Probabilistic Analysis of Complex Gaussian Elimination without Pivoting

Man-Chung Yeung [1]

Dept. 3036, 1000 East University Avenue, Laramie, WY 82071

Abstract

We consider Gaussian elimination without pivoting applied to complex Gaussian matrices $\mathbf{X} \in \mathbb{C}^{n \times n}$. We first study some independence properties of the elements of the LU factors of $\mathbf{X}$. Based on this, we then derive the probability distributions for all the $\mathbf{L}$ and $\mathbf{U}$ elements and obtain bounds for the probabilities of the occurrence of small pivots and large growth factors. Numerical experiments are presented to support the theoretical results and discussions are made to relate the results to the crucial practical problems of numerical stability of GE.

Key words: Gaussian matrix, Gaussian elimination, pivot, growth factor, density function

AMS classification: 65F05, 65G05

1 Introduction

Gaussian elimination (GE) is the most common general method for solving an $n \times n$, square, dense, unstructured linear system $\mathbf{Ax} = \mathbf{b}$. Together with partial pivoting, the method is extremely stable in practice (see, for instance, [18, p.166]) though some reports on matrices that cause GE with pivoting unstable have been published in literature (see, for example, Foster[7] and Wright[28]).

Theoretical studies of the numerical stability of GE have been made since the 1940s by a number of authors, including Turing [21], von Neumann and Goldstine [23,24], Wilkinson [25,27]. See Higham [11] for a detailed summary

Email address: myeung@uwyo.edu (Man-Chung Yeung).

[1] The work was supported in part by NSF grant DMS-0314152.

of the historical development in this topic. In [19], Trefethen and Schreiber considered the average case analysis of GE with pivoting in order to explain its practical numerical stability. Among their many experimental results, they observed that for many distributions of matrices, the matrix elements after the first few steps of Gaussian elimination with (partial or complete) pivoting are approximately normally distributed. They also found that, for $n \leq 1024$, the average growth factor (normalized by the standard deviation of the initial matrix elements) is within a few percent of $n^{2/3}$ for the partial pivoting case and approximately $n^{1/2}$ for the complete pivoting case. Recently, Spielman and Teng[17] introduced a new method called *smoothed analysis* to analyze stabilities of algorithms. *Smoothed analysis* is a hybrid of the traditional worst-case and average-case analysis, and has been applied to the stability analysis of simplex algorithm of linear programming problems[17] and GE without pivoting[15] and to other areas.

Following Trefethen and Schreiber, Yeung and Chan [29] studied the probability of small pivots and large growth factors with GE without pivoting. One reason that they studied the non-pivoting case is that, with the advent of parallel computing, there is incentive to trade off the stability of partial or complete pivoting for the higher performance of simpler but possibly less stable forms of GE, including no pivoting, since pivoting can significantly complicate the algorithm, increase data movement, and reduce speed[13,14]. Another reason is quite obvious: the non-pivoting case is far easier to analyze than the pivoting case.

In [29], Yeung and Chan supposed $\mathbf{A} \in \mathbb{R}^{n \times n}$ is a real Gaussian matrix. This choice was motivated by the empirical results of Trefethen and Schreiber mentioned earlier. Matrices of this type have also been studied by Edelman [3,5], who derived exact formulas for the probability distribution of the singular values and for the expected singular values. Yeung and Chan then derived, when GE without pivoting is performed on $\mathbf{A}$, the distributions of the elements of $\mathbf{L}$ and $\mathbf{U}$, where $\mathbf{A} = \mathbf{LU}$ is the LU factorization of $\mathbf{A}$. In [12], Olkin has a deeper insight into the distributions of $\mathbf{L}$ and $\mathbf{U}$ where he gave the joint probability density function of $\mathbf{L}$ and $\mathbf{U}$. Based on the distributions obtained of elements of the LU factors, Yeung and Chan [29] proved that (i) the probability of a pivot less than ϵ in magnitude is $O(\epsilon)$; (ii) the probabilities $Prob\left(\|\mathbf{L}\|_\infty > r\right) = O(n^3/r)$ and $Prob\left(\|\mathbf{U}\|_\infty / \|\mathbf{A}\|_\infty\right) = O\left(n^{2.5}/r\right)$. The result (ii) has been significantly improved by Sankar, Spielman and Teng[15] by considering a bunch of elements simultaneously, instead of element by element as was done in [29]. In this paper, we are motivated by their treatment of elements.

In this paper, we continue the work of [29] on GE without pivoting. This time, we assume $\mathbf{A} \in \mathbb{C}^{n \times n}$ is a complex Gaussian matrix. The techniques used to obtain results are similar — starting from Schur complement, then simplifying the expression by repeatedly using the property that normal distribution does not change under an orthogonal transformation. In §3 and §4, we explore the independence among the elements of the LU factors. We then derive the density functions for all the elements. In §5, we discuss the probability of small pivots and prove that the probability of the occurrence of a pivot less than ϵ in magnitude is $O(\epsilon^2)$. [2] In §6, we derive bounds on the probabilities of large growth factors. In particular, we show that $Prob\,(\|\mathbf{L}\|_\infty > r) = O\,(n^2(\ln r)^2/r^2)$ and $Prob\,(\|\mathbf{U}\|_\infty/\|\mathbf{A}\|_\infty > r) = O\,(n^2/r^2)$, which further imply that the expected values $E(\|\mathbf{L}\|_\infty) = O(n \ln n)$ and $E(\|\mathbf{U}\|_\infty/\|\mathbf{A}\|_\infty) = O(n)$. Techniques in [15] have been borrowed in the proofs, incorporating with the independence properties obtained of the LU factors in §3 and §4, and the concrete probability distributions of elements. In §7, we present experimental results. We observed that the probability density distributions of $\|\mathbf{L}\|_\infty$ and $\|\mathbf{U}\|_\infty/\|\mathbf{A}\|_\infty$ decay algebraically — a huge difference from that the probability density of the growth factor with partial pivoting decays exponentially[18, p.168]. In §8, discussions are made on relating the theoretical results to the crucial practical problems of numerical stability.

We hope that the series of results for GE without pivoting presented in [29] and in this paper may be useful in the analysis of, as well as providing a basis of comparison for, the partial pivoting case.

2 Notation

Throughout the paper, we use a boldface capital letter to denote a matrix (e.g., $\mathbf{X}$, $\mathbf{\Phi}$, $\hat{\mathbf{\Phi}}$) and the corresponding lower case letter with subscript ij refers to the (i, j) entry of the matrix (e.g., $x_{ij}, \phi_{ij}, \hat{\phi}_{ij}$). To refer to a portion of a matrix $\mathbf{X}$, we use our stylized MATLAB notation. For example:

- $\mathbf{X}_{(p,:)}$ — the p-th row of $\mathbf{X}$.
- $\mathbf{X}_{(p_1:p_2,q_1:q_2)}$ — the block of elements x_{ij} with $p_1 \le i \le p_2$ and $q_1 \le j \le q_2$.

A vector is denoted by a boldface lower case letter (e.g., $\mathbf{a}, \mathbf{z}$) and the corresponding lower case letter with subscript i refers to the i-th component of the vector (e.g., a_i, z_i).

For a complex variable v, its real and imaginary parts are denoted by v^R and v^I respectively.

[2] We note that Foster [6] has studied the probability of large diagonal elements in the QR factorization of a rectangular matrix $\mathbf{A}$.

3 Density Functions of u_{pq}

Let $N(0,1)$ denote the normal distribution with mean 0 and variance 1 and $\tilde{N}(0,2)$ the complex distribution of $x+yi$ where x and y are independent and identically distributed $N(0,1)$. By definition, a complex Gaussian matrix is a random matrix with independent and identically distributed elements from $\tilde{N}(0,2)$ [4].

Let $\mathbf{X}$ be an $n \times n$ complex Gaussian matrix and let $\mathbf{X} = \mathbf{LU}$, where $\mathbf{L}$ is a unit lower triangular matrix and $\mathbf{U}$ an upper triangular matrix, be the *LU* factorization of $\mathbf{X}$.[3] The (p,q)-th ($p \le q$) element u_{pq} of $\mathbf{U}$ and the elements of $\mathbf{X}$ have the following relation (see, for instance, [29]).

Lemma 1 *Let* $\mathbf{X} = \mathbf{LU}$ *be the LU factorization of* $\mathbf{X}$. *Then*

$$u_{pq} = x_{pq} - \mathbf{X}_{(p,1:p-1)}\mathbf{X}^{-1}_{(1:p-1,1:p-1)}\mathbf{X}_{(1:p-1,q)} \, .$$

The upper triangular segment $\mathbf{U}_{(p,p:n)}$ of the p-th row of $\mathbf{U}$ can therefore be written as

$$\mathbf{U}_{(p,p:n)} = \mathbf{X}_{(p,p:n)} - \mathbf{X}_{(p,1:p-1)}\mathbf{X}^{-1}_{(1:p-1,1:p-1)}\mathbf{X}_{(1:p-1,p:n)} \, . \qquad (1)$$

Proposition 2 *Suppose* $\mathbf{X}$ *is an* $n \times n$ *complex Gaussian matrix and let* $\mathbf{X} = \mathbf{LU}$ *be the LU factorization of* $\mathbf{X}$. *Then there exist an* $n \times n$ *diagonal matrix* $\mathbf{\Phi} = diag(\phi_1, \ldots, \phi_n)$ *and an* $n \times n$ *upper triangular matrix* $\mathbf{H} = (\eta_{pq})$ *with the following properties such that* $\mathbf{U} = \mathbf{\Phi}\,\mathbf{H}$:

(i) $\phi_1 = 1$ *and* $\phi_p \ge 0$ *for* $2 \le p \le n$.
(ii) $(\phi_p^2 - 1)/(p-1)$ *has the* $F_{2(p-1),2}$ *distribution for* $2 \le p \le n$,.
(iii) *All* η_{pq} *with* $1 \le p \le q \le n$ *have the* $\tilde{N}(0,2)$ *distribution.*
(iv) $\phi_p, \eta_{pp}, \eta_{pp+1}, \ldots, \eta_{pn}$ *are mutually independent for* $1 \le p \le n$.
(v) η_{pq} *is independent of* $\mathbf{H}_{(1:q-1,1:q-1)}$ *where* $1 \le p \le q \le n$.
(vi) $\eta_{p_1 1}, \eta_{p_2 2}, \ldots, \eta_{p_n n}$ *are mutually independent for any* $p_1, p_2, \ldots, p_n$ *with* $1 \le p_i \le i$ *for* $i = 1, 2, \ldots, n$.

