CONVERGENCE ANALYSIS ON A FAST ITERATIVE PHASE RETRIEVAL ALGORITHM WITHOUT INDEPENDENCE ASSUMPTION

Gen Li, Yuchen Jiao, and Yuantao Gu

Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

ABSTRACT

Phase retrieval has been an attractive problem, and many algorithms have been proposed. Randomized Kaczmarz method is a fast iterative method with good performance in both convergence rate and computational cost with theoretical analysis. However, they all assume that the iteratively updated variables and the original data are independent, which is not true in reality. In this paper, we study for the first time the convergence analysis of this method in the real case, where only a finite number of measurements are available. Specifically, we theoretically prove the linear rate of convergence for the phase retrieval via randomized Kaczmarz algorithm without independence assumption.

***Index Terms*—** Phase Retrieval, Data Reuse, Convergence Analysis, Independence Assumption, randomized Kaczmarz algorithm

1. INTRODUCTION

1.1. Background

The phase retrieval problem is to recover a vector from some magnitude measurements, which is equivalent to solving a system of quadratic equations,

$$y_r = |\langle \mathbf{a}_r, \mathbf{x}^* \rangle|, \quad r = 1, 2, \ldots, m, \tag{1}$$

where $\mathbf{x}^* \in \mathbb{C}^n$ is an unknown signal to be recovered, $\mathbf{a}_r \in \mathbb{C}^n$, y_r, and m denote the known sampling vectors, the rth measurement, and the total number of measurements, respectively. It is generally assumed that the sampling vectors are independent random variables following the distribution $\mathcal{N}\left(0, \frac{1}{2}\mathbf{I}\right) + i\mathcal{N}\left(0, \frac{1}{2}\mathbf{I}\right)$. Apparently, $\mathbf{x}^* \mathrm{e}^{i\theta}$ is also a solution for any $\theta \in [0, 2\pi)$, so the uniqueness of the solution to the phase retrieval problem can only be defined up to a global phase. It has been shown that a unique solution can be determined if $m \geq (4n-4)$ [1, 2]. In real case, that is $\mathbf{x}^* \in \mathbb{R}^n$ and $\mathbf{a}_r \in \mathbb{R}^n$, $\mathbf{a}_r \sim \mathcal{N}(0, \mathbf{I})$ independently. Then $2n-1$ measurements are sufficient.

The phase retrieval problem belongs to the class of the nonconvex quadratic programs. It has appeared frequently in science and engineering, such as X-ray crystallography [3], microscopy [4], astronomy [5], diffraction and array imaging [6], and optics [7]. Other fields of application include acoustics, blind channel estimation in wireless communications, interferometry, quantum mechanics, and quantum information [8]. Focusing on the literature on these physical science fields, one can find the phase retrieval problem commonly encountered, because most sensors can only record the intensity of the light field without the phase information.

The classical algorithms for phase retrieval are the error reduction algorithm and its generalizations [9, 10]. These algorithms alternate between the estimates of the missing phase and the unknown signals iteratively. As suggested by their names, these algorithms satisfy the residual reduction property, but they lack formal theoretical performance guarantees, although these simple iterative algorithms are often shown to be effective in empirical circumstances.

Another popular method, PhaseLift, approaches the problem through reconstructing a rank-one matrix, from which the unknown signal can be obtained [11, 12, 13]. The reconstruction can be solved using tractable semi-definite programming (SDP) based convex relaxations. PhaseLift is known to provide exact solutions (up to a global phase) to the phase retrieval problem using a near minimal number of sampling vectors [1]. However, the computational complexity and memory requirement for SDP based algorithms become prohibitive as the dimension of the signal increases.

Recently, many iterative methods arise including the alternating minimization method [14], phase retrieval via Kaczmarz method [15, 16], and the Wirtinger Flow algorithm and its variants [17, 18], which directly attack the phase retrieval problem in its original nonconvex formulation. In the random online setting, these iterative methods have been shown to achieve linear rate of convergence to the solution. Moreover, [16] establishes an exact analysis of the dynamics of the phase retrieval via Kaczmarz method in the large systems limit.

1.2. Motivation

In current theoretical works on randomized Kaczmarz algorithms for phase retrieval, the independence assumption has always been adopted to make the analysis mathematically easier.

