2009

TR-2009010: Randomized Preconditioning versus Pivoting

Victor Y. Pan

Guoliang Qian

Ai-Long Zheng

Randomized preconditioning versus pivoting *

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 known that pivoting-free Gaussian elimination is numerically unsafe but can run significantly faster than GEPP. We prove and confirm experimentally that randomized preconditioning combined with iterative refinement can to some extent replace pivoting in numerical Gaussian elimination. The resulting algorithms compute solution more rapidly than GEPP and still with high accuracy. In the case of Toeplitz, Hankel, and other structured inputs we yield acceleration from cubic to nearly linear arithmetic time. Our auxiliary estimates for the condition number of the product of fixed and random matrices can be of independent interest.

2000 Math. Subject Classification: 65F05, 65F22, 65F35

Key Words: Gaussian elimination, Pivoting, Randomized preconditioning, Toeplitz matrix computations

1 Gaussian Elimination with No Pivoting

Gaussian elimination with no pivoting (hereafter we refer to it as GENP) generally fails to produce uncorrupted numerical solution to a linear system of equations because of the propagation of rounding errors. Pivoting, that is row or column interchange, however, takes its toll. It "usually degrades the performance" [GL96, page 119] by interrupting the string of arithmetic computations with the foreign operations of comparisons. Furthermore pivoting rapidly destroys matrix structure of Toeplitz and Hankel types, which increases the running time of the solution dramatically, from nearly linear to cubic. We refer the reader to [VBHK01, Introduction] and the references cited there on various methods of solving or circumventing this problem. They rely on using normal equations and augmentation, look ahead techniques, and displacement transformation. On the latter method we also cite the original source [P90] and [P01, Sections 4.8, 4.9, and 5.6].

For which classes of input matrices can we safely avoid pivoting? In addition to the classical examples of diagonally-dominant and positive definite matrices, we point out random matrices.

*Supported by PSC CUNY Awards 69330–0038 and 61406–0039

Indeed random matrices tend to be strongly well conditioned [D88], [E88], and if they are, then numerical safety of GENP can be proved (see Sections 5 and 6).

In computational practice the input matrices are rarely random, but we prove that pre- and post-multiplication of a matrix by random multipliers randomize it enough to expect that application of GENP becomes numerically safe. Moreover in our extensive tests we observe this tendency of GENP even in the case of circulant multipliers filled with random integers $+1$ and -1.

In our tests the relative residual norms of the output were a little greater in the case of GENP with preprocessing than with GEPP but remained small enough to support rapid and noncostly iterative refinement of the resulting approximate solutions.

We apply our approach to recursive block GENP and block Gauss-Jordan elimination with no pivoting (hereafter referred to as *GJENP*). These variations support superfast solution (that is solution in nearly linear time) of structured linear systems of equations of Toeplitz, Hankel and some other types. We believe that our present demonstration of the power of randomized preconditioning shows just the tip of an iceberg. Distinct variants of our approach in [PQa], [PQb], [PY] facilitate the solution of linear systems of equations, matrix inversion based on augmentaion (cf. also our Section 7), matrix eigen-solving, and root-finding for polynomial and secular equations. Randomized preconditioning seems to be a natural means for the acceleration of structured matrix inversion via Newton's iteration, which is highly effective where the input matrix is structured and well conditioned [P01, Chapter 6], [PRW02], [CPV04], [PKRK06], [P08].

Our estimates for the condition number of the product of fixed and random matrices in Section 3 have further applications in [PQa] and [PQb] and may be of independent interest.

We organize our paper as follows. We devote the next section to definitions and auxiliary results supporting our randomized preconditioning. In Sections 3 and 8 we recal and extend the known estimates for condition numbers of random matrices and of their products with fixed matrices. Section 4 covers the basic 2×2 block triangular factorizations, Section 5 their recursive extension (which include GENP and GJENP as special cases), Section 6 the respective numerical properties and randomized preconditioning, and Section 7 the extension of our study to the case of structured input matrices and random structured preconditioners. Section 9 describes our numerical experiments, designed by the first author and performed by his coauthors, mostly by the second author. Otherwise (tests excluding), the paper is due to the first author and should be cited as his work.

2 Definitions and basic facts

2.1 General matrices

We use and extend the customary definitions (cf. [GL96]).

"GENP" and "GJENP" (resp. "GEPP" and "GJEPP") stand for "Gaussian Elimination with No Pivoting" and "Gauss-Jordan Elimination with No Pivoting" (resp. "with Partial Pivoting").

$\mathbb{C}$ (resp. $\mathbb{R}$) is the field of complex (resp. real) numbers.

M^T and M^H denote the transpose and the Hermitian transpose of an $m \times n$ matrix M, respectively ($M^H = M^T$ for a real matrix M), $\rho = \operatorname{rank} M$ denotes its rank, $\operatorname{nul} M = n - \rho$ its nullity, $N(A)$ its null space, $M^{(k\times l)}$ its $k \times l$ leading (that is northwestern) block submatrix, so that $M = M^{(m\times n)}$. Hereater we write $M^{(k)}$ for $M^{(k\times k)}$. A matrix M of a rank ρ has *generic rank profile* if all its leading blocks $M^{(k)}$ of size $k \times k$ for $k \le \rho$ are nonsingular. If in addition $\rho = n$, the matrix is *strongly nonsingular*.

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

I_n or just I denote the $n \times n$ identity matrix. $\mathbf{e}_i$ is its ith column vector, $i = 0, 1, \dots, n-1$. 0 is a matrix filled with zeros. A matrix U is unitary or orthonornmal if $U^H U = I$.

If the columns of a matrix B of full column rank span the null space $N(A)$, then B is a *null matrix basis* or *nmb* for a matrix A.

$M(n)$ flops suffice to multiply a pair of $n \times n$ matrices, $M(n) \le Cn^{2.375}$ for immense constants C [CW90], $M(n) \le (2n-1)n^2$.

2.2 Structured matrices

$Z = (z_{i,j})_{i,j=0}^{n-1}$ where $z_{i,i-1} = 1$ for all i and $z_{i,j} = 0$ for $j \neq i-1$ is the lower shift matrix. $Z_f = Z + f\mathbf{e}_0\mathbf{e}_{n-1}^T$ for any scalar f, so that $Z = Z_0$, $Z_f^n = fI$.

An f-*circulant matrix* $Z_f(\mathbf{v}) = \sum_{i=0}^{n-1} v_i Z_f^i$ is defined by its first column vector $\mathbf{v} = (v_i)_{i=0}^{n-1}$ and a scalar $f \neq 0$ and is called *circulant* if $f = 1$ and *skew circulant* if $f = -1$.

Theorem 2.1. *(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 $\mathbf{u} = (u_i)_{i=0}^{n-1}$, and $\Omega = (\omega_n^{ij})_{i,j=0}^{n-1}$ is the $n \times n$ matrix of the discrete Fourier transform at n points, $\omega_n = \exp(2\pi\sqrt{-1}/n)$ being the primitive n-th root of one.*

The theorem implies that multiplication and inversion of f-circulant matrices (wherever feasible) produce f-circulant matrices and can be performed in $O(n \log n)$ flops based on FFT.

T is an $n \times n$ Toeplitz matrix (resp. Toeplitz-like matrix of displacement rank d) if it can be nonuniquely represented as the sum of a pair of $n \times n$ circulant and skew circulant matrices (resp. as the sum $\sum_{k=1}^{d} Z_1(\mathbf{g}_k)Z_{-1}(\mathbf{h}_k)^T$ for d pairs of vectors $(\mathbf{g}_k, \mathbf{h}_k)$ of dimension n), although the standard representation of a Toeplitz matrix $T = (t_{i-j})_{i=0,j=0}^{m-1,n-1}$ is by its $m+n-1$ distinct entries $t_{1-n}, \dots, t_{m-1}$.

$J = J_n = (j_{i,k})_{i,k=0}^{n-1}$ is the reflection matrix, $j_{i,n-1-i} = 1$, $j_{i,k} = 0$ unless $i+k = n-1$. ($J^2 = I$.) H is a Hankel matrix (resp. Hankel-like matrix of displacement rank d) if and only if HJ is a Toeplitz matrix (resp. Toeplitz-like matrix of displacement rank d).

See, e.g., [P01] on these and other popular classes of structured matrices, such as Vandermonde matrices $V = (t_i^j)_{i=0,j=0}^{m-1,n-1}$, Cauchy matrices $C = (\frac{1}{s_i - t_j})_{i=0,j=0}^{m-1,n-1}$, and matrices with similar structures.

