

Article

Not peer-reviewed version

On the Convergence of Randomized Block Kaczmarz Algorithm for Solving Matrix Equation

[Lili Xing](#) , Wendi Bao , [Weiguo Li](#) *

Posted Date: 11 October 2023

doi: 10.20944/preprints202310.0686.v1

Keywords: matrix equation; randomized block Kaczmarz; randomized extended block Kaczmarz; convergence



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

On the Convergence of Randomized Block Kaczmarz Algorithm for Solving Matrix Equation

Lili Xing, Wendi Bao and Weiguo Li *

College of Science, China University of Petroleum, Qingdao 266580, China

* Correspondence: liwg@upc.edu.cn

Abstract: Randomized block Kaczmarz Method and randomized extended block Kaczmarz Method are proposed for solving matrix equation $AXB = C$, where the matrix A and B may be full rank or rank deficient. These methods are iterative methods without matrix multiplication, and are especially suitable for solving large-scale matrix equations. It is theoretically proved these methods converge to the solution or least square solution of the matrix equation. The numerical results show that these methods are more efficient than the existing algorithms for high-dimensional matrix equations.

Keywords: matrix equation; randomized block Kaczmarz; randomized extended block Kaczmarz; convergence

MSC: 65F10; 65F45; 65H10

1. Introduction

Consider the linear matrix equation

$$AXB = C, \quad (1)$$

where $A \in \mathbb{R}^{m \times p}$, $B \in \mathbb{R}^{q \times n}$ and $C \in \mathbb{R}^{m \times n}$. Such problems arise in linear control and filtering theory for continuous or discrete-time large-scale dynamical systems. If $AXB = C$ is consistent, $X^* = A^\dagger CB^\dagger$ is the minimum Frobenius norm solution. If $AXB = C$ is inconsistent, $X^* = A^\dagger CB^\dagger$ is the minimum Frobenius norm least squares solution. When the matrices A and B are small and dense, direct methods based on QR-fraction are attractive [1,2]. However, for large matrices A and B , iterative methods have attracted much attention [3–7]. Recently, Du et al. proposed the randomized block coordinate descent (RBCD) method for solving the matrix least-squares problem $\min_{X \in \mathbb{R}^{p \times q}} \|C - AXB\|_F$ without strong convexity assumption in [8]. This method requires that matrix B is full row rank. Wu et al. [9] introduced two kinds of Kaczmarz-type methods to solve consistent matrix equation $AXB = C$: relaxed greedy randomized Kaczmarz (ME-RGRK) and maximal weighted residual Kaczmarz (ME-MWRK). Although the row and column index selection strategy is time-consuming, the ideas of these two methods are suitable for solving large-scale consistent matrix equations.

In this paper, randomized Kaczmarz method [10] and randomized extended Kaczmarz method [11] are used to solve consistent and inconsistent matrix equations (1) by the product of matrix and vector.

All the results in this paper hold in the complex field. But for the sake of simplicity, we only discuss it in the real number field.

In this paper, we denote A^T , A^\dagger , $r(A)$, $R(A)$, $\|A\|_F = \sqrt{\text{trace}(A^T A)}$ and $\langle A, B \rangle_F = \text{trace}(A^T B)$ as the transpose, the Moore-Penrose generalized inverse, the rank of A , the column space of A , the Frobenius norm of A and the inner product of two matrices A and B , respectively. For an integer $n \geq 1$, let $[n] = \{1, 2, \dots, n\}$. We use I to denote the identity matrix whose order is clear from the context. In addition, for a given matrix $G = (g_{ij}) \in \mathbb{R}^{m \times n}$, $G_{i,:}$, $G_{:,j}$, $\sigma_{\max}(G)$ and $\sigma_{\min}(G)$, are used to denote i th row, j th column, the maximum singular value and the smallest nonzero singular value of G respectively. Let E_k denote the expected value conditional on the first k iterations, that is, $\mathbb{E}_k[\cdot] = \mathbb{E}[\cdot | i_0, j_0, i_1, j_1, \dots, i_{k-1}, j_{k-1}]$, where i_s and j_s ($s = 0, 1, \dots, k-1$) are the row and the

column chosen at the s th iteration. Let the conditional expectations with respect to the random row index be $\mathbb{E}_k^i[\cdot] = \mathbb{E}[\cdot | i_0, j_0, i_1, j_1, \dots, i_{k-1}, j_{k-1}, j_k]$ and with respect to the random column index be $\mathbb{E}_k^j[\cdot] = \mathbb{E}[\cdot | i_0, j_0, i_1, j_1, \dots, i_{k-1}, j_{k-1}, i_k]$. By the law of total expectation, it holds that $\mathbb{E}_k[\cdot] = \mathbb{E}_k^i[\mathbb{E}_k^j[\cdot]]$.

The organization of this paper is as follows. In Section 2, we will discuss Kaczmarz method (ME-RBK and ME-PRBK) for finding the minimal F -norm solution ($A^\dagger CB^\dagger$) of the consistent matrix Eq. (1). In Section 3, we discuss the extended Kaczmarz method (IME-REBK) for finding the minimal F -norm least-squares solution of the matrix Eq. (1). In Section 4, some numerical examples are provided to illustrate the effectiveness of our new methods. Finally, some brief concluding remarks are described in Section 5.

2. The Randomized Block Kaczmarz Method for Consistent Equation

At the k th iteration, Kaczmarz method selects randomized a row $i \in [m]$ of A and does an orthogonal projection of the current estimate matrix $X^{(k)}$ onto the corresponding hyperplane $H_i = \{X \in \mathbb{R}^{p \times q} | A_{i,:}XB = C_{i,:}\}$, that is

$$\min_{X \in \mathbb{R}^{p \times q}} \frac{1}{2} \|X - X^{(k)}\|_F^2 \text{ s.t. } A_{i,:}XB = C_{i,:}. \quad (2)$$

The Lagrangian function of the conditional optimization problem (2) is

$$L(X, Y) = \frac{1}{2} \|X - X^{(k)}\|_F^2 + \langle Y, A_{i,:}XB - C_{i,:} \rangle, \quad (3)$$

where $Y \in \mathbb{R}^{1 \times n}$ is Lagrangian multiplier. By the matrix differentiation, we get the gradient of $L(X, Y)$ and set $\nabla L(X, Y) = 0$ for finding the stationary matrix:

$$\begin{cases} \nabla_X L(X, Y)|_{X^{(k+1)}} = X^{(k+1)} - X^{(k)} + A_{i,:}^T Y B^T = 0, \\ \nabla_Y L(X, Y)|_{X^{(k+1)}} = A_{i,:} X^{(k+1)} B - C_{i,:} = 0. \end{cases} \quad (4)$$

By the first equation of (4), we have $X^{(k+1)} = X^{(k)} - A_{i,:}^T Y B^T$. Substituting this into the second equation of (4), we can get $Y = -\frac{1}{\|A_{i,:}\|_2^2} (C_{i,:} - A_{i,:} X^{(k)} B) (B^T B)^\dagger$. So the projected randomized block Kaczmarz (ME-PRBK) for solving $AXB = C$ iterates as

$$X^{(k+1)} = X^{(k)} + \frac{A_{i,:}^T}{\|A_{i,:}\|_2^2} (C_{i,:} - A_{i,:} X^{(k)} B) B^\dagger. \quad (5)$$

However, in practice, it is very expensive to calculate the pseudoinverse of large-scale matrices. Next, we generalize the average block Kaczmarz method [12] for solving linear equations to matrix equations.