Definition 1 (Independence Assumption) *In an iterative algorithm for solving phase retrieval problem* (1)*, denote* $\mathbf{x}_{t-1}$ *as the temporary estimate of* $\mathbf{x}^*$ *before the* t*th iteration, and* $\mathbf{a}$ *as the sensing vector used to measure* $\mathbf{x}^*$ *and update the estimation in the* t*th iteration. It is assumed that* $\mathbf{x}_{t-1}$ *and* $\mathbf{a}$ *are independent.*

The independence assumption holds in the ideal case, where m approaches infinity, or the scenario of *online processing*, where the sensing vector is randomly generated for every measurement obtained for real-time processing. However, in the real case, the sensing vectors and the corresponding measurements have to be repeatedly used, in that in order to reach a high precision estimate, the number of iterations of the algorithm is usually larger than the number of measurements m. Therefore, the independence assumption does not hold, for the reason that if $\mathbf{a}_r$ and y_r have contributed to

The authors are with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China. This work was partially supported by National Natural Science Foundation of China (NSFC 61531166005, 61571263), the National Key Research and Development Program of China (Project No. 2016YFE0201900, 2017YFC0403600), and Tsinghua University Initiative Scientific Research Program (Grant 2014Z01005). The corresponding author of this work is Y. Gu (E-mail: gyt@tsinghua.edu.cn).

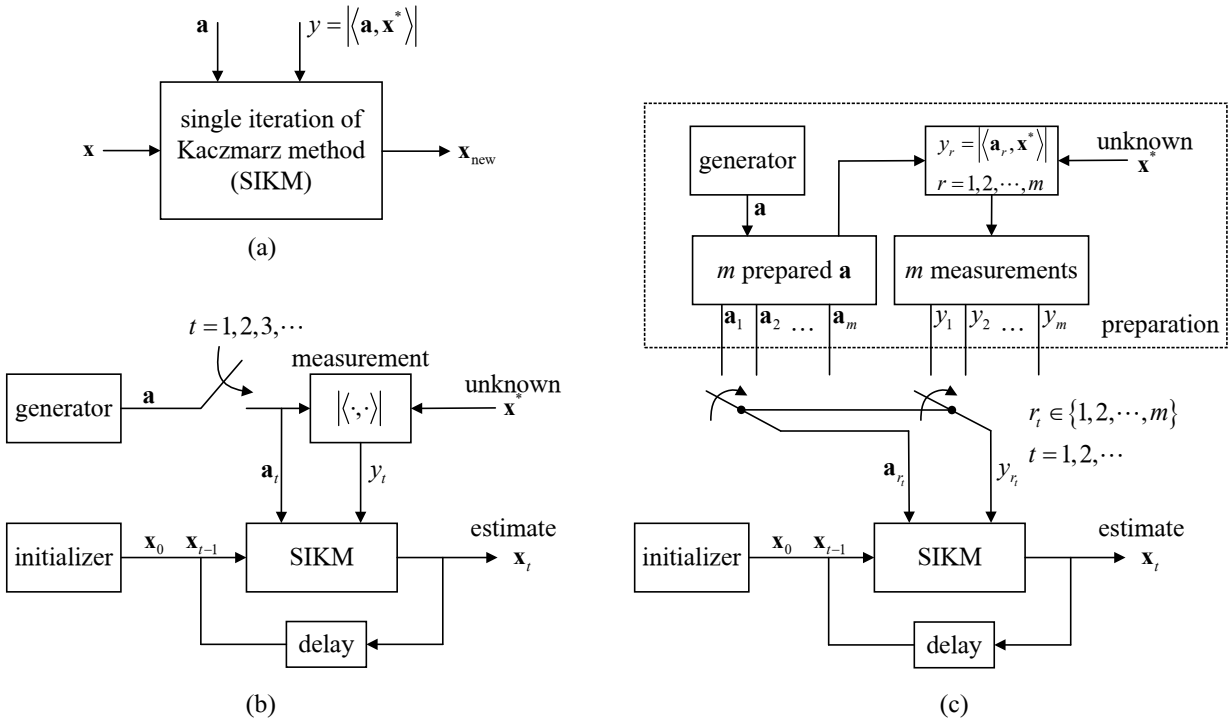

Fig. 1: Visualization of phase retrieval via the randomized Kaczmarz method with both finite and infinite measurements. (a) We introduce a functional block to denote the single iteration of Kaczmarz method (SIKM). (b) The ideal case of infinite measurements or *online processing*, where a new sensing vector is generated in each iteration and used to measure the unknown $\mathbf{x}^*$. Notice that $\mathbf{a}_t$ is independent with the temporary estimate $\mathbf{x}_{t-1}$. (c) The real case with finite measurements. Before processing, m sensing vectors are generated and used to produce m measurements. While in the t-th iteration, a pair of sensing vector and measurement, denoted by $(\mathbf{a}_{r_t}, y_{r_t})$ are randomly chosen from the set and sent to SIKM. Notice that $\mathbf{a}_{r_t}$ is dependent with $\mathbf{x}_{t-1}$ because the former may have already been used to produce the latter in previous iterations.

