On the Extended Deterministic Row-Action Method with Maximum Weighted Residual and Its Convergence Analysis

Ya-Nan Zhao, Nian-Ci Wu, and Chengzhi Liu

***Abstract*—The row-action method, owing to its straightforward form, has been widely employed in diverse fields such as computer graphics. In this paper, we first propose a novel deterministic row-action approach for computing the orthogonal decomposition of linear systems using a maximum weighted residual selection strategy. Then, we extend it to solve a given overdetermined and inconsistent linear system. We prove that the proposed methods can linearly converge to the least-squares solution with a minimum Euclidean norm. Several numerical studies are presented to corroborate our theoretical findings and demonstrate that our methods achieve faster convergence than existing deterministic extended row-action methods.**

***Index Terms*—indicator selection, row-action, orthogonal decomposition, extended method.**

I. Introduction

LARGE-SCALE linear systems frequently arise in scientific computing and engineering applications such as computerized tomography [1], machine learning [2], and signal processing [3]. This paper focuses on solving the linear system $Ax = b$ with $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, by computing the least-squares solution $x^* = \arg\min_{x\in\mathbb{R}^n} \left\{\|b - Ax\|^2\right\}$. The solution $x^* = A^\dagger b$ is unique and has minimum Euclidean norm, with $A^\dagger$ representing the Moore-Penrose pseudoinverse [4].

The least-squares solution of $Ax = b$ can be obtained by solving the consistent system $Ax = b_R = AA^\dagger b$, where b_R represents the orthogonal projection of b onto the range space of A ($\mathcal{R}(A)$) [5]. The row-action method, first introduced by Censor [6], is characterized by its low per-iteration cost and computational efficiency. For comprehensive reviews of related row-action techniques, see [7], [8]. Among these methods, the Kaczmarz iteration [9], [10] stands out due to its simplicity and clear geometric interpretation. Popa's Extended Kaczmarz (EK) approach addresses this least-squares problem through an alternating projection scheme; see [11, Algorithm (R)]. It first performs a Kaczmarz iteration for the system $Ax = b - z^{(k)}$, then updates $z^{(k)}$ to approximate $b_N = b - b_R$. The iterative scheme proceeds as follows.

$$\begin{cases} x^{(k+1)} & = x^{(k)} + \frac{b_i - z_i^{(k)} - a_i^T x^{(k)}}{\|a_i\|^2} a_i \\ & := \mathcal{P}_i(A, x^{(k)}, b - z^{(k)}) \\ z^{(k+1)} & = z^{(k)} - \frac{A_j^T z^{(k)}}{\|A_j\|^2} A_j \\ & := \mathcal{T}_j(A, z^{(k)}, 0) \end{cases} \tag{1}$$

for $k = 0, 1, 2, \cdots$, where a_i and A_j^T respectively denote the ith and jth rows of A and A^T, with indices $i \in [m]$ and $j \in [n]$ selected cyclically. Here, the symbol $[\ell]$ denotes the set of integers from 1 to ℓ.

Zouzias and Freris [12] enhanced the original framework by introducing randomized row selection, thereby developing the Randomized EK (REK) method with provable convergence guarantees. Further improvements emerged through various index selection strategies. Bai and Wu [13] proposed a partially REK method in which the index j_k employs a cyclic ordering. Their analysis demonstrated that the expected convergence rate upper bound could outperform the original REK method under certain conditions. This research direction has led to several REK variants based on different index selection strategies, such as the greedy REK method [14], the partially REK with residuals method [15, Algorithm 3.1], and the Maximum-Distance Extended Kaczmarz (MDEK) method [15, Algorithm 3.2]. Each of these approaches offers distinct advantages through their specialized selection mechanisms while maintaining the fundamental REK framework.

Since $z^{(k+1)}$ provides a better approximation to b_N than $z^{(k)}$, Du applied the RK iterate to the linear system $Ax = b - z^{(k+1)}$ in the second-half step and proposed a slightly modified REK method, termed REK-S; see [16, Algorithm 3]. The REK-S iteration consists of

$$\begin{cases} z^{(k+1)} = \mathcal{T}_j(A, z^{(k)}, 0), \\ x^{(k+1)} = \mathcal{P}_i(A, x^{(k)}, b - z^{(k+1)}), \end{cases} \tag{2}$$

where the indices i and j are chosen at random. This framework has inspired several enhanced variants. Gao and Chen introduced a hybrid approach, combining the deterministic cyclic selection for i with the residual-based random sampling for j; see [17, Method 3.1]. This method can be regarded as a modified form of the partially REK with residuals method in [15]. In contrast, Mustafa and Saha adopted a fully deterministic strategy, selecting i_k and j_k based on maximal residual criteria; see [18, Algorithm 1], and called it the Maximal Residual Extended Kaczmarz (MREK) method. These developments demonstrate that various index selection

Manuscript received June 6, 2025; revised August 7, 2025.

This work was supported in part by the National Natural Science Foundation of China under grant 12201651.

Y. Zhao is a graduate student of School of Mathematics and Statistics, South-Central Minzu University, Wuhan 430074, P. R. China (e-mail: yananzhao@mail.scuec.edu.cn).

N. Wu is an associate professor of School of Mathematics and Statistics, South-Central Minzu University, Wuhan 430074, P. R. China (e-mail: nianciwu@scuec.edu.cn).

C. Liu is an associate professor of School of Mathematics and Finance, Hunan University of Humanities, Science and Technology, Loudi 417000, P. R. China (corresponding author to provide e-mail: it-rocket@163.com).

strategies work effectively in the extended Kaczmarz framework, while preserving its core projection-based approach.

This paper presents a novel deterministic projection approach for computing the orthogonal decomposition of the right-hand side. The method employs a greedy selection strategy based on the maximum weighted residual, which we term the Maximum Weighted Residual Orthogonal Projection (MWROP) method. We provide rigorous convergence guarantees for MWROP and subsequently extend it to handle inconsistent linear systems, resulting in an enhanced variant called the Maximum Weighted Residual Extended Kaczmarz (MWREK) method. Through carefully analyzing the error recursion in single row-action methods, we derive an optimal index selection strategy for the MWREK method. Our convergence result demonstrates that MWREK achieves a faster convergence rate compared to the MREK method in [18]. Numerical experiments further validate this theoretical advantage, showing consistent performance improvements across various test cases.

