2012

TR-2012008: Randomized Matrix Computations III

Victor Y. Pan

Guoliang Qian

Ai-Long Zheng

Randomized Matrix Computations III *

Victor Y. Pan[1,2],[a], Guoliang Qian[2],[b], and Ai-Long Zheng[2],[c]

[1] Department of Mathematics and Computer Science
Lehman College of the City University of New York
Bronx, NY 10468 USA
[2] Ph.D. Programs in Mathematics and Computer Science
The Graduate Center of the City University of New York
New York, NY 10036 USA
[a] victor.pan@lehman.cuny.edu
http://comet.lehman.cuny.edu/vpan/
[b] gqian@gc.cuny.edu
[c] azheng-1999@yahoo.com

Abstract

It is well known that random matrices tend to be well conditioned, and we employ this property to advance some fundamental matrix computations. We prove effectiveness of our novel techniques of randomized preconditioning, estimate the condition numbers of random Toeplitz and circulant matrices, numerically stabilize Gaussian elimination with no pivoting and block Gaussian elimination, compute 2×2 matrix factorization where both diagonal blocks are better conditioned than an ill conditioned input matrix, and apply our dual variation of the Sherman–Morrison–Woodbury formula to low-rank matrix approximation. Our formal study and numerical tests show significant progress versus the known algorithms and should motivate further research efforts.

2000 Math. Subject Classification: 15A52, 15A12, 15A06, 65F22, 65F05, 65F10

Key Words: Random matrices, Preconditioning, Numerical rank

1 Introduction

It is well known that random matrices tend to be well conditioned [D88], [E88], [ES05], [CD05], [SST06], [B11], and we employ this property to advance some fundamental matrix computations. We estimate the condition numbers of random Toeplitz and circulant matrices, propose and analyze some new techniques of randomized preprocessing, prove that they precondition a large and important class of ill conditioned matrices A, apply these results to the approximation of the trailing singular spaces of these matrices A, their triangulation and the solution of linear systems $A\mathbf{y} = \mathbf{b}$, numerically stabilize Gaussian elimination with no pivoting and block Gaussian elimination, extend the Sherman–Morrison–Woodbury formula, and apply this extension to approximation by low-rank matrices and by structured matrices. Our analysis and experiments show substantial progress versus the known algorithms and should motivate further research efforts.

*Some results of this paper have been presented at the ACM-SIGSAM International Symposium on Symbolic and Algebraic Computation (ISSAC '2011), San Jose, CA, 2011, the 3nd International Conference on Matrix Methods in Mathematics and Applications (MMMA 2011) in Moscow, Russia, June 22-25, 2011, the 7th International Congress on Industrial and Applied Mathematics (ICIAM 2011), in Vancouver, British Columbia, Canada, July 18-22, 2011, and the SIAM International Conference on Linear Algebra, in Valencia, Spain, June 18-22, 2012

1.1 Numerically safe Gaussian elimination with no pivoting

Hereafter "flop" stands for "arithmetic operation", "expected" and "likely" mean "with probability 1 or close to 1", $\sigma_j(A)$ denotes the jth largest singular value of an $n \times n$ matrix A, and the ratio $\kappa(A) = \sigma_1(A)/\sigma_\rho(A)$ for $\rho = \text{rank}(A)$ denotes its condition number. If this number is large in context, then the matrix A is ill conditioned, that is lies near a rank deficient matrix, and then rounding errors tend to be magnified greatly in such computations as numerical solution of linear systems $A\mathbf{y} = \mathbf{b}$. Unless the number $\kappa(A)$ is large, the matrix A is well conditioned. Then it can be treated safely with the IEEE standard double or single precision provided that the matrix is nonsingular and the computations involve no ill conditioned or rank deficient auxiliary matrices.

To avoid dealing with the latter matrices in Gaussian elimination, one incorporates pivoting, that is row or column interchange. *Gaussian elimination with no pivoting* (hereafter we refer to it as *GENP*) can easily fail in numerical computations with rounding errors, except for some special input classes such as the classes of diagonally dominant and positive definite matrices. For such matrices, GENP outperforms Gaussian elimination with pivoting [GL96, page 119]. By extending our previous study in [PGMQ, Section 12.2] and [PQZa], we expand these classes dramatically by proving in Corollary 4.2 that pre- and post-multiplication of a well conditioned coefficient matrix of full rank by a square Gaussian random matrix is expected to yield safe numerical performance of GENP as well as block Gaussian elimination. The results of our tests support this theoretical finding consistently, even where we use just circulant or Householder multipliers generated by a vector or a pair of vectors, respectively, and fill these vectors with integers ± 1 for random signs $\pm$ (see our Table 8.6 and [PQZa, Table 2]).

1.2 Randomized preconditioning: the basic theorem

Given an ill conditioned matrix A, can we extend our advance by applying randomized multipliers X and Y to yield a much better conditioned matrix product XAY? No, because random square matrices X and Y are expected to be nonsingular and well conditioned [D88], [E88], [ES05], [CD05], [SST06], [B11] and because $\kappa(XAY) \geq \frac{\kappa(A)}{\kappa(X)\kappa(Y)}$. Approximate inverses are popular multipliers but only where it is not costly to compute them, that is only for some special, although important classes of matrices A. The inverses of ill conditioned matrices tend to be completely corrupted in numerical computations with double precision but can still serve as multiplicative preconditioners according to a series of experiments in [R90], although there seems to be no formal support yet for this surprising phenomenon.

We can readily produce a well conditioned matrix C by applying additive preprocessing $A \Longrightarrow C = A + P$, e.g., by choosing $P = I - A$, but it is not clear how this would help us to solve a linear system $A\mathbf{y} = \mathbf{b}$. (Here and hereafter I and I_n denote the $n \times n$ identity matrix.) Assume, however, that we are given a nonsingular ill conditioned $n \times n$ matrix A together with a small upper bound r on its numerical nullity, that is on the number of its singular values that are much smaller than the 2-norm $||A||_2$. Such matrices make up a large and important subclass of nonsingular ill conditioned matrices (cf. [CDG03] and Remarks 2.1 and 5.2), for which randomized additive preconditioning is supported by the following theorem. We prove it, extend it to rectangular matrices A, and further detail its probabilistic estimates in Section 5.

Theorem 1.1. *Suppose A is a real $n \times n$ matrix, $||A||_2 = 1$, U and V are standard Gaussian random $n \times r$ matrices for $0 < r < n$, and either $U = V$ or all $2nr$ entries of the matrices U and V are independent of each other. Then the matrix $C = A + UV^T$ is singular if $r < n - \text{rank}(A)$, but otherwise the matrix C is nonsingular with probability 1 and its condition number $\kappa(C)$ is expected to have at most the order $\sigma_1(A)/\sigma_{n-r}(A)$; consequently the matrix C is expected to be nonsingular and well conditioned if the matrix A has numerical nullity at most r.*

1.3 Randomized preconditioning: some applications

Suppose we have computed a small positive numerical nullity r of a normalized nonsingular $n \times n$ matrix A, with $||A|| = 1$ (cf. Section 7.4 on such computation), have generated standard Gaussian

random $n \times r$ matrices U and V, and have computed the matrix $C = A + UV^T$. Then by virtue of Theorem 1.1 this matrix is nonsingular with probability 1 and is expected to be well conditioned, having condition number of at most order $\sigma_1(A)/\sigma_{n-r}(A)$. Suppose it is indeed nonsingular and well conditioned, as expected. Then we can solve an ill conditioned linear system $A\mathbf{y} = \mathbf{b}$ of n equations more efficiently by reducing it to some well conditioned linear systems $C\mathbf{x} = \mathbf{f}$.

We can achieve this by combining *randomized additive preconditioning* $A \Longrightarrow C = A + UV^T$ with the Sherman–Morrison–Woodbury formula, hereafter referred to as the *SMW formula*, fast advanced algorithms of [ORO05] for accurate sums and products, and a variant of iterative refinement from [PGMQ, Section 9]. The overall computational cost decreases dramatically versus Gaussian elimination in the case of large n and an ill conditioned matrix A having a small numerical nullity (see Section 7.7). The heuristics of [R90] lead to the same progress, which we now support formally, although for the more narrow input class of matrices having small numerical nullities (cf. Remarks 5.1, 7.10, and 7.12).

Another group of applications of our randomized additive preconditioning employs the following observation: the ranges of the $n \times r$ matrices $C^{-1}U$ and $C^{-T}V$ are expected to approximate closely the left and right *trailing singular spaces* associated with the r smallest singular values of A. If this is the case indeed, as expected, then we can readily compute a 2×2 block triangulation of the matrix A where the two diagonal blocks have the sizes $(n-r) \times (n-r)$ and $r \times r$, respectively, and according to our formal study and tests, are expected to be substantially better conditioned than the matrix A (see Section 7.5). The estimated computational cost of this algorithm is about the same as for the approach employing the SMW formula.

Our *dual SMW formula* and *dual additive preprocessing* enable us to extend these randomized techniques to approximate the left and right *leading singular spaces* associated with the q largest singular values of A. The extended algorithms perform high precision computations involving only matrix multiplications and additions but no inversions or iterative refinement (cf. Remarks 7.8 and 7.9) and enable us to extend both groups of our randomized algorithms for linear systems $A\mathbf{y} = \mathbf{b}$, based on the SMW formula and 2×2 block triangulation, to the case where the coefficient matrix A has a small numerical rank (see Section 7.6 and Remark 7.9). In Sections 7.2 and 7.3 we approximate such matrices readily by low-rank matrices. This subject has been extensively studied because of numerous applications to fundamental computations with matrices and tensors (cf. [T00], [MMD08], [OT09], [HMT11]). Our techniques are quite effective and are distinct from the known methods.

If we agree to increase the size of an input matrix a little, we can alternatively apply randomized augmentation, defined, e.g., by the map $A \Longrightarrow K = \begin{pmatrix} I_r & V^T \\ -U & A \end{pmatrix}$. In the case of general $n \times n$ matrix A having norm 1 we choose standard Gaussian random $n \times r$ matrices U and V and deduce probabilistic upper bounds of order $||A||/\sigma_{n-r}(A)$ on the condition number $\kappa(K)$ (see Section 6). Our analysis enables distinct derivation of some variants of Theorem 1.1 and links closely augmentation to additive preprocessing (e.g. we extend augmentation to *dual augmentation* in (7.4) similarly to devising the dual SMW formula), although augmentation better preserves sparseness and structure of a matrix A (cf. Section 6.4).

1.4 Condition numbers of random structured matrices and numerical experiments

Estimating the condition numbers of random structured matrices is a well known challenge [SST06]. We respond to it by estimating the condition numbers of Gaussian random Toeplitz matrices. In Section 3.4 we prove that these numbers are not likely to grow to the infinity exponentially in n for $n \times n$ random Toeplitz matrices, even though the opposite has been proved in [BG05] for a large and important subclass of Toeplitz matrices. We obtain even stronger upper bounds on the condition numbers of random circulant matrices in Section 3.5. We cannot extend our proof of Theorem 1.1 to the case of Gaussian random Toeplitz matrices U and V, but such extension has been supported consistently by the results of our numerical tests. Moreover, empirically we observed that all our techniques remained as effective where instead of Gaussian random matrices we employed matrices defined by much fewer random parameters such as cirlulant matrices of Section 1.1. Other than that

our extensive tests were in good accordance with our theoretical estimates, e.g. with our estimates for the condition numbers of random Toeplitz and circulant matrices (see Tables 8.1–8.4).

Likewise in our tests randomized additive preprocessing decreased dramatically the condition number of an ill conditioned matrix having a small numerical nullity (see Table 8.5), whereas random circulant multipliers filled with ± 1 stabilized GENP numerically (see Table 8.6). Furthermore according to our test results, we accurately solved ill conditioned linear systems of equations when we combined our randomization techniques with 2×2 block triangulation (see Tables 8.7–8.10) or with the SMW formula (see Tables 8.15 and 8.16), our algorithms approximated accurately trailing and leading singular spaces of ill conditioned matrices that have small numerical nullities or ranks, and approximated a matrix by a low-rank matrix (see Tables 8.11–8.14). We have also matched the output accuracy of the customary algorithms for solving ill conditioned Toeplitz linear systems of equations but outperformed them in terms of the CPU time (see Table 8.17). We hope that our progress will motivate further work on application and extension of Theorem 1.1.

1.5 Organization of the paper and selective reading

We recall some definitions and basic results on matrix computations in the next section. We estimate the condition numbers of Gaussian random general, Toeplitz and circulant matrices in Section 3 and of randomized matrix products in Section 4. The latter estimate supports GENP with randomized multipliers. In Section 5 we prove that our randomized additive preprocessing of an ill conditioned matrix is expected to produce a well conditioned matrix. In Section 6 we study randomized augmentation, link it to randomized additive preprocessing, and apply to the solution of ill conditioned Toeplitz linear systems of equations.

In Section 7 we discuss a number of applications of randomized preprocessing to an ill conditioned matrix having a small numerical nullity or rank: we compute the numerical nullity and numerical rank of such a matrix, approximate its trailing and leading singular spaces, and extend this computation to the approximation of a matrix by low-rank matrices and by structured matrices. We also solve ill conditioned linear systems of equations by combining randomized additive preconditioning with the SMW formula or 2×2 block triangulation. We discuss structured preprocessing in Section 7.8. Corollary 5.1 and the results of Sections 3.4 and 6 motivate the design of the respective randomized algorithms and show their power, but otherwise Section 7 can be read independently of Sections 3–6.

In Section 8 we cover numerical tests, which constitute the contribution of the second and the third authors. In Section 9 we comment on the related works, our progress, and some directions for further study. In Appendix A we recall some results on structured matrices. In Appendix B we estimate the probability that under the uniform probability distribution a random matrix has full rank. In Appendix C we extend our probabilistic estimates to the case of complex matrices. Appendix D cover some results on randomized augmentation.

In our paper, unlike its introduction, we cover the general case of rectangular input matrices A.

2 Some definitions and basic results

We assume computations in the field $\mathbb{R}$ of real numbers, and comment on the extension to the field $\mathbb{C}$ of complex numbers in Appendix C.

For two scalars a and b we write $a \ll b$ and $b \gg a$ if the ratio $|b/a|$ is large. We write $a \approx b$ if $|a - b| \ll |a| + |b|$.

Hereafter "flop" stands for "arithmetic operation"; "expected" and "likely" mean "with probability 1 or close to 1", and the concepts "large", "small", "near", "closely approximate", "ill conditioned" and "well conditioned" are quantified in the context.

In the next subsections we recall and extend some customary definitions of matrix computations of [GL96], [S98].

2.1 Some basic definitions on matrix computations

$(B_1 \mid \ldots \mid B_k) = (B_j)_{j=1}^k$ is a $1 \times k$ block matrix with blocks $B_1, \ldots, B_k$.

$\mathrm{diag}(B_1, \ldots, B_k) = \mathrm{diag}(B_j)_{j=1}^k$ is a $k \times k$ block diagonal matrix with diagonal blocks $B_1, \ldots, B_k$.

$\mathbf{e}_i$ is the ith coordinate vector of dimension n for $i = 1, \ldots, n$. These vectors define the identity matrix $I_n = (\mathbf{e}_1 \mid \ldots \mid \mathbf{e}_n)$ and the reflection matrix $J_n = (\mathbf{e}_n \mid \ldots \mid \mathbf{e}_1)$, of size $n \times n$. $O_{k,l}$ is the $k \times l$ matrix filled with zeros. $\mathbf{0}_k$ is the vector $O_{k,1}$. We write I, J, O, and $\mathbf{0}$ where the size of a matrix or a vector is not important or is defined by context. Furthermore we write

$$I_{g,h} = I_g \text{ where } g \le h, \text{ whereas } I_{g,h} = (I_h \mid O_{h,g-h}) \text{ where } g > h. \tag{2.1}$$

A^T is the transpose of a matrix A. A^H is its Hermitian transpose. A matrix A is symmetric if $A = A^T$ and is symmetric positive definite if $A = B^T B$ for a nonsingular matrix B.

A real matrix Q is called *orthogonal* if $Q^T Q = I$ or $QQ^T = I$. More generally, over the complex field $\mathbb{C}$ a matrix U is called *unitary* if $U^H U = I$ or $UU^H = I$.

Fact 2.1. *[GL96, Theorem 5.2.2]. QR factorization $A = QR$ of a matrix A having full column rank into the product of an orthogonal matrix $Q = Q(A)$ and an upper triangular matrix $R = R(A)$ is unique provided that the factor R is a square matrix with positive diagonal entries.*

2.2 Range, null space, rank, nullity, nmbs, and generic rank profile

$\mathcal{R}(A)$ denotes the range of an $m \times n$ matrix A, that is the linear space $\{\mathbf{z} : \ \mathbf{z} = A\mathbf{x}\}$ generated by its columns. $\mathcal{N}(A)$ denotes its null space $\{\mathbf{v} : \ A\mathbf{v} = \mathbf{0}\}$, $\mathrm{rank}(A) = \dim \mathcal{R}(A)$ its rank, and $\mathrm{nul}(A) = \dim \mathcal{N}(A) = n - \mathrm{rank}(A)$ its nullity or right nullity, whereas $\mathrm{nul}\, A^T = m - \mathrm{rank}(A)$ is the left nullity of A, equal to $\mathrm{nul}\, A$ if and only if $m = n$. $\mathbf{v}$ is the null vector of A if $A\mathbf{v} = \mathbf{0}$.

Fact 2.2. *The set $\mathbb{M}$ of $m \times n$ matrices of rank ρ is an algebraic variety of dimension $(m+n-\rho)\rho$.*

Proof. Let M be an $m \times n$ matrix of a rank ρ with a nonsingular leading $\rho \times \rho$ block M_{00} and write $M = \begin{pmatrix} M_{00} & M_{01} \\ M_{10} & M_{11} \end{pmatrix}$. Then the $(m-\rho) \times (n-\rho)$ *Schur complement* $M_{11} - M_{10} M_{00}^{-1} M_{01}$ must vanish, which imposes $(m-\rho)(n-\rho)$ algebraic equations on the entries of M. Similar argument can be applied where any $\rho \times \rho$ submatrix of the matrix M (among $\binom{m}{\rho}\binom{n}{\rho}$ such submatrices) is nonsingular. Therefore $\dim \mathbb{M} = mn - (m-\rho)(n-\rho) = (m+n-\rho)\rho$. □

A matrix B having full column rank is said to be a *matrix basis* for its range $\mathcal{R}(B)$. Suppose a matrix B has full column rank and satisfies $\mathcal{R}(B) = \mathcal{N}(A)$ for a matrix A. Then we call B a *null matrix basis* or a *nmb* for this matrix A and write $B = \mathrm{nmb}(A)$. $\mathcal{N}(A^T)$ is the left null space of a matrix A, and similarly the map $A \to A^T$ defines left null vectors, left nmbs, and the left nullity of a matrix A; in particular an $m \times n$ matrix of a rank ρ has the left nullity $m - \rho$.

$A_k^{(k)}$ denotes the leading, that is northwestern $k \times k$ block submatrix of a matrix A.

A matrix of a rank ρ has *generic rank profile* if all its leading $i \times i$ blocks are nonsingular for $i = 1, \ldots, \rho$. If such a matrix is nonsingular itself, then it is called *strongly nonsingular*.

2.3 Norms, SVD, and low-rank approximation

$||A||_h$ is the h-norm of a matrix $A = (a_{i,j})_{i,j=1}^{m,n}$. We write $||A|| = ||A||_2$ and $||\mathbf{v}|| = \sqrt{\mathbf{v}^T \mathbf{v}} = ||\mathbf{v}||_2$ and recall from [GL96, Section 2.3.2 and Corollary 2.3.2] that

$$\max_{i,j=1}^{m,n} |a_{i,j}| \le ||A|| = ||A^T|| \le \sqrt{mn} \max_{i,j=1}^{m,n} |a_{i,j}|,$$

$$\frac{1}{\sqrt{m}} ||A||_1 \le ||A|| \le \sqrt{n} ||A||_1, \quad ||A||_1 = ||A^T||_\infty, \quad ||A||^2 \le ||A||_1 ||A||_\infty, \tag{2.2}$$

$$||AB||_h \le ||A||_h ||B||_h \text{ for } h = 1, 2, \infty \text{ and any matrix product } AB. \tag{2.3}$$

We write $A \approx B$ if $||A - B|| \ll ||A|| + ||B||$.

A matrix A is *normalized* if $||A|| = 1$. A normalized vector is orthogonal (unitary), and we call it *unit*.

Define an *SVD* or *full SVD* of an $m \times n$ matrix A of a rank ρ as follows,

$$A = S_A \Sigma_A T_A^T \tag{2.4}$$

where $S_A S_A^T = S_A^T S_A = I_m$, $T_A T_A^T = T_A^T T_A = I_n$, $\Sigma_A = \operatorname{diag}(\widehat{\Sigma}_A, O_{m-\rho,n-\rho})$, $\widehat{\Sigma}_A = \operatorname{diag}(\sigma_j(A))_{j=1}^{\rho}$, $\sigma_j = \sigma_j(A) = \sigma_j(A^T)$ is the jth largest singular value of a matrix A for $j = 1, \dots, \rho$, and we write $\sigma_j = 0$ for $j > \rho$. These values have the minimax property

$$\sigma_j = \max_{\dim(\mathbb{S})=j} \min_{\mathbf{x}\in\mathbb{S},\ ||\mathbf{x}||=1} ||A\mathbf{x}||,\ j = 1, \dots, \rho, \tag{2.5}$$

where $\mathbb{S}$ denotes linear spaces [GL96, Theorem 8.6.1]. Consequently $\sigma_\rho > 0$, $\sigma_1 = \max_{||\mathbf{x}||=1} ||A\mathbf{x}|| = ||A||$, and

$$||\operatorname{diag}(M_j)_j|| = \max_j ||M_j|| \text{ for any set of matrices } M_j. \tag{2.6}$$

Fact 2.3. *If A_0 is a submatrix of a matrix A, then $\sigma_j(A) \ge \sigma_j(A_0)$ for all j.*

Proof. [GL96, Corollary 8.6.3] implies the claimed bound where A_0 is any block of columns of the matrix A. Transposition of a matrix and permutations of its rows and columns do not change singular values, and thus we can extend the bounds to all submatrices A_0. □

2.4 Inverses, generalized inverses, and perturbation bounds

$A^+ = T_A \operatorname{diag}(\widehat{\Sigma}_A^{-1}, O_{n-\rho,m-\rho}) S_A^T$ is the Moore–Penrose pseudo-inverse of the matrix A of (2.4), and

$$||A^+|| = 1/\sigma_\rho(A) \tag{2.7}$$

for a matrix A of a rank ρ. A^{+T} stands for $(A^+)^T = (A^T)^+$, and A^{-T} stands for $(A^{-1})^T = (A^T)^{-1}$.

An $n \times m$ matrix $X = A^{(I)}$ is a left inverse of an $m \times n$ matrix A if $XA = I$ and is its right inverse if $AX = I$. A^+ is a left or right inverse $A^{(I)}$ if and only if a matrix A has full rank. $A^{(I)}$ is unique and is equal to A^{-1} if A is a nonsingular matrix.

Theorem 2.1. *Suppose C and $C + E$ are two nonsingular matrix of the same size and $||C^{-1}E|| = \theta < 1$. Then $||I - (C + E)^{-1}C|| \le \frac{\theta}{1-\theta}$ and $|||(C + E)^{-1} - C^{-1}|| \le \frac{\theta}{1-\theta}||C^{-1}||$.*

Proof. See [S98, Corollary 1.4.19] for $P = -C^{-1}E$. □

Theorem 2.2. *[GL96, Section 5.5.5]. Assume two matrices $C \in \mathbb{C}^{m\times n}$ and $\tilde{C} \in \mathbb{C}^{m\times n}$ having full rank and write $E = \tilde{C} - C$. Then $||\tilde{C}^+ - C^+|| \le 2\sqrt{n}||E|| \max\{||\tilde{C}^+||^2, ||C^+||^2\}$.*

2.5 SMW and dual SMW formulae

Theorem 2.3. *[GL96, page 50], [S98, Corollary 4.3.2]. Suppose that $U, V \in \mathbb{R}^{n\times r}$, the matrices $A \in \mathbb{R}^{n\times n}$ and $C = A + UV^T$ are nonsingular, and $0 < r < n$. Then the matrix $G = I_r - V^T C^{-1} U$ is nonsingular and we have the Sherman–Morrison–Woodbury (hereafter* SMW*) formula*

$$A^{-1} = C^{-1} + C^{-1} U G^{-1} V^T C^{-1}.$$

Corollary 2.1. *Suppose that $U_-, V_- \in \mathbb{R}^{n\times q}$, $A \in \mathbb{R}^{n\times n}$, A and $A^{-1} + U_- V_-^T$ are nonsingular matrices, and $0 < q < n$. Write*

$$C_-^{-1} = A^{-1} + U_- V_-^T,\ H = I_q + V_-^T A U_-, \tag{2.8}$$

Then the matrix H is nonsingular and the following dual SMW formula *holds,*

$$C_- = A - A U_- H^{-1} V_-^T A. \tag{2.9}$$

Proof. Apply Theorem 2.3 to the matrices A^{-1}, U_-, V_- and C_-^{-1} replacing the matrices A, U, V and C, respectively. □

2.6 Condition number, numerical rank and nullity, generic conditioning profile

$\kappa(A) = \frac{\sigma_1(A)}{\sigma_\rho(A)} = ||A||\,||A^+||$ is the condition number of an $m \times n$ matrix A of a rank ρ. Such a matrix is *ill conditioned* if $\sigma_1(A) \gg \sigma_\rho(A)$ and is *well conditioned* otherwise. See [D83], [GL96, Sections 2.3.2, 2.3.3, 3.5.4, 12.5], [H02, Chapter 15], [KL94], and [S98, Section 5.3] on the estimation of norms and condition numbers of nonsingular matrices.

An $m \times n$ matrix A has *numerical rank* q, not exceeding $\text{rank}(A)$, and has the right numerical nullity $r = n - q$ (to which we will refer just as *numerical nullity*) if the ratios $\sigma_j(A)/||A||$ are small for $j > q$ but not for $j \le q$. The *left numerical nullity* of the matrix A equals the numerical nullity $m - q$ of the $n \times m$ transpose A^T and coincides with the numerical nullity of A if and only if $m = n$.

Remark 2.1. *One can specify the adjective "small" above as "smaller than a fixed positive tolerance". The choice of the tolerance can be a challenge, e.g., for the matrix* $\text{diag}(1.1^{-j})_{j=0}^{999}$.

If a well conditioned $m \times n$ matrix A has a rank $\rho < l = \min\{m, n\}$, then almost all its close neighbours have full rank l, and all of them have numerical rank $q = \rho$. Conversely, suppose a matrix has a positive numerical rank q, set to 0 its all but q largest singular values and denote the resulting matrix by A and the original matrix by $A + E$. Then $\text{rank}(A) = q$, A is a well conditioned matrix, $||E|| = \sigma_{q+1}(A)$, and so A is a rank-q approximation to the matrix $A + E$. We can obtain this approximation from the SVD of $A + E$, but other rank-q approximations can be obtained at a lower cost from rank-revealing QRP or PRQ factorizations of the matrix $A + E$, where P is a proper permutation matrix [GE96]. One can compute such approximations by applying randomized algorithms, which are even less costly (see Section 7.3).

An $m \times n$ matrix has *generic conditioning profile* (cf. the end of Section 2.2) if it has a numerical rank q and if its leading $i \times i$ blocks are nonsingular and well conditioned for $i = 1, \dots, q$. If such a matrix has full rank (that is if $q = \min\{m, n\}$) and if it is well conditioned itself, then we call it *strongly well conditioned*.

Theorem 2.4. *Cf. [PQZa]. Suppose Gaussian elimination with no pivoting or block Gaussian elimination has been applied to a matrix A having a rank ρ (resp. numerical rank q) to compute LU factorizations of the leading block submatrices $A_j^{(j)}$ for $j = 1, \dots, \rho$ (resp. for $j = 1, \dots, q$). Then the computations involve no divisions by 0 (resp. by the values that are absolutely small relatively to the norm $||A||$) if and only if the matrix A has generic rank (resp. generic conditioning) profile.*

The claim about involving no divisions by absolutely small values is a simple extension of the well known fact about involving no divisions by 0.

2.7 Toeplitz, Hankel and f-circulant matrices

A *Toeplitz* $m \times n$ matrix $T_{m,n} = (t_{i-j})_{i,j=1}^{m,n}$ is defined by its first row and column, that is by the vector $(t_h)_{h=1-n}^{m-1}$ of dimension $m + n - 1$. We write $T_n = T_{n,n} = (t_{i-j})_{i,j=1}^{n,n}$ (see the display below).

