Average-case rigidity lower bounds[*]

Xuangui Huang[1*†] and Emanuele Viola[1*†]

[1*]Khoury College of Computer Sciences, Northeastern University,
440 Huntington Ave, Boston, 02115, MA, USA.

[*]Corresponding author(s). E-mail(s): stslxg@ccs.neu.edu; viola@ccs.neu.edu;
[†]These authors contributed equally to this work.

Abstract

It is shown that there exists $\boldsymbol{f\colon \{0,1\}^{n/2} \times \{0,1\}^{n/2} \to \{0,1\}}$ in $\mathrm{E}^{\mathbf{NP}}$ such that for every $\boldsymbol{2^{n/2} \times 2^{n/2}}$ matrix $\boldsymbol{M}$ of rank $\boldsymbol{\le \rho}$ we have $\mathbb{P}_{\boldsymbol{x},\boldsymbol{y}}[\boldsymbol{f(x,y) \neq M_{x,y}}] \boldsymbol{\ge 1/2 - 2^{-\Omega(k)}}$, whenever $\boldsymbol{\log \rho \le \delta n/k(\log n + k)}$ for a sufficiently small $\boldsymbol{\delta > 0}$, and $\boldsymbol{n}$ is large enough. This generalizes recent results which bound below the probability by $\boldsymbol{1/2 - \Omega(1)}$ or apply to constant-depth circuits.

Keywords: average-case lower bounds, matrix rigidity, correlation bounds

Starting with the seminal paper by Williams [1] a sequence of recent works have proved new lower bounds for functions in various classes which contain super-polynomial non-deterministic time [2–17], lower bounds that we do not know how to prove by other means. Two sub-sequences of results are relevant to the present work. The first is the sub-sequence establishing *average-case hardness results* for various circuit classes. The concurrent works [14, 15] proved incomparable, new average-case lower bounds against AC^0 with parity gates. Both results were improved in [16] to obtain a function that any such circuit of sub-exponential size cannot compute with a sub-exponentially small advantage over random guessing, for a uniform input.

[*]Supported by NSF awards CCF-1813930 and CCF-2114116. The authors thank the anonymous reviewers for valuable feedback, especially for pointing out a technical issue in the proof of the full version of Theorem 1 that works for all sufficiently large n in the manuscript (See Remark 1 for details).

The second is the sub-sequence constructing *rigid matrices* [18], that is, obtaining functions $f(x, y)$, where $|x| = |y| = n/2$ such that the corresponding $2^{n/2} \times 2^{n/2}$ matrix $M_{x,y} = f(x, y)$ is far from low-rank matrices in the Hamming distance. Using Probabilistically Checkable Proofs (PCP), [10] gave f such that $\mathbb{P}[M_{x,y} \neq f(x, y)] \geq \Omega(1)$ for any M of rank up to at most $2^{n^{1/4-\varepsilon}}$. Low-rank matrices are a generalization of low-degree polynomials [19], as the truth table of a degree-d polynomial over n input bits has rank at most $\binom{n}{d}$ when viewed as a matrix. Using this connection, rigidity over $\mathbb{F}_2$ can also be seen as a generalization of average-case lower bounds against AC^0 with parity gates, since such circuits of sub-exponential size can be approximated by $\mathbb{F}_2$-polynomials with sub-linear degree and sub-exponentially small error. However, the rank bound in [10] is not strong enough to improve the classic results on polynomials due to Razborov and Smolensky [20–22] which hold up to degree $\sqrt{n}$. The subsequent paper [15] achieved nearly-optimal probabilistic degree $n/\mathrm{poly}\log n$ relying on the PCP construction [23]. It also raised the question of constructing PCPs with stronger properties and showed that these would improve the rank bounds in [10] to $2^{n/\Omega(\log^2 n)}$ (under some distribution). Related PCPs were constructed in the subsequent work [17], finally obtaining f such that $\mathbb{P}[M_{x,y} \neq f(x, y)] \geq \Omega(1)$ for any M of rank up to $2^{n/\Omega(\log n)}$.

In this paper we prove a result that generalizes both sub-sequences. We simultaneously achieve the strong average-case hardness parameters of [16] and work in the general model of low-rank matrices.

Theorem 1 *There exists a family of functions $f_n \colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ in $\mathrm{E}^{\mathbf{NP}}$ such that for any rank-ρ matrix $M \in \mathbb{F}_2^{2^n \times 2^n}$, we have*

$$\mathbb{P}_{x,y}[f_n(x, y) \neq M_{x,y}] \geq 1/2 - 2^{-\Omega(k)}$$

for all large enough n, where $k(\log n + k)\log \rho \leq \delta n$ for a sufficiently small constant $\delta > 0$.

To illustrate the parameters, we can prove lower bounds whenever $k^2 \log \rho \leq \delta n$, for $k \geq \log n$. In particular we can for example bound below the probability by $1/2 - 2^{-n^{\Omega(1)}}$ for log rank $n^{0.99}$, i.e. when $\log \rho = n^{0.99}$. We can also have $\log \rho = n/\Omega(\log n)$ whenever $k = O(1)$, recovering the result from [17].

It seems within reach to improve the tradeoff between k and ρ to obtain lower bounds whenever $k \log \rho \leq \delta n$. Improving the tradeoff even further to obtain lower bounds when $k \log \rho$ is $n^{1+\Omega(1)}$ would give new *data-structure lower bounds*, for functions in $\mathrm{E}^{\mathbf{NP}}$, via a connection established in [15].

Independently, Chen and Lyu [24] proved lower bounds whenever $k^{1.5} \log \rho \leq \delta n$. Their proof proceeds in exactly the same way as ours, but in addition they prove a new derandomized XOR lemma where the seed length is just $\sqrt{k}n$ as opposed to kn in our Lemma 12. One can also plug their new XOR lemma in our proof and infer the stronger bound.

Techniques

Our proof builds on the previous works mentioned earlier. We adapt a clever approach in [16] which is based on Levin's proof of Yao's famous XOR lemma, cf. [25]. The approach shows that to prove a strong average-case hardness result it suffices to prove a mild average-case hardness result *for an intermediate model.* The intermediate model in our case consists of *rational sums of low-rank matrices.* We show that a lower bound for this model can be obtained from the rectangular PCP in [17], see Theorem 10.

A little more in detail, we prove a constant-error lower bound for rational sums of low-rank matrices by contradiction using the non-deterministic time-hierarchy theorem following [26]. We fix a unary language in $\mathbf{NTIME}(2^n) \setminus \mathbf{NTIME}(o(2^n))$, and let the lexicographically first rectangular PCP proof for this language be the hard function. Assuming that this hard function has constant correlation with a sum of low-rank matrices, we derive a contradiction by giving a quick non-deterministic algorithm. This algorithm first guesses a sum of low-rank matrices as an approximation of the hard function, i.e. the boolean proof, then performs a series of validity tests that are adapted from [16] to guarantee that this sum is bounded and close to boolean. Then the rectangular property of the PCP is exploited to make sure that when the guessed sum is plugged in as a proof, the bits that the PCP verifier probes can also be written as sums of low-rank matrices, thus the algorithm can quickly evaluate the "acceptance probability" of the guessed sum, based on the fast counting algorithm for low-rank matrices in [10, 27]. Now the boundedness and close-to-boolean properties will ensure that this "acceptance probability" is close to that of the boolean proof approximated by the guessed sum, so the algorithm can make a decision for the language based on this value.

We shall first prove our result for infinitely many input lengths n; at the end we shall explain what modifications are sufficient to obtain all sufficiently large n, using results in [16].

1 Preliminaries

For any $n \in \mathbb{N}$, define $[n] = \{1, 2, \ldots, n\}$. For any matrix M, we use $M_{i,j}$ to denote its entry on row i column j. For any real matrix M, we define its ℓ_p-norm as $\|M\|_p = (\mathbb{E}_{i,j}[|M_{i,j}|^p])^{1/p}$, while the ℓ_∞-norm is defined as $\|M\|_\infty = \max_{i,j} |M_{i,j}|$. For any two matrices A and B with the same shape, we define $A \circ B$ to be the Hadamard product (entrywise product) of them over $\mathbb{R}$, which is distributive. We use $\widetilde{O}$ to hide poly(n) terms in runtime.

Fourier Basis

For convenience we will mainly work on the Fourier basis $\{-1, 1\}$ instead of the boolean basis $\{0, 1\}$. This includes the *outputs* of most of our functions and the matrices we will be working on. To convert an $\mathbb{F}_2$ matrix into a $\{-1, 1\}$ matrix, we use the following notation: for any $n \times m$ $\mathbb{F}_2$-matrix M, we define the matrix $(-1)^M$ by $\left((-1)^M\right)_{i,j} = (-1)^{M_{i,j}}$ for all $i \in [n]$, $j \in [m]$.

We will also be working on rational sums of these functions and matrices, i.e. sums of -1's and 1's, so we use the following notion of boundedness.

Definition 1 We say a real matrix M is *bounded* if $M_{i,j} \in [-1, 1]$ for all i, j. Similarly, we say a real-valued function f is *bounded* if $f(x) \in [-1, 1]$ for all x.

For functions using the Fourier basis $\{-1, 1\}$ and more generally the range $[-1, 1]$, we have the following natural definition of correlations.

Definition 2 Let $f, g \colon \{0,1\}^n \to [-1, 1]$ be two functions. We define their *correlation* as $\mathsf{corr}(f, g) = \left|\mathbb{E}_{x\sim\{0,1\}^n}[f(x)g(x)]\right|$. We say f *ε-correlates with* g iff $\mathsf{corr}(f, g) \geq \varepsilon$.

For $f, g \colon \{0,1\}^n \to \{-1, 1\}$, we have $\mathsf{corr}(f, g) = |\Pr_x[f(x) = g(x)] - \Pr_x[f(x) \neq g(x)]|$. So if we have $\mathsf{corr}(f, g) \leq \varepsilon$, then we know $\Pr_x[f(x) \neq g(x)] \geq 1/2 - \varepsilon/2$.

In some parts we will also work on functions whose *input* basis are Fourier. For any such function $f \colon \{-1, 1\}^k \to \mathbb{R}$, we identify f with its *multilinear extension over domain* $\mathbb{R}$, defined by its Fourier expansion $f = \sum_{S\subseteq[k]} \beta_S \prod_{i\in S} x_i$, where $\beta_S \in \mathbb{R}$.

Rational sums

Now we define the intermediate model, rational sums of low-rank matrices. We need the following technical definition of bit-complexity first.

Definition 3 For any $\alpha \in \mathbb{Q}$ we define its *bit-complexity* as the maximum of the bit lengths of the denominator and numerator. For a polynomial p with rational coefficients we define its bit complexity as the maximum bit complexity among the coefficients.

Definition 4 For any given function class $\mathcal{C}$, we call the sum $\widetilde{Q} = C\sum_{i=1}^{m} b_i \cdot f_i$ *an m-sum of* $\mathcal{C}$, for $b_i \in \{-1, 1\}$ and $f_i \in \mathcal{C}$ for all $i \in [m]$ and $C \in \mathbb{Q}$. We define the *bit-complexity of* $\widetilde{Q}$ as the bit-complexity of C.