We organize the remaining part of this paper as follows. The rest of this section introduces some notation. We present the MWROP method in Section II. In Section III, we extend the MWROP method to solve inconsistent linear systems, by introducing the MWREK method along with its theoretical analysis. In Section IV, we provide numerical experiments that support our theoretical analysis and demonstrate that, in comparison with some commonly used conventional least-squares solvers, our method achieves faster convergence. We end this work with some conclusions in Section V.

Throughout the paper, all vectors are assumed to be column vectors. The identity matrix of size n is denoted by I_n. For the coefficient matrix A, the smallest non-zero singular value of A is denoted by $\sigma_{A,r}$. We define several parameters in Table I.

TABLE I
PARAMETERS WITH A BEING THE COEFFICIENT MATRIX.

η_1	$\min_{i\in[m]}\left\{\|a_i\|^2\right\}$	η_2	$\max_{i\in[m]}\left\{\|a_i\|^2\right\}$
η_3	$\min_{j\in[n]}\left\{\|A_j\|^2\right\}$	η_4	$\max_{j\in[n]}\left\{\|A_j\|^2\right\}$
θ_1	$\|A\|_F^2$	θ_2	$\|A\|_F^2-\eta_1$
θ_3	$\|A\|_F^2-\eta_3$	θ_4	$m\eta_2$
θ_5	$n\eta_4$	ρ_1	$1-\sigma_{A,r}^2/\theta_1$
ρ_2	$1-\sigma_{A,r}^2/\theta_2$	ρ_3	$1-\sigma_{A,r}^2/\theta_3$
ρ_4	$1-\sigma_{A,r}^2/\theta_5$	ρ_{COP}	$1-\det(A^TA)/\prod_{j=1}^n\|A_j\|^2$

II. A DETERMINISTIC ORTHOGONAL PROJECTION METHOD

When the system of linear equations $Ax=b$ is consistent ($b_N=0$), x^* is one of the least-squares solutions and has the minimum Euclidean norm [19]. Assume that $\widehat{e}^{(k)}:=x^{(k)}-x^*$ for $k=0,1,2,\cdots$. By the Pythagorean theorem, a direct calculation indicates that the squared error of (2) satisfies

$$\begin{aligned}\|\widehat{e}^{(k+1)}\|^2&=\|\widehat{e}^{(k)}\|^2-\frac{|b_{i_k}-a_{i_k}^Tx^{(k)}|^2}{\|a_{i_k}\|^2}\\&:=\|\widehat{e}^{(k)}\|^2-\varphi_{i_k}(A,x^{(k)},b)\end{aligned}\tag{3}$$

for $i_k\in[m]$. This implies that we may select the row index such that the corresponding loss is as large as possible. McCormick presented a deterministic greedy strategy to select the row index

$$i_k=\arg\max_{i\in[m]}\left\{\varphi_i(A,x^{(k)},b)\right\}$$

in [20, Section 2.1]. Du and Gao called it the maximal weighted residual Kaczmarz method and gave an easily computable theoretical estimate for the convergence rate of the MWRK method in [21]. This result is stated as follows. From any initial guess $x^{(0)}\in\mathcal{R}(A^T)$, the MWRK iteration sequence $\left\{x^{(k)}\right\}_{k=0}^{\infty}$ converges to x^* with the error estimate

$$\|\widehat{e}^{(k+1)}\|^2\le\rho_2^k\rho_1\|\widehat{e}^{(0)}\|^2\tag{4}$$

for $k=0,1,2,\cdots$. For more details, we refer to Theorem 3.1 in [21].

When the system $Ax=b$ is inconsistent ($b_N\neq0$), it follows from the convergence analysis of the MWRK method that

$$\widehat{e}^{(k+1)}=\left(I_n-\frac{a_{i_k}a_{i_k}^T}{\|a_{i_k}\|^2}\right)\widehat{e}^{(k)}+\frac{a_{i_k}^Tb_N}{\|a_{i_k}\|^2}a_{i_k}^T$$

with $Ax^*=b_R$ and $b=b_N+b_R$. By orthogonality, it follows that

$$\|\widehat{e}^{(k+1)}\|^2=\|\widehat{e}^{(k)}\|^2-\frac{|a_{i_k}^T\widehat{e}^{(k)}|^2}{\|a_{i_k}\|^2}+\frac{|a_{i_k}^Tb_N|^2}{\|a_{i_k}\|^2}.$$

This indicates that the MWRK method may not converge to the least-squares solution due to the existence of b_N. To address this problem, extended row-action methods convert the inconsistent linear system into a consistent or approximately consistent form. This transformation modifies b to match or closely approximate b_R. The adjusted system removes the part that causes convergence issues, while maintaining the system's fundamental structure. The primary objective reduces to computing the orthogonal decomposition of b.

Bai and Wu proposed a Cyclic Orthogonal Projection (COP) to approximate b_N by solving the homogeneous linear system $A^Tz=0$ in [13]. The iterative method is initialized with $z^{(0)}=b$, and at each iteration k, it selects an index $j\in[n]$ in a cyclic order. The update rule applies an orthogonal projection according to formula (2). For any $k=0,1,2,\cdots$, let $\widetilde{e}^{(k)}:=z^{(k)}-z^*$ with $z^*=b_N$. After k iterations, the expected convergence rate of the COP method is

$$\|\widetilde{e}^{(k+n)}\|^2\le\rho_{\text{COP}}^k\|\widetilde{e}^{(0)}\|^2.\tag{5}$$

Mustafa and Saha later introduced the Maximal Residual Orthogonal Projection (MROP) method, replacing cyclic selection with a greedy rule

$$j_k=\arg\max_{j\in[n]}\left\{|A_j^Tz^{(k)}|\right\}.$$

This variant achieves a deterministic convergence rate

$$\|\widetilde{e}^{(k)}\|^2\le\rho_4^k\|\widetilde{e}^{(0)}\|^2.\tag{6}$$

For further analysis, see Lemma 2.3 in [18].

In the orthogonal projection process, we know that

$$\begin{aligned}\|\widetilde{e}^{(k+1)}\|^2&=\|\widetilde{e}^{(k)}\|^2-\frac{|A_{j_k}^Tz^{(k)}|^2}{\|A_{j_k}\|^2}\\&=\|\widetilde{e}^{(k)}\|^2-\psi_{j_k}(A,z^{(k)},0).\end{aligned}\tag{7}$$

This implies that we can select the column index j_k such that the corresponding loss is maximized, which motivates us to propose the following MWROP method; see Algorithm 1.

Algorithm 1 The MWROP method.