A lower *triangular Toeplitz* $n \times n$ matrix $Z(\mathbf{t}) = (t_{i-j})_{i,j=1}^{n}$ (where $t_k = 0$ for $k < 0$) is defined by its first column $\mathbf{t} = (t_h)_{h=0}^{n-1}$. We write $Z(\mathbf{t})^T = (Z(\mathbf{t}))^T$. $Z = Z_0 = Z(\mathbf{e}_2)$ is the downshift $n \times n$ matrix displayed below. We have $Z\mathbf{v} = (v_i)_{i=0}^{n-1}$ and $Z(\mathbf{v}) = Z_0(\mathbf{v}) = \sum_{i=1}^{n} v_i Z^{i-1}$ for $\mathbf{v} = (v_i)_{i=1}^{n}$ and $v_0 = 0$,

$$T_n = \begin{pmatrix} t_0 & t_{-1} & \cdots & t_{1-n} \\ t_1 & t_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_{-1} \\ t_{n-1} & \cdots & t_1 & t_0 \end{pmatrix}, \quad Z = \begin{pmatrix} 0 & & \dots & & 0 \\ 1 & \ddots & & & \\ \vdots & \ddots & \ddots & & \vdots \\ & & \ddots & 0 & \\ 0 & & \dots & 1 & 0 \end{pmatrix}.$$

Combine the equations $||Z(\mathbf{v})||_1 = ||Z(\mathbf{v})||_\infty = ||\mathbf{v}||_1$ with (2.2) to obtain

$$||Z(\mathbf{v})|| \le ||\mathbf{v}||_1. \tag{2.10}$$

Theorem 2.5. *Write $T_k = (t_{i-j})_{i,j=0}^{k-1}$ for $k = n, n+1$.*

(a) Let the matrix T_n be nonsingular and write $\mathbf{p} = T_n^{-1}\mathbf{e}_1$ and $\mathbf{q} = T_n^{-1}\mathbf{e}_n$. If $p_1 = \mathbf{e}_1^T\mathbf{p} \neq 0$, then $p_1 T_n^{-1} = Z(\mathbf{p})Z(J\mathbf{q})^T - Z(Z\mathbf{q})Z(ZJ\mathbf{p})^T$.

In parts (b) and (c) below let the matrix T_{n+1} be nonsingular and write $\widehat{\mathbf{v}} = (v_i)_{i=0}^n = T_{n+1}^{-1}\mathbf{e}_1$, $\mathbf{v} = (v_i)_{i=0}^{n-1}$, $\mathbf{v}' = (v_i)_{i=1}^n$, $\widehat{\mathbf{w}} = (w_i)_{i=0}^n = T_{n+1}^{-1}\mathbf{e}_{n+1}$, $\mathbf{w} = (w_i)_{i=0}^{n-1}$, and $\mathbf{w}' = (w_i)_{i=1}^n$.

(b) If $v_0 \neq 0$, then the matrix T_n is nonsingular and $v_0 T_n^{-1} = Z(\mathbf{v})Z(J\mathbf{w}')^T - Z(\mathbf{w})Z(J\mathbf{v}')^T$.

(c) If $v_n \neq 0$, then the matrix $T_{1,0} = (t_{i-j})_{i=1,j=0}^{n,n-1}$ is nonsingular and $v_n T_{1,0}^{-1} = Z(\mathbf{w})Z(J\mathbf{v}')^T - Z(\mathbf{v})Z(J\mathbf{w}')^T$.

Proof. See [GS72] on parts (a) and (b); see [GK72] on part (c). □

$Z_f = Z + f\mathbf{e}_1^T\mathbf{e}_n$ for a scalar $f \neq 0$ denotes the $n \times n$ matrix of *f-circular shift*. An *f-circulant matrix* $Z_f(\mathbf{v}) = \sum_{i=1}^n v_i Z_f^{i-1}$ is a special Toeplitz $n \times n$ matrix defined by its first column vector $\mathbf{v} = (v_i)_{i=1}^n$ and a scalar f. f-circulant matrix is called *circulant* if $f = 1$ and *skew circulant* if $f = -1$. By replacing f with 0 we arrive at a lower triangular Toeplitz matrix $Z(\mathbf{v})$. The following theorem implies that the inverses (wherever they are defined) and pairwise products of f-circulant $n \times n$ matrices are f-circulant and can be computed in $O(n \log n)$ flops.

Theorem 2.6. *(See [CPW74].) We have $Z_1(\mathbf{v}) = \Omega^{-1}D(\Omega\mathbf{v})\Omega$. More generally, for any $f \neq 0$, we have $Z_{f^n}(\mathbf{v}) = U_f^{-1}D(U_f\mathbf{v})U_f$ where $U_f = \Omega D(\mathbf{f})$, $\mathbf{f} = (f^i)_{i=0}^{n-1}$, $D(\mathbf{u}) = \mathrm{diag}(u_i)_{i=0}^{n-1}$ for a vector $\mathbf{u} = (u_i)_{i=0}^{n-1}$, $\Omega = (\omega_n^{ij})_{i,j=0}^{n-1}$ is the $n \times n$ matrix of the discrete Fourier transform at n points, $\omega_n = \exp(\frac{2\pi}{n}\sqrt{-1})$ being a primitive n-th root of 1, and $\Omega^{-1} = \frac{1}{n}(\omega_n^{-ij})_{i,j=0}^{n-1} = \frac{1}{n}\Omega^H$.*

Hankel $m \times n$ matrices $H = (h_{i+j})_{i,j=1}^{m,n}$ can be defined equivalently as the products $H = TJ_n$ or $H = J_m T$ of $m \times n$ Toeplitz matrices T and the Hankel reflection matrices $J = J_m$ or J_n. Note that $J = J^{-1} = J^T$ and obtain the following simple fact.

Fact 2.4. *For $m = n$ we have $T = HJ$, $H^{-1} = JT^{-1}$ and $T^{-1} = JH^{-1}$ if $H = TJ$, whereas $T = JH$, $H^{-1} = JT^{-1}$ and $T^{-1} = H^{-1}J$ if $H = JT$. Furthermore in both cases $\kappa(H) = \kappa(T)$.*

By using the equations above we can readily extend any Toeplitz matrix inversion algorithm to Hankel matrix inversion and vice versa, preserving the flop count and condition numbers. E.g. $(JT)^{-1} = T^{-1}J$, $(TJ)^{-1} = JT^{-1}$, $(JH)^{-1} = H^{-1}J$ and $(HJ)^{-1} = JH^{-1}$.

2.8 Toeplitz-like, Hankel-like and some other structured matrices

Let us extend the class of Toeplitz and Hankel matrices to a more general class of structured matrices. We only employ them in Sections 7.3, 7.8 and 7.9. With every pair of $n \times n$ *operator matrices* A and B associate the class of $n \times n$ matrices M for which the rank of the Sylvester *displacement* $AM - MB$ (called the *displacement rank* of M is small in context.

The matrices T with the structure of Toeplitz type (we call them *Toeplitz-like* matrices) have small *displacement ranks* $d = d(A, B)$ for $A = Z_e$ and $B = Z_f$ and for any pair of scalars e and f. Such matrices extend the class of Toeplitz matrices, for which $d \leq 2$. Any variation of a pair of scalars e and f can change the displacement rank of a matrix by at most 2, and so the class of Toeplitz-like matrices is independent of the choice of such a pair.

Every matrix of a rank d, and in particular a displacement of a rank d, can be nonuniquely represented as the sum of d outer products $\mathbf{g}_j\mathbf{h}_j^T$ of d pairs of vectors $\mathbf{g}_j$ and $\mathbf{h}_j$ for $j = 1, \ldots, d$. Motivated by the following result we call the pair of matrices $G = G_{Z_e,Z_f}(M) = (\mathbf{g}_j)_{j=1}^d$ and $H = H_{Z_e,Z_f}(M) = (\mathbf{h}_j)_{j=1}^d$, made up of the vectors $\mathbf{g}_j$ and $\mathbf{h}_j$, a *displacement generator* of length d for the matrix M and for the operator matrices Z_e and Z_f where $e \neq f$ (cf. [P01, Example 4.4.2]).

Theorem 2.7. *If $Z_e M - MZ_f = \sum_{j=1}^d \mathbf{g}_j\mathbf{h}_j^T$ for a pair of distinct scalars e and f, then*

$$(e - f)M = \sum_{j=1}^d Z_e(\mathbf{g}_j)Z_f(J\mathbf{h}_j). \tag{2.11}$$

The theorem expresses the matrix through the generator $\{G, H\}$ by using $2dn$ parameters instead of n^2 entries. For $d \ll n$, this is a dramatic compression, which furthermore reduces multiplication of the matrix M by a vector essentially to $2d$ multiplications of circulant matrices by vectors, that is to $O(dn \log n)$ flops. Moreover we can operate with matrices by using their displacement representation, which preserves Toeplitz-like structure and can accelerate the computations dramatically where $d \ll n$. Namely, for Toeplitz-like matrices T (nonsingular), T_1 and T_2, scalars e, f, α, and β, and operator matrices $A = Z_e$, $B = Z_f$, and $C = Z_c$, we can readily compute Toeplitz-like matrices $\alpha T_1 + \beta T_2$, T^T, $T_1 T_2$ and T^{-1}. The following theorem bounds the growth of the length of the associated displacement generators and the respective flop cost.

Theorem 2.8. *Assume that $n \times n$ matrices T_1, T_2, and T have been represented with their displacement generators of lengths d_1, d_2, and d, respectively, for appropriate operator matrices $A = Z_e$ and $B = Z_f$, defining Toeplitz-like structure. Then there exist displacement generators of length d for T^{-1} (provided the matrix T is nonsingular) and T^T and of length at most $d_1 + d_2 + O(1)$ for $T_1 T_2$ and $\alpha T_1 + \beta T_2$ (for appropriate operator matrices defining Toeplitz-like structure and for any pair of scalar α and β). One can compute these generators by using $O(d^2 n \log^2 n)$ and $O(d_1 d_2 n \log n)$ flops, respectively.*

Proof. The theorem readily follows from Theorem A.1 and Corollary A.1 in Appendix A, which also define all the respective displacement generators. □

A matrix H is *Hankel-like* if its displacement rank $\mathrm{rank}(AH - HB)$ is small where $A = Z_e$ and $B = Z_f^T$ or alternatively where $A = Z_e^T$ and $B = Z_f$ for two scalars e and f. It follows that MN is a Hankel-like matrix if one of the factors is a Toeplitz-like matrix and another is a Hankel-like matrix, whereas MN is a Toeplitz-like matrix if both M and N are Hankel-like matrices or both are Toeplitz-like matrices. We can alternatively define Hankel-like matrices as the products TJ or JT where T is a Toeplitz-like matrix, or we can define Toeplitz-like matrices T as the products HJ and JH where H are Hankel-like matrices (cf. Fact 2.4). By using these properties we can readily extend our algorithms as well as expressions (2.11) (cf. [P01, Example 4.4.4]) from the case of Toeplitz and Toeplitz-like to Hankel and Hankel-like matrices, preserving the flop count.

Remark 2.2. *By choosing operator matrices A and B among f-circulant and appropriate diagonal matrices we define the important classes of matrices with the structures of Vandermonde and Cauchy types whose displacement rank* $\mathrm{rank}(AM - MB)$ *is small. This extends the classes of Vandermonde matrices $V_{\mathbf{x}} = (x_i^{j-1})_{i,j=1}^n$, having displacement rank 1 for the operator matrices $A = \mathrm{diag}(x_i)_{i=1}^n$ and $B = Z_f^T$ for a scalar f, and Cauchy matrices $C_{\mathbf{s},\mathbf{t}} = (\frac{1}{s_i - t_j})_{i,j=1}^n$, having displacement rank 1 for the operator matrices $A = \mathrm{diag}(s_i)_{i=1}^n$ and $B = \mathrm{diag}(t_j)_{j=1}^n$. Alternatively [P90], [P01], the matrices of these classes can be defined as the products UMV where M is a Toeplitz matrix, whereas U and V are properly selected among Vandermonde matrices, their transposes and the identity matrices. Similarly to the Toeplitz–Hankel link at the end of the previous subsection, this enables us to extend any successful algorithm for Cauchy-like inversion to Toeplitz-like, Hankel-like and Vandermonde-like inversion and vice versa because $(UMV)^{-1} = V^{-1} M^{-1} U^{-1}$ [P90]. Theorems 2.7 and 2.8 and other basic properties of Toeplitz-like and Hankel-like matrices can be extended to the matrices having structures of Vandermonde or Cauchy types (see [P00], [P01] or Appendix A), except that, unlike the orthogonal reversion matrix J, Vandermonde multipliers and their transposes are usually ill conditioned except for a narrow class such as the matrices Ω and Ω^H of Theorem 2.6.*

3 Ranks and conditioning of Gaussian random matrices

3.1 Random variables and Gaussian random matrices

Definition 3.1. *$F_\gamma(y) = Probability\{\gamma \le y\}$ for a real random variable γ is the* cumulative distribution function (cdf) *of γ evaluated at y. $F_{g(\mu,\sigma)}(y) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{y} \exp(-\frac{(x-\mu)^2}{2\sigma^2}) dx$ for a Gaussian random variable $g(\mu, \sigma)$ with a mean μ and a positive variance σ^2, and so*

$$\mu - 4\sigma \le y \le \mu + 4\sigma \text{ with a probability near 1.} \tag{3.1}$$

Definition 3.2. *A matrix or a vector is a* Gaussian random matrix or vector *with a mean μ and a positive variance σ^2 if it is filled with independent identically distributed Gaussian random variables, all having the mean μ and variance σ^2. $\mathcal{G}_{\mu,\sigma}^{m\times n}$ is the set of such Gaussian random $m \times n$ matrices (which are* standard *for $\mu = 0$ and $\sigma^2 = 1$). By restricting this set to Toeplitz or f-circulant matrices we obtain the sets $\mathcal{T}_{\mu,\sigma}^{m\times n}$ and $\mathcal{Z}_{f,\mu,\sigma}^{n\times n}$ of* Gaussian random Toeplitz *and* Gaussian random f-circulant matrices*, respectively.*

Definition 3.3. *$\chi_{\mu,\sigma,n}(y)$ is the cdf of the norm $||\mathbf{v}|| = (\sum_{i=1}^{n} v_i^2)^{1/2}$ of a Gaussian random vector $\mathbf{v} = (v_i)_{i=1}^{n} \in \mathcal{G}_{\mu,\sigma}^{n\times 1}$. For $y \geq 0$ we have $\chi_{0,1,n}(y) = \frac{2}{2^{n/2}\Gamma(n/2)} \int_0^y x^{n-1} \exp(-x^2/2)dx$ where $\Gamma(h) = \int_0^\infty x^{h-1} \exp(-x)dx$, $\Gamma(n+1) = n!$ for nonnegative integers n.*

3.2 Nondegeneration of Gaussian random matrices

The total degree of a multivariate monomial is the sum of its degrees in all its variables. The total degree of a polynomial is the maximal total degree of its monomials.

Lemma 3.1. *[DL78], [S80], [Z79]. For a set Δ of a cardinality $|\Delta|$ in any fixed ring let a polynomial in m variables have a total degree d and let it not vanish identically on this set. Then the polynomial vanishes in at most $d|\Delta|^{m-1}$ points.*

We assume that Gaussian random variables range over infinite sets Δ, usually being the real line or its intervals. Then the lemma implies that a nonzero polynomial vanishes with probability 0. Thus Gaussian random general, Toeplitz and circulant matrices have full rank with probability 1 because the determinants of its square submatrices are polynomials in the entries. A square Gaussian random general, Toeplitz or circulant matrix A is nonsingular with probability 1. Furthermore all entries of the matrix, its adjoint, denoted $\operatorname{adj} A$, are subdeterminants and thus are nonzeros with probability 1. Clearly this property of the adjoint can be extended to the inverse $A^{-1} = \frac{\operatorname{adj} A}{\det A}$. Hereafter, wherever this causes no confusion, we assume by default that *Gaussian random general, Toeplitz and circulant matrices have full rank, and their inverses (if defined) have nonzero entries.*

Similar properties with probability near 1 hold where the random variables are sampled under the uniform probability distribution from a finite set of a large cardinality (see Appendix A).

3.3 Conditioning of Gaussian random matrices

Besides having full rank with probability 1, Gaussian random matrices in Definition 3.2 are expected to be well conditioned [D88], [E88], [ES05], [CD05], [B11], and even the sum $M + A$ for $M \in \mathbb{R}^{m\times n}$ and $A \in \mathcal{G}_{\mu,\sigma}^{m\times n}$ is expected to be well conditioned unless the ratio $\sigma/||M||$ is small or large [SST06].

The following theorem states an upper bound proportional to y on the probability (the cdf) $F_{1/||A^+||}(y)$ that the smallest positive singular value $1/||A^+||$ of a Gaussian random matrix A is less than a nonnegative scalar y (cf. (2.7)). The bound holds for any fixed matrix B, but we mostly use it where $B = O_{m,n}$. In this case the bound can be strengthened by a factor $y^{|m-n|}$ [ES05], [CD05], but we do not use this improvement.

Theorem 3.1. *Suppose $A \in \mathcal{G}_{\mu,\sigma}^{m\times n}$, $B \in \mathbb{R}^{m\times n}$, $l = \min\{m, n\}$, and $y \geq 0$. Then $F_{1/||A^+-B||}(y) \leq 2.35\sqrt{l}y/\sigma$.*

Proof. For $m = n$ this is [SST06, Theorem 3.3]. Apply Fact 2.3 to extend it to any pair $\{m, n\}$. □

The following two theorems supply lower bounds $F_{||A||}(z)$ and $F_{\kappa(A)}(y)$ on the probabilities that $||A|| \leq z$ and $\kappa(A) \leq y$ for two scalars y and z, respectively, and a Gaussian random matrix A. We do not use the second theorem, but state it for the sake of completeness and only for square $n \times n$ matrices A. The theorems imply that the functions $1 - F_{||A||}(z)$ and $1 - F_{\kappa(A)}(y)$ decay as $z \to \infty$ and $y \to \infty$, respectively, and that the decays are exponential in $-z^2$ and proportional to $\sqrt{\log y}/y$, respectively. For small values $y\sigma$ and a fixed n the lower bound of Theorem 3.3 becomes negative, in which case the theorem becomes trivial.

We assume that $\mu = 0$ when we estimate the cdfs of $||A||$ and consequently of $\kappa(A)$ for Gaussian random matrices A, but we make no such assumption when we estimate the cdf of $1/||A^+||$.

Theorem 3.2. *[DS01, Theorem II.7]. Suppose $A \in \mathcal{G}_{0,\sigma}^{m\times n}$, $h = \max\{m,n\}$ and $z \geq 2\sigma\sqrt{l}$. Then $F_{||A||}(z) \geq 1 - \exp(-(z-2\sigma\sqrt{h})^2/(2\sigma^2))$.*

Theorem 3.3. *[SST06, Theorem 3.1]. Suppose $0 < \sigma \leq 1$, $y \geq 1$, $A \in \mathcal{G}_{0,\sigma}^{n\times n}$. Then the matrix A has full rank with probability 1 and $F_{\kappa(A)}(y) \geq 1 - (14.1 + 4.7\sqrt{(2\ln y)/n})n/(y\sigma)$.*

Proof. See [SST06, the proof of Lemma 3.2]. □

3.4 Conditioning of Gaussian random Toeplitz matrices

A matrix $T_n = (t_{i-j})_{i,j=1}^n$ is the sum of two triangular Toeplitz matrices

$$T_n = Z(\mathbf{t}) + Z(\mathbf{t}_-)^T,\ \mathbf{t} = (t_i)_{i=0}^{n-1},\ \mathbf{t}_- = (t'_{-i})_{i=0}^{n-1},\ t'_0 = 0. \tag{3.2}$$

If $T_n \in \mathcal{T}_{\mu,\sigma}^{n\times n}$, then T_n has $2n-1$ pairwise independent entries in $\mathcal{G}_{\mu,\sigma}$. Thus (2.10) implies that

$$\frac{1}{\sqrt{n}}||T_n|| \leq ||T_n||_1 \leq ||Z(\mathbf{t})||_1 + ||Z(\mathbf{t}_-)^T||_1 = ||\mathbf{t}||_1 + ||\mathbf{t}_-||_1 = ||(t_i)_{i=1-n}^{n-1}||_1 \leq \sqrt{n}||(t_i)_{i=1-n}^{n-1}||,$$

and so

$$F_{||T_n||}(y) \geq \chi_{\mu,\sigma,2n-1}(y/n). \tag{3.3}$$

Next we estimate the norm $||T_n^{-1}||$ for $T_n \in \mathcal{T}_{\mu,\sigma}^{n\times n}$, beginning with a basic lemma.

Lemma 3.2. *[SST06, Lemma A.2]. For a nonnegative scalar y, a unit vector $\mathbf{t} \in \mathbb{R}^{n\times 1}$, and a vector $\mathbf{b} \in \mathcal{G}_{\mu,\sigma}^{n\times 1}$, we have $F_{|\mathbf{t}^T\mathbf{b}|}(y) \leq y \leq \sqrt{\frac{2}{\pi}}\frac{y}{\sigma}$.*

Remark 3.1. *The latter bound is independent of μ and n; it holds for any μ even if all coordinates of the vector $\mathbf{b}$ are fixed except for a single coordinate in $\mathcal{G}_{\mu,\sigma}$.*

Theorem 3.4. *Given a matrix $T_n = (t_{i-j})_{i,j=1}^n \in \mathcal{T}_{\mu,\sigma}^{n\times n}$, assumed to be nonsingular (cf. Section 3.2), write $p_1 = \mathbf{e}_1^T T_n^{-1}\mathbf{e}_1$. Then $F_{1/||p_1 T_n^{-1}||}(y) \leq 2n\alpha\beta$ for two random variables α and β such that*

$$F_\alpha(y) \leq \sqrt{\frac{2n}{\pi}}\frac{y}{\sigma} \text{ and } F_\beta(y) \leq \sqrt{\frac{2n}{\pi}}\frac{y}{\sigma} \text{ for } y \geq 0. \tag{3.4}$$

Proof. By virtue of part (a) of Theorem 2.5 we have $p_1 T_n^{-1} = Z(\mathbf{p})Z(J\mathbf{q})^T - Z(Z\mathbf{q})Z(ZJ\mathbf{p})^T$, and so $||p_1 T_n^{-1}|| \leq ||Z(\mathbf{p})||\ ||Z(J\mathbf{q})^T|| + ||Z(Z\mathbf{q})||\ ||Z(ZJ\mathbf{p})^T||$ for $\mathbf{p} = T_n^{-1}\mathbf{e}_1$, $\mathbf{q} = T_n^{-1}\mathbf{e}_n$, and $p_1 = \mathbf{p}^T\mathbf{e}_1$. It follows that $||p_1 T_n^{-1}|| \leq ||Z(\mathbf{p})||\ ||Z(J\mathbf{q})|| + ||Z(Z\mathbf{q})||\ ||Z(ZJ\mathbf{p})||$ since $||A|| = ||A^T||$ for all matrices A. Furthermore $||p_1 T_n^{-1}|| \leq ||\mathbf{p}||_1\ ||J\mathbf{q}||_1 + ||Z\mathbf{q}||_1\ ||ZJ\mathbf{p}||_1$ due to (2.10). Clearly $||J\mathbf{v}||_1 = ||\mathbf{v}||_1$ and $||Z\mathbf{v}||_1 \leq ||\mathbf{v}||_1$ for every vector $\mathbf{v}$, and so (cf. (2.2))

$$||p_1 T_n^{-1}|| \leq 2||\mathbf{p}||_1\ ||\mathbf{q}||_1 \leq 2n||\mathbf{p}||\ ||\mathbf{q}||. \tag{3.5}$$

By definition the vector $\mathbf{p}$ is orthogonal to the vectors $T_n\mathbf{e}_2, \dots, T_n\mathbf{e}_n$, whereas $\mathbf{p}^T T_n\mathbf{e}_1 = 1$ (cf. [SST06]). Consequenty the vectors $T_n\mathbf{e}_2, \dots, T_n\mathbf{e}_n$ uniquely define the vector $\mathbf{u} = \mathbf{p}/||\mathbf{p}||$, whereas $|\mathbf{u}^T T_n\mathbf{e}_1| = 1/||\mathbf{p}||$.

The last coordinate t_{n-1} of the vector $T_n\mathbf{e}_1$ is independent of the vectors $T_n\mathbf{e}_2, \dots, T_n\mathbf{e}_n$ and consequently of the vector $\mathbf{u}$. Apply Remark 3.1 to estimate the cdf of the random variable $\alpha = 1/||\mathbf{p}|| = |\mathbf{u}^T T_n\mathbf{e}_1|$ and obtain that $F_\alpha(y) \leq \sqrt{\frac{2n}{\pi}}\frac{y}{\sigma}$ for $y \geq 0$.

Likewise the $n-1$ column vectors $T\mathbf{e}_1, \dots, T_{n-1}$ define the vector $\mathbf{v} = \beta\mathbf{q}$ for $\beta = 1/||\mathbf{q}|| = |\mathbf{v}^T T_n\mathbf{e}_n|$. The first coordinate t_{1-n} of the vector $T_n\mathbf{e}_n$ is independent of the vectors $T\mathbf{e}_1, \dots, T_{n-1}$ and consequently of the vector $\mathbf{v}$. Apply Remark 3.1 to estimate the cdf of the random variable β and obtain that $F_\beta(y) \leq \sqrt{\frac{2n}{\pi}}\frac{y}{\sigma}$ for $y \geq 0$.

Finally combine these bounds on the cdfs $F_\alpha(y)$ and $F_\beta(y)$ with (3.5). □

By employing parts (b) and (c) of Theorem 2.5 instead of its part (a), we similarly deduce the bounds $||v_0 T_{n+1}^{-1}|| \le 2\alpha\beta$ and $||v_n T_{n+1}^{-1}|| \le 2\alpha\beta$ for two pairs of random variables α and β satisfying (3.4) for $n+1$ replacing n. We have $p_1 = \frac{\det T_{n-1}}{\det T_n}$, $v_0 = \frac{\det T_n}{\det T_{n+1}}$, and $v_n = \frac{\det T_{0,1}}{\det T_{n+1}}$ for $T_{0,1} = (t_{i-j})_{i=0,j=1}^{n-1,n}$. Next we bound the geometric means of the ratios $|\frac{\det T_{h+1}}{\det T_h}|$ for $h = 1, \dots, k-1$. $1/|p_1|$ and $1/|v_0|$ are such ratios for $k = n-1$ and $k = n$, respectively, whereas the ratio $1/|v_n|$ is similar to $1/|v_0|$, under slightly distinct notation.

Theorem 3.5. *Let $T_h \ne O$ denote $h \times h$ matrices for $h = 1, \dots, k$ whose entries have absolute values at most t for a fixed scalar or random variable t, e.g. for $t = ||T||$. Furthermore let $T_1 = (t)$. Then the geometric mean $(\prod_{h=1}^{k-1} |\frac{\det T_{h+1}}{\det T_h}|)^{1/(k-1)} = \frac{1}{t} |\det T_k|^{1/(k-1)}$ is at most $k^{\frac{1}{2}(1+\frac{1}{k-1})} t$.*

Proof. The theorem follows from Hadamard's upper bound $|\det M| \le k^{k/2} t^k$, which holds for any $k \times k$ matrix $M = (m_{i,j})_{i,j=1}^k$ with $\max_{i,j=1}^k |m_{i,j}| \le t$. □

The theorem says that the geometric mean of the ratios $|\det T_{h+1}/\det T_h|$ for $h = 1, \dots, k-1$ is not greater than $k^{0.5+\epsilon(k)} t$ where $\epsilon(k) \to 0$ as $k \to \infty$. Furthermore if $T_n \in \mathcal{G}_{\mu,\sigma}^{n \times n}$ we can write $t = ||T||$ and apply (3.3) to bound the cdf of t.

3.5 Conditioning of Gaussian random circulant matrices

Next we estimate the norms of a random Gaussian f-circulant matrix and its inverse.

Theorem 3.6. *Assume $y \ge 0$ and a circulant $n \times n$ matrix $T = Z_1(\mathbf{v})$ for $\mathbf{v} \in \mathcal{G}_{\mu,\sigma}^{n \times 1}$. Then*
(a) $F_{||T||}(y) \ge \chi_{\mu,\sigma,n}(2y/n)$ for $\chi_{\mu,\sigma,n}(y)$ in Definition 3.3 and
(b) $F_{1/||T^{-1}||}(y) \le \sqrt{\frac{2}{\pi}} \frac{ny}{\sigma}$.

Proof. For the matrix $T = Z_1(\mathbf{v})$ we have both equation (3.2) and the bound $||\mathbf{t}_-||_1 \le ||\mathbf{t}||_1$, and so $||T||_1 \le 2||\mathbf{t}||_1$. Now part (a) of the theorem follows similarly to (3.3).

To prove part (b) recall Theorem 2.6 and write $B = \Omega T \Omega^{-1} = D(\Omega \mathbf{v})$, $\mathbf{u} = (u_i)_{i=0}^{n-1} = \Omega \mathbf{v}$. We have $\sigma_j(T) = \sigma_j(B)$ for all j because $\frac{1}{\sqrt{n}}\Omega$ and $\sqrt{n}\Omega^{-1}$ are unitary matrices.

