Conditional Hardness for Averaged-Condition Linear-System Solvers

July 31, 2026

Abstract

For a nonsingular N-by-N matrix A with singular values $\sigma_1(A) \geq \cdots \geq \sigma_N(A) > 0$, let $\bar{\kappa}_p(A) = \left(N^{-1}\sum_i(\sigma_i(A)/\sigma_N(A))^p\right)^{1/p}$ for finite $p > 0$. We consider high-accuracy bit-complexity algorithms for linear systems with running time either $\widetilde{O}(\mathrm{nnz}(A)\bar{\kappa}_p(A))$ or $\widetilde{O}(N^2\bar{\kappa}_p(A))$, corresponding to the two access-cost models in Problem 2.6 of Amsel et al. [ABB+26]. Assume that for dense n-by-n systems of condition number $\kappa = n^\alpha$, every high-accuracy solver requires $\widetilde{\Omega}(\min(n^\omega, n^{2+\alpha/2}))$ bit operations on an infinite hard family, and write $\beta(\alpha) = \min\{\omega, 2+\alpha/2\}$. Under this hypothesis, no sparsity-sensitive solver with time $\widetilde{O}(\mathrm{nnz}(A)\bar{\kappa}_p(A))$ exists for any fixed finite $p > 0$ satisfying

$$p < \max\left\{\frac{\beta(\alpha)-1}{\alpha}, \frac{1}{2+\alpha-\beta(\alpha)}\right\}.$$

For the dense-access model, no solver with time $\widetilde{O}(N^2\bar{\kappa}_p(A))$ exists for any fixed $p < (\beta(\alpha) - 2)/(2\alpha)$. In the mildly conditioned regime $0 < \alpha < 2\omega - 4$, this dense-access threshold is $p < 1/4$, while the sparse-access threshold includes $p < 2/\alpha$ whenever $\alpha < 2$. If the dense lower-bound hypothesis holds for arbitrarily small positive α, the sparse-access result rules out every fixed finite p, and the dense-access result rules out every fixed $p < 1/4$. The output model is a constant relative forward-error high-accuracy solution.

Note. This paper was fully AI generated by a harness using GPT 5.6 Sol Ultra from a prompt based on observations by Junzhao Yang and describing the padding-by-extra-entries transformation.

Contents

1 Introduction

Solving a linear system $Ax = b$ to high accuracy is one of the basic algorithmic tasks in numerical linear algebra. For dense n-by-n matrices, direct methods can be accelerated using fast matrix multiplication [Str69, BH74, ADW+25], while iterative methods such as conjugate gradients, LSQR, and GMRES [HS52, PS82, SS86] can exploit favorable conditioning and matrix-vector multiplication [GL13, LS13b]. These two viewpoints motivate the familiar dense benchmark $\min\{n^\omega, n^2\kappa^{1/2}\}$ for condition number κ, but they also make clear that the worst singular-value ratio is a blunt summary of the spectrum. Much of the modern algorithmic literature asks when one can replace worst-case condition-number dependence by a parameter that sees the whole singular-value distribution.

Several lines of work support this broader viewpoint, but with different models or additional structure. Randomized Kaczmarz, sketch-and-project, and coordinate-descent methods use Frobenius, trace, or sum-condition quantities for linear systems and regression in structured settings [SV09, LL10, GR15, LS13a, AKK+20]. Sketching and randomized numerical linear algebra give input-sparsity and preconditioning tools for least-squares and related problems [Sar06, RT08, CW17, HMT11, MT20]. More recent dense-system algorithms study spectra with a small number of large outliers or other spectral-tail control [DY24, DMY25, DLNR25, DNRY25]. These papers are important nearby or restricted results; their parameters are not, in general, identical to the averaged condition number studied below.

The precise object here is the finite-p averaged condition number. For a nonsingular N-by-N matrix A with singular values $\sigma_1(A) \geq \cdots \geq \sigma_N(A) > 0$, define

$$\bar{\kappa}_p(A) = \left(\frac{1}{N} \sum_{i=1}^{N} \left(\frac{\sigma_i(A)}{\sigma_N(A)} \right)^p \right)^{1/p}.$$

Problem 2.6 in the Simons workshop list of Amsel et al. [ABB+26] asks for the smallest $p > 0$ for which high-accuracy linear-system solving is possible in time $\widetilde{O}(T_A \bar{\kappa}_p(A))$, with $T_A = N^2$ in the dense-access model or $T_A = \mathrm{nnz}(A)$ in the input-sparsity-sensitive model. The closest same-problem positive result is the dense-access work of Dereziński and Sidford [DS26]; as recorded in the Simons problem statement, it gives the desired type of guarantee for all $p > 1/2$ for general systems and for all $p > 1/4$ for positive semidefinite systems. Sparse-side results such as the row-sampling and regression algorithms cited above are best viewed as partial or restricted versions of the same averaged-conditioning theme.

This paper gives a conditional lower-bound complement. The condition is an explicit dense optimality hypothesis: for dense n-by-n systems with condition number $\kappa = n^\alpha$, assume that every high-accuracy bit-complexity solver requires

$$\widetilde{\Omega}\left(\min\{n^\omega, n^{2+\alpha/2}\} \right)$$

time on an infinite hard family. This is a lower-bound assumption, not a consequence of the known upper bounds. Under it, identity padding turns a hypothetical averaged-condition solver into a faster dense solver in the ranges of p stated below.