Input: The coefficient matrix $A \in \mathbb{R}^{m\times n}$, an initial vector $z^{(0)} \in b + \mathcal{R}(A)$, and the maximum iteration number ℓ.
Output: $z^{(\ell)}$.
1: **for** $k = 0, 1, \cdots, \ell - 1$ **do**
2: compute $z^{(k+1)} = \mathcal{T}_{j_k}(A, z^{(k)}, 0)$ with $j_k = \arg\max_{j\in[n]} \left\{\psi_j(A, z^{(k)}, 0)\right\}$;
3: **endfor**

In the following, we will give a convergence analysis of the MWROP method in Theorem 1.

Theorem 1: Let $A \in \mathbb{R}^{m\times n}$ be a matrix whose columns are all nonzero. The iteration sequence $\left\{z^{(k)}\right\}_{k=0}^{\infty}$, generated by the MWROP method starting from any initial guess $z^{(0)} \in b + \mathcal{R}(A)$, converges to z^*, and satisfies the error bound

$$\|\widetilde{e}^{(k+1)}\|^2 \le \rho_3^k \rho_1 \|\widetilde{e}^{(0)}\|^2, \tag{8}$$

for $k = 0, 1, 2, \cdots$.

Proof. From formula (7), we observe that the error $\widetilde{e}^{(k)} \in \mathcal{R}(A)$ since z^* and $z^{(k)} \in b + \mathcal{R}(A)$. For the case of $k = 0$, we have

$$\begin{aligned}\|A^T z^{(0)}\|^2 &= \sum_{j\in[n]} \psi_j(A, z^{(0)}, 0)\|A_j\|^2 \\ &\le \max_{j\in[n]} \left\{\psi_j(A, z^{(0)}, 0)\right\} \|A\|_F^2.\end{aligned}$$

By Lemma 2.4 in [22], we have

$$\|A^T z^{(0)}\|^2 = \|A^T \widetilde{e}^{(0)}\|^2 \ge \sigma_{A,r}^2 \|\widetilde{e}^{(0)}\|^2.$$

After combining these two formulas together, we get

$$\|\widetilde{e}^{(1)}\|^2 \le \|\widetilde{e}^{(0)}\|^2 - \frac{\sigma_{A,r}^2}{\|A\|_F^2}\|\widetilde{e}^{(0)}\|^2 = \rho_1\|\widetilde{e}^{(0)}\|^2.$$

For the case of $k \ge 1$, since

$$A_{j_{k-1}}^T z^{(k)} = A_{j_{k-1}}^T \left(z^{(k)} - \frac{A_j^T z^{(k)}}{\|A_{j_{k-1}}\|^2} A_{j_{k-1}}\right) = 0,$$

it immediately yields that

$$\begin{aligned}\|A^T z^{(k)}\|^2 &= \sum_{j\in[n]} \psi_j(A, z^{(k)}, 0)\|A_j\|^2 \\ &= \sum_{j\in[n]/j_{k-1}} \psi_j(A, z^{(k)}, 0)\|A_j\|^2 \\ &\le \max_{j\in[n]} \left\{\psi_j(A, z^{(k)}, 0)\right\} \sum_{j\in[n]/j_{k-1}} \|A_j\|^2.\end{aligned}$$

Moreover,

$$\begin{aligned}\psi_{j_k}(A, z^{(k)}, 0) &= \max_{j\in[n]} \left\{\psi_j(A, z^{(k)}, 0)\right\} \\ &\ge \frac{1}{\sum_{j\in[n]/j_{k-1}} \|A_j\|^2}\|A^T z^{(k)}\|^2 \\ &\ge \frac{1}{\|A\|_F^2 - \eta_3}\|A^T z^{(k)}\|^2.\end{aligned}$$

Using Lemma 2.4 in [22] again, it follows that

$$\|A^T z^{(k)}\|^2 = \|A^T \widetilde{e}^{(k)}\|^2 \ge \sigma_{A,r}^2 \|\widetilde{e}^{(k)}\|^2,$$

which results in

$$\begin{aligned}\|\widetilde{e}^{(k+1)}\|^2 &\le \|\widetilde{e}^{(k)}\|^2 - \frac{1}{\|A\|_F^2 - \eta_3}\sigma_{A,r}^2\|\widetilde{e}^{(k)}\|^2 \\ &= \rho_3\|\widetilde{e}^{(k)}\|^2.\end{aligned}$$

By recursion, we obtain the conclusion in Theorem 1. The proof is complete. □

III. THE MAXIMUM WEIGHTED RESIDUAL EXTENDED KACZMARZ METHOD

Mustafa and Saha proposed the MREK method in [18] for iteratively computing the least-squares solution x^* of the large sparse inconsistent linear system. Here, we reformulate it by specifying the initial guesses $x^{(0)} \in \mathcal{R}(A^T)$ and $z^{(0)} \in b + \mathcal{R}(A)$, in Algorithm 2. An upper bound for the MREK solution error was given in [18, Theorem 2.4]. This result is precisely restated as follows.

Algorithm 2 The MREK method.

Input: The coefficient matrix $A \in \mathbb{R}^{m\times n}$, the right-hand side $b \in \mathbb{R}^m$, two initial guesses $x^{(0)} \in \mathcal{R}(A^T)$ and $z^{(0)} \in b + \mathcal{R}(A)$, and the maximum iteration number ℓ.
Output: $x^{(\ell)}$.
1: **for** $k = 0, 1, \cdots, \ell - 1$ **do**
2: compute $z^{(k+1)} = \mathcal{T}_{j_k}(A, z^{(k)}, 0)$ with $j_k = \arg\max_{j\in[n]} \left\{|A_j^T z^{(k)}|\right\}$;
3: compute $x^{(k+1)} = \mathcal{P}_{i_k}(A, x^{(k)}, b - z^{(k+1)})$ with $i_k = \arg\max_{i\in[m]} \left\{|b_i - z_i^{(k+1)} - a_i^T x^{(k)}|\right\}$;
4: **endfor**

Theorem 2: ([18, Theorem 2.4]) Let $A \in \mathbb{R}^{m\times n}$ be a matrix without any zero rows and $b \in \mathbb{R}^m$. Starting from any initial guesses $x^{(0)} \in \mathcal{R}(A^T)$ and $z^{(0)} \in b + \mathcal{R}(A)$, the iteration sequences $\left\{x^{(k)}\right\}_{k=0}^{\infty}$ and $\left\{z^{(k)}\right\}_{k=0}^{\infty}$ are generated by the MREK method for solving the inconsistent linear system $Ax = b$. Then, the error estimate satisfies

$$\begin{aligned}\|x^{(k+1)} - x^*\|^2 \le \rho_{1,\alpha}^{k+1}\|x^{(0)} - x^*\|^2 + \widetilde{c}_0 \frac{\rho_4^{k+1} - \rho_{1,\alpha}^{k+1}}{\rho_4 - \rho_{1,\alpha}} \\ \cdot \rho_4\|z^{(0)} - z^*\|^2,\end{aligned} \tag{9}$$

where the constants $\rho_{1,\alpha} = 1 - \alpha^2\sigma_{A,r}^2/\theta_4$ with $\alpha \in (0, 1)$ and

$$\widetilde{c}_0 = \frac{1}{M}\left(1 + \frac{\alpha}{1-\alpha} + \frac{\alpha^2}{m(1-\alpha)}\right).$$

