2008

TR-2008007: Degeneration of Structured Integer Matrices Modulo an Integer

Victor Y. Pan

Xinmao Wang

Degeneration of Structured Integer Matrices Modulo an Integer

Victor Y. Pan
Department of Mathematics and Computer Science
Lehman College of CUNY, Bronx, NY 10468, USA
victor.pan@lehman.cuny.edu
http://comet.lehman.cuny.edu/vpan/
Supported by NSF Grant CCR 9732206 and PSC CUNY Awards
67297–0036, 68291–0037, 69330–0038, and 69350–0038
and
Xinmao Wang
xinmao@ustc.edu.cn
Department of Mathematics,
University of Science and Technology of China,
Hefei, Anhui 230026, China

May 5, 2008

Abstract

Hensel's lifting modulo a prime q is a customary means of the solution of an integer or rational linear system of equations. In combination with some effective numerical algorithms this technique enables solution in nearly optimal time in the case of most popular structured inputs. Practically one can further benefit from choosing $q = 2^v$ for a proper positive integer v and performing binary computations within the computer precision. If the input matrix becomes singular because of the reduction modulo q, then the approach fails. For larger integers q and random integer input matrices, however, such degeneration occurs rarely according to the analysis by Brent and McKay 1987. Based on distinct techniques we show that degeneration also occurs rarely for random integer matrices with all most popular structures such as the Toeplitz, Hankel, band and rank (quasiseparable) structures. Furthermore with random small-rank modifications of an input matrix we have good chances to overcome degeneration, safely solve the new linear system, and recover the solution of the original one. The results of our extensive tests support our formal analysis.

2000 Math. Subject Classification: 68W30, 68W20

Key Words: Structured matrices, Matrix inverse modulo an integer

1 Introduction

The algorithm in [P02], [PMRW04], [PMRa], and [PMRb] for solving linear systems of equations with integer coefficients employs Hensel's symbolic lifting modulo a prime and initializes it with some effective numerical techniques. Presently such symbolic–numerical combinations towards common goals are growing in popularity [SNC07], [SNC07a], [TCS08].

The algorithm solves Toeplitz, Hankel, banded, and other most popular structured nonsingular linear systems of n equations with integer coefficients in nearly optimal time under the Boolean (bit-operation) and word operation complexity models. More precisely, this is achieved wherever the coefficient matrix M as well as its precomputed inverse M^{-1} can be multiplied by vectors in early linear arithmetic time, and in this case the algorithm involves the order of $n^2 \log^2 n$ bit-operations provided all coefficients have absolute values bounded by n^c for a constant c. This is within logarithmic factor from the information lower bound $O(n^2 \log n)$.

The algorithm and nearly optimal cost bounds are extended to computing the rank, determinant, and a basis for the null space of structured matrices and to some fundamental computations with univariate polynomials. Further applications include theoretical and practical acceleration of Wiedemann and Coppersmith's block Wiedemann algorithms for the determinant and Smith's factors of a general or sparse integer matrix [P04], [PMRa].

Practically one can further benefit from initializing Hensel's lifting modulo $q = 2^v$ for a positive integer v chosen within the computer precision and performing binary computations. If the input matrix becomes singular because of the reduction modulo q, then the approach fails, but how frequently such a problem occurs? Not very frequently on the average integer input matrix for larger integers q, according to the estimates in [BMK87].

These estimates do not apply to sparse and/or structured matrices, however, and in this area the known results are sparse. Namely, it was proved in [D60] (and later also in [KL96]) that a random integer Toeplitz/Hankel matrix is singular modulo a prime q with the probability q. A respective estimate modulo any $q > 1$ appeared in [PMRW04]. In [BKY07] D. Bleichenbacher, A. Kiayias, and M. Yung 2007 deduced similar estimates modulo a prime q for (generalized) Vandermonde matrices. This was needed for decoding Reed–Solomon codes.

We extend the above short list of the known results to the large class of all most popular structured matrices and any integer $q > 1$. Like [BKY07] and [PMRW04] we rely on the celebrated lemma by Demillo–Lipton 1978 [DL78] (also by Zippel 1979 [Z79] and Schwartz 1980 [S80]). Our resulting estimates show that the degeneration modulo a fixed larger integer q is rare on the average integer structured matrix.

We also propose small-rank random modification of the input matrix if it degenerates and show that this is likely to resolve the problem. And with

a high probability we can avoid such degeneration for any fixed nonsingular integer matrix if we choose a reasonably large random prime or its power as the modulus q (see Appendix A).

The results of our extensive numerical tests are in reasonably good agreement with our theoretical study and show that the solution modulo a reasonably large power of two is quite safe for an average general or structured integer linear system of equations and that small-rank random modifications of an input matrix is an effective means of countering the rare cases of degeneracy.

We organize our presentation as follows. We devote the next section to some basic definitions and Section 3 to estimating how frequently general and structured integer matrices are nonsingular on the average. In Section 4 we transform an integer matrix to counter its singularity modulo a fixed integer. In Section 5 we present the results of our numerical tests. In Appendix A we estimate how frequently a fixed nonsingular integer matrix becomes singular modulo the power of a random prime. In Appendix B we briefly recall some effective algorithms for integer structured matrices that require nondegeneration modulo $q > 1$. Section 5 is due to the second author. Otherwise the paper is due to the first author. Part of our work was covered earlier in the proceedings paper [P02] and technical teports [P02a], [W02], and [PMRW04].

2 Basic definitions

Definition 2.1. *We write $\mathbb{Z}$ for the ring of integers and $\mathbb{Z}_q$ for the ring of integers modulo an integer q. We write $a = z \mod q$, for three integers $q > 1$, a, and z, to denote a unique integer a such that q divides $z - a$ and $0 \le a < q$. A prime p has an order h in an integer s if p^h divides s but p^{h+1} does not.*

Definition 2.2. *$M = (m_{i,j})_{i,j=1}^{k,l} \in \mathbb{R}^{k\times l}$ is a $k \times l$ matrix with entries $m_{i,j}$ in a ring $\mathbb{R}$. $\mathbf{v} = (v_i)_{i=1}^{k} \in \mathbb{R}^{k\times 1}$ is a column vector. I is the identity matrix of a proper size. I_l is the $l \times l$ identity matrix. M^T is the transpose of M. $|M| = \max_j \sum_i |m_{i,j}|$ is the column norm of a matrix $M = (m_{i,j})_{i,j}$. $\det M$ is the determinant of a matrix M. A matrix M is* nonsingular *if $\det M = 0$.*

Definition 2.3. *$d_k = d_k(M)$ is the kth determinantal divisor of a matrix $M \in \mathbb{Z}^{n\times n}$ for $k = 1, \dots, n$, that is, the greatest common divisor (hereafter gcd) of all its $k \times k$ minors (subdeterminants). $s_0 = d_0 = 1$, $s_k = s_k(M) = d_k/d_{k-1}$ are the kth* Smith invariant factors *of M for $k = 1, \dots, n$.*

We have $s_1, \dots, s_n \in \mathbb{Z}$ and

$$|\det M| = s_1 \cdots s_n \tag{2.1}$$

(see Newman 1972 [N72]).

Next we recall most popular classes of structured matrices with displacement structure [P01] and then, in Definition 2.8, the classes of banded and rank structured (quasi-separable) matrices.

