Lattice compression of integer matrices

Zhuliang Chen , Arne Storjohann *,1

David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1

Abstract

The set of column vectors of an $m \times n$ integer matrix with $m < n$ generate an integer lattice. A generating set for the lattice consisting of only $m + k$ vectors, k a small constant, can be produced with high probability by multiplying by a random matrix with relatively small entries.

Key words: Integer lattice; Randomization

1 Introduction

Let A be an $m \times n$ integer matrix with rank m. The columns of A generate the m-dimensional lattice $\mathcal{L}(A)$ comprised of the set of all $\mathbb{Z}$-linear combinations of the columns. In this paper we propose to construct a new generating set for $\mathcal{L}(A)$ by computing AB for an integer matrix $B \in \mathbb{Z}^{n \times (m+k)}$, k a small constant, even $k = 1$. We call this *lattice compression*. The compression is successful if $\mathcal{L}(AB) = \mathcal{L}(A)$.

In case of success we can use AB in place of A in algorithms that compute invariants of $\mathcal{L}(A)$. The advantage is that AB is nearly square, with only k more columns than rows. For example, the system solving algorithm described in (Chen and Storjohann, 2005) is designed especially for the case of an input matrix with at most a few more columns than rows (e.g., $k = 30$). Although

* Corresponding author.
Email addresses: `z4chen@uwaterloo.ca` (Zhuliang Chen),
`astorjoh@uwaterloo.ca` (Arne Storjohann).
[1] Partially supported by the National Sciences and Engineering Research Council of Canada under Grant No. RGPIN 26082-03: Algorithms for the exact solution to problems in linear algebra.

applicable for an arbitrary input matrix, the cost of the algorithm grows prohibitively as the difference between the column and row dimension increases, even $n = 2m$ can be too large for example. Lattice compression extends the usability of the algorithm to arbitrary rectangular matrices.

This paper gives a thorough study of the effectiveness of lattice compression when $B \in \mathbb{Z}^{n\times(m+k)}$ has entries chosen uniformly and randomly from $\{0, 1, \ldots, \lambda - 1\}$, $\lambda \geq 2$. Our main contribution is to derive explicit bounds on the probability of success in terms of k. Let $\|A\| = \max_{ij} |A_{ij}|$ and log denote $\log_2$. We show that if λ is a multiple of thirty and at least as large as $1200m \log 2nm\|A\|$, then $\mathcal{L}(AB) = \mathcal{L}(A)$ with probability greater than 0.4, even for $k = 1$. Moreover, we show that the probability of failure decreases exponentially with respect to k: $\mathcal{L}(AB) \neq \mathcal{L}(A)$ with probability less than $0.9\left(\frac{1}{2}\right)^{k-1}$.

We also study the effect of λ on the probability of failure. On the one hand, increasing λ does not allow to drive down the probability of failure for a given choice of k. On the other hand, as k increases the upper bound $0.9\left(\frac{1}{2}\right)^{k-1}$ on the probability of failure still holds for smaller choices of λ: it suffices that $\lambda \geq 8 \times (25m \log 2nm\|A\|)^{1/\lceil k/3\rceil}$. The exponential decay in the probability of error coupled with the decreasing bound for λ means that lattice compression is a highly effective tool in practise as well as for the design of Las Vegas and Monte Carlo randomized algorithms. In the conclusions section of the paper we give some specific examples of applications.

We mention some related work. An analysis of the special case of lattice compression when $m = 1$ and $k = 2$ is given by Cooperman et al. (1999) and von zur Gathen and Shparlinski (2004). Note that $\mathcal{L}(AB) = \mathcal{L}(A)$ in this special case precisely when the gcd of the n integers in $A \in \mathbb{Z}^{1\times n}$ is equal to that of the two integers in $AB \in \mathbb{Z}^{1\times 2}$. The polynomial analogue of the case $m = 1$ and $k = 2$ is analyzed by Conflitti (2003); the case with m arbitrary and $k = 1$ is considered by Storjohann and Labahn (1995).

We now outline the paper. Section 2 characterizes a good lattice conditioner in terms of some rank properties. Section 3 recalls some bounds on the probability that a partially random matrix will have full row rank. Section 4 uses conditional probability to bound the probability that $\mathcal{L}(AB) \neq \mathcal{L}(A)$ in terms of four other probabilities. An approximation scheme to facilitate computation of these four probabilities to sufficient accuracy is also developed. Finally, Section 5 gives the main results.

An earlier version of this paper appears in (Chen, 2005, Chapter 3).

2 Characterization of good lattice compressors

Recall the situation over a field K. Let $A \in \mathsf{K}^{m\times n}$ have full row rank m. Then any set of m linearly independent columns of A provide a basis for the vector space $\mathcal{L}(A)$ generated by the columns of A. It follows that for any matrix $B \in \mathsf{K}^{n\times(m+k)}$, we have $\mathcal{L}(AB) = \mathcal{L}(A)$ precisely when $AB \in \mathsf{K}^{m\times(m+k)}$ also has rank m. The next lemma gives an alternative characterization of a good compression over K that is more suitable for estimating the probability of success when B is chosen randomly. A matrix $N \in \mathsf{K}^{n\times(n-m)}$ is called a *right nullspace basis* of A if $\mathcal{L}(N) = \{x \in \mathsf{K}^n \mid Ax = 0\}$.

Lemma 1 *Let $A \in \mathsf{K}^{m\times n}$ have rank m and right nullspace $N \in \mathsf{K}^{n\times(n-m)}$. Let B be over K with row dimension n. Then* $\mathrm{rank}(AB) = \mathrm{rank}[N|B] - (n-m)$.

