Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 185 (2026)

Approximating commutative rank of matrix spaces in NC

Foram Lakhani* Partha Mukhopadhyay†

October 28, 2026

Abstract

Given any fixed constant $0 < \varepsilon < 1$ and a matrix space $\mathcal{B} = \langle B_1, \ldots, B_m \rangle \leq \mathbb{Q}^{n \times n}$, we give a deterministic NC^3 algorithm that outputs a matrix $A \in \mathcal{B}$ such that

$$\mathsf{rank}(A) \geq (1 - \varepsilon)\mathsf{crk}(\mathcal{B}),$$

where $\mathsf{crk}(\mathcal{B})$ denotes the maximum rank of a matrix in $\mathcal{B}$. This complements the recent breakthrough of Chatterjee, Ghosh, Gurjar, Raj, and Thierauf (ECCC, TR26-100), who gave an NC algorithm for computing the noncommutative rank of symbolic matrices. For commutative rank, Bläser, Jindal, and Pandey previously gave a deterministic polynomial-time approximation scheme (ToC, 2018).

Our algorithm follows a different route from the subspace-design approach of Chatterjee, Ghosh, Gurjar, Raj, and Thierauf. It has two main ingredients. First, using a polynomial-size 4-wise independent family together with operator scaling (Gurvits'04, Garg-Gurvits-Oliveira-Wigderson'20), we obtain a scalar matrix whose rank is an absolute constant fraction of $\mathsf{crk}(\mathcal{B})$. The operator scaling is used only in the analysis. Second, we boost this constant-factor approximation to a $(1 - \varepsilon)$-approximation by analyzing the associated Schur complements using Smith normal form over a discrete valuation ring. This analysis shows that logarithmically many boosting steps suffice to obtain the desired approximation.

1 Introduction

Given a set of matrices $B_1, \ldots, B_m \in \mathbb{Q}^{n \times n}$, the rank problem asks to compute the rank of the linear matrix $B(x) = \sum_{i=1}^{m} x_i B_i$ over the rational function field $\mathbb{Q}(x_1, \ldots, x_m)$. Equivalently, if $\mathcal{B} = \langle B_1, \ldots, B_m \rangle$ is the corresponding matrix space, the problem is to determine

$$\mathrm{crk}(\mathcal{B}) = \max\{\mathrm{rank}(B) : B \in \mathcal{B}\}.$$

This problem, often referred to as Edmonds' problem, has close connections with combinatorial optimization. The classical algebraic formulations of bipartite and general matching via the Edmonds and Tutte matrices are important special cases [Edm67, Tut47]. These connections were developed further by Lovász in the study of matching and matroid problems [Lov79].

*CSE, IIT Kanpur. Email: foramp23@iitk.ac.in.

†Chennai Mathematical Institute, Chennai, India. Email: partham@cmi.ac.in.

ISSN 1433-8092

Deciding whether the matrix space has full rank matrix or equivalently checking if $\det B(x) \not\equiv 0$ is usually referred to as symbolic determinant identity testing problem SDIT. Over a sufficiently large field there is a simple randomized polynomial-time algorithm to solve the SDIT problem using the Polynomial Identity Lemma [Sch80, Zip79, DL78]. However, finding an efficient deterministic algorithm seems elusive due to its direct connection with circuit lower bounds [KI04].

Over the last few years, a different picture emerged for the noncommutative version of the problem. Here the variables in $B(x)$ are treated as noncommuting variables, and the rank is taken over the free skew field. Independent breakthrough works of Garg, Gurvits, Oliveira, and Wigderson and of Ivanyos, Qiao, and Subrahmanyam gave deterministic polynomial-time algorithms for noncommutative rank [GGOW20, IQS18]. The former proceeds through operator scaling, while the latter gives an algebraic algorithm based on shrunk subspaces and the concept of blow-up. Another algorithm using convex optimization was developed by Hamada and Hirai [HH21]. Building on the ideas of [IQS18], Chatterjee and Mukhopadhyay later gave another polynomial-time algorithm in which the rank-increment step is reduced to polynomial identity testing for noncommutative algebraic branching programs [CM23]. These results exhibit a sharp contrast between the commutative and noncommutative rank problems.

For commutative rank, a beautiful approximation result was obtained by Bläser, Jindal, and Pandey [BJP18]. They gave a deterministic polynomial-time approximation scheme: for every fixed $\varepsilon > 0$, one can find $A \in \mathcal{B}$ satisfying

$$\operatorname{rank}(A) \geq (1-\varepsilon)\operatorname{crk}(\mathcal{B}).$$

This improves on the numerical 1/2-approximation to commutative rank that follows from deterministic algorithms for noncommutative rank and the well-known inequalities [FR04]

$$\operatorname{crk}(\mathcal{B}) \leq \mathsf{ncrank}(\mathcal{B}) \leq 2\operatorname{crk}(\mathcal{B}).$$

Subsequently, Bhargava, Bläser, Jindal and Pandey extended the related ideas to higher-degree algebraic rank problems [BBJP19]. Thus, although exact deterministic computation of commutative rank remains open, it admits efficient deterministic approximation to arbitrary constant accuracy. A recent work of Chakraborty, Datta, Kusre, Mukhopadhyay, and Sinhababu gives a catalytic logspace algorithm (CL) with a polynomial-time bound (CLP) for $(1-\varepsilon)$ rank approximation [CDK+26].

Very recently, in a breakthrough, Chatterjee, Ghosh, Gurjar, Raj, and Thierauf proved that bipartite matching is in NC, and extended their method to give an NC algorithm for computing noncommutative rank [CGG+26]. Their approach is inspired by the explicit subspace designs [GK16], and also by the beautiful work of Brakensiek, Chen, Dhar, and Zhang [BCDZ26]. This result naturally raises the question of parallel complexity of the approximation question for commutative rank: can the polynomial-time approximation scheme for $\operatorname{crk}(\mathcal{B})$ also be parallelized? Our main result answers this question affirmatively over $\mathbb{Q}$. Our main result is the following.

Theorem 1.1. *For every fixed $0 < \varepsilon < 1$, and matrix space $\mathcal{B} = \langle B_1, \ldots, B_m \rangle \leq \mathbb{Q}^{n\times n}$, there is an NC^3 algorithm that computes a matrix $A \in \mathcal{B}$ such that $\operatorname{rank}(A) \geq (1-\varepsilon)\operatorname{crk}(\mathcal{B})$.*

The proof technique for Theorem 1.1 is independent of the proof technique for the NC algorithm for noncommutative rank [CGG+26].

1.1 Proof idea

The proof of Theorem 1.1 is broadly divided into two parts. In the first part, we compute a matrix A, in NC^2, from the given matrix space $\mathcal{B}$ such that $\operatorname{rank}(A) \geq c_0 \operatorname{crk}(\mathcal{B})$ for an absolute constant $0 < c_0 < 1$. Although computing the noncommutative rank gives a numerical 1/2-approximation to $\operatorname{crk}(\mathcal{B})$, this does not by itself produce a matrix $A \in \mathcal{B}$ with the corresponding rank guarantee. Our first step requires such a matrix, which we obtain using the rounding argument below. In the second part, we use a *boosting argument* to find a matrix from $\mathcal{B}$, in NC^3, whose rank is at least $(1-\varepsilon)\operatorname{crk}(\mathcal{B})$.

The argument used for the first part is analytical and inspired by the connection of quantum information theory with operator scaling [Gur04], and moment methods [Ber97]. For the analysis, it suffices first to consider a matrix space containing a nonsingular matrix. For a general matrix space of commutative rank c, we apply the same argument to the space obtained by restricting every input matrix to the rows and columns of a nonzero $c \times c$ minor. This restriction is used only in the analysis.

If a matrix space $\mathcal{D} = \langle D_1, \ldots, D_m \rangle \subseteq \mathbb{Q}^{n \times n}$ contains a nonsingular matrix, then it is also dimension non-decreasing. In other words, for every subspace U

$$\dim\langle \bigcup_{i=1}^{m} D_i(U) \rangle \geq \dim U.$$

Now we use the connection with operator scaling. Given any $0 < \eta < 1$, there exist left and right invertible matrices L and R in $\mathrm{GL}_n(\mathbb{C})$ such that the set of matrices $\{A_1, \ldots, A_m\}$ where $A_i = LD_iR$ is approximately doubly stochastic. More precisely, define $M(A) = \sum_i A_i A_i^*$ and $N(A) = \sum_i A_i^* A_i$ and

$$\mathsf{ds}(A) = \|M(A) - I\|_F^2 + \|N(A) - I\|_F^2 \leq \eta.$$

This is a measure of the distance from doubly-stochasticity. The scaling matrices are used only for the analysis. Now consider

$$S = \sum_i \epsilon_i A_i,$$

where the sign vectors $\epsilon = (\epsilon_1, \ldots, \epsilon_m) \in \{-1, +1\}^m$ are chosen from a 4-wise independent family $\mathcal{H}_m \subseteq \{-1, +1\}^m$, that can be easily shown to be constructible in NC^2. Set,

$$X = \operatorname{Tr}(S^*S), \qquad Y = \operatorname{Tr}((S^*S)^2).$$

The quantities X and Y are the sum of the square and fourth powers of the singular values of S. More precisely,

$$X = \sum_{j=1}^{\rho} \sigma_j^2 \qquad Y = \sum_{j=1}^{\rho} \sigma_j^4.$$

A routine application of Cauchy–Schwarz gives

$$X^2 \leq \rho Y = \operatorname{rank}(S) Y,$$

implying that $\operatorname{rank}(S) \geq \frac{X^2}{Y}$ [1]. Choosing the parameter η appropriately in the doubly stochastic distance measure and the 4-wise independence of the signs, one can control $\mathbb{E}[X]$

[1] Define $X^2/Y = 0$ on the event $Y = 0$, where necessarily $S = 0$ and $X = 0$.

and $\mathbb{E}[Y]$. In particular, $\mathbb{E}[X] = \Theta(n)$ and $\mathbb{E}[Y] = O(n)$. Using another application of Cauchy–Schwarz inequality, we prove that

$$\mathbb{E}[\operatorname{rank}(S)] \geq \mathbb{E}\left[\frac{X^2}{Y}\right] \geq \frac{\mathbb{E}^2[X]}{\mathbb{E}[Y]} \geq c_0 n,$$

for an absolute constant $0 < c_0 < 1$. This shows the existence of a sign vector $\epsilon \in \mathcal{H}_m$ such that

$$\operatorname{rank}(S) = \operatorname{rank}\left(\sum_i \epsilon_i A_i\right) \geq c_0 n.$$

Since L and R are invertible

$$\operatorname{rank}\left(\sum_i \epsilon_i D_i\right) = \operatorname{rank}\left(\sum_i \epsilon_i A_i\right).$$

Such an $\epsilon \in \mathcal{H}_m$ can be found by trying all of them in parallel and choosing the one with the maximum rank. Clearly, this can be performed in NC^2[Csa76, Mul87].