By combining the equations $u_i = \mathbf{e}_i^T \Omega \mathbf{v}$, the bounds $||\Re(\mathbf{e}_i^T \Omega)|| \ge 1$ for all i, and Lemma 3.2, deduce that $F_{|\Re(u_i)|}(y) \le \sqrt{\frac{2}{\pi}} \frac{y}{\sigma}$ for $i = 1, \dots, n$. We have $F_{\sigma_n(B)}(y) = F_{\min_i |u_i|}(y)$ because $B = \mathrm{diag}(u_i)_{i=0}^{n-1}$. Clearly $|u_i| \ge |\Re(u_i)|$, and part (b) of the theorem follows. □

Remark 3.2. *Our extensive experiments suggest that the estimates of Theorem 3.6 are overly pessimistic (cf. Table 8.4).*

Combining Theorem 2.6 with minimax property (2.5) implies that

$$\frac{1}{g(f)} \sigma_j(Z_1(\mathbf{v})) \le \sigma_j(Z_f(\mathbf{v})) \le g(f) \sigma_j(Z_1(\mathbf{v}))$$

for all vectors $\mathbf{v}$, scalars $f \ne 0$, $g(f) = \max\{|f|, 1/|f|\}$, and $j = 1, \dots, n$. Thus we can readily extend the estimates of Theorem 3.6 to f-circulant matrices for $f \ne 0$. In particular the estimates do not change in the case of skew circulant matrices (for which $f = -1$), and furthermore Gaussian random f-circulant matrices tend to be well conditioned unless $f \approx 0$ or $1/f \approx 0$.

4 Conditioning of randomized matrix products and generic preconditioning

Next we deduce probabilistic lower bounds on the smallest singular values of the products of fixed and random matrices. We begin with three lemmas. The first two of them easily follow from minimax property (2.5).

Lemma 4.1. *Suppose* $\Sigma = \mathrm{diag}(\sigma_i)_{i=1}^n$, $\sigma_1 \ge \sigma_2 \ge \cdots \sigma_n$, $G \in \mathbb{R}^{r\times n}$, *and* $H \in \mathbb{R}^{n\times r}$. *Then* $\sigma_j(G\Sigma) \ge \sigma_j(G)\sigma_n$, $\sigma_j(\Sigma H) \ge \sigma_j(H)\sigma_n$ *for all* j. *If also* $\sigma_n > 0$, *then* $\mathrm{rank}(G\Sigma) = \mathrm{rank}(G)$, $\mathrm{rank}(\Sigma H) = \mathrm{rank}(H)$.

Lemma 4.2. $\sigma_j(SM) = \sigma_j(MT) = \sigma_j(M)$ *for all* j *if* S *and* T *are square orthogonal matrices.*

Lemma 4.3. *[SST06, Proposition 2.2]. Suppose* $H \in \mathcal{G}_{\mu,\sigma}^{m\times n}$, $SS^T = S^TS = I_m$, $TT^T = T^TT = I_n$. *Then* $SH \in \mathcal{G}_{\mu,\sigma}^{m\times n}$ *and* $HT \in \mathcal{G}_{\mu,\sigma}^{m\times n}$.

Theorem 4.1. *Suppose* $G \in \mathcal{G}_{\mu,\sigma}^{r\times m}$, $H \in \mathcal{G}_{\mu,\sigma}^{n\times r}$, $M \in \mathbb{R}^{m\times n}$, $r(M) = \mathrm{rank}(M)$, *and* $y \ge 0$. *Then* $\max\{F_{1/||(GM)^+||}(y), F_{1/||(MH)^+||}(y)\} \le 2.35y\sqrt{\widehat{r}}/(\sigma_{r(M)}(M)\sigma)$ *for* $\widehat{r} = \min\{r, r(M)\}$, *that is the probability that the reciprocal of the norm* $||(GM)^+||$ *or* $||(MH)^+||$ *exceeds a large value* y *is proportional to* $1/y$.

Proof. With probability 1, the matrix MH has rank $\widehat{r}$ because $H \in \mathcal{G}_{\mu,\sigma}^{n\times r}$. So (cf. (2.7))

$$F_{1/||(MH)^+||}(y) = F_{\sigma_{\widehat{r}}(MH)}(y). \tag{4.1}$$

Let $M = S_M\Sigma_M T_M^T$ be full SVD where $\Sigma_M = \mathrm{diag}(\widehat{\Sigma}_M, O) = \Sigma_M\,\mathrm{diag}(I_{r(M)}, O)$ and $\widehat{\Sigma}_M = \mathrm{diag}(\sigma_j(M))_{j=1}^{r(M)}$ is a nonsingular diagonal matrix.

We have $MH = S_M\Sigma_M T_M^T H$, and so $\sigma_j(MH) = \sigma_j(\Sigma_M T_M^T H)$ for all j by virtue of Lemma 4.2, because S_M is a square orthogonal matrix. Write $H_{r(M)} = (I_{r(M)} \mid O)T_M^T H$ and observe that $\sigma_j(\Sigma_M T_M^T H) = \sigma_j(\widehat{\Sigma}_M H_{r(M)})$ and consequently

$$\sigma_j(MH) = \sigma_j(\widehat{\Sigma}_M H_{r(M)}) \text{ for all } j. \tag{4.2}$$

Combine equation (4.2) for $j = \widehat{r}$ with Lemma 4.1 for the pair (Σ, H) replaced by $(\widehat{\Sigma}_M, H_{r(M)})$ and obtain that $\sigma_{\widehat{r}}(MH) \ge \sigma_{r(M)}(M)\sigma_{\widehat{r}}(H_{r(M)})$.

We have $T_M^T H \in \mathcal{G}_{\mu,\sigma}^{n\times r}$ by virtue of Lemma 4.3, since T_M is a square orthogonal matrix; consequently $H_{r(M)} \in \mathcal{G}_{\mu,\sigma}^{r(M)\times r}$. Therefore we can apply Theorem 3.1 for $A = H_{r(M)}$ and obtain the bound of Theorem 4.1 on $F_{1/||(MH)^+||}(y)$.

One can similarly deduce the bound on $F_{1/||(GM)^+||}(y)$ or can just apply the above bound on $F_{1/||(MH)^+||}(y))$ for $H = G^T$ and M replaced by M^T and then recall that $(M^TG^T)^T = GM$. □

Theorem 4.1 implies that multiplication by standard Gaussian random matrix is unlikely to decrease the smallest positive singular value of a matrix dramatically, even though $UV = O$ for some pairs of rectangular orthogonal matrices U and V.

By combining (2.3) with Theorems 3.2 for $B = O$ and 4.1 one can probabilistically bound the condition numbers of randomized matrix products GM and MH.

The following corollary extends the bound of Theorem 4.1 for a randomized matrix product to the respective bounds for its leading blocks. This implies that *randomized multiplication of a well conditioned matrix is expected to be generic preconditioning*, that is, to ensure (with probability 1 or near 1) generic rank and conditioning profiles for the product.

Corollary 4.1. *Suppose* j, k, m, n, q *and* s *are integers,* $1 \le j \le q$, $1 \le k \le s$, $M \in \mathbb{R}^{m\times n}$, $\sigma > 0$, $G \in \mathcal{G}_{\mu,\sigma}^{q\times m}$, $H \in \mathcal{G}_{\mu,\sigma}^{n\times s}$, $\mathrm{rank}(M_j) = j$ *for* $M_j = M\begin{pmatrix} I_j \\ O_{n-j,j}\end{pmatrix}$, $\mathrm{rank}(M^{(k)}) = k$ *for* $M^{(k)} = (I_k \mid O_{k,m-k})M$, *and* $y \ge 0$. *Then (i) with probability* 1 *the matrix* GM *(resp.* MH*) has generic rank profile if* $\mathrm{rank}(M) \ge q$ *(resp. if* $\mathrm{rank}(M) \ge s$*). Furthermore (ii)* $F_{1/||((GM)_j^{(j)})^+||}(y) \le 2.35y\sqrt{j}/(\sigma_j(M_j)\sigma)$ *if* $\mathrm{rank}(M) \ge j$, $F_{1/||((MH)_k^{(k)})^+||}(y) \le 2.35y\sqrt{k}/(\sigma_k(M^{(k)})\sigma)$ *if* $\mathrm{rank}(M) \ge k$.

Proof. We immediately verify part (i) by applying the techniques of Section 3.2. To prove part (ii) apply Theorem 4.1 for every j, replacing G by $(I_j \mid O_{j,q-j})G$ and M by $M\begin{pmatrix} I_j \\ O_{n-j,j}\end{pmatrix}$. For every k apply Theorem 4.1 replacing M by $(I_k \mid O_{k,m-k})M$ and H by $H\begin{pmatrix} I_k \\ O_{s-k,k}\end{pmatrix}$. □

Combining the latter results with (2.3) and Theorem 3.2 for $B = O$ implies that for well conditioned inputs randomized multiplication is expected to serve as generic preconditioning, and if it does serve so, then both Gaussian elimination with no pivoting and block Gaussian elimination are numerically safe. This is stated in the following corollary, implied by Theorem 2.4.

Corollary 4.2. *Suppose M is a normalized well conditioned $m \times n$ matrix of full rank, $||M|| = 1$, $G \in \mathcal{G}_{0,1}^{m\times m}$ and $H \in \mathcal{G}_{0,1}^{n\times n}$. Then Gaussian elimination with no pivoting as well as block Gaussian elimination applied to the matrices GM and MH are expected to proceed using no divisions by absolutely small values.*

Remark 4.1. *We cannot extend Lemma 4.3 and consequently Theorem 4.1 and its corollaries to the case of Gaussian random Toeplitz matrices $G \in \mathcal{T}_{\mu,\sigma}^{r\times m}$ and $H \in \mathcal{T}_{\mu,\sigma}^{n\times r}$, but such extensions have been supported empirically (cf. Tables 8.6 and 8.10).*

5 Randomized additive and dual additive preconditioning

In this section we first prove Theorem 1.1 and then extend it to the cases of rectangular $m \times n$ matrices A for any pair $\{m, n\}$ and of the dual additive preprocessing, generalizing (2.8) and (2.9). The respective rank estimates are implicit in Section 3.2, and we only prove the norm and condition estimates.

5.1 Proof of Theorem 1.1

Theorem 5.1. *Suppose $A, C, S, T \in \mathbb{R}^{n\times n}$ and $U, V \in \mathbb{R}^{n\times r}$ for two positive integers r and n, $r \le n$, $A = S\Sigma T^T$ is full SVD of the matrix A (cf. (2.4)), S and T are orthogonal matrices, $\Sigma = \mathrm{diag}(\sigma_j)_{j=1}^n$, the matrix $C = A + UV^T$ is nonsingular, and so $\rho = n - r \le \mathrm{rank}(A)$ and $\sigma_\rho > 0$. Write*

$$S^TU = \begin{pmatrix} \bar{U} \\ U_r \end{pmatrix},\ T^TV = \begin{pmatrix} \bar{V} \\ V_r \end{pmatrix},\ R_U = \begin{pmatrix} I_\rho & \bar{U} \\ O_{r,\rho} & U_r \end{pmatrix},\ R_V = \begin{pmatrix} I_\rho & \bar{V} \\ O_{r,\rho} & V_r \end{pmatrix}, \tag{5.1}$$

where U_r and V_r are $r \times r$ matrices. Then

(a) *$R_U \Sigma R_V^T = \Sigma$, whereas $R_U\, \mathrm{diag}(O_{\rho,\rho}, I_r) R_V^T = S^TUV^TT$, and so*

$$C = SR_U D R_V^T T^T,\ D = \Sigma + \mathrm{diag}(O_{\rho,\rho}, I_r) = \mathrm{diag}(d_j)_{j=1}^n \tag{5.2}$$

where $d_j = \sigma_j$ for $j = 1, \dots, \rho$, $d_j = \sigma_j + 1$ for $j = \rho + 1, \dots, n$.

Furthermore suppose that the matrix A has been normalized so that $||A|| = 1$ and that the $r \times r$ matrices U_r and V_r are nonsingular, which holds with probability 1 where U and V are Gaussain random matrices (cf. Section 3.2). Write

$$p = ||R_U^{-1}||\ ||R_V^{-1}|| \text{ and } f_r = \max\{1, ||U_r^{-1}||\}\ \max\{1, ||V_r^{-1}||\}. \tag{5.3}$$

Then

(b) *the matrix C is nonsingular,*

(c) $1 - ||U||\ ||V|| \le ||C|| \le 1 + ||U||\ ||V||$,

(d) $1 \le ||R_V||\ ||R_U|| \le \sigma_\rho(A)/\sigma_n(C) \le p$,

(e) $p \le (1 + ||U||)(1 + ||V||) f_r$,

(f) $1 \le \sigma_\rho(A)/\sigma_n(C) \le (1 + ||U||)(1 + ||V||) f_r$.

Proof. Parts (a) and (b) are readily verified.

(c) Combine the relationships $||A|| - ||U||\ ||V|| \le ||C|| \le ||A|| + ||U||\ ||V^T||$, $||A|| = 1$ and $||V^T|| = ||V||$.

(d) Combine the equations $S^{-1} = S^T$, $T^{-1} = T$ and (5.2) and obtain $C^{-1} = TR_V^{-T} D^{-1} R_U^{-1} S^T$ or equivalently $D^{-1} = R_V T^T C^{-1} S R_U$. It follows that $||C^{-1}|| = ||R_V^{-T} D^{-1} R_U^{-1}||$ and $||D^{-1}|| = ||R_V T^T C^{-1} S R_U||$. Apply bound (2.3), substitute $||S|| = ||S^T|| = ||T|| = ||T^T|| = 1$ and obtain $||C^{-1}|| \le ||R_V^{-T}||\ ||D^{-1}||\ ||R_U^{-1}||$ and $||D^{-1}|| \le ||R_V||\ ||C^{-1}||\ ||R_U||$. Substitute the equations (5.3),

$||D^{-1}|| = 1/\sigma_\rho(A)$ (implied by the equations $||A|| = 1$ and (5.2)) and $||C^{-1}|| = 1/\sigma_n(C)$ and the bounds $||R_V|| \ge 1$ and $||R_U|| \ge 1$ and obtain that $1 \le \sigma_\rho(A)/\sigma_n(C) \le p$.

(e) Observe that $R_U^{-1} = \begin{pmatrix} I_\rho & -\bar{U} \\ O & I_r \end{pmatrix} \begin{pmatrix} I_\rho & O \\ O & U_r^{-1} \end{pmatrix}$, $R_V^{-1} = \begin{pmatrix} I_\rho & -\bar{V} \\ O & I_r \end{pmatrix} \begin{pmatrix} I_\rho & O \\ O & V_r^{-1} \end{pmatrix}$, $||\bar{U}|| \le ||U||$ and $||\bar{V}|| \le ||V||$. Then combine these relationships.

(f) Combine the bounds of parts (d) and (e). □

We have $\frac{\kappa(C)}{\kappa(A)} \le \frac{||C||}{||A||}\frac{\sigma_\rho(A)}{\sigma_n(C)}$, and so parts (e) and (f) together bound the ratio $\frac{\kappa(C)}{\kappa(A)}$ in terms of the norms $||U||$, $||V||$, $||U_r^{-1}||$ and $||V_r^{-1}||$ as follows,

$$\frac{\kappa(C)}{\kappa(A)} \le (1 + ||U||\,||V||)(1 + ||U||)(1 + ||V||)\ \max\{1, ||U_r^{-1}||\}\ \max\{1, ||V_r^{-1}||\}, \tag{5.4}$$

and in particular

$$\frac{\kappa(C)}{\kappa(A)} \le (1 + ||U||^2)(1 + ||U||)^2 \max\{1, ||U_r^{-1}||^2\} \text{ if } U = V. \tag{5.5}$$

Let us estimate these norms where U and V are Gaussian random matrices.

Theorem 5.2. *Suppose $W \in \mathcal{G}_{0,\sigma}^{n\times r}$ and $y \ge 2\sigma\sqrt{n}$. Then $F_{||W||}(y) \ge 1 - \exp(-(y - 2\sigma\sqrt{n})^2/(2\sigma^2))$.*

Proof. The theorem follows from Theorem 3.2 applied for $A = W$ (in which case $l = r$) and $B = O$. □

Theorem 5.3. *Suppose that A, U, V, U_r and V_r denote the five matrices of Theorem 5.1 where $U, V \in \mathcal{G}_{\mu,\sigma}^{n\times r}$. Then $\max\{F_{1/||U_r^{-1}||}(y), F_{1/||V_r^{-1}||}(y)\} \le 2.35\ y\sqrt{r}/\sigma$ for $y \ge 0$.*

Proof. Lemma 4.3 implies that $S^TU, T^TV \in \mathcal{G}_{\mu,\sigma}^{n\times r}$ because S and T are orthogonal matrices. Therefore $U_r, V_r \in \mathcal{G}_{\mu,\sigma}^{r\times r}$. Apply Theorem 3.1 for $A = U_r$ and $A = V_r$ where in both cases $m = n = r$. □

Combine bounds (5.3)–(5.5) with Theorem 5.2, for $W = U$ and $W = V$, and Theorem 5.3 and obtain Theorem 1.1.

Remark 5.1. *(Cf. Remark 7.10.) Theorem 1.1 can guide us in choosing the size $n \times r$ of the matrices U and V. Surely there is no point for applying the preprocessing $A \Longrightarrow C = A + UV^T$ to a well conditioned matrix A of full rank, whereas the output matrix C remains singular or ill conditioned where the numerical nullity ν of the matrix A exceeds r. The theorem shows that for a normalized input matrix A and two independent standard Gaussian random matrices U and V as well as for standard Gaussian random matrix U and for $V = U$ the above preprocessing is expected to produce a well conditioned matrix C of full rank as long as $\nu \le r$. The best choice $r = \nu$ minimizes the size of the matrices U and V and leads to some other simplifications. These comments can be extended to the results of Section 6 and Corollary 5.1 of the next subsection.*

5.2 Extension of Theorem 1.1 to the case of rectangular matrices

Clearly our upper bound on the ratio $||C||/||A||$ holds for any pair of m and n. Furthermore we can apply Theorem 5.2 for n replaced by m and can immediately extend Theorem 5.3 to the case where $U \in \mathcal{G}_{\mu,\sigma}^{m\times r}$. Let us also relax the assumption that $m = n$ for our upper bound on $\sigma_\rho(A)/\sigma_n(C)$.

Theorem 5.4. *Suppose that $A \in \mathbb{R}^{m\times n}$, $U \in \mathcal{G}_{0,\sigma}^{m\times r}$, and $V \in \mathcal{G}_{0,\sigma}^{n\times r}$ for three positive integers m, n and r and assume that the matrix $C = A + UV^T$ has full rank $l = \min\{m, n\}$ (and thus is nonsingular where $m = n$) and that $\rho = l - r \le \text{rank}(A)$. Keep equation (5.1) but write*

$$I_{m,n}S^TU = \begin{pmatrix} \bar{U} \\ U_r \end{pmatrix},\ I_{m,n}T^TV = \begin{pmatrix} \bar{V} \\ V_r \end{pmatrix} \tag{5.6}$$

for $I_{g,h}$ defined in (2.1) where U_r and V_r still denote $r \times r$ matrices. Keep the assumptions of parts (a)–(f) of Theorem 5.1. Then the upper bound of part (f) of Theorem 5.1 can be extended, that is, $\sigma_\rho(A)/\sigma_l(C) \le (1 + ||U||)(1 + ||V||)f_r$ where $f_r = \max\{1, ||U_r^{-1}||\}\ \max\{1, ||V_r^{-1}||\}$ as in (5.3) but for U_r and V_r defined above.

Proof. Let $A = S_A \Sigma_A T_A^T$ denote SVD of (2.4). Write $\widehat{C} = I_{m,n} S_A^T C T_A I_{n,m}^T$, $\widehat{U} = I_{m,n} S_A^T U$, $\widehat{V} = I_{n,m} T_A^T V$, $\widehat{A} = I_{m,n} S_A^T A T_A I_{n,m}^T$, and so $\widehat{C} = \widehat{A} + \widehat{U}\widehat{V}^T$. Note that $\widehat{A}$ and $\widehat{C}$ are the $l \times l$ leading submatrices of the matrices $S_A^T A T_A = \Sigma_A$ and $S_A^T C T_A$, respectively. Consequently $\widehat{A} = (\sigma_j(A))_{j=1}^l$, $\sigma_j(\widehat{A}) = \sigma_j(A)$, and $\sigma_j(\widehat{C}) \le \sigma_j(S_A^T C T_A) = \sigma_j(C)$ for all j. Furthermore $\widehat{U}$ and $\widehat{V}$ are $l \times r$ submatrices of the matrices $S_A^T U$ and $T_A^T V$, respectively, and so $\sigma_j(\widehat{U}) \le \sigma_j(U)$ and $\sigma_j(\widehat{V}) \le \sigma_j(V)$ for all j. Apply Theorem 5.3 to the $l \times l$ matrices $\widehat{A}$ and $\widehat{C}$ and obtain that $\sigma_\rho(\widehat{A})/\sigma_n(\widehat{C}) \le (1 + ||\widehat{U}||)(1 + ||\widehat{V}||) f_r$. It remains to substitute the following relationships that we have proved above, $\sigma_\rho(\widehat{A}) = \sigma_\rho(A)$, $\sigma_l(\widehat{C}) \le \sigma_l(C)$, $||U|| = \sigma_1(U) \ge \sigma_1(\widehat{U}) = ||\widehat{U}||$ and $||V|| = \sigma_1(V) \ge \sigma_1(\widehat{V}) = ||\widehat{V}||$. □

In the above proof the transitions $U \Longrightarrow \widehat{U} = I_{m,n} S_A^T U$ and $V \Longrightarrow \widehat{V} = I_{n,m} T_A^T V$ have yielded the matrices $\widehat{U}$ and $\widehat{V}$ in the classes $\mathcal{G}_{\mu,\sigma}^{l \times r}$ and therefore the matrices U_r and V_r in the classes $\mathcal{G}_{\mu,\sigma}^{r \times r}$, and so we can combine Theorem 5.4 with Theorem 5.2 for $W = U$ and $n = m$ and for $W = V$ as well as with Theorem 5.3 to yield the following probabilistic estimate for the ratio $\sigma_\rho(A)/\sigma_l(C)$.

Theorem 5.5. *Assume that $A \in \mathbb{R}^{m \times n}$, $U \in \mathcal{G}_{0,\sigma}^{m \times r}$, $V \in \mathcal{G}_{0,\sigma}^{n \times r}$, $C = A + UV^T$, and $l = \min\{m, n\}$. Then the matrix C is rank deficient if $r < l - \mathrm{rank}(A)$. Otherwise with probability 1 it has full rank (and therefore is nonsingular if $m = n$) as well as the matrices U and V, and with probability 1 the bound of Theorem 5.4 holds (reproducing part (f) of Theorem 5.1). In this bound the norms $||U||$ and $||V||$ satisfy the randomized bounds of Theorem 5.2 (for $W = U$ and m replacing n as well as for $W = V$), and the reciprocals of the norms $||U_r^{-1}||$ and $||V_r^{-1}||$ satisfy the randomized bounds of Theorem 5.3.*

Corollary 5.1. *Theorem 1.1 can be extended to the case of matrices $A \in \mathbb{R}^{m \times n}$, $U \in \mathcal{G}_{0,\sigma}^{m \times r}$ and $V \in \mathcal{G}_{0,\sigma}^{n \times r}$ for any pair of positive integers $\{m, n\}$ and for $l = \min\{m, n\}$, that is the matrix $C = A + UV^T$ is rank deficient if $r < l - \mathrm{rank}(A)$, whereas for $l - \mathrm{rank}(A) \le r < l$ it has full rank with probability 1 and is expected to have condition number of at most order $||A||/\sigma_{l-r}(A)$; consequently the matrix C is expected to be nonsingular and well conditioned if the matrix A has numerical rank at least $l - r$.*

Corollary 5.2. *Suppose that A, U and V denote the same matrices as in Corollary 5.1, $l = \min\{m, n\}$, and $C = A + UMV^T$ for a nonsingular $r \times r$ matrix M. Then Theorem 1.1 and Corollary 5.1 can be extended as follows: the matrix C is rank deficient if $r < l - \mathrm{rank}(A)$, whereas for $l - \mathrm{rank}(A) \le r < l$ it has full rank with probability 1 and is expected to have condition number of at most order $\kappa(M)||A||/\sigma_{l-r}(A)$.*

Proof. Let $M = S_M \Sigma_M T_M^T$ be SVD and rewrite $C = A + UMV^T$ as $C = A + \bar{U}\Sigma_M \bar{V}^T$ where $\bar{U} = US_M\Sigma_M$ and $\bar{V} = T_M^T V$. Note that $\bar{U} \in \mathcal{G}_{0,\sigma}^{m \times r}$ and $\bar{V} \in \mathcal{G}_{0,\sigma}^{n \times r}$ by virtue of Lemma 4.3. Now reapply the proofs of this section replacing U by $\bar{U}$ and V by $\bar{V}$ and observe that all the proofs are extended except that estimates for the norm V_r^{-1} increase by at most a factor $\kappa(\Sigma_M) = \kappa(M)$ by virtue of Lemma 4.1. □

Remark 5.2. *Parts (c) and (f) of Theorem 5.1 provide lower bounds on the norms $||C||$ and $||C^{-1}||$ in terms of the norm $||A||$ and the value $\sigma_{l-r}(A)$. In particular Theorem 3.2 implies that the norm $||C||$ is expected to exceed 1/2, for $||A|| = 1$ and $\sigma = 0.1/\sqrt{rn}$, say. Thus we can use these bounds to prove that under the assumptions of Corollary 5.2 the condition number $\kappa(C)$ is expected to have order $\kappa(\Sigma_M)/\sigma_{l-r}(A)$ (rather than just at most this order) for $l - \mathrm{rank}(A) \le r < l$ under these assumptions on $||A||$ and and σ. By writing $M = I$ we similarly strengthen Theorem 1.1 and Corollary 5.1.*

Remark 5.3. *As in Theorem 1.1 our results and proofs of this section and the next one can be applied in both cases where all entries of the matrices U and V are independent Gaussian random varaibles and where these two matrices share some or all their entries, which allows us to proceed with fewer random parameters. Empirically the presented randomized techniques work very well even under very weak randomization, where the matrices U and V are defined with much fewer random parameters (cf. Tables 8.5, 8.6, 8.10, and 8.17).*

Remark 5.4. *How large is our class of $m \times n$ matrices $\bar{A}$ having a numerical rank q? We characterize it indirectly, by noting that by virtue of Fact 2.2 the nearby matrices A of rank q form a variety of dimension $(m+n-q)q$, which increases as q increases.*

5.3 Dual additive preconditioning

For an $m \times n$ matrix A of full rank we extend (2.8) and (2.9) to define the *dual additive preprocessing*

$$A^+ \Longrightarrow C_-^+ = A^+ + U_- V_-^T. \tag{5.7}$$

Our analysis implies that the value $\kappa(C_-^+)$ (equal to $\kappa(C_-)$) is expected to have order $\sigma_{q+1}(A)/\sigma_l(A)$ provided $l = \min\{m,n\}$, $U_- \in \mathcal{G}_{0,1}^{n\times q}$, $V_- \in \mathcal{G}_{0,1}^{m\times q}$, and the norm $||A^+||$ is neither large nor small. The randomized algorithm of [D83] is expected to estimate the norm $||A^+||$ at a low computational cost. We can work with the $(m+1)\times(n+1)$ matrix $\widehat{A} = \mathrm{diag}(A, \epsilon)$ instead of the matrix A and choose a sufficiently small positive scalar ϵ such that $||\widehat{A}^+|| = 1/\epsilon$. Then we can scale the matrix $\widehat{A}$ to obtain that $||(\widehat{A}/\epsilon)^+|| = 1$.

6 Randomized augmentation

6.1 Augmentation and the SMW formula

The solution of a nonsingular linear system of n equations, $A\mathbf{y} = \mathbf{b}$ can be readily recovered from a null vector $\begin{pmatrix} -1/\beta \\ \mathbf{y} \end{pmatrix}$ of the matrix $K = (\beta\mathbf{b} \mid A)$ for a nonzero scalar β. If the matrix A has numerical nullity 1 and if the ratio $||A||/||\beta\mathbf{b}||$ is neither large nor small, then the matrix K is well conditioned for the average vector $\mathbf{b}$ [PQa, Section 13.1]. The above map $A \Longrightarrow K$ is a special case of more general augmentation

$$K = \begin{pmatrix} W & V^T \\ -U & A \end{pmatrix}, \tag{6.1}$$

which we study next, beginning with the following extension of the SMW formula.

Theorem 6.1. *Suppose equation (6.1) holds, $m = n$ and the matrices A, W and K are nonsingular. Write $S = A + UW^{-1}V^T$ and $R = I - V^T S^{-1} U W^{-1}$. Then the matrix S is nonsingular and S^{-1} is the trailing (southwestern) $n \times n$ block of K^{-1}. Furthermore*

$$A^{-1} = S^{-1} + S^{-1}UW^{-1}R^{-1}V^T S^{-1}. \tag{6.2}$$

Proof. Apply the SMW formula of Theorem 2.3 for C replaced by S, U by UW^{-1}, and G by R. □

6.2 Links to additive preprocessing and condition estimates

The following theorem is readily verified. It links additive preprocessing $A \Longrightarrow C = A + UMV^T$ to the augmentation $A \Longrightarrow K$ for K of (6.1).

