

On the probability of generating a primitive matrix

Jingwei Chen^{a,b,c}, Yong Feng^{a,b}, Yang Liu^{d,*}, Wenyuan Wu^{a,b}

^aChongqing Key Laboratory of Automated Reasoning and Cognition, Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing 400714, China

^bChongqing College, University of Chinese Academy of Sciences, Chongqing 400714, China

^cKey Laboratory of Advanced Manufacturing Technology of Ministry of Education, Guizhou University, Guiyang 550025, China

^dInformation Science and Engineering, Chongqing Jiaotong University, Chongqing 400074, China

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 size of entries bounded by λ , we study the probability that the $m \times n$ matrix extended from \mathbf{A} by choosing other $m - k$ vectors uniformly at random from $\{0, 1, \dots, \lambda - 1\}$ is still primitive. We present a complete and rigorous proof that the probability is at least a constant for the case of $m \leq n - 4$. Previously, only the limit case for $\lambda \rightarrow \infty$ with $k = 0$ was analysed in Maze *et al.* (2011), known as the natural density. As an application, we 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 and $\|\mathbf{A}\|$ is the maximal absolute value of entries of \mathbf{A} .

Keywords: integer matrix, unimodular matrix, matrix completion

2010 MSC: 15B36, 15A83

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 $\mathbf{y} \in \mathbb{Z}^k$; 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 proved that a $k \times n$ primitive matrix can always be extended to an $n \times n$ unimodular matrix over \mathbb{Z} .

*This work was partially supported by NSFC (11671377, 61903053), Youth Innovation Promotion Association of CAS, Guizhou Science and Technology Program [2020]4Y056 and Chongqing Science and Technology Program (cstc2019yszx-jcyjX0003).

*Corresponding author

Email address: liuyang13@cqjtu.edu.cn (Yang Liu)

Given a primitive matrix $\mathbf{A} \in \mathbb{Z}^{k \times n}$ with $\|\mathbf{A}\| := \max_{i,j} |a_{i,j}| \leq \lambda$, our focus in this paper will be on the probability that the $m \times n$ matrix extended from \mathbf{A} by choosing other $m - k$ vectors with entries uniformly at random from $\Lambda := \mathbb{Z} \cap [0, \lambda)$ is still primitive. In particular, we prove the following

Theorem 1. *Given a primitive matrix $\mathbf{A} \in \mathbb{Z}^{k \times n}$ with $\|\mathbf{A}\| \leq \lambda$ 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 the other rows chosen uniformly at random from Λ^n . Then the probability of that \mathbf{B} is primitive is at least*

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

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

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

when $\lambda \rightarrow \infty$, where $\zeta(\cdot)$ is the Riemann's zeta function. Theorem 1 gives an effective lower bound on the probability for finite λ , and hence will be useful in practice, especially in computer science.

From Eq. (1), the parameter k plays a very limited role for the result. Therefore, one may use a simpler but a little worse bound than that given in Eq. (1):

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

Note that this bound is independent of the parameter k . Moreover, if λ is large enough, this bound can be further simplified as

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

for some $0 < \delta < 1$. Surprisingly, this oversimplified bound only depends on s . For example, if $s = 3$ is fixed, then the resulting $(n - 4) \times n$ matrix will be primitive with a constant probability at least 0.2.

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 Eq. (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 matches the empirical probability well according to our experiments. In addition, one can 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

limit probability (natural density) of random $n \times n$ unimodular matrices is 0; see [1, Lemma 5]. These are limitations of Theorem 1. A similar situation occurs in a somewhat ‘dual’ case [2], where the probability of that m integer vectors with bounded entries generate a same lattice of rank n was studied. In [2], the ideal choice is $m = n + 1$ but a constant probability only for the case of $m \geq 2n + 1$ was rigorously proven. However, based on an extensive experimental study, we conjecture that a constant probability exists as well for $0 \leq s \leq 2$; see Conjecture 1 for detail.

As an application of Theorem 1, we present a fast Las Vegas algorithm (Algorithm 2) that completes a primitive matrix $\mathbf{A} \in \mathbb{Z}^{k \times n}$ to an $n \times n$ unimodular matrix \mathbf{U} such that $\|\mathbf{U}\| \leq n^{O(1)} \|\mathbf{A}\|$ in an expected number of $O(n^{\omega+\varepsilon} \log^{1+\varepsilon} \|\mathbf{A}\|)$ bit operations.