2.3 Matrix norms, SVDs, inverses, and condition numbers

$||A||_h$ denotes the h-norm of a matrix A, $h = 1, 2, \infty$. We write $||A||_2 = ||A||$. For $A = (a_{i,j})_{i,j=1}^{m,n}$ we have

$$\max_{i,j=1}^{m,n} |a_{i,j}| \leq ||A|| = ||A^H|| \leq \sqrt{mn} \max_{i,j=1}^{m,n} |a_{i,j}|. \tag{2.1}$$

$A = S_A \Sigma_A T_A^H$ is a *full SVD* of an $m \times n$ matrix A of a rank ρ provided $S_A S_A^H = S_A^H S_A = I_m$, $T_A T_A^H = T_A^H T_A = I_n$, $\Sigma_A = \mathrm{diag}(\widehat{\Sigma}_A, 0_{m-\rho,n-\rho})$, $\widehat{\Sigma}_A = \mathrm{diag}(\sigma_j(A))_{j=1}^{\rho}$, $\sigma_j = \sigma_j(A) = \sigma_j(A^H)$ is the jth largest singular value of a matrix A, having its minimax characterization

$$\sigma_j = \max_{\dim(\mathbb{S})=j} \min_{\mathbf{x}\in\mathbb{S}, ||\mathbf{x}||=1} ||A\mathbf{x}|| \tag{2.2}$$

where $\mathbb{S}$ denotes linear spaces [GL96, Theorem 8.6.1] and $j = 1, \dots, \rho$. It follows that

$$\sigma_1 = \max_{||\mathbf{x}||=1} ||A\mathbf{x}|| = ||A||, \quad \sigma_n = \min_{||\mathbf{x}||=1} ||A\mathbf{x}||. \tag{2.3}$$

The matrix $X = A^{(I)}$ is a left (resp. right) inverse of a matrix A if $XA = I$ (resp. $AX = I$). $A^{(I)} = A^{-1}$ for a nonsingular matrix A.

$\mathrm{cond}\, A = \sigma_1(A)/\sigma_\rho(A)$ is the condition number of a 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. The concepts "large", "ill" and "well conditioned" are quantified in the context of the computational task and computer environment.

2.4 Singular values of submatrices and matrix products

The two following theorems are used in the proofs of Theorems 3.4 and 3.5 in Section 3. The first of them follows from minimax characterization (2.2).

Theorem 2.2. *Fix four positive integers p, q, m, and n, assume that A_0 is a $p \times q$ submatrix of an $m \times n$ matrix A. Then $\sigma_j(A) \geq \sigma_j(A_0)$ for $j = 1, 2, \dots, \min\{p, q\}$.*

Theorem 2.3. *Let $A \in \mathbb{C}^{m\times r}$ and $B \in \mathbb{C}^{r\times n}$ and write $r_A = \operatorname{rank} A$, $r_B = \operatorname{rank} B$, $r_- = \min\{r_A, r_B\}$ and $r_+ = \max\{r_A, r_B\}$. Let $r_+ = r$. (In particular this holds if at least one of the matrices A and B is nonsingular.) Then* $\operatorname{rank}(AB) = r_-$, $\sigma_{r_-}(AB) \geq \sigma_{r_A}(A)\sigma_{r_B}(B)$ *and* $\operatorname{cond}(AB) \leq (\operatorname{cond} A)\operatorname{cond} B$.

Proof. Let $M = S_M \Sigma_M T_M^H$ be a full SVD where $\Sigma_M = \operatorname{diag}(\widehat{\Sigma}_M, 0_{s,t})$, $\widehat{\Sigma}_M = \operatorname{diag}(\sigma_j(M))_{j=1}^{r_M}$ for $M = A$, $s = m - r_A$, $t = r - r_A$ as well as for $M = B$, $s = r - r_B$, $t = n - r_B$. Let the matrix $\widehat{A} \in \mathbb{C}^{r_A\times r}$ (resp. $\widehat{B}^H \in \mathbb{C}^{r_B\times r}$) be obtained by deleting the zero rows of the matrix $\Sigma_A T_A^H = S_A^H A$ (resp. $\Sigma_B S_B^H = T_B^H B^H$), so that $\widehat{A}\widehat{B} \in \mathbb{C}^{r_A\times r_B}$. This pruning keeps all singular values (and therefore the ranks) of the matrices A, B, and AB intact. Clearly $\widehat{A}$ and $\widehat{B}$ are full rank matrices. Furthermore the equation $r_+ = r$ implies that at least one of the matrices $\widehat{A}$ and $\widehat{B}$ is nonsingular and the product $\widehat{A}\widehat{B}$ has full rank $r_- = \operatorname{rank}(AB)$.

Suppose $r_A = r$. Then $m \geq r_A = r_+ = r \geq r_B$. From minimax characterization (2.2) we obtain $\sigma_{r_-}(\widehat{A}\widehat{B}) = ||\widehat{A}\widehat{B}\mathbf{x}||$ where $||\mathbf{x}|| = 1$. We have $\widehat{B}\mathbf{x} \neq \mathbf{0}$ because $\widehat{B} \in \mathbb{C}^{r\times r_B}$. Therefore $\sigma_{r_-}(\widehat{A}\widehat{B}) = \sigma_A\sigma_B$ where $\sigma_A = ||\widehat{A}\mathbf{y}||$, $\mathbf{y} = \widehat{B}\mathbf{x}/||\widehat{B}\mathbf{x}||$, $\sigma_B = ||\widehat{B}\mathbf{x}||$. Since $\widehat{A} \in \mathbb{C}^{r_A\times r} = \mathbb{C}^{r_A\times r_A}$, we can apply equation (2.3) for $A = \widehat{A}$ and $n = r_A$ to obtain that $\sigma_{r_A}(\widehat{A}) = \min_{||\mathbf{z}||=1}||\widehat{A}\mathbf{z}|$. Consequently $\sigma_{r_A}(\widehat{A}) \leq \sigma_A$. Likewise $\widehat{B} \in \mathbb{C}^{r\times r_B}$, and so we can apply equation (2.3) for $A = \widehat{B}$ and $n = r_B$ to obtain that $\sigma_{r_B}(\widehat{B}) = \min_{||\mathbf{z}||=1}||\widehat{B}\mathbf{z}||$. Therefore $\sigma_B \geq \sigma_{r_B}(B)$. Consequently $\sigma_{r_-}(AB) = \sigma_{r_-}(\widehat{A}\widehat{B}) = \sigma_A\sigma_B \geq \sigma_{r_A}(A)\sigma_{r_B}(B)$.

If $r_A < r_+ = r = r_B$, then the same argument shows that $\sigma_{r_-}(AB) = \sigma_{r_-}(B^H A^H) \geq \sigma_{r_A}(B^H)\sigma_{r_B}(A^H) = \sigma_{r_A}(A)\sigma_{r_B}(B)$.

Finally recall that $||AB|| \leq ||A||\,||B||$, and so $\operatorname{cond}(AB) \leq (\operatorname{cond} A)\operatorname{cond} B$. ☐

Remark 2.1. $\operatorname{cond}(AB)$ *can be arbitrarily large even for $m \times r$ unitary matrices A and B^H if $m > r$.*

2.5 Random sampling and random matrices

$|\Delta|$ is the cardinality of a set Δ. *Random sampling* of elements from a set Δ is their selection from this set at random, independently of each other, and under the uniform probability distribution on Δ. A matrix is *random* if its entries are randomly sampled (from a fixed set Δ).

Lemma 2.1. *[DL78] (cf. also [S80], [Z79]). For a set Δ of cardinality $|\Delta|$ (in a fixed ring), let a polynomial in m variables have total degree d, let it not vanish identically on the set Δ^m, and let the values of its variables be randomly sampled from the set Δ. Then the polynomial vanishes with a probability of at most $d/|\Delta|$.*

Corollary 2.1. *An $m \times n$ matrix with entries sampled at random from a set Δ has full rank with a probability of at least $1 - l/|\Delta|$ for $l = \min\{m, n\}$. Such a matrix of a rank ρ has generic rank profile with a probability of at least $1 - (\rho + 1)\rho/|\Delta|$*

Proof. The determinant of a $r \times r$ matrix is a nonvanishing polynomial of degree r in its entries. It remains to apply Lemma 2.1 to $r \times r$ submatrices of the input matrix. ☐

Definition 2.1. $F_X(y) = Probability\{X \leq y\}$ *for a real random variable X is the* cumulative distribution function (CDF) *of X evaluated at y. $F_A(y) = F_{\sigma_l(A)}(y)$ for an $m \times n$ matrix A and an integer $l = \min\{m, n\}$. A matrix (resp. vector) is a* Gaussian random matrix *(resp. vector) with a mean μ and a variance σ^2 if it is filled with independent Gaussian random variables, all having the same mean μ and variance σ^2. If $\mu = 0$ and $\sigma^2 = 1$, this is a* standard Gaussian random matrix *(resp. vector). $F_{\mu,\sigma}(y) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{y}\exp(-\frac{(x-\mu)^2}{2\sigma^2})dx$ is the CDF for a Gaussian random variable with a mean μ and a variance σ^2.* $\Phi_{\mu,\sigma}(y) = F_{\mu,\sigma}(y) - F_{\mu,\sigma}(-y) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-y}^{y}\exp(-\frac{(x-\mu)^2}{2\sigma^2})dx$ for $y \geq 0$.

