Polynomial time ρ-locally maximum volume search

Alexander Osinsky[1]

Abstract
In this paper, we prove that a 3ρ-locally maximum volume submatrix $\hat{A} \in \mathbb{R}^{r \times r}$ in the matrix $A \in \mathbb{R}^{M \times N}$ can be found in $O\left(MNr\left(\log r + \log_\rho r\right)\right)$ operations, and a ρ-locally maximum volume submatrix for $\rho \leq 3$ can be found in $O\left(MNr^3 \log_\rho r\right)$ operations. Based on these submatrices, it is possible to construct a rank revealing LU decomposition with guarantees for the approximation accuracy in spectral and Chebyshev norms.

1 Introduction

To quickly construct low-rank matrix approximations, it is advantageous to use CUR (or CGR) cross approximation based on a small number of columns $C \in \mathbb{R}^{M \times r}$ and rows $R \in \mathbb{R}^{r \times N}$ of the original matrix $A \in \mathbb{R}^{M \times N}$. Generator U is commonly chosen as $U = \hat{A}^{-1}$, where $\hat{A} \in \mathbb{R}^{r \times r}$ is a submatrix at the intersection of rows R and columns C. Approximation $C\hat{A}^{-1}R$ has an important interpretation: it is equivalent to an incomplete LU decomposition of the matrix A with rows and columns rearranged accordingly. In other words, it is given by the first r steps of Gaussian elimination, with the choice of initial rows and columns corresponding to $\hat{A}$ submatrix. In practice, $C\hat{A}^{-1}R$ approximations are built using adaptive cross approximation (ACA) [1] or maxvol [2] algorithm. They also play an important role in the construction of tensor trains [3]. If the resulting submatrix turns out to be a submatrix of locally maximum

✉ Alexander Osinsky
osinskiy1189@gmail.com

1 Institute of Numerical Mathematics, Russian Academy of Sciences, Gubkina Street, 8, Moscow 119991, Russia

volume (we will give the exact definition later), then there exist error bounds for the $C\hat{A}^{-1}R$ approximation [4, 5].

Note that in many papers [4, 6, 7] the emphasis is placed on the submatrices of maximum volume. On the other hand, in order to achieve the corresponding estimates, it already suffices to find a submatrix of locally maximum volume. The difference between them is very significant, since finding a submatrix of the maximum volume or a volume close to the maximum is an NP-hard problem [8, 9]. Until now, it was not clear whether there is an algorithm that can find a submatrix $\hat{A} \in \mathbb{R}^{r \times r}$ of almost locally maximum volume in the matrix $A \in \mathbb{R}^{M \times N}$ of a general form in polynomial time. In this paper, such an algorithm is presented for the first time, estimates for the number of steps are given and the accuracy of the obtained approximation is proved. Obtaining a submatrix of almost locally maximum volume allows us to construct cross and incomplete LU decompositions with guaranteed accuracy in the spectral norm and Chebyshev norms.

Before we go any further, it is important to mention, that for symmetric positive semi-definite (SPSD) matrices locally maximum volume search is much easier, as one can only consider principle submatrices. A ρ-locally maximum volume principle submatrix can be found using the strong rank revealing Cholesky factorization algorithm [10]. Since any SPSD matrix $A \in \mathbb{R}^{N \times N}$ can be written in the form $B^T B$, $B \in \mathbb{R}^{N \times N}$, the search of the locally maximum volume principle submatrix is equivalent to the search for $N \times r$ locally maximum volume columns in B, which requires $O\left(r \log_\rho r\right)$ column exchanges [11, Theorem 2]. Consequently, strong rank revealing Cholesky factorization can also be achieved in $O\left(r \log_\rho r\right)$ exchanges, which require in total $O\left(Nr^2 \log_\rho r\right)$ operations, when using the algorithm from [10]. This complexity bound has been proven directly for SPSD matrices in [12] (Section 2.2.3).

2 Definitions

We will consider the problem of finding the submatrix $\hat{A}$ in terms of constructing the corresponding incomplete LU decomposition. Incomplete LU decomposition can serve as a criterion to estimate the rank of a matrix if the rows and columns are chosen appropriately. Such LU decompositions are called rank revealing LU (RRLU).

Similarly, RRQR is an abbreviation for the rank revealing QR decomposition. Strong RRQR decomposition plays an important role in constructing estimates and algorithms for RRLU decompositions.

Definition 1 Incomplete QR decomposition with column pivoting

$$A\Pi = Q \begin{bmatrix} R & B \\ & C \end{bmatrix} \in \mathbb{R}^{M \times N},$$

where $Q \in \mathbb{R}^{M \times M}$ is orthogonal and $R \in \mathbb{R}^{r \times r}$ is upper triangular is called "rank revealing", if

$$\sigma_r(R) \geq \sigma_r(A)/p_1(r, N)$$

and

$$\sigma_1(C) \leq \sigma_{r+1}(A)p_2(r, N),$$

where $p_1(r, N)$ and $p_2(r, N)$ are polynomials or some functions which can be bounded by polynomials in r and N.

It is called "spectrum revealing" if also

$$\sigma_i(R) \geq \sigma_i(A)/p_1(r, N), \quad i = \{1, \ldots, r-1\}.$$

It is called "strong" [13] if also

$$\sigma_i(C) \leq \sigma_{r+i}(A)p_2(r, N), \quad i = \{1, \ldots, \min(M-r, N-r)\}$$

and

$$\left\| R^{-1}B \right\|_C \leq f = \text{const},$$

where

$$\|A\|_C = \max_{i,j} \left| A_{ij} \right|$$

is the Chebyshev norm.

Similar definitions apply for $RRLU$.

Definition 2 Incomplete LU decomposition with row and column pivoting

$$\Pi_1 A \Pi_2 = \begin{bmatrix} L_{11} \\ L_{21} \ I_{M-r} \end{bmatrix} \begin{bmatrix} U_{11} \ U_{12} \\ \quad U_{22} \end{bmatrix} \in \mathbb{R}^{M \times N}, \tag{1}$$

where $L_{11} \in \mathbb{R}^{r \times r}$ is unit lower triangular and $U_{11} \in \mathbb{R}^{r \times r}$ is upper triangular is called "rank revealing", if

$$\sigma_r(L_{11}U_{11}) \geq \sigma_r(A)/p_1(r, M, N)$$

and

$$\sigma_1(U_{22}) \leq \sigma_{r+1}(A)p_2(r, M, N),$$

where $p_1(r, M, N)$ and $p_2(r, M, N)$ are polynomials or some functions which can be bounded by polynomials in r, M and N.

It is called "spectrum revealing" [14] if also

$$\sigma_i(L_{11}U_{11}) \geq \sigma_i(A)/p_1(r, M, N), \quad i = \{1, \ldots, r-1\}.$$

It is called "strong" [15] if also

$$\sigma_i(U_{22}) \le \sigma_{r+i}(A)p_2(r, M, N), \quad i = \{1, \dots, \min(M - r, N - r)\}, \quad (2)$$

$$\left\| L_{21}L_{11}^{-1} \right\|_C \le f = \text{const} \quad (3)$$

and

$$\left\| U_{11}^{-1}U_{12} \right\|_C \le f = \text{const.} \quad (4)$$

Construction of RRQR and RRLU decompositions can be considered equivalent to choosing r columns in the case of RRQR and r rows and columns in the case of RRLU, which are given by permutations of Π, Π_1 and Π_2 respectively. From an algorithmic perspective, searching for these permutations usually takes more time than constructing RRLU or RRQR based on rows and columns, which are already given. In particular, after choosing columns, constructing an incomplete QR decomposition takes $O\left(\text{nnz}(A)r + Mr^2\right)$ operations, where $\text{nnz}(A)$ is the number of nonzero elements of the matrix A. The construction of the factors $L = \begin{bmatrix} L_{11} \\ L_{21} \end{bmatrix}$ and $U = \begin{bmatrix} U_{11} & U_{12} \end{bmatrix}$ in the incomplete LU decomposition requires only $O\left((M + N)r^2\right)$ arithmetic operations (when the rows and columns, corresponding to L_{11} and U_{11}, are known). Due to the linear complexity of RRLU in terms of matrix sizes, algorithms for finding a large volume submatrix are often used, which also have linear complexity [1, 2]. They do not, however, guarantee high approximation accuracy precisely because that linear complexity does not allow them to consider the entire matrix and select close to optimal rows and columns.

As will be shown, to achieve strong RRLU decomposition, it suffices to choose permutations Π_1 and Π_2 so that the first r rows and columns correspond to a submatrix of ρ-locally maximum volume, $\rho = \text{const}$. Thus, the problem of constructing a strong RRLU decomposition is reduced to finding a submatrix of ρ-locally maximum volume and proving the "strong" properties of the corresponding decomposition.

Definition 3 The volume of the matrix $\hat{A} \in \mathbb{R}^{n \times r}$, $n \ge r$, is the following quantity:

$$\mathcal{V}\left(\hat{A}\right) = \sqrt{\det\left(A^T A\right)}.$$

In the case of a square matrix $(n = r)$,

$$\mathcal{V}\left(\hat{A}\right) = \left| \det \hat{A} \right|.$$

Definition 4 A full rank submatrix $\hat{A} \in \mathbb{R}^{m \times n}$ of a matrix $A \in \mathbb{R}^{M \times N}$ is said to have ρ-locally maximum volume (hereinafter, estimates with $\rho \ge 1$ are considered) if the permutation of its arbitrary i-th row and/or j-th column with any other row k and/or column l of the matrix A increases its volume by at most ρ times.