In particular, consider $\mathcal{C}$ to be the class of rank-ρ $\mathbb{F}_2$-matrices (converted to $\{-1, 1\}$-matrices), then an *m-sum of rank-ρ $\mathbb{F}_2$-matrices* $\widetilde{Q} \in \mathbb{R}^{n\times n'}$ is given by $\widetilde{Q} = C\sum_{i=1}^{m} b_i \cdot (-1)^{M^{(i)}}$ where $M^{(i)} \in \mathbb{F}_2^{n\times n'}$ are rank-ρ matrices over $\mathbb{F}_2$, i.e. $M^{(i)} = A^{(i)}B^{(i)}$ for some $A^{(i)} \in \mathbb{F}_2^{n\times\rho}$ and $B^{(i)} \in \mathbb{F}_2^{\rho\times n'}$.

To avoid confusions between functions on numbers and functions on matrices, we use the following bar notation to lift a function onto matrices entry-wisely.

Definition 5 For any sets X, Y and function $f\colon X^k \to Y$ we define its *extension over matrices* $\overline{f}\colon \left(X^{n\times m}\right)^k \to Y^{n\times m}$ that maps matrices $M^{(1)}, M^{(2)}, \ldots, M^{(k)} \in X^{n\times m}$ to a matrix $M' \in Y^{n\times m}$ defined by $M'_{i,j} = f(M^{(1)}_{i,j}, M^{(2)}_{i,j}, \ldots, M^{(k)}_{i,j})$ for all $i \in [n], j \in [m]$.

For example, the Hadamard product $A \circ B$ of matrices A and B is recovered as $\overline{f}(A, B)$ where f is multiplication.

Rectangular PCP

We need the following rectangular PCP to prove our main theorem.

Definition 6 (Rectangular PCP, [17]) For any language L, we say it has *an* $(\ell^2, r, q, p, t, s, \tau)$*-rectangular PCP verifier* V *over alphabet* $\{-1, 1\}$ if we have the following properties:

Proof. the proof π of length ℓ^2 is viewed as a matrix in $\{-1,1\}^{\ell\times\ell}$.
Randomness. the random string $R \in \{0,1\}^r$ is partitioned into three parts

$$R = (R_{\mathsf{row}}, R_{\mathsf{col}}, R_{\mathsf{shared}}) \in \{0,1\}^{r_{\mathsf{rect}}} \times \{0,1\}^{r_{\mathsf{rect}}} \times \{0,1\}^{r_{\mathsf{shared}}},$$

where $r_{\mathsf{rect}} = (1-\tau)r/2$ and $r_{\mathsf{shared}} = \tau r$.
Computation. Given input x and proof oracle $\pi \in \{-1,1\}^{\ell\times\ell}$, with randomness R, $V^{\pi}(x;\ R)$ runs as follows:

1. Use shared randomness $R_{\mathsf{shared}} \in \{0,1\}^{r_{\mathsf{shared}}}$ to:
   (a) construct a decision function $D = D(x;\ R_{\mathsf{shared}})\colon \{-1,1\}^q \times \{-1,1\}^p \to \{0,1\}$,
   (b) construct randomness parity check $(C_1, \ldots, C_p) = (C_1(x;\ R_{\mathsf{shared}}), \ldots, C_p(x;\ R_{\mathsf{shared}}))$ where each $C_i\colon \{0,1\}^{r_{\mathsf{rect}}} \times \{0,1\}^{r_{\mathsf{rect}}} \to \{-1,1\}$ is a parity function, i.e. $C_i(R_{\mathsf{row}}, R_{\mathsf{col}}) = (-1)^{\langle R_{\mathsf{row}}, u\rangle + \langle R_{\mathsf{col}}, v\rangle + b}$ for some $u, v \in \{0,1\}^{r_{\mathsf{rect}}}$ and $b \in \{0,1\}$, where $\langle x, y\rangle$ is the inner product of x and y.
2. Use row randomness $R_{\mathsf{row}} \in \{0,1\}^{r_{\mathsf{rect}}}$ to construct row locations of queries

$$i^{(1)} = i^{(1)}(x;\ R_{\mathsf{row}}, R_{\mathsf{shared}}), \ldots, i^{(q)} = i^{(q)}(x;\ R_{\mathsf{row}}, R_{\mathsf{shared}}).$$

3. Use column randomness $R_{\mathsf{col}} \in \{0,1\}^{r_{\mathsf{rect}}}$ to construct column locations of queries

$$j^{(1)} = j^{(1)}(x;\ R_{\mathsf{col}}, R_{\mathsf{shared}}), \ldots, j^{(q)} = j^{(q)}(x;\ R_{\mathsf{col}}, R_{\mathsf{shared}}).$$

4. Output the result

$$D(\pi_{i^{(1)},j^{(1)}}, \ldots, \pi_{i^{(q)},j^{(q)}}, C_1(R_{\mathsf{row}}, R_{\mathsf{col}}), \ldots, C_p(R_{\mathsf{row}}, R_{\mathsf{col}})).$$

Completeness. If $x \in L$ then $\exists \pi \in \{-1,1\}^{\ell\times\ell}, \Pr_R[V^{\pi}(x;\ R) = 1] = 1$.

Soundness. If $x \notin L$ then $\forall \pi \in \{-1,1\}^{\ell \times \ell}, \Pr_R[V^\pi(x;\ R) = 1] < s$.
Complexity The verifier V runs in time $t \geq r$, the query complexity is q and parity-check complexity is p.

Definition 7 We say an $(\ell^2, r, q, p, t, s, \tau)$-rectangular PCP verifier is *smooth* if V queries uniformly on π over the choice of randomness $R \in \{0,1\}^r$ and queries $k \in [q]$.

The above definition means that each location of the proof has equal probability of being queried by a *random* query. A stronger requirement would be that this holds for *each* query. The stronger notion is available in some PCPs (e.g. [28]), but as far as we know not for rectangular PCPs.

Lemma 2 ([17]) *For any constants $s \in (0, \frac{1}{2})$, $\tau \in (0,1)$, and language $L \in$ **NTIME**(2^n), L has a smooth $(\ell^2, r, q, p, t, s, \tau)$-rectangular PCP verifier V over alphabet $\{-1,1\}$ with the following parameters:*

- $r = n + O(\log n)$.
- $q, p = O_s(1)$.
- $\ell^2 = O_s(2^r)$.
- $t = 2^{O(\tau n)}$.

2 Fast Algorithm for "Acceptance Probability"

In this section we prove and collect several facts that allow us to quickly compute the acceptance probability of a rectangular PCP verifier when its proof is a rational sum of low-rank matrices.

First we need the following result to quickly calculate the number of 1's in low-rank matrices over $\mathbb{F}_2$ given low-rank decompositions.

Lemma 3 ([10, 27]) *Given two matrices $A \in \mathbb{F}_2^{N \times \rho}$ and $B \in \mathbb{F}_2^{\rho \times N}$ where $\rho = N^{o(1)}$, there is a deterministic algorithm that computes the number of 1's in the product matrix AB over $\mathbb{F}_2$ in time $T(N, \rho) = N^{2-\Omega(1/\log \rho)}$.*

We prove a general result on evaluating the expectation of a polynomial on sums of low-rank matrices.

Theorem 4 *Let $\left\{\widetilde{Q}_i\right\}_{i \in [k]}$ be k m-sums of rank-ρ $\mathbb{F}_2$ matrices with bit-complexity c, and let their low-rank decompositions be $\widetilde{Q}_i = C_i \sum_{j=1}^m b_{i,j} \cdot (-1)^{A^{(i,j)} B^{(i,j)}}$ where $C_i \in \mathbb{Q}$ has bit-complexity c, $b_{i,j} \in \{-1,1\}$, $A^{(i,j)} \in \mathbb{F}_2^{N \times \rho}$, and $B^{(i,j)} \in \mathbb{F}_2^{\rho \times N}$ for all $i \in [k]$ and $j \in [m]$. For any degree-d polynomial on k variables $p \colon \mathbb{R}^k \to \mathbb{R}$ that has bit complexity c' and s monomials, given the decompositions we can compute the value of $\mathbb{E}_{i,j \in [N]}\left[\left(\overline{p}\left(\widetilde{Q}_1, \ldots, \widetilde{Q}_k\right)\right)_{i,j}\right]$ in time $O\left(sm^d (T(N, d\rho) + \mathrm{poly}(c, c', d, \log N))\right)$*

if $d\rho = N^{o(1)}$. *In particular for any boolean function* $f\colon \{-1,1\}^k \to \{0,1\}$, *the value of* $\mathbb{E}_{i,j\in[N]}\left[\left(\overline{f}\left(\widetilde{Q}_1,\ldots,\widetilde{Q}_k\right)\right)_{i,j}\right]$ *can be computed in time* $O(2^k m^k(T(N,k\rho) + \mathrm{poly}(c,k,\log N)))$ *if* $k\rho = N^{o(1)}$.

Proof To calculate $\mathbb{E}_{i,j}\left[\left(\overline{p}\left(\widetilde{Q}_1,\ldots,\widetilde{Q}_k\right)\right)_{i,j}\right]$, by linearity of expectation it suffices to calculate the expectation for each monomial of p. Wlog, let the monomial $p'(x) = x_1x_2\cdots x_d$. Then by the distributive property of Hadamard products we have

$$\begin{aligned}\overline{p'}\left(\widetilde{Q}_1,\ldots,\widetilde{Q}_k\right) &= \widetilde{Q}_1 \circ \widetilde{Q}_2 \circ \cdots \circ \widetilde{Q}_d \\ &= \circ_{i=1}^{d}\left(C_i\sum_{j=1}^{m} b_{i,j}\cdot(-1)^{A^{(i,j)}B^{(i,j)}}\right) \\ &= \sum_{(j_1,j_2,\ldots,j_d)\in[m]^d}\left(\prod_{i=1}^{d} C_i b_{i,j_i}\right)\cdot\left(\circ_{i=1}^{d}(-1)^{A^{(i,j_i)}B^{(i,j_i)}}\right) \\ &= \sum_{(j_1,j_2,\ldots,j_d)\in[m]^d}\left(\prod_{i=1}^{d} C_i b_{i,j_i}\right)\cdot\left((-1)^{\oplus_{i=1}^{d}A^{(i,j_i)}B^{(i,j_i)}}\right),\end{aligned}$$

where '$\oplus$' is the addition of $\mathbb{F}_2$-matrices over $\mathbb{F}_2$. Hence by linearity of expectation, it suffices to calculate the expectation of $(-1)^{\oplus_{i=1}^{d}A^{(i,j_i)}B^{(i,j_i)}}$ for each $(j_1,\ldots,j_d)\in[m]^d$. Note that for any $\mathbb{F}_2$-matrix M we have $\mathbb{E}_{\mathsf{row},\mathsf{col}}\left[\left((-1)^M\right)_{\mathsf{row},\mathsf{col}}\right] = 1 - 2\mathbb{E}_{\mathsf{row},\mathsf{col}}[M_{\mathsf{row},\mathsf{col}}]$, thus it suffices to calculate

$$\mathbb{E}_{\mathsf{row},\mathsf{col}}\left[\left(\oplus_{i=1}^{d}A^{(i,j_i)}B^{(i,j_i)}\right)_{\mathsf{row},\mathsf{col}}\right] = \frac{1}{N^2}\cdot\ \text{number of 1's in } \oplus_{i=1}^{d}A^{(i,j_i)}B^{(i,j_i)}.$$