Theorem 6.2. *Suppose $A \in \mathbb{R}^{m\times n}$, $W \in \mathbb{R}^{r\times r}$, the matrix W is nonsingular, $l = \min\{m,n\}$, a matrix K in $\mathbb{R}^{(m+r)\times(n+r)}$ is defined by (6.1), and $C = A + UW^{-1}V^T$. Then $K = \widehat{U}\,\mathrm{diag}(C, I_r)\widehat{V}$ for $\widehat{U} = \begin{pmatrix} O_{r,m} & I_r \\ I_m & -U \end{pmatrix}$, $\widehat{V} = \begin{pmatrix} O_{n,r} & I_n \\ I_r & V^T \end{pmatrix}$, the matrix C has full rank if and only if the matrix K has full rank, and so both matrices are rank deficient for $r < l$. Furthermore $\widehat{U}^{-1} = \begin{pmatrix} U & I_n \\ I_r & O_{r,n} \end{pmatrix}$, $\widehat{V}^{-1} = \begin{pmatrix} -V^T & I_r \\ I_n & O_{n,r} \end{pmatrix}$, and if $m = n$ and the matrices C and K are nonsingular, then $C^{-1} = (O_{n,r} \mid I_n)K^{-1}(O_{n,r} \mid I_n)^T$ and $K^{-1} = \widehat{V}^{-1}\,\mathrm{diag}(C^{-1}, I_r)\widehat{U}^{-1}$.*

Corollary 6.1. *(Cf. [PQa, Remark 10.1 and Corollary 11.1].) Define three integers m, n, and l and three matrices A, K, and W as in Theorem 6.2 and suppose that $r \geq l - \mathrm{rank}(A)$, $U \in \mathcal{G}_{0,1}^{m\times r}$ and $V \in \mathcal{G}_{0,1}^{n\times r}$. Then (i) the matrix K has full rank with probability 1 and is expected to have condition number of at most order $\kappa(W)/\sigma_{l-r}(A)$. (ii) This number is expected to be of at most order $1/\sigma_{l-r}(A)$ where $W \in \mathcal{G}_{0,1}^{r\times r}$; consequently the matrix K is expected to be nonsingular and well conditioned if the matrix A has numerical rank at least $l - r$.*

Proof. To deduce part (i) combine Theorem 6.2 and Corollary 5.2 for $M = W^{-1}$ and then recall that $\kappa(W^{-1}) = \kappa(W)$. To extend part (i) to part (ii) note that a matrix in $\mathcal{G}_{0,1}^{r\times r}$ is nonsingular with probability 1 and is expected to be well conditioned (see Sections 3.2 and 3.3). □

6.3 Direct condition estimation

Let us estimate the condition number $\kappa(K)$ directly, without reducing this task to additive preprocessing.

Theorem 6.3. *Suppose $A \in \mathbb{R}^{m\times n}$, $||A|| = 1$, $W \in \mathcal{G}_{0,1}^{r\times q}$, $U \in \mathcal{G}_{0,1}^{m\times q}$, $V \in \mathcal{G}_{0,1}^{n\times r}$, K in $\mathbb{R}^{(m+r)\times(n+r)}$ defined by (6.1), $\rho = \mathrm{rank}(A)$, $l = \min\{m, n\}$, $s = \min\{m - q, n - r\}$, and $s > 0$. Then*

(i) the matrix K is rank deficient if $\rho < s$, whereas

(ii) otherwise it has full rank $l' = s + \rho$ with probability 1 and

(iii) is expected to have the condition number $\kappa(K)$ of order at most $1/\sigma_s(A)$; consequently the matrix K is expected to be nonsingular and well conditioned if the matrix A has numerical rank at least s.

Proof. The claims about the rank of the matrix K is readily verified (cf. Section 3.2). Let us assume that $\rho + r + q \geq l' = \mathrm{rank}(K)$ and estimate the condition number $\kappa(K) = ||K'||/\sigma_{l'}(K)$. By virtue of Theorem 3.2 the norms of the matrices U, V, and W are in $O(1)$, that is bounded by a constant, and so $||K|| = O(1)||$ as well because $||K|| \leq ||U|| + ||V|| + ||W|| + ||A||$ and $||A|| = 1$. It remains to estimate the value $\sigma_{l'}(K)$ from below. Furthermore it is sufficient to consider the case where $l' = m + r \leq n + q$, and so $s = m - q$. Indeed otherwise we can prove the theorem for the matrix K^T and then immediately extend it to the matrix K because $\kappa(K) = \kappa(K^T)$.

At first assume that $r = 0$. Then $l' = \mathrm{rank}(K) = m$ and V and W are empty matrices. Let $A = S_A\Sigma_A T_A^T$ be SVD of (2.4). Write $K' = S_A^T K\,\mathrm{diag}(I_q, T_A) = (U' \mid \Sigma_A)$ where $\sigma_{l'}(K) = \sigma_m(K')$ because S_A, T_A and $\mathrm{diag}(T_A, I_q)$ are square orthogonal matrices and where $U' = S_A^T U \in \mathcal{G}_{0,1}^{m\times q}$ because S_A is an orthogonal matrix (cf. Lemma 4.3).

The $m \times (n + q)$ matrix K' has $m \times m$ leading submatrix $\bar{K} = \begin{pmatrix} U_0 & \Sigma_{m-q} \\ U_1 & O_{q,m-q} \end{pmatrix}$ where $U_0 \in \mathcal{G}_{0,1}^{(m-q)\times q}$, $U_1 \in \mathcal{G}_{0,1}^{q\times q}$, $\mathrm{rank}(\bar{K}) = \mathrm{rank}(K) = m$, $\Sigma_{m-q} = \mathrm{diag}(\sigma_j(A))_{j=1}^{m-q}$, and so $\mathrm{rank}(\Sigma_{m-q}) = m - q$ and $\sigma_{l'}(K) \geq \sigma_m(\bar{K}) = 1/||\bar{K}^{-1}||$. We have

$$\bar{K}^{-1} = \begin{pmatrix} O_{q,m-q} & U_1^{-1} \\ \Sigma_{m-q}^{-1} & -\Sigma_{m-q}^{-1}U_0U_1^{-1} \end{pmatrix} = \mathrm{diag}(I_q, \Sigma_{m-q}^{-1}) \begin{pmatrix} O_{q,n} & I_q \\ I_{m-q} & -U_0 \end{pmatrix} \mathrm{diag}(I_{m-q}, U_1^{-1}).$$

Therefore $||\bar{K}^{-1}|| \leq ||\Sigma_{m-q}^{-1}||(1 + ||U_0||)||U_1^{-1}||$ where $||\Sigma_{m-q}^{-1}|| = 1/\sigma_{m-q}(A)$. Theorems 3.1 and 3.2 bound the norms $||U_1^{-1}||$ and $||U_0||$, and therefore the value $1/\sigma_{l'}(K) = ||\bar{K}^{-1}||$ is expected to have at most order $1/\sigma_{m-q}(A)$.

Now let $r > 0$. Represent the matrix K as $\begin{pmatrix} B \\ F \end{pmatrix}$ where $B = (W \mid V^T)$, $F = (-U \mid A)$ is the matrix studied above (in the case where $r = 0$), and so the value $||F^+||$ has order at most $1/\sigma_{m-q}(A)$. Let $F = S_F\Sigma_F T_F^T$ be SVD and write $K'' = \mathrm{diag}(I_r, S_F^T)KT_F = \begin{pmatrix} B_0 & B_1 \\ \bar{\Sigma}_F & O_{m,n+q-m} \end{pmatrix}$ where $B_0 \in \mathcal{G}_{0,1}^{r\times m}$, $B_1 \in \mathcal{G}_{0,1}^{r\times(n+q-m)}$, $\bar{\Sigma}_F = \mathrm{diag}(\sigma_j(F))_{j=1}^m$, $\mathrm{rank}(K'') = \mathrm{rank}(K) = m + r$, and so the matrix $\bar{\Sigma}_F$ is nonsingular and the value $1/\sigma_m(F) = ||F^+||$ is expected to have at most order $1/\sigma_{m-q}(A)$.

Delete the last $n+q-m-r$ columns of the matrix K'' and obtain the $(m+r)\times(m+r)$ submatrix $\widehat{K} = \begin{pmatrix} B_0 & \bar{B}_1 \\ \bar{\Sigma}_F & O_{m,n+q-m} \end{pmatrix}$. We have $\sigma_{l'}(K) = \sigma_{l'}(K'') \geq \sigma_m(\widehat{K})$. The Gaussian random $r \times r$ matrix $\bar{B}_1$ is nonsingular with probability 1. We assume that it is nonsingular, and then so is the matrix $\widehat{K}$ as well, and consequently $\sigma_{l'}(K) = \sigma_m(\widehat{K}) = 1/||\widehat{K}^{-1}||$. Observe that

$$\widehat{K}^{-1} = \begin{pmatrix} O_{m,n+q-m} & \bar{\Sigma}_F^{-1} \\ \bar{B}_1^{-1} & -\bar{B}_1^{-1}B_0\bar{\Sigma}_F^{-1} \end{pmatrix} = \mathrm{diag}(I_q, \bar{B}_1^{-1}) \begin{pmatrix} O_{q,n} & I_q \\ I_{m-q} & -B_0 \end{pmatrix} \mathrm{diag}(I_{m-q}, \Sigma_F^{-1}).$$

Therefore $||\widehat{K}^{-1}|| \leq ||\bar{B}_1^{-1}||(1 + ||B_0||)||\Sigma_F^{-1}||$. Theorems 3.1 and 3.2 bound the norms $||\bar{B}_1^{-1}||$ and $||B_0||$. To complete the proof of the theorem recall that the norm $||\Sigma_F^{-1}||$ is expected to have at most order $1/\sigma_{m-q}(A)$. □

Compared to Corollary 6.1, Theorem 6.3 allows rectangular matrices W in $\mathcal{G}^{r\times q}$. Combined with Theorem 6.2 it implies Corollary 5.2 restricted to the case where $M^{-1} \in \mathcal{G}_{0,1}^{r\times r}$. In Appendix D we assume that $r = l - \mathrm{rank}\, A$ and extend Theorem 6.3 by relaxing such restriction on the block W in the most interesting case where $r = l - \mathrm{rank}\, A$.

Remark 6.1. *Corollary 6.1 and Theorem 6.3 suggest that randomized augmentation and additive preprocessing are equally effective for preconditioning, but here is an example of subtle differences. The map* $A \Longrightarrow K = \begin{pmatrix} W & V^T \\ V & A \end{pmatrix}$ *cannot decrease the condition number* $\kappa(A)$ *if* K *is a symmetric and positive definite matrix; this follows from the Interlacing Property of the eigenvalues of* K *[GL96, Theorem 8.1.7]. In contrast scaled randomized symmetric additive preprocessing* $A \Longrightarrow C = A + VV^T$ *is expected to work as preconditioning for an ill conditioned symmetric positive definite matrix* A *having a small numerical nullity (cf. (5.5) and [W07]).*

Further study of randomized augmentation can be found in [PQa]. The results of [PQa, Corollary 11.1] are similar to Theorem 6.3, but [PQa] only provides a pointer to the idea of a proof.

6.4 A randomized Toeplitz solver

Let us apply Theorem 2.5 to support randomized augmentation for solving a nonsingular Toeplitz linear system $T\mathbf{y} = \mathbf{b}$ of n equations provided the matrix T has numerical nullity 1.

To compute the vector $\mathbf{y} = T^{-1}\mathbf{b}$, we first embed the matrix T into a Toeplitz $(n+1)\times(n+1)$ matrix $K = \begin{pmatrix} w & \mathbf{v}^T \\ \mathbf{f} & T \end{pmatrix}$. We write $w = \mathbf{e}_1^T T \mathbf{e}_1$ and fill the vectors $\mathbf{f} = (f_i)_{i=1}^n$ and $\mathbf{v} = (v_i)_{i=1}^n$ with appropriate entries of the matrix T except for the two coordinates f_n and v_n, which we choose at random and then scale to have the ratios $\frac{|f_n|}{||K||}$ and $\frac{|v_n|}{||K||}$ neither large nor small.

Part (b) of Theorem 2.5 expresses the inverse T^{-1} via the vectors $\mathbf{v} = K^{-1}\mathbf{e}_1$ and $\mathbf{w} = K^{-1}\mathbf{e}_{n+1}$.

In view of Section 3.2 and Appendix A, this policy is likely to produce a nonsingular matrix K whose inverse is likely to have a nonzero entry $\mathbf{e}_1^T K^{-1}\mathbf{e}_1$. Our tests were in very good accordance with these formal results and moreover consistently produced well conditioned matrices K.

To summarize, we reduce the solution of a nonsingular ill conditioned Toeplitz linear system $T\mathbf{y} = \mathbf{b}$ to computing highly accurate solutions of two linear systems $K\mathbf{x} = \mathbf{e}_1$ and $K\mathbf{z} = \mathbf{e}_{n+1}$, both expected to be well conditioned. High accuracy shall counter the magnification of the input and rounding errors, expected in the case of ill conditioned input.

In the important special case where a Toeplitz matrix T is real symmetric, we can choose real scalars w and $f_n = v_n$ to yield a real symmetric matrix $K = \begin{pmatrix} w & \mathbf{v}^T \\ \mathbf{v} & T \end{pmatrix}$. In this case the computation is simplified because $J_{n+1}K^{-1}J_{n+1} = K^{-1}$, $J_{n+1}\mathbf{e}_{n+1} = \mathbf{e}_1$, and so $K^{-1}\mathbf{e}_{n+1} = J_{n+1}K^{-1}\mathbf{e}_1$, and we only need to solve a single linear system with the matrix K. We refer to the resulting algorithm for the linear system $T\mathbf{y} = \mathbf{b}$ as **Algorithm 6.1**. In Section 8.7 we test this algorithm for solving an ill conditioned real symmetric Toeplitz linear system.

One can readily extend the approach of this section to the case of Toeplitz-like, Hankel and Hankel-like inputs and to augmenting the input matrix with r rows and r columns for $r > 1$.

For the transition back to the solution of the original problem, we can employ expression (6.2) or Theorem 2.5.

7 Applications of randomized additive preprocessing and augmentation

In this section we outline some randomized matrix algorithms that utilize the results of the three previous sections. In the next subsection we express approximations of a matrix by matrices of smaller ranks in terms of orthogonal matrix bases for its left and right leading singular spaces. In Section 7.2 we first recall the results of [PQ10] on computing a nmb and a left nmb of a rank deficient matrix and then apply Corollary 5.1 to extend these results to randomized approximation of trailing and leading singular spaces of ill conditioned matrices. We also recall two alternative techniques for this task, based on randomized augmentation and of randomized sampling, respectively. In Sections 7.3–7.6 we combine the algorithms of Sections 7.1 and 7.2 with various other techniques to compute lower-rank approximations of an ill conditioned matrix, its numerical rank and numerical nullity, its 2×2 block triangulation and inverse, and the solution of linear systems with this matrix. In Section 7.7 we solve such linear systems by combining Corollary 5.1 with the SMW formula and show significant decrease of the estimated computational cost versus Gaussian elimination. Similar decrease can rely on 2×2 block triangulation of Section 7.3. In Section 7.8 we comment on using random structured matrices in these computations, and in Section 7.9 on the links between our randomized computations and Newton's iteration for matrix inversion. Sections 7.5 and 7.6 are dual to one another, and as we said, the results of Sections 7.1 and 7.2 are used in Sections 7.3–7.7, but otherwise all subsections can be read independently of each other. To simplify the notation we allow q (resp. r) to denote either the numerical rank (resp. numerical nullity) or its upper bound where this causes no confusion.

7.1 Singular spaces and their orthogonal matrix bases

Let us introduce some techniques for the next subsections. For every integer k in the range $1 \le k < \rho$ define the partition $S_A = (S_{k,A} \mid S_{A,m-k})$ and $T_A = (T_{k,A} \mid T_{A,n-k})$ where the submatrices $S_{k,A}$ and $T_{k,A}$ are formed by the first k columns of the matrices S_A and T_A, respectively. Write $\mathbb{S}_{k,A} = \mathcal{R}(S_{k,A})$ and $\mathbb{T}_{k,A} = \mathcal{R}(T_{k,A})$. If $\sigma_k > \sigma_{k+1}$, then $\mathbb{S}_{k,A}$ and $\mathbb{T}_{k,A}$ are the left and right leading singular spaces, respectively, associated with the k largest singular values of the matrix A, whereas their orthogonal complements $\mathbb{S}_{A,m-k} = \mathcal{R}(S_{A,m-k})$ and $\mathbb{T}_{A,n-k} = \mathcal{R}(T_{A,n-k})$ are the left and right trailing singular spaces, respectively, associated with the other singular values of A. The pairs of subscripts $\{k, A\}$ versus $\{A, m-k\}$ and $\{A, n-k\}$ mark the leading versus trailing singular spaces. The left singular spaces of A are the right singular spaces of A^T and vice versa.

Next we specify some expressions for matrices of a rank q that approximate a matrix A within the error norm $\sigma_{q+1}(A)$. We obtain a low-rank approximation where the bound $\sigma_{q+1}(A)$ and the rank q are small in context.

Theorem 7.1. *Suppose A is an $m \times n$ matrix, $S_A \Sigma_A T_A^T$ is its SVD of (2.4), q is a positive integer exceeded by m and n, T and S are orthogonal matrix bases for the spaces $\mathbb{T}_{q,A}$ and $\mathbb{S}_{q,A}$, respectively. Then*

$$||A - ATT^T|| = ||A - SS^T A|| = \sigma_{q+1}(A). \tag{7.1}$$

Proof. Let us first write $P = T_{q,A} T_{q,A}^T$ and $r = n - q$ and estimate the norm $||A - AP||$. We have $AP = S_A \Sigma_A T_A^T T_{q,A} T_{q,A}^T$. Substitute $T_A^T T_{q,A} = \begin{pmatrix} I_q \\ O_{r,q} \end{pmatrix}$ and obtain $AP = S_A \Sigma_A \begin{pmatrix} T_{q,A}^T \\ O_{r,q} \end{pmatrix}$, whereas $A = S_A \Sigma_A \begin{pmatrix} T_{q,A}^T \\ T_{A,r}^T \end{pmatrix}$. Therefore

$$A - AP = S_A \Sigma_A \begin{pmatrix} O_{q,n} \\ T_{A,r}^T \end{pmatrix} = S_A \operatorname{diag}(O_q, \operatorname{diag}(\sigma_j)_{j=q+1}^n) \begin{pmatrix} O_{q,n} \\ T_{A,r}^T \end{pmatrix}.$$

Recall that S_A and $T_{A,r}$ are orthogonal matrices and deduce that $||A - AP|| = ||\operatorname{diag}(\sigma_j)_{j=q+1}^{n}|| = \sigma_{q+1}$. Similarly deduce that $||A - S_{q,A}S_{q,A}^T A|| = \sigma_{q+1}(A)$.

It remains to extend equation (7.1) from the case where $T = T_{q,A}$ and $S = S_{q,A}$ to the case where the matrices T and S are orthogonal, $\mathcal{R}(T) = \mathbb{T}_{q,A} = \mathcal{R}(T_{q,A})$ and $\mathcal{R}(S) = \mathbb{S}_{q,A} = \mathcal{R}(S_{q,A})$. The latter properties imply that $T = T_{q,A}U$ and $S = S_{q,A}V$ for some square orthogonal matrices U and V. Consequently $TT^T = T_{q,A}UU^TT_{q,A}^T = T_{q,A}T_{q,A}^T$ and $SS^T = S_{q,A}VV^TS_{q,A}^T = S_{q,A}S_{q,A}^T$, implying the desired extension. □

7.2 Computation of nmbs and approximation of singular spaces

One can compute nmbs by employing the following theorems.

Theorem 7.2. *[PQ10, Theorem 3.1]. Suppose a matrix $A \in \mathbb{R}^{m\times n}$ has rank ρ, $l = \min\{m, n\}$, $U \in \mathbb{R}^{m\times r}$, $V \in \mathbb{R}^{n\times r}$, $r = l - \rho$, and the matrix $C = A + UV^T$ has full rank l. Then every matrix $C^{(I)}U$ is a $\operatorname{nmb}(A)$ if $m \geq n$, whereas every matrix $(C^{(I)})^TV$ is a $\operatorname{nmb}(A^T)$ if $n \geq m$.*

Theorem 7.3. *Suppose that $A \in \mathbb{R}^{m\times n}$, $U \in \mathbb{R}^{m\times q}$, $V \in \mathbb{R}^{r\times n}$, $W \in \mathbb{R}^{r\times q}$, $K = \begin{pmatrix} W & V \\ U & A \end{pmatrix} \in \mathbb{R}^{(m+r)\times(n+q)}$, $K^{(I)}K = I_{n+q}$ and $W^{(I)}W = I_q$ for some matrices $K^{(I)}$ and $W^{(I)}$. Define the matrix $Y = (O_{n,q} \mid I_n)K^{(I)}\begin{pmatrix} O_{r,q} \\ U \end{pmatrix}$.*

(a) Then we have $\mathcal{N}(A) \subseteq \mathcal{R}(Y)$.

Furthermore suppose $\operatorname{rank} U \leq \nu = \operatorname{nul} A$. Then

(b) $\mathcal{N}(A) = \mathcal{R}(Y)$ and

(c) $\mathcal{N}(A) = \mathcal{R}(Z)$ for $Z = (O_{n,q} \mid I_n)K^{(I)}\begin{pmatrix} O_{m+r-\nu,\nu} \\ I_\nu \end{pmatrix}$.

Proof. See [PQa, Theorem 11.2] on parts (a) and (b). Part (c) immediately follows from part (b). □

For a matrix $A \in \mathbb{R}^{m\times n}$ having a numerical rank q, we can substitute C^+ for $C^{(I)}$, K^+ for $K^{(I)}$, and L^+ for $L^{(I)}$ in the latter theorems and combine them with Theorem 2.2 to approximate the spaces $\mathbb{S}_{A,n-q}$ and $\mathbb{T}_{A,n-q}$. Next we elaborate upon such extension where we substitute C^+ for $C^{(I)}$ into Theorem 7.2. See some alternative techniques in [PQ10] and [PQa].

Theorem 7.4. *Assume a matrix $\tilde{A} \in \mathbb{R}^{m\times n}$ having numerical rank q where $0 < q < l = \min\{m, n\}$. Write $r = l - q$ and suppose $U \in \mathbb{R}^{m\times r}$, $V \in \mathbb{R}^{n\times r}$, and the matrix $\tilde{C} = \tilde{A} + UV^T$ has full rank and is well conditioned. Then there is a scalar c independent of $\tilde{A}$, U, V, m, n and q such that $||\tilde{C}^+UX - T_{\tilde{A},r}|| \leq c\sigma_{q+1}(\tilde{A})$ for some matrix $X \in \mathbb{R}^{r\times r}$ if $m \geq n$, whereas $||\tilde{C}^{+T}VY - S_{\tilde{A},r}|| \leq c\sigma_{q+1}(\tilde{A})$ for some matrix $Y \in \mathbb{R}^{r\times r}$ if $n \geq m$.*

Proof. Cf. [PQ10, Section 7.1]. Set to zero all but the q largest singular values in the SVD of the matrix $\tilde{A}$. Apply Theorem 7.2 to the resulting matrix A and to the matrix $C = A + UV^T$ and obtain that $\mathcal{R}(C^+U) = \mathcal{N}(A) = \mathbb{T}_{\tilde{A},r}$ and $\mathcal{R}(C^{+T}V) = \mathcal{N}(A^T) = \mathbb{S}_{\tilde{A},r}$. Therefore $T_{\tilde{A},r} = C^+UX$ and $S_{\tilde{A},r} = C^{+T}VY$ for some nonsingular $r \times r$ matrices X and Y. Note that $||\tilde{C} - C|| = ||\tilde{A} - A|| = \sigma_{q+1}(A)$. Combine the latter four equations with Theorem 2.2, recall that the matrix $\tilde{C}$ is well conditioned, and obtain Theorem 7.4. □

The theorem shows that additive preprocessing $\tilde{A} \to \tilde{A} + UV^T$ enables us to approximate the trailing singular spaces $\mathbb{S}_{\tilde{A},r}$ and $\mathbb{T}_{\tilde{A},r}$. Next we apply dual additive preprocessing $\tilde{A}^+ \to \tilde{A}^+ + U_-V_-^T$ to approximate the leading singular spaces $\mathbb{S}_{q,\tilde{A}}$ and $\mathbb{T}_{q,\tilde{A}}$.

Theorem 7.5. *Assume that the matrices $\tilde{A} \in \mathbb{R}^{m\times n}$, $U_- \in \mathbb{R}^{n\times q}$, $V_- \in \mathbb{R}^{m\times q}$, $H = I_q + V_-\tilde{A}U_-^T$, and $\tilde{C}_-^+ = \tilde{A}^+ + U_-V_-^T$ have full ranks, the matrix $\tilde{A}^+$ has numerical rank q, and the matrix $\tilde{C}_-^+$ is well conditioned. Then there exists a scalar c_- independent of $\tilde{A}$, U_-, V_-, m, n and q and such that $||\tilde{C}_-U_-X_- - S_{q,\tilde{A}}|| \leq c_-\sigma_{q+1}(\tilde{A})$ for some matrix $X_- \in \mathbb{R}^{q\times q}$ if $m \leq n$, whereas $||\tilde{C}_-^TV_-Y_- - T_{q,\tilde{A}}|| \leq c_-\sigma_{q+1}(\tilde{A})$ for some matrix $Y_- \in \mathbb{R}^{q\times q}$ if $m \geq n$.*

Proof. Rewrite the SVDs $\tilde{A} = S_{\tilde{A}}\Sigma_{\tilde{A}}T_{\tilde{A}}^T$ and $\tilde{A}^+ = T_{\tilde{A}}\Sigma_{\tilde{A}}^+S_{\tilde{A}}^T$ as follows,

$$\tilde{A} = (S_{q,\tilde{A}} \mid S_{\tilde{A},m-q})\,\mathrm{diag}(\Sigma_{q,\tilde{A}}, \bar{\Sigma}_{\tilde{A}})(T_{q,\tilde{A}} \mid T_{\tilde{A},n-q})^T,$$

$$\tilde{A}^+ = (T_{q,\tilde{A}} \mid T_{\tilde{A},n-q})\,\mathrm{diag}(\Sigma_{q,\tilde{A}}^+, \bar{\Sigma}_{\tilde{A}^+})(S_{q,\tilde{A}} \mid S_{\tilde{A},m-q})^T$$

where $\Sigma_{q,\tilde{A}} = (I_q \mid O_{q,n-q})\Sigma_{\tilde{A}}$ denotes the leading $q \times q$ block of the diagonal matrix $\Sigma_{\tilde{A}}$. Apply Theorem 7.4 to the matrix $\tilde{A}^+$ replacing A and observe that $\mathbb{T}_{q,\tilde{A}} = \mathbb{S}_{\tilde{A}^+,q}$ for $m \geq n$, whereas $\mathbb{S}_{q,\tilde{A}} = \mathbb{T}_{\tilde{A}^+,q}$ for $m \leq n$. □

Apply Theorem 7.4 for $\tilde{C}^+U = B$ and $BX \approx T_{\tilde{A},r}$ and apply Theorem 7.5 for $\tilde{C}_-^TV_- = B_-$ and $B_-Y_- \approx T_{q,\tilde{A}}$ to obtain

$$X \approx B^+T_{\tilde{A},r} \text{ and } \tilde{C}^+UX - T_{\tilde{A},r} \approx BB^+T_{\tilde{A},r} - T_{\tilde{A},r}, \tag{7.2}$$

$$Y_- \approx B_-^+T_{q,\tilde{A}} \text{ and } \tilde{C}_-^TV_-Y_- - T_{q,\tilde{A},} \approx B_-B_-^+T_{q,\tilde{A}} - T_{q,\tilde{A}}. \tag{7.3}$$

We employ these relationships in Section 8.4.

By virtue of Corollary 5.1 the matrix $\tilde{C}$ of Theorem 7.4 has full rank with probability 1 and is expected to be well conditioned (as we require in Theorem 7.4) provided $U \in \mathcal{G}^{m\times r}_{0,||\tilde{A}||}$ and $V \in \mathcal{G}^{n\times r}_{0,||\tilde{A}||}$. Likewise the matrix $\tilde{C}_-$ of Theorem 7.5 has full rank with probability 1 and is expected to be well conditioned provided $U_- \in \mathcal{G}^{n\times q}_{0,||\tilde{A}^+||}$ and $V_- \in \mathcal{G}^{m\times q}_{0,||\tilde{A}^+||}$.

Theorem 7.4 shows that $\mathcal{R}(\tilde{C}^+U) \approx \mathbb{T}_{\tilde{A},r}$ if $m \geq n$ and $\mathcal{R}(\tilde{C}^{+T}V) \approx \mathbb{S}_{\tilde{A},r}$ if $m \leq n$, that is, the linear spaces $\mathcal{R}(\tilde{C}^+U)$ for $m \geq n$ and $\mathcal{R}(\tilde{C}^{+T}V)$ for $m \leq n$ approximate the right and left trailing singular spaces associated with the r smallest singular values of the matrix $\tilde{A}$, respectively. (Some of these values can vanish.) Likewise Theorem 7.5 shows that $\mathcal{R}(\tilde{C}_-U_-) \approx \mathbb{S}_{q,\tilde{A}}$ if $m \leq n$ and $\mathcal{R}(\tilde{C}_-^TV_-) \approx \mathbb{T}_{q,\tilde{A}}$ if $m \geq n$, that is, the linear spaces $\mathcal{R}(\tilde{C}_-U_-)$ for $m \leq n$ and $\mathcal{R}(\tilde{C}_-^TV_-)$ for $m \geq n$ approximate the left and right leading singular spaces associated with the q largest singular values of the matrix $\tilde{A}$, respectively.