A submatrix is said to have locally maximum volume if $\rho = 1$.

Note that, unlike the definition in [16], one row and one column can be replaced at the same time.

Taking into account the definitions above, we can now formulate the main result of this paper.

Theorem 1 *Strong RRLU decomposition of rank r of the matrix $A \in \mathbb{R}^{M \times N}$ can be constructed on the basis of a submatrix of 3ρ-locally maximum volume of size $r \times r$. Moreover, such a submatrix can be found in $O\left(MNr\left(\log r + \log_\rho r\right)\right)$ operations, and ρ-locally maximum volume submatrix with $\rho \leq 3$ can be found in $O\left(MNr^3 \log_\rho r\right)$ operations.*

Before we move on to construct a decomposition, which satisfies Theorem 1, it is worth paying attention to other existing methods for constructing RRQR and RRLU decompositions with proven bounds on the number of steps and approximation error, since they can also be used to obtain the required spectral norm approximation accuracy.

3 Strong RRQR

We start with RRQR decomposition, since it can be used as a first step to constructing RRLU decomposition. The following theorem was proved in [13].

Theorem 2 *[13] Let the submatrix $A_1 \in \mathbb{R}^{M \times r}$ consisting of r columns of the matrix $A \in \mathbb{R}^{M \times N}$ have ρ-locally maximum volume.*

Then, based on the columns A_1, one can construct strong RRQR decomposition with

$$\left\| R^{-1} B \right\|_C \leq \rho$$

and

$$p_1(r, N) = p_2(r, N) = \sqrt{1 + \rho^2 r (N - r)}.$$

Moreover, a ρ-locally maximal submatrix can be found in $O\left(\left(nnz(A)r + (M + N)r^2\right)\left(1 + \log_\rho r\right)\right)$ operations, where $nnz(A)$ is the number of nonzero elements in the matrix A.

The complexity bound was first proved in [17] and then improved in [11].

4 Rank revealing LU

Now we move on to constructing RRLU decomposition. Recall the main theorem from [16] and an algorithm for obtaining bounds from it, based on strong RRQR [13]. Note that an alternative algorithm suggested in [16] does not have a bound on the number of steps.

Theorem 3 [16] *Let the submatrix $A_1 \in \mathbb{R}^{M \times r}$ consisting of r columns of the matrix $A \in \mathbb{R}^{M \times N}$ have ρ-locally maximum volume in A. Let the submatrix $\hat{A} \in \mathbb{R}^{r \times r}$ have ρ-local maximum volume in A_1.*

Then, based on the columns A_1 and the rows corresponding to $\hat{A}$, one can construct a spectrum revealing LU decomposition with

$$p_1(r, M, N) = p_2(r, M, N) = \sqrt{1 + \rho^2 r(N - r)}\sqrt{1 + \rho^2 r(M - r)}.$$

In addition,

$$\left\| L_{21} L_{11}^{-1} \right\|_C \leq \rho$$

and

$$\left\| A_1^\dagger A \right\|_C \leq \rho,$$

where $A_1^\dagger$ denotes the Moore–Penrose pseudoinverse of A_1.

Moreover, submatrices A_1 and $\hat{A}$ can be found in $O\left((nnz(A)r + (M + N)r^2)\left(1 + \log_\rho r\right)\right)$ operations, where $nnz(A)$ is the number of nonzero elements of the matrix A.

The bound on the complexity follows directly from Theorem 2.

Since in [16] it was not proved that the decomposition is spectrum revealing, we prove it here. To do this, we will apply spectrum revealing estimates for RRQR twice.

Using the spectrum revealing property of QR based on the submatrix A_1 of ρ-locally maximum volume (Theorem 2), we obtain the inequality

$$\sigma_i(A_1) \geq \sigma_i(A) / \sqrt{1 + \rho^2 r(N - r)}. \tag{5}$$

Next, columns $A_1^T \in \mathbb{R}^{r \times M}$ contain the submatrix $\hat{A}^T \in \mathbb{R}^{r \times r}$ of ρ-locally of maximum volume, and the QR decomposition $A_1^T \Pi = Q_1 [R_1 \ B_1]$, where $Q_1 R_1 = \hat{A}^T$, according to Theorem 2 has the following spectrum revealing property:

$$\sigma_i\left(\hat{A}^T\right) \geq \sigma_i\left(A_1^T\right) / \sqrt{1 + \rho^2(M - r)}. \tag{6}$$

Since $\hat{A} = L_{11} U_{11}$, and $A_1 = \begin{bmatrix} L_{11} \\ L_{21} \end{bmatrix} U_{11}$, then, combining (6) and (5) together leads to the following inequality:

$$
\begin{aligned}
\sigma_i(L_{11} U_{11}) &\geq \sigma_i\left(\begin{bmatrix} L_{11} \\ L_{21} \end{bmatrix} U_{11}\right) / \sqrt{1 + \rho^2 r(M - r)} \\
&= \sigma_i(A_1) / \sqrt{1 + \rho^2 r(M - r)}) \\
&\geq \sigma_i(A) / \left(\sqrt{1 + \rho^2 r(M - r)}\sqrt{1 + \rho^2 r(N - r)}\right), \tag{7}
\end{aligned}
$$

which proves the spectrum revealing property of the resulting RRLU decomposition.

The columns A_1 can be found numerically using the RRQR algorithm from [13], while the submatrix $\hat{A} \in \mathbb{R}^{r \times r}$ in the already fixed columns $A_1 \in \mathbb{R}^{M \times r}$ can also be found using the maxvol [2] algorithm.

Although we are not aware of an example, which shows that Theorem 3 may not provide the strong RRLU bounds (2), it can definitely violate the condition (4), which, as we will see later, is crucial for proving the bounds (2).

For example, consider the matrix

$$
A = \begin{bmatrix} 1 + \varepsilon & \sqrt{N} & \cdots & \sqrt{N} \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{N \times N}
$$

with arbitrary small ε and its rank 1 approximation, based on the Theorem 3. First, we select the locally maximum volume column, i.e., the column with the maximum length. This is always going to be the first column. Then we select 1×1 maximum volume submatrix in this column, which is going to be the element $1 + \varepsilon$. It differs from the locally maximum volume submatrix in the first row by a factor arbitrarily close to $\sqrt{N}$. Namely, we have $f = \sqrt{N}/(1 + \varepsilon)$, which depends on the matrix size and thus contradicts the "strong" property (4). Nevertheless, it does not directly prevent (2), which may or may not hold.

5 Strong RRLU

As we have seen, rank revealing LU from the previous section provides reasonable approximation accuracy in the spectral norm, from which the Chebyshev norm bound can also be derived. Nevertheless, finding close to locally maximum volume submatrix not only additionally guarantees the "strong" properties (2)–(4), but also provides a much better Chebyshev norm bound, which does not depend on matrix sizes [5].

For the first time, it was proposed to search for a submatrix of locally maximum volume in the entire matrix in [15]. In addition, the following criterion was proved. Here Eqs. (8) and (9) provide the same bounds as Eqs. (3) and (4).

Lemma 1 [15] *Consider a matrix $A \in \mathbb{R}^{M \times N}$ with the following block structure*

$$
A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},
$$

with

$$
C = \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} \in \mathbb{R}^{M \times r},
$$

and

$$R = \begin{bmatrix} A_{11} & A_{12} \end{bmatrix} \in \mathbb{R}^{r \times N}.$$

Then submatrix $\hat{A} = A_{11} \in \mathbb{R}^{r \times r}$ has ρ-locally maximum volume if and only if the following inequalities are true:

$$\left\| \hat{A}^{-1} R \right\|_C \le \rho, \tag{8}$$

$$\left\| C \hat{A}^{-1} \right\|_C \le \rho, \tag{9}$$

$$\max_{i,j,k,l} \left| \left(\hat{A}^{-1} R \right)_{jl} \left(C \hat{A}^{-1} \right)_{ki} + \hat{A}^{-1}_{ji} \left(A - C \hat{A}^{-1} R \right)_{kl} \right| \le \rho. \tag{10}$$

Moreover, when the i-th row of the submatrix $\hat{A}$ is replaced by the k-th row of the matrix A, the volume of the submatrix is multiplied by the factor $\left| \left(C \hat{A}^{-1} \right)_{ki} \right|$; when the j-th column of the submatrix $\hat{A}$ is replaced by the l-th column of the matrix A, the volume is multiplied by the factor $\left| \left(\hat{A}^{-1} R \right)_{jl} \right|$; when replacing a row and a column, the volume is multiplied by the factor $\left| \left(\hat{A}^{-1} R \right)_{jl} \left(C \hat{A}^{-1} \right)_{ki} + \hat{A}^{-1}_{ji} \left(A - C \hat{A}^{-1} R \right)_{kl} \right|$. Hereinafter $\hat{A}^{-1}_{ji}$ denotes the ji-th element of $\hat{A}^{-1}$.

Finally, the following theorem was proved.