Note that $\oplus_{i=1}^{d}A^{(i,j_i)}B^{(i,j_i)}$ is just the product of an $N\times d\rho$ matrix and a $d\rho\times N$ matrix over $\mathbb{F}_2$, where the first matrix is obtained by concatenating the rows of $\left\{A^{(i,j_i)}\right\}_{i\in[d]}$ and the second matrix is obtained by concatenating the columns of $\left\{B^{(i,j_i)}\right\}_{i\in[d]}$. Hence by Lemma 3 the counting can be done in time $T(N,d\rho)$ if $d\rho = N^{o(1)}$. This expectation value has bit-complexity $O(\log N)$, so multiplying it by $\prod_{i=1}^{d} C_i b_{i,j_i}$ and adding to the running sum take time $\mathrm{poly}(c,d,\log N)$. We still need to multiply the result by the coefficients of the monomials in p, thus the runtime becomes $\mathrm{poly}(c,c',d,\log N)$. Therefore the total running time is $O\left(sm^d(T(N,d\rho)+\mathrm{poly}(c,c',d,\log N))\right)$.

Any boolean function f on k input bits can be written as a degree-k multilinear polynomial so there are at most 2^k monomials. Fourier analysis shows that every coefficient of this polynomial is a multiple of 2^{-k}, so its bit complexity is $O(k)$. Therefore the total running time becomes $O(2^k m^k(T(N,k\rho)+\mathrm{poly}(c,k,\log N)))$ if $k\rho = N^{o(1)}$. □

The following lemma from [29] shows that randomness parity checks can be written as low-rank matrices. For completeness we include the very short proof here.

Lemma 5 ([29, Claim B.1]) *For any parity function* $f\colon \{0,1\}^m \times \{0,1\}^m \to \{-1,1\}$ *defined by* $f(i,j) = (-1)^{\langle i,u\rangle + \langle j,v\rangle + b}$ *for some* $u, v \in \{0,1\}^m$ *and* $b \in \{0,1\}$*, we can compute in time* $O(m2^m)$ *two matrices* $A \in \mathbb{F}_2^{2^m \times 3}$ *and* $B \in \mathbb{F}_2^{3 \times 2^m}$ *such that* $f(i,j) = \left((-1)^{AB}\right)_{i,j}$ *for all* $i, j \in \{0,1\}^m$*.*

Proof The first column of A is $\langle i, u\rangle$, row-indexed by $i \in \{0,1\}^m$. The second column of A is all 1, while the third column of A is all b. The second row of B is $\langle j, u\rangle$, column-indexed by $j \in \{0,1\}^m$, while every other entry in B is 1. □

We use the following lemma to quickly calculate the "acceptance probability" of a sum of low-rank matrices $\widetilde{\pi}$.

Lemma 6 *Let* V *be any* $(\ell^2, r, q, p, t, s, \tau)$*-rectangular PCP verifier over* $\{-1,1\}$*, and* $\widetilde{V}$ *be the same as* V *but with* D *multilinearly extended over* $\mathbb{R}$*. Given* $\widetilde{\pi} = C\sum_{i=1}^{m} b_i \cdot (-1)^{A^{(i)}B^{(i)}}$ *with* $C \in \mathbb{Q}$ *of bit-complexity* $O(n)$*,* $b_i \in \{-1,1\}$*,* $A^{(i)} \in \mathbb{F}_2^{\ell \times \rho}$*, and* $B^{(i)} \in \mathbb{F}_2^{\rho \times \ell}$ *for all* $i \in [m]$*. Assuming that* $\log((q+p)\rho) = o(r)$*, we can calculate* $\mathbb{E}_R\left[\widetilde{V}^{\widetilde{\pi}}(1^n;\ R)\right]$ *in time* $\widetilde{O}\left(2^{r_{\mathsf{rect}}+r_{\mathsf{shared}}} \cdot (t + m\rho) + m^{q+p} \cdot 2^{q+p+r-\Omega(r/\log((q+p)\rho))}\right)$*.*

Using the parameters of the PCP in Lemma 2, the time bound in the above lemma becomes

$$\widetilde{O}\left(m^{O(1)}\left(2^{0.51n}\rho + 2^{n-\Omega(n/\log\rho)}\right)\right), \tag{1}$$

which is $O(2^n/n)$ when $n/\log\rho \geq \kappa(\log m + \log n)$ for a constant κ. The proof of Lemma 6 follows closely from the computation process of the PCP in Definition 6, similar to parts of the proof of Lemma 3.1 in [17].

Proof of Lemma 6 The algorithm on input $\widetilde{\pi} = \sum_{i=1}^{m} \alpha_i \cdot (-1)^{A^{(i)}B^{(i)}}$ runs as follows:

1. Initialize the result res to be 0.
2. For each $R_{\mathsf{shared}} \in \{0,1\}^{r_{\mathsf{shared}}}$:
   - (a) Compute the decision function $D = D(1^n;\ R_{\mathsf{shared}})$ and randomness parity check
     $(C_1, \ldots, C_p) = (C_1(1^n;\ R_{\mathsf{shared}}), \ldots, C_p(1^n;\ R_{\mathsf{shared}}))$.
   - (b) For each $k \in [q]$, for each $i \in [m]$,
     - (i) Compute the $2^{r_{\mathsf{rect}}} \times \rho$ matrices $A^{(k,i)}$ whose R_{row}-th row is the row of $A^{(i)}$ indexed by $i^{(k)}(1^n;\ R_{\mathsf{row}}, R_{\mathsf{shared}})$ for all $R_{\mathsf{row}} \in \{0,1\}^{r_{\mathsf{rect}}}$.
     - (ii) Compute the $\rho \times 2^{r_{\mathsf{rect}}}$ matrices $B^{(k,i)}$ whose R_{col}-th column is the column of $B^{(i)}$ indexed by $j^{(k)}(1^n;\ R_{\mathsf{col}}, R_{\mathsf{shared}})$ for all $R_{\mathsf{col}} \in \{0,1\}^{r_{\mathsf{rect}}}$.
   - (c) For each $j \in [p]$,
     - (i) Compute the $2^{r_{\mathsf{rect}}} \times 3$ matrix $A^{(q+j,1)}$ and the $3 \times 2^{r_{\mathsf{rect}}}$ matrix $B^{(q+j,1)}$ with

$$\left((-1)^{A^{(q+j,1)}B^{(q+j,1)}}\right)_{R_{\mathsf{row}},R_{\mathsf{col}}} = C_j(R_{\mathsf{row}}, R_{\mathsf{col}}) \text{ given by Lemma 5.}$$

(d) Now we define q m-sums of rank-ρ matrices, $\widetilde{Q}_k = C\sum_{i=1}^m b_i \cdot (-1)^{A^{(k,i)}B^{(k,i)}}$ for each $k \in [q]$, and p 1-sums of rank-3 matrices, $\widetilde{Q}_{q+j} = (-1)^{A^{(q+j,1)}B^{(q+j,1)}}$ for each $j \in [p]$. Apply Theorem 4 to calculate the following value and add it to res:

$$\mathbb{E}_{R_{\mathsf{row}},R_{\mathsf{col}}}\left[\left(\overline{D}(\widetilde{Q}_1,\ldots,\widetilde{Q}_q,\widetilde{Q}_{q+1},\ldots,\widetilde{Q}_{q+p})\right)_{R_{\mathsf{row}},R_{\mathsf{col}}}\right].$$

3. Return res as the value of $\mathbb{E}_R\left[\widetilde{V}^{\widetilde{\pi}}(1^n;\ R)\right]$.

Correctness of the algorithm follows from Definition 6.

Step 2(b) runs in time $O(2^{r_{\mathsf{rect}}} \cdot (t + m\rho))$, while Step 2(c) runs in $O(r2^{r_{\mathsf{rect}}})$ by Lemma 5, which is dominated by the runtime of Step 2(b) since $t \geq r$. By Theorem 4, Step 2(d) takes time $O\left(2^{q+p} \cdot m^{q+p} \cdot (T(2^{r_{\mathsf{rect}}}, (q+p)\rho) + \mathrm{poly}(n, q+p, r))\right) = O\left(2^{q+p} \cdot m^{q+p} \cdot T(2^{r_{\mathsf{rect}}}, (q+p)\rho)\right)\mathrm{poly}(n)$, if $(q+p)\rho = (2^{r_{\mathsf{rect}}})^{o(1)}$, i.e. $\log((q+p)\rho) = o(r)$. Therefore the running time of the above algorithm is

$$\begin{aligned} &O\left(2^{r_{\mathsf{shared}}} \cdot \left(2^{r_{\mathsf{rect}}} \cdot (t + m\rho) + 2^{q+p} \cdot m^{q+p} \cdot T(2^{r_{\mathsf{rect}}}, (q+p)\rho)\right)\right)\mathrm{poly}(n) \\ = \;&O\left(2^{r_{\mathsf{shared}}+r_{\mathsf{rect}}} \cdot (t + m\rho) + m^{q+p} \cdot 2^{q+p+r-\Omega(r/\log((q+p)\rho))}\right)\mathrm{poly}(n). \end{aligned}$$

□

3 Validity Tests

In this section we discuss two tests on sums of low-rank matrices $\widetilde{\pi}$ to ensure that they are close to boolean and somewhat bounded. The following close-to-boolean test simplifies a similar test in [16] due to the smoothness of the PCP verifier. Using the parameters of the PCP in Lemma 2, the time bound in the following lemma becomes $\widetilde{O}(m^4 \cdot 2^{n-\Omega(n/\log\rho)})$, which is $O(2^n/n)$ for m and ρ satisfying $n/\log\rho \geq \kappa(\log m + \log n)$ for a constant κ.

