Article:

Li, Q., Huang, L., Zhang, P. et al. (2 more authors) (2018) A Fast Gradient-Based Iterative Algorithm for Undersampled Phase Retrieval. IEEE Transactions on Aerospace and Electronic Systems, 54 (4). ISSN 0018-9251

https://doi.org/10.1109/TAES.2018.2832558

A Fast Gradient-Based Iterative Algorithm for Undersampled Phase Retrieval

Qiang Li, Lei Huang*, *Senior Member, IEEE*, Wei Liu, *Senior Member, IEEE*, Weize Sun, Peichang Zhang

***Abstract*—This work develops a fast iterative shrinkage-thresholding algorithm which can efficiently tackle the issue in undersampled phase retrieval. First, using the gradient framework and proximal regularization theory, the undersampled phase retrieval problem is formulated as an optimization in terms of least-absolute-shrinkage-and-selection-operator (LASSO) form with $(\ell_2 + \ell_1)$-norm minimization in the case of sparse signals. A gradient-based phase retrieval via majorization-minimization technique (G-PRIME) is applied to solve a quadratic approximation of the original problem, which, however, suffers a slow convergence rate. Then, an extension of the G-PRIME algorithm is derived to further accelerate the convergence rate, in which an additional iteration is chosen with a marginal increase in computational complexity. Experimental results show that the proposed algorithm outperforms the state-of-the-art approaches in terms of the convergence rate.**

***Index Terms*—Phase retrieval, proximal regularization, majorization-minimization, sparse signal.**

I. Introduction

Phase retrieval seeks to recover a signal or image from the magnitudes of linear measurements, which poses a big challenge in various application areas, such as microscopy [1], waveform optimization [2] and optical imaging [3], to name just a few.

Since phase retrieval is an inherently non-convex ill-posed inverse problem, generally it is difficult to get a closed-form solution. Using the alternating minimization technique [2], [4], the earliest iterative transform method, called the Gerchberg-Saxton algorithm was developed to solve this problem [5]. Another popular method is based on the semidefinite programming (SDP) technique and the rank-1 matrix recovery framework [3]. However, the "matrix-lifting" problem will occur in the case of high dimensional incident signals [6]. More recently, using a steepest descent method with a heuristic step, a Wirtinger Flow algorithm was proposed [7]. Besides, a novel approach, called truncated amplitude flow (TAF) algorithm, was presented in [8] by employing the magnitude-based least squares cost function.

The work described in this paper was supported by the National Natural Science Foundation of China under Grants U1501253, U1713217, 61601304, 61501300, 61601300, 61501485, the Natural Science Foundation of Guangdong Province under Grant 2015A030311030 and by the Foundation of Shenzhen Government under Grants ZDSYS201507081625213, KC2015ZDYF0023A and JCYJ20160520165659418, the Foundation of China Postdoctoral Science under Grant 2015M582414, 2017M610547, and the China Scholarship Council.

Q. Li, L. Huang, W. Sun and P. Zhang are with the College of Information Engineering, Shenzhen University, Shenzhen 518060, China. (Corresponding author: Lei Huang, lhuang8sasp@hotmail.com).

W. Liu is with the Department of Electronic and Electrical Engineering, University of Sheffield, S1 3JD Sheffield, United Kingdom.

Normally, in order to successfully recover an original signal with relatively large probability, the number of measurements M needs to be greater than the dimension N of incident signals. Theoretically, M should be at least on the order of $N\log N$ when the measurement vectors are independent and uniform on a unit sphere [6]. In practice, however, undersampled problem is often encountered, which refers to the case of $M < N$. Existing approaches attempt to tackle the underdetermined problem by introducing the sparsity assumption on incident signals [9]–[12]. In [9], it proposes that a P-sparse complex signal can be recovered successfully with $M \geq 8P - 2$ in the case of Gaussian measurement vectors. Utilizing the feasible point pursuit technique, a phase retrieval approach for DOA estimation was proposed in the presence of gain and phase errors [10], in which the DOA estimation problem was transformed into a phase retrieval formulation with sparse constraint. However, there was no analytical result about the uniqueness of the restored signal for general measurement vectors. More recently, associating compressive phase retrieval via majorization-minimization technique (C-PRIME) with the convex ℓ_1-norm penalty term encouraging sparse solution, a new phase retrieval approach was proposed in [11], where the phase retrieval problem is formulated into the LASSO form. However, its convergence rate is usually slow. Furthermore, [12] developed a sparse TAF algorithm for phase retrieval of sparse signals with a recovery guarantee.