In the MREK iteration for computing $x^{(k+1)}$, we know that

$$\begin{aligned}\widehat{e}^{(k+1)} &= x^{(k+1)} - x^* \\ &= \widehat{e}^{(k)} + \frac{(b_R)_{i_k} - a_{i_k}^T x^{(k)}}{\|a_{i_k}\|^2} a_{i_k} + \frac{(b_N)_{i_k} - z_{i_k}^{(k+1)}}{\|a_{i_k}\|^2} a_{i_k} \\ &= \widehat{e}^{(k)} - \frac{a_{i_k} a_{i_k}^T}{\|a_{i_k}\|^2}\widehat{e}^{(k)} + \frac{(b_N)_{i_k} - z_{i_k}^{(k+1)}}{\|a_{i_k}\|^2} a_{i_k} \\ &= \left(I_n - \frac{a_{i_k} a_{i_k}^T}{\|a_{i_k}\|^2}\right)\widehat{e}^{(k)} + \frac{(b_N)_{i_k} - z_{i_k}^{(k+1)}}{\|a_{i_k}\|^2} a_{i_k}\end{aligned}$$

for $k = 0, 1, 2, \cdots$. Taking the norm on both sides yields

$$\|\widehat{e}^{(k+1)}\|^2 = \|\widehat{e}^{(k)}\|^2 - \frac{|a_{i_k}^T \widehat{e}^{(k)}|^2}{\|a_{i_k}\|^2} + \frac{|(b_N - z^{(k+1)})_{i_k}|^2}{\|a_{i_k}\|^2}.$$

There exists a constant $\alpha \in (0, 1)$ such that

$$\begin{aligned} |a_{i_k}^T \widehat{e}^{(k)}|^2 &= |a_{i_k}^T x^{(k)} - (b_R)_{i_k} - (b_N)_{i_k} + z_{i_k}^{(k+1)} \\ &\quad + (b_N)_{i_k} - z_{i_k}^{(k+1)}|^2 \\ &\geq \alpha |a_{i_k}^T x^{(k)} - (b_R)_{i_k} - (b_N)_{i_k} + z_{i_k}^{(k+1)}|^2 \\ &\quad - \frac{\alpha}{1-\alpha} |(b_N)_{i_k} - z_{i_k}^{(k+1)}|^2 \\ &= \alpha |b_{i_k} - z_{i_k}^{(k+1)} - a_{i_k}^T x^{(k)}|^2 \\ &\quad - \frac{\alpha}{1-\alpha} |(b_N)_{i_k} - z_{i_k}^{(k+1)}|^2. \end{aligned}$$

Thus,

$$\begin{aligned} \|\widehat{e}^{(k+1)}\|^2 &\leq \|\widehat{e}^{(k)}\|^2 - \alpha \varphi_{i_k}(A, x^{(k)}, b - z^{(k+1)}) \\ &\quad + \left(1 + \frac{\alpha}{1-\alpha}\right) \frac{|(b_N)_{i_k} - z_{i_k}^{(k+1)}|^2}{\|a_{i_k}\|^2} \\ &\leq \|\widehat{e}^{(k)}\|^2 - \alpha \varphi_{i_k}(A, x^{(k)}, b - z^{(k+1)}) \\ &\quad + \frac{1}{\theta_2(1-\alpha)} \|z^{(k+1)} - b_N\|^2 \\ &= \|\widehat{e}^{(k)}\|^2 - \alpha \varphi_{i_k}(A, x^{(k)}, b - z^{(k+1)}) \\ &\quad + \frac{1}{\theta_2(1-\alpha)} \left(\|\widetilde{e}^{(k)}\|^2 - \psi_{j_k}(A, z^{(k)}, 0) \right). \end{aligned} \tag{10}$$

This observation motivates the selection criterion for the index pair (i_k, j_k). That is, we prioritize indices corresponding to the largest loss values φ_{i_k} and ψ_{j_k}. This optimal selection strategy forms the foundation of our proposed MWREK method, as described in Algorithm 3.

Algorithm 3 The MWREK method.

Input: The coefficient matrix $A \in \mathbb{R}^{m \times n}$, the right-hand side $b \in \mathbb{R}^m$, two initial guesses $x^{(0)} \in \mathcal{R}(A^T)$ and $z^{(0)} \in b + \mathcal{R}(A)$, and the maximum iteration number ℓ.

Output: $x^{(\ell)}$.

1: **for** $k = 0, 1, \cdots, \ell - 1$ **do**

2: compute $z^{(k+1)} = \mathcal{T}_{j_k}(A, z^{(k)}, 0)$ with $j_k = \arg\max_{j \in [n]} \left\{ \psi_j(A, z^{(k)}, 0) \right\}$;

3: compute $x^{(k+1)} = \mathcal{P}_{i_k}(A, x^{(k)}, b - z^{(k+1)})$ with $i_k = \arg\max_{i \in [m]} \left\{ \varphi_i(A, x^{(k)}, b - z^{(k+1)}) \right\}$;

4: **endfor**

Having introduced the MWREK method, we proceed to analyze its convergence properties. The following theorem characterizes its convergence behavior.

