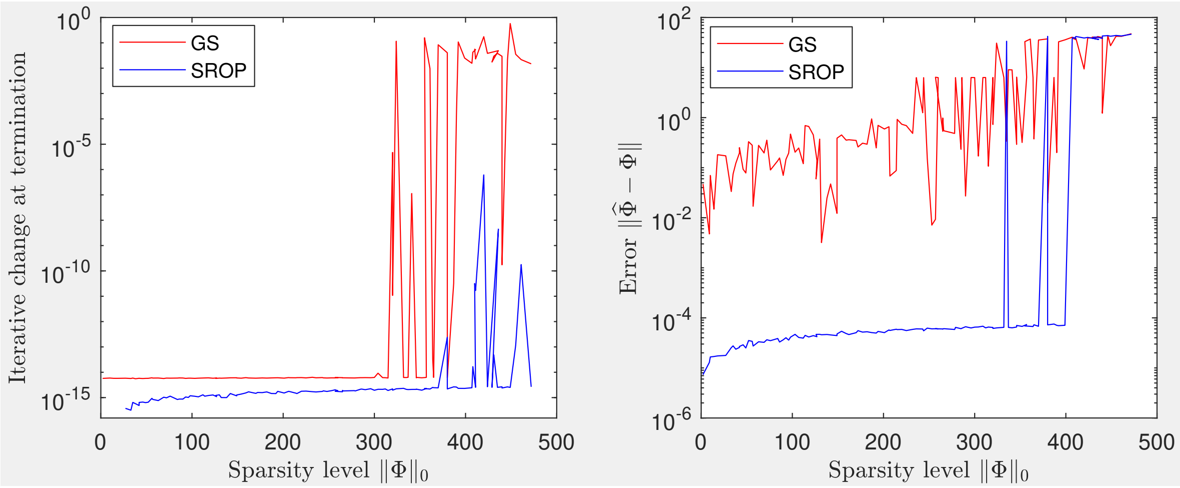

Figure 3. *Experiment shows the solvability of the SROP algorithm for SPR with different sparsity levels in the presence of noise. For sparsity level up to 330 pixels (> 10%), the algorithm (the blue curves) exhibits convergence (the left figure) and accurate restoration (the right figure). It clearly outperforms the GS method (the red curves) in both convergence speed and accuracy.*

5.3. Analysis of sparsity parameter. In application, the sparsity level of the phase object is unknown. In this section, we numerically analyze the sensitivity of the SROP algorithm with respect to the tuning of the sparsity parameter.

Experiment setup:

- simulation phase: Φ randomly generated with values in $(-\pi, \pi]$,
- sparsity level: $\|\Phi\|_0 = 319$ (the pixel totality is 3168), and
- sparsity parameter: 120 different values $s = 168 + 25k$ for $k = 1, 2, \ldots, 120$.

The performance of SROP for 120 experiments in the presence of noise is summarized in Figure 4. For a sparsity parameter smaller than the exact sparsity level of Φ (that is, $s < \|\Phi\|_0$), SROP is not able to find an accurate solution, as indicated by the high values at the beginning of the curve in the right figure. Note that in this case the feasibility problem (4.7) is inconsistent even in the noise-free setting. For a sparsity parameter from 100% up to about 500% of the exact sparsity level, the right figure shows highly accurate restoration by the algorithm. Within that range of sparsity parameter, SROP is convergent, as shown in the left figure. These results show that the SROP algorithm is not sensitive to the tuning of the sparsity parameter s.