Theorem 1.1 (Conditional hardness from identity padding). *Assume $\omega > 2$. Fix $\alpha > 0$ and set*

$$\beta(\alpha) = \min\{\omega, 2 + \alpha/2\}.$$

Assume that there is an infinite family of dense n-by-n nonsingular systems $Bx = b$ with $\sigma_n(B) = 1$ and $\sigma_1(B) = n^\alpha$ such that every high-accuracy bit-complexity solver requires $\widetilde{\Omega}(n^{\beta(\alpha)})$ time on this family.

Then there is no high-accuracy solver for all nonsingular $A \in \mathbb{C}^{N\times N}$ *with running time* $\widetilde{O}(\mathrm{nnz}(A)\bar{\kappa}_p(A))$ *for any fixed finite* $p > 0$ *satisfying*

$$p < \max\left\{\frac{\beta(\alpha)-1}{\alpha}, \frac{1}{2+\alpha-\beta(\alpha)}\right\}.$$

There is no high-accuracy solver for all nonsingular $A \in \mathbb{C}^{N\times N}$ *with running time* $\widetilde{O}(N^2\bar{\kappa}_p(A))$ *for any fixed finite* $p > 0$ *satisfying*

$$p < \frac{\beta(\alpha)-2}{2\alpha}.$$

The boundary cases are not claimed. In particular, if the dense lower-bound hypothesis holds for arbitrarily small positive α, *then no fixed finite* p *is possible in the* $T_A = \mathrm{nnz}(A)$ *model, and no fixed* $p < 1/4$ *is possible in the* $T_A = N^2$ *model.*

The proof also shows that, on hard families whose p-moment is $\Theta(n\kappa^p)$, optimizing over decoupled identity padding cannot improve the dense-access threshold beyond $p < 1/4$ in the mildly conditioned regime. This is a limitation of the padding reduction, not a lower bound against all possible dense-access algorithms.

The reduction is elementary but sensitive to the access term. Starting from a hard dense instance $Bx = b$ normalized so that $\sigma_n(B) = 1$, we append identity coordinates and solve $\mathrm{diag}(B, I_m)y = (b, 0)$. Projection onto the first n coordinates recovers a solution to the original system, while the padding changes $\bar{\kappa}_p$ according to a simple singular-value moment calculation. The sparse-access contradiction uses both PaddedSparseSolve, which pads to constant averaged condition number, and QuadraticallyPaddedSparseSolve, which pads by exactly n^2 coordinates. The dense-access contradiction uses PaddedDenseSolve and pays for the square of the padded dimension. This difference is the source of the stronger sparse threshold and the $p < 1/4$ dense threshold in the mildly conditioned regime.

The theorem deliberately avoids boundary and priority claims. It does not assert an unconditional lower bound, and it does not rule out dense-access methods at or above $p = 1/4$ by non-padding reductions. The proof does show that, on spectra with $S_p(B) = \Theta(n\kappa^p)$, merely optimizing the number of decoupled identity-padding coordinates cannot improve the dense-access padding comparison beyond this threshold.

The rest of the paper is organized as follows. The preliminaries define the averaged condition number, the high-accuracy output model, and the dense optimality hypothesis. The overview gives the padding calculation and the exponent comparisons. The final section proves the padding lemmas, the algorithmic reductions used in the contradiction, and Theorem 1.1.

2 Preliminaries

We use $\widetilde{O}(\cdot)$ and $\widetilde{\Omega}(\cdot)$ to suppress polylogarithmic factors in the relevant dimension and accuracy parameters. All matrices are over $\mathbb{C}$. For a matrix A, $\mathrm{nnz}(A)$ denotes the number of nonzero entries. If $A \in \mathbb{C}^{N\times N}$ is nonsingular, its singular values are ordered as

$$\sigma_1(A) \geq \cdots \geq \sigma_N(A) > 0,$$

and its condition number is $\kappa(A) = \sigma_1(A)/\sigma_N(A)$.

Definition 2.1 (Averaged condition number)**.** *Let $A \in \mathbb{C}^{N\times N}$ be nonsingular. For finite $p > 0$, define*

$$\bar{\kappa}_p(A) = \left(\frac{1}{N} \sum_{i=1}^{N} \left(\frac{\sigma_i(A)}{\sigma_N(A)} \right)^p \right)^{1/p}.$$

This is the finite-p averaged condition number from [ABB+26, Definition 2.5], specialized to square nonsingular systems.

Definition 2.2 (High-accuracy output model)**.** *Fix a sufficiently small constant $\varepsilon \in (0,1)$ independent of the dimension. A solver returns a high-accuracy solution to $Cx = d$ if its output z satisfies*

$$\|z - C^{-1}d\|_2 \le \varepsilon \|C^{-1}d\|_2.$$

Definition 2.3 (Averaged-condition solvers)**.** *Fix finite $p > 0$. A sparsity-sensitive p-averaged solver solves every nonsingular N-by-N system $Ax = b$ to high accuracy in bit time*

$$\widetilde{O}\left(\mathrm{nnz}(A)\bar{\kappa}_p(A)\right).$$

A dense-access p-averaged solver solves every such system in bit time

$$\widetilde{O}\left(N^2 \bar{\kappa}_p(A)\right).$$

Definition 2.4 (Dense optimality hypothesis)**.** *For $\alpha > 0$, the dense optimality hypothesis at exponent α asserts that there is an infinite sequence of dense n-by-n nonsingular linear-system instances $Bx = b$, with $\sigma_n(B) = 1$ and $\sigma_1(B) = n^\alpha$, for which every high-accuracy bit-complexity solver requires*

$$\widetilde{\Omega}\left(\min(n^\omega, n^{2+\alpha/2})\right)$$

time. Equivalently, after setting $\beta(\alpha) = \min\{\omega, 2 + \alpha/2\}$, the lower bound is $\widetilde{\Omega}(n^{\beta(\alpha)})$.