the estimate $\mathbf{x}_{t-1}$ and are sampled again in the tth iteration, then $\mathbf{x}_{t-1}$ and $\mathbf{a}_r$ are dependent. Please refer to Fig. 1(b) and (c) for a visualization of the ideal online case and the real case.

Though the independence assumption does not hold for the real model with finite measurements in (1), it still plays an important role in the available works, because it alleviates difficulty in the analysis [16, 19, 20]. It is far from accurate and adequate to explain the behavior of algorithms with finite data by using theories founded on ideal scenario of infinite measurements. This will be visualized in Fig. 2 in next section. The number of data is always finite in real-world applications, and we have to reuse the data in many cases, so theoretical analysis in the finite observation setting is of importance in both theory and application, which motivates our study.

According to the simulation results, randomized Kaczmarz method has good performance in both convergence rate and computational cost. As a general row-action method, its computational complexity is only $\mathcal{O}(n)$ per iteration [21, 22].

1.3. Main contribution

In this paper, we study the convergence for randomized Kaczmarz method for the phase retrieval problem. This is the first theoretical analysis on this method without using the independence assumption. To some surprise, we successfully proved that the linear rate of convergence to the solution still holds in the finite measurements setting. This will encourage more theoretical research on the finite measurement setting for phase retrieval. In addition, our methodology may be adopted to analyze other problems, such as low-rank matrix recovery, adaptive filtering, independent component analysis, and other algorithms for phase retrieval problem.

2. PRELIMINARY

In this work we focus on analyzing the randomized Kaczmarz algorithm for phase retrieval in real case, where $\mathbf{x}^* \in \mathbb{R}^n$ and $\mathbf{a}_r \in \mathbb{R}^n$. For random models, we assume that $\mathbf{a}_r \sim \mathcal{N}(0, \mathbf{I})$ are independent for $r = 1, 2, \cdots, m$.

The Phase Retrieval via randomized Kaczmarz algorithm was proposed in [15] and analyzed in [16]. If we know the sign of $\langle \mathbf{a}_r, \mathbf{x}^* \rangle$, according to the Kaczmarz method, $\mathbf{x}_t$ is obtained by projecting $\mathbf{x}_{t-1}$ onto the hyperplane determined by the linear equation $\langle \mathbf{a}_r, \mathbf{x} \rangle = \langle \mathbf{a}_r, \mathbf{x}^* \rangle$. For phase retrieval problem, we can use the sign of $\mathbf{a}_r^{\mathrm{T}} \mathbf{x}_{t-1}$ to evaluate the unknown sign of $\mathbf{a}_r^{\mathrm{T}} \mathbf{x}^*$. Then the iteration becomes

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \frac{y_r \mathrm{sgn}(\mathbf{a}_r^{\mathrm{T}} \mathbf{x}_{t-1}) - \mathbf{a}_r^{\mathrm{T}} \mathbf{x}_{t-1}}{\|\mathbf{a}_r\|^2} \mathbf{a}_r. \tag{2}$$

The algorithm is summarized in Algorithm 1. In order to illustrate our motivation and the independence issue, the implementation in

Algorithm 1 Phase Retrieval via Randomized Kaczmarz algorithm for real case.

Input: $\{(\mathbf{a}_r, y_r), r = 1, \ldots, m\}$, initialization $\mathbf{x}_0$ using the spectral method, $t = 1$.

Ensure: $\mathbf{x}_T$ as an estimate for $\mathbf{x}^*$.