Techniques. The essential ingredient of our proof for Theorem 1 is adapted from [3, Section 6], which is 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, which holds 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 [4, 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 [5, Section 13.2]. We give full description of the algorithm for the integer matrix case and present a detailed analysis in Section 4.

Related work. Primitive and unimodular matrices have many applications in different areas. For example, unimodular matrices can be applied to signal compression [6]; the lattice reduction algorithms [7, 8] essentially produce a series of unimodular matrices (linear transformation) to improve the quality of the input lattice basis.

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 integer matrix can be completed to a unimodular matrix. In 1956 Reiner [9] proved a row vector can be completed to a unimodular matrix if and only if it is primitive. Zhan [10] 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 [11] improved Zhan’s result by proving that if a partial integral matrix has a free diagonal then this matrix can be completed to a unimodular matrix. Duffner and Silva [12] 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] analyze the natural density of $k \times n$ primitive matrices. Guo *et al.* [13] extends 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 are considered. Note that the natural density is a limit probability when the magnitude of matrix entry tends to infinity. However, a finite version is usually needed in practice, e.g., algorithm analysis. Therefore, the result in Theorem 1 will be useful.

For algorithms, Randall [14] presented an algorithm for generating random matrix with given determinant over finite field, which naturally can be used to generate unimodular matrices over finite fields. Kalaimani *et al.* [15] and Zhou and Labahn [16] 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 Remark 1 and 2).

Roadmap. We prove Theorem 1 in Section 2. In Section 3, we present an extensive experimental study on the probability that the resulting $(n-s-1) \times n$ matrix is primitive and conjecture that a constant probability exists for the case of $0 \leq s \leq 2$. We give a full description of the determinant reduction technique for integer matrices in Section 4 and apply it to the problem of unimodular matrix completion.

2. Proof of Theorem 1

Given a primitive matrix $\mathbf{A} \in \mathbb{Z}^{k \times n}$ with $k < n$ and $\|\mathbf{A}\| \leq \lambda$, we now consider extended \mathbf{A} to an $(n-s-1) \times n$ matrix by choosing other $n-k-s-1$ vectors with entries uniformly at random from $\Lambda = \mathbb{Z} \cap [0, \lambda)$, 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 $\Lambda = \{0\}$. 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 uniformly at random from the set Λ . To prove Theorem 1, we firstly need to bound from above the probability P of the event that $\mathbf{a}_1, \dots, \mathbf{a}_{n-s-1}$ are not primitive in \mathbb{Z}^n under the assumption that \mathbf{A} is primitive.

Lemma 1. *Let all notations be as above. Then*

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

We prove Lemma 1 following the approach of Eberly *et al.* [3, 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 $\text{rank}(\mathbf{a}_1, \dots, \mathbf{a}_i) \leq i-1$ over \mathbb{Z}_p .¹

¹The definition of **MDep _{i}** here is different from that in [3, Section 6], where **MDep _{i}** denotes

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

$$\begin{aligned}
P &= \Pr[\text{MDep}_{n-s-1}] \\
&\leq \Pr[\text{MDep}_{k+1} \vee \text{MDep}_{k+2} \vee \cdots \vee \text{MDep}_{n-s-1}] \\
&= \Pr[\text{MDep}_{k+1} \vee (\text{MDep}_{k+2} \wedge \neg \text{MDep}_{k+1}) \vee \cdots (\text{MDep}_{n-s-1} \wedge \neg \text{MDep}_{n-s-2})] \\
&= \Pr[(\text{MDep}_{k+1} \wedge \neg \text{MDep}_k) \vee (\text{MDep}_{k+2} \wedge \neg \text{MDep}_{k+1}) \vee \cdots (\text{MDep}_{n-s-1} \wedge \neg \text{MDep}_{n-s-2})] \\
&\leq \sum_{i=k+1}^{n-s-1} \Pr[\text{MDep}_i \wedge \neg \text{MDep}_{i-1}] \\
&\leq \sum_{i=k+1}^{n-s-1} \Pr[\text{MDep}_i | \neg \text{MDep}_{i-1}].
\end{aligned} \tag{3}$$