Theorem 4 [15] *Based on rows $R \in \mathbb{R}^{r \times N}$ and columns $C \in \mathbb{R}^{M \times r}$ corresponding to submatrix $\hat{A} \in \mathbb{R}^{r \times r}$ of matrix $A \in \mathbb{R}^{M \times N}$ such that*

$$\left\| \hat{A}^{-1} R \right\|_C \le \sqrt{\rho},$$
$$\left\| C \hat{A}^{-1} \right\|_C \le \sqrt{\rho}$$

and

$$\left\| \hat{A}^{-1} \right\|_C \left\| A - C \hat{A}^{-1} R \right\|_C \le 2\rho \tag{11}$$

(it is also sufficient that $\hat{A}$ has $\sqrt{\rho}$-locally maximum volume in the entire matrix) one can construct a strong RRLU with

$$f = \sqrt{\rho},$$
$$p_1(r, M, N) = p_2(r, M, N) = \sqrt{(1 + 3\rho r(M - r))(1 + 3\rho r(N - r))}.$$

In [15] the result was formulated less rigorously, so we repeat its proof here.

Proof First, we prove that

$$\sigma_i (U_{22}) \leq \sqrt{(1+3\rho r (N-r)) (1+3\rho r (N-r))}\sigma_{r+i} (A) = \sigma_{r+i} (A) \, p_2 (r, M, N).$$

Like in Lemma 1, we can assume $\hat{A}$ is on the intersection of the first r rows and columns, $\hat{A} = A_{11}$, where the entire matrix A can be written in block form as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

Then $U_{22} = A_{22} - A_{21} A_{11}^{-1} A_{12}$ and submatrix $\hat{A} = A_{11}$ corresponds to the rows $R = \begin{bmatrix} A_{11} & A_{12} \end{bmatrix}$ and columns $C = \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix}$. The error of the corresponding cross approximation is

$$A - C \hat{A}^{-1} R = \begin{bmatrix} 0 & 0 \\ 0 & A_{22} - A_{21} A_{11}^{-1} A_{12} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & U_{22} \end{bmatrix}.$$

Consider the matrix

$$D = \begin{bmatrix} abA_{11} & 0 \\ 0 & A_{22} - A_{21} A_{11}^{-1} A_{12} \end{bmatrix}, \tag{12}$$

where

$$a = \sqrt{2\rho r (M - r)},$$
$$b = \sqrt{2\rho r (N - r)}.$$

Observe that

$$
\begin{aligned}
\left\| A_{11}^{-1} \right\|_2 \|A_{22} - A_{21} A_{11}^{-1} A_{12}\|_2 &= \left\| \hat{A}^{-1} \right\|_2 \|A - C \hat{A}^{-1} R\|_2 \\
&\leq r \left\| \hat{A}^{-1} \right\|_C \cdot \sqrt{(M - r) (N - r)} \|A - C \hat{A}^{-1} R\|_C \\
&\leq 2\rho r \sqrt{(M - r) (N - r)} \\
&= ab, \tag{13}
\end{aligned}
$$

where we used condition (11). Hereinafter

$$\|A\|_2 = \sigma_1 (A)$$

denotes the spectral norm.

Then, using (13), we can write the following inequalities for the singular values, corresponding to the submatrices of D:

$$\sigma_{\min} (abA_{11}) = ab\sigma_{\min} (A_{11}) = ab/ \left\| A_{11}^{-1} \right\|_2 \geq \|A_{22} - A_{21} A_{11}^{-1} A_{12}\|_2.$$

Since rank $A_{11} = r$, the first r largest singular values of D (12) correspond to the submatrix abA_{11}, so for the rest of its singular values

$$\sigma_{r+i}(D) = \sigma_i\left(A_{22} - A_{21}A_{11}^{-1}A_{12}\right).$$

Now, let us factorize matrix D as follows:

$$D = \begin{bmatrix} abA_{11} & 0 \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix} = \begin{bmatrix} aI & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} bI & -A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix}.$$

Then

$$\begin{aligned}
\sigma_i\left(A_{22} - A_{21}A_{11}^{-1}A_{12}\right) &= \sigma_{r+i}(D) \\
&\le \left\| \begin{bmatrix} aI & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix} \right\|_2 \sigma_{r+i}\left(\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \right) \left\| \begin{bmatrix} bI & -A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix} \right\|_2.
\end{aligned} \tag{14}$$

Spectral norms of the block matrices can be estimated as

$$\begin{aligned}
\left\| \begin{bmatrix} aI & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix} \right\|_2^2 &\le \|aI\|_2^2 + \left\| -A_{21}A_{11}^{-1} \right\|_2^2 + \|I\|_2^2 \\
&\le 1 + a^2 + r(M-r)\left\| A_{21}A_{11}^{-1} \right\|_C^2 \\
&\le 1 + 2\rho r(M-r) + r(M-r)\rho \\
&= 1 + 3\rho r(M-r).
\end{aligned} \tag{15}$$

Similarly,

$$\left\| \begin{bmatrix} bI & -A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix} \right\|_2^2 \le 1 + 3\rho r(N-r). \tag{16}$$

Substituting (15) and (16) into (14) leads us to the desired inequality:

$$\begin{aligned}
\sigma_i(U_{22}) = \sigma_i\left(A_{22} - A_{21}A_{11}^{-1}A_{12}\right) &\le \sqrt{(1 + 3\rho r(N-r))(1 + 3\rho r(N-r))}\sigma_{r+1}(A) \\
&= p_2(r, M, N)\sigma_{r+i}(A).
\end{aligned}$$

Next, we prove the "spectrum revealing" property for $p_1(r, M, N) = p_2(r, M, N)$:

$$\sigma_i(L_{11}U_{11}) = \sigma_i(A_{11}) \ge \sigma_i(A)/\sqrt{(1 + 3\rho r(N-r))(1 + 3\rho r(N-r))}. \tag{17}$$

To do that, we can use the following factorization:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ A_{21}A_{11}^{-1} & aI \end{bmatrix} \begin{bmatrix} A_{11} & 0 \\ 0 & \frac{1}{ab}\left(A_{22} - A_{21}A_{11}^{-1}A_{12}\right) \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ 0 & bI \end{bmatrix}.$$

Then

$$\sigma_i(A)$$
$$\leq \left\| \begin{bmatrix} I & 0 \\ A_{21}A_{11}^{-1} & aI \end{bmatrix} \right\|_2 \sigma_i\left(\begin{bmatrix} A_{11} & 0 \\ 0 & \frac{1}{ab}\left(A_{22} - A_{21}A_{11}^{-1}A_{12}\right) \end{bmatrix} \right) \left\| \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ 0 & bI \end{bmatrix} \right\|_2. \quad (18)$$

We again estimate the spectral norms, similar to (15) and (16):

$$\left\| \begin{bmatrix} I & 0 \\ A_{21}A_{11}^{-1} & aI \end{bmatrix} \right\|_2^2 \leq 1 + 3\rho r\,(M-r) \quad (19)$$

and

$$\left\| \begin{bmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & bI \end{bmatrix} \right\|_2^2 \leq 1 + 3\rho r\,(N-r). \quad (20)$$

Moreover,

$$\sigma_i\left(\begin{bmatrix} A_{11} & 0 \\ 0 & \frac{1}{ab}\left(A_{22} - A_{21}A_{11}^{-1}A_{12}\right) \end{bmatrix} \right) = \sigma_i\left(\frac{1}{ab}D \right) = \sigma_i(A_{11}) \quad (21)$$

as now we have $i \leq r$, and we have already shown that the first r singular values correspond to the submatrix A_{11}. Combining together Eqs. (18) through (21), we obtain

$$\sigma_i(A) \leq \sqrt{(1 + 3\rho r\,(N-r))(1 + 3\rho r\,(N-r))}\sigma_i(A_{11}). \quad (22)$$

Equation (17) follows directly from (22). $\square$

However, it is still not clear how to find a submatrix of ρ-locally maximum volume fast enough. The authors themselves state in [14] that there are still no effective algorithms for its search. Let us show that it is possible to reach the bounds of Theorem 4 fairly quickly.

We will start from the initial $r \times r$ submatrix, constructed using Gaussian elimination with complete pivoting. For our purposes, we will require the following result, which is also useful for adaptive cross approximation [18].

Lemma 2 [18] *Let the rows of the submatrix* $\hat{A} \in \mathbb{R}^{r \times r}$ *be obtained by Gaussian elimination with partial pivoting in the columns* $C \in \mathbb{R}^{M \times r}$. *Then*

$$\left\| C\hat{A}^{-1} \right\|_C \leq 2^{r-1}.$$

Obviously, the same result will hold for complete pivoting. In complete pivoting, we will also have the inequality for $\left\|\hat{A}^{-1}R\right\|_C \leq 2^{r-1}$. Moreover, in [19] the following result was proved.

Lemma 3 [19] *Let the rows $R \in \mathbb{R}^{M \times r}$ and columns $C \in \mathbb{R}^{r \times N}$ corresponding to submatrix $\hat{A} \in \mathbb{R}^{r \times r}$ be the first r rows and columns selected by Gaussian elimination with complete pivoting for the matrix $A \in \mathbb{R}^{M \times N}$. Then*

$$\left\|A - C\hat{A}^{-1}R\right\|_C \leq 4^r \gamma_r \sigma_{r+1}(A),$$

where $\gamma_r \leq 2\sqrt{r+1}(r+1)^{\frac{\ln(r+1)}{4}}$ is the growth factor.