For a normalized seed matrix $B \in \mathbb{C}^{n\times n}$ with $\sigma_n(B) = 1$, define its finite-p singular-value moment by

$$S_p(B) = \sum_{i=1}^{n} \sigma_i(B)^p.$$

For an integer $m \ge 0$, the identity-padded instance associated with $Bx = b$ is

$$A_m(B) = \mathrm{diag}(B, I_m), \qquad c_m(b) = (b, 0) \in \mathbb{C}^{n+m},$$

where I_m is the m-by-m identity matrix. When B and b are clear from context, we write A_m and c_m.

Fact 2.5 (Block-diagonal singular values)**.** *The singular values of a block diagonal matrix are the multiset union of the singular values of its diagonal blocks. In particular, if $\sigma_n(B) = 1$, then $A_m(B)$ has smallest singular value 1, has padded dimension $n + m$, and satisfies*

$$\mathrm{nnz}(A_m(B)) = \mathrm{nnz}(B) + m.$$

3 Overview

The proof is a reduction from the hypothesized hard family of dense linear systems. We normalize each dense instance so that $\sigma_n(B) = 1$ and $\sigma_1(B) = \kappa = n^\alpha$. This normalization is useful because the padding coordinates can then be identities: adding copies of the singular value floor does not change the original solution, but it does dilute the average defining $\bar{\kappa}_p$. For

$$A_m(B) = \mathrm{diag}(B, I_m), \qquad c_m(b) = (b, 0),$$

the exact solution is $(B^{-1}b, 0)$, so any high-accuracy solution to $A_m y = c_m$ gives a high-accuracy solution to $Bx = b$ by taking the first n coordinates. The central calculation is

$$\bar{\kappa}_p(A_m)^p = \frac{S_p(B) + m}{n + m}, \qquad S_p(B) = \sum_{i=1}^{n} \sigma_i(B)^p.$$

Thus padding trades a smaller averaged condition number against a larger input size.

This tradeoff is the main technical issue. If an assumed solver has running time $\widetilde{O}(\mathrm{nnz}(A)\bar{\kappa}_p(A))$, then identity padding only increases the sparsity term additively by m. The proof uses two choices of m. The algorithm PaddedSparseSolve takes the dense input B, b, a bound $\kappa \geq \sigma_1(B)$, and the assumed sparsity-sensitive averaged-condition solver; it sets $m = \lceil n\kappa^p \rceil$, solves the padded system, and projects back. Since $S_p(B) \leq n\kappa^p$, this makes $\bar{\kappa}_p(A_m) = O_p(1)$ and yields a dense solver with running time

$$\widetilde{O}(n^2 + n\kappa^p).$$

With $\kappa = n^\alpha$, this contradicts the dense lower-bound exponent $\beta(\alpha) = \min\{\omega, 2 + \alpha/2\}$ whenever $p < (\beta(\alpha) - 1)/\alpha$.

The first padding scale is not always the best sparse comparison, because it may add more coordinates than are needed to beat the dense lower bound. The second algorithm, QuadraticallyPaddedSparseSolve, fixes $m = n^2$. It keeps the sparsity of the padded instance at the dense input scale while only partially reducing the averaged condition number:

$$\mathrm{nnz}(A_m)\bar{\kappa}_p(A_m) = \widetilde{O}\left(n^2 \left(1 + \frac{\kappa^p}{n}\right)^{1/p}\right).$$

For $\kappa = n^\alpha$ this is $\widetilde{O}(n^{2+\max\{\alpha - 1/p, 0\}})$, which is below $n^{\beta(\alpha)}$ when $p < 1/(2 + \alpha - \beta(\alpha))$. Combining the two sparse comparisons gives the threshold in Lemma 4.7:

$$p < \max\left\{\frac{\beta(\alpha) - 1}{\alpha}, \frac{1}{2 + \alpha - \beta(\alpha)}\right\}.$$

In particular, if the dense optimality hypothesis holds for arbitrarily small positive α, this rules out every fixed finite p in the $T_A = \mathrm{nnz}(A)$ model.

The dense-access model has the same embedding but a different cost accounting. PaddedDenseSolve again pads to $m = \lceil n\kappa^p \rceil$ and projects the solution back, but an assumed dense-access solver pays $N^2\bar{\kappa}_p(A_m)$ for the padded dimension $N = n + m = O(n(1 + \kappa^p))$. The resulting dense-instance running time is

$$\widetilde{O}(n^2(1 + \kappa^p)^2) = \widetilde{O}(n^2\kappa^{2p}),$$

and comparison with the same lower-bound exponent gives the dense-access threshold

$$p < \frac{\beta(\alpha) - 2}{2\alpha}.$$

In the mildly conditioned regime $0 < \alpha < 2\omega - 4$, this is exactly $p < 1/4$.

The final ingredient explains the scope of this dense-access conclusion. The loss comes from squaring the padded dimension, so the proof checks whether a different amount of decoupled padding could avoid it. On spectra with $S_p(B) \geq cn\kappa^p$, the optimized dense-access padded cost satisfies

$$N_m^2 \bar{\kappa}_p(A_m) \geq C_{p,c} n^2 \kappa^{\min\{1,2p\}}$$

for every identity-padding size m. Therefore, in the mildly conditioned regime, this route cannot produce the polynomial speedup needed for a contradiction once $p \geq 1/4$. This is why the theorem states a route-specific limitation for identity padding rather than a lower bound against all possible dense-access reductions.

