



Phase Retrieval by the Inverse Power Method

Qi Luo^{1,2(✉)}, Hongxia Wang^{1,2}, and Jianyun Chen^{1,2}

¹ College of Science, National University of Defense Technology,
Changsha 410073, Hunan, People's Republic of China

² Beijing Institute of Graphics, Beijing 10029, People's Republic of China
luoqi_nudt@outlook.com, whx8292@hotmail.com, c jy2918@163.com
<http://www.ccb2018.xju.edu.cn/>

Abstract. Phase retrieval is to recover signals from phaseless linear measurements. The most efficient methods to tackle this problem are nonconvex gradient approaches, which however generally need an elaborate initialized guess to ensure successful reconstruction. The inverse power method is proposed to provide a more accurate initialization. Numerical experiments illustrate the higher accuracy of the proposed method over other initialization methods. And we further demonstrate the iterative use of the initialization method can obtain an even better estimate.

Keywords: Phase retrieval · Spectral method · Inverse power method

1 Introduction

Phase retrieval is to recover the signal $\mathbf{x} \in \mathbb{C}^n$ from the phaseless measurement:

$$b_i = |\langle \mathbf{a}_i, \mathbf{x} \rangle|, \quad 1 \leq i \leq m, \quad (1)$$

where the measuring vectors \mathbf{a}_i in \mathbb{C}^n . The measuring matrix $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_m]$ is assumed to be full-rank. The applications of phase retrieval exists widely in many fields of sciences and engineering, including X-ray crystallography [1], molecular imaging [2], biological imaging [3] as well as astronomy [4]. To avoid the illness of problem, it is proved that m should be larger than $2n - 1$ in real-valued case or $4n - 4$ in complex-valued case [5, 6], which we call the information limit in phase retrieval.

The nonlinear inverse problem in (1) is generally known to be NP-hard [7]. Therefore, without adopting effective heuristic function, it is almost impossible to find the solution in tolerable time when n is relatively large. The efficient and practical approaches to this problem are generally nonconvex and can be roughly categorized into two types: alternating projection algorithms (also referred as fixed point algorithms) [8, 9] and gradient-descent methods [10, 11].

Work in this paper was supported by NSF grants 61571008.

The crucial step for nonconvex methods is an elaborate initialization. Namely, one usually needs a good estimate of the true solution to ensure the probability of successful reconstruction. In fact, many nonconvex gradient-like method, e.g. the Wirtinger flow method [10] and the Amplitude flow method [12] degenerate to be convex if the initialization lies around the true solution. As more accurate initializer is proposed, the nonconvex method can perform even better when m is below the information limit in phase retrieval [13].

Related work is presented in Sect. 2. In Sect. 3, a more accurate initialization method, called the inverse power method is proposed in this paper. And through numerical experiments, the effectiveness and superior accuracy are illustrated in Sect. 4.

2 Related Work

Current initialization methods include the spectral method [8], the truncated spectral method [14] and the null vector method [15]. For the convenience of the notation, the measuring vectors are normalized to unit vectors before the initialization step, namely $\|\mathbf{a}_i\| = 1$.

The spectral method is a linear algebraic initializer by the maximization problem:

$$\mathbf{x}_{\text{spec}} = \arg \max \left\{ \|\text{diag}(\mathbf{b})\mathbf{A}^*\mathbf{x}\|^2 : \mathbf{x} \in \mathbb{C}^n, \|\mathbf{x}\| = \|\mathbf{x}_0\| \right\}, \quad (2)$$

where $\text{diag}(\mathbf{b})$ is a function that returns a square diagonal matrix with the elements of vector \mathbf{b} on the main diagonal.

