FINITE-PRECISION GRAM-SCHMIDT WALKS

Anonymous authors
Paper under double-blind review

ABSTRACT

The Gram–Schmidt Walk is a randomized vector-balancing algorithm whose sub-gaussian guarantees support applications in discrepancy, experimental design, and data compression; however, these theoretical guarantees are established in exact arithmetic, whereas implementations must approximate least-squares directions, boundary updates, and sampling probabilities in finite precision. This is important because small numerical errors can change which coordinates freeze and thereby alter the subsequent trajectory. We analyze the concentration of the perturbed Gram-Schmidt walk directly under bounded, potentially biased and history-dependent errors. For n input vectors of Euclidean norm at most one, we obtain a modified MGF bound depending on key error sources which recovers the original result as the error goes to zero. We also construct a full-column-rank instance in which bounded update errors produce bias of order $\min\{n^2\varepsilon_{\text{iter}}, n\}$, showing that updates accumulate error unavoidably under this model. Finally, we validate our findings in a variety of settings by ablating on the bit precision and problem size.

1 INTRODUCTION

Selecting a small representative subset of a dataset is an important task in algorithmic statistics (Xiao et al., 2026; Göttlich et al., 2024; Chazelle, 2000), which requires preserving many statistics with relatively few points. Discrepancy theory offers a useful perspective on this task: assign signs to data vectors so that their signed sum is small. Given some matrix $A \in \mathbb{R}^{m\times n}$, we seek to assign weights $\{1, -1\}$ to its columns such that the sum is minimized in the infinity norm. In other words, we ask how to split the columns into two groups such that both groups have roughly the same sum as measured by the infinity norm of the difference. We call this quantity the discrepancy of the matrix,

$$\operatorname{disc}(A) = \min_{x\in\{-1,1\}^n} \|Ax\|_\infty \tag{1}$$

A balanced partition can serve as a compression step if its cardinality or weights are controlled, and finding such a coloring has natural applications in machine learning such as coresets for large datasets (Karnin and Liberty, 2019; Liberty et al., 2026), attention approximation (Kochetkova et al., 2026), covariate balance in experimental design (Harshaw et al., 2024), and distribution thinning (Carrell et al., 2025), where one wishes to summarize a large dataset with a few carefully chosen points. In theoretical computer science, this problem has found applications to spectral sparsification (Lau et al., 2025), approximation algorithms (Eisenbrand et al., 2012), computational geometry (Chazelle, 2004), and Monte Carlo integration (Dick and Pillichshammer, 2014). Intuitively, if we can split a dataset into two balanced parts where the whole dataset can be summarized by one of the halves, then we can recursively repeat this procedure to find a small approximation of the dataset.

One such algorithm is the Gram-Schmidt Walk of Bansal et al. (2017) which recovers a seminal result of Banaszczyk (1998) by allowing one to sample x from a certain distribution such that Ax concentrates tightly around the zero vector. Given $B = [b_1, \ldots, b_n] \in \mathbb{R}^{m\times n}$ with $\|b_i\|_2 \le 1$ and an initial fractional assignment $z_0 \in [-1, 1]^n$, the walk returns $z \in \{1, -1\}^n$. At a high level, the algorithm takes a random walk in the unit cube $[-1, 1]^n$ with entries becoming integral as the algorithm progresses. Once an entry becomes integral we say that it is frozen, with future steps not moving along this dimension. The steps of the random walk are chosen to be mean zero while not increasing discrepancy too much among the alive entries such that the final output has small discrepancy. Harshaw et al. (2024) showed that for every $v \in \mathbb{R}^m$,

$$\mathbb{E}\left[\exp\left(\langle B(z - z_0), v\rangle\right)\right] \le \exp\left(\frac{1}{2}\|v\|_2^2\right). \tag{2}$$

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Thus every fixed linear projection of the discrepancy has Gaussian-type tails with a dimension-independent scale. This property is also useful as a component of more elaborate algorithms. For instance, Carrell et al. (2025) develop Gram-Schmidt-based thinning procedures and implementations that accelerate repeated balancing steps.

The concentration guarantees above are established for an idealized Gram–Schmidt Walk operating in exact arithmetic: specifically, the least-squares directions and boundary steps are computed exactly, and each random choice follows its prescribed probability, and these guarantees do not automatically extend to the finite precision process which would be executed on a computer that uses a small number of bits. In particular, these aforementioned theoretical guarantees do not quantify how much numerical error the walk can tolerate due to the sensitivity of its movement within the hypercube, or how much working precision is sufficient to preserve the desired discrepancy bound. This motivates a finite-precision analysis that explains how numerical errors affect concentration and provides sufficient conditions for preserving a desired discrepancy guarantee.

Indeed, implementing these steps in practice introduces errors due to finite-precision. These errors can manifest in the computed direction, the boundary update, and the probabilities of hitting each boundary of the unit hypercube, and can affect the current iterate; however, more seriously, they can also affect which coordinate freezes, which in turn changes the least-squares system (which is also solved by maintaining a Cholesky factorization for improved runtimes) that must later be solved. Therefore, in the worst-case, the perturbed trajectory can deviate significantly from the ideal trajectory. It is not clear a priori that subgaussianity should be preserved under approximate arithmetic, as small perturbations in position can potentially change which entries become frozen in a given iteration, changing the path that the walk takes over the rest of the algorithm. So, any analysis that follows the perturbed trajectory alongside one exact trajectory may lose the structure responsible for concentration. Moreover, deterministic rounding may not preserve conditional mean zero, so an implementation can accumulate systematic discrepancy even when the individual update errors are small. Numerical sensitivity is already familiar from classical Gram-Schmidt orthogonalization, which can suffer severe loss of orthogonality in finite precision, especially for nearly linearly dependent inputs (Giraud et al., 2005). The Gram-Schmidt walk introduces a further challenge: numerical errors can affect both the computed directions and the branching and freezing decisions, changing the subsequent trajectory and the distribution of the final coloring. Understanding its numerical robustness therefore requires connecting local computational errors to the concentration of the final discrepancy.

These observations expose a gap between theory and practice for the Gram–Schmidt Walk. The classical concentration guarantees assume exact arithmetic, while practical implementations use approximate least-squares directions, rounded updates, and inexact sampling probabilities. These errors can change which coordinates freeze and hence which subproblems the algorithm solves next, so the exact-arithmetic guarantees need not hold for the resulting finite-precision output distribution.

We address this gap by studying a model with updates $z_{t+1} = z_t + \hat{\delta}_t \hat{u}_t + \varepsilon_t$ for $\|\varepsilon_t\|_\infty \le \varepsilon_{\mathrm{iter}}$ where $\hat{u}_t$ and $\hat{\delta}_t$ are computed quantities, and ε_t includes rounding, clamping, and freezing corrections. Uniform budgets ε_{LS} and $\varepsilon_{\mathrm{pr}}$ additionally control least-squares image error and conditional branch-probability error, respectively. Our analysis permits biased, history-dependent errors and changes to the freezing trajectory. Section 3 specifies the complete model, including feasibility and freezing requirements. Our contributions address the theory-practice gap in two directions:

(1) **Theoretical guarantees.** We establish concentration guarantees for the Gram-Schmidt Walk under biased, history-dependent perturbations. The key technical challenge is that numerical errors can change freezing decisions and subsequent least-squares directions, while destroying the exact conditional unbiasedness used in prior analyses. We address these interacting effects through a phase-based analysis that controls accumulated conditional drift without requiring proximity to an ideal execution. While our bounds retain the sharp exact-arithmetic quadratic term in the log moment-generating function, we further reveal how the discrepancy quality depends on arithmetic precision, the bit complexity, and the error of the least-squares solver.

(2) **Empirical validation.** Our experiments illustrate the precision-dependent discrepancy in practical numerical implementations. We show that finite precision can *significantly* degrade the discrepancy achieved by the Gram-Schmidt Walk, and characterize the sensitivity of this degradation on the precision and matrix conditioning. By measuring empirical discrepancy tails, we evaluate when approximate arithmetic preserves output quality and when it is a limitation.

1.1 Related Works

Gram-Schmidt walk. Bansal et al. (2017) introduced the Gram-Schmidt walk as an efficient constructive approach to Banaszczyk's vector balancing theorem, obtaining a distribution over colorings with subgaussian discrepancy (Banaszczyk, 1998; Dadush et al., 2016; 2018). Harshaw et al. (2024) subsequently established the sharp 1-subgaussian guarantee and developed the walk's application to covariate balance in randomized experiments. These analyses exploit the geometry of the least-squares directions and the conditional mean-zero property of the updates. In contrast, we study how this guarantee changes under biased, history-dependent perturbations. Such errors can change which coordinate freezes, changing the linear systems solved in future iterations. We analyze the realized perturbed trajectory directly, preserving the guarantee due to the exact-arithmetic in the log moment-generating-function bound but also revealing an additional term due to accumulated error.

Algorithmic discrepancy theory. Several lines of work develop efficient randomized constructions for discrepancy minimization. Alweiss et al. (2020) introduced a self-balancing walk with guarantees for online vector balancing (Spencer, 1985; Bansal et al., 2020b; Kulkarni et al., 2024) and the Komlós problem (Spencer, 1994). Bansal et al. (2022) give a stochastic-process framework that recovers several discrepancy bounds, while Deng et al. (2025) accelerate discrepancy minimization through faster implementations of Edge-Walk and hereditary-discrepancy algorithms. Our focus is the arithmetic accuracy needed to preserve the output concentration of a particular balancing algorithm. A different notion of stability appears in Gamarnik et al. (2023), who establish barriers for algorithms whose outputs change little under perturbations of random input instances. Their input-stability condition concerns sensitivity of the selected coloring; our analysis permits changes in the trajectory and terminal coloring while controlling the resulting discrepancy distribution.

Finite precision and numerical stability. Classical numerical analysis relates roundoff error, backward error, and conditioning for linear-system and least-squares computation (Higham, 2002). Recent works makes precision requirements explicit in algorithms that repeatedly update linear-algebraic quantities. Ghadiri et al. (2024; 2023; 2026) analyze the backward stability of inverse maintenance and distributed optimization, and derive bit-complexity bounds for regression and linear programming. Diem (2011); Pan (1985); Anand et al. (2024) extend this perspective to dynamic matrix expressions and their determinants. For implementations based on inverse updates, Ma et al. (2026) provide related forward- and backward-error bounds for the Sherman–Morrison–Woodbury formula evaluated with approximate inverses. These results concern the accuracy and cost of numerical subroutines. The Gram–Schmidt Walk additionally requires controlling how these errors interact with adaptive freezing decisions and randomized updates. Our analysis addresses an additional question of how local numerical errors affect concentration when they also change adaptive freezing decisions and conditional update distributions, connecting the working precision and conditioning to conditional drift, accumulated discrepancy error, and ultimately concentration of the terminal coloring, without requiring exact conditional unbiasedness.

Discrepancy in learning and data compression. Karnin and Liberty (2019) connect low class discrepancy to small offline and streaming coresets. Particularly close to our setting, Carrell et al. (2025) analyze subgaussian thinning, introduce Gram–Schmidt-based thinning procedures, and give a cubic-time halving implementation based on local inverse updates. Their framework relates concentration of the balancing primitive to compression quality and supports applications to attention approximation, gradient reordering, and two-sample testing. Discrepancy also informs neural-network weight quantization (Chee et al., 2025) and attention mechanisms (Kochetkova et al., 2026). These applications motivate understanding the precision requirements of balancing algorithms. Our contribution provides a concentration guarantee for an implementation of the Gram–Schmidt primitive. In turn, these guarantees are flexible and can be converted into guarantees for repeated compression or any learning pipeline by propagating the resulting error budget through that procedure.

2 Preliminaries and Notation

We write $[n]$ for $\{1, ..., n\}$. For any vector $v \in \mathbb{R}^n$ and $i \in [n]$ we write $v(i)$ for the ith entry.

We first review a few key properties of subgaussian random variables and random vectors.

Fact 2.1 (Characterization of subgaussian random variables)**.** *For a real-valued random variable X, the following are equivalent up to constant factors in the parameter K_i (Vershynin, 2026):*

Algorithm 1: Gram–Schmidt walk (Harshaw et al., 2024)

Initialize $t \leftarrow 1$ and $z_1 \leftarrow z_0$
Choose $p \in [n]$ uniformly from $A_1 := \{i : |z_1(i)| < 1\}$

while $z_t \notin \{\pm 1\}^n$ **do**
  Create the active set $A_t \leftarrow \{i \in [n] : |z_t(i)| < 1\}$
  if $p \notin A_t$ **then**
    Sample p uniformly at random from A_t
  Let $p_t := p$ and let $U_{p_t} \leftarrow \{u \in \mathbb{R}^n : u(p_t) = 1,\ u(i) = 0 \text{ for all } i \notin A_t\}$
  Set $u_t \leftarrow \arg\min_{u \in U_{p_t}} \|\boldsymbol{B}u\|_2^2$
  Let $[-\delta_t^-, \delta_t^+] = \{\delta \in \mathbb{R} : z_t + \delta u_t \in [-1,1]^n\}$
  Sample δ_t according to $\delta_t = \begin{cases} \delta_t^+, & \text{with probability } \dfrac{\delta_t^-}{\delta_t^- + \delta_t^+}, \\ -\delta_t^-, & \text{with probability } \dfrac{\delta_t^+}{\delta_t^- + \delta_t^+}. \end{cases}$
  Set $z_{t+1} \leftarrow z_t + \delta_t u_t$, such that $\mathbb{E}[z_{t+1}|z_t] = z_t$
  Increment $t \leftarrow t + 1$
return $z_t \in \{\pm 1\}^n$

1. *There exists some K_1 such that $\Pr[|X| \geq t] \leq 2\exp\left(-t^2/K_1^2\right)$ for all $t \geq 0$,*
2. *There exists some K_2 such that $\mathbb{E}[\exp(X^2/K_2^2)] \leq 2$,*

and for X such that $\mathbb{E}[X] = 0$, the following condition is equivalent to the previous two.

3. *$\mathbb{E}[\exp(tX)] \leq \exp(K_3^2 t^2)$, for all $t \in \mathbb{R}$.*

A random vector is subgaussian with parameter K_2 if all of its unit-direction projections are subgaussian with parameter not exceeding K_2, and for the third formulation given we can equivalently show for a random vector Z that $\mathbb{E}[\exp(\langle Z, v\rangle)] \leq \exp(K_3^2\|v\|_2^2)$ for all vectors v, with the norm of v playing the role of the scalar t.