4 Proof of the Main Theorem

The first lemma is the basic padding calculation. It records both the spectral effect of appending identity coordinates and the preservation of the linear-system solution under projection onto the original coordinates.

Lemma 4.1 (Identity-padding moment bound)**.** *Fix finite $p > 0$. Let $B \in \mathbb{C}^{n\times n}$ be nonsingular with $\sigma_n(B) = 1$, and let $b \in \mathbb{C}^n$. For any integer $m \geq 0$, set $A = A_m(B)$ and $c = c_m(b)$. Then $\sigma_{n+m}(A) = 1$,*

$$\bar{\kappa}_p(A)^p = \frac{S_p(B) + m}{n + m},$$

and any high-accuracy solution y to $Ay = c$ yields, by taking its first n coordinates, a high-accuracy solution to $Bx = b$. Moreover,

$$\mathrm{nnz}(A) = \mathrm{nnz}(B) + m.$$

In particular, if $m \geq S_p(B)$, then

$$\bar{\kappa}_p(A) \leq 2^{1/p} \quad \textit{and} \quad \mathrm{nnz}(A)\bar{\kappa}_p(A) \leq 2^{1/p}\left(\mathrm{nnz}(B) + m\right).$$

Proof. By Fact 2.5, the singular values of A are the singular values of B together with m copies of 1, and the smallest singular value of A is 1. Thus

$$\sum_{j=1}^{n+m} \left(\frac{\sigma_j(A)}{\sigma_{n+m}(A)}\right)^p = \sum_{i=1}^{n} \sigma_i(B)^p + m = S_p(B) + m,$$

and dividing by $n + m$ gives the formula for $\bar{\kappa}_p(A)^p$.

If $m \geq S_p(B)$, then $n + m \geq m$ and

$$\bar{\kappa}_p(A)^p = \frac{S_p(B) + m}{n + m} \leq \frac{2m}{m} = 2.$$

Hence $\bar{\kappa}_p(A) \leq 2^{1/p}$.

The exact solution of $Ay = c$ is $(B^{-1}b, 0)$. If $y = (x, w)$ satisfies Definition 2.2, then

$$\|x - B^{-1}b\|_2 \leq \|y - (B^{-1}b, 0)\|_2 \leq \varepsilon\|(B^{-1}b, 0)\|_2 = \varepsilon\|B^{-1}b\|_2.$$

Therefore x is a high-accuracy solution to $Bx = b$. Finally, identity padding adds exactly m nonzero diagonal entries, so $\mathrm{nnz}(A) = \mathrm{nnz}(B) + m$. Multiplying by the bound on $\bar{\kappa}_p(A)$ gives the last inequality. $\square$

The next reduction pads to the full p-moment upper bound. For a dense input with $\sigma_1(B) \le \kappa$ this makes the averaged condition number constant while adding at most $n\kappa^p$ identity coordinates.

PaddedSparseSolve: Solve a dense system by padding to constant averaged condition number.

Input: A finite $p > 0$; a nonsingular matrix $B \in \mathbb{C}^{n\times n}$ and right-hand side $b \in \mathbb{C}^n$ with $\sigma_n(B) = 1$; a number $\kappa \ge \sigma_1(B)$; and a sparsity-sensitive p-averaged solver $\mathcal{S}_{\mathrm{sp}}$ as in Definition 2.3.
Output: A vector $x \in \mathbb{C}^n$.
Guarantee: The output x is a high-accuracy solution to $Bx = b$.

1. Set $m = \lceil n\kappa^p \rceil$.
2. Let $A = A_m(B)$ and $c = c_m(b)$.
3. Set $y = \mathcal{S}_{\mathrm{sp}}(A, c)$, and write $y = (x, w)$ with $x \in \mathbb{C}^n$.
4. Return x.

Lemma 4.2 (Full-moment sparse speedup)**.** *Fix finite $p > 0$ and assume that a sparsity-sensitive p-averaged solver exists. Let $B \in \mathbb{C}^{n\times n}$ be nonsingular with $\sigma_n(B) = 1$ and $\sigma_1(B) \le \kappa$. Then* PaddedSparseSolve *solves $Bx = b$ to high accuracy in time*

$$\widetilde{O}(n^2 + n\kappa^p).$$

Proof. Since $\sigma_i(B) \le \kappa$ for all i,

$$S_p(B) = \sum_{i=1}^{n} \sigma_i(B)^p \le n\kappa^p.$$

With $m = \lceil n\kappa^p \rceil$, set $A = A_m(B)$ and $c = c_m(b)$. Lemma 4.1 gives $\bar{\kappa}_p(A) \le 2^{1/p}$ and shows that projection of a high-accuracy padded solution solves $Bx = b$ to high accuracy. The padded matrix has at most $n^2 + \lceil n\kappa^p \rceil$ nonzero entries. The assumed sparsity-sensitive solver therefore takes time

$$\widetilde{O}\left(\mathrm{nnz}(A)\bar{\kappa}_p(A)\right) = \widetilde{O}(n^2 + n\kappa^p).$$

$\square$

For a fixed condition-number exponent and larger p, padding all the way to constant averaged condition number is not always the strongest comparison. The second sparse reduction fixes the padding dimension at n^2, keeping the access term at the dense input size and only partially reducing $\bar{\kappa}_p$.