Corollary 1 *Under the conditions of Lemma 3,*

$$\left\|A - C\hat{A}^{-1}R\right\|_C \leq 4^{r-1/2} r^{\frac{2+\ln r}{4}} \left\|\hat{A}^{-1}\right\|_2^{-1}.$$

Proof Consider the submatrix $\bar{A} \in \mathbb{R}^{(r-1) \times (r-1)}$, obtained after $r-1$ steps of Gaussian elimination. Applying Lemma 3 to the matrix $\hat{A}$ (instead of A), after $r-1$ steps (instead of r) we have

$$\left\|\hat{A} - \hat{C}\bar{A}^{-1}\hat{R}\right\|_C \leq 4^{r-1} \gamma_{r-1} \sigma_r\left(\hat{A}\right),$$

where $\hat{R} \in \mathbb{R}^{(r-1) \times r}$ and $\hat{C} \in \mathbb{R}^{r \times (r-1)}$ are rows and columns of $\hat{A}$, corresponding to $\bar{A}$.

Then, by definition of the growth factor, we can estimate the error at the r-th step of Gaussian elimination:

$$\begin{aligned}
\left\|A - C\hat{A}^{-1}R\right\|_C &\leq \gamma_1 \left\|\hat{A} - \hat{C}\bar{A}^{-1}\hat{R}\right\|_C \\
&\leq 2 \left\|\hat{A} - \hat{C}\bar{A}^{-1}\hat{R}\right\|_C \\
&\leq 4^{r-1/2} \gamma_{r-1} \sigma_r\left(\hat{A}\right) \\
&\leq 4^{r-1/2} r^{\frac{2+\ln r}{4}} \sigma_r\left(\hat{A}\right) \\
&= 4^{r-1/2} r^{\frac{2+\ln r}{4}} \left\|\hat{A}^{-1}\right\|_2^{-1}.
\end{aligned}$$

$\square$

Combining Lemma 2 and Corollary 1, we see that Gaussian elimination produces a ρ-locally maximum volume submatrix, albeit with quite a large ρ.

Corollary 2 *Let the rows $R \in \mathbb{R}^{M \times r}$ and columns $C \in \mathbb{R}^{r \times N}$ corresponding to submatrix $\hat{A} \in \mathbb{R}^{r \times r}$ be the first r rows and columns selected by Gaussian elimination*

with complete pivoting for the matrix $A \in \mathbb{R}^{M \times N}$. Then $\hat{A}$ is a ρ-locally maximum volume submatrix with

$$\rho \leq 3 \cdot 4^{r-1} r^{\frac{2+\ln r}{4}}.$$

Proof Comparing Lemma 2 with the conditions (8) and (9) of Lemma 1, we see that they are satisfied for $\rho = 2^{r-1}$. Let us now check the last condition, where we substitute the inequalities, provided by Lemma 2 and Corollary 1:

$$\max_{i,j,k,l} \left| \left(\hat{A}^{-1} R \right)_{ik} \left(C \hat{A}^{-1} \right)_{jl} + \hat{A}_{ij}^{-1} \left(A - C \hat{A}^{-1} R \right)_{lk} \right|$$
$$\leq \left\| C \hat{A}^{-1} \right\|_C \left\| \hat{A}^{-1} R \right\|_C + \left\| \hat{A}^{-1} \right\|_C \left\| A - C \hat{A}^{-1} R \right\|_C$$
$$\leq 4^{r-1} + 4^{r-1/2} r^{\frac{2+\ln r}{4}}$$
$$\leq 3 \cdot 4^{r-1} r^{\frac{2+\ln r}{4}}.$$

$\square$

Finally, we will need a result on the closeness of the volume of a submatrix of locally maximum volume to the volume of the maximum volume submatrix.

Lemma 4 *Let $\hat{A} \in \mathbb{C}^{r \times r}$ be a ρ-locally maximum volume in the matrix $A \in \mathbb{C}^{M \times N}$. Let $A_M \in \mathbb{C}^{r \times r}$ be a submatrix of maximum volume in A. Then*

$$\mathcal{V}(A_M) \leq \mathcal{V}(\hat{A}) \left(1 + 3\rho^2 r^2 \right)^r.$$

Proof Let us start with the observation, that there exists a submatrix $\tilde{A} \in \mathbb{C}^{m \times n}$ with $m, n \leq 2r$, which contains the submatrices $\hat{A}$ and A_M.

According to Theorem 4, when constructing LU for $\tilde{A}$ based on $\hat{A}$,

$$p_1(r, m, n) \leq \sqrt{\left(1 + 3\rho^2 r(m-r) \right) \left(1 + 3\rho^2 r(n-r) \right)} \leq 1 + 3\rho^2 r^2,$$

and therefore

$$\sigma_i(\hat{A}) \geq \sigma_i(\tilde{A}) / p_1(r, m, n) \geq \sigma_i(\tilde{A}) / (1 + 3\rho^2 r^2), \quad i = \{1, \ldots, r\}.$$

Taking the product over all singular values, we get

$$\mathcal{V}(A_M) = \prod_{i=1}^{r} \sigma_i(A_M)$$
$$\leq \prod_{i=1}^{r} \sigma_i(\tilde{A})$$

$$\leq \left(1 + 3\rho^2 r^2\right)^r \prod_{i=1}^{r} \sigma_i\left(\hat{A}\right)$$

$$\leq \mathcal{V}(\hat{A})\left(1 + 3\rho^2 r^2\right)^r.$$

$\square$

We are now ready to prove the main Theorem 1. To do that, we are going to use the following Algorithm 1.

Algorithm 1 3ρ-locally maximum volume search

Input: Matrix $A \in \mathbb{R}^{M \times N}$, required rank r, parameter $\rho \geq 1$.
Output: Submatrix $\hat{A} \in \mathbb{R}^{r \times r}$ of 3ρ-locally maximum volume.
1: Run r steps of Gaussian elimination with full pivoting to find the starting submatrix $\hat{A} \in \mathbb{R}^{r \times r}$, containing the pivots
2: Let $\hat{A}$ be on the intersection of rows $R \in \mathbb{R}^{r \times N}$ and columns $C \in \mathbb{R}^{M \times r}$
3: **loop**
4:    Find volume growth for a single row swap:
5:    $i_1, k_1 = \arg\max_{i,k} \left|\left(C\hat{A}^{-1}\right)_{ki}\right|, j_1 = l_1 = 1$
6:    $\rho_1 = \max_{i,k} \left|\left(C\hat{A}^{-1}\right)_{ki}\right|$
7:    Find volume growth for a single column swap:
8:    $j_2, l_2 = \arg\max_{j,l} \left|\left(\hat{A}^{-1}R\right)_{jl}\right|, i_2 = k_2 = 1$
9:    $\rho_2 = \max_{j,l} \left|\left(\hat{A}^{-1}R\right)_{jl}\right|$
10:    Find volume growth for the simultaneous swap of a row and a column:
11:    $i_3, j_3, k_3, l_3 = \arg\max_{i,j,k,l} \left|\hat{A}_{ji}^{-1}\right| \left|\left(A - C\hat{A}^{-1}R\right)_{kl}\right|$
12:    $\rho_3 = \left|\left(\hat{A}^{-1}R\right)_{j_3 l_3}\left(C\hat{A}^{-1}\right)_{k_3 i_3} + \hat{A}_{j_3 i_3}^{-1}\left(A - C\hat{A}^{-1}R\right)_{k_3 l_3}\right|$
13:    Select maximum estimated growth:
14:    $\rho_n = \max\left(\rho_1, \rho_2, \rho_3\right), n \in \{1, 2, 3\}$.
15:    **if** $\rho_n > \sqrt{\rho}$ **then**
16:        Swap the rows i_n and k_n
17:        Swap the columns j_n and l_n
18:        Update $\hat{A}^{-1}, C\hat{A}^{-1}, \hat{A}^{-1}R$ and $A - C\hat{A}^{-1}R$
19:    **else**
20:        **break**
21:    **end if**
22: **end loop**

It starts from Gaussian elimination, and then exchanges rows and columns based on the best of three values ρ_1, ρ_2 and ρ_3. They correspond to the estimated volume increase after exchanging a row, a column, or both a row and a column respectively. Note, that if it stops, the following inequalities are satisfied:

$$\left\| \hat{A}^{-1} R \right\|_C \leq \sqrt{\rho},$$

$$\left\| C \hat{A}^{-1} \right\|_C \leq \sqrt{\rho},$$

$$\left\| \hat{A}^{-1} \right\|_C \left\| A - C \hat{A}^{-1} R \right\|_C \leq 2\rho. \tag{23}$$