In this paper, we propose two simple and efficient undersampled phase retrieval algorithms, gradient-PRIME and Fast gradient-PRIME (G-PRIME and FG-PRIME for short, respectively) based on the gradient framework and the proximal regularization theory. It is interesting that the proposed G-PRIME algorithm turns out to have a similar closed-form solution with that of the C-PRIME approach, but our G-PRIME algorithm is based on the derivation of the gradient framework. On the basis of the G-PRIME algorithm, we extend the scheme of [13] to the phase retrieval problem to accelerate the convergence rate and the incurred additional computation is marginal.

II. Problem formulation

The problem of estimating an N-dimensional complex signal $\boldsymbol{x}$ from M magnitude-only linear measurements $\boldsymbol{y}$ is called phase retrieval. A basic phase retrieval model with intensity measurements is

$$y_i = |(\boldsymbol{A}\boldsymbol{x})_i|^2 + n_i, \quad i = 1, \cdots, M, \tag{1}$$

where $|\cdot|$ is the element-wise magnitude, y_i and complex measurement matrix $\boldsymbol{A} \in \mathbb{C}^{M\times N}$ are known beforehand and

$\boldsymbol{n} = [n_1, \cdots, n_M]^T$ denotes real-valued white Gaussian noise.

It is easy to observe that the intensity measurements are non-convex and not linear with regard to $\boldsymbol{x}$ due to the magnitude operator. Also, we consider the undersampled phase retrieval problem in this paper, which is an ill-posed inverse problem, and also assume that the incident signal is sparse, which can be found in various areas, such as imaging processing [3].

Because the additive noise to the modulus information $\{\sqrt{y_i}\}_{i=1}^M$ has a smaller variance value than that to the intensity information $\{y_i\}_{i=1}^M$ [11] when $|(\boldsymbol{Ax})_i| > 0.5$, we formulate the undersampled phase retrieval problem as

$$\min_{\boldsymbol{x}} \sum_{i=1}^{M} (y_i - |(\boldsymbol{Ax})_i|)^2 + \lambda \|\boldsymbol{x}\|_1 , \tag{2}$$

where $\|\cdot\|_1$ denotes ℓ_1 norm and the first term is a data fitting error, which should be comparable to the noise level for a successful recovery. Note that $\|\boldsymbol{x}\|_1$ in the second term is used to regularize the ill-posed phase retrieval problem and promote sparsity in $\boldsymbol{x}$. The parameter $\lambda > 0$ is a regularization penalty factor to balance the weights between the sum of measurement error and sparsity level of the estimated solution. Due to the magnitude operator, (2) is not a convex problem either, which can not be directly solved by standard convex optimization approaches.

Employing the majorization-minimization (MM) technique, [11] proposed an efficient C-PRIME method to solve a convex surrogate problem instead. The surrogate optimization problem is convex with regard to $\boldsymbol{x}$ and equivalent to the following issue

$$\boldsymbol{x} = \arg\min_{\boldsymbol{x}} \left[C \|\boldsymbol{x} - \boldsymbol{c}\|_2^2 + \lambda \|\boldsymbol{x}\|_1 \right] , \tag{3}$$

where C is a constant satisfying $C \geq \lambda_{\max}(\boldsymbol{A}^H\boldsymbol{A})$ with $\lambda_{\max}(\cdot)$ denoting the largest eigenvalue of a matrix and the vector $\boldsymbol{c}$ independent of the variable $\boldsymbol{x}$ at the k iteration is

$$\boldsymbol{c} = \boldsymbol{x}^{k-1} - \frac{1}{C}\boldsymbol{A}^H \left(\boldsymbol{A}\boldsymbol{x}^{k-1} - \sqrt{\boldsymbol{y}} \odot e^{j\,\mathrm{ang}(\boldsymbol{A}\boldsymbol{x}^{k-1})} \right) , \tag{4}$$

where $\mathrm{ang}(\cdot)$ denotes the phase angle.