QuadraticallyPaddedSparseSolve: Solve a dense system with exactly n^2 identity padding coordinates.
Input: A finite $p > 0$; a nonsingular matrix $B \in \mathbb{C}^{n\times n}$ and right-hand side $b \in \mathbb{C}^n$ with $\sigma_n(B) = 1$; a number $\kappa \geq \sigma_1(B)$; and a sparsity-sensitive p-averaged solver $\mathcal{S}_{\mathrm{sp}}$ as in Definition 2.3.
Output: A vector $x \in \mathbb{C}^n$.
Guarantee: The output x is a high-accuracy solution to $Bx = b$.

1. Set $m = n^2$.
2. Let $A = A_m(B)$ and $c = c_m(b)$.
3. Set $y = \mathcal{S}_{\mathrm{sp}}(A, c)$, and write $y = (x, w)$ with $x \in \mathbb{C}^n$.
4. Return x.

Lemma 4.3 (Quadratic-padding sparse speedup)**.** *Fix finite $p > 0$ and assume that a sparsity-sensitive p-averaged solver exists. Let $B \in \mathbb{C}^{n\times n}$ be nonsingular with $\sigma_n(B) = 1$ and $\sigma_1(B) \leq \kappa$. Then* QuadraticallyPaddedSparseSolve *solves $Bx = b$ to high accuracy in time*

$$\widetilde{O}\left(n^2\left(1 + \frac{\kappa^p}{n}\right)^{1/p}\right).$$

In particular, if $\kappa = n^\alpha$ for a fixed $\alpha > 0$, the running time is

$$\widetilde{O}\left(n^{2+\max\{\alpha-1/p,0\}}\right).$$

Proof. Choose $m = n^2$ and set $A = A_m(B)$ and $c = c_m(b)$. Since $\sigma_i(B) \leq \kappa$ for every i, Lemma 4.1 gives

$$\bar{\kappa}_p(A)^p = \frac{S_p(B) + n^2}{n + n^2} \leq \frac{n\kappa^p + n^2}{n^2} = 1 + \frac{\kappa^p}{n}.$$

Also $\mathrm{nnz}(A) \leq 2n^2$. Invoking the assumed sparsity-sensitive solver on (A, c) therefore takes time

$$\widetilde{O}\left(\mathrm{nnz}(A)\bar{\kappa}_p(A)\right) = \widetilde{O}\left(n^2\left(1 + \frac{\kappa^p}{n}\right)^{1/p}\right).$$

By Lemma 4.1, the first block of the returned vector solves $Bx = b$ to high accuracy. If $\kappa = n^\alpha$, then $(1 + \kappa^p/n)^{1/p} = O(n^{\max\{\alpha-1/p,0\}})$, giving the exponent form. □

The dense-access reduction uses the same full-moment padding, but the assumed solver pays for N^2 where N is the padded dimension. This square is the source of the weaker dense-access threshold.

PaddedDenseSolve: Solve a dense system using a dense-access averaged-condition solver on the padded instance.
Input: A finite $p > 0$; a nonsingular matrix $B \in \mathbb{C}^{n\times n}$ and right-hand side $b \in \mathbb{C}^n$ with $\sigma_n(B) = 1$; a number $\kappa \ge \sigma_1(B)$; and a dense-access p-averaged solver $\mathcal{S}_{\text{dense}}$ as in Definition 2.3.
Output: A vector $x \in \mathbb{C}^n$.
Guarantee: The output x is a high-accuracy solution to $Bx = b$.

1. Set $m = \lceil n\kappa^p \rceil$.
2. Let $A = A_m(B)$ and $c = c_m(b)$.
3. Set $y = \mathcal{S}_{\text{dense}}(A, c)$, and write $y = (x, w)$ with $x \in \mathbb{C}^n$.
4. Return x.

Lemma 4.4 (Dense-access padding speedup). *Fix finite $p > 0$ and assume that a dense-access p-averaged solver exists. Let $B \in \mathbb{C}^{n\times n}$ be nonsingular with $\sigma_n(B) = 1$ and $\sigma_1(B) \le \kappa$. Then* PaddedDenseSolve *solves $Bx = b$ to high accuracy in time*

$$\widetilde{O}(n^2(1 + \kappa^p)^2).$$

In particular, since $\kappa \ge 1$, the running time is $\widetilde{O}(n^2\kappa^{2p})$.

Proof. The bound $S_p(B) \le n\kappa^p$ implies that $m = \lceil n\kappa^p\rceil$ satisfies the hypothesis of Lemma 4.1. Thus, for $A = A_m(B)$ and $c = c_m(b)$, we have $\bar{\kappa}_p(A) \le 2^{1/p}$, and projecting any high-accuracy solution to $Ay = c$ onto the first n coordinates gives a high-accuracy solution to $Bx = b$.

The padded dimension is $N = n + m = O(n(1 + \kappa^p))$. Invoking the assumed dense-access solver on (A, c) therefore has bit time

$$\widetilde{O}\left(N^2\bar{\kappa}_p(A)\right) = \widetilde{O}(n^2(1 + \kappa^p)^2).$$

Since $\kappa \ge 1$, this is $\widetilde{O}(n^2\kappa^{2p})$. □

The next result explains why the dense-access padding comparison cannot be improved merely by choosing a different number of identity coordinates on spectra whose p-moment is as large as the condition number allows.