Definition 2.4. *$T = (t_{i,j})_{i,j}$ is a Toeplitz matrix if $t_{i,j} = t_{i+1,j+1}$ for every pair of its entries $t_{i,j}$ and $t_{i+1,j+1}$. $H = (h_{i,j})_{i,j}$ is a Hankel matrix if $h_{i,j} = h_{i-1,j+1}$ for every pair of its entries $h_{i,j}$ and $h_{i-1,j+1}$. $J = (j_{g,h})_{g,h=0}^{n-1,n-1}$ is the unit Hankel (reflection) matrix where $j_{g,n-1-g} = 1$ for $g = 0, \dots, n-1$, $j_{g,h} = 0$ for $h + g \neq n - 1$. ($J\ (v_i)_{i=0}^{n-1} = (v_{n-i-1})_{i=0}^{n-1}$, $J^2 = I$.)*

Definition 2.5. *T is a Toeplitz-like (resp. Hankel-like) matrix with the displacement rank of at most r if it can be represented as $\sum_{i=1}^{r} Z(\mathbf{g}_i)Z(\mathbf{h}_i)^T$ (resp. $\sum_{i=1}^{r} Z(\mathbf{g}_i)Z(\mathbf{h}_i)^T J$) for some vectors $\mathbf{g}_i$ and $\mathbf{h}_i$, $i = 1, \dots, r$, where $Z(\mathbf{v})$ denotes the lower triangular Toeplitz matrix defined by its first column $\mathbf{v}$.*

Remark 2.1. *TJ and JT are Hankel matrices if T is a Toeplitz matrix and are Hankel-like matrices if T is a Toeplitz-like matrix. HJ and JH are Toeplitz matrices if H is a Hankel matrix and are Toeplitz-like matrices if H is a Hankel-like matrix. Due to these observations, our study of Toeplitz and Toeplitz-like matrices can be readily extended to Hankel and Hankel-like matrices.*

Definition 2.6. *$V(\mathbf{t}) = (t_i^{j-1})_{i,j}$ is a Vandermonde matrix. A matrix V has a Vandermonde-like structure and the displacement rank at most r if this matrix or its transpose can be represented as $\sum_{i=1}^{r} \operatorname{diag}(\frac{1}{1-t_i^n})_{i=1}^{r} \operatorname{diag}(\mathbf{g}_i)V(\mathbf{t})JZ(J\mathbf{h}_i)^T$ where $\operatorname{diag}(\mathbf{v}) = \operatorname{diag}(v_i)_i$ denotes the diagonal matrix with the diagonal entries $d_{ii} = v_i$ given by the coordinates of the vector $\mathbf{v} = (v_i)_i$.*

Definition 2.7. *$C(\mathbf{s}, \mathbf{t}) = (\frac{1}{s_i - t_j})_{i,j}$ is a Cauchy matrix. A matrix C has a Cauchy-like structure and the displacement rank of at most r if it can be represented as $\sum_{i=1}^{r} \operatorname{diag}(\mathbf{g}_i)C(\mathbf{s}, \mathbf{t}) \operatorname{diag}(\mathbf{h}_i)$.*

We recall the following properties.

Fact 2.1. *The determinant of a Vandermonde matrix $V(\mathbf{t}) = (t_i^{j-1})_{i,j=1}^{n}$ equals $\prod_{i<j}(t_i - t_j)$.*

Fact 2.2. *The determinant of a Cauchy matrix $C(\mathbf{s}, \mathbf{t}) = (\frac{1}{s_i - t_j})_{i,j}$ equals $\prod_{i<j}(s_i - s_j)(t_i - t_j)/\prod_{i,j}(s_i - t_j)$.*

In the above definitions $n \times n$ structured matrices are expressed bilinearly via the entries of the $n \times r$ generator matrices $G = (\mathbf{g}_i)_{i=1}^{r}$ and $H = (\mathbf{h}_i)_{i=1}^{r}$ where r is assumed to be much smaller than n, $r \leq 2$ for Toeplitz and Hankel matrices and equals one for Vandermonde and Cauchy matrices. Similar properties hold for the structured matrices below. Fast computations with all these structured matrices largely rely on such representations.

Definition 2.8. *(Cf. [GL96].) $B = (b_{i,j})_{i,j}$ is an (l, u) banded matrix having a lower bandwidth of at most $l = l(B)$ and an upper bandwidth of at most $u = u(B)$ if $b_{i,j} = 0$ where $i > j + l$ and where $j > i + u$. An $n \times n$ matrix B is banded if $l(B) \ll n$ and $u(B) \ll n$.*

Banded matrices are a special case of the next matrix classes.

Definition 2.9. *(Cf. [EG01], [VVGM05], and the bibliography therein.) A matrix M has a lower rank l (resp. upper rank u) if this is the maximal rank of its submatrices lying below (resp. above) its diagonal. $M = (m_{ij})_{i,j=1}^{n}$ is an (l, u) rank structured matrix (also called quasiseparable of order (l, u)) if it has a lower rank l and an upper rank u. Such a matrix has a trilinear (l, u) generator for its off-diagonal entries given by the set of $l \times l$ matrices A_h and $u \times u$ matrices B_h, vectors $\mathbf{p}_h$ and $\mathbf{q}_h$ of dimension l, and vectors $\mathbf{s}_h$ and $\mathbf{t}_h$ of dimension u for $h = 2, 3, \dots, n$ such that $m_{ij} = \mathbf{p}_i^T A_{ij} \mathbf{q}_j$, $1 \le j < i \le n$, $m_{ij} = \mathbf{s}_i^T B_{ij} \mathbf{t}_j$, $1 \le i < j \le n$, where $A_{ij} = A_{i-1} \cdots A_{j+1}$ for $i - 1 > j$, $A_{i+1,i} = I_l$, $B_{ij} = B_{j-1} \cdots B_{i+1}$ for $j - 1 > i$, $B_{i,i+1} = I_u$. (Clearly an (l, u) banded matrix is (l, u) rank structured but is not (l, u) semiseparable.)*

Definition 2.10. *An (l, u) rank structured $n \times n$ matrix has a bilinear (l, u) generator given by the vectors $\mathbf{p}_i$, $\mathbf{q}_j$, $\mathbf{s}_i$, and $\mathbf{t}_j$ for all i and j, if it has a trilinear (l, u) generator of Definition 2.9 where $A_h = I_l$ and $B_h = I_u$ for all h. In this case the subdiagonal (resp. superdiagonal) part of the matrix M (together with its diagonal) can be extended to a matrix of rank l (resp. u).*

Definition 2.11. *$\Phi(S)$, the fill of an $n \times n$ matrix S, is the set of its nonzero entries, and $|\Phi(S)|$ is its cardinality. Such a matrix is sparse if $|\Phi(S)| \ll n^2$. A fill degenerates if all matrices having this fill are singular. (Clearly, banded matrices are sparse and generically have nondegenerating fills.)*

3 How likely is a random integer matrix nonsingular modulo a fixed integer?

3.1 The cases of general and Toeplitz/Hankel matrices

We first recall the following estimate by Brent and McKay 1987 where $q = p^u$ for a prime p.

Theorem 3.1. *[BMK87, Corollary 2.2]. Write $P_k(r) = (1 - r)(1 - r^2) \cdots (1 - r^k)$, $r = 1/p$, $k > 1$, $P_0(r) = 1$. Then nonsingular matrices make up the fraction $\frac{P_{n+u-1}(r)}{P_{u-1}(r)}$ of all matrices in $\mathbb{Z}_{p^u}^{n \times n}$.*

Brent and McKay also supply similar estimates for any integer $q > 1$ and specify them for $n \to \infty$ and $q = 1, \dots, 16$. Our Table 5.4 in Section 5 shows the results of our tests for nonsingularity of random integer matrices in $\mathbb{Z}_q$, for $n = 5, 10, 50, 100, q = 2^g$, and $g = 0, 1, \dots, 20$. They are in reasonable agreement with the analytic estimates in [BMK87].

Such a study is much less developed for structured integer matrices. Best known is the following result.

Theorem 3.2. *[D60]. For any pair of a prime p and a positive integer n, the fraction of $1/p$ of all Toeplitz (as well as all Hankel) matrices in $\mathbb{Z}_p^{n \times n}$ are singular.*

We deduce from Theorem 3.2 the following corollary of independent interest.

Corollary 3.1. *For any pair of a prime p and a positive integer n, consider the space of the pairs of polynomials $u(x)$ and $v(x)$ over $\mathbb{Z}_p$ such that $\deg v(x) = n$, $\deg u(x) < n$. Then the pairs of coprime polynomials make up a fraction of $1 - 1/p$ in this space.*

Proof. We prove the corollary by combining the latter theorem with Proposition 9.1 on page 158 in the book [BP94]. This proposition defines a bijection map of all pairs (h, H) of $h \in \mathbb{Z}_p$ and nonsingular Hankel matrices H in $\mathbb{Z}_p^{n\times n}$ to all pairs of coprime polynomials $u(x)$ and $v(x)$ over $\mathbb{Z}_p$ where $v(x)$ is monic, $\deg v(x) = n$, and $\deg u(x) < n$. Combine the bijection $J : H \leftrightarrow T = HJ$ with Theorem 3.2 to count the number of pairs (h, H) where H is nonsingular in $\mathbb{Z}_p^{n\times n}$ and extend this count to the number of pairs of coprime polynomials $u(x)$ and $v(x)$ over $\mathbb{Z}_p$ to obtain the corollary. □

The proof technique above has been extended in [PMRW04] to yield the following result.