Vector Balancing: Suppose we are given $B = [b_1 \dots b_n] \in \mathbb{R}^{m\times n}$, where $\|b_i\|_2 \leq 1$ for $i \in [n]$, let $z_0 \in [-1,1]^n$. Then, the linear discrepancy objective is to output $z \in \{-1,1\}^n$ such that $\|B(z - z_0)\|_\infty$ is minimized. The Gram-Schmidt Walk samples a coloring whose discrepancy vector has controlled directional tails. This problem has been studied extensively, even in the online setting where only one vector is revealed at a time and signs must be irrevocably assigned (Bansal et al., 2020a). Harshaw et al. (2024) shows that $B(z - z_0)$ is 1-subgaussian (over the randomness of z).

Gram-Schmidt Walk. The Gram-Schmidt walk (Algorithm 1) achieves its guarantee through an adaptive sequence of randomized updates. At each iteration, it chooses a pivot coordinate, selects a least-squares direction whose pivot entry is one, and moves to one of the two endpoints of the feasible line segment in the unit hypercube $[-1,1]^n$ with probabilities such that the exact update is conditionally mean 0. Here, coordinates that reach $\{1,-1\}$ are frozen, and the direction is recomputed over the remaining coordinates. The concentration proof of Harshaw et al. (2024) exploits this probabilistic structure and the orthogonality of the revealed subspaces as the set of possible pivots shrinks.

The canonical result of Harshaw et al. (2024) proves that the outputs are 1-subgaussian.

Theorem 2.2 ((Harshaw et al., 2024)). *Let $\boldsymbol{B} = [b_1 \dots b_n] \in \mathbb{R}^{m\times n}$ satisfy $\|b_i\|_2 \leq 1$, and let $z_0 \in [-1,1]^n$. Then, for all $v \in \mathbb{R}^m$, the terminal output z of the walk in exact arithmetic satisfies*

$$\mathbb{E}\exp\big(\langle B(z - z_0), v\rangle\big) \leq \exp\left(\frac{1}{2}\|v\|_2^2\right). \tag{3}$$

2.1 Technical Contributions

We analyze the algorithm under a refined model that takes into account errors in calculating and sampling the step at each iteration, while still remaining general enough to apply across machine

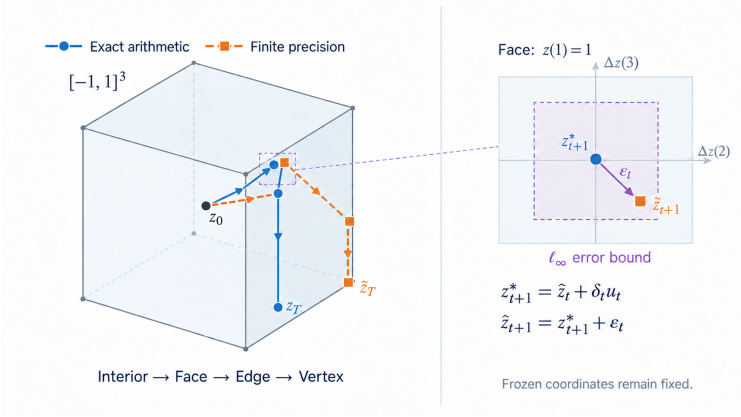


Figure 1: Finite-precision perturbations in the Gram–Schmidt walk. Left: Exact-arithmetic (solid blue) and finite-precision (dashed orange) walks start from the same point but can follow very different trajectories as coordinates successively freeze. Right: Each finite-precision update differs from the ideal update computed at its current state by a local error due to our tracked error sources.

configurations and least squares solvers. We forgo an analysis of arithmetic operation in the algorithm or considerations of overflow/underflow and instead parameterize over the key sources of error at each step. In particular we consider the error of the least squares solve ε_{LS}, the max coordinate error due to approximate arithmetic in computing the step and clamping z_t back into the hypercube ε_{iter}, and the perturbation in the distribution of $\hat{\delta}_t$ which we denote ε_{pr}. We discuss our model in more detail in section 3.

Note that the introduction of errors due to rounding can affect the path that the Gram-Schmidt walk may take, so the resultant walk may deviate significantly from the original ideal walk. Therefore, simply bounding the deviation from the original walk can lead to very weak guarantees. So, we treat this walk as its own object when bounding its subgaussianity. We show the following bound:

Theorem 2.3. *The output vector $B(z - z_0)$ satisfies the moment generating function (MGF) bound*

$$\mathbb{E}[\exp(\langle B(z-z_0), v\rangle)] \leq \exp\left(\frac{1}{2}\|v\|^2 + \left(\underbrace{n^{3/2}(2\varepsilon_{pr}+\varepsilon_{iter})}_{\textit{drift/sampling}} + \left(\underbrace{2n\varepsilon_{LS}}_{\textit{least squares}} + \underbrace{n(n+1)\varepsilon_{iter}}_{\textit{arithmetic/clamping}}\right)\right)\|v\|\right).$$

Where we can provide values for these constants depending on machine configuration; for example if we assume a floating point machine of unit roundoff u then as discussed in B.3 we can show that ε_{pr} scales according to the unit roundoff. Furthermore B.4 discusses how under special conditions, namely a backwards stable solver with full column rank B, we expect ε_{LS} to scale with condition number and unit roundoff. This is meant to give an idea for how ε_{LS} might scale in some special case rather than to hold in general.

Finite-precision errors obstruct a direct application of the analysis in Harshaw et al. (2024): they can change the freezing trajectory and destroy the conditional centering on which the phase argument relies. We overcome this by proving a phase MGF bound to explicitly account for accumulated drift along the realized perturbed trajectory, where the key step is to incorporate the pivot-coordinate update error into the scalar step so that its sum over a phase is the actual pivot displacement. This preserves the structure of the terminal displacement while allowing biased conditional probabilities, whose effect we control via a modified backward-induction argument. Combining these drift bounds across orthogonal phase subspaces produces the $n^{3/2}(2\varepsilon_{\mathrm{pr}} + \varepsilon_{\mathrm{iter}})$ contribution. The remaining terms follow from pathwise bounds on least-squares and coordinate-update error, including clamping and freezing corrections; hence, our bound keeps the original quadratic MGF term while adding an explicit linear penalty in $\|v\|$ that vanishes with error to recover the exact-arithmetic guarantee.

We note that this is not a subgaussianity guarantee in itself due to the linear term. Nevertheless, we show that an MGF bound of this form implies subgaussianity in the Orlicz norm formulation giving

a subgaussianity result: Lemma A.2 in the appendix implies

$$\|B(z - z_0)\|_{\psi_2} \le \sqrt{6} + \frac{1}{\sqrt{\log 2}} \left(n^{3/2}(2\varepsilon_{pr} + \varepsilon_{iter}) + (2n\varepsilon_{LS} + n(n+1)\varepsilon_{iter}) \right)$$

and so by the equivalence of the forms of Subgaussianity up to constants we also get Gaussian-like exponential tails on $|B(z - z_0), v|$ for every unit vector v.

Theorem 2.4 (Lower bound for bounded update errors)**.** *There exists a universal constant $c > 0$ such that, for every integer $n \ge 8$ and every $\varepsilon \ge 0$, there is a full-column-rank matrix $\boldsymbol{B} = [b_1, \dots, b_n] \in \mathbb{R}^{(n+1)\times n}$ with $\|b_i\|_2 = 1$ and an admissible error process for the Gram–Schmidt Walk, initialized at $z_0 = 0$, such that $z_{t+1} = z_t + \delta_t u_t + \varepsilon_t$ with $\|\varepsilon_t\|_\infty \le \varepsilon$, with exact step-selection probabilities, and whose terminal discrepancy $D = \boldsymbol{B} z_T$ satisfies*

$$\|\mathbb{E}D\|_2 \ge c \min\{n^2\varepsilon, n\}, \qquad \|D\|_{\psi_2} \ge c \min\{n^2\varepsilon, n\}.$$

Here $\|\cdot\|_{\psi_2}$ denotes the uncentered directional Orlicz norm. The errors are deterministic functions of the iteration, active set, and current pivot, which preserve all frozen coordinates, and the walk remains feasible and terminates after at most n iterations.

The construction proves that the $n^2\varepsilon_{iter}$ dependence is unavoidable, up to universal constants, for worst-case bias under admissible bounded coordinate errors, even when the least-squares directions and branch probabilities are exact. The same construction lower-bounds the uncentered directional Orlicz norm. This saturation at order n is consistent with $\|Bz_T\|_2 \le \sum_{i=1}^n \|b_i\|_2 \le n$ when $z_0 = 0$.

We also numerically[1] study how arithmetic precision and problem size affect discrepancy, where we illustrate the predicted quadratic accumulation of bounded update errors in Figure 4.

3 Finite-Precision Gram-Schmidt Walk

Because of the numeric issues outlined above, the existing proofs bounding the discrepancy of the Gram-Schmidt walk Bansal et al. (2017); Harshaw et al. (2024) no longer hold. The proof in Bansal et al. (2017) uses Martingales, but these properties are lost due to the errors. The more recent proof from Harshaw et al. (2024) builds on the fact that the exact solution to the L2-minimization problem can be expressed via results from an exact Gram Schmidt orthogonalization procedure. However, this too no longer holds in the presence of errors. Now, we explain the main technical ideas and tools needed to robustify the proof from Harshaw et al. (2024) against numerical issues.

Error Sources in Finite-Precision Model At the realized state z_t, let u_t denote an exact least-squares direction and let $\widehat{u}_t$ denote the computed direction. We state our bounds relative to the realized input. In our framework, we enforce $\widehat{u}_t(p_t) = 1$ and $\widehat{u}_t(i) = 0$ for $i \notin A_t$, and assume

$$\|B(\widehat{u}_t - u_t)\|_2 \le \varepsilon_{LS}. \tag{4}$$

All errors from the final arithmetic update, clamping, and explicit freezing are collected in ε_t, so that $z_{t+1} = z_t + \hat{\delta}_t \widehat{u}_t + \varepsilon_t$. The update preserves feasibility and previously frozen coordinates, freezes at least one additional coordinate, and retains the current pivot until it freezes. Let $\widehat{\delta}_t^+, \widehat{\delta}_t^- > 0$ denote the computed endpoint magnitudes, so that $\hat{\delta}_t \in \{\widehat{\delta}_t^+, -\widehat{\delta}_t^-\}$. For convenience, we cap these magnitudes by downward-rounded pivot distances, ensuring $\widehat{\delta}_t^+ \le 1 - z_t(p_t)$ and $\widehat{\delta}_t^- \le 1 + z_t(p_t)$. Thus, $\widehat{\delta}_t^+ + \widehat{\delta}_t^- \le 2$ and $|\hat{\delta}_t| \le 2$. Conditioning on the history and all computations preceding the random branch choice, let $p^{\mathrm{comp}} := \frac{\widehat{\delta}_t^-}{\widehat{\delta}_t^+ + \widehat{\delta}_t^-}$ be the probability that centers the step at the computed endpoints. We assume that the actual positive-branch probability $\widehat{p}$ satisfies $|\widehat{p} - p^{\mathrm{comp}}| \le \varepsilon_{pr}$, so

$$\left| \mathbb{E}[\hat{\delta}_t \mid \mathcal{F}_t] \right| = (\widehat{\delta}_t^+ + \widehat{\delta}_t^-)|\widehat{p} - p^{\mathrm{comp}}| \le 2\varepsilon_{pr}. \tag{5}$$

In this convention, the discrepancy increment decomposes as $B(z_{t+1} - z_t) = \hat{\delta}_t B u_t + \hat{\delta}_t B(\widehat{u}_t - u_t) + B\varepsilon_t$, where the least-squares contribution $\hat{\delta}_t B(\hat{u}_t - u_t)$ has norm at most $2\varepsilon_{LS}$.

[1] We provide our code in `https://anonymous.4open.science/r/gsw-5F80/README.md`.

Analysis framework We sketch the main ideas of Harshaw et al. (2024) and explain which pieces break due to numerical issues. We then later explain how to fix them.

Let $\sigma : [n] \to [n]$ be the reverse order in which the Gram-Schmidt Walk decides coordinates i (that is, when i is picked as a pivot or $z(i)$ becomes ± 1). Now assume we run Gram Schmidt orthogonalization on the columns of B in order σ and let $w_{\sigma(1)}, ..., w_{\sigma(n)} \in \mathbb{R}^m$ be the orthonormal vectors produced. We allow for B to not be full column rank and so the Gram-Schmidt orthogonalization procedure may produce a zero vector in the sequence; we refer to sets as orthonormal throughout our exposition with the understanding that we refer to the nonzero vectors being discussed. The vectors $w_{\sigma(i)}$ are not computed and used only for analysis purposes.

We consider the progression of the walk in pivot phases, where t_p is the iteration at which $p \in [n]$ is first selected as pivot. Within a phase, b_p refers to the current pivot column of B. When analyzing iteration $s \geq t_p$ of the phase for a pivot p, let $Q_{s,p}$ be the set of positions r decided in iteration s during the pivot phase of p. So $Q_{s,p} = \emptyset$ for $s \geq$ the end of pivot p's phase. While pivot p is technically decided at the start of iteration t_p, we charge it to iteration $t_p - 1$, i.e., $Q_{t_p-1,p}$ consists of just the pivot p. The proof of Harshaw et al. (2024) heavily relies on the following decomposition:

Lemma 3.1 (Lemma $S3.2$ in Harshaw et al. (2024))**.** *At each iteration t we have for pivot $p = p_t$*

$$Bu_t = \sum_{s=t_p-1}^{t-1} \sum_{r \in Q_s} \langle w_{\sigma(r)}, b_p \rangle w_{\sigma(r)}.$$

This is only true for the exact solution u_t and does not hold for the approximation $\hat{u_t}$ that would be produced by numeric implementation of the algorithm.