1: **while** $t \leq T$ **do**
2: Choose r randomly from $\{1, \ldots, m\}$ uniformly.
3: Update $\mathbf{x}_t$ by using (2).
4: $t \leftarrow t + 1$.
5: **end while**

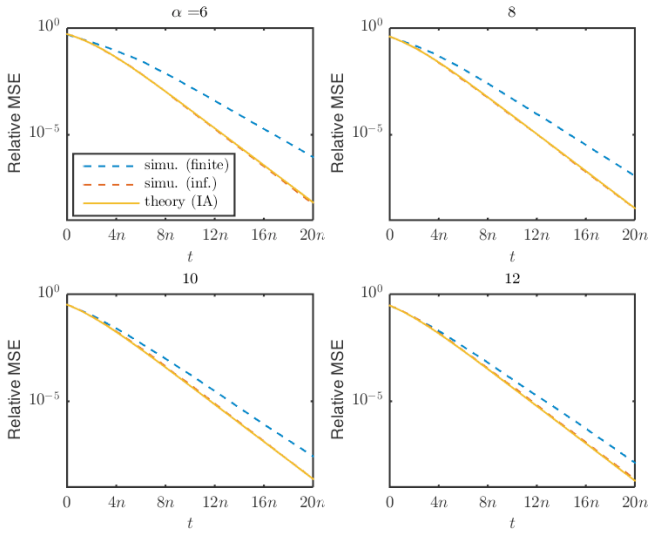

Fig. 2: Learning curves of the randomized Kaczmarz method in the numerical experiments with various α. n is taken as 256. The simulation results with finite and infinite measurements are denoted by "simu.(finite)" and "simu.(inf.)", respectively. The theory with independence assumption are denoted by "theory(IA)".

the case with finite and infinite measurements of this algorithm is further visualized in Fig. 1.

In our earlier work, [16] establishes an exact analysis of the dynamics of the algorithm in the large systems limit under the online setting. Let d_k be the squared distance between the estimate of the algorithm at the kth iteration and the true signal $\mathbf{x}^*$ and $d(t) = d_{\lfloor tn \rfloor}$. As n tends to infinity, the random sample paths of $d(t)$ will converge to a continuous time function governed by the solutions of two deterministic, coupled ordinary differential equations (ODEs) (see Proposition 2). Considering that the online setting satisfies well the independence assumption, our earlier result exactly predicted the learning curve with infinite measurement. However, it diverged obviously from those learning curves reusing a finite dataset. One may refer to Fig. 2. We read that the theory based on independence assumption in [16] coincides with the simulation results using infinite measurements. Whereas the algorithm, which has to exploit data randomly and repeatedly, converges slower than that using infinite data. When the volume of data, i.e., α, increases, the convergence will increase gradually. However, even if α reaches 12, which is the requirement of many classical methods for successful recovery, the distinction is still obvious. This highlights the importance of our work

Recently, theoretical guarantees for Phase Retrieval via Randomized Kaczmarz method were provided in (see Theorem 1.2 [19] and Theorem 1.1 [20]) with the independence assumption between the iterative variable and the sampling vectors implicitly, i.e., Lemma 2.2 [19], Section 2.1 [20]. As far as we know, there is no state-of-the-art theoretical work on randomized Kaczmarz algorithm without using the independence assumption.

3. MAIN RESULT

In Theorem 1, we provide the linear rate of convergence for the randomized Kaczmarz algorithm for the phase retrieval problem of real case.

Theorem 1 *Let $\mathbf{x}^* \in \mathbb{R}^n$ be any solution to the phase retrieval problem* (1) *in the real case and $\mathbf{x}_t$ denotes the tth iterative solution of Algorithm 1. Define $\mathbf{e}_t = \mathbf{x}_t - \mathbf{x}^*$ as the estimate error. Then for $m = \alpha n$, with probability at least $1 - \mathrm{e}^{-cn}$ where c is a constant depending on ε,*

$$\frac{\mathbb{E}\left\|\mathbf{e}_t\right\|^2 - \left\|\mathbf{e}_{t-1}\right\|^2}{\left\|\mathbf{e}_{t-1}\right\|^2/n} \leq -\left(1 - \frac{p}{\alpha}\right)\left(1 - \frac{1}{\sqrt{\alpha - p}} - \sqrt{\frac{2p}{\alpha - p}\ln\frac{\mathrm{e}\alpha}{p}}\right)^2 + \frac{3p}{\alpha}\left(1 + \frac{1}{\sqrt{p}} + \sqrt{2\ln\frac{\mathrm{e}\alpha}{p}}\right)^2 + \varepsilon, \quad (3)$$