The second part repeatedly improves the rank of the current matrix $A \in \mathcal{B}$. To analyze this improvement, we associate with A a matrix J_k, constructed from the first k coefficients of a suitable Schur complement. The parameter k is fixed depending on ε. As we explain below, preserving a constant fraction of the generic rank of J_k under suitable substitution allows us to improve the rank of A. However, the entries of J_k are polynomials of degree at most k, so the linear rounding result from the first part does not apply directly.

We therefore first explain how to extend that result to matrices whose entries are homogeneous polynomials of a fixed degree k. We will then apply this extension to a homogenized version of J_k. The key idea is *complete polarization*, which replaces a homogeneous degree k matrix $M(x)$ by a matrix linear in each of k independent variable groups $X^{(1)}, \ldots, X^{(k)}$. The polarization method is standard; see [Lan12, §2.6.4] and, for further background, [Pro07, Chapter 3, §2].

On setting all k groups equal, the polarized matrix becomes $k!M(x)$. As a result, its rank is at least the rank of $M(x)$. Applying the linear rounding result successively to the k groups gives sign vectors

$$\epsilon^{(1)}, \ldots, \epsilon^{(k)} \in \mathcal{H}_m$$

such that the evaluated polarized matrix has rank at least c_0^k times the rank of M.

By the polarization identity, this matrix is a signed sum of the 2^k ordinary matrix evaluations

$$M\left(\sum_{j \in S} \epsilon^{(j)}\right), \quad S \subseteq [k].$$

Then the sub-additivity property of rank implies that one of these evaluations has rank at least $\beta_k := \frac{c_0^k}{2^k}$ times the true rank. This gives the universal set

$$\mathcal{Z}_{k,m} := \left\{\sum_{j \in S} \epsilon^{(j)} : S \subseteq [k],\ \epsilon^{(1)}, \ldots, \epsilon^{(k)} \in \mathcal{H}_m\right\}.$$

For fixed k,

$$|\mathcal{Z}_{k,m}| \leq 2^k |\mathcal{H}_m|^k.$$

So, this family has a polynomial size and can be enumerated in parallel in NC^2.

Now we provide a sketch of the rank boosting step. In particular, we will describe the matrix J_k and explain how the polarization idea described above leads to rank improvement. Let,

$$\mathcal{B} = \langle B_1, \ldots, B_m \rangle \leq \mathbb{Q}^{n\times n}, \qquad B(x) = \sum_i x_i B_i, \qquad c = \operatorname{crk}(\mathcal{B}).$$

Suppose that the current matrix $A \in \mathcal{B}$ has rank r. After invertible row and column transformations we may write

$$A = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}, \qquad B(x) = \begin{pmatrix} X(x) & Q(x) \\ P(x) & Z(x) \end{pmatrix}.$$

These transformations are also used only in the analysis. We introduce a new formal variable t and consider the matrix $A + tB(x)$. Its upper left block is invertible over the power series ring, and its Schur complement [2] is $tH(t,x)$ where

$$H(t,x) = Z - tP(I_r + tX)^{-1}Q = h_0 + th_1 + t^2h_2 + \cdots.$$

The rank of $H(t,x)$ over $\mathbb{Q}(x)((t))$ measures the possible increase in rank from A along the symbolic direction $B(x)$. However, the rank increase may not be visible in the constant term $H(0,x) = Z(x)$ alone. We therefore retain the first k coefficients and arrange them into a single matrix

$$J_k(A, B(x)) := \begin{pmatrix} h_0 & 0 & \cdots & 0 \\ h_1 & h_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ h_{k-1} & \cdots & h_1 & h_0 \end{pmatrix}.$$

Equivalently, J_k is the matrix of multiplication by $H(t,x)$ modulo t^k. The intuition behind the construction of J_k comes from classical ideas. See for example [VDV79] and [Sch26, Appendix E]. A related construction also appears in Oki's work [Oki20, Oki23]. The matrix J_k is used only to analyze the progress of the algorithm in incresing rank. The algorithm does not construct the matrix J_k.

Let $K = \mathbb{Q}(x_1, \ldots, x_m)$. Using the theory of Smith normal form over the discrete valuation ring $K[[t]]$ ([AM69, Chs. 5,9–10]; [Lan02, Chapter §XII]; [Hun74, Chapter §VII]), and some additional simple ideas, we show that

$$\operatorname{rank}_K J_k \geq k(c-r) - r = kc - (k+1)r.$$

This suggests defining

$$\Delta_k(A) := kc - (k+1)\operatorname{rank} A.$$

The condition $\Delta_k(A) < 0$ is exactly

$$\operatorname{rank} A > \frac{k}{k+1}c.$$

To apply polarization to J_k, first introduce one extra variable z and define

$$\widehat{J}_k(z,x) := \begin{pmatrix} z^{k-1}h_0 & 0 & \cdots & 0 \\ z^{k-2}h_1 & z^{k-1}h_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ h_{k-1} & \cdots & z^{k-2}h_1 & z^{k-1}h_0 \end{pmatrix}.$$

[2] Schur complement for block-matrices is a standard concept (see, for example, [Bha07, p. 20]). It also appears in a rank-increment argument in [CM23].

Since h_j has degree $j+1$, every entry of $\widehat{J}_k$ is homogeneous of degree k in $(z, x_1, \ldots, x_m)$. Moreover, $\widehat{J}_k(1, x) = J_k(A, B(x))$, so

$$\operatorname{rank}_{\mathbb{Q}(z,x)} \widehat{J}_k(z, x) \geq \operatorname{rank}_K J_k(A, B(x)).$$

Applying the polarization argument to $\widehat{J}_k$ gives some $\widetilde{b} = (b_0, b_1, \ldots, b_m) \in \mathcal{Z}_{k,m+1}$ such that

$$\operatorname{rank} \widehat{J}_k(\widetilde{b}) \geq \beta_k \operatorname{rank}_K J_k(A, B(x)).$$

For every $\widetilde{b} \in \mathcal{Z}_{k,m+1}$, define

$$b(\widetilde{b}) := \begin{cases} (b_1/b_0, \ldots, b_m/b_0), & b_0 \neq 0, \\ (b_1, \ldots, b_m), & b_0 = 0, \end{cases} \qquad \mathcal{D}_{k,m} := \{b(\widetilde{b}) : \widetilde{b} \in \mathcal{Z}_{k,m+1}\}.$$

If $b_0 \neq 0$, homogeneity gives

$$\widehat{J}_k(\widetilde{b}) = b_0^k J_k\big(A, B(b(\widetilde{b}))\big),$$

so these two matrices have the same rank. If $b_0 = 0$, the only potentially nonzero block of $\widehat{J}_k(\widetilde{b})$ is $h_{k-1}(b_1, \ldots, b_m)$ in the bottom-left position. This block also appears in $J_k\big(A, B(b(\widetilde{b}))\big)$. Thus, in either case, some $b \in \mathcal{D}_{k,m}$, with $C = B(b)$, satisfies

$$\operatorname{rank} J_k(A, C) \geq \beta_k \operatorname{rank}_K J_k(A, B(x)) \geq \beta_k \Delta_k(A).$$

Let

$$s := \operatorname{rank}_{\mathbb{Q}(t)}(A + tC) - r.$$

The same Smith-form calculation gives $\operatorname{rank} J_k(A, C) \leq ks$, so

$$s \geq \frac{\beta_k}{k} \Delta_k(A).$$

A nonzero minor of maximum size in $A+tC$ has degree at most n. Hence some $\lambda \in \{0, \ldots, n\}$ realizes its generic rank, giving a matrix $A' = A + \lambda C$ with

$$\Delta_k(A') \leq \left(1 - \frac{k+1}{k}\beta_k\right) \Delta_k(A).$$

The family $\mathcal{D}_{k,m}$ depends only on k and m. In each round, the algorithm therefore tries all $b \in \mathcal{D}_{k,m}$ and $\lambda \in \{0, \ldots, n\}$ in parallel and selects a candidate $A + \lambda B(b)$ of maximum rank. As long as the deficit is positive, it decreases by a fixed factor depending only on k. Since it is an integer and initially at most kn, after $O_k(\log(n+1))$ rounds it is non-positive. The rank never decreases, because $\lambda = 0$ is always among the candidates. Thus

$$\operatorname{rank} A \geq \frac{k}{k+1} c.$$

Choosing $k = \lceil 1/\varepsilon - 1 \rceil$ gives $\operatorname{rank} A \geq (1-\varepsilon)c$. Each round consists of polynomially many parallel rank computations in NC^2, on rational matrices of polynomial bit length. With $O_k(\log(n+1))$ sequential rounds, the overall algorithm lies in NC^3.

AI methodology: We used GPT-5.6 to assist with several brainstorming sessions conducted over months. In particular, GPT-5.6 contributed the core idea of rank-boosting analysis using Smith normal form over discrete valuation rings during discussions of the papers by Oki [Oki20, Oki23]. It also provided clear explanations of [Gur04, Oki20, Oki23], which were instrumental to this work. The human authors have written the paper, checked the proofs, and take the full responsibility of the contents. After verification of the proofs by authors, we also use GPT-6.1 Sol to perform another round of verification. We have also used GPT-6.1 Sol to improve language, presentation, and mathematical typeset in places.

Acknowledgements We thank the organizers of the workshop *Circuits, Communication and Proofs* at ICTS, Bengaluru, for hosting the authors, where a part of this work was completed.

Organization. The rest of the paper is organized as follows. In Section 2, we provide the necessary algebraic background and also present a construction of t-wise independent family. Section 3 contains the construction of a matrix from the given matrix space whose rank is a constant fraction of the maximum possible rank. In Section 4, we prove the rank bound for the matrix J_k (which is used in Section 5). Finally, the rank boosting argument is presented in Section 5 which completes the proof of Theorem 1.1.

2 Preliminaries

In this section, we provide the necessary algebraic background and include a construction of t-wise independent distribution.

2.1 Algebraic background

We include some algebraic background related to Discrete Valuation Rings, Smith normal form which we use later in our algorithm design and analysis.

Throughout this paper, fix

$$K = \mathbb{Q}(x_1, \ldots, x_m).$$

2.1.1 Formal power series, Laurent series, and the t-adic valuation

Details can be found in [Ser79, Chapter I], [AM69, Chapter 9], and in [Lan02, Chapter §XII]. The ring

$$K[[t]] = \left\{a_0 + a_1 t + a_2 t^2 + \cdots : a_i \in K\right\}$$

is the ring of formal power series in t over the field K. Its fraction field is

$$K((t)) = \left\{\sum_{j \geq N} a_j t^j : N \in \mathbb{Z}\right\},$$

the field of formal Laurent series. For a nonzero Laurent series

$$f(t) = a_\nu t^\nu + a_{\nu+1} t^{\nu+1} + \cdots, \qquad a_\nu \neq 0,$$

define its t-adic valuation by $\operatorname{ord}_t f = \nu$. Thus every nonzero $f \in K[[t]]$ has a unique factorization

$$f(t) = t^\nu u(t), \qquad \nu \geq 0, \qquad u(0) \neq 0.$$

The factor $u(t)$ is a unit of $K[[t]]$; conversely a power series is a unit exactly when its constant term is nonzero. The maximal ideal is (t), and every nonzero ideal is (t^ν) for a unique ν. Hence $K[[t]]$ is a *discrete valuation ring* (DVR).