Lemma 7 (Close-to-Boolean Test) *Given $\widetilde{\pi} = C\sum_{i=1}^m b_i \cdot (-1)^{A^{(i)}B^{(i)}}$ with $C \in \mathbb{Q}$ of bit-complexity $O(n)$, $b_i \in \{-1,1\}$, $A^{(i)} \in \mathbb{F}_2^{\ell\times\rho}$, and $B^{(i)} \in \mathbb{F}_2^{\rho\times\ell}$. Assuming that $\rho = \ell^{o(1)}$, we can perform a test on $\widetilde{\pi}$ in time $\widetilde{O}\left(m^4 \cdot \ell^{2-\Omega(1/\log\rho)}\right)$ such that:*

- *(Completeness) If $\widetilde{\pi}$ is bounded and there is a proof $\pi \in \{-1,1\}^{\ell\times\ell}$ with $\|\pi - \widetilde{\pi}\|_1 \leq \varepsilon$, then we have $\|\pi - \widetilde{\pi}\|_2 \leq \sqrt{2\varepsilon}$, and $\widetilde{\pi}$ passes the test.*
- *(Soundness) If $\widetilde{\pi}$ passes the test, there exists a proof $\pi \in \{-1,1\}^{\ell\times\ell}$ with $\|\pi - \widetilde{\pi}\|_2 \leq 2\sqrt{2\varepsilon}$.*

Proof We use Theorem 4 to evaluate the expectation of the degree-4 univariate polynomial $f(x) = (-1-x)^2(1-x)^2$ on $\widetilde{\pi}$. We accept $\widetilde{\pi}$ if $\mathbb{E}_{i,j}[f(\widetilde{\pi}_{i,j})] \leq 8\varepsilon$, and reject otherwise. It takes time $O\left(m^4 \cdot (T(\ell, 4\rho) + \mathrm{poly}(n, \log \ell))\right) = \widetilde{O}\left(m^4 \cdot \ell^{2-\Omega(1/\log \rho)}\right)$ if $\rho = \ell^{o(1)}$.

For soundness, define $\pi \in \{-1,1\}^{\ell \times \ell}$ by $\pi_{i,j} = 1$ if $\widetilde{\pi}_{i,j} \geq 0$, and -1 otherwise, for all $i, j \in [\ell]$.

Then for all $i, j \in [\ell]$, we have $\left|(-\pi_{i,j}) - \widetilde{\pi}_{i,j}\right| \geq 1$. As $\{\pi_{i,j}, -\pi_{i,j}\} = \{-1, 1\}$, we have

$$f(\widetilde{\pi}_{i,j}) = (-1-\widetilde{\pi}_{i,j})^2(1-\widetilde{\pi}_{i,j})^2 = ((-\pi_{i,j}) - \widetilde{\pi}_{i,j})^2(\pi_{i,j} - \widetilde{\pi}_{i,j})^2 \geq (\pi_{i,j} - \widetilde{\pi}_{i,j})^2.$$

Therefore $\|\pi - \widetilde{\pi}\|_2 = \sqrt{\mathbb{E}_{i,j}[(\pi_{i,j} - \widetilde{\pi}_{i,j})^2]} \leq \sqrt{\mathbb{E}_{i,j}[f(\widetilde{\pi}_{i,j})]} \leq 2\sqrt{2\varepsilon}$.

For completeness, observe that for $\widetilde{\pi}_{i,j} \in [-1,1]$ and $\pi_{i,j} \in \{-1,1\}$ we have $\left|(-\pi_{i,j}) - \widetilde{\pi}_{i,j}\right| \leq 2$ and so $(\pi_{i,j} - \widetilde{\pi}_{i,j})^2 \leq 2\left|\pi_{i,j} - \widetilde{\pi}_{i,j}\right|$. Therefore $f(\widetilde{\pi}_{i,j}) \leq 2^2(\pi_{i,j} - \widetilde{\pi}_{i,j})^2 \leq 8\left|\pi_{i,j} - \widetilde{\pi}_{i,j}\right|$, thus $\mathbb{E}_{i,j}[f(\widetilde{\pi}_{i,j})] \leq 8\|\pi - \widetilde{\pi}\|_1 \leq 8\varepsilon$, so $\widetilde{\pi}$ passes the test. Moreover we have $\|\pi - \widetilde{\pi}\|_2 = \sqrt{\mathbb{E}_{i,j}[(\pi_{i,j} - \widetilde{\pi}_{i,j})^2]} \leq \sqrt{2\mathbb{E}_{i,j}\left|\pi_{i,j} - \widetilde{\pi}_{i,j}\right|} = \sqrt{2\|\pi - \widetilde{\pi}\|_1} \leq \sqrt{2\varepsilon}$. □

We also need to test if the sum of low-rank matrices is somewhat bounded. Ideally we would like to ensure that the sum is point-wise bounded. However the quick algorithm in Theorem 4 can only calculate expectation so it is unlikely that we can use it to get a pointwise bound. Fortunately it turns out that for our purpose we don't really need pointwise boundedness. The test we present here generalizes a similar test in [16]. We use the following notion of sampling from the lists I, J.

Definition 8 Let I, J be any two lists of the same size taking (possibly duplicate) elements from $[\ell]$. We say a real matrix $\widetilde{\pi} \in \mathbb{R}^{\ell \times \ell}$ is *power-d bounded for* (I, J) if $\mathbb{E}_{i \sim I, j \sim J}[\widetilde{\pi}_{i,j}^d] \leq 1$, where $i \sim I$ means that i is sampled from I uniformly at random.

Lemma 8 (Boundedness Test) *Let $\widetilde{\pi} = C\sum_{i=1}^m b_i \cdot (-1)^{A^{(i)}B^{(i)}}$ be an m-sum with $C \in \mathbb{Q}$ of bit-complexity $O(\log n)$, $b_i \in \{-1,1\}$, $A^{(i)} \in \mathbb{F}_2^{\ell \times \rho}$, and $B^{(i)} \in \mathbb{F}_2^{\rho \times \ell}$ for all $i \in [m]$. Let I, J be two lists of the same size taking elements from $[\ell]$. Let d be any number. Assuming that $d\rho = |I|^{o(1)}$, we can perform a test on $\widetilde{\pi}$ in time $\widetilde{O}\left(m\rho|I| + m^d|I|^{2-\Omega(1/\log(d\rho))}\right)$ such that:*

- *(Completeness) If $\widetilde{\pi}$ is bounded, it passes the test.*
- *(Soundness) If $\widetilde{\pi}$ passes the test, it is power-d bounded for (I, J).*

Jumping ahead, we will set I (and J) to be the list of the row (column, respectively) indices the verifier probes over row (column, respectively) randomness for each of the q queries and each choice of the shared randomness, so $|I| = |J| = 2^{r_{\mathrm{rect}}}$. Using the parameters of the PCP in Lemma 2, each boundedness test runs in time

$$\widetilde{O}(m^{O(1)}(2^{0.49n}\rho + 2^{0.98n - \Omega(n/\log \rho)})).$$

We will use $O(2^{0.02n})$-many boundedness tests so the total runtime is similar to (1), which becomes $O(2^n/n)$ when $n/\log\rho \geq \kappa(\log m + \log n)$ for a constant κ.

Proof We construct an m-sum $\widetilde{Q} = C\sum_{i=1}^m b_i \cdot (-1)^{A'^{(i)}B'^{(i)}}$, where $A'^{(i)} \in \mathbb{F}_2^{|I|\times\rho}$ consists of the rows of $A^{(i)}$ indexed by elements in I and $B'^{(i)} \in \mathbb{F}_2^{\rho\times|I|}$ consists of the columns of $B^{(i)}$ indexed by elements in J. This step takes time $O(m\rho|I|)$.

Note that the uniform distribution over entries of $\widetilde{Q}$ is the same as the distribution over entries of $\widetilde{\pi}$ under I, J, so we have $\mathbb{E}_{i\sim I, j\sim J}[\widetilde{\pi}_{i,j}^d] = \mathbb{E}_{i,j}\left[\left(\widetilde{Q}\right)_{i,j}^d\right]$. Hence we use Theorem 4 to evaluate the expectation of the polynomial x^d on $\widetilde{Q}$. We accept $\widetilde{\pi}$ if the value is at most 1, and reject otherwise. This step takes time $O(m^d \cdot (T(|I|, d\rho) + \mathrm{poly}(n, d, \log|I|))$ if $d\rho = |I|^{o(1)}$. Therefore the total running time is $O\left(m\rho|I| + m^d|I|^{2-\Omega(1/\log(d\rho))}\right)\mathrm{poly}(n)$.

Completeness and soundness follow from the definition. □

We need the following technical lemma for the main theorem. Intuitively it shows that if a real-valued proof is bounded and close to a boolean proof, then its "acceptance probability" is also close to that of the boolean proof. The proof of this lemma is the most involved one in this paper.

Definition 9 Let V be any $(\ell^2, r, q, p, t, s, \tau)$-rectangular PCP verifier. We say a real matrix $\widetilde{\pi} \in \mathbb{R}^{\ell\times\ell}$ is *bounded for* V if for all $R_{\mathsf{shared}} \in \{0,1\}^{r_{\mathsf{shared}}}$, and all $S \subseteq [q]$, we have

$$\mathbb{E}_{R_{\mathsf{row}}, R_{\mathsf{col}} \in \{0,1\}^{r_{\mathsf{rect}}}}\left[\prod_{k\in S} \widetilde{\pi}^2_{i^{(k)}, j^{(k)}}\right] \leq 1,$$

where $i^{(k)} = i^{(k)}(1^n;\ R_{\mathsf{row}}, R_{\mathsf{shared}})$ and $j^{(k)} = j^{(k)}(1^n;\ R_{\mathsf{row}}, R_{\mathsf{shared}})$ for all $k \in [q]$.

Lemma 9 *Let V be any smooth $(\ell^2, r, q, p, t, s, \tau)$-rectangular PCP verifier over $\{-1, 1\}$, and $\widetilde{V}$ be the same as V but with D multilinearly extended over $\mathbb{R}$. Let π be any matrix in $\{-1,1\}^{\ell\times\ell}$ and $\widetilde{\pi}$ be any matrix in $\mathbb{R}^{\ell\times\ell}$ that is bounded for V. Then we have*

$$\left|\mathbb{E}_R\left[V^{\pi}(1^n;\ R)\right] - \mathbb{E}_R\left[\widetilde{V}^{\widetilde{\pi}}(1^n;\ R)\right]\right| \leq 2^{O(q+p)} \left\|\pi - \widetilde{\pi}\right\|_2.$$

Proof Fix an arbitrary $R_{\mathsf{shared}} \in \{0,1\}^{r_{\mathsf{shared}}}$. By definition,

$$\begin{aligned}
&\mathbb{E}_{R_{\mathsf{row}}, R_{\mathsf{col}}}\left[\left|V^{\pi}(1^n;\ R) - \widetilde{V}^{\widetilde{\pi}}(1^n;\ R)\right|\right] \\
&= \mathbb{E}_{R_{\mathsf{row}}, R_{\mathsf{col}}}\Big[\Big|D(\pi_{i^{(1)}, j^{(1)}}, \ldots, \pi_{i^{(q)}, j^{(q)}}, C_1(R_{\mathsf{row}}, R_{\mathsf{col}}), \ldots, C_p(R_{\mathsf{row}}, R_{\mathsf{col}})) \\
&\qquad - D(\widetilde{\pi}_{i^{(1)}, j^{(1)}}, \ldots, \widetilde{\pi}_{i^{(q)}, j^{(q)}}, C_1(R_{\mathsf{row}}, R_{\mathsf{col}}), \ldots, C_p(R_{\mathsf{row}}, R_{\mathsf{col}}))\Big|\Big], \quad (2)
\end{aligned}$$

where $D = D(1^n;\ R_{\mathsf{shared}})$, $(C_1, \ldots, C_p) = (C_1(1^n;\ R_{\mathsf{shared}}), \ldots, C_p(1^n;\ R_{\mathsf{shared}}))$, $i^{(k)} = i^{(k)}(1^n;\ R_{\mathsf{row}}, R_{\mathsf{shared}})$ and $j^{(k)} = j^{(k)}(1^n;\ R_{\mathsf{row}}, R_{\mathsf{shared}})$ for all $k \in [q]$.