Theorem 3: Let $A \in \mathbb{R}^{m \times n}$ be a matrix without any zero rows and $b \in \mathbb{R}^m$. Starting from any initial values $x^{(0)} \in \mathcal{R}(A^T)$ and $z^{(0)} \in b + \mathcal{R}(A)$, the iteration sequences $\{x^{(k)}\}_{k=0}^{\infty}$ and $\{z^{(k)}\}_{k=0}^{\infty}$ are generated by the MWREK method for solving the inconsistent linear system $Ax = b$. Then, for $k = 0$, the following error estimate holds

$$\|x^{(1)} - x^*\|^2 \leq \rho_{2,\alpha} \|x^{(0)} - x^*\|^2 + \widetilde{c}_1 \rho_1 \|z^{(0)} - z^*\|^2 = \bar{c}; \tag{11}$$

and for $k \geq 1$

$$\begin{aligned} \|x^{(k+1)} - x^*\|^2 &\leq \bar{c} \rho_{3,\alpha}^k \|x^{(0)} - x^*\|^2 \\ &\quad + \widetilde{c}_2 \rho_1 \sum_{\ell=0}^{k} \rho_{1,\alpha}^{\ell} \rho_2^{k+1-\ell} \|z^{(0)} - z^*\|^2, \end{aligned} \tag{12}$$

where the constants $\rho_{2,\alpha} = 1 - \alpha^2 \sigma_{A,r}^2 / \theta_1$ and $\rho_{3,\alpha} = 1 - \alpha^2 \sigma_{A,r}^2 / \theta_2$,

$$\widetilde{c}_1 = \frac{\alpha^2}{(1-\alpha)\|A\|_F^2} + \frac{1}{\theta_2(1-\alpha)},$$

and

$$\widetilde{c}_2 = \frac{\alpha^2}{(1-\alpha)(\|A\|_F^2 - \eta_1)} + \frac{1}{\theta_2(1-\alpha)}.$$

Proof. When $k = 0$,

$$\begin{aligned} \|b - z^{(1)} - Ax^{(0)}\|^2 &= \sum_{i \in [m]} \varphi_i(A, x^{(0)}, b - z^{(1)}) \|a_i\|^2 \\ &\leq \max_{i \in [m]} \left\{ \varphi_i(A, x^{(0)}, b - z^{(1)}) \right\} \|A\|_F^2. \end{aligned}$$

Furthermore,

$$\begin{aligned} &\varphi_{i_0}(A, x^{(0)}, b - z^{(1)}) \\ &= \max_{i \in [m]} \left\{ \varphi_i(A, x^{(0)}, b - z^{(1)}) \right\} \\ &\geq \frac{1}{\|A\|_F^2} \|b - z^{(1)} - Ax^{(0)}\|^2 \\ &\geq \frac{1}{\|A\|_F^2} \left(\alpha \|b_R - Ax^{(0)}\|^2 - \frac{\alpha}{1-\alpha} \|b_N - z^{(1)}\|^2 \right) \\ &\geq \alpha \frac{\sigma_{A,r}^2}{\|A\|_F^2} \|\widehat{e}^{(0)}\|^2 - \frac{\alpha}{1-\alpha} \frac{1}{\|A\|_F^2} \|\widetilde{e}^{(1)}\|^2, \end{aligned}$$

where the third inequality holds because $b_R = Ax^*$ and x^*, $x^{(0)} \in \mathcal{R}(A^T)$. From formula (10), we obtain

$$\begin{aligned} \|\widehat{e}^{(1)}\|^2 &\leq \|\widehat{e}^{(0)}\|^2 - \alpha \varphi_{i_0}(A, x^{(0)}, b - z^{(1)}) \\ &\quad + \frac{1}{\eta_1(1-\alpha)} \|\widetilde{e}^{(1)}\|^2 \\ &\leq \left(1 - \alpha^2 \frac{\sigma_{A,r}^2}{\|A\|_F^2}\right) \|\widehat{e}^{(0)}\|^2 \\ &\quad + \left(\frac{\alpha^2}{(1-\alpha)\|A\|_F^2} + \frac{1}{\theta_2(1-\alpha)} \right) \|\widetilde{e}^{(1)}\|^2 \\ &\leq \rho_{2,\alpha} \|\widehat{e}^{(0)}\|^2 + \widetilde{c}_1 \rho_1 \|\widetilde{e}^{(0)}\|^2 = \bar{c}. \end{aligned}$$

When $k \geq 1$,

$$\begin{aligned} &\|b - z^{(k+1)} - Ax^{(k)}\|^2 \\ &= \sum_{i \in [m]} \varphi_i(A, x^{(k)}, b - z^{(k+1)}) \|a_i\|^2 \\ &= \sum_{i \in [m]/i_{k-1}} \varphi_i(A, x^{(k)}, b - z^{(k+1)}) \|a_i\|^2 \\ &\leq \max_{i \in [m]} \left\{ \varphi_i(A, x^{(k)}, b - z^{(k+1)}) \right\} \sum_{i \in [m]/i_{k-1}} \|a_i\|^2 \\ &\leq \varphi_{i_k}(A, x^{(k)}, b - z^{(k+1)}) (\|A\|_F^2 - \eta_1). \end{aligned}$$

It follows that

$$
\begin{aligned}
&\varphi_{i_k}(A, x^{(k)}, b - z^{(k+1)}) \\
\geq\, & \frac{1}{\|A\|_F^2 - \eta_1} \|b - z^{(k+1)} - Ax^{(k)}\|^2 \\
\geq\, & \frac{1}{\|A\|_F^2 - \eta_1} \left(\alpha \|b_R - Ax^{(k)}\|^2 - \frac{\alpha}{1-\alpha} \|b_N - z^{(k+1)}\|^2 \right) \\
\geq\, & \alpha \frac{\sigma_{A,r}^2}{\|A\|_F^2 - \eta_1} \|\widehat{e}^{(k)}\|^2 - \frac{\alpha}{1-\alpha} \frac{1}{\|A\|_F^2 - \eta_1} \|\widetilde{e}^{(k+1)}\|^2,
\end{aligned}
$$

where the fourth inequality is from the facts that $b_R = Ax^*$ and x^*, $x^{(0)} \in \mathcal{R}(A^T)$. The following results are obtained after rearrangement

$$
\begin{aligned}
&\|\widehat{e}^{(k+1)}\|^2 \\
\leq\, & \|\widehat{e}^{(k)}\|^2 - \alpha^2 \frac{\sigma_{A,r}^2}{\|A\|_F^2 - \eta_1} \|\widehat{e}^{(k)}\|^2 \\
& + \left(\frac{\alpha^2}{(1-\alpha)(\|A\|_F^2 - \eta_1)} + \frac{1}{\theta_2(1-\alpha)} \right) \|\widetilde{e}^{(k+1)}\|^2 \\
\leq\, & \left(1 - \alpha^2 \frac{\sigma_{A,r}^2}{\|A\|_F^2 - \eta_1} \right) \|\widehat{e}^{(k)}\|^2 \\
& + \left(\frac{\alpha^2}{(1-\alpha)(\|A\|_F^2 - \eta_1)} + \frac{1}{\theta_2(1-\alpha)} \right) \rho_2^k \rho_1 \|\widetilde{e}^{(0)}\|^2 \\
=\, & \rho_{3,\alpha} \|\widehat{e}^{(k)}\|^2 + \widetilde{c}_2 \rho_2^k \rho_1 \|\widetilde{e}^{(0)}\|^2.
\end{aligned}
$$

After recursion with respect to k, we get the conclusion in formula (12). The proof is complete. □

Remark 1: From the convergence analysis of the MWREK method, we can see that it solves this problem by making the inconsistent linear system consistent or nearly consistent. This requires modifying the vector b, which should equal or closely approximate b_R, its projection onto the range space of A. For approximate consistency or keeping it within an acceptable distance of b_R, the MWREK method uses a deterministic orthogonal projection iteration to remove b_R, which is the null-space component causing the convergence problems for the MWRK method. Meanwhile, the structure of the linear system remains unchanged. This guarantees solution existence and preserves the problem's key features.