The first two follow directly from the fact that $\max(\rho_1, \rho_2) \leq \rho_n \leq \sqrt{\rho}$ (inequality on line 15 is false if "break" happened), and the third one follows from

$$
\begin{aligned}
\rho_3 &= \left| \left(\hat{A}^{-1} R \right)_{j_3 l_3} \left(C \hat{A}^{-1} \right)_{k_3 i_3} + \hat{A}^{-1}_{j_3 i_3} \left(A - C \hat{A}^{-1} R \right)_{k_3 l_3} \right| \\
&\geq \left| \hat{A}^{-1}_{j_3 i_3} \left(A - C \hat{A}^{-1} R \right)_{k_3 l_3} \right| - \left| \left(\hat{A}^{-1} R \right)_{j_3 l_3} \left(C \hat{A}^{-1} \right)_{k_3 i_3} \right| \\
&= \max_{i,j} \left| \hat{A}^{-1}_{ji} \right| \max_{k,l} \left| \left(A - C \hat{A}^{-1} R \right)_{kl} \right| - \left| \left(\hat{A}^{-1} R \right)_{j_3 l_3} \left(C \hat{A}^{-1} \right)_{k_3 i_3} \right| \\
&\geq \max_{i,j} \left| \hat{A}^{-1}_{ji} \right| \max_{k,l} \left| \left(A - C \hat{A}^{-1} R \right)_{kl} \right| - \max_{j_2 l_2} \left| \left(\hat{A}^{-1} R \right)_{j_2 l_2} \right| \max_{i_1 k_1} \left| \left(C \hat{A}^{-1} \right)_{i_1 k_1} \right| \\
&= \max_{i,j} \left| \hat{A}^{-1}_{ji} \right| \max_{k,l} \left| \left(A - C \hat{A}^{-1} R \right)_{kl} \right| - \rho_2 \rho_1 \\
&= \left\| \hat{A}^{-1} \right\|_C \left\| A - C \hat{A}^{-1} R \right\|_C - \rho_2 \rho_1,
\end{aligned}
$$

so

$$\left\| \hat{A}^{-1} \right\|_C \left\| A - C \hat{A}^{-1} R \right\|_C \leq \rho_2 \rho_1 + \rho_3 \leq \rho + \sqrt{\rho} \leq 2\rho.$$

From the conditions (23) we also find that the output submatrix has 3ρ-locally maximum volume. According to Lemma 1, it is enough to satisfy the conditions (8)–(10). First two follow directly from (23) and the last condition (with 3ρ in place of ρ) follows from

$$
\begin{aligned}
\max_{i,j,k,l} &\left| \left(\hat{A}^{-1} R \right)_{jl} \left(C \hat{A}^{-1} \right)_{ki} + \hat{A}^{-1}_{ji} \left(A - C \hat{A}^{-1} R \right)_{kl} \right| \\
&\leq \max_{i,j,k,l} \left| \left(\hat{A}^{-1} R \right)_{jl} \left(C \hat{A}^{-1} \right)_{ki} \right| + \max_{i,j,k,l} \left| \hat{A}^{-1}_{ji} \left(A - C \hat{A}^{-1} R \right)_{kl} \right| \\
&\leq \left\| \hat{A}^{-1} R \right\|_C \left\| C \hat{A}^{-1} \right\|_C + \left\| \hat{A}^{-1} \right\|_C \left\| A - C \hat{A}^{-1} R \right\|_C \leq \rho + 2\rho \leq 3\rho.
\end{aligned}
$$

Algorithm 1 also does eventually stop, as the volume is multiplied at least by a factor $\max(\rho_1, \rho_2, \rho_3) > \sqrt{\rho} \geq 1$ at each iteration. Thus, it can't visit the same submatrix twice, and there is a finite number of submatrices to check.

    If we want to reach a ρ-locally maximum volume submatrix, we first run Algorithm 1 with $\rho = \sqrt{3}$ (although any constant greater than 1 would suffice) and then improve it making exchanges following Lemma 1. According to Lemma 1, Algorithm 2

eventually outputs a ρ-locally maximum volume submatrix. And it also eventually stops for the same reason.

Algorithm 2 ρ-locally maximum volume search

Input: Matrix $A \in \mathbb{R}^{M \times N}$, required rank r, parameter $\rho \geq 1$.
Output: Submatrix $\hat{A} \in \mathbb{R}^{r \times r}$ of ρ-locally maximum volume.
1: Run Algorithm 1 with $\rho = \sqrt{3}$
2: Let the output submatrix $\hat{A}$ be on the intersection of rows $R \in \mathbb{R}^{r \times N}$ and columns $C \in \mathbb{R}^{M \times r}$
3: **loop**
4:   Find maximum volume growth ($i = k$ and $j = l$ are allowed):
5:   $i, j, k, l = \underset{i,j,k,l}{\arg\max} \left| \left(\hat{A}^{-1} R \right)_{jl} \left(C \hat{A}^{-1} \right)_{ki} + \hat{A}_{ji}^{-1} \left(A - C \hat{A}^{-1} R \right)_{kl} \right|$
6:   $\rho_{\max} = \left| \left(\hat{A}^{-1} R \right)_{jl} \left(C \hat{A}^{-1} \right)_{ki} + \hat{A}_{ji}^{-1} \left(A - C \hat{A}^{-1} R \right)_{kl} \right|$
7:   **if** $\rho_{\max} > \rho$ **then**
8:     Swap the rows i and k
9:     Swap the columns j and l
10:     Update $\hat{A}^{-1}$, $C \hat{A}^{-1}$, $\hat{A}^{-1} R$ and $A - C \hat{A}^{-1} R$
11:   **else**
12:     **break**
13:   **end if**
14: **end loop**

We now estimate the number of steps required to reach 3ρ-locally maximum volume and ρ-locally maximum volume submatrices in the following proposition.

Proposition 1 *Algorithm 1 stops after $O\left(r \left(1 + \frac{1}{\log \rho} \right) \log r \right)$ iterations and has the total computational cost of $O\left(MNr \left(1 + \frac{1}{\log \rho} \right) \log r \right)$.*

Algorithm 2 performs no more than $O\left(r \log_\rho r \right)$ additional iterations and has the total computational cost of $O\left(MNr \left(1 + \frac{r^2}{\log \rho} \right) \log r \right)$.

Proof The starting submatrix $A_{(0)}$ is chosen using Gaussian elimination with complete pivoting. By Corollary 2, $A_{(0)}$ has $\rho_{(0)}$-locally maximum volume with

$$\rho_{(0)} \leq 3 \cdot 4^{r-1} r^{\frac{2 + \ln r}{4}}. \tag{24}$$

Consider now the sequence of submatrices $\hat{A} = A_{(s)}$ obtained after s steps of column and/or row exchanges. According to Lemma 4,

$$\mathcal{V}(A_M) \leq \mathcal{V}(A_{(s)}) \left(3\rho_{(s)}^2 (r^2 + 1) \right)^r,$$

$$\rho_{(s)} \geq \frac{1}{\sqrt{3(r^2 + 1)}} \left(\frac{\mathcal{V}(A_M)}{\mathcal{V}(A_{(s)})} \right)^{\frac{1}{2r}}, \tag{25}$$

where $A_M \in \mathbb{R}^{r \times r}$ is the maximum volume submatrix.

At each step of the Algorithm 1, we make an exchange according to the following rule. We select a new row k and/or a new column l that match the value of

$$
\begin{aligned}
\rho' &= \max\left(\left| \left(\hat{A}^{-1}R\right)_{j_2 l_2} \right|, \left| \left(C\hat{A}^{-1}\right)_{k_1 i_1} \right|, \right. \\
&\qquad \left. \left| \hat{A}^{-1}_{j_3 i_3}\left(A - C\hat{A}^{-1}R\right)_{k_3 l_3} + \left(\hat{A}^{-1}R\right)_{j_3 l_3}\left(C\hat{A}^{-1}\right)_{k_3 i_3} \right| \right) \\
&\geq \max\left(\max_{j,l} \left| \left(\hat{A}^{-1}R\right)_{jl} \right|, \max_{i,k} \left| \left(C\hat{A}^{-1}\right)_{ki} \right|, \right. \\
&\qquad \left. \max_{i,j} \left| \hat{A}^{-1}_{ji} \right| \max_{k,l} \left| \left(A - C\hat{A}^{-1}R\right)_{kl} \right| - \left| \left(\hat{A}^{-1}R\right)_{j_3 l_3} \right| \left| \left(\hat{C}A^{-1}\right)_{k_3 i_3} \right| \right) \\
&\geq \max\left(\max_{j,l} \left| \left(\hat{A}^{-1}R\right)_{jl} \right|, \max_{i,k} \left| \left(C\hat{A}^{-1}\right)_{ki} \right|, \right. \\
&\qquad \left. \max_{i,j} \left| \hat{A}^{-1}_{ji} \right| \max_{k,l} \left| \left(A - C\hat{A}^{-1}R\right)_{kl} \right| - \max_{j_2,l_2} \left| \left(\hat{A}^{-1}R\right)_{j_2 l_2} \right| \max_{i_1,k_1} \left| \left(\hat{C}A^{-1}\right)_{k_1 i_1} \right| \right), \quad (26)
\end{aligned}
$$

where we substituted the definitions of the indices $i_1, k_1, j_2, k_2, i_3, j_3, k_3, l_3$ from the Algorithm 1.

Such a choice increases the volume by at least ρ' times, since we have chosen $\rho' = \max\left(\rho_1, \rho_2, \rho_3\right)$.