In order to bound $\Pr[\text{MDep}_i | \neg \text{MDep}_{i-1}]$ for $k+1 \leq i \leq n-s-1$, let's consider primes $p < \lambda$ and primes $p \geq \lambda$, respectively. For convenience, denote by

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

2.1. The case of $p < \lambda$

If $\lambda = 2$, then no such prime p exists. If $\lambda = 3$, then $p = 2$. Furthermore, the event $(\text{MDep}_i | \neg \text{MDep}_{i-1})$ is that $\text{rank}(\mathbf{A}_{i-1}) = i-1$ for any prime p but $\text{rank}(\mathbf{A}_i) = i-1$ for $p = 2$. From $\text{rank}(\mathbf{A}_{i-1}) = i-1$ for any prime p it follows that there exist $i-1$ columns C_{i-1} of \mathbf{A}_i such that the submatrix consisting of the first $i-1$ rows of these $i-1$ columns has rank $i-1$ for any prime p . Now, $\text{rank}(\mathbf{A}_i) = i-1$ for $p = 2$ means that for $j \notin C_{i-1}$, the entry $a_{i,j}$ must be a linear combination of $a_{i,j'}$ over \mathbb{Z}_2 with $j' \in C_{i-1}$. However, when $\lambda = 3$ and $p = 2$

$$\Pr_{x \leftarrow \Lambda}[x \equiv 0 \pmod{p}] = \Pr_{x \leftarrow \Lambda}[x \equiv 1 \pmod{p}] = \frac{1}{2},$$

so for $\lambda = 3$ it follows that

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

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

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

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 .

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

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

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

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

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

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

If $\lambda \geq 8$, then $p = 2$ or $p = 3$ or $p = 5$ or $p = 7$, etc. Then

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

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 1. *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*

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

2.2. The case of $p \geq \lambda$

If $p \geq \lambda$ is a fixed prime, then the probability that $\text{rank}(\mathbf{A}_i) \leq i-1$ over \mathbb{Z}_p is at most $\left(\frac{1}{\lambda}\right)^{n-i+1}$, since the probability that a value chosen uniformly at random from Λ equals a given value over \mathbb{Z}_p is at most $\frac{1}{\lambda}$. So

$$\Pr[\text{MDep}_i | \neg \text{MDep}_{i-1}] \leq \sum_{p \geq \lambda} \left(\frac{1}{\lambda}\right)^{n-i+1}.$$

However, in this case, p must divide the determinant of all $i \times i$ submatrices of \mathbf{A}_i . Each of these determinants is an integer with absolute value at most $i! \cdot \lambda^i \leq (i \cdot \lambda)^i$. So the number of possible primes is at most

$$\log_\lambda (i \cdot \lambda)^i \leq i(1 + \log_\lambda i).$$

It implies that in this case

$$\Pr[\text{MDep}_i | \neg \text{MDep}_{i-1}] \leq (i(1 + \log_\lambda i)) \cdot \left(\frac{1}{\lambda}\right)^{n-i+1}. \quad (4)$$

Proof of Lemma 1. It follows from Proposition 1 and Eq. (4) that for integer $\lambda \geq 2$,

$$\Pr[\text{MDep}_i | \neg \text{MDep}_{i-1}] \leq \left(\frac{2}{3}\right)^{n-i+1} + \frac{3}{4} \left(\frac{1}{3}\right)^{n-i+1} + (i(1 + \log_\lambda i)) \cdot \left(\frac{1}{\lambda}\right)^{n-i+1}.$$

Therefore, under the assumption that $\neg \text{MDep}_k$ happens, it follows from Eq. (3) that

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

□

Now, Theorem 1 is a direct consequence of Lemma 1.

3. Experiments

In this section, we present an experimental study on the result given in Theorem 1 and on the empirical probability for the case that is not included in Theorem 1, e.g., the case of $s = 0$, which implies a fast algorithm for unimodular matrix completion.

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.

