Linear Algebra and its Applications 698 (2024) 220–260

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

A subspace constrained randomized Kaczmarz method for structure or external knowledge exploitation

Jackie Lok *, Elizaveta Rebrova

Department of Operations Research and Financial Engineering, Princeton University, Princeton, NJ 08544, USA

ARTICLE INFO

Article history:
Received 11 September 2023
Received in revised form 5 June 2024
Accepted 7 June 2024
Available online 12 June 2024
Submitted by I.C.F. Ipsen

MSC:
65F10
65F20
60B20
68W20

Keywords:
Kaczmarz algorithm
Stochastic iterative methods
Quantile methods
Corrupted linear systems

ABSTRACT

We study a version of the randomized Kaczmarz algorithm for solving systems of linear equations where the iterates are confined to the solution space of a selected subsystem. We show that the subspace constraint leads to an accelerated convergence rate, especially when the system has approximately low-rank structure. On Gaussian-like random data, we show that it results in a form of dimension reduction that effectively increases the aspect ratio of the system. Furthermore, this method serves as a building block for a second, quantile-based algorithm for solving linear systems with arbitrary sparse corruptions, which is able to efficiently utilize external knowledge about corruption-free equations and achieve convergence in difficult settings. Numerical experiments on synthetic and realistic data support our theoretical results and demonstrate the validity of the proposed methods for even more general data models than guaranteed by the theory.

© 2024 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

* Corresponding author.
E-mail addresses: jackie.lok@princeton.edu (J. Lok), elre@princeton.edu (E. Rebrova).

https://doi.org/10.1016/j.laa.2024.06.010
0024-3795/© 2024 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

1. Introduction

A ubiquitous problem across the sciences is solving large-scale systems of linear equations $\mathbf{A}\mathbf{x} = \mathbf{b}$, for which scalable and efficient iterative methods are useful when it is too slow or infeasible to solve the system directly. Instead of solving such problems obliviously, it is natural to have insights into the structural properties of the linear system of interest, such as being approximately low-rank. Moreover, external knowledge about trustworthy observations in the presence of corrupted measurements could be available.

Now, additional information about the structure of a linear system influences the choice of the most suitable solver. In this work, we take an adaptive, problem-aware approach to account for various types of auxiliary information by augmenting the iterations of a generic iterative linear solver based on a distinguished subsystem of equations.

The generic solver that we consider is the *Kaczmarz algorithm* [32], which is an iterative, row-action method for solving large-scale, typically overdetermined systems of linear equations. It is a special case of the alternating projection method that has low computational cost and storage per iteration, and can be used in the streaming setting where a single row (or block of rows) of the system can be accessed at a time. Besides its traditional applications in areas such as image reconstruction [41,28] and signal processing [8], the Kaczmarz algorithm has recently been used as a building block for more sophisticated methods to design linear solvers [14,19], and to address problems such as phase retrieval [59] and tensor recovery [9].

In each iteration of the Kaczmarz algorithm, a row $\mathbf{a}_j$ of the matrix $\mathbf{A}$ is selected, and the current iterate $\mathbf{x}^k$ is projected onto the hyperplane $\mathbf{a}_j^\mathsf{T}\mathbf{x} = b_j$ by

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \frac{b_j - \mathbf{a}_j^\mathsf{T}\mathbf{x}^k}{\|\mathbf{a}_j\|} \cdot \frac{\mathbf{a}_j}{\|\mathbf{a}_j\|}. \tag{1.1}$$

In their seminal paper, Strohmer and Vershynin [57] show that if the system of linear equations is consistent and has unique solution $\mathbf{x}^*$, then the *randomized Kaczmarz (RK) algorithm*, which samples each row independently with probability $\|\mathbf{a}_j\|^2/\|\mathbf{A}\|_F^2$ at each iteration, converges to $\mathbf{x}^*$ in expectation with an exponential rate (i.e. linearly) that depends on the geometric properties of $\mathbf{A}$ (or more precisely, its scaled condition number $\|\mathbf{A}\|_F/\sigma_{\min}(\mathbf{A})$):

$$\mathbb{E}\|\mathbf{x}^k - \mathbf{x}^*\|^2 \leq \left(1 - \frac{\sigma_{\min}(\mathbf{A})^2}{\|\mathbf{A}\|_F^2}\right)^k \cdot \|\mathbf{x}^0 - \mathbf{x}^*\|. \tag{1.2}$$

Subsequently, many variants of the randomized Kaczmarz method have been analyzed; we defer a detailed discussion of related works to Section 2 after presenting our results.

In this paper, we propose two Kaczmarz-based algorithms that can exploit (a) *approximately low-rank structure and geometric properties of the matrix in a system of linear equations to accelerate convergence*; and (b) *external knowledge about corruption-*

free equations for linear systems with arbitrary sparse corruptions to enable convergence even in the highly corrupted regime.

1.1. Setup and notation

We consider a consistent, overdetermined (i.e. tall) system of linear equations $\mathbf{A}\mathbf{x} = \mathbf{b}$, or *linear system* for short, where the rows of $\mathbf{A} \in \mathbb{R}^{m\times n}$ are denoted by $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, and $m \geq n$. We assume throughout that $\mathbf{A}$ has full rank, and denote the unique solution of the linear system by $\mathbf{x}^* \in \mathbb{R}^n$. We work in the real setting for simplicity, but everything can be generalized to the complex setting.

Vectors, oriented as columns by default, and matrices are written in boldface. The vector ℓ_2-norm is denoted by $\|\cdot\|$, and the matrix spectral and Frobenius norms are denoted by $\|\cdot\|$ and $\|\cdot\|_F$. The singular values of a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ are denoted by $\sigma_{\max}(\mathbf{A}) = \sigma_1(\mathbf{A}) \geq \sigma_2(\mathbf{A}) \geq \cdots \geq \sigma_{\min\{m,n\}}(\mathbf{A}) = \sigma_{\min}(\mathbf{A})$, and the smallest nonzero singular value is denoted by $\sigma^+_{\min}(\mathbf{A})$. The Moore-Penrose pseudoinverse of $\mathbf{A}$ is denoted by $\mathbf{A}^\dagger$. We refer to the row submatrix of $\mathbf{A}$ (resp. subvector of $\mathbf{b}$) indexed by $I \subseteq [m] := \{1, 2, \dots, m\}$ by $\mathbf{A}_I$ (resp. $\mathbf{b}_I$). The solution space $\mathbf{A}_I\mathbf{x} = \mathbf{b}_I$ of a linear system refers to the affine subspace $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}_I\mathbf{x} = \mathbf{b}_I\}$.

1.2. Methods and main results

1.2.1. The SCRK method

Fix a subset $I_0 \subset [m]$ of indices of rows of $\mathbf{A}$ with $m_0 := |I_0| < n$, and denote the remaining indices by $I_1 := [m] \setminus I_0$. We define a variant of the RK algorithm that confines the iterates within the solution space $\mathbf{A}_{I_0}\mathbf{x} = \mathbf{b}_{I_0}$, which we will refer to as the *subspace constrained randomized Kaczmarz (SCRK) method*. Each update of the SCRK algorithm consists of a projection of the current iterate $\mathbf{x}^k$ onto the solution space $\mathbf{A}_{I_0\cup\{j\}}\mathbf{x} = \mathbf{b}_{I_0\cup\{j\}}$, where the row corresponding to $j \in I_1$ is sampled according to an input probability distribution, and can be algebraically expressed by

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{A}^\dagger_{I_0\cup\{j\}}(\mathbf{b}_{I_0\cup\{j\}} - \mathbf{A}_{I_0\cup\{j\}}\mathbf{x}^k). \tag{1.3}$$

This is essentially a block Kaczmarz update [20,45], but with I_0 fixed throughout the iterations so that the iterates are confined within the selected solution space $\mathbf{A}_{I_0}\mathbf{x} = \mathbf{b}_{I_0}$. Reusing the same block allows for properties of the distinguished subsystem $\mathbf{A}_{I_0}\mathbf{x} = \mathbf{b}_{I_0}$ to be exploited, and also leads to a more efficient update formula: in Lemma 3.1, we prove that as long as $\mathbf{x}^k$ satisfies $\mathbf{A}_{I_0}\mathbf{x}^k = \mathbf{b}_{I_0}$ and $\mathbf{a}_j \notin \mathrm{Range}(\mathbf{A}^\mathsf{T}_{I_0})$, (1.3) simplifies to

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \frac{b_j - \mathbf{a}_j^\mathsf{T}\mathbf{x}^k}{\|\mathbf{P}\mathbf{a}_j\|} \cdot \frac{\mathbf{P}\mathbf{a}_j}{\|\mathbf{P}\mathbf{a}_j\|}, \tag{1.4}$$

where $\mathbf{P} := \mathbf{I} - \mathbf{A}^\dagger_{I_0}\mathbf{A}_{I_0}$ is the orthogonal projector onto $\mathrm{Null}(\mathbf{A}_{I_0}) = \mathrm{Range}(\mathbf{A}^\mathsf{T}_{I_0})^\perp$. Unlike the block update (1.3), this does not require a new pseudoinverse to be computed

at every iteration and thus can be performed faster. The SCRK method, which leverages (1.4), is summarized in Algorithm 1. For concreteness, we fix a particular sampling distribution for the rows of $\mathbf{A}_{I_1}$ that leads to an especially simple and interpretable analysis. By varying the distribution, better convergence rates may be possible [21,1].

On a conceptual level, the SCRK update (1.4) is reminiscent of the usual Kaczmarz update (1.1), with the new direction $\mathbf{P}\mathbf{a}_j$ representing the "extra information" offered by $\mathbf{a}_j$ beyond that which is already known from being in the solution space $\mathbf{A}_{I_0}\mathbf{x} = \mathbf{b}_{I_0}$.

Algorithm 1 Subspace Constrained Randomized Kaczmarz (SCRK).

procedure SCRK($\mathbf{A}, \mathbf{b}, I_0, K$)
  $\mathbf{P} = \mathbf{I} - \mathbf{A}_{I_0}^{\dagger}\mathbf{A}_{I_0}$ ▷ Orthogonal projector onto Null($\mathbf{A}_{I_0}$)
  initialize $\mathbf{x}^0 = \mathbf{A}_{I_0}^{\dagger}\mathbf{b}_{I_0}$ ▷ Initial iterate $\mathbf{x}^0$ solves $\mathbf{A}_{I_0}\mathbf{x}^0 = \mathbf{b}_{I_0}$
  for $k = 1, \dots, K$ **do**
    sample $j \in [m] \setminus I_0$ with prob. $\|\mathbf{P}\mathbf{a}_j\|^2 / \|\mathbf{A}_{I_1}\mathbf{P}\|_F^2$ ▷ Sample row in $\mathbf{A}_{I_1}$
    $\mathbf{x}^k = \mathbf{x}^{k-1} + \frac{b_j - \mathbf{a}_j^{\mathsf{T}}\mathbf{x}^{k-1}}{\|\mathbf{P}\mathbf{a}_j\|} \cdot \frac{\mathbf{P}\mathbf{a}_j}{\|\mathbf{P}\mathbf{a}_j\|}$ ▷ Project onto $\mathbf{A}_{I_0\cup\{j\}}\mathbf{x} = \mathbf{b}_{I_0\cup\{j\}}$
  end for
  return $\mathbf{x}^K$
end procedure

The following result, proved in Section 3.1, shows that the SCRK method converges linearly in expectation to the solution $\mathbf{x}^*$ of $\mathbf{A}\mathbf{x} = \mathbf{b}$ under minimal assumptions.

Theorem 1.1. *Suppose that the rows of* $\mathbf{A}$ *are partitioned into two blocks* $\mathbf{A}_{I_0}$ *and* $\mathbf{A}_{I_1}$ *of sizes* m_0 *and* $m - m_0$ *respectively. Let* $\mathbf{P} = \mathbf{I} - \mathbf{A}_{I_0}^{\dagger}\mathbf{A}_{I_0}$ *be the orthogonal projector onto* $\mathrm{Null}(\mathbf{A}_{I_0})$*, and* $\sigma_{\min}^{+}(\mathbf{A}_{I_1}\mathbf{P})$ *be the smallest non-zero singular value of* $\mathbf{A}_{I_1}\mathbf{P}$*. Then the SCRK iterates* $\mathbf{x}^k$ *from Algorithm 1 converge to the solution* $\mathbf{x}^*$ *in expectation with*

$$\mathbb{E}\|\mathbf{x}^k - \mathbf{x}^*\|^2 \le \left(1 - \frac{\sigma_{\min}^{+}(\mathbf{A}_{I_1}\mathbf{P})^2}{\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2}\right)^k \cdot \|\mathbf{x}^0 - \mathbf{x}^*\|^2. \tag{1.5}$$

In the special case that the row spaces of $\mathbf{A}_{I_0}$ and $\mathbf{A}_{I_1}$ are orthogonal (i.e. $\mathbf{A}_{I_0}\mathbf{A}_{I_1}^{\mathsf{T}} = \mathbf{0}$), the SCRK updates (1.4) reduce to the usual Kaczmarz updates (1.1) since $\mathbf{P}\mathbf{a}_j = \mathbf{a}_j$ for all $j \in I_1$, and hence we immediately deduce the following:

Corollary 1.2. *Consider the same setup as Theorem 1.1. If* $\mathbf{A}_{I_0}\mathbf{A}_{I_1}^{\mathsf{T}} = \mathbf{0}$*, then*

$$\mathbb{E}\|\mathbf{x}^k - \mathbf{x}^*\|^2 \le \left(1 - \frac{\sigma_{\min}^{+}(\mathbf{A}_{I_1})^2}{\|\mathbf{A}_{I_1}\|_F^2}\right)^k \cdot \|\mathbf{x}^0 - \mathbf{x}^*\|^2. \tag{1.6}$$

Noisy linear systems. In Section 3.2, we prove that for inconsistent systems of linear equations where a noisy measurement vector $\widehat{\mathbf{b}} \neq \mathbf{b}$ is observed, the SCRK method converges at the same rate up to an error horizon around the solution $\mathbf{x}^*$ with a radius that depends on the noise in I_0 and I_1, as well as the geometries of $\mathbf{A}_{I_0}$ and $\mathbf{A}_{I_1}\mathbf{P}$ (Theorem 3.4). This expands on a phenomenon that is known from previous analyses of Kaczmarz methods [44,45].

The analysis of inconsistent linear systems requires developing technical results involving a two-step decomposition of the block update (1.3), which provides a partial generalization of the two-subspace Kaczmarz method in [46] (see Remark 3.9).

Remark 1.3 *(Per-iteration complexity).* Each SCRK iteration can be computed in $O(m_0 n)$ flops, with calculating $\mathbf{P}\mathbf{a}_j = \mathbf{a}_j - \mathbf{A}_{I_0}^{\dagger}\mathbf{A}_{I_0}\mathbf{a}_j$ being the most expensive step. This requires directly computing $\mathbf{A}_{I_0}^{\dagger}$ or an orthonormal basis for Range$(\mathbf{A}_{I_0}^{\mathsf{T}})$[1] only once, using a method based on QR decomposition or SVD. This per-iteration cost is comparable to the $O(n)$ flops per iteration of RK if m_0 is not too large. The overall complexity is then determined by multiplying the per-iteration cost by the number of iterations required to reach a desired error, which we will elaborate upon below.

Exploiting low-rank structure with the SCRK method. The convergence rate of the SCRK algorithm depends on the geometric properties of $\mathbf{A}$ and $\mathbf{P}$: Theorem 1.1 shows that $k_\varepsilon := \kappa(\mathbf{A}_{I_1}\mathbf{P})^2 \log(1/\varepsilon)$ iterations suffice to achieve the relative error guarantee $\mathbb{E}\|\mathbf{x}^k - \mathbf{x}^*\|^2 \le \varepsilon\|\mathbf{x}^0 - \mathbf{x}^*\|^2$, where $\kappa(\mathbf{A}_{I_1}\mathbf{P}) := \|\mathbf{A}_{I_1}\mathbf{P}\|_F/\sigma_{\min}^{+}(\mathbf{A}_{I_1}\mathbf{P})$ is a scaled condition number of $\mathbf{A}_{I_1}\mathbf{P}$. For the same guarantee using RK, from (1.2), $\kappa(\mathbf{A})^2\log(1/\varepsilon)$ iterations are required, where $\kappa(\mathbf{A}) := \|\mathbf{A}\|_F/\sigma_{\min}(\mathbf{A})$.

Since $\kappa(\mathbf{A}_{I_1}\mathbf{P}) \le \kappa(\mathbf{A})$, we see that the projector $\mathbf{P}$ acts as a right preconditioner for $\mathbf{A}$, improving the convergence rate of SCRK compared to RK. In particular, we can expect a significant per-iteration improvement (and hence overall advantage) if $\|\mathbf{A}_{I_1}\mathbf{P}\|_F \ll \|\mathbf{A}\|_F$ or $\sigma_{\min}^{+}(\mathbf{A}_{I_1}\mathbf{P}) \gg \sigma_{\min}(\mathbf{A})$. We examine the connection between the geometry of $\mathbf{A}$ and convergence rates in more detail in Section 3.3.1 to show that the SCRK method is able to exploit approximately low-rank structure and geometric properties of $\mathbf{A}$ to accelerate convergence.

We also describe how a good subset I_0 of rows, if not explicitly known, can actually be *efficiently found* via a connection to low-rank matrix approximation in Section 3.3.2.

SCRK on random data and dimension reduction. In a somewhat complementary setting, we show that for "unstructured" matrices, the subspace constraint imposed by the projector $\mathbf{P}$ *acts as a form of dimension reduction that effectively increases the aspect ratio of the system to reflect the dimensionality of the solution that remains unsolved* in Section 3.4. More precisely, we prove that when $\mathbf{A} \in \mathbb{R}^{m\times n}$ is drawn from a generic class of "Gaussian-like" random matrices, the SCRK method typically converges with a rate that is approximately $1 - 1/(n - m_0)$ as long as the "effective aspect ratio" $(m - m_0)/(n - m_0)$ of the system is sufficiently large (Theorem 3.17). Note that $1 - 1/(n - m_0)$ is the best possible rate that can be achieved by the RK algorithm (with any sampling distribution) on a consistent $(m - m_0) \times (n - m_0)$ linear system (see [21]).

[1] The projector $\mathbf{A}_{I_0}^{\dagger}\mathbf{A}_{I_0}$ can be written as $\mathbf{Q}\mathbf{Q}^{\mathsf{T}}$ where $\mathbf{Q} \in \mathbb{R}^{n\times m_0}$ is a matrix whose columns form an orthonormal basis of Range$(\mathbf{A}_{I_0}^{\mathsf{T}})$, which can be computed in $O(m_0^2 n)$ flops.

1.2.2. The QuantileSCRK method

We also propose a modification of the SCRK method for solving corrupted systems of linear equations. This setting models applications where some measurements are corrupted by arbitrarily large errors, which may occur during the data collection, transmission, and storage process due to malfunctioning sensors or faulty components (for more examples, see [58,24]). Unlike the noisy setting above, the error horizon is not very meaningful since significant outliers can be introduced. Hence, the aim is to converge to the solution $\mathbf{x}^*$ *exactly* by identifying and avoiding corruptions, which may be possible if the number of corruptions is relatively small and the system is highly overdetermined.

Our model for *corrupted linear systems* is defined as follows. Let $\mathcal{C} \subseteq [m]$ and $\mathbf{b}_{\mathcal{C}} \in \mathbb{R}^n$ be a *sparse* vector of *arbitrary* (possibly adversarial) corruptions supported on $\mathcal{C}$. Moreover, suppose that we possess external knowledge in the form of a corruption-free subset $I_0 \subset [m]$ of size m_0 such that $(\mathbf{b}_{\mathcal{C}})_{I_0} = \mathbf{0}$; for example, this could reflect a set of trustworthy measurements by a reliable source, or infallible equations arising from physical laws. The goal is to reconstruct the solution $\mathbf{x}^*$ of $\mathbf{A}\mathbf{x} = \mathbf{b}$ given $\mathbf{A}$, I_0, and the corrupted measurements $\widetilde{\mathbf{b}} := \mathbf{b} + \mathbf{b}_{\mathcal{C}}$.

To achieve convergence, we take inspiration from the QuantileRK method proposed in [25], which modifies the RK algorithm so that each projection is sampled from a set of admissible rows whose residuals $|b_j - \mathbf{a}_j^{\mathsf{T}}\mathbf{x}^k|$ are smaller than the q^{th} quantile of residual sizes at each iteration for some parameter $q \in (0, 1]$. This modification is based on the heuristic that large residuals should be indicative of corrupted measurements, and small residuals lead to small steps that cannot divert the iterate too far away from the solution. We propose to exploit the auxiliary information by confining the iterates of QuantileRK within the "trusted" solution space $\mathbf{A}_{I_0}\mathbf{x} = \mathbf{b}_{I_0}$. We will refer to this procedure, summarized in Algorithm 2, as the *QuantileSCRK method*.