The truncated spectral method is a modification of the spectral method. This method first select a part of measuring vectors which have the largest inner product with the true solution \mathbf{x} . Then an estimate can be obtained by finding the vector that is most coherent with the selected set of vectors. Specifically, the truncated spectral vector method is to solve the maximization problem:

$$\mathbf{x}_{\text{t-spec}} = \arg \max \left\{ \|\text{diag}(\mathbf{1}_\tau \odot \mathbf{b})\mathbf{A}^*\mathbf{x}\|^2 : \mathbf{x} \in \mathbb{C}^n, \|\mathbf{x}\| = \|\mathbf{x}_0\| \right\}, \quad (3)$$

where \odot stands for element-wise multiplication, and $\mathbf{1}_\tau$ is the characteristic function of the set

$$\{i : b(i) \leq \tau \|\mathbf{x}_0\|\} \quad (4)$$

with some preset threshold value $\tau \in (0, 1)$.

The null vector method is proposed from another viewpoint: the orthogonality between vectors. A set of *weak* vectors that are most orthogonal to \mathbf{x} are selected from the measuring vectors. And through seeking the vector which is most orthogonal to the *weak* vectors, one can obtain an estimate of the \mathbf{x} . Mathematically speaking, the *weak* vector set is denoted as $I \subset \{1, \dots, m\}$ and

its complement I_c are defined such that $b_i \leq b_j$ for all $i \in I$ and $j \in I_c$. Then the null vector method can be formulated as

$$\mathbf{x}_{\text{null}} = \arg \min \left\{ \|\mathbf{A}_I^* \mathbf{x}\|^2 : \mathbf{x} \in \mathbb{C}^n, \|\mathbf{x}\| = \|\mathbf{x}_0\| \right\}. \quad (5)$$

To make use of the available power method, Chen convert (5) into a maximization problem [15]:

$$\mathbf{x}_{\text{null}} = \arg \max \left\{ \|\mathbf{A}_{I_c}^* \mathbf{x}\|^2 : \mathbf{x} \in \mathbb{C}^n, \|\mathbf{x}\| = \|\mathbf{x}_0\| \right\}. \quad (6)$$

Current initialization methods, (2), (3) and (6) are all about finding the leading eigenvector, which can be efficiently solved by power method. We illustrate the Algorithm 1 for computing (2) as an example.

Algorithm 1. The spectral method

Input: $\mathbf{A}, \mathbf{b}, \|\mathbf{x}_0\|$, threshold value ϵ .

Initialization: \mathbf{x}_1

- 1: **for** $k = 1, 2, \dots$ **do**
- 2: $\mathbf{x}'_{k+1} \leftarrow \mathbf{A} \text{diag}(\mathbf{b}^2) \mathbf{A}^* \mathbf{x}_k$
- 3: $\mathbf{x}_{k+1} \leftarrow \mathbf{x}'_{k+1} / \|\mathbf{x}'_{k+1}\|$
- 4: **if** $\|\mathbf{x}_{k+1} - \mathbf{x}_k\| \leq \epsilon$, **break**
- 5: **end for**

Output: $\mathbf{x}_{\text{spec}} = \mathbf{x}_{k+1} \|\mathbf{x}_0\| / \|\mathbf{x}_{k+1}\|$.

3 Algorithm

3.1 Inverse Power Method

Related work show that the null vector method has the best numerical performance compared with spectral method and the truncated spectral method [15]. However, several drawbacks remains. An important step in the null vector is setting a threshold value for picking out *weak* vectors, and the problem of how to choose the proper threshold value has not been solved yet. Moreover, the selected *weak* vectors are treated without indistinguishably, the accuracy could be improved if proper weights are added upon these vectors. Based on these considerations, our method, the inverse power method is proposed as follows:

$$\mathbf{x}_{\text{null}} = \arg \max \left\{ \left\| \text{diag}(\mathbf{b}^{-\gamma/2}) \mathbf{A}^* \mathbf{x} \right\|^2 : \mathbf{x} \in \mathbb{C}^n, \|\mathbf{x}\| = \|\mathbf{x}_0\| \right\}. \quad (7)$$

where γ is a positive number.