Lemma 1 states that $\mathcal{L}(AB) = \mathcal{L}(A)$ if and only if $[N|B]$ has full row rank. To extend the lemma to a principal ideal domain such as $\mathbb{Z}$ we need to introduce some notation and recall some concepts. Let R be a principal ideal domain, $\bar{\mathsf{R}}$ be the fraction field of R, and for $p \in \mathsf{R}$ a prime let $\mathsf{R}/(p)$ denote the residue class ring of elements modulo p. Another matrix Y over R with the same row dimension as X is a *right multiple* of X if $Y = X*$ for some matrix $*$ over R. Matrices X and Y are right multiples of each other if and only if they are *right equivalent*: in this case $\mathcal{L}(X) = \mathcal{L}(Y)$ and we denote this with $X \equiv_R Y$. *Left multiple* and *left equivalence* ($\equiv_L$) are defined similarly for the transpose situation.

Definition 2 *Let $A \in \mathsf{R}^{m\times n}$ have full row rank m. A matrix $N \in \mathsf{R}^{n\times(n-m)}$ such that $\mathcal{L}(N) = \{x \in \mathsf{R}^n \mid Ax = 0\}$ is called a* right kernel *of A.*

Any right kernel N of A has the property that $N \equiv_L I_{n-m}$.

Fact 3 *An integer matrix X satisfies $X \equiv_L I$ if and only if $X \bmod p$ has full column rank over $\mathsf{R}/(p)$ for all primes p in R. Similarly, $X \equiv_R I$ if and only if $X \bmod p$ has full row rank over $\mathsf{R}/(p)$ for all primes p.*

Any matrix $A \in \mathsf{R}^{m\times n}$ of rank m is right equivalent to a nonsingular lower triangular matrix $H \in \mathsf{R}^{m\times m}$. An example of H is the Hermite column basis. Any such H has the property that the matrix $\bar{A} := H^{-1}A$ has entries from R (as opposed to $\bar{\mathsf{R}}$) and $\bar{A} \equiv_R I_m$. For example, if $A \in \mathbb{Z}^{1\times n}$ we can choose H to be the 1×1 matrix with single entry equal to a gcd of all entries in A. The next theorem is the analogue of Lemma 1 but for lattices over a principal ideal domain instead of vector spaces over a field.

Theorem 4 *Let $A \in \mathsf{R}^{m\times n}$ with rank m and right kernel $N \in \mathsf{R}^{n\times(n-m)}$. Let $B \in \mathsf{R}^{n\times k}$. Then $AB \equiv_R A$ if and only if $[N|B] \equiv_R I_n$.*

PROOF. Let $H \in \mathsf{R}^{m \times m}$ be the Hermite column basis of A and define $\bar{A} := H^{-1}A$. Then $\bar{A}$ is over R with $\bar{A} \equiv_R I_m$. It follows that $AB \equiv_R A$ if and only if $\bar{A}B \equiv_R I_m$. For any prime p, $N \bmod p$ is a right nullspace for $\bar{A} \bmod p$ over $\mathsf{R}/(p)$, so we can apply Lemma 1 to conclude that $\bar{A}B \bmod p$ has full row rank m over $\mathsf{R}/(p)$ if and only if $[N|B] \bmod p$ has full row rank n: $\text{rank}(\bar{A}B \bmod p) = \text{rank}([N|B] \bmod p) - (n-m)$. The result now follows from Fact 3. $\square$

Corollary 5 *There exist a nonzero minor M of AB with the property that for all primes p not dividing M, the rank of $[N|B] \bmod p$ over $\mathsf{R}/(p)$ does not decrease compared to the rank of $[N|B]$ over R.*

PROOF. Let $r = \text{rank}(AB)$. Up to permuting the rows of A and columns of B we may assume without loss of generality that the leading $r \times r$ submatrix of AB is nonsingular; let M be the determinant of this submatrix. As in the proof of Theorem 4, let H be the Hermite column basis of A and define $\bar{A} := H^{-1}A$. Then $\bar{A}$ is over R with $\bar{A} \equiv_R I_m$. Because H is lower triangular, the determinant of the principal $r \times r$ submatrix of $\bar{A}B$ is a divisor of M. For any prime p not dividing M the rank of $\bar{A}B \bmod p$ over $\mathsf{R}/(p)$ will remain r. The result now follows from Lemma 1. $\square$

3 Rank properties of partly random matrices

Let $C \in \mathbb{Z}^{n \times m_1}$ be a given matrix of rank r. Let P be the probability that the augmented matrix $[C|D]$ will have full row rank when $D \in \mathbb{Z}^{n \times m_2}$ has entries chosen uniformly and randomly from $\{0, 1, \ldots, \lambda - 1\}$, $\lambda \geq 2$. An application of (Mulders and Storjohann, 2004, Lemma 13) gives

$$P \geq \prod_{i=s}^{m_2} \left(1 - \left(\frac{1}{\lambda}\right)^i\right) \text{ where } s = m_2 + r - n + 1.$$

We remark that s is exactly one more than the difference between the column dimension of D and the minimum possible number of columns that D must have if $[C|D]$ is to have any chance of having full row rank. We will estimate P in terms of s only by replacing the upper index of the product with ∞. Induction on u with base case $u = 0$ can be used to show the following technical result, the proof of which we omit.