Alternatively we can approximate the leading singular spaces $\mathbb{S}_{q,\tilde{A}}$ by $\mathcal{R}(\tilde{A}H)$ for $H \in \mathcal{G}^{n\times q}_{0,1}$ and $\mathbb{T}_{q,\tilde{A}}$ by $\mathcal{R}(\tilde{A}^TG)$ for $G \in \mathcal{G}^{m\times q}_{0,1}$ provided that the matrix $\tilde{A}$ has a numerical rank q. Indeed (cf. [HMT11, Section 10.3], [MRT11]) post-multiply the SVD of $\tilde{A}$ by H and obtain $\tilde{A}H = S_{\tilde{A}}\Sigma_{\tilde{A}}T_{\tilde{A}}^TH = S_{\tilde{A}}\Sigma_{\tilde{A}}B_{\tilde{A}}$ where $B_{\tilde{A}} = T_{\tilde{A}}^TH \in \mathcal{G}^{n\times q}_{0,1}$ by virtue of Lemma 4.3. Since $\sigma_{q+1}(\tilde{A}) \ll ||\tilde{A}||$ whereas $||\tilde{A}|| = O(\sigma_q(\tilde{A}))$, we expect to have $\Sigma_{\tilde{A}}B_{\tilde{A}} \approx \mathrm{diag}(\Sigma_{q,\tilde{A}}, O_{m-q,n-q})B_{\tilde{A}}$. Therefore $\tilde{A}H = S_{\tilde{A}}\Sigma_{\tilde{A}}B_{\tilde{A}} \approx S_{q,\tilde{A}}\,\mathrm{diag}(\Sigma_{q,\tilde{A}}, O_{m-q,n-q})B_{\tilde{A}}$, and so $\mathcal{R}(\tilde{A}G) \approx \mathbb{S}_{q,\tilde{A}}$. A similar argument shows that $\mathcal{R}(\tilde{A}^TH) \approx \mathbb{T}_{q,\tilde{A}}$.

Yet another alternative is to apply Theorem 7.3 to the matrix $\bar{K} = \begin{pmatrix} W & V^T \\ -U & A^+ \end{pmatrix} \in \mathbb{R}^{(m+r)\times(n+q)}$. Here we assume that $A^{(I)} \in \mathbb{R}^{m\times n}$ and compute an approximation $Z = (O_{n,q} \mid I_n)\bar{K}^+ \begin{pmatrix} O_{m+r-\nu,\nu} \\ I_\nu \end{pmatrix}$ to a matrix basis for the leading singular space associated with the q largest singular values of the matrix A provided that it has numerical rank q. We can approximate the respective left singular space of this matrix by applying the same technique to the transpose A^T. We call this technique *dual augmentation*.

The computation of the matrices Y and Z involves the matrix $\bar{K}^+$ but not the matrix $\bar{K}$. We can employ the following formula (we omit its straightforward verification),

$$\bar{K}^{(I)} = \begin{pmatrix} I_q & O_{q,n} \\ AU & I_n \end{pmatrix} \mathrm{diag}(S^{(I)}, I_n) \begin{pmatrix} I_r & -V^TA \\ O_{n,r} & A \end{pmatrix} \text{ for } S = W + V^TAU, \tag{7.4}$$

which we should apply for $A^{(I)} = A^+$, $K^{(I)} = K^+$, and $S^{(I)} = S^+$. In this way, similarly to the dual SMW formula, we avoid computing an inverse $A^{(I)}$ and compute just the Moore–Penrose inverse of the Schur complement $S \in \mathbb{R}^{q\times r}$.

Remark 7.1. *Approximation of the trailing and leading singular spaces as well as the computation of numerical rank and numerical nullity (see Section 7.4) are facilitated as the gaps increase between the singular values of the input matrix A. This motivates using the power transforms $A \Longrightarrow B_h = (AA^T)^h A$ for positive integers h because $\sigma_j(B_h) = (\sigma_j(A))^{2h+1}$ for all j.*

Remark 7.2. *To incorporate the power transforms into above algorithms more directly, we can write $M^{(0)} = M$ where M can stand for G, H, U, V, U_- or V_-. Then we can compute the selected sequence $G^{(i)} = Q(\tilde{A}G^{(i-1)})$, $H^{(i)} = Q(\tilde{A}^T H^{(i-1)})$, $U^{(i)} = Q(\tilde{C}^+ U^{(i-1)})$, $V^{(i)} = Q(\tilde{C}^{+T} V^{(i-1)})$, $U_-^{(i)} = Q(\tilde{C}_- U_-^{(i-1)})$, or $V_-^{(i)} = Q(\tilde{C}_-^T V_-^{(i-1)})$ for $i = 1, 2, \dots$, recursively improving the approximations $\mathcal{R}(M^{(i)}) \approx \mathbb{T}_{s,\tilde{A}}$ for $M = G$, $M = U$, $M = V_-$ or $\mathcal{R}(M^{(i)}) \approx \mathbb{S}_{s,\tilde{A}}$ for $M = H$, $M = V$, $M = U_-$.*

Remark 7.3. *In the case where $m = n$ Theorem 7.2 defines both left and right nmbs. We can reduce to this case the nmb computation for a rectangular matrix A in various ways. In particular we can apply the following properties. (a) $\mathcal{N}(A) = \mathcal{N}(A^T A)$, (b) $\mathcal{N}(A) = \mathcal{N}(B^T A)$ if $A, B \in \mathbb{R}^{m \times n}$ and the matrix B has full rank $m \le n$, and (c) $(A \mid O_{n,m-n})\mathbf{u} = \mathbf{0}_m$ if and only if $A\widehat{\mathbf{u}} = \mathbf{0}_m$ provided $m \ge n$ and $\widehat{\mathbf{u}} = (I_n \mid O_{n,m-n})\mathbf{u}$, whereas $(A^T \mid O_{n,m-n})\mathbf{v} = \mathbf{0}_n$ if and only if $\widehat{\mathbf{v}} = \mathbf{0}_n^T$ provided $m < n$ and $\widehat{\mathbf{v}} = (I_m \mid O_{n-m,m})\mathbf{v}$. Furthermore given an $m \times n$ matrix A for $m > n$, we can represent it as the block vector $A = (B_1^T \mid B_2^T \mid \dots B_h^T)^T$ where B_i are $k_i \times n$ blocks for $i = 1, \dots, h$, $\sum_{i=1}^h k_i = m$, and observe that $\mathcal{N}(A) = \cap_{i=1}^h \mathcal{N}(B_i)$, and we can compute the intersection of nmbs by applying [GL96, Theorem 12.4.1]. Clearly these comments can be extended to the tasks of the approximation of the singular spaces of ill conditioned matrices.*

7.3 Approximation by low-rank matrices and by structured matrices

Suppose that a matrix A has a numerical rank q. Then Theorem 7.1 reduces the task of the approximation of this matrix by a rank-q matrix to computing an orthogonal approximation $\tilde{S}$ or $\tilde{T}$ to an orthogonal matrix basis S or T for the singular space $\mathbb{S}_q$ or $\mathbb{T}_{q,A}$, respectively. Indeed let, say $||\tilde{S} - S|| = ||\tilde{S}^T - S^T|| = \delta$. Then we have $||A - \tilde{S}\tilde{S}^T A|| \le \delta + \delta' + \delta''$ where $\delta = ||A - SS^T A|| = \sigma_{q+1}(A)$ (cf. equation (7.1)), $\delta' = ||SS^T A - S\tilde{S}^T A|| \le \delta ||A||$ and $\delta'' = ||S\tilde{S}^T - \tilde{S}\tilde{S}^T A|| \le \delta ||A||$. Therefore we can solve the task of the approximation of the matrix A by a rank-q matrix by applying any of the three recipes from the previous section (that is by employing randomized dual additive preprocessing, randomized dual augmentation or randomized multiplication) to any of the two matrices A and A^T.

The papers [HMT11], [MRT11] use *oversampling*, that is compute the product $M = AG$ for $G \in \mathcal{G}_{0,1}^{n \times q_+}$ and for q_+ a little greater than q (thus allowing the rank of M and the computational cost to grow a little) and then estimate that the probability of obtaining a close approximation of the matrix A by the matrices $M(M^T M)^{-1} M^T A$ and $QQ^T A$ within $(1 + 9\sqrt{q_+ \min m, n})\sigma_{q+1}(A)$ is at least $1 - 3(q_+ - q)^{q - q_+}$. In view of the above, one can compute the products $N = A^T H$ for $H \in \mathcal{G}_{0,1}^{m \times q_+}$ instead, which may have some benefits for rectangular input matrices A.

By applying the same algorithms to a displacement A of a matrix W having a numerical displacement rank q, that is lying near some matrices with displacement rank q, we approximate the matrix W by one of these matrices. For a sample application we can apply such approximations to simplify Newton's structured matrix inversion (see Section 7.9).

7.4 Computation of numerical rank and numerical nullity

Algorithm 7.1. Computing numerical nullity.

INPUT: *A matrix $A \in \mathbb{R}^{m \times n}$ whose norm $||A||$ is neither large nor small, two lower and upper bounds r_- and r_+ on its numerical nullity r such that $r_- < r_+$, e.g., $r_- = 0$ and $r_+ = n$, a small positive tolerance ϵ, and a Subroutine COND that decides whether a given matrix has full rank and is well conditioned.*

OUTPUT: *FAILURE with a probability close to 0 or the numerical nullity r of the matrix A.*

INITIALIZATIONS: *Write $s_- = r_-$ and $s_+ = r_+$.*

COMPUTATIONS:

1. *Fix an integer s in the range $[s_-, s_+]$ and generate two matrices $U, V \in \mathbb{G}_{0,1}^{n\times s}$.*
2. *Compute the matrix $C = A + UV^T$ and apply to it the Subroutine COND to test whether this matrix has full rank and is well conditioned.*
3. *Set $s + 1 \to s_-$ and continue the search for the integer r in the new range $[s_-, s_+]$ if the matrix C fails the test.*
4. *Otherwise stop and output $r = s$ if $s = 0$ or if*

$$\frac{||AC^{-1}U||}{||A||\ ||C^{-1}U||} < \epsilon. \tag{7.5}$$

5. *Otherwise set $s \to s_+$ and continue the search for the integer r in the new range $[s_-, s_+]$.*
6. *Stop and output FAILURE if $s_- = s_+$.*

Correctness of the algorithm follows from Corollary 5.1 and Theorem 7.4. We can combine the algorithm with the power iterations of Remark 7.2.

Remark 7.4. *Instead of randomized additive preprocessing and Corollary 5.1 one can employ randomized augmentation and Theorem 6.3.*

Remark 7.5. *To save some tests, one can apply binary search by selecting s equal to an integer nearest to $0.5(s_- + s_+)$ at Stage 1. If one agrees to fail with a low probability, then one can drop testing bound (7.5) and instead stop and output $r = s$ where the matrix C passes the test for s but fails for $s-1$. One can further save tests by employing the aggregation methods of [PQ10] or [PQa].*

7.5 Block triangulation using approximate trailing singular spaces

Theorems 7.2–7.5 define randomized algorithms that compute nmbs of a rank deficient matrix and approximate trailing and leading singular spaces of its small norm perturbations. Next we extend these algorithms to block triangulation of such perturbations.

Algorithm 7.2. Block triangulation using approximate trailing singular spaces.

INPUT: *A matrix $A \in \mathbb{R}^{n\times n}$ whose norm $||A||$ is neither large nor small, its numerical rank q satisfying $0 < q = n - r < n$, and a Subroutine* LIN·SOLVE *that either solves a linear system of equations if it is nonsingular and well condtioned or outputs FAILURE otherwise.*

OUTPUT: *FAILURE (with a low probability) or two random matrices K_0 and L_0 in $\mathbb{R}^{n\times q}$ and two matrices K_1 and L_1 in $\mathbb{R}^{n\times r}$ such that with a probability near 1 the $q \times q$ block submatrix $W_{00} = K_0^T A L_0$ of the matrix $W = (K_0 \mid K_1)^T A (L_0 \mid L_1) = \begin{pmatrix} W_{00} & W_{01} \\ W_{10} & W_{11} \end{pmatrix}$ is nonsingular, well conditioned, and strongly dominant such that $\sigma_q(W_{00}) \gg \max\{||W_{01}||,\ ||W_{10}||,\ ||W_{11}||\}$.*

COMPUTATIONS:

1. *Generate four matrices $K_0, L_0 \in \mathcal{G}_{0,1}^{n\times q}$; $U, V \in \mathbb{G}_{0,1}^{n\times r}$. Output the matrices K_0 and L_0.*
2. *Compute the matrix $C = A + UV^T$, expected to be nonsingular and well conditioned.*
3. *Apply the Subroutine* LIN·SOLVE *to compute and to output the matrices $K_1 = C^{-T}V$ and $L_1 = C^{-1}U$. Stop and output FAILURE if so does the subroutine.*

The algorithm fails with a low probability by virtue of Theorems 1.1 and 7.4 and the results of Section 7.2. Consequently its correctness is implied by the following theorem.

Theorem 7.6. *For a matrix $A \in \mathbb{R}^{m\times n}$ and $q < l = \min\{n, m\}$, let $\sigma_q(A) \gg \sigma_{q+1}(A)$ and write $r = n - q$ and $\bar{r} = m - q$.*

(a) Suppose $L_0 \in \mathcal{G}_{0,1}^{n\times q}$, $L_1 \in \mathbb{R}^{n\times r}$, $\mathcal{R}(L_1) = \mathbb{T}_{A,r}$, and $||L_1|| = 1$. Then $||AL_1|| \le \sigma_{q+1}(A)$, whereas the condition number $\kappa(AL_0)$ is expected to have order $\sigma_1(A)/\sigma_q(A)$.

(b) Suppose $K_0 \in \mathcal{G}_{0,1}^{q\times m}$, $K_1 \in \mathbb{R}^{\bar{r}\times m}$, $\mathcal{R}(K_1) = \mathbb{S}_{A,\bar{r}}$, and $||K_1|| = 1$. Then $||K_1A|| \le \sigma_{q+1}(A)$, whereas the condition number $\kappa(K_0A)$ is expected to have order $\sigma_1(A)/\sigma_q(A)$.

Proof. To estimate $\kappa(AL_0)$ and $\kappa(K_0A)$ combine Theorems 3.2 for $B = O$ and 4.1. To estimate $||AL_1||$, write SVD $A = S_A\Sigma_A T_A^T$ of (2.4), recall that $T_A^T L_1 = (O_{n,n-r}, T_{A,r})L_1$, and obtain $AL_1 = S\Sigma T_A^T L_1 = S\Sigma(O_{n,n-r}, T_{A,r})L_1$, $||AL_1|| \le ||S\Sigma(O_{n,n-r}, T_{A,r})||\, ||L_1|| = \sigma_{q+1}(A)$. Similarly estimate $||K_1A||$. □

Block factorization

$$W = \begin{pmatrix} I & O \\ W_{10}W_{00}^{-1} & I \end{pmatrix} \begin{pmatrix} W_{00} & W_{01} \\ O & G \end{pmatrix}$$

for $G = W_{11} - W_{10}W_{00}^{-1}W_{01}$ implies that

$$W^{-1} = \begin{pmatrix} W_{00}^{-1} & -W_{00}^{-1}W_{01}G^{-1} \\ O & G^{-1} \end{pmatrix} \begin{pmatrix} I & O \\ -W_{10}W_{00}^{-1} & I \end{pmatrix}$$

and thus reduces the inversion of the matrices W and A and the solution of a linear system $A\mathbf{y} = \mathbf{b}$ to the similar operations with the matrices W_{00} and G of smaller sizes, where the matrix W_{00} is expected to be nonsingular and well conditioned.

Tables 8.11 and 8.12 demonstrate the power of this approach versus the customary algorithms.

Remark 7.6. *We expect to arrive at the matrices W_{01}, W_{10} and W_{11} having small norms. To counter the expected cancellation of the leading digits of the $2rn - r^2$ entries of these matrices, we should compute the matrices K_1 and L_1, their products by the blocks of the matrix A, and the Schur complement G with a high precision p_+. These computations involve $O(n^2r)$ flops, that is just a r/n fraction of order n^3 flops in high precision p_+ required by Gaussian elimination. See [ORO05] and [PGMQ, Section 9]).*

Remark 7.7. *We can modify Algorithm 7.2 by omitting the computation of the matrices K_0 and L_0 at Stage 1 and instead computing them later as approximations to nmbs of the matrices K_1 and L_1, respectively. E.g., we can proceed as follows: first generate random matrices $G, H \in \mathcal{G}_{0,1}^{n\times q}$ and $U, V \in \mathbb{G}_{0,1}^{n\times r}$, all having full rank with probability 1, and then successively compute the matrices $C = A + UV^T$, $K_1 = C^{-T}V$, $L_1 = C^{-1}U$, $K_0 = (I_n - K_1(K_1^TK_1)^{-1}K_1^T)G$, and $L_0 = (I_n - L_1(L_1^TL_1)^{-1}L_1^T)H$. Apart from the well conditioned computation of the matrices $C^{-T}V$ and $C^{-1}U$, this takes $O(n^2r)$ flops, and we can extend our comments in the previous remark.*

7.6 Block triangulation using approximate leading singular spaces

Suppose a square matrix A has a small positive numerical rank q and define a dual variation of Algorithm 7.2 by applying Theorem 7.5. In this case matrix inversions are limited to the $q \times q$ matrices H, $K_0^TK_0$ and $L_0^TL_0$. Alternatively we can employ any other pair of approximate bases for the left and right leading singular spaces. E.g., we can compute them as the products A^TV and AU for $V \in \mathcal{G}_{0,1}^{q\times m}$ and $U \in \mathcal{G}_{0,1}^{n\times r}$ (cf. Section 7.3). In our dual algorithm we assume that the nonsingular input matrix A has been scaled so that the norm $||A^{-1}||$ is neither large nor small. See some recipes for approximation of this norm at the end of Section 5.

Algorithm 7.3. Block triangulation using approximate leading singular spaces.

INPUT: *A nonsingular ill conditioned matrix $A \in \mathbb{R}^{n\times n}$ scaled so that the norm $||A^{-1}||$ is neither large nor small; the numerical rank q of the matrix A such that $0 < q = n - r < n$, and a Subroutine* INVERT *that either inverts a matrix if it is nonsingular and well conditioned or outputs FAILURE otherwise.*

OUTPUT: *FAILURE (with a low probability) or four matrices* $K_0,\ L_0 \in \mathbb{R}^{n\times q}$ *and* $K_1,\ L_1 \in \mathbb{R}^{n\times r}$ *such that*

$$W = \begin{pmatrix} W_{00} & W_{01} \\ W_{10} & W_{11} \end{pmatrix} = (K_0 \mid K_1)^T A (L_0 \mid L_1)$$

and the block submatrix $W_{00} = K_0^T A L_0$ *is expected to be nonsingular, well conditioned, and strongly dominant such that* $\sigma_q(W_{00}) \gg \max\{||W_{01}||,\ ||W_{10}||,\ ||W_{11}||\}$.

COMPUTATIONS**:**

1. *Generate two matrices* U_- *and* V_- *in* $\mathcal{G}_{0,1}^{n\times q}$.
2. *Compute the matrix* $H = I_q + V_- A U_-^T$ *of (2.8).*
3. *Apply the Subroutine* INVERT *to compute the matrix* H^{-1}. *Stop and output FAILURE if so does the subroutine.*
4. *Compute the matrix* $C_- = A - AU_- H^{-1} V_-^T A$ *of (2.9).*
5. *Compute and output the matrices* $K_0 = C_- U_- / ||C_- U_-||$ *and* $L_0 = C_-^T V_- / ||C_-^T V_-||$.
6. *Compute the matrices* $M \approx \mathrm{nmb}(K_0^T)$ *and* $N \approx \mathrm{nmb}(L_0^T)$ *(see our Section 7.2, [PQ10], and [PQa] on the approximation of nmbs).*
7. *Compute and output the matrices* $K_1 = M/||M||$ *and* $L_1 = N/||N||$.

The algorithm fails with a low probability by virtue of Theorems 1.1 and 7.5 and the results of Section 7.2. Complete the correctness proof by extending Theorem 7.6.

Remark 7.8. *We must compute the matrix* H *with high accuracy, but this stage involves only matrix multiplications and a matrix addition (and no inversions) and only requires* $O(qn^2)$ *high precision flops versus order* n^3 *high precision flops in Gaussian elimination. Then again we can apply the algorithms of [ORO05] for accurate computation of sums and products.*

7.7 Randomized additive preconditioning with the SMW recovery

Suppose that we seek the solution $\mathbf{y} = A^{-1}\mathbf{b}$ of a real nonsingular ill conditioned linear system $A\mathbf{y} = \mathbf{b}$ of n equations where we are given a small upper bound r on the numerical nullity of A. Assume that the norm $||A||$ is neither large nor small. Then randomized additive preprocessing $A \Longrightarrow C = A + UV^T$ for $U, V \in \mathcal{G}_{0,1}^{n\times r}$ is expected to produce a well conditioned matrix C (cf. Corollary 5.1). The SMW formula implies that $\mathbf{y} = C^{-1}\mathbf{b} + C^{-1}UG^{-1}V^TC^{-1}\mathbf{b}$ for $G = I_r - V^TC^{-1}U$. Substitute $X_U = C^{-1}U$ and $\mathbf{x_b} = C^{-1}\mathbf{b}$ and obtain

$$\mathbf{y} = \mathbf{x_b} + X_U G^{-1} V^T \mathbf{x_b} \text{ for } G = I_r - V^T X_U. \tag{7.6}$$

This reduces the computation of $\mathbf{y}$ essentially to the solution of the matrix equation $CX = (U \mid \mathbf{b})$ for $X = (X_U \mid x_{\mathbf{b}})$, computing the matrix G, and its inversion. The solution algorithm below incorporates iterative refinement.

Algorithm 7.4. Randomized Solution of a Linear System with Iterative Refinement.

INPUT: *a vector* $\mathbf{b} \in \mathbb{R}^{n\times 1}$, *a nonsingular ill conditioned matrix* $A \in \mathbb{R}^{n\times n}$, *and its numerical nullity* r *(cf. Remark 7.10).*

OUTPUT: $\tilde{\mathbf{y}} \approx A^{-1}\mathbf{b}$.

COMPUTATIONS*:*

1. *Generate two matrices* $U, V \in \mathcal{G}_{0,\sigma}^{n\times r}$.
2. *Compute the matrix* $C = A + UV^T$, *expected to be nonsingular and well conditioned.*

3. *Apply Gaussian elimination (or another direct algorithm) to compute an approximate inverse* $Y \approx C^{-1}$. *(Perform the computations in double precision. Application of the same algorithm to the original ill conditioned linear system* $A\mathbf{y} = \mathbf{b}$ *would require about as many flops but in extended precision.)*

4. *Apply sufficiently many loops of iterative refinement employing the approximate inverse* Y *to compute sufficiently accurate solution* X_U *of the matrix equation* $CX_U = U$. *(High accuracy is required to counter the cancelation of leading bits in the subsequent computation of the Schur complement* $G = I_r - V^T C^{-1} U$.*) Then recover a close approximation to the vector* $\mathbf{y} = A^{-1}\mathbf{b}$ *by applying equation (7.6).*

The algorithm reduces the original task of computations with ill conditioned matrix A to the computations with the well conditioned matrix C and $O(n^2 r)$ additional flops. To handle an ill conditioned input A, we must perform computations with extended precision to counter magnification of rounding errors, but we can confine this stage essentially to computing the Schur complement $G = I_r - V^T C^{-1} U$. This is a small fraction of the computational time of the customary algorithms for a linear system $A\mathbf{y} = \mathbf{b}$ provided the ratio r/n is small and the precision required to handle the ill conditioned matrix A is high.

Let us supply some estimates. To support iterative refienment we must use a precision p exceeding $\log_2 \kappa(C)$; for well conditioned matrices C we can assume that $p > 2\log_2 \kappa(C)$, say. Then order $p - \log_2 \kappa(C)$ new correct bits are produced per an output value by a loop of iterative refinement (see [PGMQ, Section 9]), which is reduced essentially to multiplication of the matrices C and Y by $2r$ vectors, that is to $(4n-2)nr$ flops in a low (e.g., double) precision p. The refinement algorithm outputs order rn values; by using the technique of [ORO05] one can accumulate them with high accuracy as the sums of sufficiently many low precision summands (cf. symbolic lifting in [P11]). Overall with this advanced implementation we only perform $O(rn^2 p_+/p)$ flops in low precision p at Stage 4 of Algorithm 7.4.

For comparison Gaussian elimination uses $\frac{2}{3}n^3 + O(n^2)$ flops in extended precision $p_+ \approx p_{\text{out}} + \log_2 \kappa(A)$ to output the solution to the ill conditioned linear system $A\mathbf{y} = \mathbf{b}$ with a prescribed precision p_{out}. We compute an approximate inverse Y of the well conditioned matrix C at Stage 3 by using $\frac{2}{3}n^3 + O(n^2)$ flops as well, but in the low precision p. The cost of performing Stages 1 and 2 is dominated, and so our progress is significant where $np \gg rp_+$ and p_+ greatly exceeds p.

Remark 7.9. *Given a nonsingular* $n \times n$ *matrix* A *(with* $||A^{-1}|| \approx 1$*) and a small upper bound* q *on its numerical rank, we can define a dual variation of Algorithm 7.4 as follows: generate a pair of matrices* $U_-, V_- \in \mathbb{G}_{0,1}^{n \times q}$ *and then compute the matrices* H *of (2.8) and* C_- *of (2.9) to reduce the solution of a linear system of equations* $A\mathbf{y} = \mathbf{b}$ *to computing the vector* $\mathbf{y} = (C_-^{-1} - U_- V_-^T)\mathbf{b}$ *(cf. (2.8)). The matrix* $H = I_q + V_-^T A U_-$ *must be computed with high accuracy, but this only requires* $O(qn^2)$ *flops in high precision. Furthermore, unlike the computations by means of the SMW formula, we only need matrix multiplications and an addition (and no inversions) at this stage.*

Remark 7.10. *(Cf. Remark 5.1.) There is no point for applying Algorithm 7.4 where the matrix* A *is well conditioned or has numerical nullity exceeding* r. *In the former case the preconditioning is not needed, whereas in the latter case additive preprocessing would produce an ill conditioned matrix* C. *In both cases preprocessing would give no benefits but would involve extraneous computations and additional rounding errors. In the case where* r *is equal to the numerical nullity of* A*, however, these deficiencies are overwhelmed by the benefits where we avoid performing order* n^3 *high precision flops by applying additive preconditioning* $A \Longrightarrow C = A + UV^T$. *In principle one can extend the algorithm to the case where the input integer* r *is an upper bound on the unknown numerical nullity* ν*, but for* $r > \nu$ *the Schur complement* G *is expected to be ill conditioned [PGMQ, Section 7], which would complicate numerical computations.*

Remark 7.11. *One can replace iterative refinement with the Conjugate Gradient or GMRES algorithms (cf. [GL96]). To their advantage they use no approximate inverse, but they are more sensitive to the success of preconditioning. In particular every Conjugate Gradient loop (essentially multiplication of the matrices* C *and* C^T *by two vectors) produces order of* $1/\kappa(C)$ *new correct bits*

per an output value versus $p-\log_2 \kappa(C)$ in iterative refinement. Thus we need stronger upper bounds on $\kappa(C)$ to ensure progress in the presence of rounding errors.

Remark 7.12. *Assume nonsingular ill conditioned $n \times n$ matrix A and let A_- denote the inverse A^{-1} computed with double precision. Then the matrices A_-A and AA_- are much better conditioned than the matrix A according to the experiments of [R90]. Note that both linear systems of equations $A_-A\mathbf{y} = A_-\mathbf{b}$ and $AA_-\mathbf{x} = \mathbf{b}$ where $\mathbf{y} = A_-\mathbf{x}$, are equivalent to the system $A\mathbf{y} = \mathbf{b}$. This empirical techniques is highly interesting itself and probably can be advanced by means of its recursive application. It may also accentuate the preconditioning power of our randomized preprocessing. We can compute the multiplier $A_- = C^{-1}(I_n - UG^{-1}V^TC^{-1})$ by applying the SMW formula, can drop the factor C^{-1} and write either $A_- = I_n - UG^{-1}V^TC^{-1}$ or $A_- = I_n - C^{-1}UG^{-1}V^T$ or can similarly utilize the dual SMW formula of (2.8) and (2.9).*

7.8 Randomized structured preprocessing

Would the preprocessed $n \times n$ matrices $C = A + UV^T$ inherit the structure of an $n \times n$ matrix A where $U, V \in \mathbb{R}^{n\times r}$? For a small value r the adverse impact of adding the matrix UV^T on the structure is small, e.g., the displacement rank increases by $O(r)$ (cf. [P01]).