At the k th step, we get the approximate solution $X^{(k+1)}$ by projecting the current estimate $X^{(k)}$ onto the hyperplane $H_{i,j} = \{X \in \mathbb{R}^{p \times q} | A_{i,:} X^{(k)} B_{:,j} = C_{i,j}\}$. Using the Lagrangian multiplier method, we can obtain the following Kaczmarz method for $AXB = C$:

$$X^{(k+1)} = X^{(k)} + \frac{A_{i,:}^T (C_{i,:} - A_{i,:} X^{(k)} B_{:,j}) B_{:,j}^T}{\|A_{i,:}\|_2^2 \|B_{:,j}\|_2^2}. \quad (6)$$

Inspired by the idea of average block Kaczmarz algorithm for $Ax = b$, we consider the average block Kaczmarz method for $AXB = C$ respect to B .

$$X^{(k+1)} = X^{(k)} + \lambda \frac{A_{i,:}^T}{\|A_{i,:}\|_2^2} \sum_{j=1}^n v_j^k \frac{(C_{i,:} - A_{i,:} X^{(k)} B_{:,j}) B_{:,j}^T}{\|B_{:,j}\|_2^2}, \quad (7)$$

where λ is stepsize and v_j^k be the weights that satisfy $v_j^k \geq 0$ and $\sum_{j=1}^n v_j^k = 1$. If $v_j^k = \frac{\|B_{:,j}\|_2^2}{\|B\|_F^2}$, then

$$X^{(k+1)} = X^{(k)} + \frac{\lambda}{\|B\|_F^2} \frac{A_{i,:}^T}{\|A_{i,:}\|_2^2} (C_{i,:} - A_{i,:} X^{(k)} B) B^T. \quad (8)$$

Set $\alpha = \frac{\lambda}{\|B\|_F^2} > 0$, we obtain the following randomized block Kaczmarz iteration

$$X^{(k+1)} = X^{(k)} + \frac{\alpha}{\|A_{i,:}\|_2^2} A_{i,:}^T \left((C_{i,:} - (A_{i,:} X^{(k)}) B) B^T \right), k = 0, 1, 2, \dots, \quad (9)$$

where i is selected with probability $p_i = \frac{\|A_{i,:}\|_2^2}{\|A\|_F^2}$. The cost of each iteration of this method is $4q(n+p) + p+1-2q$ if the square of the row norm of A has been calculated in advance. We describe this method as Algorithm 1, which is called the ME-RBK algorithm.

Algorithm 1 Randomized Block Kaczmarz Method for $AXB = C$ (ME-RBK)

Input: $A \in \mathbb{R}^{m \times p}$, $B \in \mathbb{R}^{q \times n}$, $C \in \mathbb{R}^{m \times n}$, $X^{(0)} = 0 \in \mathbb{R}^{p \times q}$

1: **for** $k = 0, 1, 2, \dots$, **do**

2: Pick i with probability $p_i(A) = \frac{\|A_{i,:}\|_2^2}{\|A\|_F^2}$

3: Compute $X^{(k+1)} = X^{(k)} + \frac{\alpha}{\|A_{i,:}\|_2^2} A_{i,:}^T \left((C_{i,:} - (A_{i,:} X^{(k)}) B) B^T \right)$

4: **end for**

Remark 1. Note that the problem of finding a solution of $AXB = C$ can be posed as the following linear least squares problem

$$\min_{X \in \mathbb{R}^{p \times q}} \frac{1}{2} \|AXB - C\|_F^2 = \min_{X \in \mathbb{R}^{p \times q}} \frac{1}{2} \sum_{i=1}^m \|A_{i,:} X B - C_{i,:}\|_2^2. \quad (10)$$

Define the component function

$$f_i(X) = \frac{1}{2} \|A_{i,:} X B - C_{i,:}\|_2^2,$$

then differentiate with X to get its gradient

$$\nabla f_i(X) = A_{i,:}^T (A_{i,:} X B - C_{i,:}) B^T.$$

Therefore, the randomized block Kaczmarz method (9) is equivalent to one step of the stochastic gradient decent method [13] applied to (10) with stepsize $\frac{\alpha}{\|A_{i,:}\|_2^2}$.

First, we give the following lemma whose proof can be found in [8].

Lemma 1. [8] Let $A \in \mathbb{R}^{m \times p}$ and $B \in \mathbb{R}^{q \times n}$ be any nonzero matrix. Let

$$\mathcal{M} = \{M \in \mathbb{R}^{p \times q} \mid \exists Y \in \mathbb{R}^{m \times n} \text{ s.t. } M = A^T Y B^T\}.$$

For any matrix $M \in \mathcal{M}$, it holds

$$\|AMB\|_F^2 \geq \sigma_{\min}^2(A) \sigma_{\min}^2(B) \|M\|_F^2.$$

Remark 2. $M \in \mathcal{M}$ means that $M_{:,j} \in R(A^T)$, $j = 1, 2, \dots, q$ and $(M_{i,:})^T \in R(B)$, $i = 1, 2, \dots, p$. In fact, \mathcal{M} is well defined because $0 \in \mathcal{M}$ and $A^\dagger C B^\dagger \in \mathcal{M}$.

In the following theorem, with the idea of the RK method [10], we will prove that Algorithm 1 converges to the the least F -norm solution of $AXB = C$.

Theorem 1. Assume $0 < \alpha < \frac{2}{\|B\|_2^2}$. If Eq. (1) is consistent, the sequence $\{X^{(k)}\}$ generated by the ME-RBK method starting from the initial matrix $X^{(0)} \in \mathbb{R}^{p \times q}$ in which $X_{:,j}^{(0)} \in R(A^T), j = 1, 2, \dots, q$ and $(X_{i,:}^{(0)})^T \in R(B), i = 1, 2, \dots, p$, converges linearly to $A^\dagger CB^\dagger$ in mean square form. Moreover, the solution error in expectation for the iteration sequence $X^{(k)}$ obeys

$$E \left[\|X^{(k)} - A^\dagger CB^\dagger\|_F^2 \right] \leq \rho^k \|X^{(0)} - A^\dagger CB^\dagger\|_F^2, \quad (11)$$

where $\rho = 1 - \frac{2\alpha - \alpha^2 \|B\|_2^2}{\|A\|_F^2} \sigma_{\min}^2(A) \sigma_{\min}^2(B)$, and $i \in [m]$ picked with probability $p_i(A) = \frac{\|A_{i,:}\|_2^2}{\|A\|_F^2}$.