Lemma 6 *If $s \geq 1$, $u \geq 0$ and $0 \leq x \leq 2/3$ then $\prod_{i=s}^{s+u}(1 - x^i) \geq 1 - 3x^s + 2x^{s+u}$.*

This gives the estimate $\prod_{i=s}^{s+u}(1 - x^i) \geq \prod_{i=s}^{\infty}(1 - x^i) \geq 1 - 3x^s$, independent of the upper bound $s + u$. For any $t \geq 0$, it follows that $\prod_{i=s}^{\infty}(1 - x^i) =$

$(\prod_{i=s}^{s+t-1}(1-x^i))(\prod_{i=s+t}^{\infty}(1-x^i)) \geq (\prod_{i=s}^{s+t-1}(1-x^i))(1-3x^{s+t})$. Define a function $f : \mathbb{R}_{[0,2/3]} \times \mathbb{Z}_{\geq 1} \times \mathbb{Z}_{\geq 0} \rightarrow \mathbb{R}$ by

$$f(x,s,t) = 1 - (1-3x^{s+t}) \prod_{i=s}^{s+t-1} \left(1 - x^i\right). \tag{1}$$

The sequence $f(x,s,0), f(x,s,1), f(x,s,2), \ldots$ approaches $1 - \prod_{i=s}^{\infty}(1-x^i)$ from above.

Theorem 7 *Let $C \in \mathbb{Z}^{n\times m_1}$ have rank r over $\mathbb{Q}$. Let $D \in \mathbb{Z}^{n\times m_2}$ with $m_2+r \geq n$ have entries chosen uniformly and randomly from $\{0,1,\ldots,\lambda-1\}$, $\lambda \geq 2$. The probability that $[C|D]$ does not have full row rank n over $\mathbb{Q}$ is strictly less than $f(1/\lambda, m_2+r-n+1, t)$ for any $t \geq 0$, where f is defined in (1).*

Corollary 8 *Let p be a prime and let $\bar{r}$ be the rank of C $\operatorname{mod} p$ over $\mathbb{Z}/(p)$. If $m_1+\bar{r} \geq n$ then the probability that $[C|D]$ $\operatorname{mod} p$ does not have full row rank n over $\mathbb{Z}/(p)$ is strictly less than $f(\lceil\lambda/p\rceil/\lambda, m_2+\bar{r}-n+1, t)$ for any $t \geq 0$.*

We remark that $\lceil\lambda/p\rceil$ is the maximum number of elements of $\{0,1,\ldots,\lambda-1\}$ that are in the preimage of any element of $\mathbb{Z}/(p)$ under reduction modulo p. For any prime p and $\lambda \geq 2$ we have $\lceil\lambda/p\rceil/\lambda \leq 2/3$.

4 Bounding the probability that $AB \not\equiv_R A$

Throughout this section let $A \in \mathbb{Z}^{m\times n}$ be a given matrix with full row rank m, $N \in \mathbb{Z}^{n\times(n-m)}$ be a right kernel for A, and for a given $k \geq 1$ and $\lambda \geq 2$, let $B \in \mathbb{Z}^{n\times(m+k)}$ have entries chosen uniformly from $\{0,1,\ldots,\lambda-1\}$. We begin by defining two events as follows.

- Event L: There exists at least one prime $p \geq \lambda$ such that $[N|B]$ $\operatorname{mod} p$ has rank strictly less than n over $\mathbb{Z}/(p)$.
- Event S: There exists at least one prime $p < \lambda$ such that $[N|B]$ $\operatorname{mod} p$ has rank strictly less than n over $\mathbb{Z}/(p)$.

Then $P[AB \not\equiv_R A] = P[[N|B] \not\equiv_R I_n] = P[\mathsf{L} \vee \mathsf{S}] \leq P[\mathsf{L}] + P[\mathsf{S}]$. The first equality is Theorem 4. The second is Fact 3.

We now outline our method for bounding $P[\mathsf{L}]$. Partition B into three submatrices $B = [B_1|B_2|B_3]$ according to two nonnegative integer parameters t_1 and t_2, $t_1+t_2 \leq k$, so that the column dimensions of B_1, B_2 and B_3 are $m-t_1$, t_1+t_2 and $k-t_2$ respectively. Then $[N|B_1]$ has t_1 fewer columns than rows, $[N|B_1|B_2]$ has t_2 more columns than rows, and $[N|B_1|B_3]$ has $k-t_1-t_2$ more columns than rows. Define two events as follows.

- Event L_1: There exists at least one prime $p \geq \lambda$ such that $[N|B_1] \bmod p \in \mathbb{Z}/(p)^{n\times(n-t_1)}$ has rank at most $n - t_1 - 2$ over $\mathbb{Z}/(p)$.
- Event L_2: The matrix $[N|B_1|B_2] \in \mathbb{Z}^{n\times(n+t_2)}$ does not have full row rank n.

Using standard rules for probability:

$$\begin{aligned} P[\mathsf{L}] &= P[\mathsf{L} \wedge (\mathsf{L}_1 \vee \mathsf{L}_2)] + P[\mathsf{L} \wedge \neg(\mathsf{L}_1 \vee \mathsf{L}_2)] \\ &\leq P[\mathsf{L}_1 \vee \mathsf{L}_2] + P[\mathsf{L} \wedge \neg(\mathsf{L}_1 \vee \mathsf{L}_2)] \\ &\leq P[\mathsf{L}_1] + P[\mathsf{L}_2] + P[\mathsf{L} \mid \neg(\mathsf{L}_1 \vee \mathsf{L}_2)] \end{aligned}$$

To proceed we need to derive upper bounds on $P[\mathsf{L}_1]$, $P[\mathsf{L}_2]$, $P[\mathsf{L} \mid \neg(\mathsf{L}_1 \vee \mathsf{L}_2)]$ and $P[\mathsf{S}]$. The remainder of this section establishes these bounds.