The optimization problem (3) has a simple closed-form solution at the k iteration using the soft thresholding method, i.e.,

$$\boldsymbol{x}^k = e^{j\,\mathrm{ang}(\boldsymbol{c})} \odot \max \left\{ |\boldsymbol{c}| - \frac{\lambda}{2C}, 0 \right\} , \tag{5}$$

where $\odot$ denotes the Hadamard (element-wise) product of two vectors.

The C-PRIME method has an advantage that it only needs to solve a surrogate optimization problem (3) with a simple closed-form solution at every iteration. But the convergence rate of this algorithm is low.

III. Proposed Algorithm based on gradient framework

In this section, on the basis of the C-PRIME algorithm, we first develop a G-PRIME method under the framework of gradient, which will serve as a preparation for the FG-PRIME to deal with the convergence rate problem.

A. G-PRIME Algorithm

The phase retrieval problem (3) can be cast as a second order cone programming problem. We first consider the following general formulation

$$\boldsymbol{x} = \arg\min_{\boldsymbol{x}} \left[F(\boldsymbol{x}) = f(\boldsymbol{x}) + g(\boldsymbol{x}) \right] , \tag{6}$$

where f is a smooth convex function and g is a continuous convex function which is possibly nonsmooth.

Specifically, for the convex optimization problem (3), let $f(\boldsymbol{x}) = C\|\boldsymbol{x} - \boldsymbol{c}\|_2^2$ and $g(\boldsymbol{x}) = \lambda \|\boldsymbol{x}\|_1$. One of the most popular methods for solving the problem is the iterative shrinkage-thresholding algorithm (ISTA) [14]. The iterative procedure of ISTA is

$$\boldsymbol{x}^k = e^{j\,\mathrm{ang}(\boldsymbol{a})} \odot \max \left\{ |\boldsymbol{a}| - \lambda\mu, 0 \right\} , \tag{7}$$

where μ denotes an appropriate step size and the vector $\boldsymbol{a}$ is computed as

$$\boldsymbol{a} = \boldsymbol{x}_{k-1} - 2\mu C (\boldsymbol{x}_{k-1} - \boldsymbol{c}) . \tag{8}$$

Similar to the C-PRIME algorithm, the update of $\boldsymbol{x}_k$ in the ISTA method is employed at the previous value $\boldsymbol{x}_{k-1}$. In the following section, we will consider another given quantity $\boldsymbol{\eta}$ which may be equal or not equal to $\boldsymbol{x}_{k-1}$. According to Taylor series and proximal regularization theorem [14], for a given point $\boldsymbol{\eta}$, a quadratic approximation of $F(\boldsymbol{x}) = f(\boldsymbol{x}) + g(\boldsymbol{x})$ can be written as

$$Q_L(\boldsymbol{x}, \boldsymbol{\eta}) = f(\boldsymbol{\eta}) + \langle \boldsymbol{x} - \boldsymbol{\eta}, \nabla f(\boldsymbol{\eta}) \rangle + \frac{L}{2} \|\boldsymbol{x} - \boldsymbol{\eta}\|^2 + g(\boldsymbol{x}) , \tag{9}$$

where L plays the role of a step size and $\nabla f(\cdot)$ denotes complex gradient vector. Then, we have

$$\boldsymbol{x}^k = \arg\min_{\boldsymbol{x}} \{ Q_L(\boldsymbol{x}, \boldsymbol{\eta}) \} . \tag{10}$$

Discarding the constant term about $\boldsymbol{x}$, $\boldsymbol{x}^k$ is simplified as

$$\begin{aligned} \boldsymbol{x}^k &= \arg\min \left\{ g(\boldsymbol{x}) + \frac{L}{2} \left\| \boldsymbol{x} - \left(\boldsymbol{\eta} - \frac{1}{L} \nabla f(\boldsymbol{\eta}) \right) \right\|^2 \right\} \\ &= \arg\min \left\{ \lambda \|\boldsymbol{x}\|_1 + \frac{L}{2} \left\| \boldsymbol{x} - \left[\boldsymbol{\eta} - \frac{2C}{L} (\boldsymbol{\eta} - \boldsymbol{c}) \right] \right\|^2 \right\} . \end{aligned} \tag{11}$$

