Row Selection Based on Equity of Attention in Randomized Kaczmarz Precoding for Massive MIMO Systems

TATSUKI FUKUDA
Department of Data Science, Faculty of Data Science,
Shimonoseki City University,
2-1-1 Daigakucho, Shimonoseki City, Yamaguchi, 751-8510,
JAPAN

Abstract: Randomized Kaczmarz (rKA) methods approximate the matrix inversion of Regularized Zero-Forcing (RZF) precoding in massive MIMO systems, with convergence governed by the row-selection rule. The Sainte-Laguë (SL) schedules of the author's prior work give no guarantee on the row-visitation counts if the iteration is truncated early. This paper proposes a rule based on Equity of Attention (EoA), an amortized fairness notion from search-result ranking, which at every step selects the row with the largest deficit between its target and scheduled visitation count, minimizing the ℓ_1 deviation from the target allocation step by step. In Monte Carlo simulations over i.i.d. Rayleigh channels, in both the row-projection and coordinate-descent formulations, EoA converges faster than SL whenever $M/K \gtrsim 1.25$, except at the smallest iteration budgets. SL is preferable only for $M/K \lesssim 1.2$ at large budgets, where the condition number of HH^H grows steeply.

Key-Words: Randomized Kaczmarz Algorithm, Equity of Attention, Row Selection, Massive MIMO, Regularized Zero-Forcing Precoding, Sainte-Laguë Method, Amortized Fairness, Condition Number.

Received: May 29, 2026. Revised: August 4, 2026. Accepted: September 6, 2026. Published: November 26, 2026.

1 Introduction

The transition to extremely large-scale massive MIMO has made the matrix inversion at the heart of Regularized Zero-Forcing (RZF) precoding a first-order cost concern. The randomized Kaczmarz Algorithm (rKA) replaces the exact inverse with an iterative approximation that touches one row of the regularized channel Gram matrix at a time, and prior work, including the author's own study, [1], has shown that the *order* in which rows are visited strongly affects the iterations needed to reach a target bit error rate (BER).

That study replaced the D'Hondt-based row-selection rule, [2], with the Sainte-Laguë (SL) apportionment method. As applied in that study, [1], SL allocates the T planned iterations in proportion to votes drawn at random from the rows' importance weights and, in one of its forms, visits the allocated rows in random order. The schedule therefore carries no guarantee about the counts after $t < T$ iterations, which is exactly the situation of a real-time system that must terminate early.

Equity of Attention, [3], from information retrieval, addresses such an *anytime* setting: the cumulative attention received by an item should track its cumulative relevance at every point of a sequence of rankings (Section 2.4). This paper applies its greedy allocation rule, without the ranking-specific constraints, to rKA row selection, and calls the result *Equity of Attention* (EoA) row selection. Our contributions are threefold:

1. We introduce EoA row selection in both the row-projection and coordinate-descent (Gauss–Seidel) formulations of rKA precoding; it needs only a running deficit counter per row, and we state its deficit-minimization property as a proposition.
2. We show experimentally that, in the tested i.i.d. Rayleigh settings and apart from the smallest iteration budgets, EoA converges faster than the published SL schedules whenever $M/K \gtrsim 1.25$; at $M = 64$ the advantage grows as K increases from 2 to 8, and in the coordinate-descent formulation it is far larger than SL's own advantage over uniformly random row selection.
3. We identify a narrow band close to $M/K = 1$ in which SL becomes preferable, but only at large iteration budgets. This band coincides with the steep growth of the median condition number of the Gram matrix HH^H as $M/K \to 1$.

The elements inherited from the author's prior work, [1], [2], are the RZF/rKA precoding framework and its coordinate-descent implementation, the row-importance weight, the SL row-selection procedure used as the baseline, and the Gauss–Markov CSI-error model. New in this paper are the EoA rule, its normalization and Proposition 1, the row-projection form applied to the $K \times K$ matrix A, the single-cell i.i.d. Rayleigh setting, and all simulations and their results in Sections 4 and 5. The simulator was implemented anew for this study; no figure or result of the prior papers is reproduced.

2 Related Work

2.1 Randomized Kaczmarz Algorithm and Its Extensions

The rKA was introduced as a randomized alternative to the classical Kaczmarz method, selecting the row to project onto with probability proportional to its squared norm and thereby achieving exponential expected convergence, [4]. Subsequent work has continued to refine both the convergence theory and the row (or block) selection strategy of rKA: greedy variants that deterministically select the row of maximal residual contribution, [5], semi-randomized variants that avoid computing selection probabilities, [6], [7], block and inertial extensions, [8], [9], and adaptive "selectable-set" strategies that exploit the current iterate, [10], among others, [11], [12]. The choice of row-selection rule thus remains open and actively studied — the question this paper addresses from a fairness-theoretic angle.

2.2 Apportionment-Based Row Selection for Precoding

An earlier study by the author, [2], first applied an election-apportionment rule, the D'Hondt method, to rKA row selection for RZF precoding, framing the row's importance (from its channel-norm contribution) as a vote count and the row-visitation schedule as a seat allocation. The author's prior work, [1], showed that D'Hondt's bias toward large vote counts carries over into a bias toward rows with large importance weights, and replaced it with the SL method, whose odd-divisor sequence minimizes the sum of squared deviations between vote share and seat share. The present paper continues that line of work; the apportionment literature itself keeps developing new allocation rules and monotonicity results, [13], [14], that may offer further candidates beyond SL and EoA.

2.3 Low-Complexity Precoding for Massive and Extremely Large-Scale MIMO

Iterative, matrix-inversion-free precoding continues to be an active alternative to direct RZF computation as antenna counts grow into the hundreds, both for conventional massive MIMO and for extremely large-scale (XL-)MIMO deployments, [15], [16]. A closely related line of work addresses the spatial non-stationarity that arises once the array is partitioned into subarrays, each observing only a subset of the propagation environment, [17], [18], [19]. The author's prior study, [1], used such a subarray model; the present paper does not.

