

On the Probability of Generating a Primitive Matrix*

CHEN Jingwei · FENG Yong · LIU Yang · WU Wenyuan

DOI: 10.1007/s11424-024-1313-6

Received: 20 August 2021 / Revised: 20 September 2022

©The Editorial Office of JSSC & Springer-Verlag GmbH Germany 2024

Abstract Given a $k \times n$ integer *primitive* matrix \mathbf{A} (i.e., a matrix can be extended to an $n \times n$ unimodular matrix over the integers) with the maximal absolute value of entries $\|\mathbf{A}\|$ bounded by an integer λ from above, the authors study the probability that the $m \times n$ matrix extended from \mathbf{A} by appending other $m-k$ row vectors of dimension n with entries chosen randomly and independently from the uniform distribution over $\{0, 1, \dots, \lambda-1\}$ is still primitive. The authors present a complete and rigorous proof of a lower bound on the probability, which is at least a constant for fixed m in the range $[k+1, n-4]$. As an application, the authors prove that there exists a fast Las Vegas algorithm that completes a $k \times n$ primitive matrix \mathbf{A} to an $n \times n$ unimodular matrix within expected $\tilde{O}(n^\omega \log \|\mathbf{A}\|)$ bit operations, where \tilde{O} is big- O but without log factors, ω is the exponent on the arithmetic operations of matrix multiplication.

Keywords Integer matrix, matrix completion, probabilistic algorithm, unimodular matrix.

CHEN Jingwei

Chongqing Key Laboratory of Automated Reasoning and Cognition, Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing 400714, China; Chongqing College, University of Chinese Academy of Sciences, Chongqing 400714, China; Key Laboratory of Advanced Manufacturing Technology of Ministry of Education, Guizhou University, Guiyang 550025, China. Email: chenjingwei@cigit.ac.cn.

FENG Yong

Chongqing Key Laboratory of Automated Reasoning and Cognition, Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing 400714, China; Chongqing College, University of Chinese Academy of Sciences, Chongqing 400714, China. Email: yongfeng@cigit.ac.cn.

LIU Yang (Corresponding author)

Information Science and Engineering, Chongqing Jiaotong University, Chongqing 400074, China.

Email: liuyang13@cqjtu.edu.cn.

WU Wenyuan

Chongqing Key Laboratory of Automated Reasoning and Cognition, Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing 400714, China; Chongqing College, University of Chinese Academy of Sciences, Chongqing 400714, China. Email: wuwenyuan@cigit.ac.cn.

*This research was supported by the National Key Research and Development Project under Grant No. 2020YFA0712303, National Science Foundational of China under Grant No. 61903053, Youth Innovation Promotion Association of CAS, Guizhou Science and Technology Program under Grant No. [2020]4Y056 and Chongqing Science and Technology Program under Grant Nos. cstc2021jcyj-msxmX0821, cstc2020yszx-jcyjX0005, cstc2021yszx-jcyjX0004, and 2022YSZX-JCX0011CSTB.

◊*This paper was recommended for publication by Editor FENG Ruyong.*

1 Introduction

A vector $\mathbf{x} \in \mathbb{Z}^n$ is called *primitive* if $\mathbf{x} = d\mathbf{y}$ for $\mathbf{y} \in \mathbb{Z}^n$ and $d \in \mathbb{Z}$ implies $d = \pm 1$. More generally, a matrix $\mathbf{A} \in \mathbb{Z}^{k \times n}$ with $k \leq n$ is called *primitive* if $\mathbf{x} = \mathbf{y}\mathbf{A} \in \mathbb{Z}^n$ for $\mathbf{y} \in \mathbb{Q}^k$ implies $\mathbf{y} \in \mathbb{Z}^k$; in this case we also say the k rows of \mathbf{A} are *primitive* in \mathbb{Z}^n . In particular, an $n \times n$ primitive matrix over \mathbb{Z} is also called *unimodular*, i.e., an integer square matrix with determinant ± 1 . It can be proven that a $k \times n$ primitive matrix can always be extended to an $n \times n$ unimodular matrix over \mathbb{Z} ; see, e.g., [1].

Given a primitive matrix $\mathbf{A} \in \mathbb{Z}^{k \times n}$ with $\|\mathbf{A}\| := \max_{i,j} |a_{i,j}|$ bounded by an integer λ from above, our focus in this paper will be on the probability of that the $m \times n$ matrix extended from \mathbf{A} by appending other $m - k$ vectors of dimension n with entries chosen randomly and independently from the uniform distribution over $\Lambda := \mathbb{Z} \cap [0, \lambda)$ is still primitive.

1.1 Main Results

In particular, we prove the following theorem.

Theorem 1.1 *Given a primitive matrix $\mathbf{A} \in \mathbb{Z}^{k \times n}$ with $\|\mathbf{A}\|$ bounded by an integer λ from above and an integer s with $0 \leq s \leq n - k - 2$, let $\mathbf{B} \in \mathbb{Z}^{(n-s-1) \times n}$ be a matrix with first k rows copied from \mathbf{A} and entries of the other rows chosen randomly and independently from the uniform distribution over Λ . Then the probability of the event that \mathbf{B} is primitive is at least*

$$1 - 4 \left(\frac{2}{3}\right)^{s+1} \left(1 - \left(\frac{2}{3}\right)^{n-k-s-1}\right) - \frac{2(n-s)^2}{\lambda^{s+2}} \left(1 - \frac{1}{\lambda^{n-k-s-1}}\right). \quad (1)$$

Note that Theorem 1.1 holds for $k = 0$ as well, which is the case of directly choosing $n - s - 1$ vectors with entries from the uniform distribution over Λ . Roughly speaking, for this case, it was shown by Maze, et al.^[1] that when $\lambda \rightarrow \infty$, the limit probability of that an $(n - s - 1) \times n$ integer matrix with entries random chosen from the uniform distribution over Λ is primitive is

$$\prod_{j=s+2}^n \frac{1}{\zeta(j)}, \quad (2)$$

where $\zeta(\cdot)$ is the Riemann's zeta function. In this sense, Theorem 1.1 gives an effective lower bound on the probability for finite λ , and hence will be useful in practice, especially in computer science.

From Equation (1), the parameter k plays a very limited role for the result. In fact, one may easily obtain a simpler but worse bound:

$$1 - 4 \left(\frac{2}{3}\right)^{s+1} - \frac{2(n-s)^2}{\lambda^{s+2}}. \quad (3)$$

Note that this bound is independent of the parameter k . For example, if $s = 3$ is fixed and $\lambda \geq 3(n-3)^{2/5}$, then the bound (3) implies that the resulting $(n-4) \times n$ matrix will be primitive with a probability at least 0.2.

Moreover, if λ is large enough (with respect to n), this bound can be further simplified as

$$1 - (4 + \delta) \left(\frac{2}{3}\right)^{s+1} \quad (4)$$

for some $0 < \delta < 1$. Surprisingly, this oversimplified bound only depends on s .

For given k, n and λ , one can decide the smallest integer $s \in [0, n - k - 2]$ such that the lower bound given in Equation (1) is a usable bound, i.e., between 0 and 1. For instance, s should be at least 3 for $4 \left(\frac{2}{3}\right)^{s+1} < 1$.

We remark that when $s = n - k - 2$ the probability bound given in Theorem 1.1 matches the empirical probability well according to our experiments in Section 4.

In addition, one may not further expect a constant probability for the case of $s = -1$ (that corresponds to the resulting matrix is an $n \times n$ unimodular matrix), since the natural density of random $n \times n$ unimodular matrices is 0; see [1, Lemma 5].

1.2 Implications

As an application of Theorem 1.1, we present a fast Las Vegas algorithm (Algorithm 2) that efficiently completes a primitive matrix $\mathbf{A} \in \mathbb{Z}^{k \times n}$ to an $n \times n$ unimodular matrix \mathbf{U} such that $\|\mathbf{U}\| \leq n^{O(1)} \|\mathbf{A}\|$. More specifically, we prove the following theorem.