Finally, consider the final vector $B(z_T - z_0) = B(\sum_t \delta_t u_t)$ produced in exact arithmetic. Using Lemma 3.1, Harshaw et al. (2024) proves that for any $v \in \mathbb{R}^m$

$$\mathbb{E}\left[\exp\left(\sum_{t=1}^{n} \delta_t \langle Bu_t, v \rangle\right)\right] \leq \exp\left(\frac{1}{2}\|v\|^2\right) \tag{6}$$

This is done by splitting the run of the algorithm into phases based on the pivots. For a pivot p, let $S_p \subset [n]$ be the iterations where p was pivot, and let Δ_p be all the randomness up to the point that p is chosen as pivot. Then Harshaw et al. (2024) proves

$$\mathbb{E}\left[\exp\left(\sum_{t \in S_p} \delta_t \langle Bu_t, v \rangle - \frac{1}{2}\|P_p b_p\|^2 \|P_p v\|^2\right) \middle| \Delta_p\right] \leq 1 \tag{7}$$

which can then be assembled to prove equation 6, with the $\|P_p b_p\|^2 \|P_p v\|^2$ are orthogonal projection terms accumulating to at most $\|v\|^2$. However, for $\hat{u_t}$ and $\hat{\delta_t}$ the conditional expectation is no longer bounded by 1 because $\mathbb{E}[\hat{\delta_t}] \neq 0$ and Lemma 3.1 does not hold for $\hat{u_t}$.

Bounding the error. Consider the final vector $B(z - z_0) = B(\sum_t \hat{\delta_t} \hat{u_t} + \varepsilon_t)$.

We rearrange the updates as follows:

$$\begin{aligned} Bz - Bz_0 &= B\left(\sum_{t=1}^{T} \hat{\delta_t} u_t + \hat{\delta_t}(\hat{u_t} - u_t) + \varepsilon_t\right) \\ &= B\sum_{t=1}^{T} \left(\hat{\delta_t} + \varepsilon_t(p_t)\right) u_t + B\left(\sum_{t=1}^{T} \hat{\delta_t}(\hat{u_t} - u_t) + \varepsilon_t - \varepsilon_t(p_t) u_t\right) \end{aligned} \tag{8}$$

Here the latter term can be bounded by using the definition of our errors, giving the bound

$$\left\| \sum_{t=1}^{T} B(\hat{\delta_t}(\hat{u_t} - u_t) + \varepsilon_t - \varepsilon_t(p_t) u_t) \right\|_2 \leq (2n\varepsilon_{LS} + n(n+1)\varepsilon_{iter}.$$

The calculation is given in Section B.2.

The first term in equation 8 is similar to equation 6 in that it uses the exact u_t, however, the probability $\hat{\delta_t}$ is still inexact. As a result, we prove the following weaker bound with a variable γ upper bounding the drift.

Lemma 3.2. *For $\gamma = n(2\varepsilon_{pr} + \varepsilon_{iter})$, the non-noise portion of the updates satisfy the MGF bound*

$$\mathbb{E}\left[\exp\left(\sum_{t=1}^{n}(\hat{\delta}_t + \varepsilon_t(p_t))\langle Bu_t, v\rangle\right)\right] \leq \exp(\frac{1}{2}\|v\|^2 + \sqrt{n}\gamma\|v\|)$$

The main difficulty in adapting the proof of equation 6 by Harshaw et al. (2024) to Lemma 3.2, is the reliance on the distribution of $\sum_t \delta_t$ within a pivot phase being a two point distribution, which no longer holds due to noise along the pivot unit. The high-level idea of our proof is that, while equation 7 does not hold, it does hold after adding some penalty term E_t such that

$$\mathbb{E}\left[\exp\left(\sum_{t\in S_p}\delta_t\langle Bu_t, v\rangle - \frac{1}{2}\|P_p b_p\|^2\|P_p v\|^2 - E_t\right)\middle|\Delta_p\right] \leq 1$$

The main difficulty is to find a large enough E_t for which the inequality holds, but not too large to get a bound weaker than Lemma 3.2. The full proof is long and can be found in Sections B.1 and B.3, but it turns out that $E_t = \gamma\sum_{s\geq t_p-1}\sum_{r\in Q_{s,p}}|\alpha_r\beta_r|$ for $\gamma = n(2\varepsilon_{pr} + \varepsilon_{iter})$, $\alpha_r = \langle w_{\sigma(r)}, b_p\rangle$ and $\beta_r = \langle w_{\sigma(r)}, v\rangle$ is large enough. We further have

$$\sum_p \sum_{s\geq t_p-1}^{n} \sum_{r\in Q_{s,p}} |\alpha_r\beta_r| \leq \Big(\sum_{p=1}^{n}\|P_p b_p\|^2\Big)^{1/2}\Big(\sum_{p=1}^{n}\|P_p v\|^2\Big)^{1/2} \leq \sqrt{n}\,\|v\|,$$

where P_p is the orthogonal projection matrix onto the $w_{\sigma(r)}$ decided in pivot phase p. This yields the additional term in Lemma 3.2. Assembling these two bounds on equation 8 then yields Theorem 2.3.

4 Numerical Experiments

Compute. Our walks run in parallel; we use a process pool with per-worker `SeedSequence` streams and single-threaded BLAS. The per-operation rounding makes the low-precision path $\sim 20\times$ slower than float64's. Our main figures use 1000 walks per cell on a 6-core laptop (1-3 CPU-hours).

Matrices. We use four matrix ensembles.

(i) *Clustered.* $B = u\mathbf{1}^\top + m^{-1/2}G$, with u a random unit vector and G Gaussian, column-normalized. The columns concentrate near a shared direction, which makes $B_{A\setminus p}$ increasingly ill-conditioned as the active set shrinks through $k \approx m$,

(ii) *Higgs.* The two low-level angular-momentum features of the Higgs two-sample benchmark used in low-rank thinning (Carrell et al., 2025), giving a $2 \times n$ matrix,

(iii) *Identity.* $B = I_n$, a tight control in which $d^\top Bz$ is a Rademacher sum with $\sigma = 1$.

(iv) *Adversarial lower bound.* Columns $v_i = \frac{e_0+e_i}{\sqrt{2}} \in \mathbb{R}^{n+1}$, where a bounded per-step error a applied to non-pivot coordinates in the first $\frac{n}{4}$ iterations forces an e_0-discrepancy of order $\Omega(n^2\varepsilon)$.

Metrics. For each main-sweep configuration, we run 1000 independent walks. *Discrepancy.* We report the sample mean of $\|Bz\|_1/m$ and the empirical CCDF of $\|Bz\|_\infty$ on squared-threshold, log-survival axes. *Directional concentration.* We compute empirical moment diagnostics using orders $2, 4, 6$, tail-inversion diagnostics, and ψ_2 norms over a finite direction set $\mathcal{D}$. In the n- and bit-sweeps, we report the maximum moment diagnostic and the Orlicz statistic in that same direction; error bars show the standard deviation of maxima computed on five disjoint batches. *Solver error.* We follow a float64 reference trajectory and recompute directions in low precision at identical states, recording $\Delta u_t = u_{\text{lp}} - u_{\text{fp64}}$ by the number k of free nonpivot coordinates. We summarize these coefficient errors using medians, IQRs, and the Gaussian-calibrated robust scale IQR.

Findings. At fixed arithmetic precision, our experiments show that the discrepancy inflation generally grows with the number of vectors n, with the strongest deterioration occurring at the lowest precisions. Increasing the number of retained bits mitigates this growth and brings the empirical discrepancy statistics closer to the float64 reference, suggesting that larger instances need higher precision to maintain comparable performance.

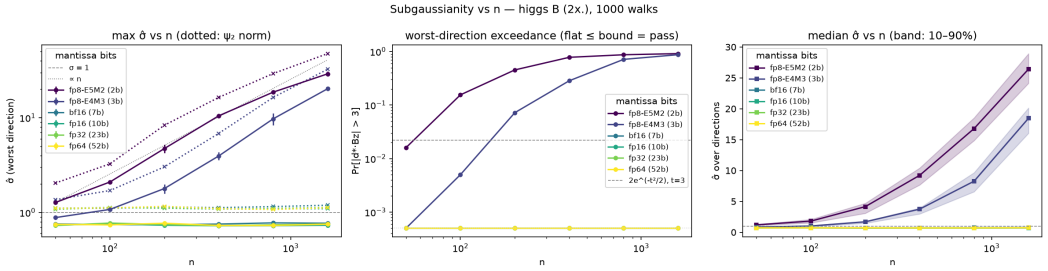


Figure 2: **Subgaussianity vs. n across fp formats on the Higgs matrix.** Left: worst-direction $\hat{\sigma}$ (solid is the moment estimate, dotted the ψ_2 norm). Formats down to bf16 stay flat near $\hat{\sigma} \approx 0.8$ for all n, while fp8 grows roughly linearly. Middle: the fraction of walks with $|d_*^\top Bz| > 3$ in the worst direction, an estimator-free check against the ideal bound $2e^{-9/2}$. Right: the median $\hat{\sigma}$ over directions with a 10–90% band, showing that only a few directions inflate.

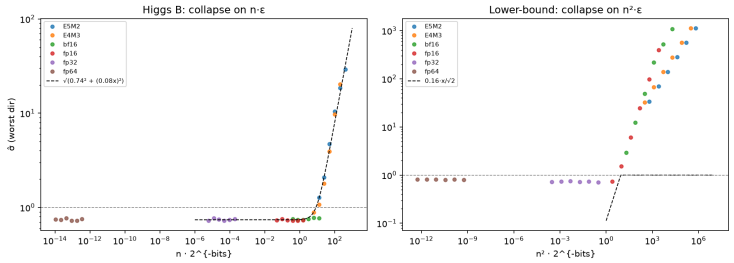


Figure 3: **Scaling collapse.** Left: every (n, b) cell on the Higgs matrix falls on one curve in $n \cdot 2^{-b}$. Right: the adversarial lower-bound matrix collapses instead onto $n^2 \cdot 2^{-b}$. Rounding error accumulates coherently, linearly in n on the Higgs data and quadratically with adversarially chosen error.

Implementation details. We compare fp16/fp32 implementations of Householder QR least squares, the GSWDESIGN Cholesky/Woodbury scheme, and the explicit-inverse Sherman–Morrison scheme from low-rank thinning's cubic walk, using common random draws. On the tested instances, QR is the most robust, with median image error less sensitive to conditioning than coefficient error. Refactoring every ten steps reduces fp16 Cholesky downdate failures to at most 1.75%, but ψ_2 inflation persists among successful runs. The explicit-inverse scheme is more sensitive to ill-conditioning, matching the squared conditioning of Gram matrices where $\kappa_2(B_S^\top B_S) = \kappa_2(B_S)^2$ of full-rank B_S. Our experiments emulate fraction-bit rounding and reductions accumulate in float64 (see Appendix D).

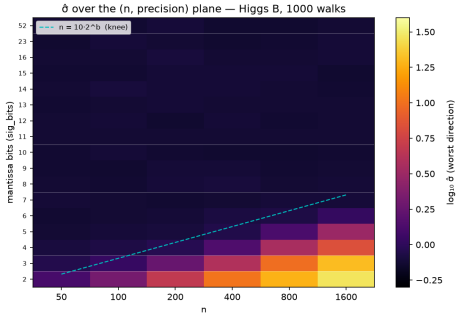


Figure 4: Estimated subgaussianity over the (n, bits) plane. Each cell is $\log_{10} \hat{\sigma}$ for the worst direction on the Higgs matrix. The float64 reference plateau goes to $\sim$8 mantissa bits at small n and ends along the dashed knee $n \approx 10 \cdot 2^b$. Above the knee, $\hat{\sigma}$ grows along lines of constant $n \cdot 2^{-b}$.

5 CONCLUSION

We analyze the Gram-Schmidt Walk under bounded, potentially biased and history-dependent perturbations. Theoretically, we provide a concentration bound that separates least-squares image error, coordinatewise update error, and conditional sampling error while retaining the exact-arithmetic quadratic term in the log-MGF. We also give a full-column-rank construction that produces a bias of order $\min\{n^2\varepsilon_{iter}, n\}$, which shows that a quadratic accumulation of update error is unavoidable before saturation. Finally, our experiments corroborate our theoretical findings on how numerical discrepancy varies with fixed precision and problem size.

AI DISCLOSURE

In this work, we disclose the use of Cognition's Devin agent to help run our suite of experiments, and we used Anthropic's Claude Fable 5.1 model to suggest improvements on an earlier version of the theoretical result. We have reviewed all AI-assisted work. LLM-generated code was verified and tested for correctness, and we take responsibility for all the final content of this work.

ETHICS STATEMENT

This work studies the worst-case numerical reliability of the Gram-Schmidt Walk, an algorithm for vector balancing. By quantifying the effects of finite precision, our results support more reliable use of discrepancy-based methods and clarify circumstances in which numerical errors can compromise output quality. Applications involving experimental design or data compression require separate consideration of data privacy, representativeness, and the consequences of the resulting decisions.

REPRODUCIBILITY STATEMENT

The theoretical assumptions and perturbation models underlying our theoretical guarantees are specified in the main body, with proofs provided in the appendices. Our experimental setup describes our instance generation, dataset preprocessing, solver and factorization-maintenance procedures, numerical precision settings, random seeds, and evaluation metrics. To support reproduction of the finite-precision experiments, we document the arithmetic and accumulation precisions, explain our compute setup, and provide our code with thorough documentation.

REFERENCES

Ryan Alweiss, Yang P. Liu, and Mehtaab Sawhney. Discrepancy minimization via a self-balancing walk, 2020. URL https://arxiv.org/abs/2006.14009.

Emile Anand, Jan van den Brand, Mehrdad Ghadiri, and Daniel J. Zhang. The Bit Complexity of Dynamic Algebraic Formulas and Their Determinants. In Karl Bringmann, Martin Grohe, Gabriele Puppis, and Ola Svensson, editors, *51st International Colloquium on Automata, Languages, and Programming (ICALP 2024)*, volume 297 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 10:1–10:20, Dagstuhl, Germany, 2024. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. ISBN 978-3-95977-322-5. doi: 10.4230/LIPIcs.ICALP.2024.10. URL https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.ICALP.2024.10.

Wojciech Banaszczyk. Balancing vectors and gaussian measures of n-dimensional convex bodies. *Random Struct. Algorithms*, 12(4):351–360, July 1998. ISSN 1042-9832.

Nikhil Bansal, Daniel Dadush, Shashwat Garg, and Shachar Lovett. The gram-schmidt walk: A cure for the banaszczyk blues, 2017. URL https://arxiv.org/abs/1708.01079.

Nikhil Bansal, Haotian Jiang, Raghu Meka, Sahil Singla, and Makrand Sinha. Online discrepancy minimization for stochastic arrivals, 2020a. URL https://arxiv.org/abs/2007.10622.