2.4 Equity of Attention and Fairness of Exposure in Rankings

Equity of attention, [3], was formalized as an individual (rather than group) fairness criterion for rankings: the attention a subject receives, cumulated over a sequence of rankings, should be proportional to its cumulated relevance. Because position bias makes this unattainable within any single ranking, the problem is cast as an online optimization minimizing the ℓ_1 discrepancy between cumulative attention and cumulative relevance at every step, and a simple greedy heuristic — always promoting the most-underserved item — performs comparably to the full ILP-based mechanism once the ranking-quality constraint is relaxed. A closely related, group-oriented notion of exposure fairness was proposed independently around the same time, [20]. Amortizing fairness over a sequence of decisions, rather than requiring it within each one, has continued to develop since, including vertical, provider-aware exposure allocation, [21], fair division formulations of ranking, [22], and non-linear or provider-tailored exposure functions, [23], [24]. This paper is, to our knowledge, the first to apply this line of work outside information retrieval, to the row-selection problem in randomized iterative linear solvers.

2.5 Robust Precoding Under Imperfect Channel State Information

Because channel state information (CSI) at the base station is never perfect, robust precoding under CSI error remains an active concern, [25], [26]. Section 4.7 evaluates EoA under this kind of estimation error, using the same Gauss–Markov error model as the author's prior work, [1].

3 Methodology

3.1 System Model and rKA Precoding

We adopt the downlink precoding formulation of the author's prior work, [1], in a single-cell setting without subarrays: a base station with M antennas serves K single-antenna users over a flat-fading channel $H \in \mathbb{C}^{K\times M}$. RZF precoding computes

$$W = H^H \left(HH^H + \alpha I_K\right)^{-1}, \tag{1}$$

where $\alpha = 1/\mathrm{SNR}_{\mathrm{linear}}$, normalized so that $\mathbb{E}\left[\|Ws\|^2\right] = K$. Writing $A = HH^H + \alpha I_K \in \mathbb{C}^{K\times K}$, rKA methods avoid the explicit inverse in Equation (1) by iteratively solving $AX = I_K$ (row-projection form) or, equivalently, K independent systems $Az_k = e_k$ (coordinate-descent / Gauss–Seidel form), visiting one row (row-projection) or one equation (coordinate-descent) of A per iteration. In the row-projection form, one iteration selects one row

of A and updates all K columns of X; in the coordinate-descent form, one iteration updates one coordinate of each z_k, each column k following its own schedule whose first entry is its own row k (warm start). In both formulations, the sequence of rows visited over T iterations is exactly what a row-selection rule must specify.

3.2 Row Selection via Sainte-Laguë Apportionment (Prior Work)

Following the author's prior work, [1], [2], each row i is assigned a raw importance ("vote") weight

$$\tilde{p}_i = 1 - \frac{\|a_i\|^2 + \xi}{\|A\|_F^2 + K\xi}, \tag{2}$$

where a_i is the i-th row of A and ξ a regularization constant ($\xi = \alpha$ here). In the coordinate-descent formulation the same weight is evaluated on the diagonal, $\tilde{p}_i = 1 - A_{ii}/\mathrm{tr}(A)$ (equivalently, with $\|h_i\|^2$ and $\|H\|_F^2$ replacing $\|a_i\|^2$ and $\|A\|_F^2$), following that work. Since $\sum_i \|a_i\|^2 = \|A\|_F^2$, the raw weights in Equation (2) satisfy $\sum_i \tilde{p}_i = K - 1$ for any ξ; we therefore use the normalized share

$$p_i = \frac{\tilde{p}_i}{\sum_{j=1}^{K} \tilde{p}_j} = \frac{\tilde{p}_i}{K-1}, \qquad \sum_{i=1}^{K} p_i = 1. \tag{3}$$

SL draws a virtual electorate of Γ votes at random across rows according to p_i, and greedily allocates the T available iterations to the rows with the largest Sainte-Laguë quotient $V_i/(2q_i + 1)$, where q_i is the number of iterations already allocated to row i and V_i its vote count. (SL depends only on the relative magnitudes of the p_i, so the normalization in Equation (3) leaves its schedule unchanged; it becomes essential for EoA below.) The two published variants serve as baselines. The row-projection variant (`sl`) follows Algorithm 1 of the prior study, [1], applied to the rows of A, with $\Gamma = 10^4$ votes (our choice) and the allocated rows visited in random order: a random permutation followed by at most five passes of swaps that reduce adjacent repetitions while preserving each row's allocated count (unlike Algorithm 1 of the prior study, [1], the complete removal of adjacent repetitions is not enforced). The coordinate-descent variant (`rka_sl`) follows the experiment code of the prior study, [1], with $\Gamma = T$ votes and the rows visited in the sequential allocation order. Because the votes are sampled, the allocated counts are proportional to the sampled V_i rather than to p_i, and the random visiting order of `sl` makes the counts of any prefix of length $t < T$ random as well; neither schedule carries a guarantee about the counts at $t < T$.

3.3 Row Selection via Equity of Attention (Proposed)

We adapt the greedy *Objective* heuristic of [3] to row selection. Reusing the same normalized importance share p_i of Equation (3) as SL (so the two methods differ only in *how* they turn p_i into a schedule, not in what they consider important), EoA builds a schedule $j_1, j_2, \ldots, j_T$. Define the target cumulative count of row i after t steps as $t\,p_i$, and let $c_i(t)$ be the number of times row i appears among $j_1, \ldots, j_t$. The normalization $\sum_i p_i = 1$ makes this target meaningful: it guarantees $\sum_i t\,p_i = t$, exactly the number of steps, so the per-row deficit $t\,p_i - c_i(t-1)$ measures under-service against a budget the counts must eventually meet. At step t, EoA chooses

$$j_t = \arg\max_i \Big(t\,p_i - c_i(t-1) \Big), \tag{4}$$

i.e., the row with the largest *deficit* between its target and scheduled count so far, ties being broken by the smallest index. Unlike SL, this rule requires neither sampling a virtual electorate nor computing a divisor sequence; it maintains only one running counter per row, making it simpler to implement than SL. The row-projection variant (`eoa`) visits $j_1, \ldots, j_T$ as they are. In the coordinate-descent variant (`rka_eoa`), the schedule is computed once from Equation (4), and for each column k its first entry is replaced by k (warm start) while the remaining entries are kept, as stated in Algorithm 1.