Proof. We first note that the blocks $\mathbf{X}_{(p,p:n)}, \mathbf{X}_{(p,1:p-1)}, \mathbf{X}_{(1:p-1,1:p-1)}$ and $\mathbf{X}_{(1:p-1,p:n)}$ of (1) are disjoint each other, that is, every two of them contain no common element of $\mathbf{X}$. Let $\mathbf{G}$ be an $(p-1)\times(p-1)$ unitary matrix, e.g., a Householder

[3] Since they just form a set of measure zero, we ignore matrices for which *LU* factorization does not exist.

matrix, such that

$$\mathbf{X}_{(p,1:p-1)}\mathbf{G} = (0, \cdots, 0, z_{p-1}) \equiv \mathbf{z}^T \tag{2}$$

with $z_{p-1} \geq 0$. From (1), we then have

$$\begin{aligned} \mathbf{U}_{(p,p:n)} &= \mathbf{X}_{(p,p:n)} - \mathbf{z}^T(\mathbf{X}_{(1:p-1,1:p-1)}\mathbf{G})^{-1}\mathbf{X}_{(1:p-1,p:n)} \\ &\equiv \mathbf{X}_{(p,p:n)} - \mathbf{z}^T\mathbf{Y}^{-1}\mathbf{X}_{(1:p-1,p:n)} \,. \end{aligned} \tag{3}$$

About the last expression of $\mathbf{U}_{(p,p:n)}$, it can be shown that all the x- and y-elements identically have the $\tilde{N}(0, 2)$ distribution while z_{p-1}^2 has the $\chi^2_{2(p-1)}$ distribution. They are mutually independent. The proof basically follows the approaches used in [4], [16] and [20].

We now decompose $\mathbf{Y}$ as

$$\mathbf{Y} = \mathbf{Q}\,\mathbf{R}$$

where $\mathbf{Q}$ is an $(p-1)\times(p-1)$ unitary matrix and $\mathbf{R}$ an $(p-1)\times(p-1)$ upper triangular matrix with positive diagonal elements. We then further have

$$\begin{aligned} \mathbf{U}_{(p,p:n)} &= \mathbf{X}_{(p,p:n)} - \mathbf{z}^T\mathbf{R}^{-1}\mathbf{Q}^{-1}\mathbf{X}_{(1:p-1,p:n)} \\ &\equiv \mathbf{X}_{(p,p:n)} - \mathbf{z}^T\mathbf{R}^{-1}\mathbf{W} = \mathbf{X}_{(p,p:n)} - \frac{z_{p-1}}{r_{p-1p-1}}\mathbf{W}_{(p-1,:)} \end{aligned} \tag{4}$$

where $\mathbf{W} = (w_{ij})$ is an $(p-1)\times(n-p+1)$ matrix. In the last equation of (4), all the x- and w-elements identically have the $\tilde{N}(0, 2)$ distribution while r_{p-1p-1}^2 and z_{p-1}^2 have the χ_2^2 and the $\chi^2_{2(p-1)}$ distributions respectively. They are mutually independent. Again the proof follows the approaches used in [4], [16] and [20].

We now set

$$\delta_p = \frac{1}{p-1}\frac{z_{p-1}^2}{r_{p-1p-1}^2} = \frac{z_{p-1}^2/2(p-1)}{r_{p-1p-1}^2/2} \,. \tag{5}$$

and define the $\mathbf{H}$ and $\mathbf{\Phi}$ in the proposition as

$$\mathbf{H}_{(p,p:n)} = \begin{cases} \mathbf{X}_{(1,:)} & \text{if } p = 1 \\ \dfrac{1}{\sqrt{1+(p-1)\delta_p}}\mathbf{X}_{(p,p:n)} - \dfrac{\sqrt{(p-1)\delta_p}}{\sqrt{1+(p-1)\delta_p}}\mathbf{W}_{(p-1,:)} & \text{if } p \geq 2\,, \end{cases}$$

$$\phi_p = \begin{cases} 1 & \text{if } p = 1 \\ \sqrt{1 + (p-1)\delta_p} & \text{if } p \geq 2\,. \end{cases}$$

Then

$$\mathbf{U}_{(p,p:n)} = \phi_p \, \mathbf{H}_{(p,p:n)}\,.$$

The variable δ_p has the $F_{2(p-1),2}$ distribution and all the η_{pk} of $\mathbf{H}_{(p,p:n)}$ for $k = p, \ldots, n$ have the $\tilde{N}(0,2)$ distribution. Again, they are mutually independent. Thus, the proof of properties (i)-(iv) is completed.

To prove property (v), we need Lemma 11 in the Appendix section. By definition, $\eta_{1q} = x_{1q}$ and

$$\eta_{pq} = \frac{1}{\sqrt{1 + (p-1)\delta_p}} x_{pq} - \frac{\sqrt{(p-1)\delta_p}}{\sqrt{1 + (p-1)\delta_p}} w_{p-1\,q-p+1} \tag{6}$$

for $p \geq 2$, where $w_{p-1\,q-p+1}$ is the $(p-1, q-p+1)$-th element of $\mathbf{W}$. Since z_{p-1} is a function of $\mathbf{X}_{(p,1:p-1)}$ by (2) and r_{p-1p-1} a function of $\mathbf{X}_{(1:p,1:p-1)}$ by (3), we have that δ_p is a function of $\mathbf{X}_{(1:p,1:p-1)}$. Moreover, since $\mathbf{W} = \mathbf{Q}^{-1}\mathbf{X}_{(1:p-1,p:n)}$ by (4) where $\mathbf{Q}$ is a unitary matrix whose elements are functions of $\mathbf{X}_{(1:p,1:p-1)}$, $w_{p-1\,q-p+1}$ can be expressed as

$$w_{p-1\,q-p+1} = \theta_1 x_{1q} + \theta_2 x_{2q} + \ldots + \theta_{p-1} x_{p-1q} \tag{7}$$

where θ_i's are functions of $\mathbf{X}_{(1:p,1:p-1)}$ with $\sum_{i=1}^{p-1} |\theta_i|^2 = 1$. Substituting (7) into (6), we have

$$\eta_{pq} = \hat{\theta}_1 x_{1q} + \hat{\theta}_2 x_{2q} + \ldots + \hat{\theta}_p x_{pq}$$

for some $\hat{\theta}$'s with $\sum_{i=1}^{p} |\hat{\theta}_i|^2 = 1$. Each of the $\hat{\theta}$'s is a function of $\mathbf{X}_{(1:p,1:p-1)}$. Now that all the upper triangular elements of $\mathbf{H}_{(1:q-1,1:q-1)}$ are functions of the block $\mathbf{X}_{(1:q-1,1:q-1)}$ which does not contain any of $x_{1q}, x_{2q}, \ldots, x_{pq}$, we have by Lemma 11

$$\begin{aligned} &Prob\left(\eta_{ij}^R < \alpha_{ij}, \eta_{ij}^I < \beta_{ij}, 1 \leq i \leq j \leq q-1, \eta_{pq}^R < \alpha, \eta_{pq}^I < \beta\right) \\ &= Prob\left(\eta_{ij}^R < \alpha_{ij}, \eta_{ij}^I < \beta_{ij}, 1 \leq i \leq j \leq q-1\right) Prob\left(\eta_{pq}^R < \alpha\right) Prob\left(\eta_{pq}^I < \beta\right) \end{aligned}$$

where α's and β's are any real numbers. Therefore, property (v) is proved.

The mutual independence of $\eta_{p_1 1}, \eta_{p_2 2}, \ldots, \eta_{p_n n}$ can be obtained by applying

Lemma 11 repeatedly to the following probability

$$\begin{aligned}
&Prob\left(\eta^R_{p_i i} < \alpha_i, \eta^I_{p_i i} < \beta_i, i = 1, \ldots, n\right)\\
= &Prob\left(\eta^R_{p_i i} < \alpha_i, \eta^I_{p_i i} < \beta_i, i = 1, \ldots, n-1\right)\ Prob\left(\eta^R_{p_n n} < \alpha_n\right)\ Prob\left(\eta^I_{p_n n} < \beta_n\right)\\
= &\ldots = \prod_{i=1}^n Prob\left(\eta^R_{p_i i} < \alpha_i\right)\ Prob\left(\eta^I_{p_i i} < \beta_i\right)
\end{aligned}$$

where α_i's and β_i's are any real numbers. □

With the help of Proposition 2, we are ready to derive the density functions of the elements of $\mathbf{U}$.

Theorem 3 *Suppose* $\mathbf{X}$ *is an* $n \times n$ *complex Gaussian matrix and let* $\mathbf{X} = \mathbf{LU}$ *be the LU factorization of* $\mathbf{X}$*. Then the density functions of* u^R_{pq}, u^I_{pq} *and* $|u_{pq}|$*, the real part, imaginary part and absolute value of the* (p, q)*-th* $(2 \le p \le q)$ *element* u_{pq} *of* $\mathbf{U}$ *respectively, are*

$$f_{u^R_{pq}}(t) = f_{u^I_{pq}}(t) = \frac{p-1}{\sqrt{2\pi}} \sum_{i=0}^{\infty} \frac{(-1)^i}{2^i i!} B(p-1, i+3/2)\, t^{2i}\,,$$

where $-\infty < t < \infty$ *and* $B(x, y)$ *is the beta function, and*

$$f_{|u_{pq}|}(t) = (-1)^p 2^p (p-1)!\, t^{1-2p} \left(\exp\left(-\frac{t^2}{2}\right)\left(1 - p - \frac{t^2}{2}\right) - \sum_{i=0}^{p-1} (-1)^i \frac{1-p+i}{2^i i!} t^{2i} \right),$$

where $0 < t$*.*

Proof. By Proposition 2, $u_{pq} = \phi_p \eta_{pq}$ where $\phi_p (\ge 0)$ and η_{pq} are independent and $(\phi_p^2 - 1)/(p-1)$ and η_{pq} have the $F_{2(p-1),2}$ and $\tilde{N}(0, 2)$ distributions respectively. Let $\delta_p = (\phi_p^2 - 1)/(p-1)$, then

$$u_{pq} = \eta_{pq} \sqrt{1 + (p-1)\delta_p}$$

and therefore

$$u^R_{pq} = \eta^R_{pq} \sqrt{1 + (p-1)\delta_p} \quad \text{and} \quad |u_{pq}| = |\eta_{pq}| \sqrt{1 + (p-1)\delta_p}\,.$$