The inverse power method is based on the following facts:

1. By minimizing the function $f_i(\mathbf{x}) = \frac{\mathbf{x}^* \mathbf{a}_i \mathbf{a}_i^* \mathbf{x}}{\|\mathbf{x}\|}$, one can obtain an arbitrary vector in the orthogonal complement space of \mathbf{a}_i .

2. Instead of selecting only a proportion of measuring vectors by comparing b_i , we carefully give various weights to measuring vectors. Specifically, the inverse power method is to minimize a combination of weighted $f_i(\mathbf{x})$:

$$f(\mathbf{x}) = \sum_i^m w_i f_i(\mathbf{x}). \quad (8)$$

The measurements b_i can also be regarded as the measure of orthogonality between \mathbf{a}_i and \mathbf{x}_0 . For arbitrary two measuring vectors \mathbf{a}_i and \mathbf{a}_j , if the b_i is smaller than b_j , then \mathbf{x} is more orthogonal with \mathbf{a}_i than \mathbf{a}_j , hence more weights is added upon the function $f_i(\mathbf{x})$. Therefore, it is reasonable to weight more $f_i(\mathbf{x})$ corresponding to smaller b_i . A natural way to satisfy this goal is weighting $f_i(\mathbf{x})$ with simple functions. Here we take the weights $w_i = b_i^{-\gamma}$ with $\gamma \geq 0$, which makes (8) and (7) equivalent. Through numerical implementations, we find that $\gamma > 1$ would ensure relative good performance.

Solving the minimization problem (7) is identical to searching for the eigenvector with the smallest eigenvalue. This can be solved efficiently through the propose inverse power approach presented in Algorithm 2. The core step of this algorithm is solving a linear system of equation, which can be implemented efficiently by the well-known Gauss-Seidel method, Jacobi method or successive over relaxation method.

Algorithm 2. The inverse power method (IPM)

Input: $\mathbf{A}, \mathbf{b}, \|\mathbf{x}_0\|$, threshod value ϵ .

Initialization: \mathbf{x}_1

- 1: **for** $k = 1, 2, \dots$ **do**
- 2: $\mathbf{x}'_k \leftarrow (\mathbf{A} \text{diag}(\mathbf{b}^{-2}) \mathbf{A}^*)^{-1} \mathbf{x}_k$
- 3: $\mathbf{x}_{k+1} \leftarrow \mathbf{x}'_k / \|\mathbf{x}'_k\|$
- 4: **if** $\|\mathbf{x}_{k+1} - \mathbf{x}_k\| \leq \epsilon$, **break**
- 5: **end for**

Output: $\mathbf{x}_{\text{IPM}} = \mathbf{x}_k \|\mathbf{x}_0\| / \|\mathbf{x}_k\|$.

3.2 Iterative Inverse Power Method

After we obtain an estimate \mathbf{x}_{IPM} by the inverse power algorithm, the similar analysis about orthogonality can be made upon the residual $\mathbf{x}_{\text{res}} = \mathbf{x}_0 - \mathbf{x}_{\text{IPM}}$. This inspires us to propose an iterative inverse power algorithm to improve the estimate.

Denote $\mathbf{b}_{\text{IPM}} = \mathbf{A}^* \mathbf{x}_{\text{IPM}}$, the residual between the true measurements $\mathbf{A}^* \mathbf{x}$ and $\mathbf{A}^* \mathbf{x}_{\text{IPM}}$ can be approximated by $\mathbf{b}_{\text{res}} := \mathbf{b} - |\mathbf{A}^* \mathbf{x}_{\text{IPM}}|$. Then the proportion of \mathbf{x}_{res} in the orthogonal complement space of \mathbf{a}_i can be characterized by $(b_{\text{IPM}})_i^{-\gamma}$. In other words, we can implement the inverse power algorithm upon the \mathbf{b}_{res} to approximate \mathbf{x}_{res} , which can be use as the search direction for better estimate.