Algorithm 1 rKA with EoA Row Selection (coordinate-descent form)

Require: Matrix $A \in \mathbb{C}^{K \times K}$, target column e_k, total iterations T
Ensure: Approximate solution $z^{(T)}$ to $Az = e_k$

```
Compute p̃_i = 1 − A_ii/tr(A) and normalize p_i = p̃_i / Σ_j p̃_j, i = 1, ..., K      ▷ Σ_i p_i = 1
c_i ← 0 for all i
for t = 1 to T do                                                                    ▷ EoA schedule
    j_t ← arg max_i (t p_i − c_i);  c_{j_t} ← c_{j_t} + 1
end for
i_1 ← k (warm start);  i_t ← j_t for t = 2, ..., T
z^(0) ← 0
for t = 1 to T do                                                                    ▷ Gauss–Seidel updates
    η ← (e_k[i_t] − (A z^(t−1))_{i_t}) / A_{i_t, i_t}
    z^(t) ← z^(t−1) + η e_{i_t}
end for
return z^(T)
```

The following proposition states the deficit-minimization property of the schedule in Equation (4).

Proposition 1. *Let $K > 1$, $p_i > 0$ and $\sum_i p_i = 1$, and let $c(t) = (c_1(t), \ldots, c_K(t))$ be the counts of the EoA schedule in Equation* (4)*. (i) Given $c(t-1)$,*

the choice j_t minimizes $\|c(t) - t\,p\|_1$ over all possible choices of the t-th row. (ii) For all $t \geq 0$ and all i, $t\,p_i - c_i(t) > -1$; consequently $\|c(t) - t\,p\|_1 < 2(K-1)$ for every t. (iii) In the coordinate-descent form, the counts of the sequence $i_1, \ldots, i_t$ actually used differ from $c(t)$ by at most one in at most two rows.

Proof. (i) Let $d_i = t\,p_i - c_i(t-1)$. Adding row j changes $\|c - t\,p\|_1$ by $g(d_j) = |d_j - 1| - |d_j|$, which is non-increasing in d_j; hence the row with the largest d_j minimizes it. (ii) The deficits d_i sum to $t-(t-1) = 1$, so the largest is at least $1/K$ and, after its count is incremented, at least $1/K - 1 > -1$; all other deficits increase by $p_i > 0$, and all deficits start at 0. Since the deficits $t\,p_i - c_i(t)$ sum to zero, the positive ones sum to the absolute sum of the negative ones, of which there are at most $K - 1$, each greater than -1. (iii) Replacing j_1 by k increases the count of row k by one and decreases that of row j_1 by one, if they differ. □

Proposition 1 concerns the schedule only: it says nothing about the resulting BER or convergence rate, and whether EoA's schedule leads to faster convergence is examined only by the Monte Carlo experiments of Sections 4 and 5.

4 Experiments

4.1 Experimental Setup

Unless otherwise noted, experiments use i.i.d. Rayleigh flat-fading channels with perfect CSI, QPSK modulation, and per-user slicer detection; each Monte Carlo trial draws one channel realization and one symbol per user. The simulator was implemented for this study following the procedures of the prior study, [1]. Each point of a curve versus the iteration count is obtained from a separate run with that iteration budget T, not by observing a single run part-way. BERs are plotted with a lower limit of 10^{-6}, and BERs of order 10^{-6} rest on only a few bit errors. Six row-selection variants are compared: plain row-projection rKA (`rk`, Strohmer–Vershynin norm-proportional selection), plain coordinate-descent rKA with uniform selection (`rka`), SL in both formulations (`sl`, `rka_sl`), and the proposed EoA rule in both (`eoa`, `rka_eoa`). Table 1 lists the configuration of each experiment in Sections 4 and 5. In its two rows for Section 4.4, the SNRs -13, -10, and -7 dB for $K = 2, 4, 8$ are the reference SNR $-10 - 10\log_{10}(M/K) + 10\log_{10} 16$ dB, and "ref." denotes a sweep from 15 dB below to 10 dB above this reference in 5 dB steps. The timing of Section 4.6 processes the channels as one batch, and Section 4.7 uses 64-QAM and CSI error $\tau \in \{0, 0.3\}$. For the condition-number measurement of Section 5.3, the last column gives the number of channel realizations, and no BER is computed. Further seeds used only in the checks of Sections 4.2 and 5.2 are not counted in the table.

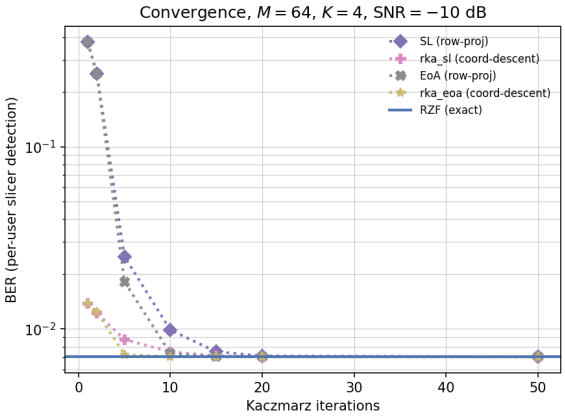


Fig. 1: Convergence (BER vs. iteration count) at $M = 64$, $K = 4$, SNR $= -10$ dB.

Source: created by the author.

Table 1. Configuration of each experiment (i.i.d. Rayleigh channel, QPSK and perfect CSI unless noted). "Per ratio": the SNR is chosen for each M/K as described in Section 5.2.