Since δ_p and η^R_{pq} are independent with the $F_{2(p-1),2}$ and the $N(0, 1)$ distributions respectively, their joint density is

$$\begin{aligned}
f(\lambda, \theta) &= f_\delta(\lambda) f_{\eta^R}(\theta)\\
&= \begin{cases} \dfrac{1}{\sqrt{2\pi}} (p-1)^p \lambda^{p-2} (1 + (p-1)\lambda)^{-p} \exp(-\theta^2/2) & \lambda > 0\\ 0 & \text{otherwise}\,. \end{cases}
\end{aligned}$$

Thus the distribution function $F_{u^R_{pq}}(\alpha)$ of u^R_{pq} is given by

$$\begin{aligned}
F_{u^R_{pq}}(\alpha) &= \iint_{u^R_{pq}\le\alpha} f(\lambda,\theta)\,d\lambda\,d\theta \\
&= \iint_{\theta\sqrt{1+(p-1)\lambda}\le\alpha} \frac{1}{\sqrt{2\pi}}(p-1)^p\lambda^{p-2}(1+(p-1)\lambda)^{-p}\exp(-\theta^2/2)\,d\lambda\,d\theta \\
&= \frac{1}{\sqrt{2\pi}}(p-1)^p\int_0^\infty d\lambda \int_{-\infty}^{\alpha/\sqrt{1+(p-1)\lambda}} \lambda^{p-2}(1+(p-1)\lambda)^{-p}\exp(-\theta^2/2)\,d\theta\,.
\end{aligned}$$

Letting $\theta = t/\sqrt{1+(p-1)\lambda}$, we obtain

$$\begin{aligned}
F_{u^R_{pq}}(\alpha) &= \frac{1}{\sqrt{2\pi}}(p-1)^p\int_0^\infty d\lambda\int_{-\infty}^{\alpha} \lambda^{p-2}(1+(p-1)\lambda)^{-p-1/2}\exp\left(-\frac{1}{2}\frac{t^2}{1+(p-1)\lambda}\right)dt \\
&= \frac{1}{\sqrt{2\pi}}(p-1)^p\int_{-\infty}^{\alpha} dt\int_0^\infty \lambda^{p-2}(1+(p-1)\lambda)^{-p-1/2}\exp\left(-\frac{1}{2}\frac{t^2}{1+(p-1)\lambda}\right)d\lambda\,.
\end{aligned}$$

Letting $s = \dfrac{1}{1+(p-1)\lambda}$,

$$F_{u^R_{pq}}(\alpha) = \frac{1}{\sqrt{2\pi}}(p-1)\int_{-\infty}^{\alpha} dt\int_0^1 s^{1/2}(1-s)^{p-2}\exp\left(-\frac{1}{2}t^2 s\right)ds\,.$$

Thus

$$\begin{aligned}
&f_{u^R_{pq}}(t) = \frac{p-1}{\sqrt{2\pi}}\int_0^1 s^{1/2}(1-s)^{p-2}\exp\left(-\frac{1}{2}t^2 s\right)ds \\
&= \frac{p-1}{\sqrt{2\pi}}\sum_{i=0}^\infty \frac{1}{i!}\left(-\frac{1}{2}t^2\right)^i\int_0^1 s^{i+1/2}(1-s)^{p-2}ds = \frac{p-1}{\sqrt{2\pi}}\sum_{i=0}^\infty\frac{(-1)^i}{2^i i!}B(p-1,i+3/2)\,t^{2i}\,.
\end{aligned}$$

To obtain the density of $|u_{pq}|$, we set

$$\zeta = |\eta_{pq}|^2\,.$$

Then

$$|u_{pq}| = \sqrt{\zeta}\sqrt{1+(p-1)\delta_p}\,.$$

ζ has the χ^2_2 distribution and is independent of δ_p. Since the joint density of δ_p and ζ is

$$f(\lambda,\theta) = \begin{cases} \frac{1}{2}(p-1)^p\lambda^{p-2}\,(1+(p-1)\lambda)^{-p}\exp(-\theta/2) & \lambda,\ \theta>0 \\ 0 & \text{otherwise}\,, \end{cases}$$

the distribution function $F_{|u_{pq}|}(\alpha)$ of $|u_{pq}|$ is

$$\begin{aligned}
F_{|u_{pq}|}(\alpha) &= \iint_{|u_{pq}|\le\alpha} f(\lambda,\theta)d\lambda d\theta \\
&= \iint_{\theta(1+(p-1)\lambda)\le\alpha^2} \frac{1}{2}(p-1)^p\lambda^{p-2}\left(1+(p-1)\lambda\right)^{-p}\exp(-\theta/2)\,d\lambda\,d\theta \\
&= \frac{1}{2}(p-1)^p\int_0^\infty d\lambda \int_0^{\alpha^2/(1+(p-1)\lambda)} \lambda^{p-2}\left(1+(p-1)\lambda\right)^{-p}\exp(-\theta/2)d\theta \\
&= (p-1)^p\int_0^\infty \lambda^{p-2}\left(1+(p-1)\lambda\right)^{-p}\left(1-\exp\left(-\frac{1}{2}\frac{\alpha^2}{1+(p-1)\lambda}\right)\right)d\lambda\,.
\end{aligned}$$

Letting $s=\dfrac{1}{1+(p-1)\lambda}$,

$$F_{|u_{pq}|}(\alpha) = 1-(p-1)\int_0^1 (1-s)^{p-2}\exp\left(-\frac{\alpha^2}{2}s\right)ds\,.$$

Therefore,

$$\begin{aligned}
&f_{|u_{pq}|}(t) = F'_{|u_{pq}|}(t) = t(p-1)\int_0^1 (1-s)^{p-2}s\exp\left(-\frac{t^2}{2}s\right)ds \\
&= t(p-1)\int_0^1(1-s)^{p-2}s\sum_{i=0}^\infty(-1)^i\frac{t^{2i}}{2^i}\frac{1}{i!}s^i ds = t(p-1)\sum_{i=0}^\infty(-1)^i\frac{t^{2i}}{2^i}\frac{1}{i!}\int_0^1(1-s)^{p-2}s^{i+1}ds \quad (8)\\
&= t(p-1)\sum_{i=0}^\infty(-1)^i\frac{t^{2i}}{2^i}\frac{1}{i!}\frac{(p-2)!(i+1)!}{(i+p)!} = t(p-1)!\sum_{i=0}^\infty(-1)^i\frac{t^{2i}}{2^i}\frac{i+1}{(i+p)!}\,.
\end{aligned}$$

Since

$$\begin{aligned}
&\sum_{i=0}^\infty x^i\frac{i+1}{(i+p)!} = \frac{d}{dx}\left(\sum_{i=0}^\infty\frac{x^{i+1}}{(i+p)!}\right) = \frac{d}{dx}\left(x^{1-p}\sum_{i=0}^\infty\frac{x^{i+p}}{(i+p)!}\right) \\
&= \frac{d}{dx}\left(x^{1-p}\left(\exp(x)-\sum_{i=0}^{p-1}\frac{x^i}{i!}\right)\right) = x^{-p}\exp(x)(1-p+x)-\sum_{i=0}^{p-1}\frac{i-p+1}{i!}x^{i-p}\,,
\end{aligned}$$

we have

$$f_{|u_{pq}|}(t) = (-1)^p2^p(p-1)!\,t^{1-2p}\left(\exp\left(-\frac{t^2}{2}\right)\left(1-p-\frac{t^2}{2}\right)-\sum_{i=0}^{p-1}(-1)^i\frac{1-p+i}{2^i i!}t^{2i}\right).$$

□

4 Density Functions of l_{pq}

Similar to the derivation of the density functions of u_{pq}, we first establish a proposition which is analogous to Proposition 2. Both the propositions will also be used in §6 where we discuss large growth factors.

Let $\mathbf{X} = \mathbf{LU}$ and $\mathbf{X}^T = \hat{\mathbf{L}}\hat{\mathbf{U}}$ be the LU factorizations of $\mathbf{X}$ and $\mathbf{X}^T$ respectively. By Proposition 2, there exist a diagonal matrix $\hat{\mathbf{\Phi}} = diag(\hat{\phi}_1, \ldots, \hat{\phi}_n)$ and an upper triangular matrix $\hat{\mathbf{H}}$ such that $\hat{\mathbf{U}} = \hat{\mathbf{\Phi}}\hat{\mathbf{H}}$, where $\hat{\mathbf{\Phi}}$ and $\hat{\mathbf{H}}$ have the properties listed in the proposition. Set

$$\hat{\mathbf{D}} = diag(\hat{\eta}_{11}, \hat{\eta}_{22}, \ldots, \hat{\eta}_{nn})$$

where $\hat{\eta}_{11}, \ldots, \hat{\eta}_{nn}$ are the diagonal elements of $\hat{\mathbf{H}}$. Then

$$\mathbf{X}^T = \hat{\mathbf{L}}\hat{\mathbf{U}} = \hat{\mathbf{L}}\hat{\mathbf{\Phi}}\hat{\mathbf{D}}\hat{\mathbf{D}}^{-1}\hat{\mathbf{H}}$$

and therefore

$$\mathbf{X} = \hat{\mathbf{H}}^T\hat{\mathbf{D}}^{-1}\hat{\mathbf{D}}\hat{\mathbf{\Phi}}\hat{\mathbf{L}}^T .$$

Note that $\hat{\mathbf{H}}^T\hat{\mathbf{D}}^{-1}$ is unit lower triangular and $\hat{\mathbf{D}}\hat{\mathbf{\Phi}}\hat{\mathbf{L}}^T$ upper triangular. By the uniqueness of the LU factorization of $\mathbf{X}$, we have

$$\mathbf{L} = \hat{\mathbf{H}}^T\hat{\mathbf{D}}^{-1} . \tag{9}$$

Therefore we arrive at the following proposition.