An elementary comparison is

$$\operatorname{ord}_t f \le \deg_t f \qquad (0 \neq f \in K[t]). \tag{1}$$

2.1.2 Smith normal form

Since a discrete valuation ring (DVR) is a principal ideal domain (PID), a matrix

$$G(t) \in K[[t]]^{p\times q}$$

of rank d over the fraction field $K((t))$ admits a Smith normal form. Thus there exist

$$U(t) \in \mathrm{GL}_p(K[[t]]), \qquad V(t) \in \mathrm{GL}_q(K[[t]])$$

such that

$$U(t)G(t)V(t) = \operatorname{diag}(t^{\nu_1}, \dots, t^{\nu_d}, 0, \dots, 0),$$

where the right-hand side is a $p \times q$ diagonal matrix and

$$0 \le \nu_1 \le \cdots \le \nu_d.$$

Each nonzero invariant factor is a unit times a power of t. Hence the unit factors are absorbed into $U(t)$ and $V(t)$. We call the integers ν_i the Smith invariant exponents. For Smith normal form over a principal ideal domain, see [Hun74, Proposition 2.11, p. 339].

For every $1 \le j \le d$,

$$\min_{\substack{|I|=|J|=j \\ \det G(t)[I,J]\neq 0}} \operatorname{ord}_t \det G(t)[I,J] = \nu_1 + \cdots + \nu_j.$$

This follows because invertible row and column transformations preserve the ideal generated by the $j\times j$ minors. In particular, some nonzero $d\times d$ minor has valuation exactly $\nu_1+\cdots+\nu_d$. This is explained in self-contained manner in Claim 4.4.

2.1.3 Truncation modulo t^k

There is a natural identification

$$K[[t]]/(t^k) \cong K[t]/(t^k).$$

Let

$$G(t) = \sum_{i\ge 0} t^i G_i \in K[[t]]^{p\times q}, \qquad G_i \in K^{p\times q},$$

where G_i is obtained by taking the coefficient of t^i in each entry of $G(t)$. Modulo t^k, only the matrices $G_0, \dots, G_{k-1}$ contribute. Then multiplication by $G(t)$ modulo t^k is a K-linear map

$$\big(K[t]/(t^k)\big)^q \longrightarrow \big(K[t]/(t^k)\big)^p.$$

Writing a vector as

$$v(t) = v_0 + tv_1 + \cdots + t^{k-1}v_{k-1},$$

the coefficients of $G(t)v(t) \bmod t^k$ are obtained by convolution. Hence, in the coefficient basis $1, t, \ldots, t^{k-1}$, the multiplication map is represented by the lower block matrix

$$\Omega_k(G) = \begin{pmatrix} G_0 & 0 & \cdots & 0 \\ G_1 & G_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ G_{k-1} & \cdots & G_1 & G_0 \end{pmatrix}. \tag{2}$$

See also [Oki20, Sec. 3.1].

2.2 Construction of t-wise independent distribution

In this section we give a self-contained construction of t-wise independent distribution over $\{-1, +1\}^m$ within NC^2. This construction is well-known [ABI86]. In our application, we use it for $t = 4$.

2.2.1 Construction of extension fields

Let $m \geq 2$, and set

$$r := \lceil \log_2 m \rceil, \qquad q := 2^r.$$

Thus $m \leq q < 2m$. We represent

$$\mathbb{F}_q \cong \mathbb{F}_2[z]/(f(z)),$$

where $f \in \mathbb{F}_2[z]$ is irreducible of degree r. All complexity bounds below are measured in m.

Lemma 2.1. *An irreducible polynomial $f \in \mathbb{F}_2[z]$ of degree r can be found deterministically in uniform NC^2. Arithmetic in $\mathbb{F}_q = \mathbb{F}_2[z]/(f)$ can also be performed in uniform NC^2.*

Proof. Enumerate all $2^r = O(m)$ monic degree-r polynomials g. For each g, test in parallel whether it is divisible by any monic polynomial h with $1 \leq \deg h \leq \lfloor r/2 \rfloor$. There are fewer than $2^{r/2+1}$ such polynomials h. Since g is reducible exactly when such a divisor exists, these tests identify all irreducible candidates.

Polynomial long division over $\mathbb{F}_2$ has depth $O(r)$., Combining the divisibility tests and selecting the lexicographically first irreducible candidate gives a uniform polynomial-size circuit of depth $O(\log^2 m)$. Field elements are represented by polynomials of degree less than r. Addition and multiplication modulo f lie in NC^2. For a nonzero element a, its inverse can be found by testing $ab = 1$ for all $q = O(m)$ elements b in parallel. □

2.2.2 The trace function

We next recall the elementary properties of the trace function

$$\mathrm{Tr} := \mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_2} : \mathbb{F}_q \longrightarrow \mathbb{F}_2.$$

It is given explicitly by

$$\mathrm{Tr}(a) = a + a^2 + a^{2^2} + \cdots + a^{2^{r-1}}.$$

Lemma 2.2. *The map $\mathrm{Tr} : \mathbb{F}_{2^r} \to \mathbb{F}_2$ has the following properties.*

1. Tr *is $\mathbb{F}_2$-linear.*

2. Tr *is not the zero map.*

3. *Consequently,* Tr *is surjective and* $|\ker(\mathrm{Tr})| = 2^{r-1} = \frac{q}{2}$. *In particular, if* U *is uniformly distributed over* $\mathbb{F}_q$*, then* $\mathrm{Tr}(U)$ *is a uniformly distributed bit.*

Proof. For $a, b \in \mathbb{F}_q$, the Frobenius identity in characteristic two gives

$$(a+b)^{2^j} = a^{2^j} + b^{2^j}.$$

It follows immediately that

$$\mathrm{Tr}(a+b) = \mathrm{Tr}(a) + \mathrm{Tr}(b),$$

and scalar linearity over $\mathbb{F}_2$ is immediate. Thus Tr is $\mathbb{F}_2$-linear.

We also note that $\mathrm{Tr}(a) \in \mathbb{F}_2$. Indeed, using $a^{2^r} = a$ for every $a \in \mathbb{F}_q$,

$$\begin{aligned}\mathrm{Tr}(a)^2 &= a^2 + a^{2^2} + \cdots + a^{2^r} \\ &= a + a^2 + \cdots + a^{2^{r-1}} = \mathrm{Tr}(a).\end{aligned}$$

The only elements of $\mathbb{F}_q$ satisfying $x^2 = x$ are 0 and 1.

It remains to show that Tr is nonzero. Consider the polynomial

$$T(X) := X + X^2 + X^{2^2} + \cdots + X^{2^{r-1}} \in \mathbb{F}_q[X].$$

Since the degree of $T(X)$ is 2^{r-1}, it can not vanish on $\mathbb{F}_q$. Hence Tr is nonzero.

Then, by the rank-nullity theorem,

$$\dim_{\mathbb{F}_2}(\ker(\mathrm{Tr})) = r - 1, \quad |\ker(\mathrm{Tr})| = 2^{r-1} = \frac{q}{2}.$$

Hence the trace of a uniformly random element of $\mathbb{F}_q$ is uniformly distributed in $\mathbb{F}_2$. □

2.2.3 t-wise independence from polynomials

Choose pairwise distinct elements $\alpha_1, \ldots, \alpha_m \in \mathbb{F}_q$; this is possible since $q \geq m$.

For $a = (a_0, \ldots, a_{t-1}) \in \mathbb{F}_q^t$, define the polynomial

$$p_a(Z) := a_0 + a_1 Z + \cdots + a_{t-1} Z^{t-1}.$$

Using this, we get the sign bit $\epsilon_i(a) := (-1)^{\mathrm{Tr}(p_a(\alpha_i))}$. Define,

$$\mathcal{H}_{m,t} := \left\{ \big(\epsilon_1(a), \ldots, \epsilon_m(a)\big) : a \in \mathbb{F}_q^t \right\} \subseteq \{-1, +1\}^m.$$

The following lemma is well-known and follows from the standard Vandermonde matrix based argument.

Lemma 2.3. *Let* a *be uniform in* $\mathbb{F}_q^t$*. Then the field-valued random variables*

$$p_a(\alpha_1), \ldots, p_a(\alpha_m)$$

are t*-wise independent and uniformly distributed over* $\mathbb{F}_q$*.*

Now we argue that the trace map converts these field-valued variables to t-wise independent random variables over $\{-1, +1\}^m$. The following standard construction is also described in [DPT24, Section 3.4, Claim 3.17], where polynomial evaluations over an extension field are followed by the field trace to $\mathbb{F}_2$.

Theorem 2.4. *Let a be chosen uniformly from $\mathbb{F}_q^t$. Then*

$$\epsilon_i(a) = (-1)^{\mathrm{Tr}(p_a(\alpha_i))}, \qquad i \in [m],$$

are t-wise independent unbiased Rademacher random variables. Equivalently, for every set of distinct indices $i_1, \ldots, i_s$, where $s \leq t$, and every $\delta_1, \ldots, \delta_s \in \{-1, +1\}$,

$$\Pr_a \left[\epsilon_{i_1}(a) = \delta_1, \ldots, \epsilon_{i_s}(a) = \delta_s\right] = 2^{-s}.$$

Proof. By Lemma 2.3,

$$\big(p_a(\alpha_{i_1}), \ldots, p_a(\alpha_{i_s})\big)$$

is uniform over $\mathbb{F}_q^s$.

Fix bits $b_1, \ldots, b_s \in \mathbb{F}_2$. By Lemma 2.2, each equation

$$\mathrm{Tr}(y_j) = b_j$$

has exactly $q/2$ solutions $y_j \in \mathbb{F}_q$. Therefore

$$\begin{aligned}\Pr\left[\mathrm{Tr}(p_a(\alpha_{i_1})) = b_1, \ldots, \mathrm{Tr}(p_a(\alpha_{i_s})) = b_s\right] \\ = \frac{(q/2)^s}{q^s} = 2^{-s}.\end{aligned}$$

□

Summarizing, we have the following corollary.

Corollary 2.5. *For every fixed constant t, there is a uniformly NC^2-constructible sample space of t-wise independent unbiased signs $(\epsilon_1, \ldots, \epsilon_m) \in \{-1, +1\}^m$ having support at most $q^t < (2m)^t = m^{O(t)}$.*

For the remainder of the paper we abbreviate the 4-wise-independent family by $\mathcal{H}_m := \mathcal{H}_{m,4}$.

3 Constant fraction rank approximation

The goal of this section is to output a matrix from the given matrix space such that the rank is at least a constant fraction of the maximum rank possible. Moreover, the computation is performed in NC^2.

Let the input matrix be

$$B(x) = \sum_{i=1}^{m} x_i B_i$$

defined over $\mathbb{Q}$. The associated matrix space is $\mathcal{B} = \langle B_1, \ldots, B_m \rangle$. The commutative rank of $\mathcal{B}$ denoted by $\mathrm{crk}(\mathcal{B})$ is same as the rank of $B(x)$ over the function field $\mathbb{Q}(x)$.