Theorem 3.3. *The fraction of at least $1 - n/q$ Toeplitz matrices in $\mathbb{Z}_q^{n\times n}$ are nonsingular for any integer $q > 1$.*

3.2 Unified treatment of structured integer matrices

The estimates for the probabilty of nonsingularity of a (generalized) Vandermonde matrix in $\mathbb{Z}_q$ for a prime q are implicit in [BKY07]. (These estimates are similar to the bounds in Theorem 3.3.) Besides this work and Theorem 3.3, we know of no other extensions of Theorem 3.2 to the cases of other structured matrices or rings $\mathbb{Z}_q$ for nonprime integers q, but we readily obtain a very general extension of this kind based on the following fundamental result, used in both [BKY07] and [PMRW04].

Theorem 3.4. *[DL78] (also cf. [Z79], [S80]). Suppose a multivariate polynomial of degree d does not vanish identically. Let the values of its variables be randomly sampled from a fixed set S. Then the polynomial vanishes with a probability of at most $\frac{d}{|S|}$.*

Corollary 3.2. *Assume integers $k > 0$ and $q > 1$ and matrices $A = (a_{i,j})_{i,j=1}^n$ where $a_{i,j}$ are polynomials over $\mathbb{Z}_q$ in some variables $t_1, \dots, t_k$ having total degrees $d_{i,j}$, $i, j = 1, \dots, n$. Write $d_i = \max_j d_{i,j}$ and $d_j = \max_i d_{i,j}$. Suppose for the variables $t_1, \dots, t_k$ ranging in $\mathbb{Z}_q$ the multivariate polynomial* $\det A$ *does not vanish identically. Then among all matrices A above the fraction of singular matrices is at most d/q for $d = \min\{\sum_{i=1}^n d_i, \sum_{j=1}^n d_j, \}$.*

Proof. Clearly, $\det A$ is a polynomial in the variables $t_1, \dots, t_k$ having degree of at most d. By assumption it does not vanish identically. It remains to invoke Theorem 3.4. □

Based on this corollary, we can estimate from above the fraction of singular matrices among all matrices of a given class. A singular matrix has vanishing determinant, which is a polynomial in the matrix entries or in other parameters defining the matrix. Our estimates apply to the classes of sparse and structured matrices that have nonsingular representatives. Many matrix classes (including Toeplitz, Toeplitz-like, banded and rank structured matrices) contain the identity matrix I which never degenerates in $\mathbb{Z}_q$. The class of sparse matrices with nongenerating fills contains permutation matrices P whose determinants equal one or -1. Likewise, the unit Hankel matrix J is a permutation matrix with $|\det J| = 1$. Thus we can apply the corollary to all respective classes of matrices. Finally, we can apply it to the classes of Vandermonde-like and Cauchy-like matrices where their basic Vandermonde and Cauchy matrices, respectively, are nonsingular modulo q (cf. Definitions 2.6 and 2.7 and Facts 2.1 and 2.2).

Let us estimate the degrees d in these cases. For Toeplitz and Hankel matrices as well as banded and other sparse matrices with a nondegenerating fill we have $d_{i,j} \le 1$ for all i and j and therefore $d = n$. Likewise, we can apply the corollary to the classes of Toeplitz-like and Hankel-like matrices as well as Vandermonde-like and Cauchy-like matrices whose basic (Vandermonde and Cauchy) matrices $V(\mathbf{t})$ and $C(\mathbf{s}, \mathbf{t})$, respectively, are fixed and are nonsingular modulo q (cf. Definitions 2.6 and 2.7). In these cases we have $d_{i,j} = 2$ for all i and j and therefore $d = 2n$. The same bound holds for (l, u) rank structured matrices having (l, u) bilinear generators. For general (l, u) rank structured matrices we have (l, u) trilinear generators, and therefore $d_{i,j} = 3$ for all i and j and $d = 3n$. Summarizing we deduce the following result.

Corollary 3.3. *The fraction of singular matrices in $\mathbb{Z}_q^{n\times n}$ for $q > 1$ is at most n/q among Toeplitz, Hankel and banded matrices and at most $2n/q$ among Toeplitz-like and Hankel-like matrices, (l, u) rank structured matrices having (l, u) bilinear generators, and Vandermonde-like and Cauchy-like matrices defined for fixed basic Vandermonde and Cauchy matrices $V(\mathbf{t})$ and $C(\mathbf{s}, \mathbf{t})$, respectively, and having representatives that are nonsingular modulo q. The fraction increases to $3n/q$ for the class of all (l, u) rank structured matrices.*

Next let us examine the class of Vandermonde matrices $V(\mathbf{t}) = (t_i^{j-1})_{i,j=1}^n$ defined by n parameters $t_1, \ldots, t_n$ and the respective class of Vandermonde--like matrices. In these cases we have $d_j = j - 1$, $d = (n - 1)n/2$ for the class of Vandermonde matrices and $d_j = j$, $d = (n+1)n/2$ for the matrices with Vandermonde-like structure. By applying Corollary 3.2, we obtain the following estimates.

Fact 3.1. *Suppose the class of Vandermonde matrices $V(\mathbf{t}) = (t_i^{j-1})_{i,j=1}^n$ in $\mathbb{Z}_q$ defined by n parameters $t_1, \ldots, t_n$ (resp. all Vandermonde-like in $\mathbb{Z}_q^{n\times n}$) has a nonsingular representative. Then the fraction of singular matrices in this class is at most $(n-1)n/(2q)$ (resp. $(n+1)n/(2q)$).*

For the class of Vandermonde matrices $V(\mathbf{t}) = (t_i^{j-1})_{i,j=1}^n$ defined by n parameters $t_1, \ldots, t_n$ consider the representative for $t_i = i - 1$, $i = 1, \ldots, n$.

In virtue of Fact 2.1 this representative is nonsingular in $\mathbb{Z}^{n\times n}$ wherever q is a prime exceeding $n-1$ or a product of such primes.

Based on Fact 2.1 one can prove that $\det V(\mathbf{t}) \mod q = 0$ if $q = p^l$ for positive integers l and smaller p unless l is quite large. We specify this claim only for $p = 2$.

Fact 3.2. *An integer Vandermonde matrix $V(\mathbf{t}) = (t_i^{j-1})_{i,j=1}^n$ is singular in $\mathbb{Z}_q$ for $q = 2^l$ and an integer $l \geq 1$ unless $l > (n-k)n/2$ where $k = \lfloor \log_2 n \rfloor$.*

Proof. Due to Fact 2.1, we just need to estimate the highest power of two that divides $\prod_{i<j}(t_i - t_j)$. Let there be exactly x even and exactly $n-x$ odd values among $t_1, \ldots, t_n$. Then there are exactly $(x-1)x/2 + (n-x-1)(n-x)/2 = x^2 - nx + (n^2-n)/2$ even numbers among the differences $t_i - t_j$ for all pairs i and $j > i$. The lower bound $(n-2)n/4$ is attained if $x = n/2$. We similarly estimate that at least $(n/2^i - 2)n/4$ of these differences are divided by 2^{i+1} for $i = 1, 2, \ldots, k-2$, and the lower bounds are attained for $t_j = j-1$, $j = 1, \ldots, n$ if $n = 2^k$. Therefore $\det V(\mathbf{t})$ is divided by 2^{w_n} for $w_n = \sum_{i=0}^{k-1}(n/2^i - 2)n/4 \geq (n-k)n/2$. □

Clearly Fact 3.2 does not extend to the matrices with Vandermonde-like structure with a displacement rank of at most $r > 1$, and this class does not seem to be too short of nonsingular representatives for all triples of n, r, and q.

We can further apply Corollary 3.2 to block matrices if they have the same block structure (of Toeplitz, Hankel, or Vandermonde types) or if they are block banded. Our estimates for the degree d in terms of the matrix dimension n would stay the same as in the case of scalars replacing blocks.