3.1. The case of $s \geq 3$

In Table 1 and 2 we study the for the case of $k = 0$, so that we can compare the empirical probability (Exp.), the lower bound given in Theorem 1 (Th. 1), and the limit probability given in Eq. (2) (Limit probability). We use three different bounds for λ , namely, $\lambda = 10^5$, $\lambda = 10^{10}$ and $\lambda = 10^{20}$. Both Table 1 and 2 shows 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. As indicated previously, $s = 3$ is the smallest s such that the lower bound given in Theorem 1 is between 0 and 1 for these experiments, while $s = n - k - 2$ is the largest possible value. Comparing Table 1 and 2 shows that larger s implies both higher success rate and better bound, which consists with the limit probability, 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 matches very well with both empirical probability and the limit probability.

Table 1: Average empirical probability versus the probability in Theorem 1 for the case of $k = 0$ and $s = 3$.

n	$\lambda = 10^5$		$\lambda = 10^{10}$		$\lambda = 10^{20}$		Limit probability
	Exp.	Th. 1	Exp.	Th. 1	Exp.	Th. 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 versus the probability in Theorem 1 for the case of $k = 0$ and $s = n - k - 2$.

n	$\lambda = 10^5$		$\lambda = 10^{10}$		$\lambda = 10^{20}$		Limit probability
	Exp.	Th. 1	Exp.	Th. 1	Exp.	Th. 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

In addition, Table 1 and 2 also shows that for different λ with same n , the theoretical bounds in column Cor. 1 are exactly the same and the data of empirical probability are almost the same. This is because that for a large enough λ , the bound given in Theorem 1 is almost independent of λ . For this reason, we fix $\lambda = 10^5$ for all other experiments.

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 uniformly at random from $[-\lambda, \lambda] \cap \mathbb{Z}$. 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 Λ 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 versus the probability in Theorem 1 for the case of $k = 1$, $s = 3$ and $\lambda = 10^5$.

n	5	10	15	20	25	30
Exp.	1.0000	0.9321	0.9283	0.9291	0.9324	0.9310
Th. 1	1.0000	0.3139	0.2235	0.2116	0.2101	0.2099

Table 4: Average empirical probability versus the probability in Theorem 1 for the case of $k = 1$, $s = n - k - 2$ and $\lambda = 10^5$.

n	5	10	15	20	25	30
Exp.	0.8919	0.9969	0.9999	1.0000	1.0000	1.0000
Th. 1	0.6049	0.9479	0.9931	0.9990	0.9998	0.9999

Table 5: Average empirical probability versus the probability in Theorem 1 for the case of $k = n/2$, $s = 3$ and $\lambda = 10^5$.

n	16	20	24	28	32	36
Exp.	0.9349	0.9338	0.9340	0.9312	0.9352	0.9333
Th. 1	0.3659	0.2792	0.2407	0.2235	0.2159	0.2125

Table 6: Average empirical probability versus the probability in Theorem 1 for the case of $k = n/2$, $s = n - k - 2$ and $\lambda = 10^5$.

n	16	20	24	28	32	36
Exp.	0.9919	0.9980	0.9996	0.9999	1.0000	1.0000
Th. 1	0.9219	0.9653	0.9845	0.9931	0.9969	0.9986

Totally, the probability bound presented in Theorem 1 is tight, especially for the case of large s . However, how to improve the theoretical bound for small s is an intriguing problem.

3.2. 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 will be negative and hence useless. Therefore, it would be very interesting and useful to obtain a lower bound for the case of $s < 3$. Unfortunately, it does not seem easy. (A somewhat ‘dual’ case in the sense of m integer vectors generating a same lattice of rank n appeared in [2], where the optimal choice is $m = n + 1$ but a constant probability only for the case of $m = 2n + 1$ was rigorously proven.) Here we present some experimental results.

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

All test examples in Table 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. Based on these numerical evidence, we formulate the following conjecture.

Conjecture 1. *Given a primitive matrix $\mathbf{A} \in \mathbb{Z}^{k \times n}$ with $\|\mathbf{A}\| \leq \lambda$ and an integer $0 \leq s \leq n - k - 2$, there exists a constant $0 < c_0 < 1$ such that the following holds: Let $\mathbf{B} \in \mathbb{Z}^{(n-s-1) \times n}$ be a matrix with first k rows copied from \mathbf{A} and the other rows with entries chosen uniformly at random from Λ . Then the probability of that \mathbf{B} is primitive is at least c_0 .*

4. Unimodular matrix completion

When completing a given $k \times n$ primitive matrix to an $n \times n$ unimodular matrix, one can first complete the input matrix to an $(n - 1) \times n$ primitive

matrix with $n - k - 1$ uniformly random vectors, which corresponds to the case of $s = 0$, and then compute a vector as the last row by, e.g., the determinant reduction technique [4, Section 15].