Furthermore, according to the soft thresholding method, we have

$$\boldsymbol{x}^k = e^{j\,\mathrm{ang}(\boldsymbol{b})} \odot \max \left\{ |\boldsymbol{b}| - \frac{\lambda}{L}, 0 \right\} , \tag{12}$$

where

$$\boldsymbol{b} = \boldsymbol{\eta} - \frac{2C}{L} (\boldsymbol{\eta} - \boldsymbol{c}) . \tag{13}$$

Then, if $\boldsymbol{\eta} = \boldsymbol{x}^{k-1}$, substituting (5) into (13) and simplifying it, we have

$$\boldsymbol{b} = \boldsymbol{x}^{k-1} - \frac{2}{L}\boldsymbol{A}^H \left(\boldsymbol{A}\boldsymbol{x}^{k-1} - \sqrt{\boldsymbol{y}} \odot e^{j\,\mathrm{ang}(\boldsymbol{A}\boldsymbol{x}^{k-1})} \right) . \tag{14}$$

In this case, the solution $\boldsymbol{x}$ depends on step size L rather than parameter C. Here, we call the algorithm as G-PRIME. It is interesting that we obtain the same solution to the problem

(3) as that of the C-PRIME algorithm but from a totally different gradient theorem. Also the C-PRIME method can be regarded as a special case of G-PRIME in the case of $\boldsymbol{\eta} = \boldsymbol{x}^{k-1}$.

It can be seen from the above analysis that the update of $\boldsymbol{x}^k$ only relies on $\boldsymbol{x}^{k-1}$ in the case of $\boldsymbol{\eta} = \boldsymbol{x}^{k-1}$, which is the same as the C-PRIME. The G-PRIME algorithm is tabulated in Algorithm 1.

Algorithm 1: G-PRIME algorithm

Input: $\boldsymbol{A}, \boldsymbol{y}, \lambda, K$
Step 1. Initial $\boldsymbol{x}^0 \leftarrow$ random complex vector,
Choose $L = 2 * \lambda_{max}(\boldsymbol{A}^H\boldsymbol{A})$.
for $k = 1, \cdots, K$ **do**
Step 2. Determine $\boldsymbol{b}$ by (14).
Step 3. Update $\boldsymbol{x}^k$ by (12).
end for
Output: $\boldsymbol{x}^K$.

B. FG-PRIME Algorithm

In order to further accelerate the convergence rate, we extend the scheme in [13] to the phase retrieval problem (10). Now assume $\boldsymbol{\eta} \neq \boldsymbol{x}^{k-1}$ and let $\boldsymbol{\eta}$ denote a specific linear combination of $\{\boldsymbol{x}^{k-1}, \boldsymbol{x}^{k-2}\}$, which is given by

$$\boldsymbol{\eta} = \boldsymbol{x}^{k-1} + \frac{\gamma^{k-1} - 1}{\gamma^k}\left(\boldsymbol{x}^{k-1} - \boldsymbol{x}^{k-2}\right), \tag{15}$$

where

$$\gamma^k = \frac{1 + \sqrt{1 + 4(\gamma^{k-1})^2}}{2}. \tag{16}$$

The recursive relationship in (16) has been proved in [14]. Compared with the G-PRIME algorithm, the FG-PRIME algorithm requires additional computations in steps (15) and (16), but it is easy to observe that this additional cost is very marginal. The proposed FG-PRIME algorithm is tabulated in Algorithm 2.

Algorithm 2: FG-PRIME algorithm

Input: $\boldsymbol{A}, \boldsymbol{y}, \lambda, K, C$
Step 1. Initial $\boldsymbol{x}^0, \boldsymbol{\eta}^1 = \boldsymbol{x}^0, \gamma^1 = 1$,
Choose $L = 2 * \lambda_{max}(\boldsymbol{A}^H\boldsymbol{A})$.
for $k = 1, \cdots, K$ **do**
Step 2. Determine $\boldsymbol{c}$ by (5) and calculate $\boldsymbol{b}$ by
$\boldsymbol{b} = \boldsymbol{\eta}^k - \frac{2C}{L}\left(\boldsymbol{\eta}^k - \boldsymbol{c}\right)$.
Step 3. Calculate $\boldsymbol{x}^k$ by (12).
Step 4. Update γ^{k+1} by (16).
Step 5. Update $\boldsymbol{\eta}^{k+1}$ by
$\boldsymbol{\eta}^{k+1} = \boldsymbol{x}^k + \frac{\gamma^k - 1}{\gamma^{k+1}}\left(\boldsymbol{x}^k - \boldsymbol{x}^{k-1}\right)$.
end for
Output: $\boldsymbol{x}_K$.