Exp.	M	K	SNR [dB]	T	Trials×seeds
Figure 1	64	4	-10	1–50	$10^5 \times 1$
Figure 2	64	4	-25 to 0	10	$10^5 \times 1$
Figure 3	64	2, 4, 8	$-13, -10, -7$	1–20	$5{\cdot}10^4 \times 1$
Figure 4	64	2, 4, 8	ref. -15 to $+10$	10	$5{\cdot}10^4 \times 1$
Figure 5	64	8	-7	2–20	$5{\cdot}10^4 \times 2$
Sec. 4.6	64	4	-10	20	$5{\cdot}10^4$
Sec. 4.7	64	4, 8	10, 12	10, 20, 50	$5{\cdot}10^4 \times 1$
Figure 6	16	16	10	2–200	$5{\cdot}10^4 \times 1$
Table 2	8–32	8	per ratio	20, 200	$2{\cdot}10^4 \times 5$
Figure 7	16–128	16	per ratio	20, 200	$2{\cdot}10^4 \times 2$
Figure 8	8–256	8, 16, 32	–	–	$2{\cdot}10^4$

Source: created by the author.

4.2 Convergence at $M = 64$, $K = 4$

At a fixed operating point ($M = 64$, $K = 4$, SNR $= -10$ dB), EoA converges faster than SL from 5 iterations on, the range that matters most for latency-constrained systems (Figure 1). At 5 iterations, `sl` and `rka_sl` achieve BER of 2.5×10^{-2} and 8.8×10^{-3}, while `eoa` and `rka_eoa` achieve 1.8×10^{-2} and 7.2×10^{-3}; the two families reach essentially the exact-RZF BER by 20 iterations. Three further seeds give 2.46–2.50×10^{-2}, 8.6–8.9×10^{-3}, 1.81–1.85×10^{-2}, and 6.9–7.3×10^{-3} at 5 iterations, respectively, with the same ordering.

4.3 Error-Floor Behavior at Low Iteration Counts

The difference is most visible when a fixed, small iteration budget is evaluated across a range of SNR (Figure 2). At 10 iterations and SNR $= -5$ dB, the coordinate-descent variants attain BER of 3.0×10^{-5}

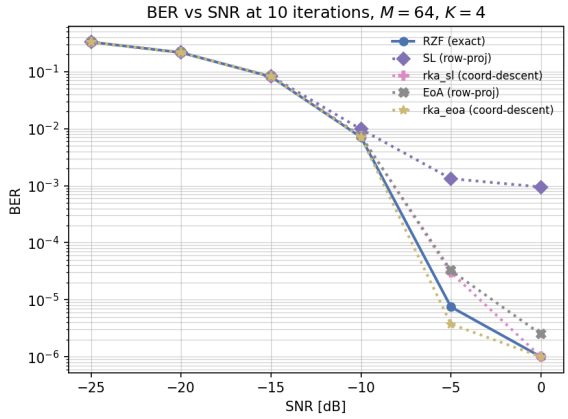


Fig. 2: BER vs. SNR at a fixed budget of 10 iterations, $M = 64$, $K = 4$.

Source: created by the author.

(`rka_sl`) versus 3.8×10^{-6} (`rka_eoa`), the latter comparable to the exact RZF solution (7.5×10^{-6}), while the row-projection variants attain 1.3×10^{-3} (`sl`) versus 3.3×10^{-5} (`eoa`). The BERs of order 10^{-6} rest on only a few bit errors and are indicative only. Ten iterations of `rka_eoa` thus reach the BER level of exact RZF, which ten iterations of `rka_sl` do not.

4.4 Scaling with the Number of Users

Repeating the comparison at $M = 64$ with $K \in \{2, 4, 8\}$ (SNR shifted to compensate for the change in array gain $10 \log_{10}(M/K)$) shows that EoA's advantage over SL grows as K increases, i.e., as M/K shrinks from 32 to 8: at $K = 8$ and 10 iterations, SNR $= -7$ dB, `rka_eoa` achieves BER of 6.0×10^{-4} against `rka_sl`'s 2.35×10^{-3}, a widening gap relative to $K = 2$ and $K = 4$ (Figure 3 and Figure 4).

4.5 Comparison Against Uniform (Non-Apportioned) Row Selection

To see how much each weighted rule improves on unweighted selection with the same update form and warm start, we compare against plain rKA (Figure 5). At $M = 64$, $K = 8$, 20 iterations: plain row-projection rKA (`rk`) attains BER of 5.9×10^{-2}, `sl` improves this to 1.5×10^{-2}, and `eoa` further to 3.1×10^{-3}. In the coordinate-descent formulation, however, plain uniform selection (`rka`) already attains 9.3×10^{-4}, close to `rka_sl`'s 8.1×10^{-4}; only `rka_eoa` (3.9×10^{-4}, matching the exact solution's 3.9×10^{-4}) improves substantially on uniform selection. In this comparison, SL's schedule thus adds little to what the coordinate-descent form with warm start already provides, whereas EoA's schedule adds a substantially larger improvement. At the smallest budgets ($T \leq 5$, fewer than $K = 8$ iterations), however, `rka` and `rka_sl` attain lower BER than `rka_eoa`.

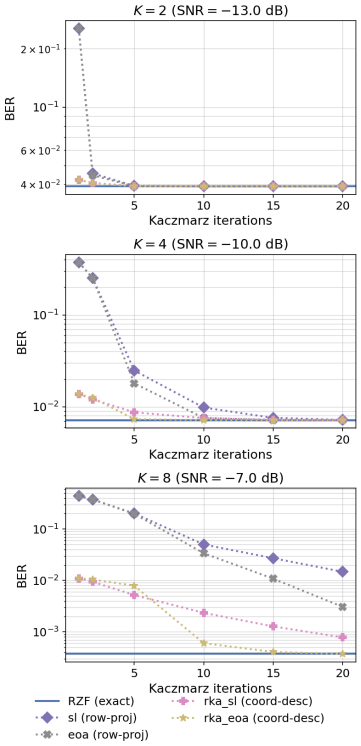


Fig. 3: Convergence for $K \in \{2, 4, 8\}$, $M = 64$.

Source: created by the author.