3.3 The case of structured rational matrices

To cover the important classes of Vandermonde-like, Cauchy and Cauchy-like matrices, where the coordinates of the basic vectors $\mathbf{s}$ and/or $\mathbf{t}$ are parameters, we extend Corollary 3.2 to matrices $A = (a_{i,j})_{i,j}$ with rational entries $a_{i,j} = b_{i,j}/c_{i,j}$ where $b_{i,j}$ and $c_{i,j}$ are pairs of coprime polynomials in a fixed set of variables $t_1, \ldots, t_k$. In this case $\det A = b/c$ where b and c are coprime polynomials in the variables $t_1, \ldots, t_k$, and we say that for a fixed set of values of these variables the matrix A degenerates if $b = 0$ and/or $c = 0$. Now Theorem 3.4 implies the following extension of Corollary 3.2.

Corollary 3.4. *Assume integers $k > 0$ and $q > 1$ and matrices $A = (a_{i,j})_{i,j=1}^n$ where $a_{i,j} = b_{i,j}/c_{i,j}$ and $b_{i,j}$ and $c_{i,j}$ are coprime polynomials over $\mathbb{Z}_q$ in some variables $t_1, \ldots, t_k$. Let $\det A = b/c$ where b and c are (nonvanishing) coprime polynomials in the variables $t_1, \ldots, t_k$ of total degrees d and δ, respectively. Then among all the above matrices A defined by the variables $t_1, \ldots, t_k$ ranging in $\mathbb{Z}_q$ the fraction of matrices that do not degenerate in $\mathbb{Z}_q$ is at least $(1 - d/q)(1 - \delta/q) = 1 - (d+\delta)/q + d\delta/q^2 > 1 - (d+\delta)/q$.*

Corollary 3.5. *For the classes of Vandermonde-like, Cauchy and Cauchy-like matrices we have $\delta = n^2$ (for all three classes) and $d = (n+1)n^3$, $d = (n-1)n$,*

and $d = (n+1)n$, respectively, and so the matrices that do not degenerate in $\mathbb{Z}_q$ make up the fractions of at least $(1-(n+1)n^3/q)(1-n^2/q) > 1-(n^2+n+1)n^2/q$, $(1-(n-1)n/q)(1-n^2/q) > 1-(2n-1)n/q$, and $(1-(n+1)n/q)(1-n^2/q) > 1-(2n+1)n/q$, respectively, as long as there are representatives of these classes that do not degenerate in $\mathbb{Z}_q$.

For the three cited classes of Vandermonde-like, Cauchy and Cauchy-like matrices we choose representatives where $s_i = i-1$ and/or $t_j = n-1+j$ for $i, j = 1, \dots, n$ and observe that they do not degenerate in $\mathbb{Z}_q$ (due to Facts 2.1 and 2.2) if q is a prime exceeding $2n-1$ (for Vandermonde-like matrices, it is sufficient if a prime q exceeds $n-1$) as well as if q is the product of such primes.

Cauchy matrices (and to a lesser extent also Vandermonde-like and Cauchy-like matrices) more frequently degenerate in $\mathbb{Z}_q$ where q is a smaller power of a smaller prime or the product of smaller number of such powers. For a simplified demonstration, consider Cauchy matrices $C(\mathbf{s}, \mathbf{t}) = (\frac{1}{s_i - t_j})_{i,j=1}^{n}$ for $q = 2$. Then no difference $s_i - t_j$ is allowed to be even. Therefore, all integers s_i must be even, whereas all integers t_j must be odd or vice versa. Then similarly to Fact 3.1, we deduce that the numerator $b = \prod_{i<j}(s_i - s_j)(t_i - t_j)$ of the determinant of the Cauchy matrix is divided by $2^{(2n-1-k)n}$ for $k = \lfloor \log_2 n \rfloor$, and so such a matrix degenerates.

3.4 Further comments

By applying Theorem 3.4 to the class of $n \times n$ structured matrices of Section 2, we yield the upper bound d/q on the probability of degeneration where d ranges from n for banded, Toeplitz and Hankel matrices to $(n+1)n/2$ for Vandermonde-like matrices, provided the basic vectors $\mathbf{s}$ and/or $\mathbf{t}$ for Vandermonde-like and Cauchy-like matrices are fixed and there are nonsingular matrices in these classes in $\mathbb{Z}_q^{n \times n}$. Therefore, the chances for the degeneration in the transition to the rings $\mathbb{Z}_q$ for larger integers q are quite remote on the average matrix in such classes. This is also confirmed by the results of our extensive experimental tests for nonsingularity of random Toeplitz, banded, and general integer matrices in $\mathbb{Z}_{p^w}^{n \times n}$ for $p = 2$, $w \le 20$, and $n \le 100$ (see Section 5).

For a prime q our upper bound on the probability of degeneration of Toeplitz and Hankel matrices is off by the factor n from Daykin's sharp bound in [D60], which is rather negligible for larger integers q, however. Our upper bounds on the probability of degeneration of structured matrices slightly exceed the sharp bounds in [BMK87], but comparison with Theorem 3.2 and the results of our tests show that degeneration is actually a little more likely for general than Toeplitz matrices.

4 Can we avoid matrix singularity modulo a fixed integer?

Suppose we have a subroutine that effectively inverts a nonsingular integer matrix modulo a fixed integer $q > 1$ (e.g., reasonably a large integer but still fitting the computer precision). Specifically, assume a fixed pair of a prime p (e.g., let $p = 2$) and an integer v defining the power $q = p^v$, let a matrix M be nonsingular modulo p, and let the selected subroutine first compute $M^{-1} \mod p$ by applying Gaussian or block Gaussian elimination or its specialization for structured matrices, by Morf 1974 and 1980 [M74] and [M80] and by Bitmead and Anderson 1980 [BA80] (cf. also Olshevsky and Pan 1998 [OP98], Pan 2001 [P01, Chapter 5], and the bibliography therein). Then let the subroutine apply Hensel's lifting to compute $M^{-1} \mod p^v$ and finally let it recover M^{-1} (cf. [MC79], [D82], [P02], [PMRW04], [PMRa], and Appendix B).

Now suppose that for such a fixed integer q we wish to apply the latter approach to invert a nonsingular integer matrix M, which becomes singular modulo q. Surely, we cannot succeed if we just apply our subroutine directly to this matrix, but we can perform modulo q the first recursive steps until we stop due to singularity. For a large class of input matrices, at this moment the remaining computations essentially amount to the simpler task of inverting some matrices of substantially smaller sizes.

Next let us describe an alternative recipe for yielding the same effect. (We can combine both recipes to enhance their power and to back up one another.) In our second recipe we rely on randomized additive preconditioning of the input matrix, that is, instead of the matrix M we can try to invert its small-rank modifications

$$M_i = M - U_i V_i, \quad i = 1, \dots, i_+. \tag{4.1}$$

Here we assume that U_i in $\mathbb{Z}_a^{n \times i}$ and V_i in $\mathbb{Z}_b^{i \times n}$ are random general or Toeplitz integer matrices or random integer matrices with the structure consistent with the structure of the input matrix; a and b are sufficiently large integers defined by n and $|M|$ (see Theorem 4.1); $i = 1, \dots, i_+$, and $i_+ = O(1)$ is a fixed relatively small positive integer. If we succeed for some $i \le i_+$, we can recover the inverse M^{-1} by applying the SMW formula, that is, the Sherman–Morrison–Woodbury formula [GL96, page 50],

$$M^{-1} = (M_i + U_i V_i)^{-1} = M_i^{-1} - M_i^{-1} U_i S_i^{-1} V_i M_i^{-1}, \quad S_i = I_i + V_i M_i^{-1} U_i.$$

The formula holds provided the matrices M_i and M are nonsingular for the pair of $n \times i$ matrices U_i and V_i^T, and in this case we have

$$\det M = (\det M_i) \det S_i, \tag{4.2}$$

so that S_i is an $i \times i$ nonsingular matrix.

Let us elaborate upon this idea assuming that a subroutine for inverting a nonsingular matrix modulo q is available. Fix two positive integers i_+ and j_+ (for simplicity the reader may first assume that $j_+ = 1$ and drop the subscripts j

below, but we have $j_+ = 15$ in Theorem 4.1). Recursively apply the subroutine to the matrices $M_{i,j} = M - U_{i,j}V_{i,j}$ for random matrices $U_{i,j}$ and $V_{i,j}$ for $i = 1$, $j = 1, \dots, j_+$; $i = 2$, $j = 1, \dots, j_+$; $\dots$, and so on. As soon as you yield the inverse modulo q of the matrix $M_i = M_{i,j}$ for some $i \le i_+$, $j \le j_+$, recover the inverse M_i^{-1} and then compute the inverse M^{-1} based on (4.1) and the SMW formula. If you have reached the bounds $i = i_+$ and $j = j_+$ without ever yielding the inverse, output FAILURE and stop. We refer to these computations as **Algorithm 4.1**, which we can extend to the solution of a linear system $M\mathbf{x} = \mathbf{b}$ by applying the SMW formula.