Algorithm 2 Quantile Subspace Constrained Randomized Kaczmarz (QuantileSCRK).

procedure QuantileSCRK($\mathbf{A}, \widetilde{\mathbf{b}}, I_0, q, K$)
  $\mathbf{P} = \mathbf{I} - \mathbf{A}_{I_0}^{\dagger}\mathbf{A}_{I_0}$ ▷ Orthogonal projector onto Null($\mathbf{A}_{I_0}$)
  initialize $\mathbf{x}^0 = \mathbf{A}_{I_0}^{\dagger}\widetilde{\mathbf{b}}_{I_0}$ ▷ Initial iterate $\mathbf{x}^0$ solves $\mathbf{A}_{I_0}\mathbf{x}^0 = \widetilde{\mathbf{b}}_{I_0}$
  for $k = 1, \ldots, K$ **do**
    $\gamma_q = q$-quantile $\left\{|\widetilde{b}_j - \mathbf{a}_j^{\mathsf{T}}\mathbf{x}^k|,\ j \in [m] \setminus I_0\right\}$ ▷ Threshold based on residuals
    $J = \left\{j \in [m] \setminus I_0 : |\widetilde{b}_j - \mathbf{a}_j^{\mathsf{T}}\mathbf{x}^k| \le \gamma_q\right\}$ ▷ Set of admissible rows
    sample $j \in J$ with prob. proportional to $\|\mathbf{P}\mathbf{a}_j\|^2$ ▷ Sample admissible row
    $\mathbf{x}^k = \mathbf{x}^{k-1} + \frac{\widetilde{b}_j - \mathbf{a}_j^{\mathsf{T}}\mathbf{x}^{k-1}}{\|\mathbf{P}\mathbf{a}_j\|} \cdot \frac{\mathbf{P}\mathbf{a}_j}{\|\mathbf{P}\mathbf{a}_j\|}$ ▷ Project onto $\mathbf{A}_{I_0\cup\{j\}}\mathbf{x} = \widetilde{\mathbf{b}}_{I_0\cup\{j\}}$
  end for
  return $\mathbf{x}^K$
end procedure

We prove the following result for QuantileSCRK, a simplified version of Theorem 4.1 that we defer the precise statement of to Section 4. It shows that for unstructured matrices modeled by continuous "Gaussian-like" random matrices, the QuantileSCRK method robustly and efficiently converges, provided that there is enough external knowledge (in

terms of m_0) to make the *effective aspect ratio* $(m - m_0)/(n - m_0)$ large enough, and the proportion of corrupted measurements, $\beta := |\mathcal{C}|/(m - m_0)$, is not too large:

Theorem 1.4 *(Simplified version of Theorem 4.1). Assuming that* $\mathbf{A}$ *is a continuous "Gaussian-like" random matrix, there exist positive constants* $R \geq 1$, $\beta_0 < 1$, c_1 *and* c_2, *which are independent of* m *and* n, *such that if* $(m-m_0)/(n-m_0) \geq R$ *and* $\beta \leq \beta_0$, *then with probability at least* $1 - e^{-c_1(m-m_0)}$ *over the randomness in* $\mathbf{A}$, *the QuantileSCRK iterates* $\mathbf{x}^k$ *from Algorithm 2 converge to the solution* $\mathbf{x}^*$ *with*

$$\mathbb{E}\|\mathbf{x}^k - \mathbf{x}^*\|^2 \leq \left(1 - \frac{c_2}{n - m_0}\right)^k \cdot \|\mathbf{x}^0 - \mathbf{x}^*\|^2. \tag{1.7}$$

Since we are interested in large-scale systems with $m, n \gg 1$, the values of the constants c_1 and c_2 are dominated by m and n (e.g. the probability guarantee is exponentially close to one for large m). Note that this result applies to almost-square matrices with $m = (1 + o(1))n$ rows provided m_0 is big enough, which lies outside the scope of existing QuantileRK theory. Experimentally, we found that the QuantileSCRK method works well for more general data models than described by the theory, such as when $\mathbf{A}$ is a structured sparse matrix in an image reconstruction problem (see Section 5.6).

Remark 1.5.

(i) (Rejection sampling). To avoid recomputing the normalizing constant $Z_J = \sum_{j \in J} \|\mathbf{P}\mathbf{a}_j\|^2$ in every iteration of Algorithm 2 for sampling a row from the admissible set J, which depends on $\mathbf{x}^k$, rejection sampling (as originally proposed in [25]) can be used: i.e. in each iteration, a row $j \in I_1$ is sampled with probability $\|\mathbf{P}\mathbf{a}_j\|/\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2$, and the projection is made if and only if $|b_j - \mathbf{a}_j^\mathsf{T}\mathbf{x}^k| \leq \gamma_q$.

(ii) (Uniform sampling). It is computationally more efficient to sample rows $\mathbf{a}_j$ uniformly at random from I_1, together with rejection sampling, in Algorithm 2. By using the threshold $\widetilde{\gamma}_q = q$-quantile $\left\{|b_j - \mathbf{a}_j^\mathsf{T}\mathbf{x}^k|/\|\mathbf{P}\mathbf{a}_j\|,\ j \in I_1\right\}$, which has been modified to capture the heterogeneity of the projected row norms, instead of γ_q, it can be shown that analogues of our results (e.g. Theorem 4.1 and Lemma 4.2) still hold, except that the relevant spectral quantities come from the matrix $\mathbf{D}\mathbf{A}_{I_1}\mathbf{P}$, where $\mathbf{D}$ is the diagonal matrix with entries $\|\mathbf{P}\mathbf{a}_j\|^{-1}$, $j \in I_1$, instead of $\mathbf{A}_{I_1}\mathbf{P}$.

1.3. Organization

Section 2 discusses related works. Section 3 analyzes the SCRK algorithm: we prove the convergence result (Theorem 1.1) in Section 3.1 and generalize it to the noisy setting in Section 3.2 (Theorem 3.4). We provide several results on using the SCRK method to exploit low-rank structure and geometric properties of $\mathbf{A}$ in Section 3.3. Furthermore, we show that the subspace constraint acts as a form of dimension reduction when $\mathbf{A}$ is a Gaussian-like random matrix in Section 3.4. Section 4 analyzes the QuantileSCRK

algorithm for solving corrupted linear systems. We provide various numerical experiments in Section 5 to complement our theoretical results, and conclude in Section 6.

2. Related works

Kaczmarz-type methods Kaczmarz-type algorithms are related to a variety of modern algorithms in (randomized) numerical linear algebra and (stochastic) optimization. Randomized Kaczmarz (RK) can be viewed as an instance of the stochastic gradient descent (SGD) algorithm with a particular step size, which, based on this connection, has led to new insights into both methods, such as highlighting the role of weighted sampling for SGD [48]. Furthermore, the RK method is one of the basic representatives of the sketch-and-project method [21], which provides a unified framework for iteratively solving linear systems – including the randomized coordinate descent method, related block variants, and the randomized Newton method – and can also be directly extended to non-linear optimization problems [22].

Recently, methods based on the Kaczmarz algorithm have been used in the design of more sophisticated linear solvers. A GMRES-type solver preconditioned by randomized and greedy Kaczmarz inner-iterations is studied in [19]. Within the sketch-and-project framework, a randomized block Kaczmarz algorithm that uses the preconditioned conjugate gradient method to perform an inexact projection in each iteration is analyzed in [14], and the method is proven to be especially computationally efficient when the matrix $\mathbf{A}$ has a flat-tailed spectrum. An iterative solver that further combines these ideas with momentum and sparse sketching matrices is analyzed in [15]. In this paper, we focus on the randomized Kaczmarz algorithm for solving systems of linear equations.

Randomized Kaczmarz The analysis of the RK algorithm in [57] spurred many developments and variants, including randomized block Kaczmarz methods [45,42] and Kaczmarz-Motzkin methods that combine sampling and greedy row selection [12,23]. It is shown that the Kaczmarz method can be extended to solve least squares problems in [62,47,38], and systems of linear inequalities in [34,5]. The duality between RK and randomized coordinate descent (also called randomized Gauss-Seidel) has motivated a unified description of the two methods and their extended versions [38]. By using ideas from optimization, RK methods with varying step sizes [48], acceleration [35], and that promote sparsity [36,53,54] have also been studied.

The SCRK method resembles the block Kaczmarz method studied in [45]: the difference is that the blocks differ by one row between iterations and thus the update simplifies so that computing new pseudoinverses is not required. Therefore, the SCRK method can offer a similar advantage from using blocks in an efficient manner if a "good" block $\mathbf{A}_{I_0}$ can be found (see Section 3.3 for a discussion of what a good block is, and how one might be found). In [46,61], a two-subspace Kaczmarz method that iteratively projects onto the solution space associated with two rows is shown to significantly outperform the RK

method when the system has correlated rows. The SCRK method can be considered as an extension of this idea to higher dimensional subspaces (see Remark 3.9).

The idea of constraining the iterates of the RK algorithm has also implicitly appeared in the design of a fast solver for Laplacian systems in the theoretical computer science literature in [33], where the row spaces of the blocks $\mathbf{A}_{I_0}$ and $\mathbf{A}_{I_1}$ are orthogonal by construction and hence Corollary 1.2 applies. The SCRK method offers a general framework for analyzing convergence when $\mathbf{A}_{I_0}$ and $\mathbf{A}_{I_1}$ are not orthogonal. Randomized sketch descent methods for solving optimization problems subject to linear constraints are also studied in [43], where in each iteration the coordinate space (corresponding to $\mathbf{x}$) is randomly sketched for dimensionality reduction and (random) projection matrices, analogous to $\mathbf{P}$, enforce the linear constraints.

Corrupted linear systems The literature on solving linear systems with arbitrary sparse corruptions is abundant: see, e.g., [2,3,11,6,58]. Such problems are often tackled within the compressed sensing and robust statistics literature using methods based on linear or SDP relaxations. The closest line of work to our approach is on iterative, row-action, corruption-avoiding algorithms. The first Kaczmarz-type method was proposed in [24], which introduced the idea that large residuals should be indicative of corrupted equations, but makes strong restrictions on the number of corrupted measurements (scaling sublinearly with m). The QuantileRK method, introduced in [25], utilizes quantile-based steps based on this residual heuristic. An important bottleneck of this method is that the linear system generally needs to be sufficiently overdetermined (i.e. $m \geq Cn$ for a large constant C) to guarantee convergence. We show that with enough external knowledge (i.e. m_0 large enough), the QuantileSCRK method works even for almost-square systems. Other works studying the QuantileRK method include [56,30,10]; in particular, we adapt a deterministic sufficient condition for convergence from [56] for QuantileSCRK (see Lemma 4.2). Another Kaczmarz-type method based on obtaining sparse least squares solutions is analyzed in [54] and demonstrated to be able to solve linear systems corrupted by impulsive noise.

3. Analysis of the SCRK method

In this section, we provide theoretical analysis of the SCRK method (Algorithm 1). Recall that $\mathbf{P} = \mathbf{I} - \mathbf{A}_{I_0}^{\dagger}\mathbf{A}_{I_0}$ is the orthogonal projector onto $\mathrm{Null}(\mathbf{A}_{I_0})$, which is equal to $\mathrm{Range}(\mathbf{A}_{I_0}^{\mathsf{T}})^{\perp}$, the orthogonal complement of the row space of $\mathbf{A}_{I_0}$.

3.1. Simplified SCRK update formula and proof of Theorem 1.1

First, we provide a proof of how the block update (1.3) simplifies to the more interpretable and computationally efficient formula (1.4).

Lemma 3.1. *Let* $\mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{A}^{\dagger}_{I_0\cup\{j\}}(\mathbf{b}_{I_0\cup\{j\}} - \mathbf{A}_{I_0\cup\{j\}}\mathbf{x}^k)$ *and* $\mathbf{P} = \mathbf{I} - \mathbf{A}^{\dagger}_{I_0}\mathbf{A}_{I_0}$. *If* $\mathbf{x}^k$ *solves* $\mathbf{A}_{I_0}\mathbf{x}^k = \mathbf{b}_{I_0}$ *and* $\mathbf{P}\mathbf{a}_j \neq \mathbf{0}$, *then*

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \frac{b_j - \mathbf{a}_j^{\mathsf{T}}\mathbf{x}^k}{\|\mathbf{P}\mathbf{a}_j\|} \cdot \frac{\mathbf{P}\mathbf{a}_j}{\|\mathbf{P}\mathbf{a}_j\|}.$$

Proof. We may assume $b_j - \mathbf{a}_j^{\mathsf{T}}\mathbf{x}^k \neq 0$, otherwise $\mathbf{x}^{k+1} = \mathbf{x}^k$. Since $\mathbf{x}^{k+1}$ is the orthogonal projection of $\mathbf{x}^k$ onto the solution space $\mathbf{A}_{I_0\cup\{j\}}\mathbf{x} = \mathbf{b}_{I_0\cup\{j\}}$, the increment $\mathbf{z} := \mathbf{x}^{k+1} - \mathbf{x}^k$ is the solution of the following optimization problem:

$$\min_{\mathbf{z}\in\mathbb{R}^n} \frac{1}{2}\|\mathbf{z}\|^2 \quad \text{subject to} \quad \mathbf{A}_{I_0}(\mathbf{x}^k + \mathbf{z}) = \mathbf{b}_{I_0},\ \mathbf{a}_j^{\mathsf{T}}(\mathbf{x}^k + \mathbf{z}) = b_j. \tag{3.1}$$

This can be solved by introducing the Lagrange multipliers $\boldsymbol{\lambda} \in \mathbb{R}^n$ and $\tau \in \mathbb{R}$ for the two constraints. Since $\mathbf{A}_{I_0}\mathbf{x}^k = \mathbf{b}_{I_0}$, the first constraint is equivalent to $\mathbf{A}_{I_0}\mathbf{z} = \mathbf{0}$, and thus $\mathbf{z}$ solves

$$\mathbf{z} + \mathbf{A}_{I_0}^{\mathsf{T}}\boldsymbol{\lambda} + \tau\mathbf{a}_j = \mathbf{0} \tag{3.2}$$

whilst satisfying $\mathbf{A}_{I_0}\mathbf{z} = \mathbf{0}$ and $\mathbf{a}_j^{\mathsf{T}}\mathbf{z} = b_j - \mathbf{a}_j^{\mathsf{T}}\mathbf{x}^k$. Since $\mathbf{P}$ is the orthogonal projector onto $\mathrm{Null}(\mathbf{A}_{I_0})$, $\mathbf{P}\mathbf{A}_{I_0}^{\mathsf{T}} = \mathbf{0}$ and the first constraint is equivalent to $\mathbf{P}\mathbf{z} = \mathbf{z}$. Hence, pre-multiplying (3.2) by $\mathbf{P}$ implies that $\mathbf{z} = -\tau\mathbf{P}\mathbf{a}_j$. Furthermore, pre-multiplying (3.2) by $\mathbf{z}^{\mathsf{T}}$ and using the constraints implies that $\tau = -\|\mathbf{z}\|^2/(b_j - \mathbf{a}_j^{\mathsf{T}}\mathbf{x}^k) = -\tau^2\|\mathbf{P}\mathbf{a}_j\|^2/(b_j - \mathbf{a}_j^{\mathsf{T}}\mathbf{x}^k)$. Solving for τ yields $\tau = (b_j - \mathbf{a}_j^{\mathsf{T}}\mathbf{x}^k)/\|\mathbf{P}\mathbf{a}_j\|^2$, which completes the proof. □

Remark 3.2. From the optimization formulation (3.1), it can also be shown that the unit direction $\mathbf{P}\mathbf{a}_j/\|\mathbf{P}\mathbf{a}_j\|$ taken from $\mathbf{x}^k$ to reach $\mathbf{x}^{k+1}$ maximizes $|\mathbf{a}_j^{\mathsf{T}}\tilde{\mathbf{z}}|^2$ over all unit vectors $\tilde{\mathbf{z}} \in \mathrm{Null}(\mathbf{A}_{I_0})$. This provides a nice geometric interpretation of the SCRK update: *the direction* $\mathbf{P}\mathbf{a}_j$ *taken to reach the solution space* $\mathbf{A}_{I_0\cup\{j\}}\mathbf{x} = \mathbf{b}_{I_0\cup\{j\}}$ *minimizes the angle from the optimal direction* $\mathbf{a}_j$ *for reaching the solution space* $\mathbf{a}_j^{\mathsf{T}}\mathbf{x} = b_j$ *within the subspace* $\mathrm{Null}(\mathbf{A}_{I_0})$; see Fig. 3.1 for an illustration. For an alternative algebraic proof of a more general version of Lemma 3.1, see Remark 3.9 later.

We will now use the simplified update formula in Lemma 3.1 to prove Theorem 1.1.

Proof of Theorem 1.1. Consider the k^{th} iterate $\mathbf{x}^k$. Suppose that $\mathbf{a}_j$ is sampled in the next iteration (with $\mathbf{P}\mathbf{a}_j \neq \mathbf{0}$). By subtracting $\mathbf{x}^*$ from both sides of (1.4) and noting that $\mathbf{a}_j^{\mathsf{T}}(\mathbf{x}^k - \mathbf{x}^*) = \mathbf{a}_j^{\mathsf{T}}\mathbf{P}(\mathbf{x}^k - \mathbf{x}^*)$ for any $j \in I_1$ since $\mathbf{x}^k - \mathbf{x}^* \in \mathrm{Null}(\mathbf{A}_{I_0})$, we have

$$\mathbf{x}^{k+1} - \mathbf{x}^* = \left(\mathbf{I} - \mathbf{v}\mathbf{v}^{\mathsf{T}}\right)(\mathbf{x}^k - \mathbf{x}^*) \quad \text{for } \mathbf{v} := \frac{\mathbf{P}\mathbf{a}_j}{\|\mathbf{P}\mathbf{a}_j\|}.$$

Since $\mathbf{v}\mathbf{v}^{\mathsf{T}}$ is an orthogonal projector, $(\mathbf{x}^{k+1} - \mathbf{x}^*) \perp \mathbf{v}\mathbf{v}^{\mathsf{T}}(\mathbf{x}^k - \mathbf{x}^*)$. Thus, by Pythagoras' theorem,

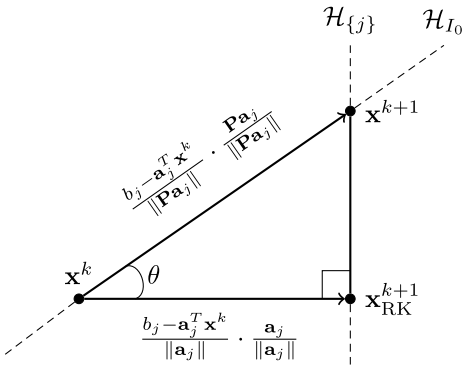


Fig. 3.1. SCRK update from the current iterate $\mathbf{x}^k$ for reaching the vector $\mathbf{x}^{k+1}$ in the solution space $\mathcal{H}_{\{j\}} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}_j^{\mathsf{T}}\mathbf{x} = b_j\}$ whilst remaining within $\mathcal{H}_{I_0} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}_{I_0}\mathbf{x} = \mathbf{b}_{I_0}\}$, compared to the RK update for reaching $\mathbf{x}^{k+1}_{\mathrm{RK}}$ alone.

$$\|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2 = \|\mathbf{x}^k - \mathbf{x}^*\|^2 - \|\mathbf{v}\mathbf{v}^{\mathsf{T}}(\mathbf{x}^k - \mathbf{x}^*)\|^2 = \|\mathbf{x}^k - \mathbf{x}^*\|^2 - \left|\mathbf{v}^{\mathsf{T}}(\mathbf{x}^k - \mathbf{x}^*)\right|^2.$$

By taking expectation (where each row $\mathbf{a}_j$ is sampled with probability $\|\mathbf{P}\mathbf{a}_j\|^2/\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2$), conditional on all the choices up to the k^{th} iteration, we obtain

$$\begin{aligned}
\mathbb{E}_k\|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2 &= \|\mathbf{x}^k - \mathbf{x}^*\|^2 - \sum_{j \in I_1 : \mathbf{P}\mathbf{a}_j \neq \mathbf{0}} \frac{\|\mathbf{P}\mathbf{a}_j\|^2}{\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2} \cdot \left| \left(\frac{\mathbf{P}\mathbf{a}_j}{\|\mathbf{P}\mathbf{a}_j\|}\right)^{\mathsf{T}} (\mathbf{x}^k - \mathbf{x}^*) \right|^2 \\
&= \|\mathbf{x}^k - \mathbf{x}^*\|^2 - \frac{1}{\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2} \|\mathbf{A}_{I_1}\mathbf{P}(\mathbf{x}^k - \mathbf{x}^*)\|^2 \\
&= \left(1 - \frac{\theta^2}{\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2}\right) \cdot \|\mathbf{x}^k - \mathbf{x}^*\|^2, \quad \text{where } \theta := \left\| \mathbf{A}_{I_1}\mathbf{P} \left(\frac{\mathbf{x}^k - \mathbf{x}^*}{\|\mathbf{x}^k - \mathbf{x}^*\|}\right) \right\|.
\end{aligned}$$

The next step is to estimate θ from below, which requires more care than a similar estimate used to prove convergence of the RK method [57] since $\mathbf{A}_{I_1}\mathbf{P}$ has a nontrivial nullspace. A similar case where the system matrix has a nontrivial nullspace was also treated in [37]. First, observe that $\mathrm{Null}(\mathbf{A}_{I_1}\mathbf{P}) = \mathrm{Null}(\mathbf{P})$. Indeed, the nontrivial inclusion $\mathrm{Null}(\mathbf{A}_{I_1}\mathbf{P}) \subseteq \mathrm{Null}(\mathbf{P})$ follows from the observation that $\mathbf{A}_{I_1}\mathbf{P}\mathbf{y} = \mathbf{0}$ implies that $\mathbf{P}\mathbf{y} \in \mathrm{Null}(\mathbf{A}_{I_0}) \cap \mathrm{Null}(\mathbf{A}_{I_1}) = \mathrm{Null}(\mathbf{A}) = \{\mathbf{0}\}$, since $\mathbf{A}$ has full rank. Therefore, since $\mathbf{x}^k - \mathbf{x}^* \in \mathrm{Null}(\mathbf{A}_{I_0})$ is orthogonal to $\mathrm{Null}(\mathbf{P}) = \mathrm{Null}(\mathbf{A}_{I_1}\mathbf{P})$,

$$\theta = \left\| \mathbf{A}_{I_1}\mathbf{P} \left(\frac{\mathbf{x}^k - \mathbf{x}^*}{\|\mathbf{x}^k - \mathbf{x}^*\|}\right) \right\|^2 \geq \min_{\substack{\mathbf{z} \in \mathrm{Null}(\mathbf{A}_{I_0}) \\ \|\mathbf{z}\| = 1}} \|\mathbf{A}_{I_1}\mathbf{z}\|^2 = \sigma_{\min}^+(\mathbf{A}_{I_1}\mathbf{P})^2. \tag{3.3}$$

This implies the following bound for one step of the SCRK method:

$$\mathbb{E}_k\|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2 \leq \left(1 - \frac{\sigma_{\min}^+(\mathbf{A}_{I_1}\mathbf{P})^2}{\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2}\right) \cdot \|\mathbf{x}^k - \mathbf{x}^*\|^2.$$

By iterating and taking the full expectation, this concludes the proof of Theorem 1.1. □

Remark 3.3. If $\mathbf{v}_\ell$ is a right singular vector of $\mathbf{A}_{I_1}\mathbf{P}$ corresponding to the ℓ^{th} largest singular value $\sigma_\ell(\mathbf{A}_{I_1}\mathbf{P})$, it can be shown that

$$\mathbb{E}\left\langle \mathbf{x}^k - \mathbf{x}^*, \mathbf{v}_\ell \right\rangle = \left(1 - \frac{\sigma_\ell(\mathbf{A}_{I_1}\mathbf{P})^2}{\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2}\right)^k \cdot \left\langle \mathbf{x}^0 - \mathbf{x}^*, \mathbf{v}_\ell \right\rangle.$$

This shows that the residual vector $\mathbf{x}^k - \mathbf{x}^*$ decays fastest in the directions corresponding to the largest singular values of $\mathbf{A}_{I_1}$ restricted to $\text{Null}(\mathbf{A}_{I_0})$. This phenomenon was proved by Steinerberger [55] for the RK method.