We can control this impact even for large values r by endowing the matrices U and V with proper structure and relying on ample empirical data that show preconditioning power of such structured additive preprocessing (see Remark 4.1 and Table 8.5). In particular in the case of a nonsingular ill conditioned Toeplitz-like $n \times n$ matrix A having numerical nullity r and norm $||A|| \approx 1$ and given with a displacement generator of a small length d, we can choose a pair of standard Gaussian random Toeplitz $n \times r$ matrices U and V and obtain a displacement generator for the matrix $C = A + UV^T$ having length $d + O(1)$ and expected to be well conditioned. By exploiting the structure we can operate with this matrix in nearly linear arithmetic time, e.g., solve a nonsingular linear system $A\mathbf{y} = \mathbf{b}$ in $O(d^2 n \log^2 n)$ flops, even where r is large (see Theorem 2.8). Randomized augmentation can preserve matrix structure even better and has about the same preconditioning power (see Section 6 and Table 8.17).

Alternative deterministic techniques of homotopy continuation also support inversion of nonsingular Toeplitz and various other structured matrices in nearly linear time (see [P01, Section 6.9], [P07], [P10]).

7.9 Preprocessing for Newton–Toeplitz iteration

Recall Newton's iteration for matrix inversion

$$X_{i+1} = X_i(2I - CX_i), \ i = 0, 1, \dots \tag{7.7}$$

Its ith loop squares the residual $I - CX_i$, that is,

$$I - CX_{i+1} = (I - CX_i)^2 = (I - CX_0)^{2^{i+1}}. \tag{7.8}$$

Therefore

$$||I - CX_{i+1}|| \le ||I - CX_i||^2 = ||I - CX_0||^{2^{i+1}}, \ i = 0, 1, \dots, \tag{7.9}$$

so that the approximations X_i quadratically converge to the inverse C^{-1} right from the start provided that $||I - CX_0|| < 1$. Newton's iteration can be incorporated into our randomized algorithms. E.g., we can use it instead of Gaussian elimination in Algorithm 7.4.

We can ensure that $||I - CX_0|| \le 1 - \frac{2n}{(\kappa(C))^2(1+n)}$ by choosing $X_0 = \frac{2nC^T}{(1+n)||C||_1||C||_\infty}$ [PS91].

Such a map $C \Longrightarrow X_0$ preserves the matrix structure of Toeplitz or Hankel type, but is the structure maintained throughout the iteration? Not automatically. In fact Newton's loop can triple the displacement rank of a matrix X_k. The papers [P92], [P93], and [P93a], however, have proposed to maintain the structure via recursive compression of the displacements (one can also say *recompression*), thus defining *Newton's structured* (e.g., Newton–Toeplitz) iteration. Recall that we can readily recover a Toeplitz-like matrix from its displacement (cf. (2.11)). According to the

compression policy proposed in the papers [P92], [P93], and [P93a], one should periodically set to 0 the smallest singular values of the displacements of the matrices X_i to keep the length of the displacements within a fixed tolerance, equal to or a little exceeding the displacement rank of the input matrix C. The computation of the SVDs of the $n \times d$ generator matrices G and H is not too costly for small ranks d, but one can further decrease the computational cost by applying rank revealing factorizations [GE96], [P00a] or randomized methods of our Section 7.3 instead of computing SVDs.

In [PBRZ99, Sections 7.5.3 and 7.5.4] Newton's iteration is linked to iterative refinement combined with recursive updating of the input matrix, and quadratic convergence is proved for the resulting algorithm. In [BM01] extension of this study has naturally led to the important concept of approximate displacement rank of a matrix.

According to the estimates in [P01], the Newton–Toeplitz iteration converges quadratically right from the start provided $||I - CX_0|| < \frac{1}{(1+||Z_e||+||Z_f||)\kappa(C)}||L^{-1}||$, $||L^{-1}|| \le c_{e,f}n$, L denotes the associated displacement operator $L: \ C \to Z_eC - CZ_f$ for $e \neq f$ or $L: \ C \to C - Z_eCZ_f^T$ for $ef \neq 1$, and $c_{e,f}$ is a constant defined by e and f. Similar bounds can be deduced for other classes of matrices with displacement structure [P01, Section 6.6], [PRW02]. [P01, Chapter 6] and [P10] have further data on the history, variations, and analysis of this approach.

Remark 7.13. *One can concurrently apply Newton–Toeplitz iteration to a number of scaled randomized small rank modifications of the input matrix. As soon as one of these applications produces the inverse, we can recover the inverse of the original matrix via the SMW formula or in case of augmentation via (6.2). Other than that no processor communication or synchronization are requiered.*

8 Numerical Experiments

Our numerical experiments with random general, Hankel, Toeplitz and circulant matrices have been performed in the Graduate Center of the City University of New York on a Dell server with a dual core 1.86 GHz Xeon processor and 2G memory running Windows Server 2003 R2. The test Fortran code was compiled with the GNU gfortran compiler within the Cygwin environment. Random numbers were generated with the random_number intrinsic Fortran function, assuming the uniform probability distribution over the range $\{x : -1 \le x < 1\}$. The tests have been designed by the first author and performed by his coauthors.

8.1 Conditioning tests

We have computed the condition numbers of random general $n \times n$ matrices for $n = 2^k$, $k = 5, 6, \dots$, with entries sampled in the range $[-1, 1)$ as well as complex general, Toeplitz, and circulant matrices whose entries had real and imaginary parts sampled at random in the same range $[-1, 1)$. We have performed 100 tests for each class of inputs, each dimension n, and each nullity r. Tables 8.2–8.4 display the test results. The last four columns of each table display the average (mean), minimum, maximum, and standard deviation of the computed condition numbers of the input matrices, respectively. Namely we have computed the values $\kappa(A) = ||A||\ ||A^{-1}||$ for general, Toeplitz, and circulant matrices A and the values $\kappa_1(A) = ||A||_1\ ||A^{-1}||_1$ for Toeplitz matrices A. We have computed and displayed in Table 8.3 the 1-norms of Toeplitz matrices and their inverses rather than their 2-norms to facilitate the computations in the case of inputs of large sizes. Relationships (2.2) link the 1-norms and 2-norms to one another, but the empirical data in Table 8.1 consistently show even closer links, in all cases of general, Toeplitz, and circulant $n \times n$ matrices A where $n = 32, 64, \dots, 1024$.

8.2 Preconditioning tests

Table 8.5 covers our tests for the preconditioning power of additive preprocessing in [PIMR10]. We have tested the input matrices of the following classes.

1n. *Nonsymmetric matrices of type I with numerical nullity r.* $A = S\Sigma_r T^T$ are $n \times n$ matrices where S and T are $n \times n$ random orthogonal matrices, that is, the factors Q in the QR factorizations

of random real matrices; $\Sigma_r = \mathrm{diag}(\sigma_j)_{j=1}^n$ is the diagonal matrix such that $\sigma_{j+1} \le \sigma_j$ for $j = 1, \dots, n-1$, $\sigma_1 = 1$, the values $\sigma_2, \dots, \sigma_{n-r-1}$ are randomly sampled in the semi-open interval $[0.1, 1)$, $\sigma_{n-r} = 0.1$, $\sigma_j = 10^{-16}$ for $j = n-r+1, \dots, n$, and therefore $\kappa(A) = 10^{16}$ [H02, Section 28.3].

1s. *Symmetric matrices of type I with numerical nullity r.* The same as in part 1n, but for $S = T$.

The matrices of the six other classes were constructed in the form of $\frac{A}{||A||} + \beta I$ where the recipes for defining the matrices A and scalars β are specified below.

2n. *Nonsymmetric matrices of type II with numerical nullity r.* $A = (W \mid WZ)$ where W and Z are random orthogonal matrices of sizes $n \times (n-r)$ and $(n-r) \times r$, respectively.

2s. *Symmetric matrices of type II with numerical nullity r.* $A = WW^T$ where W are random orthogonal matrices of size $n \times (n-r)$.

3n. *Nonsymmetric Toeplitz-like matrices with numerical nullity r.* $A = c(T \mid TS)$ for random Toeplitz matrices T of size $n \times (n-r)$ and S of size $(n-r) \times r$ and for a positive scalar c such that $||A|| \approx 1$.

3s. *Symmetric Toeplitz-like matrices with numerical nullity r.* $A = cTT^T$ for random Toeplitz matrices T of size $n \times (n-r)$ and a positive scalar c such that $||A|| \approx 1$.

4n. *Nonsymmetric Toeplitz matrices with numerical nullity* 1. $A = (a_{i,j})_{i,j=1}^n$ is a Toeplitz $n \times n$ matrix. Its entries $a_{i,j} = a_{i-j}$ are random for $i - j < n - 1$. The entry $a_{n,1}$ is selected to ensure that the last row is linearly expressed through the other rows.

4s. *Symmetric Toeplitz matrices with numerical nullity* 1. $A = (a_{i,j})_{i,j=1}^n$ is a Toeplitz $n \times n$ matrix. Its entries $a_{i,j} = a_{i-j}$ are random for $|i - j| < n - 1$, whereas the entry $a_{1,n} = a_{n,1}$ is a root of the quadratic equation $\det A = 0$. We have repeatedly generated the matrices A until we arrived at the quadratic equation having real roots.

We have set $\beta = 10^{-16}$ for symmetric matrices A in the classes 2s, 3s, and 4s, so that $\kappa(A) = 10^{16} + 1$ in these cases. For nonsymmetric matrices A we have defined the scalar β by an iterative process such that $||A|| \approx 1$ and $10^{-18}||A|| \le \kappa(A) \le 10^{-16}||A||$ [PIMR10, Section 8.2].

Table 8.5 displays the average values of the condition numbers $\kappa(C)$ of the matrices $C = A + UU^T$ over 100,000 tests for the inputs in the above classes, $r = 1, 2, 4, 8$ and $n = 100$. We have defined the additive preprocessor UU^T by a normalized $n \times r$ matrix $U = U/||U||$ where $U^T = (\pm I \mid O_{r,r} \mid \pm I \mid O_{r,r} \mid \dots \mid O_{r,r} \mid \pm I \mid O_{r,s})$, we have chosen the integer s to obtain $n \times r$ matrices U and have chosen the signs for the matrices $\pm I$ at random.

In our further tests the condition numbers of the matrices $C = A + 10^p UV^T$ for $p = -10, -5, 5, 10$ were steadily growing within a factor $10^{|p|}$ as the value $|p|$ was growing. This have showed the importance of proper scaling of the additive preprocessor UV^T.

8.3 Application of random circulant and Householder multipliers

Table 8.6 shows the results of our tests of the solution of a nonsingular well conditioned linear system $A\mathbf{y} = \mathbf{b}$ of n equations whose coefficient matrix had ill conditioned $n/2 \times n/2$ leading principal block for $n = 64, 256, 1024$. We have performed 100 numerical tests for each dimension n and computed the maximum, minimum and average relative residual norms $||A\mathbf{y} - \mathbf{b}||/||\mathbf{b}||$ as well as standard deviation. GENP applied to these systems has output corrupted solutions with residual norms ranging from 10 to 10^8. When we preprocessed the systems with circulant multipliers filled with ± 1 (choosing the n signs $\pm$ at random), the norms decreased to at worst 10^{-7} for all inputs. Table 8.6 also shows further decrease of the norm in a single step of iterative refinement. Table 2 in [PQZa] shows similar results of the tests where the input matrices were chosen similarly but so that all their leading blocks had numerical nullities 0 and 1 and where Householder multipliers $I_n - \mathbf{u}\mathbf{v}^T/\mathbf{u}^T\mathbf{v}$ replaced the circulant multipliers. Here $\mathbf{u}$ and $\mathbf{v}$ denote two vectors filled with integers 1 and -1 under random choice of the signs $+$ and $-$.

8.4 Approximation of the tails and heads of SVDs and low-rank approximation of a matrix

Table 8.7 shows the data from our tests on the approximation of trailing singular spaces of the SVD of an $n \times n$ matrix A having numerical nullity $r = n - q$ and on the approximation of this matrix with a matrix of rank $q = n - r$.

For $n = 64, 128, 256$ and $q = 1, 8, 32$ we have generated $n \times n$ random orthogonal matrices S and T and diagonal matrices $\Sigma = \mathrm{diag}(\sigma_j)_{j=1}^{n}$ such that $\sigma_j = 1/j$, $j = 1, \dots, q$, $\sigma_j = 10^{-10}$, $j = q+1, \dots, n$ (cf. [H02, Section 28.3]). Then we computed the input matrices $A = S_A \Sigma_A T_A^T$, for which $||A|| = 1$ and $\kappa(A) = 10^{10}$. Furthermore we have generated pairs of random $n \times r$ matrices U and V for $r = 1, 8, 32$, scaled them to have $||UV^T|| \approx 1$, and computed the matrices $C = A + UV^T$, $B_r = C^{-1}U$, $Y_r = B_r^+ T_{A,r}$; $B_r Y$, $B_r Y_r - T_{A,r}$, $Q = Q(B_r)$, and $AQQ^T = A - (I_n - QQ^T)A$.

Table 8.7 summarizes the data on the values $\kappa(C)$ and the residual norms $\mathrm{rn}_1 = ||B_r Y_r - T_{A,r}||$ (cf. (7.2)) and $\mathrm{rn}_2 = ||AQQ^T||$ observed in 100 runs of our tests for every pair of n and q.

We have performed similar tests on the approximation of leading singular spaces of the SVDs of the same $n \times n$ matrices A having numerical rank q and numerical nullity $r = n - q$ and on the approximation of this matrix with a matrix of rank q. In some tests we employed dual additive preprocessing to approximate matrix bases for the leading singular spaces $\mathbb{T}_{q,A}$ of the matrices A. We have generated the pairs of $n \times q$ random matrices U_- and V_- for $q = 1, 8, 32$, scaled them to have $||U_- V_-^T|| \approx ||A^{-1}|| = 10^{10}$, and successively computed the matrices $H = I_q + V_-^T A U_-$ of (2.8) and $C_- = A - AU_- H^{-1} V_-^T A$ of (2.9), $B_{q,A} = C_-^T V_-$ (all three with extended precision), $Y_{q,A} = (B_{q,A})^+ T_{q,A}$; $B_{q,A} Y_{q,A}$, $B_{q,A} Y_{q,A} - T_{q,A}$, $Q_{q,A} = Q(B_{q,A})$, and $A - Q_{q,A}(Q_{q,A})^T A$. Table 8.8 summarizes the data on the condition numbers $\kappa(C_-)$ and the residual norms $\mathrm{rn}^{(1)} = ||B_{q,A} Y_{q,A} - T_{q,A}||$ (cf. (7.3)) and $\mathrm{rn}^{(2)} = ||A - q, A(Q_{q,A})^T A||$ obtained in 100 runs of our tests for every pair of n and q.

We have also performed similar tests where we generated random $n \times q$ matrices U (for $q = 1, 8, 32$) and random Toeplitz $n \times q$ matrices $\bar{U}$ (for $q = 8, 32$) and then replaced the above matrix $B_{q,A}$ with the approximate matrix bases $Q_{q,A} = Q(AU)$ and $A\bar{U}$ for the leading singular space $\mathbb{T}_{q,A}$. Tables 8.9 and 8.10 display the results of these tests. In both cases the residual norms are equally small and are about as small as in Tables 8.7 and 8.8.

8.5 Solution of linear systems of equations by means of approximation of trailing and leading singular spaces of the SVDs

For our further tests we have chosen $n = 32, 64$ and $r = 1, 2, 4$ and for every pair $\{n, r\}$ generated 100 instances of vectors $\mathbf{b}$ and matrices A, U, and V as follows.

We have generated (a) random vectors $\mathbf{b}$ of dimension n, (b) the matrices A as the error-free products $S\Sigma T^T$ where S and T were random real orthonormal $n \times n$ matrices (generated with double precision), $\Sigma = \mathrm{diag}(\sigma_j)_{j=1}^{n}$, $\sigma_{n-j} = 10^{-17}$ for $j = 0, 1, \dots, r-1$, and $\sigma_{n-j} = 1/(n-j)$ for $j = r, \dots, n-1$, and (c) random $n \times r$ matrices U and V such that $||A| = ||U|| = ||V|| = 1$. Note that $||A^{-1}|| = 10^{17}$.

For every choice of these matrices we have solved the linear systems $A\mathbf{y} = \mathbf{b}$ by applying Algorithm 7.2. We first generated random $n \times (n-r)$ matrices K_0 and L_0 and then computed the matrices $C = A + UV^T$ (which were always nonsingular and well conditioned in our tests), $K_1 = C^{-T}V$, $L_1 = C^{-1}U$, and $W = (K_0 \mid K_1)^T A (L_0 \mid L_1) = \begin{pmatrix} W_{00} & W_{01} \\ W_{10} & W_{11} \end{pmatrix}$. In all our tests the leading principal $(n-r) \times (n-r)$ block $W_{00} = K_0^T A L_0$ was well conditioned and strongly dominated the three other blocks W_{01}, W_{10}, and W_{11} in the 2×2 block matrix W, as we expected to see by virtue of our analysis in Section 7.5. To solve the linear system $W\mathbf{x} = (K_0 \mid K_1)^T \mathbf{b}$, we first computed the dominated blocks W_{01}, W_{10}, and W_{11} with extended precision, then eliminated the subdiagonal block and computed the solution of the resulting block triangular linear system; the leading block W_{00} of its coefficient matrix was expected to be and consistently turned out to be well conditioned. Finally we computed and output the vector $\mathbf{y} = (L_0 \mid L_1)\mathbf{x}$.

Table 8.11 shows the average (mean) values of the relative residual norms $||A\mathbf{y} - \mathbf{b}||/||\mathbf{b}||$ of the output vectors $\mathbf{y}$ (these values range about 10^{-10}) as well as the minimums, maximums, and standard deviations in these tests.

For the same ill conditioned inputs the Subroutine MLDIVIDE(A,B) for Gaussian elimination from MATLAB has produced corrupted outputs, as can be seen from Table 8.12.

We have also performed similar tests for $n = 32, 64$ and $n \times n$ matrices A and vectors $\mathbf{b}$ generated as before, but for $q = n - r = 1, 2, 4$, and then we computed orthogonal matrices K_0, K_1, L_0 and L_1 by employing dual additive preprocessing and Algorithm 7.3. We first generated and scaled the pairs of random $n \times q$ matrices U_- and V_- such that $||U_-|| \approx ||V_-|| \approx 3 * 10^8$, and so $||U_-||\ ||V_-|| \approx ||A^{-1}|| = 10^{17}$. Then we successively computed the matrices H and C_- (as in Section 8.4), C_-U_- and $C_-^T V_-$ (all with extended precision), $K_0 = Q(C_-U_-)$, $L_0 = Q(C_-^T V_-)$, $K_1 = Q(\mathrm{nmb}(K_0^T))$ and $L_1 = Q(\mathrm{nmb}(L_0^T))$, and continued as in the tests for Table 8.11. We displayed the results in Table 8.13, showing the residual norms of the order 10^{-9} on the average.

Furthermore we have performed similar tests where we first generated random $n \times q$ matrices U and V and then replaced the above matrices K_0 and L_0 by $K_0 = Q(A^T V)$ and $L_0 = Q(AU)$. Table 8.14 displays the results of these tests, showing the residual norms of the order 10^{-25} on the average.

Then again for the same ill conditioned inputs the Subroutine MLDIVIDE(A,B) for Gaussian elimination from MATLAB produced corrupted outputs, similarly to the results in Table 8.12.

8.6 Solution of linear systems of equations by means of the SMW formula

We solved linear systems with the same inputs as in the previous section by applying Algorithm 7.4. Tables 8.15 and 8.16 show the test results for the same inputs as we used for tests of Table 8.11 and 8.13, respectively, except that now we have doubled the matrix size to $n = 64$ and $n = 128$.

We included Table 8.16 for completeness of our comparison of accuracy of various algorithms. For small integers q and large integers r Algorithm 7.4 slows down dramatically (see the analysis in Section 7.7), and it is much more promising to apply the algorithm of Remark 7.9.

8.7 Solution of a real symmetric Toeplitz linear system of equations with randomized augmentation

We have solved 100 real symmetric linear systems of equations $T\mathbf{y} = \mathbf{b}$ for each n where we used vectors $\mathbf{b}$ with random coordinates from the range $[-1, 1)$ and Toeplitz matrices $T = S + 10^{-9} I_n$ for a singular symmetric Toeplitz $n \times n$ matrices S having rank $n - 1$ and nullity 1 and generated according to the recipe in [PQ10, Section 10.1b].

Table 8.17 shows the average CPU time of the solutions by our Algorithm 6.1 and, for comparison, based on the QR factorization and SVD, which we computed by applying the LAPACK procedures DGEQRF and DGESVD, respectively. To solve the auxiliary Toeplitz linear system $K\mathbf{x} = \mathbf{e}_1$ in Algorithm 6.1, we first employed the Toeplitz linear solver of [KV99], [V99], [VBHK01], and [VK98] and then applied iterative refinement with double precision.

The abbreviations "Alg. 6.1", "QR", and "SVD" indicate the respective algorithms. The last two columns of the table display the ratios of these data in the first and the two other columns.

We measured the CPU time with the mclock function by counting cycles. One can convert them into seconds by dividing their number by a constant CLOCKS_PER_SEC, which is 1000 on our platform. We marked the table entries by a "-" where the tests have run too long and were not completed.

We have obtained the solutions $\mathbf{y}$ with the relative residual norms of about 10^{-15} in all three algorithms, which showed that Algorithm 6.1 employing iterative refinement was as reliable as the QR and SVD based solutions but ran much faster.

We refer the reader to [PQZC, Table 3] on similar test results for the solution of ill conditioned homogeneous Toeplitz linear systems.

Table 8.1: The norms of random general, Toeplitz and circulant matrices and of their inverses

matrix A	n	$\|\|A\|\|_1$	$\|\|A\|\|_2$	$\frac{\|\|A\|\|_1}{\|\|A\|\|_2}$	$\|\|A^{-1}\|\|_1$	$\|\|A^{-1}\|\|_2$	$\frac{\|\|A^{-1}\|\|_1}{\|\|A^{-1}\|\|_2}$
General	32	1.9×10^1	1.8×10^1	1.0×10^0	4.0×10^2	2.1×10^2	1.9×10^0
General	64	3.7×10^1	3.7×10^1	1.0×10^0	1.2×10^2	6.2×10^1	2.0×10^0
General	128	7.2×10^1	7.4×10^1	9.8×10^{-1}	3.7×10^2	1.8×10^2	2.1×10^0
General	256	1.4×10^2	1.5×10^2	9.5×10^{-1}	5.4×10^2	2.5×10^2	2.2×10^0
General	512	2.8×10^2	3.0×10^2	9.3×10^{-1}	1.0×10^3	4.1×10^2	2.5×10^0
General	1024	5.4×10^2	5.9×10^2	9.2×10^{-1}	1.1×10^3	4.0×10^2	2.7×10^0
Toeplitz	32	1.8×10^1	1.9×10^1	9.5×10^{-1}	2.2×10^1	1.3×10^1	1.7×10^0
Toeplitz	64	3.4×10^1	3.7×10^1	9.3×10^{-1}	4.6×10^1	2.4×10^1	2.0×10^0
Toeplitz	128	6.8×10^1	7.4×10^1	9.1×10^{-1}	1.0×10^2	4.6×10^1	2.2×10^0
Toeplitz	256	1.3×10^2	1.5×10^2	9.0×10^{-1}	5.7×10^2	2.5×10^2	2.3×10^0
Toeplitz	512	2.6×10^2	3.0×10^2	8.9×10^{-1}	6.9×10^2	2.6×10^2	2.6×10^0
Toeplitz	1024	5.2×10^2	5.9×10^2	8.8×10^{-1}	3.4×10^2	1.4×10^2	2.4×10^0
Circulant	32	1.6×10^1	1.8×10^1	8.7×10^{-1}	9.3×10^0	1.0×10^1	9.2×10^{-1}
Circulant	64	3.2×10^1	3.7×10^1	8.7×10^{-1}	5.8×10^0	6.8×10^0	8.6×10^{-1}
Circulant	128	6.4×10^1	7.4×10^1	8.6×10^{-1}	4.9×10^0	5.7×10^0	8.5×10^{-1}
Circulant	256	1.3×10^2	1.5×10^2	8.7×10^{-1}	4.7×10^0	5.6×10^0	8.4×10^{-1}
Circulant	512	2.6×10^2	3.0×10^2	8.7×10^{-1}	4.5×10^0	5.4×10^0	8.3×10^{-1}
Circulant	1024	5.1×10^2	5.9×10^2	8.7×10^{-1}	5.5×10^0	6.6×10^0	8.3×10^{-1}

9 Related work, our progress, and further study

Preconditioned iterative algorithms for linear systems of equations is a classical subject [A94], [B02], [G97]. The problem of creating inexpensive preconditioners for general use has been around for a long while as well.

On the study of conditioning of random matrices see [D88], [E88], [ES05], [CD05], [SST06], [B11], and the bibliography therein. Estimation of the condition numbers of random structured matrices was stated as a challenge in [SST06]. We provide such estimates for Gaussian random Toeplitz and circulant matrices in Sections 3.4 and 3.5. In particular the estimates show that the expected condition number of a Gaussian random Toeplitz $n \times n$ matrix does not grow exponentially fast as n grows to infinity. This can be surprising in view of [BG05].

Our present study of randomized preconditioning extends and substantially advances the works [PGMQ], [PIMR10], [PQZa], and [PQZC]. In our Sections 7 and 8 we outline and test some new applications of randomized preconditioning, whereas our Theorems 1.1 and 4.1 and Corollaries 4.2, 5.1 and 6.1 provide formal support for such applications. The formal support relies on using Gaussian random matrices, but empirically our algorithms remained as efficient where instead we employed Gaussian Toeplitz random matrices and quite typically where we further decreased the number of random parameters involved, e.g. where we used just circulant matrices filled with ± 1 and limited randomization to the random choice of the signs $\pm$ (see our Section 8.3 and Table 8.6 and [PQZa, Table 2]).

Besides the cited estimates for the condition numbers of Gaussian random Toeplitz and circulant matrices, our technical novelties include randomized multiplicative and additive preconditioning, the proof of their power, new augmentation and dual augmentation techniques, extension of the SMW formula to its dual version with further application to low-rank matrix approximation, randomized approximation of leading and trailing singular spaces, and 2×2 block factorizations of ill conditiooned matrices in Sections 7.5 and 7.6 where both blocks are expected to be better conditioned than the input.

Approximation by low-rank matrices (cf. Section 7.3) and the extensions to tensor decompositions are thriving research area, with numerous applications to matrix and tensor computations. One can partly trace its previous study through the papers [GTZ97], [GT01], [GOS08], [T00], [MMD08],

Table 8.2: The condition numbers $\kappa(A)$ of random matrices A

n	**input**	**min**	**max**	**mean**	**std**
32	real	2.4×10^1	1.8×10^3	2.4×10^2	3.3×10^2
32	complex	2.7×10^1	8.7×10^2	1.1×10^2	1.1×10^2
64	real	4.6×10^1	1.1×10^4	5.0×10^2	1.1×10^3
64	complex	5.2×10^1	4.2×10^3	2.7×10^2	4.6×10^2
128	real	1.0×10^2	2.7×10^4	1.1×10^3	3.0×10^3
128	complex	1.3×10^2	2.5×10^3	3.9×10^2	3.3×10^2
256	real	2.4×10^2	8.4×10^4	3.7×10^3	9.7×10^3
256	complex	2.5×10^2	1.4×10^4	1.0×10^3	1.5×10^3
512	real	3.9×10^2	7.4×10^5	1.8×10^4	8.5×10^4
512	complex	5.7×10^2	3.2×10^4	2.3×10^3	3.5×10^3
1024	real	8.8×10^2	2.3×10^5	8.8×10^3	2.4×10^4
1024	complex	7.2×10^2	1.3×10^5	5.4×10^3	1.4×10^4
2048	real	2.1×10^3	2.0×10^5	1.8×10^4	3.2×10^4
2048	complex	2.3×10^3	5.7×10^4	6.7×10^3	7.2×10^3

Table 8.3: The condition numbers $\kappa_1(A) = \frac{||A||_1}{||A^{-1}||_1}$ of random Toeplitz matrices A

n	**min**	**mean**	**max**	**std**
256	9.1×10^2	9.2×10^3	1.3×10^5	1.8×10^4
512	2.3×10^3	3.0×10^4	2.4×10^5	4.9×10^4
1024	5.6×10^3	7.0×10^4	1.8×10^6	2.0×10^5
2048	1.7×10^4	1.8×10^5	4.2×10^6	5.4×10^5
4096	4.3×10^4	2.7×10^5	1.9×10^6	3.4×10^5
8192	8.8×10^4	1.2×10^6	1.3×10^7	2.2×10^6

[OT09], [HMT11], and the bibliography therein, but much earlier advances in this area appeared in the papers [BCLR79], [B80], [B85], [B86], [BC87], directed to estimating the border rank of matrices and tensors and initially motivated by the design of fast matrix multiplication algorithms. Presently, application of tensor decompositions to the acceleration of matrix computations is a fashionable subject with applications to many important areas of modern computing (see, e.g., [T00], [MMD08], [OT09], [KB09]), but then again its earliest examples appeared in the cited papers on border rank and in [P72]. The latter paper has introduced the technique of trilinear aggregation, a basic ingredient of all subsequent fast algorithms for matrix multiplication with the inputs of both immense sizes (far beyond any practical interest) [P79], [P81], [P84], [CW90], [S10], [VW12] and realistic moderate sizes [P81], [P84], [LPS92], [K04]; historically this work was the first example of the acceleration of fundamental matrix computations by means of tensor decomposition.