For random matrices M in $\mathbb{Z}^{n\times n}$, the algorithm is likely to succeed already for reasonably small integers i_+ and j_+ due to the two following theorems in [EGV00], which relate this likelihood to the choice of the bounds i_+ and j_+. Below we write "gcd" for "the greatest common divisor".

Theorem 4.1. *[EGV00, Theorem 3.8]. For two positive integers i and n, $i < n$, a nonsingular matrix M in $\mathbb{Z}^{n\times n}$, and two integers a and b such that $b \ge 2n^2\log_2(n|M|)$, $a \ge 21n^2b$, let $U_i = U_{i,j}$ and $V_i^T = V_{i,j}^T$ denote the pairs of random matrices in $\mathbb{Z}_a^{n\times i}$ for $j = 1, 2, \dots, 15$, and in $\mathbb{Z}_b^{n\times i}$ for $j = 16$, and let the matrices $M_i = M_{i,j}$ be defined by (4.1). Then with a probability of at least $1/2$, we have $s_{n-i}(M) = \gcd(s_n(M), \gcd_{j=1}^{16}(s_n(M_{i,j})))$. To increase the probability bound above $1 - \varepsilon$ for a fixed positive ε, it is sufficient to include j_+ matrices $M_{i,j}$, $j = 1, \dots, j_+$, for every i and for a sufficiently large j_+ in $O(\log(1/\varepsilon))$.*

Theorem 4.2. *[EGV00, Theorem 6.2]. For a fixed pair of integers $\lambda > 0$ and η, let the entries of an $n \times n$ matrix M be independently sampled under the uniform probability distribution in a set of integers $\eta, \eta+1, \dots, \eta+\lambda-1$. Then* $\text{Probability}(s_{n-j}(M) > 1) \le \lambda^{-n} + 9(\frac{2}{3})^{j-1} + \frac{n^3}{\lambda^{j-1}}$.

Due to Theorems 4.1 and 4.2 (and also according to the well known statistics), with a high probability we have

$$\gcd\left(s_n(M), \gcd_{j=1}^{j_+}(s_n(M_{i,j}))\right) = 1$$

for a random $n \times n$ integer matrix M, the matrices $M_{i,j}$ defined above, some $i \le i_+$, and reasonably small integers i_+ and j_+. In fact we just need a weaker property that the above gcd is likely to be coprime with a fixed prime p, and this property has been statistically observed in our experiments with random integer general, Toeplitz, and banded matrices for $p = 2$, $q = 2^v$ (see the next section).

Remark 4.1. *The matrices S_i are expected to be of a smaller size than the input matrix M, but we can try to simplify their inversion further by applying to them Algorithm 4.1. If $q = p^v$, then the transition $M \to S_i$ (besides the matrix size reduction) decreases the order of the prime p in the value of the determinant wherever p divides* $\det M_i$ *(see equation (4.2)).*

5 Experimental computations: how frequently are random integer matrices nonsingular modulo a power of two?

In our tests we have randomly generated an $n \times n$ Toeplitz matrix $M = (t_{i-j})_{i,j}$. Its entries $t_{1-n}, \dots, t_{n-1}$ have been chosen independently of each other under the uniform random distribution on $\mathbb{Z}_q$ for $q = 2^w$ and for a positive integer w. The first column in each of Tables 5.1–5.3 shows how frequently in our tests a random $n \times n$ integer Toeplitz matrix M was nonsingular in $\mathbb{Z}_q$.

Whenever the test showed singularity, we repeated the test recursively (up to at most four times), each time adding the outer product of two random vectors to the input matrix. The *(1+i)*th column of each table, for $i = 0, 1, 2, 3, 4$, shows for how many out of 100,000 samples the results were positive for the matrices $M - U_j V_j^T$, for some $j \le i$ where $U_j, V_j^T \in \mathbb{Z}_q^{n \times l}$, $M \in \mathbb{Z}_q^{n \times n}$, $q = 2^w$. These data should motivate using Algorithm 4.1 for smaller i_+ and j_+. They should be compared with similar statistics for general, tridiagonal, and five-diagonal matrices. Table 5.4 shows such statistics, although without small rank perturbations. According to our tests in the case where $q = 2^w$, the degeneration is more likely for five-diagonal than general matrices, and is even more likely for tridiagonal matrices, but even for the latter matrices it is quite rare for larger w. We have also observed for tridiagonal matrices that the degeneration is substantially less likely where we shift from $q = 2^w$ to $q = 5^w$.

Table 5.1: Number of times the matrix $M + A_i$ for a random 20×20 Toeplitz matrix M and a random 20×20 matrix A_i of a rank of at most i is nonsingular in the ring $\mathbb{Z}_q$ for $q = 2^w$ out of 100,000 samples

$w\backslash i$	0	1	2	3	4
1	50173	59450	66672	72514	77452
2	68814	80808	87785	92256	95133
3	82971	92311	96197	98164	99136
4	90559	96899	98862	99567	99852
5	95079	98809	99671	99907	99973
6	97333	99557	99907	99981	99997
7	98643	99859	99973	99998	100000
8	99302	99948	99993	99999	100000
9	99639	99983	100000	100000	100000
10	99816	99997	100000	100000	100000
11	99903	99999	100000	100000	100000
12	99955	100000	100000	100000	100000

Table 5.2: Number of times the matrix $M + A_i$ for a random 50×50 Toeplitz matrix M and a random 50×50 matrix A_i of a rank of at most i is nonsingular in the ring $\mathbb{Z}_q$ for $q = 2^w$ out of 100,000 samples

$w\backslash i$	0	1	2	3	4
1	50054	59383	66661	72665	77581
2	68781	80792	87812	92341	95151
3	82842	92263	96282	98203	99139
4	90507	96868	98877	99589	99844
5	95132	98846	99695	99915	99976
6	97440	99597	99912	99981	99994
7	98667	99857	99972	99994	99998
8	99315	99953	99989	99997	99999
9	99653	99985	100000	100000	100000
10	99829	99997	100000	100000	100000
11	99917	99999	100000	100000	100000
12	99967	100000	100000	100000	100000

Table 5.3: Number of times the matrix $M + A_i$ for a random 100×100 Toeplitz matrix M and a random 100×100 matrix A_i of a rank of at most i is nonsingular in $\mathbb{Z}_q$ for $q = 2^w$ out of 100,000 samples

	0	1	2	3	4
1	50170	59672	66652	72460	77368
2	68969	80960	87833	92188	95130
3	82799	92261	96240	98128	99122
4	90498	96935	98884	99570	99845
5	94975	98837	99662	99893	99971
6	97255	99547	99898	99970	99991
7	98591	99827	99966	99994	99998
8	99249	99931	99989	99998	99998
9	99616	99976	99997	100000	100000
10	99804	99994	100000	100000	100000
11	99898	99998	100000	100000	100000
12	99948	100000	100000	100000	100000

Table 5.4: Number of times a random $n \times n$ general matrix M is nonsingular in the ring $\mathbb{Z}_q$ out of $100,000$ samples for $q = 2^w$

w	$n = 5$	$n = 10$	$n = 50$	$n = 100$
$w = 0$	29,986	28,781	28,940	28,781
$w = 1$	58,637	57,679	57,884	57,782
$w = 2$	77,650	76,817	77,047	77,104
$w = 3$	88,399	87,916	88,000	88,080
$w = 4$	94,102	93,888	93,943	93,921
$w = 5$	97,046	96,911	96,963	96,937
$w = 6$	98,519	98,414	98,483	98,452
$w = 7$	99,245	99,180	99,212	99,235
$w = 8$	99,634	99,598	99,590	99,620
$w = 9$	99,820	99,791	99,783	99,806
$w = 10$	99,911	99,894	99,892	99,899
$w = 11$	99,956	99,957	99,950	99,953
$w = 12$	99,977	99,977	99,978	99,980
$w = 13$	99,985	99,992	99,991	99,992
$w = 14$	99,992	99,996	99,993	99,995
$w = 15$	99,993	99,997	99,996	99,998
$w = 16$	99,995	99,999	99,999	99,998
$w = 17$	99,998	99,999	99,999	99,998
$w = 18$	99,999	100,000	99,999	99,999
$w = 19$	99,999	100,000	100,000	100,000
$w = 20$	99,999	100,000	100,000	100,000