Theorem 1.2 *Given a primitive matrix $\mathbf{A} \in \mathbb{Z}^{k \times n}$, there exists a Las Vegas algorithm that completes \mathbf{A} to an $n \times n$ unimodular matrix \mathbf{U} such that $\|\mathbf{U}\| \leq n^8 \|\mathbf{A}\|$ in an expected number of $O(n^{\omega+\varepsilon} \log^{1+\varepsilon} \|\mathbf{A}\|)$ bit operations.*

1.3 Techniques

The essential ingredient of our proof for Theorem 1.1 is adapted from [2, Section 6], which was used to analyze the expected number of nontrivial invariant factors of a random integer matrix. The main idea is to give an upper bound on the probability that the resulting $(n - s - 1) \times n$ matrix is not primitive. Based on the following lemma, the bound can be analyzed by localizing at each prime p .

Lemma 1.3 *A $k \times n$ integer matrix \mathbf{A} is not primitive if and only if there exists at least one prime number p such that the resulting matrix is not full rank over the finite field $\mathbb{Z}_p := \mathbb{Z}/p\mathbb{Z}$.*

In addition, the algorithm for unimodular matrix completion is based on the determinant reduction technique, which was originally introduced by Storjohann in [3, Section 15] for computing the determinant of a polynomial matrix, with a worked example for integer matrix followed. More details about determinant reduction for integer matrices and an iterated usage of this technique are given in [4, Section 13.2]. We give a full description of the algorithm for the integer matrix case and present a detailed analysis in Section 3.

1.4 Related Work

Primitive and unimodular matrices have many applications in different areas. For example, unimodular matrices can be applied to signal compression^[5]; the lattice reduction algorithms^[6, 7] essentially produce a series of unimodular matrices (linear transformations) to improve the basis quality of the given lattice.

In particular, generating a primitive or unimodular matrix with given rows or columns happens quite often in practice. For instance, one may need to generate unimodular matrices

with at least one column of all ones in linear programming for simplex pivoting. In the literature, there exist many results on what conditions should be satisfied for that a partial integral matrix can be completed to a unimodular matrix. In 1956, Reiner^[8] proved that a row vector can be completed to a unimodular matrix if and only if it is primitive. Zhan^[9] proved that if n entries of an $n \times n$ partial integral matrix are prescribed and these n entries do not constitute a row or a column, then this matrix can be completed to a unimodular matrix. Fang^[10] improved Zhan's result by proving that if an $n \times n$ partial integral matrix has $2n - 3$ prescribed entries and any n entries of these do not constitute a row or a column, then it can be completed to a unimodular matrix. Duffner and Silva^[11] gave necessary and sufficient conditions for the existence of unimodular matrices with a prescribed submatrix over a ring that either is Hermite and Dedekind finite or has stable range one.

For the probability analysis, Maze, et al.^[1] analyzed the natural density of $k \times n$ primitive matrices. Guo, et al.^[12] extended Maze, et al.'s result to a more general setting, where the natural density of $k \times n$ primitive matrices over all $m \times m$ ($m \geq \max\{k, n\}$) integer matrices was considered. Note that the natural density can be interpreted as a limit probability, where each matrix entry is randomly and independently chosen from a set with an upper bound but the bound tends to infinity. However, a finite version is usually preferable in practice, e.g., algorithm analysis. Therefore, the result in Theorem 1.1 will be useful.

Additionally, a somewhat "dual" case is considered in [13], where the probability of that m integral vectors with bounded entries generate a same lattice of rank n was studied. In [13], the ideal choice is $m = n + 1$, but a theoretical lower bound on the probability was only proven for $m \geq 2n + 1$. Aggarwal and Regev in [14] and Kirshanova, et al. in [15] considered a closely related problem but the entries are randomly chosen from the discrete Gaussian distribution over \mathbb{Z}^n , where m is even larger.

For algorithms, Randall^[16] presented an algorithm for generating random matrices over a finite field with a given determinant, which naturally can be used to generate unimodular matrices over finite fields. Kalaimani, et al.^[17], Zhou and Labahn^[18] discussed algorithms for unimodular completion of polynomial matrices. The unimodular matrix completion algorithm discussed in this paper works for integer matrices, which is more efficient than the standard method for this problem (see Remarks 3.4 and 3.5).

Roadmap We prove Theorem 1.1 in Section 2. Combining the determinant reduction technique for integer matrices with Theorem 1.1, we give a fast algorithm for the problem of unimodular matrix completion in Section 3. In Section 4, we present an extensive experimental study on the probability that the resulting $(n - s - 1) \times n$ matrix is primitive and discuss some interesting problems for further study.

2 Proof of Theorem 1.1

Giving a primitive matrix $\mathbf{A} \in \mathbb{Z}^{k \times n}$ with $k < n$ and $\|\mathbf{A}\| \leq \lambda$, we now consider to extend \mathbf{A} to an $(n - s - 1) \times n$ matrix by choosing other $n - k - s - 1$ vectors with entries randomly and independently chosen from the uniform distribution over $\Lambda = \mathbb{Z} \cap [0, \lambda)$, where

the integer s satisfies $0 \leq s \leq n - k - 2$. Denote by $\mathbf{a}_i = (a_{i,j})_{1 \leq j \leq n}$ the i -th row of \mathbf{A} . Then $\|\mathbf{a}_i\|_\infty := \max_j \{|a_{i,j}|\} \leq \lambda$. We always assume that $\lambda \geq 2$ for excluding the case of $A = \{0\}$. For convenience, we still use $\mathbf{a}_{k+1}, \dots, \mathbf{a}_{n-s-1} \in \mathbb{Z}^n$ to denote the random vectors with each entry chosen randomly and independently from the uniform distribution over A , and denote

$$\mathbf{A}_i = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_i \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i,1} & a_{i,2} & \cdots & a_{i,n} \end{pmatrix}, \quad i = k, k + 1, \dots, n - s - 1.$$

To prove Theorem 1.1, we firstly need to bound from above the probability P of the event that the matrix \mathbf{A}_{n-s-1} is not primitive under the assumption that the given matrix \mathbf{A} is primitive.

Lemma 2.1 *Let all notations be as above. Then*

$$P \leq 4 \left(\frac{2}{3}\right)^{s+1} \left(1 - \left(\frac{2}{3}\right)^{n-k-s-1}\right) + \frac{2(n-s)^2}{\lambda^{s+2}} \left(1 - \frac{1}{\lambda^{n-k-s-1}}\right).$$

We prove Lemma 2.1 following the approach of Eberly, et al. (see [2, Section 6]), whose original goal was to bound the expected number of invariant factors for random integer matrices. For $k \leq i \leq n - s - 1$, we define the event[†]

- **MDep_{*i*}**: There exists at least one prime number p such that $\text{rank}(\mathbf{A}_i) \leq i - 1$ over \mathbb{Z}_p .

So the assumption that \mathbf{A} is primitive is equivalent to that the event $\neg\text{MDep}_k$ happens. Under the assumption, we have

$$\begin{aligned} P &= \Pr[\text{MDep}_{n-s-1}] \\ &\leq \Pr[\text{MDep}_{k+1} \vee \text{MDep}_{k+2} \vee \cdots \vee \text{MDep}_{n-s-1}] \\ &= \Pr[\text{MDep}_{k+1} \vee (\text{MDep}_{k+2} \wedge \neg\text{MDep}_{k+1}) \vee \cdots \vee (\text{MDep}_{n-s-1} \wedge \neg\text{MDep}_{n-s-2})] \\ &= \Pr[(\text{MDep}_{k+1} \wedge \neg\text{MDep}_k) \vee (\text{MDep}_{k+2} \wedge \neg\text{MDep}_{k+1}) \\ &\quad \vee \cdots \vee (\text{MDep}_{n-s-1} \wedge \neg\text{MDep}_{n-s-2})] \tag{5} \\ &\leq \sum_{i=k+1}^{n-s-1} \Pr[\text{MDep}_i \wedge \neg\text{MDep}_{i-1}] \\ &\leq \sum_{i=k+1}^{n-s-1} \Pr[\text{MDep}_i \mid \neg\text{MDep}_{i-1}]. \end{aligned}$$

In order to bound $\Pr[\text{MDep}_i \mid \neg\text{MDep}_{i-1}]$ for $k + 1 \leq i \leq n - s - 1$, we need to introduce another useful event:

- **Dep_{*i*}**: The rows of \mathbf{A}_i are linearly dependent over \mathbb{Q} , i.e., $\text{rank}(\mathbf{A}_i) \leq i - 1$ over \mathbb{Q} .