Nikhil Bansal, Haotian Jiang, Sahil Singla, and Makrand Sinha. Online vector balancing and geometric discrepancy, 2020b. URL https://arxiv.org/abs/1912.03350.

Nikhil Bansal, Aditi Laddha, and Santosh S. Vempala. A unified approach to discrepancy minimization, 2022. URL https://arxiv.org/abs/2205.01023.

Annabelle Michael Carrell, Albert Gong, Abhishek Shetty, Raaz Dwivedi, and Lester Mackey. Low-rank thinning. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pages 6811–6848. PMLR, 2025. URL https://proceedings.mlr.press/v267/carrell25a.html.

Bernard Chazelle. *The Discrepancy Method: Randomness and Complexity*. Cambridge University Press, Cambridge, UK, 2000. ISBN 9780521770934.

Bernard Chazelle. The discrepancy method in computational geometry. In *Handbook of Discrete and Computational Geometry, 2nd Ed.*, 2004. URL https://api.semanticscholar.org/CorpusID:3198958.

Jerry Chee, Arturs Backurs, Rainie Heck, Li Zhang, Janardhan Kulkarni, Thomas Rothvoss, and Sivakanth Gopi. Discquant: A quantization method for neural networks inspired by discrepancy theory, 2025. URL https://arxiv.org/abs/2501.06417.

Daniel Dadush, Shashwat Garg, Shachar Lovett, and Aleksandar Nikolov. Towards a constructive version of banaszczyk's vector balancing theorem, 2016. URL https://arxiv.org/abs/1612.04304.

Daniel Dadush, Aleksandar Nikolov, Kunal Talwar, and Nicole Tomczak-Jaegermann. Balancing vectors in any norm. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1–10. IEEE, 2018.

Yichuan Deng, Xiaoyu Li, Zhao Song, and Omri Weinstein. Discrepancy minimization in input-sparsity time. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=TmJvacopmV.

Josef Dick and Friedrich Pillichshammer. Discrepancy theory and quasi-monte carlo integration. In *A panorama of discrepancy theory*, pages 539–619. Springer, 2014.

Claus Diem. On the notion of bit complexity. *Bull. EATCS*, 103:35–52, 2011.

Friedrich Eisenbrand, Dömötör Pálvölgyi, and Thomas Rothvoß. Bin packing via discrepancy of permutations, 2012. URL https://arxiv.org/abs/1007.2170.

David Gamarnik, Eren C Kizildağ, Will Perkins, and Changji Xu. Geometric barriers for stable and online algorithms for discrepancy minimization. In *The Thirty Sixth Annual Conference on Learning Theory*, pages 3231–3263. PMLR, 2023.

Mehrdad Ghadiri, Richard Peng, and Santosh S. Vempala. The bit complexity of efficient continuous optimization. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 2059–2070, 2023. doi: 10.1109/FOCS57990.2023.00125.

Mehrdad Ghadiri, Yin Tat Lee, Swati Padmanabhan, William Swartworth, David P. Woodruff, and Guanghao Ye. Improving the bit complexity of communication for distributed convex optimization. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, STOC 2024, page 1130–1140, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400703836. doi: 10.1145/3618260.3649787. URL https://doi.org/10.1145/3618260.3649787.

Mehrdad Ghadiri, Hoai-An Nguyen, and Junzhao Yang. Entrywise approximation for matrix inversion and linear systems. In *Proceedings of the 2026 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 5199–5210. SIAM, 2026. doi: 10.1137/1.9781611978971.189. URL https://epubs.siam.org/doi/10.1137/1.9781611978971.189.

Luc Giraud, Julien Langou, Miroslav Rozložník, and Jasper van den Eshof. Rounding error analysis of the classical Gram–Schmidt orthogonalization process. *Numerische Mathematik*, 101(1):87–100, 2005. doi: 10.1007/s00211-005-0615-4.

Simone Göttlich, Jacob Heieck, and Andreas Neuenkirch. Using low-discrepancy points for data compression in machine learning: An experimental comparison, 2024. URL https://arxiv.org/abs/2407.07450.

Christopher Harshaw, Fredrik Sävje, Daniel A. Spielman, and Peng Zhang. Balancing covariates in randomized experiments with the gram–schmidt walk design. *Journal of the American Statistical Association*, 119(548):2934–2946, 2024. doi: 10.1080/01621459.2023.2285474. URL https://doi.org/10.1080/01621459.2023.2285474.

Nicholas J. Higham. *Accuracy and Stability of Numerical Algorithms*. Society for Industrial and Applied Mathematics, second edition, 2002. doi: 10.1137/1.9780898718027. URL https://epubs.siam.org/doi/abs/10.1137/1.9780898718027.

Zohar Karnin and Edo Liberty. Discrepancy, coresets, and sketches in machine learning, 2019. URL https://arxiv.org/abs/1906.04845.

Ekaterina Kochetkova, Kshiteej Sheth, Insu Han, Amir Zandieh, and Michael Kapralov. Streaming attention approximation via discrepancy theory, 2026. URL https://arxiv.org/abs/2502.07861.

Janardhan Kulkarni, Victor Reis, and Thomas Rothvoss. Optimal online discrepancy minimization. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, STOC 2024, page 1832–1840, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400703836. doi: 10.1145/3618260.3649720. URL https://doi.org/10.1145/3618260.3649720.

Lap Chi Lau, Robert Wang, and Hong Zhou. Spectral sparsification by deterministic discrepancy walk. In *2025 Symposium on Simplicity in Algorithms (SOSA)*, pages 315–340. SIAM, 2025.

Edo Liberty, Alexandr Andoni, and Eldar Kleiner. Nearly optimal attention coresets, 2026. URL https://arxiv.org/abs/2605.05602.

Linkai Ma, Christos Boutsikas, Mehrdad Ghadiri, and Petros Drineas. A note on the stability of the sherman–morrison–woodbury formula. *SIAM Journal on Matrix Analysis and Applications*, 47(2):702–723, 2026. doi: 10.1137/25M1747087. URL https://doi.org/10.1137/25M1747087.

V. Ya. Pan. The bit complexity of matrix multiplication and of related computations in linear algebra. the segmented λ algorithms. *Computers & mathematics with applications*, 11(9):919–928, 1985.

Joel Spencer. Six standard deviations suffice. *Transactions of the American Mathematical Society*, 289(2):679–706, 1985. doi: 10.1090/S0002-9947-1985-0784009-0.

Joel Spencer. *Ten lectures on the probabilistic method*. SIAM, 1994.

Roman Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2 edition, 2026.

Lingao Xiao, Songhua Liu, Yang He, and Xinchao Wang. Unifying dataset pruning and distillation for efficient large-scale compression, 2026. URL https://arxiv.org/abs/2502.06434.

Outline of the Appendix

- Appendix A collects auxiliary lemmas used in our concentration analysis.
- Appendix B.1 gives an analysis of the single pivot phases.
- Appendix B.2 proves Theorem 2.3 by combining the phase concentration estimates with bounds on the accumulated numerical errors.
- Appendix C presents our lower-bound construction and establishes the worst-case bias caused by bounded update errors.
- Appendix D describes our experimental setup and provides additional numerical results.

A Auxiliary Lemmas

We now state a lemma from Harshaw et al. (2024) to be used in our adapted proof, to allow us to bound the expectation of a certain MGF quantity.

Lemma A.1 (Lemma $S3.5$ in Harshaw et al. (2024)). *For all $x \in [-1, 1]$ the function*

$$f_x \begin{bmatrix} \alpha & \eta \\ \eta & \beta \end{bmatrix} = \exp\left(-\frac{1}{2}\alpha\beta\right)\left[\frac{1+x}{2}\exp((1-x)\eta) + \frac{1-x}{2}\exp((-1-x)\eta)\right]$$

is monotone decreasing with respect to the Loewner order on the cone of 2×2 positive semidefinite matrices: if $M \preceq N$, then $f_x(M) \geq f_x(N)$.

Lemma A.2 (MGF Bound to Orlicz Norm). *If a random vector $X \in \mathbb{R}^m$ satisfies an MGF bound of the form*

$$\mathbb{E}[\exp(\langle X, v\rangle)] \leq \exp\left(\frac{1}{2}\|v\|^2\sigma^2 + a\|v\|\right)$$

for all $v \in \mathbb{R}^m$ and some $\sigma > 0$ then we have that $\|X\|_{\psi_2} \leq a/\sqrt{\ln 2} + \sqrt{6}\sigma$.

Proof. Recalling that we define the Orlicz norm of a random vector as the supremum of the Orlicz norm over its one dimensional scalar projections, fix some $u \in S^{m-1}$ and let $Y := \langle X, u\rangle$. We will ultimately show that $\mathbb{E}[Y^2/(a/\sqrt{\ln 2} + \sqrt{6}\sigma)^2] \leq 2$. Observe that because X satisfies the above MGF bound we have, for any $\lambda \geq 0$,

$$\mathbb{E}[e^{\lambda Y}] = \mathbb{E}[\exp\langle X, \lambda u\rangle] \leq \exp\left(\frac{1}{2}\lambda^2\sigma^2 + a\lambda\right)$$

and therefore by applying Markov's inequality in the MGF we get

$$\Pr[|Y| \geq t] = \Pr[e^{\lambda|Y|} \geq e^{\lambda t}] \leq 2\exp\left(\frac{1}{2}\lambda^2\sigma^2 + a\lambda - t\lambda\right)$$

and choose $\lambda = \frac{t-a}{\sigma^2}$ to get the following bound for $t \geq a$:

$$\leq 2\exp(-(t-a)^2/(2\sigma^2))$$

From here we define $Z := (|Y| - a)_+$ such that $Z \geq 0$. Observe that for $r > 0$ we have

$$\Pr[|Z| \geq r] = \Pr[(|Y| - a)_+ \geq r] \leq \Pr[|Y| \geq a + r] \leq 2\exp(\frac{-r^2}{2\sigma^2})$$

which allows us to show the claimed bound on the MGF of Y^2 via an application of the layer cake representation to first bound the MGF of Z^2. Let $K > 0$ be chosen later, considering the function $f(r) = e^{r^2/K^2}$ we have

$$\begin{aligned}\mathbb{E}[\exp(Z^2/K^2)] &= 1 + \int_0^\infty f'(r)\Pr[r < Z]dr \\ &= 1 + \int_0^\infty \frac{2r}{K^2}e^{r^2/K^2}\Pr[r < Z]dr\end{aligned}$$

$$\leq 1 + \int_0^\infty \frac{2r}{K^2} e^{r^2/K^2} \cdot 2\exp(-r^2/(2\sigma^2))dr$$
$$= 1 + \int_0^\infty \frac{4r}{K^2} \exp\left(-r^2\left(\frac{1}{2\sigma^2} - \frac{1}{K^2}\right)\right) dr$$
$$= 1 + \frac{4\sigma^2}{K^2 - 2\sigma^2},$$

where the final step holds for $K^2 > 2\sigma^2$. To make this at most 2 we set $K = \sqrt{6}\sigma$ which makes the above expression equal 2, ensuring convergence of the above integral. This gives $\|Z\|_{\psi_2} \leq \sqrt{6}\sigma$.

To conclude the proof we use that $|Y| \leq Z + a$ pointwise implying $\mathbb{E}[e^{Y^2}] \leq \mathbb{E}[e^{(Z+a)^2}]$ and so $\|Y\|_{\psi_2} \leq \|Z + a\|_{\psi_2} \leq \|Z\|_{\psi_2} + \|a\|_{\psi_2} \leq \sqrt{6}\sigma + a/\sqrt{\ln 2}$, where we used the triangle inequality followed by the boundedness of a in the final two steps. □

B Missing Proofs

B.1 Single Pivot Phase Analysis (Proof of Lemma 3.2)

Our proof follows the same structure as in Harshaw et al. (2024). To prove Lemma 3.2, which states

$$\mathbb{E}\left[\exp\left(\sum_{t=1}^{n}(\delta_t + \varepsilon_t(p_t))\langle Bu_t, v\rangle\right)\right] \leq \exp(\frac{1}{2}\|v\|^2 + \sqrt{n}\gamma\|v\|) \tag{9}$$

we want to split the sum into separate phases based on the used pivots p. We regard the pivot phase using pivot $p \in [n]$ as starting in iteration t_p. For $s \geq t_p - 1$ we let $Q_{s,p} \subset [n]$ be the set of all positions r decided in iteration s during the pivot phase of p. While p is technically decided in iteration t_p, for analysis we count it in $Q_{t_p-1,p}$. Further, $Q_{s,p} = \emptyset$ for every s after the iteration in which p is frozen.

Let Δ_p be all the randomness up to the start of iteration t_p point when p is picked as the pivot (the choice of p is included in Δ_p). We will analyze the contribution of pivot p's phase to equation 9 by analyzing its contribution conditional on Δ_p. In particular, given that conditioning, p, t_p, $Q_{t_p-1,p}$ are fixed. Technically, the end of the pivot p's phase is random, but for analysis is would be easier if it were fixed. To do this, we define $\delta_{t,p} = \delta_t, \varepsilon_{t,p} = \varepsilon_t$ and $= 0$ for t outside the phase. Thus we can take the sum $t = t_p...n$ for fixed end point n such that

$$\sum_{t=t_p}^{n}(\delta_{s,p} + \varepsilon_{s,p}(p_t))\langle Bu_t, v\rangle$$

is the contribution of pivot p's phase to the sum in equation 9. In particular, a sum over the pivot phase is written $\sum_{t=t_p}^{n}$ and a sum over the positions decided in the phase is written $\sum_{s=t_p-1}^{n}\sum_{r\in Q_{s,p}}$. For the later we use that $Q_{s,p} = 0$ for s outside the pivot's phase.

We reuse the notation α_r, β_r to denote $\langle w_{\sigma(r)}, b_p\rangle$ and $\langle w_{\sigma(r)}, v\rangle$, and we use x to denote $z_t(p)$, the current position in $[-1, 1]$ along the pivot unit when it improves notational clarity. Let P_p be the projection matrix onto the $\mathrm{span}(\{w_{\sigma(r)} | r$ decided during pivot p's phase$\})$

Lemma B.1. *For a pivot p and $\gamma = (2\varepsilon_{pr} + \varepsilon_{iter})n$ we have*

$$\mathbb{E}\left[\exp\left(\sum_{t=t_p}^{n}(\delta_{t,p} + \varepsilon_{t,p}(p))\langle Bu_t, v\rangle - \frac{1}{2}\|P_p b_p\|^2\|P_p v\|^2 - \gamma\sum_{s=t_p-1}^{n}\sum_{r\in Q_{s,p}}|\alpha_r\beta_r|\right)\middle|\Delta_p\right] \leq 1$$

Before proving Lemma B.1, we show that it implies Lemma 3.2. In particular, Lemma B.1 can be seen as the contribution of pivot phase p to equation 9/Lemma 3.2.