Definition 3.1. A tuple $D = (D_1, \ldots, D_m)$ of $n \times n$ complex matrices is *dimension non-decreasing* if for every subspace $U \leq \mathbb{C}^n$,

$$\dim\langle \bigcup_{i=1}^{m} D_i(U) \rangle \geq \dim U.$$

For a tuple $A = (A_1, \ldots, A_m)$ define

$$M(A) = \sum_i A_i A_i^*, \qquad N(A) = \sum_i A_i^* A_i,$$

and its distance from doubly stochasticity is given by

$$\mathrm{ds}(A) = \|M(A) - I\|_F^2 + \|N(A) - I\|_F^2.$$

Theorem 3.2. *[Gur04, Theorem 4.6] A square complex tuple of matrices is scalable if and only if it is dimension non-decreasing. Equivalently, if D is dimension non-decreasing, then for every $\eta > 0$ there exist invertible matrices L, R such that the tuple $A_i = LD_iR$ satisfies $\mathrm{ds}(A) \leq \eta$.*

The result is also implicit in [GGOW20]. Nevertheless, we include some facts underlying Theorem 3.2 in Section A mainly for completeness. We use this theorem only in the analysis. The algorithm does not compute the scaling matrices.

Since L, R are invertible matrices, we have the following rank identity

$$\operatorname{rank}\left(\sum_i \epsilon_i L D_i R\right) = \operatorname{rank}\left(\sum_i \epsilon_i D_i\right).$$

3.1 Moment method

For this sub-section, we assume that the matrix space has full rank.

Lemma 3.3. *Let $A_1, \ldots, A_m \in \mathbb{C}^{n\times n}$, and set $M = \sum_i A_i A_i^*$, $N = \sum_i A_i^* A_i$. Assume*

$$\|M - I_n\|_F^2 + \|N - I_n\|_F^2 \leq \frac{1}{16n}. \tag{3}$$

Let $\epsilon = (\epsilon_1, \ldots, \epsilon_m)$ be drawn from a 4-wise independent distribution of unbiased signs and $S = \sum_i \epsilon_i A_i$. Then with positive probability $\operatorname{rank} S \geq c_0 n$, for an absolute constant $0 < c_0 < 1$. Consequently, for at least one sign vector in the family, we get that the rank is at least $c_0 n$.

Proof. Define two *different* nonnegative random variables

$$X := \operatorname{Tr}(S^*S), \qquad Y := \operatorname{Tr}\big((S^*S)^2\big). \tag{4}$$

Let $\tau = \operatorname{Tr} M = \operatorname{Tr} N = \sum_i \|A_i\|_F^2$. Since $\mathbb{E}_{i,j}[\epsilon_i \epsilon_j] = 0$ unless $i = j$, it is easy to observe that $\mathbb{E}[X] = \tau$. Next, we prove a bound on $\mathbb{E}[Y]$.

Claim 3.4. $\mathbb{E}[Y] \leq \operatorname{Tr}(N^2) + 2\operatorname{Tr}(M^2)$

Proof. By expansion,

$$\mathbb{E}[Y] = \sum_{i,j,k,\ell} \mathbb{E}[\epsilon_i \epsilon_j \epsilon_k \epsilon_\ell] \operatorname{Tr}(A_i^* A_j A_k^* A_\ell).$$

As above, every surviving tuple belongs to one of the three pairing patterns

$$i = j,\ k = \ell; \qquad i = k,\ j = \ell; \qquad i = \ell,\ j = k,$$

and overcounting the all-equal case is enough for an upper bound. More precisely, the three pairing patterns count each all-equal tuple three times rather than once. Thus their combined contribution exceeds the exact expectation by

$$2\sum_i \operatorname{Tr}\big((A_i^*A_i)^2\big) \geq 0.$$

For the first pairing pattern, the contribution is

$$\sum_{i,k} \operatorname{Tr}(A_i^*A_iA_k^*A_k) = \operatorname{Tr}\left(\left(\sum_i A_i^*A_i\right)^2\right) = \operatorname{Tr}(N^2).$$

For the second pairing pattern, using the triangle inequality,

$$\begin{aligned}\left|\sum_{i,j} \operatorname{Tr}(A_i^*A_jA_i^*A_j)\right| &\leq \sum_{i,j} \left|\operatorname{Tr}((A_i^*A_j)^2)\right| \\ &\leq \sum_{i,j} \|A_i^*A_j\|_F^2 \\ &= \operatorname{Tr}(M^2).\end{aligned}$$

Here

$$|\operatorname{Tr}(B^2)| = |\langle B^*, B\rangle_F| \leq \|B^*\|_F\|B\|_F = \|B\|_F^2.$$

For the third pairing pattern,

$$\sum_{i,j} \operatorname{Tr}(A_i^*A_jA_j^*A_i) = \sum_{i,j} \|A_j^*A_i\|_F^2 = \operatorname{Tr}(M^2).$$

Therefore

$$\mathbb{E}[Y] \leq \operatorname{Tr}(N^2) + 2\operatorname{Tr}(M^2),$$

as claimed. □

By (3),

$$\|M - I_n\|_F \leq \frac{1}{4\sqrt{n}}, \qquad \|N - I_n\|_F \leq \frac{1}{4\sqrt{n}}.$$

Since

$$|\tau - n| = |\operatorname{Tr}(M - I_n)| \leq \sqrt{n}\|M - I_n\|_F \leq \frac{1}{4},$$

we have

$$\tau \geq \frac{3n}{4}. \tag{5}$$

Moreover

$$\begin{aligned}\operatorname{Tr}(M^2) &\leq \|M\|_F^2 \\ &\leq \left(1/4 + \sqrt{n}\right)^2 \\ &\leq \frac{25n}{16}.\end{aligned}$$

The same estimate holds for $\operatorname{Tr}(N^2)$. Then using Claim 3.4,

$$\mathbb{E}[Y] \leq \frac{75n}{16}. \tag{6}$$

If $\sigma_1, \ldots, \sigma_\rho$ are the nonzero singular values of S, then

$$X = \sum_{j=1}^{\rho} \sigma_j^2, \qquad Y = \sum_{j=1}^{\rho} \sigma_j^4.$$

Cauchy–Schwarz gives

$$X^2 \le \rho Y = \operatorname{rank}(S)\, Y. \tag{7}$$

On the event $Y = 0$, necessarily $S = 0$ and $X = 0$. We define $X/\sqrt{Y}$ and X^2/Y to be zero on this event. Observe again using Cauchy–Schwarz that

$$(\mathbb{E}[X])^2 = \mathbb{E}\left[\frac{X}{\sqrt{Y}}\sqrt{Y}\right]^2 \le \mathbb{E}\left[\frac{X^2}{Y}\right]\ \mathbb{E}[Y].$$

So,

$$\mathbb{E}[\operatorname{rank}(S)] \ge \mathbb{E}\left[\frac{X^2}{Y}\right] \ge \frac{\mathbb{E}^2[X]}{\mathbb{E}[Y]} \ge \frac{3n}{25}.$$

Hence we complete the proof with $c_0 = \frac{3}{25}$. □

3.2 The NC^2 algorithm for finding the constant rank approximation matrix

We apply the result obtained in Section 3.1 to the given matrix space whose crk is r. The main theorem is the following.

Theorem 3.5. *Let*

$$B(x) = \sum_{i=1}^{m} x_i B_i \in \mathbb{Q}[x]^{n\times n}$$

be the input matrix whose rank is r over $\mathbb{Q}(x)$. Then for some $\epsilon \in \mathcal{H}_m$, we get that $\operatorname{rank} B(\varepsilon) \ge c_0 r$*, where $\mathcal{H}_m$ be a 4-wise independent distribution over $\{-1,+1\}^m$ as obtained in Corollary 2.5, and c_0 is the constant promised by Lemma 3.3.*

Proof. Consider an $r \times r$ sub-matrix of $B(x)$ which is of full rank over $\mathbb{Q}(x)$. Let $D(x) = \sum_{i=1}^{m} x_i D_i$ be the corresponding sub-matrix. There exists $a \in \mathbb{Q}^m$ such that $D(a)$ is invertible.

We claim that the tuple $(D_1, \ldots, D_m)$ is dimension non-decreasing. Indeed, for every subspace $U \le \mathbb{C}^r$, $D(a)U \subseteq \sum_i D_i(U)$. Since $D(a)$ is invertible,

$$\dim \sum_i D_i(U) \ge \dim D(a)U = \dim U.$$

By Theorem 3.2, for $\eta = 1/16r$ there are invertible complex matrices P, Q such that

$$A_i = P D_i Q$$

satisfy (3). Lemma 3.3 gives a sign vector $\epsilon \in \mathcal{H}_m$ with

$$\operatorname{rank}\left(\sum_i \epsilon_i A_i\right) \ge c_0 r.$$

Invertibility of P, Q implies

$$\operatorname{rank}\left(\sum_i \epsilon_i D_i\right) = \operatorname{rank}\left(\sum_i \epsilon_i A_i\right) \geq c_0 r$$

Since $D(\epsilon)$ is a submatrix of $B(\epsilon)$,

$$\operatorname{rank} B(\epsilon) \geq \operatorname{rank} D(\epsilon) \geq c_0 r.$$

□

The NC^2 algorithm follows immediately by trying all $\epsilon \in \mathcal{H}_m$ in parallel followed by ordinary rank computation [Csa76, Mul87].

4 Using the properties of DVR

The main goal of this section is to connect the concepts from discrete valuation ring and Smith normal form to the rank approximation setting. This will be eventually used in Section 5 to analyze the algorithm. In particular, the normal form and the matrices J_k are used only to analyze progress. The algorithm itself works with the original rational matrices and their ordinary ranks.

Let

$$\mathcal{B} = \langle B_1, \ldots, B_m \rangle \leq \mathbb{Q}^{n \times n}, \qquad B(x) = \sum_{i=1}^{m} x_i B_i,$$

and let $c = \operatorname{crk}(\mathcal{B})$.

Fix $A \in \mathcal{B}$ of rank r. Apply the same invertible row and column transformations to A and all coefficient matrices B_i. These transformations preserve ranks and the commutative rank of the matrix space. For simplicity, retain the original notation.

$$A = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}, \qquad B(x) = \begin{pmatrix} X(x) & Q(x) \\ P(x) & Z(x) \end{pmatrix}.$$

Recall that $K = \mathbb{Q}(x_1, \ldots, x_m)$ and define the Schur complement

$$H(t, x) = Z - tP(I_r + tX)^{-1}Q = h_0 + th_1 + t^2 h_2 + \cdots.$$

Using

$$(I_r + tX)^{-1} = \sum_{\ell \geq 0} (-tX)^\ell,$$

we have

$$h_0 = Z, \qquad h_j = (-1)^j P X^{j-1} Q \quad (j \geq 1).$$

In particular, as polynomials in x, h_j is homogeneous of degree $j + 1$. For $k \geq 1$, let

$$J_k(H) \coloneqq J_k(A, B(x)) \coloneqq \begin{pmatrix} h_0 & 0 & \cdots & 0 \\ h_1 & h_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ h_{k-1} & \cdots & h_1 & h_0 \end{pmatrix}.$$

Note that the key concepts behind the construction of $J_k(H)$ is already mentioned in Section 2.1.3.