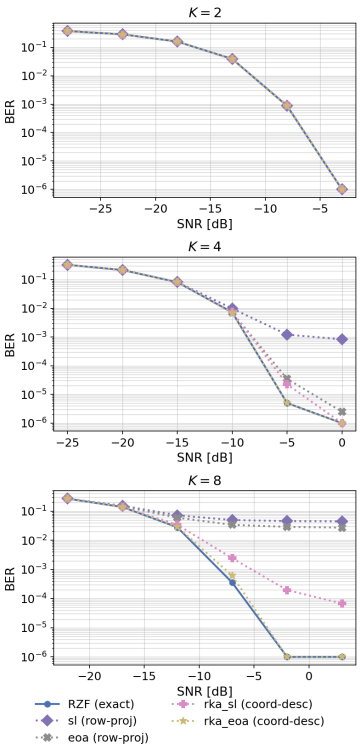


Fig. 4: BER vs. SNR at 10 iterations for $K \in \{2, 4, 8\}$, $M = 64$.

Source: created by the author.

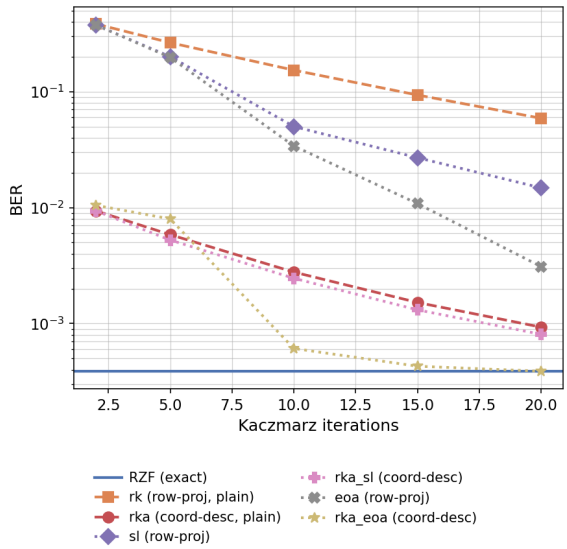


Fig. 5: All six row-selection variants at $M = 64$, $K = 8$, SNR $= -7$ dB.

Source: created by the author.

4.6 Computational Overhead

We measured the computation time per precoding call at $M = 64$, $K = 4$, 20 iterations by processing a batch of 5×10^4 channels at once in our NumPy/OpenBLAS implementation and dividing by the batch size; for every method, one warm-up run is followed by three timed runs, of which the minimum is reported. The measurements were made on two logical cores of an Intel Core i9-13980HX in a Linux virtual machine (Python 3.10, NumPy 2.2), and repeated runs differed by about 5%. The values are therefore average throughput figures, not the latency of a single call. Direct matrix inversion takes $4.2\,\mu$s; `rk` and `rka` take 9.0–9.1 μs, the SL variants 11.4–11.8 μs, and the EoA variants 9.8–11.3 μs, so all iterative variants remain within the same order of magnitude and EoA is slightly faster than SL in this implementation. Such software timings do not indicate hardware complexity.

4.7 Robustness to 64-QAM and Imperfect CSI

Repeating the $K = 4$ and $K = 8$ comparisons with 64-QAM modulation and a Gauss–Markov channel-estimation-error model ($\tau = 0.3$, matching the CSI-error parameter of the author's prior work, [1]) preserves the ordering above in the tested conditions. Under 64-QAM with perfect CSI at $M = 64$, $K = 8$ (SNR $= 12$ dB), `rka_eoa` reaches BER 3.8×10^{-6} (a handful of bit errors; indicative only) at 20 iterations while `rka_sl` remains at 1.7×10^{-2}. Under imperfect CSI ($\tau = 0.3$), where the exact-RZF BER is itself floored by residual interference at 1.8×10^{-2} ($K = 8$), `rka_eoa` attains 1.9×10^{-2} at 20 iterations — essentially the exact-precoder value — versus `rka_sl`'s 3.9×10^{-2}; at $K = 4$ (exact-RZF BER 3.1×10^{-3}) `rka_eoa` reaches 3.1×10^{-3} already at 10 iterations against `rka_sl`'s 9.9×10^{-3}. The advantage of `rka_eoa` over `rka_sl` is therefore also observed under the tested 64-QAM and imperfect-CSI conditions.

5 The $M = K$ Boundary Case

5.1 Motivation

The author's prior study, [1], evaluated SL with per-subarray blocks of $M_j = K_j = 16$, i.e., with a *square* channel matrix and no array gain from antenna oversampling, unlike the $M \gg K$ regime of Sections 4.1–4.7. This motivates examining how the row-selection rules behave as $M/K \to 1$ under the present model. Because the channel models differ, the results of this section are not compared with those of the prior study, [1].

5.2 Experimental Characterization

Figure 6 shows all six variants at $M = K = 16$, SNR $= 10$ dB. The horizontal axis is the iteration budget T on a linear scale (each point a separate run, Section 4.1), and the horizontal line is the BER of exact RZF, which does not depend on T. The three row-projection variants (`rk`, `sl`, `eoa`) are practically indistinguishable and decrease only from about 0.44 to 0.16–0.17 over 200 iterations. The coordinate-descent variants start lower, owing to the warm start, and keep decreasing: at 200 iterations `rka` reaches 3.3×10^{-2}, `rka_eoa` 2.6×10^{-2}, and `rka_sl` 1.9×10^{-2}, all still an order of magnitude above exact RZF (1.5×10^{-3}). None of the variants reaches the exact-RZF BER within 200 iterations, mainly because A is badly conditioned here: its median condition number is about 370, compared with about 1.7 at $M = 64$, $K = 4$, and the convergence of both the row-projection, [4], and the coordinate-descent iterations slows as the condition number grows. (In auxiliary runs at SNR $= 5$ and 15 dB the iterative BERs were similar, indicating that they are limited by residual interference rather than by noise.) The ordering of the coordinate-descent variants depends on the budget: at 50 iterations `rka_eoa` (6.8×10^{-2}) is better than `rka_sl` (8.2×10^{-2}), at 100 iterations the two are close (4.4×10^{-2} and 4.3×10^{-2}), and at 200 iterations `rka_sl` is better, the opposite of the Section 4 ordering.