We write D in its Fourier expansion $D(z_1,\ldots,z_{q+p})=\sum_{S\subseteq[q+p]}\beta_S\prod_{k\in S}z_k$, where for each $S\subseteq[q+p]$, $\beta_S=\mathbb{E}_{z\in\{-1,1\}^{q+p}}\left[D(z)\prod_{k\in S}z_k\right]$. For all $z\in\{-1,1\}^{q+p}$, $D(z)\in\{0,1\}$ and $\prod_{k\in S}z_k\in\{-1,1\}$, thus $|\beta_S|\le 1$ for any S. Hence by the triangular inequality we can bound (2) by

$$\begin{aligned}
&\sum_{S\subseteq[q+p]}\mathbb{E}_{R_{\mathsf{row}},R_{\mathsf{col}}}\left[\left|\left(\prod_{k\in S\cap[q]}\pi_{i^{(k)},j^{(k)}}-\prod_{k\in S\cap[q]}\widetilde{\pi}_{i^{(k)},j^{(k)}}\right)\prod_{k\in S\setminus[q]}C_k(R_{\mathsf{row}},R_{\mathsf{col}})\right|\right]\\
=\;&2^p\sum_{S\subseteq[q]}\mathbb{E}_{R_{\mathsf{row}},R_{\mathsf{col}}}\left[\left|\prod_{k\in S}\pi_{i^{(k)},j^{(k)}}-\prod_{k\in S}\widetilde{\pi}_{i^{(k)},j^{(k)}}\right|\right], \qquad (3)
\end{aligned}$$

as all the C_k's are $\{-1,1\}$-valued.

Fix any $S\subseteq[q]$. Wlog let $S=\{1,\ldots,d\}$ for some $d\le q$, then the expectation in (3) can be written as

$$\begin{aligned}
&\mathbb{E}_{R_{\mathsf{row}},R_{\mathsf{col}}}\left[\left|\prod_{u=1}^{d}\pi_{i^{(u)},j^{(u)}}-\prod_{u=1}^{d}\widetilde{\pi}_{i^{(u)},j^{(u)}}\right|\right]\\
=\;&\mathbb{E}_{R_{\mathsf{row}},R_{\mathsf{col}}}\left[\left|\sum_{v=1}^{d}\left(\prod_{u=1}^{v-1}\widetilde{\pi}_{i^{(u)},j^{(u)}}\prod_{u=v}^{d}\pi_{i^{(u)},j^{(u)}}-\prod_{u=1}^{v}\widetilde{\pi}_{i^{(u)},j^{(u)}}\prod_{u=v+1}^{d}\pi_{i^{(u)},j^{(u)}}\right)\right|\right]\\
\le\;&\sum_{v=1}^{d}\mathbb{E}_{R_{\mathsf{row}},R_{\mathsf{col}}}\left[\left|\prod_{u=1}^{v-1}\widetilde{\pi}_{i^{(u)},j^{(u)}}\prod_{u=v}^{d}\pi_{i^{(u)},j^{(u)}}-\prod_{u=1}^{v}\widetilde{\pi}_{i^{(u)},j^{(u)}}\prod_{u=v+1}^{d}\pi_{i^{(u)},j^{(u)}}\right|\right]\\
=\;&\sum_{v=1}^{d}\mathbb{E}_{R_{\mathsf{row}},R_{\mathsf{col}}}\left[\left|\prod_{u=1}^{v-1}\widetilde{\pi}_{i^{(u)},j^{(u)}}\cdot\left(\pi_{i^{(v)},j^{(v)}}-\widetilde{\pi}_{i^{(v)},j^{(v)}}\right)\cdot\prod_{u=v+1}^{d}\pi_{i^{(u)},j^{(u)}}\right|\right]\\
\le\;&\sum_{v=1}^{d}\left(\left(\mathbb{E}_{R_{\mathsf{row}},R_{\mathsf{col}}}\left[\left(\pi_{i^{(v)},j^{(v)}}-\widetilde{\pi}_{i^{(v)},j^{(v)}}\right)^2\right]\right)^{1/2}\right.\\
&\qquad\left.\cdot\left(\mathbb{E}_{R_{\mathsf{row}},R_{\mathsf{col}}}\left[\prod_{u=1}^{v-1}\widetilde{\pi}^2_{i^{(u)},j^{(u)}}\prod_{u=v+1}^{d}\pi^2_{i^{(u)},j^{(u)}}\right]\right)^{1/2}\right)\\
\le\;&\sum_{v=1}^{d}\left(\mathbb{E}_{R_{\mathsf{row}},R_{\mathsf{col}}}\left[\left(\pi_{i^{(v)},j^{(v)}}-\widetilde{\pi}_{i^{(v)},j^{(v)}}\right)^2\right]\right)^{1/2}\\
=\;&\sum_{k\in S}\left(\mathbb{E}_{R_{\mathsf{row}},R_{\mathsf{col}}}\left[\left(\pi_{i^{(k)},j^{(k)}}-\widetilde{\pi}_{i^{(k)},j^{(k)}}\right)^2\right]\right)^{1/2},
\end{aligned}$$

where the first inequality comes from the triangular inequality, the second inequality follows from the Cauchy-Schwarz inequality, and the last inequality follows from the assumptions that $\pi\in\{-1,1\}^{\ell\times\ell}$ and $\widetilde{\pi}$ is bounded for V.

Summing over S, we can bound (3) by

$$\begin{aligned}
&2^p\sum_{S\subseteq[q]}\sum_{k\in S}\left(\mathbb{E}_{R_{\mathsf{row}},R_{\mathsf{col}}}\left[\left(\pi_{i^{(k)},j^{(k)}}-\widetilde{\pi}_{i^{(k)},j^{(k)}}\right)^2\right]\right)^{1/2}\\
=\;&2^{p+q-1}\sum_{k\in[q]}\left(\mathbb{E}_{R_{\mathsf{row}},R_{\mathsf{col}}}\left[\left(\pi_{i^{(k)},j^{(k)}}-\widetilde{\pi}_{i^{(k)},j^{(k)}}\right)^2\right]\right)^{1/2}
\end{aligned}$$

$$
\begin{aligned}
&= 2^{O(p+q)}\mathbb{E}_{k\in[q]}\left[\left(\mathbb{E}_{R_{\mathsf{row}},R_{\mathsf{col}}}\left[\left(\pi_{i^{(k)},j^{(k)}}-\widetilde{\pi}_{i^{(k)},j^{(k)}}\right)^2\right]\right)^{1/2}\right]\\
&\le 2^{O(p+q)}\left(\mathbb{E}_{R_{\mathsf{row}},R_{\mathsf{col}},k}\left[\left(\pi_{i^{(k)},j^{(k)}}-\widetilde{\pi}_{i^{(k)},j^{(k)}}\right)^2\right]\right)^{1/2},
\end{aligned}
$$

where the first step uses double counting, and the last step follows from Jensen's inequality.

Therefore by averaging over R_{shared}, we have

$$
\begin{aligned}
&\left|\mathbb{E}_R\left[V^{\pi}(1^n;\ R)\right]-\mathbb{E}_R\left[\widetilde{V}^{\widetilde{\pi}}(1^n;\ R)\right]\right|\\
\le\ & \mathbb{E}_{R_{\mathsf{shared}}}\mathbb{E}_{R_{\mathsf{row}},R_{\mathsf{col}}}\left[\left|V^{\pi}(1^n;\ R)-\widetilde{V}^{\widetilde{\pi}}(1^n;\ R)\right|\right]\\
\le\ & 2^{O(p+q)}\mathbb{E}_{R_{\mathsf{shared}}}\left[\left(\mathbb{E}_{R_{\mathsf{row}},R_{\mathsf{col}},k}\left[\left(\pi_{i^{(k)},j^{(k)}}-\widetilde{\pi}_{i^{(k)},j^{(k)}}\right)^2\right]\right)^{1/2}\right]\\
\le\ & 2^{O(p+q)}\left(\mathbb{E}_{R_{\mathsf{shared}},R_{\mathsf{row}},R_{\mathsf{col}},k}\left[\left(\pi_{i^{(k)},j^{(k)}}-\widetilde{\pi}_{i^{(k)},j^{(k)}}\right)^2\right]\right)^{1/2}\\
=\ & 2^{O(p+q)}\left(\mathbb{E}_{i,j}\left[\left(\pi_{i,j}-\widetilde{\pi}_{i,j}\right)^2\right]\right)^{1/2}\\
=\ & 2^{O(p+q)}\left\|\pi-\widetilde{\pi}\right\|_2,
\end{aligned}
$$

where the first step uses triangular inequality, the second step uses the above bound for every R_{shared}, the third step comes from Jensen's inequality, and the fourth step follows from the smoothness of V. □

4 Constant hardness for rational sums of low-rank matrices

In this section we prove our main hardness result against rational sums of low-rank matrices.

Theorem 10 *There is a function $f\colon\{0,1\}^{n+O(\log n)}\to\{-1,1\}$ in $\mathrm{E}^{\mathbf{NP}}$ that does not $(1-\Omega(1))$-correlate with any bounded m-sum of rank-ρ matrices with $O(n)$ bit-complexity, for infinitely many n, as long as $n/\log\rho\ge\kappa(\log m+\log n)$ for a constant κ.*

Proof Fix L to be a unary language in $\mathbf{NTIME}(2^n)\setminus\mathbf{NTIME}(o(2^n))$ [30–32]. Let V be the smooth (ℓ^2,r,q,p,t,s,τ)-rectangular PCP verifier over alphabet $\{-1,1\}$ for L given by Lemma 2, for s and τ to be determined later. Let $\widetilde{V}$ be the same as V but with D multilinearly extended over $\mathbb{R}$.

We use the lexicographically first proof oracle π as our hard function, i.e. our algorithm $f_n\colon[\ell]\times[\ell]\to\{-1,1\}$ on input (i,j) searches bit-by-bit for the lexicographically first proof π such that $\forall R$, $V^{\pi}(1^n;\ R)=1$ if one exists, and outputs $\pi_{i,j}$. If no such proof exists, we simply output 0. Clearly $f_n\in\mathrm{E}^{\mathbf{NP}}$. Note that f_n can also be seen as a family of matrices $f_n\in\{-1,1\}^{\ell\times\ell}$.