Remark 4.1. The matrix J_k essentially collects the first k coefficients of the product $H(t)v(t)$ into a single linear map. Similar block matrices were used by Van Dooren, Dewilde, and Vandewalle [VDV79] to study the local behavior of rational matrices from their expansion coefficients. See also Schumacher [Sch26] (mainly Appendix E) for an exposition.

Next, we prove the following theorem.

Theorem 4.2.

$$\operatorname{rank}_K J_k(A, B(x)) \geq k(c-r) - r = kc - (k+1)r.$$

Proof. Let $\delta = c - r$. The following claim is standard.

Claim 4.3.

$$\operatorname{rank}_{K(t)}(A + tB(x)) = c.$$

Proof. Choose a nonzero $c \times c$ minor $B[I, J](x)$ of $B(x)$. In the corresponding minor of $A + tB(x)$, the coefficient of t^c is exactly $\det B[I, J](x)$, so that minor is nonzero. Hence

$$\operatorname{rank}_{K(t)}(A + tB(x)) \geq c.$$

For the reverse inequality, write $A = \sum_i a_i B_i$. Then

$$A + tB(x) = \sum_i (a_i + tx_i)B_i.$$

If this matrix has rank higher than c, then by specializing suitable scalar values for t and x variables, we can exhibit a matrix in $\mathcal{B}$ which has rank higher than c. □

Since $I_r + tX$ is invertible over $K(t)$, the block matrix structure shows that the rank of the Schur-complement

$$\operatorname{rank}_{K((t))} H(t, x) = c - r = \delta.$$

Next we use two standrad facts. Since $K[[t]]$ is a discrete valuation ring, it is also a PID [AM69]. The second fact is that the Smith normal form exists over every PID [Hun74, Proposition 2.11, p.339], and in particular over $K[[t]]$. Thus there exist invertible matrices

$$U(t), V(t) \in \mathrm{GL}_{n-r}(K[[t]])$$

such that

$$U(t)H(t, x)V(t) = \operatorname{diag}\big(t^{\nu_1}, \ldots, t^{\nu_\delta}, 0, \ldots, 0\big),$$

where

$$0 \leq \nu_1 \leq \cdots \leq \nu_\delta,$$

and D be the diagonal matrix $\operatorname{diag}\big(t^{\nu_1}, \ldots, t^{\nu_\delta}, 0, \ldots, 0\big)$. The necessary background behind this decomposition is already given in Section 2.1.2. We claim the following.

Claim 4.4.

$$\sum_{i=1}^{\delta} \nu_i \leq r. \tag{8}$$

Assuming the claim, we give the rest of the argument.

Let $R_k = K[t]/(t^k)$. Reduce H, U, V, D modulo t^k; for simplicity we keep the same notation. We then have

$$UHV = D \qquad \text{in } R_k^{(n-r)\times(n-r)},$$

and $U, V \in \mathrm{GL}_{n-r}(R_k)$. The reason why even after modulo by t^k the matrices U and V continue to be invertible is the following. Originally the matrices $U(t)$ and $V(t)$ are invertible over $K[[t]]$. So their determinants are units over $K[[t]]$, which remain units even after modulo by t^k.

Consider the R_k-linear map

$$T_H : R_k^{n-r} \longrightarrow R_k^{n-r}, \qquad v \longmapsto Hv.$$

Because K embeds in R_k as the constant polynomials, the same map is also K-linear after restricting scalars. As a K-vector space,

$$R_k = \mathrm{Span}_K\{1, t, \ldots, t^{k-1}\}, \qquad \dim_K R_k = k,$$

so R_k^{n-r} has K-dimension $k(n-r)$.

Write

$$v(t) = v_0 + tv_1 + \cdots + t^{k-1}v_{k-1}, \qquad v_j \in K^{n-r}.$$

Modulo t^k,

$$\begin{aligned} H(t)v(t) = {} & h_0v_0 + t(h_1v_0 + h_0v_1) \\ & + t^2(h_2v_0 + h_1v_1 + h_0v_2) + \cdots . \end{aligned}$$

Thus, with respect to the K-basis induced by $1, t, \ldots, t^{k-1}$, the matrix of the K-linear map T_H is precisely $J_k(H)$. Hence

$$\mathrm{rank}_K\, J_k(H) = \dim_K \mathrm{Im}\, T_H.$$

The identity $UHV = D$ gives, as maps on R_k^{n-r},

$$T_D = T_U \circ T_H \circ T_V.$$

Since U and V are invertible over R_k, T_U and T_V are invertible. In particular, they are invertible as K-linear maps. Therefore

$$\dim_K \mathrm{Im}\, T_H = \dim_K \mathrm{Im}\, T_D.$$

It remains to compute the latter. Since D is diagonal, T_D is the direct sum of the maps

$$R_k \longrightarrow R_k, \qquad f(t) \longmapsto t^{\nu_i} f(t),$$

for $1 \le i \le \delta$, together with zero maps. For one Smith factor t^ν,

$$1, t, \ldots, t^{k-1} \longmapsto t^\nu, t^{\nu+1}, \ldots, t^{\nu+k-1} \pmod{t^k}.$$

Its image has K-basis

$$t^\nu, t^{\nu+1}, \ldots, t^{k-1}$$

when $\nu < k$, and is zero when $\nu \ge k$. Hence its K-rank is $\max\{k-\nu, 0\}$. Since the diagonal coordinates form a direct sum, the dimensions add, yielding

$$\mathrm{rank}_K\, J_k(H) = \sum_{i=1}^{\delta} \max\{k - \nu_i, 0\}.$$

Since $\max\{k-\nu_i, 0\} \geq k-\nu_i$, so by (8) and $\delta = c-r$,

$$\begin{aligned}\operatorname{rank}_K J_k(H) &\geq \sum_{i=1}^{\delta}(k-\nu_i) = k\delta - \sum_{i=1}^{\delta}\nu_i \\ &\geq k(c-r) - r \\ &= kc - (k+1)r.\end{aligned}$$

This proves the theorem.

□

4.1 The proof of Claim 4.4

Proof. The key idea is to extend a minimum valuation $\delta \times \delta$ minor of H to a nonzero $(r+\delta) \times (r+\delta)$ minor of $A + tB(x)$, whose degree in t bounds its valuation.

Since $K[[t]]$ is a discrete valuation ring, hence a PID, Smith normal form gives matrices

$$U(t), V(t) \in \mathrm{GL}_{n-r}(K[[t]])$$

such that

$$U(t)H(t,x)V(t) = D(t) := \operatorname{diag}\big(t^{\nu_1}, \ldots, t^{\nu_\delta}, 0, \ldots, 0\big).$$

We first show that there is a nonzero $\delta \times \delta$ minor $H[I,J](t)$ satisfying

$$\operatorname{ord}_t \det H[I,J](t) = \sum_{i=1}^{\delta}\nu_i. \tag{9}$$

For a matrix M over $K[[t]]$, let

$$\mathcal{I}_\delta(M) := \langle \det M[I,J] : |I| = |J| = \delta \rangle$$

be the ideal generated by all its $\delta \times \delta$ minors.

We claim that multiplication by invertible matrices over $K[[t]]$ does not change this ideal. For example, by the Cauchy–Binet formula,

$$\det(UM)[I,J] = \sum_{S:|S|=\delta} \det U[I,S] \det M[S,J].$$

Thus every δ-minor of UM is a $K[[t]]$-linear combination of δ-minors of M, and hence

$$\mathcal{I}_\delta(UM) \subseteq \mathcal{I}_\delta(M).$$

Applying the same argument to

$$M = U^{-1}(UM)$$

gives the reverse inclusion.

Remark 4.5. Note that

$$U^{-1} = \frac{\mathsf{adj}(U)}{\det U}.$$

But $\det(U)$ is a unit in $K[[t]]$ and hence $U^{-1} \in \mathrm{GL}_{n-r}(K[[t]])$.

Therefore $\mathcal{I}_\delta(UM) = \mathcal{I}_\delta(M)$. The same argument for right multiplication shows $\mathcal{I}_\delta(MV) = \mathcal{I}_\delta(M)$. Combining both we get that,

$$\mathcal{I}_\delta(H) = \mathcal{I}_\delta(D).$$

For the diagonal Smith matrix D, the only nonzero $\delta \times \delta$ minor is $t^{\nu_1} \cdots t^{\nu_\delta} = t^{\nu_1+\cdots+\nu_\delta}$. Hence $\mathcal{I}_\delta(H) = (t^{\nu_1+\cdots+\nu_\delta})$. So some nonzero $\delta \times \delta$ minor of H has valuation exactly $\nu_1 + \cdots + \nu_\delta$, proving (9). Considering $A + tB(x)$ again, define

$$D_0(t) := I_r + tX.$$

Since $D_0(0) = I_r$, $D_0(t) \in \mathrm{GL}_r(K[[t]])$.

Consider the $(r+\delta) \times (r+\delta)$ submatrix

$$M_{I,J}(t) = \begin{pmatrix} D_0(t) & tQ_{*,J} \\ tP_{I,*} & tZ_{I,J} \end{pmatrix}$$

of $A + tB(x)$.

To expose its determinant, define

$$L(t) := \begin{pmatrix} I_r & 0 \\ -tP_{I,*}D_0(t)^{-1} & I_\delta \end{pmatrix}, \qquad R(t) := \begin{pmatrix} I_r & -D_0(t)^{-1}tQ_{*,J} \\ 0 & I_\delta \end{pmatrix}.$$

Both matrices belong to $\mathrm{GL}_{r+\delta}(K[[t]])$ and satisfy

$$\det L(t) = \det R(t) = 1.$$

A direct block multiplication gives

$$L(t)M_{I,J}(t)R(t) = \begin{pmatrix} D_0(t) & 0 \\ 0 & tH[I,J](t) \end{pmatrix}.$$

Indeed, the lower-right block is

$$tZ_{I,J} - (tP_{I,*})D_0(t)^{-1}(tQ_{*,J}) = tH[I,J](t).$$

Since the left, right matrices have determinant one,

$$\det M_{I,J}(t) = \det D_0(t)\, \det\big(tH[I,J](t)\big).$$

As $H[I,J]$ is $\delta \times \delta$,

$$\det\big(tH[I,J](t)\big) = t^\delta \det H[I,J](t).$$

Therefore

$$\det M_{I,J}(t) = \det(I_r + tX)\, t^\delta \det H[I,J](t). \tag{10}$$

Now $\det(I_r + tX)$ has constant term 1, and hence

$$\mathrm{ord}_t \det(I_r + tX) = 0.$$

Using (9) in (10), we obtain

$$\mathrm{ord}_t \det M_{I,J}(t) = \delta + \sum_{i=1}^{\delta} \nu_i.$$

On the other hand, $M_{I,J}(t)$ is an $(r+\delta)\times(r+\delta)$ matrix whose entries are affine linear in t. Hence

$$\deg_t \det M_{I,J}(t) \le r+\delta.$$

The determinant is nonzero by (10). For every nonzero polynomial $f(t)$, $\operatorname{ord}_t f \le \deg_t f$. Thus

$$\delta + \sum_{i=1}^{\delta} \nu_i = \operatorname{ord}_t \det M_{I,J}(t) \le \deg_t \det M_{I,J}(t) \le r+\delta.$$

Thus we get that, $\sum_{i=1}^{\delta} \nu_i \le r$. □

We give an illustrative example.