IV. Experimental results

In this section, we will give several examples using real-world data to demonstrate the convergence behavior of the proposed methods. For comparison, we use the implementation of several deterministic and randomized extended row-action methods including REK in [12], REK-S in [16], and MREK in [18].

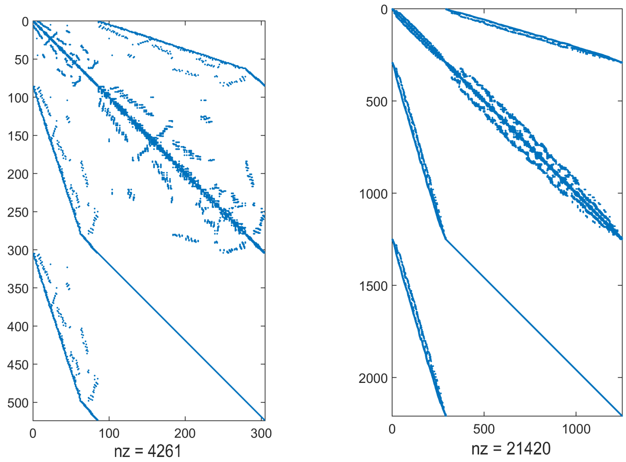
The test starts from two initial vectors $x^{(0)} = 0$ and $z^{(0)} = b$. The performance is compared in terms of iteration number (denoted by IT), relative solution error (denoted by RSE), and computing time in seconds (denoted by CPU), where RSE is defined by $\text{RSE} = \|x^{(k)} - x^*\|^2 / \|x^*\|^2$ for $k = 0, 1, 2, \cdots$ and CPU is measured using the MATLAB built-in function `tic-toc`. The algorithms are carried out on a Founder desktop PC with Intel(R) Core(TM) i5-7500 CPU 3.40 GHz.

Example 1: Consider the original least-squares problems, where the coefficient matrix

$$A = \left[A_1^T A_1 \quad A_1^T; \quad A_1 A_1^T \quad A_1; \quad A_1 \quad \Lambda \right].$$

Here the sub-matrix A_1 comes from the Suite Sparse Matrix Collection [23], including `ash219` and `ash918`, $\Lambda = \text{diag}(1 : m_1)$ is a diagonal matrix with m_1 being the row number of A_1. We generate a solution x^* by the MATLAB function `ones`. The inconsistent linear system is realized by setting the noisy right-hand side as $b = Ax^* + \delta \cdot \widehat{b}$, where $\widehat{b} = \widetilde{b} - A\widetilde{x}$ with nonzero vectors $\widetilde{b} = [(-1)^0 \ (-1)^1 \ \cdots \ (-1)^{m-1}]^T$ and $\widetilde{x}$ being in the null space of A^T generated by `null`, and the noise level $\delta = 0.01$.

Using data-sets from Example 1, comparative experiments are conducted under varying extended row-action iteration methods. We first display the sparsity pattern of matrix A in Figure 1, then present the convergence curves in Figure 2. The results demonstrate that MWREK and MREK have similar convergence behavior and successfully compute an approximate solution for all cases. In all convergent cases, the MWREK method exhibits significantly lower iteration counts and computing times compared to MREK.



(a) $A_1 =$ `ash219` (b) $A_1 =$ `ash958`

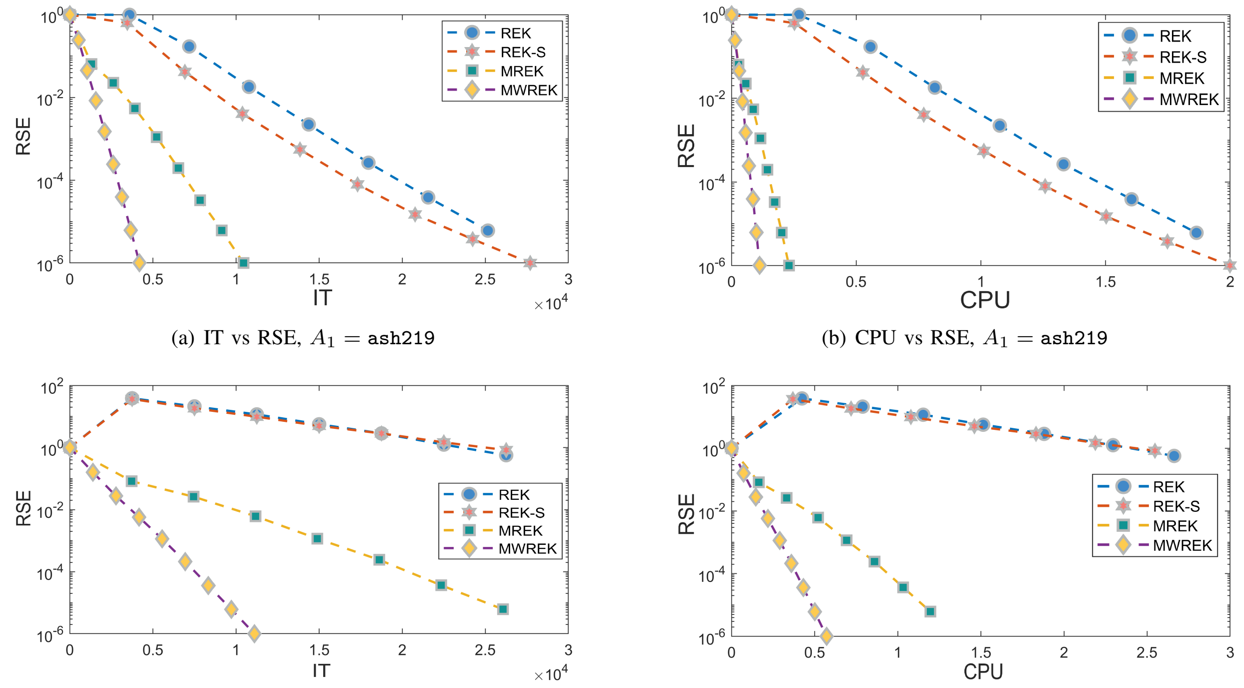
Fig. 1. Sparsity pattern of matrix A from Example 1.