An upper bound for $P[\mathsf{L}_2]$

Theorem 7 gives

$$P[\mathsf{L}_2] < f(1/\lambda, t_2 + 1, 1) < \left(1 + \frac{3}{\lambda}\right)\left(\frac{1}{\lambda}\right)^{t_2+1}. \tag{2}$$

4.1 An upper bound for $P[\mathsf{L} \mid \neg(\mathsf{L}_1 \vee \mathsf{L}_2)]$

If event L_1 is not satisfied then $[N|B_1]$ has rank at least $n-t_1-1$ over $\mathbb{Z}/(p)$. For a prime $p \geq \lambda$, let P_p be the probability that $[N|B_1|B_3] \bmod p \in \mathbb{Z}^{n\times(n+k-t_1-t_2)}$ does not have full row rank n over $\mathbb{Z}/(p)$, conditional on the assumption that L_1 is not satisfied. Then Corollary 8 gives

$$P_p < f(1/\lambda, k - t_1 - t_2, 1) < \left(1 + \frac{3}{\lambda}\right)\left(\frac{1}{\lambda}\right)^{k-t_1-t_2}. \tag{3}$$

Now assume that event L_2 is not satisfied. Then Corollary 5 gives the existence of a nonzero minor M of $A[B_1|B_2]$ such that, for all primes p not dividing M, $[N|B_1|B_2] \bmod p$ has full row rank n. Thus, under the assumption event L_1 and L_2 are false, event L can be satisfied only if there exists a prime $p \geq \lambda$ dividing M such that $[N|B_1|B_3] \bmod p$ does not have full row rank n over $\mathbb{Z}/(p)$. Let $\#P_{\geq\lambda}$ be an upper bound on the number of distinct prime divisors $p \geq \lambda$ of any nonzero minor of AB. Then

$$P[\mathsf{L} \mid \neg(\mathsf{L}_1 \vee \mathsf{L}_2)] \leq \sum_{p\geq\lambda,\, p|M} P_p < \#P_{\geq\lambda}\left(1 + \frac{3}{\lambda}\right)\left(\frac{1}{\lambda}\right)^{k-t_1-t_2}. \tag{4}$$

4.2 An upper bound for $P[\mathsf{L}_1]$

We follow closely Eberly et al. (2000, Section 6). Similar to Eberly et al. (2000) we define two types of events for $n-m \leq i \leq n$. Let Dep_i denote the event that the first i columns of $[N|B_1]$ are linearly dependent over the rationals. Let MDep_i denote the event that there exists at least one prime $p \geq \lambda$ such that the submatrix of $[N|B_1] \bmod p$ comprised of the first i columns has rank at most $i-2$ over $\mathbb{Z}/(p)$. Note that event L_1 coincides with event MDep_{n-t_1}. Since $N \equiv_L I_{n-m}$, neither event Dep_{n-m} nor event MDep_{n-m} holds. For $i > n-m$ assume that neither the event Dep_{i-1} nor the event MDep_{i-1} is satisfied.

It was shown by Eberly et al. (2000) that

$$P[\mathsf{Dep}_i|\neg(\mathsf{Dep}_{i-1} \vee \mathsf{MDep}_{i-1})] \leq P[\mathsf{Dep}_i|\neg\mathsf{Dep}_{i-1}] \leq \left(\frac{1}{\lambda}\right)^{n-i+1}. \tag{5}$$

For a prime $p \geq \lambda$, let P_p be the probability that the first i columns of $[N|B_1]$, taken modulo p, have rank at most $i-2$ over $\mathbb{Z}/(p)$, conditional on the assumption that event MDep_{i-1} is not satisfied. It was shown by Eberly et al. (2000) that $P_p \leq (1/\lambda)^{n-i+2}$. If Dep_{i-1} is not satisfied, then there exists a nonzero minor of dimension $i-1$ of the first $i-1$ columns of AB_1, and analogous to the derivation of (4) we have

$$P[\mathsf{MDep}_i|\neg(\mathsf{Dep}_{i-1} \vee \mathsf{MDep}_{i-1})] \leq \#P_{\geq\lambda} \left(\frac{1}{\lambda}\right)^{n-i+2}. \tag{6}$$

Finally, the definition of conditional probability gives

$$\begin{aligned}
P[\mathsf{L}_1] &= P[\mathsf{MDep}_{n-t_1}] \\
&\leq P\left[\vee_{i=n-m}^{n-t_1}(\mathsf{Dep}_i \vee \mathsf{MDep}_i)\right] \\
&= P\Big[\mathsf{Dep}_{n-m} \vee \mathsf{MDep}_{n-m} \\
&\qquad\qquad \vee_{i=n-m+1}^{n-t_1}(\mathsf{Dep}_i \vee \mathsf{MDep}_i \wedge \neg(\mathsf{Dep}_{i-1} \vee \mathsf{MDep}_{i-1}))\Big] \\
&\leq \sum_{i=n-m+1}^{n-t_1} P[\mathsf{Dep}_i \vee \mathsf{MDep}_i|\neg(\mathsf{Dep}_{i-1} \vee \mathsf{MDep}_{i-1})] \\
&\leq \sum_{i=n-m+1}^{n-t_1} P[\mathsf{Dep}_i|\neg(\mathsf{Dep}_{i-1} \vee \mathsf{MDep}_{i-1})] \\
&\qquad + \sum_{i=n-m+1}^{n-t_1} P[\mathsf{MDep}_i|\neg(\mathsf{Dep}_{i-1} \vee \mathsf{MDep}_{i-1})] \\
&\leq \left(1+\frac{\#P_{\geq\lambda}}{\lambda}\right) \sum_{i=n-m+1}^{n-t_1} \left(\frac{1}{\lambda}\right)^{n-i+1} < 2\left(1+\frac{\#P_{\geq\lambda}}{\lambda}\right)\left(\frac{1}{\lambda}\right)^{t_1+1}
\end{aligned} \tag{7}$$

4.3 Some upper bounds for $P[\mathsf{S}]$

Note that $P[\mathsf{S}]$ is a function of λ and k. We will derive approximation schemes F, G and H such that $P[\mathsf{S}] < F(\lambda, k, t)$ and $P[\mathsf{S} \mid \lambda \geq \lambda_{\min}] < G(\lambda_{\min}, k, t) \leq H(\lambda_{\min}, k, t, b)$. The arguments t and b are nonnegative integers. Larger values of t and b may give a better approximation. The computations in this paper were performed[2] by using Maple 10^{TM}.