Table 5.5: Number of times a random $n \times n$ tridiagonal matrix M is nonsingular in the ring $\mathbb{Z}_q$ out of $1,000,000$ samples for $q = 2^w$

w	$n = 5$	$n = 10$	$n = 50$	$n = 100$
$w = 1$	117356	17514	0	0
$w = 2$	320625	75599	0	0
$w = 3$	531878	182052	1	0
$w = 4$	703335	324629	4	0
$w = 5$	823672	478421	17	0
$w = 6$	899773	620734	64	0
$w = 7$	945210	738216	188	0
$w = 8$	970459	828122	494	0
$w = 9$	984437	891854	1324	0
$w = 10$	991892	934290	3043	0
$w = 11$	995862	961334	6210	0
$w = 12$	997903	978026	11855	0
$w = 13$	998940	987761	20951	0
$w = 14$	999455	993236	34980	1
$w = 15$	999719	996395	54784	1
$w = 16$	999859	998089	82128	7
$w = 17$	999920	999012	117742	13
$w = 18$	999962	999515	161178	22
$w = 19$	999980	999743	213241	37
$w = 20$	999988	999866	271703	72
$w = 21$	999996	999947	336451	138
$w = 22$	999999	999973	404624	289
$w = 23$	999999	999986	474595	520
$w = 24$	1000000	999992	543974	941
$w = 25$		999996	610648	1601
$w = 26$		999996	673129	2629
$w = 27$		999999	730268	4193
$w = 28$		1000000	780932	6542
$w = 29$			825245	9787
$w = 30$			862955	14404
$w = 31$			893896	20884
$w = 32$			919407	29409

Table 5.6: Number of times a random $n \times n$ tridiagonal matrix M is nonsingular in the ring $\mathbb{Z}_q$ out of $1,000,000$ samples for $q = 5^w$

w	$n = 5$	$n = 10$	$n = 50$	$n = 100$
$w = 1$	629193	453573	33036	1319
$w = 2$	902825	795859	142767	9536
$w = 3$	978126	939323	326520	36124
$w = 4$	995314	984670	537288	94676
$w = 5$	999063	996494	720272	192946
$w = 6$	999808	999263	849531	324563
$w = 7$	999964	999840	927780	473054
$w = 8$	999992	999968	968268	618013
$w = 9$	999997	999991	987267	741988
$w = 10$	1000000	999999	995305	837468
$w = 11$	1000000	1000000	998419	904365
$w = 12$	1000000	1000000	999479	947171
$w = 13$	1000000	1000000	999829	972681

Table 5.7: Number of times a random $n \times n$ tridiagonal matrix M is nonsingular in the ring $\mathbb{Z}_q$ out of $1,000,000$ samples for $q = 10^w$

w	$n = 5$	$n = 10$	$n = 50$	$n = 100$
$w = 1$	673132	462691	33063	1286
$w = 2$	934076	811362	142940	9469
$w = 3$	989846	950586	326682	36115
$w = 4$	998658	989599	537027	95083
$w = 5$	999834	998176	720163	193223
$w = 6$	999981	999689	850071	325041
$w = 7$	999997	999950	927627	473759
$w = 8$	1000000	999991	968218	618106
$w = 9$	1000000	999998	987183	742050

Table 5.8: Number of times a random $n \times n$ five-diagonal matrix M is nonsingular in the ring $\mathbb{Z}_q$ out of $1,000,000$ samples for $q = 2^w$

w	$n = 5$	$n = 10$	$n = 50$	$n = 100$
$w = 1$	291205	138605	407	0
$w = 2$	554299	353025	3189	1
$w = 3$	744030	560876	12666	28
$w = 4$	860894	725098	35279	131
$w = 5$	926953	837646	77277	534
$w = 6$	962364	908618	141802	1617
$w = 7$	980846	950327	227489	4233
$w = 8$	990280	973763	327674	9452
$w = 9$	995111	986273	433370	19223
$w = 10$	997602	992932	537207	35864
$w = 11$	998787	996416	631600	61342
$w = 12$	999400	998230	713344	97265
$w = 13$	999694	999130	781426	144238
$w = 14$	999861	999545	836186	201805
$w = 15$	999936	999768	878746	268132
$w = 16$	999969	999880	911666	340560
$w = 17$	999984	999938	936426	415820
$w = 18$	999993	999968	955325	490841
$w = 19$	999996	999986	969282	562743
$w = 20$	999998	999994	979171	628840
$w = 21$	999999	999999	986186	687759
$w = 22$	1000000	1000000	991041	738768
$w = 23$			994296	782062
$w = 24$			996454	818466
$w = 25$			997802	849199
$w = 26$			998675	874827
$w = 27$			999193	896088
$w = 28$			999547	914020
$w = 29$			999746	929000
$w = 30$			999855	942334
$w = 31$			999914	953521
$w = 32$			999957	962893

Analysis of the results of the experiments

For fixed q and n, we assume that M is singular over $\mathbb{Z}_q$ with a probability p. Next we estimate p. Let x be a random variable such that

$$x = \begin{cases} 1, & \det M = 0 \bmod q; \\ 0, & \det M \neq 0 \bmod q. \end{cases}$$

Let $x_1, \ldots, x_m$ be the observed values of x. By the Central Limit Theorem,

$$\lim_{m\to\infty} \frac{(x_1 + \ldots + x_m) - mp}{\sqrt{mp(1-p)}} = N(0,1)$$

where $N(0,1)$ is the standard normal probability distribution. Therefore, a confidence interval of probability $1-\alpha$ for p is

$$\left(\bar{x} - Z_{\alpha/2}\sqrt{\bar{x}(1-\bar{x})/m},\ \bar{x} + Z_{\alpha/2}\sqrt{\bar{x}(1-\bar{x})/m}\right)$$

where $\bar{x} = \frac{1}{m}(x_1 + \ldots + x_m)$, Z_α is defined by Probability$(N(0,1) > Z_\alpha) = \alpha$.

Example 5.1. *For $g = 8, n = 50$, we are "99.9%" sure that*

- *Probability(Toeplitz matrix M is nonsingular)* $= 0.993 \pm 0.001$*;*
- *Probability(Toeplitz matrix M is strongly nonsingular)* $= 0.731 \pm 0.005$*;*
- *Probability(general matrix M is nonsingular)* $= 0.992 \pm 0.001$*;*
- *Probability(general matrix M is strongly nonsingular)* $= 0.688 \pm 0.005$.

Appendix

A How likely is a fixed integer matrix singular in $\mathbb{Z}_{p^v}$ for a random prime p?

Hereafter ln denotes the natural logarithm $\log_e$, for $e = 2.71828128...$ Let us assume a fixed nonsingular integer matrix M and a random prime p, fix a positive integer v, and combine some known techniques to estimate the probability that p^v divides $\det M$.

We begin with some definitions and basic lemmas. Hereafter $\ln = \log_e$ stands for the natural logarithms (with the base $e = 2.718281\ldots$), $\pi(y)$ denotes the number of primes not exceeding y, $[a,b]$, (a,b), and $(a,b]$ denote closed, open, and semi-open real line intervals with the end points a and b, respectively.