Proof. For $k = 0, 1, 2, \dots$, by (9) and $A_{i,:} A^\dagger CB^\dagger B = C_{i,:}$ (consistency), we have

$$\begin{aligned} X^{(k+1)} - A^\dagger CB^\dagger &= X^{(k)} + \frac{\alpha}{\|A_{i,:}\|_2^2} A_{i,:}^T (C_{i,:} - A_{i,:} X^{(k)} B) B^T - A^\dagger CB^\dagger \\ &= (X^{(k)} - A^\dagger CB^\dagger) - \frac{\alpha}{\|A_{i,:}\|_2^2} A_{i,:}^T A_{i,:} (X^{(k)} - A^\dagger CB^\dagger) B B^T, \end{aligned}$$

then

$$\begin{aligned} \|X^{(k+1)} - A^\dagger CB^\dagger\|_F^2 &= \|X^{(k)} - A^\dagger CB^\dagger\|_F^2 + \frac{\alpha^2}{\|A_{i,:}\|_2^4} \|A_{i,:}^T A_{i,:} (X^{(k)} - A^\dagger CB^\dagger) B B^T\|_F^2 \\ &\quad - \frac{2\alpha}{\|A_{i,:}\|_2^2} \langle X^{(k)} - A^\dagger CB^\dagger, A_{i,:}^T A_{i,:} (X^{(k)} - A^\dagger CB^\dagger) B B^T \rangle_F. \end{aligned}$$

It follows from

$$\begin{aligned} &\frac{\alpha^2}{\|A_{i,:}\|_2^4} \|A_{i,:}^T A_{i,:} (X^{(k)} - A^\dagger CB^\dagger) B B^T\|_F^2 \\ &= \frac{\alpha^2}{\|A_{i,:}\|_2^2} \|A_{i,:} (X^{(k)} - A^\dagger CB^\dagger) B B^T\|_2^2 \text{ (by } \text{trace}(uu^T) = \|u\|_2^2 \text{ for any vector } u) \\ &\leq \frac{\alpha^2 \|B\|_2^2}{\|A_{i,:}\|_2^2} \|A_{i,:} (X^{(k)} - A^\dagger CB^\dagger) B\|_2^2 \text{ (by } \|u^T B^T\|_2 = \|Bu\|_2 \leq \|B\|_2 \|u\|_2), \end{aligned}$$

and

$$\begin{aligned} &\frac{2\alpha}{\|A_{i,:}\|_2^2} \langle X^{(k)} - A^\dagger CB^\dagger, A_{i,:}^T A_{i,:} (X^{(k)} - A^\dagger CB^\dagger) B B^T \rangle_F \\ &= \frac{2\alpha}{\|A_{i,:}\|_2^2} \text{trace}(B^T (X^{(k)} - A^\dagger CB^\dagger)^T A_{i,:}^T A_{i,:} (X^{(k)} - A^\dagger CB^\dagger) B) \text{ (by } \text{trace}(MN) = \text{trace}(NM)) \\ &= \frac{2\alpha}{\|A_{i,:}\|_2^2} \|A_{i,:} (X^{(k)} - A^\dagger CB^\dagger) B\|_2^2 \end{aligned}$$

that

$$\|X^{(k+1)} - A^\dagger CB^\dagger\|_F^2 \leq \|X^{(k)} - A^\dagger CB^\dagger\|_F^2 - \frac{2\alpha - \alpha^2 \|B\|_2^2}{\|A_{i,:}\|_2^2} \|A_{i,:} (X^{(k)} - A^\dagger CB^\dagger) B\|_2^2.$$

By taking the conditional expectation, we have

$$\begin{aligned} E_k \left[\|X^{(k+1)} - A^\dagger CB^\dagger\|_F^2 \right] &\leq \|X^{(k)} - A^\dagger CB^\dagger\|_F^2 - E_k \left[\frac{2\alpha - \alpha^2 \|B\|_2^2}{\|A_{i,:}\|_2^2} \|A_{i,:} (X^{(k)} - A^\dagger CB^\dagger) B\|_2^2 \right] \\ &= \|X^{(k)} - A^\dagger CB^\dagger\|_F^2 - \frac{2\alpha - \alpha^2 \|B\|_2^2}{\|A\|_F^2} \|A (X^{(k)} - A^\dagger CB^\dagger) B\|_F^2. \end{aligned}$$

By $X^{(0)} \in \mathcal{M}$ and $A^\dagger CB^\dagger \in \mathcal{M}$, we have $X^{(0)} - A^\dagger CB^\dagger \in \mathcal{M}$. Noting $A_{i,:}^T = A^T I_{i,:}$, it is easy to show that $X^{(k)} - A^\dagger CB^\dagger \in \mathcal{M}$ by induction. Then by Lemma 1 and $0 < \alpha < \frac{2}{\|B\|_2^2}$, we can get

$$\begin{aligned} E_k \left[\|X^{(k+1)} - A^\dagger CB^\dagger\|_F^2 \right] &\leq \|X^{(k)} - A^\dagger CB^\dagger\|_F^2 - \frac{2\alpha - \alpha^2 \|B\|_2^2}{\|A\|_F^2} \sigma_{\min}^2(A) \sigma_{\min}^2(B) \|X^{(k)} - A^\dagger CB^\dagger\|_F^2 \\ &= \left(1 - \frac{2\alpha - \alpha^2 \|B\|_2^2}{\|A\|_F^2} \sigma_{\min}^2(A) \sigma_{\min}^2(B) \right) \|X^{(k)} - A^\dagger CB^\dagger\|_F^2. \end{aligned} \quad (12)$$

Finally, by ((12)) and induction on the iteration index k , we obtain the estimate ((11)). \square

Remark 3. By the similar approach as used in the proof of Theorem 1, we can prove that the iterate $X^{(k)}$ generated by ME-PRBK (5) satisfies the following estimate

$$E \left[\|X^{(k)} - A^\dagger CB^\dagger\|_F^2 \right] \leq \hat{\rho}^k \|X^{(0)} - A^\dagger CB^\dagger\|_F^2,$$

where $\hat{\rho} = 1 - \frac{\sigma_{\min}^2(A) \sigma_{\min}^2(B)}{\|A\|_F^2 \|B\|_2^2}$. The convergence factor of GRK in [14] is $\rho_{GRK} = 1 - \frac{\sigma_{\min}^2(A) \sigma_{\min}^2(B)}{\|A\|_F^2 \|B\|_2^2}$. It is obvious that

$$\rho_{GRK} > \min_{0 < \alpha < \frac{2}{\|B\|_2^2}} \rho = 1 - \frac{\sigma_{\min}^2(A) \sigma_{\min}^2(B)}{\|A\|_F^2 \|B\|_2^2} = \hat{\rho}$$

and $\rho < \rho_{GRK}$ when $\frac{1 - \sqrt{1 - \frac{\|B\|_2^2}{\|B\|_F^2}}}{\|B\|_2^2} < \alpha < \frac{1 + \sqrt{1 - \frac{\|B\|_2^2}{\|B\|_F^2}}}{\|B\|_2^2}$. This means that the convergence factor of ME-PRBK is the smallest and the factor of ME-RBK can be smaller than that of GRK when α is properly selected.