where $p \in (0, \alpha)$ is a parameter determined by

$$\begin{cases} \left(1 - \frac{\alpha}{p}\frac{2\tau}{\sqrt{2\pi}}\mathrm{e}^{-\frac{\tau^2}{2}}\right)\left\|\mathbf{x}^*\right\|^2 \leq \left(1 + \frac{1}{\sqrt{p}} + \sqrt{2\ln\frac{\mathrm{e}\alpha}{p}}\right)^2 \left\|\mathbf{e}_{t-1}\right\|^2; \\ 1 - 2Q\left(\tau\right) = \frac{p}{\alpha}, \end{cases}$$

where τ is a variable and $Q(\cdot)$ is the tail probability of the standard Gaussian distribution.

Remark 1 *Notice that we do not take the expectation of $\left\|\mathbf{e}_{t-1}\right\|^2$ in the LHS of* (3)*. The bound in* (3) *holds for arbitrary $\mathbf{e}_{t-1}$ even if it is a function of the sensing vectors, which is dependent with $\{\mathbf{a}_r, r = 1, \ldots, m\}$. On the contrary, in existing works on such convergence analysis, it is always assumed that $\mathbf{e}_{t-1}$ is a random vector independent of the sensing vectors, although this assumption is unjustified for the problem.*

Remark 2 *When $\left\|\mathbf{e}_{t-1}\right\|$ is small enough, p approaches zero monotonically. Then* (3) *is reduced to*

$$\frac{\mathbb{E}\left\|\mathbf{e}_t\right\|^2 - \left\|\mathbf{e}_{t-1}\right\|^2}{\left\|\mathbf{e}_{t-1}\right\|^2/n} \leq -\left(1 - \frac{1}{\sqrt{\alpha}}\right)^2 + \frac{3}{\alpha} + \varepsilon. \quad (4)$$

We read from (4) *that when n is large enough and the initialization is good enough so that $\left\|\mathbf{e}_0\right\|$ is small enough, Algorithm 1 will converge to a solution exponentially if*

$$\alpha > \left(1 + \sqrt{3}\right)^2 = 7.46.$$

There are already some ways to get good initialization, such as the truncated spectral method [18].

PROOF Due to the limited space, we only present a sketch of the proof, which can be divided into three steps. The detailed proof is included in [23].

Step1 : Subtracting $\mathbf{x}^*$ from both sides of (2), we write the recursion formula on estimate error as

$$\mathbf{e}_t = \left(\mathbf{I} - \frac{\mathbf{a}_r\mathbf{a}_r^{\mathrm{T}}}{\left\|\mathbf{a}_r\right\|^2}\right)\mathbf{e}_{t-1} + \frac{\mathbf{a}_r^{\mathrm{T}}\mathbf{x}^*\left(\mathrm{sgn}\left(\mathbf{a}_r^{\mathrm{T}}\mathbf{x}^*\mathbf{a}_r^{\mathrm{T}}\mathbf{x}_{t-1}\right) - 1\right)}{\left\|\mathbf{a}_r\right\|^2}\mathbf{a}_r. \quad (5)$$

By introducing $\mathcal{S}$ to denote the set of sensing vectors of which the sign of $\mathbf{a}_r^{\mathrm{T}}\mathbf{x}^*$ are wrongly estimated

$$\mathcal{S}=\left\{k:\ \operatorname{sgn}\left(\mathbf{a}_k^{\mathrm{T}}\mathbf{x}_{t-1}\right)\neq\operatorname{sgn}\left(\mathbf{a}_k^{\mathrm{T}}\mathbf{x}^*\right)\right\},\tag{6}$$

and $p=|\mathcal{S}|/n$ as the ratio of *wrong* measurements number to unknown number, we discuss the estimate error on two cases of $r\in\mathcal{S}$ and $r\in\bar{\mathcal{S}}=\{1,2,\cdots,m\}\backslash\mathcal{S}$. Noticing that $\left(\mathbf{I}-\mathbf{a}_r\mathbf{a}_r^{\mathrm{T}}/\|\mathbf{a}_r\|^2\right)$ is the projection matrix of the hyperplane perpendicular to $\mathbf{a}_r$, we derive