We expect that our present paper will motivate further efforts toward development and refinement of the outlined algorithms, as well as further study of randomized augmentation techniques, its link to aggregation processes of [MP80], explored in [PQa], and specification of our methods to ill conditioned matrices that have the displacement or rank structures (cf. [KKM79], [P90], [GKO95], [P01], [VBHK01], [EG99], [VVM07], [VVM08]). Advancing our progress in Section 3.4 toward formal support of randomized Toeplitz preconditioning is another research challenge. Comparison and combination of our novel techniques with each other and various known methods, such as the homotopy continuation methods and Newton's structured iteration (cf. Section 7.9, [PKRK], [P01, Chapter 6], and [P10]) or with the techniques of [R90] (cf. our Remark 7.12) seem to be natural research directions as well.

Table 8.4: The condition numbers $\kappa(A)$ of random circulant matrices A

n	**min**	**mean**	**max**	**std**
256	9.6×10^0	1.1×10^2	3.5×10^3	4.0×10^2
512	1.4×10^1	8.5×10^1	1.1×10^3	1.3×10^2
1024	1.9×10^1	1.0×10^2	5.9×10^2	8.6×10^1
2048	4.2×10^1	1.4×10^2	5.7×10^2	1.0×10^2
4096	6.0×10^1	2.6×10^2	3.5×10^3	4.2×10^2
8192	9.5×10^1	3.0×10^2	1.5×10^3	2.5×10^2
16384	1.2×10^2	4.2×10^2	3.6×10^3	4.5×10^2
32768	2.3×10^2	7.5×10^2	5.6×10^3	7.1×10^2
65536	2.4×10^2	1.0×10^3	1.2×10^4	1.3×10^3
131072	3.9×10^2	1.4×10^3	5.5×10^3	9.0×10^2
262144	6.3×10^2	3.7×10^3	1.1×10^5	1.1×10^4
524288	8.0×10^2	3.2×10^3	3.1×10^4	3.7×10^3
1048576	1.2×10^3	4.8×10^3	3.1×10^4	5.1×10^3

Appendix

A Operations with structured matrices in terms of their displacements

The following simple theorem can be found in [P00] or [P01, Section 1.5].

Theorem A.1. *Assume five matrices A, B, C, M and N and a pair of scalars α and β. Then as long as the matrix sizes are compatible we have*

$$A(\alpha M + dN) - (\alpha M + \beta N)B = \alpha(AM - MB) + \beta(AN - NB), \tag{A.1}$$

$$A^T M^T - B^T M^T = -(BM - MA)^T, \tag{A.2}$$

$$A(MN) - (MN)C = (AM - MB)N + M(BN - NC). \tag{A.3}$$

Furthermore for a nonsingular matrix M we have

$$AM^{-1} - M^{-1}B = -M^{-1}(BM - MA)M^{-1}. \tag{A.4}$$

Corollary A.1. *Under the assumptions of Theorem A.1 we have*

$$G_{A,B}(\alpha M + \beta N) = (\alpha G_{A,B}(M) \mid \beta G_{A,B}(N)), \tag{A.5}$$

$$H_{A,B}(\alpha M + \beta N) = (\alpha H_{A,B}(M) \mid \beta H_{A,B}(N)), \tag{A.6}$$

$$G_{A,B}(M^T) = -H_{B^T,A^T}(M^T),\ H_{A,B}(M^T) = G_{B^T,A^T}(M^T), \tag{A.7}$$

$$G_{A,C}(MN) = (G_{A,B}(M) \mid MG_{B,C}(N)), \tag{A.8}$$

$$H_{A,C}(MN) = (N^T H_{A,B}(M) \mid H_{B,C}(N)), \tag{A.9}$$

$$G_{A,B}(M^{-1}) = -M^{-1}G_{B,A}(M),\ H_{A,B}(M^{-1}) = M^{-T}H_{B,A}(M) \tag{A.10}$$

and consequently

$$d_{A,B}(\alpha M + \beta N) \le d_{A,B}(M) + d_{A,B}(N), \tag{A.11}$$

$$d_{A,B}(M^T) = d_{B^T,A^T}(M), \tag{A.12}$$

$$d_{A,C}(MN) \le d_{A,B}(M) + d_{B,C}(N), \tag{A.13}$$

$$d_{A,B}(M^{-1}) = d_{B,A}(M). \tag{A.14}$$

Table 8.5: Preconditioning tests

Type	r	Cond (C)
1n	1	3.21E+2
1n	2	4.52E+3
1n	4	2.09E+5
1n	8	6.40E+2
1s	1	5.86E+2x
1s	2	1.06E+4
1s	4	1.72E+3
1s	8	5.60E+3
2n	1	8.05E+1
2n	2	6.82E+3
2n	4	2.78E+4
2n	8	3.59E+3
2s	1	1.19E+3
2s	2	1.96E+3
2s	4	1.09E+4
2s	8	9.71E+3
3n	1	2.02E+4
3n	2	1.53E+3
3n	4	6.06E+2
3n	8	5.67E+2
3s	1	2.39E+4
3s	4	1.69E+3
3s	8	6.74E+3
4n	1	4.93E+2
4n	2	4.48E+2
4n	4	2.65E+2
4n	8	1.64E+2
4s	1	1.45E+3
4s	2	5.11E+2
4s	4	7.21E+2
4s	8	2.99E+2

B Uniform random sampling and nonsingularity of random matrices

Uniform random sampling of elements from a finite set Δ is their selection from this set at random, independently of each other and under the uniform probability distribution on the set Δ.

Theorem B.1. *Under the assumptions of Lemma 3.1 let the values of the variables of the polynomial be randomly and uniformly sampled from a finite set Δ. Then the polynomial vanishes with a probability at most $\frac{d}{|\Delta|}$.*

Corollary B.1. *Let the entries of a general or Toeplitz $m \times n$ matrix have been randomly and uniformly sampled from a finite set Δ of cardinality $|\Delta|$ (in any fixed ring). Let $l = \min\{m, n\}$. Then (a) every $k \times k$ submatrix M for $k \le l$ is nonsingular with a probability at least $1 - \frac{k}{|\Delta|}$ and (b) is strongly nonsingular with a probability at least $1 - \sum_{i=1}^{k} \frac{i}{|\Delta|} = 1 - \frac{(k+1)k}{2|\Delta|}$. Furthermore (c) if the submatrix M is indeed nonsingular, then any entry of its inverse is nonzero with a probability at least $1 - \frac{k-1}{|\Delta|}$.*

Table 8.6: Relative residual norms: randomized circulant GENP for well conditioned linear systems with ill conditioned leading blocks (cf. [PQZa, Table 2])

dimension	iterations	min	max	mean	std
64	0	4.7×10^{-14}	8.0×10^{-11}	4.0×10^{-12}	1.1×10^{-11}
64	1	1.9×10^{-15}	5.3×10^{-13}	2.3×10^{-14}	5.4×10^{-14}
256	0	1.7×10^{-12}	1.4×10^{-7}	2.0×10^{-9}	1.5×10^{-8}
256	1	8.3×10^{-15}	4.3×10^{-10}	4.5×10^{-12}	4.3×10^{-11}
1024	0	1.7×10^{-10}	4.4×10^{-9}	1.4×10^{-9}	2.1×10^{-9}
1024	1	3.4×10^{-14}	9.9×10^{-14}	6.8×10^{-14}	2.7×10^{-14}

Proof. The claimed properties of nonsingularity and nonvanishing hold for generic matrices. The singularity of a $k \times k$ matrix means that its determinant vanishes, but the determinant is a polynomial of total degree k in the entries. Therefore Theorem B.1 implies parts (a) and consequently (b). Part (c) follows because a fixed entry of the inverse vanishes if and only if the respective entry of the adjoint vanishes, but up to the sign the latter entry is the determinant of a $(k-1) \times (k-1)$ submatrix of the input matrix M, and so it is a polynomial of degree $k-1$ in its entries. □

C Conditioning of random complex matrices

We have assumed dealing with real random matrices and vectors throughout the paper, but most of our study can be readily extended to the computations in the field $\mathbb{C}$ of complex numbers if we replace the transposes A^T by the Hermitian transposes A^H. All the results of Section 3.2 apply to this case equally well. Below we elaborate upon the respective extension of our probabilistic bounds on the norms and singular values.

Definition C.1. *The set $\mathcal{G}^{m \times n}_{\mathbb{C},\mu,\sigma}$ of complex Gaussian random $m \times n$ matrices with a mean μ and a variance σ is the set $\{A + B\sqrt{-1}\}$ for $(A \mid B) \in \mathcal{G}^{m \times 2n}_{\mu,\sigma}$ (cf. Definition 3.2).*

We can immediately extend Theorem 3.2 to the latter matrices. Let us extend Theorem 3.1. Its original proof in [SST06] relies on the following result.

Lemma C.1. *Suppose y is a positive number; $\mathbf{w} \in \mathbb{R}^{n \times 1}$ is any fixed real unit vector, $||\mathbf{w}|| = 1$, $A \in \mathcal{G}^{n \times n}_{\mu,\sigma}$ and therefore is nonsingular with probability* 1. *Then*

$$\text{Probability}\{||A^{-1}\mathbf{w}|| > 1/y\} \le \sqrt{\frac{2}{\pi}}\frac{y}{\sigma} \text{ for } j = 1, \dots, n.$$

The following lemma and corollary extend Lemmas 3.2 and C.1 to the complex case.

Lemma C.2. *The bound of Lemma 3.2 also holds provided $\mathbf{t} = \mathbf{q} + \mathbf{r}\sqrt{-1}$ is a fixed complex unit vector and $\mathbf{b} = \mathbf{f} + \mathbf{g}\sqrt{-1} \in \mathcal{G}^{n \times 1}_{\mathbb{C},\mu,\sigma}$ is a complex vector such that $\mathbf{f}$, $\mathbf{g}$, $\mathbf{q}$ and $\mathbf{r}$ are real vectors, $||\mathbf{t}|| = 1$, and the vectors $\mathbf{f}$ and $\mathbf{g}$ are in $\mathcal{G}^{n \times 1}_{\mu,\sigma}$.*

Proof. We have $\mathbf{t}^H\mathbf{b} = \mathbf{q}^T\mathbf{f} + \mathbf{r}^T\mathbf{g} + (\mathbf{q}^T\mathbf{g} - \mathbf{r}^T\mathbf{f})\sqrt{-1}$, and so $|\mathbf{t}^H\mathbf{b}|^2 = |\mathbf{q}^T\mathbf{f} + \mathbf{r}^T\mathbf{g}|^2 + |\mathbf{q}^T\mathbf{g} - \mathbf{r}^T\mathbf{f}|^2$. Hence $|\mathbf{t}^H\mathbf{b}| \ge |\mathbf{q}^T\mathbf{f} + \mathbf{r}^T\mathbf{g}| = |\mathbf{u}^T\mathbf{v}|$ where $\mathbf{u}^T = (\mathbf{q}^T \mid \mathbf{r}^T)$ and $\mathbf{v} = (\mathbf{f}^T \mid \mathbf{g}^T)^T$. Note that $\mathbf{v} \in \mathcal{G}^{1 \times 2n}_{\mu,\sigma}$ and $||\mathbf{u}|| = ||\mathbf{t}|| = 1$ and apply Lemma 3.2 to real vectors $\mathbf{u}$ and $\mathbf{v}$ replacing $\mathbf{b}$ and $\mathbf{t}$. □

Corollary C.1. *Suppose y is a positive number and suppose a matrix $A \in \mathcal{G}^{n \times n}_{\mathbb{C},\mu,\sigma}$ and therefore is nonsingular with probability* 1. *Then*

$$\text{Probability}\{||A^{-1}\mathbf{e}_j|| > 1/y\} \le \sqrt{\frac{2}{\pi}}\frac{y}{\sigma} \text{ for } j = 1, \dots, n.$$

Table 8.7: Tails of the SVDs and lower-rank approximations (cf. [PQ10])

r	$\kappa(C)$ or rrn_i	n	**min**	**max**	**mean**	**std**
1	$\kappa(C)$	64	$2.38 \times 10^{+02}$	$1.10 \times 10^{+05}$	$6.25 \times 10^{+03}$	$1.68 \times 10^{+04}$
1	$\kappa(C)$	128	$8.61 \times 10^{+02}$	$7.48 \times 10^{+06}$	$1.32 \times 10^{+05}$	$7.98 \times 10^{+05}$
1	$\kappa(C)$	256	$9.70 \times 10^{+02}$	$3.21 \times 10^{+07}$	$3.58 \times 10^{+05}$	$3.21 \times 10^{+06}$
1	rn_1	64	4.01×10^{-10}	1.50×10^{-07}	5.30×10^{-09}	1.59×10^{-08}
1	rn_1	128	7.71×10^{-10}	5.73×10^{-07}	1.58×10^{-08}	6.18×10^{-08}
1	rn_1	256	7.57×10^{-10}	3.2×10^{-07}	1.69×10^{-08}	5.02×10^{-08}
1	rn_2	64	4.01×10^{-10}	1.50×10^{-07}	5.30×10^{-09}	1.59×10^{-08}
1	rn_2	128	7.71×10^{-10}	5.73×10^{-07}	1.58×10^{-08}	6.18×10^{-08}
1	rn_2	256	7.57×10^{-10}	3.22×10^{-07}	1.69×10^{-08}	5.02×10^{-08}
8	$\kappa(C)$	64	$1.26 \times 10^{+03}$	$1.61 \times 10^{+07}$	$2.68 \times 10^{+05}$	$1.71 \times 10^{+06}$
8	$\kappa(C)$	128	$2.92 \times 10^{+03}$	$3.42 \times 10^{+06}$	$1.58 \times 10^{+05}$	$4.12 \times 10^{+05}$
8	$\kappa(C)$	256	$1.39 \times 10^{+04}$	$8.75 \times 10^{+07}$	$1.12 \times 10^{+06}$	$8.74 \times 10^{+06}$
8	rn_1	64	3.39×10^{-10}	2.27×10^{-06}	2.74×10^{-08}	2.27×10^{-07}
8	rn_1	128	4.53×10^{-10}	1.91×10^{-07}	1.03×10^{-08}	2.79×10^{-08}
8	rn_1	256	8.74×10^{-10}	1.73×10^{-07}	7.86×10^{-09}	1.90×10^{-08}
8	rn_2	64	1.54×10^{-09}	7.59×10^{-06}	8.87×10^{-08}	7.58×10^{-07}
8	rn_2	128	1.82×10^{-09}	7.27×10^{-07}	2.95×10^{-08}	8.57×10^{-08}
8	rn_2	256	2.62×10^{-09}	3.89×10^{-07}	2.27×10^{-08}	5.01×10^{-08}
32	$\kappa(C)$	64	$1.77 \times 10^{+03}$	$9.68 \times 10^{+06}$	$1.58 \times 10^{+05}$	$9.70 \times 10^{+05}$
32	$\kappa(C)$	128	$1.65 \times 10^{+04}$	$6.12 \times 10^{+07}$	$1.02 \times 10^{+06}$	$6.19 \times 10^{+06}$
32	$\kappa(C)$	256	$3.57 \times 10^{+04}$	$2.98 \times 10^{+08}$	$4.12 \times 10^{+06}$	$2.98 \times 10^{+07}$
32	rn_1	64	2.73×10^{-10}	3.29×10^{-08}	2.95×10^{-09}	4.93×10^{-09}
32	rn_1	128	3.94×10^{-10}	1.29×10^{-07}	7.18×10^{-09}	1.64×10^{-08}
32	rn_1	256	6.80×10^{-10}	4.00×10^{-07}	1.16×10^{-08}	4.27×10^{-08}
32	rn_2	64	2.10×10^{-09}	1.49×10^{-07}	1.55×10^{-08}	2.18×10^{-08}
32	rn_2	128	2.79×10^{-09}	3.80×10^{-07}	3.81×10^{-08}	6.57×10^{-08}
32	rn_2	256	5.35×10^{-09}	1.05×10^{-06}	5.70×10^{-08}	1.35×10^{-07}

Proof. In the case of real matrices A the corollary is supported by the argument in the proof of [SST06, Lemma 3.2], which employs the well known estimate that we state as our Lemma 3.2. Now we employ Lemma C.2 instead of this estimate, otherwise keep the same argument as in [SST06], and arrive at Corollary C.1. □

Corollary C.2. *Under the assumptions of Corollary C.1 we have* $||A^{-1}|| \le \sum_{j=1}^{n} X_j$ *where* X_j *are nonnegative random variables such that*

$$\text{Probability}\{X_j > 1/y\} \le \sqrt{\frac{2}{\pi}}\frac{y}{\sigma} \text{ for } j = 1, \dots, n.$$

Proof. Recall that for any $n \times n$ matrix B we have $||B|| = ||B\mathbf{w}||$ for some unit vector $\mathbf{w} = \sum_{j=1}^{n} w_j \mathbf{e}_j$. We have $|w_j| \le ||\mathbf{w}|| = 1$ for all j. Substitute $B = A^{-1}$ and obtain $||A^{-1}|| = ||A^{-1}\mathbf{w}|| = ||\sum_{j=1}^{n} A^{-1} w_j \mathbf{e}_j|| \le \sum_{j=1}^{n} |w_j| \, ||A^{-1}\mathbf{e}_j||$, and so $||A^{-1}|| \le \sum_{j=1}^{n} X_j$ where $X_j = ||A^{-1}\mathbf{e}_j||$ for all j. It remains to combine this bound with Corollary C.1. □

The corollary implies that Probability$\{||A^{-1}|| > 1/y\}$ converges to 0 proportionally to y as $y \to 0$, which can be viewed as an extension of Theorem 3.1 to the case of complex inputs. One can deduce similar extensions of Theorems 3.4–3.6. The resulting estimates are a little weaker than in Section 3.3, being overly pessimistic; actually random complex matrices are a little better conditioned than random real matrices (see [E88], [ES05], [CD05] and our Table 8.2).

Table 8.8: Heads of SVDs and low-rank approximations via dual additive preprocessing

q	$\kappa(C)$ or rrn_i	n	**min**	**max**	**mean**	**std**
1	$\kappa(C)$	64	$1.83 \times 10^{+02}$	$1.26 \times 10^{+06}$	$1.74 \times 10^{+04}$	$1.27 \times 10^{+05}$
1	$\kappa(C)$	128	$6.75 \times 10^{+02}$	$8.76 \times 10^{+05}$	$2.35 \times 10^{+04}$	$9.10 \times 10^{+04}$
1	$\kappa(C)$	256	$4.19 \times 10^{+03}$	$5.82 \times 10^{+05}$	$4.43 \times 10^{+04}$	$8.98 \times 10^{+04}$
1	$\mathrm{rn}^{(1)}$	64	2.43×10^{-10}	3.86×10^{-08}	2.55×10^{-09}	5.43×10^{-09}
1	$\mathrm{rn}^{(1)}$	128	4.36×10^{-10}	1.15×10^{-07}	4.45×10^{-09}	1.24×10^{-08}
1	$\mathrm{rn}^{(1)}$	256	6.40×10^{-10}	3.17×10^{-08}	4.00×10^{-09}	5.16×10^{-09}
1	$\mathrm{rn}^{(2)}$	64	8.30×10^{-10}	3.86×10^{-08}	2.81×10^{-09}	5.35×10^{-09}
1	$\mathrm{rn}^{(2)}$	128	1.21×10^{-9}	1.15×10^{-07}	4.80×10^{-09}	1.23×10^{-08}
1	$\mathrm{rn}^{(2)}$	256	1.72×10^{-9}	3.18×10^{-08}	4.53×10^{-09}	4.97×10^{-09}
8	$\kappa(C)$	64	$1.37 \times 10^{+03}$	$1.87 \times 10^{+06}$	$7.57 \times 10^{+04}$	$2.16 \times 10^{+05}$
8	$\kappa(C)$	128	$3.80 \times 10^{+03}$	$8.64 \times 10^{+06}$	$2.00 \times 10^{+05}$	$8.73 \times 10^{+05}$
8	$\kappa(C)$	256	$2.57 \times 10^{+04}$	$1.54 \times 10^{+07}$	$7.25 \times 10^{+05}$	$2.03 \times 10^{+06}$
8	$\mathrm{rn}^{(1)}$	64	1.87×10^{-9}	4.48×10^{-07}	2.29×10^{-08}	5.20×10^{-08}
8	$\mathrm{rn}^{(1)}$	128	3.04×10^{-09}	3.73×10^{-07}	2.72×10^{-08}	5.83×10^{-08}
8	$\mathrm{rn}^{(1)}$	256	3.78×10^{-09}	2.01×10^{-06}	4.81×10^{-08}	2.02×10^{-07}
8	$\mathrm{rn}^{(2)}$	64	1.30×10^{-09}	2.47×10^{-07}	1.09×10^{-08}	2.70×10^{-08}
8	$\mathrm{rn}^{(2)}$	128	1.85×10^{-09}	1.50×10^{-07}	1.36×10^{-08}	2.75×10^{-08}
8	$\mathrm{rn}^{(2)}$	256	2.19×10^{-09}	1.10×10^{-06}	2.36×10^{-08}	1.10×10^{-07}
32	$\kappa(C)$	64	$3.75 \times 10^{+03}$	$3.25 \times 10^{+07}$	$6.01 \times 10^{+05}$	$3.28 \times 10^{+06}$
32	$\kappa(C)$	128	$2.41 \times 10^{+04}$	$1.09 \times 10^{+08}$	$1.95 \times 10^{+06}$	$1.10 \times 10^{+07}$
32	$\kappa(C)$	256	$1.33 \times 10^{+05}$	$2.11 \times 10^{+10}$	$2.18 \times 10^{+08}$	$2.11 \times 10^{+09}$
32	$\mathrm{rn}^{(1)}$	64	7.78×10^{-09}	1.39×10^{-06}	8.17×10^{-08}	1.94×10^{-07}
32	$\mathrm{rn}^{(1)}$	128	9.81×10^{-09}	2.35×10^{-06}	1.17×10^{-07}	3.05×10^{-07}
32	$\mathrm{rn}^{(1)}$	256	2.05×10^{-08}	3.99×10^{-06}	1.91×10^{-07}	5.06×10^{-07}
32	$\mathrm{rn}^{(2)}$	64	1.84×10^{-09}	2.62×10^{-07}	1.85×10^{-08}	4.09×10^{-08}
32	$\mathrm{rn}^{(2)}$	128	2.47×10^{-09}	6.77×10^{-07}	2.93×10^{-08}	8.38×10^{-08}
32	$\mathrm{rn}^{(2)}$	256	5.05×10^{-09}	8.85×10^{-07}	4.38×10^{-08}	1.14×10^{-07}

D Preconditioning with randomized augmentation: additional estimates

Theorem D.1. *Suppose n and r are two positive integers, $A \in \mathbb{R}^{n\times n}$, $||A|| = 1$, $\rho = \mathrm{rank}(A)$, $A = S_A \Sigma_A T_A^T$ is the SVD of (2.4), $U, V \in \mathcal{G}_{0,1}^{n\times r}$, $W \in \mathbb{R}^{r\times r}$, $||W|| \le 1$, and K denotes the matrix of (6.1). Then the matrix K is singular if $r < n - \rho$. Otherwise it is nonsingular with probability 1. Furthermore if it is nonsingular and if $r = n - \rho$, then the condition number $\kappa(K)$ is expected to have at most order $||A||/\sigma_{n-r}(A)$.*

Proof. We proceed similarly to the proof of Theorem 6.3. Suppose $A = S_A \Sigma_A T_A^T$ is the SVD of (2.4) and write $\bar{K} = \mathrm{diag}(I_r, S_A^T) K \,\mathrm{diag}(I_r, T_A)$. Then

$$\bar{K} = \begin{pmatrix} W & \bar{V}^T \\ -\bar{U} & \Sigma_A \end{pmatrix}$$

where $\sigma_j(K) = \sigma_j(\bar{K})$ for all j and $\bar{U}, \bar{V} \in \mathcal{G}_{0,1}^{n\times r}$. Furthermore write $\Sigma_\rho = \mathrm{diag}(\sigma_j(A))_{j=1}^{\rho}$, $\Sigma_A =$

Table 8.9: Heads of SVDs and low-rank approximation with random multipliers

q	$\kappa(C)$ or rrn_i	n	**min**	**max**	**mean**	**std**
1	$\mathrm{rn}^{(1)}$	64	2.35×10^{-10}	1.32×10^{-07}	3.58×10^{-09}	1.37×10^{-08}
1	$\mathrm{rn}^{(1)}$	128	4.41×10^{-10}	3.28×10^{-08}	3.55×10^{-09}	5.71×10^{-09}
1	$\mathrm{rn}^{(1)}$	256	6.98×10^{-10}	5.57×10^{-08}	5.47×10^{-09}	8.63×10^{-09}
1	$\mathrm{rn}^{(2)}$	64	8.28×10^{-10}	1.32×10^{-07}	3.86×10^{-09}	1.36×10^{-08}
1	$\mathrm{rn}^{(2)}$	128	1.21×10^{-09}	3.28×10^{-08}	3.91×10^{-09}	5.57×10^{-09}
1	$\mathrm{rn}^{(2)}$	256	1.74×10^{-09}	5.58×10^{-08}	5.96×10^{-09}	8.47×10^{-09}
8	$\mathrm{rn}^{(1)}$	128	2.56×10^{-09}	1.16×10^{-06}	4.30×10^{-08}	1.45×10^{-07}
8	$\mathrm{rn}^{(1)}$	256	4.45×10^{-09}	3.32×10^{-07}	3.40×10^{-08}	5.11×10^{-08}
8	$\mathrm{rn}^{(2)}$	64	1.46×10^{-09}	9.56×10^{-08}	5.77×10^{-09}	1.06×10^{-08}
8	$\mathrm{rn}^{(2)}$	128	1.64×10^{-09}	4.32×10^{-07}	1.86×10^{-08}	5.97×10^{-08}
8	$\mathrm{rn}^{(2)}$	256	2.50×10^{-09}	1.56×10^{-07}	1.59×10^{-08}	2.47×10^{-08}
32	$\mathrm{rn}^{(1)}$	64	6.80×10^{-09}	2.83×10^{-06}	1.01×10^{-07}	3.73×10^{-07}
32	$\mathrm{rn}^{(1)}$	128	1.25×10^{-08}	6.77×10^{-06}	1.28×10^{-07}	6.76×10^{-07}
32	$\mathrm{rn}^{(1)}$	256	1.85×10^{-08}	1.12×10^{-06}	1.02×10^{-07}	1.54×10^{-07}
32	$\mathrm{rn}^{(2)}$	64	1.84×10^{-09}	6.50×10^{-07}	2.30×10^{-08}	8.28×10^{-08}
32	$\mathrm{rn}^{(2)}$	128	3.11×10^{-09}	1.45×10^{-06}	2.87×10^{-08}	1.45×10^{-07}
32	$\mathrm{rn}^{(2)}$	256	4.39×10^{-09}	2.16×10^{-07}	2.37×10^{-08}	3.34×10^{-08}

$\mathrm{diag}(\Sigma_\rho, O_{n-\rho,n-\rho})$, $\bar{U} = \begin{pmatrix} U_0 \\ U_1 \end{pmatrix}$ and $\bar{V} = \begin{pmatrix} V_0 \\ V_1 \end{pmatrix}$ and obtain

$$\bar{K} = \begin{pmatrix} W & V_0^T & V_1^T \\ U_0 & \Sigma_\rho & O_{\rho,n-\rho} \\ U_1 & O_{\rho,n-\rho} & O_{n-\rho,n-\rho} \end{pmatrix}$$

where $U_0, V_0 \in \mathcal{G}_{0,1}^{\rho \times r}$ and $U_1, V_1 \in \mathcal{G}_{0,1}^{(n-\rho) \times r}$. Now we can readily verify the claims about $\mathrm{rank}(K)$.

It remains to estimate the condition number $\kappa(K) = ||K||\ ||K^{-1}|| = ||\bar{K}||\ ||\bar{K}^{-1}||$ provided that $r = n - \rho$ and the matrix K is nonsingular. To bound the norm $||K||$, note that $K = \begin{pmatrix} W & O_{r,n} \\ O_{n,r} & A \end{pmatrix} + \begin{pmatrix} O_{r,n} & I_{r,r} \\ I_{n,n} & O_{n,r} \end{pmatrix} \begin{pmatrix} -U & O_{n,n} \\ O_{r,r} & V^T \end{pmatrix}$, recall that $||W|| \le ||A|| = 1$, apply bound (2.6) and obtain

$$||K|| \le 1 + \max\{||U||, ||V||\}.$$

By virtue of randomized bounds of Theorem 3.2 we expect to have the norms $||U||$ and $||V||$ in $O(1)$, that is bounded by a constant, and so $||K||$ is in $O(1)$.