[†]The definition of MDep_i here is different from that in [2, Section 6], where MDep_i denotes the event that there exists at least one prime p such that $\text{rank}(\mathbf{a}_1, \dots, \mathbf{a}_i) \leq i - 2$ over \mathbb{Z}_p .

Now we have

$$\begin{aligned}
 & \Pr[\text{MDep}_i \mid \neg\text{MDep}_{i-1}] \\
 &= \Pr[\text{MDep}_i \wedge (\text{Dep}_i \vee \neg\text{Dep}_i) \mid \neg\text{MDep}_{i-1}] \\
 &= \Pr[(\text{MDep}_i \wedge \text{Dep}_i) \vee (\text{MDep}_i \wedge \neg\text{Dep}_i) \mid \neg\text{MDep}_{i-1}] \\
 &\leq \Pr[(\text{MDep}_i \wedge \text{Dep}_i) \mid \neg\text{MDep}_{i-1}] + \Pr[(\text{MDep}_i \wedge \neg\text{Dep}_i) \mid \neg\text{MDep}_{i-1}].
 \end{aligned}
 \tag{6}$$

We first bound

$$\Pr[(\text{MDep}_i \wedge \text{Dep}_i) \mid \neg\text{MDep}_{i-1}] \leq \Pr[\text{Dep}_i \mid \neg\text{MDep}_{i-1}].$$

The latter is the probability of that $\text{rank}(\mathbf{A}_i) = i - 1$ over \mathbb{Q} on the condition that \mathbf{A}_{i-1} is primitive. From \mathbf{A}_{i-1} is primitive, it follows that there must exist $i - 1$ columns of \mathbf{A}_{i-1} such that the submatrix consisting of the first $i - 1$ rows of these $i - 1$ columns has rank $i - 1$ over \mathbb{Q} . Denote by the set of indices of these columns C_{i-1} . Now $\text{rank}(\mathbf{A}_i) = i - 1$ over \mathbb{Q} implies that for all $j \notin C_{i-1}$, the entry $a_{i,j}$ must be a linear combination of $a_{\ell,j}$'s for $\ell = 1, \dots, i - 1$ with the same rational coefficients that determined by those $a_{i,j}$'s with $j \in C_{i-1}$. However, each entry of \mathbf{a}_i are chosen randomly and independently from the uniform distribution over Λ . Thus, for each $j \notin C_{i-1}$, the likelihood that $a_{i,j}$ is equal to such a rational linear combination is either 0 or $\frac{1}{\lambda}$. Therefore,

$$\Pr[(\text{MDep}_i \wedge \text{Dep}_i) \mid \neg\text{MDep}_{i-1}] \leq \Pr[\text{Dep}_i \mid \neg\text{MDep}_{i-1}] \leq \left(\frac{1}{\lambda}\right)^{n-i+1}.
 \tag{7}$$

To bound $\Pr[(\text{MDep}_i \wedge \neg\text{Dep}_i) \mid \neg\text{MDep}_{i-1}]$, let us consider primes $p < \lambda$ and primes $p \geq \lambda$, respectively. For that, we define the following events

- $\text{MDep}_i^{(p)}$: For prime p , $\text{rank}(\mathbf{A}_i) \leq i - 1$ over \mathbb{Z}_p .
- $\text{MDep}_i^{(p < \lambda)}$: There exists a prime $p < \lambda$ such that $\text{rank}(\mathbf{A}_i) \leq i - 1$ over \mathbb{Z}_p .

2.1 The Case of $p < \lambda$

We bound $\Pr[(\text{MDep}_i^{(p < \lambda)} \wedge \text{Dep}_i) \mid \neg\text{MDep}_{i-1}]$ by $\Pr[\text{MDep}_i^{(p < \lambda)} \mid \neg\text{MDep}_{i-1}]$. The latter one is the probability that there exists a prime $p < \lambda$ such that $\text{rank}(\mathbf{A}_i) = i - 1$ over \mathbb{Z}_p under the condition that \mathbf{A}_{i-1} is primitive. Now we bound the latter one case by case.

If $\lambda = 2$, then no such prime p exists.

If $\lambda = 3$, then $p = 2$. Furthermore, the event $(\text{MDep}_i^{(p < 3)} \mid \neg\text{MDep}_{i-1})$ is that $\text{rank}(\mathbf{A}_i) = i - 1$ over \mathbb{Z}_2 assuming $\text{rank}(\mathbf{A}_{i-1}) = i - 1$ over \mathbb{Z}_p for any prime p . It follows from the assumption that there exist $i - 1$ columns of \mathbf{A}_i such that the submatrix consisting of the first $i - 1$ rows of these $i - 1$ columns has rank $i - 1$ over \mathbb{Z}_2 . Denote by the set of indices of these columns C_{i-1} . Now, $\text{rank}(\mathbf{A}_i) = i - 1$ for $p = 2$ means that for each $j \notin C_{i-1}$, the entry $a_{i,j}$ must be a linear combination of $a'_{\ell,j}$ s over \mathbb{Z}_2 for $\ell = 1, \dots, i - 1$. However, when $\lambda = 3$ and $p = 2$, we have

$$\Pr_{x \leftarrow \Lambda}[x \equiv 0 \pmod{p}] = \frac{2}{3}$$

and

$$\Pr_{x \leftarrow \Lambda} [x \equiv 1 \pmod p] = \frac{1}{3},$$

so for $\lambda = 3$ it follows that

$$\Pr \left[\text{MDep}_i^{(p < 3)} \mid \neg \text{MDep}_{i-1} \right] \leq \left(\frac{2}{3} \right)^{n-i+1}.$$

If $\lambda = 4$, then $p = 2$ or $p = 3$. Similarly, we have

$$\Pr \left[\text{MDep}_i^{(p < 4)} \mid \neg \text{MDep}_{i-1} \right] \leq 2 \left(\frac{1}{2} \right)^{n-(i-1)} \leq \left(\frac{2}{3} \right)^{n-i+1}$$

for $i < n - 1$.

If $\lambda = 5$, then $p = 2$ or $p = 3$, and further we obtain

$$\Pr \left[\text{MDep}_i^{(p < 5)} \mid \neg \text{MDep}_{i-1} \right] \leq \left(\frac{3}{5} \right)^{n-(i-1)} + \left(\frac{2}{5} \right)^{n-(i-1)} \leq \left(\frac{2}{3} \right)^{n-i+1}$$

for $i < n - 1$.

If $\lambda = 6$, then $p = 2$ or $p = 3$ or $p = 5$ and we have

$$\Pr \left[\text{MDep}_i^{(p < 6)} \mid \neg \text{MDep}_{i-1} \right] \leq \left(\frac{1}{2} \right)^{n-i+1} + 2 \left(\frac{1}{3} \right)^{n-(i-1)} \leq \left(\frac{2}{3} \right)^{n-i+1}$$

for $i < n - 1$.

If $\lambda = 7$, then $p = 2$ or $p = 3$ or $p = 5$. It follows that

$$\Pr \left[\text{MDep}_i^{(p < 7)} \mid \neg \text{MDep}_{i-1} \right] \leq \left(\frac{4}{7} \right)^{n-i+1} + \left(\frac{3}{7} \right)^{n-(i-1)} + \left(\frac{2}{7} \right)^{n-(i-1)} \leq \left(\frac{2}{3} \right)^{n-i+1}$$

for $i < n - 1$.