Proof of Lemma 3.2.

$$\mathbb{E}\left[\exp\left(\sum_{t=1}^{T}(\delta_t + \varepsilon_t(p_t))\langle Bu_t, v\rangle\right)\right]\exp\left(-\gamma\sqrt{n}\|v\| - \frac{1}{2}\|v\|^2\right)$$

$$\leq \mathbb{E}\left[\prod_p \exp\left(\sum_{t=t_p}^{n}(\delta_{t,p}+\varepsilon_{t,p}(p))\langle Bu_t,v\rangle - \frac{1}{2}\|P_pv\|^2 - \gamma\sum_{s=t_p-1}^{n}\sum_{r\in Q_{s,p}}|\alpha_r\beta_r|\right)\right]$$

$$\leq \mathbb{E}\left[\prod_p \exp\left(\sum_{t=t_p}^{n}(\delta_{t,p}+\varepsilon_{t,p}(p))\langle Bu_t,v\rangle - \frac{1}{2}\|P_pb_p\|^2\|P_pv\|^2 - \gamma\sum_{s=t_p-1}^{n}\sum_{r\in Q_{s,p}}|\alpha_r\beta_r|\right)\right] \tag{10}$$

where the first inequality holds pointwise because $\sum_p\|P_pv\|^2 \leq \|v\|^2$, as the P_p project onto mutually orthogonal subspaces, and because by Cauchy–Schwarz

$$\sum_p\sum_{s=t_p-1}^{n}\sum_{r\in Q_{s,p}}|\alpha_r\beta_r| \leq \Big(\sum_{p=1}^{n}\|P_pb_p\|^2\Big)^{1/2}\Big(\sum_{p=1}^{n}\|P_pv\|^2\Big)^{1/2} \leq \sqrt{n}\,\|v\|,$$

using orthonormality of the $\{w_{\sigma(r)}\}$ and $\|b_p\| \leq 1$ which holds deterministically.

Let

$$X_p := \exp\left(\sum_{t=t_p}^{n}(\delta_{t,p}+\varepsilon_{t,p}(p))\langle Bu_t,v\rangle - \frac{1}{2}\|P_pb_p\|^2\|P_pv\|^2 - \gamma\sum_{s=t_p-1}^{n}\sum_{r\in Q_{s,p}}|\alpha_r\beta_r|\right).$$

then by conditioning on Δ_p and Lemma B.1 we get for any $k \leq n$ that $\mathbb{E}[X_k|\Delta_k] \leq 1$ and thus

$$\begin{aligned}\mathbb{E}\left[\prod_{p=1}^{k}X_p\right] &= \mathbb{E}\left[X_k\cdot\prod_{p=1}^{k-1}X_p\right]\\ &= \mathbb{E}\left[\mathbb{E}[X_k|\Delta_k]\cdot\prod_{p=1}^{k-1}X_p\right]\\ &\leq \mathbb{E}\left[\prod_{p=1}^{k-1}X_p\right].\end{aligned}$$

Inductively this implies $\mathbb{E}\left[\prod_{p=1}^{k}X_p\right] \leq 1$. As in Harshaw et al. (2024) we note that the result holds for any fixed pivot ordering so we are free to take the above ordering for notation, with the full result following from a marginalization over all orderings.

At last, since equation 10 is bounded by 1, rearranging shows

$$\mathbb{E}\left[\exp\left(\sum_{t=1}^{T}(\delta_t+\varepsilon_t(p_t))\langle Bu_t,v\rangle\right)\right] \leq \exp\left(\frac{1}{2}\|v\|^2+\sqrt{n}\gamma\|v\|\right)$$

as desired. □

Proof of Lemma B.1. By applying Lemma 3.1 and linearity we have $\langle Bu_t,v\rangle = \sum_{s=t_p-1}^{t-1}\sum_{r\in Q_{s,p}}\alpha_r\beta_r$ for pivot $p = p_t$, and that the set of $\{w_{\sigma(r)}\}$ where position r is decided in pivot phase p is an orthonormal basis for the subspace P_p projects into, implying $\|P_pb_p\|^2 = \sum_{s=t_p-1}^{n}\sum_{r\in Q_{s,p}}\alpha_r^2$ and $\|P_pv\|^2 = \sum_{s=t_p-1}^{n}\sum_{r\in Q_{s,p}}\beta_r^2$. This yields

$$\mathbb{E}\Bigg[\exp\Bigg(\sum_{t=t_p}^{n}(\delta_{s,p}+\varepsilon_{s,p}(p))\sum_{s=t_p-1}^{t-1}\sum_{r\in Q_{s,p}}\alpha_r\beta_r - \frac{1}{2}\left(\sum_{s=t_p-1}^{n}\sum_{r\in Q_{s,p}}\alpha_r^2\right)\left(\sum_{s=t_p-1}^{n}\sum_{r\in Q_{s,p}}\beta_r^2\right) - \gamma\sum_{s=t_p-1}^{n}\sum_{r\in Q_{s,p}}|\alpha_r\beta_r|\Bigg)\Bigg|\Delta_p\Bigg].$$

To prove Lemma B.1, this expression must be bounded by 1.

For $R \in \{t_p - 1, \ldots, n\}$ let us define

$$g(R) := \mathbb{E}\Bigg[\exp\Bigg(\sum_{t=t_p}^{n} (\delta_{t,p} + \varepsilon_{t,p}(p)) \sum_{s=t_p-1}^{\min\{R,\, t-1\}} \sum_{r \in Q_{s,p}} \alpha_r \beta_r - \frac{1}{2} \Bigg(\sum_{s=t_p-1}^{R} \sum_{r \in Q_{s,p}} \alpha_r^2 \Bigg) \Bigg(\sum_{s=t_p-1}^{R} \sum_{r \in Q_{s,p}} \beta_r^2 \Bigg) - \gamma \sum_{s=t_p-1}^{R} \sum_{r \in Q_{s,p}} |\alpha_r \beta_r| \Bigg) \Bigg| \Delta_p \Bigg].$$

Observe that $g(n)$ recovers the same expression, since $\min\{n, t-1\} = t-1$ for every $t \le n$. We will prove that $g(n) \le g(n-1) \le \cdots \le g(t_p - 1) \le 1$ which then implies Lemma B.1.

Since t_p is determined by Δ_p, every summation range above is deterministic under this conditioning. We note that we're progressing not over the contributions to the MGF at each timestep in the pivot phase but rather over which decided positions r we consider for the terms involving α_r, β_r. We first point out that $g(n) \le g(n-1)$, as the first summation

$$\sum_{t=t_p}^{n} (\delta_{t,p} + \varepsilon_{t,p}(p)) \sum_{s=t_p-1}^{\min\{R,\, t-1\}} \sum_{r \in Q_{s,p}} \alpha_r \beta_r$$

is identical for $R = n$ and $R = n-1$; every $t \le n$ satisfies $\min\{n, t-1\} = \min\{n-1, t-1\} = t-1$. The other two summations in the above expectation only increase in absolute value as we sum over more terms, and so we have $g(n) \le g(n-1)$. We now proceed to show for $R \in \{t_p, \ldots, n-1\}$ we have $g(R) \le g(R-1)$, with the final inequality $g(t_p - 1) \le 1$ to be shown later via a similar technique.

Proving $g(R) \le g(R-1)$: We claim it suffices to show that for any such R we have

$$\mathbb{E}\left[\exp\left(\sum_{t=R+1}^{n} (\delta_{t,p} + \varepsilon_{t,p}(p)) \eta_R - \frac{1}{2} a_R b_R - \gamma c_R \right) \middle| \Delta_R \right] \le \mathbb{E}\left[\exp\left(\sum_{t=R+1}^{n} (\delta_{t,p} + \varepsilon_{t,p}(p)) \eta_{R-1} - \frac{1}{2} a_{R-1} b_{R-1} - \gamma c_{R-1} \right) \middle| \Delta_R \right],$$

where we use Δ_R to denote randomness up to and including iteration R (so it includes Δ_p), and the shorthand $\eta_R := \sum_{s=t_p-1}^{R} \sum_{r \in Q_{s,p}} \alpha_r \beta_r$ with a_R, b_R, c_R having the same summation but with the inner term being α_r^2, β_r^2 or $|\alpha_r \beta_r|$ respectively. The expectation on the left matches $g(R)$ except that the first summation considers only iterations $t \ge R+1$, for which $\min\{R, t-1\} = R$, and so all of the summations in the expectation involve the same terms.

To show the above inequality, we observe that the only randomness under this conditioning is in the realization of the $\delta_{t,p} + \varepsilon_{t,p}(p)$.

Note that this sum must take values in $\{1 - x, -1 - x\}$, where $x := z_{R+1}(p)$ is the position of the pivot coordinate after iteration R: by our convention each coefficient $\delta_{t,p} + \varepsilon_{t,p}(p)$ is the pivot increment $z_{t+1}(p) - z_t(p)$, so the sum telescopes to $\sum_{t=R+1}^{n} (\delta_{t,p} + \varepsilon_{t,p}(p)) = z_{n+1}(p) - x$ with $z_{n+1}(p) \in \{1, -1\}$. In the case where $\sum_{\substack{t \in S_p \\ t > R}} \delta_{t,p} + \varepsilon_{t,p}(p)$ is mean zero a calculation gives that the distribution takes value $1 - x$ with probability $\frac{1+x}{2}$, however we don't assume that the noise terms are mean zero. We do however know that they are bounded and therefore have bounded mean conditioned on Δ_R and so we introduce $\tilde{x}$ to stand in for x, assuming that we take value $1 - x$ with probability $\frac{1+\tilde{x}}{2}$ where $|\tilde{x} - x| \le \gamma$. Here $|\tilde{x} - x| \le \gamma$ holds for $\gamma = (2\varepsilon_{pr} + \varepsilon_{iter})n$ by Lemma B.2.

We can therefore take expectations and rearrange to get

$$\mathbb{E}\left[\exp\left(\sum_{t=R+1}^{n} (\delta_{t,p} + \varepsilon_{t,p}(p)) \eta_R - \frac{1}{2} a_R b_R - \gamma c_R \right) \middle| \Delta_R \right]$$

$$
\begin{aligned}
&= \frac{1+\tilde{x}}{2}\exp((1-x)\eta_R - \frac{1}{2}a_R b_R - \gamma c_R) + \frac{1-\tilde{x}}{2}\exp((-1-x)\eta_R - \frac{1}{2}a_R b_R - \gamma c_R) \\
&= \exp(-\gamma c_R)\left[\frac{1+\tilde{x}}{2}\exp((1-x)\eta_R - \frac{1}{2}a_R b_R) + \frac{1-\tilde{x}}{2}\exp((-1-x)\eta_R - \frac{1}{2}a_R b_R)\right] \\
&= \exp((\tilde{x}-x)\eta_R - \gamma c_R) \\
&\qquad \cdot \left[\frac{1+\tilde{x}}{2}\exp((1-\tilde{x})\eta_R - \frac{1}{2}a_R b_R) + \frac{1-\tilde{x}}{2}\exp((-1-\tilde{x})\eta_R - \frac{1}{2}a_R b_R)\right],
\end{aligned}
$$

where in the last step we used $1 - x = (1 - \tilde{x}) + (\tilde{x} - x)$ and $-1 - x = (-1 - \tilde{x}) + (\tilde{x} - x)$, factoring out $\exp((\tilde{x} - x)\eta_R)$. This gives that

$$
\left[\frac{1+\tilde{x}}{2}\exp((1-\tilde{x})\eta_R - \frac{1}{2}a_R b_R) + \frac{1-\tilde{x}}{2}\exp((-1-\tilde{x})\eta_R - \frac{1}{2}a_R b_R)\right] = f_{\tilde{x}}\begin{bmatrix} a_R & \eta_R \\ \eta_R & b_R \end{bmatrix}.
$$

We note that $\frac{1+\tilde{x}}{2} \in [0, 1]$ as it corresponds to the probability of a certain event and so $\tilde{x} \in [-1, 1]$ so $f_{\tilde{x}}$ is defined. We can therefore apply lemma A.1 as for any r the matrix $\begin{bmatrix} \alpha_r^2 & \alpha_r\beta_r \\ \alpha_r\beta_r & \beta_r^2 \end{bmatrix}$ is positive semi-definite and so $\begin{bmatrix} a_R & \eta_R \\ \eta_R & b_R \end{bmatrix}$ is a sum of positive semi-definite matrices.