3.2. SCRK convergence on inconsistent linear systems

In the general case, the measurement vector $\mathbf{b}$ might not be known exactly, but only accessible through a set of noisy observations $\widehat{\mathbf{b}} := \mathbf{b} + \mathbf{r}$, where $\mathbf{r}$ is an arbitrary error vector (which is considered to be small). Similar to previous analyses of Kaczmarz methods [44,45,50], we prove that if the SCRK method is used with the noisy measurements $\widehat{\mathbf{b}}$, then the iterates converge to the solution $\mathbf{x}^*$ up to an error horizon:

Theorem 3.4. *Suppose that the rows of $\mathbf{A}$ are partitioned into two blocks $\mathbf{A}_{I_0}$ and $\mathbf{A}_{I_1}$ of sizes m_0 and $m - m_0$ respectively, and assume that $\mathbf{A}_{I_0}$ has full row rank. If $\widehat{\mathbf{x}}^k$ denotes the sequence of SCRK iterates from Algorithm 1 where the noisy measurement vector $\widehat{\mathbf{b}} = \mathbf{b} + \mathbf{r}$ is used in place of $\mathbf{b}$, and the initial iterate $\widehat{\mathbf{x}}^0$ solves $\mathbf{A}_{I_0}\widehat{\mathbf{x}}^0 = \widehat{\mathbf{b}}_{I_0}$, then*

$$\mathbb{E}\|\widehat{\mathbf{x}}^k - \mathbf{x}^*\|^2 \le \left(1 - \frac{\sigma_{\min}^+(\mathbf{A}_{I_1}\mathbf{P})^2}{\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2}\right)^k \cdot \|\widehat{\mathbf{x}}^0 - \mathbf{x}^*\|^2 + \gamma_0 + \gamma_1,$$

where $\gamma_0, \gamma_1 \ge 0$ are given by

$$\gamma_0 = \frac{2\|\mathbf{r}_{I_0}\|^2}{\sigma_{\min}(\mathbf{A}_{I_0})^2} - \|\mathbf{A}_{I_0}^\dagger \mathbf{r}_{I_0}\|^2 \quad \textit{and} \quad \gamma_1 = \frac{\|\mathbf{r}_{I_1} - \mathbf{A}_{I_1}\mathbf{A}_{I_0}^\dagger \mathbf{r}_{I_0}\|^2}{\sigma_{\min}^+(\mathbf{A}_{I_1}\mathbf{P})^2}.$$

Remark 3.5 *(Error horizon).* Note that γ_0 only depends on the noise in the measurements corresponding to the fixed block I_0. In particular, if $\mathbf{r}_{I_0} = 0$, then $\gamma_0 = 0$ and γ_1 only depends on $\|\mathbf{r}_{I_1}\|^2$. On the other hand, if $\mathbf{r}_{I_1} = 0$, then γ_1 only depends on $\|\mathbf{A}_{I_1}\mathbf{A}_{I_0}^\dagger \mathbf{r}_{I_0}\|^2 = \sum_{j \in I_1} |\mathbf{a}_j^\intercal \mathbf{A}_{I_0}^\dagger \mathbf{r}_{I_0}|^2$. Note that $|\mathbf{a}_j^\intercal \mathbf{A}_{I_0}^\dagger \mathbf{r}_{I_0}|$ corresponds to the angle between the row $\mathbf{a}_j$ and $\mathbf{A}_{I_0}^\dagger \mathbf{r}_{I_0}$; the vector $\mathbf{A}_{I_0}^\dagger \mathbf{r}_{I_0}$ accounts for how noise in the measurements corresponding to the fixed block $\mathbf{A}_{I_0}$ shifts the solution space (see Lemma 3.8a). Finally, if $\mathbf{r} = 0$, then $\gamma_0 = \gamma_1 = 0$ and we recover Theorem 1.1.

Remark 3.6 *(Least squares).* Given a set of noisy measurements $\widehat{\mathbf{b}}$, our setup can easily be translated to the problem of solving the *inconsistent* system of linear equations $\mathbf{A}\mathbf{x} \approx \widehat{\mathbf{b}}$

in a least squares sense. In this setting, by defining $\mathbf{x}^*$ to be the least squares solution (i.e. $\mathbf{x}^* := \arg\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{A}\mathbf{x} - \widehat{\mathbf{b}}\|^2 = \mathbf{A}^\dagger\widehat{\mathbf{b}}$), and setting $\mathbf{b} := \mathbf{A}\mathbf{x}^*$ and $\mathbf{r} := \widehat{\mathbf{b}} - \mathbf{A}\mathbf{x}^*$, Theorem 3.4 can be applied to deduce that the SCRK iterates converge to the least squares solution up to the same error horizon.

We develop some technical results before proving Theorem 3.4. Note that with noisy measurements $\widehat{\mathbf{b}}$, the relationship $\widehat{\mathbf{x}}^k - \mathbf{x}^* \in \mathrm{Null}(\mathbf{A}_{I_0})$ does not necessarily hold anymore. Thus, it will be more convenient to work directly with the block update (1.3) instead. First, we present a decomposition of the pseudoinverse $\mathbf{A}_{I\cup J}^\dagger$ in terms of $\mathbf{A}_I$ and $\mathbf{A}_J$.

Lemma 3.7. *Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $I, J \subseteq [m]$ be two disjoint subsets of row indices. If $\mathbf{A}_{I\cup J}$ has full row rank, then the pseudoinverse $\mathbf{A}_{I\cup J}^\dagger$ admits the block representation*

$$\mathbf{A}_{I\cup J}^\dagger = \left(\mathbf{A}_I^\dagger - (\mathbf{A}_J\mathbf{P})^\dagger\mathbf{A}_J\mathbf{A}_I^\dagger \;\middle|\; (\mathbf{A}_J\mathbf{P})^\dagger\right), \tag{3.4}$$

where $\mathbf{P} = \mathbf{I} - \mathbf{A}_I^\dagger\mathbf{A}_I$ is the orthogonal projection operator onto $\mathrm{Null}(\mathbf{A}_I)$.

Proof of Lemma 3.7. First, we record the key algebraic property that will be used repeatedly:

$$\mathbf{X}^\dagger = \mathbf{X}^\mathsf{T}(\mathbf{X}\mathbf{X}^\mathsf{T})^{-1} \quad \text{for } \mathbf{X} = \mathbf{A}_{I\cup J},\ \mathbf{A}_I, \text{ or } \mathbf{A}_J\mathbf{P}. \tag{3.5}$$

This follows since all three matrices have full row rank: $\mathbf{A}_{I\cup J}$ by assumption, $\mathbf{A}_I$ as its row subset, and $\mathbf{A}_J\mathbf{P}$ from the observation that if its rows were linearly dependent then there would exist some nonzero $\boldsymbol{\alpha} \in \mathbb{R}^{|J|}$ such that $\sum_{j\in J}\alpha_j\mathbf{a}_j^\mathsf{T}\mathbf{P} = \mathbf{0}$, which would imply that $\sum_{j\in J}\alpha_j\mathbf{a}_j \in \mathrm{Null}(\mathbf{P}) = \mathrm{Range}(\mathbf{A}_I^\mathsf{T})$ and thus contradict the assumption that $\mathbf{A}_{I\cup J}$ has full row rank. We have

$$\mathbf{A}_{I\cup J}\mathbf{A}_{I\cup J}^\mathsf{T} = \left(\begin{array}{c|c} \mathbf{A}_I\mathbf{A}_I^\mathsf{T} & \mathbf{A}_I\mathbf{A}_J^\mathsf{T} \\ \hline (\mathbf{A}_I\mathbf{A}_J^\mathsf{T})^\mathsf{T} & \mathbf{A}_J\mathbf{A}_J^\mathsf{T} \end{array}\right) =: \left(\begin{array}{c|c} \mathbf{A}_{II} & \mathbf{A}_{IJ} \\ \hline \mathbf{A}_{IJ}^\mathsf{T} & \mathbf{A}_{JJ} \end{array}\right).$$

Since $\mathbf{A}_I$ has full row rank, $\mathbf{A}_{II}$ is invertible and

$$\begin{aligned} \mathbf{A}_{I\cup J}^\dagger &\overset{(3.5)}{=} \mathbf{A}_{I\cup J}^\mathsf{T}\left(\mathbf{A}_{I\cup J}\mathbf{A}_{I\cup J}^\mathsf{T}\right)^{-1} \\ &= \left(\mathbf{A}_I^\mathsf{T} \;\middle|\; \mathbf{A}_J^\mathsf{T}\right)\left(\begin{array}{c|c} \mathbf{A}_{II}^{-1} + \mathbf{A}_{II}^{-1}\mathbf{A}_{IJ}\mathbf{R}^{-1}\mathbf{A}_{IJ}^\mathsf{T}\mathbf{A}_{II}^{-1} & -\mathbf{A}_{II}^{-1}\mathbf{A}_{IJ}\mathbf{R}^{-1} \\ \hline -\mathbf{R}^{-1}\mathbf{A}_{IJ}^\mathsf{T}\mathbf{A}_{II}^{-1} & \mathbf{R}^{-1} \end{array}\right), \end{aligned} \tag{3.6}$$

where $\mathbf{R} := \mathbf{A}_{JJ} - \mathbf{A}_{IJ}^\mathsf{T}\mathbf{A}_{II}^{-1}\mathbf{A}_{IJ}$ is the Schur complement [29] of the block $\mathbf{A}_{II}$. Recall that $\mathbf{P} = \mathbf{I} - \mathbf{A}_I^\dagger\mathbf{A}_I$, and so

$$\mathbf{R} = \mathbf{A}_J\mathbf{A}_J^\mathsf{T} - \mathbf{A}_J\mathbf{A}_I^\mathsf{T}(\mathbf{A}_I\mathbf{A}_I^\mathsf{T})^{-1}\mathbf{A}_I\mathbf{A}_J^\mathsf{T} \overset{(3.5)}{=} \mathbf{A}_J[\mathbf{I} - \mathbf{A}_I^\dagger\mathbf{A}_I]\mathbf{A}_J^\mathsf{T} = \mathbf{A}_J\mathbf{P}\mathbf{A}_J^\mathsf{T},$$

which implies that

$$\mathbf{P}\mathbf{A}_J^\mathsf{T}\mathbf{R}^{-1} = \mathbf{P}\mathbf{A}_J^\mathsf{T}(\mathbf{A}_J\mathbf{P}\mathbf{A}_J^\mathsf{T})^{-1} = (\mathbf{A}_J\mathbf{P})^\mathsf{T}((\mathbf{A}_J\mathbf{P})(\mathbf{A}_J\mathbf{P})^\mathsf{T})^{-1} \overset{(3.5)}{=} (\mathbf{A}_J\mathbf{P})^\dagger. \tag{3.7}$$

Next we compute expressions for the two blocks of $\mathbf{A}_{I\cup J}^\dagger$ in (3.6). The first block, $\mathbf{A}_I^\mathsf{T}[\mathbf{A}_{II}^{-1} + \mathbf{A}_{II}^{-1}\mathbf{A}_{IJ}\mathbf{R}^{-1}\mathbf{A}_{IJ}^\mathsf{T}\mathbf{A}_{II}^{-1}] + \mathbf{A}_J^\mathsf{T}[-\mathbf{R}^{-1}\mathbf{A}_{IJ}^\mathsf{T}\mathbf{A}_{II}^{-1}]$, simplifies to

$$\mathbf{A}_I^\dagger - [-\mathbf{A}_I^\dagger\mathbf{A}_I + \mathbf{I}]\mathbf{A}_J^\mathsf{T}\mathbf{R}^{-1}\mathbf{A}_J\mathbf{A}_I^\dagger = \mathbf{A}_I^\dagger - \mathbf{P}\mathbf{A}_J^\mathsf{T}\mathbf{R}^{-1}\mathbf{A}_J\mathbf{A}_I^\dagger \overset{(3.7)}{=} \mathbf{A}_I^\dagger - (\mathbf{A}_J\mathbf{P})^\dagger\mathbf{A}_J\mathbf{A}_I^\dagger.$$

The second block, $\mathbf{A}_I^\mathsf{T}[-\mathbf{A}_{II}^{-1}\mathbf{A}_{IJ}\mathbf{R}^{-1}] + \mathbf{A}_J^\mathsf{T}[\mathbf{R}^{-1}]$, simplifies to

$$[-\mathbf{A}_I^\dagger\mathbf{A}_I + \mathbf{I}]\mathbf{A}_J^\mathsf{T}\mathbf{R}^{-1} = \mathbf{P}\mathbf{A}_J^\mathsf{T}\mathbf{R}^{-1} \overset{(3.7)}{=} (\mathbf{A}_J\mathbf{P})^\dagger.$$

Combining the two preceding displayed equations completes the proof. □

Next, we describe how the noise affects the geometry of the solution spaces.

Lemma 3.8. *Denote the true and noisy solution spaces associated with $I \subset [m]$ by*

$$\mathcal{H}_I = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}_I\mathbf{x} = \mathbf{b}_I\} \quad \text{and} \quad \widehat{\mathcal{H}}_I = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}_I\mathbf{x} = \mathbf{b}_I + \mathbf{r}_I\} \tag{3.8}$$

respectively. If $\mathbf{A}_I$ has full row rank, then $\mathcal{H}_I$ and $\widehat{\mathcal{H}}_I$ satisfy the following:

(a) $\widehat{\mathcal{H}}_I = \mathcal{H}_I + \mathbf{A}_I^\dagger\mathbf{r}_I$.
(b) $\widehat{\mathcal{H}}_I - \mathbf{x}^* = \mathrm{Null}(\mathbf{A}_I) + \mathbf{A}_I^\dagger\mathbf{r}_I$.
(c) *The vector $\mathbf{A}_I^\dagger\mathbf{r}_I$ is orthogonal to* $\mathrm{Null}(\mathbf{A}_I)$.

Proof. (a): Since $\mathbf{A}_I$ has full row rank, $\mathbf{A}_I\mathbf{A}_I^\dagger = \mathbf{I}$. Therefore, for any $\mathbf{x} \in \mathcal{H}_I$, we have $\mathbf{A}_I(\mathbf{x} + \mathbf{A}_I^\dagger\mathbf{r}_I) = \mathbf{b}_I + \mathbf{r}_I$ and so $\mathbf{x} + \mathbf{A}_I^\dagger\mathbf{r}_I \in \widehat{\mathcal{H}}_I$. Conversely, for any $\widehat{\mathbf{x}} \in \widehat{\mathcal{H}}_I$, we have $\mathbf{A}_I(\widehat{\mathbf{x}} - \mathbf{A}_I^\dagger\mathbf{r}_I) = \mathbf{b}_I$. Thus, $(\widehat{\mathbf{x}} - \mathbf{A}_I^\dagger\mathbf{r}_I) \in \mathcal{H}_I$, and so

$$\widehat{\mathbf{x}} = (\widehat{\mathbf{x}} - \mathbf{A}_I^\dagger\mathbf{r}_I) + \mathbf{A}_I^\dagger\mathbf{r}_I \in \mathcal{H}_I + \mathbf{A}_I^\dagger\mathbf{r}_I.$$

(b): Since $\mathbf{x}^* \in \mathcal{H}_I$, we have $\mathcal{H}_I - \mathbf{x}^* = \mathrm{Null}(\mathbf{A}_I)$. Together with part (a), this implies (b). Finally, (c) follows from the fact $\mathrm{Range}(\mathbf{A}_I^\dagger) = \mathrm{Range}(\mathbf{A}_I^\mathsf{T}) = \mathrm{Null}(\mathbf{A}_I)^\perp$. □

We will now prove Theorem 3.4 by using the expression for $\mathbf{A}_{I_0\cup\{j\}}^\dagger$ given in Lemma 3.7, as well as the geometry of the shifted solution spaces described by Lemma 3.8.

Proof of Theorem 3.4. Consider the k^{th} iterate $\widehat{\mathbf{x}}^k$. Suppose that $\mathbf{a}_j$ is sampled in the next iteration (with $\mathbf{P}\mathbf{a}_j \neq \mathbf{0}$ and hence $\mathbf{A}_{I_0\cup\{j\}}$ has full row rank). Then one step of

the SCRK algorithm with noisy measurements corresponds to the projection of $\widehat{\mathbf{x}}^k$ onto the noisy solution space $\widehat{\mathcal{H}}_{I_0\cup\{j\}}$, namely,

$$\widehat{\mathbf{x}}^{k+1} = \widehat{\mathbf{x}}^k + \mathbf{A}^\dagger_{I_0\cup\{j\}}(\mathbf{b}_{I_0\cup\{j\}} + \mathbf{r}_{I_0\cup\{j\}} - \mathbf{A}_{I_0\cup\{j\}}\widehat{\mathbf{x}}^k).$$

We will compare $\widehat{\mathbf{x}}^{k+1}$ with the projection of $\widehat{\mathbf{x}}^k$ onto the true solution space $\mathcal{H}_{I_0\cup\{j\}}$, denoted by

$$\mathbf{x}^{k+1} := \widehat{\mathbf{x}}^k + \mathbf{A}^\dagger_{I_0\cup\{j\}}(\mathbf{b}_{I_0\cup\{j\}} - \mathbf{A}_{I_0\cup\{j\}}\widehat{\mathbf{x}}^k).$$

Step 1. Exact computations. Note that

$$\widehat{\mathbf{x}}^{k+1} - \mathbf{x}^* = (\mathbf{x}^{k+1} - \mathbf{x}^*) + \mathbf{A}^\dagger_{I_0\cup\{j\}}\mathbf{r}_{I_0\cup\{j\}},$$

and $\mathbf{A}^\dagger_{I_0\cup\{j\}}\mathbf{r}_{I_0\cup\{j\}} \perp (\mathbf{x}^{k+1}-\mathbf{x}^*) \in \mathrm{Null}(\mathbf{A}_{I_0\cup\{j\}})$ by Lemma 3.8c. By using Pythagoras' theorem twice (and orthogonality of the true Kaczmarz projections), we have

$$\begin{aligned}\|\widehat{\mathbf{x}}^{k+1} - \mathbf{x}^*\|^2 &= \|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2 + \|\mathbf{A}^\dagger_{I_0\cup\{j\}}\mathbf{r}_{I_0\cup\{j\}}\|^2 \\ &= \|\widehat{\mathbf{x}}^k - \mathbf{x}^*\|^2 - \|\mathbf{A}^\dagger_{I_0\cup\{j\}}\mathbf{A}_{I_0\cup\{j\}}(\widehat{\mathbf{x}}^k - \mathbf{x}^*)\|^2 + \|\mathbf{A}^\dagger_{I_0\cup\{j\}}\mathbf{r}_{I_0\cup\{j\}}\|^2. \quad (3.9)\end{aligned}$$

By using Lemma 3.7 with $I = I_0$ and $J = \{j\}$, we can simplify the last two terms: firstly,

$$\begin{aligned}\mathbf{A}^\dagger_{I_0\cup\{j\}}\mathbf{r}_{I_0\cup\{j\}} &= \left(\mathbf{A}^\dagger_{I_0} - \tfrac{\mathbf{P}\mathbf{a}_j\mathbf{a}_j^\mathsf{T}\mathbf{A}^\dagger_{I_0}}{\|\mathbf{P}\mathbf{a}_j\|^2} \;\middle|\; \tfrac{\mathbf{P}\mathbf{a}_j}{\|\mathbf{P}\mathbf{a}_j\|^2}\right)\begin{pmatrix}\mathbf{r}_{I_0}\\ r_j\end{pmatrix} \\ &= \mathbf{A}^\dagger_{I_0}\mathbf{r}_{I_0} + (r_j - \mathbf{a}_j^\mathsf{T}\mathbf{A}^\dagger_{I_0}\mathbf{r}_{I_0})\cdot\frac{\mathbf{P}\mathbf{a}_j}{\|\mathbf{P}\mathbf{a}_j\|^2}. \quad (3.10)\end{aligned}$$

Next, since $\widehat{\mathbf{x}}^k - \mathbf{x}^* - \mathbf{A}^\dagger_{I_0}\mathbf{r}_{I_0} \in \mathrm{Null}(\mathbf{A}_{I_0})$ (from Lemma 3.8b) is a fixed point for $\mathbf{P}$,

$$\begin{aligned}\mathbf{A}^\dagger_{I_0\cup\{j\}}\mathbf{A}_{I_0\cup\{j\}}(\widehat{\mathbf{x}}^k - \mathbf{x}^*) &= \left(\mathbf{A}^\dagger_{I_0} - \tfrac{\mathbf{P}\mathbf{a}_j\mathbf{a}_j^\mathsf{T}\mathbf{A}^\dagger_{I_0}}{\|\mathbf{P}\mathbf{a}_j\|^2} \;\middle|\; \tfrac{\mathbf{P}\mathbf{a}_j}{\|\mathbf{P}\mathbf{a}_j\|^2}\right)\begin{pmatrix}\mathbf{r}_{I_0}\\ \mathbf{a}_j^\mathsf{T}(\widehat{\mathbf{x}}^k - \mathbf{x}^*)\end{pmatrix} \\ &= \mathbf{A}^\dagger_{I_0}\mathbf{r}_{I_0} + \frac{(\mathbf{P}\mathbf{a}_j)^\mathsf{T}(\widehat{\mathbf{x}}^k - \mathbf{x}^* - \mathbf{A}^\dagger_{I_0}\mathbf{r}_{I_0})}{\|\mathbf{P}\mathbf{a}_j\|}\cdot\frac{\mathbf{P}\mathbf{a}_j}{\|\mathbf{P}\mathbf{a}_j\|}. \quad (3.11)\end{aligned}$$