As mentioned in [15], due to the loss of phase information, the recovered signal may have a constant phase shift with respect to the original signal $\boldsymbol{x}_0$. Therefore, an accurate phase shift needs to be calculated in the following procedure. After

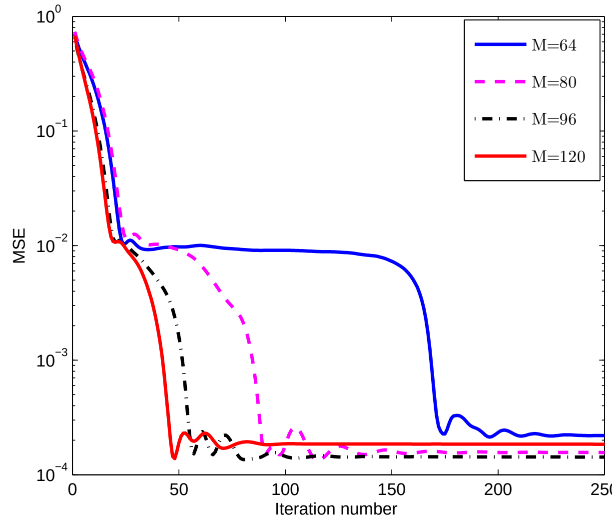


Fig. 1. MSE of the FG-PRIME algorithm versus iteration number, $M = \{64, 80, 96, 120\}$.

getting a solution $\boldsymbol{x}^*$ from the above algorithm, we define a function of mean squared error (MSE) as

$$h(\phi) = \|\boldsymbol{x}_0 - \boldsymbol{x}^* \cdot e^{j\phi}\|_2^2, \tag{17}$$

where ϕ denotes the constant phase shift. The derivative of $h(\phi)$ with respect to ϕ is

$$\nabla h(\phi) = j[(\boldsymbol{x}^*)^H\boldsymbol{x}_0 \cdot e^{-j\phi} - \boldsymbol{x}_0^H\boldsymbol{x}^* \cdot e^{j\phi}]. \tag{18}$$

Setting the derivative to zero, we have

$$e^{j\phi^*} = \frac{(\boldsymbol{x}^*)^H\boldsymbol{x}_0}{|(\boldsymbol{x}^*)^H\boldsymbol{x}_0|}. \tag{19}$$

where $e^{j\phi^*}$ is the estimation of $e^{j\phi}$. Finally, $\boldsymbol{x}^* \cdot e^{j\phi^*}$ gives the recovered signal.

IV. Simulation Results

To compare the performance of the proposed G-PRIME and FG-PRIME algorithms with existing ISTA [14], C-PRIME and C-PRIME-SQUAREM algorithms [11] for various scenarios, we present some experimental results in this section. It should be noted that the original ISTA algorithm in [14] is used to tackle the general linear inverse problem. In this paper, combining the model of the C-PRIME algorithm, the ISTA technique can solve the phase retrieval problem, which is abbreviated as the ISTA-PRIME algorithm.

We assume that the measurement matrix is standard complex Gaussian distributed, corrupted with real-valued additive white Gaussian noise and the original complex signal is generated randomly. The length N of the original complex signal is set as 128 with sparsity level $P = 8$. Moreover, the number of measurements is $M = 120$ and the signal-to-noise ratio (SNR) is 25dB unless specified otherwise. The parameter C and regularization penalty factor λ in all tested methods are set as $C = \lambda_{\max}(\boldsymbol{A}^H\boldsymbol{A})$ and $\lambda = 0.1$, respectively. We assign step size $L = 2\lambda_{\max}(\boldsymbol{A}^H\boldsymbol{A})$ for our proposed G-PRIME and