We can therefore split off the final term of the summation common to a_R, b_R, η_R and apply the fact f is operator monotone decreasing to get

$$
\begin{aligned}
&\mathbb{E}\left[\exp\left(\sum_{t=R+1}^{n}(\delta_{t,p} + \varepsilon_{t,p}(p))\eta_R - \frac{1}{2}a_R b_R - \gamma c_R\right)\middle|\,\Delta_R\right] \\
&= \exp((\tilde{x}-x)\eta_R - \gamma c_R) f_{\tilde{x}}\begin{bmatrix} a_R & \eta_R \\ \eta_R & b_R \end{bmatrix} \\
&\leq \exp((\tilde{x}-x)\eta_R - \gamma c_R) f_{\tilde{x}}\begin{bmatrix} a_{R-1} & \eta_{R-1} \\ \eta_{R-1} & b_{R-1} \end{bmatrix} \\
&= \exp((\tilde{x}-x)\eta_R - \gamma c_R) \\
&\quad \cdot \left[\frac{1+\tilde{x}}{2}\exp((1-\tilde{x})\eta_{R-1} - \frac{1}{2}a_{R-1} b_{R-1}) + \frac{1-\tilde{x}}{2}\exp((-1-\tilde{x})\eta_{R-1} - \frac{1}{2}a_{R-1} b_{R-1})\right],
\end{aligned}
$$

where we can now use a similar manipulation to factor out $\exp(x - \tilde{x})\eta_{R-1} = \exp(\tilde{x} - x)(-\eta_{R-1})$ to exchange the inner uses of $\tilde{x}$ with x, matching the expression for the original expectation.

$$
\begin{aligned}
&= \exp((\tilde{x}-x)(\eta_R - \eta_{R-1}) - \gamma c_R) \\
&\quad \cdot \left[\frac{1+\tilde{x}}{2}\exp((1-x)\eta_{R-1} - \frac{1}{2}a_{R-1} b_{R-1}) + \frac{1-\tilde{x}}{2}\exp((-1-x)\eta_{R-1} - \frac{1}{2}a_{R-1} b_{R-1})\right]
\end{aligned}
$$

Note that we have $|\tilde{x} - x| \leq \gamma$ by Lemma B.2, and so by the definitions of η_R, c_R we have $(\tilde{x} - x)(\eta_R - \eta_{R-1}) - \gamma c_R \leq \gamma \sum_{r \in Q_R} |\alpha_r \beta_r| - \gamma c_R = -\gamma c_{R-1}$. This gives

$$
\begin{aligned}
&\leq \exp(-\gamma c_{R-1})\left[\frac{1+\tilde{x}}{2}\exp((1-x)\eta_{R-1} - \frac{1}{2}a_{R-1} b_{R-1})\right. \\
&\qquad \left. + \frac{1-\tilde{x}}{2}\exp((-1-x)\eta_{R-1} - \frac{1}{2}a_{R-1} b_{R-1})\right] \\
&= \left[\frac{1+\tilde{x}}{2}\exp((1-x)\eta_{R-1} - \frac{1}{2}a_{R-1} b_{R-1} - \gamma c_{R-1})\right. \\
&\qquad \left. + \frac{1-\tilde{x}}{2}\exp((-1-x)\eta_{R-1} - \frac{1}{2}a_{R-1} b_{R-1} - \gamma c_{R-1})\right]
\end{aligned}
$$

$$= \mathbb{E}\left[\exp\left(\sum_{t=R+1}^{n}(\delta_{t,p}+\varepsilon_{t,p}(p))\eta_{R-1} - \frac{1}{2}a_{R-1}b_{R-1} - \gamma c_{R-1}\right)\middle|\,\Delta_R\right]$$

as desired. Having established this inequality we proceed to show $g(R) \le g(R-1)$ as discussed. We will apply iterated expectations and condition on the first R iterations of the algorithm, applying the lemma to the inner expectation, noticing that the resulting expression recovers $g(R-1)$ exactly. We have

$$\begin{aligned}
g(R) &:= \mathbb{E}\Bigg[\exp\Bigg(\sum_{t=t_p}^{n}(\delta_{t,p}+\varepsilon_{t,p}(p)) \sum_{s=t_p-1}^{\min\{R,\,t-1\}} \sum_{r\in Q_{s,p}} \alpha_r\beta_r \\
&\quad - \frac{1}{2}\left(\sum_{s=t_p-1}^{R}\sum_{r\in Q_{s,p}}\alpha_r^2\right)\left(\sum_{s=t_p-1}^{R}\sum_{r\in Q_{s,p}}\beta_r^2\right) - \gamma\sum_{s=t_p-1}^{R}\sum_{r\in Q_{s,p}}|\alpha_r\beta_r|\Bigg)\,\Bigg|\,\Delta_p\Bigg] \\
&= \mathbb{E}\Bigg[\mathbb{E}\Bigg[\exp\Bigg(\sum_{t=t_p}^{n}(\delta_{t,p}+\varepsilon_{t,p}(p)) \sum_{s=t_p-1}^{\min\{R,\,t-1\}} \sum_{r\in Q_{s,p}} \alpha_r\beta_r \\
&\quad - \frac{1}{2}\left(\sum_{s=t_p-1}^{R}\sum_{r\in Q_{s,p}}\alpha_r^2\right)\left(\sum_{s=t_p-1}^{R}\sum_{r\in Q_{s,p}}\beta_r^2\right) - \gamma\sum_{s=t_p-1}^{R}\sum_{r\in Q_{s,p}}|\alpha_r\beta_r|\Bigg)\,\Bigg|\,\Delta_R\Bigg]\,\Bigg|\,\Delta_p\Bigg] \\
&= \mathbb{E}\Bigg[\exp\Bigg(\sum_{t=t_p}^{R}(\delta_{t,p}+\varepsilon_{t,p}(p)) \sum_{s=t_p-1}^{t-1} \sum_{r\in Q_{s,p}} \alpha_r\beta_r\Bigg) \\
&\quad \cdot \mathbb{E}\left[\exp\left(\sum_{t=R+1}^{n}(\delta_{t,p}+\varepsilon_{t,p}(p))\eta_R - \frac{1}{2}a_Rb_R - \gamma c_R\right)\middle|\,\Delta_R\right]\,\Bigg|\,\Delta_p\Bigg] \\
&\le \mathbb{E}\Bigg[\exp\Bigg(\sum_{t=t_p}^{R}(\delta_{t,p}+\varepsilon_{t,p}(p)) \sum_{s=t_p-1}^{t-1} \sum_{r\in Q_{s,p}} \alpha_r\beta_r\Bigg) \\
&\quad \cdot \mathbb{E}\left[\exp\left(\sum_{t=R+1}^{n}(\delta_{t,p}+\varepsilon_{t,p}(p))\eta_{R-1} - \frac{1}{2}a_{R-1}b_{R-1} - \gamma c_{R-1}\right)\middle|\,\Delta_R\right]\,\Bigg|\,\Delta_p\Bigg]
\end{aligned}$$

an check that in the outer summation ($t \le R$) the index s ranges up to $t-1 = \min\{R-1, t-1\}$, whereas in the inner summation ($t \ge R+1$) it ranges up to $R-1 = \min\{R-1, t-1\}$, allowing us to recombine as follows:

$$\begin{aligned}
&= \mathbb{E}\Bigg[\mathbb{E}\Bigg[\exp\Bigg(\sum_{t=t_p}^{n}(\delta_{t,p}+\varepsilon_{t,p}(p)) \sum_{s=t_p-1}^{\min\{R-1,\,t-1\}} \sum_{r\in Q_{s,p}} \alpha_r\beta_r \\
&\quad - \frac{1}{2}\left(\sum_{s=t_p-1}^{R-1}\sum_{r\in Q_{s,p}}\alpha_r^2\right)\left(\sum_{s=t_p-1}^{R-1}\sum_{r\in Q_{s,p}}\beta_r^2\right) - \gamma\sum_{s=t_p-1}^{R-1}\sum_{r\in Q_{s,p}}|\alpha_r\beta_r|\Bigg)\,\Bigg|\,\Delta_R\Bigg]\,\Bigg|\,\Delta_p\Bigg] \\
&= \mathbb{E}\Bigg[\exp\Bigg(\sum_{t=t_p}^{n}(\delta_{t,p}+\varepsilon_{t,p}(p)) \sum_{s=t_p-1}^{\min\{R-1,\,t-1\}} \sum_{r\in Q_{s,p}} \alpha_r\beta_r \\
&\quad - \frac{1}{2}\left(\sum_{s=t_p-1}^{R-1}\sum_{r\in Q_{s,p}}\alpha_r^2\right)\left(\sum_{s=t_p-1}^{R-1}\sum_{r\in Q_{s,p}}\beta_r^2\right) - \gamma\sum_{s=t_p-1}^{R-1}\sum_{r\in Q_{s,p}}|\alpha_r\beta_r|\Bigg)\,\Bigg|\,\Delta_p\Bigg] \\
&= g(R-1)
\end{aligned}$$

as claimed. We remark that the case $g(t_p - 1) \le 1$ can be shown with the same argument as above, without needing to apply iterated expectations as this iteration only counts the pivot b_p as decided.

In particular, with $x = z_{t_p}(p)$ we have

$$\begin{aligned} g(t_p - 1) &= \mathbb{E}\left[\exp\left(\sum_{t=t_p}^{n}(\delta_{t,p} + \varepsilon_{t,p}(p))\eta_{t_p-1} - \frac{1}{2}a_{t_p-1}b_{t_p-1} - \gamma c_{t_p-1}\right)\middle|\,\Delta_p\right] \\ &= \exp((\tilde{x} - x)\eta_{t_p-1} - \gamma c_{t_p-1}) f_{\tilde{x}}\begin{bmatrix} a_{t_p-1} & \eta_{t_p-1} \\ \eta_{t_p-1} & b_{t_p-1}\end{bmatrix} \\ &\leq 1 \cdot f_{\tilde{x}}\begin{bmatrix} a_{t_p-1} & \eta_{t_p-1} \\ \eta_{t_p-1} & b_{t_p-1}\end{bmatrix} \\ &\leq f_{\tilde{x}}\begin{bmatrix} 0 & 0 \\ 0 & 0\end{bmatrix} = 1 \end{aligned}$$

which completes the proof. □

B.2 Proof of Theorem 2.3

Proof of Theorem 2.3. As discussed in equation 8 we can rewrite our output $Bz - Bz_0$ as

$$\begin{aligned} Bz - Bz_0 &= B\left(\sum_{t=1}^{T}\hat{\delta}_t u_t + \hat{\delta}_t(\hat{u_t} - u_t) + \varepsilon_t\right) \\ &= B\sum_{t=1}^{T}\left(\hat{\delta}_t + \varepsilon_t(p_t)\right)u_t + B\left(\sum_{t=1}^{T}\hat{\delta}_t(\hat{u_t} - u_t) + \varepsilon_t - \varepsilon_t(p_t)u_t\right) \end{aligned}$$

where we are able to use a modified version of the proof in Harshaw et al. (2024) to bound the MGF of the first term. We use ε_{iter} as an upper bound on the infinity norm of ε_t, deferring the choice of a concrete constant. Considering the remaining terms observe that we have

$$\|B(\hat{\delta}_t(\hat{u_t} - u_t) + \varepsilon_t - \varepsilon_t(p_t)u_t)\|_2 \leq |\hat{\delta}_t|\|B(\hat{u_t} - u_t)\|_2 + \|B\varepsilon_t\|_2 + |\varepsilon_t(p_t)|\|Bu_t\|_2$$

and from here we note that $\|Bu_t\| \leq 1$ by the fact Bu_t is some projection of the pivot unit, and $\|B\varepsilon_t\| = \|\sum_i \varepsilon_t(i)b_i\| \leq \sum_i|\varepsilon_t(i)|\|b_i\| \leq \sum_i|\varepsilon_t(i)| \leq n\varepsilon_{iter}$ which together with the fact $|\hat{\delta}_t| \leq 2$ gives

$$\sum_{t=1}^{n}\|B(\hat{\delta}_t(\hat{u_t} - u_t) + \varepsilon_t - \varepsilon_t(p_t)u_t)\|_2 \leq \sum_{t=1}^{n}(2\varepsilon_{LS} + n\varepsilon_{iter} + \varepsilon_{iter}) \leq 2n\varepsilon_{LS} + n(n+1)\varepsilon_{iter}$$

which implies that $\exp(\sum_t\langle Be_t, v\rangle) \leq \exp((2n\varepsilon_{LS} + n(n+1)\varepsilon_{iter})\|v\|)$ by the Cauchy-Schwarz inequality where e_t is the split of error term discussed in 3. This allows us to show

$$\begin{aligned} \mathbb{E}[\exp(\langle Bz - Bz_0), v\rangle] &= \mathbb{E}\left[\exp\left(\sum_t(\delta_t + \varepsilon_t(p_t))\langle Bu_t, v\rangle + \langle Be_t, v\rangle\right)\right] \\ &\leq \exp\left(\frac{1}{2}\|v\|^2 + (\sqrt{n}\gamma + (2n\varepsilon_{LS} + n(n+1)\varepsilon_{iter}))\|v\|\right) \end{aligned}$$

by applying the above bound on the e_t terms, followed by lemma 3.2. As will be shown in section B.3, we can choose $\gamma = (2\varepsilon_{pr} + \varepsilon_{iter})n$ to get a bound of

$$\exp\left(\frac{1}{2}\|v\|^2 + \left(\underbrace{n^{3/2}(2\varepsilon_{pr} + \varepsilon_{iter})}_{\text{drift/sampling}} + \left(\underbrace{2n\cdot\varepsilon_{LS}}_{\text{least squares}} + \underbrace{n(n+1)\cdot\varepsilon_{iter}}_{\text{arithmetic/clamping}}\right)\right)\|v\|\right).$$

□

B.3 CHOOSING CONSTANTS

We now discuss how to choose the constant γ used in our result and how to further decompose $\varepsilon_{pr}, \varepsilon_{LS}$ to give a more precise characterization of error in terms of machine dependent quantities.

Lemma B.2. *For each iteration R within a pivot phase for pivot p, defining $\tilde{x}$ such that $\Pr[\sum_{R+1}^{n}(\delta_{t,p} + \varepsilon_{t,p}(p)) = 1 - x \,\big|\, \Delta_R] = \frac{1+\tilde{x}}{2}$ we have $|\tilde{x} - x| \leq (2\varepsilon_{pr} + \varepsilon_{iter})n$ where $x := z_{R+1}(p)$, and in the ideal algorithm the corresponding probability is $\frac{1+x}{2}$.*

Proof. Focusing first on γ, we previously stated that γ is our upper bound on $|\tilde{x} - x|$ where $\tilde{x}$ is defined such that the probability that $\sum_{t=R+1}^{n}(\delta_{t,p} + \varepsilon_{t,p}(p))$ takes value $1 - x$, conditional on Δ_R, is $\frac{1+\tilde{x}}{2}$. We now motivate this definition further and elaborate on how we arrive at our concrete choice of γ.

Recall that, conditioned on Δ_R, $\sum_{t=R+1}^{n}(\delta_{t,p} + \varepsilon_{t,p}(p))$ obeys a two point distribution, taking values in $\{1 - x, -1 - x\}$. Therefore if we use μ to denote its conditional mean, we can write for some $q \in [0, 1]$ that

$$\begin{aligned}\mu &= q(1 - x) + (1 - q)(-1 - x)\\ &= 2q - 1 - x\end{aligned}$$

and therefore $q = \frac{1+x+\mu}{2}$. Setting $q = \frac{1+\tilde{x}}{2}$ gives that $\tilde{x} = x+\mu$ and so our assumption of $|\tilde{x}-x| \leq \gamma$ is therefore equivalent to assuming that $|\mu| \leq \gamma$. Put another way, $\tilde{x}$ is simply a notational alternative to writing $x+\mathbb{E}[\sum_{t=R+1}^{n}(\delta_{t,p}+\varepsilon_{t,p}(p)) \mid \Delta_R]$ throughout the proof, emphasizing that it is the result of perturbing x (which would be the value in the infinite precision case) slightly due to noise.