Furthermore, by Lemma 3.8c, $\mathbf{A}^\dagger_{I_0}\mathbf{r}_{I_0} \perp \mathbf{P}\mathbf{a}_j \in \mathrm{Null}(\mathbf{A}_{I_0})$, which implies that the two summands in both (3.10) and (3.11) are orthogonal. Hence, we can further expand (3.9) to show that $\|\widehat{\mathbf{x}}^{k+1} - \mathbf{x}^*\|^2$ is equal to

$$\|\widehat{\mathbf{x}}^k - \mathbf{x}^*\|^2 - \|\mathbf{A}^\dagger_{I_0}\mathbf{r}_{I_0}\|^2 - \left|\frac{(\mathbf{P}\mathbf{a}_j)^\mathsf{T}(\widehat{\mathbf{x}}^k - \mathbf{x}^* - \mathbf{A}^\dagger_{I_0}\mathbf{r}_{I_0})}{\|\mathbf{P}\mathbf{a}_j\|}\right|^2 + \|\mathbf{A}^\dagger_{I_0}\mathbf{r}_{I_0}\|^2 + \frac{|r_j - \mathbf{a}_j^\mathsf{T}\mathbf{A}^\dagger_{I_0}\mathbf{r}_{I_0}|^2}{\|\mathbf{P}\mathbf{a}_j\|^2}.$$

By canceling identical terms and taking the expectation, conditional on all the choices of the algorithm up to the k^{th} iteration (similar to the proof of Theorem 1.1), we obtain

$$\mathbb{E}_k\|\widehat{\mathbf{x}}^{k+1}-\mathbf{x}^*\|^2=\|\widehat{\mathbf{x}}^k-\mathbf{x}^*\|^2-\frac{\|\mathbf{A}_{I_1}\mathbf{P}(\widehat{\mathbf{x}}^k-\mathbf{x}^*-\mathbf{A}_{I_0}^\dagger\mathbf{r}_{I_0})\|^2}{\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2}+\frac{\|\mathbf{r}_{I_1}-\mathbf{A}_{I_1}\mathbf{A}_{I_0}^\dagger\mathbf{r}_{I_0}\|^2}{\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2}. \tag{3.12}$$

Step 2. Spectral bounds. Recall that $\text{Null}(\mathbf{A}_{I_1}\mathbf{P})=\text{Null}(\mathbf{P})$ from the proof of Theorem 1.1. Since $\widehat{\mathbf{x}}^k-\mathbf{x}^*-\mathbf{A}_{I_0}^\dagger\mathbf{r}_{I_0}\in\text{Null}(\mathbf{A}_{I_0})$, which is orthogonal to $\text{Null}(\mathbf{P})=\text{Null}(\mathbf{A}_{I_1}\mathbf{P})$, arguing as in (3.3) shows that

$$\|\mathbf{A}_{I_1}\mathbf{P}(\widehat{\mathbf{x}}^k-\mathbf{x}^*-\mathbf{A}_{I_0}^\dagger\mathbf{r}_{I_0})\|^2\geq\sigma_{\min}^+(\mathbf{A}_{I_1}\mathbf{P})^2\cdot\|\widehat{\mathbf{x}}^k-\mathbf{x}^*-\mathbf{A}_{I_0}^\dagger\mathbf{r}_{I_0}\|^2.$$

By expanding the square, $\|\widehat{\mathbf{x}}^k-\mathbf{x}^*-\mathbf{A}_{I_0}^\dagger\mathbf{r}_{I_0}\|^2$ is equal to $\|\widehat{\mathbf{x}}^k-\mathbf{x}^*\|^2+\|\mathbf{A}_{I_0}^\dagger\mathbf{r}_{I_0}\|^2-2(\widehat{\mathbf{x}}^k-\mathbf{x}^*)^\mathsf{T}\mathbf{A}_{I_0}^\dagger\mathbf{r}_{I_0}$. Since $\mathbf{A}_{I_0}^\dagger=\mathbf{A}_{I_0}^\mathsf{T}(\mathbf{A}_{I_0}\mathbf{A}_{I_0}^\mathsf{T})^{-1}$ and $\mathbf{A}_{I_0}(\widehat{\mathbf{x}}^k-\mathbf{x}^*)=\mathbf{r}_{I_0}$,

$$(\widehat{\mathbf{x}}^k-\mathbf{x}^*)^\mathsf{T}\mathbf{A}_{I_0}^\dagger\mathbf{r}_{I_0}=\mathbf{r}_{I_0}^\mathsf{T}(\mathbf{A}_{I_0}\mathbf{A}_{I_0}^\mathsf{T})^{-1}\mathbf{r}_{I_0}\leq\frac{\|\mathbf{r}_{I_0}\|^2}{\sigma_{\min}(\mathbf{A}_{I_0})^2}.$$

Hence, we can bound the second term of (3.12) from below by

$$\frac{\|\mathbf{A}_{I_1}\mathbf{P}(\widehat{\mathbf{x}}^k-\mathbf{x}^*-\mathbf{A}_{I_0}^\dagger\mathbf{r}_{I_0})\|^2}{\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2}\geq\frac{\sigma_{\min}^+(\mathbf{A}_{I_1}\mathbf{P})^2}{\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2}\left(\|\widehat{\mathbf{x}}^k-\mathbf{x}^*\|^2+\|\mathbf{A}_{I_0}^\dagger\mathbf{r}_{I_0}\|^2-\frac{2\|\mathbf{r}_{I_0}\|^2}{\sigma_{\min}(\mathbf{A}_{I_0})^2}\right).$$

By instating the definitions of γ_0 and γ_1, we have shown that

$$\mathbb{E}_k\|\widehat{\mathbf{x}}^{k+1}-\mathbf{x}^*\|^2\leq\left(1-\frac{\sigma_{\min}^+(\mathbf{A}_{I_1}\mathbf{P})^2}{\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2}\right)\cdot\|\widehat{\mathbf{x}}^k-\mathbf{x}^*\|^2+\frac{\sigma_{\min}^+(\mathbf{A}_{I_1}\mathbf{P})^2}{\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2}(\gamma_0+\gamma_1). \tag{3.13}$$

By iterating (3.13), we deduce that $\mathbb{E}\|\widehat{\mathbf{x}}^k-\mathbf{x}^*\|^2$ is upper bounded by

$$\left(1-\frac{\sigma_{\min}^+(\mathbf{A}_{I_1}\mathbf{P})^2}{\|\mathbf{A}_{\mathbf{I_1}}\mathbf{P}\|_F^2}\right)^k\cdot\|\widehat{\mathbf{x}}^0-\mathbf{x}^*\|^2+\sum_{i=0}^{k-1}\left(1-\frac{\sigma_{\min}^+(\mathbf{A}_{I_1}\mathbf{P})^2}{\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2}\right)^i\cdot\frac{\sigma_{\min}^+(\mathbf{A}_{I_1}\mathbf{P})^2}{\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2}(\gamma_1+\gamma_2).$$

We conclude by bounding the geometric series by $\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2/\sigma_{\min}^+(\mathbf{A}_{I_1}\mathbf{P})^2$. □

Remark 3.9. A natural generalization of the SCRK update (1.3) is to project onto the solution space $\mathbf{A}_{I_0\cup J}\mathbf{x}=\mathbf{b}_{I_0\cup J}$, where $J\subseteq[m]\setminus I_0$ is a block of row indices disjoint from I_0 with $|J|\geq 1$:

$$\mathbf{x}^{k+1}=\mathbf{x}^k+\mathbf{A}_{I_0\cup J}^\dagger(\mathbf{b}_{I_0\cup J}-\mathbf{A}_{I_0\cup J}\mathbf{x}^k). \tag{3.14}$$

Assuming that $\mathbf{A}_{I_0\cup J}$ has full row rank, Lemma 3.7 implies that (3.14) can be computed by the following two-step procedure (which does not require $\mathbf{x}^k$ to satisfy $\mathbf{A}_{I_0}\mathbf{x}^k=\mathbf{b}_{I_0}$):

(1) Project $\mathbf{x}^k$ onto the solution space $\mathbf{A}_{I_0}\mathbf{x} = \mathbf{b}_{I_0}$ to obtain $\mathbf{y}^k$:

$$\mathbf{y}^k = \mathbf{x}^k + \mathbf{A}_{I_0}^{\dagger}(\mathbf{b}_{I_0} - \mathbf{A}_{I_0}\mathbf{x}^k).$$

(2) Compute the new measurements $\boldsymbol{\beta}_J := \mathbf{b}_J - \mathbf{A}_J\mathbf{A}_{I_0}^{\dagger}\mathbf{b}_{I_0} \in \mathbb{R}^{|J|}$, then project $\mathbf{y}^k$ onto the solution space $\mathbf{A}_J\mathbf{x} = \mathbf{b}_J$ whilst remaining in the solution space of $\mathbf{A}_{I_0}\mathbf{x} = \mathbf{b}_{I_0}$ for $\mathbf{x}^{k+1}$:

$$\mathbf{x}^{k+1} = \mathbf{y}^k + (\mathbf{A}_J\mathbf{P})^{\dagger}(\boldsymbol{\beta}_J - (\mathbf{A}_J\mathbf{P})\mathbf{y}^k).$$

In particular, by restricting to a single row $J = \{j\}$ and imposing the condition $\mathbf{A}_{I_0}\mathbf{x}^k = \mathbf{b}_{I_0}$, we recover the simplified update formula (1.4), which provides an alternative algebraic proof of Lemma 3.1.

By further restricting to the special case $I = \{i\}$, the two-step procedure above reduces to an update of the two-subspace Kaczmarz method of [46]. Thus, the SCRK method can be seen as a partial generalization of the two-subspace Kaczmarz method, except that the subset I_0 is fixed throughout the iterations to exploit specific features of the block $\mathbf{A}_{I_0}$, and similar results concerning coherence with respect to more general subsets of equations can be obtained (see Remark 3.13).

Finally, while all the convergence results in this paper are stated for the case $|J| = 1$, we believe that similar techniques can be extended to the case of $|J| > 1$.

3.3. Exploiting structure with the SCRK method

In this section, we discuss how the SCRK method can exploit approximately low-rank structure and geometric properties of the data matrix $\mathbf{A}$ to accelerate convergence. For simplicity, we will restrict our attention to the noiseless case. Our goal is to study the per-iteration convergence rate (i.e. with $k = 1$). First, note that the SCRK rate (1.5) is as good as the RK rate (1.2). Indeed,

$$\sigma_{\min}^{+}(\mathbf{A}_{I_1}\mathbf{P}) = \min_{\substack{\mathbf{z}\in \mathrm{Null}(\mathbf{A}_{I_0}) \\ \|\mathbf{z}\|=1}} \|\mathbf{A}_{I_1}\mathbf{z}\| = \min_{\substack{\mathbf{z}\in \mathrm{Null}(\mathbf{A}_{I_0}) \\ \|\mathbf{z}\|=1}} \|\mathbf{A}\mathbf{z}\| \geq \min_{\substack{\mathbf{z}\in \mathbb{R}^n \\ \|\mathbf{z}\|=1}} \|\mathbf{A}\mathbf{z}\| = \sigma_{\min}(\mathbf{A}), \quad (3.15)$$

and $\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2 \leq \|\mathbf{P}\|^2\|\mathbf{A}_{I_1}\|_F^2 \leq \|\mathbf{A}\|_F^2$. Therefore,

$$1 - \frac{\sigma_{\min}^{+}(\mathbf{A}_{I_1}\mathbf{P})^2}{\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2} \leq 1 - \frac{\sigma_{\min}(\mathbf{A})^2}{\|\mathbf{A}\|_F^2}. \quad (3.16)$$

However, since each SCRK iteration requires more computation (as discussed in Remark 1.3), we would like to understand when the SCRK method is advantageous to RK overall. In the following, we first examine what features of the matrices $\mathbf{A}$ and $\mathbf{A}_{I_0}$ lead to such an advantage in Section 3.3.1. Furthermore, we discuss how a good subset I_0 of rows, if not explicitly given, can actually be efficiently found when $\mathbf{A}$ has approximately low-rank structure in Section 3.3.2.

3.3.1. Geometry of the matrix and convergence rates

As highlighted by (3.16), either $\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2 \ll \|\mathbf{A}\|_F^2$ or $\sigma_{\min}^+(\mathbf{A}_{I_1}\mathbf{P}) \gg \sigma_{\min}(\mathbf{A})$ leads to significant per-iteration advantage of SCRK over RK. We describe two specific motivating examples of systems with such structure before generalizing our observations in Corollary 3.12 below. For these examples, consider an arbitrary $(m - m_0)$-dimensional subspace $\mathcal{U}$ of $\mathbb{R}^n$. Let $\{\mathbf{u}_1, \dots, \mathbf{u}_{m_0}\}$ and $\{\mathbf{c}_{m_0+1}, \mathbf{c}_{m_0+2}, \dots, \mathbf{c}_n\}$ be orthonormal bases for $\mathcal{U}$ and $\mathcal{U}^\perp$ respectively, and $\varepsilon \approx 0$ be a small positive constant.

Example 3.10 *($\sigma_{\min}^+(\mathbf{A}_{I_1}\mathbf{P}) \gg \sigma_{\min}(\mathbf{A})$).* This can happen if the equations in the selected block $\mathbf{A}_{I_0}$ are almost collinear, but the system $\mathbf{A}_{I_1}\mathbf{P}$ with projected rows is well-conditioned. Let $\overline{\mathbf{u}} := \frac{1}{m_0}\sum_{i=1}^{m_0}\mathbf{u}_i$, and $\mathbf{A} \in \mathbb{R}^{n\times n}$ be the matrix where the first m_0 rows are given by $\mathbf{a}_j := (1-\varepsilon)\overline{\mathbf{u}} + \varepsilon\mathbf{u}_i$, $j = 1, \dots, m_0$, and the remaining rows are $\mathbf{c}_{m_0+1}, \dots, \mathbf{c}_n$. Choose $I_0 = [m_0]$ so that $\mathbf{P}$ is the orthogonal projection onto $\mathcal{U}^\perp$. Then $\sigma_{\min}^+(\mathbf{A}_{I_1}\mathbf{P}) = 1 \gg \varepsilon \geq \sigma_{\min}(\mathbf{A})$. Indeed, if $\mathbf{e}_i$ is the i^{th} standard basis vector, then

$$\sigma_{\min}(\mathbf{A}) = \min_{\mathbf{x}\in\mathbb{R}^n:\|\mathbf{x}\|=1}\|\mathbf{A}^\mathsf{T}\mathbf{x}\| \leq \left\|\mathbf{A}^\mathsf{T}\frac{(\mathbf{e}_1 - \mathbf{e}_2)}{\sqrt{2}}\right\| = \frac{\varepsilon}{\sqrt{2}}\|\mathbf{u}_1 - \mathbf{u}_2\| = \varepsilon.$$

Furthermore, since the rows of $\mathbf{A}_{I_1}\mathbf{P}$ form an orthonormal basis for $\mathcal{U}^\perp$, for any unit vector $\mathbf{x} = \sum_{i=m_0+1}^n \alpha_i\mathbf{c}_i$ in $\mathcal{U}^\perp$, we have $\|\mathbf{A}_{I_1}\mathbf{x}\|^2 = \sum_{i=m_0+1}^n \alpha_i^2 = 1$, and hence $\sigma_{\min}^+(\mathbf{A}_{I_1}\mathbf{P}) = \min_{\mathbf{x}\in\mathcal{U}^\perp:\|\mathbf{x}\|=1}\|\mathbf{A}_{I_1}\mathbf{x}\| = 1$.

Example 3.11 *($\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2 \ll \|\mathbf{A}\|_F^2$).* This can happen if the block $\mathbf{A}_{I_0}$ is highly correlated with the remaining rows $(\mathbf{a}_j)_{j\in I_1}$. Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be the matrix where the first m_0 rows are $\mathbf{u}_1, \dots, \mathbf{u}_{m_0}$, and the remaining rows are $\mathbf{a}_j := (1-\varepsilon)\mathbf{v}_j + \varepsilon\mathbf{c}_j$, where $\mathbf{v}_j$ is any unit vector in $\mathcal{U}$, for $j = m_0 + 1, \dots, n$. Choose $I_0 = [m_0]$, so that $\mathrm{Range}(\mathbf{A}_{I_0}^\mathsf{T}) = \mathcal{U}$ and $\mathbf{P}$ is the orthogonal projection onto $\mathcal{U}^\perp$. Then $\mathbf{P}\mathbf{a}_j = \varepsilon\mathbf{c}_j$ and $\|\mathbf{P}\mathbf{a}_j\| = \varepsilon \ll 1 = \|\mathbf{a}_j\|$ for all $j \in I_1$, and hence $\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2 = \sum_{j\in I_1}\|\mathbf{P}\mathbf{a}_j\|^2 \leq \varepsilon^2\sum_{j=1}^m\|\mathbf{a}_j\|^2 = \varepsilon^2\|\mathbf{A}\|_F^2$.

The calculations in the preceding example, together with Theorem 1.1 and (3.15), generalize to the following result:

Corollary 3.12. *Consider the same setup as Theorem 1.1, and assume that for some $\delta \in [0, 1)$,*

$$\frac{\|\mathbf{P}\mathbf{a}_j\|^2}{\|\mathbf{a}_j\|^2} \leq 1 - \delta^2 \quad \textit{for all } j \in I_1. \tag{3.17}$$

Then the SCRK iterates $\mathbf{x}^k$ converge to $\mathbf{x}^*$ in expectation with

$$\mathbb{E}\|\mathbf{x}^k - \mathbf{x}^*\|^2 \leq \left(1 - \frac{1}{1-\delta^2}\cdot\frac{\sigma_{\min}(\mathbf{A})^2}{\|\mathbf{A}\|_F^2}\right)^k \cdot \|\mathbf{x}^0 - \mathbf{x}^*\|^2. \tag{3.18}$$

Remark 3.13. Since $\mathbf{P}$ is the orthogonal projection onto $\text{Range}(\mathbf{A}_{I_0}^\mathsf{T})^\perp$, $\|\mathbf{P}\mathbf{a}_j\|^2/\|\mathbf{a}_j\|^2 = \sin^2\theta_j$ where θ_j is the principal angle between the subspaces $\text{Range}(\mathbf{a}_j)$ and $\text{Range}(\mathbf{A}_{I_0}^\mathsf{T})$. Therefore, the quantity δ in (3.17) measures the coherence between the row space of the fixed block $\mathbf{A}_{I_0}$ and each of the remaining rows $(\mathbf{a}_j)_{j\in I_1}$. A value of δ close to one means that the principal angles are uniformly small; i.e. all of the $\mathbf{a}_j$ are close to the row space of $\mathbf{A}_{I_0}$ and offer little new information by themselves. By projecting each row with $\mathbf{P}$, the shared information is effectively modded out, and thus each SCRK iteration is able to make more meaningful progress towards the solution.

In particular, if we take $\mathbf{A}_{I_0} = \mathbf{a}_i^\mathsf{T}$ to be a single row and assume that $\|\mathbf{a}_i\| = 1 = \|\mathbf{a}_j\|$, then $\|\mathbf{P}\mathbf{a}_j\|^2 = 1 - |\mathbf{a}_i^\mathsf{T}\mathbf{a}_j|^2$, where $|\mathbf{a}_i^\mathsf{T}\mathbf{a}_j|$ is the correlation between $\mathbf{a}_i$ and $\mathbf{a}_j$. It is shown in [46] that the two-subspace Kaczmarz method, which iteratively projects onto the solution space associated with two random rows, significantly improves upon RK if $\mathbf{A}$ has highly correlated rows. Thus, Corollary 3.12 quantifies a similar phenomenon for the SCRK method for higher dimensional subspaces.

3.3.2. Sampling rows to find a good subspace

Previously, we showed that bounds of the form $\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2 \ll \|\mathbf{A}\|_F^2$ using a specific choice of rows I_0 imply significant improvements in the convergence rate of the SCRK method over randomized Kaczmarz. However, what if we are not explicitly given a good set I_0, even though there is latent low-rank structure in $\mathbf{A}$ – in the sense that the matrix has $r \ll n$ dominant singular values – that can be exploited? We begin by considering a motivating hypothetical example where the row span of $\mathbf{A}_{I_0}$ is able to align perfectly with the leading right singular subspace.

Example 3.14 *($\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2 \ll \|\mathbf{A}\|_F^2$).* Let $\mathbf{A}_{(r)} := \mathbf{U}_{(r)}\mathbf{\Sigma}_{(r)}\mathbf{V}_{(r)}^\mathsf{T}$ be the best rank-r approximation of $\mathbf{A}$ (with respect to $\|\cdot\|_F$), where $\mathbf{\Sigma}_{(r)} = \text{diag}(\sigma_1(\mathbf{A}), \dots, \sigma_r(\mathbf{A}))$ is the diagonal matrix of the top r singular values of $\mathbf{A}$, and the columns of $\mathbf{V}_{(r)} \in \mathbb{R}^{n\times r}$ and $\mathbf{U}_{(r)} \in \mathbb{R}^{m\times r}$ contain the corresponding right and left singular vectors. Suppose that the row span of $\mathbf{A}_{I_0}$ equals $\text{Range}(\mathbf{V}_{(r)})$. Then

$$\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2 = \|\mathbf{A}\mathbf{P}\|_F^2 = \|\mathbf{P}\mathbf{A}^\mathsf{T}\|_F^2 = \|\mathbf{A}^\mathsf{T} - \mathbf{P}^\perp\mathbf{A}^\mathsf{T}\|_F^2 = \|\mathbf{A}^\mathsf{T} - \mathbf{A}_{(r)}^\mathsf{T}\|_F^2 = \sum_{i=r+1}^{n} \sigma_i(\mathbf{A})^2.$$

If the top r singular values of $\mathbf{A}$ are much larger than the rest, then $\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2$ is much smaller than $\|\mathbf{A}\|_F^2 = \sum_{i=1}^n \sigma_i(\mathbf{A})^2$.