Moreover, if at the s-th step we have exactly $\rho_{(s)}$-locally maximum volume submatrix with $\rho_{(s)} \geq 3$, then $\rho' \geq \sqrt{\rho_{(s)}/3}$. Indeed, if $\max_{j,l} \left| \hat{A}^{-1}R \right|_{jl} < \sqrt{\rho_{(s)}/3}$ and $\max_{i,k} \left| C\hat{A}^{-1} \right|_{ki} < \sqrt{\rho_{(s)}/3}$, then

$$
\begin{aligned}
\max_{i,j} &\left| \hat{A}^{-1} \right|_{ji} \max_{k,l} \left| A - C\hat{A}^{-1}R \right|_{kl} - \max_{j_2,l_2} \left| \left(\hat{A}^{-1}R\right)_{j_2 l_2} \right| \max_{i_1,k_1} \left| \left(C\hat{A}^{-1}\right)_{k_1 i_1} \right| \\
&\geq \max_{i,j,k,l} \left| \hat{A}^{-1}_{ji}\left(A - C\hat{A}^{-1}R\right)_{kl} + \left(\hat{A}^{-1}R\right)_{jl}\left(C\hat{A}^{-1}\right)_{ki} \right| \\
&\quad - 2\max_{j_2,l_2} \left| \left(\hat{A}^{-1}R\right)_{j_2 l_2} \right| \max_{i_1,k_1} \left| \left(C\hat{A}^{-1}\right)_{k_1 i_1} \right| \\
&\geq \rho_{(s)} - 2\frac{\rho_{(s)}}{3} = \rho_{(s)}/3 \geq \sqrt{\rho_{(s)}/3}.
\end{aligned}
$$

Taking (25) into account, we get that the ratio to the maximum volume changes as

$$
\frac{\mathcal{V}(A_M)}{\mathcal{V}(A_{(s+1)})} \leq \frac{\mathcal{V}(A_M)}{\mathcal{V}(A_{(s)})} \cdot \sqrt{\frac{3}{\rho_{(s)}}} \leq \sqrt{3\sqrt{3(r^2+1)}}\left(\frac{\mathcal{V}(A_M)}{\mathcal{V}(A_{(s)})}\right)^{1-\frac{1}{4r}}. \quad (27)
$$

Let us denote

$$
\alpha_{(s)} = \frac{\left(\frac{\mathcal{V}(A_M)}{\mathcal{V}(A_{(s)})}\right)^{\frac{1}{4r}}}{\sqrt{3\sqrt{3(r^2+1)}}}. \quad (28)
$$

From (24) and Lemma 4

$$\alpha_{(0)} \leq \sqrt{\rho_{(0)}/3} \leq 2^{r-1} r^{\frac{2+\ln r}{8}} \leq 2^r r^{\frac{2+\ln r}{8}}. \qquad (29)$$

Substituting $\alpha_{(s)}$ (28) into (27), we can estimate how $\alpha_{(s)}$ changes after each step:

$$\left(27(r^2+1)\right)^r \alpha_{(s+1)}^r \leq \left(27(r^2+1)\right)^r \alpha_{(s)}^{4r-1},$$

$$\alpha_{(s+1)} \leq \alpha_{(s)}^{1-\frac{1}{4r}},$$

$$\alpha_{(s)} \leq \alpha_{(0)}^{\left(1-\frac{1}{4r}\right)^s},$$

$$\ln \ln \alpha_{(s)} - \ln \ln \alpha_{(0)} \leq s \ln \left(1 - \frac{1}{4r}\right) \leq -\frac{s}{4r}. \qquad (30)$$

Thus, to reach, for instance, $\alpha_{(s)} \leq 3^{1/4} < 2^r r^{\frac{2+\ln r}{8}}$, one would need

$$s_1 \leq 4r \ln \ln \left(2^r r^{\frac{2+\ln r}{8}}\right) - 4r \ln \ln 3^{1/4} + 1$$
$$= O\left(r \log r\right) \qquad (31)$$

steps. After that

$$\frac{\mathcal{V}(A_M)}{\mathcal{V}(A_{(s_1)})} \leq \left(\sqrt{3\sqrt{3(r+1)}}\alpha_{(s_1)}\right)^{4r} \leq \left(81\left(r^2+1\right)\right)^r.$$

Then, while $\rho' \geq \sqrt{\rho}$, the Algorithm 1 increases the volume by at least a factor of $\sqrt{\rho}$ at every step. Since we cannot exceed the maximum volume, the algorithm requires at most

$$s_2 \leq \log_{\sqrt{\rho}} \left(81\left(r^2+1\right)\right)^r = O\left(r\frac{\log r}{\log \rho}\right) \qquad (32)$$

steps until termination.

Each step of the algorithm is an update of rank no greater than 2 for the submatrices $\hat{A}$, C, and R, since no more than one row and no more than one column is replaced. It follows that the recalculation of $C\hat{A}^{-1}$, $\hat{A}^{-1}R$ and $C\hat{A}^{-1}R$ is also of low rank and requires a total of at most $O(MN)$ operations per step, after which the choice of indices i, j, k, l based on ρ' (26) also takes at most $O(MN)$ operations. Given the constraints (31) and (32), the total number of steps is $s_1 + s_2 = O\left(r\left(1 + \frac{1}{\log \rho}\right)\log r\right)$, and the total computational complexity of the algorithm is $O\left(MNr\left(1 + \frac{1}{\log \rho}\right)\log r\right)$. Gaussian elimination with complete pivoting takes $O(MNr)$ operations, which does not affect the general asymptotics.

If it is necessary to search for an ρ-locally maximum volume with $\rho \leq 3$, then we can also make exchanges of rows and columns after 3ρ-locally maximum volume is reached for $\rho = \sqrt{3}$ (which will take $O\left(MNr \log r\right)$ operations), but then checking

$$\max_{i,j,k,l} \left| \left(\hat{A}^{-1}R\right)_{jl} \left(C\hat{A}^{-1}\right)_{ki} + \hat{A}_{ji}^{-1} \left(A - C\hat{A}^{-1}R\right)_{kl} \right| > \rho,$$

which is what Algorithm 2 does.

Since in this case we guarantee an increase in volume by at least ρ times, with Lemma 4 the number of additional steps can be bounded by

$$s_3 \leq \log_\rho \left(1 + 3\left(3\sqrt{3}\right)^2 r^2\right)^r = \log_\rho \left(1 + 81r^2\right)^r = O\left(r \log_\rho r\right).$$

However, now at each step it is required to check all MNr^2 variants of the quadruple i, j, k, l, so these (at most) s_3 steps are r^2 times more expensive. Thus, the total cost of the search for ρ-locally maximum volume is $O\left(MNr\left(1 + \frac{r^2}{\log \rho}\right) \log r\right)$. For $\rho \leq 3$ it simplifies to $O\left(MNr^3 \log_\rho r\right)$. $\qquad\qquad\square$

Note that, as shown in [5], cross approximation based on a ρ-locally maximum volume submatrix also allows achieving high accuracy in the Chebyshev norm (which does not depend on matrix sizes). Let us formulate this fact in terms of the LU decomposition.

Theorem 5 [5] *Let an incomplete LU decomposition* (1) *be constructed based on the submatrix* $\hat{A} = L_{11}U_{11} \in \mathbb{R}^{r \times r}$ *of ρ-locally maximal volume. Then*

$$\|A - LU\|_C \leq \rho \left(r + 1\right)^2 \|E\|_C, \tag{33}$$

where matrix E is the error of the best approximation in the Chebyshev norm.

So, finding 3ρ-locally maximum volume submatrix also guarantees a good Chebyshev norm bound.

An algorithm, which allows to guarantee the bound (33) was recently suggested in [20]. Namely, it was shown that a factor of ρ in the Chebyshev norm can be reached without considering all MNr^2 possible exchanges. Following [20, Lemma 3.15], we show that it is possible to quickly reach a factor of ρ instead of 3ρ in Chebyshev norm, even when we only guarantee to find a 3ρ-locally maximum volume submatrix. It should be noted, however, that the algorithm, described in [20], is asymptotically slower and that reaching Chebyshev norm bound does not, by itself, guarantee either "spectrum revealing" or "strong" RRLU properties (see Definition 2), which are necessary to obtain a ρ-locally maximum volume submatrix (see Theorem 4).

To guarantee the bound (33), we calculate the value ρ_3 in the Algorithm 1 more efficiently, by directly finding the maximum possible volume multiplier over all i and j. We write the result as Algorithm 3.

Algorithm 3 3ρ-locally maximum volume search with better Chebyshev norm error guarantees