Example 2: Consider the following least-squares problem, characterized by the coefficient matrix

$$A = \left[A_1^T A_1 \quad A_1^T; \quad A_1 A_1^T \quad A_1; \quad A_1 \quad \Lambda \right],$$

where the sub-matrix A_1 is generated using the MATLAB function `randn` with dimensions $m_1 \times n_1$, and $\Lambda = \text{diag}(1 : m_1)$ represents a diagonal matrix. The solution x^* is created using the MATLAB function `ones`. An inconsistent linear system is constructed by defining the noisy right-hand side as $b = Ax^* + \delta \cdot \widehat{b}$, where $\widehat{b} = \widetilde{b} - A\widetilde{x}$ involves nonzero vector $\widetilde{b} = [(-1)^0, (-1)^1, \ldots, (-1)^{m-1}]^T$ and $\widetilde{x}$, which lies in the null space of A^T and is generated by the MATLAB function `null`. The noise level is specified as $\delta = 0.01$.

The numerical results for Example 2, obtained using the REK, REK-S, MREK, and MWREK methods, are presented in Table II. We consider four different data sizes: $(m_1, n_1) = (500, 50)$, $(500, 150)$, $(1000, 50)$, and $(1000, 150)$. The results indicate that the MWREK method achieves comparable accuracy while significantly reducing the computation time



(a) IT vs RSE, $A_1 =$ `ash219`

(b) CPU vs RSE, $A_1 =$ `ash219`

(c) IT vs RSE, $A_1 =$ `ash958`

(d) CPU vs RSE, $A_1 =$ `ash958`

Fig. 2. The number of iteration steps and computing times of REK, REK-S, MREK, and MWREK for Example 1.

compared to the REK, REK-S, and MREK methods. This advantage becomes increasingly pronounced as the data size increases.

Example 3: We solve the image reconstruction problem chosen from AIR Tools II toolbox [24]. We select 2-dimension X-ray tomography as the test problem, which is shown in Figure 3(a). It is represented by $[A, b, x^*] = \mathrm{paralleltomo}(N, \theta, p)$, which generates a coefficient matrix A, an exact solution x^*, and an exact right-hand side $b = Ax^*$ by adjusting the input parameters, such as the size of the discrete domain (N), the projection angles in degrees (θ), and the number of rays (p). The inconsistent linear system is realized by setting the noisy right-hand side as $b = Ax^* + \delta \cdot \widehat{b}$, where $\widehat{b} = \widetilde{b} - A\widetilde{x}$ with nonzero random vector $\widetilde{b}$ generated by `randn` and $\widetilde{x}$ being in the null space of A^T generated by `null`, and the noise level $\delta = 0.01$.

In this tomography test problem, we set the input parameters as $N = 50$, $\theta = 0 : 3 : 360$, and $p = 50$, resulting in a coefficient matrix of size 6050×2500. The numerical results for the iteration number, relative solution error, and computing time, provided by REK, REK-S, MREK, and MWREK, are depicted in Figure 4. It is evident that MWREK exhibits significantly lower iteration counts and computing times compared to REK, REK-S, and MREK. Therefore, MWREK substantially outperforms the other tested methods. In Figure 3(c), we present the $N \times N$ images of the approximate tectonic phantom obtained by MWREK. The image accurately converges to the exact solution.

V. Conclusions

The row-action method, known for its simplicity, has found widespread application in various domains, including computer graphics. This paper introduces a novel deterministic row-action approach for computing the orthogonal decomposition of linear systems. The method employs a maximum weighted residual selection strategy to enhance efficiency. Furthermore, the approach is extended to address overdetermined and inconsistent linear systems. Theoretical analysis demonstrates that the proposed methods can linearly converge to the least-squares solution with a minimum Euclidean norm. To validate the effectiveness of the new methods, several numerical studies are conducted. The results show that the proposed methods achieve a faster convergence rate compared to existing deterministic extended row-action methods.

References

[1] Y. Censor, "The mathematics of computerized tomography," *Inverse Problems*, vol. 18, no. 1, pp. 283, 2002.

[2] K. W. Chang, C. J. Hsieh, and C. J. Lin, "Coordinate descent method for large-scale l_2-loss linear support vector machines," *The Journal of Machine Learning Research*, vol. 9, no. 46, pp. 1369-1398, 2008.

[3] C. Byrne, "A unified treatment of some iterative algorithms in signal processing and image reconstruction," *Inverse Problems*, vol. 20, no. 1, pp. 103, 2003.

[4] D. S. Purushothama, "A new outer inverse of a matrix and its characterizations," *IAENG International Journal of Applied Mathematics*, vol. 55, no. 2, pp. 431-435, 2025.

[5] M. Hanke and W. Niethammer, "On the acceleration of Kaczmarz's method for inconsistent linear systems," *Linear Algebra and Its Applications*, vol. 130, pp. 83-98, 1990.

[6] Y. Censor, "Row-action methods for huge and sparse systems and their applications," *SIAM Review*, vol. 23, no. 4, pp. 444-466, 1981.

[7] Y. Censor and S. A. Zenios, "Parallel Optimization: Theory, Algorithms, and Applications," Oxford University Press, New York, USA, 1997.

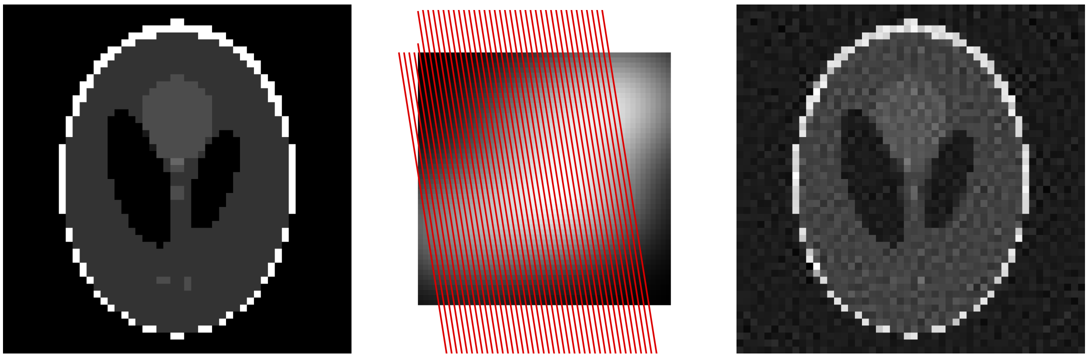
[8] R. Escalante and M. Raydan, "Alternating Projection Methods," Society for Industrial and Applied Mathematics, Philadelphia, USA, 2011.