Note that in general, such a subset of rows does not exist in $\mathbf{A}$. This raises the following question: can we efficiently find a small subset I_0 of rows of $\mathbf{A}$ so that the row span of $\mathbf{A}_{I_0}$ is a good approximation of the top r-dimensional right singular subspace of $\mathbf{A}$? This is known as the problem of finding an approximate CX decomposition in the randomized numerical linear algebra literature. Algorithms have been proposed that sample rows of $\mathbf{A}$ according to their Euclidean norms [16] or their leverage scores $(\ell_j)_{j\in[m]}$ [17], where

ℓ_j is the squared ℓ_2 norm of the j^{th} row of $\mathbf{U}_{(r)}$, using the same notation as the example above (for more details, see [39]). The following summarizes guarantees for these two sampling schemes proved in [17,4] and [16]:

Theorem 3.15 *(Theorem 1 [17] and [4]; Theorem 4 [16]). Suppose that c rows of $\mathbf{A}$ are independently sampled, where row j is selected with probability p_j in each trial. Let I_0 be the set of indices of the sampled rows, and $\mathbf{P}$ be the orthogonal projection onto* $\mathrm{Null}(\mathbf{A}_{I_0})$.

(i) *If $p_j = \ell_j / r$ and $c = O(r \log r/\varepsilon^2)$, then with probability at least 0.9,*

$$\|\mathbf{AP}\|_F^2 \leq (1+\varepsilon)\|\mathbf{A} - \mathbf{A}_{(r)}\|_F^2 = (1+\varepsilon)\sum_{i=r+1}^{n} \sigma_i(\mathbf{A})^2. \tag{3.19}$$

(ii) *If $p_j = \|\mathbf{a}_j\|^2/\|\mathbf{A}\|_F^2$ and $c = O(r/\varepsilon^2)$, then with probability at least 0.9,*

$$\|\mathbf{AP}\|_F^2 \leq \|\mathbf{A} - \mathbf{A}_{(r)}\|_F^2 + \varepsilon\|\mathbf{A}\|_F^2 = (1+\varepsilon)\sum_{i=r+1}^{n} \sigma_i(\mathbf{A})^2 + \varepsilon\sum_{i=1}^{r} \sigma_i(\mathbf{A})^2. \tag{3.20}$$

Theorem 3.15 implies that sampling $c \approx r \log r$ rows of $\mathbf{A}$ produces a subspace that tames the leading r singular values of $\mathbf{A}$ with high probability. In practice, it has been observed that a modest oversampling factor (i.e. c is a small constant times r) usually suffices [17]. The relative-error bound (3.19) is better than the additive-error bound (3.20); however it is more costly because it requires the estimation of the leverage scores (see, e.g., [18,26]). By combining Theorem 3.15 and (3.15) with the SCRK convergence result (Theorem 1.1), we deduce the following:

Corollary 3.16. *Suppose that $I_0 \subset [m]$ contains $m_0 = O(r \log r/\varepsilon^2)$ rows of $\mathbf{A}$, randomly sampled according to the leverage scores of $\mathbf{A}$ relative to its best rank-r approximation as described in Theorem 3.15, and partition $\mathbf{A}$ into blocks $\mathbf{A}_{I_0}$ and $\mathbf{A}_{I_1}$ with $I_1 = [m] \setminus I_0$. Then with probability at least* 0.9 *over the sampling of I_0, the SCRK iterates $\mathbf{x}^k$ satisfy*

$$\mathbb{E}\|\mathbf{x}^k - \mathbf{x}^*\|^2 \leq \left(1 - \frac{\sigma_n(\mathbf{A})^2}{(1+\varepsilon)\sum_{i=r+1}^{n} \sigma_i(\mathbf{A})^2}\right)^k \cdot \|\mathbf{x}^0 - \mathbf{x}^*\|^2.$$

Thus, if the rank of $\mathbf{A}$ is effectively less than r (in the sense $\sum_{i=r+1}^{n} \sigma_i(\mathbf{A})^2 \ll \|\mathbf{A}\|_F^2$), then the SCRK method with iterates constrained to the solution space corresponding to $m_0 = O(r \log r)$ randomly sampled rows significantly improves upon RK. Note that similar results are known for the sketch-and-project method [13]. Moreover, the effective rank of a large-scale matrix $\mathbf{A}$ can be estimated in a data-driven manner by sketching [40].

3.4. SCRK on random data and dimension reduction

Previously, we discussed how the SCRK method accelerates the iterative solver when the matrix $\mathbf{A}$ has approximately low-rank structure. In this section, we consider a somewhat complementary setting to study the effect of the subspace constraint when $\mathbf{A}$ *is unstructured and homogeneous*: namely, when $\mathbf{A}$ is drawn from a class of generic random matrices (precisely defined below) whose rows behave like independent standard Gaussian vectors. Such a matrix is typically well-conditioned as long as its aspect ratio m/n is large enough, and hence the corresponding linear system is easily solved using randomized Kaczmarz. However, for almost-square systems with an aspect ratio close to one, the convergence rate is far from optimal.

First, we review some definitions from probability theory (we refer to [60] for more details). If $\mathbf{a} \in \mathbb{R}^n$ is a random vector, we say that $\mathbf{a}$ is *mean-zero* if $\mathbb{E}[\mathbf{a}] = \mathbf{0}$, and $\mathbf{a}$ is *isotropic* if $\mathbb{E}\left[\mathbf{a}\mathbf{a}^{\mathsf{T}}\right] = \mathbf{I}$. We say that a scalar random variable X is K*-subgaussian* if its subgaussian norm $\|X\|_{\psi_2} := \inf_{t>0}\{\mathbb{E}\left[\exp(X^2/t^2) \leq 2\right]\}$ is bounded by $K > 0$; informally, this means that X concentrates around its mean with a light, exponentially decaying tail. Furthermore, a random vector $\mathbf{a}$ is K-subgaussian if all of its one-dimensional marginals are K-subgaussian: $\|\mathbf{a}\|_{\psi_2} := \sup_{\mathbf{z}\in\mathbb{R}^n:\|\mathbf{z}\|=1} \|\langle \mathbf{z}, \mathbf{a}\rangle\|_{\psi_2} \leq K$.

As before, we will continue to assume that $\mathbf{A}$ has full rank (almost surely). For our model, we will allow $\mathbf{A}_{I_0} \in \mathbb{R}^{m_0\times n}$ to be arbitrary, and we assume that $\mathbf{A}_{I_1} \in \mathbb{R}^{(m-m_0)\times n}$ is a random matrix that satisfies the following:

A1. The rows of $\mathbf{A}_{I_1}$ are independent, mean-zero, isotropic, and K-subgaussian random vectors.

The canonical example for our model is a standard Gaussian matrix $\mathbf{A}_{I_1}$ whose entries are independent standard normal random variables. In this special case, exact computations are often possible. More generically, Assumption A1 models unstructured matrices containing homogeneous data (that is centered and isotropic) with light tails.

Our main result in this section shows that for such matrices, the subspace constraint imposed by $\mathbf{P}$ acts as a form of dimension reduction, typically resulting in a near-optimal convergence rate of approximately $1 - 1/(n - m_0)$ as long as the "effective aspect ratio" $(m - m_0)/(n - m_0)$, which may be much larger than m/n, is large enough.

Theorem 3.17. *Suppose that the rows of* $\mathbf{A}$ *are partitioned into two blocks* $\mathbf{A}_{I_0}$ *and* $\mathbf{A}_{I_1}$ *of sizes* m_0 *and* $m - m_0$ *respectively, where* $\mathbf{A}_{I_0}$ *is arbitrary and* $\mathbf{A}_{I_1}$ *is a random matrix that satisfies Assumption A1. There exists constants* $c, R > 0$ *(only depending on* K*) such that if*

$$r := \frac{m - m_0}{n - m_0} \geq R,$$

then for any $\varepsilon \in (0,1)$*, with probability at least* $1 - 3\exp\left\{-c\varepsilon^2\left(1-\sqrt{\frac{R}{r}}\right)(m-m_0)\right\}$ *over the randomness in* $\mathbf{A}_{I_1}$*, the SCRK iterates* $\mathbf{x}^k$ *satisfy*

$$\mathbb{E}\|\mathbf{x}^k - \mathbf{x}^*\|^2 \leq \left(1 - \frac{(1-\varepsilon)^2}{1+\varepsilon}\left(1-\sqrt{\frac{R}{r}}\right)^2 \cdot \frac{1}{n-m_0}\right)^k \cdot \|\mathbf{x}^0 - \mathbf{x}^*\|^2. \tag{3.21}$$

In the special case where $\mathbf{A}_{I_1}$ *is standard Gaussian, this result holds with* $R = 1$*.*

The values of R and c depend on the precise distributional properties of the random matrix, and are, importantly, independent of m and n. Note that for tall, large-scale systems with $m, n \gg 1$ and $r \gg 1$, the requirement $r \geq R$ is not difficult to meet, and taking $\varepsilon \approx 0$ shows that the convergence rate is approximately $1 - 1/(n - m_0)$ with a probability guarantee that is exponentially close to one.

3.4.1. Proof of Theorem 3.17

To study the typical convergence rate with a random matrix, we will obtain tail bounds for $\sigma^+_{\min}(\mathbf{A}_{I_1}\mathbf{P})$ and $\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2$. The first lemma is deterministic, and shows that instead of studying the non-zero singular values of the $(m - m_0) \times n$ matrix $\mathbf{A}_{I_1}\mathbf{P}$, we can study the singular values of a thinner $(m - m_0) \times (n - m_0)$ matrix after rotating.

Lemma 3.18. *Let* $\mathbf{X} \in \mathbb{R}^{m\times n}$ *be a matrix, and* $\mathbf{P} \in \mathbb{R}^{n\times n}$ *be an orthogonal projection onto a d-dimensional subspace of* $\mathbb{R}^n$*. Suppose that the columns of* $\bar{\mathbf{Q}} \in \mathbb{R}^{n\times d}$ *form an orthonormal basis for* Range$(\mathbf{P})$*. Then the non-zero singular values of* $\mathbf{XP} \in \mathbb{R}^{m\times n}$ *and* $\mathbf{X}\bar{\mathbf{Q}} \in \mathbb{R}^{m\times d}$ *are the same.*

Proof. Note that $\mathbf{P} = \bar{\mathbf{Q}}\bar{\mathbf{Q}}^\mathsf{T}$. Let $\mathbf{X}\bar{\mathbf{Q}} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\mathsf{T}$ be a compact singular value decomposition of $\mathbf{X}\bar{\mathbf{Q}}$, which means that $\mathbf{\Sigma}$ is a square diagonal matrix containing the rank$(\mathbf{X})$ non-zero singular values of $\mathbf{X}\bar{\mathbf{Q}}$, and $\mathbf{U}, \mathbf{V}$ are rectangular matrices with orthonormal columns. Observe that $\mathbf{XP} = \mathbf{U}\mathbf{\Sigma}(\mathbf{V}')^\mathsf{T}$ where $\mathbf{V}' := \bar{\mathbf{Q}}\mathbf{V}$ also has orthonormal columns. This allows us to conclude the desired result since the non-zero singular values of $\mathbf{XP}$ are presented in the same matrix $\mathbf{\Sigma}$. □

The following probabilistic result shows that the smallest non-zero singular value of $\mathbf{A}_{I_1}\mathbf{P}$ can be lower bounded with very high probability.

Lemma 3.19. *Let* $\mathbf{P}$ *be an orthogonal projection onto a fixed* $(n - m_0)$*-dimensional subspace. Suppose that the random matrix* $\mathbf{A}_{I_1} \in \mathbb{R}^{(m-m_0)\times n}$ *satisfies Assumption A1. Then there exists an absolute constant* $C > 0$ *such that for all* $s > 0$*, with probability at least* $1 - 2e^{-s^2(m-m_0)}$*, the smallest and largest non-zero singular values of* $\mathbf{A}_{I_1}\mathbf{P}$ *satisfy*

$$\sigma^+_{\min}(\mathbf{A}_{I_1}\mathbf{P}) \geq \sqrt{m-m_0} - CK^2(\sqrt{n-m_0} + s\sqrt{m-m_0}) \quad \text{and}$$
$$\sigma_{\max}(\mathbf{A}_{I_1}\mathbf{P}) \leq \sqrt{m-m_0} + CK^2(\sqrt{n-m_0} + s\sqrt{m-m_0}).$$

In the case where $\mathbf{A}_{I_1}$ is Gaussian, the inequalities hold with CK^2 replaced by one.

Proof. Let $\bar{\mathbf{Q}} \in \mathbb{R}^{(m-m_0)\times(n-m_0)}$ be a matrix whose columns form an orthonormal basis for Range($\mathbf{P}$). By Lemma 3.18, the smallest and largest non-zero singular values of $\mathbf{A}_{I_1}\mathbf{P} \in \mathbb{R}^{(m-m_0)\times n}$ and $\mathbf{B} := \mathbf{A}_{I_1}\bar{\mathbf{Q}} \in \mathbb{R}^{(m-m_0)\times(n-m_0)}$ are equal. It can be directly checked that the rows of $\mathbf{B}$ are also independent, mean-zero, isotropic, K-subgaussian random vectors in $\mathbb{R}^{n-m_0}$. Hence, using a standard tail bound for the extremal singular values $\sigma_{\min}(\mathbf{B})$ and $\sigma_{\max}(\mathbf{B})$ of the random matrix $\mathbf{B}$ (see [60, Theorem 4.6.1]) implies the claimed inequalities. In the Gaussian case, the precise constants can be computed using Gaussian concentration tools (see [60, Corollary 7.3.3, Exercise 7.3.4]). □

Remark 3.20. The minimum restricted singular value of random matrices has also been studied in the context of universality laws for randomized dimension reduction in [49]. If the entries of $\mathbf{A}_{I_1}$ are independent random variables satisfying some mild regularity conditions, then [49, Theorem II] establishes that $\sigma^+_{\min}(\mathbf{A}_{I_1}\mathbf{P}) \approx \sqrt{m-m_0} - C\sqrt{n-m_0}$ with high probability since Range($\mathbf{P}$) is a $(n-m_0)$-dimensional subspace in $\mathbb{R}^n$. Thus, $\sigma^+_{\min}(\mathbf{A}_{I_1}\mathbf{P})$ is of comparable order for a wide class of distributions. However, the maximum restricted singular value is not necessarily universal.

Next, our goal is to obtain tail bounds for $\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2 = \sum_{j\in I_1}\|\mathbf{P}\mathbf{a}_j\|^2$. In the setting where the rows $\mathbf{a}_j$ of $\mathbf{A}_{I_1}$ are mean-zero, K-subgaussian random vectors, it is proved in [31] that the Euclidean norms $\|\mathbf{a}_j\|$ are $O(K\sqrt{n})$-subgaussian. The next lemma states that the norms of the projected vectors $\|\mathbf{P}\mathbf{a}_j\|$ are $O(K\sqrt{n-m_0})$-subgaussian.

Lemma 3.21. *Let $\mathbf{a}$ be a mean-zero, K-subgaussian random vector in $\mathbb{R}^n$, and $\mathbf{P}$ be an orthogonal projection onto a fixed d-dimensional subspace. Then the subgaussian norm of $\|\mathbf{P}\mathbf{a}\|$ is bounded by $CK\sqrt{d}$ for some absolute constant $C > 0$.*

The proof uses the following geometric observation about unit spheres of subspaces, which we record for later reference. We say that $\mathcal{N}$ is an *ε-net* of a set $S \subseteq \mathbb{R}^n$ if $\mathcal{N} \subseteq S$ and every point in S is within distance ε of some point in $\mathcal{N}$. It is known that there exists an ε-net of the d-dimensional unit sphere $\mathbb{S}^{d-1} := \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\| = 1\}$ with cardinality bounded by $(1+2/\varepsilon)^d$ for any d (see, e.g., [60, Corollary 4.2.13]). Thus, if $\mathcal{U}$ is a d-dimensional subspace of $\mathbb{R}^n$, then by identifying $\mathcal{U} \cong \mathbb{R}^d$ (using the fact that rotations are isometries) and obtaining a net of $\mathbb{S}^{d-1}$, we deduce the following:

Lemma 3.22. *Let $\mathcal{U}$ be a d-dimensional subspace of $\mathbb{R}^n$. Then for any $\varepsilon > 0$, there exists an ε-net $\mathcal{N}$ of $\mathcal{U} \cap \mathbb{S}^{n-1}$ with cardinality $|\mathcal{N}| \le (1+2/\varepsilon)^d$.*

Proof of Lemma 3.21. The proof is similar to the proof of [31, Lemma 1]; we provide it for completeness. First, it can be checked that $\mathbf{P}\mathbf{a}$ is also a mean-zero K-subgaussian random vector. Next, by Lemma 3.22, we can fix a 1/2-net $\mathcal{N}$ of Range($\mathbf{P}$) $\cap\, \mathbb{S}^{n-1}$ with cardinality $|\mathcal{N}| \le 5^{n-m_0}$. By using these observations, we will show that

$$\mathbb{P}\left(\|\mathbf{Pa}\| \geq t\right) \leq 2e^{-t^2/(10K^2(n-m_0))} \quad \text{for all } t \geq 0. \tag{3.22}$$

Indeed, for any realization of $\mathbf{Pa}$, there exists $\mathbf{v} \in \mathcal{N}$ such that $\left\| \frac{\mathbf{Pa}}{\|\mathbf{Pa}\|} - \mathbf{v} \right\| \leq 1/2$, and we can write

$$\|\mathbf{Pa}\| = \langle \mathbf{v}, \mathbf{Pa} \rangle + \left\langle \frac{\mathbf{Pa}}{\|\mathbf{Pa}\|} - \mathbf{v}, \mathbf{Pa} \right\rangle \leq \langle \mathbf{v}, \mathbf{Pa} \rangle + \left\| \frac{\mathbf{Pa}}{\|\mathbf{Pa}\|} - \mathbf{v} \right\| \cdot \|\mathbf{Pa}\| \leq \langle \mathbf{v}, \mathbf{Pa} \rangle + \frac{\|\mathbf{Pa}\|}{2}$$

to deduce that $\|\mathbf{Pa}\| \leq 2\langle \mathbf{v}, \mathbf{Pa} \rangle$. Therefore, because $\langle \mathbf{v}, \mathbf{Pa} \rangle$ is a mean-zero, K-subgaussian random variable, a union bound implies that for all $t \geq 0$,

$$\mathbb{P}\left(\|\mathbf{Pa}\| \geq t\right) \leq \mathbb{P}\left(\exists \mathbf{v} \in \mathcal{N} : \langle \mathbf{v}, \mathbf{Pa} \rangle \geq \frac{t}{2}\right) \leq 5^{n-m_0} \cdot e^{-t^2/(4K^2)}.$$

We claim that this implies (3.22). If $t^2 \leq 4\log(5)K^2(n-m_0)$, then (3.22) trivially holds. Otherwise, if $t^2 = 4\log(5)K^2(n-m_0) + s$ for $s > 0$, then

$$5^{n-m_0} \cdot e^{-t^2/(4K^2)} = e^{-s/(4K^2)} \leq e^{-s/(6\log(5)K^2(n-m_0))} \leq 2e^{-t^2/(10K^2(n-m_0))}.$$

Thus, the tail bound (3.22) holds for all $t \geq 0$, which implies that (by, e.g., [60, Proposition 2.5.2]), $\|\mathbf{Pa}\|$ has subgaussian norm bounded by $CK\sqrt{n-m_0}$ for some absolute constant $C > 0$. □

Lemma 3.23. *Consider the same setup as Lemma 3.19. Then there exists an absolute constant $c > 0$ such that for all $\varepsilon > 0$, with probability at least $1 - e^{-c\min\{\varepsilon, \varepsilon^2\}(m-m_0)/K^4}$,*

$$\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2 \leq (1+\varepsilon)(m-m_0)(n-m_0). \tag{3.23}$$

Proof. Since $\mathbf{P}$ is an orthogonal projection onto an $(n-m_0)$-dimensional subspace and $\mathbf{a}_j$ is isotropic, using the cyclic property of trace implies that for all $j \in I_1$,

$$\mathbb{E}\|\mathbf{Pa}_j\|^2 = \mathbb{E}\left[\mathrm{tr}(\mathbf{a}_j^\mathsf{T}\mathbf{Pa}_j)\right] = \mathrm{tr}(\mathbb{E}\left[\mathbf{a}_j\mathbf{a}_j^\mathsf{T}\right]\mathbf{P}) = \mathrm{tr}(\mathbf{P}) = n - m_0.$$

Therefore,

$$\mathbb{E}\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2 = \mathbb{E}\left[\sum_{j \in I_1} \|\mathbf{Pa}_j\|^2\right] = |I_1| \cdot \mathbb{E}\|\mathbf{Pa}_1\|^2 = (m-m_0)(n-m_0).$$

Now, the random variables $\|\mathbf{Pa}_j\|$ are independent and, by Lemma 3.21, $O(K\sqrt{n-m_0})$-subgaussian. Hence, by centering and Bernstein's inequality [60, Theorem 2.8.1], there exists an absolute constant $c > 0$ such that

$$\mathbb{P}\left(\|\mathbf{A}_{I_1}\mathbf{P}\|_F^2 - (m-m_0)(n-m_0) \geq t\right) \leq \exp\left(-\frac{c}{K^4}\min\left\{\frac{t^2}{(m-m_0)(n-m_0)^2}, \frac{t}{n-m_0}\right\}\right)$$

for all $t \geq 0$. By choosing $t = \varepsilon(m-m_0)(n-m_0)$, we obtain (3.23). □

The main result of this section now easily follows from the tail bounds for $\mathbf{A}_{I_1}\mathbf{P}$.