For a prime $p < \lambda$, let P_p be the probability that $[N|B] \bmod p \in \mathbb{Z}^{n\times(n+k)}$ does not have full row rank n over $\mathbb{Z}/(p)$. Then $P[\mathsf{S}] \leq \sum_{p<\lambda} P_p$. Corollary 8 gives $P_p < f(\lceil\lambda/p\rceil/\lambda, k+1, t)$. We can estimate

$$\frac{\lceil\lambda/p\rceil}{\lambda} \leq \frac{1}{\lambda} + \frac{1}{p} - \frac{1}{p\lambda},$$

which is $\leq 2/3$ for all choices of $1 < p < \lambda$ and is non-increasing in λ. If we assume that λ is bounded from below by $\lambda_{\min}$ then

$$P_p \leq f\left(\frac{\lceil\lambda/p\rceil}{\lambda}, k+1, t\right) \leq f\left(\frac{1}{\lambda_{\min}} + \frac{1}{p} - \frac{1}{p\lambda_{\min}}, k+1, t\right). \tag{8}$$

The last inequality uses the observation that $f(x, s, t) \leq f(y, s, t)$ for $x \leq y$. An upper bound for given values of k and λ is given by

$$P[\mathsf{S}] < F(\lambda, k, t) := \sum_{p<\lambda} f\left(\frac{\lceil\lambda/p\rceil}{\lambda}, k+1, t\right). \tag{9}$$

To get an estimate valid for all $\lambda \geq \lambda_{\min}$, note that

$$P[\mathsf{S} \mid \lambda \geq \lambda_{\min}] \leq \sum_{p<\lambda_{\min}} P_p + \sum_{\lambda_{\min}\leq p<\lambda} P_p \tag{10}$$

and estimate

$$\begin{aligned}\sum_{\lambda_{\min}\leq p<\lambda} P_p &\leq \sum_{\lambda_{\min}\leq p<\lambda} f\left(\frac{1}{\lambda} + \frac{1}{p} - \frac{1}{p\lambda}, k+1, 0\right)\\ &\leq \sum_{\lambda_{\min}\leq p<\lambda} f(2/p, k+1, 0)\end{aligned}$$

[2] A text file containing definitions of F, G and H as well as the computations used in the three lemmas of this section can be found on the second authors home page.

$$
\begin{aligned}
&\leq \sum_{\lambda_{\min} \leq p < \lambda} 3\left(\frac{2}{p}\right)^{k+1} \\
&< \sum_{i=\lambda_{\min}/2}^{\infty} 3\left(\frac{1}{i}\right)^{k+1} < \left(\frac{3}{k}\right)\left(\frac{1}{\lambda_{\min}/2-1}\right)^{k} \qquad (11)
\end{aligned}
$$

Substituting (8) and (11) into (10) gives

$$
\begin{aligned}
P[\mathsf{S} \mid \lambda \geq \lambda_{\min}] &< G(\lambda_{\min}, k, t) \\
&:= \sum_{p<\lambda_{\min}} f\left(\frac{1}{\lambda_{\min}} + \frac{1}{p} - \frac{1}{p\lambda_{\min}}, k+1, t\right) \\
&\qquad + \left(\frac{3}{k}\right)\left(\frac{1}{\lambda_{\min}/2-1}\right)^{k} \qquad (12)
\end{aligned}
$$

We can improve the approximation by using the tighter estimate (9) for values of λ in the range $\lambda_{\min} \leq \lambda < \lambda_{\min} + b$, some $b \geq 0$.

$$
\begin{aligned}
P[\mathsf{S} \mid \lambda \geq \lambda_{\min}] &\leq \max\{P[\mathsf{S} \mid \lambda_{\min} \leq \lambda < \lambda_{\min}+b], P[\mathsf{S} \mid \lambda \geq \lambda_{\min}+b]\} \\
&< H(\lambda_{\min}, k, t, b) \\
&:= \max\left\{\max_{\lambda_{\min} \leq \lambda < \lambda_{\min}+b}\{F(\lambda, k, t)\}, G(\lambda_{\min}+b, k, t)\right\}. \qquad (13)
\end{aligned}
$$

The estimates F, G and H derived above give the following result.

Lemma 9 $P[\mathsf{S}] < 2\left(\frac{2}{3}\right)^k$.

PROOF. $F(3,k,0) = 2\left(\frac{2}{3}\right)^k$ and $F(4,k,1) < \left(\frac{5}{2}\right)\left(\frac{1}{2}\right)^k < 2\left(\frac{2}{3}\right)^k$ using $k \geq 1$, proving the result for all $\lambda \leq 4 \wedge k \geq 1$. Direct evaluation verifies that $H(5,k,10,20) < 2\left(\frac{2}{3}\right)^k$ for $1 \leq k \leq 5$, proving the result for $\lambda \geq 5 \wedge 1 \leq k \leq 5$. Finally, consider that $G(5,k,0) = \left(\frac{9}{5}\right)\left(\frac{3}{5}\right)^k + \left(\frac{7}{5}\right)\left(\frac{7}{15}\right)^k + \left(\frac{3}{k}\right)\left(\frac{2}{3}\right)^k$, which for $k \geq 5$ is bounded by $G(5,5,0)\left(\frac{2}{3}\right)^{k-5} < 2\left(\frac{2}{3}\right)^k$, proving the result for $\lambda \geq 5 \wedge k \geq 5$. $\square$

Under the assumption that λ is a multiple of six we can estimate $P_2 < f(1/2, k+1, t)$ and $P_3 < f(1/3, k+1, t)$. This leads to the following result.