---

Input: Matrix $A \in \mathbb{R}^{M \times N}$, required rank r, parameter $\rho \geq 1$.
Output: Submatrix $\hat{A} \in \mathbb{R}^{r \times r}$ of 3ρ-locally maximum volume.
1: Run r steps of Gaussian elimination with full pivoting to find the starting submatrix $\hat{A} \in \mathbb{R}^{r \times r}$, containing the pivots
2: Let $\hat{A}$ be on the intersection of rows $R \in \mathbb{R}^{r \times N}$ and columns $C \in \mathbb{R}^{M \times r}$
3: **loop**
4:   Find volume growth for a single row swap:
5:   $i_1, k_1 = \arg\max\limits_{i,k} \left| \left(C\hat{A}^{-1} \right)_{ki} \right|, \; j_1 = l_1 = 1$
6:   $\rho_1 = \max\limits_{i,k} \left| \left(C\hat{A}^{-1} \right)_{ki} \right|$
7:   Find volume growth for a single column swap:
8:   $j_2, l_2 = \arg\max\limits_{j,l} \left| \left(\hat{A}^{-1} R \right)_{jl} \right|, \; i_2 = k_2 = 1$
9:   $\rho_2 = \max\limits_{j,l} \left| \left(\hat{A}^{-1} R \right)_{jl} \right|$
10:   Find volume growth for k and l corresponding to the maximum error:
11:   $k_3, l_3 = \arg\max\limits_{k,l} \left| \left(A - C\hat{A}^{-1} R \right)_{kl} \right|$
12:   $i_3, j_3 = \arg\max\limits_{i,j} \left| \left(\hat{A}^{-1} R \right)_{jl_3} \left(C\hat{A}^{-1} \right)_{k_3 i} + \hat{A}_{ji}^{-1} \left(A - C\hat{A}^{-1} R \right)_{k_3 l_3} \right|$
13:   $\rho_3 = \left| \left(\hat{A}^{-1} R \right)_{j_3 l_3} \left(C\hat{A}^{-1} \right)_{k_3 i_3} + \hat{A}_{j_3 i_3}^{-1} \left(A - C\hat{A}^{-1} R \right)_{k_3 l_3} \right|$
14:   Select maximum estimated growth:
15:   $\rho_n = \max\left(\rho_1, \rho_2, \rho_3 \right), \; n \in \{1, 2, 3\}.$
16:   **if** $\rho_n > \sqrt{\rho}$ **then**
17:     Swap the rows i_n and k_n
18:     Swap the columns j_n and l_n
19:     Update $\hat{A}^{-1}$, $C\hat{A}^{-1}$, $\hat{A}^{-1} R$ and $A - C\hat{A}^{-1} R$
20:   **else**
21:     **break**
22:   **end if**
23: **end loop**

Proposition 2 *Algorithm 3 requires at most* $O\left(MNr \left(1 + \frac{1}{\log \rho} \right) \log r \right)$ *operations, guarantees the same bounds as Algorithm 1 and also allows to construct an incomplete LU decomposition, such that*

$$\| A - LU \|_C \leq \rho (r+1)^2 \| E \|_C. \tag{34}$$

Proof In Algorithm 3, calculating ρ_3 allows additionally checking the condition

$$\left| \left(\hat{A}^{-1} R \right)_{jl_3} \left(C\hat{A}^{-1} \right)_{k_3 i} + \hat{A}_{ji}^{-1} \left(A - C\hat{A}^{-1} R \right)_{k_3 l_3} \right| \leq \sqrt{\rho} \tag{35}$$

for all i and j and for

$$k_3, l_3 = \arg\max\limits_{k,l} \left| \left(A - C\hat{A}^{-1} R \right)_{kl} \right|.$$

If, given k_3 and l_3, it is possible to find i and j such that the condition (35) is violated, then the increase in volume with the corresponding replacement is at least $\sqrt{\rho}$ times, according to the Lemma 1. Only if this growth is higher than all other options for ρ' (26), then we make this replacement. We note right away that this will not increase the cost of the algorithm, since the cost of going through all pairs of i and j requires only $O(r^2)$ arithmetic operations, which does not increase the asymptotic cost of one step of the algorithm. In addition, the bound on the number of steps also does not change, as each increase in volume is not lower than before.

Now we show that when the algorithm stops, the inequality (34) holds. We argue from the contrary: if the inequality for the Chebyshev norm fails, then, in particular, it fails for those k and l where the error reaches its maximum. It is these new row k_3 and new column l_3 that we are checking. Consider the submatrix $\tilde{A} \in \mathbb{R}^{(r+1)\times(r+1)}$, which is the extension of the submatrix $\hat{A}$ by the found row k_3 and column l_3. If inequality (34) is false, then

$$\left\| \tilde{A} - \tilde{L}\tilde{U} \right\|_C > \rho(r+1)^2 \|E\|_C \geq \rho(r+1)^2 \|\tilde{E}\|_C,$$

where $\|\tilde{E}\|_C$ is the error of the best approximation of $\tilde{A}$ in Chebyshev norm, submatrices $\tilde{L} \in \mathbb{R}^{(r+1)\times r}$ and $\tilde{U} \in \mathbb{R}^{r\times(r+1)}$ of L and U correspond to the rows and columns of $\tilde{A}$. Then from Theorem 5 we get that the submatrix $\hat{A}$ does not have ρ-locally maximum volume inside $\tilde{A}$, and hence there is a replacement that includes only the row k_3 or only the column l_3, increasing the volume by more than ρ times (which contradicts (23)), or simultaneous replacement of some row i and column j by k_3 and l_3, which increases the volume by more than ρ times, which contradicts the additional check (35) that we introduced into the algorithm. This contradiction shows that after termination the inequality (34) must be satisfied.   □

6 Numerical experiments

In this section, we demonstrate the effectiveness of various methods for constructing rank revealing LU decompositions. The purpose of this chapter is to compare the proposed method for finding submatrices of locally maximum volume with other methods, which search for a large volume submatrix.

The fastest method based on volume maximization is the adaptive cross approximation (ACA) [1]. It uses an incomplete Gaussian elimination with partial pivoting. Each new row and column are chosen corresponding to the search for the maximum element in the current row/column of the error. At each step, only one row and one column are searched, and the final row is chosen as the starting row for the next step. The rows and columns thus added are used to construct the $C\hat{A}^{-1}R$ decomposition, also known as the cross or skeleton [6].

As an alternative to ACA, one can also use an incomplete Gaussian elimination with the choice of the next maximum element in the entire matrix (complete pivoting). This guarantees the stability of the method, but greatly increases its computational complexity.

Table 1 Estimates of the complexity and accuracy of approximation for various methods of incomplete LU decomposition of the square matrix $A \in \mathbb{R}^{N \times N}$

Method	$\frac{\|A-LU\|_C}{\|E_{\text{best}}\|_C}$	$\frac{\|A-LU\|_2}{\sigma_{r+1}(A)}$	Complexity
ACA [18]	–	–	$O\left(Nr^2\right)$
maxvol [2], r swaps	–	–	$O\left(Nr^2\right)$
Gauss, complete pivoting	–	–	$O\left(N^2r\right)$
Theorem 3	$\left(1+\rho^2 r(N-r)\right)N$	$1+\rho^2 r(N-r)$	$O\left(N^2r\log_\rho r\right)$
Algorithm 3	$\rho(r+1)^2$	$1+3\rho r(N-r)$	$O\left(N^2r\log_\rho r\right)$

The parameter ρ is assumed to be bounded $\rho = \text{const} > 1$

The maxvol algorithm [2] is more efficient, but also slightly more expensive compared to the ACA method. It allows replacing the chosen rows and columns of the submatrix. After termination, it guarantees that the found submatrix will have ρ-locally maximum volume inside its corresponding rows and columns. To limit its running time, one can directly prohibit more than r row and column swaps.

Finally, we will look at the rank revealing LU decompositions, based on Theorems 3 and 4.

Complexities and known estimates for all the above methods are shown in Table 1 for the case of a square matrix $A \in \mathbb{R}^{N \times N}$. From Theorem 3 one can derive the following bound on the Chebyshev norm

$$
\begin{aligned}
\|A - LU\|_C &\leq \|A - LU\|_2 \\
&\leq \left(1 + \rho^2 r(N-r)\right)\sigma_{r+1}(A) \\
&\leq \left(1 + \rho^2 r(N-r)\right)\|E_{\text{best}}\|_2 \\
&\leq \left(1 + \rho^2 r(N-r)\right)N\|E_{\text{best}}\|_C,
\end{aligned}
$$

where E_{best} is the matrix of the lowest rank r approximation error in the Chebyshev norm.

To test these methods, we will use the RANDSVD ensemble [21], which consists of matrices of the form

$$
A = USV, \quad U, S, V \in \mathbb{R}^{N \times N}, \tag{36}
$$

where S is a fixed diagonal matrix of singular values and U, V are independent random orthogonal matrices.

Since in practice the algorithms stop after a small number of replacements even for $\rho = 1$, we will use $\rho = 1$, although the obtained estimates for the number of steps do not hold in this case.

Table 2 Average errors and computation times of rank $r = 20$ approximation for 100 random matrices of the form (36) of size $N = 1000$ with singular values as in (37)

Method	$\|A - LU\|_F$	$\|A - LU\|_2$	$\|A - LU\|_C$	Time (s)
ACA [18]	3.37×10^{-6}	3.13×10^{-6}	3.88×10^{-8}	1.14×10^{-3}
maxvol [2], r swaps	1.68×10^{-6}	1.52×10^{-6}	1.87×10^{-8}	3.28×10^{-3}
maxvol [2], unlimited steps	1.36×10^{-6}	1.21×10^{-6}	1.45×10^{-8}	6.57×10^{-3}
Gauss, complete pivoting	1.46×10^{-6}	1.27×10^{-6}	1.52×10^{-8}	4.61×10^{-2}
Theorem 3, $\rho = 1$	$\mathbf{1.23 \times 10^{-6}}$	$\mathbf{1.07 \times 10^{-6}}$	$\mathbf{1.31 \times 10^{-8}}$	6.78×10^{-2}
Algorithm 3, $\rho = 1$	1.26×10^{-6}	1.10×10^{-6}	1.35×10^{-8}	7.83×10^{-2}

Lowest average errors among the algorithms are indicated in bold

Table 2 shows the results of rank 20 approximations of random RANDSVD matrices with singular values

$$\sigma_k(A) = 1/2^k. \tag{37}$$

It can be seen that, in order to achieve a high-quality approximation, it is usually not necessary to iterate over all rows and columns, and the faster ACA and maxvol algorithms are sufficient. At the same time, although the estimate of Theorem 3 does not imply achieving high accuracy in the Chebyshev norm, and the found submatrix may differ from the submatrix of locally maximum volume, the corresponding algorithm allows one to achieve approximations of higher accuracy and in less time than search for a locally maximum volume in the entire matrix. At the moment, we do not have a theoretical justification for this feature.