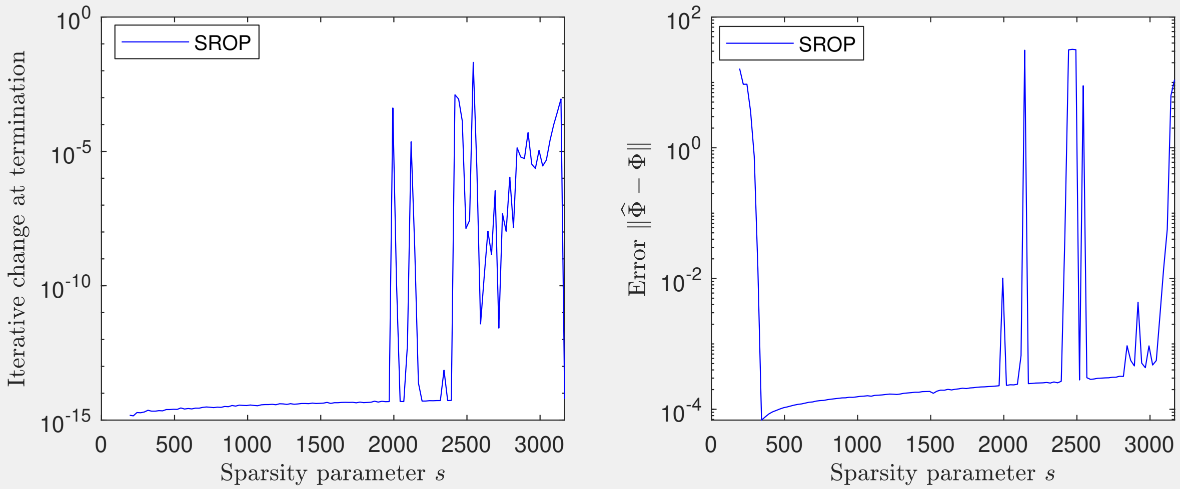

Figure 4. *Experiment shows the stability of the SROP algorithm with respect to the sparsity parameter in the presence of noise. For a sparsity parameter from* 100% *up to* 500% *of the exact sparsity level, the algorithm exhibits convergence (the left figure) and highly accurate restoration (the right figure).*

6. Conclusion. Phase retrieval with sparse phase constraint (SPR) was investigated in this paper for the first time. The problem encompasses a number of important applications in optical science where the phase object occupies a small portion of the whole optical field. A new regularization scheme was proposed for capturing the sparse phase constraint of the SPR problem. The combination of this regularization scheme and the classical GS method forms the so-called SROP algorithm for solving the SPR problem in the setting that only one PSF image is given. Numerical results clearly demonstrated the effectiveness and efficiency of the SROP algorithm for this class of phase retrieval problems. From the viewpoint of projection methods, the SROP algorithm is equivalent to the cyclic projection for solving an associated three-set feasibility problem. Analyzing regularity properties of the involved sets and using the recently developed convergence analysis scheme based on the theory of almost averaged operators, we formulated a linear convergence criterion for the SROP algorithm.

Acknowledgment. The authors thank the two anonymous reviewers for their careful reading and constructive comments that have helped to improve their manuscript.

REFERENCES

[1] 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. Amer. A, 19 (2002), pp. 1334–1345, https://doi.org/10.1364/JOSAA.19.001334.

[2] H. H. Bauschke, D. R. Luke, H. M. Phan, and X. Wang, *Restricted normal cones and sparsity optimization with affine constraints*, Found. Comput. Math., 14 (2014), pp. 63–83, https://doi.org/10.1007/s10208-013-9161-0.

[3] E. Brunner, C. C. de Visser, and M. Verhaegen, *Nonlinear spline wavefront reconstruction from Shack-Hartmann intensity measurements through small aberration approximations*, J. Opt. Soc. Amer. A, 34 (2017), pp. 1535–1549, https://doi.org/10.1364/JOSAA.34.001535.

[4] A. L. Dontchev and R. T. Rockafellar, *Implicit Functions and Solution Mappings*, 2nd ed., Springer-Verlag, New York, 2014, https://doi.org/10.1007/978-1-4939-1037-3.

[5] J. R. Fienup, *Phase retrieval algorithms: A comparison*, Appl. Opt., 21 (1982), pp. 2758–2769, https://doi.org/10.1364/AO.21.002758.

[6] J. R. Fienup, *Phase retrieval algorithms: A personal tour*, Appl. Opt., 52 (2013), pp. 45–56, https://doi.org/10.1364/AO.52.000045.

[7] R. W. Gerchberg and W. O. Saxton, *A practical algorithm for the determination of phase from image and diffraction plane pictures*, Optik, 35 (1972), pp. 237–246.