To locate this reversal, we sweep fourteen M/K values between 1 and 8 (still $K = 16$ fixed) for the coordinate-descent family (`rka`, `rka_sl`, `rka_eoa`), averaging two random seeds of 2×10^4 channel realizations per point (Figure 7); the left and right panels show the budgets $T = 20$ and $T = 200$. Because the exact-RZF BER varies strongly with M/K, the SNR is chosen per ratio:

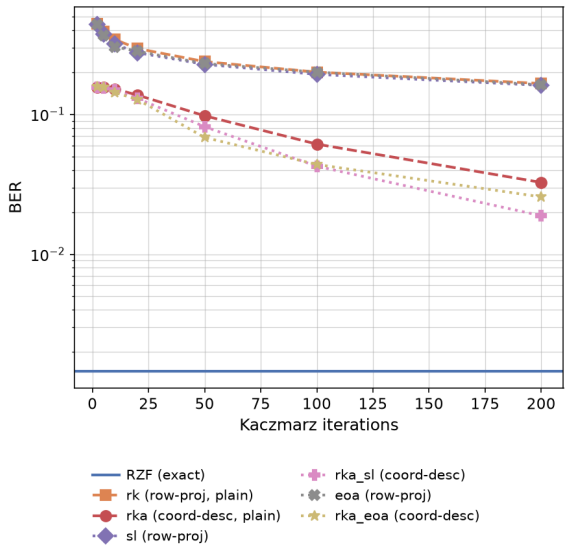


Fig. 6: Convergence at $M = K = 16$, SNR $= 10$ dB; the horizontal line is exact RZF.

Source: created by the author.

from -15 to 35 dB in 5 dB steps, the value at which exact RZF comes closest to a BER of 10^{-3} in a preliminary run (10^4 trials, 20 iterations). The chosen SNR decreases from 10 dB at $M/K = 1$ to -10 dB for $M/K \geq 4$; because of the coarse grid, the realized exact-RZF BER ranges from about 2×10^{-4} to 10^{-2}, which causes the zigzag of the exact-RZF curve. Absolute BER is therefore not comparable across M/K, only the gap between variants at a common ratio. Two observations follow. First, the reversal is not confined to $M/K = 1$: at 200 iterations `rka_sl` beats `rka_eoa` for every ratio in $\{1, 1.0625, 1.125, 1.1875\}$, and EoA regains its advantage from $M/K = 1.25$ onward, where `rka_sl` and `rka_eoa` come within about 15% of the exact-RZF BER. Re-running $M/K \leq 1.5$ with two further seeds confirmed both the reversal and its end point (paired t-statistics over four seeds: $|t| \geq 17$ for $M/K \leq 1.1875$ and $t \geq 5.5$ for $1.25 \leq M/K \leq 1.5$). Note, however, that the chosen SNR also changes between $M/K = 1.1875$ (5 dB) and 1.25 (0 dB), so the location of the end point reflects both M/K and the SNR selection and should not be read as a universal threshold. Second, the reversal is absent at smaller budgets: at 20 iterations EoA is at least as good as SL across the *entire* sweep, including $M/K = 1$. The unfavorable regime is therefore not "$M = K$" as such, but the conjunction of a narrow band $M/K \lesssim 1.2$ *and* a large iteration budget. A smaller auxiliary sweep at $K = 8$ (Table 2) shows the same pattern. Its ratios $M/K \in \{1, 1.25, 1.5, 2, 4\}$ ($M = 8, 10, 12, 16, 32$) are a subset of the $K = 16$ grid chosen to cover both the reversal band and the regular regime, with the SNR chosen in the same way. At 20 iterations `rka_eoa` is better than

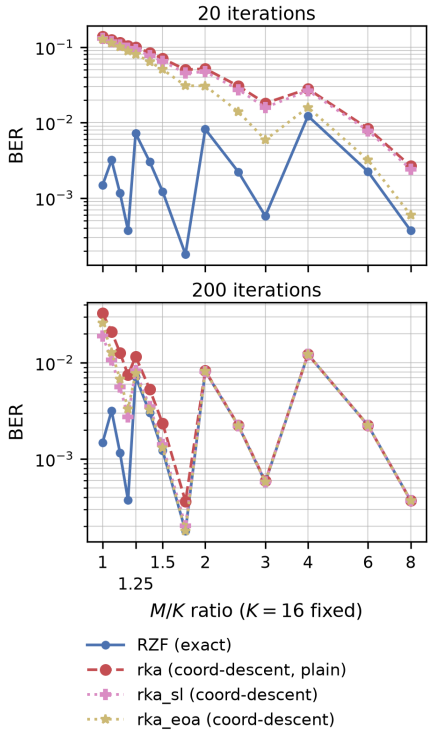


Fig. 7: BER vs. M/K ($K = 16$) at 20 and 200 iterations, with the SNR set per ratio (Section 5.2).

Source: created by the author.

`rka_sl` at every ratio, whereas at 200 iterations `rka_sl` is better only at $M/K = 1$, the two being statistically indistinguishable (both essentially converged) for $M/K \geq 1.25$.

Table 2. Auxiliary $K = 8$ sweep: mean BER over five seeds of 2×10^4 trials, with the SNR [dB] set per ratio (Section 5.2). *: significantly lower of the pair (paired t-test, $p < 0.01$). Seed-to-seed standard deviation: below 2% (20 iterations) and 15% (200 iterations) of the mean.

		20 iterations		200 iterations	
M/K	SNR	`rka_sl`	`rka_eoa`	`rka_sl`	`rka_eoa`
1.00	15	9.5×10^{-2}	9.2×10^{-2} *	1.2×10^{-2} *	1.4×10^{-2}
1.25	5	5.2×10^{-2}	4.4×10^{-2} *	2.8×10^{-3}	2.7×10^{-3}
1.50	5	2.9×10^{-2}	2.1×10^{-2} *	2.5×10^{-4}	2.5×10^{-4}
2.00	0	1.5×10^{-2}	7.0×10^{-3} *	1.8×10^{-3}	1.8×10^{-3}
4.00	-5	5.6×10^{-3}	2.9×10^{-3} *	2.6×10^{-3}	2.6×10^{-3}

Source: created by the author.