If $\lambda \geq 8$, then $p = 2$ or $p = 3$ or $p = 5$ or $p = 7$, etc. Then for $i < n - 1$, we have

$$\begin{aligned} \Pr \left[\text{MDep}_i^{(p < \lambda)} \mid \neg \text{MDep}_{i-1} \right] &\leq \left(\frac{1}{2} \right)^{n-i+1} + \left(\frac{3}{8} \right)^{n-(i-1)} + \left(\frac{1}{4} \right)^{n-(i-1)} + \sum_{p \geq 7} \left(\frac{2}{p-1} \right)^{n-i+1} \\ &\leq \left(\frac{2}{3} \right)^{n-i+1} + \left(\frac{1}{3} \right)^{n-i+1} \cdot \sum_{p \geq 7} \frac{4}{(p-1)^2} \\ &\leq \left(\frac{2}{3} \right)^{n-i+1} + 4 \cdot \left(\frac{1}{3} \right)^{n-i+1} \cdot \sum_{j=6}^{\infty} \frac{1}{j^2} \\ &\leq \left(\frac{2}{3} \right)^{n-i+1} + 4 \cdot \left(\frac{1}{3} \right)^{n-i+1} \cdot \left(\zeta(2) - 1 - \frac{1}{4} - \frac{1}{9} - \frac{1}{16} - \frac{1}{25} \right) \\ &\leq \left(\frac{2}{3} \right)^{n-i+1} + \frac{3}{4} \left(\frac{1}{3} \right)^{n-i+1}, \end{aligned}$$

where $\zeta(\cdot)$ is Riemman’s zeta function and the fact that $\lceil \lambda/p \rceil / \lambda$ does not increase with respect to λ and that $\lceil \lambda/p \rceil / \lambda \leq 2/(p - 1)$ for $p < \lambda$ were used. Therefore, for the case of $p < \lambda$ we proved the following proposition.

Proposition 2.2 *Let $\lambda \geq 2$ be an integer and $k+1 \leq i \leq n-3$, and suppose that the event $\neg\text{MDep}_{i-1}$ happens. The probability that there exists any prime $p < \lambda$ such that $\text{rank}(\mathbf{A}_i) \leq i-1$ over \mathbb{Z}_p is at most $\left(\frac{2}{3}\right)^{n-i+1} + \frac{3}{4} \left(\frac{1}{3}\right)^{n-i+1}$. In particular,*

$$\Pr \left[(\text{MDep}_i^{(p < \lambda)} \wedge \neg\text{Dep}_i) \mid \neg\text{MDep}_{i-1} \right] \leq \left(\frac{2}{3}\right)^{n-i+1} + \frac{3}{4} \left(\frac{1}{3}\right)^{n-i+1}. \tag{8}$$

2.2 The Case of $p \geq \lambda$

Now our goal is to bound

$$\begin{aligned} & \Pr \left[(\text{MDep}_i^{(p \geq \lambda)} \wedge \neg\text{Dep}_i) \mid \neg\text{MDep}_{i-1} \right] \\ &= \Pr \left[\left(\bigvee_{p \geq \lambda} \text{MDep}_i^{(p)} \right) \wedge \neg\text{Dep}_i \mid \neg\text{MDep}_{i-1} \right] \\ &= \Pr \left[\bigvee_{p \geq \lambda} (\text{MDep}_i^{(p)} \wedge \neg\text{Dep}_i) \mid \neg\text{MDep}_{i-1} \right] \\ &\leq \sum_{p \geq \lambda} \Pr \left[(\text{MDep}_i^{(p)} \wedge \neg\text{Dep}_i) \mid \neg\text{MDep}_{i-1} \right], \end{aligned} \tag{9}$$

where p ranges all primes at least λ .

If $p \geq \lambda$ is a fixed prime, then the probability that $\text{rank}(\mathbf{A}_i) \leq i-1$ over \mathbb{Z}_p under the condition $\text{rank}(\mathbf{A}_{i-1}) = i-1$ over \mathbb{Q} is at most $\left(\frac{1}{\lambda}\right)^{n-i+1}$, since the probability that a value chosen randomly and independently from the uniform distribution on $\Lambda = \mathbb{Z} \cap [0, \lambda)$ equals a given value in \mathbb{Z}_p is either 0 or $\frac{1}{\lambda}$, and hence at most $\frac{1}{\lambda}$. So we have

$$\Pr \left[(\text{MDep}_i^{(p)} \wedge \neg\text{Dep}_i) \mid \neg\text{MDep}_{i-1} \right] \leq \Pr \left[\text{MDep}_i^{(p)} \mid \neg\text{MDep}_{i-1} \right] \leq \left(\frac{1}{\lambda}\right)^{n-i+1}. \tag{10}$$

Furthermore, we claim that

$$\Pr \left[(\text{MDep}_i^{(p)} \wedge \neg\text{Dep}_i) \mid \neg\text{MDep}_{i-1} \right] = 0 \tag{11}$$

for all primes $p > (i \cdot \lambda)^i$.

In fact, if there exists a prime $p > (i \cdot \lambda)^i$ that makes the event $(\text{MDep}_i^{(p)} \wedge \neg\text{Dep}_i)$ happen assuming that \mathbf{A}_{i-1} is primitive, then p must divide all $i \times i$ minors of \mathbf{A}_i . However, although $a_{i,j}$'s ($j = 1, 2, \dots, n$) are chosen randomly and independently from the uniform distribution on Λ , once $\neg\text{Dep}_i$ happens, there exists at least one nonsingular $i \times i$ submatrix in \mathbf{A}_i over \mathbb{Q} . The absolute value of the determinant of such a submatrix is at most $i! \cdot \lambda^i \leq (i \cdot \lambda)^i$. Combining $p > (i \cdot \lambda)^i$, p divides the determinant, and the absolute value of the determinant is bounded by $(i \cdot \lambda)^i$, we have the determinant must be zero, which contradicts with non-singularity of the submatrix. Also, note that the bound $(i \cdot \lambda)^i$ is independent of the choice of $a_{i,j}$'s when such $a_{i,j}$'s make $\neg\text{Dep}_i$ happen (Instead, when such $a_{i,j}$'s do not make $\neg\text{Dep}_i$ happen, the probability $\Pr[(\text{MDep}_i \wedge \text{Dep}_i) \mid \neg\text{MDep}_{i-1}]$ is already discussed in (7)).

Now, the number of possible primes $p \geq \lambda$ that divides an integer $d \leq (i \cdot \lambda)^i$ is at most $\log_\lambda(i \cdot \lambda)^i \leq i(1 + \log_\lambda i)$, where d is in fact the greatest common divisor of all nonsingular $i \times i$ submatrix of \mathbf{A}_i over \mathbb{Q} . Although the integer d may vary as $a_{i,j}$'s are chosen randomly and independently from the uniform distribution on Λ , it is always bounded by $(i \cdot \lambda)^i$. Thus, the number of possible primes that we need to consider is at most $i(1 + \log_\lambda i)$. Thus, combining Equations (9)–(11), we have

$$\begin{aligned} \Pr \left[(\text{MDep}_i^{(p \geq \lambda)} \wedge \neg \text{Dep}_i) \mid \neg \text{MDep}_{i-1} \right] &\leq \sum_{p \geq \lambda} \Pr \left[(\text{MDep}_i^{(p)} \wedge \neg \text{Dep}_i) \mid \neg \text{MDep}_{i-1} \right] \\ &\leq (i(1 + \log_\lambda i)) \cdot \left(\frac{1}{\lambda} \right)^{n-i+1}. \end{aligned} \tag{12}$$

Therefore, it follows from Equations (8) and (12) that

$$\begin{aligned} &\Pr \left[(\text{MDep}_i \wedge \neg \text{Dep}_i) \mid \neg \text{MDep}_{i-1} \right] \\ &= \Pr \left[(\text{MDep}_i^{(p < \lambda)} \wedge \neg \text{Dep}_i) \mid \neg \text{MDep}_{i-1} \right] + \Pr \left[(\text{MDep}_i^{(p \geq \lambda)} \wedge \neg \text{Dep}_i) \mid \neg \text{MDep}_{i-1} \right] \\ &\leq \left(\frac{2}{3} \right)^{n-i+1} + \frac{3}{4} \left(\frac{1}{3} \right)^{n-i+1} + (i(1 + \log_\lambda i)) \cdot \left(\frac{1}{\lambda} \right)^{n-i+1}. \end{aligned} \tag{13}$$