3. The Randomized Extended Block Kaczmarz Method for Inconsistent Equation

In [11,15,16], the authors proved that the Kaczmarz method does not converge to the least squares solution of $AX = b$ when $AX = b$ is inconsistent. Analogically, if the matrix Eq. (1) is inconsistent, the above ME-PRBK method dose not converge to $A^\dagger CB^\dagger$. The following theorem give the error bound of the for the inconsistent matrix equation.

Theorem 2. Assume that the consistent equation $AXB = C$ has a solution $X^* = A^\dagger CB^\dagger$. Let $\hat{X}^{(k)}$ denote the k th iterate of the ME-PRBK method applied to the inconsistent $AXB = C + W$ for any $W \in \mathbb{R}^{m \times n}$ starting from the initial matrix $X^{(0)} \in \mathbb{R}^{p \times q}$ in which $X_{:,j}^{(0)} \in R(A^T), j = 1, 2, \dots, q$ and $(X_{i,:}^{(0)})^T \in R(B), i = 1, 2, \dots, p$. In exact arithmetic, it follows that

$$E \left[\|\hat{X}^{(k)} - A^\dagger CB^\dagger\|_F^2 \right] \leq \rho^k \|X^{(0)} - A^\dagger CB^\dagger\|_F^2 + \frac{1 - \rho^k}{1 - \rho} \frac{\|WB^\dagger\|_F^2}{\|A\|_F^2}, \quad (13)$$

Proof. Set $H_i = \{X | A_i X B = C_i\}$, $\hat{H}_i = \{X | A_i X B = C_i + W_i\}$. Let Y denote the iterate of the PRBK method applied to the consistent $AXB = C$ at k th step, that is

$$Y = \hat{X}^{(k)} + \frac{A_{i,:}^T}{\|A_{i,:}\|_2^2} (C_{i,:} - A_{i,:} \hat{X}^{(k)} B) B^\dagger.$$

It follows from

$$\begin{aligned}\langle Y - A^\dagger C B^\dagger, \hat{X}^{(k+1)} - A^\dagger C B^\dagger \rangle_F &= \langle Y - A^\dagger C B^\dagger, \frac{A_{i,:}^T}{\|A_{i,:}\|_2^2} W_i B^\dagger \rangle_F \\ &= \text{trace} \left((B^\dagger)^T W_i^T \frac{A_{i,:}}{\|A_{i,:}\|_2^2} (Y - A^\dagger C B^\dagger) \right) \\ &= \text{trace} \left(W_i^T \frac{A_{i,:} Y B - A_{i,:} A^\dagger C B^\dagger B}{\|A_{i,:}\|_2^2} (B^T B)^\dagger \right) = 0\end{aligned}$$

and

$$\|\hat{X}^{(k+1)} - Y\|_F^2 = \frac{1}{\|A_{i,:}\|_2^4} \|A_{i,:}^T W_i B^\dagger\|_F^2 = \text{trace} \left((B^\dagger)^T W_i^T A_{i,:} A_{i,:}^T W_i B^\dagger \frac{1}{\|A_{i,:}\|_2^2} \right) = \frac{\|W_i B^\dagger\|_2^2}{\|A_{i,:}\|_2^2}$$

that

$$\begin{aligned}\|\hat{X}^{(k+1)} - A^\dagger C B^\dagger\|_F^2 &= \|Y - A^\dagger C B^\dagger\|_F^2 + \|\hat{X}^{(k+1)} - Y\|_F^2 \\ &= \|Y - A^\dagger C B^\dagger\|_F^2 + \frac{\|W_i B^\dagger\|_2^2}{\|A_{i,:}\|_2^2}.\end{aligned}\tag{14}$$

By taking the conditional expectation on both sides of (14), we can obtain

$$\begin{aligned}E_k \left[\|\hat{X}^{(k+1)} - A^\dagger C B^\dagger\|_F^2 \right] &= E_k \left[\|Y - A^\dagger C B^\dagger\|_F^2 \right] + E_k \left[\frac{\|W_i B^\dagger\|_2^2}{\|A_{i,:}\|_2^2} \right] \\ &\leq \left(1 - \frac{\sigma_{\min}^2(A) \sigma_{\min}^2(B)}{\|A\|_F^2 \|B\|_2^2} \right) \|\hat{X}^{(k)} - A^\dagger C B^\dagger\|_F^2 + \frac{\|W B^\dagger\|_F^2}{\|A\|_F^2}\end{aligned}$$

The inequality is obtained by Remark 3. Applying this recursive relation iteratively, we have

$$\begin{aligned}E \left[\|\hat{X}^{(k+1)} - A^\dagger C B^\dagger\|_F^2 \right] &\leq \rho E \left[\|\hat{X}^{(k)} - A^\dagger C B^\dagger\|_F^2 \right] + \frac{\|W B^\dagger\|_F^2}{\|A\|_F^2} \\ &\leq \rho^2 E \left[\|\hat{X}^{(k-1)} - A^\dagger C B^\dagger\|_F^2 \right] + (\rho + 1) \frac{\|W B^\dagger\|_F^2}{\|A\|_F^2} \\ &\leq \dots \\ &\leq \rho^{k+1} \|\hat{X}^{(0)} - A^\dagger C B^\dagger\|_F^2 + (\rho^k + \dots + \rho + 1) \frac{\|W B^\dagger\|_F^2}{\|A\|_F^2} \\ &= \rho^{k+1} \|X^{(0)} - A^\dagger C B^\dagger\|_F^2 + \frac{1 - \rho^{k+1}}{1 - \rho} \frac{\|W B^\dagger\|_F^2}{\|A\|_F^2}.\end{aligned}$$

This completes the proof. \square

Next, we use the idea of randomized extended Kaczmarz method (see [16–18] for details) for solving the least squares solution of the inconsistent Eq. (1). At each iteration, $Z^{(k)}$ is the k th iterate of

ME-RBK applied to $A^T Z B^T = 0$ with the initial guess $Z^{(0)}$, and $X^{(k)}$ is one-step ME-RBK update for $AXB = C - Z^{(k)}$. we can get the following randomized extended block Kaczmarz iteration

$$\begin{cases} Z^{(k+1)} = Z^{(k)} - \frac{\alpha}{\|A_{:,j}\|_2^2} A_{:,j} ((A_{:,j}^T Z^{(k)}) B^T) B, \\ X^{(k+1)} = X^{(k)} + \frac{\alpha}{\|A_{i,:}\|_2^2} A_{i,:}^T \left((C_{i,:} - Z_{i,:}^{(k+1)} - (A_{i,:} X^{(k)}) B) B^T \right), \end{cases} \quad (15)$$

where $\alpha > 0$ is the step size, and i and j are selected with probability $p_i = \frac{\|A_{i,:}\|_2^2}{\|A\|_F^2}$ and $\hat{p}_j(A) = \frac{\|A_{:,j}\|_2^2}{\|A\|_F^2}$, respectively. The cost of each iteration of this method is $4n(q+m) + m + 1 - 2n - q$ for updating $Z^{(k)}$ and $(4q+1)(n+p) + 1 - 2q$ for updating $X^{(k)}$ if the square of the row norm and the column norm of A have been calculated in advance. We describe this method as Algorithm 2, which is called the ME-REBK algorithm.