5.3 Discussion

To characterize this band, we measured the condition number of the $K \times K$ Gram matrix HH^H for M/K from 1 to 8 and $K \in \{8, 16, 32\}$, using 2×10^4 channel realizations per point (Figure 8; $M = \text{round}(K \cdot M/K)$, with one point per distinct M). The

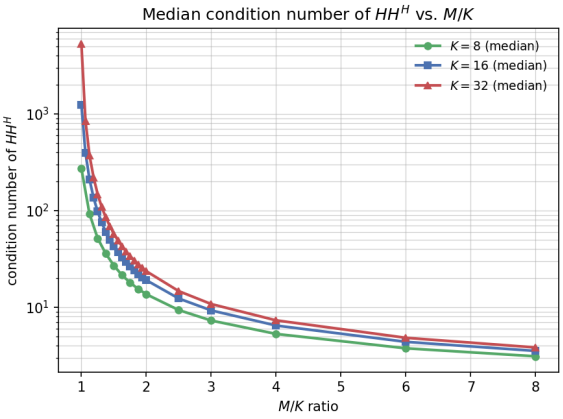


Fig. 8: Median condition number of HH^H vs. M/K for $K \in \{8, 16, 32\}$.

Source: created by the author.

median condition number rises steeply as $M/K \to 1$ — e.g., for $K = 16$, from 3.5 at $M/K = 8$ to 19 at $M/K = 2$ to 394 at $M/K = 1.0625$ to 1231 at $M/K = 1$ — with the same qualitative shape for all three K. This is consistent with the asymptotic Marchenko–Pastur description, under which the smallest eigenvalue of HH^H approaches zero as $M/K \to 1$; our finite-K measurements do not establish more than this consistency. The reversal band above ($M/K \lesssim 1.2$) coincides with the region where this condition number exceeds roughly 10^2 for $K = 16$. It measures a property of the channel only: the condition number of HH^H, not of the regularized matrix $A = HH^H + \alpha I_K$ actually solved, whose condition number $(\lambda_{\max} + \alpha)/(\lambda_{\min} + \alpha)$ is smaller (e.g., a median of about 370 instead of 1231 at $M = K = 16$, 10 dB), and not of any row-selection schedule. The coincidence of the reversal band with the steep region is therefore an empirical correlation, not a causal explanation; we do not derive a convergence-rate bound linking the condition number to the EoA/SL ordering, and leave a precise characterization of the crossover to future work (Section 7).

6 Discussion and Limitations

The only analytical result, Proposition 1, concerns the EoA schedule itself; it does not guarantee a lower BER, a faster convergence rate, or a particular crossover threshold. All performance statements in Sections 4 and 5 are Monte Carlo observations for i.i.d. Rayleigh channels, with the SNR selection of Section 5.2 near the boundary. The author's prior study, [1], used a subarray-based non-stationary channel with imperfect CSI, so its results are not directly comparable with the present ones; all comparisons in this paper use SL re-implemented under identical conditions.

7 Conclusion and Future Work

This paper introduced Equity-of-Attention-based row selection for randomized Kaczmarz precoding. Its schedule minimizes the ℓ_1 deviation from the target allocation one step at a time (Proposition 1), whereas the published SL schedules carry no guarantee on the counts of a truncated schedule. In simulations over i.i.d. Rayleigh channels, EoA converged faster than the published SL for $M/K \gtrsim 1.25$ apart from the smallest iteration budgets, and SL was preferable only inside a narrow band $M/K \lesssim 1.2$ at large budgets. Future work includes a condition-number-based account of this reversal, an adaptive EoA/SL switching rule, subarray-level spatial correlation and non-stationarity, residual-based stopping criteria, and further modulation orders and user counts.

Declaration of Generative AI and AI-assisted Technologies in the Writing Process

During the preparation of this work the author used Grammarly for language editing. After using this service, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

References:

[1] Tatsuki Fukuda. Fairness-based row selection in randomized Kaczmarz precoding for massive MIMO systems. *WSEAS Trans. Comput. Res.*, vol. 13, pp. 634–643, 2025, https://doi.org/10.37394/232018.2025.13.56

[2] Tatsuki Fukuda. Randomized Kaczmarz algorithm applied D'Hondt method for extremely massive MIMO wireless communication systems. *WSEAS Trans. Comput. Res.*, vol. 12, pp. 494–502, 2024, https://doi.org/10.37394/232018.2024.12.48

[3] Asia J. Biega, Krishna P. Gummadi, and Gerhard Weikum. Equity of attention: Amortizing individual fairness in rankings. In *Proc. 41st Int. ACM SIGIR Conf. Res. Develop. Inf. Retr. (SIGIR)*, Ann Arbor, MI, USA, 2018, pp. 405–414, https://doi.org/10.1145/3209978.3210063

[4] Thomas Strohmer and Roman Vershynin. A randomized Kaczmarz algorithm with exponential convergence. *J. Fourier Anal. Appl.*, vol. 15, no. 2, pp. 262–278, 2009, https://doi.org/10.1007/s00041-008-9030-4

[5] Zhong-Zhi Bai and Wen-Ting Wu. On greedy randomized augmented Kaczmarz method for solving large sparse inconsistent linear systems. *SIAM J. Sci. Comput.*, vol. 43, no. 6, pp. A3892–A3911, 2021, https://doi.org/10.1137/20M1352235

[6] Yutong Jiang, Gang Wu, and Long Jiang. A semi-randomized Kaczmarz method with simple random sampling for large-scale linear systems. *Adv. Comput. Math.*, vol. 49, no. 2, Art. no. 20, 2023, https://doi.org/10.1007/s10444-023-10018-2

[7] Gang Wu and Qiao Chang. A semi-randomized block Kaczmarz method with simple random sampling for large-scale consistent linear systems. *Commun. Appl. Math. Comput.*, vol. 7, no. 5, pp. 2120–2136, 2025, https://doi.org/10.1007/s42967-024-00447-1

[8] Cun-Qiang Miao and Wen-Ting Wu. On greedy randomized average block Kaczmarz method for solving large linear systems. *J. Comput. Appl. Math.*, vol. 413, Art. no. 114372, 2022, https://doi.org/10.1016/j.cam.2022.114372