Lemma 10 $P[\mathsf{S} \mid \lambda \in 6\mathbb{Z}] < \left(\frac{4}{5}\right)\left(\frac{1}{2}\right)^{k-1}$.

PROOF. Direct evaluation verifies that $F(6,k,3) < \left(\frac{4}{5}\right)\left(\frac{1}{2}\right)^{k-1}$ for $1 \leq k \leq$

31, proving the result for $\lambda = 6 \wedge 1 \leq k \leq 30$. We can estimate $P[\mathsf{S} \mid \lambda \geq 6\,] < G(6,k,0)$ where G is as in (12), but if $\lambda \in 6\mathbb{Z}$ we can replace the sum with the tighter estimate $f(1/2, k+1, 0) + f(1/3, k+1, 0) + f(1/6 + 1/5 - 1/30)$. This gives $P[\mathsf{S} \mid \lambda \in 6\mathbb{Z}] < \left(\frac{3}{2}\right)\left(\frac{1}{2}\right)^k + 2\left(\frac{1}{3}\right)^k + \left(\frac{3}{k}\right)\left(\frac{1}{2}\right)^k$ which for $k \geq 31$ is bounded by

$$\left(\left(\frac{3}{2}+\frac{3}{k}\right)\left(\frac{1}{2}\right)^k + 2\left(\frac{1}{3}\right)^k\right)\Bigg|_{k=31}\left(\frac{1}{2}\right)^{k-31} < \left(\frac{4}{5}\right)\left(\frac{1}{2}\right)^{k-1},$$

proving the result for $\lambda = 6\mathbb{Z} \wedge k \geq 31$. $\square$

For $k = 1$ the last lemma gives the bound $P[\mathsf{S} \mid \lambda \in 6\mathbb{Z}] < \frac{4}{5}$. We can slightly improve the bound by assuming that λ is a large enough multiple of thirty. Let P_{30} be the probability that $[N|B] \bmod p$ does not have full row rank over $\mathbb{Z}/(p)$ for at least one prime $p \in \{2,3,5\}$. Then conditional on λ being a multiple of thirty, we have $P_{30} \leq 1 - (1-P_2)(1-P_3)(1-P_5)$. For $\lambda \in 30\mathbb{Z}$ define F_{30} similar to F in (9) but with the first three terms of the sum replaced with the previous bound for P_{30}. Similarly, define G_{30} similar to G in (12) but with first three terms of the sum replaced by the tighter estimate $1 - (1 - f(1/2, k+1, t))(1 - f(1/3, k+1, t))(1 - f(1/5, k+1, t))$. Finally, for both λ and b multiples of thirty, define H_{30} similar to H in (13) but with F and G replaced by F_{30} and G_{30} respectively, and the max ranging only over $\lambda \in [\lambda_{\min}, \lambda_{\min} + b)$ which are multiples of 30. Direct evaluation of $H_{30}(1200, 1, 30, 6000)$ gives the following result.

Lemma 11 $P[\mathsf{S} \mid \lambda \in 30\mathbb{Z} \wedge \lambda \geq 1200] < \frac{3}{5} - 0.00356$.

The upper bound $P[\mathsf{S} \mid \lambda \in 30\mathbb{Z} \wedge \lambda \geq 1200] < 0.6$ of the last lemma cannot be substantially improved with further analysis. When $k = 1$, it follows from Theorem 4 and (Mulders and Storjohann, 2004, Lemma 13) that the probability of failure with respect to at least one of the first three primes is equal to $1 - \prod_{i=2}^{m+1}\left(1 - \frac{1}{2^i}\right)\left(1 - \frac{1}{3^i}\right)\left(1 - \frac{1}{5^i}\right)$, which for any $m \geq 3$ is greater than 0.5.

5 Some upper bounds on $P[AB \not\equiv_R A]$

Recall that $\log \equiv \log_2$ and that $\|A\| = \max_{ij}|A_{ij}|$. In the following theorem we use a function $\omega : \mathbb{Z}_{\geq 2} \to \mathbb{Z}$ such that the product of the first $\lfloor \omega(N) \rfloor$ primes is at least as large as N. An obvious choice is $\omega(N) = \log N$. Robin (1983) shows that we take $\omega(N) = \frac{\ln N}{\ln \ln N} + 1.45743\frac{\ln N}{(\ln \ln N)^2}$.

Theorem 12 *Let $A \in \mathbb{Z}^{m\times n}$ have rank m. For $k \geq 1$ and $\lambda \geq 2$ let $B \in \mathbb{Z}^{n\times(m+k)}$ have entries chosen uniformly and randomly from $\{0, 1, \ldots, \lambda - 1\}$.*

Let P be the probability that $AB \not\equiv_R A$. Then

(1) $P < \frac{3}{5}$, *if* $\lambda \in 30\mathbb{Z}$ *and* $\lambda \geq 1200m \log 2nm\|A\|$.

(2) $P < \left(\frac{9}{10}\right)\left(\frac{1}{2}\right)^{k-1}$, *if* $\lambda \in 6\mathbb{Z}$ *and* $\lambda \geq \lambda_{\min}(k) := 8 \times (25m \log 2nm\|A\|)^{1/\lceil k/3 \rceil}$.