Given $\mathbf{A} \in \mathbb{Z}^{n \times n}$, the determinant reduction technique introduced in [4, 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. An $m \times n$ matrix \mathbf{H} with integer entries is in *Hermite normal form* (HNF) if \mathbf{H} is upper triangular with the following properties:

1. the first r rows of \mathbf{H} are the non-zero rows of \mathbf{H} ,
2. for each row i , if h_{i,j_i} is its first nonzero entry, then $h_{i,j_i} > 0$ and $j_1 < j_2 < \dots < j_r$,
3. for each $1 \leq k < i \leq r$, the entries h_{k,j_i} of the j_i -th column of \mathbf{H} satisfy $0 \leq h_{k,j_i} < h_{i,j_i}$.

Given a matrix $\mathbf{A} \in \mathbb{Z}^{k \times n}$, one can always compute a matrix \mathbf{H} in HNF from \mathbf{A} via certain unimodular transformations. The matrix \mathbf{H} is unique, denoted by $\text{HNF}(\mathbf{A})$, although the unimodular transformations are usually not unique.

The determinant reduction technique was originally presented for integral polynomial matrix in [4, Section 15], with a worked example for integer matrix. Here, we give full details and analyses for the integer matrix case.

Algorithm 1 (Determinant reduction)

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

Output: A matrix $\mathbf{B} \in \mathbb{Z}^{n \times n}$, with \mathbf{B} equal 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 $\mathbf{C}_{n-1}^T \mathbf{u} = \mathbf{0}$.
 - 2: Call an extended gcd algorithm to compute $\mathbf{b} \in \mathbb{Z}^n$ such that $\mathbf{u}^T \mathbf{b} = 1$.
 - 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 $\overline{\mathbf{q}} := \lceil \overline{\mathbf{A}}^{-1} \overline{\mathbf{b}} \rceil$. //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 $\mathbf{b} - \mathbf{A}\mathbf{q}$.
 - 6: **return** \mathbf{B} .
-

Proposition 2. *Given an $n \times n$ integer matrix \mathbf{A} , Algorithm 1 correctly computes an $n \times n$ integer matrix \mathbf{B} satisfying \mathbf{B} equals \mathbf{A} except for possibly the last column, $\|\mathbf{B}\| = O(n^2 \|\mathbf{A}\|)$, and the last diagonal entry of $\text{HNF}(\mathbf{B})$ equals one, and requires at most $O(n^{\omega+\varepsilon} \log^{1+\varepsilon} \|\mathbf{A}\|)$ bit operations.*

Proof. Without loss of generality, we can assume that the first $n - 1$ columns of \mathbf{A} are linearly independent. This assumption implies that the vector \mathbf{u} produced in Step 1 is the unique primitive vector in $\ker \mathbf{C}_{n-1}^T \cap \mathbb{Z}^n$. Let \mathbf{U} be an arbitrary unimodular matrix such that $\text{HNF}(\mathbf{A}) = \mathbf{U}\mathbf{A}$. Then the uniqueness of \mathbf{u} implies that \mathbf{u}^T 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} will have HNF with the last diagonal entry one. The whole algorithm can be expressed as the following equation

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

Step 3-5 are to reduce the size of $\bar{\mathbf{b}}$ by the columns of $\bar{\mathbf{A}}$. It follows from $\mathbf{C}_{n-1}\mathbf{u} = \mathbf{0}$ that all entries of $\mathbf{u}^T\mathbf{A}$ are zero except for possibly the last one. However, the last entry of \mathbf{q} equals zero, so we have $\mathbf{u}^T\mathbf{A}\mathbf{q} = 0$, and hence

$$\mathbf{u}^T(\mathbf{b} - \mathbf{A}\mathbf{q}) = \mathbf{u}^T\mathbf{b} = 1, \quad (5)$$

which implies the last diagonal of $\text{HNF}(\mathbf{B})$ must be one.