The iterative inverse power method is presented in Algorithm 3. The assignment $\mathbf{x}_{\text{IPM}} \leftarrow \text{sign}(\mathbf{x}_{\text{IPM}}^* \mathbf{x}_s) \mathbf{x}_{\text{IPM}}$ adds a phase factor $e^{i\theta}$ upon \mathbf{x}_{IPM} , where $\theta = \arg \min_{\theta} \|e^{i\theta} \mathbf{x}_{\text{IPM}} - \mathbf{x}_s\|$. The step size is restricted to be less than l such that the elements of \mathbf{b}_{res} are always positive, hence ensuring the weights used in the inverse power method are always positive. Otherwise, the negative weights will invalidate the inverse power method. The relaxation parameter α controls the amount of each adjustment.

Algorithm 3. The iterative inverse power method (IIPM)

Input: $\mathbf{A}, \mathbf{b}, \alpha, \|\mathbf{x}_0\|, \epsilon$, maximum number of iterations: K

Initialization: $\mathbf{x}_s = \mathbf{0}$

- 1: **for** $k = 1, 2, \dots, K$ **do**
- 2: $\mathbf{b}_{\text{res}} \leftarrow \mathbf{b} - |\mathbf{A}^* \mathbf{x}_s|$
- 3: $\mathbf{x}_{\text{IPM}} \leftarrow \text{IPM}(\mathbf{A}, \mathbf{b}_{\text{res}}, \|\mathbf{x}_0\|, \epsilon)$
- 4: $\mathbf{x}_{\text{IPM}} \leftarrow \text{sign}(\mathbf{x}_{\text{IPM}}^* \mathbf{x}_s) \mathbf{x}_{\text{IPM}}$
- 5: $l \leftarrow \max \{l \in \mathbb{R}^+ : |\mathbf{a}_i^* (\mathbf{x}_s + l \mathbf{x}_{\text{IPM}})| \leq b_i, \quad i = 1, \dots, m.\}$
- 6: $\mathbf{x}_s \leftarrow \mathbf{x}_s + \alpha l \mathbf{x}_{\text{IPM}}$
- 7: **end for**

Output: $\mathbf{x}_{\text{IIPM}} = \mathbf{x}_s \|\mathbf{x}_0\| / \|\mathbf{x}_s\|$.

4 Numerical Experiments

In this section, we implement several numerical experiments to compare the performance of our algorithms with other initialization algorithms. For comparison, we define the relative error between recovered signal $\hat{\mathbf{x}}$ and the true signal \mathbf{x} as:

$$\text{RE} = \frac{\text{dist}(\hat{\mathbf{x}}, \mathbf{x}_0)}{\|\mathbf{x}\|^2}, \quad (9)$$

where the function $\text{dist}(\hat{\mathbf{x}}, \mathbf{x}_0) = \min_{\theta} \|e^{i\theta} \hat{\mathbf{x}} - \mathbf{x}_0\|$.

Figure 1 compares the inverse power method with three initialization schemes mentioned in Sect. 2. The relative errors of the returned initialization estimate under different oversampling rate m/n are presented. The measuring vectors are i.i.d. $\mathcal{N}(\mathbf{0}, \mathbf{I}_n/2) + i\mathcal{N}(\mathbf{0}, \mathbf{I}_n/2)$. Obviously, all methods perform better as oversampling rate increases. The spectral and truncated spectral method exhibit generally the worst performance compared with other methods. The null vector method enjoys a better performance than the spectral and truncated spectral method, and it achieves similar performance with the inverse method. However, all the inverse power methods performs better than the null vector method as oversampling rate increases. The inverse power methods with $\gamma = 3$ and $\gamma = 3.5$ have the lowest relative error over all oversampling rate. And interestingly, the inverse power method performs almost the same with $\gamma \geq 2$ when oversampling rate is sufficiently large ($m/n > 25$).