We now discuss the value of γ. To provide a value for γ it suffices to bound

$$\left|\mathbb{E}\left[\sum_{t=R+1}^{n}(\delta_{t,p} + \varepsilon_{t,p}(p)) \,\middle|\, \Delta_R\right]\right| = \left|\sum_{t=R+1}^{n}\mathbb{E}\big[\delta_{t,p} + \varepsilon_{t,p}(p) \,\big|\, \Delta_R\big]\right|$$

Let $\mathcal{F}_t$ contain the history and all computations preceding the random branch choice at iteration t, and let $\mathcal{H}_t$ contain the history through the completed update at iteration t. Thus $\Delta_R = \mathcal{H}_R$ and $\mathcal{H}_R \subseteq \mathcal{F}_t$ for $t > R$. Write $I_{t,p}$ for the indicator that iteration t belongs to the phase of pivot p, noting that this indicator is $\mathcal{F}_t$-measurable. Before termination, $\widehat{p}_t$ differs from the probability centering the computed endpoints by at most $\varepsilon_{\mathrm{pr}}$, so

$$\left|\mathbb{E}[\widehat{\delta}_t \mid \mathcal{F}_t]\right| = (\widehat{\delta}_t^{+} + \widehat{\delta}_t^{-})|\widehat{p}_t - p_t^{\mathrm{comp}}| \leq 2\varepsilon_{\mathrm{pr}}.$$

After termination the padded step has zero conditional mean. Moreover, $|\varepsilon_t(p)| \leq \varepsilon_{\mathrm{iter}}$, including when the update error depends on the sampled branch. Therefore,

$$\left|\mathbb{E}\Big[I_{t,p}\big(\widehat{\delta}_t + \varepsilon_t(p)\big) \mid \mathcal{F}_t\Big]\right| \leq I_{t,p}(2\varepsilon_{\mathrm{pr}} + \varepsilon_{\mathrm{iter}}).$$

The tower property and the triangle inequality now give

$$\begin{aligned}|\mu| &= \left|\mathbb{E}\left[\sum_{t=R+1}^{n} I_{t,p}(\widehat{\delta}_t + \varepsilon_t(p)) \mid \mathcal{H}_R\right]\right|\\ &\leq \sum_{t=R+1}^{n}\mathbb{E}\Big[\Big|\mathbb{E}\Big[I_{t,p}(\widehat{\delta}_t + \varepsilon_t(p)) \mid \mathcal{F}_t\Big]\Big| \mid \mathcal{H}_R\Big]\\ &\leq (n - R)(2\varepsilon_{\mathrm{pr}} + \varepsilon_{\mathrm{iter}}) \leq \gamma\end{aligned}$$

Since $\widetilde{x} = x + \mu$, this proves $|\widetilde{x} - x| \leq \gamma$ for $\gamma := n(2\varepsilon_{\mathrm{pr}} + \varepsilon_{\mathrm{iter}})$.

□

Lemma B.3 (Bounding sampling error, arithmetic error, in terms of machine precision)**.** *On a floating point machine with b random bits, we have $\varepsilon_{pr} \leq \frac{2u}{1-2u} + 2^{-b}$ for unit roundoff u.*

Proof. We note that in the following discussions u refers to unit roundoff while u_t is the step direction at time t. In the following discussion we condition on the history and all computations preceding the random branch choice at iteration t. Let $p_t^{\mathrm{comp}} := \frac{\hat{\delta}_t^-}{\hat{\delta}_t^- + \hat{\delta}_t^+}$, where $\hat{\delta}_t^+ \geq 0$, $\hat{\delta}_t^- \geq 0$, and $\hat{\delta}_t^+ + \hat{\delta}_t^- > 0$. Here, p_t^{comp} is the exact ratio of the stored endpoint magnitudes, and it hence centers the step at the computed endpoints.

We assume binary floating-point arithmetic with L significand bits (including the leading bit), rounding to nearest, and no overflow or underflow in the operations below. Here the unit roundoff is $u = 2^{-L}$. Compute $s_t = \mathrm{fl}(\hat{\delta}_t^- + \hat{\delta}_t^+)$, $r_t = \mathrm{fl}(\hat{\delta}_t^- / s_t)$, and $q_t = \min\{1, \max\{0, r_t\}\}$.

Then, by the standard floating-point model (in Higham's lemma, (Higham, 2002)), there exists $\theta_{1,t}, \theta_{2,t}$ with $|\theta_{i,t}| \leq u$ such that $r_t = p_t^{\mathrm{comp}} \frac{1+\theta_{2,t}}{1+\theta_{1,t}}$. Since clipping to $[0,1]$ cannot increase the error and $p_t^{\mathrm{comp}} \in [0,1]$, for $2u < 1$ we have

$$\begin{aligned} |q_t - p_t^{\mathrm{comp}}| &\leq |r_t - p_t^{\mathrm{comp}}| \\ &= p_t^{\mathrm{comp}} \frac{|\theta_{2,t} - \theta_{1,t}|}{|1 + \theta_{1,t}|} \\ &\leq \frac{2u}{1-u} \\ &\leq \frac{2u}{1-2u}. \end{aligned}$$

To sample using b fresh independent unbiased random bits, form $k_t = \lfloor 2^b q_t \rfloor$ exactly, draw J_t uniformly from $\{0, \ldots, 2^b - 1\}$ independently of the conditioned history, and choose the positive branch if $J_t < k_t$. The resulting conditional positive-branch probability is $\hat{p}_t = k_t / 2^b$, so we have that $|\hat{p}_t - q_t| < 2^{-b}$. Then, the triangle inequality therefore gives, uniformly over all iterations and histories, that

$$|\hat{p}_t - p_t^{\mathrm{comp}}| \leq \frac{2u}{1-2u} + 2^{-b}.$$

Therefore, we can set $\varepsilon_{\mathrm{pr}} = \frac{2^{1-L}}{1-2^{1-L}} + 2^{-b}$.

□

Lemma B.4 (Least-squares image error from backward stability)**.** *Suppose that $B \in \mathbb{R}^{m \times n}$ has full column rank and $\|b_i\|_2 \leq 1$ for every i. Write $\kappa_B := \|B\|_2 / \sigma_{\min}(B)$. At iteration t, let $S_t = A_t \setminus \{p_t\}$, $M_t = B_{S_t}$ and $b = b_{p_t}$. The exact nonpivot coefficients solve $x_t \in \arg\min_x \|M_t x + b\|_2$. Suppose that the computed coefficients $\hat{x}_t$ satisfy the following normwise backward-error guarantee, uniformly over all iterations and execution histories: for some perturbations E_t, f_t,*

$$\hat{x}_t \in \arg\min_x \|(M_t + E_t)x + b + f_t\|_2,$$

where $\|E_t\|_2 \leq \eta \|M_t\|_2$ and $\|f_t\|_2 \leq \eta \|b\|_2$, where $\eta \geq 0$ is deterministic. Form $u_t, \hat{u}_t$ by placing $x_t, \hat{x}_t$ on S_t, setting the pivot entry to one, and setting all inactive entries to zero. If $\eta \kappa_B < 1$, then we have that $\|B(\hat{u}_t - u_t)\|_2 \leq \varepsilon_{\mathrm{LS}}$ for every t, where

$$\varepsilon_{\mathrm{LS}} \leq \frac{3\eta\kappa_B}{1 - \eta\kappa_B}.$$

In particular, if $\eta\kappa_B \leq \frac{1}{2}$, one may instead take $\varepsilon_{\mathrm{LS}} = 6\eta\kappa_B$.

Proof. We clarify that the assumptions of full column rank on B and backwards stability in solvers are not true in general and that this result is meant to provide an example for what ε_{LS} scales like in certain scenarios. In particular the Higgs matrix in the experiments is an example of a wide matrix that lacks full column rank. When $S_t = \emptyset$, both directions equal the pivot unit vector and the image error is zero. Otherwise, suppress the iteration index and write $M = M_t$, $\widetilde{M} = M + E$, and $q = \eta\kappa_2(M)$. Since M consists of columns of B, we have that $\sigma_{\min}(M) \geq \sigma_{\min}(B)$, with $\|M\|_2 \leq \|B\|_2$ and $\kappa_2(M) \leq \kappa_B$. Thus $q < 1$, and

$$\sigma_{\min}(\widetilde{M}) \geq \sigma_{\min}(M) - \|E\|_2$$

$$\geq (1-q)\sigma_{\min}(M) > 0.$$

Let P and $\widetilde{P}$ be the orthogonal projectors onto the column spaces of M and $\widetilde{M}$. These spaces have equal dimension, so the equal-rank projector identity gives

$$\begin{aligned}\|P-\widetilde{P}\|_2 &= \|(I-\widetilde{P})P\|_2 \\ &= \|(I-\widetilde{P})EM^\dagger\|_2 \\ &\leq \frac{\|E\|_2}{\sigma_{\min}(M)} \leq q,\end{aligned}$$

where $M^\dagger$ denotes the Moore-Penrose pseudoinverse and $(I-\widetilde{P})M = -(I-\widetilde{P})E$. Then, least-squares optimality gives $Mx_t = -Pb$, $\widetilde{M}\hat{x}_t = -\widetilde{P}(b+f)$, and $\|\hat{x}_t\|_2 \leq \frac{(1+\eta)\|b\|_2}{(1-q)\sigma_{\min}(M)}$.

Hence, we have that

$$\begin{aligned}\|B(\hat{u}_t - u_t)\|_2 &= \|M(\hat{x}_t - x_t)\|_2 \\ &= \|(P-\widetilde{P})b - \widetilde{P}f - E\hat{x}_t\|_2 \\ &\leq \left(q+\eta+\frac{q(1+\eta)}{1-q}\right)\|b\|_2 \\ &= \frac{\eta+2q-q^2}{1-q}\|b\|_2 \\ &\leq \frac{3q}{1-q} \\ &\leq \frac{3\eta\kappa_B}{1-\eta\kappa_B},\end{aligned}$$

where we used $\|b\|_2 \leq 1$, $\kappa_2(M) \geq 1$, and $\eta \leq q$, which completes the proof. The simplified bound comes from substituting $\eta\kappa_B \leq \frac{1}{2}$. □

Remark B.5. *For a solver whose backward-error guarantee is $\eta \leq C_{\mathrm{LS}}(m,n)u$, with the dimension-dependent factor chosen uniformly over all active systems, the sufficient precision condition $C_{\mathrm{LS}}(m,n)u\kappa_B \leq \frac{1}{2}$ therefore permits $\varepsilon_{\mathrm{LS}} = 6C_{\mathrm{LS}}(m,n)\kappa_B u$. Its contribution to the linear term in Theorem 2.3 is bounded by*

$$2n\varepsilon_{\mathrm{LS}} \leq 12nC_{\mathrm{LS}}(m,n)\kappa_B u.$$

In particular the discussion in B.3, B.4 and the relevant discussions imply a bound of

$$\begin{aligned}\mathbb{E}[\exp(\langle B(z-z_0), v\rangle)] \leq \exp\Bigg(\frac{1}{2}\|v\|^2 + \Bigg(n^{3/2}\left(2\left(\frac{2u}{1-2u}+2^{-b}\right)+\varepsilon_{iter}\right) \\ + \left(12nC_{\mathrm{LS}}(m,n)\kappa_B u + n(n+1)\varepsilon_{iter}\right)\Bigg)\|v\|\Bigg).\end{aligned}$$

C Lower Bound

Theorem C.1 (Lower bound for bounded update errors)**.** *There exists a universal constant $c > 0$ such that, for every integer $n \geq 8$ and every $\varepsilon \geq 0$, there is a full-column-rank matrix $\boldsymbol{B} = [b_1, \ldots, b_n] \in \mathbb{R}^{(n+1)\times n}$ with $\|b_i\|_2 = 1$ and an admissible error process for the Gram–Schmidt Walk, initialized at $z_0 = 0$, such that $z_{t+1} = z_t + \delta_t u_t + \varepsilon_t$ with $\|\varepsilon_t\|_\infty \leq \varepsilon$, with exact step-selection probabilities, and whose terminal discrepancy $D = \boldsymbol{B}z_T$ satisfies*

$$\|\mathbb{E}D\|_2 \geq c\min\{n^2\varepsilon, n\}, \qquad \|D\|_{\psi_2} \geq c\min\{n^2\varepsilon, n\}.$$

Here $\|\cdot\|_{\psi_2}$ denotes the uncentered directional Orlicz norm. The errors are deterministic functions of the iteration, active set, and current pivot, which preserve all frozen coordinates, and the walk remains feasible and terminates after at most n iterations.

Proof. Let $e_0, e_1, \ldots, e_n$ be the standard basis of $\mathbb{R}^{n+1}$, and for all $i \in [n]$, set $b_i = \frac{e_0+e_i}{\sqrt{2}}$. We use a zero index based notation here for convenience; this convention is local to the lower bound. Each column has Euclidean norm one, and

$$\boldsymbol{B}^\top \boldsymbol{B} = \frac{1}{2}\left(\boldsymbol{I}_n + \mathbf{1}\mathbf{1}^\top\right)$$

is positive definite, so $\boldsymbol{B}$ has full column rank.

We first identify the exact update direction. For an active set A of size k and pivot $p \in A$, every feasible direction satisfies $u(p) = 1$ and $u(j) = 0$ for $j \notin A$. Moreover,

$$2\|\boldsymbol{B}u\|_2^2 = \left(\sum_{j\in A} u(j)\right)^2 + \sum_{j\in A} u(j)^2.$$

Minimizing this strictly convex quadratic gives $u(j) = -\frac{1}{k}$ for $j \in A \setminus \{p\}$ and $\sum_{j\in A} u(j) = \frac{1}{k}$. In particular, we have $\langle Bu, e_0\rangle = \frac{1}{\sqrt{2}\,k}$. Set $L = \left\lfloor \frac{n}{4} \right\rfloor$ with $a = \min\left\{\varepsilon, \frac{1}{8n}\right\}$. Writing A_t and p_t for the active set and pivot before iteration t, define

$$\varepsilon_t(j) = \begin{cases} a, & 0 \le t < L, \quad j \in A_t \setminus \{p_t\}, \\ 0, & \text{otherwise.} \end{cases}$$

Thus $\|\varepsilon_t\|_\infty \le \varepsilon$, and the errors vanish on the pivot and all previously frozen coordinates.