Now for the sake of contradiction, we assume that f_n $(1-\varepsilon)$-correlates with a bounded m-sum of rank-ρ matrices $\widetilde{\pi}$ with bit-complexity $O(n)$, for a constant

ε to be determined later. We will show that $L \in \mathbf{NTIME}(2^n/n)$, thus deriving a contradiction.

Algorithm

The nondeterministic algorithm for L goes as follows:

1. Guess $\widetilde{\pi} = C\sum_{i=1}^{m} b_i \cdot (-1)^{A^{(i)}B^{(i)}}$ by guessing $b_i \in \{-1,1\}$, matrices $A^{(i)} \in \mathbb{F}_2^{\ell\times\rho}$, and $B^{(i)} \in \mathbb{F}_2^{\rho\times\ell}$ for all $i \in [m]$ and $C \in \mathbb{Q}$ with bit-complexity $O(n)$.
2. Perform the close-to-boolean test in Lemma 7 for ε on $\widetilde{\pi}$, reject if it doesn't pass.
3. For each $R_{\mathsf{shared}} \in \{0,1\}^{r_{\mathsf{shared}}}$, $k \in [q]$:
   (a) Compute the lists
$$\begin{aligned} I^{(k)} &= \left[i^{(k)}(1^n;\ R_{\mathsf{row}}, R_{\mathsf{shared}}) | R_{\mathsf{row}} \in \{0,1\}^{r_{\mathsf{rect}}}\right], \\ J^{(k)} &= \left[j^{(k)}(1^n;\ R_{\mathsf{col}}, R_{\mathsf{shared}}) | R_{\mathsf{col}} \in \{0,1\}^{r_{\mathsf{rect}}}\right]. \end{aligned}$$
   (b) For each $2 \le d \le 2q$:
      (i) Perform the boundedness test in Lemma 8 for d and $(I^{(k)}, J^{(k)})$ on $\widetilde{\pi}$, reject if it doesn't pass.
4. Use Lemma 6 to calculate $\mathbb{E}_R[\widetilde{V}^{\widetilde{\pi}}(1^n;\ R)]$, and accept if $\mathbb{E}_R[\widetilde{V}^{\widetilde{\pi}}(1^n;\ R)] > \gamma$ where the constant γ is to be determined later, otherwise reject.

Runtime

Step 1 takes time $O(m\ell\rho + n)$.

By Lemma 7, Step 2 takes time $\widetilde{O}\left(m^4 \cdot \ell^{2-\Omega(1/\log\rho)}\right)$ if $\rho = \ell^{o(1)}$.

Step 3(a) takes time $\widetilde{O}\left(2^{r_{\mathsf{rect}}} \cdot t\right)$, and we have $|I| = 2^{r_{\mathsf{rect}}}$. Therefore by Lemma 8, Step 3(b) takes time $\widetilde{O}\left(qm\rho 2^{r_{\mathsf{rect}}} + qm^{2q}2^{2r_{\mathsf{rect}}-\Omega(r/\log(q\rho))}\right)$, if $q\rho = (2^{r_{\mathsf{rect}}})^{o(1)}$, i.e. $\log(q\rho) = o(r)$. Hence the total runtime for Step 3 is

$$\begin{aligned} &\widetilde{O}\left(2^{r_{\mathsf{shared}}} \cdot \left(2^{r_{\mathsf{rect}}} \cdot (t + qm\rho) + q \cdot m^{2q} \cdot 2^{2r_{\mathsf{rect}}-\Omega(r/\log(q\rho))}\right)\right) \\ = \ &\widetilde{O}\left(2^{r_{\mathsf{shared}}+r_{\mathsf{rect}}} \cdot (t + qm\rho) + qm^{2q} \cdot 2^{r-\Omega(r/\log(q\rho))}\right). \end{aligned}$$

By Lemma 6, Step 4 runs in time

$$\widetilde{O}\left(2^{r_{\mathsf{shared}}+r_{\mathsf{rect}}} \cdot (t + m\rho) + m^{q+p} \cdot 2^{q+p+r-\Omega(r/\log((q+p)\rho))}\right)$$

if $\log((q+p)\rho) = o(r)$.

For the algorithm to run in time $O(2^n/n)$, it suffices to satisfy all the above requirements and make all the runtime to be $O(2^n/n)$. For convenience we take logarithms on all the time bounds. In summary, it is sufficient to satisfy the following conditions:

1. $\log(m\ell\rho + n) < n - \log n$.
2. $\rho = \ell^{o(1)}$.
3. $\log(m^4\ell^2) - \Omega(\log\ell/\log\rho) + O(\log n) < n - \log n$.
4. $\log(q\rho) = o(r)$.

5. $r_{\mathsf{shared}} + r_{\mathsf{rect}} + \log(t + qm\rho) + O(\log n) < n - \log n$.
6. $\log q + 2q \log m + r - \Omega\left(\frac{r}{\log(q\rho)}\right) + O(\log n) < n - \log n$.
7. $\log((q + p)\rho) = o(r)$.
8. $(q + p)(\log m + 1) + r - \Omega\left(\frac{r}{\log((q+p)\rho)}\right) + O(\log n) < n - \log n$.

We are going to set parameters to meet these conditions at the end.

Correctness

We first prove the following claim.

Claim 11 *If $\widetilde{\pi}$ passes all the tests in the definition of the algorithm then it is bounded for V.*

Proof Fix any $R_{\mathsf{shared}} \in \{0,1\}^{r_{\mathsf{shared}}}$ and $S \subseteq [q]$. Let $d = |S|$. By Hölder's inequality, we get

$$\mathbb{E}_{R_{\mathsf{row}},R_{\mathsf{col}}}\left[\prod_{k\in S} \widetilde{\pi}^2_{i^{(k)},j^{(k)}}\right] \leq \prod_{k\in S}\left(\mathbb{E}_{R_{\mathsf{row}},R_{\mathsf{col}}}\left[\widetilde{\pi}^{2d}_{i^{(k)},j^{(k)}}\right]\right)^{1/d}$$
$$= \prod_{k\in S}\left(\mathbb{E}_{i\sim I^{(k)},j\sim J^{(k)}}\left[\widetilde{\pi}^{2d}_{i,j}\right]\right)^{1/d}.$$

We have $2d \leq 2q$, therefore the boundedness tests in Step 3 can guarantee that all the terms in the product are bounded by 1, hence $\mathbb{E}_{R_{\mathsf{row}},R_{\mathsf{col}}}\left[\prod_{k\in S}\widetilde{\pi}^2_{i^{(k)},j^{(k)}}\right] \leq 1$, thus by definition $\widetilde{\pi}$ is bounded for V. □

If $x = 1^n \in L$, let $\widetilde{\pi} \in \mathbb{R}^{\ell\times\ell}$ be any bounded m-sum of rank-ρ matrices with bit-complexity $O(\log n)$ that $(1-\varepsilon)$-correlates with the hard function f_n, which is the lexicographically first proof π in this case. We can assume wlog $\mathbb{E}_{i,j}[\pi_{i,j}\widetilde{\pi}_{i,j}] \geq 1-\varepsilon$, otherwise we can simply use $-\widetilde{\pi}$. Note that for any $x \in \{-1,1\}$, $y \in [-1,1]$ we have $|x-y| = 1-xy$. As $\pi_{i,j} \in \{-1,1\}$, $\widetilde{\pi}_{i,j} \in [-1,1]$, we have

$$\|\pi - \widetilde{\pi}\|_1 = \mathbb{E}_{i,j}\left[\left|\pi_{i,j} - \widetilde{\pi}_{i,j}\right|\right] = 1 - \mathbb{E}_{i,j}\left[\pi_{i,j}\widetilde{\pi}_{i,j}\right] \leq \varepsilon.$$

Hence $\|\pi - \widetilde{\pi}\|_2 \leq \sqrt{2\varepsilon}$, and moreover $\widetilde{\pi}$ passes the close-to-boolean test. Since $\widetilde{\pi}$ is bounded by assumption it is also bounded for V. Therefore by Lemma 9 we have

$$\begin{aligned}\mathbb{E}_R\left[\widetilde{V}^{\widetilde{\pi}}(1^n;\ R)\right] &\geq \mathbb{E}_R[V^{\pi}(1^n;\ R)] - \left|\mathbb{E}_R\left[V^{\pi}(1^n;\ R)\right] - \mathbb{E}_R\left[\widetilde{V}^{\widetilde{\pi}}(1^n;\ R)\right]\right| \\ &\geq 1 - 2^{O(p+q)}\|\pi - \widetilde{\pi}\|_2 \\ &\geq 1 - 2^{O(p+q)}\sqrt{\varepsilon}.\end{aligned}$$

If $x = 1^n \notin L$, then for any guessed m-sum of matrices $\widetilde{\pi}$ that passes all the tests, by soundness there exists a boolean proof $\pi \in \{-1,1\}^{\ell\times\ell}$ such that $\|\pi - \widetilde{\pi}\|_2 \leq 2\sqrt{2\varepsilon}$, and by the above claim we know that $\widetilde{\pi}$ is bounded for V. Therefore by Lemma 9 we have

$$\begin{aligned}\mathbb{E}_R[\widetilde{V}^{\widetilde{\pi}}(1^n;\ R)] &\leq \mathbb{E}_R[V^{\pi}(1^n;\ R)] + \left|\mathbb{E}_R\left[V^{\pi}(1^n;\ R)\right] - \mathbb{E}_R\left[\widetilde{V}^{\widetilde{\pi}}(1^n;\ R)\right]\right| \\ &\leq s + 2^{O(p+q)}\|\pi - \widetilde{\pi}\|_2\end{aligned}$$

$$\leq s + 2^{O(p+q)}\sqrt{\varepsilon}.$$

If there is a gap between $s + 2^{O(p+q)}\sqrt{\varepsilon}$ and $1 - 2^{O(p+q)}\sqrt{\varepsilon}$, we can set the parameter γ in the last step of the algorithm to be any value between these two. Assuming that this gap can be created and that Conditions 1-8 are all met, the above nondeterministic algorithm computes L in time $O(2^n/n)$, a contradiction to our choice of L.

Setting constant parameters

We choose an arbitrary small constant s so both p,q are constants by Lemma 2. To make sure that $s + 2^{O(p+q)}\sqrt{\varepsilon} < 1 - 2^{O(p+q)}\sqrt{\varepsilon}$, we set ε to be a constant smaller than $\left(\frac{1-s}{2^{O(q+p)}}\right)^2$.