$$\begin{aligned}\|\mathbf{e}_t\|^2=&\|\mathbf{e}_{t-1}\|^2-\left(\frac{\mathbf{a}_r^{\mathrm{T}}\mathbf{e}_{t-1}}{\|\mathbf{a}_r\|}\right)^2\mathbb{I}_{r\in\bar{\mathcal{S}}}\\&+\left(4\left(\frac{\mathbf{a}_r^{\mathrm{T}}\mathbf{x}^*}{\|\mathbf{a}_r\|}\right)^2-\left(\frac{\mathbf{a}_r^{\mathrm{T}}\mathbf{e}_{t-1}}{\|\mathbf{a}_r\|}\right)^2\right)\mathbb{I}_{r\in\mathcal{S}},\end{aligned}\tag{7}$$

where $\mathbb{I}$ is the indicator function. Then take the expectation of (7) with respect to r which is uniformly distributed in $\{1,\ldots,m\}$, we have

$$\begin{aligned}\mathbb{E}\|\mathbf{e}_t\|^2=&\|\mathbf{e}_{t-1}\|^2-\frac{1}{m}\sum_{k\in\bar{\mathcal{S}}}\left(\frac{\mathbf{a}_k^{\mathrm{T}}\mathbf{e}_{t-1}}{\|\mathbf{a}_k\|}\right)^2\\&+\frac{1}{m}\sum_{k\in\mathcal{S}}\left(-\left(\frac{\mathbf{a}_k^{\mathrm{T}}\mathbf{e}_{t-1}}{\|\mathbf{a}_k\|}\right)^2+4\left(\frac{\mathbf{a}_k^{\mathrm{T}}\mathbf{x}^*}{\|\mathbf{a}_k\|}\right)^2\right).\end{aligned}\tag{8}$$

Notice that here we cannot simply take expectation of (7) with respect to $\mathbf{a}_r$ with $\mathbf{e}_{t-1}$ fixed. The reason is that $\mathbf{a}_r$ is dependent on $\mathbf{e}_{t-1}$ and is not a Gaussian random vector any more conditioning on $\mathbf{e}_{t-1}$. This is the key point of our approach of abandoning the independence assumption.

Now we are ready to transform the second and the third item in the RHS of (8) to the problems of estimating eigenvalues of random matrices.

Step2 : For the second item, we can get

$$\frac{1}{m}\sum_{k\in\bar{\mathcal{S}}}\left(\frac{\mathbf{a}_k^{\mathrm{T}}\mathbf{e}_{t-1}}{\|\mathbf{a}_k\|}\right)^2\geq\frac{1}{z_{\mathrm{M}}^2}\frac{\alpha-p}{\alpha}\lambda_{\mathrm{m}}\left(\mathbf{A}_{\bar{\mathcal{S}}}\right)\|\mathbf{e}_{t-1}\|^2,\tag{9}$$

where $\lambda_{\mathrm{m}}(\cdot)$ denotes the smallest eigenvalue of a matrix, and

$$\begin{aligned}\mathbf{A}_{\bar{\mathcal{S}}}=&\frac{1}{|\bar{\mathcal{S}}|}\sum_{k\in\bar{\mathcal{S}}}\mathbf{a}_k\mathbf{a}_k^{\mathrm{T}},\\z_{\mathrm{M}}=&\max_{k=1,\cdots,m}\|\mathbf{a}_k\|.\end{aligned}$$

For the third item, noticing that for $k\in\mathcal{S}$, we read

$$\left|\mathbf{a}_k^{\mathrm{T}}\mathbf{x}^*\right|\leq\left|\mathbf{a}_k^{\mathrm{T}}\mathbf{e}_{t-1}\right|,$$

then we have

$$\begin{aligned}&\frac{1}{m}\sum_{k\in\mathcal{S}}\left(-\left(\frac{\mathbf{a}_k^{\mathrm{T}}\mathbf{e}_{t-1}}{\|\mathbf{a}_k\|}\right)^2+4\left(\frac{\mathbf{a}_k^{\mathrm{T}}\mathbf{x}^*}{\|\mathbf{a}_k\|}\right)^2\right)\\\leq&\frac{3}{m}\sum_{k\in\mathcal{S}}\left(\frac{\mathbf{a}_k^{\mathrm{T}}\mathbf{e}_{t-1}}{\|\mathbf{a}_k\|}\right)^2\\\leq&\frac{3}{z_{\mathrm{m}}^2}\frac{p}{\alpha}\lambda_{\mathrm{M}}\left(\mathbf{A}_{\mathcal{S}}\right)\|\mathbf{e}_{t-1}\|^2,\end{aligned}\tag{10}$$

where $\lambda_{\mathrm{M}}(\cdot)$ denotes the largest eigenvalue of a matrix, and

$$\begin{aligned}\mathbf{A}_{\mathcal{S}}=&\frac{1}{|\mathcal{S}|}\sum_{k\in\mathcal{S}}\mathbf{a}_k\mathbf{a}_k^{\mathrm{T}},\\z_{\mathrm{m}}=&\min_{k\in1,\cdots,m}\|\mathbf{a}_k\|.\end{aligned}$$