4.2 An example

Example 4.6. *Let*

$$R_5 := K[t]/(t^5),$$

with K-basis

$$1, t, t^2, t^3, t^4.$$

Consider first multiplication by t^2:

$$m_{t^2} : R_5 \to R_5, \qquad f(t) \mapsto t^2 f(t).$$

On the basis elements,

$$1 \mapsto t^2, \qquad t \mapsto t^3, \qquad t^2 \mapsto t^4, \qquad t^3 \mapsto 0, \qquad t^4 \mapsto 0.$$

Hence, with respect to the basis

$$1, t, t^2, t^3, t^4,$$

the corresponding K-matrix is

$$T_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

Therefore $\operatorname{rank}_K T_2 = 3$.

Similarly, multiplication by t^3,

$$m_{t^3} : R_5 \to R_5, \qquad f(t) \mapsto t^3 f(t),$$

satisfies

$$1 \mapsto t^3, \qquad t \mapsto t^4, \qquad t^2, t^3, t^4 \mapsto 0.$$

Thus its matrix is

$$T_3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix},$$

and hence $\operatorname{rank}_K T_3 = 2$.

Now suppose the truncated Smith diagonal is

$$D(t) = \begin{pmatrix} t^2 & 0 \\ 0 & t^3 \end{pmatrix}$$

over R_5. *As an* R_5*-linear map it acts on* R_5^2. *After viewing* R_5^2 *as a* 10*-dimensional vector space over* K, *the corresponding* K*-linear map is represented by the block diagonal matrix*

$$\widetilde{D} = \begin{pmatrix} T_2 & 0 \\ 0 & T_3 \end{pmatrix}.$$

Therefore $\operatorname{rank}_K \widetilde{D} = \operatorname{rank}_K T_2 + \operatorname{rank}_K T_3 = 5$.

5 The $(1-\varepsilon)$-approximation for the matrix rank in NC^3

In this section, we complete the proof of Theorem 1.1. Let,

$$\mathcal{B} = \langle B_1, \ldots, B_m \rangle \leq \mathbb{Q}^{n\times n}, \qquad B(x) = \sum_{i=1}^{m} x_i B_i,$$

and write $c := \operatorname{crk}(\mathcal{B}) = \operatorname{rank}_{\mathbb{Q}(x)} B(x)$. Let the current scalar matrix be $A = B(a) \in \mathcal{B}$, and $r := \operatorname{rank} A$.

We have already proved the following in Theorem 3.5. There is an absolute constant c_0 and, for every number of variables s, a polynomial-size family $\mathcal{H}_s \subseteq \{\pm 1\}^s$ that is uniformly enumerable in deterministic parallel time, such that for every homogeneous rational linear pencil $L(y) = \sum_{i=1}^{s} y_i L_i$ of generic rank R has some $\epsilon \in \mathcal{H}_s$ with $\operatorname{rank} L(\epsilon) \geq c_0 R$.

Also, we have already proved,

$$\operatorname{rank}_{\mathbb{Q}(x)} J_k(A, B(x)) \geq kc - (k+1)r.$$

Define

$$\Delta_k(A) := kc - (k+1)\operatorname{rank} A. \tag{11}$$

If $\Delta_k(A) \leq 0$, then

$$\operatorname{rank} A \geq \frac{k}{k+1} c,$$

and we get approximation of the quality [BJP18] by fixing $k = \lceil 1/\varepsilon - 1 \rceil$.

We first state a result, and assuming that we complete the rest of the argument.

Lemma 5.1. *Fix* k. *There is a polynomial-size, deterministically (in parallel) computable family* $\mathcal{D}_{k,m} \subseteq \mathbb{Q}^m$ *and a constant* $\beta_k := \frac{c_0^k}{2^k}$ *with the following property. For every matrix space* $\mathcal{B}$ *and every current scalar matrix* $A \in \mathcal{B}$, *some* $b \in \mathcal{D}_{k,m}$, *with* $C := B(b)$, *satisfies*

$$\operatorname{rank} J_k(A, C) \geq \beta_k \operatorname{rank}_{\mathbb{Q}(x)} J_k(A, B(x)). \tag{12}$$

Moreover, for every coordinate of $b \in \mathcal{D}_{k,m}$ *the numerator and the denominator are bounded by* k.

Assuming Lemma 5.1, we complete the argument.

Lemma 5.2. *Let* $s := \operatorname{rank}_{\mathbb{Q}(t)}(A + tC) - r$. *Then,*

$$\operatorname{rank} J_k(A, C) \leq ks. \tag{13}$$

Proof. Use left and right transformations to write

$$A = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Let $H_C(t)$ be defined as in Section 4, so that the Schur complement of the upper-left block of $A + tC$ is $tH_C(t)$. Since the upper-left block is invertible over $\mathbb{Q}(t)$, $\operatorname{rank}_{\mathbb{Q}((t))} H_C(t) = s$. Thus over the DVR $\mathbb{Q}[[t]]$, the nonzero Smith factors of H_C are $t^{\nu_1}, \ldots, t^{\nu_s}$ up to units. After reducing modulo t^k, the i^{th} term contributes $\max\{k - \nu_i, 0\}$, as we already discussed in the proof of Theorem 4.2. Hence,

$$\operatorname{rank} J_k(A, C) = \sum_{i=1}^{s} \max\{k - \nu_i, 0\} \leq \sum_{i=1}^{s} k = ks.$$

□

The following lemma is standard and the proof is similar to the proof of Claim 4.3. This follows from the simple fact that the determinant as a polynomial in t has degree at most n.

Lemma 5.3. *For any* $A, C \in \mathbb{Q}^{n \times n}$, *there exists* $\lambda \in \{0, 1, \ldots, n\}$ *such that*

$$\operatorname{rank}(A + \lambda C) = \operatorname{rank}_{\mathbb{Q}(t)}(A + tC).$$

5.1 The boosting step

We are handling the case,

$$\operatorname{rank}_{\mathbb{Q}(x)} J_k(A, B(x)) \geq \Delta_k(A).$$

By Lemma 5.1, some $b \in \mathcal{D}_{k,m}$, with $C = B(b)$, satisfies

$$\operatorname{rank} J_k(A, C) \geq \beta_k \Delta_k(A).$$

Recall that $s = \operatorname{rank}_{\mathbb{Q}(t)}(A + tC) - r$. Lemma 5.2 gives

$$s \geq \frac{1}{k} \operatorname{rank} J_k(A, C) \geq \frac{\beta_k}{k} \Delta_k(A).$$

Lemma 5.3 now provides $\lambda \in \{0, \ldots, n\}$ such that

$$\operatorname{rank}(A + \lambda C) = r + s \geq r + \frac{\beta_k}{k} \Delta_k(A).$$

Continuing like this, let A_j be the current matrix after round j. Let $r_j := \operatorname{rank} A_j$, and define

$$\Delta_k^{(j)} := \Delta_k(A_j) = kc - (k+1)r_j. \tag{14}$$

Here k is fixed throughout the run, only j changes.

If $\Delta_k^{(j)} > 0$, the preceding argument gives a next matrix $A_{j+1} \in \mathcal{B}$ satisfying

$$r_{j+1} \geq r_j + \frac{\beta_k}{k} \Delta_k^{(j)}.$$

Therefore

$$\begin{aligned}\Delta_k^{(j+1)} &= kc - (k+1)r_{j+1} \\ &\le kc - (k+1)\left(r_j + \frac{\beta_k}{k}\Delta_k^{(j)}\right) \\ &= \left(1 - \frac{k+1}{k}\beta_k\right)\Delta_k^{(j)}.\end{aligned}$$

Let

$$\rho_k := 1 - \frac{k+1}{k}\beta_k. \tag{15}$$

For fixed k, the number $\rho_k \in (0,1)$ is the same in every round. Thus, as long as the deficit is positive,

$$\Delta_k^{(j+1)} \le \rho_k \Delta_k^{(j)}. \tag{16}$$

Theorem 5.4. *For every fixed $k \ge 1$, after $T = O_k(\log n)$ rounds, the algorithm outputs $A \in \mathcal{B}$ satisfying*

$$\operatorname{rank} A \ge \frac{k}{k+1}\operatorname{crk}(\mathcal{B}).$$

Proof. Start with the matrix A_0 produced by Theorem 3.5. Then

$$\Delta_k^{(0)} = kc - (k+1)\operatorname{rank} A_0 \le kn.$$

Choose T so that $\rho_k^T kn < 1$. Since k is fixed and $\rho_k < 1$ is a fixed constant, such a T is $O_k(\log n)$.

Suppose, for contradiction, that

$$\Delta_k^{(j)} > 0 \qquad \text{for every } 0 \le j \le T.$$

Then the contraction estimate applies in every round and gives

$$0 < \Delta_k^{(T)} \le \rho_k^T \Delta_k^{(0)} \le \rho_k^T kn < 1.$$

But

$$\Delta_k^{(T)} = kc - (k+1)r_T$$

is an integer, a contradiction. Hence for some $j \le T$ one has

$$\Delta_k^{(j)} \le 0,$$

which is equivalent to

$$r_j \ge \frac{k}{k+1}c.$$

The actual deterministic update simply choose in parallel, among all candidates λ considered in a round, one of maximum ordinary rank. □

To summarize, for fixed k, the algorithm enumerates $\mathcal{D}_{k,m}$ once. In each round, it computes $A_j + \lambda B(b)$ and its rank for every $b \in \mathcal{D}_{k,m}$ and every $\lambda \in \{0, 1, \dots, n\}$, all in parallel, and selects a candidate of maximum rank.

We briefly account for bit complexity. Let L bound the bit length of the numerators and denominators of the input entries, and let

$$A' = \sum_{i=1}^{m} \epsilon_i B_i, \qquad \epsilon_i \in \{\pm 1\},$$

be the matrix output by Theorem 3.5. After T updates, the output matrix is

$$A' + \sum_{j=1}^{T} \lambda_j B(b^{(j)}) = \sum_{i=1}^{m} \left(\epsilon_i + \sum_{j=1}^{T} \lambda_j b_i^{(j)} \right) B_i \tag{17}$$

where $\lambda_j \in \{0, \ldots, n\}$ and $b^{(j)} \in \mathcal{D}_{k,m}$. It is shown in Section 5.2 that every coordinate of a vector in $\mathcal{D}_{k,m}$ has numerator and denominator at most k in absolute value. Hence from Equation 17 it is easy to see that the final bit complexity bound for each matrix entry is given by

$$O_k(mL + mT + \log(mn)) .$$

Since $T = O_k(\log n)$, all candidate matrices have bit length polynomial in the input size.

Since each round can be performed in NC^2, using parallel rank computation [Csa76, Mul87], and there are $O_k(\log(n+1))$ rounds, the overall algorithm lies in NC^3.