[8] R. Hesse and D. R. Luke, *Nonconvex notions of regularity and convergence of fundamental algorithms for feasibility problems*, SIAM J. Optim., 23 (2013), pp. 2397–2419, https://doi.org/10.1137/120902653.

[9] R. Hesse, D. R. Luke, and P. Neumann, *Alternating projections and Douglas–Rachford for sparse affine feasibility*, IEEE Trans. Signal. Process., 62 (2014), pp. 4868–4881, https://doi.org/10.1109/TSP.2014.2339801.

[10] A. Y. Kruger, *About regularity of collections of sets*, Set-Valued Anal., 14 (2006), pp. 187–206, https://doi.org/10.1007/s11228-006-0014-8.

[11] A. Y. Kruger, D. R. Luke, and N. H. Thao, *About subtransversality of collections of sets*, Set-Valued Var. Anal., 25 (2017), pp. 701–729, https://doi.org/10.1007/s11228-017-0436-5.

[12] A. Y. Kruger, D. R. Luke, and N. H. Thao, *Set regularities and feasibility problems*, Math. Program., 168 (2018), pp. 279–311, https://doi.org/10.1007/s10107-016-1039-x.

[13] A. Y. Kruger and N. H. Thao, *Quantitative characterizations of regularity properties of collections of sets*, J. Optim. Theory Appl., 164 (2015), pp. 41–67, https://doi.org/10.1007/s10957-014-0556-0.

[14] D. R. Luke, *Finding best approximation pairs relative to a convex and prox-regular set in a Hilbert space*, SIAM J. Optim., 19 (2008), pp. 714–739, https://doi.org/10.1137/070681399.

[15] D. R. Luke, J. V. Burke, and R. G. Lyon, *Optical wavefront reconstruction: Theory and numerical methods*, SIAM Rev., 44 (2002), pp. 169–224, https://doi.org/10.1137/S003614450139075.

[16] D. R. Luke, M. Teboulle, and N. H. Thao, *Necessary conditions for linear convergence of iterated expansive, set-valued mappings*, Math. Program., 180 (2020), pp. 1–31, https://doi.org/10.1007/s10107-018-1343-8.

[17] D. R. Luke, N. H. Thao, and M. K. Tam, *Quantitative convergence analysis of iterated expansive, set-valued mappings*, Math. Oper. Res., 43 (2018), pp. 1143–1176, https://doi.org/10.1287/moor.2017.0898.

[18] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation*. I. *Basic Theory*, Grundlehren Math. Wiss. 330, Springer-Verlag, Berlin, 2006, https://doi.org/10.1007/3-540-31247-1.

[19] S. Mukherjee and C. Seelamantula, *An iterative algorithm for phase retrieval with sparsity constraints: Application to frequency domain optical coherence tomography*, in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2012, pp. 553–556, https://doi.org/10.1109/ICASSP.2012.6287939.

[20] E. J. R. Pauwels, A. Beck, Y. C. Eldar, and S. Sabach, *On Fienup methods for sparse phase retrieval*, IEEE Trans. Signal Process., 66 (2018), pp. 982–991, https://doi.org/10.1109/TSP.2017.2780044.

[21] R. A. Poliquin, R. T. Rockafellar, and L. Thibault, *Local differentiability of distance functions*, Trans. Amer. Math. Soc., 352 (2000), pp. 5231–5249, https://doi.org/10.1090/S0002-9947-00-02550-2.

[22] R. T. Rockafellar and R. J. Wets, *Variational Analysis*, Grundlehren Math. Wiss. 317, Springer-Verlag, Berlin, 1998, https://doi.org/10.1007/978-3-642-02431-3.

[23] Y. Shechtman, A. Beck, and Y. C. Eldar, *GESPAR: Efficient phase retrieval of sparse signals*, IEEE Trans. Signal Process., 62 (2014), pp. 928–938, https://doi.org/10.1109/TSP.2013.2297687.

[24] 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 (2015), pp. 87–109, https://doi.org/10.1109/MSP.2014.2352673.

[25] N. H. Thao, *A convergent relaxation of the Douglas–Rachford algorithm*, Comput. Optim. Appl., 70 (2018), pp. 841–863, https://doi.org/10.1007/s10589-018-9989-y.