Proposition 4 *Suppose $\mathbf{X}$ is an $n \times n$ complex Gaussian matrix and let $\mathbf{X} = \mathbf{LU}$ be the LU factorization of $\mathbf{X}$. Then there exist an $n \times n$ diagonal matrix $\mathbf{\Psi} = diag(\psi_1, \ldots, \psi_n)$ and an $n \times n$ lower triangular matrix $\mathbf{\Xi} = (\xi_{pq})$ with the following properties such that $\mathbf{L} = \mathbf{\Xi}\,\mathbf{\Psi}^{-1}$:*

(i) *$\psi_1, \psi_2, \ldots, \psi_n$ are mutually independent and all have the $\tilde{N}(0,2)$ distribution.*
(ii) *All ξ_{pq} with $1 \le q \le p \le n$ have the $\tilde{N}(0,2)$ distribution.*
(iii) *ξ_{pq} is independent of $\mathbf{\Xi}_{(1:p-1,1:p-1)}$ where $1 \le q \le p \le n$.*
(iv) *$\xi_{1q_1}, \xi_{2q_2}, \ldots, \xi_{nq_n}$ are mutually independent for any $q_1, q_2, \ldots, q_n$ with $1 \le q_j \le j$ for $j = 1, 2, \ldots, n$.*

Proof. Set $\mathbf{\Psi} = \hat{\mathbf{D}}$ and $\mathbf{\Xi} = \hat{\mathbf{H}}^T$ in (9). The desired properties follow from Proposition 2. □

We remark that $\psi_p = \xi_{pp}$ for $p = 1, \ldots, n$ since $\mathbf{L}$ is unit along its diagonal.

Theorem 5 *Suppose* $\mathbf{X}$ *is an* $n \times n$ *complex Gaussian matrix and let* $\mathbf{X} = \mathbf{LU}$ *be the LU factorization of* $\mathbf{X}$. *Then the density functions of* l_{pq}^R, l_{pq}^I *and* $|l_{pq}|$, *the real part, imaginary part and absolute value of the* (p, q)*-th* $(1 \le q < p \le n)$ *element* l_{pq} *of* $\mathbf{L}$ *respectively, are*

$$f_{l_{pq}^R}(t) = f_{l_{pq}^I}(t) = \frac{1}{2}\left(1 + t^2\right)^{-3/2},$$

where $-\infty < t < \infty$ *and*

$$f_{|l_{pq}|}(t) = \frac{2t}{(1 + t^2)^2},$$

where $0 < t$.

Proof. By Proposition 4, we have

$$l_{pq} = \frac{\xi_{pq}}{\psi_q} = \frac{\xi_{pq}}{\xi_{qq}}$$

where ξ_{pq} and ξ_{qq} have the $\tilde{N}(0, 2)$ distribution and are independent. Thus,

$$l_{pq} = \frac{\xi_{pq}^R + i\,\xi_{pq}^I}{\xi_{qq}^R + i\,\xi_{qq}^I} = \frac{\xi_{pq}^R \xi_{qq}^R + \xi_{pq}^I \xi_{qq}^I + i\left(\xi_{pq}^I \xi_{qq}^R - \xi_{pq}^R \xi_{qq}^I\right)}{{\xi_{qq}^R}^2 + {\xi_{qq}^I}^2}$$

and therefore,

$$\begin{aligned} l_{pq}^R &= \frac{\xi_{pq}^R \xi_{qq}^R + \xi_{pq}^I \xi_{qq}^I}{{\xi_{qq}^R}^2 + {\xi_{qq}^I}^2} = \frac{\xi_{pq}^R \xi_{qq}^R + \xi_{pq}^I \xi_{qq}^I}{\sqrt{{\xi_{qq}^R}^2 + {\xi_{qq}^I}^2}} \Big/ |\xi_{qq}| \equiv \zeta / |\xi_{qq}| \\ &= \frac{1}{\sqrt{2}} \frac{\zeta}{|\xi_{qq}|/\sqrt{2}} \equiv \frac{1}{\sqrt{2}}\phi. \end{aligned}$$

$|\xi_{qq}|^2$ and ζ have the χ_2^2 and $N(0, 1)$ distributions respectively and are independent. Hence ϕ has the student's t distribution with 2 degrees of freedom. Thus,

$$f_{l_{pq}^R}(t) = \frac{1}{2}\left(1 + t^2\right)^{-3/2},$$

where $-\infty < t < \infty$. Moreover,

$$|l_{pq}| = |\xi_{pq}|/|\xi_{qq}| = \sqrt{\frac{|\xi_{pq}|^2/2}{|\xi_{qq}|^2/2}} \equiv \sqrt{\delta},$$

where δ has the $F_{2,2}$ distribution. Hence

$$f_{|l_{pq}|}(t) = \begin{cases} \dfrac{2t}{(1 + t^2)^2} & t > 0 \\ 0 & \text{otherwise}. \end{cases} \qquad \square$$

5 Probability of small pivot

If one of the pivot elements u_{pp} is zero, Gaussian elimination will breakdown. Even though no pivot is zero, small pivots can cause some elements of the *LU* factors large, and then lead to a possible loss of accuracy in finite precision computation.

In this section, we describe the probability of the occurrence of small pivots. To make the statements below neatly, we use a shorthand notation. For given $\epsilon > 0$ and $1 \leq p \leq n$, we define

$$E_{p,\epsilon} = \{\mathbf{X} \in \mathbb{C}^{n\times n} | \, |u_{pp}| < \epsilon\} \, .$$

Then the event that at least one u_{pp} has $|u_{pp}| < \epsilon$ is naturally denoted by $\bigcup\limits_{p=1}^{n} E_{p,\epsilon}$.

Lemma 6 *Suppose* $\mathbf{X}$ *is an* $n \times n$ *complex Gaussian matrix and let* $\mathbf{X} = \mathbf{LU}$ *be the LU factorization of* $\mathbf{X}$*. Let* $\epsilon > 0$ *and* $1 \leq p \leq n$ *be given. Then*

$$Prob(E_{p,\epsilon}) \leq \frac{1}{2p}\,\epsilon^2 \, .$$

Proof. From (8), we have

$$f_{|u_{pq}|}(t) = t(p-1)\int_0^1 (1-s)^{p-2} s \exp\left(-\frac{t^2}{2}s\right) ds \leq t(p-1)\int_0^1 (1-s)^{p-2} s\, ds = \frac{t}{p}$$

where $p \geq 2$, and from which the desired result follows.

For the case where $p = 1$, it is sufficient to note that (i) $u_{11} = x_{11}$ and (ii) $|u_{11}|^2 = |x_{11}|^2$ is χ_2^2-distributed. Therefore,

$$Prob(E_{1,\epsilon}) = Prob(|x_{11}| < \epsilon) = \int_0^\epsilon t \exp\left(-\frac{t^2}{2}\right) dt \leq \frac{\epsilon^2}{2} \, .$$

□

It is worth indicating that $\frac{1}{2p}$ is the least upper bound for $\frac{Prob(E_{p,\epsilon})}{\epsilon^2}$ since $\lim_{\epsilon\to 0} \frac{Prob(E_{p,\epsilon})}{\epsilon^2} = \frac{1}{2p}$.

Theorem 7 *Suppose X is an $n \times n$ complex Gaussian matrix and let $X = LU$ be the LU factorization of X. Then*

$$Prob(\bigcup_{p=1}^{n} E_{p,\epsilon}) \leq \frac{\epsilon^2}{2} \sum_{p=1}^{n} \frac{1}{p}\,. \tag{10}$$

Proof. Since $Prob(\bigcup_{p=1}^{n} E_{p,\epsilon}) \leq \sum_{p=1}^{n} Prob(E_{p,\epsilon})$, (10) follows from Lemma 6. □

The coefficient of ϵ^2 is a rather slow-growing function of n. In fact, it is $O(\ln n)$. So, if ϵ is small enough, (10) will certainly give a satisfying bound for the desired probability. Moreover, the right-hand side of (10) is quadratic in ϵ.

6 Probability of large growth factor

When Gaussian elimination is performed for the solution of a linear system $\mathbf{Ax} = \mathbf{b}$, the computed *LU* factors $\tilde{\mathbf{L}}$ and $\tilde{\mathbf{U}}$ of $\mathbf{A}$ are produced. Then, by solving two corresponding triangular systems, we obtain the solution $\tilde{\mathbf{x}}$ to the linear system. The computed solution $\tilde{\mathbf{x}}$ satisfies a perturbed system

$$(\mathbf{A} + \mathbf{E})\,\tilde{\mathbf{x}} = \mathbf{b}$$

with

$$|\mathbf{E}| = |\mathbf{L}||\mathbf{U}|O(\epsilon_{machine}) \tag{11}$$

where $\epsilon_{machine}$ is the *machine epsilon* (for the definition of $\epsilon_{machine}$, see, for instance, [18, p.98]) and where, for any matrix $\mathbf{M}$, we use $|\mathbf{M}|$ to denote the matrix obtained by taking the absolute value of each element of $\mathbf{M}$. From this, it follows that

$$\|\mathbf{E}\|_\infty = \|\mathbf{L}\|_\infty \|\mathbf{U}\|_\infty O(\epsilon_{machine}) \tag{12}$$

(see, for instance, [18, p.164]). We therefore define the growth factors ρ_L and ρ_U to be

$$\rho_L = \|\mathbf{L}\|_\infty\,, \;\; \rho_U = \|\mathbf{U}\|_\infty / \|\mathbf{A}\|_\infty\,.$$

We remark that this definition of growth factors differs from the traditional one

$$\rho = \frac{\max_{i,j} |u_{ij}|}{\max_{i,j} |a_{ij}|} \qquad \text{or} \qquad \rho = \frac{\max_{i,j,k} |a_{ij}^{(k)}|}{\max_{i,j} |a_{ij}|} \tag{13}$$

(see, for instance, [2,9,11,18,26]). It may be easier to analyze ρ_L and ρ_U than ρ probabilistically. The following theorem provides probabilistic bounds on the sizes of ρ_L and ρ_U. The proof basically follows the proofs of Theorems 4.3 and 4.4 in [15] except that some independence properties of the LU factors are exploited and the concrete distributions of elements are used.

Theorem 8 *Suppose* $\mathbf{X}$ *is an* $n \times n$ *complex Gaussian matrix and let* $\mathbf{X} = \mathbf{LU}$ *be the LU factorization of* $\mathbf{X}$. *Then there exist positive constants* c *and* d, *independent of* α, r *or* n, *such that*

$$Prob(\rho_L > r) \leq \frac{1}{2} n^2 e^{-\alpha} + c \frac{\alpha\, n^2}{r^2} (2 \ln r - \ln \alpha)$$

for $r \geq 2$ *and* $0 < \alpha < (r-1)^2/e^2$, *and*

$$Prob(\rho_U > r) \leq d \frac{n^2}{r^2}$$

for $r \geq 2$.