We use the two following lemmas (the first of them is immediately verified).

Lemma 2.2. *For positive y, y_1, and y_2 we have*
$F_X(y) \le F_{X_1}(y_1) + F_{X_2}(y_2)$ *if* $X \le \min\{X_1 y/y_1, X_2 y/y_2\}$,
$1 - F_X(y) \le 2 - F_{X_1}(y_1) - F_{X_2}(y_2)$ *if* $X \ge \max\{X_1 y/y_1, X_2 y/y_2\}$.

Lemma 2.3. *We have*

a) $\frac{y}{\sigma}\sqrt{\frac{2}{\pi}}\exp(-\frac{(|\mu|+y)^2}{2\sigma^2}) \le \Phi_{\mu,\sigma}(y) \le \frac{y}{\sigma}\sqrt{\frac{2}{\pi}}$ *for* $y \ge 0$,

b) $\Phi_{\mu,\sigma}(y) \le \frac{y}{\sigma}\sqrt{\frac{2}{\pi}}\exp(-\frac{(|\mu|-y)^2}{2\sigma^2})$ *for* $0 \le y < |\mu|$,

c) $\Phi_{\mu,\sigma}(y) \approx \frac{y}{\sigma}\sqrt{\frac{2}{\pi}}\exp(-\frac{\mu^2}{2})$ *for* $y \approx 0$,

d) $1 - \Phi_{\mu,\sigma}(y) \le \frac{\sigma}{\sqrt{2\pi}}(\frac{1}{y-\mu}\exp(-\frac{(y-\mu)^2}{2\sigma^2}) + \frac{1}{y+\mu}\exp(-\frac{(y+\mu)^2}{2\sigma^2}))$ *for* $y \ge |\mu| + \sigma$,

e) $1 - \Phi_{\mu,\sigma}(y) \le \frac{1}{a}\sqrt{\frac{2}{\pi}}\exp(-\frac{a^2}{2})$ *for* $y \ge |\mu| + a\sigma,\ a \ge 1$,

f) $k(1 - \Phi_{\mu,\sigma}(y)) \le \frac{1}{a+b}\sqrt{\frac{2}{\pi}}\exp(-\frac{(a+2b)a}{2})$ *for* $y \ge |\mu| + (a+b)\sigma$, $b = \sqrt{2\ln k}$, $k \ge 2$, $a + b \ge 1$.

Proof. Parts a)–c) are immediately verified. Observe that $\Phi_{\mu,\sigma}(y) = \Phi_{0,1}(\frac{y-\mu}{\sigma}) = F_{0,1}(\frac{y-\mu}{\sigma}) - F_{0,1}(\frac{-y-\mu}{\sigma})$ and $F_{0,1}(\frac{-y-\mu}{\sigma}) = 1 - F_{0,1}(\frac{y+\mu}{\sigma})$. Deduce that $1 - \Phi_{\mu,\sigma}(y) = 2 - F_{0,1}(\frac{y-\mu}{\sigma}) - F_{0,1}(\frac{y+\mu}{\sigma})$. Combine this equation with the bound $1 - F_{0,1}(z) \le \frac{1}{z\sqrt{2\pi}}\exp(-\frac{z^2}{2})$ for $z = \frac{y\pm\mu}{\sigma} \ge 1$ from [SST06, Lemma A.1] and obtain part d). Part e) follows from part d) and implies part f). $\square$

3 Extremal singular values of random matrices and of their products with fixed matrices

Gaussian random matrices (cf. Definition 2.1) tend to be well conditioned [D88], [E88], and even perturbations by such a matrix A is expected to make a matrix M well conditioned if the norms $||A||$ and $||M||$ have the same order [SST06]. Next we specify and extend the respective estimates using a constant $c \le 2.35$ from [SST06].

Theorem 3.1. *Assume an $m \times n$ matrix A filled with d random variables $X_1, \dots, X_d$. ($d = m + n - 1$ for random Toeplitz and Hankel matrices A, $d = mn$ for random general matrix A.) Write $F_-(y) = \min_{i=1}^d F_{|X_i|}(y)$ for $y \ge 0$. Then a) $F_-(y) = \Phi_{\mu,\sigma}(y)$ where $X_1, \dots, X_d$ are Gaussian random variables with a mean μ and a variance σ, whereas b) $F_-(y) = y/a$ (resp. $F_-(y) = (y/a)^2$) where $0 \le y \le a$ and the random variables $X_1, \dots, X_d$ are uniformly distributed on the real line segments $[-a, a]$ or $[0, a]$ (resp. the circle $\{x : |x| \le a\}$ on the complex plane). Furthermore we have c) $1 - F_{||A||}(y) \le (1 - F_-(y/\sqrt{mn}))d$, which is a trivial bound unless $F_-(y/\sqrt{mn}) > 1 - 1/d$, and d) $F_{||A||}(y) \ge (F_-(y/\sqrt{mn}))^d$ if the d random variables are independent of each other.*

Proof. Parts a) and b) are immediately verified. Part d) follows because bounds (2.1) imply that $||A|| \le y$ if $|X_i| \le y/\sqrt{mn}$ for all i. Apply Lemma 2.2 to deduce part c). $\square$

Theorem 3.2. *(See [DS01, Theorem II.7].) Suppose $A \in \mathbb{T}^{n\times n}$ is a Gaussian random matrix with mean zero and variance σ^2. Then $F_{||A||}(y) \ge 1 - \exp(-x^2/2)$ for $x = y/\sigma - 2\sqrt{n} \ge 0$.*

Theorem 3.3. *(See [SST06, Theorem 3.3].) Suppose $M \in \mathbb{T}^{m\times n}$, $\bar{U} \in \mathbb{T}^{m\times m}$, and $\bar{V} \in \mathbb{T}^{n\times n}$ are three fixed matrices, $\bar{U}$ and $\bar{V}$ are unitary matrices, $A \in \mathbb{T}^{m\times n}$ is a Gaussian random matrix independent of the matrix M and having mean zero and a variance σ^2, $W = \bar{U}(A + M)\bar{V}$, $l = \min\{m, n\}$, and $y \ge 0$. Then $F_W(y) \le cy\sqrt{l}/\sigma$.*

Combining the two latter theorems implies the following result.

Corollary 3.1. *(See [SST06, Theorem 3.1].) Under the assumptions of Theorem 3.3, let $||M|| \le \sqrt{l}$. Then $F_{\mathrm{cond}\,W}(y) \ge 1 - (c_1 + c_2\sqrt{(\ln y)/n})n/(y\sigma)$ for $c_1 = 14.1$, $c_2 = 4.7\sqrt{2}$, and all $y \ge 1$.*

On a further improvement of this bound by the factor of $\sqrt{\log n}$, see [W04].

Let us employ our results in Section 2.4 to extend the estimates of Theorem 3.3 to yield probabilistic lower bounds on the smallest singular values of the products of fixed and random matrices.

Theorem 3.4. *Suppose that $G \in \mathbb{T}^{q\times m}$, $H \in \mathbb{T}^{n\times r}$, and a random matrix $W \in \mathbb{T}^{m\times n}$ have full rank ρ with probability one. Write $r_G = \operatorname{rank} G$ and $r_H = \operatorname{rank} H$. Let $y \ge 0$. Then $F_{GW}(y) \le F_W(y/\sigma_{r_G}(G))$ if $r_G = m$, whereas $F_{WH}(y) \le F_W(y/\sigma_{r_G}(H))$ if $r_H = n$.*

Proof. The theorem follows from Theorem 2.3. □

In view of Remark 2.1, we cannot merely drop the above assumptions that $r_G \ge m$ and $r_H \ge n$, but the next theorem (employing Theorem 3.4) circumvents the problem. We use this theorem only for $\bar{U} = 0$ and $\bar{V} = 0$.