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

$$\|\bar{\mathbf{b}} - \bar{\mathbf{A}}\bar{\mathbf{q}}\|_\infty = \left\| \bar{\mathbf{b}} - \bar{\mathbf{A}} \lceil \bar{\mathbf{A}}^{-1}\bar{\mathbf{b}} \rceil \right\|_\infty = \left\| \bar{\mathbf{b}} - \bar{\mathbf{A}} \left(\bar{\mathbf{A}}^{-1}\bar{\mathbf{b}} + \boldsymbol{\varepsilon} \right) \right\|_\infty \leq \|\bar{\mathbf{A}}\boldsymbol{\varepsilon}\|_\infty \leq \frac{n-1}{2} \|\bar{\mathbf{A}}\|,$$

where $\|\cdot\|_\infty$ is the ℓ_∞ -norm of a vector and $\|\boldsymbol{\varepsilon}\|_\infty \leq \frac{1}{2}$ is used. Denote $\mathbf{u} = \begin{pmatrix} \bar{\mathbf{u}} \\ u_n \end{pmatrix}$. Without loss of generality, we can assume that

$$\|\mathbf{u}\|_\infty = u_n. \quad (6)$$

Otherwise there must exist a permutation matrix \mathbf{P} such that $\mathbf{u}^T\mathbf{P}\mathbf{P}^{-1}\mathbf{C}_{n-1}^T = \mathbf{0}$, $\mathbf{u}^T\mathbf{P}$ satisfies Eq. (6), and $\mathbf{P}^{-1}\mathbf{C}_{n-1}^T$ still corresponds to the first $n-1$ columns of \mathbf{A} but with a certain column permutation. From Eq. (5), we have

$$u_n(b_n - \bar{\mathbf{a}}^T\bar{\mathbf{q}}) = 1 - \bar{\mathbf{u}}^T(\bar{\mathbf{b}} - \bar{\mathbf{A}}\bar{\mathbf{q}}),$$

combining Eq. (6), which gives

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

where $\|\cdot\|_1$ is the ℓ_1 -norm of a vector. 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 (Step 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 [17] and an extended gcd computation (Step 2) that can be accomplished within $O(n^{2+\varepsilon} \log^{1+\varepsilon} \|\mathbf{A}\|)$ bit operations [5, Section 13.2]. Totally, Algorithm 1 costs at most $O(n^{\omega+\varepsilon} \log^{1+\varepsilon} \|\mathbf{A}\|)$ bit operations. \square

Corollary 1. *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. One can use the matrix $\begin{pmatrix} \mathbf{A} \\ \mathbf{0} \end{pmatrix}^T$ as the input for Algorithm 1. Then the transpose of the output matrix will be a unimodular completion of \mathbf{A} . \square

Remark 1. 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 for determinant in [5].

An algorithm named Iterated Determinant Reduction was presented in [5, Section 13.2]. Given a 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

$$\mathbf{P} = \begin{pmatrix} \mathbf{0} & \mathbf{I}_{n-1} \\ 1 & 0 \end{pmatrix}$$

from right.

Now we give our algorithm for unimodular matrix completion as follows.

Algorithm 2 (Unimodular matrix completion)

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

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

- 1: Set $\lambda := \|\mathbf{A}\|$ and $\mathbf{B} := \mathbf{A}^T$.
 - 2: Repeat
 - 3: Complete \mathbf{B} as an $n \times n$ matrix with entries chosen uniformly at random from $\{0, 1, \dots, \lambda-1\}$.
 - 4: Repeat the following 4 times:
 - 5: Set \mathbf{B} as the output of Algorithm 1 with input \mathbf{B} .
 - 6: Set $\mathbf{B} := \mathbf{B}\mathbf{P}$.
 - 7: Until $\det(\mathbf{B}) = \pm 1$
 - 8: **return** $(\mathbf{B}\mathbf{P}^4)^{-T}$.
-

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 4 diagonal entries of the Hermite normal form of \mathbf{B} be one. If \mathbf{B}_{n-4}^T is primitive at the end of Step 3, then the output of the algorithm is a unimodular matrix completed from the input matrix \mathbf{A} . Thanks to Theorem 1, after Step 3, \mathbf{B}_{n-4}^T is primitive with probability at least 0.2. The main cost part of Algorithm 2 is calling Algorithm 1 a constant times. As a consequence, we have the following

Theorem 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.*

Remark 2. Without the help of Theorem 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.