Algorithm 2 Randomized Extended Block Kaczmarz Method for $AXB = C$ (ME-REBK)

Input: $A \in R^{m \times p}$, $B \in R^{q \times n}$, $C \in R^{m \times n}$, $X^{(0)} = 0 \in R^{p \times q}$, $Z^{(0)} = C$
1: **for** $k = 0, 1, 2, \dots$, **do**
2: Pick j with probability $\hat{p}_j(A) = \frac{\|A_{:,j}\|_2^2}{\|A\|_F^2}$
3: Compute $Z^{(k+1)} = Z^{(k)} - \frac{\alpha}{\|A_{:,j}\|_2^2} A_{:,j} ((A_{:,j}^T Z^{(k)}) B^T) B$
4: Pick i with probability $p_i(A) = \frac{\|A_{i,:}\|_2^2}{\|A\|_F^2}$
5: Compute $X^{(k+1)} = X^{(k)} + \frac{\alpha}{\|A_{i,:}\|_2^2} A_{i,:}^T \left((C_{i,:} - Z_{i,:}^{(k+1)} - (A_{i,:} X^{(k)}) B) B^T \right)$
6: **end for**

Theorem 3. Assume $0 < \alpha < \frac{2}{\|B\|_2^2}$. Let $\{Z^{(k)}\}$ denote the k th iterate of ME-RBK applied to $A^T Z B^T = 0$ starting from the initial matrix $Z^{(0)} \in R^{m \times n}$ in which $Z_{:,j}^{(0)} \in C_{:,j} + R(A)$, $j = 1, 2, \dots, n$ and $(Z_{i,:}^{(0)})^T \in (C_{i,:})^T + R(B^T)$, $i = 1, 2, \dots, m$. Then $\{Z^{(k)}\}$ converges linearly to $C - AA^+CB^+B$ in mean square form, and the solution error in expectation for the iteration sequence $X^{(k)}$ obeys

$$E \left[\left\| Z^{(k)} - (C - AA^+CB^+B) \right\|_F^2 \right] \leq \rho^k \left\| Z^{(0)} - (C - AA^+CB^+B) \right\|_F^2, \quad (16)$$

where the j th column of A is selected with probability $\hat{p}_j(A) = \frac{\|A_{:,j}\|_2^2}{\|A\|_F^2}$.

Proof. The approach of proof is the same as that used in the proof of Theorem 1, so we omit it. \square

Theorem 4. Assume $0 < \alpha < \frac{2}{\|B\|_2^2}$. The sequence $\{X^{(k)}\}$ is generated by the ME-REBK method for $AXB = C$ starting from the initial matrix $X^{(0)} \in R^{p \times n}$ and $Z^{(0)} \in R^{m \times n}$, where $X_{:,j}^{(0)} \in R(A^T)$, $j = 1, 2, \dots, q$, $(X_{i,:}^{(0)})^T \in R(B)$, $i = 1, 2, \dots, p$, $Z_{:,j}^{(0)} \in C_{:,j} + R(A)$, $j = 1, 2, \dots, n$ and $(Z_{i,:}^{(0)})^T \in (C_{i,:})^T + R(B^T)$, $i = 1, 2, \dots, m$. For any $\varepsilon > 0$, it holds

$$\begin{aligned} E \left[\left\| X^{(k)} - A^+CB^+ \right\|_F^2 \right] &\leq \frac{(1+\varepsilon)^{k+1} - (1+\varepsilon)}{\varepsilon^2} \frac{\alpha^2 \|B\|_2^2 \rho^k}{\|A\|_F^2} \left\| Z^{(0)} - (C - AA^+CB^+B) \right\|_F^2 \\ &\quad + (1+\varepsilon)^k \rho^k \left\| X^{(0)} - A^+CB^+ \right\|_F^2, \end{aligned} \quad (17)$$

where $i \in [m]$, $j \in [p]$ are picked with probability $p_i(A) = \frac{\|A_{i,:}\|_2^2}{\|A\|_F^2}$ and $\hat{p}_j(A) = \frac{\|A_{:,j}\|_2^2}{\|A\|_F^2}$, respectively.

Proof. Let $X^{(k)}$ denote the k th iterate of ME-REBK method for $AXB = C$, and $\tilde{X}^{(k+1)}$ be the one-step Kaczmarz update for the matrix equation $AXB = AA^\dagger CB^\dagger B$ from $X^{(k)}$, i.e.,

$$\tilde{X}^{(k+1)} = X^{(k)} + \frac{\alpha}{\|A_{i,:}\|_2^2} A_{i,:}^T (A_{i,:} A^\dagger CB^\dagger B - A_{i,:} X^{(k)} B) B^T.$$

We have

$$\tilde{X}^{(k+1)} - A^\dagger CB^\dagger = X^{(k)} - A^\dagger CB^\dagger - \frac{\alpha}{\|A_{i,:}\|_2^2} A_{i,:}^T A_{i,:} (X^{(k)} - A^\dagger CB^\dagger) B B^T$$

and

$$X^{(k+1)} - \tilde{X}^{(k+1)} = \frac{\alpha}{\|A_{i,:}\|_2^2} A_{i,:}^T (C_{i,:} - Z_{i,:}^{(k+1)} - A_{i,:} A^\dagger CB^\dagger B) B^T.$$

For any $\varepsilon > 0$, by triangle inequality and Young's inequality, we can get

$$\begin{aligned} \|X^{(k+1)} - A^\dagger CB^\dagger\|_F^2 &= \|(X^{(k+1)} - \tilde{X}^{(k+1)}) + (\tilde{X}^{(k+1)} - A^\dagger CB^\dagger)\|_F^2 \\ &\leq (\|X^{(k+1)} - \tilde{X}^{(k+1)}\|_F + \|\tilde{X}^{(k+1)} - A^\dagger CB^\dagger\|_F)^2 \\ &\leq \|X^{(k+1)} - \tilde{X}^{(k+1)}\|_F^2 + \|\tilde{X}^{(k+1)} - A^\dagger CB^\dagger\|_F^2 \\ &\quad + 2\|X^{(k+1)} - \tilde{X}^{(k+1)}\|_F \|\tilde{X}^{(k+1)} - A^\dagger CB^\dagger\|_F \\ &\leq (1 + \frac{1}{\varepsilon}) \|X^{(k+1)} - \tilde{X}^{(k+1)}\|_F^2 + (1 + \varepsilon) \|\tilde{X}^{(k+1)} - A^\dagger CB^\dagger\|_F^2. \end{aligned} \quad (18)$$