Lemma 4.5 (Dense-access padding barrier). *Fix finite $p > 0$ and a constant $c \in (0, 1]$. Let $B \in \mathbb{C}^{n\times n}$ be nonsingular with $\sigma_n(B) = 1$, $\sigma_1(B) \le \kappa$, and*

$$S_p(B) = \sum_{i=1}^{n} \sigma_i(B)^p \ge cn\kappa^p.$$

Assume κ is large enough that $c\kappa^p \ge 1$. For every integer $m \ge 0$, set $A_m = \operatorname{diag}(B, I_m)$ and $N_m = n + m$. Then

$$N_m^2\bar{\kappa}_p(A_m) \ge C_{p,c}n^2\kappa^{\min\{1,2p\}},$$

where $C_{p,c} > 0$ depends only on p and c. Consequently, for $\kappa = n^\alpha$, dense-access identity padding has cost at least $n^{2+\alpha\min\{1,2p\}}$ up to constants depending only on p and c on such spectra.

Proof. Let $s = S_p(B)/n$ and $r = m/n$. By Lemma 4.1,

$$N_m^2 \bar{\kappa}_p(A_m) = n^2(1+r)^{2-1/p}(s+r)^{1/p}.$$

Also $s \geq c\kappa^p$, and the assumption $c\kappa^p \geq 1$ gives $s \geq 1$.

First suppose $p \geq 1/2$. Then $2 - 1/p \geq 0$, and therefore

$$(1+r)^{2-1/p}(s+r)^{1/p} \geq s^{1/p} \geq c^{1/p}\kappa.$$

This is the claimed lower bound because $\min\{1, 2p\} = 1$ in this case.

Now suppose $0 < p < 1/2$. Put $a = 2 - 1/p < 0$ and $b = 1/p$, so $a + b = 2$. If $0 \leq r \leq s$, then $1 + r \leq 1 + s \leq 2s$, and since $a < 0$,

$$(1+r)^a(s+r)^b \geq (2s)^a s^b = 2^a s^2.$$

If $r \geq s$, then $r \geq 1$, so $1 + r \leq 2r$, and again

$$(1+r)^a(s+r)^b \geq (2r)^a r^b = 2^a r^2 \geq 2^a s^2.$$

Thus in all cases

$$(1+r)^{2-1/p}(s+r)^{1/p} \geq 2^{2-1/p}s^2 \geq 2^{2-1/p}c^2\kappa^{2p}.$$

Taking $C_{p,c} = \min\{c^{1/p}, 2^{2-1/p}c^2\}$ proves the lemma. ☐

The preceding barrier is about the identity-padding route. The following claim states the corresponding scope in the mildly conditioned regime.

Claim 4.6 (Dense-padding scope)**.** *Within reductions that only append decoupled padding coordinates with singular values at least $\sigma_n(B)$ and then project the solution back to the original block, the dense-access padding route cannot yield a conditional contradiction for $p \geq 1/4$ in the mildly conditioned regime from condition-number-only hard families with $S_p(B) = \Theta(n\kappa^p)$. This is a route-specific barrier, not an impossibility theorem for all reductions.*

Proof. For any fixed number of appended decoupled coordinates, choosing their singular values equal to $\sigma_n(B) = 1$ minimizes the added contribution to the p-moment while keeping the smallest singular value unchanged. Thus identity padding is the best decoupled padding of this kind for the dense-access quantity $N^2\bar{\kappa}_p$.

Consider a hard family with $\kappa = n^\alpha$ and $S_p(B) = \Theta(n\kappa^p)$. By Lemma 4.5, every identity-padding choice has dense-access cost at least

$$n^{2+\alpha\min\{1,2p\}}$$

up to constants depending on p and polylogarithmic factors. In the mildly conditioned regime, $\beta(\alpha) = 2 + \alpha/2$. If $p \geq 1/4$, then $\min\{1, 2p\} \geq 1/2$, so this lower bound on the padded solver's own cost is at least $n^{\beta(\alpha)}$. Therefore this route cannot produce the polynomial speedup needed to contradict the dense optimality hypothesis.

The obstruction is not exhaustive outside this route. A stronger dense-access hardness result could follow from a dense lower bound for hard families with $S_p(B) = o(n\kappa^p)$, from a solve-preserving transformation that reduces the relevant spectral moment without adding $\Omega(n\kappa^p)$ dimensions, or from a stronger lower-bound hypothesis than the $\widetilde{\Omega}(n^2\kappa^{1/2})$ term. ☐

We now compare the two sparse reductions with the dense optimality hypothesis. The full-moment padding gives the exponent $1 + \alpha p$ in the added spectral term, while the n^2 padding gives exponent $2 + \max\{\alpha - 1/p, 0\}$.

Lemma 4.7 (Sparse-access conditional hardness)**.** *Assume $\omega > 2$. Fix $\alpha > 0$, set*

$$\beta(\alpha) = \min\{\omega, 2 + \alpha/2\},$$

and assume the dense optimality hypothesis at exponent α. Then there is no sparsity-sensitive p-averaged solver for any fixed finite $p > 0$ satisfying

$$p < \max\left\{\frac{\beta(\alpha) - 1}{\alpha}, \frac{1}{2 + \alpha - \beta(\alpha)}\right\}.$$

When $0 < \alpha < 2\omega - 4$, the two thresholds inside the maximum are $1/\alpha + 1/2$ and $2/\alpha$, so the larger threshold is $2/\alpha$ whenever also $\alpha < 2$. When $\alpha \geq 2\omega - 4$, the thresholds are $(\omega - 1)/\alpha$ and $1/(2 + \alpha - \omega)$. The boundary case at the larger threshold is not claimed. Consequently, if the dense optimality hypothesis holds for arbitrarily small positive values of α, then no such solver exists for any fixed finite $p > 0$.

Proof. Fix finite $p > 0$ below the displayed maximum, and suppose for contradiction that a sparsity-sensitive p-averaged solver exists.