Theorem 3.5. *Suppose $\mathbb{T} \in \mathbb{C}$ or $\mathbb{T} \in \mathbb{R}$, $G \in \mathbb{T}^{r_G\times m}$, $H \in \mathbb{T}^{n\times r_H}$, $X \in \mathbb{T}^{m\times n}$, $\tilde{U} \in \mathbb{T}^{r_G\times n}$, $\tilde{V} \in \mathbb{T}^{m\times r_H}$, $\operatorname{rank} G = r_G < m$, $\operatorname{rank} H = r_H < n$, the assumptions of Theorem 3.3 hold for the matrix X replacing W, and $y \ge 0$. Then*
a) $F_{GX+\tilde{U}}(y) \le cy\sqrt{l}/(\sigma_{r_G}(G)\sigma)$ and
b) $F_{XH+\tilde{V}}(y) \le cy\sqrt{l}/(\sigma_{r_H}(H)\sigma)$.

Proof. Let $\bar{G}$ maximize the value $\sigma_{r_G}(\bar{G})$ among all $r_G \times r_G$ block submatrices $\bar{G}$ of the matrix G, which implies that the matrix $\bar{G}$ is nonsingular. Write $G = \bar{G}(G_1, I_{r_G}, G_2)$, $X^T = (X_1^T, \bar{X}^T, X_2^T)$, and $GX+\bar{U} = \bar{G}(A+M)$ where $A = \bar{X}$ and $M = G_1X_1+G_2X_2+\bar{U}$. Observe that the assumptions of Theorem 3.4 hold for G replaced by $\bar{G}$ and W by $A+M$. Deduce that $F_{GX}(y) \le F_{A+M}(y/\sigma_{r_G}(\bar{G}))$. Obtain Theorem 3.5a by combining this estimate with the bounds $F_{A+M}(y) \le cy\sqrt{l}/\sigma$ (implied by Theorem 3.3) and $\sigma_{r_G}(\bar{G}) \le \sigma_{r_G}(G)$ (implied by Theorem 2.2). Obtain Theorem 3.5b by applying Theorem 3.5a to the matrices X^T, H^T, $(XH)^T$, and V replacing the matrices X, G, GX, and U, respectively. □

Corollary 3.2. *Suppose k, m, and n are integers, $1 \le k \le m \le n$, $\mathbb{T} \in \mathbb{C}$ or $\mathbb{T} \in \mathbb{R}$, $G,\ H^T \in \mathbb{T}^{m\times n}$, $\operatorname{rank} G = \operatorname{rank} H = n$, X is a standard Gaussian random matrix with a mean μ and a variance σ, and $y \ge 0$. Then*
a) $F_{(GX)^{(k)}}(y) \le cy\sqrt{k}/(\sigma_n(G)\sigma)$ and
b) $F_{(XH)^{(k)}}(y) \le cy\sqrt{k}/(\sigma_n(H)\sigma)$.

Corollary 3.3. *Under the assumptions of Corollary 3.2 choose a scalar $z \ge 2\sigma\sqrt{n}$. Then*
a) $F_{\operatorname{cond}(GX)^{(k)}}(yz||G||) \ge 1 - \exp(\frac{(z-2\sigma\sqrt{n})^2}{2\sigma^2}) - cy\sqrt{k}/(\sigma_n(G)\sigma)$ and
b) $F_{\operatorname{cond}(XH)^{(k)}}(yz||H||) \ge 1 - \exp(\frac{(z-2\sigma\sqrt{n})^2}{2\sigma^2}) - cy\sqrt{k}/(\sigma_n(H)\sigma)$.

Proof. a) Theorem 3.2 for $y = z$ and Corollary 3.2a) together imply that $F_{||X||\ \sigma_k(GX)^{(k)}}(yz) \ge 1 - \exp(\frac{(z-2\sigma\sqrt{n})^2}{2\sigma^2}) - cy\sqrt{k}/(\sigma_n(G)\sigma)$. Recall that $||(GX)^{(k)}|| \le ||GX|| \le ||G||\ ||X||$ and obtain part a). Part b) is proved similarly. □

Our next definition extends the class of strongly nonsingular matrices.

Definition 3.1. *A matrix M is* strongly well conditioned *if all its leading submatrices $M^{(k)}$ are well conditioned.*

Suppose G and H are well conditioned matrices of full rank. Corollary 3.3 implies that their *preconditioning with a Gaussian random matrix X is expected to yield strongly well conditioned matrices GX and XH.* If the above matrix G (resp. H) is ill conditioned but has a $k \times l$ well conditioned submatrix F in its row set α_k and column set β_l, then one can obtain from the corollary that any $k \times l$ submatrix of the matrix GX in its row set α_k (resp. of the matrix XH in its column set β_l) is expected to be strongly well conditioned provided X is standard Gaussian random matrix in $\mathbb{T}^{n\times n}$. Furthermore one can deduce that the $k \times l$ leading submatrix of the matrix YGX (resp. XHY) for standard Gaussian random matrix $Y \in \mathbb{T}^{m\times m}$ is expected to be strongly well conditioned. This property holds even in the *Hermitian preprocessing* where $X = Y^H$ is standard Gaussian random matrix, in which case we can preserve the Hermitian structure of the input matrix.

In the next sections we elaborate upon application of such preconditioning as a substitution for pivoting in the basic and block recursive versions of GEPP and GJEPP.

4 Basic block factorizations

Let n, k, and $r = n - k$ be positive integers and let $M_{00} = M^{(k)}$ be the nonsingular $k \times k$ leading block of an $n \times n$ matrix $M = \begin{pmatrix} M_{00} & M_{01} \\ M_{10} & M_{11} \end{pmatrix}$, represented as a 2×2 block matrix. Then a single step of block GENP and block GJENP outputs the block factorizations

$$M = \begin{pmatrix} I_k & 0 \\ M_{10}M_{00}^{-1} & I_r \end{pmatrix} \begin{pmatrix} M_{00} & M_{01} \\ 0 & S \end{pmatrix} \tag{4.1}$$

and

$$M = \begin{pmatrix} I_k & 0 \\ M_{10}M_{00}^{-1} & I_r \end{pmatrix} \begin{pmatrix} M_{00} & 0 \\ 0 & S \end{pmatrix} \begin{pmatrix} I_r & M_{00}^{-1}M_{01} \\ 0 & I_r \end{pmatrix} \tag{4.2}$$

where

$$S = S(M^{(k)}, M) = M_{11} - M_{01}M_{00}^{-1}M_{01} \tag{4.3}$$

denotes the Schur complement of $M_{00} = M^{(k)}$ in M. These factorizations are infeasible (resp. numerically unstable) if the block M_{00} is singular (resp. ill conditioned). If the matrices M and S are nonsingular, then we have

$$M^{-1} = \begin{pmatrix} M_{00}^{-1} & -M_{00}^{-1}M_{01}S^{-1} \\ 0 & S^{-1} \end{pmatrix} \begin{pmatrix} I_k & 0 \\ -M_{10}M_{00}^{-1} & I_r \end{pmatrix} \tag{4.4}$$

and

$$M^{-1} = \begin{pmatrix} I_r & -M_{00}^{-1}M_{01} \\ 0 & I_r \end{pmatrix} \begin{pmatrix} M_{00}^{-1} & 0 \\ 0 & S^{-1} \end{pmatrix} \begin{pmatrix} I_k & 0 \\ -M_{10}M_{00}^{-1} & I_r \end{pmatrix}. \tag{4.5}$$

Note that S^{-1} is the trailing (southeastern) block of M^{-1}.

5 Recursive (block) factorizations, GENP and GJENP

We can extend factorizations (4.1)–(4.5) to the diagonal blocks M_{00}, S, M_{00}^{-1} and S^{-1} as long as these blocks are not scalars, are nonsingular, and have nonsingular leading blocks. If these assumptions hold recursively, we can recursively extend the factorization process until it ends where all diagonal blocks turn into scalars. At this point we can multiply together all lower triangular factors in (4.1) (resp. all lower triangular factors as well as all upper triangular factors in (4.2)) to arrive at the unique LU (resp. LDU_1) factorization of the matrix M, where the factor D is diagonal, the factor U is upper triangular, and the factor L is (resp. both factors L and U_1^T are) unit lower triangular. The computation of the recursive factorizations (4.1) and (4.2) represents block GENP and block GJENP, respectively, and is completely defined by the choices of the dimensions k and r in all steps. These processes turn into GENP and GJENP, respectively, if we always choose $k = 1$ in all steps, that is choose the size 1×1 for all the leading blocks involved. We call the recursive process *balanced* if $|k - r| \leq 1$ in all its steps.

Due to equations (4.4) and (4.5) the same processes also define recusive factorizations of the inverse M^{-1} if M is a nonsingular matrix.