Lemma A.1. *(See also (A.4).) If $y > 114$, then $1 < \frac{\pi(y)}{y} \ln y < 1.25$.*

Proof. See Rosser and Schoenfeld 1962 [RS62]. □

Lemma A.2. *Let $y \geq 114$, then $\pi(y) - \pi(\frac{y}{20}) > (1/\tilde{\beta})\frac{y}{\ln y}$ for*

$$\tilde{\beta} = \frac{1}{1-\tilde{\alpha}} = 1.2049303\ldots, \quad \tilde{\alpha} = \frac{\ln 114}{16\ln 5.7} = 0.17007650\ldots. \tag{A.1}$$

Proof. By Lemma A.1, we have $\pi(y) - \pi(\frac{y}{20}) > \frac{y}{\ln y} - \frac{1.25y}{20\ln(y/20)}$. Observe that $\frac{\ln(y/20)}{\ln y}$ is monotone increasing as y grows. So $\frac{1.25}{20\ln(y/20)} \leq \frac{\tilde{\alpha}}{\ln y}$ for $\tilde{\alpha}$ in (A.1) and $y \geq 114$. Combine the above estimates. □

Lemma A.3. *(Cf. Corollary 7.8.2 in [P01].) Let y, v, h, and k be positive integers such that*

$$y \geq 114, \quad 0 < h^{1/k} \leq y/20. \tag{A.2}$$

Let p be a random prime selected in the range $(y/20, y]$ under the uniform probability distribution. Then Probability$(h \bmod p^v = 0) < \frac{\tilde{\beta}k\ln y}{vy}$ *for $\tilde{\beta}$ in (A.1).*

Proof. Suppose that in the above range there are exactly l distinct primes whose vth powers divide h. Then the product of these powers also divides h, and therefore we have $h \geq (\frac{y}{20})^{vl}$ because each of the l primes lying in the range $[y/20, y]$ is at least as large as $\frac{y}{20}$. On the other hand, $h \leq (\frac{y}{20})^k$ by assumption. Therefore, $vl \leq k$, that is, $l \leq k/v$. Compare the latter upper bound on l with the lower bound in Lemma A.2 on the overall number of primes in the range $(\frac{y}{20}, y]$. □

Theorem A.1. *(Cf. Corollary 7.8.3 in [P01].) Suppose that ϵ is a positive number, v is a positive integer, $M \in \mathbb{Z}^{n\times n}$ is nonsingular, and a prime p is randomly sampled from the range $(y/20, y]$ under the uniform probability distribution in this range where $y = \frac{n\xi\ln|M|}{v\epsilon} \geq 114$ and $\xi = \frac{16\ln 114}{16\ln 5.7 - \ln 114} = 16\tilde{\alpha}\tilde{\beta} = 3.278885\ldots$ for $\tilde{\alpha}$ and $\tilde{\beta}$ in (A.1). Then we have*

$$P = \text{Probability}((\det M) \bmod p^v = 0) < \epsilon. \tag{A.3}$$

Proof. Write $h = |\det M|$, $k = \frac{n\ln|M|}{\ln(y/20)}$, so that $h \leq |M|^n$ and $k\ln\frac{y}{20} \geq \ln h$, which implies (A.2). Apply Lemma A.3 and deduce that

$$P < \frac{\tilde{\beta}k\ln y}{uy} = \frac{\tilde{\beta}n\ln|M|}{v\ln(y/20)}\frac{\ln y}{y} = \frac{v\epsilon\tilde{\beta}n\ln|M|}{vn\ln|M|}\frac{\ln y}{\xi\ln(y/20)} = \frac{\epsilon\tilde{\beta}\ln y}{\delta\ln(y/20)}.$$

Note that

$$\frac{\ln y}{\ln(y/20)} \leq \frac{\ln 114}{\ln 5.7}$$

for $y \geq 114$. Therefore

$$P < \epsilon\frac{\tilde{\beta}\ln 114}{\xi\ln 5.7} = \frac{16\tilde{\alpha}\tilde{\beta}}{\xi}\epsilon = \epsilon.$$

□

To extend the above results to smaller y, one can exploit the known extensions of Lemma A.1, e.g.,

$$1 + \frac{1}{2\ln y} < \pi(y)\frac{\ln y}{y} < 1 + \frac{3}{2\ln y} \tag{A.4}$$

for $y \geq 59$ [GG03, Theorem 18.7]. Refined estimates for $\pi(y)$ can be found in Karatsuba 1990 [K90].

Let us extend Theorem A.1 to any integer q instead of $q = p^v$. We rely on the following observation.

Lemma A.4. *Let p and q be coprime and let u, v, and h be three positive integers. Then $p^u q^v$ divides h if and only if both p^u and q^v divide h.*

Corollary A.1. *Let p_1, ..., p_h be h distinct primes sampled randomly and independently in the ranges $(y_i/20, y_i]$, $i = 1, \ldots, h$, respectively, under the uniform probability distribution. Here $y_i = \frac{\xi n}{2v_i\epsilon}\ln|M| \geq 114$ for ξ in Theorem A.1 and $i = 1, \ldots, h$; the matrix $M \in \mathbb{Z}^{n\times n}$ is nonsingular; $v_1, \ldots, v_h$ are positive integers, and*

$$2h - 2 \leq \frac{y_i}{\tilde{\beta}\ln y_i} \tag{A.5}$$

for $\tilde{\beta}$ in Lemma A.2 and for all i. Then we have

$$P = \text{Probability}(p_1^{v_1}\cdots p_h^{v_h} \text{ divides } \det M) \leq \epsilon^h.$$

Proof. Corollary A.1 follows from Lemma A.4 and Theorem A.1 for $y = y_i$ and $v = 2v_i$. The primes $p_1, \ldots, p_{i-1}$ are excluded from the range $(y_i/20, y_i]$ for every i; this decreases the overall number of primes in this range but less than by twice for $i \leq h$ because of (A.5) and Lemma A.2. The effect of this decrease on the probability estimates is overweighed by the increase of v from v_i to $2v_i$. □

Remark A.1. *Under the assumptions of Theorem A.1 and Corollary A.1, an integer matrix M is strongly nonsingular in $\mathbb{Z}_q^{n\times n}$ for $q = p^v$ or $q = p_1^{v_1}\cdots p_k^{v_k}$ with a probability which is within the factor of n from the respective bounds in Theorem A.1 and Corollary A.1.*

B Hensel's lifting modulo a prime and prime power

Hensel's lifting [MC79], [D82] is one of the most effective techniques for the exact solution of linear systems $M\mathbf{x} = \mathbf{b}$ of equations with integral or rational input values.

It begins with computing the inverse M^{-1} modulo a random prime p, e.g., from the range $(y, y/20)$ for $y = O(n\ln|M|)$ in Theorem A.1. Such a choice of a prime p ensures that a nonsingular integer matrix M is likely to remain

nonsingular modulo p. The computation of (a short generartor for) the inverse M^{-1} modulo p (with the precision $\lceil \log_2 p \rceil$) is inexpensive for a prime p in the above range.

Then in h lifting steps the solution modulo p^h is computed. Every lifting step amounts essentially to multiplication of the matrices M and M^{-1} by two vectors with the precision in $O(\log(p + |M| + |\mathbf{b}|))$.

Finally, the exact rational solution is readily reconstructed provided h is large enough (of the order of $O(n \log(n|M| + |\mathbf{b}|))$).

It is practicallly attractive to perform lifting in binary, that is, modulo powers of two. The respective modification of lifting was worked out in [P02] (cf. also [PMRW04]) and enabled its substantial acceleration, according to the extensive tests by the authors of the papers [PMRa], [PMRb] as well as by J.-C. Dumas and A. Urbanska in Grenoble, France. Instead of the initialization with the matrix $M^{-1} \mod p$ for a random prime p, binary lifting in [P02] and [PMRW04] uses a matrix Q such that

$$MQ \mod (qs) = qI \tag{B.1}$$

where q and s are appropriate powers of two. If $(\det M)$ is an odd integer, then the latter equation holds for $q = 1$ and $Q = M^{-1} \mod s$, but even if $\gcd(\det M, s) = b > 1$, we can satisfy equation (B.1) as long as b divides q.

To keep the cost of the initialization stage reasonably low, we should not use too high a precision $\lceil \log_2(qs) \rceil$. To yield matrix Q that satisfies equation (B.1), however, we should not allow qs to divide $s_n(M)$, Smith's largest factor of an input matrix M. We can exclude this problem by choosing qs exceeding $(n|M|)^{n/2} \geq |\det M| \geq s_n(M)$ (cf. Definition 2.3 and equation (2.1)), but according to the analysis and experiments in this paper much smaller integers qs are sufficient for the average integer matrix with the structures of Toeplitz--like, Hankel-like, band and other types that we studied, that is on the avearage one can satisfy equation (B.1) even for $qs = 2^k$ and k of the order of $\log_2 n$. This precision level is low enough to enable the solution with binary Hensel's lifting at a nearly optimal Boolean cost.