Proof of Theorem 3.17. Suppose that the random matrix $\mathbf{A}_{I_1}\mathbf{P}$ satisfies the events in Lemma 3.19, using $s = \varepsilon\left(1 - \frac{CK^2}{\sqrt{r}}\right)\frac{1}{CK^2}$ and relabeling C^2K^4 by R, and Lemma 3.23. If this occurs, which holds with the claimed probability after simplifying, the convergence result (3.21) then directly follows from the SCRK convergence result, Theorem 1.1. □

4. Analysis of the QuantileSCRK algorithm

In this section, we consider the QuantileSCRK method for solving corrupted linear systems (Algorithm 2). Recall that in our model, we are given a corrupted measurement vector $\widetilde{\mathbf{b}} := \mathbf{b} + \mathbf{b}_{\mathcal{C}}$, where $\mathbf{b}_{\mathcal{C}}$ is a sparse vector of arbitrary corruptions supported on $\mathcal{C} \subseteq [m]$, as well as a corruption-free subset $I_0 \subseteq [m]$ of size m_0 such that $(\mathbf{b}_{\mathcal{C}})_{I_0} = \mathbf{0}$. Our goal is to reconstruct the solution $\mathbf{x}^*$ of the linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Our main result in this section is Theorem 4.1, which shows that the QuantileSCRK method is able to converge robustly and efficiently when $\mathbf{A}$ is an unstructured random matrix as long as the effective aspect ratio $(m-m_0)/(n-m_0)$ is tall enough and the proportion of corrupted measurements $|\mathcal{C}|/(m-m_0)$ is not too large. Specifically, we consider the class of "Gaussian-like" random matrices previously considered in Section 3.4, and assume that $\mathbf{A}_{I_1}$ is a random matrix that satisfies Assumption A1 in addition to the following continuity assumption:

A2. Each row of $\mathbf{A}_{I_1}$ either has a log-concave distribution[2] or has independent entries with bounded probability densities.[3]

The class of log-concave distributions is a generalization of the standard Gaussian distribution that allows for some dependence between the entries of a random vector; for example, the uniform distribution over any convex body in $\mathbb{R}^n$ is log-concave. For more details and examples, we refer to [52].

Assumption A2 is essentially needed for technical reasons for our proof of Theorem 4.1. Empirically, convergence is observed even if $\mathbf{A}$ has random discrete entries [25], or if $\mathbf{A}$ is a structured, sparse matrix in an imaging problem (see Section 5.6). The assumption

[2] A log-concave distribution in $\mathbb{R}^n$ has a probability density f that satisfies $f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \geq f(\mathbf{x})^{\lambda} f(\mathbf{y})^{1-\lambda}$ for all $\lambda \in [0,1]$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

[3] By scaling, we may assume without loss of generality that the densities are bounded by one.

of having independent coordinates with bounded densities in each row was previously considered in [25], and we extend the model by allowing for log-concave distributions.

We can now state our main result:

Theorem 4.1. *Suppose that the rows of* $\mathbf{A}$ *are partitioned into blocks* $\mathbf{A}_{I_0}$ *and* $\mathbf{A}_{I_1}$ *of sizes* m_0 *and* $m - m_0$ *respectively, where* $\mathbf{A}_{I_0}$ *is arbitrary and* $\mathbf{A}_{I_1}$ *is a random matrix that satisfies Assumptions A1 and A2. Suppose that the corrupted measurement vector* $\widetilde{\mathbf{b}} = \mathbf{b} + \mathbf{b}_{\mathcal{C}}$ *is observed,* $(\mathbf{b}_{\mathcal{C}})_{I_0} = \mathbf{0}$*, and a quantile parameter* $q \in (0,1)$ *is fixed. There exist constants* $\beta_0 \in (0,1)$ *and* $R \geq 1$ *(only depending on* q *and* K*) such that if*

$$\frac{m - m_0}{n - m_0} \geq R \quad \text{and} \quad \beta := \frac{|\mathcal{C}|}{m - m_0} \leq \beta_0, \tag{4.1}$$

then for some constants $c_1, c_2 > 0$ *(only depending on* q*,* β*, and* K*), with probability at least* $1 - 6e^{-c_1(m-m_0)}$ *over the randomness in* $\mathbf{A}_{I_1}$*, the QuantileSCRK iterates* $\mathbf{x}^k$ *from Algorithm 2 converge to the solution* $\mathbf{x}^*$ *in expectation with*

$$\mathbb{E}\|\mathbf{x}^k - \mathbf{x}^*\|^2 \leq \left(1 - \frac{c_2}{n - m_0}\right)^k \cdot \|\mathbf{x}^0 - \mathbf{x}^*\|^2. \tag{4.2}$$

As mentioned previously, the values of the constants c_1 and c_2 are dominated in large-scale systems with $m, n \gg 1$, and the requirement $(m - m_0)/(n - m_0) \geq R$ is not difficult to meet if the system is tall and there is enough external knowledge (i.e. $m \gg n$, $m_0 \gg 1$). In addition, we believe that it should be possible to obtain sharper theoretical estimates for β_0 and R.

The strategy to prove Theorem 4.1 is to combine a deterministic sufficient condition for the convergence of QuantileSCRK, adapting a result for QuantileRK proved by [56], with probabilistic results for the spectra of the projected random matrix $\mathbf{A}_{I_1}\mathbf{P}$. First, we define some spectral quantities that will be needed. For $\alpha \in (0,1]$, define

$$\sigma^+_{\alpha,\min}(\mathbf{A}_{I_1}\mathbf{P}) := \inf_{\substack{T \subseteq I_1 \\ |T| = \alpha(m-m_0)}} \sigma^+_{\min}((\mathbf{A}_{I_1}\mathbf{P})_T). \tag{4.3}$$

For simplicity, we will assume throughout that $\alpha(m - m_0)$ is an integer. This quantity, which represents the uniform minimum singular value over all row submatrices of $\mathbf{A}_{I_1}\mathbf{P}$ with $\alpha(m-m_0)$ rows, has appeared in previous analyses of the QuantileRK algorithm [25, 56], and quantifies whether there are any poorly-conditioned row submatrices that are particularly susceptible to corruptions. Similarly, define

$$Z_\alpha := \sup_{\substack{T \subseteq I_1 \\ |T| = \alpha(m-m_0)}} \|(\mathbf{A}_{I_1}\mathbf{P})_T\|_F^2. \tag{4.4}$$

Together, (4.3) and (4.4) provide a uniform upper bound for the scaled condition numbers of all row submatrices of $\mathbf{A}_{I_1}\mathbf{P}$ possibly containing the uncorrupted, admissible rows

(i.e. whose residuals are smaller than the quantile threshold). This is the key step which guarantees that the expected improvement from moving in an uncorrupted direction offsets the expected deterioration caused by a corruption.

4.1. A deterministic condition for convergence

The following lemma provides a deterministic condition that guarantees the convergence of the QuantileSCRK algorithm for any arbitrary sparse corruption vector $\mathbf{b}_{\mathcal{C}}$, which may be of independent interest. This is adapted from a similar condition for convergence of the QuantileRK algorithm that was proved by Steinerberger [56].

Lemma 4.2. *Recall that $\mathcal{C} \subseteq I_1$ are the indices of the corrupted measurements, and $\beta = |\mathcal{C}|/(m - m_0)$. Suppose that $\beta < q < 1 - \beta$. Define*

$$C_{q,\beta} := \frac{1}{Z_{q-\beta}} \left\{ \sigma^{+}_{q-\beta,\min}(\mathbf{A}_{I_1}\mathbf{P})^2 - \sigma_{\max}(\mathbf{A}_{I_1}\mathbf{P})^2 \left(\frac{\beta}{1-q} + 2\sqrt{\frac{\beta}{1-q}} \right) \right\}, \tag{4.5}$$

where $\sigma^{+}_{q-\beta,\min}(\mathbf{A}_{I_1}\mathbf{P})$ is defined in (4.3), and $Z_{q-\beta}$ in (4.4). If $C_{q,\beta} > 0$, or equivalently,

$$\frac{\sigma^{+}_{q-\beta,\min}(\mathbf{A}_{I_1}\mathbf{P})^2}{\sigma_{\max}(\mathbf{A}_{I_1}\mathbf{P})^2} > \frac{\beta}{1-q} + 2\sqrt{\frac{\beta}{1-q}}, \tag{4.6}$$

then the QuantileSCRK iterates $\mathbf{x}^k$ from Algorithm 2 converge to the solution $\mathbf{x}^*$ in expectation with

$$\mathbb{E}\|\mathbf{x}^k - \mathbf{x}^*\|^2 \leq (1 - C_{q,\beta})^k \cdot \|\mathbf{x}^0 - \mathbf{x}^*\|^2. \tag{4.7}$$

The proof of Lemma 4.2 follows the same strategy as [56] with a minor improvement in the condition (4.6) for convergence. For completeness, we provide the full details.

Proof. Consider the iterate $\mathbf{x}^k$. Recall that $J = J(q,k)$ is the set of indices of the admissible rows that satisfy $|b_j - \mathbf{a}_j^{\mathsf{T}}\mathbf{x}^k| \leq \gamma_q = q\text{-quantile}\left\{|b_j - \mathbf{a}_j^{\mathsf{T}}\mathbf{x}^k| : j \in I_1\right\}$, with $|J| = q(m - m_0)$. Let $S := \mathcal{C} \cap J$ be the indices of the corrupted yet admissible rows, which satisfies $0 \leq |S| \leq \beta(m - m_0)$. Recall that the row j is sampled from J with probability equal to $\|\mathbf{P}\mathbf{a}_j\|^2/Z_J$, where $Z_J := \sum_{j\in J}\|\mathbf{P}\mathbf{a}_j\|^2$ is the normalizing constant. Conditional on all the choices up to the k^{th} iteration, we have

$$\mathbb{E}_k\|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2 = \left[1 - \frac{\sum_{j\in S}\|\mathbf{P}\mathbf{a}_j\|^2}{Z_J}\right] \mathbb{E}_{\{j\in J\setminus S\}}\|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2 \tag{4.8}$$

$$+ \frac{\sum_{j\in S}\|\mathbf{P}\mathbf{a}_j\|^2}{Z_J} \mathbb{E}_{\{j\in S\}}\|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2, \tag{4.9}$$

where $\mathbb{E}_{\{j\in S\}}$ denotes the expectation further conditional on $j \in S$, and similarly for $\mathbb{E}_{\{j\in J\setminus S\}}$. We proceed to estimate the two summands (4.8) and (4.9) individually.

Step 1: Lower bounding the improvement from selecting an uncorrupted equation. Conditional on sampling an admissible, uncorrupted equation $j \in J\setminus S$, the improvement is given by one step of the SCRK method applied to the row submatrix $(\mathbf{A}_{I_1}\mathbf{P})_{J\setminus S}$. Thus, by Theorem 1.1,

$$\mathbb{E}_{\{j\in J\setminus S\}}\|\mathbf{x}^{k+1}-\mathbf{x}^*\|^2 \le \left(1-\frac{\sigma^+_{\min}((\mathbf{A}_{I_1}\mathbf{P})_{J\setminus S})^2}{\sum_{j\in J\setminus S}\|\mathbf{P}\mathbf{a}_j\|^2}\right)\cdot\|\mathbf{x}^k-\mathbf{x}^*\|^2.$$

Since $|J\setminus S| \ge (q-\beta)(m-m_0)$, by using the definition of $\sigma^+_{q-\beta,\min}(\mathbf{A}_{I_1}\mathbf{P})$ in (4.3) together with the fact that adding rows to a matrix can only increase its minimum singular value, we obtain the following upper bound for the first term (4.8):

$$\left(1-\frac{\sum_{j\in S}\|\mathbf{P}\mathbf{a}_j\|^2}{Z_J}\right)\cdot\|\mathbf{x}^k-\mathbf{x}^*\|^2-\frac{\sigma^+_{q-\beta,\min}(\mathbf{A}_{I_1}\mathbf{P})^2}{Z_J}\cdot\|\mathbf{x}^k-\mathbf{x}^*\|^2. \tag{4.10}$$

Step 2: Upper bounding the deterioration from selecting a corrupted equation. The second term (4.9) represents the possible deterioration from selecting a corrupted yet admissible row that may take $\mathbf{x}^k$ further away from the solution $\mathbf{x}^*$. By expanding the square, it is equal to

$$\begin{aligned}
&\frac{\sum_{j\in S}\|\mathbf{P}\mathbf{a}_j\|^2}{Z_J}\sum_{j\in S}\frac{\|\mathbf{P}\mathbf{a}_j\|^2}{\sum_{i\in S}\|\mathbf{P}\mathbf{a}_i\|^2}\left\|\mathbf{x}^k-\mathbf{x}^*+\frac{b_j-\mathbf{a}_j^\mathsf{T}\mathbf{x}^k}{\|\mathbf{P}\mathbf{a}_j\|^2}\mathbf{P}\mathbf{a}_j\right\|^2\\
&=\frac{\sum_{j\in S}\|\mathbf{P}\mathbf{a}_j\|^2}{Z_J}\|\mathbf{x}^k-\mathbf{x}^*\|^2+\frac{1}{Z_J}\sum_{j\in S}|b_j-\mathbf{a}_j^\mathsf{T}\mathbf{x}^k|^2+\frac{2}{Z_J}\sum_{j\in S}(b_j-\mathbf{a}_j^\mathsf{T}\mathbf{x}^k)(\mathbf{P}\mathbf{a}_j)^\mathsf{T}(\mathbf{x}^k-\mathbf{x}^*)\\
&\le\frac{\sum_{j\in S}\|\mathbf{P}\mathbf{a}_j\|^2}{Z_J}\|\mathbf{x}^k-\mathbf{x}^*\|^2+\frac{1}{Z_J}|S|\gamma_q^2+\frac{2}{Z_J}\gamma_q\sqrt{|S|}\,\|(\mathbf{A}_{I_1}\mathbf{P})_C(\mathbf{x}^k-\mathbf{x}^*)\|,
\end{aligned}\tag{4.11}$$

where the definition of the quantile γ_q and Cauchy-Schwarz is used for the inequality.

Step 3: Bounding the q-quantile of a sample. Since any uncorrupted row $\mathbf{a}_j$ with $j \in I_1\setminus\mathcal{C}$ satisfies $\mathbf{a}_j^\mathsf{T}\mathbf{x}^* = b_j$, we have

$$b_j-\mathbf{a}_j^\mathsf{T}\mathbf{x}^k=\mathbf{a}_j^\mathsf{T}(\mathbf{x}^*-\mathbf{x}^k)=(\mathbf{P}\mathbf{a}_j)^\mathsf{T}(\mathbf{x}^*-\mathbf{x}^k),$$

recalling that $\mathbf{x}^*-\mathbf{x}^k \in \mathrm{Null}(\mathbf{A}_{I_0})$. Since there are at least $(1-q)(m-m_0)-(\beta(m-m_0)-|S|)=(1-q-\beta)(m-m_0)+|S|$ uncorrupted equations in I_1 whose residuals are larger than γ_q, we have

$$((1-q-\beta)(m-m_0)+|S|)\gamma_q^2\le\sum_{j\in I_1\setminus\mathcal{C}}|b_j-\mathbf{a}_j^\mathsf{T}\mathbf{x}^k|^2\le\sum_{j\in I_1}\left|(\mathbf{P}\mathbf{a}_j)^\mathsf{T}(\mathbf{x}^*-\mathbf{x}^k)\right|^2$$

$$= \|\mathbf{A}_{I_1}\mathbf{P}(\mathbf{x}^k - \mathbf{x}^*)\|^2 \leq \sigma_{\max}(\mathbf{A}_{I_1}\mathbf{P})^2 \cdot \|\mathbf{x}^k - \mathbf{x}^*\|^2.$$

Therefore, the q-quantile of the sizes of the residuals can be bounded by

$$\gamma_q \leq \frac{\sigma_{\max}(\mathbf{A}_{I_1}\mathbf{P})}{\sqrt{(m-m_0)(1-q-\beta)+|S|}} \cdot \|\mathbf{x}^k - \mathbf{x}^*\|. \tag{4.12}$$

Step 4: Conclude. Combining (4.10) and (4.11) with the bound on γ_q shows that the expected relative improvement $\mathbb{E}_k\|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2/\|\mathbf{x}^k - \mathbf{x}^*\|^2$ is upper bounded by

$$\begin{aligned} &1 - \frac{1}{Z_J}\left(\sigma^+_{q-\beta,\min}(\mathbf{A}_{I_1}\mathbf{P})^2 - \frac{|S|\gamma_q^2}{\|\mathbf{x}^k - \mathbf{x}^*\|^2} - \frac{2\gamma_q\sqrt{|S|}\,\|(\mathbf{A}_{I_1}\mathbf{P})_C(\mathbf{x}^k - \mathbf{x}^*)\|}{\|\mathbf{x}^k - \mathbf{x}^*\|^2}\right) \\ &\leq 1 - \frac{1}{Z_J}\left(\sigma^+_{q-\beta,\min}(\mathbf{A}_{I_1}\mathbf{P})^2 - \left[\frac{|S|\cdot\sigma_{\max}(\mathbf{A}_{I_1}\mathbf{P})^2}{\theta+|S|} + \frac{2\sqrt{|S|}\cdot\sigma_{\max}(\mathbf{A}_{I_1}\mathbf{P})^2}{\sqrt{\theta+|S|}}\right]\right), \end{aligned} \tag{4.13}$$

where $\theta := (m-m_0)(1-q-\beta)$. Now, we can upper bound Z_J by $Z_{q-\beta}$ from (4.4). Next, consider the function $f(x) = \frac{x}{\theta+x} + \frac{2\sqrt{x}}{\sqrt{\theta+x}}$, and observe that $f(|S|)$ appears in the upper bound (4.13). Since $f'(x) > 0$ for all $x > 0$, the upper bound is increasing in $|S|$. Because $|S| \leq \beta(m-m_0)$, we conclude that the most pessimistic bound, independent of $|S|$ and J (and hence k), is obtained by setting $|S| = \beta(m-m_0)$, which implies that

$$\mathbb{E}_k\|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2 \leq (1 - C_{q,\beta}) \cdot \|\mathbf{x}^k - \mathbf{x}^*\|^2, \quad \text{where } C_{q,\beta} \text{ is defined in (4.5).}$$

To ensure that the mean squared error contracts after each step, it suffices for $C_{q,\beta}$ to be positive: this is exactly secured by the condition (4.6). By iterating, we obtain (4.7). □

Note that Lemma 4.2 only provides a sufficient condition for convergence in the worst case (see [56] for further discussion). Empirically, convergence is observed for larger values of β because the corruptions are quickly detected and trapped beyond the threshold. The dependence on $|S|$ in (4.13) shows that if the number of admissible, corrupted equations is small, then far less is demanded of the spectral quantities of $\mathbf{A}_{I_1}\mathbf{P}$ for the mean squared error to contract. For similar reasons, the QuantileRK method also empirically outperforms currently available theoretical convergence guarantees [25,10].

4.2. Proof of Theorem 4.1

To prove Theorem 4.1, our strategy will be to show that the ratio of $\sigma_{\max}(\mathbf{A}_{I_1}\mathbf{P})$ and $\sigma^+_{q-\beta,\min}(\mathbf{A}_{I_1}\mathbf{P})$ is of the same order with high probability. Together with the condition (4.6) for convergence in Lemma 4.2, this implies that the QuantileSCRK method will efficiently converge if the proportion of corruptions is small enough.

First, we show that $\sigma^+_{q-\beta,\min}(\mathbf{A}_{I_1}\mathbf{P})$ can be lower bounded with high probability as long as the effective aspect ratio $(m-m_0)/(n-m_0)$ is tall enough. This is proved using a similar technique as [25, Proposition 3.4].

Lemma 4.3. *Let $\mathbf{P}$ be an orthogonal projection onto a fixed $(n-m_0)$-dimensional subspace, $\alpha \in (0,1]$, and $\mathbf{A}_{I_1} \in \mathbb{R}^{(m-m_0)\times n}$ be a random matrix that satisfies Assumptions A1 and A2. Then there exist absolute constants $C, \theta > 0$ such that if*

$$\frac{m-m_0}{n-m_0} \geq \frac{24}{\alpha}\log\left(\frac{36\theta(1+CK^2)}{\alpha^{3/2}}\right), \tag{4.14}$$

then with probability at least $1-3e^{-\alpha(m-m_0)/24}$,

$$\inf_{\substack{T\subseteq I_1\\ |T|=\alpha(m-m_0)}} \sigma^+_{\min}((\mathbf{A}_{I_1}\mathbf{P})_T) \geq \frac{\alpha^{3/2}}{32}\sqrt{m-m_0}. \tag{4.15}$$

Proof. Recall that $\mathbb{S}^{n-1} = \{\mathbf{x}\in\mathbb{R}^n : \|\mathbf{x}\| = 1\}$ denotes the unit sphere. By Lemma 3.22, we can fix an ε-net $\mathcal{N}$ of $\mathrm{Range}(\mathbf{P})\cap\mathbb{S}^{n-1}$ with cardinality $|\mathcal{N}| \leq (3/\varepsilon)^{n-m_0}$ for some $\varepsilon \in (0,1]$ to be chosen later. Fix any $T \subseteq I_1$ with $|T| = \alpha(m-m_0)$. Since for any $\mathbf{z}\in\mathrm{Range}(\mathbf{P})\cap\mathbb{S}^{n-1}$, there exists $\mathbf{x}\in\mathcal{N}$ such that $\|\mathbf{z}-\mathbf{x}\| \leq \varepsilon$, using the reverse triangle inequality and $\|(\mathbf{A}_{I_1}\mathbf{P})_T\| \leq \|\mathbf{A}_{I_1}\mathbf{P}\|$ implies that

$$\sigma^+_{\min}((\mathbf{A}_{I_1}\mathbf{P})_T) \geq \inf_{\mathbf{z}\in\mathrm{Range}(\mathbf{P})\cap\mathbb{S}^{n-1}}\|(\mathbf{A}_{I_1}\mathbf{P})_T\mathbf{z}\| \geq \inf_{\mathbf{x}\in\mathcal{N}}\|(\mathbf{A}_{I_1}\mathbf{P})_T\mathbf{z}\| - \varepsilon\|\mathbf{A}_{I_1}\mathbf{P}\|. \tag{4.16}$$

Firstly, by Lemma 3.19 (with $s=1$), we have that with probability at least $1-2e^{-(m-m_0)}$,

$$\|\mathbf{A}_{I_1}\mathbf{P}\| \leq (1+CK^2)\sqrt{m-m_0}. \tag{4.17}$$

Next, our goal is to define an event $\mathcal{E}$ on which a good bound for $\inf_{\mathbf{x}\in\mathcal{N}}\|(\mathbf{A}_{I_1}\mathbf{P})_T\mathbf{z}\|$ that is independent of T holds. More precisely, for every $j\in I_1$ and $\mathbf{x}\in\mathcal{N}$, define the "bad" event

$$\mathcal{E}_j^{\mathbf{x}} := \{|\langle\mathbf{a}_j,\mathbf{x}\rangle| \leq \alpha/(4\theta)\},$$

where θ is some constant to be specified later. Let $\mathcal{E}^{\mathbf{x}}$ be the "good" event where less than $\alpha(m-m_0)/2$ of the events $(\mathcal{E}_j^{\mathbf{x}})_{j\in I_1}$ occur, and $\mathcal{E} := \bigcap_{\mathbf{x}\in\mathcal{N}}\mathcal{E}^{\mathbf{x}}$. Observe that $\langle\mathbf{P}\mathbf{a}_j,\mathbf{x}\rangle = \langle\mathbf{a}_j,\mathbf{P}\mathbf{x}\rangle = \langle\mathbf{a}_j,\mathbf{x}\rangle$ since $\mathbf{x}\in\mathrm{Range}(\mathbf{P})$. Therefore, on $\mathcal{E}$, at least half of the rows of $(\mathbf{A}_{I_1}\mathbf{P})_T$ have nontrivial correlation with any $\mathbf{x}\in\mathcal{N}$, which implies that

$$\inf_{\mathbf{x}\in\mathcal{N}}\|(\mathbf{A}_{I_1}\mathbf{P})_T\mathbf{x}\| = \inf_{\mathbf{x}\in\mathcal{N}}\sqrt{\sum_{j\in T}|\langle\mathbf{P}\mathbf{a}_j,\mathbf{x}\rangle|^2} \geq \sqrt{\frac{\alpha(m-m_0)}{2}\cdot\frac{\alpha^2}{16\theta^2}} \geq \frac{\alpha^{3/2}}{6\theta}\sqrt{m-m_0}. \tag{4.18}$$

To balance (4.17) and (4.18), we choose $\varepsilon = \alpha^{3/2}/(12\theta(1+CK^2))$. Therefore, if both events $\mathcal{E}$ and (4.17) hold, then (4.16) implies that the desired bound (4.15) holds.