Theorem 5.1. *Suppose the recursive block factorization process based on (4.1) or (4.2) has been completed for an $n \times n$ input matrix M. Then (a) this takes $O(n^3)$ flops for the GENP process and (b) $O(M(n))$ flops of level 3 BLAS for the balanced process (cf. Section 2 and [GL96]). If M is a Toeplitz-like or Hankel-like matrix having a displacement rank d and represented with its displacement generator of lengths $O(r)$, then displacement representations of lengths $O(d)$ for all auxiliary matrices involved can be computed by using (c) $O(dn^2)$ flops in the case of the GENP process and (d) $O(d^2 n \log^2 n)$ flops in the case of the balanced processes. (e) Given a complete recursive factorization of a matrix M based on equations (4.1)–(4.3), we can compute the value* $\det M$ *in $n - 1$ flops and the vector $M^{-1}\mathbf{b}$ in $O(n^2)$ flops for general nonsingular matrix M and in $O(nd \log n)$ flops for a Toeplitz-like or Hankel-like matrix M having a displacement rank d.*

Proof. See, e.g., [GL96] on part (a), [AHU74] on part (b), [GKO95] on part (c), and [M80], [BA80], [P01, Chapter 5] on part (d). □

The output LU and LDU_1 factorizations do not depend on the order of decreasing the block sizes to 1×1. The auxiliary Schur complements in these factorizations are completely defined by their locations in the matrix M.

Theorem 5.2. *In every step of the recursive block factorization process based on (4.1) or (4.2) every diagonal block of a block diagonal factor is either a leading block of the input matrix M or the Schur complement*

$$S(M^{(h)}, M^{(k)}) = (S(M^{(h)}, M))^{(h)} \tag{5.1}$$

for some integers h and k, $0 < h < k \leq n$.

Corollary 5.1. *(Cf. Corollary 2.1.) a) The recursive block factorization process based on equations (4.1) or (4.2) can be completed if and only if the input matrix M is strongly nonsingular. b) Generally, for any input matrx M, this process completes the recursive block factorization of the strongly nonsingular leading block of this matrix that has the largest size $\rho \times \rho$, $\rho \leq \operatorname{rank} M$. The flop estimates in Theorem 5.1 with n replaced by ρ can be applied to these restricted computations. c) For matrices M having generic rank profile $\rho = \operatorname{rank} M$, $S(M^{(\rho)}, M) = 0$, and a nmb of the matrix M is given by the matrix*

$$F = \begin{pmatrix} -M_{00}^{-1} M_{01} \\ I_r \end{pmatrix}. \tag{5.2}$$

6 Numerical behavior of recursive block GENP and GJENP and randomized preconditioning

Theorem 6.1. *Let us write $N = ||M||$ and $N_- = \max_{k=1}^{n} ||(M^{(k)})^{-1}||$. Then $\operatorname{cond} D \leq (1 + N_- N) N N_-$ for every diagonal block in the recursive factorization of the matrix M in Section 4.*

Proof. For the blocks M_{00} of M and S^{-1} of M^{-1} we surely have $||M_{00}|| \leq N$, $||M_{00}^{-1}|| \leq N_-$, and $||S^{-1}|| \leq ||M^{-1}|| \leq N_-$. We also have $||S|| \leq N + N_- N^2$, due to (4.3). Now the claimed bound follows from Theorem 5.2. □

Clearly the bound $(1 + N_- N) N N_-$ is not large if and only if the matrix M is strongly well conditioned, and if so, recursive GENP and GJENP for this matrix are numerically stable.

Our study implies that GENP and GJENP (as well as recursive GENP and GJENP) for a matrix $M^{(\rho)}$ where $\rho = \operatorname{rank} M$ are feasible and well conditioned if and only if the matrix $M^{(\rho)}$ is strongly nonsingular and strongly well conditioned. Furthermore these properties are expected to hold for the matrices $XMY^{(\rho)}$ if the matrix $M^{(\rho)}$ is nonsingular and well conditioned and if X and Y are square standard Gaussian random matrices (cf. Corollary 3.3).

7 Random structured input matrices and preconditioners. Inversion via randomized augmentation

In the case of structured input matrices the standard and recursive GENP and GJENP can be performed in nearly linear arithmetic time (see Theorem 5.1), but these superfast algorithms are prone to numerical stability problems [B85]. Can we extend our formal study in the previous sections to support structured preconditioners? The answer is negative, except for the results in the next section on circulant matrices and the norm bounds in Theorems 3.1 and 3.2, which hold for or can be readily extended to sparse and structured matrices. In particular no good estimates for the smallest positive singular values seem to be available for random Toeplitz, Hankel, banded and other sparse matrices yet, whereas random Vandermonde and Cauchy matrices tend to be ill conditioned (cf. [GI88]). Respectively we have no extension of Theorems 3.3–3.5 and Corollaries 3.1–3.3 to structured matrices of these classes. Our extensive experiments (see Section 9) show, however, that random Toeplitz (and therefore random Hankel) matrices tend to be well conditioned, although less eagerly than random general and circulant matrices (cf. Tables 9.1–9.3).

Multiplication by Toeplitz and Hankel matrices does not destroy Toeplitz and Hankel structure but a little spoils it. Let us outline *randomized augmentation* techniques, which perfectly preserve the input structure (cf. [PQb]). Let $K \rightarrow \begin{pmatrix} W & V \\ B & A \end{pmatrix}$ for a positive r and nonsingular matrices $A \in \mathbb{C}^{n\times n}$, $W \in \mathbb{C}^{r\times r}$, and $K \in \mathbb{C}^{(n+r)\times(n+r)}$. (If the matrix A is Hermitian, we can keep this property for the matrix K by choosing $B = V^H$ and a Hermitian matrix W.) According to the formal study in [PQb], we can expect that $\operatorname{cond} K^{(h)}$ has the order of the ratio $\sigma_1(A^{(h)})/\sigma_{h-r}(A^{(h)})$ for $h = r+1, \dots, n$ in the case of random and properly scaled matrices B, V, and W. In the experiments in [PQb] this property was consistently observed also where B, V, and W were scaled random structured matrices, which allowed us to keep the structure of the matrix A intact in the transition $A \rightarrow K$. For a large class of well conditioned but not strongly well conditioned matrices A this transition is expected to yield strongly well conditioned matrices K, even where the positive integers r are small. In this case the inverses K^{-1} can be safely and rapidly computed by means of recursive block GENP or GJENP, and then we can readily recover the inverse A^{-1} as follows.

Recall that $K^{-1} = \begin{pmatrix} X & Y \\ Z & F \end{pmatrix}$ for $F = (A - BW^{-1}V)^{-1}$. Apply the Sherman–Morrison–Woodbury formula [GL96, page 50] and obtain that $A^{-1} = F - FUG^{-1}V^HF$ for $U = BW^{-1}$ and $G = I_r - V^HFU$. The original inversion problem is reduced to the case of the matrices W and G of smaller size.

Remark 7.1. *If A is a Toeplitz matrix, then to yield a Toeplitz matrix K, we must fix the Toeplitz matrix W. We can, however, vary the matrix W if we apply our approach to invert the Hankel matrix $H = JA$ and then obtain $A^{-1} = H^{-1}J$.*

8 The norms of random f-circulant matrices and of their inverses

In this section we estimate the norms of random f-circulant matrices and their inverses. The estimates show that such matrices can be expected to be well conditioned for f of the order $1 \pm c/n$ and positive constants c.

Theorem 8.1. *Let an $n \times n$ circulant matrix $A = Z_1(\mathbf{v}) = \sum_{i=0}^{n-1} v_i Z_f^i$ be defined by a Gaussian random vector $\mathbf{v} = (v_i)_{i=0}^{n-1}$ having a mean μ and a variance σ^2. Then we have*

$$F_{||A||}(y) \geq (n-1)(\Phi_{0,2\sigma\sqrt{n}}(y) - 1) + \Phi_{\mu n,\sigma\sqrt{n}}(y),$$

$$F_A(y) \leq (n-1)\Phi_{0,\sigma\sqrt{n/\sqrt{2}}}(y) + \Phi_{\mu n,\sigma\sqrt{n}}(y)$$

for all nonnegative y and for $\Phi_{\mu,\sigma}(y)$ in Definition 2.1.

Proof. Represent the matrix A as in Theorem 2.1 and write $B = \Omega A \Omega^{-1} = D(\Omega\mathbf{v})$. Recall that $\frac{1}{\sqrt{n}}\Omega$ is a unitary matrix, so that $\sigma_j(A) = \sigma_j(B)$ for all j.