Notice that these two bounds (9) and (10) hold no matter what dependence $\mathbf{e}_{t-1}$ and $\mathbf{a}_r$ have. Actually, as what we mentioned above, $\mathbf{e}_{t-1}$ and $\mathbf{a}_r$ are often dependent, which makes this operation correct and necessary. The inappropriate independence assumption is successfully abandoned for the first time. This is different from all the previous theoretical works.

Step3 : Substituting (9) and (10) into (8), we have

$$\mathbb{E}\|\mathbf{e}_t\|^2\leq\left(1-\frac{1}{z_{\mathrm{M}}^2}\frac{\alpha-p}{\alpha}\lambda_{\mathrm{m}}\left(\mathbf{A}_{\bar{\mathcal{S}}}\right)+\frac{3}{z_{\mathrm{m}}^2}\frac{p}{\alpha}\lambda_{\mathrm{M}}\left(\mathbf{A}_{\mathcal{S}}\right)\right)\|\mathbf{e}_{t-1}\|^2.$$

Notice that the dependence problem disappears in this step, since the minimal or maximal eigenvalues of $\mathbf{A}$ can be estimated only by itself, and is independent from $\mathbf{e}_{t-1}$. That is, according to random matrices theory [24], we have

$$\mathbb{P}\left(\lambda_{\mathrm{m}}\left(\mathbf{A}_{\bar{\mathcal{S}}}\right)<\left(1-\frac{1}{\sqrt{\alpha-p}}-\varepsilon_3\right)^2\right)<\mathrm{C}_m^{(\alpha-p)n}\mathrm{e}^{-\frac{(\alpha-p)n\varepsilon_3^2}{2}},$$

and

$$\mathbb{P}\left(\lambda_{\mathrm{M}}\left(\mathbf{A}_{\mathcal{S}}\right)>\left(1+\frac{1}{\sqrt{p}}+\varepsilon_2\right)^2\right)<\mathrm{C}_m^{pn}\mathrm{e}^{-\frac{pn\varepsilon_2^2}{2}}.$$

Then with probability at least

$$1-2m\mathrm{e}^{-n\left(\frac{\varepsilon_1^2}{4}-\frac{\varepsilon_1^3}{6}\right)}-\mathrm{C}_m^{pn}\mathrm{e}^{-\frac{pn\varepsilon_2^2}{2}}-\mathrm{C}_m^{(\alpha-p)n}\mathrm{e}^{-\frac{(\alpha-p)n\varepsilon_3^2}{2}},$$

we have

$$\begin{aligned}\mathbb{E}\|\mathbf{e}_t\|^2\leq&\|\mathbf{e}_{t-1}\|^2\left(1-\left(1-\frac{p}{\alpha}\right)\frac{1}{n(1+\varepsilon_1)}\left(1-\frac{1}{\sqrt{\alpha-p}}-\varepsilon_3\right)^2\right.\\&\left.+\frac{3p}{\alpha}\frac{1}{n\left(1-\varepsilon_1\right)}\left(1+\frac{1}{\sqrt{p}}+\varepsilon_2\right)^2\right).\end{aligned}$$

After simplification, we complete the proof. ∎

4. CONCLUSION

In this paper, we reveal the linear rate of convergence of the randomized Kaczmarz method in the real case with finite measurements. Different from the previous theoretical works, the inappropriate independence assumption is abandoned in the analysis for this method for the first time. A good upper bound of the rate of convergence is given. Moreover, our methodology could be adopted to analyze other problems.

5. REFERENCES

[1] R. Balan, P. Casazza, and D. Edidin, "On signal reconstruction without phase," *Applied and Computational Harmonic Analysis*, vol. 20, no. 3, pp. 345–356, 2006.