Verifying Conditions 1-8

Now fix any ρ and m such that $n/\log\rho \geq \kappa(\log m + \log n)$ for a large constant κ to be determined. We are going to verify that Conditions 1-8 are satisfied. Note that by Lemma 2 we have $r = n + O(\log n)$, $\log \ell = n/2 + O(\log n)$, and $\log t = O(\tau n)$. First, $\log\rho \leq n/(\kappa \log n) = o(r)$ and similarly $\rho = \ell^{o(1)}$, so Conditions 2, 4, and 7 are all satisfied. As both $\log\rho$ and $\log m$ are at most n/κ, for Condition 1 we have

$$\log(m\ell\rho) + \log n \leq 2n/\kappa + n/2 + O(\log n) < n - \log n,$$

for κ sufficiently large, while for Condition 5 we have

$$(1+\tau)r/2 + \log t + \log(qm\rho) + O(\log n) \leq (1/2 + O(\tau))n + \log\rho + \log m < n - \log n,$$

for κ sufficiently large and τ sufficiently small. For Condition 8 we have

$$\begin{aligned} & (q+p)(\log m + 1) + r - \Omega\left(\frac{(1-\tau)r}{\log((q+p)\rho)}\right) + O(\log n) \\ \leq\ & O(\log m) + n + O(\log n) - \Omega(n/\log\rho) \\ \leq\ & n + O(\log m + \log n) - \Omega(\kappa(\log m + \log n)) \\ <\ & n - \log n, \end{aligned}$$

for κ sufficiently large. Similarly Conditions 3 and 6 are also satisfied, and we are done. □

5 Correlation bounds via XOR Lemma

In this section we adapt the approach in [16] to our setting, and then prove the main result in this paper, Theorem 1. We first show the following XOR Lemma.

Definition 10 For any boolean function $f\colon \{0,1\}^n \to \{-1,1\}$ and number k, we define $f^{\oplus k}\colon \{0,1\}^{nk} \to \{-1,1\}$ by $f^{\oplus k}(x_1,\ldots,x_k) = \prod_{i=1}^{k} f(x_i)$ for all $x_1,\ldots,x_k \in \{0,1\}^n$.

Lemma 12 *Let $f\colon \{0,1\}^n \to \{-1,1\}$ be any boolean function. Let rational $\varepsilon < 1$ have constant bit-complexity, and for any number $k \geq 1$, let $\varepsilon_k = (\frac{1+\varepsilon}{2})^{k-1}\varepsilon$. Assume* *that $f^{\oplus k}$ ε_k-correlates with some function $h\colon \{0,1\}^{nk} \to [-1,1]$. Then f ε-correlates with a bounded m-sum of restrictions of h (by fixing some inputs), where $m = O\left(\frac{n}{\varepsilon_k^2}\right)$ and the bit-complexity is $O(k + \log n)$.*

Proof We prove it by induction on k. For $k = 1$ it is trivial as h is bounded. Now we assume that the hypothesis holds for $k-1$, and we are proving for k.

For all $x_1 \in \{0,1\}^n$, define $g(x_1) = \mathbb{E}_{y\sim\{0,1\}^{n(k-1)}}\left[f^{\oplus k-1}(y)h(x_1, y)\right]$, where we use y for $(x_2, \ldots, x_k)$ for convenience. If there exists $x_1 \in \{0,1\}^n$ such that $|g(x_1)| \geq \varepsilon_{k-1}$, then we know that $f^{\oplus k-1}$ ε_{k-1}-correlates with h' defined by $h'(y) = h(x_1, y)$, so we can use the induction hypothesis for $k-1$ to get a bounded m-sum of functions obtained by fixing inputs of h', thus by fixing inputs of h.

Otherwise, for all $x_1 \in \{0,1\}^n$ we have $|g(x_1)| \leq \varepsilon_{k-1} = \frac{2\varepsilon_k}{1+\varepsilon}$. We take m i.i.d. samples $y_1, \ldots, y_m$ uniformly from $\{0,1\}^{n(k-1)}$ for $m = O\left(\frac{n}{(\varepsilon_k)^2}\right)$, then define $\widetilde{g}(x_1) = \mathbb{E}_{i\in[m]}\left[f^{\oplus k-1}(y_i)h(x_1, y_i)\right]$. By Chernoff bound,

$$\Pr_{y_1,\ldots,y_m}\left[|g(x_1) - \widetilde{g}(x_1)| \geq \frac{1-\varepsilon}{(1+\varepsilon)^2}\varepsilon_k\right] \leq 2^{-n-1}.$$

By union bound, there exists a fixed assignment to $y_1, \ldots, y_m$ such that for all $x_1 \in \{0,1\}^n$,

$$|g(x_1) - \widetilde{g}(x_1)| \leq \frac{1-\varepsilon}{(1+\varepsilon)^2}\varepsilon_k, \tag{4}$$

thus $|\widetilde{g}(x_1)| \leq |g(x_1)| + |g(x_1) - \widetilde{g}(x_1)| \leq \left(\frac{2}{1+\varepsilon} + \frac{1-\varepsilon}{(1+\varepsilon)^2}\right)\varepsilon_k = \frac{3+\varepsilon}{(1+\varepsilon)^2}\varepsilon_k$.

Let $r = \frac{3+\varepsilon}{(1+\varepsilon)^2}\varepsilon_k$. We define $\widetilde{h}$ by

$$\widetilde{h}(x_1) = \frac{\widetilde{g}(x_1)}{r} = \frac{1}{mr}\sum_{i=1}^{m} f^{\oplus k-1}(y_i)h(x_1, y_i).$$

Now $|\widetilde{h}(x_1)| \leq 1$ for all x_1. We can write $\widetilde{h}$ as $C\sum_{i=1}^m b_i h_i$, where

$$\begin{aligned}
&C = \frac{1}{mr},\\
&b_i = f^{\oplus k-1}(y_i), \forall i \in [m],\\
&h_i\colon x_1 \mapsto h(x_1, y_i), \forall x_1 \in \{0,1\}^n, \forall i \in [m],
\end{aligned}$$

which is a bounded $O(n/\varepsilon_k^2)$-sum of functions that can be obtained by fixing inputs of h. The bit-complexity of m is $O(k + \log n)$, and $O(k)$ for r, thus the bit-complexity of $\widetilde{h}$ is $O(k + \log n)$.

What remains is to prove that $\mathsf{corr}(f, \widetilde{h}) \geq \varepsilon$. By the definition of g and assumption we have $\mathsf{corr}(f, g) = \mathsf{corr}(f^{\oplus k}, h) \geq \varepsilon_k$. Therefore by the definition of $\widetilde{h}$, the fact that $f(x_1) \in \{-1, 1\}$ for all x_1, and (4), we have

$$\begin{aligned}
\mathsf{corr}(f, \widetilde{h}) &= \frac{\mathsf{corr}(f, \widetilde{g})}{r}\\
&\geq \frac{1}{r}\left(\mathsf{corr}(f, g) - \mathbb{E}_{x_1}|g(x_1) - \widetilde{g}(x_1)|\right)
\end{aligned}$$

$$
\begin{aligned}
&\geq \frac{\varepsilon_k - \frac{1-\varepsilon}{(1+\varepsilon)^2}\varepsilon_k}{\frac{3+\varepsilon}{(1+\varepsilon)^2}\varepsilon_k} \\
&= \varepsilon.
\end{aligned}
$$

□

Now we can prove our main theorem. We will first prove a weaker version that only works for infinitely many n, then show how to extend it to the full version that works for all sufficiently large n.

Proof of Theorem 1 for infinitely many n Let $f\colon \{0,1\}^{n+O(\log n)} \to \{-1,1\}$ in $\mathrm{E}^{\mathbf{NP}}$ be the function given by Theorem 10. We know that f does not ε-correlate with any bounded m-sum of rank-ρ matrices with $O(n)$ bit-complexity for infinitely many n, for a rational constant ε with constant bit-complexity.

Let $k \leq n$. We set $\varepsilon_k = (\frac{1+\varepsilon}{2})^{k-1}\varepsilon = 2^{-\Theta(k)}$, and $F = f^{\oplus k}$ so $N = k(n+O(\log n))$. For the sake of contradiction we assume that F ε_k-correlates with some rank-ρ matrix h for all such N. We view a matrix as the truth table of a function, so when we take restrictions on the function, we are taking some rows and columns of the matrix but keeping its dimensions, thus the rank doesn't increase. By Lemma 12 we know that f ε-correlates with a bounded m-sum of rank-ρ matrices $\widetilde{h}$, where $m = O(n/\varepsilon_k^2) = n2^{O(k)}$ and the bit complexity is $O(k + \log n) = O(n)$.

To get a contradiction for infinitely many n we still need to verify that $n/\log\rho \geq \kappa(\log m + \log n)$ for a sufficiently large constant κ given by Theorem 10. We have $\log n = \log N - \log k - O(\log\log n)$, thus

$$\log m + \log n = 2\log n + O(k) = O(\log N + k).$$

Let $c > 0$ be the constant hidden in the last big-O. We take $\delta = 1/2c\kappa$. Then

$$\kappa(\log m + \log n)\log\rho \leq c\kappa(\log N + k)\log\rho \leq \frac{N}{2k} = \frac{n + O(\log n)}{2} < n.$$

This shows that for infinitely many N, the correlation of $F\colon \{0,1\}^N \to \{-1,1\}$ and any $\{-1,1\}$-matrix $h = (-1)^{h'}$, where h' is a rank-ρ $\mathbb{F}_2$-matrix, is at most ε_k. Thus we have $\mathbb{P}_{x,y}\left[F(x,y) \neq \left((-1)^{h'}\right)_{x,y}\right] \geq 1/2 - 2^{-\Omega(k)}$. Converting the output basis back to the boolean basis, we get the weaker version of Theorem 1 that only works for infinitely many N. □

To prove the full version of Theorem 1 that works for all sufficiently large n, we need the following *refuter* from [16].

Theorem 13 *There is a constant $c > 0$ such that the following holds.*

For any non-decreasing time-constructible function $T(n)$ such that $n \leq T(n) \leq 2^{\mathrm{poly}(n)}$, there is an $\mathbf{NTIME}(T(n))$ *language L and an algorithm R such that:*

Input. *The input for R is a pair $(M, 1^n)$ where M is a nondeterministic algorithm running in time $\leq cT(n)/\log T(n)$.*
Output. *For any fixed M, for all large enough n, $R(M, 1^n)$ outputs a string x such that $|x| \in [n, n+T(n)]$ and $L(x) \neq M(x)$.*
Complexity. *R is a deterministic algorithm running in $O(T(n)\cdot T(T(n)+n))$ time with an* $\mathbf{NP}$ *oracle.*

We also need a lemma on padding rigid matrices from [10].