By taking the conditional expectation on the both sides of (18), we have

$$E_k \left[\|X^{(k+1)} - A^\dagger CB^\dagger\|_F^2 \right] \leq (1 + \frac{1}{\varepsilon}) E_k \left[\|X^{(k+1)} - \tilde{X}^{(k+1)}\|_F^2 \right] + (1 + \varepsilon) E_k \left[\|\tilde{X}^{(k+1)} - A^\dagger CB^\dagger\|_F^2 \right]. \quad (19)$$

It follows from

$$\begin{aligned} \|X^{(k+1)} - \tilde{X}^{(k+1)}\|_F^2 &= \left\| \frac{\alpha}{\|A_{i,:}\|_2^2} A_{i,:}^T (C_{i,:} - Z_{i,:}^{(k+1)} - A_{i,:} A^\dagger CB^\dagger B) B^T \right\|_F^2 \\ &= \frac{\alpha^2}{\|A_{i,:}\|_2^2} \text{trace} \left(B (C_{i,:} - Z_{i,:}^{(k+1)} - A_{i,:} A^\dagger CB^\dagger B)^T (C_{i,:} - Z_{i,:}^{(k+1)} - A_{i,:} A^\dagger CB^\dagger B) B^T \right) \\ &\leq \frac{\alpha^2 \|B\|_2^2}{\|A_{i,:}\|_2^2} \|C_{i,:} - Z_{i,:}^{(k+1)} - A_{i,:} A^\dagger CB^\dagger B\|_2^2 \end{aligned}$$

that

$$\begin{aligned} E_k \left[\|X^{(k+1)} - \tilde{X}^{(k+1)}\|_F^2 \right] &\leq \alpha^2 \|B\|_2^2 E_k^j E_k^i \left[\frac{\|C_{i,:} - Z_{i,:}^{(k+1)} - A_{i,:} A^\dagger CB^\dagger B\|_2^2}{\|A_{i,:}\|_2^2} \right] \\ &= \alpha^2 \|B\|_2^2 E_k^j \left[\frac{1}{\|A\|_F^2} \sum_{i=1}^m \|C_{i,:} - Z_{i,:}^{(k+1)} - A_{i,:} A^\dagger CB^\dagger B\|_2^2 \right] \\ &= \frac{\alpha^2 \|B\|_2^2}{\|A\|_F^2} E_k \left[\|Z^{(k+1)} - (C - AA^\dagger CB^\dagger B)\|_F^2 \right]. \end{aligned}$$

It yields

$$\begin{aligned} E \left[\left\| X^{(k+1)} - \tilde{X}^{(k+1)} \right\|_F^2 \right] &\leq \frac{\alpha^2 \|B\|_2^2}{\|A\|_F^2} E \left[\left\| Z^{(k+1)} - (C - AA^\dagger CB^\dagger B) \right\|_F^2 \right] \\ &\leq \frac{\alpha^2 \|B\|_2^2}{\|A\|_F^2} \rho^{k+1} \left\| Z^{(0)} - (C - AA^\dagger CB^\dagger B) \right\|_F^2 \quad (\text{by Lemma 3}). \end{aligned} \quad (20)$$

By $X^{(0)} - A^\dagger CB^\dagger \in \mathcal{M}$, we have $X^{(k)} - A^\dagger CB^\dagger \in \mathcal{M}$. Then, by using Theorem 1, we can get

$$\begin{aligned} E_k [\left\| \tilde{X}^{(k+1)} - A^\dagger CB^\dagger \right\|_F^2] &= E_k \left[\left\| X^{(k)} - A^\dagger CB^\dagger - \frac{\alpha}{\|A_{i,:}\|_2^2} A_{i,:}^T A_{i,:} (X^{(k)} - A^\dagger CB^\dagger) B B^T \right\|_F^2 \right] \\ &\leq \rho \left\| X^{(k)} - A^\dagger CB^\dagger \right\|_F^2, \end{aligned}$$

then

$$E \left[\left\| \tilde{X}^{(k+1)} - A^\dagger CB^\dagger \right\|_F^2 \right] \leq \rho E \left[\left\| X^{(k)} - A^\dagger CB^\dagger \right\|_F^2 \right]. \quad (21)$$

Combining (19), (20) and (21) yields

$$\begin{aligned} E \left[\left\| X^{(k+1)} - A^\dagger CB^\dagger \right\|_F^2 \right] &\leq (1 + \frac{1}{\varepsilon}) E \left[\left\| X^{(k+1)} - \tilde{X}^{(k+1)} \right\|_F^2 \right] + (1 + \varepsilon) E \left[\left\| \tilde{X}^{(k+1)} - A^\dagger CB^\dagger \right\|_F^2 \right] \\ &\leq (1 + \frac{1}{\varepsilon}) \frac{\alpha^2 \|B\|_2^2 \rho^{k+1}}{\|A\|_F^2} \left\| Z^{(0)} - (C - AA^\dagger CB^\dagger B) \right\|_F^2 \\ &\quad + (1 + \varepsilon) \rho E \left[\left\| X^{(k)} - A^\dagger CB^\dagger \right\|_F^2 \right] \\ &\leq (1 + \frac{1}{\varepsilon}) \frac{\alpha^2 \|B\|_2^2 \rho^{k+1}}{\|A\|_F^2} \left\| Z^{(0)} - (C - AA^\dagger CB^\dagger B) \right\|_F^2 [1 + (1 + \varepsilon)] \\ &\quad + (1 + \varepsilon)^2 \rho^2 E \left[\left\| X^{(k-1)} - A^\dagger CB^\dagger \right\|_F^2 \right] \\ &\leq \dots \\ &\leq (1 + \frac{1}{\varepsilon}) \frac{\alpha^2 \|B\|_2^2 \rho^{k+1}}{\|A\|_F^2} \left\| Z^{(0)} - (C - AA^\dagger CB^\dagger B) \right\|_F^2 \sum_{i=0}^k (1 + \varepsilon)^i \\ &\quad + (1 + \varepsilon)^{k+1} \rho^{k+1} \left\| X^{(0)} - A^\dagger CB^\dagger \right\|_F^2 \\ &= \frac{(1 + \varepsilon)^{k+2} - (1 + \varepsilon)}{\varepsilon^2} \frac{\alpha^2 \|B\|_2^2 \rho^{k+1}}{\|A\|_F^2} \left\| Z^{(0)} - (C - AA^\dagger CB^\dagger B) \right\|_F^2 \\ &\quad + (1 + \varepsilon)^{k+1} \rho^{k+1} \left\| X^{(0)} - A^\dagger CB^\dagger \right\|_F^2. \end{aligned}$$

This completes the proof. \square

Remark 4. Replacing B^T in (15) with B^\dagger , we obtain the following projection-based randomized extended block Kaczmarz method (ME-PREBK) iteration

$$\begin{cases} Z^{(k+1)} = Z^{(k)} - \frac{\alpha}{\|A_{:,j}\|_2^2} A_{:,j} ((A_{:,j}^T Z^{(k)}) B^T) (B^\dagger)^T, \\ X^{(k+1)} = X^{(k)} + \frac{\alpha}{\|A_{i,:}\|_2^2} A_{i,:}^T \left((C_{i,:} - Z_{i,:}^{(k+1)}) - (A_{i,:} X^{(k)}) B \right) B^\dagger, \end{cases} \quad (22)$$

4. Numerical Experiments

In this section, we will present some experiment results of the proposed algorithms for solving various matrix equations, and compare them with ME-RGRK and ME-MWRK in [9] for consistent matrix equations and RBCD in [8] for inconsistent matrix equations. All experiments are carried out by using MATLAB (version R2020a) on a DESKTOP-8CBRR86 with Intel(R) Core(TM) i7-4712MQ CPU @2.30GHz 2.29GHz, RAM 8GB and Windows 10.