Proof of Lemma 2.1 It follows from Equations (6), (7), and (13) that for integer $\lambda \geq 2$,

$$\begin{aligned} \Pr[\text{MDep}_i \mid \neg \text{MDep}_{i-1}] &\leq \Pr[(\text{MDep}_i \wedge \text{Dep}_i) \mid \neg \text{MDep}_{i-1}] + \Pr[(\text{MDep}_i \wedge \neg \text{Dep}_i) \mid \neg \text{MDep}_{i-1}] \\ &\leq \left(\frac{1}{\lambda} \right)^{n-i+1} + \left(\frac{2}{3} \right)^{n-i+1} + \frac{3}{4} \left(\frac{1}{3} \right)^{n-i+1} + (i(1 + \log_\lambda i)) \cdot \left(\frac{1}{\lambda} \right)^{n-i+1} \\ &\leq \left(\frac{2}{3} \right)^{n-i+1} + \frac{3}{4} \left(\frac{1}{3} \right)^{n-i+1} + (i(1 + \log_\lambda i) + 1) \cdot \left(\frac{1}{\lambda} \right)^{n-i+1}. \end{aligned}$$

Therefore, under the assumption that $\neg \text{MDep}_k$ happens, it follows from Equation (5) that

$$\begin{aligned} P &= \Pr[\text{MDep}_{n-s-1}] \\ &\leq \sum_{i=k+1}^{n-s-1} \Pr[\text{MDep}_i \mid \neg \text{MDep}_{i-1}] \\ &\leq \sum_{i=k+1}^{n-s-1} \left(\left(\frac{2}{3} \right)^{n-i+1} + \frac{3}{4} \left(\frac{1}{3} \right)^{n-i+1} + (i(1 + \log_\lambda i) + 1) \cdot \left(\frac{1}{\lambda} \right)^{n-i+1} \right) \\ &\leq 2 \cdot \sum_{i=k+1}^{n-s-1} \left(\frac{2}{3} \right)^{n-i+1} + (n-s)(1 + \log_\lambda(n-s-1)) \sum_{i=k+1}^{n-s-1} \frac{1}{\lambda^{n-i+1}} \\ &\leq 4 \left(\frac{2}{3} \right)^{s+1} \left(1 - \left(\frac{2}{3} \right)^{n-k-s-1} \right) + \frac{(n-s)^2}{(\lambda-1)\lambda^{s+1}} \left(1 - \frac{1}{\lambda^{n-k-s-1}} \right) \\ &\leq 4 \left(\frac{2}{3} \right)^{s+1} \left(1 - \left(\frac{2}{3} \right)^{n-k-s-1} \right) + \frac{2(n-s)^2}{\lambda^{s+2}} \left(1 - \frac{1}{\lambda^{n-k-s-1}} \right). \end{aligned} \quad \blacksquare$$

Now, Theorem 1.1 is a direct consequence of Lemma 2.1.

Remark 2.3 According to the above proof, we use localization at each prime p and count the number of possibilities over \mathbb{Z}_p . So the set A can be any set containing λ contiguous integers, as in [2, Section 6].

3 Unimodular Matrix Completion

When completing a given $k \times n$ primitive matrix to an $n \times n$ unimodular matrix, we first complete the input matrix to an $n \times n$ matrix with $n - k$ vectors whose entries are chosen randomly and uniformly from the uniform distribution over A , and then rectify the last four vectors via the determinant reduction technique (see [3, Section 15]). Repeat the above process until the resulting matrix is unimodular.

Equation (3) (a consequence of Theorem 1.1) with $s = 3$ and $\lambda \geq 3(n - 3)^{2/5}$ guarantees that the completed $(n - 4) \times n$ matrix is still primitive with a probability at least a constant, say 0.2. Then after a constant number of repeat, it is expected that we will obtain at least one $(n - 4) \times 4$ primitive matrix, and the determinant reduction technique applies to produce a unimodular matrix.

3.1 Hermite Normal Form

A matrix $\mathbf{A} \in \mathbb{Z}^{m \times n}$ of rank r has a (row) *Hermite normal form* (HNF) if there exists a square unimodular matrix \mathbf{U} such that $\mathbf{H} = \mathbf{U}\mathbf{A}$ satisfies the following: There exist indices $1 \leq i_1 < i_2 < \dots < i_r \leq n$ such that for $j = 1, \dots, r$, $h_{j,i_j} > 0$, $h_{j,k} = 0$ if $k < i_j$ and $0 \leq h_{\ell,i_j} < h_{j,i_j}$ if $\ell < j$; the bottom $m - r$ rows of \mathbf{H} are zero.

The matrix \mathbf{H} is unique, denoted by $\text{HNF}(\mathbf{A})$, although the unimodular transformations are usually not unique. With the definition of HNF, we can character the primitive matrix as the following lemma.

Lemma 3.1 A $k \times n$ integer matrix \mathbf{A} with $k < n$ is primitive iff $\text{HNF}(\mathbf{A}^T) = \begin{pmatrix} I_k \\ \mathbf{0} \end{pmatrix}$.

One can find a proof of this lemma in, e.g., [1].

3.2 Determinant Reduction

Given a nonsingular $\mathbf{A} \in \mathbb{Z}^{n \times n}$, the determinant reduction introduced in [3, Section 15] computes a matrix $\mathbf{B} \in \mathbb{Z}^{n \times n}$, obtained from \mathbf{A} by replacing the last column, such that the last diagonal entry in the Hermite normal form of \mathbf{B} is one. The determinant reduction was originally presented for integral polynomial matrix in [3, Section 15], with a worked example for integer matrix. Here, we give full details and analyses for the integer matrix case.

Proposition 3.2 Given a nonsingular $n \times n$ integer matrix \mathbf{A} , Algorithm 1 correctly computes an $n \times n$ integer matrix \mathbf{B} within $O(n^{\omega+\varepsilon} \log^{1+\varepsilon} \|\mathbf{A}\|)$ bit operations, where \mathbf{B} satisfies the following:

- \mathbf{B} equals \mathbf{A} except for possibly the last column;
- The last diagonal entry of $\text{HNF}(\mathbf{B})$ equals one;
- $\|\mathbf{B}\| = O(n^2 \|\mathbf{A}\|)$.

Algorithm 1 (Determinant reduction)

Input: A nonsingular integer matrix $\mathbf{A} \in \mathbb{Z}^{n \times n}$.

Output: A matrix $\mathbf{B} \in \mathbb{Z}^{n \times n}$, with \mathbf{B} equals to \mathbf{A} except for possibly the last column, $\|\mathbf{B}\| \leq n^2 \|\mathbf{A}\|$, and the last diagonal entry of $\text{HNF}(\mathbf{B})$ equal to one.

- 1: Set \mathbf{C}_{n-1} to be the matrix consisting of the first $n - 1$ columns of \mathbf{A} . Compute a primitive vector $\mathbf{u} \in \mathbb{Z}^n$ such that $\mathbf{u}\mathbf{C}_{n-1} = \mathbf{0}$.
- 2: Call an extended gcd algorithm to compute $\mathbf{b} \in \mathbb{Z}^n$ such that $\mathbf{u}\mathbf{b}^T = 1$.
- 3: Set $\overline{\mathbf{A}}$ to be the $(n - 1) \times (n - 1)$ principal submatrix of \mathbf{A} , and $\overline{\mathbf{b}}$ the vector consisting of the first $n - 1$ entries of \mathbf{b} .
- 4: Compute $\overline{\mathbf{q}}^T := \lceil \overline{\mathbf{A}}^{-1} \overline{\mathbf{b}}^T \rceil$. For a vector \mathbf{v} , $\lceil \mathbf{v} \rceil$ means each entry rounded.
- 5: Set $\mathbf{q} := (\overline{\mathbf{q}}, 0)$ and set \mathbf{B} to be \mathbf{A} except replacing the last column by $\mathbf{b}^T - \mathbf{A}\mathbf{q}^T$.
- 6: **return** \mathbf{B} .