Proof. The theorem is obvious when $n = 1$. So we assume $n \geq 2$ in the following. By Proposition 4, we have

$$\mathbf{L}_{(p,1:p-1)} = [\xi_{p1}/\psi_1, \xi_{p2}/\psi_2, \ldots, \xi_{pp-1}/\psi_{p-1}] \ .$$

Hence

$$\|\mathbf{L}_{(p,1:p-1)}\|_\infty = \frac{|\xi_{p1}|}{|\psi_1|} + \frac{|\xi_{p2}|}{|\psi_2|} + \cdots + \frac{|\xi_{pp-1}|}{|\psi_{p-1}|} \ .$$

Therefore

$$\begin{aligned} &Prob\left(\|\mathbf{L}\|_\infty > r\right) = Prob\left(\max_{2 \leq p \leq n} \|\mathbf{L}_{(p,1:p-1)}\|_\infty > r-1\right) \\ &\leq Prob\left(\theta_n \sum_{i=1}^{n-1} \frac{1}{|\psi_i|} > r-1\right) \\ &= Prob\left(\theta_n > \sqrt{2\alpha}, \theta_n \sum_{i=1}^{n-1} \frac{1}{|\psi_i|} > r-1\right) + \\ &\quad Prob\left(\theta_n \leq \sqrt{2\alpha}, \theta_n \sum_{i=1}^{n-1} \frac{1}{|\psi_i|} > r-1\right) \\ &\leq Prob\left(\theta_n > \sqrt{2\alpha}\right) + Prob\left(\sqrt{2\alpha} \sum_{i=1}^{n-1} \frac{1}{|\psi_i|} > r-1\right), \end{aligned} \tag{14}$$

where $\theta_n = \max_{1\le q<p\le n} |\xi_{pq}|$. Since ξ_{pq} has the $\tilde{N}(0,2)$ distribution, $|\xi_{pq}|^2$ is distributed χ_2^2. Hence

$$Prob\left(\theta_n > \sqrt{2\alpha}\right) \le \sum_{1\le q<p\le n} Prob\left(|\xi_{pq}| > \sqrt{2\alpha}\right) = \frac{1}{2}n(n-1)\exp\left(-\alpha\right). \quad (15)$$

To bound the second probability of (14), we would like to use Chebyshev's inequality [1, p.463]. However, since the variance $\sigma^2\left(\frac{1}{|\psi_i|}\right) = \infty$, we have to apply the inequality indirectly as follows.

By Jensen's inequality [10, Theorem 19, p.28],

$$\begin{aligned} Prob\left(\sqrt{2\alpha}\sum_{i=1}^{n-1}\frac{1}{|\psi_i|} > r-1\right) &= Prob\left(\sum_{i=1}^{n-1}\frac{1}{|\psi_i|} > \frac{r-1}{\sqrt{2\alpha}}\right) \\ &\le Prob\left(\sum_{i=1}^{n-1}\frac{1}{|\psi_i|^\beta} > \left(\frac{r-1}{\sqrt{2\alpha}}\right)^\beta\right) \end{aligned}$$

for $\forall\beta$ with $1/2 < \beta < 1$. Since $\psi_1, \psi_2, \ldots, \psi_{n-1}$ are independent and identically distributed $\tilde{N}(0,2)$ variables, we have $E\left(\left(\sum_{i=1}^{n-1}\frac{1}{|\psi_i|^\beta}\right)^2\right) = \frac{n-1}{2^\beta}\left[\Gamma(1-\beta)+(n-2)(\Gamma(1-\frac{\beta}{2}))^2\right]$. Therefore

$$Prob\left(\sum_{i=1}^{n-1}\frac{1}{|\psi_i|^\beta} > \left(\frac{r-1}{\sqrt{2\alpha}}\right)^\beta\right) \le \frac{\alpha^\beta(n-1)}{(r-1)^{2\beta}}\left[\Gamma(1-\beta)+(n-2)\left(\Gamma(1-\frac{\beta}{2})\right)^2\right]$$

by Chebyshev's inequality. For the right-hand side, we can find some constant $\hat{c} > 0$ such that

$$\begin{aligned} \Gamma(1-\beta)+(n-2)\left(\Gamma(1-\frac{\beta}{2})\right)^2 &= \frac{1}{1-\beta}\frac{\pi(1-\beta)}{\sin\pi(1-\beta)}\frac{1}{\Gamma(\beta)}+(n-2)\left(\Gamma(1-\frac{\beta}{2})\right)^2 \\ &\le \hat{c}\left(\frac{1}{1-\beta}+(n-2)\right) \le \frac{\hat{c}}{2}\frac{n}{1-\beta}, \end{aligned}$$

where we used $\Gamma(1-\beta)\Gamma(\beta) = \frac{\pi}{\sin\pi\beta}$[8, 8.334, p.946] in the first equation. Thus

$$Prob\left(\sqrt{2\alpha}\sum_{i=1}^{n-1}\frac{1}{|\psi_i|} > r-1\right) \le \frac{1}{2}\hat{c}\,n^2\frac{\alpha^\beta}{(1-\beta)(r-1)^{2\beta}}$$

for $\forall\beta \in (1/2, 1)$. As a function of β, the right-hand side of the above inequality attains its minimum value at the point $\beta^* = 1 - \frac{1}{2\ln(r-1)-\ln\alpha} \in (1/2, 1)$. By taking this value as the bound, we then have

$$Prob\left(\sqrt{2\alpha}\sum_{i=1}^{n-1}\frac{1}{|\psi_i|} > r-1\right) \le \frac{1}{2}e\,\hat{c}\,n^2\frac{\alpha}{(r-1)^2}\left[2\ln(r-1)-\ln\alpha\right]. \quad (16)$$

The desired inequality of ρ_L now follows by combining (14), (15) and (16).

To prove the inequality about ρ_U, we use Proposition 2. By the proposition,

$$\mathbf{U}_{(p,p:n)} = \phi_p \mathbf{H}_{(p,p:n)} = \phi_p \left[\eta_{pp}, \eta_{pp+1}, \ldots, \eta_{pn}\right]$$

where $\phi_p, \eta_{pp}, \eta_{pp+1}, \ldots, \eta_{pn}$ are mutually independent, and $(\phi_p^2-1)/(p-1)$ has the $F_{2(p-1),2}$ distribution while the η's identically have the $\tilde{N}(0,2)$ distribution.

Let $\delta_p = (\phi_p^2 - 1)/(p-1)$. Then

$$\begin{aligned} \|\mathbf{U}_{(p,p:n)}\|_\infty &= |\phi_p| \sum_{q=p}^{n} |\eta_{pq}| = \sqrt{1+(p-1)\delta_p} \sum_{q=p}^{n} |\eta_{pq}| \\ &\le \sqrt{n-p+1}\sqrt{1+(p-1)\delta_p}\sqrt{\sum_{q=p}^{n} |\eta_{pq}|^2} \equiv \sqrt{n-p+1}\sqrt{1+(p-1)\delta_p}\sqrt{\zeta_p}\,. \end{aligned}$$

The variable ζ_p has the $\chi^2_{2(n-p+1)}$ distribution and is independent of δ_p. Hence

$$\begin{aligned} & Prob\left(\sqrt{1+(p-1)\delta_p}\sqrt{\zeta_p} > \hat{r}\right) \\ &= \iint_{\sqrt{1+(p-1)\lambda}\sqrt{\theta}>\hat{r}} (p-1)^p \lambda^{p-2}(1+(p-1)\lambda)^{-p} \frac{1}{2^{n-p+1}(n-p)!} \theta^{n-p} e^{-\theta/2} d\lambda d\theta \\ &= \frac{(p-1)^p}{2^{n-p+1}(n-p)!} \int_0^\infty d\lambda \int_{\hat{r}^2/(1+(p-1)\lambda)}^\infty \lambda^{p-2}(1+(p-1)\lambda)^{-p}\theta^{n-p}e^{-\theta/2}d\theta \end{aligned}$$

for $\forall \hat{r} > 0$. Let $s = (p-1)\lambda$. Then

$$\begin{aligned} & Prob\left(\sqrt{1+(p-1)\delta_p}\sqrt{\zeta_p} > \hat{r}\right) \\ &= \frac{p-1}{2^{n-p+1}(n-p)!} \int_0^\infty ds \int_{\hat{r}^2/(1+s)}^\infty s^{p-2}(1+s)^{-p}\theta^{n-p}e^{-\theta/2}d\theta \\ &\le \frac{p-1}{2^{n-p+1}(n-p)!} \int_0^\infty ds \int_{\hat{r}^2/(1+s)}^\infty (1+s)^{-2}\theta^{n-p}e^{-\theta/2}d\theta\,. \end{aligned}$$

Let $t = \hat{r}^2/(1+s)$. We further have

$$\begin{aligned} Prob\left(\sqrt{1+(p-1)\delta_p}\sqrt{\zeta_p} > \hat{r}\right) &\le \frac{p-1}{2^{n-p+1}(n-p)!\,\hat{r}^2} \int_0^{\hat{r}^2} dt \int_t^\infty \theta^{n-p}e^{-\theta/2}d\theta \\ &\le \frac{p-1}{2^{n-p+1}(n-p)!\,\hat{r}^2} \int_0^{\infty} dt \int_t^\infty \theta^{n-p}e^{-\theta/2}d\theta\,. \end{aligned}$$

By integration by parts,

$$\int_0^\infty dt \int_t^\infty \theta^{n-p} e^{-\theta/2} d\theta = \int_0^\infty t^{n-p+1} e^{-t/2} dt = 2^{n-p+2}(n-p+1)!\,.$$

Hence

$$Prob\left(\sqrt{1+(p-1)\delta_p}\sqrt{\zeta_p} > \hat{r}\right) \leq \frac{2(n-p+1)(p-1)}{\hat{r}^2}$$

and therefore

$$\begin{aligned} Prob\left(\|\mathbf{U}_{(p,p:n)}\|_\infty > \hat{r}\right) &\leq Prob\left(\sqrt{n-p+1}\sqrt{1+(p-1)\delta_p}\sqrt{\zeta_p} > \hat{r}\right) \\ &\leq \frac{2(n-p+1)^2(p-1)}{\hat{r}^2} \end{aligned} \tag{17}$$

for $\forall \hat{r} > 0$. From (17), we can now obtain a bound for $Prob\,(\rho_U > r)$ as follows.