All computations are started from the initial guess $X^{(0)} = 0$, the step size $\alpha = \frac{1.5}{\|B\|_2^2}$, and terminated once the relative error (RE) of the solution, defined by

$$RE = \frac{\|X^{(k)} - X^*\|_F^2}{\|X^*\|_F^2}$$

at the the current iterate $X^{(k)}$, satisfies $RE < 10^{-6}$ or exceeds maximum iteration $K = 50000$, where $X^* = A^\dagger CB^\dagger$. We report the average number of iterations (denoted as "IT") and the average computing time in seconds (denoted as "CPU") for 20 trials repeated runs of the corresponding method. Three examples are tested, and A and B are generated as follows.

- Type I: For given m, p, q, n , the entries of A and B are generated from a standard normal distribution, i.e., $A = randn(m, p)$, $B = randn(q, n)$.
- Type II: Like [14], for given m, p , and $r_1 = rank(A)$, we construct a matrix A by $A = U_1 D_1 V_1^T$, where $U_1 \in R^{m \times r_1}$ and $V_1 \in R^{p \times r_1}$ are orthogonal columns matrices, $D_1 \in R^{r_1 \times r_1}$ is a diagonal matrix whose first $r - 2$ diagonal entries are uniformly distributed numbers in $[\sigma_{\min}(A), \sigma_{\max}(A)]$, and the last two diagonal entries are $\sigma_{\max}(A), \sigma_{\min}(A)$. The entries of B are generated by similar method with parameters $q, n, r_2 = rank(B)$.
- Type III: The real-world sparse data come from the Florida sparse matrix collection [19].

4.1. Consistent Matrix Equation

Given A, B , we set $C = AX^*B$ with $X^* = randn(p, q)$ to construct a consistent matrix equation. In Tables 1–3, we report the average IT and CPU of the ME-RGRK, ME-MWRK, ME-RBK and ME-PRBK methods for solving consistent matrix. In the following tables, the item '>' represents that the number of iteration steps exceeds the maximum iteration (50000), and the item '-' represents that the method does not converge.

From these tables, we can see that the ME-RBK and ME-PRBK methods vastly outperform the ME-RGRK and ME-MWRK methods in terms of both IT and CPU times regardless of whether the matrices A and B are full column/row rank or not. As the increasing of matrix dimension, the CPU time of ME-RBK and ME-PRBK methods is increasing slowly, while the running time of ME-RGRK and ME-MWRK increases dramatically.

In addition, when the matrix size is small, the ME-PRBK method is competitive, because of the pseudoinverse is less expensive and the number of iteration steps is small. When the matrix size is large, the matrix is large, the ME-RBK method is more challenging because it does not need to calculate the pseudoinverse (see the last line in Table 1).

Table 1. IT and CPU of ME-RGRK, ME-MWRK, ME-RBK and ME-PRBK for the consistent matrix equations with Type I.

<i>m</i>	<i>p</i>	<i>q</i>	<i>n</i>		ME-RGRK	ME-MWRK	ME-RBK	ME-PRBK
100	40	40	100	IT	49707	27579	8609.9	1152.8
				CPU	0.78	2.01	0.21	0.04
40	100	100	40	IT	>	49332.7	6449.2	1507.2
				CPU	>	2.15	0.34	0.08
500	100	100	500	IT	>	32109	4049.4	1866.1
				CPU	>	57.61	0.53	0.19
1000	200	300	2000	IT	>	>	39941.2	4450.4
				CPU	>	>	7.01	0.72

Table 2. IT and CPU of ME-RGRK, ME-MWRK, ME-RBK and ME-PRBK for the consistent matrix equations with Type II.

<i>m</i>	<i>p</i>	<i>r</i> ₁	$[\sigma_{\min}(A), \sigma_{\max}(A)]$	<i>q</i>	<i>n</i>	<i>r</i> ₂	$[\sigma_{\min}(B), \sigma_{\max}(B)]$		ME-RGRK	ME-MWRK	ME-RBK	ME-PRBK
100	40	20	[1,2]	40	100	40	[1,2]	IT	4361.0	1987.1	503.1	332.4
								CPU	0.06	0.17	0.01	0.008
100	40	20	[1,5]	40	100	20	[1,5]	IT	22423.2	6439.8	9307.6	1056.3
								CPU	0.33	0.57	0.20	0.02
1000	200	100	[1,2]	100	1000	50	[1,2]	IT	20055.5	7047.8	2587.4	1674.3
								CPU	78.45	70.56	0.42	0.23
1000	100	50	[1,5]	200	1000	200	[1,5]	IT	>	>	28024.4	2833.6
								CPU	>	>	4.72	0.42

Table 3. IT and CPU of ME-RGRK, ME-MWRK, ME-RBK and ME-PRBK for the consistent matrix equations with Type III.

<i>A</i>	<i>B</i>		ME-RGRK	ME-MWRK	ME-RBK	ME-PRBK
divorce	ash219 ^T	IT	43927.8	14164.4	10993.4	3873.5
		CPU	1.15	1.35	0.36	0.13
divorce	ash219	IT	40198.7	17251.4	11557.4	3124.7
		CPU	0.63	0.80	0.43	0.11
ash219	ash958 ^T	IT	>	>	6042.3	2267.0
		CPU	>	>	1.04	0.36
ash219	ash958	IT	>	>	5745.4	2114.2
		CPU	>	>	1.22	0.42

4.2. Inconsistent Matrix Equation

To construct an inconsistent matrix equation, we set $C = AX^*B + R$, where X^* and R are random matrices which are generated by $X^* = randn(p, q)$ and $R = \delta * randn(p, q), \delta \in (0, 1)$. Numerical results of the RBCD, IME-REBK and IME-PREBK methods are listed in Table 4–6. From these tables, we can see that the IME-PREBK method is better than the RBCD method in terms of IT and CPU time, especially when the $\frac{\sigma_{\max}}{\sigma_{\min}}$ is large (see the last line in Table 5). The IME-REBK method is not competitive for B with full row rank because it needs to solve two equations. However, when B has not full row rank, the RBCD method does not converge, while the IME-REBK and IME-PREBK methods do.