Proof The singularity of \mathbf{A} implies that the vector \mathbf{u} produced in Step 1 is the unique primitive vector in $\{\mathbf{x} \in \mathbb{Z}^n : \mathbf{x}\mathbf{C}_{n-1} = \mathbf{0}\}$. Let \mathbf{U} be an arbitrary unimodular matrix such that $\mathbf{H} := \text{HNF}(\mathbf{A}) = \mathbf{U}\mathbf{A}$. Then the uniqueness of \mathbf{u} implies that \mathbf{u} must be the last row of \mathbf{U} . By construction of \mathbf{b} in Step 2, the matrix obtained from \mathbf{A} by replacing the last column with \mathbf{b}^T will have HNF with the last diagonal entry one. The whole algorithm can be expressed as the following equation

$$\mathbf{B} = \begin{pmatrix} \overline{\mathbf{A}} & \overline{\mathbf{b}}^T \\ \overline{\mathbf{a}} & b_n \end{pmatrix} \cdot \begin{pmatrix} \mathbf{I}_{n-1} & -\overline{\mathbf{q}}^T \\ & 1 \end{pmatrix} = \begin{pmatrix} \overline{\mathbf{A}} & \overline{\mathbf{b}}^T - \overline{\mathbf{A}}\overline{\mathbf{q}}^T \\ \overline{\mathbf{a}} & b_n - \overline{\mathbf{a}}\overline{\mathbf{q}}^T \end{pmatrix},$$

where $\overline{\mathbf{A}}$ is the $(n - 1) \times (n - 1)$ principal submatrix of \mathbf{A} , $\overline{\mathbf{b}}$ is the vector consisting of the first $n - 1$ entries of \mathbf{b} , and $\overline{\mathbf{a}}$ is the vector consisting of the first $n - 1$ entries of the last row of \mathbf{A} . Steps 3–5 are to reduce the size of $\overline{\mathbf{b}}$ by the columns of $\overline{\mathbf{A}}$. It follows from $\mathbf{u}\mathbf{C}_{n-1} = \mathbf{0}$ that all entries of $\mathbf{u}\mathbf{A}$ are zero except for possibly the last one. However, the last entry of \mathbf{q} equals zero, so we have $\mathbf{u}\mathbf{A}\mathbf{q}^T = 0$, and hence

$$\mathbf{u}(\mathbf{b}^T - \mathbf{A}\mathbf{q}^T) = \mathbf{u}\mathbf{b}^T = 1, \tag{14}$$

which implies that $\mathbf{U}\mathbf{B}$ is exactly \mathbf{H} except for the last column replaced by $(*, \dots, *, 1)^T$, and hence the last diagonal of $\text{HNF}(\mathbf{B})$ must be one.

We now consider the size of $\mathbf{b}^T - \mathbf{A}\mathbf{q}^T$. First,

$$\begin{aligned} \|\overline{\mathbf{b}}^T - \overline{\mathbf{A}}\overline{\mathbf{q}}^T\|_\infty &= \|\overline{\mathbf{b}}^T - \overline{\mathbf{A}} \lceil \overline{\mathbf{A}}^{-1} \overline{\mathbf{b}}^T \rceil\|_\infty \\ &= \|\overline{\mathbf{b}}^T - \overline{\mathbf{A}} (\overline{\mathbf{A}}^{-1} \overline{\mathbf{b}}^T + \boldsymbol{\varepsilon}^T)\|_\infty \\ &\leq \|\overline{\mathbf{A}}\boldsymbol{\varepsilon}^T\|_\infty \\ &\leq \frac{n-1}{2} \|\overline{\mathbf{A}}\|, \end{aligned} \tag{15}$$

where $\|\cdot\|_\infty$ is the ℓ_∞ -norm of a vector and $\|\boldsymbol{\varepsilon}^T\|_\infty \leq \frac{1}{2}$ is used. In fact, Step 4 is essentially the same as Babai’s rounding algorithm^[19]. Denote $\mathbf{u} = (\overline{\mathbf{u}}, u_n)$. Without loss of generality,

we can assume that

$$\|\mathbf{u}\|_\infty = u_n. \tag{16}$$

Otherwise there must exist a permutation matrix \mathbf{P} such that $\mathbf{uPP}^{-1}\mathbf{C}_{n-1} = \mathbf{0}$, \mathbf{uP} satisfies Equation (16), and $\mathbf{P}^{-1}\mathbf{C}_{n-1}$ still corresponds to the first $n - 1$ columns of \mathbf{A} but with a certain column permutation. From Equation (14), we have $u_n(b_n - \bar{\mathbf{a}}\bar{\mathbf{q}}^T) = 1 - \bar{\mathbf{u}}(\bar{\mathbf{b}}^T - \bar{\mathbf{A}}\bar{\mathbf{q}}^T)$, combining Equation (16), which gives

$$|b_n - \bar{\mathbf{a}}\bar{\mathbf{q}}^T| = \frac{1}{|u_n|} \cdot |1 - \bar{\mathbf{u}}(\bar{\mathbf{b}}^T - \bar{\mathbf{A}}\bar{\mathbf{q}}^T)| \leq 1 + \frac{\|\bar{\mathbf{u}}\|_\infty}{|u_n|} \cdot \|\bar{\mathbf{b}}^T - \bar{\mathbf{A}}\bar{\mathbf{q}}^T\|_1 \leq 1 + \frac{(n-1)^2}{2} \|\bar{\mathbf{A}}\|,$$

where $\|\cdot\|_1$ is the ℓ_1 -norm of a vector, the first inequality follows from that $|\mathbf{xy}^T| \leq \|\mathbf{x}\|_\infty \cdot \|\mathbf{y}\|_1$ holds for any two vectors \mathbf{x} and \mathbf{y} of the same dimension, and the second inequality follows from Equations (15), (16), and $\|\mathbf{x}\|_1 \leq n\|\mathbf{x}\|_\infty$ for all n -dimensional vector \mathbf{x} . Therefore, the resulting matrix \mathbf{B} satisfies $\|\mathbf{B}\| \leq n^2\|\mathbf{A}\|$.

The cost of Algorithm 1 consists in nonsingular rational linear system solving (Steps 1 and 4) that can be finished by a Las Vegas algorithm in an expected number of $O(n^{\omega+\varepsilon} \log^{1+\varepsilon} \|\mathbf{A}\|)$ bit operations^[20] and an extended gcd computation (Step 2) that can be accomplished within $O(n^{2+\varepsilon} \log^{1+\varepsilon} \|\mathbf{A}\|)$ bit operations (see [4, Section 13.2]). Totally, Algorithm 1 costs at most $O(n^{\omega+\varepsilon} \log^{1+\varepsilon} \|\mathbf{A}\|)$ bit operations. ■

Corollary 3.3 *Given an $(n-1) \times n$ primitive matrix \mathbf{A} , there exists a Las Vegas algorithm which completes \mathbf{A} to an $n \times n$ unimodular matrix in an expected number of $O(n^{\omega+\varepsilon} \log^{1+\varepsilon} \|\mathbf{A}\|)$ bit operations.*

Proof Since \mathbf{A} is primitive, Lemma 3.1 implies that $\text{HNF}(\mathbf{A}^T)$ has the form $(\mathbf{I}_{n-1} \ \mathbf{0})$. Now, one can use the matrix $\mathbf{B} = \begin{pmatrix} \mathbf{A} \\ \mathbf{a}_n \end{pmatrix}^T$ as the input for Algorithm 1, where \mathbf{a}_n can be an arbitrary integer vector such that the resulting square matrix is nonsingular. Then, $\text{HNF}(\mathbf{B})$ has the same form as the identity matrix of order n except for the last column. Thus, the transpose of the output matrix of Algorithm 1 will be a unimodular completion of \mathbf{A} . ■

Remark 3.4 To complete an $(n - 1) \times n$ primitive matrix to an $n \times n$ unimodular matrix, a standard method is the following: Firstly compute n determinants of all $(n - 1) \times (n - 1)$ submatrix of the input matrix and then invoking an extended euclidean algorithm will give the information of the last row. However, this standard method can be finished in an expected number of $O(n^{\omega+1+\varepsilon} \log^{1+\varepsilon} \|\mathbf{B}\|)$ bit operations, even using the fast Las Vegas algorithm in [4] for computing the determinant of an integer matrix in an expected number of $O(n^{\omega+\varepsilon} \log^{1+\varepsilon} \|\mathbf{B}\|)$ bit operations.

3.3 Proof of Theorem 1.2

We prove Theorem 1.2 by Algorithm 2, which uses the Iterated Determinant Reduction presented in [4, Section 13.2]. Given a nonsingular matrix $\mathbf{A} \in \mathbb{Z}^{n \times n}$, the Iterated Determinant Reduction technique computes a matrix $\mathbf{B} \in \mathbb{Z}^{n \times n}$, obtained from \mathbf{A} by replacing the last d column, such that the last d diagonal entries in the Hermite normal form of \mathbf{B} is one.