[9] S. Kaczmarz, "Angenäherte auflösung von systemen linearer gleichungen," *Bulletin International de l'Académie Polonaise des Sciences A*, vol. 35, pp. 355-357, 1937.

[10] Y. Liu, S. Liu, and Z. Zhang, "Accelerated greedy randomized Kaczmarz algorithm for solving linear systems," *IAENG International Journal of Computer Science*, vol. 50, no. 3, pp. 1009-1015, 2023.

TABLE II
NUMERICAL RESULTS GIVEN BY REK, REK-S, MREK, AND MWREK FOR EXAMPLE 2.

Methods	$m_1 = 500, n_1 = 50$		$m_1 = 500, n_1 = 150$		$m_1 = 1000, n_1 = 50$		$m_1 = 1000, n_1 = 150$	
	IT	CPU	IT	CPU	IT	CPU	IT	CPU
REK	212704	13.094	148683	9.786	745157	67.629	348860	32.643
REK-S	193007	12.018	142652	9.823	725887	65.506	324553	32.060
MREK	26916	0.302	31796	0.443	90564	2.641	61976	2.052
MWREK	**10814**	**0.141**	**20036**	**0.288**	**18721**	**0.493**	**24653**	**0.797**



(a) Exact solution (b) Parallel-beam (c) Approximate solution

Fig. 3. The exact solution of the test problem (a). The geometries of parallel-beam (b) with a few of rays from Example 3. The approximate solutions obtained by MWREK (c).

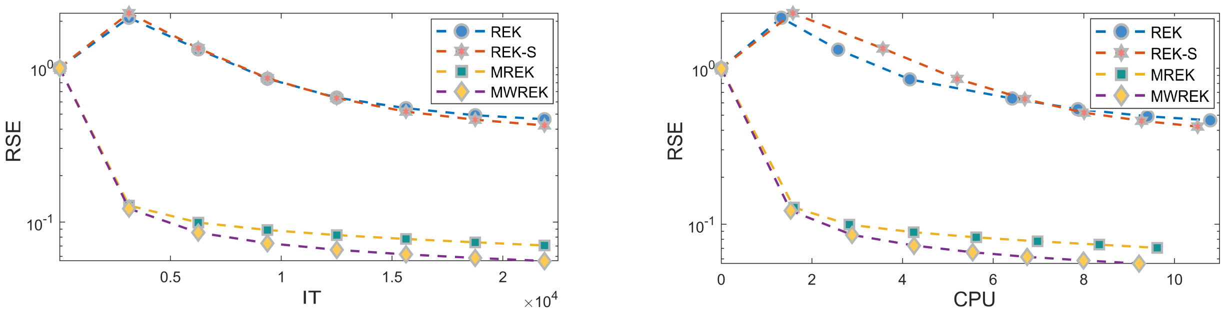


Fig. 4. The number of iteration steps and computing times of REK, REK-S, MREK, and MWREK for Example 3.

[11] C. Popa, "Least-squares solution of overdetermined inconsistent linear systems using Kaczmarz's relaxation," *International Journal of Computer Mathematics*, vol. 55, no. 1-2, pp. 79-89, 1995.

[12] A. Zouzias and N. M. Freris, "Randomized extended Kaczmarz for solving least squares," *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 2, pp. 773-793, 2013.

[13] Z. Z. Bai and W. T. Wu, "On partially randomized extended Kaczmarz method for solving large sparse overdetermined inconsistent linear systems," *Linear Algebra and Its Applications*, vol. 578, pp. 225-250, 2019.

[14] N. Chen and D. Zhu, "The greedy randomized extended Kaczmarz algorithm for noisy linear systems," *Journal of Applied Analysis and Computation*, vol. 13, no. 2, pp. 913–927, 2023.

[15] W. Bao, Z. Lv, F. Zhang, and W. Li, "A class of residual-based extended Kaczmarz methods for solving inconsistent linear systems," *Journal of Computational and Applied Mathematics*, vol. 416, pp. 114529, 2022.

[16] K. Du, "Tight upper bounds for the convergence of the randomized extended Kaczmarz and Gauss-Seidel algorithms," *Numerical Linear Algebra with Applications*, vol. 26, no. 3, pp. e2233, 2019.

[17] C. X. Gao and F. Chen, "Modified partially randomized extended Kaczmarz method with residual for solving large sparse linear systems," *Applied Numerical Mathematics*, vol. 212, pp. 215-222, 2025.

[18] A. Mustafa and M. Saha, "Maximal residual extended Kaczmarz and Gauss-Seidel methods-convergence properties and applications," *Journal of Computational and Applied Mathematics*, vol. 448, pp. 115920, 2024.

[19] Z.-Z. Bai and W.-T. Wu, "On greedy randomized Kaczmarz method for solving large sparse linear systems," *SIAM Journal on Scientific Computing*, vol. 40, no. 1, pp. A592-A606, 2018.

[20] S. F. McCormick, "The methods of Kaczmarz and row orthogonalization for solving linear equations and least squares problems in Hilbert space," *Indiana University Mathematics Journal*, vol. 26, no. 6, pp. 1137-1150, 1977.

[21] K. Du and H. Gao, "A new theoretical estimate for the convergence rate of the maximal weighted residual Kaczmarz algorithm," *Numerical Mathematics: Theory, Methods and Applications*, vol. 12, no. 2, pp. 627-639, 2019.

[22] K. Du, W.-T. Si, and X.-H. Sun, "Randomized extended average block Kaczmarz for solving least squares," *SIAM Journal on Scientific Computing*, vol. 42, no. 6, pp. A3541-A3559, 2020.

[23] T. A. Davis and Y. Hu, "The University of Florida sparse matrix collection," *ACM Transactions on Mathematical Software*, vol. 38, no. 1, pp. 1-25, 2011.

[24] P. C. Hansen and J. S. Jørgensen, "AIR Tools II: Algebraic iterative reconstruction methods, improved implementation," *Numerical Algorithms*, vol. 79, no. 1, pp. 107-137, 2018.