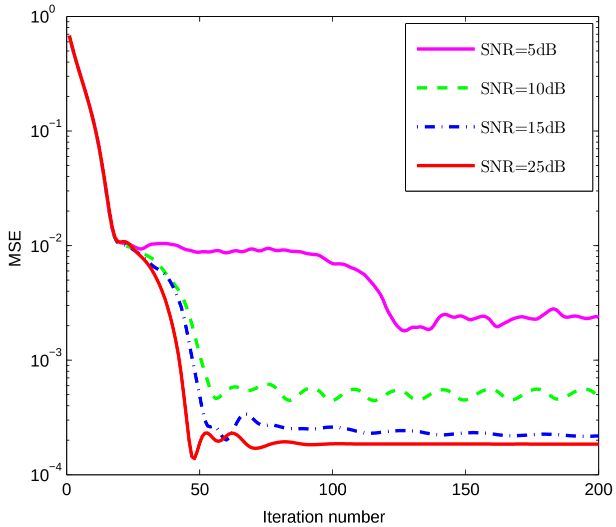


Fig. 2. MSE of the FG-PRIME algorithm versus iteration number, SNR=$\{5, 10, 15, 25\}$ dB.

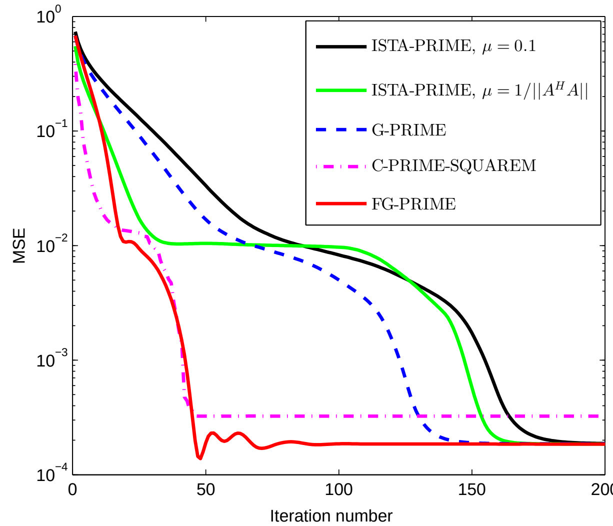


Fig. 3. MSE versus iteration number, M=120, SNR=25 dB.

FG-PRIME algorithms. For the ISTA method, the step size μ should satisfy $\mu \in (0, 1/||\boldsymbol{A}^H \boldsymbol{A}||]$. The other parameters are initialized as in Algorithms 1 and 2.

Firstly, for the proposed FG-PRIME algorithm, we test its MSE performance with different number of measurements. Fig. 1 shows that the MSE performances versus iteration number in the cases of $M = \{64, 80, 96, 120\}$. It is observed that all the MSE curves can converge close to 2×10^{-4}, which agrees with the statement in [9] that the original signal can be recovered successfully when measurements satisfy $M \geq 8P - 2$. Furthermore, Fig. 1 also indicates that, as the number of measurements increases, the convergence of MSE curves becomes faster. Specifically, the MSE curves of M=64 and M=80 approach the steady state when the number of iterations reaches 200 and 130, respectively. Moreover, the MSE curves of M=$\{$96, 120$\}$ have converged before the number of measurements reaches 100. Furthermore, the case for M=120 has the highest convergence rate and its steady-state value is slightly larger than those of the MSE curves for M=$\{$80, 96$\}$.

Then, we consider the MSE performance of the FG-PRIME algorithm under different SNRs. The MSE curves of FG-PRIME versus iteration number for SNR=$\{5, 10, 15, 25\}$ dB are shown in Fig. 2, where all the MSE curves decrease rapidly. We observe that as the SNR increases, the MSE for SNR=$\{10, 15, 25\}$ dB converges fast and has a lower converged value, in which the case of SNR=25 dB is the fastest and it also has the lowest steady-state value close to 2×10^{-4}.

Fig. 3 depicts the MSE performance of the ISTA-PRIME ($\mu = 0.1, 1/||\boldsymbol{A}^H \boldsymbol{A}||$), G-PRIME, C-PRIME-SQUAREM and FG-PRIME algorithms. As mentioned in Section III, the C-PRIME algorithm has the same solution as the G-PRIME algorithm. So the MSE curve of C-PRIME is not shown in Fig. 3. It is obvious that the ISTA-PRIME algorithm has the slowest convergence rate, the C-PRIME and G-PRIME algorithms converge when the iteration number reaches 150. Moreover, the MSE curves of C-PRIME-SQUAREM and FG-PRIME have the fastest convergence rate, because the C-PRIME-SQUAREM and FG-PRIME algorithms have used more *a priori* information. Moreover, compared with the C-PRIME-SQUAREM algorithm, the FG-PRIME algorithm has a lower steady-state value.