We conclude the proof by estimating the norm $||\bar{K}|| = ||K||$. We readily verify that

$$\bar{K}^{-1} = \begin{pmatrix} O_{r,r} & O_{r,n-r} & -U_1^{-1} \\ O_{n-r,r} & \Sigma_{n-r}^{-1} & -\Sigma_{n-r}^{-1} U_0 U_1^{-1} \\ V_1^{-T} & -V_1^{-T} V_0^T \Sigma_{n-r}^{-1} & V_1^{-T}(W + V_0^T \Sigma_{n-r}^{-1} U_0) U_1^{-1} \end{pmatrix}.$$

Apply bound (2.6) and deduce that

$$||K^{-1}|| = ||\bar{K}^{-1}|| \le N_1 + N_2 + N_3$$

where $N_1 = \max\{||V_1^{-T}||, ||\Sigma_{n-r}^{-1}||, ||U_1^{-1}||\}$, $N_2 = \max\{||V_1^{-T} V_0^T \Sigma_{n-r}^{-1}||, ||\Sigma_{n-r}^{-1} U_0 U_1^{-1}||\}$ and $N_3 = ||V_1^{-T}(W + V_0^T \Sigma_{n-r}^{-1} U_0) U_1^{-1}||$. Recall that $||W|| \le 1$, $||\Sigma_{n-r}^{-1}|| = 1/\sigma_{n-r}(A)$, $||V_0^T|| = ||V_0||$, and $||V_1^{-T}|| = ||V_1^{-1}||$ and deduce that

$$N_1 = \max\{||V_1^{-1}||, ||U_1^{-1}||, 1/\sigma_{n-r}(A)\},$$

Table 8.10: Heads of SVDs and low-rank approximations with random Toeplitz multipliers

q	$\mathrm{rrn}^{(i)}$	n	**min**	**max**	**mean**	**std**
8	$\mathrm{rrn}^{(1)}$	64	2.22×10^{-09}	7.89×10^{-06}	1.43×10^{-07}	9.17×10^{-07}
8	$\mathrm{rrn}^{(1)}$	128	3.79×10^{-09}	4.39×10^{-05}	4.87×10^{-07}	4.39×10^{-06}
8	$\mathrm{rrn}^{(1)}$	256	5.33×10^{-09}	3.06×10^{-06}	6.65×10^{-08}	3.12×10^{-07}
8	$\mathrm{rrn}^{(2)}$	64	1.13×10^{-09}	3.66×10^{-06}	6.37×10^{-08}	4.11×10^{-07}
8	$\mathrm{rrn}^{(2)}$	128	1.81×10^{-09}	1.67×10^{-05}	1.90×10^{-07}	1.67×10^{-06}
8	$\mathrm{rrn}^{(2)}$	256	2.96×10^{-09}	1.25×10^{-06}	2.92×10^{-08}	1.28×10^{-07}
32	$\mathrm{rrn}^{(1)}$	64	6.22×10^{-09}	5.00×10^{-07}	4.06×10^{-08}	6.04×10^{-08}
32	$\mathrm{rrn}^{(1)}$	128	2.73×10^{-08}	4.88×10^{-06}	2.57×10^{-07}	8.16×10^{-07}
32	$\mathrm{rrn}^{(1)}$	256	1.78×10^{-08}	1.25×10^{-06}	1.18×10^{-07}	2.03×10^{-07}
32	$\mathrm{rrn}^{(2)}$	64	1.64×10^{-09}	1.26×10^{-07}	9.66×10^{-09}	1.48×10^{-08}
32	$\mathrm{rrn}^{(2)}$	128	5.71×10^{-09}	9.90×10^{-07}	5.50×10^{-08}	1.68×10^{-07}
32	$\mathrm{rrn}^{(2)}$	256	4.02×10^{-09}	2.85×10^{-07}	2.74×10^{-08}	4.48×10^{-08}

Table 8.11: Relative residual norms: ill conditioned linear systems via nmb approximation and block triangulation

n	**r**	**min**	**max**	**mean**	**std**
32	1	1.49×10^{-13}	1.36×10^{-9}	4.25×10^{-11}	1.56×10^{-10}
32	2	3.70×10^{-13}	2.13×10^{-8}	3.83×10^{-10}	2.35×10^{-9}
32	4	9.33×10^{-13}	1.08×10^{-8}	3.37×10^{-10}	1.26×10^{-9}
64	1	1.11×10^{-12}	6.87×10^{-9}	2.03×10^{-10}	7.49×10^{-10}
64	2	1.53×10^{-12}	1.21×10^{-8}	5.86×10^{-10}	1.77×10^{-9}
64	4	2.21×10^{-12}	1.27×10^{-7}	1.69×10^{-9}	1.28×10^{-8}

$$N_2 \le \max\{||V_1^{-1}||\ ||V_0||, ||U_0||\ ||U_1^{-1}||\}/\sigma_{n-r}(A),$$

$$N_3 \le ||V_1^{-1}||\ ||U_1^{-1}||(1 + ||V_0||||U_0||/\sigma_{n-r}(A)).$$

Apply Theorems 3.1 and 3.2 to estimate the norms $||U||$, $||V||$, $||V_0||$, $||U_0||$, $||U_1^{-1}||$ and $||V_1^{-1}||$. Combine all the above bounds to estimate the norm $||K^{-1}||$. □

Corollary D.1. *Suppose n and r are two positive integers, $A \in \mathbb{R}^{n\times n}$, $||A|| = 1$, the matrix A has numerical rank $n - r$, $U, V \in \mathcal{G}_{0,1}^{n\times r}$, $W \in \mathbb{R}^{r\times r}$, $||W|| \le 1$, and the matrix K is defined by (6.1). Then this matrix has full rank with probability 1 and is expected to have condition number $\kappa(K)$ of at most order $||A||/\sigma_{l-r}(A) = 1/\sigma_{n-r}(A)$.*

Proof. We immediately obtain the corollary by combining Theorems D.1 and 2.1. □

Let us extend our analysis to the case of rectangular matrices $A \in \mathbb{R}^{m\times n}$.

Theorem D.2. *Suppose m, n, and r are three positive integers, $l = \min\{m, n\}$, $A \in \mathbb{R}^{m\times n}$, $||A|| = 1$, the matrix A has numerical rank $l - r$, $U \in \mathcal{G}_{0,1}^{m\times r}$, $V \in \mathcal{G}_{0,1}^{n\times r}$, $W \in \mathbb{R}^{r\times r}$, $||W|| \le 1$, K is the $(m+r\times(n+r)$ matrix defined by equation (6.1). Then this matrix has full rank with probability 1 and is expected to have condition number $\kappa(K)$ of at most order $||A||/\sigma_{l-r}(A) = 1/\sigma_{l-r}(A)$.*

Proof. Suppose $A = S_A\Sigma_A T_A^T$ is the SVD of (2.4) and write $\bar{K} = \mathrm{diag}(I_r, S_A^T)K\,\mathrm{diag}(I_r, T_A)$, as in the proof of Theorems 6.3 and D.1. Then write

$$\widehat{K} = I_{m+r,n+r}\bar{K}I_{n+r,m+r} \tag{D.1}$$

Table 8.12: Relative residual norms: ill conditioned linear systems with MLDIVIDE(A,B)

n	**r**	**min**	**max**	**mean**	**std**
32	1	6.34×10^{-3}	7.44×10^{1}	1.74×10^{0}	7.53×10^{0}
32	2	2.03×10^{-2}	1.32×10^{1}	9.19×10^{-1}	1.62×10^{0}
32	4	4.57×10^{-2}	1.36×10^{1}	1.14×10^{0}	1.93×10^{0}
64	1	3.82×10^{-3}	9.93×10^{0}	1.03×10^{0}	1.66×10^{0}
64	2	1.96×10^{-2}	1.27×10^{2}	3.09×10^{0}	1.40×10^{1}
64	4	7.13×10^{-3}	6.63×10^{0}	8.23×10^{-1}	1.20×10^{0}

Table 8.13: Relative residual norms: ill conditioned linear systems with dual additive preprocessing and block triangulation

n	q	**min**	**max**	**mean**	**std**
32	1	2.33×10^{-14}	2.28×10^{-06}	2.31×10^{-08}	2.28×10^{-07}
32	2	3.40×10^{-13}	4.93×10^{-08}	9.11×10^{-10}	5.71×10^{-09}
32	4	5.97×10^{-13}	1.63×10^{-07}	2.22×10^{-09}	1.64×10^{-08}
64	1	3.90×10^{-14}	2.78×10^{-05}	2.81×10^{-07}	2.78×10^{-06}
64	2	3.53×10^{-13}	3.76×10^{-08}	1.13×10^{-09}	4.72×10^{-09}
64	4	3.54×10^{-12}	2.53×10^{-07}	5.19×10^{-09}	2.83×10^{-08}

for $I_{g,h}$ of (2.1). Observe that $\kappa(K) = \kappa(\bar{K})$,

$\widehat{K} = \begin{pmatrix} \bar{K} \\ \widehat{U} \end{pmatrix}$, $\widehat{U} = (U_2 \mid O_{m-n,n})$, $U_2 \in \mathcal{G}_{0,1}^{(m-n)\times r}$ if $m \geq l = n$, whereas

$\widehat{K} = (\bar{K} \mid \widehat{V}^T)$, $\widehat{V}^T = (V_2^T \mid O_{n-m,m})$, $V_2 \in \mathcal{G}_{0,1}^{(n-m)\times r}$ if $n \geq l = m$.

It follows that $\sigma_l(\bar{K}) \geq \sigma_l(\widehat{K})$ (cf. Fact 2.3) and $||\widehat{K}|| \leq ||\bar{K}|| + ||F||$ for $F = U_2$ or $F = V_2$. In both cases $F \in \mathcal{G}_{0,1}^{|n-m|\times r}$, and so we can expect that $||\bar{K}|| = O(||\widehat{K}||)$ because $||\widehat{K}|| \geq ||\Sigma_A|| = ||A|| = 1$. Corollary D.1 implies that with probability 1 the $(l+r)\times(l+r)$ matrix $\widehat{K}$ is nonsingular, and so $\mathrm{rank}(K) = \mathrm{rank}(\bar{K}) = \mathrm{rank}(\widehat{K}) = l + r$, that is the matrix K has full rank.

Furthermore Corollary D.1 implies that the condition number $\kappa(\widehat{K})$ is expected to have order $||A||/\sigma_{n-r}(A)$. It remains to extend this bound to $\kappa(K)$. Recall that $\kappa(\widehat{K}) = ||\widehat{K}||/\sigma_l(\widehat{K})$ and $\kappa(K) = \kappa(\bar{K}) = ||\bar{K}||/\sigma_l(\bar{K})$ and combine the above equations with the bounds $\sigma_l(\bar{K}) \geq \sigma_l(\widehat{K})$ and $||\bar{K}|| = O(||\widehat{K}||)$. □

Acknowledgement: Our research has been supported by NSF Grant CCF-1116736 and PSC CUNY Awards 63153–0041 and 64512–0042. We are also very grateful to Professor Dr. Siegfried M. Rump and Mr. Jesse Wolf for helpful comments.

References

[A94] O. Axelsson, *Iterative Solution Methods*, Cambridge Univ. Press, England, 1994.

[B80] D. Bini, Border Rank of $p \times q \times 2$ Tensors and the Optimal Approximation of a Pair of Bilinear Forms, in *Lecture Notes in Computer Science*, **85**, 98–108, Springer, 1980.

[B85] D. Bini, Tensor and Border Rank of Certain Classes of Matrices and the Fast Evaluation of Determinant, Inverse Matrix and Eigenvalues, *Calcolo*, **22**, 209–228, 1985.

[B86] D. Bini, Border Rank of $m \times n \times (mn-q)$ Tensors, *Linear Algebra and Its Applications*, **79**, 45–51, 1986.

Table 8.14: Relative residual norms: ill conditioned linear system with random multipliers and block triangulation

n	q	**min**	**max**	**mean**	**std**
32	1	7.08×10^{-30}	4.00×10^{-23}	4.52×10^{-25}	4.01×10^{-24}
32	2	7.49×10^{-30}	2.29×10^{-21}	2.77×10^{-23}	2.33×10^{-22}
32	4	1.46×10^{-28}	1.63×10^{-07}	4.83×10^{-25}	2.73×10^{-24}
64	1	1.13×10^{-29}	1.01×10^{-24}	2.31×10^{-26}	1.11×10^{-25}
64	2	6.60×10^{-29}	6.90×10^{-24}	1.45×10^{-25}	7.73×10^{-25}
64	4	2.60×10^{-28}	1.41×10^{-21}	1.61×10^{-23}	1.42×10^{-22}

Table 8.15: Relative residual norms: ill conditioned linear systems with the SMW formula; the case of small numerical nullity

n	**r**	**min**	**max**	**mean**	**std**
64	1	1.18×10^{-15}	6.30×10^{-13}	2.37×10^{-14}	7.45×10^{-14}
64	2	3.42×10^{-15}	1.94×10^{-10}	2.15×10^{-12}	1.94×10^{-11}
64	4	6.66×10^{-15}	1.25×10^{-10}	1.82×10^{-12}	1.25×10^{-11}
128	1	5.79×10^{-15}	4.85×10^{-12}	1.21×10^{-13}	4.96×10^{-13}
128	2	1.45×10^{-14}	1.85×10^{-11}	5.23×10^{-13}	1.88×10^{-12}
128	4	8.41×10^{-14}	4.75×10^{-11}	2.89×10^{-12}	5.95×10^{-12}

[B02] M. Benzi, Preconditioning Techniques for Large Linear Systems: a Survey, *J. of Computational Physics*, **182**, 418–477, 2002.

[B11] C. Beltrán, Estimates on the Condition Number of Random, Rank-deficient Matrices, *IMA Journal of Numerical Analysis*, **31, 1**, 25–39, 2011.

[BC87] D. Bini, M. Capovani, Tensor Rank and Border Rank of Band Toeplitz Matrices, *SIAM J. on Computing*, **2**, 252–258, 1987.

[BCLR79] D. Bini, M. Capovani, G. Lotti, F. Romani, $O(n^{2.7799})$ Complexity for $n \times n$ Approximate Matrix Multiplication, *Information Processing Letters*, **8**, 234–235, 1979.

[BG05] A. Böttcher, S. M. Grudsky, *Spectral Properties of Banded Toeplitz Matrices*, SIAM Publications, Philadelphia, 2005.

[BM01] D. A. Bini, B. Meini, Approximate Displacement Rank and Applications, in *AMS Conference "Structured Matrices in Operator Theory, Control, Signal and Image Processing"*, Boulder, 1999 (edited by V. Olshevsky), *American Math. Society*, 215–232, Providence, RI, 2001.

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

[CD05] Z. Chen, J. J. Dongarra, Condition Numbers of Gaussian Random Matrices, *SIAM. J. on Matrix Analysis and Applications*, **27**, 603–620, 2005.

[CDG03] B. Carpentieri, I. S. Duff, L. Giraud, A Class of Spectral Two-level Preconditioners, *SIAM J. Scientific Computing*, **25, 2**, 749–765, 2003.

[CPW74] R. E. Cline, R. J. Plemmons, and G. Worm, Generalized Inverses of Certain Toeplitz Matrices, *Linear Algebra and Its Applications,* **8**, 25–33, 1974.

Table 8.16: Relative residual norms: ill conditioned linear systems with the SMW formula; the case of small numerical rank

n	**q**	**min**	**max**	**mean**	**std**
64	1	6.54×10^{-14}	4.68×10^{-11}	3.29×10^{-12}	8.29×10^{-12}
64	2	1.59×10^{-13}	9.20×10^{-11}	5.77×10^{-12}	1.43×10^{-11}
64	4	3.16×10^{-13}	8.61×10^{-11}	1.22×10^{-11}	2.35×10^{-11}
128	1	1.01×10^{-12}	1.02×10^{-10}	1.86×10^{-11}	2.65×10^{-11}
128	2	1.63×10^{-12}	9.18×10^{-11}	2.40×10^{-11}	2.84×10^{-11}
128	4	3.99×10^{-12}	9.12×10^{-11}	3.91×10^{-11}	2.80×10^{-11}

Table 8.17: The CPU time (in cycles) for solving an ill conditioned real symmetric Toeplitz linear system

n	**Alg. 6.1**	**QR**	**SVD**	**QR/Alg. 6.1**	**SVD/Alg. 6.1**
512	56.3	148.4	4134.8	2.6	73.5
1024	120.6	1533.5	70293.1	12.7	582.7
2048	265.0	11728.1	–	44.3	–
4096	589.4	–	–	–	–
8192	1304.8	–	–	–	–

[CW90] Coppersmith, S. Winograd, Matrix Multiplicaton via Arithmetic Progressions. *J. Symbolic Comput.*, **9, 3**, 251–280, 1990.

[D83] J. D. Dixon, Estimating Extremal Eigenvalues and Condition Numbers of Matrices, *SIAM J. on Numerical Analysis*, **20, 4**, 812–814, 1983.

[D88] J. Demmel, The Probability That a Numerical Analysis Problem Is Difficult, *Math. of Computation*, **50**, 449–480, 1988.

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

[DS01] K. R. Davidson, S. J. Szarek, Local Operator Theory, Random Matrices, and Banach Spaces, in *Handbook on the Geometry of Banach Spaces* (W. B. Johnson and J. Lindenstrauss editors), pages 317–368, North Holland, Amsterdam, 2001.

[E88] A. Edelman, Eigenvalues and Condition Numbers of Random Matrices, *SIAM J. on Matrix Analysis and Applications*, **9**, **4**, 543–560, 1988.

[EG99] Y. Eidelman, I. Gohberg, On a New Class of Structured Matrices, *Integral Equations and Operator Theory*, **34**, 293–324, Birkhäuser, Basel, 1999.

[ES05] A. Edelman, B. D. Sutton, Tails of Condition Number Distributions, *SIAM J. on Matrix Analysis and Applications*, **27**, **2**, 547–560, 2005.

[G97] A. Greenbaum, *Iterative Methods for Solving Linear Systems*, SIAM, Philadelphia, 1997.

[GE96] M. Gu, S. C. Eisenstat, Efficient algorithms for computing a strong rank-revealing QR factorization, *SIAM Journal on Scientific Computing*, **17**, 848–869, 1996.

[GK72] I. Gohberg, N. Y. Krupnick, A Formula for the Inversion of Finite Toeplitz Matrices, *Matematicheskiie Issledovaniia* (in Russian), **7**, **2**, 272–283, 1972.

[GKO95] I. Gohberg, T. Kailath, V. Olshevsky, Fast Gaussian Elimination with Partial Pivoting for Matrices with Displacement Structure, *Math. of Comp.,* **64(212)**, 1557–1576, 1995.

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

[GOS08] S. A. Goreinov, I. V. Oseledets, D. V. Savostyanov, E. E. Tyrtyshnikov, N. L. Zamarashkin, How to Find a Good Submatrix, Research Report 08-10, ICM HKBU, Kowloon Tong, Hong Kong, 2008.

[GS72] I. Gohberg, A. Sementsul, On the Inversion of Finite Toeplitz Matrices and Their Continuous Analogs, *Matematicheskiie Issledovaniia* (in Russian), **7**, **2**, 187–224, 1972.

[GT01] S. A. Goreinov, E. E. Tyrtyshnikov, The Maximal-volume Concept in Approximation by Low-rank Matrices, *Contemporary Mathematics*, **208**, 47–51, 2001.

[GTZ97] S. A. Goreinov, E. E. Tyrtyshnikov, N. L. Zamarashkin, A Theory of Pseudo-skeleton Approximations, *Linear Algebra and Its Applications*, **261**, 1–22, 1997.

[H02] N. J. Higham, *Accuracy and Stability in Numerical Analysis*, SIAM, Philadelphia, 2002 (second edition).

[HMT11] N. Halko, P. G. Martinsson, J. A. Tropp, Finding Structure with Randomness: Probabilistic Algorithms for Constructing Approximate Matrix Decompositions, *SIAM Review*, **53, 2**, 217–288, 2011.

[K04] I. Kaporin, The Aggregation and Cancellation Techniques as a Practical Tool for Faster Matrix Multiplication, *Theoretical Computer Science*, **315**, **2–3**, 469–510, 2004.

[KB09] T. G. Kolda, B. W. Bader, Tensor Decompositions and Applications, *SIAM Review*, **51, 3**, 455–500, 2009.

[KKM79] T. Kailath, S. Y. Kung, M. Morf, Displacement Ranks of Matrices and Linear Equations, *Journal Math. Analysis and Appls*, **68, 2**, 395–407, 1979.

[KL94] C. S. Kenney, A. J. Laub, Small-Sample Statistical Condition Estimates for General Matrix Functions, *SIAM J. on Scientific and Statistical Computing*, **15**, 36–61, 1994.

[KV99] P. Kravanja, M. Van Barel, Algorithms for Solving Rational Interpolation Problems Related to Fast and Superfast Solvers for Toeplitz Systems, *SPIE*, 359–370, 1999.

[LPS92] J. Laderman, V. Y. Pan, H. X. Sha, On Practical Algorithms for Accelerated Matrix Multiplication, *Linear Algebra and Its Applications*, **162–164**, 557–588, 1992.

[MMD08] M. W. Mahoney, M. Maggioni, P. Drineas, Tensor-CUR Decompositions for Tensor-based Data, *SIAM J. on Matrix Analysis and Applications*, **30, 2**, 957–987, 2008.

[MP80] W. L. Miranker, V. Y. Pan, Methods of Aggregations, *Linear Algebra and Its Applications*, **29**, 231–257, 1980.

[MRT11] P.-G. Martinsson, V. Rokhlin, M. Tygert, A randomized algorithm for the decomposition of matrices, *Applied and Computational Harmonic Analysis*, **30**, 47–68, 2011.

[ORO05] T. Ogita, S. M. Rump, S. Oishi, Accurate Sum and Dot Product, *SIAM Journal on Scientific Computing*, **26**, **6**, 1955–1988, 2005.

[OT09] I. V. Oseledets, E. E. Tyrtyshnikov, Breaking the Curse of Dimensionality, or How to Use SVD in Many Dimensions, *SIAM J. Scientific Comp.*, **31**, **5**, 3744–3759, 2009.

[P72] V. Y. Pan, On Schemes for the Evaluation of Products and Inverses of Matrices (in Russian), *Uspekhi Matematicheskikh Nauk*, **27**, **5 (167)**, 249–250, 1972.

[P79] V. Y. Pan, Fields Extension and Trilinear Aggregating, Uniting and Canceling for the Acceleration of Matrix Multiplication, *Proceedings of the 20th Annual IEEE Symposium on Foundations of Computer Science (FOCS 79)*, 28–38, IEEE Computer Society Press, Long Beach, California, 1979.

[P81] V. Y. Pan, New Combinations of Methods for the Acceleration of Matrix Multiplications, *Computers and Mathematics (with Applications)*, **7**, 73–125, 1981.

[P84] V. Y. Pan, How Can We Speed up Matrix Multiplication? *SIAM Review,* **26**, **3**, 393–415, 1984.

[P90] V. Y. Pan, On Computations with Dense Structured Matrices, *Math. of Computation*, **55, 191**, 179–190, 1990. Proceedings version in *Proc. Intern. Symposium on Symbolic and Algebraic Computation (ISSAC'89)*, 34–42, ACM Press, NY, 1989.

[P92] V. Y. Pan, Parallel Solution of Toeplitz-like Linear Systems, *J. of Complexity*, **8**, 1–21, 1992.

[P93] V. Y. Pan, Concurrent Iterative Algorithm for Toeplitz-like Linear Systems, *IEEE Transactions on Parallel and Distributed Systems*, **4, 5**, 592–600, 1993.

[P93a] V. Y. Pan, Decreasing the Displacement Rank of a Matrix, *SIAM Journal on Matrix Analysis and Applications*, **14**, **1**, 118–121, 1993.

[P00] V. Y. Pan, Nearly Optimal Computations with Structured Matrices, *Proceedings of 11th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'2000)*, 953–962, ACM Press, New York, and SIAM Publications, Philadephia, 2000.

[P00a] C.–T. Pan, On the Existence and Computation of Rank-revealing LU Factorization, *Linear Algebra and Its Applications*, **316**, 199–222, 2000.

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

[P07] V. Y. Pan, New Homotopic/Factorization and Symmetrization Techniques for Newton's and Newton/Structured Iteration, *Computer and Math. with Applications*, **54**, 721–729, 2007.

[P10] V. Y. Pan, Newton's Iteration for Matrix Inversion, Advances and Extensions, pp. 364–381, in *Matrix Methods: Theory, Algorithms and Applications* (dedicated to the Memory of Gene Golub, edited by V. Olshevsky and E. Tyrtyshnikov), World Scientific Publishing, New Jersey, ISBN-13 978-981-283-601-4, ISBN-10-981-283-601-2 (2010).

[P11] V.Y. Pan, Nearly optimal solution of rational linear systems of equations with symbolic lifting and numerical initialization. *Computers & Math. with Applications*, **62**, 1685–1706, 2011.

[PBRZ99] V. Y. Pan, S. Brahnam, B. Murphy, R. E. Rosholt, Newton's Iteration for Structured Matrices and Linear Systems of Equations, *SIAM Volume on Fast Reliable Algorithms for Matrices with Structure* (T. Kailath and A. H. Sayed, editors), 189–210, SIAM Publications, Philadelphia, 1999.

[PGMQ] V. Y. Pan, D. Grady, B. Murphy, G. Qian, R. E. Rosholt, A. Ruslanov, Schur Aggregation for Linear Systems and Determinants, *Theoretical Computer Science*, *Special Issue on Symbolic–Numerical Algorithms* (D. A. Bini, V. Y. Pan, and J. Verschelde editors), **409**, **2**, 255–268, 2008.

[PKRK] V. Y. Pan, M. Kunin, R. Rosholt, H. Kodal, Homotopic Residual Correction Processes, *Math. of Computation*, **75**, 345–368, 2006.

[PIMR10] V. Y. Pan, D. Ivolgin, B. Murphy, R. E. Rosholt, Y. Tang, X. Yan, Additive Preconditioning for Matrix Computations, *Linear Algebra and Its Applications*, **432**, 1070–1089, 2010.

[PQ10] V. Y. Pan, G. Qian, Randomized Preprocessing of Homogeneous Linear Systems of Equations, *Linear Algebra and Its Applications*, **432**, 3272–3318, 2010.

[PQa] V. Y. Pan, G. Qian, Solving Linear Systems of Equations with Randomization, Augmentation and Aggregation II, Tech. Report TR 2012007, *Ph.D. Program in Computer Science*, *Graduate Center, the City University of New York*, 2012.
Available at http://www.cs.gc.cuny.edu/tr/techreport.php?id=434

[PQZa] V. Y. Pan, G. Qian, A. Zheng, Randomized Preprocessing versus Pivoting, *Linear Algebra and Its Applications*, in print. http://dx.doi.org/10.1016/j.laa.2011.02.052

[PQZC] V. Y. Pan, G. Qian, A. Zheng, Z. Chen, Matrix Computations and Polynomial Root-finding with Preprocessing, *Linear Algebra and Its Applications*, **434**, 854–879, 2011.

[PRW02] V. Y. Pan, Y. Rami, X. Wang, Structured Matrices and Newtons Iteration: Unified Approach, *Linear Algebra and Its Applications*, **343/344**, 233–265, 2002.

[PS91] V. Y. Pan, R. Schreiber, An Improved Newton Iteration for the Generalized Inverse of a Matrix, with Applications, *SIAM Journal on Scientific and Statistical Computing*, **12, 5**, 1109–1131, 1991.

[R90] S. M. Rump, Approximate Inverses of Almost Singular Matrices Still Contain Useful Information, Technical Report 90.1, *Forschungsschwerpunkt Informations-und Kommunikationstechnik, TU Hamburg-Harburg*, 1990.

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

[S98] G. W. Stewart, *Matrix Algorithms, Vol I: Basic Decompositions*, SIAM, 1998.

[S10] A. J. Stothers, On the Complexity of Matrix Multiplication, Ph.D. Thesis, University of Edinburgh, 2010.

[SST06] A. Sankar, D. Spielman, S.-H. Teng, Smoothed Analysis of the Condition Numbers and Growth Factors of Matrices, *SIAM J. on Matrix Analysis*, **28**, **2**, 446–476, 2006.

[T00] E. E. Tyrtyshnikov, Incomplete Cross Approximation in Mosaic Skeleton Method, *Computing*, **64**, 367–380, 2000.

[V99] M. Van Barel, A Superfast Toeplitz Solver, 1999.
Available at http://www.cs.kuleuven.be/˜ marc/software/index.html

[VBHK01] M. Van Barel, G. Heinig, P. Kravanja, A Stabilized Superfast Solver for Nonsymmetric Toeplitz Systems, *SIAM J. on Matrix Analysis and Applications*, **23**, **2**, 494–510, 2001.

[VK98] M. Van Barel, P. Kravanja, A Stabilized Superfast Solver for Indefinite Hankel Systems, *Linear Algebra and its Applications*, **284**, **1–3**, 335–355, 1998.

[VVM07] R. Vandebril, M. Van Barel, N. Mastronardi, *Matrix Computations and Semiseparable Matrices: Linear Systems* (Volume 1), The Johns Hopkins University Press, Baltimore, Maryland, 2007.

[VVM08] R. Vandebril, M. Van Barel, N. Mastronardi, *Matrix Computations and Semiseparable Matrices: Eigenvalue and Singular Value Methods* (Volume 2), The Johns Hopkins University Press, Baltimore, Maryland, 2008.

[VW12] V. Vassilevska Williams, Multiplying Matrices Faster than Coppersmith–Winograd, *Proc. 44th Annual ACM Symposium on Theory of Computing (STOC 2012)*, 887–898, ACM Press, New York, 2012.

[W07] X. Wang, Affect of Small Rank Modification on the Condition Number of a Matrix, *Computer and Math. (with Applications)*, **54**, 819–825, 2007.

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