[2] A. Conca, D. Edidin, M. Hering, and C. Vinzant, "An algebraic characterization of injectivity in phase retrieval," *Applied and Computational Harmonic Analysis*, vol. 38, no. 2, pp. 346–356, 2015.

[3] R. W. Harrison, "Phase problem in crystallography," *JOSA a*, vol. 10, no. 5, pp. 1046–1055, 1993.

[4] J. Miao, T. Ishikawa, Q. Shen, and T. Earnest, "Extending x-ray crystallography to allow the imaging of noncrystalline materials, cells, and single protein complexes," *Annu. Rev. Phys. Chem.*, vol. 59, pp. 387–410, 2008.

[5] C. Fienup and J. Dainty, "Phase retrieval and image reconstruction for astronomy," *Image Recovery: Theory and Application*, pp. 231–275, 1987.

[6] O. Bunk, A. Diaz, F. Pfeiffer, C. David, B. Schmitt, D. K. Satapathy, and J. F. van der Veen, "Diffractive imaging for periodic samples: retrieving one-dimensional concentration profiles across microfluidic channels," *Acta Crystallographica Section A: Foundations of Crystallography*, vol. 63, no. 4, pp. 306–314, 2007.

[7] A. Walther, "The question of phase retrieval in optics," *Journal of Modern Optics*, vol. 10, no. 1, pp. 41–49, 1963.

[8] 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 processing magazine*, vol. 32, no. 3, pp. 87–109, 2015.

[9] R. W. Gerchberg, "A practical algorithm for the determination of the phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237–246, 1972.

[10] J. R. Fienup, "Phase retrieval algorithms: a comparison," *Applied optics*, vol. 21, no. 15, pp. 2758–2769, 1982.

[11] E. J. Candes, T. Strohmer, and V. Voroninski, "Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming," *Communications on Pure and Applied Mathematics*, vol. 66, no. 8, pp. 1241–1274, 2013.

[12] E. J. Candès and X. Li, "Solving quadratic equations via phaselift when there are about as many equations as unknowns," *Foundations of Computational Mathematics*, vol. 14, no. 5, pp. 1017–1026, 2014.

[13] I. Waldspurger, A. d'Aspremont, and S. Mallat, "Phase recovery, maxcut and complex semidefinite programming," *Mathematical Programming*, vol. 149, no. 1-2, pp. 47–81, 2015.

[14] P. Netrapalli, P. Jain, and S. Sanghavi, "Phase retrieval using alternating minimization," in *Advances in Neural Information Processing Systems*, 2013, pp. 2796–2804.

[15] K. Wei, "Solving systems of phaseless equations via kaczmarz methods: A proof of concept study," *Inverse Problems*, vol. 31, no. 12, p. 125008, 2015.

[16] G. Li, Y. Gu, and Y. M. Lu, "Phase retrieval using iterative projections: Dynamics in the large systems limit," in *Communication, Control, and Computing (Allerton), 2015 53rd Annual Allerton Conference on*. IEEE, 2015, pp. 1114–1118.

[17] E. J. Candes, X. Li, and M. Soltanolkotabi, "Phase retrieval via wirtinger flow: Theory and algorithms," *IEEE Transactions on Information Theory*, vol. 61, no. 4, pp. 1985–2007, 2015.

[18] Y. Chen and E. Candes, "Solving random quadratic systems of equations is nearly as easy as solving linear systems," in *Advances in Neural Information Processing Systems*, 2015, pp. 739–747.

[19] Y. S. Tan and R. Vershynin, "Phase retrieval via randomized kaczmarz: Theoretical guarantees," *arXiv preprint arXiv:1706.09993*, 2017.

[20] H. Jeong and C. S. Güntürk, "Convergence of the randomized kaczmarz method for phase retrieval," *arXiv preprint arXiv:1706.10291*, 2017.

[21] S. Kaczmarz, "Angenaherte auflosung von systemen linearer gleichungen," *Bull. Int. Acad. Sci. Pologne, A*, vol. 35, pp. 355–357, 1937.

[22] T. Strohmer and R. Vershynin, "A randomized solver for linear systems with exponential convergence," *Lecture Notes in Computer Science*, vol. 4110, p. 499, 2006.

[23] G. Li, Y. Jiao, and Y. Gu, "Linear convergence of an iterative phase retrieval algorithm with data reuse," 2017.

[24] R. Vershynin, "Introduction to the non-asymptotic analysis of random matrices," *Eprint Arxiv*, 2010.