Lemma 14 ([10, Lemma II.7]) *Let $\mathbf{1}_m$ be the all-ones $m \times m$ matrix. For any square matrix A, A ε-correlates with some rank-ρ matrix if and only if $A \otimes \mathbf{1}_m$ ε-correlates with some rank-ρ matrix, where $\otimes$ is the tensor product of matrices.*

Proof of Theorem 1 We show how to remove the "infinitely often" part from the previous proof. Fix $T(n)$ to be the half-exponential function where $T(T(n)) = 2^n$. We are going to use the general version of Lemma 2, Theorem 8.2 in [29] that works for $\mathbf{NTIME}(T(n))$ languages. Most importantly, we have $r = \log T(n) + O(\log\log T(n)) + O(\log n)$ and $2\log \ell = r + O(1)$. Then the proof of Theorem 10 shows that we have the following results:

1. For any $\mathbf{NTIME}(T(n))$ language L, let V be the smooth $(\ell^2, r, q, p, t, s, \tau)$-rectangular PCP verifier over alphabet $\{-1, 1\}$ for L given by the generalized version of Lemma 2, for small constants s and τ. We define the function $f_{L,x} \colon [\ell] \times [\ell] \to \{-1, 1\}$ such that on input (i, j) it searches bit-by-bit for the lexicographically first proof π such that $\forall R$, $V^{\pi}(x;\ R) = 1$ if one exists, and outputs $\pi_{i,j}$. Clearly $f_{L,x} \in \mathrm{E}^{\mathbf{NP}}$. Note that $f_{L,x}$ can also be seen as an $\ell \times \ell$ matrix.
2. For any $\mathbf{NTIME}(T(n))$ language L, there exists an explicit nondeterministic algorithm that decides if $x \in L$ in time $O(T(n)/\log T(n))$, for any input x such that $|x| = n$ and $f_{L,x}$ $(1 - \Omega(1))$-correlates with a bounded m-sum of rank-ρ matrices $\widetilde{\pi}$ with bit-complexity $O(n)$, as long as $\log T(n)/\log \rho \geq \kappa(\log m + \log n)$ for a constant κ.

We consider the language L for $\mathbf{NTIME}(T(n))$ from Theorem 13. Similarly as before, by combining Item 2 with Lemma 12, there exists an explicit nondeterministic $O(T(n)/\log T(n))$-time algorithm M deciding if $x \in L$ for any input x such that $|x| = n$ and $f_{L,x}^{\oplus k}$ ε_k-correlates with some rank-ρ matrix h, as long as $\log T(n)/\log \rho \geq \kappa(\log m + \log n)$ for $m = n2^{O(k)}$.

We aim to use the refuter R from Theorem 13 to get a contradiction. For all large enough n, R on $(1^n, M)$ will output an x such that $|x| \in [n, n + T(n)]$ and $L(x) \neq M(x)$. Now the input length of $f_{L,x}^{\oplus k}$ is $k \cdot 2\log \ell = k(r + O(1)) = k(\log T(|x|) + O(\log\log T(|x|)) + O(\log |x|)) \leq 2k \log T(T(n)) = 2kn$. We view $f_{L,x}^{\oplus k}$ as a matrix and use Lemma 14 to get a function F_x with input length $2kn$ such that F_x ε_k-correlates with some rank-ρ matrix iff $f_{L,x}^{\oplus k}$ ε_k-correlates with some rank-ρ matrix. Therefore if F_x ε_k-correlates with some rank-ρ matrix then M can decide if $x \in L$, a contradiction.

Our final hard function f works as follows. On input of length $N = 2kn$ it runs R on $(1^n, M)$ to get an x, then run F_x. Then for all large enough N, f does not ε_k-correlate with any rank-ρ matrices. R runs in $O(T(n) \cdot T(T(n))) = 2^{O(n)} = 2^{O(N/k)}$ time with an $\mathbf{NP}$ oracle and $f_{L,x} \in \mathrm{E}^{\mathbf{NP}}$, thus $f \in \mathrm{E}^{\mathbf{NP}}$. The condition $\log T(n)/\log \rho \geq \kappa(\log m + \log n)$ in Item 2 can be verified similarly as in the previous proof for a sufficiently small δ. □

Remark 1 Note that Item 2 in the above proof only works because we choose a superpolynomial $T(n)$. Previously in the manuscript we mistakenly chose $T(n)$ to be a polynomial, which was pointed out by one of the anonymous reviewers. Such $T(n)$ won't work because there won't be a large enough gap between $\log T(n)/\log \rho$ and $\kappa(\log m + \log n)$. The "$+\log n$" term here comes from the "$+O(\log n)$" terms in Conditions 1-8 in the proof of Theorem 10, which are the results of the following two overheads: 1) is the $O(\log n)$ overhead of the length of randomness from the rectangular PCP, Theorem 8.2 in [29]; 2) is the overhead of arithmetic computations in Theorem 4 that we hide in $\widetilde{O}$ later. 1) can be eliminated when $T(n)$ is polynomial by Remarks 5.3 and 8.3 in [29], as suggested by the above-mentioned anonymous reviewer. However we don't know how to eliminate 2) thus it prevents us to use a polynomial $T(n)$, among other reasons. By choosing $T(n)$ to be half-exponential, we create a large enough gap between $\log T(n)/\log \rho$ and $\kappa(\log m+\log n)$ so that it can be separated, and the maximum input length of $f_{L,x}$ (where x is produced by the refuter R) can be bounded by $2\log T(T(n)) = 2n$ so the final input length N is similar to the one in the previous proof, thus the verification of $\log T(n)/\log \rho \geq \log T(n)/\log \rho$ is also similar.

References

[1] Williams, R.: Nonuniform ACC circuit lower bounds. J. of the ACM **61**(1), 2–1232 (2014)

[2] Williams, R.: Guest column: a casual tour around a circuit complexity bound. SIGACT News **42**(3), 54–76 (2011)

[3] Williams, R.: Natural proofs versus derandomization. In: ACM Symp. on the Theory of Computing (STOC) (2013)

[4] Williams, R.: New algorithms and lower bounds for circuits with linear threshold gates. In: ACM Symp. on the Theory of Computing (STOC) (2014)

[5] Alman, J., Chan, T.M., Williams, R.: Polynomial representations of threshold functions and algorithmic applications. In: IEEE Symp. on Foundations of Computer Science (FOCS) (2016)

[6] Tamaki, S.: A satisfiability algorithm for depth two circuits with a sub-quadratic number of symmetric and threshold gates. Electronic Colloquium on Computational Complexity (ECCC) **23**, 100 (2016)

[7] Chen, R., Oliveira, I.C., Santhanam, R.: An average-case lower bound against ACC^0. In: LATIN. Lecture Notes in Computer Science, vol. 10807, pp. 317–330 (2018)

[8] Murray, C., Williams, R.R.: Circuit lower bounds for nondeterministic quasi-polytime: an easy witness lemma for NP and NQP. In: STOC, pp.

890–901 (2018)

[9] Rajgopal, N., Santhanam, R., Srinivasan, S.: Deterministically counting satisfying assignments for constant-depth circuits with parity gates, with implications for lower bounds. In: Symp. on Math. Foundations of Computer Science (MFCS). LIPIcs, vol. 117, pp. 78–17815 (2018)

[10] Alman, J., Chen, L.: Efficient construction of rigid matrices using an NP oracle. In: IEEE Symp. on Foundations of Computer Science (FOCS), pp. 1034–1055 (2019)

[11] Chen, L.: Non-deterministic quasi-polynomial time is average-case hard for ACC circuits. In: IEEE Symp. on Foundations of Computer Science (FOCS) (2019)

[12] Chen, L., Williams, R.R.: Stronger connections between circuit analysis and circuit lower bounds, via PCPs of proximity. In: IEEE Conf. on Computational Complexity (CCC), pp. 19–11943 (2019)

[13] Vyas, N., Williams, R.: Lower bounds against sparse symmetric functions of ACC circuits: Expanding the reach of #SAT algorithms. In: Symp. on Theoretical Aspects of Computer Science (STACS) (2020)

[14] Chen, L., Ren, H.: Strong average-case circuit lower bounds from nontrivial derandomization. In: ACM Symp. on the Theory of Computing (STOC) (2020)

[15] Viola, E.: New lower bounds for probabilistic degree and AC0 with parity gates. Electronic Coll. on Computational Complexity (ECCC) **27**, 15 (2020)

[16] Chen, L., Lyu, X., Williams, R.: Almost everywhere circuit lower bounds from non-trivial derandomization. In: IEEE Symp. on Foundations of Computer Science (FOCS) (2020)

[17] Bhangale, A., Harsha, P., Paradise, O., Tal, A.: Rigid matrices from rectangular PCPs. In: IEEE Symp. on Foundations of Computer Science (FOCS) (2020)

[18] Valiant, L.G.: Graph-theoretic arguments in low-level complexity. In: 6th Symposium on Mathematical Foundations of Computer Science. Lecture Notes in Computer Science, vol. 53, pp. 162–176 (1977)

[19] Servedio, R.A., Viola, E.: On a special case of rigidity. Available at http://www.ccs.neu.edu/home/viola/ (2012)

[20] Razborov, A.: Lower bounds on the dimension of schemes of bounded

depth in a complete basis containing the logical addition function. Akademiya Nauk SSSR. Matematicheskie Zametki **41**(4), 598–607 (1987). English translation in Mathematical Notes of the Academy of Sci. of the USSR, 41(4):333-338, 1987.

[21] Smolensky, R.: Algebraic methods in the theory of lower bounds for Boolean circuit complexity. In: 19th ACM Symp. on the Theory of Computing (STOC), pp. 77–82 (1987)

[22] Smolensky, R.: On representations by low-degree polynomials. In: 34th IEEE IEEE Symp. on Foundations of Computer Science (FOCS), pp. 130–138 (1993)

[23] Ben-Sasson, E., Viola, E.: Short PCPs with projection queries. In: Coll. on Automata, Languages and Programming (ICALP) (2014)

[24] Chen, L., Lyu, X.: Inverse-exponential correlation bounds and extremely rigid matrices from a new derandomized XOR lemma. In: ACM Symp. on the Theory of Computing (STOC) (2021)

[25] Goldreich, O., Nisan, N., Wigderson, A.: On Yao's XOR-lemma. In: Studies in Complexity and Cryptography. Lecture Notes in Computer Science, vol. 6650, pp. 273–301 (2011)

[26] Williams, R.: Improving exhaustive search implies superpolynomial lower bounds. In: 42nd ACM Symp. on the Theory of Computing (STOC), pp. 231–240 (2010)

[27] Chan, T.M., Williams, R.: Deterministic APSP, orthogonal vectors, and more: Quickly derandomizing Razborov-Smolensky. In: ACM-SIAM Symp. on Discrete Algorithms (SODA), pp. 1246–1255 (2016)

[28] Paradise, O.: Smooth and strong PCPs. In: ACM Innovations in Theoretical Computer Science conf. (ITCS), pp. 2–1241 (2020)

[29] Bhangale, A., Harsha, P., Paradise, O., Tal, A.: Rigid matrices from rectangular PCPs. Electron. Colloquium Comput. Complex. **TR20-075** (2020)

[30] Cook, S.A.: A hierarchy for nondeterministic time complexity. J. of Computer and System Sciences **7**(4), 343–353 (1973)

[31] Seiferas, J.I., Fischer, M.J., Meyer, A.R.: Separating nondeterministic time complexity classes. J. of the ACM **25**(1), 146–167 (1978)

[32] Zák, S.: A turing machine time hierarchy. Theoretical Computer Science **26**, 327–333 (1983)