V. Conclution

Two undersampled phase retrieval algorithms based on the gradient framework have been derived. For the non-convex objective function of phase retrieval, a quadratic approximation of the original problem was tackled by the proposed G-PRIME technique. Then, in order to further accelerate the convergence rate, the FG-PRIME algorithm is developed, with more *a priori* information exploited. Numerical results have confirmed that the FG-PRIME algorithm is superior to other existing algorithms in terms of convergence rate. In our future work, we will consider adaptive step-size L and test the performance of the proposed algorithm using large-dimensional data.

References

[1] J. Miao, T. Ishikawa, B. Johnson, E. H. Anderson, B. Lai, and K. O. Hodgson, "High resolution 3D x-ray diffraction microscopy," *Phys. Rev. Lett.*, vol. 89, no. 8, pp. 088303, Aug. 2002.

[2] L. K. Patton, and B. D. Rigling, "Phase retrieval for radar waveform optimization," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 4, pp. 3287-3302, Oct. 2012.

[3] Y. Shechtman, Y. Eldar, O. Cohen, H. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: A contemporary overview," *IEEE Signal Process. Mag.*, vol. 32, no. 3, pp. 87-109, May 2015.

[4] P. Netrapalli, P. Jain, and S. Sanghavi, "Phase retrieval using alternating minimization," *IEEE Trans. Signal Process.*, vol. 63, no. 18, pp. 4814-4826, Sep. 2015.

[5] R. W. Gerchberg, and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, no. 2, pp. 237-246, 1972.

[6] E. J. Candès, T. Strohmer, and V. Voroninski, "Phaselift: Exact and stable signal recovery frommagnitude measurements via convex programming," *Commun. Pure Appl. Math.*, vol. 66, no. 8, pp. 1241-1274, Aug. 2013.

[7] E. J. Candès, X. Li, and M. Soltanolkotabi, "Phase retrieval via Wirtinger Flow: Theory and algorithms," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1985-2007, Apr. 2015.

[8] G. Wang, G. B. Giannakis, and Y. Eldar, "Solving random systems of quadratic equations via truncated amplitude flow," *IEEE Trans. Inf. Theory*, vol. 64, no. 2, pp. 773-794, Feb. 2018.

[9] X. Li, and V. Voroninski, "Sparse signal recovery from quadratic measurements via convex programming," *SIAM Journal on Mathematical Analysis*, vol. 45, no. 5, pp. 3019-3033, Sep. 2013.

[10] W. Wang, R. Wu, J. Liang, and H. C. So, "Phase retrieval approach for DOA estimation with array errors." *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 5, pp. 2610-2620, Oct. 2017.

[11] T. Qiu, and D. P. Palomar, "Undersampled sparse phase retrieval via majorization-minimization," *IEEE Trans. Signal Process.*, vol. 65, no. 22, pp. 5957-5969, Nov. 2017.

[12] G. Wang, L. Zhang, G. B. Giannakis, M. Akcakaya, and J. Chen, "Sparse phase retrieval via truncated amplitude flow," *IEEE Trans. Signal Process.*, vol. 66, no. 2, pp. 479-491, Jan. 2018.

[13] Y. E. Nesterov, "A method for solving the convex programming problem with convergence rate $O(1/k^2)$," *Dokl. Akad. Nauk SSSR*, vol. 269, pp. 543-547, 1983.

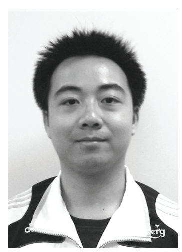
[14] A. Beck, and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sci.*, vol. 2, no. 1, pp. 183-202, Mar. 2009.

[15] T. Qiu, P. Babu, and D. P. Palomar, "PRIME: Phase retrieval via Majorization-Minimization," *IEEE Trans. Signal Process.*, vol. 64, no. 19, pp. 5174-5186, Oct. 2016.

Qiang Li received the B.Eng. degree in electronic engineering from Henan University of Science and Technology, China, in 2010, and the Ph.D. degree in control science and engineering from Harbin Engineering University, China, in 2015. From 2015-2016, he was employed at Shenzhen University, China, where he worked on frequency diverse array and space-time adaptive processing. He joined communication group of the University of Sheffield, UK, in 2016, as a visiting scholar, where he researched the phase retrieval algorithms for antenna array.