[9] Yansheng Su, Deren Han, Yun Zeng, and Jiaxin Xie. On greedy multi-step inertial randomized Kaczmarz method for solving linear systems. *Calcolo*, vol. 61, no. 4, Art. no. 68, 2024, https://doi.org/10.1007/s10092-024-00621-0

[10] Yotam Yaniv, Jacob D. Moorman, William Swartworth, Thomas Tu, Daji Landis, and Deanna Needell. Selectable set randomized Kaczmarz. *Numer. Linear Algebra Appl.*, vol. 30, no. 1, Art. no. e2458, 2023, https://doi.org/10.1002/nla.2458

[11] Yanjun Zhang and Hanyu Li. Greedy Motzkin–Kaczmarz methods for solving linear systems. *Numer. Linear Algebra Appl.*, vol. 29, no. 4, Art. no. e2429, 2022, https://doi.org/10.1002/nla.2429

[12] Zhong-Zhi Bai and Lu Wang. On convergence rates of Kaczmarz-type methods with different selection rules of working rows. *Appl. Numer. Math.*, vol. 186, pp. 289–319, 2023, https://doi.org/10.1016/j.apnum.2023.01.013

[13] Julian Chingoma, Ulle Endriss, Ronald de Haan, Adrian Haret, and Jan Maly. Apportionment with weighted seats. In *Proc. 28th Eur. Conf. Artif. Intell. (ECAI)*, Bologna, Italy, 2025, Frontiers in Artificial Intelligence and Applications, IOS Press, pp. 3831–3838, https://doi.org/10.3233/FAIA251265

[14] Michael A. Jones, Brittany Ohlinger, and Jennifer Wilson. Sequential apportionment from stationary divisor methods. arXiv:2512.20686, 2025, https://doi.org/10.48550/arXiv.2512.20686

[15] Bokai Xu, Zhe Wang, Huahua Xiao, Jiayi Zhang, Bo Ai, and Derrick Wing Kwan Ng. Low-complexity precoding for extremely large-scale MIMO over non-stationary channels. In *Proc. IEEE Int. Conf. Commun. (ICC)*, Rome, Italy, 2023, pp. 6516–6522, https://doi.org/10.1109/ICC45041.2023.10279327

[16] Mengyu Liu, Cunhua Pan, Kangda Zhi, Hong Ren, Cheng-Xiang Wang, Jiangzhou Wang, and Yonina C. Eldar. Low-complexity iterative precoding design for near-field multiuser systems with spatial non-stationarity. *IEEE Trans. Signal Process.*, vol. 74, pp. 1309–1324, 2026, https://doi.org/10.1109/TSP.2026.3672691

[17] Zhe Wang, Jiayi Zhang, Hongyang Du, Dusit Niyato, Shuguang Cui, Bo Ai, Mérouane Debbah, Khaled B. Letaief, and H. Vincent Poor. A tutorial on extremely large-scale MIMO for 6G: Fundamentals, signal processing, and applications. *IEEE Commun. Surveys Tuts.*, vol. 26, no. 3, pp. 1560–1605, 2024, https://doi.org/10.1109/COMST.2023.3349276

[18] Kangda Zhi, Cunhua Pan, Hong Ren, Kok Keong Chai, Cheng-Xiang Wang, Robert Schober, and Xiaohu You. Performance analysis and low-complexity design for XL-MIMO with near-field spatial non-stationarities. *IEEE J. Sel. Areas Commun.*, vol. 42, no. 6, pp. 1656–1672, 2024, https://doi.org/10.1109/JSAC.2024.3389128

[19] Yuhao Chen and Linglong Dai. Non-stationary channel estimation for extremely large-scale MIMO. *IEEE Trans. Wireless Commun.*, vol. 23, no. 7, pp. 7683–7697, 2024, https://doi.org/10.1109/TWC.2023.3343740

[20] Ashudeep Singh and Thorsten Joachims. Fairness of exposure in rankings. In *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, London, United Kingdom, 2018, pp. 2219–2228, https://doi.org/10.1145/3219819.3220088

[21] Tao Yang, Zhichao Xu, and Qingyao Ai. Vertical allocation-based fair exposure amortizing in ranking. In *Proc. Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr. Asia Pacific Region (SIGIR-AP)*, Beijing, China, 2023, pp. 234–244, https://doi.org/10.1145/3624918.3625313

[22] Yuta Saito and Thorsten Joachims. Fair ranking as fair division: Impact-based individual fairness in ranking. In *Proc. 28th ACM SIGKDD Conf. Knowl. Discovery Data Mining (KDD)*, Washington, DC, USA, 2022, pp. 1514–1524, https://doi.org/10.1145/3534678.3539353

[23] Yiteng Tu, Weihang Su, Shuguang Han, Yiqun Liu, and Qingyao Ai. Equity vs. equality: Optimizing ranking fairness for tailored provider needs. In *Proc. 49th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr. (SIGIR)*, Melbourne, VIC, Australia, 2026, pp. 1767–1777, https://doi.org/10.1145/3805712.3809647

[24] Xuancheng Li, Tao Yang, Yujia Zhou, Qingyao Ai, and Yiqun Liu. Beyond exposure: Optimizing ranking fairness with non-linear time-income functions. arXiv:2602.03345, 2026, https://doi.org/10.48550/arXiv.2602.03345

[25] Yafei Wang, Xinping Yi, Hongwei Hou, Wenjin Wang, and Shi Jin. Robust symbol-level precoding for massive MIMO communication under channel aging. *IEEE Trans. Wireless Commun.*, vol. 23, no. 9, pp. 10864–10878, 2024, https://doi.org/10.1109/TWC.2024.3376796

[26] Zijiao Guo, Vaskar Sen, and Honggui Deng. Robust low-complexity WMMSE precoding under imperfect CSI with per-antenna power constraints. *Sensors*, vol. 26, no. 1, Art. no. 159, 2026, https://doi.org/10.3390/s26010159

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The author solely contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflicts of Interest

The author has no conflicts of interest to declare that are relevant to the content of this article.

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)

This article is published under the terms of the Creative Commons Attribution License 4.0. https://creativecommons.org/licenses/by/4.0/deed.en_US