Despite the efficiency of ACA and maxvol on average, it is easy to give a lot of examples where it is necessary to consider the entire matrix to build a high accuracy approximation. Consider, for example, an arbitrary matrix with a hidden block structure

$$B = \Pi_1 \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix} \Pi_2 + E, \tag{38}$$

where E is some noise and Π_1 and Π_2 are permutations of rows and columns. In the worst case, when the starting submatrix is located entirely in one of the blocks, replacing the row or column will not allow leaving the current block and moving to another. This is due to the fact that any column in the rows corresponding to B_1 submatrix will either be inside B_1 or consist entirely of noise. Only the simultaneous replacement of a row and a column allows the algorithm to "see" elements from another block. Even if the initial submatrix is not entirely in one of the blocks, such a structure still significantly complicates the search for the optimal submatrix and leads to an increase in the approximation error.

Table 3 Average errors and computation times of rank $r = 20$ approximations for 100 random matrices of the form (38) of size $N = 1000$ with singular values of blocks as in (39)

Method	$\|A - LU\|_F$	$\|A - LU\|_2$	$\|A - LU\|_C$	Time (s)
ACA [18]	5.41×10^{-1}	4.69×10^{-1}	1.05×10^{-2}	1.20×10^{-3}
maxvol [2], r swaps	5.39×10^{-3}	4.82×10^{-3}	1.02×10^{-4}	3.39×10^{-3}
maxvol [2], unlimited steps	4.42×10^{-3}	3.87×10^{-3}	8.01×10^{-5}	7.73×10^{-3}
Gauss, complete pivoting	1.55×10^{-3}	1.06×10^{-3}	2.26×10^{-5}	4.90×10^{-2}
Theorem 3, $\rho = 1$	$\mathbf{1.41 \times 10^{-3}}$	$\mathbf{9.34 \times 10^{-4}}$	$\mathbf{2.06 \times 10^{-5}}$	6.90×10^{-2}
Algorithm 3, $\rho = 1$	1.45×10^{-3}	9.97×10^{-4}	2.19×10^{-5}	8.64×10^{-2}

Lowest average errors among the algorithms are indicated in bold

We will take the blocks B_1 and B_2 also from the RANDSVD ensemble with singular values

$$\sigma_k (B_1) = \sigma_k (B_2) = 1/2^k, \tag{39}$$

and E as a random Gaussian matrix in which each element has zero mean and variance $\mathbb{E} |E_{ij}|^2 = 10^{-14}$, so that $\|E\|_F \approx 10^{-4}$. Let us again look for an approximation of rank 20 of such a matrix. The optimum in this case is to take 10 rows and columns from the submatrix B_1 and 10 more from the submatrix B_2, so that the error in the spectral norm and the Frobenius norm will be about 10^{-3}. The results of the same algorithms for the matrix B are presented in Table 3. As we can see, the adaptive cross and maxvol methods in this case work much worse and do not allow building an approximation with high accuracy.

Overall, the choice of the rank revealing LU decomposition algorithm depends on the requirements for the speed and accuracy of the method. Adaptive cross approximation [1] is significantly faster than all other methods, but also leads to the largest error. The maxvol algorithm [2] reduces the error several times, while maintaining the same asymptotic cost as the adaptive cross. On average, it can even outperform complete pivoting in accuracy (see Table 2). However, like the adaptive cross, it does not provide any guarantees for the accuracy of the approximation, and can partially lose its accuracy if the approximated matrix has some hidden structure. The algorithm corresponding to Theorem 3 guarantees the best accuracy in terms of the spectral norm and also shows the best results in approximation accuracy among all the given algorithms. However, the presence of the factor N in the Chebyshev norm error estimate does not make this bound very useful. The algorithm for finding a locally maximum volume, proposed in this paper, although looks a little less efficient in practice, allows for high accuracy of the approximation in the Chebyshev norm, as well as the "strong" property (Definition 2) of the resulting RRLU decomposition.

The algorithms from Tables 2 and 3 are realised in Fortran and are available in https://github.com/RodniO/Projective-volume-low-rank (file `ExampleLU.f90`).

7 Conclusion

The obtained estimates show that, in contrast to submatrices of almost maximum volume, the search for which is an NP-hard problem [8, 9], submatrices close to locally maximum volume can be found in polynomial time, which ensures that the various bounds that are often associated with maximum volume submatrices are satisfied. In particular, it is possible to quickly construct cross approximations with guarantees of accuracy in the Chebyshev norm and the spectral norm.

References

1. Bebendorf, M., Rjasanow, S.: Adaptive low-rank approximation of collocation matrices. Computing **70**, 1–24 (2003). https://doi.org/10.1007/s00607-002-1469-6
2. Goreinov, S.A., Oseledets, I.V., Savostyanov, D.V., Tyrtyshnikov, E.E., Zamarashkin, N.L.: How to find a good submatrix. In: Matrix Methods: Theory, Algorithms, Applications, pp. 247–256 (2010). https://doi.org/10.1142/9789812836021_0015
3. Oseledets, I.V., Tyrtyshnikov, E.E.: TT-cross approximation for multidimensional arrays. Linear Algebra Appl. **432**, 70–88 (2010). https://doi.org/10.1016/J.LAA.2009.07.024
4. Goreinov, S.A., Tyrtyshnikov, E.E.: The maximal-volume concept in approximation by low-rank matrices. Contemp. Math. **268**, 47–51 (2001). https://doi.org/10.1090/conm/280
5. Goreinov, S.A., Tyrtyshnikov, E.E.: Quasioptimality of skeleton approximation of a matrix in the Chebyshev norm. Dokl. Math. **83**(3), 1–2 (2011). https://doi.org/10.1134/S1064562411030355
6. Goreinov, S.A., Tyrtyshnikov, E.E., Zamarashkin, N.L.: A theory of pseudo-skeleton approximations. Linear Algebra Appl. **261**, 1–21 (1997). https://doi.org/10.1016/S0024-3795(96)00301-1
7. Osinsky, A.I., Zamarashkin, N.L.: Pseudo-skeleton approximations with better accuracy estimates. Linear Algebra Appl. **537**, 221–249 (2018). https://doi.org/10.1016/j.laa.2017.09.032
8. Çivril, A., Magdon-Ismail, M.: On selecting a maximum volume sub-matrix of a matrix and related problems. Theor. Comput. Sci. **410**(47–49), 4801–4811 (2009). https://doi.org/10.1016/j.tcs.2009.06.018
9. Çivril, A., Magdon-Ismail, M.: Exponential inapproximability of selecting a maximum volume sub-matrix. Algorithmica **65**(1), 159–176 (2013). https://doi.org/10.1007/s00453-011-9582-6
10. Gu, M., Miranian, L.: Strong rank revealing Cholesky factorization. Electron. Trans. Numer. Anal. **17**, 76–92 (2004)
11. Osinsky, A.: Rectangular maximum volume and projective volume search algorithms. Preprint at arXiv:1809.02334v1 (2018)
12. Massei, S.: Some algorithms for maximum volume and cross approximation of symmetric semidefinite matrices. BIT Numer. Math. **62**, 195–220 (2022). https://doi.org/10.1007/s10543-021-00872-1
13. Gu, M., Eisenstat, S.C.: Efficient algorithms for computing a strong rank-revealing QR factorization. SIAM J. Sci. Comput. **17**(4), 848–869 (1996). https://doi.org/10.1137/0917055
14. Anderson, D.G., Gu, M.: An efficient, sparsity-preserving, online algorithm for low-rank approximation. In: Precup, D., Teh, Y.W. (eds.) Proceedings of the 34th International Conference on Machine Learning, vol. 70, pp. 156–165. PMLR, Sydney (2017)
15. Miranian, L., Gu, M.: Strong rank revealing LU factorizations. Linear Algebra Appl. **367**, 1–16 (2003). https://doi.org/10.1016/S0024-3795(02)00572-4
16. Pan, C.-T.: On the existence and computation of rank revealing LU factorizations. Linear Algebra Appl. **316**, 199–222 (2000). https://doi.org/10.1016/S0024-3795(00)00120-8
17. Boutsidis, C.: Topics in matrix sampling algorithms. PhD thesis, Rensselaer Polytechnic Institute (May 2011)

18. Bebendorf, M.: Approximation of boundary element matrices. Numer. Math. **86**, 565–589 (2000). https://doi.org/10.1007/PL00005410

19. Cortinovis, A., Kressner, D., Massei, S.: On maximum volume submatrices and cross approximation for symmetric semidefinite and diagonally dominant matrices. Linear Algebra Appl. **593**, 251–268 (2020). https://doi.org/10.1016/j.laa.2020.02.010

20. Cortinovis, A.: Fast deterministic and randomized algorithms for low-rank approximation, matrix functions, and trace estimation. PhD thesis, EPFL (June 2022)

21. Osinsky, A.I., Zamarashkin, N.L.: On the accuracy of cross and column low-rank maxvol approximations in average. Comput. Math. Math. Phys. **61**, 786–798 (2021). https://doi.org/10.1134/S0965542521050171