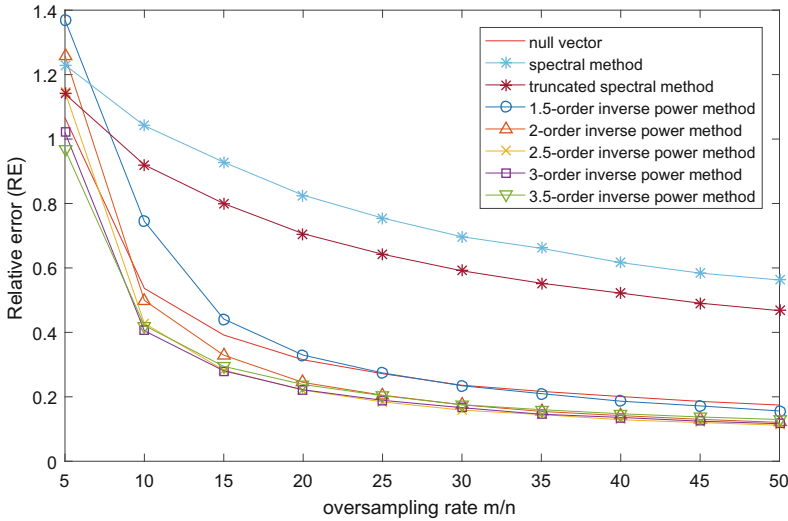


Fig. 1. Relative error of estimates by different initialization schemes under different oversampling rate m/n . $n = 256$, and m/n varies from 5 to 50. The parameter γ of the inverse power method varies from 1.5 to 3.5.

Figure 2 illustrates the effectiveness of the proposed iterative inverse power method. Overall, the iterative inverse power method perform considerably better than the single inverse power method. And better accuracy is achieved as oversampling rate increases. Let α correspond to the step size of each update. Different α can lead to different rates of convergence. Apparently, a larger step size can cause a faster converging rate. It only takes about 10 iterations to converge for $\alpha = 0.9$, while for the $\alpha = 0.1$, dozens of iterations are implemented before converging. However, a larger α can skip the optimum solution, which has been widely discussed in optimization theory [16]. As shown in Fig. 2b, for $\alpha = 0.9$, the relative error is 0.48, much higher than 0.32 for $\alpha = 0.6$, and 0.31 for $\alpha = 0.3$.

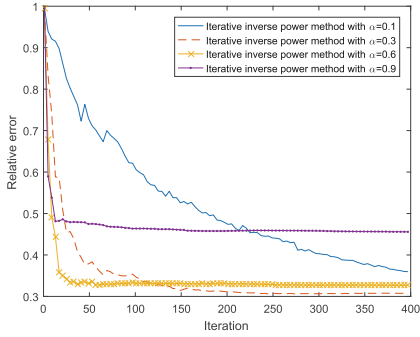
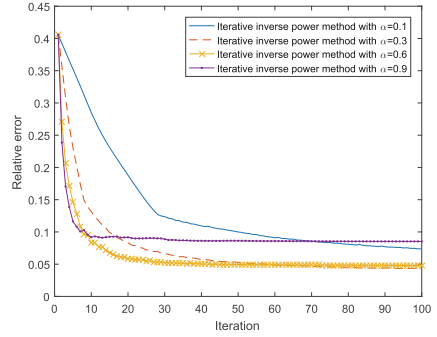
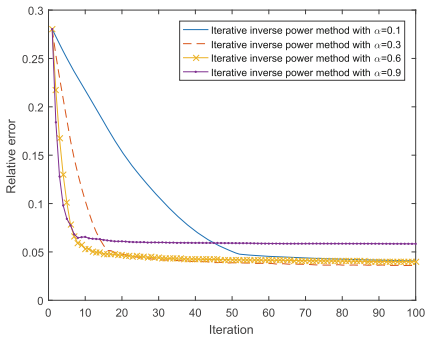
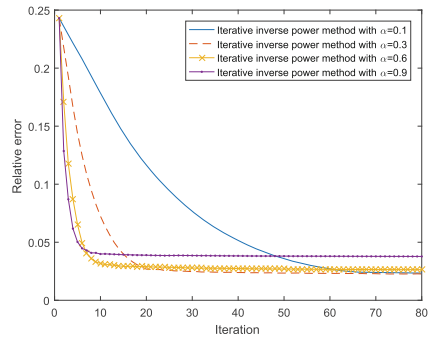
(a) $m/n = 5$ (b) $m/n = 10$ (c) $m/n = 15$ (d) $m/n = 20$