$$\begin{aligned} &Prob\,(\rho_U > r) = Prob\left(\max_{2\leq p\leq n} \|\mathbf{U}_{(p,p:n)}\|_\infty > r\|\mathbf{X}\|_\infty\right) \\ &\leq Prob\left(\max_{2\leq p\leq n-1} \|\mathbf{U}_{(p,p:n)}\|_\infty > r\|\mathbf{X}\|_\infty\right) + Prob\left(\|\mathbf{U}_{(n,:)}\|_\infty > r\|\mathbf{X}\|_\infty\right) \\ &\leq Prob\left(\max_{2\leq p\leq n-1} \|\mathbf{U}_{(p,p:n)}\|_\infty > r\|\mathbf{X}_{(n,:)}\|_\infty\right) + Prob\left(\|\mathbf{U}_{(n,:)}\|_\infty > r\|\mathbf{X}\|_\infty\right) \\ &\leq \sum_{p=2}^{n-1} Prob\left(\|\mathbf{U}_{(p,p:n)}\|_\infty > r\|\mathbf{X}_{(n,:)}\|_\infty\right) + Prob\,(|u_{nn}| > r\|\mathbf{X}\|_\infty)\,. \end{aligned} \tag{18}$$

Note that $\|\mathbf{U}_{(p,p:n)}\|_\infty$ and $\|\mathbf{X}_{(n,:)}\|_\infty$ are independent variables if $p < n$. A similar argument to Lemma C.4 of [15] together with (17) implies that

$$Prob\left(\|\mathbf{U}_{(p,p:n)}\|_\infty > r\|\mathbf{X}_{(n,:)}\|_\infty\right) \leq \frac{2(n-p+1)^2(p-1)}{r^2} E\left(\frac{1}{\|\mathbf{X}_{(n,:)}\|_\infty^2}\right)$$

for $2 \leq p \leq n-1$. Therefore

$$\begin{aligned} &\sum_{p=2}^{n-1} Prob\left(\|\mathbf{U}_{(p,p:n)}\|_\infty > r\|\mathbf{X}_{(n,:)}\|_\infty\right) \leq \frac{2}{r^2} E\left(\frac{1}{\|\mathbf{X}_{(n,:)}\|_\infty^2}\right) \sum_{p=2}^{n-1} (n-p+1)^2(p-1) \\ &\leq \frac{2}{r^2}\frac{\hat{d}}{n^2} \sum_{p=2}^{n-1} (n-p+1)^2(p-1) = \frac{\hat{d}}{6n^2\, r^2}(n-2)(n-1)(n^2+3n+6) \end{aligned} \tag{19}$$

for some constant $\hat{d} > 0$ by Lemma 12. For the second probability of (18), we have

$$u_{nn} = x_{nn} - \mathbf{z}^T \mathbf{R}^{-1} \mathbf{Q}^{-1} \mathbf{X}_{(1:n-1,n)}$$

by (4). Hence

$$
\begin{aligned}
|u_{nn}| &\leq |x_{nn}| + \frac{z_{n-1}}{r_{n-1n-1}} \|\mathbf{X}_{(1:n-1,n)}\|_2 \leq |x_{nn}| + \frac{z_{n-1}\sqrt{n-1}}{r_{n-1n-1}} \|\mathbf{X}_{(1:n-1,n)}\|_\infty \\
&\leq \|\mathbf{X}\|_\infty + \frac{z_{n-1}\sqrt{n-1}}{r_{n-1n-1}} \|\mathbf{X}\|_\infty = \|\mathbf{X}\|_\infty \left(1 + (n-1)\sqrt{\delta_n}\right)
\end{aligned}
$$

where $\delta_n = \frac{1}{n-1} \frac{z_{n-1}^2}{r_{n-1n-1}^2}$ by (5) and has the $F_{2(n-1),2}$ distribution. Thus

$$
Prob(|u_{nn}| > r\|\mathbf{X}\|_\infty) \leq Prob\left(1 + (n-1)\sqrt{\delta_n} > r\right) \leq \frac{(n-1)^2}{(r-1)^2}. \qquad (20)
$$

The desired result now follows from (18), (19) and (20). □

Corollary 9 *Let* $\mathbf{X} = \mathbf{LU}$ *be the LU factorization of an* $n \times n$ *complex Gaussian matrix* $\mathbf{X}$. *Then*

(i) *there exists a constant* $c > 0$, *independent of* r *or* n, *such that*

$$
Prob\,(\rho_L > r) \leq c\, n^2 \frac{(\ln r)^2}{r^2}
$$

for $r \geq 7$.

(ii) $\lim_{n\to\infty} Prob\,(\rho_L > \alpha_n n \ln n) = \lim_{n\to\infty} Prob\,(\rho_U > \alpha_n n) = 0$ *for* $\forall\{\alpha_n\}$ *with* $\lim_{n\to\infty} \alpha_n = \infty$.

Proof. For the proof of part (i), let $\alpha = 2\ln r$ in Theorem 8. It is easy to verify that the condition $0 < \alpha < (r-1)^2/e^2$ is satisfied when $r \geq 7$. Part (ii) follows from part (i) and Theorem 8 respectively. □

Corollary 10 *Let* $\mathbf{X} = \mathbf{LU}$ *be the LU factorization of an* $n \times n$ *complex Gaussian matrix* $\mathbf{X}$. *Then there exist constants* $c > 0$ *and* $d > 0$ *such that*

$$
E(\rho_L) \leq c\, n \ln n \qquad \text{and} \qquad E(\rho_U) \leq d\, n
$$

for $\forall n \geq 2$.

Proof. Let $f_{\rho_L}(t)$ be the density function of ρ_L. Corollary 9 asserts that $\int_r^\infty f_{\rho_L}(t)dt \leq \hat{c}\, n^2 \frac{(\ln r)^2}{r^2}$ for some constant $\hat{c}$ and $\forall r \geq 7$. By integration by

Table 1
Probabilities of small pivot with sample sizes of 10^6 for each pair (n, ϵ).

n	ϵ	*Frequency*	*Empirical probability*	*Theoretical bound*
25	10^{-1}	19056	1.9056×10^{-2}	1.9080×10^{-2}
25	10^{-2}	177	1.77×10^{-4}	1.9080×10^{-4}
50	10^{-1}	22375	2.2375×10^{-2}	2.2496×10^{-2}
50	10^{-2}	203	2.03×10^{-4}	2.2496×10^{-4}
75	10^{-1}	24108	2.4108×10^{-2}	2.4507×10^{-2}
75	10^{-2}	235	2.35×10^{-4}	2.4507×10^{-4}

parts, we then have

$$E(\rho_L) = \int_0^\infty t f_{\rho_L}(t)dt = \int_0^\infty \{\int_t^\infty f_{\rho_L}(s)ds\}dt = (\int_0^r + \int_r^\infty)\{\int_t^\infty f_{\rho_L}(s)ds\}dt$$
$$\leq r + \hat{c}\, n^2 \int_r^\infty \frac{(\ln t)^2}{t^2} dt = r + \hat{c}\, n^2 \,\frac{1}{r} \left[(\ln r)^2 + 2\ln r + 2\right].$$

The desired result of ρ_L follows by setting $r = n \ln n$. The result of ρ_U can be proved similarly. □

7 Numerical Experiments

We present numerical results to support the theorems obtained in the previous sections. All the experiments were performed in MATLAB 6.1.0.450 Release 12.1.

Experiment 1. In our first experiment, 10^6 complex Gaussian matrices of dimension $n = 31$ were selected independently. Then LU factorization without pivoting was applied to each of the matrices and statistics on $l^R_{13,12}$, $|l_{13,12}|$, $u^R_{12,12}$, $|u_{12,12}|$, $u^R_{31,31}$ and $|u_{31,31}|$ were accumulated. The data are plotted in Figures 1 and 2, together with the corresponding density functions indicated by Theorems 3 and 5.

Experiment 2. The purpose of our second experiment was to test formula (10). Complex Gaussian matrices of several dimensions n were selected independently, with sample sizes all being 10^6. Two values of ϵ were used and we obtained the results in Table 1. The frequency column of the table provides the numbers of matrices which, in their LU factors, have at least one u_{pp} less than ϵ in magnitude. By comparing with the theoretical probabilities, we conclude that the bound given by (10) is a fairly tight one.

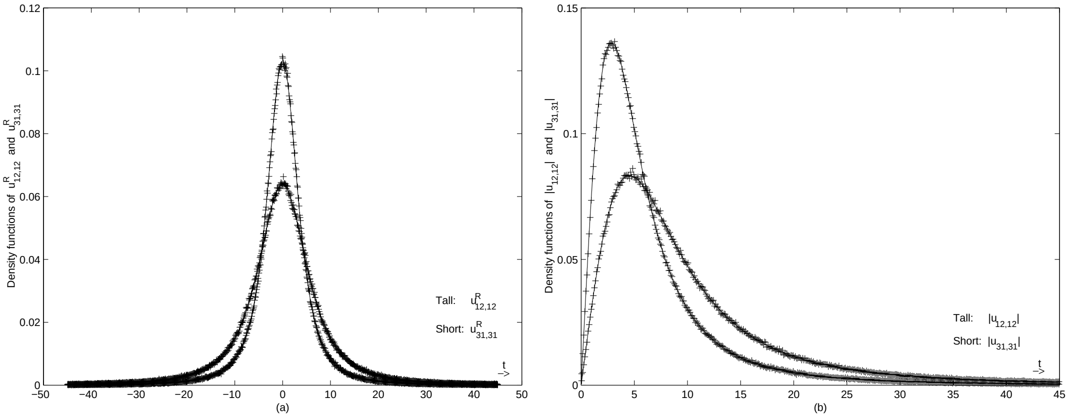

Fig. 1. (a) Distributions of $u^R_{12,12}$ (tall) and $u^R_{31,31}$ (short): observed (plus), predicted (solid). (b) Distributions of $|u_{12,12}|$ (tall) and $|u_{31,31}|$ (short): observed (plus), predicted (solid).