Lei Huang (M'07-SM'14) was born in Guangdong, China. He received the B.Sc., M.Sc., and Ph.D. degrees in electronic engineering from Xidian University, Xian, China, in 2000, 2003, and 2005, respectively. From 2005 to 2006, he was a Research Associate with the Department of Electrical and Computer Engineering, Duke University, Durham, NC, USA. From 2009 to 2010, he was a Research Fellow with the Department of Electronic Engineering, City University of Hong Kong, and a Research Associate with the Department of Electronic Engineering, The Chinese University of Hong Kong. From 2011 to 2014, he was a Professor with the Department of Electronic and Information Engineering, Harbin Institute of Technology Shenzhen Graduate School. Since 2014, he has been with the College of Information Engineering, Shenzhen University, where he is currently a Distinguished Professor and the Director of the Shenzhen Key Lab of Advanced Navigation Technology. His research interests include array signal processing, statistical signal processing, sparse signal processing, and their applications in radar, navigation, and wireless communications. He has been on the editorial boards of the IEEE Transactions on Signal Processing (2015-present), Elsevier-Digital Signal Processing (2012-present) and IET Signal Processing (2017-present). Moreover, he has been an elected member of the Sensor Array and Multichannel (SAM) Technical Committee of the IEEE Signal Processing Society (2016-present). Besides, he was elected an IET Fellow in 2018.

Wei Liu received his B.Sc. in Space Physics in 1996, L.L.B. in Intellectual Property Law in 1997, both from Peking University, China, M.Phil. from the Department of Electrical and Electronic Engineering, University of Hong Kong, in 2001, and Ph.D. in 2003 from the School of Electronics and Computer Science, University of Southampton, UK. He then worked as a postdoc in the same group and later in the Communications and Signal Processing Group, Department of Electrical and Electronic Engineering, Imperial College London. In September 2005, he joined the University of Sheffield, UK, as a Lecturer and now he is a Senior Lecturer. His research interests are mainly in sensor array signal processing, blind signal processing, multirate signal processing and their various applications such as wireless communications, sonar, radar, satellite navigation, and speech enhancement, etc. He has published one research monograph ("Wideband Beamforming: Concepts and Techniques", John Wiley & Sons, 2010), three book chapters, 105 journal papers and 135 conference papers. He is a senior member of IEEE, an Associate Editor for two journals "IEEE Transactions on Signal Processing", and "IEEE Access", an editorial board member for the journal "Frontiers of Information Technology & Electronic Engineering", an elected member of the Digital Signal Processing Technical Committee (DSPTC) of the IEEE Circuits and Systems Society and the Sensor Array and Multichannel Technical Committee (SAMTC) of the IEEE Signal Processing Society, and a fellow of the Higher Education Academy of the UK. He is a General Co-Chair for the 2018 IEEE Sensor Array and Multi-Channel Signal Processing Workshop, which will be held at the University of Sheffield in July 2018.



Weize Sun (S11-M15) was born in Guangdong, China. He received the B.Sc. degree in electronic information science and technology from SUN YATSEN University, Guangdong, China, in 2009 and the Ph.D. degree from The City University of Hong Kong in electronic engineering in 2013. Since 2015, he has been with the Department of Information Engineering, Shenzhen University, where he is currently an Assistant Professor. His research interests include multidimensional signal processing, matrix and tensor reconstruction and completion, machine learning and their applications.

Peichang Zhang received the B.Eng. degree (with first-class honor) in electronic engineering from the University of Central Lancashire, Preston, U.K., in 2009 and the M.Sc. and Ph.D degree in wireless communications from the University of Southampton, Southampton, U.K., in 2010 and 2015, respectively. He is now with the College of Information and Engineering, Shenzhen University, China. His research interests include antenna selection, coherent and non-coherent detection, iterative detection, as well as channel estimation.

Figure Captions:

Fig. 1. MSE of the FG-PRIME algorithm versus iteration number, $M = \{64, 80, 96, 120\}$.

Fig. 2. MSE of the FG-PRIME algorithm versus iteration number, SNR=$\{5, 10, 15, 25\}$ dB.

Fig. 3. MSE versus iteration number, M=120, SNR=25 dB.