It remains to bound the probability of the event $\mathcal{E}$, for which we will combine an anti-concentration result with a Chernoff bound. By using either [51, Theorem 1.2] if the row $\mathbf{a}_j$ has independent entries with bounded densities (with $\theta \geq 2\sqrt{2}$), or [7, Theorem 8] if $\mathbf{a}_j$ has a log-concave distribution (increasing the value of θ based on the absolute constant in this result), we deduce that $\mathbb{P}\left(\mathcal{E}_j^{\mathbf{x}}\right) \leq \alpha/4$ for all $j \in I_1$. Hence, a standard Chernoff bound implies that $\mathbb{P}\left(\mathcal{E}^{\mathbf{x}}\right) \geq 1 - e^{-\alpha(m-m_0)/12}$ for all $\mathbf{x} \in \mathcal{N}$, and a union bound shows that $\mathcal{E}$ fails to hold with probability less than

$$|\mathcal{N}| \cdot \exp\left(\frac{-\alpha(m-m_0)}{12}\right) \leq \exp\left((n-m_0)\log\left(\frac{3}{\varepsilon}\right) - \frac{\alpha(m-m_0)}{12}\right) \leq e^{-\alpha(m-m_0)/24},$$

where the condition (4.14) is used for the final inequality. Combining this with the probability bound for (4.17) to hold completes the proof. □

Next, the following lemma bounds $Z_{q-\beta}$ from above with high probability.

Lemma 4.4. *Let $\mathbf{P}$ be an orthogonal projection onto a fixed $(n-m_0)$-dimensional subspace, $\alpha \in (0,1]$, and $\mathbf{A}_{I_1} \in \mathbb{R}^{(m-m_0)\times n}$ be a random matrix that satisfies Assumptions A1 and A2. Then there exists an absolute constant $c > 0$ such that with probability at least $1 - e^{-c\alpha(m-m_0)/K^4}$,*

$$\sup_{\substack{T\subseteq I_1 \\ |T|=\alpha(m-m_0)}} \|(\mathbf{A}_{I_1}\mathbf{P})_T\|_F^2 \leq \left(2 + \frac{K^4}{c}\log\left(\frac{e}{\alpha}\right)\right)\alpha(m-m_0)(n-m_0). \tag{4.19}$$

Proof. For all fixed $T \subseteq I_1$ with $|T| = \alpha(m-m_0)$, Lemma 3.23 applied to the submatrix $(\mathbf{A}_{I_1}\mathbf{P})_T$ implies that there exists an absolute constant $c > 0$ such that for all $\varepsilon > 0$,

$$\mathbb{P}\left(\|(\mathbf{A}_{I_1}\mathbf{P})_T\|_F^2 \geq (1+\varepsilon)\alpha(m-m_0)(n-m_0)\right) \leq e^{-c\min\{\varepsilon^2,\varepsilon\}\alpha(m-m_0)/K^4}.$$

Hence, by a union bound over all $\binom{m-m_0}{\alpha(m-m_0)} < e^{\alpha(m-m_0)\log(e/\alpha)}$ such subsets T, we deduce that the probability that the event (4.19) does not hold is not greater than $\exp\left\{-\alpha(m-m_0)\left(\frac{c\varepsilon}{K^4} - \log\left(\frac{e}{\alpha}\right)\right)\right\}$ for $\varepsilon \geq 1$. In particular, choosing $\varepsilon = 1 + \frac{K^4}{c}\log\left(\frac{e}{\alpha}\right)$ leads to the claimed probability guarantee. □

By combining our tail bounds for $\sigma_{\max}(\mathbf{A}_{I_1}\mathbf{P})$ and $\sigma^+_{q-\beta,\min}(\mathbf{A}_{I_1}\mathbf{P})$ as well as $Z_{q-\beta}$, we can now prove Theorem 4.1.

Proof of Theorem 4.1. In this proof, the various constants of the form $c_1, C_1, \ldots$ that appear only depend on K. By Lemma 3.19, $\sigma_{\max}(\mathbf{A}_{I_1}\mathbf{P}) \leq C_1\sqrt{m-m_0}$ with probability at least $1 - 2e^{-c_1(m-m_0)}$. By Lemma 4.3, $\sigma^+_{q-\beta,\min}(\mathbf{A}_{I_1}\mathbf{P}) \geq C_2(q-\beta)^{3/2}\sqrt{m-m_0}$ with

probability at least $1 - 3e^{-c_2(q-\beta)(m-m_0)}$, given that the condition (4.1) is satisfied. Therefore, if both of these events hold, then

$$\frac{\sigma^+_{q-\beta,\min}(\mathbf{A}_{I_1}\mathbf{P})^2}{\sigma_{\max}(\mathbf{A}_{I_1}\mathbf{P})^2} \geq \left(\frac{C_2}{C_1}\right)^2 (q-\beta)^3.$$

Hence, by Lemma 4.2 we deduce that the QuantileSCRK algorithm converges if

$$\left(\frac{C_2}{C_1}\right)^2 (q-\beta)^3 > \frac{\beta}{1-q} + 2\sqrt{\frac{\beta}{1-q}}.$$

Since q is fixed and the right-hand side can be made arbitrarily small by decreasing β, it follows that this condition is satisfied as long as β is sufficiently small. Finally, Lemma 4.4 implies that $Z_{q-\beta} \leq C_{q,\beta}(m-m_0)(n-m_0)$ with probability at least $1-e^{-c_3(q-\beta)(m-m_0)}$ for some constant $C_{q,\beta} > 0$ that only depends on q, β, and K. If all of these events hold, then Lemma 4.2 implies that QuantileSCRK converges with the claimed rate (4.2). The proof is completed after simplifying the probability bound. □

5. Numerical experiments

In this section, we present numerical experiments that demonstrate various features of the SCRK method (Algorithm 1) and the QuantileSCRK method (Algorithm 2). For the plots with random, simulated data that follow, the lines represent the median over 200 trials, and the shaded regions indicate the 0.1- and 0.9-quantiles around the corresponding medians. The log relative error refers to the quantity $\log(\|\mathbf{x}^k - \mathbf{x}^*\|/\|\mathbf{x}^0 - \mathbf{x}^*\|)$. The experiments were performed on a MacBook Air M1 with 8GB RAM using Python 3.11.

5.1. SCRK method for systems with correlated rows

In Fig. 5.1, we compare the performance of the SCRK method on a system with highly correlated rows for various sizes m_0 of I_0. It is known that RK performs poorly in this setting [45,46]. The entries of $\mathbf{A} \in \mathbb{R}^{2,000\times 1,000}$ are independently and uniformly distributed on $[0.9, 1.1]$, and the solution $\mathbf{x}^* \in \mathbb{R}^{1,000}$ is a standard Gaussian vector. The same initial iterate starting in the solution space corresponding to the biggest block (i.e. $m_0 = 200$) is used for each variation.

As predicted by Corollary 3.12, the SCRK method with I_0 as the first m_0 rows of $\mathbf{A}$ outperforms RK for any $m_0 \geq 1$ since the pairwise row correlations of $\mathbf{A}$ are bounded from below. Moreover, increasing m_0 increases the rate of convergence (see Theorem 3.17). However, since increasing m_0 leads to heavier iterations and a higher initial cost from computing $\mathbf{A}^\dagger_{I_0}$ (see Remark 1.3), the optimal block size for a given target error and time budget is not necessarily the largest as highlighted by Fig. 5.1 (right).

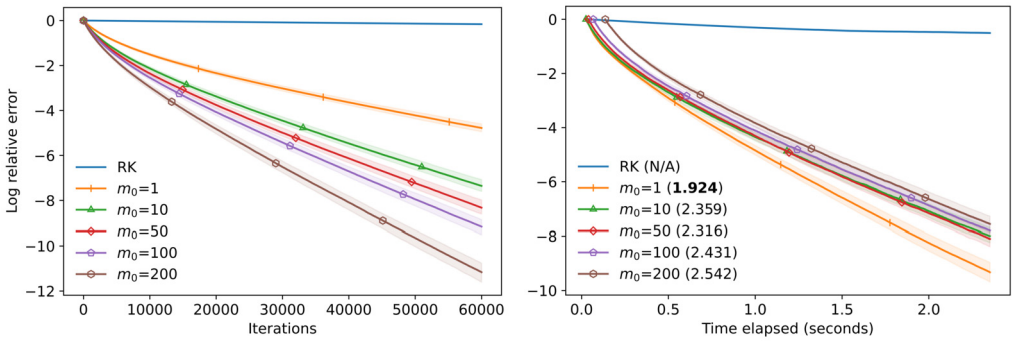


Fig. 5.1. Performance of SCRK on a system with highly correlated rows for various sizes m_0 of $\mathbf{A}_{I_0}$. *(Left)* Log relative error at each iteration. *(Right)* Log relative error against time elapsed, including the initial cost of precomputing $\mathbf{A}_{I_0}^{\dagger}$ for each m_0. The time taken to reach a log relative error of less than -8 is reported in brackets (N/A indicates that this was not reached in 30 seconds).

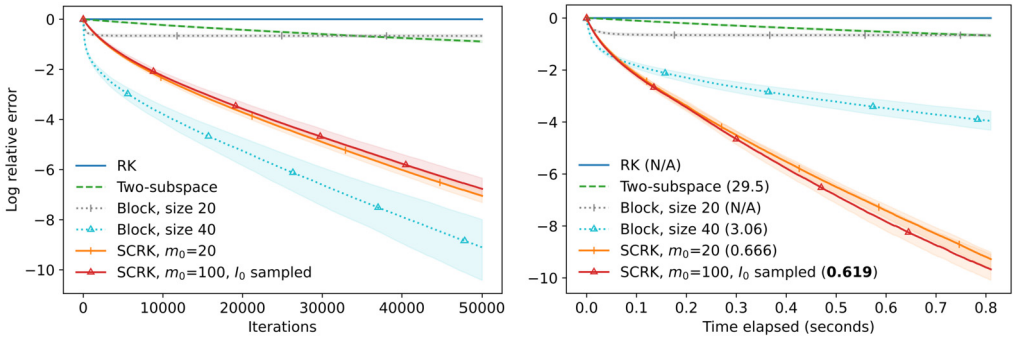


Fig. 5.2. Performance of SCRK on a coherent system with low-rank structure using a "perfect" block (with $m_0 = 20$) and a randomly sampled block (with $m_0 = 100$) as described in the main text. The two-subspace Kaczmarz method [46] and randomized block Kaczmarz method [45] (with two block sizes) are also included. *(Left)* Log relative error at each iteration. *(Right)* Log relative error against time elapsed, not including the initial costs of precomputing pseudoinverses for SCRK and block Kaczmarz. The time taken to reach a log relative error of less than -8 is reported in the brackets (N/A indicates that this was not reached in 30 seconds).

5.2. SCRK method for systems with low-rank structure

In Fig. 5.2, we consider the performance of the SCRK method on a structured matrix $\mathbf{A} \in \mathbb{R}^{2{,}000\times 1{,}000}$, constructed as in Example 3.11. The first $r = 20$ rows of $\mathbf{A}$ are normalized standard Gaussian vectors. The remaining $m - r$ rows $(\mathbf{a}_j)_{j>r}$ are equal to $\mathbf{a}_j := (1-\varepsilon)\mathbf{a}'_j + \varepsilon \mathbf{c}_j$, where $\varepsilon = 0.1$, $\mathbf{a}'_j$ is sampled from $\{\mathbf{a}_1, \dots, \mathbf{a}_r\}$, and $\mathbf{c}_j$ is sampled from $\text{span}(\{\mathbf{a}_1, \dots, \mathbf{a}_r\})^{\perp}$ and normalized; i.e. $\mathbf{a}_j$ mainly consists of a row from the special top block plus some noise in the orthogonal direction. The solution $\mathbf{x}^* \in \mathbb{R}^{1{,}000}$ is a standard Gaussian vector.

The SCRK algorithm is run with two choices of I_0: the first uses the "perfect" block of size $m_0 = 20$ with the rows $\{\mathbf{a}_1, \dots, \mathbf{a}_r\}$ that generate the coherence structure. The second variant uses a block of $m_0 = 5r = 100$ rows of $\mathbf{A}$ sampled (without replacement)

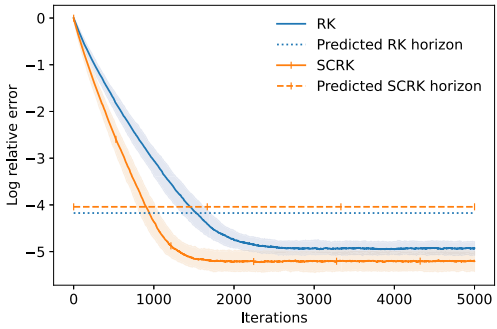


Fig. 5.3. Convergence paths for the SCRK (with I_0 equal to the first $m_0 = 25$ rows) and RK methods on a noisy system. The dashed/dotted lines indicate the predicted error horizons $\gamma_0 + \gamma_1$ from Theorem 3.4 and $\gamma = \|\mathbf{r}\|^2/\sigma_{\min}(\mathbf{A})^2$ from [44] respectively.

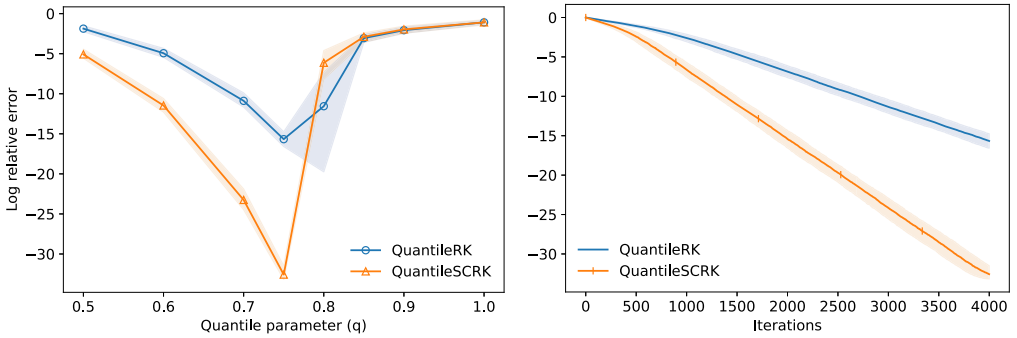
uniformly at random. This represents the case where the source of coherence is unknown, but sampling the effective rank of $\mathbf{A}$ (with an appropriate oversampling factor) should find a good block $\mathbf{A}_{I_0}$ as predicted by Theorem 3.15. Indeed, Fig. 5.2 shows that both choices of I_0 converge effectively: the dramatic improvement in the per-iteration convergence rate of SCRK over RK shown by the left plot is explained by the (inverse) scaled condition number $\sigma^+_{\min}(\mathbf{A}_{I_1}\mathbf{P})/\|\mathbf{A}_{I_1}\mathbf{P}\|_F = 9.33 \times 10^{-3}$ of $\mathbf{A}_{I_1}\mathbf{P}$ with $m_0 = 20$ (and similarly 9.56×10^{-3} with $m_0 = 100$) being significantly larger than the corresponding quantity $\sigma_{\min}(\mathbf{A})/\|\mathbf{A}\|_F = 3.29 \times 10^{-5}$ for $\mathbf{A}$ (see Section 3.3).

The two-subspace Kaczmarz method [46] and randomized block Kaczmarz method [45] (using equally-sized blocks of size 20 and 40 chosen uniformly at random, and precomputed pseudoinverses) are also included. The same initial iterate as SCRK with $m_0 = 20$ is used. It is known that these algorithms perform well in systems with highly correlated rows, such as the one previously considered in Fig. 5.1. However, Fig. 5.2 shows that the effectiveness of two-subspace Kaczmarz and block Kaczmarz with blocks of size 20 that are "too small" is impeded by the coherence structure of $\mathbf{A}$.

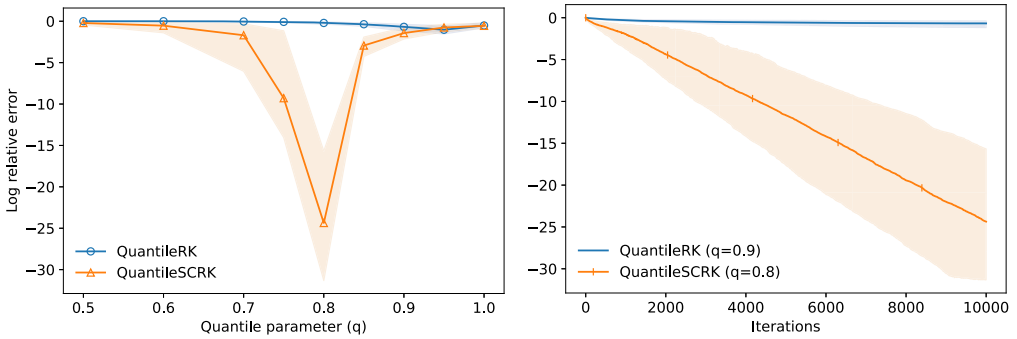
On the other hand, block Kaczmarz with blocks of size 40 that are "large enough" (relative to $r = 20$ for this problem) converges effectively. While Fig. 5.2 (left) shows that it converges with a greater per-iteration rate than SCRK (since it effectively uses 40 new rows in each iteration instead of just one), Fig. 5.2 (right) shows that the lighter iterations of the SCRK method actually make it more efficient on a time basis.

5.3. SCRK method for noisy systems

In Fig. 5.3, we consider the performance of the SCRK algorithm on a noisy system to demonstrate the validity of the error horizon predicted by Theorem 3.4. The rows of $\mathbf{A} \in \mathbb{R}^{300\times 100}$ are independent normalized standard Gaussian vectors, the solution $\mathbf{x}^* \in \mathbb{R}^{100}$ is a standard Gaussian vector, and the entries of the noise vector $\mathbf{r}$ are independently and uniformly distributed on $[-0.01, 0.01]$.



(a) Tall system $\mathbf{A}^{500\times 50}$ with $c = 100$, $m_0 = 20$, $k = 4,000$, and $q_{\mathrm{RK}} = q_{\mathrm{SCRK}} = 0.75$.



(b) Almost-square systems $\mathbf{A}^{130\times 100}$ with $c = 10$, $m_0 = 75$, $k = 10,000$, $q_{\mathrm{RK}} = 0.9$, and $q_{\mathrm{SCRK}} = 0.8$.

Fig. 5.4. Performance of the QuantileSCRK method, given a corruption-free block of size m_0, compared to the QuantileRK method [25] on Gaussian systems with different aspect ratios and c corrupted measurements. *(Left)* Log relative error after k iterations for various values of the quantile parameter q. *(Right)* Convergence paths using the best quantile parameters q_{RK} and q_{SCRK}.

5.4. QuantileSCRK algorithm

In Fig. 5.4, we compare the performance of the QuantileSCRK and QuantileRK [25] methods on Gaussian systems $\mathbf{A}$ with different aspect ratios, where the measurements are corrupted by a sparse vector with c non-zero entries independently and uniformly distributed on $[-1, 1]$. The rows of $\mathbf{A}$ are independent normalized standard Gaussian vectors, and the solution $\mathbf{x}^*$ is a standard Gaussian vector.

Tall systems are considered in Fig. 5.4a, where $100/500 = 20\%$ (resp. $100/480 \approx 20.8\%$) of the rows of $\mathbf{A}$ (resp. $\mathbf{A}_{I_1}$) correspond to corrupted measurements. These plots replicate the finding that the QuantileRK method converges effectively for tall, Gaussian-like matrices even in the presence of numerous corruptions [25], and also show that exploiting information about corruption-free measurements using the QuantileSCRK method accelerates convergence (see Theorem 4.1).