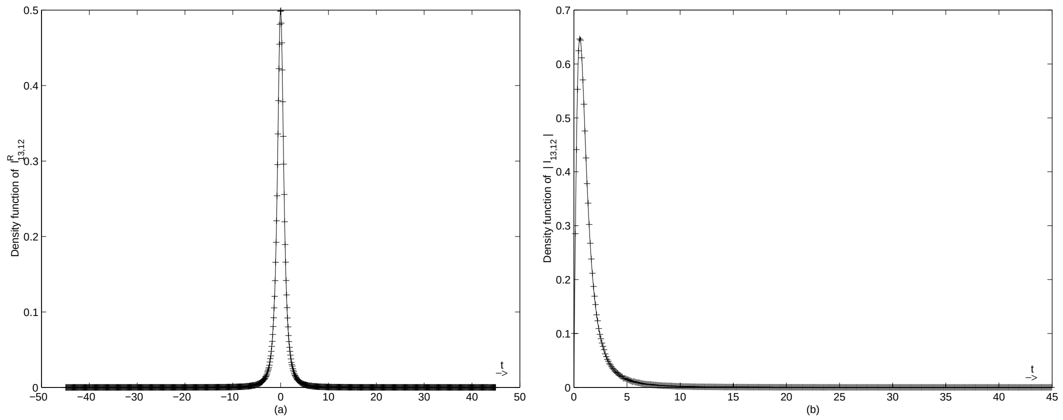

Fig. 2. (a) Distribution of $l^R_{13,12}$: observed (plus), predicted (solid). (b) Distribution of $|l_{13,12}|$: observed (plus), predicted (solid).

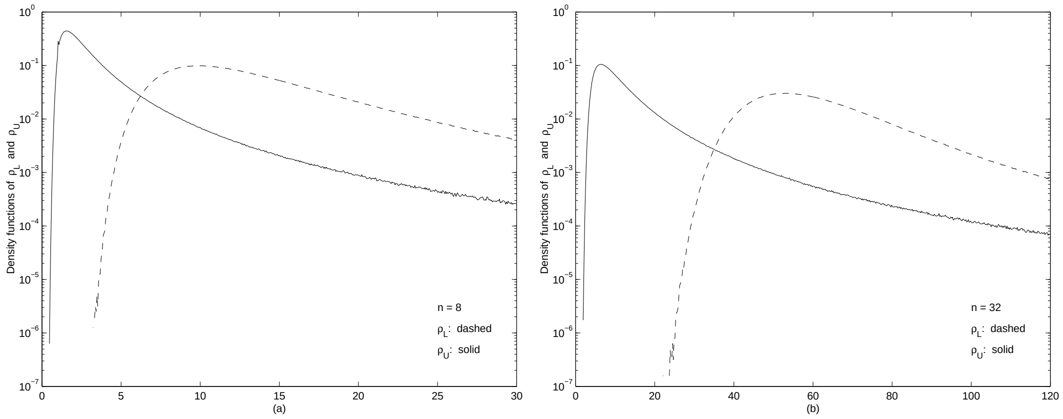

Fig. 3. Density functions of ρ_L (dashed) and ρ_U (solid). (a) $n = 8$. (b) $n = 32$.

Table 2
Computed expected values $\tilde{E}(\rho_L)$ and $\tilde{E}(\rho_U)$ with sample sizes of 10^6 for each n.

n	2	4	8	16	32	64	128
$\tilde{E}(\rho_L)$	2.5723	6.3435	14.4969	31.0962	63.7324	126.8528	248.3116
$\tilde{E}(\rho_L)/n$	1.2862	1.5859	1.8121	1.9435	1.9916	1.9821	1.9399
$\tilde{E}(\rho_U)$	1.1416	1.7643	3.3463	6.7372	13.8291	28.3944	58.4784
$\tilde{E}(\rho_U)/n$	0.5708	0.4411	0.4183	0.4211	0.4322	0.4437	0.4569

Experiment 3. We computed the expected values $E(\rho_L)$ and $E(\rho_U)$ with sample sizes of 10^6 for each n. The data were recorded in Table 2. It appears that $E(\rho_L) = c_L\, n$ and $E(\rho_U) = c_U\, n$ where c_L first increases and then decreases while c_U decreases first and then increases. With far smaller sample sizes for $n = 256, 512$ and 1024, further experiments showed that c_L continues to decrease. However, it seemed that c_U starts to decrease after $n = 512$. Theoretically, $E(\rho_L) \geq E(\mathbf{L}_{(n,:)}) = \sum_{q=1}^{n-1} E(|l_{nq}|) + 1 = \frac{\pi}{2}(n-1) + 1$ by Theorem 5. However, we failed to obtain any lower bound for $E(\rho_U)$.

We plotted in Fig. 3 the experimental density functions of ρ_L and ρ_U for $n = 8$, 32, based on sample sizes of about 3×10^7 matrices for each n. The functions appear to decrease algebraically with ρ_L and ρ_U and it seems that this characteristic remains while the means move to the right as n is increased. By contrast, with partial pivoting, the density of ρ defined in (13) appears to decrease exponentially with ρ [18, p.168].

8 Discussions

If a small pivot is encountered, then we can expect that large elements will appear in computed $\tilde{\mathbf{L}}$ and $\tilde{\mathbf{U}}$. As a result, the perturbation matrix $\mathbf{E}$ may contain large elements because of (11). Since there is nothing in GE without pivoting to manage the occurrence of small pivots, such a possibility exists. In the case of complex Gaussian matrices, the probability of the appearance of small pivots is about the square of the size of the small pivots (see Theorem 7). By comparison, the corresponding probability in the real Gaussian matrix case is just linear in the size of small pivots [29]. Thus, in general, we can expect that the $\mathbf{E}$ in the complex case has smaller elements in magnitude than the $\mathbf{E}$ does in the real case.

Instability of GE without pivoting can arise if one or both of the factors $\|\mathbf{L}\|_\infty$ and $\|\mathbf{U}\|_\infty$ is large relative to $\|\mathbf{A}\|_\infty$ since $\|\mathbf{E}\|_\infty/\|\mathbf{A}\|_\infty$ is then large by (12). On average, we have $\rho_L\rho_U \approx n^2$ for complex Gaussian matrices $\mathbf{X}$ (see Experiment 3). This implies that $\|\mathbf{E}\|_\infty/\|\mathbf{X}\|_\infty \approx n^2 O(\epsilon_{machine}) \gg O(\epsilon_{machine})$.

Therefore, we expect that GE without pivoting is backward instable when n is very large [18, p.164]. In [19], Trefethen and Schreiber developed a statistical model of the growth factor ρ defined by (13) for the pivoting cases. They observed that, on average, the growth factor ρ, normalized by the standard deviation of the initial matrix elements, is about $n^{2/3}$ for partial pivoting and about $n^{1/2}$ for complete pivoting based on various distributions of random matrices, including real Gaussian matrices. Even though their observations were made based on real matrices, we believe the (normalized) ρ is about the same or even smaller for complex matrices. Therefore, we have $\|\mathbf{E}\|_\infty/\|\mathbf{X}\|_\infty \approx n^{2/3}O(\epsilon_{machine})$ for partial pivoting and $\|\mathbf{E}\|_\infty/\|\mathbf{X}\|_\infty \approx n^{1/2}O(\epsilon_{machine})$ for complete pivoting. This is an improvement on without pivoting by a factor of more than n. When n is large, the difference will become huge between without pivoting and with pivoting in reliability.

The probability of large growth factors decreases exponentially with pivoting [18, p.168], but just algebraically if there is no pivoting (see Experiment 3). So, it is highly unlikely for the growth factors with pivoting to be far away from their expected values. By contrast, the growth factors without pivoting are likely to be very large quite regularly. This huge difference reflects the fact that, in practice, instability has never arisen for GE with pivoting applied to random matrices. However,without pivoting, one expects to lose a number of digits of accuracy quite regularly and a lot of digits on rare occasions.

Assuming that the elements $\breve{l}_{pq}$ $(p \geq q)$ of a lower or unit lower $n \times n$ random $\breve{\mathbf{L}}$ are independent and have the same symmetric, strictly stable distribution, Viswanath and Trefethen [22] proved that the 2-norm condition number $\kappa_n = \|\breve{\mathbf{L}}\|_2 \|\breve{\mathbf{L}}^{-1}\|_2$ grows exponentially with n. Such exponentially ill-conditioned $\breve{\mathbf{L}}$'s, however, are simply not present in the LU factors of random matrices, regardless of whether one pivots or not (see the Explanation section of Lecture 22, [18]). The LU factors of a complex Gaussian matrix indeed possess some degree of independence among their elements (see Propositions 2 and 4). For example, the $\mathbf{L}$-factor can be written as $\mathbf{L} = \mathbf{\Xi}\,\mathbf{\Psi}^{-1}$ where $\mathbf{\Psi}$ is a diagonal matrix with independent diagonal elements and $\mathbf{\Xi}$ a lower triangular matrix of which elements from different rows are mutually independent. However, elements from the same row of $\mathbf{\Xi}$ are correlated. These row element correlations help prevent the appearance of exponentially ill-conditioned $\mathbf{L}$-factors (note that $\|\mathbf{\Psi}\|_2 \|\mathbf{\Psi}^{-1}\|_2$ is almost impossible to grow exponentially because the diagonal elements of $\mathbf{\Psi}$ have the $\tilde{N}(0, 2)$ distribution).

9 Appendix

Lemma 11 *Suppose $x_1, \ldots, x_m, y_1, \ldots, y_n$ are independent $N(0, 1)$ random variables. Let $\mathbf{x} = [x_1, \ldots, x_m]^T$ and $\mathbf{y} = [y_1, \ldots, y_n]^T$. If*

(i) $v_i = f_i(x_1, \ldots, x_m)$, *where* $i = 1, \ldots, k$, *are real-value functions of* $\mathbf{x}$,
(ii) $\mathbf{Q} \in \mathbb{R}^{n\times n}$ *is an orthogonal matrix whose elements are functions of* $\mathbf{x}$,
(iii) $\mathbf{z} = \mathbf{Q}\,\mathbf{y}$,

then

$$\begin{aligned} & Prob(v_1 < \alpha_1, \ldots, v_k < \alpha_k, z_1 < \beta_1, \ldots, z_j < \beta_j) \\ = & Prob(v_1 < \alpha_1, \ldots, v_k < \alpha_k) \textstyle\prod_{i=1}^{j} Prob(z_i < \beta_i) \end{aligned}$$

for any $1 \le j \le n$ *and any real numbers* $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_j$.