We now verify via mathematical induction that during the first L iterations, both feasible endpoints are determined by the pivot and exactly that coordinate freezes. We argue by induction. Suppose this holds before iteration $t < L$, so that $|A_t| = n - t$. The two candidate pivot endpoints correspond to signed steps $\delta = 1 - z_t(p_t)$ and $\delta = -1 - z_t(p_t)$, each of magnitude at most two.

Any $j \in A_t \setminus \{p_t\}$ has been an active non-pivot coordinate in every preceding iteration: a previous pivot would already have frozen. Hence, for either candidate step, its ideal updated value and its perturbed updated value both have absolute value at most

$$\begin{aligned} \sum_{s=0}^{t}\left(\frac{2}{n-s} + a\right) &\le \frac{2L}{n-L+1} + La \\ &\le \frac{2}{3} + \frac{1}{32} = \frac{67}{96} < 1. \end{aligned}$$

Previously frozen coordinates remain unchanged, while the pivot reaches 1 or -1 and receives zero error. Therefore both candidate endpoints are feasible before and after adding the error. Since extending either candidate step would move the pivot outside the cube, these are precisely the exact feasible endpoints. Every other active coordinate remains strictly interior, so exactly the pivot freezes. This completes the induction. After iteration $L - 1$, all errors vanish and the algorithm continues as an exact Gram-Schmidt Walk. Thus the entire process is admissible and terminates at some $T \le n$. Now, it remains to bound the accumulated discrepancy. Let

$$Y = \langle Bz_T, e_0\rangle.$$

We condition on the history before selecting the step, including the current pivot. The exact step-selection probabilities ensure that $\mathbb{E}[\delta_t \mid \mathcal{F}_t] = 0$. Since u_t is determined by this history, each ideal discrepancy increment has conditional mean zero. Padding terminated runs with zero updates therefore gives

$$\begin{aligned} \mathbb{E}Y &= \mathbb{E}\sum_{t=0}^{T-1}\langle B\varepsilon_t, e_0\rangle \\ &= \frac{a}{\sqrt{2}}\sum_{t=0}^{L-1}(n-t-1). \end{aligned}$$

For $n \ge 8$, we have $L \ge n/8$ and $n - L \ge 3n/4$. Consequently,

$$\mathbb{E}Y \ge \frac{a}{\sqrt{2}}L(n-L)$$

$$\geq \frac{3}{32\sqrt{2}} an^2.$$

Furthermore, $an^2 = \min\{n^2\varepsilon, \frac{n}{8}\} \geq \frac{1}{8}\min\{n^2\varepsilon, n\}$, which proves the lower bound on $\|\mathbb{E}D\|_2$ as e_0 is a unit vector. Finally, for every $K > 0$, Jensen's inequality gives

$$\mathbb{E}\exp(Y^2/K^2) \geq \exp\big((\mathbb{E}Y)^2/K^2\big).$$

Using $\|Y\|_{\psi_2} = \inf\{K > 0 : \mathbb{E}\exp(Y^2/K^2) \leq 2\}$, it follows that

$$\begin{aligned}\|D\|_{\psi_2} &\geq \|Y\|_{\psi_2} \\ &\geq \frac{|\mathbb{E}Y|}{\sqrt{\log 2}} \\ &= \Omega\big(\min\{n^2\varepsilon, n\}\big),\end{aligned}$$

which completes the proof. □

D Experimental Details

For the distributional experiments in Figures 5–7, we generate $R = 1000$ independent walks per configuration, parallelized across CPU processes on a six-core laptop with single-threaded BLAS. For each tested direction d, we form projections $X_r(d) = d^\top B z^{(r)}$ and estimate their concentration scale using

$$\widehat{\sigma}(d) = \max_{j\in\{1,2,3\}} \left(\frac{R^{-1}\sum_{r=1}^{R} X_r(d)^{2j}}{(2j-1)!!} \right)^{1/2j}.$$

We also compute empirical ψ_2 norms by bisection, replacing the expectation in the Orlicz definition with the sample average. We select the direction maximizing the moment diagnostic over the finite test set and evaluate its Orlicz norm and standardized empirical survival curve.

Experimental limitations. Our implementation does not satisfy all the hypotheses of Section 3. We round endpoint ratios to nearest without enforcing the downward pivot caps, use tolerance-based activity and snapping, permit early termination when the computed endpoint sum is small, and convert the terminal state to signs. Our implementation computes directions using a rounded copy of the input and replaces nonfinite solver coefficients by zero, while reporting discrepancy with the original matrix B_0.

Metrics. We report exploratory concentration diagnostics over a finite set of unit directions. The moment statistic uses orders $2, 4, 6$; neither it, nor the empirical Orlicz norm, is a certified upper bound on the population subgaussian parameters, but are fair estimators. The two scales also have different normalizations: for $G \sim N(0, \sigma^2)$, we have that $\|G\|_{\psi_2} = \sqrt{8/3}\,\sigma$, whereas a Rademacher variable has Orlicz norm $1/\sqrt{\log 2}$. Our projections are uncentered, so increases can reflect bias and spread. The reported direction maximizes the moment diagnostic over the tested set using the same walks. The error bars we present are the standard deviations of maxima over five batches, and zero exceedance counts are displayed at $1/2R$ solely for plotting on a logarithmic axis; this display floor is not an estimated nonzero tail probability. For instance, with $R = 1000$, the sampled tails do not resolve probabilities substantially below 10^{-3}.

Discussion. Figure 5 shows decreasing estimated scales as precision increases, followed by a plateau; its directional bands summarize variation across tested directions. Figure 6 compares the rescaled empirical tails across problem sizes, suggesting broadly similar behavior over the observed range without establishing a common population tail shape. Figure 7 compares rounding with independently injected noise and examines candidate rescalings; the differing curves do not establish universal accumulation laws. Finally, Figure 8 uses 150 walks with $m = 30$, $n = 200$, and eight retained fraction bits. Along each float64 reference trajectory, we recompute directions in low precision at identical states and summarize their coefficient differences by the number of free nonpivot coordinates. The median and interquartile ranges reveal a pronounced error spike near $k = m$.

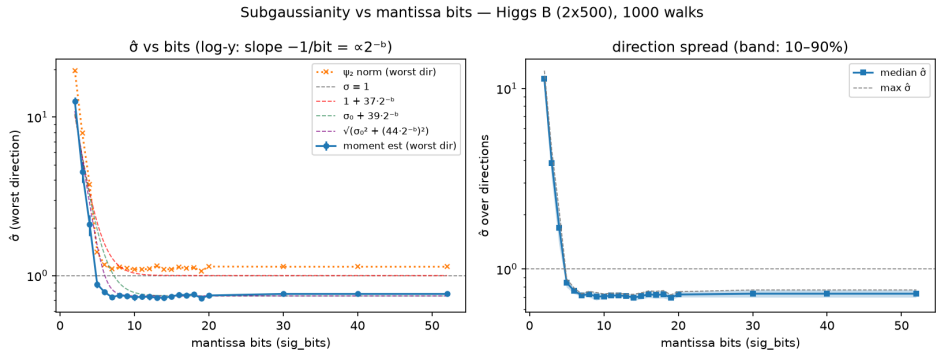


Figure 5: Worst-direction $\hat{\sigma}$ vs. mantissa bits at fixed $n = 500$. The plateau ends near $b \approx 8$. Three one-parameter fits describe the low-precision side nearly equally well; of these, the quadrature form $\sqrt{\sigma_0^2 + (44 \cdot 2^{-b})^2}$ is the one consistent with the n-sweep. The dotted ψ_2 norm sits about $2\times$ above the moment estimate throughout. Right: the 10–90% direction band.

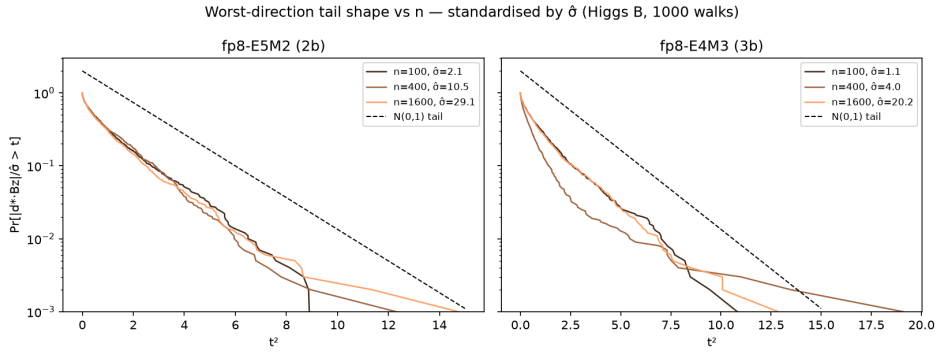


Figure 6: **Standardized empirical tails at two and three retained fraction bits for $n = 100, 400, 1600$.** We plot $\widehat{\Pr}\{|d_*^\top Bz|/\widehat{\sigma} > t\}$ using the direction and scale selected from the same samples. The dashed curve is $2e^{-t^2/2}$, which we use as a reference envelope. The curves show broadly similar behavior over part of the observed range, with deviations in the empirical far tail.

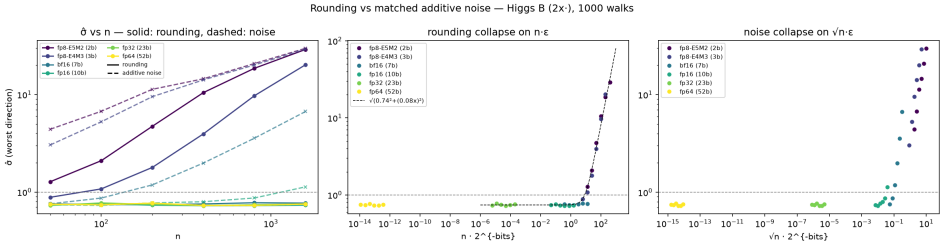


Figure 7: **Rounding is not additive noise.** Solid curves run the walk's arithmetic at b mantissa bits, dashed curves inject i.i.d. noise of the same magnitude 2^{-b} per step. Rounding is reused coherently and grows like $n\varepsilon$, so the cells collapse on the $n\varepsilon$ axis (middle). Independent noise accumulates like $\sqrt{n}\,\varepsilon$ and does not collapse (right).

D.1 DTYPE-BASED SOLVER COMPARISONS

In this subsection, a walk is classified as failed if it encounters a nonfinite direction or branch probability, a Cholesky or rank-one downdate breakdown, or exhausts the step budget before reaching a complete signing.

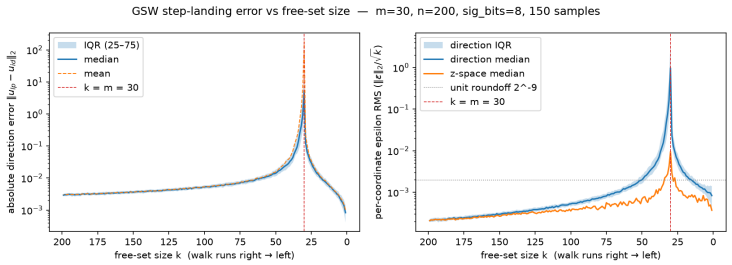


Figure 8: Coefficient differences on the float64 reference trajectory (150 walks, $m = 30$, $n = 200$, 8 mantissa bits). The error stays a small fraction of the unit roundoff for most of the walk, then spikes by orders of magnitude when the free set reaches $k \approx m$ and $B_{A\setminus p}$ becomes nearly singular. We show the median and IQR because the rare blowups dominate the mean. This diagnostic compares coefficient vectors, not the image-error quantity $\|B(\widehat{u}_t - u_t)\|_2$, and it is evaluated on reference rather than perturbed trajectories.

The experiments we provide here are separate from the main fraction-width sweeps. They use dtype-based float64, float32, and float16 arithmetic, including pairwise reductions in the selected dtype. The test families comprise Gaussian matrices, matrices with controlled condition numbers, near-collinear column pairs, and Harshaw designs, with $n \in \{50, 200, 500\}$. Attempted sample counts vary by solver and configuration. The repository reports 6.9 hours of wall-clock time with 32 workers and 211 CPU-hours of walks for the EPYC grid; this compute budget is separate from that of the main precision sweeps.

Here, the failure rates include all attempted walks. Orlicz and bias diagnostics exclude failed walks and therefore describe outputs conditional on successful execution, not unconditional reliability. Discrepancy is evaluated using the original float64 matrix. Table 1 reproduces selected configurations from the stored EPYC-grid summaries, and the complete configuration tables are provided in the repository.

n	Solver	Attempts	Failures	Fail. (%)	Successes	$\widehat{K}_{\mathrm{succ}}$
50	QR	2000	0	0.00	2000	0.756
50	Cholesky	2000	1948	97.40	52	1.24
50	Cholesky, $k = 10$	2000	0	0.00	2000	1.37
50	Explicit inverse	2000	0	0.00	2000	0.754
200	QR	1000	0	0.00	1000	0.653
200	Cholesky	2000	2000	100.00	0	—
200	Cholesky, $k = 10$	2000	1	0.05	1999	6.72
200	Explicit inverse	2000	0	0.00	2000	1.19
500	QR	20	0	0.00	20	0.582
500	Cholesky	2000	2000	100.00	0	—
500	Cholesky, $k = 10$	2000	35	1.75	1965	19.2
500	Explicit inverse	300	0	0.00	300	3.36

Table 1: Float16 solver comparisons on correlated Harshaw designs with $d = 5$ and $\phi = 0.01$. QR recomputes the least-squares direction from scratch; $k = 10$ denotes Cholesky refactoring every ten steps. $\widehat{K}_{\mathrm{succ}}$ is the maximum empirical Orlicz estimate over 20 fixed random unit directions and the top and bottom left singular-vector directions, using only successful walks. A dash indicates no successful walks. These are selected stress configurations, not averages over the full grid. The QR estimate at $n = 500$ uses only 20 walks.

For these configurations, refactoring substantially reduces the observed failure rate but does not eliminate discrepancy inflation among successful walks. At $n = 500$, its 35 failures comprise 26 Cholesky breakdowns and nine downdate breakdowns. Across the 36 float16 Harshaw configurations with refactoring, the largest observed failure rate is 1.75%; all remaining configurations have rates at most 0.05%. These observations are specific to the tested configurations. Neither zero observed failures nor a small successful-run Orlicz estimate establishes an unconditional performance guarantee.