Moreover, $\lambda = 0$ is among the candidates in every round, so the rank never decreases. Hence, once $\Delta_k(A_j) \leq 0$, it remains non-positive in all subsequent rounds. Choosing $k = \lceil 1/\varepsilon - 1 \rceil$ gives the desired $(1-\varepsilon)$-approximation.

5.2 Proving Lemma 5.1

In this section we prove Lemma 5.1. The main difficulty in proving Lemma 5.1 is that the matrix J_k is not a linear (or affine linear). If that is the case, then we will be able to apply the construction of 4-wise independence, and we will be done. So we try to find a way to *linearize* (in a certain sense) the matrix J_k without paying too much for the rank. A standard approach to such problems is the use of polarization; see [Pro07, Chapter 3, §2] and [Lan12, §2.6.4].

Throughout this proof, A and $B_1, \ldots, B_m$ are fixed. For brevity, write

$$J_k(x) := J_k(A, B(x)).$$

Proof. The first observation is that the entries of J_k are not homogeneous. In particular, the degree of the terms in h_j is $j+1$. We consider a new variable z and multiply each h_j by z^{k-j-1}, to obtain a new matrix $\widehat{J}_k$. Notice that each term in $\widehat{J}_k$ is homogeneous of degree k over the set of variables $Y = \{z, x_1, x_2, \ldots, x_m\}$. We first compare generic ranks. Let

$$R := \operatorname{rank}_{\mathbb{Q}(x_1,\ldots,x_m)} J_k(x_1, \ldots, x_m),$$

and choose a nonzero $R \times R$ minor $p(x)$ of $J_k(x)$. Let $\widehat{p}(z, x)$ denote the corresponding minor of $\widehat{J}_k(z, x)$. Since the entries of $\widehat{J}_k$ are polynomials,

$$\widehat{p}(1, x) = p(x) \neq 0.$$

Therefore $\widehat{p}(z, x)$ is itself a nonzero polynomial, and hence

$$\operatorname{rank}_{\mathbb{Q}(Y)} \widehat{J}_k(Y) \geq \operatorname{rank}_{\mathbb{Q}(x_1,\ldots,x_m)} J_k(x_1, \ldots, x_m). \tag{18}$$

Consider k copies of the variable set Y and denote them by $\{Y^{(1)}, \ldots, Y^{(k)}\}$. In particular,

$$Y^{(j)} = \{z^{(j)}, x_1^{(j)}, \ldots, x_m^{(j)}\}.$$

Next, we use the concept of *Polarization* [Lan12, §2.6.4],

$$\mathrm{Pol}(\widehat{J}_k(Y^{(1)},\dots,Y^{(k)})) \coloneqq \sum_{S\subseteq[k]} (-1)^{k-|S|}\widehat{J}_k\left(\sum_{j\in S} Y^{(j)}\right),$$

and,

$$\sum_{j\in S} Y^{(j)} = \left(\sum_{j\in S} z^{(j)}, \sum_{j\in S} x_1^{(j)}, \dots, \sum_{j\in S} x_m^{(j)}\right).$$

By the sub-additive property of rank function, we get that there exists $S \subseteq [k]$ such that

$$\mathrm{rank}_{\mathbb{Q}(Y^{(1)},\dots,Y^{(k)})}\,\widehat{J}_k\left(\sum_{j\in S} Y^{(j)}\right) \geq \frac{1}{2^k}\,\mathrm{rank}_{\mathbb{Q}(Y^{(1)},\dots,Y^{(k)})}(\mathrm{Pol}(\widehat{J}_k(Y^{(1)},\dots,Y^{(k)}))). \tag{19}$$

Now we observe a few basic properties.

Lemma 5.5.

$$\mathrm{rank}_{\mathbb{Q}(Y^{(1)},\dots,Y^{(k)})}(\mathrm{Pol}(\widehat{J}_k(Y^{(1)},\dots,Y^{(k)}))) \geq \mathrm{rank}_{\mathbb{Q}(Y)}(\widehat{J}_k(z,x_1,\dots,x_m)).$$

Proof. By homogeneity of degree k,

$$\mathrm{Pol}(\widehat{J}_k)(Y,\dots,Y) = \left(\sum_{S\subseteq[k]} (-1)^{k-|S|}|S|^k\right)\widehat{J}_k(Y) = k!\,\widehat{J}_k(Y).$$

Let

$$R \coloneqq \mathrm{rank}_{\mathbb{Q}(Y)}\,\widehat{J}_k(Y),$$

and choose a nonzero $R\times R$ minor $p(Y)$ of $\widehat{J}_k(Y)$. Let $q(Y^{(1)},\dots,Y^{(k)})$ be the corresponding minor of $\mathrm{Pol}(\widehat{J}_k)$. Then the polynomial substitution $Y^{(1)} = \cdots = Y^{(k)} = Y$ gives

$$q(Y,\dots,Y) = (k!)^R p(Y) \neq 0.$$

Hence q is not the zero polynomial, proving the rank inequality. □

Next, we prove that the entries in $\mathrm{Pol}(\widehat{J}_k(Y^{(1)},\dots,Y^{(k)}))$ are linear in $Y^{(1)},\dots,Y^{(k)}$.

Lemma 5.6. *Every entry of* $\mathrm{Pol}(\widehat{J}_k)(Y^{(1)},\dots,Y^{(k)})$ *is homogeneous of degree one in each variable group* $Y^{(i)}$ *separately.*

Proof. It is enough to argue entrywise, and then monomial by monomial. Fix a monomial m' occurring in the expansion of

$$\sum_{S\subseteq[k]} (-1)^{k-|S|}\, m\left(\sum_{j\in S} Y^{(j)}\right),$$

and let $T \subseteq [k]$ be the set of variable buckets actually used by m'. The monomial m' can occur only from terms with $T \subseteq S$, and with the same coefficient in each such term. Its total coefficient is therefore multiplied by

$$\sum_{\substack{S\subseteq[k]\\ T\subseteq S}} (-1)^{k-|S|}.$$

Writing $S = T \cup U$ gives

$$\sum_{S \supseteq T} (-1)^{k-|S|} = (-1)^{k-|T|} \sum_{U \subseteq [k] \setminus T} (-1)^{|U|} = (-1)^{k-|T|}(1-1)^{k-|T|}.$$

Thus the coefficient vanishes unless $T = [k]$. Since every entry of $\widehat{J}_k$ is homogeneous of total degree k, any surviving monomial has total degree k and uses all k buckets. Consequently it uses each bucket exactly once. Hence every entry of the polarization is homogeneous linear in each bucket separately, i.e. multilinear of multidegree $(1, \ldots, 1)$ over the buckets $Y^{(1)}, \ldots, Y^{(k)}$. □

Putting everything together. Let

$$P(Y^{(1)}, \ldots, Y^{(k)}) := \mathrm{Pol}(\widehat{J}_k(Y^{(1)}, \ldots, Y^{(k)}))$$

and

$$R_0 := \mathrm{rank}_{\mathbb{Q}(Y^{(1)}, \ldots, Y^{(k)})} P(Y^{(1)}, \ldots, Y^{(k)}).$$

By Lemma 5.6, P is homogeneous linear in $Y^{(1)}$, with coefficients that are polynomial matrices in $Y^{(2)}, \ldots, Y^{(k)}$.

Choose a nonzero $R_0 \times R_0$ minor of P. Since $\mathbb{Q}$ is infinite, we can specialize (only for the purpose of proof) $Y^{(2)}, \ldots, Y^{(k)}$ to rational values $a^{(2)}, \ldots, a^{(k)}$ so that this minor remains a nonzero polynomial in $Y^{(1)}$. Hence,

$$\mathrm{rank}_{\mathbb{Q}(Y^{(1)})} P(Y^{(1)}, a^{(2)}, \ldots, a^{(k)}) = R_0.$$

Theorem 3.5, now applied to this linear pencil over the first bucket, gives some $\varepsilon^{(1)} \in \mathcal{H}_{m+1}$ such that

$$\mathrm{rank}\, P(\epsilon^{(1)}, a^{(2)}, \ldots, a^{(k)}) \geq c_0 R_0.$$

Therefore the corresponding minor of

$$P(\epsilon^{(1)}, Y^{(2)}, \ldots, Y^{(k)})$$

is not identically zero, and so

$$\mathrm{rank}_{\mathbb{Q}(Y^{(2)}, \ldots, Y^{(k)})} P(\epsilon^{(1)}, Y^{(2)}, \ldots, Y^{(k)}) \geq c_0 R_0.$$

Repeating the same argument bucket by bucket, we obtain

$$(\epsilon^{(1)}, \ldots, \epsilon^{(k)}) \in \mathcal{H}_{m+1}^k$$

such that

$$\mathrm{rank}\, P(\epsilon^{(1)}, \ldots, \epsilon^{(k)}) \geq c_0^k R_0. \tag{20}$$

Now consider the identity:

$$P(\epsilon^{(1)}, \ldots, \epsilon^{(k)}) = \sum_{S \subseteq [k]} (-1)^{k-|S|} \widehat{J}_k \left(\sum_{j \in S} \epsilon^{(j)} \right).$$

By rank sub-additivity, some $S \subseteq [k]$ satisfies

$$\begin{aligned} \mathrm{rank}\, \widehat{J}_k \left(\sum_{j \in S} \epsilon^{(j)} \right) &\geq \frac{1}{2^k} \mathrm{rank}\, P(\epsilon^{(1)}, \ldots, \epsilon^{(k)}) \\ &\geq \frac{c_0^k}{2^k} R_0. \end{aligned} \tag{21}$$

Combining Lemma 5.5 with (18), this gives

$$\operatorname{rank} \widehat{J}_k\left(\sum_{j\in S} \epsilon^{(j)}\right) \geq \frac{c_0^k}{2^k} \operatorname{rank}_{\mathbb{Q}(x)} J_k(x_1, \ldots, x_m). \tag{22}$$

Now define the family

$$\mathcal{Z}_{k,m+1} := \left\{ \sum_{j\in S} \epsilon^{(j)} : S \subseteq [k],\ (\epsilon^{(1)}, \ldots, \epsilon^{(k)}) \in \mathcal{H}_{m+1}^k \right\}.$$

For fixed k,

$$|\mathcal{Z}_{k,m+1}| \leq 2^k |\mathcal{H}_{m+1}|^k = \operatorname{poly}(m),$$

and the family is uniformly enumerable in deterministic NC^2.

For every

$$\widetilde{b} = (b_0, b_1, \ldots, b_m) \in \mathcal{Z}_{k,m+1}$$

put into $\mathcal{D}_{k,m}$ the vector

$$b = \begin{cases} (b_1/b_0, \ldots, b_m/b_0), & b_0 \neq 0, \\ (b_1, \ldots, b_m), & b_0 = 0. \end{cases}$$

This remains a polynomial-size deterministically (in parallel) enumerable family for fixed k.