First suppose

$$p < \frac{\beta(\alpha) - 1}{\alpha}.$$

Apply Lemma 4.2 with $\kappa = n^\alpha$. The derived high-accuracy solver for dense systems with $\sigma_n(B) = 1$ and $\sigma_1(B) = n^\alpha$ runs in time

$$\widetilde{O}(n^2 + n(n^\alpha)^p) = \widetilde{O}(n^2 + n^{1+\alpha p}).$$

By the strict inequality on p, there is $\delta > 0$ such that $1 + \alpha p \leq \beta(\alpha) - \delta$. Also $\beta(\alpha) > 2$, because $\omega > 2$ and $\alpha > 0$. Hence, for $\eta = \min\{\beta(\alpha) - 2, \delta\} > 0$, the derived running time is $\widetilde{O}(n^{\beta(\alpha)-\eta})$. This contradicts the dense optimality lower bound $\widetilde{\Omega}(n^{\beta(\alpha)})$ on the same hard family.

It remains to handle the case where the second threshold is witnessed:

$$p < \frac{1}{2 + \alpha - \beta(\alpha)}.$$

The denominator is positive because $\beta(\alpha) \leq 2 + \alpha/2 < 2 + \alpha$. Apply Lemma 4.3 with $\kappa = n^\alpha$. The derived solver runs in time

$$\widetilde{O}\left(n^{2+\max\{\alpha-1/p,0\}}\right).$$

The strict inequality on p gives $1/p > 2 + \alpha - \beta(\alpha)$, and therefore

$$\alpha - \frac{1}{p} < \beta(\alpha) - 2.$$

If $\alpha - 1/p \leq 0$, the derived exponent is $2 < \beta(\alpha)$. If $\alpha - 1/p > 0$, the derived exponent is $2 + \alpha - 1/p < \beta(\alpha)$. In both subcases the derived running time is $\widetilde{O}(n^{\beta(\alpha)-\eta})$ for some $\eta > 0$, again contradicting the dense optimality hypothesis.

Since p is below the maximum of the two thresholds, it satisfies at least one of the two strict inequalities just treated. The displayed specializations follow by evaluating $\beta(\alpha)$. If $0 < \alpha < 2\omega - 4$, then $\beta(\alpha) = 2 + \alpha/2$, so the two thresholds are $1/\alpha + 1/2$ and $2/\alpha$. If $\alpha \geq 2\omega - 4$, then $\beta(\alpha) = \omega$, so the two thresholds are $(\omega - 1)/\alpha$ and $1/(2 + \alpha - \omega)$.

For the final consequence, fix any finite $p > 0$. If the dense optimality hypothesis holds for arbitrarily small positive α, choose such an α with $\alpha < \min\{1/p, 2\omega - 4\}$. Then $p < 1/\alpha < 1/\alpha + 1/2$, so the preceding argument rules out the chosen p. $\square$

The dense-access comparison is analogous, but the padded dimension contributes the exponent $2 + 2\alpha p$.

Lemma 4.8 (Dense-access conditional hardness). *Assume $\omega > 2$. Fix $\alpha > 0$, set*

$$\beta(\alpha) = \min\{\omega, 2 + \alpha/2\},$$

and assume the dense optimality hypothesis at exponent α. Then there is no dense-access p-averaged solver for any fixed finite $p > 0$ satisfying

$$p < \frac{\beta(\alpha) - 2}{2\alpha}.$$

Equivalently, this rules out every fixed $p < 1/4$ when $0 < \alpha < 2\omega - 4$, and every fixed $p < (\omega - 2)/(2\alpha)$ when $\alpha \geq 2\omega - 4$. The boundary case $p = (\beta(\alpha) - 2)/(2\alpha)$ is not claimed. Consequently, if the dense optimality hypothesis holds for arbitrarily small positive values of α, then no dense-access solver exists for any fixed $p < 1/4$.

Proof. Fix finite $p > 0$ with $p < (\beta(\alpha) - 2)/(2\alpha)$. Suppose for contradiction that a dense-access p-averaged solver exists. Apply Lemma 4.4 with $\kappa = n^\alpha$. The derived high-accuracy solver for dense systems with $\sigma_n(B) = 1$ and $\sigma_1(B) = n^\alpha$ runs in time

$$\widetilde{O}(n^2 (n^\alpha)^{2p}) = \widetilde{O}(n^{2 + 2\alpha p}).$$

By the strict inequality on p, there is $\delta > 0$ such that $2 + 2\alpha p \leq \beta(\alpha) - \delta$. Therefore the derived running time is $\widetilde{O}(n^{\beta(\alpha) - \delta})$, contradicting the dense optimality lower bound $\widetilde{\Omega}(n^{\beta(\alpha)})$ on the same hard family.

The displayed equivalences follow by evaluating $\beta(\alpha)$. If $0 < \alpha < 2\omega - 4$, then $\beta(\alpha) = 2 + \alpha/2$, and hence $(\beta(\alpha) - 2)/(2\alpha) = 1/4$. If $\alpha \geq 2\omega - 4$, then $\beta(\alpha) = \omega$, and the threshold is $(\omega - 2)/(2\alpha)$.