Almost-square systems are considered in Fig. 5.4b, where $10/130 \approx 7.7\%$ (resp. $10/55 \approx 18.2\%$) of the rows of $\mathbf{A}$ (resp. $\mathbf{A}_{I_1}$) correspond to corrupted measurements. It is clear that the QuantileRK method is unable to make any progress in this setting. On the

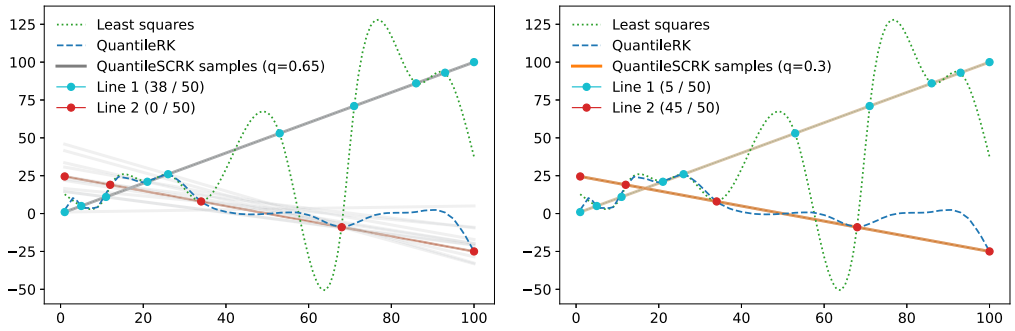


Fig. 5.5. Solving a discretized differential equation for a line in the plane given two sets of inconsistent initial conditions as described in the main text. *(Left)* 50 outputs after 10,000 iterations of QuantileSCRK with $q = 0.65$ (translucent gray lines); 38 out of the 50 converged to Line 1 (in the sense $\|\mathbf{A}_{\text{Line 1}}\mathbf{x}^k - \mathbf{b}_{\text{Line 1}}\|_2 < 10^{-3}$). *(Right)* 50 outputs after 10,000 iterations of the QuantileSCRK algorithm with $q = 0.3$ (translucent orange lines); 45 out of the 50 converged to Line 2.

other hand, the QuantileSCRK method converges for q around 0.8, which demonstrates that exploiting external knowledge in the form of a large block $\mathbf{A}_{I_0}$ corresponding to corruption-free measurements can enable convergence in such challenging settings.

5.5. Systems of differential equations with inconsistent initial conditions

We consider the problem of numerically solving a system of differential equations given competing data for the initial conditions as another application of the QuantileSCRK method. After discretization via a finite difference scheme, two types of equations arise: the first describe the underlying law and can be considered to be known exactly, and the second type encode the initial conditions, which can be obtained from real data with potentially faulty measurements. Thus, the problem can be viewed as one about detecting and disregarding the "corrupted" equations coming from inconsistent initial conditions, given that the majority of the equations of the first type can be "trusted".

In Fig. 5.5, we consider the linear system obtained from discretizing the differential equation $y'' = 0$ for a line illustrate this idea. The top 98×100 block is a Toeplitz matrix with entries $1, -2, 1$ along the diagonal before normalization, which we take to be $\mathbf{A}_{I_0}$. We consider two sets of initial conditions corresponding to two lines: Line 1 being $y = x$ with 10 initial conditions, and Line 2 being $y = 25 - x/2$ with 5 initial conditions.

The plots show that solving this system using least squares or QuantileRK produces poor solutions. However, using QuantileSCRK with a careful choice of the quantile parameter enables convergence to one line or the other as the algorithm is able to find a set of consistent initial conditions: when $q = 0.65$, QuantileSCRK converges to Line 1 a majority of the time (left), and when $q = 0.3$, QuantileSCRK converges to Line 2 instead (right). We also observed that the initial iterate $\mathbf{x}^0$ has a significant biasing effect on which solution is preferred for convergence.

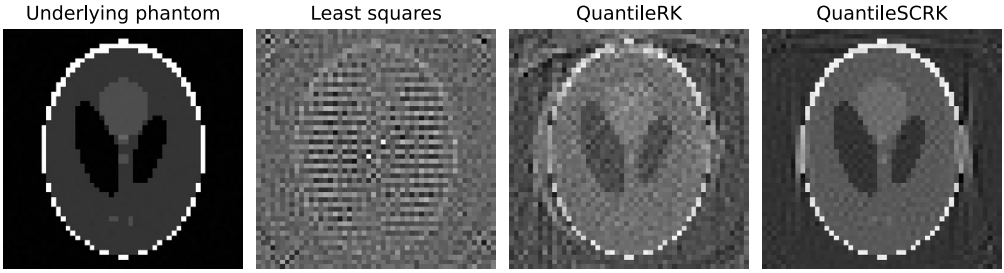


Fig. 5.6. Reconstructions of the Shepp-Logan phantom from $\mathbf{A} \in \mathbb{R}^{4,500\times 2,500}$ and $\widetilde{\mathbf{b}} \in \mathbb{R}^{4,500}$ with $c = 1,125$ corruptions as described in the main text. The QuantileSCRK method, given a corruption-free block $\mathbf{A}_{I_0}$ of size $m_0 = 500$, and the QuantileRK method were both run using $q = 0.7$ for $60m = 270,000$ iterations, obtaining a final ℓ_2 error $\|\mathbf{x}^k - \mathbf{x}\|_2$ of 3.47 and 6.85 respectively.

5.6. CT image reconstruction

Finally, we investigate the performance of QuantileSCRK on a realistic dataset. We consider the Shepp-Logan phantom, generated using the Air Tools II package [27] with parameters $N = 50$ (the image is $N \times N$), $\theta = \{0, 2, 4, \ldots, 178\}$ (angles used), and $\rho = 50$ (number of parallel rays). The image is encoded by the measurement matrix $\mathbf{A} \in \mathbb{R}^{4,500\times 2,500}$ and measurements $\mathbf{b} \in \mathbb{R}^{4,500}$. A subset I_0 of $m_0 = 500$ rows of $\mathbf{A}$ was randomly chosen to be corruption-free (e.g. corresponding to trustworthy measurements), and a random set of $c = 1,125$ of the remaining measurements were corrupted by quantities uniformly distributed in $[2, 6]$ to produce the corrupted measurements $\widetilde{\mathbf{b}}$.

In Fig. 5.6, we show various reconstructions given $\mathbf{A}$ and the corrupted measurements $\widetilde{\mathbf{b}}$. It is clear that the least squares solution of the corrupted linear system is very poor. The QuantileRK method, initialized from zero, achieves a noisy reconstruction that does not recover the fine details. Using the QuantileSCRK method with the corruption-free block $\mathbf{A}_{I_0}$ achieves the best reconstruction, even though a significant proportion (25%) of the measurements have been corrupted.

6. Conclusion and future directions

In this paper, we introduced the subspace constrained randomized Kaczmarz (SCRK) method for solving consistent, overdetermined systems of linear equations $\mathbf{Ax} = \mathbf{b}$, which provides a framework for studying the dynamics of the randomized Kaczmarz algorithm when the iterates are confined to a selected solution space $\mathbf{A}_{I_0}\mathbf{x} = \mathbf{b}_{I_0}$. We described the convergence rate of the SCRK method in terms of the spectral properties of the matrix $\mathbf{A}_{I_1}\mathbf{P}$, where $\mathbf{P}$ is the orthogonal projector onto $\mathrm{Null}(\mathbf{A}_{I_0})$. We also demonstrated, both theoretically and empirically, how the SCRK method can exploit approximately low-rank structure to accelerate convergence.

We also proposed the QuantileSCRK method for solving corrupted linear systems, which is able to exploit external knowledge about corruption-free subsystems. In addition to theoretical convergence analysis, we demonstrated numerically that it is able to

converge for almost-square corrupted linear systems, where existing iterative methods are ineffective, and that it can be useful for solving differential equations with inconsistent initial conditions and image reconstruction from significantly corrupted measurements.

The framework of subspace constrained iterations raises many possible future directions. For example, since our analysis showed that the SCRK updates simplify to a version of the usual Kaczmarz updates with skewed step directions and the projector $\mathbf{P}$ acts as a right preconditioner for $\mathbf{A}$ to improve the convergence rate, it seems plausible that similar ideas could be applied to related solvers such as the sketch-and-project algorithm [21] or iterative projection methods for solving systems of linear inequalities [34,5]. It would also be interesting to develop and analyze a QuantileSCRK method in which the trusted solution space is built up adaptively in a data-driven way, based on the information accumulated during the iteration process, which could lead to an effective way for solving corrupted linear systems even in the absence of external knowledge.

Declaration of competing interest

There is no competing interest.

Data availability

Data will be made available on request.

Acknowledgements

The authors acknowledge partial support by NSF DMS-2309685. ER was also partially funded by NSF DMS-2108479. ER is grateful to Y. Kevrekidis and M. Dereziński for the helpful and inspiring discussions. We would like to thank the referees for their valuable comments and suggestions, which have improved the clarity and exposition of the paper.

References

[1] A. Agaskar, C. Wang, Y.M. Lu, Randomized Kaczmarz algorithms: exact MSE analysis and optimal sampling probabilities, in: IEEE Global Conference on Signal and Information Processing (GlobalSIP), 2014, pp. 389–393.
[2] E. Amaldi, V. Kann, The complexity and approximability of finding maximum feasible subsystems of linear relations, Theor. Comput. Sci. 147 (1–2) (1995) 181–210, https://doi.org/10.1016/0304-3975(94)00254-G.
[3] E. Amaldi, P. Belotti, R. Hauser, Randomized relaxation methods for the maximum feasible subsystem problem, in: Integer Programming and Combinatorial Optimization, Springer, Berlin, Heidelberg, 2005, pp. 249–264.
[4] C. Boutsidis, M.W. Mahoney, P. Drineas, An improved approximation algorithm for the column subset selection problem, in: ACM-SIAM Symposium on Discrete Algorithms (SODA), 2009, pp. 968–977.
[5] J. Briskman, D. Needell, Block Kaczmarz method with inequalities, J. Math. Imaging Vis. 52 (2015) 385–396, https://doi.org/10.1007/s10851-014-0539-7.
[6] E.J. Candes, M. Rudelson, T. Tao, R. Vershynin, Error correction via linear programming, in: IEEE Symposium on Foundations of Computer Science (FOCS), 2006, pp. 668–681.

[7] A. Carbery, J. Wright, Distributional and L^q norm inequalities for polynomials over convex bodies in $\mathbb{R}^n$, Math. Res. Lett. 8 (2001) 233–248, https://doi.org/10.4310/MRL.2001.v8.n3.a1.
[8] C. Cenker, H.G. Feichtinger, M. Mayer, H. Steier, T. Strohmer, New variants of the POCS method using affine subspaces of finite codimension with applications to irregular sampling, in: Visual Communications and Image Processing '92, vol. 1818, 1992, pp. 299–310.
[9] X. Chen, J. Qin, Regularized Kaczmarz algorithms for tensor recovery, SIAM J. Imaging Sci. 14 (4) (2021) 1439–1471, https://doi.org/10.1137/21M1398562.
[10] L. Cheng, B. Jarman, D. Needell, E. Rebrova, On block accelerations of quantile randomized Kaczmarz for corrupted systems of linear equations, Inverse Probl. 39 (2) (2023) 024002, https://doi.org/10.1088/1361-6420/aca78a.
[11] A. Dalalyan, P. Thompson, Outlier-robust estimation of a sparse linear model using ℓ_1-penalized Huber's M-estimator, in: Advances in Neural Information Processing Systems, vol. 32, 2019.
[12] J.A. De Loera, J. Haddock, D. Needell, A sampling Kaczmarz–Motzkin algorithm for linear feasibility, SIAM J. Sci. Comput. 39 (5) (2017) S66–S87, https://doi.org/10.1137/16M1073807.
[13] M. Dereziński, E. Rebrova, Sharp analysis of sketch-and-project methods via a connection to randomized singular value decomposition, SIAM J. Math. Data Sci. 6 (1) (2024) 127–153, https://doi.org/10.1137/23M1545537.
[14] M. Dereziński, J. Yang, Solving dense linear systems faster than via preconditioning, in: ACM Symposium on Theory of Computing (STOC), 2024.
[15] M. Dereziński, D. LeJeune, D. Needell, E. Rebrova, Fine-grained analysis and faster algorithms for iteratively solving linear systems, arXiv preprint, arXiv:2405.05818, 2024.
[16] P. Drineas, R. Kannan, M.W. Mahoney, Fast Monte Carlo: algorithms for matrices II: computing a low-rank approximation to a matrix, SIAM J. Comput. 36 (1) (2006) 158–183, https://doi.org/10.1137/S0097539704442696.
[17] P. Drineas, M.W. Mahoney, S. Muthukrishnan, Relative-error CUR matrix decompositions, SIAM J. Matrix Anal. Appl. 30 (2) (2008) 844–881, https://doi.org/10.1137/07070471X.
[18] P. Drineas, M. Magdon-Ismail, M.W. Mahoney, D.P. Woodruff, Fast approximation of matrix coherence and statistical leverage, J. Mach. Learn. Res. 13 (1) (2012) 3475–3506, https://doi.org/10.5555/2503308.2503352.
[19] Y.-S. Du, K. Hayami, N. Zheng, K. Morikuni, J.-F. Yin, Kaczmarz-type inner-iteration preconditioned flexible GMRES methods for consistent linear systems, SIAM J. Sci. Comput. 43 (5) (2021) S345–S366, https://doi.org/10.1137/20M1344937.
[20] T. Elfving, Block-iterative methods for consistent and inconsistent linear equations, Numer. Math. 35 (1980) 1–12, https://doi.org/10.1007/BF01396365.
[21] R.M. Gower, P. Richtárik, Randomized iterative methods for linear systems, SIAM J. Matrix Anal. Appl. 36 (4) (2015) 1660–1690, https://doi.org/10.1137/15M1025487.
[22] R.M. Gower, D. Kovalev, F. Lieder, P. Richtárik, RSN: randomized subspace Newton, in: Advances in Neural Information Processing Systems, vol. 32, 2019.
[23] J. Haddock, A. Ma, Greed works: an improved analysis of sampling Kaczmarz–Motzkin, SIAM J. Math. Data Sci. 3 (1) (2021) 342–368, https://doi.org/10.1137/19M1307044.
[24] J. Haddock, D. Needell, Randomized projection methods for linear systems with arbitrarily large sparse corruptions, SIAM J. Sci. Comput. 41 (5) (2019) S19–S36, https://doi.org/10.1137/18M1179213.
[25] J. Haddock, D. Needell, E. Rebrova, W. Swartworth, Quantile-based iterative methods for corrupted systems of linear equations, SIAM J. Matrix Anal. Appl. 43 (2) (2022) 605–637, https://doi.org/10.1137/21M1429187.
[26] N. Halko, P.-G. Martinsson, J.A. Tropp, Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions, SIAM Rev. 53 (2) (2011) 217–288, https://doi.org/10.1137/090771806.
[27] P. Hansen, J. Jørgensen, AIR tools II: algebraic iterative reconstruction methods, improved implementation, Numer. Algorithms 79 (2018) 107–137, https://doi.org/10.1007/s11075-017-0430-x.
[28] G.T. Herman, Fundamentals of Computerized Tomography. Advances in Computer Vision and Pattern Recognition, Springer-Verlag, London, 2009.
[29] R.A. Horn, C.R. Johnson, Matrix Analysis, 2 edition, Cambridge University Press, Cambridge, 2012.
[30] B. Jarman, D. Needell, QuantileRK: solving large-scale linear systems with corrupted, noisy data, in: 55th Asilomar Conference on Signals, Systems, and Computers, 2021, pp. 1312–1316.
[31] C. Jin, P. Netrapalli, R. Ge, S.M. Kakade, M.I. Jordan, A short note on concentration inequalities for random vectors with SubGaussian norm, arXiv preprint, arXiv:1902.03736, 2019.

[32] S. Kaczmarz, Angenäherte Auflösung von Systemen linearer Gleichungen, Bull. Int. Acad. Pol. Sci. Lett. 35 (1937) 355–357.
[33] J.A. Kelner, L. Orecchia, A. Sidford, Z.A. Zhu, A simple combinatorial algorithm for solving SDD systems in nearly-linear time, in: ACM Symposium on Theory of Computing (STOC), 2013, pp. 911–920.
[34] D. Leventhal, A.S. Lewis, Randomized methods for linear constraints: convergence rates and conditioning, Math. Oper. Res. 35 (3) (2010) 641–654, https://doi.org/10.1287/moor.1100.0456.
[35] J. Liu, S.J. Wright, An accelerated randomized Kaczmarz algorithm, Math. Comput. 85 (297) (2016) 153–178, https://doi.org/10.1090/mcom/2971.
[36] D.A. Lorenz, S. Wenger, F. Schöpfer, M.A. Magnor, A sparse Kaczmarz solver and a linearized Bregman method for online compressed sensing, in: IEEE International Conference on Image Processing (ICIP), 2014, pp. 1347–1351.
[37] D.A. Lorenz, S. Rose, F. Schöpfer, The randomized Kaczmarz method with mismatched adjoint, BIT Numer. Math. 58 (2018) 1079–1098, https://doi.org/10.1007/s10543-018-0717-x.
[38] A. Ma, D. Needell, A. Ramdas, Convergence properties of the randomized extended Gauss–Seidel and Kaczmarz methods, SIAM J. Matrix Anal. Appl. 36 (4) (2015) 1590–1604, https://doi.org/10.1137/15M1014425.
[39] M.W. Mahoney, Randomized algorithms for matrices and data, Found. Trends Mach. Learn. 3 (2) (2011) 123–224, https://doi.org/10.1561/2200000035.
[40] M. Meier, Y. Nakatsukasa, Fast randomized numerical rank estimation for numerically low-rank matrices, Linear Algebra Appl. 686 (2024) 1–32, https://doi.org/10.1016/j.laa.2024.01.001.
[41] F. Natterer, The Mathematics of Computerized Tomography, Classics in Applied Mathematics, Society for Industrial and Applied Mathematics, 2001.
[42] I. Necoara, Faster randomized block Kaczmarz algorithms, SIAM J. Matrix Anal. Appl. 40 (4) (2019) 1425–1452, https://doi.org/10.1137/19M1251643.
[43] I. Necoara, M. Takáč, Randomized sketch descent methods for non-separable linearly constrained optimization, IMA J. Numer. Anal. 41 (2) (2021) 1056–1092, https://doi.org/10.1093/imanum/draa018.
[44] D. Needell, Randomized Kaczmarz solver for noisy linear systems, BIT Numer. Math. 50 (2010) 395–403, https://doi.org/10.1007/s10543-010-0265-5.
[45] D. Needell, J.A. Tropp, Paved with good intentions: analysis of a randomized block Kaczmarz method, Linear Algebra Appl. 441 (2014) 199–221, https://doi.org/10.1016/j.laa.2012.12.022.
[46] D. Needell, R. Ward, Two-subspace projection method for coherent overdetermined systems, J. Fourier Anal. Appl. 19 (2013) 256–269, https://doi.org/10.1007/s00041-012-9248-z, Associated technical report, arXiv:1204.0279.
[47] D. Needell, R. Zhao, A. Zouzias, Randomized block Kaczmarz method with projection for solving least squares, Linear Algebra Appl. 484 (2015) 322–343, https://doi.org/10.1016/j.laa.2015.06.027.
[48] D. Needell, N. Srebo, R. Ward, Stochastic gradient descent, weighted sampling, and the randomized Kaczmarz algorithm, Math. Program. 155 (2016) 549–573, https://doi.org/10.1007/s10107-015-0864-7.
[49] S. Oymak, J.A. Tropp, Universality laws for randomized dimension reduction, with applications, Inf. Inference 7 (3) (2018) 337–446, https://doi.org/10.1093/imaiai/iax011.
[50] E. Rebrova, D. Needell, On block Gaussian sketching for the Kaczmarz method, Numer. Algorithms 86 (1) (2021) 443–473, https://doi.org/10.1007/s11075-020-00895-9.
[51] M. Rudelson, R. Vershynin, Small ball probabilities for linear images of high-dimensional distributions, Int. Math. Res. Not. 2015 (19) (2015) 9594–9617, https://doi.org/10.1093/imrn/rnu243.
[52] A. Saumard, J.A. Wellner, Log-concavity and strong log-concavity: a review, Stat. Surv. 8 (2014) 45–114, https://doi.org/10.1214/14-SS107.
[53] F. Schöpfer, D. Lorenz, Linear convergence of the randomized sparse Kaczmarz method, Math. Program. 173 (2019) 509–536, https://doi.org/10.1007/s10107-017-1229-1.
[54] F. Schöpfer, D.A. Lorenz, L. Tondji, M. Winkler, Extended randomized Kaczmarz method for sparse least squares and impulsive noise problems, Linear Algebra Appl. 652 (2022) 132–154, https://doi.org/10.1016/j.laa.2022.07.003.
[55] S. Steinerberger, Randomized Kaczmarz converges along small singular vectors, SIAM J. Matrix Anal. Appl. 42 (2) (2021) 608–615, https://doi.org/10.1137/20M1350947.
[56] S. Steinerberger, Quantile-based random Kaczmarz for corrupted linear systems of equations, Inf. Inference 12 (1) (2023) 448–465, https://doi.org/10.1093/imaiai/iaab029.
[57] T. Strohmer, R. Vershynin, A randomized Kaczmarz algorithm with exponential convergence, J. Fourier Anal. Appl. 15 (2009) 262–278, https://doi.org/10.1007/s00041-008-9030-4.

[58] C. Studer, P. Kuppinger, G. Pope, H. Bolcskei, Recovery of sparsely corrupted signals, IEEE Trans. Inf. Theory 58 (5) (2012) 3115–3130, https://doi.org/10.1109/TIT.2011.2179701.
[59] Y.S.T. Tan, R. Vershynin, Phase retrieval via randomized Kaczmarz: theoretical guarantees, Inf. Inference 8 (1) (2019) 97–123, https://doi.org/10.1093/imaiai/iay005.
[60] R. Vershynin, High-Dimensional Probability: An Introduction with Applications in Data Science, Cambridge University Press, Cambridge, 2018.
[61] W.-T. Wu, On two-subspace randomized extended Kaczmarz method for solving large linear least-squares problems, Numer. Algorithms 89 (2022) 1–31, https://doi.org/10.1007/s11075-021-01104-x.
[62] A. Zouzias, N.M. Freris, Randomized extended Kaczmarz for solving least squares, SIAM J. Matrix Anal. Appl. 34 (2) (2013) 773–793, https://doi.org/10.1137/120889897.