(3) $P < \left(\frac{1}{2}\right)^{e}$ *where* $e = (k - 8 - 2\lceil \log \omega(N) \rceil/3$, $N = (nm\|A\|)^m$, *if* $\lambda = 2$.

In the last section we showed that

$$P[AB \not\equiv_R A] \leq \overbrace{P[\mathsf{L}_1] + P[\mathsf{L}_2] + P[\mathsf{L} \mid \neg(\mathsf{L}_1 \vee \mathsf{L}_2)]}^{P[\mathsf{L}]} + P[\mathsf{S}]$$

and derived bounds for the probabilities on the right hand side in terms of the given parameters k and λ, but also in terms of choosable parameters t_1 and t_1 such that $t_1 + t_2 \leq k$. We begin by making some remarks about the asymptotic behaviour of these probabilities in terms of of k and λ. The previous section showed that $P[\mathsf{S}] = O(2^{-k})$ if $\lambda \in 6\mathbb{Z}$, while $P[\mathsf{L}_1]$, $P[\mathsf{L}_2]$ and $P[\mathsf{L}|\neg(\mathsf{L}_1 \vee \mathsf{L}_2)]$ are $O(\lambda^{-k})$. Below we will see that a careful choice of t_1 and t_2 gives $P[\mathsf{L}] = O(\lambda^{-k/3})$. Thus, for $\lambda \geq 8$ we have $P[\mathsf{L}] = O(2^{-k})$ also, giving $P[AB \not\equiv A] = O(2^{-k})$ if $\lambda \in 6\mathbb{Z}$ and $\lambda \geq 8$.

The bounds $P[\mathsf{L}_1]$ and $P[\mathsf{L} \mid \neg(\mathsf{L}_1 \vee \mathsf{L}_2)]$ also depend on $\#P_{\geq\lambda}$, a bound on the number of distinct prime divisors $p \geq \lambda$ that can divide any nonzero minor of AB.

Lemma 13 $\#P_{\geq\lambda} \leq m \log 2nm\|A\|$.

PROOF. Let M be any minor of AB. The number of divisors of M at least as large as λ is bounded by $\log_\lambda M$. Using $\|AB\| \leq n\|A\|(\lambda - 1)$ gives $M \leq m!(n(\lambda - 1)\|A\|)^m \leq (mn\lambda\|A\|)^m$. Now note that $\log_\lambda mn\lambda\|A\| = \log_\lambda 2mn\|A\|$. □

From (2), (4) and (7) we get

$$P[\mathsf{L}] < 2\left(\frac{1}{\lambda} + \frac{\#P_{\geq\lambda}}{\lambda^2}\right)\left(\frac{1}{\lambda}\right)^{t_1} + \left(\frac{1}{\lambda} + \frac{3}{\lambda^2}\right)\left(\frac{1}{\lambda}\right)^{t_2} + \#P_{\geq\lambda}\left(\frac{1}{\lambda} + \frac{3}{\lambda^2}\right)\left(\frac{1}{\lambda}\right)^{k-t_1-t_2-1}. \quad (14)$$

If we choose $t_1 = t_2 = \lceil k/3 \rceil - 1$ then $k - t_1 - t_2 - 1 \geq \lceil k/3 \rceil - 1$ and we can estimate

$$P[\mathsf{L}] < \left(\frac{3 + \#P_{\geq\lambda}}{\lambda} + \frac{3 + 5\#P_{\geq\lambda}}{\lambda^2}\right)\left(\frac{1}{\lambda}\right)^{\lceil k/3 \rceil - 1}. \quad (15)$$

Note that the upper bound (15) is nonincreasing in both λ and k. Substituting $k = 1$ and $\lambda = 1200\#P_{\geq\lambda}$ into the bound and simplifying gives the following result, valid for all $k \geq 1$ but mainly interesting for $k = 1$.

Lemma 14 *$P[\mathsf{L} \mid \lambda \geq 1200\#P_{\geq\lambda}] < 0.00334.$*

Using $P[AB \not\equiv A] \leq P[\mathsf{L}] + P[\mathsf{S}]$, the first part of the theorem now follows from the previous two lemmas together with Lemma 11. The second part of the theorem follows from the next lemma together with Lemma 10.

Lemma 15 *$P[\mathsf{L} \mid \lambda \geq \lambda_{\min}(k)] < \left(\frac{1}{10}\right)\left(\frac{1}{2}\right)^{k-1}.$*

PROOF. Let $x = \lceil k/3 \rceil$. Then $k \leq 3x$. If $\lambda \geq \lambda_{\min}(k)$ then the upper bound (15) simplifies to $(3+\#P_{\geq\lambda})/\lambda^x + (3+5\#P_{\geq\lambda})/\lambda^{x+1} < 4/(8^x \times 25) + 8/(8^{x+1} \times 25) = \left(\frac{1}{10}\right)\left(\frac{1}{2}\right)^{3x-1}$. □

Now consider the case $\lambda = 2$. There are no primes < 2 so $P[\mathsf{S} \mid \lambda = 2] = 0$. Define $\bar{t}_1 := t_1 - (3 + \lceil \log \#P_{\geq 2} \rceil)$, $\bar{t}_2 := t_2 - 2$ and $\bar{k} := k - t_1 - t_2 - 1 - (2 + \lceil \log \#P_{\geq 2} \rceil)$. Substituting $\lambda = 2$ into (14) gives

$$
\begin{aligned}
P &< \left(1 + \frac{\#P_{\geq 2}}{2}\right)\left(\frac{1}{2}\right)^{t_1} + \left(\frac{5}{4}\right)\left(\frac{1}{2}\right)^{t_2} + \left(\frac{5\#P_{\geq 2}}{4}\right)\left(\frac{1}{2}\right)^{k-t_1-t_2-1} \\
&\leq \left(\frac{3}{16}\right)\left(\frac{1}{2}\right)^{\bar{t}_1} + \left(\frac{5}{16}\right)\left(\frac{1}{2}\right)^{\bar{t}_2} + \left(\frac{5}{16}\right)\left(\frac{1}{2}\right)^{\bar{k}}
\end{aligned}
$$