Table 4. IT and CPU of RBCD, IME-REBK and IME-PREBK for the inconsistent matrix equations with Type I.

<i>m</i>	<i>p</i>	<i>q</i>	<i>n</i>		RBCD	IME-REBK	IME-PREBK
100	40	40	100	IT CPU	17212.2 0.69	21270.5 1.57	2469.4 0.19
40	100	100	40	IT CPU	- -	24708.3 2.38	2174.6 0.21
500	100	100	500	IT CPU	6059.1 4.97	7352.8 7.28	2512.3 2.70
1000	200	300	2000	IT CPU	14209.2 152.32	12490.4 246.56	5183.5 99.68

Table 5. IT and CPU of RBCD, IME-REBK and IME-PREBK for the inconsistent matrix equations with Type II.

<i>m</i>	<i>p</i>	<i>r</i> ₁	$[\sigma_{\min}(A), \sigma_{\max}(A)]$	<i>q</i>	<i>n</i>	<i>r</i> ₂	$[\sigma_{\min}(B), \sigma_{\max}(B)]$		RBCD	IME-REBK	IME-PREBK
100	40	20	[1,2]	40	100	40	[1,2]	IT CPU	1035.9 0.04	762.2 0.05	384.5 0.03
100	40	20	[1,5]	40	100	20	[1,5]	IT CPU	- -	16224.7 1.23	1507.8 0.12
1000	200	100	[1,2]	100	1000	50	[1,2]	IT CPU	- -	4067.5 24.23	2217.8 13.67
1000	200	100	[1,5]	100	1000	100	[1,5]	IT CPU	> >	48269.6 365.78	4328.0 30.25

Table 6. IT and CPU of RBCD, IME-REBK and IME-PREBK for the inconsistent matrix equations with Type III.

<i>A</i>	<i>B</i>		RBCD	IME-REBK	IME-PREBK
divorce	ash219 ^T	IT CPU	> >	20026.3 2.26	4308.4 0.48
divorce	ash219	IT CPU	- -	19199.1 1.49	4026.2 0.31
ash219	ash958 ^T	IT CPU	22313.4 15.69	10823.5 19.03	2561.8 5.89
ash219	ash958	IT CPU	- -	10020.7 15.83	2363.5 4.72

5. Conclusions

In this paper we have proposed a randomized block Kaczmarz algorithm for solving the consistent matrix equation and its extended version for the inconsistent case. Theoretically, we have proved the proposed algorithms converge linearly to the unique minimal *F*-norm solution or least-squares solution (i.e., A^+CB^+) without requirements on *A* and *B* have full column/row rank. Numerical results show the effectiveness of the algorithms. Since the proposed algorithms only require one row or one column of *A* at each iteration without matrix-matrix product, it is suitable for the scenarios where the matrix *A* is too large to fit in memory or matrix multiplication is considerably expensive.

Author Contributions: Conceptualization, L.X.; Methodology, L.X. and W.B.; Validation, L.X. and W.L.; Writing—original draft preparation, L.X.; Writing—review and editing, L.X., W.B and W.L.; Software, L.X.; Visualization, L.X. and W.L. All authors have read and agreed to the published version of the manuscript

Funding: This research received no external funding.

Institutional Review Board Statement: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability Statement: The datasets that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: The authors are thankful to the referees for their constructive comments and valuable suggestions, which have greatly improved the original manuscript of this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fausett, D.W.; Fulton, C.T. Large least squares problems involving Kronecker products. *SIAM J. Matrix Anal. Appl.* **1994**, *15*, 219–227.
2. Zha, H.Y. Comments on large least squares problems involving Kronecker products. *SIAM J. Matrix Anal. Appl.* **1995**, *16*, 1172–1172.
3. Peng, Z.Y. An iterative method for the least squares symmetric solution of the linear matrix equation $AXB = C$. *Appl. Math. Comput.* **2005**, *170*, 711–723.
4. Ding, F.; Liu, P.X.; Ding, J. Iterative solutions of the generalized Sylvester matrix equations by using the hierarchical identification principle. *Appl. Math. Comput.*, **2008**, *197*, 41–50.
5. Huang, G.X.; Yin, F.; Guo, K. An iterative method for the skew-symmetric solution and the optimal approximate solution of the matrix equation $AXB = C$. *J. Comput. Appl. Math.*, **2008**, *212*, 231–244.
6. Wang, X.; Li, Y.; Dai, L. On hermitian and skew-hermitian splitting iteration methods for the linear matrix equation $AXB = C$. *Comput. Math. Appl.*, **2013**, *65*, 657–664.
7. Shafiei, S.G.; Hajarian, M. Developing Kaczmarz method for solving Sylvester matrix equations. *J. Franklin Inst.*, **2022**, *359*, 8991–9005.
8. Du, K.; Ruan, C.C.; Sun, X.H. On the convergence of a randomized block coordinate descent algorithm for a matrix least squares problem. *Appl. Math. Lett.*, **2022**, *124*, 107689.
9. Wu, N.C.; Liu, C.Z.; Zuo, Q. On the Kaczmarz methods based on relaxed greedy selection for solving matrix equation $AXB=C$. *J. Comput. Appl. Math.*, **2022**, *413*, 114374.
10. Strohmer, T.; Vershynin, R. A randomized Kaczmarz algorithm with exponential convergence. *J. Fourier Anal. Appl.* **2009**, *15*, 262–278.
11. Zouzias, A.; Freris, N.M. Randomized extended Kaczmarz for solving least squares. *SIAM J. Matrix Anal. Appl.*, **2013**, *34*, 773–793.
12. Ion, N. Faster randomized block kaczmarz algorithms. *SIAM J. Matrix Anal. Appl.*, **2019**, *40*, 1425–1452.
13. Nemirovski, A.; Juditsky, A.; Lan, G.; Shapiro, A. Robust stochastic approximation approach to stochastic programming. *SIAM J. Optimiz.* **2009**, *19*, 1574–1609.
14. Niu, Y.Q.; Zheng, B. On global randomized block Kaczmarz algorithm for solving large-scale matrix equations. *arXiv*, **2022**, arxiv-2204.13920.
15. Needell, D. Randomized Kaczmarz solver for noisy linear systems. *BIT Numer Math.* **2010**, *50*, 395–403.
16. Ma, A.; Needell, D. A. Ramdas, Convergence properties of the randomized extended Gauss-Seidel and Kaczmarz methods. *SIAM J. Matrix Anal. Appl.*, **2015**, *36*, 1590–1604.
17. Du, K. Tight upper bounds for the convergence of the randomized extended Kaczmarz and Gauss-Seidel algorithms. *Numer. Linear Algebra Appl.*, **2019**, *26*, e2233.
18. Du, K.; Si, W.T.; Sun, X.H. Randomized extended average block kaczmarz for solving least squares. *SIAM J. Sci. Comput.*, **2020**, *42*, A3541–A3559.
19. Davis, T.A.; Hu, Y. The university of Florida sparse matrix collection. *Math. Softw.*, **2011**, *38*, 1–25.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.