Write $\mathbf{u} = (u_i)_{i=0}^{n-1} = \Omega\mathbf{v}$. We have $\sigma_1(B) = \max_j |u_j|$ and $\sigma_n(B) = \min_j |u_j|$ because $B = \operatorname{diag}(u_i)_{i=0}^{n-1}$. Our next goal is to define Gaussian random variables s_i and t_i such that

$$|s_i| \leq |u_i| \leq |t_i| \tag{8.1}$$

for all i and thus (cf. Lemma 2.2)

$$F_{||A||}(y) = F_{\max_i |u_i|}(y) \geq 1 - \sum_i (1 - F_{|u_i|}(y)) \geq 1 - n + \sum_i F_{|s_i|}(y), \tag{8.2}$$

$$1 - F_A(y) = 1 - F_{\min_i |u_i|}(y) \geq 1 - \sum_i F_{|u_i|}(y) \geq 1 - \sum_i F_{|t_i|}(y). \tag{8.3}$$

Clearly $u_0 = \sum_{j=0}^{n-1} v_j$ is a Gaussian random variable with the mean μn and variance $n\sigma^2$, so that $s_0 = t_0 = u_0$,

$$F_{|s_0|} = F_{|t_0|} = F_{|u_0|}(y) = \Phi_{\mu n, \sigma\sqrt{n}}(y). \tag{8.4}$$

Now represent the vector $\mathbf{u}$ as $\bar{\mathbf{u}} + \tilde{\mathbf{u}}\sqrt{-1}$ where $\bar{\mathbf{u}} = (\bar{u}_i)_{i=0}^{n-1} = \mathrm{Re}\,\mathbf{u}$, $\tilde{\mathbf{u}} = (\tilde{u}_i)_{i=0}^{n-1} = \mathrm{Im}\,\mathbf{u}$, and $\mathrm{Re}\,\mathbf{z}$ and $\mathrm{Im}\,\mathbf{z}$ stand for the real and imaginary parts of a complex vector $\mathbf{z}$, respectively. $\bar{\mathbf{u}}$ and $\tilde{\mathbf{u}}$ are real Gaussian random vectors because they are filled with real linear functions $\bar{u}_i = \sum_{j=0}^{n-1} v_j \cos\frac{2\pi ij}{n}$ and $\tilde{u}_i = \sum_{j=0}^{n-1} v_j \sin\frac{2\pi ij}{n}$ of independent Gaussian random variables $v_0, \ldots, v_{n-1}$. For all positive i the Gaussian random variables $\bar{u}_i$ and $\tilde{u}_i$ have the same mean zero (because $\Omega(1,1,\ldots,1)^T = (n,0,\ldots,0)^T)$) and the variances $\bar{\sigma}_i^2 = \sigma^2 \sum_{j=0}^{n-1}(\cos\frac{2\pi ij}{n})^2$ and $\tilde{\sigma}_i^2 = \sigma^2 \sum_{j=0}^{n-1}(\sin\frac{2\pi ij}{n})^2$, respectively. Therefore $\bar{\sigma}_i^2 + \tilde{\sigma}_i^2 = n\sigma^2$ and thus $\sigma\sqrt{n/\sqrt{2}} \leq \max\{|\bar{\sigma}_i|, |\tilde{\sigma}_i|\} \leq \sigma\sqrt{n}$ for all i. Now note that $|u_i|/2 \leq \max\{|\bar{u}_i|, |\tilde{u}_i|\} \leq |u_i|$ and arrive at the desired Gaussian random variables s_i and t_i satisfying bounds (8.1) for all positive i and having the mean zero and the variances $\sigma^2 n/\sqrt{2}$ and $4\sigma^2 n$, respectively. Therefore $|\Phi_{0,\sigma\sqrt{n/\sqrt{2}}}(y) \leq F_{|u_i|}(y) \leq \Phi_{0,2\sigma\sqrt{n}}(y)$ for all positive i. Combine these bounds with relationships (8.2)–(8.4) and obtain the theorem. $\square$

Corollary 8.1. *Under the assumptions of Theorem 8.1 we have*

a) $F_{||A||}(y) \geq 1 - \sqrt{\frac{2}{\pi}}\left(\frac{1}{a+b}\exp(-\frac{(a+2b)a}{2}) + \frac{1}{c}\exp(-\frac{c^2}{2})\right)$

for $y \geq \max\{2(a+b)\sigma\sqrt{n}, |\mu| + c\sigma\sqrt{n}\}$, $a+b \geq 1$, $b = \sqrt{2\ln(n-1)} > 0$, $c \geq 1$,

b) $F_A(y) \leq ((n-1)2^{1/4} + 1)\frac{y}{\sigma}\sqrt{\frac{2}{\pi\, n}}$ for $y \geq 0$.

Proof. Combine Theorem 8.1 with Lemma 2.3 a), e), f). $\square$

Theorem 2.1 implies that $\frac{1}{g(n)}\sigma_j(Z_1(\mathbf{v})) \leq \sigma_j(Z_f(\mathbf{v})) \leq g(n)\sigma_j(Z_1(\mathbf{v}))$ for all vectors $\mathbf{v}$, scalars f, and $j = 1,2,\ldots,n$ provided $g(n) = \max\{1, |f|^{2n}\}\ \max\{1, \frac{1}{|f|^n}\}$. This enables us to extend the estimates of Theorem 8.1 to f-circulant matrices for $f \neq 0$. In particular these estimates do not change in the case of skew circulant matrices (for which $f = -1$) and show that $n \times n$ f-circulant matrices can be expected to be well conditioned where $\log g(n) = O(\log n)$.

Remark 8.1. *Our extensive experiments (cf. Table 9.2) suggest that the estimates of Theorem 8.1 are rather pessimistic. This is probably because they rely on the bounds of Lemma 2.2, which are crude in this application.*

9 Numerical Experiments

Our numerical experiments with random general, Toeplitz and circulant matrices have been performed in the Graduate Center of the City University of New York. We conducted the tests 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 intrincic Fortran function, assuming the uniform probability distribution over the range $\{x : -1 \leq x < 1\}$.

9.1 Conditioning tests

We computed the condition numbers of $n \times n$ random general, Toeplitz, and circulant matrices for $n = 2^k$, $k = 5, 6, \ldots$, with the entries sampled in the range $[-1, 1)$. We also performed the same tests with complex general matrices whose entries had real and imaginary parts sampled at random in the same range $[-1, 1)$. We computed the values $\mathrm{cond}_1 T = ||T||_1\ ||T^{-1}||_1$ for Toeplitz matrices T and the values $\mathrm{cond}\, M = ||M||\ ||M^{-1}||$ for general and circulant matrices M. We performed $m = 100$ tests for each dimension n and represented the test results in Tables 9.1–9.3.

9.2 Solution of linear systems of equations with multiplicative preconditioning

We solved 1000 linear systems of equations $M\mathbf{x} = \mathbf{b}$ for each input class for vectors $\mathbf{b}$ with random coordinates from the range $[-1, 1)$ and matrices M specified below.

Input Matrices:

$M = \begin{pmatrix} M_k & A \\ B & C \end{pmatrix}$ is an $n \times n$ matrix, M_k is a $k \times k$ matrix, A, B, and C are random Toeplitz matrices such that $||A|| \approx ||B|| \approx ||C|| \approx ||M_k|| \approx 1$, $n = 2^s$, $s = 5, 6, 7, 8, 10$, and $k = n/2$. We have chosen the matrices M_k as follows.

1) *General matrices* $M_k = U\Sigma V^H$ where $\Sigma = \mathrm{diag}(\sigma_i)_{i=1}^k$, $\sigma_i = 1$ for $i = 1, \ldots, k - h$, $\sigma_i = 0$ for $i = k - h + 1, \ldots, k$, $h = 4$, and U and V are $k \times k$ random orthonormal matrices, that is the factors Q in the QR factorization of $k \times k$ random matrices where the factor R has positive diagonal entries (cf. [H02, Section 28.3]).

2) *Toeplitz-like matrices M_k with nullity $h = 4$.* $M_k = c(T, TS)$ for random Toeplitz matrices T of size $k \times (k - h)$ and S of size $(k - h) \times h$, for $h = 4$, and for a positive scalar c such that $||M_k|| \approx 1$.

Multiplicative preconditioners:

(a) $n \times n$ circulant multipliers, each defined by its first column with the n entries $+1$ and -1 chosen at random,

(b) Householder multipliers $H = \prod_{i=1}^h H_i$, $H_i = I - 2\mathbf{v}_i\mathbf{v}_i^T/\mathbf{v}_i^T\mathbf{v}_i$, with the vectors $\mathbf{v}_i$ filled with the integers $+1$ and -1 chosen at random for all i.