Note that we can choose $\bar{t}_1 = \bar{t}_2 = \bar{k} = 0$ when $k = 8 + 2\lceil \log \#P_{\geq\lambda} \rceil$. Each increase in k beyond $8 + 2\lceil \log \#P_{\geq\lambda} \rceil$ allows to increase one of $\bar{t}_1$, $\bar{t}_2$ or $\bar{k}$. If all the exponents $\bar{t}_1$, $\bar{t}_2$, and $\bar{k}$ equal zero, the bound simplifies to $\frac{13}{16} < \left(\frac{1}{2}\right)^{0/3}$. If $\bar{t}_1 = 1$ and $\bar{t}_2 = \bar{k} = 0$ the bound simplifies to $\frac{23}{32} < \left(\frac{1}{2}\right)^{1/3}$. Finally, if $\bar{t}_1 = \bar{t}_2 = 1$ and $\bar{k} = 0$ the bound simplifies to $\frac{9}{16} < \left(\frac{1}{2}\right)^{2/3}$. We conclude that $P < \left(\frac{1}{2}\right)^{(k-8-2\lceil \log \#P_{\geq 2} \rceil)/3}$ for all k. Finally, note that minors of AB are bounded in magnitude by $(nm\|A\|)^m$ so we can take $\#P_{\geq 2} = \omega((nm\|A\|)^m)$. This proves the last part of theorem.

6 Conclusions

We have analyzed the probability that $AB \equiv_R A$ when $B \in \mathbb{Z}^{n \times (m+k)}$ has entries chosen uniformly and randomly from $\{0, 1, \ldots, \lambda - 1\}$. For convenience, we have assumed throughout the paper that the given matrix $A \in \mathbb{Z}^{m\times n}$ has full row rank m but this is not required. If A has rank $r < m$ then the bounds derived in Theorem 12 still hold, even with each occurrence of m replaced by r.

Compression to a matrix with k extra columns will fail with probability $O(2^{-k})$ provided λ is a multiple of six and $\lambda = \Omega((m \log nm\|A\|)^{3/k})$. Most of our effort was devoted to deriving explicit bounds, for example when $k = 1$. On the one hand, considering only the first three primes shows that the probability that a randomized compression to an $m \times (m + k)$ matrix with $k = 1$ will be successful is strictly less than 0.5 whenever $r = m \geq 3$. On the other hand, part one of Theorem 12 states that compression with $k = 1$ will be successful with probability greater than 0.4 provided λ is a multiple of thirty and $\lambda \geq 1200m \log 2nm\|A\|$.

We use lattice compression in the implementation of the linear solving algorithm described in Chen and Storjohann (2005). A typical choice for k in this algorithm is $k = 20$. For any input matrix A with $m \leq 10000$, $n \leq 20000$ and $\|A\| \leq 2^{100}$, part two of Theorem 12 states that choosing $\lambda = 96$ and $k = 20$ will give a probability of success greater than 0.999998. For $\lambda = 2$, in which case B is a random $\{0, 1\}$-matrix, part three of Theorem 12 states that the probability of success will be greater than $1 - \left(\frac{1}{2}\right)^{(k-42)/3}$.

References

Chen, Z., 2005. A BLAS based C library for exact linear algebra on integer matrices. Master's thesis, School of Computer Science, University of Waterloo.

Chen, Z., Storjohann, A., 2005. A BLAS based C library for exact linear algebra on integer matrices. In: Kauers, M. (Ed.), Proc. Int'l. Symp. on Symbolic and Algebraic Computation: ISSAC '05. ACM Press, New York, pp. 92–99.

Conflitti, A., 2003. On computation of the greatest common divisor of several polynomials over a finite field. Finite Field Appl. 9, 423–431.

Cooperman, G., Feisel, S., von zur Gathen, J., Havas, G., 1999. GCD of many integers (Extended Abstract). In: Proceedings of the Fifth International Computing and Combinatorics Conference, Tokyo, 1999, Lecture Notes in Computer Science. Vol. 1627. Springer-Verlag, Berlin, pp. 310–317.

Eberly, W., Giesbrecht, M., Villard, G., 2000. Computing the determinant and Smith form of an integer matrix. In: Proc. 31st Ann. IEEE Symp. Foundations of Computer Science. pp. 675–685.

von zur Gathen, J., Shparlinski, I. E., 2004. GCD of random linear forms. In:

Algorithms and Computation: 15th International Symposium, ISAAC 2004. LNCS 3341. Springer Verlag, pp. 464–469.

Mulders, T., Storjohann, A., 2004. Certified dense linear system solving. Journal of Symbolic Computation 37 (4), 485–510.

Robin, G., 1983. Estimation de la fonction de Tchebychef Θ sur le k-ième nombre premier et grandes valeurs de la fonction $\omega(n)$ nombre de diviseurs premiers de n. Acta Arith. 42, 367–389.

Storjohann, A., Labahn, G., 1995. Preconditioning of rectangular polynomial matrices for efficient Hermite normal form computation. In: Levelt, A. H. M. (Ed.), Proc. Int'l. Symp. on Symbolic and Algebraic Computation: ISSAC '95. ACM Press, New York, pp. 119–125.