Every coordinate of $\widetilde{b}$ is an integer of absolute value at most k. Hence for every coordinate of $b \in \mathcal{D}_{k,m}$ the numerator and the denominator are bounded by k in absolute value. Choose $\widetilde{b} = (b_0, b_1, \ldots, b_m)$ guaranteed by Equation 22, and let b be its corresponding vector in $\mathcal{D}_{k,m}$. If $b_0 \neq 0$, homogeneity of the blocks gives

$$\widehat{J}_k(\widetilde{b}) = b_0^k J_k(A, B(b)).$$

Since $b_0 \neq 0$, it also gives that

$$\operatorname{rank} J_k(A, B(b)) = \operatorname{rank} \widehat{J}_k(\widetilde{b}) \geq \frac{c_0^k}{2^k} \operatorname{rank}_{\mathbb{Q}(x)} J_k(A, B(x)).$$

If $b_0 = 0$, every homogenized block $z^{k-j-1}h_j$ with $j < k-1$ vanishes. The only potentially nonzero block is $h_{k-1}(b_1, \ldots, b_m)$ in the bottom-left position. Hence

$$\operatorname{rank} \widehat{J}_k(\widetilde{b}) = \operatorname{rank} h_{k-1}(b_1, \ldots, b_m).$$

Since $b = (b_1, \ldots, b_m)$ in this case, $J_k(A, B(b))$ contains this block, and therefore

$$\operatorname{rank} J_k(A, B(b)) \geq \operatorname{rank} \widehat{J}_k(\widetilde{b}).$$

Together with Equation 22, this proves Lemma 5.1 with $\beta_k = c_0^k/2^k$. □

This also completes the proof of Theorem 1.1.

Open problem

Extending our result to sufficiently large finite fields is a natural question. The polynomial-time approximation scheme of Bläser, Jindal, and Pandey [BJP18] works over fields of cardinality greater than n. Can one obtain a deterministic NC algorithm that, for every fixed $0 < \varepsilon < 1$ and every matrix space $\mathcal{B} \leq \mathbb{F}_q^{n \times n}$ with $q > n$, outputs $A \in \mathcal{B}$ satisfying

$$\operatorname{rank} A \geq (1 - \varepsilon) \operatorname{crk}(\mathcal{B})?$$

Our proof over $\mathbb{Q}$ uses operator scaling and polarization. If $\mathsf{char}(F_q) > k$, then $k!$ term appearing in the proof of Lemma 5.5 gives no obstruction. However, we still need to find an alternative of the rank rounding method which currently uses operator scaling. Finding suitable alternatives to these ingredients over finite fields remains an interesting direction.

References

[ABI86] Noga Alon, László Babai, and Alon Itai. A fast and simple randomized parallel algorithm for the maximal independent set problem. *Journal of Algorithms*, 7(4):567–583, 1986.

[AM69] M. F. Atiyah and I. G. Macdonald. *Introduction to Commutative Algebra*. Addison-Wesley Publishing Company, Reading, Massachusetts, 1969.

[BBJP19] Vishwas Bhargava, Markus Bläser, Gorav Jindal, and Anurag Pandey. A deterministic PTAS for the algebraic rank of bounded degree polynomials. In *Proceedings of the Thirtieth Annual ACM–SIAM Symposium on Discrete Algorithms (SODA 2019)*, pages 647–661. SIAM, 2019.

[BCDZ26] Joshua Brakensiek, Yeyuan Chen, Manik Dhar, and Zihan Zhang. From random to explicit via subspace designs with applications to local properties and matroids. In Aditya Bhaskara and Artur Czumaj, editors, *Proceedings of the 58th Annual ACM Symposium on Theory of Computing, STOC 2026, Salt Lake City, UT, USA, June 22-26, 2026*, pages 619–630. ACM, 2026.

[Ber97] Bonnie Berger. The fourth moment method. *SIAM Journal on Computing*, 26(4):1188–1207, 1997.

[Bha07] Rajendra Bhatia. *Positive Definite Matrices*, volume 44 of *Texts and Readings in Mathematics*. Hindustan Book Agency, New Delhi, 2007.

[BJP18] Markus Bläser, Gorav Jindal, and Anurag Pandey. A deterministic PTAS for the commutative rank of matrix spaces. *Theory of Computing*, 14(3):1–21, 2018.

[CDK+26] Srijan Chakraborty, Samir Datta, Aryan Kusre, Partha Mukhopadhyay, and Amit Sinhababu. Maximum matching and related problems in catalytic logspace. *CoRR*, abs/2604.24275, 2026.

[CGG+26] Abhranil Chatterjee, Sumanta Ghosh, Rohit Gurjar, Roshan Raj, and Thomas Thierauf. Bipartite matching is in NC. *Electronic Colloquium on Computational Complexity (ECCC)*, TR26-100, 2026.

[CM23] Abhranil Chatterjee and Partha Mukhopadhyay. The noncommutative Edmonds' problem re-visited, 2023.

[Csa76] Lajos Csanky. Fast parallel matrix inversion algorithms. *SIAM Journal on Computing*, 5(4):618–623, 1976.

[DL78] Richard A. DeMillo and Richard J. Lipton. A probabilistic remark on algebraic program testing. *Information Processing Letters*, 7(4):193–195, 1978.

[DPT24] Dean Doron, Edward Pyne, and Roei Tell. Opening up the distinguisher: A hardness to randomness approach for BPL = L that uses properties of BPL. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, STOC '24, pages 2039–2049. ACM, 2024.

[Edm67] Jack Edmonds. Systems of distinct representatives and linear algebra. *Journal of Research of the National Bureau of Standards, Section B*, 71B(4):241–245, 1967.

[FR04] Marc Fortin and Christophe Reutenauer. Commutative/noncommutative rank of linear matrices and subspaces of matrices of low rank. *Séminaire Lotharingien de Combinatoire*, 52:B52f, 2004.

[GGOW18] Ankit Garg, Leonid Gurvits, Rafael Oliveira, and Avi Wigderson. Algorithmic and optimization aspects of brascamp–lieb inequalities, via operator scaling. *Geometric and Functional Analysis*, 28(1):100–145, 2018.

[GGOW20] Ankit Garg, Leonid Gurvits, Rafael Oliveira, and Avi Wigderson. Operator scaling: Theory and applications. *Foundations of Computational Mathematics*, 20(2):223–290, 2020.

[GK16] Venkatesan Guruswami and Swastik Kopparty. Explicit subspace designs. *Combinatorica*, 36(2):161–185, 2016.

[Gur04] Leonid Gurvits. Classical complexity and quantum entanglement. *Journal of Computer and System Sciences*, 69(3):448–484, 2004.

[HH21] Masaki Hamada and Hiroshi Hirai. Computing the nc-rank via discrete convex optimization on CAT(0) spaces. *SIAM Journal on Applied Algebra and Geometry*, 5(3):455–478, 2021.

[Hun74] Thomas W. Hungerford. *Algebra*, volume 73 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1974.

[IQS18] Gábor Ivanyos, Youming Qiao, and K. Venkata Subrahmanyam. Constructive non-commutative rank computation is in deterministic polynomial time. *Computational Complexity*, 27(4):561–593, 2018.

[KI04] Valentine Kabanets and Russell Impagliazzo. Derandomizing polynomial identity tests means proving circuit lower bounds. *Computational Complexity*, 13(1–2):1–46, 2004.

[Lan02] Serge Lang. *Algebra*, volume 211 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 3rd edition, 2002.

[Lan12] J. M. Landsberg. *Tensors: Geometry and Applications*, volume 128 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.

[Lov79] László Lovász. On determinants, matchings, and random algorithms. In Lothar Budach, editor, *Fundamentals of Computation Theory (FCT 1979)*, pages 565–574, Berlin, 1979. Akademie-Verlag.

[Mul87] Ketan Mulmuley. A fast parallel algorithm to compute the rank of a matrix over an arbitrary field. *Combinatorica*, 7(1):101–104, 1987.

[Oki20] Taihei Oki. On solving (non)commutative weighted edmonds' problem. In *47th International Colloquium on Automata, Languages, and Programming (ICALP 2020)*, volume 168 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 89:1–89:14. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020.

[Oki23] Taihei Oki. Computing valuations of the dieudonné determinants. *Journal of Symbolic Computation*, 116:284–323, 2023.

[Pro07] Claudio Procesi. *Lie Groups: An Approach through Invariants and Representations*. Universitext. Springer, New York, 2007.

[Sch80] Jacob T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *Journal of the ACM*, 27(4):701–717, 1980.

[Sch26] Johannes M. Schumacher. The mathematical history behind the Granger–Johansen representation theorem. *Oxford Bulletin of Economics and Statistics*, 88(4):587–602, 2026.

[Ser79] Jean-Pierre Serre. *Local Fields*, volume 67 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1979.

[Tut47] W. T. Tutte. The factorization of linear graphs. *Journal of the London Mathematical Society*, s1-22(2):107–111, 1947.

[VDV79] Paul Van Dooren, Patrick Dewilde, and Joos Vandewalle. On the determination of the Smith–MacMillan form of a rational matrix from its laurent expansion. *IEEE Transactions on Circuits and Systems*, 26(3):180–189, 1979.

[Zip79] Richard E. Zippel. Probabilistic algorithms for sparse polynomials. In *Symbolic and Algebraic Computation (EUROSAM 1979)*, volume 72 of *Lecture Notes in Computer Science*, pages 216–226. Springer, 1979.

A The operator scaling result

In this section, we mainly records some facts used in the proof of Theorem 3.2 for completeness. All concepts and results are from [Gur04, GGOW20]. For a tuple

$$D = (D_1, \dots, D_m), \qquad D_i \in \mathbb{C}^{r \times r},$$

write

$$T_D : M_r(\mathbb{C}) \longrightarrow M_r(\mathbb{C}), \qquad T_D(X) := \sum_{i=1}^{m} D_i X D_i^*.$$

Thus, T_D is a completely positive, and in particular positive, linear operator.

Recall that the tuple D is *dimension non-decreasing* if

$$\dim\left(\sum_{i=1}^{m} D_i(U)\right) \geq \dim U \qquad \text{for every subspace } U \leq \mathbb{C}^r.$$

The following result is well known.

Lemma A.1. *[GGOW20, Theorem 1.4] The tuple D is dimension non-decreasing if and only if the positive operator T_D is rank non-decreasing, i.e.,*

$$\operatorname{rank} T_D(X) \geq \operatorname{rank} X \qquad \textit{for every } X \succeq 0.$$

For a tuple $A = (A_1, \ldots, A_m)$ define

$$M(A) := \sum_i A_i A_i^*, \qquad N(A) := \sum_i A_i^* A_i,$$

and

$$\mathsf{ds}(A) := \|M(A) - I_r\|_F^2 + \|N(A) - I_r\|_F^2.$$

Using the connection with the operator T_D the following theorem is proved in [Gur04]. Also, see [GGOW18, Section 2] for a more comprehensive explanation.

Theorem A.2. *[Gur04, Theorem 4.6] If $D = (D_1, \ldots, D_m)$ is dimension non-decreasing, then for every $\eta > 0$ there exist invertible matrices $L, R \in \mathrm{GL}_r(\mathbb{C})$ such that, for $A_i := LD_iR$, one has $\mathsf{ds}(A) \leq \eta$.*

This is essentially the restatement of Theorem 3.2.

ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il