For the final consequence, fix any finite $p < 1/4$. If the dense optimality hypothesis holds for arbitrarily small positive values of α, choose such an α with $\alpha < 2\omega - 4$. The first case then rules out this p. $\square$

Proof of Theorem 1.1. The theorem is the combination of Lemmas 4.7 and 4.8. The additional statement about arbitrarily small positive α is exactly the final consequence in Lemma 4.7 for $T_A = \mathrm{nnz}(A)$ and in Lemma 4.8 for $T_A = N^2$. $\square$

References

[ABB+26] Noah Amsel, Yves Baumann, Paul Beckman, Peter Bürgisser, Chris Camaño, Tyler Chen, Edmond Chow, Anil Damle, Michał Dereziński, Mark Embree, Ethan N. Epperly, Robert Falgout, Mark Fornace, Anne Greenbaum, Chen Greif, Diana Halikias, Zhen Huang, Elias Jarlebring, Yiannis Koutis, Daniel Kressner, Rasmus Kyng, Jörg Liesen, Jackie Lok, Raphael A. Meyer, Yuji Nakatsukasa, Kate Pearce, Richard Peng, David Persson, Eliza Rebrova, Ryan Schneider, Rikhav Shah, Edgar Solomonik, Nikhil Srivastava, Alex Townsend, Robert J. Webber, and Jess Williams. Linear systems and eigenvalue problems: Open questions from a simons workshop, 2026. i, 1, 3

[ADW+25] Josh Alman, Ran Duan, Virginia Vassilevska Williams, Yinzhan Xu, Zixuan Xu, and Renfei Zhou. More asymmetry yields faster matrix multiplication. In *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2025*, pages 2005–2039. SIAM, 2025. 1

[AKK+20] Naman Agarwal, Sham Kakade, Rahul Kidambi, Yin-Tat Lee, Praneeth Netrapalli, and Aaron Sidford. Leverage score sampling for faster accelerated regression and erm. In *Proceedings of the 31st International Conference on Algorithmic Learning Theory*, volume 117 of *Proceedings of Machine Learning Research*, pages 22–47. PMLR, 2020. 1

[BH74] James R. Bunch and John E. Hopcroft. Triangular factorization and inversion by fast matrix multiplication. *Mathematics of Computation*, 28(125):231–236, 1974. 1

[CW17] Kenneth L. Clarkson and David P. Woodruff. Low-rank approximation and regression in input sparsity time. *Journal of the ACM*, 63(6):54:1–54:45, 2017. 1

[DLNR25] Michał Dereziński, Daniel LeJeune, Deanna Needell, and Elizaveta Rebrova. Fine-grained analysis and faster algorithms for iteratively solving linear systems. *Journal of Machine Learning Research*, 26(144):1–49, 2025. 1

[DMY25] Michał Dereziński, Christopher Musco, and Jiaming Yang. Faster linear systems and matrix norm approximation via multi-level sketched preconditioning. In *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2025*, pages 1972–2004. SIAM, 2025. 1

[DNRY25] Michał Dereziński, Deanna Needell, Elizaveta Rebrova, and Jiaming Yang. Randomized Kaczmarz methods with beyond-Krylov convergence, 2025. 1

[DS26] Michał Dereziński and Aaron Sidford. Approaching optimality for solving dense linear systems with low-rank structure. In *Proceedings of the 2026 Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2026*, pages 925–938. SIAM, 2026. 1

[DY24] Michał Dereziński and Jiaming Yang. Solving dense linear systems faster than via preconditioning. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024*, pages 1118–1129. ACM, 2024. 1

[GL13] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, fourth edition, 2013. 1

[GR15] Robert M. Gower and Peter Richtárik. Randomized iterative methods for linear systems. *SIAM Journal on Matrix Analysis and Applications*, 36(4):1660–1690, 2015. 1

[HMT11] Nathan Halko, Per-Gunnar Martinsson, and Joel A. Tropp. Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Review*, 53(2):217–288, 2011. 1

[HS52] Magnus R. Hestenes and Eduard Stiefel. Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, 49(6):409–436, 1952. 1

[LL10] Dennis Leventhal and Adrian S. Lewis. Randomized methods for linear constraints: Convergence rates and conditioning. *Mathematics of Operations Research*, 35(3), 2010. 1

[LS13a] Yin Tat Lee and Aaron Sidford. Efficient accelerated coordinate descent methods and faster algorithms for solving linear systems. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science, FOCS 2013*, pages 147–156. IEEE Computer Society, 2013. 1

[LS13b] Jörg Liesen and Zdeněk Strakoš. *Krylov Subspace Methods: Principles and Analysis*. Oxford University Press, 2013. 1

[MT20] Per-Gunnar Martinsson and Joel A. Tropp. Randomized numerical linear algebra: Foundations and algorithms. *Acta Numerica*, 29:403–572, 2020. 1

[PS82] Christopher C. Paige and Michael A. Saunders. LSQR: An algorithm for sparse linear equations and sparse least squares. *ACM Transactions on Mathematical Software*, 8(1):43–71, 1982. 1

[RT08] Vladimir Rokhlin and Mark Tygert. A fast randomized algorithm for overdetermined linear least-squares regression. *Proceedings of the National Academy of Sciences*, 105(36):13212–13217, 2008. 1

[Sar06] Tamás Sarlós. Improved approximation algorithms for large matrices via random projections. In *47th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2006*, pages 143–152. IEEE Computer Society, 2006. 1

[SS86] Youcef Saad and Martin H. Schultz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing*, 7(3):856–869, 1986. 1

[Str69] Volker Strassen. Gaussian elimination is not optimal. *Numerische Mathematik*, 13(4):354–356, 1969. 1

[SV09] Thomas Strohmer and Roman Vershynin. A randomized Kaczmarz algorithm with exponential convergence. *Journal of Fourier Analysis and Applications*, 15(2):262–278, 2009. 1