Proof. Define the following space regions:

$$\begin{aligned} &\Omega = \{(\mathbf{x}^T, \mathbf{y}^T)^T \in \mathbb{R}^{m+n} | v_1 < \alpha_1, \ldots, v_k < \alpha_k, z_1 < \beta_1, \ldots, z_j < \beta_j\}\,, \\ &\Omega_1 = \{\mathbf{x} \in \mathbb{R}^m | v_1 < \alpha_1, \ldots, v_k < \alpha_k\}\,, \quad \Omega_{\mathbf{x}} = \{\mathbf{y} \in \mathbb{R}^n | z_1 < \beta_1, \ldots, z_j < \beta_j\}\,, \\ &\Omega_2 = \{\mathbf{z} \in \mathbb{R}^n | z_1 < \beta_1, \ldots, z_j < \beta_j\}\,, \qquad \hat{\Omega}_i = \{(\mathbf{x}^T, \mathbf{y}^T)^T \in \mathbb{R}^{m+n} | z_i < \beta_i\}\,. \end{aligned}$$

Then

$$\begin{aligned} & Prob(v_1 < \alpha_1, \ldots, v_k < \alpha_k, z_1 < \beta_1, \ldots, z_j < \beta_j) \\ = & \iint\limits_{\Omega} \frac{1}{(\sqrt{2\pi})^{m+n}} \exp\{-\frac{1}{2}\left(\|\mathbf{x}\|_2^2 + \|\mathbf{y}\|_2^2\right)\}\, d\mathbf{x} d\mathbf{y} \\ = & \frac{1}{(\sqrt{2\pi})^{m+n}} \int\limits_{\Omega_1} d\mathbf{x} \int\limits_{\Omega_{\mathbf{x}}} \exp\{-\frac{1}{2}\left(\|\mathbf{x}\|_2^2 + \|\mathbf{y}\|_2^2\right)\}\, d\mathbf{y} \\ = & \frac{1}{(\sqrt{2\pi})^{m+n}} \int\limits_{\Omega_1} d\mathbf{x} \int\limits_{\Omega_2} \exp\{-\frac{1}{2}\left(\|\mathbf{x}\|_2^2 + \|\mathbf{z}\|_2^2\right)\}\, d\mathbf{z} \\ = & \frac{1}{(\sqrt{2\pi})^{m+j}} \int\limits_{\Omega_1} \exp\{-\frac{1}{2}\|\mathbf{x}\|_2^2\}\, d\mathbf{x} \prod_{i=1}^{j} \int\limits_{-\infty}^{\beta_i} \exp\{-\frac{1}{2} z_i^2\}\, dz_i\,. \end{aligned}$$

Similarly, one can show that

$$\begin{aligned} Prob(z_i < \beta_i) &= \iint\limits_{\hat{\Omega}_i} \frac{1}{(\sqrt{2\pi})^{m+n}} \exp\{-\frac{1}{2}\left(\|\mathbf{x}\|_2^2 + \|\mathbf{y}\|_2^2\right)\}\, d\mathbf{x} d\mathbf{y} \\ &= \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\beta_i} \exp\{-\frac{1}{2} z_i^2\}\, dz_i\,. \end{aligned}$$

Therefore, the desired equation holds. $\Box$

Lemma 12 *Suppose* $x_1, x_2, \ldots, x_n$ *with* $n \ge 2$ *are independent* $\tilde{N}(0, 2)$ *random variables. Then there exists a constant* $d > 0$ *such that*

$$E\left(\frac{1}{(\sum_{i=1}^{n} |x_i|)^2}\right) \le d\frac{1}{n^2}\,.$$

Proof. Since x_i has the $\tilde{N}(0,2)$ distribution, the density function of $|x_i|$ is

$$f(\lambda) = \lambda \exp(-\lambda^2/2)$$

where $\lambda > 0$. Hence

$$E\left(\frac{1}{(\sum_{i=1}^n |x_i|)^2}\right) = \int_0^\infty \cdots \int_0^\infty \frac{\prod_{i=1}^n \lambda_i}{(\sum_{i=1}^n \lambda_i)^2} \exp(-\frac{1}{2}\sum_{i=1}^n \lambda_i^2)\, d\lambda_1 \cdots d\lambda_n\,.$$

By Arithmetic-geometric inequality[8, 11.116, p.1126],

$$E\left(\frac{1}{(\sum_{i=1}^n |x_i|)^2}\right) \leq \frac{1}{n^2}\int_0^\infty \cdots \int_0^\infty \prod_{i=1}^n \lambda_i^{1-\frac{2}{n}} \exp(-\frac{1}{2}\sum_{i=1}^n \lambda_i^2)\, d\lambda_1 \cdots d\lambda_n$$

$$= \frac{1}{n^2}\left(\int_0^\infty \lambda^{1-\frac{2}{n}} \exp(-\lambda^2/2)\, d\lambda\right)^n = \frac{1}{2n^2}\left(\int_0^\infty t^{-1/n}\exp(-t)dt\right)^n = \frac{1}{2n^2}\left[\Gamma\left(1-\frac{1}{n}\right)\right]^n.$$

From [8, 8.342, p.948], we can have $\lim_{n\to\infty}\left[\Gamma\left(1-\frac{1}{n}\right)\right]^n = \gamma$ where $\gamma = 1.781072\ldots$ is the Euler's constant[8]. Thus there is a constant $\hat{d} > 0$ such that $\left[\Gamma\left(1-\frac{1}{n}\right)\right]^n \leq \hat{d}$ for all $n \geq 2$. Therefore, $E\left(1/(\sum_{i=1}^n |x_i|)^2\right) \leq \frac{\hat{d}}{2n^2}$ for $\forall n \geq 2$. □

Acknowledgements

The author is grateful to the anonymous referee for giving him very informative and helpful comments on an earlier version of this paper. Also, Prof. Alan Edelman of MIT provided the authors of [29] with several important references and suggestions based on which the research there could be carried out. Those references and suggestions proved to be still invaluable to the work of this paper and the author wishes to thank him again. The author would also like to thank Prof. Shaun Wulff of UW for helpful discussions.

References

[1] P. Bickel and K. Doksum, *Mathematical Statistics*, Holden-Day, Inc., 1977.

[2] J. W. Demmel, *Applied Numerical Linear Algebra*, SIAM Publications, Philadelphia, PA, 1997.

[3] A. Edelman, *Eigenvalues and condition numbers of random matrices*, SIAM J. Matrix Anal. Appl., 9 (1988), 543-560.

[4] ——, *Eigenvalues and condition numbers of random matrices*, Ph.D. Thesis, Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139.

[5] ——, *The distribution and moments of the smallest eigenvalue of a random matrix of Wishart type*, Linear Alg. Appl., 159 (1991) 55-80.

[6] L. V. Foster, *The probability of large diagonal elements in the QR factorization*, SIAM J. Sci. Stat. Comput., 11 (1990), 531-544.

[7] ——, *Gaussian elimination with partial pivoting can fail in practice*, SIAM J. Matrix Anal. Appl., 15 (1994), 1354-1362.

[8] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 5th ed., Academic Press, 1994.

[9] G. H. Golub and C. F. Van Loan, *Matrix Computations*, third edition, The Johns Hopkins University Press (Baltimore), 1996.

[10] G. H. Hardy, J. E. Littlewood and G. Polya, *Inequalities*, second edition, Cambridge University Press, 1952.

[11] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, SIAM Publications, Philadelphia, PA, 1996.

[12] I. Olkin, *The 70th anniversary of the distribution of random matrices: A survey*, Linear Algebra Appls., 354 (2002), 231-243.

[13] D. S. Parker, *Random butterfly transformations with applications in computational linear algebra*, Technical Report CSD-950023, UCLA Computer Science Department, 1995.

[14] D. S. Parker and D. Le, *How to eliminate pivoting from Gaussian elimination — by randomizing instead*, Technical Report CSD-950037, UCLA Computer Science Department, 1995.

[15] A. Sankar, D. A. Spielman and S. H. Teng, *Smoothed Analysis of the Condition Numbers and Growth Factors of Matrices*, avaliable at http://math.mit.edu/~spielman/SmoothedAnalysis/index.html

[16] J. W. Silverstein, *The smallest eigenvalue of a large-dimensional Wishart matrix*, Ann. Prob. 13 (1985), 1364-1368.

[17] D. A. Spielman and S. H. Teng, *Smoothed Analysis of Algorithms: Why the Simplex Algorithm Usually Takes Polynomial Time*, In Proceedings of the 33rd Annual ACM Symposium on the Theory of Computing (STOC '01), pages 296-305, 2001.

[18] L. N. Trefethen and D. Bau, III, *Numerical Linear Algebra*, SIAM Publications, Philadelphia, PA, 1997.

[19] L. N. Trefethen and R. S. Schreiber, *Average-case stability of Gaussian elimination*, SIAM J. Matrix Anal. Appl., Vol. 11, No. 3, pp. 335-360, July 1990.

[20] H. F. Trotter, *Eigenvalue distributions of large Hermitian matrices; Wigner's semi-circle law and a theorem of Kac, Murdock, and Szegő*, Advances in Math. 54 (1984), 67-82.

[21] A. M. Turing, *Rounding-off errors in matrix processes*, Quart. J. Mech. Appl. Math., 1(1948), pp. 287-308.

[22] D. Viswanath and L. N. Trefethen, *Condition numbers of random triangular matrices*, SIAM J. Matrix Anal. Appl., Vol. 19, No. 2, pp. 564-581, April 1998.

[23] J. von Neumann and H. H. Goldstine, *Numerical inverting of matrices of high order*, Bull. Amer. Math. Soc., 53 (1947), pp. 1021-1099; von Neumann's Collected Works, Vol. 5, A. H. Taub, ed., Pergamon, Elmsford, NY, 1963.

[24] ——, *Numerical inverting of matrices of high order, Part II*, Proc. Amer. Math. Soc., 2 (1951), pp. 188-202; von Neumann's Collected Works, Vol. 5, A. H. Taub, ed., Pergamon Press, Elmsford, NY, 1963.

[25] J. H. Wilkinson, *Error analysis of direct methods of matrix inversion*, J. Assoc. Comput. Mach., 8 (1961), pp. 281-330.

[26] ——, *Rounding Errors in Algebraic Processes*, Prentice Hall, Englewood Cliffs, NJ, 1963.

[27] ——, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.

[28] S. J. Wright, *A collection of problems for which Gaussian elimination with partial pivoting is unstable*, SIAM J. Sci. Comput., 14 (1993), pp. 231-238.

[29] M. Yeung and T. Chan, *Probabilistic analysis of Gaussian elimination without pivoting*, SIAM J. Matrix Anal. Appl., Vol. 18, No. 2, pp. 499-517, April 1997.