The algorithm consists of d times calling of Algorithm 1, each followed with multiplying by a permutation matrix $\mathbf{P} = \begin{pmatrix} 0 & \mathbf{I}_{n-1} \\ 1 & 0 \end{pmatrix}$ from right. Now we are ready to present our algorithm for unimodular matrix completion (Algorithm 2).

Algorithm 2 (Unimodular matrix completion)

Input: A primitive matrix $\mathbf{A} \in \mathbb{Z}^{k \times n}$ with $k < n$ and $n \geq 5$.

Output: A unimodular matrix $\mathbf{U} \in \mathbb{Z}^{n \times n}$, with \mathbf{U} equal to \mathbf{A} except for the last $n - k$ rows and $\|\mathbf{U}\| \leq n^8 \|\mathbf{A}\|$.

- 1: Repeat
 - 2: Set $\lambda := \max\{\|\mathbf{A}\|, \lceil 3(n - 3)^{2/5} \rceil\}$ and $\mathbf{B} := \mathbf{A}^T$.
 - 3: Complete \mathbf{B} as an $n \times n$ matrix with entries chosen randomly and independently from the uniform distribution over $\{0, 1, \dots, \lambda - 1\}$.
 - 4: Repeat the following four times:
 - 5: Set \mathbf{B} as the output of Algorithm 1 with input \mathbf{B} and then set $\mathbf{B} := \mathbf{B}\mathbf{P}$.
 - 6: Until $\det(\mathbf{B}) = \pm 1$
 - 7: **return** $\mathbf{U} := (\mathbf{B}(\mathbf{P}^{-1})^4)^T$.
-

Denote \mathbf{B}_i by the submatrix of \mathbf{B} consisting of the first i columns. The goal of the loop in Step 4 is to make the last four diagonal entries of the Hermite normal form of \mathbf{B} be one. First, we assume that \mathbf{B}_{n-4}^T is primitive at the end of Step 3. Then Lemma 3.1 implies that

$$\text{HNF}(\mathbf{B}) = \begin{pmatrix} \mathbf{I}_{n-4} & * & * & * & * \\ & * & * & * & * \\ & & * & * & * \\ & & & * & * \\ & & & & * \end{pmatrix}.$$

After running Algorithm 1 once, the HNF of the resulting \mathbf{B} has the following form:

$$\text{HNF}(\mathbf{B}) = \begin{pmatrix} \mathbf{I}_{n-4} & * & * & * & \mathbf{0} \\ & * & * & * & 0 \\ & & * & * & 0 \\ & & & * & 0 \\ & & & & 1 \end{pmatrix}, \quad \text{HNF}(\mathbf{B}\mathbf{P}) = \begin{pmatrix} \mathbf{I}_{n-3} & * & * & * \\ & * & * & * \\ & & * & * \\ & & & * \end{pmatrix}.$$

Therefore, at the end of execution of Step 4, we have $\text{HNF}(\mathbf{B}) = \mathbf{I}_n$ for the resulting \mathbf{B} , which implies that \mathbf{B} is unimodular.

Thanks to Equation (3) with $s = 3$ and $\lambda \geq 3(n - 3)^{2/5}$, after each execution of Step 3, \mathbf{B}_{n-4}^T is primitive with probability at least 0.2. Therefore, after c (a constant) executions of Step 3, the probability that there is at least one \mathbf{B}_{n-4}^T is primitive is at least $1 - 0.8^c$. So after a constant number of executions of the loop Steps 1–6, it is expected that the output of

Algorithm 2 is a unimodular completion of \mathbf{A} . After each execution of Step 5, it follows from Proposition 3.2 that the magnitude of the new column is bounded by $O(n^2\|\mathbf{B}\|)$. Therefore, after four executions of Step 5, the magnitude of the new column is bounded by $O(n^8\|\mathbf{A}\|)$. So the correctness of Algorithm 2 is proved.

The main cost part of Algorithm 2 is a constant number of executions of the loop Steps 1–6. According to Proposition 3.2 and the algorithm for computing determinant in [4], each execution of the loop can be finished in an expected number of $O(n^{\omega+\varepsilon} \log^{1+\varepsilon} \|\mathbf{A}\|)$ bit operations. As a consequence, Theorem 1.2 is proved. ■

Remark 3.5 Without the help of Theorem 1.1, one may use the Iterated Determinant Reduction algorithm $n - k$ times for unimodular completion. This results in an algorithm that completes \mathbf{A} to an $n \times n$ unimodular matrix \mathbf{U} with $\|\mathbf{U}\| \leq n^{2(n-k)}\|\mathbf{A}\|$ in an expected number of $O((n - k)n^{\omega+\varepsilon} \log^{1+\varepsilon} \|\mathbf{A}\|)$ bit operations.

4 Experiments and Discussions

In this section, we present an experimental study on the empirical probability for the case that is included in Theorem 1.1 ($s \geq 3$), and also for the case that is not included in Theorem 1.1, e.g., the case of $s < 3$. In all the following experimental data, each empirical probability (labeled Exp.) are obtained by running 10,000 random tests on computer algebra system Maple and counting the success rate, i.e., the proportion of primitive matrices among all resulting matrices.

4.1 The Case of $s \geq 3$

In Tables 1 and 2 we study the for the case of $k = 0$, so that we can compare the empirical probability (column labeled Exp.), the lower bound given in Theorem 1.1 (column labeled Th. 1), and the limit probability given in Equation (2) (column labeled Limit probability). We use three different bounds for λ , namely, $\lambda = 10^5$, $\lambda = 10^{10}$ and $\lambda = 10^{20}$.

Table 1 Average empirical probability vs the probability in Theorem 1.1 ($k = 0$ and $s = 3$)

n	$\lambda = 10^5$		$\lambda = 10^{10}$		$\lambda = 10^{20}$		Limit probability
	Exp.	Th. 1.1	Exp.	Th. 1.1	Exp.	Th. 1.1	
5	0.9652	0.7366	0.9662	0.7366	0.9639	0.7366	0.9643
10	0.9335	0.2792	0.9291	0.2792	0.9306	0.2792	0.9334
15	0.9292	0.2190	0.9278	0.2190	0.9312	0.2190	0.9325
20	0.9338	0.2110	0.9349	0.2110	0.9345	0.2110	0.9325

Table 2 Average empirical probability vs the probability in Theorem 1.1 ($k = 0$ and $s = n - k - 2$)

n	$\lambda = 10^5$		$\lambda = 10^{10}$		$\lambda = 10^{20}$		Limit probability
	Exp.	Th. 1.1	Exp.	Th. 1.1	Exp.	Th. 1.1	
5	0.9652	0.7366	0.9662	0.7366	0.9639	0.7366	0.9643
10	0.9995	0.9653	0.9990	0.9653	0.9990	0.9653	0.9990
15	0.9999	0.9954	1.0000	0.9954	1.0000	0.9954	0.9999
20	1.0000	0.9993	1.0000	0.9993	1.0000	0.9993	0.9999

Both Tables 1 and 2 show that the empirical probability is relatively near to the limit probability, both of which are much better than our theoretical lower bound given in Theorem 1.1.

As indicated previously, $s = 3$ is the smallest s such that the lower bound given in Theorem 1.1 is between 0 and 1 for these experiments, while $s = n - k - 2$ is the largest possible value.

Comparing Table 1 with Table 2, it shows that larger s implies that both higher experimental success rate and better theoretical bound, however, smaller s implies the resulting matrix is closer to a unimodular matrix. In particular, for the case of $s = n - k - 2$, the probability bound given in Theorem 1.1 matches very well with both empirical probability and the limit probability.

In addition, Tables 1 and 2 also show that for different λ with same n , the theoretical bounds in column Th. 1 are almost the same and the data of empirical probability are almost the same as well. This is because that for a large enough λ , the bound given in Theorem 1.1 is almost independent of λ , as indicated by the oversimplified bound given in (4). For this reason, we fix $\lambda = 10^5$ for all other experiments (Tables 3–10).