Tables

Tables 9.4 and 9.5 show the test results for the solution of linear systems where we apply GENP with randomized structured preconditioning. For all tests the tables display min, max, and average values of the relative residual norm $||M\mathbf{x} - \mathbf{b}||/||\mathbf{b}||$ and the standard deviations. For GENP with preconditioning we show these data before we performed iterative refinement and after the first and sometimes also the third step of it. We continued iterative refinement until we decreased the output residual norms to the level of 10^{-14} (achieved by GEPP). The columns **iterations** in our tables show the respective numbers of steps of iterative refinement.

Due to the singularity of the leading block M_k, the relative residual norms in GENP without preconditioning stayed in the range $[10, 10^8]$ and were too large to allow iterative refinement. With our randomized structured preconditioning, however, these norms were always small enough to allow rapid iterative refinement to the level achieved in GEPP.

Our tests have also showed that with the growth of the input size the growth of the relative norm was limited.

We repeated our computations where the input entries of the matrix M were truncated to the single precision. In this case double-precision multipication by our multipliers was error-free. The residual norms remained essentially the same as for the double precision input. This showed that the rounding errors have made all their impact at the elimination stage rather than at the stage of multiplication by preconditioners.

9.3 Computation of nmbs

Input Matrices:

We first defined auxiliary matrices $\widehat{M} = \begin{pmatrix} M_k & A \\ B & C \end{pmatrix}$ where the blocks M_k, A, B, and C were defined as in Section 9.2 but for n replaced by $n = 2^s - r$, $r = 4$. Then we defined the input matrices $M = \begin{pmatrix} \widehat{M} & E \\ G & K \end{pmatrix}$ where the block G was a random $r \times (2^s - r)$ Toeplitz matrix, $\begin{pmatrix} E \\ K \end{pmatrix} = \begin{pmatrix} \widehat{M} \\ G \end{pmatrix} T$, and T was a random $(n - r) \times r$ Toeplitz matrix.

We generated **multiplicative preconditioners** in the same way as in Section 9.2. Tables 9.6 and 9.7 display the results of the computation of nmbs for the matrices M based on GJENP with weakly random preconditioning (cf. (5.2)). The same format as in Tables 9.4 and 9.5 is used, except that now the tables display the data on the relative residual norms $||MB||/(||M||||B||)$.

Table 9.1: condition numbers of random matrices

size	**type**	**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 9.2: condition numbers $\mathrm{cond}_1 T$ of random Toeplitz matrices T

size	**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

References

[AHU74] A. V. Aho, J. E. Hopcroft, J. D. Ullman, *The Design and Analysis of Algorithms.* Addison-Wesley, Reading, MA, 1974.

[B85] J. R. Bunch, Stability of Methods for Solving Toeplitz Systems of Equations, *SIAM J. Sci. Stat. Comput.*, **6(2)**, 349–364, 1985.

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

[CPV04] G. Codevico, V. Y. Pan, M. Van Barel, Newton-like Iteration Based on Cubic Polynomials for Structured Matrices, *Numerical Algorithms*, **36**, 365-380, 2004.

Table 9.3: condition numbers of random circulant matrices

size	**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

[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.

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

[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 Teory, 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.

[GI88] W. Gautschi, G. Inglese, Lower Bounds for the Condition Number of Vandermonde Matrices, *Numerische Math.*, **52**, 241–250, 1988.

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

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

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

[KVB99] P. Kravanja, M. Van Barel, Algorithms for Solving Rational Interpolation Problems Related to Fast and Superfast Solvers for Toeplitz Systems, *SPIE*, 359–370, 1999.

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

[P90] V. Y. Pan, On Computations with Dense Structured Matrices, *Math. of Computation*, **55(191)**, 179–190, 1990. Proceedings version in *Proc. of Intern. Symposium on Symbolic and Algebraic Computation (ISSAC89)*, 34-42, ACM Press, New York, 1989.

Table 9.4: residual norms of the solutions of Toeplitz-like linear systems

multiplier	**size**	**iterations**	**min**	**max**	**mean**	**std**
circulant	32	0	7.8×10^{-15}	1.6×10^{-10}	3.6×10^{-12}	1.8×10^{-11}
circulant	32	1	8.3×10^{-16}	5.7×10^{-12}	7.4×10^{-14}	5.7×10^{-13}
circulant	64	0	5.9×10^{-14}	1.6×10^{-9}	2.4×10^{-11}	1.6×10^{-10}
circulant	64	1	1.7×10^{-15}	7.3×10^{-13}	4.9×10^{-14}	1.2×10^{-13}
circulant	128	0	3.1×10^{-13}	1.9×10^{-8}	3.5×10^{-10}	2.1×10^{-9}
circulant	128	1	5.2×10^{-15}	1.3×10^{-10}	1.6×10^{-12}	1.3×10^{-11}
circulant	256	0	2.7×10^{-12}	3.6×10^{-9}	1.7×10^{-10}	4.6×10^{-10}
circulant	256	1	8.8×10^{-15}	2.8×10^{-12}	1.6×10^{-13}	3.5×10^{-13}
circulant	1024	0	4.0×10^{-10}	3.8×10^{-9}	1.5×10^{-9}	1.5×10^{-9}
circulant	1024	1	1.2×10^{-13}	5.1×10^{-13}	2.3×10^{-13}	1.9×10^{-13}
Householder	32	0	4.9×10^{-11}	1.8×10^{-7}	5.8×10^{-9}	2.0×10^{-8}
Householder	32	1	4.6×10^{-16}	1.1×10^{-13}	6.1×10^{-15}	1.7×10^{-14}
Householder	64	0	3.2×10^{-10}	9.2×10^{-7}	4.1×10^{-8}	1.3×10^{-7}
Householder	64	1	8.3×10^{-16}	1.4×10^{-12}	2.8×10^{-14}	1.5×10^{-13}
Householder	128	0	2.6×10^{-9}	9.9×10^{-6}	2.6×10^{-7}	1.1×10^{-6}
Householder	128	1	1.4×10^{-15}	1.5×10^{-10}	1.5×10^{-12}	1.5×10^{-11}
Householder	128	3	1.5×10^{-15}	1.5×10^{-12}	3.1×10^{-14}	1.5×10^{-13}
Householder	256	0	1.1×10^{-8}	2.7×10^{-5}	8.8×10^{-7}	3.2×10^{-6}
Householder	256	1	2.8×10^{-15}	4.8×10^{-11}	8.2×10^{-13}	5.0×10^{-12}
Householder	256	3	2.8×10^{-15}	1.3×10^{-12}	5.9×10^{-14}	1.8×10^{-13}
Householder	1024	0	5.0×10^{-6}	3.4×10^{-5}	1.6×10^{-5}	1.3×10^{-5}
Householder	1024	1	1.6×10^{-12}	1.2×10^{-11}	6.7×10^{-12}	4.7×10^{-12}
Householder	1024	3	1.9×10^{-14}	1.7×10^{-13}	7.1×10^{-14}	6.5×10^{-14}

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

[P08] V. Y. Pan, Newton's Iteration for Matrix Inversion, Advances and Extensions, in *Matrix methods: Theory, Algorithms and Applications* (dedicated to the Memory of Gene H. Golub, edited by Vadim Olshevsky and Eugene Tyrtyshnikov), pages 364–381, World Scientific Publishing, 2008.

[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.

[PKRK06] V. Y. Pan, M. Kunin, R. E. Rosholt, H. Kodal, Homotopic Residual Correction Processes, *Math. of Computation*, **75**, 345–368, 2006.

[PQa] V. Y. Pan, G. Qian, Randomized Preprocessing of Homogeneous Linear Systems of Equations, Technical Report TR 2009008, *Ph.D. Program in Computer Science*, *Graduate Center, the City University of New York*, 2009.

Available at http://www.cs.gc.cuny.edu/tr/techreport.php?id=352

[PQb] V. Y. Pan, G. Qian, Solving Linear Systems with Randomized Augmentation, Technical Report TR 2009009, *Ph.D. Program in Computer Science*, *Graduate Center, the City University of New York*, 2009.