Fig. 2. Relative error for the iterative inverse power method versus iteration count with various step size. α under different oversampling rates.

5 Conclusion

This paper develops a new initialization method, which can efficiently give a more accurate estimate of the original signal for phase retrieval. And through iterative use of the proposed inverse power method, a better performance is achieved. This work is helpful for solving various imaging problems based on phase retrieval.

References

1. Miao, J., Charalambous, P., Kirz, J., Sayre, D.: Extending the methodology of x-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens. *Nature* **400**(6742), 342 (1999)
2. Shechtman, Y., Eldar, Y.C., Cohen, O., Chapman, H.N., Miao, J., Segev, M.: Phase retrieval with application to optical imaging: a contemporary overview. *IEEE Sig. Process. Mag.* **32**(3), 87–109 (2015)

3. Stefik, M.: Inferring dna structures from segmentation data. *Artif. Intell.* **11**(1–2), 85–114 (1978)
4. Fienup, C., Dainty, J.: Phase retrieval and image reconstruction for astronomy. *Image Recovery Theory Appl.* **231**, 275 (1987)
5. Balan, R., Casazza, P., Edidin, D.: On signal reconstruction without phase. *Appl. Comput. Harmonic Anal.* **20**(3), 345–356 (2006)
6. Bodmann, B.G., Hammen, N.: Stable phase retrieval with low-redundancy frames. *Adv. Comput. Math.* **41**(2), 317–331 (2015)
7. Pardalos, P.M., Vavasis, S.A.: Quadratic programming with one negative eigenvalue is NP-hard. *J. Glob. Optim.* **1**(1), 15–22 (1991)
8. Netrapalli, P., Jain, P., Sanghavi, S.: Phase retrieval using alternating minimization. In: *Advances in Neural Information Processing Systems*, pp. 2796–2804 (2013)
9. Elser, V.: Phase retrieval by iterated projections. *JOSA A* **20**(1), 40–55 (2003)
10. Candes, E.J., Li, X., Soltanolkotabi, M.: Phase retrieval via wirtinger flow: theory and algorithms. *IEEE Trans. Inf. Theory* **61**(4), 1985–2007 (2015)
11. Zhang, H., Liang, Y.: Reshaped wirtinger flow for solving quadratic system of equations. In: *Advances in Neural Information Processing Systems*, pp. 2622–2630 (2016)
12. Wang, G., Giannakis, G.B., Eldar, Y.C.: Solving systems of random quadratic equations via truncated amplitude flow. *IEEE Trans. Inf. Theory* **64**, 773–794 (2017)
13. Wang, G., Giannakis, G., Saad, Y., Chen, J.: Solving most systems of random quadratic equations. In: *Advances in Neural Information Processing Systems*, pp. 1865–1875 (2017)
14. Chen, Y., Candes, E.: Solving random quadratic systems of equations is nearly as easy as solving linear systems. In: *Advances in Neural Information Processing Systems*, pp. 739–747 (2015)
15. Chen, P., Fannjiang, A., Liu, G.R.: Phase retrieval by linear algebra. *SIAM J. Matrix Anal. Appl.* **38**(3), 854–868 (2017)
16. Nesterov, Y.: *Introductory Lectures on Convex Optimization: A Basic Course*, vol. 87. Springer, New York (2013). <https://doi.org/10.1007/978-1-4419-8853-9>