Tables 3–6 are for the case of $k > 0$, for which there does not exist a known limit probability. We generate the initial primitive matrix as follows: We first generate a $k \times n$ matrix, whose entries are chosen randomly and independently from the uniform distribution over $[-\lambda, \lambda] \cap \mathbb{Z}$. If the matrix is not primitive, we regenerate a new matrix until it is eventually primitive. For each initial matrix, we complete it with uniformly random entries from A to an $(n - s - 1) \times n$ matrix 10,000 times and count the success rate. From these experiments, we can observe a similar phenomenon as the case of $k = 0$.

Table 3 Average empirical probability vs the probability in Theorem 1.1 ($k = 1, s = 3$)

n	10	15	20	25	30
Exp.	0.9321	0.9283	0.9291	0.9324	0.9310
Th. 1.1	0.3139	0.2235	0.2116	0.2101	0.2099

Table 4 Average empirical probability vs the probability in Theorem 1.1 ($k = 1, s = n - k - 2$)

n	5	10	15	20	25	30
Exp.	0.8919	0.9969	0.9999	1.0000	1.0000	1.0000
Th. 1.1	0.6049	0.9479	0.9931	0.9990	0.9998	0.9999

Table 5 Average empirical probability vs the probability in Theorem 1.1 ($k = n/2, s = 3$)

n	16	20	24	28	32	36
Exp.	0.9349	0.9338	0.9340	0.9312	0.9352	0.9333
Th. 1.1	0.3659	0.2792	0.2407	0.2235	0.2159	0.2125

Table 6 Average empirical probability vs the probability in Theorem 1.1 ($k = n/2, s = n - k - 2$)

n	16	20	24	28	32	36
Exp.	0.9919	0.9980	0.9996	0.9999	1.0000	1.0000
Th. 1.1	0.9219	0.9653	0.9845	0.9931	0.9969	0.9986

Totally, the probability bound presented in Theorem 1.1 is tight, especially for the case of large s . However, how to improve the theoretical bound for small s is an intriguing problem.

4.2 The Case of $s < 3$

For $s < 3$, the lower bound on that the resulting $(n - s - 1) \times n$ matrix is primitive given in Theorem 1.1 will be negative and hence useless. Therefore, it would be very interesting and useful to obtain a lower bound for the case of $s < 3$. Here we present some experimental results.

All test examples in Tables 7–9 are generated with the same method described in the last subsection, i.e., the input matrices are randomly chosen. In Table 10, all test examples are fixed to a $1 \times n$ primitive row $(1, 1, \dots, 1)$, but still with $\lambda = 10^5$. All of these tables show that for $0 \leq s \leq 2$, the empirical probability of that the resulting matrix is primitive is relatively high, similar with that of the case $s \geq 3$ shown in the last section. However, an effective theoretical lower bound on the probability for this case is left open.

Table 7 Average empirical probability for the case of $s = 2$ and $\lambda = 10^5$

n	16	20	24	28	32	36
$k = 0$	0.8599	0.8543	0.8575	0.8604	0.8628	0.8643
$k = 1$	0.8654	0.8618	0.8580	0.8671	0.8646	0.8620
$k = \frac{n}{2}$	0.8609	0.8611	0.8651	0.8621	0.8682	0.8662

Table 8 Average empirical probability for the case of $s = 1$ and $\lambda = 10^5$

n	16	20	24	28	32	36
$k = 0$	0.7201	0.7177	0.7124	0.7227	0.7125	0.7110
$k = 1$	0.7103	0.7141	0.7154	0.7129	0.7118	0.7192
$k = \frac{n}{2}$	0.7168	0.7212	0.7210	0.7226	0.7106	0.7154

Table 9 Average empirical probability for the case of $s = 0$ and $\lambda = 10^5$

n	16	20	24	28	32	36
$k = 0$	0.4365	0.4363	0.4353	0.4435	0.4434	0.4377
$k = 1$	0.4323	0.4337	0.4345	0.4451	0.4336	0.4440
$k = \frac{n}{2}$	0.4385	0.4371	0.4290	0.4330	0.4427	0.4330

Table 10 Average empirical probability for the case of $k = 1$ and $\lambda = 10^5$ with fixed $(1, 1, \dots, 1) \in \mathbb{Z}^n$ as the $k \times n$ matrix to be completed

n	16	20	24	28	32	36
$s = 0$	0.4330	0.4354	0.4342	0.4381	0.4284	0.4407
$s = 1$	0.7163	0.7120	0.7209	0.7198	0.7218	0.7159
$s = 2$	0.8629	0.8641	0.8649	0.8673	0.8616	0.8568

Conflict of Interest

WU Wenyuan is an editorial board member for Journal of Systems Science & Complexity and was not involved in the editorial review or the decision to publish this article. All authors declare that there are no competing interests.

References

- [1] Maze G, Rosenthal J, and Wagner U, Natural density of rectangular unimodular integer matrices, *Linear Algebra and Its Applications*, 2011, **434**(5): 1319–1324.

-
- [2] Eberly W, Giesbrecht M, and Villard G, On computing the determinant and Smith form of an integer matrix, *Proceedings of the 41st Annual Symposium on Foundations of Computer Science*, IEEE Computer Society, Los Alamitos, 2000, 675–685.
- [3] Storjohann A, High-order lifting and integrality certification, *Journal of Symbolic Computation*, 2003, **36**(3): 613–648.
- [4] Storjohann A, The shifted number system for fast linear algebra on integer matrices, *Journal of Complexity*, 2005, **21**(4): 609–650.
- [5] Phoong S M and Lin Y P, Application of unimodular matrices to signal compression, *Proceedings of 2002 IEEE International Symposium on Circuits and Systems*, IEEE, Piscataway, 2002, 837–840.
- [6] Lenstra A K, Lenstra H W, and Lovász L, Factoring polynomials with rational coefficients, *Mathematische Annalen*, 1982, **261**(4): 515–534.
- [7] Schnorr C P, A hierarchy of polynomial time lattice basis reduction algorithms, *Theoretical Computer Science*, 1987, **53**(2–3): 201–224.
- [8] Reiner I, Unimodular complements, *The American Mathematical Monthly*, 1956, **63**(4): 246–247.
- [9] Zhan X, Completion of a partial integral matrix to a unimodular matrix, *Linear Algebra and Its Applications*, 2006, **414**(1): 373–377.
- [10] Fang M, On the completion of a partial integral matrix to a unimodular matrix, *Linear Algebra and Its Applications*, 2007, **422**(1): 291–294.
- [11] Duffner M G and Silva F C, On the existence of unimodular matrices with a prescribed submatrix, *Linear Algebra and Its Applications*, 2017, **515**: 321–330.
- [12] Guo X, Hou F, and Liu X, Natural density of integral matrices that can be extended to invertible integral matrices, *Linear and Multilinear Algebra*, 2016, **64**(9): 1878–1886.
- [13] Fontein F and Wocjan P, On the probability of generating a lattice, *Journal of Symbolic Computation*, 2014, **64**: 3–15.
- [14] Aggarwal D and Regev O, A note on discrete gaussian combinations of lattice vectors, *Chicago Journal of Theoretical Computer Science*, 2013, **2016**: 1–11.
- [15] Kirshanova E, Nguyen H, Stehlé D, et al., On the smoothing parameter and last minimum of random orthogonal lattices, *Designs, Codes and Cryptography*, 2020, **88**(5): 931–950.
- [16] Randall D, *Efficient Generation of Random Nonsingular Matrices*, Technical Report UCB/CSD-91-658, EECS Department, University of California, Berkeley, 1991.
- [17] Kalaimani R K, Belur M N, and Sivasubramanian S, Generic pole assignability, structurally constrained controllers and unimodular completion, *Linear Algebra and Its Applications*, 2013, **439**(12): 4003–4022.
- [18] Zhou W and Labahn G, Unimodular completion of polynomial matrices, *Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation*, Ed. by Katsusuke Nabeshima, ACM, New York, 2014, 413–420.
- [19] Babai L, On Lovász’ lattice reduction and the nearest lattice point problem, *Proceedings STACS’85*, *Lecture Notes in Computer Science*, Ed. by Mehlhorn K, Springer, Heidelberg, 1985, **182**: 13–20.
- [20] Mulders T and Storjohann A, Certified dense linear system solving, *Journal of Symbolic Computation*, 2004, **37**(4): 485–510.