Acknowledgements. We thank Richard Isaac for suggesting a format for the statistical tests reported in Section 5, Jean-Guillaume Dumas and Anna Urbanska for their data on the CPU time for binary and nonbinary lifting, a reviewer for helpful comments and pointing out the paper [BKY07], and Aggelos Kiayias and Moti Yung for sending us a reprint of this paper.

References

[BA80] R. R. Bitmead, B. D. O. Anderson, Asymptotically Fast Solution of Toeplitz and Related Systems of Linear Equations, *Linear Algebra and Its Applications*, **34**, 103–116, 1980.

[BGY80] R. P. Brent, F. G. Gustavson, D. Y. Y. Yun, Fast Solution of Toeplitz Systems of Equations and Computation of Padé Approximations, *J. Algorithms*, **1**, 259–295, 1980.

[BKY07] D. Bleichenbacher, A. Kiayias, M. Yung, Decoding Interleaved Reed-Solomon Codes over Noisy Channels, *Theoretical Computer Science (Special Issue: Selected Papers from ICALP 2003)*, **379**, **3**, 348–360, 2007.

[BMK87] R. P. Brent, B. D. McKay, Determinants and Ranks of Random Matrices over $\mathbb{Z}_m$, *Discrete Math.*, **66**, 35–49, 1987.

[BP94] D. Bini, V. Y. Pan, *Polynomial and Matrix Computations, Volume. 1: Fundamental Algorithms*, Birkhäuser, Boston, 1994.

[D60] D. E. Daykin, Distribution of Bordered Persymmetric Matrices in a Finite Field. *J. Reine und Angewandte Math.,* **203**, 47–54, 1960.

[D82] J. D. Dixon, Exact Solution of Linear Equations Using p-adic Expansions, *Numerische Math.*, **40**, 137–141, 1982.

[DL78] R. A. Demillo, R. J. Lipton, A Probabilistic Remark on Algebraic Program Testing, *Information Processing Letters*, **7(4)**, 193–195, 1978.

[EG99] Y. Eidelman, I. Gohberg, Linear Comlexity Inversion Algorithms for a Class of Structured Matrices, *Integral Equations and Operator Theory*, **35**, 28–52, Birkhäuser, Basel, 1999.

[EG01] Y. Eidelman, I. Gohberg, Fast Inversion Algorithms for a Class of Block Structured Matrices, *Contemporary Mathematics*, **281**, 17–38, 2001.

[EGV00] W. Eberly, M. Giesbrecht, G. Villard, On Computing the Determinant and Smith Form of an Integer Matrix, *Proc. 41st Annual Symposium on Foundations of Computer Science (FOCS'2000)*, 675–685, IEEE Computer Society Press, Los Alamitos, California, 2000.

[GG03] J. von zur Gathen, J. Gerhard, *Modern Computer Algebra*, Cambridge University Press, Cambridge, UK, 2003 (second edition).

[GL96] G. H. Golub, C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, Maryland, 1996 (third addition).

[K90] A. A. Karatsuba, The Distribution of Prime Numbers, *Russian Math. Surveys*, **45**, 99–171, 1990.

[KL96] E. Kaltofen, A. Lobo, On Rank Properties of Toeplitz Matrices over Finite Fields, *Proceedings of International Symposium on Symbolic and Algebraic Computation (ISSAC'96)*, 241–249, ACM Press, New York, 1996.

[M74] M. Morf, Fast Algorithms for Multivariable Systems, Ph.D. Thesis, *Department of Electrical Engineering, Stanford University,* Stanford, CA, 1974.

[M80] M. Morf, Doubling Algorithms for Toeplitz and Related Equations, *Proceedings of IEEE International Conference on ASSP*, 954–959, IEEE Press, Piscataway, New Jersey, 1980.

[MC79] R. T. Moenck, J. H. Carter, Approximate Algorithms to Derive Exact Solutions to Systems of Linear Equations, *Proceedings of EUROSAM, Lecture Notes in Computer Science*, **72**, 63–73, Springer, Berlin, 1979.

[N72] M. Newman, *Integral Matrices*, Academic Press, New York, 1972.

[OP98] V. Olshevsky, V. Y. Pan, A Unified Superfast Algorithm for Boundary Rational Tangential Interpolation Problem and for Inversion and Factorization of Dense Structured Matrices, *Proc. 39th Annual IEEE Symposium on Foundation of Computer Science (FOCS 98)*, 192–201, IEEE Computer Society Press, Los Alamitos, California, 1998.

[P01] V. Y. Pan, *Structured Matrices and Polynomials: Unified Superfast Algorithms*, Birkhäuser/Springer, Boston/New York, 2001.

[P02] V. Y. Pan, Can We Optimize Toeplitz/Hankel Computations? *Proc. of the Fifth International Workshop on Computer Algebra in Scientific Computing (CASC'02)*, Yalta, Crimea, Sept. 2002 (E. W. Mayr, V. G. Ganzha, E. V. Vorozhtzov, Editors), 253–264, *Technische Universität München*, Germany, 2002.

[P02a] V. Y. Pan, Nearly Optimal Toeplitz/Hankel Computations, Technical Reports 2002001, 2002002, and 2002017. *Ph.D. Program in Computer Science, The Graduate Center, CUNY*, New York, 2002. Available at http://www.cs.gc.cuny.edu/tr/techreport.php?id=352

[P04] V. Y. Pan, On Theoretical and Practical Acceleration of Randomized Computation of the Determinant of an Integer Matrix, *Zapiski Nauchnykh Seminarov POMI* (in English), **316**, 163-187, St. Petersburg, Russia, 2004. Also available at http://comet.lehman.cuny.edu/vpan/

[PMRW04] V. Y. Pan, B. Murphy, R. E. Rosholt, X. Wang. Toeplitz and Hankel Meet Hensel and Newton Nearly Optimal Algorithms and Their Practical Acceleration with Saturated Initialization, Technical Report 2004013, *PhD Program in Computer Science, The Graduate Center, CUNY*, New York, 2004. Available at http://www.cs.gc.cuny.edu/tr/techreport.php?id=352

[PMRa] V. Y. Pan, B. Murphy, R. E. Rosholt, Unified Nearly Optimal Algorithms for Structured Integer Matrices and Polynomials, Technical Report 2008003, *PhD Program in Computer Science, The Graduate Center, CUNY*, New York, 2008. Available at http://www.cs.gc.cuny.edu/tr/techreport.php?id=352

[PMRb] V. Y. Pan, B. Murphy, R. E. Rosholt, Unified Nearly Optimal Algorithms for Structured Integer Matrices, *Numerical Methods for Structured Matrices and Applications: Georg Heinig memorial volume*, Birkhäuser Verlag, in press.

[PW02] V. Y. Pan, X. Wang, Acceleration of Euclidean Algorithm and Extensions, *Proceedings of the International Symposium on Symbolic and Algebraic Computation (ISSAC'02)*, (Teo Mora editor), 207–213, ACM Press, New York, 2002.

[RS62] J. B. Rosser, L. Schoenfeld, Approximate Formulas of Some Functions of Prime Numbers, *Illinois J. of Math.*, **6**, 64–94, 1962.

[S80] J. T. Schwartz, Fast Probabilistic Algorithms for Verification of Polynomial Identities, *Journal of ACM*, **27(4)**, 701–717, 1980.

[SNC07] *Symbolic-Numeric Computation*, (Dongming Wang and Li-Hong Zhi, editors), Birkhäuser, Basel/Boston, 2007.

[SNC07a] *Proceedings of the 3rd Intern. Workshop on Symbolic-Numeric Computation (SNC2007)*, July 2007, London, Ontario, Canada (J. Verschelde and S. Watt, eds.), ACM Press, New York, 2007.

[TCS08] *Theoretical Computer Science*, Special Issue on Symbolic–Numerical Algorithms (D. A. Bini, V. Y. Pan, and J. Verschelde, editors), (in press).

[VVGM05] R. Vandebril, M. Van Barel, G. Golub, N. Mastronardi, A Bibliography on Semiseparable Matrices, *Calcolo*, **42, 3–4,** 249–270, 2005.

[W02] X. Wang, How Frequently is the Matrix Nonsingular, Technical Report 2002018, *PhD Program in Computer Science, The Graduate Center, CUNY*, New York, 2002.

[Z79] R. E. Zippel, Probabilistic Algorithms for Sparse Polynomials, *Proceedings of EUROSAM'79, Lecture Notes in Computer Science*, **72**, 216–226, Springer, Berlin, 1979.