References

- [1] G. Maze, J. Rosenthal, U. Wagner, Natural density of rectangular unimodular integer matrices, *Linear Algebra and its Applications* 434 (5) (2011) 1319–1324, doi: 10.1016/j.laa.2010.11.015.
- [2] F. Fontein, P. Wocjan, On the probability of generating a lattice, *Journal of Symbolic Computation* 64 (2014) 3–15, doi: 10.1016/j.jsc.2013.12.002.
- [3] W. Eberly, M. Giesbrecht, G. Villard, Computing the determinant and Smith form of an integer matrix, in: *Proceedings of the 41st Annual Symposium on Foundations of Computer Science* (12–14 November 2000, Redondo Beach, USA), IEEE Computer Society, Los Alamitos, 2000, pp. 675–685, doi: 10.1109/SFCS.2000.892335.
- [4] A. Storjohann, High-order lifting and integrality certification, *Journal of Symbolic Computation* 36 (3) (2003) 613–648, doi: 10.1016/S0747-7171(03)00097-X.
- [5] A. Storjohann, The shifted number system for fast linear algebra on integer matrices, *Journal of Complexity* 21 (4) (2005) 609–650, doi: 10.1016/j.jco.2005.04.002.
- [6] S.-M. Phoong, Y.-P. Lin, Application of unimodular matrices to signal compression, in: *Proceedings of 2002 IEEE International Symposium on Circuits and Systems* (May 26–29, 2002, Phoenix-Scottsdale, USA), IEEE, Piscataway, 2002, pp. 837–840, doi: 10.1109/ISCAS.2002.1009971.
- [7] A. K. Lenstra, H. W. Lenstra, L. Lovász, Factoring polynomials with rational coefficients, *Mathematische Annalen* 261 (4) (1982) 515–534, doi: 10.1007/BF01457454.
- [8] C.-P. Schnorr, A hierarchy of polynomial time lattice basis reduction algorithms, *Theoretical Computer Science* 53 (2–3) (1987) 201–224, doi: 10.1016/0304-3975(87)90064-8.
- [9] I. Reiner, Unimodular complements, *The American Mathematical Monthly* 63 (4) (1956) 246–247, doi: 10.2307/2310351.
- [10] X. Zhan, Completion of a partial integral matrix to a unimodular matrix, *Linear Algebra and its Applications* 414 (1) (2006) 373–377, doi: 10.1016/j.laa.2005.10.013.
- [11] M. Fang, On the completion of a partial integral matrix to a unimodular matrix, *Linear Algebra and its Applications* 422 (1) (2007) 291–294.

- [12] M. G. Duffner, F. C. Silva, On the existence of unimodular matrices with a prescribed submatrix, *Linear Algebra and its Applications* 515 (2017) 321–330, doi: 10.1016/j.laa.2016.11.015.
- [13] X. Guo, F. Hou, X. Liu, Natural density of integral matrices that can be extended to invertible integral matrices, *Linear and Multilinear Algebra* 64 (9) (2016) 1878–1886, doi: 10.1080/03081087.2015.1127316.
- [14] D. Randall, Efficient generation of random nonsingular matrices, Tech. Rep. UCB/CSD-91-658, EECS Department, University of California, Berkeley, available at <http://www2.eecs.berkeley.edu/Pubs/TechRpts/1991/6147.html> (Nov 1991).
- [15] R. K. Kalaimani, M. N. Belur, S. Sivasubramanian, Generic pole assignability, structurally constrained controllers and unimodular completion, *Linear Algebra and its Applications* 439 (12) (2013) 4003–4022, doi: 10.1016/j.laa.2013.10.004.
- [16] W. Zhou, G. Labahn, Unimodular completion of polynomial matrices, in: K. Nabeshima (Ed.), *Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation* (July 23–25, 2014, Kobe, Japan), ACM, New York, 2014, pp. 413–420, doi: 10.1145/2608628.2608640.
- [17] T. Mulders, A. Storjohann, Certified dense linear system solving, *Journal of Symbolic Computation* 37 (4) (2004) 485–510, doi: 10.1016/j.jsc.2003.07.004.