Available at http://www.cs.gc.cuny.edu/tr/techreport.php?id=352

Table 9.5: residual norms of the solutions of linear systems in the case of general matrices M_k

multiplier	**size**	**iterations**	**min**	**max**	**mean**	**std**
circulant	32	0	9.1×10^{-15}	5.4×10^{-12}	5.7×10^{-13}	1.1×10^{-12}
circulant	32	1	1.1×10^{-15}	7.8×10^{-14}	9.7×10^{-15}	1.4×10^{-14}
circulant	64	0	4.7×10^{-14}	8.0×10^{-11}	4.0×10^{-12}	1.1×10^{-11}
circulant	64	1	1.9×10^{-15}	5.3×10^{-13}	2.3×10^{-14}	5.4×10^{-14}
circulant	128	0	2.8×10^{-13}	2.1×10^{-10}	1.6×10^{-11}	3.1×10^{-11}
circulant	128	1	4.3×10^{-15}	1.6×10^{-12}	6.6×10^{-14}	1.8×10^{-13}
circulant	256	0	1.7×10^{-12}	1.4×10^{-7}	2.0×10^{-9}	1.5×10^{-8}
circulant	256	1	8.3×10^{-15}	4.3×10^{-10}	4.5×10^{-12}	4.3×10^{-11}
circulant	1024	0	1.7×10^{-10}	4.4×10^{-9}	1.4×10^{-9}	2.1×10^{-9}
circulant	1024	1	3.4×10^{-14}	9.9×10^{-14}	6.8×10^{-14}	2.7×10^{-14}
Householder	32	0	5.5×10^{-15}	2.3×10^{-11}	1.0×10^{-12}	3.1×10^{-12}
Householder	32	1	4.1×10^{-16}	1.8×10^{-13}	4.5×10^{-15}	1.8×10^{-14}
Householder	64	0	2.9×10^{-14}	1.8×10^{-10}	3.6×10^{-12}	1.9×10^{-11}
Householder	64	1	5.8×10^{-16}	3.6×10^{-13}	9.0×10^{-15}	3.8×10^{-14}
Householder	128	0	1.2×10^{-13}	9.1×10^{-10}	2.4×10^{-11}	1.0×10^{-10}
Householder	128	1	1.2×10^{-15}	4.8×10^{-13}	1.9×10^{-14}	6.0×10^{-14}
Householder	256	0	1.1×10^{-12}	3.2×10^{-8}	4.5×10^{-10}	3.2×10^{-9}
Householder	256	1	2.0×10^{-15}	6.4×10^{-13}	2.7×10^{-14}	8.2×10^{-14}
Householder	1024	0	3.2×10^{-11}	2.7×10^{-9}	8.6×10^{-10}	1.3×10^{-9}
Householder	1024	1	1.6×10^{-14}	9.5×10^{-14}	4.3×10^{-14}	3.7×10^{-14}

[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.

[PY] V. Y. Pan, X. Yan, Additive Preconditioning, Eigenspaces, and the Inverse Iteration, *Linear Algebra and Its Applications*, **430**, 186–203, 2009.

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

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

[VB99] M. Van Barel, A Supefast 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 Journal on Matrix Analysis and Applications*, **23**, **2**, 494–510, 2001.

[VBK98] M. Van Barel, P. Kravanja, A Stabilized Superfast Solver for Indefinite Hankel Systems, *Linear Algebra and its Applications*, **284**, **1–3**, 335–355, 1998.

[W04] M. Wschebor, Smoothed Analysis of $\kappa(a)$, *J. of Complexity*, **20**, 97–107, 2004.

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

Table 9.6: residual norms of approximate nmbs in the case of Toeplitz-like matrices M_k

multiplier	size	iterations	min	max	mean	std
circulant	32	0	4.4×10^{-13}	6.0×10^{-9}	2.0×10^{-10}	7.9×10^{-10}
circulant	32	1	4.4×10^{-17}	2.3×10^{-15}	3.2×10^{-16}	3.9×10^{-16}
circulant	64	0	6.1×10^{-13}	6.2×10^{-9}	2.3×10^{-10}	8.9×10^{-10}
circulant	64	1	3.9×10^{-17}	3.1×10^{-15}	3.0×10^{-16}	4.2×10^{-16}
circulant	128	0	1.9×10^{-13}	3.8×10^{-9}	1.3×10^{-10}	4.7×10^{-10}
circulant	128	1	3.6×10^{-17}	2.5×10^{-15}	3.7×10^{-16}	4.2×10^{-16}
circulant	256	0	1.7×10^{-12}	1.4×10^{-8}	3.8×10^{-10}	1.8×10^{-9}
circulant	256	1	2.1×10^{-17}	2.1×10^{-14}	5.8×10^{-16}	2.2×10^{-15}
circulant	1024	0	9.3×10^{-12}	5.4×10^{-11}	3.2×10^{-11}	2.2×10^{-11}
circulant	1024	1	1.7×10^{-16}	2.2×10^{-16}	2.0×10^{-16}	2.2×10^{-17}
circulant	1024	3	6.2×10^{-17}	4.7×10^{-16}	2.0×10^{-16}	1.8×10^{-16}
Householder	32	0	3.5×10^{-15}	2.6×10^{-5}	2.7×10^{-7}	2.6×10^{-6}
Householder	32	1	1.1×10^{-20}	1.9×10^{-9}	1.9×10^{-11}	1.9×10^{-10}
Householder	32	3	1.9×10^{-20}	5.5×10^{-15}	2.4×10^{-16}	6.6×10^{-16}
Householder	64	0	3.3×10^{-14}	4.3×10^{-7}	1.4×10^{-8}	5.1×10^{-8}
Householder	64	1	1.7×10^{-20}	8.1×10^{-15}	5.7×10^{-16}	1.4×10^{-15}
Householder	128	0	7.5×10^{-14}	4.1×10^{-7}	1.5×10^{-8}	4.7×10^{-8}
Householder	128	1	1.7×10^{-20}	4.2×10^{-14}	1.1×10^{-15}	4.8×10^{-15}
Householder	256	0	9.7×10^{-14}	1.4×10^{-6}	3.7×10^{-8}	1.6×10^{-7}
Householder	256	1	3.5×10^{-20}	1.6×10^{-13}	3.4×10^{-15}	1.7×10^{-14}
Householder	1024	0	1.8×10^{-11}	1.2×10^{-6}	3.3×10^{-7}	5.6×10^{-7}
Householder	1024	1	6.0×10^{-18}	1.3×10^{-12}	3.2×10^{-13}	6.4×10^{-13}
Householder	1024	3	1.5×10^{-19}	1.0×10^{-14}	3.2×10^{-15}	4.6×10^{-15}

Table 9.7: residual norms of approximate nmbs in the case of general matrices M_k

multiplier	size	iterations	min	max	mean	std
circulant	32	0	6.4×10^{-13}	8.0×10^{-9}	2.1×10^{-10}	9.9×10^{-10}
circulant	32	1	4.5×10^{-17}	1.8×10^{-15}	2.3×10^{-16}	2.6×10^{-16}
circulant	64	0	9.4×10^{-13}	5.3×10^{-9}	2.0×10^{-10}	7.5×10^{-10}
circulant	64	1	2.4×10^{-17}	4.1×10^{-15}	3.9×10^{-16}	5.9×10^{-16}
circulant	128	0	1.1×10^{-12}	2.3×10^{-9}	1.2×10^{-10}	3.0×10^{-10}
circulant	128	1	3.6×10^{-17}	3.8×10^{-15}	3.7×10^{-16}	6.1×10^{-16}
circulant	256	0	2.2×10^{-12}	5.3×10^{-8}	9.8×10^{-10}	6.1×10^{-9}
circulant	256	1	4.4×10^{-17}	6.2×10^{-15}	4.8×10^{-16}	8.6×10^{-16}
circulant	1024	0	1.0×10^{-11}	1.1×10^{-9}	3.3×10^{-10}	5.4×10^{-10}
circulant	1024	1	1.9×10^{-16}	6.3×10^{-16}	3.5×10^{-16}	2.0×10^{-16}
Householder	32	0	1.8×10^{-16}	3.3×10^{-11}	6.4×10^{-13}	3.4×10^{-12}
Householder	32	1	5.7×10^{-21}	1.8×10^{-17}	6.1×10^{-19}	2.3×10^{-18}
Householder	64	0	1.7×10^{-16}	2.9×10^{-11}	6.3×10^{-13}	3.2×10^{-12}
Householder	64	1	5.6×10^{-21}	5.7×10^{-17}	1.5×10^{-18}	6.7×10^{-18}
Householder	128	0	2.4×10^{-16}	9.9×10^{-11}	1.5×10^{-12}	1.0×10^{-11}
Householder	128	1	3.5×10^{-21}	6.9×10^{-17}	2.3×10^{-18}	7.6×10^{-18}
Householder	256	0	5.2×10^{-16}	5.8×10^{-11}	1.1×10^{-12}	5.9×10^{-12}
Householder	256	1	3.0×10^{-21}	9.4×10^{-17}	4.0×10^{-18}	1.1×10^{-17}
Householder	1024	0	6.3×10^{-14}	1.2×10^{-12}	5.6×10^{-13}	4.5×10^{-13}
Householder	1024	1	1.9×10^{-18}	4.1×10^{-17}	1.9×10^{-17}	1.8×10^{-17}